Original Article

An Efficient Approach for Driver Drowsiness Detection at Moderate Drowsiness Level Based on Electroencephalography Signal and Vehicle Dynamics Data

Sara Houshmand[1], Reza Kazemi[1], Hamed Salmanzadeh[2]

Departments of [1]Mechanical Engineering and [2]Industrial Engineering, KN. Toosi University of Technology, Tehran, Iran

Abstract

Background: Drowsy driving is one of the leading causes of severe accidents worldwide. In this study, an analyzing method based on drowsiness level proposed to detect drowsiness through electroencephalography (EEG) measurements and vehicle dynamics data. **Methods:** A driving simulator was used to collect brain data in the alert and drowsy states. The tests were conducted on 19 healthy men. Brain signals from the parietal, occipital, and central parts were recorded. Observer Ratings of Drowsiness (ORD) were used for the drowsiness stages assessment. This study used an innovative method, analyzing drowsiness EEG data were in respect to ORD instead of time. Thirteen features of EEG signal were extracted, then through Neighborhood Component Analysis, a feature selection method, 5 features including mean, standard deviation, kurtosis, energy, and entropy are selected. Six classification methods including K-nearest neighbors (KNN), Regression Tree, Classification Tree, Naive Bayes, Support vector machines Regression, and Ensemble Regression are employed. Besides, the lateral position and steering angle as a vehicle dynamic data were used to detect drowsiness, and the results were compared with classification result based on EEG data. **Results:** According to the results of classifying EEG data, classification tree and ensemble regression classifiers detected over 87.55% and 87.48% of drowsiness at the moderate level, respectively. Furthermore, the classification results demonstrate that if only the single-channel P4 is used, higher performance can achieve than using data of all the channels (C3, C4, P3, P4, O1, O2). Classification tree classifier and regression classifiers showed 91.31% and 91.12% performance with data from single-channel P4. The best classification results based on vehicle dynamic data were 75.11 through KNN classifier. **Conclusion:** According to this study, driver drowsiness could be detected at the moderate drowsiness level based on features extracted from a single-channel P4 data.

Keywords: *Driving simulator, drowsy driving, electroencephalography signal, feature extraction, signal classification, supervised learning methods, vehicle dynamics*

Submitted: 23-May-2021 **Revised:** 25-Aug-2021 **Accepted:** 28-Oct-2021 **Published:** 10-Nov-2022

Introduction

Drowsy driving is an important cause of fatal accidents. Drowsy drivers have a low level of consciousness and cognition about their environment. Making the right decisions becomes increasingly difficult for the driver. Drunk driving is just as dangerous as drowsy driving.[1]

Drowsiness can be determined by the drivers' faces gestures, vehicle dynamics, and physiological signals. Physiological signals can be determined regardless of the driver's driving ability or the environment they are driving in.[2,3] The combination of electroencephalography (EEG), electrooculography, and driving quality signals was used by Noori *et al.* to detect driver drowsiness.[4] EEG signals are the best descriptors of sleep state compared to other physiological methods for detecting driver drowsiness.[2,4,5] In addition, the logarithm of energy of the signal as well as Higuchi's and Petrosian's fractal dimension was used to detect drowsiness using the two-tailed *t*-test method.[6] There are four major bands in EEG signals, and some actions have their own spectral domains. Four frequency bands of EEG signals are delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (12–30 Hz). During relaxation state and closed eyes, the brain's parietal and occipital regions produce alpha

How to cite this article: Houshmand S, Kazemi R, Salmanzadeh H. An efficient approach for driver drowsiness detection at moderate drowsiness level based on electroencephalography signal and vehicle dynamics data. J Med Sign Sens 2022;12:294-305.

Address for correspondence:
Dr. Sara Houshmand, Department of Mechanical Engineering, KN. Toosi University of Technology, Tehran, Iran. E-mail: houshmandsara8@gmail.com

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.jmss_124_21

Quick Response Code:

waves.[7] As opposed to this, the frontal area generates beta waves that involve thinking and making a decision.[8] There is a higher rate of alpha waves being emitted in a drowsy state compared to alertness.[9]

EEG signals can be analyzed by various methods, such as fast Fourier transform (FFT) and the wavelet transform. Due to the non-stationary nature of the EEG signal, it is essential to consider both frequency and time domain features simultaneously to perform reliable analysis. Time-frequency analysis includes the wavelet transform, which is an extremely powerful method. Using wavelet transforms and FFT, Akin demonstrated the major advantage of wavelet transforms over Fourier transforms in detecting disease in the brain that allows both time and frequency information to be obtained.[10] By passing the time-series signal through both high and low pass filters, the wavelet decomposition can be obtained. Continuous wavelets and discrete wavelets are the two forms of the wavelet transform. In processing and analyzing EEG signals, the Discrete Wavelet Transform (DWT) has been prominently utilized. Low computation times and ease of implementation are two benefits of DWT.[11] For epilepsy EEG signal classification, several studies have used DWT[12,13] and also there have been several studies using DWT to detect Alzheimer's.[14] EEG signals contain information about the signal that can be gleaned from their features. As part of the visual perception and mental imagery of paintings tasks, Shourie classified participants based on EEG features such as energy, average, standard deviation, and entropy.[15]

It is extremely challenging to process the EEG signal due to artifacts. A number of events can result in these artifacts, such as blinking eyes and heart pulses. In other words, many physiological signals appear on EEG data including electrocardiograms, electromyography, and electrooculograms. EEG signals can be removed from artifacts using independent component analysis (ICA). The researchers used ICA to detect and eliminate the eye artifacts and to detect signals from two categories of danger and information words.[16]

Scientists use multiple types of classifiers for the classification of EEG signals. Based on neuro-fuzzy algorithms and K-nearest neighbors (KNN), distinguishing Alzheimer's from other brain disorders is possible.[17] KNN classifier was used to recognize emotional cues using an autoregressive model of EEG signals during visual and auditory inductions.[18] Support vector machines (SVM) were used by Rasekhi *et al*. to classify epileptic patients' EEG data into preictal and non-preictal seizure classes.[19] In order to detect motor imagery signals, Wang and Zhang used the Naive Bayes classifier.[20] In Prasad *et al*.'s study, the EEG data in the lower gamma band (30–40 Hz) were classified by logistic Ensemble Regression.[21]

For monitor driver drowsiness levels, using a limited number of electrodes is desired.[22-25] In[22,23] a few numbers of channels, and in[24,25] a single-channel were applied. Nevertheless, the analysis of EEG channels over a number of channels shows how the EEG signal changed during the transition from being awake to extreme drowsiness. EEG data obtained from a single channel can be analyzed more quickly with high accuracy. These data are preferable for real-world application. During the transition from being awake to extreme drowsiness, the EEG signal of midline sites (e.g., F4, C4, P4, and O2) undergoes significant changes in its frequency and amplitude.[26] The correlation analyses performed by Lin and colleagues evaluate the relationship between the smoothed driving error and the log sub-band power spectrum of the EEG signals. Their results indicate that P4 and PZ are suitable channels for tasks related to drowsiness detection.[27]

Using spectral features of multiple independent brains sources, a study proposed a perceptual function integration system to detect the driver's alertness and drowsiness states. The activation of the different cortical sources of the brain is highly correlated with changes in alertness state, which was demonstrated by the analysis of the data.[28] Correa *et al*. designed a neural network to detect driver drowsiness. With the aid of spectral, and wavelet analyses, the study developed a method to automatically detect the drowsiness stage. To distinguish alertness from drowsiness, 19 features were computed from only one EEG channel. Seven parameters were chosen by using Wilks' lambda criteria to use as input to a neural network classifier.[29] Hajinoroozi *et al*. developed a novel channel-wise CNN based on raw EEG data and ICA for the evaluation of driver fatigue.[30]

In our last study, we designed a novel convolutional neural network to detect drowsiness based on alpha spindle detection.[31] This study aims to propose a more applicable and practical method for driver drowsiness detection in real-world use. In this method, the EEG signal features in respect to ORD, in other words in respect to drowsiness level, were investigated for drivers at alert and drowsy state. To get the result, EEG data classified into alert and drowsy categories using KNN, SVM, Decision Tree machines, Naive Bayes, and Ensemble Regression method. Major contributions of this study are as follows:

- Based on EEG recordings, moderate drowsiness is distinguishable from fatigue and extreme drowsiness;
- 91% accuracy can be achieved in detecting drowsiness using five EEG features of single-channel P4 which are easily and quickly can be computed and made this method more applicable for real-world application

- An EEG pattern was mainly observed, first ascending up from alertness to the moderate drowsiness state and then descending to extreme drowsiness state;
- The elapsed time does not provide a reliable indicator of driving drowsiness. Thus, this paper replaces the observer rating of drowsiness with the time.

Materials and Methods

Drowsiness assessment

By observing the driver, experts can assess driver drowsiness. Three observers who are experts in drowsiness measure drive behavior and facial signs to calculate the Observer Rate of Drowsiness (ORD). There are five levels of drowsiness, respectively: Alertness, slight drowsiness, moderate drowsiness, very drowsiness, and extreme drowsiness. This study analyzed EEG signals using drowsiness scores instead of time because, during driver drowsiness tests, the driver may begin to become alert after several epochs of being drowsy. The ORD levels are shown in Table 1.

Driving simulator

Nasir Semi 003 was used to conduct drowsiness tests. C++ programming and UnrealTM 4 graphics engine are used to develop the driving simulator's 14 DOF dynamic model. Sample rates of 30 Hz were employed to record dynamic data about the vehicle. The steering wheel, shift sticks, and pedals allow the driver to control the simulator. The driving simulator and data collection tools are shown in Figure 1.

There is a 67 km quasi-circle highway with three lanes and designated so that a drowsy driver can exit the road if no torque is applied. Pedestrians and other vehicles were not on the road, so there were no sharp turns to confuse drowsy drivers. It is allowed to drive at a maximum speed of 100 km/h and a minimum speed of 80 km/h.

Participants

Tests for drowsiness were conducted on nineteen healthy and licensed male subjects. Their ages ranged from 26 to 50. The driving scene from the third-person perspective and the driving path are shown in Figures 2(a) and 2(b), respectively. None of the participants suffer from sleep disorders, nor are they addicted to alcohol, drugs, or cigarettes. Before each test, all participants receive a complete description of the test protocol. A questionnaire was used to collect information about their lifestyle, sleep, and health. On the night before the test, participants were asked to sleep at their usual bedtime and not to drink any tea or coffee. Maintenance of Wakefulness Test (MWT) a few days prior to the driving tests identified participants with insufficient ability to withstand drowsiness during the given time period. The purpose of the MWT test is to keep subjects alert for 40 min without engaging in any activity. The test is stopped if subjects fall asleep three times in the first stage of sleep, or once in any of the other stages. MWT test results revealed that three subjects were excluded from the driving trials based on their abnormal behavior. Therefore, nineteen subjects underwent driving tests. The test protocols have been approved by the ethics committee of the cognitive and science council's Institutional Review Board (Grant No. 1307). According to the Declaration of Helsinki, the procedures were followed.

Figure 1: Nasir Semi 003™ driving simulator[3]

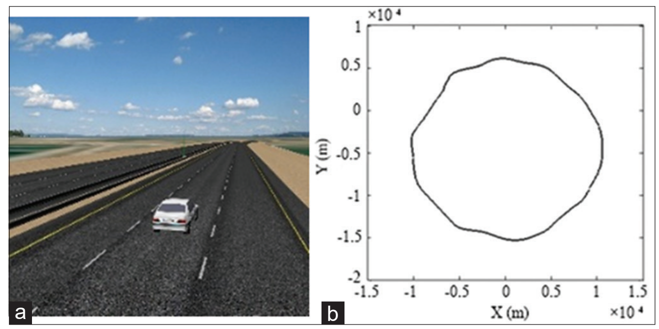

Figure 2: The following are the maps: (a) Driving scene from the third-person perspective; (b) 67-km closed-loop driving expedition[3]

Table 1: The levels of observer ratings of drowsiness[31]

Drowsiness level	1	2	3	4	5
Driver status	Not drowsy	Slightly drowsy	Moderately drowsy	Very drowsy (fatigue)	Extremely drowsy

Table 2: lists the categories of each classifier

Types	Instance-based learning	Logic-based algorithms	Support vector machines	Statistical learning algorithms
Classifier	KNN	Regression tree, Ensemble regression	SVM, SVM regression model	Naive Bayes model

KNN - K-nearest neighbors; SVM - Support vector machine

Signal preprocessing

(C3, C4, P3, P4, O1, O2) EEG channels were evaluated and Cz was used as a reference. Preprocessing and denoising the EEG data are necessary due to the low signal-to-noise ratio. EEG data is tested for outliers using Grubbs' outlier test. There is a 4th-order zero-lag Butterworth band-pass filter applied to remove out-of-band noise (between 0.1 and 31 Hz) after removing outliers; the frequency of the power line interferences is 50 Hz. Data were divided into 30 s epochs after denoising. Figure 3 shows the flow chart of signal preparation for classification.

Feature selection

In this paper, Neighborhood Component Analysis (NCA) method is used for feature selection. With NCA classification problems can be resolved with the greatest level of accuracy. Further, NCA is capable of reducing computational costs through features selection and dimensional reduction. As part of an NCA analysis, S represents the training set:[32]

$$S = \{(a_i, b_i),\ i = 1,2,...n,\ a_i \text{ is a features vector and } b_i \text{ is drowsiness state} = \{\text{“alerts”, “drowsy”}\} \quad (1)$$

Where the feature vector is a 13-dimensional vector consist of extracted features including (variance, standard deviation, shape factor, RMS, range, geomean, energy, average, entropy, and power spectrum for each frequency band [delta, theta, alpha, beta]). There is accuracy for classification:[32]

$$accuracy(f_w) = \sum_{i=1}^{n} \sum_{j=1\, j'i}^{k} \frac{exp\left(\frac{-\sum_{k=1}^{m} f_{wk}^{2} \left|a_{ik} - a_{jk}\right|}{á}\right)}{\left(\sum_{j=1}^{m} exp\left(\frac{-\sum_{k=1}^{m} f_{wk}^{2} \left|a_{ik} - a_{jk}\right|}{â}\right)\right)} \quad (2)$$

$b_{ij} b_i$

Where f_w denote the feature weights, and b_{ij} is 1 when $b_i = b_j$ and otherwise it is 0. α and β are inputs parameter. And finally, the last step is finding the maximum of the derivation of *accuracy* ($f_{w)}$ function.[32]

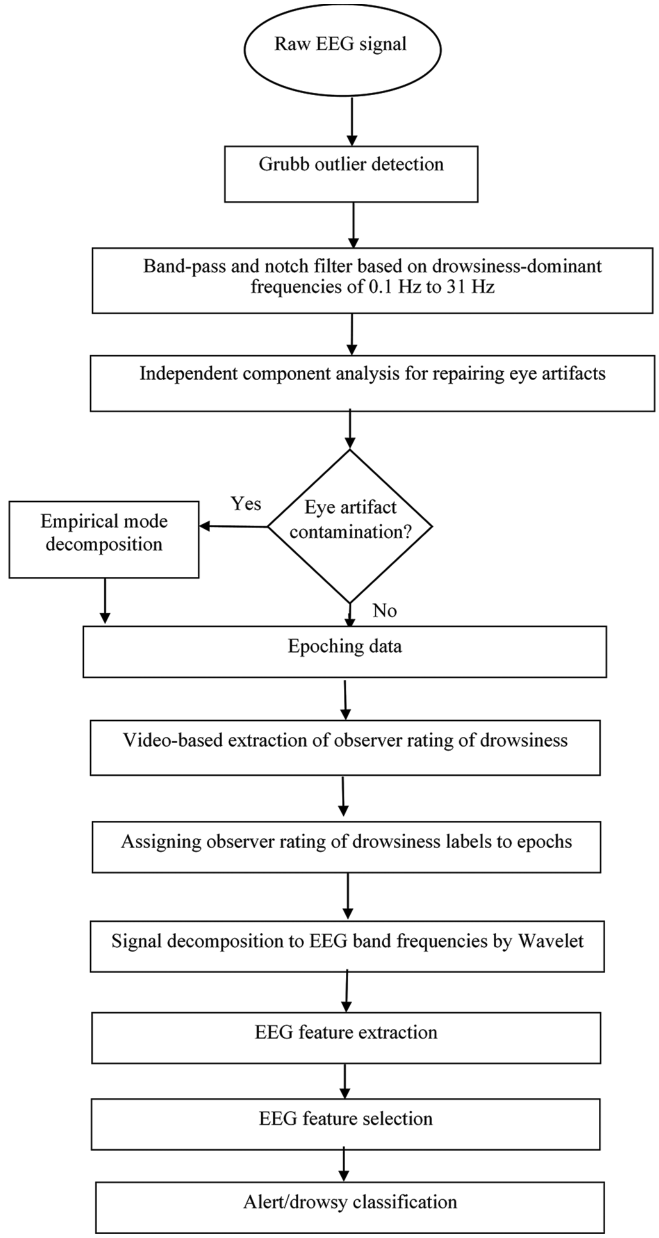

Figure 3: Flowchart of signal preparation for classification

$$\frac{accuracy(f_w)}{\partial w_k} = 2\left(\frac{1}{\alpha} \sum_{i=1} \Big(\frac{exp\left(\frac{-\sum_{k=1}^{m} f_{wk}^{2} \left|a_{ik}\right|}{á}\right)}{\left(\sum_{j=1}^{m} exp\left(\frac{-\sum_{k=1}^{m} f_{wk}^{2} \left|a_{ik}\right|}{â}\right)\right)} \sum_{l \neq i} \frac{exp\left(\frac{-\sum_{k=1}^{m} f_{wk}^{2} \left|a_{ik}\right|}{á}\right)}{\left(\sum_{j=1}^{m} exp \left(\frac{-\sum_{k=1}^{m} f_{wk}^{2} \left|a_{ik}\right|}{â}\right)\right)} - \sum_{j=1} b_{ij} \frac{exp\left(\frac{-\sum_{k=1}^{m} f_{wk}^{2} \left|a_{ik}\right|}{á}\right)}{\left(\sum_{j=1}^{m} exp \left(\frac{-\sum_{k=1}^{m} f_{wk}^{2} \left|a_{ik}\right|}{â}\right)\right)} \left|a_{ik} - a_{jk}\right| \left|a_{ik} - a_{jk}\right| \right) - \gamma) f_w \quad (3)$$

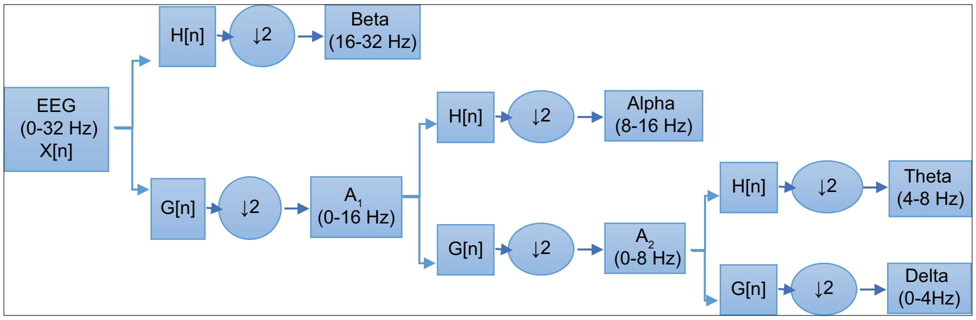

Figure 4: Decomposition of EEG signal by using discreet wavelet analysis

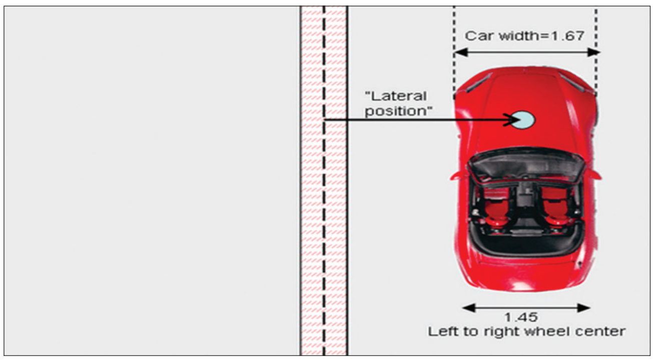

Figure 5: The lateral position of the vehicle[36]

Where γ is regularization parameter. By setting α = β = 2 and γ = 1, the most reasonable feature weights can be achieved.[32]

Signal decomposition

DWT can be used for signal decomposition. DWT is based on a method of decomposing signals with multiple resolutions. In wavelet theory, the wavelet transform can be modeled as a combination of low-pass filters G and high-pass filters H. G(z) represents the z-transform of the filter G. Complementary high-pass filters can be defined as follows:[33]

$$H(z) = zG(-z^{-1}) \tag{4}$$

The following formula can be used to determine a sequence of filters:[33]

$$G_{i+1}(z) = G\left(z^{2^i}\right)G_i(z),\;\; H_{i+1}(z) = H\left(z^{2^i}\right)H_i(z) \tag{5}$$

Time-domain expression of a two-scale relation is:[33]

$$g_{(i+1)}(k) = [g]_{(\uparrow 2^i)}\, g_i(k) h_{(i+1)}(k) = [h]_{(\uparrow 2^i)}\, h_i(k) \tag{6}$$

Where $[*]_{\uparrow m}$ indicates the up-sampling by a factor of m. By decomposing the signal into a coarse approximation and detail information, the DWT analyzes the signal at different frequency bands, with different resolutions. Two sets of functions are used by the DWT, known as scaling functions and wavelet functions, which are related to low-pass and high-pass filters. By using sequences of high-pass and low-pass filters on the time domain signal, the signal is decomposed into different frequency bands. Figure 4 illustrates a schematic representation of the multi-resolution decomposition of signal X(n). There are two digital filters including a High-pass filter, H(.) and a low-pass filter, G(.) and two downsamplers in each stage of the scheme.

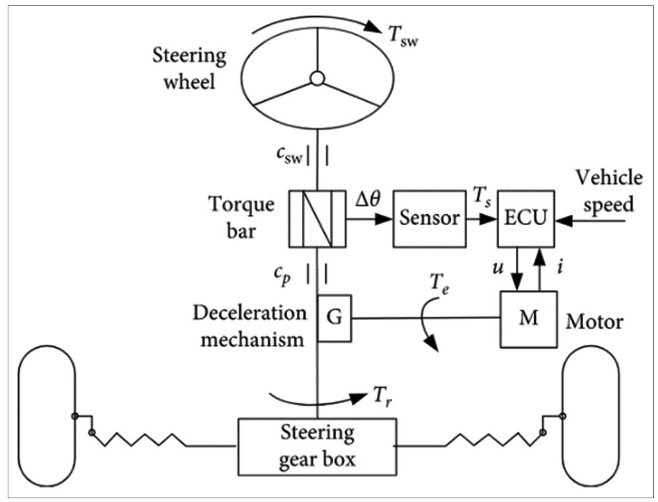

Figure 6: The simplified model of the electric power which includes a steering angle sensor[37]

Classification

There are four types of classification algorithms including logic-based algorithms, statistical learning algorithms, perceptron-based techniques, and instance-based learning.[34] The problem with perceptron-based techniques is they have an implicit underlying probability model and they are inefficient for predicting driving behavior, for this reason, they are ignored. Data are labeled with a real value in logic-based algorithms; the decision being made is to predict a value for new unpredicted data. There are statistical learning algorithms, which determine the probability of an instance belonging to each class.[35] The instance-based learning algorithm uses data labeling to classify participants into either drowsy or alert. Each classifier and its combination are selected from the categories. In this study, six classification methods were used for supervised learning. KNN, regression tree, SVM, SVM regression model, Naive Bayes model, and ensemble of learners for regression were used as classifiers. Table 2 lists the categories of each classifier.

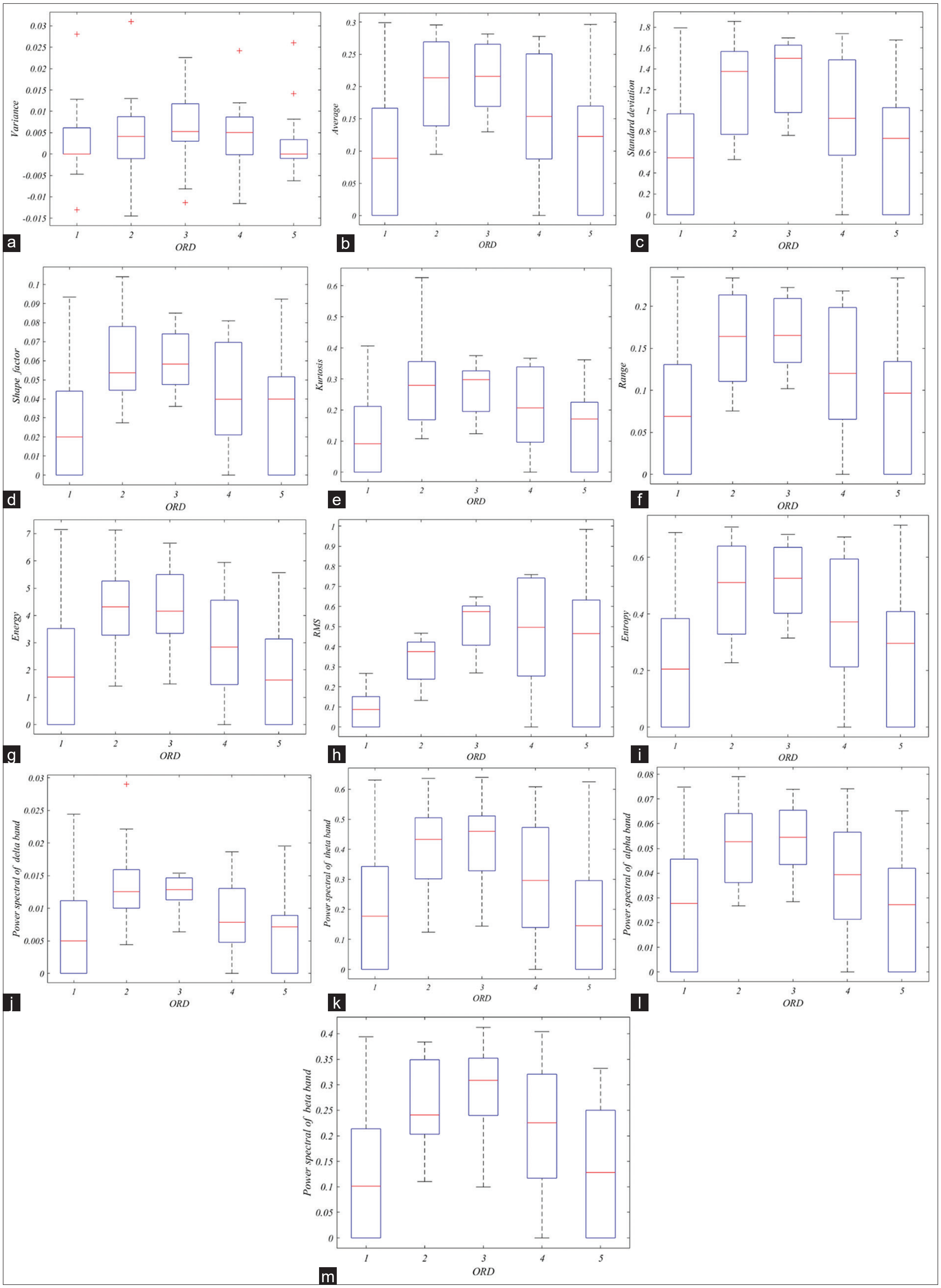

Figure 7: The variation of features during transition from alertness to drowsiness state including (a) variance, (b) average, (c) standard deviation, (d) shape factor, (e) kurtosis, (f) range, (g) energy, (h) RMS, (i) entropy, and power spectral of (j) delta, (k) theta, (l) alpha, (m) beta bands

Regression-tree classifier is an effective classification technique is decision tree. In this method, every element of the classification domain is referred to as a class. An internal (non-leaf) node of a decision tree or classification tree is labeled with an input feature. SVM is surprised learning model, which builds a hyperplane or set of hyperplanes in an infinitum or high dimensional space. Using different kernel functions as decision functions, SVM controls the sparseness of data by controlling the margin. Naive Bayes is an effective classifier, based on the maximum likelihood principle. The class label c can be assigned to an unknown sample with features x, i.e., choose the class with the highest posterior probability based on observed data. In KNN classifier, a neighbor is chosen based on the smallest difference in the intended property of a set of adjacent objects. KNN classifier is easy to interpret, and it has a low calculation time, high predictive power.

Vehicle dynamic data

To detect driver drowsiness based on vehicle dynamic data, the lateral position of the vehicle and steering angle were selected. The mean of them was fed into classifiers to classify data into two categories "alert" and "drowsy."

The lateral position

To determine the lateral position of a vehicle, the distance between the centerline of the road and a vehicle's position has been used. As show in Figure 5, the lateral position of a vehicle is indicated.

Steering angle

The Steering Angle Sensor is an important part of a vehicle's safety system. It transmits the steering wheel's rate of turn, wheel angle, and other important data to the specific vehicle's computer. Figure 6 indicates the simplified model of the electric power steering system, which includes a steering angle sensor.

Results

Feature extraction of EEG signal

Participants in the VR-based highway-driving study included nineteen healthy people. Data from the EEG channels (C3, C4, P3, P4, O1, O2]) of subjects used to detect driver drowsiness. Thirteen features of EEG signal were investigated, including variance, standard deviation, shape factor, RMS, range, kurtosis, energy, average, entropy, and power spectrum for each frequency band (delta, theta, alpha, beta). During the transition from alertness to drowsiness state, the change in these features is indicated in Figure 7.

Feature selection was carried out using NCA in this study. In this method, significant features are those with a high weight to classify data. An irrelevant feature has negligible weight. Figure 8 displays the weights associated with each feature in EEG classification into "drowsy" or "alert." Figure 8 shows that feature weights of mean, standard deviation, kurtosis, Energy, and Entropy are more than 1.5. Therefore, NCA analysis of EEG data suggests that these features are more preferred than others for classifying alertness and drowsiness.

Feature extraction of vehicle dynamic data

To detect driver drowsiness based on vehicle dynamic data, lateral position of the vehicle and steering angle were selected. The mean of them was fed into classifiers including KNN, Regression Tree, SVM, Classification Tree, Naive Bayesian, SVM Regression, Ensemble Regression to classify data into two categories "alert" and "drowsy."

During the transition from alertness to drowsiness, fluctuations of the lateral position of the vehicle are

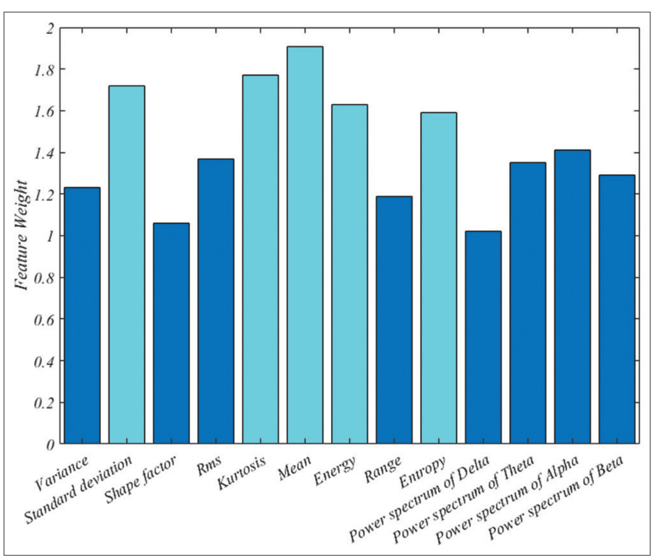

Figure 8: Weights of extracted features for drowsiness classification

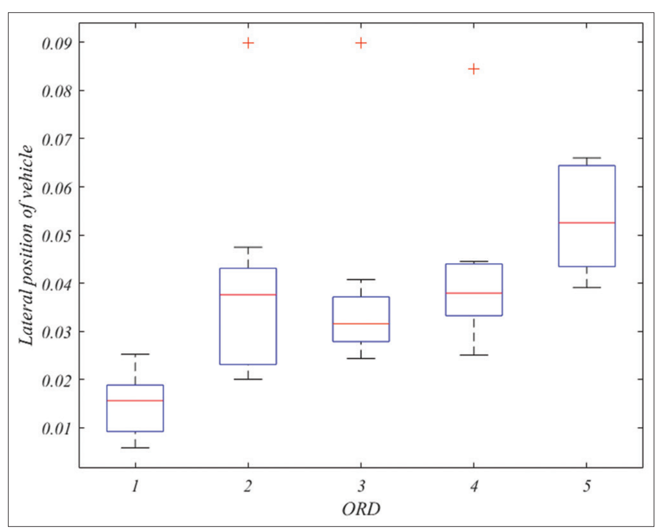

Figure 9: The variation of the lateral position of the vehicle during transition from alertness to drowsiness state

indicated in Figure 9. As shown by the figure, drivers in the awake state are more likely to drive freely with greater confidence since they know they can control the vehicle. When drivers are slightly drowsy, they try to drive in the middle of the road because they are aware that their ability to control the car is reduced, so they drive more carefully. Once the driver's ORD has increased to moderate and extreme drowsiness, they cannot resist sleep anymore. They will no longer be able to hold the vehicle on the road, so they will end up driving in higher lateral positions.

The change in steering angle during the transition from alertness to drowsiness state is shown in Figure 10. By considering this figure and figure (lateral position), the awake driver also has a high lateral position, but he can control the car by applying a greater steering angle. On the other hand, when they are moderately and extremely drowsy, they have insufficient steering angles, which result in them going off the road. But in the slight drowsiness, according to figure (lateral position), try to drive in the centerline, for staying on the road they have to applying a smaller steering angle.

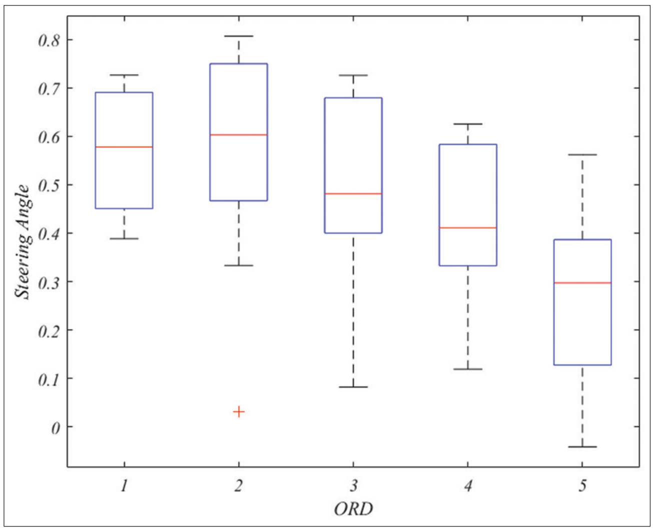

Figure 10: The variation of the steering angle during transition from alertness to drowsiness state

Classification result based on EEG data

The extracted features including mean, standard deviation, kurtosis, energy, and entropy serve as the input parameters

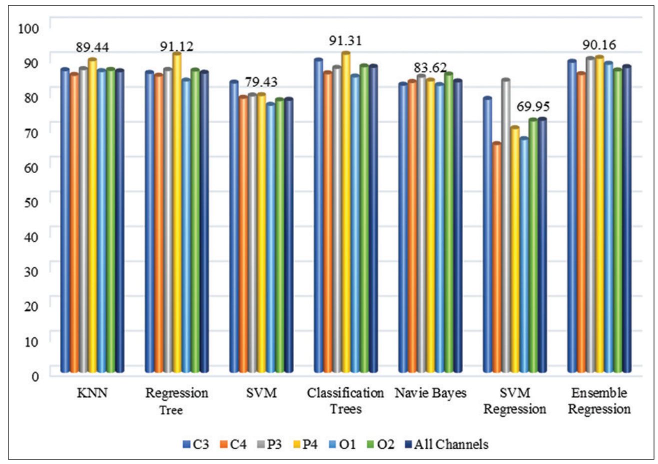

Figure 11: The result of using classifiers including the KNN, Regression Tree, SVM, Classification Tree, Naive Bayes, SVM-Regression, Ensemble Regression classifier for driver drowsiness detection based on single-channel C3, single-channel C4, single-channel P3, single-channel P4, single-channel O1, single-channel O2, and data of all channels {C3, C4, P3, P4, O1, O2} together

Table 3: The classification result of all channels

Subjects	Accuracy of classifiers						
	KNN	Regression tree	SVM	Classification tree	Navie Bayes	SVM regression	Ensemble regression
1	78.33	81.67	76.67	86.67	83.33	71.67	85.00
3	82.13	85.80	79.10	87.31	89.58	76.14	84.41
4	81.24	87.24	85.96	88.52	82.11	71.10	87.24
5	78.45	76.73	72.39	78.58	74.99	66.27	79.25
6	94.44	92.59	57.41	98.15	94.44	57.41	96.30
7	93.46	93.82	86.94	94.91	86.21	80.05	94.91
8	77.78	70.83	65.28	72.22	48.61	50.16	72.22
9	74.43	70.57	55.72	72.69	55.49	41.93	75.10
10	85.01	83.76	64.32	85.85	80.98	55.54	87.93
11	72.22	79.72	72.64	80.97	78.33	79.44	79.72
12	95.45	95.45	80.30	95.45	94.70	57.58	93.18
13	99.26	98.89	87.41	99.26	97.78	91.85	99.26
14	96.94	97.79	97.43	97.53	96.32	96.27	97.89
15	96.35	97.33	96.60	95.58	95.33	96.52	96.64
16	89.39	75.76	75.76	75.76	83.33	63.64	86.36
18	96.64	96.46	96.05	95.68	95.77	95.26	96.56
19	78.33	78.03	79.27	80.45	80.02	78.67	78.45
Average	85.85	86.29	78.09	87.55	83.38	72.36	87.48

KNN - K-nearest neighbors; SVM - Support vector machine to the KNN, Regression Tree, SVM, Classification Tree, Naive Bayesian, SVM Regression, Ensemble Regression classifiers. Ninety percent of each subject's data were used as training data, and the rest as testing data. The single-channel P4 has been proven to effectively detect driver drowsiness in several studies.[27,31] Drowsiness is detected in the first step based on the data of all channels. After that, only data of single-channel P4 were used, and the results were compared. The classification results for All channels for each subject are shown in Table 3.

Figure 11 shows the results of using various classifiers, including KNN, Regression Tree, SVM, Classification Tree, Naive Bayes, SVM Regression, and Ensemble Regression. Firstly, each channel's data is classified, and then all of the channels' data is used for classification. Figure 11 shows the classification results for every channel (such as C3, C4, P3, P4, O1, O2) and the classification results for all channels taken together (C3, C4, P3, P4, O1, O2). Figure 11 compares the results from the different channels to demonstrate the P4 channel's potential to detect drowsiness. It is still the single-channel P4 that provides the best classification results even after taking into account all the channel data. Based on the single-channel P4, the classification performance was found to be: 89.44%, 91.12%, 79.43%, 91.31%, 83.62%, 69.95%, and 90.16%, respectively. Furthermore, the classification results reveal that regression and classification tree should have the most success in classifying drowsy data.

Classification result based on vehicle dynamic data

The mean of lateral position and steering angle were used to classify vehicle dynamic data into two categories "alert" and "drowsy." The KNN, Regression Tree, SVM, Classification Tree, Naive Bayesian, SVM Regression, Ensemble Regression classifiers were employed on data. 90 percent of each subject's driving data were used as training data and the rest as testing data. The classification results for vehicle dynamic data for each subject are shown in Table 4.

Figure 12 shows the results of using various classifiers, including KNN, Regression Tree, SVM, Classification Tree, Naive Bayes, SVM Regression, and Ensemble Regression. Based on the vehicle dynamic data, the classification performance was found to be: 75.11%, 55.01%, 71.98%, 68.73%, 51.68%, 68.84%, and 59.05%, respectively.

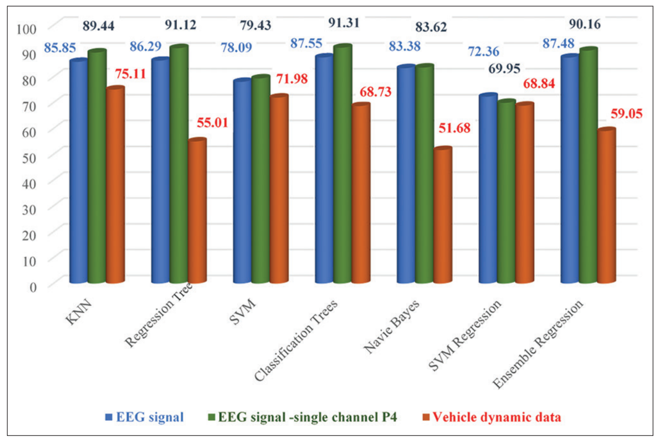

Figure 12: The result of using classifiers including the K-nearest neighbors, Regression Tree, Support vector machines, Classification Tree, Naive Bayes, Support vector machines-Regression, Ensemble Regression classifier for driver drowsiness detection based on EEG signal of (C3, C4, P3, P4, O1, O2) channels and single channel P4, and vehicle dynamic data

Table 4: The classification result of vehicle dynamic data

Subjects	Accuracy of classifiers						
	KNN	Regression tree	SVM	Classification tree	Navie Bayes	SVM regression	Ensemble regression
1	67.33	53.67	76.67	63.67	55.67	74.33	57.43
3	73.13	58.1	74.31	68.8	65.14	73.58	60.41
4	68.24	64.96	72.52	70.24	46.1	71.11	59.24
5	70.45	60.39	59.58	60.73	28.27	62.99	52.25
6	78.44	35.41	79.15	79.59	25.41	75.44	69.3
7	82.46	62.94	81.91	79.82	71.05	75.21	56.91
8	64.78	35.28	61.22	49.83	33.16	33.61	48.22
9	57.43	35.72	59.69	48.57	15.93	41.49	41.1
10	72.01	43.32	75.85	63.76	32.54	62.98	57.93
11	60.22	43.64	57.97	57.72	68.44	59.33	45.72
12	82.45	65.3	86.45	75.45	33.58	76.7	61.18
13	86.26	69.41	76.26	79.89	82.85	82.78	67.26
14	88.94	70.43	82.53	87.79	65.27	80.32	66.89
15	87.35	69.6	80.58	78.33	76.52	83.33	74.64
16	78.39	47.76	53.76	55.76	56.64	66.33	60.36
18	86.64	68.05	80.68	80.46	72.26	80.77	72.56
19	72.33	51.27	64.45	68.03	49.67	70.02	52.45
Average	75.11	55.01	71.98	68.73	51.68	68.84	59.05

KNN - K-nearest neighbors; SVM - Support vector machine

Furthermore, the classification results reveal that KNN has the most performance in classifying drowsy data about 75%.

Discussion

This study investigated the behavior of the EEG signal during driving and drowsiness in healthy individuals. Participants completed questionnaires about nicotine and alcohol addiction, their health, and their daily sleep schedules. There were no addictions or sleep disorders among the participants. A driving simulator test was conducted on 19 male participants. We measured the drowsiness levels of drivers using the ORD method, which is nonintrusive and nondistractive.

Detection of driver drowsiness must be done in the beginning of serious drowsiness signs. This time is the golden time which early enough to effective prevention of drowsy driving dangers and fatal crashes. In the moderate level of drowsiness (ORD = 3), a driver still can control the vehicle, and the vehicle is more likely to be in its lane [as shown in Figure 9]. Having a small amount of warning will quickly make him aware of his surroundings and he/she can effectively control his vehicle and prevent accidents. However, drowsy driver loses control of vehicle at higher level of drowsiness (ORD ≥4), and usually drives with the higher lateral position [Figure 9] and drives outside the road, so not only they will not have enough time to prevent an accident, but they will likely overreact to the warning, such as applying a very strong steering force that overturns the car.

With respect to ORD, five features are extracted. They are including mean, standard deviation, kurtosis, energy, and entropy, which drastically change at the moderate level of drowsiness compared to their amounts in the alertness state. Different types of classifiers incorporate these features. Using single-channel P4 data, the classification tree, regression tree, and ensemble regression classifiers achieve a performance that's higher than 90%, and when using all channels' data (C3, C4, P3, P4, O1, O2), the performance is higher than 86%.

For real-world applications, single-channel data are preferred over multiple channels since it facilitates faster data processing; meanwhile, a single channel is more convenient for drivers and less intrusive. Besides, using a single channel is intended to reduce computational complexity. The main purpose of this article is to detect driver drowsiness using a single-channel P4 sensor. This simplifies computation, reduces cost, and is less intrusive to drivers because data is collected from only one channel. Further, a single channel drowsiness detection method is more practical in real-world, as placing many electrodes requires expert knowledge about electrode locations, an inefficient approach.

In our previous study[31] drowsiness was detected through a convolutional neural network based on the detect alpha spindle. It also needed to compute the continuous wavelet analysis of the signal for detection of alpha spindles. Due to its heavy computations, the previous method has a notable delay in the real-world application. In the detection of driver drowsiness, time has a significant role because the vehicle has dynamic, and it takes time to control it, so drowsiness must be detected without any considerable delay and as fast as possible. This proposed method can detect drowsiness based on features of the raw EEG signals; therefore, its computation can be done faster. In consequence, this method is more desired, practicable, and feasible in real-world and online applications.

Several features made this study unique, and it hardly can be compared with other studies in the field of detect drowsiness based on EEG signal:

- This study uses ORD to measure drowsiness level, which in contrast with other methods, doesn't disturb the driver and doesn't affect their drowsiness level. In ORD method, driver drowsiness level measured based on the scores that three dedicated expert and trained observers assigned to the driver based on their facial and driving behavioral signs. All other studies used intrusive methods for measure drowsiness level. They use reaction-time based method or driver self-assessment method
    - Most of the other studies try to measure driver drowsiness level based on their reaction time and ignoring the important fact that these events increase the awareness level of drivers. In fact, when an event happens the reaction time of driver is completely related to him/her driving skills and they are not in a drowsy state anymore
    - The other studies use driver self-assessment method, ask driver to assess his/her drowsiness level, this method obviously increase driver awareness level and can be considered as a warning method to prevent drowsiness.
- All other studies are consisting of event-reaction-based scenarios. However, most of the driver drowsiness accidents not happen in the crowded roads with lots of obstacles and usually happen in boring monotonous roads. Our driving simulator tests simulated more realistic conditions of accidents due to drowsiness
- The moderate drowsiness detected in this study, whereas the drowsiness studies usually aim to detect fatigue at a level after moderate drowsiness, in fatigue state ORD = 4 while in moderate drowsiness ORD = 3
- Last, but not least, in this study data were analyzed based on ORD, in other words, based on the level of drowsiness. Because time does not provide a reliable indicator of driving drowsiness.

This study achieved a high performance in detecting moderate drowsiness, especially the result of using single-channel P4 is entirely desirable. The moderate drowsiness level is a level before fatigue. At the fatigue level, the drivers start to lose control of the vehicle, and symptoms of drowsiness appear in the vehicle dynamic data features such as lateral position. But before that, at moderate drowsiness, there are very weak symptoms of drowsiness in the vehicle dynamic data. However, drowsiness symptoms of drowsiness are detectable in EEG signals from the beginning of drowsiness. In this study, the classification results based on EEG data were compared to vehicle dynamic data to obtain a more accurate assessment of the results. The results of the classification based on EEG signal show incredibly higher performance than the classification based on vehicle dynamic data in detecting moderate drowsiness. So, these results show EEG is more powerful to detect drowsiness earlier than vehicle dynamic data.

Conclusion

In this study, the EEG signals were used for driver drowsiness detection at the moderate level of drowsiness (ORD = 3). EEG data were analyzed in respect to ORD instead of time. Several supervised learning methods were used to classify EEG features, including regression-based and multiclass classifiers. Nineteen healthy controls participated in driving simulator tests. The EEG data were collected from the (C3, C4, P3, P4, O1, O2) channels, and the reference channel was Cz. The EEG signals are divided into 30 s epochs, and ORD was used to measure drowsiness levels for each EEG epoch. Five features, including mean, standard deviation, kurtosis, energy, and entropy, were extracted for each epoch and used as input to classifiers. Results showed that the classification tree and regression classifiers were 87 percent accurate with all channels and 91% accurate with single-channel P4. These results show Drowsiness can be detected most accurately based on single-channel P4 data. Besides, the result of comparing drowsiness detection methods based on vehicle dynamic data and EEG data show, EEG data have significantly higher potential to detect drowsiness at the moderate level. Single-channel P4 provide an opportunity to develop an EEG drowsiness detection device that is practical and applicable. Because P4 is located near the ear and auditory cortex, it is possible to in feature design a drowsiness detection device that collects data specifically from P4. Furthemore, in feature works, In-Ear EEG devices can use to collect data from the P4 channel to detect driver drowsiness in real-world use.

Acknowledgments

The authors would like to express their gratitude and appreciation for Nahvi A and Mahmoodi M, whose guidance, support, and encouragement have been invaluable throughout this study. We also wish to thank the KNTU Virtual Reality lab team, who has been a great source of support.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

References

1. Williams LR, Davies DR, Thiele K, Davidson JR, MacLean AW. Young drivers' perceptions of culpability of sleep-deprived versus drinking drivers. J Safety Res 2012;43:115-22.
2. Sahayadhas A, Sundaraj K, Murugappan M. Detecting driver drowsiness based on sensors: A review. Sensors (Basel) 2012;12:16937-53.
3. Mahmoodi M, Nahvi A. Driver drowsiness detection based on classification of surface electromyography features in a driving simulator. Proc Inst Mech Eng H 2019;233:395-406.
4. Noori SM, Mikaeili M. Driving drowsiness detection using fusion of electroencephalography, electrooculography, and driving quality signals. J Med Signals Sens 2016;6:39-46.
5. Borghini G, Astolfi L, Vecchiato G, Mattia D, Babiloni F. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci Biobehav Rev 2014;44:58-75.
6. Mardi Z, Ashtiani SN, Mikaili M. EEG-based drowsiness detection for safe driving using chaotic features and statistical tests. J Med Signals Sens 2011;1:130-7.
7. Tuladhar AM, Huurne NT, Schoffelen JM, Maris E, Oostenveld R, Jensen O. Parieto-occipital sources account for the increase in alpha activity with working memory load. Hum Brain Mapp 2007;28:785-92.
8. Keune PM, Hansen S, Sauder T, Jaruszowic S, Kehm C, Keune J, *et al*. Frontal brain activity and cognitive processing speed in multiple sclerosis: An exploration of EEG neurofeedback training. Neuroimage Clin 2019;22:101716.
9. Rahma ON, Rahmatillah A. Drowsiness analysis using common spatial pattern and extreme learning machine based on electroencephalogram signal. J Med Signals Sens 2019;9:130-6.
10. Akin M. Comparison of wavelet transform and FFT methods in the analysis of EEG signals. J Med Syst 2002;26:241-7.
11. Shaker MM. EEG waves classifier using wavelet transform and Fourier transform. Brain 2006;2:169-174.
12. Li M, Chen W, Zhang T. Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble. Biomed Signal Proc Control 2017;31:357-65.
13. Sharmila A, Geethanjali P. DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers. Ieee Access 2016;4:7716-27.
14. Ghorbanian P, Devilbiss DM, Verma A, Bernstein A, Hess T, Simon AJ, *et al.* Identification of resting and active state EEG features of Alzheimer's disease using discrete wavelet transform. Ann Biomed Eng 2013;41:1243-57.
15. Shourie N. Cepstral analysis of EEG during visual perception and mental imagery reveals the influence of artistic expertise. J Med Signals Sens 2016;6:203-17.
16. Imani E, Pourmohammad A, Bagheri M, Mobasheri V. ICA-based imagined conceptual words classification on EEG signals. J Med Signals Sens 2017;7:130-44.
17. Kashefpoor M, Rabbani H, Barekatain M. Automatic diagnosis of mild cognitive impairment using electroencephalogram spectral features. J Med Signals Sens 2016;6:25-32.
18. Hatamikia S, Maghooli K, Nasrabadi AM. The emotion

recognition system based on autoregressive model and sequential forward feature selection of electroencephalogram signals. J Med Signals Sens 2014;4:194-201.

19. Rasekhi J, Mollaei MR, Bandarabadi M, Teixeira CA, Dourado A. Epileptic seizure prediction based on ratio and differential linear univariate features. J Med Signals Sens 2015;5:1-11.
20. Wang H, Zhang Y. Detection of motor imagery EEG signals employing Naïve Bayes based learning process. Measurement 2016;86:148-58.
21. Prasad PD, Halahalli HN, John JP, Majumdar KK. Single-trial EEG classification using logistic regression based on ensemble synchronization. IEEE J Biomed Health Inform 2014;18:1074-80.
22. Simon M, Schmidt EA, Kincses WE, Fritzsche M, Bruns A, Aufmuth C, *et al.* EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. Clin Neurophysiol 2011;122:1168-78.
23. Johnson RR, Popovic DP, Olmstead RE, Stikic M, Levendowski DJ, Berka C. Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. Biol Psychol 2011;87:241-50.
24. da Silveira TL, Kozakevicius AJ, Rodrigues CR. Automated drowsiness detection through wavelet packet analysis of a single EEG channel. Expert Syst Appl 2016;55:559-65.
25. Picot A, Charbonnier S, Caplier A. Monitoring drowsiness on-line using a single encephalographic channel. Carlo Alexandre Barros Mello 2009;1:145-164.
26. Wright KP Jr., Badia P, Wauquier A. Topographical and temporal patterns of brain activity during the transition from wakefulness to sleep. Sleep 1995;18:880-9.
27. Lin CT, Wu RC, Liang SF, Chao WH, Chen YJ, Jung TP. EEG-based drowsiness estimation for safety driving using independent component analysis. IEEE Trans Circuits Syst I Regul Pap 2005;52:2726-38.
28. Chuang CH, Huang CS, Ko LW, Lin CT. An EEG-based perceptual function integration network for application to drowsy driving. Knowl Based Syst 2015;80:143-52.
29. Correa AG, Orosco L, Laciar E. Automatic detection of drowsiness in EEG records based on multimodal analysis. Med Eng Phys 2013;36:244-9.
30. Hajinoroozi M, Mao Z, Jung TP, Lin CT, Huang Y. EEG-based prediction of driver's cognitive performance by deep convolutional neural network. Signal Proc Image Commun 2016;47:549-55.
31. Houshmand S, Kazemi R, Salmanzadeh H. A novel convolutional neural network method for subject-independent driver drowsiness detection based on single-channel data and EEG alpha spindles. Proc Inst Mech Eng H 2021;235:1069-78.
32. Yang W, Wang K, Zuo W. Neighborhood component feature selection for high-dimensional data. J Comput 2012;7:161-8.
33. Akansu AN, Haddad RA. Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets. Boston, MA: Academic Press; 1992.
34. Kotsiantis SB, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering. 2007;160:3-24.
35. Kotsiantis SB, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. Emerg Artif Intell Appl Comput Eng 2007;160:3-24.
36. Anund A, Kircher A, Tapani A. The effect of milled rumble strips versus virtual rumble strips on sleepy drivers: A driving simulator study. Swedish Road Administration and EU; 2009.
37. Na S, Li Z, Qiu F, Zhang C. Torque control of electric power steering systems based on improved active disturbance rejection control. Math Probl Eng 2020;2020:6509607.