Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2021) 29: 1429 – 1444

doi:10.3906/elk-2008-83

Using EEG to detect driving fatigue based on common spatial pattern and support vector machine

Li WANG, David JOHNSON, Yingzi LIN*
Department of Mechanical and Industrial Engineering, College of Engineering,
Northeastern University, Boston, USA

Received: 19.08.2020 • **Accepted/Published Online:** 01.12.2020 • **Final Version:** 31.05.2021

Abstract: To investigate the correlation between electroencephalogram (EEG) and driving fatigue states, this study used machine learning algorithms to detect driving fatigue based on EEG. 14 channels of EEG data were collected from thirty-four healthy subjects in this research at Northeastern University. Each subject participated in two scenarios (baseline and fatigue scenarios). Subjective ratings of fatigue levels were also obtained from the subjects using the NASA-Task Load Index (TLX). The common spatial pattern (CSP) algorithm was used to extract features from the raw EEG data. The support vector machine (SVM) was used as the classifier in the design of the machine learning algorithm. A grid search cross validation was exploited to find optimal hyperparameter settings. The best classification result was 90%, obtained by using all 14 EEG channels and linear kernel of SVM. The experimental results proved that a machine learning algorithm was able to reliably classify driving fatigue states using EEG data. This study demonstrated that CSP and SVM were promising in detecting driving fatigue, and therefore, they could be strong foundations for future efforts to reduce traffic accidents and save thousands of human lives.

Key words: Driving fatigue, driving safety, electroencephalogram, common spatial pattern, support vector machine

1. Introduction

Many factors affect a driver's ability to perform well and avoid mistakes, such as fatigue, illness, stress, and many others [1]. Of these, fatigue is the most significant and avoidable factor. In the United States from 2005 to 2009, there were an estimated 83,000 crashes each year related to drowsy driving [2]. In 2014, there were 846 drowsiness-driving related fatalities, which comprised two percent of the overall total of driving fatalities [2].

Fatigued driving is common across significant groups of drivers. Preventing driving fatigue has many significant aspects. Of individuals who reported being fatigued while driving, nearly half were only driving for one hour or less. While drowsy driving affects over a third of all drivers, drowsy driving is also common among commercial motor vehicle (CMV) drivers [2]. A U.S. Department of Transportation study showed that 13 percent of CMV drivers were considered to have been drowsy at the time of their crashes [3]. This is concerning in terms of how large and damaging commercial vehicles can be, and the large number of hours those vehicles and their drivers spend on the road at one time. By detecting and preventing drowsy driving, accidents related to fatigued driving of CMVs can be considerably reduced.

In addition, the costs of fatigue-related accidents can be reduced by identifying fatigue. The National Highway Traffic Safety Administration (NHTSA) estimates that these accidents account for over $12 billion in

*Correspondence: yi.lin@northeastern.edu

damages each year, which come in the form of increased insurance premiums, damage to vehicles and property, injuries to the driver or others involved in the accident, or death. Numerous studies have been done to improve driving conditions, such as intelligent human assistance systems, smart driving, and driving cognitive assistance, etc. [4–7]. It is essential that we create a machine learning algorithm to detect driving fatigue.

The current literature surrounding use of electroencephalogram (EEG) to detect fatigue is inconsistent, and there is little consensus on what specific factors indicate fatigue. In the studies reviewed, it has been suggested that:

- Delta and theta waves increase with fatigue increasing [8],
- Alpha spindle parameters increase with fatigue increasing [9],
- Alpha band relative energy increases with fatigue increasing [10],
- Alpha wave increases with fatigue increasing [11],
- Alpha band decreases with extreme fatigue increasing [12],
- A change in Delta and theta waves indicates a transition to a fatigue state [13]

The findings of some of these studies contradict one another, making it evident that there is little consensus on the individual factors that best identify driving fatigue. Because of the lack of consensus, this study hypothesized that it would be advantageous to use a method that relies on a common spatial pattern algorithm (CSP) and support vector machine (SVM) to analyze the raw data independently, instead of relying on any specific groups of wavebands and limited human bias.

The CSP algorithm has been reliably used for dimension reduction since 1990 and it has been widely used with multichannel EEG data [14–17]. Some studies have used CSP to reduce EEG data dimensionality and extract features in brain-computer interface applications [18, 19]. Another study used CSP to decode movement target direction in nonhuman primates [20]. CSP has been demonstrated to be a successful method for dimensionality reduction.

SVM has been widely used to classify many types of bio-signals [21–23]. It has been shown that SVM has performed well in detecting contaminants in electromyography (EMG) signals [24]. Some studies have detected seizures successfully using EEG data based on SVM [25, 26]. Another study used the SVM to create a recognition model to detect driving anger [27]. Hundreds of studies have shown that SVM was a valid method to classify EEG data.

In this paper, we propose a machine learning method for detecting driving fatigue using EEG data collected on 14 channels. To achieve a better signal quality, independent component analysis was used to remove noise in the signal. CSP was used as the feature extraction method, and SVM was used as the classifier. The remaining parts of this paper are organized as follows: Section 2 shows the material and method of this study. Section 3 represents the experimental results. Section 4 is the discussion and conclusion of the research. Some main contributions of this work are as follows:

- The present study achieved a classification accuracy of 90% by using CSP and SVM for driving fatigue detection.
- CSP proved to be an efficient and victorious technique for EEG dimensionality reduction and feature extraction.
- It proved that EEG could be a promising tool for driving fatigue detection in real driving tasks.

2. Material and methods

2.1. Apparatus

A driving simulator in the Intelligent Human-Machine Systems (IHMS) Laboratory at Northeastern University was used to set up a virtual environment for the experiment. A 14-channel Emotiv EPOC+ (San Francisco, California) was used for EEG data collection at a sample rate of 128 Hz. The 14 channels are listed as AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 according to the 10-20 system [28], as shown in Figure 1. The driving simulator used in this study was from extreme competition controls Inc (ECCI). It is a mid-level unit comprised of two pedals (brake and accelerator) and a steering wheel. The virtual driving environment was created in Unity 3D (San Francisco, CA), as shown in Figure 2.

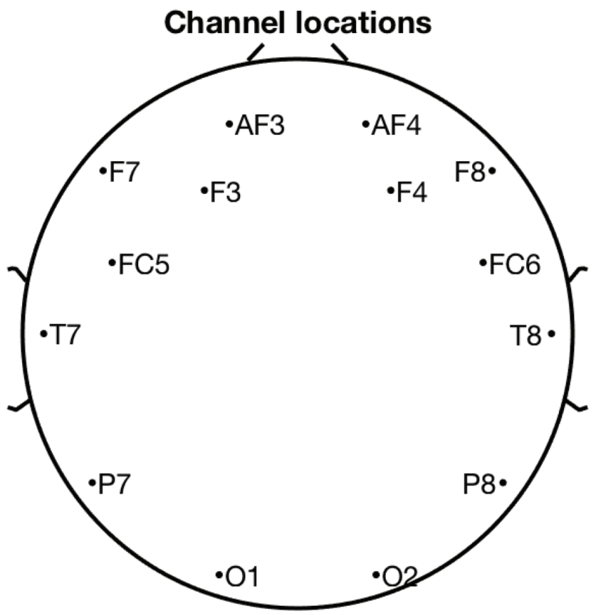

Figure 1. The EEG channel locations used in this study.

2.2. Experimental design

There were two driving scenarios designed for this experiment. The purpose of the first scenario was to collect baseline data on the subjects, while the second scenario was designed to trigger driving fatigue. This experimental design was approved by the Northeastern University IRB (#17-01-25).

a) Baseline/nonfatigue scenario: In the baseline scenario, subjects were asked to drive along the outer perimeter roads in the virtual environment as shown in Figure 3. The speed limit for this scenario was 60 MPH. This scenario lasted for 5 min.

b) Fatigue scenario: In the fatigue scenario, subjects were asked to follow the road signs in the virtual environment with repeating random turns, simulating completing mundane tasks to help create a boring environment that invited a quicker onset of fatigue. The speed limit for this scenario was also 60 MPH. This scenario lasted for 25 min. Some research suggested that 30 min driving is enough to trigger fatigue [29]. The road design of the fatigue scenario is shown in Figure 4.

Figure 2. Driving simulator and virtual environment.

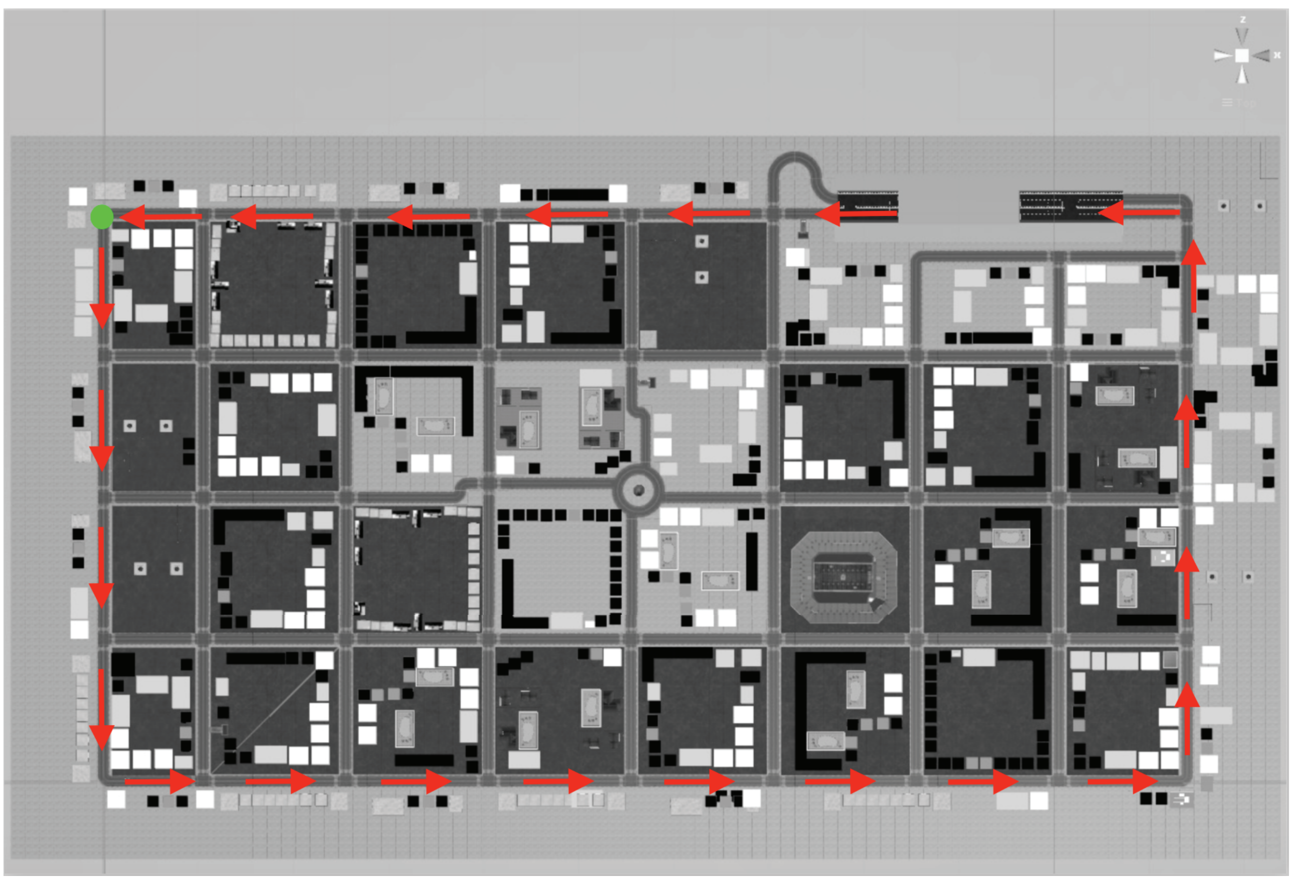

Figure 3. Top view of virtual environment (road map design of the baseline scenario). Green dot: starting point; red arrows: driving direction.

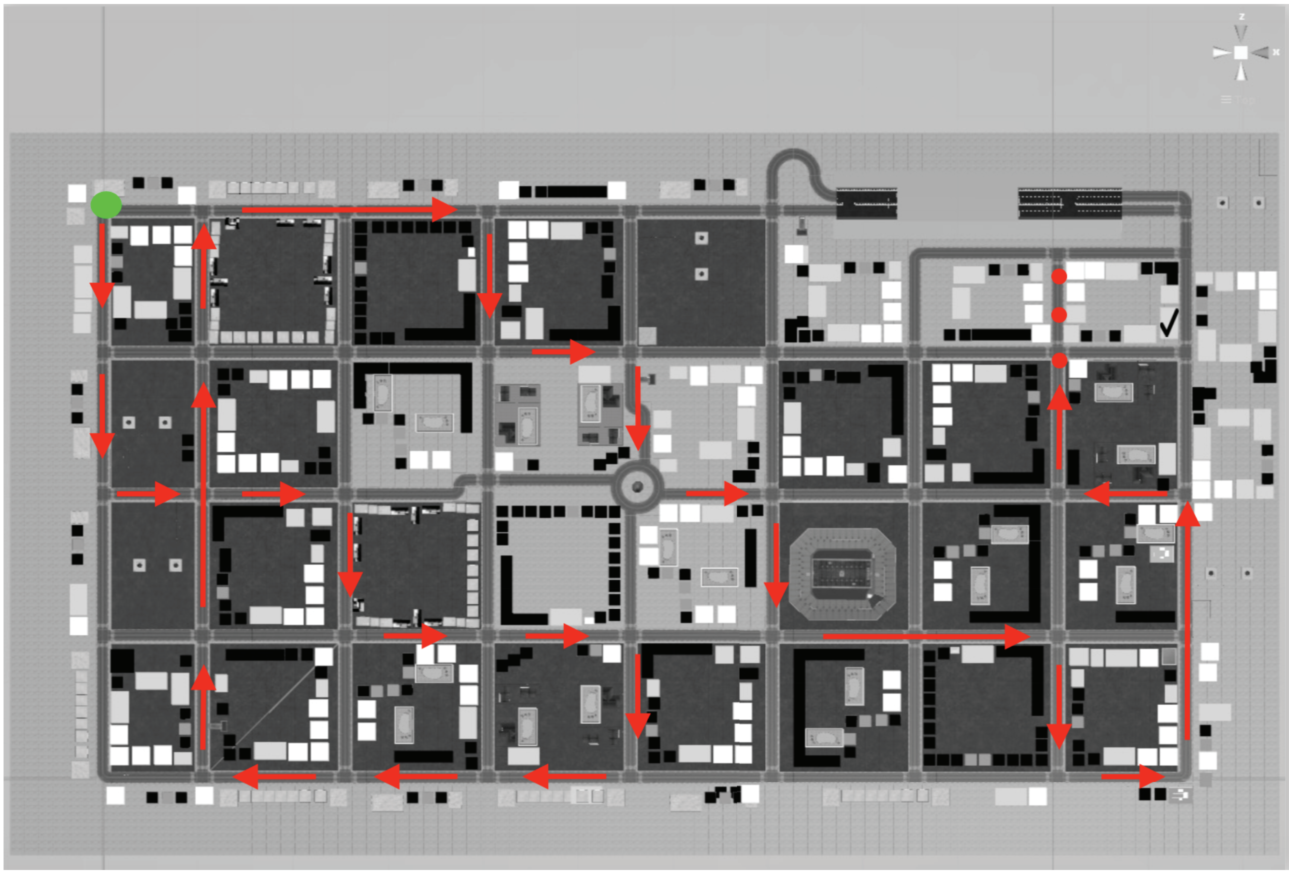

Figure 4. Top view of virtual environment (road map design of the fatigue scenario). Green dot: starting point, red arrows: a demonstration of the random driving direction.

2.3. Subject selection

Thirty-four subjects took part in this study (age mean = 22.5 years, SD = 2.09 years, 10 females). All subjects were asked to sign a consent form before the experiment. The subjects were required to have a valid US driver's license to take part in this experiment. The subjects were instructed not to consume any caffeinated beverages, such as coffee or energy drinks, and avoid similar stimulants for 24 h prior to the experiment. The subjects were advised to have a meal heavy in carbohydrates before the experiment since some research demonstrated that carbohydrate-heavy meals help to trigger fatigue [30, 31]. The experiments took place in the evening, after working hours.

2.4. Data acquisition

In order to measure fatigue, we asked the subjects to rate their fatigue level on a scale of 0–10, with 10 being the most fatigued and 0 being not at all fatigued, at two times during the study. The first time was right before the subjects sat down to begin driving. This was to set a baseline for each subject. The second time was immediately after completion of the driving task, in order to gauge the level of fatigue induced by driving in the simulated environment. The self-rated fatigue level survey was based on the NASA-Task Load Index (TLX) [32].

The EEG signal was collected across 14 channels by using an Emotiv device (Emotiv Inc., San Francisco, CA, USA) set to a sampling rate of 128 Hz with 2 reference channels. The EEG locations were based on the 10–20 system [28]. Each recording was split into two sectors. The first 3 min were placed in the first sector, and the last 3 min were placed in the second sector. The first sector was used as a representation of baseline data, and the second sector was used as a representation of fatigue data. In total, 6 min of EEG data from each subject were used for further preprocessing.

2.5. Data preprocessing

Using the raw EEG data of the 34 subjects from which to pull from, we selected 200 baseline data samples and 200 fatigue data samples, making 400 total samples. Baseline samples were selected from each subject's first sector data, and fatigue samples were selected from each subject's second sector data. The 400 data samples were selected based on visual inspection of the EEG data. The EEG data with obvious artifacts were removed, as shown in Figure 5. Each of the 400 samples contained 10 s of the data (1280 data points) for each of the 14 channels. The EEG data were filtered by using a band-pass filter from 1 Hz to 40 Hz. Signal noises caused by eye blinks, eye movements were removed by using independent component analysis (ICA) in Matlab with the EEGLAB toolkit (Mathworks, Natick, MA; UC San Diego, CA, USA). The EEG data were also converted to the frequency domain with the EEGLAB toolbox for preliminary analysis.

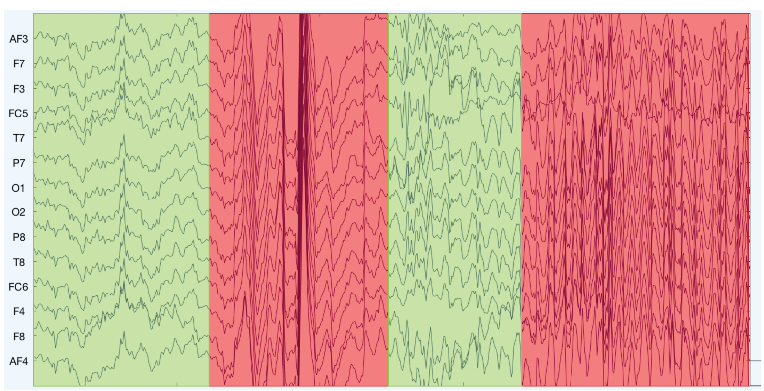

Figure 5. Five seconds of EEG data sample (red parts: artifacts, removed from the data; green parts: kept for further analysis).

2.6. Feature extraction

The features of the EEG data were extracted using CSP. The steps of the CSP method are shown below:

An $N \times T$ matrix V represents the raw data of a single EEG sample, in which N is the number of recording channels and T is the number of data points in each channel. The normalized spatial covariance,

$$R_i = \frac{V_i V_i^{'}}{trace(V_i V_i^{'})} \tag{1}$$

where $V_i^{'}$ is the transpose of matrix V_i, $i \in \{B(baseline), F(fatigue)\}$ and $trace(X)$ is the sum of the diagonal elements of matrix. The composite spatial covariance R can be obtained from,

$$R = \overline{R_B} + \overline{R_F} \tag{2}$$

where $\overline{R_1}$ is the average normalized covariance of data samples in one class and $\overline{R_2}$ is the average normalized covariance of data samples in the other class. The composite spatial covariance R can be factored as,

$$R = U_0 \Sigma U_0^{'} \tag{3}$$

where U_0 is the matrix of the eigenvector and Σ is the diagonal matrix of eigenvalues. The whitening transformation matrix P can be represented as,

$$P = \Sigma^{-1/2} U_0^{'} \tag{4}$$

Transform the average covariance matrices as,

$$S_B = P\overline{R_B}P^{'}, \; S_F = P\overline{R_F}P^{'} \tag{5}$$

And by factoring S_B and S_F,

$$S_B = U\Sigma_B U^{'}, \; S_F = U\Sigma_F U^{'} \tag{6}$$

where Σ_B and Σ_F are the eigenvalue matrices of S_B and S_F respectively. S_B and S_F share the same common eigenvectors U, and

$$\Sigma_B + \Sigma_F = I \tag{7}$$

In order to simplify the eigenvector matrix U, only the first and last five eigenvectors are selected to build the eigenvector matrix U_1 based on the ranking of their corresponding eigenvalues in a descending order. The projection matrix can be calculated from

$$W = U_1 P \tag{8}$$

By applying the projection matrix to the raw EEG data,

$$Z_B = WV_B, \; Z_F = WV_F \tag{9}$$

The covariance matrices (R_{ZB} and R_{ZF}) of Z_B and Z_F are

$$R_{ZB} = Z_B Z_B^{'}, \; R_{ZF} = Z_F Z_F^{'} \tag{10}$$

The final EEG signal features f_B, f_F are obtained by normalizing the covariance matrices R_{ZB} and R_{ZF}. The dimension of each sample of raw EEG data was 14 channels $\times$ 1280 data points. After the CSP filtering process, each feature was a 10×1 vector. The CSP algorithm was coded in Python 3.7 (Wilmington, DE, USA) programming language.

2.7. Classification

The SVM was initially introduced by Vapnik and colleagues in 1992 [33]. SVM is a two-class, supervised machine learning classification method, which is defined by a hyperplane [34, 35]. Figure 6 shows an example of a hyperplane generated by SVM. There were two classes of data points with the circled points as the support vectors in the plot.

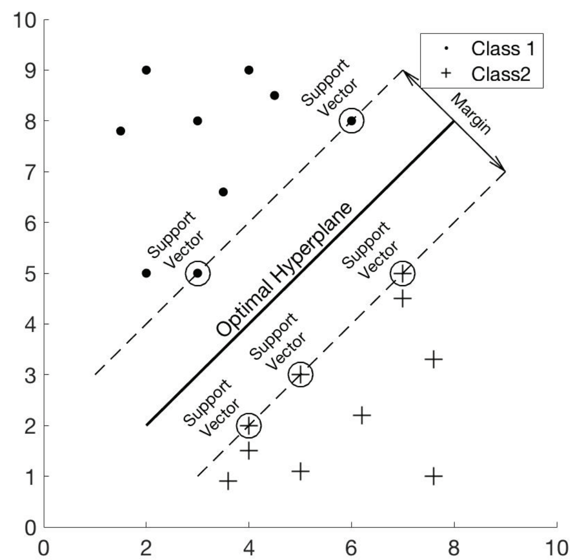

Figure 6. Description of SVM model, showing optimal hyperplane, support vectors, and margin.

The principle idea of SVM is to find an optimal hyperplane that separates two different data sets with the maximum margin possible between the two data sets. If we consider that there is a data set with two categories of data $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \cdots, (x_n, y_n)\}$, in which x_i is a sample and y_i is the label of x_i ($y_i \in \{-1, 1\}$), then the training process of SVM finds the parameters w and b of equation $y = wx + b$, which defines the hyperplane. To find the optimal hyperplane, the maximum margin value is solved by the equation:

$$\min_{w,b,\xi} \left(\frac{1}{2} ww' + C \sum_{i=1}^{k} \xi_i \right), \tag{11}$$

where ξ_i is the slack variable and C is the cost factor.

When there are many features, using kernel functions can make the SVM classification more efficient. Four kinds of kernels were explored in this study, namely linear, radial basis function (RBF), polynomial, and sigmoid kernels.

The training and test split was 80/20 for the present study. A grid search cross validation was utilized to find the optimal hyperparameters of the SVM classifier on the training set. A 5-fold cross validation was used in this grid search. Different types of hyperparameters were explored. For instance, C is the cost factor for the SVM. *Gamma* is the kernel coefficient for 'rbf', 'poly', and 'sigmoid' kernels. The settings for the SVM hyperparameters were shown in Table 1. Python 3.7 programming language was used for the SVM and grid search algorithm.

Table 1. The settings for the SVM hyperparameters.

Hyperparameter	settings (number of settings)	
Kernel type	'rbf', 'poly', 'sigmoid' (3)	'linear' (1)
C	0.1, 1, 10, 100, 1000 (5)	0.1, 1, 10, 100, 1000 (5)
Gamma	1, 0.1, 0.01, 0.001, 0.0001 (5)	*N/A*
	(3)*(5)*(5) = 75	(1)*(5) = 5

3. Results and discussion

3.1. Subject self-reported fatigue level

Subjects were asked to report their fatigue level before and after the driving tasks. Each subject's fatigue level was measured by a survey on a 0–10 scale based on the NASA-TLX, with 10 being the most fatigued, and 0 being not at all fatigued. The average fatigue level for the baseline was 3.9 and the fatigue scenario had an average fatigue level of 7. All subjects reported a higher fatigue level after the experiment, which demonstrated that the scenario design triggered fatigue successfully (paired T-test, $P < 0.01$, degrees of freedom (DOF)=33) (Minitab, Version17, State College PA). The fatigue level of each subject, both before and after the driving task, is shown in Figure 7.

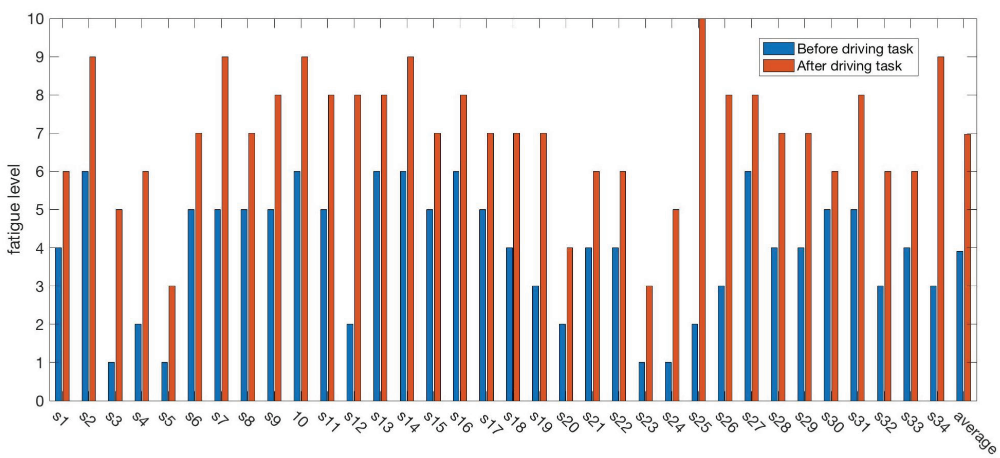

Figure 7. The fatigue level of each subject (before and after the driving task).

All the data collection happened in the late afternoon and in the evening, which was beneficial for triggering driving fatigue. Subjects' self-reported fatigue level was used as a reference to see if the driving tasks created driving fatigue in participants. Another study [11] used eye blinks as the indicator of driving fatigue. Using eye blinks to indicate driving fatigue could be inaccurate, since many other reasons could cause eye blinks, such as dry eye, refractive surgery, and dry contact lenses [36]. Another advantage of using the self-reported fatigue level from 0 to 10 is that we can design new classification models that can predict a fatigue level from anywhere within the 0 to 10 range that was desired.

3.2. Frequency domain analysis

The power spectrum plots and topographies of the EEG in baseline and fatigue states were generated with the EEGLAB toolbox. Figure 8 shows an example of five topographies and one power spectrum plot created for both baseline and fatigue EEG data. The 5 topographies were chosen from delta band (1–3 Hz), theta band (4–8 Hz), alpha band (8–13 Hz), beta band (13–30 Hz), and gamma band (30–50 Hz), respectively. The EEG power spectrum plots show that the power of the EEG signal under the fatigue state increased around 13 Hz to 15 Hz, compared to that under the baseline state. The EEG alpha band power increased under the fatigue state compared to the baseline state, which was consistent with [10] and [11] but contradicted [12]. According to [8], delta and theta band power increased with fatigue level increasing, which was not observed in this study. Based on the topographies, the EEG power near channel T7 escalated under the fatigue condition within the delta band. The theta and alpha band topographies did not show any change from the baseline state to the fatigue state. The EEG power fell near channel O1 from the baseline state to the fatigue state within the beta band. The gamma band topography showed an increase of EEG power near channel O2 and P8 under the fatigue state.

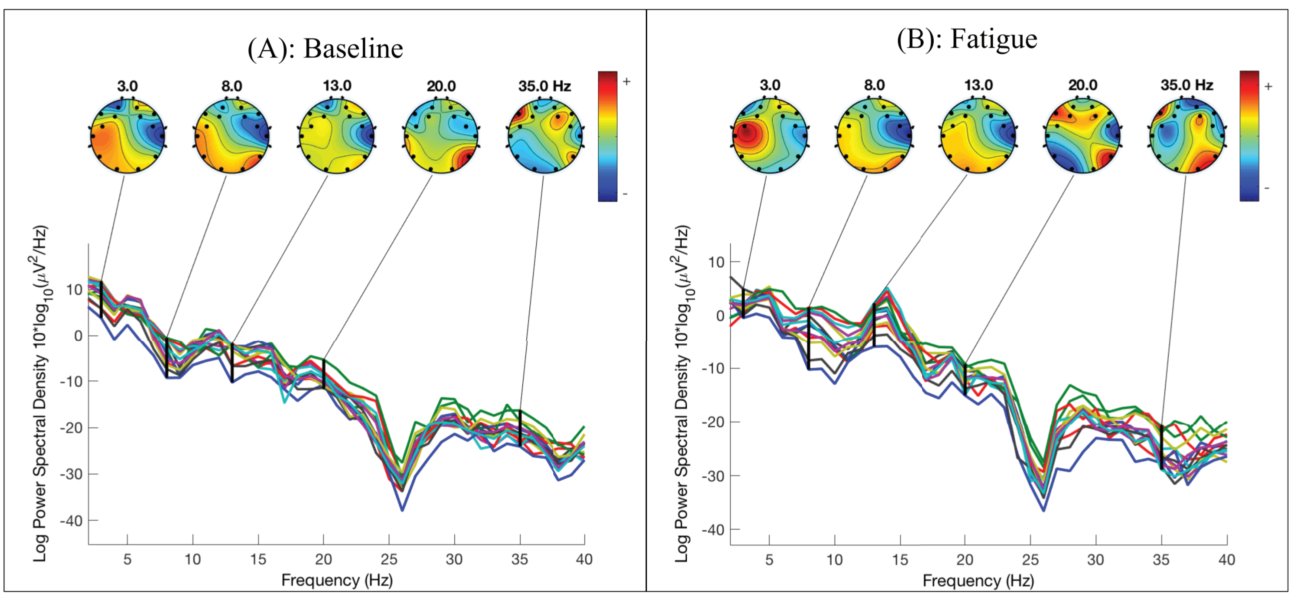

Figure 8. The power spectrum and topographies of the EEG data in baseline (A) and fatigue (B) states.

3.3. Extracted feature results

CSP was used in this study to extract the EEG features from the raw data. An example of fatigue and nonfatigue features is shown in Figure 9. After applying the transformation matrix to the raw data and then normalizing, it is clear that the feature value of the nonfatigue state has an increasing trend, while the feature value of the fatigue state has a decreasing trend.

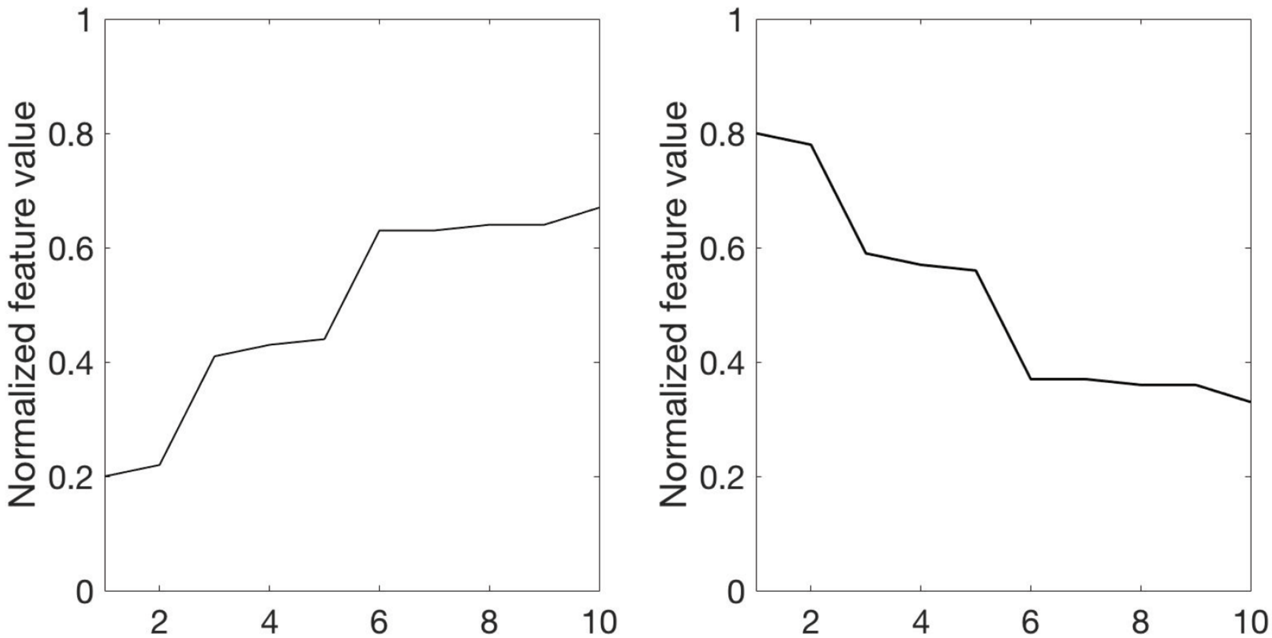

Figure 9. An example of fatigue and nonfatigue features extracted by CSP (left: feature of nonfatigue state; right: feature of fatigue state).

3.4. Classification results

3.4.1. Grid search

From Table 1, there were 80 combinations of the hyperparameter settings. The grid search ran 400 iterations on the 320 training samples with a 5-fold cross validation. The grid search cross validation on the training set shows that the optimal hyperparameter setting was 'linear' as kernel type and 10 as C (cost factor).

Four kinds of kernels were tested iin this study: linear, RBF, polynomial, and sigmoid. The performance scores of different kernel types during grid search are shown in Figure 10. During the grid search, the highest level of accuracy was obtained using the linear kernel at 87.12%. The RBF, polynomial, and sigmoid kernels had performance accuracies of 81.7%, 79.3%, and 79%, respectively in the grid search.

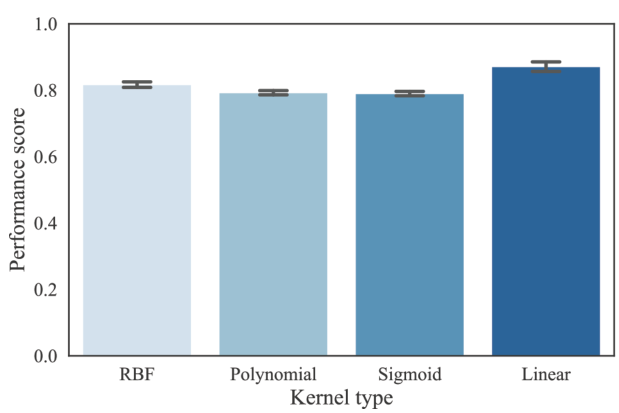

Figure 10. A bar plot showing the performance scores of different kernel types in grid search.

Figure 11 shows the running time of different kernel types during grid search. The linear kernel had the best running time performance of 0.001375 s per iteration. The polynomial, sigmoid, and RBF kernels had running time performance of 0.001672, 0.001676, and 0.001800 s per iteration respectively, in the grid search. It is well known that EEG is high-dimensional data, which made it hard for classifiers to perform the classification. The CSP algorithm tremendously reduced the dimensionality of the EEG data. As shown in Figure 9, the data features obtained from the CSP algorithm were linearly separable. Using the SVM linear kernel also speeded up the classification process.

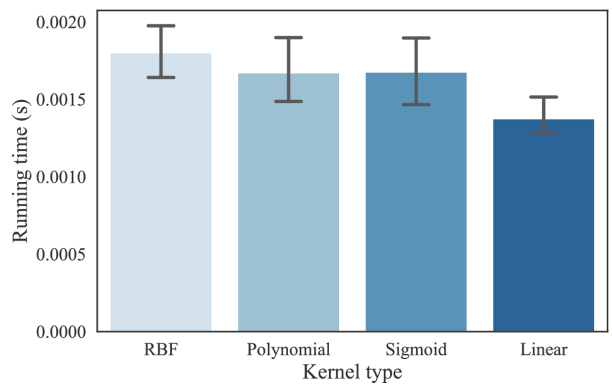

Figure 11. A bar plot showing the running time of different kernel types in grid search.

3.4.2. Final classification

Using the optimal hyperparameter settings ($C = 10$, linear kernel), we ran the SVM on the test data set. The performance accuracy on the test data set was 90%. The confusion matrix of the final classification results is shown in Figure 12.

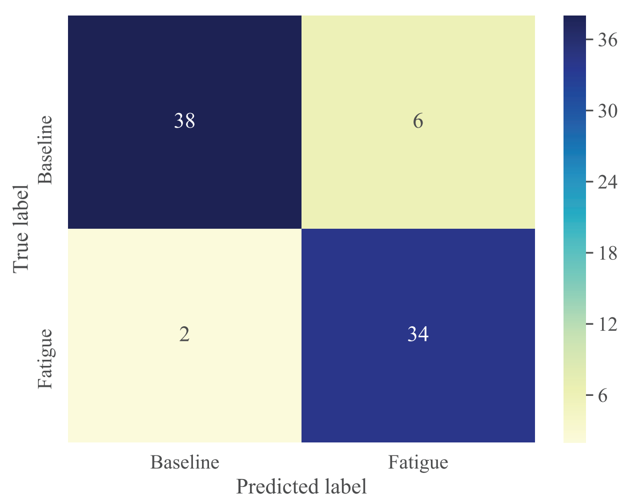

Figure 12. A confusion matrix showing the final classification results.

The final classification performance was also evaluated using f1-score, precision, sensitivity, and specificity. The evaluation results are shown in Table 2. We also compared our classification performance with other similar literature studies. The comparison of different studies is shown in Table 3. Chaudhuri and Routray obtained an accuracy of 86% by using standardized low-resolution brain electromagnetic tomography (sLORETA) and SVM, but the classification running time was not reported in this study. In [38], deep generic model (DGM)-SVM was utilized to classify EEG data. Their best classification accuracy was 73.29%, but the sensitivity and specificity were 91.10 % and 55.48 %, respectively, which means their algorithm was not very reliable in finding the nonfatigue samples. Xiong et al. [39] achieved a performance accuracy of 85% by using kernel principal component analysis (KPCA)-SVM with EEG signal. The KPCA method extracted 17 features for each sample and the average classification time was 0.2 s. The KPCA method was a common dimensionality reduction technique. However, CSP had a better performance in dimensionality reduction (10 features versus 17 features) and resulted in a faster classification speed. Hu et al. study [40] used independent component analysis with reference (ICA-R) and SVM with 40 features, accomplishing a classification accuracy of 86%; however, the average running time was not reported in this study. With the EEG signal, the classification accuracy was around 80% to 90%. However, by combining EEG with heart rate variability (HRV), [41] achieved a classification performance of 91%, which suggests using multiple signal modalities could potentially improve classification performance.

$$Precision = \frac{Truepositive}{Truepositive + Falsepositive} \tag{12}$$

$$Recall = \frac{Truepositive}{Truepositive + Falsenegtive} \tag{13}$$

$$F1 = \frac{2}{Precision^{-1} + Recall^{-1}} \tag{14}$$

Table 2. The evaluation of the final classification performance.

	F1-score	Precision	Sensitivity	Specificity
Baseline	0.9	0.95	0.86	0.94
Fatigue	0.89	0.85	0.94	0.86

Table 3. The comparison of different similar studies, regarding method, signal, running time, and accuracy.

	Method	Signal	Average running time (s)	Accuracy
This study	CSP + SVM	EEG	0.001375	0.9
[37] (2019)	sLORETA + SVM	EEG	N/A	0.86
[38] (2016)	DGM-SVM	EEG	N/A	0.7329
[39] (2013)	KPCA-SVM	EEG	0.2	0.85
[40] (2013)	ICA-R	EEG	N/A	0.86
[41] (2010)	SVM	EEG + HRV	N/A	0.91

The original EEG data included 14 channels (AF3, AF4, F7, F8, F3, F4, FC5, FC6, T7, T8, P7, P8, O1, O2) in total. However, fewer EEG channels mean faster computational speed, easier application for a real driving scenario, and cheaper EEG devices. To find out if fewer EEG channels can achieve better performance, different combinations of EEG channels were also tested in the present study. The first subset was from the frontal and temporal areas. The second subset was from the frontal and parietal areas. The third subset was from the frontal and occipital areas. The same classification techniques were applied to these subsets. Table 4 shows the classification accuracy of different combinations of channels that were tested in the study.

Table 4. Classification accuracy using different channel combinations.

Area	Channel name	Accuracy
Frontal + temporal area	AF3, AF4, F7, F8, F3, F4, FC5, FC6, T7, T8	87.5%
Frontal + parietal area	AF3, AF4, F7, F8, F3, F4, FC5, FC6, P7, P8	81.25%
Frontal + occipital area	AF3, AF4, F7, F8, F3, F4, FC5, FC6, O1, O2	82.5%
All 14 channels	AF3, AF4, F7, F8, F3, F4, FC5, FC6, T7, T8, P7, P8, O1, O2	90%

With all of the channels used, the classification accuracy was 90%, while other combinations of channels output lower levels of accuracy with limited variances between them. Reducing the number of EEG channels didn't improve classification accuracy. It is suggested that more EEG channels should be collected in future studies, and more combinations of channels should be tested. It has so far been demonstrated that the more channels that are used, the higher the accuracy that the model produces.

4. Limitations and conclusion

The present study investigated the effects of driving fatigue on EEG responses in a virtual environment. It showed that EEG with machine learning algorithms can detect drivers' driving fatigue successfully. However, there are a few limitations in this study. First, more channels of EEG are recommended in future studies. Second, 400 data samples (200 baseline data samples and 200 fatigue data samples) were obtained from 34 subjects in this research. In future studies, it is recommended to use a larger number of data samples and also recruit elderly people. Third, future studies should focus on real driving scenarios instead of using driving simulators.

In this study, a machine learning algorithm was created to classify EEG signals under baseline and fatigue states. CSP was used as the feature extraction method and SVM was used as the classifier. A grid search cross validation was utilized to find the optimal hyperparameter settings for the SVM. The highest accuracy was 90% using self-collected EEG data in this study. It demonstrated that the machine learning method was a promising way to detect driving fatigue and help decrease traffic accidents caused by driving fatigue.

Acknowledgment

The work has been supported by US National Science Foundation (NSF) grant #1838796, 0954579 and 133524. The authors would like to thank all the members who have participated in this work. Y.L. gave the idea, L.W. did the experiments, L.W., D.J., and Y.L. interprated the results and wrote the paper.

References

[1] Taylor AH, Dorn L. Stress, fatigue, health, and risk of road traffic accidents among professional drivers: the contribution of physical inactivity. Public Health 2006; 27: 371-391.

[2] National Highway Traffic Safety Administration. Traffic Safety Facts: Drowsy Driving (Rep. No. DOT-HS-811-449). Washington, DC, USA: NHTSA. 2011.

[3] Craft R, Blower D. The large truck crash causation study. In: 2004 R&T Stakeholder Forum, Arlington, VA. 2004.

[4] Yang G, Lin Y, Bhattacharya P. A driver fatigue recognition model based on information fusion and dynamic Bayesian network. Information Sciences 2010; 180 (10): 1942-54.

[5] Lin Y, “i-DRIVE, are we ready?,” Journal of Automotive Safety and Energy 2016; 7 (1):14-24.

[6] Lin Y. Toward intelligent human machine interactions. Mechanical Engineering 2017; 139 (06): S4-8.

[7] Cai H, Lin Y. Coordinating cognitive assistance with cognitive engagement control approaches in human–machine collaboration. IEEE Transactions on Systems, Man, and Cybernetics 2011; 42 (2): 286-94.

[8] Lal SK, Craig A. A critical review of the psychophysiology of driver fatigue. Biological Psychology 2001; 55(3): 173-94.

[9] Simon M, Schmidt EA, Kincses WE, Fritzsche M, Bruns A et al. EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. Clinical Neurophysiology 2011; 122 (6): 1168-78.

[10] Kar S, Bhagat M, Routray A. EEG signal analysis for the assessment and quantification of driver’s fatigue. Transportation Research Part F: Traffic Psychology and Behaviour 2010; 13 (5): 297-306.

[11] Yeo MV, Li X, Shen K, Wilder-Smith EP. Can SVM be used for automatic EEG detection of drowsiness during car driving?. Safety Science 2009; 47(1): 115-24.

[12] Lal SK, Craig A, Boord P, Kirkup L, Nguyen H. Development of an algorithm for an EEG-based driver fatigue countermeasure. Journal of Safety Research 2003; 34 (3): 321-8.

[13] Lal SK, Craig A. Driver fatigue: electroencephalography and psychological assessment. Psychophysiology 2002; 39 (3): 313-21.

[14] Lu H, Eng HL, Guan C, Plataniotis KN, Venetsanopoulos AN. Regularized common spatial pattern with aggregation for EEG classification in small-sample setting. IEEE Transactions on Biomedical Engineering 2010; 57 (12): 2936-46.

[15] Ramoser H, Muller-Gerking J, Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Transactions on Rehabilitation Engineering 2000; 8 (4): 441-6.

[16] Koles ZJ, Lazar MS, Zhou SZ. Spatial patterns underlying population differences in the background EEG. Brain Topography 1990; 2 (4): 275-84.

[17] Müller-Gerking J, Pfurtscheller G, Flyvbjerg H. Designing optimal spatial filters for single-trial EEG classification in a movement task. Clinical Neurophysiology 1999; 110 (5): 787-98.

[18] Ang KK, Chin ZY, Wang C, Guan C, Zhang H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. Frontiers in Neuroscience 2012; 6:39.

[19] Wang Y, Gao S, Gao X. Common spatial pattern method for channel selelction in motor imagery based brain-computer interface. In: 2005 IEEE Engineering in Medicine and Biology Conference; 2006. pp. 5392-5395

[20] Ince NF, Gupta R, Arica S, Tewfik AH, Ashe J et al. High accuracy decoding of movement target direction in non-human primates based on common spatial patterns of local field potentials. PloS One 2010; 5 (12): e14384.

[21] Tarabalka Y, Fauvel M, Chanussot J, Benediktsson JA. SVM-and MRF-based method for accurate classification of hyperspectral images. IEEE Geoscience and Remote Sensing Letters 2010; 7 (4): 736-40.

[22] Moavenian M, Khorrami H. A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification. Expert Systems with Applications 2010; 37 (4): 3088-93.

[23] Palaniappan R, Sundaraj K, Sundaraj S. A comparative study of the svm and k-nn machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals. BMC Bioinformatics 2014; 15 (1): 223.

[24] Fraser GD, Chan AD, Green JR, MacIsaac DT. Automated biosignal quality analysis for electromyography using a one-class support vector machine. IEEE Transactions on Instrumentation and Measurement 2014; 63 (12): 2919-30.

[25] Liu Y, Zhou W, Yuan Q, Chen S. Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2012; 20 (6): 749-55.

[26] Gardner AB, Krieger AM, Vachtsevanos G, Litt B. One-class novelty detection for seizure analysis from intracranial EEG. Journal of Machine Learning Research 2006; 7 (7): 1025-44.

[27] Wan P, Wu C, Lin Y, Ma X. A recognition model of driving anger based on physiological features by ROC curve analysis. In: Transportation Research Board 95th Annual Meeting, Washington DC, USA; 2016. pp.16-5500

[28] Jasper HH. The ten-twenty electrode system of the International Federation. Electroencephalography and Clinical Neurophysiology 1958; 10: 370-5.

[29] Artaud P, Planque S, Lavergne C, Cara H, Tarriere C, et al. An on-board system for detecting lapses of alertness in car driving. In: International Technical Conference on the Enhanced Safety of Vehicles; 1995. pp. 350-359.

[30] Spring B, Maller O, Wurtman J, Digman L, Cozolino L. Effects of protein and carbohydrate meals on mood and performance: interactions with sex and age. Journal of Psychiatric Research 1982; 17 (2): 155-67.

[31] Lieberman HR. Behavioral Changes Caused by Nutrients1. Nutrition and Neurobiology 1986; 38: 219-224

[32] Hart, SG, Lowell ES. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology 1988; 52: 139-183.

[33] Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In Proceedings of the fifth annual workshop on Computational Learning Theory 1992; 7(1): 144-152.

[34] Vapnik V. The nature of statistical learning theory. Springer Science & Business Media; 2013.

[35] Cortes C, Vapnik V. Support vector machine. Machine learning 1995; 20 (3): 273-97.

[36] McMonnies CW. Incomplete blinking: exposure keratopathy, lid wiper epitheliopathy, dry eye, refractive surgery, and dry contact lenses. Contact Lens and Anterior Eye 2007; 30 (1): 37-51.

[37] Chaudhuri A, and Routray A. Driver fatigue detection through chaotic entropy analysis of cortical sources obtained from scalp EEG signals. IEEE Transactions on Intelligent Transportation Systems 2019; 21 (1): 185-198.

[38] San PP, Ling SH, Chai R, Tran Y, Craig A et al. EEG-based driver fatigue detection using hybrid deep generic model. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2016. pp. 800-803.

[39] Xiong YJ, Zhang R, Zhang C, Yu XL. A novel estimation method of fatigue using EEG based on KPCA-SVM and complexity parameters. Applied Mechanics and Materials 2013; 373: 965-969.

[40] Hu S, Zheng G, Peters B. Driver fatigue detection from electroencephalogram spectrum after electrooculography artefact removal. IET Intelligent Transport Systems 2013; 7 (1): 105-113.

[41] Zhang C, Yu X. Estimating mental fatigue based on electroencephalogram and heart rate variability. Polish Journal of Medical Physics and Engineering 2010; 16 (2): 67-84.