2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Vigilance Estimating in SSVEP-Based BCI Using Multimodal Signals*

Kangning Wang[1, 2], Shuang Qiu[2], Wei Wei[2], Chuncheng Zhang[2], Huiguang He[2], *Senior Member, IEEE*, Minpeng Xu[1, 3], *Member, IEEE*, and Dong Ming[1, 3], *Senior Member, IEEE*

***Abstract*—Brain-computer interface (BCI) is a communication system that allows a direct connection between the human brain and external devices. With the application of BCI, it is important to estimate vigilance for BCI users. In order to investigate the vigilance changes of the subjects during BCI tasks and develop a multimodal method to estimate the vigilance level, a high-speed 4-target BCI system for cursor control was built based on steady-state visual evoked potential (SSVEP). 18 participants were recruited and underwent a 90-min continuous cursor-control BCI task, when electroencephalogram (EEG), electrooculogram (EOG), electrocardiography (ECG), and electrodermal activity (EDA) were recorded simultaneously. Then, we extracted features from the multimodal signals and applied regression models to estimate vigilance. Experimental results showed that the differential entropy (DE) feature could effectively reflect the change of vigilance. The vigilance estimation method, which integrates DE and EOG features into the support vector regression (SVR) model, achieved a better performance than the compared methods. These results demonstrate the feasibility of our methods for estimating vigilance levels in BCI.**

I. Introduction

A brain-computer interface (BCI) is a system that measures central nervous system (CNS) activity and converts it into artificial output. It replaces, restores, enhances, supplements, or improves natural CNS output, and thereby changes the ongoing interactions between the CNS and its external or internal environment [1]. BCIs can be used to improve the quality of life of users with the loss of neuromuscular control [2] as well as provide help for healthy operators. Thereinto, steady-state visual evoked potential (SSVEP)-based BCI has excellent signal-to-noise ratio (SNR), high information transfer rate (ITR) and learnability [3, 4] and has received increasing interest. In the field of SSVEP-based BCI, the encoding methods and decoding methods of SSVEP are very important directions, whose goal is to improve performance [4]. Recently, some studies have found that the absence of vigilance is one of the most important factors leading to the decline of human performance in human-computer interaction, for example, driving tasks [5, 6]. Thus, in SSVEP-based BCI, as one human-computer interaction system, the performance of users may be affected by their vigilance. Therefore, it is necessary to explore the relationship between the performance and vigilance status of users under long-term SSVEP-based BCI tasks, and further to develop methods to estimate vigilance in BCI tasks.

Vigilance is a vital cognitive state and is usually defined as the ability of organisms to maintain long-term attention to stimuli or tasks [7]. In the past few years, some modalities for estimating vigilance levels have been used in simulated vehicle-driving tasks or flight tasks [5, 8]. Among these modalities, physiological signals have been found to be relevant for different vigilance levels [6]. Thereinto, electroencephalogram (EEG) is considered as a reliable indicator for vigilance estimation because of its advantages of high temporal resolution and relatively direct brain information presentation [9-11]. Some studies reported that the change in the cognitive state was usually accompanied by the significant changes of the power spectrum of EEG frequency bands, such as delta, theta, alpha, beta, and gamma bands [12, 13]. Shi et al. found that the power spectral density (PSD) in the posterior brain region is an effective feature for estimating the driver's vigilance states [10]. Based on PSD and differential entropy (DE) extracted from EEG signals, Wu et al. developed a novel regression network to detect driver's vigilance [14]. These studies proved the feasibility of using EEG signals to detect drivers' vigilance levels.

Electrooculogram (EOG) is another widely used physiological signal in the estimation of vigilance. It has the characteristics of easy setup and high SNR. Studies found that EOG features (e.g., slow eye movements and spontaneous eyelid closures) are efficient indicators for vigilance estimation. Ma et al. used EOG features to detect the vigilance level of subjects during a monotonous task [15]. Zhang et al. applied a novel electrode placement method on the forehead and extracted forehead EOG features to detect driving vigilance [16]. In short, some features of EOG signals could be used to represent the loss of vigilance in the driving or flight tasks.

Electrocardiography (ECG) signals contain information highly related to vigilance status. As one of the most useful features in ECG-based studies, heart rate variability (HRV) is an important indicator to evaluate the autonomic nervous system (ANS) status, such as fatigue status [17]. Recently, HRV has been adopted in vigilance estimation research and received a satisfactory performance [18]. In addition, electrodermal activity (EDA) has also been used in the estimating of vigilance. It has been shown to have a meaningful relationship to visual attention [19].

*Research supported in part by the Beijing Municipal Natural Science Foundation Grant 4214078, in part by the Beijing Science and Technology Program under Grant Z201100004420015, and in part by the National Natural Science Foundation of China under Grant 61976209, Grant 61906188. (Corresponding authors: Shuang Qiu. e-mail: shuang.qiu@ia.ac.cn).

[1]Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300072, China.

[2]Research Center for Brain-Inspired Intelligence, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing 100190, China.

[3]College of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China.

On the other hand, multimodal approaches for vigilance estimation have been developed. Awais et al. recognized the alert and drowsy state with high accuracy by using the support vector machine (SVM) from EEG and ECG signals [20]. Zheng et al. proposed a multimodal approach to estimate vigilance in a simulated driving environment [13]. The fusion of EEG features and EOG features has been utilized for estimating vigilance, and received an improved performance over EEG and EOG alone. These results demonstrated that different modalities contain complementary information and features from different modalities can be integrated to construct a more efficient vigilance estimation method [13].

Researchers have studied vigilance state during simulated vehicle-driving tasks or flight tasks. However, to our knowledge, previous studies have not explored the vigilance state during BCI tasks. Also, the single- or multi-modal vigilance estimating approach for SSVEP-based BCI has not been reported. In this paper, we designed a 4-target brain-controlled cursor platform based on SSVEP and conducted an online cursor control experiment to collect the multimodal signals (EEG signals, EOG signals, ECG signals, and EDA signals) of subjects during a long-term BCI task. Then, the widely used features of the multimodal signals were extracted and analyzed. Next, the useful features were fed into regression models to estimate vigilance levels. We found that the vigilance levels of subjects have changed in the experiment and some features have a close relationship with the vigilance in the BCI task. The experimental results suggest that an acceptable approach to estimate vigilance for SSVEP-based BCI is established in our study.

II. METHOD

A. Subjects

18 healthy subjects (five females, 26±4 years old) with normal or corrected-to-normal vision were recruited in the experiment. The subjects attended the experiment during early afternoons (13:00) or nights (18:30) and were asked to avoid drinking alcohol or coffee before the experiment. All of them signed the informed consent forms before the experiment. The experiment was performed according to the standards of the Declaration of Helsinki, and the study was approved by the Research Ethics Committee of Chinese Academy of Sciences, China.

B. Experiment

Each participant underwent a 90 min sustained cursor-control task based on SSVEP-based BCI, and the procedure is shown in Figure 1. During the task, 4 stimuli flickered at the same time and lasted until the end. The subjects needed to focus on one of the stimulus targets independently, and the BCI system decoded SSVEP to control the cursor move along the path step-by-step in the designed scene. The duration of each command (step) was about 1.2 s, and the position of the cursor was updated once. The circular control path we designed in this study is simple, single, and monotonous, which is more likely to cause vigilance decrement of the subjects.

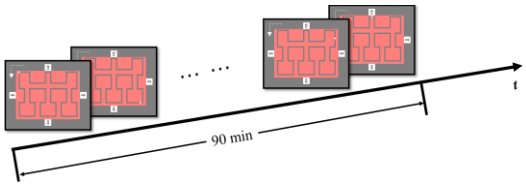

Figure 1. Procedure and stimulus interface of the experiment.

C. SSVEP-based BCI

In this study, we designed a SSVEP-based BCI system. The visual flicks were modulated by the sampled sinusoidal stimulation method [21] in the stimulus interface to evoke SSVEP, as shown in Figure 1. The stimulus interface was presented on a 19-inch LED screen with a resolution of 1280×1024 pixels and a refresh rate of 60 Hz. Each stimulus was presented within a 100×100 pixels square, which corresponded to the basic control intentions (commands) of moving-downward, turning-left, moving-upward, and turning-right. In order to get a stronger SSVEP response and avoid harmonic interference, the stimulus frequencies were set at 8 Hz, 9 Hz, 10 Hz, and 11 Hz, respectively.

To achieve frequency detection, filter bank canonical correlation analysis (FBCCA), a no-training algorithm, was adopted in our study [4]. FBCCA is an extended method to improve the accuracy of canonical correlation analysis (CCA) in the frequency detection of SSVEP. It decomposes the full frequency range of EEG data into sub-bands and calculates the correlation coefficients between each sub-band and the reference signal, respectively [4]. The maximum weighted sum of the correlation coefficients is used to determine the classified results of SSVEP.

The details of frequency detection are as follows. First, the EEG data streams were acquired from the EEG acquisition server in real-time via TCP/IP. The data of eight channels (O1, Oz, O2, PO3, POz, PO4, PO7, and PO8) over the occipital and parietal areas were selected for target identification. Then the selected data were processed by a band-pass filter (1-100 Hz) and downsampled to 250 Hz. Next, the 2-s epoch data were fed into the FBCCA algorithm for classification. In this study, a pseudo classification result was calculated within a 2 s time window with a 0.4 s moving step. In order to ensure the classification accuracy, we conducted an available result judgment in each of three-pseudo classification results (1.2 s). If the three pseudo results are the same, an available classification result was acquired. Otherwise, there is no available result output and wait for the next round of three-pseudo classification results. So, the interval of each available classification result was about 1.2 s. In FBCCA, we set N=7 (the number of sub-bands) and the frequency of the n-th sub-band range from n×8 Hz to 88 Hz. The better empirical values of a, b, and N_h (the number of harmonics) were 1.25, 0.25, and 5, respectively. All the above parameters were set according to [4]. Finally, the available classification results were transformed into commands and sent to the designed interface via TCP/IP.

D. Multimodal data acquisition

During the whole experiment, EEG signals, ECG signals, EOG signals, EDA signals, and eye movement signals of the participant were acquired synchronously. The sampling rate of these signals was 1000 Hz. The EEG signals and EOG

signals were continuously recorded using a 64-channel Synamps2 system (Neuroscan, Inc.) and the 64 EEG electrodes were placed on the scalp according to the international 10-20 system. EOG signals were recorded from the 4 electrodes placed on the left and right outer canthus and above and below the left eye, respectively. The impedance of EEG and EOG electrodes was maintained below 10 kΩ, and the reference and ground electrodes were placed on the M1 (left mastoid) and AFz channels, respectively. The ECG signals were measured using three electrodes (according to the Einthoven lead-II configuration) with an ECG amplifier (BIOPAC System, Inc.). The EDA signals were collected using an EDA amplifier (BIOPAC System, Inc.) and the two electrodes were placed on the middle phalanges of the middle and ring fingers. Besides, the eye movement signals were measured by EyeLink eye tracker (SR Research Ltd., Canada) with a PC for doing the image processing.

E. Data preprocessing and feature extraction

The EEG data were first re-referenced to the average signal of left and right mastoids, and then processed with a band-pass filter between 1 and 50 Hz, down-sampled to 200 Hz. To remove the artifacts, an Automatic Artifact Removal (AAR) toolbox [22] was applied. And then, the EEG data were divided into non-overlapping 4-s segments and then the baseline of each segment was extracted and removed.

The widely used features, PSD and DE, have shown valid and good performance in the areas of emotion recognition and fatigue driving detection. Thus, we used PSD and DE in this study. PSD feature and DE feature of each EEG segment were calculated using Short-term Fourier Transforms (STFT) from delta band (1-4 Hz), theta band (4-8 Hz), alpha band (8-13 Hz), beta band (13-30 Hz), and gamma band (30-50 Hz). Following the previous study, the DE can simply be calculated by the following formulation [13]:

$$h(\boldsymbol{X}) = \frac{1}{2}\log(2\pi e\sigma^2), \quad (1)$$

where the variable $\boldsymbol{X}$ obeys the Gaussian distribution $N(\mu, \sigma^2)$.

The EOG data were first re-referenced to the average signal of the left and right mastoids, and then filtered with a 30 Hz low-pass filter to eliminate the noise. Finally, the filtered data were down-sampled to 200 Hz and split into 4-s, non-overlapping segments. After the above preprocessing, the pure horizontal EOG (HEOG) and vertical EOG (VEOG) data were acquired. To extracted the features, we first applied the wavelet transform method [23] to achieve reliable edge detection of VEOG and HEOG and then used the peak detection algorithm to detect the blink events and saccade events. Finally, we extracted a total of 36 EOG features from the detected blink events and saccade events, refer to the literature [13] for more details.

For ECG data, a band-pass filter between 2 and 40 Hz was used to remove noise. Next, the data were down-sampled to 200 Hz and split into 4-s, non-overlapping, detrended segments. To extract the heart rate (HR) and HRV features for every segment, the RR intervals were calculated first with a sliding window (56-s overlapping). Then a total of 25 features about HR, time-domain HRV, frequency-domain HRV were obtained, refer to [24] for details.

To reduce the noise and remove the artifacts, a low-pass filter with a cutoff frequency at 10 Hz was applied for the EDA data. Next, the filtered EDA data were down-sampled to 200 Hz and segmented to non-overlapping epochs of 4 s. In the end, 8 EDA features were obtained. They are the maximum, minimum, mean, standard deviation, variance, and quadratic sum of EDA amplitude, the average and root mean square of the first derivative of EDA amplitude. On the whole, there were 2×(5×62) (bands×channels) EEG features, 36 EOG features, 25 ECG features, and 8 EDA features for each segment. And all the features were normalized to the range from zero to one before being input into the regression models.

F. Vigilance labeling

PERCLOS, the PERcentage of eye CLOSure over time, has been used as the label (ground truth) to estimate vigilance in this study. According to the eye movements information (such as blink duration, saccade duration, fixation duration) acquired from the eye tracker, the PERCLOS index values were calculated from the percentage of the durations of blinks and 'CLOS' over a specified time interval as follows [13]:

$$PERCLOS = \frac{blink + CLOS}{blink + fixation + saccade + CLOS}, \quad (2)$$

where 'CLOS' denotes the duration of the eye closures. And the PERCLOS values (ranges from 0 to 1) were further smoothed (sliding average).

G. Regression

As a human intrinsic mental state that involves temporal evolutions, the vigilance of users is a dynamic changing process [13]. The outcome of the regression algorithm is to predict continuous values, thus, we used the regression algorithm for vigilance estimation. In this study, three regression models, support vector regression (SVR), Bayesian ridge regression (BRR), and gradient boosting regressor (GBR), were performed independently for each participant. SVR is a supervised learning method and constructs a hyperplane that optimally predicts the distribution of samples. GBR is an integrated model with high performance and good stability. In this paper, we used a SVR model with a linear kernel. In addition, the minimum-Redundancy-Maximum-Relevance (mRMR) algorithm [25], a powerful feature selection method based on minimum redundancy and maximum relevance conditions, was adopted to realize the multimodal features selection. The three algorithms were implemented based on the Python package scikit-learn. In the aspect of performance evaluation, the root mean square error (RMSE) and correlation coefficient (COR) are conventional metrics of regression models and were applied in our study. The results were verified by a 5-fold cross-validation method and we concatenated the predicted results and labels of the five sessions and calculated the evaluation metrics.

III. Results and Discussions

A. Relationship between representative features and vigilance

Figure 2 shows the PERCLOS, the classification accuracy of SSVEP-based BCI tasks, the maximum blink rate of EOG, the HR of ECG, and the mean value of EDA of one subject

throughout the experiment. It can be seen from the PERCLOS subplot that the vigilance level changed significantly during the whole experiment. For example, in the intervals 601-680, the value of PERCLOS is 1 when the subject was asleep. Following the traditional definition of vigilance level based on PERCLOS value [13, 14], the PERCLOS values in the range of 0-0.35 and 0.36-1 represent the high-vigilance level and the low-vigilance level, respectively. As shown in Figure 2, the classification accuracy has a close relationship with PERCLOS. In the low-vigilance level, the classification accuracy is relatively low. This indicates that the performance of subjects is related to vigilance level. Compared with the high-vigilance level, the subject has a lower value of HR and a higher mean amplitude of EDA in the low-vigilance level. As for the maximum blink rate, its value is zero during the intervals 601-680, since the subject was asleep. Therefore, these features are related to vigilance and have the potential to be used to estimate vigilance.

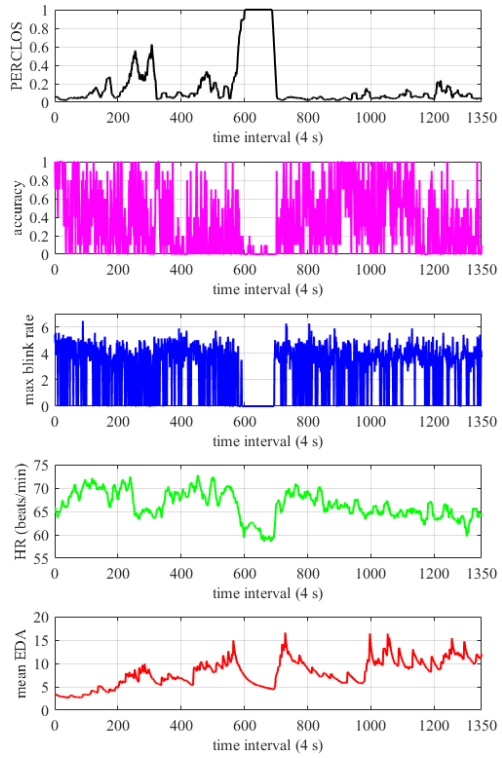

Figure 2. The PERCLOS and some features of one subject in the experiment.

For the same subject, the averaged PSD of five bands in the high-vigilance level (intervals 1-80 on Figure 2) and low-vigilance level (intervals 601-680 on Figure 2) were calculated respectively, and the topographical maps based on the averaged PSD are depicted in Figure 3. The distribution of the topographical map based on average PSD is obviously different in the different frequency bands and different vigilance levels. In the high-vigilance level, the delta-band spectrum power is mainly distributed in the prefrontal region, while the alpha-band power is mainly distributed in the occipital and parietal regions. It is consistent with the brain activity of the subjects when they performed the SSVEP task in the high-vigilance level. In the low-vigilance level, the subject was asleep and there are the increased PSD of delta and theta bands in the frontal region. These results support that the increasing trend for the ratio of slow and fast waves of EEG activities reflects decreasing attentional demands [26].

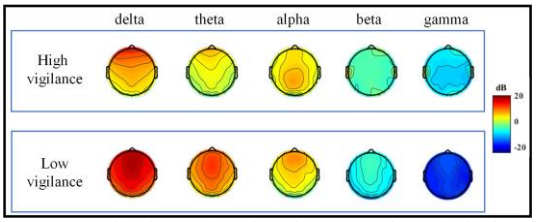

Figure 3. Topographical maps based on the averaged PSD in high-vigilance level and low-vigilance level.

B. *Single-modal based vigilance estimation*

In order to keep consistent with the processing of labels, we smoothed the predicted results in the same way and then calculated the evaluation metrics between the results and labels. To compare the performance of different modalities under different regression models, the mean RMSE and mean COR of all subjects are shown in Table I. Among the single modality, the EEG modality achieves the best performance for vigilance estimation in terms of both RMSE and COR. Within the EEG features, DE outperforms the PSD with higher COR and lower RMSE in all models. The reason why the EEG modality performs better than other modalities may be that it has more dimensional features. And the better performance of DE features in EEG modality may be attributed to its stability [27]. In the aspect of models, the linear SVR model performs better than the other models in the EEG modality (DE and PSD) while the BRR model achieves the better performance in EOG, ECG, and EDA modality. Overall, the combination of DE features and linear SVR model provides the best performance with RMSE of 0.081 and COR of 0.487.

TABLE I. RESULTS OF THE SINGLE-MODAL FEATURES.

Single-modal		Linear SVR		BRR		GBR	
		RMSE	COR	RMSE	COR	RMSE	COR
EEG	DE	**0.081**	**0.487**	0.082	0.470	0.083	0.448
	PSD	**0.085**	**0.427**	0.086	0.412	0.085	0.372
EOG	/	0.093	0.131	**0.091**	**0.167**	0.092	0.116
ECG	/	0.099	0.048	**0.099**	**0.064**	0.099	-0.033
EDA	/	**0.097**	-0.007	0.101	**0.011**	0.099	-0.119

C. *Multimodal based vigilance estimation*

Since the EEG features achieve the best performance, we further combined EEG features (DE or PSD) with the features of other modalities respectively to investigate the effect of multimodal signals on vigilance estimation. In order to compare in the same dimension, the mRMR algorithm was applied to selected the features from multimodal signals with the same dimension (310) as DE features of the single modality. The performance of the multimodal features is shown in Table II.

Compared to the PSD-based multimodal methods, the DE-based multimodal methods have better performance. Moreover, EOG-based multimodal methods outperform the ECG-based and EDA-based with higher COR and lower RMSE. Also, the best multimodal method is the fusion of DE and EOG. In the DE-based multimodal methods, the linear SVR model achieves better performance than BRR and BGR models. In the PSD-based multimodal methods, the BRR model achieves better performance when compared with other models. The best result is obtained by combining the

fusion of DE and EOG features with the linear SVR model, whose COR and RMSE are 0.511 and 0.080 respectively. Combined with Table I and Table II, the performance of the multimodal methods is better than that of the single-modal methods with higher COR and lower RMSE. The results demonstrate that modality fusion can effectively enhance the performance of vigilance estimation in comparison with a single modality in the BCI task.

TABLE II. RESULTS OF THE MULTIMODAL FEATURES.

Multimodal	Linear SVR		BRR		GBR	
	RMSE	COR	RMSE	COR	RMSE	COR
DE+EOG	**0.080**	**0.511**	0.080	0.505	0.082	0.481
DE+ECG	**0.081**	**0.489**	0.082	0.484	0.084	0.441
DE+EDA	**0.082**	**0.474**	0.082	0.468	0.084	0.427
PSD+EOG	0.083	0.461	**0.083**	**0.462**	0.083	0.434
PSD+ECG	0.085	0.443	**0.085**	**0.445**	0.086	0.374
PSD+EDA	0.085	0.416	**0.085**	**0.424**	0.084	0.399

IV. CONCLUSIONS

The present study tried to estimate the vigilance levels during a BCI task. We designed the cursor-control BCI based on SSVEP and acquired the multimodal signals, which included EEG, EOG, ECG, and EDA, from subjects during a long-term cursor-control task. The features of these signals were extracted and the regression models were trained using the single modality features and multimodal features to realize the estimation of vigilance. The experimental results show that the performance during the BCI task was related to the vigilance of the BCI user. Among multimodal signals, EEG signals can efficiently estimate vigilance in the BCI task. Moreover, the multimodal methods can significantly enhance the performance of vigilance estimation compared with the single modality methods. Combining the fusion of DE and EOG features with the SVR model has achieved a satisfactory performance of vigilance estimation in BCI tasks. These experimental results demonstrate the feasibility and efficiency of our proposed approach based on multimodal signals and regression models.

REFERENCES

[1] J. Wolpaw and E. W. Wolpaw, "Brain-computer interfaces, principles and practise," *Oxford University Press*, 2012.

[2] J. N. Mak and J. R. Wolpaw, "Clinical applications of brain-computer interfaces: current state and future prospects," *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 187-199, Dec. 2009.

[3] M. Xu, X. Xiao, Y. Wang, H. Qi, T. -P. Jung, and D. Ming, "A brain-computer interface based on miniature-event-related potentials induced by very small lateral visual stimuli," *IEEE Trans. Biomed. Eng.*, vol. 65, pp. 1166-1175, May. 2018.

[4] X. Chen, Y. Wang, S. Gao, T. -P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface," *J. Neural Eng.*, vol. 12, p. 046008, Jun. 2015.

[5] F. Sauvet et al., "In-flight automatic detection of vigilance states using a single EEG channel," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 12, pp. 2840-2847, Dec. 2014.

[6] W. Zheng et al., "Vigilance estimation using a wearable EOG device in real driving environment," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 1, pp. 170-184, Jan. 2020.

[7] B. S. Oken, M. C. Salinsky, and S. Elsas, "Vigilance, alertness, or sustained attention: physiological basis and measurement," *Clin. Neurophysiol.*, vol. 117, no. 9, pp. 1885-1901, Sept. 2006.

[8] Z. Guo, Y. Pan, G. Zhao, S. Cao and J. Zhang, "Detection of driver vigilance level using EEG signals and driving contexts," *IEEE Trans. Reliab.*, vol. 67, no. 1, pp. 370-380, Mar. 2018.

[9] C. -T. Lin, R. -C. Wu, S. -F. Liang, W. -H. Chao, Y. -J. Chen and T. -P. Jung, "EEG-based drowsiness estimation for safety driving using independent component analysis," *IEEE Trans. Circuits Syst. I-Regul. Pap.*, vol. 52, no. 12, pp. 2726-2738, Dec. 2005.

[10] L. Shi and B. Lu, "EEG-based vigilance estimation using extreme learning machines," *Neurocomputing*, vol. 102, pp. 135-143, Feb. 2013.

[11] S. Khessiba, A. G. Blaiech, K. B. Khalifa, A. B. Abdallah, and M. H. Bedoui, "Innovative deep learning models for EEG-based vigilance detection," *Neural Comput. Appl.*, vol. 33, pp. 6921-6937, Jun. 2021.

[12] S. Makeig, T. -P. Jung, and T. Sejnowski, "Awareness during drowsiness: dynamics and electrophysiological correlates," *Canadian Journal of Experimental Psychology*, vol. 54, pp. 266-273, Dec. 2000.

[13] W. Zheng and B. Lu, "A multimodal approach to estimating vigilance using EEG and forehead EOG," *J. Neural Eng.*, vol. 14, p. 026017, Feb. 2017.

[14] W. Wu, W. Sun, Q. M. J. Wu, Y. Yang, H. Zhang, W. Zheng, et al., "Multimodal vigilance estimation using deep learning," *IEEE T. Cybern*., Early Access, 2020.

[15] J. Ma, L. Shi, and B. Lu, "Vigilance estimation by using electrooculographic features," in *Proc. 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Buenos Aires, 2010, pp. 6591-6594.

[16] Y. Zhang, X. Gao, J. Zhu, W. Zheng, and B. Lu, "A novel approach to driving fatigue detection using forehead EOG," in *Proc. 2015 7th International IEEE/EMBS Conference on Neural Engineering* (*NER*), Montpellier, 2015, pp. 707-710.

[17] S. -J. Jung, H. -S. Shin, and W. -Y. Chung, "Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel," *IET Intell. Transp. Syst.*, vol. 8, no. 1, pp. 43-50, Feb. 2014.

[18] F. Zhou, A. Alsaid, M. Blommer, R. Curry, R. Swaminathan, D. Kochhar, et al., "Driver fatigue transition prediction in highly automated driving using physiological features," *Expert Syst. Appl.*, vol. 147, p. 113204, Jun. 2020.

[19] M. I. Posner, C. R. Snyder, and B. J. Davidson, "Attention and the detection of signals," *J. Exp. Psychol.-Gen.*, vol. 109, pp. 160-174, Jun. 1980.

[20] M. Awais, N. Badruddin, and M. Drieberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability," *Sensors*, vol. 17, no. 9, p. 1991, Aug. 2017.

[21] N. V. Manyakov, N. Chumerin, A. Robben, A. Combaz, M. van Vliet, and M. M. Van Hulle, "Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain-computer interfacing," *J. Neural Eng.*, vol. 10, p. 036011, Apr. 2013.

[22] G. Gomez-Herrero et al., "Automatic removal of ocular artifacts in the EEG without an EOG reference channel," in *Proc. Proceedings of the 7th Nordic Signal Processing Symposium*, Reykjavik, 2006, pp. 130-133.

[23] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 741-753, Apr. 2011.

[24] Y. Hsu, J. Wang, W. Chiang and C. Hung, "Automatic ECG-based emotion recognition in music listening," *IEEE Trans. Affect. Comput.*, vol. 11, no. 1, pp. 85-99, Jan. 2020.

[25] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *J. Bioinform. Comput. Biol.*, vol. 03, no. 2, pp. 185-205, Apr. 2005.

[26] B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, "Using EEG spectral components to assess algorithms for detecting fatigue," *Expert Syst. Appl.*, vol. 36, no. 22, pp. 2352-2359, Mar. 2009.

[27] R. Duan, J. Zhu, and B. Lu, "Differential entropy feature for EEG-based emotion classification," in *Proc. 6th International IEEE/EMBS Conference on Neural Engineering* (*NER*), San Diego, 2013, pp. 81-84.