Estimating prediction horizon of driver fatigue using Euclidean distance-based similarity score between electroencephalograms

Rodney Petrus Balandong[1#], Syaimaa Solehah Mohd Radzi[2], Zulkifli Yunus[3], Mohamad Zul Hilmey Makmud[1], Tang Tong Boon[2]

1 Green Technologies and Advanced Matter (GreAt) Research Group, Faculty of Science and Natural Resources, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, MALAYSIA
2 Centre for Intelligent Signal and Imaging Research, Universiti Teknologi PETRONAS, Persiaran UTP, 32610 Seri Iskandar, Perak, MALAYSIA.
3 Petronas Twin Towers, Tower 1, Kuala Lumpur City Centre, Kuala Lumpur, Kuala Lumpur, 50088, MALAYSIA.
Corresponding author. E-Mail: balandong@ums.edu.my; Tel: +6088-320000; Fax: +6088-435324.

ABSTRACT Driver fatigue is one of the major causes of road accidents. While numerous electroencephalography (EEG) related methodologies have been proposed for automatic fatigue detection, very little attention has been given to explore the use of EEG in the estimation of the prediction horizon of driver fatigue. This paper proposed a novel framework based on the similarity score measured by the Euclidean distance in the brain oscillatory rhythmic patterns to determine how far ahead the decrement in driver's vigilance could be detected. A new metric for the confidence level of the estimation was also suggested to quantify prediction reliability. The proposed framework was assessed using the data from a driving simulation experiment involving 20 healthy female subjects with mean age of 22 and found that the prediction horizon can be extended up to 56s solely based on EEG features. In conclusion, this study demonstrated how the EEG features can be used for the estimation of prediction horizon in driver fatigue management.

KEYWORDS: Electroencephalography, Prediction Horizon, Drowsiness Detection, Driver Fatigue, Kernel Density Estimation.
Received 30 October 2024 **Revised** 16 December 2024 **Accepted** 31 December 2024 **Online** 4 January 2025

Original Article

INTRODUCTION

Continuous intensive driving without a break demands a significant amount of physical and mental resources that can eventually lead to driver fatigue (Getzmann *et al.*, 2018). Fatigue reduces the driver's vigilance and subsequently impairs their abilities in controlling the vehicle (Kong *et al.*, 2017). According to the National Highway Traffic Safety Administration of the United States (US), fatigued driving is associated with 100,000 crashes in the US, leading to over 1,500 deaths and 71,000 injuries each year (NSF, 2024).

Prediction horizon is defined as the time window starting from the detection of a fatigued state in the driver to the onset of the road accident. Within the context of accident prevention, a longer prediction horizon is desirable so that drivers have adequate lead time in taking the necessary intervention measures. In two published studies, several vehicle variables were utilized as a feature set for the driver fatigue detection system to predict unintentional lane departure by 0.2-0.6s (Ambarak *et al.*, 2017), and 60s (McDonald *et al.*, 2013) in advance, albeit at an area under the receiver operating characteristic curve of 0.7. In a study by Murata, correlation and regression analysis linked the changes of physiological factors including electroencephalography, heart rate, and eye movement with behavioural (neck bending angles, foot pressure, and back pressure) to achieve 20s of prediction horizon (Murata, 2016).

Electroencephalograph (EEG) oscillatory activity reflects thalamocortical activity in the brain (Olbrich *et al.*, 2014). Our previous study (Radzi *et al.*, 2019) confirmed that the EEG power spectrum is a reliable marker for distinguishing between alert and fatigued driving states. Building on this, the

present study revisited the same dataset and represents the first attempt to use EEG signals for estimating prediction horizons. We proposed a novel framework employing a Euclidean distance-based similarity score, calculated between non-overlapping EEG features at various time windows and the time window of road excursion onset. Kernel density estimation was used to model the similarity score, yielding a "confidence level" metric. This framework was tested using data from a driving simulation experiment to identify EEG features across brain regions and frequency band ratios that provide the longest prediction horizon.

BACKGROUND THEORY

Similarity Between Two Time Windows

This study aimed to examine the similarity in the neural oscillatory patterns between the preceding event and the road excursion onset windows as a potential means to estimate the prediction horizon. By comparing with the road excursion onset, the similarity score could be a useful indicator of the starting point of the vigilance decrement. Based on this concept, the Euclidean distance method was used to measure similarity between two-time windows. This method measured the geometric distance between two points (neural oscillatory parameters). In a Euclidean plane, if $m = (x_p, y_p)$ and $n = (x_q, y_q)$, the distance could be computed using Equation (1) (Liberti *et al.*, 2014).

$$d(m,n) = \sqrt{(x_p - x_q)^2 + (x_p + x_q)^2} \quad (1)$$

Then, the distance values can be normalized using unity-based normalization to restrict the range of values between 0 and 1 which is written as

$$\hat{d} = \frac{d - d_{min}}{d_{max} - d_{min}} \quad (2)$$

Using Equation (3), the normalized distance values were transformed into similarity values, denoted as $\hat{d}$. Generally, the higher the $\hat{d}$ value is, the more likely the mental state at that particular time window is showing a higher similarity to that of the road excursion onset window.

$$\hat{d} = 1 - \hat{d} \quad (3)$$

As this study only aimed to identify whether there was a high or low similarity, the value was then binarized. This was calculated by applying a threshold (T_b) to all the similarity values. In other words, only similarity values larger than T_b (i.e., high similarity value) would be categorized as '1' (thereafter, they were labeled as 'case'). Otherwise, they were removed from subsequent processes. Binarization not only reduces individual differences and improves the generalization, but it also allows easier interpretation of similarities between windows.

By using the window with label W_0 as the referencing window, the aforementioned process was repeated between the referencing window W_0 against each of the preceding window W_{-264}, W_{-260}, W_{-256}, ..., W_{-4} for each epoch and each participant. For simplicity, throughout this paper, the comparison between W_0 against each of the preceding window W_{-264}, W_{-260}, W_{-256}, ..., W_{-4} would be abbreviated as time 264s, 260s ,256s, ... ,4s, respectively.

Preliminary analysis showed that, as expected, the cases were unevenly distributed due to variability among subjects. In order to better identify the most common time instance of detecting

EEG patterns similar to those at events of interest (road edge excursion) amidst the variability among subjects, we need to model the histograms of the cases. Here, we proposed to use Kernel Density Estimate (KDE), which is a non-parametric method, to smoothen and model the probability density of the binary data. The modelling can be achieved by overlaying an identified density function (the kernel) across the binary data to generate smooth histograms that preserve important density features at multiple scales. Let X be a random variable and the random sample $\{X_1,\ldots,X_n\}$, then the density $f(x)$ can be calculated as

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{X_i - x}{h}\right) \quad (4)$$

where h is the bandwidth that controls the degree of smoothing, $k(u)$ is a kernel function, and this study employs the gaussian kernel. The estimator $\hat{f}(x)$ represents the total observations around the point x. Such that, $\hat{f}(x)$ is having a higher value if there is a higher number of observations around x, and a lower value of $\hat{f}(x)$ when there is only a few X_i near the x.

Evaluation Of Prediction Horizon

In addition to modeling prediction horizons, this study also aims to investigate which pair of electrode and frequency band ratio that could provide the furthest prediction horizon. In particular, we are more interested in obtaining a range of potential prediction horizon for each of the electrodes and frequency bands ratio, instead of extracting singular value. To be able to make comparisons, a metric must be developed to quantify the quality of the prediction horizon. Thus, we proposed a new metric R_{data} that denoted as the ratio of the sample count (m) within a specific sigma level against the total number of samples (n). The R_{data} helped determine the best channel-feature pair that provided the best prediction horizon range. Sigma (σ) level was akin to the standard deviation in between the peak of the density distribution, or in other words, the prediction horizon dispersion.

For simplicity, we set each σ to be a one time-window, i.e. 4s. As shown in Figure 1, the sigma level of two (2σ) indicated the prediction horizon range to be two time-windows from the highest peak to the lower (152s) and upper (136s) time series respectively. Let $\{X_1, \ldots, X_n\}$ be a set of finite data samples, and the samples within an enclosed sigma level are $\{X_1, \ldots, X_m\}$, then R_{data} can be obtained as using Equation (5).

$$\mathrm{R_{data}} = \frac{\mathrm{m}}{\mathrm{n}} \quad (5)$$

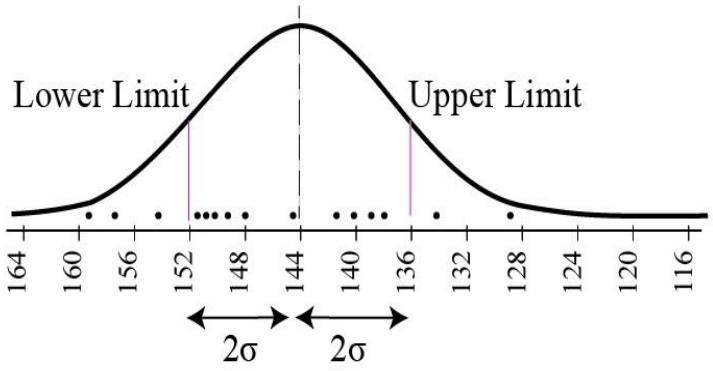

Figure 1. Illustration of different sigma levels from the peak 144s. Sigma level of two indicates that the prediction horizon range is two levels from the peak to the lower (144s to 152s) and upper (144s to 136s) limits respectively.

METHODOLOGY

EEG Dataset

The dataset contained electrophysiological data from 32 healthy subjects (28 females and four males) with a mean age of 22 years old. The study was approved by the Universiti Teknologi MARA Research Ethics Committee and all the subjects provided informed consent before participating in the study. The study involved the collection of continuous measurements during a simulated session in which the participants were required to keep the vehicles at the center of the road as accurately as possible at a constant driving speed. At the instance the vehicle drifted beyond the edge of the road, the participant was instructed to make corrective steering adjustments to return to the centre of the road. The road edge excursion event during the sustained-attention driving task was logged into the STISIM software. The EEG signals were recorded using a 64-channel EEGOSPORTS (ANT Neuro, Hengelo, Netherlands) with 17 active electrode positions at F_{p1}, F_{p2}, F_{pz}, F_3, F_z, F_4,C_3, C_Z, C_4, P_3, P_Z, P_4, T_7, T_8, O_1, and O_2. All the EEG electrodes were referred to the C_{pz} and placed in accordance with a modified International 10–20 system of electrode. For the purposes of this manuscript, only the dataset from 20 females was utilized in the analysis, as provided and permitted by the data owner.

Segmenting Epochs of Interest

The onset of crossing a road edge was used as the reference point to identify the instance at which the event of interest occurs. These time instances were used as references to segment the continuous time-series EEG signals into epochs of 270s. Specifically, each epoch referred to the period of 270s preceding a road edge excursion event. The choice of using 270s as the epoch size was based on the finding that a driver's vigilance fluctuated within a cycle length of greater than 240s (Makeig *et al*., 1993). Also, this value was selected to maximize the number of the epochs for analysis while ensuring no overlapping between two road edge excursion events within the same time-series data.

Prior to segmenting the time series into epochs, the EEG data were band-pass filtered (1-50 Hz) and down-sampled to 200Hz to improve the signal-to-noise ratio and to reduce the dimensionality of the data. Bad channels in the recordings were removed using the Autoreject module. This was followed by an independent component analysis (ICA) to remove any components that correlated strongly with eye movements and muscle artifacts. In addition, noisy channels were also flagged and ameliorated using the Autoreject module.

The continuous time series was segmented into a series of 67 temporal windows of 4s length each without overlapping. The choice of window length was based on previous studies in which the duration of 10-16 cycles for the lowest EEG band (for θ band, about 4s) was accepted as an adequate choice to capture the dynamic of fast and slow brain rhythms (Dimitriadis *et al*., 2017). Each of the windows was assigned with a label, with the window closest and furthest to the event onset labeled as W_0 and W_{-256} respectively. The relative power for three standard frequency bands, namely θ (3–7 Hz), α (8 – 13Hz), and β (14–30Hz), for each of the temporal windows were calculated via fast Fourier transform. Then, the ratio indices were calculated from the three standard frequency bands for each of the temporal windows. The α/θ and $(\theta + \alpha)/\beta$ were selected for their known sensitivity to the change of driving fatigue. It should be note that, the α/θ and $(\theta + \alpha)/\beta$ were sometime referred as TA and TAB, respectively in the following text.

RESULT

In this study, the parameter T_b and sigma level played an important role in determining the outcome of the prediction horizon analysis. The selection of T_b was directly related to the sensitivity

and specificity of the detection of vigilance decrement. Clearly, if T_b was higher, then the evaluation would be more specific to the similarity between the corresponding windows. The sigma level selection might affect how wide the prediction horizon range around the average prediction horizon would be. Narrow sigma level means the proposed prediction horizon range is based on data points that are clustered closely around the density distribution peak. In contrast, a wide sigma level refers to wider prediction horizon range which might include possible outliers. We first investigated the performance and pattern of R_{data} with varying T_b (0.90-0.99 with step size of 0.01) and sigma levels (1-30 with step size of 1). In most of the channel-feature pairs, no substantial changes in R_{data} was found when increasing the T_b by 0.01 unit while keeping the sigma level constant. This suggests that most of the channel and feature pairs have a similarity value higher than 0.99. On the other hand, varying the sigma level while maintaining the T_b resulted in a proportional increase of R_{data} value. This is because widening the sigma level means a larger sample number (data points) within the area under the sigma level while the total sample is maintained. In view of this observation, the values for T_b and sigma level were selected to be 0.99 and 15 (half of the sigma evaluation range), respectively, for the subsequent analysis.

Considering a total of 16 channels and two features, there was a total number of 32 channel-feature pairs. It might be useful to examine the top five channel feature pairs and to sort the performance results in descending order with the highest R_{data} value on the top of the list which is shown in Table 1. The top five electrodes (channels) were O_2, F_4, T_8, P_4, and P_3. The experiment results suggested that the O_2-TA (R_{data} = 0.58228) was the best channel-pair with the highest R_{data} value, followed by F_4-TA (R_{data} = 0.51685) and T_8-TAB (R_{data} = 0.49180) channel-pair. The result also indicated that the variation of R_{data} magnitude across the five top-ranked channel-feature pairs was little, except for the top two channel-pair. On the contrary, the peak of the prediction horizon showed a substantial variation in performance across the five top-ranked channel-feature pairs with the lowest and highest prediction horizon peaks of 52s and 100s respectively. In terms of prediction horizon range, P_3-TA (R_{data} = 0.48438) gave the furthest range of 100s, followed by P_4-TA (R_{data}= 0.48649), and F_4-TA (R_{data} = 0.51685), with the peak prediction at 84s and 80s, respectively. Despite being the two top channel-feature pairs, both O_2-TA and F_4-TA, each provided a prediction horizon of 56s (R_{data} = 0.58228), and 80s (R_{data} = 0.51685), respectively.

Table 1. Five top-ranked channel -feature pairs sorted according to the R_{data} in descending order.

No	Ch	Feature	R_{data}	$Lim_{Low}(s)$	$Lim_{Upper}(s)$	Peak (s)
1	O_2	TA	0.58228	100	12	56
2	F_4	TA	0.51685	124	36	80
3	T_8	TAB	0.49180	96	8	52
4	P_4	TA	0.48649	128	40	84
5	P_3	TA	0.48438	144	56	100

In this study, the initial assumption was that the similarity ($\hat{S}$) magnitude would be increased toward the accident onset. This assumption can be visualized clearly via the overlay of the histogram and KDE plot. As shown in Figure 2a for the O_2-TA pair, the frequency and density of event between the 264s to 104s prediction horizon were lower, as expected. Subsequently, there was a sharp increase in the frequency and density of events from 104s to 64s. Finally, the number of frequencies of event remained stable towards the accident onset. In contrary, the density distribution displayed a decrease from 64s towards accident onset. Such decrease behaviour at the end of the continuum is expected since the weight of the Gaussian kernel values decreases with distance from the central. A similar

pattern was also displayed by the T_8-TAB pair (Figure 2b) whereby there was an increase in the number of frequencies from 264s to 104s, followed by relatively stable number of the events between 104s to 24s, and finally accompanied by a steep increased number of events from 24s to 4s. The density distribution showed a bi-modal behaviour but with a progressive increased density as the prediction horizon progress closer to the accident onset and reach the highest peak density at 52s. The stark difference between the plots in Figure 2a and Figure 2b is the number of events covered under the highest peak of each plot, 56s, and 52s, respectively. The number of events under the highest peak reflects the confidence level (hence, the R_{data} magnitude) in such prediction horizon.

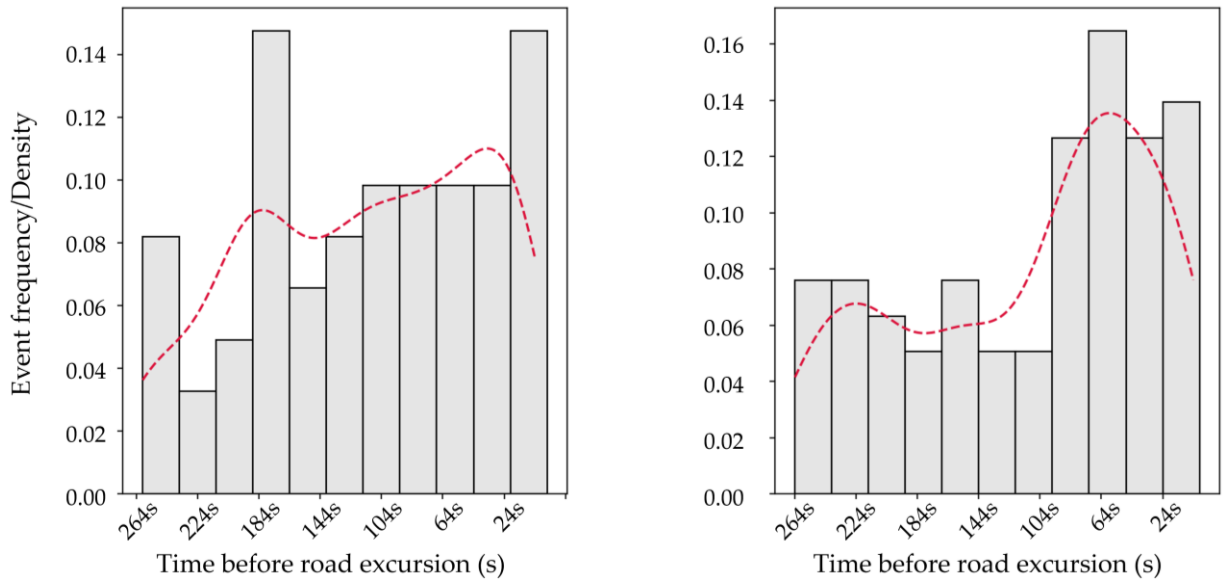

Figure 2. Overlay of histogram and KDE plot for the (a) O_2-TA and (b) T_8-TAB pair. The histogram is represented by rectangles, while the red dashed line denotes the corresponding KDE plot.

DISCUSSION

Sustained cognitive attention involves multiple cortical and subcortical brain regions. The frontal region plays a major role in coordinating executive functions; the central and parietal regions are involved in sensory activities; the temporal region coordinates specific functions including visual memory, verbal memory, and auditory tasks, whereas the occipital region is involved in visual information reception. Table 1 lists the top five channel-feature pairs. Interestingly, multiple brain regions including dorsolateral prefrontal cortex (F_4), temporal (T_8), parietal (P_3 & P_4), occipital (O_2) were all listed in the top five list. These regions are consistent with the strategy employed by multiple studies that relied upon electrode placement at the regions such as frontal (Kartsch *et al.*, 2018), occipital (Kartsch *et al.*, 2018), and central (Qian *et al.*, 2017) as the main brain regions for the development of automatic mental driving fatigue detection.

In terms of the R_{data}, O_2-TA scored the highest value. In (Chuang *et al.*, 2015), an increased alpha and theta power was detected in the occipital region with the increasing fatigue level. It is known that alpha oscillation in the occipital cortex amplifies significantly during eye closure and decreases during eye-opening. Hence, there is probably a similar amplitude of eye closure and fatigue level before the onset of an accident. The result showed that the pair F_4-TA with the peak prediction range of 196s displayed the second highest value of R_{data}. The importance of the F_4-TA pair coincided with another study (Zhao *et al.*, 2011) that reported a significant TA increase on the F_4 with an increasing fatigue level.

As discussed in the introduction, several studies addressed the prediction horizon. Among the studies, only one of the studies (Murata, 2016) employed the EEG feature in their methodology. Despite utilizing a multimodal input, the system only achieved 20s of the prediction horizons. Among the remaining studies, the framework proposed by Larue *et al*. (2015) achieved the longest prediction horizon of up to ten minutes with a one-minute increment. Our study utilized only EEG features and reported the peak of the prediction horizon varied with the highest and lowest prediction horizon peak of 100s and 52s respectively.

This study has several limitations that warrant consideration. First, the use of a driving simulator, while ensure a controlled environment, but does not fully represent a real-world driving condition, such as varied road environments and external stressors that may influence the generalizability of the findings. Additionally, the participant sample may lack diversity in terms of age and gender which subsequently limiting the applicability of the results across different populations. Moreover, the computational complexity of the proposed framework, involving extensive preprocessing and analysis steps, presents an obstacle for a real-time implementation in onboard systems with limited processing power. Addressing these limitations in future research will be essential for broader application and validation of the findings.

Future research should prioritize real-world validation studies to test the proposed EEG-based fatigue prediction framework under natural driving conditions. Apart from that, future study may incorporate diverse driving environments inclusive varying road types, weather conditions, and traffic scenarios, as well as involving participants from broader demographic groups to ensure generalizability. Additionally, cross modality fusion between the EEG with other fatigue detection methods, such as heart rate variability, eye tracking, or behavioural monitoring may create a robust multimodal system. This approach would leverage the complementary strengths of each modality and therefore increase the system's reliability.

CONCLUSION

This paper explored the ability of EEG features to predict driver fatigue ahead of road accidents. The study aim was achieved via the newly proposed novel framework based on the Euclidean distance method and kernel density estimation integrated with the sigma level. The analysis showed that EEG power spectral ratio coupled with appropriate electrode placement could inform the driver's performance decrement as early as in between 56s before the onset of an accident. Furthermore, this paper offered a new research perspective about the EEG prediction horizon within the transportation domain and extended the requirement for real-time automated EEG-based fatigue detection.

ACKNOWLEDGEMENTS

This research was supported by the Ministry of Higher Education (MOHE) under the Fundamental Research Grant Scheme - Early Career Researcher (FRGS-EC) with the project code (FRGC026-2024). The authors are also grateful to the Centre for Intelligent Signal and Imaging Research (CISIR), Universiti Teknologi PETRONAS for the facilities provided.

REFERENCES

[1] Ambarak, J. M., Ying, H., Syed, F.& Filev, D. 2017. A neural network for predicting unintentional lane departures. *2017 IEEE International Conference on Industrial Technology (ICIT)*. 22-25 March, 2017, Toronto, Ontario, Canada. pp 492-497.

[2] Chuang, C. H., Huang, C. S., Ko, L. W.& Lin, C. T. 2015. An EEG-based perceptual function integration network for application to drowsy driving. *Knowledge-Based Systems*, *80*, 143-152.

[3] Dimitriadis, S. I., Salis, C., Tarnanas, I. & Linden, D. E. 2017. Topological filtering of dynamic functional brain networks unfolds informative chronnectomics: A novel data-driven thresholding scheme based on orthogonal minimal spanning trees (OMSTs). *Frontiers in Neuroinformatics,* 11, Article 28.

[4] Getzmann, S., Arnau, S., Karthaus, M., Reiser, J. E. & Wascher, E. 2018. Age-Related Differences in Pro-active Driving Behavior Revealed by EEG Measures. *Frontiers in Human Neuroscience,* 12, Article 321.

[5] Kartsch, V. J., Benatti, S., Schiavone, P. D., Rossi, D. & Benini, L. 2018. A sensor fusion approach for drowsiness detection in wearable ultra-low-power systems. *Information Fusion,* 43, 66-76.

[6] Kong, W., Zhou, Z., Jiang, B., Babiloni, F. & Borghini, G. 2017. Assessment of driving fatigue based on intra/inter-region phase synchronization. *Neurocomputing,* 219, 474-482.

[7] Larue, G. S., Rakotonirainy, A. & Pettitt, A. N. 2015. Predicting Reduced Driver Alertness on Monotonous Highways. *IEEE Pervasive Computing,* 14(2), 78-85.

[8] Liberti, L., Lavor, C., Maculan, N. & Mucherino, A. 2014. Euclidean Distance Geometry and Applications. *SIAM Review,* 56(1), 3-69.

[9] Makeig, S. & Inlow, M. 1993. Lapse in alertness: coherence of fluctuations in performance and EEG spectrum. *Electroencephalography and Clinical Neurophysiology,* 86(1), 23-35.

[10] McDonald, A. D., Lee, J. D., Schwarz, C.& Brown, T. L. 2013. Steering in a Random Forest. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 56(5), 986-998.

[11] Murata, A. 2016. Proposal of a method to predict subjective rating on drowsiness using physiological and behavioral measures. *IIE Transactions on Occupational Ergonomics and Human Factors,* 7323, 1-13.

[12] National Sleep Foundation. 2024. *Drowsy driving: facts and stats* (https://www.thensf.org/drowsy-driving-prevention/). Las access on 16 December 2024.

[13] Olbrich, E., Landolt, H. P. & Achermann, P. 2014. Effect of prolonged wakefulness on electroencephalographic oscillatory activity during sleep. *Journal of Sleep Research, 23*(3), 255-262.

[14] Qian, D., Wang, B., Qing, X., Zhang, T., Zhang, Y., Wang, X. & Nakamura, M. 2017. Bayesian Nonnegative CP Decomposition-Based Feature Extraction Algorithm for Drowsiness Detection. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25*(8), 1297-1308.

[15] Radzi, S. S. M., Asirvadam, V. S. & Yusoff, M. Z. 2019. Fractal dimension and power spectrum of electroencephalography signals of sleep inertia state. *IEEE Access,* 7, 185879-185892.

[16] Zhao, C., Zheng, C., Zhao, M., Liu, J. & Tu, Y. 2011. Automatic classification of driving mental fatigue with EEG by wavelet packet energy and KPCA-SVM. *International Journal of Innovative Computing, Information and Control, 7*(3), 1157-1168.