

# Predictive Analysis of Driver Drowsiness Progression: Multi-Level Drowsiness Classification Using Physiological Signals

Natchira Dachoponchai<sup>1</sup>, Yodchanan Wongsawat<sup>1</sup>, and Jetsada Arnin<sup>1\*</sup>

<sup>1</sup> Department of Biomedical Engineering, Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand

E-mail: natchira.dac@student.mahidol.edu Tel: +66-892165777

E-mail: yodchanan.won@mahidol.ac.th Tel: +66-24414255

\*E-mail: jetsada.arn@mahidol.ac.th Tel: +66-835666542

**Abstract**— Drowsiness poses a significant challenge to cognitive and motor functions, compromising safety in critical tasks such as driving and increasing the risk of traffic accidents. Existing driver drowsiness detection systems inadequately address the gradual progression of drowsiness, focusing solely on binary classifications of drowsiness. This study aims to develop a neural network model that utilizes physiological signals, including EEG and ECG, to detect multiple levels of a driver's current drowsiness (Alert, Moderately Drowsy, or Extremely Drowsy). EEG and ECG data were collected from ten participants during 1-hour simulated driving experiments, supplemented by video recordings for once-per-minute drowsiness assessments through Observer Rating of Drowsiness (ORD) by two observers, which served as the ground truth. The neural network trained on 2-channel prefrontal EEG frequency-domain features, heart-rate variability (HRV) features, and driving time achieved accuracies of 92%, 77%, and 77% for Alert, Moderately Drowsy, and Extremely Drowsy, respectively. This high performance, with reliance on minimal electrodes and simple architecture, supports its feasibility for real-time applications as an early warning system for critical drowsiness in order to promote driver safety.

## I. INTRODUCTION

Drowsiness is a transitional state to falling asleep, characterized by fluctuating levels of alertness and the onset of sleep, manifested as difficulty in maintaining wakefulness and attentiveness during activities. It is a major concern in many daily-life scenarios, especially in attention-demanding tasks such as driving. Monotonous driving, lowers physiological arousal, sensorimotor functions, and information processing, thereby reducing the driver's capacity to react to sudden and critical situations on the road and making drivers prone to drowsiness [1]. Cognitive and behavioral changes as a result of drowsiness can not only reduce task efficiency but also compromise safety, potentially leading to life-threatening consequences [2].

In recent years, intensive research has been conducted into techniques for detecting driver drowsiness to create a more

objective evaluation. The current limitation of driver drowsiness detection systems lies in emphasising the binary classification of human mental states as *alert* or *drowsy*. Since drowsiness is a dynamic process, defining a threshold to separate mental states into only two extreme categories is too simplistic and not practical. Increasing the number of detection levels allows for a more comprehensive capture of the gradual progression of drowsiness, which is essential for timely and effective intervention. Therefore, this study primarily focuses on the three-level classification of driver drowsiness to identify the intermediate stages between extremely alert and drowsy.

Driver drowsiness detection systems can be classified into four main types: biological-based, image-based, vehicle-based, and hybrid-based methods [3-6]. Physiological measures achieve the highest accuracy among these categories, directly representing our functional state. EEG and ECG are two commonly used physiological signals in drowsiness detection, offering insights into the underlying neurological and cardiovascular systems. EEG, a brain electrical signal, is considered as the gold standard to indicate drowsiness as it represents the central nervous system activity [7]. The frequency shift of EEG signals is found to correlate with drowsiness. The study by Zhao *et al.* shows the rise in low-frequency power caused by drowsiness after driving for 90 minutes [8]. Various feature extraction methods are developed to obtain significant parameters underlying the complex and non-stationary dynamics of the brain [9]. The change in alertness affects the central nervous system and the autonomous nervous system, which is linked to HRV that can be extracted from ECG [10]. During drowsiness, our body becomes restful, where parasympathetic activity increases and sympathetic activity decreases. This relaxing response lowers the heart rate and increases the intervals between heartbeats, allowing high-frequency HRV to occur [11]. The previous study suggests that combining EEG and ECG features enhances

binary drowsiness detection model performance [12], making them the primary parameters for model development in this study.

Practicality is also important, apart from detection accuracy. This study aims to minimize the system's complexity to ensure feasibility for further real-world applications. We focus on reducing the electrode requirements to enhance the potential for developing wearable devices or gadgets and simplifying the detection model for real-time processing.

Thus, the main objective of this study is to develop the physiological-based multi-level driver drowsiness classification model using minimal EEG channels and individual ECG leads. Since NHTSA reports young adults aged 16-24 as a high-risk group for falling asleep while driving [13], young adult participants were chosen as representative EEG and ECG data acquisition groups in a controlled and safe simulated monotonous driving environment.

## II. MATERIALS AND METHODS

The development of a physiological-based machine learning model for multi-level driver drowsiness classification involves three phases: *Data Acquisition*, including driving simulator setup and data collection; *Data Preparation*, including data preprocessing, feature extraction and selection, and ground truth labeling; and *Model Development*, as shown in Fig. 1.

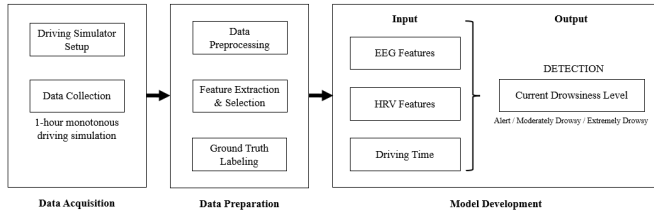


Fig. 1 Overview of Multi-Level Driver Drowsiness Classification Model Development

### A. Driving Simulator Setup

Data Acquisition was performed using a 1-hour driving simulation in the BCI Laboratory at Mahidol University with the 'City Car Driving' simulator and Logitech G27 USB Racing Wheel to mimic a monotonous driving environment. Physiological signals including EEG and ECG lead II were collected via a 16-channel g.USBamp RESEARCH device. Participant behaviors were monitored through video recordings using two cameras to capture overall body movements and facial expressions for drowsiness level labeling. Ethics approval for this experiment was granted by Mahidol University (MU-CIRB 2023/283.1209).

### B. Data Collection

Data for model development was obtained from 10 healthy young adults aged 20-24 with valid driver's licenses and no known neurologic or cardiac conditions that could interfere with EEG and ECG signal interpretation. Individuals who took medications that may affect neurological or cardiac function,

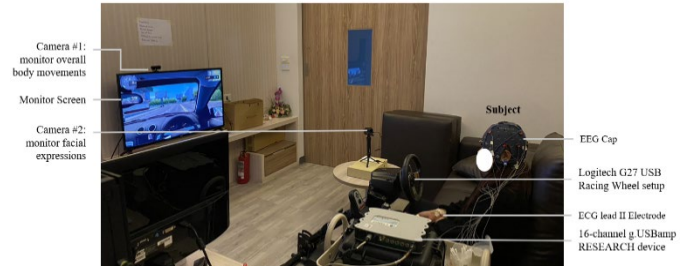


Fig. 2 Driving Simulator Setup at BCI Laboratory at Mahidol University

had a recent history of substance abuse in the past six months, consumed alcohol or caffeine within 8 hours before the study, or experienced recent traumatic events that might affect sleep patterns or mental state were excluded.

All participants were instructed to maintain a constant speed in a specified lane during a one-hour drive. To maintain a controlled environment and ensure the authenticity of participants' drowsiness progression, no more than three observers were present during the experiment, and their interaction with the subjects was minimized.

### C. Data Preprocessing

Data preprocessing was executed using MATLAB's Signal Processing and EEGLAB Toolboxes. EEG preprocessing involves Butterworth band-pass filtering with a low-pass filter at 40 Hz (order 9) and a high-pass filter at 0.5 Hz (order 6), followed by Artifact Subspace Reconstruction (ASR) to remove artifacts. ECG preprocessing includes Butterworth band-pass filtering at 1-40 Hz (order 3) and QRS detection via the Pan & Tompkins Algorithm. Both EEG and ECG were then split into 1-minute segments with 50% overlap for further analysis.

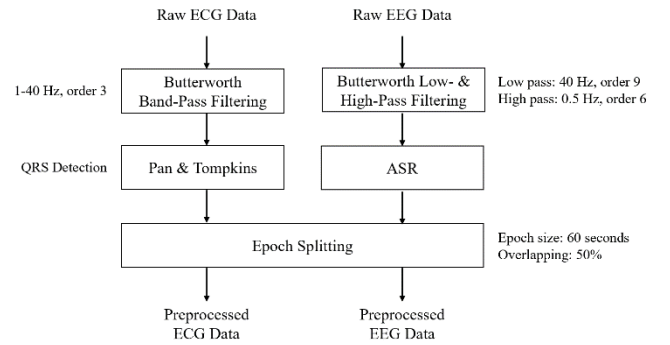


Fig. 3 Data Preprocessing Pipeline

### D. Feature Extraction and Selection

This study focuses on the frequency domain features of both EEG and ECG signals. Features were extracted from each EEG and ECG data segment using MATLAB's Signal Processing and EEGLAB Toolboxes. The feature set includes:

*EEG Features:* Relative Delta Power (0.5-4.0 Hz), Relative Theta Power (4.0-8.0 Hz), Relative Alpha Power (8.0-12.0 Hz), Relative Beta Power (12.0-30.0 Hz), Alpha/Beta Ratio, and (Theta+Alpha)/Beta Ratio

*ECG Features:* HRV features including Relative Low-Frequency (LF) Power (0.04-0.15 Hz), Relative High-Frequency (HF) Power (0.14-0.40 Hz), and LF/HF Ratio

Channels O1, O2, Fp1, and Fp2 were chosen as the primary EEG channels in this study due to their strong correlation with drowsiness, as indicated by previous research [12, 14] and their suitability for wearable device designs such as headbands. The feature sets from each channel were initially trained separately to determine the performance of single EEG channel usage. Subsequently, the combination of high-performing channels was examined to assess their integrated effects.

#### E. Ground Truth Labeling

ORD was applied for ground truth labeling, which involves assessing drowsiness levels by observing participants' behaviors and facial expressions from video recordings [15]. The 60-minute video data from the driving phase were divided into 60 segments, each one minute. For each video segment, two independent raters assigned a drowsiness level, categorized into five stages: Not Drowsy (scored as 1), Slightly Drowsy (2), Moderately Drowsy (3), Very Drowsy (4), and Extremely Drowsy (5). Then, the assessment scores from two raters were averaged, in order to minimize subjective bias. These labels for each minute of the sessions were aligned to the EEG and ECG segments at the same timestamps. In contrast, the segment that overlaps the labels was calculated by the mean of two consecutive labels that the segment overlaps. For the feasibility of model training, the average scores were used to categorize drowsiness into three simplified levels: Alert (average score  $\leq 2$ ), Moderately Drowsy ( $2 < \text{average score} \leq 3$ ), and Extremely Drowsy (average score  $> 3$ ). The cut-point of 3 for Extremely Drowsy was chosen because any segment with an average score above that must be labeled as 'Very Drowsy' or 'Extremely Drowsy' by at least one rater, which is critically harmful to driving. While not safe for proper driving, the Moderately Drowsy stage is an early warning stage before progressing to the critical Extremely Drowsy stage.

#### F. Model Development

Multilayer Perceptron (MLP) was used to predict simplified drowsiness levels as Alert, Moderately Drowsy, and Extremely Drowsy from EEG and HRV features. Due to the progressive nature of drowsiness in a monotonous environment, driving time was included as one of the input features to enhance model accuracy. The dataset was split into an 80% Training set and a 20% Testing set, maintaining class proportions. The model architecture was simplified to ensure practicality for further real-time applications, consisting of two hidden layers, each followed by a dropout layer and batch normalization layer. 5-fold cross-validation was performed to minimize overfitting due to the limited sample size, and class weights were applied to handle class imbalance as some subjects might rarely experience extreme drowsiness under the simulated

environment. Hyperparameters were tuned using Random Search, varying the optimizer, learning rate, dropout rates, batch size, epochs and units per layer. Several models were trained to compare performances across different feature sets. EEG features were used in every model, with comparisons made based on different EEG channels and the inclusion of HRV features. Additionally, the models with and without driving time as a feature were compared to investigate its impact on model performance. StandardScaler was used to rescale all features except driving time to preserve the consistent temporal properties.

### III. RESULTS

#### A. Comparison of Feature Sets

Table 1 shows the class-wise accuracy and F1 scores for models trained on different feature sets. The results demonstrate that including driving time and HRV features improves the models' performance across all EEG channels (O1, O2, Fp1, Fp2).

Table 1 Class-wise Accuracy and F1-scores for Models Trained on Different Feature Sets

EEG Channel	HRV features	Driving time	Class-wise Accuracy (%)			F1-score
			1	2	3	
O1	No	No	48.96	48.68	70.59	0.537
		Yes	92.71	53.95	20.59	0.635
	Yes	No	48.96	59.21	73.53	0.579
		Yes	78.13	61.84	73.53	0.720
O2	No	No	33.05	51.32	50.00	0.445
		Yes	79.66	59.21	47.06	0.687
	Yes	No	46.61	52.63	73.53	0.545
		Yes	73.73	78.95	50.00	0.725
Fp1	No	No	55.08	53.95	64.71	0.578
		Yes	83.05	65.79	79.41	0.772
	Yes	No	61.86	51.32	73.53	0.609
		Yes	83.05	75.00	85.29	<b>0.813</b>
Fp2	No	No	64.60	52.86	81.82	0.650
		Yes	84.07	57.14	86.36	0.759
	Yes	No	69.03	57.14	72.73	0.670
		Yes	84.07	52.86	72.73	0.728

Class 1: Alert, Class 2: Moderately Drowsy, Class 3: Extremely Drowsy

For all EEG channels, including driving time helps increase the class-wise accuracy of all classes, while adding HRV features results in less bias, more balanced class accuracies, and higher F1 scores. Among all EEG channels used, the Fp1 channel showed the most promising performance, with the model including both features achieving the highest overall F1-score of 0.813 and being the only one that reached an accuracy of at least 75% for all classes. When considering pairs of EEG channels, Fp1 and Fp2 generally provide higher accuracy than O1 and O2, representing a higher significant correlation with drowsiness classification. Thus, the combination of the former

pair of EEG with both HRV features and driving time included is investigated.

### B. Combination of Fp1 and Fp2 EEG channels

The new MLP is trained using an optimal combination of features, including EEG features from Fp1 and Fp2 channels, HRV features, and driving time. The predicted results, visualized in Fig. 4, indicate high true positive rates for Alert (92%), Moderately Drowsy (77%), and Extremely Drowsy (77%). The confusion matrix represents the model's high effectiveness in identifying the Alert state. For the Moderately Drowsy class, most errors are misclassified as a higher level of drowsiness, with small proportions falsely classified as Alert. This is acceptable for a system designed to proactively and early detect higher levels of drowsiness rather than miss detections. Regarding the Extremely Drowsy class, 23% of samples are misclassified as Moderately Drowsy, indicating some confusion between these two classes. However, none of the actual Extremely Drowsy data points are detected as Alert, demonstrating the model's ability to distinguish between extreme classes and ensure safety in drowsiness detection during the extremely drowsy state.

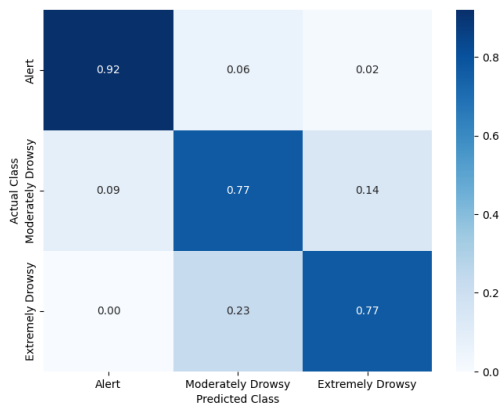


Fig. 4 Confusion Matrix for the Final Model Using Fp1 and Fp2 EEG Channels, HRV Features, and Driving Time

## IV. DISCUSSION

### A. EEG channels

The Fp1 and Fp2 channels demonstrated superior performance in classifying drowsiness levels compared to O1 and O2 channels. This can be attributed to their placement over the prefrontal cortex, a brain region associated with executive functions and alertness. As drowsiness progresses, the frequency of EEG waves in this area tends to shift from higher frequencies, like beta waves, to lower frequencies, including alpha and delta waves, due to increased neuron synchronization [16]. Additionally, changes in blink rates, which affect EEG signals in Fp1 and Fp2, may also occur. These factors make frequency domain features from these complementary channels in the prefrontal cortex significant for neural networks to

capture useful information from EEG and effectively classify drowsiness levels.

In contrast, the O1 and O2 channels are located over the occipital lobe and primarily process visual information. During relaxed states with eyes closed, sensory input reduction can lead to synchronous EEG and shifts to low-frequency ranges [17]. However, subjects must remain visually attentive during driving tasks despite cognitive declines. Therefore, Fp1 and Fp2 are more likely to reflect drowsiness-related changes during driving. Combining features from both channels can be more effective, especially when there is a loss of EEG signal or other issues.

### B. Combination of EEG and HRV Features

ECG signals provide valuable information about heart rate variability, which reflects autonomic nervous system activity that changes as drowsiness progresses. Biological signals vary more distinctly between extreme states, such as being extremely alert and drowsy. The transition period, however, is challenging for the model to identify due to the gradual nature of drowsiness.

The study shows that integrating HRV features effectively enhances the model's ability to identify this intermediate state. This improvement is likely due to the more holistic view of physiological changes provided by ECG, enabling the model to detect subtle changes during the transition period between extreme states. Moreover, HRV features help mitigate some limitations of EEG signals, which are more susceptible to artifacts and noise that might obscure the neural patterns associated with drowsiness. Therefore, combining EEG and HRV features as a multimodal approach significantly enhances the drowsiness classification process by the neural network compared to using EEG features alone.

### C. Driving Time

Adding driving time as a feature substantially impacts the model's performance. Drowsiness is a progressive condition, and including driving time helps the model understand the temporal context of the physiological signals. This temporal information allows the model to accurately capture the gradual transition between alertness and drowsiness and discern patterns and trends in the physiological data corresponding to different drowsiness stages. This is particularly important for detecting the Moderately Drowsy state, where transitions are more subtle and gradual. This finding also aligns with the previous study, which supports the positive impact of driving time on model performance [18].

### D. Model Practicality

The practicality of the final model, trained using features from Fp1 and Fp2 EEG channels, frequency-domain HRV features, and driving time, is considered based on model performance and the feasibility of wearable technology.

In terms of model performance, the model effectively distinguishes between different levels of drowsiness, including the *Moderately Drowsy* state, with no instances of Extremely Drowsy being misclassified as Alert. This capability enables early warnings that allow users to be notified and aware of their declining mental readiness for driving before reaching a critical stage of drowsiness.

For application in wearable technology, the model's reliance on minimal electrodes, including two EEG channels in the prefrontal area, with HRV features, makes it feasible to design devices that capture those signals during driving. EEG monitoring devices can take the form of headbands for less intrusiveness and a more comfortable setup for continuous monitoring, while heart rate can be monitored via wristbands or smartwatches. Using minimal features and simple model architecture reduces the complexity, making real-time processing more feasible.

## V. CONCLUSIONS

The study suggests that combining EEG and HRV features enhances the performance of 3-level drowsiness classification, particularly in identifying the intermediate state between alertness and critical drowsiness. Moreover, implementing driving time preserves the temporal pattern of drowsiness progression, leading to better overall accuracy. The final model based on 2-channel prefrontal EEG features, HRV features, and driving time yields high accuracy across all classes. Although there are some misclassifications, most errors in *Moderate Drowsy* detection are falsely predicted as *Extremely Drowsy*, which is acceptable for early warnings of critical drowsiness.

The model's use of minimal electrodes and simple architecture supports its practicality for real-time applications, contributing to driver safety. Future developments could include expanding the variety of driving conditions and the participants' age range to improve model generalization, developing wireless, real-time systems, and incorporating personalized participant data for improved accuracy. Additionally, pre-driving data could be applied to identify physiological markers correlating with drowsiness onset and progression, offering potential safety measures before driving.

## VI. ACKNOWLEDGMENT

This work was technically supported by the BCI Laboratory at Mahidol University and funded by the Faculty of Medicine Ramathibodi Hospital, Mahidol University.

## REFERENCES

- [1] D. J. Mascord and R. A. Heath, "Behavioral and physiological indices of fatigue in a visual tracking task," *Journal of Safety Research*, vol. 23, no. 1, pp. 19-25, 1992/03/01/ 1992, doi: [https://doi.org/10.1016/0022-4375\(92\)90036-9](https://doi.org/10.1016/0022-4375(92)90036-9).
- [2] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neuroscience & Biobehavioral Reviews*, vol. 44, pp. 58-75, 2014/07/01/ 2014, doi: <https://doi.org/10.1016/j.neubiorev.2012.10.003>.
- [3] Y. Albadawi, M. Takruri, and M. Awad, "A Review of Recent Developments in Driver Drowsiness Detection Systems," *Sensors*, vol. 22, no. 5, p. 2069, 2022. [Online]. Available: <https://www.mdpi.com/1424-8220/22/5/2069>.
- [4] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood, "A Survey on State-of-the-Art Drowsiness Detection Techniques," *IEEE Access*, vol. 7, pp. 61904-61919, 2019, doi: 10.1109/ACCESS.2019.2914373.
- [5] M. Shaik, "A systematic review on detection and prediction of driver drowsiness," *Transportation Research Interdisciplinary Perspectives*, vol.21, 06/23 2023, doi: 10.1016/j.trip.2023.100864.
- [6] V. Saini and R. Saini, "Driver drowsiness detection system and techniques: a review," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp. 4245-4249, 2014.
- [7] S. S. Poorna, A. Deep, K. Hariharan, R. R. Jain, and S. Krishnan, "A Transfer Learning Approach for Drowsiness Detection from EEG Signals," in *Innovations in Computational Intelligence and Computer Vision*, Singapore, M. K. Sharma, V. S. Dhaka, T. Perumal, N. Dey, and J. M. R. S. Tavares, Eds., 2021// 2021: Springer Singapore, pp. 369-375.
- [8] C. Zhao, M. Zhao, J. Liu, and C. Zheng, "Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator," *Accident Analysis & Prevention*, vol. 45, pp. 83-90, 2012/03/01/ 2012, doi: <https://doi.org/10.1016/j.aap.2011.11.019>.
- [9] I. Stancin, M. Cifrek, and A. Jovic, "A Review of EEG Signal Features and Their Application in Driver Drowsiness Detection Systems," *Sensors*, vol. 21, no. 11, p. 3786, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/11/3786>.
- [10] F. Shaffer and J. P. Ginsberg, "An Overview of Heart Rate Variability Metrics and Norms," (in English), *Frontiers in Public Health*, Review vol. 5, 2017-September-28 2017, doi: 10.3389/fpubh.2017.00258.
- [11] K. Fujiwara *et al.*, "Heart Rate Variability-Based Driver Drowsiness Detection and Its Validation With EEG," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 6, pp. 1769-1778, 2019, doi: 10.1109/TBME.2018.2879346.
- [12] M. Awais, N. Badruddin, and M. Driberg, "A Hybrid Approach to Detect Driver Drowsiness Utilizing Physiological Signals to Improve System Performance and Wearability," (in eng), *Sensors*, vol. 17, no. 9, Aug 31 2017, doi: 10.3390/s17091991.
- [13] *NHTSA Drowsy Driving Research and Program Plan*, 2016. [Online]. Available: [https://www.nhtsa.gov/sites/nhtsa.gov/files/drowsydriving\\_strategicplan\\_030316.pdf](https://www.nhtsa.gov/sites/nhtsa.gov/files/drowsydriving_strategicplan_030316.pdf).

- [14] S. Chaabene, B. Bouaziz, A. Boudaya, A. Hökelmann, A. Ammar, and L. Chaari, "Convolutional Neural Network for Drowsiness Detection Using EEG Signals," *Sensors*, vol. 21, no. 5, p. 1734, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/5/1734>.
- [15] W. W. Wierwille and L. A. Ellsworth, "Evaluation of driver drowsiness by trained raters," (in eng), *Accid Anal Prev*, vol. 26, no. 5, pp. 571-81, Oct 1994, doi: 10.1016/0001-4575(94)90019-1.
- [16] T. Kirschstein and R. Köhling, "What is the Source of the EEG?," *Clinical EEG and Neuroscience*, vol. 40, no. 3, pp. 146-149, 2009, doi: 10.1177/155005940904000305.
- [17] F. L. Britton JW, Hopp JLet al., F. L. St. Louis EK, Ed. *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants*. Chicago: American Epilepsy Society, 2016.
- [18] C. Jacobé de Naurois, C. Bourdin, A. Stratulat, E. Diaz, and J. L. Vercher, "Detection and prediction of driver drowsiness using artificial neural network models," (in eng), *Accident Analysis & Prevention*, vol. 126, pp. 95-104, May 2019, doi: 10.1016/j.aap.2017.11.038.