

Article

Classification of Drowsiness and Alertness States Using EEG Signals to Enhance Road Safety: A Comparative Analysis of Machine Learning Algorithms and Ensemble Techniques

Masoud Sistaninezhad ¹, Saman Rajebi ¹, Siamak Pedrammehr ², Arian Shajari ³,
Hussain Mohammed Dipu Kabir ^{4,5,*}, Thuong Hoang ⁶, Stefan Greuter ⁷ and Houshyar Asadi ³

¹ Department of Electrical Engineering, Seraj University, Tabriz 5137894797, Iran; b.sistaninezhad.ms@seraj.ac.ir (M.S.); s.rajebi@seraj.ac.ir (S.R.)

² Faculty of Design, Tabriz Islamic Art University, Tabriz 5164736931, Iran; s.pedrammehr@tabriziau.ac.ir

³ Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Warrn Ponds, VIC 3216, Australia; ashajari@deakin.edu.au (A.S.); houshyar.asadi@deakin.edu.au (H.A.)

⁴ Artificial Intelligence and Cyber Futures Institute, Charles Sturt University, Bathurst, NSW 2795, Australia

⁵ Rural Health Research Institute, Charles Sturt University, Orange, NSW 2800, Australia

⁶ Faculty of Science Engineering and Built Environment, Deakin University, Warrn Ponds, VIC 3216, Australia; thuong.hoang@deakin.edu.au

⁷ School of Communication and Creative Arts, Deakin University, Burwood, VIC 3125, Australia; stefan.greuter@deakin.edu.au

* Correspondence: hkabir@csu.edu.au

Abstract

Drowsy driving is a major contributor to road accidents, as reduced vigilance degrades situational awareness and reaction control. Reliable assessment of alertness versus drowsiness can therefore support accident prevention. Key gaps remain in physiology-based detection, including robust identification of microsleep and transient vigilance shifts, sensitivity to fatigue-related changes, and resilience to motion-related signal artifacts; practical sensing solutions are also needed. Using Electroencephalogram (EEG) recordings from the MIT-BIH Polysomnography Database (18 records; >80 h of clinically annotated data), we framed wakefulness–drowsiness discrimination as a binary classification task. From each 30 s segment, we extracted 61 handcrafted features spanning linear, nonlinear, and frequency descriptors designed to be largely robust to signal-quality variations. Three classifiers were evaluated—k-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Decision Tree (DT)—alongside a DT-based bagging ensemble. KNN achieved 99% training and 80.4% test accuracy; SVM reached 80.0% and 78.8%; and DT obtained 79.8% and 78.3%. Data standardization did not improve performance. The ensemble attained 100% training and 84.7% test accuracy. While these results indicate strong discriminative capability, the training–test gap suggests overfitting and underscores the need for validation on larger, more diverse cohorts to ensure generalizability. Overall, the findings demonstrate the potential of machine learning to identify vigilance states from EEG. We present an interpretable EEG-based classifier built on clinically scored polysomnography and discuss translation considerations; external validation in driving contexts is reserved for future work.

Keywords: alertness; drowsiness; EEG; classification; machine learning; ensemble techniques



Academic Editor: Paolo Bellavista

Received: 2 August 2025

Revised: 7 November 2025

Accepted: 19 November 2025

Published: 24 November 2025

Citation: Sistaninezhad, M.; Rajebi, S.; Pedrammehr, S.; Shajari, A.; Dipu Kabir, H.M.; Hoang, T.; Greuter, S.; Asadi, H. Classification of Drowsiness and Alertness States Using EEG Signals to Enhance Road Safety: A Comparative Analysis of Machine Learning Algorithms and Ensemble Techniques. *Computers* **2025**, *14*, 509. <https://doi.org/10.3390/computers14120509>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Road accidents are one of the most frequent and devastating occurrences that can lead to loss of life and financial damage [1]. There are several factors that contribute to road accidents, such as the drivers' mental and physical states, technical malfunctions of equipment, the influence of drugs, driver fatigue and drowsiness, and other related factors [2]. Drowsiness is a condition that occurs when one transitions from being awake to asleep, and vice versa. It is characterized by a decrease in alertness and ability to make quick decisions [3]. This can be particularly dangerous for drivers, as statistics from around the world have shown that driver fatigue is a leading cause of traffic accidents. To address this, car companies have invested millions of dollars in developing warning systems that would detect drowsiness. As the number of on-road vehicles continues to increase, the problem of drowsy driving is becoming more prevalent [4]. Therefore, it is crucial to identify signs of driver drowsiness and develop intelligent drowsiness detection systems to prevent and decrease the occurrence of accidents. The accurate detection of drowsiness is a primary objective in the advancement of novel driver assistance systems and advanced drowsiness detection methods. Drowsiness can be assessed in three ways: based on behavior, vehicle, and physiological criteria [5].

Contemporary driver-assistance platforms including those from Tesla, Volvo, and Mercedes-Benz incorporate camera-based fatigue monitoring to detect distraction and drowsiness. Despite their practicality, such systems are vulnerable to environmental variation (lighting, viewpoint, occlusion) and may pose privacy concerns. Physiological sensing, particularly EEG, offers a complementary pathway by capturing neural correlates of drowsiness that are less dependent on external conditions. Accordingly, the proposed framework is designed to augment existing in-vehicle monitoring solutions.

The EEG data analyzed here originate from the MIT-BIH Polysomnography Database collected in controlled clinical sleep laboratories, not driving settings. As such, this study is a methodological investigation to develop and evaluate an EEG-based classifier that can subsequently be adapted and tested in more ecologically valid driving scenarios.

One way to evaluate drowsiness in drivers is by observing their behavior, such as yawning, head movements, blink rate, and duration of eyelid closure. Using a camera to capture images of drivers' faces is a noninvasive and convenient method, but it can be affected by environmental factors such as lighting, camera movements, and viewing angles. However, drivers may be uncomfortable with direct camera monitoring, and adjusting the camera to monitor the driver's alertness or drowsiness involves fine-tuning parameters such as viewing angle, focal length, lens distortion correction, and coordinating with the installation environment including ambient light and positioning. This precise adjustment of hardware and software is needed for each driver. Furthermore, relying solely on behavioral criteria may not always accurately detect drowsiness. Also, evaluating behavioral criteria alone may not provide accurate results in detecting drowsiness. Vehicle-based metrics make use of one or more sensors that are embedded inside the car and can continuously monitor the driver's head, hand, and eye movements [6,7]. Car-based criteria comprise various factors such as the position of the wheels, hand movements, speed, acceleration of the vehicle, etc. These criteria are noninvasive, but they highly depend on the driver's driving skills, road conditions, and vehicle characteristics [8]. Although the advantage of these criteria is that they are noninvasive, the disadvantage is that they take a long time to detect the presence of a problem in the driver, and they are unable to prevent an accident in real driving conditions [9]. Physiological criteria are an alternative and complement to vehicle and behavioral criteria. These criteria include various indicators, such as heart rate, brain activity, breathing rate, etc., which can be obtained from sensors recording signals, including Electrocardiogram (ECG), Electroencephalogram (EEG), Photoplethysmography

(PPG), etc. [10,11]. It is possible to analyze the level of drowsiness through heart activity. Heart Rate Variability (HRV) is used to measure the variance of intervals between heartbeats, which is a valid indicator of the activity of the cardiovascular system in different human physiological states [12]. One of the drawbacks of medical signal-based methods is the necessity for sensors and cables to be attached to the body for recording [13–15]. However, this problem can be effectively addressed by incorporating new wireless sensors, such as smart watches and wearables. The brain’s electrical activity exhibits distinct wave patterns in certain areas, and the differences in these patterns between alertness and drowsiness have been extensively researched. Several techniques have been developed to process the EEG signal and identify drowsiness [16–19].

To provide context for our contribution, Table 1 presents a comparative overview of selected EEG-based drowsiness detection studies. The table outlines key methodologies, feature types, reported accuracy rates, and associated limitations. This comparison highlights the effectiveness of our ensemble-based model, which achieves competitive performance using exclusively handcrafted EEG features.

Table 1. Comparative analysis of the proposed method with related literature in terms of methods, feature extraction approaches, classification accuracy, and key remarks.

Reference	Method(s) Used	Features/Approach	Accuracy (%)	Remarks
Kiyimik et al. [7]	Wavelet + ANN	EEG wavelet features	84.1	Early work; limited feature types
De Gennaro et al. [8]	EEG quantitative features	Sleep onset EEG changes	~80	General EEG analysis
Delimayanti et al. [9]	High-dimensional FFT + ML	Sleep stage EEG classification	~83	Uses FFT spectrum only
Kim. [10]	EEG + ML	Frequency domain analysis	78.2	Focus on sleep staging
Chen et al. [20]	EEG + CCA + EOG + ELM	Ensemble of EEG/EOG features	95.6	Hybrid EEG + EOG, high complexity
Sistaninezhad et al. [21]	PPG + ML	Smartphone PPG-based drowsiness	89.2	Alternative to EEG
Christensen et al. [22]	Data-driven EEG + sleep staging	Insomnia EEG	~85	Insomnia-specific
This study (EL model)	61 handcrafted EEG features + Bagging Ensemble (DT)	EEG only, nonlinear + frequency + time features	84.7 (Test) 100 (Train)	Superior or comparable accuracy with single-modality EEG features

The level of vigilance is tightly linked to activity in specific brain regions [23]. We propose a lightweight, interpretable EEG-based method for drowsiness detection using data from the MIT-BIH Polysomnography database, which contains multi-channel recordings from clinical sleep studies. To promote computational simplicity and near-real-time feasibility, all experiments use a single EEG derivation (C4–A1)—a choice supported by prior evidence that single-channel EEG, when paired with robust processing, can capture drowsiness-related neural dynamics [7,11].

Signals were preprocessed and segmented into 30 s epochs, and 61 handcrafted features—covering linear, nonlinear, and frequency-based descriptors—were extracted. These features served as inputs to KNN, SVM, DT, and a bagging ensemble. The objective is an accurate, efficient, and explainable pipeline suitable for driver-monitoring contexts.

Consistent with the nonstationary and nonlinear nature of EEG during sleep, the selected features capture evolving temporal and complexity patterns associated with

alertness–drowsiness states. (Section 2 reviews prior work; Section 3 details the dataset; Section 4 describes the methodology; Section 5 reports evaluations; Section 6 concludes with implications and future directions.)

Contributions:

1. A large-scale, balanced dataset of 6212 labeled 30 s EEG segments drawn from >80 h of MIT-BIH polysomnography.
2. Design and extraction of 61 handcrafted features (linear, nonlinear, frequency-based) chosen for robustness to signal noise/quality, offering broader coverage than narrowly statistical or purely deep-learning approaches.
3. A comparative analysis across multiple classifiers (KNN/SVM/DT) and a DT-based bagging ensemble, all trained and evaluated on the same data.
4. Bayesian hyperparameter optimization and performance reporting with six metrics—Accuracy, Precision, Sensitivity, Specificity, F1, and MCC—to support robust evaluation.

We posit that combining statistical, frequency, and model-based EEG features can reliably separate alert from drowsy states while preserving interpretability. Although EEG is sensitive to subtle neural changes, real-world deployment faces challenges (user comfort, electrode placement, motion artifacts). Accordingly, this work is an exploratory feasibility study under controlled conditions, identifying salient EEG features and evaluating interpretable ML models as a foundation for future simplified or hybrid solutions that pair physiological markers with practical deployment strategies. For deployment, single-channel designs are attractive because they reduce computational load and can be implemented with wearable or in-cabin sensors; such engineering integration is discussed as future work rather than a component of the present clinical analysis.

Because drowsiness exhibits consistent EEG signatures—elevated theta, reduced beta, transient alpha bursts, and lower nonlinear complexity—we frame the task as a physiology-centric EEG state-classification problem rather than a driving-performance study. Clinically annotated transitions enable precise labeling for an interpretable baseline model, and we outline steps for adaptation to simulator/on-road data where domain shift and artifact profiles differ.

2. Literature Review and Study Contributions

2.1. Related Work on EEG-Based Drowsiness Detection

Previous studies have explored the use of photoplethysmography (PPG) signals to examine the connection between heart rate variability and stress levels in individuals. A crucial parameter in this context is pulse transit time, which correlates closely with blood pressure changes [13,18,22]. Additionally, the pulsatile blood flow causes subtle color variations on the skin surface. To utilize this effect, certain methods deploy low-frame-rate cameras to capture facial movements, enabling the estimation of pulse transit time and subsequent reconstruction of the PPG signal [18,22–24]. These techniques have shown promising applicability in vehicular environments. PPG offers a noninvasive means to assess vascular properties such as arterial stiffness, elasticity, and microvascular blood volume fluctuations [21]. The cardiac cycle produces a pressure wave that propagates blood through tissues, causing volume changes detectable by illuminating the skin with a light-emitting diode (LED) and measuring the transmitted light via a photodiode sensor. Figure 1 illustrates the architecture of a typical PPG signal acquisition system [16].

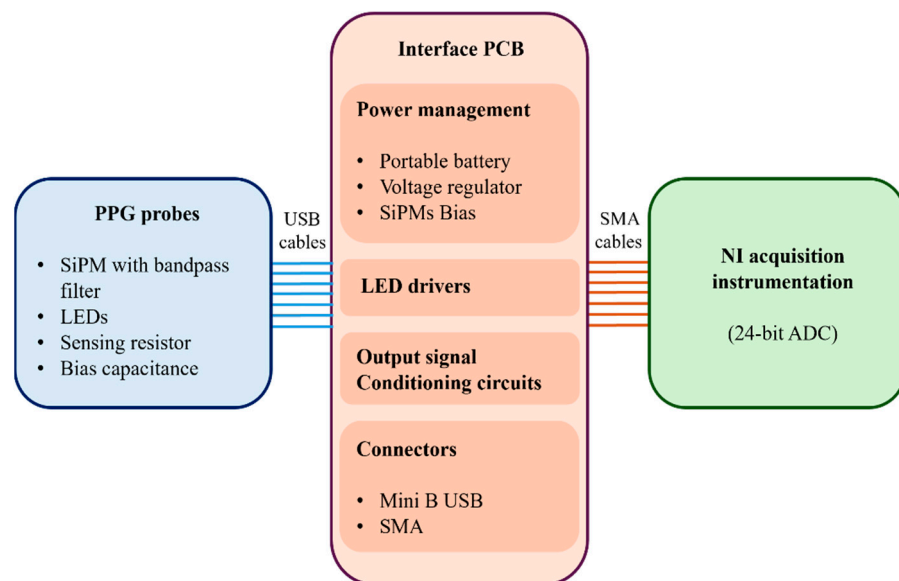


Figure 1. Block diagram of PPG signal sampling system.

Drowsiness can be inferred from heart rate variability (HRV), estimated from troughs of the PPG signal. A practical advantage of PPG is single-hand acquisition, which avoids requiring both hands on the steering wheel; however, such approaches often rely on relatively complex hardware (e.g., steering-wheel ECG/PPG sensors) and continuous skin contact, which may be impractical in real use [24]. Alternative lines of work reconstruct PPG from facial video to estimate HRV and driver state, but these methods are sensitive to camera placement, lighting, and per-user calibration, and purely behavioral/facial cues can be unreliable due to inter-subject variability and algorithmic constraints. In contrast, the present study focuses exclusively on EEG, offering a direct physiological assessment of vigilance that does not depend on visual inputs or vehicle-embedded sensors.

In the literature, two broad families of drowsiness evaluation have been reported. The first uses EEG, which underpins applications in gaming, psychotherapy, drowsiness assessment, and certain neurorehabilitation contexts [25]. Prior EEG work commonly employs frequency-domain features—e.g., PSD or wavelet-based descriptors—and trains ANN classifiers, with accuracies around 84.1% in some reports. For example, Chen et al. [20] and Delimaynati et al. [9] combined wavelet-band and Fourier-based spectral EEG features with EOG eyelid-movement cues and classified them using Extreme Learning Machine (ELM), a fast single-hidden-layer approach that sets input weights randomly and solves output weights by least squares; they reported accuracies up to 95.6%. (ELM and the referenced multimodal features are part of prior work and are not used in the present study.)

A second stream integrates traditional signal processing with deep-learning feature extraction to improve accuracy. Reported features include energy distribution, zero-crossing velocity, spectral entropy, and instantaneous frequency structures [7,9,10]. In these studies, alpha-band activity is often isolated from EDF-formatted PhysioNet EEG, PSD powers in delta/theta/alpha are estimated (typically via FFT), and classifiers such as ANN and SVM are evaluated with accuracy and ROC metrics. A representative block diagram of such pipelines appears in Figure 2.

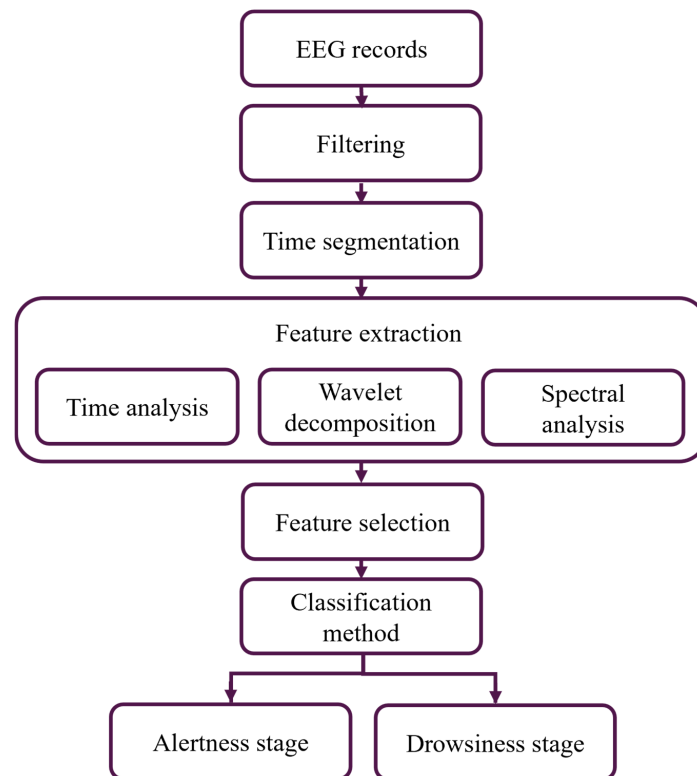


Figure 2. Process of the proposed algorithm.

In prior work, feature extraction commonly combines time-domain statistics (e.g., mean, variance, Hjorth parameters) with frequency-domain analyses (e.g., band power and power spectral density, PSD) and time–frequency methods (e.g., wavelet decomposition). Such multi-domain pipelines provide a comprehensive characterization of temporal and spectral patterns underlying transitions between alertness and drowsiness, followed by a classification stage. Figure 3 illustrates representative EEG traces for alert versus drowsy conditions.

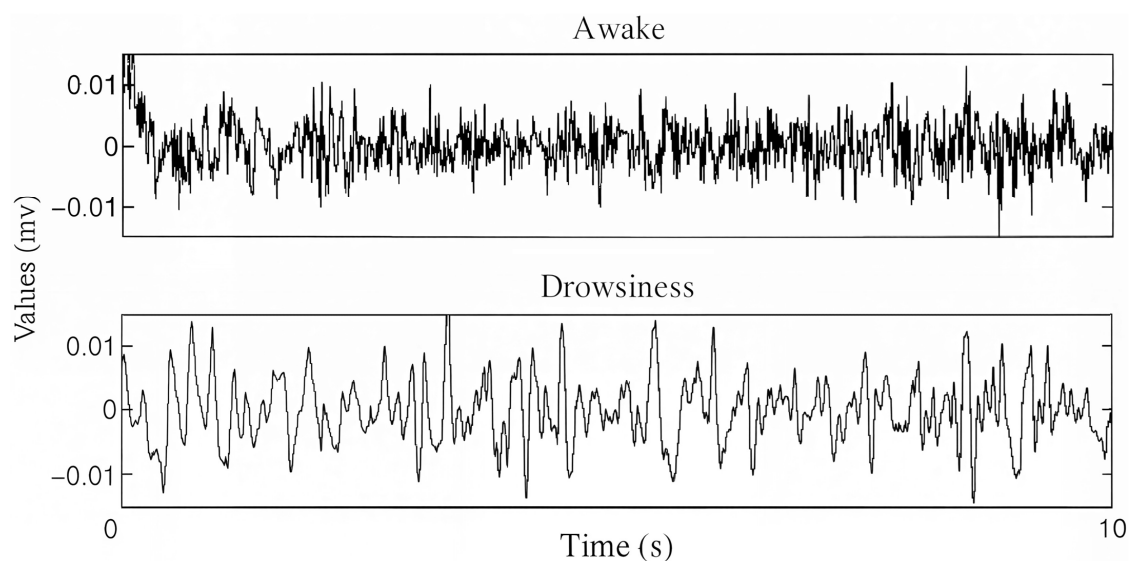


Figure 3. A simple EEG signal recorded from one participant, using the C4–A1 channel (C4 electrode referenced to A1) during alertness and drowsiness states. (Note: No visible power line interference is present due to the controlled recording environment and use of low-pass filtering and detrending.).

In related work, EEG preprocessing has often used a two-stage filtering pipeline: a bidirectional Butterworth filter followed by a low-pass stage (cutoff within 0.5–60 Hz), with adaptive filters to suppress biological artifacts (e.g., speech, eye movements) and power-line interference [22]. While some studies segmented EEG into 5 s epochs for PSD estimation under a quasi-stationarity assumption, that segmentation strategy was not adopted in the present study.

For feature selection and interpretability, prior work [22] applied Linear Discriminant Analysis (LDA) with a stepwise (forward/backward) procedure based on Lambda Prediction (LW) to rank feature importance. Using the most discriminative features, Artificial Neural Networks (ANNs) were trained: 21 three-layer architectures were explored with input sizes of 8/12/27 features, 10–40 hidden neurons, tansig output activation, and Levenberg–Marquardt training. Data were split 70%/30% for train/test, with validation after each training cycle; results (metrics and confusion matrices) are reported in Tables 2 and 3 of that study [22].

Table 2. The confusion matrix and classification results of the LDA.

Features No.	State	Confusion Matrix		Classification Results
		A (%)	D (%)	Accuracy (%)
19 Features	A	70.5	29.5	72.9
	D	24.6	75.4	
13 Features	A	70.3	29.7	72.9
	D	24.5	75.5	
7 Features	A	67.9	32.1	71.7
	D	24.5	75.5	

A: Alertness; D: Drowsiness.

Table 3. The results of ANN.

Hidden Layer Neurons No.	No. of Features in the Input Vector						
	19			13		7	
	State	A	D	A	D	A	D
10	A	86.79	13.21	87.67	12.33	86.81	13.19
	D	15.50	84.50	18.42	81.58	17.21	82.79
15	A	87.96	12.04	85.81	14.19	86.22	13.78
	D	16.42	83.58	16.64	83.36	23.04	76.96
20	A	86.96	13.04	85.66	14.34	86.86	13.14
	D	15.53	84.47	16.50	83.50	16.50	83.50
25	A	88.30	11.70	86.47	13.53	86.56	13.44
	D	16.18	83.82	15.02	84.98	15.61	84.39
30	A	87.33	12.67	86.05	13.95	87.02	12.98
	D	14.47	85.53	18.51	81.49	15.69	84.31
35	A	87.11	12.89	85.58	14.42	87.46	12.54
	D	15.13	84.87	14.86	85.14	16.14	83.86
40	A	87.39	12.61	87.02	12.98	86.82	13.18
	D	15.07	84.93	16.00	84.00	15.42	84.58

A: Alertness; D: Drowsiness.

Note: Techniques such as PSD segmentation, LDA-based feature selection, and ANN classification are discussed solely as part of prior literature and are not incorporated into the methodology of the present study.

2.2. Contribution of This Study

Building upon the strengths and addressing the limitations of previous EEG-based drowsiness detection research, this study introduces a lightweight and interpretable framework utilizing 61 handcrafted features extracted from time-domain statistics, frequency-domain energy distributions, and model-based parameters to classify states of drowsiness and alertness. Unlike deep learning models that demand extensive datasets and high computational resources, our approach emphasizes simplicity and transparency by employing machine learning classifiers such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), and ensemble bagging methods. Diverging from prior techniques that depend on multiple EEG channels or auxiliary modalities, we demonstrate that effective classification can be achieved using a single EEG electrode, facilitating real-time implementation in embedded or wearable devices. Our comparative analysis further substantiates that ensemble learning markedly enhances classification accuracy while maintaining computational efficiency.

While previous studies, applied Linear Discriminant Analysis (LDA) for feature selection, our approach intentionally avoids LDA or any dimensionality reduction methods to retain complete interpretability of the extracted features.

3. Data Description

This study uses the MIT-BIH Polysomnography Database, comprising multi-physiological overnight recordings acquired at the Beth Israel Hospital Sleep Laboratory (Boston) for the monitoring of obstructive sleep apnea and evaluation of Continuous Positive Airway Pressure (CPAP) therapy. For analysis, EEG was segmented into 30 s epochs and mapped to two classes: Alertness, defined as expert-labeled Wake (W) with predominant beta activity; and Drowsiness, defined as stages N1–N3, with N1 reflecting microsleep-like episodes relevant to driving contexts. REM epochs were excluded because their physiology is not directly aligned with the wake-to-sleep transition of interest. Using this binary scheme, we assembled a balanced set of 6212 EEG epochs. Overall, the database provides > 80 h of polysomnography across four-, six-, and seven-channel montages. Each record includes beat-by-beat labeled ECG and EEG/respiratory channels annotated for sleep staging and apnea events [26,27]. The dataset has been widely employed in research on sleep staging and vigilance, and its clinically curated annotations underpin numerous EEG studies [26,27].

Scope note: EEG was obtained in a clinical polysomnography paradigm to support EEG-based state classification (wakefulness vs. drowsiness), rather than assessment of driving behavior; implications for deployment in vehicles are addressed in the Discussion. Recordings were sampled at 256 Hz with 16-bit resolution. Although multi-channel signals are available, the present analysis focuses on the C4–A1 derivation.

4. Methodology

4.1. EEG Acquisition

Scalp EEG from the MIT-BIH Polysomnography Database was acquired in a clinical sleep laboratory using standard PSG instrumentation and expert scoring. The analysis focused on the C4–A1 derivation, a central lead commonly used for drowsiness studies due to reduced ocular contamination and sensitivity to sleep-onset dynamics. Signals were recorded with clinical grounding and digitized at the database's native specifications (256 Hz, 16-bit). Sleep stages were scored by clinical experts according to standard

guidelines, and labels were mapped to Wake vs. drowsiness as defined in the labeling section. To attenuate acquisition-related artifacts, EEG was band-pass filtered 0.5–30 Hz (relevant alertness/drowsiness bands); no notch filter was applied given minimal power-line contamination in this dataset and adequate suppression by the band-pass.

4.2. Signal Preprocessing

EEG preprocessing used the MIT-BIH Polysomnography data segmented into 30 s, expert-labeled epochs. For computational simplicity and real-time suitability, analysis was restricted to the C4–A1 channel. The pipeline comprised:

- Band-pass filtering (0.5–30 Hz, 4th-order, zero-phase Butterworth) to retain components relevant to alertness/drowsiness while suppressing slow drift and high-frequency noise; no notch filter was applied given minimal power-line contamination and adequate attenuation by the band-pass.
- Segmentation and labeling: 30 s epochs kept their clinical labels; Wake (W) epochs were treated as alert, and N1–N3 as drowsy; REM was excluded.

All steps follow established EEG practices and were implemented with custom MATLAB R2024a scripts. The focus on C4–A1 reflects its sensitivity to sleep-onset dynamics with reduced ocular contamination, consistent with prior drowsiness literature [7,11].

To reduce baseline drift, signals were mean-centered and de-trended (least-squares line removal) within each labeled epoch following band-pass filtering. Processing parameters were derived from the training portion only during evaluation to prevent information leakage. This filtering configuration aligns with common EEG protocols in drowsiness and sleep studies [7,20].

4.3. Feature Extraction

From each 30 s EEG epoch (C4–A1), we extracted 61 handcrafted features capturing temporal statistics, nonlinear dynamics, and model-based descriptors. The feature set comprised:

- Time-domain statistics: mean, variance, skewness, kurtosis, Hjorth parameters, zero-crossing rate, and related waveform-shape indices.
- Nonlinear/complexity metrics: Shannon entropy, fractal dimension, Hurst exponent, and Detrended Fluctuation Analysis (DFA).
- Model-based descriptors: low-order autoregressive (AR) coefficients, signal-energy measures, and dominant-frequency estimates.

Clarification on spectral methods and non-stationarity. No Power Spectral Density (PSD) or other Fourier-based descriptors were computed in this study; any PSD discussion appears only in the literature review to contextualize prior work. The present feature set is limited to time-domain statistics, nonlinear/complexity metrics, and low-order AR descriptors, which are less sensitive to EEG non-stationarity.

AR-order selection. The AR model order was fixed at 10 based on (i) empirical sweeps (orders 4–14) showing no consistent performance gain beyond 10 but increasing parameter variance, and (ii) established EEG practice for 256 Hz data below 30 Hz, where 8–12 coefficients provide a balanced trade-off between spectral fidelity and stability.

Given the nonstationary nature of EEG signals, the 30 s EEG epochs were divided into 500 ms rectangular windows with a 400 ms overlap to ensure local stationarity. Within each short window, the signal can be reasonably considered quasi-stationary, allowing reliable estimation of autoregressive (AR) parameters.

A 10th-order AR model was fitted to each window using the Yule–Walker method implemented in MATLAB’s Signal Processing Toolbox. This window length (500 ms) provided an optimal balance between temporal resolution and the statistical reliability

of parameter estimation, given the sampling rate of over 200 Hz. Such a setting ensures sufficient data points per window for stable parameter estimation while maintaining sensitivity to short-term EEG dynamics.

For each window, ten AR coefficients (a_1 – a_{10}) were obtained. To summarize these parameters at the segment level, coefficients corresponding to the same order (e.g., a_1 across all windows) were averaged across all windows, resulting in ten representative AR features per 30 s EEG segment. This averaging approach preserves the overall temporal structure of the signal while minimizing the influence of transient fluctuations.

All coefficients were z-score normalized prior to aggregation to ensure consistency across different windows and subjects. This procedure ensures that AR modeling was performed under locally stationary conditions and that the resulting features robustly capture the underlying temporal dynamics of the EEG signal. The adopted framework aligns with established practices in EEG autoregressive and spectral modeling literature, which recommend sub-second windowing for reliable AR estimation on nonstationary EEG data.

All features were computed from the preprocessed signals described in Section 4.1. Windowing was applied only for the AR-based (model-derived) descriptors as detailed above, where 500 ms overlapping windows were used to ensure local stationarity. In contrast, all other features—time-domain, statistical and nonlinear/complexity measures—were directly extracted from the full 30 s EEG epochs without further segmentation. To improve numerical stability, both min–max and z-score normalization were applied, and all features were evaluated individually and jointly as classifier inputs. No feature selection or dimensionality-reduction step was applied; all 61 features were retained to preserve interpretability and traceability to the underlying EEG mechanisms (Table 4).

Table 4. Extracted features categorized by domain and feature name.

Domain	Feature Name
Time-Domain	Clearance Factor, Crest Factor, Impulse Factor, Kurtosis, Mean, Peak Value, Root Mean Square (RMS), Signal-to-Noise and Distortion Ratio (SINAD), Signal-to-Noise Ratio (SNR), Shape Factor, Skewness, Standard Deviation, Total Harmonic Distortion (THD), Minimum, Median, Maximum, First Quartile (Q1), Third Quartile (Q3), Interquartile Range (IQR)
Detrended Time-Domain	Clearance Factor, Crest Factor, Impulse Factor, Kurtosis, Mean, Peak Value, RMS, SINAD, SNR, Shape Factor, Skewness, Std. Deviation, THD, Minimum, Median, Maximum, Q1, Q3, IQR.
Model-Based	AR Coefficient 1, Dominant Frequency, Damping Ratio, Mean Squared Error (MSE), Mean Absolute Error (MAE), Akaike Information Criterion (AIC), Residual Mean, Residual Variance, Residual RMS, Residual Kurtosis
Model-Based (Detrended)	AR Coefficient 1, Dominant Frequency, Damping Ratio, MSE, MAE, AIC, Residual Mean, Residual Variance, Residual RMS, Residual Kurtosis.
Frequency Domain	Peak Amplitude, Peak Frequency

Each 30 s EEG segment was processed using a custom MATLAB R2024a implementation of standard signal-processing routines. The extracted features were organized into three principal categories: (1) time-domain statistics (e.g., mean, RMS, variance/kurtosis/skewness, SNR); (2) model-based time-series descriptors derived from a 10th-order AR model, including dominant frequency, damping ratio, and residual-error statistics; and (3) nonlinear/complexity measures such as entropy-based indices and DFA, as detailed in Table 4.

Signals were detrended after band-pass filtering, and features were computed from the resulting preprocessed C4–A1 epochs. The extracted feature set was subsequently used to train and evaluate well-established classifiers—KNN, SVM, DT, and a DT-based bagging ensemble (EL)—commonly adopted in EEG classification research [28,29]. Performance was assessed using Accuracy, Precision, Sensitivity, Specificity, F1-Score, and Matthews Correlation Coefficient (MCC).

4.4. Classification Algorithms

ML algorithms are extensively employed for classification in medical diagnosis. This research assesses the efficacy of SVM, KNN, DT, and DT-based bagging EL algorithms in the classification of drowsiness and alertness. The selection of these algorithms was based on their unique operational features and prevalence in research. The following is a comprehensive overview of the functioning of each algorithm.

4.4.1. Support Vector Machine (SVM)

The SVM is a powerful tool that is used for classifying data effectively [24]. Its operational principle is to find a linear separator between different classes of data, with the aim of maximizing the distance from each class. This technique works well for data with only two classes, but for data with multiple classes, it must be calculated in pairs. Figure 4 shows the schematic of SVM algorithm.

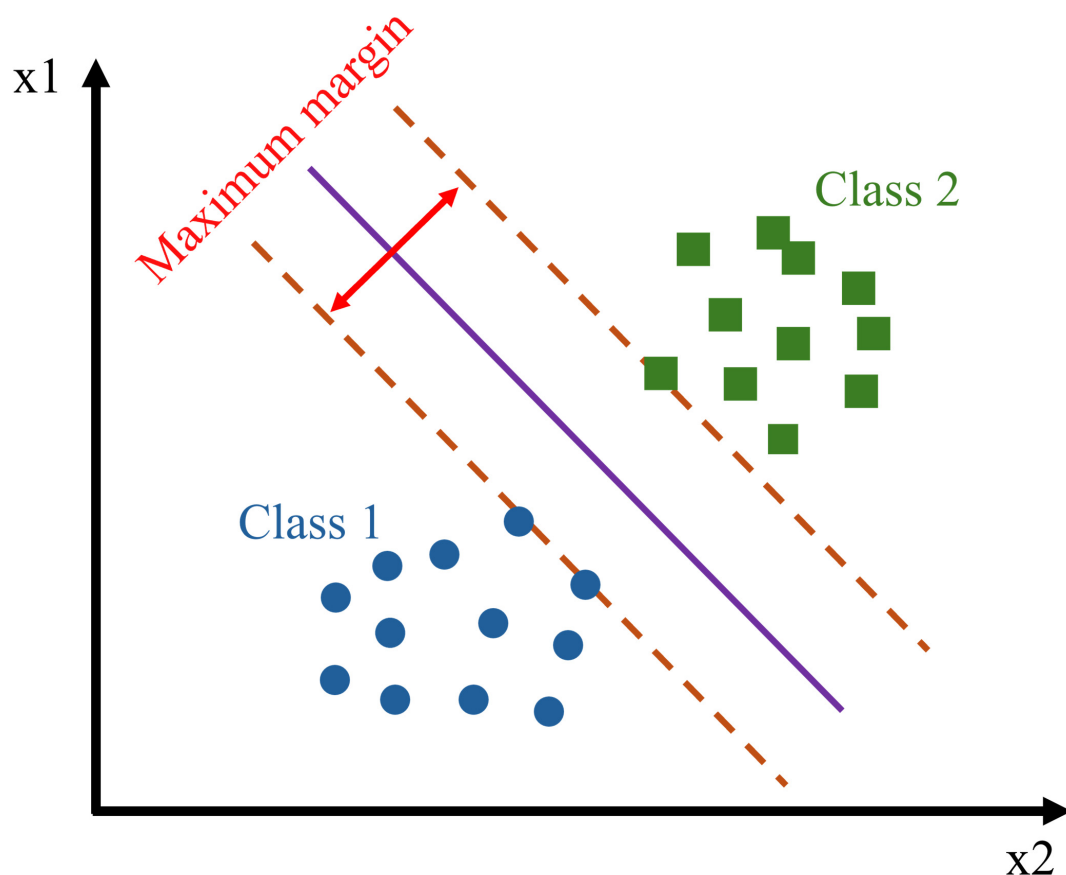


Figure 4. A simple scheme of SVM algorithm.

4.4.2. K-Nearest-Neighbor (KNN)

KNN is a statistical technique used for classification and regression. K refers to the closest training samples in the data space [30]. An unlabeled test sample is classified among

its KNNs in the training set. Various methods are used to calculate neighborhood distance or assign weight to different neighbors.

4.4.3. Decision Tree (DT)

A DT is a hierarchical model that plays a crucial role in decision-making processes [31]. It considers chance events, resource costs, and utility, and is presented as a tree structure featuring nodes that represent decisions or conditions. The algorithm uses specific criteria such as entropy or Gini Impurity to categorize data into different categories. Due to their high interpretability and readability, DTs are essential in data mining, artificial intelligence decision-making, and ML.

4.4.4. Bagging Ensemble Learning

Bagging is an effective EL technique that aims to minimize learning errors by utilizing a group of homogenous ML models [32,33]. The objective of this technique is to reduce variance, which subsequently results in lower classification or regression errors [34,35]. The process involves selecting the number and type of basic models, followed by choosing training data for each model using the bootstrap approach, which entails random sampling with replacement. The training data may be selected multiple times, and it is divided into several subsets based on the bootstrap approach. The basic models are then trained separately on each of these subsets, resulting in different knowledge and opinions for a specific input. Although the basic models are of the same type, each model is trained with a different set, leading to variations in knowledge of each model. During the testing process, the trained models provide an output estimation for new data based on their knowledge. Finally, all models are combined to estimate the new data output. In a classification problem, the models use a simple voting mechanism to determine the new data's class, with the class having the most votes considered the winner. In contrast, in a regression problem, the models use a simple averaging of their opinions instead.

During preliminary experiments, we assessed several feature-selection and dimensionality-reduction strategies; however, none yielded consistent improvements over the complete 61-feature set, so their details are omitted here for brevity and all features were retained. In the final workflow, no dimensionality-reduction or algorithmic feature-selection step was applied; keeping all 61 handcrafted features preserves interpretability and full physiological coverage. Potential overfitting was addressed via model-specific hyperparameter constraints (e.g., SVM margin/kernel settings, KNN neighborhood size, and tree depth/min-samples limits in DT and the DT-based ensemble) and by reporting performance on a held-out test set averaged across repeated splits.

5. Results and Discussions

This study distinguishes alert versus drowsy states using EEG from the MIT-BIH polysomnography dataset. Signals were segmented into 30 s epochs and labeled as Wake (alert) or N1–N3 (drowsy) as outlined in Section 3; REM was excluded. In total, a balanced set of 6212 segments was analyzed. From each segment, 61 features were extracted (time-domain linear/nonlinear plus a small set of frequency descriptors).

Data were partitioned with stratified sampling (70% train/30% test) to preserve class balance. To reduce variability from a single split, the procedure was repeated across five independent runs with different seeds, and average performance is reported. The split was performed at the signal level (i.e., subject-dependent), so segments from the same individual could appear in both train and test sets. We trained KNN, SVM, and DT models with hyperparameters tuned via Bayesian optimization. KNN achieved the highest training accuracy (99%) but generalized less effectively (80.4% test). SVM showed more

balanced train–test behavior. A DT-based bagging ensemble (EL) yielded the best overall test performance—accuracy 84.7% and F1 84.9%—surpassing single classifiers on accuracy, sensitivity, and F1.

The ensemble’s perfect training accuracy is not a practical advantage; it indicates overfitting driven by dataset size and feature dimensionality. Accordingly, test accuracy (84.7%) is the appropriate indicator of performance, and validation on independent datasets is required to assess generalizability. The ensemble’s strength stems from bagging, which reduces variance and enhances model diversity—useful under noisy, high-dimensional EEG conditions—yet the notable train–test gap ($\approx 100\%$ vs. 80.4–84.7%) confirms some overfitting.

To mitigate overfitting, we (i) compared multiple classifiers (KNN/SVM/DT and the DT-based ensemble), (ii) conducted Bayesian hyperparameter tuning, and (iii) examined normalization configurations and alternative random splits. None surpassed the ensemble on the held-out test sets. The observed gap is attributable to (a) limited inter-class diversity, (b) 61-dimensional feature space, and (c) noise/ambiguity during transitions between alert and drowsy states.

Recommendations: Future work should (1) expand to larger, more diverse EEG cohorts, (2) introduce regularization within the ensemble framework, and (3) consider dimensionality management strategies where appropriate. We also assessed data standardization as a potential remedy, but it did not improve performance. Consequently, the DT-based bagging ensemble was adopted, delivering 100% training accuracy and 84.7% test accuracy. On test data, the ensemble achieved higher Precision, Sensitivity, F1, and MCC, while KNN obtained the highest Specificity. Summary metrics are given in Table 5, with confusion matrices in Table 6.

Table 5. Evaluation of all algorithms.

Classifier	Process	Metrics (%)					
		Accuracy	Precision	Sensitivity	Specificity	F1-Score	MCC
KNN	Train	99	100	98.1	100	99	98
	Test	80.4	83.2	76.3	84.5	79.6	61
SVM	Train	80	84.2	73.7	86.2	78.6	60
	Test	78.8	81.7	74.4	83.2	77.9	58
DT	Train	79.8	80.8	78.1	81.5	79.4	60
	Test	78.3	79.5	76.4	80.2	77.9	57
EL	Train	100	100	100	100	100	100
	Test	84.7	84.4	85.4	84.1	84.9	69.42

■ The highest score of train data; ■ The highest score of test data.

We acknowledge that a subject-dependent split can yield optimistic estimates because individual-specific EEG patterns may appear in both training and test sets. While this design is common in early-stage feasibility studies, subject-independent validation will be required for practical clinical/automotive deployment to ensure generalizability.

Development-time trials. During preliminary experiments we applied 10-fold cross-validation, class rebalancing (under/oversampling), and algorithmic feature selection/dimensionality reduction. None produced consistent or meaningful gains across models, so—for focus and brevity—their detailed results are omitted. Final results are therefore reported under a stratified 70/30 hold-out protocol repeated five times, with the full 61-feature set retained to preserve interpretability.

Table 6. Confusion matrix of all algorithms on test data.

Features No.	Predicted Class (%)			
	State		Alertness (%)	Drowsiness (%)
KNN	A	A	83.2	22.1
	D	D	16.8	77.9
SVM	A	A	81.7	23.7
	D	D	18.3	76.3
DT	A	A	79.5	22.9
	D	D	20.5	77.1
EL	A	A	84.4	14.9
	D	D	15.6	85.1

A: Alertness; D: Drowsiness.

The pronounced train–test gap, most evident for KNN and DT, indicates potential overfitting that was not substantially reduced by the explored configurations (feature selection, k-fold schemes, or hyperparameter tuning). Among all methods, the ensemble learning (EL) approach showed greater stability, principally due to bootstrap aggregation reducing variance—highlighting the effectiveness of ensembles for noisy, high-dimensional EEG. Future work should investigate stronger regularization, simplified model architectures, and individualized learning to enhance generalization. A limited re-referencing check (C4 to linked mastoids and to a common average) showed no material difference relative to C4–A1; accordingly, we report results for C4–A1.

Ecological Validity and Domain Shift to Driving

Our analysis focuses on EEG-based recognition of wakefulness–drowsiness transitions using clinically annotated labels. The task-agnostic neural signatures employed—elevated theta, reduced beta, transient alpha bursts, and lower entropy/complexity—have been reported across resting, vigilance, and simulator paradigms, supporting the use of clinical EEG to establish an interpretable baseline for drowsiness detection.

Nonetheless, clinical sleep studies and simulator/on-road driving differ in artifact profiles (EOG/EMG/motion), sensory load, and vigilance dynamics. To enable translation, we recommend: (i) channel-matched acquisition with robust artifact mitigation; (ii) subject-independent evaluation on driving-drowsiness datasets (simulator or on-road); and (iii) lightweight domain adaptation (e.g., feature re-centering or covariance alignment) without modifying the core classifier. Such external validation is a necessary next step.

6. Conclusions and Future Work

This study leveraged the MIT-BIH Polysomnography dataset (18 records; >80 h of clinically annotated overnight data) to distinguish alertness vs. drowsiness from single-channel EEG (C4–A1). We extracted 61 handcrafted features spanning time-domain, nonlinear, and frequency descriptors, and evaluated KNN, DT, SVM, and a DT-based ensemble (EL). Consistent with the nonstationary and nonlinear nature of EEG during sleep, the feature set captured relevant dynamics and enabled reliable discrimination: the EL achieved 84.7% test accuracy with $F1 = 84.9\%$, outperforming single classifiers.

The findings confirm that a lightweight, interpretable pipeline can detect vigilance states from clinical EEG with low computational overhead, making it suitable for embedded/edge inference. At the same time, the train–test gap (perfect training accuracy

vs. lower test accuracy) indicates overfitting, emphasizing the need for larger, more heterogeneous cohorts to establish generalizability.

Using a clinical PSG database is appropriate for a methodological baseline because (i) drowsiness-related EEG markers—elevated theta/alpha and reduced beta—are physiologically consistent across lab and driving settings, (ii) high-quality annotations provide precise labels, and (iii) such datasets are well-established in prior work (e.g., Chen et al. [20]; Christensen et al. [22]). Our aim here was preclinical method development, not direct in-vehicle deployment.

Implications and future work: For real-world translation, priorities include: (1) subject-independent evaluation on simulator/on-road driving datasets; (2) multimodal integration with EOG, EMG, HRV, and respiration to boost robustness; (3) development of artifact-resistant, minimally intrusive sensors (e.g., dry-electrode headbands/ear-EEG) to address motion and comfort; (4) regularization and simplified model architectures, plus personalization to individual baselines; and (5) lightweight domain adaptation (e.g., feature re-centering/covariance alignment) without altering the core classifier. The proposed EEG module should be viewed as complementary to existing camera-based fatigue monitoring to increase resilience to environmental variability.

Recent deep learning systems (e.g., HATNet, IEEE TCYB 2025; ~90–92% test accuracy) illustrate the upper bound in accuracy but often require substantial compute and GPU-supported inference. In contrast, our ensemble of decision trees with engineered features attained 84.7% while providing full interpretability and fast, resource-efficient inference, a practical advantage for embedded automotive platforms.

Finally, single-channel modeling was intentional to support deployment and interpretability. Although multi-channel montages may offer incremental gains (e.g., spatial filtering with CAR/Laplacian), our sensitivity checks suggest the present conclusions do not hinge on a particular reference; systematic multi-channel validation is left for future work.

Author Contributions: M.S.: Validation, Methodology, Formal analysis, Software, Conceptualization. S.R.: Software, Validation, Methodology, Supervision. S.P.: Writing—original draft, Writing—review and editing, Methodology, Supervision, Project Administration. A.S.: Writing—review and editing, Validation. H.M.D.K.: Writing—review and editing, Validation. T.H.: Writing—review and editing, Validation, Supervision. S.G.: Writing—review and editing, Supervision. H.A.: Writing—review and editing, Validation, Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset utilized in this study is openly accessible through the Kaggle platform at the following link: “<https://www.physionet.org/content/slpdb/1.0.0/>” (accessed on 20 February 2025).”

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Touahmia, M. Identification of risk factors influencing road traffic accidents. *Eng. Technol. Appl. Sci. Res.* **2018**, *8*, 2417–2421. [CrossRef]
2. Lee, H.; Lee, J.; Shin, M. Using wearable ECG/PPG sensors for driver drowsiness detection based on distinguishable pattern of recurrence plots. *Electronics* **2019**, *8*, 192. [CrossRef]
3. Cui, J.; Lan, Z.; Sourina, O.; Müller-Wittig, W. EEG-based cross-subject driver drowsiness recognition with an interpretable convolutional neural network. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *34*, 7921–7933. [CrossRef]
4. Banks, S.; Catcheside, P.; Lack, L.C.; Grunstein, R.R.; McEvoy, R.D. The Maintenance of Wakefulness Test and driving simulator performance. *Sleep* **2005**, *28*, 1381–1385. [CrossRef]

5. Morelli, D.; Rossi, A.; Cairo, M.; Clifton, D.A. Analysis of the impact of interpolation methods of missing RR-intervals caused by motion artifacts on HRV features estimations. *Sensors* **2019**, *19*, 3163. [\[CrossRef\]](#)
6. Garber, Z. Night Encounters: Theologizing Dialogue. *Shofar Interdiscip. J. Jew. Stud.* **1996**, *15*, 38–53. [\[CrossRef\]](#)
7. Kiymik, M.K.; Akin, M.; Subasi, A. Automatic recognition of alertness level by using wavelet transform and artificial neural network. *J. Neurosci. Methods* **2004**, *139*, 231–240. [\[CrossRef\]](#)
8. De Gennaro, L.; Ferrara, M.; Bertini, M. The boundary between wakefulness and sleep: Quantitative electroencephalographic changes during the sleep onset period. *Neuroscience* **2001**, *107*, 1–11. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Delimayanti, M.K.; Purnama, B.; Nguyen, N.G.; Faisal, M.R.; Mahmudah, K.R.; Indriani, F.; Kubo, M.; Satou, K. Classification of brainwaves for sleep stages by high-dimensional FFT features from EEG signals. *Appl. Sci.* **2020**, *10*, 1797. [\[CrossRef\]](#)
10. Kim, J. A comparative study on classification methods of sleep stages by using EEG. *J. Korea Multimed. Soc.* **2014**, *17*, 113–123. [\[CrossRef\]](#)
11. Motin, M.A.; Karmakar, C.; Palaniswami, M.; Penzel, T. Photoplethysmographic-based automated sleep–wake classification using a support vector machine. *Physiol. Meas.* **2020**, *41*, 075013. [\[CrossRef\]](#)
12. Abdollahi, M.; Jafarizadeh, A.; Asbagh, A.G.; Sobhi, N.; Pourmoghtader, K.; Pedrammehr, S.; Asadi, H.; Alizadehsani, R.; Tan, R.S.; Acharya, U.R. Artificial Intelligence in Assessing Cardiovascular Diseases and Risk Factors via Retinal Fundus Images: A Review of the Last Decade. *arXiv* **2023**, arXiv:2311.07609. [\[CrossRef\]](#)
13. van Meulen, F.B.; Grassi, A.; Van den Heuvel, L.; Overeem, S.; van Gilst, M.M.; van Dijk, J.P.; Maass, H.; van Gastel, M.J.H.; Fonseca, P. Contactless camera-based sleep staging: The healthbed study. *Bioengineering* **2023**, *10*, 109. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Sun, C.; Chen, C.; Fan, J.; Li, W.; Zhang, Y.; Chen, W. A hierarchical sequential neural network with feature fusion for sleep staging based on EOG and RR signals. *J. Neural Eng.* **2019**, *16*, 066020. [\[CrossRef\]](#)
15. Jafari, Z.; Rajebi, S.; Haghipour, S. Using the Neural Network to Diagnose the Severity of Heart Disease in Patients Using General Specifications and ECG Signals Received from the Patients. *Adv. Sci. Technol. Eng. Syst. J.* **2020**, *5*, 882–892. [\[CrossRef\]](#)
16. Jafari, Z.; Yousefi, A.M.; Rajabi, S. Using different types of neural networks in detection the body's readiness for blood donation and determining the value of each of its parameters using genetic algorithm. *Innovaciencia* **2020**, *8*, 1–10. [\[CrossRef\]](#)
17. Yousefi, V.; Kheiri, S.; Rajebi, S. Evaluation of K-nearest neighbor, bayesian, perceptron, RBF and SVM neural networks in diagnosis of dermatology disease. *Int. J. Tech. Phys. Probl. Eng.* **2020**, *42*, 114–120.
18. Cen, L.; Yu, Z.L.; Tang, Y.; Shi, W.; Kluge, T. Deep learning method for sleep stage classification. In Proceedings of the Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, 14–18 November 2017; Proceedings, Part II. Springer: Cham, Switzerland, 2017; pp. 796–802.
19. Cho, T.; Sunarya, U.; Yeo, M.; Hwang, B.; Koo, Y.S. Deep-ACTINet: End-to-end deep learning architecture for automatic sleep–wake detection using wrist actigraphy. *Electronics* **2019**, *8*, 1461. [\[CrossRef\]](#)
20. Chen, X.; Wang, Y.; Gao, S.; Jung, T.P.; Gao, X. Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface. *J. Neural Eng.* **2015**, *12*, 046008. [\[CrossRef\]](#)
21. Sistaninezhad, M.; Jafarizadeh, A.; Rajebi, S.; Pedrammehr, S.; Alizadehsani, R.; Gorriz, J.M. Morning Anxiety Detection Through Smartphone-Based Photoplethysmography Signals Analysis Using Machine Learning Methods. In Proceedings of the Artificial Intelligence for Neuroscience and Emotional Systems, IWINAC 2024, Olhão, Portugal, 4–7 June 2024; Lecture Notes in Computer Science. Springer: Berlin/Heidelberg, Germany, 2024.
22. Christensen, J.A.E.; Wassing, R.; Wei, Y.; Ramautar, J.R.; Lakbila-Kamal, O.; Jennum, P.J.; Van Someren, E.J. Data-driven analysis of EEG reveals concomitant superficial sleep during deep sleep in insomnia disorder. *Front. Neurosci.* **2019**, *13*, 445494. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Cajochen, C.; Munch, M.; Kobialka, S.; Krauchi, K.; Steiner, R.; Oelhafen, P.; Orgul, S.; Wirz-Justice, A. High sensitivity of human melatonin, alertness, thermoregulation, and heart rate to short wavelength light. *J. Clin. Endocrinol. Metab.* **2005**, *90*, 1311–1316. [\[CrossRef\]](#)
24. Schölkopf, B.; Smola, A.J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, USA, 2002.
25. Stojanovic, R.; Karadaglic, D. A LED–LED-based photoplethysmography sensor. *Physiol. Meas.* **2007**, *28*, 19. [\[CrossRef\]](#)
26. Ichimaru, Y.; Moody, G. Development of the polysomnographic database on CD-ROM. *Psychiatry Clin. Neurosci.* **1999**, *53*, 175–177. [\[CrossRef\]](#)
27. Jafari, B.; Mohsenin, V. Polysomnography. *Clin. Chest Med.* **2010**, *31*, 287–297. [\[CrossRef\]](#)
28. Roy, R.N.; Charbonnier, S.; Bonnet, S. Detection of mental fatigue using an EEG-based brain–computer interface. *Int. J. Hum.-Comput. Interact.* **2016**, *32*, 1023–1034.
29. Nasehi, S.; Pourghassem, H. A novel fuzzy decision-making based feature selection algorithm for classification of human alertness levels using EEG signals. *Eng. Appl. Artif. Intell.* **2013**, *26*, 1608–1616.
30. Ertuğrul, Ö.F.; Tağluk, M.E. A novel version of k nearest neighbor: Dependent nearest neighbor. *Appl. Soft Comput.* **2017**, *55*, 480–490. [\[CrossRef\]](#)

31. Ashari, A.; Paryudi, I.; Tjoa, A.M. Performance comparison between Naïve Bayes, decision tree, and k-nearest neighbor in searching alternative design in an energy simulation tool. *Int. J. Adv. Comput. Sci. Appl.* **2013**, *4*, 33–39. [\[CrossRef\]](#)
32. Pham, B.T.; Tien Bui, D.; Prakash, I. Landslide susceptibility assessment using bagging ensemble based alternating decision trees, logistic regression, and J48 decision trees methods: A comparative study. *Geotech. Geol. Eng.* **2017**, *35*, 2597–2611. [\[CrossRef\]](#)
33. Dong, X.; Yu, Z.; Cao, W.; Shi, Y.; Ma, Q. A survey on ensemble learning. *Front. Comput. Sci.* **2020**, *14*, 241–258. [\[CrossRef\]](#)
34. Sun, Q.; Pfahringer, B. Bagging ensemble selection. In Proceedings of the AI 2011: Advances in Artificial Intelligence: 24th Australasian Joint Conference, Perth, Australia, 5–8 December 2011; Proceedings, AI'11. Springer: Berlin/Heidelberg, Germany, 2011; pp. 251–260.
35. Bühlmann, P. Bagging, boosting and ensemble methods. In *Handbook of Computational Statistics: Concepts and Methods*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 985–1022.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.