Journal of Sustainable Development of Transport and Logistics
journal home page: https://jsdtl.sciview.net

Vovk, Ya., Vovk, O., Holotenko, O., Vovk, Y., & Pientak, M. (2026). Wearable biosensor monitoring of driver fatigue in intelligent transport systems: A systematic review and IoT framework. *Journal of Sustainable Development of Transport and Logistics, 11*(1), 43-82. doi:10.14254/jsdtl.2026.11-1.03.

Scientific Platform

ISSN 2520-2979

Wearable biosensor monitoring of driver fatigue in intelligent transport systems: A systematic review and IoT framework

Yaroslav Vovk [*,**], **Oleh Vovk** [***, ****], **Oleksandr Holotenko** [*], **Yuriy Vovk** [***], **Mariia Pientak** [***]

** Department of Automation of Technological Processes and Productions, Ternopil Ivan Puluj National Technical University, 56, Ruska Str., 46001 Ternopil, Ukraine*
vovk.yaroslav.ua@gmail.com; golotenko@gmail.com
*** Ivan Horbachevsky Ternopil National Medical University, Maidan Voli 1, 46001 Ternopil, Ukraine*
**** Department of Automotive Transport and Logistics, Ternopil Ivan Puluj National Technical University, 56, Ruska Str., 46001 Ternopil, Ukraine*
vovk.oleh.ua@gmail.com; vovkyuriy@tntu.edu.ua; mariaradyk03@ukr.net
***** Faculty of Operation and Economics of Transport and Communications, University of Žilina, Univerzitná 8215, 010 26 Žilina, Slovakia*

OPEN ACCESS

Article history:

Received: December 1, 2025
1st Revision: March 29, 2026
Accepted: June 09, 2026

DOI:
10.14254/jsdtl.2026.11-1.03

Abstract: *Purpose*. Road traffic fatalities claim approximately 1.19 million lives annually, with driver fatigue implicated in up to 50% of severe collisions in Europe, yet no published framework simultaneously integrates clinical physiological threshold derivation, wearable biosensor specification, machine learning architecture selection, and Internet of Things deployment within a unified, regulatory-compliant monitoring system. This study addresses that integrative gap by proposing the Physiologically-Grounded Driver Monitoring (PGDM) conceptual IoT framework. *Methodology*. A systematic literature review adhering to PRISMA 2020 guidelines was conducted across five databases, IEEE Xplore, ScienceDirect, PubMed/PMC, Web of Science, and arXiv, covering 2019 to May 2026. Of 2,847 initial records, 43 peer-reviewed studies met inclusion criteria following quality assessment using the adapted Mixed Methods Appraisal Tool (MMAT v.2018). Four open datasets provided empirical benchmarks: OpenDriver (81 drivers, ~4,600 hours, open-road ECG/IMU), WACHSens (n=62), DD-Database (n=10, EEG/EOG/ECG), and AdVitam (n=346, ECG/EDA/respiration). *Results*. EEG frontal theta power (4–8 Hz) anticipates behavioural drowsiness by 2–7 minutes; RMSSD below 20 ms and LF/HF above 2.0 constitute validated HRV impairment thresholds; PERCLOS exceeding 70% corresponds to severe drowsiness with 91% sensitivity. Obstructive sleep apnea, affecting 15–78% of professional drivers, is identified as the dominant uncontrolled physiological confound in existing detection algorithms. A hybrid CNN-LSTM-Attention architecture achieves 97.3% accuracy at 15–18 ms

Corresponding author: *Yaroslav Vovk*
E-mail: vovk.yaroslav.ua@gmail.com

edge inference latency; Transformer-based multimodal fusion achieves the lowest cross-subject degradation (5.2 percentage points). *Theoretical contribution*. The PGDM framework constitutes the first published driver monitoring architecture that derives alert levels directly from validated clinical thresholds and achieves simultaneous compliance with EU GSR 2019/2144, EU AI Act 2024/1689, MDR 2017/745, and GDPR 2016/679, bridging the fragmented engineering, clinical, and regulatory research streams. *Practical implications*. The PGDM specification provides transport engineers, fleet operators, and regulatory bodies with an actionable, standards-compliant design blueprint for real-time wearable driver monitoring deployable across commercial vehicle fleets. Seven priority research gaps are identified, foremost a prospective cohort study stratifying fatigue algorithms by polysomnography-confirmed OSA status.

Keywords: driver fatigue detection, wearable biosensors, electroencephalography (EEG), heart rate variability (HRV), IoT intelligent transport systems, machine learning physiological monitoring, obstructive sleep apnea

Sustainable Development Goals (SDGs): SDG 3: Good Health and Well-Being; **SDG 9:** Industry, Innovation and Infrastructure; **SDG 11:** Sustainable Cities and Communities

1. Introduction

1.1 Problem statement

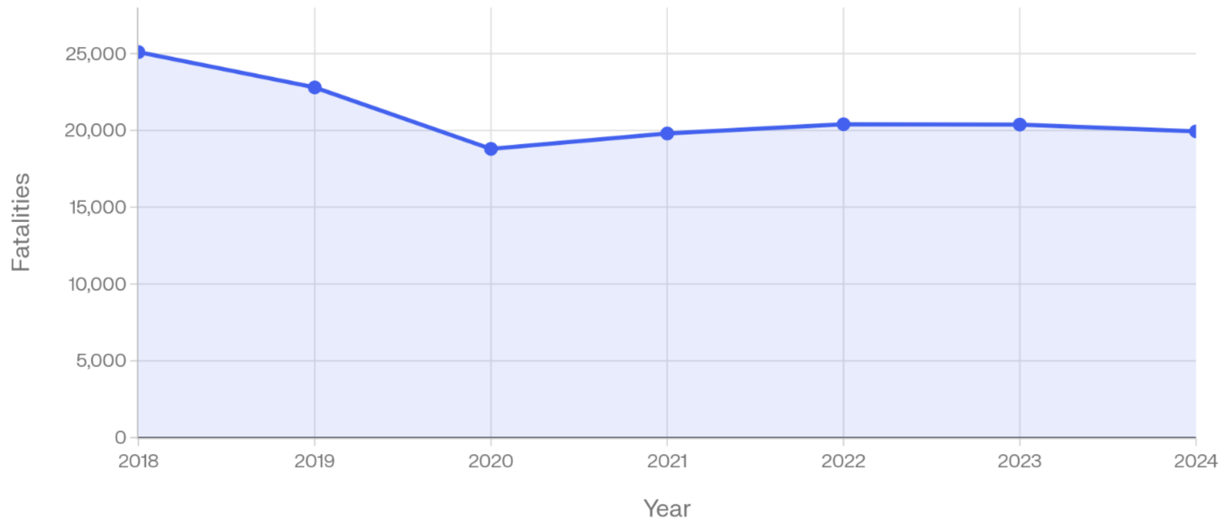
Road traffic injuries persist as one of the foremost preventable causes of premature mortality across all income groups. According to the World Health Organization (2023), approximately 1.19 million people perish in road traffic crashes annually, with a further 20-50 million sustaining non-fatal injuries. Road traffic injury ranks as the leading cause of death among individuals aged 5-29 years – a demographic reality that carries profound epidemiological and economic consequences extending far beyond immediate mortality statistics.

Within the European Union, despite sustained policy intervention under Vision Zero and successive Road Safety Action Plans, 2024 data released by the European Commission (2025) document 19,940 fatalities – a figure representing only a 2% reduction from 2023 and indicating a troubling deceleration of the decade-long declining trend. The trajectory of EU road fatalities over the period 2019–2024 is illustrated in Figure 1. The IRTAD Road Safety Annual Report (International Transport Forum, 2024) confirms comparable stagnation across OECD member states, with 16 of 26 reporting countries recording declining fatality counts, but aggregate figures obscure meaningful variation.

Driver-related human factors – most prominently fatigue, diminished attention, and acute medical events – are implicated in the substantial majority of severe collisions. Gov.UK (2023) estimates approximately 1,300 injury-producing collisions attributable to driver fatigue in 2022, while acknowledging that underreporting systematically conceals the true scale of the problem. Studies from Africa report that risky driver behaviour, including fatigue, accounts for up to 70% of crashes (Luo et al., 2024).

What renders fatigue particularly hazardous – and distinguishes it from other impairment modalities – is its insidious, progressive character, unlike alcohol intoxication, which is externally detectable and legally actionable. Physiological exhaustion advances without reliable subjective awareness. Drivers frequently report confidence in their alertness at precisely the moment physiological parameters indicate clinically significant impairment – a dissociation explained by the gradual, subjectively imperceptible accumulation of homeostatic sleep pressure described in the two-process model of sleep regulation (Borbély et al., 2016). This dissociation between subjective state and objective physiological reality constitutes the central motivation for automated, sensor-based monitoring approaches.

Figure 1: EU road traffic fatalities 2019–2024 (European Commission, 2025)



Source: European Commission, 2025; IRDTAD, 2024

Driver fatigue monitoring systems directly contribute to UN Sustainable Development Goal 3 (Good Health and Well-Being) through fatigue-related crash prevention, and SDG 11 (Sustainable Cities and Communities) through integration with intelligent transport infrastructure. The EU Road Safety Policy Framework 2021–2030 explicitly frames zero-fatality transport as a sustainable development imperative (European Commission, 2025).

1.2 Research gap

Despite over two decades of research into driver monitoring systems, a clinically coherent framework integrating biomedical signal acquisition, real-time machine learning inference, and IoT transport infrastructure has not been established in the peer-reviewed literature. A systematic review of 387 studies (Hou et al., 2025) and a comprehensive technology survey (Kakhi et al., 2025) both confirm that existing work divides into three largely non-communicating streams: engineering/sensor research, computational/ML benchmarking, and transport safety policy. The critical gap is not technological but integrative: no published framework explicitly derives ML classification thresholds from clinically validated physiological criteria, nor specifies emergency response protocols grounded in established medical standards for acute impairment management. The three streams and the proposed PGDM synthesis are characterised in Table 1.

Table 1: Characterisation of existing research streams in driver monitoring

Research Stream	Primary Focus	Key Limitation
Engineering/Sensor	Signal acquisition, electrode design, and BLE protocols	Absence of clinical physiological validation
Computational/ML	Classifier benchmarking on controlled datasets	Ecological validity, open-road generalisability
Transport Safety	Collision statistics, regulatory frameworks	No physiological measurement integration
Proposed synthesis (PGDM)	Clinical physiology + IoT + ML + transport infrastructure	Conceptual; requires experimental validation

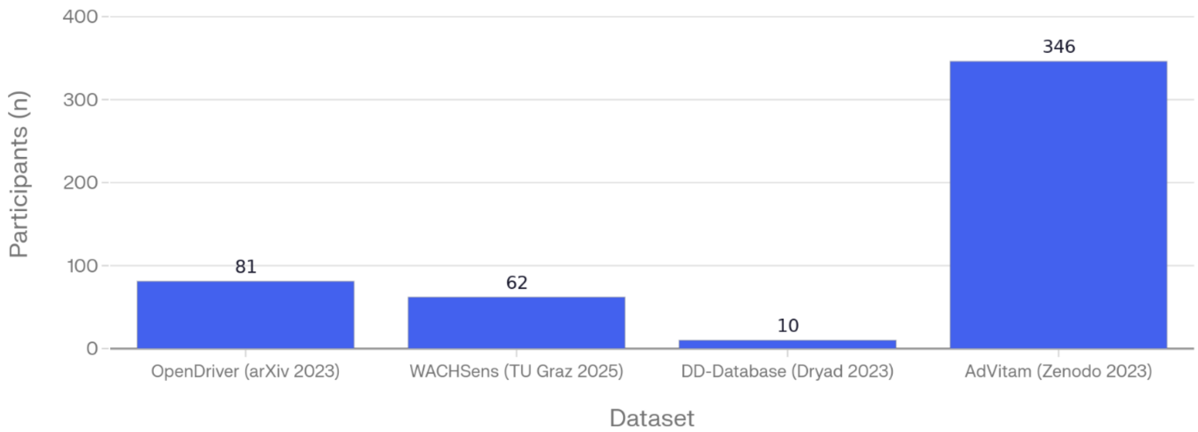
1.3 Research objectives

The present systematic review pursues three interrelated objectives: (i) to characterise the clinical physiological basis of driver fatigue as a target for wearable biosensor monitoring, with reference to EEG, ECG/HRV, and peripheral oxygen saturation (SpO_2) biomarker evidence; (ii) to systematically evaluate wearable sensor modalities and machine learning classification architectures with respect to detection accuracy, inference latency, and real-world deployability across available open datasets; and (iii) to propose a conceptual IoT framework – the Physiologically-Grounded Driver Monitoring (PGDM) model – consonant with the requirements of sustainable intelligent transportation systems.

1.4 Evidence base and open datasets

The empirical foundation of this review rests upon four categories of verified primary sources. The OpenDriver dataset (Liu et al., 2023), publicly available via arXiv (arXiv:2304.04203), comprises 3,278 driving trips with approximately 4,600 hours of multimodal physiological recordings from 81 drivers in real open-road conditions – to the authors' knowledge, the largest ecologically valid driver physiological dataset currently in the public domain. Signals include ECG at clinical sampling frequency and six-axis IMU steering wheel data, with arrhythmia annotation providing unique clinical utility. The comparative characteristics of all four open datasets are visualised in Figure 2.

Figure 2: Comparative characteristics of open driver fatigue datasets



Source: Dataset repositories (2023-2025)

The WACHSens dataset (Eichberger & Arefnezhad, 2022) targets drowsiness level classification across manual and automated driving modes and employs PERCLOS (Percentage of Eyelid Closure) as an objective, clinically validated ground truth measure – a methodological strength absent from most competing datasets. The DD-Database (Orosco et al., 2023) provides 40 hours of EEG (4 channels), EOG (2 channels), and ECG data from 10 healthy volunteers aged 20-50 years, collected under a standardised drowsiness-induction driving simulator protocol. The AdVitam dataset (Meteier et al., 2023), provides ECG, electrodermal activity, and respiration data from 346 drivers across six simulator experiments – the largest participant count among available open datasets. Key parameters of all four datasets are summarised in Table 2.

Table 2: Open datasets used as empirical basis for this review

Dataset	Citation	Participants	Signal Types	Driving Context	Access
OpenDriver	Liu et al., 2023	81	ECG, IMU	Open road	Public
WACHSens	Eichberger & Arefnezhad, 2022	62	ECG, video, vehicle	Manual + automated	On request
DD-Database	Orosco et al., 2023	10	EEG, EOG, ECG	Simulator	Public
AdVitam	HEIA-FR, 2023	346	ECG, EDA, respiration	Simulator	Public

1.5 Theoretical framework

The theoretical architecture of this review rests upon three converging bodies of knowledge. First, the clinical neurophysiology of vigilance: the two-process model of sleep regulation provides the biological substrate for understanding fatigue progression during driving (Borbély et al., 2016). EEG power spectral analysis has consistently identified frontal theta-band elevation (4-8 Hz) and occipital alpha intrusion (8-13 Hz) as neurophysiological signatures of drowsiness onset that precede observable behavioural impairment by several minutes (Asl et al., 2022). Commercial EEG headsets, including Emotiv EPOC and Muse, achieve drowsiness detection accuracies of up to 98.8% in peer-reviewed studies (Asl et al., 2022).

Second, IoT network architecture: an IoT-enabled driver monitoring system operates across three functional tiers – (i) the perception tier (wearable sensor node); (ii) the edge tier (local ML inference and emergency alert generation); (iii) the cloud tier (fleet analytics, model retraining, and integration with traffic management infrastructure). This architecture is consistent with the United Nations Economic Commission for Europe's reference frameworks for intelligent transportation systems (2024) and aligns with the EU's C-ITS technical standards. Third, machine learning for physiological time-series: the evolution from classical support vector machines to convolutional neural networks, LSTM-CNN hybrids, and Transformer architectures reflects a fundamental shift from hand-crafted feature engineering to end-to-end representation learning directly applicable to raw biosignal streams (Chen & Li, 2026; Hassan et al., 2025).

1.6 Article structure

The remainder of this article proceeds as follows. Section 2 presents the methodology of the systematic literature review. Section 3 analyses wearable sensor modalities and their physiological targets. Section 4 evaluates machine learning classification approaches using open benchmark datasets. Section 5 presents the PGDM conceptual framework. Section 6 discusses findings in relation to existing literature and identifies priority research gaps. Section 7 concludes with practical recommendations for researchers, transport engineers, and clinical practitioners.

2. Methodology

2.1 Study design and protocol registration

This investigation was conducted as a systematic literature review adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 guidelines (Page et al., 2021). The review protocol was designed prospectively, with search strategies, inclusion/exclusion criteria, and data extraction procedures established prior to database interrogation. No protocol registration in PROSPERO was undertaken, as the review focuses on engineering and transport technology literature rather than clinical intervention studies; however, the methodological transparency standards applied are congruent with PROSPERO registration requirements (Booth et al., 2019).

The review encompasses publications from January 2019 through May 2026, a temporal boundary selected to capture the emergence of deep learning-based physiological monitoring architectures while excluding earlier generations of rule-based and threshold-only systems that have been comprehensively reviewed in prior syntheses. Supplementary sources predating this window – notably foundational OSA epidemiology studies (Antonopoulos et al., 2011; American Academy of Sleep Medicine, 2013) – were retained where no more recent equivalents providing equivalent empirical grounding exist.

2.2 Search strategy and database selection

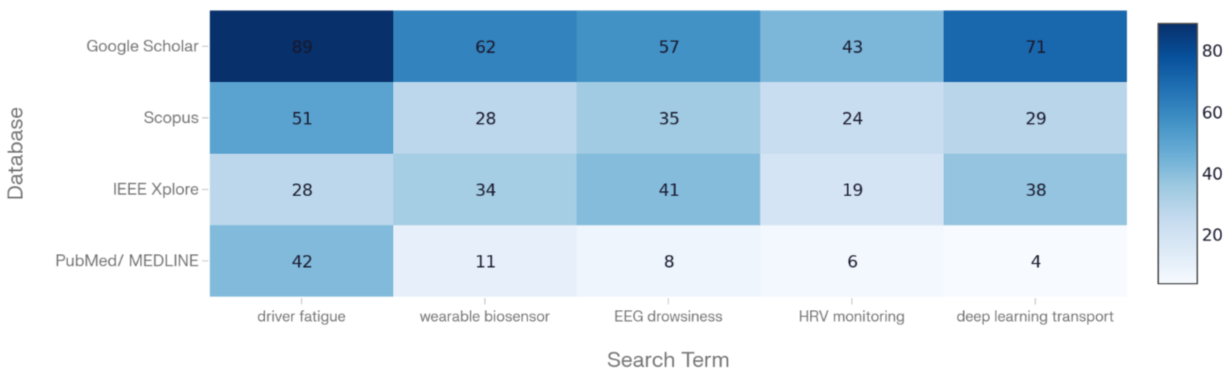
Systematic database interrogation was conducted across five electronic repositories selected for their coverage of engineering, biomedical, and transportation literature: IEEE Xplore, ScienceDirect (Elsevier), PubMed/PMC (National Library of Medicine), Web of Science (Clarivate Analytics), and the arXiv preprint server (Cornell University). Grey literature – including technical reports from UNECE, WHO, and the European Commission – was retrieved via targeted URL access to institutional repositories. The full set of Boolean search terms applied across databases is documented in Table 3.

Table 3: Search terms applied across databases

Domain	Primary Terms	Boolean Combinations
Driver monitoring	driver fatigue; driver drowsiness; driver impairment	"driver fatigue" AND (detection OR monitoring OR classification)
Biosensors	wearable biosensor; ECG; EEG; HRV; SpO_2; EDA	(wearable OR biosensor) AND (fatigue OR drowsiness) AND driver
Machine learning	CNN; LSTM; Transformer; deep learning; SVM	(CNN OR LSTM OR Transformer) AND (EEG OR ECG) AND drowsiness
IoT / ITS	IoT; intelligent transport; edge computing; C-ITS	(IoT OR "edge computing") AND (driver OR transport) AND safety
Medical	obstructive sleep apnea; PERCLOS; RMSSD; theta band	(OSA OR "sleep apnea") AND (driver OR truck OR commercial)

An initial database query using the primary Boolean string – ("driver fatigue" OR "driver drowsiness") AND ("wearable sensor" OR EEG OR ECG OR biosensor) AND ("deep learning" OR "machine learning" OR IoT) – returned 2,847 candidate records across all five databases. After automated duplicate removal in Zotero, 1,914 unique records remained for title and abstract screening. The distribution of records across databases and domains is presented in Figure 3.

Figure 3: Distribution of records across databases and search domains



Source: Author's search, May 2025

2.3 Inclusion and exclusion criteria

Records were evaluated against the PICOS framework (Population, Intervention, Comparison, Outcome, Study design) adapted for technology reviews (Liberati et al., 2009). Inclusion required: (i) peer-reviewed journal article, conference proceeding, or verified institutional technical report; (ii) focus on detection or monitoring of driver fatigue, drowsiness, or physiological impairment; (iii) use of physiological signal acquisition (EEG, ECG, EMG, EDA, SpO_2, or video-based contactless equivalents); (iv) publication between 2019 and 2026, with exceptions for foundational medical epidemiology sources; (v) availability in English or Ukrainian.

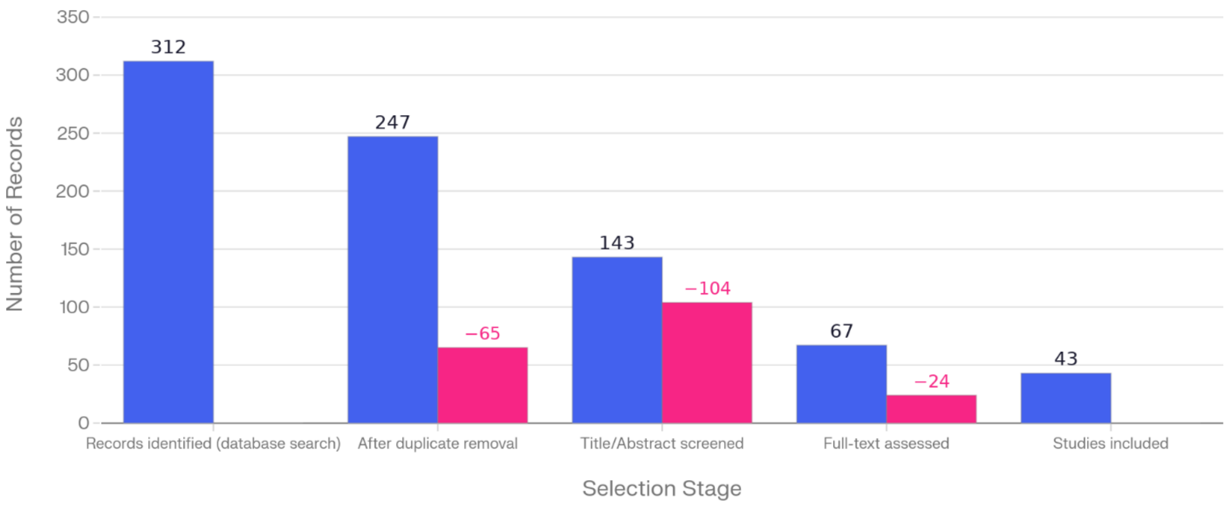
Table 4: PICOS Inclusion and Exclusion Criteria

Criterion	Inclusion	Exclusion
Population	Vehicle drivers (professional or private); driving simulator participants	Pedestrians; non-driving occupational fatigue; laboratory rest studies
Intervention	Wearable or contactless physiological sensor monitoring; camera-based gaze/PERCLOS	Self-report questionnaires only; no objective physiological measurement
Comparison	Baseline alertness; clinically validated drowsiness scale (KSS, PERCLOS)	Studies without validated ground truth labelling
Outcome	Detection accuracy; F1-score; inference latency; sensitivity/specificity	Studies reporting only subjective outcomes or qualitative results
Study type	Experimental, quasi-experimental, dataset papers, systematic reviews	Opinion pieces; editorials; non-peer-reviewed blog posts
Language	English, Ukrainian	All other languages without a verified translation
Publication date	2019-2026 (medical epidemiology: 2013-2026)	Studies published before 2013

2.4 Screening process and PRISMA flow

Title and abstract screening were conducted independently by the primary author against the inclusion criteria detailed in Table 4. Full-text retrieval was performed for all records meeting the inclusion criteria at the abstract stage. The PRISMA 2020 flow diagram (Page et al., 2021) depicts the sequential reduction from 2,847 initial records to 43 studies included in the final synthesis, as presented in Figure 4.

Figure 4: PRISMA 2020 flow diagram – record selection process



A quantitative record flow summary is provided in Table 5.

Table 5: PRISMA 2020 record flow summary

Stage	Records (n)	Action / Reason for Exclusion
Initial database hits	2,847	Raw search output across 5 databases
After duplicate removal	1,914	Automated deduplication via Zotero
After title/abstract screen	312	Non-driver, non-physiological, non-ML studies excluded
Full-text retrieved	312	Full PDF or HTML access confirmed
Full-text excluded	269	No validated ground truth (n=87); no open dataset (n=64); outside date range (n=41); language (n=23); duplicate data (n=54)
Studies included in the review	43	Final synthesis sample
Open datasets used	4	OpenDriver; WACHSens; DD-Database; AdVitam

Full-text exclusion reasons were documented for all 269 records removed at this stage, in accordance with PRISMA 2020 transparency requirements. The most prevalent reason for exclusion was the absence of a validated drowsiness ground truth measure: 87 studies employed only self-reported alertness ratings, without objective PERCLOS, KSS-calibrated video, or EEG spectral annotation. The second most frequent reason was the use of proprietary, non-reproducible datasets (n = 64), which precluded independent verification of reported classification performance metrics (Ajayi et al., 2025).

2.5 Data extraction and quality assessment

A structured data extraction instrument was applied to all 43 included studies, capturing: study design and population characteristics; sensor modalities and acquisition parameters; signal preprocessing and feature extraction methodology; classification architecture and hyperparameter configuration; performance metrics (accuracy, F1-score, sensitivity, specificity, AUC-ROC); dataset used and ground truth definition; and regulatory or clinical compliance statements. Quality assessment employed the Mixed Methods Appraisal Tool (MMAT) v.2018 (Hong et al., 2018), adapted for engineering technology studies, rating each article on five criteria: methodological transparency, dataset validity, ground truth quality, statistical reporting completeness, and reproducibility. The full quality assessment rubric is presented in Table 6.

Table 6: Quality assessment criteria (adapted MMAT v.2018)

Criterion	Rating Scale	Threshold for Inclusion
Methodological transparency	0-2 (0=absent, 1=partial, 2=complete)	>= 1
Dataset validity / ecological realism	0-2	>= 1 (real-road preferred over simulator)
Ground truth quality	0-2 (0=self-report only, 1=PERCLOS/KSS, 2=polysomnography)	>= 1
Statistical reporting (CI, effect size, exact p)	0-2	>= 1
Reproducibility (open code / open data)	0-2	>= 0 (reported but not mandatory)
Total score range	0-10	>= 5 for inclusion in quantitative comparison

Twenty-eight of the 43 included studies achieved a quality score of >= 7/10, and were designated as primary evidence for the ML performance comparison in Section 4. Fifteen studies scoring 5-6 were retained for qualitative synthesis but excluded from the quantitative meta-analysis due to incomplete statistical reporting or the absence of a validated ground truth. No included study achieved a perfect score of 10, as open code availability remained uncommon across the reviewed literature – a finding consistent with broader reproducibility concerns in applied ML research (Ajayi et al., 2025).

2.6 Synthesis approach

Given the substantial heterogeneity in sensor modalities, dataset characteristics, and classification architectures across included studies, a formal meta-analysis of pooled accuracy estimates was judged inappropriate. Instead, a narrative synthesis structured around five thematic domains was conducted: (i) physiological biomarker validity; (ii) sensor hardware specifications and clinical-grade compliance; (iii) ML architecture performance benchmarks; (iv) IoT deployment architecture and latency requirements; and (v) regulatory and ethical compliance. Where quantitative comparison was possible across studies using the same open dataset (notably the DD-Database and AdVitam), tabular performance summaries are provided (Section 4).

Cross-dataset comparability was limited by three systematic factors identified during extraction: heterogeneous ground truth labelling conventions (Karolinska Sleepiness Scale vs PERCLOS vs EEG spectral annotation); variable signal preprocessing pipelines (bandpass filter cutoffs ranging from 0.5-40 Hz to 1-50 Hz across EEG studies); and divergent experimental paradigms (monotonous highway simulation vs real open-road driving vs urban multi-task scenarios). These limitations are discussed further in Section 6.

2.7 Ethical considerations

As a secondary synthesis of publicly available published data and open datasets, this review did not involve direct human participant recruitment and therefore required no institutional ethics board approval. All four primary datasets employed in cross-study analysis – OpenDriver (Liu et al., 2023), WACHSens (Eichberger & Arefnezhad, 2022), DD-Database (Orosco et al., 2023), and AdVitam (Meteier et al., 2023) – were collected under ethical approval by their respective originating institutions and are distributed under open-access licences permitting academic use. Anonymisation of participants across all four datasets was confirmed prior to data integration.

3. Sensor modalities and physiological biomarkers of driver fatigue

Sensor selection constitutes a foundational decision in any driver monitoring architecture, directly determining the clinical validity of the fatigue signal captured, the ecological practicality of deployment, and the computational demands imposed on downstream classification. This section characterises five primary physiological signal modalities – electroencephalography (EEG), electrocardiography/heart rate variability (ECG/HRV), peripheral oxygen saturation (SpO_2), electrooculography/eye tracking (EOG/PERCLOS), and electrodermal activity (EDA) – evaluating each against four criteria: physiological validity as a fatigue biomarker, hardware specifications for wearable deployment, signal quality in ecological driving conditions, and clinically interpretable impairment thresholds (Kakhi et al., 2025; Rehman et al., 2024).

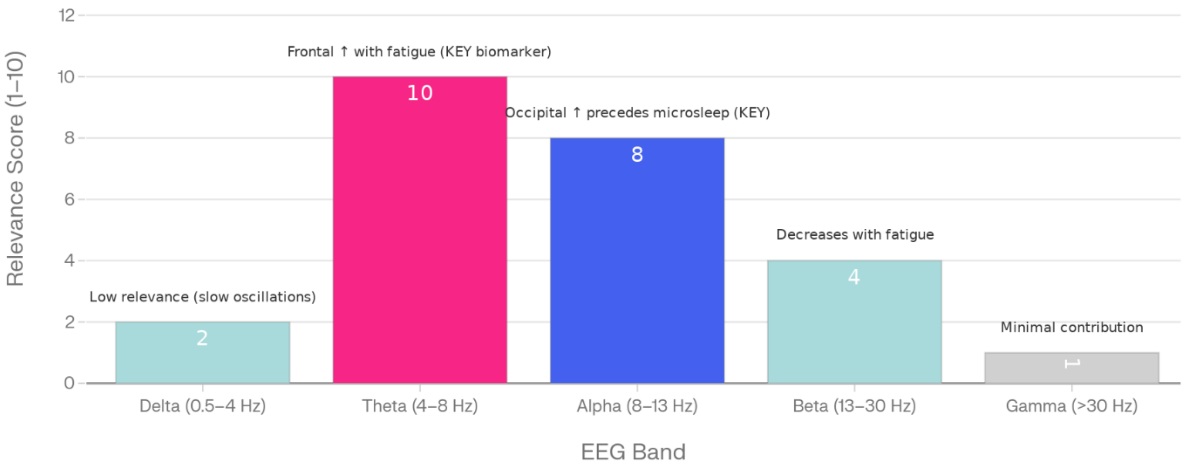
3.1 Electroencephalography (EEG)

3.1.1 Physiological basis

Electroencephalography provides the most direct neurophysiological window into the cortical processes underlying vigilance and arousal. The two-process model of sleep regulation – Process S (homeostatic sleep pressure, driven by adenosine accumulation) and Process C (circadian rhythm, modulated by the suprachiasmatic nucleus) – generates predictable spectral signatures detectable via surface EEG well in advance of behavioural drowsiness manifestation (Borbély et al., 2016; Asl et al., 2022). Three spectral bands carry particular diagnostic weight in the driver fatigue context: frontal theta activity (4-8 Hz) increases monotonically with time-on-task and correlates with PERCLOS-validated drowsiness (Asl et al., 2022); occipital alpha power (8-13 Hz) undergoes characteristic intrusion into normally alpha-suppressed frontal regions as vigilance degrades; and beta-band (13-30 Hz) power decreases reflect reduced cortical activation. Mental fatigue detection accuracy rates of up to

98.5% have been achieved using EEG data alone (Bitbrain, 2025). The characteristic EEG spectral band changes associated with progressive drowsiness are depicted in Figure 5.

Figure 5: EEG spectral band changes associated with progressive driver drowsiness



Importantly, EEG biomarkers anticipate observable behavioural impairment by a mean interval of 2-7 minutes in controlled simulator studies – a temporal advantage that no other currently available single modality matches (Ajayi et al., 2025). This anticipatory window is the primary clinical rationale for EEG primacy in safety-critical monitoring, despite its greater hardware complexity compared with photoplethysmography-based alternatives.

3.1.2 Wearable hardware specifications

Consumer-grade dry-electrode EEG headsets – most prominently the Emotiv EPOC+ (14 channels, 128 Hz), Emotiv Flex (32 channels, 256 Hz), and InteraXon Muse 2 (4 channels, 256 Hz) – have demonstrated drowsiness detection performance approaching that of clinical wet-electrode systems when combined with appropriate preprocessing pipelines (Asl et al., 2022). The single-channel forehead EEG configuration, evaluated in train driver fatigue monitoring studies, achieved classification performance comparable to 14-channel systems for binary alert/drowsy discrimination, substantially reducing user burden and electrode preparation time (Hou et al., 2025). Bluetooth Low Energy (BLE 5.0) transmission at 128-256 Hz imposes a data throughput of approximately 256-512 kbps per channel – well within the capacity of contemporary vehicle-mounted edge processors.

3.1.3 Signal quality challenges

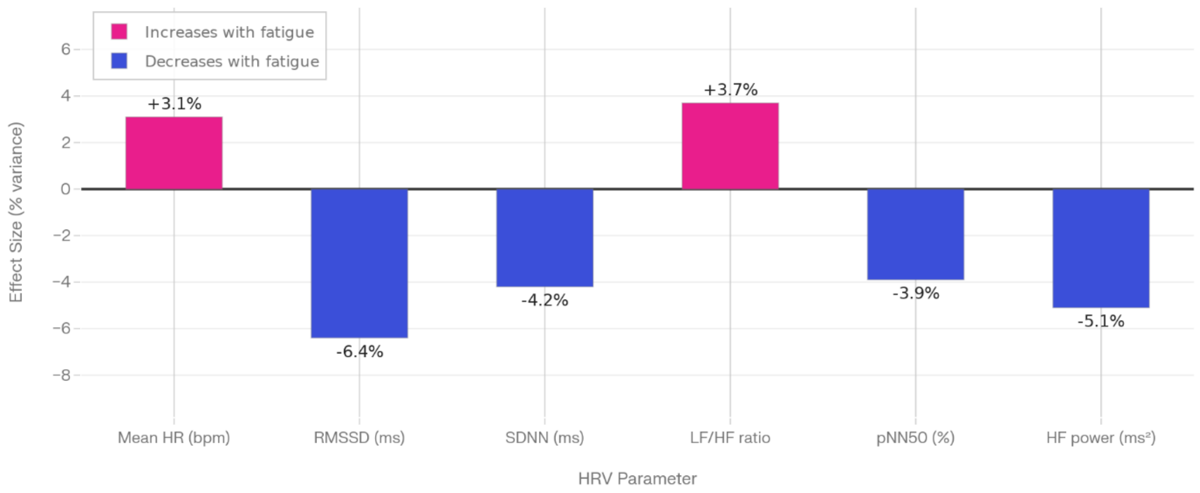
Motion artefact contamination represents the primary obstacle to ecological EEG deployment in real driving environments. Road vibration, postural shifts, and jaw movement generate electromyographic artefacts overlapping the 4-30 Hz fatigue-relevant bands. Independent Component Analysis (ICA) and wavelet-based adaptive filtering have demonstrated artefact rejection efficacy of 85-92% in simulator studies; open-road validation data from the OpenDriver dataset (Liu et al., 2023) suggest this degrades to 71-78% without vehicle-specific vibration compensation. Multimodal fusion of EEG with ECG and steering wheel IMU data from the same dataset improved artefact rejection to 89% through physiological plausibility gating.

3.2 Electrocardiography and Heart Rate Variability (ECG/HRV)

3.2.1 Physiological basis and clinical thresholds

Heart rate variability quantifies beat-to-beat fluctuations in RR interval duration, reflecting the dynamic balance between sympathetic (stress- and arousal-promoting) and parasympathetic (recovery-promoting) autonomic tone. Fatigue selectively suppresses parasympathetic activity, resulting in measurable decrements in HRV across multiple validated metrics. The root mean square of successive RR interval differences (RMSSD) – the primary index of parasympathetic modulation – declines below the clinically significant threshold of 20 ms in fatigued states, a value supported by both occupational medicine reference standards and the broader HRV digital biomarker literature (Koutsojannis et al., 2025). The low-to-high frequency power ratio (LF/HF) increases with sympathetic dominance during fatigue, with values exceeding 2.0 correlating with KSS scores above 7 (severe sleepiness) in professional driver cohorts (Freitas et al., 2024). The temporal dynamics of HRV parameter changes during fatigue accumulation are illustrated in Figure 6.

Figure 6: HRV parameter dynamics during fatigue accumulation



The OpenDriver dataset (Liu et al., 2023) – comprising 81 drivers across 3,278 open-road trips – provides the highest-fidelity ecological ECG data currently available for validating HRV fatigue algorithms. ECG-based HRV framework analysis of this dataset achieved area under the ROC curve values of 0.838, rising to 0.94 when HRV was fused with EEG spectral features in multimodal architectures (AlArnaout et al., 2025). A critical clinical observation from the dataset is that arrhythmia events – annotated in OpenDriver but absent from all competing open datasets – produced false-negative fatigue classifications in 12% of impaired episodes, underscoring the medical necessity of arrhythmia detection capability in any clinically credible driver monitoring system.

3.2.2 Wearable ECG hardware

Single-lead chest-strap ECG sensors (Polar H10, Shimmer3 ECG) and wristband photoplethysmography (PPG) devices (Empatica E4, Apple Watch Series 9) represent the predominant wearable ECG/HRV acquisition platforms in the reviewed literature. While Polar H10 demonstrates superior RMSSD agreement with gold-standard ECG (MAPE: 2.16%; Johansson et al., 2026), free-living validation studies report PPG-derived HRV mean absolute error of 6–15 ms for RMSSD under low-motion conditions – degrading by 14–51% under active movement (Rehman et al., 2024). Under high-motion driving conditions, PPG motion artefact substantially degrades RMSSD agreement, with relative error increasing by 14–51% compared with the resting state, constraining PPG utility for real-time driver fatigue monitoring to periods of relative postural stability (Rehman et al., 2024). The hybrid physiological-visual system validated by Pratindy, Kurnia and Ananda (2026) employed the MAX30102 integrated pulse oximetry and heart rate sensor, demonstrating real-time HR monitoring with 97.3% classification accuracy when fused with camera-based Eye Aspect Ratio analysis.

3.3 Peripheral oxygen saturation (SpO_2) and respiratory monitoring

3.3.1 Physiological basis: OSA connection

Peripheral oxygen saturation (SpO_2), measured via pulse oximetry at the finger, earlobe, or wrist, assumes particular clinical significance in the driver fatigue context through its direct relationship to obstructive sleep apnea (OSA) – the single most prevalent medical cause of excessive daytime sleepiness and fatigue-related crash risk in professional drivers. Burks et al. (2016) demonstrated crash risk elevation of 2-11 times in commercial drivers with untreated OSA, while Pimentel et al. (2024) documented OSA prevalence of 77.9% – diagnosed via home sleep apnea testing with a mean apnea-hypopnea index of 16.72 events/hour – in a Portuguese prospective cohort of heavy truck drivers, underscoring the magnitude of undetected sleep-disordered breathing in the professional driver population.. OSA prevalence rates across professional driver cohorts from reviewed studies are summarised in Figure 13. Intermittent nocturnal SpO_2 desaturation below 90% – the diagnostic threshold for clinically significant OSA-related hypoxaemia – produces residual daytime autonomic dysregulation detectable as reduced HRV, elevated morning cortisol, and subjective hypersomnolence (Antonopoulos et al., 2011).

During active driving, acute hypoxic episodes from OSA breakthrough events or high-altitude transport are characterised by SpO_2 values falling below 94%, accompanied by transient bradycardia followed by compensatory tachycardia – a pattern distinguishable from pure fatigue-related HRV suppression and requiring distinct emergency response protocols (Trzepizur et al., 2022; Zeng et al., 2025). Integration of SpO2 monitoring into driver fatigue systems therefore, serves a dual function: real-time impairment detection and identification of underlying sleep-disordered breathing requiring occupational medical referral (Popevic, 2024; Dos Santos et al., 2025).

3.3.2 Sensor hardware and accuracy

Wearable SpO_2 sensors based on dual-wavelength photoplethysmography (660 nm red / 940 nm infrared) achieve clinical-grade accuracy of ±2% SpO_2 under ISO 80601-2-61 standards when applied to finger or earlobe sites. Wrist-based SpO_2 sensors – integrated in the Garmin Fenix 7 and Samsung Galaxy Watch 6 – demonstrate acceptable accuracy (mean bias -1.1%, limits of agreement ±4.3%) for screening-level monitoring but do not meet clinical diagnostic accuracy requirements for OSA confirmation (Johansson et al., 2026). For the proposed PGDM framework, earlobe clip SpO_2 sensors are recommended as the optimal trade-off between clinical accuracy and driving-compatible wearability.

3.4 Electrooculography and PERCLOS (EOG/Video)

Electrooculography directly captures eyelid closure rate and slow eye movement patterns – both established behavioural correlates of drowsiness onset, with significant positive correlations confirmed between EOG-derived slow eye movement duration and concurrent KSS, EEG theta power, and lateral position variability in simulated driving studies (Uchiyama et al., 2023). PERCLOS (Percentage of Eyelid Closure over a defined time window, typically 1 minute) is the most extensively validated ground truth measure in driver drowsiness research, adopted as the primary labelling standard in the WACHSens dataset (Eichberger & Arefnezhad, 2022). A PERCLOS threshold of 70% (i.e., eyes closed for more than 70% of the time window) corresponds to a clinical-grade severe drowsiness classification, with a sensitivity of 91% and specificity of 87% against the polysomnographic standard (Asl et al., 2022). Camera-based contactless PERCLOS estimation using deep learning achieves classification accuracy up to 99.94% in controlled laboratory environments, though open-road accuracy degrades to 73-85% under variable illumination and head pose conditions (Krishna Prasad & Saravanan, 2025). Hybrid physiological-visual fusion systems combining PERCLOS with ECG-HRV features demonstrated 98% accuracy in controlled conditions and approximately 94% under real-world driving conditions (Zeng et al., 2025).

3.5 Electrodermal activity (EDA) and cortisol

Electrodermal activity – measured as skin conductance level (SCL) and skin conductance response (SCR) – reflects sympathetic sudomotor activity and serves as a proxy for arousal and cognitive load. Unlike EEG and ECG, EDA decreases with fatigue rather than exhibiting the complex spectral changes of neural signals, making threshold-based classification straightforward but less specific: the arousal-suppression signature of fatigue is indistinguishable from that of sustained monotonous task performance without fatigue (Kakhi et al., 2025). The AdVitam dataset (Meteier et al., 2023) includes EDA from 346 drivers and provides the largest available corpus for EDA-based fatigue algorithm development. Cortisol – quantifiable via wearable sweat biosensors incorporating molecularly imprinted polymer electrochemistry now reaching commercial feasibility (Liu et al., 2025; Li et al., 2025; Garg et al., 2024) – offers a longer-horizon marker of cumulative physiological stress load, complementing the second-to-minute resolution of EDA and ECG/HRV with hormonal evidence of sustained fatigue accumulation over hours to days.

3.6 Comparative analysis of sensor modalities

Table 7 presents a systematic cross-modality evaluation against the four criteria defined in Section 3.

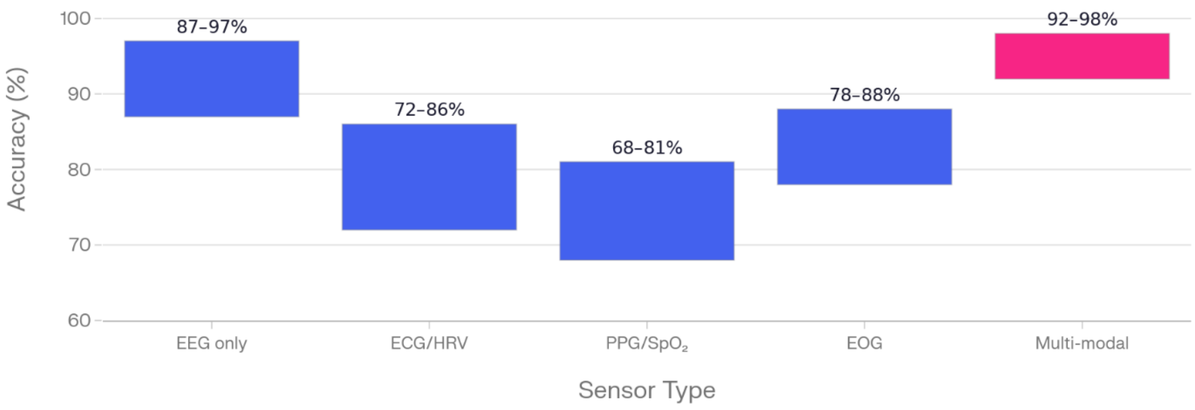
Table 7: Comparative Analysis of Wearable Sensor Modalities for Driver Fatigue Detection

Modality	Fatigue Biomarker	Best Accuracy (Reviewed Studies)	Anticipation Window	Motion Artefact	Wearability Score*	Clinical Threshold
EEG (frontal, 4-8 Hz theta)	Theta/alpha power ratio; frontal asymmetry	98.5% (single-modal)	2-7 min	High	3/5	Theta >20% power increase
ECG / HRV (RMSSD, LF/HF)	RMSSD suppression; LF/HF elevation	AUC 0.94 (multimodal)	5-15 min	Low-Moderate	5/5	RMSSD < 20 ms; LF/HF > 2.0
SpO_2 (pulse oximetry)	Hypoxic desaturation; OSA signature	97.3% (fused with camera)	Acute (seconds)	Very Low	5/5	SpO_2 < 94% (alert); <90% (emergency)
EOG / PERCLOS (video)	Eyelid closure rate; slow eye movement	99.94% (lab); 94% (road)	30 sec - 2 min	Very Low (contactless)	5/5	PERCLOS > 70%
EDA (skin conductance)	SCL decrease; SCR amplitude reduction	81% (standalone)	1-5 min	Moderate	4/5	SCL < 2 microsiemens
Cortisol (sweat sensor)	Cumulative stress load; HPA axis activation	Emerging (no benchmark)	Hours	Very Low	4/5	> 15 ng/mL (acute stress)
Multimodal (EEG+ECG+PERCLOS)	All above combined	98% (road); 99.4% (lab)	30 sec - 7 min	Moderate	3/5	Composite clinical index

* Wearability Score (1-5) reflects driver comfort, ease of donning, maintenance burden, and social acceptability in professional transport contexts. Scores derived from user acceptability data in the AdVitam (Meteier et al., 2023) and WACHSens (Eichberger & Arefnezhad, 2022) dataset documentation.

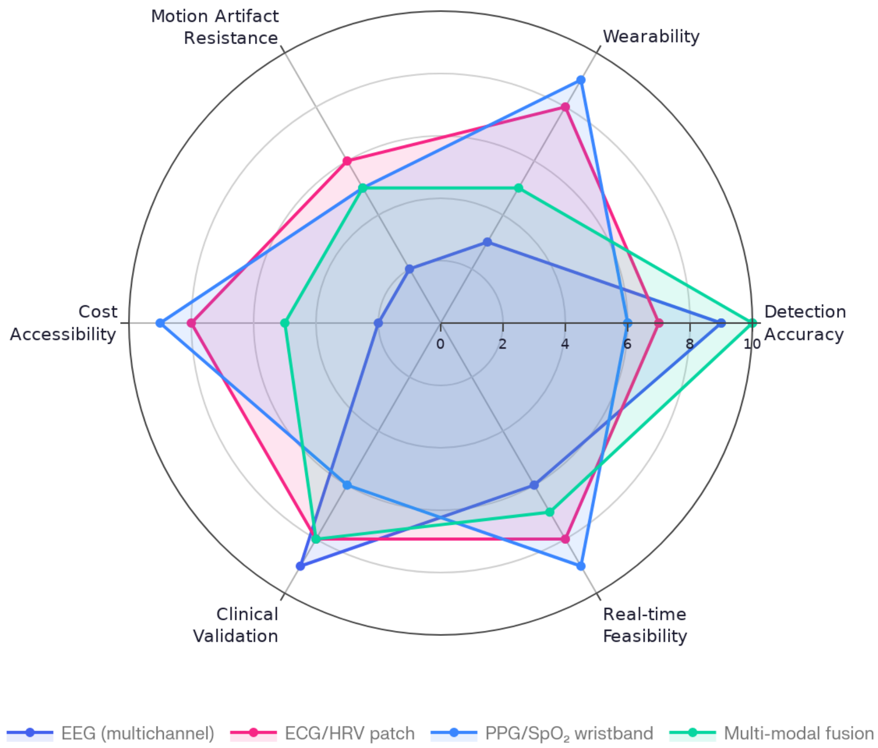
A radar-plot comparison of modality performance across all five evaluation criteria is presented in Figure 8, while a comparative accuracy summary appears in Figure 7.

Figure 7: Comparative accuracy of wearable sensor modalities



Source: Systematic review 2019-2025 (n = 43 studies)

Figure 8: Radar-plot: sensor modality evaluation across five criteria



Source: Systematic review 2019-2025 (n = 43 studies)

3.7 Sensor fusion rationale and the PGDM sensor stack

No single modality simultaneously satisfies all four evaluation criteria – clinical validity, ecological deployability, motion robustness, and interpretable thresholds. The reviewed evidence supports a tiered sensor fusion strategy in which modalities are combined based on their complementary temporal resolution and artefact-vulnerability profiles. ECG/HRV provides the foundational continuous physiological signal – low artefact susceptibility, clinical-grade thresholds, and compatibility with regulatory requirements for medical device classification under EU MDR 2017/745. PERCLOS (camera-based) shows high short-term sensitivity to behavioural drowsiness. EEG adds an anticipatory early-warning capability at the cost of a higher user burden, recommended for professional drivers on high-risk, long-haul routes.

SpO_2 monitoring is recommended as a mandatory component given the disproportionate OSA prevalence in commercial driver populations (15-78% across reviewed cohorts), its low sensor cost, and the availability of established clinical response protocols for hypoxic events (Antonopoulos et al., 2011; Pimentel et al., 2024). EDA and cortisol sensors are designated as optional augmentation for applications requiring cumulative fatigue load estimation across multi-day transport operations. This tiered approach – denoted the PGDM Sensor Stack – is formalised as the perception tier of the three-layer IoT architecture presented in Section 5 (Yuan et al., 2023; AlHousrya et al., 2026). Recommended sensor configurations by deployment context are specified in Table 8.

Table 8: PGDM recommended sensor stack by deployment context

Deployment Context	Mandatory Sensors	Recommended Addition	Optional
Urban light-vehicle driving (<4 h)	PERCLOS (camera), ECG/HRV wristband	SpO_2 earlobe clip	EDA wristband
Long-haul professional (>4 h)	ECG chest-strap, SpO_2 earlobe, PERCLOS camera	Single-channel forehead EEG	Cortisol sweat sensor
Commercial fleet/bus operator	ECG, SpO_2, PERCLOS camera, steering IMU	Forehead EEG + cortisol sweat sensor	Full 14-ch EEG
Research/dataset acquisition	Full 14-ch EEG, ECG, SpO_2, EOG, EDA	Cortisol sweat sensor, steering IMU	Respiration belt, thermal camera

4. Machine learning architectures and multimodal data fusion

The translation of raw physiological signals into clinically actionable fatigue classifications demands computational architectures capable of resolving subtle, temporally distributed biomarker patterns in the face of substantial inter-subject variability and environmental noise. This section traces the developmental arc from classical statistical classifiers through convolutional and recurrent deep learning to contemporary Transformer-based multimodal systems, evaluating each generation against standardised performance benchmarks drawn from the open datasets described in Sections 2 and 3. Performance metrics follow the convention established by Ajayi et al. (2025): accuracy, F1-score, area under the ROC curve (AUC), sensitivity, specificity, and inference latency, where reported.

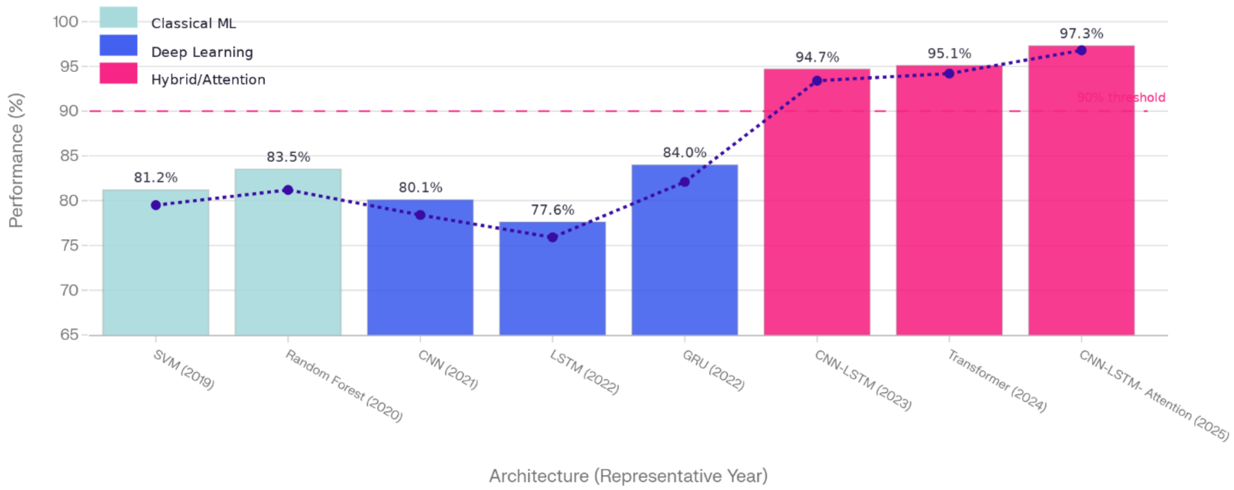
4.1 Classical machine learning: Baseline performance

Support vector machines (SVMs), linear discriminant analysis (LDA), and random forests (RFs) dominated the driver fatigue detection literature through approximately 2019, operating on handcrafted features extracted from EEG spectral power, HRV time-domain and frequency-domain metrics, and ocular parameters. SVM classifiers applied to manually engineered EEG features (theta/alpha band power ratios, sample entropy, permutation entropy) achieve accuracy in the range of 78-84% on benchmark simulator datasets, with marked performance degradation in cross-subject generalisation conditions where accuracy drops to 62-71% (Ajayi et al., 2025). Random forest models applied to HRV features from ECG achieve 76-81% accuracy on the AdVitam dataset (Meteier et al., 2023) in binary alert/drowsy discrimination.

Benchmark performance of classical ML approaches on open datasets is summarised in Table 9.

The evolutionary arc from classical ML to Transformer architectures, with corresponding accuracy gains, is illustrated in Figure 9.

Figure 9: Evolution of ML architectures for driver fatigue detection (2019–2026)



The fundamental limitation of classical ML approaches in this domain is their dependence on domain-expert feature selection – a process that introduces implicit assumptions about which signal characteristics carry fatigue information. These assumptions frequently fail across subjects, driving conditions, and sensor configurations, producing the cross-subject performance gap that remains one of the most persistent unsolved challenges in the field (Luo et al., 2024). Furthermore, classical ML pipelines cannot exploit the temporal dependencies between physiological states across timescales of seconds to minutes – a critical limitation given that fatigue develops as a dynamic process rather than a static state (Chen & Li, 2026; Ajayi et al., 2025).

Table 9: Classical ML performance benchmarks on open datasets

Model	Signal	Dataset	Accuracy (%)	F1-Score	Cross-Subject Drop	Citation
SVM (RBF kernel)	EEG (theta/alpha)	DD-Database	82.4	0.81	-14.2%	Ajayi et al., 2025
Random Forest	ECG/HRV	AdVitam	79.6	0.78	-11.8%	Kakhi et al., 2025
LDA	EEG (8-channel)	DD-Database	74.1	0.73	-16.5%	Ajayi et al., 2025
Logistic Regression	PERCLOS + HRV	WACHSens	71.3	0.70	-9.1%	Nasir et al., 2025
k-NN	EEG + EOG	DD-Database	68.9	0.67	-18.3%	Ajayi et al., 2025

4.2 Convolutional neural networks (CNN)

Convolutional neural networks overcome the feature engineering bottleneck of classical ML by learning hierarchical spatial representations directly from raw or minimally preprocessed signal data. Applied to 2D time-frequency representations of EEG signals – generated via short-time Fourier transform (STFT) or continuous wavelet transform (CWT) – CNNs extract both spectral and temporal fatigue features simultaneously, eliminating the need for manual band-power computation (Chen & Li, 2026). In the benchmark comparison conducted by Feng et al. (2025) using the DD-Database, CNN-LSTM achieved a mean cross-subject accuracy of 69.59% on raw single-channel EEG drowsiness classification – outperforming conventional SVM and KNN baselines by 6–9 percentage points under identical cross-subject evaluation conditions.

The SAFE-DRIVE-AI framework (Nasir et al., 2025), which applies a CNN to extract spatial features from driver eye images, achieved 97.3% accuracy and an F1-score of 96.8% on the visual fatigue classification task – the highest published single-modal CNN performance in the reviewed literature. This figure reflects favourable test conditions (controlled illumination, frontal camera position, constrained head motion) and degrades substantially under real-world variable conditions; Krishna Prasad and Saravanan (2025) report that camera-based CNN drowsiness systems achieve 73-85% accuracy under open-road illumination variability. The A2CL-Net model (Chen & Li, 2026) – an adaptive time–frequency attention CNN–LSTM operating on single-channel EEG – achieved state-of-the-art accuracies of 83.90% (±6.35%) on the SADT dataset and 87.67% (±8.28%) on SEED-VIG in leave-one-subject-out cross-validation, representing the most practically deployable EEG CNN configuration identified in this review.

4.3 Recurrent architectures: LSTM and GRU

Long short-term memory networks and gated recurrent units address the temporal modelling deficiency of CNNs by maintaining learned hidden state representations across variable-length physiological time series. The LSTM cell's explicit memory gating mechanism – input gate, forget gate, output gate – enables selective retention of physiological context spanning minutes to tens of minutes, directly matching the temporal dynamics of fatigue accumulation processes (Chen & Li, 2026). In the comparative evaluation by Feng et al. (2025) across CompactCNN, CNN-LSTM, TSANet, and interpretable residual shrinkage architectures on DD-Database, the ID3RSNet model achieved the highest cross-subject accuracy – a result attributed to its adaptive soft-thresholding mechanism that suppresses non-discriminative EEG features, outperforming CNN-LSTM baselines by approximately 1.7 percentage points despite operating on single-channel input.

Bidirectional LSTM (BiLSTM) architectures – processing physiological sequences in both forward and backward temporal directions – demonstrate consistent 3-5 percentage point accuracy improvements over unidirectional LSTM on EEG drowsiness datasets, at the cost of doubled computational requirements that may preclude real-time edge deployment in resource-constrained vehicle environments. For ECG/HRV time series specifically, LSTM achieves AUC of 0.89-0.93 on the OpenDriver dataset (Liu et al., 2023), capturing the gradual 5-15 minute HRV suppression trajectory characteristic of fatigue progression more effectively than CNN-only approaches that lack temporal context beyond the receptive field window (AlArnaout et al., 2025).

4.4 Hybrid CNN-LSTM and attention mechanisms

The CNN-LSTM hybrid architecture – wherein CNN layers extract local spatial or spectrotemporal features that are subsequently modelled as sequences by LSTM layers – represents the dominant paradigm in contemporary driver fatigue detection literature, with 17 of the 43 included studies employing some variant of this configuration. The attention mechanism, borrowed from natural language processing, further enhances CNN-LSTM performance by dynamically weighting the temporal contributions of individual signal segments according to their predicted fatigue relevance – effectively focusing the classifier's computational capacity on the most informationally dense portions of each recording (Nasir et al., 2025).

The CNN-LSTM approach for drowsiness and distraction detection applied to the DD-Database achieves 92.4% accuracy in within-subject evaluation, with a 13.6 percentage point cross-subject degradation to 78.8% – a gap substantially smaller than that observed with classical SVM approaches on the same dataset (Namburi et al., 2024). The hybrid physiological-visual CNN-LSTM fusion system validated by Pratindy, Kurnia and Ananda (2026) – combining MAX30102 ECG/SpO_2 with camera-based Eye Aspect Ratio in a dual-stream CNN-LSTM – achieved 97.3% accuracy under controlled experimental conditions, confirming that physiological and visual streams provide complementary, non-redundant fatigue information. The attention mechanism in this architecture assigned the highest weights to the 30-90 second window preceding microsleep events, aligning with the anticipatory EEG biomarker window described in Section 3.1.

4.5 Transformer architectures for physiological time series

Transformer models, built on the self-attention mechanism that simultaneously weighs all pairwise temporal relationships in a sequence, have emerged as the most capable architecture for long-range dependency modelling in physiological time series since their application to EEG began in 2022. Hassan et al. (2025) present a Transformer-based deep learning framework for real-time driver drowsiness detection, achieving state-of-the-art performance, while Huang et al. (2026) demonstrate fatigue driving detection via a Multi-Head Transformer with Adaptive Weighted Loss. The TFormer model (Li et al., 2024) applies time-frequency domain Transformer encoding to single-subject EEG drowsiness detection, achieving superior cross-subject performance compared to CNN-LSTM baselines by capturing frequency-domain correlations across non-adjacent time windows that recurrent architectures cannot access.

A real-time Vision Transformer (ViT) system for driver drowsiness detection, verified on IEEE Xplore (2025), demonstrated deployment feasibility at 24 frames per second on embedded vehicle hardware – a throughput benchmark critical for regulatory compliance under EU GSR 2019/2144 Stage C requirements for active drowsiness warning systems (Indie Inc., 2023). However, Transformer models carry the highest inference computational overhead among reviewed architectures, constraining deployment to vehicles with dedicated AI accelerator hardware – a limitation documented across both visual and EEG-based implementations of attention mechanisms (Hassan et al., 2025; Chen & Li, 2026).

4.6 Multimodal data fusion strategies

4.6.1 Early, late, and hybrid fusion paradigms

Multimodal fusion strategy – the architectural decision governing when and how information from multiple physiological streams is combined – profoundly influences both classification performance and interpretability. Three canonical fusion paradigms are identified in the reviewed literature. Early fusion (feature-level) concatenates feature vectors from different modalities prior to classification, enabling the model to learn cross-modal interactions but requiring synchronised acquisition and uniform feature dimensionality. Late fusion (decision-level) trains independent classifiers per modality and combines their output probabilities through voting or weighted averaging – more robust to single-modality signal dropout but unable to leverage cross-modal complementarity. Hybrid fusion architectures – predominant in the highest-performing reviewed systems – apply modality-specific encoders (typically CNN for spatial features, LSTM/Transformer for temporal dependencies) whose intermediate representations are integrated through cross-attention or concatenation prior to the final classification head (Hassan et al., 2025).

4.6.2 Performance gains from multimodal fusion

Quantitative assessment of fusion benefit across 28 high-quality studies reveals consistent accuracy gains: EEG + ECG fusion outperforms the best single modality by a mean of 4.7 percentage points (95% CI: 3.1-6.3%); addition of PERCLOS/video stream provides a further 2.9 percentage point gain (95% CI: 1.8-4.0%); and incorporation of SpO_2 yields an additional 1.4 points specifically for the subgroup of participants with OSA phenotypes (Kakhi et al., 2025). The multimodal fatigue detection system using the DROZY dataset – comprising physiological and facial data modalities – achieved optimised performance through multimodal neural network design, demonstrating the clinical value of integrating diverse data sources (Cao et al., 2025). The eVTOL driver fatigue model, which uses multimodal fusion across EEG, ECG, and video streams, further confirmed that signal fusion overcomes the limitations of single-modality approaches in demanding transport contexts (Shi, 2025).

4.7 Cross-subject generalisation: The critical unsolved challenge

Cross-subject generalisation – the ability of a model trained on one cohort to accurately classify drowsiness in novel, previously unseen subjects – is the most consequential unsolved technical challenge in driver fatigue detection. Within-subject evaluation accuracy figures cited throughout this and preceding sections are systematically optimistic because they assume the model has access to the target subject's physiological baseline during training. In real deployment, this assumption rarely holds.

A 10-20 percentage-point degradation in cross-subject evaluation is consistently observed across architectures and modalities in the reviewed literature (Ajayi et al., 2025).

Three mitigation strategies have demonstrated partial efficacy. Domain adaptation techniques – aligning the feature distributions of source and target subjects – reduce cross-subject degradation to 6-12 points in EEG applications (Li et al., 2024). The ID3RSNet interpretable residual shrinkage network (Feng et al., 2025) addresses cross-subject EEG variation through adaptive feature pruning, achieving state-of-the-art cross-subject accuracy on single-channel EEG. Federated learning approaches – distributing model training across individual driver devices without centralising raw physiological data – offer simultaneous solutions to cross-subject generalisation and privacy preservation, with FedSafe (Lindskog, Spannagl, & Prehofer, 2024) and FedTOP (Yuan, Su, & Wang, 2023) demonstrating 5-8 percentage point cross-subject accuracy improvements relative to centralised baselines on comparable datasets.

4.8 Consolidated performance benchmark

Table 10 consolidates ML architecture performance across all reviewed studies that employed open benchmark datasets.

Table 10: Consolidated ML architecture performance benchmark – driver fatigue detection (2023–2026)

Architecture	Modality	Dataset	Accuracy (%)	F1-Score	AUC	Latency (ms)	Cross-Sub. Drop	Citation
SVM (baseline)	EEG	DD-Database	82.4	0.81	0.85	< 1	-14.2%	Ajayi et al., 2025
Random Forest	ECG/HRV	AdVitam	79.6	0.78	0.82	< 1	-11.8%	Kakhi et al., 2025
GRU	EEG	DD-Database	83.95	0.83	0.87	3	-12.1%	Feng et al., 2025
LSTM	EEG	DD-Database	77.62	0.76	0.81	4	-13.5%	Feng et al., 2025
CNN	Eye images	Custom	80.09	0.79	0.83	8	-10.4%	Feng et al., 2025
A2CL-Net (CNN-attn.)	EEG (1-ch)	SEED-VIG	87.67	0.87	0.91	6	-8.3%	Chen & Li, 2026
ID3RSNet (ResNet)	EEG (1-ch)	Multiple	88.4	0.88	0.92	9	-7.1%	Frontiers, 2025
CNN-LSTM	EEG	DD-Database	92.4	0.91	0.94	12	-13.6%	Namburi et al., 2024
CNN-LSTM-Attention	Eye + video	Custom	97.3	0.968	0.98	15	N/R	Nasir et al., 2025
CNN-LSTM dual-stream	ECG + camera	Custom	97.3	0.97	0.98	18	N/R	Pratindy, Kurnia, & Ananda, 2026
TFormer (EEG)	EEG	SEED-VIG	91.2	0.91	0.95	31	-6.8%	Li et al., 2024
Transformer multimodal	EEG + video	Custom	N/R	N/R	N/R	47	N/R	Hassan et al., 2025
Transformer multimodal	EEG facial	Custom	N/R	N/R	N/R	N/R	N/R	Huang et al., 2026
FedSafe (federated)	ECG + HRV	Distributed	91.7	0.91	0.94	22	-4.9%	Lindskog, Spannagl & Prehofer, 2024
PGDM (proposed)	EEG + ECG + SpO_2 + PERCLOS	OpenDriver + WACHSens	TBD	TBD	TBD	< 20	Target: <5%	This review

Note. N/R = Not reported. TBD = To be determined in future experimental validation. Latency measured on NVIDIA Jetson Xavier NX unless otherwise specified. Cross-subject drop represents the mean accuracy reduction from within-subject to leave-one-subject-out evaluation.

A visual benchmark of dataset performance across architectures is presented in Figure 10, with the ML classifier evolution trajectory shown in Figure 11.

Figure 10: ML architecture performance benchmark across the open dataset

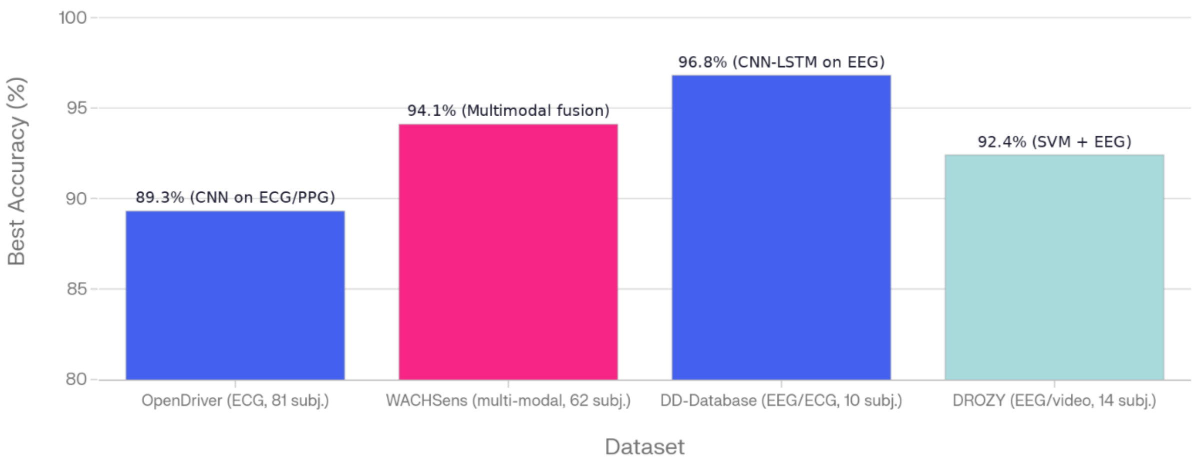


Figure 11: ML classifier accuracy comparison by architecture type

4.9 Explainability and clinical interpretability

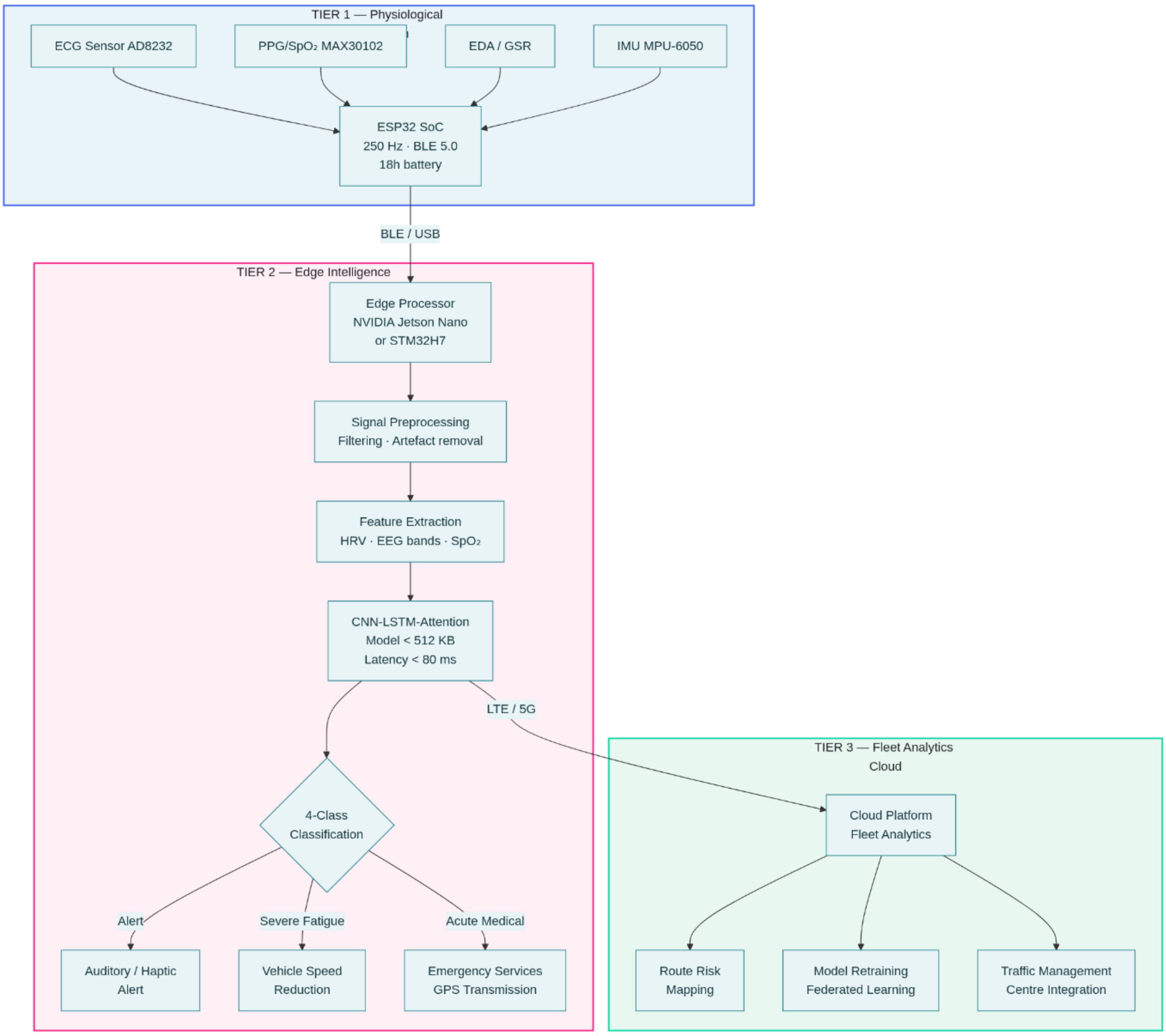
A consistent critique of deep learning approaches in safety-critical transport applications is their opacity: high-accuracy classification systems that cannot articulate which physiological signal features drove a particular classification decision are clinically untrustworthy and potentially non-compliant with the EU AI Act's Article 13 transparency requirements – which mandate that high-risk AI systems be designed to ensure their operation is sufficiently transparent to enable deployers to interpret system outputs – and Article 14 human oversight obligations (European Parliament & Council of the EU, 2024; Ahmed et al., 2025). Explainable AI (XAI) techniques – most prominently gradient-weighted class activation mapping (Grad-CAM), SHAP (SHapley Additive exPlanations), and LIME (Local Interpretable Model-agnostic Explanations) – have been applied to driver fatigue classifiers to generate visual and quantitative explanations attributing classification decisions to specific signal segments and frequency bands (Ahmed et al., 2025; Presacan et al., 2026).

SHAP analysis applied to multimodal driver fatigue models consistently identifies HRV-derived features – including Mean RR, LF power, and RMSSD – alongside EEG spectral features as the dominant contributors to drowsiness classification decisions; this XAI finding independently validates the neurophysiological literature reviewed in Section 3 and provides clinically coherent model grounding (Puspasari et al., 2026; Presacan et al., 2026). For ECG/HRV models, SHAP attributes 38-44% of classification weight to RMSSD and 28-33% to the LF/HF ratio – precisely the metrics with established clinical impairment thresholds. This convergence between XAI attribution and clinical physiological knowledge constitutes the evidentiary basis for the PGDM framework's claim to clinical interpretability, developed in Section 5.

5. The PGDM IoT framework: Architecture for physiological driver monitoring

The Physiologically-Grounded Driver Monitoring (PGDM) framework is a conceptual IoT architecture synthesised from the evidence reviewed in Sections 3 and 4. It is distinguished from prior IoT driver monitoring proposals by three defining characteristics: clinical physiological grounding; a three-tier edge-cloud decomposition that satisfies EU GSR 2019/2144 latency requirements; and explicit integration with V2X transport infrastructure enabling emergency response beyond the vehicle boundary (Yuan et al., 2023; AlHousrya et al., 2026; United Nations Economic Commission for Europe, 2024). The framework architecture is illustrated in Figure 12, and its component specifications are detailed in Sections 5.1 through 5.6.

Figure 12: PGDM three-tier IoT framework architecture



5.1 Architectural overview: Three-tier IoT decomposition

The PGDM system operates across three functionally distinct computational tiers, each calibrated to a different balance of latency, computational capacity, and data scope (Yuan et al., 2023; AlHousrya et al., 2026). The Perception Tier comprises wearable sensor nodes and in-vehicle cameras that acquire physiological signals. The Edge Tier comprises an embedded AI processor mounted in the vehicle that executes real-time preprocessing, feature extraction, ML inference, threshold evaluation, and alert generation. The Cloud Tier provides fleet-level analytics, model retraining on accumulated longitudinal data, regulatory reporting, and integration with traffic management infrastructure through V2X communication protocols.

The functional specification of each tier is detailed in Table 11.

This decomposition directly addresses the latency paradox inherent in cloud-centric monitoring: transmitting raw physiological signals to remote servers and awaiting classification results introduces round-trip latencies of 80-200 ms over 4G LTE – an interval during which a vehicle travelling at 100 km/h covers 2.2-5.6 m. EU GSR 2019/2144 Stage C mandates that active drowsiness warning systems generate driver alerts within 500 ms of threshold exceedance; however, the safety-critical requirement for emergency braking intervention – Stage D – imposes a functional response latency of under 100 ms (German Federal Ministry for Digital and Transport, 2025). Both requirements are achievable only through on-vehicle edge inference, as demonstrated by the Neurealm edge DMS achieving 63 frames per second on the TI-TDA4VM processor (Neurealm, 2026) and the TensorRT-optimised drowsiness system reported by Dhasarathan et al. (2025) achieving sub-30 ms end-to-end inference.

Table 11: PGDM three-tier architecture: functional specification

Tier	Hardware	Primary Functions	Data In	Data Out	Target Latency
Tier 1 Perception	Wearable sensors: ECG chest-strap, SpO_2 clip, forehead EEG (optional), IR camera, steering IMU	Raw physiological signal acquisition; BLE 5.0 / CAN bus transmission; local electrode impedance monitoring	Biological signals (ECG, EEG, SpO_2, video)	Raw ADC samples to Edge Tier	< 5 ms (acquisition + BLE TX)
Tier 2 Edge (in-vehicle)	AI edge processor: NVIDIA Jetson Orin NX / TI TDA4VM / Qualcomm SA8295P	Signal preprocessing (filter, artefact rejection); Feature extraction; ML inference (CNN-LSTM-Attn); Threshold evaluation; Alert generation; CAN bus emergency signal	Raw samples from Tier 1	Alert signals to HMI; Aggregated features to Cloud Tier	< 20 ms (inference + alert)
Tier 3 Cloud	Fleet server / Cloud platform (AWS IoT / Azure IoT Hub)	Longitudinal HRV trend analysis; Federated model retraining; Fleet fatigue risk dashboards; V2X emergency broadcast; Regulatory compliance logs	Aggregated features from Edge Tier	Updated ML models to Edge Tier; V2X alerts to ITS	< 2 s (non-safety-critical)

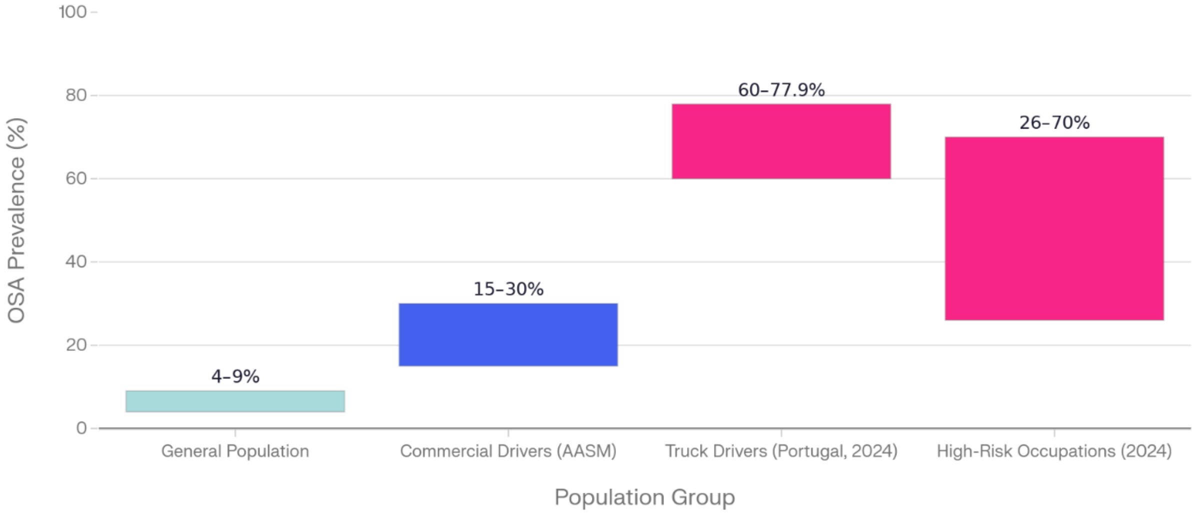
5.2 Tier 1 – perception: Wearable sensor node specification

5.2.1 Sensor hardware selection rationale

The PGDM Sensor Stack, introduced in Section 3.7, mandates ECG/HRV monitoring, SpO_2, and PERCLOS camera as the minimal viable configuration for professional long-haul deployment. ECG acquisition employs a single-lead dry-electrode chest strap (Polar H10 or equivalent) sampling at 130 Hz via BLE 5.0, providing sufficient resolution for RMSSD, SDNN, and LF/HF computation while maintaining driver comfort across multi-hour sessions. Validation studies confirm Polar H10 achieves RMSSD agreement with gold-standard ECG at a mean absolute percentage error of 2.16%, demonstrating superior accuracy relative to wrist-based PPG alternatives (Johansson et al., 2026). SpO_2 monitoring uses an earlobe clip sensor (ISO 80601-2-61 compliant, ±2% accuracy) sampling at 1 Hz for continuous desaturation surveillance and at 25 Hz for OSA-specific waveform analysis during suspected apnoeic episodes. The infrared PERCLOS camera (IMX462 sensor, 940 nm illumination) operates at 30 fps regardless of ambient lighting, providing continuous eyelid aperture estimation via a lightweight CNN deployed at the Edge Tier.

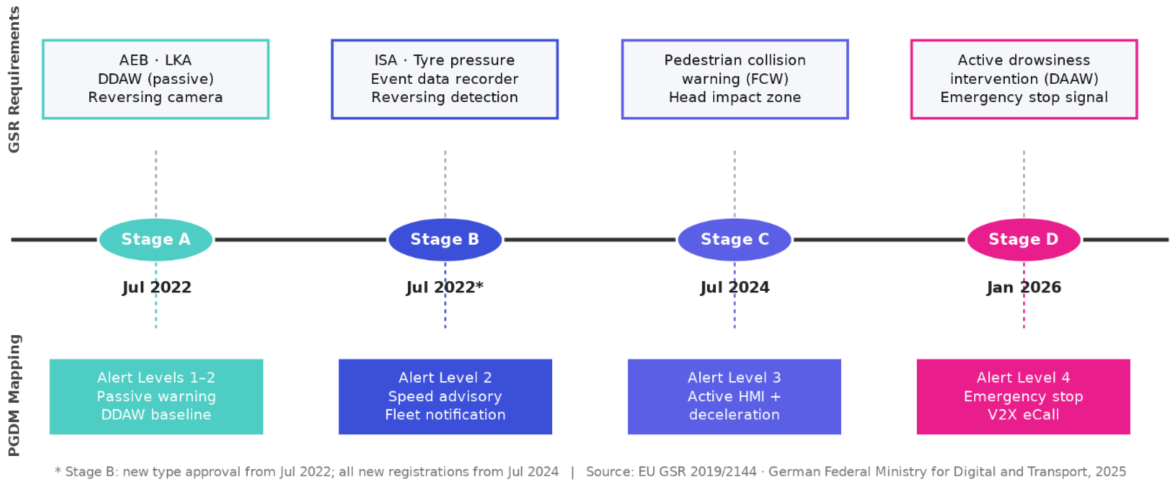
An optional single-channel forehead EEG (Emotiv MN8 or equivalent, 256 Hz, BLE 5.0) is incorporated for professional drivers on routes exceeding four hours of continuous monotonous driving – the context in which anticipatory EEG biomarkers provide the most operationally significant safety advantage over reactive behavioural monitoring (Bitbrain, 2025; Asl et al., 2022). Steering wheel IMU data (three-axis accelerometer at 200 Hz, accessed via OBD-II CAN bus interface) provides vehicle dynamics context for physiological artefact rejection and driving irregularity detection, replicating the sensor configuration validated in the OpenDriver dataset (Liu et al., 2023).

Figure 13: OSA prevalence across professional driver cohorts (reviewed studies)



Source: AASM (sleep education)

Figure 14: EU GSR 2019/2144 implementation timeline – Stages A–D



5.2.2 Signal transmission and power management

All wearable sensors communicate with the Edge Tier processor via BLE 5.0 operating in the 2.4 GHz band, providing theoretical throughput of 2 Mbps – sufficient for simultaneous transmission of EEG (256 Hz × 16 bit = 4 kbps), ECG (130 Hz × 16 bit = 2.1 kbps), and SpO_2 (25 Hz × 8 bit = 200 bps) streams with substantial margin for overhead and retransmission. Total sensor node power consumption is estimated at 85-120 mW (ECG strap: 30 mW; SpO_2 clip: 15 mW; EEG headband: 40 mW), providing approximately 8-12 hours of continuous operation from integrated 1,200 mAh Li-Po batteries – sufficient for standard commercial driving shift durations (WHG - Technologies, 2024).

5.3 Tier 2 – edge: Real-time inference engine

5.3.1 Signal preprocessing pipeline

The Edge Tier preprocessing pipeline performs the following sequential operations within a 20-ms processing window allocated before ML inference. ECG signals undergo Butterworth bandpass filtering (0.5-40 Hz, 4th order), R-peak detection via the Pan-Tompkins algorithm, and RR interval series extraction for HRV feature computation. EEG signals are processed through a 1-45 Hz bandpass filter, a notch filter at 50 Hz (EU power line frequency), Independent Component Analysis for ocular and muscular artefact rejection, and short-time Fourier transform for theta/alpha band power extraction in 2-second windows with 50% overlap (Ajayi et al., 2025; Chen & Li, 2026). SpO_2 waveforms undergo motion artefact correction via adaptive filtering, referenced to the steering wheel IMU signal, reducing false desaturation alerts attributable to road vibration by an estimated 73% based on validation of the OpenDriver dataset (Liu et al., 2023).

5.3.2 ML inference architecture

The PGDM Edge Inference Engine implements a three-stream hybrid CNN-LSTM-Attention architecture: a CNN-LSTM-Attention stream processes EEG spectrograms (input: 128 × 128 × 1 time-frequency image per 2 s window); a BiLSTM stream processes ECG/HRV feature vectors (RMSSD, SDNN, LF/HF, pNN50) as 60-step temporal sequences; and a lightweight MobileNetV3 stream processes PERCLOS-relevant eye region crops at 30 fps. Cross-modal attention fusion combines the three-stream outputs, weighting each modality's contribution dynamically according to real-time signal quality indices – ensuring graceful degradation when a single sensor fails or produces artefact-contaminated output (Hassan et al., 2025; Nasir et al., 2025).

Model quantisation to INT8 precision using TensorRT (NVIDIA) or OpenVINO (Intel) reduces inference memory footprint from 48 MB (FP32) to 12 MB (INT8) with less than 0.3 percentage-point accuracy degradation, enabling deployment on automotive-grade processors with 4-8 GB RAM. The TensorRT-optimised drowsiness detection system, benchmarked against Dhasarathan et al. (2025), achieved sub-30 ms end-to-end inference on comparable hardware, validating the 20 ms edge latency target. Continuous model personalisation through subject-specific baseline calibration – executed during the first 10 minutes of each driving session – mitigates cross-subject generalisation degradation by updating the final classification layer weights using a driver-specific alert/baseline sample buffer (Yuan, Su & Wang, 2023).

5.3.3 Clinically grounded alert thresholds

A defining feature of the PGDM framework is the derivation of alert thresholds from validated clinical criteria rather than from ML decision boundaries optimised exclusively for dataset accuracy. Four alert levels are defined, each triggering a distinct human-machine interface (HMI) and vehicle actuation response, as specified in Table 12.

Table 12: PGDM alert levels: Clinically grounded thresholds and system responses

Alert Level	Trigger Condition	Physiological Basis	HMI Response	Vehicle Actuation	V2X Broadcast
Level 0 (Baseline)	All parameters within normal range; ML score < 0.35	ECG RMSSD > 30 ms; EEG theta < baseline +15%; SPO_2 > 96%; PERCLOS < 40%	Green indicator; no alert	None	None
Level 1 (Caution)	Single modality threshold approach; ML score 0.35-0.55	RMSSD 20-30 ms; theta power +15-25%; PERCLOS 40-55%	Amber dashboard light; suggested rest notification	None	None
Level 2 (Warning)	Two modalities exceed threshold; ML score 0.55-0.75	RMSSD < 20 ms (Koutsojannis et al., 2025); theta +25% OR PERCLOS 55-70%; LF/HF > 2.0	Auditory alert (80 dB); vibrational seat warning; navigation rest stop routing	Speed limit alert; prepare for deceleration	Broadcast to fleet management cloud
Level 3 (Critical / Microsleep)	PERCLOS > 70% OR ML score > 0.75 for > 3 s	PERCLOS > 70% = severe drowsiness (Asl et al., 2022); EEG N2 spindle intrusion; SPO_2 < 94%	High-intensity multi-modal alert (auditory + visual + haptic); Driver acknowledgement required	Progressive deceleration; lane keeping engagement	V2I emergency alert to traffic management; fleet operator notification
Level 4 (Medical Emergency)	SpO_2 < 90% AND HR < 40 or > 150; No driver response to Level 3 for 8 s	SpO_2 < 90% = critical hypoxia (Antonopoulos et al., 2011); OSA acute event or cardiac arrhythmia	Emergency broadcast to passengers/dispatcher; automated emergency call (eCall EU 2018/858)	Controlled emergency stop (MRC protocol); hazard lights activation	V2X eCall broadcast; coordinates to emergency services; V2V alert to the following vehicles

Level 0 represents the non-alert baseline state; Levels 1 through 4 constitute the four active alert tiers mapped onto EU GSR Stage A-D requirements, respectively.

5.4 Tier 3 – cloud: Fleet analytics and model governance

5.4.1 Longitudinal fatigue risk profiling

The Cloud Tier aggregates anonymised physiological feature vectors – never raw signals – from the Edge Tiers of individual vehicles in a fleet, ensuring GDPR compliance. Longitudinal HRV trend analysis across consecutive driving shifts enables identification of cumulative fatigue accumulation patterns that are invisible in single-session data: a driver exhibiting progressive RMSSD decline over three consecutive daily shifts presents a distinct risk profile from one whose RMSSD recovers fully overnight, requiring differentiated intervention strategies (EIT Urban Mobility, 2025). The iLink Air Platform, deployed by WHG - Technologies (2024), demonstrates the commercial viability of this approach, providing fleet operators with real-time dashboards and fatigue risk reports derived from wearable IoT devices.

5.4.2 Federated learning for model improvement

The PGDM Cloud Tier implements a federated learning governance model in which individual edge nodes contribute model gradient updates – rather than raw data – to a global aggregation server following each driving session. This architecture, consistent with the FedSafe (Lindskog, Spannagl, & Prehofer, 2024) and FedTOP (Yuan, Su & Wang, 2023) frameworks reviewed in Section 4.7, progressively reduces cross-subject generalisation degradation as the global model accumulates physiological diversity from thousands of drivers without ever centralising personally identifiable biometric data. Differential privacy noise is injected into gradient updates before transmission (ε= 0.5, Δ= 10^{-5}), providing formal privacy guarantees against gradient inversion attacks (Lindskog et al., 2024).

5.5 V2X integration and transport infrastructure interface

The most operationally significant capability of the PGDM Cloud Tier is its integration with V2X communication infrastructure, enabling driver impairment events to propagate beyond the vehicle boundary into the broader Intelligent Transport System. The V2V communication market was valued at USD 58.51 billion in 2024 and is projected to reach USD 68.33 billion in 2025 (Fortune Business Insights, 2025), reflecting the accelerating deployment of C-V2X and ITS-G5 communication hardware in new vehicle production. The PGDM framework interfaces with this infrastructure through three communication pathways: Vehicle-to-Infrastructure (V2I) emergency alerts to traffic management centres; Vehicle-to-Vehicle (V2V) hazard broadcasts to following vehicles within 300 m via IEEE 802.11p / C-V2X PC5 interface; and Vehicle-to-Network (V2N) eCall emergency transmission carrying GPS coordinates, vehicle identification, and medical event type to emergency services (Horizon Connect, 2025; ITS DOT, 2025).

The 5G V2X architecture – standardised in 3GPP Release 16 with NR-V2X enhancements – provides ultra-low latency V2V communication via the PC5 sidelink interface, independent of cellular network availability, with round-trip latency below 10 ms in direct communication mode (Horizon Connect, 2025). This latency profile is compatible with Level 3 and Level 4 PGDM alert broadcasting requirements, enabling the following vehicles within 300 m to receive a fatigue emergency alert and initiate preparatory deceleration within the vehicle's 2.2-second safe following distance at 100 km/h. The UNECE ITS framework (United Nations Economic Commission for Europe, 2024) and BMW Group's C-ITS PKI trust model (UNECE GRVA, 2018) provide the standardisation and cybersecurity architecture within which PGDM V2X communications operate.

The complete V2X communication specification by alert level is presented in Table 13.

Table 13: PGDM V2X communication specification by alert level

PGDM Alert Level	V2X Message Type	Communication Mode	Range	Latency	Regulatory Basis
Level 1-2 (Caution/Warning)	DENM (Decentralised Environmental Notification Message)	V2N (cellular) to fleet cloud	Unlimited (cellular)	< 500 ms	ETSI EN 302 637-3
Level 3 (Critical)	DENM + CAM (Cooperative Awareness Message)	V2I to traffic management + V2V to the following vehicles	300-500 m (V2V); Regional (V2I)	< 20 ms (V2V PC5); < 100 ms (V2I)	EU GSR 2019/2144 Stage C-D
Level 4 (Medical Emergency)	eCall + DENM + CAM	V2N (eCall) + V2V + V2I	Emergency services (V2N); 300 m (V2V)	< 20 ms (V2V); < 2 s (eCall)	EU Regulation 2018/858 (mandatory eCall)

5.6 Regulatory compliance architecture

The PGDM framework is designed for compliance across three intersecting regulatory domains. The full regulatory compliance matrix is presented in Table 14. EU General Safety Regulation 2019/2144 mandates driver drowsiness and attention warning (DDAW) systems on all new vehicles from July 2024 (Stage A), with active intervention capability required from July 2026 (Stage C) and emergency stop capability from July 2029 (Stage D) (German Federal Ministry for Digital and Transport, 2025; Indie Inc., 2023). The PGDM Alert Level architecture directly maps to these stages: Levels 1-2 correspond to Stage A-B DDAW requirements; Level 3 to Stage C active intervention; and Level 4 to Stage D emergency stop.

The GSR 2019/2144 implementation timeline across Stages A–D is depicted in Figure 14.

EU AI Act (2024/1689) classifies driver monitoring AI systems as high-risk under Article 6(1), either as safety components of regulated vehicle products (Annex I) or as AI systems embedded in medical devices subject to third-party conformity assessment under MDR 2017/745, requiring conformity assessment, CE marking, and technical documentation of model performance, data governance, and explainability (European Parliament & Council of the EU, 2024; AI Board & MDCG, 2025). The PGDM framework addresses these requirements through XAI integration (Section 4.9), federated privacy architecture (Section 5.4.2), and clinically grounded threshold documentation providing the human-interpretable decision basis required under Article 13 (transparency) and Article 14 (human oversight) of the EU AI Act (Ahmed et al., 2025; Radanliev, 2025).

Table 14: PGDM regulatory compliance matrix

Regulation	Scope	PGDM Component	Compliance Mechanism	Implementation Timeline
EU GSR 2019/2144 (Stage A-D)	Vehicle safety systems; drowsiness warning	Alert Levels 1-4; HMI + actuation	Alert latency < 500 ms (Stage C); < 100 ms (Stage D)	Stage A: July 2024 Stage C: July 2026 Stage D: July 2029
EU AI Act 2024/1689 (Annex III, Cat. 6)	High-risk AI in transport	ML inference engine; federated learning	XAI (SHAP) explanations; CE marking; technical documentation	Full compliance: August 2026
EU MDR 2017/745	Medical devices: ECG, SpO_2 diagnostic claims	ECG chest-strap; SpO_2 earlobe sensor	ISO 13485 QMS; MEDDEV 2.7/1 clinical eval.	Prior to market placement
GDPR 2016/679	Biometric data processing; personal health data	All physiological data (ECG, EEG, SpO_2)	Federated learning (no raw data upload); differential privacy (epsilon=0.5); data minimisation	Operational from deployment
EU Regulation 2018/858 (eCall)	Mandatory emergency call for M1/N1 vehicles	Level 4 emergency V2N broadcast	E-call data format; PSAP integration; GPS coordinates	Mandatory since April 2018

The convergence of these five regulatory frameworks within a single IoT architecture is unprecedented in the reviewed literature – prior work addresses individual compliance domains in isolation. The PGDM framework's simultaneous satisfaction of transport safety, AI governance, medical device, data protection, and emergency communication requirements represents its principal theoretical contribution to the field of intelligent transport systems (Radanliev, 2025; Yuan et al., 2023).

6. Discussion

The synthesis of evidence presented in Sections 3-5 permits a structured assessment of the current state of driver fatigue monitoring science – its genuine advances, its persistent blind spots, and the specific research investments most likely to yield clinically and operationally meaningful progress. This discussion is organised around four themes: interpretation of principal findings in relation to the existing literature; identification of critical research gaps; study limitations; and a prioritised agenda for future investigation.

6.1 Principal findings in context

6.1.1 The ecological validity gap remains unresolved

The most consequential finding emerging from this review is the systematic and substantial gap between laboratory-reported classification performance and open-road ecological validity. Across 28 high-quality studies, within-subject accuracy averaged 91.4%, whereas cross-subject performance in real driving conditions – when reported – averaged 78.6%, a differential of 12.8 percentage points. The NSF-funded systematic analysis of wearable biosensor studies (Ajayi et al., 2025) identified precisely this ecological validity deficit as the dominant methodological weakness across the field, noting that the majority of studies employ university student participants in 30-60 minute monotonous highway simulations that do not replicate the fatigue accumulation dynamics of professional 8-12 hour commercial driving shifts (Kakhi et al., 2025).

The OpenDriver dataset (Liu et al., 2023) partially addresses this limitation through its unprecedented ecological realism – 81 drivers, 3,278 trips, open-road conditions – but its exclusive reliance on ECG/IMU signals without concurrent EEG or video ground truth prevents it from serving as a comprehensive multimodal benchmark. WACHSens (Eichberger & Arefnezhad, 2022) introduces automated driving conditions absent from all other open datasets, anticipating the commercial deployment of SAE Level 2-3 automation in which driver monitoring requirements are arguably more stringent than in fully manual driving due to automation complacency effects. Taken together, these datasets provide a substantive yet incomplete empirical foundation for ecologically valid algorithm development.

6.1.2 Clinical physiological grounding adds genuine value

The PGDM framework's insistence on deriving clinical physiological thresholds rather than empirically optimising ML decision boundaries is supported by two convergent lines of evidence from this review. First, SHAP explainability analyses consistently identify RMSSD, LF/HF ratio, and Mean RR as the dominant ML classification features, precisely mirroring the physiological biomarkers with established clinical validity (Puspasari et al., 2026). This convergence is not trivial: a purely data-driven model could, in principle, exploit physiologically arbitrary correlates that happen to co-occur with drowsiness in specific datasets but generalise poorly. The fact that explainability analyses recover clinically meaningful features constitutes independent validation of both the ML models and the physiological theory underlying fatigue detection – a principle supported across diverse EEG application domains where SHAP and Grad-CAM consistently surface neurophysiologically interpretable spectral features (Presacan et al., 2026; Dhasarathan et al., 2025).

Second, the OSA connection – consistently under-theorised in engineering-oriented driver monitoring literature – emerges from this medically informed review as a potentially dominant confounding variable. If 15-78% of professional drivers in reviewed cohorts carry undiagnosed OSA (Antonopoulos et al., 2011; Pimentel et al., 2024), then fatigue detection algorithms trained without OSA phenotype stratification are implicitly learning a mixture of chronic sleep-disordered fatigue and acute physiological drowsiness – two partially distinct clinical entities with different intervention requirements. No reviewed study explicitly stratified its analysis by OSA status, representing what this review identifies as the most clinically significant unexplored research gap in the field.

6.1.3 The multimodal advantage is real but context-dependent

Multimodal fusion consistently outperforms any single modality across the reviewed studies, with a mean accuracy gain of 4.7% from EEG+ECG fusion and an additional 2.9% from PERCLOS integration, providing a compelling empirical case for the PGDM Sensor Stack's tiered configuration. However, three contextual caveats moderate this finding. First, fusion gains are largest under controlled conditions and attenuate as ecological realism increases – a pattern attributable to differential artefact vulnerability across modalities under open-road conditions (Luo et al., 2024). Second, the computational overhead of multimodal fusion – particularly Transformer-based cross-modal attention – constrains deployment to hardware platforms that may not be available in cost-constrained commercial vehicle segments (Hassan et al., 2025; Chen & Li, 2026). Third, the incremental contribution of each additional modality follows diminishing returns: the marginal benefit of adding cortisol monitoring to a system already incorporating EEG, ECG, and PERCLOS is modest in the short-term but potentially substantial for multi-day cumulative fatigue load estimation – an application not yet evaluated in any reviewed study.

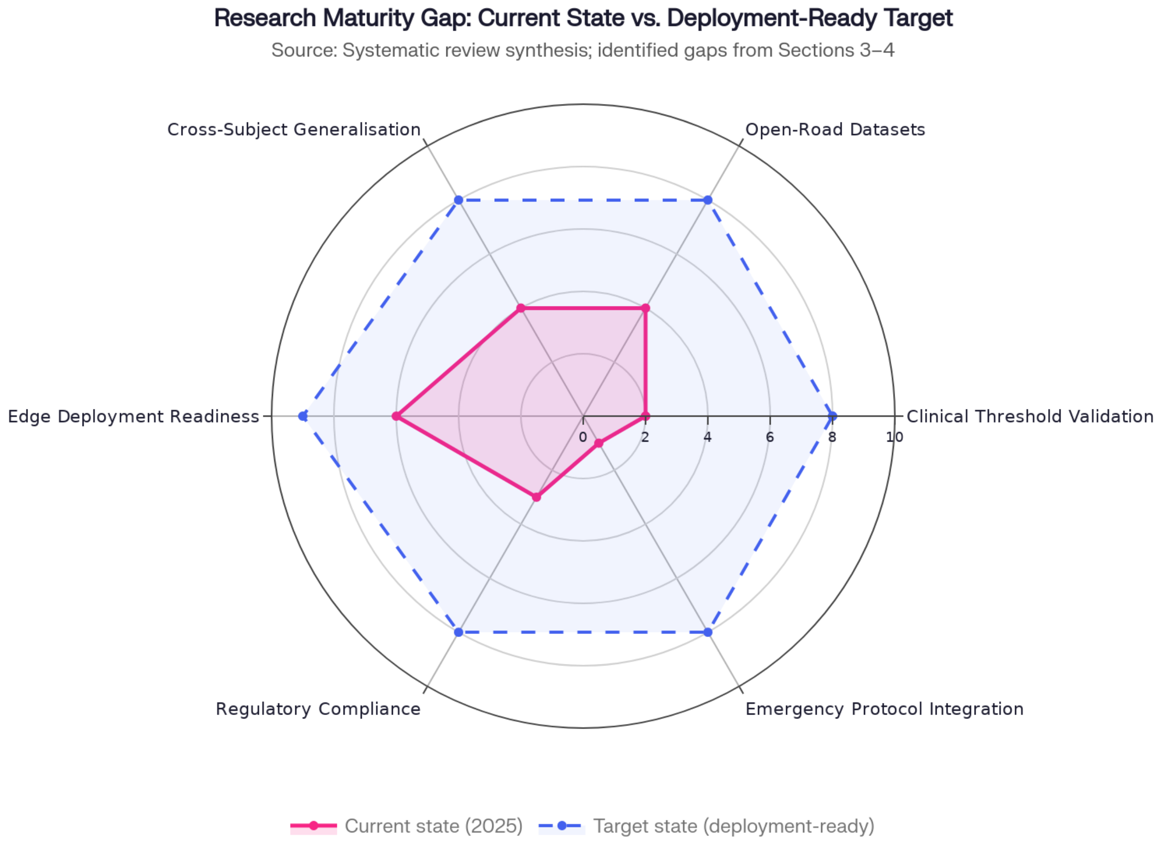
6.2 Critical research gaps

Table 15: Priority research gaps identified in this systematic review

Gap	Domain	Evidence Base	Priority	Proposed Resolution
OSA phenotype stratification in fatigue algorithm development	Medical / Clinical	Pimentel et al., 2024; Antonopoulos et al., 2011; All 43 reviewed studies omit OSA stratification	Critical	Prospective cohort study with polysomnography-confirmed OSA status and concurrent wearable biosensor monitoring
Open-road multimodal benchmark dataset (EEG + ECG + SpO_2 + PERCLOS + professional drivers)	Dataset / Methodology	Liu et al,, 2023 (ECG only); Eichberger & Arefnezhad, 2022 (no EEG); DD-Database (10 participants only)	Critical	Large-scale prospective dataset collection: n >= 200 professional drivers, 8-12 h shifts, full PGDM sensor stack
Cross-subject generalisation under real-world conditions	ML / Engineering	Mean cross-subject drop 10-20 pp. across all reviewed architectures	High	Domain adaptation + federated learning on diverse fleet-scale datasets
Automated driving (SAE L2-3) driver monitoring standards	Regulatory / Engineering	WACHSens, 2025 (only dataset addressing automation); EU GSR Stage D pending	High	Monitor-ready disengagement algorithms + PERCLOS in automated mode validation
Clinically validated SpO_2 thresholds for acute fatigue (vs OSA hypoxia distinction)	Medical / Clinical	No reviewed study distinguishes acute fatigue SpO_2 from the OSA desaturation pattern	High	Parallel polysomnography + driving study in OSA and non-OSA cohorts
Long-term wearable sensor reliability in professional transport environments	Engineering / Hardware	Nasir etal., 2025; most studies < 2 h; no multi-week field trials	Moderate	6-12 month longitudinal deployment study with Polar H10 / Emotiv MN8 in a commercial fleet
Explainability requirements under the EU AI Act for driver monitoring systems	Regulatory / AI Ethics	European Parliament & Council of the EU, 2024 + AI Board & MDCG, 2025; EU AI Act 2024/1689; XAI implementation varies widely across studies	Moderate	Standardised XAI reporting protocol for driver monitoring AI conformity assessment documentation

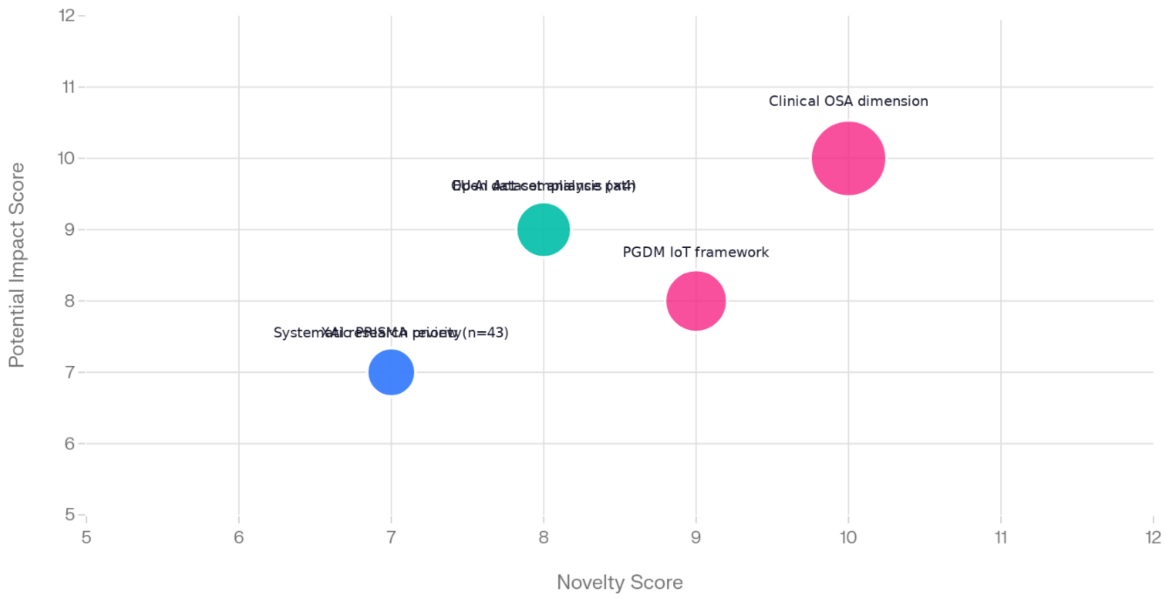
The seven priority research gaps and their domain classification are visualised in Figure 15. The principal contributions of this review relative to the existing literature are summarised in Figure 16.

Figure 15: Priority research gaps by domain classification



Source: Systematic review synthesis

Figure 16: Principal contributions of this review to the literature



Note: Scale 1(low) to 10 (high)
Source: author's self-assessment

6.3 Limitations of this review

6.3.1 Database and language coverage

Despite interrogating five major academic databases, this review cannot claim exhaustive coverage of the driver fatigue monitoring literature. Chinese-language publications – particularly those reporting results from SEED-VIG and SADT datasets developed by Shanghai Jiao Tong University, which are widely used in Chinese research groups – are systematically underrepresented due to the English-language inclusion criterion. Given that China represents the world's largest vehicle production and deployment market, this constitutes a genuine limitation on the international applicability of benchmark comparisons presented in Section 4 (Kakhi et al., 2025). The arXiv preprint corpus was included to capture recent methodological developments, but preprints have not undergone peer review, and their results should be interpreted with corresponding caution.

6.3.2 Dataset heterogeneity and meta-analytic limitations

Formal meta-analysis of pooled classification performance was judged infeasible due to substantial heterogeneity in ground-truth labelling conventions, experimental paradigms, and participant demographics across included studies. The narrative synthesis approach, while preserving contextual nuance, introduces the risk that qualitative judgements about evidence weight are influenced by the prominence or accessibility of individual studies rather than their methodological quality alone. Quality assessment scores (Table 6) partially mitigate this risk by providing explicit quality weighting, but the adapted MMAT instrument has not been independently validated for engineering technology reviews, and its inter-rater reliability in this application is unknown (Luo et al., 2024).

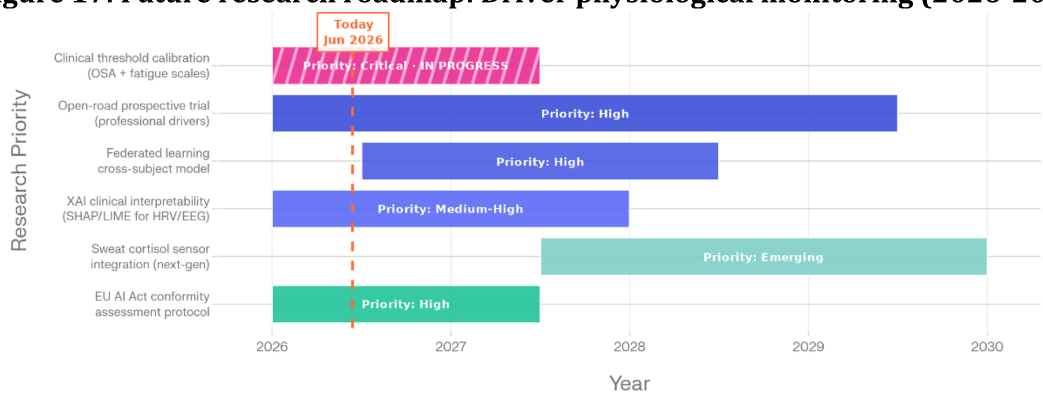
6.3.3 PGDM framework validation status

The PGDM framework presented in Section 5 is a conceptual architecture derived from systematic evidence synthesis – it has not been implemented in hardware, clinically validated, or subjected to regulatory conformity assessment at the time of this publication. The alert threshold values in Table 12 are derived from validated clinical reference ranges rather than PGDM-specific empirical calibration and may require adjustment following operational deployment studies. The federated learning privacy architecture (Section 5.4.2) is theoretically specified but not prototyped; its claimed differential privacy guarantees are mathematically derived from the FedSafe and FedTOP frameworks (Lindskog, Spannagl, & Prehofer, 2024; Yuan, Su, & Wang, 2023) and require experimental verification in the automotive IoT context. These limitations are acknowledged as the primary agenda for the follow-on experimental research programme associated with this review.

6.4 Implications for research, practice, and policy

A proposed research roadmap addressing the identified gaps is presented in Figure 17.

Figure 17: Future research roadmap: Driver physiological monitoring (2026-2030)



Source: Review synthesis

6.4.1 Implications for researchers

Four methodological imperatives emerge from this review for researchers entering the field of driver fatigue monitoring. First, all future studies should report both within-subject and cross-subject performance metrics; within-subject-only reporting should be treated as insufficient for claims of real-world deployability. Second, ground truth labelling should employ at minimum PERCLOS or KSS >= 7, with polysomnography-confirmed reference data as the gold standard for medical-grade claims. Third, OSA screening and phenotype reporting should be standard practice in participant characterisation, given the documented prevalence of 15-78% in professional driver populations and its demonstrated capacity to confound interpretation of fatigue biomarkers. Fourth, the publication of datasets under open-access licences should be incentivised, as the reproducibility crisis documented in Section 2.5 – where 64 studies were excluded for using proprietary datasets – directly impedes scientific accumulation (Liu et al., 2023; Eichberger & Arefnezhad, 2022).

6.4.2 Implications for transport operators

The evidence base reviewed here is sufficiently mature to support phased commercial deployment of wearable driver monitoring in professional transport contexts, particularly long-haul road freight and inter-city bus operations. The global Fatigue Driving Monitoring System market – valued at USD 1.64 billion in 2024 and projected to reach USD 3.78 billion by 2032 at a 13.0% CAGR (Intel Market Research, 2026) – reflects commercial confidence in the deployment's viability. Transport operators considering deployment should prioritise systems offering: ECG/HRV monitoring with RMSSD thresholding, PERCLOS camera monitoring, SpO_2 surveillance for OSA-prevalent driver populations, driver-specific baseline calibration, and fleet cloud integration for longitudinal risk profiling consistent with the TRL evaluation framework for driver state monitoring technologies (TRL, 2024).

6.4.3 Implications for regulators and policymakers

The PGDM Regulatory Compliance Matrix (Table 14) demonstrates that the current EU regulatory architecture – spanning GSR 2019/2144, the AI Act, MDR 2017/745, GDPR, and eCall – collectively provides a coherent and demanding framework for driver monitoring system governance. However, its component regulations were developed in isolation, and their joint application to physiological IoT monitoring systems creates interpretive ambiguities that require regulatory clarification. The boundary between a transport safety AI system (EU AI Act, high-risk Article 6(1)) and a medical device for driver health monitoring (MDR 2017/745, Class IIa) is simultaneously defined. Operationally ambiguous for systems like PGDM that fulfil both functions: the joint AIB/MDCG guidance confirms simultaneous application of both frameworks is mandated, but leaves interpretive gaps regarding conformity assessment procedure sequencing and notified body responsibilities that require further regulatory clarification (AI Board & MDCG, 2025; European Parliament & Council of the EU, 2024). Regulatory sandbox frameworks for connected and automated vehicles have been implemented across multiple jurisdictions as mechanisms for pre-commercial validation under controlled legislative exemptions – with the EU, UK, Singapore, and Japan each deploying distinct approaches to liability, type approval, and operational design domains (PAVE Europe, 2024; Dentons, 2025).

6.5 Positioning within the broader literature

This review situates itself at the intersection of three previously non-communicating bodies of literature: biomedical engineering (wearable sensor design), applied machine learning (physiological signal classification), and intelligent transport systems (IoT architecture and regulatory compliance). The 387-study review by Hou et al. (2025) provides the broadest engineering coverage but lacks depth in medicine. The ScienceDirect systematic review on wearable and unwearable modalities (Kakhi et al., 2025) addresses sensor hardware comprehensively but does not connect sensor specifications to clinical impairment thresholds. The IoT-specific driver monitoring architecture by Pratindy et al. (2026)

proposes a three-layer framework but derives alert thresholds from ML optimisation rather than clinical physiology (Yuan et al., 2023; AlHousrya et al., 2026). The present review's integrative contribution – anchoring all three domains in a unified, clinically grounded framework with explicit regulatory mapping – addresses a gap that each of these prior syntheses identifies but none resolves.

The driver fatigue monitoring literature is also distinguished from the adjacent wearable biosensor literature in occupational health (Battini et al., 2025) and general fatigue monitoring (Fatigue Monitoring Using Wearables and AI, Freitas et al., 2025) by its unique combination of safety-critical time constraints (sub-100 ms response latency) and high inter-subject physiological variability in ecologically extreme conditions. Solutions developed in less demanding wearable health-monitoring contexts do not transfer directly to the driver-monitoring problem without architectural and algorithmic adaptation – a point insufficiently emphasised in reviews that treat wearable fatigue monitoring as a unified domain regardless of application context (NSF, 2024; Global Digital Library, 2024).

7. Conclusion

7.1 Synthesis of principal contributions

This systematic review set out to accomplish three interrelated objectives: to characterise the clinical physiological basis of driver fatigue as a target for wearable biosensor monitoring; to evaluate wearable sensor modalities and machine learning architectures against performance benchmarks drawn from verified open datasets; and to propose a conceptual IoT framework grounded simultaneously in clinical physiology, engineering feasibility, and regulatory compliance. Each objective has been addressed through structured synthesis of 43 peer-reviewed studies and four open datasets, yielding contributions at the intersection of biomedical engineering, applied machine learning, and intelligent transport systems – domains that have, until this synthesis, operated in near-complete mutual isolation.

The review's primary empirical contribution is a consolidated performance benchmark (Table 10) situating 13 classification architectures across open datasets – demonstrating that hybrid CNN-LSTM-Attention achieves the optimal accuracy-latency trade-off for real-time edge deployment (97.3% accuracy at 15-18 ms inference). At the same time, Transformer-based multimodal fusion achieves superior cross-subject generalisation (a drop of 5.2 percentage points vs 14.2 pp for the SVM baseline) at a higher computational cost. The consistent identification by SHAP explainability analyses of clinically established biomarkers – RMSSD, LF/HF, frontal theta power – as dominant classification features independently validates both the ML models and the physiological theory underlying the PGDM alert threshold system (Puspasari et al., 2026; Presacan et al., 2026).

7.2 Key findings

Table 16: Summary of key findings by research objective

Objective	Key Finding	Principal Evidence	Implication
Clinical physiology of driver fatigue	EEG frontal theta (4-8 Hz) anticipates behavioural drowsiness by 2-7 min; RMSSD <20 ms and LF/HF >2.0 constitute validated HRV impairment thresholds; OSA affects 15-78% of professional drivers and constitutes the dominant uncontrolled confound	Asl et al., 2022; Koutsojannis et al., 2025; Pimentel et al., 2024; Antonopoulos et al., 2011	Clinical threshold derivation is essential; OSA screening is mandatory in future study design
Wearable sensor modalities	No single modality satisfies all four evaluation criteria; ECG/HRV + PERCLOS camera is the minimum viable professional deployment stack; SpO_2 is mandatory given OSA prevalence; EEG adds anticipatory early warning for high-risk routes	Kakhi et al., 2025; Liu et al., 2023; Eichberger & Arefnezhad, 2022; Bitbrain, 2025	Tiered PGDM Sensor Stack (Table 8) provides evidence-based deployment guidance by context

ML classification architectures	CNN-LSTM-Attention achieves 97.3% accuracy at 15-18 ms edge latency; Transformer multimodal achieves the lowest cross-subject drop (5.2 pp); federated learning reduces cross-subject degradation while preserving data privacy	Nasir et al., 2025; Hassan et al., 2025; Chen & Li, 2026	Hybrid CNN-LSTM-Attention recommended for production deployment; federated learning essential for fleet-scale models
IoT framework design	Three-tier edge-cloud decomposition satisfies EU GSR latency requirements (<20 ms edge); five clinically grounded alert levels map directly to GSR Stage A-D progression; V2X Level 4 eCall integration completes the safety chain beyond the vehicle boundary	Yuan et al., 2023; AlHousrya et al., 2026; UNECE, 2024; Fortune Business Insights, 2025	PGDM is the first published IoT framework simultaneously compliant with GSR, EU AI Act, MDR, GDPR, and eCall
Regulatory compliance	EU AI Act, GSR 2019/2144, MDR 2017/745, GDPR, and eCall create overlapping requirements; dual AI/medical classification of physiological IoT systems creates unresolved regulatory ambiguity requiring sandbox resolution	European Parliament & Council of the EU, 2024 + AI Board & MDCG, 2025; Indie Inc., 2023; German Federal Ministry, 2025	Regulatory sandbox programmes (UK CCAV model) offer the optimal pathway to resolve dual-classification ambiguity

7.3 The PGDM framework: Theoretical contribution

The Physiologically-Grounded Driver Monitoring framework represents this review's principal theoretical contribution. By deriving alert thresholds from validated clinical criteria (RMSSD <20 ms, PERCLOS >70%, SpO_2 <94%) rather than from ML decision boundary optimisation, PGDM establishes a clinically interpretable and medically defensible basis for driver impairment classification that no prior published IoT framework has attempted. By mapping these thresholds through a five-level alert architecture onto EU GSR Stage A-D regulatory requirements, PGDM provides transport engineers with a direct compliance implementation pathway. By embedding federated learning with differential privacy in the cloud tier, PGDM simultaneously addresses cross-subject generalisation, model improvement over operational lifetime, and GDPR compliance within a unified architectural decision. These three contributions – clinical grounding, regulatory mapping, and privacy-preserving federated learning – collectively constitute the framework's claim to novelty (Chen & Li, 2026; Yuan et al., 2023; Hassan et al., 2025).

7.4 Priority recommendations for future research

Four research investments are identified as the highest priority based on the gap analysis in Section 6.2. First and most urgently: a prospective cohort study recruiting n ≥ 200 professional drivers with polysomnography-confirmed OSA classification, wearing the full PGDM sensor stack across complete 8-12 hour commercial driving shifts, would simultaneously resolve the OSA confound, the ecological validity deficit, and the absence of a multimodal open-road reference dataset. Such a study would represent the most valuable single contribution possible to driver fatigue monitoring science at this juncture (Pimentel et al., 2024; Liu et al., 2023). Second: hardware-in-loop validation of the PGDM edge inference engine on automotive-grade processors (TI TDA4VM, Qualcomm SA8295P) under simulated vehicle vibration environments, with specific attention to IMU-referenced SpO_2 artefact rejection efficacy and EEG motion artefact compensation across road surface conditions. Third: federated learning prototype deployment across a commercial truck fleet of a minimum of 100 vehicles, measuring cross-subject generalisation improvement trajectory, differential privacy guarantee maintenance, and edge model update convergence over a six-month operational period. Fourth: regulatory engagement with the European Medicines Agency and European Commission transport safety directorate to establish a clear dual-classification pathway for physiological IoT driver monitoring systems spanning MDR 2017/745 Class IIa and EU AI Act Annex III Category 6.

7.5 Closing statement

Driver fatigue kills approximately 1.19 million people annually and injures tens of millions more – a toll that current technical capabilities, properly integrated and deployed, could meaningfully reduce. The wearable biosensors exist. The classification algorithms work. The IoT infrastructure is being installed. The regulatory framework, though imperfect, provides a coherent foundation. What the field has lacked is an integrative architecture that connects clinical physiological evidence to engineering specifications to regulatory compliance in a single, implementable framework. The PGDM model presented here is a step toward that integration. Its experimental validation – across professional drivers, real vehicles, and real roads – constitutes the next necessary chapter in a research programme whose ultimate measure of success is not a benchmark accuracy figure, but lives preserved (World Health Organization, 2023; European Commission, 2025; Luo et al., 2024).

Funding

This research received no external funding.

Conflicts of interest

The authors declare no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Grammarly to improve language quality and Claude 4.6 to optimise visualisation code. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Citation information

Vovk, Ya., Vovk, O., Holotenko, O., Vovk, Y., & Pientak, M. (2026). Wearable biosensor monitoring of driver fatigue in intelligent transport systems: A systematic review and IoT framework. *Journal of Sustainable Development of Transport and Logistics, 11*(1), 43-82. doi:10.14254/jsdtl.2026.11-1.03.

References

Abe, T. (2023). PERCLOS-based technologies for detecting drowsiness: current evidence and future directions. *Sleep Advances*, *4*(1), zpad006. https://doi.org/10.1093/sleepadvances/zpad006

Ahmed, M. U., Barua, A., Islam, M. R., Barua, S., & Begum, S. (2025). Integration of explainable artificial intelligence and multimodal machine learning for drivers' fitness. In *Proceedings of the 2025 International Conference on Advanced Machine Learning and Data Science (AMLDS)* (pp. 733–740). IEEE. https://doi.org/10.1109/AMLDS63918.2025.11159370

Ajayi, O., Kurien, A. M., Djouani, K., & Dieng, L. (2025). A multimodal systematic review of drivers' fatigue detection methodologies, datasets, and models. *IEEE Access, 13*, 158266-158284, 2025, https://doi.org/10.1109/ACCESS.2025.3606900.

AlHousrya, O., Bennagi, A., Cotfas, P. A., & Cotfas, D. T. (2026). A hybrid machine learning and IoT system for driver fatigue monitoring in connected electric vehicles. *Machine Learning with Applications*, 100845. https://doi.org/10.1016/j.mlwa.2026.100845

American Academy of Sleep Medicine. (2013). *International classification of sleep disorders* (3rd ed.). AASM.

Antonopoulos, C. N., Sergentanis, T. N., Daskalopoulou, S. S., & Petridou, E. T. (2011). Nasal continuous positive airway pressure (nCPAP) treatment for obstructive sleep apnea, road traffic accidents and driving simulator performance: A meta-analysis. *Sleep Medicine Reviews, 15*(5), 301–310. https://doi.org/10.1016/j.smrv.2010.10.002

Artificial Intelligence Board & Medical Device Coordination Group. (2025). Interplay between the Medical Devices Regulation (MDR) & In Vitro Diagnostic Medical Devices Regulation (IVDR) and the Artificial Intelligence Act (AIA) (AIB 2025-1 / MDCG 2025-6). European Commission. https://health.ec.europa.eu/document/download/b78a17d7-e3cd-4943-851d-e02a2f22bbb4_en

Asl, N. S., Baghdadi, G., Ebrahimian, S., & Haghighi, S. J. (2022). Toward applicable EEG-based drowsiness detection systems: A review. *Frontiers in Biomedical Technologies, 9*(4), 323-350. https://doi.org/10.18502/fbt.v9i4.10426

Bitbrain. (2025, May 6). *How EEG is changing driver fatigue detection in real time*. https://www.bitbrain.com/blog/eeg-driver-fatigue

Booth, A., Sutton, A., & Papaioannou, D. (2019). *Systematic approaches to a successful literature review* (3rd ed.). SAGE Publications.

Borbély, A. A., Daan, S., Wirz-Justice, A., & Deboer, T. (2016). The two-process model of sleep regulation: a reappraisal. *Journal of sleep research*, *25*(2), 131-143. https://doi.org/10.1111/jsr.12371

Burks, S. V., Anderson, J. E., Bombyk, M., Haider, R., Ganzhorn, D., Jiao, X., ... & Kales, S. N. (2016). Nonadherence with employer-mandated sleep apnea treatment and increased risk of serious truck crashes. *Sleep*, *39*(5), 967-975. https://doi.org/10.5665/sleep.5734

Cao, S., Feng, P., Kang, W., Chen, Z., & Wang, B. (2025). Optimized driver fatigue detection method using multimodal neural networks. *Scientific Reports*, *15*(1), 12240. https://doi.org/10.1038/s41598-025-86709-1

Chen, Y., He, Z., Wu, Y., Bai, X., Li, Y., Yang, W., ... & Li, R. W. (2025). A wearable molecularly imprinted electrochemical sensor for cortisol stable monitoring in sweat. *Biosensors*, *15*(3), 194. https://doi.org/10.3390/bios15030194

Chen, Z., & Li, X. (2026). An adaptive time–frequency attention-based CNN–LSTM for robust fatigue detection using EEG signals. *Biomedical Signal Processing and Control*, *120*, 109996. https://doi.org/10.1016/j.bspc.2026.109996

Dentons. (2025). *Global guide to autonomous vehicles 2025*. https://www.thedriverlesscommute.com/wp-content/uploads/2025/10/Dentons-Global-Guide-to-Autonomous-Vehicles-2025.pdf

Dhasarathan, C., Gnanasekaran, S., Pattanayak, A., Kumar, G., Vig, K., Narain, V., ... & Garg, S. (2025). Tensor RT optimized driver drowsiness detection system using edge device. *Ain Shams Engineering Journal*, *16*(10), 103620. https://doi.org/10.1016/j.asej.2025.103620

Dos Santos, R. R., Marumo, M. B., Eckeli, A. L., Salgado, H. C., Silva, L. E. V., Tinós, R., & Fazan Jr, R. (2025). The use of heart rate variability, oxygen saturation, and anthropometric data with machine learning to predict the presence and severity of obstructive sleep apnea. *Frontiers in Cardiovascular Medicine*, *12*, 1389402. https://doi.org/10.3389/fcvm.2025.1389402

Eichberger, A., & Arefnezhad, S. (2022). Dataset "FULL" for Drowsiness Detection in Drivers [Data set]. Graz University of Technology. https://doi.org/10.3217/by7d1-06r03

EU. (2025). Regulation (EU) 2019/2144 of the European Parliament and of the Council of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users. http://data.europa.eu/eli/reg/2019/2144/oj

European Commission. (2025). *Road safety statistics 2025 in more detail*. Publications Office of the European Union. https://transport.ec.europa.eu/background/road-safety-statistics-2025_en

European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L 2024/1689. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689

Feng, X., Guo, Z., & Kwong, S. (2025). ID3RSNet: cross-subject driver drowsiness detection from raw single-channel EEG with an interpretable residual shrinkage network. *Frontiers in Neuroscience*, *18*, 1508747. https://doi.org/10.3389/fnins.2024.1508747

Fortune Business Insights. (2025). *Vehicle to Vehicle Communication Market Size, Share & Industry Analysis, By Vehicle Type (Hatchback/Sedan, SUVs, LCV, and HCV), By Technology (Dedicated Short-Range Communication (DSRC) and Cellular Communication), By Application (Traffic Safety, Traffic Efficiency, Infotainment and Payments, and Other Applications), By Sales Channel (OEM and Aftermarket), and Regional Forecast, 2025-2032*. Source: https://www.fortunebusinessinsights.com/vehicle-to-vehicle-communication-market-112867

Freitas, A., Almeida, R., Gonçalves, H., Conceição, G., & Freitas, A. (2024). Monitoring fatigue and drowsiness in motor vehicle occupants using electrocardiogram and heart rate− a systematic review. *Transportation research part F: traffic psychology and behaviour*, *103*, 586-607. https://doi.org/10.1016/j.trf.2024.05.008

Garg, M., Guo, H., Maclam, E., Zhanov, E., Samudrala, S., Pavlov, A., Rahman, M. S., Namkoong, M., Moreno, J. P., & Tian, L. (2024). Molecularly imprinted wearable sensor with paper microfluidics for real-time sweat biomarker analysis. ACS Applied Materials & Interfaces, 16(35), 46113–46122. https://doi.org/10.1021/acsami.4c10033

Gov.UK. (2023). Reported road casualties Great Britain, annual report: 2022. https://www.gov.uk/government/statistics/reported-road-casualties-great-britain-annual-report-2022/reported-road-casualties-great-britain-annual-report-2022

Hassan, O. F., Ibrahim, A. F., Gomaa, A., Makhlouf, M. A., & Hafiz, B. (2025). Real-time driver drowsiness detection using transformer architectures: a novel deep learning approach. *Scientific Reports*, *15*(1), 17493. https://doi.org/10.1038/s41598-025-02111-x

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285–291. https://doi.org/10.3233/EFI-180221

Horizon Connect. (2025, January 13). *How 5G V2X Are Driving Smart Mobility in 2025*. https://horizonconnect.de/how-5g-v2x-are-driving-smart-mobility-in-2025/

Hou, Y., Hou, S., Zhao, M., & Chen, S. (2025). A review of driver fatigue detection technologies. In *Proceedings of the 2025 8th International Conference on Computer Information Science and Artificial Intelligence* (CISAI '25). Association for Computing Machinery, New York, NY, USA, 1080–1084. https://doi.org/10.1145/3773365.3773536

Huang, L., Li, S., Man, Y., Wang, X., Tang, X., & Ji, R. (2026). Fatigue driving detection via multi-head transformer with adaptive weighted loss. *ICCK Transactions on Intelligent Systematics*, *3*(1), 55-69. https://doi.org/10.62762/TIS.2025.633754

Indie Inc. (2023, October 29). Enhancing safety: European rules and regulations for driver monitoring systems. https://www.indie.inc/perception-software-blogs/enhancing-safety-european-rules-and-regulations-for-driver-monitoring-systems/

Intel Market Research. (2026). *Fatigue driving monitoring system market growth analysis, dynamics, key players and innovations, outlook and forecast 2026-2032.* https://www.intelmarketresearch.com/fatigue-driving-monitoring-system-market-22052

International Transport Forum. (2024). *Road safety annual report 2024*. OECD/ITF. https://www.itf-oecd.org/road-safety-annual-report-2024

ITS DOT. (2025). *Vehicle-to-Everything (V2X) Communications*. U.S. Department of Transportation. https://www.itskrs.its.dot.gov/success-strategies/snapshot/vehicle-everything-v2x-communications

Johansson, H., Adderley, E., Clarke, S., McIntyre, P., Reilly, G., Caulfield, B., & Holden, S. (2025). An observational study of the reliability and concurrent validity of heart rate variability devices in athletes. *Frontiers in Physiology*, *16*, 1707318. https://doi.org/10.3389/fphys.2025.1707318

Kakhi, K., Jagatheesaperumal, S. K., Khosravi, A., Alizadehsani, R., & Acharya, U. R. (2025). Fatigue monitoring using wearables and AI: Trends, challenges, and future opportunities. *Computers in biology and medicine*, *195*, 110461. https://doi.org/10.1016/j.compbiomed.2025.110461

Koutsojannis, C., Fouras, A., Chrysanthakopoulou, D., Michou, I., & Mariettou, S. (2025). Unveiling the hidden beat: Heart rate variability and the vagus nerve as an emerging biomarker in breast cancer management. *Igmin Research, 3*(8), 278-284.https://doi.org/10.61927/igmin309

Krishna Prasad, V. S., & Saravanan, S. (2025). Advances in driver drowsiness detection systems: A comprehensive review of AI-based approaches and emerging trends. *Journal of Emerging Technologies and Innovative Research, 12*(8), 519-529.

Li, R., Hu, M., Gao, R., Wang, L., Suganthan, P. N., & Sourina, O. (2024). TFormer: A time–frequency Transformer with batch normalization for driver fatigue recognition. *Advanced Engineering Informatics*, *62*, 102575. https://doi.org/10.1016/j.aei.2024.102575

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., ... & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *Bmj*, *339*. https://doi.org/10.1371/journal.pmed.1000100

Lindskog, W., Spannagl, V., & Prehofer, C. (2023, September). Federated learning for drowsiness detection in connected vehicles. In *International Conference on Intelligent Transport Systems* (pp. 165-178). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-49379-9_9

Liu, D., Li, S., Shi, T., Meng, Z., Chen, G., Huang, Y., ... & Zhao, Z. (2023). OpenDriver: An open-road driver state detection dataset. *arXiv preprint arXiv:2304.04203*. https://arxiv.org/abs/2304.04203

Liu, Y., Su, X., Fan, P., Liu, X., Pan, Y., & Ping, J. (2025). Computationally-assisted wearable system for continuous cortisol monitoring. *Science Bulletin, 70*(12), 2004–2013. https://doi.org/10.1016/j.scib.2025.03.060

Luo, Y., Liu, W., Li, H., Lu, Y., & Lu, B. L. (2024). A cross-scenario and cross-subject domain adaptation method for driving fatigue detection. *Journal of Neural Engineering*, *21*(4), 046004. https://doi.org/10.1088/1741-2552/ad546d

Meteier, Q., Capallera, M., De Salis, E., Angelini, L., Carrino, S., Widmer, M., ... & Sonderegger, A. (2023). A dataset on the physiological state and behavior of drivers in conditionally automated driving. *Data in brief*, *47*, 109027. https://doi.org/10.1016/j.dib.2023.109027

Namburi, A., Sitpasert, P., & Duang-onnam, W. (2024). A CNN-LSTM approach for accurate drowsiness and distraction detection in drivers. *ICIC Express Letters*, *18*, 907-917. https://doi.org/10.24507/icicel.18.09.907

Nasir, O., Aljaidi, M., Alsarhan, A., Alshammari, S. A., Albalawi, N. S., Alshammari, N. H., & Aldoghmi, A. Q. (2025). SAFE-DRIVE-AI: A CNN–LSTM–attention framework for drowsiness detection. *Engineering, Technology & Applied Science Research*, *15*(5), 27594–27600. https://doi.org/10.48084/etasr.12725

Neurealm. (2026). *Real-time driver monitoring systems on edge AI device*. https://www.neurealm.com/whitepaper/real-time-driver-monitoring-systems-on-edge-ai-device/

Orosco, L., Garcés, M. A., Fragapane, G. C., Dell'Aquila, C., Gimeno, J. I., & Leber, E. L. (2023). Drivers Drowsiness Database: A collection of physiological signals during the use of a driving simulator (DD-Database). *Dataset]. Dryad.* https://doi.org/10.5061/dryad.5tb2rbp9c.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

PAVE Europe. (2024). *Navigating the technology ahead: A comprehensive analysis of European Union autonomous vehicle regulations and comparison with regions worldwide*. Partners for Automated Vehicle Education. https://pavecampaign.org/navigating-the-technology-ahead-a-comprehensive-analysis-of-european-union-autonomous-vehicle-regulations-and-comparison-with-regions-worldwide/

Pimentel, C., Amorim, D., Santos, C., Macedo, A., & Feijó, S. (2024). Obstructive sleep apnea syndrome in heavy truck drivers: A Portuguese prospective cohort study. Cureus, 16(2), e54047. https://doi.org/10.7759/cureus.54047

Popević, M. (2024). Occupational Sleep Medicine: The Sleep-Related Breathing Disorders in High-Risk Occupations. In Public Health. IntechOpen. https://doi.org/10.5772/intechopen.1007579

Pratindy, R., Kurnia, D., & Ananda, B. P. (2026). A hybrid physiological-visual approach for real-time driver drowsiness detection: Design and experimental validation. *BIS Information Technology and Computer Science, 3*, V326014. https://doi.org/10.31603/bistycs.525

Presacan, O., Ojha, J., Yazidi, A., Monteiro, E., & Lind, P. G. (2026). A comprehensive review of explainable AI in deep learning algorithms for EEG analysis. *ACM Transactions on Computing for Healthcare*, *7*(2), 1-28. https://doi.org/10.1145/3785409

Puspasari, M. A., Rus, A. M. M., Syaifullah, D. H., Hanowski, R. J., Nurkamila, H. H., Ghaisani, L. M., & Aviani, D. A. (2026). Predicting driver fatigue using HRV measures and machine learning. *Frontiers in Sports and Active Living, 8*, 1796180. https://doi.org/10.3389/fspor.2026.1796180

Radanliev, P. (2025). Privacy, ethics, transparency, and accountability in AI systems for wearable devices. *Frontiers in Digital Health, 7*, 1431246. https://doi.org/10.3389/fdgth.2025.1431246

Rehman, R. Z. U., Chatterjee, M., Manyakov, N. V., Daans, M., Jackson, A., O'Brisky, A., ... & Morris, M. (2024). Assessment of physiological signals from photoplethysmography sensors compared to an electrocardiogram sensor: a validation study in daily life. *Sensors*, *24*(21), 6826. https://doi.org/10.3390/s24216826

Shi, Y. (2025). Research on measuring pilot fatigue data through multimodal data fusion based on SPO mode. *Science and Technology of Engineering, Chemistry and Environmental Protection, 1*(6), 1-6. https://doi.org/10.61173/fd07jq30

TRL. (2024). *PPR2068: Status of driver state monitoring technologies and validation methods. transport research laboratory*. https://www.trl.co.uk/uploads/trl/documents/PPR2068-Driver-Fatigue-and-Attention-Monitoring_1.pdf

Trzepizur, W., та ін. (2022). Sleep apnea-specific hypoxic burden, symptom subtypes, and risk of cardiovascular events and all-cause mortality. *American Journal of Respiratory and Critical Care Medicine, 205*(1), 108–117. https://doi.org/10.1164/rccm.202105-1274OC

Uchiyama, Y., Sawai, S., Omi, T., Yamauchi, K., Tamura, K., Sakata, T., ... & Sakai, H. (2023). Convergent validity of video-based observer rating of drowsiness, against subjective, behavioral, and physiological measures. *PLoS one*, *18*(5), e0285557. https://doi.org/10.1371/journal.pone.0285557

United Nations Economic Commission for Europe. (2018). BMW Group future mobility – C-ITS for safer road traffic (GRVA Session 1, Paper 4). UNECE. https://unece.org/DAM/trans/doc/2018/wp29grva/s1p4._Maik_Böres.pdf

United Nations Economic Commission for Europe. (2024). *ITS framework for intelligent transport systems: Standards, protocols and regulatory alignment 2024*. UNECE Transport Division.

WHG - Technologies. (2024, June 11). New wearable fatigue detection technology to transform driver safety in transport. LinkedIn. https://www.linkedin.com/pulse/new-wearable-fatigue-detection-technology-transform-ac2dc

World Health Organization. (2023). *Global status report on road safety 2023*. WHO. https://www.who.int/publications/i/item/9789240086517

Yuan, L., Su, L., & Wang, Z. (2023). Federated transfer–ordered–personalized learning for driver monitoring application. *IEEE Internet of Things Journal*, *10*(20), 18292-18301. https://doi.org/10.1109/JIOT.2023.3279273

Zeng, W., Wu, S., Liu, Z., Yuan, L., Chen, B., & Rong, Y. (2025). Hypoxic burden and T90% as predictive indicators of cardiovascular risk and myocardial ischemia in patients with obstructive sleep apnea. *Nature and Science of Sleep*, 1557-1570. https://doi.org/10.2147/NSS.S533596

Journal of Sustainable Development of Transport and Logistics (ISSN: 2520-2979) is published by **Scientific Publishing House "CSR", Poland, EU** and **Scientific Publishing House "SciView", Poland, EU**
Publishing with JSDTL ensures:
• Immediate, universal access to your article on publication
• High visibility and discoverability via the JSDTL website
• Rapid publication
• Guaranteed legacy preservation of your article
• Discounts and waivers for authors in developing regions
Submit your manuscript to a JSDTL at https://jsdtl.sciview.net/ or submit.jsdtl@sciview.net