Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 8 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i08.59843

PAPER

AI-Based Mobile Coaching Architectures for Driver Behavior and Road Safety: A Systematic Mapping and Design Framework

Manuel Hilario[1] (✉),
Pervis Paredes[1],
Henry Maquera[2],
Shirley Martínez[1],
Jhosep Velasque[1]

[1]Federico Villarreal National University, Lima, Peru

[2]National University of Central Peru, Huancayo, Peru

fhilario@unfv.edu.pe

ABSTRACT

A systematic mapping of artificial intelligence (AI)-based mobile driver coaching for behavior change and road safety in land transportation has been conducted following PRISMA, resulting in nine empirical studies (n = 9). The evidence is organized through a taxonomy covering intervention goals, AI techniques, data sources, deployment patterns, outcome and engagement metrics, and privacy/dependability safeguards. Across the mapped interventions, a substantial share targets specific risky behaviors, particularly phone-related distraction, reporting relative reductions of up to 21% versus controls in controlled evaluations. Deployment designs are predominantly hybrid, combining on-device inference for timely detection with server-side components for aggregation and program features. Based on the extracted evidence, a system-oriented design framework and practical guidelines are proposed to support researchers and mobile system designers in selecting architectures and evaluation metrics for interactive driver coaching. Key limitations include heterogeneous outcome definitions and predominantly short- to medium-term follow-up, which constrain comparability and generalization across settings.

KEYWORDS

artificial intelligence (AI), mobile driver coaching, road safety, telematics, dependability

1 INTRODUCTION

Road safety and the operational performance of land transport remain shared priorities for the public and private sectors; however, evidence shows that traffic accidents and their collateral effects, such as mortality, injuries, costs, and loss of productivity, persist despite efforts in regulation, control, and infrastructure [1]. Furthermore, various reports rank them among the leading causes of death and years of healthy life lost worldwide, with particularly severe impacts in low- and middle-income countries, where more than 90% of deaths occur even though they

Hilario, M., Paredes, P., Maquera, H., Martínez, S., Velasque, J. (2026). AI-Based Mobile Coaching Architectures for Driver Behavior and Road Safety: A Systematic Mapping and Design Framework. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(8), pp. 149–171. https://doi.org/10.3991/ijim.v20i08.59843

Article submitted 2025-11-28. Revision uploaded 2026-01-27. Final acceptance 2026-03-03.

account for a smaller share of the vehicle fleet [2]. In this context, coaching drivers, managers, and users through mobile technologies continues to play a strategic role because it can reduce risky behavior, improve driving decisions, and accelerate the adoption of good practices in everyday life [3]. However, its effectiveness increasingly depends on personalizing content. Among young drivers, feedback based on events detected by smartphone sensors and progressive goals shows decreases in risky behavior by providing timely feedback and linking learning to real-world operating contexts, i.e., to what actually happens on the road [4]. Therefore, this work focuses on artificial intelligence (AI)-based mobile coaching for driver behavior change in land transportation, understood as interactive driver-coaching systems that, through brief sessions and post-trip feedback, seek to modify behaviors and consolidate safe driving practices in authentic operating contexts. It also explicitly establishes the difference between safety or telematics solutions aimed at detection or surveillance without behavioral or engagement metrics and those mobile coaching interventions that do report measurable behavior change and signs of adoption.

Artificial intelligence deployed on mobile platforms offers a concrete pathway to personalization at scale. Smartphones, by virtue of their ubiquity, integrated sensors, and connectivity, make it possible to record driving patterns (acceleration, braking, speed, trajectories) and contextual signals (time, weather, road type, traffic density). With this, it is possible to build dynamic profiles and deliver situated recommendations before, during, and after the trip [5].

This technological promise imposes strict requirements for design, safety, and privacy. In terms of user experience and safety of use, any mobile driver-coaching application should be safety-by-design: avoiding distraction while the vehicle is in motion, reducing manual interaction, prioritizing deferred feedback, and ensuring accessibility for diverse user profiles [6]. In terms of privacy, the processing of behavioral and performance data requires principles of minimization, transparency, access control, and traceability [7]. In terms of dependability, the platform should sustain inference latencies compatible with the delivery of useful feedback, operate with heterogeneous connectivity (including offline conditions), and manage energy consumption so as not to erode sustained adoption [8].

This study maps and synthesizes AI-based mobile driver coaching for behavior change and road safety in ground transportation and proposes an analytical framework to structure the evidence. To this end, the framework integrates three axes: (i) intervention-goal and design taxonomy; (ii) deployment and operation patterns; and (iii) attributes of dependability, safety, and privacy, together with user experience metrics relevant to adherence, effective use of mobile coaching systems, and user engagement [9]. This approach makes it possible to compare heterogeneous approaches using uniform criteria, identify gaps, and derive practical design guidelines.

First, regarding intervention goal and design taxonomy, we organize systems according to their primary intervention goals: hazard perception, regulatory compliance, defensive driving, eco-driving, logistics and fleet operation, or attention to specific collectives (e.g., novice or professional drivers) [10]. We cross these objectives with the AI techniques employed (rules, linear and tree-based models, support vector machines, deep neural networks, and reinforcement learning) [11] and with user engagement mechanisms (gamification, progressive goals, comparative feedback, and incentives) [12]. In addition, we describe the data sources that support behavior assessment and feedback (device sensors, vehicle data, environmental context, expert annotations, and incident and near-miss reports) and their quality [13].

In the second axis: deployment and operation, we characterize where and how inference and learning occur: on the device (on-device), at the edge, or in the cloud, as well as hybrid combinations with co-inference. We examine the architectural implications of each pattern (latency, bandwidth, resilience, operational costs) and their suitability for driver-coaching use cases. We also discuss mechanisms to preserve privacy and data sovereignty, for example, federated learning and the associated trade-offs (client heterogeneity, load balancing, energy consumption, and participation biases) [14].

In the third axis: dependability, safety, and privacy, we propose a set of attributes and metrics to evaluate AI-based mobile coaching platforms in transportation: service dependability, availability and operational continuity (including fault tolerance and offline functioning), information security (confidentiality, integrity, and traceability), privacy protection (data minimization, anonymization, and access control), and functional safety (not introducing additional hazards to the user) [15]. Complementarily, we address user experience metrics (e.g., SUS or UMUX-Lite), perceived mental workload, and adherence, given their direct relationship to program adherence and effective behavior change.

On this basis, we pose the following research questions that guide the mapping: What intervention goals and outcome metrics are prioritized in current AI-based mobile coaching platforms for ground transportation? Which AI techniques and data sources demonstrate greater potential to modulate risky behaviors without increasing distraction? Which deployment patterns best balance latency, availability, energy consumption, and privacy in real-world scenarios? Which interface and architectural design practices foster sustained adoption and user trust? A comparative answer to these questions enables an informed discussion of the effectiveness and limits of existing solutions. These questions also inform the discussion of feedback timing, privacy-by-design trade-offs, and deployment cost constraints in real-world mobile coaching.

The contributions of this work are as follows: (1) a structured mapping of the empirical evidence base on AI-based mobile coaching for driver behavior change in ground transportation (n = 9), covering applications, algorithms, and platforms; (2) a design taxonomy that integrates intervention goals, AI techniques, data sources, deployment patterns, engagement mechanisms, and dependability/safety/privacy requirements; and (3) a set of design guidelines and operational metrics that can guide the implementation and evaluation of future solutions, with an emphasis on minimizing distraction, protecting privacy, and sustaining service availability [16]; [17]; [18].

Organization of the article. Section 2 synthesizes background on road safety programs, mobile driver-coaching solutions, and AI techniques applied to transportation. Section 3 details the mapping protocol and the extraction/coding scheme. Section 4 presents the results: taxonomy, deployment patterns, and attributes of dependability, safety, and privacy. Section 5 discusses implications, limits, and design opportunities. Finally, Section 6 concludes and indicates lines of future work.

2 LITERATURE REVIEW

This section reviews the literature that connects AI, AI-based mobile coaching, and ground transportation. Its purpose is to delineate a framework consistent with the introduction, identifying how AI is applied in transportation systems, what contributions AI-based mobile coaching makes to the education, coaching, and support of drivers and managers, and what coaching capabilities, risks, and safety measures are reported in recent studies. It also highlights trends toward more dependable and user-centered environments. This review is scoped to AI-enabled mobile interventions that deliver feedback or coaching to drivers and report behavioral, engagement, or safety-related outcomes. Broader AI-in-education and automated-driving literature is referenced only insofar as it informs the design of interactive mobile driver coaching (e.g., feedback loops, behavior-change mechanisms, and safety-by-design constraints).

2.1 Integration of AI in ground transportation

The adoption of AI in ground transportation encompasses tasks of environmental perception, driver assistance, trajectory analysis, risk detection, and decision support. These functions rely on the integration of data from heterogeneous sensors, cameras, radar, or infrared, processed by models capable of classifying scenes, estimating states, and proposing actions [19]. The accuracy of such systems is affected by external variables such as weather conditions, lighting, or traffic density and by the quality and representativeness of the training and test data [20], [21]. Likewise, limitations in the coverage of rare cases require human supervision during labeling and validation. Consequently, the robustness of performance is related to the diversity and quality of the data, in addition to the use of complementary strategies such as synthetic generation, domain adaptation, and evaluation in real scenarios to broaden generalization capacity [21].

For driver-coaching purposes, AI-based developments facilitate the implementation of telematics and analytics systems aimed at studying driving habits. These systems record variables such as acceleration, braking, compliance with limits, anticipation of risks, and responsible use of mobile devices [22]. Unlike operational models that act on vehicle systems, this coaching-oriented application employs the same sensing infrastructure to generate personalized feedback and recommendations that promote safer and more deliberate driving decisions [23]. A key distinction in this review is between systems that primarily detect or monitor driving events and those that operationalize coaching through measurable engagement and behavior-change mechanisms. While detection-focused telematics can achieve high recognition performance, coaching-oriented interventions require an explicit feedback loop (goals, reinforcement, reflection) and reporting of behavioral or adherence outcomes to demonstrate educational and safety value.

2.2 Expanded access to training via mobile devices

Mobile devices have transformed access to training and practice, enabling formative and contextual practice experiences in diverse environments. Their portability, processing capacity, and connectivity support brief yet recurrent interactions based on strategies such as microlearning, spaced practice, and active information retrieval [24]. These methodologies strengthen retention and transfer of knowledge, provided that instructional design maintains continuity between sessions, sets clear goals, and adapts to the user's routines through reminders and personalized trajectories [24]. In parallel, online platforms facilitate the management of time, objectives, and activities, driving the gradual integration of competencies across different training levels, from initial instruction to professional development programs [25]; [24].

The utility of this strategy is conditioned by design quality, which must reduce usage friction, adapt to real travel contexts, and prevent distractions while driving. In practice, this implies shifting most interaction to before or after the trip, reserving in-route notifications for brief, non-intrusive messages directly linked to user safety [26].

2.3 AI for personalized training and feedback

AI supports personalized driver coaching by modeling behavioral patterns from mobile and telematics signals and adapting feedback timing, content, and difficulty to the driver's context. In mobile coaching, this typically involves detecting events (e.g., harsh braking, speeding, phone handling), estimating risk levels, and delivering feedback that reinforces safer choices through goal setting, progress tracking, and tailored recommendations [27]; [28]. When interactive training modules are included (e.g., hazard-perception micro-scenarios), AI can further personalize practice schedules and feedback based on performance, supporting retention and transfer while keeping interaction non-intrusive during driving [29].

2.4 Privacy, security, and ethical considerations in mobile driver coaching

In driver-coaching contexts, these issues are amplified because location traces and driving behaviors can be sensitive and re-identifiable, requiring strict minimization, purpose limitation, and transparent governance. The expansion of AI systems in education and coaching introduces significant challenges in terms of privacy, equity, and the ethics of data processing. The collection of sensitive information related to user behavior, performance, or context requires policies of minimization, explicit purposes, and traceability of operations. Model opacity and informational asymmetries between developers and users can generate biases and compromise the fairness of outcomes, so the availability of explanations and clear data management policies is essential to maintain trust [30]; [31].

Moreover, excessive or poorly designed personalization can provoke dependence on digital assistants, constrain deep learning strategies, and increase administrative workload due to false positives in the detection of fraud or plagiarism. Potential adverse effects on teacher autonomy and professional judgment have also been noted when these technologies are applied without a well-defined governance structure [32].

In the context of mobile platforms, approaches that prioritize local processing and distributed training schemes, among them federated learning and its variants, gain relevance. These strategies help preserve privacy by avoiding the transfer of raw data to central servers. However, they introduce new challenges related to heterogeneity among participating devices, energy consumption, and coordination of updates, making it necessary to implement robust synchronization policies and continuous risk assessment mechanisms [33]; [34]; [35].

2.5 Horizons of AI in ground transportation and mobile driver coaching

The interaction between AI and AI-based mobile coaching in the field of ground transportation opens a landscape of coaching innovation oriented toward real contexts and greater coaching efficacy and behavior change. Even so, operational challenges remain that recent research considers priorities. First, the evaluation of interventions must rely on indicators that reflect their impact on road safety, such as hazard perception, reduction of risky behaviors, observance of rules, and the decrease of near misses. Second, it is essential to assess technical attributes of the platforms that affect sustained adoption: end-to-end latency, availability under limited connectivity conditions, energy consumption, and resilience to failures.

Lastly, to evaluate the user experience, standardized criteria are required that articulate usability, mental workload, and adherence, using validated tools such as technology acceptance instruments and brief usability scales. Although automated driving is not the focus of this study, related advances in perception and decision-making illustrate the maturity of sensing and inference techniques that can be adapted to mobile coaching, provided that user-centered interaction and safety constraints are preserved [36]. The main challenge lies in transferring such advances to the driver-coaching domain, ensuring user-centered designs, privacy safeguards, and transparency mechanisms that strengthen institutional trust [37].

3 METHODOLOGY

This study was conducted as a PRISMA-guided systematic mapping focused on AI-based mobile driver coaching for behavior change and road safety in ground transportation. The methodology aimed to (i) identify and classify empirical interventions and (ii) derive a design-oriented framework and a core set of behavioral, engagement, and operational metrics to support future evaluations.

3.1 Approach and research questions

The study was designed as a systematic mapping that prioritized breadth of evidence over experimental depth. After screening and full-text assessment, nine empirical studies (n = 9) met the eligibility criteria and were retained for extraction and coding. As a result, the information could be synthesized into comparable categories and read with a design orientation applied to AI-based mobile coaching in driving and land transportation contexts. Likewise, primary empirical reports were retained to ensure consistent and comparable extraction. The questions that have guided the research are:

a) What intervention goals have AI-enabled mobile solutions addressed in the context of ground transportation?
b) What AI techniques and deployment patterns have predominated, and which data sources (e.g., smartphone sensors, OBD/ECU, telematics, video) have worked best in real-world scenarios?
c) What behavioral, engagement, and adoption metrics have been used to evaluate effectiveness and user trust, and how have the results been reported (e.g., effect sizes, error rates, adherence)?
d) What privacy and security risks and safeguards have been documented, and how have they been integrated into system design and evaluation?

Operationalization. In line with the above, these questions were translated into the extraction dimensions used in the review–intervention goal, AI technique, data source, deployment pattern, metrics, and main result; in addition, privacy/security and relevance to design were recorded as auxiliary fields.

3.2 Thematic scope and eligibility criteria

To narrow down the scope of the study, several complementary axes were defined: first, a deliberately broad population and context covering solutions aimed at drivers, trainees, fleet managers, and other relevant users of land transport, considering both initial training and continuous improvement [38] and [20] second, the interventions and phenomena of interest focused on the technological layer, mobile applications with AI, coaching algorithms, and formative feedback, along with architectures that enable personalization and evaluation [39]. In terms of types of evidence, priority was given to peer-reviewed articles, trial protocols, and reviews with explanatory value; large-scale technical reports were only considered when they presented verifiable controlled experiments [38]. Finally, strict criteria aligned with this scope were applied for filtering, preserving traceability to the included results at all times.

Operational inclusion criteria (I)

- I1 (Domain): Land transport (drivers, trainees, fleet managers, or other road system users).
- I2 (Training link): Direct relationship with training or behavior change (e.g., road safety, eco-driving, distraction reduction).
- I3 (AI-based coaching): The solution uses machine learning approaches (e.g., supervised/deep/RL) and/or advanced rule-based or hybrid inference explicitly employed to personalize coaching/feedback, assessment, or behavior-change support (not solely event logging).
- I4 (Mobile/Embedded): Use of mobile/embedded data or interaction (smartphone/tablet, telematics, OBD/ECU, in-vehicle video).
- I5 (Empirical evidence): Recent publication with verifiable methodological details and observable results (experiments, pilots, field studies, or real datasets).

Operational exclusion criteria (E)

- E-a (Irrelevant domain): Sectors outside land transport (e.g., aviation or maritime).
- E-b (No training link): “Pure” automation/safety without training or behavioral change objectives or metrics.
- E-c (No AI-based coaching): Purely descriptive logging or fixed-threshold/rule-only monitoring that does not implement adaptive/personalized coaching or does not report behavioral/engagement outcomes aligned with coaching.
- E-d (No mobile/embedded): Exclusively desktop or back-office solutions without a mobile component.
- E-e (Insufficient evidence): Proposals without operational evaluation, abstracts without full text, editorials, or lack of access to full text.

3.3 Sources and search strategy

To maximize retrieval of relevant studies, an inclusive search string was applied across multiple digital libraries with extensive indexing in transportation, traffic psychology, computer science, and public health [40]. Searches were conducted in

Scopus as the primary database and in parallel in IEEE Xplore, ACM Digital Library, and Web of Science using equivalent queries with analogous terms. Backward and forward snowballing was performed, complemented by monitoring registered trial projects [41]. The strategy aimed to balance recall and precision under the pre-defined eligibility criteria; however, the review is not intended to be exhaustive, and the findings should be interpreted as a structured mapping of the empirical evidence captured by the protocol and time window (n = 9) [42].

```
TITLE-ABS-KEY(
("AI-based mobile coaching" OR "m-learning" OR "microlearning" OR "learning analytics")
AND (driver* OR driving OR "road user*" OR bus OR truck OR taxi OR rail OR tram OR metro OR "public transport" OR "ground transportation")
AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network*" OR "computer vision" OR "reinforcement learning" OR "natural language processing")
AND (smartphone OR "mobile app*" OR "in-vehicle" OR telematic* OR "on-board" OR OBD OR ECU)
AND (feedback OR "behaviour change" OR "learning outcome*" OR assessment OR "skill acquisition")
)
AND (PUBYEAR > 2009 AND PUBYEAR < 2026)
AND (DOCTYPE(ar) OR DOCTYPE(cp))
```

Operationalization: The above chain has been directly mapped to the RQs and extraction dimensions (intervention goal, AI technique, data source, deployment, metrics, and main result), maintaining privacy/security and relevance to the design as auxiliary fields.

3.4 Limitations of the search and evidence base

Four limitations should be acknowledged. First, the included studies were limited (n = 9), which constrains breadth and supports interpretation as a structured mapping of the evidence captured by the protocol and time window rather than an exhaustive representation of the field. Second, the included studies reported behavioral outcomes using heterogeneous metrics and reporting practices, which limits direct comparability across interventions and prevents meaningful aggregation beyond descriptive synthesis. Third, the scope was restricted to land transportation, so the resulting taxonomy and design implications should not be extrapolated to other mobility contexts without additional evidence. Fourth, the strategy prioritized English-language publications and major indexed sources, which may underrepresent evidence published in other languages or regional outlets.

3.5 Screening and data extraction procedure

Following PRISMA, the selection and extraction process was structured in two complementary phases: in Phase 1 (initial identification and screening), records were compiled from Scopus, IEEE Xplore, ACM Digital Library, and Web of Science, and after removing duplicates, titles and abstracts were screened by preliminarily applying criteria I1–I5; in Phase 2 (full-text evaluation), potentially eligible reports were reviewed in their entirety, and the criteria were strictly applied, recording a single coded reason E-x (E-a irrelevant domain, E-b no training link, E-c no AI-based coaching, E-d no mobile/embedded component, and E-e insufficient evidence or no access), with any discrepancies resolved by consensus. The final set included nine primary studies, reflecting the strict exclusion criteria and the decision to retain only reports with verifiable empirical metrics on driver behavior, user engagement, or system performance, prioritizing the quality and comparability of evidence over sheer volume of publications. To ensure numerical consistency in the PRISMA flow, counts were cross-checked across stages, and the decision flow was verified as internally consistent.

Table 1. PRISMA 2020 flowchart of the study selection process

Stage	Description	n
A	Records identified in databases (Scopus/IEEE/ACM/WoS)	482
B	Records after deduplication	385
C	Screened records (title/abstract)	385
D	Full-text reports evaluated	54
E	Excluded from the full text (due to E-x)	45
F	Studies included in the mapping	9

Notes: Excluded in screening (title/abstract) X = C – D = 331. **Breakdown E**. E-a = 7; E-b = 14; E-c = 9; E-d = 8; E-e = 7. (Total E = 45).

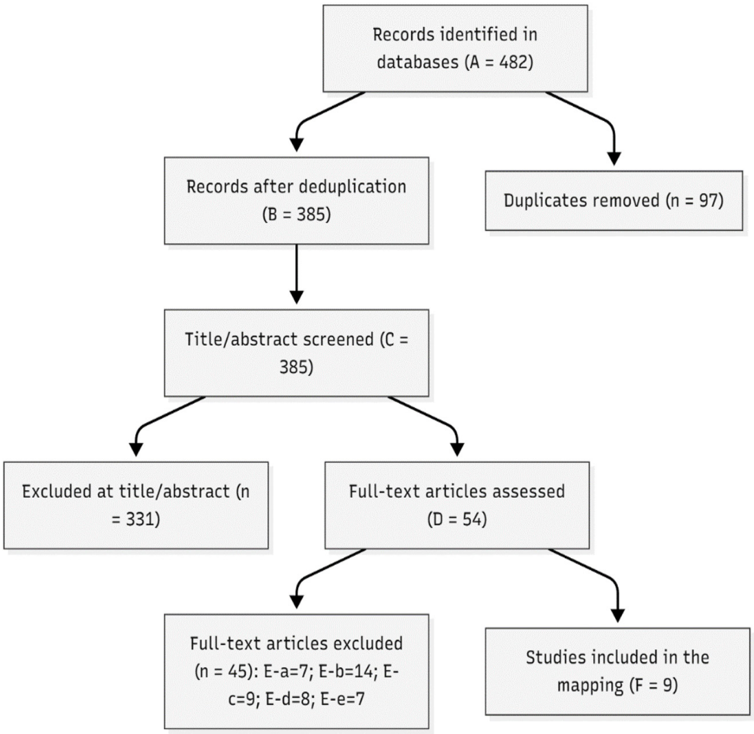


Fig. 1. PRISMA 2020 flow diagram of study selection

3.6 Taxonomy construction

The taxonomy was constructed inductively, i.e., it emerged directly from the evidence reported by the screened and ultimately included cases. To this end, an extraction template aligned with the research questions was used, which, in each study, recorded (i) the intervention goal, (ii) the AI technique used, (iii) the data sources, (iv) the deployment pattern (on-device, in-vehicle/edge, or cloud/hybrid),

(v) the evaluation metrics, and (vi) the main result, along with privacy/security notes and their relevance to the design. Through open coding and consensus refinement, semantically equivalent labels were harmonized, for example: distraction due to phone use, speed management, hazard perception, eco-driving, and compliance/reports. Where possible, units were normalized to relative variation (%) and the time window of the intervention was documented.

The resulting framework achieved a careful balance between the technical specificity required for documentation and the practical usefulness required by teams designing and evaluating AI-based mobile coaching solutions in transportation, in line with the findings of design-oriented literature [38]; [5]. Specifically, the taxonomy articulated five coaching-oriented clusters with families of AI techniques, including classic ML, deep learning/vision, NLP, reinforcement learning, and hybrids, and also

Table 2. Initial taxonomy of representative studies

ID	Study	Type of Contribution	Intervention Goal	AI Technique	Data and Sensors
1	[44] – Safer Driver App (RCT)	Mobile application with behavioral coaching	Reduce phone-related distractions and promote attentiveness while driving	Telematics analytics with rules and lightweight machine learning for detecting phone use and events	Accelerometer, gyroscope, GPS, device interaction events
2	[45] – Personalized feedback in young drivers	Mobile app with telematics and reporting	Decrease risky behaviors in novice drivers	Event detection models from smartphone signals	Harsh acceleration/braking, speed, location
3	[46] – FEEDBACK trial (RCT protocol)	Large-scale randomized trial protocol with incentives	Reduce crashes and speeding via telematics and messages	Telematics analytics with detection rules and risk thresholds	Smartphone and OBD device; speed and event data
4	[47] – Safe driving mobile intervention in young drivers	Application with behavioral intervention	Reduce intention to speed and normalize safe habits	Behavior analytics with feedback rules and goals	Phone use signals, location, and speed
5	[48] – Incentive trial to reduce phone use	Randomized trial with app and incentive schemes	Decrease manual phone use while driving	Detection of interactions and distraction patterns with lightweight ML	Unlock events, screen touches, motion, and speed
6	[49] – Review on distraction detection with smartphones	Technical systematic review	Map methods to detect distraction that can feed driver coaching	Classification with ML on accelerometer, gyroscope, GPS, and camera	Smartphone sensors and, in some cases, video signals
7	[50] – Online hazard-perception training	Automated course with performance analytics	Improve hazard perception and decision-making	Analytics on responses and content adaptation by performance	Labeled traffic clips, response times, and accuracy
8	[51] – Reinforcement learning for eco-driving	Algorithm applicable to efficiency coaching	Reduce consumption and emissions via efficient driving recommendations	Reinforcement learning with states derived from traffic and trajectory	Trajectories, speed, and traffic conditions
9	[52] – Offline RL with real trajectories	Offline-training algorithm	Train efficient-driving policies from historical data for subsequent recommendations	Offline reinforcement learning with later validation	Historical trajectories, traffic profiles, and consumption

related them to deployment decisions (on-device/in-vehicle, edge, and cloud/hybrid). This combination facilitated a comparable reading of the results and adoption signals without losing sight of the privacy and security considerations documented by the primary studies.

3.7 Quality assessment and threats to validity

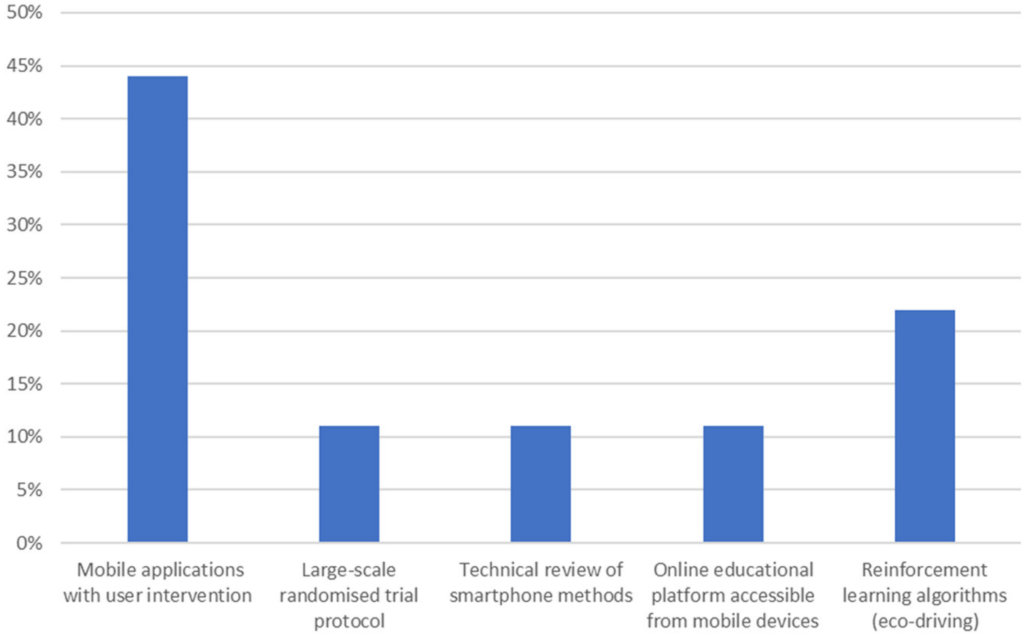
To evaluate the quality of the reviewed studies, we considered design clarity, the validity of the metrics used, sample size, and transferability to real contexts [43]. Nevertheless, threats to validity were identified, such as metric heterogeneity, the absence of long-term follow-up, dependence on self-reports, and device variability.

Predominant Deployment	Behavioral, Engagement, and Operational Metrics	Main Results or Central Evidence	Privacy and Security	Relevance for Design
On-device with feedback panel; server for aggregation	Frequency and duration of manual phone use while driving; risk events; app adherence	Significant reduction in manual use versus control app; evidence in young population	Personal data minimization, explicit consent, aggregated reporting	Shows that mobile telematics coaching can modify specific behavior without constant human intervention
On-device app with analytics and periodic summaries	Risk event rates, speeding, temporal trend of indicators	Significant reductions in risky behaviors versus the no-feedback group	Anonymization in reports and dashboards; protection of sensitive locations	Reinforces value of post-trip summaries and progressive goals in high-risk populations
Hybrid architecture with on-device app and backend for evaluation	Reported crashes, speed-limit violations, risk exposure, adherence to incentives	Provides robust methodological design targeting hard outcomes and longitudinal evaluation	Data governance, role-based access control, secure storage	Serves as a template for effectiveness trials with real impact metrics
On-device app with motivational dashboard	behavioral intention, self-reported habits, pre/post changes	Signals of improvement in intention and self-reported practices	Informed consent and location-permission controls	Illustrates usefulness of goals and comparison with user's own history
On-device app and backend for incentive calculation	Sustained reduction in manual use; subgroup and incentive-type analyses	Reduction observed versus control; greater effect in users with high baseline	Minimum data-retention policies and encrypted storage	Evidence that combining feedback and rewards improves adherence
Local processing assisted by a server for training and validation	Detection accuracy per sensor/method; false-positive rates	Summary of best sensor–algorithm combinations for road use cases	Privacy recommendations via local processing and video minimization	Technical basis for detection modules integrable into mobile coaching apps
Online platform accessible via mobile	Hazard-perception test scores, sustained improvement, and transfer	Improvements in performance with signals of operational safety impact	Management of performance data with pseudonymized profiles	Demonstrates structured hazard-perception training can be integrated into mobile itineraries
On-device or edge inference with parameters updated from server	Energy savings versus rule baselines; stability and convergence analysis	Efficiency gains without degrading travel time in simulated and real scenarios	Local parameters and periodic updates to minimize data exposure	Transferable pattern for eco-driving advice within mobile coaching apps
Server-side training and on-device policy deployment	Policy improvement over baseline; robustness to distribution shift	Real-world data enable useful policies for later coaching	Separation of raw user data and model deployment without re-identification	Favors designs combining centralized training with local inference

These limitations are key and are discussed in detail when interpreting the results and proposing future design guidelines.

4 RESULTS

The Results section synthesizes the nine studies previously classified in Table 2. Patterns are reported across five thematic blocks: contributions and objectives, AI techniques and data sources, deployment patterns, evaluation metrics, and privacy/security/dependability attributes. Summary tables and figures are provided to enable direct comparison across studies.



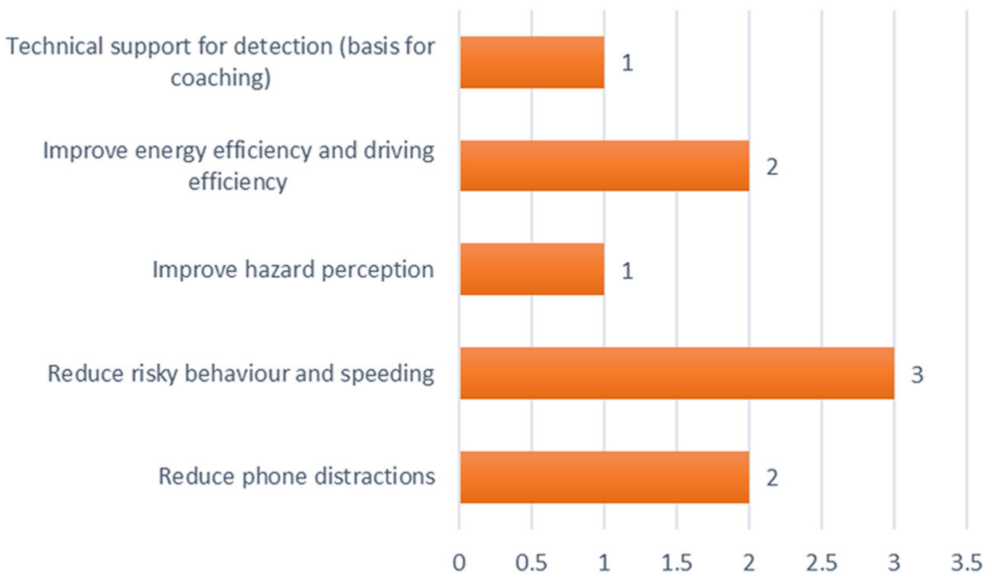


Fig. 2. Distribution of mapped studies (n = 9): (a) by type of contribution and (b) by intervention goal

Note: Bars indicate study counts; study IDs correspond to Table 2.

4.1 Overview of contributions and objectives

Across the mapped studies, intervention-oriented mobile applications primarily target phone-related distraction and speeding. Additional contributions include algorithmic studies focused on efficiency (e.g., eco-driving optimization) and training-oriented approaches centered on hazard perception. A technical review contributes methodological coverage of smartphone-based detection approaches that are commonly used as inputs for mobile coaching analytics and feedback mechanisms.

4.2 AI techniques, data sources, and deployment patterns

Table 3. Predominant AI techniques and data sources reported in the mapped studies (n = 9)

AI Technique	Studies	Most Common Data Sources
Telematics analytics with rules and lightweight models	[44], [45], [46], [47], [48]	Accelerometer, gyroscope, GPS, speed, phone interaction events
Multi-sensor classification on smartphone	[49]	Accelerometer, gyroscope, GPS, and camera, where applicable
Reinforcement learning for eco-driving	[51], [52]	Trajectories, speed, traffic conditions, estimated consumption
Educational performance analytics	[50]	Responses to clips, reaction times, accuracy

Note: Study IDs correspond to Table 2.

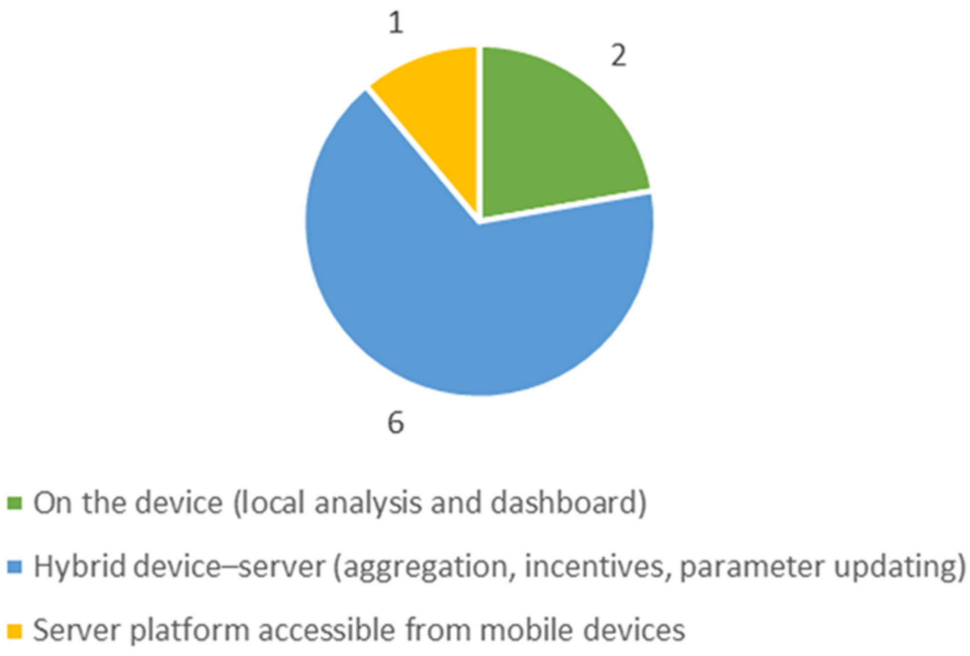


Fig. 3. Deployment patterns across mapped studies (n = 9): hybrid device–server (6), on-device (2), and server-based (1)

Hybrid deployments combining on-device inference with server-side components are the most frequently reported pattern, typically to support aggregation, incentives, parameter updates, or longer-term monitoring. Purely on-device approaches are reported when the focus is immediate event detection and low-friction feedback. Server-hosted platforms appear mainly in structured training settings where centralized learning analytics are required.

4.3 Behavioral, engagement, and operational metrics

Table 4. Types of outcome and process metrics reported in the mapped studies (n = 9)

Type of Metric	Description	Studies
Risk behavior	Manual phone use, speeding, harsh braking, and acceleration	[44], [45], [46], [47], [48]
Educational performance	Hazard-perception scores, response times, accuracy rate	[50]
Operational efficiency	Energy savings, policy stability, convergence	[51], [52]
Adoption and adherence	Usage frequency, session completion, response to incentives	[44], [46], [48]

Note: Study IDs correspond to Table 2. The Studies column lists sources report each metric family.

The metrics reveal a clear predominance of those that evaluate the driver's risk behavior. Reported outcome measures are dominated by behavioral risk metrics (e.g., manual phone use, speeding, harsh maneuvers). Engagement and adherence indicators are also reported in several interventions (e.g., usage frequency, session completion, response to incentives). Educational performance metrics are reported primarily in hazard-perception training studies, whereas operational efficiency metrics are reported mainly in optimization-oriented contributions.

4.4 Privacy, security, and dependability

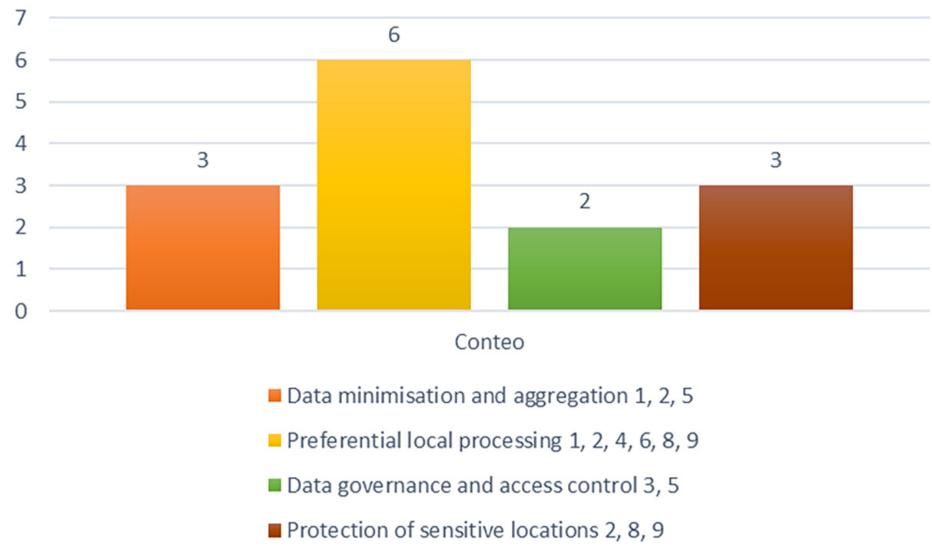


Fig. 4. Reported privacy and security practices across mapped studies (n = 9)

Note: Bar heights indicate study counts; legend lists supporting study IDs. Numbers in the legend denote study IDs as listed in Table 2.

The reported privacy and security practices focus on protecting user data through complementary strategies. Data minimization and aggregation seek to capture only strictly necessary information and to present it in aggregate reports to avoid unnecessary exposure, while preferential local processing allows detection and inference to be performed directly on the device, thereby reducing transfer and leakage risks. In turn, data governance and access control are supported by role policies, consent, and encryption that ensure responsible and secure information handling. Finally, the protection of sensitive locations is achieved through anonymization or pseudonymization of trajectories, which prevents direct identification of users from their movement patterns. Taken together, these practices strengthen security and trust in data processing, balancing the utility of information with the safeguarding of privacy.

Table 5. Coverage of dependability attributes in the mapped studies (n = 9)

Attribute	Reported Explicitly	Studies	Observation
End-to-end latency	0/9	–	No quantitative measurements are presented
Offline operation	3/9	[44], [45], [47]	Mentioned in general terms; no degradation testing
Battery impact	0/9	–	Not documented with comparable metrics
Availability and continuity	2/9	[46], [50]	Operational allusions without target rates
Fault resilience	0/9	–	No controlled recovery tests

Note: "Reported explicitly" indicates quantitative reporting rather than qualitative mention; study IDs correspond to Table 2.

The table shows that dependability attributes in the reviewed studies are sparsely covered and, in most cases, are not reported explicitly or with clear metrics. Aspects such as end-to-end latency, battery impact, and fault resilience are not documented at all. Offline operation is mentioned in some studies, but only in general terms and without degradation testing, whereas availability and continuity are referred to superficially, without setting target rates. Overall, there is a lack of rigorous, controlled evaluations of these attributes, which limits a realistic understanding of the performance of the systems analyzed.

4.5 Strength of the evidence

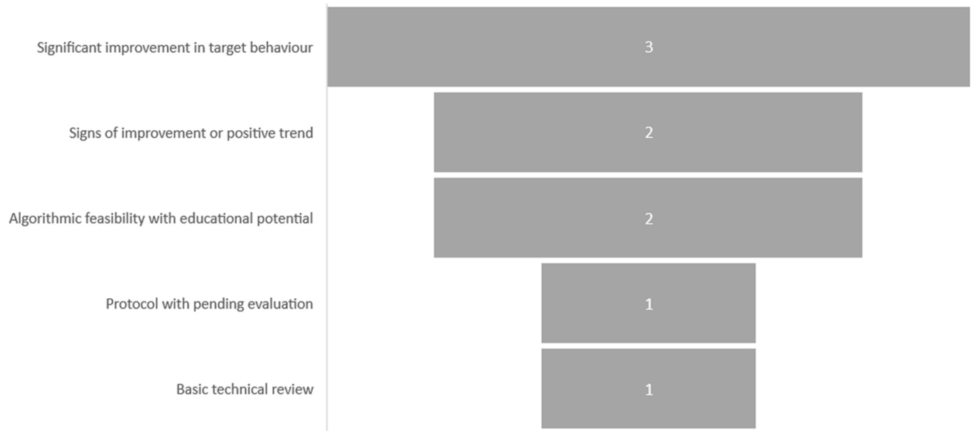


Fig. 5. Strength of evidence classification for the mapped studies (n = 9)

Note: Counts indicate the number of studies per evidence tier.

Across the nine studies, three interventions report statistically significant improvements in target behaviors under controlled comparisons, while two report positive trends with methodological constraints. Two studies contribute algorithmic evidence on eco-driving or efficiency, one protocol reports a large-scale design with pending outcomes, and one technical review supports smartphone-based detection for coaching.

5 DISCUSSION

This section interprets the mapped evidence through the three analytical axes (intervention goals and design, deployment patterns, and dependability/safety/privacy) and situates the findings within recent literature to derive practical implications for interactive mobile driver coaching.

The mapped evidence (n = 9) suggests two overarching patterns. First, telematics-enabled interventions that deliver personalized feedback have been associated with measurable behavioral improvements. Second, eco-driving modules are increasingly aligned with deployment choices that favor local inference over remote processing. Together, these patterns support a mobile coaching itinerary that prioritizes post-trip feedback and treats privacy-by-design as a core requirement [49]; [50].

A cross-cutting finding is the need to "align times": the coaching intervention should occur at moments that maximize its usefulness and minimize interference with driving. In practice, this translates into privileging post-trip summaries and recommendations, activating non-intrusive predictive alerts when the vehicle is stopped, and avoiding interactions that demand visual or motor attention while the vehicle is in motion. The design should also consider the diversity of contexts (urban/interurban environments, private or fleet vehicles, and novice or professional drivers) and accessibility (language, digital literacy, lighting conditions, and disabilities) [16]. This timing principle also operationalizes safety-by-design by reducing interaction demands during vehicle motion.

Another key tension emerges between personalization and privacy. The promise of adapting content and levels of difficulty requires understanding user performance and context; however, capturing and processing those data entail operational and reputational risks. Therefore, we emphasize strategies that reduce the exposure of personal data (local processing, anonymization, and secure aggregation) and increase transparency (permission-control dashboards, purpose explanations, and audit reports) [17]. For transferability, these safeguards should be reported as concrete design features (e.g., data-retention rules, access-control roles, and auditability) rather than as generic compliance statements. This combination is essential to earn and sustain trust, a prerequisite for any measurable behavior and road-safety impact.

From the perspective of adoption and scalability, operational cost also matters. Solutions that depend exclusively on the cloud may face latency bottlenecks and computing and transfer costs; on the other hand, processing entirely on the device may be limited by computing capacity and battery consumption. Hybrid designs that combine local inference with edge or cloud services for heavy tasks and for global training make it possible to balance performance, costs, and privacy, provided that robust synchronization and update policies are implemented [18].

With respect to phone distraction, our findings on the reduction of manual use are congruent with key findings in the literature. Indeed, [48] showed, in a large-scale randomized clinical trial, that the combination of comparative social feedback and financial incentives yields relative decreases of between 15% and 21% versus the control. Additionally, [44] reinforced the effectiveness of mobile coaching by observing a significant reduction in the duration of distraction with a specific application. In sum, this evidence aligns with effects reported in real-world evaluations while acknowledging differences in contexts and outcome measures across studies.

External validity remains constrained by contextual factors that are unevenly reported across studies, including geography, regulatory enforcement, fleet versus private driving, and driver demographics (e.g., novice versus professional). These factors may moderate both baseline risk and responsiveness to feedback and should be documented systematically to support comparison and generalization.

For the issue of young drivers and speeding, the improvement identified when using personalized reports and goals is consistent with the evidence base. For example, [45] observed significant reductions in risky behavior among novice drivers when applying feedback based on events captured directly by smartphone sensors. In this sense, the study by [47], a three-month randomized trial using a safe-driving application, complements the above by finding a decrease in the intention to exceed the speed limit. This effectiveness is attributed to key engagement mechanisms such as progress indicators, graduated goals, and comparison with the user's own history. Therefore, these methodological and design elements are mechanisms that are also reflected in the taxonomy and may support sustained engagement.

In the line of hazard perception, the mapped evidence supports the incorporation of two fundamental elements: spaced-practice modules and automated assessments accessible via mobile. The literature provides strong support for this, as [50] demonstrated that an online course produced improvements in response times and performance, achieving an explicit bridge between measurable learning and the effective reduction of road risk. Accordingly, this evidence justifies that the minimum battery of educational assessment we propose should be robust, including not only behavioral and adherence metrics but also a standardized and specific measure of hazard perception.

Regarding detection and analytics on smartphones, our methodological inclination towards local processing and data minimization aligns with best practice. In their review, [49] catalog algorithms and accuracy and recommend prioritizing on-device computation to mitigate the exposure and biases that arise when transferring sensitive data. Consequently, this strong external evidence reinforces the hybrid deployment pattern that we have observed. In this model, primary inference is strategically maintained at the edge to optimize latency and privacy, while the cloud is reserved solely for less sensitive functions such as aggregation, incentive calculation, and parameter updates, achieving a key balance with educational utility.

For efficiency and eco-driving, our findings converge directly with the reinforcement-learning stream. [51], for example, demonstrated that policies learned with DDPG, PPO, and SAC in mixed traffic improve efficiency without compromising travel times. This result provides direct support for integrating an eco-driving advice module that uses local inference with periodic parameter updating. Complementarily, [52] showed that offline reinforcement learning (using real trajectories) enables training of these policies without exposing raw data. In sum, this approach is not only coherent with our priority of privacy by design but also offers a practical route for implementing these algorithms in mobile coaching itineraries.

Finally, the gap that remains under-documented in the applied literature is highlighted: that relating to system dependability, end-to-end latency, battery impact, offline functioning, and resilience. Notwithstanding the above, [46] provides useful methodological guidance by establishing a robust protocol to measure hard outcomes at scale that integrates data governance and transparent incentive calculation.

The issue is that most studies do not report operational metrics that are comparable, which significantly limits their transferability. For this reason, and based on our findings, we propose that it is essential to incorporate a minimum block of operational indicators alongside the educational and road-safety metrics already proposed in order to strengthen external validity and accelerate institutional adoption of these technologies.

6 CONCLUSIONS AND RECOMMENDATIONS

This review consolidated an operational taxonomy and a set of comparable metrics and, on this basis, proposed system-oriented design guidelines that combine on-device detection, personalized feedback, and, where applicable, incentives, consistent with the maturity of smartphone analytics and with evidence on brief, structured hazard-perception training that can be integrated into driver itineraries without compromising safety [53]; [49]; [50]. Consequently, priority is given to on-device deployments or hybrid options with privacy-by-design, local processing, and distributed training, in addition to recommending comparative evaluations with educational, behavioral, and operational indicators.

However, significant limitations remain, such as the heterogeneity of metrics and mostly short- or medium-term monitoring, which requires standardizing definitions and units, as well as publishing dependability, latency, energy consumption, and availability profiles to facilitate adoption and scalability [54].

In relation to phone distraction, the convergence between our findings and recent randomized trials is clear: both the multicenter study with comparative social feedback and incentives and the trial with a specific application report significant reductions in manual phone use relative to controls [55]. This reinforces that progress dashboards, graduated goals, and rewards increase adherence and amplify the behavioral effect in real contexts [48]; [44]. Among young drivers, feedback based on events detected by smartphone sensors and progressive goals shows decreases in risky behavior and in the intention to exceed speed limits, consistent with instructional design that prioritizes post-trip summaries and comparison with one's own history as drivers of sustained engagement [56]; [45]; [47]. Regarding efficiency and eco-driving, policies learned through reinforcement learning improve consumption indicators without penalizing travel times and offer a practical pathway for deploying on-device recommendations with periodic parameter updates, especially when trained with real trajectories in offline schemes that minimize exposure of raw data [51]; [52].

On the basis of this evidence, the design of mobile driver-coaching platforms should consolidate a minimum set of comparable measurements that makes it possible to assess, beyond behavioral change, the operational robustness of the system. It is appropriate to combine behavioral indicators such as the percentage change in manual phone use, speeding episodes per hundred kilometers, and a normalized rate of harsh maneuvers with standardized educational metrics for hazard perception and adherence, as well as an operational block that includes end-to-end latency with percentiles, battery impact per hour of use, effective availability in limited-connectivity scenarios, and in-field version stability. The scarcity of reports on these dependability attributes in the applied literature suggests an immediate opportunity for methodological improvement, which can draw on pragmatic designs with longitudinal follow-up and clear data governance, as proposed by large-scale evaluation protocols focused on hard outcomes [46].

In parallel, user trust is sustained when the architecture prioritizes local processing and inference near the edge; when aggregation and incentive calculation occur in a deferred manner on the server; and when the data policy is explicit regarding purposes, retention, and access controls, in line with technical recommendations for smartphone-based detection that favor minimization and on-device computation [49].

Taken together, the results and their contrast with the literature support a concrete roadmap. This roadmap is intended to be actionable for three stakeholder groups: researchers (comparative evidence and reporting standards), developers and system designers (deployment and UX patterns), and policymakers/fleet operators (adoption requirements, governance, and scalable evaluation). The roadmap prioritizes: local detection and inference to reduce exposure and perceived latency; non-intrusive and preferably post-trip feedback to avoid distractions; progress dashboards and graduated goals with the option of incentives to sustain adherence; brief hazard-perception modules integrated into the mobile itinerary to reinforce transferable learning; and evaluation with a common core of educational, behavioral, and operational metrics that raises comparability across proposals. With these components, AI-enabled mobile driver behavior change can become an effective instrument to reduce risks, increase technological literacy, and sustain tangible improvements in ground transportation while preserving privacy by design and strengthening external validity through consistent operational measurements [48]; [44]; [50]; [49]; [46]; [51]; [52]; [45]; [47]. The first components are directly grounded in the mapped evidence, whereas the operational reporting block (latency, energy, availability, resilience) is advanced as a pragmatic requirement to address the documented dependability gap and strengthen transferability.

7 DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, ChatGPT (OpenAI) was used to improve readability and language. After using this tool, the content was reviewed and edited as necessary, with full responsibility assumed for the content of the publication.

8 REFERENCES

[1] R. Goel, "Effectiveness of road safety interventions: An evidence and gap map," *Campbell Systematic Reviews*, vol. 20, no. 1, p. e1367, 2024. https://doi.org/10.1002/cl2.1367

[2] K. Bhalla, D. Mohan, and B. O'Neill, "How much would low- and middle-income countries benefit from addressing the key risk factors of road traffic injuries?" *International Journal of Injury Control and Safety Promotion*, vol. 27, no. 1, pp. 83–90, 2020. https://doi.org/10.1080/17457300.2019.1708411

[3] S. Waring, L. Almond, and L. Halsall, "Examining the effectiveness of an education-based road safety intervention and the design and delivery mechanisms that promote road safety in young people," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 105, pp. 336–349, 2024. https://doi.org/10.1016/j.trf.2024.07.019

[4] S. Von and A. Bresges, "Effectiveness of educational interventions in minimizing reactance in traffic accident prevention," *Frontiers in Education*, vol. 8, p. 1276380, 2023. https://doi.org/10.3389/feduc.2023.1276380

[5] E. Mantouka, E. Barmpounakis, E. Vlahogianni, and J. Golias, "Smartphone sensing for understanding driving behavior: Current practice and challenges," *International Journal of Transportation Science and Technology*, vol. 10, no. 3, pp. 266–282, 2020. https://doi.org/10.1016/j.ijtst.2020.07.001

[6] A. Kashevnik, I. Lashkov, and A. Gurtov, "Methodology and mobile application for driver behavior analysis and accident prevention," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 6, pp. 2427–2436, 2020. https://doi.org/10.1109/TITS.2019.2918328

[7] D. Levshun, A. Chechulin, and I. Kotenko, "Security and privacy analysis of smartphone-based driver monitoring systems from the developer's point of view," *Sensors*, vol. 22, no. 13, p. 5063, 2022. https://doi.org/10.3390/s22135063

[8] W. Feng, S. Yang, Y. Gao, N. Zhang, R. Ning, and S. Lin, "Reverse offloading for latency minimization in vehicular edge computing," in *Proc. IEEE Int. Conf. Communications (ICC)*, 2021, pp. 1–6. https://doi.org/10.1109/ICC42927.2021.9500532

[9] J. Yu and X. Yu, "Assessing the user acceptance of mobility-as-a-service platforms: A case study of Shenzhen, China," *Industrial Management & Data Systems*, vol. 126, no. 1, pp. 323–343, 2025. https://doi.org/10.1108/IMDS-10-2024-1017

[10] S. Haghshenas, V. Astarita, S. Shaffiee Haghshenas, and G. Guido, "Artificial Intelligence-powered driver behavior analysis for fuel consumption optimization: A pathway to greener roads," in *Proc. 10th Int. Conf. Control, Decision and Information Technologies (CoDIT)*, 2024, pp. 116–122. https://doi.org/10.1109/CoDIT62066.2024.10708512

[11] K. Yuan, Y. Huang, L. Guo, H. Chen, and J. Chen, "Human feedback enhanced autonomous intelligent systems: A perspective from intelligent driving," *Autonomous Intelligent Systems*, vol. 4, p. 9, 2024. https://doi.org/10.1007/s43684-024-00071-z

[12] Z. Marafie *et al.*, "AutoCoach: An intelligent driver behavior feedback agent with personality-based driver models," *Electronics*, vol. 10, no. 11, p. 1361, 2021. https://doi.org/10.3390/electronics10111361

[13] Y. Takefuji, M. Tano, M. Shigehara, and S. Sato, "Artificial intelligence abnormal driving behavior detection for mitigating traffic accidents," *Computers & Industrial Engineering*, vol. 198, p. 110667, 2024. https://doi.org/10.1016/j.cie.2024.110667

[14] W. Wu, L. He, W. Lin, and R. Mao, "Accelerating federated learning over reliability-agnostic clients in mobile edge computing systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 6, pp. 1539–1551, 2021. https://doi.org/10.1109/TPDS.2020.3040867

[15] H. Hu, L. Wang, Q. Hu, Y. Bu, and Y. Zhang, "Deep-learning-based mobile group intelligence perception mechanism oriented to user privacy and data security in the Internet of Things," *IEEE Wireless Communications*, vol. 29, no. 3, pp. 60–67, 2022. https://doi.org/10.1109/MWC.007.2100444

[16] Y. Peng, G. Song, M. Guo, L. Wu, and L. Yu, "Investigating the impact of environmental and temporal features on mobile phone distracted driving behavior using phone use data," *Accident Analysis & Prevention*, vol. 180, p. 106925, 2023. https://doi.org/10.1016/j.aap.2022.106925

[17] A. Siddiqua, S. Sabeer, R. S. Rao, S. Ahuja, S. Aggarwal, and M. Lourens, "Machine learning-driven educational ethics considerations: Striking a balance between privacy and personalization," in *Proc. 10th IEEE Uttar Pradesh Section Int. Conf. Electrical, Electronics and Computer Engineering (UPCON)*, 2023, pp. 1748–1753. https://doi.org/10.1109/UPCON59197.2023.10434880

[18] W. Ejaz *et al.*, "Generative AI driven edge cloud system for intelligent road infrastructure inspection," *ICT Express*, vol. 11, no. 3, pp. 325–333, 2025. https://doi.org/10.1016/j.icte.2025.06.004

[19] K. Hussain, C. Moreira, J. Pereira, S. Jardim, and J. Jorge, "A comprehensive literature review on modular approaches to autonomous driving: Deep learning for road and racing scenarios," *Smart Cities*, vol. 8, no. 3, p. 79, 2025. https://doi.org/10.3390/smartcities8030079

[20] J. Li, R. Xu, J. Ma, Q. Zou, J. Ma, and H. Yu, "Domain adaptive object detection for autonomous driving under foggy weather," in *Proc. 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2023, pp. 612–622. https://doi.org/10.1109/WACV56688.2023.00068

[21] T. Zhang *et al.*, "Unsupervised domain adaptation for self-driving from past traversal features," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, 2023, pp. 4040–4046. https://doi.org/10.1109/ICCVW60793.2023.00436

[22] T. Fonseca and S. Ferreira, "Monitoring technologies for truck drivers: A systematic review of safety and driving behavior," *Applied Sciences*, vol. 15, no. 12, p. 6513, 2025. https://doi.org/10.3390/app15126513

[23] V. Shirole, A. K. Shahade, and P. V. Deshmukh, "A comprehensive review on data-driven driver behaviour scoring in vehicles: Technologies, challenges and future directions," *Discover Artificial Intelligence*, vol. 5, p. 26, 2025. https://doi.org/10.1007/s44163-025-00244-6

[24] L. Martinengo *et al.*, "Spaced digital education for health professionals: Systematic review and meta-analysis," *Journal of Medical Internet Research*, vol. 26, p. e57760, 2024. https://doi.org/10.2196/57760

[25] W. K. Monib, A. Qazi, and R. A. Apong, "Microlearning beyond boundaries: A systematic review and a novel framework for improving learning outcomes," *Heliyon*, vol. 11, no. 2, p. e41413, 2025. https://doi.org/10.1016/j.heliyon.2024.e41413

[26] Q. Zhong, J. Zhi, Y. Xu, P. Gao, and S. Feng, "Assessing driver distraction from in-vehicle information system: An on-road study exploring the effects of input modalities and secondary task types," *Scientific Reports*, vol. 14, p. 20289, 2024. https://doi.org/10.1038/s41598-024-71226-4

[27] L. Y. Tan, S. Hu, D. J. Yeo, and K. H. Cheong, "Artificial intelligence-enabled adaptive learning platforms: A review," *Computers and Education: Artificial Intelligence*, vol. 9, p. 100429, 2025. https://doi.org/10.1016/j.caeai.2025.100429

[28] W. Strielkowski *et al.*, "AI-driven adaptive learning for sustainable educational transformation," *Sustainable Development*, vol. 33, no. 5, pp. 1–12, 2025. https://doi.org/10.1002/sd.3221

[29] M. Martínez, X. Rigueira, A. Ana, J. Martínez, I. Ocarranza, and D. Kreibel, "Impacto de la inteligencia artificial en los métodos de evaluación en la educación primaria y secundaria: Revisión sistemática de la literatura," *Revista de Psicodidáctica*, vol. 28, no. 2, pp. 93–103, 2023. https://doi.org/10.1016/j.psicod.2023.06.001

[30] Y. Yan, H. Liu, and T. Chau, "A systematic review of AI ethics in education: Challenges, policy gaps, and future directions," *Journal of Global Information Management*, vol. 33, no. 1, pp. 1–50, 2025. https://doi.org/10.4018/JGIM.386381

[31] Y. R. Marín *et al.*, "Ethical challenges associated with the use of artificial intelligence in university education," *Journal of Academic Ethics*, vol. 23, no. 4, pp. 2443–2467, 2025. https://doi.org/10.1007/s10805-025-09660-w

[32] C. Zhai, S. Wibowo, and D. Li, "The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review," *Smart Learning Environments*, vol. 11, no. 28, 2024. https://doi.org/10.1186/s40561-024-00316-7

[33] A. Arouj and A. M. Abdelmoniem, "Towards energy-aware federated learning via collaborative computing approach," *Computer Communications*, vol. 221, pp. 131–141, 2024. Available: https://doi.org/10.1016/j.comcom.2024.04.012

[34] C. Chen *et al.*, "Advances in robust federated learning: Heterogeneity, attacks and aggregations," *IEEE Transactions on Neural Networks and Learning Systems*, 2025. https://arxiv.org/html/2405.09839v1

[35] N. Saeed *et al.*, "Comprehensive review of federated learning challenges: Data, systems, fairness," *Journal of Big Data*, vol. 12, no. 195, 2025. https://doi.org/10.1186/s40537-025-01195-6

[36] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: Perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, p. 43, 2023. https://doi.org/10.1186/s41239-023-00411-8

[37] S. Wang, F. Wang, Z. Zhu, J. Wang, T. Tran, and Z. Du, "Artificial intelligence in education: A systematic literature review," *Expert Systems with Applications*, vol. 252, p. 124167, 2024. https://doi.org/10.1016/j.eswa.2024.124167

[38] J. Passmore, B. Olafsson, and D. Tee, "A systematic literature review of artificial intelligence (AI) in coaching: Insights for future research and product development," *Journal of Work-Applied Management*, 2025. https://doi.org/10.1108/JWAM-11-2024-0164

[39] M. Siami, M. Naderpour, and J. Lu, "A mobile telematics pattern recognition framework for driving behavior extraction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 1459–1472, 2021. https://doi.org/10.1109/TITS.2020.2971214

[40] Y. Adeoye, E. F. Onotole, T. Ogunyankinnu, G. Aipoh, A. A. Osunkanmibi, and J. Egbemhenghe, "Artificial intelligence in logistics and distribution: The function of AI in dynamic route planning for transportation, including self-driving trucks and drone delivery systems," *World Journal of Advanced Research and Reviews*, vol. 25, no. 2, pp. 155–167, 2025. https://doi.org/10.30574/wjarr.2025.25.2.0214

[41] N. C. Tiglao, N. M. Tiglao, E. Sanciangco, and M. A. Y. Tacderas, "Crowdsourcing and bus telematics for promoting fuel efficiency and eco-driving practices on the EDSA Busway," *Transportation Research Procedia*, vol. 82, pp. 450–458, 2025. https://doi.org/10.1016/j.trpro.2025.03.045

[42] P. Ruer *et al.*, "Improving driving behavior with an insurance telematics mobile application," in *Lecture Notes in Computer Science*, 2020. https://doi.org/10.1007/978-3-030-59506-7_27

[43] T. Bobrovskaya *et al.*, "Sample size for assessing a diagnostic accuracy of AI-based software in radiology," *Siberian Journal of Clinical and Experimental Medicine*, vol. 39, no. 4, pp. 122–130, 2024. https://doi.org/10.29001/2073-8552-2024-39-3-188-198

[44] C. Van Vliet *et al.*, "The safer driver app decreases mobile phone-induced distracted driving: Evidence from a randomized controlled trial," *Journal of Road Safety*, vol. 35, no. 2, pp. 3–13, 2024. https://doi.org/10.33492/JRS-D-24-2-2133200

[45] L. Meuleners, M. Fraser, M. Stevenson, and P. Roberts, "Personalized driving safety: Using telematics to reduce risky driving behaviour among young drivers," *Journal of Safety Research*, vol. 86, pp. 164–173, 2023. https://doi.org/10.1016/j.jsr.2023.05.007

[46] M. Stevenson *et al.*, "FEEDBACK trial: A randomized control trial to investigate the effect of personalized feedback and financial incentives on reducing the incidence of road crashes," *BMC Public Health*, vol. 23, p. 2035, 2023. https://doi.org/10.1186/s12889-023-16886-z

[47] D. Vankov, R. Schroeter, and A. Rakotonirainy, "Provisional drivers intend to speed less: The positive outcome for young drivers of a safe-driving app randomized trial," *Transportation Research Interdisciplinary Perspectives*, vol. 21, p. 100877, 2023. https://doi.org/10.1016/j.trip.2023.100877

[48] M. K. Delgado *et al.*, "Feedback and financial incentives for reducing cell phone use while driving: A randomized clinical trial," *JAMA Network Open*, vol. 7, no. 7, pp. 1–13, 2024. https://doi.org/10.1001/jamanetworkopen.2024.20218

[49] E. Papatheocharous, C. Kaiser, J. Moser, and A. Stocker, "Monitoring distracted driving behaviours with smartphones: An extended systematic literature review," *Sensors*, vol. 23, no. 17, p. 7505, 2023. https://doi.org/10.3390/s23177505

[50] M. S. Horswill, A. Hill, L. Silapurem, and M. O. Watson, "A thousand years of crash experience in three hours: An online hazard perception training course for drivers," *Accident Analysis & Prevention*, vol. 152, p. 105969, 2021. https://doi.org/10.1016/j.aap.2020.105969

[51] Z. Yang, Z. Zheng, J. Kim, and H. Rakha, "Eco-driving strategies using reinforcement learning for mixed traffic in the vicinity of signalized intersections," *Transportation Research Part C: Emerging Technologies*, vol. 165, p. 104683, 2024. https://doi.org/10.1016/j.trc.2024.104683

[52] X. Shi, J. Zhang, X. Jiang, J. Chen, W. Hao, and B. Wang, "Learning eco-driving strategies from human driving trajectories," *Physica A: Statistical Mechanics and Its Applications*, vol. 633, p. 129353, 2024. https://doi.org/10.1016/j.physa.2023.129353

[53] M. Hilario, P. Paredes, J. Mayhuasca, M. Liendo, and S. Martínez, "Evaluation of the impact of artificial intelligence on the systems audit process," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 15, no. 3, pp. 184–202, 2024. https://doi.org/10.58346/JOWUA.2024.I3.013

[54] A. Arouj and A. M. Abdelmoniem, "A systematic literature review of driver inattention and distraction," *International Journal of Interactive Mobile Technologies*, vol. 16, no. 18, pp. 4–22, 2022. https://doi.org/10.3991/ijim.v16i18.33075

[55] T.-K. Nguyen, V. T. Pham, D. V. Nguyen, and L. Do, "Development of a smartphone application for safe car driving using Google API and built-in sensor," *International Journal of Interactive Mobile Technologies*, vol. 14, no. 2, pp. 178–191, 2020. https://doi.org/10.3991/ijim.v14i02.11118

[56] Y. A. Alqudah, B. Sababha, E. Qaralleh, and T. Yousseff, "Machine learning to classify driving events using mobile phone sensors data," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 2, pp. 124–136, 2021. https://doi.org/10.3991/ijim.v15i02.18303

9 AUTHORS

Manuel Hilario is with the Federico Villarreal National University, Lima, Peru (E-mail: fhilario@unfv.edu.pe).

Pervis Paredes is with the Federico Villarreal National University, Lima, Peru (E-mail: pparedes@unfv.edu.pe).

Henry Maquera is with the National University of Central Peru, Huancayo, Peru (E-mail: hmaquera@uncp.edu.pe).

Shirley Martínez is with the Federico Villarreal National University, Lima, Peru (E-mail: 2020101086@unfv.edu.pe).

Jhosep Velasque is with the Federico Villarreal National University, Lima, Peru (E-mail: 2020101139@unfv.edu.pe).