International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 10 | Oct 2026 www.irjet.net p-ISSN: 2395-0072

CogniRoute: A Survey and Proposed Framework for Multi-Modal Driver Drowsiness Detection and Fatigue-Aware Route Optimisation

Mr. Avinash Changdev Taskar[1], Rohan Kshirsagar[2], Yash Deore[3], Jayesh Deore[4], Vedika Bhadane[5]

[1]Assistant Professor, Department of Computer Science and Engineering, Sandip University, Nashik, India
[2, 3, 4, 5]Student, Department of Computer Science and Engineering, Sandip University, Nashik, Maharashtra, India

---

Abstract - *Driver fatigue is a major contributor to traffic accidents worldwide. Still, most driver-monitoring systems use only a single sensing modality and do not adjust the trip in response to imminent fatigue. This paper surveys existing approaches to automated drowsiness detection, including ocular geometry proxies like Eye Aspect Ratio and PERCLOS, deep sequence models such as CNN-LSTM, and physiological and behavioural indicators such as PPG and steering-pattern analysis. It also surveys classical route-optimization algorithms such as Dijkstra's and A* search, identifying a research gap: fatigue-aware cost functions are rarely incorporated, and multi-modal fusion is seldom combined with a closed-loop routing response. Based on this survey, the paper proposes CogniRoute, a software-only prototype framework that combines a real webcam-based visual drowsiness score with realistically simulated, literature-calibrated physiological and behavioural channels to form a fused drowsiness index, which in turn modifies a fatigue-aware A* routing engine to suggest alternate, rest-friendly routes when the driver begins to feel drowsy. This revised paper specifies the fusion and routing formulations mathematically, names the datasets intended for evaluation, and sets out a concrete, reproducible evaluation plan with baselines, ablations, statistical analysis, robustness testing, and privacy considerations. As no experiments have yet been run, all quantitative claims below describe the intended methodology rather than achieved results.*

Key Words: driver drowsiness detection, eye aspect ratio, PERCLOS, CNN-LSTM, multimodal sensor fusion, fatigue-aware routing, A* search, road safety.

1. INTRODUCTION

Fatigue reduces driver alertness, increasing the risk of road accidents through loss of concentration and control in a manner comparable to alcohol consumption. Despite extensive research, most driver-monitoring systems today suffer from two shortcomings: they typically use only one sensing channel (usually a camera observing the driver's face) and they do not modify the trip in response to imminent fatigue. This paper makes two contributions: first, it surveys the literature on vision-based, physiological, and behavioural indicators of driver drowsiness, as well as classical route-optimization algorithms. Second, it proposes CogniRoute a software-only prototype framework that (a) extracts real visual features from a webcam, (b) fuses them with simulated physiological (PPG) and behavioural (steering/IMU) indicators into a single drowsiness index, and (c) uses this index to modify a routing algorithm to suggest alternate, fatigue-aware routes.

CogniRoute is developed by a four-member, time-constrained final-year project team. The staged, modular development strategy described in Section III was adopted because it is well suited to a small team working under a fixed academic deadline, not because the project is undertaken by a single person: it lets the team validate the vision pipeline, the fusion logic, and the routing engine independently before integrating them, distributing this work across members with different areas of ownership (vision and deep learning, signal simulation, routing and graph algorithms, backend/dashboard, and evaluation).

2. RELATED WORK

2.1 Vision-Based Drowsiness Detection

Much of the early work in this space centres on the geometry of the eye. Soukupová and Čech's Eye-Aspect-Ratio (EAR) [1] compares vertical to horizontal eye-opening distances between facial landmarks and remains a simple, effective way to flag blinking. Wierwille's PERCLOS [3] takes a related but distinct view, tracking the percentage of time the eye stays at least 80% closed over a sliding window rather than looking at individual blinks. Both rely on first finding the face reliably, a problem Viola and Jones [2] addressed with their cascade-classifier approach; today's systems more commonly use the MediaPipe Face Mesh framework [7] for the same purpose. What these geometric proxies miss, however, is how drowsiness unfolds gradually over minutes rather than in a single frame. This has pushed researchers toward sequence models Long-Short-Term-Memory networks [6] have been used to pick up on slowly changing blink frequency, duration, and the onset of micro-sleeps, usually paired with a CNN front-end that extracts spatial features from

cropped eye or face images, giving the CNN-LSTM architecture used on benchmarks such as NTHU-DDD [11]. The trade-off is that these models are data-hungry: they need substantial labelled sequence data to train well, which is one reason gradual-onset drowsiness detection remains harder to deploy than simple blink counting.

2.2 Physiological and Behavioural Indicators

Vision alone struggles in low light or when the driver's head is turned away from the camera, which is why physiological and behavioural signals have drawn attention as complementary channels. On the physiological side, heart-rate variability (HRV) derived from photoplethysmography (PPG) is a reasonably well-established fatigue indicator — parasympathetic activity tends to rise as alertness drops — and wearable datasets such as WESAD [13], which include real chest- and wrist-worn PPG/BVP recordings, offer a useful reference for what these signals actually look like statistically. On the behavioural side, steering-wheel position and the frequency of small corrective movements say something similar: a tired driver tends to correct less often but more abruptly to stay centred in the lane. UAH-DriveSet [14] is a helpful reference here too, since it records smartphone IMU and GPS data under explicitly labelled normal, drowsy, and aggressive driving. Neither channel needs a clear line of sight to the driver's face, which is precisely what makes them useful alongside a vision-based system rather than instead of one.

2.3 Route Optimization and Fatigue-Aware Routing

On the routing side, the underlying algorithms are much older and more settled. Dijkstra's algorithm [4] and A* search [5] remain the workhorses of commercial routing, usually sped up in practice with contraction hierarchies. Traffic-aware routing, which folds live congestion data into edge weights, is a fairly standard extension by now — but almost nothing in the literature adjusts the route itself as a driver's fatigue changes mid-trip. That gap is what this paper tries to address, by adding a driver-fatigue-aware penalty to A*'s edge weights: long, monotonous stretches are penalized and paths near rest-stop opportunities are rewarded once the driver is flagged as fatigued. Section III-B gives the exact formulation.

2.4 Research Gap

Vision-based approaches to drowsiness detection have been extensively studied but are single-channel and hence susceptible to lighting and occlusion challenges. Physiological and behavioural channels have been investigated, but there is limited work on their fusion with vision-based channels in a single deployable system. Fatigue-aware routing, when considered at all, is often a static property rather than a dynamically updated function of the driver's current drowsiness level. CogniRoute addresses this gap by (a) combining three channels into a single, explicitly-defined drowsiness index and (b) using this index to modify the routing algorithm's edge costs to suggest rest-stop-friendly alternate routes.

3. PROPOSED METHODOLOGY AND SYSTEM ARCHITECTURE

CogniRoute is a software-only prototype for now, requiring only a webcam as a sensor. Physiological and behavioural channels are simulated in software, calibrated against real recordings as described in Section IV, to allow testing of the vision-fatigue fusion logic in isolation. The system is designed to be modular, and the vision pipeline can be replaced with real MAX30100 (PPG) and MPU6050 (IMU) sensor drivers in a future iteration.

The proposed system consists of six modules. First, a webcam captures video of the driver, which is processed using MediaPipe Face Mesh to extract visual features such as Eye-Aspect-Ratio (EAR), Mouth-Aspect-Ratio (MAR), and PERCLOS. Second, a CNN-LSTM classifier is trained to predict a continuous vision-based drowsiness score V from these features. Third, simulated physiological (PPG-derived) and behavioural (steering/IMU-derived) signals are generated as described in Section IV to produce scores P and B respectively. Fourth, a fusion module combines the three channels' scores into a single weighted drowsiness index D, with configurable, justified weights (Section III-A). Fifth, when D crosses a threshold, an on-screen alert is triggered and a fatigue-aware routing algorithm (Section III-B) suggests alternate routes that minimize monotony and maximize rest opportunities. Finally, a logging module stores pseudonymized trip details in a SQLite database and displays relevant analytics on a web dashboard.

3.1 Multimodal Signal Normalization and Fusion

Each channel produces a raw score on its own scale (e.g., PERCLOS as a percentage, HRV-derived features in physiological units, steering-correction rate in corrections per minute). Every raw score x is first min–max normalized to a common [0,1] range using:

$$x_{norm} = (x - x_{min}) / (x_{max} - x_{min}) \quad (1)$$

where xmin and xmax are the calibration-set minimum and maximum for that specific feature (e.g., derived from the labelled training partitions of NTHU-DDD/YawDD for

the visual score, and from the simulator's calibrated operating range for the physiological and behavioural scores). Let V, P, B ∈ [0,1] denote the normalized visual, physiological, and behavioural drowsiness scores. The fused drowsiness index D is defined as the convex combination:

$$D = wvV + wpP + wbB, \text{ with } wv + wp + wb = 1 \quad (2)$$

Weight-selection rationale: the weights are not fixed arbitrarily. An initial prior is set from the literature reviewed in Section II the visual channel is treated as the most directly validated indicator under adequate lighting (wv = 0.5 initially), with the physiological and behavioural channels contributing supporting evidence (wp = wb = 0.25 initially). These initial weights are then tuned by a grid search over the simplex wv + wp + wb = 1 on a held-out validation split, selecting the combination that maximizes F1-score against ground-truth drowsiness labels (Section VII). When a channel becomes unavailable or unreliable for example, the visual channel under camera failure or severe occlusion its weight is set to zero and the remaining weights are re-normalized to sum to one, so the fusion degrades gracefully rather than failing (Section VII-H).

3.2 Fatigue-Aware A* Routing Formulation

The routing engine retains the standard A* evaluation function:

$$f(n) = g(n) + h(n) \quad (3)$$

where g(n) is the accumulated cost from the origin to node n and h(n) is an admissible heuristic (straight-line distance to the destination), exactly as in unmodified A*. The proposed contribution modifies how the edge cost that feeds g(n) is computed. For an edge (u,v) with baseline cost c(u,v) (travel time or distance), the fatigue-aware edge cost is:

$$c'(u,v) = c(u,v) \cdot (1 + \lambda D \cdot D \cdot M(u,v)) - \lambda R \cdot D \cdot S(v) \quad (4)$$

Here D ∈ [0,1] is the current fused drowsiness index from Eq. (2); M(u,v) ∈ [0,1] is a monotony score for edge (u,v), computed from segment length, exit/intersection density, and curvature variance along that edge (long, straight, low-exit-density stretches score close to 1); S(v) ∈ [0,1] is a rest-stop suitability score for the node v that the edge leads into (Section III-C); and λD and λR are non-negative tunable weights that respectively control the strength of the fatigue penalty and the strength of the rest-stop reward. The term c′(u,v) is clamped to a minimum of a small positive ε to preserve non-negativity, which is required for A*'s optimality guarantee to hold with the given heuristic. To keep an unmodified-A* baseline available for comparison (Section VII-B), the fatigue penalty is only activated once D exceeds a threshold θ (e.g., θ = 0.6): for D < θ, λD and λR are effectively set to 0 and Eq. (4) reduces to c′(u,v) = c(u,v), i.e., standard A*.

3.3 Rest-Friendly Route Identification

Rest-stop suitability S(v) is computed offline for every node v in the road graph from static points-of-interest (POI) data (e.g., OpenStreetMap tags such as highway=rest_area, amenity=fuel, and tourism=parking) obtained through the routing backend (Section VI). For each node, S(v) is a weighted, normalized count of qualifying POIs found within a fixed search radius r of that node, with category-specific weights reflecting suitability for a fatigued driver to stop safely (a signed, staffed rest area is weighted more heavily than an unattended parking bay). S(v) is precomputed and cached against the road graph so that it can be looked up in constant time during route search rather than recomputed per query, keeping the added routing latency bounded (Section VII-G). During route scoring, S(v) enters the edge cost through Eq. (4), so that, once a driver is flagged as fatigued, the search is biased away from long monotonous stretches and toward paths whose nodes have higher rest-stop suitability, without requiring a hard detour constraint.

4. SIMULATION DESIGN AND REALISM

Because CogniRoute is currently software-only, the physiological (PPG-derived) and behavioural (steering/IMU-derived) channels are synthetic. This section specifies the simulation assumptions, ranges, and noise model, and states how simulation validity is limited.

4.1 Physiological Signal Simulation

The physiological score P is generated from a synthetic heart-rate-variability (HRV) proxy under the standard assumption, drawn from the literature reviewed in Section II-B, that reduced alertness is associated with declining HRV complexity and a shift toward parasympathetic dominance. The simulator models the underlying HRV proxy as a slow, monotonic drift over elapsed time-on-task (representing accumulating fatigue) superimposed on a circadian-like sinusoidal component and additive Gaussian sensor noise, i.e., a signal of the form drift(t) + sinusoid(t) + $N(0, \sigma p^2)$, subsequently mapped to [0,1] via Eq. (1). Both the amplitude of the drift and the noise variance σp^2 are moment-matched (mean and variance) against baseline- and stress-condition BVP/PPG segments from the WESAD dataset [13] rather than chosen arbitrarily, so

that the simulated signal's operating range and noise floor are anchored to a real wearable-PPG recording.

4.2 Behavioural Signal Simulation

The behavioural score B is derived from a synthetic steering/IMU proxy. Consistent with the assumption that fatigued drivers make larger, less frequent lane-keeping corrections, steering-angle-like time series are generated as a mean-reverting (Ornstein–Uhlenbeck) process whose correction interval and correction magnitude scale with a latent fatigue level, with additive Gaussian noise representing IMU sensor noise. The mapping between latent fatigue level and correction interval/magnitude, together with the noise-floor parameters, is calibrated by moment-matching against the drowsy- and normal-labelled trips of the UAH-DriveSet naturalistic driving dataset [14], which records comparable smartphone-IMU signals under an explicit drowsy label.

4.3 Assumptions, Limitations, and Threats to Validity

- Assumptions: fatigue produces a monotonic, gradually accumulating drift in the underlying physiological and behavioural proxies rather than abrupt jumps; sensor noise is approximately Gaussian and stationary within a short window.

- Ranges: after normalization, all three channel scores are bounded to [0,1]; the moment-matching procedure in IV-A and IV-B constrains simulated means/variances to lie within the envelope observed in the reference real datasets rather than allowing unconstrained parameter choices.

- Limitations: the simulator cannot reproduce inter-subject physiological variability, motion or skin-tone artifacts affecting real PPG, confounds such as caffeine, medication, or illness, or real vehicle–road dynamics affecting IMU signals.

- Threat to validity: because P and B are synthetic, results obtained purely from simulated channels (Section VII) demonstrate the correctness of the fusion and routing logic, not the on-road efficacy of physiological or behavioural drowsiness sensing; end-to-end efficacy claims are deferred until the simulated channels are replaced with the real MAX30100/MPU6050 sensors noted in Section IX.

5. PRIVACY AND RESPONSIBLE DATA USE

CogniRoute processes webcam video to extract fatigue-related facial features. To limit privacy exposure:

- (i) Video frames are processed in memory by the vision pipeline in real time, and only derived low-dimensional features (EAR, MAR, PERCLOS, and the CNN-LSTM drowsiness score) are retained by default — raw frames are discarded immediately after feature extraction and are not written to disk;

- (ii) If raw-frame logging is ever enabled for debugging, it is opt-in, stored locally with encryption at rest, and subject to a short, user-configurable retention period;

- (iii) The system performs fatigue-state estimation only and does not perform biometric identity matching against any reference database;

- (iv) Any participant whose video is recorded for evaluation purposes provides informed consent under the relevant institutional research-ethics process, with the right to withdraw their data;

- (v) The public datasets used for evaluation (NTHU-DDD, YawDD, WESAD, UAH-DriveSet) are used strictly under their respective research-only licenses and are not redistributed; and

- (vi) Trip logs stored in the SQLite database (route, timestamps, drowsiness index) are keyed by a pseudonymous session identifier rather than by personally identifying information, and can be deleted by the user from the dashboard.

6. TOOLS AND TECHNOLOGIES

The vision pipeline is built in Python on top of OpenCV [8] and MediaPipe [7], with the CNN-LSTM classifier trained using TensorFlow/Keras [9]. The physiological and behavioural signal generators from Section IV are implemented in NumPy/SciPy, calibrated against the reference datasets described there. A FastAPI service exposes the system, backed by a SQLite database for logging trips in the prototype (Firebase would be a natural swap for a production deployment). For routing, the plan is to sit on top of the Open Source Routing Machine (OSRM) [10] or a comparable free routing API, using it as the underlying road graph and adding the precomputed rest-stop suitability scores from Section III-C on top. The dashboard itself is left flexible — a Streamlit app if time is tight, or a small React.js single-page app if there's room to build one.

7. EVALUATION PLAN

No experiments have been run yet; this section specifies a concrete, reproducible plan rather than reporting results. Table I summarizes the four evaluation stages, each detailed below.

7.1 Dataset and Preprocessing

The vision-based drowsiness classifier will be trained and evaluated on the NTHU Drowsy Driver Detection dataset (NTHU-DDD) [11], which provides infrared video from 36 subjects across normal, yawning, slow-blink, microsleep, and laughing conditions under day and night lighting with frame-level drowsiness annotation, and cross-validated on the YawDD dataset [12], which provides 107-subject RGB video labelled for normal, talking, and yawning behaviour from dashboard and mirror camera angles. NTHU-DDD is selected because its frame-level, multi-condition labels (including night/IR and eyewear variation) directly support the robustness evaluation in Stage 4, while YawDD provides an independent, differently-collected dataset to test generalization across camera angle and population. Preprocessing follows the face-crop and landmark-extraction pipeline used in prior CNN-LSTM work on NTHU-DDD [11]: MediaPipe Face Mesh landmark extraction, EAR/MAR/PERCLOS feature computation per frame, and windowing into fixed-length sequences for the CNN-LSTM classifier, with the official training/validation/test subject splits preserved to avoid subject leakage.

7.2 Baseline Methods

- Vision-only detection: the CNN-LSTM classifier operating on V alone, without physiological/behavioural fusion.
- Threshold baseline: simple EAR/PERCLOS thresholding, without any learned classifier.
- Standard A*: unmodified A* search (Eq. 3 with $c' = c$), with no fatigue penalty or rest-stop reward.
- Routing without the fatigue penalty: A* using Eq. (4) with $\lambda D = \lambda R = 0$, i.e., equivalent to standard A*, included as the ablation counterpart to the full fatigue-aware configuration.

7.3 Ablation Study

The fusion module is evaluated in four configurations: visual only (V); visual + physiological (V+P); visual + behavioural (V+B); and the full multimodal fusion (V+P+B). Each configuration is evaluated with the same weight-tuning procedure described in Section III-A restricted to the active channels, isolating the marginal contribution of each channel to detection performance and to routing outcomes.

7.4 Quantitative Metrics

Drowsiness-detection metrics: accuracy, precision, recall, F1-score, and the full confusion matrix, computed per configuration and per robustness condition. Routing metrics: total distance, total travel time, cumulative fatigue exposure (the time-integral of D along the trip), rest-stop exposure (count and spacing of passed rest-suitable nodes), and route (re)computation time.

7.5 Statistical Analysis

Each experimental configuration is repeated across $N = 30$ runs (k-fold cross-validation folds for the classifier; independent simulated-noise seeds and held-out trips for the routing/fusion experiments), and results are reported as mean ± standard deviation. Normality of the resulting metric distributions is checked with the Shapiro–Wilk test; where normality holds, paired t-tests are used to compare fatigue-aware A* against standard A* and each ablation configuration against full fusion, otherwise the Wilcoxon signed-rank test is used. Significance is assessed at $\alpha = 0.05$ with a Bonferroni correction applied across the family of metric comparisons in Section VII-D.

7.6 Safety vs. Travel-Time Trade-off

Fatigue-aware routing is expected to increase distance and/or travel time relative to the shortest-path baseline whenever λD and λR are active. This trade-off is quantified as the percentage increase in distance/travel time relative to standard A*, plotted against the corresponding reduction in cumulative fatigue exposure and improvement in rest-stop spacing, while sweeping λD and λR. This trade-off curve lets an operator choose a bounded travel-time-increase ceiling (e.g., capped at a configurable percentage) that is deemed an acceptable cost for reduced continuous, monotonous driving exposure, consistent with the road-safety rationale in Section II-C; the paper does not assume a specific ceiling is correct and treats it as a configurable operating point to be justified against future on-road data.

7.7 Computational Latency

End-to-end latency is measured as the sum of: (i) per-frame visual-inference latency, from webcam capture to the CNN-LSTM drowsiness score; (ii) fusion-module latency, from the three channel scores to D; and (iii) route-recomputation latency, from D crossing the threshold θ to a new route being returned. Each stage is timestamped with wall-clock measurements averaged over repeated runs; the target is a total pipeline latency under one second, consistent with real-time alerting.

7.8 Robustness Evaluation

Robustness is evaluated under: low-light/night conditions and eyewear variation (using the corresponding NTHU-DDD subsets); simulated face occlusion; missing or noisy physiological signal (simulated dropped-sample gaps and increased noise variance in the calibrated simulator of Section IV-A); simulated camera failure, in which the visual weight wv is set to zero and the physiological/behavioural weights are re-normalized as described in Section III-A; and increased sensor-noise floors for the behavioural channel. For each condition, the resulting degradation in detection F1/accuracy and in the stability of D is reported against the Stage 1/2 baselines.

Table -1: Suggested Evaluation Structure

Stage	Experiments / Comparisons	Main Measurements
1. Detection	NTHU-DDD and YawDD (VII-A); vision vs. EAR/PERCLOS baseline	Accuracy, precision, recall, F1, latency
2. Ablation	Visual only; V+P; V+B; full V+P+B fusion	F1/recall gain, sensor dropout resilience
3. Routing	Standard A* vs. fatigue-aware A* (Eq. 4); rest-stop terms	Distance, travel time, fatigue exposure, compute time
4. Robustness	Low light IR, occlusion, eyewear, sensor noise/failure	Degradation vs. baseline; re-normalization

8. CHALLENGES

The most significant challenge in a software-only prototype is producing simulators for the physiological and behavioural channels that are realistic enough to meaningfully validate the fusion module (Section IV) while remaining anchored to real reference data rather than arbitrary constants. A related challenge is the reliability of the vision channel under low-light or occluded conditions, a known limitation of vision-based drowsiness detection that is directly targeted by the robustness evaluation in Section VII-H. The training data for the CNN-LSTM classifier may also be limited or imbalanced across drowsy/alert classes, requiring careful handling (e.g., class weighting) to avoid bias. Finally, the routing algorithm depends on an external API (Section VI), which may introduce latency or downtime that must be accounted for in the latency measurements of Section VII-G.

9. FUTURE ENHANCEMENTS

Future work includes replacing the simulated physiological and behavioural channels with real MAX30100 (PPG) and MPU6050 (IMU) sensors on a Raspberry Pi to create a full hardware prototype, re-running the evaluation plan of Section VII against real, rather than simulated, signals. The system can be enhanced to support fleet-level monitoring with a centralized operator dashboard. Other physiological channels, such as consumer-grade EEG headbands, can also be explored for richer fatigue estimation. Reinforcement learning can be used to adapt the routing weights λD and λR to long-term driver fatigue patterns rather than a single snapshot. Finally, the system can be extended to work with vehicle CAN-bus data and OEM Advanced Driver-Assistance Systems (ADAS) for closed-loop safety-critical interventions, subject to the additional safety validation such integration would require.

10. CONCLUSION

This paper surveyed the literature on vision-based, physiological, and behavioural indicators of driver drowsiness, as well as classical route-optimization algorithms, and identified a research gap: existing systems rarely combine multiple indicators of drowsiness, and routing algorithms rarely modify the suggested route in response to a driver's changing fatigue level. The paper proposed CogniRoute — a software-only prototype framework that combines a real webcam-based visual drowsiness score with literature-calibrated simulated physiological and behavioural channels into an explicitly-defined fused drowsiness index (Eq. 2), and a fatigue-aware A* routing formulation (Eq. 4) that biases route selection toward rest-friendly, less monotonous paths as this index rises. The paper specified the normalization and fusion mathematics, the simulation assumptions and their calibration against real reference datasets, and a concrete, reproducible evaluation plan with named datasets, baselines, ablations, statistical analysis, latency measurement, robustness testing, and privacy safeguards. As no experiments have yet been conducted, these contributions describe a proposed framework and evaluation methodology rather than validated real-world performance; by requiring only a webcam as a sensor and being upgradable to real physiological and behavioural channels, the design aims to turn passive driver monitoring into a preventive road-safety approach that a time-constrained final-year project team can realistically build and evaluate in stages.

ACKNOWLEDGEMENT

The authors would like to thank Sandip University, Nashik, and the Department of Computer Science and Engineering for providing the academic infrastructure, computing facilities, and guidance necessary to conduct this survey and architectural design.

REFERENCES

[1] T. Soukupová and J. Čech, "Real-Time Eye Blink Detection Using Facial Landmarks," in Proc. 21st Computer Vision Winter Workshop (CVWW), 2016.

[2] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2001.

[3] W. W. Wierwille, "Overview of Research on Driver Drowsiness Definition and Detection," in Proc. 14th Int. Technical Conf. Enhanced Safety of Vehicles (ESV), 1994.

[4] E. W. Dijkstra, "A Note on Two Problems in Connexion with Graphs," Numerische Mathematik, vol. 1, pp. 269-271, 1959.

[5] P. E. Hart, N. J. Nilsson, and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," IEEE Trans. Systems Science and Cybernetics, vol. 4, no. 2, pp. 100-107, 1968.

[6] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[7] Google AI, "MediaPipe Documentation," developers.google.com/mediapipe, accessed 2026.

[8] OpenCV Library Documentation, docs.opencv.org, accessed 2026.

[9] TensorFlow Documentation, tensorflow.org, accessed 2026.

[10] Open Source Routing Machine (OSRM) Project Documentation, project-osrm.org, accessed 2026.

[11] C.-H. Weng, Y.-H. Lai, and S.-H. Lai, "Driver Drowsiness Detection via a Hierarchical Temporal Deep Belief Network," in Proc. ACCV 2016 Workshop on Driver Drowsiness Detection, 2016.

[12] S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, and B. Hariri, "YawDD: A Yawning Detection Dataset," in Proc. ACM Multimedia Systems Conf. (MMSys), 2014.

[13] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection," in Proc. 20th ACM Int. Conf. Multimodal Interaction (ICMI), 2018.

[14] E. Romera, L. M. Bergasa, and R. Arroyo, "Need Data for Driver Behaviour Analysis? Presenting the Public UAH-DriveSet," in Proc. IEEE Int. Conf. Intelligent Transportation Systems (ITSC), 2016.

BIOGRAPHIES

Mr. Avinash Changdev Taskar: Assistant Professor, Department of Computer Science and Engineering, Sandip University, Nashik, Maharashtra, India. Area of interest includes Machine Learning, Artificial Intelligence, and Intelligent Transportation Systems.

Rohan Kshirsagar: Undergraduate Scholar, Department of Computer Science and Engineering, Sandip University, Nashik, Maharashtra, India. Research interests include Computer Vision, Deep Learning, and Embedded Systems.

Yash Deore: Undergraduate Scholar, Department of Computer Science and Engineering, Sandip University, Nashik, Maharashtra, India. Focuses on Machine Learning pipelines and signal processing.

Jayesh Deore: Undergraduate Scholar, Department of Computer Science and Engineering, Sandip University, Nashik, Maharashtra, India. Specializes in graph algorithms, spatial routing frameworks, and backend architectures.

Vedika Bhadane: Undergraduate Scholar, Department of Computer Science and Engineering, Sandip University, Nashik, Maharashtra, India. Specializes in data analysis, dashboard systems, and multimodal evaluation.