Driver Drowsiness Detection Using Computer Vision with Adaptive Methods on Android

Paulo Roberto Varjal de Melo, Lucas Maciel Sousa, Denilson José do Bom Jesus Silva de Lima, and Fernando Buarque de Lima Neto

[1] PPGEC, Faculdade de Engenharia (POLI), Universidade de Pernambuco (UPE), Recife, PE, Brazil

{prvm,lms3,djbjsl,fbln}@ecomp.poli.br

Abstract. *Traffic accidents associated with driver drowsiness represent a relevant challenge for road safety. This work presents an Android system for real-time drowsiness detection, with on-device processing, using the smartphone's front camera. The approach integrates multiple computer vision indicators (EAR, MAR, PERCLOS, blink frequency, blink duration, and yawn frequency), where EAR and MAR drive blink/eye-closure and yawn event detectors whose outputs feed the four fused metrics (PERCLOS, blink frequency, blink duration, and yawn frequency). The system also features user personalization, based on initial calibration and the use of data labeled by the user to adjust individual reference parameters. An evaluation with 40 videos from the UTA-RLDD dataset, corresponding to 20 users, using a temporal protocol, obtained accuracies of 97.1% for the alert state and 98.9% for the drowsy state, demonstrating the technical feasibility of the mobile application.*

1. Introduction

Road safety remains a major contemporary challenge: the World Health Organization estimates 1.19 million preventable deaths per year worldwide [World Health Organization 2026], costing around 3% of GDP. In Brazil, which ranks among the countries with the highest absolute numbers of road traffic deaths [World Health Organization 2024], open data from the Federal Highway Police indicate that more than 25% of highway deaths are associated with driver drowsiness [Federal Highway Police (Brazil) 2026], with estimated annual losses of BRL 50 billion [Instituto de Pesquisa Econômica Aplicada 2020].

Driver drowsiness compromises essential functions for safe driving, including attention, vigilance, and reaction time. Studies indicate that its effects may be comparable to driving under the influence of alcohol, significantly increasing crash risk [World Health Organization 2026, Jeong et al. 2022]. In this context, computer vision techniques have been widely investigated because they allow non-invasive monitoring of physiological–behavioral signals related to drowsiness, such as eye opening, prolonged eye closure, blink frequency, mouth opening, yawning occurrence, and blink duration.

This article presents an Android system for continuous driver drowsiness detection using the smartphone's front camera. Rather than isolated event detection, the system

integrates multiple physiological–behavioral signals (*EAR*, *MAR*, *PERCLOS*, blink frequency, yawn frequency, and average blink duration) within a unified temporal structure, with on-device processing and user personalization.

The main contribution of this work lies in the combination of four elements: (i) real-time execution directly on the mobile device, without reliance on external infrastructure; (ii) temporal fusion of multiple complementary indicators through a sliding window; (iii) rapid same-session adaptation of reference parameters based on initial calibration and user-labeled data; and (iv) an initial experimental evaluation on videos from the UTA-RLDD dataset, using a temporal protocol that separates learning, forgetting, and testing phases.

The results are presented as initial evidence of technical feasibility under controlled conditions, motivating future evaluations in realistic driving scenarios.

2. Background and Related Work

Driver drowsiness compromises cognitive, motor, and behavioral functions essential for safe driving, including attention, vigilance, and reaction time [Bittencourt et al. 2005, Singh et al. 2023, Das and Bhargava 2023, El-Nabi et al. 2024, Bajaj et al. 2021].

Computer vision-based methods are widely used for drowsiness monitoring because they allow non-invasive analysis of facial and behavioral signs [Bergasa et al. 2006, Albadawi et al. 2022, Abe 2023, El-Nabi et al. 2024], extracting indicators such as eye opening, blink frequency, mouth opening, and blink duration through facial landmark models such as MediaPipe-based pipelines [Lugaresi et al. 2019, Baul et al. 2025, Florez et al. 2024].

Frequently employed indicators include the *Eye Aspect Ratio* (EAR) [Soukupová and Čech 2016], *Mouth Aspect Ratio* (MAR) [Florez et al. 2024], and *Percentage of Eye Closure* (PERCLOS) [Dinges et al. 1998, Abe 2023], complemented by blink frequency, yawn frequency, and blink duration [Kolus 2024, Schleicher et al. 2008, Häkkänen et al. 1999].

Isolated visual events may be insufficient to characterize the driver's state [Das and Bhargava 2023, Kashevnik et al. 2020, Albadawi et al. 2022], since blinks and mouth openings can occur due to transient factors [Albadawi et al. 2022, El-Nabi et al. 2024], while yawn absence does not guarantee alertness [Yang et al. 2021]. Temporal analysis mitigates this by aggregating recent evidence and smoothing transient oscillations [Ghoddoosian et al. 2019, Bergasa et al. 2006], with weights adjustable through constrained optimization such as *SLSQP* [Kraft 1994, Joshy and Hwang 2024]. Furthermore, individual variability in blink patterns, facial expressions, and fatigue response [Arakawa 2021, Albadawi et al. 2022, Kolus 2024] motivates personalized reference parameters.

Dasgupta et al. [Dasgupta et al. 2019] proposed a three-stage Android system using PERCLOS from the front camera, speech-based voiced-to-unvoiced ratio, and a touch-response verification for final classification.

This work proposes the temporal, personalized, and on-device integration of these indicators on Android. *EAR* and *MAR* serve as event detectors for blinks and yawns, respectively, whose outputs feed the four directly fused metrics: *PERC-*

LOS, blink frequency, blink duration, and yawn frequency. Compared to Dasgupta et al. [Dasgupta et al. 2019], which relied on a single behavioral indicator, the proposed approach integrates multiple complementary signals in a sliding temporal window with personalized references (Section 3).

3. Proposed System

This section describes the proposed Android system for real-time driver drowsiness detection using the smartphone's front camera. The system extracts multiple physiological–behavioral indicators from the driver's face, analyzes them within a sliding temporal window, and estimates the driver's operational state as **Alert** or **Drowsy**.

The system classifies the driver into two operational states: **Alert**, when the driver maintains attention and reaction capacity compatible with safe driving; and **Drowsy**, a higher-risk condition associated with physiological and behavioral changes, reduced attention, vigilance, and response capacity. Its general flow is illustrated in Figure 1.

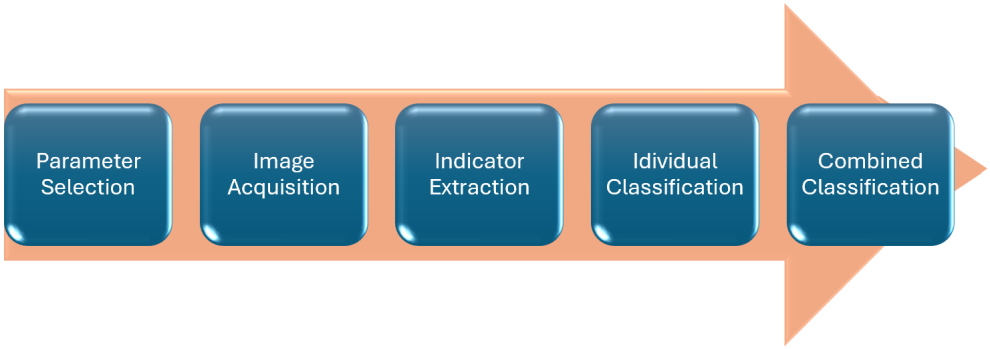


Figure 1. General operation flow of the drowsiness detection system.

3.1. Selection, Calibration, and Parameter Personalization

The system supports two modes: registered (with personalization) and anonymous. Both modes require an initial 5-second calibration to estimate the user's baseline *EAR* value, distinguishing open-eye from closed-eye conditions.

For registered users, manual labeling during use ("Alert" or "Drowsy") drives progressive personalization: the system stores aggregated indicator statistics and adjusts per-indicator reference ranges, centroids, and comparison values. Labels from one user are never used for another.

3.2. Image Acquisition and Indicator Extraction

Monitoring is performed through continuous image capture by the front camera of the Android smartphone. All processing required for indicator extraction and drowsiness estimation is executed on the device itself, without reliance on external infrastructure.

To obtain the indicators used in classification, the system employs MediaPipe Face Landmarker, accessed through the MediaPipe Tasks API [Google AI Edge 2025]. This model estimates facial landmarks in real time from images or video streams, enabling the computation of metrics related to the driver's alertness level.

3.3. Sliding Temporal Window and Continuous Inference

Classification is updated once per second using a 2-minute sliding window: each update incorporates the newest observation and discards the oldest, as illustrated in Figure 2. This strategy focuses the decision on recent behavior, diluting brief oscillations while capturing progressive drowsiness trends.

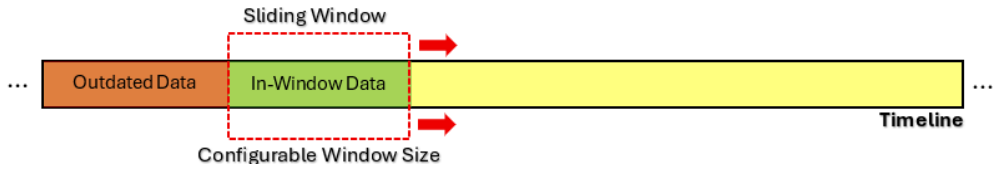


Figure 2. Representation of the sliding temporal window used in continuous inference. Old data are discarded, while new observations become part of the analysis window.

3.4. Individual Indicator Classification

Each indicator is computed in real time and analyzed within the sliding window against per-indicator reference ranges for the "Alert" and "Drowsy" states. For registered users, manually labeled data adjust these references to capture driver-specific patterns (e.g., natural blink frequency, typical *PERCLOS*); for anonymous users, general references are used.

The individual classification of each indicator produces a probability associated with the alert state. The following formal definitions underpin the indicators used in this work.

The *Eye Aspect Ratio* (EAR) quantifies eye opening from six facial landmarks [Soukupová and Čech 2016]:

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2 \cdot \|p_1 - p_4\|} \tag{1}$$

where p_1–p_6 are the six per-eye landmarks returned by MediaPipe. A similarly structured *Mouth Aspect Ratio* (MAR) is computed from mouth landmarks to support yawn detection [Florez et al. 2024]. The *Percentage of Eye Closure* (PERCLOS) measures the fraction of time the eyes remain closed within a temporal window [Dinges et al. 1998]:

$$PERCLOS = \frac{t_{closed}}{t_{total}} \tag{2}$$

where t_{closed} is the accumulated duration of eye-closure frames and t_{total} is the window duration.

Each indicator is first classified against per-indicator reference ranges for the "Alert" and "Drowsy" states. When the observed value falls clearly within one reference range, the indicator receives a high-confidence classification for that state. When the ranges overlap, the per-indicator alert probability is estimated from the Euclidean distance to the class centroids:

$$P_A = 1 - \frac{d_{obs,A}}{d_{obs,A} + d_{obs,D}} \tag{3}$$

where $d_{obs,A}$ and $d_{obs,D}$ are the distances from the observed (window-aggregated) value to the centroids of the Alert and Drowsy reference classes, respectively. When the observed value is clearly within one reference range, P_A is saturated to 1.0 or 0.0.

Each indicator thus provides partial evidence about the driver's state. These individual pieces of evidence are combined in the global classification, as described next.

3.5. Combined Classification and Fusion Weights

The global classification is obtained by fusing the probabilities produced by the individual indicators. For this purpose, a weighted average of the alert probabilities provided by the indicators is used, on a scale from 0% to 100%. This fusion seeks to combine complementary signals, reducing dependence on a single indicator and allowing different manifestations of drowsiness to contribute to the final decision.

Global fusion weights (3 for blink frequency, blink duration, and *PERCLOS*; 5 for yawn presence; 1 for yawn absence) were defined a priori based on preliminary development tests, before the evaluation described in Section 4. The asymmetric yawn treatment reflects that yawn presence is strong evidence of drowsiness, while absence does not guarantee alertness [Yang et al. 2021].

After the tests were completed, the indicator values and the classifications generated second by second were used in a complementary optimization analysis. This optimization, based on SLSQP, aimed to evaluate whether the empirical weight configuration was close to a combination that maximized performance on the evaluated set. The SLSQP analysis was computed on the same test outputs and therefore does not constitute independent validation; it only indicates that the empirical weights were close to a configuration maximizing accuracy on this specific set. Thus, the main results reported in this work use the empirical weights, while the optimized weights are presented as a posterior calibration analysis of the fusion strategy. This optimization does not alter the temporal protocol, the indicators used, the personalized parameters of each user, or the continuous inference mechanism of the system.

Figure 3 presents the Android application running in real time on an authorized test video. The original system interface is in Portuguese: "Piscadas" corresponds to blink frequency, "Bocejos" to yawn frequency, "DMP" to average blink duration, and "Sonolento" to the state classified as drowsy. The global classification shown at the bottom results from the weighted fusion of the indicators computed within the sliding temporal window.

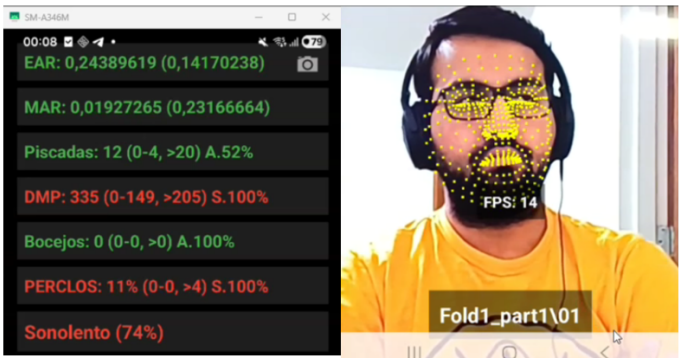


Figure 3. Android application screen during real-time execution on a video of Subject 01 from the UTA-RLDD dataset, whose image is indicated as publishable.

3.6. Validation and Tests

The system was designed for real-time use, with images continuously acquired from the smartphone's front camera and processed directly on the Android device. In this study, videos from the UTA-RLDD dataset were used as a controlled input source, enabling evaluation under reproducible conditions while preserving the same processing flow, temporal window, personalization mechanism, and classification procedure.

The system also targets longitudinal use, in which users accumulate labeled samples of the "Alert" and "Drowsy" states to progressively build more robust personalized references. However, evaluating this scenario would require long videos captured under real or simulated driving conditions. As such datasets are still limited, the UTA-RLDD dataset [Ghoddoosian et al. 2019] was adopted for an initial feasibility assessment, using videos of approximately 10 minutes labeled as "Alert" and "Drowsy"

Each selected user has two videos, one per state. Given their limited duration, the protocol evaluates the method under restrictive conditions, with few samples for personalization. For each user, only three windows labeled as "Alert" and three as "Drowsy" were used to estimate personalized references before testing.

A temporal protocol with five consecutive 2-minute segments was adopted (Figure 4). The first three segments (0–6 minutes) were used for learning: segments from the "Alert" video contributed to alert references, and those from the "Drowsy" video to drowsy references. The fourth segment (6–8 minutes) served as a forgetting phase, resetting moving-average histories to reduce temporal leakage. The fifth segment (8–10 minutes) was reserved for testing.

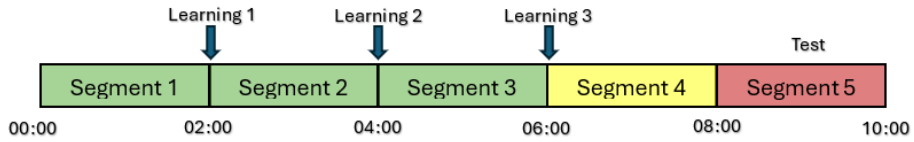


Figure 4. Temporal evaluation protocol adopted in this study, with learning, forgetting, and testing phases.

During testing, classification is continuous: each final segment produces 120 classifications, one per second, compared with the video's global label.

Although more restrictive than the intended use, this setup allows assessing the system's ability to learn user-specific patterns quickly and classify subsequent segments without overlap between learning and testing. Results should be interpreted as initial evidence of feasibility and personalization potential, not as definitive validation under real driving conditions.

Since the dataset was not recorded in real or simulated driving, only suitable videos were selected. Users with videos shorter than 10 minutes or with prolonged downward gaze, e.g., using a phone or computer, were excluded. The binary formulation (Alert vs. Drowsy) was adopted because the system targets two operational states; the intermediate "low vigilance" class in UTA-RLDD has no corresponding label in the application, and videos in that class would be ambiguous for the global video label used in this protocol. The system is intended for healthy individuals and may not be suitable for people with sleep disorders or atypical facial features.

3.7. Reproducibility Parameters

Table 1 summarizes the key parameters and thresholds used in the system, to support reproducibility.

Table 1. Key system parameters for reproducibility.

Parameter	Value
Sliding window duration	120 s
Update rate	1 Hz (1 classification/s)
Calibration duration	5 s
Learning windows per class	3 (2 min each)
Forgetting phase	2 min (6–8 min)
Test phase	2 min (8–10 min)
EAR landmark set	6 per-eye points (MediaPipe)
Eye-closure detection	EAR below threshold (calibrated)
Yawn detection	MAR above threshold (calibrated)
Fusion weights (empirical)	Blink freq. = 3, PERCLOS = 3, Blink duration = 3, Yawn = 5/1

4. Analysis of the Results

The experimental evaluation of this work focused on assessing the feasibility of an Android application executed directly on the device, with user-specific parameter adaptation and continuous classification from public videos. Although previous studies have used the UTA-RLDD dataset for drowsiness detection [Ghoddoosian et al. 2019, Tamanani et al. 2021, Essahraui et al. 2025], they differ from the present work in problem formulation, evaluation protocol, processing architecture, personalization strategy, and reported metrics. Therefore, the results are analyzed mainly through the behavior of the individual indicators and the combined classification within the adopted temporal protocol, rather than as a strict benchmark comparison.

For context, Tamanani et al. [Tamanani et al. 2021] report 96.8% accuracy on UTA-RLDD under a 5-fold cross-validation protocol, and Essahraui et

al. [Essahraui et al. 2025] report approximately 98.9% on the same dataset. Direct comparison is limited by differences in protocol: those works split videos into randomly drawn folds, whereas the present evaluation uses a strict temporal protocol in which personalized references are estimated from only three 2-minute windows per class before an unseen temporal segment is classified second by second. Under these more restrictive conditions, the combined classification reached 97.1%/98.9%, i.e., results comparable to prior art rather than superior, which we interpret as evidence of feasibility of the on-device, personalized approach.

4.1. Classification Results

During the testing phase, the system generated one classification per second in the final 2-minute segment of each video. Thus, each video produced 120 successive classifications, which were compared with the video's global label. The accuracy percentage assigned to each video corresponds to the proportion of these second-by-second classifications that matched the reference label. The results presented in Table 2 correspond to the average of these percentages for videos labeled as "Alert" and "Drowsy".

Table 2. System performance: videos labeled "Alert" and "Drowsy" — 20 users.

Indicator	"Alert" Videos	"Drowsy" Videos
Blink frequency	85.1%	64.2%
PERCLOS	95.7%	92.8%
Yawn frequency	98.6%	34.9%
Blink duration	81.2%	86.0%
Combined classification	**97.1%**	**98.9%**

For the videos labeled as "Alert", yawn frequency showed the highest individual performance (98.6%), followed by *PERCLOS* (95.7%), blink frequency (85.1%), and blink duration (81.2%). This result indicates that, in the evaluated set, the absence or low occurrence of yawns was strongly associated with videos in the alert state. The combined classification reached 97.1%, maintaining high performance by integrating evidence from different indicators.

Although the fused classifier (97.1%) is slightly below yawn frequency alone (98.6%) on Alert-labeled videos, the same indicator drops to 34.9% on Drowsy-labeled videos, whereas the fused classifier remains at 98.9% in that state. The fusion thus trades a small margin in one state for substantially more consistent performance across both states, which is the relevant property for continuous monitoring.

For the videos labeled as "Drowsy", *PERCLOS* (92.8%) and blink duration (86.0%) were the most effective individual indicators, while blink frequency showed intermediate performance (64.2%) and yawn frequency had the lowest individual performance (34.9%). This behavior is related to a characteristic of the dataset used: among the videos labeled as "Drowsy", only about one third presented yawning episodes. Thus, although the occurrence of yawns is relevant evidence of drowsiness, their absence is not sufficient to characterize the alert state. This reinforces the limitation of using this indicator in isolation.

The combined classification reached 98.9% for videos labeled as "Drowsy", outperforming all individual indicators in this state. Taken together, the results indicate that

indicator fusion was the most consistent alternative in the evaluated protocol. While some signals were more informative in a specific state, such as yawn frequency in "Alert" videos and *PERCLOS* in "Drowsy" videos, the temporal integration of indicators reduced dependence on isolated events and produced high performance in both evaluated states.

The 120 second-by-second decisions per video are heavily autocorrelated, as adjacent 120-second windows overlap by 119 seconds. Accordingly, the video ($n = 40$), not the individual second ($n = 4{,}800$), is treated as the unit of analysis for any summary statistic or comparison.

After the main tests were completed, a complementary optimization analysis of the fusion weights was performed. This optimization was applied only to the combined classifier, using the values recorded second by second during the experiments, without changing the individual indicators, the users' personalized parameters, or the temporal protocol. The objective was to verify whether the empirically defined weights were close to a configuration that maximized performance on the evaluated set.

The constrained optimization based on *SLSQP* estimated the following weights: blink frequency = 2.52, *PERCLOS* = 2.75, yawns = 0.97 for absence and 5.61 for presence, and average blink duration = 3.29. With these weights, class-wise accuracy reached 99.2% for "Drowsy" (+0.3%) and 97.6% for "Alert" (+0.5%). These gains were incremental in relation to the empirical weights, indicating that the configuration originally used was already close to an optimized solution for the evaluated set.

5. Conclusions

This work presented an Android system for real-time driver drowsiness detection through the temporal interpretation of multiple physiological–behavioral indicators extracted by computer vision. The proposal integrates information related to the eyes and mouth, including *EAR*, *MAR*, *PERCLOS*, blink frequency, yawn frequency, and average blink duration, to continuously estimate the driver's operational state as "Alert" or "Drowsy".

The main contribution of this work lies in the temporal and personalized integration of these indicators into a system executed directly on Android devices, without reliance on external infrastructure. Unlike approaches centered only on the isolated detection of visual events, such as blinks or yawns, the system combines complementary evidence in a sliding temporal window and uses personalized reference parameters based on initial calibration and labeled sessions from the registered user. This structure allows the classification to consider individual driver patterns and reduce dependence on punctual events.

The results obtained in the initial experimental evaluation indicated high performance in both analyzed states. The combined classification reached 97.1% accuracy in videos labeled as "Alert" and 98.9% in videos labeled as "Drowsy", outperforming the individual indicators in the "Drowsy" state and maintaining high performance in the "Alert" state. The complementary optimization analysis of the fusion weights produced only incremental gains, reaching 97.6% for "Alert" and 99.2% for "Drowsy", which indicates that the empirical configuration used in the main tests was already close to an optimized solution for the evaluated set.

Despite the promising results, some limitations should be considered. The evalu-

ation was conducted with videos from the UTA-RLDD dataset, which were not captured under real or simulated driving conditions. Therefore, the results may not fully reflect typical in-cabin situations, such as vibration, partial facial occlusions, intense lighting variations, and natural gaze changes during driving. In addition, learning and testing segments come from the same recording session (same lighting, pose, camera, and ground-truth state); the results therefore evidence same-session personalization capacity rather than cross-session generalization. No cross-session evaluation is presented, so the generalization of personalized references beyond the calibration session remains unknown. This study also did not systematically evaluate participant representativeness or system performance across subgroups, such as age, sex, skin tone, ethnicity, facial accessories, or other facial characteristics. This limitation is partially mitigated by the use of MediaPipe Face Landmarker, whose documentation (the successor of MediaPipe Face Mesh) reports evaluations with different user profiles [Google 2022], and by the system's personalization mechanism, which adjusts individual references based on initial calibration and user-labeled data. Still, these factors do not replace a specific validation by subgroups. The 20 evaluated individuals were selected based on video duration and visual behavior compatible with the adopted protocol; the use of curated subjects and binary classification (two extreme classes, excluding the intermediate "low vigilance" class) means the headline accuracy figures should be interpreted as evidence of feasibility rather than direct superiority over prior works that use three classes or cross-validation protocols. The results should not be interpreted as evidence of robustness for all user profiles.

As future work, we intend to expand the evaluation with videos captured in real or simulated driving environments, including greater variability of users, lighting, posture, accessory use, and facial occlusion conditions. We also intend to conduct longitudinal evaluations with multiple labeled sessions from the same driver, incorporate additional drowsiness indicators, integrate complementary physiological sensors, and improve the mobile solution in terms of usability, energy efficiency, and continuous operation.

References

Abe, T. (2023). Perclos-based technologies for detecting drowsiness: current evidence and future directions. *Sleep Advances: A Journal of the Sleep Research Society*, 4.

Albadawi, A., Nigam, A., Guo, W., Alam, M. S., Ouyang, J., and Wang, H. (2022). A review of recent driver drowsiness detection systems using computer vision. *Sensors*, 22(13):5064.

Arakawa, T. (2021). Trends and future prospects of the drowsiness detection and estimation technology. *Sensors*, 21(23):7921.

Bajaj, J. S., Kumar, N., and Kaushal, R. K. (2021). Comparative study to detect driver drowsiness. In *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pages 678–683, Greater Noida, India.

Baul, S., Rana, M. R., Trisna, N. J., and Alam, F. B. (2025). Development of a real-time driver's drowsiness detection system using MediaPipe Face Mesh. *International Journal of Engineering and Manufacturing (IJEM)*, 15(5):46–57.

Bergasa, L. M., Nuevo, J., Sotelo, M. A., Barea, R., and López, M. E. (2006). Real-time system for monitoring driver vigilance. *IEEE Transactions on Intelligent Transportation Systems*, 7(1):63–77.

Bittencourt, L. R. A., Silva, R. S., Santos, R. F., Pires, M. L. N., and Mello, M. T. (2005). Sonolência excessiva. *Revista Brasileira de Psiquiatria*, 27(Supl I):16–21.

Das, S. and Bhargava, R. (2023). Vision-based fatigue detection in drivers using multi-facial feature fusion. In *2023 15th International Conference on Developments in eSystems Engineering*, pages 375–380. IEEE.

Dasgupta, A., Rahman, D., and Routray, A. (2019). A smartphone-based drowsiness detection and warning system for automotive drivers. *IEEE Transactions on Intelligent Transportation Systems*, 20:4045–4054.

Dinges, D. F., Mallis, M. M., Maislin, G., Powell, J. W., and Grace, R. (1998). Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Technical Report FHWA-MCRT-98-006, Federal Highway Administration (FHWA), Washington, DC.

El-Nabi, S. A., El-Shafai, W., El-Rabaie, E.-S. M., Ramadan, K. F., El-Samie, F. E. A., and Mohsen, S. (2024). Machine learning and deep learning techniques for driver fatigue and drowsiness detection: a review. *Multimedia Tools and Applications*, 83(3):9441–9477.

Essahraui, S., Lamaakal, I., El Hamly, I., Maleh, Y., Ouahbi, I., El Makkaoui, K., Filali Bouami, M., Pławiak, P., Alfarraj, O., and Abd El-Latif, A. A. (2025). Real-time driver drowsiness detection using facial analysis and machine learning techniques. *Sensors*, 25(3):812.

Federal Highway Police (Brazil) (2026). PRF open data. https://www.gov.br/prf/pt-br/acesso-a-informacao/dados-abertos/dados-abertos-da-prf. Accessed: 22 Jun. 2026.

Florez, R., Palomino-Quispe, F., Alvarez, A. B., Coaquira-Castillo, R. J., and Herrera-Levano, J. C. (2024). A real-time embedded system for driver drowsiness detection based on visual analysis of the eyes and mouth using convolutional neural network and mouth aspect ratio. *Sensors*, 24(19):6261.

Ghoddoosian, R., Galib, M., and Athitsos, V. (2019). A realistic dataset and baseline temporal model for early drowsiness detection. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 178–187, Long Beach, CA, USA.

Google (2022). Model card: MediaPipe Face Mesh V2. https://storage.googleapis.com/mediapipe-assets/Model%20Card%20MediaPipe%20Face%20Mesh%20V2.pdf. Accessed: 22 Jun. 2026.

Google AI Edge (2025). Face Landmark Detection Guide. https://ai.google.dev/edge/mediapipe/solutions/vision/face_landmarker. Accessed: 22 Jun. 2026.

Häkkänen, H., Summala, H., Partinen, M., Tiihonen, M., and Silvo, J. (1999). Blink duration as an indicator of driver sleepiness in professional bus drivers. *Sleep*, 22(6):798–802.

Instituto de Pesquisa Econômica Aplicada (2020). Custos dos acidentes de trânsito no brasil. Technical report, Instituto de Pesquisa Econômica Aplicada, Brasília. https://repositorio.ipea.gov.br/items/98e21d94-320d-42bf-950b-7eb1a7f90911. Accessed: 22 Jun. 2026.

Jeong, M., Kim, D., Park, S., and Paik, J. (2022). Drowsy status monitoring system based on face feature analysis. In *2022 International Conference on Electronics, Information, and Communication (ICEIC)*, pages 1–4, Jeju, Korea, Republic of.

Joshy, A. J. and Hwang, J. T. (2024). PySLSQP: A transparent Python package for the SLSQP optimization algorithm modernized with utilities for visualization and post-processing. *Journal of Open Source Software*, 9(103):7246.

Kashevnik, A., Lashkov, I., and Gurtov, A. (2020). Methodology and mobile application for driver behavior analysis and accident prevention. *IEEE Transactions on Intelligent Transportation Systems*, 21(6):2427–2436.

Kolus, A. (2024). A systematic review on driver drowsiness detection using eye activity measures. *IEEE Access*, 12:97969–97993.

Kraft, D. (1994). Algorithm 733: TOMP—Fortran modules for optimal control calculations. *ACM Transactions on Mathematical Software*, 20(3):262–281.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M., Lee, J., Chang, W.-T., Hua, W., Georg, M., and Grundmann, M. (2019). MediaPipe: A framework for perceiving and processing reality. In *Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR) 2019*.

Schleicher, R., Galley, N., Briest, S., and Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: Looking tired? *Ergonomics*, 51(7):982–1010.

Singh, N. T., Saurav, Pathak, N., Raizada, A., and Shukla, S. (2023). Real-time driver drowsiness detection system using cascaded convnet framework. In *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, pages 828–833, Coimbatore, India.

Soukupová, T. and Čech, J. (2016). Real-time eye blink detection using facial landmarks. In *Proceedings of the 21st Computer Vision Winter Workshop (CVWW)*, pages 1–8.

Tamanani, R., Muresan, R., and Al-Dweik, A. (2021). Estimation of driver vigilance status using real-time facial expression and deep learning. *IEEE Sensors Letters*, 5(5):1–4.

World Health Organization (2024). Global status report on road safety 2023: Country and territory profiles. https://www.who.int/publications/i/item/9789240087712. Accessed: 22 Jun. 2026.

World Health Organization (2026). Road traffic injuries. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries. Accessed: 22 Jun. 2026.

Yang, H., Liu, L., Min, W., Yang, X., and Xiong, X. (2021). Driver yawning detection based on subtle facial action recognition. *IEEE Transactions on Multimedia*, 23:572–583.