



OPEN ACCESS

EDITED BY

Ayşe ÜNAL,
Siirt University, Türkiye

REVIEWED BY

Saeed Mohsen,
King Salman International University, Egypt
Vijay Sharma,
Manipal University Jaipur, India

*CORRESPONDENCE

Carlos Andrés Mugruza-Vassallo,
✉ cmugruza@yahoo.com

RECEIVED 31 July 2025

REVISED 09 October 2025

ACCEPTED 17 October 2025

PUBLISHED 11 November 2025

CITATION

Lozano-Reyes GS and Mugruza-Vassallo CA
(2025) A vision-based drowsiness detection
system for railway operators using lightweight
convolutional neural networks.
Front. Future Transp. 6:1677442.
doi: 10.3389/ffutr.2025.1677442

COPYRIGHT

© 2025 Lozano-Reyes and Mugruza-Vassallo.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

A vision-based drowsiness detection system for railway operators using lightweight convolutional neural networks

Guisella Stefany Lozano-Reyes¹ and
Carlos Andrés Mugruza-Vassallo^{2*}

¹Universidad Tecnológica de Lima Sur (UNTELS), Lima, Peru, ²Grupo de Investigación en Computación y Neurociencia Cognitiva, Universidad Tecnológica de Lima Sur (UNTELS), Lima, Peru

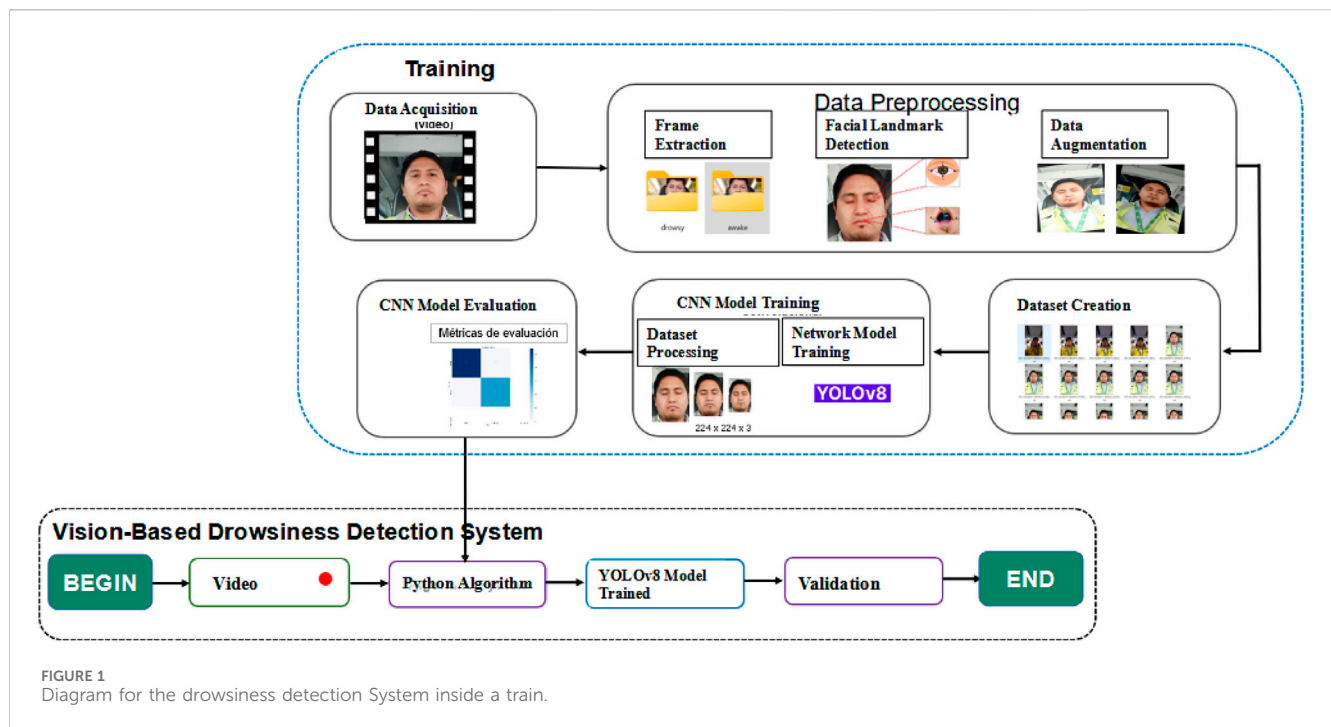
This research addresses the challenge of monitoring railway driver drowsiness using a real-time, vision-based system powered by convolutional neural networks, specifically the YOLOv8 architecture including attention mechanisms. The core idea is to keep the eye on subtle facial features like eyelid closure durations as indicators of fatigue. The model is designed to be lightweight for fast processing, which is critical for real-time applications. To build the model, a custom dataset of 6,991 frames was compiled. It also boosted the dataset's diversity using data augmentation, improving the model's robustness against real-world variability. And it paid off: the system hit an overall accuracy of 96.8%, precision of 97.28%, and recall of 97.46%, which is impressive, especially under different lighting conditions. The system works best in low sunlight. When strong solar glare kicks in, detection dips, showcasing the impact environmental factors can have on vision-based systems. In short, this study highlights how deep learning can realistically enhance railway safety by alerting operators before drowsiness leads to incidents. For future work, the plan was to toughen up the system to handle tough lighting better and explore combining vision with other sensor types (e.g., electroencephalography) for a fuller fatigue picture. Discussion about particular cognitive brain computer interface and health issues as anemia for further studies.

KEYWORDS

attention mechanisms, drowsiness detection, railway safety, convolutional neural networks (CNN), you only look once (YOLO)v8, computer vision, fatigue monitoring, real-time systems

1 Introduction

The system's built around Convolutional Neural Networks (CNNs) processing live video feeds from cameras aimed at the train operator's face. It detects classic drowsiness signs like yawning, slow eyelid closure, and irregular blinking—key fatigue cues backed by research (Fakhri et al., 2024). Recent advances like the attention mechanism and transformer-based architectures, exemplified by attention-centric You Only Look Once v8 (YOLOv8) models, enable more efficient and precise focus on critical facial features in real time, enhancing fatigue detection performance without sacrificing speed (Nimma et al., 2025). The hardware is a neat integration of high-res cameras and image processors running these sophisticated algorithms continuously while the train's in motion (Alstom, 2020; Albadawi et al., 2024). Also, other deep learning models have been developed for neck



measurement for kinematic analysis (e.g., Garrosa et al., 2023), which may be impractical for a long time in professional railway settings.

Why is it important? Human error tied to fatigue is a serious risk in train operation. Drowsiness is also associated with a gradual reduction in responsiveness, decreased selective attention, and errors in short-term memory (Kamran et al., 2019). Going in line with technology trends (Mugruza-Vassallo and Miñano-Suarez, 2016), this tech not only boosts safety but also fits international trends toward AI-driven monitoring in transport, nudging us toward smarter, automated safety measures. Other studies have already shown AI's promise in spotting driver fatigue across cars, trucks, and buses (Alstom, 2017). So adapting these ideas specifically for trains—and for places like the Lima Metro—is a logical and valuable move.

Fitzharris et al. (2017) reported that in a commercial truck fleet, the use of in-cabin alerts and company-wide real-time feedback reduced fatigue events by 66%–95%, with fatigue episodes occurring later and for shorter durations. The goal here is straightforward: create a reliable, scalable system that can help prevent the typically incorrect train parking at boarding platforms, that delays commercial service and accidents tied to operator tiredness, ultimately supporting safer urban transit worldwide.

The aim of this work was to successfully contribute to railway operations in different time periods (morning, afternoon and night) using a lightweight model (CNN with ~2.7 million parameters) to achieve a high level of accuracy (>96%) in a dataset of train drivers.

The manuscript is organized as follows. Section 1 introduced and reviewed the indexes for cabin alerts in train systems. Section 2 outlines the methodology, including a description of the experimental test setup, data acquisition, data processing (training and validation) and deployment. Section 3 presented the results and meeting requirements and design issues for

Dataset and Performance, Evaluation Metrics, Temporal Detection, Real-World Testing in the monitoring system. The discussion is presented in Section 4. Finally, the conclusion drawn from the results is presented in Section 5.

2 Methods

The system's workflow revolves around YOLOv8, a CNN model known for balancing solid accuracy with fast inference—which you need when time is of the essence (Figure 1 shows the setup inside a train).

2.1 Experimental setup and data acquisition

The Linea Uno train cabin was used as the setting for some train drivers' gestures after a few rounds of driving the train. Videos were captured using a mobile device in MP4 format at 30 frames per second (fps) with a resolution of 478 by 850 pixels. The device was placed in front of the driver and focused primarily on his or her face under different lighting conditions.

It collected facial videos under controlled conditions simulating drowsiness indicators like eye closure and yawning, placing a high-def camera device about 60–65 cm from the conductor's face to capture consistent images. The threshold of 833 ms for eye closure aligned well with known fatigue markers (Dinges and Grace, 1998), accurately flagging drowsiness episodes in the videos. They then manually labeled key regions (eyes and mouth) to train the model accurately. We ensured different lightning conditions using real operating settings in the morning, afternoon and night. The protocol was approved by the Institutional Review Board at UNTELS (VIII PTM-TSP-FIG-2024).

TABLE 1 Training hyperparameters across YOLOv8 variants.

Model	Epochs	Batch size	Learning rate	Image size	Optimizer
YOLOv8n-cls	500	1,024	0.01	224 × 224	AdamW
YOLOv8s-cls	100	256	0.01	224 × 224	AdamW
YOLOv8m-cls	500	1,024	0.01	224 × 224	AdamW

2.2 Data preprocessing

We preprocessed video frames extracted at 30 fps following three stages:

Feature Extraction and Downsampling: In total 17,476 frames were acquired. For feature extraction one-sixth (2,913 images) were selected to reduce redundancy and facilitate computation.

Facial Landmark Detection: Twelve key facial landmarks were detected using OpenCV and MediaPipe in *cvzone* library (CVzone, n.d.) identifying four points for each eye and mouth. To calculate eye closure a ratio threshold <18 and yawning threshold >35 following fatigue detection literature (Khabarлак and Koriashkina, 2021).

Image normalization, Resizing and Data Augmentation: All the images were normalized, resized to 640 × 640 pixels, and went through augmentations—rotations, brightness tweaks, slight shifts—to reach rotation between −15° and +15°, brightness between −20% and +20% to mimic real environmental changes and avoid overfitting. Roboflow was used for data augmentation. The process begins with loading the original 2,913 images, which are classified into two categories: Awake and Drowsy, as shown in Figure 1, top right.

2.3 Dataset creation

In this study, data augmentation allowed expansion to 6,991 images and the final dataset was divided for training (87%), validation (6.5%) and testing (6.5%). For training images were resized to 224 × 224 × 3 pixels using YOLOv8.

2.4 Model training

In Python 3.8 and PyTorch, following the approach of Simonyan and Zisserman (2014) with an 80/20 train-validation split, they fine-tuned hyperparameters (see Table 1) like learning rate and epochs until the model converged nicely. It also used regularization methods to keep performance solid on unseen data.

2.5 Validation and evaluation

Testing was done on separate video sequences reflecting real-world conditions. Then, performance metrics measuring accuracy, precision, recall, and false positive rates (Metz, 1979) as well as confusion matrix and loss function (Hicks et al., 2022). These are done also by current CNN deep learning works (in optical recognition by Zayed, et al., 2024; X-ray to seek COVID-19 in chest by Mohsen et al., 2024; ECG compression by Hassan and

Mohsen, 2025; brain tumor detection by Mohsen et al., 2023; electroencephalography - EEG emotion recognition by Mohsen and Alharbi, 2021) to truly gauge reliability, defined in Equations 1–4:

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN}) \quad (1)$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (2)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (3)$$

$$\text{F1-score} = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall}) \quad (4)$$

2.6 System deployment

The model was embedded into hardware (NVIDIA A100 40GB GPU accessed remotely via SSH) capable of rapid processing, with an alert system that triggers visual or sound warnings when drowsiness signs are detected—all designed for minimal latency to alert operators promptly. We also extracted model complexity (layers, parameters, GFLOPs) from each as Qian and Liu (2024) plotted for attention mechanisms for lightweight image classification. Here, five additional videos were tested in real system settings, with manually annotation in collaboration with two experienced train drivers operating at *Linea Uno*.

3 Results

The system performed remarkably well:

- **Dataset and Performance:** After augmentations, the dataset grew from 2,913 to 6,991 frames, with about 87% reserved for training. Training stabilized by epoch 145 with a low loss of 0.06787. Figure 2 was obtained as a result of the three training sessions, showing the relationship between the loss function and the epoch for each variant of the YOLOv8 model. These scatter plots illustrate how the network adjusts its weights during each training epoch to optimize predictions (see Figures 2a–c), achieving high accuracy and minimizing loss (see Figures 2d–f).

The results show that the network reached optimal performance at different epochs for each variant. Figure 2a shows that the network achieved its best performance in epoch 145, with a training loss value of 0.06787 as well as validation loss curves in each model. No significant improvements were subsequently observed, indicating that the network had reached its limit. Figure 2b shows optimal performance in epoch 100 with a loss value of 0.04662. Figure 2c shows optimal performance in epoch 64 with a loss value of 0.05495.

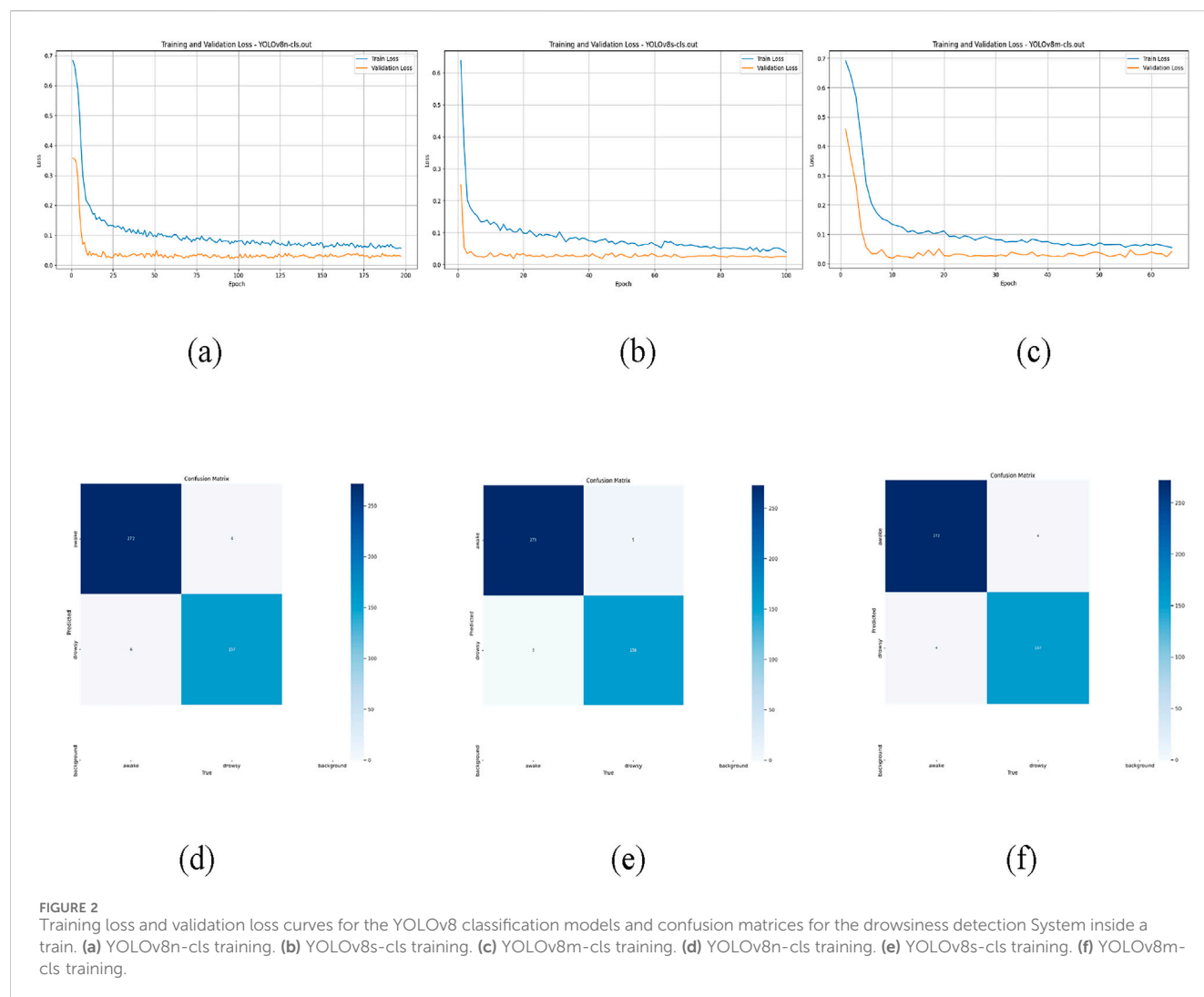


FIGURE 2

Training loss and validation loss curves for the YOLOv8 classification models and confusion matrices for the drowsiness detection System inside a train. (a) YOLOv8n-clc training. (b) YOLOv8s-clc training. (c) YOLOv8m-clc training. (d) YOLOv8n-clc training. (e) YOLOv8s-clc training. (f) YOLOv8m-clc training.

- **Evaluation Metrics:** The accuracy hovered around 97% during mornings and nights but dipped slightly to about 95.4% in the afternoon—likely due to sunlight affecting image quality (details on Table 2). Similarly, precision and recall stayed in the high 90s, indicating very reliable drowsiness detection with few false alarms.
- **Temporal Detection:** The system successfully tracked transitions from alert to drowsy states based largely on an eye closure threshold set at 833 ms (a proven standard in literature, e.g., Dinges and Grace, 1998). Figure 3 lays out these changes visually.
- **Real-World Testing:** We recorded and tested five participants in real train environments, making videos across shifts—morning, afternoon, night—with data showing high accuracy and precision despite environmental challenges (Table 2 outlines detailed results).

Figure 4 illustrates the average metrics per shift in real setting. Confirming the system's high efficacy in diverse lighting and environmental conditions, with a noted decrease during the afternoon due to solar glare affecting facial feature detection.

- **System Reliability and Limitations:** The biggest hiccup was about a 6% false alarm rate in afternoon tests, attributed to solar reflections on the dashboard messing with facial recognition. The trained model integrated for the system showed a lightweight inference pipeline capable of processing video in real time for the YOLOv8 nano model (see Figure 5). This highlights a clear avenue for future improvement.

In short, with average accuracy near 96.8%, this system looks well-suited for the tough demands of real-time fatigue detection in rail settings.

4 Discussion

The results indicate that YOLOv8 achieves a balanced accuracy of 96.8% in detecting drowsiness among train drivers, as evaluated by expert-validated in real settings. This outcome aligns with findings reviewed by Disha and Upadhyaya (2025), who emphasize in driver fatigue monitoring. Therefore, it is



FIGURE 3
Drowsiness detection system inside a train. . (a) Frame 1. (b) Frame 1. (c) Frame 1. (d) Frame 1. (e) Frame 1. (f) Frame 1. (g) Frame 1. (h) Frame 1. (i) Frame 1.

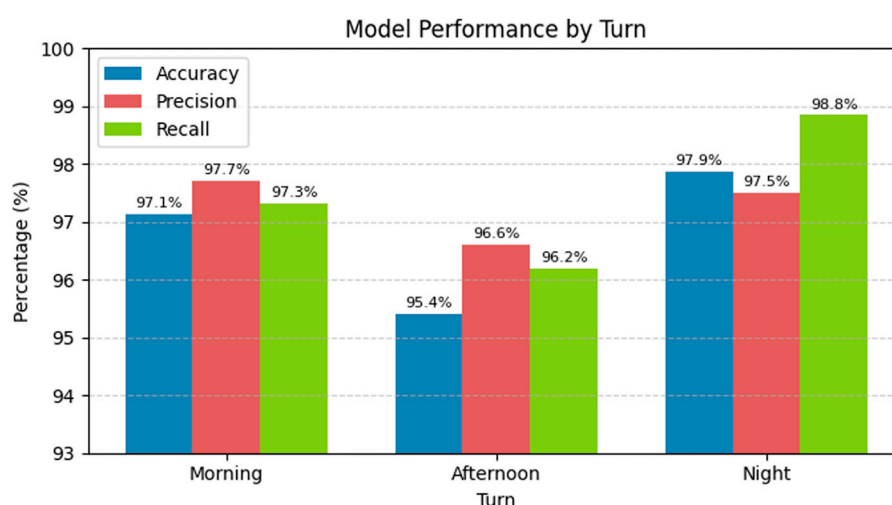
noteworthy that the present results of performance evaluation employed manually annotated videos labeled with the assistance of an experienced train operation team. Similarly, Wang et al. (2021) also support this method by applying CNNs for eye classification accuracy, building upon the foundational CNN architecture

proposed by Simonyan and Zisserman (2014) testing small matrices and several layers.

Performance varied across different lighting conditions as expected, with night shifts achieving nearly 98% accuracy, and afternoon shifts the lowest at about 95%, largely due to sunlight

TABLE 2 Results of testing in real train at Linea Uno - Lima ... No testing ... Validation in real train.

Video	Shift	Number of photograms	Drowsiness		Awake		Accuracy	Precision	Recall
			TP	FN	TN	FP			
V1	Morning	1,890	1,110	12	743	25	98.04%	97.79%	98.93%
V2	Morning	2,250	1,230	55	935	30	96.22%	97.62%	95.72%
V3	Afternoon	1,740	1,050	45	600	45	94.83%	95.89%	95.89%
V4	Afternoon	1,620	1,020	37	535	28	95.99%	97.33%	96.50%
V5	Night	2,250	1,290	15	900	45	97.87%	97.51%	98.85%

FIGURE 4
Accuracy, precision and recall averages in each turn.

reflections complicating facial feature extraction, as [Zhao et al. \(2024\)](#) observed. Such lighting issues are common challenges in computer vision systems, even those utilizing infrared cameras using SVM classifiers ([Travieso-Gonzalez et al., 2021](#)), emphasizing the importance of lighting-robust models.

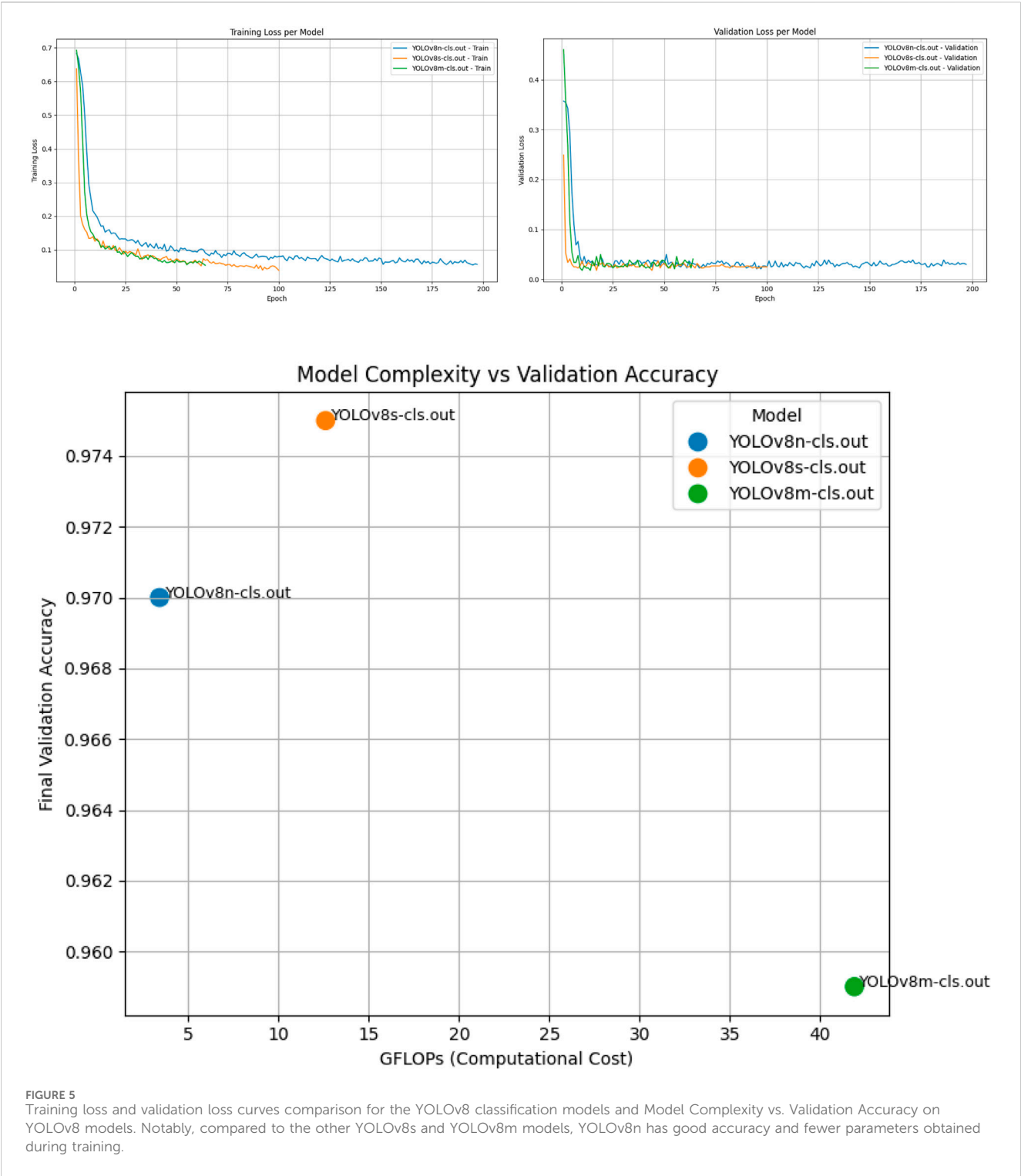
The lightweight YOLOv8n-cls models, comprising ~2.7 million parameter model, allowed real-time processing without sacrificing accuracy, a characteristic by [Kausar and Aishwarya \(2016\)](#) and [Howard et al. \(2017\)](#) in earlier models (VGG, Inception V2 and Mobilenet) using small CNN architectures and efficient and precise focus recently reported ([Nimma et al., 2025](#)).

The 833 ms eye closure threshold corresponds well with established fatigue indicators identified by [Dinges and Grace \(1998\)](#) and subsequently employed by [Zhang et al. \(2017\)](#), effectively detecting drowsiness episodes.

False positives under conditions of intense glare remain a limitation. Future developments could include infrared imaging, advanced image preprocessing for virtual reality and 3D face reconstruction ([Wen et al., 2021](#); [Yang et al., 2024](#)), and integration of additional data such as EEG signals (using wavelet transform by [Tuncer et al., 2021](#)) or vehicle telemetry to further reduce false alarms.

Overall, these findings support the practical application of vision-based fatigue detection systems for train drivers, contributing to improved accident prevention and railway safety, consistent with previous studies utilizing eye blink analysis ([Fakhri et al., 2024](#)), Bi-LSTM-SVM adaptive algorithms ([Chen and Zheng, 2023](#)), and LSTM and CNN approaches to EMG and cognitive state analysis ([Yu et al., 2024](#)).

Railway conductors face a physically demanding job that requires strength, agility, and stamina—from lifting heavy items to climbing between train cars—while also staying mentally sharp throughout long, often irregular shifts. They work in all kinds of weather and must be constantly alert to ensure safety, quickly responding to any issues or emergencies. This combination of physical effort and sustained focus can lead to fatigue and strain, making it essential for conductors to maintain good fitness, rest well, and manage health factors that might worsen tiredness, like anemia ([Arnold & Itkin LLP., n.d.](#); also in other sector [Chowdhury and Nuruzzaman, 2023](#)). Next work is pointing to cognitive computing in drivers, in a recent work in postpartum women, anemia was shown to be impair cognitive processing in 3D video scenes ([Cajas-Shao et al., in press](#)).



Railway conductors face a physically demanding job that requires strength, agility, and stamina, from lifting heavy items to climbing between train cars, while also staying mentally sharp throughout long, often irregular shifts. They work in all kinds of weather and must be constantly alert to ensure safety, quickly responding to any issues or emergencies, pointing to a basic brain computer interface (BCI) to help drivers. This combination of physical effort and sustained focus can lead

to fatigue and strain, making it essential for conductors to maintain good fitness, rest well, and manage health factors that might worsen tiredness, like anemia (Arnold & Itkin LLP., n.d; also in other sector Chowdhury and Nuruzzaman, 2023). Next work is pointing to cognitive computing in drivers, in a recent work in postpartum women, anemia was shown to be impair cognitive processing in 3D video scenes (Cajas-Shao et al., in press).

5 Conclusion

To wrap up, this study successfully designed and tested a vision-based drowsiness detection tool tailored for railway drivers, leveraging the YOLOv8 CNN. With a rich training set of nearly 7,000 frames and effective data augmentation, the system achieved strong accuracy (96.8%), precision (97.28%), and recall (97.46%) in real-world environments.

It reliably differentiated alert versus fatigued states across different shifts, though performance dipped slightly with afternoon solar lighting. Setting appropriate detection thresholds allowed timely and accurate alerts, bolstering safety.

So, this technology appears very promising for real-time fatigue monitoring in the rail industry, which could significantly reduce risks related to driver tiredness. Next steps focus on improving lighting resilience and exploring multi-sensor fusion for even better accuracy.

5.1 Future works

System Deployment: First, to improve the vision-based system, different methods to enhance difficult lightning conditions must be studied, either incorporating infrared cameras or other image preprocessing algorithms (Yang et al., 2024). Second, to evolve a multimodal fusion approach, such as EEG to have a comprehensive and reliable fatigue assessment (Cao et al., 2025; Yu et al., 2024). Therefore two directions are on minimally intrusive EEG systems (reviewed by Balam, 2024) and optimizing drowsiness index (although with 14 parietal electrodes by Di Flumeri et al., 2024). Other methods may use visual and auditory linear models (Mugruza-Vassallo, 2016) and BCI may send a warning signal to train drivers as some automobile systems have been launched.

Health factors like anemia can impair cognitive processing and reaction times (Cajas-Shao et al., in press). Bearing in mind the high percentage of anemia children in Peru during the last 40 years, probably around 40% of drivers got anemia. Suggest longer reaction times to visual stimuli. Therefore, EEG testing in drivers who got drowsiness alarms and those who did not plus some anemia medical records would be the basis to better understand individual susceptibility to fatigue and drowsiness. The model was embedded into the idea of being incorporated on hardware capable of rapid processing, with an alert system that triggers visual or sound warnings when drowsiness signs are detected and then studied in a dataset.

The rapid development of YOLO architectures, now reaching YOLOv12, is well recognized. Our decision to use YOLOv8 was guided by a comprehensive view of the model's evolution alongside the specific needs of our project. Our laboratory's work on vision-based drowsiness detection dates back to early investigations of driver distraction in 2013 (Arriaga and Mugruza-Vassallo, 2013, internal report) and has since progressed through the evaluation of various CNN architectures such as Xception, VGG16, and Inception V3 (Mamani-Diaz et al., 2019). Previous trials with YOLOv4 and YOLOv5 in 2021 did not deliver satisfactory outcomes for our application, despite positive reports with YOLOv3 in other studies (Xiao et al., 2022). At the outset of this research in early 2024, YOLOv8 presented a notable advancement, building on YOLOv7's architectural enhancements, including transformer-like components (Gomaa and Abdalrazik, 2024) and offering a stable, extensively documented framework with dedicated classification models (e.g.,

YOLOv8n-cls) essential for our real-time deployment objectives (Alif and Hussain, 2024). Although YOLOv9 was released in February 2024, it initially did not include a nano-version, which was critical for our lightweight system design. Consequently, we adopted YOLOv8 to establish a reliable baseline. The later introduction of nano-versions in YOLOv10 and subsequent releases (Sapkota et al., 2025) supports our approach and confirms that our methodology is readily adaptable to these newer, more efficient models. Thus, our findings for drowsiness serve not as a fixed conclusion but as a vital proof-of-concept that lays the groundwork for immediate future research leveraging the latest YOLO architectures to enhance brain-computer interface and railway safety technologies.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://github.com/Lozano993/Somnolencia-en-conductores-de-tren>.

Ethics statement

The studies involving humans were approved by IRB-VRIN-Universidad Nacional Tecnológica de Lima Sur. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

GL-R: Conceptualization, Data curation, Formal Analysis, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review and editing. CM-V: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The present authors express thanks to UNTELS support through the request for this publication OFICIO 0693-2025-UNTELS-VRI-II for additional expenses to RCU-125-2024-UNTELS under the project “Tareas cognitivas en estudiantes con anemia y exploración de respuesta cognitivas en ElectroEncefalograma”.

Acknowledgements

The present authors express gratitude to internal reviewers Prof. Ricardo Palomares, Jorge Lopez, and Felix Illesca for helping with the

constructive critique, as well as Mayte Rojas and Gloria Castro from the Instituto de Investigación at UNTELS, and the former thesis of Gilbert Arriaga (2013-2014) at UTP for earlier development that inspired the analysis that was reported in the undergraduate dissertation at the UNTELS Lozano Reyes, G. S. (2025). The present authors express thanks to UNTELS support through RCU-125-2024-UNTELS the project “Tareas cognitivas en estudiantes con anemia y exploración de respuesta cognitivas en ElectroEncefalograma”. This project helps with the experimental setup and most of the analysis that is reported here and envision a future BCI integration to analyse control and anemia railway drivers.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Albadawi, A. (2024). Graph-based online monitoring of train driver states via facial and behavioral analysis. Available online at: <https://arxiv.org/abs/2505.08800>.
- Alif, M. A. R., and Hussain, M. (2024). YOLOv1 to YOLOv10: a comprehensive review of YOLO variants and their application in the agricultural domain. *arXiv preprint arXiv:2406.10139*. Available online at: <https://arxiv.org/html/2406.10139v1>.
- ALSTOM (2017). ALSTOM. Obtenido de. Available online at: <https://www.alstom.com/es/press-releases-news/2017/10/alstom-envia-el-primer-tren-adicional-para-la-linea-1-del-metro-de-lima-peru>.
- ALSTOM (2020). Código de ética. Available online at: https://www.alstom.com/sites/alstom.com/files/2020/07/08/Alstom_CodeofEthics_2020_ES_0.pdf.
- Arnold & Itkin LLP. (n.d.). Toxic hazards for railroad workers: exposure from the transportation of hazardous materials. Available online at: <https://www.arnolditkin.com/railroad-accidents/railroad-worker-injuries/toxic-hazards/>
- Balam, V. P. (2024). Systematic review of single-channel EEG-based drowsiness detection methods. *IEEE Transactions on Intelligent Transportation Syst.* 25 (11), 15210–15228. doi:10.1109/TITS.2024.3442249
- Bowler, N., and y Gibson, H. (2015). Fatigue and its contributions to railway incidents. Available online at: <https://trid.trb.org/View/1542424>.
- Cajas Cerna, S. P., Portilla-Fernández, J. A., and Mugruza-Vassallo, C. A. The impact of postpartum anemia on cognitive function: a study using 3D video game-based assessment of reaction times in women. *Front. Psychol. - Cognitive Sci.* doi:10.3389/ffpsyg.2025.1598851
- Cao, S., Feng, P., Kang, W., Chen, Z., and Wang, B. (2025). Optimized driver fatigue detection method using multimodal neural networks. *Sci. Rep.* 15, 12240. doi:10.1038/s41598-025-86709-1
- Chen, L., and Zheng, W. (2023). Research on railway dispatcher fatigue detection method based on deep learning with multi-feature fusion. *Electronics* 12 (10), 2303. doi:10.3390/electronics12102303
- Chowdhury, A., and Nuruzzaman, M. (2023). Design, testing, and troubleshooting of industrial equipment: a systematic review of integration techniques for US manufacturing plants. *Rev. Appl. Sci. Technol.* 2 (01), 53–84. doi:10.63125/893et038
- cvzone (n.d.). cvzone (Version 2023) [Computer software]. Available online at: <https://github.com/cvzone/cvzone>.
- Dinges, D. F., and Grace, R. C. (1998). Perclos: a valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Available online at: <https://api.semanticscholar.org/CorpusID:141471330>.
- Disha, Y., and Upadhyaya, A. (2025). A YOLO and machine learning-based framework for real-time driver drowsiness detection. *Int. J. Innovative Res. Technol.* 11 (11), 5262–5269.
- Fakhri, M. (2024). Train driver drowsiness detection using deep learning approach. *Int. J. Railw. Res.* doi:10.22068/ijrare.344
- Fitzharris, M., Liu, S., Stephens, A. N., and Lenné, M. G. (2017). The relative importance of real-time in-cab and external feedback in managing fatigue in real-world commercial transport operations. *Traffic Inj. Prev.* 18 (Suppl. 1), S71–S78. doi:10.1080/15389588.2017.1306855
- Garrosa, M., Ceccarelli, M., Díaz, V., and Russo, M. (2023). Experimental validation of a driver monitoring system. *Machines* 11 (12), 1060. doi:10.3390/machines11121060
- Gomaa, A., and Abdalrazik, A. (2024). Novel deep learning domain adaptation approach for object detection using semi-self building dataset and modified YOLOv4. *World Electr. Veh. J.* 15 (6), 255. doi:10.3390/wevj15060255
- Hassan, A. M. A., and Mohsen, S. (2025). Compression of electrocardiogram signals using compressive sensing technique based on curvelet transform toward medical applications. *Multimed. Tools Appl.* 84, 11203–11219. doi:10.1007/s11042-024-19328-z
- Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., et al. (2022). On evaluation metrics for medical applications of artificial intelligence. *Sci. Rep.* 12 (1), 5979. doi:10.1038/s41598-022-09954-8
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., and Weyand, T. (2017). Mobilenets: efficient convolutional neural networks for mobile vision applications. *arXiv Preprint arXiv:1704.04861*.
- Kamran, M. A., Mannan, M. M. N., and y Jeong, M. Y. (2019). Drowsiness, fatigue and poor sleep's causes and detection: a comprehensive study. *IEEE Access* 7, 167172–167186. doi:10.1109/access.2019.2951028
- Kausar, F., and Aishwarya, P. (2016). “Artificial neural network: framework for fault tolerance and future,” in *2016 international conference on electrical, electronics, and optimization techniques (ICEEOT)* (IEEE), 6042–6052. doi:10.1109/ICEEOT.2016.7754760
- Khabaralak, K., and Koriashkina, L. (2021). Fast facial landmark detection and applications: a survey. *arXiv Preprint arXiv:2101.10808*.
- Lozano Reyes, G. S. (2025). Propuesta de un sistema de visión artificial para detectar somnolencia en conductores de tren Metrópolis basado en redes neuronales convolucionales [Trabajo de suficiencia profesional, Universidad Nacional Tecnológica de Lima Sur, Facultad de Ingeniería y Gestión, Escuela Profesional de Ingeniería Electrónica y Telecomunicaciones]. *Univ. Nac. Tecnológica Lima Advisor Mugruza Vassallo, (ORCID 0000-0002-9262-7198)*.
- Metz, C. E. (1979). *Applications of ROC analysis in diagnostic image evaluation (No. CONF-790783-1)*. Chicago, IL: Chicago Univ., IL; Franklin McLean Memorial Research Inst. Available online at: <https://www.osti.gov/servlets/purl/5616875>.
- Mohsen, S., and Alharbi, A. G. (2021). “EEG-based human emotion prediction using an LSTM model,” in *2021 IEEE international midwest symposium on circuits and systems (MWSCAS)* (IEEE), 458–461. doi:10.1109/MWSCAS47672.2021.9531707
- Mohsen, S., Abdel-Rehim, W. M. F., Ahmed, E., and Mohamed Kasem, H. (2023). A convolutional neural network for automatic brain tumor detection. *Proc. Eng. Technol. Innov.* 24, 15–21. doi:10.46604/peti.2023.10307
- Mohsen, S., Scholz, S. G., and Elkaseer, A. (2024). Detection of COVID-19 in chest X-Ray images using a CNN model toward medical applications. *Wirel. Pers. Commun.* 137, 69–87. doi:10.1007/s11277-024-11309-7
- Mugruza Vassallo, C. A. (2016). “Different regressors for linear modelling of ElectroEncephaloGraphic recordings in visual and auditory tasks,” in *Wearable and implantable Body Sensor Networks (BSN), 2016 IEEE 13th International Conference on* (IEEE), 260–265. doi:10.1109/BSN.2016.7516270

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. For generating APA references.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Mugruza Vassallo, C. A., and Miñano Suarez, S. (2016). "Academia and patents at information and communications technology in South-America productivity," in *Information communication and management (ICICM), international conference on (IEEE)*, 24–29. doi:10.1109/INFOCOMAN.2016.7784209
- Mugruza-Vassallo, C. A., Granados-Domínguez, J. L., Flores Benites, V., and Córdova-Berrios, L. (2022). Different Markov chains modulate visual stimuli processing in a Go-Go experiment in 2D, 3D and augmented reality. *Front. Hum. Neurosci.* 16, 955534. doi:10.3389/fnhum.2022.955534/
- Naseri, R., and Mohammadzadeh, S. (2020). Nonlinear train-track-bridge interaction with unsupported sleeper group. *Int. J. Railw. Res.* 7 (1), 11–28. doi:10.22068/IJRARE.7.1.11
- Nimma, D., Al-Omari, O., Pradhan, R., Ulmas, Z., Krishna, R. V. V., El-Ebiary, T. Y. A. B., et al. (2025). Object detection in real-time video surveillance using attention based transformer-YOLOv8 model. *Alexandria Eng. J.* 118 (Suppl. C), 482–495. doi:10.1016/j.aej.2025.01.032
- Qian, C., and Liu, B. (2024). NCANet: integrating normalized channel attention for enhanced lightweight image classification. *Preprint*. doi:10.21203/rs.3.rs-4008005/v1
- Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv Preprint arXiv:1409.1556*.
- Travieso-González, C. M., Alonso-Hernández, J. B., Canino-Rodríguez, J. M., Pérez-Suárez, S. T., Sánchez-Rodríguez, D., and Ravelo-García, A. G. (2021). Robust detection of fatigue parameters based on infrared information. *IEEE Access*. doi:10.1109/ACCESS.2021.3052770
- Tuncer, T., Dogan, S., and Subasi, A. (2021). EEG-based driving fatigue detection using multilevel feature extraction and iterative hybrid feature selection. *Biomed. Signal Process. Control* 68, 102591. doi:10.1016/j.bspc.2021.102591
- Viola, P., and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proc. 2001 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. CVPR* 1, 1–511–518. doi:10.1109/CVPR.2001.990517
- Wang, Y., Li, Z., Zhang, H., Wang, K., and Zhang, W. (2021). Tactile and thermal sensors built from carbon-polymer nanocomposites—A critical review. *Sensors* 21 (4), 1234. doi:10.3390/s21041234
- Wen, L., Zhou, J., Huang, W., and Chen, F. (2021). A survey of facial capture for virtual reality. *IEEE Access* 10, 6042–6052. doi:10.1109/ACCESS.2021.3138200
- Xiao, W., Liu, H., Ma, Z., Chen, W., Sun, C., and Shi, B. (2022). Fatigue driving recognition method based on multi-scale facial landmark detector. *Electronics* 11 (24), 4103. doi:10.3390/electronics11244103
- Yang, X., Taketomi, T., Endo, Y., and Kanamori, Y. (2024). "Makeup prior models for 3D facial makeup estimation and applications," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2165–2176. doi:10.48550/arXiv.2403.17761
- Yu, L., Yang, X., Wei, H., Liu, J., and Li, B. (2024). Driver fatigue detection using PPG signal, facial features, head postures with an LSTM model. *Heliyon* 10 (21), e39479. doi:10.1016/j.heliyon.2024.e39479
- Yu, J., Wang, J., Zhu, C., Meng, W., Lu, B., and Zhou, Z. (2024). "Hybrid CNN-LSTM-Transformer model for robust muscle fatigue detection during rehabilitation using sEMG signals," in *Proceedings of ICARM*. doi:10.1109/icarm62033.2024.10715764
- Zayed, M. M., Mohsen, S., Alghuried, A., Hijry, H., and Shokair, M. (2024). IoUT-Oriented an efficient CNN model for modulation schemes recognition in optical wireless communication systems. *IEEE Access* 7, 186836–186855. doi:10.1109/access.2024.3515895
- Zhang, X., Li, J., Liu, Y., Zhang, Z., Wang, Z., Luo, D., et al. (2017). Design of a fatigue detection system for high-speed trains based on driver vigilance using a wireless wearable EEG. *Sensors* 17 (3), 486. doi:10.3390/s17030486
- Zhao, J., Xue, B., and Zhang, M. (2024). SGformer: boosting transformers for indoor lighting estimation from a single image. *Comput. Vis. Media* 10 (4), 671–686. doi:10.1007/s41095-024-0447-8