Fusion-Based Driver State Monitoring: A Multimodal Approach to Drowsiness Mitigation

Suhrit Ghosh Advisor: Kevin Crowthers, Ph.D.

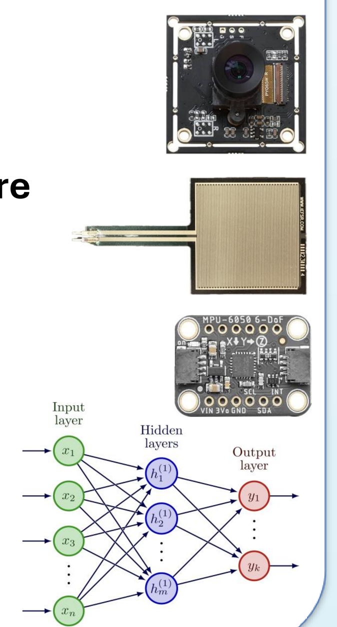
Background

Drowsy Driving

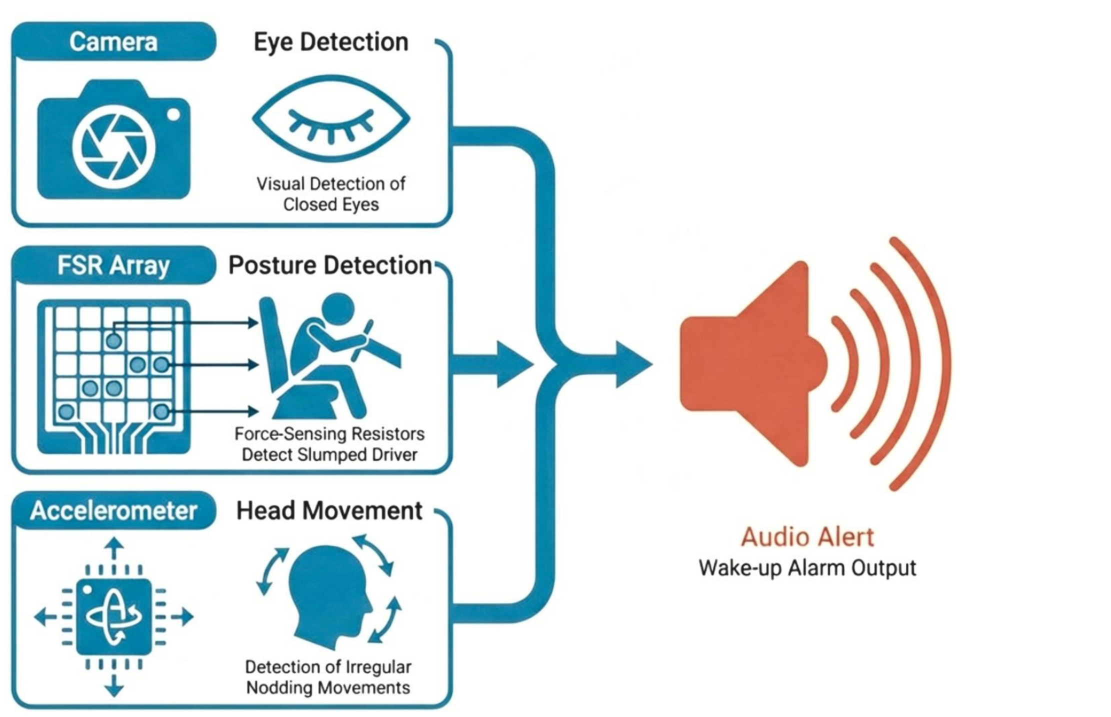
- Each year, ~**100,000 crashes** are caused by **fatigued drivers** (National Highway Traffic Safety Administration, n.d.)
- Driver monitoring systems are not robust, relying on simple image processing and vehicle-based data (Insurance Institute for Highway Safety [IIHS], 2024)
- Vehicle-based systems detect drowsiness *after* hazardous driving performance has **already occurred**

Mitigation

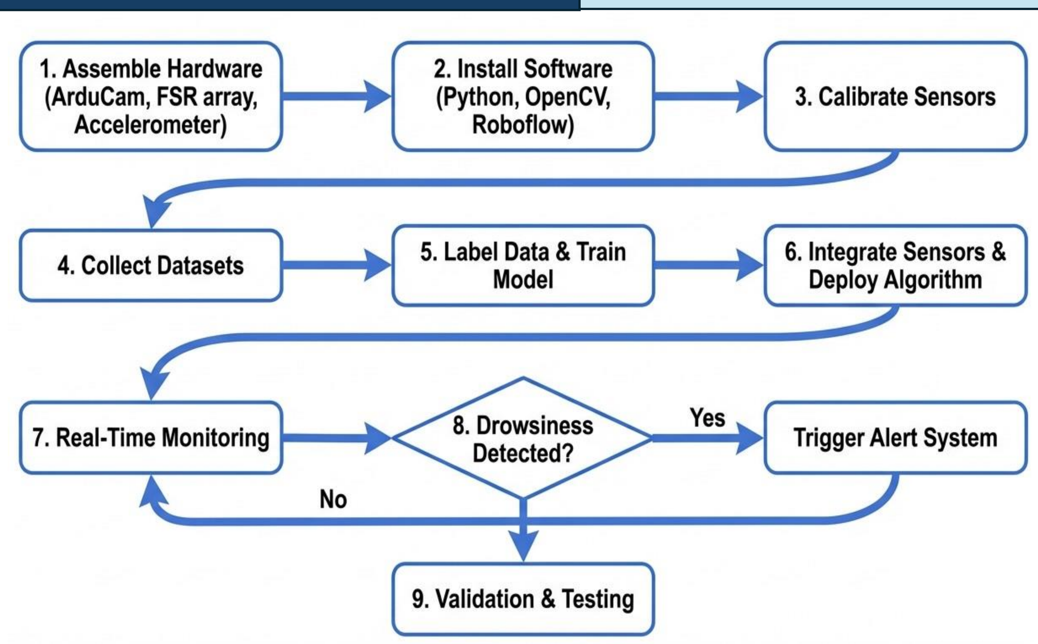
- A **camera** can be programmed to detect **eyelid closure** by implementation of a **Machine-Learning model**.
- **Force-sensing resistors (FSRs)** can monitor **postural shifts** by lack of applied pressure
- **Accelerometer** can monitor **head movement**
- **YOLOv8 Machine Learning** allows for **increased accuracy and reliability** of visual detection methods, processing patterns from continuous data



Visual Abstract



Methodology



Decision Matrix

Criteria	Computer Vision	FSR Array	Accelerometer	Vehicle Input	EEG
Cost	4 (Low)	5 (Low)	3 (Low)	2 (High)	1 (High)
Ease of Implementation	5	4	5	1	1
Hindrance	5 (Non-contact)	4 (Non-obtrusive)	2 (Contact)	3 (Non-contact)	1 (Highly obtrusive)
Prediction capability	4	3	3	4	5
Total Score	18	16	13	10	8
Accept/Reject	Accept	Accept	Accept	Reject	Reject

Engineering Need

Driver fatigue is one of the leading causes of vehicular accidents, and existing single-module detection methods for mitigation currently lack the accuracy and robustness required for reliable real-time safety.

Objective

The objective was to develop a Driver Fatigue Intervention system that integrates a fusion sensor inputs and machine learning for accurate, real-time detection and mitigation of early-stage driver drowsiness.

Results

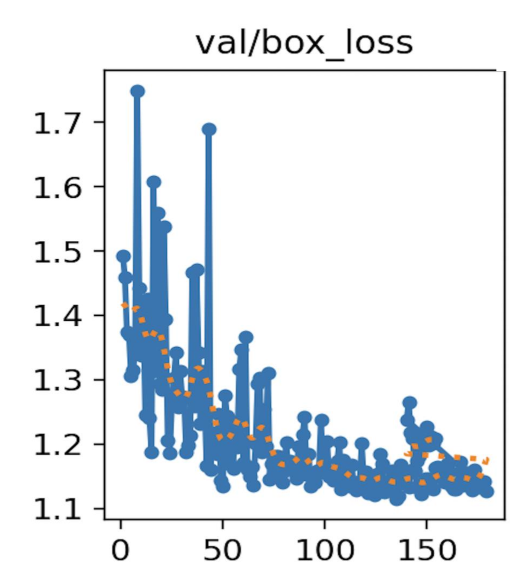


Figure 1: This graph tracks the model's ability to **precisely locate the driver's face and eyes**. The steady decline in validation loss shows this spatial localization.

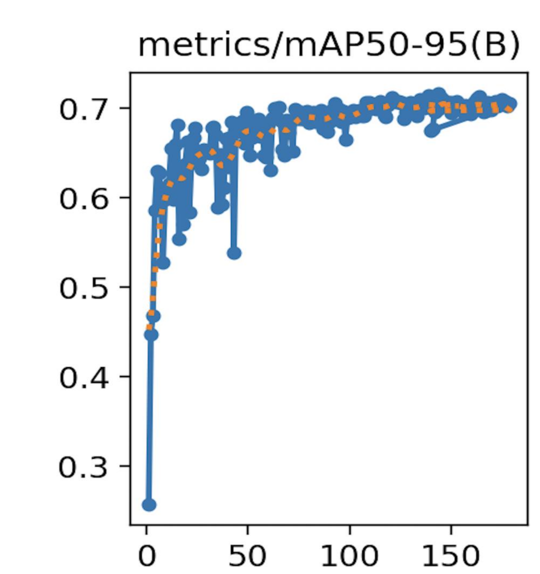


Figure 2: This graph depicts the Mean Average Precision (mAP) across multiple Intersection-over-Union (IoU) thresholds. The 0.7 mark signifies **correct detection** of facial features **even under challenging conditions**.

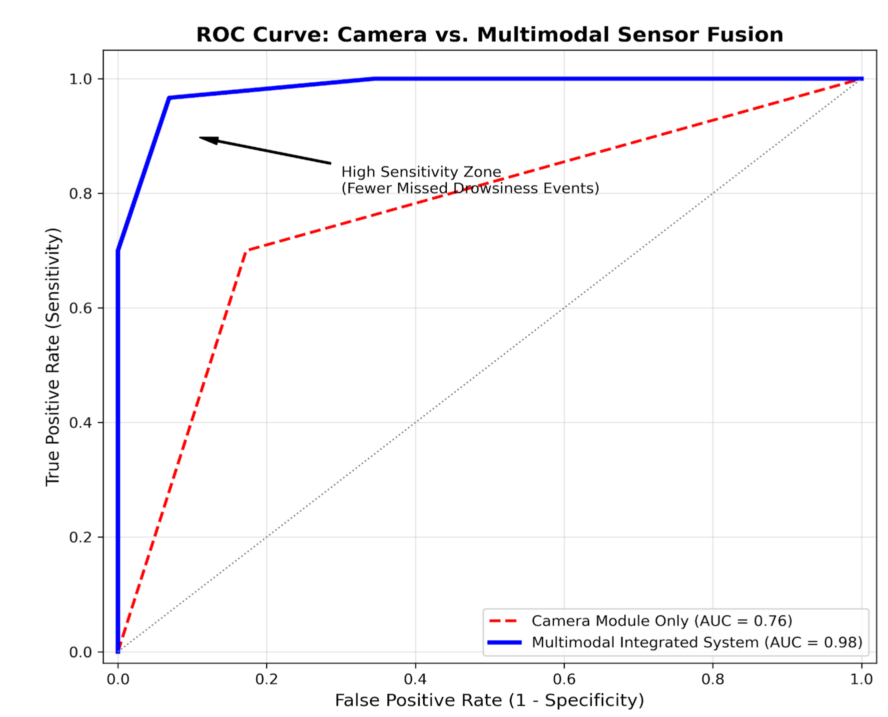


Figure 3: Receiver Operating Characteristic (ROC) analysis comparing stand-alone computer vision performance against the Integrated Multimodal System. The Integrated System achieved an **AUC of 0.98**, indicating **reliability** under varied stressors.

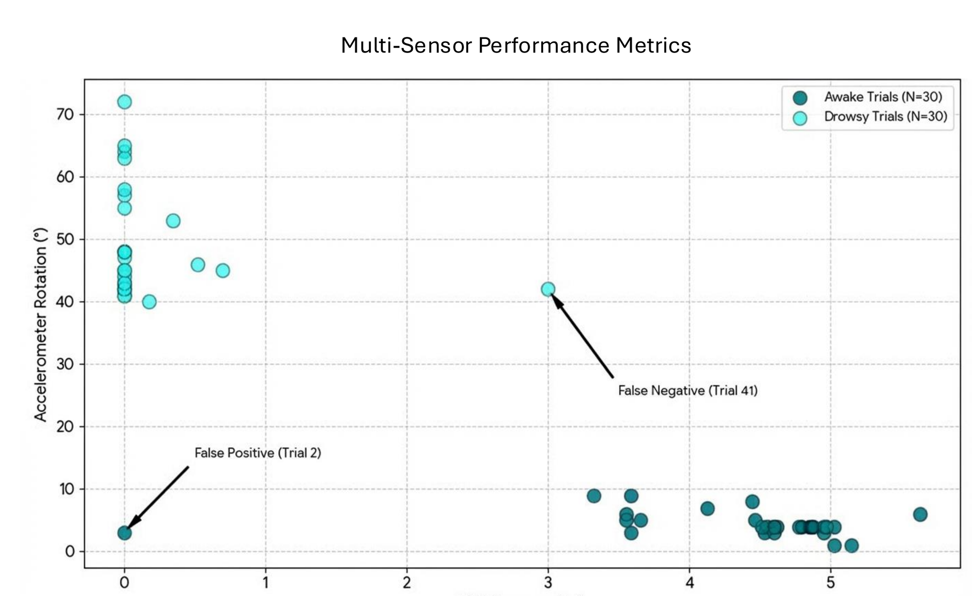


Figure 4: The FSR and Accelerometer readings during Awake and Drowsy states are depicted. The drowsy state signifies a **clear physical slump that differs from active driving posture**. Student's T-test was conducted to verify a significant difference between states: **p < 0.0001.** There is convincing statistical evidence that the sensor readings during the awake state are significantly different from the drowsy state.

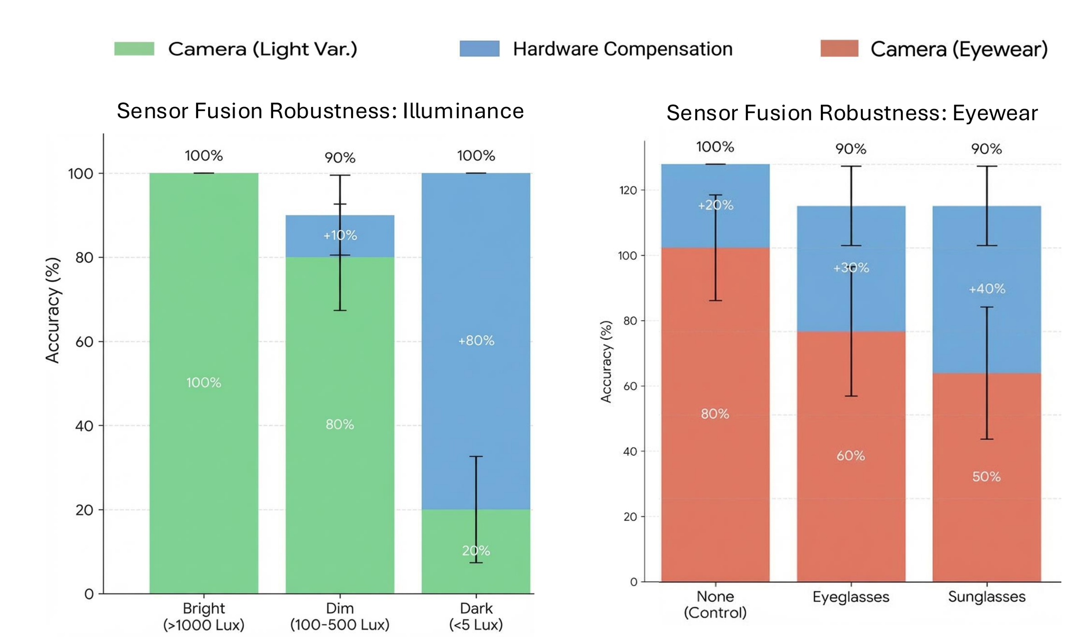


Figure 6: System accuracy across lighting conditions. The integrated system signifies **high reliability despite obstructions**. Fisher's Exact Test was conducted to verify significant performance differences between the integrated system and standalone software: ****p<0.01.** There is convincing evidence that the sensor fusion is significantly more robust against illuminance interference.

Figure 7: System accuracy across eyewear types. (***p<0.05)** There is convincing evidence that the sensor fusion is significantly more robust against eyewear interference.

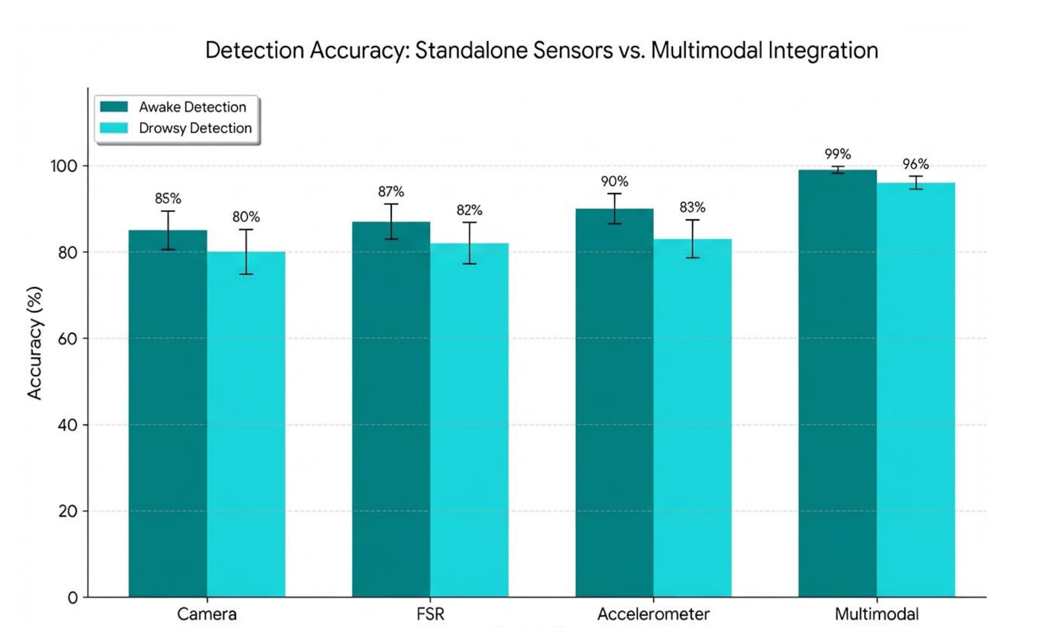


Figure 5: The detection accuracy of individual sensors compared to the integrated multimodal system. Fisher's Exact Test was conducted for the trials to verify a significant difference in accuracy: ***p<0.05.** There is convincing evidence that the **multimodal integration is significantly more accurate** than each individual module's performance.

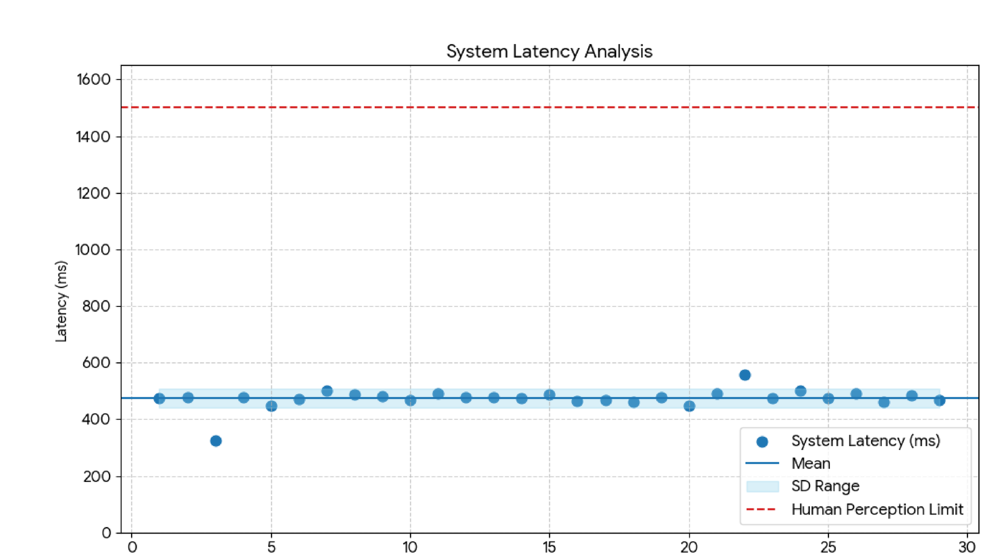
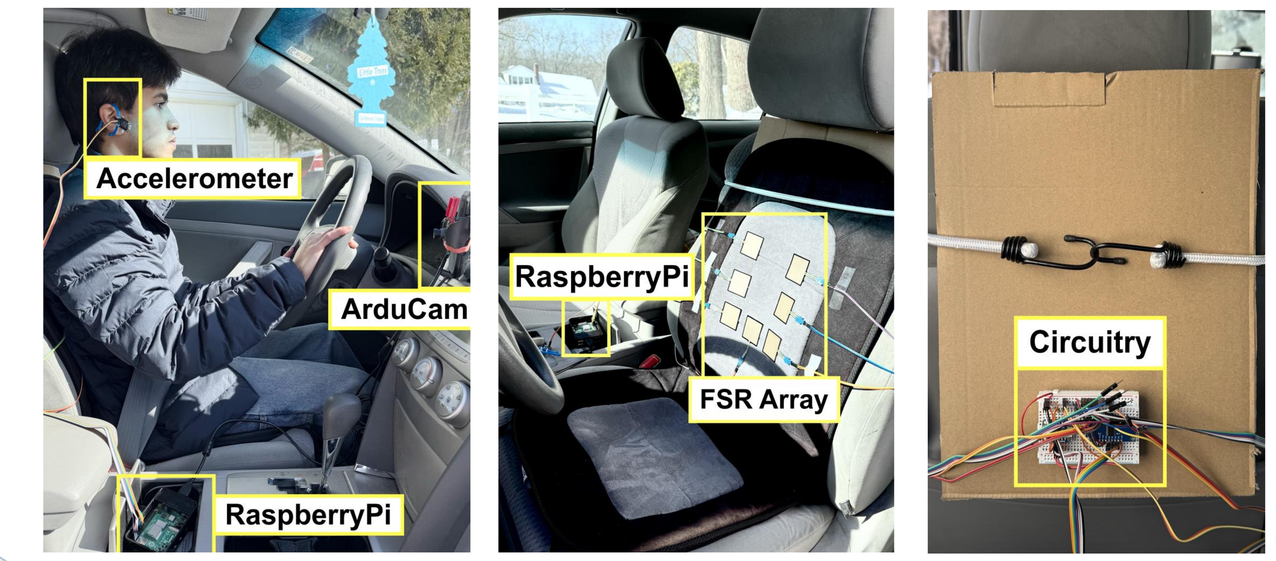


Figure 8: The system **latency** in milliseconds over multiple trials. The mean latency of **473.88 ms** signifies a response time well under the human standard (1500 ms). A One-sample T-test was conducted to verify a significant difference between system speed and the 1500 ms human perception threshold: ****p < 0.0001.** There is convincing statistical evidence that the system response time is significantly faster than human perception.

Final Design



Conclusions

- **Integrated Detection:** Developed a **real-time, low-cost** multimodal system combining **YOLOv8 computer vision** with **FSR** and **accelerometer** hardware.
- **Robust Tracking:** Achieved precise facial tracking via YOLOv8, maintaining vision accuracy across diverse seating positions and head tilts.
- **Superior Accuracy:** Sensor fusion significantly outperformed standalone modules, achieving an **AUC of 0.98.**
- **Environmental Resilience:** Maintained high reliability during lighting variation and eyewear use; physical sensors compensate for visual obstructions.
- **Validated Biomarkers:** Slumping and head tilting are statistically significant indicators of drowsiness.
- **Low-Latency Intervention:** System achieved a response time of **473.88 ms**, enabling **immediate safety intervention**.

Driver is reawakened

Safer and controlled driving

Potential accidents avoided

More lives saved

Future Directions

- **Expanded Dataset Diversity:** Incorporate a wider range of driver demographics, clothing types, and headwear to improve YOLOv8 model performance and minimize any possible source of overfitting.
- **Accelerometer:** Implement wireless connectivity.
- **Heart Rate Monitor:** Integrate a module to monitor heart rate patterns to detect changes indicative of drowsiness.
- **Haptic Alerts:** Integrate a haptic intervention system, such as seat or steering wheel vibrations to immediately alert the driver upon detection.

References

Ahmed, M., Masood, S., Ahmad, M., & Abd El-Latif, A. A. (2022). Intelligent driver drowsiness detection for traffic safety based on multi CNN deep model and facial subsampling. *IEEE Transactions on Intelligent Transportation Systems*, *23*(10), 19743–19752. https://doi.org/10.1109/TITS.2021.3134222

Bhadoriya, A. S., Vegamoor, V., & Rathinam, S. (2022). Vehicle Detection and Tracking Using Thermal Cameras in Adverse Visibility Conditions. *Sensors (Basel, Switzerland)*, *22*(12), 4567. https://doi.org/10.3390/s22124567

Dasgupta, A., George, A., Happy, S. L., & Routray, A. (2013). A vision-based system for monitoring the loss of attention in automotive drivers. *IEEE Transactions on Intelligent Transportation Systems*, *14*(4), 1603–1610.

Ed-Doughmi, Y., Idrissi, N., & Hbali, Y. (2020). Real-Time System for Driver Fatigue Detection. Based on a Recurrent Neuronal Network. *Journal of imaging*, *6*(3), 8. https://doi.org/10.3390/jimaging6030008

Ponn, T., Kröger, T., & Diermeyer, F. (2020). Performance Analysis of Camera-based Object Detection for Automated Vehicles. *Sensors (Basel, Switzerland)*, *20*(13), 3699. https://doi.org/10.3390/s20133699

U.S. Department of Transportation. (n.d.). *Drowsy Driving*. NHTSA. https://www.nhtsa.gov/