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

Grant Proposal

Suhrit Ghosh

Massachusetts Academy of Math and Science at Worcester Polytechnic Institute

Worcester, Massachusetts

Author Note

I would like to thank my project advisor Dr. Crowthers for his guidance through the course of this project, as well as my parents for their support.

Executive Summary

The overall aim of this project is to develop and construct a multimodal driver fatigue intervention system that integrates sensor inputs and machine learning to provide accurate real-time detection and mitigation of driver drowsiness in the early stages. This involves developing a hybrid system that analyses data from three modules. The first is a computer vision module that detects behavioural fatigue (Dasgupta et al., 2013). The second module consists of a force-sensing resistor (FSR) to monitor shifts in driver posture that are characteristic of drowsiness, such as slumping. The third module is comprised of an accelerometer to detect irregular head movement. The machine learning algorithm analysis of the combined data from these three modules is expected to offer a significant improvement in detection accuracy and reliability over existing single-module methods (Ed-Doughmi et al., 2020). This ensures the proper addressing of the issue of driver drowsiness to effectively reduce traffic-related fatalities.

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

Introduction

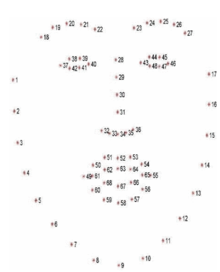
Driver fatigue is a major cause of vehicular accidents despite existing single-module detection methods; they often lack the accuracy and robustness required for reliable real-time safety systems. The goal of this project is to develop and construct a Multimodal Driver Fatigue Intervention system that integrates various sensor inputs and machine learning to provide accurate, real-time detection and mitigation of early-stage driver drowsiness. This project aims to develop a solution to the issue that simple image processing techniques alone are often insufficient to address the complexity of driver state monitoring (Jose et al., 2021), a limitation that is frequently highlighted in surveys of driver behaviour analysis (Wang et al., 2022).

Fatigue and Behavioural Indicators

Early detection systems of driver fatigue primarily rely on behavioural indicators. One such indicator is eyelid closure, with methods focused on the Eye Aspect Ratio (EAR) technique, proving effective for initial screening (Sathasivam et al., 2020). The EAR value is calculated using facial feature markings (as depicted in Figure 1 below), for which a score of less than 0.25 inclusive indicates a drowsy state (Dasgupta et al., 2013).

Figure 1

Facial Mapping



Note. By Jose et al., 2021, depiction of facial landmark predictor.

However, achieving high accuracy in detecting drowsiness on the driver's face visually in real-world driving environments require moving beyond basic image processing to incorporation of more advanced methods, such as multi-Convolutional Neural Network (CNN) Deep Models, which utilise facial subsampling for enhanced robustness (Ahmed et al., 2022). The research in this area essentially established that the driver could be monitored using vision-based systems, although with limitations.

Multimodal Sensor Integration

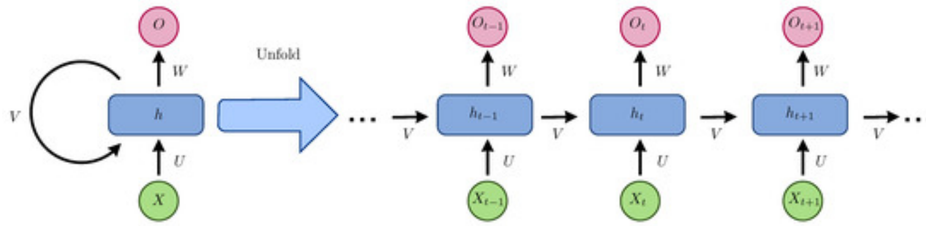
A multimodal approach is essential to overcome the limitations of purely visual systems. Vision-based monitoring often faces challenges with changes in lighting or other obscurities (Wang et al., 2022). This project integrates inputs from an accelerometer for head movement analysis and a posture sensor for seat data (applied pressure measurements). This fusion positional data is critical for reliability of the design, for one sensor to compensate for the weaknesses and shortcomings of another, minimising the output of spurious data. Furthermore, emerging technologies such as Infrared 3D cameras could be utilised to monitor the state of the driver, providing additional data (Wang & Zhao, 2022).

Machine Learning and Sensor Fusion

The combination of data from these sensors requires complex algorithms. Machine learning algorithms are a viable approach, processing continuous input from the camera, accelerometer, and posture pressure FSR to identify patterns between physical indicators. Additionally, the deployment of models utilising CNN and RNNs (as depicted in Figure 2 below) ensures that the system can learn the unique patterns of drowsiness specific to the driver (Ahmed et al., 2022).

Figure 2

Recurrent Neural Networks



Note. By Ed-Doughmi et al., 2020, depiction of a recurrent neural network.

This integrated machine learning approach which incorporates Recurrent Neuronal Networks (Ed-Doughmi et al., 2020) leads to an effective comprehension of the state of the driver, and the driver can subsequently be alerted to avert a potential collision if necessary.

Section II: Specific Aims

This proposal's objective is to construct and validate a multimodal driver fatigue intervention system.

The work proposed will combine data from three sources – a computer vision system to detect eye closure in a state of drowsiness, an array of force-resisting sensors (FSRs) for identifying a slumped posture, and an accelerometer to detect sudden head movement – to determine if the driver is in a state of drowsiness or not.

Specific Aim #1: Develop and train a computer vision module to accurately detect a drowsy state. This is measured by accuracy in classifying open and closed eyes.

Specific Aim #2: Position FSRs in an array across the backrest of the driver's seat in the car to collect data on posture variation.

Specific Aim #3: Setup the accelerometer to effectively analyse head tilt in degrees.

Specific Aim #4: Achieve data fusion between the three modules. This is measured by successful data capture simultaneously from all three sensors with a high accuracy rate.

Specific Aim #5: Develop and implement an alert system that successfully activates and alerts the driver within 1.5 seconds of a drowsy state detection.

The expected outcome of this work is that a fully functional driver drowsiness detection system is developed, which will analyse combined data output from all three modules to effectively determine the state of the driver, and subsequently trigger the alert to awaken the driver if necessary.

Section III: Project Goals and Methodology

Relevance/Significance

With increases in the number of automotive accidents due to the driver being in a state of drowsiness, the proposed multimodal drowsy driver detection system directly addresses this issue by attempting to awake the driver if they are on the verge of falling asleep. By doing so, the driver regains proper control of the vehicle and awareness of their surroundings while on the road. As a result, a potential accident may be effectively averted. Furthermore, this device is relatively low-cost, and can be used in any car.

Innovation

The successful implementation of such driver-assistive safety technologies relies on the system's ability to demonstrate consistent accuracy in a driving environment. The proposed multimodal driver state monitoring system is essentially an innovation, as it overcomes the critical limitations of current commercial solutions through integration of sensor fusion.

Commercial driver drowsiness detection systems currently deployed in vehicles—such as Mercedes-Benz's Attention Assist—primarily rely on vehicle-based data. These systems only account for changes in steering wheel position, lane position, and general vehicle orientation on the road to infer driver fatigue (Fletcher Jones Motorcars, n.d.). While these are somewhat effective, a significant limitation is that they detect drowsiness *after* hazardous driving performance has already occurred. Other systems utilise *only* image-based monitoring to monitor the driver's face for signs of fatigue. Such systems, however, are vulnerable in various light conditions, do not work effectively when the driver is partially obscured, and often collect spurious data.

The core innovation of this project is the development of a hybrid system that is robust and consistent in its detection, not found in the previously mentioned solutions. By integrating three data input modules, the points of failure for each are mitigated. The following chart provides an overview of the advantages and improvements over current systems of each module proposed for this project:

Table 1

Sensor Module Information

Sensor Module	**Primary Data Type**	**Advantage**	**Improvement over Current Systems**
Computer Vision (Camera)	**Behavioural**	Non-contact, high-quality predictive tracking of eye state using CNN-based object detection.	Provides visual data using advanced CNN machine learning systems, and is reinforced by other sensors when visual data fails.
Force-Sensing Resistor (FSR) Array	**Positional** (Driver Posture)	Monitors subtle shifts in driver posture (slumping, leaning).	Direct improvement over vehicle-based systems, assessing the driver's body state *before* it affects steering, enabling earlier detection of fatigue.
Accelerometer	**Kinematic** (Head Movement)	Detects irregular or exaggerated head movements characteristic of microsleeps or severe fatigue.	Provides a physical, non-visual confirmation of major lack of driver attention, serving as independent validation of the visual/posture data.

Note. The features, advantages, and improvement points of three means of data input.

The combination of data from these modules being processed by the machine learning algorithm allows for the learning of complex, individual-specific correlations to identify drowsiness patterns with higher confidence. This fusion allows for more accurate and less faulty detection of driver drowsiness than previous designs, and furthermore, allows for early intervention by alert to prevent a collision.

Methodology

Specific Aim #1: Develop and train the computer vision module to accurately detect a drowsy state.

Justification and Feasibility. In terms of feasibility of using image processing for driver fatigue detection, it provides a simple and non-intrusive method to identify a key behavioural indicator of drowsiness. Techniques such as machine learning based object detection are a standard way of determining eye closure. The major justification for this aim lies in improving robustness and accuracy beyond simple image processing, a limitation of current systems. By focusing on the development of more advanced methods, such as the utilisation of CNN deep model training and validation, the system can achieve enhanced performance. The visual data provided to the final algorithm is then reliable, accounting for changes in light, glare, etc.

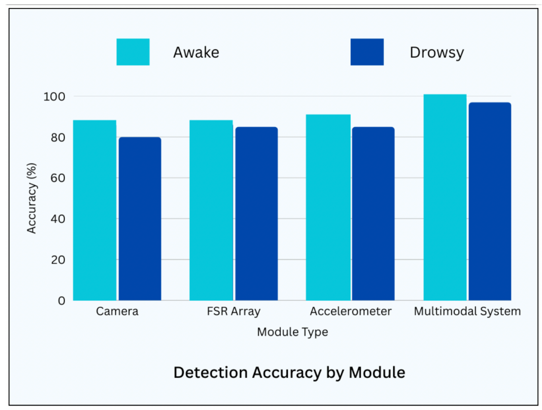


Figure 1: Accuracy rates for each module tested independently for both awake and drowsy scenarios, as well as combined testing of all three modules.

Summary of Preliminary Data. The camera vision system achieved an overall accuracy of above 80%, with rates of detection of awake states of the driver being slightly higher than those of drowsy states. Of the three modules, the camera-based object detection system had the lowest rates of accuracy.

Expected Outcomes. The expected outcome is a fully developed and effectively trained computer vision module capable of object detection in real-time, with a 95%+ level of accuracy and robustness across diverse testing conditions. The module will be ready to send data to the machine learning algorithm in the final system.

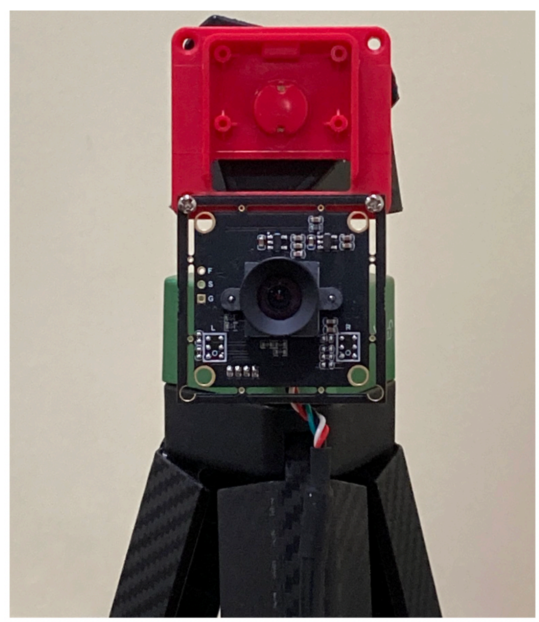


Figure 2: Adafruit ArduCam for object detection.

Potential Pitfalls and Alternative Strategies. The primary risk involves achieving a high accuracy level in the classification of open or closed eyes across diverse conditions. A poorly trained model might generate an unacceptable number of false positives or negatives, contributing to unreliability (Ponn et al., 2020). The capability of the camera hardware itself could also hinder real-time performance. To address these issues, a pre-trained deep learning model for facial recognition could be used as a supplement to the manually created dataset model, increasing accuracy rates. Furthermore, the camera could be replaced with a higher quality infrared camera, noted for its mitigation of poor lighting, glare, and darkness, and improved image quality (Bhadoriya et al., 2022).

Specific Aim #2: Position FSRs in an array across the backrest of the driver's seat in the car to collect data on posture variation.

Justification and Feasibility. The force-sensing resistor (FSR) array provides crucial data input that does not rely on visual means. This overcomes the fundamental limitations of purely image-based systems (such as excess false positives due to environmental conditions). Monitoring shifts in driver posture, such as slumping, is characteristic of drowsiness and provides an essential positional indicator. The continuous pressure distribution captured by the FSRs offers a direct measure of body fatigue, complementing behavioural data from eye closure detection. The FSR array is feasible due to the FSRs' low cost, refined sensitivity, and flexible form, allowing for integration into the seat fabric for continuous data collection without being obtrusive to the driver. This dedicated positional data serves as critical validation for the visual module when the latter is compromised by previously mentioned hindrances.

Summary of Preliminary Data. Drowsiness was clearly characterised by a near-total loss of seat back pressure. FSR readings displayed approximately 0 lb. for the drowsy states, and readings up to 22 lb. for the awake states. The FSR array overall achieved detection accuracies of above 80% for the both drowsy and awake states, with the awake state having a slightly higher accuracy.

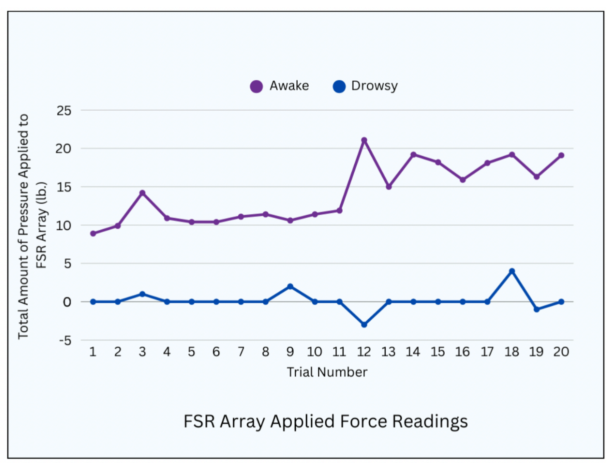


Figure 3: The FSR Array demonstrates a significant loss of seatback pressure during drowsiness, correlating with the driver slumping forward.

Expected Outcomes. The expected outcome is a functional FSR array integrated into a driver's seat that reliably provides continuous and quantitative data on changes in weight distribution at pressure points. This will be used as the validated source of positional data for input into the final sensor fusion algorithm, allowing for a comprehensive detection of fatigue-based changes in the driver's state indicated by postural shifts.

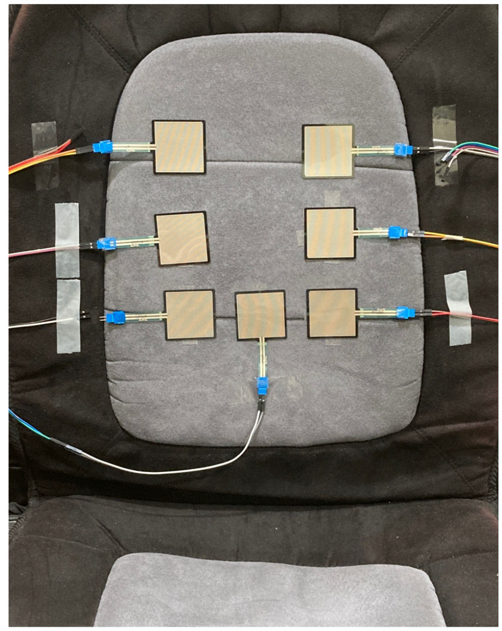


Figure 4: A 7-unit FSR array attached to the driver's seat back to detect pressure or lack of pressure in certain points of the lumbar and thoracic regions of the driver's back.

Potential Pitfalls and Alternative Strategies. Spurious data, such as a driver shifting position for comfort or reaching for some vehicle control could be misinterpreted by the system, potentially leading to occasional false readings. If such irrelevant detections are made, the algorithm would be improved to only flag sustained positional changes correlating with fatigue slumping patterns.

Specific Aim #3: Setup the accelerometer to effectively analyse head movement.

Justification and Feasibility. The accelerometer serves as a third sensor, intended to detect irregular head movement, such as nodding off. This is necessary to interpret an especially dangerous phase of fatigue—microsleeps—which may involve loss of control, but not detected as fatigued eye closure or slumping. Accelerometers are quite small and low-powered devices, allowing them to be easily positioned on the driver's head.

Summary of Preliminary Data. The accelerometer characterised drowsiness by a forward head tilt of 20 degrees or more from the standard resting position. The accelerometer-based detection overall achieved detection accuracies of slightly above 80% for the both drowsy and awake states, with the awake state having a slightly higher accuracy.

Expected Outcomes. The expected outcome is a calibrated accelerometer system that effectively captures and processes pitch, yaw, and roll data to identify movement patterns that indicate drowsiness, specifically the slow drift and jolt motions associated with microsleep. This will be used as a second source of positional data for input to the final sensor fusion algorithm.

Potential Pitfalls and Alternative Strategies. Normal, intentional head movements (checking side mirrors, looking at drive controls) could be misinterpreted as fatigue, causing a risk of false positive. To combat this, the accelerometer data would be processed in conjunction with the camera module's visual data. If the visual module does not detect eye closure while the accelerometer detects head movement as that of a fatigued state, the accelerometer data would be weighted lower for such instances. Likewise, the converse is applicable as well, where the visual model's positive result would be weighted lower than the accelerometer if it were to not detect fatigue. However, another hindrance of the accelerometer is that there would be physical wires from the accelerometer running to the RaspberryPi, which is not optimal. Nevertheless, this could be resolved by implementing an intermediate wireless module to send data.

Specific Aim #4: Achieve data fusion between the three modules.

Justification and Feasibility. Data fusion is the central innovation of the system that overcomes limitations of current detection methods. The fusion of data from the camera module, FSR array, and accelerometer allows the system to construct a robust detected driver state. This integrated approach is essential for enhancing detection accuracy and reliability. This is indeed feasible, due to the utilisation of machine learning (ML) algorithms, incorporating recurrent neural networks (RNNs) designed to process and identify patterns within time-based multi-dimensional data input.

Summary of Preliminary Data. Testing of the combined sensor fusion system yielded the highest rates of accuracy: 99% and 96% for the awake and drowsy states respectively. Essentially, by combining visual (camera), force (FSR), and positional (accelerometer) data, the system maintains safety even if one sensor's data becomes spurious or noisy.

Expected Outcomes. The expected outcome is the successful development and implementation of an ML-based sensor fusion algorithm. This is measured by the successful and simultaneous capturing of data from all three modules with high correlation between the algorithm's output and the actual state of drowsiness. The output, if positive, would be used to trigger the final alert. As for overfitting, data could be diversified through augmentation and split into groups for training, especially for the camera module data that would be fed to the algorithm.

Potential Pitfalls and Alternative Strategies. The most significant challenge is ensuring synchronisation and minimising latency and overfitting for the detection and output. Such software-related issues could be mitigated with hardware-based correction by a micro-controller unit for consistent time-stamping. Overfitting

Specific Aim #5: Develop and implement an alert system that successfully activates and alerts the driver within 1.5 seconds of a drowsy state detection.

Justification and Feasibility. The final component of this project is the intervention system to prevent an accident at an earlier point. Rapid mitigation is the ultimate goal of the system, waking the driver in time. The feasibility of achieving a fast system is dependent on the processing speed of the fusion algorithm developed in the previous specific aim. The alert system itself is quite feasible due to readily available technology for audio-based output.

Expected Outcomes. The expected outcome is a functional driver drowsiness detection and intervention system. The system will successfully trigger the alert system upon a validated detection of drowsiness in the driver. The alert system directly addresses the issue of drowsy driving by waking up the driver in time.

Potential Pitfalls and Alternative Strategies. A potential risk is the issue of the driver becoming desensitised to frequent alarms (depending on the frequency of driver drowsiness while driving), causing them to become compelled to disable the system entirely. As a solution to this issue, a multi-staged alert system could be implemented. The system could begin with haptic warnings, such as steering wheel vibration. If drowsy driving persists, the warning will consequently escalate to an audio alarm. The machine learning model could also potentially be adapted to learn the driver's unique response to the alarms, maxmising effectiveness and minimising irritation of the driver.

Section III: Resources/Equipment

The construction of the Multimodal Driver Fatigue Intervention system required a high-definition vision module for real-time facial monitoring and an array of force-sensitive resistor (FSR) posture sensors to collect seat pressure data. To account for the irregular head movements characteristic of microsleeps, an accelerometer was integrated into the sensor suite. The computing environment was established on a high-performance computer utilizing Roboflow for the management of a diverse, annotated dataset exceeding 6,000 images, which served as the foundation for training the multi-Convolutional Neural Network (CNN) architecture. Additional materials included the hardware interfaces such as an Analog-Digital System converter and multiplexer necessary to bridge the sensor array with the processing unit and the physical components of the alert system.

Section V: Ethical Considerations

For the safety of myself as well as other drivers, this drowsy river detection system shall not be tested on public roads in a moving vehicle. Simulating a drowsy scenario in real life is as equally hazardous as a real scenario involving a drowsy driver, and it is not my intention to bring harm to anybody.

Section VI: Timeline

October:

- Finalise research
- Select all components
- Setup the RaspberryPi OS and install all necessary coding libraries

November:

- Order/receive all hardware
- Calibrate/test camera, accelerometer, and FSR sensors
- Code foundation for basic detection system
- Collect preliminary data for alert/drowsy states needed for model training

December:

- Finalise code for sensor fusion
- Create dataset for training
- Train machine learning model
- Integrate alert system
- Conduct control group testing and establish baseline F1 score
- Conduct experimental group testing

January:

- Finish experimental group testing
- Perform all data analysis
- Calculate final scores
- Complete statistical analysis
- Write final report
- Prepare presentation and documentation materials

Section VIII: 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

Jose, J., Vimali, J. S., Ajitha, P., Gowri, S., Sivasangari, A., & Jinila, B. (2021). Drowsiness Detection System for Drivers Using Image Processing Technique. In *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1527–1530). IEEE. https://doi.org/10.1109/ICOEI51242.2021.9452864

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

Sathasivam, S., Mahamad, A. K., Saon, S., Sidek, A., Som, M. M., & Ameen, H. A. (2020). Drowsiness detection system using Eye Aspect Ratio technique. In *2020 IEEE Student Conference on Research* *and Development (SCOReD)* (pp. 448–452). IEEE. https://doi.org/10.1109/SCOReD50371.2020.9251035

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

Wang, J., He, T., Lu, Z., Yang, Y., & Wei, R. (2022). A survey on driver behavior analysis from in-vehicle cameras. *IEEE Transactions on Intelligent Transportation Systems*, *23*(8), 10186–10209. https://doi.org/10.1109/TITS.2021.3126231

Wang, X., & Zhao, Q. (2022). A new method for monitoring the state of drivers-occupant based on infrared 3D camera. In *2022 China Automation Congress (CAC)* (pp. 6944–6949). IEEE. https://doi.org/10.1109/CAC57257.2022.10055849

Fletcher Jones Motorcars. (n.d.) What Is Mercedes-Benz ATTENTION ASSIST®? *Fletcher Jones Motorcars*, www.fjmercedes.com/mercedes-benz-attention-assist/