DRIVER SITUATION AWARENESS & FATIGUE EVALUATION SYSTEM (DRIVESAFE) VESTEL MOBILITY A.Ş.

PROJECT TEAM
Ertuğ Kayra Alemdar
Ömer Burak Avcıoğlu
Orkun Bayri
Yusuf Berkan Gökçe
Emirhan Yağcıoğlu
Selin Yurttaş

COMPANY MENTOR
Berkay Bana
ACADEMIC MENTOR
Assoc. Prof. Cem Tekin
TEACHING ASSISTANT
Mehmet Efe Lorasdağı

Abstract. In recent years, the rising number of traffic accidents caused by driver distraction and fatigue has emphasized the need for in-vehicle monitoring systems to enhance road safety. DriveSafe is a cutting-edge Driver Monitoring System (DMS) developed in collaboration with Vestel, designed to identify and react to high-risk driver behaviors such as sleep, drowsiness, prolonged distraction, and unresponsiveness. The system utilizes of a single-camera setup integrated with advanced deep-learning models that run on an NVIDIA Jetson TX2 platform. By analyzing facial and eye movements in real-time, DriveSafe accurately detects early signs of fatigue and driver distraction—achieving over 90% accuracy—and triggers warning mechanisms instantly to help prevent potential hazards. The DriveSafe system incorporates a suite of specialized machine learning and computer vision models, including those for face detection, eye detection eye state classification, and eye tracking. The models are developed and optimized to work under a variety of scenarios, as well as diverse driver characteristics. Built in accordance wih Euro NCAP standards, DriveSafe is designed to be both cost-effective and easily integrated into existing vehicle architectures. Precise testing and iterative improvements guarantee that DriveSafe delivers high performance and robust safety.

PROJECT DESCRIPTION

DriveSafe is a real-time Driver Monitoring System (DMS) developed in collaboration with Vestel to enhance road safety. It addresses the pressing issue of traffic accidents caused by driver distraction, fatigue, and unresponsiveness, factors responsible for a significant portion of global accidents. The system continuously analyzes the driver's facial and eye behavior to detect signs of risk and immediately issues warnings to prevent potential hazards. The expected outcome is a cost-efficient, deployable, and EURO NCAP-compliant [1] DMS prototype capable of real-time operation and high detection accuracy.

The project motivation stems from the growing demand for intelligent vehicle safety systems, driven by the increasing number of road accidents involving distracted or fatigued drivers. Advanced solutions such as Tobii AutoSense and Smart Eye have demonstrated the effectiveness of driver monitoring through sophisticated sensing technologies and multi-camera configurations. While these systems offer high accuracy and extensive monitoring capabilities, DriveSafe explores an alternative approach by utilizing a single-camera setup to deliver essential driver state detection features. This strategy emphasizes cost-effectiveness and integration simplicity, aiming to extend the benefits of driver monitoring to a broader range of vehicles while still aligning with Euro NCAP standards.

DriveSafe offers a novel, affordable alternative by employing a single-camera setup combined with embedded machine-learning models deployed on the NVIDIA Jetson TX2 platform. Unlike existing systems, our solution reduces hardware complexity while preserving essential functionality, such as eye state detection, eye tracking, and fatigue detection.

The essential system requirements of DriveSafe are gathered solely from NCAP regulations ensuring full compliance with established standards. These requirements set the design and performance specifications of the system. The general requirements ensure that DriveSafe remains active by default at the start of each journey and cannot be easily deactivated by the driver. The software requirements ensure that all DriveSafe's detection algorithms operate with over 75% accuracy and complete each system loop in under 1 second, enabling reliable and real-time driver monitoring.

The Noise Variable Requirements ensure that the DriveSafe system maintains consistent performance across a diverse range of driver and environmental conditions. These requirements guarantee reliable operation for users of varying ages, sex, body stature, and skin tone. They also ensure robustness under typical occlusion scenarios, including different lighting conditions, the use of sunglasses, and the presence of facial hair.

The DriveSafe system detects two main types of driver distraction: long distraction and short distraction. The detection requirements for long distraction ensure that the system identifies when a driver continuously looks away from the forward road view for 3 seconds or longer, using gaze tracking through an eye-tracking algorithm. The detection requirements for short distraction ensure that the system captures repeated, brief glances away from the road, specifically when a driver looks away for a cumulative total of 10 seconds within a 30-second window. In both cases, when the defined thresholds are exceeded, the system initiates a visual or auditory warning to alert the driver.

The DriveSafe system is designed to detect three distinct stages of driver fatigue: drowsiness, and sleep. The detection requirements for drowsiness ensure that the system classifies a driver as drowsy when their Karolinska Sleepiness Scale (KSS) score exceeds 7, indicating reduced alertness. Additionally, the DriveSafe system monitors for unresponsive drivers. The detection requirements for unresponsive driver behavior ensure that the system responds if their eyes stay closed for more than 6 seconds. In both fatigue and unresponsiveness cases, detection is carried out using the eye closure detection algorithm. These requirements have been fully addressed in the design and implementation of the DriveSafe system.

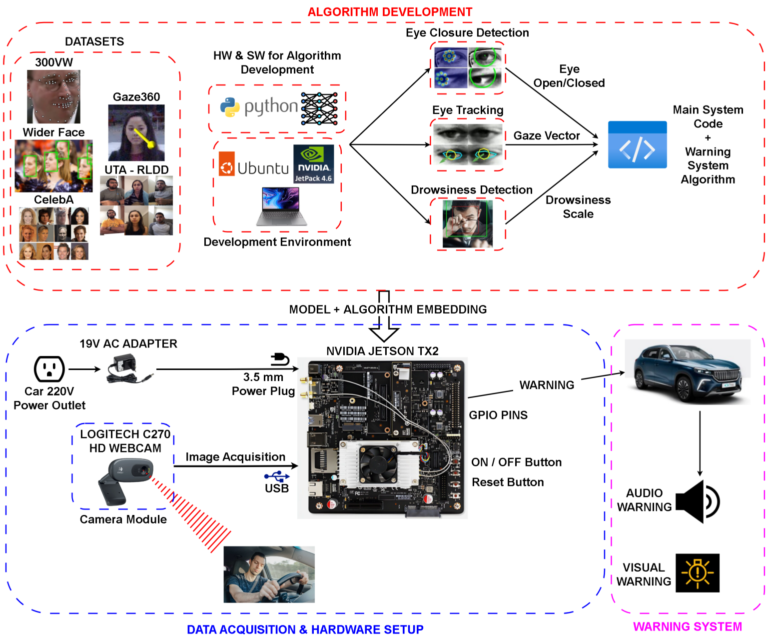

FIGURE 1. Big picture representation of DriveSafe

Figure 1 presents a high-level visual representation of the DriveSafe system, including its components, processes, and overall project structure. The system is composed of three main subsystems: algorithm development, data acquisition, and hardware setup, and the warning system. The algorithm development subsystem is responsible for designing and implementing the core deep learning and computer vision models: Eye Closure Detection, Eye Tracking, and Drowsiness Detection. These models are developed using Python and machine learning frameworks and then integrated into an embedded environment based on NVIDIA JetPack 4.6 for deployment on the NVIDIA Jetson TX2 platform. The data acquisition and hardware subsystem captures real-time driver images using a USB 3.0 Logitech C270 HD webcam. Captured frames are processed on the Jetson TX2, which operates on a 12V AC adapter, allowing compatibility with in-vehicle power systems. This subsystem is connected to the warning subsystem via the GPIO interface. The warning subsystem issues real-time feedback to the driver based on model outputs. Visual alerts are provided through an LED, while auditory alerts are delivered via a buzzer, ensuring immediate response to critical driver states.

MILESTONES

- **Model Completion:** The models are implemented and evaluated by performing the necessary research and dataset collection.
- **Model Embedding & Algorithm Design:** Developed models are embedded on the Nvidia Jetson TX2 platform and related algorithms are implemented.
- **Prototype Completion:** The prototype is tested both on python and in the cabin. The prototype is put in the vehicle, the algorithm under testing is toggled ON, then the driver simulates both nominal and dangerous behaviors. The system tries to classify them. In order to pass, model needs to achieve a 75% accuracy rate as defined in the requirements.
- **Final Product Delivery:** The final product meets all the previously specified requirement criteria and delivering the final product that meets these criteria.

DESIGN DESCRIPTION

To distinguish driver behavior between nominal and dangerous, the system first needed to detect the driver. For this, we use the camera placed on the dashboard. This camera is placed so that it provides a frontal view of the driver while leaving the passengers outside the frame as much as possible. When the system is run, the camera starts capturing in-cabin frames. Each frame is passed through many processing steps, and in the end, several parameters belonging to the driver are obtained.

We first apply a face detection algorithm in the frames gathered by this camera, Sample and Computation Redistribution for Efficient Face Detection (SCRFD) [2]. This provides sufficient performance while incurring small computation cost. If there is a case where more than one face is detected, a simple elimination step checks face locations and proximity. It deletes the passenger detections, thus the system gets the driver's face reliably.

After the driver's face is cropped from the frame, an eye detector neural network of our own design is used to find the driver's eye landmarks, which we use to both crop the eye images, and find the head tilt angle. This network is based on a MobileNetV2 backbone. While it is not designed for localization, the backbone provided good results after last pooling layers were removed to regain feature localization capability. Off the shelf facial landmark detectors or more specialized landmark detectors struggled to detect the eyes accurately as these models try to detect all facial features at once and eyes are not the model's main focus. The cropped eyes are given to another model, a eye state detector, to evaluate driver's eye openness. The head tilt angle is used further down the line by the gaze tracker model.

The eye state detector model is a simple CNN we designed to assign an openness score the inputted left and right eye images of the driver. A total eye openness score is obtained by checking both eyes and head orientation. This value represents how open the driver's eyes are in current frame. This value is obtained for all frames in the video feed, which is then used by warning algorithms to generate sleep and unresponsiveness state warnings.

The head tilt angle is used to rotate the cropped driver face image to horizontal alignment. Then a gaze tracking algorithm, based on L2CS-Net [3], produces driver face pitch and yaw angles which represent driver's eyes gaze angles with respect to the camera. These angles are vectorized, then subjected to a 2D-to-3D projection, converting the 2D gaze vector inside the image to a 3D gaze vector in the vehicle coordinate system. This vector is used to estimate the point on the front view plane the driver is looking at. This point is obtained for all frames in the video feed, as well.

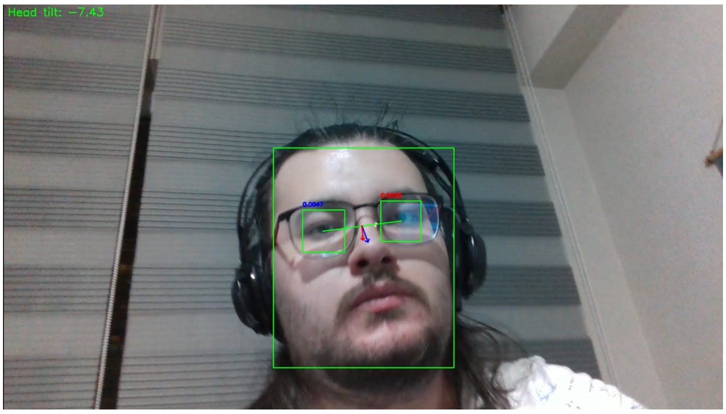

FIGURE 2. An evaluated frame

Upon receiving the eye-state, gaze-tracking, and fatigue outputs, the Warning Algorithm evaluates these signals into five different driver states; short distraction, long distraction, sleep, unresponsive driver, and drowsy driver, and issues visual (LED) and audible (buzzer) alerts accordingly. For long distraction, gaze vectors that fall outside the central 75% of the windshield toward the driver's side set a timer; if this situation persists for ≥ 3 seconds, a continuous LED/buzzer warning is activated until the driver's gaze returns inside the zone. For short distraction, all off-road gaze events within a rolling 30 seconds window are accumulated; when their total duration exceeds 10 seconds, the system issues a 5 seconds alert. In the sleep state, the eye-state detector starts a closure timer when eyes are closed for more than 3 seconds the system triggers a sleep warning, and more than 6 seconds escalates to an unresponsive driver warning. Finally, fatigue is assessed via a PERCLOS [4] metric over the past 15 minutes, comparing the percentage of eye closure to a personalized baseline; exceeding 1.8× this baseline activates the fatigue alert.

The system performance was evaluated by first evaluating performance for each model, then for the combined system. Model performances were evaluated by running tests on datasets not used to train the model. The combined system performance was tested by simulating different driving conditions and driver behavior on recording, then testing the system on nominal and dangerous driving scenarios.

RESULTS AND PERFORMANCE EVALUATION

The main goals of the system as stated in the requirements and milestones, were checked by testing each danger scenario in cabin by simulating several instances of each scenario with different drivers.

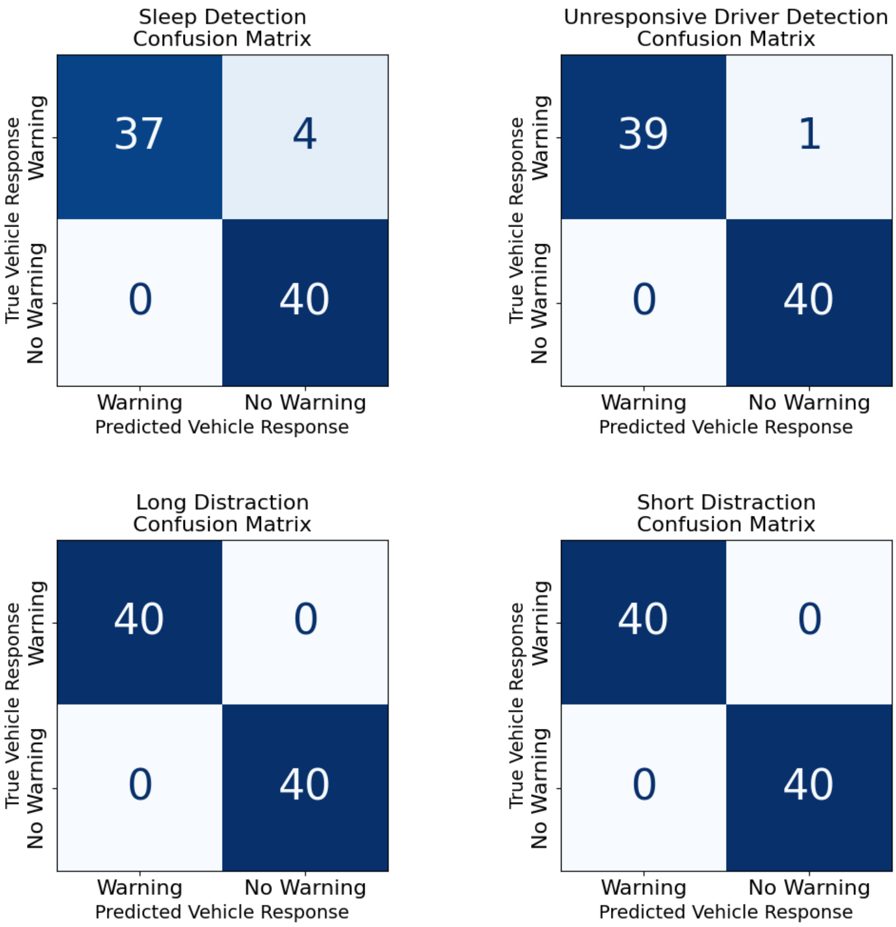

FIGURE 3. System detection performance for different scenarios

To evaluate the system performance, both the distraction and fatigue detection algorithms were tested independently. Video data was collected from four group members with varying facial features, heights, and other characteristics. Each participant contributed 10 nominal and 10 non-nominal videos for each scenario, where in the non-nominal videos, the driver intentionally looked away from the road for over 3 seconds for long distraction testing, and for over 10 seconds in a 30 second time frame for long distraction testing. Similarly, the driver closed their eyes continuously for 3 seconds for the sleep scenario, and for 6 seconds for the unresponsiveness scenario. The results show over 90% accuracy in fatigue scenarios, and a 100% accuracy in detecting long distraction cases. This high accuracy is believed to be influenced by the relatively small test sample. However, such accuracy was anticipated due to the design of the system: gaze locations and eye openness values are tracked over sliding windows, and a warning is issued if at least 80% of the frames within that window indicate the driver is looking away from the road, or have their eyes closed. As a result, while individual frames may contain false detections, the aggregated decision over time significantly boosts overall reliability and accuracy. In addition, during the tests and demonstrations, the warning system was successfully able to issue a visual and an auditory warning to the driver when a long distraction state was detected.

CONCLUSIONS AND FUTURE DIRECTIONS

DriveSafe, a real-time Driver Monitoring System (DMS), has been successfully developed to detect distraction, fatigue, and unresponsiveness using a single-camera setup combined with embedded deep-learning models. The system has been implemented on the NVIDIA Jetson TX2 platform and fulfills key design and performance requirements outlined by Euro NCAP, including real-time operation with more than 75% detection accuracy. Multiple detection models have been designed, trained, and integrated into the embedded environment, and a visual and auditory alert system has been implemented to notify the driver of critical conditions. The system has been validated under various environmental conditions and with a wide range of driver characteristics, demonstrating its reliability and robustness. For future development, the system may be extended to improve performance under extreme occlusion scenarios, such as facial coverings, headwear, or low-transmittance eyewear. Additional behavioral cues, such as yawning detection, head pose estimation, and facial expression analysis, can be incorporated to enhance fatigue and distraction classification. Integration with OEM vehicle systems and compliance with automotive safety standards represent important next steps for enabling real-world deployment. As an intelligent and modular safety solution, DriveSafe holds significant potential to reduce accident risks, promote responsible driving, and advance the development of safer vehicle technologies.

REFERENCES

[1] European New Car Assessment Programme (Euro NCAP), "Assessment Protocol – Safety Assist Safe Driving – Implementation 2023," Feb. 2024. [Online]. Available: https://www.euroncap.com/media/80158/euro-ncap-assessment-protocol-sa-safe-driving-v104.pdf. [Accessed: 18 Apr. 2025].

[2] J. Guo, J. Deng, A. Lattas, and S. Zafeiriou, "Sample and Computation Redistribution for Efficient Face Detection," in *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2021, pp. 1021-1030, doi: 10.1109/WACV48630.2021.00105.

[3] A. T. A. Mohamed and F. Lu, "L2CS-Net: Fine-Grained Gaze Estimation in Unconstrained Environments," in *2021 IEEE International Conference on Image Processing (ICIP)*, 2021, pp. 3927-3931, doi: 10.1109/ICIP42928.2021.9506466.

[4] Office of Motor Carrier Research and Standards, PERCLOS: A Valid Psychophysiological Measure of Alertness as Assessed by Psychomotor Vigilance, Washington, DC: U.S. Dept. of Transportation, 1998.

BEHIND THE SCENES