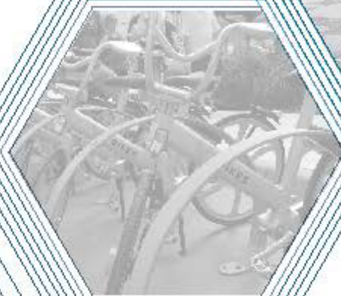

DEVELOPMENT OF A MONITORING SYSTEM FOR DRIVER READINESS IN PROLONGED AUTOMATED DRIVING

Shuchisnigdha Deb
Chen Kan
Anurag Pande
David Noyce

FINAL REPORT

Center for Transportation Equity, Decisions & Dollars

DEVELOPMENT OF A MONITORING SYSTEM FOR DRIVER READINESS IN PROLONGED AUTOMATED DRIVING

FINAL PROJECT REPORT

By:

Shuchisnigdha Deb
Chen Kan
The University of Texas at Arlington
Anurag Pande
Cal Poly State University
David A. Noyce
University of Wisconsin-Madison

Sponsorship:
CTEDD

For:
Center for Transportation, Equity, Decisions and Dollars **(CTEDD)**
USDOT University Transportation Center
The University of Texas at Arlington
Woolf Hall, Suite 325
Arlington TX 76019 United States
Phone: 817-272-5138 | Email: c-tedd@uta.edu

In cooperation with United States Department of Transportation's Office of Research, Development, and Technology (RD&T)

Acknowledgments

This work was supported by a grant from the Center for Transportation Equity, Decisions, and Dollars (CTEDD) funded by U.S. Department of Transportation Research and Innovative Technology Administration (OST-R) and housed at The University of Texas at Arlington.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The Center for Transportation, Equity, Decisions and Dollars (CTEDD), the U.S. Government and matching sponsor assume no liability for the contents or use thereof.

Technical Report Documentation Page

<table>
<tr><td>1. Report No.
CTEDD 022-04</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Development of A Monitoring System for Driver Readiness in Prolonged Automated Driving</td><td colspan="2">5. Report Date
August 2024</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)
Shuchisnigdha Deb, 0000-0001-5626-2825
Chen Kan, 000-0003-1885-0516
Anurag Pande, 0000-0002-3456-7932
David A. Noyce, 0000-0001-5887-8391</td><td colspan="2">8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Department of Industrial, Manufacturing, and Systems Engineering
The University of Texas at Arlington</td><td colspan="2">10. Work Unit No. (TRAIS)</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
USDOT - 69A3551747134</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Organization Name and Address
Center for Transportation, Equity, Decisions and Dollars (CTEDD)
USDOT University Transportation Center
The University of Texas at Arlington
Woolf Hall, Suite 325
Arlington TX 76019 United States</td><td colspan="2">13. Type of Report and Period Covered
Final Report</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code</td></tr>
<tr><td colspan="4">15. Supplementary Notes
Additional project information is available at https://ctedd.uta.edu/research-projects/development-of-a-monitoring-system-for-driver-readiness-in-prolonged-automated-driving. Project performed under a grant from the U.S. Department of Transportation's University Transportation Centers (UTC) Program.</td></tr>
<tr><td colspan="4">16. Abstract: Fatal traffic crashes have increased significantly, largely due to human error, which automated vehicle technology aims to reduce. However, challenges such as driver fatigue and the need for quick intervention in case of system failures must be urgently addressed to ensure safety. This study aims to develop a driver fatigue monitoring system to detect different driver fatigue levels during prolonged automated driving. The proposed fatigue monitoring system integrates deep learning, computer vision, and machine learning techniques, leveraging postural and behavioral data. A driving simulator experiment was conducted to collect data on eye aspect ratio (EAR), mouth opening ratio (MOR), percentage of eye closure (PERCLOS), and driver postural information. Computer vision techniques were utilized to extract these features from visual data automatically. The study's key finding is that postural data is the most critical factor in detecting driver fatigue. Among the evaluated machine learning algorithms, the random forest algorithm demonstrated the best performance, achieving an accuracy of 0.97 in detecting driver fatigue. Combining postural data with physical measures such as EAR, MOR, and PERCLOS proved highly effective in accurately identifying driver fatigue levels. This integrated approach offers a promising solution for enhancing the safety and reliability of automated driving systems by effectively monitoring and addressing driver fatigue.</td></tr>
<tr><td colspan="2">17. Key Words: Driver monitoring, partial automation, fatigue, driver readiness, takeover request.</td><td colspan="2">18. Distribution Statement
No restrictions. This document is available to the public through the Transport Research International Documentation (TRID) repository.</td></tr>
<tr><td>19. Security Classification (of this report)
Unclassified.</td><td>20. Security Classification (of this page)
Unclassified.</td><td>21. No. of Pages
59</td><td>22. Price
N/A</td></tr>
</table>

Table of Contents

List of Tables

List of Figures

Abstract

Fatal traffic crashes have increased significantly, largely due to human error, which automated vehicle technology aims to reduce. However, challenges such as driver fatigue and the need for quick intervention in case of system failures must be urgently addressed to ensure safety. This study aims to develop a driver fatigue monitoring system to detect different driver fatigue levels during prolonged automated driving. The proposed fatigue monitoring system integrates deep learning, computer vision, and machine learning techniques, leveraging postural and behavioral data. A driving simulator experiment was conducted to collect data on eye aspect ratio (EAR), mouth opening ratio (MOR), percentage of eye closure (PERCLOS), and driver postural information. Computer vision techniques were utilized to extract these features from visual data automatically. The study's key finding is that postural data is the most critical factor in detecting driver fatigue. Among the evaluated machine learning algorithms, the random forest algorithm demonstrated the best performance, achieving an accuracy of 0.97 in detecting driver fatigue. Combining postural data with physical measures such as EAR, MOR, and PERCLOS proved highly effective in accurately identifying driver fatigue levels. This integrated approach offers a promising solution for enhancing the safety and reliability of automated driving systems by effectively monitoring and addressing driver fatigue.

Chapter I: Introduction

Fatal traffic crashes have risen significantly over the last several years. According to the National Center for Statistics and Analysis (NCSA), there were about 6.1 million traffic crashes reported by the police in the United States during the year 2021. These accidents resulted in approximately 42,939 deaths and around 2.5 million injuries (NCSA, 2023). Compared to 2020, these statistics show a 10% rise in fatalities and a 9.4% increase in injuries (NCSA, 2023). Literature and research consistently demonstrate that human error is the primary cause of traffic crashes (Zipper, 2021; Smith, 2013). The National Highway Traffic Safety Administration (NHTSA) believes that vehicle automation technology, which will gradually remove human drivers from vehicle control, will significantly improve roadway safety (NHTSA, 2018).

In this context, the world is witnessing an unprecedented rise in automated vehicle technology. With human error responsible for nine out of ten severe traffic crashes, automated vehicle technologies offer the potential to save thousands of lives, alleviate congestion, enhance mobility, and improve productivity (NHTSA, 2017). However, it is imperative to acknowledge the potential unintended challenges associated with this innovation before its widespread adoption. Despite advancements, automated vehicle technology still faces hurdles in efficiently navigating all possible driving situations, making swift driver intervention necessary in case of system failures (DiMatteo et al., 2020). Currently, in the United States, users can drive under SAE Level 2; however, advancements to SAE Level 3 would improve safety, economy, and society (NHTSA, 2017; NHTSA, 2018). Society of Automotive Engineers (SAE) defines SAE Level 3 or Conditional Driving Automation (CAD) as a vehicle driving itself, while the driver, who can detach himself or herself from driving tasks, must remain available to take over and respond effectively to any situation (NHTSA, 2017; NHTSA, 2018). Because SAE Level 3 requires minimal driver interaction with the system, it is necessary to ensure drivers are not fatigued and have a robust and reliable system that can alert drivers when the vehicle requires the driver to take control.

Driving is a complex task that requires continuous attention and monitoring (*situational awareness)*. Prolonged and monotonous monitoring of automated driving of SAE level 3 frees drivers from the driving task to a great extent and may cause boredom, daydreaming, and drowsiness among drivers, potentially inducing fatigue and compromising their readiness to take control of the vehicle when required. Vehicle automation-based driver fatigue can be both active and passive. Active fatigue is associated with a high cognitive workload. In contrast, passive fatigue develops when there is a requirement for "system monitoring with either rare or even no overt perceptual-motor requirements" (Desmond & Hancock, 2001, p. 601). This research focuses on passive drive fatigue and addresses a critical safety need since collisions involving partially automated vehicles show that responsible human-in-vehicle may become inattentive or drowsy (DiMatteo et al., 2020), enough to significantly delay the takeover or not be able to take over if the driving situation demands it.

In 2021, 8.2% of all fatal traffic crashes involved distracted or inattentive drivers, which is a 12% increase from 2020 crashes (NHTSA, 2023). Additionally, fatalities from drowsy driving accounted for 1.6% of total fatalities, showing an 8.2% increase from the previous year (NHTSA, 2023). Such incidents are more severe than other vehicular crashes as a fatigued driver tends not to perform evasive and preventive measures due to depleted levels of cognition to avoid the crash. Studies have shown that driver fatigue degrades aspects of cognitive functions, including perception, attention, and reaction time (Depestele et al., 2020; Fonseca et al., 2018). The American Automobile Association (AAA) Foundation for Traffic Safety has conducted a study and revealed that each year there are about 328,000 drowsy driving-related crashes (National Safety Council, 2019). Of these incidents, approximately 6,400 crashes are fatal, and 109,000 crashes result in injuries. The researchers indicated that the occurrence of drowsy driving fatalities can be more than 350% greater than reported. Therefore, it is crucial and imperative to focus on fatigued driving-related research and ensure that automated driving for prolonged hours would not deteriorate traffic safety.

This research proposed and researched a driver monitoring system that employs two high-definition webcams to detect and record various driver features such as eye-aspect ratio (EAR), mouth-opening ratio (MOR), head position, hand-to-steering wheel distance, and driver posture data. Understanding these features and their correlation to driver fatigue allows us to predict a driver's fatigue level and send warning signals to alert the driver accordingly.

This report includes two different studies described in the following two chapters. Chapter Two describes fatigue-related research and driving simulator study to explore the correlation of different driver features with fatigue. Chapter Three describes the driving simulator study conducted to identify the most promising warning modality to alert drivers. Chapter Four presents the conclusions and recommendations for future research.

Chapter II: Classification of Driver Fatigue for Prolonged Automated Driving

This chapter presents a summary of the literature review on the safety impacts of driver fatigue, existing technologies in vehicles, methods for fatigue detection, and the driving simulator study conducted to investigate the correlation between driver features like mouth aspect ratio (MAR), eye aspect ratio (EAR), and postural data of the driver (head-to-headrest distance and hand-to-steering wheel distance) and fatigue.

LITERATURE REVIEW

The literature review section discusses existing research on driver fatigue detection, focusing on integrating behavioral and postural data using computer vision. A systematic literature search was conducted using the databases IEEE Xplore, Web of Science, ACM Digital Library, and Science Direct, employing keywords: “driver fatigue detection,” “factors,” “computer vision,” “deep learning,” and "machine learning." Boolean operators “AND” and “OR” have been applied during the literature search process as per requirements. Key factors considered in previous works include applying various postural and behavioral features and the efficacy of computer vision techniques. Additionally, the review examines various machine learning and deep learning models applied in these studies, highlighting their performance and limitations.

Features Considered in Fatigue Detection Models

Recently, researchers have explored various physiological, postural, and behavioral features for developing driver fatigue detection models because of the convenience of extracting these complex data. Physiological features have widely been used for detecting driver fatigue, leveraging signals like Electroencephalography (EEG) and Electrocardiography (ECG). For instance, Huang et al. (2018) grouped features into physiological measures, including EEG, ECG, and electrooculogram (EOG), offering a comprehensive approach to detecting fatigue during driving. However, these methods often require intrusive equipment, potentially causing discomfort in drivers. Similarly, Gwak et al. (2018) and Wang et al. (2018) focused on EEG signals, which, while accurate, are complex to interpret and implement in real-world applications. The physiological approach, though effective, faces practical challenges in terms of user-friendliness and integration of equipment and sensors into everyday driving environments.

On the other hand, postural features, including head and body movements, provide non-intrusive indicators of fatigue. Therefore, several researchers have considered postural data to develop the fatigue detection model. While postural data is promising, ensuring accuracy and minimizing false alarms are still challenging. Savaş & Becerikli (2018) included head movement detection in their studies. However, the individual differences in natural head movements have not been considered, which could result in false positives. Wijaya et al. (2022) also considered nodding and head movements using a comprehensive dataset. Similarly, it remains unclear whether their findings apply to diverse populations. Anber et al. (2022) classified head positions into specific categories, providing a thorough analysis. This can help to cover the variability of head positions in populations. In another study, Mahmoodi & Nahvi (2019) used surface electromyography for postural data. Nevertheless, the practicality of considering this feature in real-world settings is questionable due to the need for muscle sensors.

Facial features, such as eye and mouth movements, have become prominent in fatigue detection research. Studies by Savaş & Becerikli (2018) measured eye closure duration and yawning frequency with other features to model the driver fatigue detection system. Gu et al. (2018) and Zhao et al. (2018) also considered the eye and mouth states for similar types of research. However, these features can be influenced by various environmental conditions and individual differences in facial expressions. The inclusion of slow blink rates and yawning in the National Tsing Hua University dataset by Wijaya et al. (2022) has added robustness to the model development. In this study, the drivers were asked to drive in the morning and at night, considering the environmental conditions. Facial features are intuitive and less intrusive, yet their effectiveness heavily depends on the reliability of the underlying image-processing algorithms.

Computer Vision Techniques in Fatigue Detection

Since the driver fatigue detection model is effectiveness depends on the reliability of image processing, the application of computer vision for data extraction has gained significant attention from researchers. Numerous studies have employed a variety of methodologies to monitor facial expressions and behaviors indicative of fatigue (Sikander and Anwar, 2018). Various computer vision techniques have used experimental setups, such as real and simulated driving environments. Despite the inherent challenges, researchers have aimed to capture realistic data in real driving environments. Du et al. (2022) and Huang et al. (2018) utilized camera setups to monitor facial expressions in actual driving conditions, including subjects wearing and not wearing glasses and using mobile phones. These studies provide valuable insights but face challenges, such as varying lighting conditions and the need to maintain driver comfort and safety. Zhao et al. (2018) and Fatima et al. (2020) emphasized non-intrusive camera placements to minimize driver distraction while capturing facial expressions. While these real-world studies are crucial for validating the effectiveness of fatigue detection methods, developing universally applicable systems is challenging because of the variability in environmental conditions, the inability to investigate all types of traffic conditions, and individual driver behaviors.

Simulated driving environments offer a controlled setting for data collection, allowing researchers to systematically study driver fatigue without the risks associated with real-world driving. Savaş & Becerikli (2018) employed facial expression detection software to monitor eye closure and yawning frequency in driving simulators. Similarly, Wijaya et al. (2018) and Guo et al. (2019) used infrared cameras to capture high-resolution facial expressions even in low-light conditions. Studies by Gwak et al. (2018) and de Naurois et al. (2018) utilized driving simulators with various environmental settings to create a realistic driving experience. While simulated environments offer many benefits, the controlled nature of these simulations may not fully replicate the complexities and unpredictability of real-world driving. Therefore, Ji et al. (2019) and Rajkar et al. (2022), in their CEW and Yawn DD datasets, used data from real-world driving, video samples, and simulation-based driving to enhance the robustness of detection algorithms.

Machine Learning and Deep Learning Algorithms for Fatigue Detection

Machine Learning (ML) and Deep Learning (DL) based implementations for fatigue classification are data-intensive and use advanced modeling techniques like Neural Networks (NN), Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Support Vector Machines (SVM), and the ensemble methods such as Random Forest (RF), Decision Tree (DT) and Gradient Boosted Regression Tree (GBRT). For the past two decades, the applications of various ML and DL algorithms have significantly contributed to driver fatigue detection systems. Savaş and Becerikli (2018) utilized SVM with OpenCV and Dlib libraries, achieving a notable 97.93% accuracy in detecting driver fatigue by analyzing features like eye closure percentage, yawning frequency, and head detection. This study highlighted the effectiveness of SVM in handling detection models, demonstrating its potential for real-time applications. Similarly, Gwak et al. (2018) explored multiple ML algorithms, including logistic regression, SVM, k-nearest neighbor (KNN), and RF, with RF achieving the highest accuracy of 81.4%. Zandi et al. (2019) employed both RF and non-linear SVM for the binary classification of vigilance states, with RF achieving accuracy rates between 88.37% and 91.18%, outperforming SVM. Wang et al. (2018) compared multiple models, including KNN, SVM, GBRT, and DT, with the GBDT achieving the highest accuracy of 94.3%. These ML models demonstrate the potential of non-intrusive, behavior-based fatigue detection systems. However, despite their promising results, ML models often face limitations in handling complex and high-dimensional data, hindering their robustness and adaptability in real-world scenarios. ML models typically rely on manually generated features, which may not fully capture the intricate patterns associated with driver fatigue, leading to potential inaccuracies in diverse driving conditions.

In many cases, DL algorithms have shown exceptional performance in feature extraction for driver fatigue detection due to their ability to learn complex features from large datasets. Gu et al. (2018) developed a CNN model with multi-scale pooling (MSP-Net), achieving high accuracies of 98.05% and 98.85% on different datasets, demonstrating the model's ability to generalize across various conditions. Similarly, Huang et al. (2018) introduced the P-Fatigue Detection Convolutional Network (P-FDCN), which demonstrated robust eye state recognition with accuracies of 94.9% and 95.1%. Wijaya et al. (2022) developed a CNN-based model that achieved an accuracy of 70% for validation and 56% for testing, indicating areas for improvement in robustness and accuracy. Another significant contribution is by Xiao et al. (2019), who proposed a convolutional recurrent network architecture combining CNN and Long Short-Term Memory (LSTM) units, achieving a 95.83% accuracy by learning spatial representations and temporal dynamics of eye features. Ji et al. (2019) introduced a Mouth State Recognition Network, achieving 98.42% detection accuracy on a public eye dataset and 97.93% on an open mouth dataset. Zhao et al. (2018) proposed a framework using facial dynamic fusion information and a deep belief network (DBN), achieving 96.7% accuracy. These DL models, particularly CNNs, effectively captured spatial features from facial images, making them highly suitable for real-time fatigue detection applications. However, the black-box nature of DL models can pose challenges in understanding and interpreting their decision-making processes, which is crucial for ensuring the reliability and safety of fatigue detection systems.

Hybrid models combining multiple ML and DL techniques have also been explored to enhance detection accuracy and robustness. Guo et al. (2019) introduced a hybrid approach using convolutional neural networks and long short-term memory (LSTM) algorithms, achieving an average accuracy of 84.85%. This hybrid approach used CNNs for spatial feature extraction and LSTMs for temporal sequence learning, effectively capturing both spatial and temporal aspects of driver behavior. Another hybrid model by Gao et al. (2019) proposed an EEG-based spatial-temporal convolutional neural network (ESTCNN), which integrated temporal information processing with dense layers to fuse spatial features, resulting in a 97.37% accuracy. Rajkar et al. (2022) utilized CNN models with OpenCV's Haar cascade algorithm, achieving 96% accuracy in real-time detection. The study by Hu and Min (2018) employed an ensemble learning method using gradient-boosting decision trees (GBDT), achieving a 94.0% recognition rate with EEG signals. These hybrid approaches demonstrate the potential to overcome the limitations of standalone ML or DL models by combining their strengths.

Several innovative models have been proposed to address the challenges of real-time driver fatigue detection, focusing on novel architectures and feature extraction methods. Képešiová et al. (2020) utilized a two-stage approach combining artificial ANNs and CNNs, achieving 98.02% validation accuracy. This method demonstrated the potential of staged processing to refine feature extraction and improve detection accuracy. A more recent study by Sun et al. (2023) introduced a three-stream FFF-CNN model for robust fatigue detection on low-quality inputs, achieving 98.35% accuracy. This model employed a global face stream and two local eye streams, incorporating feature fusion modules to enhance stability and reduce noise. Anber et al. (2022)

proposed a non-invasive approach using AlexNet CNN, achieving an impressive 99.65% accuracy by combining features from head position and mouth movements. Furthermore, You et al. (2019) introduced a deep cascaded convolutional neural network (DCCNN) with an eye assessment parameter (EAR), achieving a 94.8% accuracy. Fatima et al. (2020) used SVM and Adaboost for eye-state classification, achieving 96.5% and 95.4% accuracy, respectively. Deng and Wu (2019) proposed the MCNN-KCF and multitask convolutional neural networks (MTCNN), achieving an average of 92% accuracy. The study by He et al. (2020) introduced a two-stage convolutional neural network, achieving 93.83% classification accuracy and 94.7% on Raspberry Pi 4. Table 1 represents a summary of some previous significant works on driver fatigue detection systems. These models highlight the continuous advancements in fatigue detection technologies, aiming for higher accuracy, robustness, and practical applicability in real-world scenarios. However, despite their high accuracy, these models often require extensive pre-processing and may struggle with varying lighting conditions and occlusions, impacting their real-world effectiveness.

Research Gaps

Despite significant advancements in driver fatigue detection using ML and DL models, several research gaps still need to be addressed. One major gap is the generalizability of these models across diverse driving environments and conditions. Factors such as varying lighting conditions, different vehicle types, and diverse driver demographics can significantly impact the accuracy and reliability of fatigue detection systems. Another notable research gap is integrating multimodal data and developing comprehensive systems that combine behavioral and postural information. While individual models using PERCLOS, EAR, MAR, and postural data have shown promise, their standalone applications may not provide a holistic view of the driver's state. Therefore, current research has considered integrating behavioral and postural features under varying driving conditions to develop the driver fatigue detection system.

Table 1: Features and accuracy in driver fatigue detection and classification models

Research Group	*Type of Model*	*Features Used*	*Accuracy*
Deng et al. (2019)	CNN	Duration of blinking, Duration of eye closure, and yawning	92%
Hu et al. (2018)	Gradient boosting decision tree model	Single Channel EEG signals	94%
Gu et al. (2018)	Hierarchical CNN	PERCLOS, FOM	98.02% (Precision)
Alparslan et al. (2020)	CNN	Eye-closedness	94%
Li et al. (2017)	RNN	Entropy in steering wheel angle	83.25%
Chellapa et al. (2016)		Pulse rate, yawning, body temperature, drooping eyelids	80.55%
You et al. (2019)	DCCNN and SVM	EAR	94.80%
Zhou et al. (2021)	XGBoost	Average heart rate, Average breathing rate, heart rate variability	0.996 (R^2 - value)
Zhang et al. (2017)	CNN	PERCLOS, blink frequency	91.45%
Savaş and Becerikli (2018)	SVM	PERCLOS, # of yawns, mouth opening, count of eye blinking, and head detection	97.93%
Li et al. (2021)	System input-based classification	The grip on the steering wheel	86.6%
Lu et al. (2021)	XGBoost	EMG signals	90%
Jia et al., (2018)	CSI Variations	WiFi signals	89.6% (Single Driver Case)
Wang et al. (2018)	Wearable device	Dry EEG signals	-
Boon-Leng et al. (2015)	SVM	EMG, Galvanic skin response	90% (Precision)

METHODS

This study aims to design and evaluate an accurate and robust classification model to detect efficient driver fatigue, can be implemented at scale with ease, and is non-intrusive. This research study underwent rigorous review and received approval from the Institutional Review Board's Ethics Committee at The University of Texas at Arlington.

Participants

Twenty licensed drivers, each with at least one year of driving experience, were selected as participants. To recruit this diverse group, the research team utilized flyers and targeted emails directed at students from the University of Texas at Arlington (UTA). A detailed screening questionnaire was employed to evaluate potential participants on several criteria, including their driving behavior, attitudes toward automated technology, and demographic background. The exclusion criteria were carefully applied to ensure the integrity of the research data. Individuals who used eyeglasses or corrective lenses for driving, those with pacemakers, anyone with less than one year of driving experience, or those susceptible to simulation or motion sickness were not considered for participation. The selection process prioritized achieving a diverse group in terms of gender, age, and ethnicity, aiming to enhance the generalizability of the study's outcomes.

Apparatus

Figure 1: Settings of the simulator study

The study utilized a three-degree-of-freedom (3DOF) motion-base RDS-1000 single-seat sedan driving simulator from Realtime Technologies. The system is also outfitted with SAE level 3 automated driving capabilities to mimic an actual vehicle's control experience closely. The visual environment for the simulator was projected onto three 65-inch TV screens, each display features a resolution of 1920 × 1080 pixels, providing a comprehensive field of view of 205° horizontally and 38° vertically. The virtual environments and scenarios were designed using the Internet Scene Assembler (ISA), a sophisticated Virtual Reality Modeling Language (VRML) authoring tool. Vehicle movements within these environments were managed by SimCreatorDX, ensuring a realistic driving experience. The entire simulator setup was housed in a room designed for quietness and controlled lighting conditions to minimize external distractions and influences.

Advanced vision-based features were analyzed, including PERCLOS (a measure of eyelid closure over time), the mouth-opening ratio (MAR), and head position, utilizing the GoPro HERO10 camera positioned to face the driver. In addition, the Logitech 720p HD webcam, placed behind the driver at an inclined angle and an approximate height of 8 feet, was used to capture detailed postural data. This setup facilitated the collection of critical postural information, including the distance of the hands from the steering wheel and the distance of the head from the headrest, enriching the behavioral analysis of the driving simulation study. The setting of the simulator study is shown in Figure 1.

Driving Scenarios

During the experiment, participants were exposed to a carefully designed 45-minute-long driving scenario that closely replicated the condition of the Interstate-30 highway near Dallas, Texas (see Figure 2). The scenario was programmed to maintain simulated vehicles' speed at 70 mph, modeling a driving situation with clear lanes ahead and ongoing traffic in the next lane and in the opposite direction. Each scenario had a straight four-lane highway with divided traffic lanes, having a constant road geometry and lane width. The virtual driving environment was set for daytime driving with adequate visibility settings. There were no weather-related disturbances in the virtual environment, such as gusty winds, fog, or rain. The settings incorporated various highway features, such as multiple exits and entrance ramps, overhead bridges, bridges with curvature, and commercial and high-rise buildings, to enrich the driving experience and realism of the simulation. In order to avoid simulation sickness in the participants, the entire drive was broken into three 15-minute-long sessions with two 5-minute breaks in the middle.

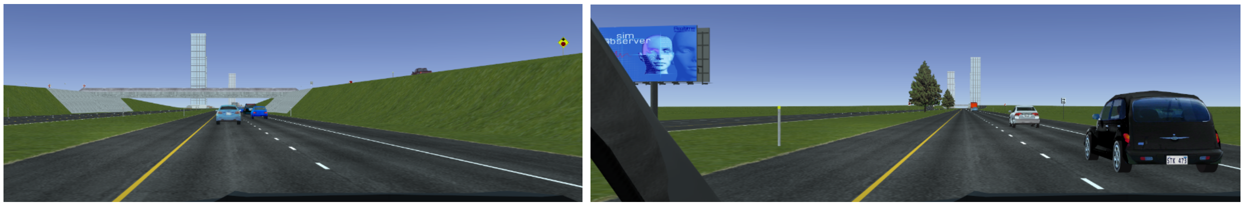

Figure 2: Simulated driving scenario

From the onset of each scenario, participants were instructed to engage the simulator's automated driving mode within the first ten seconds. This action activated the vehicle’s automatic lane changing and speed control features, providing comprehensive support for lateral and longitudinal vehicle maneuvering without manual driver intervention.

Study Protocol

After the arrival of each participant, their body temperature (safety measure for COVID-19) and driver’s license were checked to ensure participation eligibility. Then, each participant read and signed the consent form. Before performing a 5-minute test drive, participants were asked to complete a simulation sickness questionnaire (SSQ, Kennedy et al., 1993) to monitor their well-being and determine simulation sickness. The scale has 16 items, each of which can be rated from 0 to 4. In the actual drive, participants engaged in the driving simulation lasting 45 minutes, divided into three 15-minute sessions with interposed breaks. During these breaks, participants were asked to complete SSQ again to identify any developing sickness for their exposure to the simulation. Each time the participants took SSQ, a cumulative score of 5 or more indicated simulation sickness, prompting the immediate ending of the participant’s involvement in the experiment. However, no participants were withdrawn from the study due to simulation sickness. Figure 3 shows the flow diagram of the study protocol. During all of these sessions, two researchers were present in the lab. One handled the apparatus and data collection process, while the other communicated with the participants. The participants were compensated $30 for their participation.

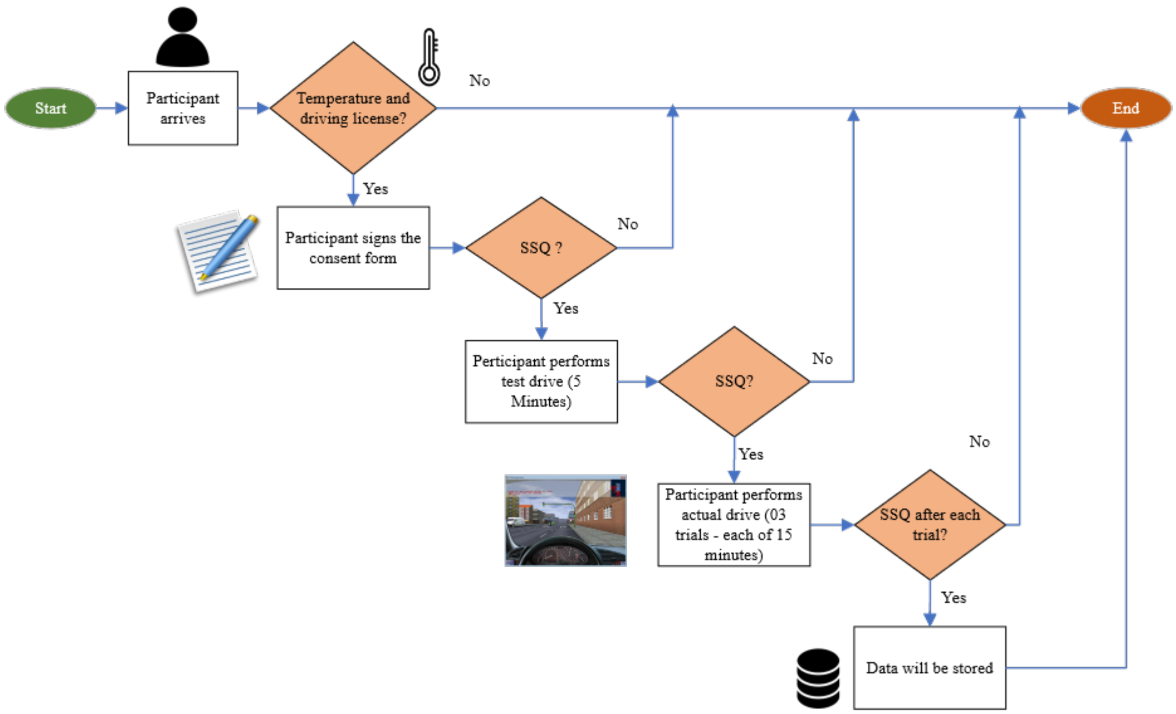

Figure 3: Flow diagram of the study protocol

Data Set Description

This study investigates the development of a comprehensive system for driver state monitoring using visual cues. Although various studies have explored different types of input variables for modeling driver fatigue monitoring systems, the current study considered eye-aspect ratio (EAR) and mouth-aspect ratio (MAR) as the behavioral input factors, and head-pose estimation, head-to-headrest distance, hand-to-steering wheel distance, deviation of head-pose estimation as the postural input factors to detect and classify varying levels of driver fatigue. During the study, video cameras were used to collect the recording of the input factors.

Figure 4: Feature extraction using computer vision

The EAR, a scalar value representing the ratio of eye width to eye height, can effectively predict a driver's eye states, serving as a critical measure for assessing alertness and safety on the road. Additionally, Ling et al. (2021) confirmed EAR's efficacy in evaluating a driver's eye state with impressive accuracy in complex driving environments. When the eye is open, the value is almost constant, but it approaches zero when the eye is closed (Zhu et al., 2022). The MAR is another key indicator for determining if a driver is yawning, an important sign of fatigue, by measuring the scalar quantity of the mouth's dimensions (Martinez & Huang, 2022). Head-pose estimation is used to determine the driver’s gaze direction and can depict the driver's distraction and engagement in any non-driving related tasks (NDRTs). This estimation and its deviation can be decomposed into three variables viz, yaw (y-axis), pitch (x-axis), and roll (z-axis). Pitch indicates the angles when the driver is moving the face in an up-and-down direction. Yaw indicates the angle when the driver is turning face left or right, and roll indicates the angle when the face is tilting (see Figure 4). Head-to-headrest distance is defined as the Euclidean distance between the driver’s head and the headrest. Hand-to-steering wheel distance is defined as the Euclidean between the driver’s hand and the steering wheel. The deviation of head-pose estimation is to determine if the driver's gaze is focused on the road. The deviation is calculated as the difference between the average value of the head-pose estimation variable during the first 150 frames of the experiment and the head-pose estimation for each frame.

The response variables for the experiment are class labels: 0,1,2,3, depicting the severity of the fatigue level experienced by the driver during the experiment. The response values are manually assigned for each data point using the results of the Karolinska Sleep Scale (KSS), a 10-point scale where 1 indicates extreme alertness and 10 represents extreme sleepiness to the point of struggling to stay awake. Each researcher assigned an individual rating for each block of video data (window labeling). A rating of 1, 2, and 3 on the KSS indicated that the driver was in an alert state, and a fatigue class label of 0 was assigned for all the data points. A rating of 4 on the KSS implied the driver was rather alert, and a fatigue class label of 1 was assigned, respectively. Ratings 5,6,7 were given on the KSS if the driver exhibited signs of fatigue, such as yawning or non-attentive driving postures, and a fatigue class label of 2 was assigned for these KSS ratings. If the driver showed signs of severe fatigue, such as frequent yawning, drooping eyes, and visible struggle to stay awake, a rating of 8, 9, or 10 was given on the KSS depending upon the severity of the fatigue signs and consequently received a fatigue class label of 3. The final class label for each data point was calculated using the rule of majority voting. The response variable levels are as follows: 0 – Alert, 1- Alert, but showing signs of fatigue, 2- Fatigued, but making attempts to stay awake, and 3 – Fatigued, should not drive.

RESULTS

Data Preprocessing and Extraction

For each run in the experiment, the video data was recorded at a rate of 30 frames per second (fps), resulting in 81000 data points for each experimental run. Various deep-learning techniques were implemented on the recorded video data to extract the dependent variable data. Dlib, a modern C++ library, was used for image processing. The pre-trained facial landmark detector of dlib was used to extract x-y coordinates of 68 key facial landmarks in the facial region. The get_frontal_face_detector() function in dlib, which uses a Histogram of Oriented Gradients (HOG) alongside a Linear Support Vector Machine classification function, was utilized to detect faces in the video data.

The '*shape_predictor()*' function in dlib was used to extract key points of facial landmarks. This function identified six landmark points for each eye and mouth, using 68 face landmarks. These points were consequently used to calculate the EAR and MAR values. This approach had the advantage of being computationally inexpensive. Wide range head pose estimation network (WHEnet), a novel end-to-end head-pose estimation network that excels in wide range estimation, was used to extract and calculate the head-pose estimation data. WHENet uses YOLOv4 to crop images of subjects' heads and compute accurate head pose Euler angles relative to the camera. A custom-trained YOLOv7 object detection algorithm was used to extract postural feature data. We collected images with bounding box annotations for the markers on the steering wheel, hands, and headrest from multiple runs of lab researchers and other participants. YOLOv7 was preferred for its fast and efficient object detection performance on edge devices. It is an anchor-based model that predicts potential bounding boxes and applies non-maximal suppression to get the best fit. The model was fine-tuned on the dataset for 100 epochs, with an input dimension of 640x640 pixels. A learning rate of 0.01, a momentum of 0.937, and a weight decay of 0.0005 to optimize the learning process, speed up convergence, and prevent overfitting. The model was initialized with pre-trained weights from Yolo. The algorithm was applied to segment the hands, two key positions on the steering wheel, and the headrest. The Euclidean distances were calculated from these segmented positions' center points and stored in a dictionary with the class and confidence score. This method allowed us to collect and organize important feature data systematically.

Each participant’s video data was divided into multiple blocks of 3-minute length video data (windowing). Windowing ensures that analyses are consistent across the participants. Several research studies have implemented a machine-learning approach to model similar types of data (Lever et al., 2016; Mohsenzadeh et al., 2020). The collected data was thoroughly examined for duplicates and missing values. Python and R programming languages were utilized to conduct the analysis. Several instances of duplicate data values were found. This can be attributed to the fact that the video data is collected at a rate of 30 frames per second (FPS), translating to 30 data points for each feature vector per second. Since the driver's position might remain constant for a few seconds, duplicate values are generated. The duplicate values were identified and removed, ensuring the data entries' uniqueness for subsequent analysis stages. The preprocessing analysis also revealed that the collected data had severe class imbalances, where certain class levels were significantly under-represented in the data. The performance of standard classification algorithms is often poor when learning from imbalanced data (Li et al., 2013). The class imbalance can potentially induce bias in the statistical analysis of the data, necessitating the consideration of resampling strategies. There is abundant literature discussing various techniques to tackle the issue of class imbalance. The current study implemented the random under-sampling (RUS) algorithm to tackle this issue. The random under-sampling algorithm was chosen over other oversampling techniques as it would decrease the overall size of the training dataset and would be computationally beneficial. However, the RUS algorithm has the obvious disadvantage of throwing out points from the dataset, which could contain useful information.

However, the huge amounts of data collected in this study allowed us to have a sufficiently sized dataset after implementing the RUS algorithm, with each class having around 40,000 instances, providing us with enough data for the statistical model to learn. The data was also shuffled after implementing the RUS algorithm to randomize the data well to mitigate any potential sequential order effects and to avoid any unwanted bias and dependencies in the statistical learning models. Figure 5 highlights the stark class imbalance before implementing the RUS algorithm.

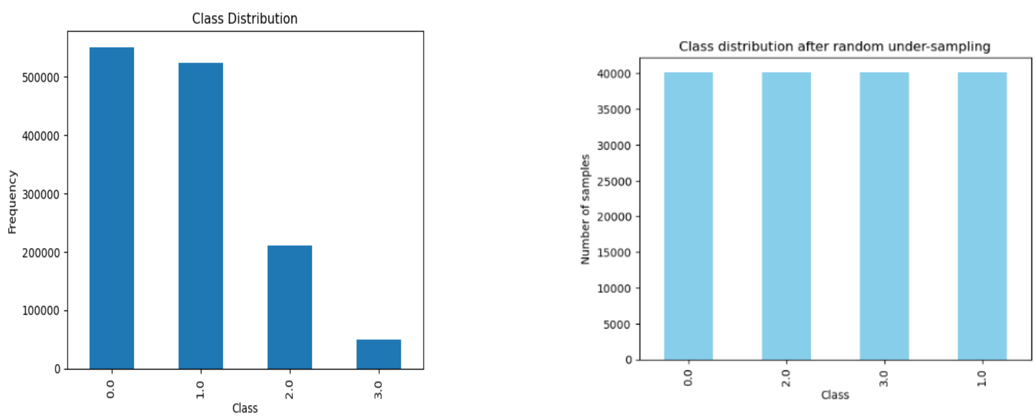

Figure 5: Class distribution

Algorithm Classification Performance

The data was randomly split into training (70%) and testing (30%) sets. A 10-fold cross-validation approach was also employed as a control to prevent the models from over-fitting the training dataset (Abbas & Alsheddy, 2020). This method enhances the reliability of our findings. Several classification algorithms were compared, such as Decision Trees, Random Forest, Support Vector Machines, Naïve Bayes Classifier, and k-nearest Neighbors. Hyper-parameter tuning for suitable models was also employed to achieve the maximum possible accuracy for each algorithm. The algorithms were compared across their accuracy and confusion matrices for their predictions on the testing data set. Figure 6 illustrates the accuracy of the comparison of the various machine-learning approaches. The random forest classification algorithm outperformed the other classification algorithms and had an overall accuracy of 0.97. The worst-performing algorithm was the Naïve Bayes classifier, with an overall accuracy rate of 0.474. Variable importance calculations were performed post-algorithm analysis to extract additional insights from the best-performing algorithm.

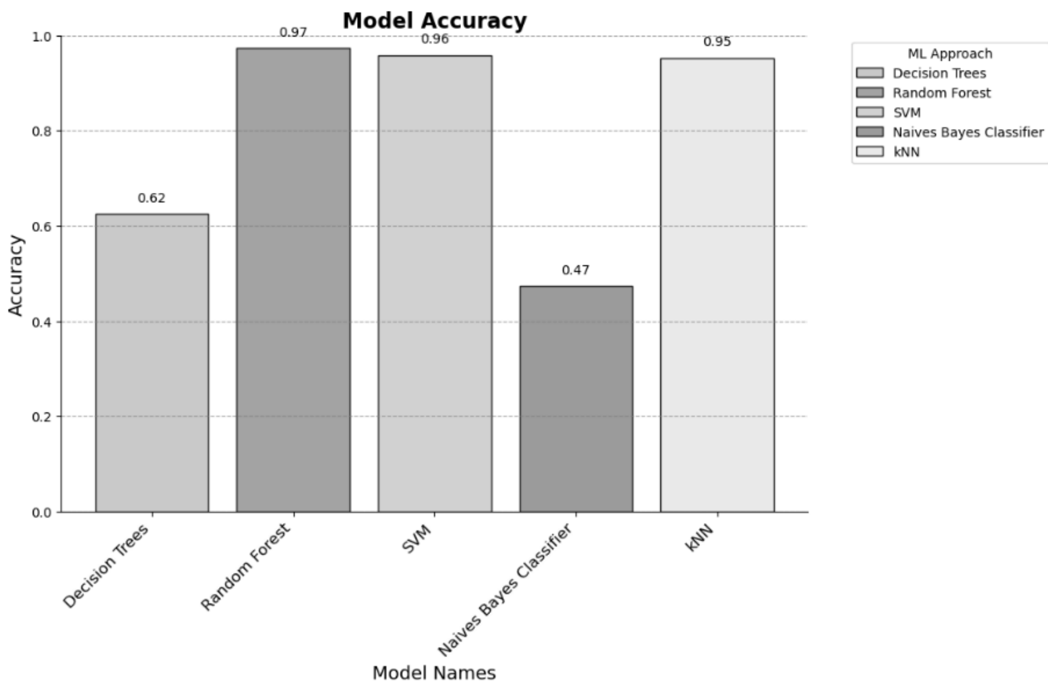

Figure 6: Accuracy comparison of classification algorithms

Although the Support Vector Machine (SVM) and Random Forest (RF) algorithms demonstrated similar accuracies, we chose Random Forest for this study due to its lower computational demands. Non-linear SVMs are known to be computationally intensive and less practical for larger datasets. In contrast, Random Forest offers a more efficient and scalable solution, making them more suitable for our analysis (Alshaqaqi et al., 2013). Table 2 presents the hyper-parameter optimization values for various algorithms used in the study. For the Random Forest algorithm, the number of trees and the number of variables randomly sampled as candidates at each split (mtry) were optimized. Similarly, for the Support Vector Machine (SVM), the kernel type, regularization parameter (C), and gamma values were fine-tuned. These optimizations were critical for achieving the best performance for each model, ensuring that each algorithm was configured to handle the dataset's specific characteristics effectively.

Table 2: Features and accuracy in driver fatigue detection and classification models

Model	Validation Approach	Hyperparameter values/ Algorithm Settings	Model Accuracy
Decision Trees	10-Fold Cross Validation	Complexity parameter (cp) = .01	0.64
Random Forest	10-Fold Cross Validation	number of trees (ntree)	0.97
Support Vector Machine	10-Fold Cross Validation	Radial basis function(kernel), Regularization parameter (C) = 7.130, Kernel Coefficient (gamma) = 0.0980	0.89
Naïve Bayes Classifier	10-Fold Cross Validation	Variable Smoothing	0.47
k-Nearest Neighbor	10-Fold Cross Validation	Metric = Manhattan Distance, # neighbors = 3, weights = distance	0.95

DISCUSSIONS

Fatigued driving remains one of the largest safety risks in ground transportation systems. Although previous research studies have advanced fatigue monitoring and driver state detection systems, few attempt multiclass classification (Ansari et al., 2022), and even fewer attempts to compare the performance of various classification algorithms to detect driver state and fatigue (Alshaqaqi et al, 2013). The results of this research study can help identify the important features of driver fatigue modeling algorithm development and illustrate the importance of advanced statistical modeling techniques for inferring driver fatigue states. The purpose of the current research study was to develop real-time fatigue monitoring system that does not implement the use of intrusive equipment for data acquisition and can be designed to distinguish multiple severities of fatigue.

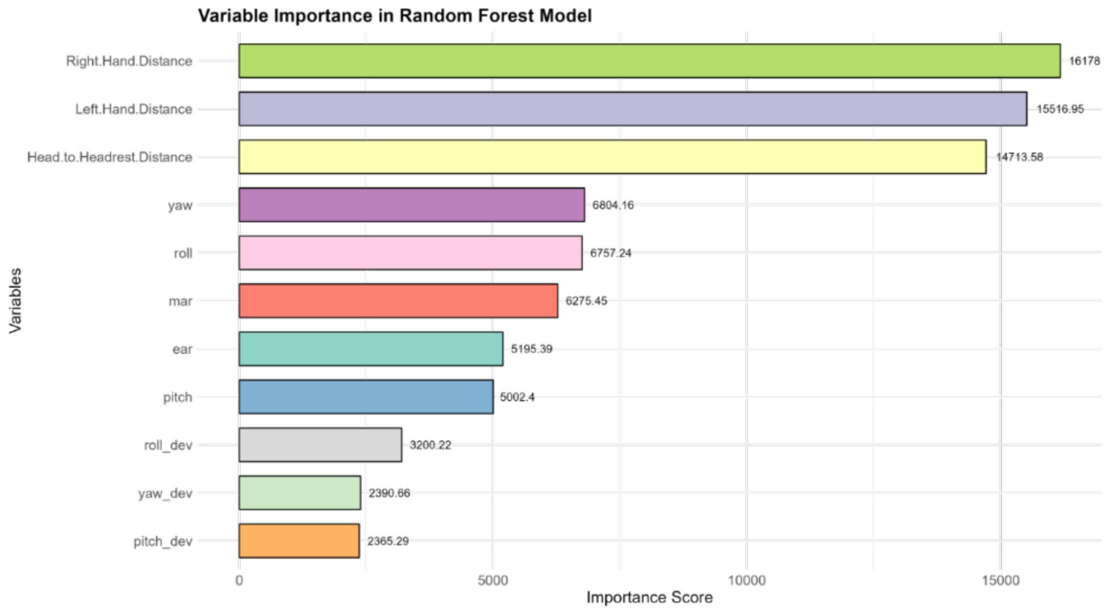

Figure 7: Variable importance plot for RF model

The findings provide significant insights into detecting driver fatigue through a hybrid model integrating computer vision, deep learning, and machine learning algorithms. The results have demonstrated that combining postural data with behavioral data can effectively identify the various driver fatigue levels. Savaş & Becerikli (2018) also found the fatigue detection system effective by integrating both types of features. The study has also compared various machine learning classification algorithms and revealed that the RF model outperformed others, achieving an accuracy of 97%. The variable importance plot in Figure 7 illustrates the significance of various features in predicting driver fatigue using an RF model. The top three variables—right-hand Hand Distance, Left left-hand distance, and hand-to-head distance—are the most influential, suggesting that the distances between the driver's hands and the steering wheel are critical factors. Ansari et al. (2022) also identified that the position of the hands-on steering wheel is one of the key factors in determining driver cognitive fatigue. Other significant features include the yaw, roll, and pitch angles, along with their standard deviations, indicating that head movements and their variability also play crucial roles. This analysis highlights the importance of both static postural data and dynamic movement data in effectively predicting driver fatigue, which is vital for designing effective driver monitoring systems in the field of human factors.

The accuracy and results of the various algorithms show that the driver’s postural behavior data combined with features like PERCLOS, MAR, and EAR are effective for differentiating and detecting different levels of driver fatigue. Alshaqaqi et al. (2013) found that the drowsy states of drivers can be effectively identified using these features related to the eyes and mouth. The current study's findings add to a body of evidence from research that seeks to infer driver fatigue states by combining and comparing various statistical models and physical features like postural data, MARS, and PERCLOS. The algorithms and models in the current research study are designed to detect and distinguish four driver fatigue states non-intrusively. Most prior work focuses on detecting driver fatigue using intrusive data acquisition techniques and majorly detected only two states of fatigue - “high fatigue” vs. “low/no fatigue” (Savas & Becerikli, 2018).

Chapter III: Comparing Effects of Environments and Warnings on Driver’s Time to Takeover

This chapter describes the driving simulator experiment conducted to study the effect of different types of fatigue on drivers' takeover performance. Therefore, our experiment includes two opposing environments: high workload-induced fatigue and low workload-induced fatigue. The experiment requires the driver to continuously monitor and wait for a take-over request (TOR) to react and take over the driving from the automated system.

LITERATURE REVIEW

Fatigue, Automated Driving, and Takeover Performance

There have been multiple studies relating fatigue and automated driving. Korber et al. (2015) found a causal relationship between automation and passive fatigue from the participants’ Dundee Stress State Questionnaire (DSSQ). The effects of non-driving related tasks (NDRT) on passive fatigue were discussed and evaluated in several studies; however, there are conflicting results in the literature. Jarosh et al. (2017) suggested no significant difference in takeover performance between a simple monitoring task and a quiz task. On the contrary, Schomig et al. (2015) concluded that drivers who do not participate in NDRTs attain the highest levels of drowsiness. Because the focus of this study is to assess fatigue levels induced by the environment only, our experiment will not include NDRTs.

The environment, an external factor to the vehicle, poses some unforeseen challenges that need further study. Saxby et al. (2013) found that active fatigue reduced the takeover performance of the driver when compared to passive fatigue. Their study simulated active fatigue by exposing the drivers to frequent wind gusts, which required them to frequently correct the steering wheel to counter the wind and remain within the highway lane. They induced passive fatigue by letting the driver sit in fully autonomous mode without interfering with the vehicle (Saxby et al., 2013).

Feldhütter et al. (2019) conducted a study to determine if long periods of monotonous states increase a driver's fatigue level and if that affects take-over performance. The automated driving lasted 60 minutes, and a takeover request (TOR) was issued. One group was allowed to participate in NDRTs (alert group) while the other monitored the environment (fatigue group). The study found that individual drivers experience different fatigue levels on similar periods of long, automated driving. Although results show no significant changes in the take-over time between the alert and fatigue groups, the authors suggest future work where the take-over request is given once a driver reaches certain fatigue levels (Feldhütter et al., 2019).

Various authors have developed methods to measure fatigue levels during automated driving. Niu & Ma (2022) developed a 3-level fatigue system to determine the driver's fatigue status. Once the system detected the fatigue level, the system sent an auditory signal according to the NHTSA guidelines. The beep notifications were classified into low-, medium-, and high-frequencies (Niu & Ma, 2022). Another study was conducted using eye blink duration and the Karolinska Sleepiness Scale (KSS) to determine whether the drowsiness levels of the driver significantly increased after 30 minutes of driving. Eyeblink duration and KSS proved highly correlated and, thus, a useful system to detect driver drowsiness (Wu et al., 2019). Another study showed that a head-mounted eye tracker would effectively detect eye-blinking activity resulting from automated driving (Schmidt, 2017). Our fatigue detection system used four driver fatigue levels, which were detected based on the eye aspect ratio, mouth-opening ratio, gaze direction, and hands on the steering wheel.

Tavakoli et al. (2021) investigated the environmental factors affecting driver fatigue, such as weather, traffic density, noise levels, road type, and passengers. They concluded that the average driver is more distracted in the city than on the highway. The highway induces monotony as there is less movement in the environment. Another study found no significant changes in take-over performance in an environment involving a two-car crash vs. no crash ahead of the driver (Bourrelly et al., 2019). However, results show a significant influence of high traffic density on the take-over time and take-over quality in a highway setting. A more complex traffic scenario showed a higher takeover time (Radlmayr et al., 2014). Tavakoli et al. (2021) also found that gaze direction and head movement can change between clear or cloudy weather. The data shows that clear weather has higher gaze and head movement, so the driver is more alert. During a different study, takeover time (TOT) was slightly longer in the dark than in the daytime. A dark environment may evoke the circadian rhythm, leading to higher fatigue levels than in the daytime. There was no significant difference in takeover quality between daytime and dark environments (Shi & Bengler, 2022). Self-reported surveys indicated driver fatigue increased when the environment involved rain (Yu et al., 2016). Rain also increased cognitive workload and affected the takeover time during another study (Li et al., 2018). Both studies suggest that higher fatigue levels are, among other reasons, due to reduced visibility of mental workload when assessing wet conditions.

TOR Warning System Design

During SAE Level 3 vehicle operation, the driver must be prepared to take over when the system sends a signal, regardless of the fatigue level. Since automation leads to passive fatigue, developing a reliable Takeover Request (TOR) signal is necessary to trigger the driver to take over effectively (Korber et al., 2015). Several research papers on multimodal takeover requests have been studied (Yun et al., 2020; Huang et al., 2019) and determined a convenient model for the TOR signal. Multimodal warning appeared to be more effective. Several studies have evaluated all combinations of Visual (V), Auditory (A), and Tactile (T) warnings and have found that redundant multimodal signals decrease response time (Politis et al., 2016; Petermeijer et al., 2017).

Table 3: Summary of literature review for experimental design investigating warning type for take-over performance (V-Visual, A-Audible, H-Haptic)

Paper	Population	Type	Independent	Dependent	Warning Type
(Niu & Ma, 2022)	China Male: 24 Female: 6 Age Mean: 25.5 yrs	Car Racing Simulator	Fatigue Levels (3 levels) Warning Signals (Beep)	**Objective Measures** Gazing Behavior (eye gazing at road per minute) Braking Behavior (mean brake) Average Speed (mean speed) **Subjective Measures:** Overall Workload (NASA-TLX) Trust (Questionnaire)	A
(Wu et al., 2019)	Japan Male: 60 Female: 55 Age Mean: 44.6 yrs	AIST Simulator	Age Experimental Condition (Auto-31 or Auto-Manual-Auto)	**Objective Measures** Time-Steer (turning of steering wheel) Time-Brake (brake pedal is pressed) Reaction Time (min Time-Steer or Time-Brake) steer (Manoeuvrings smoothness) TTC (time to collision) **Subjective Measures:** Karolinska Sleepiness Scale (KSS) Eyeblink Duration	A & V (multimode)
(Feldhütter et al., 2019)	Germany Male: 27 Female: 15 Age Mean: 46.0 yrs	BMW Simulator	Condition (Natural Load Condition vs Underload Condition) Time Intervals	**Objective Measures** PERCLOS (%) Blink Frequency (Count) Take Over Time (s) Maximum Longitudinal Accelerations (m/s^2) Maximum Lateral Accelerations (m/s^2) Minimal time-to-collisions (s) Securing Behavior (%)	A (Double Beep)
(Korber et al., 2015)	Germany Male: 18 Female: 2 Age mean: 23.3 yrs	Static Driving Simulator	Vigilance	**Objective Measures** Vigilance task (Reaction Time) Eye Tracking (Blink frequency, Blink duration, PERCLOS) Mind Wandering (DSSQ)	A
(Yun et al., 2020)	Korea Male: 25 Female: 16 Age Mean: 26.2 yrs	Full-Scale driving simulator	Multimodal TOR warning design and TOR events	**Objective Measures** Reaction time TOR Time to lane change (TTL) Vehicle control metrics (SDLP and SRR) Physiological metrics (SCR and AHR)	Unplanned ODD (V, A, VA, H, AH, VH, VAH) Planned ODD (VA2, VAH2)

Paper	Population	Type	Independent	Dependent	Warning Type
(Jarosch et al., 2017)	Germany Male: 47 Female: 9 Age Mean: 30.10 yrs	Simulator with the motion system	Response time due to NDR	Karolinska Sleepiness Scale (KSS) PERCLOS Blink Rate Blink Duration	V
(Schomig et al., 2015)	Germany 16 total Age Mean: 30.5 yrs	Simulator with the motion system	Drowsiness level	**Objective measure:** Drowsiness Index (Blinking duration, eyelid opening level, blinking frequency)	V (Screen)
(Saxby et al., 2013)	USA Male: 42 Female:66 Age Mean: 19.92 yrs	Static Simulator driving	Active and Passive Fatigue	**Subjective measures:** Mental workload (NASA-TLX) Pre-task and Post task DSSQ, Task Load Index (TLX) Appraisal for Life Events (ALE) Coping inventory for Task Situations (CITS)	V (Arrows on display)
Bourrelly et al., 2019)	France Male: 15 Female: 15 Age Mean 46 yrs	SHERPA Simulator	Time Traffic Condition (Critical, Not Critical)	**Objective Measures** Reaction Times Car Trajectories **Subjective Measures** Drowsiness (Likert Scale) Manoeuvre Performance Effectiveness of TOR Ease-of-use Adequacy of 10s time frame Safety of take-over manoeuvre. Trust Automated System	V & A (multimodal)
(Huang et al., 2019)	USA 16 people Age Mean: 22.8 yrs	Static Driving Simulator	Effect of Multimodal Signal on TOR	**Objective Measures:** Response time Road sign detection accuracy Pupil Diameter **Subjective Measures:** NASA-TLX	V, A, T, VA, VT, AT, VAT
(Shi & Bengler, 2022)	Germany Male: 21 Female:15 Age Mean: 42.3 yrs	Real Driving Setting	NDRTs (Tetris vs Read/Write vs Film) Dark vs Daylight	**Objective Measures:** Take Over Time Takeover Quality (min/max lateral/longitudinal acceleration) Time to Collision **Subjective Measures:** Flow Experience	V (HMI Icon) & A (Tone) (Multimodal)
(Roche et al., 2019)	USA Male: 20 Female: 20 Age Mean: 27 yrs	Simulator	NDRT Modality TOR Design (A vs AV)	**Objective Measures:** Take over time Time to Collision Minimum acceleration Lateral position Steering wheel angle **Subjective Measures:** Workload	A AV

Paper	Population	Type	Independent	Dependent	Warning Type
(Radlmayr et al., 2014)	USA Male: 38 Female: 10 Age Mean: 33.5 yrs	BMW 5 simulator	Group (Baseline, n-back, SuRT) NDRTs Traffic Situation	**Objective Measures:** Take over time Steering wheel angle Brake pedal application Longitudinal acceleration Time to Collision # of collisions Tactile Detection Response Task (DRT) **Subjective Measures:** Complexity (questionnaire)	A (high-pitched tone) & V (icon) (Multimodal)
(Kuehn et al., 2017)	Germany Male: 38 Female:22 Age: 20-76 yrs	Simulator	Secondary Task (manual, monitored, automated + secondary task) Take over situation	**Objective Measures:** Reaction time	V (red hand/steering wheel) & A (multimodal)

Auditory: When comparing auditory TORs and visual-auditory TORs, drivers presented with auditory TORs proved faster takeovers and longer time to collisions. Auditory TORs had better overall takeover behavior and lower subjective workload (Roche et al., 2019). The NHTSA (2016) guidelines dictate that the tone should be 15-30bd louder than the ambient noise but not louder than 90 db. Another study by Lin et al. (2009) shows that a 1750 Hz tone will incite a quicker reaction. Hence, regarding Yun et al. (2020) and Lin et al. (2009), in our experiment, the participants will hear a "beep, beep, beep" sound at 1750 Hz with a tone interval of 30ms.

Visual: Visual cues require visual attention. Visual signs are not useful, but augmented with other modal signals improve steering and braking performance (Bazilinskyy et al., 2017). During a study, a visual-only signal was compared to a bi-modal Visual-Auditory signal system during automated driving (Praetorius, 2020). The results showed higher effectiveness through shorter reaction times for the Visual-Auditory takeover requests. A screen in the center console flashed a signal to the driver to place hands on the steering wheel. The audio was a one-second

1000 Hz sinus tone. Bazilinsky et al. (2017) and NHTSA's report on human factor guidance shows that red, orange, and yellow visual signals elicit quicker drivers' reactions. Furthermore, Politis et al. (2016) found that a simple design like a circular icon has maximum effect. Because of our simulator's limitations, we will be using a white, squared icon flashing on the bottom left of the screen.

Tactile: The findings of Yun et al. (2020) show that haptic modality can elicit more immediate TOR, whereas auditory modality can elicit more stable TOR. Additionally, Spence et al. (2008) found that haptic feedback improves response time, increases attention, and enhances situational awareness. Based on the study conducted by Huang et al. (2020), warning cues with tactile signals yielded faster takeover times. Petermeijuer et al. (2017) looked into drivers' reaction time, availability, and pleasantness concerning tactile modality in the event of a take-over request in autonomous vehicles. They found that alerting via both tactile and auditory modalities was more effective than solely based on tactile modality. We developed a bi-modal TOR Warning System based on the literature to compare the known effects (Table 4).

Table 4: TOR warning system

	Auditory	Visual
TOR Warning System Design	– Beep – 1750 Hz tone – < 90db – Interval: 30 ms	– Square white icon – Bottom left of the frontal screen

Table 5: Environmental factors used in takeover performance studies.

Times	Environment	Findings
(Shi & Bengler, 2022)	A sudden vehicle in front of the car Daylight Dark hours	TOT longer for Dark Hours

(Radlmayr et al., 2014)	Four traffic scenarios A sudden car crash with flashing warning lights (Dangerous situation) 1. Obstacles in middle lane, left and right lane have traffic 2. Obstacle in right lane, no other vehicles present 3. Obstacle in left lane, no other vehicles present 4. Obstacle in middle lane, no other vehicles present	Complexity Affects Take over quality and time. #1 showed higher levels of criticality.
(Kuehn et al., 2017)	Five critical scenarios 1. Change lane or exit highway, no traffic 2. No road markings, vehicle in front travelling same speed, sudden brake, moderate traffic 3. Sensor failure, vehicle in front travelling same speed, sudden brake, moderate traffic 4. Roadworks, high traffic density, stationary vehicle appear 5. Extreme weather conditions, sudden heavy rain, vehicle in front travelling same speed, sudden brake, moderate traffic	
(Zhang et al., 2022)	Two way two lane road- three scenarios: 1- Straight bush 2- Semi open Chevron shaped bush 3- Zig-Zag shaped Chevron bush These are divided into two heights - 1.5m (1) and 0.5m(2 and 3)	Semi open chevron landscape induces better alertness in drivers.
(Tavakoli et al., 2021)	3 Different Environments: Four different types of roads (City, Country, 2 lanes on each side and 3 lanes on each side) Weather (Clear,Cloudy and Rainy) Front Passenger (Passenger/no passenger)	Clear weather has higher gaze and head movement. Average driver is more distracted in city driving.
(Guo et al., 2022)	Four types of landscapes- 1. Enclosed space landscape 2. Semi Closed space landscape 3. Semi Open space landscape 4. Open space landscape.	Drivers attention was more focussed in enclosed landscape.

Measuring Takeover Performance

Simply giving a signal alert to the driver does not imply a well-performed takeover maneuver. According to Pipkorn et al. (2024), drivers may look away from the road during automated driving in favor of tasks unrelated to driving. However, when a change in control is necessary, drivers could do so without being aware of the traffic situation or a potential threat. Therefore, we must establish guidelines for measuring take-over performance in this experiment.

Multiple studies have developed systems to measure take-over performance or how well a driver can resume control over the vehicle when required. Takeover performance can be adequately measured in metrics such as Gaze Reaction Time and Takeover Time; however, these characteristics depend on factors such as Time Budget and Traffic Density (Gold, 2016). The definition of taking control can be broken down into four different reactions: orientation, readiness, action, and stabilization. Readiness to act is described as coming in contact with the system, which can be measured by taking overtime or time to take over (Schomig et al., 2015). In our experiment, we will use readiness to measure takeover performance. The TOT will begin when the TOR is issued and end when the driver places his or her foot on the brake pedal. The literature review is summarized in Table 3, and the environmental factors used in the literature are presented in Table 5.

METHOD

Driving Simulator Experiment

Research suggests that the environment will affect the takeover time, as observed by Zhang et al. (2022). Specifically, takeover time will increase in a high workload-induced fatigue environment. The objective is to assess if the environment (high workload-induced fatigue vs. low workload-induced fatigue) affects the driver's takeover performance after a takeover request is issued in SAE Level 3 vehicles.

A half-factorial experimental design assessed the response time after the takeover request signal in high workload versus low workload scenarios. The response time will depend on the fatigue levels and the environment. To assess the effect of fatigue on driver's performance, the experiment was run for 45 minutes. According to Reinermann et al. (2008), the brain showed a decline in cerebral blood flow (CBFV), which indicates a loss of alertness after 36 minutes of monotonous driving. Participants in the test run are healthy adults with no history of sleep disorders or psychological conditions that might affect the study. Before the experiment began, the moderator ran through the simulator control setup and ensured the participant was comfortable.

Each participant was randomly assigned one of the environments. The participants encountered a TOR signal at minute 15 (control measurement) and minute 45. The moderator checked on the well-being of the participants every 15 minutes throughout the experiment. The timeline is shown in Figure 17.

Independent Variable

There are two independent variables—workload type and fatigue level (moderating variable type). Participants' fatigue level and takeover response time will be recorded after the TOR signal (Table 3) is played. The workload type and TOR warning system are the between-subjects factors, and the fatigue level is the within-subjects factor.

Dependent Variable

Response Time: The timer starts as soon as the beep (auditory signal) starts, and it stops when the driver puts his or her foot on the brake pedal. The brake coefficient will be measured, and the driver's response time, i.e., the time between the beep start and pedal press, will be recorded and logged.

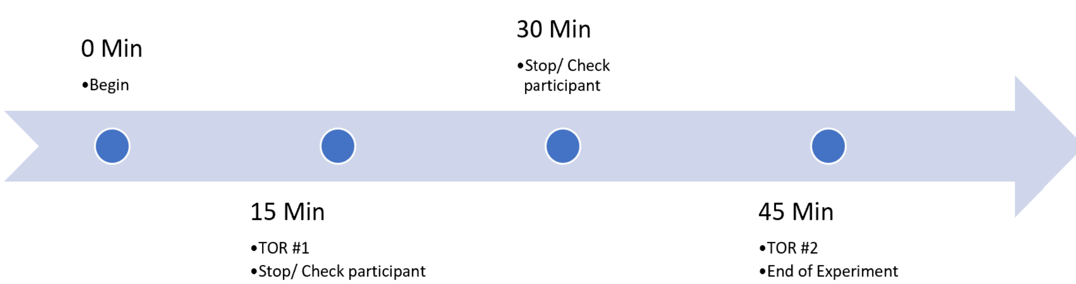

Figure 8: Timeline for Experiment Procedure. The experiment will stop every 15 mins to check on participants' well-being.

Apparatus

This project used RealTime Technology's RDS-1000 motion-based driving simulator. This high-fidelity simulator features a quarter-cab design. Three 65-inch displays display a virtual environment. The simulator comes equipped with SimCreator and SimCreatorDX and a driver automation system.

Environment

An open landscape with long, tangent sections on a four-lane divided highway with a speed limit of 70 mph was used to develop the driving simulation environment. Research shows that reaction time for straight roads is significantly longer during an open landscape scenario, possibly due to the drivers' frequent gaze behavior (Guo et al., 2019). The region next to the shoulder was covered with green vegetation. According to the findings of Yao et al. (2020), the color green gives a pleasant viewing experience for drivers and reduces visual fatigue. The urban terrain remained flat, and asphalt was used for the road surface. We designed a two-environment system to compare the effects of the environment on the drivers' takeover performance. These environments and their salient features are described below:

Environment 1 (low workload-induced fatigue)

The first environment had an open, clear, blue sky. It was during the daytime, and there was high visibility. There was light traffic volume (trucks and cars).

Environment 2 (high workload-induced fatigue)

The second environment had rain and a cloudy sky. It was nighttime, and visibility was reduced. There was heavy vehicle traffic (trucks and cars).

TOR Scenario

The subjects were in a car in full autonomy mode at 70mph. During this simulation, the driver sat back and maintained awareness of the environment while experiencing various fatigue levels. At minute 15:00, a lead car started to slow down until it came to a complete stop. The TOR signal was issued, and the driver was asked to avoid a collision. The time to collision was 6 seconds (Eriksson & Stanton, 2017). The lead car sped up after 6 seconds. The experiment resumed. At minute 45:00, a lead car started to slow down until it completely stopped. The TOR signal was issued, and the driver was asked to avoid a collision. The time to collision was 6 seconds (Eriksson & Stanton, 2017).

RESULTS AND DISCUSSIONS

The paper aimed to compare the effects of two workloads on the driver's performance while driving a vehicle equipped with SAE level 3 automated driving. A combination of visual and auditory warning signals to warn the drivers of the impending collision. Participants were asked to respond to the obstruction as best as their decision suggested. The results indicate that the environment-related workload affected the driver's reaction times.

Eight subjects aged 20-28 (five male and three female) participated in the experiment. None of the participants indicated infractions within the past five years. They all said they drove between 0-2 times/day with a total drive time of 30-45 mins/day. The participants sometimes drive when they are tired and often drive at night. On average, the participants felt neutral towards the usage of an automated system in vehicles and its ability to help drivers quickly respond to unsafe driving conditions. However, they mainly agreed when asked if such a system could help prevent accidents.

Twelve 12 readings were gathered from the study. When analyzing the video footage of the drivers:50% of the subjects kept their hands on the steering wheel at all times 75% of the subjects avoided the collision by either swerving (45%) or braking (55%). These results indicate that drivers' reactions differ and are unpredictable. About a quarter (25%) of the subjects collided with the obstacle by either reacting too late or not reacting at all. After the experiment, these participants indicated they had high trust in the safety system of the simulated vehicle and thought the system would avoid the collision. Over half (58%) of the total subjects reacted before the TOR warning signal. Out of these subjects, 42% saw the obstacle beforehand, started braking before the TOR signals were issued, and were driving in the daytime. This perception-reaction shows that the environment affects the spatial awareness of the participant as it affects their response time and response reaction.

Among all subjects, 33% moved their legs but did not press the brakes when the obstacle appeared. This shows that participants were confident in their driving skills to maneuver around the obstacle rather than panic and stop. Finally, 50% of the participants did not go through the entire experiment as they experienced simulation sickness (According to the Questionnaire).

At the end of the experiment, each participant was provided with a self-reporting questionnaire, the Situational Awareness Rating Technique (SART). The participants were asked to rate dimensions on a scale of 1 (least) to 7 (highest). The average SART score was 3, which indicates that the participants' awareness level was very low after the experiment.

Findings and Limitations

In our experiment, the TOT started when the TOR (A-V) was issued and ended when the driver reacted; however, 58% of the readings showed that the participants had started their reaction before the TOR (A-V) warning signal because they had already been alerted by the jerking motion. The vehicle simulator was programmed to stop or slow down if it encountered an obstacle. During the experiments, the simulator would detect the vehicle in front of it and attempt to slow down to avoid collision. During this time, the simulator would jerk. When the simulator realized it could not control the situation, it gave the driver the TOR (A-V) so the driver could perform the necessary steps to take over safely. The jerk right before the TOR might have affected the drivers' reaction to the TOR (A-V). In other words, by the time the TOR (A-V) was given, the jerk had already alerted the driver.

The results from the simulator must be used with caution when making general statements in real driving. For further studies, we suggest developing a warning system with a tactile signal and another without a tactile signal and comparing its effects with various fatigue levels. Because different drivers react differently to the same scenario, takeover times need to have multiple measurements.

Chapter IV: Conclusions and Recommendations

The study's results suggest that postural data metrics combined with measures like PERCLOS, EAR, and MAR are most promising for detecting and classifying driver fatigue. These outcomes can be utilized for both technology development and the study of fatigue monitoring algorithms. The development of driver fatigue warning systems can leverage the results of this study to detect and distinguish between driver fatigue levels. A major shortcoming of the study is that it relies upon computer-vision techniques for extracting data from the recorded video data of the driver. It may have inferior performance in low-lighting conditions or cases where a driver wears sunglasses or a facemask limiting the approach to extract feature measures like MAR and EAR. This is a noteworthy limitation, as this limits the use of this approach during night-time driving. The data for the study was collected in a simulator-based environment. Although the data for this study was collected in a realistic environment, several studies have observed differences in physical and behavioral measures between simulators and naturalistic environments (Engström, Johansson, & Östlund, 2005). One of the other limitations of this research study is that it did not implement any experimental design for the systematic collection of data to study the effects of different factors on the different states of driver fatigue. These limitations can be addressed by subsequent naturalistic studies implementing an experimental design for data collection with varied driving scenarios.

A driving simulator experiment was conducted to study the effect of different types of fatigue on driver takeover performance. Results indicated differences between takeover responses, although a rigorous analysis could not be completed due to many participants experiencing simulator sickness and not completing the experiment. Future research should further validate the performance of Random Forest models in estimating driver fatigue using driver features like mouth aspect ratio (MAR), eye aspect ratio (EAR), and postural data of the driver (head-to-headrest distance and hand-to-steering wheel distance). Also, additional research is needed to study the interaction between driving environment, driver fatigue, warning modality, and driver response to takeover requests.

References

1. National Center for Statistics and Analysis. (2023, October). Summary of motor vehicle traffic crashes: 2021 data (Traffic Safety Facts. Report No. DOT HS 813 515). National Highway Traffic Safety Administration. Available on https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813515. Accessed on 02/26/2024.
2. Zipper, D. (2021). The Deadly Myth That Human Error Causes Most Car Crashes. Available on https://www.theatlantic.com/ideas/archive/2021/11/deadly-myth-human-error-causes-most-car-crashes/620808/. Accessed on 02/26/2024.
3. Smith, B.W. (2013). Human Error as a cause of vehicle crashes. Available on https://cyberlaw.stanford.edu/blog/2013/12/human-error-cause-vehicle-crashes.
4. NHTSA. (2018). Human Factors Design Guidance for Level 2 And Level 3 Automated Driving Concepts. https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13494_812555_l2l3automationhfguidance.pdf. Accessed on 02/26/2024.
5. NHTSA. (2017). Automated Driving Systems 2.0: A Vision for Safety. Available on https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf. Accessed on 02/26/2024.
6. DiMatteo, J., Berry, D. M., & Czarnecki, K. (2020). Requirements for Monitoring Inattention of the Responsible Human in an Autonomous Vehicle: The Recall and Precision Tradeoff. In *REFSQ Workshops*.
7. Desmond, P. A., & Hancock, P. A. (2001). Active and passive fatigue states. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 455–465). Lawrence Erlbaum Associates Publishers.
8. NHTSA. (2023). Overview of Motor Vehicle Traffic Crashes in 2021. Available on https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813435. Accessed on 03/06/2024.
9. Depestele, S., Ross, V., Verstraelen, S., Brijs, K., Brijs, T., van Dun, K., & Meesen, R. (2020). The impact of cognitive functioning on driving performance of older persons in comparison to younger age groups: A systematic review. *Transportation research part F: traffic psychology and behaviour*, *73*, 433-452.
10. Fonseca, A., Kerick, S., King, J. T., Lin, C. T., & Jung, T. P. (2018). Brain network changes in fatigued drivers: a longitudinal study in a real-world environment based on the effective connectivity analysis and actigraphy data. *Frontiers in human neuroscience*, *12*, 418.
11. National Safety Council. (2019). *Drivers are Falling Asleep Behind the Wheel. https://www.nsc.org/road/safety-topics/fatigued-driver*
12. Huang, R., Wang, Y., & Guo, L. (2018). P-FDCN based eye state analysis for Fatigue Detection. *2018 IEEE 18th International Conference on Communication Technology (ICCT)*.
13. Gwak, J., Shino, M., & Hirao, A. (2018). Early detection of driver drowsiness utilizing machine learning based on physiological signals, behavioral measures, and driving performance. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*.

14. Wang, P., Min, J., & Hu, J. (2018). Ensemble classifier for driver's fatigue detection based on a single EEG channel. *IET Intelligent Transport Systems*, *12*(10), 1322-1328.
15. Savas, B. K., & Becerikli, Y. (2018). Real time driver fatigue detection based on SVM algorithm. *2018 6th International Conference on Control Engineering & Information Technology (CEIT)*.
16. Savaş, B. K., & Becerikli, Y. (2020). Real time driver fatigue detection system based on multi-task ConNN. *Ieee Access*, *8*, 12491-12498.
17. Wijaya, D., Stanley, M., Wilman, P., Lucky, H., & Chowanda, A. (2022). A fatigue detection model based on Convolutional Neural Network. *2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*.
18. Anber, S., Alsaggaf, W., & Shalash, W. (2022). A hybrid driver fatigue and distraction detection model using AlexNet based on facial features. *Electronics*, *11*(2), 285.
19. Mahmoodi, M., & Nahvi, A. (2019). Driver drowsiness detection based on classification of surface electromyography features in a driving simulator. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, *233*(4), 395-406.
20. Gu, W. H., Zhu, Y., Chen, X. D., He, L. F., & Zheng, B. B. (2018). Hierarchical CNN-based real-time fatigue detection system by visual-based technologies using MSP model. *IET Image Processing*, *12*(12), 2319-2329.
21. Zhao, L., Wang, Z., Wang, X., & Liu, Q. (2018). Driver drowsiness detection using facial dynamic fusion information and a DBN. *IET Intelligent Transport Systems*, *12*(2), 127-133.
22. Sikander, G., & Anwar, S. (2018). Driver fatigue detection systems: A review. *IEEE Transactions on Intelligent Transportation Systems*, *20*(6), 2339-2352.
23. Du, G., Zhang, L., Su, K., Wang, X., Teng, S., & Liu, P. X. (2022). A multimodal fusion fatigue driving detection method based on heart rate and PERCLOS. *IEEE transactions on intelligent transportation systems*, *23*(11), 21810-21820.
24. Fatima, B., Shahid, A. R., Ziauddin, S., Safi, A. A., & Ramzan, H. (2020). Driver fatigue detection using viola jones and principal component analysis. *Applied Artificial Intelligence*, *34*(6), 456-483.
25. Guo, J. M., & Markoni, H. (2019). Driver drowsiness detection using hybrid convolutional neural network and long short-term memory. *Multimedia tools and applications*, *78*, 29059-29087.
26. de Naurois, C. J., Bourdin, C., Bougard, C., & Vercher, J. L. (2018). Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness. *Accident Analysis & Prevention*, *121*, 118-128.
27. Ji, Y., Wang, S., Zhao, Y., Wei, J., & Lu, Y. (2019). Fatigue state detection based on multi-index fusion and state recognition network. *Ieee Access*, *7*, 64136-64147.
28. Rajkar, A., Kulkarni, N., & Raut, A. (2022). Driver drowsiness detection using deep learning. *Applied Information Processing Systems: Proceedings of ICCET 2021*, 73-82, Springer Singapore.

29. Zandi, A. S., Quddus, A., Prest, L., & Comeau, F. J. (2019). Non-intrusive detection of drowsy driving based on eye tracking data. *Transportation research record*, *2673*(6), 247-257.
30. Xiao, Z., Hu, Z., Geng, L., Zhang, F., Wu, J., & Li, Y. (2019). Fatigue driving recognition network: fatigue driving recognition via convolutional neural network and long short-term memory units. *IET Intelligent Transport Systems*, *13*(9), 1410-1416.
31. Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., & Zuo, S. (2019). EEG-based spatio–temporal convolutional neural network for driver fatigue evaluation. *IEEE transactions on neural networks and learning systems*, *30*(9), 2755-2763.
32. Hu, J., & Min, J. (2018). Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model. *Cognitive neurodynamics*, *12*, 431-440.
33. Képešiová, Z., Cigánek, J., & Kozák, Š. (2020, January). Driver drowsiness detection using convolutional neural networks. *2020 Cybernetics & Informatics (K&I)*. IEEE.
34. Sun, Z., Miao, Y., Jeon, J. Y., Kong, Y., & Park, G. (2023). Facial feature fusion convolutional neural network for driver fatigue detection. *Engineering Applications of Artificial Intelligence*, *126*, 106981.
35. You, F., Li, X., Gong, Y., Wang, H., & Li, H. (2019). A real-time driving drowsiness detection algorithm with individual differences consideration. *Ieee Access*, *7*, 179396-179408.
36. Deng, W., & Wu, R. (2019). Real-time driver-drowsiness detection system using facial features. *Ieee Access*, *7*, 118727-118738.
37. He, H., Zhang, X., Jiang, F., Wang, C., Yang, Y., Liu, W., & Peng, J. (2020). A real-time driver fatigue detection method based on two-stage convolutional neural network. *IFAC-PapersOnLine*, *53*(2), 15374-15379.
38. Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 4690-4699.
39. Alparslan, K., Alparslan, Y., & Burlick, M. (2020). Towards evaluating driver fatigue with robust deep learning models. *arXiv preprint arXiv:2007.08453*.
40. Li, Z., Li, S. E., Li, R., Cheng, B., & Shi, J. (2017). Online detection of driver fatigue using steering wheel angles for real driving conditions. *Sensors*, *17*(3), 495.
41. Chellappa, Y., Joshi, N. N., & Bharadwaj, V. (2016, August). Driver fatigue detection system. *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*, 655-660. IEEE.
42. Zhou et al., “Predicting Driver Fatigue in Automated Driving with Explainability.” arXiv, Mar. 02, 2021. Accessed: Feb. 11, 2023. [Online]. Available: http://arxiv.org/abs/2103.02162.
43. Zhang, F., Su, J., Geng, L., & Xiao, Z. (2017, February). Driver fatigue detection based on eye state recognition. *2017 International Conference on Machine Vision and Information Technology (CMVIT)*, 105-110. IEEE.
44. Li, R., Chen, Y. V., & Zhang, L. (2021). A method for fatigue detection based on Driver's steering wheel grip. *International Journal of Industrial Ergonomics*, *82*, 103083.

45. Lu, J., Zheng, X., Tang, L., Zhang, T., Sheng, Q. Z., Wang, C., ... & Zhou, W. (2021). Can steering wheel detect your driving fatigue?. *IEEE Transactions on Vehicular Technology*, *70*(6), 5537-5550.
46. Jia, W., Peng, H., Ruan, N., Tang, Z., & Zhao, W. (2018). WiFind: Driver fatigue detection with fine-grained Wi-Fi signal features. *IEEE Transactions on Big Data*, *6*(2), 269-282.
47. Boon-Leng, L., Dae-Seok, L., & Boon-Giin, L. (2015, November). Mobile-based wearable-type of driver fatigue detection by GSR and EMG. *TENCON 2015-2015 IEEE Region 10 Conference*, 1-4. IEEE.
48. Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, *3*(3), 203-220.
49. Ling, Y., Luo, R., Dong, X., & Weng, X. (2021). Driver eye location and state estimation based on a robust model and data augmentation. *IEEE Access*, *9*, 67219-67231.
50. Zhu, S., Lakshminarasimhan, K. J., Arfaei, N., & Angelaki, D. E. (2022). *Author Response: Eye Movements Reveal Spatiotemporal Dynamics of Visually-Informed Planning in Navigation*.
51. Martinez, K. D., & Huang, G. (2022). In-vehicle human machine interface: Investigating the effects of tactile displays on information presentation in automated vehicles. *IEEE Access*, *10*, 94668-94676.
52. Lever, J., Krzywinski, M., & Altman, N. (2016). Points of significance: model selection and overfitting. *Nature methods*, *13*(9), 703-705.
53. Mohsenzadeh Karimi, S., Kisi, O., Porrajabali, M., Rouhani-Nia, F., & Shiri, J. (2020). Evaluation of the support vector machine, random forest and geo-statistical methodologies for predicting long-term air temperature. *ISH Journal of Hydraulic Engineering*, *26*(4), 376-386.
54. Li, N., Jain, J. J., & Busso, C. (2013). Modeling of driver behavior in real world scenarios using multiple noninvasive sensors. *IEEE transactions on multimedia*, *15*(5), 1213-1225.
55. Abbas, Q., & Alsheddy, A. (2020). Driver fatigue detection systems using multi-sensors, smartphone, and cloud-based computing platforms: a comparative analysis. *Sensors*, *21*(1), 56.
56. Alshaqaqi, B., Baquhaizel, A. S., Ouis, M. E. A., Boumehed, M., Ouamri, A., & Keche, M. (2013). Driver drowsiness detection system. *2013 8th international workshop on systems, signal processing and their applications (WoSSPA)*, 151-155. IEEE.
57. Ansari, S., Du, H., Naghdy, F., & Stirling, D. (2022). Automatic driver cognitive fatigue detection based on upper body posture variations. *Expert Systems with Applications*, *203*, 117568.
58. Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manufacturing*, *3*, 2403-2409.
59. Jarosch, O., Kuhnt, M., Paradies, S., & Bengler, K. (2017, June). It's out of our hands now! Effects of non-driving related tasks during highly automated driving on drivers' fatigue. In *Driving Assessment Conference* (Vol. 9, No. 2017). University of Iowa.

60. Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., & Othersen, I. (2015). The interaction between highly automated driving and the development of drowsiness. *Procedia Manufacturing*, *3*, 6652-6659.
61. Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *Journal of experimental psychology: applied*, *19*(4), 287.
62. Feldhütter, A., Hecht, T., Kalb, L., & Bengler, K. (2019). Effect of prolonged periods of conditionally automated driving on the development of fatigue: With and without non-driving-related activities. *Cognition, Technology & Work*, *21*, 33-40.
63. Niu, J., & Ma, C. (2022). Is it good or bad to provide driver fatigue warning during take-over in highly automated driving?. *Transportation research record*, *2676*(2), 762-774.
64. Wu, Y., Kihara, K., Takeda, Y., Sato, T., Akamatsu, M., & Kitazaki, S. (2019). Effects of scheduled manual driving on drowsiness and response to take over request: A simulator study towards understanding drivers in automated driving. *Accident Analysis & Prevention*, *124*, 202-209.
65. Schmidt, J., Laarousi, R., Stolzmann, W., & Karrer-Gauß, K. (2018). Eye blink detection for different driver states in conditionally automated driving and manual driving using EOG and a driver camera. *Behavior research methods*, *50*, 1088-1101.
66. Tavakoli, A., Balali, V., & Heydarian, A. (2021). How do Environmental Factors Affect Drivers' Gaze and Head Movements?.
67. Bourrelly, A., Jacobé de Naurois, C., Zran, A., Rampillon, F., Vercher, J. L., & Bourdin, C. (2019). Long automated driving phase affects take-over performance. *IET Intelligent Transport Systems*, *13*(8), 1249-1255.
68. Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014, September). How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 58, No. 1, pp. 2063-2067). Sage CA: Los Angeles, CA: Sage Publications.
69. Shi, E., & Bengler, K. (2022). Non-driving related tasks' effects on takeover and manual driving behavior in a real driving setting: A differentiation approach based on task switching and modality shifting. *Accident Analysis & Prevention*, *178*, 106844.
70. Li, S., Blythe, P., Guo, W., & Namdeo, A. (2018). Investigation of older driver's takeover performance in highly automated vehicles in adverse weather conditions. *IET Intelligent Transport Systems*, *12*(9), 1157-1165.
71. Yun, H., & Yang, J. H. (2020). Multimodal warning design for take-over request in conditionally automated driving. *European transport research review*, *12*, 1-11.
72. Huang, G., Steele, C., Zhang, X., & Pitts, B. J. (2019, November). Multimodal cue combinations: a possible approach to designing in-vehicle takeover requests for semi-autonomous driving. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 63, No. 1, pp. 1739-1743). Sage CA: Los Angeles, CA: SAGE Publications.
73. Politis, I. (2016). *Effects of modality, urgency and situation on responses to multimodal warnings for drivers* (Doctoral dissertation, University of Glasgow).
74. Petermeijer, S., Bazilinskyy, P., Bengler, K., & De Winter, J. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied ergonomics*, *62*, 204-215.

75. Roche, F., Somieski, A., & Brandenburg, S. (2019). Behavioral changes to repeated takeovers in highly automated driving: effects of the takeover-request design and the nondriving-related task modality. *Human factors*, *61*(5), 839-849.
76. Kuehn, M., Vogelpohl, T., & Vollrath, M. (2017). Takeover times in highly automated driving (level 3). In *25th International technical conference on the enhanced safety of vehicles (ESV) national highway traffic safety administration*.
77. NHTSA. (2016). Human factors design guidance for driver-vehicle interfaces. Washington, D.C: National Highway Traffic Safety Administration, U.S. Department of Transportation.
78. Lin, C. T., Huang, T. Y., Liang, W. C., Chiu, T. T., Chao, C. F., Hsu, S. H., & Ko, L. W. (2009). Assessing effectiveness of various auditory warning signals in maintaining drivers' attention in virtual reality-based driving environments. *Percept Mot Skills, 108(3)*, 825–835. https://doi.org/10. 2466/pms.108.3.825-835.
79. Bazilinskyy, P., Eriksson, A., Petermeijer, B., & de Winter, J. (2017). Usefulness and satisfaction of takeover requests for highly automated driving. *In Road Safety and Simulation International Conference* (RSS2017); October 17-19, 2017; The Hague, The Netherlands.
80. Praetorius, L. (2020). *An Integrative Model of Drivers' Take-over Time in Semi-Automated Driving* (Doctoral dissertation).
81. Pipkorn, L., Dozza, M., & Tivesten, E. (2024). Driver visual attention before and after take-over requests during automated driving on public roads. *Human factors*, *66*(2), 336-347.
82. Gold, C., Körber, M., Lechner, D., & Bengler, K. (2016). Taking over control from highly automated vehicles in complex traffic situations: The role of traffic density. *Human factors*, *58*(4), 642-652.
83. Zhang, Yu-Le & Zhu, Shou-Lin. (2022). The influence of landscape intervention used as an alertness maintaining 'tool' on driving behaviour. *IET Intelligent Transport Systems. 16*, 394-407. 10.1049/itr2.12150.
84. Reinerman, L. E., Warm, J. S., Matthews, G., & Langheim, L. K. (2008, September). Cerebral blood flow velocity and subjective state as indices of resource utilization during sustained driving. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, No. 18, pp. 1252-1256). Sage CA: Los Angeles, CA: SAGE Publications.
85. Yao, X., Ji, B., Li, M., Men, Y., & Jin, X. (2020, February). Research on the impact of road landscape color on the driving fatigue of drivers. In *IOP Conference Series: Earth and Environmental Science* (Vol. 440, No. 4, p. 042044). IOP Publishing.
86. Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: noncritical transitions to and from manual control. *Human factors*, *59*(4), 689-705.
87. Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation research part F: traffic psychology and behaviour*, *8*(2), 97-120.

Appendix A:

Appendix B: Technology Transfer

An Appendix should be included in this final report to document the Technology Transfer activities conducted during the project term, accomplishments towards T2 adoption and implementation by relevant stakehold ers, as well as any relevant post-project T2 plans.

Title	Conference/Audience	Delivery
Driver Fatigue in Prolonged Automated Driving: Research Gaps and Future Directions	Proceedings of the IISE Annual Conference & Expo, 2023	Conference meeting
Identification of driver fatigue through physiological measures	Undergrad and graduate students in Industrial Engineering at UTA	Class lecture
Meeting with Stakeholders	City of Arlington, City of Madison, Association of Unmanned Vehicle Systems International	Virtual meeting
Classification of Driver Fatigue for Prolonged Automated Driving	TRB Annual Meeting, 2025	Conference meeting

Center for Transportation Equity, Decisions & Dollars

The **Center for Transportation, Equity, Decisions and Dollars (CTEDD)** is a USDOT University Transportation Center, leading **transportation policy research** that aids in **decision making** and **improves economic development** through **more efficient, and cost-effective use of existing transportation systems,** and offers **better access to jobs and opportunities.** We are leading a larger consortium of universities focused on **providing outreach** and **research to policy makers**, through **innovative methods** and **educating future leaders of the transportation field.**

University of Texas Arlington

Cal Poly San Luis Obispo

Georgia Tech

USF University of South Florida

Wisconsin University of Wisconsin–Madison

C-TEDD@UTA.EDU
CTEDD.UTA.EDU
@UTACTEDD
@ C__TEDD
@ CENTER FOR TRANSPORTATION, EQUITY, DECISIONS AND DOLLARS
817.272.5138