Detection of Driver Cognitive Distraction Using Machine Learning Methods

by

Apurva Misra

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2020

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Autonomous vehicles seem to be closer than expected on their timeline. However, there is still a decade of driving manual as well as semi-autonomous vehicles before we can experience completely automated vehicles on the road. Hence, the number of deaths due to driving accidents will take a while to drop, and we require alternative ways to prevent them.

Driver distraction is one of the primary causes of accidents. Driver distraction has posed a significant problem since the first car appeared on our roadways. According to WHO findings, 1.25 million people lose their lives every year due to road traffic crashes. One of the major causes of traffic crashes is distracted driving. As a result, there is a profound need and necessity to continuously observe driver state and provide appropriately informed alerts to distracted drivers. As defined by the National Highway Traffic Safety Administration (NHTSA), there are several types of distractions including cognitive, visual and manual distractions, which may be distinguished from each other based upon the resources required to perform the task. Cognitive distraction refers to the "look but not see" situations when the drivers' eyes are focused on the forward roadway, but his/her mind is not. Typically, cognitive distractions can result from fatigue, conversation with a co-passenger, listening to the radio, or other similarly loading secondary tasks that do not necessarily take a driver's eyes off the roadway. This makes it one of the hardest distractions to detect as there are no visible clues whether the driver is distracted. In this thesis, we have identified features from different sources such as pupil size, heart rate, acceleration that are relevant to classify distracted and non-distracted drivers through collection and analysis of driving data collected from participants over multiple driving scenarios. The Machine Learning methods used dealt with classification including, but not limited to Random Forest, Decision Trees, and SVM. A reduced feature set including pupil area, pupil vertical and horizontal motion was found while maintaining an average accuracy of 90% across different road types. Also, the impact of road types on driver behaviour is identified. Information about dominant features which affect the classification would aid early detection of distracted driving, and mitigation through the development of effective warning systems. The algorithm could be personalized to the specific driver depending on their reaction to driving situations. It would enable a safer and more comfortable driving experience.

Acknowledgements

My masters work and this thesis as a result would not be possible without the support of a lot of people. Firstly, I would like to thank my supervisor Dr. Shi Cao for guiding me in my professional career and for being an awesome mentor. Thank you for giving me the opportunity to work with you.

I would like to thank Dr. Hamid Tizhoosh and Dr. Ning Jiang for being my thesis readers. I am grateful for their support, feedback and comments on this thesis to help and improve my research.

I am also thankful to the people who participated in this research.

To my family, Aditya, my mom and dad, my heartfelt appreciation for your continued support over the years.

To my friends, Aakash, Anmol, Georgette, Jing, Kanav, and Surya, thank you for always being there. Finally, thank you, the reader, for reading my work. I hope you find my thesis insightful.

Dedication

This is dedicated to my parents.

Table of Contents

List of Figures

List of Tables

Abbreviations

- **ML** Machine Learning
- **SVM** Support Vector Machine
- **kNN** k-Nearest Neighbours
- **RBF** Radial Basis Function
- **RF** Random Forest
- **DT** Decision Tree
- **EDA** Electrodermal Activity
- **HR** Heart Rate
- **KL** Kullback-Leibler
- **CNN** Convolutional Neural Network
- **IQR** Interquartile range
- **GPS** Global Positioning System
- **IVIS** In-Vehicle Information System
- **NB** Naïve Bayes
- **NHTSA** National Highway Traffic Safety Administration
- **PERCLOS** Percentage of eyelid closure

Chapter 1

Introduction

1.1 Motivation

The increase in popularity of in-vehicle technology has emphasized the importance of vehicle safety and driver experience. The risk of crash and near crash among novice and experienced drivers increases significantly while using a cellphone [1]. Figure 1.1 shows the statistics on phone use from a survey conducted in 2015 by the National Highway Traffic Safety Administration (NHTSA). Distraction is not just limited to cellphone use and expands to conversations with a fellow passenger, adjusting the radio, being lost in thought [2]. Due to distracted driving, fatalities from crashes have increased 28% after the year 2005 and have become a public safety hazard [3].

Vehicle safety aspect has been enabled through vehicle sensors that observe the surrounding environment as well as the driver behavior, which is further used to give alerts to drivers for latent collisions. While the driver assistance aspect is equipped through In Vehicle Information System (IVIS) and systems such as Cruise Control which enable drivers to relax. Vehicle technology has been enriched with an advanced observation of the human driver in the vehicle to bring about a more personalized and safe experience.

Furthermore, driver state observation is necessary for the transition of control from one level to another in autonomous driving. The levels depict the difference in the extent of response and control the driver has over the autonomous vehicle- level 0 (No automation) to level 5 (complete automation), according to SAE J3016 (2018). For instance, driver state observation will aid in deciding when and how to inform the drivers to take control

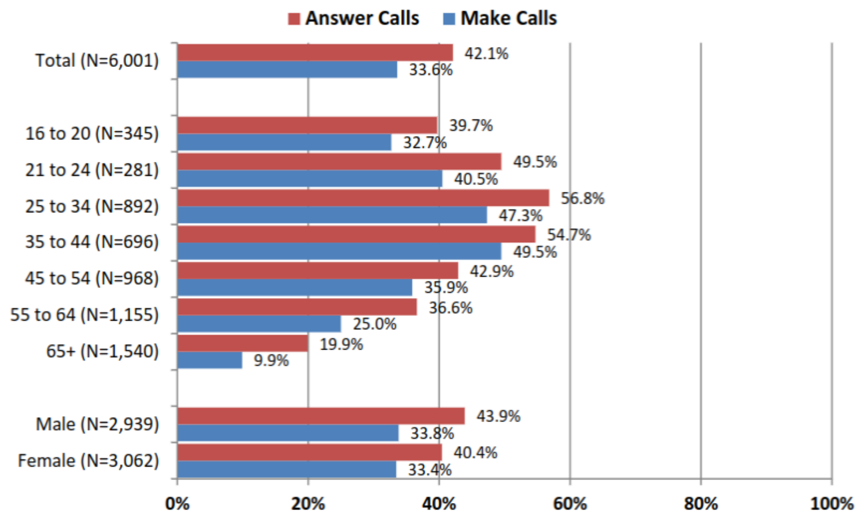


Figure 1.1: Answer or make phone calls while driving by age and gender(% atleast sometimes). Source: [4]

optimally [5] . An ameliorated alert timing will mitigate latent hazards and lead to a safer driving experience.

1.2 Research Objective and Questions

The growth of manually driven and partially automated vehicles has created a need for intelligent transportation systems. Driver assistive systems to inform drivers about their distraction status and to implement a safe transfer of control between the human and autonomous agent. The driver needs to be alert and situationally aware of his/her surroundings while in control of a vehicle, especially in the case of "cognitive distraction" which is difficult to detect through physical changes. The research questions below involve understanding and mitigating cognitive distraction while driving and are the focus of this thesis.

1. Training and development of an ML model to classify distracted from non-distracted driving.

2. Identification of features among the three modalities: vehicle kinematics, physiological measurements, and eye-tracking that enable detecting distracted driving.

3. Examining the effect of road type on distracted and non-distracted driving.

1.3 Thesis Organization

The remainder of the thesis is structured as follows:

1. Chapter 2 provides an overview of the literature, focusing on two aspects - driver distraction and its types, and proposed solutions for detecting them.

2. In Chapter 3, the objectives of this study are defined, and the study experiment protocol is outlined. The method and the equipment used for data collection are presented and discussed.

3. In Chapter 4, preprocessing, feature engineering and algorithms used are elaborated upon.

4. In Chapter 5, feature importance and the relevance of features from different sources is discussed.

5. Chapter 6 summarizes the discussion and conclusions.

Chapter 2

Background

2.1 Review of the Literature

2.1.1 Driver distraction and its types.

In Cognitive Ergonomics, "attention" is treated as a single resource or multiple resources that are utilized during human information processing. The complete model of human cognition consists of sensory inputs from the environment, which is defined as "sensation". Sensation is the converted physical stimuli from the environment into neural signals automatically, requiring no attention. This is followed by the perception of these sensations, specifically interpreting, forming meaningful mental representations. Based on the above observations and knowledge in Long Term Memory as well as Working Memory, an appropriate response is selected and executed, as shown in Figure 2.1.

As seen, the attention resource is an integral part of human information processing, and its divided use or lack of use can lead to inattention. Driver inattention can be classified into two main types [7]-

1. Distraction
2. Mental fatigue

Attention can be disrupted in various ways such as stress, which leads to the tunneling of attention, multitasking causes divided attention resource and others. Further, driver inattention is another factor leading to reduced Situation Awareness. Situation Awareness

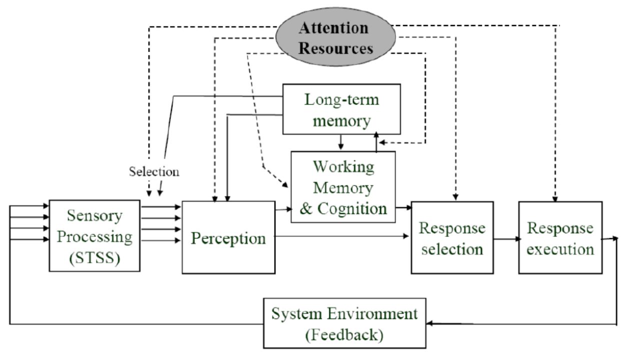


Figure 2.1: Descriptive model of human information processing. Source: [6]

is the state of the level of fitness between memory and task requirements. It consists of three steps that are (a) perception of elements in the current situation (b) comprehension of current situation (c) projection of future status. Situation Awareness is important as it is a measure of driver awareness regarding the task and the environment, leading to a faster assessment of situation and response times. Situation Awareness is affected by fatigue and driver distraction leading to unsafe driving.

A framework for discussing the sources of driver distraction is mentioned in [8] and is shown in Figure 2.2. The sources are "visual" such as when the driver takes his/her eyes off the road for some other task; "cognitive" in which the driver is not processing the information required for safe driving such as having a conversation on a cell phone and "manual" in which the driver is doing something else in the car cabin with his/her hands not being on the steering wheel, for example, drinking water. The three distractors can occur independently or can co-occur while performing an activity. The inner-circle in Figure 2.2 shows activities which pose low level of demands from the visual, cognitive and manual resources while the middle circle represents situations requiring moderate demand on resources. The outer circle depicts concurrent task which poses high demands to the visual, manual and cognitive resources such as interacting with a touchscreen device to retrieve information from the internet. Hence, the crash risk is higher for multitasking activities in the outer circle than in the inner circle. Additionally, the category of "looking but not seeing" is a type of cognitive distraction, called change blindness. That is, drivers take in the sensory inputs but do not process the information.

Earlier driver distraction was defined to be associated with any secondary task (task not

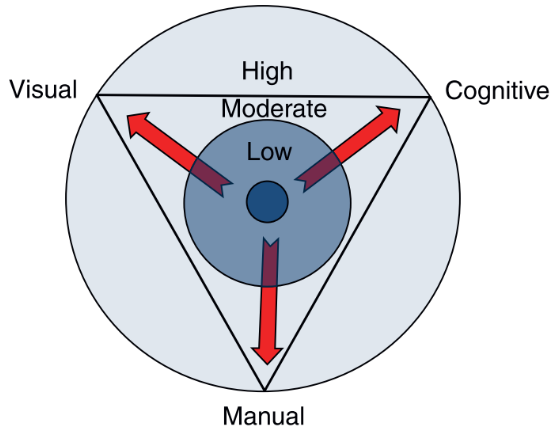


Figure 2.2: Framework for conceptualizing the sources of driver distraction. Source: [8]

related to driving), but this definition has changed over time with the development of complex In-Vehicle Information System (IVIS) and handheld devices. According to the "100-Car Naturalistic Driving Study" conducted by National Highway Traffic Safety Administration (NHTSA) the secondary tasks that contributed to the highest number of crashes or near-crashes were cell phones, internal distraction, and passenger related secondary tasks (primarily conversations).

According to NHTSA's estimate for the year 2015, 72,000 police-reported crashes involved drowsy drivers. Fatigue or tiredness is the inability to continue an activity for longer due to its monotonous nature or physical limitations caused by sleep deprivation, drugs, age, and others [9]. Fatigue can also be defined as the state of mind when an individual continues working beyond a point of decline in task efficiency [10]. Driver fatigue is caused by long haul driving or uneventful and monotonous driving. Driver fatigue can also be caused by sleep deprivation and other personal factors. The definition of driver fatigue does not change among individuals, but the rate at which fatigue is reached varies. Therefore, it is necessary for drivers to know their personal fatigue limit to avoid fatigue-related accidents [11]. Fatigue is considered as a separate category of distraction from cognitive driver distraction. Ways to detect

2.1.2 Ways to detect driver distraction

Driver distraction can be measured through many modalities due to the changes it causes in driver behaviour. These changes such as heart rate, pupil movement, vehicle acceleration can be recorded through sensors and other sources to deduce distraction. It is generally preferred to have a non-intrusive method of observation such as a camera on the dashboard instead of a wearable for the driver to detect a combination of indicators. This aids in retrieving accurate observation without adding to driver distraction due to wearables restricting driver motion or itself becoming a source of distraction. The various categories of measurements for driver distraction detection and their descriptions are given below.

Objective Measures

1. Driver Biological Measures: Measuring and inferencing from physical and physiological signals from the driver's body constitute Driver Biological Measures. This method is also called bio-signal processing. In Figure 2.3 an overview of physiological changes to the human body is shown along with sensors to observe them concerning the lungs and heart [12].

   Bio-signals help with detecting emotions which hinder rational thinking and be-

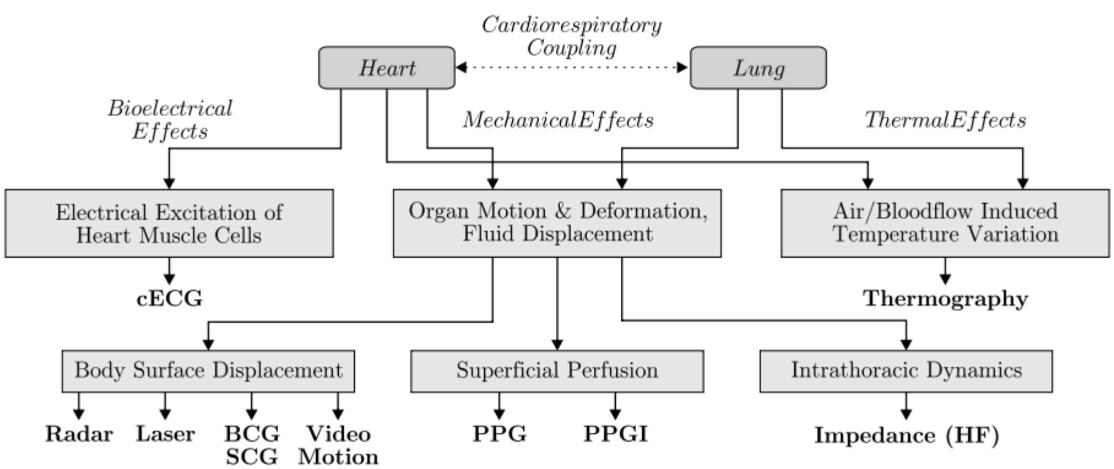


Figure 2.3: Overview of physiological sources, effect and respective sensors. Here, BCG:Ballistocardiograph, SCG:Superior Cervical Ganglion, HF:High Frequency, PPGI:Photoplethysmography Imaging, ECG:Electrocardiography. Source: [12]

haviour [13] with 81% recognition accuracy on eight classes of emotions. Also, using a set of integrated and ambient sensors can be used for in vehicle health monitoring [14]. The relationship between heart rate variability and stressful driving has been studied through simulating a stressful driving environment and observing Electrocardiography (ECG) and Photoplethysmograph (PPG) [15]. ECG signals are the electrical activity of the heart and are generally obtained through placing electrodes on the skin of the individual. PPG uses illumination of skin to measure changes in light absorption to determine heart rate. Most of the methodologies include placing electrodes on the steering wheel, external wearable devices or integration with the car seat for ECG and PPG recording [16], [17], [18], [19].

To have the least obtrusive (no contact) experience for drivers, optical methods are used to gain information from the absorption, reflection, and transmission of radiation and can also be used for respiratory monitoring. It avoids the loss of information due to no contact of drivers to the sensors but has a limitation of environment lighting and varies with the camera quality used [20].

2. Vehicle kinematics: Significant effects of driver distraction are observed on driver's vehicle control, such as drivers adapting to drive at a slower speed to increase available response time when distracted [21], the correlation between steering wheel angle and lane position affected by driver drowsiness [19] and others. Vehicle Kinematics are observed by making drivers perform additional secondary tasks such as a cell phone conversation, navigation control, and playing a radio with varying workload to cause driver distraction on a simulator[22], [23], [24]. Vehicle kinematic data is retrieved through Controller Area Network (CAN)-Bus in cars while simulators have inbuilt system for the same. The retrieved data can be further processed to get high-level signals thus containing more information [24]. The numerous features available through this data can aid in driver distraction detection to varying degree. In [25], the SVM-RFE technique is applied to generate the ranking of features most representative of the driver state, as shown in Table 2.1 with an 80% reduction in false warning without missing out on any critical warnings. Whereas [26] states that braking and turning events are better at characterizing a driver compared to acceleration events. Hence, maybe a combination of these features is a more accurate criterion for detecting driver distraction.

3. Driver Physical Measures: In driver physical behaviour, eye data, head rotation, head nodding, facial features are some of the few that are used quite extensively. Eye data, for example, has a number of features like fixation duration, blink percentage, gaze

Rank	Feature
1	Gas pedal position
2	Brake flag
3	Yaw rate
4	Turn signal
5	Longitudinal acceleration
6	Range rate
7	Speed
8	Heading
9	Range

Table 2.1: Ranking of vehicle kinematics. Source: [25]

deviation, Percentage of Eyelid Closure (PERCLOS), and others. Hence, observing one or all of these features will depend on the purpose and requirement of the system. For instance, [27] states that in eye-tracking behaviour PERCLOS is a very effective drowsiness indicator.

In [28], driver distraction detection and classification is implemented using colour and depth map data from the Kinect sensor. Eye behaviour that is gaze detection and blinking, arm position, head orientation, and facial expressions are merged together and fed into a classifier for accurate sorting. Similarly, in [18] eye movement monitoring is implemented and through a webcam collects frames at a specific rate and sends it to a smartphone to fuse it with other data. Whereas [29] uses a bit more intrusive approach for detecting driver cognitive distraction by using an eye-tracking system to capture the gaze vector, which requires the participants to not wear spectacles or eye make-up achieveing an accuracy of 81.1%.

On the whole, having multiple measures has shown to be more accurate in detecting driver behaviour instead of using just a single type, as asserted in [27], [28] achieving 90% accuracy for distraction detection. These measures are further processed, weighted depending on their effect on driver behaviour before being fed into algorithms such as Support Vector Machines, Dynamic Bayesian Networks, Neural Networks, AdaBoost classifier, Hidden Markov Model among others for distraction detection and recognition [19], [28], [18], [23], [30].

2.1.3 Machine Learning methodologies

Machine learning is the development of algorithms and statistical models seeking to learn patterns from data which otherwise are intractable to develop using classical rule-based programming. Machine learning eliminates the aspect of developing rules and learns from data, as the data changes its ability to adapt to new patterns. There are three classes of ML models:

1. Supervised Learning: Here, data is available along with the labels which aid in the learning process.
2. Unsupervised Learning: The data has no labels, and hence the model has no direction in its learning and tries on its own to learn structures in the data, for example, clustering.
3. Reinforcement Learning: It is about the model learning by trial and error in an environment while getting feedback in the form of a reward or punishment signal.

Most of the ML work involves labeled data and thus uses supervised learning [31]. It is fundamentally learning the mapping between inputs and output labels. Therefore, when a new data point is presented to the model, its able to predict an output label.

As mentioned in Section 2.1.2, the sources of data vary within the qualitative and quantitative space for driver distraction detection and its shown that multiple data sources furthers the accuracy of the system. Research has included Kinect, cameras to collect arm position, head orientation, and facial expression to develop module for detecting driver distraction. Since the source of data is image-based, it generally requires the use of Convolution Neural Networks (CNN), a class of neural networks shown to work well with images as they take the spatial positioning into account[32],[33]. Alternatively, most of the experiments include data collection from multiple sensors like breathing, driving simulator, heart rate monitor, eye tracker, and their synchronisation through resampling, followed by some level of pre-processing, secondary feature generation and finally classification. The classification step is dominantly carried out using SVM in literature [34], [35], [36], [37]. Additionally, logistic regression, decision trees, random forest, kNN, Adaboost, and Neural Networks are also used [38], [39]. Each of these classifiers provide varying levels of explainability and complexity, hence are chosen based on the research objective of the experiment.
Given below are some of the essential classifiers from literature in detail.

SVM

SVM stands for Support Vector Machine and is a popular binary classifier which provides the most optimum boundary between two classes. In Figure 2.4, given labeled training examples, triangles and circles, the algorithm generates an optimal hyperplane that can categorize unseen examples. In two-dimensional space, a hyperplane is just a line while in three-dimensional space, its a plane, subsequently the hyperplane dimension keeps incrementing. The dotted lines are called support vectors and are generated by the closest class points. They are used to keep the maximum margin from the boundary to create optimal separation.

Kernel: As shown in Figure 2.5, there are data points of two classes (circles and squares) in a 2-D plane and they can only be separated through a non-linear boundary such as the black circle illustrated. The data points can be projected to a higher dimension via a function Φ, as shown in the figure on the right (added z-dimension) to make the data linearly separable. The function is written in the form of a kernel function (dot product) $K(xi, xj) = \Phi(xi)^T\Phi(xj)$ used in the SVM calculation. There are numerous kernels such as linear, polynomial, gaussian, radial basis function.

Regularization: The regularization parameter is also known as the “C parameter” and is used to decide the number of miscalssifications allowed while forming the solution. For large values of C the optimizer searches for a small margin hyperplane and vice versa. This parameter is used to find a balance between underfitting and overfitting.

The widespread use of SVM can be attributed to its guaranteed optimality due to the nature of convex optimization, the solution is a global minimum and not a local minimum.

k Nearest Neighbours

k Nearest Neighbours is a lazy learning algorithm; there is no training step. All training examples are stored and at the testing step new examples are classified based on similarity with training examples nearby. The parameter k decides the number of training examples to take into consideration while labeling the new example. The class is assigned based on the majority voting of the k nearest neighbours.

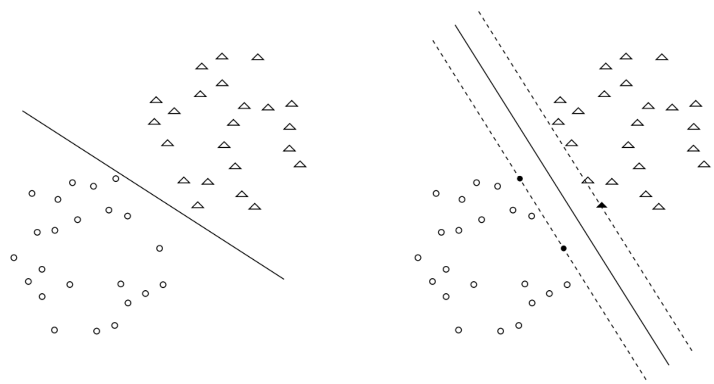


Figure 2.4: The hyperplane on the right gives the most optimal separation while the one on the left does distinguish but not in an optimal way. Source: [40]

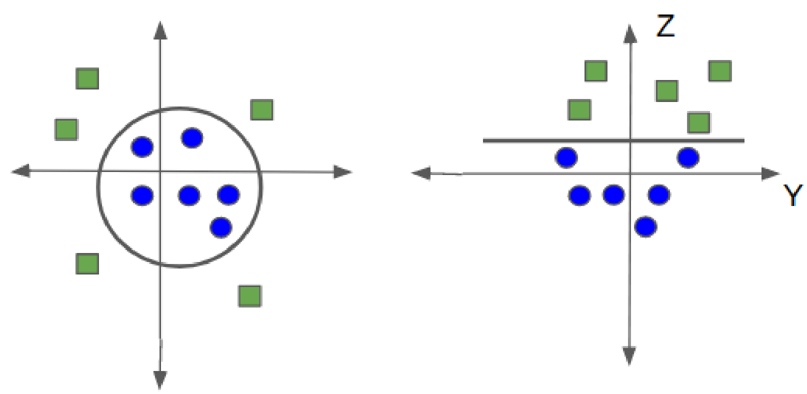


Figure 2.5: Use of kernel to project data in higher dimensions to obtain a linear boundary.

Naïve Bayes

The Naive Bayes classifier is a probabilistic machine learning model based on Bayes theorem given in equation 2.1.

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)} \tag{2.1}$$

The probability of A happening given B has already occurred. *P(A)* is called the prior probability, *P(B/A)* is the likelihood and *P(A/B)* is the posterior probability. It is called naïve as it assumes the conditional independence of every pair of features given the class, which almost is never satisfied in real-world datasets.

Decision Trees

Decision trees are one of the most transparent and explainable classifiers. The decision tree is based on greedy search and hence does not guarantee a globally optimal tree. A decision tree construction involves choosing features for splitting data into subsets having a more homogeneous nature (same labels) [41]. The selection of these features is accomplished through a variety of heuristics, namely, information gain, entropy, and gini-impurity. All of them try to reduce the entropy (a measure of uncertainty in a specific distribution) in the data by splitting it based on the chosen features and values.

There are various algorithms for DT construction based on the heuristic utilized for feature selection as well as measures taken to reduce overfitting, for instance, early-stopping, post-pruning. Decision trees provide an interpretable and simple model with almost no requirement for data preparation.

Random Forest: A collection of trees is called a forest. It is based on the ideology that a collective decision outperforms any individual constituent models. There are two key concepts necessary in building a RF:

1. Random sampling of training data points while building trees
2. Random subsets of features considered when splitting nodes

Each tree learns from a random sample of data points which are drawn with replacement (bootstrapping). Accordingly, each tree is trained on different samples and produces an uncorrelated forest of trees whose collective decision is superior compared to an individual decision. Predictions are made by averaging the output of the whole forest.

Extra Tree classifier: Extra-Tree method stands for extremely randomized trees with the objective to further randomize tree building with the context of input features. It avoids the idea of bootstrapping and chooses random cut points for features at each node. It uses averaging to improve the predictive accuracy and control overfitting.

Artificial Neural Network

It is a collection of non-linear units, also called neurons arranged in a multi-layer network used to model the relationship between data and its labels. A single neuron is equivalent to logistic regression when used along with a sigmoid activation function, a three-layer neural network is shown in Figure 2.6. In simple terms, the learning is carried out by searching for an optimal solution in the search space to minimize a loss function. The loss function is computed on the observed output from the model and the expected output. As the learning happens, the loss computation evolves. There are a variety of loss functions utilized based on the algorithm and the application, for example: Mean absolute error, Mean squared error, and KL divergence. The activation functions in neurons are chosen such that gradient vanishing can be avoided and in the last layer it is based on the expected output type. The algorithm, loss function, and evaluation metric are chosen based on data type, labels, and improved with multiple iterations.

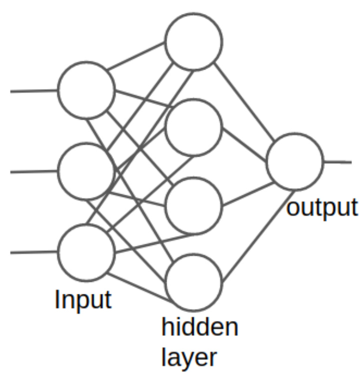


Figure 2.6: A neural network with three layers, information is assumed to be travelling forward.

2.2 Gap in the literature

In literature, the secondary tasks used to bring about distraction have been either a mathematical task, n-back task, a clock task, or a combination of the three [42], [43], [44], [45], which provide a different scale of cognitive workload in comparison to actual driving distractions. Additionally, there are some studies which involve a spoken task between the experimenter and the participant, but due to the subjective nature of questions it results in an inconsistent administration of workload [46].

A typical cognitive distraction observed by a driver is quite different from the above. In this study, we use a distraction task to fill in this gap and make it more similar to a conversation with a fellow passenger.

The only study with a focus group mean age (19.5) representative of novice drivers [47], demonstrates that distracted driving can be distinguished from focused driving using eye-tracking data, but this study is limited because it only included 30 participants. There has been a lack of research with an emphasis on the detection of cognitive driver distraction for novice drivers and even then, it is restricted to a small group of participants. We mainly aimed at the age group of 18-23 years as they are more susceptible to cognitive distraction [1] with an increased sample size and trained our models to classify their driving.

Road types used for driving research has shown an impact on the data as it affects the driving environment and speed limit. A lot of cognitive driver distraction research have restricted use to a few road types or the same road type multiple times [37], [48] in their experiments. There has not been a comprehensive look into different road types and their effect on driver distraction and mitigation. This thesis explores the aspect of road types and whether it might have an effect on driver distraction as an additional feature.

As mentioned in Section 2.1.2, there are multiple ways to detect driver distraction ranging from driver physical measures to subjective measures, which are again varied based on ways of administration. It also mentions that hybrid measures are more accurate in detecting driver behaviour, however, there has not been any research on cognitive driver distraction in which all three (eye metrics, vehicle kinematics, and physiological data) sources were utilised in combination to differentiate between distracted and non-distracted driving. This thesis is an attempt to identify features among the three sources, which leads to a finer classifier and aids in filtering features that contribute more to identifying driver distraction.

The next chapter details the methodology and design for the experiment. Additionally, it explains the technicalities of the equipment used.

Chapter 3

Human Experiments

3.1 Method and Materials

3.1.1 Participants

This study consisted of 40 participants recruited through flyers and emails sent to various departments at the University of Waterloo. Participants were included from the age group of 18-23 years with a valid Canadian full G driver's license and having driving experience under 15,000 km. It was asked of the participants to have good or corrected vision with contacts/glasses, and individuals with known vertigo or motion sickness were not eligible to participate as they are prone to develop simulator sickness.

There were 40 participants (14 females and 26 males) with a mean age of 20.5 (female=20.78 and male=20.34). The average DBQ (Driver Behaviour Questionnaire), included in Appendix A, score was 0.71 where scale ranges from 0 (good driver) to 5 (bad driver), with the female participants having a mean score of 0.68 and the male participants had a mean score of 0.72. The mean age of participants when they received their full G Canadian driving license was 18.87, with 27 participants having less than 50km of driving accomplished in the week before engaging in the study. Twenty-four participants had less than 5,000 km of driving experience in the past 12 months, thus suggesting that the participants were novice drivers.

The eye-tracking component of the study was crucial, and 20 participants noted that they were not wearing any corrective eye wear while 11 participants wore contacts and nine wore glasses during the study.

The study took around 50 mins on average and was remunerated with $20. This study was granted ethics clearance (ORE # 40678) through the University of Waterloo Office of Research Ethics and was conducted as stated in the approved protocols.

3.1.2 Procedure

If a participant chooses to participate, they must sign the consent form. After which they will be asked to fill in a Motion Sickness Susceptibility Questionnaire (MSSQ), included in Appendix A, based on which they might be requested to discontinue from the study to avoid simulator-based sickness and will be remunerated for the time spent based on the scoring obtained in the MSSQ questionnaire. Once the participant has qualified (score below 23) the simulator sickness scoring, they will be asked to fill in demographic details and DBQ which assesses their driving on a scale of 0 (good driver) to 5 (bad driver).

Before starting the experiment, participants have to drive a car simulator through a training scenario consisting of a suburban road with no traffic and multiple turns, to establish competence in handling the equipment. All this while they will be wearing the physiological sensor to collect baseline data. After the training, the participant wears the eye-tracker, which has to be calibrated along with the physiological sensor. The experimental flow is shown in Figure 3.1.

Participants have to drive through six road scenarios consisting of different environments and speed limits, elaborated upon in the following sections. Each of these scenarios are approximately of equal length; speed and other vehicle variables are observed throughout these scenarios. The eye tracker and the wrist band are wiped clean using antiseptic wipes between each participant's use.

3.1.3 Apparatus

In this study, data to evaluate driver behaviour was observed from multiple modalities such as physiological, eye-tracking, and vehicle kinematics. The drive consisted of following directions from a navigation window and heeding the traffic regulations. The vehicle was reset to the previous path if they missed a turn in the navigation and could resume driving from the new position. All the equipment was within a closed room with no windows.

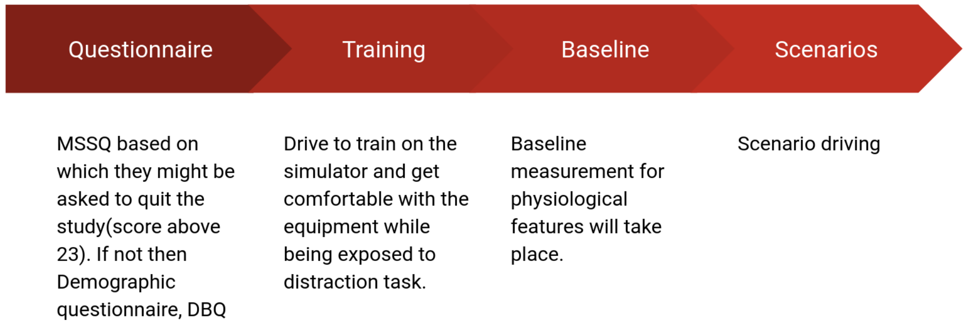


Figure 3.1: Experimental flow for each participant.

The lights were kept on throughout the experiments, and there was a presence of only the experimenter other than the participant inside the room while the experiment was going on. The participants were made aware that they could ask to stop the experiment at any point in time, depending on their comfort level with the equipment.

The following equipment was used for this study.

1. Carnetsoft Driving Simulator: It has 210 degrees surround graphics with a resolution of 5760X1080, consisting of 3 screens - left, center, and right, as shown in Figure 3.2. It has realistic shadows, lighting, and animations. Animations of people and animals and unexpected situations can be controlled to assess hazard anticipation. The density of traffic participants in scenarios can be controlled. Simulator data such as acceleration, and lateral position was collected at 10 Hz and is elaborated upon further in Chapter 4.

2. Dikablis Glasses 3: The eye-tracker needs 4-point calibration and has an accuracy of 0.1-0.3 degrees. The eye-tracking frequency is 60Hz (per eye), and the scene camera recording frequency is 30 Hz with a resolution of 768x576 px.

3. E4 Empatica wrist band: It is an unobtrusive physiological monitoring band which collects data on PPG (photoplethysmography), EDA (Electrodermal activity) at 4Hz, 3-axis accelerometer at 32Hz, and skin temperature at 4Hz. The band can connect to any computing device with a Bluetooth connection and transfer data in real-time.

4. Speakers: Provides sound from the road, wind, tires, engine noise, and the distraction task, which consisted of audio played periodically and controlled based on the driver's location in the scenario.

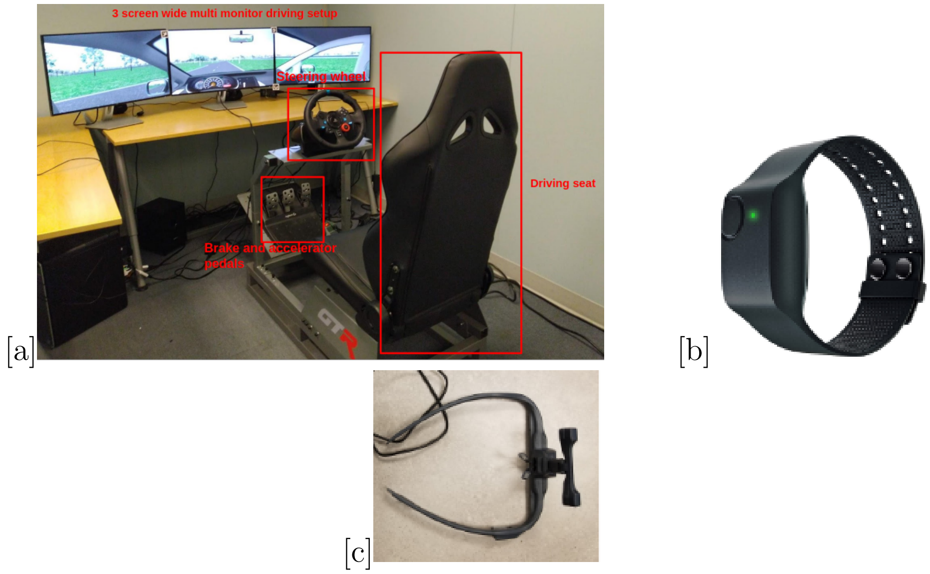


Figure 3.2: a) Driving simulator with the red bounding boxes indicating the parts b) E4 Empatica wrist band, Source: [49] c) Dikablis Eye tracker

3.1.4 Experimental Design

Scenario design was developed based on the literature of drivers' scanning and mitigating patterns for latent hazards. Latent hazards are potentially dangerous threat that may cause an accident but will not lead to one in these simulator based scenarios. Scenarios covered various road types from sub-urban to highway to explore different driving patterns. They were implemented with daylight to avoid the effects of ambient lighting conditions. Each of these scenarios had a speed limit, which was conveyed through the signage in the simulation as well as mentioned before each drive. Each scenario has a latent hazard present, and this zone is further referred to as the critical zone in the thesis centred from the latent hazard location. Latent hazards varied based on the road type and the surrounding environment. Each participant drove all six scenarios, presented to them in a pseudo-random order, the three scenarios containing distractions were chosen pseudo-randomly while maintaining the total of 20 drivers performing a scenario with distraction and 20 drivers performing a scenario without distraction. No driver experienced a scenario twice.

The following are description of the scenarios.

1. Work zone scenario (110 km/h): There is a work zone in the emergency lane of a two-lane highway, two lanes in each direction. There is light traffic in the opposite lane, which was separated by a divider. The latent hazard is a worker hidden in the work zone behind a bulldozer [50],[51].

2. Curve scenario (80 km/h): Two trucks are parked on either side of a curved segment in a sub-urban road type, which makes it harder to perceive oncoming traffic and hazards hidden behind the trucks. There is no other traffic in this scenario, and the latent hazard is a pedestrian hidden behind the truck on the right [52].

3. Stop-controlled intersection scenario (50 km/h): Stop-controlled four-way intersection is to be navigated by the driver in an urban environment where the line of sight of either periphery at the intersection is severely limited by the placement of trucks with the stop signage obscured by vegetation. There are no other traffic participants in the scenario[50],[51].

4. Pedestrian crossing (50 km/h): A crosswalk at an intersection of a two-lane city road with one lane in each direction. A truck is parked on the left lane and the latent hazard is a driver hidden behind the truck. There are no other traffic participants in the scenario [53].

5. School zone (50 km/h): A sub-urban two-lane road with one in each direction having a crossing in a school zone with early signage cautioning about school children. There is vegetation blocking a pedestrian trying to cross at the crosswalk with the presence of multiple people playing in the park on the other side of the road [53].

6. Parked vehicles (50 km/h): A two-lane road with one in each direction, and the driver has to move straight through along a line of parked cars to the right. There are no other traffic participants and the latent hazard is a car with its turn signal on trying to pull out into the path of the driver [53].

Scenario / Type	Highway	Curved	Parking	School zone	Pedestrian crossing	Stop controlled
Distraction	20	20	20	20	20	20
Non-distraction	20	20	20	20	20	20

Figure 3.3: Experimental design.

Cognitive Distraction task

The cognitive distraction task was a spoken task, but instead of interacting with the experimenter, a series of statements sounded through the speakers after a constant period of five seconds, to which the driver responded. It initiated in the path preceding the critical zone and terminated in the following path after the critical zone. This secondary task acted as an alternative to conversations carried out while driving vehicles.

Before the start of the secondary task, there was a beep to alert the driver and a beep after all the sentences were completed and answered to. The sentences were similar to the grammatical reasoning tasks used by [54] and are considered to provide a comparative workload as a hands-free cellphone call.

The sentences were about four to five words long, and after hearing the sentence the participant was supposed to answer in the five-second period before the start of the next sentence. For example, a sentence and the answers are given below-
Statement: The rat drove the car
Expected response: Rat, Car, No

After each sentence, the participant was required to list out aloud the “subject”, “object”, “yes/no” - depending on whether the sentence was plausible. A positive example would be-
Statement: Ron fixed the door
Expected response: Ron, Door, Yes

The next chapter looks into the data preprocessing steps. It explores features through visualisation and statistical methods as well as elaborates upon the data gathering steps.

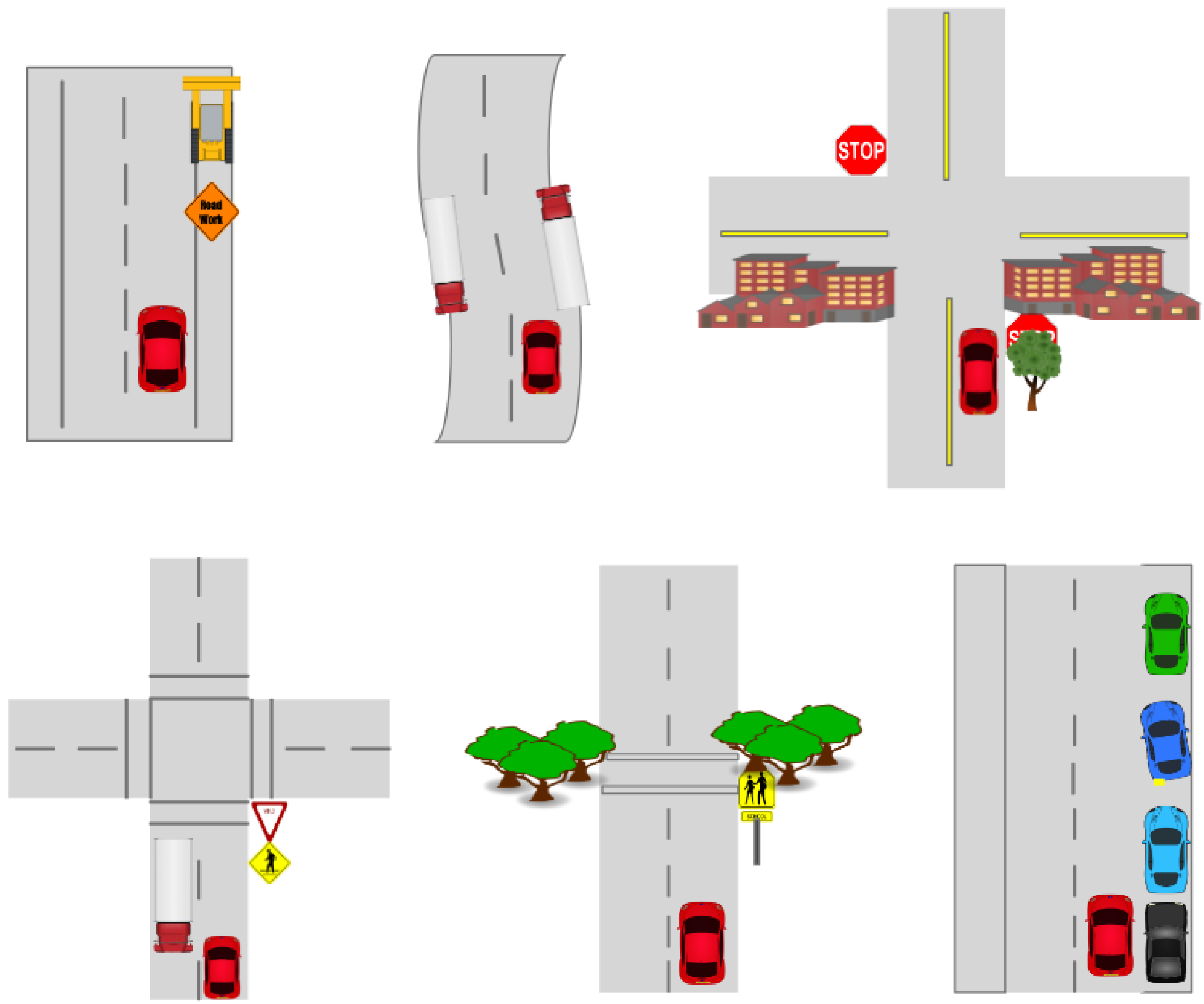


Figure 3.4: (Top left to right): 1. Work zone scenario, 2. Curved scenario with trucks parked, 3. Stop controlled intersection with limited visibility, 4. Pedestrian crossing with a parked truck, 5. School zone scenario, 6. Parking zone scenario.

Chapter 4

Data Collection and Description

4.1 Data Description

4.1.1 Vehicle Kinematics Data

Each scenario was approximately 2 mins long, and a script file was written to generate a CSV file to store kinematics data for each participant's scenarios. The vehicle kinematics were sampled at 10 Hz to maintain a manageable size of the generated data file; a value is recorded for each variable every 100 ms. The data variables sampled in this study are given in Table 4.1 along with engineered features.

The following steps were utilized for data cleaning:

1. Dropping columns which were not needed for the analysis.
2. Generating timestamps using the initial timestamp based on the sampling frequency of 10 Hz.
3. Storing the timestamp intervals when the participant enters the critical zone and drives out of the critical zone and filtering the data from other sources based on it.
4. Removing the data points which mirror resetting, when the drivers did not follow the navigation.

ID	Feature	Description
1	Velocity	Velocity in m/s
2	Acceleration	Acceleration in m/s^2
3	Lateral velocity	Lateral velocity in m/s,Left(+ve values) and Right(-ve values)
4	Lateral Position	Left (+ve), Right (-ve) wrt center line of right lane
5	rpm	Engine rotations per minute
6	Steer	Steering wheel angle (in degrees)
7	Wheel angle	Front wheel angle (in degrees)
8	Heading	Heading wrt road(in degrees)
9	TLC	Time to line crossing, Left (+ve) Right (-ve)
10	Gas	Accelerator pedal position (0 to 100)
11	Brake	Brake pedal position (0 to 100)
12	Steering speed	Steering wheel rotation velocity in degrees/s
13	Steering error	Deviation b/w actual angle and required angle in degrees
14	Longitudnal velocity	Secondary feature (in m/s)
15	Steering standard deviation	Secondary feature (in degrees)
16	SDLP	Secondary feature
17	Steering error mean	Secondary feature (in degrees)

Table 4.1: Vehicle kinematics measures from the driving simulator and their descriptions.

A windowing step was performed to reduce the frequency of data points and map information from within a window to a single data point, hence condensing the information. A non-overlapping window of 1-second duration was used and a number of secondary features were generated based on literature [55] to get a finer classification. The features generated are given below:

1. Longitudinal velocity: It was derived from the total velocity vector and lateral velocity vector. It represents the vehicle velocity in the direction of advancement.

2. Steering standard deviation: It is the standard deviation of the steering wheel angle (Steer in Table 4.1).

3. SDLP: It is the standard deviation of lateral position, considered to be a relevant feature as cognitive load has shown to lower SDLP [56].

4. Steering error mean: It is the mean of the steering error at index 13 shown in Table 4.1, which is the deviation between the expected steer movement and the one performed by the driver.

Standard scaling was used to normalize the data to attain columns with zero mean and unit variance based on different dataset splits. It removes the effect of units used in the measurements as well as the range of each column leading to a faster convergence to solutions by a few algorithms. This step was performed selectively based on the algorithm's need such as SVM, kNN, NB, and SVM-RBF.

4.1.2 Physiological sensor data

The participant wore the wrist band while they were training on the simulator to correspond to the baseline measurements as well as while they were driving through the experimental scenarios. The measurement duration could be controlled through a button on the wrist band, the measurement was stopped once after the training and initiated again for the experimental scenarios hence recording all the six experimental scenarios in one attempt. The physiological sensor transfers the data to a system in real-time which can be further converted into folders corresponding to each participant containing CSV files of Heart Rate, Electrodermal Activity, Temperature, Accelerometer data.

The following steps were performed for data cleaning -

1. Generating timestamps based on the initial timestamp and the corresponding frequency of sampling for the datatype.
2. Filtering data points based on the interval of the critical zone obtained from driving simulator data.
3. Generated features to be used for further analysis, showing the absolute change of physiological variables compared to their baseline average HR change, EDA change, temperature change.

A windowing step similar to the previous one was performed to bring down the frequency to 1Hz, this step was necessary for the synchronization of physiological data with driving simulator data. The standard scaling was performed based on the algorithm's input requirement on all the features such as SVM, kNN, NB, and SVM-RBF.

4.1.3 Eye-tracker data

The Dikablis eye-tracker used in this study was wired and connected to an adapter, which was then connected to a laptop. There was a wide range of data variants available for analysis through the D-lab software. Data could also be processed and visualized within the software, and it also provided video of the participant's view with gaze cross-head overlap. The raw gaze data file was exported as CSV file from the software and analyzed.

The eye tracker consists of multiple cameras, as shown in Figure 4.1. The two eye cameras can be adjusted to track the pupil generating a gray-scale video available in D-lab software. The scene camera is also adjustable with an opening angle of up to 90 degrees which generates a RGB video file.

The fluctuation of a participant's head in its position moves the scene coordinate system along with it, as the eye tracker moves along with the head. This leads to the gaze position getting affected by head movement. D-lab has a special feature to eliminate the effect of head movement by using markers, QR-codes that can be recognized by D-lab. The markers create a new coordinate system, and the gaze is calculated with respect to the markers, eliminating the effect of head movement.

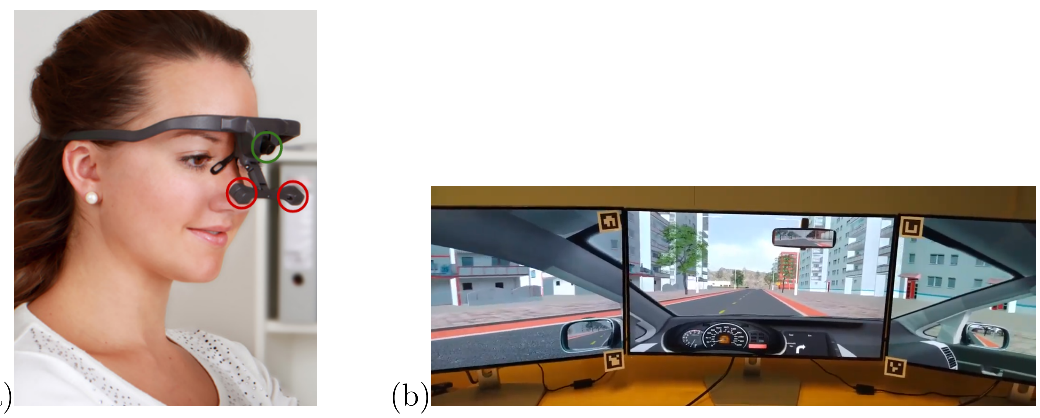
(a) (b)


Figure 4.1: a) Cameras present on the eye tracker: Red circles represent eye-cameras which track the pupils, green circle represents scene camera which records the view of the participant. Source: [57] b) The four markers around the middle screen used in this study.

Eye-trackers are capable of generating variant features based on the data captured using the cameras. Different algorithms utilizing various thresholds on velocity and displacement are employed for feature generation. Given below is a brief description of the features used in this study:

ID	Category	Features	Description
1	Pupil X position	Pupil X	X Position in [px] of detected pupil centre in coordinate system of eye-camera
		Right Pupil X	Pupil X for only Right eye
		Left Pupil X	Pupil X for only Left eye
2	Pupil Y position	Pupil Y	Y Position in [px] of detected pupil centre in coordinate system of eye-camera
		Right Pupil Y	Pupil Y for only Right eye
		Left Pupil Y	Pupil Y for only Left eye
3	Pupil Area	Pupil Right Area Pupil Left Area	Size of the detected pupil in [px] of the the eye-camera
4	Pupil Width	Pupil Right Width Pupil Left Width	Width of detected pupil in [px] of the eye-camera
5	Pupil Height	Pupil Right Height Pupil Left Height	Height of detected pupil in [px] of the eye-camera
6	Saccades	Pupil Right Saccade Pupil Left Saccade	0 or 1, depends on whether a saccade was detected
7	Saccades Duration	Pupil Right saccade dur Pupil Left saccade dur	0 or 1, depends on whether a saccade was detected
8	Saccades Angle	Pupil Right Saccade angle Pupil Left Saccade angle	Angle of saccade in degrees
9	Fixations	Pupil Right Fixation Pupil Left Fixation	0 or 1, depends on fixation detected in current frame
10	Fixations Duration	Pupil Right Fixation dur Pupil Left Fixation dur	Duration of fixation in [s]
11	Blink	Blink	Secondary Feature
		Blink ratio	Secondary Feature
12	Saccade	Saccade ratio	Secondary Feature
13	Fixation	Fixation ratio	Secondary Feature

Table 4.2: Features obtained through eye-cameras and its description.

1. Gaze point: One gaze point equals one raw sample captured by the eye tracker; it corresponds to the coordinates where the eyes are looking for a particular point in time.

2. Fixation: Alignment of the eyes such that the image of the fixated area of interest falls on the fovea for a given time period (duration from 100ms-300ms), it corresponds to the gaze point maintained at a consistent position for a certain amount of time. It is an indicator of user attention.

3. Saccades: Brief fast movements of the eyes that change the point of fixation, it refers to eyes moving in jumps.

In this study, the raw data file was retrieved from D-lab and analyzed. The eye-tracking data was sampled at 60 Hz, and the timestamps were represented in Coordinated Universal Time (UTC). It consisted of two data streams, eye data and field data, described below:

Eye-data

Eye-data is based on the image of the eye cameras and measurements on its coordinate system. The features for the left and right eye are shown in Table 4.2. The features from the left eye were dropped from further analysis due to their high correlation (measure of how strongly pairs of variables are related) to right eye features.

The saccade and fixation detection is performed by D-lab using a velocity-based algorithm with a threshold of 100 degrees/second; movement speed higher than the value is interpreted as saccade, and movement speed lower than this value is interpreted as fixation.

Secondary features were generated according to literature to enhance the classification task along with windowing as given below:

1. Blink: It is predicted by using the pupil values; if 0 then there is no pupil detected (blink), this allows generating a binary feature indicating whether a frame contained a blink or not.

2. Blink ratio: It is the ratio of number of blinks in a window to the length of the window (in datapoints) [58].

3. Fixation ratio: The ratio of number of fixations detected in a window with the length of the window (in datapoints).

4. Saccade ratio: The ratio of the number of saccades detected in the window with the length of the window (in datapoints).

Field-data

Field-data is based on the image of the scene-camera, and the measurements are carried out in its coordinate system. It consists of the X and Y gaze coordinates in the scene camera coordinate system as 'Scene Cam Original Gaze X' and 'Scene Cam Original Gaze Y'. Secondary features were generated to account for the gaze distribution in the horizontal and vertical direction using standard deviation - 'SceneXstd', 'SceneYstd'; described in Table 4.3.

ID	Feature	Description
1	Scene Cam Original Gaze X	X Position of the Gaze given in [px]
2	Scene Cam Original Gaze Y	Y Position of the Gaze given in [px]
3	SceneXstd	Secondary Feature
4	SceneYstd	Secondary Feature

Table 4.3: Features obtained from scene camera and its description.

4.2 Exploration

An exploratory analysis was done on the features from all the three sources to gather insights and look for unique patterns. It is a way to summarize all their main characteristics numerically and visually. In particular, looking into correlations between features, supplements our insight between their relationships and is also necessary for removing correlated features to improve model performance.

4.2.1 Eye-tracking data

Data files from a few participants were empty due to equipment failure and hence discarded. The available data used for analysis is described in Table 4.4. Each scenario's data was treated as a separate data frame for analysis and the correlation between features for each scenario was generated. A threshold of absolute value of 0.8 was used to observe

highly correlated features. Some common characteristics which were expected from the pre-processing step were noticed, for instance, pupil height and width being correlated to pupil area; blink correlated to blink ratio; fixation ratio correlated with fixations; Pupil X correlated with Scene Cam Gaze X. Also, there were some unusual correlations, such as SceneYstd positively correlated with SceneXstd and blink ratio negatively correlated with fixation. The correlation matrix is illustrated using a heat map in Figure 4.2 for "curved" scenario, lighter shades correspond to positive correlation while the darker shades represent negative correlation.
A t-test was performed on the data to assess the significance in difference between the two means, and the p-values for the features for the corresponding road types are given in Table 4.5.

Scenario	Participant:dis	Participant:non-dis	Datapoints
Highway	20	20	2446
Curved	20	20	3113
Parking zone	20	19	2452
Stop controlled	20	19	2495
School zone	19	18	1386
Pedestrian crossing	20	19	2459

Table 4.4: Data available from participants and their length after windowing step for eye-tracker, here "dis" stands for distraction.

4.2.2 Vehicle kinematics data

Data files from the driving simulator had lesser data loss compared to the eye-tracker with only one empty file for a participant, the number of data points for each scenario after the windowing step is shown in Table 4.6. The driving variables showed typical correlations that were expected, such as steer being highly correlated to wheelangle. The visualisations for different features were also generated and are shown for a few features in Figure 4.3. There were not any apparent differences between the visualisations of the treatment which could be perceived. A t-test was performed on the data to assess the significance in difference between the two means, the p-values for the features for each road type is described in Table 4.7.

Features	Highway	Curved	Parking	Stop	School	Pedestrian
Pupil X	0.173	0.537	0.898	0.393	0.537	0.195
Pupil Y	0.105	0.969	0.706	0.789	0.325	0.711
Pupil Right Fixation	0.732	0.478	0.045*	0.132	0.122	0.683
Pupil Right Fixation dur	0.933	0.71	0.098	0.043*	0.092	0.629
Pupil Right Area	0.468	0.212	0.288	0.098	0.574	0.956
Pupil Right Height	0.536	0.24	0.414	0.229	0.529	0.802
Pupil Right Width	0.688	0.259	0.412	0.29	0.297	0.733
Right Pupil X	0.121	0.683	0.027*	0.244	0.521	0.17
Right Pupil Y	0.169	0.926	0.228	0.86	0.407	0.54
Pupil Right Saccade	0.138	0.838	0.653	0.704	0.73	0.827
Pupil Right Saccade angle	0.062	0.267	0.456	0.288	0.348	0.296
Pupil Right Saccade duration	0.171	0.953	0.675	0.55	0.743	0.276
Scene Cam Original Gaze X	0.168	0.806	0.779	0.793	0.906	0.152
Scene Cam Original Gaze Y	0.953	0.59	0.175	0.83	0.946	0.396
Blink	0.357	0.28	0.014*	0.068	0.09	0.347
SceneXstd	0.761	0.769	0.007*	0.742	0.042*	0.794
SceneYstd	0.205	0.984	0.001*	0.229	0.106	0.585

Table 4.5: P-values for eye-tracking features, '*' refers to p-value<0.05.

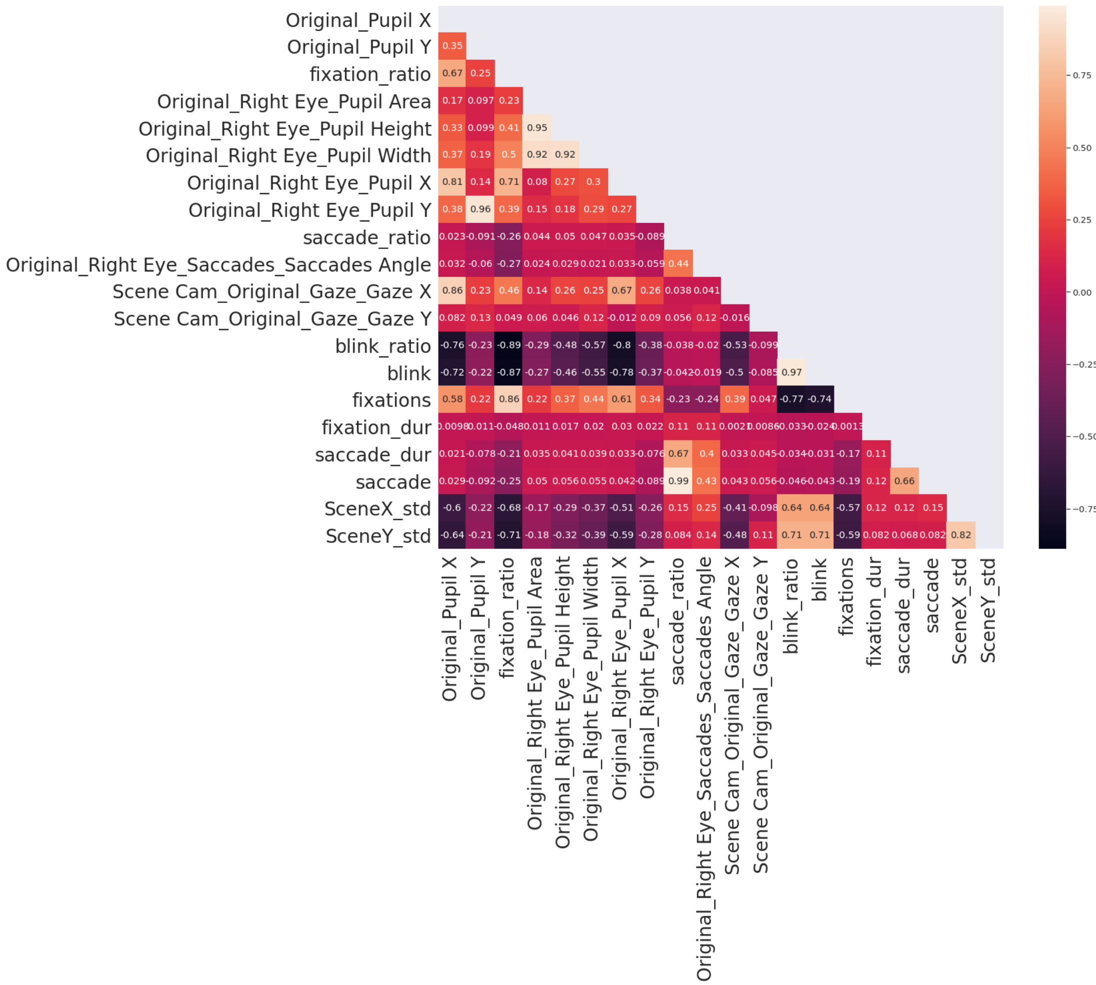


Figure 4.2: Heat map of correlation between eye-tracking features for “curved” scenario.

4.2.3 Physiological data

Physiological sensor had no data loss but due to the missing data file from the driving simulator, the time interval for the critical zone was not available for that participant and hence, that particular file was excluded from further analysis. There were no high correlations among the features. Moreover, the preliminary stats showed the presence of an outlier, that is the maximum value of the feature being quite different from the median value, which is also observed in Figure 4.4 corresponding to a particular participant.

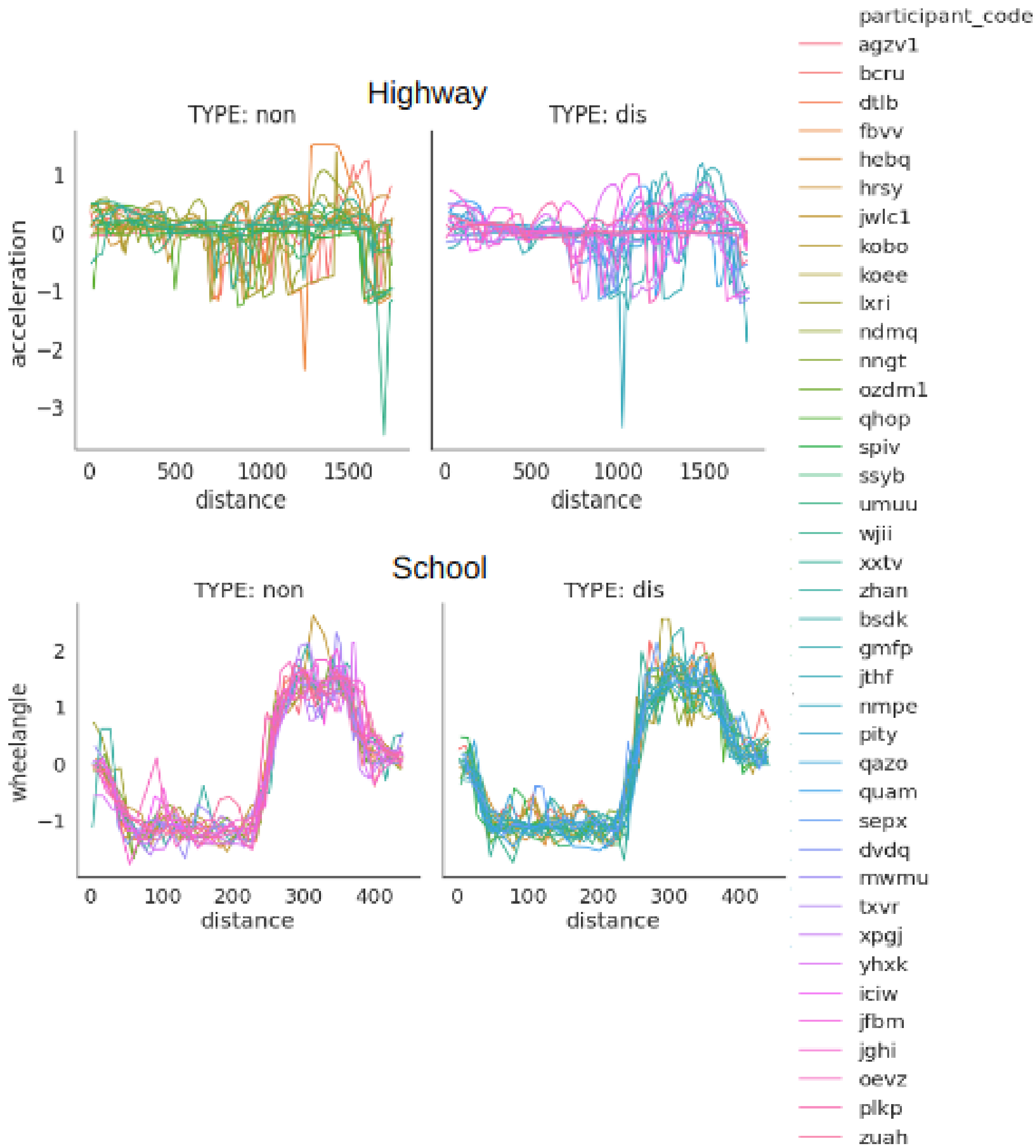


Figure 4.3: Acceleration for highway scenario illustrating a comparison between treatment of secondary task. Wheelangle for school zone scenario illustrating a comparison between treatment of secondary task.

Scenario	Participant:dis	Participant:non-dis	Datapoints
Highway	20	20	2445
Curved	20	20	3118
Parking zone	20	20	2499
Stop controlled	20	20	2625
School zone	20	19	1453
Pedestrian crossing	20	20	2501

Table 4.6: Data available from participants and their length after windowing step for driving simulator, here “dis” stands for distraction.

Additionally, observing the visualisations of features comparing the groups treated to the secondary task and otherwise demonstrated that the EDA-change varied significantly towards the end of the drive for participants subjected to distraction while the other group's EDA-change remained consistent as shown in Figure 4.4. However, when a t-test was performed on the data to assess the significance in difference between the two means, the p-value was not significant. The t-test was performed after inspecting for equality of variance to choose between Welch's or Student's t-test.

4.2.4 All modalities

All the modalities of data were combined to form a single dataset with 40 features, including driving simulator (17), eye-tracker (20) and physiological sensor (3). Resampling of sources to 1 Hz frequency enabled the synchronisation of data points. Since the missing data files from the sources did not overlap, the length of the combined dataset was shorter than expected for each road type shown in Table 4.8.

The next chapter looks into the implementation of ML algorithms to form boundaries to discriminate between distracted and non-distracted data points. Additionally, it explores the features which contribute towards classification.

Features	Highway	Curved	Parking	Stop	School	Pedestrian
Velocity	0.274	0.484	0.889	0.013*	0.729	0.287
Acceleration	0.971	0.059	0.731	0.645	0.148	0.863
Lateral velocity	0.507	0.151	0.473	0.763	0.26	0.906
Lateral position	0.256	0.086	0.208	0.164	0.407	0.119
rpm	0.188	0.689	0.442	0.031*	0.564	0.91
Steer	0.147	0.894	0.971	0.281	0.084	0.9
Wheel angle	0.144	0.894	0.971	0.281	0.083	0.899
Heading	0.568	0.44	0.186	0.383	0.01*	0.843
Gas	0.55	0.469	0.75	0.042*	0.854	0.574
Brake	0.601	0.422	0.722	0.26	0.048*	0.551
Steering speed	0.014*	0.684	0.032*	0.138	0.071	0.888
Steering error	0.369	0.765	0.956	0.367	0.77	0.627
Longitudnal velocity	0.278	0.493	0.867	0.013*	0.736	0.298
Steering standard deviation	0.121	0.086	0.563	0.892	0.36	0.032*
SDLP	0.004*	0.959	0.236	0.705	0.512	0.39
Steering error mean	0.369	0.765	0.956	0.367	0.77	0.627

Table 4.7: P-values for vehicle kinematics features, '*' refers to p-value<0.05.

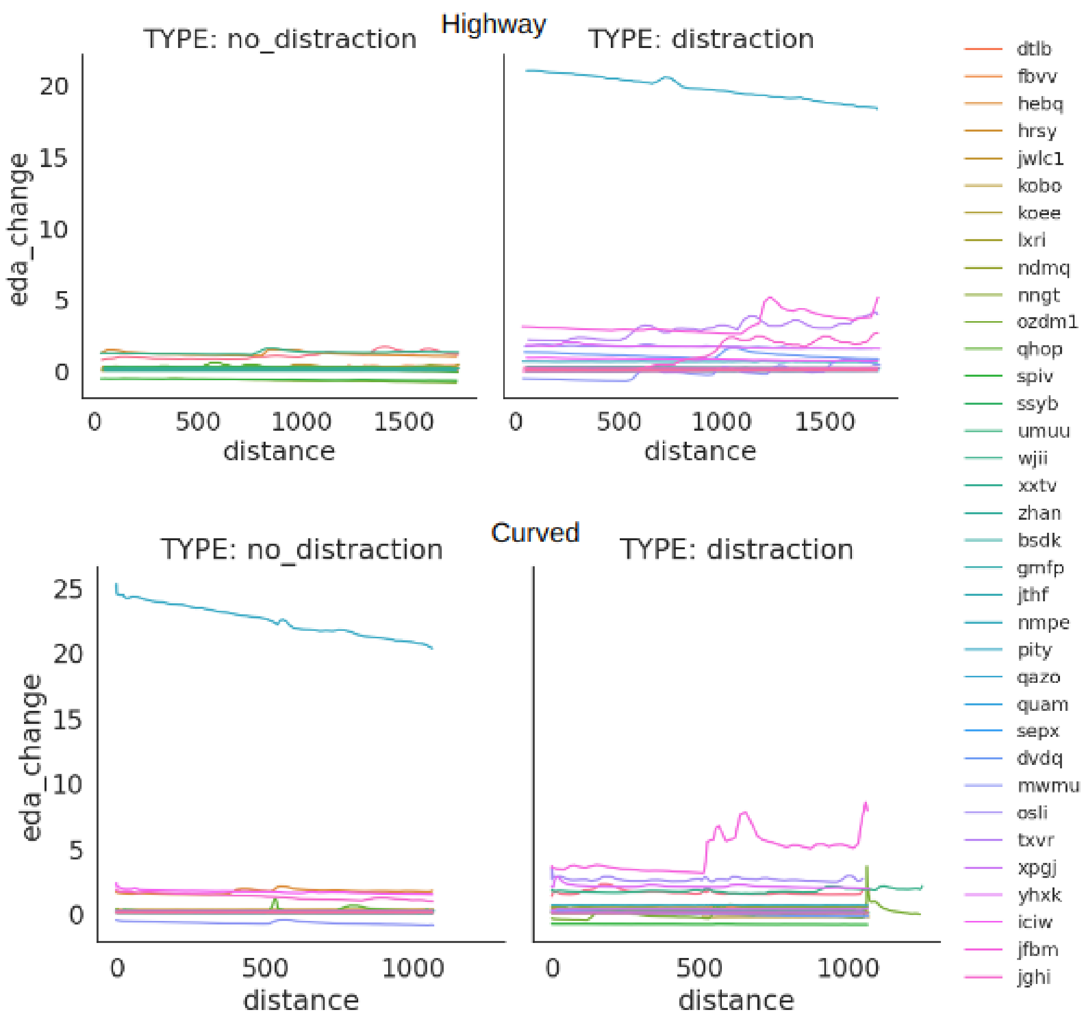


Figure 4.4: EDA for highway scenario illustrating a comparison between the treatment of a secondary task. EDA for curved scenario illustrating a comparison between the treatment of a secondary task.

Scenario	Participant:dis	Participant:non-dis	Datapoints
Highway	20	20	2406
Curved	20	20	3041
Parking zone	20	19	2410
Stop controlled	20	19	2453
School zone	19	18	1348
Pedestrian crossing	20	19	2418

Table 4.8: Data available from participants and their length after windowing step for all sources synchronised together, here “dis” stands for distraction.

Chapter 5

Data Analysis and Results

Multiple classification algorithms were utilized on different subsets of the data with a focus on obtaining a fine discriminator between distracted and non-distracted driving. Equally important was finding the features which indicate the most difference between distracted and non-distracted driving data. This chapter is organized into sections based on the source of the dataset used for analysis. The "Analysis" section contains sub-sections about splitting the dataset based on modalities (vehicle kinematics, physiological, and eye-tracking). Analyzing various ways of splitting the data modalities (group of features) can assist in finding the most effective features supporting the need for reduced data collection to reach high accuracy. The dataset was also resolved on road types, that is, all the scenarios treated as a single dataset and the scenarios being treated as separate datasets. Analysing different road types versus all road types combined could aid in observing the effect of road types.

5.1 Algorithms

The algorithms mentioned in Section 2.1.3 were trained in Python [59] utilizing the scikit-learn package [60] and trained on the datasets. The algorithms were trained on the complete dataset using cross-validation to avoid overfitting and to increase generalizability. The algorithms were then evaluated using the accuracy metric (number of correct predictions/total number of predictions) to compare their performance. The accuracy metric was chosen because of the balanced nature of the classes coupled with the high numbers for true positives and true negatives in confusion matrices. Furthermore, tree-based algorithms were used to identify features which influence the discrimination the most, particularly the ones which contain the most information about the difference between data in the two categories. Section 5.2.4 provides details about the essential features assisting the classification task.

5.2 Analysis

5.2.1 Three data modalities (Vehicle kinematics, Physiological and Eye-tracking)

In the first stage, all the scenarios and sources of data were treated together and classification was carried out using the algorithms introduced in Section 2.1.3. The dataset had 14076 data points and 40 features including driving simulator (17), eye-tracker (20) and physiological sensor (3). The dataset was shuffled and split into 80% training data and 20% test data. A 10 fold cross-validation was implemented on the training dataset to choose the best parameters for the respective algorithms. The parameters chosen are given in Table 5.1. The SVM algorithms have the parameters “c” whose lower values correspond to simpler models and higher values correspond to complex models, “gamma” corresponds to the inverse of the radius of influence of samples selected by the model as support vectors; “k” parameter in kNN represents the number of neighbours accounted for while assigning a class; in tree models “depth” corresponds to the maximum depth of the tree (to limit the tree from overfitting), “minimum samples leaf”, “minimum sample split” are also parameters for controlling the complexity of the model and correspond to the minimum number of samples required at a node and the minimum number of samples required to split respectively; “estimators” are the number of trees that the RF builds in a model.

Algorithm	Parameter
SVM Linear	c=1, gamma=0.1
SVM RBF	c=10, gamma=0.1
kNN	k=1
Decision Tree	depth=20, minimum samples leaf=0.001
Random Forest	depth=20, estimators=200, minimum sample split=0.001

Table 5.1: Hyperparameter tuning through 10-fold cross-validation.

The dataset was separated based on road types - “highway, curved, school zone, parking zone, pedestrian, stop-controlled intersection” to observe the effect of road types on driver distraction classification. The same split described previously for training and test was implemented followed by classification using the algorithms, incorporating the chosen parameters from Table 5.1. The length of each scenario is given in Table 4.8, and the features were the same as the previous section. In Figure 5.1, the accuracy of different road types and the combined dataset with respect to each algorithm is illustrated. It can be observed that the accuracies improve drastically when the data is split based on road types. It was also found that Random Forest performed the best, so the comparison for different combinations of data modalities was carried out using Random Forest and is given in Table 5.2.

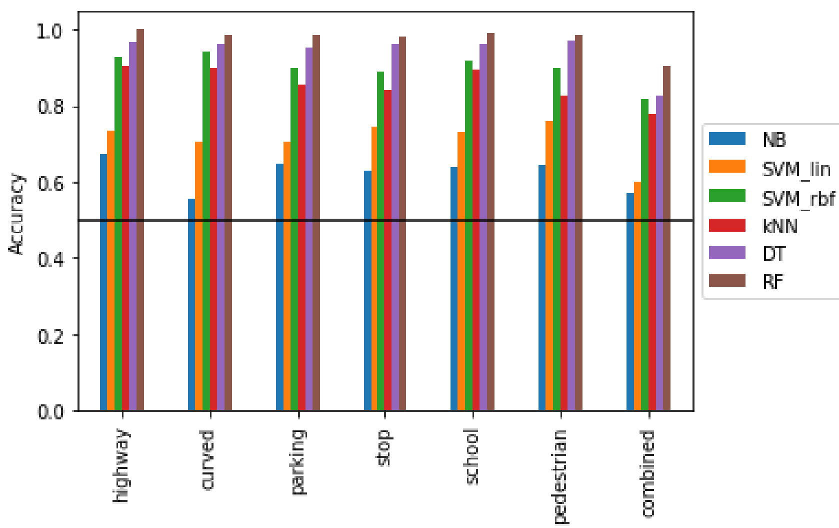


Figure 5.1: The accuracies for the scenarios and the combined set using the chosen parameters for the three modalities. The black line represents the accuracy of a random classifier.

5.2.2 Two data modalities

Different combinations of two data modalities were utilized to trim the models to simpler versions requiring less expense for data processing. The two modality combinations analyzed are - “Eye+Vehicle”, “Vehicle+Physiological”, and “Eye+Physiological”. A similar procedure as given in the previous sections was carried out for road types combined together and also separated while applying parameters given in Table 5.1. Physiological data when combined with vehicle kinematics and eye-tracking results in notable increase

Road type \ Data	All	Eye+Veh	Veh+Physio	Eye+Physio	Eye	Veh	Physio
Combined	91.08%	79.6%	91.2%	90.3%	71.9%	61.7%	96.7%
Highway	99.37%	95.9%	97.3%	98.5%	97.1%	68.09%	99.5%
Curved	98.35%	93.4%	96.7%	98.5%	93%	64.9%	99.8%
Parking	97.71%	90.8%	95.5%	99.3%	92.6%	60.4%	100%
Stop	98.98%	90.9%	96.3%	99.3%	92.1%	67.8%	99%
School	100%	91.3%	96.4%	98.8%	92%	68.7%	99.6%
Pedestrian	99.38%	92.6%	96.7%	99.3%	92.6%	62.6%	99.7%

Table 5.2: Accuracies of scenarios for different combination of modalities using RF model, here "veh" stands for vehicle kinematics features.

in accuracy.

As seen in Table 5.2, there is a slight drop in accuracy when vehicle data was added with eye-tracking data for a few road types compared to using eye-tracking data alone. The results are shown in Figure 5.2. The accuracies for the combined scenarios remained lower compared to when separated into road types, following the trend shown in Figure 5.1.

5.2.3 One data modality (Vehicle kinematics, Physiological or Eye-tracking)

One of the primary objectives of this study was to filter out dominant features, which would lead to a simpler model while maintaining high accuracy. Hence, single data modalities were tested on the algorithms to have a fair comparison of their effectiveness.

Vehicle kinematics

The algorithms were trained using data combined together and also split into different road types. Each scenario was of length given in Table 4.6 with 17 features given in Table 4.1. The data was shuffled and divided into training and test sets with the test set being 20% of the whole data. The algorithms utilized parameters given in Table 5.1 to maintain consistency for comparison. The classification accuracies for vehicle kinematics data was

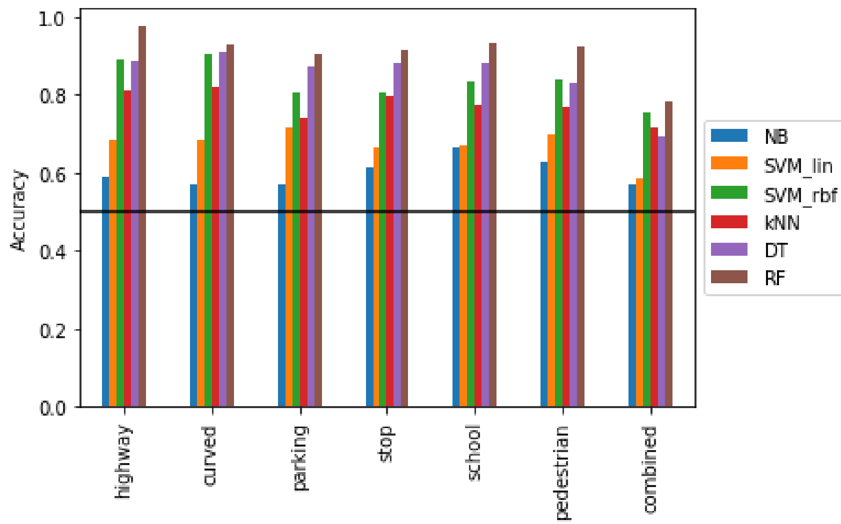


Figure 5.2: The accuracies for the scenarios and the combined set using the chosen parameters for eye tracking and vehicle kinematics incorporated together. The black line represents the accuracy of a random classifier.

in the sixties for the best performing model-RF as shown in Table 5.2 and was even lower for other models, hence omitted here.

Eye tracking data

Each scenario's data was of length given in Table 4.4 and contained 20 features from Table 4.2 and Table 4.3. It was shuffled and split into training and test sets, with test being 20% of the whole data. The algorithms were trained using the parameters given in Table 5.1. The data from different road types was combined and analyzed in a similar way to observe the effect of road types on the results. The results for the best performing model are shown in Table 5.2. The classification accuracy shows a remarkable increase compared to vehicle kinematics data.

Physiological data

A similar pipeline, as described in the above sections was followed, the scenarios were of the lengths given in Table 4.6. There were three features: EDA-change, HR-change, and temperature-change. The data was split into 80% training and 20% test, and the algorithms were trained and tested using the parameters given in Table 5.1. The data from different road types were joined together to form the combined set, trained, and tested on the algorithms as given above. The accuracies are shown in Figure 5.3; kNN, DT and RF are resulting in the best performance with accuracies reaching up to 100%. This result has been further explored at the end of the chapter.

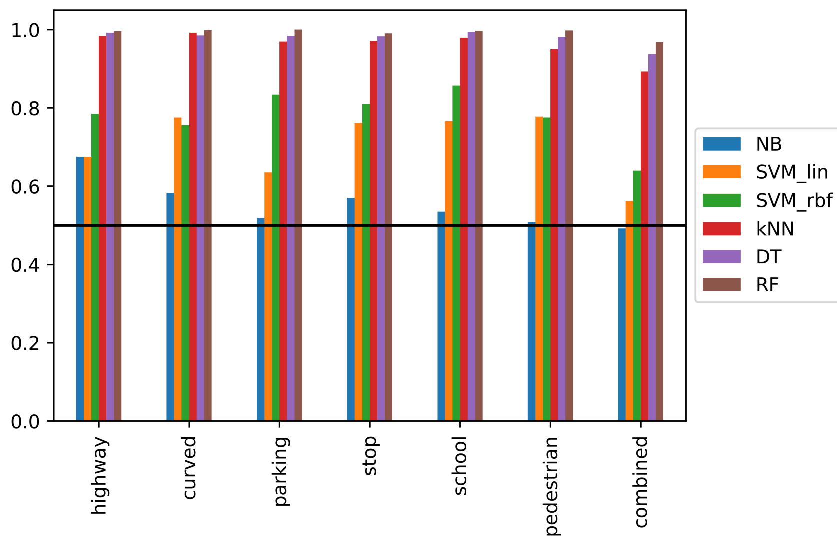


Figure 5.3: The accuracies for the scenarios and the combined set using the chosen parameters for physiological data. The black line represents the accuracy of a random classifier.

The analysis shows that the physiological modality performed the best followed by eye-tracking and then vehicle kinematics. Vehicle kinematics data seems to be noisy while physiological data is explored more later in the chapter.

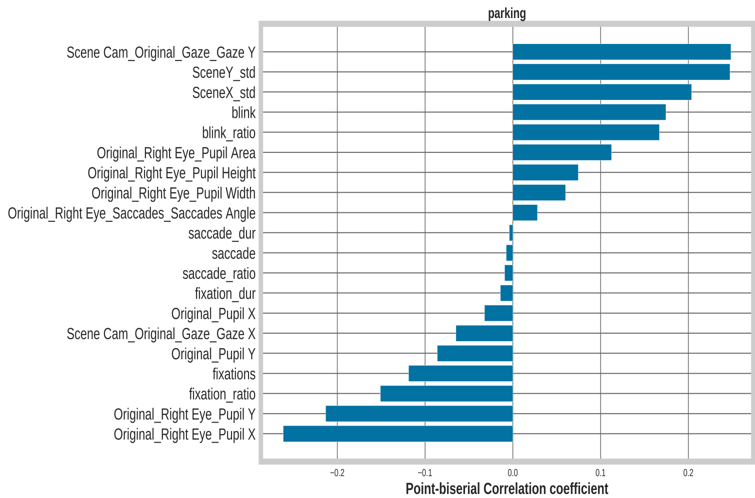
5.2.4 Feature selection

One of the key factors for feature selection is to eliminate features that are not informative, which involves selecting features that have less correlation to each other and high correlation with class labels. It helps to remove redundant features which do not contribute to the model performance. Since physiological data had only three features and vehicle kinematics data did not show much promise, the focus for feature selection was on eye-tracking data. In this study, the labels were binary hence the Point-biserial correlation coefficient was used to find the correlation between features and labels. One of the patterns that can be observed from the graphs is that features related to saccade and fixation had nearly zero correlation with the label. The correlating coefficient varies between the scenarios and only a few have been included in Figure 5.4. None of the features had a significantly high negative or positive correlation with the label.

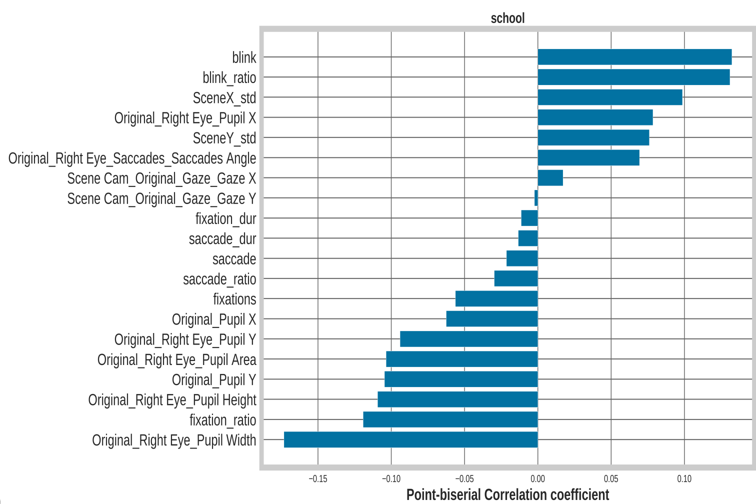
One of the methods for feature selection comes under the category of wrapper methods, where subsets of features are generated and evaluated against a classifier. The classification accuracy is used to evaluate the features, and they are optimized for the classifier utilized. In this study, there was a keen focus on tree-based algorithms because of their transparency. ExtraTrees classifier from Scikit learn was used to find the feature importance and is visualised in Figure 5.5. It can be seen that features related to spatial and size features for the eye show significant contribution while fixations and saccade have minimal relevance.

Furthermore, a Random Forest based wrapper method was utilized to identify essential features while maintaining high accuracy. Random forest classifier, along with permutation feature importance was used to select the reduced feature set. Permutation feature importance is a model inspection technique in which one feature is garbled at a time, and the decrease in classifier accuracy is observed for that action. This technique indicates the dependency of the model on a particular feature. This technique has an issue when features are strongly correlated with each other; as a result, when one feature is garbled, the other correlated feature is able to provide the necessary information to the model. This results in lower importance being given to the correlated features. This problem has been avoided by clustering correlated features and choosing one feature from each cluster for the correlation task as it is able to provide similar information to the other features in its cluster. Hierarchical clustering was performed on the features using Spearman rank-order correlation and a threshold was chosen to pick a single feature from each cluster. In Figure 5.6, the clustering illustrates some expected characteristics, that is, the features corresponding to the x-axis are clustered together, similarly for the y-axis; features corresponding to saccades were clustered together, but fixation duration also showed high correlation with them; height, width and area of the pupil were clustered together; blink and blink ratio formed clusters with sceneX std and sceneY std. This pattern was observed through all the scenarios. It was desired to have a simple model learning from a small number of informative features while maintaining the most accuracy.

The above method selected the following features having an accuracy shown in Table 5.3

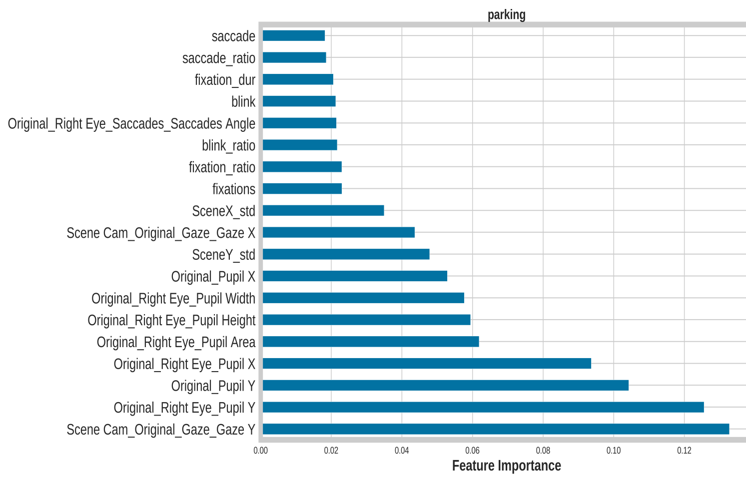


(a)

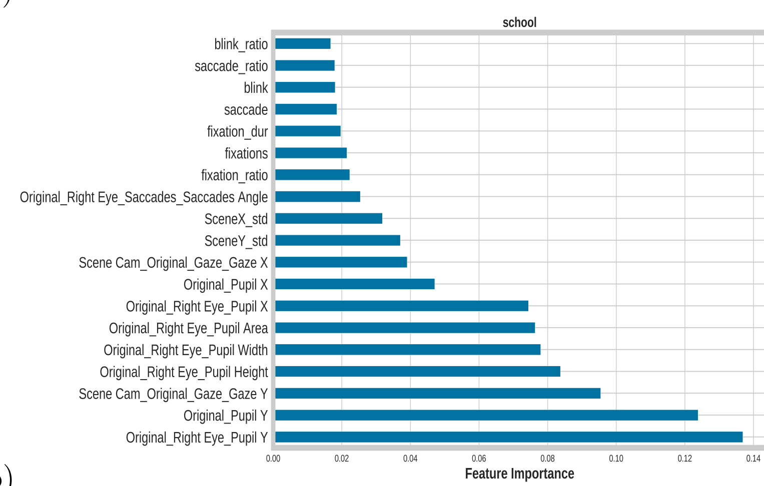


(b)

Figure 5.4: a) Correlation coefficient for features in parking zone scenario. b) Correlation coefficient for features in school zone scenario.

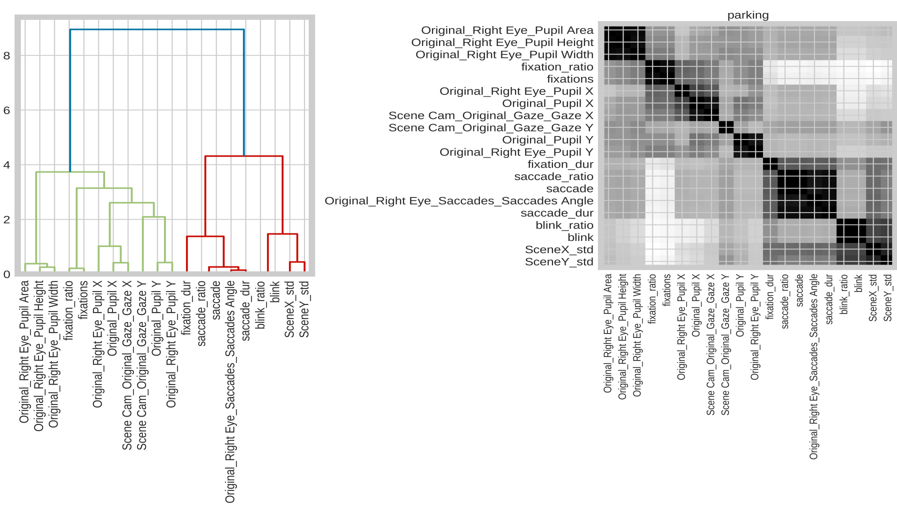


(a)

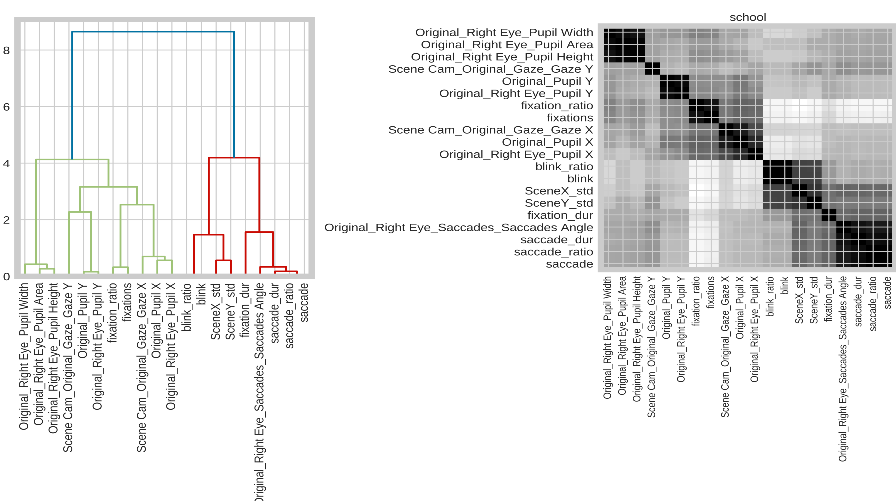


(b)

Figure 5.5: a) Feature importance using ExtraTree classifier for parking zone scenario. b) Feature importance using ExtraTree classifier for school zone scenario.



(a)



(b)

Figure 5.6: a) Hierarchical clustering of features and the corresponding correlation heatmap for parking zone scenario. b) Hierarchical clustering of features and the corresponding correlation heatmap for school zone scenario.

for respective scenarios.

- Original Pupil X
- Original Pupil Y
- Original Right Eye Pupil Area
- Scene Cam Original Gaze Y

Performing a t-test on the selected features to compare the two classes did not show any significance alluding to the univariate nature of this test. The violin plots illustrating the features comparing the two groups is shown in Figure 5.7.

Furthermore, when the "Original Pupil X" feature was dropped, and the decrease in accuracy observed was not significant, which illustrated the importance of vertical motion of pupil and gaze to detect driver distraction. The updated accuracies are given in Table 5.4, and these are a common set of features that can be used in all road types to detect driver distraction.

Scenarios	Accuracy
Highway	93%
Curved	95%
Parking	93%
Stop	91%
School	91%
Pedestrian	86%

Table 5.3: Accuracies of scenarios for reduced set of features using RF - horizontal eye motion, vertical eye motion and pupil size.

Likewise, a similar analysis for feature selection was carried out for vehicle kinematics data and the features selected were - 'velocity', 'acceleration', 'lateral velocity', 'lateral position', and 'steer'. An analysis of physiological data was not considered necessary as the features were limited to just three in number and all of them were uncorrelated to each other and hence considered to be important features.

Scenarios	Accuracy
Highway	91%
Curved	89%
Parking	89%
Stop	90%
School	94%
Pedestrian	87%

Table 5.4: Accuracies of scenarios for reduced set of features using RF consisting of vertical eye motion and pupil size.

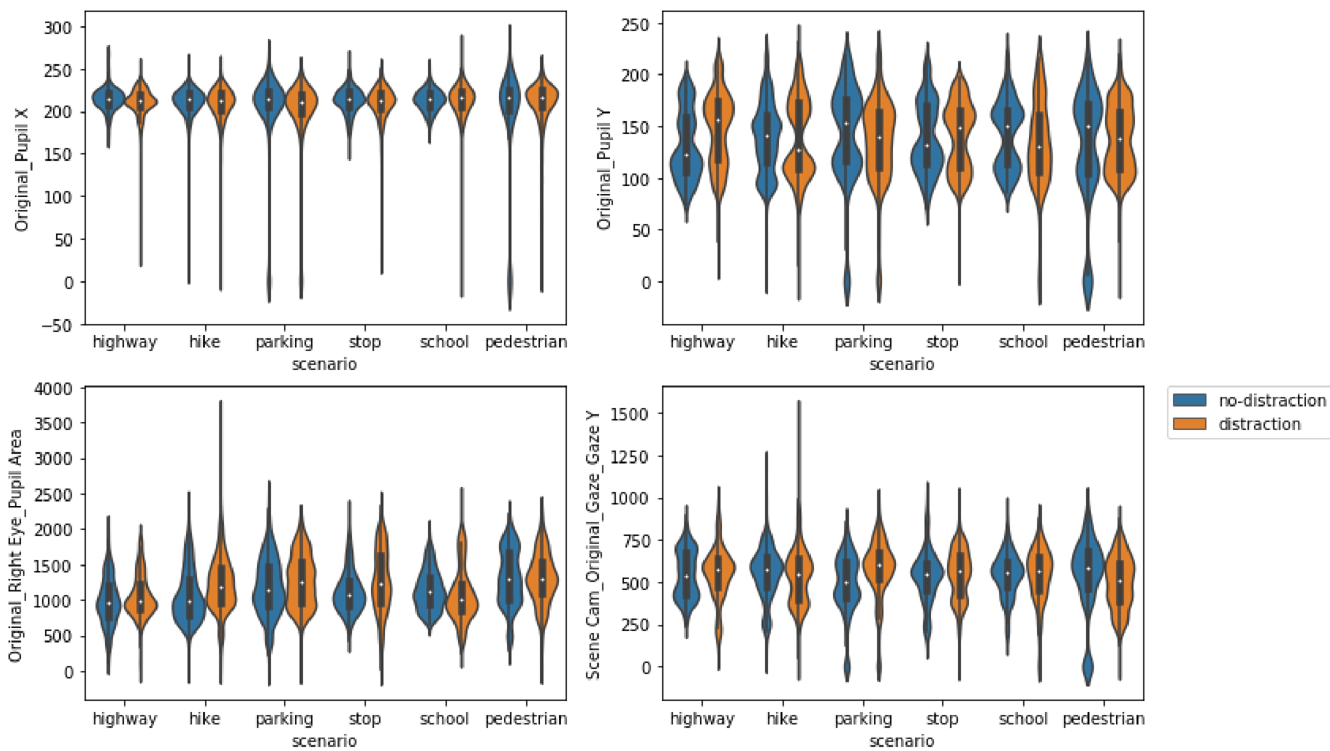


Figure 5.7: Violin plots of selected eye-tracking features comparing two classes. It represents the distribution shape of data with the white dot corresponding to the median and the thick black line representing the Interquartile range (IQR), thin black line extending to 1.5x the IQR range.

5.3 Additional physiological data analysis

The physiological data gave unprecedented accuracies which suggest an unexplained issue with the dataset, hence requiring further analysis. The physiological data was tuned to find the best parameters using 10-fold cross-validation and showed that SVM-RBF, kNN, DT and RF are giving accuracies reaching up to 100%, which indicates a non-linear boundary. NB and SVM-linear performed close to random as given in Table 5.5. However, when a t-test was performed to discriminate between the distracted and non-distracted group of drivers, none of the features showed any significance. The analysis of the data may imply that the training and the testing tasks were very similar leading to the repetition of the same task and resulting in unreasonably high accuracy.

It could also be attributed to the fact that time series physiological data is being treated as independent individual data points here, which may not be the right approach to analyse physiological data. Furthermore, the short duration of data for each participant's driving scenario might have limited the capture of change in physiological data.

Road type \ Classsifier	NB	SVM-lin	SVM-RBF	kNN	DT	RF
Combined	50.1%	51.04%	82.58%	88.27%	89.41%	94.24%
Highway	64.8%	58.59%	96.89%	98.13%	98.34%	99.79%
Hike	59.77%	56.97%	98.19%	98.85%	98.85%	99.67%
Parking	52.53%	60.24%	98.58%	96.34%	99.39%	99.39%
Stop	59.92%	68.2%	97.49%	97.88%	98.07%	99.42%
School	50.69%	57.69%	94.05%	95.8%	97.55%	99.65%
Pedestrian	55.66%	62.34%	95.74%	94.33%	97.77%	99.19%

Table 5.5: The accuracies for the scenarios and the combined set using re-tuned parameters for physiological data.

Chapter 6

Discussion and Conclusion

6.1 Discussion

This study brought into focus the effect of road types on distraction classification. All the scenarios combined into a single dataset gave a much lower accuracy than treating them separately, as seen in Table 5.2, which is especially evident in the case of eye-tracking data. This finding is supported by [37] which shows that the addition of driving context from outside generates an improved driver distraction monitoring system. The complexity of the driving environment has varied effects on driving; factors like urban driving, highway driving, traffic density, and speed limits have an impact on driver behavior and should be accounted for [61]. A system for driver distraction detection can include GPS data providing the location of the car and the road type in addition to the driver's behavioral data to generate more accurate predictions.

An analysis comparing different combinations of modalities was performed to identify the feature set contributing the most information towards the classification task. To have a fair comparison, the same algorithms were trained for all datasets. The results showed that the addition of physiological and eye-tracking data to the dataset resulted in a notable increase in accuracy.

Accordingly, the study emphasised the importance of eye-data to enhance driver distraction identification due to cognitive workload. The dominant features identified illustrated the importance of pupil movement data and gaze dispersion in both horizontal and vertical directions. It also highlighted the need to include pupil size measures. The effect of cognitive tasks on gaze dispersion was shown to be statistically significant in [43],[62]. Also, [62] shows that changes in visual attention can act as an early indicator of driver distraction even before vehicle control is affected, which was observed in this study as well. [62] also does a statistical analysis showing a significant change in the mean central location of vertical gaze between driving with 0-back task and pre-task baseline driving, corroborating the significance of vertical pupil measure seen here. It was also seen here that features related to saccadic movements had no effect in classification, which may be because saccadic measures are not an indicator of a driver's ability to acquire visual information or interference of it due to cognitive distraction. Furthermore, this study used a method for feature selection which was not affected by multicollinearity; it identified the primary features necessary to achieve good accuracy. Although the eye-trackers used in this study were wearables, the advent of wireless eye-trackers that can be installed on the dashboard will make the transition to an eye-tracking based driver-assistive system smoother.

The inclusion of physiological data showed significant improvement in accuracy for all combinations of sources in Table 5.2, notably, the combination of eye-tracking data with physiological data. This confirmed the relevance of physiological data such as EDA, HR and temperature as good indicators to differentiate distracted and non-distracted driving behaviour. It could also be attributed to the similarity between training and test datasets due to the short duration of data length for each participant, hence not having much variation between the training and test datasets. Further studies are required to confirm the applicability of the physiological data. It could be seen that SVM-RBF, kNN, decision tree, and random forest performed the best while NB, SVM-linear performed almost at random as shown in Table 5.5. It may be because the physiological data had only three features (small feature set) and was treated as independent data points even though it belonged to a time series. The addition of physiological data into other modalities increased accuracies substantially as the ML models concentrated on the physiological features while classifying as was seen from their attributed feature importance. On the other hand, it is shown in [63] that physiological features such as respiration, electrocardiogram, skin conductance, and body temperature have given 100% accuracy for identifying drowsiness and stress, which lead to cognitive distraction. In [63], for most drivers studied, skin conductivity and HR are most closely related to driver stress levels; also, significant individual differences in the mean of the skin conductance was found. These findings suggest that a driver distraction detection could be made more robust by personalizing it to the respective driver's physiological metrics. With the continual increase in sensor capability and the decrease in costs, the inclusion of physiological sensors in vehicles may become widespread and allow systems to draw on these sensors to improve driver-assistive technology. A major draw-back of these systems, however, could be the intrusiveness of the devices used to carry out the measurements compromising driver privacy and introducing the inconvenience of an additional physical device.

The vehicle kinematics data did not show much promise because of the limited fidelity of the driving simulator and the distraction task not being distracting enough. It may also be explained by the Cognitive Control Hypothesis, which states that cognitive load leaves automatic performance unaffected. If the driving task is too simple, it might be automatized, leading to no observed effects due to secondary tasks [64]. Notwithstanding, the following vehicle kinematics features were deemed the most important: 'velocity', 'acceleration', 'lateral position', and 'steer'. Results from [65] confirm that distraction has a significant influence on lateral vehicle control, driving speed, and steering wheel angle. Drivers try to compensate for limited cognitive resources by altering the respective driving features.

Overall comparison between the performance of the six algorithms: NB, SVM linear, SVM-RBF, kNN, DT, RF showed that RF achieved the best accuracies for all datasets attributing to its ensemble of classifiers. NB performed the worst because of its assumption of samples being independent, which was not the case for the dataset. In this study, there was a significant emphasis on tree classifiers because of their interpretable nature and simplicity to understand, which is in direct contrast to Neural Networks' black box nature, and which were not utilized in this study.

6.1.1 Contribution

- In the future, vehicles will be required to be intelligent and more responsive to the user. Early indicators of driver distraction can help in changing information output by IVIS to enable safer driving. In this study, these indicators were obtained from three sources of driver data: Physiological, Eye-tracking, Vehicle kinematics. None of the earlier studies included a comparative analysis of all three sources.

- It was shown that road types such as urban, highway, and sub-urban have an effect on driver behaviour and can influence the impact of workload on driving, as shown from the comparison of datasets split based on road types and combined. None of the earlier studies focused on showing the effect of road types on driver behavioral data.

- The reduced feature set was found while maintaining high accuracy, showing the feasibility of a simpler model for driver distraction detection. Earlier studies did try finding important features but did not show the quantitative effect after feature reduction.

- Among the ML models used, RF was found to be the most accurate. The reduced feature set and information about the current road type utilized on RF can lead to a simpler and more accurate system for driver distraction prediction. [55] also showed that RF performed best across the driving and physiological datasets as compared to other algorithms.

6.1.2 Limitations

Being a lab-based study of short length, it does not provide a comparative environment to expand the results to real-world, uncontrolled situations. The context of monitored experimentation likely affects drivers' performance along with the use of a driving simulator and wearable sensors, as opposed to more naturalistic driving conditions. The window size was limited to 1 second because of the short duration of observations and it may be possible that it is too short a duration to observe significant changes. To develop a personalized driver-based classifier as well as obtain robust physiological data, it would be necessary to have a longer duration of driving on different road types. In further studies, more extended duration scenarios for different road types should be utilized and windows of varied length used for the analysis to account for long-term effects. The models in this study assumed that drivers had a similar structural response to distraction and were treated together; in the future, a longer duration study could help train personalized driver models. A more realistic driving simulator would also contribute positively to data quality. The other limitation was the sample not being completely randomized as the study was advertised and conducted in a University setting. It restricted the participant pool to University students and could only approximate a completely randomized control trial.

6.2 Conclusion

This study was conducted with young drivers, and the results demonstrate their reaction to distracted driving in comparison with normal driving without any secondary task. The study accounted for various road types and consisted of a controlled communication task similar to real-life situations as a secondary task. Various machine learning techniques were utilized for cognitive distraction classification and the results suggest that physiological and eye-tracking data are good indicators of driver behaviour. Conversely, vehicle kinematics data did not contribute to the classification between distracted and non distracted driving, which may be attributed to the limited fidelity of the driving simulator or the limitation of vehicle kinematics data for predicting distraction. The study showed that a communication task replicating the workload of a conversation with a co-passenger can cause changes in driver state in most aspects. These changes may be detected using the relevant features found in the study for early mitigation and development of effective warning systems.

References

[1] Sheila G Klauer, Feng Guo, Bruce G Simons-Morton, Marie Claude Ouimet, Suzanne E Lee, and Thomas A Dingus. Distracted driving and risk of road crashes among novice and experienced drivers. *New England journal of medicine*, 370(1):54–59, 2014.

[2] Santokh Singh. Distracted driving and driver, roadway, and environmental factors. Technical report, 2010.

[3] Fernando A Wilson and Jim P Stimpson. Trends in fatalities from distracted driving in the united states, 1999 to 2008. *American journal of public health*, 100(11):2213–2219, 2010.

[4] Paul Schroeder, Melanie Wilbur, Reyna Pena, SRBI Abt, et al. National survey on distracted driving attitudes and behaviors-2015. Technical report, United States. National Highway Traffic Safety Administration, 2018.

[5] Natasha Merat, A. Hamish Jamson, Frank C.H. Lai, Michael Daly, and Oliver M.J. Carsten. Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27(PB), 2014.

[6] Christopher D Wickens, Justin G Hollands, Simon Banbury, and Raja Parasuraman. *Engineering psychology and human performance*. Psychology Press, 2015.

[7] Yanchao Dong, Zhencheng Hu, Keiichi Uchimura, and Nobuki Murayama. Driver inattention monitoring system for intelligent vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems*, 12(2):596–614, 2011.

[8] David L. Strayer, Jason M. Watson, and Frank A. Drews. *Cognitive Distraction While Multitasking in the Automobile*, volume 54. Elsevier Inc., 1 edition, 2011.

[9] European Transport Safety Council Etsc. *The Role of Driver Fatigue in Commercial Road Transport Crashes*. 2001.

[10] Ivan D. Brown. Driver Fatigue. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 36(2):298–314, 1994.

[11] Ping Huang Ting, Jiun Ren Hwang, Ji Liang Doong, and Ming Chang Jeng. Driver fatigue and highway driving: A simulator study. *Physiology and Behavior*, 94(3):448–453, 2008.

[12] Christoph Brüser, Christoph Hoog Antink, Tobias Wartzek, Marian Walter, and Steffen Leonhardt. Ambient and unobtrusive cardiorespiratory monitoring techniques. *IEEE Reviews in Biomedical Engineering*, 8:30–43, 2015.

[13] Rosalind W Picard, Elias Vyzas, and Jennifer Healey. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. 23(10):1175–1191, 2001.

[14] Stephan Heuer, Bhavin Chamadiya, Adnene Gharbi, Christophe Kunze, and Manfred Wagner. Unobtrusive in-vehicle biosignal instrumentation for advanced driver assistance and active safety. *Proceedings of 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES 2010*, (December):252–256, 2010.

[15] Bit Lee, Jung Kim, Yun Kim, Hyun Baek, Myung Ryu, and Kwang Park. The relationship between hrv parameters and stressful driving situation in the real road. pages 198 – 200, 12 2007.

[16] Bhavin Chamadiya, S. Heuer, U. G. Hofmann, and M. Wagner. Towards a capacitively coupled electrocardiography system for car seat integration. In Jos Vander Sloten, Pascal Verdonck, Marc Nyssen, and Jens Haueisen, editors, *4th European Conference of the International Federation for Medical and Biological Engineering*, pages 1217–1221, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[17] Bit Lee, Jong Min Choi, Jung Soo Kim, Yun Seong Kim, Hyun Jae Baek, Myung Suk Ryu, Ryang Hee Sohn, and Kwang Park. Nonintrusive biosignal measurement system in a vehicle. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2007:2303–6, 02 2007.

[18] Boon Giin Lee and Wan Young Chung. Driver alertness monitoring using fusion of facial features and bio-signals. *IEEE Sensors Journal*, 12(7):2416–2422, 2012.

[19] Li Li, Klaudius Werber, Carlos F. Calvillo, Khac Dong Dinh, Ander Guarde, and Andreas König. Multi-sensor soft-computing system for driver drowsiness detection. *Advances in Intelligent Systems and Computing*, 223:129–140, 2014.

[20] Steffen Leonhardt, Lennart Leicht, and Daniel Teichmann. Unobtrusive vital sign monitoring in automotive environments—A review. *Sensors (Switzerland)*, 18(9):1–38, 2018.

[21] Panagiotis Papantoniou, Eleonora Papadimitriou, and George Yannis. ScienceDirect-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY. ScienceDirect Review of driving performance parameters critical for distracted driving research-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY. ScienceDirect Review of driving performance parameters critical for distracted driving research-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY. *Transportation Research Procedia*, 25:1796–1805, 2017.

[22] Rayka Mohebbi, Rob Gray, and Hong Z. Tan. Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone. *Human Factors*, 2009.

[23] Lisheng Jin, Qingning Niu, Haijing Hou, Huacai Xian, Yali Wang, and Dongdong Shi. Driver cognitive distraction detection using driving performance measures. *Discrete Dynamics in Nature and Society*, 2012, 2012.

[24] M Wollmer, C Blaschke, T Schindl, B Schuller, B Farber, S Mayer, and B Trefflich. Online Driver Distraction Detection Using Long Short-Term Memory. *Intelligent Transportation Systems, IEEE Transactions on*, 12(2):574–582, 2011.

[25] Seyed Mehdi Iranmanesh, Hossein Nourkhiz Mahjoub, Hadi Kazemi, and Yaser P. Fallah. An Adaptive Forward Collision Warning Framework Design Based on Driver Distraction. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–10, 2018.

[26] Minh Van Ly, Sujitha Martin, and Mohan M. Trivedi. Driver classification and driving style recognition using inertial sensors. *IEEE Intelligent Vehicles Symposium, Proceedings*, (Iv):1040–1045, 2013.

[27] Iván García Daza, Luis Miguel Bergasa, Sebastián Bronte, Jose Javier Yebes, Javier Almazán, and Roberto Arroyo. Fusion of optimized indicators from advanced driver

assistance systems (ADAS) for driver drowsiness detection. *Sensors (Switzerland)*, 14(1):1106–1131, 2014.

[28] Céline Craye and Fakhri Karray. Driver distraction detection and recognition using RGB-D sensor. pages 1–11, 2015.

[29] Yulan Liang, Michelle L. Reyes, and John D. Lee. Real-time detection of driver cognitive distraction using support vector machines. *IEEE Transactions on Intelligent Transportation Systems*, 8(2):340–350, 2007.

[30] Tianchi Liu, Yan Yang, Guang Bin Huang, Yong Kiang Yeo, and Zhiping Lin. Driver Distraction Detection Using Semi-Supervised Machine Learning. *IEEE Transactions on Intelligent Transportation Systems*, 17(4), 2016.

[31] VALÉRIE BÉCAERT. Supervised machine learning in action. https://www.elementai.com/news/2019/unsupervised-vs-supervised-machine-learning#:~:text=Supervised%20learning%20has%20many%20applications,text%20that%20have%20been%20translated., 2019.

[32] Céline Craye and Fakhri Karray. Driver distraction detection and recognition using rgb-d sensor. *arXiv preprint arXiv:1502.00250*, 2015.

[33] Hesham M Eraqi, Yehya Abouelnaga, Mohamed H Saad, and Mohamed N Moustafa. Driver distraction identification with an ensemble of convolutional neural networks. *Journal of Advanced Transportation*, 2019, 2019.

[34] Yulan Liang, Michelle L Reyes, and John D Lee. Real-time detection of driver cognitive distraction using support vector machines. *IEEE transactions on intelligent transportation systems*, 8(2):340–350, 2007.

[35] MH Kutila, Maria Jokela, Tapani Mäkinen, Jouko Viitanen, Gustav Markkula, and TW Victor. Driver cognitive distraction detection: Feature estimation and implementation. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 221(9):1027–1040, 2007.

[36] Ali Darzi, Sherif M Gaweesh, Mohamed M Ahmed, and Domen Novak. Identifying the causes of drivers' hazardous states using driver characteristics, vehicle kinematics, and physiological measurements. *Frontiers in neuroscience*, 12:568, 2018.

[37] Yuan Liao, Shengbo Eben Li, Wenjun Wang, Ying Wang, Guofa Li, and Bo Cheng. Detection of driver cognitive distraction: A comparison study of stop-controlled intersection and speed-limited highway. *IEEE Transactions on Intelligent Transportation Systems*, 17(6):1628–1637, 2016.

[38] Li Li, Klaudius Werber, Carlos F Calvillo, Khac Dong Dinh, Ander Guarde, and Andreas König. Multi-sensor soft-computing system for driver drowsiness detection. In *Soft computing in industrial applications*, pages 129–140. Springer, 2014.

[39] Amardeep Sathyanarayana, Sandhya Nageswaren, Hassan Ghasemzadeh, Roozbeh Jafari, and John HL Hansen. Body sensor networks for driver distraction identification. In *2008 IEEE international conference on vehicular electronics and safety*, pages 120–125. IEEE, 2008.

[40] Seong-Whan Lee and Alessandro Verri. *Pattern Recognition with Support Vector Machines: First International Workshop, SVM 2002, Niagara Falls, Canada, August 10, 2002. Proceedings*, volume 2388. Springer, 2003.

[41] S Rasoul Safavian and David Landgrebe. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674, 1991.

[42] Pradipta Biswas and Gowdham Prabhakar. Detecting drivers' cognitive load from saccadic intrusion. *Transportation research part F: traffic psychology and behaviour*, 54:63–78, 2018.

[43] Ying Wang, Bryan Reimer, Jonathan Dobres, and Bruce Mehler. The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation research part F: traffic psychology and behaviour*, 26:227–237, 2014.

[44] Joanne L Harbluk, Y Ian Noy, Patricia L Trbovich, and Moshe Eizenman. An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. *Accident Analysis & Prevention*, 39(2):372–379, 2007.

[45] Manbir Sodhi, Bryan Reimer, and Ignacio Llamazares. Glance analysis of driver eye movements to evaluate distraction. *Behavior Research Methods, Instruments, & Computers*, 34(4):529–538, 2002.

[46] Charlotte Desmet and Kevin Diependaele. An eye-tracking study on the road examining the effects of handsfree phoning on visual attention. *Transportation research part F: traffic psychology and behaviour*, 60:549–559, 2019.

[47] Sandra P Marshall. Identifying cognitive state from eye metrics. *Aviation, space, and environmental medicine*, 78(5):B165–B175, 2007.

[48] Yuan Liao, Shengbo Eben Li, Wenjun Wang, Ying Wang, Guofa Li, and Bo Cheng. The impact of driver cognitive distraction on vehicle performance at stop-controlled intersections. In *2015 IEEE Intelligent Vehicles Symposium (IV)*, pages 946–951. IEEE, 2015.

[49] E4 empatica. https://www.empatica.com/en-int/research/e4/.

[50] Foroogh Hajiseyedjavadi, Tingru Zhang, Ravi Agrawal, Michael Knodler, Donald Fisher, and Siby Samuel. Effectiveness of visual warnings on young drivers hazard anticipation and hazard mitigation abilities. *Accident Analysis & Prevention*, 116:41–52, 2018.

[51] Siby Samuel, Avinoam Borowsky, Shlomo Zilberstein, and Donald L Fisher. Minimum time to situation awareness in scenarios involving transfer of control from an automated driving suite. *Transportation research record*, 2602(1):115–120, 2016.

[52] Yalda Ebadi, Ganesh Pai Mangalore, and Siby Samuel. Impact of cognitive distractions on drivers' anticipation behavior in vehicle-bicycle conflict situations. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 62, pages 1898–1902. SAGE Publications Sage CA: Los Angeles, CA, 2018.

[53] Willem Vlakveld, Matthew RE Romoser, Hasmik Mehranian, Frank Diete, Alexander Pollatsek, and Donald L Fisher. Do crashes and near crashes in simulator-based training enhance novice drivers' visual search for latent hazards? *Transportation research record*, 2265(1):153–160, 2011.

[54] Alan D Baddeley. A 3 min reasoning test based on grammatical transformation. *Psychonomic science*, 10(10):341–342, 1968.

[55] Anthony D McDonald, Thomas K Ferris, and Tyler A Wiener. Classification of driver distraction: A comprehensive analysis of feature generation, machine learning, and input measures. *Human factors*, page 0018720819856454, 2019.

[56] Yuan Liao, Shengbo Eben Li, Wenjun Wang, Ying Wang, Guofa Li, and Bo Cheng. The impact of driver cognitive distraction on vehicle performance at stop-controlled intersections. *IEEE Intelligent Vehicles Symposium, Proceedings*, 2015-Augus(Iv):946–951, 2015.

[57] Dikablis glasses. http://www.jalimedical.com/dikablis-glasses-eye-tracker.html.

[58] Miguel Ángel Recarte, Elisa Pérez, Ángela Conchillo, and Luis Miguel Nunes. Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish journal of psychology*, 11(2):374–385, 2008.

[59] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.

[60] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[61] Yung-Ching Liu and Tsun-Ju Wu. Fatigued driver's driving behavior and cognitive task performance: Effects of road environments and road environment changes. *Safety Science*, 47(8):1083–1089, 2009.

[62] Bryan Reimer. Impact of Cognitive Task Complexity on Drivers' Visual Tunneling. *Transportation Research Record: Journal of the Transportation Research Board*, 2138:13–19, 2009.

[63] Jennifer A Healey and Rosalind W Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on intelligent transportation systems*, 6(2):156–166, 2005.

[64] Johan Engström, Gustav Markkula, Trent Victor, and Natasha Merat. Effects of cognitive load on driving performance: The cognitive control hypothesis. *Human factors*, 59(5):734–764, 2017.

[65] Jianwei Niu, Xiai Wang, Xingguo Liu, Dan Wang, Hua Qin, and Yunhong Zhang. Effects of mobile phone use on driving performance in a multiresource workload scenario. *Traffic injury prevention*, 20(1):37–44, 2019.

[66] Salma Elmalaki, Huey-Ru Tsai, and Mani Srivastava. Sentio: Driver-in-the-loop forward collision warning using multisample reinforcement learning. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, SenSys '18, pages 28–40, New York, NY, USA, 2018. ACM.

[67] Aisha Siddiqa and Anis Jahan. the Relationship Between. *Printing Area International Research journal*, 002(43053):198–200, 2018.

[68] Y. C. Liu. Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveler information systems. *Ergonomics*, 44(4):425–442, 2001.

[69] Xingwei Wu, Linda Ng Boyle, and Dawn Marshall. Drivers' avoidance strategies when using a Forward Collision Warning (FCW) system. *Proceedings of the Human Factors and Ergonomics Society*, 2017-Octob(2016):1939–1943, 2017.

[70] Jianqiang Wang, Chenfei Yu, Shengbo Eben Li, and Likun Wang. A Forward Collision Warning Algorithm with Adaptation to Driver Behaviors. *IEEE Transactions on Intelligent Transportation Systems*, 17(4):1157–1167, 2016.

[71] Maria Victoria Martinez, Ines Del Campo, Javier Echanobe, and Koldo Basterretxea. Driving Behavior Signals and Machine Learning: A Personalized Driver Assistance System. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2015-Octob(September):2933–2940, 2015.

[72] Chenyi Chen, Ari Seff, Alain Kornhauser, and Jianxiong Xiao. Deepdriving: Learning affordance for direct perception in autonomous driving. In *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[73] Luis M. Bergasa, Daniel Almeria, Javier Almazan, J. Javier Yebes, and Roberto Arroyo. DriveSafe: An app for alerting inattentive drivers and scoring driving behaviors. *IEEE Intelligent Vehicles Symposium, Proceedings*, (April):240–245, 2014.

[74] Peter I.J. Wouters and John M.J. Bos. Traffic accident reduction by monitoring driver behaviour with in-car data recorders. *Accident Analysis and Prevention*, 32(5):643–650, 2000.

[75] Lex Fridman. Human-centered autonomous vehicle systems: Principles of effective shared autonomy. *CoRR*, abs/1810.01835, 2018.

[76] Uber telematics. https://eng.uber.com/telematics/.

[77] Axa drive app. https://itunes.apple.com/sk/app/axa-drive-2/id1093063218?mt=8.

[78] Carmela Troncoso. Pripayd: Privacy friendly pay-as-you-drive insurance. *WPES'07*, 2007.

APPENDICES

Appendix A

Questionnaires

The questionnaires filled by the participants before starting the experiment are given below-

Date:
Participant ID:

USER PERFORMANCE LAB
PRE-STUDY QUESTIONNAIRE

This is strictly confidential questionnaire. Only a randomly generated participant ID number, assigned by the research administrator will be on this questionnaire. No information reported by you here will be traced back to you personally in any way. **You can skip any questions you do not feel comfortable answering.**

Section1: Demographics

Gender: Male ____ Female____ Other____ Prefer not to say____

Date of birth: ____(MM)/_____(YYYY) **Age:**______

Section 2: Driving History

Approximately how old were you when you got your driver's license?
____Years____Months

About how many kms did you drive in the past week?
____Less than 50 ____Less than 100 ____100-200 ____200-300 ____300-500
____500 or more

About how many kms did you drive in the past 12 months?
___Less than 5000 ___5000 to 10,000 ____10,001-15,000 ___15,001-20,000
___More than 20,000

Do you usually wear glasses or contacts while driving?
____No ____Yes, glasses ____Yes, contacts

Is there anything related to your background or health, including any medications that might cause you to drive much better or worse than other drivers?
____Yes ____No

If yes, please describe:____

USER PERFORMANCE LAB
Driver Behavior Questionnaire

This is a ***strictly confidential*** questionnaire. Only a randomly generated participant ID number, assigned by the research administrator, will be on this questionnaire. No information reported by you here will be traced back to you personally in any way. **You can skip any questions you do not feel comfortable answering.**

Participant ID: **Date:**

There are 24 driving behavior statements and you are expected to rate them on a scale ranging **from 0 (rarely engage in this behavior) to 5 (engage in this behavior nearly all the time)** based on your own experience.

Please see the following page for the full questionnaire.

Statements	Rarely 0	1	2	3	4	Always 5
Try to pass another car that is signaling a left turn.						
Select the wrong turn lane when approaching an intersection.						
Fail to “Stop” or “Yield” at a sign, almost hitting a car that has right of way.						
Misread signs and miss your exit.						
Fail to notice pedestrians crossing when turning onto a side street.						
Drive very close to a car in front of you as a signal that they should go faster or get out of the way.						
Forget where you parked your car in a parking lot.						
When preparing to turn from a side road onto a main road, you pay too much attention to the traffic on the main road so that you nearly hit the car in front of you.						
When you backup, you hit something that you did not observe before but was there.						
Pass through an intersection even though you know that the traffic light has turned yellow and may go red.						
When making a turn, you almost hit a cyclist or pedestrian who has come up on your right side.						
Ignore speed limits late at night or very early in the morning.						
Forget that your lights are on high beam until another driver flashes his headlights at you.						
Fail to check your rear-view mirror before pulling out and changing lanes.						
Have a strong dislike of a particular type of driver, and indicate your dislike by any means that you can.						
Become impatient with a slow driver in the left lane and pass on the right.						
Underestimate the speed of an oncoming vehicle when passing.						
Switch on one thing, for example, the headlights, when you meant to switch on something else, for example, the windshield wipers.						
Brake too quickly on a slippery road, or turn your steering wheel in the wrong direction while skidding.						
You intend to drive to destination A, but you “wake up” to find yourself on the road to destination B, perhaps because B is your more usual destination.						
Drive even though you realize that your blood alcohol may be over the legal limit.						
Get involved in spontaneous, spur-of-the-moment, races with other drivers.						
Realize that you cannot clearly remember the road you were just driving on.						
You get angry at the behavior of another driver and you chase that driver so that you can give him/her a piece of your mind.						

Motion Sickness Susceptibility Questionnaire Short-form (MSSQ-Short)

1. Please State Your **Age** .... Years. *2.* Please State Your Sex (tick box) **Male** [ ] 1 **Female** [ ] 2

This questionnaire is designed to find out how susceptible to motion sickness you are, and what sorts of motion are most effective in causing that sickness. Sickness here means feeling queasy or nauseated or actually vomiting.

Your CHILDHOOD Experience Only (before 12 years of age), for each of the following types of transport or entertainment please indicate:

3. **As a CHILD (before age 12),** how often you **Felt Sick or Nauseated** (tick boxes):

	Not Applicable - Never Travelled	**Never Felt Sick**	**Rarely Felt Sick**	**Sometimes Felt Sick**	**Frequently Felt Sick**
Cars					
Buses or Coaches					
Trains					
Aircraft					
Small Boats					
Ships, e.g. Channel Ferries					
Swings in playgrounds					
Roundabouts in playgrounds					
Big Dippers, Funfair Rides					
	t	0	1	2	3

Your Experience over the LAST 10 YEARS (approximately), for each of the following types of transport or entertainment please indicate:

4. **Over the LAST 10 YEARS,** how often you **Felt Sick or Nauseated** (tick boxes):

	Not Applicable - Never Travelled	**Never Felt Sick**	**Rarely Felt Sick**	**Sometimes Felt Sick**	**Frequently Felt Sick**
Cars					
Buses or Coaches					
Trains					
Aircraft					
Small Boats					
Ships, e.g. Channel Ferries					
Swings in playgrounds					
Roundabouts in playgrounds					
Big Dippers, Funfair Rides					
	t	0	1	2	3

Scoring the MSSQ- Short

Section A (Child) (Question 3)

Score the number of types of transportation not experienced (i.e., total the number of ticks in the ‘t’ column, maximum is 9).

Total the sickness scores for each mode of transportation, i.e. the nine types from ‘cars’ to ‘big dippers’ (use the 0-3 number score key at bottom, those scores in the ‘t’ column count as zeroes).

MSA = (total sickness score child) x (9) / (9 - number of types not experienced as a child)

Note 1. Where a subject has not experienced any forms of transport a division by zero error occurs. It is not possible to estimate this subject’s motion sickness susceptibility in the absence of any relevant motion exposure.
Note 2. The Section A (Child) score can be used as a pre-morbid indicator of motion sickness susceptibility in patients with vestibular disease.

Section B (Adult) (Question 4)

Repeat as for section A but using the data from section B.

MSB = (total sickness score adult) x (9) / (9 - number of types not experienced as an adult)

Raw Score MSSQ-Short

Total the section A (Child) MSA score and the section B (Adult) MSB score to give the MSSQ-Short raw score (possible range from minimum 0 to maximum 54, the maximum being unlikely)

MSSQ raw score = MSA + MSB

Percentile Score MSSQ-Short

The raw to percentile conversions are given below in the Table of Statistics & Figure, use interpolation where necessary.

Alternatively a close approximation is given by the fitted polynomial where y is percentile; x is raw score

$y = a.x + b.x^2 + c.x^3 + d.x^4$

a = 5.1160923 b = -0.055169904
c = -0.00067784495 d = 1.0714752e-005

Table of Means and Percentile Conversion Statistics for the MSSQ-Short (n=257)

Percentiles Conversion	Raw Scores MSSQ-Short		
	Child Section A	Adult Section B	Total A+B
0	0	0	0
10	.0	.0	.8
20	2.0	1.0	3.0
30	4.0	1.3	7.0
40	5.6	2.6	9.0
50	7.0	3.7	11.3
60	9.0	6.0	14.1
70	11.0	7.0	17.9
80	13.0	9.0	21.6
90	16.0	12.0	25.9
95	20.0	15.0	30.4
100	23.6	21.0	44.6
Mean	7.75	5.11	12.90
Std. Deviation	5.94	4.84	9.90

Table note: numbers are rounded

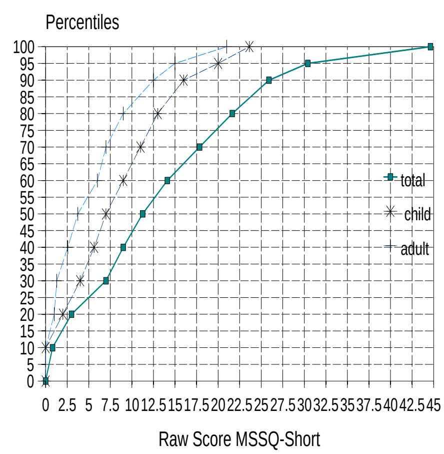


Figure: Cumulative distribution Percentiles of the Raw Scores of the MSSQ-Short (n=257 subjects).

Reference Note

For more background information and references to the original Reason & Brand MSSQ and to its revised version the ‘MSSQ-Long’, see:

Golding JF. Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. **Brain Research Bulletin**, 1998; 47: 507-516.
Golding JF. (2006) Predicting Individual Differences in Motion Sickness Susceptibility by Questionnaire. **Personality and Individual differences, 41:** 237-248.