A Context-Aware Classification System for Monitoring Driver's Distraction Levels

Adebamigbe Alex Fasanmade

Faculty of Computing, Engineering and Media (CEM)

De Montfort University(DMU)
Leicester – United Kingdom

This dissertation is submitted for the degree of

Doctor of Philosophy in Computer Science

2021

DECLARATION

I declare that the work in this thesis is original and has not been submitted in whole or part for consideration for any other degree. This work is submitted for the award of Doctor of Philosophy at Computing Engineering and Media (CEM) at De Montfort University, United Kingdom.

Adebamigbe Alex Fasanmade
2021

ACKNOWLEDGEMENTS

I am very grateful for the support of my supervisory team, Associate Prof Ali Al-Bayatti & Peter Clarke, for their guidance throughout this project. Their guidance and instructions made my experience much more productive and memorable. I am grateful to my family for their financial and moral support throughout this journey. In addition, thanks to my colleagues for their support and for not making this journey a lonely one over the years.

LIST OF RELEVANT PUBLICATIONS

1. A. Fasanmade, S. Aliyu, Y. He, A. H. Al-Bayatti, M. S. Sharif, and A. S. Alfakeeh, "Context-Aware Driver Distraction Severity Classification using LSTM Network," in *2019 International Conference on Computing, Electronics Communications Engineering (iCCECE)*, Aug. 2019, pp. 147–152, doi: 10.1109/iCCECE46942.2019.8941966
2. A. Fasanmade *et al*, "A Fuzzy-Logic Approach to Dynamic Bayesian Severity Level Classification of Driver Distraction Using Image Recognition," *IEEE Access*, vol. 8, pp. 95197–95207, 2020, doi: 10.1109/ACCESS.2020.2994811.

LIST OF ARTICLES TO BE SUBMITTED

1. A Fasanmade A, Al-Bayatti, Morden J and S, O, Aliyu MDDRA: A Novel Context-Aware Quantitative Risk Assessment Model for Degree of Drivers Distraction Image-Based label detection using Machine Learning.
2. A Fasanmade A, Al-Bayatti, Morden, J, S.O Aliyu A Multiclass Context-Aware Driver Distraction Severity Classification using a Hybrid CNN-DBN-LSTM Network.

ABSTRACT

Understanding the safety measures regarding developing self-driving futuristic cars is a concern for decision-makers, civil society, consumer groups, and manufacturers. The researchers are trying to thoroughly test and simulate various driving contexts to make these cars fully secure for road users. Including the vehicle' surroundings offer an ideal way to monitor context-aware situations and incorporate the various hazards. In this regard, different studies have analysed drivers' behaviour under different case scenarios and scrutinised the external environment to obtain a holistic view of vehicles and the environment. Studies showed that the primary cause of road accidents is driver distraction, and there is a thin line that separates the transition from careless to dangerous. While there has been a significant improvement in advanced driver assistance systems, the current measures neither detect the severity of the distraction levels nor the context-aware, which can aid in preventing accidents. Also, no compact study provides a complete model for transitioning control from the driver to the vehicle when a high degree of distraction is detected.

The current study proposes a context-aware severity model to detect safety issues related to driver's distractions, considering the physiological attributes, the activities, and context-aware situations such as environment and vehicle. Thereby, a novel three-phase Fast Recurrent Convolutional Neural Network (Fast-RCNN) architecture addresses the physiological attributes. Secondly, a novel two-tier FRCNN-LSTM framework is devised to classify the severity of driver distraction. Thirdly, a Dynamic Bayesian Network (DBN) for the prediction of driver distraction. The study further proposes the Multiclass Driver Distraction Risk Assessment (MDDRA) model, which can be adopted in a context-aware driving distraction scenario. Finally, a 3-way hybrid CNN-DBN-LSTM multiclass degree of driver distraction according to severity level is developed. In addition, a Hidden Markov Driver Distraction Severity Model (HMDDSM) for the transitioning of control from the driver to the vehicle when a high degree of distraction is detected.

This work tests and evaluates the proposed models using the multi-view TeleFOT naturalistic driving study data and the American University of Cairo dataset (AUCD). The evaluation of the developed models was performed using cross-correlation, hybrid cross-correlations, K-Folds validation. The results show that the technique effectively learns and adopts safety measures related to the severity of driver distraction. In addition, the results also show that while a driver is in a dangerous distraction state, the control can be shifted from driver to vehicle in a systematic manner.

KEYWORDS

Autonomous driving vehicles, Safety measures, Cars, Manufacturers, Simulate, Road, Traffic, Users, Environment, Context-awareness, Monitor, Technology, Careless driving, Behaviour, Distraction level, Severity, Accident, Classification, Face orientation, Diver's hands, Activities, Deep learning techniques, Fast Convolution Neural Network, Long Short-Term Memory, TELEFOT data, Bayesian, Artificial Intelligence, Computer Vision, Hidden Markov Model.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADAS	Advanced Driver Assistance System
FCNN	Fast Convolution Neural Network
LSTM	Long Short-Term Memory
NDS	Naturalistic Driving Study
AUCDDD	American University of Cairo Driver Distraction Dataset
KNN	K Nearest Neighbour
RCNN-LSTM	Recurrent Convolution Neural Networks – Long Short Terms Memory
SVM	Support Vector Machine
AI	Artificial Intelligence
AUC	Area Under the Curve
KNN	K Nearest Neighbour
RCNN-LSTM	Recurrent Convolution Neural Networks – Long Short Terms Memory
SVM	Support Vector Machine
NHTSA	National Highway Traffic Safety Administration
WLAN	Wireless Local Area Networks
CNN	Convolutional Neural Network
DBN	Dynamic Bayesian Network
NADS	National Advanced Driving Simulator
ITS	Intelligent Transport System

CHAPTER 1. INTRODUCTION

1.1 Introduction

The advent of Intelligent Transportation Systems (ITS) has not only revolutionised how safety information is gathered and shared, but it has also increased road safety. A critical requirement for improving road traffic safety is data availability, such as the causes of driver distraction and safety information on blind spots, emergency brake lights, accidents on the road, prevailing weather conditions and collision warnings [1]. ITS have greatly improved how such information is gathered and shared. Additionally, these systems aid in sharing critical vehicle information, such as signal intersections, acceleration, and the speed and direction of movement of vehicles. Although the availability of information and the adoption of ITS have greatly improved road safety, drivers still need to react to changing context-aware information on the road such as road condition, dual carriageways, urban roads, weather conditions, and other fast-moving vehicles, necessitating research into real-time context-aware systems that aid in accident prevention.

The US National Highway Traffic Safety Administration (NHTSA) cites driver distraction as one of the significant causes of road traffic accidents [2]. A major cause of driver distraction – according to NHTSA – is the presence of multiple in-vehicle electronic devices, promoting the agency to published guidelines discouraging excessive distraction [2]. Excess electronic devices, coupled with an ever-increasing amount of information presented on vehicle user interfaces, are a significant cause of distraction, occupying the driver's attention to dangerous levels that can easily cause accidents. Modern-day infotainment systems divert the driver's visual attention as they require complex operations [3]. A considerable amount of safe driving inputs is visual, while the outputs are predominantly manual activities, including feet and hand movements on the accelerator, steering wheel, and gear shift. Additionally, gazing at in-vehicle dashboard-monitors makes the driver take their eyes off the road, in most cases accompanied by removing a hand from the steering wheel to manipulate the in-vehicle display [4].

Visual input dynamically affects the perception of the driver's behaviour within a context-aware driving environment. Consequently, the impact and influence on the driver's behaviour and decision-making processes can be hugely detrimental to a driver's safety on the road. Research has shown that human beings can only hold the same concentration level for more than three hours at a time. Therefore, the European driving law obliges drivers, especially those driving HGVs, to have a rest break every three hours. This safety measure is monitored by a tracker installed in the vehicle.

Nevertheless, various studies have been carried out to understand drivers' behaviours in different environments, especially concerning existing context-aware safety systems that have demonstrated some limitations regarding a clear-cut differentiation between careless and normal driving behaviour. This research proposes an Advanced Driver Assistance System (ADAS) to capture a driver's context state and distractions to alert drivers in potentially risky situations. This gives the driver time to change their behaviour to avoid a critically dangerous situation; this encompasses the primary motivation for using this software.

This introductory part of this thesis encompasses the research background, which sheds light on the field of context-aware research. This is followed by the research motivations, which present the reasons for this study, identifying the research gaps, thereby underlining the primary deficiencies in previous research, and the research questions, wherein the principal interrogations supporting this research are asked. The research objectives present the primary orientation of this research and are followed by the research aims, highlighting the achievable milestones of this study. The contributions to the present research are presented to highlight the research achievements and delineate its limitations. Finally, the different steps of this thesis are detailed and presented for maximum clarity.

1.2 Research Background

In current society, driving is considered necessary, especially for families and commuting. However, driving introduces significant consequences, such as car crashes or other forms of accidents related to human error and people control these vehicles.

Estimates show that human error accounts for almost 94% of road traffic accidents, while 75% of accidents are attributed to the bad decisions made by the driver [5]. Studies on deaths resulting from road traffic accidents showed that 55% of deaths were careless driving. Indeed, a driver taking their gaze from the road for 5-6 seconds at a speed of 55 mph will travel the length of a football pitch [3], thus underlining how dangerous it can be for a driver to lose their focus on the road and lose control of their vehicle. Thus, a few seconds of distracted behaviour can have severe consequences.

Critically, the driver's behaviour can be significantly affected by in-vehicle devices, such as infotainment systems. Indeed, radio-sets, Compact Disc (CD) players, mobile phones connected to the car dashboard, among others, are an enormous source of distraction, causing drivers to perform actions that lead to dangerous and unacceptable driving behaviour, thereby breaching UK/EU driving laws. Therefore, detecting and monitoring driver behaviour are of paramount importance to avoid catastrophic situations. Driver behaviour detection thereby finds wide usage in designing and developing autonomous driving software and intelligent vehicle applications.

Context-aware information has a significant influence on driver reflexes, instinctive reactions and driving behaviour. This context-aware changes dynamically, and so do the perceptions of drivers and the associated driving risks and hazards. Mitigating driving risks in such an environment requires using a real-time sensing, context-aware system with the ability to detect and learn driver behaviours dynamically. Implementing such a system requires a clear definition of the context, context-aware information, and components of the context-aware application. Understanding the context will aid in selecting the context to be used in the application, subject to the naturalistic driving data available.

In ADAS, the analysis of context-aware information relating to driver distraction plays a critical role in warning or alerting the driver when the distraction level is potentially dangerous. This can be when the driver’s attention has been distracted by in-vehicle electronics. Entertainment leading to distraction plays a big part in our everyday lifestyle, and this is usually transferred in our driving environment to a smaller scale, including CD players or music and sound systems. The reality is that they are now part of the in-vehicle environment, and thus drivers must become accustomed to them and find a way to ensure they stay in control by managing the distraction level inside the car environment.

Distraction inside a car can critically affect and reduce a driver’s alertness, concentration, and reaction time. To handle driver distraction, researchers worldwide have developed intelligent systems, such as Intelligent Driver Assistance Systems (IDAS), to improve driving safety and reduce accidents. The use of IDAS to prevent road accidents is part of driver monitoring or vehicle-oriented accident prevention measures. Meanwhile, road transportation challenges such as faulty road facilities and traffic jam reduction could be monitored using ADAS.

Therefore, the development of ADAS appears to be the only effective way to handle accidents and help drivers remain focused while at the steering wheel. This succinct analysis of the research background will help to detect the essential motivation factors of this research.

1.3 Careless Driving and Dangerous Driving

The classification of careless driving behaviour cannot be carried out without defining careless driving behaviour or what normal driving behaviour entails. There is a need to define the terminologies used in various classifications of driving behaviour, such as normal, careless, dangerous, and abnormal driving. The Oxford Dictionary defines carelessness as "not giving sufficient attention or thought to avoid harm or errors" [6].

According to the United Kingdom Crown Prosecution Service (CPS), a driving offence resulting in a fatality is considered dangerous or careless driving. It further states that a driver may be engaging in dangerous driving whilst feeling that they are driving safely [7]. However, careless driving is different from dangerous driving. The CPS listed the attributes that will cause a person to be classified as engaging in dangerous driving behaviour, namely fast racing, aggressive driving, ignoring traffic lights, violating road signs, dangerous overtaking, ignoring vehicle faults, unfit driving, drowsiness, distractions (e.g., using hand-held phones, reading, infotainment system control, cigarette lighting, passenger communications) [7].

From a legal and regulatory perspective, careless driving is defined by section 3ZA of the Road Traffic Act (RTA) 1988, which states that "a person is to be regarded as driving without due care and attention if (and only if) the way he drives falls below what would be expected of a competent and careful driver" [8]. A driving offence is committed when an individual's "…driving falls below the standard expected of a competent and careful driver" or when one drives a vehicle in a public place "…without due care and attention, or reasonable consideration for other persons using the road or place" [7].

Government agency ThinkDirect states that certain driving behaviours result in distractions, such as mobile phones [9]. In addition, National Highway Traffic Safety Administration (NHTSA) and Ranney et al [1] define driver distraction as any activity that diverts a driver's attention away from the task at hand, including visual, cognitive, auditory, and other elements [10]. Furthermore, potential causes of driver distraction are other passengers, external stimuli, and in-vehicle technologies [11].

However, it should be noted that some of the attributes mentioned above can be further classified as contextually internal or external. Internal attributes influence driver behaviour within the vehicle. They can be detected by pervasive technologies, such as sensors, while the external attributes primarily affect the vehicle dynamics as detected by roadside cameras or vehicle side cameras. The taxonomy and degree of driving behaviour attributes are significant to the severity of the impacts of such driving behaviour.

According to Smith [2], the infotainment systems of a vehicle is an attack vector that could be compromised by an attacker [12]. In a context attack scenario, a third-party remote attack could passively monitor the communication channels of the vehicle and remotely control its infotainment system such that the driver is prompted to act in a way that could distract them, thus resulting in careless driving behaviour.

Both internal or external vehicle components can cause drivers to behave distracted or aggressive, leading them to be classified as careless drivers. Such components are the seat, mirror, Global Positioning System (GPS), infotainment system, window, and gearstick. There are elements of driving behaviour that should be detected and analysed before a driver can be classified as a careless driver. The list of driving behaviours is not exhaustive and is taken from the CPS and police charging standards. The types of careless driving behaviour are driving too close, inattention (lapses), fatigue, nodding off, mobile phone use, talking to passengers, failure to see traffic lights or signs, unsafe overtaking, and failure to see other vehicles or pedestrians [13]. According to driving law and CPS [7], inattention, or having more than a momentary lapse in attention, indicates careless driving, whilst anything significantly more than a momentary lapse of attention indicates dangerous driving. The degree of inattention (distraction) in any incident can be a subjective judgement. However, it could be argued that some of the behaviours mentioned above indicate careless driving behaviour, but in extreme cases could be classified as dangerous driving. Further, it could be argued that a careless driving state can change from being careless to dangerous, based on the severity level of driver distraction.

Does this raise the question of when driver behaviour can be regarded as usual, careless, or dangerous? There is a need to propose a metric for the degree of careless driving and assigning a severity level to possible incidents, thereby facilitating the development of an ADAS system based on the severity level of careless driving. Drawing knowledge from those as mentioned above and as a further contribution of this work, the most applicable definition of careless driving in the modern context of ITS is:

“Careless driving behaviour is a driving act that entails a deviation from normal driving behaviour, either by the driver actions or emanating from an entity, such as a malicious cyber attacker, pedestrian, or an environment that is influencing the driver behaviour, leading them not to give reasonable consideration to others, resulting in careless driving that can cause a casualty.”

This research proposes a novel approach to adopting naturalistic driving data to classify careless driving behaviour (driver inattention). Table 1 identifies the different event and distraction types that constitute careless driving and dangerous driving. However, one of the arguments in this work is that some of these distractions can transition from careless to dangerous driving when a degree or multiclass distraction is considered.

Table 1-1: Careless and Dangerous Driving Distractions (Ref: RTA 1988)

CARELESS DRIVING	DANGEROUS DRIVING
Driving too Close	Fast Racing
Inattention Lapses of Fatigue	Aggressive Driving
Nodding Off (Eyes Closed)	Ignoring Traffic Lights
Mobile Phone Use	Violating Road Signs
Talking to Passengers	Dangerous Overtaking
Failure to See Traffic Lights	Ignoring vehicle faults
Unsafe Overtaking	Drowsiness Eyes Closed
Failure to See other vehicle and Pedestrians	Distraction (Handheld Phone)
	Inattention Lapses

1.4 Research Motivation

Understanding and monitoring drivers' behaviour in a real-life situation and in real-time can be a game-changer in saving people's lives regarding accidents caused by human error or driver distraction. Employing systems that can fully control vehicles is seen by many as crucial to reducing human involvement and responsibility in accidents due to the latter being distracted by in-vehicle infotainment.

According to Braunagel [14], ADAS can enable the vehicle to take over lateral and longitude control to support the driver in certain situations where driving may be difficult or riskier for human drivers. Partially automation of driving exists in various vehicle models, with conditionally automated driving functions still being developed. Braunagel further stated that the responsibility of controlling a vehicle always lies with the driver, even for partially automated vehicles. Occasionally, control can be transferred to the autonomous system to allow the driver to perform secondary tasks, such as reading or communicating [14],[15].

However, passing controls to the autonomous system is not a licence to perform secondary tasks and is still closely regulated, even in fully autonomous vehicles.

The employed approach entails driver monitoring through in-vehicle monitoring systems that collect and analyse the driver's state and context-aware information. Still, situations arise where the intelligent features of a vehicle need to take over from the driver. Such situations may arise where a vehicle detects the possibility of a collision because the driver is distracted and consequently is not giving their full attention to the driving activity. Therefore, this shows the importance of measuring and classifying the driver's distraction according to severity.

Previous research on these subjects analyses the accident's impact according to the severity of the injuries sustained by the victims or the number of casualties [16],[17]. Other research has focused on communication protocols in VANET, leading to research limitations in context-awareness attacks in vehicle networks. It is essential to mention that an incident before applying a safety measure is not the best form of prevention, and it is best to eliminate the drivers' behaviour that led to that accident. The advent of autonomous vehicles has made monitoring driver activity possible via sensors and pervasive technologies, such as cameras incorporated into ADAS.

This research is from the driving behaviour perspective and the close interaction and influence of the driver's behaviour on context-aware information. However, proactive is preferable to the reaction after an accident, especially when human life is the price to pay for an incomplete or ill-designed system [18],[19]. This evaluation of the main reasons and motivating factors for engaging in this task of using context-aware systems to understand and reduce human involvement, and most importantly, reduce human error in vehicle accidents, will enable the critical gaps in the previous research to be identified.

1.5 Research Gap Identification

Several research studies have been conducted concerning autonomous vehicles, and the understanding and monitoring of drivers' behaviour to ensure their involvement in accidents is reduced. However, preliminary research has been conducted on driver distraction classification and severity level of distraction in the context-aware situation. In addition, the context-aware behaviour of the driver, when the context situation changes, is crucial, which this study uses as a pivotal framework.

Bruanagel et al [20] developed an automated recognition activity in autonomous driving scenarios for driver take-over readiness, not when the vehicle needs to take over or transition from driver to vehicle. A typical example may be losing a lane marking or reaching the end of the right road; such situations require the driver to resume the responsibility for driving. The limitation in the methodology is the focus on eye-tracking and head posture and the use of a driving simulation. The study focused on driver's inattention, which can be used in monitoring drivers. However, there are situations where there is a need to transition from driver to vehicle in a semi-autonomous vehicle. The vehicle does have a limited amount of time to take such intelligent decisions, and thus, the question is at what moment and threshold should a semi-autonomous vehicle ready to take over from the driver or transition from the driver to the vehicle; especially in ADAS?

The measurement of the degree of driver distraction can help reduce accidents and their impacts. In contrast, vehicle mobility is dynamic, and the context-awareness of a vehicle changes in real-time. Thus, there is a need for an effective and efficient intelligent safety system to predict driver behaviour based on context-awareness. Observing the gaps mentioned earlier in the literature, this work presents approaches that can further integrate and improve ADAS to solve the identified gaps in ITS. Therefore, in line with the above analysis, the following gaps have been detected in the literature.

The first gap in the literature is related to the fact that an inadequate model or framework was used to differentiate a driver's careless behaviour. Most driving behaviour research has been carried out using simulation, such as VEINS, measuring different parameters that include braking events and traffic flow [2], [25], [26]. To address the issues presented above, we have analysed and developed algorithms that satisfy both types of behaviour by using naturalistic driving data from TeleFOT vehicle behaviour, which is developed using algorithms for different attack types and context scenarios; Furthermore, video image coding and analysis frames detect drivers' behaviour based on Deep Learning (DL) CNN.

The second gap in the literature concerns the lack of a system that can classify a driver's distraction into careless, dangerous, and safe driving behaviour. The approach used to address this gap is to review the literature on Artificial Intelligence (AI) and machine learning (ML) techniques. The combinations of ML techniques, such as DL CNN, Support Vector Machines (SVM), and Naive Bayesian classifiers, are applied. The systems are developed using the Python programming language. Further research can help in this direction in that the identification and justification of AI or ML techniques can be used to detect and identify both types of behaviour.

The third gap discovered in the literature relates to the lack of a system that can simulate driver behaviour and measure driver distraction [6], [7]. To solve this, an extensive literature review and identification of the parameters that satisfy the requirement of measuring driver distraction; and developing an algorithm that can predict and prevent strange behaviour.

The fourth gap is related to limited and context-aware driver behaviour [8], [9]. To address this gap in research, the following approach is adopted concerning the development of algorithms that consider context-aware. Further research appears necessary to develop safety issue scenarios covering the broader context-aware vehicular environment, thus considering dual carriageways, urban roads, weather conditions, and other factors.

The fifth gap in this study is related to the lack of optimisation techniques to analyse the degree of driver distraction or reaction. The following approach can be used to tackle the fifth gap identified in the literature: a heuristic analysis and pattern recognition analysis in a system that monitors the driver's behaviour in real-time, considering the intent, reaction level, and engagement with any source of distraction. Finally, concerning further research, the following suggestion is made optimisation algorithms that can enable efficient decision-making by automated vehicles. In addition, the integration of an optimised system with safety control will ensure a more accurate prediction and more effective prevention of accidents.

The sixth gap, referring to Bruanagel et al [10], relates to developing an automated recognition activity during autonomous driving scenarios for driver take-over readiness, not when the vehicle needs to take over. Thus, it is essential to develop an ADAS system that enables the vehicle to take over from the driver when a certain degree of driver distraction severity level has been reached in a semi-autonomous vehicle.

The seventh gap is that UK CPS classifies some distractions as either careless or dangerous. However, it could be argued that while some of these distractions are classified as careless, when a multiclass distraction is considered, they can quickly become dangerous [11], meaning the degree of classification can change when multiclass distraction is considered.

This detailed presentation of the gaps in the literature will lead to the generation of the research questions presented below.

1.6 Research Questions

The theory challenges our assumptions in an essential and even significant way for understanding a phenomenon [12]. Nevertheless, forming research questions is usually based on the identification of gaps in the literature. For the current research, this raises a certain number of questions concerning the classification of drivers' behaviour; these are formulated as follows:

This research seeks to address the following questions:

RQ1: Can we develop algorithms to estimate careless driver behaviour/dangerous distractions and design a mathematical model for measuring the degree of driver distraction?

RQ2: What threshold is safe for different severity levels of driver distraction; on what basis the severity levels should be assigned?

RQ3: How do we implement a multiclass risk assessment model with a safety framework for the degree of driving distraction?

1.7 Research Aims and Objectives

This thesis aims to develop a robust context-aware safety mechanism for the detection and classification of driver's distraction into severity levels. The research objectives are intimately related to the research aims, given that the latter provides a more comprehensive view/indication/definition of the former. Therefore, the research objectives are further detailed to clarify the smaller achievable tasks during this study. The current research objectives are defined as follows:

1. To develop algorithms to estimate careless driver behaviour/dangerous distractions.
2. To develop a metric and mathematical model for measuring the degree of driver distraction.
3. To define a threshold for different severity levels of driver distraction.
4. To develop a novel multiclass risk assessment model for the degree of driving distraction.
5. To design and develop a safety framework based on the severity level of driver distraction.

Tackling the different highlighted objectives will enable the researcher to generate contributions during this study, advancing the body of research on self-driving cars.

1.8 Main Contributions

This research aims to develop a context-aware system to monitor driving behaviour by classifying it into careless or normal driving behaviour. After this study, it becomes apparent that the following contributions were made to the scope of the literature in this area:

- The development of a novel ADAS framework that classifies driver distractions into severity levels to aid vehicle take-over.
- The development and evaluation of a mathematical model that classifies driver behaviour according to severity levels using thresholds.
- The definition of a threshold safety system classifies driver behaviour into careless and dangerous driving, thus enabling an autonomous vehicle to take over from the driver or know when it is safe to return autonomy to the driver.
- A novel MDDRA risk assessment model for the classification of driver distractions using ML algorithms.
- The development of a novel 3-phase parallel Fast-CNN architecture to address each physiological attribute.
- The development of a context-aware situation and using output from the parallel FCNN via a novel three-tier FCNN-DBN-LSTM that detects and classify driver's distraction into the severity level of distractions.
- The development of a Fuzzy-Logic-DDBN model for the classification of driver distraction.
- Development of a Hidden Markov Driver Distraction Severity Model (HMDDSM) for classification of driver distractions.

The main contributions of this research are achieved based on a series of studies carried out in this thesis; these are presented in the following, with a succinct description of the content of each chapter.

1.9 Thesis Organisation

The present thesis, dealing with a context-aware safety system for detecting and improving dangerous driver behaviours, encompasses nine chapters and is built around a clear and logical structure that coordinates and links these chapters. The organisation presented below thus supports the thesis.

Chapter one is the opening chapter and deals with the research background related to the study of context-aware safety systems, discussing, in general, the detection and classification of driver behaviour with regards to the level of danger it can carry. The research motivations and questions are addressed, wherein the researcher tries to explain the reasons for this research. This is followed by pertinent questions that have arisen during this study and helped the researcher guide this work. In addition, by inference, the research aims, and objectives are deducted from the research questions, which are a more detailed form of the research aims. Moreover, the research's contributions that encompass the prominent achievements made during this research are likely to be included in the literature in this research area. Finally, the research organisation shows how the different chapters connect to create a coherent narrative for the reader.

Chapter two deals with the literature review, giving a more comprehensive view of this study and the literature related to the context-aware identification of driver behaviour, especially distinguishing normal behaviours from careless ones. This chapter concerns three main domains: ITS, computer vision, and DL areas of research. Regarding the ITS, the drivers' behaviours are scrutinised via behaviour profiling and through simulation, such as using the National Advanced Driving Simulator (NADS). The involvement of AI and ML techniques is necessary to distinguish normal from dangerous driving. The detection of cognitive drivers' behaviour is performed using facial geometry-based eye region detection, enabling the inference of the driver's concentration and risk level. Finally, the relationship and impact of the detection and classification of drivers' behaviour on autonomous vehicles are scrutinised, whereby the design and implementation of context-aware systems are also closely monitored, and their importance in previous research is established.

Chapter three carefully investigates the methodology adopted in the present study. In this regard, a quantitative data analysis approach is applied in a time-series dataset, images, and videos, monitoring driver behaviour through these recorded data. The chapter also highlights the main algorithms, techniques and models supporting this research. These mainly revolve around algorithms, such as Convolutional Neural Network (CNN) as the DL algorithm, Fuzzy-logic, Hidden Markov Model, DBN as the AI algorithm, and Long Short-Term Memory (LSTM) as the computer vision algorithm, which is related to the analysis of driver behaviour. The choice of the research method is also justified and is developed around two methods, namely simulation and Field Operational Testing (FOT). Indeed, one of the primary datasets used in this study is TeleFOT, one of the most extensive European datasets related to monitoring and improving autonomous and cooperative systems in the ITS context.

Chapter four, aims to tackle the main reasons for carrying out this study. It starts by succinctly presenting the main algorithms involved in this study and their usage. This includes selecting regions of interest (RoIs) that will be used for image recognition and classification, DBN is then used to model the dynamic state of the driver's behaviour in a context-aware system, and finally, the LSTM for the classification and predictive analysis. Therefore, three main aims are identified in this research study and are highlighted to facilitate the definition of the research objectives. However, the extension of this chapter is in chapter 7.

Chapter five is the research questions chapter and focuses on the main questions to define the present study's direction. Indeed, the question formulation establishes the primary orientation of the research to ensure more rationality and clarity in how the study is conducted. With regards to the current study, different questions are formulated and addressed as per the following points: i) the possibility of classification of driver distraction into safe, careless, or dangerous, ii) the possibility of distinguishing between these two distractions, iii) and finally the use of mathematical models and technology, including algorithms, to implement the models to understand and monitor drivers' behaviour. This chapter involves the use of the Mamdani Fuzzy-Logic Dynamic Bayesian Network model.

Chapter six presents a novel-risk assessment model Multiclass Driver Distraction Risk Assessment (MDDRA) model for the driver's distractions. This chapter involves a risk assessment model that covers context-aware in-vehicle, vehicle and environment parameters. The metrics to measure the driver distraction level and a safety framework based on the distraction severity level are also highlighted. Finally, an optimisation solution for the degree of distraction, helping the in-vehicle decision-making process, was also developed. The MDDRA involves applying the ML technique to classify driver's distractions. Further, this also entails validation using the cross-correlation test and Kruskal Wallis test.

Chapter seven is the hallmark of the research as it introduces a novel context-aware model called the Hidden Markov Driver Distraction Severity Model (HMDDSM), which integrates the MDDRA developed in chapter 6 into the detection and classification. Here is the introduction of a hybrid CNN-LSTM-DBN model to detect and classify the driver's distraction. Adopting Fast-Recurrent Neural Network (Fast-RCNN) from a pre-trained network Resnet in detecting context-aware constitutes distractions.

Chapter eight is the chapter that deals with evaluation and comparison with works of other authors. Here there is very significant performance in the developed algorithms compared with other works. The results in this research compared with other works are promising. A critical analysis and reflection, and justification of a few instances where other works outperformed our model inaccuracy due to the few parameters used in their works.

Chapter nine, as the closing chapter, presents the contribution of the current thesis. Several contributions were made, focusing on the main algorithms used to improve the driver behaviour classification and monitoring. Moreover, an algorithm is used to predict driver behaviours to control what is classified as dangerous or careless driving behaviour compared to the previous research in the same area. Meanwhile, different metrics are defined and developed to help assess the driver distraction level and distinguish between dangerous and normal driving behaviour. Limitations of the research and future work were highlighted as well.

CHAPTER 2. LITERATURE REVIEW

2.1 Synopsis

This research cuts across three main domains: the Intelligent Transportation System (ITS), Computer Vision, Deep Learning (DL). This section covers the significant and most recent literature related to the three domains mentioned above. The review of this literature takes place in the following sections.

2.2 Artificial Intelligence and Driver's Behaviour Classification

2.2.1 Introduction

It would be unimaginable decades ago to think of using a computer to diagnose a disease from blood samples or to have a car without a driver. Surgeons can now perform clinical surgery with higher precision using laser surgery technology [32]. This seems not to shock people anymore and looks quite normal and even natural nowadays. However, there is a need to accept and recognise that technological advancement has made so many things possible and so easy that it is now possible to look back and admit that technology has come a very long way [32]. However, other researchers support that human intelligence is not fully understood; therefore, it will be challenging to develop an intelligence that can imitate human intelligence. Nevertheless, the areas where much progress has been made include AI's research field [32].

AI in computer science deals with the area of research where machines are programmed to simulate human intelligence. Authors believe that it is also associated with the cases where the machine mimicking or exhibiting human-like traits of learning, reasoning and then developing capability in problem-solving, knowledge representation, etc., [32]– [34] [026]. Stephen Hawking affirms that the advent of artificial intelligence could be the worst event in the history of civilisation unless humans can control its development and expansion [32]. In this 21st century, AI is everywhere in daily lives and activities. It can be found in smartphones, cars, learning materials, production lines, cities, etc.

Nonetheless, Artificial Intelligence encompasses different subclasses, including ML, neural networks, and DL. It has been applied in different areas, including computer science, mathematics, physics, chemistry, medicine, aeronautics, banking, military, postage, neuroscience, transport, and aviation, nearly all aspects of human life, with more areas covered discoveries and work undertaken. Chakraborty [35] even further believes that artificial intelligence will transform every aspect of future life. The following subsection below will present some of its current applications.

2.2.2 Machine Learning

Machine learning (ML) is a subclass of AI, where a system can learn from experience and improve without any further programming. ML spans different research fields such as Artificial Neural Networks (ANN), DL, etc. [36]. One of the main focuses of ML is to develop computer programs where the machine will access data and use it to learn by itself. Indeed, with the world wide web showing faster development, a massive amount of data is available in all the fields of research, and one of the main focuses of computer scientists is to build programs that will help in analysing this available data, build models and infer helpful knowledge for all [33], [36], [37]. Some of the essential algorithms built and used so far encompass Support Vector Machines (SVM) and Naïve Bayes for classification, Bayesian Decision Trees, Self-Organising Maps for clustering, Principal Component Regression, etc. [36]. For example, in a support vector machine, an upper plane is needed to segregate two individuals: sick and those belonging to the healthy control group. This can also be applied to fully concentrated drivers and drivers who are very distracted by infotainment devices. The SVM classifier will create a separation boundary between the different classes [33].

Nevertheless, ML is also confronted with ever-growing data, especially security-related data, for network security issues [37]. Data collection issues will become even more complex due to the inadaptability of most current systems to the soon available 5G network. The diverse origin of data to be collected and analyse and the variety of their format will add to data analysis problems [37]. Once data collection has been solved, the research shows that context-aware systems behaviours can be adapted to the user's context. Therefore, correlating user modelling and context-awareness is essential to develop systems and services adapted to users' needs [38]. This development in the ML arena can benefit users in several areas, such as DL.

2.2.3 Deep Learning and Neural Network

Deep Learning is a subclass of ML that uses artificial neural networks; hence, the learning method can be supervised, semi-supervised or unsupervised. So far, there have been applications in different research areas such as natural language processing (NLP), machine vision, computer vision, speech recognition, machine translation, and drug design. Deriving from artificial neural networks (ANNs), DL is more about the unbounded number of layers and limited sizes related to system optimisation. This can be seen as an ANN optimisation model that is more disconnected and detached from the traditional biological structural model of connectivity, which is related to the original view and base of ANNs [39]. Therefore, in a DL application, such as in the field of image recognition, the input might have a matrix representation, whereby different layers might represent a different or specific part of the image. For example, the image of a driver in a car might have the following distribution: layer-1 dealing with the surrounding of the car, layer-2 dealing with the road, layer-3 dealing with the interior of the car, and layer-4 dealing with the driver's head positioning. In this manner, different layers will be dealing with as many details as needed for the analysis outcomes, providing what could represent a different level of abstractions and hence, improving the quality of the image and picture [39].

The precision of the image and a perfect object is essential with the classification of images. Furthermore, another essential factor is the localisation of the object contained within the images. A prominent model used in the detection of objects is the Deep Neural Network (DNN). According to Szegedy et al [39], the classification and localisation of objects are challenging. Szegedy et al developed a DNN that detects large object instances with varying sizes in the same image using limited computing resources. The DNN predicts bounding boxes of multiple objects in each image. The method entailed a DNN-based regression that outputs a binary mask of the object bounding box. A generic architecture for localisation is based on seven layers, with five being convolutional and the last two being fully connected. Furthermore, the SoftMax classifier as the last layer is used to generate a binary object mask.

Using some levels of tuning in terms of human involvement, DL can allow the developed model to magnify part of the image under investigation, giving some level of autonomy to the researcher in terms of the research aims and objectives. Nevertheless, DL has several applications in another area of research, namely the deep vision research field.

2.2.4 Deep Vision

The notion of deep vision is a computer technology allowing information to be extracted from images, video footage or camera images. Deep vision enables the application of CNN, whereby deep vision algorithms can convert images into shapes and movement and enable information extraction from images and videos using event automatic image analysis [24]. Different technologies have been proposed in terms of the applications of Deep Vision.

Researchers proposed a wearable size version of deep vision software dealing with different executions of several deep vision models, such as CNNs. For example, DeepEye can run multiple cloud-scale DL models, whereby rich analysis can be performed locally using an embedded processor without the need to offload the image file on the cloud. In this manner, the software reduces the computational overhead cost, as the heavy convolutional layers load from memory and reduce the connection between layers. In addition, the execution framework includes memory caching and minimises the memory bottlenecks by applying a model compression technique, which in turn considerably reduces the need for ample memory space [40].

On the other hand, Xu et al [41] presented DeepCache, another deep vision technology used for learning inference in a continuous mobile vision. DeepCache tackles mobile vision challenges, whereas trade-offs about other challenges must be overcome, including cacheability, overhead, and, more importantly, the loss of model accuracy [41]. The model is an application of deep vision in DL whereby the model refrains from using video heuristics because of the difficulty related to the data interpretation. Nevertheless, many advantages are related to the development of the model, including saving inference execution times of 18-47% while reducing energy consumption by 20% [41]. Guo et al [42] developed a robotic grasp detection system based on images that can predict and accelerate the robot's detection speed.

Since robots cannot intuitively detect a grasp location area for a given object, the current deep vision system allows the robot to learn from the image. The model allows rectangular potential grasp areas to be created, refined with a score attributed to each graspable location in real-time at the speed of 80 frames per second. Datasets are built in this manner from those images, allowing a comparison of different rectangles to produce the best detection performance, whereby the robot learns quickly to detect the best graspable areas of each object [42].

A significant challenge in object detection is that CNNs require a fixed-size (224 x 224) input image, which is stimulated, thus reducing image recognition accuracy. He et al [43] proposed using a pooling strategy, spatial pyramid pooling (SPP), to resolve the earlier challenge. Convolutional layers and fully connected layers involve the use of cropping or warping at the initial stage. He et al used Spatial Pyramid Pooling (SPP) in generating fixed-length output, regardless of input size with single window size [43]. SPP makes it possible to generate arbitrarily sized images for testing and accepts images with varying sizes or scales during the training. The approach involved the convolutional layers using sliding filters, with their outputs having the same ratio at the inputs. The output is referred to as feature maps, which entails spatial positions.

2.3 COMPUTER VISION

2.3.1 Computer Vision Algorithms

Computer vision is the field of research involving computers, digital videos and images, whereby the former will gain more understanding and cognition from the latter [44]. Different algorithms have been developed in the computer vision field of research. Szeliski [44] looked at framed family photographs and stated that it could be easy to count and name everyone on each photograph and, further, guess their deep emotions at the time from their facial expressions [44]. On the other hand, perceptual psychologists have spent decades trying to understand how the visual system works; although developing optical illusions to tease apart some of its principles, reaching a satisfactory solution can be elusive[26], [27]]. Researchers have been trying to develop methods and techniques to recover three-dimensional shapes from images. Advanced techniques enable an individual to be tracked in a complex environment, dealing with a combination of faces to determine people's names in a photograph, etc. However, the truth is that to develop a system that will allow a computer to reach the same level of interpretation and accuracy as a two-year-old child remains elusive[25].

Nevertheless, it is also true to admit that significant advancements have been made in the development of computer vision algorithms, with applications in different fields of research, such as algorithms in distributed computing, cognition understanding, computer surveillance, intelligent environment, robot coordination, space exploration, and so on. Indeed, camera-equipped sensor nodes communicating with a wireless network can provide ease, flexibility and robustness to nodes failure [47]. Computer vision algorithms involve the collection and analysis of data from videos and images for various purposes. This includes decision making, monitoring, etc. The initial step to this work is the object detection and recognition stage, where cameras can track the object in motion. Tron and Vidal [47] supported the inference level on the relationship between the human-object interaction, human behaviour, multi-objects, etc., and the context environment can be critical in the decision-making process.

Nevertheless, different applications areas exist, and they are dealt with in the following subsection.

2.3.2 Application Areas

Computer vision algorithms (CVA) have been applied in different areas and can have several applications. For example, a Camera Sensor Network (CSN) is used to continuously monitor scenes such as the workplace, disaster zones, exploration areas such as space exploration, etc. In these cases, the function of the CVA will be to direct the cameras to record images data and transmit these images to a server for a centralised analysis [47]. In this case scenario, the images will go through different stages, including the transmission, processing, and analysis of a large volume of data. The process can be effective if each camera can process its data with other sources (cameras, etc.). This usually creates problems such as noise with overlapping images from different sources. Distributed algorithms can be a solution for noise management, with a global analysis of the scenes obtained with these distributed algorithms fusing all local and external image data [47].

On the other hand, computer vision techniques have allowed consistent advancement in areas such as computer vision technologies integrated into Unmanned Aerial Vehicles (UAV), which fall under the category of autonomous vehicles; these applications have enabled effective management of the aerial perception issues. This includes obstacle detection and avoidance, visual navigation algorithms, and aerial decision-making. One well-known and even controversial utilisation of this combined system is the Remotely Piloted Aerial System (RPAS), also known as drones [48]. Recent applications of drones include mail and postal delivery, such as the Amazon delivery drones, which are meant to make delivery faster; drones are also used in targeting enemies on battlefields, etc. Other forms of UAV include helicopters, tricopters, quadcopters, etc., which were initially designed for military purposes on battlefields and to save human lives [49]. However, for their ability to operate in dangerous areas and situations, such as scientific exploration, rescue operations in disasters, etc., certain special helicopters and Vertical Take-Off and Landing (VTOL) rotor-craft, including quad, Hexa, octo-rotors, are being used [50]. It is essential to mention that military applications are one of the most critical areas of application. For example, missile guidance, where the missile is sent to a combat area rather than to a specific target, is made clear by more information provide by locally acquired data (xx). This presents some of the main applications of computer vision algorithms, however, many more exist and and reviewing all these applications is beyond the scope of the current study.

2.3.3 Example Standard Dataset

Computer vision is presented as a fast-growing area of research that is primarily applied in the industry. The interaction between the human agent and machine has changed with Computer Vision technologies that have transformed this relationship to benefit industries. Building a trusted and solid work relationship using robust DL algorithms for Computer Vision requires a high-quality dataset to train the algorithm [51]. A list of standard datasets used for a computer vision project includes the following, most of which will be succinctly reviewed in this subsection.

2.3.3.1 CIFAR-10

This popular computer vision dataset was used for object recognition encompassing 60,000 colour images (32 × 32) in 10 classes. This series of images is divided into 50,000 training images and 10,000 testing images. The classes encompass the following: aeroplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck. The CIFAR-100 dataset is composed of superclass and classes [51]. They can be accessed through the following link: http://www.cs.toronto.edu/~kriz/cifar.html

It is essential to mention that the different classes are mutually exclusive, meaning that overlapping them is impossible. For example, a distinction is made between dog and deer, aeroplane and automobile, etc. [51].

2.3.3.2 Cityscapes

This corresponds to an open-sourced large-scale dataset for computer vision projects containing a diverse set of stereo video sequences recorded in street scenes from 50 different worldwide. It encompasses a high-quality pixel level of 50,000 annotated frames and a more extensive set of 20,000 weakly annotated frames. The datasets are used to train DL neural networks for the performance assessment of vision algorithms [034]. It can be assessed from the following link: https://www.cityscapes-dataset.com/

The dataset is free for research and other related purposes, including teaching, scientific publications, etc. Benchmarks are available for performance measurement, whereby the results can be sent to an evaluation server [51].

2.3.3.3 Fashion MNIST

This corresponds to an image dataset for computer vision encompassing 60,000 samples for the training set and 10,000 samples for the testing set, whereby the example is a (28 × 28) grayscale image of 10 classes. A benchmark system using Scikit-learn covers 129 classifiers with different features. The dataset is utilised by the AI-ML community as a benchmark for the algorithm's validation and can be accessed from GitHub on the following link: https://github.com/zalandoresearch/fashion-mnist

Researchers believe that this dataset is the whole first set used, and if a project fails on MNIST, it is likely to not work in any other dataset. However, they are planning to replace MNIST because of the following reasons: first, it is too easy, and Convolution Networks and ML algorithms can quickly achieve 97% or more; second, it is overused; and third, the dataset cannot present modern CV tasks, as research has warned [51].

2.3.3.4 ImageNet

This is a popular image dataset classified based on the WordNet hierarchy in the field of computer vision. Therefore, it provides access to an image database, whereby the WordNet hierarchy allows a cleanly complete organisation of these images. Themes or concepts describable by words or sentences are called synonym sets or synsets. The database constitutes 100,000 synsets in WordNet, with ImageNet creating 1,000 images representing the illustration of each synset in the WordNet. It also corresponds to several 1,000,000 sorted images related to different themes in the WordNet hierarchy [51]. It can be accessed from this link: http://www.image-net.org/.

The ImageNet project answers a call from researchers in the computer vision field and academia asking for more available data in this field of research. In this respect, researchers have developed more complex algorithms to index, organise, retrieve, update, etc., multimedia data. This will be a means for helping researchers in the field to make a large-scale image database available for different purposes, including research purposes [51].

2.4 Intelligent Transportation Systems

2.4.1 The Notion of Intelligent Transport

The notion of the intelligent transport system (ITS) relates to the usage of different types of advanced applications that provide better transport combined with an effective traffic management system, allowing traffic users to make an informed decision regarding the usage of transport networks system. Researchers believe that the intelligent transport system has brought several network changes and improvements in the last two decades. This includes a drastic improvement of the security in transportation, the multiplication of choices for travellers. Furthermore, thanks to the data availability from an enormous variety of sources, the researcher's work are now more complete and accessible [52]. The same researchers believe that this can significantly change ITS development, which is currently a necessity. Indeed, due to the increase of vehicles on the road and the related issues such as road congestion, increased pollution level, illness such as heart disease, but also the increased number of road accidents, there is a need to come up with a more intelligent transportation management plan, whereby AI, ML, and DL can play a leading role by bringing more innovative solutions. Research shows that three-fourths of road accidents are caused by human error [53].

Given that the transportation system plays an increasingly important role in a country development, its economic strength depends on its effectiveness and efficiency. Indeed, it is reported that 40% of the population worldwide spend at least 1 hour in the transport system every day [54]. Human daily activities are so transport-dependent that this has increased the number/level of congestion in big cities, leading to increased pollution and consequent health problems such as heart disease [53]. Therefore, it is crucial to implement a country-level strategy that helps mitigate and reduce the transportation system's impact and human involvement on issues in modern society. This section of the thesis analyses the different algorithms and their role in reducing problems linked to the increased need for a transportation system in daily life.

2.4.2 Intelligent Transport Systems Using Pervasive Technologies

Naturalistic data from vehicles has been collected using onboard sensors that collect driver braking behaviours to reduce collisions and prevent accidents. Research has shown that most car crash fatalities and injuries happen at roads junction or intersections. Therefore, many car manufacturers seeking to enforce security measures are planning to release an automated braking system that will reduce the car speed at the approach to an intersection to reduce car crashes [44]. Nevertheless, the researchers have decided in this study to develop a probabilistic model for human driving behaviour, which will distinguish between possible and probable scenarios. However, the system’s total safety probability p is still correlated to the surrounding of the vehicle, such as the other drivers’ (1-p) probability, which is related to their behaviour [55] [45]. That is perfectly understandable because a vehicle can be parked, and another driver can still crash into it, causing causalities.

In another study, researchers tried to monitor drivers' behaviour in a normal and driving task condition. Different devices were used to perform this monitoring process, including cameras, microphones, and Controlled Area Network-Bus (CAN-Bus). Indeed, from the frontal video camera and the car CAN-Bus data, the researchers would be able to differentiate/discriminate between two main conditions, such as average and task driving conditions [23]. Controller Area Network Bus (CAN Bus) provides other information, such as steering wheel angle, brake value, and vehicle speed [23], [56]. The data thus collected non-invasively using CAN-Bus, video camera and microphone arrays are made available to model and understand drivers' behaviour behind the steering wheel. A binary K-Nearest Neighbour classifier was used for the analysis of the data collected. Researchers found that the information provided by the frontal cameras made significant difference in segregating between the two conditions, with a small improvement in the classification result due to the CAN-Bus system [23].

Another group of researchers also used a non-invasive system to monitor the driver's attention level, given that they perform secondary tasks. Using real driving scenarios using a multimodal approach of driving behaviour monitoring, based on the secondary task performed while the attention level was controlled, showed a variation in the driver's attention due to this secondary task performance. Support vector machine using K-Nearest Neighbour (KNN) combined to a sequential floating forward selection (SFFS), where the last one was used as a dimensionality reduction algorithm [57].

On the other hand, invasive sensors have been used to monitor drivers' distraction levels. For example, the driver's head pose, and eye gaze were used to understand and monitor his behaviour. Such monitoring tools include the use of monocular, infrared (IR) and stereo cameras to track driver distractions have been considered. Other forms of invasive sensors were used in the data collection process and included Electroencephalography (EEG), Electrocardiography (ECG), and Electrooculography (EOG,) which were necessary to estimate relevant biometric signals associated with distraction. In addition, a UTDrive platform was also used in buildings a multimodal database to gather actual driving conditions [57].

Nevertheless, a survey was also built to extract driving features of the internal state of the driver, which could result in uncomfortable and dangerous situations for him and other road users. Statistical analysis was used to separate normal and abnormal driving behaviour, making this dataset available and a fascinating document in analysing the drivers' behaviours [57]. Driver-controlled behavioural data were analysed, and abnormal driving behaviour was categorised using the jerk information, which ultimately allowed the group of researchers to classified abnormal and normal driving behaviours [58]. This research was conducted on a Mitsubishi Precision Co. Inc., DS6000 driving simulator, allowing the collection of this different type of data related to jerk behaviour [58].

It has been stated that accidents occur primarily due to human error; thus, it is necessary to understand human behaviour to reduce accidents. It could be argued that the elimination of dangerous driving by identifying careless driver's behaviour is essential. However, the driver's behaviour, sickness, drowsiness, time constraints, and the environment [61]. Therefore, investigating the relationship between the internal state of the driver and the environment could give a holistic picture of the accidents and the conditions surrounding them. The study also showed that one of the main factors influencing human driving behaviour is the instability of the internal state of the driver. In order to have data under realistic driving conditions, such as freestyle, everyday driving and time constraints, researchers have developed a model using questionnaires to correlate driving behaviour and the internal state of the driver [61]. However, this research was also limited by the lack of a thorough understanding and information on the driver's internal state, as admitted by the researchers.

Profiling drivers' behaviour has been performed to identify risky driving manoeuvresand improve drivers' efficiency. In addition, monitoring and profiling of drivers have been applied in car insurance areas to have fair insurance premiums for customers. The driver data collection was performed through the use of mobile devices and telematics boxes. The researcher recorded data related to turning right and left manoeuvres, acceleration and deceleration manoeuvres. Finally, bumps and potholes in the road make this Android mobile a fair and accurate device in helping drivers be safe on the road for themselves and other cars in the surrounding [59]. However, factors and metrics such as weather information, distance, speed distribution, and road topology (lanes, troughs, road conditions) could affect the driving behaviour, thus misleading vehicles to be detected as careless by the safety mechanisms. The data collected did not consider those, which can be considered some of the limitations of this study.

Drivers' behaviour profiling through smartphone analysis of everyday driving and risky conditions by collecting driving traces from one point to another [59], [60]. In contrast, in another study/research, it was proposed to detect risky behaviour by focusing on events rather than road traces and providing feedback, thus, allowing drivers to adapt their driving based on events [59]. Driver feedback dissemination has been used in the monitoring and correction of driver behaviours and bad driver habits. This was achieved by collecting driving data using phones in predefined vehicle positions and analysing acceleration, lane changing events, and braking. Driver classification was based on features such as acceleration variation jerk and maximum acceleration threshold [61]. However, the driver feedback dissemination collection methodology does have its limitations due to the phone's positioning, which could be slightly modified by device vibration or user manipulation. In addition, the solution, as mentioned earlier, uses fixed thresholds in combination with different metrics such as braking events, acceleration, steering events, and manoeuvres in profiling the drivers' behaviour. In contrast, this time, the features used include the acceleration, global positioning systems (GPS) data in the detection of road quality, such as potholes, bumps and traffic flow (stop-go or fluid) – events that can trigger unpredictable drivers' behaviour [62].

Other algorithms have been used to distinguish between normal driver's behaviour and risky driving behaviour with metrics such as smooth acceleration and magnetometer data. The findings include detecting events such as braking, acceleration, aggressive steering and sudden manoeuvres [63]. It could be argued that risky driving behaviours can lead to a situation that could cost human lives. The Bayesian technique was used to classify drivers' behaviour into risky or safe/average driver's behaviour [64].

2.4.3 Impact of ADAS on Drivers' behaviour

The impact of ADAS on the driver's behaviour can be immensely positive in bringing tremendous changes in helping to improve human/driver's behaviour, especially their involvement in road accidents. According to the research, significant factors influencing the cause of road accidents are vehicle, human and road conditions. One way to reduce road accidents is by profiling the vehicle driver using ADAS [59]. ADAS relies on driving data from infrastructures and vehicles to infrastructures (V2I) collected through sensors, mobile devices and vehicles [65]. It is also stated that human error is the major influencing factor that causes road accidents, whilst the other two factors are road infrastructure and vehicle capabilities [66]. In addition, aggressive driving behaviour is different from driving manoeuvres or events during a journey, like harsh braking and acceleration, lane changing and rapid turnings. It could be argued that the driver's action, such as performing certain events, influences vehicle behaviour. According to researchers, in a VANET, the vehicle moves around with no boundaries on direction and speed, resulting in arbitrary motion, which poses great difficulties to researchers [67]. In contrast, sophisticated malware could control the vehicle, thus forcing it to act recklessly. This gives rise to concerns and asks for a possibility to differentiate between different occurrences.

Smartphones have been used in assessing, estimating, and evaluating driver's behaviours [61], [62], [64], [66]. The sensory data can generate driving behaviour profiles and categorise drivers according to events, such as acceleration, braking, turning, and lane-changing during a journey [68]. Drivers' behaviour profiles based on potential blackspots using data of vehicles that had previously passed a given road section for road safety analysis was used. This is a clustering approach in the categorization of drivers into specific driving styles and profiling of driving events that occur during a journey [64], [69], [70]. The crowd-sourcing model has also been adopted to collect driving data in vehicle-to-infrastructure implementation, and analysis in real-time localised driving information is being broadcasted to vehicles within a range of infrastructure.

Pattern-matching techniques in analysing driving events from acceleration data, whereby a smartphone was placed in a vehicle lying flat while pointing in the same direction of travel as the vehicle, profiling drivers. The result also showed a classification of drivers into aggressive (careless) and average [59], [65]. Driver's behaviour can impact vehicle dynamics in various ways, and it can even disrupt vehicular network formation, such as platooning in a VANET. Vehicle to vehicle (V2V) communication can be attacked to disrupt the vehicle's behaviour, thus leading to communication that could endanger other moving vehicles on the road. For example, an attacker might perform a Sybil attack, resulting in the vehicle deviating from its platoon [71].

2.4.4 Drivers' Behaviour Simulation

Driving research has previously been done using a sophisticated driving research simulator called National Advanced Driving Simulator (NADS) [10]. However, the NADS is limited regarding realistic driving conditions that did not consider the context environment. Furthermore, driver behaviour simulation in traffic conditions has been analysed using microscopic models. Driver's actions are influenced by factors such as traffic movements and road causalities. The microscopic model analysis studies individual driver behaviour, connecting rear and front vehicles and cumulative macroscopic traffic [25]. The parameters of individual drivers have been used to represent unique driving behaviour; rules are further applied by relating the traffic state observed by the driver to the decision made by them.

On the other hand, it could be argued that having a predefined rule will not capture a naturalistic driving behaviour before human decision making can be randomised. The Artificial Neural Network (ANN) technique on predefined driving rules using car following models involved measuring traffic state to drivers' actions. However, ANN requires sufficient real-time data to capture the correlations between driver actions and traffic states effectively.

In contrast, it should be noted that Naturalistic Truck Driving Study (NTDS) data conducted in Virginia Tech Transportation institute collected was used in the findings of truck drivers' responses to cars following a traffic situation [26]. The research findings showed that driver behaviour differs in context, and driving actions could be influenced by the type of vehicle driven [29], [30]. Nevertheless, using only truck driver data could be regarded as some of the approach's limitations, as mentioned earlier.

On the other hand, Lim and Yang [55] collected abnormal driver data using state-of-the-art sensor technology in a simulated driving environment to estimate driver states, such as cognitive distraction, visual distraction, driver drowsiness, and workload [72]. Detection algorithms in the detection of driver states, such as DL CNNs, have been applied in different fields of research, such as speech recognition and computer vision [27], [28], [73]. The researchers found out that DL was a promising approach compared to the dynamic Bayesian networks. However, Lim and Yang's [55] proposed approach adopted a CNN technique performed solely on the image plane but did not consider driving dynamics and driver's reaction. In addition, the model did not consider the context-awareness of the vehicle and the driver's perception. Such omissions and choices could be considered as limitations to their research approach.

Video coding of animal behaviour required analysing each frame individually by detecting animal body parts with further ML technique CNNs [74]. In this study, the participants were giving predefined conditions to adhere to before the simulation took place. This is believed to introduce some bias in the study. Nonetheless, the research was conducted in a simulated environment that did not offer the same conditions as a natural driving environment, whereby other road users' presence can significantly impact the driver state. These are some limitations to this approach, including the need to analyse accurate driving data. A preferred solution to the limitation, as mentioned earlier, could be the use of DL CNNs that can recognise behavioural states from Images. In addition, there is also a need to estimate the degree of careless driver's behaviour that can have severe consequences such as accidents.

2.4.5 Classification Techniques of Drivers' Behaviour

Statistical observation techniques method using time series analysis has been used to classify driver's behaviour in the differentiation between everyday driving from dangerous driving behaviour. Time-series data analysis of vehicle states with corresponding timestamps of fixed sampling rate and pattern emissions in profiling naturalistic driver's behaviour [75]. The technique is based on manoeuvres and lane trajectories of a perceived vehicle in classifying a driver's behaviour. Furthermore, a probabilistic model such as Hidden Markov Models (HMM) is applied in this classification task to classify dangerous driving behaviour and normal driving behaviour. The probabilistic technique takes into consideration driver's behaviour such as driver distraction, fatigue driving, etc. This probabilistic technique can be Gaussian Mixture Model, Bayesian Network or the aforementioned Hidden Markov Models (HMM). The limitation of the adopted technique is its inability to detect driver's behaviour events leading to chaotic manoeuvre that constitutes a dangerous driving behaviour. It could be argued that vehicle behaviour is impacted by their surrounding vehicles, making the driving environment challenging to predict and increasing the system's complexity.

The Probabilistic Bayesian technique has been applied in detecting driver's behaviour by observing driving manoeuvres and traffic context. A probabilistic model is efficient in distinguishing between probable and possible driver's behaviour. Behavioural distribution of driver's behaviour with an algorithm that computes the number of times driver hits the brake correcting vehicle speed at the proximity of intersections. Forghani et al [55] proposed using a probabilistic model designing in-vehicle driver-assisted systems that warn drivers to prevent collisions. The technique adopted involves simulation using naturalistic data, but it was limited to the isolation of the surrounding vehicles' impact on driver's behaviour.

Furthermore, several assumptions made limited the efficiency of the safety system to the prevention of only rear-end collision. The stochastic model of driver's behaviour based on convex Markov chains (CMC) in tracking potential driver distractions and the predictions of car trajectory based on the estimation of driver's behaviour in prevention of collision had been adopted. Therefore, looking at the probability of random events using this stochastic model could be of paramount importance in studying driver's behaviour. Nevertheless, other intelligent models were also considered in various studies, with different results, some of which improved precedent research.

For example, the **Statistical Gaussian Mixture Model (GMMs**) has been used to identify multimodal features that can be used to separate normal driving behaviour from task driving conditions. Classification of the degree of secondary tasks according to the degree of distraction has been observed. GMM was used in quantifying actual deviations in drivers' behaviour from expected standard driving patterns [57]. A regression model was proposed as a metric for characterizing the attention level of the driver.

On the other hand, **Hidden Markov Models (HMM)** has been used to model dynamic processes of drivers during the phase transition period at high speed signalized intersections [76]**.** However, using a single deterministic model of driver's behaviour is not efficient due to the unpredictability of human behaviour. Using stochastic techniques such as Markov Chain (MC), human behaviour-based prediction under varying distractions and environment was employed as a supplemental technique. In addition, a probabilistic driver model that predicts driver trajectories using a Convex-MC (CMC) model was used in another study to correct some of the limitations of the previous study [77], [78]. The mentioned above Markov chain entails the transition probabilities with convex uncertainty sets. Future Prediction of driver's behaviour relies on vehicle environment, driver's state and previous data such as steering manoeuvrescollected using a car simulator.

Furthermore, **Fuzzy Logic Methods** can be used in simulating and predicting car-following behaviour by employing an improved neuro-fuzzy inference system (ANFIS) model in analysing the reaction delays of drivers. In simulating the prediction of car-following behaviour, the primary inputs to the ANFIS model were the reaction delays obtained from real-world data sets. Results from the simulation proved that the proposed model is highly realistic and compatible with real-world data.

In a bid to effectively predict car-following behaviour for different lead and following vehicle types, the Neural Network Model [15] was proposed. The model's performance was analysed with the use of real-world data from six vehicle types. Vehicle type following behaviour prediction was done using a multilayer feed-forward Backpropagation network (MLFF-BPN), where the inputs to the model were the vehicle type. The integration of the model into the simulation aided in studying the macroscopic behaviour of the model.

Regarding the attacks in the vehicle, this implies that the taxonomy of attack will vary according to the vehicle type.

Trajectory Prediction is also an optimization technique that has been used in the detection of driving manoeuvresin real-time. Strategic behavioural trajectories deploy **a** knowledge-based cognitive architecture to model human drivers behaviour or predict future actions and cognitive load of human drivers behaviour [79]–[81].

ML algorithms such as Bayesian classifiers, Artificial neural networks (ANNs), Decision trees and Support vector machines (SVM) have been applied to gather driving behaviour**.** The algorithms allow real-time monitoring of vehicle behaviour, real-time predictions, and notably coping with overlapping inter differences (average and careless vehicles) [82].

Support Vector Machine (SVM) has been used in a study to monitor driving situations such as unintentional lane departure prediction associated with an alert system to warn the driver to ensure he improves his behaviour [83]. A binary SVM is deployed in the classification of the time series of selected variables. Training and testing of driver experiment data were performed using VIRTTEX, a device hydraulically powered in 6 degrees of freedom moving driving simulator at Ford Motor Company. The SVM classifier achieved the following performance: sensitivity = 99.774% and specificity = 99.999% in predicting driver lane departure for each of the 22 drivers sampled. It should be noted that a non-linear kernel was used for the SVM classifier to achieve the separability of the driving data due to the complexity of the data structure.

Predictive modelling such as sparse Bayesian learning in the classification of driver's intent using lane change analysis from a camera view of the driver, internal vehicle sensors, lane position and trajectory. Indeed, a study using the Bayesian framework that assesses the criticality of the situation based on data monitoring the vehicle and its surroundings and the human behaviour related to the prediction that the driver intended to break to avoid a catastrophe or not. Using actual data from test driving performed by 28 different drivers over 22 hrs in driving scenarios and in other case scenarios, a probabilistic model was built for the system preventative measures to be constructed based on different levels of distraction severity level and the driver behaviour, including his intent. One of the frameworks developed allowed the fusing of the predictive driver behaviour information with the vehicle, its surrounding information, and braking assistance. The results show that the framework is fit for assessing the criticality of the situation and the need for an intelligent vehicle safety system to intervene [78], [84], [85].

A survey on prediction models to detect driver's behaviour has been conducted to ascertain driver's actions and ensure the safety of people and compliance with driving regulations. Onboard diagnostic (OBD) information has been used to collect the vehicle speed change rate, throttle change rate, engine speed change rate, and engine load calculation. The data was analysed using AdaBoost algorithms to create the driving behaviour classification model [86], [87].

It could be argued that the probabilistic approach relies on previous data to predict the outcomes. Relying on previous data to predict the future can sometimes be flawed; therefore, it is essential to get an algorithm to develop a real-time monitoring model. Nevertheless, different classification methods have been used to detect careless driving behaviour, but little has been done to prevent and classify maliciously infected vehicles. In addition, there are possible ways in which a maliciously infected vehicle could lead the vehicle to behave carelessly thus, resulting in false classification into careless driver's behaviour.

2.4.6 Driver Perceptions of the Road Environment

Previous researchers have studied the relationship between accidents and human errors, and the results show that over 600,000 traffic accidents per year have led to severe problems. The number of yearly accidents is more than 70 million, and the eradication of traffic accidents has not been achieved [88]. In order to eradicate the persistent occurrence of accidents, several techniques have been applied to assist drivers by providing driving safety support systems and autonomous vehicles. Technology advancement has risen over the years in the area above, but there is still a need for limited human interaction to aid the decisions.

According to Imamura and Asakawa [73], drivers' ability to perceive hazards influences driving behaviour. Pattern analysis of driver pedal with that good pedal has detected dangerous driving behaviour in intersections [89]. Furthermore, risk perception was considered from risk acceptance (knowledge) and behavioural readiness (ability). The study further suggests classifying hazard perception in a driving situation with pedestrians' fault, driver's fault, hiding and environmental factors. In addition, the classification of driver's behaviour tendency has been listed as near-miss, disrespect accident, overconfidence, lack of confidence, and good balance. Another type of classification was performed that considers the risk awareness of the driver using heart rate changes in passenger and driver. Risk perception in driver's ability in hazard situations by their knowledge for prediction and capability in accident avoidance was also investigated [89]. However, it could be argued that the driver will react under specific circumstances. It is essential to ascertain whether a reaction is expected in deciding whether the driver might choose to swerve around an obstacle or whether automatic emergency braking should be initiated [90], [91]. Nevertheless, the study showed some limitations because the time of driver reaction to an eventual source of distraction should be ascertained, but the degree of reaction to the event or distraction from the ordinary course of driving needs to be measured. In preventing accidents, optimising the degree of driver's reaction could help vehicles make intelligent decisions, such as prioritising and triggering the safety control mechanism.

Driver's aggression has been one of the significant factors in accident-related incidents in the driving environment. Aggressive behaviour from another driver or a pedestrian, such as hostile gestures, angry epithets, and strong words that can cause anger, elevated blood pressures, has led to accidents. In addition, the level of perception could differ from individual to individual, while factors such as race, age, gender could also be discriminative. A study has shown that gender differences influence anger in emotional processing in the driving environment [92]. The study showed that males are more sensitive to influence factors of cognitive component compared to women. Males are more likely to perform hostile actions such as sudden braking, overtaking, fuzzy road signs, succession red light, whilst females are more likely to be sensitive towards horn urging, encountered, careless curse and exhaust emissions. Group-specific targets have been used to detect the driver's perception using a Drink and Drive simulation method. The simulation involved a younger audience who were intoxicated to drive a simulated car (gamification) in a static position. The result shows that the reaction and perception of drivers under the influence of alcohol could trigger intrinsic motivation to engage the young participants in gaming [93]. However, the simulation mentioned above was not carried out and did not achieve a result that could be a precise scientific finding related to drinking and driving, thus being an explicit limitation to this study.

Tannahill et al [94] stated that there are factors and parameters to be considered in the estimation of drivers' behaviour and perception, such as environmental conditions, namely wind (speed), temperature, rain (wet road, windscreen wipers), time of day (daylight/dark). Their proposed system included the parameters mentioned above in developing a driver alerting system using real-time range estimation.

Pugeault and Bowde [95] proposed a method to detect driver's pre-attentiveness using a novel vision-based approach to autonomous driving that can predict and anticipate drivers' behaviour in real-time. The analysis entails analysis of visual scenes most predictive driving context or driver’s actions. Furthermore, the research focused on the driver’s behaviour on urban roads that requires different visual skills [96]. The research results showed that the vision-based approach could detect, predict, and even anticipate driver behaviour using preventive vision only in real-time. In addition, the model can detect the driver’s action related to braking and turning in over 80% of cases and estimate the driver’s steering angle accurately.

2.4.7 Detection of Cognitive Drivers’ Behaviour

Facial Geometry-Based Eye Region Detection

The detection of cognitive distraction and sleepiness can be inferred from the facial geometry-based eye region detection by analysing the frequency of eye closure and eye blinking when the driver is zoning out. The analysis of the gaze estimation and driver’s deviation from the frontal view of driving or a context-aware situation can be used to infer the driver’s cognitive distraction. Face recognition is the most straightforward approach to detect gaze estimation or driver’s attention. The driver’s awareness can be measured using a method that tracks the pupil and estimates the driver's reaction time to events, while the vision field correlates strongly with the driver’s gaze [95].

2.4.7.1 Gaze Estimation: Eye Tracking

Detection and tracking facial features from face images with different facial expressions under various face orientations in real-time has been used in detecting cognitive distractions [97]. Eyes play a significant role in understanding a driver’s intentions and emotional states. The technique adopted measures the electric potential of the skin around the eyes, but this technique is intrusive and results in profound user acceptance. Alternatively, a non-intrusive eye tracker involves a camera that tracks and profiles the driver eye gaze and blinks rate in real-time then estimates and profiles the driver’s eyes blink rate. The cognitive distraction is detected by deviation from the driver profiled blink rate and the degree of reaction of the drivers towards events such as distraction or near-crashes. The reaction level of the driver to an event from innate cognitive distraction can be used to detect zoning out [97].

2.4.7.2 Wearable Technology: Wearable glove system

Innate emotions such as stressed mental workload could lead to temporal loss of concentration and vehicle control. Stress response could result from psychological thought, observing physiological reactions from respiration and heartbeat using biosensors that could disclose driver mental stress conditions in different driving conditions, including driving on highway and city [98], [99]. The innate emotions detection system uses the wearable glove system and consists of sensor module, hardware processing units, PPG sensor, inertial motion unit (IMU) sensor, MCU processing unit, analysis module, and alarm module implemented in end terminal application [99]. The studies showed to be more effective as monitoring systems with regards to monitoring driver's cognitive behaviour and, for example, monitoring the stress level in real-time by detecting the physiological signal and steering wheel motion. The study established a strong impact of the stress on driver's behaviour with over 95% accuracy level obtained by the SVM classifier.

2.4.7.3 Electroencephalogram (EEG)

Neurophysical signal such as Electroencephalogram (EEG) and brain activity has been used to understand the precursors of cognitive driving distraction at the psychological level [100]. EEG power spectrum is used to analyse spatial and temporal brain signal dynamics in monitoring driver states. The classifier was used on features extracted is Nearest Neighbour Decision Tree, Naïve Bayes, Random Forest and Support Vector Machines (SVM). In addition, Electrocardiogram (ECG), if embedded in the driver's seat, can be used to detect cognitive distractions.

Nevertheless, it should be noted that the proposed classifier will not detect cognitive distraction because cognitive detection requires sensors and having contact with the driver. This implies that this research will only focus on visual features using secondary data (TeleFOT data). Significantly, the classifier could be further applied in developing a driver monitoring system that can detect drivers induced by a stroke based on reaction time or driver state, which can be detected visually. For example, a stroke with symptoms such as coordination problems due to facial drooping, arm weakness, numbness and stiffness, legs stiffness, speech difficulty for a certain period can be inferred through image recognition technique. The argument could be the risk of a high level of false positives error. However, to reduce this type of error, the classifier could be tuned to detect the presence of multiple symptoms before triggering the alert system. The use mentioned above case will rely on image recognition and will require no contact.

2.4.8 Autonomous Vehicles and their Relationship to the Drivers' Behaviour

Autonomous vehicles are the future technology regarding the new type of vehicles on city roads and in the countryside. It is essential to understand their functioning concerning accidents limitation and avoidance. Research carried out so far has shown that autonomous vehicles do have limitations in perceiving obstacles and difficulty in identifying objects such as bicycles [101]. There will be a need for human intervention to subvert any form of accidents from happening in such occurrences. Thus, the driver state needs to be monitored in real-time. In addition, looking at the level of autonomy, achieving whole autonomy level for all the countries and terrains might take some time to get to perfection, achieving a zero-accident tolerance. Therefore, drivers must remain in control or take-back control indeed failing. This boundary might be challenging to establish unless specified by the vehicle design specification, user manual including.

Nevertheless, in certain conditions, it might also be challenging to allow a driver who shows apparent signs of tiredness, fatigue, etc., the detection of the driver's behaviour will be necessary to avoid irreparable happening. For example, in some case scenarios, it will not be ideal for handing back the complete control of the vehicle to a driver who is showing apparent signs of tiredness or might be highly distracted by inside vehicle infotainment or such things. There should be a decent warning system to alert drivers of the risk he is taking and the danger he is exposing other road users to, in which case he shouldn't if he was prompt to take over from the automated driving system. Thus, driver monitoring is still paramount. To even go further, the software should detect this type of case scenario, hence never allow or prompt this type of driver to take over from the automated driving system.

The current projections in the field were that by 2018 there would be hands-off driving on certain motorways by autonomous vehicles for a maximum of around 3 minutes at a time [102]. This level of autonomy will be at level 4 (high automated vehicle), a more ADAS. For example, Telsa vehicles would warn if the user removed his hands from the steering wheel for more than 5 seconds in a hands-off driving situation. Presently, the regulations indicate or oblige users always to hold the steering wheel for all autonomous vehicles. In addition, the drivers will still be expected to take control in some circumstances if the technology fails. As a result of technology failure, Uber recently suspended their autonomous driving project due to an accident in Arizona [103]. Tesla pulled back 53,000 vehicles due to electronic braking systems failure in the Tesla Model X and Model S vehicles [104].

However, it is essential to understand that research in the field should never stop, hence as a researcher, there should be continuous improvement of the technology and avoid these failures that will happen in the early days of applying this technology. Indeed, the application areas are so vast and could be a game-changer for human society. For example, the current research can also be integrated into autonomous vehicle technology and applied to health-related issues, such as the driver behind the steering wheel is suddenly having a stroke. The example, as mentioned earlier, is a way the vehicle could detect and respond to such dangerous situations that could be crucial to driver safety and the safety of other road users, instead of focusing only on the management of the potential driver’s behaviour. This shows that the driver’s behaviour stated just as expected, careless or dangerous is not enough and should include some health-related issues that can benefit all, including the driver and the vehicle surrounding.

Presently, autonomous vehicles cannot be fully integrated within the UK infrastructure, but an isolated test is being carried out on the motorway in Milton Keynes and Coventry In the UK [105]. The current road signage and layout standards vary widely depending on where you drive in road infrastructure. A brand new dual carriageway integrating several critical environmental factors needs to be laid out before autonomous cars can work adequately [102]. Furthermore, in a narrow countryside road environment, where no white lines defining the road's edge are visible, it will be challenging for an autonomous vehicle to work correctly. With such a road marking condition, the autonomous vehicles will have to revert the complete car control to the driver, and the driver’s state should be monitored. In addition, the roads will be covered with UK national speed limit such as 60mph in certain conditions. Therefore, for an autonomous vehicle to try to reach this speed limit, there should be several factors that have to be included and considered, making sure that extreme security conditions are met before it can take on this type of challenge irrespective of the twistiness of the roads it is driving on. All these scenarios will need to be trialled before vehicles full autonomy is reached as expected by all stakeholders, including road users and manufacturers. The classifier developed will have the ability to monitor the context environment and the driver state in real-time [75].

Another critical stage in the projected timeline is moving closer to full autonomy by 2021 and having motorways where the autonomous vehicle can take complete control and allow the driver to carry out reading tasks [102]. However, in a situation where the car is off such motorways, the classifier will be helpful for ADAS systems to aid drivers when they are in control of the vehicle to determine whether it is safe to give the car/vehicle control back to the driver by monitoring the driver's behaviour in real-time. Nevertheless, there is a risk that wrong users or drivers could exploit the automated driving features by drinking behind the steering. Such case scenarios should be monitored by the software and avoid giving complete control to such a driver.

In state-of-the-art manufactured vehicles, for example, Volkswagen indicates that by 2021, autonomous vehicles can be used alongside non-autonomous vehicles on United Kingdom roads, keeping in mind that those vehicles would have restrictions as to what the drivers of the autonomous vehicle can and cannot do because of the risk of accidents due to other drivers on the road. Consequently, it will be necessary to detect and improve the system based on the dangerous driving behaviours of non-autonomous vehicles drivers and pedestrians. A far as research is concerned, two main types of self-driving cars will be on the road, including the highly automated vehicle where the driver will take control from time to time. The government believes that the existing licence and laws should still apply for this type of self-driving car in the UK. On the other hand, fully automated vehicles will need a new licence and legislation to be written to ensure the vehicle responsibility is total; however, the car owner should make sure that the software is up to date to avoid cyber-attack and software corruption [75]. Until that happens, for now, for the test on the UK road, the manufacturers and organisations wishing to carry out tests should make sure that manual override always exists with a test pilot sitting in the vehicle. Therefore, the drivers of a highly automated vehicle should not be allowed to read a book or to catch up with emails behind the steering wheel because of a probable system failure or an average (non-automated) car driver's dangerous behaviour and the subsequent fatal injuries in the case of an accident [75].

In another recent study, experts revealed that by 2025, fully autonomous level 5 driverless cars would be available, but a steering wheel will still be present, implying that there will be the possibility for the driver to take complete control in some cases of system failure. Therefore, the driver's behaviour classification and analysis will still be required [102]. In contrast, according to SAE International, the level of automation timeline suggest level 5 autonomy will not be achieved until the 2030s; the forecast in the UK for production of fully autonomous (level4/5) vehicles is expected to be in 2025, and this will be 4% of the total number of the vehicle in the country [106].

The widespread use of autonomous vehicles in developed countries will increase the number of users, leading to massive cars traffic and congestion where human intervention and control will still be required in some cases. On the other hand, the widespread autonomous cars can bring another issue related to youths avoiding taking a driving test. Indeed, it could be argued that youths are highly prone to distraction behind the steering wheel; thus, the classifier will help monitor the severity level of driver's behaviour in real-time concerning this segment of the driver population.

Nonetheless, the widespread use of autonomous vehicles is further away than is generally thought to be the case, though there will be areas of automation; however, for the majority of road networks, although technically possible, it might take longer than planned for self-driving cars to become a common thing that is widespread. Areas such as busy town centres or other complex urban scenes will most likely remain non-autonomous for longer. Developing countries such as South Africa, one of the most developed countries in Africa, will have to wait until 2040 before autonomous vehicles become a reality [107]. Therefore, the proposed classifier could be widely used in developing countries such as South Africa, Nigeria, Ivory Coast, Morocco, Thailand, etc., where level 1 to 4 automatic vehicle is scarcely used.

The position and assumption that a fully autonomous vehicle will become a reality in a decade or two in the future is a debatable and robust argument. However, the analysis of driver behaviour in terms of usage of systems such as ADAS, intelligent algorithms, etc., will be still helpful in monitoring driver states whereby autonomous vehicles give control back to the driver in some context environment such as countryside roads are narrow without lane markers. Furthermore, the state of driver or passenger in autonomous vehicles can be monitored by an algorithm such as a classifier for health benefits, including a stroke detection where, in addition to taking over from the driver, the autonomous vehicle can make intelligent decisions, such as alerting emergency services or driving itself to the nearest hospital. The algorithm or classifier will still be handy if integrated into the autonomous vehicle with those mentioned earlier. Nevertheless, humans will still make decisions regardless of the type of vehicle use. Therefore, they should not rule out the possibility to choose between the following options, 1) a manual vehicle (level 0-2), 2) an automatic vehicle (level 3 to 4), and 3) a fully autonomous vehicle (level 5).

2.5 Context-Aware Systems

According to Bolchini et al [93], a context is a process of attributing meaning to a defined environment through the experience learnt over a certain period [108]. A context is any information that can be characterized as an entity's situation [109]. Context has a significant impact on the ways machines and humans act and how they interact with things. In addition, context change can result in the change of the environment and the transformation that people will live and experience. Context is also an active process dealing with the way humans weave their experience within their environment to give it meaning [108].

Shilit and Thiemer [96] supported that the context entails information about users locations, identities, objects presents in the surrounding environment [110]. VANET discussion has been prevalent around communications mechanisms in-vehicle networks (V2I, V2V). There have been some limitations in the context-aware information, such as road types and weather, impacting a driver's behaviour. However, researchers have used synonyms such as situation, background, and situation to describe context [111]. Furthermore, a computer system knowledge representation of its user’s environment has been ascribed to be context [112].

Brown [100], stated that context is based on users' location and identifies features present in the user’s environment [113]. In contrast**,** environment features or context can also play a key role in how a user interacts with the features that can be sensed. In addition, from an attacker’s perspective, context features can dictate the taxonomy of attacks in vehicle networks. To justify those, as mentioned earlier, in network communication, the physical environment in a particular situation is the context. Context can be subdivided into three i.) location of a user, ii) with whom the user is, and iii) available resources for the user [114].

Context can be the situation of a particular place or the impact the information about an environment can have on its user [115], [116]. Therefore, it could be argued that an entity will behave following its context-aware. Thus, in this case, the driver will behave following the driving situation, and the driving will be impacted by context-aware information such as pedestrians, road type, vehicles, weather.

According to Gartner [103], Context is when something exists or happens [117]. In a traditional network penetration testing audit, one of the strategies adopted is to map out the design of the computer network before launching an attack using reconnaissance tools such as Nmap to probe the network. Dey and Abowd 1999, [106], [107] stated that computer sciences perceive context as just user locations [109]. Dey 1998 [98], defined context as information related to the features of an entity in a state [111]. An entity can be an object, a person or a place related to a present state. According to Bolchini [93], advanced context models can support context-aware applications used for interfaces.

2.5.1 Design of Context-Aware Systems

This section describes the design, architecture, and framework behind context-aware systems. Context-aware systems can be implemented using different approaches that satisfy conditions and requirements, such as sensor location, user amount, and device resources. The design of context-aware system architecture is shaped under the methodology used to collect context data [118], [119].

According to Chen [110], there are three techniques used in the acquisition of context-aware, which are detailed below:

2.5.1.1 Direct sensor access

This technique entails embedded sensor devices with information-gathering capabilities. No added layer is needed to gain and process sensor data, and the sensor drivers are integrated into the application. However, the limitation of this technique is its unsuitability for distributed systems due to the components that do not enable the management of concurrent sensor access.

2.5.1.2 Middleware infrastructure

The middleware architecture comprises layers of context-aware applications, employing encapsulation to hide low-level architectural and sensing details. The general architecture enhances scalability, reusability and extensibility and is a better approach than direct sensor access. This is because it has a modular design and employs strict encapsulation.

2.5.1.3 Context Server

The context server approach provides a distributed open-access feature that extends the middleware architecture by introducing a remote-access managing component [120].

2.5.2 Context-Aware Applications in ADAS Systems

In-vehicle components/devices, such as an onboard infotainment system, can cause driver distractions in that drivers are obliged to interact with them, leading to dangerous and unlawful driving behaviour. Hence, detecting drivers' distraction levels is a crucial part of autonomous driving and smart vehicles. In addition, driver behaviour is also affected by the driving context, with the context-aware influencing the perception and risk of the driver, underlining the need for a context-aware system that can identify and learn the behaviour of the driver in real-time. To this end, there is a need first to develop a definition of the context and what components form a context-aware application. For example, ADAS use in-vehicle monitoring to evaluate driver distraction in certain situations, subsequently alerting the drivers and passengers to inherently dangerous scenarios.

Being distracted is normal and can cause a decline in concentration, alertness, and reaction time, especially when driving. The development of ADAS has taken driver distraction into account, intending to avoid accidents and enhance road safety; the approaches used by ADAS can be categorised as either the vehicle-oriented approach or the driver-monitoring approach. Braunagel [14], showed how ADAS could allow the vehicle to take control from the driver when the situation requires it, proposes a system to classify driver distraction that will contribute to this process, especially in the case of semi-autonomous vehicles. Furthermore, Braunagel underlined the continuous responsibility of the driver in semi-autonomous vehicles. This responsibility is given to automated vehicles under certain conditions, allowing the driver to engage in secondary tasks, e.g., entertainment or resting [14], [15]. However, such a transfer during secondary tasks is regulated and not fully authorized, even in entirely autonomous vehicles.

Moreover, ADAS alerts the driver when they have removed their hands from the steering wheel for autonomous vehicles. Consequently, drivers are forced by the vehicle to take control of the driving task when necessary. Regarding the driver's readiness to perform such a takeover, Braunagel also developed an approach to ease the transition, namely driver monitoring, e.g., gaze guidance or increased deceleration [14]. Nonetheless, there are several scenarios in which the vehicles must control the driver, such as driver distraction or a lack of focus; hence, information on the severity of the driver distraction is critical. There is currently a significant gap in traffic accident risk assessment based on the predicted accident severity using Recurrent Neural Networks (RNN) [140]–[143]. Thus, instead of a mere detection approach,

2.6 Long Short Term-Memory in Driver Distractions

Predicting unsafe driving behaviour will strengthen safe driving practices. For example, research has shown that 20% of traffic accidents on monotonous roads are attributable to driver drowsiness originating from sleep deprivation. In detecting driver behaviour, both visual and non-visual features can be drawn upon. The former refers to eye movements and facial expressions, while the latter measures heart rate variability (HRV), grip pressure, and galvanic skin response [144]. Chakraborty and Nakano [145] used a driving simulator to explore cognitive distraction in drivers, evaluating normal and drivers' secondary cognitive tasks under several driving scenarios with various road conditions. The experimental results were recorded in time-series data, e.g., speed, accelerator strokes, and brake strokes, captured by onboard sensors and analysed by data mining algorithms. Yan et al [146] used a CNN to learn the features of a driver's state, e.g., mouth, ears and eyes, and then predict their state of mind. The features were captured using a training dataset comprising four activities: everyday driving, eating, phone use, and falling asleep. The Face++ Research toolkit was used to localize drivers' facial landmarks to enable feature detection. The study presented a classification accuracy of 95.56% for the abovementioned driver features [146]. Meanwhile, Le et al [147] detected such objects as phones and hands via an advanced DL approach.

The authors' proposed DL technique used Multiple Scale Faster-RCNN with an integrated standard Region Proposal Network (RPN) with maps entailing convolution feature maps, including Regions of Interest (RoI) pooling, conv4 and con3. They used data from the SHRP-2 database, leading to reduced cost of testing, improved accuracy, and independent facial landmarking. Higher accuracy was achieved by MS-FRCNN based on DL compared to the similar yet faster R-CNN. Donahue et al [148] highlighted how RNN has become increasingly crucial in interpreting images in recurrent models that have sequences, i.e. time-series data, as well as a visual representation. They also proposed using a Long-term Recurrent Convolutional Networks (LRCNs) architecture to facilitate visual recognition that can combine CNNs with long-range temporal recursion. Their architecture takes into account three difficulties of vision, i.e., activity recognition, image description, and video description, instantiating the sequential learning task of sequential input, static output $(x_1, x_2, \ldots, x_t) \rightarrow (y_1, y_2, \ldots, y_t)$, static input sequential output $\left(x \rightarrow (y_1, y_2, \ldots, y_t)\right)$, and sequential input and output $(x_1, x_2, \ldots, x_t) \rightarrow (y_1, y_2, \ldots, y_t)$. At this moment, it is possible to apply the sequential input approach to NDS time-series data, e.g., speed or acceleration.

2.7 Fuzzy Logic in Driver Distractions

By enabling designers to model system controls with high complexity, fuzzy logic offers a non-complex method to reduce the uncertainty of knowledge-based systems concretely. For example, human behaviour, such as highly unpredictable driver behaviour, contains many uncertainties and is often measured using fuzzy logic. Research has shown that 95% of traffic accidents result from driver distraction due to abnormal behaviour [150]. As driver distractions vary substantially, estimating the severity of a particular distraction event is vital in the development of ADAS [1], and a system that can well predict driver distraction would play a vital role in the prevention of traffic accidents; however, predicting driver distraction is challenging driver distractions are very difficult to predict.

In previous research, Ohn-Bar et al [151] classified the activities of drivers based on head, eye and hand movements, captured via a Multiview vision framework based on two videos, one for the hands and the other for the head. However, this study only focused on one activity as well as hand control. Similarly, most studies on driver distraction detect and recognise activities primarily through individual activities instead of taking multi-class distractions into account; this leads to a reduced strength accident-prevention system. In related research, a system employing a fuzzy-logic model based on acceleration data from vehicle dynamics (vehicle jerks) could predict the severity of vehicle crashes [4]. In contrast, the system proposed here to detect and classify multi-class distractions draws on the factors of face orientation, hand position, distraction activity and prior driver distraction. Taking such factors into account is crucial to strengthening ADAS. In addition, a naturalistic driving study (NDS) is utilised as the driving dataset in place of distractions as perceived by the driver as it offers a more precise approach to measuring relevant activities or values. It is possible to classify multi-class distractions according to their level of severity. If the driver is focused on the road, has both hands on the wheel, knows road traffic signs and weather conditions, and follows traffic laws, they drive safely.

In contrast, using only one hand to manipulate the wheel, using a phone, conversing with a passenger, glancing sideways and not focusing on the road ahead can be considered distracted or careless driving behaviour. Drivers are increasingly engaging in several such distracted driving behaviours at once, which can substantially adversely affect their driving. This highlights the urgency of developing a classification system for driver distraction severity, mainly as there is no clear distinction between careless and dangerous driving.

NDS videos comprise image sequences (frames) that describe driver behaviour, facilitating the measurement of distracted behaviour using various metrics. A method is developed to predict driver distraction by drawing on such driving data images, combining various metrics utilising Image-Based Discrete Dynamic Bayesian Fuzzy Logic (Fuzzy Logic-DDB). Using this metric to validate the driver distraction severity model enables the classification of the severity level of driver distraction in situations necessitating a semi-autonomous vehicle taking over, thus contributing to better ADAS. Sato and Akamatsu [152] showed that the difficulty of a driving task follows the driver's capacity and the demands posed by the task. An increased task burden alters the perception of the driver, thereby temporarily reducing their ability. The authors also used fuzzy logic to describe specific driving behaviours based on driver perceptions and conditions, e.g., in terms of physical space, e.g., the felt speed and relative distance, and shifting traffic or road conditions. Nonetheless, this study only considered the consequences of distraction events, not how ADAS could be improved to mitigate distractions.

Aksjonov et al [153] developed a fuzzy inference system based on simple matrix operations as a new technique to measure driver distraction when engaging in secondary tasks. The authors simulated various driver activities and evaluated the consequent performance regarding staying in the lane and maintaining the vehicle's speed. However, they only considered text messaging as a secondary distraction. Meanwhile, Aksjonov et al [154], using a separate vehicle simulator for each driver, developed a new model of driver performance based on a neuro-fuzzy inference system that can be adapted to any driver. Their model had two inputs, namely the road speed limit and road curvature, which allowed speed errors and lane deviations to be predicted. Including 18 participants holding valid driving licences in their experiment, they used an Artificial Neural Network (ANN) that had 500 neurons as well as a neuro-fuzzy inference system (ANFIS) with a membership function (MF), generating 81 rules after training. Eighty thousand nodes were gathered for the individual drivers, and the training and testing data respectively comprised 67% and 33%. They found that both ANFIS and ANN produced similar results, with the ANN performing better in prediction accuracy. Their input used three Membership Functions (MFs) while the system had two class inputs and one output, using nine rules for the fuzzy logic evaluator.

Aksjonov et al, [155] further proposed a technique to identify normal driving behaviour and evaluate errors due to secondary tasks and total distraction, using fuzzy logic algorithms to distinguish the two. The authors observed drivers being distracted by using their cell phones and measured how this affected their ability to follow speed limits and stay in their lane. They found that phone usage was responsible for 20% of driver distractions. Using the first publicly available dataset that contained more distraction identifiers than alternatives, Eraqi et al [156] developed a system to identify driver distraction. This is based on a set of CNNs that was genetically weighted, whereby the classifier set weighted with a genetic algorithm offered significantly more classification confidence. Furthermore, they examined how different visual elements, such as face or hand position and skin segmentation, can affect distraction detection. Their eventual model was able to achieve 86.64% accuracy in real-time.

Aboueknaga et al [157] estimated driver posture using the distracted driver dataset, introducing a novel system that achieved 95.98% accuracy. While their CNN algorithm classified posture according to face or hand regions, among others, they did not account for the influence of multi-class distractions, which could have a severe effect on the distraction level. Riaz et al [158] used artificial human driver emotions to develop a system to evaluate driver distraction via fuzzy logic. They proposed that emotions take priority in the driver's decision-making process. Furthermore, they developed the Enabled Cognitive Driver Assistance Model (ECDAM) to calculate the external factors and the driver's distraction level. Once driver distraction crosses a certain threshold, the model initiates two sounds to alert the driver to appropriate actions.

Meanwhile, Munyazikwiye et al [159] developed a model to predict the severity of a vehicle crash based on acceleration and other vehicle data. The crash dynamics were analysed using fuzzy logic, whereby they used the acceleration signal to create two inputs, namely car jerk and kinetic energy. They demonstrated that car jerk contributes more to the crash's severity than the vehicle's kinetic energy. Nevertheless, preventing driver distractions that affect the vehicle dynamics in the lead-up to a crash is crucial to reducing the crash's overall impact.

Upadhya and Vinothina [160] defined various distraction parameters using fuzzy logic to analyse the likelihood of a traffic accident. Their factors included driver age, vehicle speed, driver's alcohol consumption, and infotainment system usage. They revealed that various distractions contribute to accidents. However, they did not examine which distractions play a vital role, which could be considered a limitation of their research. Kim et al [161] developed a fuzzy logic system to predict and make decisions about the intentions of pedestrians based on their position, distance and direction of movement captured via computer vision. Their consequent pedestrian protection system was able to reduce the risk level of pedestrians. Despite this, there is still a need for a system that can correlate the driver's behaviour in responding to the pedestrian's behaviour. Salleh et al [162] proposed using ANFIS to create an estimation model that yields highly accurate results in various fields, including medicine, transportation and engineering. However, due to its complex structures, ANFIS is limited by its high computational cost. The authors suggested the removal of the fourth layer to mitigate the complexity.

Dobbins and Fairclough [163] used Mamdani-based fuzzy logic to define various driving context categories by monitoring driver stress. They used only two context inputs, namely traffic density and speed. They demonstrated the suitability of estimating the stress level based on human activity recognition (HAR) and the cognitive perspective via computer vision, electroencephalogram (ECG) and DL. Taken together, this underlines the possibility of preventing behaviour that can promote aggressive driving, e.g., speeding. Erdogan and Yavuz [164] further proposed integrating fuzzy logic into ADAS, specifically in lane tracking assistance, collision avoidance and Cruise Control (ACC). They drew on the monitoring of two key factors, namely vehicle speed and driver stress levels. However, their approach did not consider that high vehicle speeds do not inevitably lead to increased stress, nor did they consider other potentially confounding distractions, such as a driver talking on the phone while speeding, to get home more quickly. Such potentially influential emotions can be identified via image recognition, classifying these distractions according to their severity. The current study uses an NDS dataset with various driver activities, e.g., conversing with a passenger, using a phone for talking or texting, operating the radio. The primary focus was passenger conversation, texting, and talking on the phone, as these activities have been identified as common driver distractions. Furthermore, as these distractions can have erratic driving behaviour according to the context, multi-class distraction activity is also considered.

2.8 Justification of Metrics

2.8.1 Face Orientation

Dong et al [17] measured drivers' fatigue from facial expressions and eye activity using physiological features. One of the metrics used is the number of times drivers touch their faces. It was stated that when drivers are tired, they exhibit less frequent head motions. Thus, measuring the frequency of face turns during a journey, the beginning duration depends on deduction from a consecutive number of frames. Hu et al [18] stated that careless driving is a significant cause of road accidents and tracked using images' face orientation and facial features. Infrared image technology was used in face region detection and facial feature detection. However, a single driver distraction is used, which is a limitation.

Sato et al [19] inferred from driver body information states using driver distraction state, concertation state, and distraction state. The measures were looking out for near misses when approaching an intersection. Time-series of different eye-gaze movements and face orientations before the collision near-miss was logged. Rasouli et al 2018 [20] analyzed pedestrian behaviour at crossing points under various weather and road types. Their findings show that there is a strong correlation in the head orientation of pedestrians before crossing intention. Pedestrians make an inference about traffic dynamics (vehicle seed), crosswalk (width), and pedestrian demographic impact pedestrian behaviour after the initial purpose of crossing has been displayed. The result shows interrelation in context elements, and one factor may decrease or increase the influence of other factors.

Fasanmade et al [21] used a multi-class distraction to classify driver distraction into severity levels using driver physiological features. The approach involved the use of an image processing rule-based fuzzy logic system. It is found that a combination of face orientation and eye glance does increase the degree of driving distraction. Furthermore, results show that a driver's distraction can transition from careless driving to dangerous driving when a certain threshold is reached, and multi-class distraction occurs.

2.8.2 Hand State

The hands are vital in the perform driving tasks such as controlling steering and changing gears. During operating, the state of writing is critical in every changing context. Thus, monitoring its state is crucial in the prevention of accidents. Das et al [22] introduced a naturalistic driving study using bounding boxes and hands annotation to detect driving hands. The validation checked for false positives that may arise from illumination conditions, no-hand objects of similar colour, occlusion, and truncation. For the detection, Aggregate Channel Features (ACF) was used as the detector, and the hand detector's accuracy was measured using precision call (PR) to evaluate parameter performance. The initial results suffered from missed detections and false positives; however, cross dataset comparison yielded better accuracy. Dong et al [23] stated that fatigued drivers assume more comfortable hand positions on the steering wheel. However, Carsten and Brookhius noted that the impact of cognitive distraction on driving performance differs from visual distraction. Visual distraction adversely impacts a driver's steering ability and lateral vehicle control, particularly car following.

Le et al [24], [25] used a novel multiple-scale region-based fully CNNs (MSRFCN) for human regions detection in illumination and low-resolution conditions. They used a pre-trained network called the "Oxford" hand dataset and compared it with several hand detection approaches. The proposed MS-FRCN algorithm achieved an average precision and average recall of 95.1% and 94.5%. Besides, there is an improvement of AP / AR of 7% and 13%, respectively, to classify left and right hands.

2.8.3 Eye Glances

According to an article by Taub [26], just two seconds of looking away from the road can be the difference between life-or-death. According to Peugeot's latest eye-tracking study, Car drivers take their eyes off the road in a one-hour trip about two miles when travelling in urban traffic. The French carmaker observed several drivers on 25 parallel six-mile drives, utilizing driver-used special glasses to investigate precisely where their eyes appeared when operating a range of SUV styles. The outcome results found that drivers' eyes off the road constitute about 7%. A one-hour driving speed of 30 mph is equivalent to driving 3,350 meters with an eye glance off the road [27].

Yuan et al [28] suggested classifying existing driving situations and forecasting off-road vehicle situations using Hidden Markov Model (HMM). The experiment was done using a driving simulator analysis involving 26 driver participants in three driving scenarios rural, urban, and motorway. Three different occlusion durations (0-s, 1-s, and 2-s) were added to measure the eyes-off-road durations. Results revealed that existing driving situations could be optimally defined using glance position sequences, with up to 89.3% accuracy. Moreover, the motorway was distinguishable with over 90% precision. Moreover, in the driver's eyes-off-road period estimation, using HMM-based algorithms with two inputs as look duration and look position sequences gives the highest accuracy rate of 92.7%.

Vehicles of 42 newly approved adolescent drivers are fitted with sensors, accelerometers, Global Positioning System(s) (GPS) to collect data continuously for 18 months period. Crashes and near-crashes (CNCs) situations were reported through the investigation of significantly elevated gravitational force incidents. Analysis of video has a duration of 6 seconds previous to each CNC, and randomly sampled non-CNC road fragments were coded for the period of eye glances off the front road and occurrence of secondary mission participation. The chance (odds ratio) of CNC due to eye-glance activity was determined by contrasting the prevalence of secondary task participation and the length of off-road eyes before CNC with the prevalence and period of off-road looks non-CNC road segments. Crash incidence improved with the period of single most prolonged glimpse during all secondary tasks (OR = 3.8 for >2 s) and wireless secondary task presence (OR = 5.5 for >2 s). The single most extended glimpse offered a constant estimation of an accident's likelihood than absolute eyes off the forward roadway [29].

2.8.4 Road Type (Urban, Highway)

Doshi et al [30] developed an algorithm that includes critical vehicle data such as the status of brake switch, throttle position, and wheel speed; and uses the inputs in calculating several parameters, namely; shifts per a given time interval, throttle variations, mean velocity, acceleration among others. The resulting calculated parameters help the algorithm identify the road type on which the assigned vehicle is travelling. This road identification process is achieved through parameter comparisons with reference values that define various road types. Additionally, the algorithm identifies a driver type using driver inputs such as gear shift patterns and a driver's handling of brake and accelerator pedals. Doshi's algorithm attained a "Receiver Operating Characteristic (ROC)" value of 85 % accuracy in road-type identification.

Chai et al [31] conducted a behavioural analysis on road rage in China. The study showed an inverse proportionality between cases of road rage and lanes number on a given road. Thus, with fewer lanes, there are higher incidences of road rage. Additionally, the study revealed that road rage increased with an increase in the number of non-motorized vehicles. Road rage involved fewer trucks; daytimes had fewer incidences involving non-motorized vehicles, while more trucks were involved in road range on highways. A limitation in the study is that a small sample size of data was used, lacked demographic and environmental variables, necessitating a more detailed analysis in the future. To characterize road types and measure the degree of aggressiveness of drivers. Messeguer et al [32] designed and implemented a neural network-based algorithm to assist drivers by pointing out unacceptable driving behaviours as offering driving tips that would help improve fuel economy. Test results proved neural networks' ability to achieve a degree of precision in the classification of driver and road types.

With context-aware playing a critical role in the accurate performance of various road classification and driver distraction identification algorithms, the useful context-aware collection is vital. Rakotonirainy et al and Khan [33], [34] proposed a context-aware system for real-time collection and analysis of context-aware; related to a vehicle, the immediate environment, and the driver. The system also gathered information from filled questionnaires. A Bayesian network model was employed to analyse the context-aware through a learning model, facilitating the observation and prediction of a driver's future moves. The model attained a high accuracy in predicting a driver's future behaviours and warning other road users.

Methods for recognizing and classifying road traffic accidents' severity play essential roles in understanding accidents, causes, and possible mitigation strategies. To that effect, Jianfeng et al [35] designed a set theory-based accident recognition and classification method; that supports vector machines. Their model employed rough set theory in calculating the significance of driving environment, road, vehicle, and human attributes, with their results, show the model's ability to improve recognition accuracy and reduce computational workloads. This chapter's limitation is that they have not considered human physical behaviour, i.e., the driver's face orientation.

2.8.5 Weather

Cai et al [13] developed a travel weather warning system (TWWS) similar to the Road Weather Information System (RWIS) for sharing weather safety information and disseminate safety warnings to drivers. This system is made up of risk estimate models that are based on extensive weather-related crash data. Weather-related data were collected using questionnaires, where drivers identified various risks while driving under different driving conditions. The severity of each wager is measured on a four-point scale that ranges from slight to catastrophic. Metrics such as the intensity of rain, traffic volume were considered. Malin et al 2019 [36] stated that rainy weather is a significant factor in traffic incidents, and the risk of accidents increases with poor road weather conditions. Sherretz and Farhar [37] stated a positive linear correlation between rainfall and the frequency of road traffic crashes. Bergel-Hayat et al [38] revealed a significant correlation between an aggregate number of traffic accident injuries and weather variables. They observed that the correlations between these two parameters varied depending on road types.

Brodsky and Hakkert [39] proposed measuring the risk of a road accident during rainy weather. The method shows a drastic rise in road traffic accident injuries during rainy weather compared to dry weather. The increased dangers, mainly when wet conditions follow a long dry season, are well known to drivers, as found by Knapper [40], who, through sampling, found that drivers were aware and could recognize the risks. In assigning weights associated with weather conditions, expert opinion is needed on whether there is a correlation in weather scenarios with collision studies.

2.8.6 Speed

Maintaining correct speed continues to be a challenge to many drivers. It is said that drivers who perform abuses of driving, such as speeding, crash more. Stradling and Auberlet [25], [26] illustrated that vehicle trajectory variations may disclose valuable details on how spatial restrictions impact the behaviour of drivers (e.g., lateral location and speed). The findings revealed that the lateral location variability was more significant on the one side while driving on the crest vertical curve measure before encountering oncoming vehicles and narrower lane width. On the other hand, it was reduced according to the perceptual procedure used. Another study investigates the impact on drivers' vehicles' speed perception of multiple factors such as image size, speed, road shape, driving experience, and gender.

Wu et al [164] examine the most miniature image scale (38% of the actual field of view), and speed calculations were the most reliable. The driving velocity was gradually undervalued as the image scale grew. Participants with driving expertise correctly measured the driving speed on both wide and narrow roads. However, those without driving experience had more considerable underestimates on broader roads. Furthermore, environmental conditions concerning speed performance have been highlighted by Bellis et al [165], who challenged the current policies and suggest they can intervene through teaching drivers about the relationship between inverse illuminance-speeding and measuring how better vehicle headlights and intelligent road lighting will attenuate speed. The real-world's speeding actions and its association with illuminance, an environmental property described as the incidence of luminous flux on a surface. Manser and Hancock [166] addressed the need to ascertain if the visual pattern and wall tunnel texture impact driving performance since maintaining correct speed continue to be a challenge to many drivers. The findings show that the relationship of speed by drivers and their reaction is impacted by the visual pattern and the tunnel walls' texture.

2.8.7 Vehicle

Mishra and Bajaj and Kamar and Patra [46], [47] used the ML technique to predict drivers' driving patterns and their impact on social behaviour using CCTV cameras installed to monitor traffic. The observation was carried out during the day, and the metrics for measurement were instances of traffic violations due to aggressive patterns. Lee and Kum [48] proposed a feature-based lateral position estimation algorithm, which employs lateral positioning and stereo vision, irrespective of changes to viewpoints and obstructions -

resulting from pixel-wise feature extraction. The algorithm extracted vehicle images through image filtering, thresholding, and removing the ground portions from images captured from cameras. The algorithm's detection component employed a deep CNN with a speeded-up robust feature (SURF) to match successive image frames. They estimated the lateral position of ground points involved an inverse perspective mapping (IPM) algorithm. The testing and validation were done using urban and highway to attain zero mean error and standard deviation of 0.25m in lateral position estimation.

Xu et al [49] detected driver behaviour using car-following behaviours, which could change due to distraction, fatigue, drivers' habit, and surrounding traffic. On-Road trajectory data obtained in Beijing was used, and as a metric, distinctive driver states and car-following models were observed. This led to the prediction of the driver's velocity control with improved accuracy. Mittal [50] used object detection and a faster R-CNN model to detect different scale and size vehicles. An evaluation was performed using the FLIR_ADAS dataset for both RGB and thermal images. Gong et al [51] proposed using the YOLOv3 algorithm to detect the vehicle in thermal images. This led to a 65% higher accuracy and speed than the original YOLOv3-tiny.

2.8.8 Pedestrians

Kharjul et al [52], Introduces the implementation of an active protection automobile pedestrian identification device to minimize the amount and intensity of vehicle-pedestrian collisions. The authors present a pedestrian identification approach dependent on photos in this framework to segment pedestrian candidates from the picture. The method used is the Ada-Boost algorithm and cascading algorithm. They are confirming whether each claimant is a pedestrian. The Support Vector Machine (SVM) is specialized in identifying classifiers. The system sends input features mined from both the sample grey images and edge images to the device used for SVM training. Taiwan and Yamada [53], [54] developed a tool for calculating a driver's knowledge and behaviour about pedestrian's position at the crosswalk and cross, especially at the direction of a left or right turn at an intersection. An appraisal carried out using objective evidence on automobiles' driving behaviour on public roads has been published.

Rangesh et al [55] examine the behaviour of pedestrians instead. In particular, from a solely vision-based point of view, the authors concentrate on detecting pedestrians engaging in secondary behaviours involving their mobile phones and related hand-held multimedia devices, suggesting a pipeline integrating articulated human pose prediction and utilizing gradient-based picture features to detect the presence/absence of a smartphone in either hand of a pedestrian. A belief network encodes knowledge from multiple streams and their dependency on each other. This network is then used to forecast a likelihood score that suggests a subject's engagement with the device.

Phan et al [56], Whenever a person emerges in front of the car, the authors intend to research the driver's actions. Also, two static parameters-based methods, which include Necessary Deceleration Parameter and Time-To-Collision were included in the problem and compared to the proposed approach, a technique applied to driving behaviour using the Hidden Markov Model (HMM) algorithm is used in characterizing the driver knowledge of pedestrians and the driver unawareness of pedestrians. Compared with basic ones, the outcome indicates a significant enhancement of the HMM-based process.

2.8.9 Illumination (Day, Night)

Clarke et al [57] observed that the rate and severity of road traffic crashes are influenced by data time. In their study, the visibility conditions under investigation included rainy and night driving, with the control test being dry daytime driving. Their findings on the increased rate and severity of crashes at night and rainy weather correspond with the conclusions of [58], which shows the risk of fatal crashes increased by a factor of four on night driving, as compared to daytime.

2.8.10 Passenger talk

Hole [166] asserted that talkative passengers appear to be less distracting than phone conversations, possibly because this passenger also functions as a second pair of eyes for the driver, thus moderating the degree of their interaction in the event of road hazards. Hence, less weight is assigned if the driver’s face is oriented towards the road. However, if the face is oriented away from the road to converse with the passenger directly, more weight is assigned, although less than for texting or phone usage, as outlined above. Reviewing empirical studies from the 1968-2012 period, Ferdinand and Menachemi [167] created a logistic regression model to reveal any relationship between drivers engaging in a secondary task and their performance. They revealed that around 29.2% of driver distractions could be attributed to conversations with a passenger [167].

Meanwhile, Foss and Goodwin [168] explored the issue of driver distraction in adolescents. They collected data on 52 high school student participants via unobtrusive event-triggered data recorders that, when triggered, captured 20 seconds of audio, video, and vehicle kinematic information. They revealed that the largest single source of distraction was electronic devices, at 6.7%, followed by adjusting the vehicle's controls, at 6.2%, and grooming, at 3.8% [168]. The authors also estimated the driver distractions using a statistical approach, identifying, and counting how many distraction events occurred.

Overall, there is a consensus that most driver distractions emanate from three different sources: physical, cognitive, and visual activities. Physical activities include using a phone, texting, and operating an infotainment system. Meanwhile, recognising distractions that could affect drivers' cognitive abilities is crucial as these substantially impact drivers' decision-making. Texting, which could be considered a visual activity, is one such distraction. In addition, while driving is simultaneously a visual and cognitive activity, the visual aspect is paramount to perception or decision-making. Cognitive distraction can encompass conversing with a passenger or using a phone, while the subject matter and nature of the conversation can also have a considerable effect. It is possible to have a multi-level distraction that includes all three inputs; this strengthens the distraction's severity level and degree. For instance, texting implies all types of distraction input happening concurrently, thereby representing a considerable threat to the driver's behaviour. Distraction can also vary throughout the journey and can be measured using time-series data to assess the distraction frequency and duration in addition to the driver's engagement level with the distraction's source.

2.8.11 Texting

The NHTSA highlights texting as the distraction that most severely contributes to traffic accidents. Using a test case, the NHTSA demonstrated that 5 seconds of texting is equivalent to the driver closing their eyes while driving across a football field at 55 mph [169]. In light of this, Madden and Lenhart [13,14] showed that 28% of teens in their survey reported using mobile devices while driving, which critically affect their ability to drive. While 52% of teens said that texting while driving is not very common, they reported using a phone to conduct a conversation while driving. The survey findings underlined the dangers of taking one's eyes off the road to text or otherwise use a phone.

2.8.12 Phone Usage

Meanwhile, Hole [166] also found that hands-free phone usage represents a distraction similar to holding a phone because drivers visually imagine what is being discussed in the conversation. The author showed that the discussion type significantly impacts the driver's mental processing and facial expression, potentially raising the distraction level. The research used the duration of use, type of discussion, and frequency of use during the journey. In contrast, an in-person conversation uses multiple non-verbal cues that reduce the mental demands of the conversation than it would be if held over the phone. Specifically, conversations over the phone are typically a lot more stressful as the need to imagine the discussion visually puts extra demands on the brain's processing capacity, meaning that the driver may miss road hazards.

Finally, Drews et al [172] explored how phone conversations and conversing with passengers while driving differed, specifically how drivers can tackle driving demands while holding either a phone conversation, an in-person conversation, or when there is no distraction. They demonstrated that more errors occur when using a phone compared to conversing with a passenger, in particular, a phone conversation impacts the driver's abilities, with their speech coordination reducing as part of a response to increased traffic demands. This indicates that conversations with passengers are unlike conversations via the phone, this is because the traffic conditions can become a conversation topic, thereby helping both vehicle occupants increase their awareness of their surroundings and the driving conditions themselves also directly affect the conversation, i.e., its complexity, thus mitigating the conversation's adverse effects on the driver's focus. In the current study, the weight of data according to the activity's potential risk; thus, texting is considered the most dangerous, followed by using a phone and conversing with passengers. Nevertheless, there is a need to consider the possibility of an instance were talking to a passenger is combined with an additional distraction, thus representing a risk level equivalent to that of texting.

2.9 Risk Assessment Analysis in Drivers Distraction

Since newer technology, distracted drivers are one of the most significant problems occurring in road-related accidents. Intelligent transportation will soon allow vehicle takeover to semi-autonomous level 4, i.e., when out of necessity or by option, the vehicle takes control from the driver to commence driving activity. With vehicle takeover forthcoming, drivers will become heavily reliant on allowing the vehicle to perform more in-vehicle tasks, drivers will become more relaxed, and distractions will occur more often, opening too many risks to the driver. Moreover, relevant context vehicle information will be utilised to help the driver. Such context-aware accounts for many areas relevant to the driver, including vehicle performance and environmental conditions, which directly affect the driver's safety. This implies a need for an ADAS that can mitigate the risks before an accident occurs and provide a qualitative- and quantitative-based risk assessment.

An article by the European Commission for Mobility and road transport safety presents the fact that a significant percentile of road accidents occur when a driver is distracted, with common distractions such as handheld mobile devices, using the radio, eating, talking to passengers, smoking and glancing at in-vehicle navigation systems [1]. According to Kulkarni and Shinde [2], in-vehicle interfaces can also cause an overload to the driver. Additionally, fatigued drivers present a significant risk on the road. In recent years, drivers' eyes became an efficient metric for measuring driver distraction, and driver's ability in placing their eyes on the road during driving is crucial. A statistical analysis by the Department for Transport shows that out of the 1,456 cases in fatal car accidents, 383 of those cases involved careless tendencies by pedestrians, while 110 of the cases resulted from drivers' reduced attention on the road [3]. Inexperienced drivers are a significant factor accounted for causing the number of road accidents to surge. Young inexperienced drivers are particularly at risk, while skilled drivers may change their tactics in good time and predict different driving scenarios [4] [5]. In comparison, the higher crash incidence by young drivers is attributed to low cognitive ability [6] and a loss of attention due to distractions [7].

Furthermore, if knowledge transfer – particularly driving perception – were transferable from experienced to novice drivers, expectations would be that the novice drivers would better identify and mitigate driving risks, translating to lower crash incidences. Risk mitigation is difficult to model as official accident reports are relatively undetermined due to the possibility of numerous definitions of distractions or a country simply not collecting the data [8]. Disparagingly, driver's distractions can be impacted by the context-aware situation in which driving occurs. Thus, a novel context-aware risk model which uses intelligent image recognition to detect and form a risk matrix to profile drivers into distraction classifications can reduce the occurrence of an accident. The drivers can be classified into three groups, i.e., safe, careless, and dangerous. Capturing the driver's behaviour is crucial in risk mitigation, developing context-aware ADAS systems that may influence the risk levels and prevent accidents. Moreover, a real-time novel risk assessment determines a driver's risk profile and the development of the driver's distractions to work with multiple driving context influences such as auditory, visual, cognitive, and biomechanical distractions simultaneously.

Risk assessment can be defined as an evaluation process applied in evaluating adverse effects that arise from a natural phenomenon, an activity or a substance [9]. Benedict stated that risk constitutes the likelihood and probability of the incidents [196], [197]. Relative Risk Ratio has been used in quantifying vehicle crashing risks under bad weather conditions; this requires a large dataset of crashes arising from adverse weather conditions [198]. However, using a risk matrix that combines probability and consequences has overcome the former method in popularity [196]. A risk matrix can be used to determine the level of driving risk. Understandably, most risk indicators related to drivers' distractions has been modelled after the crash event. According to Cai et al [156], certain studies have shown that driver's subjective assessments of driving risks - particularly those related to various weather scenarios - are consistent with collision-based studies. In [156] the authors assumed that the driver's perceived risks are consistent with the actual crash statistics; specifically for incidences related to rainy conditions. The main flaw of modelling driving risk assessment via post-crash data is the fact that it is a reaction strategy rather than prevention.

Furthermore, different factors could impact the driving capability that could be extracted from the driving context. The driving context-aware that impact the driver can be from the driver, vehicle, and environment. The context-aware comprises weather, road, speed, manoeuvres, pedestrians, drivers state, braking. However, inadequate data and facilities ensure an efficient and robust risk assessment model for driving context. This research proposes using a Naturalistic Driving Study (NDS) TeleFOT that is complete enough for the environment, vehicle, and driver monitoring. The proposed approach[64] used the mathematical model as:

$$C_i = \sum_j^J x X_{ij}\, \beta_j + \varepsilon_i, \quad C_i = 1,2, \dots \dots, M \qquad (2.1)$$

Where (C_i) denotes a discrete model-dependent variable that represents the level of distraction's impact on driving. This variable's various impact levels include minor impact, overall impact, profound impact, and disastrous impact. The 'i' included in this variable represents the i^{th} driver with non-observable ε_i Variables may include the volume of traffic, vehicle type, road type, and rain intensity. A non-observable variable was selected to fit a logistic distribution for generating a continuous latent variable (C_i), denoting the influence on driving.

Another proposed approach is the Rank Order Cluster Analysis adopts that driving risk R_i is sorted in ascending as indicated by $R_1, R_2, R_3, \ldots\ldots R_n$. Consideration of categories (G), including $R_i, R_{(i+1)}, \ldots \ldots R_j$ and satisfying $j > i$, which can be denoted as $G = \{ i, i+1, \ldots\ldots, j\}$. Consequently, diameter of $G, D\ (I, j)$, is calculated from the equation:

$$\mathrm{D(i,j)} = \sum_{t=1}^{j} (R_{(t)} - R_{(G)})^2 \tag{2.2}$$

Where R_G represents the mean driving risk, where the driving hazard is segmented into k segments expressed as:

$$G_1 = \{ i_1, i_1, +1, \ldots\ldots, i_2 - 1\}, \quad G_2 = \{ i_2, i_2, +1, \ldots\ldots, i_3 - 1\} \ldots\ldots \quad G_k = \{ i_k, k, +1, \ldots\ldots, i_{k+1} - 1\}. \tag{2.3}$$

Where the variable i satisfies the condition: 1 = { $i_1 < i_2, < \ldots < i_k, < i_{k+1}$ = n+1.

There is also a minimal loss function with a recursion relationship represented by the formula:

$$\mathrm{L\ [b(n,k)]} = \sum_{t=1}^{k} D(i_{t,} i_{t+1,} - 1 \tag{2.4}$$

Where b(n,k) denotes a special classification method:

$$\mathrm{L\ [b(n,k)]} = \min_{k \le j \le n} \{L[P(j-1, k-1)] + D(j,n)\} \tag{2.5}$$

Furthermore, $P\ (n, k)$ denotes the method to minimize the loss function. Where n and k are given, $P\ (n, k)$ depicts the optimal driving risk categories.

One of the patents held by MOVON Corporation [64] is a method to ensure the safety of drivers using a lane departure warning system based on image processing using a mono camera installed inside the car. A distinctive feature of the system is that it successfully processes several road conditions, including undesirable situations such as changing the width of the road lane, the radius of its curve, the direction of the road and the complete absence of a road surface.

The following observable parameters can characterize signs of attention deficit and fatigue in the driver: PERCLOS (PERcentage of eye CLOSure - a percentage of time the driver's eyes are closed) [65], turning the head to the left/right to the body, tilting the head forward relative to the body (the moment when the driver "nodding off"), duration of blinking of the eyelids, the frequency of the blinking of the eyelids, the degree of openness of the person's mouth (signs of yawning). In particular, for PERCLOS, there was a discrete number of parameters defined, namely: P70, i.e., 70% of the time the eyes were closed; P80, i.e., 80% of the time eyes were closed; and EYEMEAS (EM), the mean square percentage of the eyelid closure rating [65]. Furthermore, general information describing the vehicle driver helps explicitly identify the driver among all drivers who installed and using a particular monitoring software package but also helps to improve the search and ratio drivers with similar characteristics (general patterns among groups would help predict developing situations). This can be accessed via database, with weight coefficient applied since this is a “common” behaviour, but not this individual driver’s behaviour.

Ginting H et al [59] adopted the Likert scale in modelling individual coronary heart disease anxiety into different levels. The scale was used in implementing a 5-pointer scale. Lopez-Fernandez et al [60] used a scale in assessing problematic internet entertainment use scale for adolescents. The scale adopted is a self-administered scale for measuring behavioural addiction of online social network users and video gamers to the degree of severity. Drawing from the knowledge, this study formulates the distraction severity levels. Furthermore, the ratings of the severity level of distractions are designed using a 5-point scale as seen in Table 2 Driving Severity levels below, deduced using the Likert Scale [61]–[63].

Table 2-1: Driving severity levels

Consequence	**Severity (0.0 – 1.0)**	**Risk Colour**	**Severity Levels**	**Distraction Class**
No Distraction is observed.	0.0	Light Green	No Impact	Safe
A Slight Distraction Observed	0.1-0.25	Green	Slight Impact	Safe
Noticeable Distraction	0.25-0.399	Yello	Low	Safe

		w		
Substantial Level of distraction detected	0.4-0.599	dark yello w	Medium	Careless
Frequent level of distraction	0.6-0.79	Orang e	High	Dangerous
Casualty Prone	0.8-0.9	Dark Orang e	Very High	Dangerous
Severe Casualty Prone	0.9-1.0	Red	Extreme	Very Dangerous

2.10 Summary

In this chapter, a complete review of the state-of-art context-aware safety systems related to driver distractions. Throughout the chapter, different terms and themes were reviewed, giving a clearer idea of the work undertaken by others in the area to understand the research field better. For example, DL, deep vision, and computer vision were scrutinised, with some main algorithms developed in the field, leading to some up-to-date applications in research, industry, and all aspects of everyday life. The next chapter deals with the methodology chapter, where the focus will be on the methods, including technologies and algorithms used to complete this thesis.

CHAPTER 3. METHODOLOGY DEEP LEARNING AND COMPUTER VISION

3.1 Methodology Overview

The term methodology requires adopting a common approach that involves research leading to the research design [204]. On the other hand, research methodology is viewed as a strategic move favouring the research outcomes [205]. Wahyuni [206], supports that depending on the aims and objectives pursued by the research undertaken, the methodology chosen varies, and one must choose the methodology that best applies to his objectives. It is essential to mention that concerning research methodologies, two main choices exist. This includes the qualitative research methodology and the quantitative research methodology. The qualitative research methodology is related to research dealing with more theoretical assertion, with scarce or no use of experiment with logical steps leading to a conclusion using software to research outcomes. Conversely, [207] sustains that the quantitative research approach is more related to the inductive approach to quantitative study than empirical research studies, where the focus is on understanding human behaviours. Through this approach, experiments are completed, protocols are followed, numerical data collection is performed, performances are measured throughout the study, and analysed the results [208] .

For the current research, the quantitive research methodology applied to the context-aware safety system for detecting and improving dangerous drivers' behaviour related to their distraction levels is chosen to accomplish this research. The rest of this chapter has the following structure: section 1) relates to the introductory part of the chapter, section 2) deals with the research methodology, where, after a brief definition of the chosen methodology is presented, section 3) offers a clear presentation of the research methods applied, section 4) highlights the research design to be implemented, section 5) provides a succinct summary of the whole chapter.

3.2 Research Methodology

3.2.1 Definition

The notion of methodology refers to utilising research approaches and methods that lead to the steps applied in the research overall planning, including the research design and mainly objective dependent. Therefore, the research methodology chosen will depend on the aim of your research; hence, it will vary based on the above, which also means that the researcher should select the research methodology that leads to the optimum outcome. Concerning the research approach, the author has shown that two main approaches exist, including qualitative and quantitative research approaches, and it is down to the researcher to select the one that applies to it. That means that the researcher should address the choice of the methodology that will support the design and implementation of his research strategies. On the other hand, different authors have established that the research method is related to the techniques and tools used to investigate a given topic. This includes using either an experiment, case study, etc., as a research method to carry out a research study [209].

3.2.2 Quantitative Research Approach

The current thesis discusses utilising Artificial intelligence and ML techniques to classify driving behaviour in the current thesis. Research has shown that different research approaches can be used to carry out a study [207]. However, there are two primary questions and issues to be solved for this research. The first question will address how driving behaviour is classified, while the second will relate to measuring the degree of driving behaviour. The proposed research approach will adopt naturalistic driving data obtained from United Kingdom Field Operational Test TeleFOT [210] and will be further analysed using experimental research to guide research. There are two possible steps involved in the demonstration of drivers’ behaviour. The first step is related to the analysis and evaluation using simulations, and the second step is dealing with the analysis and verification using operational field testing (FOT) [211]. Driver’s behaviour detection has mainly involved simulation measuring of traffic events, braking event that does not provide adequate context-aware that will not fully explain drivers' perception of driving in a naturalistic environment. Implementing an actual vehicle is very expensive and will involve human participation; thus, there is a need for an alternative approach [211]. In addition, physical implementation will entail safety and ethical issues. However, research and the government has invested a lot in transportation systems. The government and other consurtiums has sponsored naturalistic driving data such as TeleFOT and UDRIVE [210], [212], [213]. An experiment involving the combination of Simulation and naturalistic data from Field Operation Trust (FOT) will be adopted to satisfy the research question.

3.2.3 Rationale Of Research Approach Selected

The quantitative research approach was chosen for this thesis for different reasons; the main ones are related to the fact that empirical research was conducted. Hence data were gathered, process and analysed. Some of the steps involved are highlighted below:

First, quantitative data were collected, including several driving hours, mainly using video recording or obtained from United Kingdom Field Operational Test TeleFOT [210]. The reason for this data extraction was to predict drivers' behaviour in different driving case scenarios.

Second, quantitative data pre-processing was carried out. Indeed, techniques were used to convert video recording into images or vice-versa.

Third, quantitative analysis meant evaluating measurable and verifiable data to predict driver distractions [208]. Mainly, drivers' behaviour detection also meant the involvement of simulation to measure traffic events, driving in a naturalistic environment with less context-awareness information available. The drivers' behaviours detection can be done through the collection and analysis of driving data.

3.3 Research Method (Simulation Research-Based)

3.3.1 Research Method Selected

The research method utilised in this thesis is the simulation research method, which is supported by different reasons. Indeed, a secondary naturalistic driving study data was utilised. Extracted, were different drivers selected that drove in different weather conditions with various driving styles. There were different categories of drivers selected, drivers qualified as "experienced" for those driving more than once, and "novice" for those driving only once were used to test different cars. In the NDS data adopted the driving sessions were recorded using inside and outside cameras, and the videos data enhanced recorded version will be used to study a context-aware safety system for improving dangerous driver's behaviours. In addition, the videos data were converted into several images, which were then cleaned and analysed using different image processing techniques.

3.3.2 The Rationale of the Research Method Selected

The experimental research method was selected for different reasons that are highlighted below.

First, an extraction of data from the NDS was performed selecting the drivers with different driving styles in various weather conditions.

Second, videos data collected were converted into image data thanks to different image converter software such as ImageJ, FFMPEG, Video Proc, VLC player, etc.

The research method employed is based on the computer vision in intelligent transportation systems (ITS) that entails collecting and analysing data from videos and images used in decision making. The initial stage is the object detection and recognition stage, where algorithms can track the object in motion. The level of inference on the relationship between an object (human-object interaction, human behaviour, multi-objects) and context-aware is critical in decision making. The computer Vision domain of AI falls within the following areas: perception, visual sensing, and reasoning. In the intelligent transportation systems (ITS) context, drivers' reactions to the environment, such as in-vehicle and the outer vehicle, can determine their behaviours on the road. Therefore, the behavioural study will be developed based on collecting different driving sequences of several drivers. The data thus extracted will generally be video data that can later be converted into different formats, such as the conversion of videos into images, etc.

3.4 Pilot Project: Image-Based Driver Activity Detection

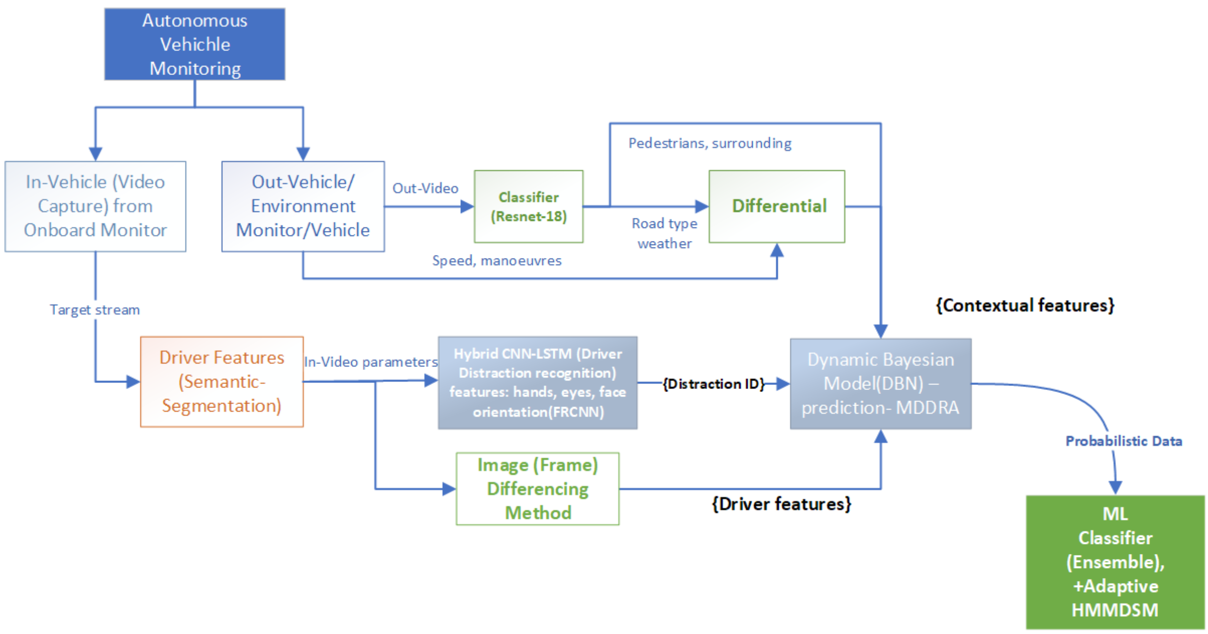

Figure 3-1 CNN-LSTM Process Flow

This section describes the experimental design and data processing for driver behaviour recognition. A CNN LSTM architecture uses a CNN for feature extraction on a given set of input data, combined with a Long Short-Term Memory Network, which supports sequence prediction. By design, a CNN LSTM was intended for handling visual time series prediction and generation of textual description from a given input sequence of images and video frames. CNN LSTM-DBN handles activity recognition through the generation of textual descriptions of activities identified in sequences of images. The process flow for a CNN LSTM network comprises ten stages, each represented by a box in figure 3-1 above. The process flow starts with autonomous vehicle monitoring and ends at the ML classifier; in between, two process flows make up the internal operations of the model. The two flows represent the in-vehicle and out-vehicle flows. Each of the process flows are as follows.

3.4.1 In-Vehicle Process Flow

- **In-vehicle video capture:** the first stage of this process flow involves collecting the vehicle's interior data with a video from the onboard monitor.
- **Semantic Segmentation:** the second stage of the in-vehicle process flow entails semantic segmentation of the driver's features. This component extracts driver features such as driver activity, number of hands on the wheel, and face orientation off the road. Extracted driver features are then fed into a hybrid CNN-LSTM model as in-vehicle parameters.
- **Hybrid CNN-LSTM:** the in-video frames from the semantic segmentation section are fed to the hybrid CNN-LSTM model. The CNN layer performs feature extraction on input data, while the LSTM performs sequence prediction and activity recognition. The model is tasked with identifying the type of distraction a driver experiences. Any identified activity is compared with historical data to recognise the distraction type and give it a distraction identifier fed to the dynamic Bayesian network. In essence, the hybrid CNN-LSTM performs driver distraction recognition by analysing the extracted driver features where fuzzy sets for classification of the distraction by severity level are extracted. The fuzzy sets of distractions inform the model of the distraction ID, which is then fed to the Dynamic Bayesian model together with driver features, extracted by the deferential stage of the model.

- **Image Frame Differencing:** a copy of the video stream from the semantic segmentation section is fed to a frame differencing component, identifying and extracting the driver’s features.

3.4.2 Out Vehicle Environment Monitoring Process Flow

The second process flow making up the entire CNN-LSTM Process Flow is the external

- **Out-vehicle/Environment monitors**: this component collects data about the vehicle and the vehicle's external environment, such as speed, manoeuvres, and a video recording of the road and pedestrians. Two streams of data are obtained here, external video and vehicle data which includes speed and manoeuvres.
- **Fast R-CNN:** the video recorded from the outside of the video is fed to a faster R-CNN, which analyses the frames to extract information related to road type, weather and identify pedestrians and the surrounding. Fast R-CNN detects regions that have objects of interest.
- **Differential:** the differencing component receives data about driving speed**,** road type, weather and driving manoeuvres. The component relates the different variables; speed, road type and manoeuvres to generate critical context-aware.

3.4.3 Dynamic Bayesian Model

The dynamic Bayesian model takes in three variables; distraction ID and driver features from the in-vehicle monitoring stream and context-awareness from the out-vehicle streams. With the three key inputs, the Dynamic Bayesian model performs severity classification by relating the variables to each other over adjacent time steps, outputting probabilistic data, which forms the basis of operations of the ML classifier.

3.4.4 ML Classifier

The last component of the model is the ML classifier, which takes in the probabilistic data, and performs prediction of the class of given data points, resulting in distraction classification. For this case, the classifier performs severity classification; given the outputs of the dynamic Bayesian network model, the classier approximates and maps the level of distraction on a severity scale. The severity of distraction acts as the basis of whether the system takes over the vehicle's operations or not.

Each TeleFOT Image is 1280 x720 with a width of 1280 pixels and 720 pixels, as depicted in Figure 3.2, which illustrates a sample of participant BL001, an enhanced image. The image was split into four frames using MATLAB representing In-vehicle (frontal view, side view) and Outer-vehicle (front and rear view), as shown in Figure 3.3. The driver point was determined based on a significant pre-defined point around the driver body region. The region around the head will indicate where the driver is; see Figure 3.4 and Figure 3.5. This enables us to perform further image segmentation to improve the accuracy of a head detection algorithm that ensures human recognition. We further classify the images into some of the distraction events that our algorithm will detect.

Figure 3-2: Image Enhancement Single Hand on wheel vs. Double Hands on the wheel

2. Autonomous Vehicle Monitoring Sample Image in the In-vehicle and Out-vehicle

Figure 3-3: In-Vehicle vs. Outer-Vehicle

3. Driver Feature Semantic Segmentation

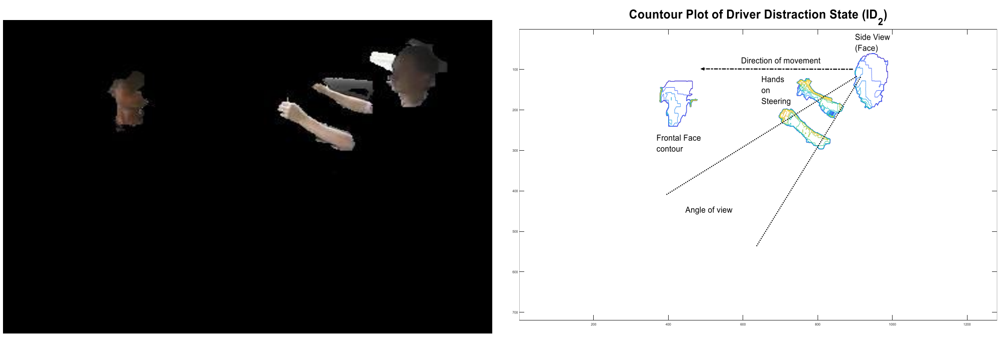

Figure 3-4: Image Segmentation (left) and the Contour Plot Driver state (right)

3.5 Contribution

The contribution of this study is as follows. First, a novel DL severity level of drivers' distraction identifies the degree of drivers' distraction and ensures a transition vehicle take-over. The existing approach focuses only on driver activity, not how severe the driver's behaviour is to make intelligent decisions. Our proposed system is and entails a hybrid-DL technique, namely CNN-DBN-LSTM.

Second, the neural network (NN) technique LSTM and DBN will be applied on time-series vehicles with frames from drivers' distractions. The LSTM tracks changes and the sequence to sequence of driver distractions. The Dynamic Bayesian network enables the tracking of ongoing events and prediction of driver distractions. The model will be trained with pre-trained network Resnet-18 to recognise driving environment features. Third, a driver distraction risk assessment MDDRA model for semi-autonomous vehicles. Fourth, we automated the enhancement and segmentation of driver features from the raw images and extracted drivers' distractions to achieve a high-level accuracy leading to a context-aware HMDDSM method for using the Hidden Markov Model to transition the driver to the vehicle.

3.6 Research Design

3.6.1 Introduction

Some authors present the research design as the necessary steps of structuring a study, particularly a research study and a scientific work, to tackle research questions and bring the required answer to the research questions. Authors affirm that research design is related to the study's logical structure [214]. The necessity to identify and shape the evidence that will allow answering the research questions unambiguously will depend on the area of research and how to plan the empirical research planning. In addition, the research design also highlights how the research investigation is being conducted [215]. Therefore, the research design is central to the scientific research inquiry, allowing the researcher to carry out his research work by avoiding bias, random error and error variance related to the research [216]. The organisation of the current research design session can be structured as presented below; firstly, an introduction, followed by a generality on data collection, next the TeleFOT NDS data were visited, next gathering the TeleFOT data were dealt with.

3.6.2 Data Collection Techniques (Conversion Image to Video – Time-Series Data)

The table below shows the trial type (baseline, experienced and novice) for all participants. There are 27 different participants from the table below, with some of the participants repeated over different conditions.

Table 3-1:Drivers Participants table

Trail type		
Baseline	Experienced	Novice
001	001	
002		
003		
004	004	
005		
006	006	
007	007	
		029
033	033	033
		034
		036
037	037	037
042	042	042
		043
		047
		059
	061	061
063	063	063
064	064	064
067	067	067
		071
074	074	074
		079
	080	080
081	081	081
083	083	083

088	088	088

Some participants only appear once, some twice and some completed all three trials as depicted in the table above. However, the maximum number of different participants is 27.

3.6.3 Video-Based Activity Recognition

The sampling rate of the image was generated at 24 frames per second using the FFMPEG software. This will ensure the tracking of events such as eyes glances that can occur within milliseconds. Xing et al [217] stated that eyes glance at the mirror can last from 0.5 to 1 second. An example of images split and the segment is below in Figure 3.6.

Figure 3-5: sample image enhanced and segmented

3.6.4 Times-Series Data – Content Awareness System / Analysis

The Pre-Processing of the time series data of the vehicle involves using the Race Technology software used to extract the vehicle time series data in CSV format. The following parameters were extracted from the CSV: time (s), long acceleration, latitude acceleration, vector acceleration, speed, distance, position X, position Y, video frame, video CPU, GPS latitude, and GPS longitude. In Figure 3.7, an example is depicted below:

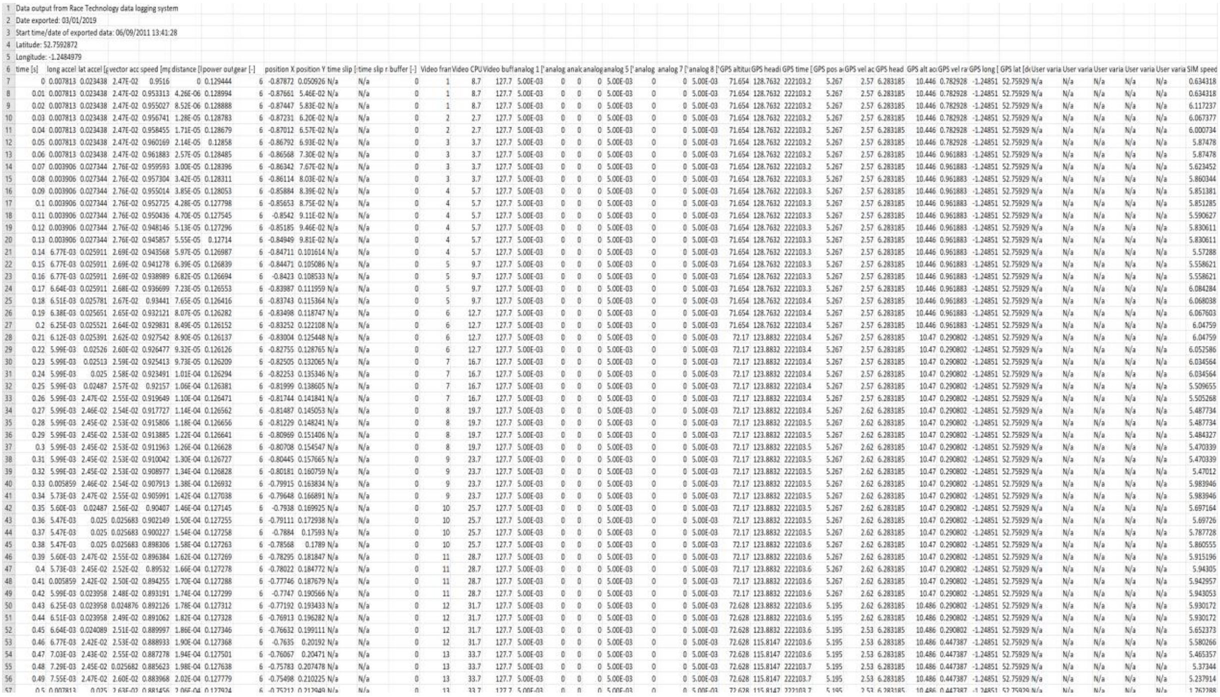

Figure 3-6: The algorithm suitable for time-series data

3.6.5 The TeleFOT Naturalistic Driving Study (NDS) Dataset

The Race Technology Data Analysis Software is the main application that comes with nearly all Race Technology products. The software can compute lap times and simple lap and sector times during the duration of a journey. The Race technology data analysis software is software used in the analysis of naturalistic driving data. Figure 3.8 below depicts an instance of a sample of the naturalistic driving video data analysis below. Furthermore, the capabilities of the Race technology software have features such as Controller Area Networks (CAN) with up to 100 variables with decoding of raw CAN, track maps, virtual dashboard, the complex calculation (speed and throttle averages), exporting data to spreadsheet and Matlab. The analysis involves annotating the driver’s behaviour that constitutes the attributes related to careless behaviour from indicators annotated from playing the video frame by frame.

3.6.6 TeleFOT NDS Data Gathering the

3.6.6.1 About TeleFOT NDS Data

TeleFOT comprises the most significant European Field Operation Test (FOT**)** regarding the functionality of in-vehicle aftermarket and nomadic devices. The TeleFOT project commenced in 2008 and ended in 2012. The project's primary purpose was to improve Autonomous systems and cooperative systems in the Intelligent Transportation Systems environment [212], [218]. The project further addresses Field Operational Test (FOTs) issues such as distraction, driving congestion, traffic efficiency, and travel speed controls. TeleFOT also involved large scale trial conduction with a significant number of vehicles instrumented with data loggers. The large scale FOT (LFOT) was carried out at eight individual test sites in Europe, namely UK, Finland, Sweden, Germany, Greece, Italy, and Spain. The Large scale FOT conducted involved vehicle collecting and recording driving data such as speed measuring, vehicle dynamics and vehicle positions. In 2011, TeleFOT –UK Detailed Field Operational Test (DFOT3) was launched to collect naturalistic drivers' behaviour without any predefined condition in the United Kingdom. The test location was mainly in the East Midlands (Leicester, Coventry, Nottingham) area of the UK and partnership with Loughborough University [210]. Initially, the challenging milestone in the methodology would have been getting real driving data.

The summary of TeleFOT Naturalistic Driving Data (NDS) is as follows:

- There are 27 individual driving participants
- Consistent route for all drivers, including predominantly urban with a range of junction and traffic types.
- Constant vehicle for all drivers using Ford Sedan 2008 MY model
- Approximately 50 minutes to 1 hour of driving for each data ‘packet’ (some drivers have repeated drives, some only one drive)
- Driving data packets up to 50 instances.
- Naturalistic driving with predefined conditions applied to be that constant vehicle and 1 hour.
- Linked four-channel video data of the front, rear, back, and side of the vehicle
- video data with 100Hz Global Positioning Service (GPS) and accelerometer data

3.6.6.2 TeleFOT NDS Video Channels

There in Figure 3.7 is a 4 tiled video channels data provisioned for both outside the vehicle and inside the vehicle [210]. Video data is also synchronised to GPS and accelerometer data of the vehicle. Thus, at every frame vehicle accelerometer is logged.

Figure 3-7: Showing 4 video channels data provisioned for both outside and inside the vehicle

3.6.6.3 TeleFOT NDS Data Formatting

To get the data out of the RUN Race Technology File, one will need to use the export function – again, this will be in the manual for the software. one can export it into excel or a Matlab file format. All the variables that the software exports will probably not be used for context-aware. A lot will just record null values as there was not anything connected to the data channel. As a start, the following variables will be considered:

Time [s] refers to the accumulated time from the first data record until the end of the records. This is recorded in seconds; however, as the data frequency is at 100Hz, we will need to run through 100 data records during one second.

Long accel [g] – This records pure 100Hz data from a tri-axis accelerometer; the data is in g and is not filtered or interpolated in any way. Longitudinal accel will record both positive and negative acceleration (i.e., acceleration and braking).

Lataccel [g] - This records pure 100Hz data from a tri-axis accelerometer; the data is in g and is not filtered or interpolated in any way. Lateral accel will record both positive and negative acceleration (i.e., cornering left and right).

Speed – Recorded in miles per hour and calculated from the GPS signal. Unlike other systems, this is the vehicle speed over the ground, not wheel speed or road speed.

Video frame – This is the matched video for each line of data. As the data is recorded at 100Hz and the video at 25Hz, the data will show blocks of four video frames (i.e. 100Hz / 25Hz)

GPS long [degs] – This is the longitudinal GPS position of the vehicle for each line in the data. The data is interpolated as the GPS sensor records at 20Hz and the data at 100Hz – therefore, the software interpolates vehicle positions between the actual records.

GPS lat [degs] - This is the lateral GPS position of the vehicle for each line in the data. The data is interpolated as the GPS sensor records at 20Hz and the data at 100Hz – therefore, the software interpolates vehicle positions between the actual records.

3.6.6.4 TeleFOT NDS Data Processing

Data will be extracted from the video data using an export function to log data into an excel format. In addition, some of the driving data variables that will be explored are time, acceleration, speed, video frame, GPS longitude and latitude. Ekambaram et al 2016 [219] extracted significant distraction of eye glances using only 10% (10 drivers selected) of the TeleFOT data. Some of the events that constitute eye glance distractions are (eyes off-road, at objects), eye closure. The research focused on the drivers' faces and eyes in identifying the object or field of attention. On average, the participants had eyes off normal activity (looking forward) for only 7% of the total test duration, the highest being 13% and the lowest 4%. This result is from approximately 23 minutes of analysed data, from a trial lasting over an hour drawn from only nine drivers. Results further show several minutes of remarkably accurate eye glances reading found within hours of recordings of the naturalistic driving data. Ekambaram et al [219] further stated that manual review of the TeleFOT data video recording shows other distraction events available such as mouth events (Talking to the passenger, biting nails), Hand gestures(waving), Hand distractions (hand on leg, seat belt removal and adjustments), Head Movements (neck and head position) etc. Ciscal-Terry et al [220] used TeleFOT data to analyse drivers' eye movements. Morris et al 2011 [221] deduced drivers' distractions from TeleFOT data by analysing the percentage of eyes that glances off-road. Franzen et al 2012 [222] used TeleFOT data to analyse wider distractions from outside the vehicle. Deducing from the TeleFOT data available, the events to be measured can be further narrowed down. For example, the severity level of a single event (eye glances on distraction sources) can be analysed. Another possible approach could be the severity level amongst multiple events (hands gesture, hand distractions, mouth events, head movements, eye glances).

This research would focus on the frame-by-frame image processing from the naturalistic driving video data captured at a minimum threshold of 24 frames per second (FPS). However, it should be noted that the duration of distraction events to be detected will guide the number of frames to be filtered and analysed. Realistically, the minimum severity threshold required for an event to be considered careless or dangerous driving will be justified. For example, detecting an event instance such as hand gesture event (seat belt adjustment, wave to passers, panel adjustment) for 10 seconds. The frames realised in this period of 10 seconds would be 24 x 10 seconds = 240 frames. Distractions and inattention events that can be detected from the TeleFOT are aided through the driver's eyes, mouth, and hands (Region of Interests (ROI)). Events such as eye glances, hand gestures, and mouth events can be significantly detected; the degree or threshold of the events mentioned above will classify driving behaviour. Furthermore, a severity model will also be developed. In this research, the events to be detected and classified will be narrowed down to eye glance, face orientation and hand gestures.

3.6.6.5 TeleFOT Video Coding Taxonomy

The video coding taxonomy in the table below has been drawn up from video analysis of 1 hr 04 minutes Naturalistic driving data around the city of Leicester. In addition, the glossary of data variables for fatal and accidents causation database has been consulted to complete the distraction[219,221,222] [223]. However, it should be noted that the list is non-exhaustive as the research is still at its preliminary stage.

Table 3-2: Showing the Naturalistic Driving data around the city of Leicester [210,219, 220, 221]

Distraction				
Distraction type	Code	Description or Notes	Occurrence ratings	Primary (P) or Secondary (S) distraction
Left Mirror	20	glancing to the left side mirror	-	P
Left Window	30	glance sideways at the left window	-	P
Wave to passing drivers	42	The driver looks at and waves to a passing vehicle either as a greeting or gesture of thanks.	-	S
Right Mirror	40	Any glance to the right-side mirror	5	P
Rear view Mirror	60	Any glance to the rear-view mirror	8	P

Interior Object	80	Looking at an identifiable object in the vehicle other than a cell phone. Such includes items as food etc	2	S
Talk to Passenger	11	The driver is talking to a passenger sitting in the passenger's seat	3	S
look at Passenger	12	The driver is looking at (and talking) to the passenger	4	S
Forward	10	Glancing out the straight-forward windshield.	5	P
adjust seat belt	13	Adjusting seat belt. Assumes driver is looking at and may reach for an object	-	P
Right Window	50	Any glance to the right-side window	-	P
Adjusting clothing	15	The driver puts on or takes off - part of the clothing.	-	S
Adjust in seat	24	Driver adjusting his position in the driver's seat	-	S
Look at GPS navigation systems	90	The driver interacts with an after-market GPS device that is mounted on the windscreen.	6	S
Look at outside vehicle or person,	40	The driver looks outside the vehicle to another person or vehicle at any critical situations	-	S
Look out rear	41	The driver turns around and looks out the rear window. It must be apparent that the driver is looking out the window	-	P
Hand gestures	43	The driver uses hand gestures, usually during speech (can include pointing)	-	S
Hands-on gearstick	44	Driver rests a hand on brake or gear sticks	-	S
Hand on leg	45	Driver rests a hand on the leg	-	S
Arm on windows	46	Driver relaxes arm on the window	-	S

3.6.7 TeleFOT Data Sampling Size

The role of this sampling is to have an idea of the data type used and how using different sample sizes can be determinant for the analysis to be conducted on this dataset. Indeed, the different data types gathered and the data size will substantially affect the analysis process. Therefore, sampling is an essential part of the data analysis to come.

Table 3-3: Showing a TeleFOT Data Sampling size

TELEFOT PARTICIPANTS	Baseline (BL), Experienced (E), Novice (N)	VIDEO LENGTH	IMAGE STATISTICS	DATA POINT (IMAGE STATISTICS X 4)
001	BL001	01:13:00	105,109	420436
	E001	00:33:40	48,485	193940
033	BL033	01:10:55	106,398	425592
	E033	00:38:13	57,334	229336
	N033	00:18:20	27,512	110048
074	BL074	00:33:45	48,605	194420
	E074	00:44:41	64,360	257440
	N074	01:33:13	134,219	536876
081	BL081	00:33:43	48,562	194248
	E081	00:34:24	49,556	198224
	N081	01:39:59	106,534	426136
083	BL083	00:33:43	48,562	194248
	E083	00:35:26	51,039	204156
	N083	00:57:58	83,470	333880
088	BL088	00:35:00	50,405	201620
	E088	00:42:17	60,904	243616
	N088	01:29:05	128,271	513084
TOTAL			1,219,325	4,877,300

3.6.8 TeleFOT Splitting Dataset (Training, Testing, and Evaluation)

The compelling feature of the network is its simplicity of the network layer in convolutional structure. However, it needs to be trained with considerable training data:

objects of different events must occur at every frame analysed. For the training, we generate thousands of images from each video at 25 frame MPS at the size of 1280x720. The TeleFOT data has been subdivided into the following 75% constitutes the training data, 15% testing and 10% for evaluation. We divided the images into the training images into 40% positive samples and 60% negatives samples. Positive samples are a sample with the event of interest to be identified within the object bounding boxes. The negative sample is realised when the bounding box does not intersect the object of interest. For the training, the crop is sampled in a distributed manner to ensure the whole image of the object is detected. The TeleFOT NDS dataset is quite large, weighing in at 138GB for the training images, 13GB for the testing images, and 6.3GB for the validation images. The test directory contains (as the name applies) 100,000 images (100 data points for each of the 1,000 classes) for our testing split.

Training the classifier for object detection requires several samples from each image, 40% for positive and 60% for negatives. Negative samples constitute bounding boxes with less than similarity than ground truth object boxes with a degree of 0.2(20%) similarity. Positives samples should fall within the object bounding box with 0.6(60%) similarity. Localization is complicated than classification when training; however, to address the issue, starting with a model with high-quality weights is crucial. The network will first be trained for the classification and weights of layers reused. Localization involves fine-tuning the convolutional layers in the whole network.

3.6.9 Dataset 2: American University Cairo Driver Distraction Dataset - v2

We used secondary data from American University in Cairo Driver Distraction Dataset (AUCDDD) V2 obtained from the Machine Intelligence group at the American University in Cairo [129,130]. The dataset is the first publicly available dataset for distracted driver detection. The study involves 44 participants from 7 different countries: Egypt (37), Germany (2), USA (1), Canada (1), Uganda (1), Palestine (1), and Morocco (1). Out of all participants, 29 were males, and 15 were females. Some drivers participated in more than one recording session with different periods, driving conditions, and wearing different clothes. Videos were shot in 5 different cars: Proton Gen2, Mitsubishi Lancer, Nissan Sunny, KIA Carens, and a prototyping car. We extracted 14,478 frames distributed over the following classes: Safe Driving (2,986), Phone Right (1,256), Phone Left (1,320), Text Right (1,718), Text Left (1,124), Adjusting Radio (1,123), Drinking (1,076), Hair or Makeup (1,044), Reaching Behind (1,034), and Talking to Passenger (1,797). The dataset satisfies research question 2, which contains distractions classified into careless or Dangerous driving. The first process is to perform a data cleaning by manually inspecting the video files with the eye and giving a distraction label for each frame. The transitional actions between each consecutive distraction type are manually removed. The Table below shows samples for the ten classes in our secondary dataset.

Table 3-4: Dataset 2 – AUCDDD - Distracted Driver Dataset v2

DISTRACTION EVENT CLASSES	FRAMES STATISTICS
Safe Driving c0	2,986
Phone Right c1	1,256
Phone Left c2	1,320
Text Right c3	1,718
Text Left c4	1,124
Adjusting Radio c5	1,123
Drinking c6	1,076
Hair or Makeup c7	1,044
Reaching Behind c8	1,034
Talking to Passengers c9	1,797

3.6.10 Justification of Context-Aware drivers' distractions Measurements

Wierwille et al [224] measured the response time in braking to measure the drivers' behaviour. Smith et al, [225], [225], the present research examined associations between poor driving behaviour (DB), driving when fatigued (DF), risk-taking (RT) and road traffic accidents (RTAs). The study involved a cross-sectional online survey of clients of an insurance company. The survey measured DB (speeding, distraction, lapses of attention and aggression), RT and frequency of driving when fatigued (DF, driving late at night, prolonged driving, driving after a demanding working day and driving with a cold). Speed Metrics using maximum speed (speed limit) and mean speed standard deviations. The use of driving speed and speed violations. This entails monitoring drivers' behaviour data and, in correlation, braking events during driving can be used. Incidents can be used to measure driver's behaviour and driver's aggression. However, the application area of this research is to prevent events that can lead to accidents.

Nabi et al [226] studied a behavioural pattern that could impact human behaviour called Type A behaviour pattern (TABP), characterised by impatience, time urgency, and hostility also linked with coronary heart disease. It has been debated that TABP is linked with risky driving behaviours, resulting in road traffic accidents RTAs. The methodology involved using participants reported maximum speed limits in different road types such as rural roads and highways.

Simmons-Morton et al, 2011 [227] stated gravitation force (g-force) elevation resulting from sudden deceleration or acceleration and hard turns are essential measures of risky driving. The high g-force rapid acceleration and deceleration can also reduce the amount of time to respond to hazards and increase the loss of vehicle control. Risky driving associated with elevated g -force events were assessed: longitudinal deceleration/hard braking (≤ -0.45 g); longitudinal acceleration/rapid starts (≥ 0.35 g), hard left (≤ -0.50 g) and hard right turns (≥ 0.50 g), and yaw (± 6 degrees within 3 seconds). A collision avoidance rule is meant to manoeuvre by giving 3 to 4 seconds distance between the vehicles before braking events. However, this is suitable under excellent weather conditions and normal traffic conditions.

3.7 AI and ML Techniques

3.7.1 Deep Learning

Deep Learning is a sub-topic of ML that relies on learning representation from data which uses learning successive layers for increasingly meaningful representations. DL allows computational models composed of multiple processing layers to learn data representations with multiple levels of abstraction. The layers are like neural networks structured in layers one after another. DL has been applied to image classification, speech recognition, handwriting transcription, text processing, speech recognition, and digital assistants. According to Goodfellow [228], some problematic representation can require a nearly human-level understanding of the data using DL. This has been applied primarily in cases whereby neural networks were shallow and can represent only one or two layers of representations. Using this neural network has been reduced by methods such as Support Vector Machines (SVMs) or Random Forests. DL enables computer systems to build complex concepts out of more straightforward concepts.

3.7.2 CNN

3.7.2.1 Definition

DL-CNN is a multilayer perception (MLP) mathematical function that maps some set of the input value to output values. These functions are actualized by composing many more specific functions [228]. CNN method is entirely an unsupervised feature learning that solves complex problems.

Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio; however recurrent nets have emphasized sequential data such as text and speech [229]. Machine-learning systems are used to identify objects in images, transcribe speech into text, match news items, posts or products with users' interests, and select relevant search results. Increasingly, these applications make use of a class of techniques called DL.

Conventional machine-learning techniques were limited in their ability to process biological data in their raw form. For decades, constructing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input. For classification tasks, higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations. An image, for example, comes in the form of an array of pixel values, and the learned features in the first layer of representation typically represent the presence or absence of edges at particular orientations and locations in the image [229]. In addition to beating records in image recognition and speech recognition, it has beaten other machine-learning techniques at predicting the activity of potential drug molecules8, analysing particle accelerator data, reconstructing brain circuits, and predicting the effects of mutations in non-coding DNA on gene expression and disease. Perhaps more surprisingly, DL has produced promising results for various tasks in natural language understanding, particularly topic classification, sentiment analysis, question answering and language translation.

3.7.2.2 CNN Architecture: Convolutional Layers

Convolution neural network is applying ML technique in computer vision-related problems such as object recognition. CNN consist of pooling layers and alternating convolution, as shown in Figure 3.9. Multiple weighted inputs are be assigned with a bias to the learning features that involve image segmentation [230].

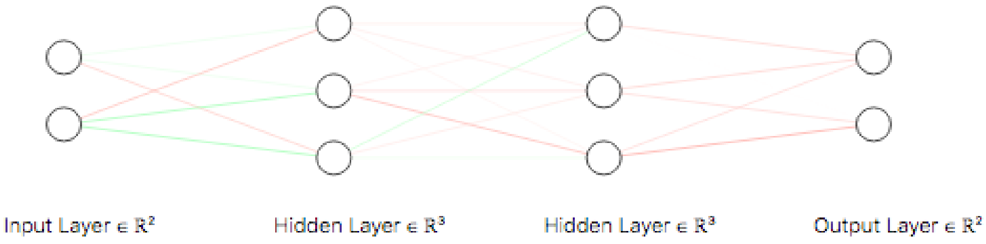

Figure 3-8: CNN Layers

Multiple hidden layers can be densely connected, and that makes inference challenging. Hinton and Osindero [231] eliminates the challenge using a method called complementary priors. This is achieved through a fast and greedy algorithm that learns deep one layer when given that the top two layers form an undirected associative memory. The fast, greedy algorithm slows the learning procedures that fine-tune the weight using a contrastive wake-sleep algorithm. In this research, the driver's behaviour can be of multiple labels, and context perceptions are dynamic and real. However, as the number of parameters increases, scalability becomes difficult. In addition, variation learning requires all the parameters to be learned together. The fast and greedy learning algorithm can quickly learn multiple set of parameters, with the capability to learn from the deep network with millions of parameters and many hidden layers.

Furthermore, deep belief networks are unsupervised learning and can be applied to labelled data through learning models that generate both data and labels.

3.8 Long Short-Term Memory

3.8.1 Definition

LSTM layer is a recurrent neural network (RNN) layer that supports time-series and sequence data in a network. The layer performs additive interactions, which can help improve gradient flow over long sequences during training. LSTM layers are best used for learning dependencies from distant time steps. The LSTM is used in the prediction of the drivers' distractions

3.8.2 Sequence Input Layers

The sequence input layer inputs sequence data to a network. The sequence input layer is created using the sequence input layer. An LSTM layer is a recurrent neural network (RNN) layer that supports a network's time-series and sequence data. The layer performs additive interactions, which can help improve gradient flow over long sequences during training. LSTM layers are best used for learning dependencies from distant time steps. LSTM layers are best used for learning dependencies from distant time steps. The learnable weights of the LSTM network are input weights W, recurrent weights R, and bias b.

$$W = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_o \end{bmatrix} \quad R = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_o \end{bmatrix} \qquad b = \begin{bmatrix} b_i \\ b_f \\ b_g \\ b_o \end{bmatrix} \tag{3.1}$$

Where i, f, g, and o denote the input gate, forget gate, layer input, and output gate. The cell state at time step t is given by,

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \tag{3.2}$$

Where $\odot$ denotes the element-wise multiplication of vector (Hadamard product).

The hidden output state at time step t is given by,

$$h_t = o_t \odot \tanh(c_t) \tag{3.3}$$

$$i_t = \sigma(W_i x_t + R_i h_{t-1} + b_i) \tag{3.4}$$

$$f_t = \sigma(W_f x_t + R_f h_{t-1} + b_f) \tag{3.5}$$

$$g_t = tanh(W_g x_t + R_g h_{t-1} + b_g) \tag{3.6}$$

$$o_t = \sigma(W_o x_t + R_o h_{t-1} + b_o) \tag{3.7}$$

$$\sigma \; is \; the \; sigmoid \; function \; where \; \sigma(x) = (1 + e^{-x})^{-1} \tag{3.8}$$

3.9 Dynamic Bayesian Network

DBN is a directed acyclic graph representing conditional independence between a set of random variables, which deals with uncertain information and probabilistic inference upon receiving evidence. It consists of nodes representing the random variables and arcs representing the conditional independence between variables.

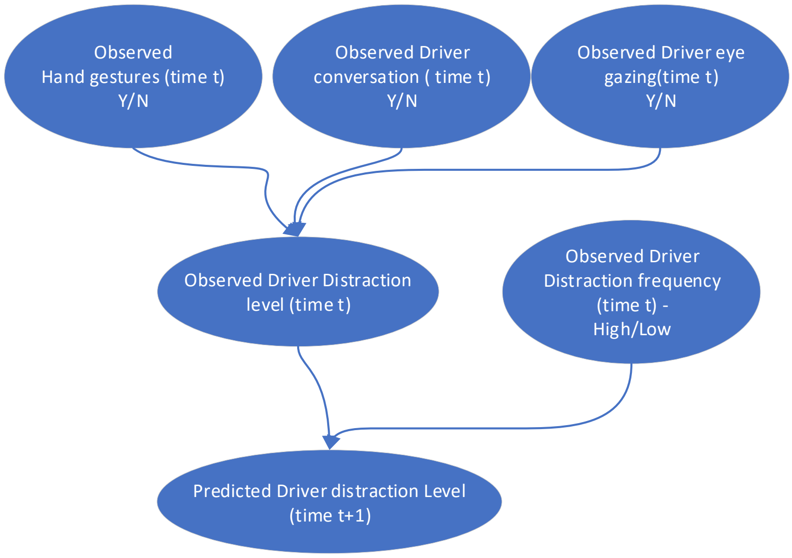

Figure 3-9: Dynamic Bayesian Network: Driver Distraction Leve

3.10 Severity Level Drivers' Behaviour System Framework

Learning and predicting driving behaviour are challenging due to complex factors needed to model context-aware with the driver's behaviour. In Figure 3.10 above, we present an architecture that could be used to learn, analyse and make decisions based on driver's context-aware and Drivers' behaviour.

3.10.1.1 Context-aware Driver Distraction Severity Classification Architecture

Figure 3.11 presents the complete Architecture of the proposed Context-aware Driver Distraction Severity Classification model.

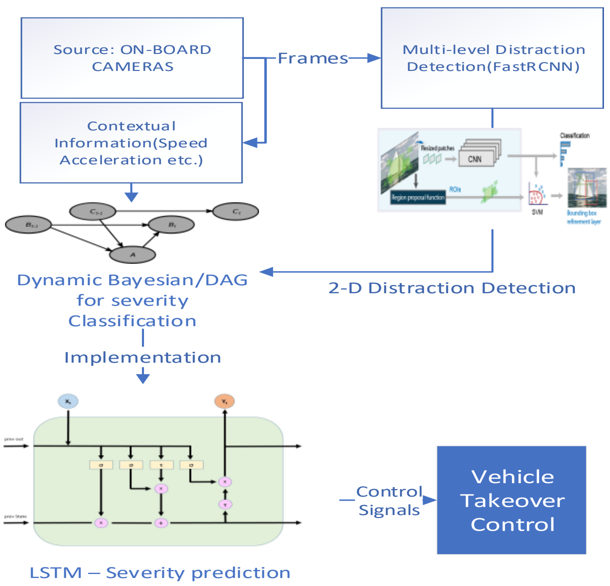

Figure 3-10: The Context-aware Driver Distraction Severity Classification Architecture.

3.11 Fuzzy Logic

Fuzzy Logic Methods using an improved neuro-fuzzy inference system (ANFIS) model could be used to simulate and predict the car-following behaviour based on the reaction delay of the driver-vehicle unit. The reaction delay is used as an input in this model, while other model inputs and outputs were chosen concerning this parameter. Using the real-world's collected data, the performance of the model was evaluated. This model was also compared with the responses of existing ANFIS car following models[232]. The simulation results showed that the proposed model has very close compatibility with the real-world data and reflects the traffic flow situation in a more realistic way, which was a significant improvement.

3.12 Hidden Markov Model

A hidden Markov model (HMM) is a statistical model used to describe the transformation of observable events that rely on internal factors, which cannot be observed directly. The invisible factor underlies the observation is called a 'state', and the observed event is called a 'symbol'[13], [17].

3.13 Imagenet Evolution

It has been founded that DL broke out from the use of ImageNet evolution. Deeper neural networks are difficult to train; He et al [233] presented a residual learning framework to enable the training of extensive networks. The approach involved using the ImageNet dataset to evaluate residual nets with a depth of 152 layers. This approach resulted in a 3.57% error on the ImageNet test set. The result also presents an analysis of the classification task on CIFAR-10 with 100 and 1000 layers. According to Szegedy et al [234], a deep CNN architecture named inception was used for the classification and detection in ImageNet Large-Scale Visual Recognition. Improved results in ML increase computational power, large datasets, but it is mainly due to efficient algorithms. In a large dataset, increasing the number of layers and dropouts is to address overfitting. The inception method uses an optimal sparse structure that improves dense blocks through neural networks for computer vision. The limitation to this object detection methodology is the lack of utilizing context nor performing bounding box regression. Krizzhevsky et al [235] adopted a deep CNN to classify 1.2 million images in the ImageNet into 1000 different classes. The test data achieved an error rate of 37.5% and 17.0% at top-1 and top-5, respectively.

Furthermore, it has been stated that Deep CNN is an effective neural network for up to 60 million parameters, and 650,000 neurons can be achieved using 5 CNN layers. However, there is other new pre-trained network architecture such as ResNet (5,10,18,50), GoogleNet and MobileNet. However, Resnet-18 will be adopted for this research due to the quality of the dataset and computational resources. Most important, the ResNet gives the best accuracy compared with other models [223].

3.14 Deep Learning (DL)

3.14.1DL Technologies Supporting the Design CNN Architecture

DL methods are learning methods that use multiple levels of representation, obtained by composing nonlinear but straightforward modules that transform the representation at one level given raw input into a representation at a higher, slightly more abstract level. Using a general-purpose learning procedure rather than manual labelling, the layers are learned from data LeCun et al [236]. According to Goodfellow et al [237], DL is a subfield of ML -DL entails perceptron algorithm, which will be used in automatic learning and assigning weights in the classification of inputs as depicted in figure 3.12 below. This approach will be applied in solving the research gaps. For example, weights will be assigned to drivers' distraction to classification and transitioning of vehicle take-over.

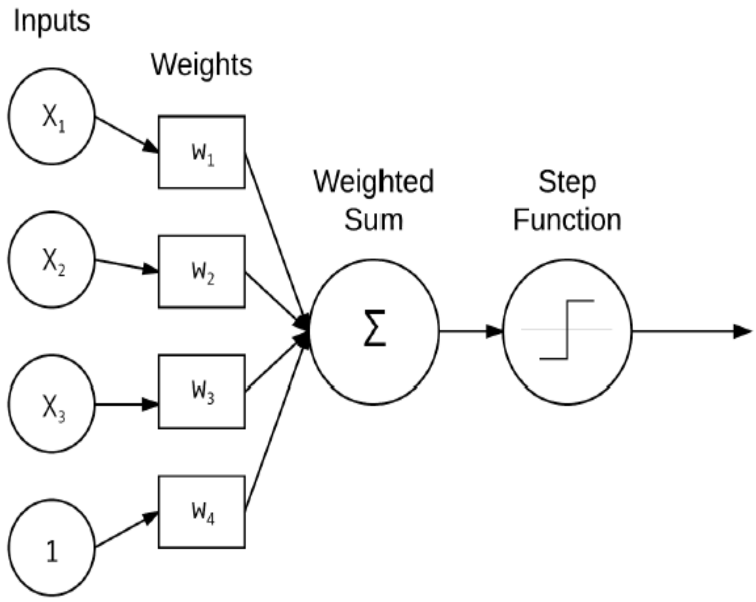

Figure 3-11: Simple Perceptron Network Architecture (Rosebrock, 2017)

3.14.2 AI Technologies Supporting the Design

3.14.2.1 Convolution Neural Networks

Yan et al [238] proposed a CNN-based model trained with six classes of labelled data, which learns, detects and predicts driver distraction and fatigue from analysing a driver's ears, mouth and eyes movements. Their model achieved 95.56% performance accuracy. Similarly, Le et al [121] developed a Multiple Scale Faster-RCNN approaches for detecting cell phone usage by drivers and steering wheel handling. The method achieved state-of-the-art accuracy in detecting cell phone usage. Yuen et al [239] used a CNN-based method to observe faces, localize landmarks, and estimate head pose. Yuen et al [239] improved face detection by using a deep CNN-based approach, which performs discrete head pose estimation and performed better compared to baseline methods. Muhlbacher-Karrer et al [240] proposed another CNN-based method that detects the state of a driver; distraction, tiredness and stress. The approach detects a driver's hand actions on a steering wheel using a "capacitive-based wireless hand detection sensor". Vora et al [241] introduced a CNN system that performs driver gaze zone estimation. Yan et al [120] recognized driver's inattention using a CNN to learn and predict driver state features such as the eyes, mouth and ear. The detection of the features as mentioned earlier was done by training dataset that consists of four activities everyday driving, cell phone usage, eating and falling asleep. Detection was achieved using a Face++ Research toolkit that localizes the facial landmarks on drivers. Results yielded a 95.56% accuracy in classifying the driver's mouth, ear and eye [120]. Le et al [121] used an advanced DL approach that detects objects such as hands, cell-phone usage.

3.14.2.2 Fast Recurrent Convolution Neural Network (Fast RCNN)

Le et al [121] proposed a DL technique that features a Multiple Scale Faster-RCNN integrated with a standard Region Proposal Network (RPN), which features maps that entails convolution feature maps such as ROI pooling, conv4 and con3. The data adopted is from SHRP-2 databases, and results yielded a reduced testing cost, better accuracy and independent facial landmarking. The DL-based MS-FRCNN achieved higher accuracy than similar Faster R-CNN. Donahue et al [122] stated that Recurrent Neural networks had gained recognition in the image interpretation for recurrent models for tasks with sequences (time series data) and visual representation. They also proposed a Long-term Recurrent Convolutional Networks (LRCNs) architecture for visual recognition that combines convolutional layers and long-range temporal recursion. The architecture considers three vision difficulties (activity recognition, image description, and video description) and instantiates the following sequential learning task, namely Sequential input, static output ($(x_1, x_2, ...x_t) \mapsto (y_1, y_2, ...y_t)$), Static input sequential output ($x \mapsto (y_1, y_2, ...y_t)$) Sequential input and output ($(x_1, x_2, ...x_t) \mapsto (y_1, y_2, ...y_t)$). The aforementioned sequential input approach can be applied in time-series data such as speed, acceleration in the Naturalistic driving study (NDS).

3.15 Summary

This chapter carefully investigates the methodology adopted in the present study. In this regard, a quantitative data analysis approach is applied in a time-series dataset, images, and videos, monitoring driver behaviour through these recorded data. The chapter also highlights the main algorithms, techniques and models supporting this research. These mainly revolve around algorithms, such as Convolutional Neural Network (CNN) as the DL algorithm, Fuzzy-logic, Hidden Markov Model, DBN as the AI algorithm, and Long Short-Term Memory (LSTM) as the computer vision algorithm, which is related to the analysis of driver behaviour. The choice of the research method is also justified and is developed around two methods, namely simulation and Field Operational Testing (FOT). Indeed, one of the primary datasets used in this study is TeleFOT, one of the most extensive European datasets related to monitoring and improving autonomous and cooperative systems in the ITS context.

CHAPTER 4. CONTEXT-AWARE DRIVER DISTRACTION SEVERITY CLASSIFICATION USING LSTM

4.1 Chapter Objectives

- To introduce a novel multi-event driver distraction detection system based on events information (RoI) and using context-aware parameters.
- To propose a severity level classification system of driver behaviour applying LSTM to implement a probabilistic model.
- To propose a method to predict driver distraction activity from times series data of participants from naturalistic driving.
- The design of proposed deep recurrent neural network using the non-linear autoregressive with exogenous inputs.

4.2 Synopsis

ADAS play a vital role in ensuring the safety of passengers and drivers in both private vehicles and public transportation systems. This chapter entails applying a DL method to classify driver distraction behaviour based on parameters of context-awareness, namely speed, manoeuvre, and event type. Driver distraction is examined via video coding taxonomy using event information on regions of interest, including eye gaze estimation, facial orientation, and hand gestures. Meanwhile, the severity of driver distraction is classified using a novel probabilistic (Bayesian) model drawn from an LSTM. Furthermore, there is an approach for further classification of driver distraction severity using frame-based context data derived from the multi-view TeleFOT naturalistic driving study (NDS) dataset. The presented methodology enables driver distraction severity to be predicted using recurrent deep neural network layers trained on time-series data.

This chapter's contribution is its proposal of a frame-based metric to measure driver distraction severity via linear transformation; the classification of severity levels through LSTM; the experimental validation of a frame-based model of severity; and the use of naturalistic driving study data to develop and test an classification system that will enable the vehicle to take over from the driver based on the driver distraction severity level. This system would be a helpful component in ADAS. This chapter also involves using a Dynamic Bayesian Network model to predict the driver's distraction.

4.3 Introduction

A proposed prevention system based on data from the secondary Naturalistic Driving Study (NDS) TeleFOT to driver distraction reduces the likelihood of traffic accidents. To this end, 27 subjects are assessed, and the TeleFOT data usage is explored to identify any existing events in the dataset. Specifically, using linear transformation to propose a frame-based driver distraction severity metric and develop an architecture to classify driver distraction in levels of severity based on LSTM. The validation of the proposed model via experimentation using naturalistic driving to develop and test the classification system for vehicle take-over based on the driver distraction severity level. This system will contribute to the existing ADAS. The focus is on driver distraction monitoring via context-awareness, analysed using LSTM in a Recurrent Neural Network (RNN). It will be possible to apply the system to the evaluation of driver behaviour, thereby facilitating systems that can prevent or correct driver distraction according to the distraction severity level. This evaluation can be subjective according to the event duration and frequency; hence, the proposed system takes both driver distractions and context-awareness information into account.

4.4 Context-Aware Driver Distraction Severity

Table 4.1 presents the distractions detected in participant BL_001 in the Virginia Tech Transport Institute (VTTI) standard distraction taxonomy [242]. Each distraction event (ID) is assigned a unique number and coded by the frame number. Therefore, each distraction is identified in each frame. The system of taxonomy coding is derived from the VTTI, and the limitations in capturing driver distraction are addressed through the development of the minimum required attention (MIRA) theory. The code sections are not sequential in BL_001's video coding data (context-aware), and only the analysed distraction codes are logged. Table 4.1 is derived from the MIRA standard video coding taxonomy.

Table 4-1: Driver Distractions

Distraction type	Distraction Type ID	Description	MIRA #
Left mirror	2	Any left side-mirror glance	
Left window	3	Any glance to the left-side window (looking at junctions, else 40)	
Right mirror	4	Any glance at the right side-mirror	7
Right window	5	Any glance to the right-side window (looking at junctions, else	

		40)	
Rear-view mirror	6	Any glance at the rear-view mirror	8
Instrument cluster	7	Any glance at the instrument cluster located beneath the dashboard, e.g., speedometer, control stalks, and steering wheel	
Interior object	8	Any glance at an object in the vehicle is different from a mobile phone. Objects may include personal items brought in by the participant	2
Look at passenger	12	Driver looking at (and talking to) passenger.	3
Look outside vehicle either through windscreen or side window	40	The driver looks at another vehicle, person, animal, or undetermined object outside the vehicle (not checking at junctions, else 3 or 5)	

In classifying the driver distraction severity, there will be considerations of the following context-aware variables:

- **Distraction Type:** The driver distraction events/occurrences are listed in Table 4.1.
- **Speed:** Recorded in miles per hour and derived from the GPS signal; here, vehicle over the ground speed, not wheel speed or road speed, is the case in other systems.
- **Manoeuvres:** Indicating the vehicle stopped ('S'), turning ('T') or otherwise in Table 4.2.

Table 4-2: Manoeuvres

Manoeuvre (Column 9)			
Type	Code	Notes	
Stopping or Stopped	S(1)	Code only when the vehicle is coming to a stop or is very slow-moving; once the vehicle moves off, stop coding—no S codes to be used with the following.	P
Turning	T(2)	This identifies the moment the vehicle enters the actual manoeuvre – the first movement on the steering wheel or when the vehicle crosses the giveaway line; always ends with the last record for each 'event'.	P

4.5 Context-Aware Driver Distraction Severity Classification Architecture

In Figure 4.1 below is the architecture for the context-aware driver distraction severity architecture.

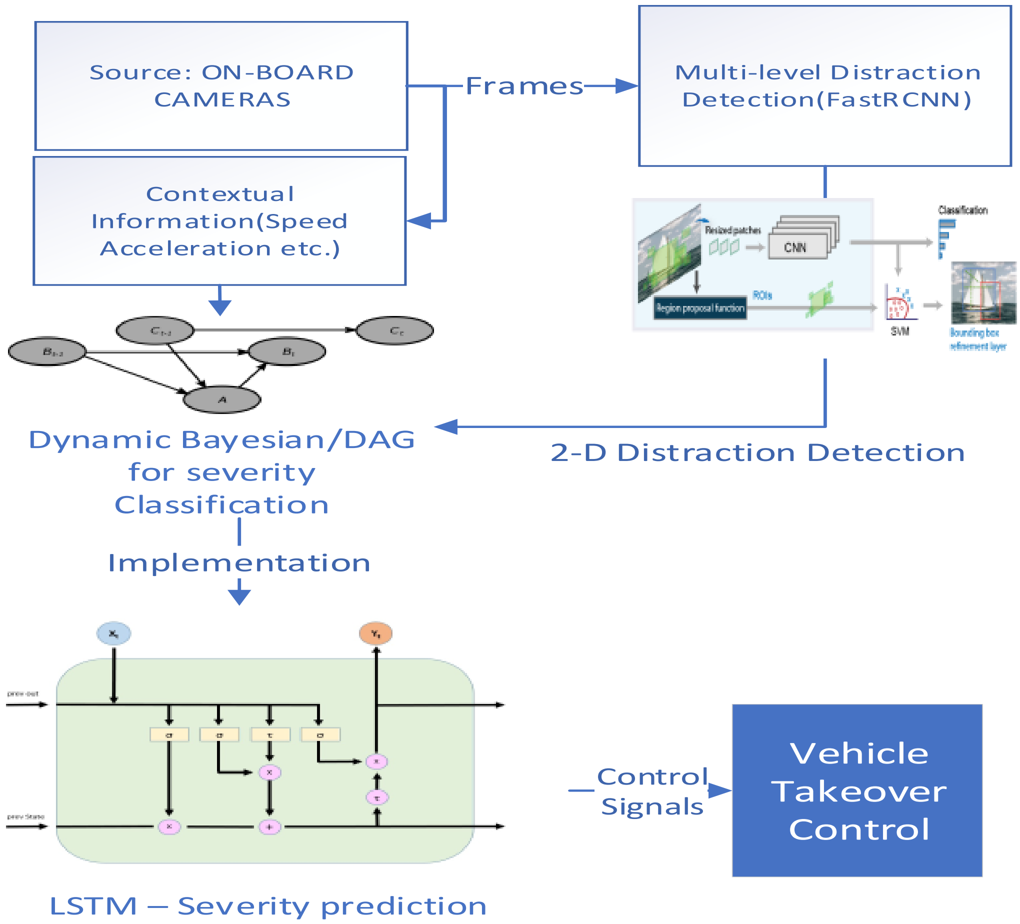

Figure 4-1: The Context-aware Driver Distraction Severity Classification Architecture

4.6 Dynamic Bayesian Network

According to the DBN, the driver distraction severity level is classified as seen in figure 4.2.

- **Distraction Type:** The driver distraction event/occurrence identified as a distraction.
- **Speed:** Recorded in miles per hour and estimated from the GPS signal. Unlike previous systems, this refers to vehicle speed over the ground rather than wheel speed or road speed.
- **Manoeuvres:** This indicates whether the vehicle is stopped ('S'), turning ('T') or otherwise.

$$Pr(S_t) = Pr(S_{t-1}, D_t) = Pr\big((S_{t-1}|D_t)\big) \times \Pr(D_t) \quad (4.1)$$

Where S_t, S_{t-1}, D_t are severity probability at time t, severity probability at t-1, and distraction probability at t.

computation of the frame-based severity metric denoted S_t To classify driver distraction using the direct acyclic graph (DAG).

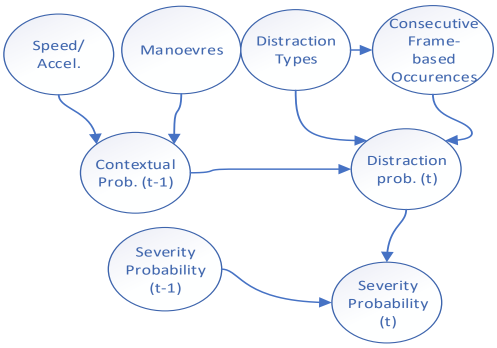

Figure 4-2: The Dynamic Bayesian network for severity classification.

4.7 LSTM-Based Driver Distraction Severity Classification

An LSTM layer represents an RNN layer supporting time-series and sequence data within a given network, conducting summative interaction that promotes gradient flow throughout long sequences during the training process. These layers are highly suitable for pattern learning or capturing dependencies based on distance (time) steps. In the LSTM network for driver distraction severity, the learnable weights are the input weights W, the recurrent weights R, and the bias b. The sequence input layer, created using the sequence input layer, inputs time-series data into the driver distraction severity LSTM network.

$$W = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_0 \end{bmatrix} \quad R = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_0 \end{bmatrix} \qquad b = \begin{bmatrix} b_i \\ b \\ b_g \\ b_0 \end{bmatrix} \quad (4.2)$$

I, f, g, and o are the input gate, forget gate, layer input, and output gate. The frame cell state at time step t is given by:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (4.3)$$

Where ⊙ denotes the element-wise multiplication of the vector (Hadamard product).

The hidden output state at the time (frame) step t is given by:

$$h_t = o_t \odot \tanh(c_t) \tag{4.4}$$

$$i_t = \sigma(W_i x_t + R_i h_{t-1} + b_i) \tag{4.5}$$

$$f_t = \sigma(W_f x_t + R_f h_{t-1} + b_f) \tag{4.6}$$

$$g_t = tanh(W_g x_t + R_g h_{t-1} + b_g) \tag{4.7}$$

$$o_t = \sigma(W_o x_t + R_o h_{t-1} + b_o) \tag{4.8}$$

The σ is the sigmoid function, where

$$\sigma(x) = (1 + e^{-x})^{-1} \tag{4.9}$$

4.8 LSTM Architecture

The LSTM architecture below in figure 4.3, showing the features forget, update, and output.

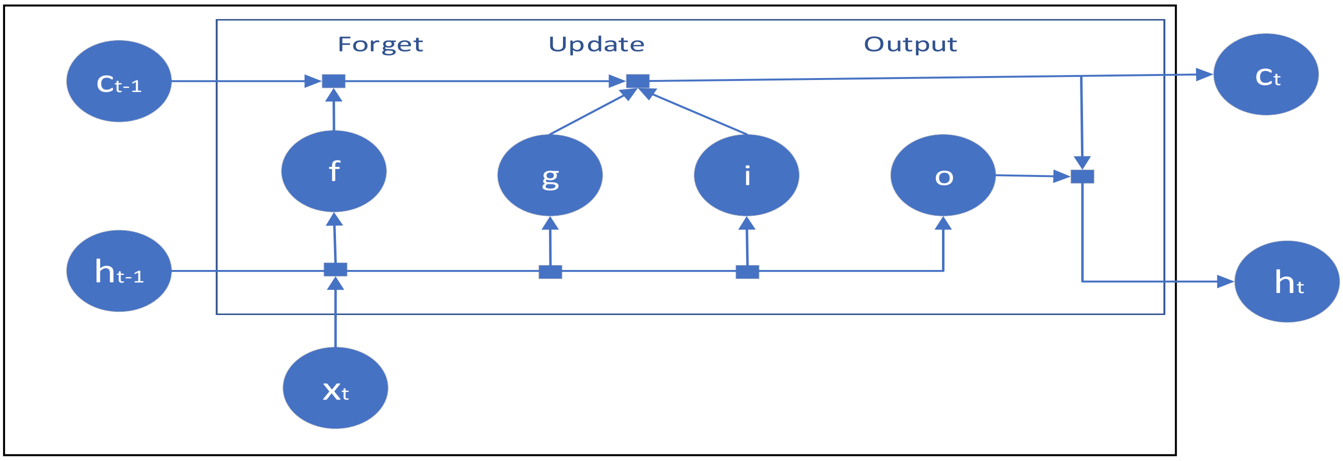

Figure 4-3: Frame-based Data Flow at Time Step t (LSTM Layer)

4.9 Experimental Results

Based on experiments in MATLAB 2019a, the development of the driver distraction severity model adopts the use of a non-linear autoregressive exogenous (NARX) neural network or LSTM for classification, accuracy, and precision (see Figure 4.3). The RNN uses ten hidden neurons and a delay of 2. Figure. 4.4. presents the response from the deep RNN (LSTM network) developed to classify driver distraction severity based on the distraction taxonomy in conjunction with participant BL_001's naturalistic data in figure 4.4. Table 4.4 presents the dataset selection based on the distractions listed in Table 4.1 above (section 4.1).

Table 4-3: Statistical Frequency of Distraction in the Video Frames of Driver BL001

Distraction type	Distraction Type ID	Description	Statistics of Frame
Look outside the vehicle, either through windscreen or side window	40	The driver looks at another vehicle, person, animal, or undetermined object outside the vehicle (not checking at junctions, else 3 or 5)	3643
Interior Object	8	Any glance at an object in the vehicle that is not a mobile phone; the object may include personal items brought in by the participant	148
Right Mirror	4	Any glance at the right side-mirror	1024
Left Mirror	2	Any glance at the left side-mirror	179
Rear-view Mirror	6	Any glance at the rear-view mirror	38
Look at Passenger	12	Driver looking at (and talking to) a passenger	148
Left Window	3	Any glance at the left-side window (looking at junctions, else 40)	44
Right Window	5	Any glance at the right-side window (looking at junctions, else 40)	439
Instrument Cluster	7	Any glance at the instrument cluster beneath the dashboard, e.g., the speedometer, control stalks, and steering wheel.	61

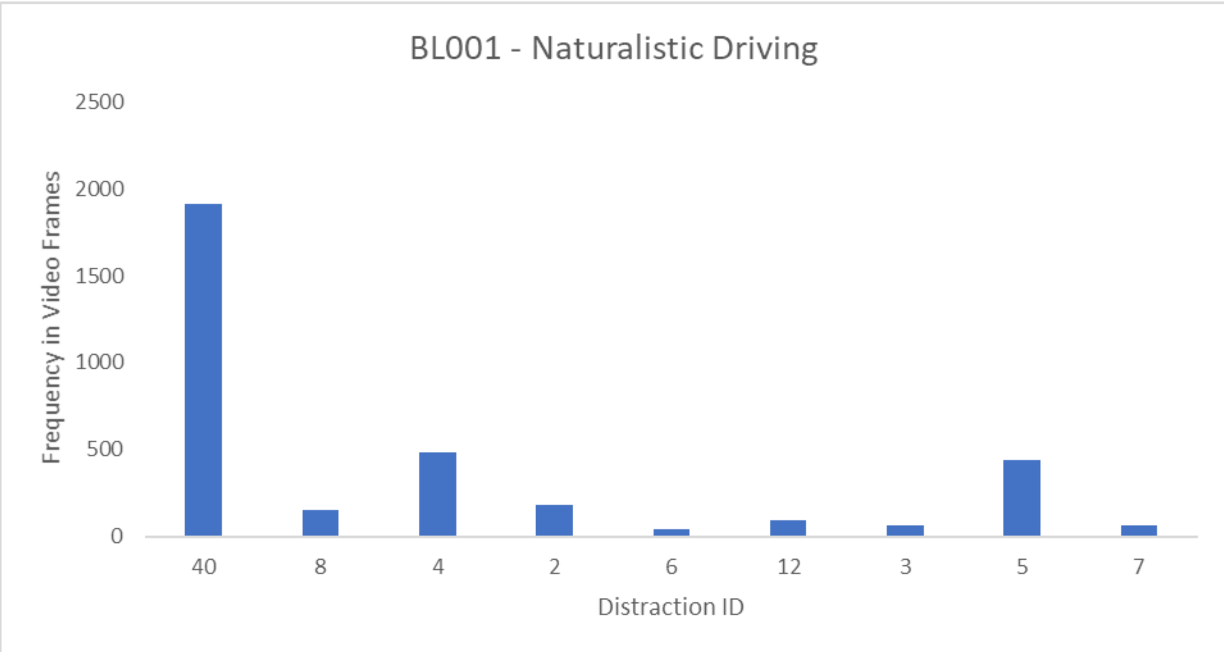

Figure 4-4: Naturalistic Driving study for participant BL001

4.9.1 TeleFOT Data Selection

The input ***(training data)*** represents a 1912 x 4 matrix consisting of dynamic data, i.e. the 1912 timesteps of 4 elements. The selected data are context, namely speed/acceleration, manoeuvres, distraction type, and event information. Meanwhile, the target of 'severity' is a 1912 x 1 matrix, constituting the severity classification's computed probabilities, i.e. 1912 timesteps of 1 element. The Levenberg-Marquardt (trainlm) is adopted as the training algorithm, requiring more memory yet less time. Once the generalization ceases to improve, indicated through an increased mean square error (MSE) of the validation samples, the training automatically stops.

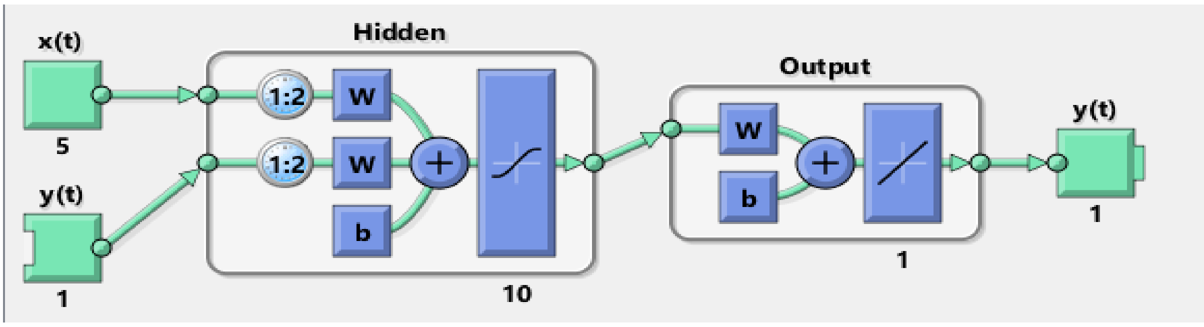

Figure 4-5: The LSTM Network Implementation

4.9.2 Validation And Testing of The TeleFOT Data

Training: Table III gives the training data provided to the network, which is adjusted based on its MSE. To enhance the quality of the results, 75% of the frame-based training data were selected to train the LSTM.

Validation: These data are utilised to estimate the network generalization and cease training once the generalization has stopped improving. 15% of the data are used in the validation.

Testing: These data do not affect the training and thus supply an independent measure to evaluate the network's performance during and after training. 15% of the data are used for testing, as presented in Table 4.5.

Table 4-4: LSTM Design and Implementation

	Target Values	Mean Square Error (MSE)	Regression R
Training	1338	6.67462e-4	9.90168e-1
Validation	287	5.41263e-4	9.92007e-1
Testing	287	9.62103e-4	9.86073e-1

4.9.3 Times-Series Response of The LSTM Network for Driver Distraction Severity

Figure 4.6 portrays the time series response plot. The small prediction errors support the adequacy of the prediction responses obtained for the intelligent filter.

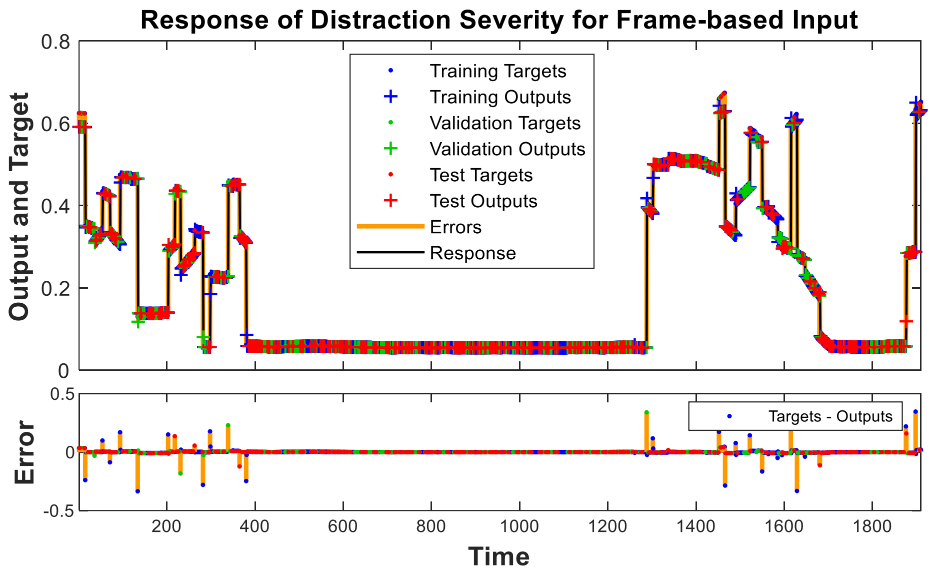

Figure 4-6: Time-Series Response of the LSTM Network for Driver Distraction Severity

Figure 4.7 shows the Auto-correlation error plot, which signifies the input-error correlations. The value of correlation that stands out is zero lag; the confidence unit (degree of confidence) falls below 0, which depicts is a strong correlation (positive correlation).

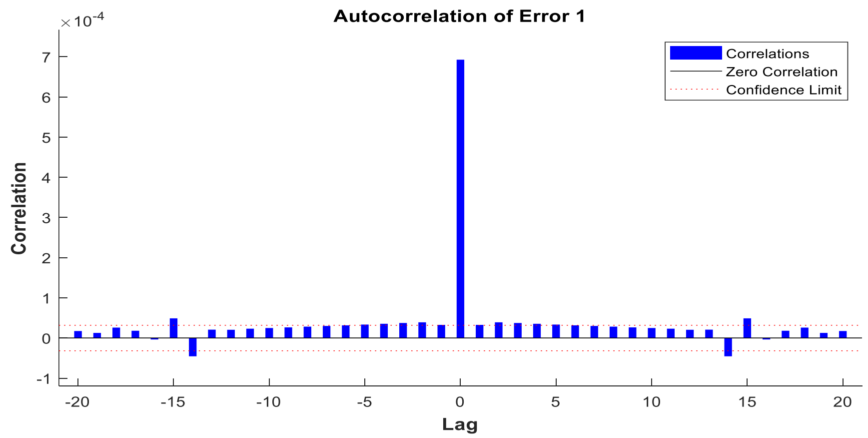

Figure 4-7: Error Auto-Correlation plot

Figure 4.8 depicts the input-output correlation errors regarding the target variable. The filter utilizes the initial values to predict the appropriate output and how the errors correlate with the input sequence concerning the target variable. As the autocorrelation values have about zero-correlation Lag, there is a 95% confidence limit which makes the prediction contains no errors.

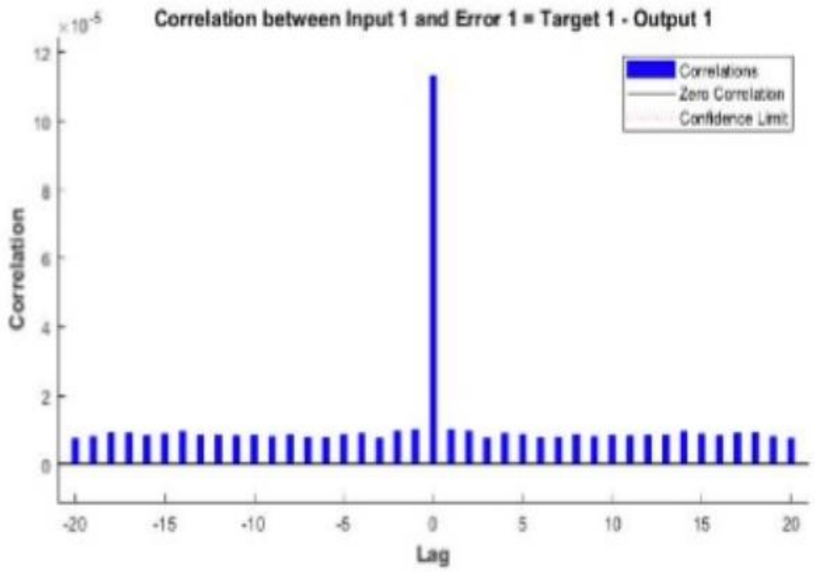

Figure 4-8: Input Error Autocorrelation

4.9.4 Performance

4.9.4.1 Performance/Response:

The network's performance (MSE) starts at 0.0327 and, after the 27^{th} epochs, stops at 0.000541. Figure 4.8 above graphically presents the driver distraction severity model's response, comparing the (training, validation and testing) targets of the time-series (frame-based) data and the actual outputs. Following the 27th epoch, the error validation is repeated many times. As the error shows no sign of reducing, the test is halted at 35 epochs.

As shown in Figure 4.9, the error repeat that begins at epoch 27th shows data over-fitting. Hence, epoch 27th is chosen as the base, with its weights selected as the final weights. Furthermore, six iterations are run in the validation check to enhance the filter's performance; as the error does not reduce, the testing is halted at epoch 35th.

4.9.4.2 Training/Validation Accuracy

The Levenberg-Marquardt training algorithm needs more memory but less time to perform the training. It also improves performance by using the gradient-descent method. In training, the accuracy begins at 0.523 and started repeat at the 27th epoch with an accuracy of 99.0168. Once the generalization ceases to improve, the training was stopped based on the validation samples' MSE and Accuracy. This occurs at epoch 32, with a validation check time of 6 secs and validation Accuracy of 98.60%, as shown in Figure 4.9.

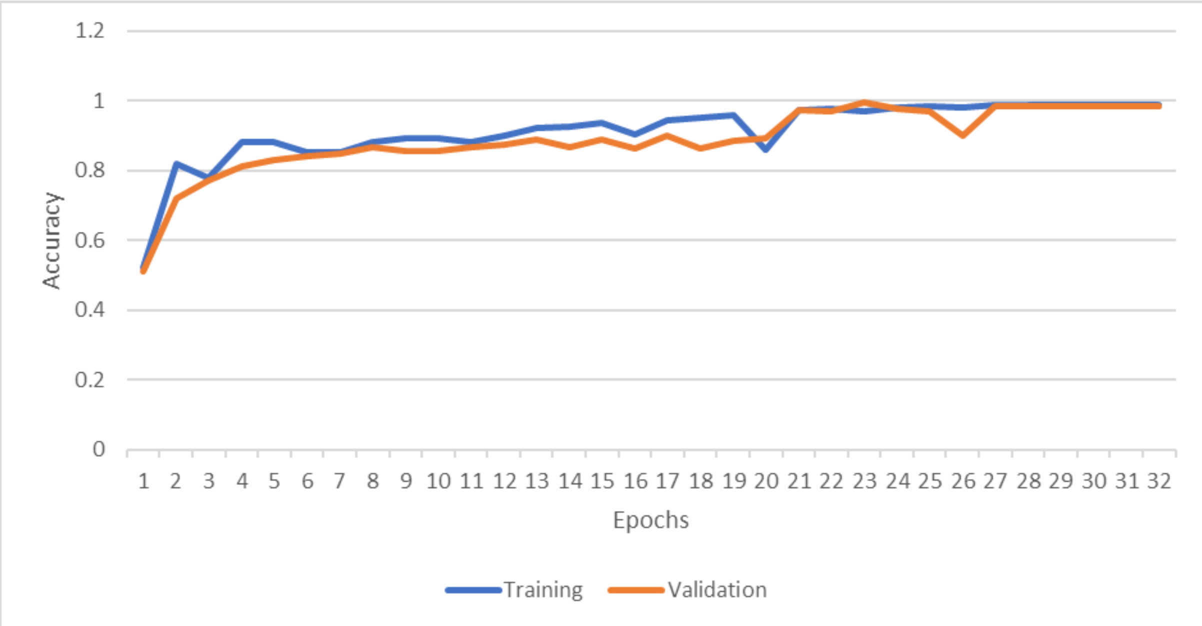

Figure 4-9: LSTM Training and Validation Accuracy using the Gradient Descent Method.

Figure 4.10 presents the MSE plot against the epochs, demonstrating an improved performance for every iteration between 1 and 27. Nevertheless, based on the MSE, performance shows no improvement between iterations 27 and 32. The best validation performance and MSE begin at 0.881 and decline to an error value of 0.000541. Notably, the three lines respectively depict the training, validation and testing steps. In this case, to avoid over-fitting the dataset, the training cycle is ongoing until the point at which training reduces the validation cycle's prediction errors.

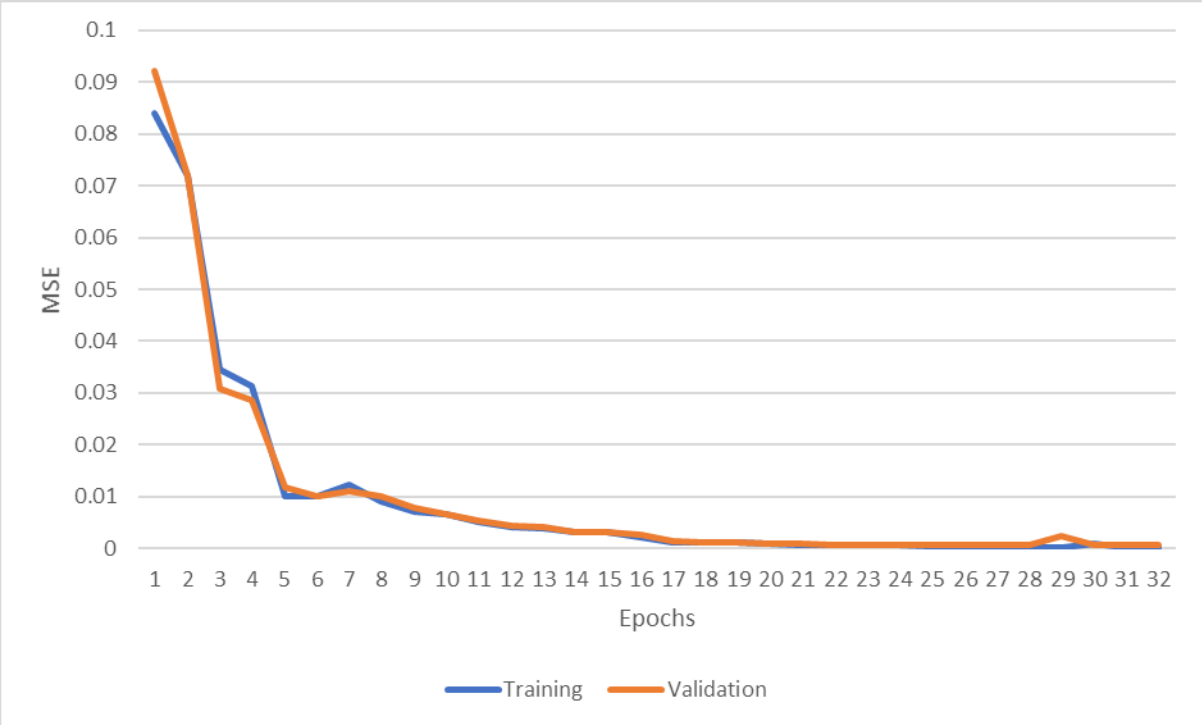

Figure 4-10: MSE vs Epochs

The samples were divided into 1338 points for training, 287 points for validation, and 287 for testing. Figure 4.11 presents an error histogram with 20 bins for the training, testing and validation. As the values are well-distributed around 0, there are no fitting errors.

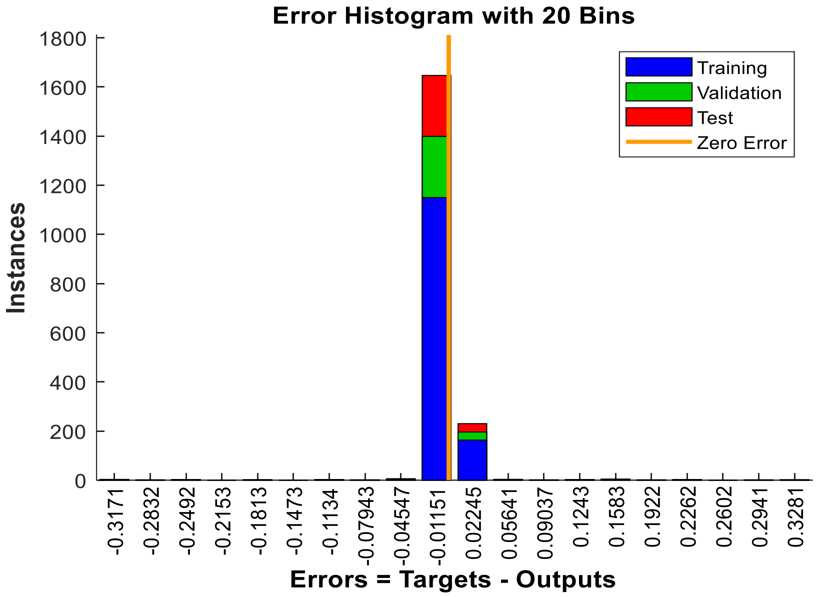

Figure 4-11: Error histogram with 20 bins

Figure 4.12 performs a comparison between the output variables and target variables in the training, validation and testing steps. “Target variables” is “Measured Comprehensive Strength, “Output values” is “Predicted Comprehensive Strength”, while “R” gives the efficiency of the model. “R” shows that the model has acceptable accuracy in the training and validation cycles. Thus, the R values for training, validation and testing are, respectively, 6.15265e-1, 5.80725e-1 and 5.97079e-1, evidencing the model's efficiency. Meanwhile, because there is a non-linear relationship between the input and the variables that configure the parameters, there are fewer errors, and the prediction’s accuracy is perfect

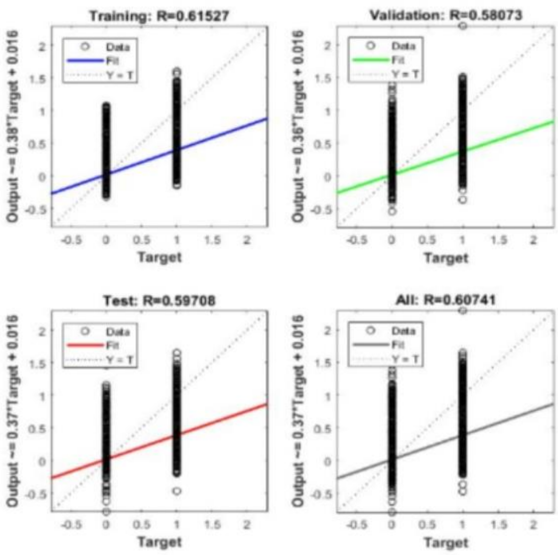

Figure 4-12: Training validation and testing

4.10 Summary

The development of a driver distraction severity prediction system substantially enhances the use of DL algorithms in classifying driver distractions. Based on LSTM, the testing dataset of this system shows a good MSE while also demonstrating effective classification. The proposed approach's assessment is based on a sample of the user testing data, comprising 25% of the images taken for one participant. The following chapter reviews the context data and determines how to automatically detect driver distraction to classify its severity based on a hybrid CNN-LSTM. In addition, the algorithm is reworked to incorporate more context-aware on the environment. Finally, the chapter explores the future applicability of integrating the algorithm into ADAS in semi-autonomous vehicles to allow these transitioning from the driver in situations that demand it.

CHAPTER 5. A FUZZY-LOGIC APPROACH TO DYNAMIC BAYESIAN SEVERITY LEVEL CLASSIFICATION OF DRIVER DISTRACTION

5.1 Synopsis

As distractions are a significant factor causing traffic accidents by affecting both the driver's behaviour the dynamics of the vehicle, their detection and classification are critical to preventing traffic accidents. For example, knowledge of the severity of driver distraction can help develop techniques to prevent accidents, such as transferring control to a level 4 semi-autonomous vehicle once a high driver distraction severity has been identified. Strengthening ADAS is key to enhancing road safety for all users. Here, drawing on a contiguous set of video frames from the Naturalistic Driving AUCDDD, a new technique is proposed to predict driver distraction severity based on an expert knowledge rule system. A Multi-class distraction system would be developed that incorporates face orientation, driver activities, the driver's hands, and any previous driver distraction to develop a severity classification model using a discrete dynamic Bayesian (dDDB).

Meanwhile, the severity levels of multiple classes of distractions are classified as safe, careless or dangerous driving using a Mamdani-based fuzzy system. This allows a semi-autonomous vehicle to take over from the driver if a high driver distraction severity is reached. Findings indicate that some forms of driver distraction can quickly shift from careless driving to dangerous driving in a multi-class distraction context.

This chapter makes the following main contributions:

- A rule-based driver distraction detection and classification system
- A severity classification system based on a Dynamic Bayesian Fuzzy logic model
- A system for classifying driver distraction based on the severity level, i.e. safe, careless and dangerous driving.

5.2 Dataset And Data Transformation

The dataset is extracted from the AUCDDD V2, was obtained via the Machine Intelligence group at the American University in Cairo (MI-AUC) [243]. This is the first publicly available distracted driver detection and was drawn from a study of 44 participants (29 male, 15 female) in seven countries, namely Egypt (37), Germany (2), USA (1), Canada

(1), Uganda (1), Palestine (1), and Morocco (1). Some of the participants were recorded more than once, i.e., under various driving conditions, and dressed differently. The recordings were made in five car models: Proton Gen2, Mitsubishi Lancer, Nissan Sunny, KIA Carens, and a prototype car. This research involved the extraction of 14,478 frames from the classes of safe driving (2,986), phone right (1,256), phone left (1,320), text right (1,718), text left (1,124), operating radio (1,123), drinking (1,076), hair or makeup (1,044), reaching behind (1,034), and talking to passenger (1,797).

The video files were manually inspected, and a distraction label was assigned to each frame. Any transitional actions that occurred between each consecutive type of distraction type were removed by hand. Table 5.1 presents three of the ten dataset classes that were utilised in this study. The chosen frame statistics are those containing driver activities such as Phone rights, Text rights and talking to passengers in sequence for a given period.

Table 5-1: Distraction Events Classes and Frame Number

DISTRACTION EVENT CLASSES	FRAME NUMBER
Phoning	1,256
Texting	1,718
Talking	1,797

5.3 Selection And Extraction of Distraction Features

The images in the dataset are labelled based on the driver's activities observed in the video following the extraction of the features according to the distraction class. The images are subsequently tabulated in the form of ground truth labels and regions of interest (RoI) by employing MATLAB’s 2019b Image Labeler Toolbox and Graphical User Interface (GUI) editor; these are then placed in fuzzy sets to classify each distraction according to its level of severity. In total, 150 images receive a label with at least one of three observed behaviours, i.e. face orientation, driver activity, and hands-on the wheel.

The ground truth label for a driver talking to a passenger as per the dataset is presented in Figure 5.1. The driver engages in a multi-class activity: talking to the passenger with their face oriented away from the road and both hands holding the wheel. Meanwhile, Figure 5.2 presents the driver similarly engaging in a multi-class activity, talking to the passenger, again with the face oriented away from the road, yet with only one hand on the wheel. Figure 5.3

shows the driver engaging in multi-class activity, namely talking to the passenger with their face oriented towards the road and both of their hands at the wheel. Figure 5.4 depicts an additional multi-class activity, namely texting with the face oriented towards the road and one hand at the wheel, while Figure 5.5 involves speaking on the phone with the face oriented away from the road and only one hand no the wheel. On a few occasions, the driver was observed to be having their face oriented away from the road with both hands on the wheel while talking on the telephone for 1 second (25 fps).

Figure 5-1: Ground truth label of driver activity: talking to a passenger, face orientation off-road, both hands on the wheel

Figure 5-2: Ground truth label of driver activity: talking to a passenger, single hand on the wheel

Figure 5-3: Ground truth label of driver activity: talking to a passenger, face orientation, both hands on the wheel

Figure 5-4: Ground truth label of driver activity: texting, face orientation on the road, a single hand on the wheel

Figure 5-5: Ground truth label of driver activity: phoning, face orientation off-road, a single hand on the wheel

5.4 Dynamic Bayesian Model Used to Classify Severity

Adopting a dynamic Bayesian model to classify severity based on distractions existing in the dataset in terms of those physiological features that the algorithm can detect. The fuzzy set has four inputs: hands, face orientation, driver activity, and last driver activity. The first frame of change is set at r = 0; r increases if there is no alteration to the distraction profile compared to the previous frame. Thus, the r-value is the first time a distraction occurs. A distraction's severity is calculated by (f_{n-1}. α), where α is the distraction likelihood function determining for how long the distraction has repeated; β_0 is the likelihood of the first occurrence of a distraction in a frame; and f_{n-1} is the existing evidence.

5.4.1 Dynamic Bayesian Fuzzy-Logic Model

Drawing on probability distribution components, namely the likelihood of future distraction and prior beliefs or observations of previous distractions in the dataset, construct the formal distraction severity. Thus, one may generate a distraction severity predictive system through the proper use of dynamic Bayesian methodology.

5.4.2 Distraction Type Likelihood Function

To assess the distraction type likelihood function, the probability of the same distraction type pattern occurring over a specific number of sequential (i.e., contiguous) frames is computed using

$$\alpha_r = \beta_0 + \left(1 - \frac{1}{r}\right), \qquad (5.1)$$

where β_0 is the likelihood that a new distraction will first occur during the exponential function $\left(1 - \frac{1}{r}\right)$ is the likelihood of its continuing to occur in subsequent frames, where *r* > 0.

5.4.3 Observation of Driver Distraction Features

Previous evidence drawn from ground truth labelling of the belief comprises the second probability component in the driver distraction severity classification model, thereby allowing the driver distraction features to be observed. Defining this probability function as:

$$f(x) \leftarrow fo_x^1 \oplus da_x^2 \oplus ha_x^3 \oplus \dots O_x^n . \qquad (5.2)$$

weighting the likelihood of a particular distraction severity level using the normalizing constant τ_α, i.e., taking into account how significantly each observatory dataset element is thought to contribute to the classification of the distraction severity level (τ_α= number of observable events).

Here, face orientation fo_x^1, driver activity da_x^2And hands-on the wheel ha_x^3 are all normalized between the interval [0,1], thereby representing existing evidence of the distraction features of the driver, that is, facial orientation, activity (talking, phoning, texting), and hand gestures (either one or two hands at the wheel). Finally, formulating the prediction of the overall distraction severity level classification as a discrete Dynamic Bayesian network (DDBN) model:

$$S_t(x) = \begin{cases} f_{t-1}(x)\alpha_r / \tau_\alpha , r \geq 2 \\ f_t(x)\beta_0 / \tau_\alpha , \quad r = 1 . \\ \quad 0, \quad at\ t = 0 \end{cases} \qquad (5.3)$$

This model is used to generate the test dataset based on the larger Distracted Driver Dataset. At this moment, there is the assumption that a severity probability of zero for the first timestamp (t = 0) in the video frame. If this is the first occurrence of the distraction feature pattern, i.e., $r = 1$, then the severity is computed using only the probability; the severity probability is computed for future occurrences using the abovementioned DDBN model. The thus transformed test data represent the groundwork for assessing the novel inference system based on fuzzy logic to determine the severity of certain driver activities that cause distraction.

5.5 Fuzzy-Based Dynamic Bayesian Model

The degree of distraction severity can shift from careless driving to dangerous driving if a secondary distraction occurs within a given time. There should be a justifiable minimum threshold for a distraction's severity level to be classed as either "safe," "careless," or "dangerous". For example, a 10-second hand gesture (adjusting the seat belt or control panel, waving at pedestrians) could be considered careless. Here, outlined are various measures to describe driving performance based on several physiological features: face orientation, hands, and type of distraction, with the latter, refers explicitly to talking, phoning, and testing because of the associated cognitive distraction. For instance, the severity of a multi-class distraction can be measured based on the length of the driver's conversation and additional factors, e.g., hand gestures or face orientation. The frame rate is a result of this used to calculate the time, that is, 25 fps. Thus, the driver's conversation length is measured using a sequence of frames containing the "talking" distraction type. The coding is designed to allow the classification decision to be made once the threshold of 125 consecutive frames (equivalent to 5 seconds) has been achieved.

5.6 Implementation and Results

Figure 5.6 presents the system developed here, which is derived from the Mamdani fuzzy inference model. The Mamdani approach is frequently employed in expert knowledge acquisition as it clarifies the human experience with more excellent intuition, making it ideal for examining decision-making that contains uncertainties demanding the knowledge of human experts.

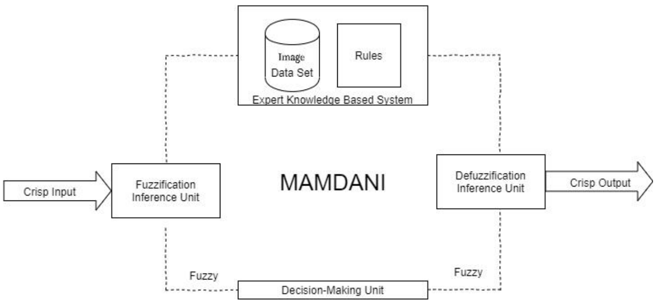

Figure 5-6: Mamdani inference model

Using the Mamdani approach to simulate actual driver performance and behaviour while driving. At this moment, giving each input a value and a certain number of MFs and then comparing the remaining inputs. The multi-inference Mamdani fuzzy model strives to detect multi-class distractions, enabling it to classify driving as safe, careless, or dangerous. The previous literature was used to produce the rule generation process, and previous studies on the specific distraction types were used to justify the weighting of each distraction. The labelling of the RoI comprises the feature extraction method and is integrated with the generated fuzzy rules. The distraction training data comprise classes containing the activities of talking, using the phone, and texting. The dataset is subsequently divided into subclasses, i.e., one hand at the wheel, speaking to the passenger, and orienting the face away from the road; the same procedure is performed on the testing data used in the validation. The rules are then inputted into the fuzzy inference engine, which employs Mamdani inference in line with the model architecture to detect distractions. MATLAB 2019b ground truth labelling is used to pre-process the data for feature extraction. In addition, each classification is assigned the MFs, associations, and rules. Finally, the rules for each driver distraction classification are tested using the testing data.

The distraction severity level measures the extent to which a driver distraction event affects driving performance. Classifying driver distractions into levels of severity plays a crucial role in transferring control to a semi-autonomous vehicle once a set distraction threshold has been reached. Following the abovementioned steps, the fuzzification process decomposes input and output into at least one fuzzy set. While it is possible to use several curve and table types, the most typical are triangular or trapezoidal-shaped MFs, as they can be more easily represented in the embedded controllers. Figure 5.7 presents the system of fuzzy sets for input using triangular MFs, whereby each fuzzy set is distributed across a region of input (or output) values plotted against membership. The scope is restricted to those activities that induce driver distractions, employing four parameters to detect the level of severity: face orientation fo_x, driver activity da_x, number of hands on the wheel ha_xAnd previous driver activity Pda_x.

Table 5-2: Driving severity level for membership functions

Description	Membership Function Range	Example of Driver Membership Functions	Distraction Severity Level
No distraction is observed	0 - 0.25	Talking to the passenger, two hands on wheel or single hand on the wheel, face orientation on road	Safe
Substantial level of distraction detected	0.25 - 0.75	Texting for less than 2 seconds, a single hand on the wheel	Careless
High level of distraction	0.75 - 1	Texting for more than 2 seconds but less than 5 seconds, a single hand on the wheel	Dangerous

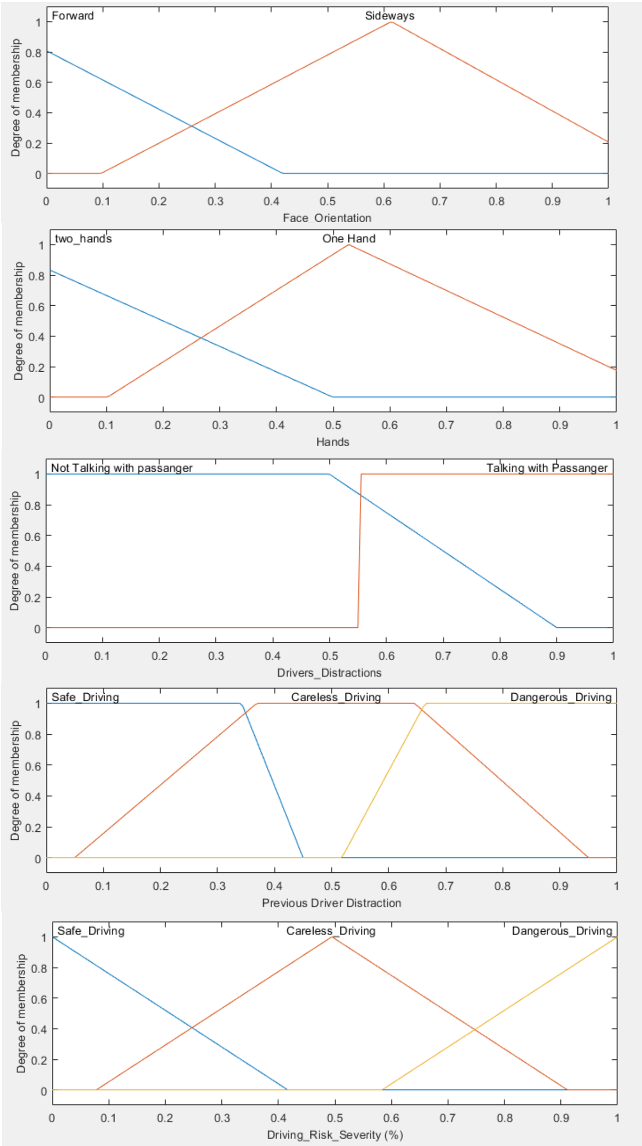

Figure 5-7: Inputs and membership functions.

A. MULTI-CLASS DRIVER DISTRACTION SEVERITY SCALE

Driver distraction is primarily analysed via the detection of the driver's activities. However, as driving uses physiological features with different levels of coordination, the effect of these actions are not always the same. Furthermore, they hypothesise that the severity level classification can cause any driver distraction to have a different impact. Testing this hypothesis by drawing on the previous literature in chapter 2 to justify the metrics for the various distraction types found in the dataset [17,18]. The severity level's category comprises the output of elements as represented by the MFs: safe driving = 0 - 0.25, referring to safe driving behaviour with a credible false distraction and an acceptable distraction event, e.g., changing gears; careless driving = 0.25 - 0.75, referring to a multi-class distraction event or a distraction combination; and dangerous driving = 0.75 - 1.0, referring to a highly critical distraction.

5.7 Rule-Based System

A rule base on fuzzy logic controls the output variable. A fuzzy rule is a simple IF-THEN rule that has both a condition and a conclusion. Table 5.2 above presents an example of fuzzy rules that classify the driver's distraction into severity levels. In Figure 5.3. A sample of 3 of 16 rules for the Mamdani fuzzy logic inference system to detect the severity level of a driver distraction is presented in the following:

Table 5-3: Fuzzy Rule Base

Rule	Face Orientation	Driver Activity	Hands	Previous Driver Activity	Severity
1 (system 1)	Forward	No talking to the passenger	Two Hands	Safe Driving	0-0.25 Safe Driving
9 (system 1)	Sideways	Talking to passenger	Two Hands	Safe Driving	0.25-0.75 Dangerous Driving
16 (system 1)	Forward	Talking with passenger	Single Hand	Safe Driving	0.75-1 Dangerous Driving
1 (system 2)	Forward	Not texting passenger	Two hands	Safe Driving	0-0.25 Safe Driving
9 (system 2)	Sideways	Texting with passenger	Two Hands	Safe Driving	0.25-0.75 Dangerous Driving
16 (system 2)	Forward	Texting with passenger	Single Hand	Safe Driving	0.75-1 Dangerous Driving
1 (system 3)	Forward	Not phoning passenger	Two Hands	Safe Driving	0-0.25 Safe Driving
9 (system 3)	Sideways	Phoning passenger	Two Hands	Safe Driving	0.25-0.75 Dangerous Driving
16 (system 3)	Forward	Phoning passenger	Single Hand	Safe Driving	0.25-0.75 Dangerous Driving

5.8 Results and Discussion

This section covers the frame-based rule-based fuzzy logic for driver distraction severity classification in terms of the outcome. The results for driver distraction are assessed through testing an unobserved dataset without fuzzy rules.

5.8.1 Surface Plots

Figure 5.8(A) presents a plot comparing face orientation with the driver's previous activity. At this moment, a yellow plateau region can be discerned, indicating uniform driver distraction severity. The sheer increase in blue is due to the orientation of the face shifting at around 0.4; in other words, the driver's face – and hence, gaze – is no longer on the road, resulting in a higher distraction severity level. The region with the blue curve demonstrates the driver's distraction with their face oriented towards the road, primarily between 0 and 0.4. Subsequently, there is a shift to a higher distraction severity level, with the driver orienting their face away from the road. This type of distraction is different even when the participant knows the road, particularly with multi-class distractions, e.g., the driver looks sideways more often. In addition, findings show that careless driving occurred more frequently than dangerous driving. A driver engaged in a conversation while orienting their face away from the road for longer than 5 seconds; such behaviour represents a critically severe distraction and can cause a fatal accident.

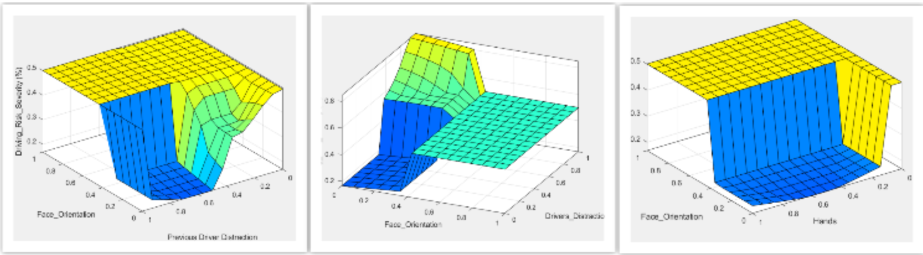

Figure 5-8: A,B,C. Surface plots for talking

Figure 5.8(B) compares face orientation with the driver activity of talking, whereby certain sections emerge. While dark blue signifies safe driving, cyan appears when the driver begins to engage the passenger in conversation; this represents a high level of distraction and can result in careless driving. However, it becomes dangerous driving once the driver stops looking at the road with a higher severity level.

Figure 5.8(C) graphs the position of the hands against the driver's face orientation. The curved area coloured blue indicates a sharp increase in the level of severity. Furthermore, the region coloured yellow portray enhanced distraction severity level due to the face being oriented away from the road.

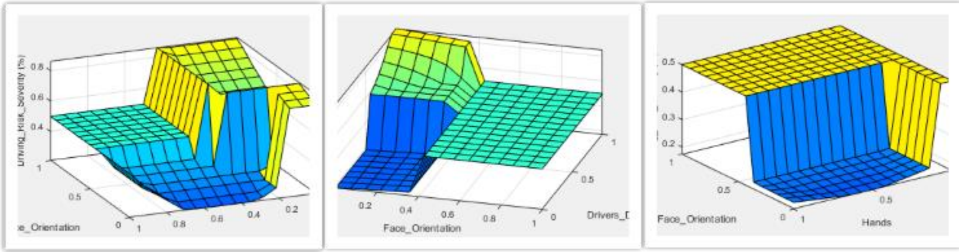

Figure 5-9: A,B,C. Surface plots for phoning

Figure 5.9(A) shows the impact on the distraction severity resulting from the driver's face orientation in phoning for a long duration. There is a further sharp increase at 0.4 as the driver takes their eyes off the road. Figure 5.9(B) indicates that the driver frequently orients their face away from the road when speaking on the phone, causing a higher level of distraction severity. Figure 5.9(C) portrays the face oriented away from the road, continuing until 0.4, when it changes and is directed towards the road. Furthermore, a brief occurrence of only one hand on the wheel is observed during the activity. Subsequently, there are occasions when there are no hands on the wheel; this represents a dramatic increase in the severity of the driver distraction.

Figure 5.10(A) depicts how the orientation of the driver's face contributes substantially to the distraction severity level of texting. This level is further increased by having a combination of texting and face orientation off the road. Meanwhile, as shown in Figure 5.10(B), the driver continuously engages in the texting activity for 2 seconds, increasing the severity level even further. Figure 5.10(C) similarly shows the driver having no hands on the wheel at 0.3, in addition to orienting their face away from the road, thereby sharply increasing the driver distraction severity level; this is classed as dangerous driving. Taken together, these plots demonstrate the correlation between the driver distraction severity level and the activity, i.e., talking and texting, as the probability of the driver's eyes being off the road increases.

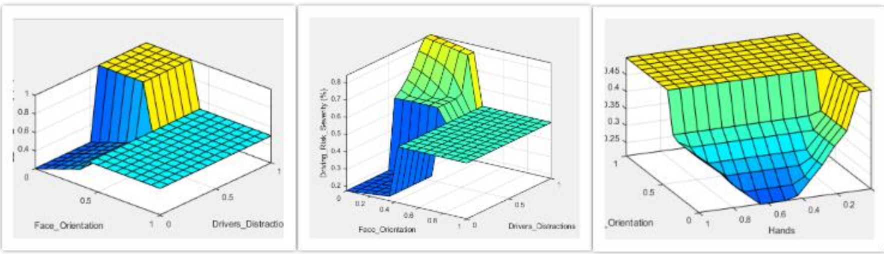

Figure 5-10: A,B,C. Surface plots for texting

Table 5.2 presents the definitions of the input values collected from the image labels dataset. The values are extracted from the labels and transferred into binary values, whereby

0 = false and 1 = true. The driver's previous activity is identified based on the previous frame.

5.8.2 Root Mean Squared Error

Tables 5.5, 5.6, and 5.7 respectively give the input test data for the multi-level distractions, the distraction severity level of the previous frames, and the outputs of the defuzzification methods. Adoption of the following defuzzification methods: Smallest of Maxima (SOM), with the defuzzified value being used as the element with the lowest membership values; Middle of Maxima (MOM), with the defuzzified value being used as the element with the median membership values; largest of Maxima (LOM), with the largest element from all membership values; centroid defuzzification, i.e. returning the centre of the area under the curve; and bisector, referring to the vertical line splitting the region into two sub-regions that each have an equal area.

Table 5-4: Driving Severity Levels for The Membership Functions

Face Orientation*(fo)*	Driver Activity (*da*)	Hands (ha)	Previous Driver Activity (*pda*)
0	1	1	0
0	0	1	0.06666666
0	0	1	0.06666666
0	1	1	0.33333333
0	1	1	0.44444444
0	1	1	0.5
0	1	1	0.53333333
0	1	1	0.55555555
0	1	1	0.57142857

Table 5.5 presents the values related to using the phone analysed in this scenario, with both MOM and centroid defuzzification yielding the results with the highest accuracy. LOM and SOM perform less well regarding the driving severity level because they only select extreme cases, generating an exaggerated crisp value. Specifically, LOM leads to an extremely high value, and SOM generates an extremely low value; these do not match the severity levels observed in either the weights or the MFs.

Table 5-5: Driving distraction severity defuzzification crisp output values for talking, using multiple methods

CENTROID	BISECTOR	MOM	SOM	LOM
0.494678671	0	0.495	0.12	0.87

0.466961833	0.47	0.495	0.1	0.89
0.466961833	0.47	0.495	0.1	0.89
0.596267826	0.64	0.82	0.64	1
0.71235178	0.76	0.82	0.64	1
0.807455156	0.81	0.82	0.64	1
0.807455156	0.81	0.82	0.64	1
0.81177008	0.81	0.83	0.66	1

Table 5-6: Driving distraction severity defuzzification crisp values for phoning, using multiple methods.

CENTROID	BISECTOR	MOM	SOM	LOM
0.494667	0.49	0.495	0.13	0.86
0.470227	0.47	0.495	0.11	0.88
0.470227	0.47	0.495	0.11	0.88
0.591258	0.63	0.825	0.65	1
0.708797	0.76	0.825	0.65	1
0.809211	0.81	0.825	0.65	1
0.809211	0.81	0.825	0.65	1
0.81177	0.81	0.83	0.66	1

Table 5-7: Driving distraction severity defuzzification crisp output values for texting, using multiple methods

CENTROID	BISECTOR	MOM	SOM	LOM
0.494679	0.49	0.495	0.12	0.87
0.467124	0.47	0.495	0.1	0.89
0.467124	0.47	0.495	0.1	0.89
0.588455	0.63	0.82	0.64	1
0.706695	0.75	0.82	0.64	1
0.806618	0.81	0.82	0.64	1
0.806618	0.81	0.82	0.64	1
0.81177	0.81	0.83	0.66	1

Table 5-8: Driving distraction severity levels for the membership functions

Defuzzification	RMSE	Driver Activity

Method	Value	
CENTROID	0.32	Talking
CENTROID	0.31	Texting
CENTROID	0.32	Phoning

Similar results based on the dataset are reported in Table 5.6; in this case, the centroid, bisector and MOM defuzzification crisp values best match the weights. The test demonstrates that the severity of the distraction rises with increasing duration. Recalling Table 5.4, it can be known that the number of continuous frames impacts the subsequent severity level. This is observed in the PDA column: the numbers steadily increase, yet they decrease with the severity level when that activity ceases. The crisp value output for the activity of texting is presented in Table 5.7. While the values are to those found for using the phone and talking, this activity has the highest severity level. Centroid, bisector, and MOM are the most accurate defuzzification methods here.

The calculation of the root mean squares is based on the dataset for the timeframes 1-47, and the severity levels of the driver distractions were estimated by computing the Root Mean Square Error (RMSE), based on the model of previous distraction severity:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{da,i} - X_{pda,i})x^2}{n}}, \quad (4)$$

da, i is the predicted value of driver activity, PDA, i is the previous driver activity (cf. Tables 6.5-6.7), and n is the data. According to the observed timeframes, the predicted value for the output defuzzification method is Centroid, as it most accurately reads the weights that have been assigned to the rules. Table 5.8 gives the results for the RMSE value with the most accurate error prediction for the previous and present severity frames. After comparing the Sugeno and Mamdani approaches, the latter shows better performance in this context regarding complexity, restrictive rules, accuracy, and modelling structure. Mamdani has a substantial advantage over Sugeno as it does not need all possible rule combinations to build the fuzzy rule base. Hence, Mamdani can non-linearly relate inputs with outputs via occasions of sharp transitions in distraction severity that range from high to low or low to high, as captured by the fuzzy membership functions. The outcome is a shift, with the semi-autonomous vehicle taking over from the driver once a given threshold has been reached.

In contrast, unsupervised learning based on classification techniques via a set of rules can profile the driver based on the distraction severity level. The classification methods construct rules by identifying patterns in the previous data of the driver or by predicting instances of driver distraction, particularly if the driver has already been monitored and profiled in a relevant setting. In addition, it should be possible to combine hybrid fuzzy-DL methods, such as CNNs.

5.9 Classifying the Distraction Severity Level from Careless to Dangerous

Certain driving behaviour elements must be identified and assessed before classifying a driver as careless. Adopted from the Crown Prosecution Service and Police charging standard, this non-exhaustive list of careless driving behaviours includes driving too close, lapses in attention, fatigue, falling asleep, using a phone, talking to passengers, missing traffic lights or signs, unsafe overtaking, and missing other vehicles or pedestrians [246]. Based on the driving law and CPS 1996 [11], inattention, referring to more than a very brief lack of attention, signifies careless driving and a longer lack of attention is considered dangerous driving. While the degree of distraction, i.e., inattention, could be considered subjective, in some instances, some of these behaviours imply careless driving, while in other – extreme – cases, they would be classed as dangerous driving. In addition, careless driving can shift to dangerous driving according to the driver's distraction severity level.

Thus, the question emerges at what point a driver's behaviour should be considered normal, careless, or dangerous. This highlights the need for a metric of the careless driving degree, thereby allowing a severity level to be assigned to potential incidents. This would facilitate the development of an ADAS system determined by the severity level of careless driving behaviour. Based on the existing literature, as well as the contribution above, the following is the most fitting way to define careless driving in the modern context of intelligent transport systems (ITS):

"Careless driving behaviour is a driving act that entails a deviation from normal driving behaviour, either by driver actions or emanating from an entity, such as a malicious cyber attacker, pedestrians, or the environment, which could be influencing the driver's behaviour and leading them not to give reasonable consideration to others, and thus, resulting in careless driving that can cause a casualty."

This chapter highlights a novel approach by drawing on real driving data to identify careless driving behaviour, specifically driver inattention, thereby providing realistic results. Table 5.9 compares the various events and distraction types that characterise careless and dangerous driving.

Table 5-9: Careless and Dangerous Driving Classification (RTA 1988)

CARELESS DRIVING	DANGEROUS DRIVING
Driving too Close	Fast Racing
Inattention Lapses of Fatigue	Aggressive Driving
Nodding Off (Eyes Closed)	Ignoring Traffic Lights
Mobile Phone Use	Violating Road Signs
Talking to Passengers	Dangerous Overtaking
Failure to See Traffic Lights	Ignoring vehicle faults
Unsafe Overtaking	Drowsiness Eyes Closed
Failure to See other vehicle and Pedestrians	Distraction (Handheld Phone)
	Inattention Lapses

Table 5.9 demonstrates that talking to passengers in conjunction with a multi-level distraction can cause a shift from careless driving to dangerous driving. The outcome of this study thus confirms the fact that careless driving can soon lead to dangerous driving.

5.10 Summary

In this chapter, there is an introduction of a method of evaluating driver distraction using fuzzy set theory. A rule-based fuzzy system was derived from an NDS dataset to detect multi-class distractions in image sequences. The severity levels of these multi-class distractions were calculated by combining the driver's activity, face orientation, number of hands on the wheel and previous activity. The inference system was able to classify a multi-class distraction's severity using various metrics, including the distraction's type, duration, and frequency. Findings show that the fuzzy logic inference system can detect and classify such multi-class distractions into either safe, careless, or dangerous driving. This method could be employed in developing ADAS to address the issue of driver distraction. Notably, while the previous literature demonstrates that texting and talking on the phone are more dangerous than careless driving behaviour, findings show that, as part of a multi-class distraction, talking to a passenger with the face oriented away from the road is nearly as dangerous as texting with the face oriented away from the road. This is because drivers who engage in conversations with their passengers tend to look away from the road, leading to a similar distraction degree for both activities. This chapter has contributed to the body of literature by facilitating the determination of a threshold at which a semi-autonomous vehicle should take over from the driver to become a level 4 semi-autonomous vehicle. In future research, the aim is to employ a neural network to classify driver distractions.

CHAPTER 6. MDDRA: A NOVEL CONTEXT-AWARE QUANTITATIVE RISK ASSESSMENT MODEL FOR SEVERITY LEVEL CLASSIFICATION

6.1 Introduction

Since the emergence of new in-vehicle technologies, distracted drivers have become a central problem in road accidents. Meanwhile, intelligent transportation systems will soon allow vehicles to control a semi-autonomous level 4 (or the level approved by the authorities) out of either necessity or driver choice. Thus, drivers may become more reliant on having the vehicle perform in-vehicle tasks, meaning they will become more relaxed and more distracted, thereby opening up many risks. In light of this, relevant context-aware (including vehicle performance and environmental conditions, which have direct implications for driver safety) can be utilized to help the driver account for many situations. This implies a need for an ADAS to mitigate risks before an accident occurs by providing a qualitative- and quantitative-based risk assessment.

The European Commission for Mobility and Road Transport Safety highlights that a significant proportion of road accidents occur when the driver is distracted, with common distractions encompassing handheld mobile devices, using the radio, eating, talking to passengers, smoking, and glancing in-vehicle navigation systems [247]. According to Kulkarni and Shinde [248], in-vehicle interfaces can also overload the driver. Additionally, fatigued drivers present a significant risk on the road. In recent years, the driver's eyes have become an efficient metric for measuring driver distraction, and the driver's ability to keep their eyes on the road is crucial. A statistical analysis by the Department for Transport shows that out of 1,456 fatal car accidents, 383 involved careless tendencies by pedestrians, while 110 resulted from drivers' reduced attention on the road [249]. Inexperienced drivers are another significant factor that has caused the number of road accidents to surge. Young and inexperienced drivers are particularly at risk, unlike skilled drivers, who adjust their driving strategy in time and predict different driving scenarios [250]. Compared to young drivers, the higher crash incidence is attributed to low cognitive ability [251] and a loss of attention due to distraction [252].

So there is a strong need for driver's risk assessment [7], which can provide an easy control shift to automatic driving, especially when the driver is intoxicated, unconscious or not available in the situation. Although, risk mitigation is tricker and challenging to model as official accident reports are relatively undetermined due to the possibility of numerous definitions of distractions or a country simply not collecting the data [253]. Furthermore, driver distraction can be influenced by the situation in which the driving occurs. Thus, there is a significant gap in the available mechanism that accommodates the context-aware risk model. The model should be concise enough with intelligent image recognition to detect and form a risk matrix to profile drivers into distraction classifications. This can reduce the occurrence of an accident by a significant margin. A plethora of literature is available [254]–[259] on the importance and urgency of driving risk mitigation techniques to prevent driving behaviour-related accidents and shift the control. For a false proof robust alert system, the precise classification of driving behaviour is needed. However, to the best of our knowledge, the current works lack complexity, rigidness, synthesized dataset, are more focused on a particular side of perspective (vehicle, driver, or environment), false-positive classes, and low accuracy.

Drivers can be classified into three groups, namely safe, careless, and dangerous drivers. Capturing driver behaviour is crucial to risk mitigation and developing context-aware ADAS may influence the risk levels and prevent accidents. Moreover, a real-time novel risk assessment determines both a driver's risk profile and the development of potential driver distraction, simultaneously working with multiple driving context influences, such as auditory, visual, cognitive, and biomechanical distractions.

Consequently, the critical contributions of this study are:

- Development of a definition of a severity level for driver distraction.
- A frame-by-frame analysis of driver behaviour severity level in an ADAS.
- A proposed model for characterizing driver behaviour considering context factors such as speed, acceleration, and surrounding vehicles.
- Development of the MDDRA model for driving behaviour and its evaluation using ML.

6.2 Risk Assessment Related to Driver's Distraction

Risk assessment can be defined as a process evaluating the adverse effects of a natural phenomenon, activity, or substance [260]. Berdica stated that risk constitutes the likelihood and probability of an incident occurring [261], [262]. Relative risk ratio has been used to quantify vehicle crashing risks under bad weather conditions; its calculation requires a large dataset of crashes arising from adverse weather conditions [263]. However, using a risk matrix, which combines probability and consequences, has overcome the former method in popularity [261]. A risk matrix can be used to determine the level of driving risk. Understandably, most risk indicators related to driver distractions have been modeled after crash events. However, the main flaw in modeling driving risk assessment via post-crash data is that it is a reaction strategy rather than a prevention method.

According to Cai et al [264], certain studies have shown that a driver's subjective assessment of the driving risks – particularly those related to various weather scenarios – is consistent with collision-based studies. In [264], the authors assumed that the driver's perceived risks are consistent with the actual crash statistics, especially for incidences related to rainy conditions. Various factors can impact driving capability; these can be extracted from the driving context, i.e., the driver, vehicle, and environment, including the weather, road, speed, manoeuvres, pedestrians, driver state, and braking. However, there is currently a lack of adequate data and facilities to ensure the development and implementation of an efficient and robust risk assessment model for the driving context. In response to this, this chapter proposes using the Naturalistic Driving Study TeleFOT, which is sufficiently complete for the environment, vehicle, and driver monitoring. The proposed approach uses the following mathematical model

$$C_i = \sum_j^J x X_{ij} \quad \beta_j + \ \varepsilon_i, \quad C_i = 1,2, \dots \dots, M \tag{6.1}$$

Where (C_i) denotes a discrete model-dependent variable that represents the level of a distraction's impact on driving. This variable's various impact levels include minor impact, overall impact, profound impact, and disastrous impact. The 'i' included in this variable represents the i^{th} driver with non-observable ε_i Variables, including the volume of traffic, vehicle type, road type, and rain intensity. A non-observable variable is selected to fit a logistic distribution for generating a continuous latent variable (C_i) denoting the influence on driving. Another proposed approach is the Rank Order Cluster Analysis, which sorts driving risk R_i in ascending order, as indicated by $R_1, R_2, R_3, \dots \dots, R_n$. Consideration of categories

(G), including $R_i, R_{(i+1)}, \ldots, R_j$ and satisfying $j > i$ can be denoted as $G = \{i, i+1, \ldots\ldots, j\}$. Consequently, the diameter of $G, D\ (I, j)$, is calculated from the equation:

$$\mathrm{D}(i,j) = \sum_{t=1}^{j} (R_{(t)} - R_{(G)})^2 \tag{6.2}$$

Where R_G represents the mean driving risk, and the driving hazard is segmented into k segments, expressed as;

$$G_1 = \{ i_1, i_1, +1, \ldots\ldots, i_2 - 1\},\ G_2 = \{ i_2, i_2, +1, \ldots\ldots, i_3 - 1\}, \ldots\ldots,\ G_k = \{ i_k, k, +1, \ldots., i_{k+1} - 1\} \tag{6.3}$$

Where the variable i satisfies the following condition:

$$I = i_1 < i_2, < \ldots. < i_k, < i_{k+1} = n + 1 \tag{6.4}$$

There is also a minimal loss function with a recursion relationship represented by the equation:

$$L[b(n,k)] = \sum_{t=1}^{k} D(i_{t,} i_{t+1,} - 1) \tag{6.5}$$

Where $b(n,k)$ denotes a special classification method. This loss function can be further explained as:

$$\underset{k \le j \le n}{L[b(n,k)]} = min\{L[P(j-1, k-1)] + D(j,n)\} \tag{6.6}$$

Here, $P\ (n,k)$ denotes the method to minimize the loss function; where n and k are given, $P\ (n,k)$ depicts the optimal driving risk categories. However, our proposed model assumes that driving is a discrete and time-series event; therefore, it takes the risk in the previous frame to compute the severity level of driving risk in the current frame. Furthermore, we consider the sequence of occurrence and the duration of the distracted driving event in computing risk. We address the limitations, thereby enhancing our proposed model.

6.3 The Multi-Class Driver Distraction Risk Assessment (MDDRA) Model

This section examines the perceived severity of naturalistic driving, thereby showing varying levels of human-perceived severity. Driver distraction can have a different impact on driving behaviour, and it can thus be classified as safe, careless, or dangerous. Basing on the previous literature, we developed and tested our hypothesis that driver behaviour based on driver distraction has different severity levels, as seen in the justification of metrics in chapter 2 and table 6-1. We then justify the weighting metrics for the distractions present in the TeleFOT dataset. The following observable parameters can characterize signs of attention deficit and fatigue in the driver: PERCLOS (PERcentage of eye CLOSure, i.e., the percentage of the time the driver's eyes are closed) [265], turning the head to the left/right to the body, tilting the head forward relative to the body (the moment when the driver is "nodding off"), duration and frequency of blinking, and the degree of openness of the person's mouth (a sign of yawning). In particular, for PERCLOS, there was a discrete number of parameters defined, namely P70, which is the proportion of time for which the eyes were closed of at least 70%; P80, which is the proportion of time for which the eyes were closed of at least 80%; and EYEMEAS (EM), which is the mean square percentage of the eyelid closure rating [265].

Furthermore, general information describing a driver helps to not only explicitly identify that driver among all other drivers who installed and used a particular monitoring software package, but it also helps to improve the search for and classification of drivers with similar characteristics (general patterns among groups would help to predict developing situations). This can be accessed via the database, with a weight coefficient applied since this is a "common" behaviour rather than an individual driver's behaviour.

Ginting et al [266] adopted a 5-point Likert scale to model anxiety about individual coronary heart disease at different levels. Lopez-Fernandez et al [267] also used a scale in assessing problematic internet entertainment among adolescents. The scale adopted was a self-administered scale for measuring the degree of severity of the behavioural addiction of online social network users and video gamers. Based on this, the formulation of the distraction severity levels. At this moment, the ratings of the severity level of distractions were designed using a 5-point Likert-type scale, as seen in Table I below [268]–[270].

The proposed model considered the severity level of driver distraction based on an observation of their driving history. While this can be unpredictable, we opted to analyze the driver's behaviour frame by frame to obtain intricate details. The following steps were taken:

- Decompose the video to a frame-by-frame level.
- Study each frame to assess its severity level.
- Aggregate the previous severity level of frames to the current frame severity level.
- Provide a precise class of severity based on the calculated severity level.

The following, outlining the essential aspects of our model for accessing the severity level of driver distraction. We acquired the knowledge and data by observing and analyzing individual frames from the input source. We began by formulating the risk assessment based on driver behaviour according to P = {p1, p2, p3, p(n..)}, as described in Table 6-1. Each parameter Pi is characterized by some set of action $A = \{a1, a2, a3, \ldots, a3\}$, with each action *ai* having a weight Wi.

Table 6-1: Parameters & Weightings

#	Parameter	Maximum Weight	Action	Weight
1	State of Hand	2	Double hands	0
			Single hand	1
			No hands	2
2	Road Type	3	Urban	1
			Dual	2
			Highway	3
3	Face Orientation	2	On road	1
			Off road	2
4	Illumination	1	Day	1
			Night	2
5	Eye Gaze	2	Eyes on road	0
			Eyes off-road	1
			Eyes shut	2
6	Weather	3	Dry	1
			Rain	2
			Snow	3

7	Manoeuvres	2	Stopped	0
			Turning	1
			Moving	2
8	Surroundings	2	Vehicle not present	0
			Vehicle present	1
9	Pedestrians	2	Pedestrian not present	0
			Pedestrian present	1
10	Speed	Urban 30 mph Single carriage 60 mph Dual carriage/motorway 70 mph	(Speed * Road Type)/210	Urban Single carriage Dual Carriage Highway

The next stage identified the severity levels according to severity rates, respective colour for identification, and classification label. For instance, if the severity is 0.0, the risk colour will be right green; no distraction from the driver has been observed, and it will have no impact on the driver's life. While, if the severity level is 0.9 or above, the risk colour will be red, and it will mean that a severe causality can be expected, and it is hazardous to keep driving. Table 6-2 provides these details along with the relevant consequences.

Table 6-2: Driving severity levels

Severity (0.0 – 1.0)	**Risk Color**	**Severity Levels**	**Distraction Class**	**Consequences**
0.0	Light Green	No Impact	Safe	No distraction observed
0.1-0.25	Green	Slight Impact	Safe	Slight distraction observed
0.25-0.399	Yellow	Low	Safe	Noticeable/substantial distraction
0.4-0.599	Dark Yellow	Medium	Careless	Level of distraction detected
0.6-0.79	Orange	High	Dangerous	Frequent level of distraction
0.8-0.9	Dark Orange	Very High	Dangerous	Casualty prone
0.9-1.0	Red	Extreme	Extremely Dangerous	Severe casualty Prone

6.3.1 Risk Assessment Matrix

An approach to the computation of risk assessment in a quantitative model uses a Risk Assessment Matrix's graphical tool. The risk matrix involves calculating the magnitude of the potential consequences scaled on the vertical axis (levels of probability) of these consequences occurring; technically, the probability of these consequences occurs on the horizontal axis. This facilitates an increase in the visibility of risk and impacts on the decision-making. The risk is computed by calculating the $Consequence \times Likelihood$ of Occurrence Likelihood: The likelihood depicts the probability of a driver's distraction being related to their context-awareness. Consequences/Severity Level: The occurrence of multi-class context-aware distractions is classified into severity levels of distraction.

6.3.2 Probability

Probability is the measure of the likelihood that an event will occur. For example, a possible aggregation can measure the number of times a driver experiences a particular distraction during a driving course. The driver may be profiled according to the distraction severity level at the end of the driving course.

6.3.3 Likelihood

The likelihood levels can be described as frequency values (duration course) and state values (every frame). Four impact levels are considered in this chapter, namely no impact, low impact, medium impact, and high impact; when an effect has no impact, the likelihood score is one, and the likelihood of that distraction observes no distraction or a distraction that has not currently occurred. When a slight distraction is detected, the impact is low, with a score of 2. A medium result is considered when a minor distraction has occurred, and the score is then set to 3; 4 implies a medium to significant distraction occurrence. More impacts can be seen in Tables 6-3 below.

Table 6-3: Severity Risk Matrix

CONSEQUENCES					
Extreme	7	7	14	21	28
Very High	6	6	12	18	24
High	5	5	10	15	20
Medium	4	4	8	12	16
Low	3	3	6	9	12
Slight/Very low	2	2	4	6	8
No Impact	1	1	2	3	4

The risk assessment values and their likelihood are explained in Table IV.

Table 6-4: Severity Risk Assessment Matrix

Risk Assessment Matrix	Likelihood
1	No distraction is observed or occurred yet
2	A slight distraction has been observed
3	A minor distraction has occurred
4	A medium or major distraction has occurred

The proposed model is implemented using a weighted average of the parameters to compute the severity levels per frame, as depicted in Table 6-4. These weights are capped by the maximum number that a parameter can take. For example, we take "State of Hand" as a parameter and grade it as 0 - double hands, 1- single hand, 2- no hands. If the value of a given frame for this parameter is x, then the weighted value is $\frac{x}{2}$ since the maximum value, this parameter can take 2. Let us generalize this for any parameter x_i with a maximum value m_i as follows:

Severity level => $\sum_{i=0}^{n} \frac{x_i}{m_i}$ Where n is the number of parameters we took into consideration.

6.3.4 Special Considerations

One of the patents held by MOVON Corporation [271] to ensure drivers' safety is a lane departure warning system based on image processing using a mono camera installed inside the car. A distinctive feature of the system is that it successfully processes several road conditions, including undesirable situations such as changing the width of the road lane, the radius of its curve, the direction of the road, and the complete absence of a road surface.

We realize that speed depends on the road type; hence, multiplying it with the weight of its road for speed. There is consideration of road types in the UK as this conforms to the source of the dataset. For the metric of road types, the threshold is defined according to the speed limit allowable on the road type, i.e., urban, single carriage, and motorway at 30 mph, 60 mph, and 70 mph, respectively. Furthermore, we define the following context data:

- Vehicle (V) and driver data with probabilities $P(V) = \{v1, v2, \dots, vm\}$
- Environmental data with probabilities $P(E) = \{e1, e2, \dots, en\}$,
- Speed a
- Surrounding $P(S)$

- Pedestrians $P(Pe)$

We formulate the following equations. The speed is computed as described in equation 7, for example, given that the national speed limit of UK is 70 mph and the maximum road type weight, and the score is 3:

$$\frac{(Speed)(Road\ Type)}{(Max\ Speed \times Max\ Road\ Type)} \tag{6.7}$$

We understand there are different data points in each frame; thus, the severity level of a given frame with k data points is:

$$S(fi) = \frac{1}{\mathrm{k}} \times (\sum_{i=0}^{m} \mathrm{P(Vi)} + \sum_{j=0}^{n} P(Ei) + a + P(S) + P(Pe)) \tag{6.8}$$

We now compute the aggregate severity (S*) of a given frame given the last $i - t$ frames. This is achieved by taking the average of the current frame's severity score compared to the severity score of the last $i - t$ frames:

$$\mathrm{S}(fi) = \frac{1}{\mathrm{t}}(\left(S(fi) + \sum_{j=t}^{i-1}(S(fj)\right)) \tag{6.9}$$

The verification and validation processes for the proposed model typically include both computational and physical aspects. To assess the degree of adequacy of the numerical modelling, the following steps can be performed: 1) Determine the order of convergence of numerical solutions in comparison with a numerical solution using a reduced number of parameters; and 2) assess the sensitivity of the sampling algorithm to various uncertainties, including parameter constraints, grid adaptation to real measurements and boundary conditions. Furthermore, validation assumes a careful comparison of the numerical calculation results of the phenomenon under study with experimental data to obtain an answer to the question "is the numerical solution correct?". Thus, a comparative analysis of the model with all the conditions, including the uncertainties associated with missing parameters and boundary conditions from the real world and computational points of view, is carried out. A few methods can be used for model validation purposes:

a. Evaluate the loss function; the squared error loss function applied to each training dataset can be used. This is referred to as L2 Loss, which is the square of the differences between the actual and the predicted values:

$$\mathrm{B.L} = \left(\mathrm{y} - \mathrm{f(x)}\right)^2 \tag{6.10}$$

b. Simulate the system to compare the real output $y(t)$ with the (noise-free) model output; this could involve a Bayesian approach.
c. Investigate frequency response, poles, zeros, and their uncertainty.
d. Analyze prediction errors (residuals) via a cross-correlation test: Are the residuals uncorrelated with the input?
e. Apply the model to unseen data (cross-validation). This strategy may be helpful since it establishes the robustness of the proposed model. It can also provide the basis for the hybrid cross-correlation validation since there is a need to separately investigate how the inputs and outputs are correlated and how this correlation is affected by our modelling scheme;
f. Apply an "Inverse Problem," i.e., acquire a solution to the problem and solve the inverse case to obtain the output parameters. This will help to validate the assigned weight coefficients and the overall parameterization scheme. In our case, the reliability of our modelling is tested by the following methods in section F below:

Cross-correlation test to analyze the residuals.

Hybrid cross-correlation test on the data obtained over two separate datasets, with the analysis, separately applied to the inputs and outputs.

6.4 Experimental Methods

A discrete-time model is proposed for the application of ML to detect the pattern in time-series driver distraction data. Consequently, the development of a model for predicting a driver's severity level based on distraction. The MDDRA model architecture illustrates the state flow of the data and system modules that constitute the entire system. The architecture is made up of six states:

6.4.1 Six States of MDDRA architecture

6.4.1.1 Data Collection

Figure 6-1 below shows that there is a collection of all required data (from vehicle and driver) using multiple sensors and video recorders. Sensor-based data, including road type, driver movements, driver face and head direction, vehicle speed, weather, and the surrounding driving environment, are collected on a real-time basis.

6.4.1.2 Object Extraction

This architecture module extracts distraction state information (gaze at something else, Overspeed, etc.), including the changing state of the distraction, and feeds it into a probabilistic model for labelling.

6.4.1.3 Data Labelling

The probabilistic model is further applied to the labelled extracted data, then used to train the system's core engine before the ML model is applied.

6.4.1.4 Real-time Monitoring

Context-aware real-time data from the real-time driving video streams of the internal and external sensors of the vehicle are monitored. The data are further analyzed, and feature extraction of both the driver and vehicle state-based data is performed; this is then fed into the ML model.

6.4.1.5 ML Model

The ML model takes in state-based data (eye gaze, state of the hand, speed, face orientation, manoeuvre) and training datasets to predict the level of distraction. The resultant model is the probability of the occurrence of driver distraction in the current distraction frame state $P(C_{t+1}|S_{t+1})$, measured as the state transition from the previous frame state, denoted as $P(S_{t+1}|C_t)$. If the severity of distraction is high, vehicle takeover operations take effect.

6.4.1.6 Vehicle Takeover

The severity level informs the decision to perform a vehicle takeover of the distraction detected by the ML. If the distraction passes the threshold, i.e., transitions from careless to dangerous, then the decision for the vehicle to transition from driver to vehicle is triggered.

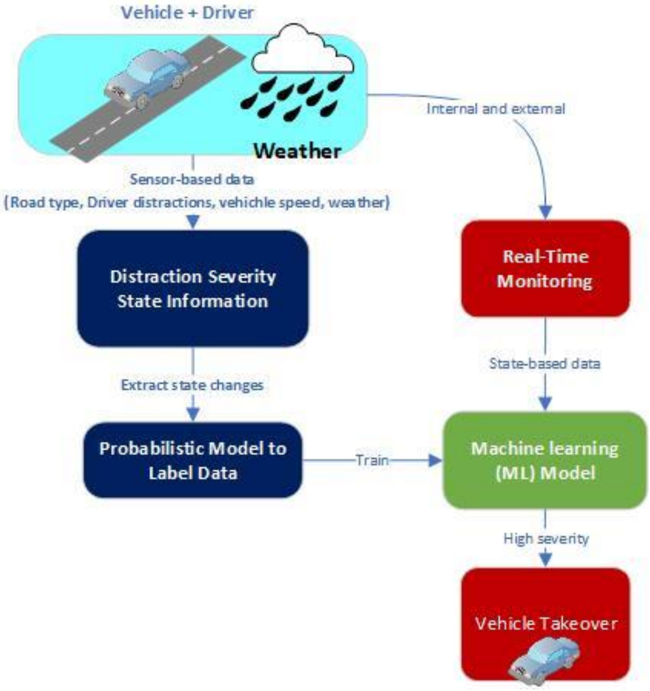

Figure 6-1: MDDRA Model Architecture

As depicted in Figure 6.2 below, a DBN is an extension of a Bayesian network that uses the time (dynamic) concept in modelling sequential time-series observations. It also uses a probabilistic inference model in handling uncertain information. An acyclic graphic represents the conditional independent and latent temporal variables discretely and continuously. For this case, the inference from the DBN is derived from three fundamental classes of nodes. Namely, driver features nodes, distraction identifiers, and context data. These inputs are represented in this model by nodes such as the state changes of the driver, consisting of 5 central nodes, namely face orientation, speed, manoeuvres, eye gaze, and state of the hands. The environmental changes node, consisting of road type, weather, and time of day, forms part of the context input data into the model and data on pedestrians and the surrounding environment. The final input is the distraction identifier derived from the analysis of the driver features by a hybrid CNN-LSTM. The output of this acyclic graph is a severity score, which measures the degree of the driver's distraction extracted from the driver features and context-aware.

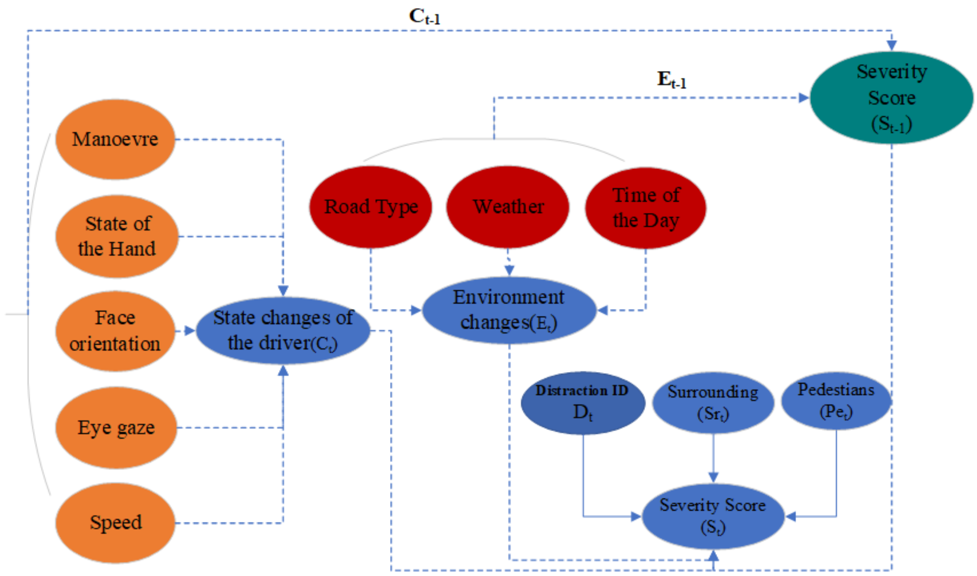

Figure 6-2: Context-aware probabilistic model for severity classification

6.4.2 Dataset

The TeleFOT Naturalistic driving study dataset is a European Field Operation Test (FOT) [272]. The project was designed to enhance research on intelligent transportation systems [273] [274]. The experiment was conducted in the UK and involved 27 participants [274]. Each driving video consists of four video channels that monitor in-vehicle and out-vehicle parameters, including face orientation, eye gaze, and hand position. The dataset consists of time-series data.

Figure 6-3: TeleFOT Dataset

6.4.3 Probabilistic Data Model

Considering the driver distraction state's changes frame by frame, as depicted in Figure 6.2 above. Technically, our proposed Context-aware probabilistic model for severity classification can be described as the probability of the occurrence of driver distraction in the current frame state $P(C_{t+1}|S_{t+1})$ from the previous frame state $P(S_{t+1}|C_t)$. *So*, there exists the probability of the occurrence of distractions in the environmental state $P(E_{t+1}|S_{t+1})$, if the current frame state S_{t+1} changes from the previous frame state S_t. The proposed extended dynamic Bayesian model includes several environmental variables such as road type, weather, and day. In order to compute the probability severity scores $P(S_{t+1}|C_{t+1}, E_{t+1}, Sr_{t+1}, Pe_{t+1}, S_t)$, here is a utilizing the dynamic linear model (Eq. 11) and which is a combination of the state change of driver distraction C_{t+1}, environmental changes E_{t+1}, distraction identification D_{t+1}, pedestrians Pe_{t+1}, and surroundings Sr_{t+1}.

$$P\ (S_{t+1}|C_{t+1}, E_{t+1}, Sr_{t+1}, Pe_{t+1}, S_t) \\ = P\ (C_{t+1}|S_{t+1}).P(E_{t+1}|S_{t+1}).P(Sr_{t+1}|S_t).P(Pe_{t+1}|S_{t+1}). \\ P(D_{t+1}|S_{t+1}).P(S_{t+1}|S_t) \qquad (6.11)$$

6.4.4 Interdependencies Test

We can apply the developed interdependency test for road type and its impact on driving speed. For example, in Table 6-4, the regression analysis coefficient is calculated as 0.529134, implying a significantly positive relationship. We assume that the driver would drive within the UK speed limit. The dataset of the driver may be more biased towards a degree of severity compared to other databases. Thus, it is necessary to validate the model using a regression model to test the interdependencies. We perform a correlation analysis between driver distraction and the severity classification of the distraction. Also, we conduct a multi-linear regression analysis to estimate the influence of driver distraction on the degree of severity classification.

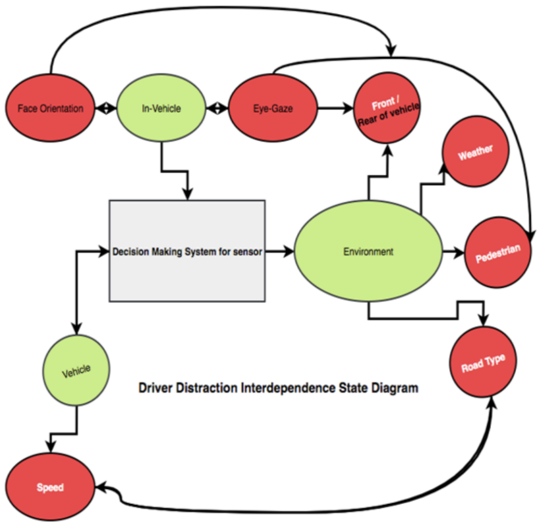

Figure 6-4: Distraction Interdependence State Diagram

6.4.5 Data Normalization

Logging and normalization of the vehicle's speed synchronously with the distraction state frame; thus, the time-series data of the vehicle are correlated at every frame. Subsequently, a regression analysis is conducted to validate our hypothesis, as seen in the results section. The severity level classification of the baseline drivers (baseline drivers in the original dataset) is likely to have a lower mean than experienced drivers. Furthermore, regression analysis indicates the mean of the in-vehicle parameters, mean vehicle data, and mean environmental data to produce the safe severity level. Meanwhile, in professional drivers, the safe severity level is likely to be more than the baseline. However, in other parameters, like per frame, severity means an aggregate severity means. In contrast, the statistical analysis of all the parameters contributes significantly towards the severity level considered safe, careless, or dangerous. However, only the vehicle speed distribution across the journey and its relation to the road type and driver distraction severity level has a substantial impact, with an intercept of 0.556982, as seen in Table 6-6.

6.4.6 Results of the Model Validation Procedure

To provide basic information about the variables in the dataset, the descriptive statistics for one of the simulated events (driver 1, event 1) are presented in Table 6-5. The mean, median, kurtosis, and skewness values are calculated by using equations 12-16.

$$mean = \frac{\text{Sum (x)}}{\text{Count (n)}} \tag{6.12}$$

$$\text{Median(x)} = \begin{cases} X\left[\frac{n}{2}\right] & ;\ \text{if n is even} \\ \frac{\left(X\left[\frac{n-1}{2}\right] + X\left[\frac{n+1}{2}\right]\right)}{2} & ;if\ n\ is\ odd \end{cases} \tag{6.13}$$

Where X is an ordered list of values in the data set, and n denotes the values in the data set (count). For a univariate data $y_1, y_2, \ldots, y_N$ the formula for skewness is:

$$g_1 = \frac{\sum_{i=1}^{N} \frac{(y_i - \bar{y})^3}{N}}{s^3} \tag{6.14}$$

Where $\bar{y}$ Shows the mean, s is the standard deviation, and N is the number of data points. Note that in computing the skewness, the s is computed with N in the denominator rather than $N - 1$.

$$\text{kurtosis} = \frac{\sum_{i=1}^{N} \frac{(y_i - \bar{y})^4}{N}}{s^4} \tag{6.15}$$

$$S = \sqrt{\sum \frac{(y_i - \bar{y})^2}{N}} \tag{6.16}$$

The results of these metrics suggest that this is a symmetrical distribution. This reflects how the data were modelled. It would be valuable to deploy this model using real data from the video sensor to access the accurate distribution of parameters, such as face orientation and eye gaze, and then analyze the results.

Table 6-5: Descriptive statistics

Mean	0.513049625
Standard Error	0.007304311
Median	0.508023896
Standard Deviation	0.118456024

Sample Variance	0.01403183
Kurtosis	0.057262351
Skewness	0.217203269
Range	0.725482175
Minimum	0.162361126
Maximum	0.887843301
Sum	134.9320515
Count	263
Largest (1)	0.887843301
Smallest (1)	0.162361126
Confidence Level (95.0%)	0.014382625

In order to validate the model, its predictions are tested using correlation analysis, as suggested in Section 6.3.4. This technique is typically used to test relationships between quantitative or categorical variables. Correlation coefficients have a value of between -1 and 1. A "0" value means no relationship between the variables, while -1 or 1 means a perfect negative or positive correlation (negative or positive correlation here refers to the type of graph the relationship will produce).

Table 6-6:Correlation coefficients

State of Hand	0.425847
Road Type	0.363796
Face Orientation	0.420461
Time of day	0.224532
Eye Gaze	0.296584
Weather	0.247372
Maneuver	0.323121
Speed	0.053056
Surrounding	0.441935
Pedestrians	0.255076

From table 6-6 above, it is clear that there is a positive correlation with all but one of the parameters used in the model, namely the speed of the vehicle. The model is also tested across multiple events, and the results demonstrate a consistent lack of correlation with vehicle speed. This might indicate a need for a wider speed span in the dataset or better represent the model's influence if this does not affect the results. The speed span is evenly distributed due to driving conforming to the speed obtainable on the road.

6.5 Results and Analysis

Implementing our model and architecture in Figure 6.1 was carried out to determine which ML model will best predict driver distraction to aid vehicle take over decision-making.

Furthermore, to avoid biases, the experiment results were determined using different ML algorithms on the dataset. This is analyzed using the scatter plot and confusion matrix of the predicted class.

6.5.1 Interdependency Test using Regression Analysis

The regression analysis was performed on the following three context-aware features:

a. The in-vehicle features related to the driver distraction such as hand moment, gaze;
b. The vehicle features such as vehicle speed, manoeuvres;
c. The environmental features such as pedestrians, vehicles, weather.

There is a need to test the association and interdependencies between a pair of distractions. We applied a regression to the prediction of driver distraction divided into severity levels. Here, we further tested the relationship between distractions by classifying distraction into either in-vehicle, context-aware, or environmental distraction.

6.5.1.1 Driver Distraction

The driver distraction features consist of state of hand, face orientation, and eye gaze. Figure 6.5 depicts driver distraction, showing a strong relationship between eye gaze on the road (Eye Gaze 0), face orientation off-road, single hand on the wheel, and a high severity level score of distraction leading to dangerous driving. Eyes shut, face orientation on the road, and double hands-on wheel also significantly impact the severity score.

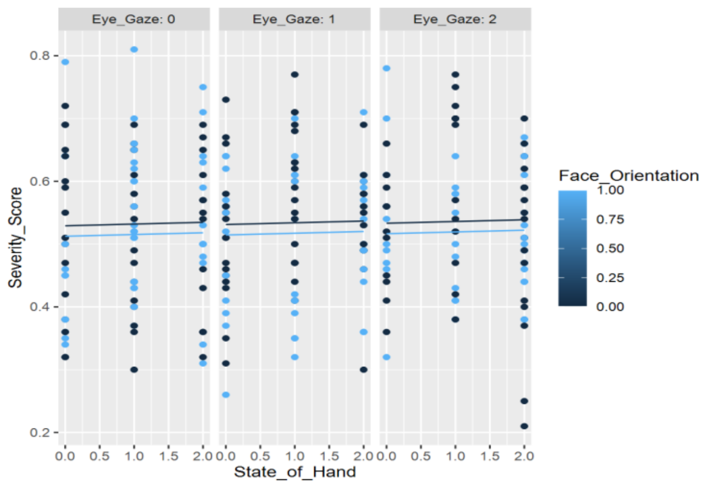

Figure 6-5: In-Vehicle State of Hand, Eye Gaze and Face Orientation.

Table 6-7 presents the in-vehicle distractions regarding the prediction of the severity score. Based on the P-value of 0.758, the probability of the state of hands to predict distraction severity is low. The intercept of 0.529134, which is highly significant, suggests a relationship between the state of hands, face orientation, and eye gaze. The statistical predictors use the t-statistics and P-values of each distraction. The lower P-value of 0.249 for face orientation shows a highly significant predictor of the severity score. The coefficient of determination is 0.005668.

Table 6-7: Driver Distractions Regression analysis

Intercepts	**Estimated**	**Standard Error**	**T-Value**	**P Value**
State of Hand	0.015526	0.015526	34.080	0.758
Face Orientation	0.1-0.25	0.014414	-1.156	0.249
Eye Gaze	0.002045	0.008981	0.228	0.820
Intercept	0.529134	0.015526	34.080	<2e-16

6.5.1.2 Environmental Distraction

Figure 6.6 shows that dry weather, a dual carriageway, and a bright day achieved the maximum dangerous severity score, while rainy weather, double carriageway, and night produced a slightly riskier situation. Snowy conditions on the highway and night had the highest degree of influence on the severity score.

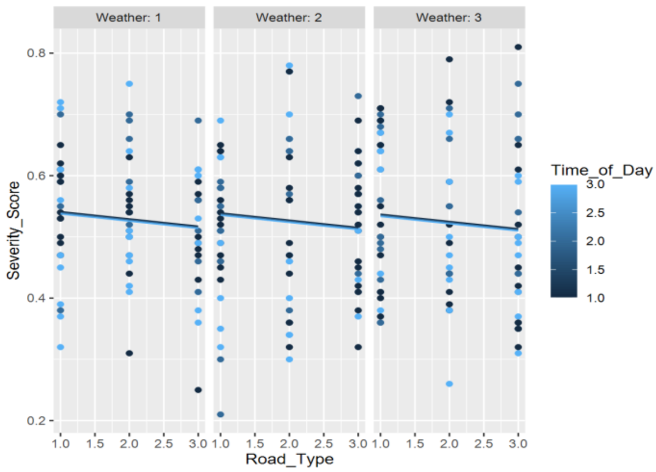

Figure 6-6: Road Type, Time of Day, and Weather

The results in Table 6-8 show environmental distraction in the prediction of the severity score. The low P-value of 0.175 for the road type shows that road type significantly impacts the prediction. The environment intercept of 0.556982 showed a significant association between the severity score and the outcome distraction road type, time of day, and weather.

However, the least average P-value of 0.5897 was the least compared to the other distraction classifications. The least residual standard error with 0.1165, the smallest of all the residual standard errors, indicates that this model best fits the data.

Table 6-8:Environment regression analysis

Intercepts	Estimated	Standard Error	T-Value	P Value
Road Type	-0.011838	0.008702	-1.360	0.175
Time of Day	-0.011848	0.008579	-0.215	0.830
Weather	-0.002086	0.009035	-0.231	0.818
Intercept	0.556982	0.031870	17.477	<2e-16

Figure 6.7 shows five instances of vehicle presence and pedestrian presence resulting in a hazardous distraction classification, which means if there are vehicles or pedestrians present in the surroundings, the chances of driver's distraction are significant.

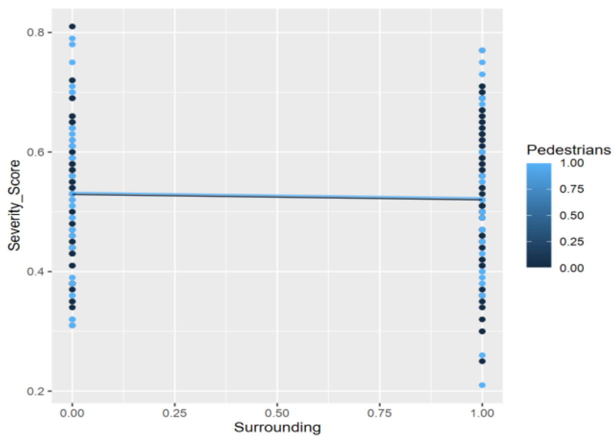

Figure 6-7: Pedestrian and Surrounding (Vehicle Presence).

Table 6-9 presents the environmental distractions to the prediction of the severity score. Based on the P-value of 0.532, the probability of the surroundings influencing the prediction is low. The intercept of 0.529157, which is significant, suggests a relationship between surroundings and pedestrians. The statistical predictors use each distraction's t-statistics and P-values. The lower P-value of 0.532 for the surroundings (vehicle presence) is a highly significant predictor of the severity score. However, the P-value of 0.830 of pedestrians suggests no association between pedestrians and the severity score.

Table 6-9: External distractions regression analysis

Intercepts	Estimated	Standard Error	T-Value	P Value
Surrounding	-0.009065	0.014502	-0.625	0.532
Pedestrians	0.003121	0.014487	0.215	0.830
Intercept	0.529157	0.013145	40.255	<2e-16

6.5.1.3 Vehicle Distractions

Vehicle distractions include manoeuvres and speed. Figure 6.8 shows the distraction within the speed range of 23 mph to 26.2 mph due to a high frequency of speed manoeuvres. There are a few outliers with very high severity and very high danger levels.

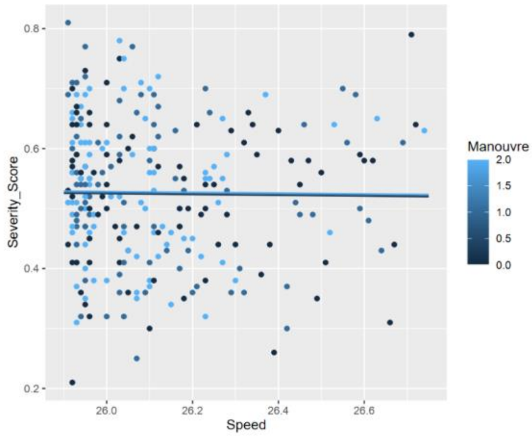

Figure 6-8: Vehicle, Speed and Manoeuvre.

The results in Table 6-10 show the influence of vehicle distractions on predicting a severity score. Based on the P-value of 0.855, the probability of speed influencing the prediction is low because the driver stays within the speed limit. However, during manoeuvres, there is a higher degree of significance. The intercept of 0.695812, which is highly significant, suggests a strong relationship between speed and manoeuvre. The statistical predictors use the t-statistics and P-values of each distraction. The lower P-value of 0.815 for manoeuvres shows a highly significant predictor of the severity score.

Table 6-10:Vehicle Distractions Regression analysis

Intercepts	Estimated	Standard Error	T-Value	P Value
Speed	-0.009065	0.035983	-0.183	0.855
Manoeuvre	0.002050	0.008758	0.234	0.815
Intercept	0.695812	0.941042	0.739	0.460

In this case, the driver is tested with the previous severity score and the predicted actual severity score of the next video frame; in this experiment, the driver's overall performance is tested throughout the drive, whereby there are a total of 262 frames, which is the equivalent of approximately 11 seconds. In Figure 6.9, we can assume that the driver has maintained a primarily constant careless driving behaviour. Furthermore, the regression analysis shows a strong correlation between the previous severity and the current severity scores.

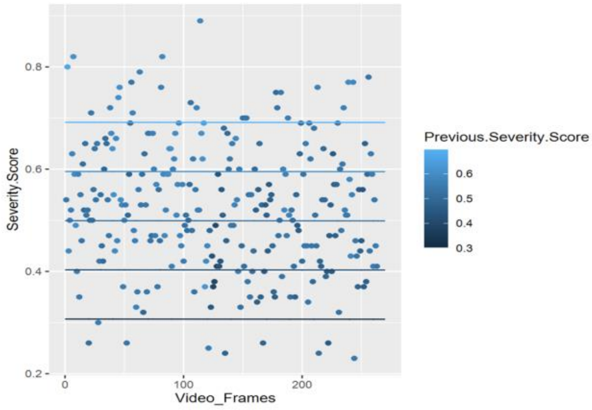

Figure 6-9: Severity Score, Previous Severity Score and Video Frames

The results presented in Table 6-11 depict the influence of in-vehicle distractions on predicting the severity score. Based on the P-value with the lower value of 0.990, the state of hand's probability predicts a low score. The intercept of 1.828e, which is highly significant, suggests a relationship between the sequence of video frames and the previous severity score.

Table 6-11: Severity Score Regression analysis

Intercepts	Estimated	Standard Error	T-Value	P Value
Video Frames	-1.329e-06	1.024e-04	-0.013	0.990
Previous Severity Score	9.618e-01	9.618e-01	5.812	1.8e-08
Intercept	1.828e-02	1.828e-02	0.196	0.845

6.5.2 ML Model

Different ML models are implemented, such as discriminant, naïve Bayes, SVM, K-Means Nearest Neighbour (KNN), and Ensemble ML. To better evaluate the performance, the classification results are compared in Table 6-12.

Table 6-12: Classification Performance

Classifiers	Model Type	Accuracy	Prediction Speed(~obs/sec)	Training Time (sec)
KNN	Fine KNN	79.1	2700	4.4574
	Medium KNN	78.3	2500	3.5617
	KNN Coarse	59.3	2500	4.4974
	KNN Cosine	80.6	2600	4.368
	KNN Cubic	76.4	2000	4.239
	KNN Weighted KNN	80.6	2500	3.975
Discriminant	Linear Discriminant	90.9	2700	3.5265
	Quadratic Discriminant	82.9	2500	5.2346
Naïve Bayes	Gaussian Naïve Bayes	93.2	3000	5.0814
	Kernel Naïve Bayes	90.1	1500	5.9402
SVM	Linear SVM	92.0	2400	4.9151
	Quadratic SVM	92.4	2300	4.8007
	Cubic SVM	92.4	2300	4.6915
	Fine Gaussian SVM	58.6	2200	5.7229
	Medium Gaussian SVM	85.2	2100	5.5983
	Coarse Gaussian SVM	77.2	2300	5.4722
Ensemble	Boosted Trees	58.6	3600	4.5331
	Bagged Trees	96.2	1000	6.3019
	Subspace Discriminant	92.4	780	6.8675
	Subspace KNN	79.8	600	6.7319
	RUSBoosted Trees	74.5	2900	4.6438

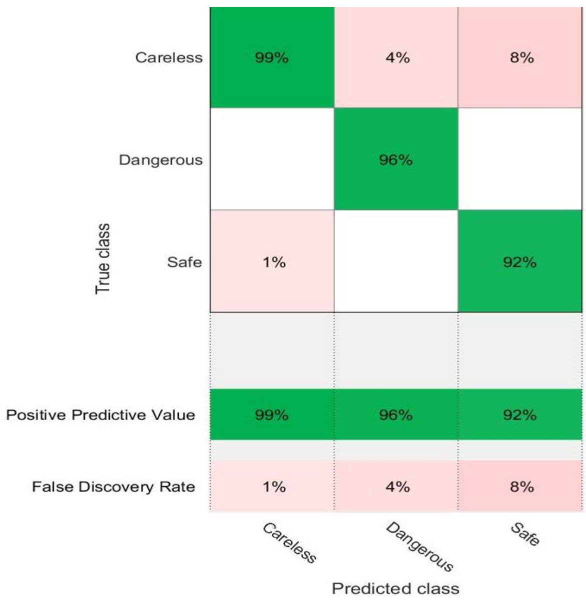

Figure 6-10: Confusion Matrix

In figure 6.10, the first two diagonal cells show the percentage of correct classification by the trained network. For example, 142 frames are correctly classified as careless. This corresponds to 99% of the 262 frames. Similarly, 80 cases are correctly classified as dangerous, corresponding to 96% of all the edges. Three dangerous and three safe instances are incorrectly classified, corresponding to 12% of all 264 frames in the data. Similarly, one of the careless structures is incorrectly classified, corresponding to 1% of all the data. Out of 148 careless predictions, 99% are correct, and 1% are wrong. Out of 80 dangerous predictions, 96% are correct, and 4% are wrong. Out of 35 safe cases, 92% are correctly predicted as safe, and 8% false.

6.5.3 Scatter Plot

Figure 6.11 shows a strong, linear association between the previous severity score and the observed severity score. In this case, the driver appears to progress from safe driving to dangerous driving, and the scatter plots made 355 correct predictions from the 263 total observations. In Figure 6.10, the consideration of the use of three observations, namely safe, careless and dangerous. The example predictors are given the severity score of 0.8373, which is predicted with a hazardous class and entails ten predictors. Furthermore, the results show that the intercept of the severity score of 0.6 and 8 observations is realised, considered careless. The severity score of 0.39 and 4 comments refers to a safe driving prediction, a correct classification.

The scatter plot predicts 74 safe driving instances, 142 careless driving instances, and 134 dangerous driving instances. In total, we have 355 total predictions for the Ensemble Bagged Trees. Finally, 13 numbers return a safe classification with 62 predictors. Other classification models, such as linear SVM, produced nine negative predictions, while Gaussian Naïve Bayes returned nine cases of negative predictions. In this case, the authors consider adopting Ensemble Bagged Trees to provide the best accuracy, with seven negative predictions.

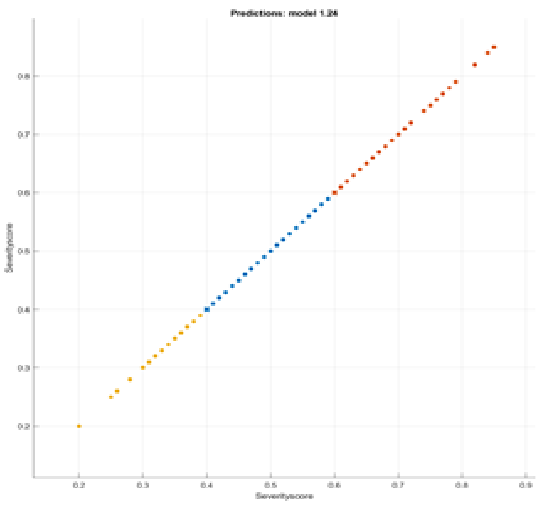

Figure 6-11: Scatter Plot

The Kruskal–Wallis rank was obtained using ML algorithms to confirm accuracy, training time, and prediction time; these are presented in Table 6-13. It can be observed that ensemble learning with the Bagged model obtained the highest mean rank of 21 compared to the other variants of that model and the other state-of-the-art ML algorithms. However, the mean rank for prediction and training time are 3 and 19, respectively. This phenomenon indicates that Bagged's complex fitness function helps extract rich feature vectors for classification. As opposed to Bagged Trees, ensemble learning with the Boosted Trees model obtained the highest mean rank of 21 than the other variants of that model and other state-of-the-art ML algorithms. This phenomenon shows that Boosted Trees' linear fitness function helps extract poor feature vectors for classification. Furthermore, the linear discriminant variant obtained the lowest mean rank of 1 in training time, with 90% accuracy.

The linear functions were evaluated using the previous severity score and the next video frame's expected total severity score. Gaussian Naïve Bayes appeared as the second-best algorithm for performance compared to others, except for Bagged Trees. The average mean rank gained by Gaussian Naïve Bayes is 20, with 93.2% accuracy and a z-score of 1.49.

Table 6-13: Kruskal- Wallis ranks were obtained using ML algorithms to confirm accuracy, training, and prediction.

		KRUSKAL–WALLIS AVE RANKS			Z
MODEL	MEDIAN	ACCURACY	PREDICTION SPEED(APPX - OBS/SECONDS)	TRAINING TIME (SECONDS)	SCORE
BAGGED TREES	96.2	21	3	19	1.65
BOOSTED TREES	58.6	1.5	21	8	-1.57
COARSE GAUSSIAN SVM	77.2	6	9	15	-0.83
CUBIC SVM	92.4	18	9	10	1.16
FINE GAUSSIAN SVM	58.6	1.5	7	17	-1.57
FINE KNN	79.1	8	17.5	6	-0.5
GAUSSIAN NAÏVE BAYES	93.2	20	20	13	1.49
KERNEL NAÏVE BAYES	90.1	14	4	18	0.5
KNN COARSE	59.3	3	13.5	7	-1.32
KNN COSINE	80.6	10.5	16	5	-0.08
KNN CUBIC	76.4	5	5	4	-0.99
KNN WEIGHTED KNN	80.6	10.5	13.5	3	-0.08
LINEAR DISCRIMINANT	90.9	15	17.5	1	0.66
LINEAR SVM	92	16	11	12	0.83
MEDIUM GAUSSIAN SVM	85.2	13	6	16	0.33
MEDIUM KNN	78.3	7	13.5	2	-0.66
QUADRATIC DISCRIMINANT	82.9	12	13.5	14	0.17
QUADRATIC SVM	92.4	18	9	11	1.16
RUSBOOSTED TREES	74.5	4	19	9	-1.16
SUBSPACE DISCRIMINANT	92.4	18	2	21	1.16
SUBSPACE KNN	79.8	9	1	20	-0.33

6.5.4 Time Complexity

Safety in intelligent transportation systems (ITS) is critical and having a fast ML model to make an efficient decision is crucial for road user safety. The linear discriminant gave the shortest training time of 3.5265, allowing faster decision-making. However, when predicting speed observed per second, the results showed ~3600 obs/sec.

Each classifier's residual value is examined by calculating the mean difference of training time and prediction time. Figure 6.12 shows that training time is an independent factor and indicates no significant deviation from the prediction time. Residual values (cf. y-axis) show that the prediction was exceedingly low. Fitted values (refer to the x-axis) show that the prediction was significantly accurate; 0 on the y-axis indicates a 100% correct positive rate.

Figure 6.12 shows that the fitted line's intercept and slope values are projections for the distribution's position and residual parameters, respectively. Simultaneously, the percentage on the y axis is helpful for the probability curve since the sample variance approximates the accuracy, prediction time, and training time obtained using several ML algorithms. Furthermore, the histogram of residual values indicates the distance between the observed prediction time from the mean of each classifier's total time for training. The significant residual value between -200 and +300 (refer to figure 6.12) shows an optimal configuration for the proposed framework when employing ML variants.

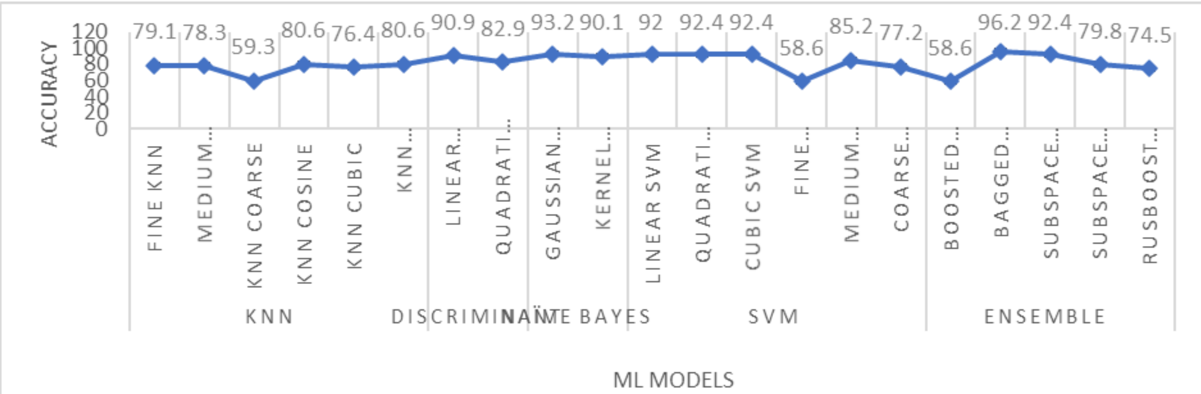

Figure 6-12: Show the comparison of accuracy across multiple ML models

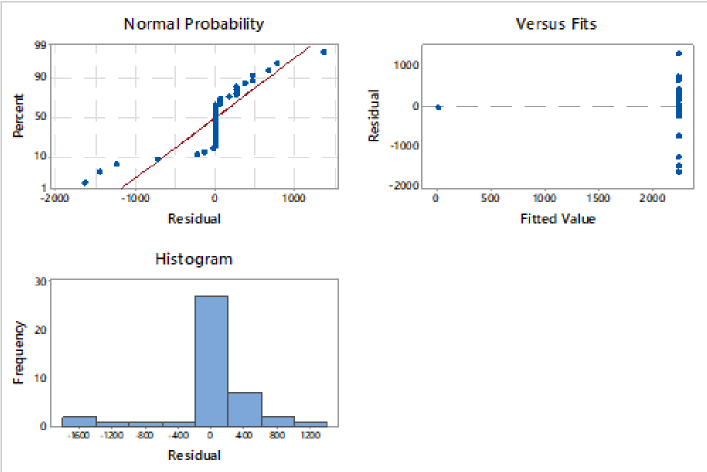

Figure 6-13: Residual Plots for Prediction Speed (~obs/sec) vs Training Time (sec)

6.6 Discussion

This chapter deals with the problem by employing ML. The authors have proposed a novel and robust Multi-class Driver Distraction Risk Assessment (MDDRA) model. The model has tackled the driver with almost possible variants such as the current state of hand, which means whether the driver uses double hands, single hands, or no hands at all. Similarly, the type of the road on which the vehicle is running, the face orientation is on the road or off-road, whether it is a daytime or nighttime, the eye gaze of the driver, if the weather is dry, rain, or snowy, what is the current manoeuvre, the surrounding vehicles, speed of the vehicle, speed of the surrounding vehicle, and the pedestrians. The suggested model, MDDRA, considers vehicle, driver, and environmental data during a journey to categorize drivers into a risk matrix such as safe, careless, and dangerous.

The proposed model offers flexibility to adjust parameters and weights to consider each event's specific severity level. Real-world data was collected using the Field Operation Test (TeleFOT), which consisted of drivers using the same routes in the East Midlands, UK. The results have a massive potential to reduce road accidents caused by driver's distractions. We have also tested the correlation of driver's distraction (In-vehicle, vehicle, and environment distractions) on severity classification against continuous driver's distraction severity score. Furthermore, we have applied several ML techniques to classify and predict driver's distraction according to severity levels to aid transitioning from driver to vehicle.

As implemented with different ML models such as Discriminant, Naïve Bayes, Support Vector Machine (SVM), K-Means Nearest Neighbour (KNN) Ensemble ML for classification. The above figure shows the comparison of accuracy by applying these models.

It can be seen that the Bagged Trees-based Ensemble model has provided the highest accuracy of 96.2% for classification, while fine Gaussian SVM and Boosted Trees-based ensemble methods have resulted in the lowest accuracy of 58.6% for the classification task. The comparison of various ML models is shown in Figure 6.13.

The graph in Figure 6.14 compares the accuracy value of the proposed MDDRA, the work of Mengtao Zhu et al [275], the work proposed by Yanli Ma et al [276], and the work of Tianchi Liu et al [277]. It can be seen that the proposed model has outperformed the current state of the arts in the multi-class distraction prediction. Moreover, the model has achieved an accuracy of 96.21%, while the current state of the art claimed accuracy of 95.87%, which is lower than our proposed methodology. Although Tianchi Liu et al [277] have achieved slightly higher accuracy, they have worked on a binary classification problem. The multi-class classification is a more complex task than a simple binary classification model, the model state-of-the-art with excellent results in more than eight classes. Furthermore, the proposed model has provided fast results as high as 3600 observations per second, making the proposed model accurate but robust in terms of speed.

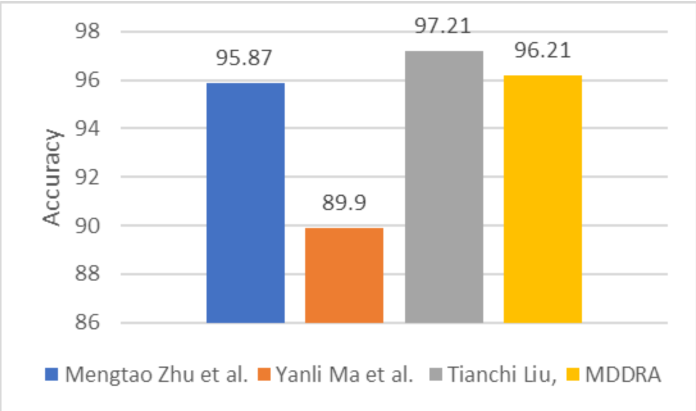

Figure 6-14: Shows the comparison of accuracy provided by MADDRA with the current state of the arts

6.7 Summary

A plethora of literature on the importance and urgency of driving risk mitigation techniques to prevent driving behaviour-related accidents. For a false proof robust alert system, the precise classification of driving behaviour is needed. However, to the best of our knowledge, the current works lack complexity, rigidness, synthesized dataset, are more focused on a particular side of perspective (vehicle, driver, or environment), false-positive classes, and low accuracy. This chapter aimed to provide a novel Multi-Class Driver Distraction Risk Assessment model that considers the vehicle, driver, and environmental data during a journey to categorize the driver on a risk matrix as safe, careless, or dangerous. The

MDDRA model offers flexibility in adjusting the parameters and weights to consider each event's specific severity level. Real-world data were collected using the Field Operation Test (TeleFOT), consisting of drivers using the same routes in the East Midlands, United Kingdom (UK). The results showed that it is possible to reduce road accidents caused by driver distraction. We also tested the correlation between distraction (driver, vehicle, and environment) and the classification severity based on a continuous distraction severity score.

Furthermore, we applied ML techniques to classify and predict driver distraction according to severity levels to aid the transition of control from the driver to the vehicle (vehicle takeover) when a situation is deemed risky. The experimental results obtained using various ML algorithms have shown improved results than the baseline and previous literature. The algorithm with the best performance was Ensemble Bagged Trees, which gave an accuracy of 96.2%.

However, this approach's limitation is that DL will produce better results regarding speed performance than an ML technique. The in-vehicle regression analysis result had a higher degree of correlation and was highly significant. The MDDRA model can be adjusted to fit any distraction risk assessment considering the driver, vehicle, and environmental contexts. In assigning weights to pedestrians on the road, we did not consider accidents or vehicles are hitting the pedestrians. However, the results of the regression show that vehicle distraction constitutes a higher level of significance. Factors such as sample size and data spread may have influenced the regression analysis's P-value results. Confidence intervals around the sample statistics would yield a better result than P-values alone. In addition, adopting CNN-DBN-LSTM techniques in detecting and classifying multi-class driver distraction would yield more effective and efficient results. Finally, considering the accuracy over time-complexity, the best ML model adopted is the Bagged Trees.

CHAPTER 7. A MULTI-CLASS CONTEXT-AWARE DRIVER DISTRACTION SEVERITY CLASSIFICATION USING AN HYBRID CNN-DBN-LSTM NETWORK

7.1 Synopsis

ADAS is a critical component in semi-autonomous vehicles and vital to the safety of vehicle drivers and public road transportation systems. In this chapter, presented is a hybrid DL technique that detects and classifies drivers' distractions using a multi-class Context-Aware drivers' distractions (event types-hand state, face orientation, eye glances) in combination with several context-awareness parameters: speed, weather, manoeuvre, surroundings, GPS position, accelerometer, and road type. Furthermore, a novel probabilistic DBN model based on the Fast-Recurrent CNN and LSTM network is developed to detect and classify driver's distraction into severity levels and use frame-based context data from the multi-view TeleFOT naturalistic driving study (NDS) data monitoring to classify the severity of driver distractions. The proposed methodology entails FRCNN trained to detect the driver's distraction, recurrent neural network layers LSTM trained to predict driver distraction severity from time-series data, and a probabilistic DBN calculates probability from probability with changing times and frames. This chapter entails multi-class distractions that, when combined with context-aware, leads to a severity level that can be further classified into safe, careless or dangerous driving. The model involves a Hidden Markov Driver Distraction Severity Model (HMDDSM) for transitioning the driver to the vehicle when a distraction level is reached. Validation of these results was performed using a cross-validation method applied to an unseen driver dataset

7.2 Background

Intelligent Transportation Systems is highly utilized to share drivers' behaviour and vehicle safety information such as collision warning, weather condition, accident occurrence, emergency brake light, and blind-spot warning [1]. In addition, vehicle information such as direction, speed, acceleration, signal intersections is also shared to prevent accidents. However, drivers do react to context-aware while driving. Thus, there is a need for real-time context-aware systems to prevent accidents.

According to (NHTSA) driver distraction is a crucial contribution to many road traffic accidents. National Highway Traffic Safety Administration (NHTSA) identified increased distraction from in-vehicle electronic devices and published guidelines to discourage excessive distraction by electronic devices in vehicles [2]. Furthermore, the vehicle user interface presents information overload to drivers, leading to distractions and causing accidents. Infotainment Systems are highly automated and requires a complex operation. Thus, diversion of visual attention of the driver away from observing his driving environment is crucial [3]. Driving is predominantly visual and manually by the hands (steering wheel and gear shift) and my feet (acceleration, braking). However, the driver inputs(eye gaze, hands) are often positioned in different states and sometimes perform tasks simultaneously [4]. Thus, a limitation is that the driver's input that constitutes distraction can have a different severity level.

It has been estimated that 94% of accidents result from drivers error, and about 75% is from drivers decision errors [5]. Furthermore, in research and survey conducted about the causes of road accidents, 55% were due to careless driving.

Critically, driver distractions could be influenced by in-vehicle components (In-vehicle devices), thus making the driver perform an act that leads to careless driving behaviour thereby, breaching driving laws. For example, infotainment system operation while driving could result in driver distraction. Drivers distractions detection is vital for many different applications in the domain of intelligent vehicles and autonomous driving.

Driving Context influences the behaviour and reaction of drivers. Also, context-aware changes affect the driver’s perceptions and risk levels. These challenges need a real-time context-aware system that can be applied to detect and learn driver’s behaviour in real-time. There is a need to define a context and the components of a context-aware application to implement such a system. In ADAS, capturing driver’s distraction in scenarios such as in-vehicle monitoring can be used to alert humans inside the vehicle when dangerous situations arise. Distraction is part of people’s everyday lives, and it reduces reaction time, concentration and alertness in a driving environment. Drivers distractions have led to ADAS development to improve driving safety and reduce accidents. Prevention of traffic accidents using ADAS can be categorised into driver monitoring or vehicle-oriented approach. According to Braunagel [6], stated ADAS system could aid the vehicle to take over in longitude and control situations which have led to our proposed systems for a severity model for drivers distractions to aid vehicle situation most especially in Semi-autonomous vehicles. Brauagel, further stated that the driver is responsible for the vehicle in semi-autonomous vehicles all the time. The driver's responsibility is transferred to an automated vehicle in some scenarios thus, enabling the driver to perform secondary tasks (reading, watching movies, sleeping) [6], [7]. Performing secondary tasks are still being regulated and not fully authorized, even in fully autonomous vehicles.

Furthermore, ADAS in autonomous vehicles has been designed to alert the driver when hands are not steering. This led to drivers taking over situations whereby the vehicle forces driver to take over driving tasks. Bruanagel resolved the above readiness of the driver in easing the transition of the driver taking over control without reducing the driver’s take-over readiness [6]. The approach used entail driver monitoring through features such as gaze guidance or increased decelerations.

Nevertheless, there can be scenarios where the vehicle needs to take over from the driver; this is probably due to the driver being distracted and not giving utmost concentration to driving activity; thus, having a degree of driving distraction according to severity level is crucial. A significant gap is the risk assessment of road accidents using severity prediction of traffic accidents with Recurrent Neural Network (RNN) [8]–[11]. The proposed prevention system rather than a detection approach leads to a system to prevent distraction that can lead to accidents. In this research, the utilization of secondary naturalistic driving study (NDS) data TeleFOT with 27 subjects and explore some of the TeleFOT data usages to determine the events in the TeleFOT data.

The main contributions of this chapter are:

- Proposed a frame-based severity metric of Drivers distractions using a linear transformation.
- Proposed architecture for a Multi-classification of drivers’ distractions into severity level using CNN and LSTM.
- Dynamic Bayesian Network model for forecasting and prediction of driver distraction.
- Integration of the MDDRA risk assessment model
- Hidden Markov Model Driver Distraction Severity Model (HMDDSM)
- Validation of frame-based severity model using cross-validation.

An approach towards a classification system of vehicles transitioning from driver to vehicle according to driver’s distraction severity level will be developed and tested on naturalistic driving study data. In addition, having such a system can be helpful in ADAS systems. This chapter focused on driver distraction monitoring using Context-aware drivers’ distraction and analysed with LSTM a Recurrent Neural Network (RNN). The proposed systems can be further applied in adaptation to driver's behaviour, leading to a preventive and corrective system for drivers' distractions based on the severity. Though, this can be subjective to the frequency and duration in which the event occurred. The proposed severity classifier considers the driver's distractions and considers context-aware information.

7.2.1 TeleFOT Data Analysis

TeleFOT NDS comprises the most significant European Field Operation Test (FOT) regarding the functionality of in-vehicle aftermarket and nomadic devices. The project's primary purpose was to improve Autonomous systems and cooperative systems in the Intelligent Transportation Systems environment [11], [23]. The FOT involved vehicle collecting and recording driving data such as speed measuring, vehicle dynamics and vehicle positions. This chapter considered the TeleFOT NDS study in the UK jurisdiction, launched in 2011 to collect naturalistic driver behaviour without any predefined condition in the United Kingdom. The test location was mainly in the East Midlands (Leicester, Coventry, Nottingham) area of the UK and partnership with Loughborough University [279]. The TeleFOT NDS study involved 27 participants(subjects), with some participants, repeated over different conditions. The trial type conditions are Baseline, Experienced and Novice.

7.2.2 Data Sampling Size

This driver sampling size will evaluate the developed algorithm using the driver not involved in the training.

Table 7-1:Data Sampling Size

TeleFOT PARTICIPANTS	**Baseline (BL), Experienced(E), Novice (N)**	**VIDEO LENGTH**	**IMAGE STATISTICS**	**DATA POINT (IMAGE STATISTICS X 4)**
001	BL001	01:13:00	105,109	420436
	E001	00:33:40	48,485	193940
033	BL033	01:10:55	106,398	425592
	E033	00:38:13	57,334	229336
	N033	00:18:20	27,512	110048
074	BL074	00:33:45	48,605	194420
	E074	00:44:41	64,360	257440

	N074	01:33:13	134,219	536876
081	BL081	00:33:43	48,562	194248
	E081	00:34:24	49,556	198224
	N081	01:39:59	106,534	426136
083	BL083	00:33:43	48,562	194248
	E083	00:35:26	51,039	204156
	N083	00:57:58	83,470	333880
088	BL088	00:35:00	50,405	201620
	E088	00:42:17	60,904	243616
	N088	01:29:05	128,271	513084
TOTAL			1,219,325	4,877,300

7.2.3 Context-Aware Threshold Mathematical Model for the Degree of Careless to Dangerous Driving

Two approaches could be implemented in the classification of the distraction, namely threshold detection and profile-based detection. The threshold detection involves tracking the duration of events and the number of occurrences of the specific distraction type during driving. In addition, if the duration and occurrence of distraction surpass a reasonable number as described in the justification section of metrics above, then a level of distraction is assumed based on the threshold. However, some distractions can instantly reach an optimum severity level; thus, there is a need for a detection and classification system. Another approach is a profile-based that characterises the driver's past behaviour and detects significant deviations from the expected safe driving profile of the driver.

Furthermore, a profile may consist of a set of parameters. Just a single parameter may not be sufficient to classify the driver distraction; thus, a multi-class distraction event detection and classification is needed—Matrix Table Metrics Weightings Threshold Severity Level as depicted in chapters 5 and 6. In section 7.2.4-6 are tables depicting the metric tables for the distractions considered.

7.2.4 Driver Context In-vehicle (Hand State, Eye's Gaze, and Face Orientation)

Table 7-2: Driver Context-Aware

Distraction Type	Distraction Type States	Threshold	Weight
Hand	Single Hand	Duration	1
	Double Hands	Duration	2
	No Hand	Duration	0
Face Orientation	Face Orientation On the road	Normal	1
	Face Orientation Off road	Duration of glance	2
Eyes Gaze	Eyes on Road	Normal	1
	Eyes off Road	Duration of glance off-road	2
	Eyes Shut	Duration of event	3

7.2.5 Environment Context-aware

Table 7-3: Environment Context-aware

Environment	Values	Thresholds
Road type	Urban	0 - 30
	Highway and Motorway	30 -70
	Dual carriageway	70 >
Weather	Day	Degree of brightness
	Night	dark
Manoeuvres	Stopped	Static
	turning	Speed towards a turning
Surrounding	Vehicle	Front/rear or vehicle
	Pedestrian	Front/rear and vehicle state.

7.2.6 Vehicle Context-aware

Table 7-4: Vehicle Context-aware

	Value	**Thresholds**	**Measures**	**Weight**
Speed	0-30mph	Road Type Urban	Vehicle state, Speed, Location tracking GPS, Positioning	Urban 30mph-
	30-70	Road Type Highway	speed limit and road type	Single carriage 60mph
	70	Dual Carriage	speed limit and road type	Dual carriage/motorway 70mph

7.3 Multi-Class Driver Distraction Risk Assessment (MDDRA)

This section will examine the perceived severity of naturalistic driving, thus showing varying levels of human-perceived severity. The driver's distraction does have a different impact that can be classified into safe, careless or dangerous. This is achieved by testing our hypothesis that driver's behaviour and driver's distraction having different severity levels by inferring from literature as seen in section A above. Then justification of weights, metrics for distractions present in the TeleFOT dataset.

The MDDRA involves a weighted average of the parameters to compute the severity levels per frame, as depicted in Table III.

1. These weights are capped by the maximum number a parameter can take.
   a. For example, taking "State of Hand" as a parameter, grade it as follows: (0 - double hands, 1- single hand, 2- no hands). If the value of a given frame for this parameter is x, then the weighted value is x/2 since the maximum value this parameter can take is 2.
2. Let us generalize this for any parameter x_i with maximum value m_i as follows:
   a. Severity Level => $\sum_{i=0}^{n} \frac{x_i}{m_i}$ where n is the number of parameters, we considered
3. Special considerations

a. Notification that speed should depend on the road type; hence, multiply it with the weight of its Road for speed. Consideration is given to road type in the UK, which conforms to the source of the dataset. For the metric of road types, we defined the threshold according to the speed limit allowable on the road type urban, single carriage and motorway with 30mph, 60mph and 70mph, respectively.

Furthermore, we defined the following context data as follows:

- Vehicle(V) and Driver data with probabilities P(V) = $\{v_1, v_2, \ldots, v_m\}$
- Environment data with probabilities P(E) = $\{e_1, e_2, \ldots, e_n\}$,
- Speed a
- Surrounding P(S)
- Pedestrians P(Pe)

The equation is formulated as follow:

1. The speed is computed as follows:
   a. Give the national speed limit of UK is 70mph, and the maximum road type score is 3

Average Speed = (Speed * Road Type) / (Max Speed * Max Road Type)

$$\alpha = \frac{speed * RoadType}{210} \tag{7.1}$$

1. There are different data points in each; thus, the severity level of a given frame with k data points is below.

$$S(fi) = \frac{1}{\mathrm{k}} * \left(\sum_{i=0}^{m} \mathrm{P(Vi)} + \sum_{j=0}^{n} P(Ei) + a + P(S) + P(Pe) \right) \tag{7.2}$$

2. Now compute the aggregate severity (S*) of a given frame given the last $i - t$ frames.

   This is achieved by taking the average of the current frame's severity score than the severity score of last $i - t$ frames

$$\mathrm{S}^{*}(fi) = \frac{1}{\mathrm{t}}((S(fi) + \sum_{j=t}^{i-1}(S(fj)) \tag{7.3}$$

Table 7-5: Parameters & Weightings

#	Parameter	Maximum Weight	Distraction State Type	Weight
1	State of Hand	2	Double hands	0
			Single hand	1
			No hands	2
2	Road Type	3	Urban	1
			Dual	2
			Highway	3
3	Face Orientation	2	On road	1
			Off road	2
4	Illumination	1	Day	1
			Night	2
5	Eye Gaze	2	Eyes on road	0
			Eyes off-road	1
			Eyes shut	2
6	Weather	3	Dry	1
			Rain	2
			Snow	3
7	Manoeuvre	2	Stopped	0
			Turning	1
			Moving	2
8	Surroundings	2	Vehicle not present	0
			Vehicle present	1
9	Pedestrians	2	Pedestrian not present	0
			Pedestrian's present	1
10	Speed	Urban 30mph- Single carriage 60mph Dual carriage/motorway 70mph	(Speed * Road Type)/300	Urban Single carriage Dual Carriage Highway

7.4 Distraction Detection and Methodologies

Here presented a detailed description of our detection and classification approach, which entail a hybrid algorithm CNN-DBN-LSTM respectively. The classification driving distractions from images entails combining a pre-trained image classification model with an

LSTM network. The approach is a sequence-to-sequence of images labelled into the class of the distraction type. We manually extract the features using a single label approach detection a single distraction type in the CNN-LSTM method. In contrast, we used a multi-label approach to extract features for each frame in the fuzzy logic section. We train an LSTM for the prediction and classification into severity levels. The approach of the methodology is depicted as follows. In addition, the DBN will be used in the prediction of the severity of distractions.

7.4.1 Data Pre-processing

Many software tools were used for data pre-processing. These tools aid in generating video data, conversions, thresholding, splitting the video data and conversions to formats that could be processed. The tools used includes:

7.4.1.1 Race Technology Software

TeleFOT driving data videos was generated via the Race Technology Software, which included times-series data of the vehicle. This aids in knowing the vehicle data at every frame.

7.4.1.2 Matlab Tools

Colour Threshold: this tool is used in segmenting image pixels based on colours to make it easier to analyse the image. The tool converts a given image into a binary image that the algorithm can then handle.

Image Acquisition: this tool helps in retrieving images to be analysed from the source for further processing. The tool has advanced capabilities to extract image frames from a video stream, to facilitate processing.

Image Batch Processor: to facilitate quick batch processing of images, the image batch processor was used. The tool facilitates the processing of images from a folder, thus speeding up the process.

Image Labeller: the labeller tool facilitates marking rectangular regions of interest on images with scene labels, pixel ROI labels, and polyline ROI labels. This tool was applied in the labelling of the region of interest (RoI).

Image Region Analyser: this tool was used to measure various properties of an image and display tabulated information and create other binary images by filtering regions of interest.

Feature extraction: there is noise removal involved to increase the accuracy of eye gaze or glances detection.

7.4.1.3 ETL data processing

Image Segmenter: the tool was used to segment the images and create segmentation masks using automatic, semi-automatic and manual algorithms. The tool aided in segmenting the labelled ROI on images as depicted in 7.1 and 7.2, respectively. Each TeleFOT Image is 1280x720 with a width of 1280 pixels and a height of 720 pixels as depicted in 1a and 1b, which illustrates a sample of participant BL001, an enhanced image.

Image Splitting: The image was split into four frames using MATLAB representing In-vehicle (frontal view, side view) and Outer-vehicle (front and rear view). The input images of the CNN and contains the raw pixel values of the images. The local receptive fields (LRF) comprise four inputs relative to the TeleFOT datasets; the drivers view in the dataset is depicted as follows: Inside Frontal View (IFV), Inside Side View (ISV), Outside Rear View (ORV) and Frontal View (OFV). A mathematical representation of the LRF is as follows:

Views (V) = { IFV, ISV, ORV, OFV }

LRF $\subseteq$ { IFV, ISV, ORV, OFV }

LRF $\subseteq$ P(V) - $\emptyset$ Possible outcomes is $2^n - 1$ where n denotes the number of views and the possible outcome is 15 states.

Figure 7-1: Image Enhancement:

The driver point was determined based on a significant pre-defined point around the driver body region. The region around the head will be used in indicating where the driver is. This enables us to perform further image segmentation to improve the accuracy of a head detection algorithm that ensures human recognition. We further classify the images into some of the distraction events that our algorithm will detect.

Figure 7-2: Image Enhancement Single Hand-on wheel vs. Double Hands-on wheel

7.4.2 Autonomous Vehicle Monitoring Sample Image in the In-vehicle and Out-vehicle

Figure 7.3 shows the vehicle monitoring image samples from the In-vehicle and the view of the Out-vehicle (Figure 7.4).

Figure 7-3: In-Vehicle

Figure 7-4: Outer-Vehicle

7.5 Driver Feature Semantic Segmentation

Figure 7.5 depicts the image enhancement of the sample image and feature extraction alongside the driver's state plot.

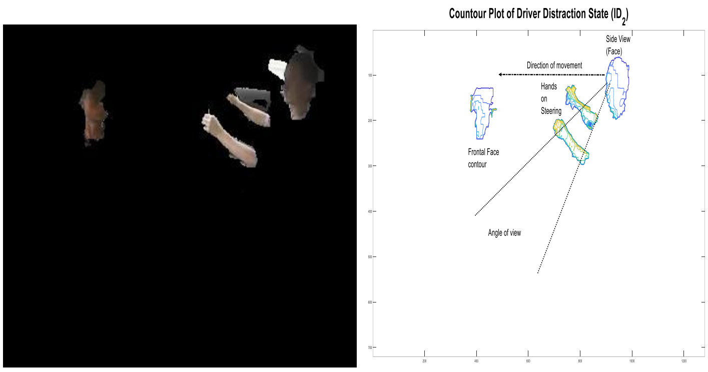

Figure 7-5: Image Segmentation Contour Plot of Driver state

Image Viewer: as the name suggests, the tool finds usage in visualizing the images.

FFMPEG: this is open-source with many libraries for handling various multimedia streams and files. For this project, FFMPEG was used to convert video to images at a rate of 25 fps.

VIDEOPROC: this is a video editing software suite that enables video cutting, cropping, merging, rotating and compressing. The tool was used in the splitting of video into equal lengths for training. In addition, the tool was used in the enhancement of the driver's video quality into a 4K resolution since the driving videos are dated.

IMAGEJ: Image Sequence is a Java-based image processing software. This tool was used to convert the sequence of frames to video at the rate of 25 fps in this project. Afterwards, the converted images were saved into AVI format.

7.5.1.1 Model Preparations and Transfer lLarning

Several pre-trained network models can be adopted as the CNN Architectures, multi-layer neural networks designed to recognise visual patterns from images. ConvNets have millions of parameters and a lot of hidden layers. For this, Resnet18 and Resnet50 are out-of-the-box types of CNN classifiers.

7.6 CNN-LSTM-DBN Process Flow

A CNN LSTM architecture uses a CNN for feature extraction on a given input data set, combined with a Long Short-Term Memory Network, which supports sequence prediction. By design, a CNN LSTM was intended for handling visual time series prediction and generation of textual description from a given input sequence of images and video frames. CNN LSTM handles activity recognition by generating textual descriptions of activities identified in sequences of images. The process flow for a CNN LSTM network comprises ten stages, each represented by a box in figure 7.6 below. The process flow starts with autonomous vehicle monitoring and ends at the ML classifier; in between, two process flows make up the internal operations of the model. The two flows represent the in-vehicle and out-vehicle flows. Each of the process flows are as follows

7.6.1 In-Vehicle Video Capture

The first stage of this process flow involves collecting the vehicle's interior data with a video from the onboard monitor.

7.6.2 Semantic Segmentation

The second stage of the in-vehicle process flow entails semantic segmentation of the driver’s features. This component extracts driver features such as driver activity, number of hands on the wheel, and face orientation off the road. Extracted driver features are then fed into a hybrid CNN-LSTM model as in-vehicle packets.

7.6.3 Hybrid CNN-LSTM

The in- vehicle packets from the semantic segmentation section are fed to the hybrid CNN-LSTM model. The CNN layer performs feature extraction on input data, while the LSTM performs sequence prediction and activity recognition. The model is tasked with identifying the type of distraction a driver experiences. Any identified activity is compared with historical data to recognise the distraction type and give it a distraction identifier fed to the dynamic Bayesian network. In essence, the hybrid CNN-LSTM performs driver distraction recognition by analysing the extracted driver features, where fuzzy sets for classification of the distraction by severity level are extracted. The fuzzy sets of distractions inform the model of the distraction ID, which is then fed to the Dynamic Bayesian model together with driver features, extracted by the deferential stage of the model.

7.6.4 Frame Differencing

A copy of the video stream from the semantic segmentation section is fed to a frame differencing component, which identifies and extracts the driver's feature of hands on the steering wheel. Out-vehicle environment monitoring process flow: the second process flow making up the entire CNN-LSTM Process Flow is the external

7.6.5 Out-vehicle/Environment monitors

This component collects data about the vehicle and the vehicle's external environment, such as speed, manoeuvres, and a video recording of the road and pedestrians. Two streams of data are obtained here, external video and vehicle data which includes speed and manoeuvres.

7.6.6 Faster R-CNN

The video recorded from the outside is fed to a faster R-CNN, which analyses the frames to extract information related to road type, weather and identify pedestrians and the surrounding. Faster R-CNN detects regions that have objects of interest.

7.6.7 Differential

The differencing component receives data about driving speed**,** road type, weather and driving manoeuvres. The component relates the different variables; speed, road type and manoeuvres to generate critical context-aware.

7.6.8 Dynamic Bayesian Model

The dynamic Bayesian model takes in three variables; distraction ID and driver features from the in-vehicle monitoring stream and context-aware from the out-vehicle streams. With the three key inputs, the Dynamic Bayesian model performs severity classification by relating the variables to each other over adjacent time steps, outputting probabilistic data, which forms the basis of operations of the ML classifier.

7.6.9 ML Classifier

The last component of the model is the ML classifier, which takes in the probabilistic data, and performs prediction of the class of given data points, resulting in distraction classification. For this case, the classifier performs severity classification; given the outputs of the dynamic Bayesian network model, the classier approximates and maps the level of distraction on a severity scale. The severity of distraction acts as the basis of whether the system takes over the vehicle's operations or not.

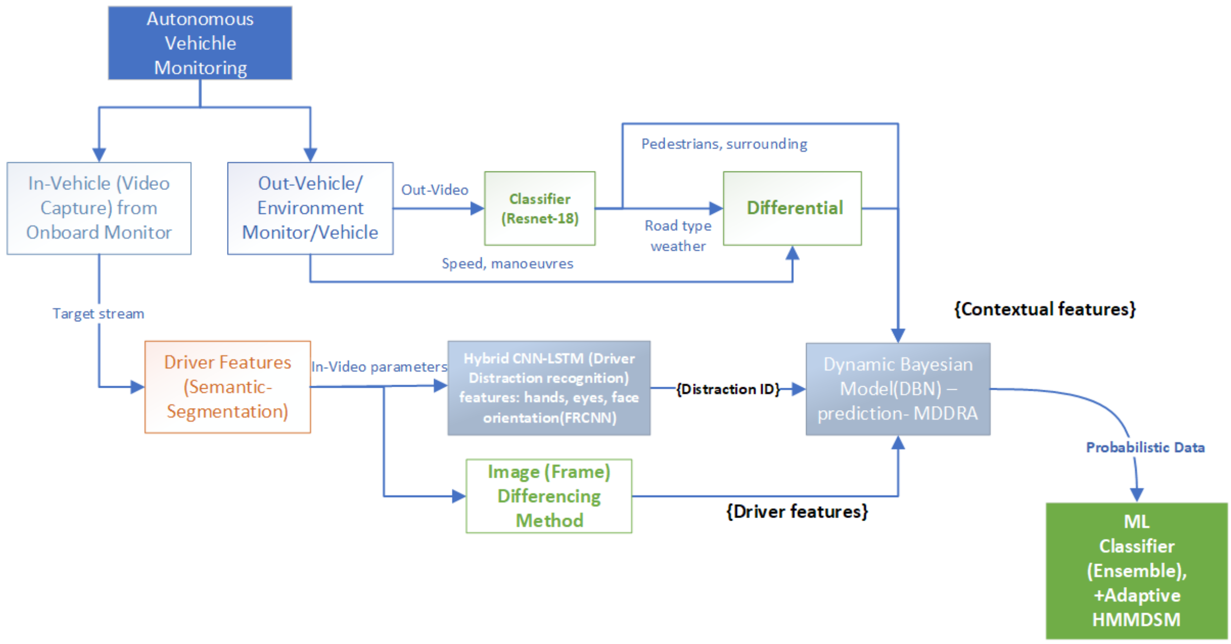

Figure 7-6: CNN-LSTM-DBN Process Flow

Figure 7.7 detailed the proposed process flow chart for in-vehicle (Figure 7.7a) and out vehicle (Figure 7.7b).

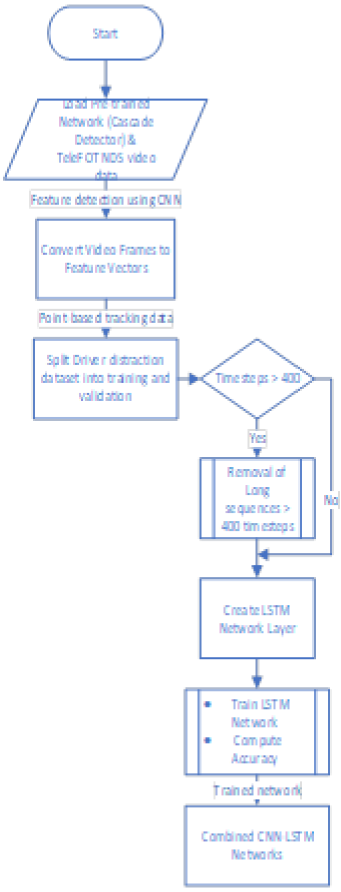

Figure 7-7a: In-vehicle

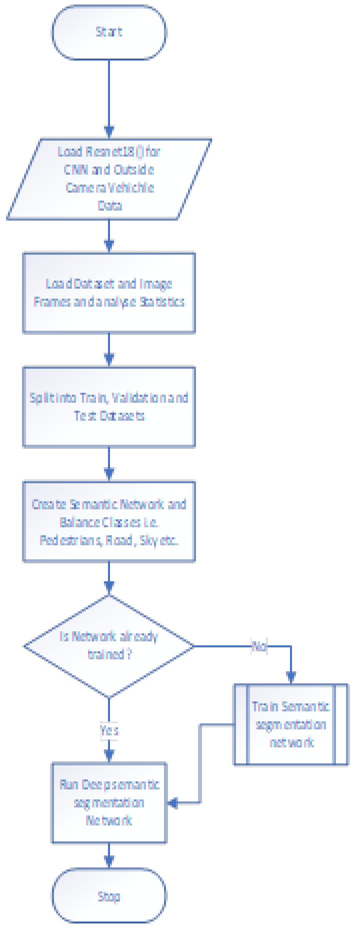

Figure 7-8b: Outer-vehicle (left)

7.7 Dynamic Bayesian Network (Data) Model

The dynamic Bayesian Network Model is made up of seven components. The first component is a Fast R-CNN, designed to handle context-awareness and physiological data. The component receives a video stream from the outside of the vehicle, analyses it to detect objects in the video stream, such as pedestrians, neighbouring vehicles and any other external object around the moving vehicle. By design, the Fast R-CNN uses two networks for object detection; a region proposal network (RPN), which generates region proposals, and the second network in RPN regions to detect objects.

7.7.1 Bounding Boxes

The second component of the architecture employs the bounding boxes approach; each detected object is annotated using bounding boxes annotators. These annotations are placed around key driver features such as the driver's face, eyes, hands, external objects such as pedestrians, obstacles and vehicles on the road.

7.7.2 Frame-by-Frame Face and Eye's Tracking

having established objects of interest and outlined them using bounding boxes, this component keeps track of frame-by-frame changes in the state of the detected and boxed objects. Any change is tracked and fed to the dynamic Bayesian network model.

7.7.3 CNN-LSTM

This system component analyses video streams from a vehicle's interior and exterior to identify distractions experienced by a driver. The CNN layer performs feature extraction on input data, while the LSTM enables sequence prediction. The CNN-LSTM is, therefore, able to perform activity recognition which is critical for this architecture. In this case, CNN-LSTM is employed to analyse context data and identify driver distraction and send the type of distraction to the dynamic Bayesian network model in a distraction ID. In essence, this component informs the DBN of the kind of distraction that a driver is experiencing at any given time.

7.7.4 Dynamic Bayesian Network Model

Information about identified objects that have already been annotated using bounding boxes and tracked changes are fed to this module, the dynamic Bayesian network model. Dynamic Bayesian model relates the variables to each other over adjacent time steps, often called a two-time-slice BN. For this case, the DBN takes in three variables; the identified objects with their bounding boxes tracked changes of the objects in the form of frame-by-frame object tracking information and a distraction identifier from the CNN-LSTM module. The complete architecture of the DBN model is given below in Figure 7.8.

7.7.5 ML Classifier

The last step in this architecture entails distraction classification. In classification, prediction of the class of given data points is performed. For this case, the classifier performs severity classification; given the outputs of the dynamic Bayesian network model, the classier approximates and maps the level of distraction on a severity scale. Depending on how severe the distraction is, the system may take corrective measures, including vehicle takeover.

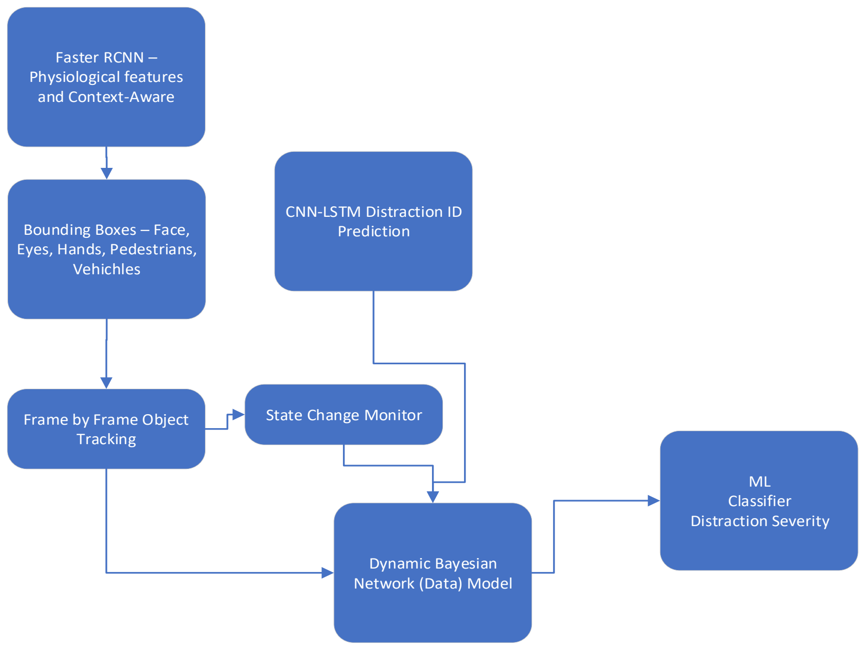

Figure 7-9: Dynamic Bayesian Network (Data) Model

7.8 Context-Aware Architecture

CNN LSTM is built as a 3-tier architecture system, consisting of a Sensing layer, a logic layer and an application layer. The proposed 3-tier architecture system can be seen in Figure 7.9.

7.8.1 The sensing layers

Being the lowest layer handles data collection from in-vehicle and out-vehicle video capture and sensors components. In-Vehicle data is collected using a dashboard camera, which video streams to a CNN and LSTM layer that handles driver features extraction. On the other hand, the out-vehicle data is collected from various sensors and video recorders, including accelerometer, speed gauge, weather data, maps and video recording. The out-vehicle data forms the context data, while the in-vehicle constitutes the driver features.

7.8.2 Logic Layer

The context and driver feature data streams are fed to the Dynamic Bayesian model in the logic layer of the architecture. The Bayesian model takes historical data and, with several computation steps, performs severity classification, outputting probabilistic data, which forms the basis of operations of the ML classifier and for fuzzy regression inference.

7.8.3 Application Layer

The ensemble and the validator feed the application layer components, making the prudent decision on vehicle takeover. The regression fuzzy, on the other hand, feeds the application layer with the severity classification probability.

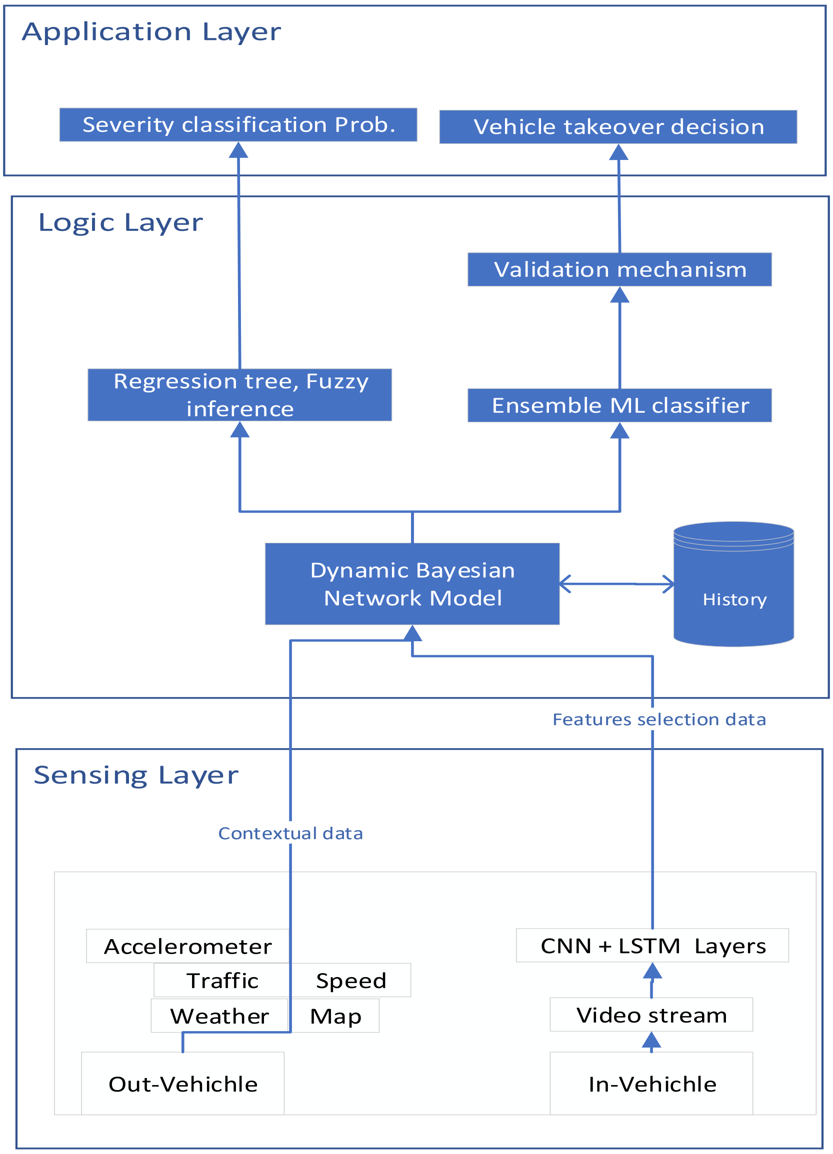

Figure 7-10: 3-tier Context-aware architecture for driver distraction

7.9 Novel DL-Based Driver Distraction Severity Model

The severity threshold will be deduced using metrics such as time, frequency and behaviour. The drivers' behaviour can be profiled based on previous data, and prediction of drivers' behaviour can be inferred from the previous. However, there are uncertainties in driver's behaviour which is hard to predict, such as cognitive distractions. Thus, we adopted a proactive approach in forecasting the distraction severity level.

Video-Based CNN-LSTM Based driving distraction detection and classification

The conversion of the sequence of images of the driver's distractions into a single video at the streaming rate of 25 frames per seconds video can be considered a generalization of the image data in which a temporal component is inherent to a sequence of images. We generalised the 2-dimensional spatial convolutions into 3-dimensional Spatio-temporal convolutions; thus, each frame in the video can be considered an image, and one, therefore, receives a sequence of images in time. Each of the image sizes is 224 x 224 x 3, and a total of no. Frames received. The size of the video segment is 224 x 224 x frame no x 3. The 3-dimensional capture enables us to capture the colour channel since the dataset is now a video. The sequential data set (e.g., text) requires 1- dimensional convolutions, an image data set requires 2-dimensional convolutions, and a video data set requires 3-dimensional convolutions. However, 3-dimensional convolutions add only a limited amount to what one can achieve by averaging the classification of individual frames by image classifiers.

Furthermore, motion adds a tiny amount to information available in the individual frame for classification purposes. Finally, the 3-dimensional CNN is suitable for relatively shorter video segments (half seconds) but might not be suitable for longer videos. In longer videos, a better approach combines recurrent neural networks (LSTM) with CNNs. For example, the 2-dimensional convolutional over individual frames, but a recurrent network carries over states from one frame to the next. Another approach is adopting 3-dimensional CNNs over short segments of video and linking them up with recurrent units. Thus, this helps in identifying actions over longer time horizons.

In addition, the use of LSTM has been an idea in the case of storing information from previous values and exploit the time dependencies between the samples. We compute the current frame severity considering the severity in the previous frame, which is the combination of CNN and LSTM. The Dataset is a naturalistic driving study dataset collected over successive periods characterized as a Time Series. The model developed is based on CNN-LSTM, which detects and passes the previous hidden state to the next step of the sequence. Lastly, we applied DBN for the prediction of distraction [280].

ALGORITHM 7.1: Multi-class Distraction Severity Classification and Takeover System

Begin

Input: TeleFOT Data (Driving data) $V \leftarrow IV, OV\}$

// In and out video streams from data acquisition devices.

Output: SC (severity classification)

For $\forall$ i in nFrames $\in V$ **do**

$Frame\_i \leftarrow IV(i)$

$Frame\_j \leftarrow OV(j)$

$[Frame\_i, Frame\_j] \leftarrow Preprocessing\_method\ (Frame\_i, Frame\_j)$

// Image Split, Segmentation, enhancement,

//noise removal, ROIs

$IV(i) \leftarrow Frame\_i$

$OV(j) \leftarrow Frame\_j$

end for

$[O_{Eye}, O_{Face}] \leftarrow \text{trackerKLT}(IV)$

// detect and track driver, distraction type eyes and face orientation.

$D_{ID} \leftarrow \text{FastRCNN\_LSTM}(IV)$

$X_t^{in} \leftarrow [D_{ID}, O_{Eye}, O_{Face}]$

$[O_{surr}, O_{Ped}, O_{road}, , \ldots] \leftarrow ResNet18(OV)$

$X_t^{env} \leftarrow [O_{surr}, O_{Ped}, O_{road}, , \ldots]$

$X_t \leftarrow \{X_t^{in}, X_t^{env}, X_t^{speed}\}$

// the feature vector for distraction severity

$y_t \leftarrow MDDRA(X_t)$

// using a novel risk assessment method, compute the severity score (class)

$y_t^p \leftarrow \sum_{j=1}^{p} W_{_j}\, y_{t-j} + \alpha + \nabla t + BX_t + \epsilon_t$. // Probabilistic (Data) Model

$y \leftarrow \underset{c_j \in C}{\text{argmax}} \sum_{h_i \in H} P(c_j|h_i).P(T|h_i).P(h_i)$

// Apply trained ensemble Machine learning (Bayes) classifier

```
    // Training a Hidden Markov Driver distraction model using the Baum-Welch and
     Viterbi algorithm to analyse Takeover's decision-making.
End
```

7.9.1 Face and Eye’s Point-tracking using Kanade Lucas Tomasi (KLT)

The point tracker object in our framework tracks a set of points (x-y coordinates) using the Kanade-Lucas-Tomasi (KLT) feature-tracking algorithm. Then, apply the point tracker for face and eye tracking, video stabilization, and camera motion estimation. It works well for tracking faces and eye’s that do not change shape and those that exhibit a specific visual texture (Figure 7.10). The point tracker is often used for short-term tracking as part of a larger tracking framework.

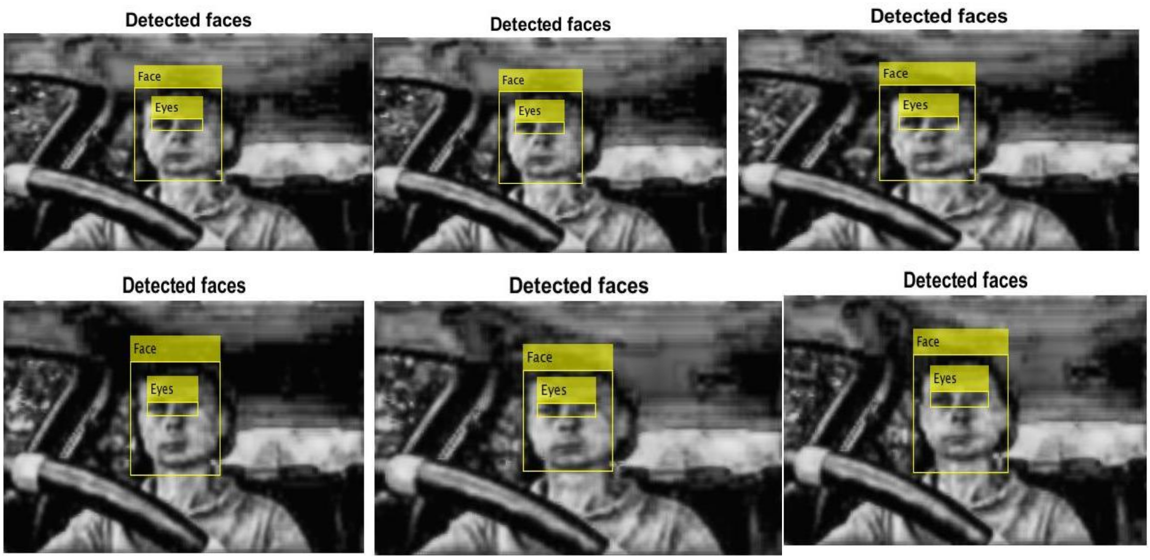

Figure 7-11: Face and eyepoint tracking

7.9.2 Semantic segmentation for Context-awareness

A semantic segmentation network for context objects classifies every pixel in a distraction frame from OV, which is segmented by class. An essential application for semantic segmentation is road segmentation for autonomous driving; we used the CamVid dataset [2] from the University of Cambridge and distracted driver video frames captured from OV for training. This dataset is a collection of images containing street-level views obtained while driving and from TeleFOT datasets. The complete dataset provides pixel-level labels for 32 semantic classes, including car, pedestrian, and road. This creates the Deeplab v3+ network with weights initialized from a pre-trained Resnet-18 network. ResNet-18 is an efficient network that is well suited for applications with limited processing resources. The Labelled pixels for the training of the Semantic segmentation network can be seen in Figure 7.11.

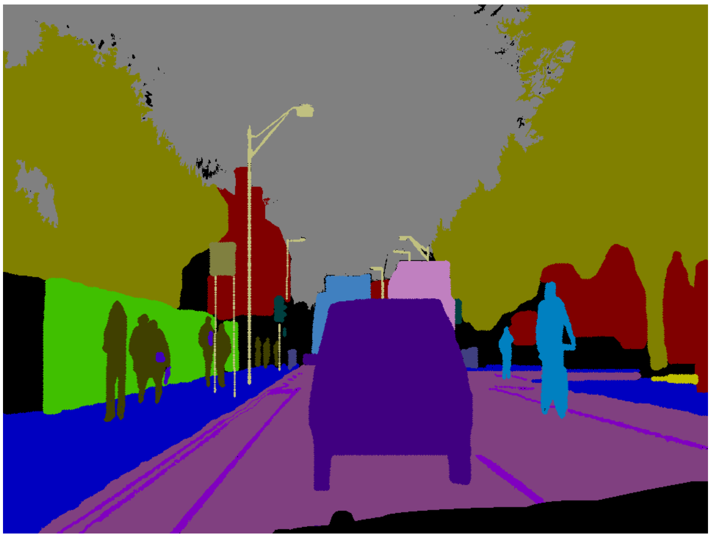

Figure 7-12: Labelled pixels for the training of Semantic segmentation network

7.9.3 Training of Semantic segmentation network

The training of Semantic segmentation network, i.e., Fast R-CNN was done, the accuracy graph (Figure 7.12), loss graph (Figure 7.13) can be seen below. While Figure 7.14 shows the frequency of occurrences of environment detections in the training set.

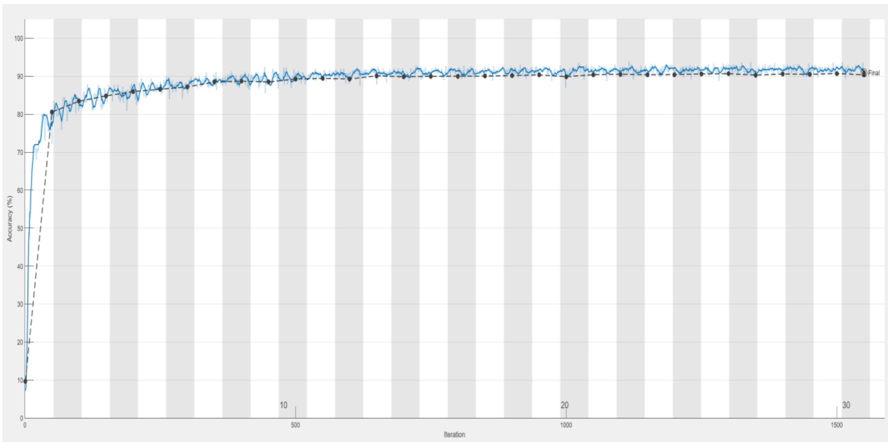

Figure 7-13: Fast R-CNN training model

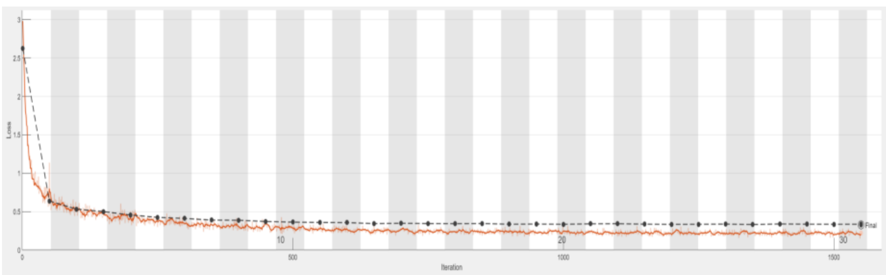

Figure 7-14: Fast R-CNN loss function during training

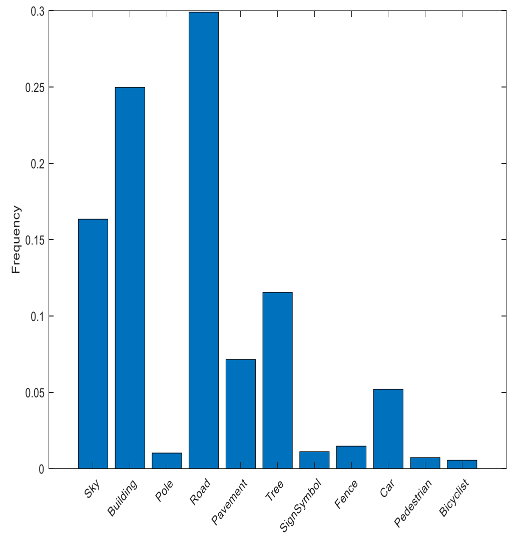

Figure 7-15: frequency of occurrences of environment detections in the training set

Table 7-6: Validation Results

GlobalAccuracy	MeanAccuracy	MeanIoU	WeightedIoU	MeanBFScore
0.89416	0.86257	0.66641	0.83109	0.69922

Table 7-7: Fast R-CNN classification accuracy

	Accuracy	**The intersection of union (IoU)**	**MeanBEScore**
Sky	0.93824	0.90821	0.90645
Building	0.81932	0.79479	0.64927
Pole	0.76297	0.24525	0.58579
Road	0.94568	0.93048	0.81708
Pavement	0.89163	0.74731	0.76321
Tree	0.88847	0.77611	0.73491
SignSymbol	0.76303	0.42155	0.53403
Fence	0.81325	0.58934	0.58114
Car	0.92007	0.79514	0.75216
Pedestrian	0.85778	0.47054	0.63498
Bicyclist	0.88784	0.65182	0.60337

7.9.3.1 Blob Processing: State of Hands on the Steering Wheel

The image below (Figure 7.16) provides an overview of blob processing for the state of hands.

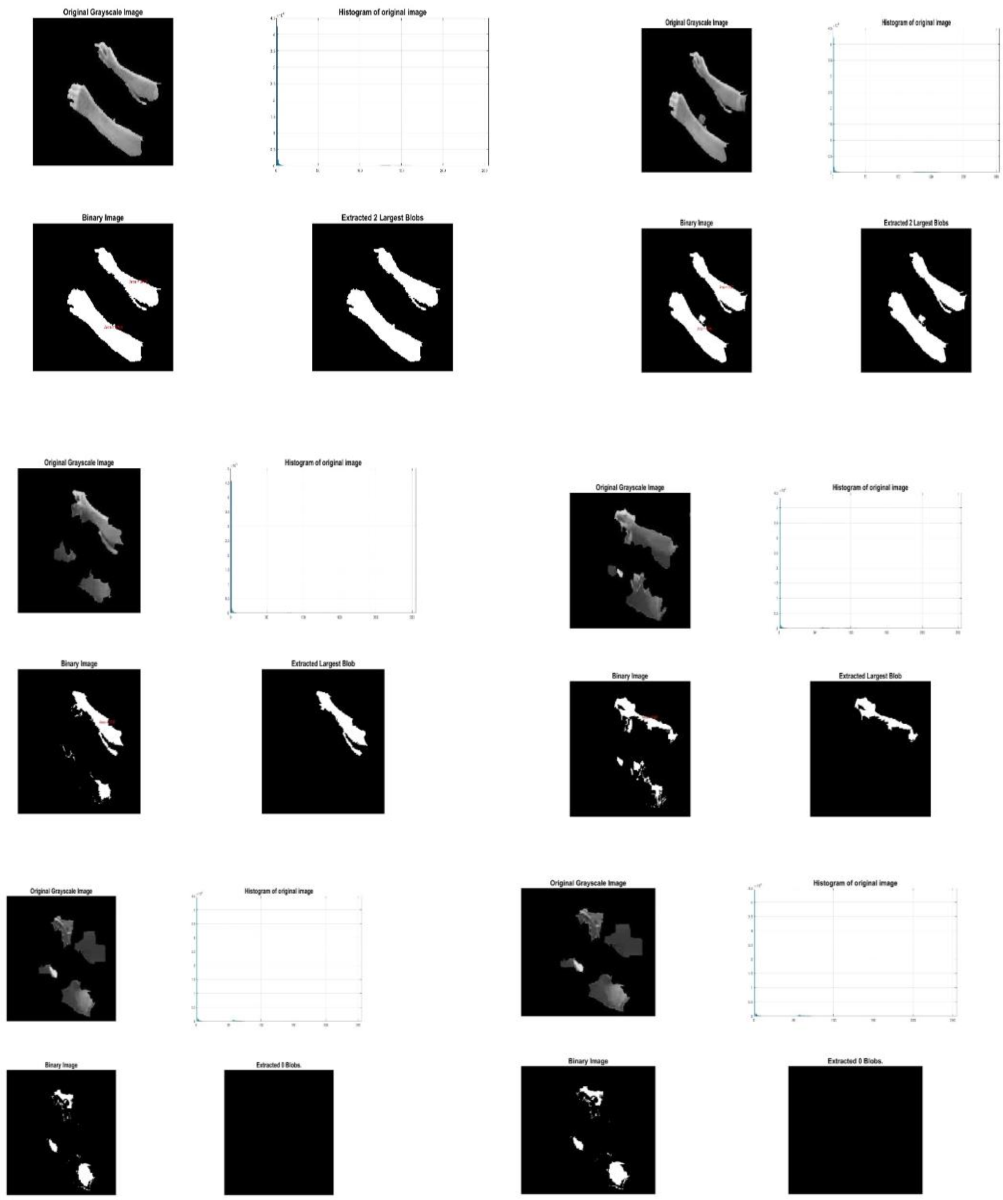

Figure 7-16: blob processing state of hands

7.9.3.2 CNN-LSTM-DBN Classifier

The dynamic Bayesian Network classifier is made up of seven components.

The first component is a Faster RCNN, designed to handle context-awareness and physiological data. The component receives a video stream from the outside of the vehicle, analyses it to detect objects in the video stream, such as pedestrians, neighbouring vehicles and any other external object around the moving vehicle. By design, the Faster R-CNN uses two networks for object detection; a region proposal network (RPN) that generates region proposals and a second network that takes in the regions from RPN to detect objects.

Bounding Boxes: The second component of the architecture employs the bounding boxes approach; each detected object is annotated using bounding boxes annotators. These annotations are placed around key driver features such as the driver's face, eyes, hands, external objects such as pedestrians, obstacles and vehicles on the road.

Frame-by-frame face and eye's tracking: having established objects of interest and outlined them using bounding boxes, this component now keeps track of frame-by-frame changes in the state of the detected and boxed objects. Any change is tracked and fed to the dynamic Bayesian network model.

CNN-LSTM: this system component analyses video streams from the interior and exterior to identify distractions experienced by a driver. The CNN layer performs feature extraction on input data by design, while the LSTM enables sequence prediction. The CNN-LSTM is, therefore, able to perform activity recognition which is critical for this architecture. In this case, CNN-LSTM is employed to analyse context data and identify driver distraction and send the type of distraction to the dynamic Bayesian network model in a distraction ID. In essence, this component informs the DBN of the kind of distraction that a driver is experiencing at any given time.

Dynamic Bayesian Network model: information about identified objects that have already been annotated using bounding boxes, as well as tracked changes are fed to this module; the dynamic Bayesian network model. Dynamic Bayesian model relates the variables to each other over adjacent time steps; often called a two-time-slice BN. For this case, the DBN takes in three variables; the identified objects with their bounding boxes tracked changes in the form of frame-by-frame object tracking information and a distraction identifier that comes from the CNN-LSTM module. Figure 7.17 provides a Dynamic Bayesian Network (Data) Classifier.

ML Classifier: the last step in this architecture entails distraction classification. In classification, prediction of the class of given data points is performed. For this case, the classifier performs severity classification; given the outputs of the dynamic Bayesian network model, the classier approximates and maps the level of distraction on a severity scale. Depending on how severe the distraction is, the system may take corrective measures, including vehicle takeover.

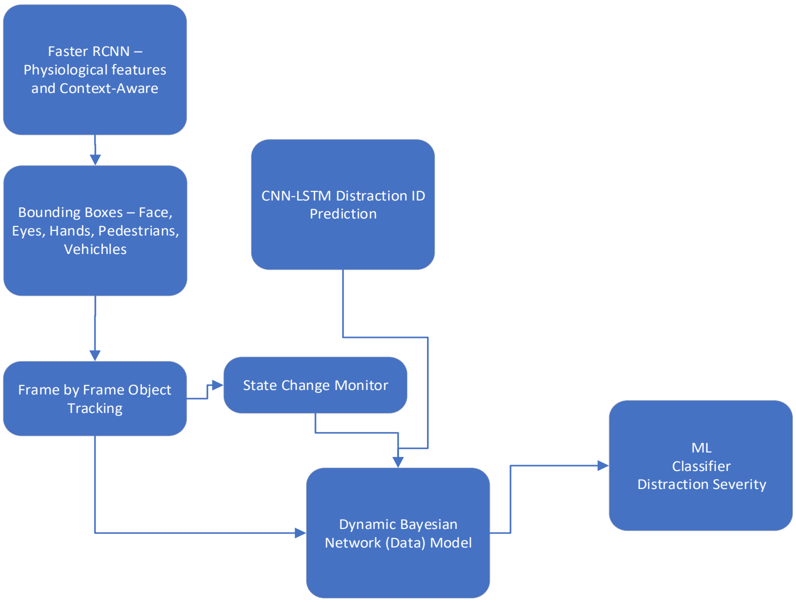

Figure 7-17: Dynamic Bayesian Network (Data) Classifier

7.9.3.3 HMDDSM: A Model for Decision-Making in Vehicle Transitioning

The architecture shown below in Fig 7.18 is for individual driving, applying the ML learning classification of severity. The model (Fig. 7.19 and Fig.7.20 adapt to the driver's severity level of the driver's distraction behaviour. However, if the drivers are not distracted frequently, the vehicle takes no-decision. We can now compute for a driver the decision for transitioning as seen in the model below.

The following components and assumptions specify the Hidden Markov Driver Distraction Severity Model (HMDDSM):

$Q = \{q1_, q2\}$ *a set of* $N = 2$ *driver distraction severity states – High and Low*

$A = a11 \ldots aij \ldots aNN$ a transition probability matrix A each aij representing the probability of moving from severity state i to state j, s.t. $\sum_{j=1}^{2} a_{\mathrm{i},\mathrm{j}} = 1\ \forall i$.

$O = o1o2 \ldots oT$ is the sequence of T observations (distinct) each one drawn from a timed vocabulary $V = v1, v2, \ldots,$

$B = bi\ (ot)$ is a sequence of observation likelihoods, also called emission probabilities, each expressing the probability of an observation ot being generated from a severity classification state i from symbol set $\{k1, k2\}$.

$\pi = \{\pi 1, \pi 2\}$ an initial probability distribution over driver distraction severity states. πi is the probability that the Markov chain will start in state i. Some states j can have $\pi\ j = 0$. This means that they cannot be initial severity states. Also, $\sum_{i=1}^{2} \pi i = 1$.

- The probability of a particular driver severity state depends only on the previous severity state as with a first-order Markov chain: $P(qi|q1 \ldots qi-1) = P(qi|qi-1)$.

- The probability of an output observation oi depends only on the driver severity state that produced the observation qi and not on any other severity states or any other observations: $P\ (oi|q1 \ldots qi, \ldots, qT, o1, \ldots, oi, \ldots, oT) = P(oi|qi)$.

We solve the HMDDSM problem for vehicle transition in three steps:

- HMDDSM Learning using the Baum-Welch algorithm: Given an observation sequence O for the driver and the set of possible severity states in the HMM, learn the HMDDSM parameters A and B.
- Computing Likelihood for vehicle transition: Given an HMDDSM $\lambda = (A, B)$ and an observation sequence O, determine the likelihood $P\ (O|\lambda)$.

- Decoding the driver severity states using Viterbi algorithm: Given as input an HMDDSM $\lambda = (A, B)$ and a sequence of observations $O = o_1, o_2, \ldots, o_T$, find the most probable sequence of severity states $Q = q_1, q_2, \ldots, q_T$ To support decision making for the switch from driver to ADAS.

Tables 7.8 and 7.9 below show the transmission and emission probabilities for the hidden Markov driver distraction severity model of driver 001 in the TeleFOT data)

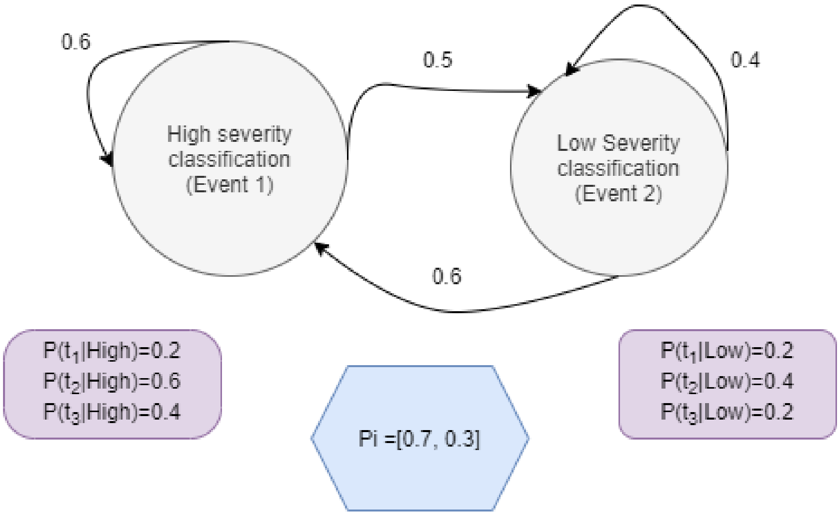

Figure 7-18: Hidden Markov Driver distraction model for Decision-making in-vehicle transitioning (between driver and ADAS).

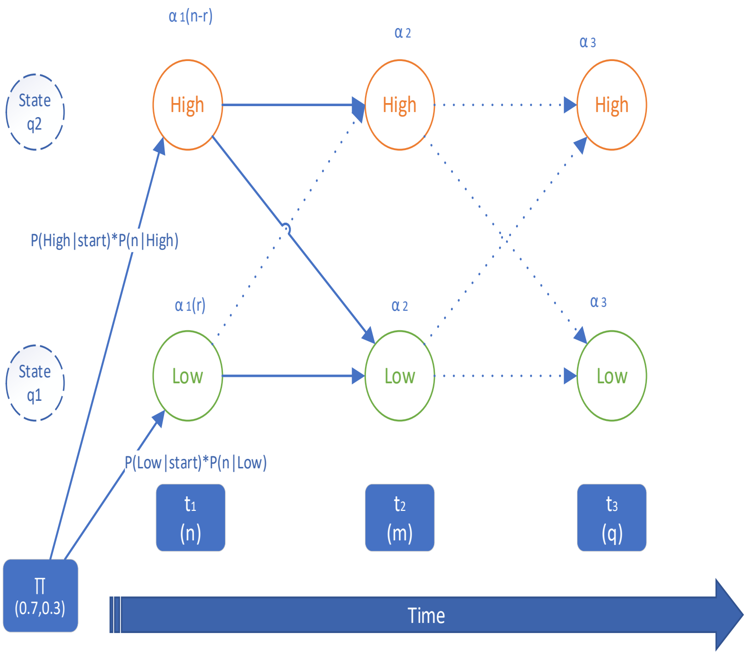

Figure 7-19: Observation likelihood for distraction severity events in vehicle transitioning between driver and semi-autonomous vehicle

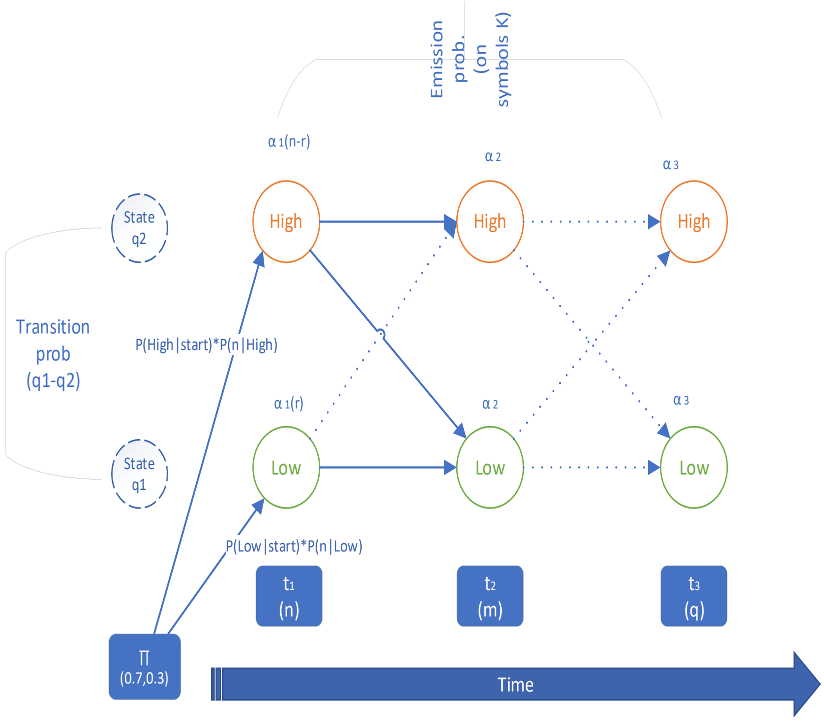

Figure 7-20: Estimating the transmission and emission probabilities (HMM)

Table 7-8: Estimated transmission (TR) probabilities for distraction severity of Driver 001

0.9767	0.0233
0.1538	0.8462

Table 7-9: Estimated emission (E) probabilities for distraction severity of Driver 001

0.1379	0.1609	0.1839	0.1724	0.2299	0.1149
0.0769	0.1538	0.0769	0.0769	0.2308	0.3846

7.10 Summary

A hidden Markov model has been adopted to decide how much dangerous driving is detected and more petite than careless driving. Thus, if less towards careless then driver continues. If more dangerous overall, then autonomous vehicles take over. The internal mechanism, time observation data based on how long or time-series. How long of the probabilities at the same time. The decision making ensures that the one that is more advantageous the car leans on. If the after running the frame-data and if severity less than careless then the driving continues and if dangerous the transitioning of drivers to the vehicle in semi-autonomous vehicles.

CHAPTER 8. THE EVALUATION, COMPARISON AND DISCUSSION WITH RELATED WORK

8.1 Case Study 1: Evaluation of Single Class Based on DL Models

This case study will complete evaluation of the outcome of proposed context-aware driver distraction severity classification using LSTM, DBN-LSTM and provide a comparison with related works.

8.1.1 Performance and Response

The network's performance (MSE) starts at 0.0327 and, after ten iterations or epochs, stops at 0.000541. Figure 8.1 graphically presents the driver distraction severity model's response; a comparison is performed between the (training, validation and testing) targets of the time-series (frame-based) data and the actual outputs. Following epoch 5, the error validation is repeated five times. As the error shows no sign of reducing, the test is halted at ten epochs. As shown in Figure 8.1, the error repeat that begins at epoch 4 shows data over-fitting. Hence, epoch five is chosen as the base, with its weights selected as the final weights. Furthermore, six iterations are run in the validation check to enhance the filter's performance; as the error does not reduce, the testing is halted at epoch 10.

8.1.2 Training and Validation

The Levenberg-Marquardt training algorithm needs more memory but less time to perform the training. It also improves performance by using the gradient-descent method. In training, the gradient begins at 0.141 and stops at 0.000219. Once the generalization ceases to improve, the training automatically stops based on the MSE of the validation samples. This occurs at epoch 10, with a validation check time of 6 secs, as shown in Figure 8.1.

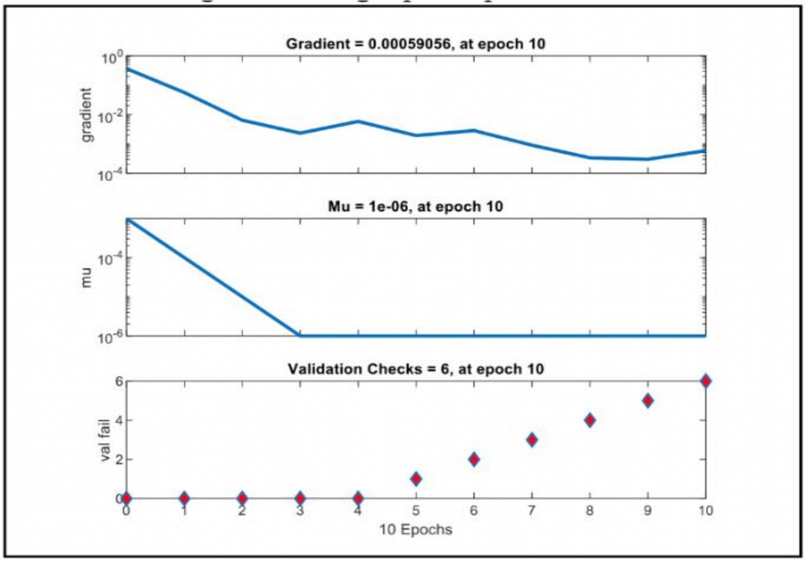

Figure 8-1: LSTM Training Using the Gradient Descent Method.

8.1.3 Discussion Related to The Combination of DBN-LSTM

As already discussed in the literature section, a massive amount of work has been done on driver's distraction classification and driving severity identification; however, having a complete algorithm that deals with the problem is crucial. Chapter 4 of this study uses time-series data provided by TeleFOT to monitor driver's distraction. A Context-Aware drivers distraction model is proposed using DBN-LSTM. The proposed systems can be further applied in adaptation to driver's behaviour, leading to a preventive and corrective system for drivers' distractions based on the severity. Though, this can be subjective to the frequency and duration in which the event occurred. However, the MDDRA model can keep track of distraction state changes and the frequency and duration of distraction. The proposed severity classifier considers the driver's distractions and considers context-aware information.

Figure 8.2 below provided the comparative analysis of our proposed Context-Aware DBN-LSTM model with the most recent work of Kouchak, S. M., & Gaffar, A. [281] in the driver's distraction task by using time-series data. The graph compares MSE for training, test, and validation for the baseline model, provided by them, the Stacked LSTM model, and our proposed Context-Aware DBN-LSTM model. The Blue bar shows the baseline model from Kouchak, S. M., & Gaffar, [281] 's work; the orange bar shows the stacked LSTM, while the grey bar shows our proposed methodology. It can be seen that the proposed MSE is far less than the recent state-of-the-art works.

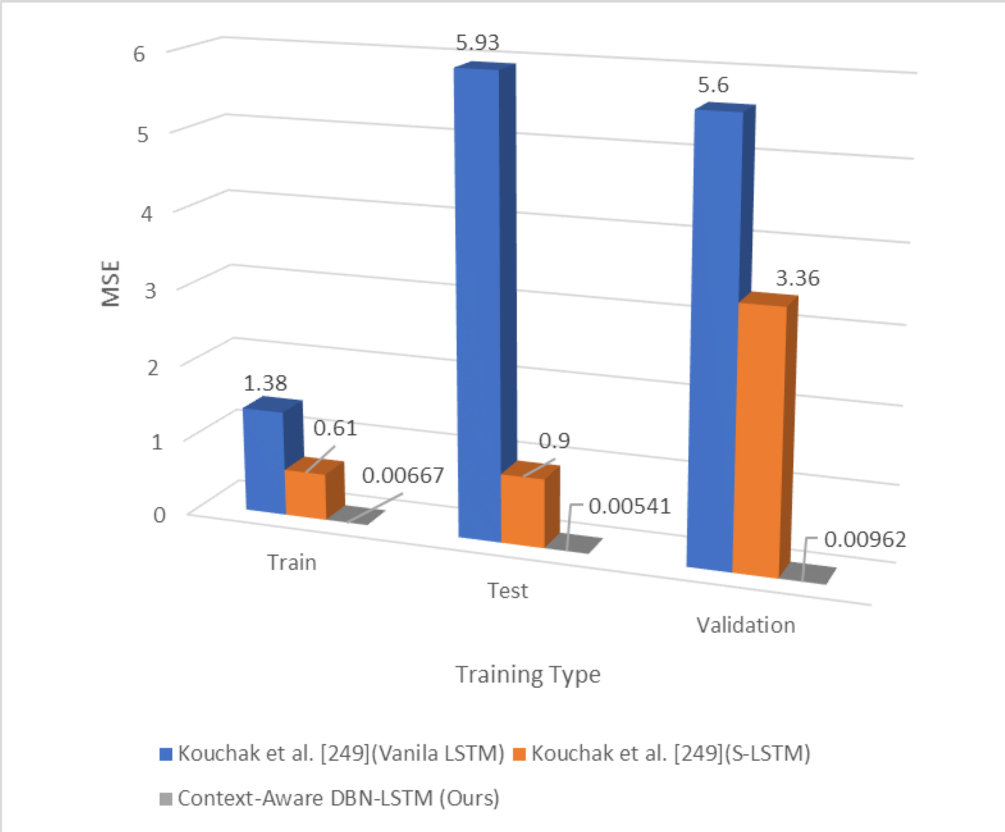

Figure 8.2: Shows the comparison of our proposed method with the current state of the arts.

Similarly, Figure 8.3 below presents a detailed comparison of Jie Chen et al [282] work, Wollmer, Martin, et al [283] work, Olabiyi, Oluatobi, et al [284] 's work, and proposed work for the task of driver's distraction. The graph clearly shows that the accuracy of the proposed Context-Aware DBN-LSTM model is superior to all the previous models. The proposed work is unique as previous studies have utilized LSTM but for different serval purposes. Jie Chen et al [282] suggested a driver distraction recognition method by utilizing the power of temporal context with the LSTM network's help. Olabiyi, Oluatobi, et al [284] suggested a similar methodology for driver distraction prediction by exploiting a Recurrent Neural Network. Finally, Wollmer, Martin, et al [283] discussed an LSTM based driver's distraction method. It can be seen from the figure below that the proposed model outperformed all the previous method's accuracy.

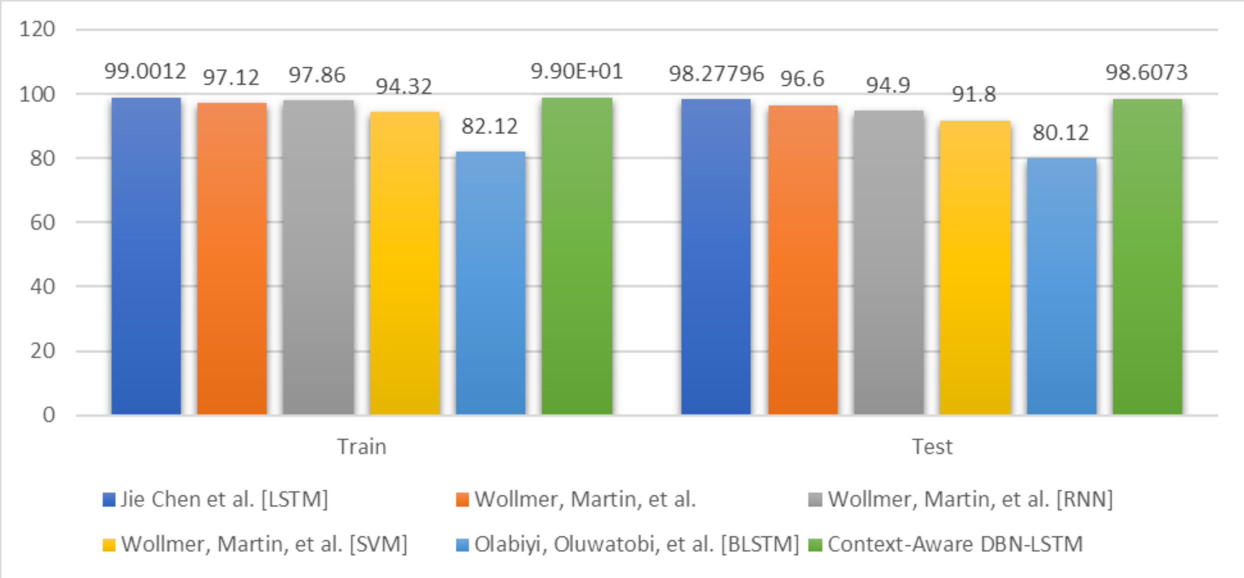

Figure 8.3: The comparative analysis of proposed methodology with the recent approaches in terms of accuracy

8.2 Case study 2: Evaluation Using Different Fuzzy Logic Model Sugeno

This evaluation entails comparing the Mamdani fuzzy logic model adopted in chapter 5 against the ANFIS Sugeno Model. The results of the Mamdani are depicted in Table 8.2 below.

Table 8-1: Training data 9/47 frames

Face Orientation*(fo)*	Driver Activity (*da*)	Hands (ha)	Previous Driver Activity (*pda*)
0	1	1	0
0	0	1	0.06666666
0	0	1	0.06666666
0	1	1	0.33333333
0	1	1	0.44444444
0	1	1	0.5
0	1	1	0.53333333
0	1	1	0.55555555
0	1	1	0.57142857

The gradual increase is evident from the Mamdani defuzzification methods, most importantly, centroid and bisector have clear visibility starting from 0.5 danger, gradually increasing further to 0.70 then finally reaching 0.80, this shows that by using the previous driver activity for a prolonged amount of time the severity of danger that the driver is in will be increased.

Table 8-2: Mamadani

CENTROID	BISECTOR	MOM	SOM	LOM
0.494678671	0	0.495	0.12	0.87
0.466961833	0.47	0.495	0.1	0.89
0.466961833	0.47	0.495	0.1	0.89
0.596267826	0.64	0.82	0.64	1
0.71235178	0.76	0.82	0.64	1
0.807455156	0.81	0.82	0.64	1
0.807455156	0.81	0.82	0.64	1
0.81177008	0.81	0.83	0.66	1

In the Sugeno comparison, the adoption of ML type of Fuzzy neural network (FNN), also known as Neuro-Fuzzy Inference System (ANFIS) Seguno, was applied to the same input image-based parameters for the Mamdani based system. The output of the ANFIS has been compared to show similar results and slightly better performance in few instances over Mamdani. In the Sugeno method, the Wtaver defuzzification method proved to also have a gradual increase with some numbers not rising to the amount that would react fast enough for level 4 semi-autonomous takeover.

An example of this would be that in the centroid defuzzification method table 8.3 of Mamdani, the 5th value down “0.71235178” would be equivalent to “0.5” on the weighted average (Wtaver) method in the Sugeno Table 8.4; this could be very severe as the transition would not happen as can be seen in chapter 5. This implies that the Membership function (MF) output the transfer to dangerous driving happens at 0.75; therefore, the ANFIS trained data would ignore this level of drivers severity, leading to an accident. Furthermore, the Weighted Sum(Wtsum).

Table 8-3: Sugeno Deffuzication Method Wtaver and Wtsum

Wtaver	Wtsum
0.479	0.5
0.479	0
0.5	0

0.499	0.613
0.5	0.788
0.807	0.877
0.809	0.693
0.811	0.793

Figure 8.4 below depicts the ANFIS neural network designed with 100 Hidden Neurons with a delay of 4.

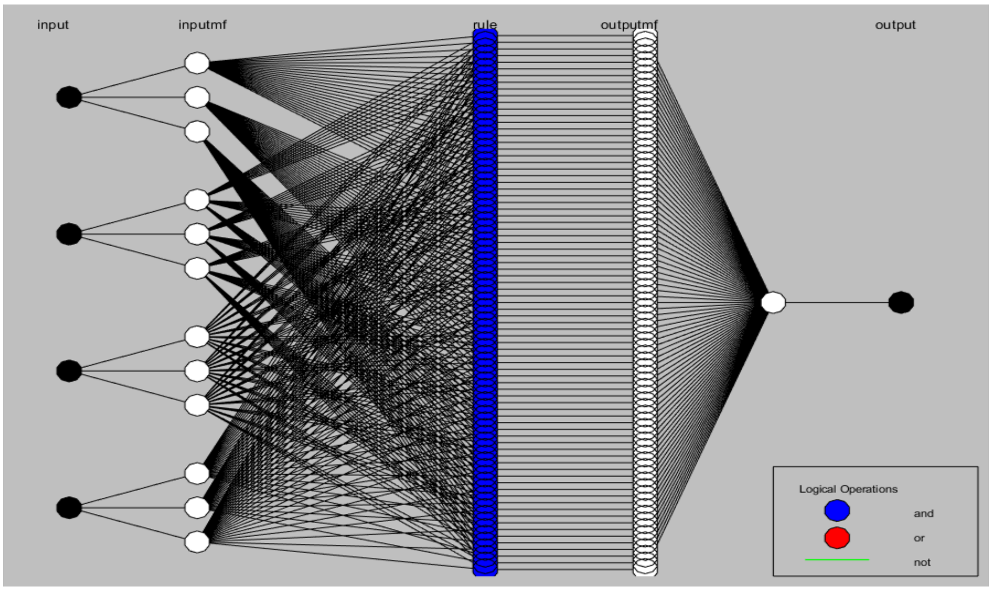

Figure 8.4: ANFIS Neural Network Implementation

The inputs (training data) is the same as the input of Mamdani, representing discrete data of 47 consecutive frames consisting of a driver with distractions such as face orientation, eyes gaze, and hands state.

8.2.1 Training

The performance of the ANFIS systems is described in Table 8.4 below, with the network showing its mean square error (MSE). 75% of the dataset was assigned for training purposes.

8.2.2 Testing

Training constitutes 15% of the dataset, and it is used.

8.2.3 Validation

The validation is used in measuring network generalization and stops when generalization stops improving. 15% of the dataset is used for validation.

Table 8-4: Training, Testing, Validation

	Target Values
Training	26
Validation	13
Testing	13

8.3 Results & Discussion

8.3.1 Performance Measures

The network's performance Mean Square Error (MSE) started at 0.037, stopped at 0.025968 after 19 epochs, and stopped training at 100 epochs. In Fig.11, the response of the severity distraction model is depicted graphically whereby the targets (training, validation and testing) of the consecutive image-based data is compared with the actual outputs.

8.3.2 Training/Validation

The training algorithm Sugeno ANFIS required less time to train but needed more memory. The algorithm applies the gradient-descent method to improve performance. In training, as seen in Figure 8.5, the gradient started at 0.037 and stopped at 0.025. Training automatically stops when the generalization stops improving at 100 epochs, as indicated by an increase in the mean square error of the validation samples; this happened at epoch 19, as seen in Figure below. The plot of MSE against the epochs showing the improvement in performance at every iteration from 1 to 19. However, the performance does not improve based on the MSE from iterations 19 to 100.

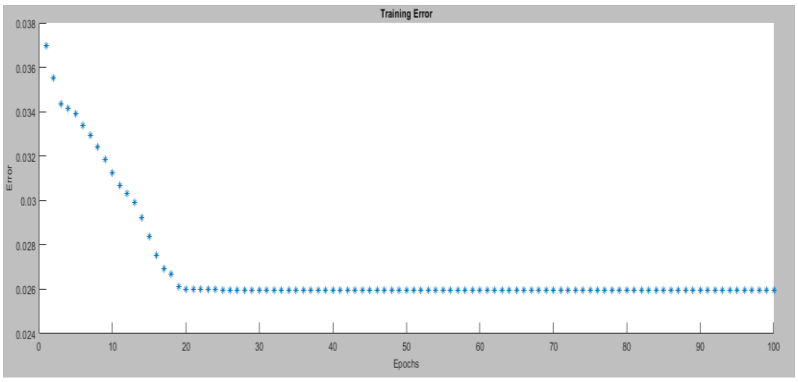

Figure 8.5 Training Performance

In Figure. 8.6 and Figure 8.7 Training and Testing output are shown below the 39 matches picked by the Sugeno and the six missed data points in blue as shown in the figure. The level of accuracy of the Sugeno is relatively close to Mamdani

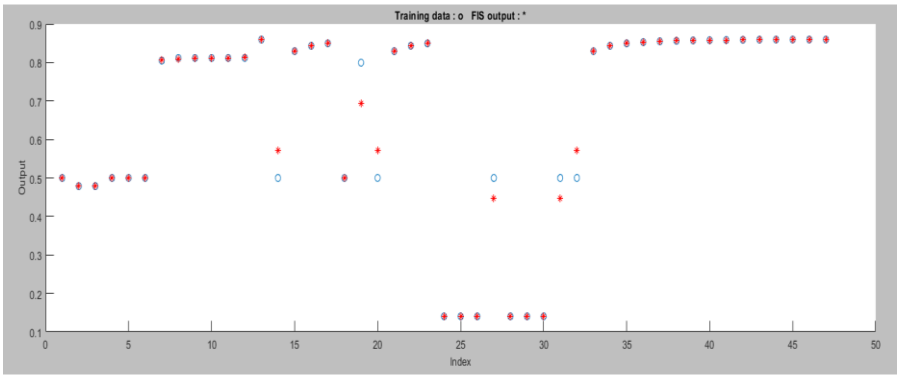

Figure 8.6 Training data FIS output

The testing data output result is shown below in Figure depicts the average testing error to be 0.25968 and generated with a Linear type Membership Function (MF) and Generate FIS of trapmf.

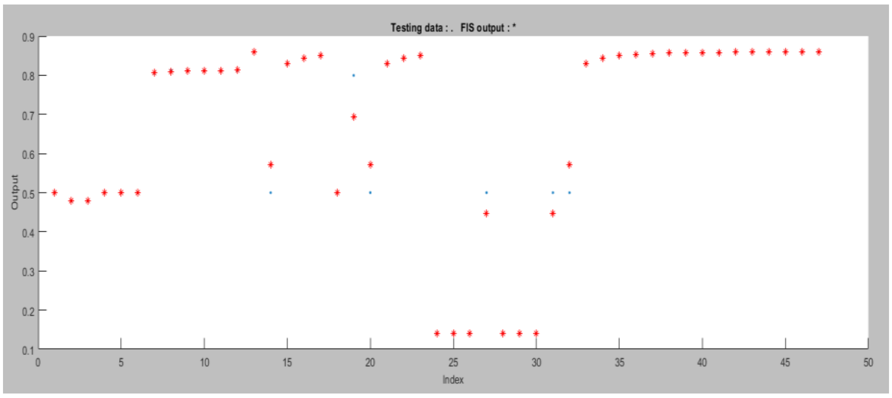

Figure 8.7: Testing data output

8.3.3 Comparison of Sugeno and Mamdani

Table 8.6 shows Sugeno RMSE results below if compared with that of Mamdami Table 8.7. The Wtaver produced the same results for the phoning and talking.

Table 8-5: Sugeno RMSE

Defuzzification Method	RMSE Value
Wtsum	0.291

Wtaver	0.327

Table 8-6: driving distraction severity levels for the membership functions

Defuzzification Method	RMSE Value	Driver Activity
CENTROID	0.32	Talking
CENTROID	0.31	Texting
CENTROID	0.32	Phoning

However, Mamdani is a better choice in this context since rules and weights were created based on expert knowledge. This approach is the idea in the context of semi-autonomous drivers where drivers can be different. For example, a comparison of how dangerous a driver is compared with another driver is possible. For instance, driving for 1-minute driving for Mamdani is sufficient, unlike Sugeno, which may require more than 5 minutes of sufficient data. However, ANFIS can take crisp input in membership function (MF) and generate rules.

In contrast, the rules generated by the fuzzy inference systems (FIS) were 81 in number compared with 48 of Mamdani. This may lead to a high computational cost, complex structure and gradient learning. Thus, a less rule-based Fuzzy system requiring less computational resources to make decisions is the best approach. ANFIS may lead to false positives (FP) if overfitting, unlike the Mamdani system, thus, the accuracy.

Comparing these approaches suggests that the Mamdani approach is superior in restrictive rules, modelling structure, and accuracy. A clear advantage Mamdani has over Sugeno is that not all possible rule combination is required to construct the fuzzy rule base. Thus, Mamdani can relate inputs and outputs in a non-linear manner through instances of sharp transitions through high to low and low to high value captured by the fuzzy membership functions. The actual outcome is to change from semi-autonomous take over from the driver when a certain threshold is reached.

8.3.4 Discussion Related to the Discrete Dynamic Bayesian

In Chapter 6, the methodology involves using an expert knowledge rule system to predict the severity of distraction in a contiguous set of video frames using the Naturalistic Driving American University of Cairo distraction Dataset. A multiclass distraction system comprises the face orientation, drivers' activities, hands, and previous driver distraction; a severity classification model is developed as a discrete dynamic Bayesian (DDB). Furthermore, a

Mamdani-based fuzzy system was implemented to detect multiclass distractions into a severity level of safe, careless, or dangerous driving. Thus, if a high level of severity is reached, the semi-autonomous vehicle will take control. Although much literature is available on the topic, most recent and state-of-art are based on deep, complex neural networks that require high computation complexity and a massive amount of energy resources. At the same time, few works have exploited the similar idea of fuzzy logic for the prescribed task and claimed excellent results. The table below shows the compression of the top recent fuzzy-based methods with our proposed methodology.

Table 8-7: Presents the collective comparison of the proposed model with the current state of the arts

Reference	Method	RMSE	Classes
Aksjonov, Andrei, et al [285]	ANN with Fuzzy logic	0.52	Phone only
Riaz, Faisal, et al [286]	Cognitive Agent-Based Computing with Fuzzy logic	0.48	Phone only
Aksjonov, Andrei, et al [287]	ML with Fuzzy Logic	0.38	Radio, Media, Telephone, Navigation
Ou, Chaojie, et al [288]	DL and Fuzzy Inferencing	0.49	Phone only
Fuzzy-logic-DDBN	Mamdani-based fuzzy system	0.32	Talking, Texting, Phoning

Table 8.8 above shows a comparison of RMSE and the number of tackled classes. It can be seen that the proposed Fuzzy-logic DDBN methodology has been applied to three classes of distractions and provided the lowest RMSE. For a clear view, the results from the relevant studies have been visualized as bars below in Figure 8.8.

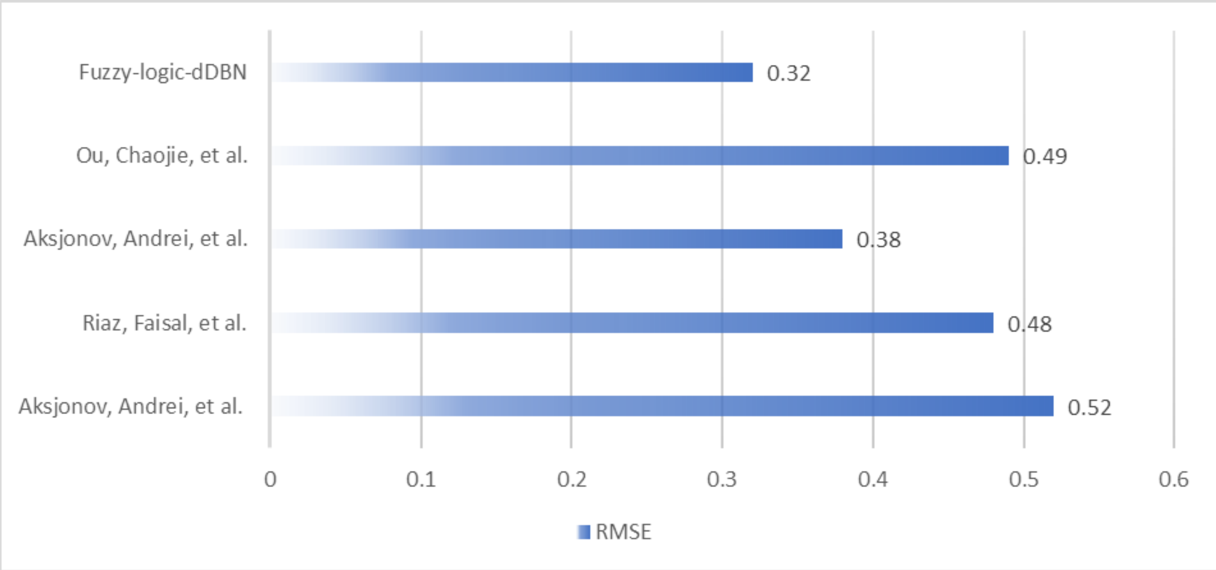

Figure 8.8 shows the comparative analysis of the proposed model in terms of RMSE with current literature

As it can be seen that the developed fuzzy-logic DDBN method has outperformed others and provided excellent results with minimal computation complexity and excellent energy efficiency.

8.4 Case Study 3: Evaluation for the MDDRA

The evaluation of chapter 6 Statistics validation and evaluation of the model performed the test using Kruskal Wallis and cross-correlation validation. To provide basic information about variables in a dataset, descriptive statistics for one of the simulated events (driver 1, event 1) is presented in Table 8.9 below. From this table, values of Mean, Median, Kurtosis and Skewness suggest that this is a case of symmetrical distribution. This might reflect on how the data were modelled as it is manually labelled before training and applied.

Table 8-8:Descriptive statistics

Mean	**0.513049625**
Standard Error	0.007304311
Median	0.508023896
Standard Deviation	0.118456024
Sample Variance	0.01403183
Kurtosis	0.057262351
Skewness	0.217203269
Range	0.725482175
Minimum	0.162361126
Maximum	0.887843301
Sum	134.9320515

Count	263
Largest(1)	0.887843301
Smallest(1)	0.162361126
Confidence Level (95.0%)	0.014382625

Invalidating the model, the model predictions were tested using correlation analysis as suggested in Section 6.3 above. This technique is used in testing the relationship between categorical variables or quantitative variables. In addition, correlation coefficients with a value between -1 and 1 are ideal. However, a value of 0 denotes no relationship at all. On the other hand, -1 and 1 imply a perfect negative or positive correlation.

Table 8-9: Correlation coefficients

State of Hand	0.425847
Road Type	0.363796
Face Orientation	0.420461
Time of day	0.224532
Eye Gaze	0.296584
Weather	0.247372
Manoeuvre	0.323121
Speed	0.053056
Surrounding	0.441935
Pedestrians	0.255076

From table 8.10 above, it is clear that there is a positive correlation with all but one parameter used in the model. This parameter is the velocity of the vehicle.

The model was also tested across multiple events, and the results demonstrated a consistent lack of correlation with velocity. This might indicate either a need for a wider velocity span to be present in the dataset or, if this will not affect results, to better represent velocity influence in the model.

Kruskal Wallis rank obtained using ML algorithms to confirm the performance in terms of accuracy, training time, and prediction time are presented in Table 8.11 below. It can be observed that ensemble learning with the Bagged model obtained the highest mean rank of 21 compared to the other variants of that model and other state-of-the-art ML algorithms. However, the mean rank for prediction and training time are 3 and 19, respectively. This ML algorithm indicates that Bagged's complex fitness function helps extract rich feature vectors to classify. As opposed to Bagged Trees, ensemble learning with Boosted Trees model obtained the highest mean rank of 21 compared to the other variants of that model and other state-of-the-art ML algorithms. This entails that the linear fitness function used in Boosted helps to extract short feature vectors to classify. Furthermore, the Linear Discriminant variant obtained the lowest mean rank of 1 in training time with 90 % accuracy.

Linear functions are evaluated using the previous severity score and the next video frame's expected total severity score. Gaussian Naïve Bayes appeared as the 2nd best algorithm that performed well except for Bagged Trees. The average mean rank gained by Gaussian Naïve Bayes is 20, with 93.2 % accuracy and a z-score of 1.49.

Table 8-10: Kruskal- Wallis ranks obtained using ML algorithms to confirm the performance concerning accuracy, training time and prediction time

		Kruskal-Wallis Average Ranks			
Model	**Median**	**Accuracy**	**Prediction Speed(~obs/sec)**	**Training Time (sec)**	**Z score**
Bagged Trees	96.2	21	3	19	1.65
Boosted Trees	58.6	1.5	21	8	-1.57
Coarse Gaussian SVM	77.2	6	9	15	-0.83
Cubic SVM	92.4	18	9	10	1.16
Fine Gaussian SVM	58.6	1.5	7	17	-1.57
Fine KNN	79.1	8	17.5	6	-0.5
Gaussian Naïve Bayes	93.2	20	20	13	1.49
Kernel Naïve Bayes	90.1	14	4	18	0.5
KNN Coarse	59.3	3	13.5	7	-1.32
KNN Cosine	80.6	10.5	16	5	-0.08
KNN Cubic	76.4	5	5	4	-0.99
KNN Weighted KNN	80.6	10.5	13.5	3	-0.08
Linear Discriminant	90.9	15	17.5	1	0.66
Linear SVM	92	16	11	12	0.83
Medium Gaussian SVM	85.2	13	6	16	0.33
Medium KNN	78.3	7	13.5	2	-0.66
Quadratic Discriminant	82.9	12	13.5	14	0.17
Quadratic SVM	92.4	18	9	11	1.16
RUSBoosted Trees	74.5	4	19	9	-1.16
Subspace Discriminant	92.4	18	2	21	1.16
Subspace KNN	79.8	9	1	20	-0.33

8.4.1 DISCUSSION RELATED TO THE MDDRA RISK ASSESSMENT MODEL

In chapter 6 ML model was deployed for the classification of drivers' distraction. The authors have proposed a novel and robust Multiclass Driver Distraction Risk Assessment (MDDRA) model. The model has tackled the driver with almost possible variants such as the current state of hand, which means whether the driver uses double hands, single hands, or no hands at all. Similarly, the type of the road on which the vehicle is running, the face orientation is on the road or off-road, whether it is a daytime or night time, the eye gaze of the driver, if the weather is dry, rain, or snowy, what is current manoeuvre is, the surrounding vehicles, speed of the vehicle, speed of the surrounding vehicle, and the pedestrians. The suggested model, MDDRA, considers vehicle, driver, and environmental context-aware situations during a journey to categorize drivers into risk matrices such as safe, careless, and dangerous. The proposed model offers flexibility to adjust parameters and weights to consider each event's specific severity level. Real-world data was collected using the Field Operation Test (TeleFOT), which consisted of drivers using the same routes in the East Midlands, UK. The results have a massive potential to reduce road accidents caused by driver's distractions. Also, a test of the correlation of driver's distraction (In-vehicle, vehicle, and environment distractions) on severity classification against continuous driver's distraction severity score was performed.

Furthermore, several ML techniques are adopted to classify and predict driver's distraction according to severity levels to aid transitioning from driver to vehicle. Figure 8.9 shows all implemented ML models such as Discriminant, Naïve Bayes, Support Vector Machine (SVM), K-Means Nearest Neighbour (KNN) Ensemble ML task of classification. The above figure shows the comparison of accuracy by applying these models. It can be seen that the Bagged Trees-based Ensemble model has provided the highest accuracy of 96.2% for classification, while fine Gaussian SVM and Boosted Trees-based ensemble methods have resulted in the lowest accuracy of 58.6% for the classification task.

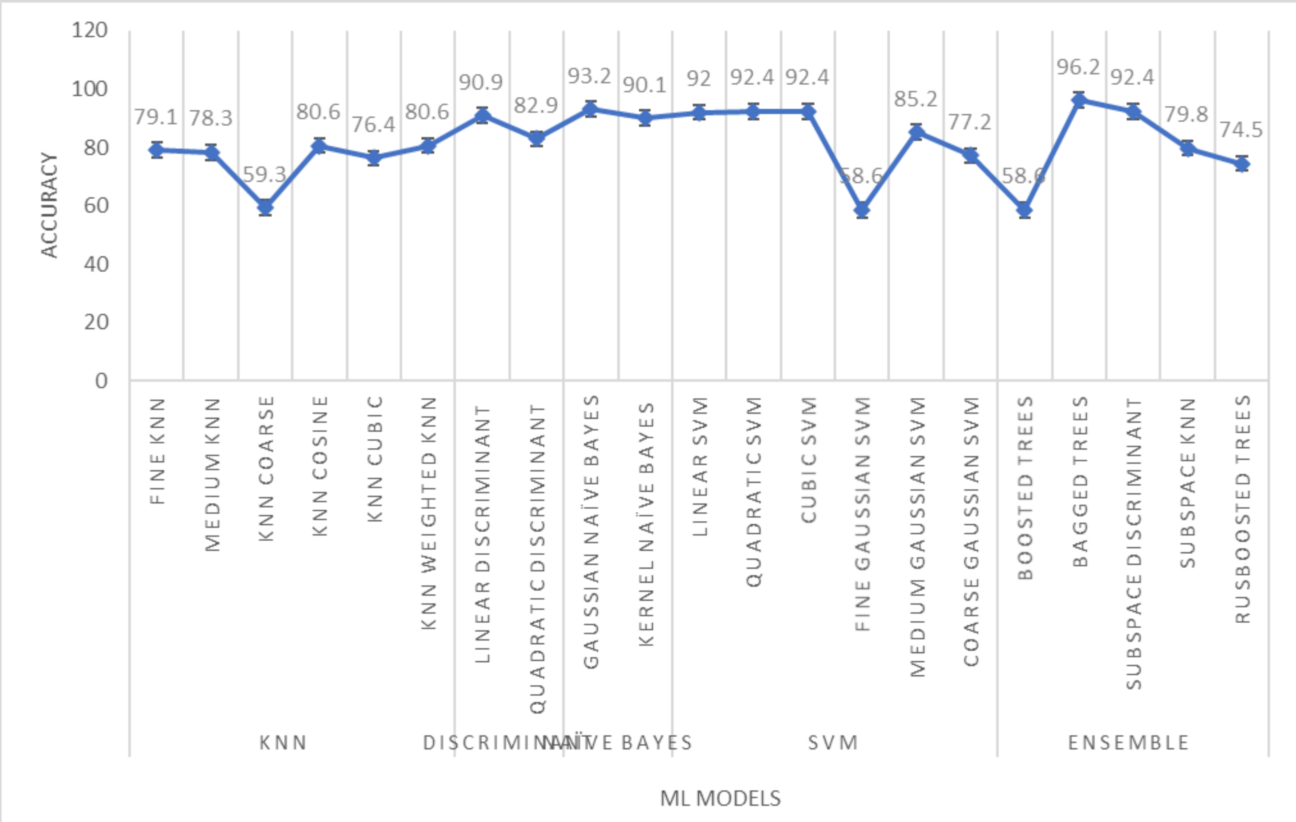

Figure 8.9: Show the comparison of accuracy across multiple ml models

Figure 8.10 above is a comparison of the optimum classifier accuracy result of the proposed MDDRA, with the works of Mengtao Zhu et al [275], Yanli Ma et al [276], and Tianchi Liu et al [277]. It can be seen that the proposed model has outperformed the current state-of-the-art in the multiclass distraction prediction. Moreover, the model has achieved an accuracy of 96.21%, while the current state-of-the-art claimed accuracy of 95.87%, which is lower than our proposed model. Although Tianchi Liu et al [277] have achieved slightly higher accuracy(97.21%), they have worked on a binary classification problem. As multiclass classification is a more complex task than a simple binary classification model, the MDDRA model state-of-the-art yielded excellent results in more than eight classes. Furthermore, the proposed model has provided fast results as high as 3600 observations per second, making the proposed model accurate but robust in terms of speed.

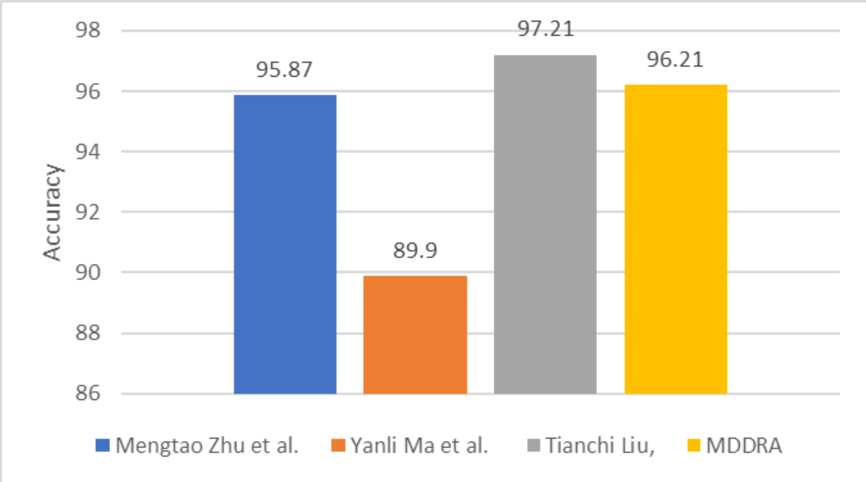

Figure 8.10: Shows the comparison of accuracy provided by MDDRA with the current state-of-the-art.

8.4.2 Evaluation of Model Using Unknown Driver and Applying K-Folds Validation

The evaluation in the section is the outcome of chapter 7, which uses K-folds cross-validation for evaluation. The evaluation results are as depicted in Table 8.12 below, and the validation accuracy resulted in an accuracy of 90.92% with a learning rate of 9e-05. Despite the quality of the dataset, the model still yielded a high result.

Table 8-11: Validation results

Validation Accuracy	Epochs	Iterations	Maximum Iterations	Learning Rate
90.92%	30/30	1550 /1560	1560	9e-05

Figure 8.11 below shows the training accuracy of the model, and its starting accuracy is from 10%

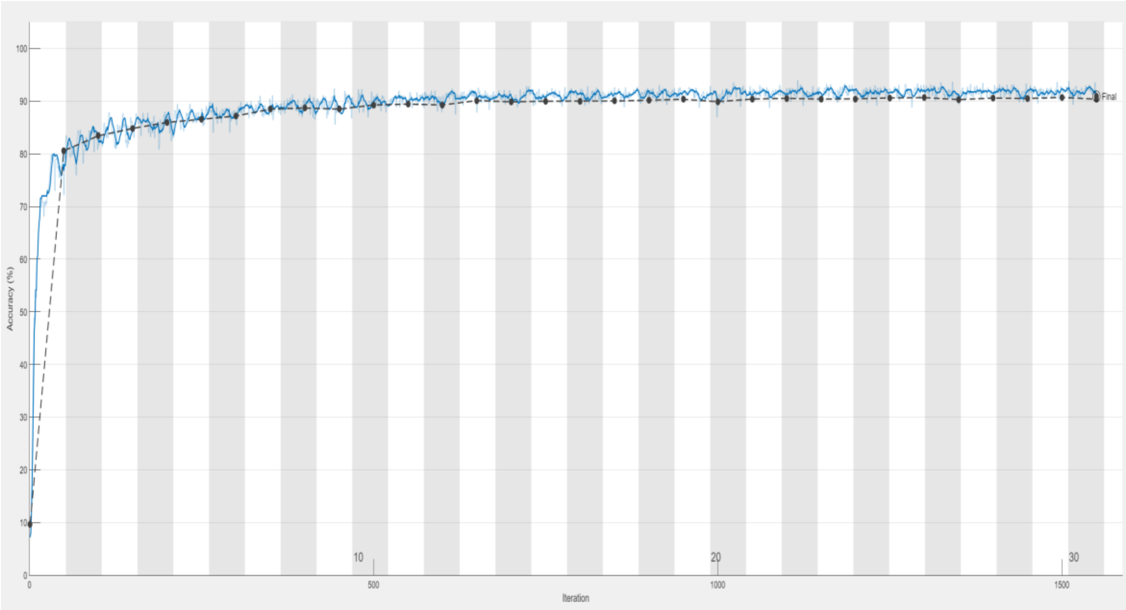

Figure 8.11: Fast-RCNN Training model

Depicted below in Figure 8.12 is the Fast RCNN training model; the loss flattens out when it got ten and gradually stabilizes from 10 epochs onwards on the horizontal axis.

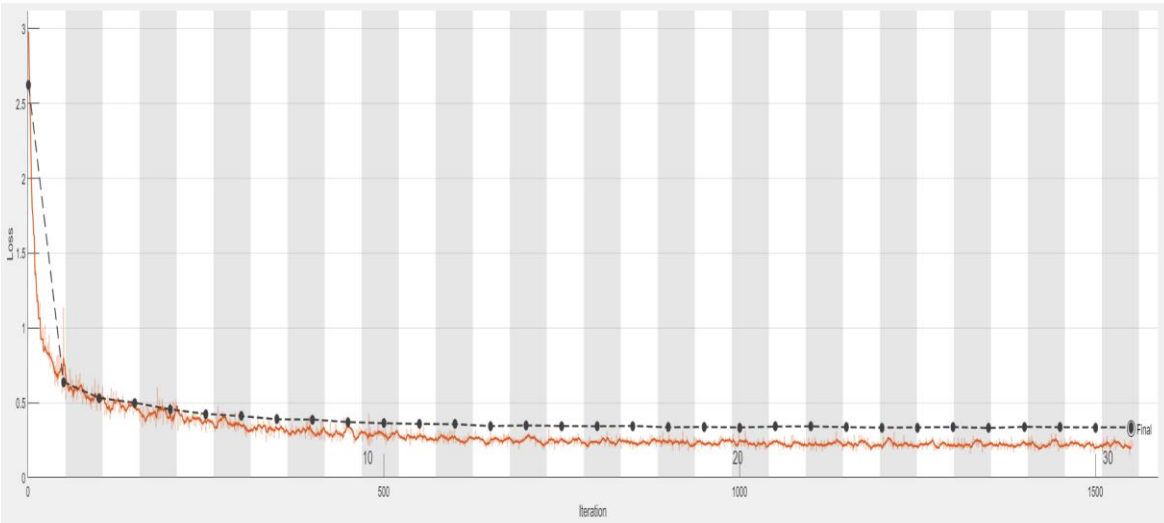

Figure 8.12: Fast-RCNN Loss Function

8.4.3 Discussion Related to the Hybrid Model Cnn-Lstm-Dbn

Chapter 7 presents a hybrid DL technique that detects and classifies drivers' distractions using a multiclass Context-Aware drivers' distraction (event types-hand state, face orientation, eye glances), in combination with several context-aware parameters: speed, weather, manoeuvre, surroundings, GPS position, accelerometer, and road type. Furthermore, a novel probabilistic DBN model based on the Fast-Recurrent CNN (FRCNN) and Long short-term memory (LSTM) network is developed to detect and classify driver's distractions into severity levels. The proposed methodology entails DL CNN trained to detect the driver's distraction, recurrent neural network layers LSTM trained to predict driver distraction severity from time-series data, and a probabilistic DBN calculates severity from probability with changing times and frames. This research entails multiclass distractions that, when combined with context-aware, leads to a severity level that can be further classified into safe, careless, or dangerous driving.

Figure 8.13 below compares the work of Li Li et al [289], Arief Koesdwiady et al [290], Duy Tran et al [291], and Yang Xing et al [292] with our proposed Hybrid CNN-DBN-LSTM model. It can be seen that the proposed model outperforms others in terms of accuracy. The method is a unique blend of CNN-LSTM-DBN. In this methodology, a hybrid CNN-LSTM architecture exploits a CNN power for feature extraction on a given set of input data, combined with a Long Short-Term Memory Network, which supports sequence prediction. By design, a CNN-LSTM was intended for handling visual time series prediction and generation of textual description from a given input sequence of images and video frames. In particular, CNN-LSTM handles activity recognition by generating textual descriptions of activities identified in sequences of images.

The results are robust and accurate because the proposed model is an intelligent fusion of various latest architecture that combines the power and effectively solves the problem. For example, in a hybrid CNN-LSTM for feature extraction and tie series prediction, the identified activity is compared with historical data to recognize the distraction type and give it a distraction identifier fed to the dynamic Bayesian network. In essence, the hybrid CNN-LSTM performs driver distraction recognition by analyzing the extracted driver features, where fuzzy sets for classification of the distraction by severity level are extracted. The fuzzy sets of distractions inform the model of the distraction ID, which is then fed to the Dynamic Bayesian model and driver features extracted by the model's deferential stage. Finally, the Out-vehicle environment monitoring process flow comprises Faster R-CNN to identify road type, weather and track pedestrians and the surrounding. A Differential component that receives data about driving speed, road type, weather, and driving manoeuvres, a Dynamic Bayesian model performs severity classification by relating the variables to each other over adjacent time steps, outputting probabilistic data, which forms the basis of operations of the ML classifier. The complete setup makes the proposed model one of a kind with robust and accurate results compared to the current literature.

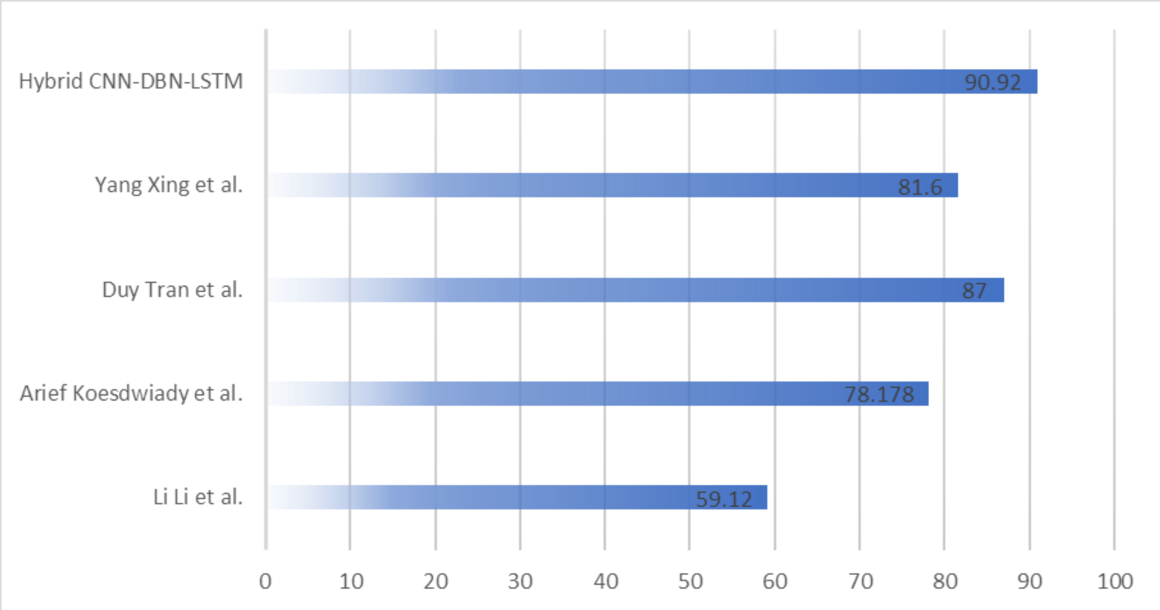

Figure 8.13 Comparison of accuracy provided by Hybrid CNN-DBN-LSTM with the current state of the arts

Figure 8.14 below shows a resultant picture of the applied Fast-RCNN model, and the results are robust and accurate. The model has detected face orientation and eyes so the eye glance and face orientation can be tracked easily.

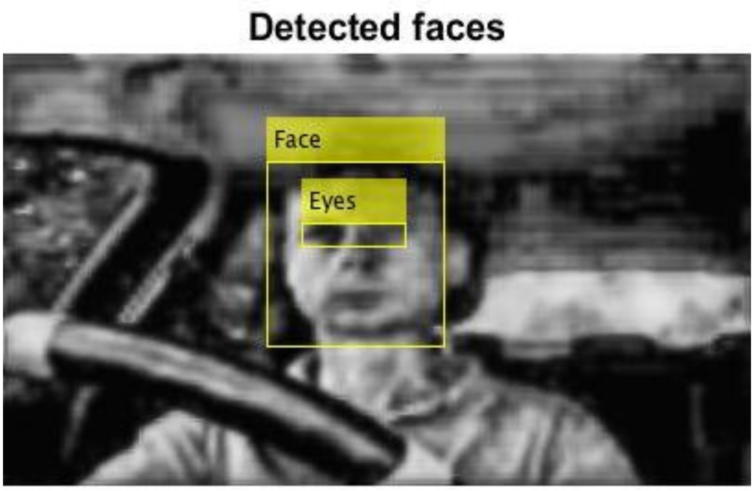

Figure 8.147: Detected Faces using Fast-RCNN

Similar results have been identified for the detection of the out-environment. Table 8.13 below provides the accuracies of detected classes.

Table 8-12: Context-aware environment detection and classification

Class	Accuracy	Class	Accuracy
Sky	0.90821	Tree	0.88847
Building	0.81932	Sign Symbol	0.76303
Pole	0.76297	Fence	0.81325

Road	0.94568	Car	0.92007
Pavement	0.89163	Pedestrian	0.85778
Sky	0.90821	Tree	0.88847

The accuracy of the proposed classes is comparable with the original faster-RCNN network. As the number of classes is far less than the original COCO dataset, that is why the employed model has shown such a robust performance than the original faster-RCNN.

8.5 Summary

There has been a comparison with numerous existing works and how the developed model outperforms the work of others in most cases. However, other algorithms achieved higher accuracy because of the distractions and parameters they applied their algorithm and model were less than the proposed models in this research.

CHAPTER 9. CONCLUSION AND FUTURE WORK

9.1 Introduction

There is a rise in the adoption of artificial intelligence in autonomous systems, mainly in autonomous vehicles and computer vision applications. In autonomous vehicles, every year, the world hinges closer to level 5 in a present-day society with the increasing adoption of mobile devices that can be used for entertainment and often used while driving, leading to a high distraction rate. Thus, needing A.I. systems in managing our lives as humans, the application area of the developed methods in this research is not limited to A.V. but also in flying cars, healthcare, and commercial industries. This research has its applications and limitations, which has been covered in this chapter.

9.2 Contribution to Knowledge

- The development of a novel ADAS framework that classifies driver distractions into severity levels to aid vehicle take-over.
- The development and evaluation of a mathematical model that classifies driver behaviour according to severity levels using thresholds.
- The definition of a threshold safety system classifies driver behaviour into careless and dangerous driving, thus enabling an autonomous vehicle to take over from the driver or know when it is safe to return autonomy to the driver.
- A novel MDDRA risk assessment model for the classification of driver distractions using ML algorithms.
- The development of a novel 3-phase parallel Fast-CNN architecture to address each physiological attribute.
- The development of a context-aware situation and using output from the parallel FCNN via a novel three-tier FCNN-DBN-LSTM that detects and classify driver's distraction into the severity level of distractions.
- The development of a Fuzzy-Logic-DDBN model for the classification of driver distraction.
- Development of a Hidden Markov Driver Distraction Severity Model (HMDDSM) for classification of driver distractions.

9.3 Research Limitations

9.3.1 Lack of Cognitive Distraction

Cognitive distraction does occur when the driver is not paying attention to the course of driving. Chapter 5, distraction such as talking to a passenger, a cognitive distraction, was identified in the dataset. However, this is our DL algorithms detected a visual kind of distraction. Furthermore, cameras and sensors embedded with a computer vision algorithm can detect when drivers talk to passengers. Moreover, non-visual related cognitive distraction can be detected using electroencephalogram (EEG) and Electrocardiogram (ECGs), which detect heart rhythms that can inform the research other factors that can impact the driver's behaviour. Electroencephalogram can be used in the detection of the electrical activity of the brain.

In contrast, it could be further argued that a degree of reaction to an external context can infer cognitive distraction. There are three main types of distractions Visual, Manual, and Cognitive. Visual distraction constitutes taking your eyes off the road. Manual entails the driver taking hands off the wheel, and cognitive constitutes taking your mind off the driving course. Thus, the researchers have only covered visual and manual-related distractions. A notable instance of cognitive distraction, such as talking to the passenger, was studied. Almahasneh et al [293], examined cognitive distraction using a simulator collecting behavioural data via EEG. The experiment involves two simulated driving sessions undergone by forty-two participants. The results showed that driving performance decreased during the execution of distractor tasks.

9.3.2 Lack of Vehicle Dynamics Detection Limitation

One factor impacting ADAS systems' accuracy is vehicle dynamics, which may occur due to road models and road surfaces. Thus, in this research, vehicle dynamics have not been considered. However, to investigate the impact of the vehicle dynamics model, implementation testbeds could be adopted. Two vehicle dynamics models that could be adopted in future work are namely Dymola and VeSyMA.

9.3.3 Ineffectiveness of Vehicle Braking System Due to The Proximity of The Vehicle to Pedestrian's

The Uber accident was due to radar detection and object at a distance, but a flaw is that the vehicle's braking system is not triggered promptly. A possible model to calibrate pedestrians' proximity to the vehicle and the vehicle speed using frame rate can prevent accidents and enable the vehicle braking system to kick in early.

9.3.4 Real-Life Deployment in Vehicle

An algorithm needs to be thoroughly tested; otherwise, the ADAS cannot be deployed in the real world, as it can involve a threat to human life. Moreover, proper testing could help mitigate this threat, and a Proof of Concept (PoC) would help mitigate deployment issues concerning computational resources.

The mitigation technique that could be implemented to address the computational resource issues could be using embedded systems, for example, Raspberry Pi. The Raspberry Pi uses limited computational power, although it has a small CPU that is unsuitable for DL-based applications, crucial for safety-critical and timely decisions required with high accuracy. Nvidia has established a simulator equipped with GPUs to mitigate this deficiency by including more sensors-based deployment capabilities like LiDAR and RADAR. The Nvidia system can be easily deployed in different context scenarios. In addition, the Virtual Reality concept can be utilised in the simulations in the form of a context-aware environment before the real-time deployment.

Moreover, the Nvidia Drive Constellation can be made available to a Drive Sim for sensors, Constellation vehicle, a software stack of autonomous vehicles. The Nvidia Drive Constellation is well designed robust simulator for testing an autonomous vehicle before road deployment [294]. It is a cost-effective approach that could help to save resources in terms of cost and human lives. This test could not be conducted as secondary data, and the cost of acquiring massive resources was not possible. Moreover, such expensive systems are not publicly available to researchers and academia.

9.3.5 Dataset Limitations

The TeleFOT dataset is outdated. The quality of the camera used in the data collection is not advanced as what is available today. Essentially, this made it challenging to detect eye gaze estimation. Thus, the picture quality of the pixels in the dot per inch is not up to 5 megapixels given the data's age. Enhancement with the data processing tool did not yield the best even after converting into a 4K resolution. However, this is solved using the segmentation technique. Besides, some of the data of the participants are corrupted.

Besides the AUCDDD, there is not vehicle data present in the dataset. Thus, only the driver distraction features were observed.

9.4 Alternative Data Source

One of the research limitations of this research is the dataset's quality because the dataset is dated. However, some of the data sources were contracted during this study. However, the data sources were out of the budget available, and some were undergoing examination. Thus, there was no access to researchers that were not part of the consortium of the study. Future research students may consider this source for the data set. Some data sources are UDRIVE, 100-car study, Naturalistic teen driving study, SHRP2, Oxford Robot Car Data Set, 8-truck naturalistic driving study, AMUSE, UAH Drive set.

9.5 Future Work

Future work will involve deploying the algorithm in ADAS, and another possible future work is enhancing the DL algorithm using a combination of Fuzzy logic and DL techniques. Besides, the speed of the model in making decisions will be analyzed. Fuzzy logic has limitations when it comes to incomplete data. Thus, a DL approach could be a better approach. There is possible integration of cognitive distraction and vehicle dynamics in future work.

9.6 Potential Application Area

9.6.1 Prevention Systems of Driver With Malicious Intent using Vehicle Take-Over

Additionally, an incident that had provided a baseline for this project was the recent protest of "***Black lives Matter," in which*** the New York Police Department (NYPD) deliberately hit protesters in New York City [295]. This incident should have been avoided if some vehicle take-over methods can prevent such malicious drivers' intent amongst the NYPD police. The proposed algorithm in this thesis can detect aggressive driving when multi-class distraction is considered and promptly take over the driver's control. As well as a FERC model using in the detection of facial expressions could be integrated to increase the classification accuracy and prevent the increase in false positives alarms.

9.6.2 Possible Real-Life Implementation and Deployment of ADAS Algorithm

The developed architecture and model have been presented to Roll Royce and Airbus. The possible suggestion of deploying the model for monitoring of pilots in the cockpit. In addition, we are presently implementing the model into the Nvidia AVX product for proof of concept. Evaluation of the accuracy and validation, confidence level to the accuracy of result using Bayesian deep learning can be adopted in the future. Future work may entail integrating Federated Learning, which can be integrated with Reinforcement learning to ensure the model learning in real-time with little data. The semi-autonomous vehicle can continue to improve or fine-tunes even when the vehicle is parked and can transition to fully autonomous with time. This implies that the vehicle will be generating data on the fly, which improves the reaction. The use of edge computing rather than sending data to the cloud is also the best, with data being pruned and weights being normalised to increase speed, accuracy(reduced loss function and noise removal) and reduce complexity.

9.6.3 Hybrid Techniques Fuzzy- DL Technique

The proposed model to measure the degree of driver distraction in semi-autonomous vehicles to aid transitioning of control to the autonomous vehicle has been done in this work. Furthermore, the proposal of an enhanced ADAS safety of vehicle drivers using Fuzzy logic rule-based multi-class drivers' distraction for classifying driver's distraction into severity levels from safe, careless, and dangerous driving when a degree of distraction is reached has been developed. However, having a hybrid system that combines a neuro-fuzzy convolutional-based approach would be a better approach. The result shows instances of correlation that drivers' distraction transitions from being careless to dangerous driving in a multi-class distraction context.

9.7 Conclusion

ADAS has been a critical component in vehicles and vital to the safety of vehicle drivers and public road transportation systems. This first part of the proposed thesis presents a DL-based technique that classifies drivers' distraction behaviour using three context-aware parameters: speed, manoeuvre, and event type—using a video coding taxonomy, studying drivers' distractions based on events information from Regions of Interest (RoI), such as hand gestures, facial orientation, and eye gaze estimation. Furthermore, a novel probabilistic (dynamic Bayesian network) model based on the Long short-term memory (LSTM) network is developed for classifying driver's distraction severity. This thesis also proposes using frame-based context data from the multi-view TeleFOT naturalistic driving study (NDS) data monitoring to classify the severity of driver distractions. The proposed methodology entails recurrent deep neural network layers trained to predict driver distraction severity from time-series data.

It is well established that the detection and classification of driver distractions are crucial in preventing road accidents. These distractions impact both driver behaviour and vehicle dynamics. Knowing the degree of driver distraction can aid in accident prevention techniques, including transitioning control to a semi-autonomous vehicle level when a high distraction severity level is reached. Thus, ADAS enhancement is critical in-vehicle drivers' and other road users' safety. In the second part of the thesis, a novel methodology is introduced, using an expert knowledge rule system to predict the severity of distraction in a contiguous set of video frames using the Naturalistic AUCDDD. A multi-class distraction system comprises the face orientation, drivers' activities, hands, and previous driver distraction. A severity classification model is developed as a discrete dynamic Bayesian.

Furthermore, a Mamdani-based fuzzy system was implemented to detect multi-class distractions into a severity level of safe, careless, or dangerous driving. Thus, if a high level of severity is reached, the semi-autonomous vehicle will take control. The result further shows that some driver's distractions may quickly transition from careless to dangerous driving in a multi-class distraction context.

Similarly, risk mitigation techniques are crucial to preventing driving behaviour-related accidents; the third part of the thesis provides a novel Multi-Class Driver Distraction Risk Assessment model. MDDRA considers vehicle, driver, and environmental data during a journey to categorize drivers into a risk matrix such as safe, careless, and dangerous. The model offers flexibility to adjust parameters and weights to consider each event into a specific severity level. Real-world data was collected using the Field Operation Test (TeleFOT), which consisted of drivers using the same routes in the East Midlands, U.K. The results conclude that it is possible to reduce road accidents caused by driver's distractions. Also, the correlation of driver's distraction (In-vehicle, vehicle, and environment distractions) is tested on severity classification against continuous driver's distraction severity score. The applied ML techniques classify and predict driver's distraction according to severity levels to aid transitioning from driver to vehicle. The algorithm that gave the best performance is Ensemble Bagged Trees which observed an accuracy of 96.2%.

The thesis's final chapter provides ADAS, a critical component in semi-autonomous vehicles and vital to vehicle drivers and public road transportation systems. In the last chapter, the present is a hybrid DL technique that detects and classifies drivers' distractions using a multi-class Context-Aware drivers' distraction (event types-hand state, face orientation, eye glances) in combination with several context-awareness parameters: speed, weather, manoeuvre, surroundings, GPS position, accelerometer, and road type. Furthermore, a novel probabilistic DBN model based on the FRCNN and LSTM network is developed for detection and classifying driver's distraction into severity levels. The thesis presents frame-based context data from the multi-view TeleFOT NDS data monitoring to classify the severity of driver distractions. Our proposed methodology entails DL-CNN trained to detect the driver's distraction, recurrent neural network layers LSTM trained to predict driver distraction severity from time-series data, and a probabilistic DBN that calculates severity from probability with changing times and frames. This research entails multi-class distractions that, when combined with context-aware, leads to a severity level that can be further classified into safe, careless, or dangerous driving.

Furthermore, an HMM is used in the take-over of transitioning from driver to semi-autonomous vehicle. The model is called the Hidden Markov Driver Severity Model (HMDDSM). Validation of these results was performed using a k-folds validation method applied to an unseen driver dataset.

REFERENCES

[1] T. A. Ranney, E. Mazzae, R. Garrott, and M. J. Goodman, “NHTSA driver distraction research: Past, present, and future,” in *Driver distraction internet forum*, 2000, vol. 2000. Accessed: Jan. 07, 2017. [Online]. Available: http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv17/Proceed/00177.pdf

[2] C. Smith, *The car hacker’s handbook: a guide for the penetration tester*. San Francisco: No Starch Press, 2016.

[3] L. Chong, M. Abbas, and A. Medina, “Simulation of Driver Behavior with Agent-Based Back-Propagation Neural Network,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2249, pp. 44–51, Dec. 2011, doi: 10.3141/2249-07.

[4] D. R. Large, E. van Loon, G. Burnett, and S. Pournami, “Applying NHTSA task acceptance criteria to different simulated driving scenarios,” 2015, pp. 117–124. doi: 10.1145/2799250.2799254.

[5] I. I. Phase and C. Field, “The 100-Car Naturalistic Driving Study,” 2006, Accessed: Jul. 08, 2016. [Online]. Available: http://erso.swov.nl/knowledge/Fixed/55_fatique/ref%2029%20dingus.pdf

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105. Accessed: Jan. 18, 2017. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-w

[7] O. E. David and N. S. Netanyahu, “DeepSign: Deep learning for automatic malware signature generation and classification,” in *2015 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2015, pp. 1–8. doi: 10.1109/IJCNN.2015.7280815.

[8] I. M. Berry, “The effects of driving style and vehicle performance on the real-world fuel consumption of US light-duty vehicles,” Massachusetts Institute of Technology, 2010. Accessed: Jan. 15, 2017. [Online]. Available: http://www.web.mit.edu/sloan-auto-lab/research/beforeh2/files/IreneBerry_Thesis_February2010.pdf

[9] A. McGordon, J. E. W. Poxon, C. Cheng, R. P. Jones, and P. A. Jennings, “Development of a driver model to study the effects of real-world driver behaviour on

the fuel consumption," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 225, no. 11, pp. 1518–1530, Nov. 2011, doi: 10.1177/0954407011409116.

[10] C. Braunagel, E. Kasneci, W. Stolzmann, and W. Rosenstiel, "Driver-Activity Recognition in the Context of Conditionally Autonomous Driving," Sep. 2015, pp. 1652–1657. doi: 10.1109/ITSC.2015.268.

[11] "Dangerous Driving: Fact Sheets: News Centre: The Crown Prosecution Service." http://www.cps.gov.uk/news/fact_sheets/dangerous_driving/ (accessed Jan. 05, 2017).

[12] M. ALVESSON and J. SANDBERG, "GENERATING RESEARCH QUESTIONS THROUGH PROBLEMATIZATION," *The Academy of Management Review*, vol. 36, no. 2, pp. 247–271, 2011.

[13] C. Ahmet, *Artificial Intelligence: How Advanced Machine Learning Will Shape The Future Of Our World*. Independently published, 2018.

[14] M. Bamba, "DEVELOPMENT AND APPLICATION OF NOVEL COMPUTATIONAL INTELLIGENCE TECHNIQUES TO THE MULTIVARIATE ANALYSIS OF METABOLOMICS BIOFLUIDS DATASETS," *undefined*, 2017. /paper/DEVELOPMENT-AND-APPLICATION-OF-NOVEL-COMPUTATIONAL-Bamba/01c3e0a697e4cd5e6c370200f696d78389704878 (accessed Jan. 19, 2021).

[15] J. Frankenfield, "How Artificial Intelligence Works," *Investopedia*. https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp (accessed Jan. 19, 2021).

[16] "Artificial Intelligence for All: by Utpal Chakraborty | Waterstones." https://www.waterstones.com/book/artificial-intelligence-for-all/utpal-chakraborty//9789389328509 (accessed Jan. 19, 2021).

[17] E. Alpaydin, *Introduction to machine learning*, 2nd ed. Cambridge, Mass: MIT Press, 2010.

[18] H. Lin, Z. Yan, and Y. Fu, " security-related data collection with context awareness," *Journal of Network and Computer Applications*, vol. 126, pp. 88–103, Jan. 2019, doi: 10.1016/j.jnca.2018.11.002.

[19] H. E. Byun and K. Cheverst, "Exploiting User Models and Context-Awareness to Support Personal Daily Activities," in *Personal Daily Activities, Workshop in UM2001 on User Modelling for Context-Aware Applications*, 2001, pp. 13–16.

[20] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in *Advances in Neural Information Processing Systems*, 2013, pp. 2553–2561. Accessed:

Jun. 21, 2017. [Online]. Available: http://papers.nips.cc/paper/5207-deep-neural-networks-for-object-detection

[21] A. Mathur, N. D. Lane, S. Bhattacharya, A. Boran, C. Forlivesi, and F. Kawsar, “DeepEye: Resource Efficient Local Execution of Multiple Deep Vision Models using Wearable Commodity Hardware,” in *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, New York, NY, USA, Jun. 2017, pp. 68–81. doi: 10.1145/3081333.3081359.

[22] M. Xu, M. Zhu, Y. Liu, F. X. Lin, and X. Liu, “DeepCache: Principled Cache for Mobile Deep Vision,” *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, pp. 129–144, Oct. 2018, doi: 10.1145/3241539.3241563.

[23] “Deep Vision extracts information from images,” *TNO*. /en/focus-areas/artificial-intelligence/techniques/deep-vision/ (accessed Jan. 19, 2021).

[24] K. He, X. Zhang, S. Ren, and J. Sun, “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition,” *arXiv:1406.4729 [cs]*, Jun. 2014, doi: 10.1007/978-3-319-10578-9_23.

[25] R. Szeliski, “Computer Vision: Algorithms and Applications,” p. 979.

[26] S. Livingstone, “Online risk, harm and vulnerability: reflections on the evidence base for child Internet safety policy,” *ZER: Journal of Communication Studies*, vol. 18, no. 35, Art. no. 35, Nov. 2013.

[27] L. Tijerina, E. B. Parmer, and M. J. Goodman, “Individual Differences and in-Vehicle Distraction While Driving: A Test Track Study and Psychometric Evaluation,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 43, no. 18, pp. 982–986, Sep. 1999, doi: 10.1177/154193129904301808.

[28] R. Tron and R. Vidal, “Distributed Computer Vision Algorithms,” *IEEE Signal Processing Magazine*, 2011, doi: 10.1109/MSP.2011.940399.

[29] A. Al-Kaff, D. Martn, F. Garca, A. de la Escalera, and J. Mara Armingol, “Survey of computer vision algorithms and applications for unmanned aerial vehicles,” *Expert Syst. Appl.*, vol. 92, no. C, pp. 447–463, Feb. 2018, doi: 10.1016/j.eswa.2017.09.033.

[30] K. Dalamagkidis, K. Valavanis, and L. A. Piegl, *On Integrating Unmanned Aircraft Systems into the National Airspace System: Issues, Challenges, Operational Restrictions, Certification, and Recommendations*, 2nd ed. Springer Netherlands, 2012. doi: 10.1007/978-94-007-2479-2.

[31] S. Adams and C. Friedland, "A SURVEY OF UNMANNED AERIAL VEHICLE (UAV) USAGE FOR IMAGERY," 2011. /paper/A-SURVEY-OF-UNMANNED-AERIAL-VEHICLE-(-UAV-)-USAGE-Adams-Friedland/fd8e960ca48e183452335743c273ea41c6930a75 (accessed Jan. 19, 2021).

[32] A. Choudhury, "10 Open Datasets You Can Use For Computer Vision Projects," *Analytics India Magazine*, Jul. 31, 2019. https://analyticsindiamag.com/10-open-datasets-you-can-use-for-computer-vision-projects/ (accessed Jan. 19, 2021).

[33] J. Zhang, F. Y. Wang, K. Wang, W. H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: A survey," *IEEE Intelligent Transportation Systems Magazine*, vol. 12, no. 4, pp. 1624–1639, Dec. 2011, doi: 10.1109/TITS.2011.2158001.

[34] L. Malta, C. Miyajima, and K. Takeda, "A Study of Driver Behavior Under Potential Threats in Vehicle Traffic," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 10, pp. 201–210, Jul. 2009, doi: 10.1109/TITS.2009.2018321.

[35] J. M. Sussman, *Perspectives on intelligent transportation systems (ITS)*. New York: Springer Science+Business Media, 2005.

[36] M. Forghani, J. M. McNew, D. Hoehener, and D. D. Vecchio, "Design of Driver-Assist Systems Under Probabilistic Safety Specifications Near Stop Signs," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 1, pp. 43–53, Jan. 2016, doi: 10.1109/TASE.2015.2499221.

[37] J. J. Jain and C. Busso, "Analysis of driver behaviors during common tasks using frontal video camera and CAN-Bus information," in *2011 IEEE International Conference on Multimedia and Expo*, Jul. 2011, pp. 1–6. doi: 10.1109/ICME.2011.6011950.

[38] S. Choi, J. Kim, D. Kwak, P. Angkititrakul, and J. H. Hansen, "Analysis and classification of driver behavior using in-vehicle can-bus information," in *Biennial Workshop on DSP for In-Vehicle and Mobile Systems*, 2007, pp. 17–19. Accessed: Jul. 06, 2016. [Online]. Available: https://www.researchgate.net/profile/Pongtep_Angkititrakul/publication/228936619_Analysis_and_Classification_of_Driver_Behavior_using_In-Vehicle_CAN-Bus_Information/links/00b7d51c0e8f32c704000000.pdf

[39] N. Li, J. J. Jain, and C. Busso, "Modeling of Driver Behavior in Real World Scenarios Using Multiple Noninvasive Sensors," *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1213–1225, Aug. 2013, doi: 10.1109/TMM.2013.2241416.

[40] M. R. Othman, Z. Zhang, T. Imamura, and T. Miyake, "A study of analysis method for driver features extraction," in *IEEE International Conference on Systems, Man and*

Cybernetics, 2008. SMC 2008, Oct. 2008, pp. 1501–1505. doi: 10.1109/ICSMC.2008.4811498.

[41] G. Castignani, T. Derrmann, R. Frank, and T. Engel, “Driver Behavior Profiling Using Smartphones: A Low-Cost Platform for Driver Monitoring,” *IEEE Intelligent Transportation Systems Magazine*, vol. 7, no. 1, pp. 91–102, Spring 2015, doi: 10.1109/MITS.2014.2328673.

[42] G. Castignani, R. Frank, and T. Engel, “An evaluation study of driver profiling fuzzy algorithms using smartphones,” in *2013 21st IEEE International Conference on Network Protocols (ICNP)*, Oct. 2013, pp. 1–6. doi: 10.1109/ICNP.2013.6733681.

[43] M. Fazeen, B. Gozick, R. Dantu, M. Bhukhiya, and M. C. González, “Safe Driving Using Mobile Phones,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1462–1468, Sep. 2012, doi: 10.1109/TITS.2012.2187640.

[44] P. Mohan, V. N. Padmanabhan, and R. Ramjee, “Nericell: rich monitoring of road and traffic conditions using mobile smartphones,” in *Proceedings of the 6th ACM conference on Embedded network sensor systems*, 2008, pp. 323–336. Accessed: Jun. 26, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=1460444

[45] H. Eren, S. Makinist, E. Akin, and A. Yilmaz, “Estimating driving behavior by a smartphone,” in *2012 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2012, pp. 234–239. doi: 10.1109/IVS.2012.6232298.

[46] D. A. Johnson and M. M. Trivedi, “Driving style recognition using a smartphone as a sensor platform,” in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct. 2011, pp. 1609–1615. doi: 10.1109/ITSC.2011.6083078.

[47] C. Saiprasert, S. Thajchayapong, T. Pholprasit, and C. Tanprasert, “Driver behaviour profiling using smartphone sensory data in a V2I environment,” in *2014 International Conference on Connected Vehicles and Expo (ICCVE)*, Nov. 2014, pp. 552–557. doi: 10.1109/ICCVE.2014.7297609.

[48] Y.-J. Kweon and K. M. Kockelman, “Driver attitudes and choices: speed limits, seat belt use, and drinking-and-driving,” in *Journal of the Transportation Research Forum*, 2010, vol. 45. Accessed: Jun. 26, 2016. [Online]. Available: http://journals.library.oregonstate.edu/index.php/trforum/article/view/611

[49] V. D. Khairnar and S. N. Pradhan, “Comparative study of simulation for vehicular ad-hoc network,” *arXiv preprint arXiv:1304.5181*, 2013, Accessed: May 09, 2016. [Online]. Available: http://arxiv.org/abs/1304.5181

[50] N. AbuAli, “Advanced vehicular sensing of road artifacts and driver behavior,” in *2015 IEEE Symposium on Computers and Communication (ISCC)*, Jul. 2015, pp. 45–49. doi: 10.1109/ISCC.2015.7405452.

[51] Y. Saito, M. Itoh, and T. Inagaki, “Driver Assistance System With a Dual Control Scheme: Effectiveness of Identifying Driver Drowsiness and Preventing Lane Departure Accidents,” *IEEE Transactions on Human-Machine Systems*, vol. PP, no. 99, pp. 1–12, 2016, doi: 10.1109/THMS.2016.2549032.

[52] N. Akhtar, K. Pandey, and S. Gupta, “Mobile Application for Safe Driving,” in *2014 Fourth International Conference on Communication Systems and Network Technologies (CSNT)*, Apr. 2014, pp. 212–216. doi: 10.1109/CSNT.2014.49.

[53] D. Vitelli, “Security Vulnerabilities of Vehicular Platoon Network”, Accessed: Sep. 25, 2016. [Online]. Available: http://www.ingegneria-informatica.unina.it/sites/default/files/elaborati_tesi/2016/05/Elaborato%20Vitelli%20Daniele%20n46001378.pdf

[54] T. A. Ranney, E. Mazzae, R. Garrott, and M. J. Goodman, “NHTSA driver distraction research: Past, present, and future,” in *Driver distraction internet forum*, 2000, vol. 2000. Accessed: Jan. 07, 2017. [Online]. Available: http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv17/Proceed/00177.pdf

[55] S. Lim and J. H. Yang, “Driver state estimation by convolutional neural network using multimodal sensor data,” *Electronics Letters*, vol. 52, no. 17, pp. 1495–1497, 2016, doi: 10.1049/el.2016.1393.

[56] S. Lim and J. H. Yang, “Driver state estimation by convolutional neural network using multimodal sensor data,” *Electronics Letters*, vol. 52, no. 17, pp. 1495–1497, 2016, doi: 10.1049/el.2016.1393.

[57] Y. Bengio, “Learning Deep Architectures for AI,” *Foundations and Trends® in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009, doi: 10.1561/2200000006.

[58] U. Stern, R. He, and C.-H. Yang, “Analyzing animal behavior via classifying each video frame using convolutional neural networks,” *Scientific Reports*, vol. 5, p. 14351, Sep. 2015, doi: 10.1038/srep14351.

[59] P. Liu, A. Kurt, K. Redmill, and U. Ozguner, “Classification of Highway Lane Change Behavior to Detect Dangerous Cut-in Maneuvers,” 2015. Accessed: Jul. 06, 2016. [Online]. Available: http://docs.trb.org/prp/16-5124.pdf

[60] K. Tang, S. Zhu, Y. Xu, and F. Wang, “Modeling Drivers’ Dynamic Decision-Making Behavior During the Phase Transition Period: An Analytical Approach Based on

Hidden Markov Model Theory," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 206–214, Jan. 2016, doi: 10.1109/TITS.2015.2462738.

[61] D. Sadigh *et al.*, "Data-driven probabilistic modeling and verification of human driver behavior," *Formal Verification and Modeling in Human-Machine Systems*, 2014, Accessed: Jul. 06, 2016. [Online]. Available: http://people.eecs.berkeley.edu/~dsadigh/Papers/sadigh-fvmhs2014-arxiv.pdf

[62] J. C. Mccall and M. M. Trivedi, "Human Behavior Based Predictive Brake Assistance," in *2006 IEEE Intelligent Vehicles Symposium*, 2006, pp. 8–12. doi: 10.1109/IVS.2006.1689597.

[63] T. Gindele, S. Brechtel, and R. Dillmann, "A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments," in *2010 13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Sep. 2010, pp. 1625–1631. doi: 10.1109/ITSC.2010.5625262.

[64] A. Houenou, P. Bonnifait, V. Cherfaoui, and W. Yao, "Vehicle trajectory prediction based on motion model and maneuver recognition," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov. 2013, pp. 4363–4369. doi: 10.1109/IROS.2013.6696982.

[65] J. B. Tomas-Gabarron, E. Egea-Lopez, and J. Garcia-Haro, "Vehicular Trajectory Optimization for Cooperative Collision Avoidance at High Speeds," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1930–1941, Dec. 2013, doi: 10.1109/TITS.2013.2270009.

[66] G. S. Larue, A. Rakotonirainy, and A. N. Pettitt, "Predicting Reduced Driver Alertness on Monotonous Highways," *IEEE Pervasive Computing*, vol. 14, no. 2, pp. 78–85, Apr. 2015, doi: 10.1109/MPRV.2015.38.

[67] A. A. Albousefi *et al.*, "A support vector machine approach to unintentional vehicle lane departure prediction," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, Jun. 2014, pp. 299–303. doi: 10.1109/IVS.2014.6856602.

[68] J. C. McCall and M. M. Trivedi, "Driver Behavior and Situation Aware Brake Assistance for Intelligent Vehicles," *Proceedings of the IEEE*, vol. 95, no. 2, pp. 374–387, Feb. 2007, doi: 10.1109/JPROC.2006.888388.

[69] A. Bender, J. R. Ward, S. Worrall, and E. M. Nebot, "Predicting Driver Intent from Models of Naturalistic Driving," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Sep. 2015, pp. 1609–1615. doi: 10.1109/ITSC.2015.262.

[70] M. K. Kumar and V. K. Prasad, “Driver Behavior Analysis and Prediction Models: A Survey,” *IJCSIT) International Journal of Computer Science and Information Technologies*, vol. 6, no. 4, Accessed: Jul. 06, 2016. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.695.5255&rep=rep1&type=pdf

[71] S. I. Ao, International Association of Engineers, and International Conference on Software Engineering ; (Hong Kong) : 2015.03.18-20, Eds., *The 2015 IAENG International Conference on Artificial Intelligence and Applications, the 2015 IAENG International Conference on Bioinformatics, the 2015 IAENG International Conference on Computer Science, the 2015 IAENG International Conference on Control and Automation, the 2015 IAENG International Conference on Data Mining and Applications, the 2015 IAENG International Conference on Imaging Engineering, the 2015 IAENG International Conference on Internet Computing and Web Services, the 2015 IAENG International Conference on Scientific Computing, the 2015 IAENG International Conference on Software Engineering*. Hong Kong: IAENG, 2015.

[72] H. Nishiuchi, “Report Traffic Safety Education and Awareness Activities in Japan”, Accessed: Jan. 21, 2017. [Online]. Available: http://www.iatss.or.jp/common/pdf/en/iatss/composition/FY2014_Report_JP_En.pdf

[73] T. Imamura, T. Ogi, Z. Zhang, and T. Miyake, “Real-time implementation of estimation method for driver’s intention on a driving simulator,” in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2014, pp. 1904–1909. doi: 10.1109/SMC.2014.6974199.

[74] N. Tobe, T. Imamura, Z. Zhang, T. Miyake, and Y. Asakawa, “Study on Classification Method of Risk Perception Based on the Driving Knowledge and Behavior,” in *2015 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2015, pp. 1261–1266. doi: 10.1109/SMC.2015.225.

[75] M. Sieber and B. Färber, “Driver perception and reaction in collision avoidance: Implications for ADAS development and testing,” in *2016 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2016, pp. 239–245. doi: 10.1109/IVS.2016.7535392.

[76] H. j Gunther, O. Trauer, and L. Wolf, “The potential of collective perception in vehicular ad-hoc networks,” in *2015 14th International Conference on ITS Telecommunications (ITST)*, Dec. 2015, pp. 1–5. doi: 10.1109/ITST.2015.7377190.

[77] C. Chai and X. Wu, "Gender Differences in Driving Anger Emotional Susceptibility," in *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, Dec. 2015, vol. 2, pp. 590–593. doi: 10.1109/ISCID.2015.136.

[78] F. Gaibler, S. Faber, S. Edenhofer, and S. von Mammen, "Drink #38;amp; Drive: A Serious but Fun Game on Alcohol-Induced Impairments in Road Traffic," in *2015 7th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games)*, Sep. 2015, pp. 1–5. doi: 10.1109/VS-GAMES.2015.7295771.

[79] V. R. Tannahill, K. M. Muttaqi, and D. Sutanto, "Driver alerting system using range estimation of electric vehicles in real time under dynamically varying environmental conditions," *IET Electrical Systems in Transportation*, vol. 6, no. 2, pp. 107–116, 2016, doi: 10.1049/iet-est.2014.0067.

[80] N. Pugeault and R. Bowden, "How Much of Driving Is Preattentive?," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5424–5438, Dec. 2015, doi: 10.1109/TVT.2015.2487826.

[81] N. Pugeault and R. Bowden, "How Much of Driving Is Preattentive?," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5424–5438, Dec. 2015, doi: 10.1109/TVT.2015.2487826.

[82] P. S. Rani, P. Subhashree, and N. S. Devi, "Computer vision based gaze tracking for accident prevention," in *2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, Feb. 2016, pp. 1–6. doi: 10.1109/STARTUP.2016.7583976.

[83] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, Jun. 2005, doi: 10.1109/TITS.2005.848368.

[84] D. S. Lee, T. W. Chong, and B. G. Lee, "Stress Events Detection of Driver by Wearable Glove System," *IEEE Sensors Journal*, pp. 1–1, 2016, doi: 10.1109/JSEN.2016.2625323.

[85] V. Alizadeh and O. Dehzangi, "The impact of secondary tasks on drivers during naturalistic driving: Analysis of EEG dynamics," in *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*, 2016, pp. 2493–2499. Accessed: Mar. 20, 2017. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7795957/

[86] P. Fairley, "The Self-Driving Car's Bicycle Problem," *IEEE Spectrum: Technology, Engineering, and Science News*, Jan. 31, 2017. http://spectrum.ieee.org/cars-that-

think/transportation/self-driving/the-selfdriving-cars-bicycle-problem (accessed Aug. 10, 2017).
[87] "Driverless cars: everything you need to know about autonomous vehicles," *Auto Express*. http://www.autoexpress.co.uk/car-tech/85183/driverless-cars-everything-you-need-to-know-about-autonomous-vehicles (accessed Aug. 31, 2017).
[88] "Uber to Suspend Autonomous Tests After Arizona Accident," *Bloomberg.com*, Mar. 25, 2017. Accessed: Apr. 13, 2017. [Online]. Available: https://www.bloomberg.com/news/articles/2017-03-25/uber-autonomous-vehicle-gets-in-accident-in-tempe-arizona
[89] "Tesla Safety Update," Apr. 20, 2017. https://www.tesla.com/support/tesla-safety-update (accessed May 03, 2017).
[90] "Connected and Autonomous Vehicles: The future? - 115.pdf." https://publications.parliament.uk/pa/ld201617/ldselect/ldsctech/115/115.pdf (accessed Aug. 31, 2017).
[91] "Connected and Autonomous Vehicles – The UK Economic Opportunity - CRT036586F-Connected-and-Autonomous-Vehicles-–-The-UK-Economic-Opportu...1.pdf." https://www.smmt.co.uk/wp-content/uploads/sites/2/CRT036586F-Connected-and-Autonomous-Vehicles-%E2%80%93-The-UK-Economic-Opportu...1.pdf (accessed Aug. 31, 2017).
[92] admin, "Do Self-Driving Cars Have A Future In South Africa?," *VUT Research*, Oct. 17, 2016. http://www.vut-research.ac.za/do-self-driving-cars-have-a-future-in-south-africa/ (accessed Aug. 31, 2017).
[93] C. Bolchini, C. A. Curino, E. Quintarelli, F. A. Schreiber, and L. Tanca, "A data-oriented survey of context models," *ACM Sigmod Record*, vol. 36, no. 4, pp. 19–26, 2007.
[94] C. Bolchini, C. A. Curino, E. Quintarelli, F. A. Schreiber, and L. Tanca, "A data-oriented survey of context models," *ACM Sigmod Record*, vol. 36, no. 4, pp. 19–26, 2007.
[95] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *International Symposium on Handheld and Ubiquitous Computing*, 1999, pp. 304–307. Accessed: Oct. 10, 2016. [Online]. Available: http://link.springer.com/chapter/10.1007/3-540-48157-5_29
[96] N. A. Bradley and M. D. Dunlop, "Towards a Multidisciplinary Model of Context to Support Context-Aware Computing," p. 44.

[97] B. N. Schilit and M. M. Theimer, “Disseminating active map information to mobile hosts,” *IEEE Network*, vol. 8, no. 5, pp. 22–32, Sep. 1994, doi: 10.1109/65.313011.

[98] A. K. Dey, G. D. Abowd, and A. Wood, “CyberDesk: A framework for providing self-integrating context-aware services,” *Knowledge-Based Systems*, vol. 11, no. 1, pp. 3–13, 1998.

[99] “Enhanced Reality Fieldwork.” https://www.cs.kent.ac.uk/pubs/1998/616/content.html (accessed Oct. 10, 2016).

[100] M. G. Brown, “Supporting user mobility,” in *Mobile Communications*, Springer, 1996, pp. 69–77. Accessed: Oct. 10, 2016. [Online]. Available: http://link.springer.com/chapter/10.1007/978-0-387-34980-0_8

[101] M. G. Brown, “Supporting user mobility,” in *Mobile Communications*, Springer, 1996, pp. 69–77. Accessed: Oct. 10, 2016. [Online]. Available: http://link.springer.com/chapter/10.1007/978-0-387-34980-0_8

[102] B. Schilit, N. Adams, and R. Want, “Context-aware computing applications,” in , *Workshop on Mobile Computing Systems and Applications, 1994. Proceedings*, Dec. 1994, pp. 85–90. doi: 10.1109/MCSA.1994.512740.

[103] S. Loke, *Context-Aware Pervasive Systems: Architectures for a New Breed of Applications*. CRC Press, 2006.

[104] D. Preuveneers and Y. Berbers, “Internet of things: A context-awareness perspective,” *The Internet of Things: From RFID to the Next-Generation Pervasive Networked Systems*, pp. 287–307, 2008.

[105] “The Future of Information Security Is Context Aware and .” http://docplayer.net/3392728-The-future-of-information-security-is-context-aware-and-.html (accessed Oct. 10, 2016).

[106] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, “Towards a Better Understanding of Context and Context-Awareness,” in *Handheld and Ubiquitous Computing*, vol. 1707, H.-W. Gellersen, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 304–307. doi: 10.1007/3-540-48157-5_29.

[107] A. K. Dey, “Understanding and Using Context,” p. 4.

[108] A. Armand, D. Filliat, and J. Ibañez-Guzman, “Ontology-based context awareness for driving assistance systems,” in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, Jun. 2014, pp. 227–233. doi: 10.1109/IVS.2014.6856509.

[109] M. Erfani, “Towards an ontology-based context-aware meta-model for the software domain,” in *2015 IEEE 22nd International Conference on Software Analysis, Evolution*

and Reengineering (SANER), Mar. 2015, pp. 605–606. doi: 10.1109/SANER.2015.7081892.

[110] H. Chen, T. Finin, and A. Joshi, “An ontology for context-aware pervasive computing environments,” *The Knowledge Engineering Review*, vol. 18, no. 3, pp. 197–207, Sep. 2003, doi: 10.1017/S0269888904000025.

[111] H. Chen, T. Finin, and A. Joshi, “UMBC ebiquity,” 2004, Accessed: Nov. 09, 2016. [Online]. Available: http://ebiquity.umbc.edu/resource/html/id/75/An-Intelligent-Broker-for-Pervasive-Context-Aware-Systems

[112] C. Braunagel, W. Rosenstiel, and E. Kasneci, “Ready for Take-Over? A New Driver Assistance System for an Automated Classification of Driver Take-Over Readiness,” *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 4, pp. 10–22, 2017, doi: 10.1109/MITS.2017.2743165.

[113] C. Braunagel, D. Geisler, W. Rosenstiel, and E. Kasneci, “Online Recognition of Driver-Activity Based on Visual Scanpath Classification,” *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 4, pp. 23–36, 2017, doi: 10.1109/MITS.2017.2743171.

[114] M. Sameen and B. Pradhan, “Severity Prediction of Traffic Accidents with Recurrent Neural Networks,” *Applied Sciences*, vol. 7, no. 6, p. 476, Jun. 2017, doi: 10.3390/app7060476.

[115] S. Alkheder, M. Taamneh, and S. Taamneh, “Severity Prediction of Traffic Accident Using an Artificial Neural Network: Traffic Accident Severity Prediction Using Artificial Neural Network,” *J. Forecast.*, vol. 36, no. 1, pp. 100–108, Jan. 2017, doi: 10.1002/for.2425.

[116] C. Wang, M. A. Quddus, and S. G. Ison, “Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model,” *Accident Analysis & Prevention*, vol. 43, no. 6, pp. 1979–1990, Nov. 2011, doi: 10.1016/j.aap.2011.05.016.

[117] M. M. Kunt, I. Aghayan, and N. Noii, “PREDICTION FOR TRAFFIC ACCIDENT SEVERITY: COMPARING THE ARTIFICIAL NEURAL NETWORK, GENETIC ALGORITHM, COMBINED GENETIC ALGORITHM AND PATTERN SEARCH METHODS,” *TRANSPORT*, vol. 26, no. 4, pp. 353–366, Jan. 2012, doi: 10.3846/16484142.2011.635465.

[118] H. B. Kang, “Various Approaches for Driver and Driving Behavior Monitoring: A Review,” in *2013 IEEE International Conference on Computer Vision Workshops*, Dec. 2013, pp. 616–623. doi: 10.1109/ICCVW.2013.85.

[119] B. Chakraborty and K. Nakano, “Automatic detection of driver’s awareness with cognitive task from driving behavior,” Oct. 2016, pp. 003630–003633. doi: 10.1109/SMC.2016.7844797.

[120] C. Yan, H. Jiang, B. Zhang, and F. Coenen, “Recognizing driver inattention by convolutional neural networks,” Oct. 2015, pp. 680–685. doi: 10.1109/CISP.2015.7407964.

[121] T. H. N. Le, Y. Zheng, C. Zhu, K. Luu, and M. Savvides, “Multiple Scale Faster-RCNN Approach to Driver’s Cell-Phone Usage and Hands on Steering Wheel Detection,” Jun. 2016, pp. 46–53. doi: 10.1109/CVPRW.2016.13.

[122] J. Donahue *et al.*, “Long-term Recurrent Convolutional Networks for Visual Recognition and Description,” *arXiv:1411.4389 [cs]*, Nov. 2014, [Online]. Available: http://arxiv.org/abs/1411.4389

[123] J. Hu, L. Xu, X. He, and W. Meng, “Abnormal Driving Detection Based on Normalized Driving Behavior,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 6645–6652, Aug. 2017, doi: 10.1109/TVT.2017.2660497.

[124] E. Ohn-Bar, S. Martin, A. Tawari, and M. M. Trivedi, “Head, Eye, and Hand Patterns for Driver Activity Recognition,” in *2014 22nd International Conference on Pattern Recognition*, Aug. 2014, pp. 660–665. doi: 10.1109/ICPR.2014.124.

[125] T. Sato and M. Akamatsu, “Applying Fuzzy Logic to Understanding Driving Modes in Real Road Environments,” in *Fuzzy Logic - Tool for Getting Accurate Solutions*, E. P. Dadios, Ed. InTech, 2015. doi: 10.5772/59438.

[126] A. Aksjonov, P. Nedoma, V. Vodovozov, E. Petlenkov, and M. Herrmann, “A method of driver distraction evaluation using fuzzy logic: Phone usage as a driver’s secondary activity: Case study,” in *2017 XXVI International Conference on Information, Communication and Automation Technologies (ICAT)*, Sarajevo, Oct. 2017, pp. 1–6. doi: 10.1109/ICAT.2017.8171599.

[127] A. Aksjonov, P. Nedoma, V. Vodovozov, E. Petlenkov, and M. Herrmann, “A Novel Driver Performance Model Based on Machine Learning,” *IFAC-PapersOnLine*, vol. 51, no. 9, pp. 267–272, 2018, doi: 10.1016/j.ifacol.2018.07.044.

[128] A. Aksjonov, P. Nedoma, V. Vodovozov, E. Petlenkov, and M. Herrmann, “Detection and Evaluation of Driver Distraction Using Machine Learning and Fuzzy Logic,” *IEEE*

Trans. Intell. Transport. Syst., vol. 20, no. 6, pp. 2048–2059, Jun. 2019, doi: 10.1109/TITS.2018.2857222.

[129] H. M. Eraqi, Y. Abouelnaga, M. H. Saad, and M. N. Moustafa, “Driver Distraction Identification with an Ensemble of Convolutional Neural Networks,” *Journal of Advanced Transportation*, 2019. https://www.hindawi.com/journals/jat/2019/4125865/ (accessed Aug. 10, 2019).

[130] Y. Abouelnaga, H. M. Eraqi, and M. N. Moustafa, “Real-time Distracted Driver Posture Classification,” *arXiv:1706.09498 [cs]*, Jun. 2017, Accessed: Oct. 02, 2019. [Online]. Available: http://arxiv.org/abs/1706.09498

[131] F. Riaz, S. Khadim, R. Rauf, M. Ahmad, S. Jabbar, and J. Chaudhry, “A validated fuzzy logic inspired driver distraction evaluation system for road safety using artificial human driver emotion,” *Computer Networks*, vol. 143, pp. 62–73, Oct. 2018, doi: 10.1016/j.comnet.2018.06.007.

[132] B. B. Munyazikwiye, H. R. Karimi, and K. G. Robbersmyr, “Fuzzy logic approach to predict vehicle crash severity from acceleration data,” in *2015 International Conference on Fuzzy Theory and Its Applications (iFUZZY)*, Yilan, Taiwan, Nov. 2015, pp. 44–49. doi: 10.1109/iFUZZY.2015.7391892.

[133] S. M. Upadhya and V. Vinothina, “Fuzzy Logic Based Approach for Possibility of Road Accidents,” *J. Phys.: Conf. Ser.*, vol. 1427, p. 012011, Jan. 2020, doi: 10.1088/1742-6596/1427/1/012011.

[134] K. Kim, Y. Kim, and S. Kwak, “Vision Sensor Based Fuzzy System for Intelligent Vehicles,” *Sensors*, vol. 19, no. 4, p. 855, Feb. 2019, doi: 10.3390/s19040855.

[135] M. Salleh, N. Talpur, and K. Hussain, “ Neuro-Fuzzy Inference System: Overview, Strengths, Limitations, and Solutions,” Aug. 2017, pp. 527–535. doi: 10.1007/978-3-319-61845-6_52.

[136] C. Dobbins and S. Fairclough, “Detecting and Visualizing Context and Stress via a Fuzzy Rule-Based System During Commuter Driving,” in *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, Kyoto, Japan, Mar. 2019, pp. 499–504. doi: 10.1109/PERCOMW.2019.8730600.

[137] A. Ondogan and H. S. Yavuz, “Fuzzy Logic Based Cruise Control for Low-Speed Following,” in *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Ankara, Turkey, Oct. 2019, pp. 1–5. doi: 10.1109/ISMSIT.2019.8932776.

[138] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," *arXiv:1501.00092 [cs]*, Dec. 2014, Accessed: Nov. 20, 2018. [Online]. Available: http://arxiv.org/abs/1501.00092

[139] Z. Hu, N. Uchida, Y. Wang, and Y. Dong, "Face orientation estimation for driver monitoring with a single depth camera," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2015, pp. 958–963. doi: 10.1109/IVS.2015.7225808.

[140] K. Sato, M. Ito, H. Madokoro, and S. Kadowaki, "Driver body information analysis for distraction state detection," in *2015 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, Nov. 2015, pp. 13–18. doi: 10.1109/ICVES.2015.7396886.

[141] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Understanding Pedestrian Behavior in Complex Traffic Scenes," *IEEE Trans. Intell. Veh.*, vol. 3, no. 1, pp. 61–70, Mar. 2018, doi: 10.1109/TIV.2017.2788193.

[142] A. Fasanmade *et al.*, "A Fuzzy-Logic Approach to Dynamic Bayesian Severity Level Classification of Driver Distraction Using Image Recognition," *IEEE Access*, vol. 8, pp. 95197–95207, 2020, doi: 10.1109/ACCESS.2020.2994811.

[143] N. Das, E. Ohn-Bar, and M. M. Trivedi, "On Performance Evaluation of Driver Hand Detection Algorithms: Challenges, Dataset, and Metrics," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Sep. 2015, pp. 2953–2958. doi: 10.1109/ITSC.2015.473.

[144] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," in *Computer Vision – ECCV 2014*, vol. 8692, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 184–199. Accessed: Nov. 20, 2018. [Online]. Available: http://link.springer.com/10.1007/978-3-319-10593-2_13

[145] T. H. N. Le, C. Zhu, Y. Zheng, K. Luu, and M. Savvides, "Robust hand detection in Vehicles," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, Dec. 2016, pp. 573–578. doi: 10.1109/ICPR.2016.7899695.

[146] T. H. N. Le, K. G. Quach, C. Zhu, C. N. Duong, K. Luu, and M. Savvides, "Robust Hand Detection and Classification in Vehicles and in the Wild," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA, Jul. 2017, pp. 1203–1210. doi: 10.1109/CVPRW.2017.159.

[147] E. A. Taub, "2-Second Rule for Distracted Driving Can Mean Life or Death (Published 2018)," *The New York Times*, Sep. 27, 2018. Accessed: Dec. 18, 2020.

[Online]. Available: https://www.nytimes.com/2018/09/27/business/distracted-driving-auto-industry.html

[148] "Driver Eye-Tracking Research | Latest News | Peugeot UK." https://www.peugeot.co.uk/news/drivers-take-their-eyes-off-the-road-7-of-the-time/ (accessed Dec. 18, 2020).

[149] W. Yuan, Z. Liu, and R. Fu, "Predicting Drivers' Eyes-Off-Road Duration in Different Driving Scenarios," *Discrete Dynamics in Nature and Society*, Nov. 01, 2018. https://www.hindawi.com/journals/ddns/2018/3481628/ (accessed Jan. 22, 2021).

[150] B. G. Simons-Morton, F. Guo, S. G. Klauer, J. P. Ehsani, and A. K. Pradhan, "Keep Your Eyes on the Road: Young Driver Crash Risk Increases According to Duration of Distraction," *Journal of Adolescent Health*, vol. 54, no. 5, Supplement, pp. S61–S67, May 2014, doi: 10.1016/j.jadohealth.2013.11.021.

[151] P. Doshi, D. Kapur, R. Iyer, and A. Chatterjee, "Smart mobility: Algorithm for road and driver type determination," in *2017 IEEE Transportation Electrification Conference (ITEC-India)*, Dec. 2017, pp. 1–4. doi: 10.1109/ITEC-India.2017.8333895.

[152] C. Chai, R. Zhao, J. Shi, J. Zhou, and Z. Yang, "The effect of road environment and type of vehicles on road rage in China," in *2019 12th International Symposium on Computational Intelligence and Design (ISCID)*, Dec. 2019, vol. 2, pp. 200–203. doi: 10.1109/ISCID.2019.10129.

[153] A. Rakotonirainy, "Design of context-aware systems for vehicle using complex systems paradigms," 2005, Accessed: Oct. 28, 2016. [Online]. Available: http://eprints.qut.edu.au/2700

[154] I. Khan and S. Khusro, "Towards the Design of Context-Aware User Interfaces to Minimize Drivers' Distractions," *Mobile Information Systems*, Jun. 27, 2020. https://www.hindawi.com/journals/misy/2020/8858886/ (accessed Jul. 10, 2020).

[155] X. Jianfeng, G. Hongyu, T. Jian, L. Liu, and L. Haizhu, "A classification and recognition model for the severity of road traffic accident," *Advances in Mechanical Engineering*, vol. 11, no. 5, p. 168781401985189, May 2019, doi: 10.1177/1687814019851893.

[156] X. Cai, C. Wang, S. Chen, and J. Lu, "Model Development for Risk Assessment of Driving on Freeway under Rainy Weather Conditions," *PLoS ONE*, vol. 11, no. 2, p. e0149442, Feb. 2016, doi: 10.1371/journal.pone.0149442.

[157] F. Malin, I. Norros, and S. Innamaa, “Accident risk of road and weather conditions on different road types,” *Accident Analysis & Prevention*, vol. 122, pp. 181–188, Jan. 2019, doi: 10.1016/j.aap.2018.10.014.

[158] L. A. Sherretz and B. C. Farhar, “An Analysis of the Relationship Between Rainfall and the Occurrence Of Traffic Accidents,” *Journal of Applied Meteorology and Climatology*, vol. 17, no. 5, pp. 711–715, May 1978, doi: 10.1175/1520-0450(1978)017<0711:AAOTRB>2.0.CO;2.

[159] R. Bergel-Hayat, M. Debbarh, C. Antoniou, and G. Yannis, “Explaining the road accident risk: Weather effects,” *Accident Analysis & Prevention*, vol. 60, pp. 456–465, Nov. 2013, doi: 10.1016/j.aap.2013.03.006.

[160] H. Brodsky and A. S. Hakkert, “Risk of a road accident in rainy weather,” *Accident Analysis & Prevention*, vol. 20, no. 3, pp. 161–176, Jun. 1988, doi: 10.1016/0001-4575(88)90001-2.

[161] D. J. Andrey, “Weather Information and Road Safety,” p. 39.

[162] S. G. Stradling, “Car driver speed choice in Scotland,” *Ergonomics*, vol. 50, no. 8, pp. 1196–1208, Aug. 2007, doi: 10.1080/00140130701318681.

[163] F. Rosey and J.-M. Auberlet, “Trajectory variability: Road geometry difficulty indicator,” *Safety Science*, vol. 50, no. 9, pp. 1818–1828, Nov. 2012, doi: 10.1016/j.ssci.2012.04.003.

[164] C. Wu, D. Yu, A. Doherty, T. Zhang, L. Kust, and G. Luo, “An investigation of perceived vehicle speed from a driver’s perspective,” *PLoS ONE*, vol. 12, no. 10, p. e0185347, Oct. 2017, doi: 10.1371/journal.pone.0185347.

[165] E. de Bellis, M. Schulte-Mecklenbeck, W. Brucks, A. Herrmann, and R. Hertwig, “Blind haste: As light decreases, speeding increases,” *PLOS ONE*, vol. 13, no. 1, p. e0188951, Jan. 2018, doi: 10.1371/journal.pone.0188951.

[166] M. P. Manser and P. A. Hancock, “The influence of perceptual speed regulation on speed perception, choice, and control: Tunnel wall characteristics and influences,” *Accident Analysis & Prevention*, vol. 39, no. 1, pp. 69–78, Jan. 2007, doi: 10.1016/j.aap.2006.06.005.

[167] A. Mishra and P. Bajaj, “Driver’s Behaviour Monitoring on Urban Roads of a Tier 2 City in India,” in *2015 7th International Conference on Emerging Trends in Engineering Technology (ICETET)*, Nov. 2015, pp. 134–140. doi: 10.1109/ICETET.2015.13.

[168] A. Kumar and R. Patra, “Driver drowsiness monitoring system using visual behaviour and machine learning,” in *2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, Penang, Malaysia, Apr. 2018, pp. 339–344. doi: 10.1109/ISCAIE.2018.8405495.

[169] E. S. Lee and D. Kum, “Feature-based lateral position estimation of surrounding vehicles using stereo vision,” in *2017 IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA, USA, Jun. 2017, pp. 779–784. doi: 10.1109/IVS.2017.7995811.

[170] D. Xu, H. Zhao, F. Guillemard, S. Geronimi, and F. Aioun, “Scene-aware driver state understanding in car-following behaviors,” in *2017 IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA, USA, Jun. 2017, pp. 1490–1496. doi: 10.1109/IVS.2017.7995920.

[171] U. Mittal, R. Potnuru, and P. Chawla, “Vehicle Detection and Classification using Improved Faster Region Based Convolution Neural Network,” in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Jun. 2020, pp. 511–514. doi: 10.1109/ICRITO48877.2020.9197805.

[172] J. Gong, J. Zhao, F. Li, and H. Zhang, “Vehicle detection in thermal images with an improved yolov3-tiny,” in *2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*, Jul. 2020, pp. 253–256. doi: 10.1109/ICPICS50287.2020.9201995.

[173] R. A. Kharjul, V. K. Tungar, Y. P. Kulkarni, S. K. Upadhyay, and R. Shirsath, “Real-time pedestrian detection using SVM and AdaBoost,” in *2015 International Conference on Energy Systems and Applications*, Oct. 2015, pp. 740–743. doi: 10.1109/ICESA.2015.7503447.

[174] K. Tateiwa and K. Yamada, “Estimating driver awareness of pedestrians in crosswalk in the path of right or left turns at an intersection from vehicle behavior,” in *2015 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2015, pp. 952–957. doi: 10.1109/IVS.2015.7225807.

[175] K. Tateiwa, A. Nakamura, and K. Yamada, “Study on estimating driver awareness of pedestrians while turning right at intersection based on vehicle behavior utilizing driving simulator,” in *2016 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2016, pp. 388–393. doi: 10.1109/IVS.2016.7535415.

[176] A. Rangesh, E. Ohn-Bar, K. Yuen, and M. M. Trivedi, “Pedestrians and their phones-detecting phone-based activities of pedestrians for autonomous vehicles,” in *Intelligent*

Transportation Systems (ITSC), 2016 IEEE 19th International Conference on, 2016, pp. 1882–1887.

[177] M. T. Phan, V. Fremont, I. Thouvenin, M. Sallak, and V. Cherfaoui, “Estimation of driver awareness of pedestrian based on Hidden Markov Model,” in *2015 IEEE Intelligent Vehicles Symposium (IV),* Seoul, South Korea, Jun. 2015, pp. 970–975. doi: 10.1109/IVS.2015.7225810.

[178] D. D. Clarke, P. Ward, C. Bartle, and W. Truman, “Young driver accidents in the UK: The influence of age, experience, and time of day,” *Accident Analysis & Prevention*, vol. 38, no. 5, pp. 871–878, Sep. 2006, doi: 10.1016/j.aap.2006.02.013.

[179] A. F. Williams, “Teenage drivers: patterns of risk,” *Journal of Safety Research*, vol. 34, no. 1, pp. 5–15, Jan. 2003, doi: 10.1016/S0022-4375(02)00075-0.

[180] S. of Psychology, “Think talking on your hands-free while driving is safe? Think again, says new research,” *The University of Sussex*. http://www.sussex.ac.uk/broadcast/read/35831 (accessed Oct. 03, 2019).

[181] A. O. Ferdinand and N. Menachemi, “Associations Between Driving Performance and Engaging in Secondary Tasks: A Systematic Review,” *Am J Public Health*, vol. 104, no. 3, pp. e39–e48, Mar. 2014, doi: 10.2105/AJPH.2013.301750.

[182] R. D. Foss and A. H. Goodwin, “Distracted Driver Behaviors and Distracting Conditions Among Adolescent Drivers: Findings From a Naturalistic Driving Study,” *Journal of Adolescent Health*, vol. 54, no. 5, Supplement, pp. S50–S60, May 2014, doi: 10.1016/j.jadohealth.2014.01.005.

[183] andrew.currin.ctr@dot.gov, “Distracted Driving,” *NHTSA*, Sep. 08, 2016. https://www.nhtsa.gov/risky-driving/distracted-driving (accessed Jun. 23, 2019).

[184] 1615 L. St NW, Suite 800Washington, and D. 20036USA202-419-4300 | M.-857-8562 | F.-419-4372 | M. Inquiries, “Teens and Distracted Driving,” *Pew Research Center: Internet, Science & Tech*, Nov. 16, 2009. https://www.pewresearch.org/internet/2009/11/16/teens-and-distracted-driving/ (accessed Jan. 07, 2020).

[185] 1615 L. St NW, Suite 800Washington, and D. 20036USA202-419-4300 | M.-857-8562 | F.-419-4372 | M. Inquiries, “Teens and Mobile Phones,” *Pew Research Center: Internet, Science & Tech*, Apr. 20, 2010. https://www.pewresearch.org/internet/2010/04/20/teens-and-mobile-phones/ (accessed Jan. 07, 2020).

[186] F. A. Drews, M. Pasupathi, and D. L. Strayer, “Passenger and cell phone conversations in simulated driving.,” *Journal of Experimental Psychology: Applied*, vol. 14, no. 4, pp. 392–400, 2008, doi: 10.1037/a0013119.

[187] H. L. THANH, “Distraction,” *Mobility and transport - European Commission*, Dec. 16, 2016. https://ec.europa.eu/transport/road_safety/users/young-people/distraction_en (accessed Oct. 03, 2019).

[188] A. S. Kulkarni and S. B. Shinde, “A review paper on monitoring driver distraction in real time using computer vision system,” in *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, Apr. 2017, pp. 1–4. doi: 10.1109/ICEICE.2017.8191851.

[189] “Reported Road Casualties Great Britain: 2017,” p. 376.

[190] “Facts on Young Car Drivers,” p. 11.

[191] D. Crundall, “Hazard prediction discriminates between novice and experienced drivers,” *Accident Analysis & Prevention*, vol. 86, pp. 47–58, Jan. 2016, doi: 10.1016/j.aap.2015.10.006.

[192] N. R. Council (US), I. of Medicine (US), and and T. R. B. (US) P. C. for a W. on C. from the B. and S. S. in R. and P. T. M. Crashes, *The Anatomy of Crashes Involving Young Drivers*. National Academies Press (US), 2007. Accessed: Dec. 13, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK9672/

[193] E. Delbridge, “Causes and Effects of Distracted Driving,” *The Balance*. https://www.thebalance.com/causes-and-consequences-of-distracted-driving-527085 (accessed Oct. 03, 2019).

[194] “The Battle for Attention: Driver Distraction – A Review of Recent Research and Knowledge | Blurbs New | Blurbs | Main.” http://www.trb.org/Main/Blurbs/173581.aspx (accessed Dec. 21, 2020).

[195] M. A. Kamrin, *Reporting on risk : a journalist’s handbook on environmental risk assessment*. Ann Arbor, MI : Produced by Foundation for American Communications and National Sea Grant College Program, 1995. Accessed: Dec. 14, 2020. [Online]. Available: http://archive.org/details/reportingonriskj0000kamr

[196] K. Berdica, “An introduction to road vulnerability: What has been done, is done and should be done,” *Transport Policy*, vol. 9, pp. 117–127, Apr. 2002, doi: 10.1016/S0967-070X(02)00011-2.

[197] K. Berdica and L.-G. Mattsson, “Vulnerability: A Model-Based Case Study of the Road Network in Stockholm,” in *Critical Infrastructure: Reliability and Vulnerability*,

A. T. Murray and T. H. Grubesic, Eds. Berlin, Heidelberg: Springer, 2007, pp. 81–106. doi: 10.1007/978-3-540-68056-7_5.

[198] W. M. ElDessouki, J. N. Ivan, E. N. Anagnostou, A. W. Sadek, and C. Zhang, "USING RELATIVE RISK ANALYSIS TO IMPROVE CONNECTICUT FREEWAY TRAFFIC SAFETY UNDER ADVERSE WEATHER CONDITIONS," p. 43.

[199] H. Ginting, G. Näring, and E. S. Becker, "Attentional bias and anxiety in individuals with coronary heart disease," *Psychology & Health*, vol. 28, no. 11, pp. 1306–1322, 2013, doi: 10.1080/08870446.2013.803554.

[200] O. Lopez-Fernandez, M. Freixa-Blanxart, and M. Honrubia-Serrano, "The Problematic Internet Entertainment Use Scale for Adolescents: Prevalence of Problem Internet Use in Spanish High School Students," *Cyberpsychology, behavior and social networking*, vol. 16, Dec. 2012, doi: 10.1089/cyber.2012.0250.

[201] J. Hartley, "Some thoughts on Likert-type scales," *International Journal of Clinical and Health Psychology*, vol. 14, no. 1, pp. 83–86, Jan. 2014, doi: 10.1016/S1697-2600(14)70040-7.

[202] R. Likert, "A technique for the measurement of attitudes," *Archives of Psychology*, vol. 22 140, pp. 55–55, 1932.

[203] C. León-Mantero, J. C. Casas-Rosal, C. Pedrosa-Jesús, and A. Maz-Machado, "Measuring attitude towards mathematics using Likert scale surveys: The weighted average," *PLOS ONE*, vol. 15, no. 10, p. e0239626, Oct. 2020, doi: 10.1371/journal.pone.0239626.

[204] J. Jonker and B. W. Pennink, "The Essence of Methodology," in *The Essence of Research Methodology: A Concise Guide for Master and PhD Students in Management Science*, J. Jonker and B. Pennink, Eds. Berlin, Heidelberg: Springer, 2009, pp. 21–41. doi: 10.1007/978-3-540-71659-4_2.

[205] C. Ezeah and E. Simbabure, "Perception-Based Analysis (PBA) of Climate Change Impacts on the Forest and Agricultural Ecosystem of Shropshire, United Kingdom," *Journal of Geoscience and Environment Protection*, vol. 3, no. 5, Art. no. 5, Jul. 2015, doi: 10.4236/gep.2015.35013.

[206] D. Wahyuni, "The Research Design Maze: Understanding Paradigms, Cases, Methods and Methodologies," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 2103082, Jun. 2012. Accessed: Sep. 30, 2021. [Online]. Available: https://papers.ssrn.com/abstract=2103082

[207] J. Dudovsky, “Research Approach - Research-Methodology,” 2011. http://research-methodology.net/research-methodology/research-approach/ (accessed Sep. 01, 2016).

[208] S. L. McGregor and J. A. Murnane, “Paradigm, methodology and method: Intellectual integrity in consumer scholarship,” *International Journal of Consumer Studies*, vol. 34, no. 4, pp. 419–427, 2010.

[209] Z. Zainal, “Case Study As a Research Method,” *undefined*, 2007, Accessed: Oct. 01, 2021. [Online]. Available: https://www.semanticscholar.org/paper/Case-Study-As-a-Research-Method-Zainal/8a545bb8172b9d0de6ab8068ab63aefdfc76555d

[210] “TeleFOT - UK DFOT3 - FOT-Net WIKI.” http://wiki.fot-net.eu/index.php/TeleFOT_-_UK_DFOT3 (accessed Jan. 04, 2017).

[211] W. Liang *et al.*, “Vehicular Ad Hoc Networks: Architectures, Research Issues, Methodologies, Challenges, and Trends, Vehicular Ad Hoc Networks: Architectures, Research Issues, Methodologies, Challenges, and Trends,” *International Journal of Distributed Sensor Networks, International Journal of Distributed Sensor Networks*, vol. 2015, 2015, p. e745303, Aug. 2015, doi: 10.1155/2015/745303, 10.1155/2015/745303.

[212] S. E. R. Franzén, M. Karlsson, A. Morris, and R. Welsh, “Widening the use of the FOT methodology. Development based on experiences from the TeleFOT project,” in *Chalmers Publication Library (CPL)*, 2012, vol. 48, pp. 1826–1836. doi: 10.1016/j.sbspro.2012.06.1157.

[213] “Home - UDRIVE.” http://www.udrive.eu/ (accessed Nov. 21, 2016).

[214] “Research Design in Social Research - David de Vaus - Google Books.” https://books.google.co.uk/books/about/Research_Design_in_Social_Research.html?id=9yurQt7T65oC&redir_esc=y (accessed Oct. 01, 2021).

[215] R. C. M. Philip, M. R. Dauvermann, H. C. Whalley, K. Baynham, S. M. Lawrie, and A. C. Stanfield, “A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders,” *Neuroscience and Biobehavioral Reviews*, vol. 36, no. 2, pp. 901–942, Feb. 2012, doi: 10.1016/j.neubiorev.2011.10.008.

[216] S. C. Akagha, “Nwankwo et al”, Accessed: Oct. 01, 2021. [Online]. Available: https://www.academia.edu/4397148/Nwankwo_et_al

[217] Y. Xing, C. Lv, H. Wang, D. Cao, E. Velenis, and F. Wang, “Driver Activity Recognition for Intelligent Vehicles: A Deep Learning Approach,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5379–5390, Jun. 2019, doi: 10.1109/TVT.2019.2908425.

[218] "TeleFOT - FOT-Net WIKI." http://wiki.fot-net.eu/index.php/TeleFOT (accessed Jan. 04, 2017).

[219] K. Ekambaram, J. Lenard, and S. Reed, "Digital and video analysis of eye-glance movements during naturalistic driving from the ADSEAT and TeleFOT field operational trials-results and challenges," 2016, Accessed: Jan. 31, 2017. [Online]. Available: https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/21623

[220] W. Ciscal-Terry, M. Dell'Amico, N. S. Hadjidimitriou, and M. Iori, "An analysis of drivers route choice behaviour using GPS data and optimal alternatives," *Journal of Transport Geography*, vol. 51, pp. 119–129, Feb. 2016, doi: 10.1016/j.jtrangeo.2015.12.003.

[221] A. Morris *et al.*, "Summary of final results from WP4. 3 through to WP4. 7 as at M54: TeleFOT [Field Operational Tests of Aftermarket and Nomadic Devices in Vehicles] D4. 2.3 b," 2011, Accessed: Jan. 31, 2017. [Online]. Available: https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/12269

[222] S. E. R. Franzén, M. Karlsson, A. P. Morris, and R. Welsh, "Widening the Use of the FOT Methodology Development Based on Experiences from the TeleFOT Project," *Procedia - Social and Behavioral Sciences*, vol. 48, pp. 1826–1836, 2012, doi: 10.1016/j.sbspro.2012.06.1157.

[223] S. Reed and A. Morris, "Building the European Road Safety Observatory. SafetyNet. Deliverable 5.5, Glossary of data variables for fatal and accident causation databases," 2008, Accessed: Jan. 27, 2017. [Online]. Available: https://dspace.lboro.ac.uk/dspace/handle/2134/4372

[224] "Wierwille and Connor 1983 collected 20 workload measures while pilots performed | Course Hero." https://www.coursehero.com/file/p691o72/Wierwille-and-Connor-1983-collected-20-workload-measures-while-pilots-performed/ (accessed Oct. 01, 2021).

[225] A. P. Smith, "A UK survey of driving behaviour, fatigue, risk taking and road traffic accidents," *BMJ Open*, vol. 6, no. 8, p. e011461, Aug. 2016, doi: 10.1136/bmjopen-2016-011461.

[226] H. Nabi, S. M. Consoli, J.-F. Chastang, M. Chiron, S. Lafont, and E. Lagarde, "Type A Behavior Pattern, Risky Driving Behaviors, and Serious Road Traffic Accidents: A Prospective Study of the GAZEL Cohort," *Am J Epidemiol*, vol. 161, no. 9, pp. 864–870, May 2005, doi: 10.1093/aje/kwi110.

[227] B. G. Simons-Morton *et al.*, “The Effect of Passengers and Risk-Taking Friends on Risky Driving and Crashes/Near Crashes Among Novice Teenagers,” *J Adolesc Health*, vol. 49, no. 6, pp. 587–593, Dec. 2011, doi: 10.1016/j.jadohealth.2011.02.009.

[228] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT Press, 2016. Accessed: Jun. 21, 2017. [Online]. Available: http://books.google.com/books?hl=en&lr=&id=omivDQAAQBAJ&oi=fnd&pg=PR5&dq=%22Estimators,+Bias+and+Variance+.+.+.+.+.%22+%22Example:+Learning+XOR+.%22+&ots=MLV6arpANU&sig=gGjm_TTGBFADY7Y9Q93DtBapwoY

[229] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[230] W. Dong, J. Li, R. Yao, C. Li, T. Yuan, and L. Wang, “Characterizing Driving Styles with Deep Learning,” *arXiv preprint arXiv:1607.03611*, 2016, Accessed: Jan. 27, 2017. [Online]. Available: https://arxiv.org/abs/1607.03611

[231] G. E. Hinton, S. Osindero, and Y.-W. Teh, “A fast learning algorithm for deep belief nets,” *Neural computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[232] F. Riaz, S. Khadim, R. Rauf, M. Ahmad, S. Jabbar, and J. Chaudhry, “A validated fuzzy logic inspired driver distraction evaluation system for road safety using artificial human driver emotion,” *Computer Networks*, vol. 143, pp. 62–73, Oct. 2018, doi: 10.1016/j.comnet.2018.06.007.

[233] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778. Accessed: May 03, 2017. [Online]. Available: http://www.cv-foundation.org/openaccess/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

[234] C. Szegedy *et al.*, “Going deeper with convolutions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9. Accessed: Jun. 21, 2017. [Online]. Available: http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Szegedy_Going_Deeper_With_2015_CVPR_paper.html

[235] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105. Accessed: Jan. 18, 2017. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-w

[236] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[237] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT Press, 2016. Accessed: Jun. 21, 2017. [Online]. Available: http://books.google.com/books?hl=en&lr=&id=omivDQAAQBAJ&oi=fnd&pg=PR5&dq=%22Estimators,+Bias+and+Variance+.+.+.+.+.%22+%22Example:+Learning+XOR+.%22+&ots=MLV6arpANU&sig=gGjm_TTGBFADY7Y9Q93DtBapwoY

[238] C. Yan, H. Jiang, B. Zhang, and F. Coenen, “Recognizing driver inattention by convolutional neural networks,” in *2015 8th International Congress on Image and Signal Processing (CISP)*, Oct. 2015, pp. 680–685. doi: 10.1109/CISP.2015.7407964.

[239] K. Yuen, S. Martin, and M. M. Trivedi, “Looking at faces in a vehicle: A deep CNN based approach and evaluation,” Nov. 2016, pp. 649–654. doi: 10.1109/ITSC.2016.7795622.

[240] S. Muhlbacher-Karrer *et al.*, “A Driver State Detection System—Combining a Capacitive Hand Detection Sensor With Physiological Sensors,” *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 4, pp. 624–636, Apr. 2017, doi: 10.1109/TIM.2016.2640458.

[241] S. Vora, A. Rangesh, and M. M. Trivedi, “On generalizing driver gaze zone estimation using convolutional neural networks,” Jun. 2017, pp. 849–854. doi: 10.1109/IVS.2017.7995822.

[242] V. Neale, S. Klauer, T. Dingus, J. Sudweeks, and M. Goodman, “An Overview of the 100-Car Naturalistic Study and Findings,” p. 10.

[243] H. Eraqi, “Distracted Driver Dataset.” http://heshameraqi.github.io/distraction_detection (accessed Feb. 20, 2020).

[244] R. Likert, “A technique for the measurement of attitudes,” *Archives of Psychology*, vol. 22 140, pp. 55–55, 1932.

[245] R. P. L. Buse and W. R. Weimer, “A metric for software readability,” in *Proceedings of the 2008 international symposium on Software testing and analysis - ISSTA ’08*, Seattle, WA, USA, 2008, p. 121. doi: 10.1145/1390630.1390647.

[246] “Draft guidance on charging offences arising from driving incidents.” https://www.cps.gov.uk/consultations/draft_driving_2012_consultation.html#a25 (accessed Jan. 28, 2017).

[247] "ADAS | Mobility and transport." https://ec.europa.eu/transport/road_safety/users/young-people/distraction (accessed May 25, 2021).

[248] A. S. Kulkarni and S. B. Shinde, "A review paper on monitoring driver distraction in real time using computer vision system," in *Proceedings - 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering, ICEICE 2017*, 2017, vol. 2017-Decem, pp. 1–4. doi: 10.1109/ICEICE.2017.8191851.

[249] R. R. Herrera, C. Holloway, D. Z. Morgado Ramirez, B. Zhang, and Y. Cho, "Breathing Biofeedback Relaxation Intervention for Wheelchair Users in City Navigation," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2020, vol. 2020-July, pp. 4575–4578. doi: 10.1109/EMBC44109.2020.9176144.

[250] P. Ventsislavova and D. Crundall, "The hazard prediction test: A comparison of free-response and multiple-choice formats," *Safety Science*, vol. 109, pp. 246–255, 2018, doi: 10.1016/j.ssci.2018.06.004.

[251] T. M. Senserrick, M. J. Kallan, and F. K. Winston, "Child passenger injury risk in sibling versus non-sibling teen driver crashes: A US study," *Injury Prevention*, vol. 13, no. 3, pp. 207–210, Jun. 2007, doi: 10.1136/ip.2006.014332.

[252] C. A. R. Insurance, C. A. R. I. Claims, and B. Y. E. Delbridge, "Causes and Effects of Distracted Driving," pp. 1–10, 2020.

[253] N. Kinnear and A. Stevens, "The battle for attention Driver distraction – a review of recent research and knowledge," 2015.

[254] C. Li, S. H. Chan, and Y. T. Chen, "Who make drivers stop? Towards driver-centric risk assessment: Risk object identification via causal inference," in *IEEE International Conference on Intelligent Robots and Systems*, Oct. 2020, pp. 10711–10718. doi: 10.1109/IROS45743.2020.9341072.

[255] J. Michaels *et al.*, "Driving simulator scenarios and measures to faithfully evaluate risky driving behavior: A comparative study of different driver age groups," *PLoS ONE*, vol. 12, no. 10, Oct. 2017, doi: 10.1371/journal.pone.0185909.

[256] J. Liu, F. Jin, Q. Xie, and M. Skitmore, "Improving risk assessment in financial feasibility of international engineering projects: A risk driver perspective," *International Journal of Project Management*, vol. 35, no. 2, pp. 204–211, 2017, doi: 10.1016/j.ijproman.2016.11.004.

[257] J. L. Yin and B. H. Chen, "An Advanced Driver Risk Measurement System for Usage-Based Insurance on Big Driving Data," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 4, pp. 585–594, 2018, doi: 10.1109/TIV.2018.2874530.

[258] H. Mao, F. Guo, X. Deng, and Z. R. Doerzaph, "Decision-adjusted driver risk predictive models using kinematics information," *Accident Analysis and Prevention*, vol. 156, 2021, doi: 10.1016/j.aap.2021.106088.

[259] A. Tefferi and A. M. Vannucchi, "Genetic Risk Assessment in Myeloproliferative Neoplasms," *Mayo Clinic Proceedings*, vol. 92, no. 8. pp. 1283–1290, 2017. doi: 10.1016/j.mayocp.2017.06.002.

[260] J. D. Paradis, "Reporting on risk: A journalist's handbook on environmental risk assessment," *Journal of Criminal Justice*, vol. 23, no. 4, p. 391, 1995, doi: 10.1016/0047-2352(95)90043-8.

[261] K. Berdica, "An introduction to road vulnerability: What has been done, is done and should be done," *Transport Policy*, vol. 9, no. 2, pp. 117–127, 2002, doi: 10.1016/S0967-070X(02)00011-2.

[262] K. Berdica and L. G. Mattsson, "Vulnerability: A model-based case study of the road network in Stockholm," in *Advances in Spatial Science*, no. 9783540680550, Springer International Publishing, 2007, pp. 81–106. doi: 10.1007/978-3-540-68056-7_5.

[263] W. ElDessouki, J. Ivan, and E. Anagnostou, "Using relative risk analysis to improve connecticut freeway traffic safety under adverse weather conditions," 2004.

[264] X. Cai, C. Wang, S. Chen, and J. Lu, "Model development for risk assessment of driving on freeway under rainy weather conditions," *PLoS ONE*, vol. 11, no. 2, Feb. 2016, doi: 10.1371/journal.pone.0149442.

[265] D. Sommer and M. Golz, "Evaluation of PERCLOS based current fatigue monitoring technologies," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, 2010, pp. 4456–4459. doi: 10.1109/IEMBS.2010.5625960.

[266] H. Ginting, G. Näring, and E. S. Becker, "Attentional bias and anxiety in individuals with coronary heart disease," *Psychology and Health*, vol. 28, no. 11, pp. 1306–1322, Nov. 2013, doi: 10.1080/08870446.2013.803554.

[267] O. Lopez-Fernandez, M. Freixa-Blanxart, and M. L. Honrubia-Serrano, "The problematic internet entertainment use scale for adolescents: Prevalence of problem internet use in Spanish high school students," *Cyberpsychology, Behavior, and Social Networking*, vol. 16, no. 2, pp. 108–118, Feb. 2013, doi: 10.1089/cyber.2012.0250.

[268] G. B. Casal, J. Carlos, S. Normas, P. La, and D. E. C. Cl, “International Journal of Clinical and Health Psychology,” *International Journal of Clinical and Health Psychology*, vol. 2, no. 1. pp. 525–532, 2002.

[269] T. C. CHIVERS, W. J. ROGERS, and M. E. WILLIAMS, “A TECHNIQUE FOR THE MEASUREMENT OF GAS-LEAKAGE.,” *psycnet.apa.org*. 1974.

[270] C. León-Mantero, J. C. Casas-Rosal, C. Pedrosa-Jesús, and A. Maz-Machado, “Measuring attitude towards mathematics using Likert scale surveys: The weighted average,” *PLoS ONE*, vol. 15, no. 10 October, Oct. 2020, doi: 10.1371/journal.pone.0239626.

[271] J. F. Liu, Y. F. Su, M. K. Ko, and P. N. Yu, “Development of a vision-based driver assistance system with lane departure warning and forward collision warning functions,” in *Proceedings - Digital Image Computing: Techniques and Applications, DICTA 2008*, 2008, pp. 480–485. doi: 10.1109/DICTA.2008.78.

[272] “TeleFOT - UK DFOT3 - FOT-Net WIKI.” https://wiki.fot-net.eu/index.php/TeleFOT_-_UK_DFOT3 (accessed May 25, 2021).

[273] M. M. Kunt, I. Aghayan, and N. Noii, “Prediction for traffic accident severity: Comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods,” *Transport*, vol. 26, no. 4, pp. 353–366, Dec. 2011, doi: 10.3846/16484142.2011.635465.

[274] J. Morton, T. A. Wheeler, and M. J. Kochenderfer, “Analysis of Recurrent Neural Networks for Probabilistic Modeling of Driver Behavior,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1289–1298, 2017, doi: 10.1109/TITS.2016.2603007.

[275] M. Zhu, Y. Li, and Y. Wang, “Design and experiment verification of a novel analysis framework for recognition of driver injury patterns: From a multi-class classification perspective,” *Accident Analysis and Prevention*, vol. 120, pp. 152–164, Nov. 2018, doi: 10.1016/j.aap.2018.08.011.

[276] Y. Ma, G. Gu, B. Yin, S. Qi, K. Chen, and C. Chan, “Support vector machines for the identification of real-time driving distraction using in-vehicle information systems,” *Journal of Transportation Safety and Security*, 2020, doi: 10.1080/19439962.2020.1774019.

[277] T. Liu, Y. Yang, G. Bin Huang, Y. K. Yeo, and Z. Lin, “Driver Distraction Detection Using Semi-Supervised Machine Learning,” *IEEE Transactions on Intelligent*

Transportation Systems, vol. 17, no. 4, pp. 1108–1120, Apr. 2016, doi: 10.1109/TITS.2015.2496157.

[278] "TeleFOT - FOT-Net WIKI." http://wiki.fot-net.eu/index.php/TeleFOT (accessed Jan. 04, 2017).

[279] "TeleFOT - UK DFOT3 - FOT-Net WIKI." http://wiki.fot-net.eu/index.php/TeleFOT_-_UK_DFOT3 (accessed Jan. 04, 2017).

[280] C. C. Aggarwal, *Neural Networks and Deep Learning: A Textbook*. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-94463-0.

[281] S. Kouchak, A. G.-I. T. on Intelligent, and undefined 2020, "Detecting Driver Behavior Using Stacked Long Short Term Memory Network With Attention Layer," *ieeexplore.ieee.org*.

[282] J. Chen *et al.*, "Fine-Grained Detection of Driver Distraction Based on Neural Architecture Search," *IEEE Transactions on Intelligent Transportation Systems*, 2021, doi: 10.1109/TITS.2021.3055545.

[283] M. Wöllmer *et al.*, "Online Driver Distraction Detection Using Long Short-Term Memory," *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, vol. 12, no. 2, 2011, doi: 10.1109/TITS.2011.2119483.

[284] O. Olabiyi, E. Martinson, V. Chintalapudi, and R. Guo, "Driver action prediction using deep (bidirectional) recurrent neural network," *arXiv*. arXiv, Jun. 07, 2017.

[285] A. Aksjonov *et al.*, "A Method of Driver Distraction Evaluation Using Fuzzy Logic A Method of Driver Distraction Evaluation Using Fuzzy Logic Phone usage as a driver's secondary activity: Case study," 2017.

[286] F. Riaz, S. Khadim, R. Rauf, M. Ahmad, S. Jabbar, and J. Chaudhry, "A validated fuzzy logic inspired driver distraction evaluation system for road safety using artificial human driver emotion," *Computer Networks*, vol. 143, pp. 62–73, Oct. 2018, doi: 10.1016/j.comnet.2018.06.007.

[287] A. Aksjonov, V. Vodovozov, and E. Petlenkov, "Detection and Evaluation of Driver Distraction Using Machine Learning and Fuzzy Logic Re:creation: Virtual and Augmented Reality Laboratory View project ITEAM-Interdisciplinary Training Network in Multi-Actuated Ground Vehicles View project," *ieeexplore.ieee.org*, doi: 10.1109/TITS.2018.2857222.

[288] C. Ou, C. Ouali, S. M. Bedawi, and F. Karray, "Driver behavior monitoring using tools of deep learning and fuzzy inferencing," in *IEEE International Conference on Fuzzy Systems*, Oct. 2018, vol. 2018-July. doi: 10.1109/FUZZ-IEEE.2018.8491511.

[289] L. Li, B. Zhong, C. H. Jr, … Y. L.-A. A. &, and undefined 2020, "Detection of driver manual distraction via image-based hand and ear recognition," *Elsevier*.

[290] A. Koesdwiady, S. M. Bedawi, C. Ou, and F. Karray, "End-to-end deep learning for driver distraction recognition," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10317 LNCS, pp. 11–18. doi: 10.1007/978-3-319-59876-5_2.

[291] Y. Xing, C. Lv, H. Wang, … D. C.-I. transactions on, and undefined 2019, "Driver activity recognition for intelligent vehicles: A deep learning approach," *ieeexplore.ieee.org*.

[292] D. Tran, H. M. Do, W. Sheng, H. Bai, and G. Chowdhary, "Real-time detection of distracted driving based on deep learning; Real-time detection of distracted driving based on deep learning," *IET*, vol. 12, no. 10, pp. 1210–1219, Dec. 2018, doi: 10.1049/iet-its.2018.5172.

[293] H. Almahasneh, W.-T. Chooi, N. Kamel, and A. S. Malik, "Deep in thought while driving: An EEG study on drivers' cognitive distraction," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 26, pp. 218–226, Sep. 2014, doi: 10.1016/j.trf.2014.08.001.

[294] "Virtual-Based Safety Testing for Self-Driving Cars from NVIDIA DRIVE Constellation.," *@NVIDIA*. https://www.nvidia.com/en-gb/self-driving-cars/drive-constellation/ (accessed Jul. 25, 2020).

[295] "NYPD vehicle rams Brooklyn crowd during protest," *The Independent*, May 31, 2020. https://www.independent.co.uk/news/world/americas/brooklyn-new-york-george-floyd-protests-nypd-vehicle-a9540871.html (accessed Jul. 26, 2020).

APPENDIX

Appendix 1: Detect and Track Faces

```
%% detectAndTrackFaces
% Automatically detects and tracks multiple faces in a webcam-acquired
% video stream.



%% Instantiate video device, face detector, and KLT object tracker
%vidObj = webcam;
vidObj= VideoReader('C:\Users\bami\Documents\MATLAB\Dataset\2\28205-28467.avi');
faceDetector = vision.CascadeObjectDetector('EyePairBig'); % Finds faces by default
%eyeDetector = vision.CascadeObjectDetector; % Finds faces by default
tracker = MultiObjectTrackerKLT;

%% Get a frame for frame-size information
frame = read(vidObj,1);
frameSize = size(frame);

%% Create a video player instance
videoPlayer  = vision.VideoPlayer('Position',[240 200 fliplr(frameSize(1:2)+30)]);

%% Iterate until we have successfully detected a face
bboxes = [];
%bbox1=[];
while isempty(bboxes)&& hasFrame(vidObj)
    framergb = readFrame(vidObj);
    frame = rgb2gray(framergb);
    bboxes = faceDetector.step(frame);
    %bbox1 = eyeDetector.step(frame);
end
```

```
tracker.addDetections(frame, bboxes);
%tracker.addDetections(frame, bbox1);

%% And loop until the player is closed
frameNumber = 0;

disp('Press Ctrl-C to exit...');
while hasFrame(vidObj)

    framergb = readFrame(vidObj);
    frame = rgb2gray(framergb);

    if mod(frameNumber, 10) == 0
        % (Re)detect faces.
        %
        % NOTE: face detection is more expensive than imresize; we can
        % speed up the implementation by reacquiring faces using a
        % downsampled frame:
        % bboxes = faceDetector.step(frame);
        bboxes = 2 * faceDetector.step(imresize(frame, 0.5));
        if ~isempty(bboxes)
            tracker.addDetections(frame, bboxes);
            %tracker1.addDetections(frame, bbox1);
        end
    else
        % Track faces
        tracker.track(frame);

    end

    % Display bounding boxes and tracked points.
    displayFrame = insertObjectAnnotation(framergb, 'rectangle',...
        tracker.Bboxes, tracker.BoxIds);
    displayFrame = insertMarker(displayFrame, tracker.Points);
```

```
        videoPlayer.step(displayFrame);

        frameNumber = frameNumber + 1;
    end

    %% Clean up
    release(videoPlayer);
```

Appendix 2: Deep Segmentation

```
%% CNN
resnet18();
%% Load Pre-trained Data
pretrainedURL =
'https://www.mathworks.com/supportfiles/vision/data/deeplabv3plusResnet18CamVid.mat';
pretrainedFolder = fullfile(tempdir,'pretrainedNetwork');
pretrainedNetwork = fullfile(pretrainedFolder,'deeplabv3plusResnet18CamVid.mat');
if ~exist(pretrainedNetwork,'file')
    mkdir(pretrainedFolder);
    disp('Downloading pretrained network (58 MB)...');
    websave(pretrainedNetwork,pretrainedURL);
end
%%
imageURL =
'http://web4.cs.ucl.ac.uk/staff/g.brostow/MotionSegRecData/files/701_StillsRaw_full.zip';
labelURL =
'http://web4.cs.ucl.ac.uk/staff/g.brostow/MotionSegRecData/data/LabeledApproved_full.zip';

outputFolder = fullfile(tempdir,'CamVid');
labelsZip = fullfile(outputFolder,'labels.zip');
imagesZip = fullfile(outputFolder,'images.zip');

if ~exist(labelsZip, 'file') || ~exist(imagesZip,'file')
```

```
    mkdir(outputFolder)

    disp('Downloading 16 MB CamVid dataset labels...');
    websave(labelsZip, labelURL);
    unzip(labelsZip, fullfile(outputFolder,'labels'));

    disp('Downloading 557 MB CamVid dataset images...');
    websave(imagesZip, imageURL);
    unzip(imagesZip, fullfile(outputFolder,'images'));
end
%%
imgDir = fullfile(outputFolder,'images','701_StillsRaw_full');
imds = imageDatastore(imgDir);
%%
I = readimage(imds,559);
I = histeq(I);
imshow(I)
%%
    classes = [
        "Sky"
    "Building"
    "Pole"
    "Road"
    "Pavement"
    "Tree"
    "SignSymbol"
    "Fence"
    "Car"
    "Pedestrian"
    "Bicyclist"
    ];
%%
labelIDs = camvidPixelLabelIDs();
labelDir = fullfile(outputFolder,'labels');
```

```
pxds = pixelLabelDatastore(labelDir,classes,labelIDs);
%%
C = readimage(pxds,559);
cmap = camvidColorMap;
B = labeloverlay(I,C,'ColorMap',cmap);
imshow(B)
pixelLabelColorbar(cmap,classes);
%%
tbl = countEachLabel(pxds);
frequency = tbl.PixelCount/sum(tbl.PixelCount);bar(1:numel(classes),frequency)
xticks(1:numel(classes))
xticklabels(tbl.Name)
xtickangle(45)
ylabel('Frequency');
%%
[imdsTrain, imdsVal, imdsTest, pxdsTrain, pxdsVal, pxdsTest] = partitionCamVidData(imds,pxds);
numTrainingImages = numel(imdsTrain.Files);
numValImages = numel(imdsVal.Files);
numTestingImages = numel(imdsTest.Files);
% Specify the network image size. This is typically the same as the traing image sizes.
imageSize = [720 960 3];

% Specify the number of classes.
numClasses = numel(classes);

% Create DeepLab v3+.
lgraph = deeplabv3plusLayers(imageSize, numClasses, "resnet18");
%%
imageFreq = tbl.PixelCount ./ tbl.ImagePixelCount;
classWeights = median(imageFreq) ./ imageFreq;
pxLayer = pixelClassificationLayer('Name','labels','Classes',tbl.Name,'ClassWeights',classWeights);
lgraph = replaceLayer(lgraph,"classification",pxLayer);
```

```
%%% Define validation data.
pximdsVal = pixelLabelImageDatastore(imdsVal,pxdsVal);

% Define training options.
options = trainingOptions('sgdm', ...
    'LearnRateSchedule','piecewise',...
    'LearnRateDropPeriod',10,...
    'LearnRateDropFactor',0.3,...
    'Momentum',0.9, ...
    'InitialLearnRate',1e-3, ...
    'L2Regularization',0.005, ...
    'ValidationData',pximdsVal,...
    'MaxEpochs',30, ...
    'MiniBatchSize',8, ...
    'Shuffle','every-epoch', ...
    'CheckpointPath', tempdir, ...
    'VerboseFrequency',2,...
    'Plots','training-progress',...
    'ValidationPatience', 4);
%%
augmenter = imageDataAugmenter('RandXReflection',true,...
    'RandXTranslation',[-10 10],'RandYTranslation',[-10 10]);
pximds = pixelLabelImageDatastore(imdsTrain,pxdsTrain, ...
    'DataAugmentation',augmenter);

doTraining = true;
if doTraining
    [net, info] = trainNetwork(pximds,lgraph,options);
else
    data = load(pretrainedNetwork);
    net = data.net;
end
```

```
%%
I = readimage(imdsTest,35);
C = semanticseg(I, net);
B = labeloverlay(I,C,'Colormap',cmap,'Transparency',0.4);
imshow(B)
pixelLabelColorbar(cmap, classes);
expectedResult = readimage(pxdsTest,35);
actual = uint8(C);
expected = uint8(expectedResult);
imshowpair(actual, expected);
%%
iou = jaccard(C,expectedResult);
table(classes,iou);
%%
pxdsResults = semanticseg(imdsTest,net, ...
    'MiniBatchSize',4, ...
    'WriteLocation',tempdir, ...
    'Verbose',false);
metrics = evaluateSemanticSegmentation(pxdsResults,pxdsTest,'Verbose',false);
metrics.DataSetMetrics
metrics.ClassMetrics
```

Appendix 3: Hidden Markov driver severity model

```
%% Hidden Markov driver severity model for vehichle transition
% transition probabilities between low severity states and high severity states
% emmission probability on symbols {k1,k2} from hugh to high, high to low
% based on observations t={t1,t2,t3}

trans = [0.95,0.05;    % Based on TeleFOT Data for Driver 001
         0.10,0.90];
 emis = [1/6 1/6 1/6 1/6 1/6 1/6;
    1/10 1/10 1/10 1/10 1/10 1/2];
[seq,states] = hmmgenerate(100,trans,emis);
```

```
[estTR,estE] = hmmtrain(seq,trans,emis);
[estimateTR,estimateE] = hmmestimate(seq, states);
estimatesStates = ...
  hmmviterbi(seq,estimateTR,estimateE,...
          'Statenames',{'Low';'High'});
```

Appendix 4: Noise Removal

```
function Ioutput=Removenoise(I)
PSF = fspecial('gaussian',3,5);
%PSF = fspecial('disk',2); % optimise
%PSF = fspecial('prewitt');
%PSF = fspecial('sobel');
%PSF= fspecial('motion',4,6);
%PSF= fspecial('laplacian',0.65);
INITPSF=ones(size(PSF));
%output=deconvblind(I, INITPSF, 8);
Ioutput=deconvlucy(I, INITPSF, 8);
```

Appendix 5: CNN LSTM

```
%%
netCNN=googlenet;
cnnLayers=layerGraph(netCNN);
%%
inputSize = netCNN.Layers(1).InputSize(1:2);
averageImage = netCNN.Layers(1).Mean;

inputLayer = sequenceInputLayer([inputSize 3], ...
   'Normalization','zerocenter', ...
   'Mean',averageImage, ...
   'Name','input');
%%
```

```
layerNames = ["data" "pool5-drop_7x7_s1" "loss3-classifier" "prob" "output"];
cnnLayers = removeLayers(cnnLayers,layerNames);

%%
layers = [
    inputLayer
    sequenceFoldingLayer('Name','fold')];

lgraph = addLayers(cnnLayers,layers);
lgraph = connectLayers(lgraph,"fold/out","conv1-7x7_s2");
%%
lstmLayers = netLSTM.Layers;
lstmLayers(1) = [];
%%
layers = [
    sequenceUnfoldingLayer('Name','unfold')
    flattenLayer('Name','flatten')
    lstmLayers];

lgraph = addLayers(lgraph,layers);
lgraph = connectLayers(lgraph,"pool5-7x7_s1","unfold/in");
%%
lgraph = connectLayers(lgraph,"fold/miniBatchSize","unfold/miniBatchSize");
%%
analyzeNetwork(lgraph);
net = assembleNetwork(lgraph);
```

Appendix 6: Image enhancement

```
% The base settings
function BI2 = image_enhancement(A)

Lab=rgb2lab(A);
Linv=imcomplement(Lab(:,:,1) ./100);
```

```
Lenv=imcomplement(imreducehaze(imadjust(Linv),'Method','approxdcp','ContrastEnhancement','boost', 'AtmosphericLight',0.95,'BoostAmount',0.9));
%Lenv=imcomplement(imreducehaze(Linv,'Method','approxdcp','ContrastEnhancement','none', 'AtmosphericLight', 0.95));
Labenv(:,:,3)=Lab(:,:,3) * 4.5;
Labenv(:,:,2)=Lab(:,:,2) * 3.5;
Labenv(:,:,1)=adapthisteq(Lenv) .* 99;
C=lab2rgb(Labenv);
BI2=imguidedfilter(C,A);
%BI=trim_image3(BI2);
```

Appendix 7: Split Frame

```
function out = SplitFrame(X)

%LowerHalf = [0.5 359.5 1275 361];
LowerHalf1=[0.5 359.5 633 361];
%UpperHalf = [5.16101694915278 0.5 1275.33898305085 360.762711864407];
out=imcrop (X, LowerHalf1);

%UpperHalf = [5.16101694915278 0.5 1275.33898305085 360.762711864407];
%UpperHalf1=[0.5 0.5 633 361];
%UpperHalf2= [647.5 0.5 633 361];
%LowerHalf = [0.5 359.5 1275 361];
%LowerHalf1=[0.5 359.5 633 361];
%LowerHalf2= [647.5 359.5 633 361];
```

Appendix 8:Bayesian Series Model

```
function [PosteriorMdl,X] = bayestimeseriesmodel(x,y)
%Main contribution: predictor using MDDRA severity score(IEEE Access)
% x, y time series data
numseries = 3; % enter number of series here
```

```
numlags = 4;  % time lag
PriorMdl = bayesvarm(numseries,numlags);
%PriorMdl = bayesvarm(numseries,numlags,'ModelType','conjugate')
[PosteriorMdl,Summary] = estimate(PriorMdl,x,y);
% Access the 95% equitailed credible interval of the regression coefficient
X=Summary.CI95(:,:);

%% yF = forecast(PosteriorMdl,XF); to be used for driver severity score forecast
```

Appendix 9: fuzzy logic dbn

```
%%
%%filename = ('IntelligentEnvironment.xls');
%%testData = xlsread(filename)
filename = ('newdatasheet.xls');
testData = xlsread(filename);

% Declare a new FIS
a = newfis('AlexSystem');

% Input variable: Time of Year (Days)
a = addvar(a, 'input', 'Face_Orientation', [0 1]);
a = addmf(a, 'input', 1, 'Forward', 'trimf', [-0.9 -0.1 0.4199 0.9]);
a = addmf(a, 'input', 1, 'Sideways', 'trimf', [0.0973 0.6136 1.1 1.9]);



% Input variable: Time of Day (Mins)
a = addvar(a, 'input', 'Drivers_Distractions', [0 1]);
a = addmf(a, 'input', 2, 'Not Texting', 'trapmf', [-0.9 -0.1 0.4991 0.9]);
a = addmf(a, 'input', 2, 'Texting', 'trapmf', [0.55 0.554 1.05 1.45]);
```

```
% Input variable: Outdoor Temp (C)
a = addvar(a, 'input', 'Hands', [0 1]);
a = addmf(a, 'input', 3, 'two_hands', 'trimf', [-0.9 -0.1 0.4991 0.9]);
a = addmf(a, 'input', 3, 'One Hand', 'trimf', [0.102 0.5273 1.1 1.9]);

% Input variable: Drivers Distractions Temp
a = addvar(a, 'input', 'Previous Driver Distraction', [0 1]);
a = addmf(a, 'input', 4, 'Safe_Driving', 'trapmf', [-0.45 -0.05 0.3424 0.45]);
a = addmf(a, 'input', 4, 'Careless_Driving', 'trapmf', [0.05 0.369 0.645 0.95]);
a = addmf(a, 'input', 4, 'Dangerous_Driving', 'trapmf', [0.518 0.664 1.041 1.42]);



% Output variable: Heating (%)
a=addvar(a,'output','Driving_Risk_Severity (%)',[0 1]);

a = addmf(a, 'output', 1, 'Safe_Driving', 'trimf', [-0.416666666666667 0 0.416666666666667]);
a = addmf(a,'output',1,'Careless_Driving','trimf',[0.0780423280423282 0.494708994708995 0.911375661375662]);
a = addmf(a,'output',1,'Dangerous_Driving','trimf',[0.583333333333333 1 1.41666666666667]);

% Create rules for the FIS, the last value is for AND or OR



rule1 = [1 1 1 1 1 1 1];
rule2 = [1 1 2 1 2 0.35 1];
rule3 = [1 2 1 2 3 0.44 1];
rule4 = [1 2 2 3 3 1 1];
rule5 = [1 1 1 3 2 0.85 1];
rule6 = [1 1 2 2 1 0.6 1];
```

```
rule7 = [2 1 1 1 1 0.35 1];
rule8 = [2 1 2 1 2 0.60 1];
rule9 = [2 2 1 2 3 0.70 1];
rule10 = [2 2 2 3 3 1 1];
rule11 = [2 2 1 3 2 0.85 1];
rule12 = [2 2 1 2 1 0.60 1]; %Mrn & C = High
rule13 = [1 2 1 1 2 0.35 1];
rule14 = [1 2 2 1 2 0.60 1];
rule15 = [1 1 1 1 2 0.20 1];
rule16 = [1 2 2 2 3 0.85 1];


% Pass the rules to an rule59 = [3 0 5 1 1 1];array
%ruleList = [rule1;rule2;rule3;rule4;rule5;rulrule59 = [3 0 5 1 1 1];e6;rule7;rule8;rule9];
ruleList = [rule1; rule2; rule3; rule4;...
rule5; rule6; rule7; rule8; rule9; rule10;...
rule11; rule12; rule13; rule14; rule15; rule16;];

% Add the rules to the FIS
a = addrule(a,ruleList);

% Print the rules to the workspace
rules = showrule(a)

% Set the defuzzification method
%a.defuzzMethod = 'centroid';
%a.defuzzMethod = 'bisector';
%a.defuzzMethod = 'mom';
%a.defuzzMethod = 'som';
a.defuzzMethod = 'lom';

for i=1:size(testData,1)
eval = evalfis([testData(i, 1), testData(i, 2), testData(i, 3) , testData(i, 4) ], a);
```

```
fprintf('%d) In(1): %.2f, In(2) %.2f, In(3) %.2f, In(4) : %.2f => Out: %.2f \n\n',i,testData(i, 1),testData(i, 2),testData(i, 3),testData(i, 4), eval);
xlswrite('newdatasheet.xls', eval, 1, sprintf('F%f',i+1));
end

ruleview(a)

figure(1)
subplot(3,2,1), plotmf(a, 'input', 1)
subplot(3,2,2), plotmf(a, 'input', 2)
subplot(3,2,3), plotmf(a, 'input', 3)
subplot(3,2,4), plotmf(a, 'input', 4)
subplot(3,2,5), plotmf(a, 'output', 1)



surfview(a)
[System]
Name='AlexSystem2NewTOOLBOXwtexting'
Type='mamdani'
Version=2.0
NumInputs=4
NumOutputs=1
NumRules=12
AndMethod='min'
OrMethod='max'
ImpMethod='min'
AggMethod='max'
DefuzzMethod='centroid'

[Input1]
Name='Face_Orientation'
Range=[0 1]
NumMFs=2
MF1='Forward':'trimf',[-1 0 1]
```

MF2='Sideways':'trimf',[-0.00269541778975735 0.997304582210243 1.99730458221024]

[Input2]

Name='Drivers_Distractions'

Range=[0 1]

NumMFs=2

MF1='not_talking_with_passanger':'trapmf',[-0.9 -0.1 0.499119718309859 0.9]

MF2='texting':'trapmf',[0.55 0.554 1.05 1.45]

[Input3]

Name='Hands'

Range=[0 1]

NumMFs=2

MF1='Two_Hands':'trimf',[-1 0 1]

MF2='One_Hand':'trimf',[0.00179694519317164 1.00179694519317 2.00179694519317]

[Input4]

Name='Previous_Risk_Severity'

Range=[0 1]

NumMFs=3

MF1='Safe_Driving':'trapmf',[-0.45 -0.05 0.342429577464789 0.45]

MF2='Careless_Driving':'trapmf',[0.05 0.369 0.645 0.95]

MF3='Dangerous_Driving':'trapmf',[0.518 0.664 1.04137323943662 1.42]

[Output1]

Name='Driving_Risk_Severity'

Range=[0 1]

NumMFs=3

MF1='Safe_Driving':'trimf',[-0.416666666666667 0 0.416666666666667]

MF2='Careless_Driving':'trimf',[0.0780423280423282 0.494708994708995 0.911375661375662]

MF3='Dangerous_Driving':'trimf',[0.583333333333333 1 1.41666666666667]

[Rules]

1 1 1 1, 1 (1) : 1

1 1 2 1, 2 (1) : 1

1 2 1 2, 3 (1) : 1

1 2 2 3, 3 (1) : 1

1 1 1 3, 2 (1) : 1

1 1 2 2, 1 (1) : 1

2 1 1 1, 1 (1) : 1

2 1 2 1, 2 (1) : 1

2 2 1 2, 3 (1) : 1

2 2 2 3, 3 (1) : 1

2 2 1 3, 2 (1) : 1

2 2 1 2, 1 (1) : 1