Physiological Measurements based Automatic Driver Cognitive Distraction Detection

by

Afizan Azman

A Doctoral Thesis
Submitted in partial fulfilment of the requirements
for the award of

Doctor of Philosophy
of
Loughborough University

May, 2013

ABSTRACT

Vehicle safety and road safety are two important issues. They are related to each other and road accidents are mostly caused by driver distraction. Issues related to driver distraction like eating, drinking, talking to a passenger, using IVIS (In-Vehicle Information System) and thinking something unrelated to driving are some of the main reasons for road accidents. Driver distraction can be categorized into 3 different types: visual distraction, manual distraction and cognitive distraction. Visual distraction is when driver's eyes are off the road and manual distraction is when the driver takes one or both hands off the steering wheel and places the hand/s on something that is not related to the driving safety. Cognitive distraction whereas happens when a driver's mind is not on the road. It has been found that cognitive distraction is the most dangerous among the three because the thinking process can induce a driver to view and/or handle something unrelated to the safety information while driving a vehicle. This study proposes a physiological measurement to detect driver cognitive distraction. Features like lips, eyebrows, mouth movement, eye movement, gaze rotation, head rotation and blinking frequency are used for the purpose. Three different sets of experiments were conducted. The first experiment was conducted in a lab with faceLAB cameras and served as a pilot study to determine the correlation between mouth movement and eye movement during cognitive distraction. The second experiment was conducted in a real traffic environment using faceAPI cameras to detect movement on lips and eyebrows. The third experiment was also conducted in a real traffic environment. However, both faceLAB and faceAPI toolkits were combined to capture more features. A reliable and stable classification algorithm called Dynamic Bayesian Network (DBN) was used as the main algorithm for analysis. A few more others algorithms like Support Vector Machine (SVM), Logistic Regression (LR), AdaBoost and Static Bayesian Network (SBN) were also used for comparison. Results showed that DBN is the best algorithm for driver cognitive distraction detection. Finally a comparison was also made to evaluate results from this study and those by other researchers. Experimental results showed that lips and eyebrows used in this study are strongly correlated and have a significant role in improving cognitive distraction detection.

To Mak and Abah

For their love

For their understanding

For their sacrifice

For everything…Thank you.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

List of publications	208

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

Chapter 1-Introduction

Statistic 1:

According to the National Highway Traffic Safety Administration (NHTSA) of the U.S. Department of Transportation, various factors were found to contribute to road accidents. Driver behaviour frequently played a role in the car crashes and some of the findings were surprising. Below is a summary of driver behaviours relating to road accidents [1]:

- Approximately 41% of drivers made some type of recognition error, due to factors such as inattention, internal or external distractions, inadequate surveillance, etc.

- Approximately 34% of drivers made decision errors, including driving aggressively, driving too fast, etc.

- Approximately 10% of drivers made performance errors, such as over compensation, improper directional control, etc.

Statistic 2:

According to the Crash Index by AAMI (Australia Insurance Company) in its report "Absent Minds Take Their Toll"- Nationally, almost half of all drivers attribute previous crashes to driver inattention (44 per cent), followed by speeding (17 per cent), fatigue (11 per cent) and alcohol (9 per cent) [2].

Scenario 1:

PORTSMOUTH - A three-car chain-reaction crash snarled lunch hour traffic on Woodbury Avenue Friday and sent a Maine man to Portsmouth Regional Hospital. Police say inattentiveness by one driver, who was looking for a cigarette or a lighter, was to blame [3].

Introduction

Statistics 1, 2 and Scenario 1 are examples of driver's inattention contributing to road accidents which had happened in America and Australia. From the statistics and scenario given, it is clearly shown that a driver's attention is a very important issue in road safety. In UK, the Department for Transportation (DfT) has reported in 2007, that 92% of passenger travel is by road [4]. This passenger road transport can be categorized into several different modes: car, van, taxi, bus and coach, motorbikes, bicycle and even railroad. With a high percentage of passengers on the road, safety issues become very crucial.

GB passenger transport by mode

	1996	2006
	billion passenger km	
Car, van, taxi	622	686
Bus and coach	43	50
Rail	39	55
Motorbikes	4	6
Bicycle	4	5
Air	6	10
Total	718	812

Source: DfT 2007

Road accounts for 92% of all passenger travel, while rail serves less than 7% of our travel needs. Train usage is limited to a fraction of all travel, despite significant investment to improve rail services since privatisation and an increase of over 40% in the number of people using rail over the last decade.

92% of passenger travel is by road

Figure 1.1: DfT report 2007 [4].

Driving safety issues include inattention, fatigue, concentration, driver's behaviour, alcohol, drug, using mobile phone, listening to the radio/music and many more. From all those issues, the most important issue that has to be considered seriously is the driver's concentration and distraction.

Driver distraction can be defined as a process that draws away drivers' attention from the road. It disturbs the driving process and vehicular control. Justification for this definition might be found in Webster's Third New International Dictionary (1965) definition: from the Latin dis (apart) +

trahere (draw or pull), "to draw or cause to turn away from an original position, goal, purpose, direction, association or interest [5]. Webster also listed synonyms to distraction as divide, separate, harass and confound. From the Fernald (1947) book of synonyms the terms are disturb, perturb, remove, detach, steal, withdraw, purloin and confuse. Thus, all of those terms can imply a compromise of safety and also can connote with what distraction means to a control engineer in terms of its effect on system performance.

Generally, concentration and attention are two similar words. Concentration is when someone does something with focus, i.e. to give one's attention. Concentration and concentrating can also be defined as an ability to re-focus attention if and when one is much distracted or immediately after rest or more breaks, on the same task or different tasks; and the strength of the ability to concentrate is actually by mental conditioning. Attention, on the other hand, is part of focus, concentration, a component of intelligence. Attention usually only lasts for a few seconds. Some tests suggested that a person can only provide one's attention to 90 seconds, however, 30 seconds is regarded as the maximum duration of giving an attention. Attention is the act or process of focusing on one or more particulars in the content of one's consciousness in a way to give an essential or priority to restricting one's input and ignoring unwanted aspects [5].

1.1   Motivation

It has been the desire of all countries in the world possessing a network of transportation system, to hope for a special method or technique that can be implemented to reduce the number of road crashes and accidents. In UK, in 2008, the casualties in accidents reported to police were 7 % lower than those from the year 2007 [6]. It means the government has achieved its target of reducing the number of accidents on the road, by enforcing several road policies like speed limit and rules for motorists. In this article, studying why accidents happened could help steer improvements in policy, and road or vehicle design. Ministers in UK government are proposing an expert panel to be set up to study fatal accident patterns and if this trend and effort continues in UK, it will soon restore the country's position as having one of the safest road systems in the world. In a report prepared by DfT (Department for Transportation) in UK, it showed that in 1996, around 622 billion passengers had used car, van and taxi, 43 billion had used bus or coach, and 8 billion had used motorbikes or bicycles. However, these figures had increased significantly

by 2006 (ten years after). The number of passengers using car, van and taxi became 686 billion, which was a 10% increase after ten years. Bus or coach passengers increased to 50 billion, whereas motorbikes and bicycles reached to 11 billion [4].

Thus, it shows that once the number increases, the responsibility of every road user increases as well. This responsibility is to be shared by all road users and the government. Many facilities, equipments, services, and technologies need to be developed to investigate the driving safety and to assist road users. A road accident investigation body should not just blame others or point fingers to road users but it should lead other road investigators and researchers to improve the policy, regulation, road design, vehicle design and even traffic management in such a way that promotes a reduction in the number of accidents. Research organization says lessons are not being learned from the deaths on Britain's road [7]. Therefore, road crash expert teams are asked to develop a system or technology that can remind and make road users in Britain to learn about the importance of staying in focus while driving and being vigilant on the road.

A mitigation system that can reduce a driver's level of distraction is not necessary to avoid accidents or road crashes. However, any system that can help to put back a driver's attention on track and ensures that the driver is vigilant every time is necessary at this moment. Since this research also focuses on vehicle safety design, it is crucial for a promising strategy to be proposed to minimize driver distractions. National Highway Traffic Safety Administration (NHTSA) has estimated that driver distraction and inattention contribute to 20-30 percent of road accidents, numbering to about 1.5 million crashes a year. Thus, based on the vast number of crashes that had occurred, a vehicle with an adaptive interface technology has been developed. SAVE-IT (SAfety VEhicle using adaptive Interface Technology) is a vehicle incorporating an adaptive interface technology to help driver to mitigate the distraction. Electronic devices, stand-alone systems and portable hand held devices mounted in cars can potentially distract drivers. Therefore, SAVE-IT is a good starting point to monitor a driver's vigilance level. A central monitoring system is a promising approach to achieving and increasing the benefits of advanced vehicle technologies without creating adverse safety consequences due to driver distraction and information overload. Such a system would integrate the data obtained by sensors to control the information flow to the driver through an adaptive driver-vehicle interface [8]. With this adaptive interface technology, a driver's level of state can be monitored. Any movements especially on the driver's face area can be observed. Much information can be

captured from a face. Movements from eyes, mouth, lips, eyebrows, or even head orientation can tell about a driver's state in terms of cognitive distraction.

Generally there are three types of driver distraction: Visual, Manual and Cognitive. Visual distraction and manual distraction are easier to be detected, unlike cognitive distraction it requires a few modalities to determine the state of cognitive distraction. Visual distraction is when the driver is not looking on the road in front and manual distraction is when the driver's hand is not on the steering wheel or handles anything in the vehicle which is related to the driving safety. Cognitive distraction is an internal distraction and difficult to detect. It occurred when a driver's mind is used to think about anything else than a driving safety issue. For a visual distraction detection, it is easy to determine whether a driver is visually distracted or not by looking at the eyes movement. Percentage of Eye Closure (PERCLOS) might be the easiest way to determine visual distraction. Detecting cognitive distraction can be accomplished by using physiological measurement. Physiological measurement can refer to any organs, cells, or bio molecules that are used to indicate certain state of a human being. Many recent researches in detecting driver cognitive distraction focused only on one type of physiological measurement. Most of the information are captured on eye movement, including pupil diameter, gaze direction, eye fixation, eye saccade, eye smooth pursuit, blinking rate, blinking duration and PERCLOS [24][36]. It is a good idea to fuse this eye movement with other facial features like mouth, eyebrows and lips to improve the performance accuracy in detecting the cognitive distraction level.

1.2   Research Objective

Cognitive distraction is still an evolving topic in the research area of driver distraction as well as vehicle and road safety. Cognitive distraction is an internal distraction to the driver when the driver's mind is off the road and not thinking anything related to the driving safety. Many sub-areas are still widely opened to be discovered in this type of attention impairment. RAC foundation in UK said that Britain needs a road accident investigation body to work and study on the causes of road crashes and casualties [7]. It aims to find out why collisions have happened rather than who was to blame. Even though many appliances, equipment and systems have been

developed to increase vehicle and road safety, they are still insufficient in decreasing the number of crashes on the road.

A promising strategy to minimize driver distraction is really top notch. The British government is really looking forward to any organization, institution or personnel to develop a system and make a serious study to reduce the number of crashes on the road. When there is an air crash, train crash or tragedy at sea, many specialists and investigators are called in to investigate the causes.

This thesis aims for several purposes. First, it aims to study and understand cognitive distraction: its definition, implication and measurement. Ideally, cognitive distraction is related to a psychology research area. Thus, an in depth understanding about what is cognition in human is very important. Cognition is a theory of psychology that attempts to explain human behavior by understanding what is happening in the thought processes. The assumption is that, humans are logical beings who make choices with the most sense to them [9]. Thus, when a driver is driving and he/she is involved with a thought process, he/she is intentionally doing something that is not related or contributed to his/her safety while driving. A driver might be in serious danger in moments if he/she is driving and processing information for a long time of period, because he/she is thinking about something that is unrelated with his/her current situation while on the road. This could lead into a slower response to incoming hazards, larger lane variation, abrupt steering control and less efficient visual perception.

Secondly, the paper aims to recognize new physiological measurements that can be used to detect cognitive distraction. Actually there are four basic measurements to detect a driver's cognitive distraction: primary task, secondary task, physiological measurement and driving performance behavior. These four types of measurement are explained briefly in chapter 3. Recent research has found only a few types of measurement from human physiology in determining a driver's cognitive state. This includes measurements on eye gaze direction, pupil diameter, eye movements (fixation, saccade and smooth pursuit), head pose and heart wave (using ECG waveform). However, in choosing the best type of physiological measurement, it is very important to consider which measurement type is not intrusive but yet can give an accurate result. A good physiological sensor system should not disorient driver behaviour when driving on a real road environment.

Once the features have been identified, algorithms to classify between distracted state and undistracted state are necessary to study. There are many available algorithms for classification like Naïve Bayesian Network, Linear Regression, Support Vector Machine, J48 Decision Tree, Bayesian Network and many others. However for driver cognitive distraction detection, a few good algorithms have been recognized including Support Vector Machine (SVM), AdaBoost, Logistic regression and also Bayesian Network (BN). Those algorithms and modeling techniques will be compared later on to identify which one can give a more precise and accurate result. The model performance measures include:

a)   Accuracy rate/successful rate

b)   Sensitivity rate

A few other functions like Precision, Recall and F-measure are also computed in order to compare those algorithms performance. The developed algorithms are built with the parameters and features chosen from the physiological measures as mentioned above.

Another minor objective which can also be considered important in this thesis is to propose a better experimental setup for collecting data on driver cognitive distraction in a real time environment. Recently, many studies are involved with in-door data collection for this research. Therefore, the reliability and applicability of the data collected is questionable. An appropriate real time experiment needs to be designed to avoid any problem in data collection like missing data, noise and incorrect data collection setup. The in-door experimental setup which was used by some researchers can serve as significant references to design this as an out-door real time experiment. The most important part in the experimental design is where the cognitive state needs to be triggered. Cognitive distraction is an internal distraction, and it cannot be identified and recognized easily like visual or manual distraction.

1.3   Problem Overview

There are several research gaps available and need to be covered in this study. Driver cognitive distraction detection is still a new research in driving safety. Even driver's attention and concentration has been studied by many researchers a few years ago, however, the detection part has not been fully investigated. The applications and devices to detect driver's visual distraction

have been developed and used by many automobile companies like Toyota, Mercedes Benz and Volvo. Unfortunately, there is not yet a system being designed to detect driver's cognitive distraction. This is due to the lack of information regarding the detection part.

Almost every research that studies on driver's cognitive distraction combines two different types of modalities to detect driver's cognitive state. Physiological measurement is the most popular method to detect driver's cognitive state, however, this method is usually combined with driver's performance measurement like steering angle and lane keeping. This driver's performance measurement is possibly dangerous to drivers because it can undermine the driver's driving performance. By tracking the driver's driving performance, it might cause the driver to focus only on what is being measured and ignore the other important activities while driving like looking at the rear mirror, estimating the distance with the vehicles in front and others.

From the physiological measurements used to detect driver's cognitive distraction, many studies only focused on the eye movement and information like PERCLOS, blinking frequency, blinking duration, gaze rotation, eyelid movement, fixation, saccade and smooth pursuit. There are actually a few more available features and parameters that can be helpful to detect driver's cognitive distraction like lips, eyebrows and head orientation. These three parameters can obviously be seen and moved when a driver is cognitively distracted. Even a normal web camera can be used to detect these parameters. Unlike the detection of blinking frequency, to detect pupil diameter and gaze rotation, more expensive and sensitive infrared cameras are required.

A proper experimental setup with proper types of distraction is also required to trigger driver's cognitive distraction. Available studies nowadays mostly focused on in-door experimental setups. Therefore, the data collected might be questionable because it was not collected from a real driving environment.

1.4   Contribution

Strategies are required to minimize driver distraction. One way is by developing adaptive distraction mitigation system which is able to adjust its function and provide assistance to reduce the distraction based on the state of the drivers. The system should be able to identify user's state of mind and adapt to it. Driver's state information is collected by a range of sensors in real time.

Based on the information, an appropriate mitigation strategy can be performed accordingly to maintain driver's performance in driving. Many mitigation strategies like warning systems, blocking distraction sources and providing feedback upon the distraction are available to the drivers. However, these mitigation strategies are only accurate if the information captured from the driver's state is also accurate. Therefore, an immaculate method for information collection and algorithms for detection are required.

This study is to propose new features for driver's cognitive distraction detection. Therefore, the proposed study was executed with three different sets of experiment. These three experiments had different objectives. The first experiment is to propose new features which can only be proven theoretically and can have an impact in detecting driver's cognitive distraction. The second experiment is to check for the correlation and relationship between those proposed features, and the last experiment is to fuse these proposed features with existing features used in other studies. This features fusion is targeted to improve the performance of accuracy to detect driver's cognitive distraction.

Once the information is collected, algorithms need to be developed. Algorithms developed in this thesis are able to work in a real driving environment. Prior and posterior information are considered to give a better result in detection. The main algorithm in this thesis is the Dynamic Bayesian Network algorithm. In this algorithm, the probability of information from previous time step is also calculated.

Mainly, this thesis focuses on the development of the algorithms. A Dynamic Bayesian Network with fusion of features from different cameras is developed to satisfy the requirement for a real time experiment. Accurately identifying distraction in driving is a complex critical challenge. Appropriate algorithm is required to detect real time distraction. Unfortunately, nowadays, the algorithms to detect cognitive distraction in real time have not been well developed. Therefore, to address this need, the main goal of this thesis is to develop detection algorithms for cognitive distraction in driving. A few other algorithms like Static Bayesian Network, Support Vector Machines and Regression are also used for performance comparison with Dynamic Bayesian Network

Towards the end of this study, the following main contributions will have been made:

a) Strong correlation relationship between lips and eyebrows (or mouth and eye) is found when a driver is cognitively distracted. This contribution is explained in chapter 5.

b) A data fusion between lips, eyebrows and existing features like blinking, head rotation and gaze rotation has improved the accuracy performance for driver cognitive distraction detection. Results and comparison are highlighted for this finding in chapter 6.

c) Different classification algorithms like Dynamic Bayesian network, Static Bayesian Network, Support Vector Machine and Logistic Regression algorithms are used. Results and comparison can also be found in chapter 6.

d) Dynamic Bayesian Network algorithm for driver cognitive distraction detection with new features (lips and eyebrows) is presented in the network model.

Other than three main contributions above, below are other possible contributions which can be found in this thesis.

a) Investigation is made on how cognitive distraction is able to influence driver's performance.

b) An appropriate and better experiment is set up for driver cognitive distraction detection in real traffic environment.

c) A number of algorithms are used to detect driver's cognitive distraction. Comparisons between those algorithms are also made.

1.5   Thesis Layout

For clarity in presentation, this thesis is arranged as follows:

Chapter 2 explains background and literature review about driving safety. This chapter covers differences between active safety and passive safety definition, factors that lead to road crashes (alcohol, drug, age, and fatigue), driver attention status and types of distraction. It proposes the relationship between lips and eyebrows with thinking expression and cognition, which have been found in psychological and sociological area.

Chapter 3 describes in details three different sets of experiment conducted for this study. Information like distraction setup, participant's details, equipment setup and pictures taken during the experiments can be found in this chapter. At the end of this chapter a comparison is made between the experimental setup from this study with other experimental setup from different studies and researches.

Chapter 4 covers details about physiological measurements and Bayesian Networks. Several types of used physiological measurements like pupil diameter and gaze rotation are explained in here. Concept, formula and algorithm for Bayesian Networks are explained in this chapter too. Two types of Bayesian Networks: Static and Dynamic, which are used in this study, are discussed with detailed description and examples. SBN and DBN models developed in this study are also described in here. Since there are three different experiments, the three different models of SBNs and DBNs are explained here as well.

Chapter 5 explains results of the correlation studied between mouth and eye movement, and also lips and eyebrows. Different types of correlation methods are used for different experiments. The first section explains the correlation between mouth movement (height and width) with eye movement (height and width). The final section explains the correlation between eyebrows and lips. A numbers of graphs and tables are presented in this chapter. These graphs and tables are used to show the positive relationship between the new features proposed in this study. Scatter diagram and linear regression are presented to show those relationships of the features. It occurred that a positive relationship between mouth movement and eye movement (and later lips and eyebrows) is plotted on the graph and computed from the table when a driver is cognitively distracted.

Chapter 6 is the highlight of the thesis. This chapter discusses five types of algorithms and quantitative techniques for classification: Support Vector Machine, AdaBoost, Logistic Regression, Static Bayesian Network and Dynamic Bayesian Network. Several types of performance measurements are also presented in this chapter like accuracy rate, sensitivity rate, precision, recall and many others. At the end of this chapter, the comparison between those algorithms is made. A comparison between algorithms used in this study and algorithms used by other studies and researches is also presented.

Chapter 7 presents the concluding chapter, which highlights the contributions made within this study and also suggests a list of improvements which can be accomplished for future works.

At the end of this thesis a list of publications published and submitted for this study and bibliography are listed.

Chapter 2

Driving Safety and Cognitive Distraction

"A tree never hits an automobile except in self-defences [10]."

(American Proverb)

"Any man who can drive safely while kissing a pretty girl is simply not giving the kiss the attention it deserves [10]."

(Albert Einstein)

2.1   Overview

This chapter presents definition and context of driving safety, driving model and driving distractions. There are three major types of distraction. However, only cognitive distraction is explained in detail in this chapter. A state of mind phenomenon called "Looked but Failed to see," which is related to driver cognitive distraction is also explained. Cognitive distraction is highly linked with attention, focus, and vigilance. Attention is a person's ability to maintain focus and remain alert to target stimuli over a period of time. As the above quotations illustrate, the study of vigilance and attention came into prominence since a hundred years ago. Road accidents which happen every day and everywhere are actually caused by human errors. It is the driver's inattention and lack of focus while driving that causes the accidents. Every activity while driving should get its own attention. Failure to stay in focus and driving without attention for a period of time are common examples that can lead to an accident.

2.2   Driving Safety Issue

When it comes to the definition of a driver's attention and concentration, it is all about a driver's looking, doing and thinking only on everything that relates to information while driving. While driving, a driver's eyes must be on the road and he/she should observe the environment clearly [11]. A driver should be aware of other drivers, riders, cyclists or even pedestrians who are using the same road. A driver should also be alert of any signboards, traffic lights, bill boards and even his/her surrounding nature like trees, hills, or even drains while driving. When doing something, a driver's hands must always grab the steering at a proper angle. Traditionally, drivers tend to grab the steering wheel at the 10-2 position. Even law enforcement agencies also train their officers to place their hands on that favourite position. Regardless of what has been trained to enforcement officers, some drivers like to use 9-3 position or as low as 8-4 position. The American Automobile Association (AAA) prefers to use 9-3 position. For an optimum control and smooth steering inputs, the proper steering wheel hand position puts the left hand at 9 o'clock and the right hand at 3 o'clock [12]. Thus, with a proper angle of gripping the steering wheel, a driver can safely manage to control his/her car, despite any unpredictable event occurring while driving. For instance, when a cat suddenly crosses the road, the driver can still control his/her car safely. The driver can also avoid from hitting other vehicles, avoid crashing into something, and even avoid from making other dangerous manoeuvres. Generally, the driver's hands are supposed to be firmly located on the steering wheel. However, with so many In Vehicle Information System (IVIS) applications in a vehicle, such as radio, cassette/CD/DVD/MP3 player, navigation system, cellular phone and others, the hands might not be on the steering wheel. Those IVIS applications sometimes require the driver to manually handle a device and end up with only one hand gripping the steering wheel. This manual diversion can also distract a driver's concentration while driving. While driving, a driver is not only required to observe the road and handle the car properly, but also required to think only on something related to driving. This seems and sounds impossible because in a real time driving environment, a driver can easily be sidetracked into thinking about something else. There are several examples that usually happen in the real world when a driver is driving a moving car but

does not realize what he/she is currently involved with doing. In psychology, this is referred to as a subconscious mind phenomenon and in a driving environment, this situation is in line with Baumann and Krems' claim of inattention blindness or "looked-but-did-not-see" phenomenon or "looked but failed to see" (LBFTS) [13].

Driver concentration is a very serious issue when it comes to driving safety. Nevertheless, to put a driver into a full focused attention while driving is not an easy thing to do. With so many distractions coming from inside or outside the vehicle, a driver's concentration can be easily jeopardized. The question here is how can a driver actually concentrate while driving? There are some tips and rule of thumbs that drivers can follow to help them concentrate while driving. Several advices proposed by the Hyderabad Traffic Police on how a driver can sustain his/her concentration while driving are listed below [14]:

a) Always concentrate while driving, keeping one's eyes on the road and one's hands on the steering wheel/handle.

b) Never drink (alcohol) and drive.

c) Never use mobile phones while driving.

d) Do not drive in a distracted mood.

e) Drive only when one is emotionally stable.

f) If possible, do not involve oneself in conversations with co-passengers.

g) Use the horn sparingly and only to remind others of one's presence.

h) Do not eat or drink while driving. Always consume food when the vehicle is safely stationary.

i) Even having *gutka* or *pan masala* can cause serious distraction at times.

j) Do not play with kids in one's vehicle.

k) Never install a TV screen on the dash board.

l) Avoid using ear phones when riding a bike.

m) Do not drive when one feels sleepy or after being sleepless.

Driver inattention or distraction is actually a major contributor to accidents or crashes on the road. BBC News has reported that in 2008, there were 2,538 people killed on Britain's roads. Even this figure was 14% less than the figure from 2007, but the government is keen to restore UK's position as one of the safest countries in the world [15]. This means, every road users are invited to improve their usage of roads into a convenient and safe place to drive a vehicle on.

Improving the design of roads, learning and practicing wise and proper driving behaviours and instilling awareness of road safety are some of the main initiatives that road users need to do in order to achieve the country's target as to be one of the safest countries in the world Crashworthiness Data System which had been initiated by The National Highway Traffic Safety Administration in 1988 has contributed a big role in collecting data and information on passenger vehicles and road-safety. Their data, prepared by the American Automobile Association (AAA) has categorized driver's type of distractions into 13 categories [16]:

Figure 2.1: Driver's distractions [16]

a) Eating or drinking

b) Outside person, object or event

c) Adjusting radio, cassette, or CD

d) Other occupants in vehicle

e) Moving object in vehicle

f) Smoking related

g) Talking, Dialling or Listening on cellular/mobile phone

h) Using device/object brought into vehicle

i) Using device/controls integral to vehicle

j) Adjusting climate controls

k) Other distractions

l) Inattentive/lost in thought

m) Unknown distraction

Since, attention can only last up to 30 or 90 seconds, a driver always needs to be reminded if he/she is no longer focusing on the driving. There are many reasons that can make a driver lose his/her concentration while driving. The most well-known reasons for a driver's distraction or inattention come from either fatigue, handling/moving/using object in the vehicle, or lost in thought [16].

In BBC News report on 14th of May 2009, it has been reported in 2008 that around 45 people a year were killed in air, train or marine accidents all around UK. However in the same year, more than 2,900 died on the road [7]. Many investigators have been called to ascertain the causes of the accidents. It has been found that road crashes are contributed by many factors: vehicle condition, In-Vehicle Information System (IVIS), road condition, weather and even the drivers' behaviours. From these causes, driver error represents a dominant cause, since drivers are the only people responsible for the total control of their vehicles.

A research led by Dr. Trent Victor of the Human Systems Integration Department, Volvo Technological Development Corporation, mentioned that to assist drivers effectively, Volvo believes they need to be able to [17]:

a) collect real-time data on drivers' visual behaviour,

b) recognize what the driver is doing (using contextual information such as manoeuvres, actions and states)

c) predict what the driver would likely to do next, and

d) design an interface to assist the driver.

There are a few numbers of studies on the factors contributing to road crashes and one of the most significant studies was mentioned in [18]. The study was using epidemiological methods and it investigated on how frequently various factors contributed to road crashes. The causal factors are definite, probable or possible. From the study it has found that:-

a) 34.9% are from roadway factors.

b) 9.1% are from vehicle factors.

c) 92.9% are from driver error factors.

Although most road crashes are attributed to multiple causes, driver error always represents a dominant one. This is because drivers are responsible for operating vehicles and avoiding crashes. Thus, it clearly shows that driver error is the main contributor to road crashes. There are few examples that can be related to driver error: rear-end collisions, roadway departures crashes, the inability to detect hazards and the failure to control the vehicle properly.

All the factors mentioned above are the performance breakdowns which result from the impairment of drivers' attention. Drivers are easily distracted whenever the conditions surrounding them are not comfortable or suitable for them to drive. This may include poor condition of the vehicle, poor weather, or even poor driving performance and skills.

Driving safety issues are usually related to the attention impairments: fatigue, aging, distraction, and intoxicating substances i.e. drugs/alcohol. Driver's distractions basically are factored by three distinctive factors:

a) Drivers' state affecting factors

  (i)   Fatigue

  (ii)   Monotony

  (iii)   Intoxicating substances (drugs/alcohol)

b) Drivers' trait factors

  (i)   Experience

  (ii)   Age

(c) Environmental factors

  (i)   Road environment demands

  (ii)   Traffic demands

  (iii)   Vehicle ergonomics

  (iv)   Automation

  (v)   Feedback

Based on the provisional data published by the Institute of Alcohol Studies in UK, in the year 2007, a total number of 14,480 drivers were involved with drinking and driving, where 460 drivers were fatally killed, 1760 were seriously injured, and 12 260 were slightly injured [19]. In

the UK, the legal alcohol driving limit is a concentration of 80mg of alcohol per 100ml of blood – about two pints of lager, although this varies according to body size, fitness and how quickly one consumed the drinks. The strength of the drinks one consumes also affects one's drink driving limit. A unit of alcohol is a small glass of wine or half a pint of beer [20].

Fatigue is often cited in the accidents caused by either a young driver or a truck driver. These drivers easily get exhausted while driving because they tend to adopt a risky mentality of driving at night with a lack of decent-quality sleep. Since, the drivers are sleep deprived, they easily get drowsy on the road. Sleepiness can reduce a driver's reaction time (a critical element of safe driving). It also reduces vigilance, alertness and concentration, so much that the ability to perform attention-based activities (such as driving) is impaired. The speed at which information is processed is also reduced by sleepiness. The quality of decision-making may also be affected. A recent study by the Sleep Research Centre at Loughborough University indicates that driver fatigue causes up to 20% of accidents on monotonous roads [21]. Sleep Related Vehicle Accidents (SRVA) comprised 16% on major roads in South West England, and over 20% on Midland motorways. This suggests that there are several thousand casualties each year from accidents caused by drivers falling asleep at the wheel [22].

The factor of age can result in a longer response and reaction time to hazards and narrower field of attention, especially to older drivers. Aging changes the physical functioning, vision, perception, and processing abilities that could make driving unsafe.

Distraction has become increasingly important nowadays, especially when the In-Vehicle Information System (IVIS) like MP3/CD/Radio Cassette player, Navigator System and Video Screen Display have been embedded in many vehicles manufactured. These IVIS can simply undermine a driver's concentration while driving. Drivers are easily distracted visually, manually and even cognitively. In the 100-Car Study, driver inattention contributed to nearly 80% of crashes and 65% of near-crashes [23]. IVIS has been proven to induce distractions to drivers. As mentioned before, there are 3 major types of distraction: visual distraction, manual distraction and cognitive distraction [24].

Figure 2.2: Types of Distraction

Visual distraction occurs when the driver's eyes are not looking on the road/off the road. Manual distraction occurs when the driver's hands are not on the steering wheel for quite a while due to handling something not related to the driving process and cognitive distraction occurs when the driver's mind is off the road, not thinking of anything related to safe driving. Thus, distraction occurs when a person is not in an attentive state or drawing attention or focus on something else. In simple words, distraction is an obstacle to attention.

Visual distraction is always related to the driver's fatigue. The sleepiness while driving makes the driver not completely looking on the road. This may lead to a very serious accident. In year of 2008, Australian Associated Motor Insurers Limited (AAMI) has listed driver's fatigue as one of the major contributions to road crashes. 11% of Australian drivers are reported with fatigue while driving and contributed to the statistic of road crashes [2]. In UK, failed to look properly which also considered under visual distraction, was the most frequently reported contributory factor and was reported in 42% of all accidents reported to the police in 2011 [4][21]. A few number of research papers were also discussed about this visual distraction with several types of detection method. Some of the research papers were using physiological measurements and some were using performance measurements, such as in [38] the research study has used mouth movement to detect driver's fatigue. The mouth's height and width were used to determine whether a driver is sleepy or not. The movement was detected by using a dashboard-mounted (Charge-Coupled Device) CCD camera. The experiment to collect the data was completed in a lab setting. Research in [26][27] and [28] were also studied about driver's visual distraction. In [28] they have conducted the experiment with the usage of cell phone on a driver. The research has mentioned that 60% of drivers are likely to use cell phone while driving either to answer the phone call or texting. By using cell phone while driving is basically initiated from a manual

distraction. The driver's hand will be off the steering wheel and later leads to the visual distraction when the driver's eyes are no longer looking on the road. The research also has suggested that the use of cell phones disrupts the driving performance by diverting the driver's attention toward an engaging cognitive context other than the external environment which immediately associated with the driving.

Manual distraction is another type of driver distraction which is always related to the driver's vehicle handling. This usually occurs when a driver is manipulating something in the vehicle other than the wheel. Several times, manual distraction is related to the usage of a cell phone while driving. In [4] and [21] has reported that in year of 2011 in Great Britain, 374 accident reports were reported caused by the drivers using the cell phone while driving. Usually, by having a manual distraction the drivers will delay their respond towards any event while driving for example, driving too slow because only using one hand to handle the steering wheel and late turn when supposedly to make a turn.

Visual and manual distractions are mostly easy to be detected. The movement on the eyes and hands are easily to be detected because they are observable easily. These types of distraction might occur when a driver is driving either in rural, urban, intermediate, or even in a car park area. Visual distraction is the major distraction to occur in driving and it is also the easiest type of distraction to be detected. There are a number of measurements which can be used to detect a driver's attention or distraction visually: blinking rate and frequency, PERCLOS, eyelid movement, pupil diameter and gaze direction. Yet, cognitive distraction is still a new type of distraction and has just been studied recently. Thus, to detect cognitive distraction, especially on a mobile vehicle and real driving environment is not an easy task to carry out. Perhaps there are only few available options and measurements that have been identified as a method to detect cognitive distraction. Visual distraction is really straightforward to be detected. It reasonably occurs when a driver looks away from the main road. Simply, it occurs when the driver's eyes are off the road. Unlike visual distraction, cognitive distraction occurs only when a driver is thinking or processing information that is not related to driving safety issues or current task of vehicular control. A few studies have been conducted to detect driver cognitive distraction [24][36][45] and [46]. Most of the detections were made through physiological measurements such as gaze rotation, saccade, eye's fixation, blinking frequency and duration, pupil diameter and head rotation. Some of the studies were also combined the physiological measurements with

performance measurements at which they were looking at keeping lane position and steering wheel angle. However the data collections in the previous studies were mostly from lab simulation. This is the major difference between those existing study and this study. Data collections in this study were collected from a real traffic. In this study data cleansing process was done by deleting the dirty data. Data transformation method is used for the cleansing process. Data transformation allowed the mapping of the data from its original format (continuous) into the format expected (discrete) by the applications used in this study; Bayesian Network Toolbox (BNT). Data transformation allowed conversions or translations functions as well as normalizing the values in order to confirm minimum and maximum values.

Initially the data is audited by checking for any abnormalities in the data like missing values and incorrect data. Then the workflow specification is taking over by removing all the dirty data. The correctness of the removal is then checked and verified in workflow execution. Number of total data from one participant with other participant is to ensure be similar after the cleansing process is taken. Data analysis for this study was conducted by using BNT and WEKA application. BNT was basically used for Dynamic Bayesian Network (DBN) and Static Bayesian Network (SBN) developed for this study. WEKA whereas was used for Support Vector Machine (SVM), AdaBoost and Logistic Regression (LR).

Hence, it is a good direction if a few more identification can be recognized to detect a driver's cognitive distraction. As mentioned earlier, distraction in driving can be categorized into three general types: visual, manual and cognitive. A report prepared by the American Automobile Association (AAA) for Traffic Safety categorized driver attention into five categories [16]:

    a) Attentive

    b) Distracted

    c) Looked but did not see

    d) Sleepy or fell asleep

    e) Unknown

Attentive and distracted are opposite to each other. Attentiveness happens when a driver really puts focus on the road, physically and mentally. Attention is a part of focus, concentration and a component of intelligence. Attention usually only lasts for a few seconds. Tests suggested that attention can last up to 90 seconds but 30 seconds is regarded as the maximum time for focusing attention. To gain attention while driving, the driver's hands must totally be on the steering

wheel at a proper angle, eyes must be looking forward on the road and are always alert and aware of the surrounding environment and mind is always thinking about safety. Drivers must always be in a conscious state of mind in order to always stay alert of any incoming hazards. Attention can be influenced by either sensory factors (exogenous) or cognitive factors (endogenous). Exogenous might come from either visual or manual type of distraction, when drivers' eyes and hands are not in a proper position to maintain safety.

"Looked but did not see" or Looked but failed to see (LBFTS) is a normal phenomenon to happen among drivers. An investigation made at Melbourne Metropolitan area found that 30% of fatal accidents occurred at intersections/junctions [25]. Police investigated and discovered that the contributory factors in these accidents came from the offending drivers, who claimed to have looked in the appropriate direction, but failed to see the other vehicle. A closer look at the evidence suggests that the real cause of this type of accident may lie not in the limitations of the low-level visual system, but in the more cognitive system involving attention that crucially affected the driver's expectations of what they are likely to see. In [25], have asked the drivers to mention anything that attracted their attention while they were driving, and found that 30% to 50% of driver's attention was given to aspects of the environment which are not related to driving, especially billboard advertising. Only 15% to 20% attention was given to road signs, which was deemed insufficient to ensure that all or even most traffic control devices were noticed.

Sleepiness or falling asleep is another major problem in driving safety. Sleepiness is usually related to fatigue. This "fatigue" is an ill-defined phenomenon. Sleep has a number of stages that are associated with different patterns of activity produced by the brain. Sleep can fall into two main types: REM (Rapid Eye Movement) sleep and non-REM sleep. However, in the context of driving, very often drivers experience micro-sleeps with a very brief period lasting between 0.5 seconds to 1.5 seconds, during which they appear to be asleep, both behaviourally and physiologically [25].

2.3   Driver Cognitive Distraction

As mentioned above, driving distraction can be categorized into three categories: visual, manual and cognitive. From these three types of distraction, cognitive distraction is the most dangerous for the following two reasons: (a) it is the most difficult type of distraction to be detected since it is an internal distraction and (b) it can initiate a failure to operate the vehicle properly and may lead to the other two types of distractions, visual and manual.

Cognitive distraction is a situation that might lead or shift a person from putting his/her attention of doing something. Compared to visual distraction, cognitive distraction is more difficult to measure and detect. Cognitive distraction is an internal distraction and it is impossible to be observed directly from external behaviour. Unlike visual and manual distraction, cognitive distraction does not have a direct indicator to show when the mind is off the road. Visual distraction is usually related to cognitive distraction. A driver is usually distracted cognitively after he/she has had a visual distraction. It is possible for a cognitive distraction to bring more hazards to a driver than visual and manual distraction to do the same. For instance, cognitive distraction can delay a driver's response, being slow in applying the brakes during emergencies, missing more traffic red lights and being unable to stay at a safe distance from a leading vehicle. Thus, when a driver is not thinking about anything related to his/her current safety, it can undermine the driver's general safety during a critical event.

It has been studied that driver's cognitive distraction mostly happens when drivers operate or handle the IVIS. Performing a manual operation in a vehicle can be considered as a secondary task. In most laboratory tasks, the division between primary task and secondary task is possibly clear. Unfortunately, this is not possible in a real driving setting. In traffic, a driver's behaviours are quite often related to either primary or secondary tasks. Monitoring rear traffic through the rear mirror can be considered as a primary task and not a secondary task. Primary and secondary tasks can be confusing to be distinguished especially when the tasks are integrated. In a car-following (tailgating) event, primary tasks are restricted to speed and lateral vehicle control. Secondary tasks are non-continuous tasks, such as keeping headway of a lead vehicle and looking into the mirrors at intervals [26]. In chapter 4, secondary task is defined under driver

distraction measurement. It is related to an activity like checking the mirrors, following a leading vehicle and etc.

Drivers might also be distracted cognitively when they briefly glance away from the roadway. A visual distraction on the drivers can lead them to think about something that is not related to the driving task. As described before, when drivers are "busy" reading an advertisement on the roadway, they tend to be at a high risk of failing to identify, recognize and process crucial information on traffic sign boards.

A cognitive distraction caused by using auditory email systems, performing mathematic calculation or holding hands-free cell phone conversations can delay driver response to hazards by an average of 130ms [27]. Drivers react slower to the leading vehicle's braking event and miss more traffic signals when they perform tasks with cognitive distraction while driving.

Cognitive distraction can divert attentive resources from driving and it is associated with two effects:

a) Cognitive distraction can undermine the process of consolidating attended information. Results in [28] has shown that cognitive distraction which is associated with the usage of cell phone conversations has impaired the implicit perceptual memory and explicit memory in recognizing items from the road environment. The results suggested that some visual information was not being processed appropriately when drivers were cognitively distracted and it leads to undermine the driving performance in detecting targets across the entire visual scene.

b) Cognitive distraction also changes the pattern of eye gaze distribution. This point will be discussed in chapter 4 in this thesis.

Nevertheless, these types of distraction are usually linked to the driver behaviour. Driver behaviour plays an important role in driving safety. It is the driver's task to regulate his/her activities and adjust them to deal with the problems arising from the interactions between different components or elements in his/her surroundings. A number of functions, especially cognitive ones have to be performed in order to negotiate with these difficulties. In [29], they have found that whenever a driver failed to encounter his/her major malfunctions within the system; it can lead to a system malfunction and prevent the driver from attaining his/her regulating objective. This results in a functional failure and is commonly termed "human error".
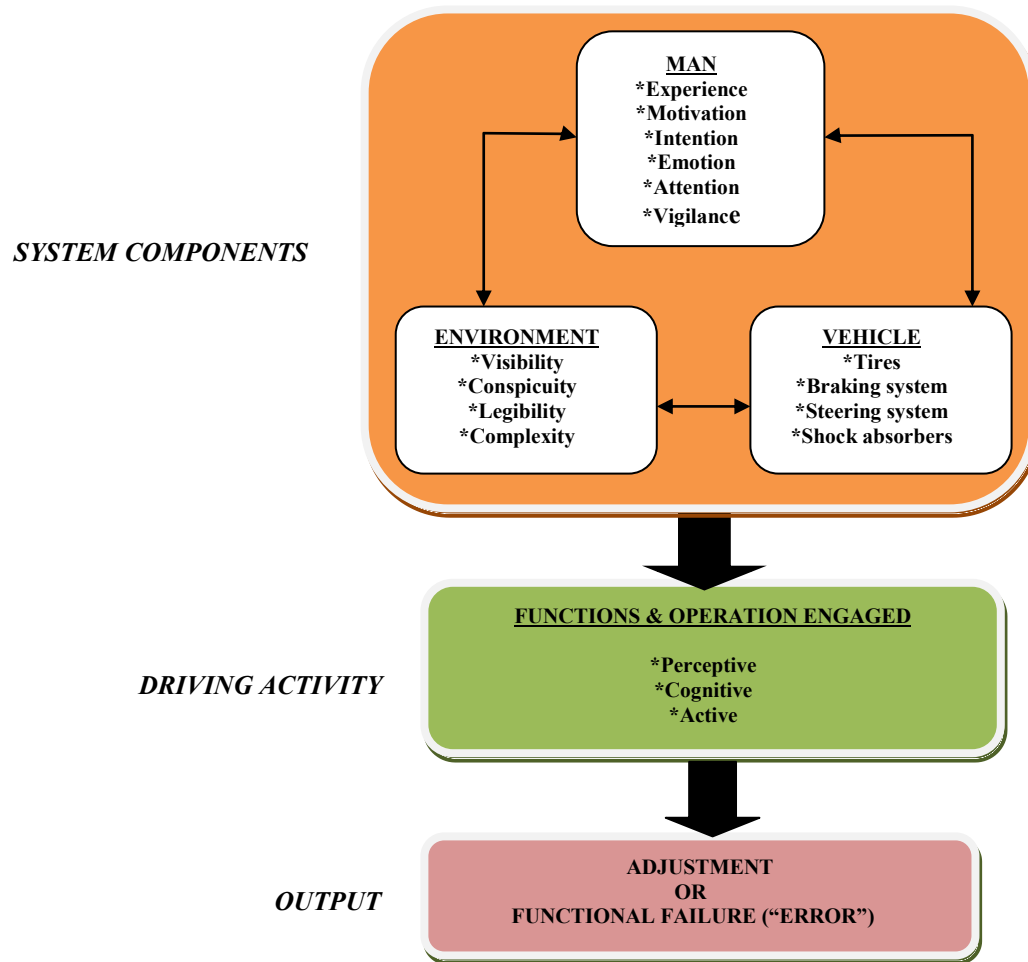
Figure 2.3: Driving system components

Figure 2.3 showed that human/driver is the most important element in the system, human behaviours involved with emotion, experience, attention, intention, vigilance and others are required to operate the driving activity. A particular attention is required to the role of the driver. As the world is far too rich for human perceptive capacities, in order to adapt to this complexity, information is determined by filters [29]. Cognitive filter is a final filter, yet the least known filter. Nevertheless it is actually the most important filter in a driving system. Cognitive filter will impede integration of information and the knowledge associated with them. Failing to do this might lead a driver into the state mentioned before, LBFTS. Research by Department for Transport [30] has found that human perceptual errors were the predominant contributory factor, where it contributed 46% of all categories of contributory factor recorded. Within this perceptual factor, LBFTS was the third most frequently recorded contributory factor, with 17% accounting

from all other factors. Inattention was the highest one with 28%, followed by misjudging others' path or speed with 21%. Surprisingly LBFTS was recorded almost 23% more frequently happening in daylight rather than in darkness. The information suggests that LBFTS is actually derived from failures of attention, perception and cognition rather than being of sensory origin. Almost 17% of daylight accidents' drivers were recorded having LBFTS. There are two types of errors that meet the criteria for LBFTS: (a) when a driver searches the traffic environment over-selectively, for example looking at a large vehicle but overlooking pedestrians or cyclists (b) when a driver searches for features which distinguish hazardous from non-hazardous objects, for instance location, orientation and speed limit but fail to integrate those features into a coherent danger image.

This state of mind causes some of the drivers to lose their attention and fail to recognize their surrounding and sometimes fail to respond to them. In the experiments of this study, it happened several times. Some of the drivers in the experiments mentioned that, they lost their focus due to the distractions from their surroundings or they failed to respond to the given question appropriately because they were focusing too much on the shuttle bus moving in front of them or the incoming vehicle. They were struggling with a decision whether to overtake the bus or to keep tailing behind it and some of them even failed to respond appropriately when stopping before a zebra crossing line or driving onto road humps/bumps along the route due to the diversion of attention. The drivers were more concerned about the bus in front of them as they assumed it was the most hazardous object and this made them neglect other surroundings. The best situation to describe this state of mind is during lane changing and it is well described in [31].

This is related to what was mentioned in [32] that cognitive distraction is able to diminish the driver's capacity to adapt their behaviour to changing workload. This usually happens when the workload is already high and tasks are time constrained. Therefore, the driver tends to overlook incoming or possible hazards. Like the example given above, overtaking may have a higher task demand than following a leading vehicle. Therefore, drivers may have a lesser capacity to adapt during overtaking and one study found that drivers drove faster than usual when they are distracted by a cognitive task. This shows that drivers failed to adapt and compensate appropriately for second cognitive demands. Their minds were overloaded with both the tasks given during the experiment and the decision to overtake.

Some drivers tended to increase their distance with the leading vehicle in the car-following scenario when they were engaged cognitively while simultaneously demanded with the secondary tasks.

Cognitive processing monitors the environment to maintain awareness of other vehicles and allows for decision making on when to change lanes based on its current mental state. It is supposedly to have sufficient amount of attention in use to occupy driving control, situation awareness and non-related driving tasks.

> *IF*   my current goal is to check for a lane change
>
>   & my current speed is less than the desired speed
>
>   & there is a lead car blocking further acceleration
>
>   & there is enough room in the other lane
>
> *THEN*   initiate a lane change for passing

If one of the condition in the IF statement above is not fulfilled, the driver can be in an unsafe state and this is due to the cognitive loading which requires more attention for one single purpose at that time and ignores the others.

It is also mentioned in [33] that a theory about cognitive mechanisms in guiding a deliberation process is involved with making decisions under uncertainty. Drivers under uncertainty usually require a lot of cognitive loading before a decision can be made. They need to integrate both information from the external environment and information from the individual's associative memory as determinants before any possible action can be evaluated in the deliberation process and then to be taken.

TRACE project [29] has made a classification towards human functional failures in road accidents as shown in Figure 2.4. The aim for that classification model is to sort out the different orders of malfunction phenomena which intervene in the genesis of driving accidents. Surprisingly, from the classification model made, it has been found that from delineation of functional failures found in depth from accident data, in overall, overstretching the cognitive capacities is one of the three types of driving capacities degradation. The other two are: (a) the loss of psycho-physiological capacities and (b) the alteration of the sensorimotor.

Overall failures or also called global failures correspond to the degradation of the whole functional chain, and the outcome is the loss of control of a situation. These cases in making the whole functions to necessitate driving seem to have been undermined in the mechanism and may lead to an accident. At this level, it actually refers to the general capacity of the individual to manage the situation encountered, including when the information is collected, treatments to be operated, decisions to be set and actions to be undertaken. It depends on the driver's ability to handle a given situation. Global failures are believed to be found in the parameters or factors indicating a psycho-physiological and cognitive state of the drivers scarcely compatible with the functional demands required during the driving activity. These factors can be caused by different things like fatigue, alcohol/drugs, fitness and others.

Failure in psycho-physiological capacities is referred to a loss of awareness of the driver resulted from being ill, falling asleep because of tiredness or high blood-alcohol level in the body. This failure obviously causes an interruption in the driving activity as a whole and results in a failure to control the vehicle too.

Alteration of the sensorimotor happens when a driver is disorganized. This is typically resulted from a high degree of intoxication whether due to alcohol, medicine or drugs. This impairment reveals that the drivers have the most difficulty in their journey in which they seem to be unable to manage at any functional level.

Overstretching cognitive capacities impairment is linked to a general loss of skills related to the driving activity. Drivers will find that their abilities are overstretched at the moment they encounter a difficulty. As a result, this sometimes leads to absurd manoeuvres.
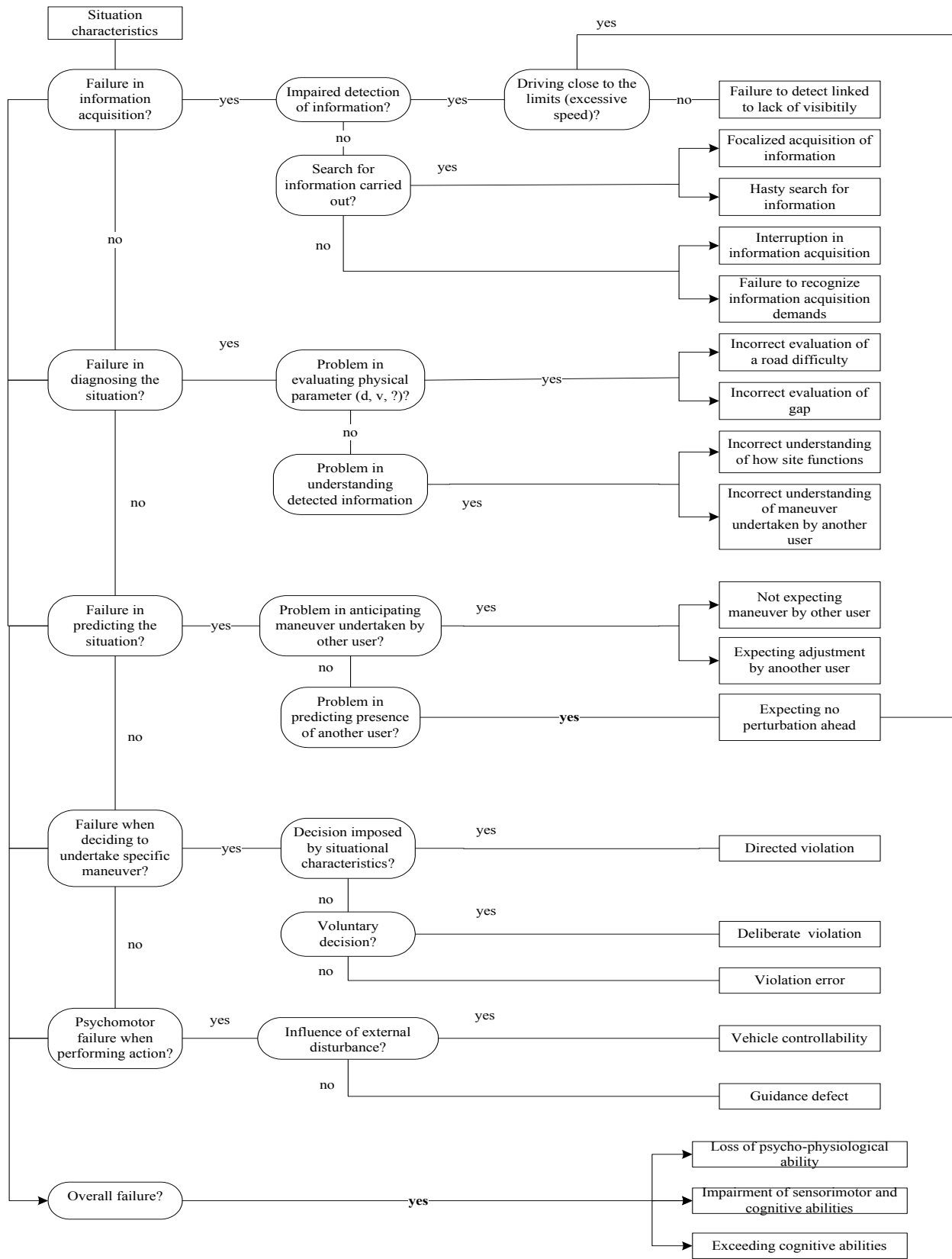
Figure 2.4: Model for human functional failures classification [34]

Also mentioned in [34] is that cognitive mental workload is relatively related to all processors including visual and manual. A new type of cognitive activity modelling which is used to account workload effects in human performances was developed by Kieras and Meyer. The Executive Process Interactive Control (EPIC) is a composite model which was developed in 1997 and is usually represented in the form of computer programs. The programs are applied to the tasks and are used to predict various aspects of human performance. EPIC was developed to provide a comprehensive computational theory of multiple task performance. Like the driving activity, which requires multiple tasks at once, this EPIC model is suitable to explain how a cognitive processor handles other processors at a time. Output from this model is intended to provide quantitative prediction of the mental workload levels and performance effects. The model's adheres to the assumptions inherent in the stage of theoretic models of human performance in [35]. The EPIC model essentially uses production system modelling techniques to emulate operation of the main cognitive information processor. The model assumes that there are several separate perceptual or motor processors with distinct operating characteristics- vocal, manual and ocular. The architecture for EPIC is shown in Figure 2.5 below. In multitask situations, task demand and performance decrement is greater when:

a) tasks are more complex

b) when the person has limited prior exposure to the task

c) when two concurrent tasks are similar

d) when the two tasks make demands on the same cognitive structures

e) when the individual's reaction to task loading is variable and differs upon a number of factors like nervous system activity, fatigue, alcohol, stress and etc.

Thus, when the tasks required excessive cognitive workload, it makes the driving activity as whole a failure and causes accidents. Cognitive distraction is subtle, inconsistent and easily affects the driver's behaviour.
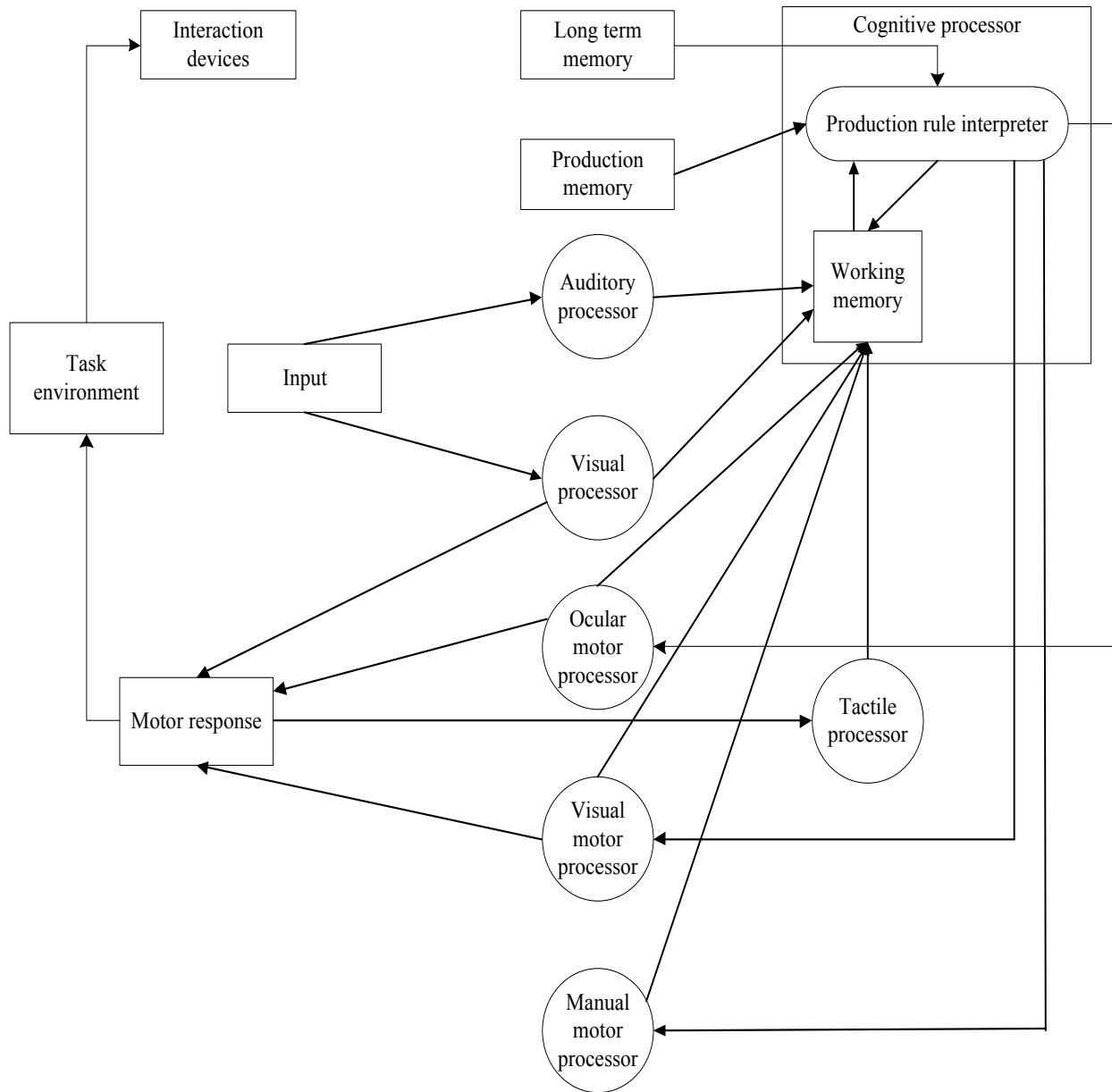
Figure 2.5: EPIC model [35]

As mentioned before, there is no exact indicator and measurement for cognitive distraction. It may be useful to determine cognitive distraction state from multiple performance measures in a relatively long period of time. Many researchers nowadays, use driving performance measurements like lane position keeping and steering variability as their indicators for driver cognitive distraction. However, in this study, driving performance measurement is not used because they often reflect the consequences of the distraction. Besides, this type of measurement usually mitigates the distraction too late after the driving performance has been degraded. Thus, identifying driver's physiological movement which has a close relationship with a driver's attention can really help to identify the driver's distraction. In [36] a long period of time to indicate cognitive distraction is around 30 seconds, which is an appropriate time to produce a good detection algorithm.

2.4   Lips and Eyebrows

In this study, new features are proposed to detect driver cognitive distraction. Lips and eyebrows are chosen as new parameters because they are potentially easy to measure and they are also capable of improving the accuracy and sensitivity rate in classification algorithms to detect driver cognitive distraction. To prove the ability of these new features, an initial experiment was conducted with faceLAB in a lab simulation. Mouth movement consisting of mouth width and mouth height and eye movement consisting of eye width and eye height from both left and right eyes were the parameters captured in this initial experiment. Other parameters like head movement, gaze rotation and pupil diameter were also captured and used. Correlations between these features found a strong relationship when a driver is cognitively distracted. Based on this positive result, the experiment was then extended to a real time environment. The second experiment only used faceAPI cameras and the main features to be captured are lips and eyebrows. Head rotation again has been retained in the second experiment. However, eye data and nose data have been eliminated because both left and right eye landmarks were only captured at the beginning of the tracking and their movements have not been tracked during the experiment. Nose landmarks are not useful in this study because nose movement is very small. Thus, this data was ignored.

A few papers have studied about lips, eyebrows and distraction. In [37], they have found that lips movement can be used to detect a driver yawning or having a conversation while driving. A very significant finding can be found in that paper on how lips can be used to detect driver fatigue. Eyebrows also can be used to determine whether a person is concentrating or focusing [38]. They have found that eyebrows are significantly and efficiently useful for the human perception of focus. However, their experiments were not conducted within a driving environment. Nevertheless, they have proven that perception of focus and eyebrows are correlated to each other.

When a person is cognitively distracted or thinking, lips and eyebrows are the two obvious moving features on a person's face. It is easy to distinguish between a normal facial expression and a thinking facial expression because movements on eyebrows and lips can be obviously seen. Figures 2.6 and 2.7 below were taken in an office and the subject was asked to give a normal expression and a thinking expression [39]. Both figures simply distinguished between non-distracted facial expressions from distracted facial expressions in a less dangerous environment.

Figure 2.6: Normal/Non-distracted face
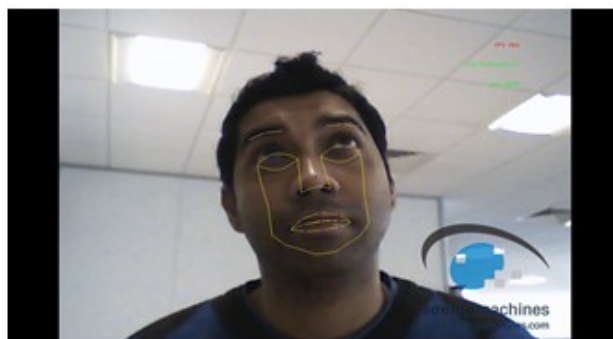
Figure 2.7: Distracted face

People in cognition, often furrow or perplex their brows when concentrating and also purse their lips when conducting mental searches [40]. The thinking facial expression is usually with the eyes looking up or down and it is usually combined with the use of biting the lower lip as an addition to the idea that a person is struggling to think of something. A person with cognition usually will have behaviours like in Table 2.1 on his lips and eyebrows,

Lips	Eyebrows
Moving lips- Lips moving in the shape of words but without making sounds mean the person is thinking. Rolling in the lips sign of uncertainty and accompanying with lowered eyebrows.	Furrowed brow - Concentration may also be shown in the forehead as the eyebrows are brought together as the listener seeks to hear and understand the other person.
Protruding lips- Both lips pressed together and pushed out generally indicate doubt. If the finger touches them, it may indicate internal thinking.	Middle-lowered eyebrow- When the middle of the eyebrows is pulled down so they slope inwards, this shows that the person is intense concentration.

Table 2.1: Lips and eyebrow behaviors when a person in cognition [41].

Ekman and Matsumoto [42] mentioned that facial musculature is fairly unique. They conclude that somatic muscles in the body are attached to one side to bone and the other to skin making facial movements to be specialized for expression. Face is one of the body parts where some muscles are not attached to any bone at all. For instance, the orbicularis oculi is a muscle surrounding the eyes and orbicularis oris is the muscle in the lips. Thus, there is no one to one correspondence between structure and function in some facial muscles, meaning that, facial expression is only possible with a group of facial movements. Therefore, lips and eyebrows in this study are mainly used and combined with other proven features like head, gaze and pupil to bring better description and prediction on driver cognitive distraction detection.

2.5   Conclusion

Since 1920s, when the first radio receivers were introduced as an option for cars, driver information overload as a source of distraction has been studied widely for traffic safety [30]. After radio receivers, mobile phones were the next device of distraction and this again has received a widespread attention to many people especially those involved with the transportation system. Furthermore, this issue becomes more critical when an increasing number of advanced driver assistance systems, IVIS and nomadic devices are introduced to entertain and assist the drivers.

Driver distraction is an important safety problem. Between 13% and 50% of crashes are attributed to driver distraction. The increasing use of IVIS has exacerbated the problem by introducing additional sources of distraction. Enabling drivers to benefit from IVISs without diminishing safety is an important challenge. Blaming the manufacturers of all IVIS products is not the solution to all the road crashes and accidents. A solution to this problem would be to work together and make a proper assessment to all the accidents which had occurred and find the real causes. As mentioned above, driver distraction is the main attribute to accidents, thus a proper design of road safety and vehicle safety must be learnt and studied.

Distractions for drivers can be categorized into three major categories: visual, manual and cognitive. Visual is when the eyes of the driver are off the road. Manual is when driver's hands are not on the steering and cognitive is when the driver's mind is not thinking about the safety of road and vehicle. From all those three types of distraction, cognitive distraction is the most dangerous and most difficult type of distraction to detect as it is an internal type of distraction. Many driving models like EPIC model and Human Functional Failures Classification model have mentioned cognitive distraction as one of the main contributing factors in the driving activity. Mainly, driver's cognition is very important in this sense. The lack or excess in cognitive load may bring the driver into an accident. Thus, an advanced appropriate method and technique are required to capture this type of distraction.

Nowadays, there are many road and vehicle safety projects focusing on cognitive distraction like TRACE project (Traffic Accident Causation in Europe) and SAVE-IT (SAfety VEhicle using

adaptive Interface Technology) project which develops a vehicle incorporating adaptive interface technology to mitigate driver distraction and evaluate its safety benefits [43].

There are a few numbers of available parameters or features for cognitive distraction detection. Eye movements are the most popular one. In this thesis, new features including lips and eyebrows are introduced. Lips and eyebrows are easier to be detected than other type of features like gaze rotation and pupil because they can be seen obviously and do not require special cameras like infrared. They are also reliable because their movements are similar every time a person is cognitively distracted. Furthermore, when a person is in cognition the shape or movement on lips and eyebrows are different with other types of feeling or expression.

Chapter 3

Experimental Setup to Assess Driver Cognitive Distraction

This chapter discusses the experimental setup to assess driver cognitive distraction. Three different experiments have been conducted. Every experiment contained two main groups of data: distracted/task and non-distracted/control.

3.1 Background

Several papers which studied driver distraction have explained methods and ways for driver distraction data collection. Some of the studies were done by constructing simulations in a lab and some studies were done by collecting data on a real road environment. Each method and approach has its own advantages and disadvantages.

All experiments in this study used faceLAB and faceAPI machines. faceLAB machine uses infrared cameras to captures features from the eyes and head. Features captured from faceLAB include gaze rotation, blinking frequency and duration, head rotation and position, pupil diameter, eye movement, mouth movement and many more. On the other hand, faceAPI is a machine with the ability to capture features like lips, eyebrows, head rotation, head position and nose. faceAPI only uses a normal web camera to capture the facial features.

In this thesis, the first experiment was conducted with faceLAB machine in a lab setting. The second experiment was conducted with faceAPI toolkit on a real road environment and the final experiment was executed with the combination of both machines: faceLAB and faceAPI. The final experiment was also conducted on a real road environment. Both machines captured different features on the driver's facial expression.

Data collected from the real road environment has a bigger impact on this study than that of simulation because the data was realistic and affected the contextual information. At the end of this chapter, a comparison of experiment setup between existing studies and ours about driver cognitive distraction detection is made. Different experimental setups provide different types of result.

3.2 First Experiment

The first experiment was conducted in early February 2010. It was only a simple lab simulation where the faceLAB equipment was set up in a lab. Six participants were involved and they were asked to watch a screen displaying video footage taken around Loughborough town in Leicestershire, United Kingdom and were distracted by some external events simultaneously.

3.2.1 Experiment overview

The experiment objective was to prove whether there is a good correlation between mouth and eye movement when a driver is cognitively distracted. The experiment contained two types, with and without the audio tasks, and was conducted in a laboratory setup. Six participants were voluntarily involved in the experiment. A video sequence of a real road environment around Loughborough town, Leicestershire, UK was shown to all participants. As pointed out by the American Automobile Association (AAA) in their report about passenger vehicles and road safety, there are mainly 13 different types of distractions [16]. However in this experiment the participants were distracted with only some specific types of distractions such as trying to recognize a pedestrian or talking with other passengers in the vehicle, listening to music or memorizing a seen object or an event from the video. Mostly, the distractions occurred from verbal distraction where they were asked a few questions. Cognitive distraction was detected within a few seconds before the participants responded to the questions. It took around 10 seconds to 50 seconds for them to give a response to a question.

3.2.2 Participants details

All the 6 participants are experienced drivers. On average, they have been driving for five years. Therefore, all the participants are familiar with the driving conditions and environment. The participants are aged between 28 to 38 years old and all of them were familiar with the route shown on the video. Summary about the participants is shown in the Table 3.1 below:

Gender	3 females 3 males
Age	28 to 38 years old $\mu = 30.67$ $\sigma = 3.45$
Occupations	2 are Engineers 4 are PhD students
Driving License	1-UK driving license 5-International driving license

Table 3.1 : Participants details-first experiment

All participants are not from UK. Therefore most of them hold an international driving license. All participants are from a country with right-hand driving. Thus, all the participants were familiar with the structure of the driving activity. At some time, they only faced problem in recognizing the signboards on the road, since not every signboard is similar to their countries. Some questions in the video were based on the signboard recognition.

Five participants have been driving in UK and have their own cars. Only one participant does not have a car but sometimes drives his friend's. Each participant showed different levels of driving frequency.
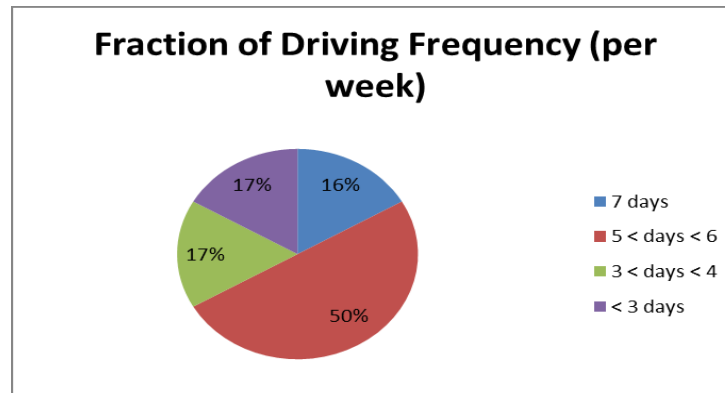
Figure 3.1: Driving Frequency (per week)

More than 50% of the participants drive their car between 5 to 6 days per week because they need to drive to their work places.

3.2.3   Experiment setup

Two experiments were conducted for each participant. The first experiment was a controlled experiment, where the participants were only required to watch the video and there was no kind of distraction. The second one was executed with several types of distractions/tasks generally with two types of task: with audio and without audio. Both experiments were conducted to trigger cognition from the participants. Generally, the distractions were all about the video displayed on the screen. Since cognitive distraction is hard to be triggered through this experiment, both visual and auditory senses were tested from every participant. Figure 3.2 below showed the experiment setup division:
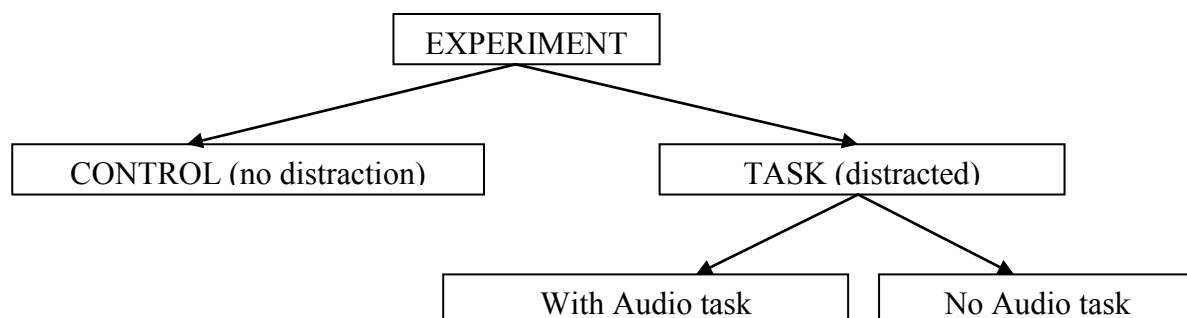
Figure 3.2: First Experiment setup

The video was 8 minutes and 49 seconds long and every participant watched the same video twice. The first watch was for the control and the second watch was for the task. The total time for every participant was 17 minutes and 38 seconds. Therefore, the total data collected from this experiment from all 6 participants was 105 minutes and 48 seconds in length.

The equipment used in this experiment were:

a) faceLAB equipment (infrared cameras and a laptop with relevant softwares)

b) Desk and chair

c) PC Desktop

d) PC screen monitor

e) Video file: video of a real road environment to be displayed on screen

(The video included road, other vehicles, pedestrians, cyclist, signboards, and traffic lights.)

f) Audio file: recorded radio on air, played on the speaker

(The audio recorded a song played from the radio channel, Heart.fm, and a short talk from the deejay.)

g) Speakers

h) Video camera to record the experiment

Speakers and the audio file were used during the distracted experiment with audio task. Figures below showed some pictures taken for this first experiment setup. Participants were required to sit and watch the video twice as mentioned above. The video and audio were played continuously. However the video was started a little bit earlier than the audio because the audio was representing a recorded radio which will be the audio task to the participants.

Figure 3.3: faceLAB equipment for the experiment
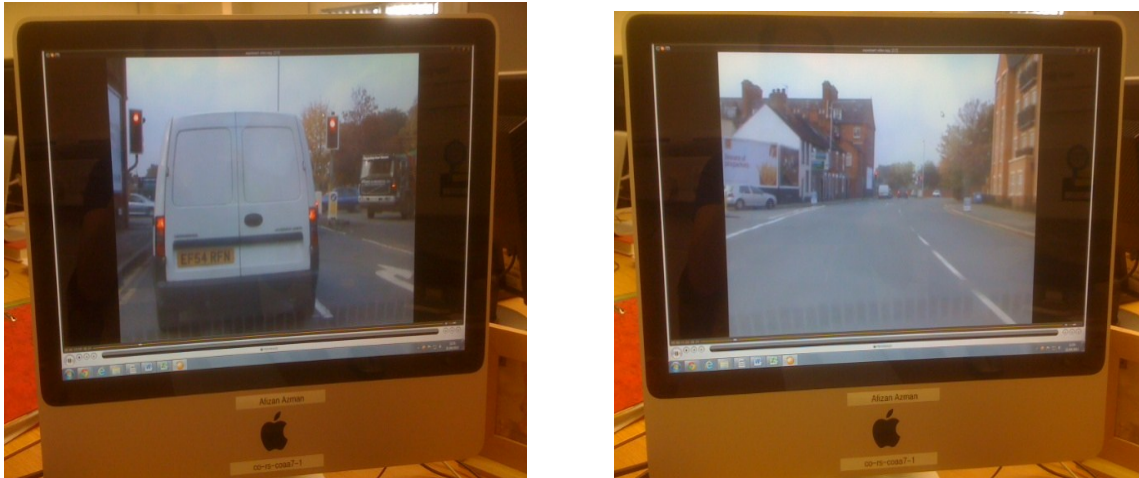
Figure 3.4: Video screens

Basically the experiment was conducted in the sequence below:

a) A video (recorded the road and traffic information from a car on a real road) was displayed on screen.

b) Participant was asked to watch the video

c) FaceLAB cameras were calibrated and the equipments were operated by the experimenter.

d) First experiment started with the task/distracted experiment. Two sets of task experiment were run. With and without audio tasks.

e) A few questions were asked to the participants with the assumption that these questions can trigger the participant's cognition. While answering questions about the video, participants were also asked about general questions like arithmetic or general knowledge questions. This was done to test their situational awareness.

f) Once the questions have been asked, the annotation button was pressed from the faceLAB machine to indicate the time frame for when the question started, when the question stopped, when the participant's facial expression changed and when the participant's facial expression returned to normal.

g) The same video was re-watched by the participants for the second experiment (control/no distraction). This set of experiment was used to compare the participant's facial expression during the control set of experiment and the task experiment.

Several steps have to be followed when using faceLAB. The flowchart below defines the steps involved when occupying the faceLAB cameras. faceLAB cameras need to be calibrated before

it could be used. To calibrate the camera is the only existing hassle of using faceLAB equipment. Since faceLAB cameras are infrared, they do not have any problems with lighting/exposure. Usually, even with insufficient source of light, the infrared cameras are still able to track and capture the facial expression on the drivers. Three major steps are defined when using faceLAB: system initialization, tracking process and storage. Data stored in faceLAB are stored in five different data output:

World output data - includes data from the world model such as the objects' name and index which intersect with the head and/or gaze vectors; including screen intersection data.

a) Timing output data - includes data such as the frame number; time stamps; delay; and annotation labels.

b) Eye output data - includes eye data such as the closure of the left and right eye; blink, PERCLOS and saccade data; the quality of the gaze and the angle direction; as well as the position of the eyeballs and pupils.

c) Features output data - includes image feature data such as the pupil positions in both camera images; and the rectangle feature area for the face, eye and mouth.

d) Head output data - includes data such as the position and rotation of the head; the head model quality; and the eye ball centre position relative to the head.
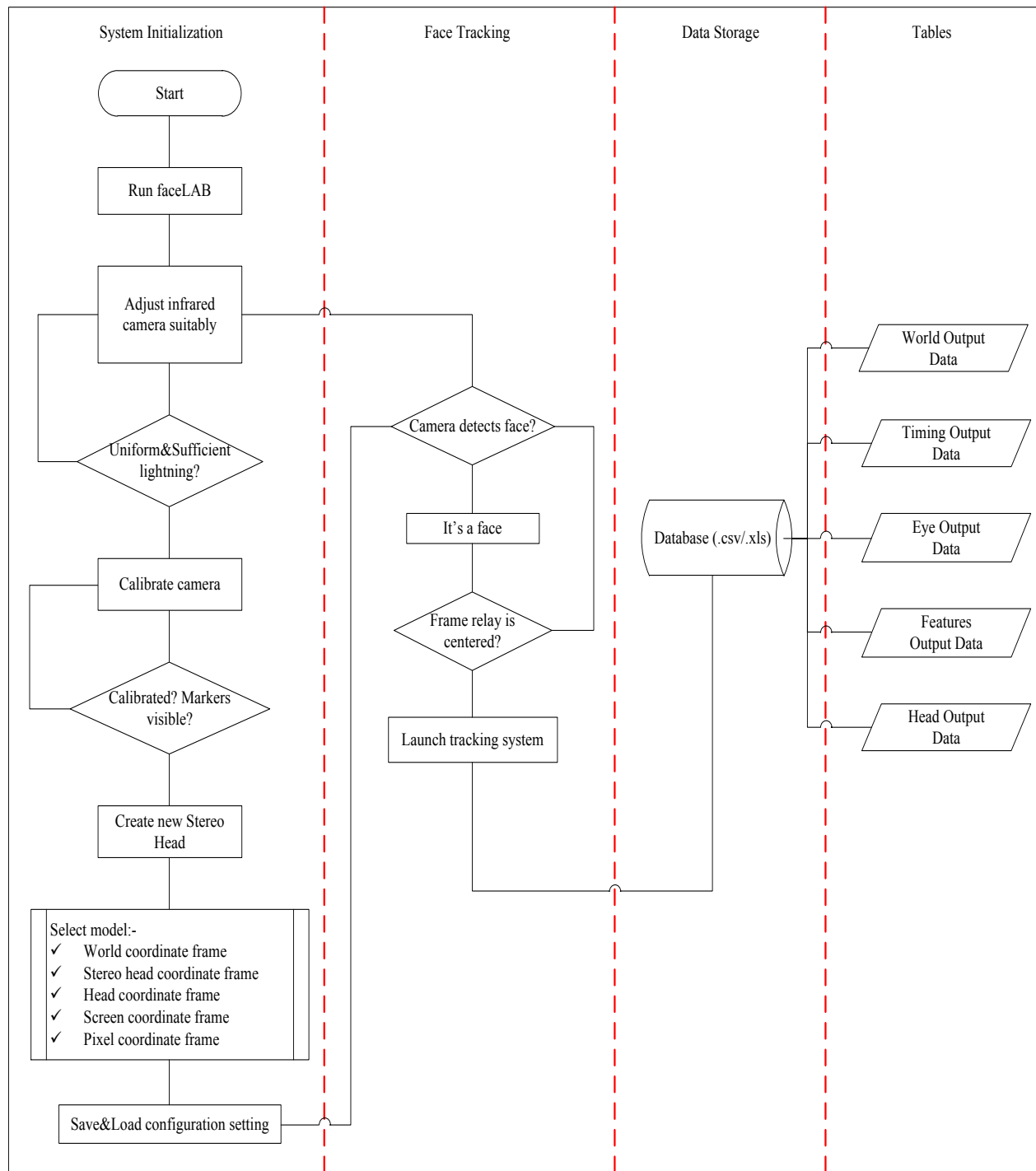
Figure 3.5: faceLAB flowchart

3.2.4   Distraction setup

As mentioned before, there were two types of task contained in the experiment: with audio task and without audio task. Questions in this experiment which were either with or without audio task have been categorized into the following groups: Open questions, Radio related questions, Safety related questions and Simple Arithmetic questions. In the audio task experiment, while watching the video, the participants were also required to listen to a recorded streaming radio. This was done to create a somewhat simulated real driving experience. However, in the non-audio task, the participants only watched the video and were only distracted by several questions from the other three different categories. The result from [44] studying on devices comprehension and eye tracking has found that there is a strong evidence showing a consistent coordination between questions being asked and eye movements. The *annotation_id* button from FaceLab was used to annotate the time frame number when a distraction started and ended. The experiments were based on Yulan and Miyaji experiment setup [24][36][45]. The questions for this experiment can be divided as shown in Figure 3.6:
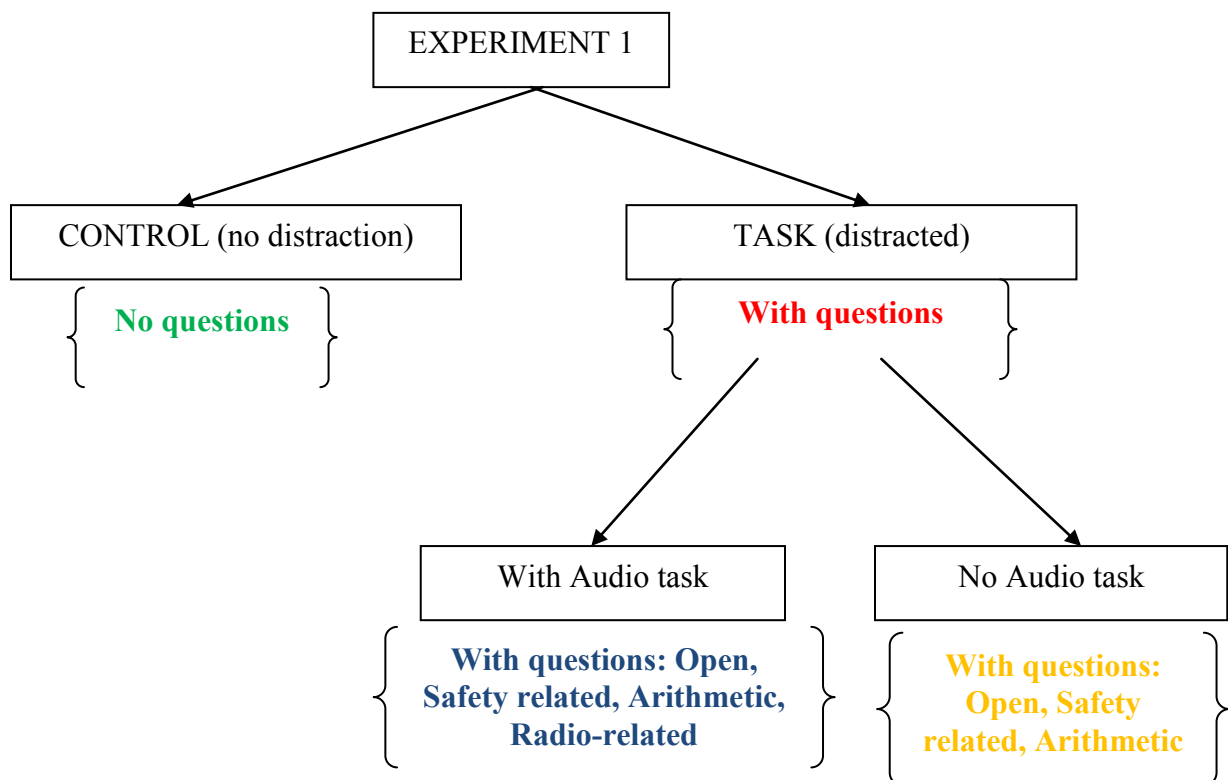
Figure 3.6: Question division

In the distracted task (with audio) the experiment was conducted with an audio file being played to the participants. Participants were required to listen to a recorded radio station (Heart.fm). They were required to listen to the song played while simultaneously watching the video on screen. The experimenter distracted the participants with several questions, either based on the audio played or the other three categories of questions.

In [24][36] experiments, they distracted their participating drivers by asking the Auditory Stock Ticker Exchange questions and in [45], their participants were bombarded with arithmetic and conversational tasks. A meta-analysis of 23 studies in [27] has found that cognitive distraction caused from using auditory e-mail, performing math calculation or holding hand-free cell phone conversations has delayed driver response to hazards by an average of 130ms. These findings were used to create auditory task in the first experiment as well as for the second and final experiment. Study [46] has found a major result where the auditory task caused a greater increase in gaze concentration during cognitive loading than the visual task. The auditory task is believed to make the driver to focus part of the attention on hard thinking.

A timeline on how this task experiment with and without audio was run has been pictured in Figure 3.7 below:
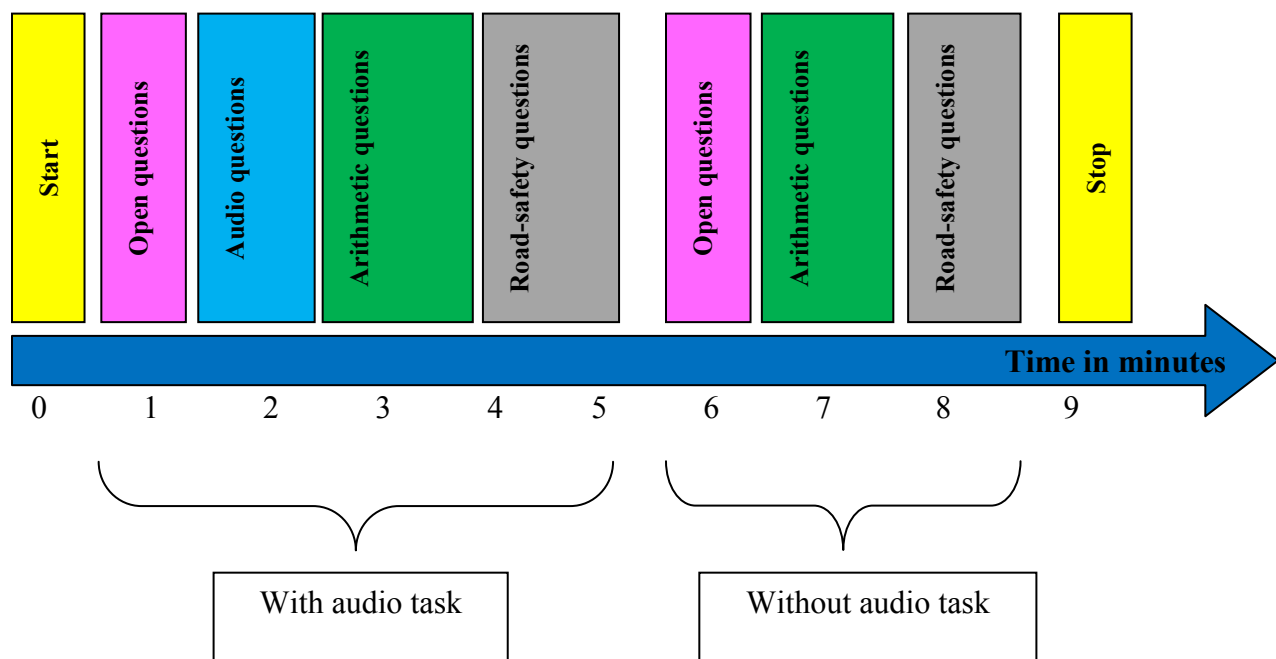
Figure 3.7: Questions time line

Result from audio task was found to cause bigger movements on the participant's facial expression than the one without any audio task. This is because the participants had been using both auditory and visual human sensors when they were with cognition. Auditory sensor was used for both to listen to the recorded radio and also to listen to the questions from the experimenter. However, in the task without audio, the auditory sensor only focused on waiting and listening to the questions given by the experimenter. Visual sensor was fully occupied in both tasks with and without audio.

As seen in the timeline above, the experiment began with an audio task and several open questions were asked first. Open questions are general questions like:

a) How are you doing?

b) What did you have for breakfast today?

c) Could you sleep last night?

Then the distractions continued with audio questions. Audio was played to the participants since the beginning of the experiment. Questions for the audio task are:

a) What is the title of the song?

b) What year do you think the song was created and why?

c) Do you know who the singer for that song is?

d) What is today's temperature as told by the deejay?

e) What do you think about the song, do you like it and why? (Elaborate briefly)

Next are the arithmetic questions. These arithmetic questions have been asked on three levels of difficulty: easy, medium and hard. The questions are as follows:

a) What is 2+10+10+2? (easy)

b) How much is $\sqrt{16} + 2 + (-6)$? (medium)

c) 1 mile is equal to 1.61 kilometres. How many kilometres is 13 miles? (hard)

Participants were also asked about the video they watched. Video was captured during a normal driving activity around Loughborough town. Questions from this category were based on the objects, people, landmarks, signboards, pedestrians, vehicles in front, oncoming vehicles and others which were basically road-safety related.

a) How many traffic lights have you seen in the video?

b) How many traffic lights are in green?

c) Can you read the plate number of the white van on the screen now?

d) Do you see a big lorry coming into your road?

e) What is the speed limit on this road? Do you see the signboard about the speed limit?

Second type of distracted/task experiment was the one without audio task. There was no recorded audio for the participants. Participants were only required to answer questions from the other three categories of questions. Open, road-safety related and arithmetic questions in this set of experiment were different from those in audio task experiment.

Open questions are:

a) What is the date and day for today?

b) What is the date for 8 days from today?

c) What are you going to eat tonight for your dinner?

The arithmetic questions in this set of experiment were also average questions just like the audio task. However, there is no different level of questions in this task without audio.

a) The time difference between Malaysia and UK is 7 hours behind during summer and 8 hours behind during winter. If the time in Malaysia is 2.14pm, what is the time in UK during summer?

b) What is 20 times 90?

Road-safety related questions are as follows:

a) At this roundabout, do you have to stop? Why you have to stop?

b) Are there any pedestrians crossing the road? Do you have to slow down your car and allow them to cross here? What colour of shirts are they wearing?

c) What is the registration number for the green Citroen Xsara car in front of you?

After all the questions had been asked, the participants were asked to continue watching the video until being told to stop. Once the video stopped and the participants were told to take a break, the faceLAB equipment was also stopped and the data was saved. After a while, the participants were again asked to watch the same video and this time no questions were asked meaning that there was no distraction at all. This was the control experiment and it was used for data comparison when participants were distracted and not distracted. Their facial movements were to be compared.

3.2.5 Summary and findings from First Experiment

The first experiment was conducted in a lab with six volunteers involved. The experiment was run with faceLAB machine as the main equipment. The objective for the first experiment was to find a relationship between eye and mouth movements. Mouth movement and eye movement were analysed on their height and width. A strong relationship between those two features has been found and discussed in Chapter 4. This positive result leads to the second experiment by finding the relationship between lips and eyebrows. The experiment was constructed with two different sets: distracted/task and not distracted/control. Task set of experiment was then divided into two types: with audio and without audio.

3.3 Second Experiment

After the first experiment successfully proved a strong correlation between eye movement and mouth movement [47] the second experiment was conducted to check on the lips and eyebrows. The experiment was executed firstly in December 2011. However, since December was a winter season and on some days the condition of the road and light were not really helpful, the experiment was extended until March 2012.

3.3.1 Experiment overview

The main objective of the experiment was to study new features: lips and eyebrows for driver cognitive distraction detection. The second objective was to study their correlation and regression. Result for this objective is discussed in chapter 5. The third objective is to use different types of classification algorithms like Support Vector Machine, Logistic Regression, Static Bayesian Network, AdaBoost and Dynamic Bayesian Network for driver cognitive distraction detection. Result for this objective is discussed in chapter 6. The second experiment used faceAPI machine to capture the features on the driver's facial expression. The experiment

was conducted on a real road environment. Ten experienced drivers had volunteered to join the experiment.

In the experiment, drivers were required to drive the car from Holywell Park to the Haslegrave Building and turn back to Holywell Park. This route is in Loughborough University, LE11 3TU, United Kingdom. The route's distance is approximately 3.1 miles and one lane per direction. The route consists of 9 humps, 7 zebra crossing and 10 shuttle bus stops. The participants only needed to turn left or right and use signals when they needed to drive out from the parking area (starting point) and make way back to the starting point.

The route map for the experiment is shown in Figure 3.8 below. The route is highlighted with a red line. Black arrows shows the direction from the starting point to the turning point where the driver needed to drive back to the end point. Starting point and ending point were at the same point, P1. P2 was the turning point.

Figure 3.8: Route map

During the experiments, oncoming traffic was present. However, the overall traffic density was moderate. Drivers had driven during daytime under different weather temperatures. The experiments had run during early spring and the temperatures were around 5 to 10 degree Celsius. Overall, 10 runs of attentive driving were captured and another 12 runs while the drivers were distracted. 2 runs from the distracted were discarded because the drivers had mistakenly taken wrong turns. The wrong runs affected the experiments because the situational awareness questions had to be changed. For instance, the questions about shuttle bus stops, humps, signboards and zebra crossings had to be changed due to the wrong turns taken by the drivers. Non-distracted experiment lasted around 18 to 24 minutes for each run and the distracted experiment lasted around 20 to 26 minutes, depending on the drivers' speed. Therefore, on average, for the distracted and the non-distracted set, the drivers drove around 23 minutes for each set. The average speed for every driver was around 12 to 20 miles per hour. The speed limit on campus is 15 miles per hour. The car used in this experiment was an automatic transmission car. All drivers had not faced any difficulty in driving the car.

Each driver was required to drive along the specified route twice. The first round was the control experiment where the driver was not distracted with any type of distraction. In this first round, the driver was asked to familiarize with the car and the route. This round has been labeled as control/baseline/non distracted experiment. The second round was the task experiment where the driver was distracted with both manual and visual distractions. As in [33], driving is a multitasking activity that requires the drivers to divide their attention to various driving and non-driving related tasks. Thus, to operate this multitasking activity, the drivers need to be experienced, well in emotion, motivated, vigilant and highly attentive while driving.

This second round was denoted as the distracted/task experiment. They were asked to drive down the same route while performing some tasks. Tasks were randomly selected for each driver. However, every driver participated with an equal number of tasks and questions.

3.3.2   Participants Details

10 experienced drivers were involved in this second experiment. However in second experiment all drivers are male. Their ages are between 21 years old to 36 years old. On average all drivers have been driving for 4 years. Unlike in the first experiment, in this second experiment one

driver was not really familiar with UK rules of driving, since the driver was from a country with left hand drive. Details about the participants are listed in the Table 3.2 below:

Gender	10 males
Age	21 to 36 years old $\mu = 29.5$ $\sigma = 5.08$
Occupations	5 are students 5 are employees
Nationality	3-UK driving license 7-International driving license

Table 3.2: Participants details-second experiment

Most of participants hold an international driving license. Nine participants are from a country with right-hand drive and one is from a left-hand drive country. Therefore, not every driver is familiar with UK's road driving. However, it turned out that every driver had safely driven the car and their data was also useable. From 5 students, 3 are undergraduate students and 2 are PhD students. 5 workers are include 1 medical doctor, 1 company manager, 1 university lecturer and 2 engineers. Out of the ten drivers, only six drivers have their own car and have been driving in UK. Four drivers drive their friend's car occasionally. Figure 3.9 below shows the fraction of driving frequency per week for every driver involved:
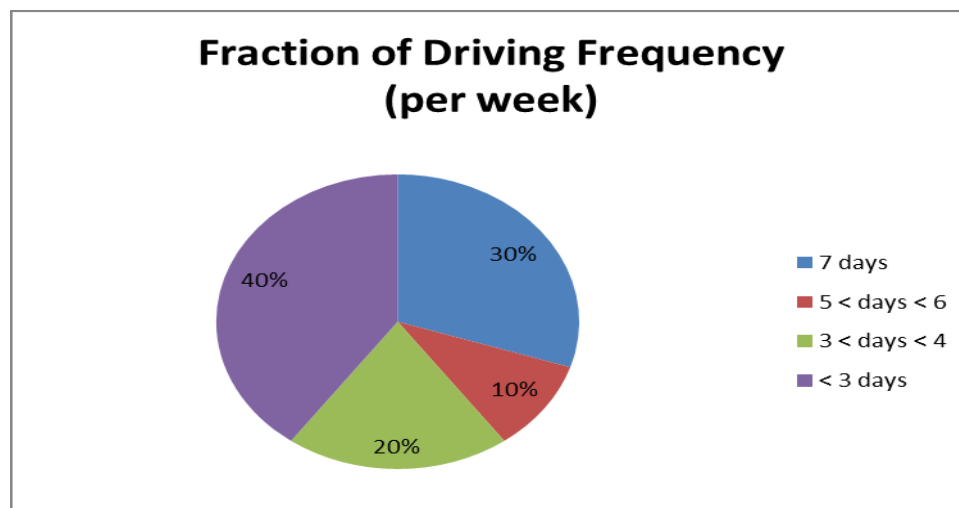
Figure 3.9: Driving Frequency (per week)

Since four drivers did not own a car in UK, 40% of the drivers do not frequently drive in UK. 30% of the drivers drive their cars every day.

3.3.3 Vehicle Setup

The second experiment used faceAPI machine to capture driver's expression. faceAPI is an application programming interface, not like faceLAB which is a standalone facial feature extraction system. However, faceAPI is easier to handle than faceLAB because it does not require any camera calibration. faceAPI can use any type of normal web camera to capture the facial features. Two different brands of web cameras were used in this experiment: ProLink and Logitech web cameras. ProLink web camera has 2 LED lights and therefore it was used whenever there was insufficient lighting during the experiment. Both web cameras were attached on the car's dashboard. A Nissan Micra 1.0, with an automatic transmission was used in this experiment.

Basically this second experiment required the equipment below:

a) faceAPI dongle

b) Laptop

c) Normal web camera with an appropriate length of wire cable

d) Duct tape

e) Software called ManyCam was used to capture static image and video during experiment

Since the drivers were allowed to adjust the seats accordingly to their preferences, the web camera also needed to be adjusted to ensure that it could capture the full face of a driver.

Figure 3.10: 1) faceAPI toolkit setup, 2) faceAPI initializing, 3) driver is driving and 4) after the experiment

As been mentioned, faceAPI is easier to setup compared to faceLAB because faceAPI does not require any camera calibration. Nevertheless, this advantage brought a problem in the tracking system especially when there was a problem with insufficient lighting. Whenever the experiment needs to be conducted with insufficient light place or time, the facial landmarks were hardly to be detected accurately. The landmarks lines were sometimes detected at the wrong position on

the participants face. Thus, to solve this problem, a high quality web camera with LED lights was used. The LEDs provided extra light required to capture landmarks on the face. Since a normal web camera was used for this faceAPI machine, the capability to capture the features precisely was sometimes challenging. A uniform and sufficient lighthing is important when using faceAPI because the facial landmarks will only can be detected accurately when the lighting is sufficient. faceAPI data are with two standards: coordinate frame and face landmark.
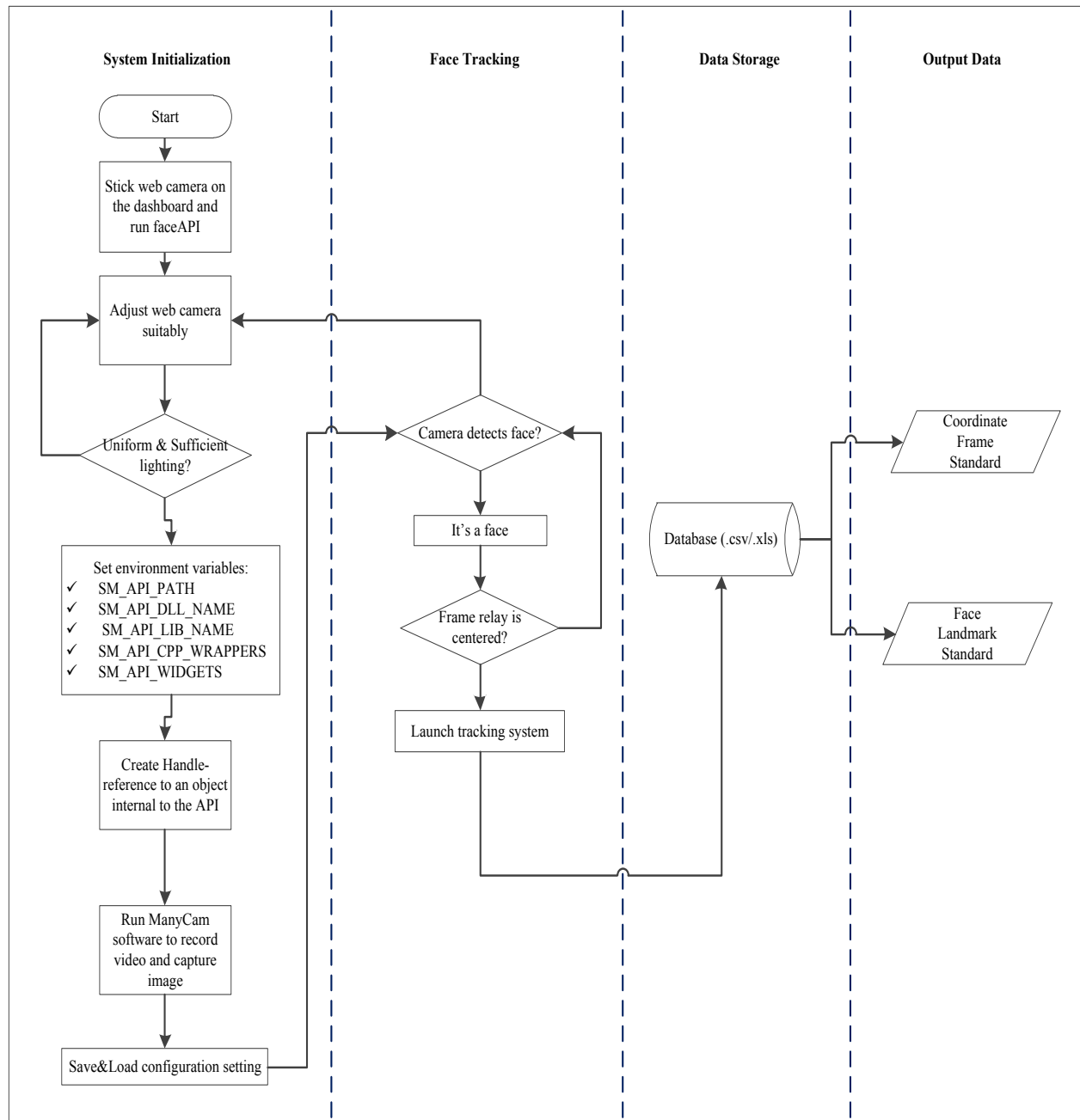
Figure 3.11: faceAPI flowchart

3.3.4   Distraction Setup

Like in the first experiment, the second experiment was also conducted with two types of set up: distracted/task and not distracted/control. In the distracted setup basically ten typical tasks were chosen as distraction conditions and these tasks were applied to all drivers:

a) Identify pedestrians (medium).

b) Recognize the road signboard, banner and poster (medium).

c) Identify road speed limit and respond accordingly (medium).

d) Converse with experimenter (low).

e) Use mobile phone (iPhone) to find words, contact numbers or use the phone's applications like Weather, Calculator or Messages (high).

f) Recognize model and registration number from the vehicle in-front (medium).

g) Read from a pamphlet. Find contact number and price on the pamphlet (high).

h) Answer some simple arithmetic questions (high).

i) Collect a bottle on the floor and pass it to the experimenter (high).

j) Adjust the radio settings (high).

Cognitive distraction was triggered from the above tasks. These ten types of distractions were generally from three different sets of cognitive loading as defined in [48]: Low cognitive loading, medium cognitive loading and high cognitive loading. Drivers drove in this experiment for about 20 to 26 minutes, therefore, on average every driver drove the car for 23 minutes each.
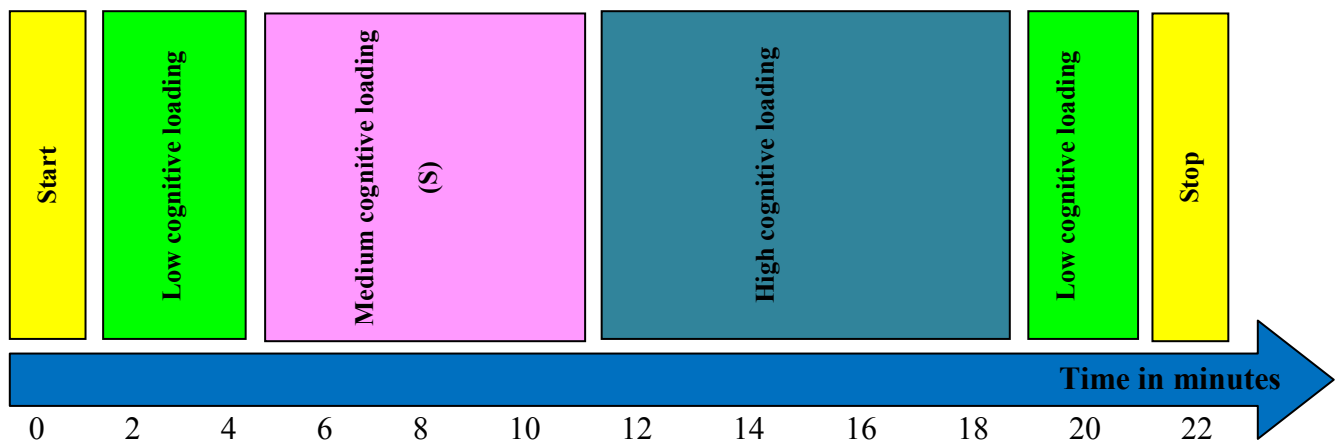
Figure 3.12: Cognitive loading time line

All participants were asked to do the same tasks. Planning and time arrangement are important to ensure the participants can complete the task within the time given. Before the experiment starts, the participants were told that they are required to answer and complete all tasks and there is no right or wrong answer to the questions or tasks. High cognitive loading took more time than the other types of cognitive loading because the drivers took longer time to think and reply to the high cognitive loading questions. Different events or tasks occurred at every minute during the experiment. The (S) symbol represents a situational awareness questions related to a road signboard, road landmark, pedestrian, trees, zebra crossing, bumper, other vehicles and other objects seen along the route. As per said, during the experiment, the experimenter asked the drivers with several questions to trigger the cognitive loading. Questions are categorized as None, Low, Medium and High on their difficulty level. At the beginning of the experiment, the drivers had not been asked anything and this happened at the start and stop time of the experiment.

The conversation started with Low Cognitive Loading question. This was only to warm up the drivers before Medium and High Cognitive Loading questions were asked. Low level questions would be imposed by a small talk, with questions such as:

   a) How are you doing? What are you doing today?

   b) What have you eaten for your breakfast or lunch?

   c) How was your day yesterday?

   d) What's your favourite class?

   e) Where do you live? (describe)

   f) What's your favorite food? How do you make one?

   g) What type of movie do you like? What was the last movie that you have watched?

The conversation with the low level questions would take in its natural course, where the experimenter did not require the participants to answer them with a complicated thinking process. The questions from this level would have been used as conversation starters if the conversation had come to a natural stop.

Medium question is related to the situation awareness. Situation awareness or SA is the perception of environmental elements with respect to space and time [48]. It involves the driver's comprehension of the meaning on safe driving. The drivers' perception on their surrounding situation or environment is crucial because it decides what decisions the drivers have to take at

the time. Situation awareness involves being aware of what is happening around the vicinity to understand how events, information and the drivers' own actions will impact the goals and objectives. It is found that, because lacking in SA, many drivers are involved with accidents respectively due to the driver's error [48]. To test this SA on every driver in this experiment, they were asked about the signboards, landmarks, bumpers, buildings, trees, registration number of the car in front, cyclists or pedestrians during the experiment. To fulfil this medium cognitive loading, the experiment included an interactive conversation with the drivers. The conversation were not scripted, but rather allowed to flow freely. The surrounding situation was possibly well known to the drivers, because they all had driven on the chosen route. The drivers also had to go through control experiment before they were to partake in the task experiment. In a control experiment, the drivers were required to drive the car following the route first before the task experiment could take place. Therefore, all of the drivers had not faced difficulty to respond to this medium cognitive loading task.

Some examples of the medium cognitive loading are as follows:

a) Can you read the registration number of the car in front of you? What type of car do you think it is?

b) Can you read what has been written on the signboard?

c) What is the speed limit here? Do you need to slow down or you need to speed up a little?

d) Can you count how many people at the bust stop?

e) Do you see any pedestrians? Do you think they want to cross the road? Do you want to stop?

f) Can you count how many cyclist have you passed and let me know the answer at the end of this experiment.

g) Can you find out what is the name of the building on your left? And what do you think that building is for?

High cognitive loading questions were involved with high thinking process. Some drivers took more than a minute to give an answer to the question given. The most and biggest facial expression on the drivers could be found during this period of cognitive loading. Drivers' lips and eyebrows had moved and changed at the maximum reach. High cognitive loading involved high level loading and manual distraction. One type of distraction was to find information about Loughborough Leisure Centre on a pamphlet. From the pamphlet the drivers were asked to read

about studio etiquette, workout class descriptions, find contact details and find information about the classes' time table and price. They also needed to calculate the price for more than two members to join the class. Thus, they were required to do multiplication for different prices. Driving and using a phone has been considered as the most contributory factor for an accident on the road. Other than getting information from the pamphlet, the drivers were also distracted with the task of using an iPhone. They were asked to open note application on the iPhone and find certain words in the application. There are many words listed in the note application, therefore, they were required to use the scroll up and down function on the iPhone to get the words. They were also required to find contact numbers using the iPhone and again they were required to use scroll up and down function to find the contact numbers. By using a cell phone while driving is considered as driver distraction. It is usually initialized by manual distraction and can lead to visual and cognitive distraction as mentioned earlier. Checking weather, using calculator and also texting are common examples of application used by the drivers on the cell phone [28].
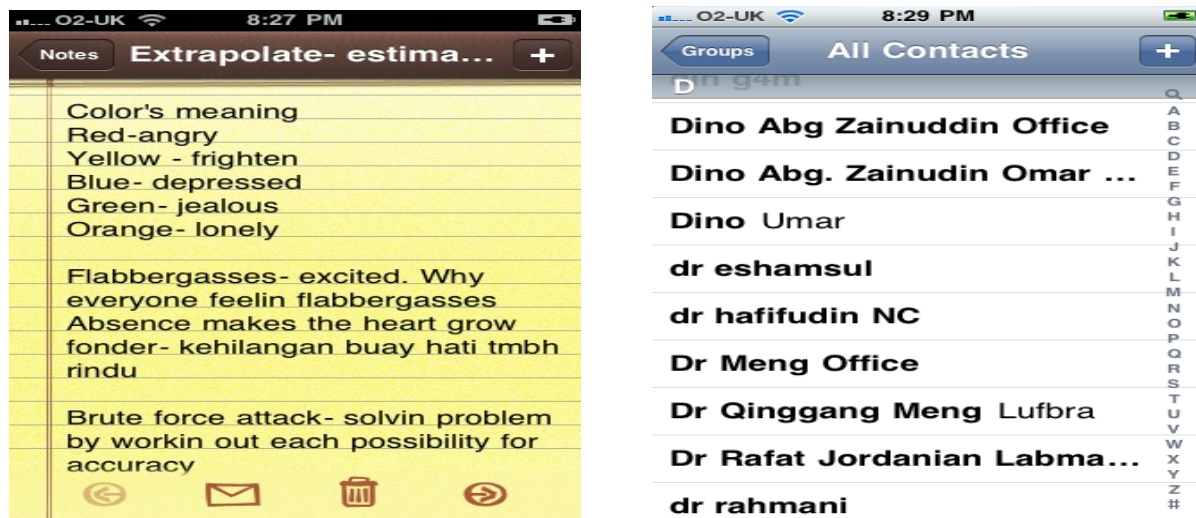
Figure 3.13: Notes and Contacts in iPhone

Like in the first experiment, the drivers for second experiment were also asked arithmetic questions. The arithmetic questions involved subtraction, addition, multiplication and division. As in other exercises, this arithmetic exercise also did not really consider too much on whether the participants gave a right answer or not.

Some of the example questions used for high cognitive loading are as follows:

a) Can you find what is the price for Body Attack class on Wednesday ?(pass the leisure centre pamphlet)

b) Can you tell me the phone number/address/website for the leisure centre?

c) Use the calculator on the phone calculate :

- 145/7 + 93
- 83-29 * 4.2
- 99+45-71

d) Use the phone find Dr. Qinggang Meng Lufbra number. Read it for me.

e) Could you please take out my ID card from my wallet under your steering and read me my ID card number.

f) Could you please turn on the wiper and wipe the windscreen and turn on the heater for me please and then turn them off again.

g) Do you know what an odometer is? Could you please read the odometer values for this car?

Some of the high cognitive loading was accomplished by manual distraction like adjusting the car's temperature, using the mobile phone, setting the navigator and others. Therefore, with the existing visual on the road and manual distraction created to the drivers, cognitive distraction can reach the maximum level. During high cognitive loading, all three types of driver distraction: visual, manual and cognitive are occupied. Thus, the movement on the driver's facial expression was higher than in manual and low cognitive loading. Towards the end of the experiment, low cognitive loading questions were given to hint the drivers that the experiment was ending.

Since the tasks imposing cognitive loading do not require a huge specific technical knowledge, the imposed cognitive loading should not differ too much between drivers. This is because all drivers are university graduates or students and they should have good working knowledge in both English and Mathematics. If there is any variation, it must be caused due to personal background like origin, primary education background and age. However, these variations are expected to be minimal. This cognitive loading task was purposely designed to examine the drivers' response time (RT) and situational awareness (SA). These two indications are useful to indicate the time takes for every driver to respond to the distraction and whether the drivers are alert or not to their surroundings. The table below shows a summary of RT and SA.

COGNITIVE LOADING	Task	Control
None	N/A	SA & RT
Low	RT	N/A
Medium	SA & RT	N/A
High	RT	N/A

SA- Situation Awareness; RT- Respond Time; N/A- Not available

Table 3.3: Cognitive Loading and its requirements

For each driver, as shown in Figures 3.14 and 3.15, a snapshot picture was taken for every round of driving. These pictures can be used to distinguish the differences between facial expression during distraction and non distraction. Basically, even without the data analysis, there were obvious movements on the drivers' facial features during distraction pictures in comparison with non distraction pictures.

The drivers' eyebrows and lips showed a significant different movement between distracted and undistracted expressions. During cognitively distracted experiment, the drivers' eyebrows were pulled down at the middle and made a slope inwards. Lips were also pursed in and pressed in together. During normal expression (without distraction), drivers' eyebrows and lips were stretched out and there were no stressing muscles around those features. Details about lips and eyebrows movement can be found in chapter 2.

Figure 3.14: Normal face & faceAPI control screen

Figure 3.15: Distracted face & faceAPI control screen

3.3.5 Summary and findings from Second Experiment

As been mentioned before, the second experiment was conducted for three main objectives. The first objective was to study new features: lips and eyebrows for driver cognitive distraction detection. The second objective was to analyze their correlation and regression. Result for this objective is discussed in chapter 5. The third objective was to use different types of classification algorithms like Support Vector Machine, Logistic Regression, Static Bayesian Network, AdaBoost and Dynamic Bayesian Network for driver cognitive distraction detection. Result for this objective is discussed in chapter 6.

The second experiment was conducted with faceAPI machine. Normal web cameras were used and attached on the dashboard of the car. Unlike faceLAB, faceAPI is easier to handle and install. There is no camera calibration required.

10 drivers were involved in this experiment. The experiment took place in Loughborough University and had run during early spring term. Just like first experiment, the second experiment was also constructed with two sets of experiment: task and control. However, the task experiment was different from the first one. Questions in task experiment were categorized into three classes of cognitive loading: low, medium and high.

3.4 Third/Final Experiment

In the final experiment, both faceLAB and faceAPI were installed in the car. The experiment was conducted on real road traffic in Loughborough University campus. The experiment took place from 27th April 2012 until 30th April 2012 and resumed again on 6th May 2012 until 7th May 2012.

3.4.1 Experiment Overview

This final experiment involved two different purposes. The main purpose for this experiment was to conduct a driver cognitive distraction experiment and this experiment had run during day time from 9am until 7pm. The second experiment was for driver fatigue or sleepiness detection. This experiment had run from 1am until 7am.

Weather condition during the experiment was rainy, cloudy with strong winds. On average, the temperature during the experiment was around 7 degree Celsius to 11 degree Celsius. The same route like in the second experiment was used in this experiment. Every driver was required to drive the route twice. The first round was for control/non-distracted or baseline data collection and the second round was with distraction or task.

The speed limit for every driver was 15 miles/hour. On average, each driver drove the car at a speed between 12 miles/hour to 22 miles/hour. The traffic along the route was just like normal traffic. Shuttle bus was operating normally during the experiment and the traffic was heavier especially during day time than at night. Students were also commuting and using the road with bicycle or walking. Therefore, a high vigilance was required during the experiment.

This final experiment was conducted with a Ford Mondeo car. The car is with a manual transmission. The car belongs to Ergonomics and Safety Research Institute (ESRI) from Loughborough University. Both faceLAB and faceAPI cameras were installed in the car for information fusion. Since this final experiment has two major objectives, for cognitive distraction detection and fatigue detection, data from both cameras were extracted separately.

Notes and observation were also taken during the whole experiment. The driver's behaviour during the experiment was noted. Pictures were also taken during and after the completion of the experiment for every driver.

3.4.2   Participants Details

For cognitive distraction objective, 14 experienced drivers had participated in the experiment voluntarily. 8 drivers had participated in the fatigue experiment. In cognitive distraction experiment, each driver was asked to read and fill up the consent form. Important details like the driver's age, occupation, driving license and years of driving experience were asked in the form. Driver's names are anonymous. However, summary from their details can be found in Table 3.4 below:

Driver	Age	Occupation	Driving License	Driving experience
1	31	Research student	International	≥ 10
2	33	Research student	United Kingdom	≥ 10
3	22	Undergraduate student	United Kingdom	≥ 5
4	26	Research student	International	≥ 5
5	22	Undergraduate student	International	≥ 5
6	34	Research student	United Kingdom	≥ 10
7	29	Research student	United Kingdom	≤ 5
8	22	Postgraduate student	United Kingdom	≤ 5
9	30	Research student	International	≥ 10
10	23	Undergraduate student	United Kingdom	≤ 5
11	45	Research student	United Kingdom	≥ 20
12	25	Research student	United Kingdom	≥ 5
13	24	Unemployed	United Kingdom	≥ 5
14	34	Engineer	United Kingdom	≥ 10

Table 3.4: Driver's summary

All drivers are familiar with manual car transmission. Therefore, no one had difficulty to operate and handle the car during the experiment. All drivers are male and they have more than five years of driving experience. Basically, 11 drivers have driving experience of more than 5 years. Mean for the driver's age is, $\mu = 28.71$ and the standard deviation is, $\sigma = 6.20$.
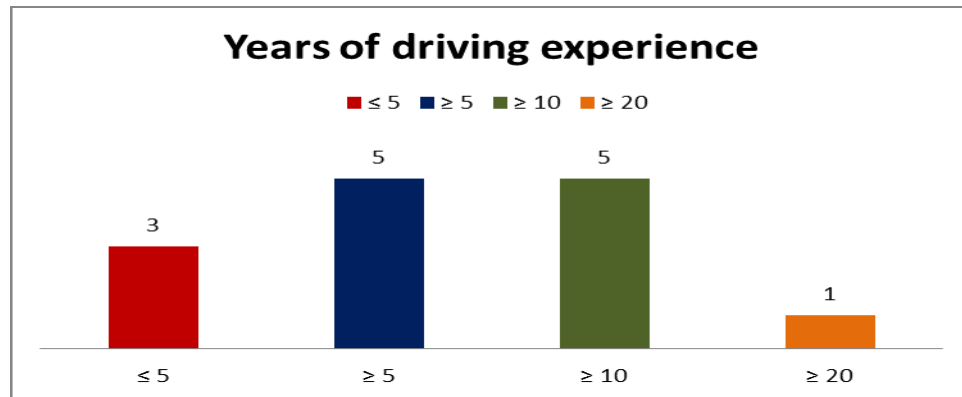
Figure 3.16: Driver's years of driving experience

Based on the observation made, most drivers have more than 5 years of driving experience. Two drivers have been found to have difficulty of adapting the rules on the road, since they are drivers from a country with left-hand drive. Initially, both drivers had difficulty in keeping to their lane, however, they adapted quickly to the driving condition.

When asked about the driver's frequency of driving per week, 5 of them had been driving almost every day. Four of the drivers rarely drive per week (<3 days).
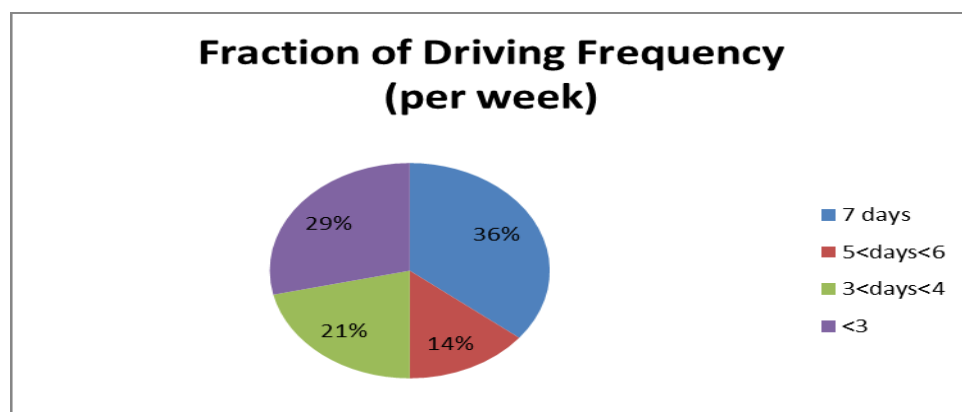
Figure 3.17: Driving frequency per week

Pictures from every driver were taken during and after the experiment run. Below are the pictures taken after the experiment run.

Figure 3.18: Driver's pictures

3.4.3 Vehicle Setup

Generally, this final experiment was conducted by combining the first and the second experiment setup. There are a few numbers of equipment available and used during the experiment:

a) faceLAB equipment

b) faceAPI equipment

c) Two laptops for each equipment

d) Four mounted video cameras in the car

e) GPS tracking system

f) Plug, cable and card adaptor for power supply to laptops and cameras

g) Calibration board

h) Digital camera

Just like in the first experiment, since this final experiment had also used faceLAB cameras, a calibration was required. Nevertheless, the calibration was not necessary to be done for every driver. Throughout the experiment, calibration was made only five times. Surprisingly, the infrared cameras tracking the driver's facial features worked better in rainy days during the experiment run. Pupil diameter was the hardest features to be tracked especially when there was very bright sunshine. Extreme light affected the detection of pupil diameter. faceLAB was trickier to set up than the faceAPI because the infrared cameras used were very sensitive to light. faceAPI on the other hand did not require any calibration. Normal web camera with LED lights was used in this final experiment just like in the second experiment. The LED lights were useful whenever there was insufficient light to detect driver's facial landmarks especially during the night time experiment.

During the whole experiment, three experimenters had been involved in operating and handling the experiment. The first experimenter is the leader who provided the distractions, set up the vehicle and monitored the whole experiment. The other two experimenters are the assistant and they were graduate students who were doing their final year project and Masters Dissertation in driver fatigue detection. From the three experimenters, unfortunately only two of them were required to be in the car during the experiment run. This was due to the limited space in the car. One person was required to handle faceLAB laptop and another person was to handle the faceAPI laptop. Leader and assistants had alternately sit in the car and conducted the experiment.

Experimenter 1   Volunteer Driver

Experimenter 2

Figure 3.19: People involved in the experiment

The faceLAB adapter cable was very short, therefore faceLAB laptop had operated from the passenger seat next to the driver and the laptop for faceAPI was set up at the passenger seat in the back.

faceLAB laptop
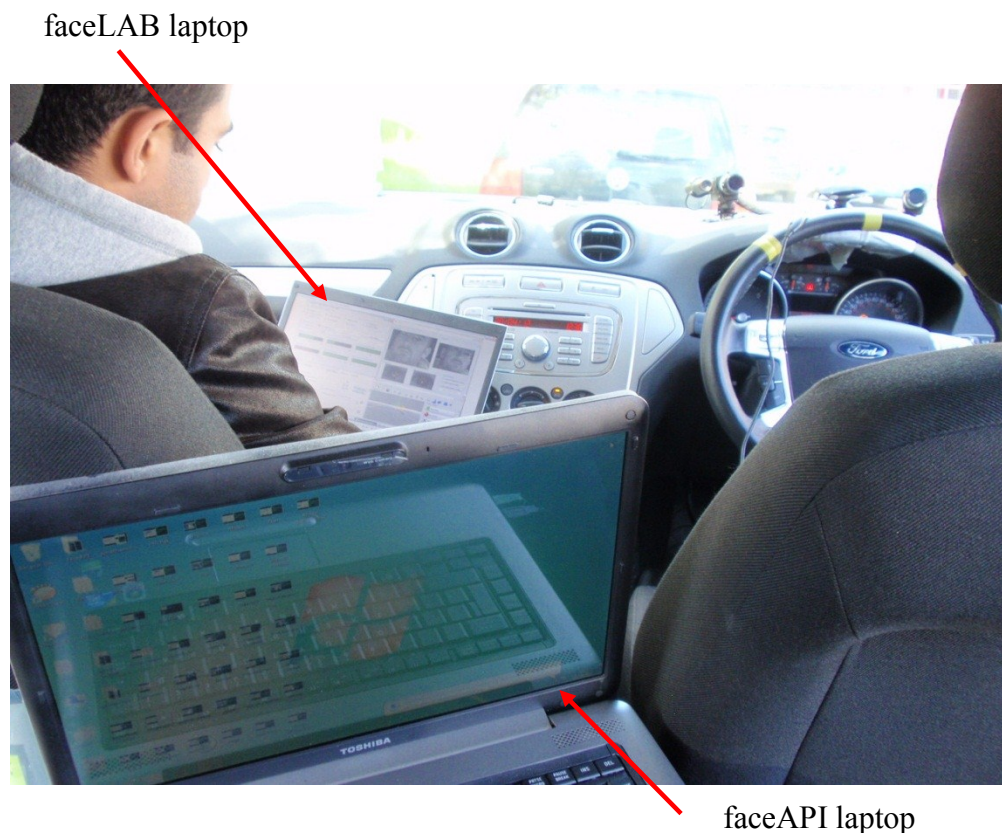
faceAPI laptop

Figure 3.20: Laptops for faceLAB and faceAPI

Both laptops are from Toshiba and they both were in good conditions. They operate the Windows 7 Operating System.

Just like being mentioned before, faceLAB cameras and faceAPI camera are different types of cameras. faceLAB uses infrared cameras and faceAPI runs with just a normal web camera. Both cameras were installed on the dashboard of the car. There was no unwanted and excessive light from both cameras which could distract and undermine the driver's vision. Thus, all drivers were not distracted at all with the existence of the cameras on the dashboard. The only problem with the faceAPI camera was that sometimes it moved around on the dashboard and sometimes fell down during the experiment because it was unstably stickered onto the dashboard. However, this

only happened a few times and to solve the problem, the experimenters had always ensured the camera was stickered firmly on the dashboard.
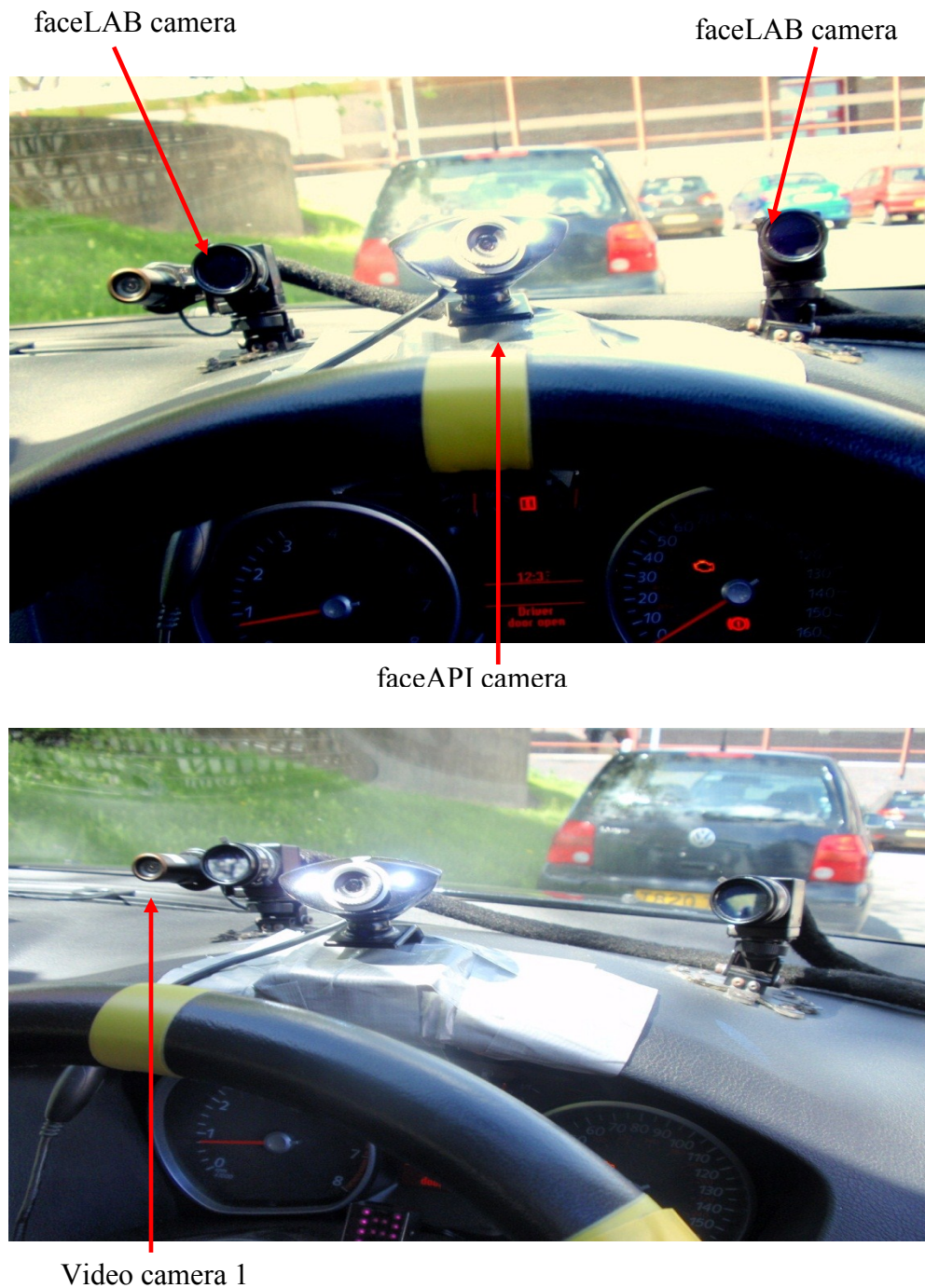
Figure 3.21: Cameras in the car

There were four video cameras mounted in the car and they were used to take video from different angles: (a) on the dashboard to capture driver's front view, (b) on top from the passenger seat next to the driver, (c) at the front mirror and (d) at the rear mirror. Figure below shows a screenshot taken from the video captured during the experiment.

Figure 3.22: Pictures taken from 4 video cameras in the car

Video camera A was used to capture the drivers from the front and video camera B was used to capture the drivers from the passenger seat. Video cameras C and D were used to capture video from outside the car from the front and rear mirrors. Video captures were only for the record purpose for this study. No specific data was collected from the videos captured. Two compact

flash memory cards were used to store the videos captured in this experiment and the box to insert the cards was installed in the car's boot. Start and stop button on the box was required to be pressed before and after the experiment.

As been mentioned before, a calibration was required before the experiment can be run because it would help the process of accurately determining the camera's extrinsic and intrinsic parameters. For stereo calibration, these parameters depend on the relative positioning of the cameras. Camera calibration in faceLAB allows the stereo-head to operate as a 3D measurement device. There are basically three steps involved in the camera calibration process:

a) Capture data for calibration - use camera calibration wizard

b) Process the calibration - use calibration key which is either provided with the system or by sending the calibration to the Seeing Machines Company.

c) Install the cameras' parameters into faceLAB.

The stereo-head takes measurements in 3D by seeing points on the subject's face from both of the camera views. If certain parameters used to model the cameras are known, the 3D coordinates of a point can be computed from the matched pair of image points. Basically the purpose of camera calibration is to determine the cameras' parameters and the parameters describe (a) imaging properties of each camera, most noticeably their focal lengths and (b) position and orientation (pose) of each camera relative to each other. Since the faceLAB cameras are very sensitive, every time the cameras are repositioned, or even bumped or knocked slightly, they should be recalibrated. The same applies if the camera zooms have changed. However, since both cameras and the infrared were screwed tightly on the dashboard of the car the calibration was re-done only because the driver's heights are different, therefore the camera's positions are required for a slight adjustment.

Figure 3.23: Camera calibration during the experiment

Two boxes above in Figure 3.23 showed the calibration process. The other two boxes are the views captured from front and rear cameras.

3.4.4   Distraction Setup

Distraction setup in the final experiment was similar to the second experiment. Three different levels of questions were asked to the driver to create cognitive distraction. Cognitive distraction mostly occurred when the driver was visually and manually distracted during the driving activity. Therefore, some questions or instructions during the experiment required the drivers to observe their surroundings and sometimes to take off one of their hands from the steering wheel in order to accomplish certain tasks to initiate the cognitive distraction.
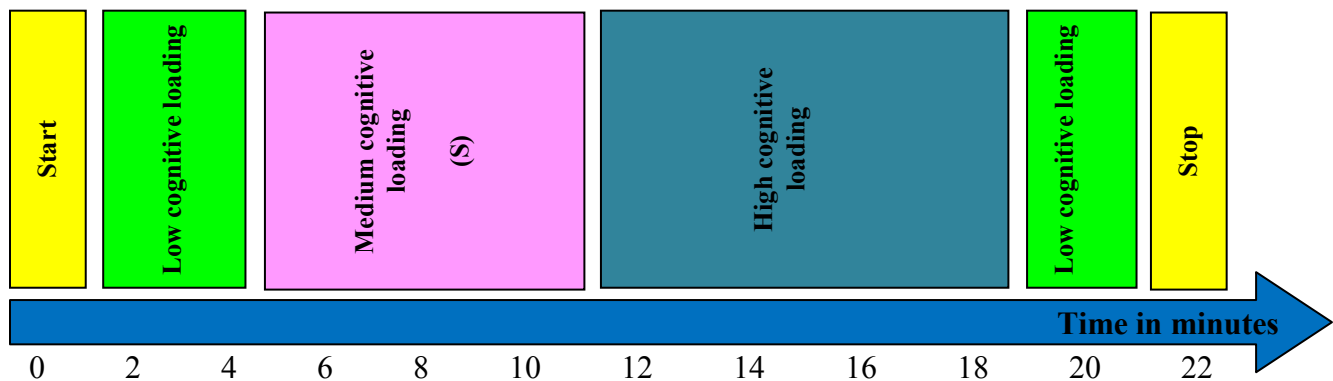
Figure 3.24: Cognitive loading timeline

Just like the other two previous experiments, this final experiment was also categorized into two: (a) control/not-distracted and (b) task/distracted. Roughly, on average, the total time taken for every driver to complete the control route was around 22.24 minutes and 26.94 minutes to complete the task route. Task round was a little bit longer than the control because the drivers required more time to respond to the questions and tasks given.

Before the experiment was conducted, all drivers had been explained about the experiment. They were asked to read the brief information about the experiment on the consent form before they could proceed with the experiment. For task experiment, they had been informed that there was no correct or wrong answer. Drivers were only required to respond to the questions and tasks appropriately.

Questions and tasks created in this final experiment were amended slightly to improve the impact of creating the cognitive distraction. For example, in high cognitive loading, drivers were given up to five questions related to arithmetic. From the observations, it showed that 9 out of 14 drivers were highly cognitive distracted when the arithmetic questions were asked. Two of them had even stopped the car for a few seconds to take time into giving an appropriate response. Sample of the questions and tasks for this experiment are given in the figures below:

LOW COGNITIVE LOADING

[1] How old are you? What is your date of birth?

[2] Have you had your breakfast? If you did, what had you have?

[3] What are you studying here? Explain briefly.

[4] What is your favourite food? Explain briefly how to make one.

[5] Where do you live in UK? Explain your neighbourhood briefly. Do you know your neighbours?

[6] What is your mobile number? How much have you bought your mobile?

[7] What is your favourite movie? Tell me about the movie. Who are the actors in the movie?

Figure 3.25: Low Cognitive Loading

MEDIUM COGNITIVE LOADING

[1] Can you read the registration number for the car in front of you? What type of car do you think it is?

[2] Can you read what's on the signboard?

[3] What is the speed limit here? Do you need to slow down or you need to speed up a little?

[4] Can you count how many people at the bus stop?

[5] Do you see any pedestrians? Do you think they want to cross the road? Do you have to stop?

[6] What does the road signboard means on your left? Do you have to take any action?

[7] Can you find out what is the name of the building on your left? And what do you think is that building for?

Figure 3.26: Medium Cognitive Loading

<div style="border: 1px solid red; border-radius: 20px;">

<u>HIGH COGNITIVE LOADING</u>

[1]Could you please read this paper for me please? (reading a pamphlet)

[2]Can you find the information from the pamphlet: price/email address/contact?

[3]Use the calculator on the phone to calculate :
- 145/7 + 93
- 83-29 * 4.2
- 99+45-71

[4]Use the phone to find Dr. Qinggang Meng Lufbra number. Read it for me.

[5]Could you please turn on the radio/heater please? Tune the volume/temperature to certain amount. And turn them off.

[6]Do you know what is odometer? Could you please read the odometer values for this car?

[7]Arithmetic questions:
- 456+65
- 20% from 450
- 864 / 4
- 555 * 5

</div>

Figure 3.27: High Cognitive Loading

Questions and tasks among the drivers varied. Sometimes, extra questions were asked based on relevant situations and reasons. On top of these three levels of cognitive loading, all drivers were required to do a memory test. At the beginning of the task experiment, they were required to count and remember either one of these:

a) How many shuttle buses have you seen/passed/followed along the drive?

b) How many bus stops have you driven/passed through?

c) How many speed bumps/zebra crossings have you driven/crossed through?

This memory test was an extra task to every driver and they were required to give the answer at the end of the experiment. Memory task created more cognitive distraction for the drivers. As

mentioned in the second experiment, the route consisted of 9 speed bumps, 7 zebra crossings and 10 shuttle bus stops on both sides of the road.

Based on the observations, several dangerous behaviours were identified during the task experiment, especially during the situational awareness loading and high cognitive loading. Some of the drivers had:

Dangerous behavior	Out of 14 drivers
Missed to look and did not see a car coming from the right when stopping at the junction when they wanted to make a turn to the left.	2
Ignored and did not see the pedestrians who intended to cross the zebra crossing.	5
Exceeded the speed limit. Some drivers drove up to 22 miles per hour.	9
Wanted to overtake a stopping shuttle bus but did not realize upcoming vehicle from the front.	2
Accidently stop and wait closely behind the stopping shuttle bus	3
Failed to slow down on the speed bumps	4
Missed or forgot the route and made an emergency turn without giving signal	1
Nearly hit the road stands on the road side	1

Table 3.5: Driver's dangerous behavior based on observations

Similarly to the second experiment, during this final experiment, all drivers' photos were taken during control experiment and task experiment. Based on the photos, the difference between control facial expression and distracted facial expression are obvious.

One sample each was taken from faceLAB and faceAPI cameras for each control and task experiment. From both cameras, it showed an obvious difference between control expression and distracted expression. Lips and eyebrows were the most obvious features that can be seen on the driver's facial expression. Lips will be puckered in when a person is thinking. The width and the height of the lips get smaller when a person is thinking. Eyebrows are pressed into the middle when a person is with cognition. Head rotation was physically changed when the drivers were cognitively distracted. Naturally, based on the observation from the experiment, one of the eyebrows was raised up based on which direction the head was tilted. If the head is tilted down

to the left, the right eyebrow will rise up and if the head is tilted down to the right, the left eyebrow will rise up. However if the head is straight forward, both eyebrows are usually pressed into the middle and both will have similar height. Nevertheless, this observation is based on the screenshots taken and it is correctly classified for 9 out of the 14 drivers. Thinking is a universal expression therefore, the facial expression involved with the features are naturally done and might change from time to time and different from a person to another person.
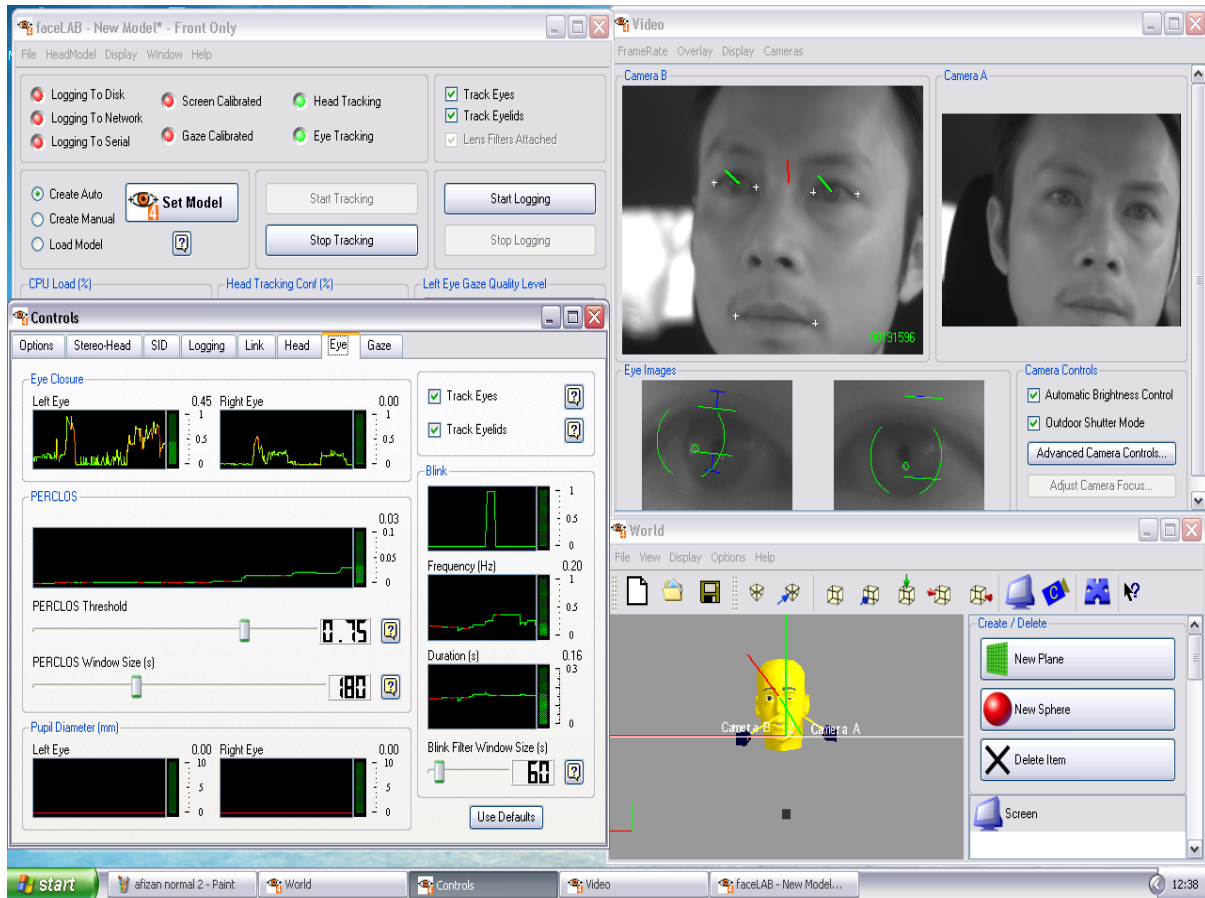
Figure 3.28: Screenshot for control expression on faceLAB cameras
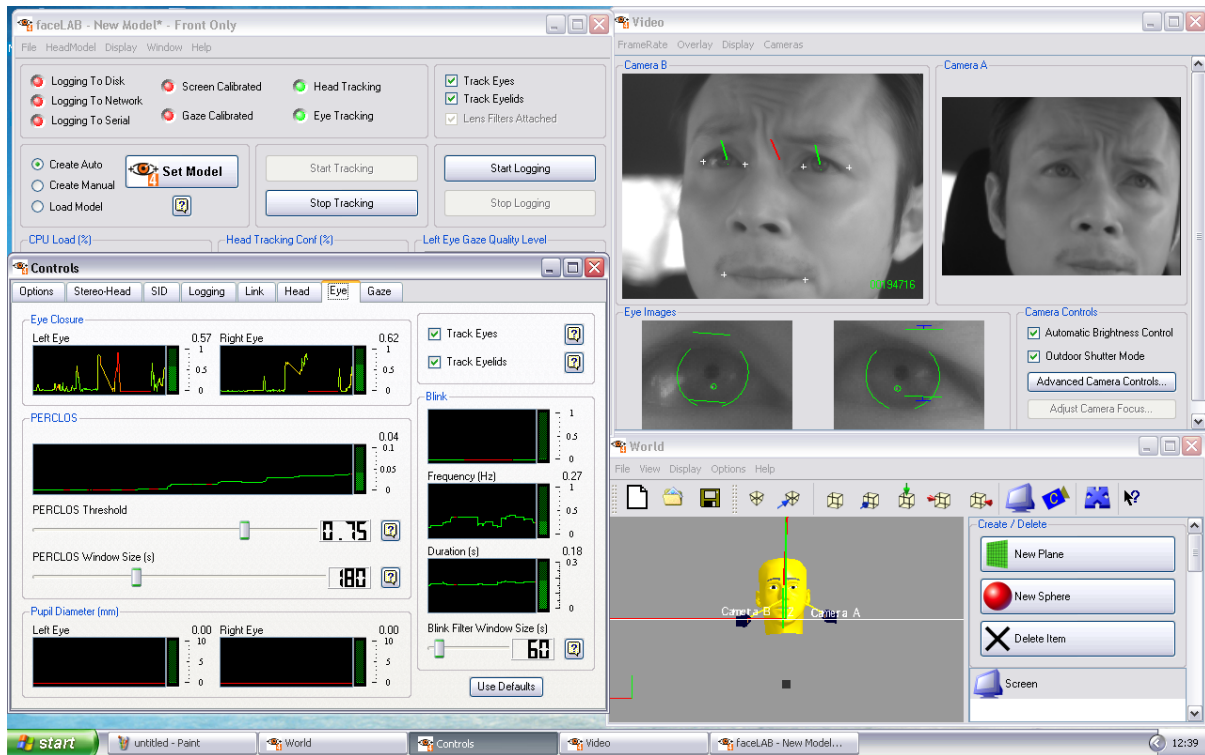
Figure 3.29: Screenshot for distracted expression on faceLAB cameras
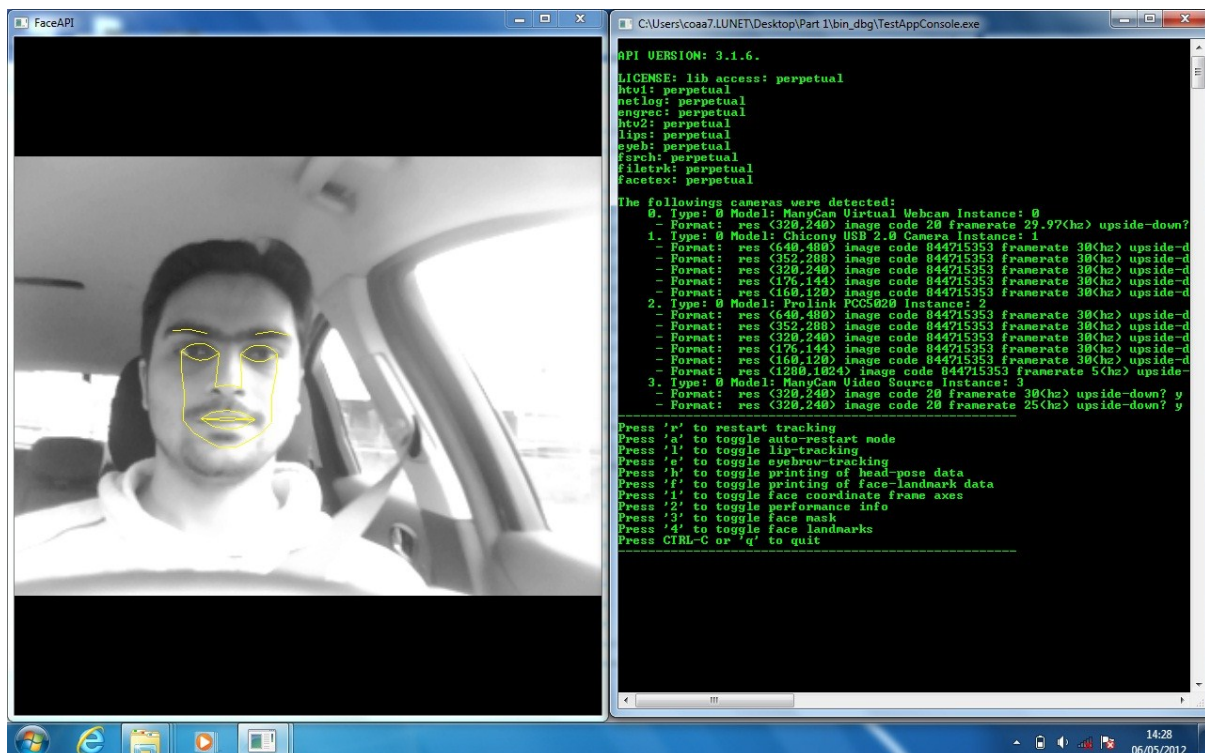
Figure 3.30: Screenshot for control expression on faceAPI camera
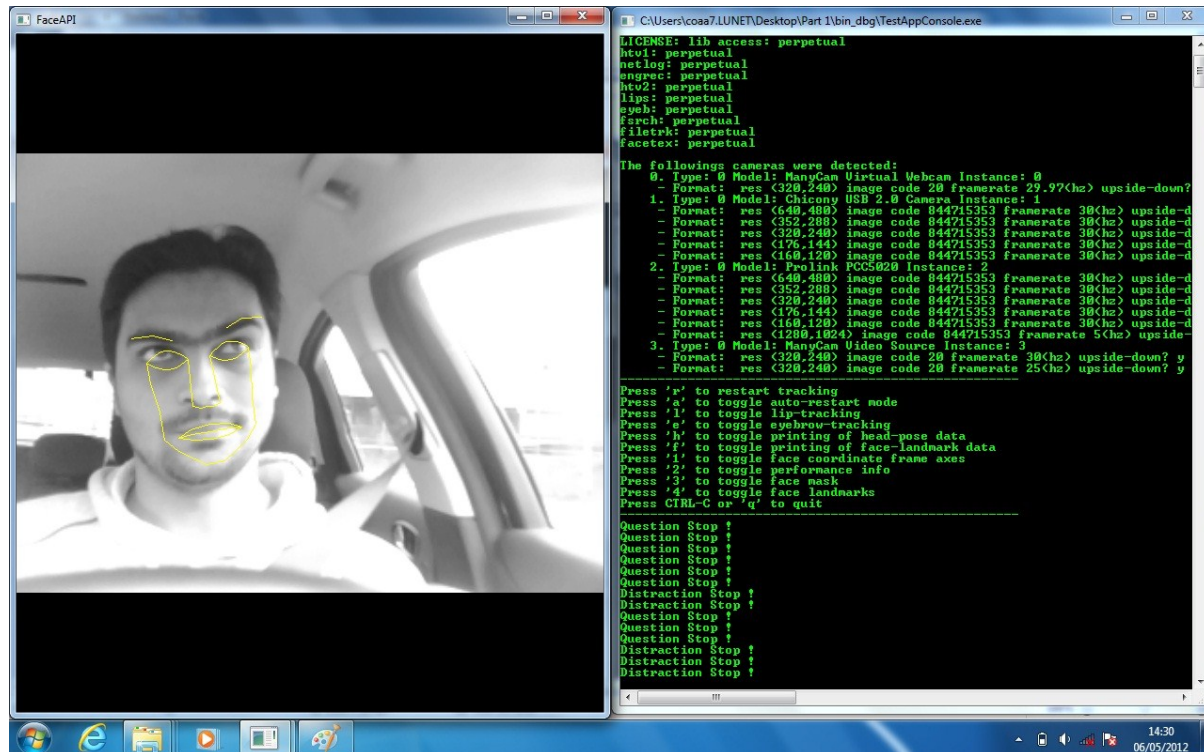
Figure 3.31: Screenshot for distracted expression on faceAPI camera

3.4.5   Summary and findings from final experiment

Final experiment was run in real time traffic like in the second experiment. There were two objectives for this final experiment:

a) To detect driver cognitive distraction

b) To detect driver fatigue

14 volunteer drivers participated in the cognitive distraction experiment and 8 drivers participated in the fatigue experiment. All drivers are experienced drivers with either UK driving license or international driving license. Most of the drivers have an experience of driving more than 5 years.

This final experiment was run in Loughborough University campus with two sets of experiment: control and distracted. On average, each set of experiment took around 22 to 23 minutes for each driver. Three different levels of questions and tasks were asked to the drivers during the

distracted experiment. Most of the drivers were highly distracted cognitively with high cognitive loading questions especially the arithmetic questions.

Major contribution from this final experiment was the fusion of the faceLAB camera with the faceAPI camera and the combination provided information fusion from both cameras. Both systems were installed in a Ford Mondeo car with a manual transmission. Videos were also captured during the experiment for the record purpose of this study. Four video cameras were installed in the car from different angles and directions.

Besides the captured data, observations were also made during the experiment, which mainly focused on the driver's behaviour while running the experiment.

Besides the captured data from the equipments, observations were also made during the experiment. Observations mainly focused on the driver's behaviour while running the experiment.

It has been found that lips and eyebrows are obviously seen on the driver's facial expression when they were cognitively distracted. Lips were puckered in and the eyebrows were pressed into the middle when they were distracted cognitively. It has been found that head tilting and eyebrow's raise are related. However, this is only on the observation made during the experiment.

3.5   Comparison of Experimental Setup

Comparison of the experimental setup was done between different studies in detecting driver's cognitive distraction. Up to date, there have not been many researchers doing a study in driver cognitive distraction. There are only few studies about driver cognitive distraction. While driver fatigue or sleepiness detection is perfected by many automobile companies like Toyota and Mercedes, there are no devices or technology for driver safety to detect cognitive distraction detection. Therefore, more studies are needed to improve the detection. Most previous studies on driver cognitive distraction were conducted with a lab simulator. Thus, the data collected was not from a real driving environment.

AUTHOR/S	PARAMETERS	PERIPHERALS	EXPERIMENT
Yulan Liang, Michelle L. Reyes, John D. Lee	a) Blinking frequency b) Fixation c) Saccade d) Smooth pursuit. e) Steering wheel angle f) Lane position g) Steering error	faceLAB cameras and vehicle sensors	Driving simulator
Masahiro Miyaji, Haruki Kawanaka, Koji Oguri	a) Pupil diameter b) Gaze rotation c) Head movement d) Heart rate	faceLAB and Electrocardiogram	Driving simulator
M H Kutila, M Jokela, J Viitanen, G Markkula, T W Victor	a) Gaze rotation b) Head angle c) Lane keeping	faceLAB and vehicle sensors (lane tracker, speedometer, brake sensors and etc.)	Real Traffic
Afizan Azman, Qinggang Meng, Eran Edirisinghe	a) Mouth movement b) Eye movement c) Head rotation d) Gaze rotation e) Pupil diameter	faceLAB	Lab setting
Afizan Azman, Qinggang Meng, Eran Edirisinghe	a) Head rotation b) Head position c) Eyebrow d) Lips	faceAPI	Real Traffic
Afizan Azman, Qinggang Meng, Eran Edirisinghe	a) Head rotation b) Eyebrow c) Lips d) Blinking Frequency e) Gaze rotation	faceAPI and faceLAB	Real Traffic

Table 3.6: Experimental setup comparison

3.6 Conclusions

Three different experiments were conducted in this study. The first experiment was set up in a simple lab setting where 6 volunteers had participated. It was purposely constructed to examine

the correlation and relationship between mouth movement and eye movement: height and width. Besides that, it was also a useful training to use faceLAB cameras. Different tasks and questions were designed in the first experiment. They were used to create a cognitive distraction to the participants.

The second experiment was conducted on a real traffic/road environment. 10 volunteer drivers were involved in the second experiment. Since result for correlation between mouth movement and eye movement in the first experiment was found to be strong, in the second experiment, faceAPI was used to capture lips and eyebrows. Other than head rotation, lips and eyebrows were chosen to be included as new parameters for cognitive distraction because these two features are the most obvious features to change when a person is with cognition. The second experiment was conducted in Loughborough University campus, LE11 3TU, Leicestershire, UK. Each driver was requested to drive a car to follow the route twice: the first round was for a control experiment and the second round for the distracted experiment. During the distracted experiment, the drivers were distracted with three levels of questions and tasks.

The final experiment was conducted similarly like in second experiment. However, this final experiment was run with the fusion of the cameras from the first and the second experiment. Both faceLAB and faceAPI were combined in the final experiment. This final experiment had two main objectives: to detect driver fatigue and to detect driver cognitive distraction. A different car was used in this final experiment. 14 different drivers were involved in the cognitive distraction experiment and 8 participated in driver fatigue experiment. Setup for final experiment was similar to the second experiment. All drivers were required to respond to three levels of questions or tasks.

All questions were set based on the guidelines given in [48]. The study was conducted for driver's cognitive distraction. 3 sets of cognitive workload were given to the participants. The questions may be varied but the levels of the question's difficulty were followed exactly like what have been done in that study. As been mentioned earlier, the answer given by the participants are not necessary to be exactly correct. They were only required to answer the question as best as possible. A sufficient of time is given to them to answer the questions. From all three experiments conducted for this study, all participants were participated finely. They had answered the questions as per requested and accomplished the tasks appropriately and accordingly.

Chapter 4

Physiological Measurement and Bayesian Network for Driver Cognitive Distraction Detection

"The purpose I mean is, to shew what reason we have for believing that there are in the constitution of things fix laws according to which things happen, and that, therefore, the frame of the world must be the effect of the wisdom and power of an intelligent cause; and thus to confirm the argument taken from final causes for the existence of the Deity. It will be easy to see that the converse problem solved in this essay is more directly applicable to this purpose; for it shews us, with distinctness and precision, in every case of any particular order or recurrency of events, what reason there is to think that such recurrency or order is derived from stable causes or regulations in nature, and not from any irregularities of chance [49]."

(Richard Price, 1763)

This chapter will discuss about methods to detect driver cognitive distraction generally and physiological measurement specifically, since it was used in this study. Bayesian network (BN) definition, concepts and implementation are also explained in this chapter. The discussion will begin with the introduction to the Bayesian network itself and upwards to the implementation of Bayesian Network in this study. Two types of BNs were used and described in this chapter: static and dynamic. Reverend Thomas Bayes was the father of Bayes's Theorem. Upon his death in 1761, Richard Price had edited and published Bayes's most famous accomplishment since Bayes

had never published his work. The Bayes' theorem had become the solution to the 'inverse probability' problem and has been used widely in many other problems and it is used in this driver cognitive distraction detection study.

4.1 Background

There are three basic methods to detect driver's cognitive distraction: (a) performance measure (primary tasks and secondary tasks), (b) physiological measurement and (c) rating scales. In this study, physiological measurement has been chosen and solely used from the initial experiment until the end. Physiological measurement was chosen because it is more reliable and cannot undermine drivers while the driving activity is running. Many researchers mainly used physiological measurement in their research and some of the research has combined physiological measurement with other types of detection methods [51]. Main objective to choose physiological measurements solely is because this method won't undermine the driver's performance since the cameras used in the experiment had not connected to the driver's body and won't disturb the driver's driving task. This chapter is used to explain the motivation and purpose to detect driver's distraction. A good driver's distraction detection is necessary because a human being is very dynamic. A person is usually won't produce a real behaviour if they knew that will be observed or an instrument is connected to their body. Thus, proper and non-intrusive distraction detection is required. Therefore, physiological measurement is the best method to detect this distraction. A few numbers of features were selected and proposed to be used as physiological measurements in this study, such as lips, eyebrows, gaze rotation, blinking frequency and head rotation.

A few numbers of classification algorithms are available to use once the data has been collected. Among of them, Support Vector Machine (SVM), Logistic Regression (LR), AdaBoost and Static Bayesian Network (SBN) are the most popular ones. In this study, Bayesian Network is the main algorithm used for classification and analysis. Bayesian Network Toolbox (BNT) by Kevin Murphy was used to develop this Bayesian Network Model [50] and focus is more onto Dynamic Bayesian Network (DBN). The classification algorithms were used as for analysis part. SVM, LR, SBN and AdaBoost are used for comparison to DBN.

4.2 Driver Cognitive Distraction Detection Methods

As been discussed before, drivers can be distracted in three general ways: visual, cognitive and manual. Manual and visual distractions are easily measured and detected. However, to detect and measure cognitive distraction requires more effort than the other two types of distraction. An increase in traffic density might increase the complexity of the driving task too.

Driver's attention need to be divided between the application, equipment and device like steering, gear stick, mirrors, radio and many more in vehicle and the primary task of longitudinal and lateral vehicle control. Thus, cognitive workload assessment is of interest to many applications and systems. Many problems can be solved if the assessment can help to understand what a person is actually thinking about. Lie detection system, patient monitoring systems and even vigilant driving monitoring systems are a few applications and systems that require an advance technology to understand a person body gestures, facial expression, mimic and body or organ movements. For instance, a human body gesture can reflect what is actually in a person's mind. Some systems can understand and respond to a person's gestures even when the person is not giving any voice command. To design these types of systems, the major problem is to measure the movement displayed or shown by the person. There are a few available measurements which can be used to measure cognitive workload for drivers [51]:

(a) Performance measure (primary tasks and secondary tasks)

(b) Physiological measurement

(c) Rating scales

(a) Performance measure:

Study in [26] defined the primary task of the driver as maintenance of safe control over the vehicle. One of the major subtasks in vehicle control is lateral position control. Thus, a measure of driving deviations from the centre of a lane is a good means to assess primary-task performance in car driving. The task of keeping a vehicle between the lines of a lane is largely a psychomotor task which involves the eye-hand coordination. Primary task measurement is more onto making the drivers maintain the centre of lane and also maintaining a safe vehicle speed. Primary tasks are continuous tasks. Drivers are required to keep their car on the centre lane of a roadway and ensure the car is driving along with a proper and permissible speed limit. Failure to maintain the lane and speed can possibly lead to a crash or an accident.

If primary tasks measurements are continuous tasks in driving, secondary tasks measurement are non-continuous tasks. Examples are like car following (tailgating) and mirror checking. The main parameter in car following performance is the delay in reaction to speed changes of the lead vehicle. This can be obtained by computing a coherence analysis on the speed signals of the lead and following cars. Coherence is a measure of the accuracy of car following performance, while modulus indicates the amount of overreaction to speed changes by the following car [26]. Since secondary tasks are non-continuous tasks, car following and maintaining the speed in reaction to the speed changes of the lead vehicle do not occur all the time. The following drivers are not supposed to keep their speed at a certain limit because it can induce an accident if the leading drivers change their speed while the following drivers do not. The same goes with mirror checking action as this is another good example of an embedded secondary task that is specific for car driving. Generally when the mirror is being checked, two variables are measured respectively: glance frequency and glance duration, both represent different aspects of driver behaviour. Duration appears to be sensitive to difficulty of information intake, while glance frequency represents visual activity in terms of checking behaviour, both inside and outside of the vehicle (checking the speedometer, side mirrors checking).

(b)Physiological Measurement:

Physiological measurement measures and monitors a range of physiological parameters usually in major organ systems like face, head, mouth, eyes, skin, heart and others providing information on the extent of disability and the response to therapeutic interventions. Compared to other types of measurement, physiological measurement needs to perform in real time. Real time cognitive assessment would be a useful tool for building and testing more detailed cognitive models and the measures obtained could be used in real time to modify a task relative to the changing cognitive measures. Physiological measurements can be intrusive or unobtrusive. The unobtrusive physiological measures are performed and objectively analysed in a real time without impacting user performance and user ratings. Many studies nowadays focus on this type of measurement especially on eye movement. In [52] they have shown the difference between normal driving, visual distraction and cognitive distraction on the gaze distribution of a driver. The shading level presents the gaze density in Figure 4.1. Cognitive distraction changes the pattern of spatial and temporal eye gaze distribution [53]. By increasing the cognitive loading, it will make the eye's fixation longer, hence the gaze is more on concentration and less frequent in

glances to the mirrors and speedometer. During cognitive distraction as in Figure 4.1 below, the gaze tends to concentrate more densely in the centre of the driving scene and the area scanned shrinks compared to normal driving. This gaze concentration may reflect the competition for attention resources between cognitive demands of the secondary tasks and the eye movement. Gaze distribution was also used in [45][46]. From Figure 4.1 below, it is shown that during ordinary normal driving, the frontal focal points were scattered widely and bigger to the peripheral area which is presented as the inner most circle in each figure (a), (b) and (c) in the figure below. However, when the cognitive task was loaded or imposed, the areas of the frontal focal points were reduced within a narrower range. These can be seen in the figure (b) and (c) from the figure below.
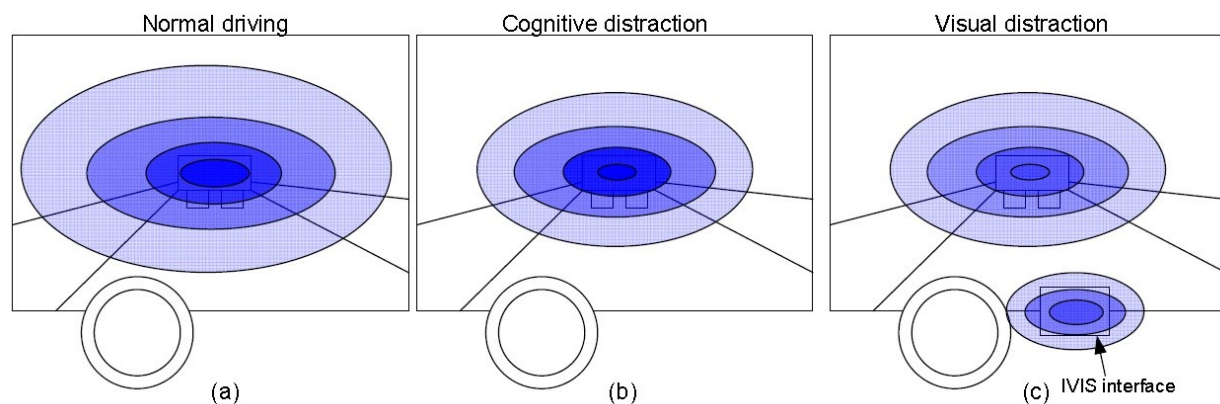
Figure 4.1: Gaze distribution in (a) normal driving, (b) cognitive distraction and (c) visual distraction) [45].

Pupil diameter is also popular eye information for driver distraction detection. In [45], they used pupil diameter for cognitive distraction in a driving simulation. The average value of pupil diameter was compared between arithmetic task and ordinary driving activity and they have found that from arithmetic task, an amount of increase was 13.1% compared to the ordinary one. Study in [45][46] also used head orientation as one of their pattern recognition features for driver cognitive distraction detection. Head movement was increased when the cognitive loading was imposed. The comparison between ordinary driving and driving with cognitive loading is shown in Figure 4.2 below. It is believed that this increased movement on the head was due to the compensatory action by which the driver attempts to obtain a field of wider view, in which may be thought as vestibule ocular reflex.

(Ordinary)   (Arithmetic)
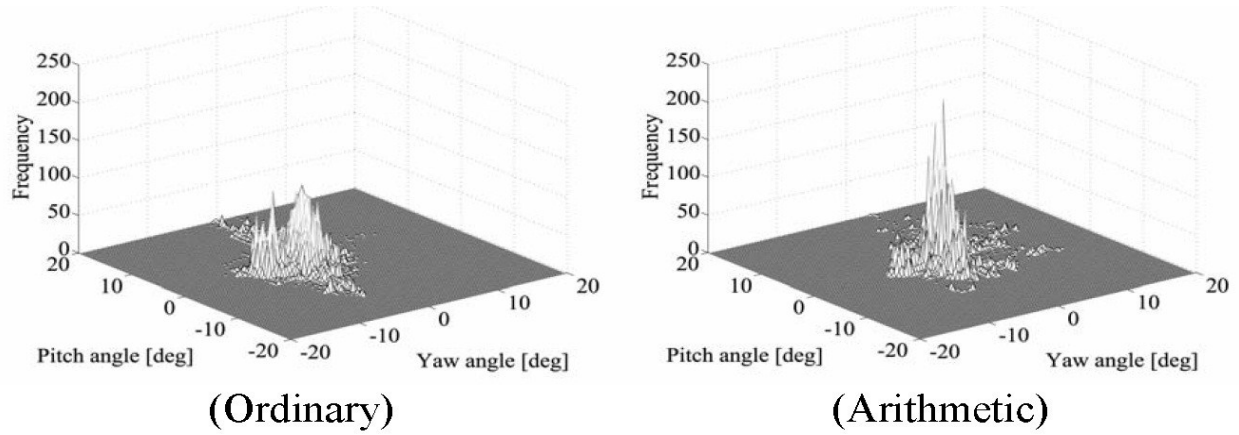
Figure 4.2: Head orientation comparison [45].

Since the movement of the head acts as a compensation for the movement of the eyes, the concentration of gaze distribution will also change the direction of the head. Therefore the value for head rotation angle was combined with the gaze rotation angle as in the formula used in [45] below:

$$x(i) = \sqrt{x_{pitch}(i)^2 + x_{yaw}(i)^2}$$

$x_{pitch}(i)$ = the pitch angle;

$x_{yaw}(i)$ = the yaw angle;

$x(i)$ = the combined gaze and head rotation angle.

Gaze distribution, pupil diameter and head orientation are the existing pattern recognition features for driver cognitive distraction. They are used with the proposed features in this study: eye height and width, mouth height and width, lips and eyebrows. The data fusion between those features gave a better improvement in detecting driver cognitive distraction. Details about the results from the data fusion with Support Vector Machine, Logistic Regression, AdaBoost and Bayesian Network are discussed in chapter 6.

Unlike visual distraction, cognitive distraction is more subtle, inconsistent and has relatively extended effects on driver behaviour. Therefore, there is no overt indicator to identify cognitive distraction from the performance measures mentioned. Even some studies used driving performance measures like lane position and steering wheel angle, unfortunately this driving performance method often reflects consequences of distraction and may be too late to mitigate the distraction when degradation in the driving occurs. Therefore, physiological measurement is

the best type of measure because it has close relationship with the driver's attention and it may help to identify the distraction before it may undermine the driving safety.

(c) Rating Scales:

The last method is by using rating scales. These measurements are subjective measures usually taken after the activity is completed. Even when interruptions occur, it will not degrade task performance, and ratings can be taken between or during the tasks of an activity. Subjective measures are mainly to assess the user's perceptions of the task and can provide information on the cognitive state of the user. However, since the rating scales are taken only after the activity, real-time cognitive state assessment is not possible to be done.

All those three methods are able to add cognitive assessment to the understanding of the user's cognitive state. However, only performance measures and physiological measurements can be measured in real time. From many types of physiological measurement, eye is one of the promising features used to detect driver distraction. Thus in this study, eye movement and other proven features like head orientation are to be combined with mouth, lips and eyebrows movement to give a better result in accuracy and sensitivity of detection. Eye movements include gaze direction, pupil diameter, and eyes height and eyes width. Table 4.1 below briefly compares the performance measures and physiological measures in four different categories.

CHARACTERISTICS	PERFORMANCE MEASURE	PHYSIOLOGICAL MEASURE
Selection	- Lane keeping, - brake behaviour, - steering angle, - stroke of pedal and brakes, - speed and acceleration, - turn signal, - wiper activation, - geographical position of the vehicle, - distance with the following vehicle	- Eye's movement, - gaze direction, - head movement, - lips biting, - shoulder stiffness, - galvanic skin conductivity (GSR), - face pose, body gesture, - eyes' fixation number, - eye's fixation duration, - pupil dilation, - heart rate, - blood pressure and circulation, - hand grips on steering wheel, - blinking rate and duration, - search patterns, - body temperature

Availability	Usually available at the end of a task	Available in real time
Affection	Can't be used to detect affection	Can detect affection
Sensitivity	Less sensitive- it is not really easy to determine that the driver is distracted or not just because the driver can't keep his lane.	More sensitive- when measure incorrect actions. Easily detect incorrect actions. For instance, if pupil dilated that means the driver is cognitively distracted, and this should indicate a warning, because the driver is doing an incorrect action while driving

Table 4.1: Performance Measurement and Physiological Measurement Comparison

It is known that these two methods of measurement are the most acceptable type of detection for driver's behaviour. Many researchers like in [24][36][45] used these two types of measurement.

4.3   Data Fusion

To produce more accurate and precise decision whether drivers are distracted or not while driving, a system must be developed with an integration of multiple data from multiple sources. Data captured by faceLAB and faceAPI can produce data in a variety of times and a variety of features from a driver's face. There are a few methods available to examine whether drivers are distracted or not as mentioned above. However, the best way to address this problem to ensure that the variety of data can be integrated is by constructing a data fusion system. Kutila et al. in [46] has mentioned that cognitive distraction may occur rapidly and can steal the driver's attention. One parameter alone will not reveal the distraction but rather by fusing many parameters, the robustness of the detection can be improved. Data fusion system can be used to integrate data from multiple sources, align data sets, correlate related variables, and combine them to support detection or classification decisions. Data captured need to be processed and the information need to be understood. Information understanding is done through data fusion and data mining, both are complementary processes that contribute automated process technologies to a variety of application domains. The automated processes are involved with abductive-inductive (learning and discovery) and deductive (detection) process. Abductive-inductive process refers to a data mining process whereas deductive is a data fusion process. Data fusion

and data mining implement fundamental reasoning processes that are mutually supportive in modelling targets and they detect instances in observed data [54]. Abduction creates a hypothesis of a model for specific sets of data, and this hypothesis is used to explain the specific set. Induction extends the model hypothesis done in abduction for the representative sets of data to make a general explanation. Deduction applies learned models as templates to infer the existence of an instance of an object in a set of data [55].

	General properties	Implementation
Abduction	Create model hypothesis for specific sets of data to explain that specific set.	Mining (discovery of models)
Induction	Extend model hypotheses for representative sets of data to make a general assertion or explanation	
Deduction	Apply models to create hypotheses to detect and classify (explain) the existence of target	Fusing (detection)

Table 4.2: Data mining and data fusion techniques

Data fusion benefit is that such a system is typically structured in a hierarchical form and can integrate information at different levels of abstraction. System can aggregate the sensor data to measure driver performance at the most concrete level and use these measures to infer driver state at higher levels. Moreover, data fusion can continuously refine assessments at different level of hierarchy independently to support real-time detection.

As defined briefly before, fusion is when different sources of information are combined to improve the performances of a system. The most obvious illustration of fusion is the use of various sensors typically to detect a target image. More generally, the different inputs may originate from a single sensor at different moments (fusion in time). This happens when data is captured in a sequence of time, where each single time frame produces different data, for instance Dynamic Bayesian Network. It can even happen in a single sensor at a given moment but on which several nodes in the network make various processing. This is a fusion of "opinion" when different nodes from different type of characteristics are joined together to create one complete and meaningful network.

Fusion may be useful for several objectives such as detection, recognition, identification, tracking, change detection, decision making and others. Using an efficient fusion scheme, significant advantages may be expected such as [54]:

a)   Improved confidence in decisions due to the use of complementary information

b)   Improved performance to countermeasures

c)   Improved performance in adverse environmental conditions.

Fusion processes are often categorized as low, intermediate or high level fusion depending on the processing stage at which fusion takes place. Low level fusion, also called data fusion, combines several sources of raw data to produce new raw data that is expected to be more informative and synthetic than the input. Intermediate level fusion or feature level fusion, various features such as edges, corners, lines and texture parameters are combined into a feature map that may then be used by further processes. There are two general approaches that can be used to implement data fusion to detect driver distraction: top down approach and bottom up approach.

 The examples for the top down approach are like Multiple Resource Theory (MRT) as in [25]. This approach identifies driver state based on only known characteristics of the driver and mechanisms of the driver's behaviour. This model only captures the trend of several performance measures, and not the changes of all measures that might be available for driver state estimation. More critically, such approach requires an accurate model of driver behaviour which is often difficult to develop. Another drawback of this approach is its theoretical basis may not be complete enough to describe a complex human behaviour especially when driving under cognitive distraction.

To encounter the problems in top down approach, a bottom up approach can be implemented by applying data mining methods to extract relationships between performance indicators and driver state from data rather than from theories of cognition [25]. As discussed before, data mining involves the abduction and induction process. Data mining searches large volumes of data for unknown patterns. In this driver distraction problem, driver behaviours when being cognitively distracted are studied. Movements especially on driver's face are learnt and recognized and could be used as hypothesis for a model or template. Once a model is discovered, the hypothesis will be used to detect and classify the target. There are many available data mining techniques: regression, decision tree, neural network, Bayesian network, Hidden Markov model (HMM), Kalman filter, Adaboost and Support Vector Machine. So far, for driver cognitive distraction detection, only regression, logistic and linear [36;47;56], AdaBoost [45], SVM [36;46;57] and Bayesian Network [24;58] are suitable algorithms and have been implemented in other research studies. Cognitive distraction may depend on more complex techniques and integrate more

indicators, which require a bottom up technique. When building detection algorithms, individual differences are important to consider. Driving data may contain two sources of variation:

a) Intra difference- Within a driver: Occurs when a driver exhibits different driving behaviours under different circumstances, such as attentive driving and distracted driving.

b) Inter difference- Between drivers: Occurs when different drivers with definite different behaviours under different circumstances are being compared.

4.4   Bayesian Network

4.4.1   Static Bayesian Network

Detecting driver's cognitive distraction is not as detecting driver's visual and manual distractions. Most current approaches or algorithms to detect driver cognitive distraction are static or discrete, and they cannot provide a continuous measure of distraction. The challenge to detect cognitive distraction is to integrate a large number of data from a few numbers of parameters, in a logical manner and comprehensively infer the driver's cognitive state. This cannot be done precisely with the top down approach, because this theory driven approach is not feasible. Although some theories have been used to make predictions regarding driver distraction, unfortunately, they are often better at describing rather than predicting. Thus, to predict cognitive distraction, a bottom up data mining technique is required. Bayesian Network is one of the best techniques because it covers both static and dynamic models and the data can be discrete or continuous [24] [36].

As been mentioned before, Bayesian network has been used in this study. Several parameters or features were captured by faceLAB and faceAPI. Different sensors for different parameters created the data fusion technique. In this study, faceLAB has been used to capture: (a) mouth height, (b) mouth width, (c) eye height, (d) eye width, (e) head rotation/movement, (f) gaze rotation, (g) pupil diameter and (h) blinking frequency whereas faceAPI has been used to capture: (a) lips height, (b) lips width, (c) eyebrow height, (d) eyebrow width, (e) head position and also can be used to capture (f) head rotation/movement like from the faceLAB camera.

Bayesian Network is simply a graphical model with a set of random variables making up the nodes of the network. Nodes in BN can be either discrete or continuous and usually they are connected with directed links or arrows. If there is an arrow from node X to node Y, X is said to be a parent of Y. Each node Xi has a conditional probability distribution P(Xi|Parents (Xi)) that quantifies the effect of the parents on the node. Thus, it is a network model for representing conditional independencies or probabilities between that set of random variables. Network graph must be an acyclic graph with direction or sometimes called as DAG (directed acyclic graph) [59].

Consider a network with four random variables, A, B, C and D. These four random variables can be factorized as a product of conditional probabilities:

$$P(A,B,C,D) = P(A) \bullet P(B \mid A) \bullet P(C \mid A,B) \bullet P(D \mid A,B,C)$$

The above equation shows that each variable potentially depends on every other variables. However, if:

$$P(A,B,C,D) = P(A) \bullet P(B) \bullet P(C \mid A) \bullet P(D \mid B,C) \qquad (1)$$

it implies a set of conditional independence relationship. A variable X is conditionally independent from Y given Z if $P(X,Y \mid Z) = P(X \mid Z) \bullet P(Y \mid Z)$ for all X, Y and Z such that P(Z) ≠ 0. It can be shown that given the values of B and C, the values of A and D are independent:

$$P(D,A \mid B,C) = \frac{P(A,B,C,D)}{P(B,C)}$$

$$= \frac{P(A) \bullet P(B) \bullet P(C \mid A) \bullet P(D \mid B,C)}{\int P(A) \bullet P(B) \bullet P(C \mid A) \bullet P(D \mid B,C) dAdD}$$

$$= \frac{P(A) \bullet P(C \mid A) \bullet P(D \mid B,C)}{P(C)}$$

$$= P(A \mid C) \bullet P(D \mid B,C)$$

Therefore, to represent the factorization in (1) an arc from A to C is possible to draw but not from A to D. This factorization network can be seen in the Figure 4.3:
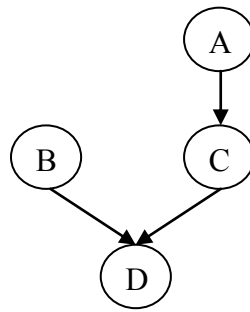
Figure 4.3: Bayesian Network for (1)

Node A is the parent for node C since there is a directed arc from A to C. C is therefore a child node to node A. However, node A is not a parent to node B. The descendants of a node are its children, children's children and so on [60]. Node D's parents are node B and node C.

Directed and undirected path in Bayesian Network is important. A directed path from node A to B is a sequence of nodes starting from node A and ending with node B such that each node in the sequence is a parent node to the following node in the sequence. An undirected path of node A to B is a sequence of nodes from node A and ending to node B such that each node in the sequence is either a parent or a child of the following node.

Semantically, each node is conditionally independent from its non-descendants given by its parents. Like in Figure 4.3, two disjoint sets of nodes A and B are conditionally independent given that C, if C is the d-separated A and B. d-separates definition is if along undirected path between a node in A and a node in B there is a node D, such that:

a) D has converging arrows. That is, D is a child of both the previous and the following nodes in the path and neither node D nor its descendants are in node C, or

b) Node D does not have converging arrow and D is in C.

Therefore, it is easier to infer many independent relations in the graphical model without explicitly implying Bayes' rule. For instance, like in Figure 4.3 above, A is conditionally independent from B, given that the set of X={C, D}. Since C ϵ X is along the only path between A and B, and C does not have the converging arrows. Unfortunately, A cannot be inferred as conditionally independent from B given D.

Since a strict ordering of the variables implies, the connections between variables define a directed acyclic graph (DAG). The undirected graphical model like Markov Model is another

type of Bayesian Network which represents probability distributions but with a different set of semantics.

A BN graph *G*, is said to be an independency map for a distribution *P* if every d-separates displayed in *G* corresponds to a valid conditional independence relation in *P*. The absence of arcs in BN implies conditional independence relations which can be exploited to obtain efficient algorithms. These algorithms are used to compute marginal and conditional probabilities between the nodes in *G*.

Let say evidence is observed. Evidence is the value of some variables in the network. Belief propagation is a single connected network in which the undirected graph has no loops, is objectively to update the marginal probabilities of all the variables in the network with this new evidence. This is done by message passing technique like in Figure 4.4 below. Each node, *n*, will pass a message to its parents and to its children about the evidence. *n* will separate the graph and therefore the evidence will also separate into two mutually exclusive sets: $e^+(n)$ and $e^-(n)$.

a) $e^+(n)$ consists of the parents of *n*, the nodes connected to *n* through its parents and *n* itself.

b) $e^-(n)$ consists of the children of n and the nodes connected to n through its children.
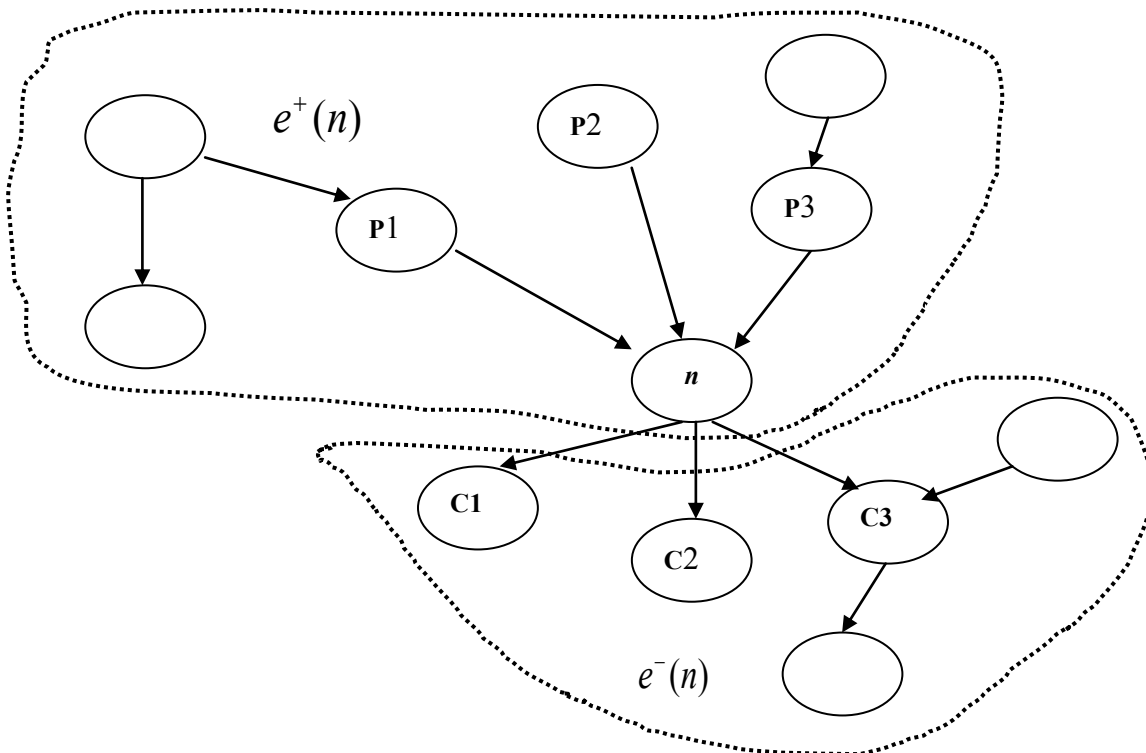
Figure 4.4: Message passing in network [60].

The message from n to each of its children is the probability of each setting of n given the evidence observed in the set $e^+(n)$. The message passing from n to each of its parents is the probability, given that every setting of the parent, of the evidence is observed in the set $e^-(n) \cup \{n\}$.

Marginal probability of a node is proportional to the product of the messages obtained from its parents. It is weighted by the conditional probability of the node given by its parents and the message obtained from its children. Thus, if the parents of n are $\{p_1,........,p_k\}$ and the children are $\{c_1,........,c_x\}$, the marginal probability is given by:

$$P(n \mid e)\alpha \left[\sum_{\{p_1,........p_k\}} P(n \mid p_1,........,p_k) \prod_{i=1}^{k} P(p_i \mid e^+(p_i)) \right] \prod_{j=1}^{x} P(c_j, e^-(c_j) \mid n)$$

where the summation extends all the settings of $\{p_1,........,p_k\}$.

For instance, given the evidence $e = \{B=b, D=d\}$

$$P(C \mid B=b, D=d)\alpha \left[\int P(C \mid A) \bullet P(A)dA \right] P(D=d, B=b \mid C)$$

$$\alpha P(C) \bullet P(D=d \mid B=b, C) \bullet P(B=b)$$

P(A) is the message passed from A to C since $e^+(A) = \Theta$ and P(D=d, B=b | C) is the message passed from D to C.

Variables in the evidence set are referred as observed variables and those which are not referred as hidden variables.

BN is usually developed by combining a prior knowledge about conditional independencies between the variables like a data set of observations.

Prior knowledge can be elicited by causality. Causality is when a variable has a direct causal effect on another variable. Thus, that variable will be parent in the network. Dynamic Bayesian network is specified by a temporal order and the direction of the causality.

As been stated previously, BN can be either SBN or DBN. Both SBN and DBN in this study were constructed by Bayesian Network Toolbox written by Kevin Murphy [50]. MATLAB is the platform for this toolbox. Static Bayesian Network only works with variables from a single slice of time frame. Thus SBN cannot work for analysing an evolving system that changes over time. In the initial experiment, parameters were collected with faceLAB. The SBN graph for the first

experiment was like in Figure 4.5 below: 8 variables have been chosen. All variables are with continuous variables. However all variables later were discretized using a decision tree algorithm to maximize information gain [24].
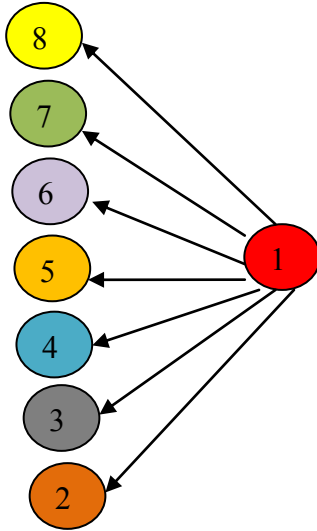
Figure 4.5: SBN for initial experiment

No.	Parameters
1	Distraction
2	Mouth Height
3	Mouth Width
4	Eye Height
5	Eye Width
6	Head Rotation
7	Gaze Rotation
8	Pupil Diameter

Table 4.3: Parameters for initial experiment

Variable 1-distraction is the only hidden variable in this network. Variables 2 until 8 are the parameters in this network and they are all observed. faceLAB provides and tracks many other parameters and features like blinking frequency, face rectangle, PERCLOS and saccade. However, in this study, based on literature comparison and initial experiment data, only head rotation, gaze rotation and pupil diameter are selected and will be combined with the proposed features: eye movement (height and width), mouth movement (height and width), lips and eyebrows.

Simply to create a Bayesian Network, the random sampling process will generate events from a network that has no evidence which is related to it. Each variable must be in topological order. As in the diagram shown in the Figure 4.5 above all nodes or variables have been linked in an order. The probability distribution from which the value is sampled is conditioned on the values already assigned to the variable's parents. The sampling algorithm below generates sample from the prior joint distribution specified by the network. First, let $S(x_1,....,x_n)$ be the probability that a specific event is generated. Thus, the sampling algorithm will give :-

$$S(x_1,.....,x_n) = \prod_{i=1}^{n} P(x_i \mid parents(X_i)).$$

This is because every sampling step depends on the parent values and the equation in the algorithm satisfies the probability of the event according to Bayesian net's representation of the joint distribution:- $S(x_1,.....x_n) = P(x_1,.....,x_n)$. Thus, to create a Static Bayesian Network the algorithm is just like in the sample algorithm 1 below:

```
Input      : n, number of nodes;
             dag, directed acyclic graph-connecting the nodes;
             ns, node size;
             CPD=tabular_CPD the parameters conditional probability
             distribution;
             evidence;
output     : mk_bnet, create a static Bayesian network
             m=marginal_nodes, marginal distribution on nodes
% For details see http://www.cs.berkeley.edu/~murphyk/Bayes/
1   N = number of nodes;
2   Create the DAG:
3   dag = zeros(N,N);
4   A = 1; B = 2;C = 3; D = 4; %define the nodes
5   dag(A,[B C]) = 1;
6   dag(B,D) = 1;
    % 1=false=not distracted ; 2=true=distracted
7   false = 1; true = 2;
8   Node size:
9   ns = 2*ones(1,N);
10  Create the bnet for driver cognitive distraction:
11  bnet = mk_bnet(dag, ns, 'names', {'A','B','C','D'}, 'discrete', 1:N);
12  Parameters represented by CPD objects:
13  bnet.CPD{A} = tabular_CPD(bnet, A, [0.5 0.5]);
14  bnet.CPD{B} = tabular_CPD(bnet, B, [0.8 0.5]);
15  joint = zeros(2,2,2,2);
16  for a=1:2
17    for b=1:2
18      for c=1:2
19       for d=1:2
20          joint(a,b,c,d) = CPD{A}(a) * CPD{B}(a,b) * CPD{C}(a,c) *
            CPD{D}(b,c,d);
21         end
22       end
23     end
24  end
25  inference engine- use Junction tree:
26  engine = jtree_inf_engine(bnet);
27  Add in evidence:
28  evidence = cell(1,N);
29  engine, ll] = enter_evidence(engine, evidence);
30  Marginal distribution:
31  m = marginal_nodes(engine, [B C D]);
32  Joint distributions:
33  evidence = cell(1,N);
34  engine, ll] = enter_evidence(engine, evidence);
35  m = marginal_nodes(engine, [B C D]);
```

Algorithm 4.1: Sample of SBN algorithm

With BN process explained before and the algorithm shown above, the computation process for a SBN for this study can be computed like below. Assume there are only five nodes in the network. The first node is the hidden variable which belongs to the cognitive distraction and denoted as C. The other four are the observed nodes which are referred as the parameters observed in this experiment and they are denoted as W, X, Y and Z. Node C is the parent node. Nodes W, X, Y and Z are the children nodes from node C. Therefore, the network for this example will look like Figure 4.6 below:
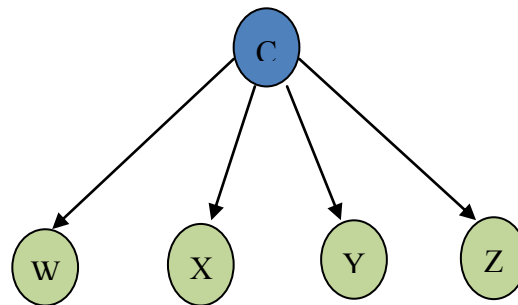
Figure 4.6: SBN example

Let say W is the head orientation, X is the mouth movement, Y is the eye movement and Z is the gaze rotation. Consider the evidence for the variables are as following:

P(C=T)	P(C=F)
0.5	0.5

Table 4.4: Node C

W	C	
	T	F
T	0.8	0.1
F	0.2	0.9

Table 4.5: Node W given that C

X	C	
	T	F
T	0.5	0.4
F	0.5	0.6

Table 4.6: Node X given that C

Y	C	
	T	F
T	0.7	0.6
F	0.3	0.4

Table 4.7: Node Y given that C

Z	C	
	T	F
T	0.8	0.2
F	0.2	0.8

Table 4.8: Node Z given that C

T is when the variable is True and F is when the variable is False. A tree diagram can be used to represent the evidence in a more understandable presentation.
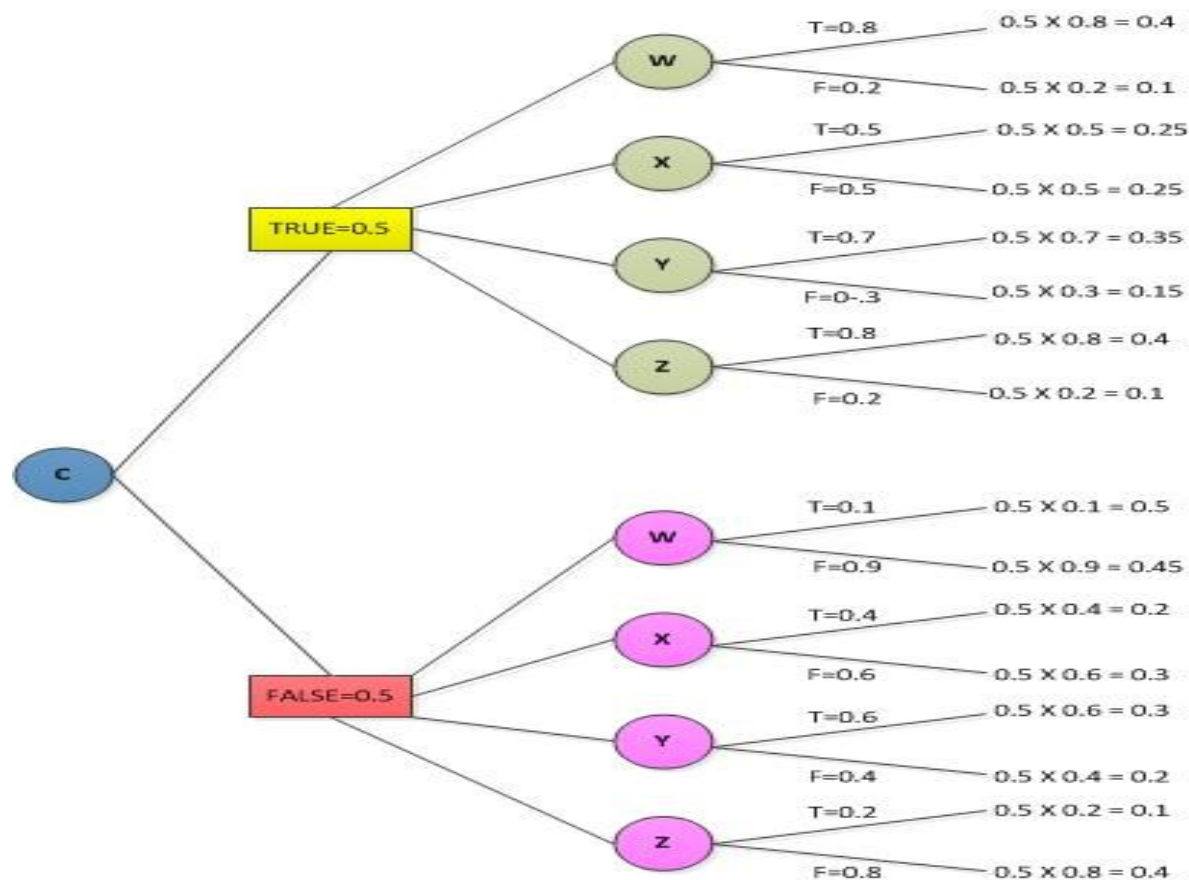
Figure 4.7: SBN tree diagram

The tree diagram above, shown on how a computation on conditional probability distribution can be computed on a five nodes network of SBN. Consider an example of probability for mouth movement when the cognitive distraction is given, the probability value for this event is equal to 0.25. Since the number of observed nodes in this example is equal to 4, therefore, the computational complexity is equal to $2^4 = 16$. The computation on probability distribution can be more complex if the network has two layers or more.

The second experiment in this study was done with faceAPI toolkit. Different parameters or features were captured in this second experiment. Therefore, different model of Bayesian Network was also constructed. However, the computation and operation in this second experiment is similar to the first experiment. 9 nodes are constructed in the network. Node 1 still

presents cognitive distraction and it is the only hidden node in the network. The other 8 nodes are as in Figure 4.8 and Table 4.9 below:
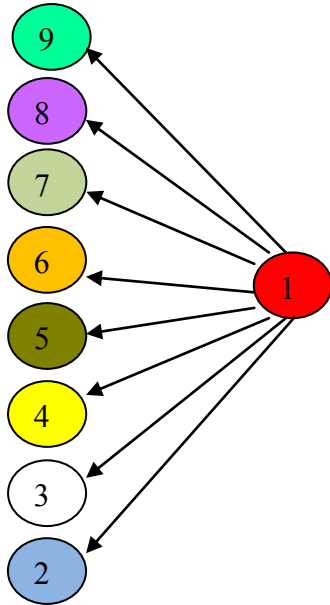
No.	Parameters
1	Distraction
2	Lips Height
3	Lips Width
4	Right Eyebrow Height
5	Right Eyebrow Width
6	Left Eyebrow Height
7	Left Eyebrow Width
8	Head Rotation
9	Head Position

Figure 4.8: SBN for second experiment   Table 4.9: Parameters for second experiment

Eyebrows are included with both sides left and right in this model because with faceAPI the face landmarks are captured specifically. FLMs from both sides of the eyebrows were then computed to give the eyebrows width and height. Details about face landmarks in faceAPI are explained in the next chapter. Both eyebrows contained outer, inner and the centre landmarks. Like eyebrows, FLMs from lips are also used to compute lips height and width. The inner and outer lips information is used for this purpose.

Imagine if the network is repeated in a sequence of time. Thus, SBN cannot be used anymore and need to be replaced with a better Bayesian network call Dynamic Bayesian Network.

Static Bayesian network for the final experiment basically contained 11 nodes. Since the final experiment had camera fusion between faceLAB and faceAPI, the data fusion is said to have happened in this final experiment. Data from faceLAB was combined with data from faceAPI. Data from faceLAB are: blinking frequency, gaze rotation and head rotation. Data from faceAPI are: lips and eyebrows. Gaze rotation and eyebrows are taken from both left and right eyes. Lips are taken from inner contour and outer contour and finally head rotation is covered by α (head

rotation at X axis), β (head rotation at Y axis) and d γ (head rotation at Z axis) angles. The SBN for this final experiment is shown in the figure below:
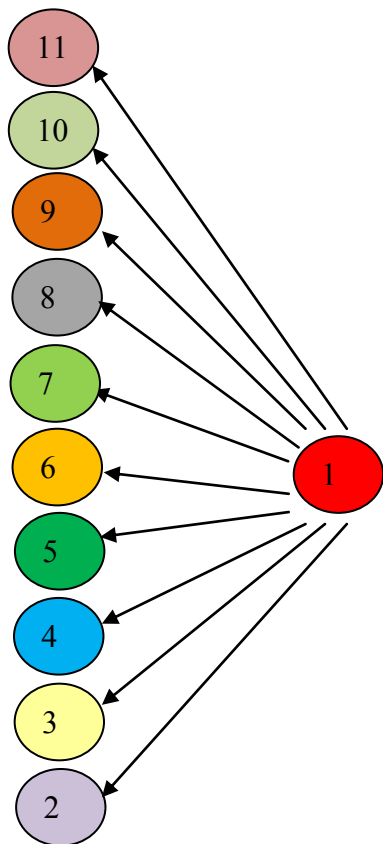
No.	Parameters
1	Distraction
2	Lips Height
3	Lips Width
4	Right Eyebrows Height
5	Right Eyebrows Width
6	Left Eyebrows Height
7	Left Eyebrows Width
8	Blinking frequency
9	Right Gaze Rotation
10	Left eye gaze rotation
11	Head rotation

Figure 4.9: SBN for third experiment   Table 4.10: Parameters for third experiment

Imagine if the network is repeated in a sequence of times. Thus, SBN cannot be used anymore and need to be replaced with a better Bayesian network call Dynamic Bayesian Network.

4.4.2   Dynamic Bayesian Network

Dynamic Bayesian Network is used to describe how variables influence each other over time based on the model derived from past data. A DBN can be described as a Markov Model Chain. Markov Model is a subset model from DBN. DBN has many states with a discrete time approximation of differential equation with time frames. DBN will illustrate the probabilities of

one variable changing another, and how each of the individual variables will change over time. Once the network has been established between the time frames, a model can be developed based on the data. The model then can be used to predict future responses by the system. The ability to predict future responses can also be used to explore different alternatives for the system and determine which alternative gives the desired results. This ability is really good for this study. DBN also provides a suitable environment for a general network that does not depend on time. Once the DBN has been established with its time frames, the network can be collapsed to remove the time component and show the general relationship between the variables. DBN is an interconnected time frame of SBNs. The nodes at certain times can affect the nodes at a future time frame, but the nodes in the future can not affect the nodes in the previous time slice. The causal links across the time frames are referred to as temporal links, where they give DBN an unambiguous direction of causality. For the convenience of computation, the variables in DBN are assumed to have a finite number of states that the variable can have and this builds up a conditional probability table (CPT) in a DBN. This CPT can be constructed to express the probabilities of each child node derived from condition of its parent nodes.
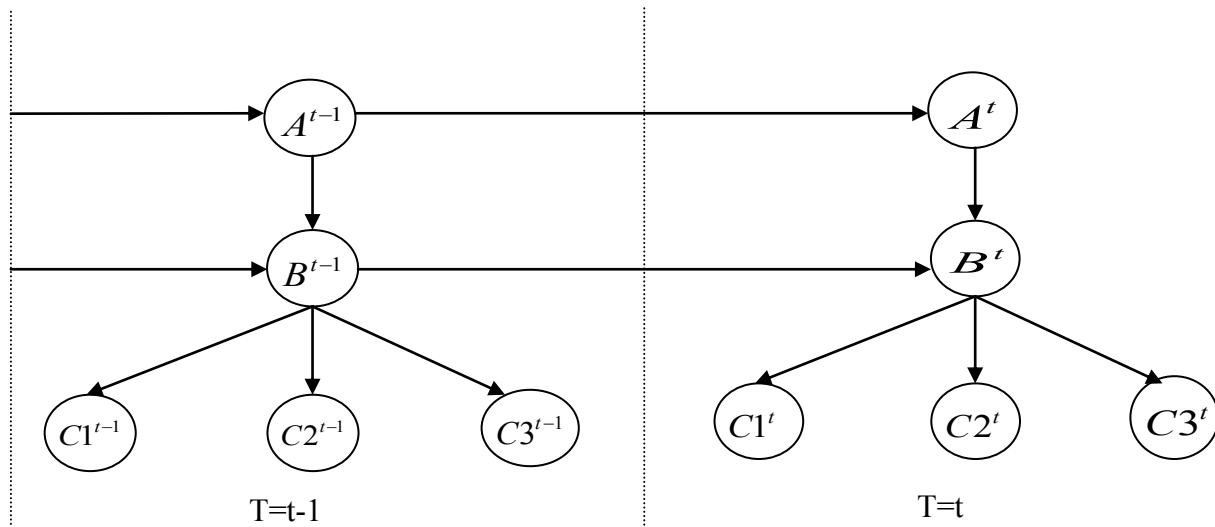
Figure 4.10: Dynamic Bayesian Network

[36] and [24] showed that BNs could identify driver distraction reliably with an average of 86.4% from DBN and 78% from SBN compared to SVM with 81.1% and Regression algorithm with only 72.7%. DBNs are considered as time dependencies of driver behaviour and produced more sensitive models than SBNs. Longer training sequences are suspected to give better results

in DBNs performance too. Thus, it is a definite good choice to focus onto BNs as algorithms to detect driver cognitive distraction because BNs can extract information from a simple to a complex behavioral data.

Many assumed that HMM (Hidden Markov Model) is similar to DBN. Unfortunately, they are not similar. As been mentioned DBN is a DAG model for stochastic processes. DBN generalizes HMM model by representing hidden and observed state in terms of state variables, which might have complex interdependencies. Therefore, a graphical structure provides an easy and better way to specify the conditional independencies. On the other hand, HMM is actually the simplest kind of DBN. HMMs can only have one discrete hidden node and one discrete or continuous observed node in every time slice as like in the figure below:
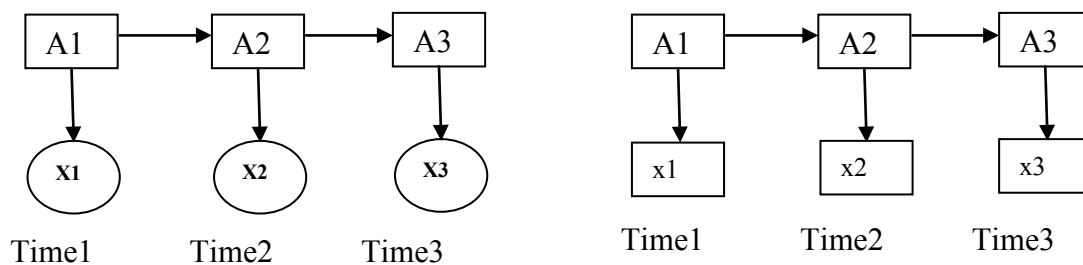
Figure 4.11: HMM examples

Node in the rectangle is a discrete variable and the node in a circle is a continuous variable. It is clear that every hidden Markov model can be represented as a DBN with a single state variable and a single evidence variable. However the difference between HMM and DBN is that DBN can decompose the state of a complex system into its constituent variables, and it is able to take advantage of sparseness in the temporal probability model. Suppose, DBN has 10 Boolean state variables, where each of which has four parents in the preceding time slice. Then, the DBN transition model has probabilities and unfortunately, a corresponding HMM has states and therefore and later a trillion probabilities in the transition matrix. Thus, HMM is less advantageous than DBN because:

a) HMM requires much more space compared to DBN

b) HMM requires huge transition matrix, thus the inference will be more expensive and complicated

c) HMM is unsuitable for large problems because of its problem to learn such a huge number of parameters

In this study, the DBN cognitive model used information of distractions to define driver cognitive state either distracted or not distracted respectively as the hypothesis (hidden) node. The evidences used in this BN models will sometimes be called as instances. The model contained 8 nodes of instances. In [61], a fatigue model is constructed also using the DBN model. However in his model, additional contextual information is added to the model. The contextual information for the fatigue model refers to the symptoms of fatigue such as work environment, weather, temperature, noise, sleep time, sleep quality and etc.

This contextual information is not included in this study because, the information is assumed to be fixed in the model. Similar to a driver when he/she is fatigued, he/she tends to exhibit various visual behaviours that deviate from the nominal behaviours. These behaviours typically can capture the cognitive state of a person which can be seen in the eye movement and head movement. In DBN a transition of data for every parameter from the hidden nodes and observed nodes are checked at every single time slice, t. It contains temporal information from previous time and current time to calculate the probability of the next occurrence. Bayesian approach learning starts with some prior knowledge about the model structure, the set of arcs in the BNs and model parameters. This initial knowledge is represented in a form of prior probability distribution and updated using data to obtain a posterior probability distribution over models and parameters. Assume that a prior distribution over models structure is P(M) and a prior distribution over parameters for each model is P(p | M), a data set D is used to form a posterior distribution over models using Bayes Rule [60]

$$P(M \mid D) = \frac{\int P(D \mid p, M) P(p \mid M) dp P(M)}{P(D)}$$

which integrates out the uncertainty in the parameters. For a given model structure, the posterior distribution over the parameters is computed by:

$$P(p \mid M, D) = \frac{P(D \mid p, M) P(p \mid M)}{P(D \mid M)}$$

A Dynamic Bayesian Network basically contains a pair of information $(G; \Theta)$ that models the temporal process by specifying a probability distribution for $X^0, \ldots, X^T$; $P(X^0, \ldots X^T \mid G; \Theta)$ where $X^t = \{X^t_1, \ldots, X^t_I\}$ is a set of I discrete random variables that represent the state of a temporal process at a discrete time point t [61]. G is a directed acyclic

graph (DAG) with the nodes correspond to the random variables in X^0 and X^1. Θ, is a set of parameters specifying a conditional probability distribution for each node X_i^t in G given its parents $Pa(X_i^t)$ in G, $P(X_i^t | Pa(X_i^t), G, \Theta)$. DBNs also can be defined by a pair of $(B_o, B_\rightarrow)$ where B_0 is a prior knowledge or network and $B_\rightarrow$ is a transition network. In DBN, joint distribution is important, given two random variables X and Y defined on the same probability space, the joint distribution for X and Y defines the probability of events defined in terms of both X and Y. For a given DBN model the joint distribution for X over X[0], X[1], ..., X[T] is:

$$P_{DBN}(x[0], x[1], x[2], \ldots x[T]) = P_{B_o}(X[0]) \prod_{t=0}^{T-1} P_{B_\rightarrow}(x[t+1] \mid x[t])$$

where $P_{B_\rightarrow}$ is a transition model. Thus, the algorithm to create a DBN is as in the sample of DBN algorithm 2 below:

```
Input      :      ss, number of nodes;
                  intra, intra connection between nodes;
                  inter, inter connection between nodes;
                  onodes, observed nodes;
                  hnodes, hidden nodes;
                  dnodes, discrete nodes;
                  CPD and file input;
                  eclass, parameter tying;
                  evidence;
                  T, time slices;
Output     :      mk_dbn, create a Dynamic Bayesian network
                  m=marginal_nodes, marginal distribution on nodes

1   DBN for driver cognitive distraction:
2   function bnet = driver_dbn(discrete_obs, obs_leaves)
    % bnet = driver_dbn (discrete_obs, obs_leaves)
    % If discrete_obs = 1 (default), the leaves are binary
    % If obs_leaves = 1, all the leaves are observed
3   Network topology:
4   Number of nodes = n
5   ss = number of nodes;
6   Intra connection:
7   intra = zeros(ss);
    intra(1,n) = 1;
    intra(2,n) = 1;
8   Inter connection:
9   inter = zeros(ss);
10  inter(1,n) = 1;
11  Observed leaves:obs_leaves = observed nodes
12  if obs_leaves
13            onodes = n:n++; % leaves
```

```
14          else
15                  onodes = [n]; % throw in some other nodes
16          end
17  end
18  Hidden node:
19  hnodes = 1:n;
20  Discrete node:
21  dnodes = 1:ss;
22  Parameter tying…details in [Bayesian Network Toolbox [50]]:
    We assume the distributions P(X(t)|X(t-1))and P(Y(t)|X(t))are
    independent of t for t > 1. Define the equivalence class here.
23  eclass1 = 1:j;
24  eclass2 = [j:k j:l];
25  Create the DBN:
26  bnet = mk_dbn(intra, inter, ns, 'discrete', dnodes, 'eclass1', eclass1,
    'eclass2', eclass2,'observed', onodes);
27  CPD tabular table:
28  for i=1:max(eclass2)
29          bnet.CPD{i} = tabular_CPD(bnet, i);
30  end
31  Learning parameters with learn_params_dbn_em:
32  [bnet2,LLtrace]=learn_params_dbn_em(engine);
33  Inference engine-jtree_dbn_inf_engine:
34  engine=jtree_dbn_inf_engine(bnet);
35  Evidence and T is referring to the time slice:
36  evidence=cell(n,T);
37  evidence(onodes,:)=num2cell(evidence2(onodes, :));
    %converting the evidence on observed nodes from number to cell.
38  Marginal probability distribution:
39  for i= 1:T
40          marg=marginal_nodes(engine, 1, i);
41  for z=1:2 %z is referring to the first node/hidden node/binary 1=not
    distracted, 2=distracted
42          marg.T(z)
43  Threshold:
44  if (marg.T(z)>threshold)
45          if (evidence (1,i)==z)
46                  truevar=truevar + 1; %compute the true variable
47          end
48  end
49  end
50  Evidence on the engine:
51  [engine, loglik]=enter_evidence(engine, evidence);% evidence is put
    into the inference engine (jtree_dbn_inf_engine)
52  end
```

Algorithm 4.2: Sample of DBN algorithm

It can be seen that, to construct a DBN, the algorithm is almost similar like in SBN. However, DBN is required a time slice, thus the variable T in the sample algorithm above is represented as the time slice in the network. Marginal probability distribution is computed in each time slice. The equivalence classes also have been used in this DBN algorithm. The equivalence classes are

used to define the interconnection between nodes in the network. By default, eclass1 = 1:ss, and eclass2 = (1:ss)+ss, where ss = slice size = the number of nodes per slice. This equivalence classes are actually a part of parameter tying process. For a network which has a repeated structure either chains or grids like DBN, it is very common to assume that the parameters are all the same at every node and at every time slices. Parameter tying reduces the amount data needed for learning. Tied parameters eliminate the one to one correspondence between nodes and conditional probability distributions (CPDs). Each CPD specifies the parameters for a whole equivalence class of nodes.

Just like in SBN, DBN in this study was also built for two sets of experiments. The first set of DBN was for the initial experiment with faceLAB camera. This first network contained 8 nodes: Distraction, Mouth Height, Mouth Width, Eye Height, Eye Width, Head Rotation, Gaze Rotation and Pupil Diameter. Rather than having one network, in DBN, the network is repeated in T time slices. faceLAB and faceAPI used 60 time slices or time frames equal to 1 second (60 frames/second). In Figure 4.12 below, it only showed an example of the DBN network up to 3 time slices. The first time slice (*n-1*) is the previous time slice. The second time slice (*n*) is current time slice and the third time slice (*n+1*) refers to the next time slice. Parameters in DBN are all the same in every time slice.

First node in the network is a distraction node which is binary either distracted or not distracted and BN model used this information to define the hypothesis (hidden) node. The children nodes in the network are the parameters with size from 1 to 10.
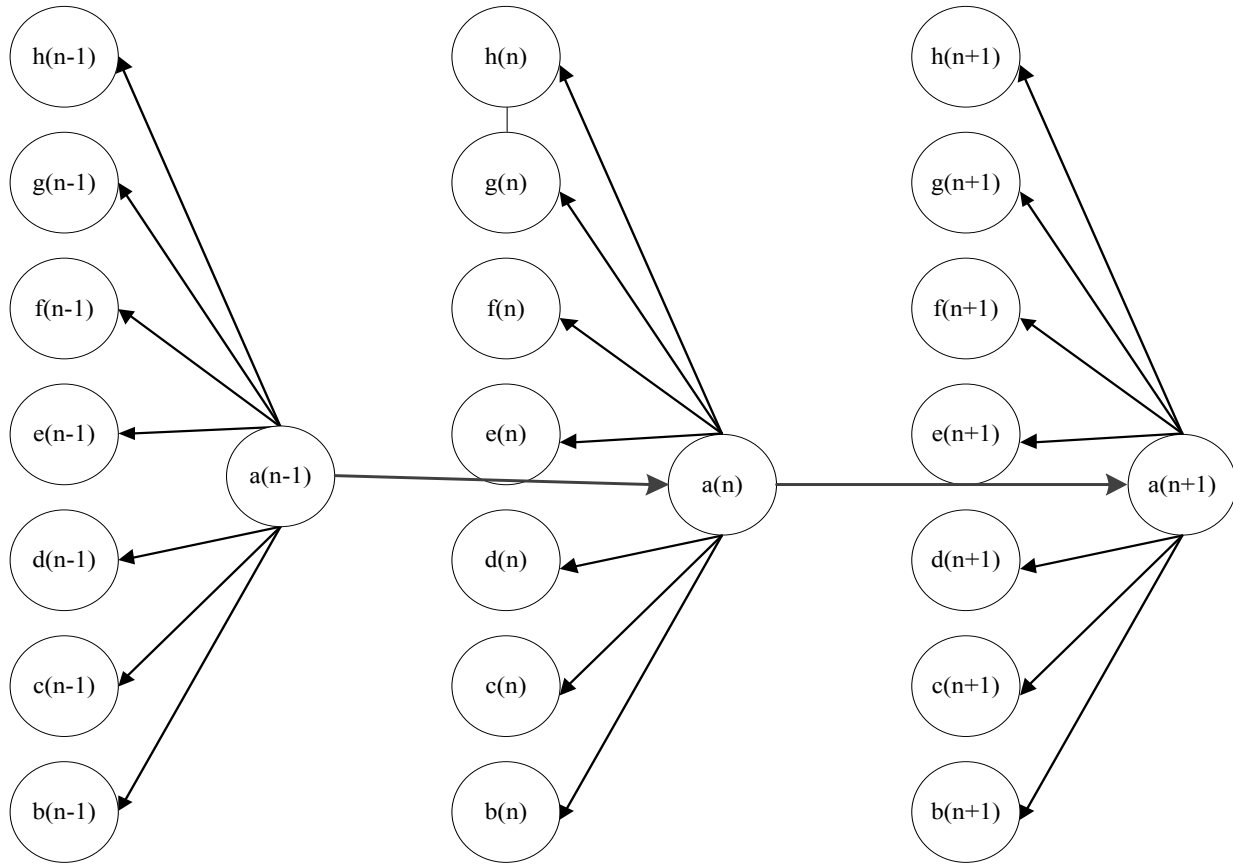
Figure 4.12: DBN for initial experiment

Data collected by faceLAB and faceAPI was measured in a continuous value. However, since DBN can be more complex with continuous value, in this study, all the parameter values were converted to discrete values by using Decision Tree Algorithm. Paper [24] also converted their data into discrete to eliminate the complexity and to maximize the information gain. A variety of BN models were trained and tested for every participant in this study. Both experiments followed the typical data mining procedure. Two-thirds of the participants were chosen at random to form the training set and one-third was for the testing set. The training of all BN models was included with parameter learning. Parameter learning justifies the conditional probability distribution for those connections made. Second experiment's network contained 9 nodes as shown in Figure 4.13: Distraction, Lips Inner, Lips Outer, Right Eyebrow Height, Right Eyebrow Width, Left Eyebrow Height, Left Eyebrow Width, Head Position and Head Rotation. Similarly, like in the first experiment, all nodes are discrete and are repeated in every time slice. Information from

inner lips and outer lips were used to compute the lips width and height. Information from right eyebrows and left eyebrows were also used to compute eyebrows' width and height.
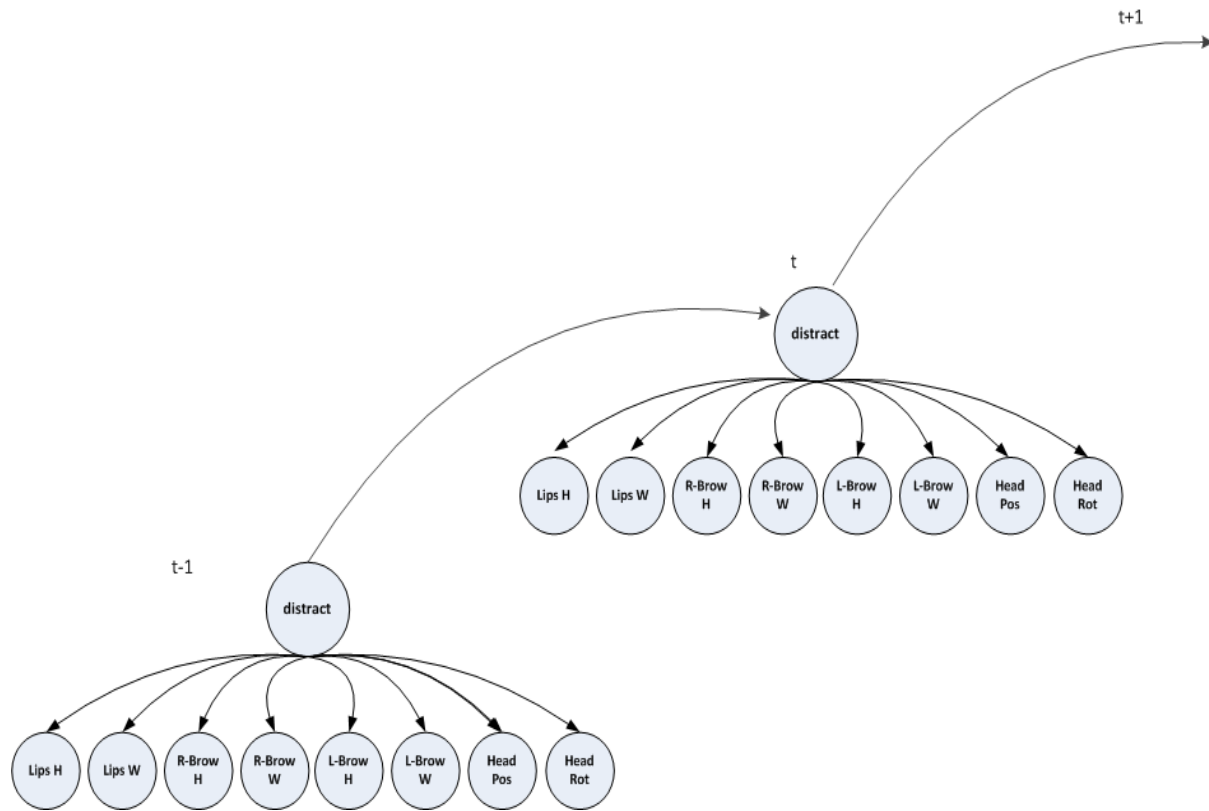
Figure 4.13: DBN for second experiment

The BN models in this study were constructed with Bayesian Network Toolbox (BNT) by Kevin Murphy [50]. The toolbox is in Matlab platform and it is accompanied by other learning packages. As in the algorithm before, DBN model has to define its intra connection and inter connection between nodes in the network. Intra connection is when the links are within a time slice and will only be presented in the first time slice. After the first time slice, the nodes were linked only with those in the previous time slice and this is called as inter connection. For DBN, those two structures, inter connection and intra connection are required. However, only one structure was trained in SBNs.

The final experiment contained 11 nodes as in Figure 4.14: distraction, blinking frequency, right eye gaze rotation, left eye gaze rotation, head rotation, lips height, lips width, right eyebrow height, right eyebrow width, left eyebrow height and left eyebrow width. Each of the eleven nodes is repeated in every time slice.
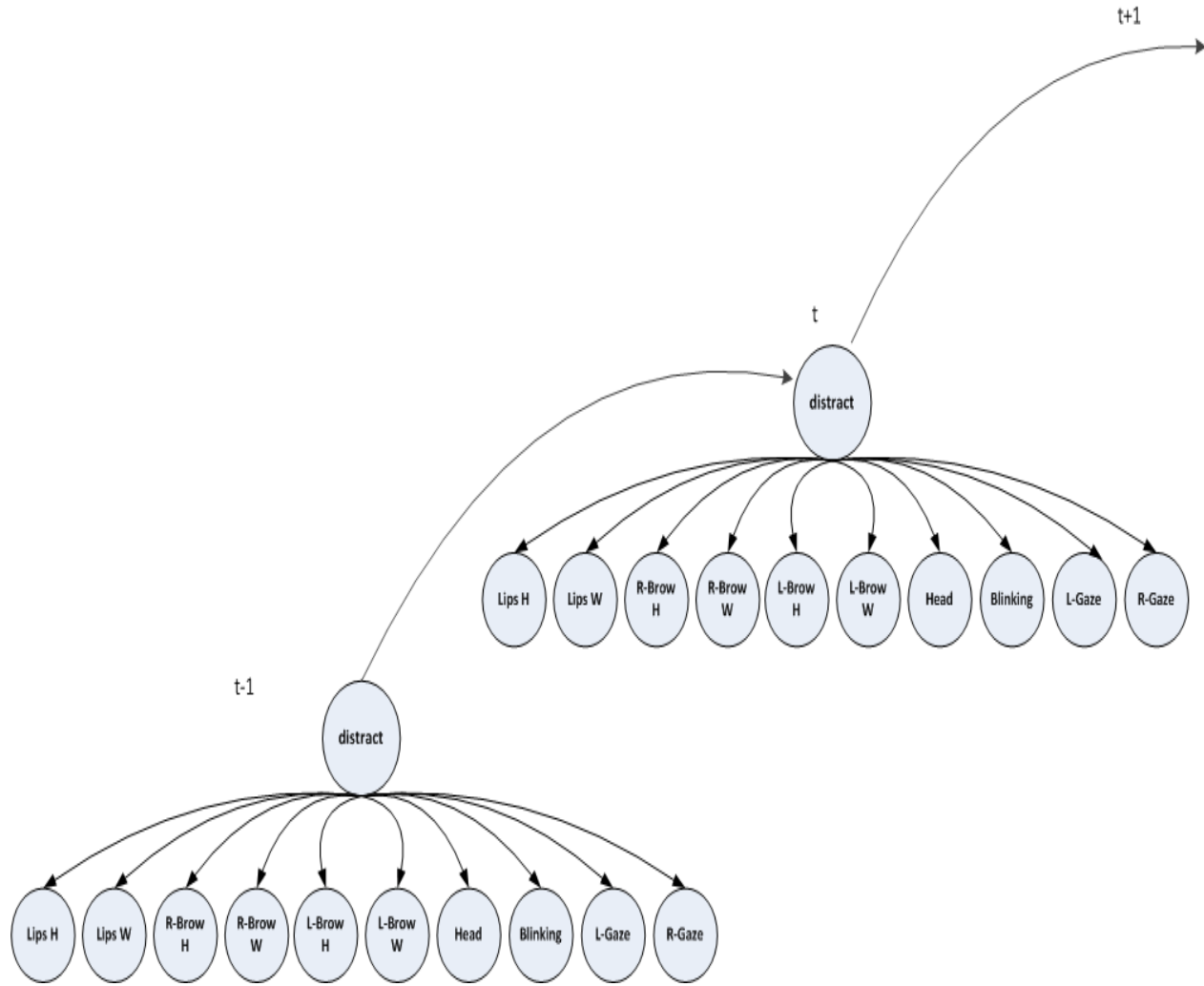
Figure 4.14: DBN for final experiment

Model performance was compared based on important performance measurements: model accuracy. Accuracy was the summarization of the parameters performance used in training and testing. Results for this model performance are discussed in chapter 6. For DBNs, the window size is the time slices for the model that is evidence at time slice t, represented the summarized measures across a window at ending time slice t; evidence at time slice $t+1$ was summarized across the next window, which is started at t and ended at $t+$ window size. Longer window size can improve model performance [24][36]. If a total quantity of training data was fixed, the number of training instances decreased with each increasing window size. Sequence length in DBN was also considered. A sequence length contained multiple and consecutive training instances. With longer sequences length, it might provide more comprehensive and stable time

dependent relationships between driver distraction and the performance. Longer sequence length will also give a better model performance. In this study, the sequence length was 60s. This is because data captured from faceLAB and faceAPI were captured with 60 frames of data at every one second. Thus, it will be easier and better if the sequence length is made to 60s.

4.5   Inference Engine for BN

Inference for Bayesian Networks has already existed. Inference is the act of reasoning from factual knowledge or evidence. In DBN, given a sequence of observations, a full BN can be represented by replicating slices of the network until the network is large enough to be observed and learned. This technique is called unrolling technique. Slices are added after the last observation. Once the DBN is unrolled, any inference algorithms like variable elimination algorithm, joint tree algorithm, particle filtering algorithm can so on be implemented. The easiest technique for inference in DBN is Naïve Approach. However with this approach, the unrolled DBN for the desired number of time slices is treated as a static BN. Naïve Approach has these behaviours that make it impossible to be applied for DBN:-

a) Smoothing- to query from a previous time slice

b) Filtering- to query at the current time

c) Prediction- to query at a future time

Those queries are essential for Dynamic network. Even Naïve Approach is simple and is "already" implemented but it typically cannot tell beforehand the number of time slices which is "desired" and with many time slices the unrolled DBN can become very huge and will make the static inference algorithms run out of memory and will take too long to be processed. Variable-elimination algorithm and junction-tree algorithm are among the best algorithms and mostly used. For this study, an exact inference; junction-tree algorithm is used. Junction-tree for DBN is variety. The variations include the particulars of secondary structure used, message passing scheme and the message format itself. This algorithm also uses a static junction-tree algorithm as a subroutine and this version of algorithm has been developed for every dynamic problem and it can solve the smoothing, filtering and prediction problem which have been mentioned before in Naïve Approach. The outline for this junction-tree algorithm is as below:-

Initialization

– On initialization, we create two junction trees, J1 and Jt

  • J1 is the junction tree for the initial time slice and is created from time slice 1 of the 2TBN

  • Jt is the junction tree for each subsequent time slice and is created from time slice 2 of the 2TBN and the outgoing interface of time slice 1

– Time is initialized to 0

Queries

– Marginals of nodes at the current time slice can be queried

  • If current time = 0, queries are performed on "_1" nodes in J1

  • If current time > 0, queries are performed on "_2" nodes in Jt

Evidence Application

– Evidence can be applied to any node in the current Time slice

  • If current time = 0, evidence is applied to "_1" nodes in J1

  • If current time > 0, evidence is applied to "_2" nodes in Jt

Advance

– Increment time counter

– Use outgoing interface from active time slice to do inference in next time slice

Since the outgoing interface d-separates the past from the future, this ensures that when we do inference in the next time slice we are taking everything that has occurred "so far" into account.

Algorithm 4.3: Junction tree outline

4.6　Conclusions

From the three types of existing measurement for driver cognitive distraction detection, physiological measurement was used in this study. This type of detection was used mainly because it would not undermine the driving activity.

Bayesian Network was used in this study since it is one of the most appropriate algorithms to be constructed for real time activity. This chapter basically describes the nature of Bayesian Network for both Static and Dynamic and their algorithms. With many types of inference engines, learning parameters and functions provided in the BNT toolbox, to develop a Bayesian Network is not really a simple thing. Advantages and disadvantages of Bayesian Network can be summarised as below:

The advantages of Bayesian Network [50] and [60]:

a) Bayesian Network can visually represent all the relationships between the variables in the system with connecting arcs.

b) Bayesian Network can easily recognize the dependence and independence between various nodes.

c) Bayesian Network can handle situations where the data set is incomplete since the model accounts for dependencies between all variables. Sometime due to the system malfunction where it is not feasible or practical to measure all variables due to system constraints (costs, not enough sensors, contextual information etc.) the data captures is missing, thus BN handles this situation by referring to the prior and posterior knowledge of the network to solve this missing data problem.

d) Bayesian Network can help to model noisy systems.

e) Bayesian Network can be used for any system model - from all known parameters to no known parameters.

The limitations of Bayesian Network:

a) All branches must be calculated in order to calculate the probability of any one branch.

b) The quality of the results of the network depends on the quality of the prior beliefs or model. A variable is only a part of a Bayesian Network if one believes that the system depends on it.

Chapter 5

Correlation and Regression between Features

This chapter presents results from correlation and regression on features used in this study. In an initial experiment, the relationship between mouth movement and eye movement was studied. Later experiments captured lips and eyebrows and their relationship was also studied. Both experiments used correlation and regression statistical modelling for relationship analysis. Correlation and regression on those features showed a strong and positive relationship between the features and cognitive distraction.

5.1 Overview

Correlation and regression are similar but different. They are likely related to each other. Basically, correlation is a method of analysing relationship or strength between two or more variables where both of them are random variables, whereas, regression is used to describe a linear or non-linear relationship where independent/factor/scalar variables are used to predict or estimate the dependent/response/explanatory variables.

A regression model is an application from linear model. Linear model is a standard methodology to make inferences on one or more quantitative variables. In regression model, the values for response variables are identified with numeric values of one or more factor variables. Making an inference on differences among the means of populations is not a primary purpose of a statistical analysis of a regression model. However it is rather to make inferences about the relationship of the mean of the response variable towards the factor variables. The inferences are made through the parameter of the model: intercept value and the slope value.

These values can be used to predict or explain the behaviour of the response variable. In this study, for example, it can be used to explain about lips movement when eyebrows move.

As been mentioned, examining the strength of the linear relationship between two or more random variables should be done with correlation model. Unlike a regression model, correlation model specifies the joint distribution of both variables and not just the conditional distribution between response variable and fixed factor variable.

The correlation model that is normally used is the normal correlation model. This model specifies that two variables, response and factor have what is known as bivariate normal distribution. This distribution is defined by five parameters: the means of response variable and factor variable, the variances of response variable and factor variable, and the correlation coefficient, r. This r value measures the strength of a linear relationship between the two variables [62].

All data are collected using faceAPI and faceLAB and the data have been through a few steps in preparing them for an analysis. Data collected were grouped and annotated as control (without distraction/task) and distracted (with distraction/task). Using faceLAB, annotation button has been used to annotate these two tasks: 1 as distracted and 0 as control. Data from control tasks are data collected when the drivers are free from asking questions and doing any activity in the car. Thus, no such distraction is created to indicate the driver as distracted. Even the driver might be distracted cognitively by the situation and surrounding during this control experiment, it should be deemed as minimal distraction as compared to when the driver is distracted with the questions and tasks. During the experiment run, the safety of the drivers and people surrounding was also took into account. The distractions created during the experiment, were basically a normal type of distractions. Once data has been collected, the data has been through for data cleansing. Any missing data or inaccurate and incomplete data will be transcribed to make it accessible and possible for further analysis. Consistencies among data between one driver to another is important. Thus, number of control data and distracted data from each driver is equal. Half of the total number of data was for control data and another was for distracted data.

There are several methods available for correlation and regression. Different methods are used for different types of data or purposes.

5.2 Background

This section provides background of the linear regression model, scatter diagram and normal correlation model. They are used to examine a relationship between variables.

5.2.1 Linear Regression

There are several types of regression model. However, the simplest regression model is the linear regression model or sometimes called simple linear regression model. This regression model is used to check on the relationship between two variables. One is response variable and another one is factor variable. The formula for linear regression model is written as [63]:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

$\beta_0 + \beta_1 x$, is the deterministic portion of the model and ε is the error variable. This portion specifies that for any value of the factor variable, x, the population mean of the response variable, y, can be described with a straight line function, $\beta_0 + \beta_1 x$. This equation is the general equation for linear regression and can be presented as in Figure 5.1 below:
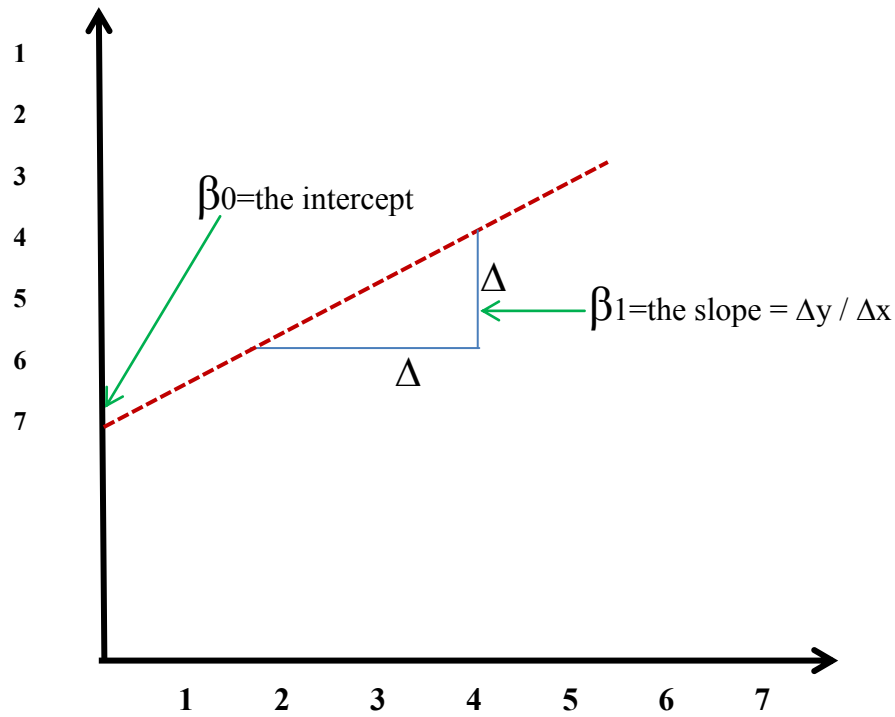
Figure 5.1: Linear regression equation

Usual notation for the general expression of a straight line, the parameter β_0 as the intercept is the value of the mean of the response variable when x is zero. The parameter β_1, the slope is the change in the mean of the response variable associated with a unit change in x. These parameters are called as regression coefficients.

After performing an analysis, linear regression statistics can be used for prediction. It can be used to predict the response variable when the factor variable is known. For instance, movement on mouth can be predicted once movement on eye is known. Regression is a statistical modelling which goes beyond the correlation. It contains prediction capabilities. Regression is used every day on an intuitive level. For example, in normal life, a person driving a big luxury car is thought to be financially successful and wealthy. Quantitative regression can be precise and useful for predictive purposes if a mathematical function is developed upon it. A regression can be applied in this example to find the best formula that fits the relationship. The formula can be used in the future to predict values for the response variable when only the factor variable is known. In this study, linear regression has been used to predict cognitive distraction on a driver when lips and eyebrows move.

Linear regression can provide a regression line known as the least squares line. This line is a plot of the expected value of the response variable for all values of the factor variable and the line can minimize the squared residuals. Regression line is the one that best fits the data. It can be represented on a scatter diagram. The scatter diagram can be plotted to find whether the relationship between two variables are with positive or negative relationship. Positive relationship indicates that the two variables are correlated to each other and negative relationship indicates that the two variables are strongly not related to each other. Since linear regression model is not usually the best predictor, there is also an error term in the equation, ε. Usually, t-statistic is used to determine the significance of the slope of the linear regression. It can be tested against a t-distribution to determine how probably true it is of the coefficient. The t-statistic for the significance of the slope is an essential test. It is used to determine if the regression model is usable or not. If the slope is significantly different than zero, then the regression model is usable to predict the response variable for any value of the factor variable. If the slope is zero, thus, it has no prediction ability because for every value of the factor variable, the prediction for the response variable would be the same. Nevertheless, knowing the value of the factor variable

would not improve the ability to predict the response variable. Therefore, if the slope is significantly zero, the model is useless for predictions.

In [62], formula for linear regression (Y') model for variables X and Y is given as:

$$Y' = r\left(\frac{SD_y}{SD_x}\right)\left(x - \bar{x}\right) + \bar{y}$$

r = The correlation coefficient between X and Y

SD_y = Standard Deviation on the Y variable

SD_x = Standard Deviation on the X variable

x = A raw score predicting Y

$\bar{x}$ = The mean of X variable

$\bar{y}$ = The mean of Y variable

Linear regression model gives correlation coefficient. Correlation coefficient will be explained in details in subsection of correlation. Since linear regression and correlation are similar in concept, the correlation coefficient for both is similar in definition. The relationship between the two variables, response and factor is determined by the correlation coefficients. The larger the correlation coefficient is, the stronger the relationship between the two variables is. Therefore, if the relationship between those two variables is stronger, the more accurate the predictions are likely to be made.

If the correlation coefficient is large, the standard error of estimate will be small. It means, while the prediction may not be perfect, it is likely to be very close. On the other hand, if the correlation is small, the standard error of estimate will be large and it means that there is a greater error in the prediction. Therefore, as long as the correlation coefficients are large, the predictions are ensured to be pretty accurate.

Usually, r-squared or coefficient of determination is also used in linear regression model. The r-squared is the square of the correlation coefficient value. Its value may vary from 0 to 1, $0 < r^2 < 1$. The r-squared has the advantage over the correlation coefficient in that it may be interpreted directly as the proportion of the variance in the response variable that can be accounted for by the regression equation. For instance, if the correlation coefficient between cognitive distraction and features is 0.72, the r-squared or coefficient of determination value is

0.5184 and it means 52% of the variance in the response variable can be explained with the regression equation and the other 48% is unexplained.

The r-squared is very useful because it gives the proportion of the variance of one variable that is predictable from the other variable. It determines how certain one variable can be in making predictions from a certain model. The r-squared is a measure of how well the regression line represents the data. If the regression line passes exactly through every point on the scatter diagram, it would be able to explain all of the variation. The further the line away from the points, the less it is able to be explained.

The linear regression in this study was built using Java programming. The pseudo-code for the linear regression is as in algorithm 5.1 below. Sxx, Syy and Sxy are the sum of squares of the variables x and y. A is the intercept and B is the slope.

```
''*************************************************

'* n   is the Number of X and Y pairs
'* X() Array of X values
'* Y() Array of Y values
'*
'* Sxx, Syy and Sxy are the Sum of Squares
'* i  Loop Counter
'*
'* A is the Intercept
'* B is the Slope

'*************************************************
Calculation of Mean

For i=0 to n-1
X_Bar=X_Bar+X(i)
Y_Bar=Y_Bar+Y(i)
Next i
X_Bar=X_Bar/n
Y_Bar=Y_Bar/n
Sum of Squares
For i=0 to n-1
     Sxx=Sxx + (X(i)-X_Bar)^2
     Sxy=Sxy + (X(i)-X_Bar)*(Y(i)-Y_Bar)
Next i
Slope and Intercept
B=Sxy/Sxx
A=Y_Bar – B * X_Bar
```

Algorithm 5.1: Linear Regression Pseudo code [62].

5.2.2   Scatter Diagram

Scatter diagram is used for the relationship between two variables analysis. One variable, usually the factor variable is plotted on the horizontal axis and the other variable, response variable is plotted on the vertical axis. The relationship patterns can be shown graphically from the pattern of their intersecting points on the diagram itself. Regularly, a scatter diagram is used to prove the cause-and-effect relationship between variables. Even if the diagram shows a relationship between variables, it does not prove anything about one variable's causes or effects of the other. Rather than to examine theories about cause-and-effect relationship, a scatter diagram also uses to root causes of an identified problem.

Once the scatter diagram is drawn, a fit line can be made in such a way that the line represents a reasonable approximation to the points in the diagram. By viewing the best fit line and the nature of the scatter diagram, it can tell more about whether the variables are related or not as in Figure 5.2 [64]:

a) A direct (positive) linear relationship between variables. The best fit line is linear and has a positive slope. Both x and y variables increase together.

b) An inverse (negative) linear relationship between the variables. The best fit line is linear and has a negative slope. Y variable decreases as x variable increases.

c) A direct curvilinear relationship. Sometimes called complex correlation, is when the relationship can either be positive or negative. The value of y seems to be related to variable x, but the relationship is not easily determined.

d) No correlation occurs when there is no demonstrated connection between the two variables.
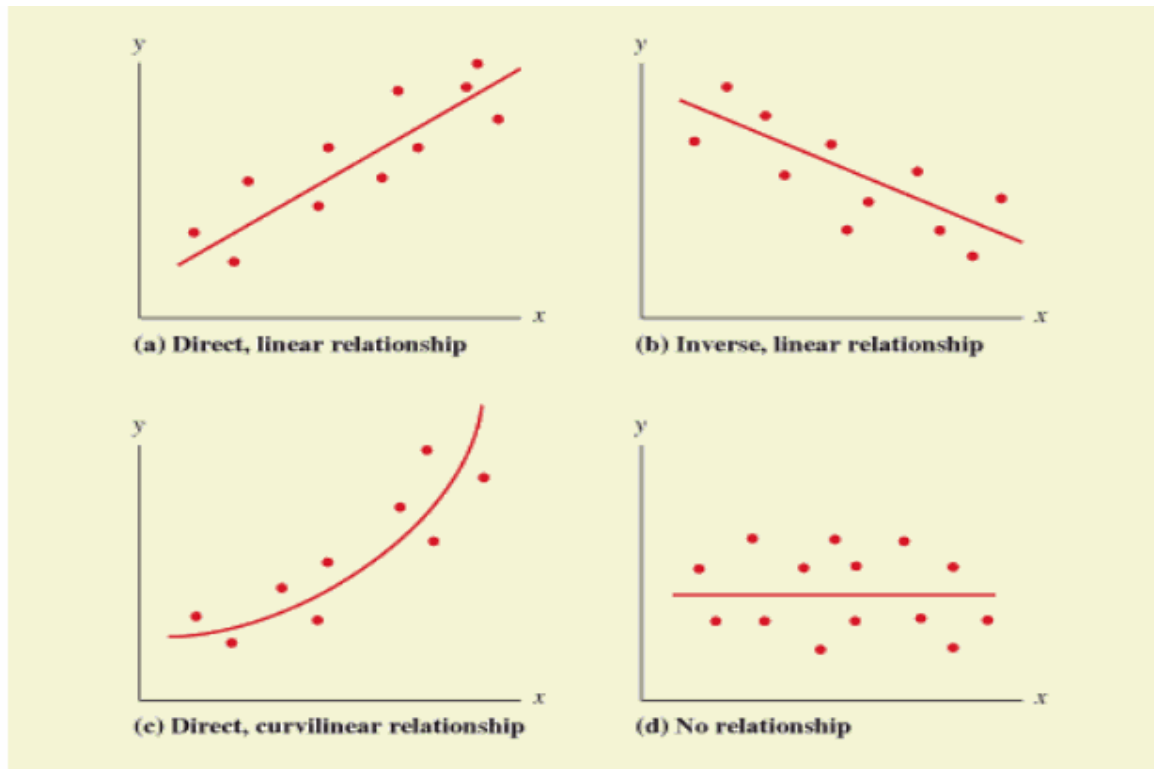
Figure 5.2: Scatter diagram definition [64].

Basically, if the dots on the scatter diagram tend to go from the lower left to the upper right it means that as one variable goes up the other variable tends to go up as well. This is called a "positive relationship".

On the other hand, if the dots on the scatter diagram tend to go from the upper left corner to the lower right corner of the scatter diagram, it means that as values on one variable go up values on the other variable go down and this is called a "negative relationship".

Scatter diagram has also been used in this study for correlation purposes. The first experiment which captured mouth movement and eye movement and the second experiment which captured lips and eyebrows used scatter diagram to determine the features' relationship with the cognitive distraction.

5.2.3 Correlation Coefficient

Correlation refers to the process of establishing the existence of a relationship between two variables. In subsections 5.2.1 and 5.2.2, relationship between two variables can also be determined with scatter diagram and linear regression. Using a scatter diagram is only appropriate to get a general idea about the relationship between the two variables. However this method has problems like:

a) It can be tedious and time consuming to plot.

b) It does not exactly tell the strength of the relationship.

c) It is difficult to make a prediction about the variables solely based on looking at the scatter diagram.

Thus, correlation coefficient is also used in this study to give a better idea about the relationship strength between variables.

The most popular correlation coefficient analyses are: Spearman's correlation coefficient and Pearson's correlation coefficient. Spearman's technique is used when calculating a correlation coefficient for ordinal data. Pearson's technique is used when an interval or ratio-type data is used. This study used Pearson's technique.

The mathematical formula for correlation coefficient is given as [65]:

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\left(\sum x^2\right) - \left(\sum x\right)^2}\ \sqrt{n\left(\sum y^2\right) - \left(\sum y\right)^2}}$$

r = the correlation coefficient

n = the sample size

x and y are the variables

Where, basically, the value of r-value varies from -1 to +1. A minus one indicates a perfect negative correlation, while a plus one indicates a perfect positive correlation. A correlation of zero means there is no relationship between the two variables.

In [63] and [65] the correlation coefficient, r, has the following properties:

a) Its value is between +1 and -1 inclusive. A positive correlation coefficient implies a direct relationship while a negative coefficient implies an inverse relationship.

b) Values of +1 and -1 signify an exact direct and inverse relationship. A plot of the values of x and y can be described exactly on a straight line with a positive or negative slope.

c) A correlation of zero indicates there is no linear relationship between variables. However, this condition does not necessarily imply that there is no relationship, because correlation only measures the strength of a straight line relationship. Knowing the correlation of the data is zero is showing that there is no linear relationship between the data. This might probably because the data is having non-linear relationship like quadratic relationship or exponential relationship. When two variables are with correlation coefficient of zero, it also indicates that the data are independent.

d) The correlation coefficient is symmetric with respect to the two variables. It is thus a measure of the strength of a linear relationship between any two variables, even if one is a factor variable in regression setting.

The sign of *r*-value provides the information about the direction of the relationship between those two variables. The magnitude of the *r*-values showed the strength of the relationship between these two variables as in Figure 5.3 [66]:
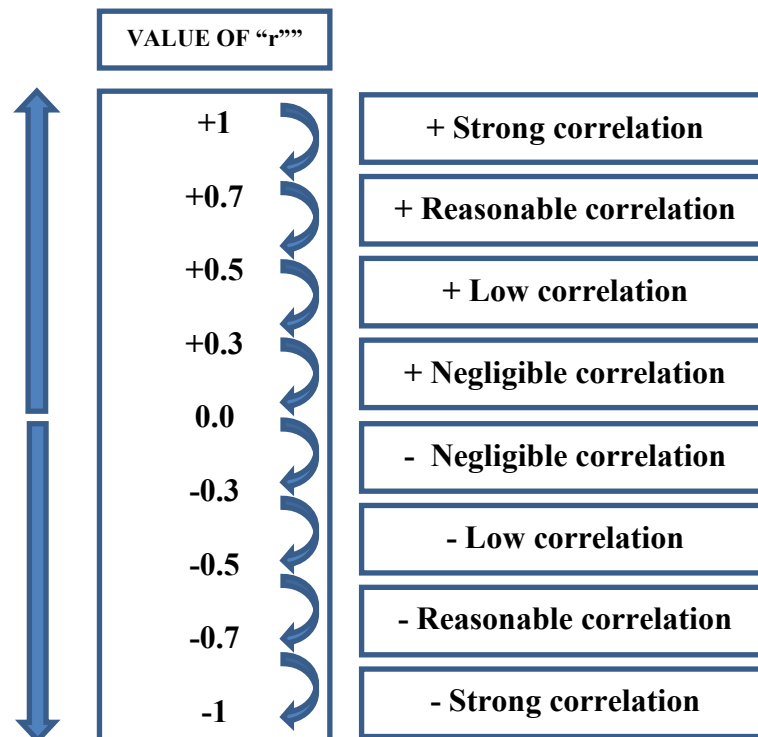
Figure 5.3: Correlation coefficient magnitude and direction

5.3   Proposed features: Mouth movement and Eye movement

Many studies in driver cognitive distraction detection, have investigated the patterns of eye movements. Eye movements can give much information based on fixation, saccade, eyelid movement (height and width), blinking duration, blinking frequency, pupil diameter, gaze direction and rotation etc.   In [45] they used blinking, saccade, eyelid movement and pupil diameter, whereas in [24][36] they used the characteristics of fixations, saccades and smooth pursuits to recognize the patterns of eye movements. In this study, eye movement was also tracked for driver cognitive distraction detection. Information like eye's height and eye's width were captured by faceLAB cameras from both the left and right eye. Information from eye's height and width were used to determine their relationship with the mouth movement. Other information from eyes like blinking frequency, pupil diameter and gaze rotation will be used later for accuracy of distraction detection.

Other than eye movement, in psychological studies, it has been proven that mouth movement is also a good indicator of a human's state of mind. Mouth movement can also be thought of as a form of body language. Body language can be used to obtain information about whether a person is distracted or not. Study by [38] has monitored the relationship between mouth movement and driver fatigue or distraction using a camera. Normally the mouth is hardly open when the driver is alert. The maximum width (W_{max}) and maximum height (H_{max}) can indicate different levels of distraction. The height (H_m) between the top lip and the bottom lip varies greatly when one is talking, yawning or even thinking. Thus, the mouth movement can be represented in the feature vector as $Z = (W_{max}, H_{max}, H_m)$. Thus, in this study, mouth height and mouth width were tracked.

Since mouth and eye movements are useful for distraction detection, their relationship to each other was examined here.

Data from eye and mouth movements were captured from the FeatureSetsByCamera class from faceLAB as in Figure 5.4. FeatureSetsByCamera contains data which is related to the position of facial features in each camera image. Mouth and eye movements are under FeatureSet class.

Their height and width information are contained in left eye rectangle, right eye rectangle and mouth rectangle. FeatureSet contains four FeatureRect (feature rectangle) objects [67]:

a) The rectangular region surrounding the entire face, in normalized image coordinates

b) The rectangular region surrounding the right eye, in normalized image coordinates

c) The rectangular region surrounding the left eye, in normalized image coordinates

d) The rectangular region surrounding the mouth, in normalized image coordinates
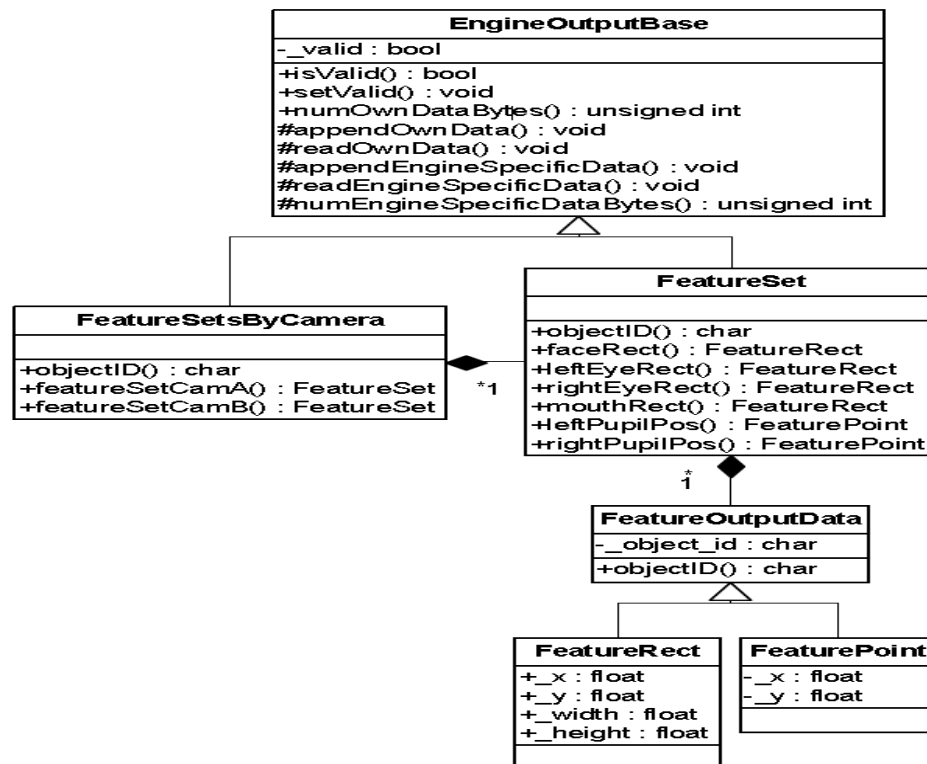
Figure 5.4: FeatureSetByCamera class

Normalized image coordinates from the FeatureSet above are a type of camera matrix used in faceLAB cameras. It assumes that the focal length is equal to 1 and the image coordinates are measured in a coordinate system where the origin is located at the intersection between axis and the image plane. It has the same units as the 3D coordinate system. The resulting image coordinates are the one referred as normalized image coordinates.

This initial experiment was run in a lab and the features were tracked and captured by faceLAB system. Data captured was annotated and grouped into distracted or not distracted. By using

faceLAB a special button was created as a function to annotate and indicate when the driver is distracted with questions or tasks. This annotation button was pressed manually by the experimenter every time when the driver was distracted. This annotation button will produce an extra column in the data file with values 1 or 0. As mentioned before, 1 was annotated as distracted (with task) and 0 annotated for control (without task). Details about experimental setup are explained in experiment setup chapter. Mouth and eye movements used for this relationship determination were from distracted group. Six participants had volunteered in this experiment.

5.3.1   Experimental Results

As mentioned before, the relationship between mouth movement and eye movement when a driver is cognitively distracted was determined by using two different methods: (a) Correlation Coefficient model, Pearson-r and (b) Scatter diagram. Pearson-r, sometimes also called bivariate correlation coefficient was chosen because it is suitable for the continuous data type from mouth movement and eye movement. Pearson correlation is defined as the covariance of the two variables divided by the product of their standard deviations.

Surprisingly, results received from Pearson-r in this experiment either fall in a reasonable correlation range or a strong correlation range. Mouth movement and eye movement data are related to the position of facial features in each camera image. Thus, the position will show an appropriate movement on both eye and mouth whenever the driver is distracted cognitively.

Movements from eye and mouth were captured based on their width and height information: width:float and height:float as can be found from FeatureSetByCamera class. + and − signs refer to the direction of the movement. leftEyeRect() : FeatureRect is left eye captured from Feature Rectangle, rightEyeRect() is right eye rectangle and mouthRect() is mouth rectangle.

Correlations

		REYERECT_CA_H	LEYERECT_CA_H	REYERECT_CB_H	LEYERECT_CB_H	MOUTRECT_CA_H	MOUTRECT_CB_H
REYERECT_CA_H	Pearson Correlation	1.000	.704**	.955**	.873**	.923**	.937**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	183292.000	183292	183292	183292	183292	183292
LEYERECT_CA_H	Pearson Correlation	.704**	1.000	.518**	.924**	.776**	.628**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	183292	183292.000	183292	183292	183292	183292
REYERECT_CB_H	Pearson Correlation	.955**	.518**	1.000	.749**	.868**	.940**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	183292	183292	183292.000	183292	183292	183292
LEYERECT_CB_H	Pearson Correlation	.873**	.924**	.749**	1.000	.899**	.829**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	183292	183292	183292	183292.000	183292	183292
MOUTRECT_CA_H	Pearson Correlation	.923**	.776**	.868**	.899**	1.000	.955**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	183292	183292	183292	183292	183292.000	183292
MOUTRECT_CB_H	Pearson Correlation	.937**	.628**	.940**	.829**	.955**	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	183292	183292	183292	183292	183292	183292.000

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5.1: Left and Right Eye's Height and Mouth's Height correlations

Table 5.1 shows a relationship between the variables right eye's height and left eye's height with the mouth's height. Sig. (2-tailed) is the p-value associated with the correlation. The footnote under the correlation table explains that the correlation is significant at the 0.01 level (2-tailed). From the table:

a)   REYERECT is Right Eye Rectangle

b)   LEYERECT is Left Eye Rectangle

c)   MOUTRECT is Mouth Rectangle

d)   CA is Camera A

e)   CB is Camera B

f)   H is Height

g)   Sig.(2-tailed) is Significance value with 2 tailed

h)   N is number of data

It clearly showed that both the variables have a strong or good relationship. Right eye's height magnitude ranged from 0.868 (the lowest) to 0.940 (the highest). However, the left eye's height magnitude ranged from 0.628 (the lowest) to 0.899 (the highest). The *r*-values in Table 5.1 mean that whenever the eye's height changes or eye moves, the mouth's height also relatively changes and mouth moves. Another interesting conclusion that can be made from the result is that the right eye's height is more related by 53.94% to the mouth's height as compared to the left eye's height with 46.05%. This can be explained by the 'Eyes Direction Language' in Neurolinguistic Programming (NLP) study [68]. In that research, many psychologists believe that when a person eye's direction is to the left or right, it indicates that they are thinking about something. It is important to recognize the direction of someone's eyes when they are thinking. Looking to the left usually indicates that a person is reminiscing or trying to remember something. On the other hand, looking to the right indicates that the person is doing more creative thoughts, and this is often interpreted as a potential sign that the person may appear to be deceitful in some situations. For instance, the person might be creating a version of events or might be trying to find a solution to a problem. Often, a person looks to the right more than to the left when thinking because the person is usually involved with visually and auditory remembered images and internal dialogues of "talking to oneself".

The Sig. (2-tailed) from both table 5.1 and 5.2 has indicates that there is significance between mouth rectangle and eye rectangle. The relationships between those features are positive. For example relationship between mouth's heights from camera A (MOUTRECT_CA_H) with right eye's height from camera A (REYERECT_CA_H) is 92.3% (0.923) which means that as one variable goes up or down, the other variable will follow.

Empty grid in the table at Sig. (2-tailed) which is with a diagonal line shows the same variable relationship. For instance, the grid between REYERECT_CA_H (at rows) and REYERECT_CA_H (at columns) shows the same variable relationship and no significance value is able to compute.

Correlations

		REYERECT_CA_W	LEYERECT_CA_W	REYERECT_CB_W	LEYERECT_CB_W	MOUTRECT_CA_W	MOUTRECT_CB_W
REYERECT_CA_W	Pearson Correlation	1.000	.710**	.962**	.880**	.946**	.958**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	183292.000	183292	183292	183292	183292	183292
LEYERECT_CA_W	Pearson Correlation	.710**	1.000	.521**	.931**	.788**	.635**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	183292	183292.000	183292	183292	183292	183292
REYERECT_CB_W	Pearson Correlation	.962**	.521**	1.000	.754**	.887**	.957**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	183292	183292	183292.000	183292	183292	183292
LEYERECT_CB_W	Pearson Correlation	.880**	.931**	.754**	1.000	.916**	.843**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	183292	183292	183292	183292.000	183292	183292
MOUTRECT_CA_W	Pearson Correlation	.946**	.788**	.887**	.916**	1.000	.963**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	183292	183292	183292	183292	183292.000	183292
MOUTRECT_CB_W	Pearson Correlation	.958**	.635**	.957**	.843**	.963**	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	183292	183292	183292	183292	183292	183292.000

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5.2: Left and Right Eye's Width and Mouth's Width correlations

Table 5.2 shows a correlation between variables right eye's width and left eye's width with respect to the mouth's width. All magnitudes for *r*-value from the eye's width are also in a group of 0.7 or more which indicates a strong relationship. Right eye's widths ranged from 0.887 to 0.958 and Left eye's widths ranged from 0.635 and 0.916. Similar to the right eye's height, right eye's width is also highly correlated to mouth's width. Right eye is 54.08% correlated to the mouth's width and left eye is only 45.92%. The right eye's width being bigger than the left eye's width is also due to the similar reasons explained when discussing about eye's height and mouth's height from Table 5.1. The above result justifies that eye and mouth is relatively correlated to each other, whenever a driver is cognitively distracted. Since the normal expression of when a driver is not cognitively distracted, the possible correlation coefficient between eye and mouth movement is close to 0, its relationship is excluded from the experiment.

Relationship between mouth movement and eye movement was also illustrated by drawing a scatter diagram. The plots showed a positive relationship between eyes movement and mouth movement. Both left eye and right eye gave a direct positive relationship with mouth movement. Scatter diagram in Figure 5.5, the x-axis refers to Right Eye Rectangle captured from Camera A for Width and y-axis is Mouth Rectangle captured from Camera A for Width. Mouth and eye movement's values captured in this experiment use pixel images and their values are normalized values ranged from 0 to 1. The line drawn on the scatter diagram showed a positive linear relationship between mouth width and right eye's width.

Figure 5.5: Scatter diagram right eye and mouth (width & camera A)

Scatter diagram in Figure 5.6, the x-axis refers to Right Eye Rectangle captured from Camera A for Height and y-axis is Mouth Rectangle captured from Camera A for Height. The line drawn on the scatter diagram showed a positive linear relationship between mouth height and right eye's height.
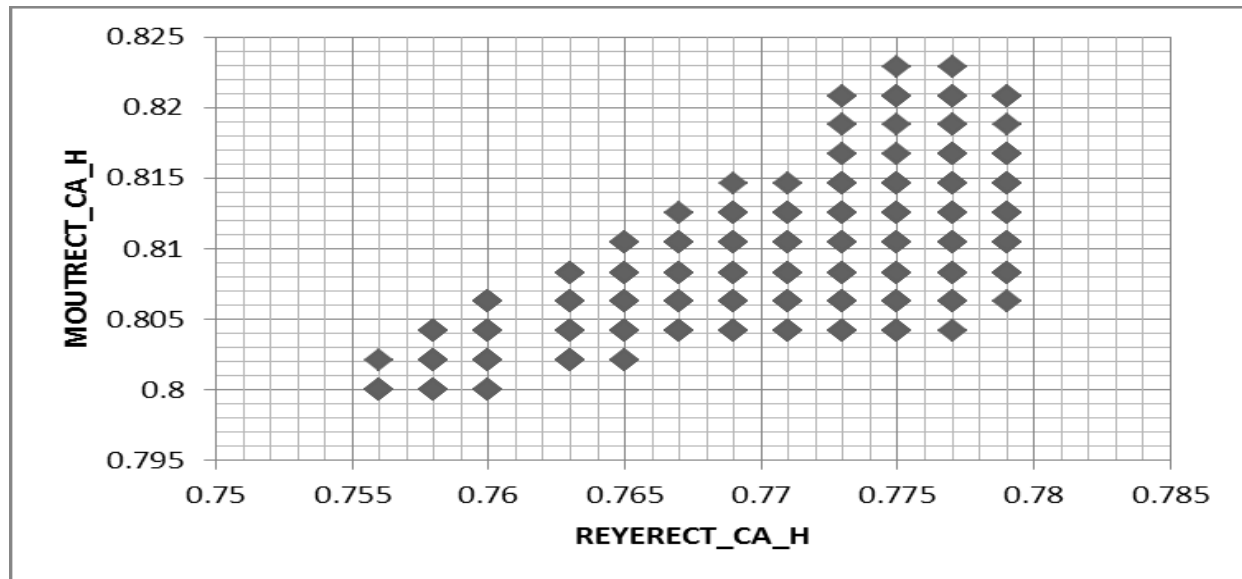
Figure 5.6: Scatter diagram right eye and mouth (height & camera A)

Scatter diagram in Figure 5.7, the x-axis refers to Left Eye Rectangle captured from Camera A for Width and y-axis is Mouth Rectangle captured from Camera A for Width. The line drawn on the scatter diagram showed a positive linear relationship between mouth width and left eye's width.
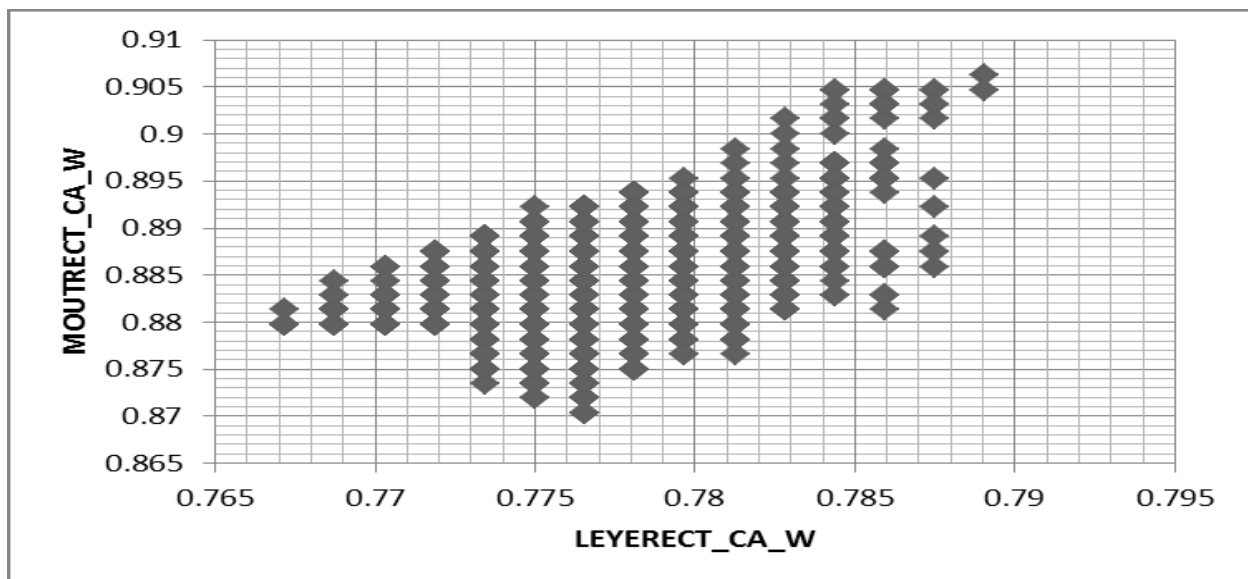
Figure 5.7: Scatter diagram left eye and mouth (width & camera A)

Scatter diagram in Figure 5.8, the x-axis refers to Left Eye Rectangle captured from Camera A for Height and y-axis is Mouth Rectangle captured from Camera A for Height. The line drawn on the scatter diagram showed a positive linear relationship between mouth height and left eye's height.
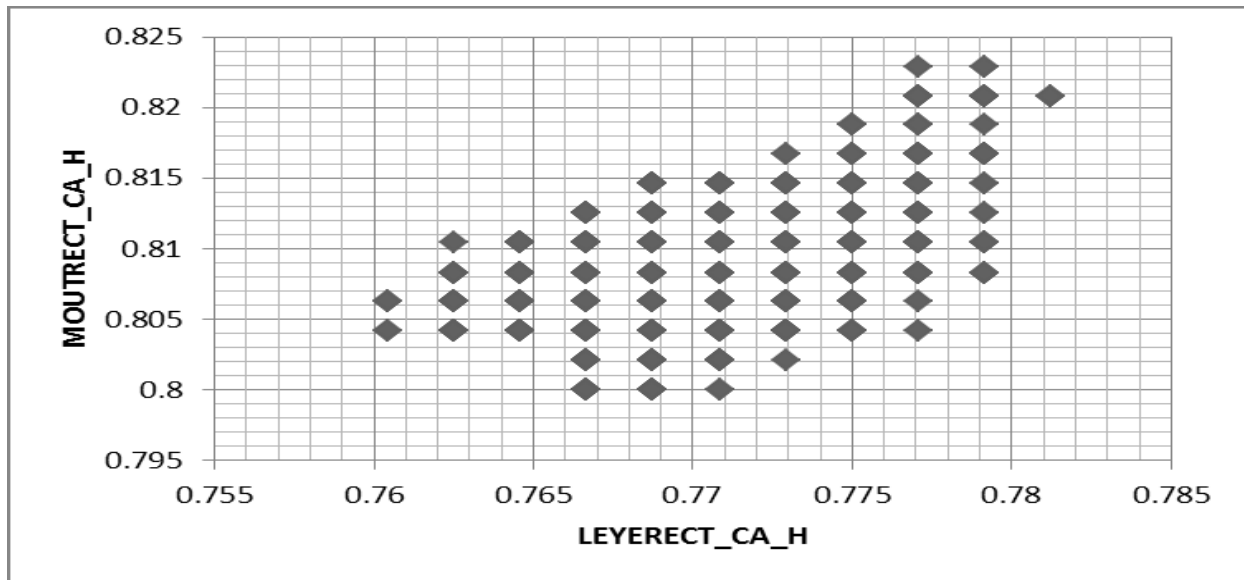
Figure 5.8: Scatter diagram left eye and mouth (height & camera A)

Scatter diagram in Figure 5.9, the x-axis refers to Right Eye Rectangle captured from Camera B for Width and y-axis is Mouth Rectangle captured from Camera B for Width. The line drawn on the scatter diagram showed a positive linear relationship between mouth width and right eye's width.
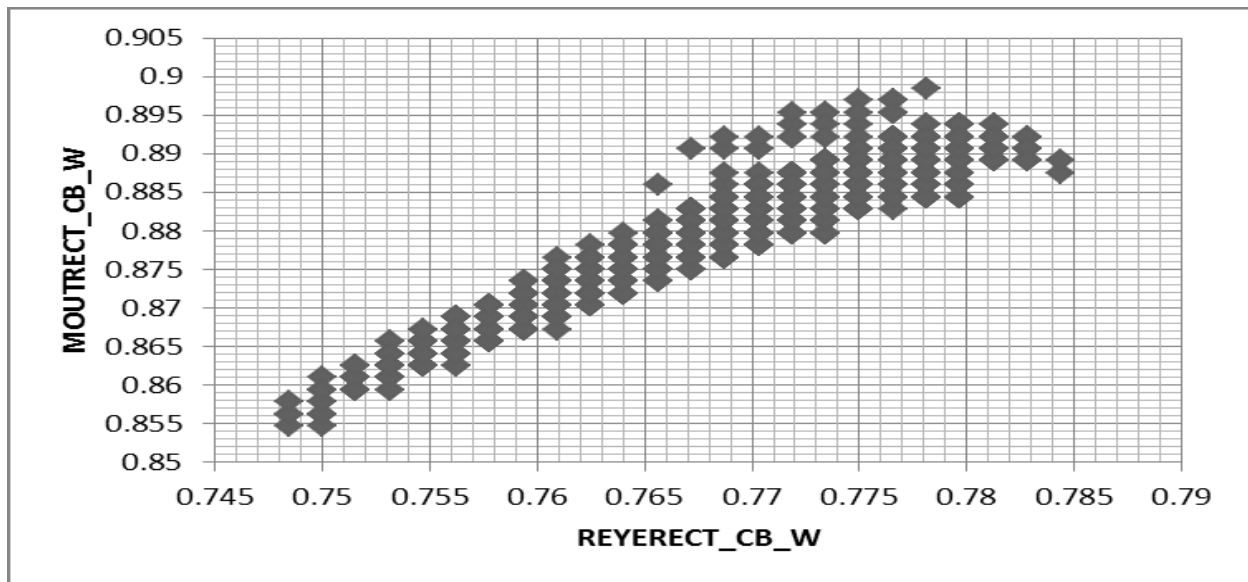
Figure 5.9: Scatter diagram right eye and mouth (width & camera B)

Scatter diagram in Figure 5.10, the x-axis refers to Right Eye Rectangle captured from Camera B for Height and y-axis is Mouth Rectangle captured from Camera B for Height. The line drawn on the scatter diagram showed a positive linear relationship between mouth height and right eye's height.
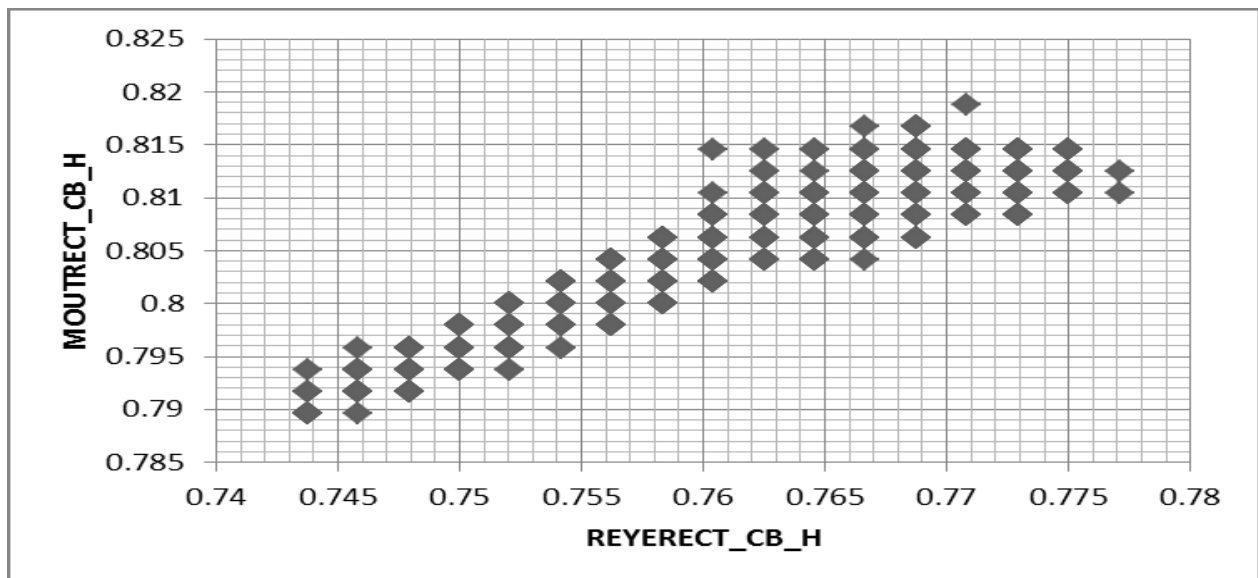
Figure 5.10: Scatter diagram right eye and mouth (height & camera B)

Scatter diagram in Figure 5.11, the x-axis refers to Left Eye Rectangle captured from Camera B for Width and y-axis is Mouth Rectangle captured from Camera B for Width. The line drawn on the scatter diagram showed a positive linear relationship between mouth width and left eye's width.
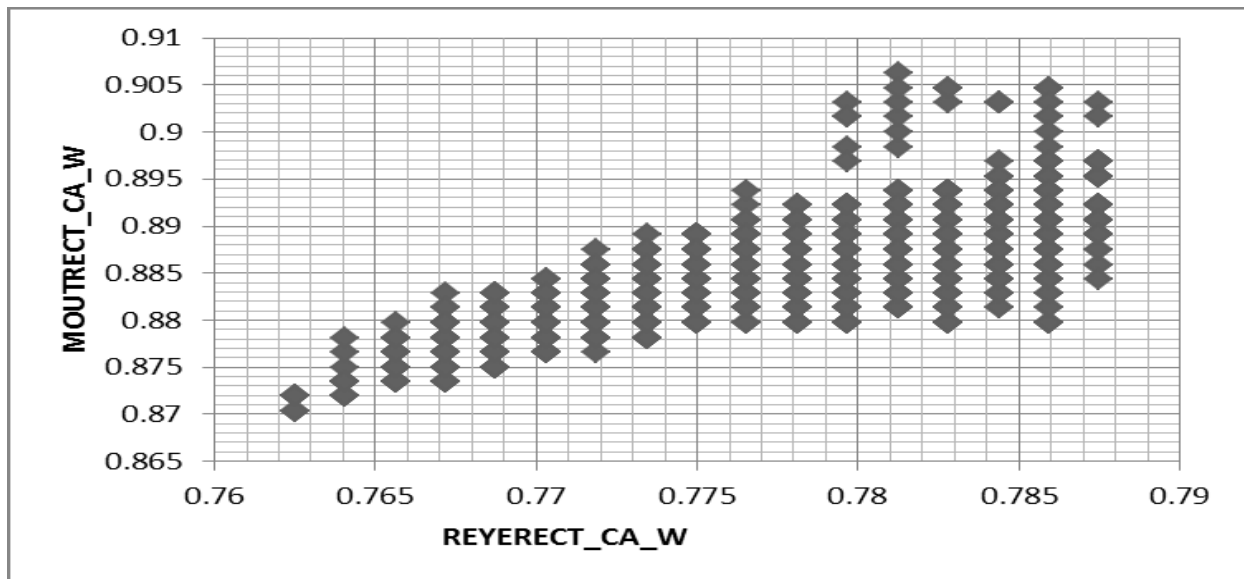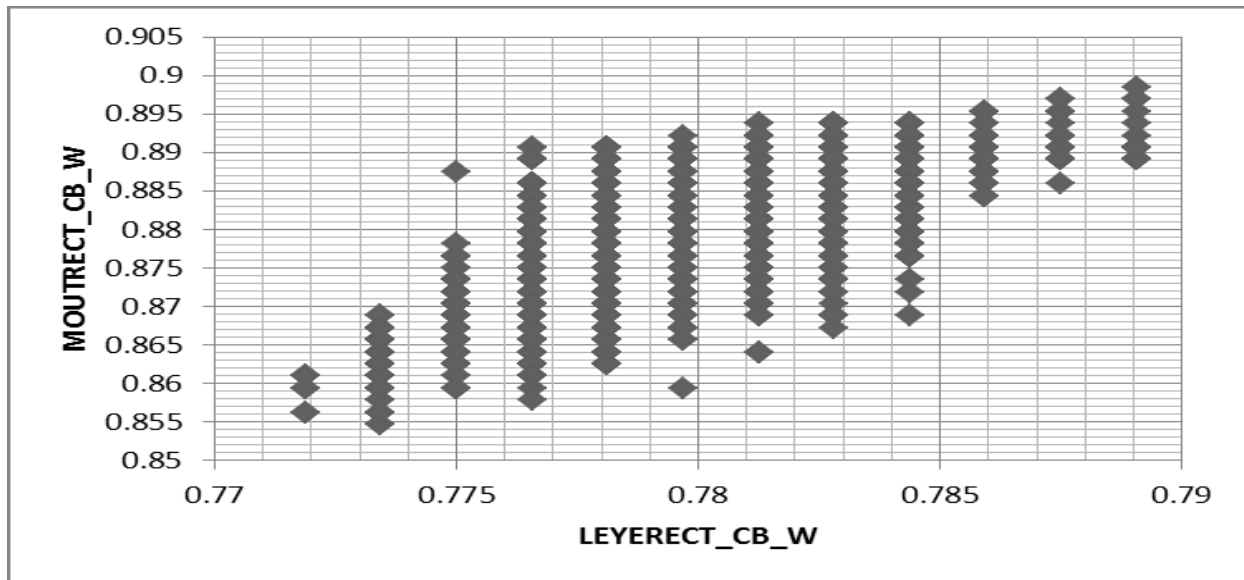

Figure 5.11: Scatter diagram left eye and mouth (width & camera B)

Scatter diagram in Figure 5.12, the x-axis refers to Left Eye Rectangle captured from Camera B for Height and y-axis is Mouth Rectangle captured from Camera B for Height. The line drawn on the scatter diagram showed a positive linear relationship between mouth height and left eye's height.
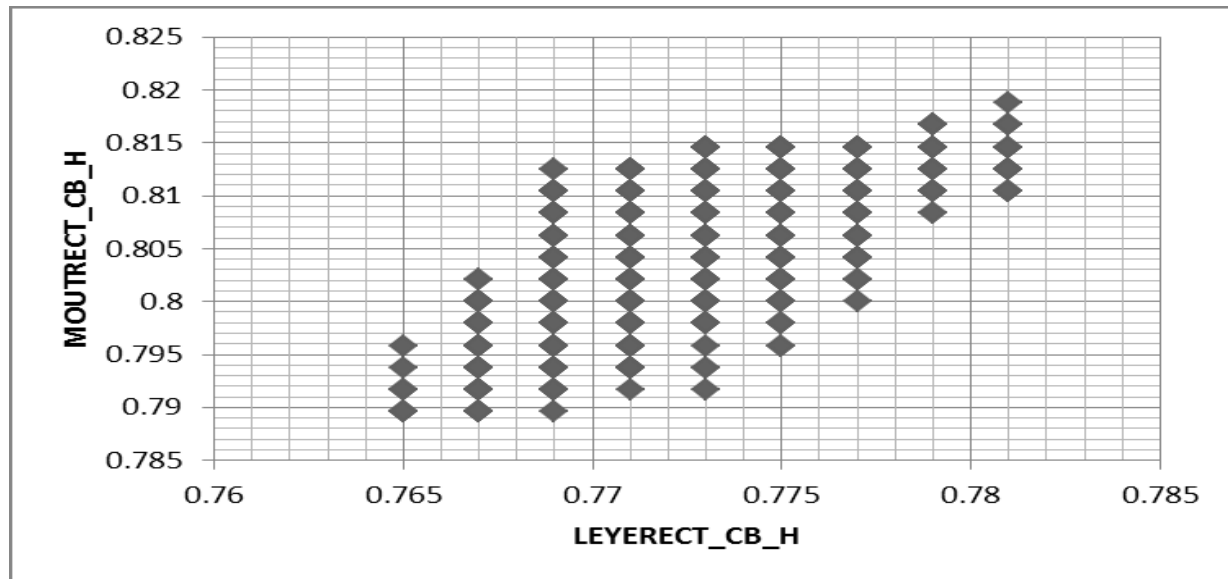
Figure 5.12: Scatter diagram left eye and mouth (height & camera B)

A straight line on every figure is a trend line or best fit line which can be drawn on the dots/points and consist as many dots/points as possible. If the dot/points can make a straight line going from the origin out to high x-axis values and y-axis values, then the variables are said to have a positive correlation. However if the line goes from a high-value on the y-axis down to a high-value on the x-axis, the variables then can have a negative correlation.

Data from faceLAB cameras contained both camera A and camera B for both right and left eyes. Values of the data are ranged from 0 to 1. All eight scatter diagrams showed a positive linear relationship between those two variables, mouth and eye. This proved that, eye and mouth has a strong relationship. The strength of the relationship can be found from the Pearson-r correlation where the correlation value (r-value) is always greater than 0.5. Scatter diagram is only used to find the occurrence of a relationship between the two variables and not the magnitude of the relationship.

As been mentioned earlier in section 5.1, all scatter diagrams are used to show positive relationship is occurred between the two variables or features (mouth movement and eye movement).

5.4   Proposed features: Lips and Eyebrows

The initial experiment with faceLAB which captured mouth movement and eye movement has proven that they have a strong relationship to each other when a driver is cognitively distracted. Based on the positive result between mouth movement and eye movement, the experiment is then extended with new features, lips and eyebrows. These features were tracked with other equipment called faceAPI toolkit. Details about the experiment setup can be found in Chapter 3. In this experiment, 10 volunteer drivers were involved.

Lips and eyebrows were chosen because many researchers in psychology area have proven that, lips and eyebrows also move when a person is with cognition and this is not limited to the driving environment only [41]. Lips and eyebrows are actually the most obvious features that can be seen on a person's facial expression when the person is thinking. Features like pupil diameter, gaze rotation, blinking frequency were also proven that they are also significantly related to cognitive distraction. However, those features could hardly be seen by the naked eyes. Specific cameras like infrared cameras are required to capture those features' information, but lips and eyebrows only require normal web cameras for their tracking.

faceAPI captured less features compared to faceLAB. However, faceAPI has an advantage over faceLAB where it can track the movements on lips and eyebrows precisely. Basically, faceAPI can be used to track head orientation, nose, eyebrows, lips and eye. faceAPI uses face landmark (FLM) standard that contains Interior Point, Mask Point and Feature Point when the camera detects a face. The whole landmark points can be seen as in Figure 5.13 below [69]. Table 5.3 lists the whole FLM groups by faceAPI.
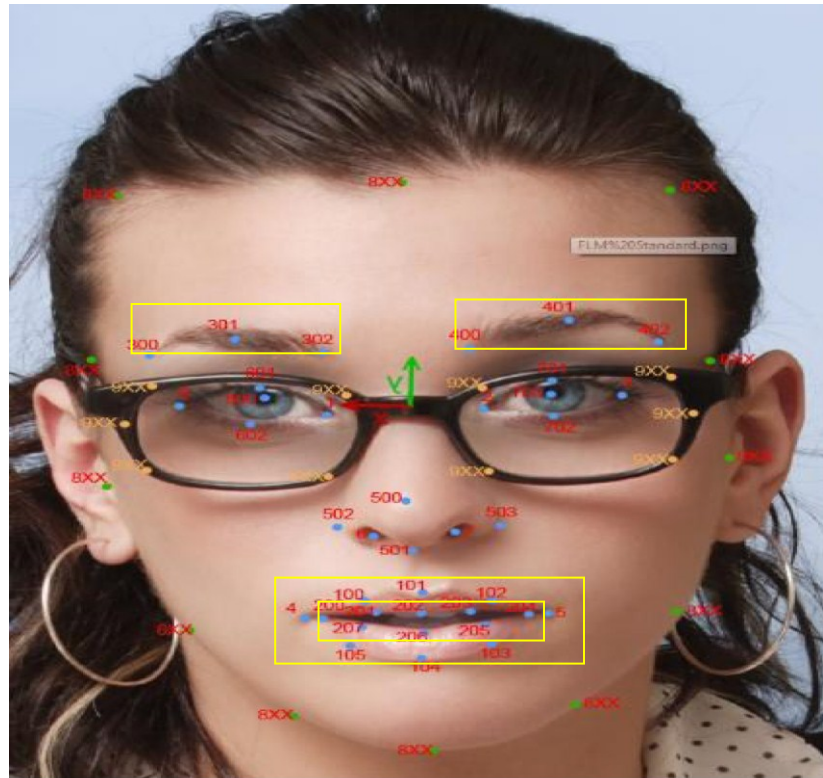
Figure 5.13: faceAPI face landmarks (FLM) [69]

0-99	reference
100-199	Outer lip contour
200-299	Inner lip contour
300-399	Right eyebrow
400-499	Left eyebrow
500-599	Nose contour
600-699	Right eye
700-799	Left eye
800-899	Mask
900-999	Glasses frame

Table 5.3: faceAPI FLM groups

However, for this study, only head orientation, eyebrows and lips are used. Other features like gaze rotation and blinking frequency are not able to be captured by faceAPI toolkit because faceAPI only uses a normal web camera. Thus, those features are excluded in this second

experiment. However, they are used later in the final experiment. By referring to Figure 5.14 below, both right and left eyebrows have three different muscles representing three different face landmarks (FLM):- 300, 301, 302 for right eyebrows (presented by red dots in colour) and 400, 401, 402 for left eyebrows (presented by black dots in colour). Lips can be categorized into two groups of FLM. Outer lip contour (presented by brown dots in colour): 100, 101, 102, 103, 104 and 105. Inner lip contour (presented by pink dots in colour): 200, 201, 202, 203, 204, 205, 206 and 207. Figure 5.14 shows a full figure of the used landmarks.
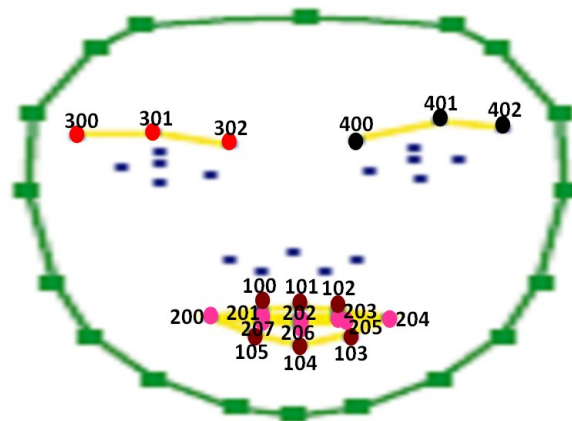
Figure 5.14: Used landmarks

Landmarks for eyes and nose have been ignored. Both left and right eyes landmarks were only captured at the beginning of the tracking and their movements were not tracked during the experiment. Nose's landmarks are not useful in this study because nose movement is very small. Thus, its data was also ignored. Table 5.4 below shows detail descriptions about the chosen FLM.

	FLM	DESCRIPTION	SIDE	POSITION
	100	Outer lip	Right	Upper
	101	Outer lip	Centre	Upper
	102	Outer lip	Left	Upper
	103	Outer lip	Left	Lower
LIPS	104	Outer lip	Centre	Lower
	105	Outer lip	Right	Lower
	200	Corner	Right	-
	201	Inner lip	Right	Upper
	202	Inner lip	Centre	Upper
	203	Inner lip	Left	Upper

	204	Corner	Left	-
	205	Inner lip	Left	Lower
	206	Inner lip	Centre	Lower
	207	Inner lip	Right	Lower
EYEBROWS	300	Eyebrow	Right	outer
	301	Eyebrow	Right	centre
	302	Eyebrow	Right	inner
	400	Eyebrow	Left	inner
	401	Eyebrow	Left	centre
	402	Eyebrow	Left	outer

Table 5.4: Used FLM descriptions

5.4.1   Experimental Results

The relationship between lips and eyebrows was also studied with two methods. However, this time, linear regression model and scatter diagram were used rather than Pearson-*r*, since linear regression and scatter diagram are easier to see, simpler to understand and easier to use for analysis purposes. On top of that, Pearson-r value also has a similar value to a slope value in a linear regression computation.

Linear regression is an approach to modeling the relationship between a scalar variable Y and one or more explanatory variables denoted X.   As explained in the background section, the equation for linear regression can also be simply presented as [62],

Y = mX+ b; where m is the slope and b is the intercept.

A driver's distraction or no distraction was the scalar variable whereas lips and eyebrows were the explanatory variables. Each driver had an average of 1380 seconds total of data frames and then they had been divided into two groups: Distracted and Not Distracted. In previous experiments, the correlation between eye movement and mouth movement on both height and width were analyzed with Pearson-*r* correlation. On average, the correlation *r* value between those features was around 0.7635 and had a very strong relationship between them.

In this experiment, linear regression model was used for two reasons. First, it was used to get the correlation coefficient between type of experiments (distracted or not distracted) with the features (lips and eyebrows). Thus, it can be used to see how lips data is varied with eyebrows

data when drivers were distracted or not distracted. Second, it was used to predict the value of lips from eyebrows or value of eyebrows from lips. For example, it can be used to predict the movement of lips when the eyebrows move at a certain value. Thus, if the movement on lips increases, the movement on eyebrows will also increase. Significantly, lips and eyebrows proportionally move to each other.

For the second reason, each driver set of data was trained and tested with linear regression model and the correlation coefficient (*r*-value), mean absolute error and root mean squared error were taken.

Mean absolute error is a quantity used to measure the predictions with the true value [70]. Root mean squared error is used to compute the differences between values predicted by a model and the values actually observed [71].

Linear regression model for eyebrows and lips was examined with three relationships. First relationship was between right eyebrows and lips. Second relationship was between left eyebrows and lips and the last group of relation was between lips and eyebrows. In the third relationship, movement values on both left and right eyebrows were summed and averaged. For all three relationships, lips values which consist of inner and outer contour have been averaged and considered as one variable/feature.

As in Table 5.5, the first group of relationship between right eyebrows and lips, the correlation coefficient in the linear regression model was 0.65 on average when the drivers were distracted and only 0.28 on average when the drivers were not distracted cognitively. The correlation coefficient falls in a reasonable relationship category. The mean absolute error when distracted was 0.23 whereas when not distracted it was 0.02. The root mean squared error for distraction was 0.30. The root mean squared error for non distraction was 0.07. The difference between correlation coefficients in distraction and non distraction was 37.27%. When cognitively distracted, the relationship between the right eyebrow and lips of every driver was greater than when the driver was not distracted.

Rightbrow & Lips	DISTRACTED-2				NO DISTRACTED-1			
	Correlation coefficient	Mean absolute error	Root mean squared error	Total Number of Instances	Correlation coefficient	Mean absolute error	Root mean squared error	Total Number of Instances
Driver1	0.60	0.19	0.27	690	0.44	0.03	0.07	690
Driver2	0.65	0.17	0.26	690	0.58	0.03	0.06	690
Driver3	0.73	0.18	0.26	690	0.24	0.02	0.06	690
Driver4	0.57	0.32	0.39	690	0.12	0.01	0.08	690
Driver5	0.58	0.24	0.31	690	0.20	0.02	0.07	690
Driver6	0.69	0.25	0.32	690	0.19	0.02	0.07	690
Driver7	0.76	0.23	0.31	690	0.45	0.03	0.08	690
Driver8	0.51	0.38	0.43	690	0.17	0.02	0.07	690
Driver9	0.79	0.15	0.22	690	0.21	0.02	0.07	690
Driver10	0.65	0.19	0.27	690	0.20	0.02	0.07	690
AVERAGE	0.65	0.23	0.30	690	0.28	0.02	0.07	690

Table 5.5: Linear regression model for right eyebrow and lips

The second group of relationship is between the left eyebrow and lips. Table 5.6 below shows that when the drivers were cognitively distracted, the correlation coefficient on average was 0.62. Correlation coefficient when the drivers were not distracted was 0.29. The difference percentage was 33.50%. Again, the correlation coefficient for the second relationship was also in a reasonable correlation between the lips and left eyebrow. Mean absolute error for distracted was 0.24 when distracted and only 0.02 when not distracted. Root mean squared error was 0. 31when distracted and 0.07 when not distracted.

Leftbrow & Lips	DISTRACTED-2				NO DISTRACTED-1			
	Correlation coefficient	Mean absolute error	Root mean squared error	Total Number of Instances	Correlation coefficient	Mean absolute error	Root mean squared error	Total Number of Instances
Driver1	0.47	0.20	0.29	690	0.41	0.02	0.07	690
Driver2	0.66	0.17	0.25	690	0.56	0.03	0.06	690
Driver3	0.71	0.18	0.27	690	0.21	0.01	0.06	690
Driver4	0.58	0.32	0.38	690	0.15	0.02	0.08	690
Driver5	0.47	0.25	0.34	690	0.33	0.02	0.07	690
Driver6	0.68	0.25	0.33	690	0.16	0.02	0.07	690
Driver7	0.76	0.25	0.31	690	0.43	0.03	0.08	690
Driver8	0.60	0.35	0.40	690	0.18	0.02	0.07	690
Driver9	0.68	0.20	0.26	690	0.24	0.02	0.07	690
Driver10	0.59	0.20	0.28	690	0.19	0.02	0.07	690
AVERAGE	0.62	0.24	0.31	690	0.29	0.02	0.07	690

Table 5.6: Linear regression model for left eyebrow and lips

On average, during distracted experiment for the third relationship between lips and eyebrows, driver correlation coefficient was 0.71598and during not distracted experiment it was only 0.31 as in Table 5.7. The difference between distracted and not-distracted was around 40.07% which is a relevant difference to distinguish between a driver's is distraction or non distraction. Again, 0.71 indicated that lips and eyebrows have a strong positive linear relationship with cognitive distraction via a firm linear rule.

Brows & Lips	DISTRACTED-2				NO DISTRACTED-1			
	Correlation coefficient	Mean absolute error	Root mean squared error	Total Number of Instances	Correlation coefficient	Mean absolute error	Root mean squared error	Total Number of Instances
Driver1	0.65	0.18	0.25	690	0.44	0.03	0.07	690
Driver2	0.69	0.17	0.24	690	0.58	0.03	0.06	690
Driver3	0.74	0.19	0.26	690	0.23	0.02	0.06	690
Driver4	0.59	0.22	0.31	690	0.16	0.02	0.08	690
Driver5	0.63	0.23	0.30	690	0.33	0.02	0.07	690
Driver6	0.71	0.25	0.31	690	0.21	0.02	0.07	690
Driver7	0.83	0.20	0.27	690	0.45	0.03	0.08	690
Driver8	0.65	0.32	0.38	690	0.21	0.02	0.07	690
Driver9	0.79	0.14	0.21	690	0.23	0.02	0.07	690
Driver10	0.77	0.15	0.22	690	0.22	0.02	0.07	690
AVERAGE	0.71	0.20	0.28	690	0.31	0.02	0.07	690

Table 5.7: Linear regression for lips and eyebrows

With correlation coefficient being equal to 0.71, the coefficient of determination or *r*-squared is equal to 49.84%. Mean absolute error was 0.20 when distracted and only 0.02 when not distracted. Root mean squared error was 0.28 when distracted and 0.07 when not distracted.

From the three relationships, it shows that driver cognitive distraction can be better detected when information from lips is combined with both left eyebrow and right eyebrow. Correlation coefficient between lip and both eyebrows was in a strong relationship whereas correlation coefficient between lip and either right or left eyebrows both were only in a reasonable relationship. Like in the initial experiment when mouth movement and eye movement was compared, right eye gave a bigger r-value compared to left eye. In this experiment, right eyebrow was also greater than the left eyebrow in correlation coefficient. Despite the difference was not huge, it significantly justified the difference clearly.

Linear regression model also gave a linear regression equation between lips and eyebrows. This equation is useful to predict the value for lips from eyebrows or vice versa. Since the second purpose was to find the best line to fit both features (lips and eyebrows), the data in the distracted experiment was combined together and this gave a total of 6900 data frames (690frames x 10drivers). If lip is X variable and eyebrow is Y variable, the best line to fit both variables is:-

$$Eyebrow = (0.9002 * Lip) + 0.5109$$

The line is a positive line with a slope equal to 0.9002. Correlation coefficient for this second purpose is 0.72 which again falls in a strong correlation between those two features.

If eyebrow is the X variable and lip is the Y variable, the best line to fit both variables is:-

$$Lip = (1.1109 * Eyebrow) -0.5675$$

The line is again a positive line with a slope equal to 1.1109. The Y-intercept for this equation is equal to -0.5675.

The relationship between lips and eyebrows was also analyzed with scatter diagram. Lips and eyebrows data for this scatter diagram was taken from the distracted group. Assume lips on the x-axis and eyebrows on the y-axis as in Figure 5.15. Scatter diagram was only drawn for the third relationship where the value for lips was averaged and considered as one variable, and where right and left eyebrows were averaged and considered as one variable too.
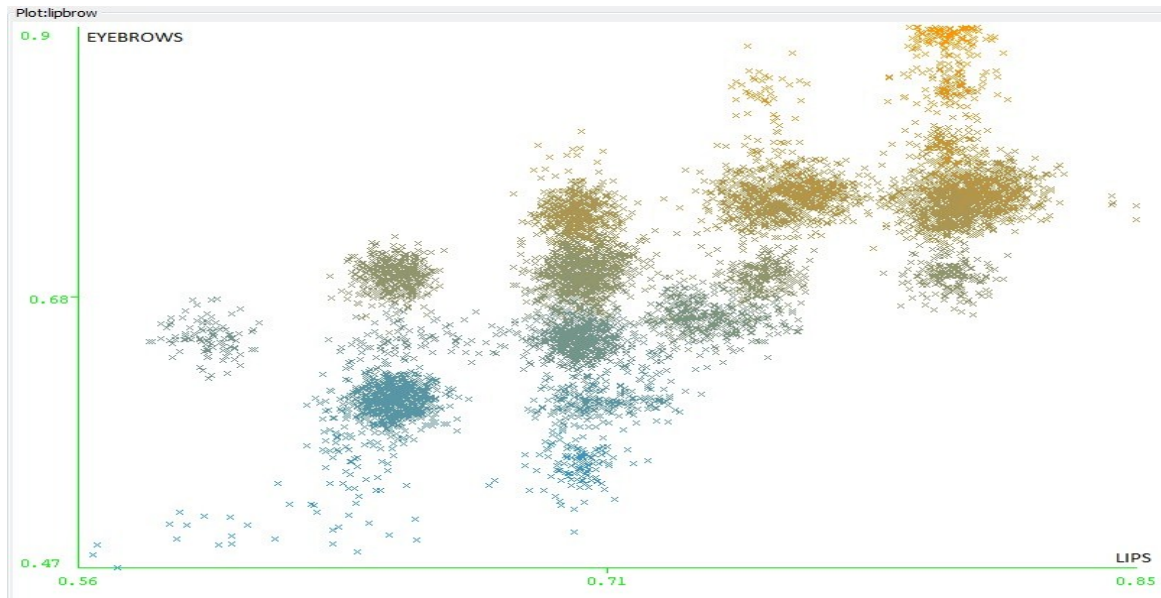
Figure 5.15: Scatter diagram for lips vs eyebrows when distracted

If the axis is changed the other way around, the scatter diagram between eyebrows and lips is also gives a linear positive relationship.
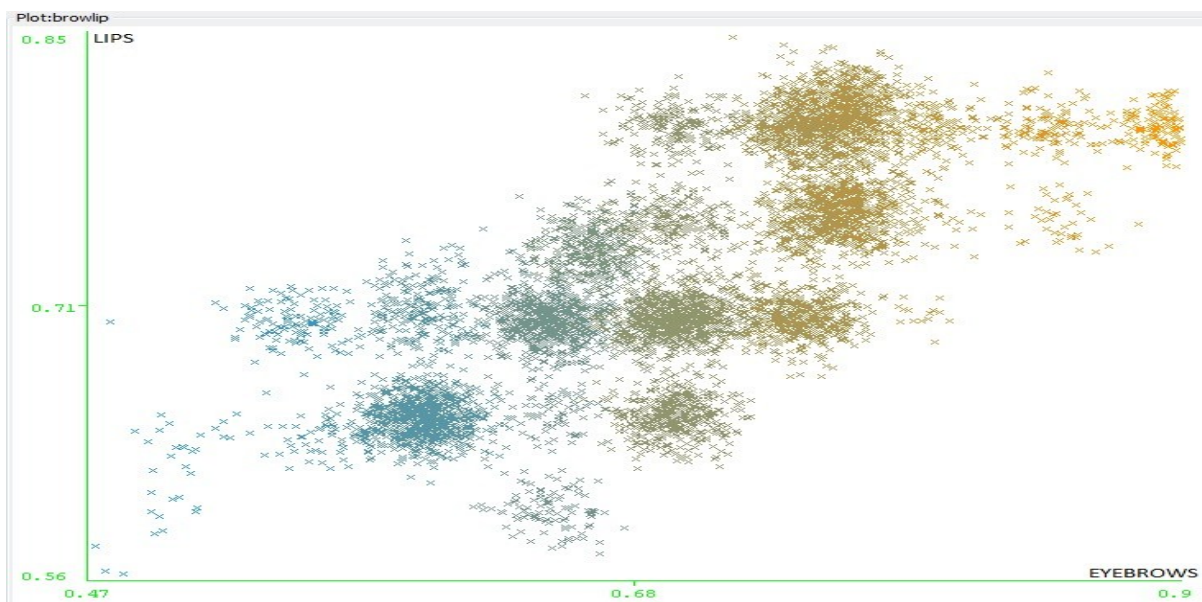
Figure 5.16: Scatter diagram for eyebrows vs lips when distracted

Both scatter diagrams above have showed that whenever the driver is cognitively distracted, lips and eyebrows will have a positive relationship between them, meaning that, when the lips are moving up or down, eyebrows will also moving up and down. Both scatter diagrams above were taken during tasks or distracted data. When a driver was with control experiment, at which no distraction or task is given to the driver, the scatter diagram for the relationship between lips and eyebrows showed a zero relationship or no relationship. It is showed in the Figure 5.17 below.
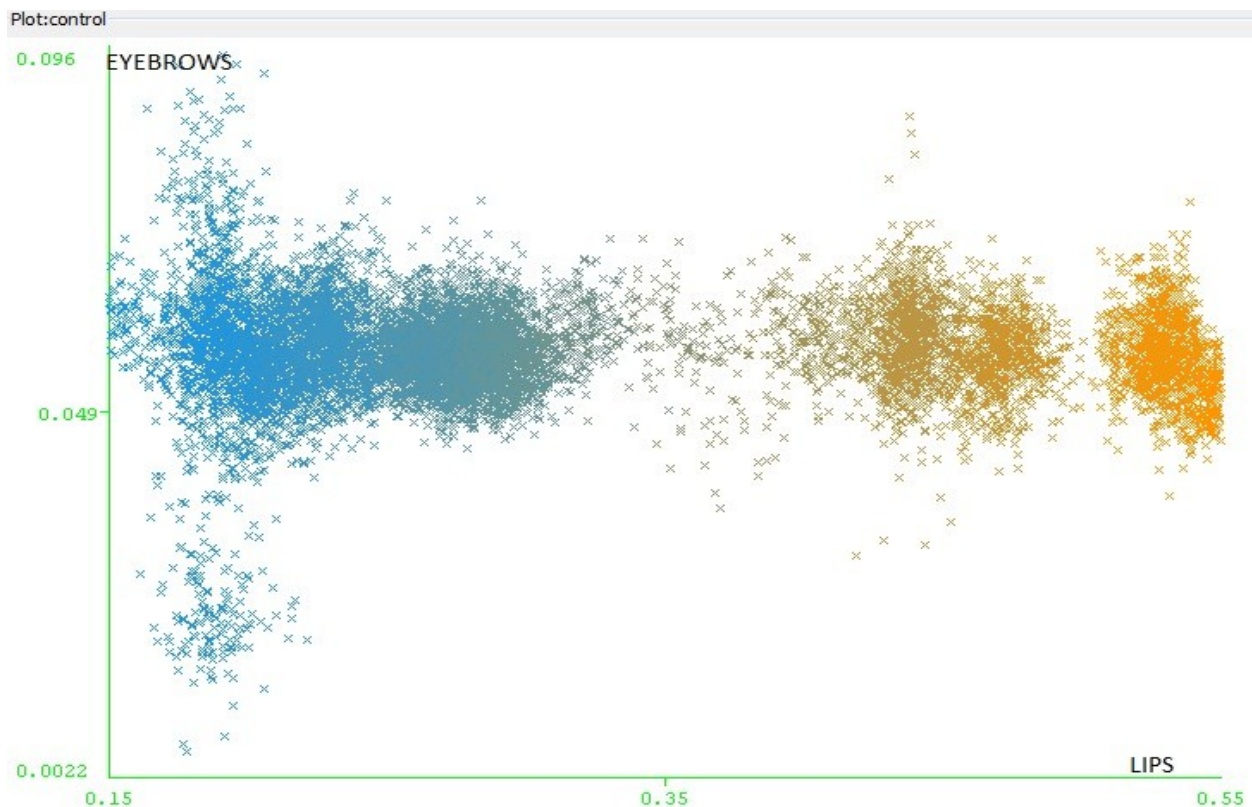
Figure 5.17: Scatter diagram for eyebrows vs lips in control experiment

5.5 Conclusions

This chapter investigated the correlation between facial features by using a correlation and regression models, in particular, the correlation between mouth movement and eye movement, as well as lips and eyebrows. In the first set of experiments, mouth movement and eye movement from six participants were tracked by faceLAB cameras. The movements referred to height and width for both features and the tracking involved two sets of cameras, A and B. Pearson-r analysis was used to examine the relationship between those two variables. It showed that, the r-values were all greater than 0.70 which means mouth movements and eye movements have a strong and positive relationship to each other when a driver is cognitively distracted. Scatter diagram was also drawn to see the relationship between those two variables. Again, all eight diagrams between mouth and eye showed a direct positive relationship.

Initial experiment has been extended where the features were changed to specific features to replace mouth and eye. Based on the positive results from the initial experiment, the second experiment used lips and eyebrows as new features. In the second experiment, a linear regression model was used instead of Pearson-r analysis. Linear regression model was used for two purposes. The first purpose is to find the correlation coefficient between lips and eyebrows and the second purpose is to find the best line that can fit both lips and eyebrows for prediction. For correlation coefficient, the relationship between lips and eyebrows was categorized into three relationships: right eyebrow and lips, left eyebrow and lips and both eyebrows and lips. Correlation coefficient for right eyebrow and left eyebrow with lips were in a reasonable relationship. However, correlation coefficient for eyebrows and lips was in a strong relationship. This means that, to detect driver cognitive distraction, lips should be correlated with both eyebrows rather than to examine them separately. Scatter diagram was also drawn to display the relationship between lips and eyebrows. Nevertheless, the scatter diagram for the second experiment only presented the average values of both lips and eyebrows from all the ten drivers. Scatter diagram for the second experiment also demonstrated positive relationship just like in the initial experiment.

Chapter 6

Detect Driver Cognitive Distraction based on Classification Algorithms

This chapter discusses about findings from different classification algorithms to detect driver cognitive distraction. Major computation was made on accuracy and sensitivity values. Five different types of classification algorithms are used: Support Vector Machine, Static Bayesian Network, AdaBoost, Logistic Regression and the main classification algorithm, Dynamic Bayesian Network.

6.1 Background

As been mentioned before, cognitive distraction is a complex state of mind where it is very subtle and presents inconsistent effects on the driver's behaviour. Therefore, using a general model for driver cognitive detection for every driver would have created a bias detection. It may eliminate the intra differences caused by cognitive distraction into inter differences between drivers.
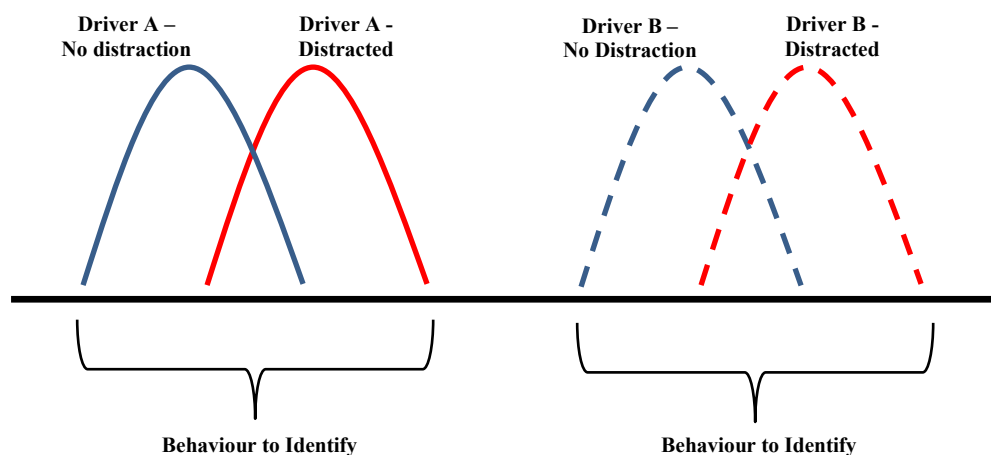
Figure 6.1: Behaviour variation between individual drivers and different distraction states.

Figure above shows different drivers with different state of distraction. The distribution curves present the behaviour of a certain driver under a certain distraction state.

To estimate visual distraction on a driver, several studies have been done with different predictive and detection models. For example, [27] proposed a predictive model to predict crash risk using mean dwell duration of in-vehicle device glances and frequency of relevant road events expressed by the product of road event rate and speed.

$$AccRate = MDD_{ivd} \times [(ColEvent + Curve + Steer) \times Vel + Turb]$$

$AccRate = Accuracy\ Rate$

$MDD_{ivd} = Mean\ Dwell\ Duration\ of\ in\ vehicle\ device\ glances$

$ColEvent = Collision\ Event$

$Curve = Road\ Curvature$

$Steer = Driver\ Steering\ Control\ and\ Speed$

$Vel = Velocity$

$Turb = Pertubations$

This predictive model was used to summarize crash factors caused by drivers' visual inattention. Just like many other models from different studies, the model for visual distraction had focused on the eye glances and eye movement.

Detecting cognitive distraction is not as easy as visual distraction detection because, cognitive distraction is more complex. Recently, classification on discrete state of cognitive distraction detection has been used. However most of those classification techniques are not suitable for a dynamic situation or continuous measure of distraction. As discussed before, to detect cognitive distraction, data fusion from multiple features or parameters is necessary. Integration of several performance measures over relatively long periods of time and personalization for different drivers are also sometimes required.

Model theories like ACT-R (Adaptive Control of Thought-Rational) in [72] have been used to make predictions regarding driver cognitive distraction, but they are actually better at describing the cognitive state of a driver rather than predicting. Therefore, the theories are not suitable to be used solely as a method to detect driver cognitive distraction. Thus, bottom-up data mining approach is the best solution to overcome this matter. Some studies with bottom-up data mining

algorithm like in [73] used a decision tree algorithm to estimate driver's cognitive workload from eye glances and driving performance. [36] used Support Vector Machine (SVM) and Bayesian Network (BN) to detect driver cognitive distraction from eyes and driving performance. AdaBoost algorithm was used in [45] and it captured only physiological measurement: gaze orientation, head orientation, pupil diameter and heart rate. In this study, all the existing algorithms, SVM, BN, AdaBoost and Logistic Regression are applied with new and different parameters. Their performances with accuracy rate are compared in this chapter.

6.2 Data Capture from Real Car Driving for Different Data Mining Algorithms

Before discussing about the results obtained from all the data mining algorithms mentioned above, Figure 6.2 below shows how the data captured from a real car driving using faceLAb and faceAPI, are used for different algorithm testing in this study.
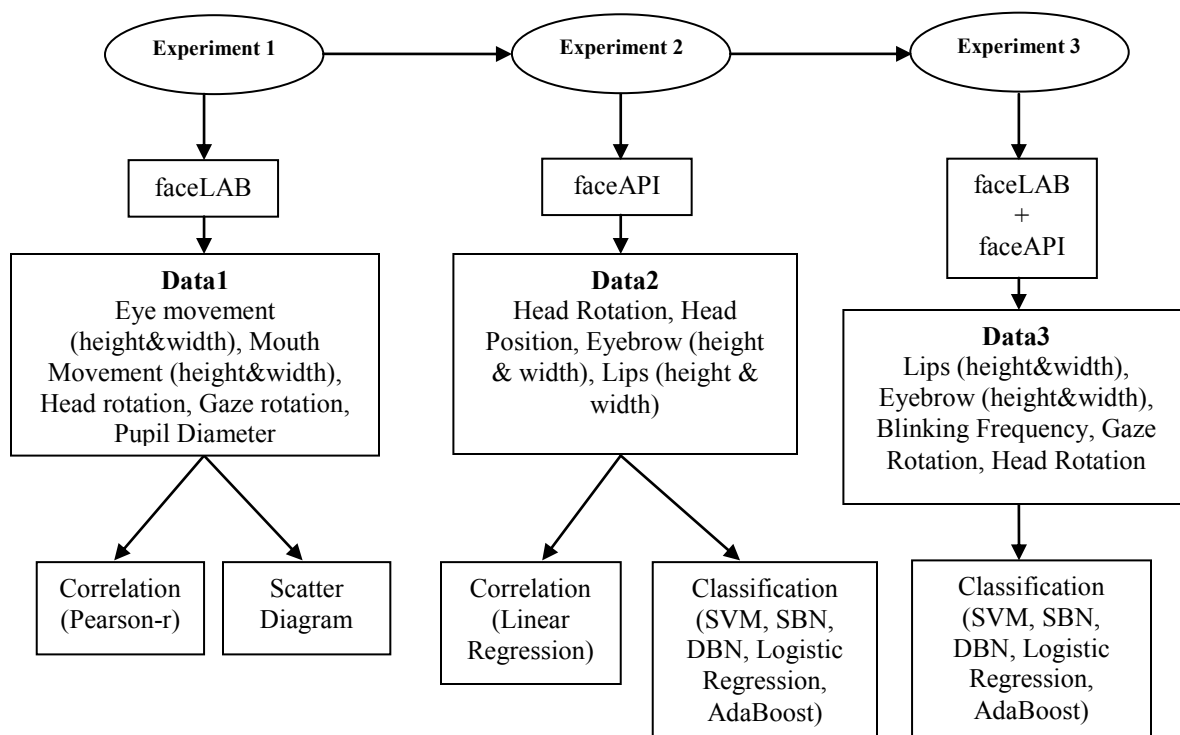
Figure 6.2: Data Division from each experiment: Equipment used and Result analysis methods

Basically every set of experiment contained its own data set. As discussed in Chapter 3 (experiment), initial experiment was conducted with faceLAB equipment. Data from first experiment was used to determine a correlation and relationship between eye movement and mouth movement. Second experiment was conducted for two major purposes: (a) correlation and regression and (b) classification. Since the first experiment has proven that eye movement and mouth movement are strongly correlated, data from the second experiment was used to prove the relationship between lips and eyebrows. Results for correlation and regression from the first and second experiment were discussed in chapter 5 (correlation and regression). Parameters from the second experiment, lips and eyebrows are found to be strongly correlated. Classification algorithms like SVM, SBN, DBN, Logistic Regression and AdaBoost are used to classify the parameters from the second experiment. Results for this classification are described in this chapter. Finally, the last set of data was produced from the data fusion captured by faceLAB and faceAPI in the third experiment. Selective parameters used in the first and second experiment were combined in the last experiment. Classification results for the final experiment are discussed in this chapter. Accuracy rate and sensitivity rate are the important results produced by these algorithms.

6.3 Classification Algorithms in Data Mining

In data mining, classification is a function that assigns items in a collection to target categories or classes. Classification is to find the accuracy to on predicting the target class for each case in the data. Classification algorithm is used when the characteristics of the data class is known. For example, in this study, the data class is defined as distracted or not distracted over the duration of the experiment. Thus, the class assignments are known as distracted or not distracted. Classification is discrete and it does not imply any order. The simplest type of classification problem is a binary classification, just like in this study: distracted or not distracted. Other examples of binary classification are like between high credit rating or low credit rating, fatigue or not fatigue and etc. Multiclass classification has more than two values, for example: low, medium, high and unknown rating [76].

In the model building or training process, a classification algorithm will find relationships between the values of the predictors and the value of the target. Different classification

algorithms use different techniques for finding those relationships. The relationships are later summarized in a model which can be applied to a different data set in which the class assignments are unknown. Classification models are then tested by comparing the predicted values to known target values in a set of test data. The historical data for a classification project is typically divided into two data sets: training the model and testing the model.

Task drives and control drives were used to determine the driver cognitive state as distracted and not distracted respectively. These two states were annotated during the data collection process. The classification algorithm especially in Bayesian Network models used this information to define the hypothesis node (the parent node in the network). The evidence or sometimes called as instances (the children nodes) are used by the BN models were the all measurements from head rotation, gaze rotation, blinking frequency, lips and eyebrows (from third experiment) and head rotation, head position, lips and eyebrows (from the second experiment). The evidence were discretized and summarized over time. Blinking frequency data captured from faceLAB cameras were averaged to calculate the mean of blinking frequency.

A variety of BN, SVM, LR and AdaBoost models were trained and tested for each participant from second and third experiments. Following the common practice in data mining [24], two-thirds of the evidences (equivalent to approximately to 29-33 minutes of driving from each participant) were chosen at random to form the training set and the rest (equivalent to approximately to 15-17 minutes of driving from each participant) formed the testing set. The training of BN models included the parameter learning and estimation. Parameter learning identifies the conditional probability for those connections in the network. The BN models were trained by using a Bayesian Network Toolbox (BNT) with a Matlab platform [50]. The structures of BNs were constrained so that the training procedure can be computationally feasible. For Dynamic Bayesian Network (DBN) two structures in the network were constructed: intrastructure and interstructure. Intrastructure is when the hidden node is connected to the next time step's hidden node. Interstructure is when the hidden node is connected to the evidence nodes. Model performance in BN was done with the summarization of the parameters used to create the evidences for training and testing. Sequence length for the parameters was considered for DBN models. A sequence contained multiple and consecutive training evidences. The longer sequences will provide more comprehensive and stable time-dependent relationships between

driver distraction and driver performance. The tested sequence length for this study was 60 seconds like been mentioned earlier in chapter 4, section 4.42.

This chapter will explain several types of classification algorithms used to classify data between distracted and not distracted, cognitively. Four main classification algorithms are used for this purpose: Support Vector Machine, AdaBoost, Logistic Regression and Bayesian Network (Static and Dynamic).

6.3.1 Support Vector Machine (SVM) for Cognitive Distraction Detection

SVM is an algorithm used to learn from examples in order to assign labels to objects. For example, in this study, SVM was used to determine whether the facial features movement from drivers give an expression for cognitively distracted or not distracted. In essence, SVM is a mathematical entity that uses an algorithm to maximize a particular mathematical function with respect to the given collection of data. SVM has a good generalization property and classification performance is not restricted to any specific application of the data. SVM in this study focussed on a binary classification where the data has only two types: distracted or not distracted, thus, basically a linearly separable binary classification of SVM has been used. In theory, to solve this type of SVM classification, with L training points, where each input x_i had D attributes and it is in one of the two classes $y_i = \pm 1$, therefore the training data is of the form of:

$$\{x_i, y_i\} \text{ where } i = 1.....L, y_i \in \{-1,1\}, x \in \Re^D$$

Assuming, the data is linearly separable, meaning that a line can be drawn on a graph of x_1 versus x_2 separating the two classes when $D=2$ and a hyperplane on graphs of $x_1, x_2.....x_D$ for when $D > 2$, the hyperplane can be described by $w.x + b = 0$ where:

a) $\mathbf{w}$ is normal to the hyperplane

b) and $\dfrac{b}{\|w\|}$ is the perpendicular distance from the hyperplane to the origin.

SVM is to orientate the hyperplane in such a way as to be as far as possible from the closest members of both classes like in the figure below:
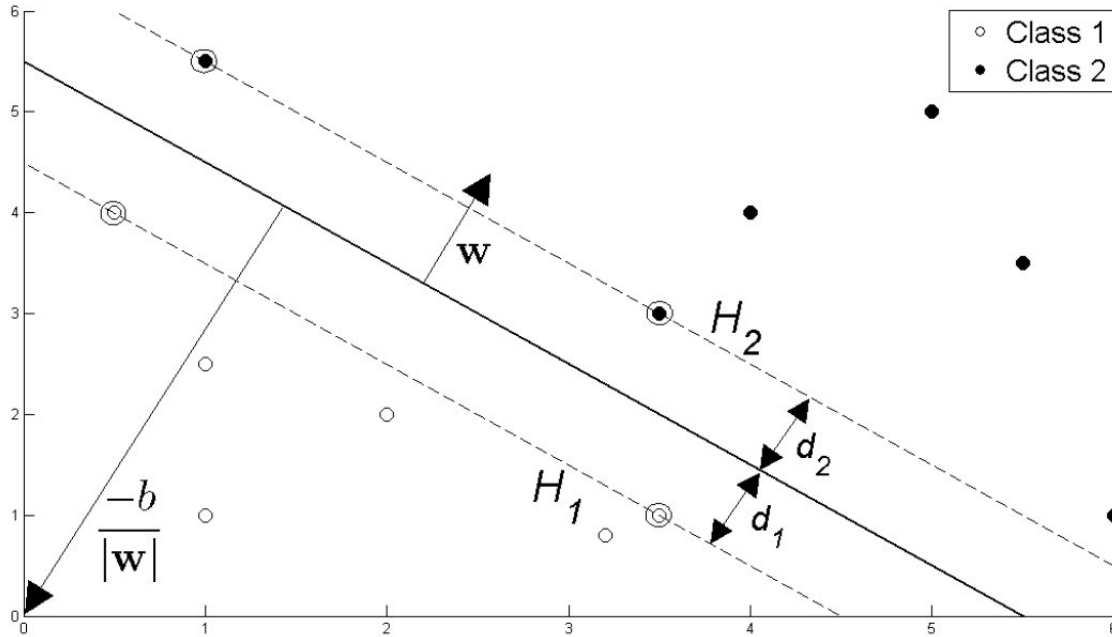
Figure 6.3: Hyperplane through two linearly separable classes [74]

In general, to use this linearly separable binary classification problem the steps below are required to be followed. Details about how this linearly separable binary classification is computed can be found in [74]

a) Create $\mathbf{H}$ where $\mathbf{H}_{ij} = y_i y_j x_i \bullet x_j$

b) Find α so that

$$\sum_{i=1}^{L} \alpha_i - \frac{1}{2} \alpha^T \mathbf{H} \alpha$$

Is maximized, subject to the constraints

$$\alpha_i \geq 0 \forall_i \text{ and } \sum_{i=1}^{L} \alpha_i y_i = 0$$

c) Calculate $w = \sum_{i=1}^{L} \alpha_i y_i x_i$

d) Determine the set of Support Vectors S by finding the indices such that $\alpha_i > 0$

e) Calculate $b = \dfrac{1}{N_s} \sum_{s \in S} (y_s - \sum_{m \in S} \alpha_m y_m x_m \cdot x_s)$

f) Each new point x' is classified by evaluating $y' = \text{sgn}(w \cdot x' + b)$

In SVM, its kernel functions can be divided into linear or non-linear functions. To construct linearly discriminated feature space, the simple canonical dot product is formulated as

$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$$

where, x_i & x_j are pattern vectors and Φ is a mapping function. Appropriate mapping function is required to map the data where a linear separation is possible.

The boundary between two data sets is rather easy to determine because the kernel is simple. However, a linear kernel is not useful for non-linear data sets. Usually, for non-linear data set a kernel function called radial basis function (RBF) is used.

SVM algorithm is easy to understand especially when applying for binary classification. In training state, the border is adapted to meet the requirement of the margin between distracted and not distracted training examples.

Several studies for driver cognitive distraction detection like [24][36][45][46] used SVM as their classification algorithms.

6.3.2 AdaBoost for Cognitive Distraction Detection

Adaptive Boosting or AdaBoost is an algorithm like SVM used for learning. AdaBoost is adaptive because the subsequent classifiers built are amended in favour of the misclassified instances from the previous classifiers. There are several general characteristics of AdaBoost:

a) It is a linear classifier with all its desirable properties
b) Its output converges to the log likelihood ratio
c) It has good generalization properties
d) It is a feature selector with a principled strategy
e) It is close to sequential decision making

AdaBoost is an algorithm for constructing a strong classifier as linear combination from simple and weak classifiers $h_t(x)$.

$$f(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$$

$h_t(x)$ is a weak classifier, hypothesis or feature. $H(x) = sign(f(x))$ is a strong or final classifier or hypothesis where *sign()* is.

The computational complexity to select h_t is independent of t and all information about previously selected classifiers captured in D_t. [75] has written the AdaBoost algorithm as below:

Given : $(x_1, y_1), \ldots \ldots (x_m, y_m); x_i \in \aleph, y_i \in \{-1,1\}$
Initialize weights $D_1(i) = 1/m$
For t=1,\ldots\ldots\ldots,T:
  1) Weak Learn-returns the weak classifier $h_t : \aleph \to \{-1,1\}$ with minimum error distribution D_t

  2) Choose $\alpha_t \in \Re$

  3) Update

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

  where Z_t is a normalization factor chosen so that D_{t+1} is a distribution
Output the strong classifier:

$$\mathrm{H}(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$$

Algorithm 6.1: AdaBoost algorithm by Singer & Schapire [75]

AdaBoost is a simple classification algorithm to implement. Its feature selection is on very large sets of features and the algorithm is fairly good for generalization. However, AdaBoost algorithm sometimes can over fit in the presence of noise which pose the disadvantage of this learning algorithm. A paper by [45] proposed the idea to detect driver cognitive distraction using AdaBoost algorithm. However, in this study, it has been found that, AdaBoost is not really a better algorithm than SVM as mentioned in the paper.

6.3.3 Logistic Regression for Cognitive Distraction Detection

Logistic regression is a statistical modelling work with odds rather than proportions. It means that, the odds are the ratio of the proportions for two possible outcomes: distracted and not distracted, low and high, success and fail. If $\hat{p}$ is the proportion for distracted, therefore 1- $\hat{p}$ is the proportion for not distracted. Therefore, odds can be computed as:

$$ODDS = \frac{\hat{p}}{1-\hat{p}}$$

Logistic regression has a minor relationship with simple linear regression. In simple linear regression as explained in chapter 5, it modelled the mean μ of the response variable y as a linear function of the explanatory variable: $\mu = \beta_0 + \beta_1 x$ where β_1 is the slope of the line and β_0 is its intercept. However, in logistic regression, it models the mean of the response variable p in terms of an explanatory variable x. p and x can be related through the equation of $p = \beta_0 + \beta_1 x$. Unfortunately, this is not good enough to describe the logistic regression model because as long as $\beta_1 \neq 0$, an extreme value of x will give values of $\beta_0 + \beta_1 x$ that are inconsistent with the fact that $0 \leq p \leq 1$.

Therefore, the logistic regression solution is to transform the ODDS using the natural log (log ODDS).

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

Relationship between p and x for some different values of β_1 and β_0 can be seen as in the graph below [76]
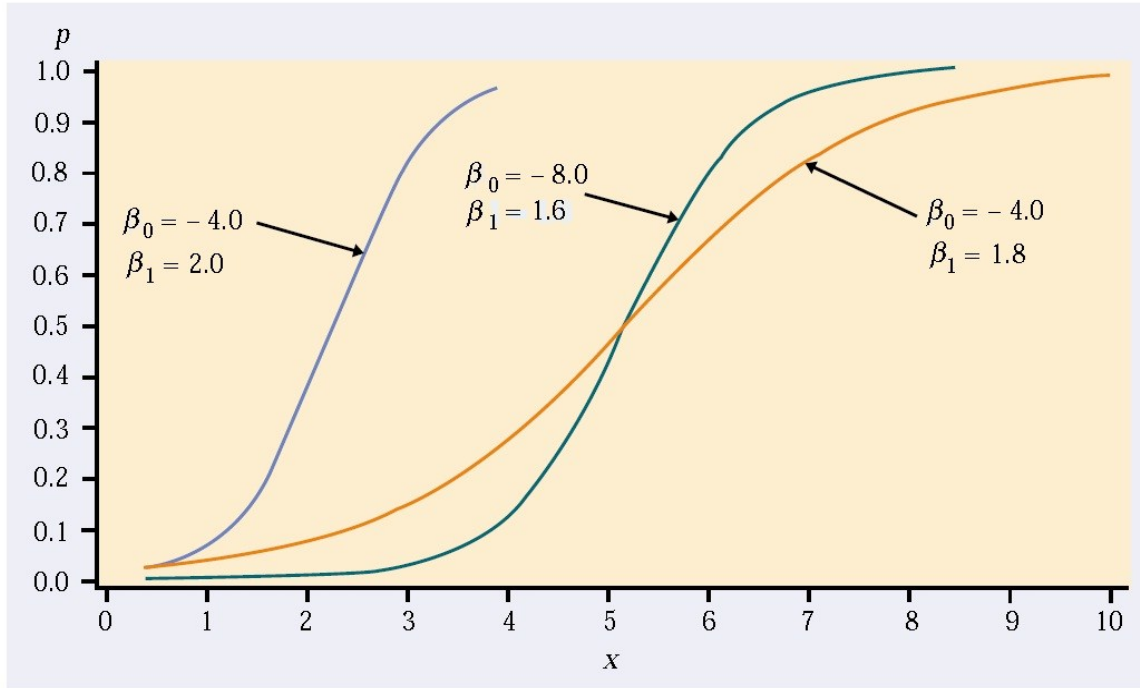
Figure 6.4: Logistic regression examples [76]

p is a binomial proportion and x is the explanatory variable in logistic regression model. The parameters for this model are β_1 and β_0. Logistic regression has been used in [24][36] and it was just used as a comparison algorithm with SVM, SBN and DBN.

6.4   Classification Algorithm on First Experiment Data

Data from the first experiment was purposely used to examine a correlation or relationship between mouth movement and eye movement based on their height and width. Results for this correlation and relationship check between those two features are explained in chapter 5.

However, for the purpose of learning about classification algorithms, data from the first experiment was also used for certain functions in the classification. Since, the experiment had only run for 8 minutes and 49 seconds each for control and task, with only 6 participants involved in this lab setting experiment, the data collected is very limited to run for classification algorithms. 8 specific features were tracked in this experiment: distraction, mouth height, mouth

width, eye height, eye width, head rotation, gaze rotation and pupil diameter. Those 8 features were applied to all six participants involved in this experiment.

Initially, from the literature review there are many available features to be selected for driver's cognitive distraction detection. However for this study, the process of selecting the relevant features for use in the model construction are done by searching technique for proposing new features, along with an evaluation measure which scores the different feature subsets. There are basically three available feature selection algorithms: wrappers, filters and embedded. In this study a filter method has been used to select features. Filter methods is fast to compute yet still able to capture the usefulness of the feature set. Common measures including Pearson correlation coefficient and inter/intra class structure distance. Thus, based on the results received from correlation coefficient accomplished in chapter 5 it has been found that lips, eyebrows are useful and more informative as new features to be proposed and combined with the existing features like head rotation, gaze rotation and blinking frequency. The main assumption when using this feature selection technique is that the data may contains many redundant or irrelevant features. Redundant features are which provide no more information than the currently selected features and irrelevant features provide no more useful information.

Since, classification is not a main objective in the first experiment, only accuracy rate or correctly classified data was computed on the data from the first experiment. Firstly, the data was analysed in WEKA to compute the accuracy rate using SVM, SBN, Logistic Regression and AdaBoost.

The accuracy rate was computed individually for every participant and it had run randomly among participants.

Participants	LR	ADA	SBN	SVM
1	66.00	70.00	71.60	77.10
2	67.30	73.30	72.50	76.40
3	72.40	72.20	73.70	76.00
4	83.20	78.50	79.50	81.20
5	84.80	85.60	81.20	84.80
6	60.60	62.20	64.40	64.40
AVERAGE	72.39	73.60	73.84	76.60

Table 6.1: Accuracy rate on data from first experiment by participants

Based on the Table 6.1 above, SVM is still the best classification algorithm with an average of 76.64%. SBN average % was 73.84%, AdaBoost was 73.64% and Logistic Regression was 72.39%. Logistic regression is the lowest result because large sample sizes are required for LR to provide sufficient numbers of the response variable. With only small sizes, the test has low power and is unlikely to detect the deviations [62][63].

DBN was also used for the first time, for the first experiment data. DBN is a classification algorithm which requires a time slice for its classification. Since the data collected was very short, the computation was only possible to be computed using cumulative accuracy. DBN was computed on each participant and the data was cumulated one after another to give a cumulative accuracy. Results are shown in Figure 6.5 below.
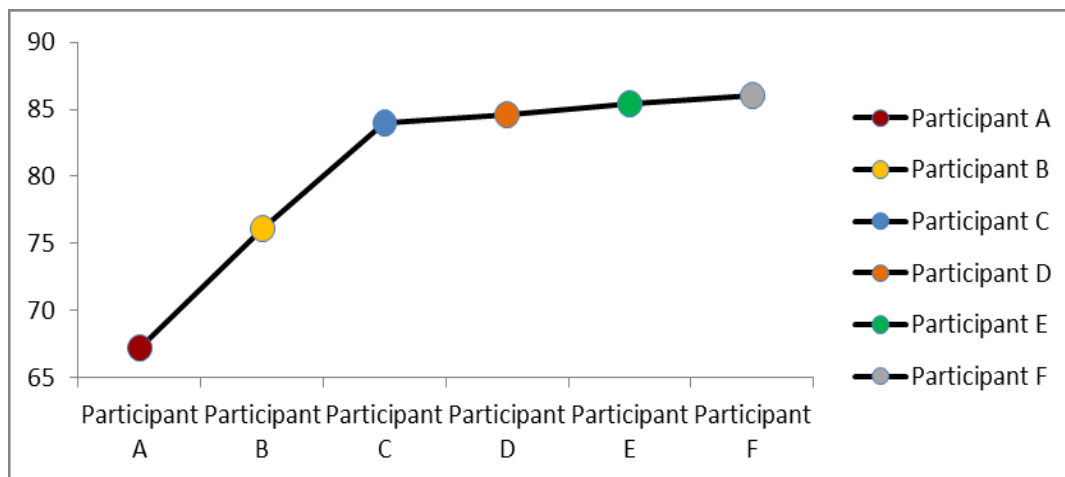
Figure 6.5: DBN cumulative accuracy

6.5   Classification Algorithm on Second Experiment Data

In the second experiment, all features captured by faceAPI include Head Position, Head Rotation, Lips and Eyebrows. faceAPI also provides tracking for nose and eye movement for both left and right eyes. However, data from nose was excluded because nose movement was really insignificant since the movement was very small. Data from eyes were also removed because faceAPI only tracked initial position from both eyes. It did not track the movement of the eyes dynamically from time to time. For data accuracy of distraction detection or

hit/successful rate computation with SVM, SBN and Logistic Regression, distracted data were sub-divided into two groups. The first group of data consist of lips, eyebrows, head rotation and head position. The second group of data consist of only head position and head rotation (without lips and eyebrows). Lips and eyebrows consist of several face landmarks (FLMs). Eyebrows have data from outer, center and inner FLMS, whereas lips have outer and inner outliers FLMs. However, to develop model from this second experiment, eyebrows were divided into two: left and right. On the other hand, lips were divided also into two: outer and inner. A standard deviation has been used just like in [45][46][57] to combine those face landmarks as in (1), (2), (3), (4), (5), (6), (7), (8), (9) and (10).

$$a(i) = \sqrt{a_{100}(i)^2 + a_{101}(i)^2 + a_{102}(i)^2 + a_{103}(i)^2 + a_{104}(i)^2 + a_{105}(i)^2} \tag{1}$$

$$\sigma(i) = \sqrt{\frac{1}{30} \sum_{j=i-30}^{i} (a(j) - \bar{a})^2} \tag{2}$$

Here, $a(i)$ is the combined outer lips, $a_{100...105}(i)$ is the outer lips face landmarks and $\sigma(i)$ here is the standard deviation of the outer lips.

$$b(i) = \sqrt{b_{200}(i)^2 + b_{201}(i)^2 + b_{202}(i)^2 + b_{203}(i)^2 + b_{204}(i)^2 + b_{205}(i)^2 + b_{206}(i)^2 + b_{207}(i)^2} \tag{3}$$

$$\sigma(i) = \sqrt{\frac{1}{40} \sum_{j=i-40}^{i} (b(j) - \bar{b})^2} \tag{4}$$

Here, $b(i)$ is the combined inner lips, $b_{200...207}(i)$ is the inner lips face landmarks and $\sigma(i)$ here is the standard deviation of the inner lips.

$$c(i) = \sqrt{c_{300}(i)^2 + c_{301}(i)^2 + c_{302}(i)^2} \tag{5}$$

$$\sigma(i) = \sqrt{\frac{1}{15} \sum_{j=i-15}^{i} (c(j) - \bar{c})^2} \tag{6}$$

Here, $c(i)$ is the combined right eyebrow, $c_{300...302}(i)$ is the right eyebrow face landmarks and $\sigma(i)$ here is the standard deviation of the right eyebrow.

$$d(i) = \sqrt{d_{400}(i)^2 + d_{401}(i)^2 + d_{402}(i)^2} \tag{7}$$

$$\sigma(i) = \sqrt{\frac{1}{15}\sum_{j=i-15}^{i}(d(j)-\bar{d})^2} \qquad (8)$$

Here, $d(i)$ is the combined left eyebrow, $d_{400...402}(i)$ is the left eyebrow face landmarks and $\sigma(i)$ here is the standard deviation of the left eyebrow.

In [45], the formula for standard deviation was also used for head rotation. However, the formula given in the paper was incorrect. Head rotation captured by the stereo camera system, faceLAB, only has 3 values (α, β, γ) and not 4 values as stated in the formula given. Therefore, in this study, the head rotation formula of standard deviation in order to combine the angles is given as below:

$$h(i) = \sqrt{h_{\alpha}(i)^2 + h_{\beta}(i)^2 + h_{\gamma}(i)^2} \qquad (9)$$

$$\sigma(i) = \sqrt{\frac{1}{3}\sum_{j=i-3}^{i}(h(j)-\bar{h})^2} \qquad (10)$$

Here, $h(i)$ is the combined head orientation, $h_{\alpha,\beta,\gamma}(i)$ is the head rotation with three different Euler angles (α, β, γ) and $\sigma(i)$ here is the standard deviation of the head orientation.

In the formulas above, the value of 30, 40, 15 and 3 refer to the number of FLMs captured. For instance, for outer lips, there are 30 FLMs captured and this can be seen in the faceAPI data file. Each FLM is captured from different angles and different axis. Those values are then combined and given the total number of FLMs measured for a standard deviation.

Data which contained all features under Distracted was used for distraction detection with lips and eyebrows. Data which contained only head position and head rotation was used for distraction detection without lips and eyebrows. Description about data divisions can be seen from the block diagram in the Figure 6.6 below. Data were also divided for scatter diagram and linear regression purposes and the results and discussion about these purposes can be found in chapter 5. Three different types of classification algorithms were used initially: SBN, SVM, and Logistic Regression, for accuracy or successful rate comparison. A rate for correctly classified instances from the data was computed.
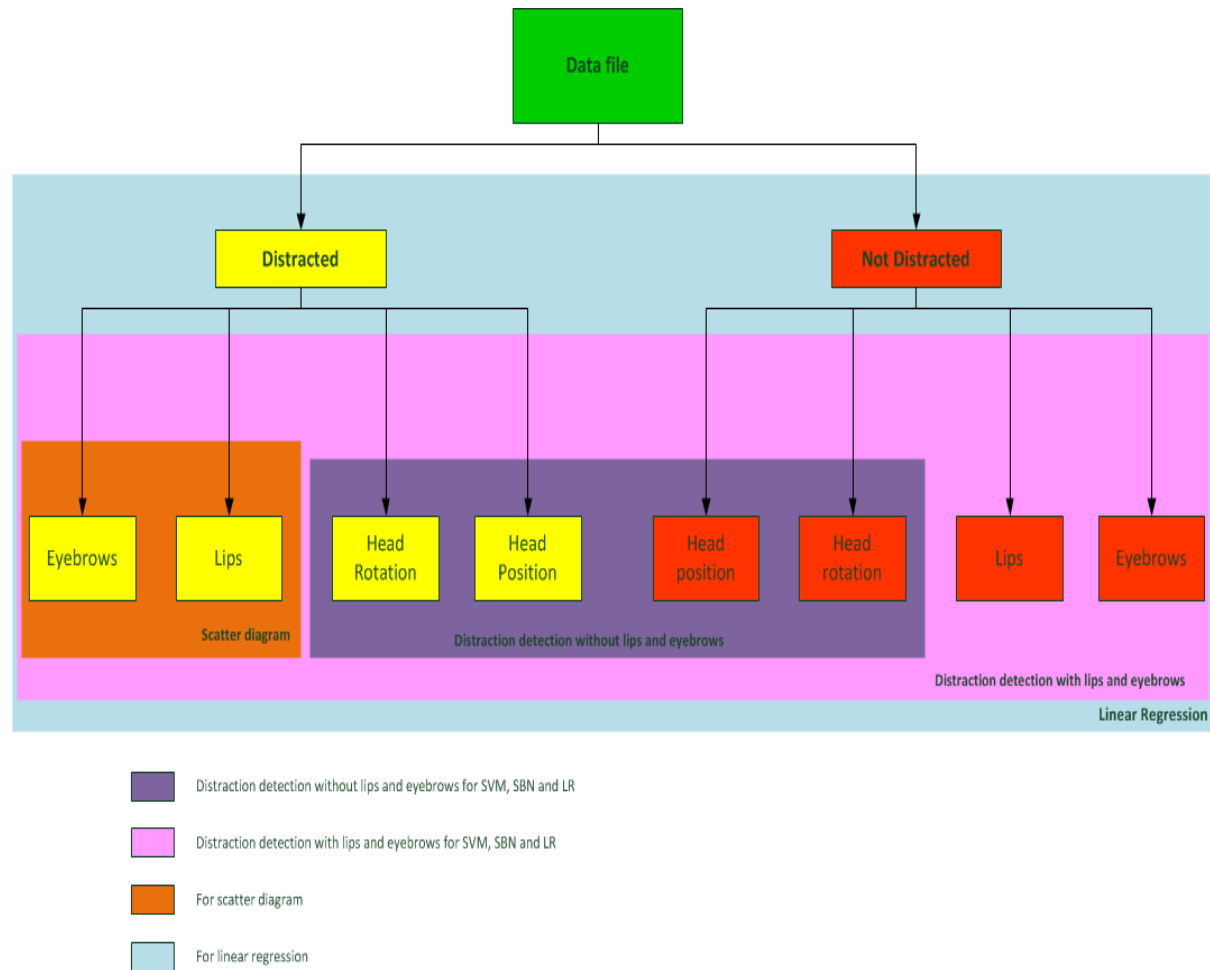
Figure 6.6: Data division

As mentioned before, data from the second experiment was first classified using those four algorithms in two different subsets of data: (a) with lips and eyebrows and (b) without lips and eyebrows as in the figure below. Figure 6.7 shows that there is a significant difference between the group with lips and eyebrows and the one without lips and eyebrows in every classification algorithms. Support Vector Machine was the best classification method to classify whether the drivers were cognitively distracted. With SVM, data with lips and eyebrows was 79.58% correctly classified and without lips and eyebrows was 71.19%. Thus, it showed that with lips and eyebrows, cognitive distraction in a driver can be classified correctly 8.39% more to the data set without lips and eyebrows. Logistic regression had an accuracy of 75.85% with lips and eyebrows and 68.41% without lips and eyebrows. Logistic regression is the less accurate data

mining in this study. SBN accuracy was 77.57% with lips and eyebrows and 69.99% without lips and eyebrows.
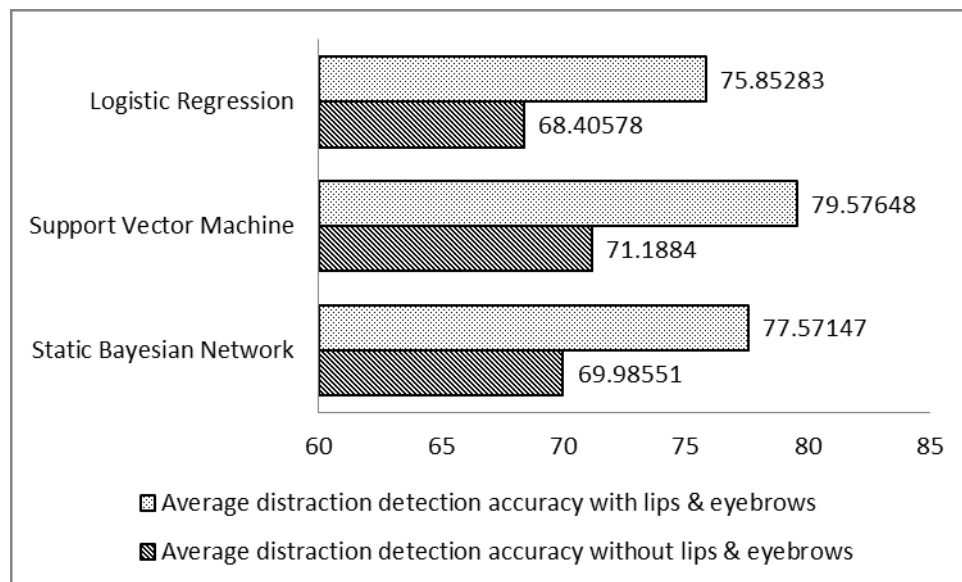
Figure 6.7: Accuracy of distraction detection comparison

It is clearly shown that, without lips and eyebrows, the cognitive distraction detection for the drivers are less accurate compared to the one with lips and eyebrows. Even when head rotation has been used and proven in many research studies [45][46][57] as one of the major features or parameters for cognitive distraction detection, it is still not good enough to give a good performance to detect cognitive distraction. [45] has mentioned, in order to improve the cognitive distraction detection performance, new parameters or features need to be added to the proven parameter. Cognitive distraction may occur rapidly and can steal driver's attention easily. One parameter alone does not reveal the distraction. However, by fusing a number of promising parameters, the robustness of the detection can be improved [46][57]. In the same subset of data, with WEKA, several more functions like TP rate (true positive), FP rate (false positive), F-measure (f-score), Precision and Recall have been computed from SVM, SBN, and Logistic regression.

TP rate calculates the true positive rate with respect to a particular class and this is defined as correctly classified positives divide with total positives [77]:

correctly classified positives
----
total positives

FP rate calculates the false positive rate with respect to a particular class and computed as incorrectly classified negatives divide with total negatives.

incorrectly classified negatives
----
total negatives

Precision is calculated as correctly classified positives divide with total predicted as positive.

correctly classified positives
--- -----------------------------
total predicted as positive

Recall is also equal to TP rate value and its calculation is the same as TP rate.

correctly classified positives
----
total positives

Finally, f-measure or f-score is calculated as:

$$fmeasure = 2\left[\frac{recall * precision}{recall + precision}\right]$$

Results for those functions are presented in the table below. The results are taken as average from every driver.

AVERAGE	COGNITIVE DETECTION with LIPS & EYEBROW		
	LOGISTIC	SVM	SBN
Accuracy %	75.90	79.60	77.60
TP rate	0.76	0.80	0.77
FP rate	0.24	0.20	0.23
F-Measure	0.75	0.80	0.77
Precision	0.75	0.80	0.77
Recall	0.76	0.80	0.77

Table 6.2: Cognitive detection results with lips and eyebrows

Classification with AdaBoost algorithm was excluded for this first computation, because the results produced were not reliable and not affirmative. Thus, to avoid any confusion, its results were removed. Results above have also been presented in the graph below:
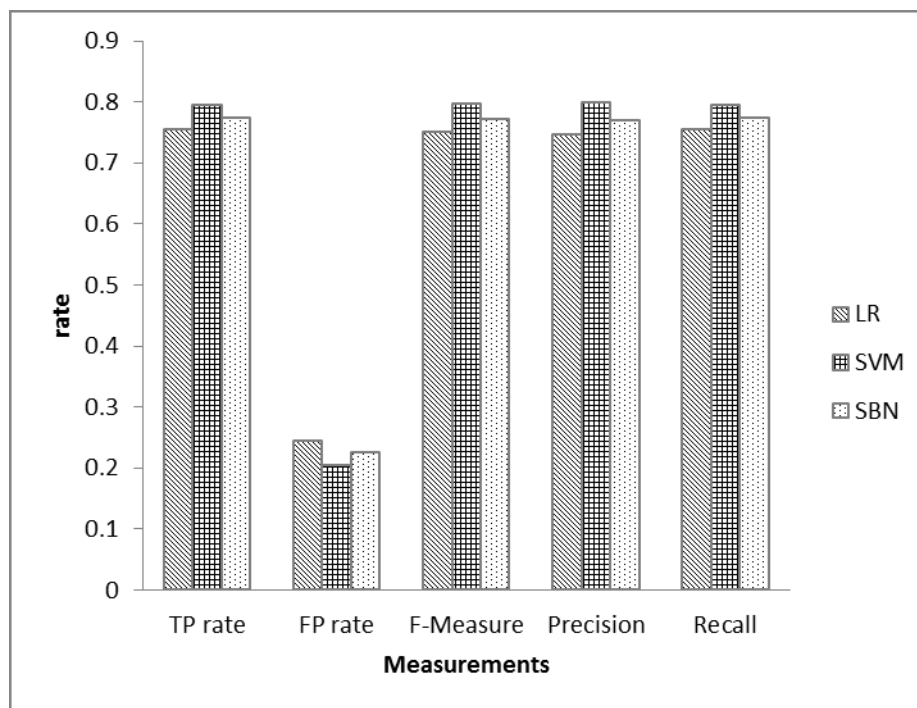
Figure 6.8: Functions comparison of Logistic Regression, SVM and SBN

From Figure 6.8 it can be seen that, SVM's TP rate, F-measure, Precision and Recall values are the highest among those three. Its FP rate is the lowest.

SVM's TP rate was 0.80, SBN was 0.77 and Logistic regression was 0.76. TP rate for SVM was higher by 0.03 from SBN and 0.04 from Logistic regression. False positive rate for SVM

whereas was the lowest with 0.20, followed by SBN with 0.23 and Logistic regression with 0.24. Precision values for SVM was 0.80, SBN was 0.77 and Logistic regression was 0.75. Recall is also similar to the TP rate values. F-measure or sometimes called F-score is related to precision and recall. Its values were 0.80 for SVM, 0.77 for SBN and 0.75 for Logistic regression. If seen from the results, values for TP rate, Precision, F-Measure and Recall were not really that much of a difference. From these three algorithms used, it can be seen that SVM is always the best to detect driver cognitive distraction. This is probably because SVM has a good generalization property and classification performance compared to the other two algorithms. SVM used for this study got even easier because the computation has been done for a binary classification, where it was only used to decide whether the driver is cognitively distracted or not cognitively distracted.

In this second part of analysis, the data was taken from both distracted/task and not distracted/control. The data file for the second part consisted of 50% from distracted data and 50% from not distracted data. The analysis was done with SVM, SBN, Logistic regression and also AdaBoost.

The whole data was also analysed to individual drivers and each result for every driver in every algorithm was presented in the table below. One third of the driver's data has been grouped for testing and two thirds as training data set.

DRIVER	In %			
	SVM	ADA	SBN	LR
1	84.20	88.20	76.80	77.40
2	86.50	76.50	77.10	72.80
3	81.40	81.10	79.30	73.60
4	81.70	77.20	78.80	79.00
5	81.70	73.10	79.10	75.10
6	74.10	76.90	79.10	73.10
7	78.80	82.00	79.20	75.60
8	79.90	71.70	68.00	77.20
9	76.70	80.70	79.30	76.50
10	70.70	79.70	79.00	78.30

Table 6.3: Accuracy rate in percentage on data from second experiment by individual drivers

From Table 6.3, on average from ten drivers, SVM is still the best classification algorithm with 79.58% followed by AdaBoost 78.72%, SBN with 77.57% and Logistic regression with 75.85%. Again, Logistic regression is the lowest algorithm to correctly calculate the classified instances from the data. The difference between SVM average of accuracy rate and AdaBoost average of accuracy rate was only 0.8607. All classification algorithms' results in this second experiment, including AdaBoost's result are better than from the first experiment. This is because the data collected from the second experiment was set up in a real time experiment with better cognitive loading tasks. The distraction is more compared to the first experiment.

Next part of analysis in this second experiment was done on another classification algorithm, Dynamic Bayesian Network. In the first experiment, the accuracy rate from DBN was done in a cumulative accuracy because the data collected was very limited and short. The cumulative accuracy rate for first experiment can be seen in Figure 6.5.

DBN model was developed by using a Bayesian Network Toolbox by Kevin Murphy [50] with a Matlab platform. As mentioned previously, 6 groups of face landmarks captured from faceAPI were used. Features used are inner lips, outer lips, right eyebrow, left eyebrow, head position and head rotation. These six groups of FLMs were used, computed and made as certain nodes in the DBN model and its network can be seen in Figure 6.9 below. From the six groups of FLMs, the network contained 9 nodes in the network. Distraction is the only hidden node in the network and it refers to whether the driver is distracted or not. The value for this distraction node is binary. Description about each node in the network can be found in Table 6.4.
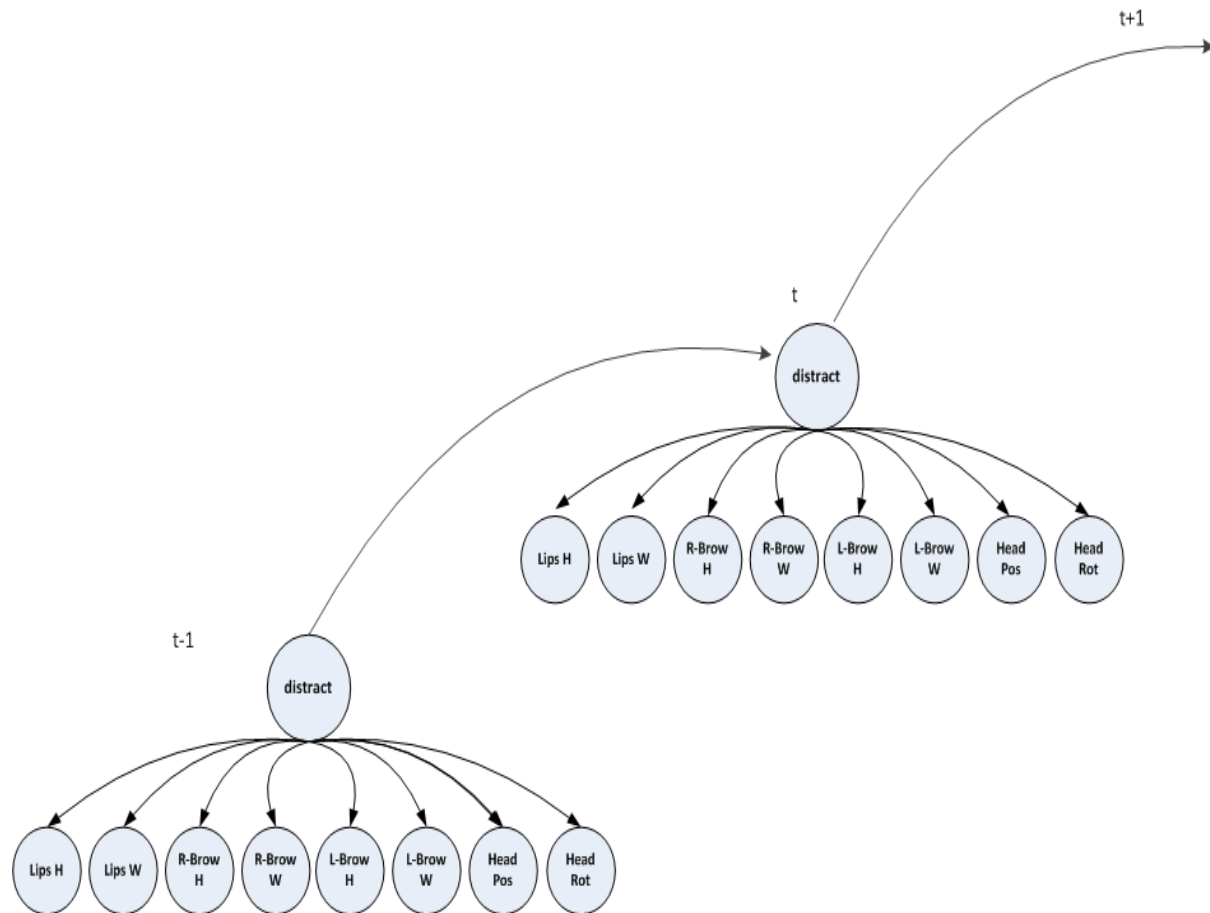
Figure 6.9: DBN network for second experiment

Notation	Description
t	time
distract	Distraction
HeadRot	Head Rotation
HeadPos	Head Position
R-Brow H	Right Eyebrow Height
R-Brow W	Right Eyebrow Width
L-Brow H	Left Eyebrow Height
L-Brow W	Left Eyebrow Width
Lips H	Lips Height
Lips W	Lips Width

Table 6.4: DBN model's (second experiment) notation and description

Data from all the ten drivers was used to train and test the proposed cognitive distraction detection method. Features used to develop this DBN model include inner and outer lips, right and left eyebrows, head position and head rotation which are similar to SVM, SBN and Logistic Regression features. The accuracy rate from each driver increases from the first driver to the next driver because the algorithm continues from the first driver to the 10th driver. Additionally, DBN model has prior and posterior knowledge which are used for the learning process. Prior knowledge is a knowledge that is gained prior to experience. It is sufficient to arrive at such knowledge by simply using the powers of reasoning. For instance, in cognitive distraction detection experiment in this study, prior knowledge occurs when a driver saw a hump/road bump signboard. Immediately, based on the driver's prior knowledge, the signboard tells the driver to be alert because the road bump is just upfront and it will cause the driver to slow down. The road bump creates a reason for the driver to slow down. A posterior knowledge is a knowledge which can only be experienced after facing certain experiences. Some observations are sometimes necessary to gain such knowledge. For instance, the driver for this study can only gain real experience of driving a manual Ford Mondeo car and share his/her knowledge about driving the car after he/she has completed his/her experiment participation. From the results in Table 6.5 below, it has been found that the accuracy rate is 93.62%. The accuracy rate is better than a study in [45] and another study in [46][57] especially when the experiment involved truck drivers. Both studies in [45] and [46][57] had not used DBN but they have used Support Vector Machine and AdaBoost machine learning algorithms. Clear comparison from these studies can be found in Table 6.13.

The threshold is set based on all of the data distribution. Threshold is set at 0.5 for marginal distribution computation. Since the data value is ranged from 0(minimum) and 1(maximum), thus 0.5 is the best separation value between being strongly or weakly distracted. The computation of true value based on the threshold value is shown in the sample code below [50]:

```
Threshold: % to calculate true value marginal prob. must > set threshold
if (marg.T(z)>threshold)
        if (evidence (1,i)==z)
                %z is referring to the type of distraction:
                1=not distracted, 2=distracted
        truevar=truevar + 1; %compute the true variable
        end
end
```

marg.T(z) = is the marginal distribution computed from the evidence

z = refers to the type of the distractions

i = refers to the time slice (T)

truevar = computes the true variable

Given the evidence from the parameter or feature used in this study, the marginal distribution is calculated. Marginal distribution is the probability distribution of the variable contained in the subset. In this case, the computation of marginal distribution is compared with the set threshold. If the value of the marginal distribution is bigger than the set threshold, its value will be computed as a true value, which later is considered as part of the accuracy percentage computation or successful rate.

Ground truth data was annotated during the experiment manually with the button 1 (distracted(and 0 (not distracted/control) and in the DBN algorithm as 2 and 1 respectively. The probability distribution computed by marginal distribution is grouped or labelled as either cognitively distracted (labelled as 2 in the DBN algorithm) or cognitively not distracted (labelled as 1 in the DBN algorithm). Thus, during the marginal distribution computation based on the given evidence, the probability is computed and the value will fall between 0.00 to 1.00. Thus, for instance if the value computed by the system/model is more than 0.5 (which is the best separation value) then the system is predicted the value as cognitively distracted. Later, the value will be checked with the ground truth data. If the ground truth data was also set as distracted thus it is considered as the true value. However, if the predicted value from the system is different with the ground truth data, thus it will give a false value.

In the DBN model shown in Figure 6.9, the distracted node refers to the parent node in the network. The node holds a binary value and it has been labelled as 1 or 2 for not distracted or distracted, respectively. If the marginal distribution computation is grouped as cognitively distracted (2), and the value is greater than the set threshold value, it is considered as a true positive. However, if the value of the marginal distribution is cognitively not distracted and the value is greater than the set threshold it is considered as false positive. This case is considered as a type 1 error in probability and statistic. The true value will only be computed if it is in the true positive event. Other cases are considered as a false positive event.

	ACCURACY OF COGNITIVE DISTRACTION DETECTION	
	Driver 1	72.50
	Driver 2	75.90
	Driver 3	78.00
	Driver 4	78.40
DRIVERS	Driver 5	80.20
	Driver 6	91.20
	Driver 7	92.50
	Driver 8	92.50
	Driver 9	93.30
	Driver 10	93.60

Table 6.5: DBN accuracy for second experiment

DBN was then compared with the other types of classification algorithms used before: SVM, SBN, and Logistic Regression. The results and comparisons made can be seen in Figure 6.10 below. Previously, SVM is the highest algorithm correctly classified driver cognitive distraction. SBN was the second and LR was the last. However, with DBN, the accuracy rate or correctly classified data was closer to 94%.
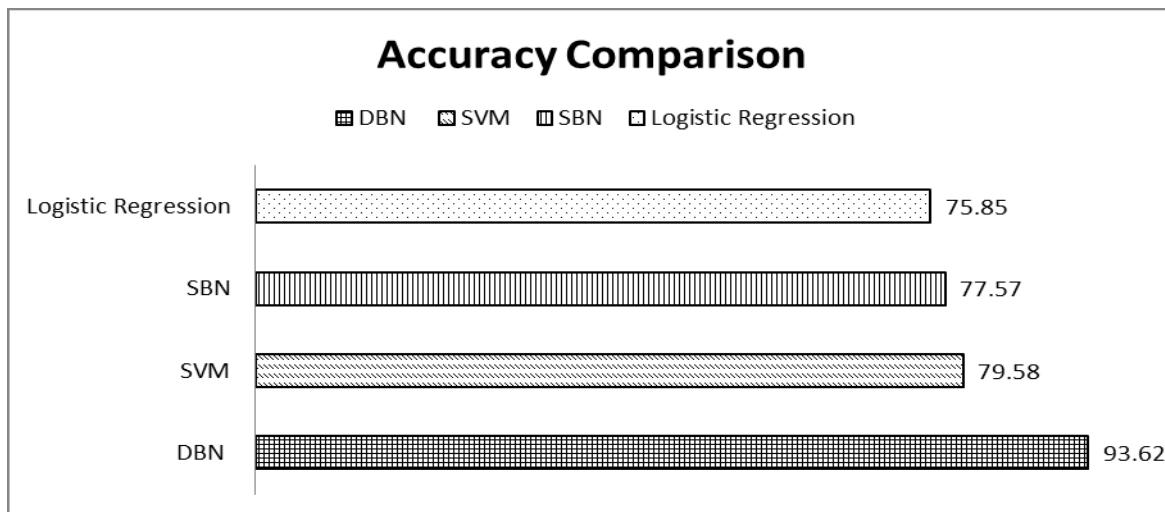
Figure 6.10: Comparison on every classification algorithms used in second experiment

Nevertheless, if to compare with SVM, SBN, and Logistic Regression, DBN is far better classification algorithm. DBN in this study is the best algorithm classified the data between distracted and not distracted because the data collected was in a real traffic and a real time

sequence were involved. Thus, a dynamic classification algorithm is the suitable and the best algorithm to classify the distractions.

Instead of computing TP-rate, Precision, Recall, FP-rate and F-measure like in the second part of analysis, the DBN algorithm was also used to compute sensitivity value. In information retrieval, sensitivity is also called as recall. The computation for sensitivity for each individual driver can be seen as in the table below. Sensitivity was also computed as in Table 6.6 below. Like in accuracy rate, the sensitivity average from the 10th driver is 98.69.

		SENSITIVITY OF COGNITIVE DISTRACTION DETECTION
DRIVERS	Driver 1	75.80
	Driver 2	82.00
	Driver 3	82.90
	Driver 4	83.60
	Driver 5	84.60
	Driver 6	85.90
	Driver 7	87.30
	Driver 8	90.60
	Driver 9	96.20
	Driver 10	98.70

Table 6.6: DBN sensitivity for 2nd experiment

From the three parts of analysis made in the second experiment, it can be concluded that, DBN is the best algorithm for classification followed by SVM. Logistic regression in any case is always the worst algorithm compared to the other four algorithms.

6.6 Classification Algorithm on Third Experiment Data

The third experiment is a fusion of a number of facial/head features captured by two systems: faceAPI and faceLab. Basically, there are four parameters from faceLAB:

a) Head rotation

b) Right gaze rotation

c) Left gaze rotation

d) Blinking frequency

and another six parameters from faceAPI:

a) Lips Height (from inner and outer lips)

b) Lips Width (from inner and outer lips)

c) Right eyebrow Height (from right eyebrow)

d) Right eyebrow Width (from right eyebrow)

e) Left eyebrow Height (from left eyebrow)

f) Left eyebrow Width (from left eyebrow)

were captured in this final experiment. 14 drivers were involved in the third experiment, therefore there are 28 data sets where 14 each are from the control data set and task data sets are collected from this experiment. Just like in the second experiment, inner lips and outer lips captured by faceAPI were used to compute lips height and width. Right eyebrows and left eyebrows landmarks are used to compute eyebrows width and height

Since, correlation has been made in both the first and second experiments, the third experiment focused only on the classification. For classification purposes, the same algorithms were used: SVM, SBN, AdaBoost, Logistic regression and last but not least DBN. Just like the second experiment, the analysis on SVM, SBN, AdaBoost and Logistic regression were computed for individual drivers.

TP rate, FP rate, Precision, Recall and F-measure were also computed. For DBN the analysis was done only for individual drivers. In DBN analysis, similarly to the second experiment, the algorithm was used to compute the accuracy rate or correctly classified percentage.

The relationships between the threshold level and the cognitive states are presented in the Table 6.7 below. The table below is important especially for DBN, the algorithm needs the information in the table to compute the accuracy and sensitivity from DBN.

	> threshold	< threshold
Distracted	Correctly Classified = Distracted > threshold	Incorrectly Classified = Distracted < threshold
Not Distracted	Incorrectly Classified = Not Distracted > threshold	Correctly Classified = Not Distracted < threshold

Table 6.7: Classification Relationship for Driver Cognitive Distraction Detection

The function relationships were also used for data analysis with WEKA for SVM, SBN, AdaBoost and Logistic regression.

First part of the analysis on the data from third experiment was done on SVM, SBN, AdaBoost and Logistic regression. Their correctly classified instances were computed in WEKA [77].

The analysis part was also done by individual drivers and the accuracy rates on the individual drivers were shown in the Table 6.8 below:

Accuracy rate by individual drivers (in %)				
DRIVER	SBN	SVM	ADA	LR
1	72.80	81.30	71.00	71.80
2	87.30	90.00	79.50	77.50
3	77.80	80.40	70.00	73.40
4	78.90	79.90	78.80	80.20
5	81.90	82.00	80.60	77.40
6	78.40	87.10	77.00	74.90
7	81.40	83.80	79.60	70.30
8	76.70	80.20	77.50	76.60
9	86.30	86.00	84.80	77.80
10	80.30	85.10	80.00	76.70
11	72.60	84.30	76.00	79.80
12	75.30	79.10	79.90	76.00
13	77.20	78.10	77.00	77.80
14	81.30	79.90	79.70	78.30
AVERAGE	79.20	82.70	78.00	76.30

Table 6.8: Accuracy rate on data from third experiment by individual drivers

On average, SVM still gives the best performance between those four algorithms. SVM average for accuracy rate was 82.6589%, SBN was 79.1546%, AdaBoost was 77.9682% and Logistic regression was 76.3182%. From the table above, almost every driver's accuracy rates are higher when using SVM algorithm. This showed that SVM is better to detect driver cognitive distraction when compared with logistic regression, static Bayesian Network and AdaBoost.

TP rate, FP rate, Precision, Recall and F-measure or F-score were also computed from the third experiment data sets. The differences between these computations were, in the second experiment, this computation was made only on the distracted data set and AdaBoost algorithm

was removed from the comparison with other algorithm because of its unreliable state. However, for the third experiment, every four algorithms were used and the data sets included distracted and not distracted data set.

Just like from the second experiment, TP rate, Precision, Recall and F-measures values for SVM are always of the highest values and the FP rate was the lowest one. The values for these functions can be found in the Table 6.9 below:

AVERAGE	Cognitive detection from third experiment			
	LOGISTIC	SBN	SVM	ADA
Accuracy %	76.32	79.15	82.66	77.96
TP rate	0.76	0.79	0.83	0.78
FP rate	0.24	0.21	0.17	0.22
F-Measure	0.76	0.80	0.81	0.79
Precision	0.76	0.80	0.83	0.79
Recall	0.76	0.79	0.83	0.78

Table 6.9: Data Mining results from third experiment

Since all values in the Table 6.9 were taken from the average values, the Accuracy %, TP-rate, F-measure, Precision and Recall values looked almost similar. The values were also presented in the graph as in the Figure 6.11 below:
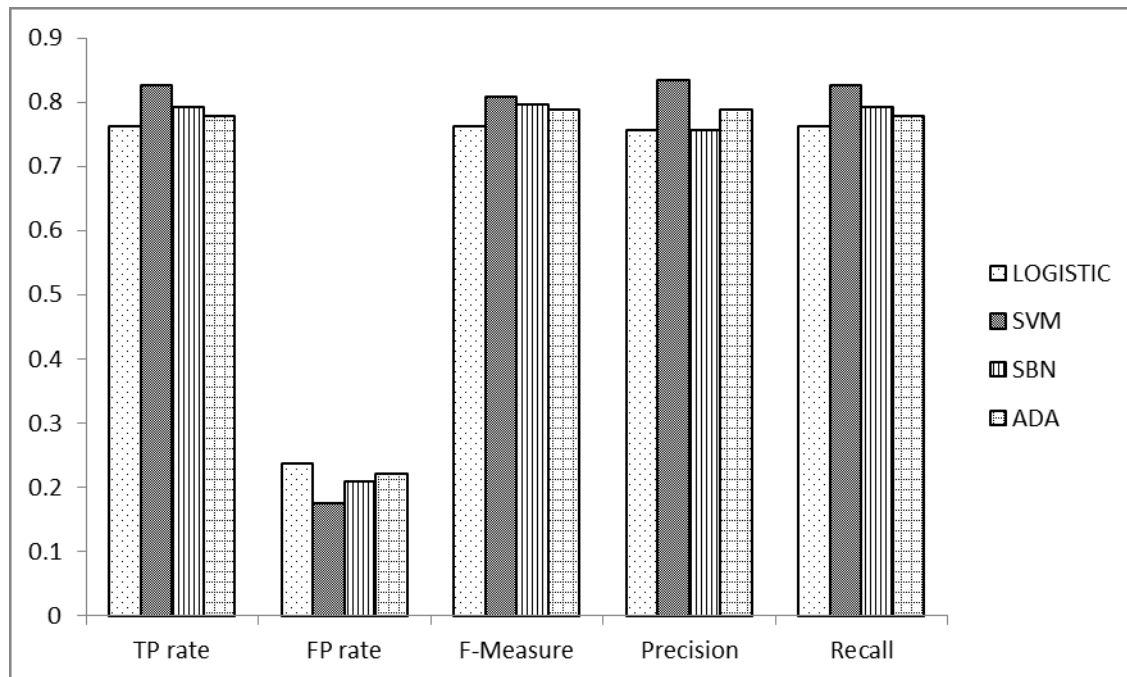
Figure 6.11: Functions comparison on SVM, SBN, AdaBoost and Logistic Regression

Further, the DBN model has been extended for the last experiment. Features from DBN model earlier were basically captured by faceAPI. Their movements can be captured by normal web camera. However, the features in the extended version model needs to be captured with infrared camera. The additional features captured by faceLAB machine are gaze rotation (left and right eyes) and blinking frequency. The extended model contained 11 nodes: distraction, blinking frequency, right eye gaze rotation, left eye gaze rotation, head rotation, lips height, lips width, right eyebrow height, right eyebrow width, left eyebrow height and left eyebrow width as in Figure 6.12.
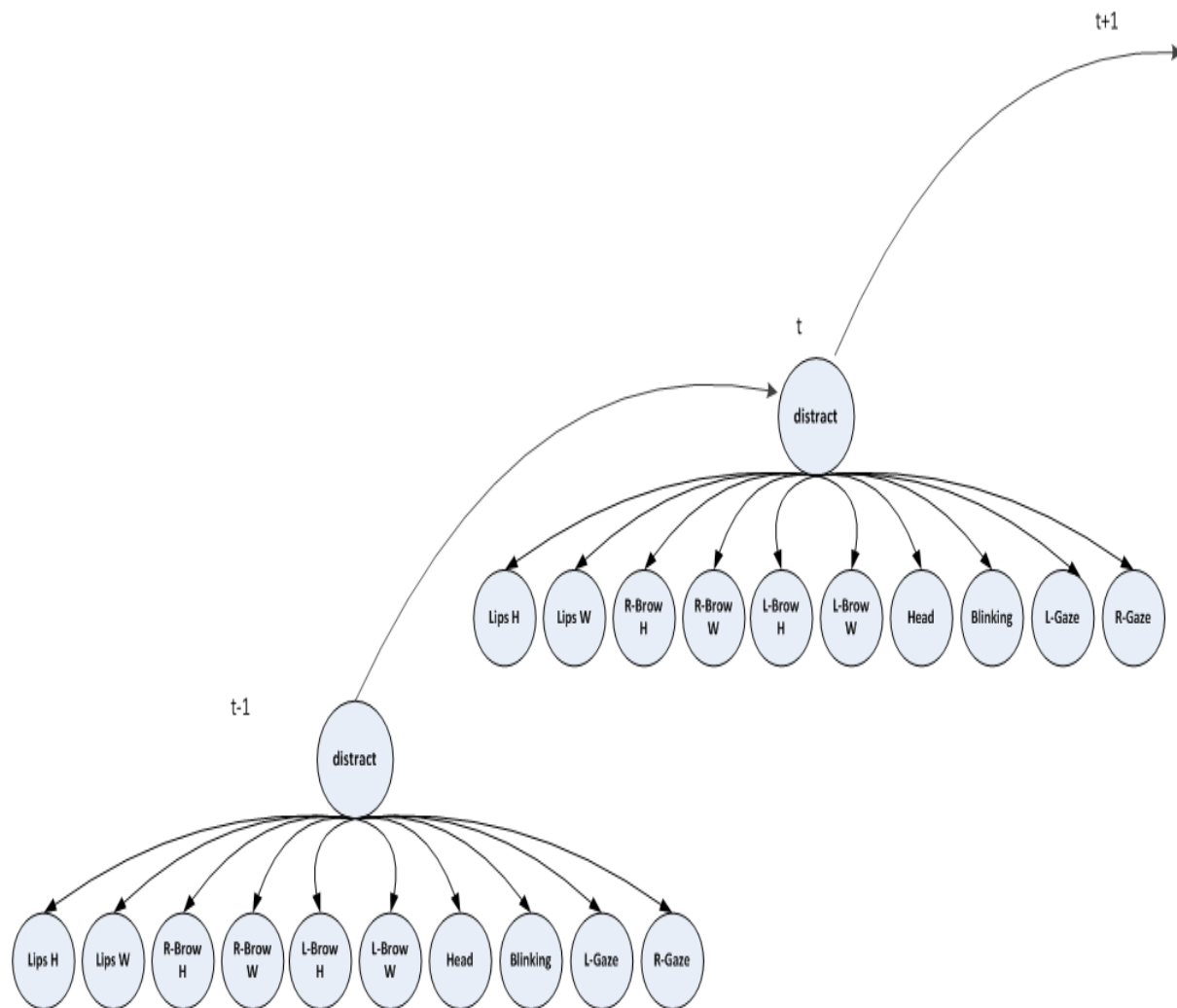
Figure 6.12: Extended DBN model with 11 nodes

From chapter 5, it has showed that lips and eyebrows are correlated to each other when a distraction is occurred. However the relationship between lips and eyebrows with other nodes in

the evidences are independence. This means that, when the lips or eyebrows is moving up or down, other nodes from the evidence is not move up or down respectively. This concept is also applied to the DBN model in third experiment.

Just like previous experiments, inner lip and outer lips information has been used to measure height and width of the lips. Height and width for eyebrows were also measured from Right and Left Eyebrows. Description for each node can be found as in Table 6.10 below:

Notation	Description
t	time
distract	Distraction
blink	Blinking Frequency
R-gaze	Right Eye Gaze Rotation
L-gaze	Left Eye Gaze Rotation
Head	Head Rotation
R-Brow H	Right Eyebrow Height
R-Brow W	Right Eyebrow Width
L-Brow H	Left Eyebrow Height
L-Brow W	Left Eyebrow Width
Lips H	Lips Height
Lips W	Lips Width

Table 6.10: DBN model's (final experiment) notation and description

From this new DBN model, head position has been removed because its information was not really useful to detect cognitive distraction because it only provided the position or coordinate of the head and did not give any movement value. In [45][46]and [57] they only used head rotation for detection.

Results in this final experiment were compared. Two groups of features are received as in Table 6.11. The first group imitated the features taken from the previous experiment where only lips, eyebrows and head were taken. The second group contained all the features captured from the combination of the machines: lips, eyebrows, head, blinking and gaze.

DRIVER	Accuracy rate comparison between 1st Group and 2nd Group	
	1st Group	*2nd Group*
	Lips, eyebrows & head	Lips, eyebrows, head, gaze & blinking
driver 1	75.60	81.70
driver 2	75.90	82.60
driver 3	76.40	83.00
driver 4	77.90	83.30
driver 5	78.30	83.90
driver 6	79.90	84.60
driver 7	80.10	84.90
driver 8	80.80	85.50
driver 9	81.00	86.70
driver 10	83.00	87.00
driver 11	83.50	89.50
driver 12	84.20	89.90
driver 13	85.70	90.60
driver 14	85.80	90.90

Table 6.11: DBN model accuracy for final experiment (1st and 2nd Groups)

Again, data from each driver is divided to 1/3 for the testing and 2/3 for the training. From the first group results, the accuracy rate is almost 85.80%. From the second group, the accuracy rate is 90.87% which is higher compared to 1st Group. Results received above also showed that, when the features were fused with other promising features (blinking and gaze), the accuracy rate is able to be enhanced. By comparing both groups as in Figure 6.13, when more reliable features are combined, the accuracy rate is boosted up to 90.87% from 85.8%. People said that quality is better than quantity. Even in this comparison, the quantity of features taken from the 2nd group is more than 1st group, but it also showed that this 2nd group is not only greater in quantity, but also

better in quality. Blinking frequency and gaze rotation is added up as additional features in the third experiment because those features are proven to be the best physiological measurement and had been used in some research papers like in [36], [45] and [46].
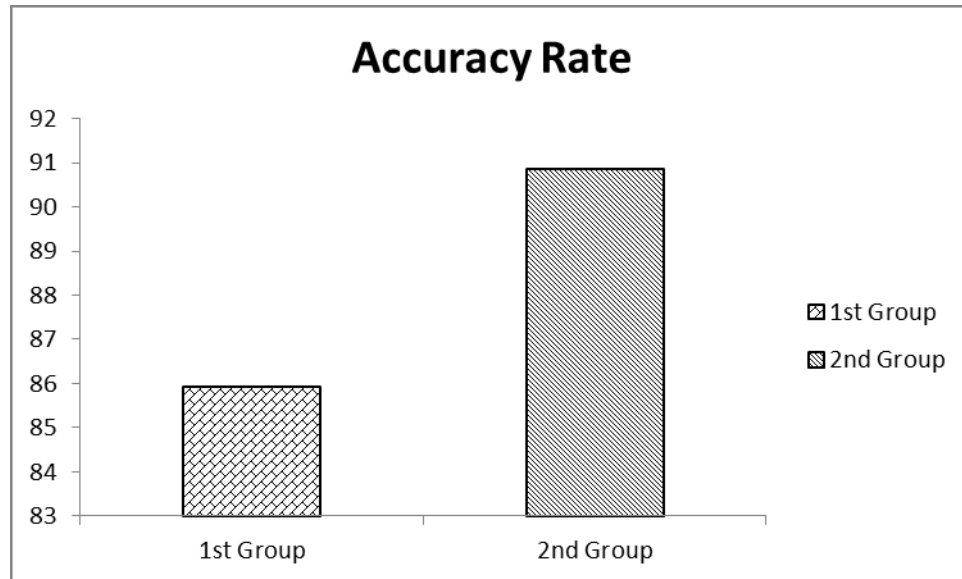
Figure 6.13: Accuracy comparison between 1st Group and 2nd Group

Whenever features from the first group combined with blinking and gaze rotation, the accuracy gets higher. Next a comparison was made between those features with lips and eyebrows and without lips and eyebrows as in Table 6.12 below. Group A contains Head Rotation, Blinking Frequency and Gaze Rotation whereas group B contains Lips, Eyebrows, Head Rotation, Blinking Frequency and Gaze Rotation.

DRIVER	Group A	Group B
	Head, gaze & blinking	Lips, eyebrows, head, gaze & blinking
driver 1	77.60	81.70
driver 2	78.80	82.60
driver 3	79.30	83.00
driver 4	79.70	83.30
driver 5	79.90	83.90
driver 6	80.10	84.60
driver 7	80.60	84.90
driver 8	80.90	85.50
driver 9	81.20	86.70
driver 10	82.00	87.00
driver 11	83.70	89.50
driver 12	84.60	89.90
driver 13	84.80	90.60
driver 14	85.90	90.90

Table 6.12: DBN model accuracy for final experiment (Group A and Group B)

From group A, the accuracy rate is 85.92%. This shows that as in Figure 6.14 below, with lips and eyebrows added in the cognitive distraction, the accuracy rate will be higher. With additional features lips and eyebrows, the accuracy rate is improved by 4.95%, from 85.92% to become 90.87%. This result indicates that when lips and eyebrows are not included in the cognitive distraction detection, the accuracy rate is lower. Table 6.11 and Table 6.12 present different purposes. Even when Table 6.11 basically compares lips and eyebrows which are combined with head rotation, the accuracy rate result is lower than when lips, eyebrows and head rotation are combined with blinking frequency and gaze rotation. Table 6.12 compares between the accuracy rate of when cognitive distraction is detected with lips and eyebrows and without lips and

eyebrows. Accuracy rate is not necessary depending on the number of sample/variable/feature/parameter used. It is more towards the number of reliable and useful feature/variable/sample/parameter to use.
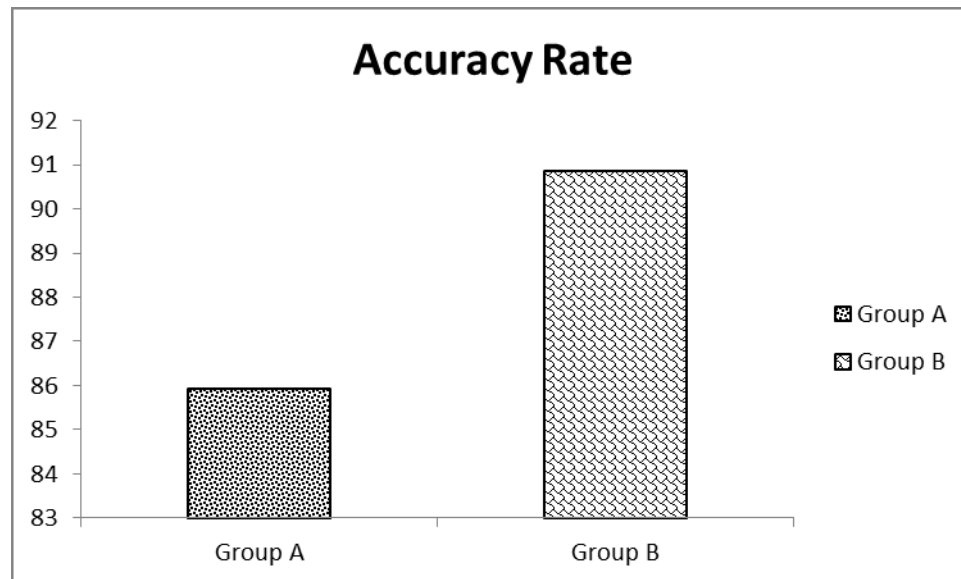
Figure 6.14: Accuracy comparison between Group A and Group B

Based on the comparisons made, it clearly shows that with lips and eyebrows added to blinking, gaze and head rotation, the performance of driver cognitive distraction detection can improve. Nevertheless, compared to SVM, SBN and LR, DBN is a better classification algorithm than the other three.

Finally in the third experiment a comparison was made between all algorithms used for driver cognitive distraction with all features captured: lips height, lips width, right eyebrow height, right eyebrow width, left eyebrow height, left eyebrow width, head rotation, blinking, right gaze rotation and left gaze rotation as seen in Figure 6.15 below:
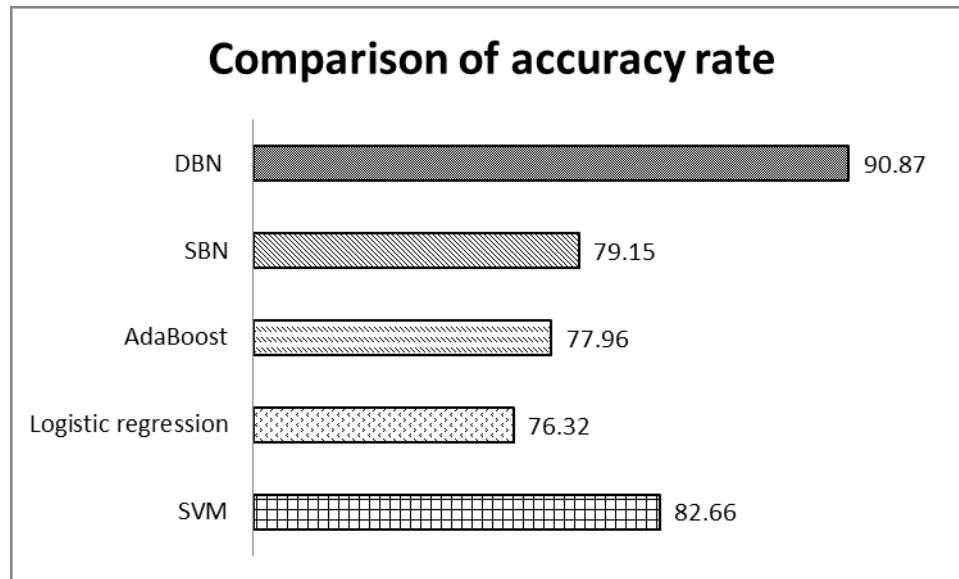
Figure 6.15: Comparison on all 11 nodes from all algorithms

From the graph above, it can be found again, that when lips and eyebrows were added into the detection system, the accuracy performance increased. Just like in the second experiment, the DBN in the final experiment has the highest accuracy rate with 90.87%. DBN is the best classification algorithm compared with other types of algorithm like SVM, Logistic regression, AdaBoost and Static Bayesian Network. The difference between DBN and SVM algorithm is 8.21% which relatively shows that DBN is better than SVM. DBN is an adaptive modelling considering information in temporal domain, where the prior knowledge present from previous learning is considered for future computation or posterior knowledge.

Figure 6.16 below shows a comparison between features in the third experiment. As been mentioned before, the first group in the third experiment captured lips height, lips width, right eyebrow height, right eyebrow width, left eyebrow height, left eyebrow width and head rotation. In the second group, the features used are lips height, lips width, right eyebrow height, right eyebrow width, left eyebrow height, left eyebrow width, head rotation, blinking frequency, left gaze rotation and right gaze rotation.

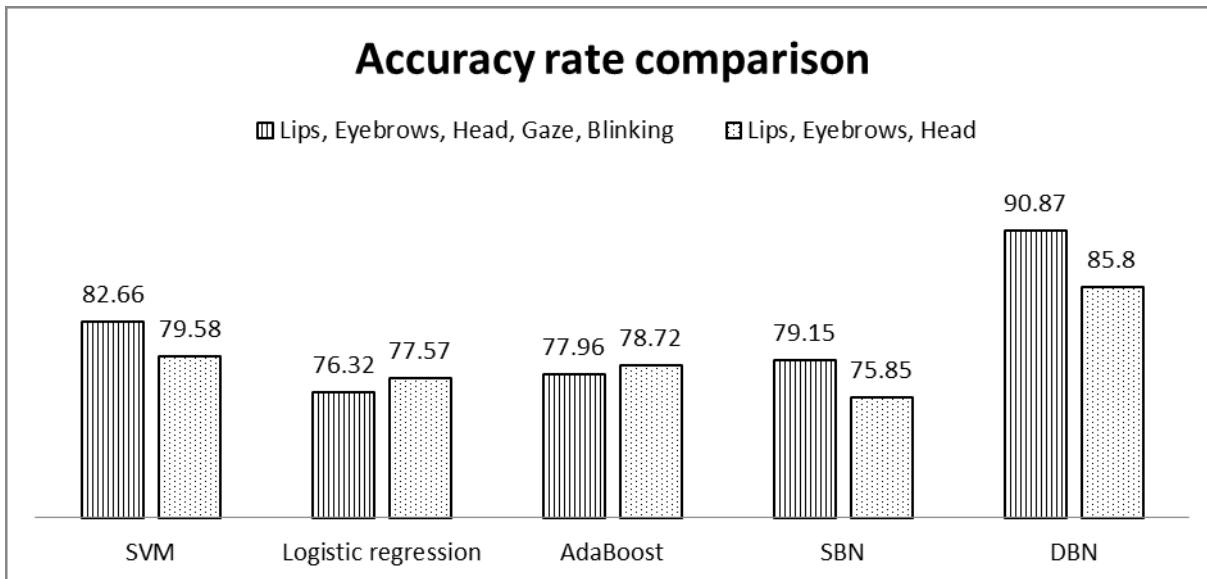Figure 6.16: Accuracy rate comparison between groups in the 3rd experiment

As can be seen in the figure above, almost every classification algorithm used for this thesis increased in a group where the features are lips, eyebrows, head, gaze and blinking. Only logistic regression and AdaBoost accuracy rates slightly decreased when all features are used. SVM, SBN and all DBNs increased. DBN showed the greatest improvement with 5.07%. SVM increased by 3.08%. SBN increased by 3.3%. Logistic regression decreased slightly by 1.25% and the same goes to AdaBoost with a decrease of 0.76%. This comparison shows that whenever lips and eyebrows were added to the distraction detection, the accuracy rate especially for Bayesian Network algorithm whether dynamic or static can really improve. Results analysis made by using Logistic regression and AdaBoost sometimes are unstable in this study. This might be a reason why Logistic Regression and AdaBoost algorithm can not be the best algorithm to detect cognitive distraction detection and probably not very good in computing accuracy rate for a dynamic real time event. LR is basically based on the linear regression and it is hardly to characterize the relationship between various features. For instance, in this study the relationship between distraction nodes with the evidences node is varied. Lips and eyebrows are correlated to each other, however the relationship between lips and eyebrows with other nodes like head rotation and blinking frequency is independence. AdaBoost is an algorithm with random sample based and thus its results may vary from one implementation to another especially when the amount of the data is large and quality of the data is varied significantly.

6.7   Comparison between Studies

Result from this thesis is also compared with results from other studies in the literature. First of all, setup for experiments from each research is shown in Table 6.13 below. It can be seen that not many researchers collect their data in a real time environment, where the driver is required to drive on a real road. A major benefit of collecting data from real environment compared to collecting data from lab is that the real environment data is more realistic and acceptable. This is because the data collected has to consider contextual information which could affect the data collection process. Table 6.13 below has shown that experiments conducted for this thesis are better compared to other researchers' experiments in terms of the number of participants and duration of time. Almost every researcher's experiment compared in this thesis used equipment produced by Seeing Machine Company.

Author/s	Parameters	Peripheral	Duration	Partici pant	Location
Yulan Liang, Michelle L. Reyes, John D. Lee [24][36]	• Blinking frequency • Fixation • Saccade • Smooth pursuit. • Steering wheel angle • Lane position • Steering error	faceLAB cameras and vehicle sensors	15 minutes	9	Driving simulator
Masahiro Miyaji, Haruki Kawanaka, Koji Oguri [45]	• Pupil diameter • Gaze rotation • Head movement • Heart rate	faceLAB and Electrocardio gram	8 minutes (3 min break)	8	Driving simulator
M H Kutila, M Jokela, J Viitanen, G Markkula, T W Victor [46][57]	• Gaze rotation • Head angle • Lane keeping	faceLAB and vehicle sensors	Approxi mately 45 minutes	12	Real Traffic
Afizan Azman, Qinggang Meng, Eran Edirisinghe [56]	• Mouth movement • Eye movement • Head rotation • Gaze rotation • Pupil diameter	faceLAB	17 minutes	6	Lab setting
Afizan Azman, Qinggang Meng, Eran Edirisinghe	• Head rotation • Head position • Eyebrow • Lips	faceAPI	46 minutes	10	Real Traffic
Afizan Azman, Qinggang Meng, Eran Edirisinghe	• Head rotation • Eyebrow • Lips • Blinking Frequency • Gaze rotation	faceAPI and faceLAB	50 minutes	14	Real Traffic

Table 6.13: Studies comparison on features and experiments' nature

From the table above, it can be seen that most of the existing research was done in a lab setting or done in a driving simulator. The types of physiological measurement used for each research were also different. However, for this research, only the best and reliable features are considered. Other types of measurement that were excluded are driving performance, pupil diameter and heart rate. Driver behaviour performance was not used in this research and some other researches because it can actually undermine the driving performance by mitigating driver's attention as mentioned in [33]. Driving performance will make the driver put their focus only on the performance measurement like keeping the lane in the middle of the road and keeping the right steering angle. This behaviour will make the driver put less focus on other factors in driving like the environment surrounding, pedestrian, and vehicle condition etc. Pupil diameter was used initially in this research. However, when the experiment has to be done in real traffic within a real environment, the data captured from pupil diameter was not applicable anymore and therefore unreliable for analysis. This is due to the exposure from sunlight which made the pupil diameter difficult to detect. Perhaps, the pupil diameter is only suitable for a lab environment where the light condition is constant and adjustable. Unfortunately, it is not suitable for a real environment setup. Thus, this feature was also excluded from this research. Heart rate measurement is rare and is an unsuitable type of measurement in a real traffic setup. This is because this type of measurement is intrusive and it obviously will mostly distract the driver [90] [47]. An intrusive measurement is not a preferable type of measurement by many researchers especially if the data collection needs to be done in real traffic.

Table 6.14 and Figure 6.17 below show the results comparison between this paper's third experiment with other researchers' results.

First Author	Learning Algorithm	% Accuracy rate
Yulan Liang	Dynamic Bayesian Network	86.40 [24]
Yulan Liang	Static Bayesian Network	78.00 [24]
Yulan Liang	Support Vector Machine	81.10 [36]
Yulan Liang	Logistic Regression	72.70 [36]
Masahiro Miyaji	AdaBoost	62.40-truck driver 86.00-commercial car driver [45]
Masahiro Miyaji	Support Vector Machine	72.40 [45]
M.H. Kutila	Support Vector Machine	68.00-truck driver 86.00-passenger car [46][57]
Afizan Azman	Dynamic Bayesian Network	90.87
Afizan Azman	Support Vector Machine	82.66
Afizan Azman	Logistic Regression	76.32
Afizan Azman	AdaBoost	77.96
Afizan Azman	Static Bayesian Network	79.15

Table 6.14: % Accuracy rate from different researches
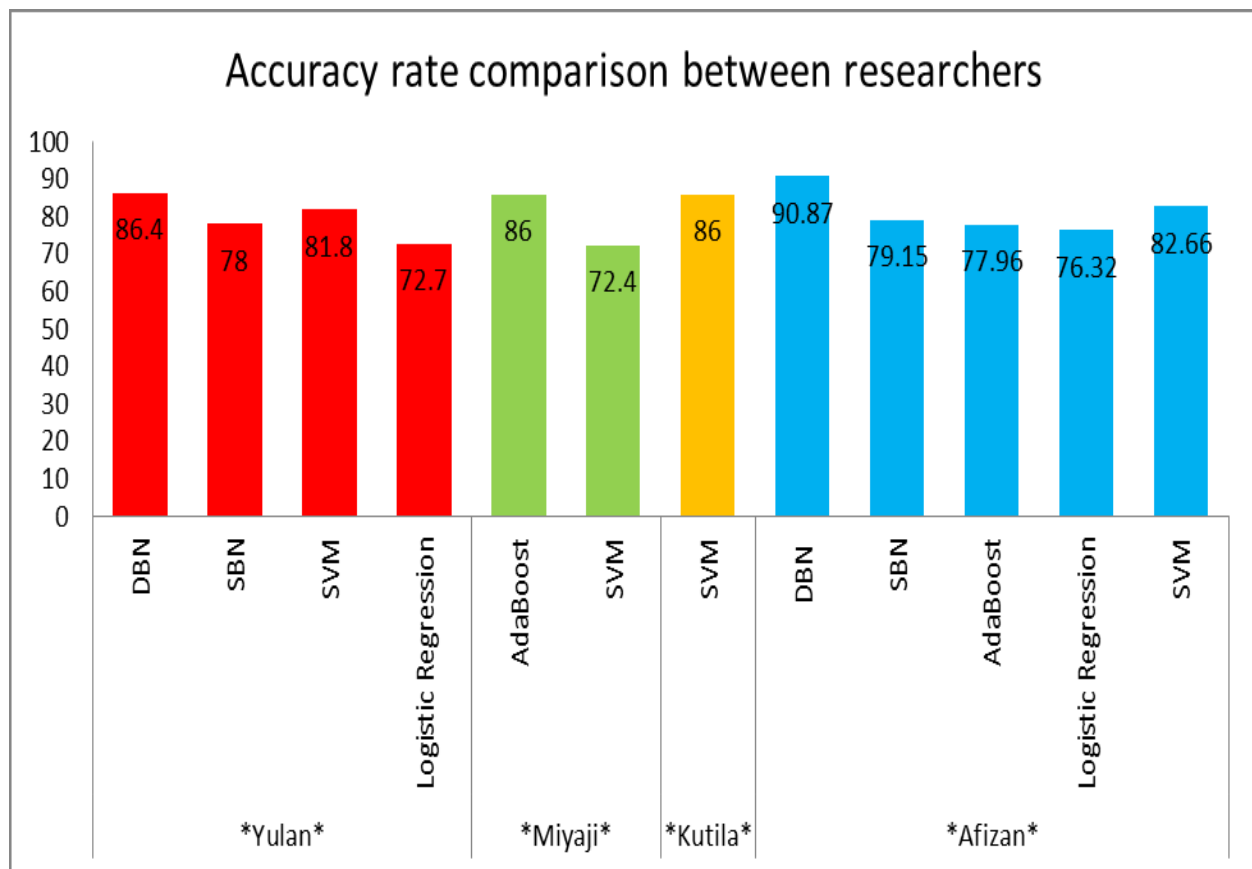
Figure 6.17: Results comparison based on different authors and algorithms

From Figure 6.17 above, results with DBN from Yulan as well as from this study show the best performance for detecting driver cognitive distraction. However, Yulan's experiment was not conducted in real traffic and she has combined physiological measurement with driving performance measurement. The second best algorithm is Support Vector Machine. Results from this study, Yulan and Kutila have shown respective indication that SVM is the second best after DBN. Kutila's experiment was conducted in a real traffic environment. However, his measurements were collected from physiological and driving performance types of measurement as in table 6.13. Static Bayesian Network is the third best algorithm for detecting driver cognitive distraction. Results received from different studies showed that different experiment set up, number of participants involved and algorithms used are important to ensure the best detection of driver cognitive distraction.

6.8 Conclusions

Different classification algorithms have been used in all experiments conducted in this research. DBNs provide a better and viable detection method for driver cognitive distraction. DBN models developed from this research are created based on the real traffic environment. Results comparison made between DBN and other traditional algorithms like SVM, Logistic Regression and Static Bayesian Network has showed that DBN outperforms them.

A number of comparisons have been made. Comparisons have shown that when lips and eyebrows information is fused into the detection system, it can improve accuracy. Comparisons have also been made between studies where a few research projects have been compared to this research. Result shows that DBN is still the best algorithm especially when the data is collected in a real traffic environment. Another finding is whenever lips and eyebrows were added into DBN it also outperformed other types of measurement made by other researchers.

Most inner lips in control are smaller movements than the outer because, when drivers are thinking, the inner lips are squeezed/puckered in. In a normal expression without cognitive distraction, the lips are stretched out. This is a situation where both lips are pressed together and pushed out. The information gathered from inner lips and outer lips are used to compute lips height and width.

During cognitive distraction, the right eyebrows are squeezed in. This makes the movements on right eyebrows smaller than during the control one. Middle lowered eyebrows- the middle of the eyebrows is pulled down so they slope inwards. The movement on eyebrows make FLMs move as well. The difference between initial position and moved position is used to compute the width and height for eyebrows.

One of the major contributions from this research is when lips and eyebrows are combined with head, gaze and blinking, it improves the average accuracy for driver cognitive distraction detection. This showed that lips and eyebrows are not just physically easy to detect but they are also capable in helping to improve the existing system in detecting driver cognitive distraction.

Chapter 7

Conclusions

This study proposed a system to detect driver cognitive distraction. New facial features were introduced to improve the performance of the system. Even though there were several features and parameters like eye movement, head rotation, and heart rate that had been used in previous work and study about driver's cognitive distraction detection, the gap in improving the measurement's performance is still wide open. A few algorithms were also used to analyze cognitive distraction detection data. However, not every algorithm provided a reliable result. This chapter contains two parts: (a) Summary and main contributions concluding important findings from every main chapters and (b) suggestions and future work to improve driver cognitive distraction detection in terms of features and parameters, algorithms and experimental setup.

7.1   Summary and main contributions

It has been proven that one facial feature alone does not clearly reveal the distraction. By fusing many features, the robustness of the cognitive distraction detection can be improved. Relationship between the lips and eyebrows are strongly related to each other when a driver is cognitively distracted. Data collected from this study was from real traffic environments. Thus, data collected was more reliable compared to data obtained from experiments run in lab. Three different experiments were conducted for different purposes. Five different algorithms were studied and implemented for the analysis part and every algorithm showed an improvement from the first experiment run to the final experiment.

Experimental setup details for each of the three experiments for this work have been explained in chapter 3. Different experiments brought in different objectives and purposes. First experiment

was basically to run a simple lab setup and it was purposely designed for checking the initial relationship and correlation between eye movement (height and width) with mouth movement (height and width). Data was captured by faceLAB machine. Results from this experiment were explained in chapter 4. The second experiment was created and designed for a real traffic environment. A different facial feature capture system called faceAPI from Seeing Machine's company was used in this second experiment. Features captured in this second experiment were: lips (height and width), eyebrows (height and width), head position and head rotation. Drivers for this experiment were required to drive an automatic transmission car using a similar route. Distraction setup, vehicle setup, drive details and route used for this experiment were discussed clearly in chapter 3. The third experiment was also run in a real traffic environment. However, in the final experiment both machines, faceLAB and faceAPI were installed together in the same vehicle. The data from both machines are then fused together in order to give better performance results. Features captured in this third experiment were: lips (height and width), eyebrows (height and width), head orientation, blinking frequency and gaze rotation. For this experiment, collaboration with Ergonomics and Safety Research Institute (ESRI) from Loughborough University was accomplished. Results from the data collection in the second and third experiment were basically discussed in chapter 6.

Chapter 4 in this work explains clearly about physiological measurement implemented for driver cognitive distraction detection. There are actually three basic measurements to detect a driver's cognitive distraction: primary and secondary task, rating scales and physiological measurement. However, the physiological measurement method was the only focus in this work. This is because, physiological measurement used for this work is non-intrusive and it did not undermine the drivers' driving performance, unlike primary and secondary task methods. This chapter has also explained in depth about Bayesian Networks. Its definition, types, concepts and examples were explained clearly. Two types of Bayesian Networks were used in this work: Static and Dynamic. These two models of Bayesian Networks from every experiment being run were presented in this chapter too. Since every experiment applied different parameters or features, the models of Bayesian Networks created for this work were also different. Algorithms for both DBN and SBN were also presented in this chapter.

Chapter 5 basically presented the results received from the correlation and linear regression analysis. These statistical modeling methods were used to determine the relationship between

eye movements and mouth movements and specifically between lips and eyebrows. It is found that the correlations between lips and eyebrows and between mouth and eye are in a strong relationship, mostly above 0.70. The results for this correlation and relationship check are presented in a number of graphs, tables and diagrams. They are found to have a positive relationship. Two sections of correlation and relationship analysis were done. The first section is for the first experiment where mouth movement (height and width) and eye movement (height and width) were used. The results were presented with Pearson-r correlation and scatter diagram. Mostly, every result is with strong and positive relationships to each other. The second section was to study the relationship between lips and eyebrows captured in the second experiment. However, for this section linear regression was used instead of Pearson-r. Since there are varieties of options available to check on correlation and relationship, different methods were used. Again, correlation coefficient received from linear regression was also above 0.7 which can be categorized as a strong relationship.

Chapter 6 covers classification algorithms for cognitive distraction detection. Five different algorithms: Support Vector Machine, AdaBoost, Logistic regression, Static Bayesian Network and Dynamic Bayesian Network are implemented. Since the first experiment objective was to check the correlation and relationship between parameters, only simple and easy classification analysis was done here. Furthermore, the data captured from the first experiment was very limited; therefore not much can be done with classification algorithm. For the second experiment, the classification algorithms were used to compare between data with lips and eyebrows and data without lips and eyebrows in SVM, SBN, AdaBoost and Logistic Regression. With SVM, data with lips and eyebrows was 79.58% correctly classified and data without lips and eyebrows was 71.19% correctly classified. Thus, it is clear that data pertaining to lips and eyebrows has higher classification accuracy (of the order of 8.39%) as compared to the data set without lips and eyebrows. Logistic regression had an accuracy of 75.85% with lips and eyebrows and 68.41% without lips and eyebrows. It can be seen that, LR is less accurate. Finally, SBN accuracy was 77.57% with lips and eyebrows and 69.99% without lips and eyebrows. Comparison on accuracy rate, sensitivity rate, precision, recall, f-measure, true positive rate and false positive rate are used. WEKA version 3.6 was used to get those results. Main algorithm in this work is the Dynamic Bayesian Network. From the second experiment results, DBN accuracy rate was 93.62%. SVM whereas was 79.58%, followed by SBN with 77.58%, and Logistic

Regression with 75.85%. This showed that DBN is the best algorithm when lips and eyebrows are used for cognitive distraction detection. The third experiment was done by combining features. Similar algorithms were used again for performance comparison. Basically, the results received from the algorithms implemented showed that the accuracy rates are mostly better than other existing studies made by other researchers. DBN has been found as the best algorithm compared to the other algorithms used in this work. Logistic regression always remains the weakest algorithm to detect driver cognitive distraction. Initial analysis on accuracy rate between SVM, SBN, AdaBoost and Logistic Regression has found that again SVM was the best algorithm between them with 82.66%, SBN with 79.15%, AdaBoost with 77.96% and Logistic regression with 76.32%. Similarly like the results from second experiment, the results received from the third experiment ranked the algorithms in a similar order like in the second experiment. DBN used in this third experiment has also shown an improvement in accuracy rate. Comparing DBN algorithm to compute accuracy rate, from group when only head, gaze rotation and blinking frequency were used, the average accuracy is 85.92%. However when lips and eyebrows were added into the measurement, the accuracy rate got higher. With lips and eyebrows the accuracy rate is 90.87%, improved by 4.95%. DBN has outperformed other algorithms including SVM. In final conclusion, it is shown from this research that lips and eyebrows are capable of improving the performance for driver cognitive distraction detection. With appropriate algorithms and experimental setup, the detection can reach its maximum performance.

7.2   Suggestions and Future Work

There are still a few numbers of improvements that can still be done to enhance driver cognitive distraction detection system.

In this research, only physiological measurements are used. In the future, this study can be improved by adding another type of cognitive distraction detection like driver behaviour performance. Steering angle, steering error rate and lane keeping are few examples of driver behaviour performance which can be added into the system for better improvement.

In terms of equipment, to add more features especially from driver behaviour performance, more cameras and sensors are required to be installed in the car. However, to ensure that data can be collected with appropriate amount, good cameras and sensors are required. Cameras and sensors must be capable to capture the measurement and movement in any weather type, temperature, time and road condition. Otherwise, it can affect the data collection process.

Other than that, the experiment setup can perhaps be improved by adding more participants and extending the duration of driving. By making the driving duration longer, it can help to improve the amount of data collected. More types of distraction can be created during the experiment. More participants will also help to enhance the amount of data to be collected. Perhaps, the participants involved in the future experiment can be from various backgrounds and ages. This might help researchers to find out which group of drivers could be more cognitively distracted while driving. Existing guidelines for experiment setup for driver's cognitive distraction is very limited. Many issues like tasks and distraction created to distract the driver is sometimes not normally occur in daily life as the distraction should be occur naturally rather than be created.

Algorithms used in this research have all been used by other researchers. Yet, DBN is the best algorithm to detect cognitive distraction. DBN is very suitable for a real time experiment because it can measure prior and posterior data. However, perhaps there might be other suitable and easier algorithms which can compete with DBN for this purpose. For instance, in [78] they have used a Long Short Term Memory Recurrent Neural Network for their data analysis. The experiment conducted for their experiment was also in a real time environment. However, the experiment was only conducted for a very short time, lasting not more than 10 minutes and was

only done in a straight route with light traffic. In terms of algorithm development, a post processing of the collected data which involved with data cleansing process, data preparation for testing and training, features selection are important to be studied and learnt before any algorithms can be used. Novel algorithms for classification like Dynamic SVM or AdaBoost with DBN are possible to be developed. Therefore the comparison on the algorithms can be more reliable and comparable between dynamic algorithms. Relationship between features selected for this study either with linear or non-linear relationship is also needed to be characterized properly. Scatter diagram used in this study only purposely used to show the positive relationship exists between features but not enough information to indicate the best fit relationship between those features. Thus, further exploration on data relationship for non-linear like exponential function, Gaussian function and quadratic function are possible to be included.

Thus from this study a proper development of driving assistance systems can be developed to assist driver who with cognitive distraction while driving. An appropriate warning system shall be developed in vehicle to warn driver with cognitive distraction, thus it can avoid the driver from being involved with any road crashes or accidents.

REFERENCES

[1]   Treat, J.R., Tumbas, N.S., McDonald, S.T., Shinar, D., Hume, R.D., Mayer, R.E., Stansifer, R.L., & Castellan, N.J. (1979). Tri-level study of the causes of traffic accidents: Executive summary. (Executive Summary No. (Contract No. DOT HS 034-3-535). DOT HS 805 099). Washington, DC: U.S: Department of Transportation, NHTSA.

[2]   Kelly, R. (2008). Absent minds take their toll No. A01867 08/08/08 A). Australia: Australia Insurance Company (AAMI).

[3]   Dinan, E. (2009). Driver inattention caused 3 car crash, say police [WWW page]. Available: www.seacostonline.com [May, 2011].

[4]   Road Users' Alliance. (2008). Road file, 2008/09. United Kingdom: Road Users' Alliance.

[5]   Concentration (2002) [WWW page]. Available: www.webster-dictionary.org/definition/concentration [2009].

[6]   Howard, A. (2009), UK road deaths reach record low. BBC NEWS [WWW page], Available: http://news.bbc.co.uk/1/hi/uk/8118341.stm [May, 2011].

[7]   Symonds, T. (2009). Calls for expert road crash teams-BBC News [WWW page]. Available: http://news.bbc.co.uk/2/hi/uk_news/8048971.stm [14th May, 2009].

[8]   Michael, P. & Mary D. S. (2007). Safety vehicle using adaptive interface technology (SAVE-IT) [WWW page]. Available: http://www.volpe.dot.gov/hf/roadway/saveit/index.html [2009].

[9]   Fritscher, L. (2009). Cognitive theory [WWW page]. Available: http://phobias.about.com/od/glossary/g/cognitivethedef.htm [Dec, 2010].

[10]   English saying: Saying and proverb in english. (2012) [WWW page]. Available: http://www.englishclub.com/ref/Sayings/ [April, 2012].

[11]   Sheridan, T. B. (2004). Driver distraction from a control theory perspective. Human Factors: The Journal of the Human Factors and Ergonomics Society Winter, Vol. 46, 587-599.

[12]   Bondurant, B. & Sanow, E. J. (2000). Bob Bondurant on police and pursuit driving (illustrated, revised ed.) MotorBooks/MBI Publishing Company [WWW page]. Available:

http://books.google.co.uk/books?id=AQGrUvif5OwC&dq=hands+on+the+steering+wheel &q=steering+wheel#v=snippet&q=steering%20wheel&f=false [Dec, 2010].

[13] Baumann, M. & Krems, J. F. (2007). Conceptual Framework and Modelling Architectures. In Pietro C. Cacciabue (Eds.), Modelling driver behaviour in automotive environments: Critical issues in driver interactions with intelligent transport systems (pp. 59-121). London, UK: Springer Ltd.

[14] Hyderabad Traffic Police. Distractions while driving (n.d.) [WWW page]. Available: http://htp.gov.in/DrivingDistractions.html [6th Aug, 2009].

[15] Howard, A. (2005). We could be doing 80 on the motorway (Life & Style, Motoring) [WWW page]. Available:
http://www.independent.co.uk/life-style/motoring/comment/andrew-howard-we-could-be-doing-80-on-the-motorway-492589.html [July, 2009].

[16] Stutts, J. C. , Reinfurt, D. W. , Staplin, L. & Rodgman, E. A. (2001).
The Role of Driver Distraction in Traffic Crashes - No. (202/638-5944). New York: AAA Foundation for Traffic Safety.

[17] Victor, T. (2000). A technical platform for driver inattention research. Volvo technical report for project NUTEK Dnr 1P21-99-4131. Goteborg, Sweden: Volvo.

[18] Institute for Research in Public Safety. (1973). Tri-level study of the causes of traffic accidents:Interim report No. 1. (DOT-H S-034-3-535-73-TAC). Washington D.C. US: U.S. Department of Transportation National Highway Traffic Safety Administration.

[19] Alcohol and driving facts: Driving over the limit (n.d) [WWW page]. Available: http://www.keepthedoctoraway.co.uk/ [August, 2009].

[20] Institute of Alcohol Studies (2009). Drinking & Driving. United Kingdom: IAS Factsheet.

[21] The Royal Society for the Prevention of Accidents (RoSPA) (2009). Driver fatigue and road accidents. [WWW page]. Available:
http://www.rospa.com/roadsafety/advice/driving/fatigue_facts.htm [2011].

[22] Horne, J.A. & Reyner, L.A. (1995) Sleep related vehicle accidents. British Medical Journal, 310, 565-567.

[23] Klauer, S. G. , Neale, V. L., Dingus, T. A., Ramsey, D., Sudweeks, J. (2005). Driver Inattention: A Contributing Factor to Crashes and Near-Crashes. Human Factors and Ergonomics Society Annual Meeting Proceedings, Surface Transportation, 5, 1922-1926.

[24] Liang, Y., Lee, J. D. & Reyes, M. L. (2007). Nonintrusive detection of driver cognitive distraction in real time using Bayesian networks. Transportation Research Record: Journal of the Transportation Research Board, No. 2018, 1-8.

[25] Hole, G. (2007). The psychology of driving. New Jersey, US: Lawrence Erlbaum Associates, Inc.

[26] De Waard, D. (1996). The measurement of drivers' mental workload. PhD thesis, Traffic Research Centre VSC, University of Groningen, The Netherlands.

[27] Horrey, W.J., Wickens, C.D., & Consalus, K.P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. Journal of Experimental Psychology: Applied, 12(2), 67-78.

[28] Strayer, D. L., Drews, F. A. & Johnston, W. A. (2003). Cell phone induced failures of visual attention during simulated driving. Journal of Experimental Psychology: Applied, 9(1), 23-32.

[29] Elslande P. V. & Fouquet, K. (May 2007). Analyzing 'human functional failures' in road accidents. TRACE report D5.1. US: Traffic Accident Causation in Europe & Information Society Technology.

[30] Rumar, K. (1985). The role of perceptual and cognitive filters in observed behavior. In Evans L. & Schwing R.C. (Eds.), Human Behavior and Traffic Safety (pp. 151-170). New York: Plenum Press.

[31] Brown, I. (February 2006). Review of the 'looked but failed to see' accident causation factor Report No. (60). UK: Department for Transport.

[32] Liu, Y. C. (2001). Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. Egonomics, 44(4), 425-442.

[33] Salvucci, D. D. , Boer, E. R. & Liu, A. (2001). Toward an integrated model of driver behavior in a cognitive architecture. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1779, 1-25.

[34] Embrey, D. , Blackett, C. , Marsden, P. , & Peachey, J. (2006). Development of a human cognitive workload assessment tool. Dalton, Lancashire: Human Reliability Associates.

[35] DiDomenico, A. T. (2003). An investigation on subjective assessments of workload and postural stability under conditions of joint mental and physical demands. PhD thesis, Faculty of the Virginia Polytechnic Institute and State University.

[36] Liang Y., Reyes M.L. and Lee J.D. (2007). Real-time Detection of Driver Cognitive Distraction Using Support Vector Machines. In Proc. IEEE Transactions on Intelligent Transportation Systems, Vol. 8 No. 2, 340-350.

[37] Stern, E. (1999). Reactions to congestion under time pressure. Transportation Research Part C: Emerging Technologies, Vol. 7(2–3), 75-90.

[38] Rongben, W., Lie, G., Bingliang, T. & Lisbeng, J. (2004). Monitoring mouth movement for driver fatigue or distraction with one camera. In: The 7th IEEE Intelligent Transportation Systems Conference, Washington, USA, 314-319.

[39] Krahmer, E., Ruttkay, Z. & Swerts, M. (2002). Pitch, eyebrows and the perception of focus. In: Speech Prosody 2002, Aix-en-Provence, France, 443-446.

[40] Azman, A., Meng, Q., Edirisinghe, E. & Azman, H. (September 2011). Non-intrusive physiological measurement for driver cognitive distraction: Eye and mouth movement. International Journal of Advanced Computer Science, Vol. 1 No. 3, 92-99.

[41] Straker, D. (2010). Changing minds: In detail (2nd Edition). UK: Syque Press.

[42] Matsumoto, D., Ekkmann, P. (2008). Facial expression. In: C. Blakemore & S. Jennett (Eds). Oxford Companion to the Body. London: Oxford University Press.

[43] Eby W. D. (2005). SAfety VEhicles using adaptive interface technology (SAVE-IT): A program overview. In: International Distracted Driving Conference, Canada, 2-5.

[44] Lu. S., Olde, B., Cooper, E. & Graesser, A. C. (2001). Individual differences in reasoning about broken devices: An eye tracking study. In: The 23rd Annual Cognitive Science Society Conference, Scotland, 582-587.

[45] Miyaji, M., Danno, M., Kawanaka, H. & Oguri, K. (2008). Driver's cognitive distraction detection using AdaBoost on pattern recognition basis. In: Proceedings of the 2008 IEEE International Conference on Vehicular Electronics and Safety, Columbus, OH, USA, 51-56.

[46] Kutila, M. H., Jokela, M., Ma¨kinen, T., Viitanen, J., Markkula, G. & Victor, T. W. (2007). Driver cognitive distraction detection: Feature estimation and implementation. IMechE Part D: Journal of Automobile Engineering, Vol. 220, 1027-1040.

[47] Azman, A., Meng, Q. & Edirisinghe, E. (2010). Non intrusive physiological measurement for driver cognitive distraction detection: Eye and mouth movements. In: 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), Chengdu, China, 595-599.

[48] Marinova, M. , Devereaux, J. , Hansman, R. J (n.d). Experimental studies of driver cognitive distraction caused by cell phopne use [WWW page]. Available: from http://ttt.media.mit.edu/impact/marinovadeverauxhans.pdf [Dec, 2011].

[49] Bayes, T. & Price, R. (1763). An essay towards solving a problem in the doctrine of chances. UK: Philosophical Transactions of the Royal Society of London.

[50] Murphy, K. (2005, October 31st). Software packages for graphical models / bayesian networks [WWW page]. Available: http://www.cs.ubc.ca/~murphyk/Bayes/bnsoft.html [July, 2009].

[51] Ikehara, C. S. & Crosby, M. E. (2005). Assessing cognitive load with physiological sensors. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Hawaii, 295-305.

[52] Wierwille, W. W., & Tijerina, L. (1998). Modeling the relationship between driver invehicle visual demands and accident occurrence. In: The 6th International Conference on Vision in Vehicles, Derby, UK, 233-243.

[53] Recarte, M. A. & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. Journal of Experimental Psychology: Applied, Vol. 6(1), 31-43.

[54] Waltz, E. L. (1998). Information understanding: Integrating data fusion and data mining process. In: Proceedings of the 1998 IEEE International Symposium on Circuits and Systems ISCAS '98, US, 553-556.

[55] Mitchell, T. M. (1997). In Eris M. Munson (Eds.), Machine learning (International Editions). Singapore: McGraw Hill Companies, Inc.

[56] Azman, A., Meng, Q. & Edirisinghe, E. (2010). Correlation between eye movements and mouth movements to detect driver cognitive distraction. In: Brain Inspired Cognitive Systems 2010 Conference, Madrid, Spain, 1-12.

[57] Kutila, M., Jokela, M., Markkula, G. & Rue, M. R. (2007). Driver distraction detection with a camera vision system. In: IEEE International Conference on Image Processing, 2007, ICIP 2007, US, 201- 204.

[58] Guhe, M., Liao, W., Zhu, Z., Ji, Q., Gray, W. D. & Schoelles, M. J. (2005). Non-intrusive measurement of workload in real-time. In: Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society, Florida, US, 1157-1161.

[59] Pearl, J. (1988). Probabilistic reasoning in intelligent systems (1 edition). US: Morgan Kaufmann Inc.

[60] Ghahramani, Z. (1998). Learning dynamic Bayesian networks. Adaptive Processing of Sequences and Data Structures, 168-197.

[61] Ji, Q., Lan, P., & Looney, C. (2006). A probabilistic framework for modeling and real-time monitoring human fatigue. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 36(5), 862-875.

[62] Montgomery, D. C., Peck, E. A. & Vining, G. G. (April 2001). Introduction to linear regression analysis (3rd edition). US: Wiley-Interscience.

[63] Freund, R. J., Wilson, W. J., & Sa, P. (2006). Regression analysis. London: Academic Press.

[64] Weiers, R. M. (2010). Introduction to business statistics (7th ed.) US: Cengage Learning.

[65] Correlation coefficient. (2012) [WWW page]. Available: http://mathbits.com/MathBits/TIsection/Statistics2/correlation.htm [February, 2012].

[66] Leech, N., Barrett, K. & Morgan, G. A. (2007). SPSS for intermediate statistics. US: Lawrence Erlbaum Associates Inc.

[67] faceLab and faceAPI from seeing machines. (2009) [WWW page]. Available: http://www.seeingmachines.com/product/facelab/ [August, 2009].

[68] Bandler, R. & Grinder, J. (1979). Frogs into princes: Neuro linguistic programming. UK: Real People Press.

[69] Face API: The real time face tracking toolkit for developers and OEMs. (2008) [WWW page]. Available: http://www.seeingmachines.com/pdfs/brochures/faceAPI-Brochure.pdf [August, 2009].

[70] Hyndman, R. J. & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting. Vol. 22(4), 679-688.

[71] Armstrong, J. S. & Collopy, F. (1990). Error measures for generalizing about forecasting methods, Empirical comparisons. US: Wharton School, University of Pennsylvania, Marketing Department.

[72] Anderson, J. R. (2007). How can the human mind occur in the physical universe? (1st edition). US: Oxford University Press.

[73] Zhang,Y., Owechko, Y., & Zhang, J. (2004). Driver cognitive workload estimation: A data-driven perspective. In: Proceedings for the 7th International IEEE Conference on Intelligent Transportation Systems, US, 642-647.

[74] Fletcher, T. (2009). Support vector machine, explained. Retrieved March [WWW page}. Available: http://www.tristanfletcher.co.uk/SVM%20Explained.pdf [2012].

[75] Schapire, R. E. & Freund, Y. (2012). Boosting. Cambridge, US: The MIT Press.

[76] Agresti, A. (2007). An introduction to categorical data analysis (2nd illustrated edition). US: John Wiley & Sons.

[77] Hall, M. , Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter, 11(1), 10-18.

[78] Wöllmer, M., Blaschke, C. , Schindl, T. , Schuller, B., Färber, B. Mayer, S. & Trefflich, B. (2011). Online driver distraction using long short term memory. IEEE Journal of Transactions on Intelligent Transportation Systems, Vol. 12 No.2, 574-582.

[79] What's the difference between a confidence interval and a credible interval? (2012) [WWW page]. Available: http://stats.stackexchange.com/questions/2272/whats-the-difference-between-a-confidence-interval-and-a-credible-interval [July, 2012].

[80] Shetty, S., Song, M., Yang, H. & Matthews, L. (2009). Learning Distributed Bayesian Network Structure Using Majority-based Method. Journal of Computational Methods in Sciences and Engineering, Vol. 9 Issue 1, 53-68.

List of Publications:

Conference Publications-

1) Azman, A. and Meng, Q., "Correlation between Eye Movements and Mouth Movements to Detect Driver Cognitive Distraction", Brain-Inspired Cognitive Systems, BICS2010, Madrid, Spain, July 2010, ISBN: 978-84-614-1870-1. (Accepted & Published)

2) Azman, A., Meng, Q., Ahmad, A. and Jayothisa, C.M., "An emphirical study to examine sex differences in cognitive distraction among drivers", Proc. of the Int. Conf. on Advanced Computer Theory and Engineering (ICATE), 3, ICATE (2010), Chengdu, China, August 2010, pp.585-589, ISBN:978-1-4244-6539-2. DOI: 10.1109/ICATE.2010.5579555. (Accepted & Published)

3) Azman, A., Meng, Q. and Edirisinghe, E.A. , "Non-Intrusive Physiological Measurement for Driver Cognitive Distraction Detection Eye and Mouth Movement", Proc. of the Int. Conf. On Advanced Computer Theory and Engineering (ICACTE), IEEE, Chengdu, China, August 2010, pp.595-599, ISBN: 978-1-4244-6540-8. (Accepted & Published)

4) Azman, A., Meng, Q., Edirisinghe, E.A. and Azman, H., "Eye and Mouth Movements Extraction for Driver Cognitive Distraction Detection", IEEE Business Engineering and Industrial Applications Colloquium 2012, 7-8 April 2012. (Accepted)

Journal Publication-

1) Azman, A., Meng, Q., Edirisinghe, E.A. and Azman, H., "Non-Intrusive Physiological Measurement for Driver Cognitive Distraction Detection Eye and Mouth Movements", International Journal of Advanced Computer Science (IJACSci), 1(3), August 2011, pp.92-99, ISSN: 2251-6379. (Accepted & Published)

2) Azman, A., Meng, Q., Edirisinghe, E.A., Hassan, A.A., "Lips, Eyebrows and Driver Cognition Distraction", International Journal of Transportation Research Part C: Emerging Technologies, Elsevier, ISSN: 0968-090X. (To be submitted)