UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
INGENIERÍA MECÁNICA – CAMPO DISCIPLINARIO: MECATRÓNICA

"DISEÑO DE UN SISTEMA DE SEGURIDAD AUTOMOTRIZ"
("DESIGN OF AN AUTOMOTIVE SAFETY SYSTEM")

TESIS QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN INGENIERÍA

PRESENTA:
JORGE LUIS PARDO GAYTÁN

TUTOR PRINCIPAL
DR. ALEJANDRO C. RAMÍREZ REIVICH

MÉXICO, CD.MX. ENERO 2019

JURADO ASIGNADO:

Presidente: Dr. López Parra Marcelo

Secretario: Dr. Borja Ramírez Vicente

Vocal: Dr. Ramírez Reivich Alejandro C.

1 er. Suplente: Dra. Corona Lira María del Pilar

2 do. Suplente: M.I. Velázquez Silva Anahí

Lugares donde se realizó la tesis: Ciudad Universitaria, Cd. Mx., México

TUTOR DE TESIS:

DR. ALEJANDRO C. RAMÍREZ REIVICH

FIRMA

Contents

Abstract 7

Acknowledgements 8

I. First Stage: GlobalDrive 10

1. Introduction 11

1.1. Problem Description 11

1.2. Safety Cloud Objectives 11

2. Background 12

2.1. Definitions 12

2.2. Advanced Driver-Assistance Systems 14

2.3. State of the Art and Benchmarking 18

2.3.1. Existing ADAS and safety-related Apps for Cars 19

2.3.2. State of the Art in Motorcycle's Safety. 21

2.3.3. Augmented Reality and Projection Systems 23

2.4. Accident Statistics and survey 25

2.5. Artificial Intelligence 32

2.5.1. Background 32

2.5.2. Machine Learning and Deep Learning 34

2.5.3. Deep Learning 36

2.5.4. Deep Learning Libraries 54

2.5.5. Cloud Services 54

2.5.6. Machine Learning and Deep Learning in the automotive industry 55

3. Concept Development Process 58

4. Concepts 59

5. Final Concept 60

II. Second Stage: Igniting the spark with AI. 62

6. Data & Deep Learning Models 64

6.1. Datasets 64

6.2. Tests & Results 64

III. Third Stage: UX between human and AI 74

7. **Study Description** 76

8. **Future Work** 78

8.1. Future work for Stage I and II 78

8.2. Future work for Stage III 78

9. **Discussion** 80

9.1. Technology 80

9.2. Regulations and Standards 80

9.3. Test and Validation 81

9.4. Human Factors 81

9.5. Thesis Outcome and Conclusion 83

References 87

Appendix 104

Appendix A – Benchmarking Description. 104

External Devices 104

Existing Smartphone Applications 106

Abstract

The Self-Driving car has been attracting a great deal of attention in recent years, thrust by huge developments from major car companies. The design, implementation, testing, and deployment of self-driving features are changing road safety and will help to create a worldwide safer mobility environment in the next years. Unfortunately, the current price of these systems is really high, and only the premium cars have the State of the Art safety systems. Therefore, most of the cars on the roads do not have Advance Driver-Assistance Systems (ADAS). This thesis describes the design process of an automotive safety system which addresses vehicles without any driving automation system.

The thesis project is divided in three stages. The objectives of this thesis are addressed to show the individual development of each stage and to describe how the three stages are related, although all of them are focused in different areas.

The first stage describes the design methodology for the conceptual design of an automotive safety system. The second stage is about exploring new horizons, new opportunity areas and how to improve the conceptual design. It will be shown how Artificial Intelligence, especially Deep Learning approaches, could generate better results for some of the system functions, therefore, the overall performance of the system will be upgraded. The third stage is about how to address and study one of the most important human factors, the “trust” issue. Extensive evidence indicates that User Experience topics as acceptance, perceived usefulness and intention to use systems or products are strongly modulated by trust. As a result, this stage contributed to the conceptual design of a study for measuring the correlation between drivers’ trust and drivers’ gaze behavior on an advanced driving simulator.

The work on this thesis will benefit future research and product design process concerning the implementation of a retrofittable automotive safety technology, allowing to create a new driving experience which will help to increase the worldwide road safety levels.

Acknowledgements

I would like to thank my supervisor Dr. Alejandro Ramirez not only for supporting me with scientific and technical questions, but also for his helpful advice and guidance in difficult situations and times. Dr. Alejandro consistently allowed this thesis to be my own work, but steered me in the right the direction whenever he thought I needed it. Thank you for always guiding me whenever I ran into a trouble spot or had a question about my research.

I would also like to thank the Mexican teaching team (Dr. Alejandro Ramírez, Dr. Vicente Borja, I.D. Yesica Escalera, Arq. Arturo Treviño and Dr. Marcelo López) who supported us here in Mexico during the GlobalDrive project. Of course, thank you Johannes Wallner for your support and guidance during the GlobalDrive project. Without their passionate participation and input, the GlobalDrive project could not have been successfully conducted.

A very warm “thank you” to my teammates, now my friends, from the GlobalDrive project. This accomplishment would not have been possible without them. Thank you.

I must express my very profound gratitude to my family (especially to my mother) and friends for always being there for me. Thank you for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. I wouldn’t be where I am today without all your support. It is always great to come back home!

Finally, a big “thank you” goes to my supervisor at CARRS-Q (Centre for Accident Research and Road Safety - Queensland), Dr Ronald Schroeter. Your feedback and guidance allowed me to successfully culminate this thesis project. Thank you, Dr Ronny, for giving me a new vision and perspective of the automotive safety research. It was an honour to work with you!

I. First Stage: GlobalDrive

GlobalDrive is an international student project offered by the chair of Automotive Technology at Technical University of Munich (TUM). Three teams made up of four German students and four students from a partner university are built to work on three different problems. Additionally, each team will be supported by an industry partner. Without being narrowed down by specific requirements from the side of university or industry, the students are encouraged to be open-minded to come up with new creative solutions addressing their project's challenge.

For the project called "***Safety Cloud***", the partner university was Universidad Nacional Autónoma de México and Autoliv B.V. & Co. KG [112], the industrial partner. I worked in this team with my Mexican teammates (2016: Arenkkar Fernández, Alan Reyes and Pamela Esquivia. 2017: Anahí Velázquez, Carlos Canchola and Pamela Esquivia) and my German teammates (Antonia Hüfner, Clarissa Böker, Julian Kreuzer, and Simon Pöllmann). The team had two supervisors, one on the German side (Johannes Wallner) and a Mexican teaching team (Dr. Alejandro Ramírez, Dr. Vicente Borja, I.D. Yesica Escalera, Ar. Arturo Treviño and Dr. Marcelo López).

In this stage, it will be shown how was the GlobalDrive design process to create a conceptual design for an automotive safety system. In this stage, research was done about driver's concerns and issues, road accident statistics, current Advance Driver-Assistance Systems (ADAS), benchmarking, interviews and surveys. With this research, it was possible to find some factors that have high relevance to the drivers, some issues and situations that drivers in Mexico City are affronting every day and how to address these issues.

Universidad Nacional Autónoma de México

Technische Universität München

Autoliv

1. Introduction

1.1. Problem Description

Road safety is about people. Every day, millions of people use the road infrastructure: by walking, driving or riding a bicycle. Last year, almost 1.25 million people did not make it home; between 20 and 50 million people came home with a life changing or severe injury [37]. Every fatal or severe crash on the roads is a tragedy. It is a shared responsibility to take road safety seriously.

While human suffering cannot be measured, monetary costs to society from automobile accidents are estimated in the hundreds of billions of dollars each year for health care, rehabilitation, road infrastructure and loss of income.

Connecting road users by cloud-based functions can be the key to overcome these tragedies. Using intelligent assisting systems to increase the information that is available to drivers and vehicles will help to reduce fatalities and accidents.

In the future, smart cars could collect and analyse data from each other, the cloud and the transportation infrastructure to provide the right information, at the right time and in the right way to keep drivers safe and take proactive measures to keep traffic moving efficiently and safely. There is still a long way to go to accomplish the latter, so improving the safety level of cars without any Driver-Assistance System is an important topic.

1.2. Safety Cloud Objectives

The goal of this GlobalDrive project was to propose the conceptual design of a system capable of generating a new safety experience within cars that are not equipped with Advance Driver-Assistance Systems. With the use of User-Centred Design methodologies, opportunity areas were discovered and will open the possibilities to improve the functions of products that are already on the market.

2. Background

For a better understanding, the most important terms will be defined in the next sub-section. It is important to mention that all this research was not made in the first stage of the project. At the beginning of every stage, a new round of research was made to get new useful information and use it for new insights and improvements.

2.1. Definitions

Safety

Safety involves physical, spiritual, financial, political, emotional, occupational, psychological, educational and more fields. To start understanding the topic some definitions are given below:

Oxford Dictionary defines safety as: “The condition of being protected from or unlikely to cause danger, risk, or injury a modifier denoting something designed to prevent injury or damage” [109].

In the German Industry standard DIN 31000 safety is defined as: A behaviour of a system towards humans. Safety describes the degree of "absence of danger" [12].

Vehicle Safety

Automotive safety is the study and practice of design, construction, equipment and regulation to minimize the occurrence and consequences of traffic collisions [36].

Generally, vehicle safety can be divided into two parts: active and passive safety. “Active safety” refers to technology assisting in the prevention of a crash and “passive safety” to components of the vehicle (primarily airbags, seatbelts and the physical structure of the vehicle) that help to reduce the severity of accidents for occupants.

Cloud

So far, there is no uniform definition of cloud-computing. One of the most common definitions is given by the US National Institute of Standards and Technology: Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [21].

Another definition comes from IBM: Cloud computing often referred to as simply "the cloud," is the delivery of on-demand computing resources over the internet on a pay-for-use basis [17]. A shorter definition can be noted as Cloud-computing is the provision of IT-resources through a network [11].

Nowadays there are three types of cloud computing [17]:

- *Software as a Service (SaaS):* Or called Cloud-based applications, run on distant computers "in the cloud". Owned and operated by others, connecting users' computers via the internet.
- *Platform as a Service (PaaS):* Provides a cloud environment with everything required to support the entire lifecycle of building and delivering cloud applications, with the pros of no cost and complexity of buying and managing underlying hardware, software, provisioning and hosting.
- *Infrastructure as a Service (IaaS):* Provides processing resources, storage, networks, and other fundamental computing resources where the consumer can deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

Artificial Intelligence

The Oxford Dictionary defines Artificial Intelligence (AI) as: "the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages" [108].

Machine Learning

Machine Learning is the practice of using algorithms to parse data, learn from it, and then decide or predict about something in the real world. Instead of hand-coding software routines with a specific set of instructions to achieve a task, the machine is "trained" using large amounts of data and algorithms that give it the ability to learn how to perform the task [71].

Artificial Neural Networks

Commonly called Neural Networks, are inspired by the biology understanding of the human brains, the only difference between the biologic approach and the Machine Learning approach is that unlike a biological brain where any neuron can connect to any other neuron within a certain physical distance, these artificial Neural Networks have discrete layers, connections and directions of data propagation [71].

Deep Learning

Deep Learning is a Machine Learning technique that teaches computers to do what comes naturally to humans: learn by example. Deep Learning models are also called as Deep Neural Networks. The term “deep” usually refers to the number of hidden layers in the neural network. Deep Neural Networks can have as many as hundreds hidden layers [76]. In the next sections, Deep Learning will be introduced, and it will be explained the way it works.

2.2. Advanced Driver-Assistance Systems

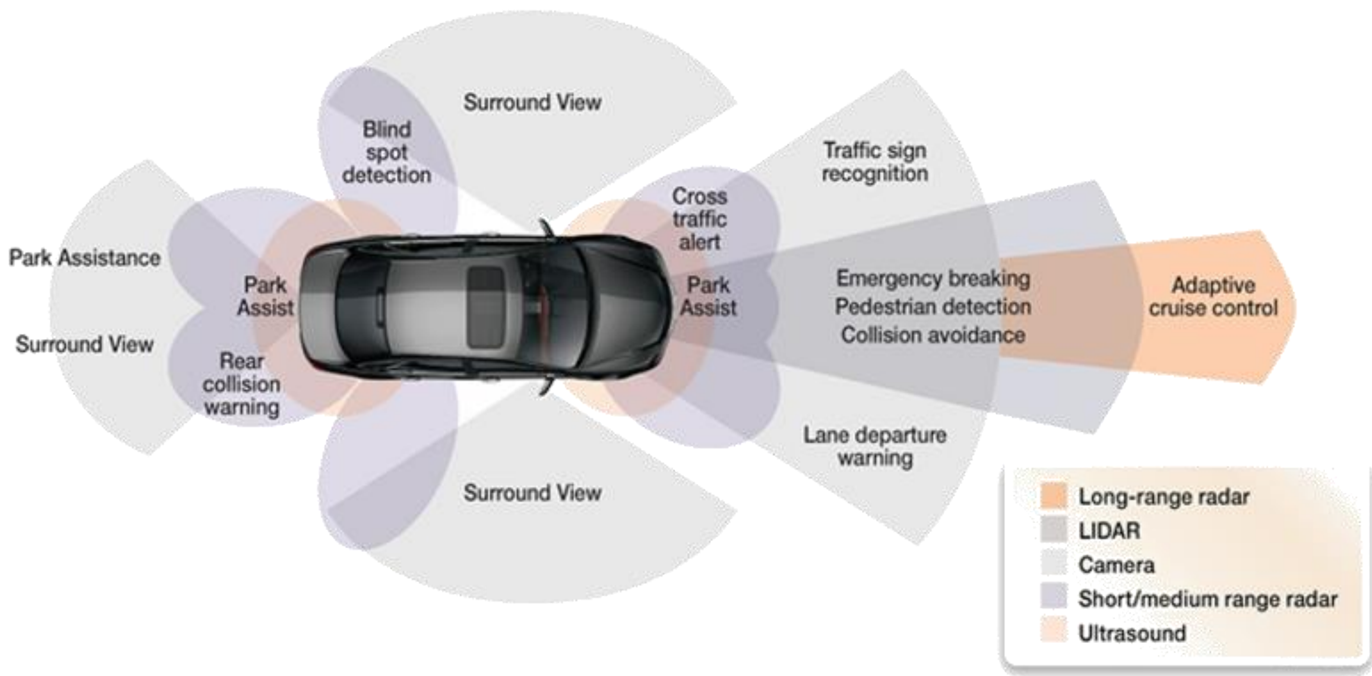


Figure 1: Overview about existing ADAS [113].

Park Assist

Using vision-based systems and ultrasonic sensors to offer information about the immediate surrounding that is mostly used for parking assist functions. Some systems can judge the size of a parking space and inform the diver about the suitability. Furthermore, some can steer into a parking space parallel as well as bay parking. This helps to reduce material damages [161].

Driver Drowsiness Detection

As mentioned above, drowsiness is a major cause of accidents. There are three different ways of measuring a driver's drowsiness [13]. These measures are:

- Vehicle-based measurements. Metrics as a deviation from lane position, movement of the steering wheel, pressure on the acceleration pedal, etc., are constantly monitored and any change that crosses a specified threshold indicates a significantly increased probability that the driver is drowsy.
- Behavioural measurements. Behaviour of the driver, including yawning, eye closure, eye blinking, head pose, etc., is monitored through a camera and the driver is alerted if any of these drowsiness conditions are detected.
- Physiological measurements: Physiological signals such as electrocardiogram (ECG), electromyogram (EMG), and electroencephalogram (EEG) are analysed to determine the probability of a drowsy driver scenario.

Brake Assist

The brake assist increases braking pressure in emergency situations. It faces the problem that most of the drivers fail to brake with enough force when faced with an emergency. By interpreting the speed and force the brake pedal is pressed, the system detects if the driver is trying to execute an emergency stop and if so, the system overrides and fully applies the brakes [162,164].

Collision Avoidance System / Collision Mitigation System

Collision avoidance system is a generic term for automobile systems to reduce the severity of rear-end collisions. It usually uses radar and sometimes camera and laser to detect an imminent crash. After the detection, the systems either provide a warning to the driver or act autonomously by steering or braking the car.

Collision mitigation systems offer a more detailed description of actions performed by a braking system. It distinguishes three stages of warning depending on the Time To Collision (TTC): first, it only warns optically (TTC~3s). At TTC~2s it additionally warns acoustically and brakes with 25% of the maximum braking force. And at TTC<1s it prepares the driver for the crash by pulling the seatbelts and brake with 60% of the maximum braking force [163].

Emergency Stop Signal

The emergency stop signal is emitted, when the car detects that the deceleration is high enough to be considered as an emergency or heavy braking situation. Both turning signals are repeatedly turn on and off.

Adaptive Cruiser Control

Mostly using the same sensors as the collision avoidance system, the ACC is designed to avoid accidents by maintaining the distance to a vehicle ahead, using only on-board sensors. The traffic ahead is watched by a small radar unit, usually located behind the front grille or under the bumper, that looks on to the preceding vehicle in the lane. Based on the radar signals, a control unit computes the distance to the preceding vehicle, the relative speed, and the lateral position on multi-lane roads to determine which car to track. Apart from radar, using a laser or an optical system based on stereoscopic cameras is also possible. Regardless of the technology, the ACC slows the car down or speeds it up to keep pace with the preceding vehicle. It works at the day as well as by night, but its abilities are limited by heavy rain, fog, or snow. Also, the system does not detect stationary obstructions.

The cooperative adaptive cruise control extends the automation of navigation by using information gathered from fixed infrastructure like satellites and roadside beacons or mobile infrastructure [110,111,163].

Lane Departure Warning

The Lane Departure Warning systems are using a camera that is mostly mounted up in the windshield, to watch the road ahead. The processing software breaks down the digitized image identifying straight or dashed lines as the lane markings. Since the driver is supposed to centre the car between the lane markings, the system monitors the vehicle's distance to the road markings, predicts unplanned lane crossing and warns the driver. If the driver indicates his lane change by using the turning lights, he is not alerted by the system, because an intentional lane change is assumed. In some cases, the lane departure warning is realized by a set of laser or infrared sensors. Also, a rear-facing camera can be used to watch the road markings behind the car [35].

Lane Keep Assistant

The Lane Keeping Assistant is a more evolved version of the Lane Departure Warning but uses the same sensing. As a difference, the Lane Keeping Assistant is not only warning the driver but directly intervenes when unintentionally leaving the lane. If the vehicle reaches the lane markings, the assistant automatically corrects the course, or it even keeps the vehicle centred in the road. The corrections can be affected by braking the opposite front wheel to turn the car or by steering the car back into the lane. The driver can always overdrive the car's guideline even though the system is adding up to 80% of the maximum steering torque to keep the vehicle on the road [35, 20]. This is only executed when:

- The driver has contact to the steering wheel.
- Turning lights are not activated.
- The driver’s steering angle indicate that he is trying to change lane

Lane Change Assistant / Blind Spot Monitors

With adding an extra camera to the side of the mirrors or radar sensors concealed in the rear bumper, vehicles and VRUs in the blind spot of the car can be detected. The driver can then be warned when trying to change the line [165].

2.3. State of the Art and Benchmarking

Advanced Driver Assistance Systems are becoming increasingly important and extensive in modern cars. One of their main purposes is to improve vehicle safety and thereby reduce traffic deaths. Statistics show that there is a clear correlation between the improvements of safety related ADAS and the reduction of traffic deaths. Figure 2 shows that the number of persons killed in car accidents reduced by 70 percent in Germany within ten years. The figure 2 also shows that a growing number of assistance systems became widespread among middle and high-class cars.

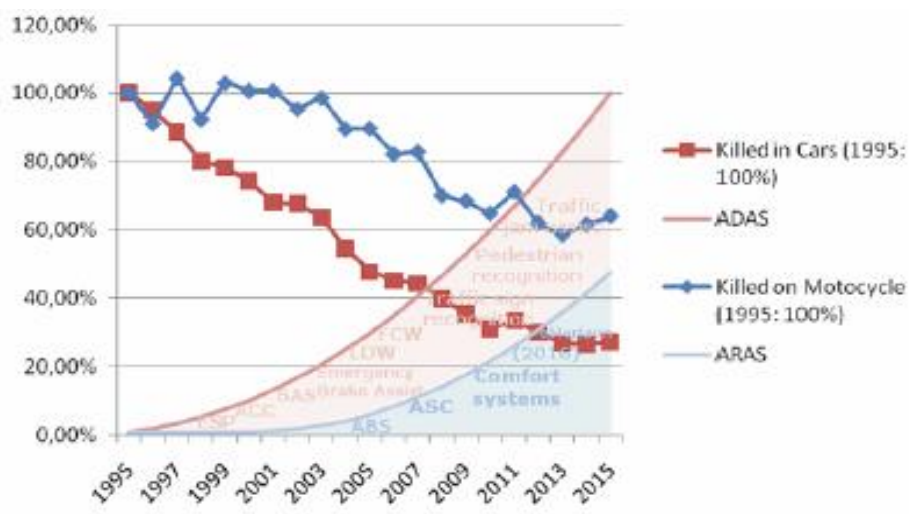


Figure 2: correlation between Traffic deaths and assistance systems [29]

While there are already many existing safety systems for cars, they are either only providing a subset of possible safety functionalities or they are too expensive. This section will present some of the existing solutions, highlighting their drawbacks which the new concepts aim to solve.

2.3.1. Existing ADAS and safety-related Apps for Cars.

According to [52] there are currently 45.1 million cars on the road in Germany with an average age of 9.2 years and both numbers increase in countries like Mexico, India or Venezuela. Based on this average age and keeping in mind that the ACC system for example has only been built into high-end cars from 1999 [49] on, it becomes obvious, that most cars on the road today cannot enjoy the safety benefit of recently developed ADAS systems.

Original Equipment Manufacturers (OEMs) usually offer a set of powerful safety features built into their high-end cars. As one example, the Mercedes Intelligent Drive system [22] includes collision prevention, an active lane keeping, and an active blind spot assistant as well as an ACC and night view assistant. The system is furthermore capable of actively braking if necessary. Besides relying on multiple car sensors these services also use expensive radar technology. Standing to the benefit of this system is only possible when buying a specific car model, as it is built-in and therefore not possible to retrofit it to older cars.

External devices in the market such as Nauto [23], Caruma [9] or Carvi [10] aim to close this gap by providing devices that may easily be built into any car. Some of them also include emergency services and different add-ons, trying to satisfy the driver. However, they come at a considerably high cost and many of them are not clear enough to understand and interpret the indications in an easy, non-distracting way.

One of the main problems of the existing devices like CARVI is the use of the smartphone as a way to get information or to interact with the driver. This is followed by some different issues, since not every driver owns a smartphone or wants to use it in this way. Also, Carvi only provides a subset of features: it just monitors the road, therefore never checking on the driver’s state.

Some wearable devices are designed to monitor the driver’s sate. The wearables AdvicyDrive [41] and StopSleep [42] are specifically designed to detect drowsy or sleeping drivers. While AdvicyDrive is based on measured heartbeat, StopSleep analyses the skin conductivity to infer on brain activity. However, the devices are only of help regarding a tired driver but do not consider any other potential risk for accidents such as unexpected situations occurring on the road ahead. Furthermore, StopSleep may disturb the driver while driving as it hinders one hand from freely moving the fingers.

The cheapest safety functionalities on the market are offered by smartphone apps. As they are easy to develop and distribute to customers, there exist multiple ones on the app market or in research projects designed for different requirements: the project CarSafe [2], from Dartmouth college, monitors both the road and the driver, proving that this is possible based on both front and rear-facing camera of the phone. The app DriveSafe [19] evaluates the driver's performance and builds up a scored driving profile. Additionally, the road ahead is checked for dangerous situations which will trigger warnings. ACoDriver [3] similarly assists in driving by identifying road signs, checking for lane changes or the distance to the car in front. Other apps, like iOnRoad [61], also offer lane-departure warning functions.

Besides this group of apps that are serving as an additional "co-pilot", there exist multiple ones displaying position-based information to the driver based on driver's reporting on the road. The most vivid traffic community is probably offered by Waze [34]. Their maps provide drivers with multiple warnings and information in addition to the best route based on the current traffic. However, they rely solely on users reports and no additional car information is analysed.

This makes the use of the app unappealing in areas without an already existing community. All of the presented apps are constrained in their accuracy and in the dimension of their provided functionalities due to the limited resources of the smartphone. Most important they only provide a subset of possible functionalities and rely on the phone alone being able to process all data, which might be difficult on old phones for computationally expensive tasks such as image processing.

One common problem of the existing solutions is the quantity of information that is shown to the driver (i.e., the output to the driver) by a smartphone or by an external display. If the way to deliver information to the driver is incorrectly addressed, it will imply a distraction while driving. Another disadvantage of the existing solutions is not only the way to show the alerts to the driver but also the way to determine whether the system should emit the alert and with what intensity. Most of the commercial devices just use a common "alert" or "no alert".

More detailed information about the mentioned products or apps, it is showed in the Appendix A – Benchmarking description.

2.3.2. State of the Art in Motorcycle's Safety.

Until now, the deploying of technologies developed for cars into Powered Two Wheelers (PTW) is a highly complex task. For instance, a system that shows how complex a transfer between the types of vehicles can be is the emergency braking system that is built into modern and mostly higher-priced cars. It represents systems with high complexity. Cars and their physics do not depend on the passenger's behaviour. In contrast to that, an assistive system in a motorcycle can only work properly and without harming the rider if his or her behaviour is suitable to the intervention.

Especially cooperative systems attend to have a huge potential such that BMW Motorrad, Honda, and Yamaha founded the Connected Motorcycle Consortium(CMC) [174] in 2015 and settled that there will be one cooperative system in at least one of each of the company's motorcycles until 2020. With this goal in mind, there is a lot of research in progress to achieve this.

One example was presented by Continental at the CES2017, called eHorizon-concept [171], that was originally developed for cars is now adapted to motorcycles. The instrument cluster does not only offer information about the current speed and the status of the motorcycle but also about potentially dangerous situations including construction sites, loose chippings, or oil on the road.

This data is retrieved from the cloud which is filled by the input of other drivers. And therein lies the drawback: The system only works if enough motorcyclists participate actively in the community. The basic principle is to indicate every dangerous situation during driving by pressing a button.

After the ride, it must be specified which kind of threat has been seen. An entry is made in the dedicated app, which causes effort for the rider without direct benefit. Again, on CES2017 Honda revealed Cooperative Mobility Ecosystem" [173] which builds the framework for dedicated short-range communication to implement cooperative motorcycle safety functions.

One assistive system that is partly already available and partly in development in many companies is the eCall system for motorcycles. There are different approaches to transfer the well-known function from cars to motorcycles. As the common crash-sensors used in cars (acceleration, pressure, and sound sensors) are not easily transferable to the changed driving physics of PTW. BMW Motorrad, for example, introduced a system that relies on lean angle and acceleration sensors to detect an accident [172]. Together with a mobile phone module, a set of data can be forwarded to an emergency central. A button on the handle bar also allows indicating an accident manually. The rider is tried to be contacted by a phone call that comes through a speaker and microphone module on the handle bar.

Figure 3: Motorcycle eCall by BMW Motorrad [172]

The German helmet producer Schuberth also invented a system called Rider eCall and introduced it to the market in 2013 [26]. The basic principle was equal to the BMW Motorrad system, only with having the acceleration sensors in the helmet.

Unfortunately, they had to withdraw it from the market only one year later because they had too many false alarms [27]. This already shows the complexity of correct accident detection. For retrofit system, dGuard [167] claims to offer the first eCall system for motorcycles. The system is similar to the BMW Motorrad system but suitable for every bike. Using vision-based environment perception systems is so far not widespread among PTW producers. Nevertheless, BMW Motorrad developed a vision system for traffic sign detection, which is either mounted on the helmet or in the front of the motorcycle. Camera-systems in or on motorcycles are so far mainly used to record the ride for entertainment and sometimes analyse afterwards.

Besides well-known action-camera producers like GoPro [15] that can easily be mounted on different spots using suction cups, a company called Waylens [18] offers a camera system that combines high-resolution footage with additional information from OBD-II. A comparable system for PTW could not be found, even though the multiple usage of unspecialized cameras on motorcycles shows the big market potential.

Navigation systems that are specifically designed for PTW are available in various shapes and implementations. Alongside many retrofit systems that are comparable to systems used in cars, there are several built-in systems in motorcycles. As they are specifically designed for this purpose, they all work well and give back the needed information to the driver. Thereby the most important issue is to transmit information in an easy and understandable way without disturbing the driver.

Concepts try this by mainly using two transmission channels the human body offers: acoustic and optic. The easiest and hence most common way is to give back feedback visually. This does include not only ordinary displays but also innovative systems like head-up displays (HUD) in helmets in many shapes. Some of them also come with acoustic signals to avoid distraction. All the commonly available systems resemble in the fact that they do not offer information to make the ride safer but only to make it more convenient and faster.

2.3.3. Augmented Reality and Projection Systems

An Augmented Reality head-up display shows information in the windscreen for the driver. As part of Lane-Departure Warning systems, actual road markings can be emphasized by overlaying with a virtual image. Another augmented reality solution is the holographic navigator from WayRay [6], which combines the use of a camera and the GPS signal to navigate the driver by displaying indicators on the road ahead [14].

Figure 4: Head-up display [16]

Figure 5: Holographic navigator [6]

On the downside, this technique does not work without actual road marks to detect. However, in most situations, drivers are capable of detecting road marks themselves. Therefore, safety would be further improved, if road marks were displayed especially on streets without any markings or in situations with difficult weather conditions when the marks cannot be seen.

Except detecting road markings, there are also approaches to improving the safety by projecting road markings onto the street with laser light. Most commonly this is used in bicycles lights, for example in the Bike Lane Safety Light from XFIRE.

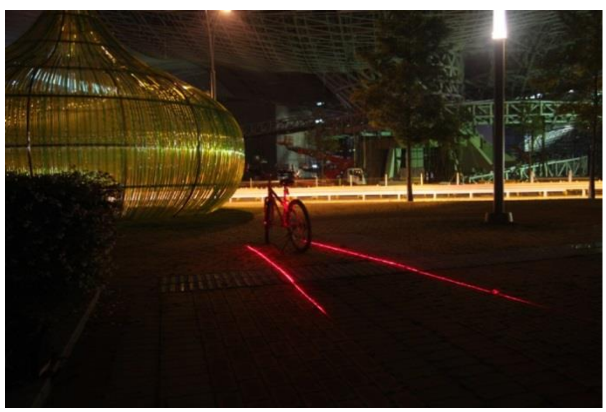


Figure 6: Bike Lane Safety Light from XFIRE [8]

Also, automobile applications are conceivable, like in the F015 research vehicle from Mercedes-Benz. This self-driving concept car is equipped with a laser projection system to communicate with the surrounding [31].

Figure 7: Mercedes-Benz F015 [31]

While these solutions generate bicycle lanes or crosswalks for pedestrians on the street, the safety for pedestrians could also be enhanced by projected warnings. While detecting pedestrians and predicting their path is susceptible to slip-ups, a visual warning on the road around the car could raise attention without disturbing the driver with false warnings.

Another method of warning pedestrians rather than the driver is currently discussed with acoustic warnings for electric vehicles. While constant acoustic signals increase the noise level, acoustic signals are generally jeopardized to be superimposed by other sounds like ambient noise or music.

2.4. Accident Statistics and survey

Road traffic injuries can be prevented. Governments need to take action to address road safety holistically, that requires involvement from multiple sectors (transport, police, health, education) and that addresses the safety of roads, vehicles, and road users themselves.

Effective interventions include designing safer infrastructure and incorporating road safety features into land-use and transport planning; improving the safety features of vehicles; and improving post-crash care for victims of road crashes. Interventions that target road user behaviour are equally important, such as setting and enforcing laws relating to key risk factors and raising public awareness.

Even though cars get increasingly equipped with modern safety systems, road traffic is still dangerous and causes tremendous damages. In 2015 the World Health Organization (WHO) revealed an extensive status report on road safety, showing that worldwide the total number of road traffic deaths has plateaued at 1.25 million per year. Figure 8 shows that road traffic injuries still cause the most deaths among young people. Besides the burden this means for the surviving dependents of the victim, it produces enormous economic damage.

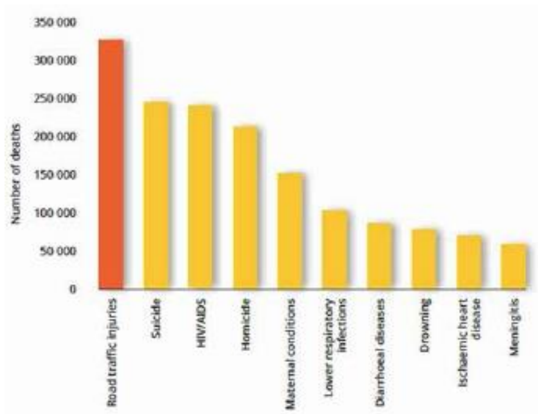


Figure 8: Top ten causes of death among people aged 15–29 years, 2012 [37]

Germany and Mexico are both below the average of killed persons per capita. The German average of 4.3 is especially remarkable in respect to an average of 9.3 killed inhabitants per 100,000 persons in whole Europe. Reasons for the low rate are manifold: High safety standards for cars, good road conditions and the ability to afford modern cars with sophisticated active and passive safety systems.

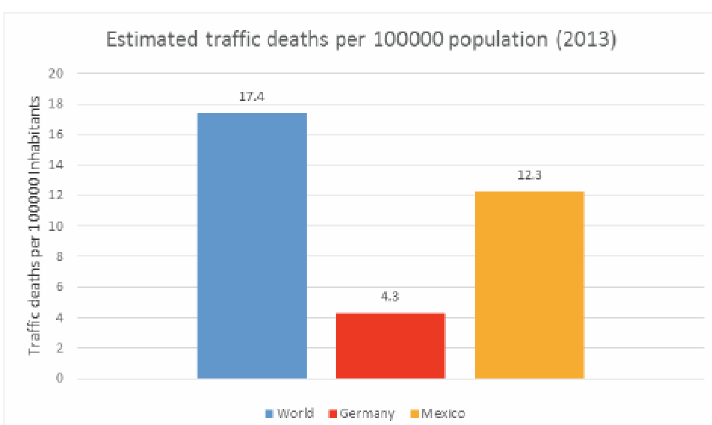


Figure 9: Traffic deaths per 100000 in Germany and Mexico (Data from [37])

Figure 9 is a result of the development in the past 20 years. As evaluated by the German statistical federation and shown in Figure 10 the number of traffic deaths decreased from 9454 in 1995 to 3377 in 2014. Unfortunately, the number increased again to 3459 in 2015.

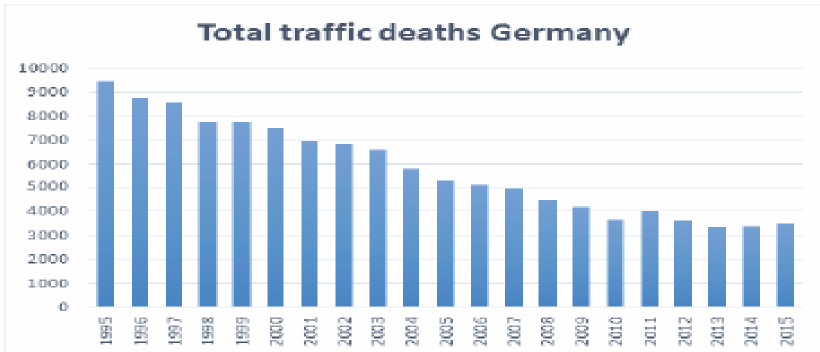


Figure 10: Total traffic deaths in Germany 1995-2015 [29]

To get a better understanding of the reasons why road accidents still happen and even tend to increase again, details of accidents need to be investigated. 1.2 errors per accident were detected in Germany in 2015 [28]. In other words, this means on average every accident is related to a human error. The fact that humans are possibly the most dangerous source of error results in the accelerating growth of Automated Vehicles. Elon Musk even goes as far to state that “People may outlaw driving cars because it's too dangerous” [175]. Up to now, the technology cannot completely replace the cognitive capabilities of human drivers. However, modern assisting systems can guide the way to automated driving and by that already offer safety benefits for the drivers. To identify the biggest potentials of cloud-based assisting systems, the most frequent human errors are evaluated.

Analysing road accidents caused by human, errors can be classified by answering the following questions: What is happening, or which mistakes are made and why is it happening?

For the first question, the German Federal Statistical Office offers detailed information about what causes the most deaths on the roads. These numbers are compared to a survey among 111 participants who formed an order of what they think causes the most traffic deaths. This order is shown in the red boxes. The position (among these factors) in the statistics is shown in the blue boxes.

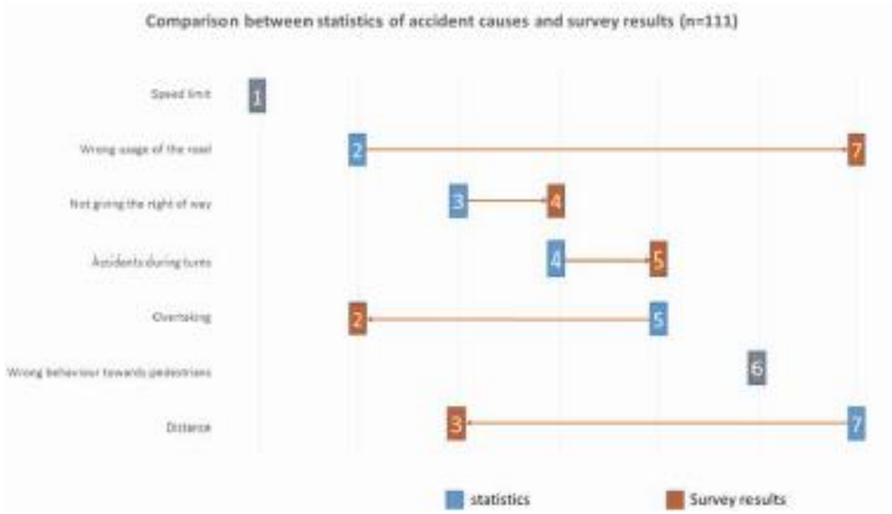


Figure 11: Comparison between statistics of accident causes and survey results [28]

While speeding is determined as the most dangerous misconduct by both – the survey and the statistics – the more interesting findings can be seen where the results differ a lot. While having too little distance to the preceding car is ‘only’ responsible for 161 fatalities in 2015, the participants see it as the third most dangerous error.

A similar imbalance can be seen in overtaking: while 282 traffic deaths caused by errors during overtaking mean place five in the statistics, the subjective feeling of the participants is that it is the second most dangerous. On the other hand, the wrong usage of the road is seen as the least dangerous but in fact, causes 473 traffic deaths and is with this the second most dangerous error. Summarizing the above, it can be said that the subjective feeling of being endangered differs a lot from the objective, assignable dangers highlighted by statistics.

After a clearer view on what is happening on the road, the question arises: Why are those mistakes happening? What key risk factors lead to those accidents?

During the development process, the team identified four main problems that contribute to crashes: environmental influences, distraction, inappropriate driver’s behaviour, and drowsiness. In the following, it is tried to estimate their influence on road safety.

From the team's point of view, the environmental influences include two mayor types of dangers: bad roads and different weather conditions. The term 'bad roads' is not only related to worn streets with potholes but also to road-designs that do not offer sufficient safety for vulnerable road users (VRU). In Germany, 155 accidents with fatalities occurred in 2015 because of the road conditions.

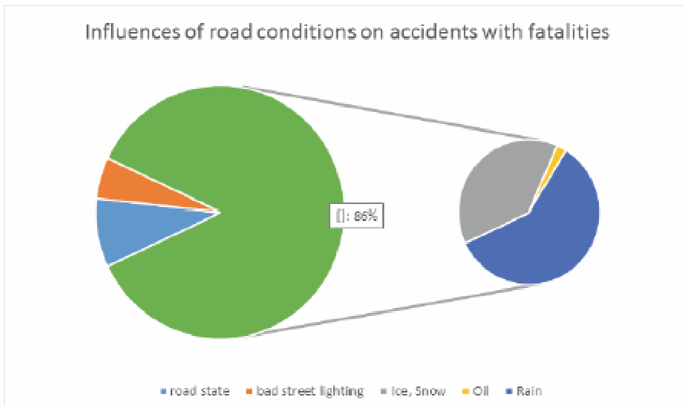


Figure 12: Influences of road conditions on accidents with fatalities [29]

As Figure 12 shows, most accidents (86%) where the road conditions had an influence happened because of slipperiness. Herein rainy and icy roads had the biggest proportion. Another 74 accidents with fatalities happened because of other weather conditions that did not influence the road conditions directly, especially by visibility impairments caused by the sun or fog. Another big issue that causes many accidents is the distraction of drivers. Before showing statistics of how crucial drivers' attention is, distraction is going to be defined. Combining various approaches, some key elements which have been thought about in defining driver distraction can be revealed:

- There is a diversion of attention away from driving or safe driving;
- Attention is diverted towards a competing activity, inside or outside the vehicle, which may or may not be driving-related;
- The competing activity may compel or induce the driver to divert attention towards it;
- There is an implicit, or explicit, assumption that safe driving is adversely affected [24].

There are three main types of distraction [30]:

Visual

Visual distraction occurs when a driver looks at anything other than the road ahead. As an example, drivers who check their children’s seatbelt are visually distracted.

Manual

Manual distraction is when the driver takes one or both hands off the wheel for any reason. Some common examples include eating and drinking in the car, adjusting the GPS, or trying to get something from a purse, wallet, or briefcase.

Cognitive

Cognitive or mental distraction is when a driver's mind is not focused on driving. Talking to another passenger or being preoccupied with personal, family, or work-related issues are some examples.

Using the smartphone during driving is especially dangerous as it includes all three kinds of distraction. The number of smartphone users is constantly increasing. The combination of those two facts is also alarming the WHO. Stating that 69% of drivers in the United States of America had used their mobile phone while driving within the previous 30 days, it shows how common the usage of the smartphone during driving is. Among the survey participants, 42% confirmed this fact. Concurrently, 34% indicated that they think that using the smartphone during driving influences their road safety the most. These numbers show that drivers are aware of the dangers smartphone usage during driving has, yet still many of them use it. It is reported that texting while driving caused 1.3 million crashes in one year [1].

Besides the accidents caused by a lack of attention on the driving task, inappropriate driver’s behaviour provokes many incidents. It is supposed that this term mostly refers to aggressive driving style. As this kind of misbehaviour is mostly intentional, it is hard to overcome it with technological approaches. Governments, therefore, therefore created various approaches to show the hazards this behaviour could cause. The foundation for Traffic Safety in the US surveyed almost 200.000 fatal crashes from 2003 to 2007. In more than 55% of the cases, at least one aggressive action was reported. In very most of the cases (30,7%) speeding over limits or too fast for the condition was the causing reason [4]. According to the WHO, ignoring speed limits is most common among male and young drivers. The given statistic also shows that an adult pedestrian has less than a 20% chance of dying if struck by a car at less than 50 km/h but almost a 60% risk of dying if hit at 80 km/h [37]. However, the report also states that speed limits can only be truly effective with enforcement, assisting systems that act contrary to the driver's intention will not be used by the drivers. 32% of the participants in the safety-cloud survey indicated that they think that misbehaviour of other drivers influences their driving safety the most.

The last key risk factor is drowsiness. In the year 2009, the US National Sleep Foundation (NSF) reported that 54% of adult drivers have driven a vehicle while feeling drowsy and 28% of them actually fell asleep [13].

The Allgemeiner Deutscher Automobilclub (ADAC) underlined that sleeping drivers are responsible for one-fourth of all fatal crashes on highways. Furthermore, a study by researchers in Australia showed that being awake for 18 hours produced an impairment equal to a blood alcohol concentration (BAC) of .05, and .10 after 24 hours; .08 is considered legally drunk. In fact, the number of crashes caused by drowsiness might even be higher as there is no measurement tool like there is for intoxication. Comprehensively drowsiness can be identified as one main cause of accidents.

All the facts support that drivers are responsible for a lot of accidents. Except for the intentional inappropriate behaviour, drivers are making mistakes whenever the level of stress is too high (overtaxed by the amount of information) or too low (boredom or mental underload). As the Yerkes-Dodson law states, the ideal performance is possible when the stress load remains on a medium level. The proposed assisting system should therefore always help to keep the stress for the driver on a medium level.

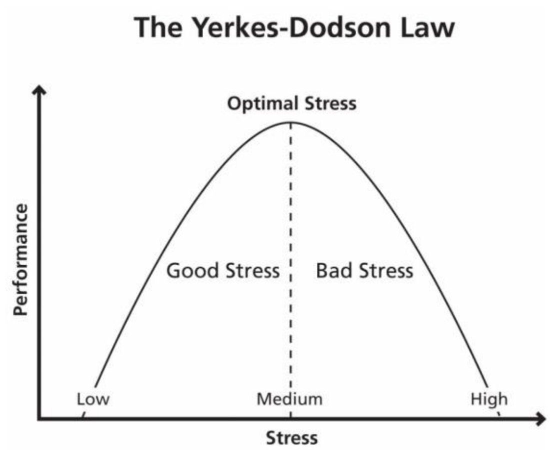


Figure 13: Yerkes-Dodson Law [38]

2.5. Artificial Intelligence

2.5.1. Background

Self-learning machines are the essence of Artificial Intelligence (AI). Concepts already date back more than 50 years. However, AI itself gained wider functional applicability only in the past few decades, with the rise of Machine Learning and Deep Learning. Only recently, technological advances have enabled successful implementation at industrial scale. AI is finally bringing a multitude of capabilities to machines which were long thought to belong exclusively to human, like processing natural language or visual information, recognizing patterns, and decision making.

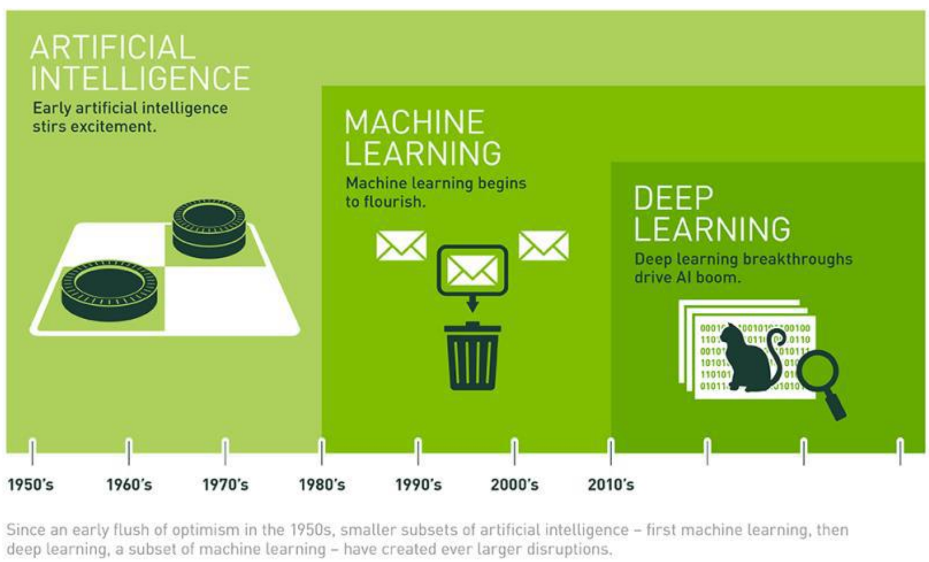


Figure 14: Artificial Intelligence timeline [71]

The essence of intelligence is learning. Just as humans learn how to identify audio-visual patterns, play a game or drive a car, machines can be trained to perform such tasks based on learning algorithms. One commonly method for training consists of providing to the machine labelled data [72]. This type of training is also named supervised training. Such machines are then said to possess a level of AI, then they can, with the previous training, get the correct label to a previously unknown data set with sufficient accuracy.

Well-trained AI can perform certain tasks at the same skill level as humans. The trained AI has the additional advantages of high scalability and no need for pauses. AI can discover some data patterns that are too complex for human experts to recognize. In some specific applications such as computer vision, AI has already achieved performance levels surpassing that of humans (e.g., in skin cancer diagnostics, heart disease diagnostics and others). The beginning of AI dates to the 1950s. In the last years, established IT companies have begun bridging the gap between science and business applications. Nowadays, adoption of AI has become increasingly easier due to open source available algorithms and libraries, relatively inexpensive cloud-based computing power and the proliferation of sensors generating data. In the industrial sector, the AI application is supported by the increasing adoption of devices and sensors connected through the Internet of Things (IoT).

Vehicles, assembly lines, production machines or devices carried by human workers generate enormous amounts of data per week. With the use of AI, it is possible to enable the use of such data for highly value-adding tasks like predictive maintenance or performance optimization with high levels of accuracy. Hence, the combination of IoT and AI is expected to kick off the next wave of performance improvements, especially in the industrial sector. Given its growing accessibility, broadening applications and specific relevance to the industrial sector, AI is a hot topic for leading researchers, investors and companies [72].

The overall interest in AI could lead to a considerable willingness to pay for those features in automotive/mobility technology. There are high expectations around the topic, with the average of all consumers expecting full autonomy to be widespread on the road already in about five years [74].

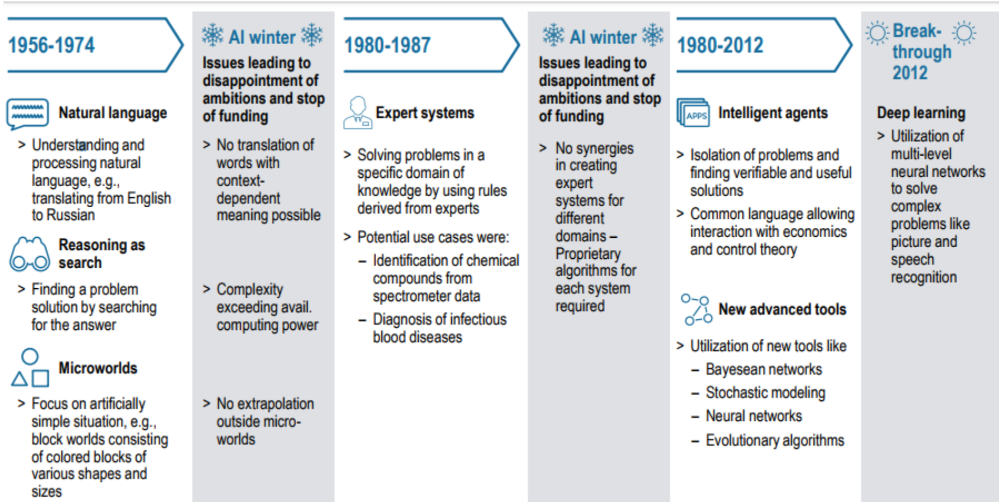


Figure 15. - Artificial Intelligence simplified story [73].

2.5.2. Machine Learning and Deep Learning

Machine Learning came from minds of the early AI academic community and the algorithmic approaches over the past years (e.g. decision tree learning, Bayesian networks).

For many years, computer vision was one of the most used application areas for Machine Learning. However, to achieve good results it required a lot of hand-coding. Engineers and scientists developed and used hand-coded classifiers. With the developed algorithms, it was possible to analyse some images and “learn” to determine whether it was a stop sign, a car, a pedestrian or a tree.

These classifiers had good results in ideal and controlled conditions, but some algorithms have problems in accuracy with hard conditions when objects are not perfectly visible (for instance, a foggy day), or when a tree obscures or occludes part of the objects. Due to those problems, Machine Learning algorithms were too brittle and too prone to error.

Another approach from the early Machine Learning community was the Artificial Neural Networks, came and mostly went over for many years. The Neural Networks (NN) have been shunned from the academic community until recently. This was due to the expensive computational resources that NN use, making them an unpractical approach. NN have discrete layers, connections and directions of data propagation.

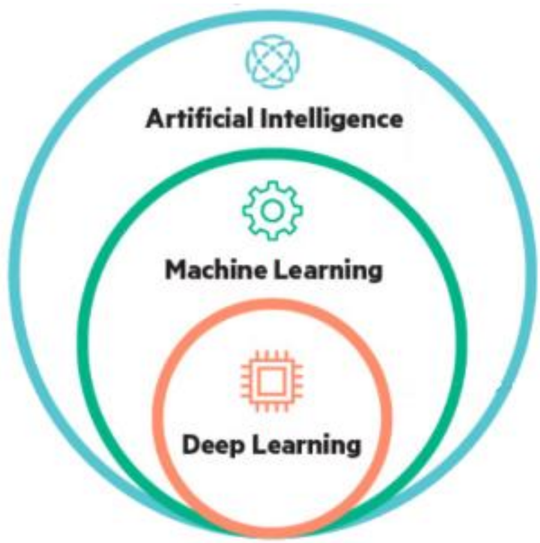


Figure 16. – The relationship between Artificial Intelligence, Machine Learning and Deep Learning [75]

Each neuron assigns a weighting to its input, this means how correct or incorrect it is relative to the task being performed. The final output is then determined by the total of those weightings.

Deep Learning is a specialized form of Machine Learning and both fall under the broad category of Artificial Intelligence. A Machine Learning pipeline starts with relevant features being manually extracted from images. The features are then used to create a model that categorizes the objects in the image. With a Deep Learning pipeline, relevant features are automatically extracted from images. Also, Deep Learning performs “end-to-end learning”, this means a network is given raw data and a task to perform, such as classification, and it learns how to do this automatically [76]. Deep Learning focuses even more narrowly on a subset of Machine Learning tools and techniques and applies them to solving just about any problem which requires human or artificial “thought” [89].

Deep Learning algorithms scale with data which is another key difference, whereas shallow learning converges. Shallow learning refers to Machine Learning methods that do not improve at a certain level of performance when more examples and training data are added to the network.

A key advantage of Deep Learning Networks is that the models often continue to improve as the size of the data increases. In Machine Learning, manually features and classifiers are selected to sort images. With Deep Learning, feature extraction and classification steps are automatic [76].

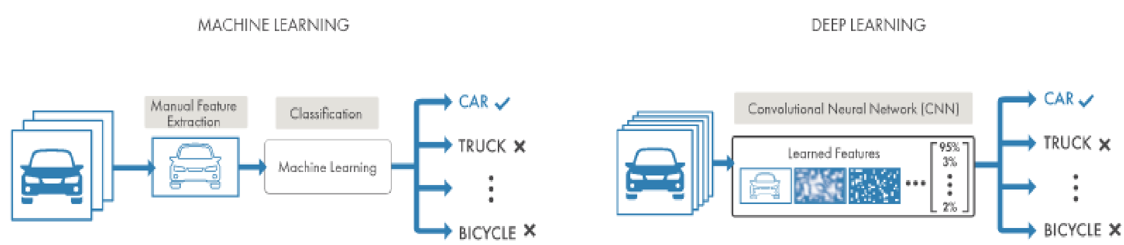


Figure 17. – Comparing a Machine Learning approach (left) and a Deep Learning approach (right) to categorize vehicles [76]

2.5.3. Deep Learning

2.5.3.1. Overview

Deep Learning is the current state of the art in AI, the potential of which has only recently been unlocked through advances in computing power and the availability of big data [74]

Deep Learning is a Machine Learning technique that teaches computers to do what comes naturally to humans: learn by example. Deep Learning is a key technology behind driverless cars, enabling them to recognize a stop sign or to distinguish a pedestrian from a lamppost. In Deep Learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep Learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labelled data and neural network architectures that contain many layers. Deep Learning has been the most successful approach to many areas where Machine Learning is applied. There are two main reasons it has only recently become useful [76]:

1) Deep Learning requires large amounts of labelled data. For example, driverless car development requires millions of images and thousands of hours of video. I.e., millions of driven miles to get good results.

2) Deep Learning requires substantial computing power. High-performance GPUs have a parallel architecture that is efficient for Deep Learning.

Deep Learning applications are used in industries from Automated Driving to medical devices. Deep Learning breaks down tasks in ways that make all kinds of machine assists seem possible [71]. Some examples are:

- Automated Driving
- Aerospace and Defence
- Medical Research
- Industrial Automation
- Electronics

Most of the Deep Learning methods use neural network architectures; therefore, Deep Learning models are also called as Deep Neural Networks. The term "deep" usually refers to the number of hidden layers in the neural network. Deep Neural Networks can have as many as hundreds hidden layers [76].

The advantage of multiple layers is that they can learn features at various levels of abstraction. For example, if a deep convolutional neural network is trained to classify images, the first layer will train itself to recognize very basic things like edges, the next layer will train itself to recognize collections of edges such as shapes, the next layer will train itself to recognize collections of shapes like eyes or noses, and the next layer will learn even higher-order features like faces [90].

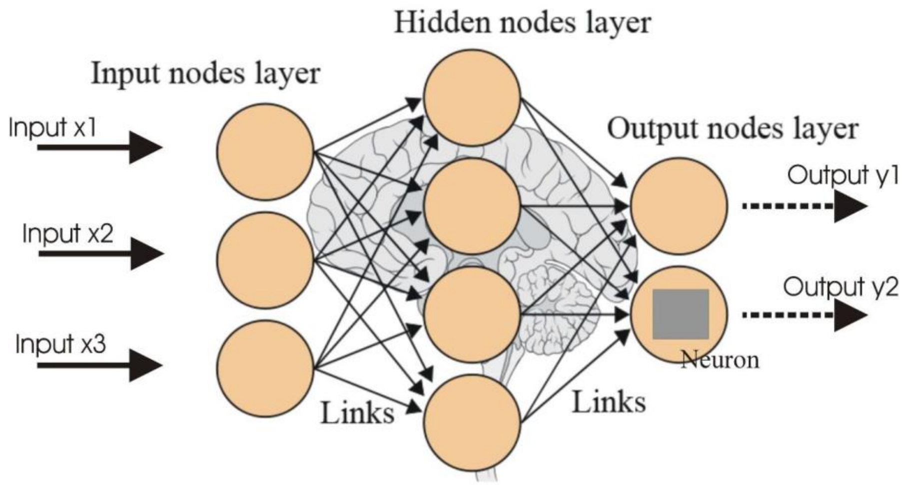


Figure 18. – Neural Networks architecture [77]

Multiple layers are much better at generalizing because they learn all the intermediate features between the raw data and the high-level classification.

Specialized neural networks have emerged for different use cases, for example, convolutional neural networks (CNN), which pre-process and tile image regions for improved image recognition. Conversely, recurrent neural networks add a hidden layer that is connected with itself for better speech recognition applications. Promising advances have been made in automatically learning features (also referred to as representation learning), through auto-encoders, sparse coders and other techniques. This is particularly important as labelled data is difficult to obtain and the costs for feature engineering are high [78].

AI is the present and the future. Deep Learning is showing great promise when it comes to developing the automated, self-teaching systems which are revolutionizing not only products, many industries as well.

2.5.3.2. Convolutional Neural Networks

Convolutional Neural Networks (CNNs), also named ConvNets, are arguably the most popular Deep Learning architecture and have played an important role in the history of Deep Learning. ConvNets were some of the first deep models to perform well. ConvNets were some of the first working deep networks trained with backpropagation. Through backpropagation, losses calculated during training can be used to stimulate changes in network parameters, allowing the mastering of complex and varied tasks [115]. It is not entirely clear why convolutional networks succeeded when general backpropagation networks were considered to have failed. ConvNets models carried the torch for the rest of Deep Learning and paved the way to the acceptance of neural networks in general.

ConvNets were also some of the first neural networks to solve important commercial applications and remain at the forefront of commercial applications of Deep Learning today. For example, in the 1990s, the neural network research group at AT&T developed a convolutional network for reading checks. By the end of the 1990s, this system deployed by NEC was reading over 10 percent of all the checks in the United States. Later, several OCR and handwriting recognition systems based on convolutional networks were deployed by Microsoft [79].

The most recent interest in ConvNets started with AlexNet in 2012 and it has grown exponentially ever since. In a few years, researchers have some progress about the depth of the layers, from 8-layer AlexNet to 152-layer ResNet [72].

ConvNets provide a way to specialize neural networks to work with data that has a clear grid-structured topology and to scale such models to very large size. This approach has been the most successful on a two-dimensional image topology [79]. To process one-dimensional sequential data, another powerful specialization of the neural networks framework is used: recurrent neural networks; nevertheless, those type of networks will not be included in this research.

ConvNets are now the go-to model on every image related problem (i.e., object detection, classification, segmentation). ConvNets are successfully applied to some tasks like recommender systems, natural language processing, and other tasks. The main advantages of ConvNets in comparison to their predecessors are two: The first one is that the model automatically detects the important features without any human supervision. For example, when the model is fed with many pictures of cats and dogs, the model learns distinctive and important features for each class by itself. The second advantage is that ConvNets are computationally efficient too.

2.5.3.2.1. Architecture Overview

ConvNets take advantage of the fact that the input consists of images and they constrain the architecture more sensibly. The layers of a ConvNet are arranged, unlike a regular NN, in 3 dimensions: width, height, depth (depth, in this case, refers to the third dimension of an activation volume, not to the total numbers of layers in the network). This means that the input images in whatever dataset, are an input volume of activations, and the volume has dimensions width x height x depth. The neurons in a layer will only be connected to a small region of the layer before it, instead of all the neurons in a fully-connect manner. For instance, if the image dataset used for training the model has 10 classes (i.e., CIFAR-10 [91]), the final output layer would have dimensions 1 x 10 x 10 for that image dataset. This is because, by the end of the ConvNet, the architecture will reduce the full image into a single vector of class scores, arranged along the depth dimension [80]. The figure 19 is a visualization of the latter:

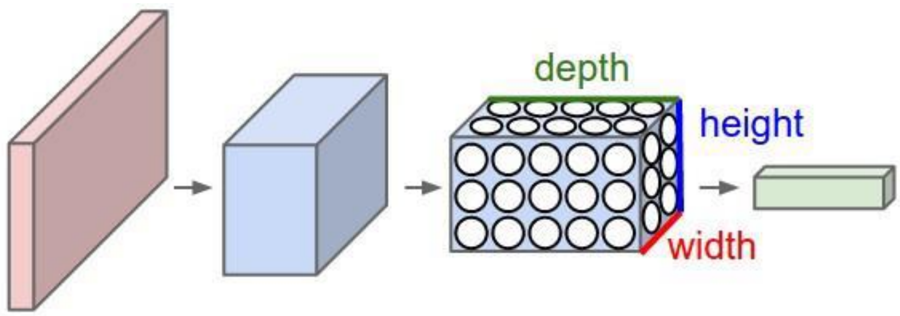


Figure 19. - Explanation of the activation volume [80]

A simple ConvNet is a sequence of layers, and every layer of a Convolutional Network transforms one volume of activations to another through a differentiable function.

To build ConvNets architectures, three main types of layers are used:

- Convolutional Layer.
- Pooling Layer.
- Fully-Connected Layer.

The layers will be stacked to form a full Convolutional Network architecture. A simple ConvNet for CIFAR-10 classification will be used as an example architecture: [INPUT - CONV - RELU - POOL - FC].

- The INPUT will hold the raw pixel values of the image, For this case an image of width 32, height 32 and with three colour channels (R,G,B).
- The CONV layer will compute the output of neurons that are connected to local regions in the input. Each computing is a dot product between their weights and a small region which they are connected to in the input volume. If it is decided to use a filter of size 12, this may result in volume such as [32x32x12].
- The RELU (REctified Linear Unit) layer will apply an elementwise non-saturating activation function, such as the *f(x) = max(0,x),* thresholding at zero. It has been shown that the network can train several times faster using this non-saturating function, as compared to using saturating functions such as *f(x) = tanh(x)* or the sigmoid function [114]. Also, this function leaves the size of the volume unchanged ([32x32x12]).
- The POOL layer will perform a down sampling operation along the spatial dimensions (width, height), resulting in volume such as [16x16x12].
- The FC (i.e. fully-connected) layer will compute the class scores, resulting in volume of size [1x1x10], where each of the 10 numbers correspond to a class score, such as among the 10 categories of CIFAR-10 (also the MNIST has 10 classes). As with ordinary Neural Networks and as the name implies, each neuron in this layer will be connected to all the numbers in the previous volume.

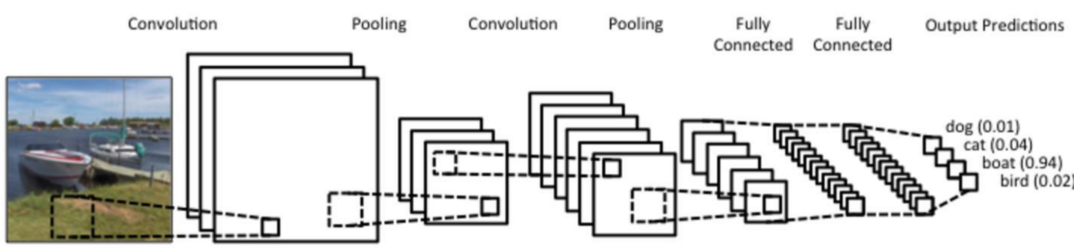


Figure 20. - Convolutional Neural Network architecture [81].

Thereby, ConvNets transform the original image layer by layer from the original pixel values to the final class scores. Some layers contain parameters and other do not. The CONV/FC layers perform transformations that are a function of not only the activations in the input volume, but also of the parameters (the weights and biases of the neurons). The RELU/POOL layers will implement a fixed function. The parameters in the CONV/FC layers will be trained with gradient descent so that the class scores that the CNN computes are consistent with the labels in the training set for each image [80].

The main building block of Convolutional Networks is the convolutional layer. Convolution is a mathematical operation to merge two sets of information. In this case, the convolution is applied to the input data using a convolution filter to produce a feature map [82].

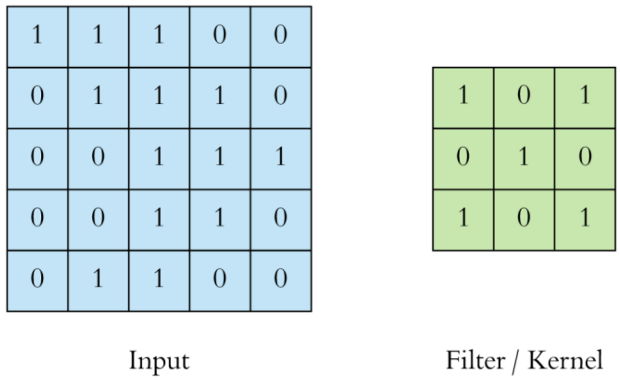


Figure 21. – Convolutional layer (Input) and a 3x3 Convolutional Filter (Kernel) in 2D [82]

The convolution operation is performed by sliding the filter over the input. At every location, element-wise matrix multiplication is done and after that, sum the result. This sum goes into the feature map. The green area in figure 21, is where the convolution operation takes place. This area is called the receptive field. The receptive field will have the size of the filter, in this case 3x3.

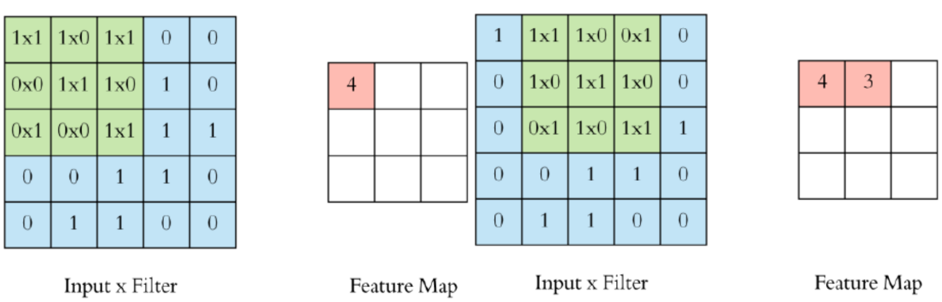


Figure 22. – Convolution operation in 2D [82].

Figure 22 is an example of a convolution operation in 2D using a 3x3 filter. These convolution operations are performed in 3D. In the real world, an image is represented as a 3D matrix with dimensions of height, width and depth, where depth corresponds to colour channels (R, G, B). A convolution filter has a specific height and width, like 3x3 or 5x5, and by design, it covers the entire depth of its input, so it needs to be 3D as well. Multiple convolutions are performed on an input, each using a different filter and resulting in a distinct feature map. After that, all these feature maps are stacked together, and that becomes the final output of the convolution layer [82].

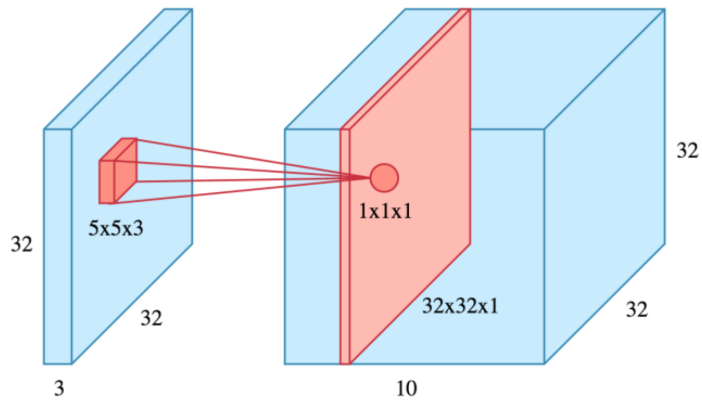


Figure 23. – Convolution in 3D [82].

In this case (figure 23), the input is a 32x32x3 image, and a filter of size 5x5x3 is used (the depth of the convolution filter must match the depth of the image, in this case both are 3). When the filter is at a certain position, it covers a small volume of the input and the program performs the convolution operation. The filter slides over the input like in 2D and performs the convolution at every location aggregating the result in a feature map. This feature map is of size 32x32x1, shown as the red slice on the right.

If the architecture uses 10 different filters, the output would have 10 feature maps of size 32x32x1 and stacking them along the depth dimension would give the final output of the convolution layer: a volume of size 32x32x10 (large blue box on the right). If a padding it is used, the height and width of the feature map are unchanged and still the same size, in this case 32.

The stride and padding specify how much the filter would be moved at each step. Usually the stride value is 1. It is possible to have bigger strides if it is desired to have less overlap between the receptive fields. The result of using a bigger stride is having a feature map smaller since potential locations are skipped over. The padding is used if it is desired to maintain the same dimensionality. There are 2 options, either pad with zeros or the values on the edge. When the padding surrounds the input with zeros, this is usually called zero padding.

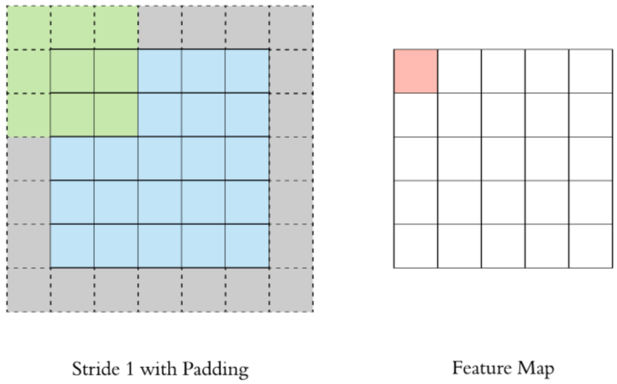


Figure 24. - Stride 1 with padding [82].

In figure 24, the grey area around the input is the padding. When padding is used, the dimensionality of the feature map matches the input. Padding is commonly used in ConvNets to preserve the size of the feature maps. Otherwise they would shrink at each layer, which is not desirable. The 3D convolution operation showed above is using padding with a stride of 1, in this way the height and width of the feature map were the same as the input (both 32x32), and only the depth changed.

After a convolution layer, it is usually performed the pooling, this helps to reduce the dimensionality. The pooling enables to reduce the number of parameters, which both shortens the training time and combats overfitting. Pooling layers down sample each feature map independently, reducing the height and width, keeping the depth intact. The most common type of pooling is “max pooling” which just takes the max value in the pooling window.

Pooling has no parameters. It slides a window over its input, and simply takes the max value in the window. The window size and stride are specified in the pooling operation. By halving the height and the width, the number of weights is reduced to 1/4 of the input. Considering that is typically dealt with millions of weights in ConvNets architectures, the reduction operation helps to reduce the computer processing times.

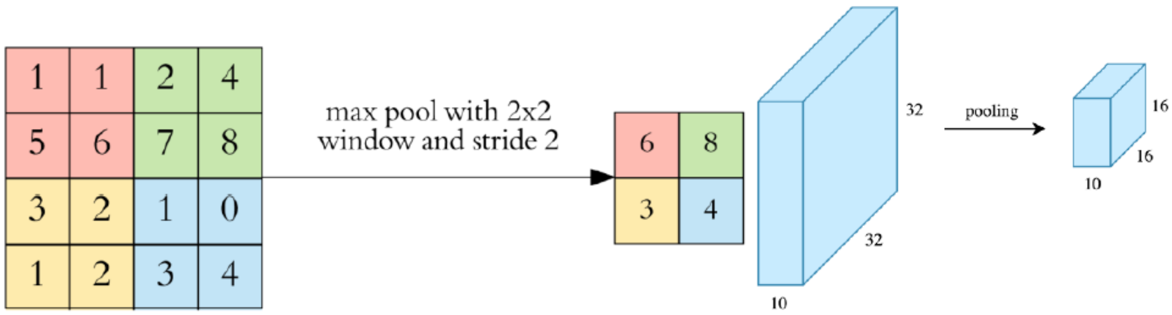


Figure 25.- Using max pooling with 2x2 window and stride with a value of 2 [82]

The windows are no overlapping because the window size and stride are 2. If the input to the pooling layer has the dimensionality 32x32x10, using the same pooling parameters described above, the result will be a 16x16x10 feature map. Both the height and width of the feature map are halved, but the depth does not change because pooling works independently on each depth slice the input.

After the convolution plus pooling layers, a couple of fully connected layers is added to wrap up the ConvNet architecture, the last layer is the fully connected layer. The output of both convolution and pooling layers are 3D volumes, but a fully connected layer expects a 1D vector of numbers. The output of the of the final pooling layer needs to be flattened to a vector, and that vector becomes the input to the fully connected layer. Flattening is simply arranging the 3D volume of numbers into a 1D vector [80].

A ConvNet model can be considered as a combination of two components: feature extraction part and the classification part. The convolution plus pooling layers perform feature extraction. The fully connected layers then act as a classifier on top of these features and assign a probability for the input image being a dog, a cat, a cheetah, etc. [82].

Topics like batch normalization, vanishing gradients, dropout, initialization techniques, non-convex optimization, biases, choices of loss functions, data augmentation, regularization methods, computational considerations, backpropagation, modifications of backpropagation, and a few more were not discussed in this research, but are important topics in the design and performance of Deep Learning architectures and models.

2.5.3.2.2. Case Studies

There are several architectures in the field of Convolutional Networks that have a name. The most common are [80]:

- **LeNet.** The first successful applications of ConvNets were developed by Yann LeCun in 1990's. The best known is the LeNet architecture and it was used to read zip codes, digits, etc.
- **AlexNet.** The network developed by Alex Krizhevsky, Ilya Sutskever and Geoff Hinton, popularized ConvNets in Computer Vision. AlexNet had very similar architecture to LeNet, but was deeper, bigger, and featured Convolutional Layers stacked on top of each other (before this model, it was common to only have a single CONV layer always immediately followed by a POOL layer).
- **ZF Net.** The ILSVRC [92] 2013 winner was a ConvNet from Matthew Zeiler and Rob Fergus, the ZFNet (short for Zeiler & Fergus Net). This model was an improvement on AlexNet by tweaking the architecture hyperparameters.
- **GoogLeNet.** The ILSVRC 2014 winner was a ConvNet from Szegedy et al. from Google. Its main contribution was the development of an Inception Module that dramatically reduced the number of parameters in the network (4M, compared to AlexNet with 60M). Also, this paper uses Average Pooling instead of Fully Connected layers at the top of the CNN, eliminating a large number of parameters that do not seem to matter much. There are also several follow-up versions to the GoogLeNet, most recently Inception-v4 [93].
- **VGGNet.** The network developed by Karen Simonyan and Andrew Zisserman is the VGGNet. The main contribution of this model was in showing that the depth of the network is a critical component of good performance. Their final best network contains 16 CONV/FC layers, featuring an extremely homogeneous architecture that only performs 3x3 convolutions and 2x2 pooling from the beginning to the end. This model is more expensive to evaluate and uses a lot more memory and parameters (140M). Most of these parameters are in the first fully connected layer, and it was since found that these FC layers can be removed with no performance downgrade, significantly reducing the number of necessary parameters.

- **ResNet.** The Residual Network was developed by Kaiming He et.al. This model was the winner of ILSVRC 2015. It features special *skip connections* and heavy use of batch normalization. The architecture is also missing fully connected layers at the end of the network. ResNets are currently by far state of the art ConvNets models and are the default choice for using ConvNets in practice.

2.5.3.3. Transfer Learning

In practice, very few people train an entire ConvNet from scratch (with random initialization), because it is relatively rare to have a dataset of sufficient size. Instead, it is common to pre-train a ConvNet on a very large dataset (e.g., ImageNet [94], which contains 1.2 million images with 1000 categories), and then use the ConvNet either as initialization or a fixed feature extractor for the task of interest. The three major Transfer Learning scenarios are [83]:

- **Convolutional Neural Network as fixed feature extractor**. Take a ConvNet model pre-trained on any database (i.e., ImageNet [94]), remove the last fully-connected layer and then treat the rest of the Convolutional Network as a fixed feature extractor for the new dataset. After that, train a classifier that aligns to the system new task for the new dataset.

- **Fine-tuning the Convolutional Neural Network**. This method is to not only replace and retrain the classifier on top of the ConvNets on the new dataset but to also fine-tune the weights of the pre-trained network model by continuing the backpropagation. It is possible to fine-tune all the layers of the ConvNet, or it is possible to keep some of the earlier layers fixed (due to overfitting concerns) and only fine-tune some higher-level portion of the network. This is motivated by the observation that the earlier features of a Convolutional Network contain more generic features (e.g. edge detectors or colour blob detectors) that should be useful to many tasks, but later layers of the model become progressively more specific to the details of the classes contained in the original dataset. For example, in the case of ImageNet, this dataset contains many dog breeds. Therefore a significant portion of the representational power of the ConvNet may be devoted to features that are specific to differentiating between dog breeds.

- **Pre-trained models.** Since modern ConvNets take from 2 to 3 weeks to train across multiple GPUs on many big datasets (i.e. ImageNet), it is common to see people release their final ConvNets checkpoints for the benefit of others who can use the networks for fine-tuning. For example, the Caffe and Tensorflow libraries have a “model zoo” where people share their network weights.

2.5.3.4. Object Detection

In a daily basis, a person classifies and detects hundreds of objects. The human perception allows undertaking complex, dynamic tasks such as driving in a seemingly easy effort. Emulating the efficiency with which the human brain is processing the environment at driving, is a huge step in the current development of Self-Driving cars functionalities and Advanced Driver-Assistance Systems. With the rise of these systems, fast and accurate models for object detection are rising in demand.

Object detection refers to the capability of a system to locate and identify each of the objects in an image or scene. In the automotive industry, object detection is a core part of many collision warning systems. This task requires to give information about the location of a classified object(s) usually in the form of bounding box coordinates in the image pixel grid. In the beginning, ConvNets were originally shown to have promising results for classification tasks. However, some modifications were made to the most successful approaches of image classification models, allowing these new models to perform object detection tasks.

Implementing a model which is capable of predicting the two-dimensional location of objects in the road, can be integrated into the final concept functions. This could allow the possibility that the general performance of the system will be improved. Nevertheless, factors like robust performance, high inference speed, and memory requirements are really important to considerate in automated object detection systems.

2.5.3.4.1. Models

2.5.3.4.1.1. R-CNN

The Region-based CNN [120] combines the Selective Search [116] algorithm for scanning the image and create region proposals, then uses a ConvNet to find and classify objects in the proposed regions. Then, the last step is to take the output of each ConvNet and feed it into a Support Vector Machine (SVM) to classify the region and a linear regressor to adapt and tight the shapes of the bounding box of the object, if the object exists [121]. This model is trained on 2012 ImageNet dataset. However, this model has a high computational cost making it very slow. Figure 26 describes the steps taken.

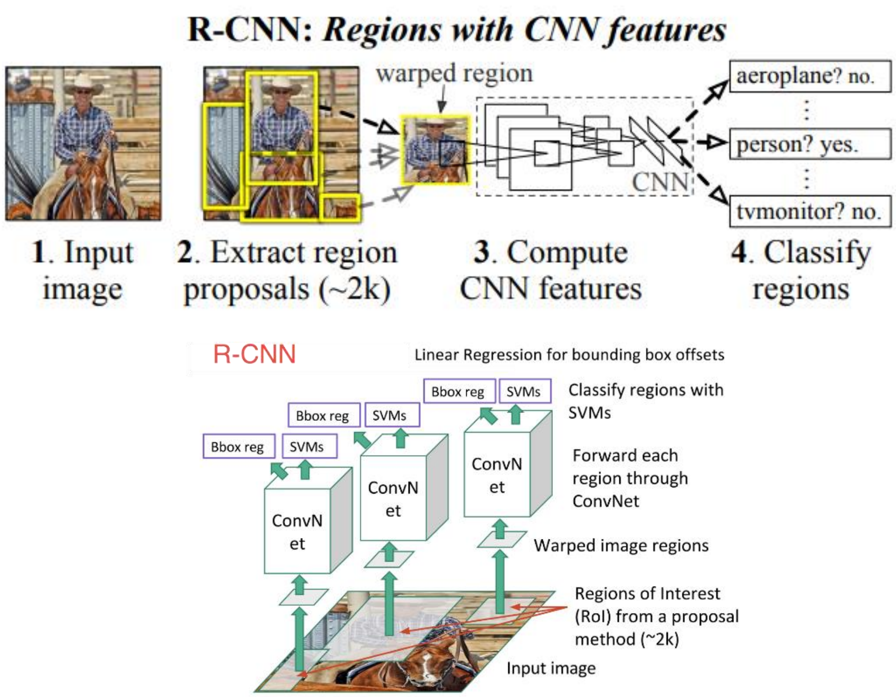


Figure 26.- R-CNN architecture [120, 121].

2.5.3.4.1.2. Fast R-CNN

The Fast Region-based CNN was the immediate successor of R-CNN. The purpose of this model was to improve the high computational cost related to the high numbers of models necessary to analyse all-region proposals. One of the improvements was the Region-of-Interest (ROI) pooling layer. In this model, the entire image is first processed by a ConvNet (feature extractor), producing a single set of features maps for the whole image. Then, these feature maps are the input for the fully connected layers and output a SoftMax classifier, which displays classes and confidence scores. Lastly, the linear regressor outputs the bounding-boxes [122].

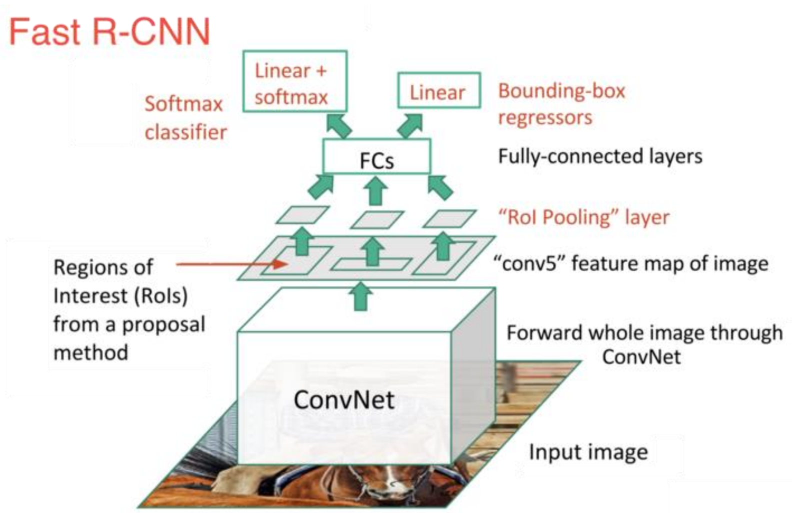


Figure 27.- Fast R-CNN improvement description [121].

2.5.3.4.1.3. Faster R-CNN

Faster R-CNN is composed of two modules. The first module is the Region Proposal Network (RPN) which proposes regions, and the second module is the Fast R-CNN detector. RPN is a Fully Convolutional Network (FCN) and can be trained end-to-end to specifically generate detection proposals. A window of size 3x3 slides all the feature maps and outputs a features vector linked to two fully-connected layers, one for box-regression and one for box-classification. Multiple region proposals are predicted by the fully-connected layers. A maximum of "k" regions is fixed thus the output of the box-regression layer has a size of 4k (coordinates of the boxes, their height and width) and the output of the box-classification layer size of 2k (“objectness” scores to detect an object or not in the box).

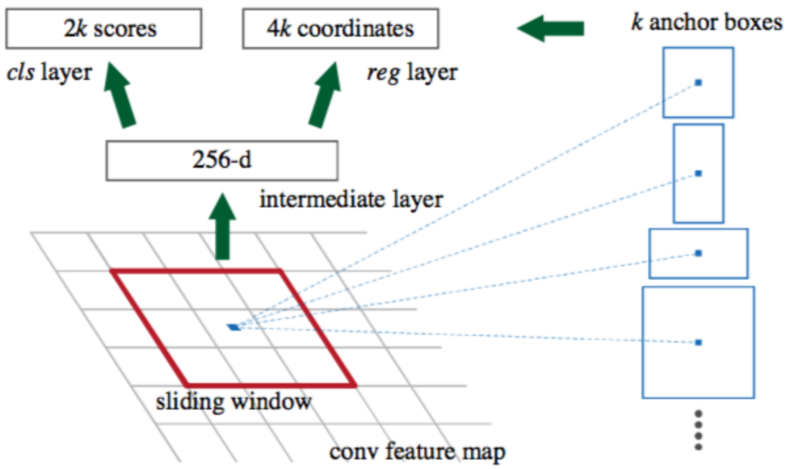


Figure 28.- Sliding window location and description [121]

The “k” region proposals detected by the sliding window are called anchors. When the anchor boxes are detected, they are selected by applying a threshold over the “objectness” score to keep only the relevant boxes. These anchor boxes and the feature maps computed by the initial CNN model feeds a Fast R-CNN model [123, 125]. Figure 29 shows the Faster R-CNN architecture.

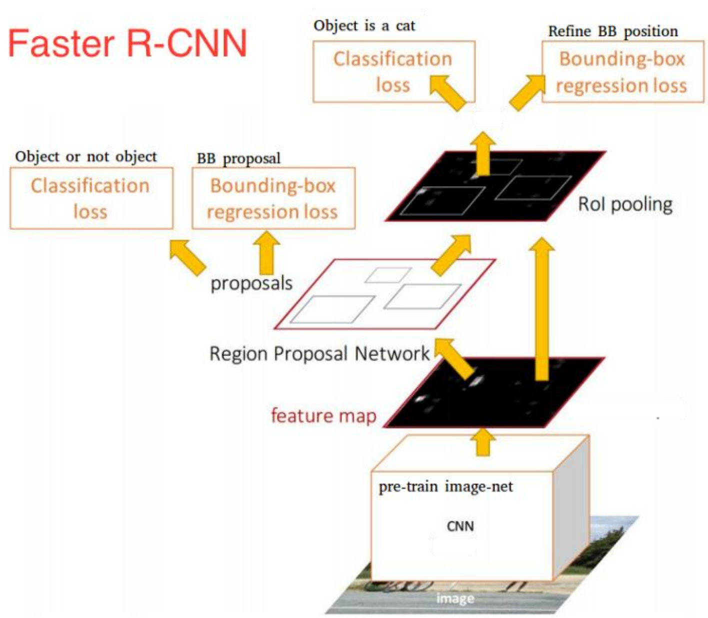


Figure 29.- Faster R-CNN architecture [121]

2.5.3.4.1.4. Mask R-CNN

Mask R-CNN [126] is an intuitive extension of Faster R-CNN. This model extends Faster R-CNN to object instance segmentation. It is added a third branch to the Faster R-CNN model, this branch is for predicting an object mask in parallel with the other 2 existing branches, the RPN (localisation) and the classifier. It uses RPN to generate bounding boxes proposals and produces the three outputs at the same time for each Region of Interest (RoI). The mask branch is a small fully-connected layer applied to each RoI, predicting a segmentation mask in a pixel-to-pixel manner.

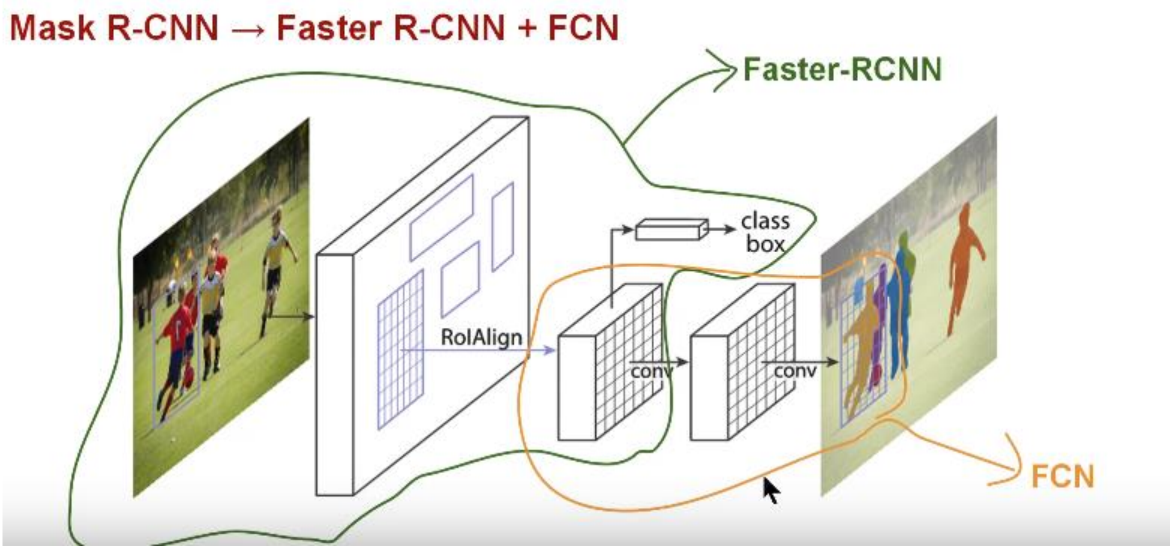


Figure 30.- Mask R-CNN architecture [127]

Pixel-level segmentation requires more fine-grained alignment than bounding boxes. The Mask R-CNN model proposes the RoIAlign layer, which improves the RoI pooling layer. The RoIAlign layer provides scale-equivariance and translation-equivariance with the region proposals, with this the RoI is better mapped into the original image [127].

2.5.3.4.1.5. SSD

The Single Shot Detector (SSD) [124] provides enormous speed gains over Faster R-CNN. These models predict all at once the bounding boxes and the class confidence with and end-to-end ConvNet architecture. This model passes through the input image, multiple convolutional layers with different sizes of the filter (3x3, 5x5, 10x10). Features maps from convolutional layers at different positions of the network are used to predict the bounding boxes. After this, they are processed by specific convolutional layers with 3x3 filters called “extra feature layers” to produce a set of bounding boxes similar to the anchor boxes of Fast R-CNN model.

Each box has 4 parameters: The width, height, and coordinates of the centre. It produces a vector of probabilities corresponding to the confidence over each class of object [125].

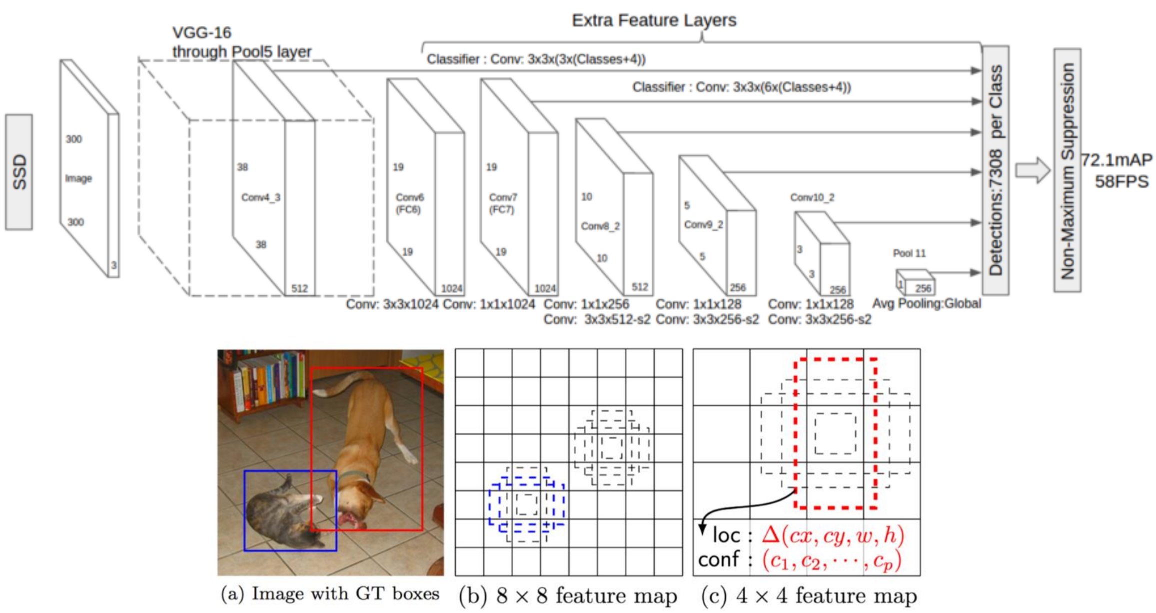


Figure 31.- SSD architecture [124]

The Non-Maximum Suppression (NMS) method is also used at the end of the SSD model to keep the most relevant bounding boxes, this means that groups together highly over-lapping boxes into a single box. The Hard-Negative Mining (HNM) technique is used due to the large number of negative boxes predicted. This technique consists in selecting only a subset of negative examples (boxes) with the highest training loss (i.e., false positives) during the training. The boxes are ordered by confidence, and the top is selected depending on the ratio between the negative and the positive which is at most 1/3 [121,125].

Ideally, confidence scores can be used to estimate the probability that an object lies within the bounding box. Confidence in this context should not be confused with the notion of confidence in the context of statistics (e.g., confidence intervals). These scores are normalised between 0-1 and are often erroneously referred to as probabilities in the literature (e.g. [123]).

2.5.4. Deep Learning Libraries

Figure 32 describes the different layers than can be found in a Deep Learning system. Nowadays, Graphic Processing Units (GPUs) have been proven to scale NNs in a good way but are presenting limitations when the input images have a large size. Some libraries rely on GPUs for optimizing and running the training of NNs, but other libraries have CPU and GPU versions. Some higher-level Deep Learning libraries emerged in different languages like Python, Java, R and among others. Examples of these libraries are shown in figure 32.

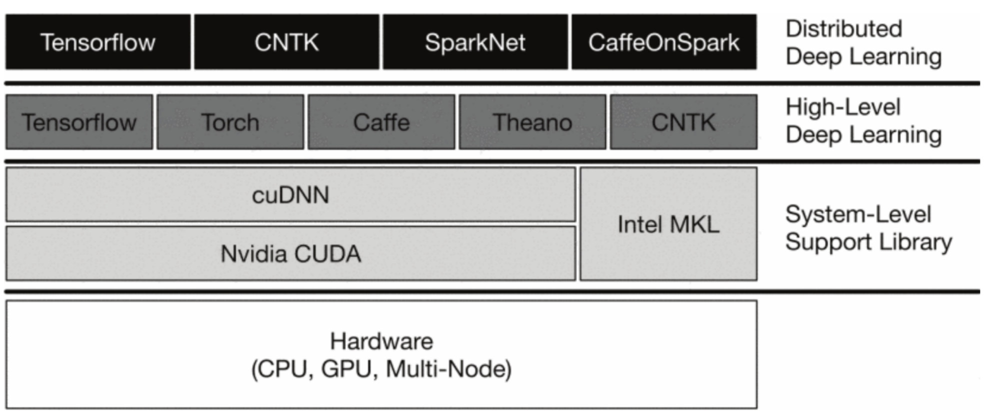


Figure 32. - Deep Learning software and hardware [78].

2.5.5. Cloud Services

Cloud computing provides comprehensive services for computing processing, backend services for applications and data storage services. For these reasons, it is becoming increasingly a viable platform for developing and implementing end-to-end

Deep Learning products or applications. Lately, an increasing number of Infrastructure-as-a-Service (IaaS) offerings with GPU support exists: for example, Amazon Web Services (AWS) [95] provide the hardware necessary for deep training and exploration while removing the necessity of obtaining a physical system for computation. All services such as GPU computing and data storage utilize the cloud and can, therefore, be managed accordingly. AWS is not the only provider with IaaS, it is possible to find similar systems with similar capabilities with other companies like Microsoft Azure Cloud [97] or Paperspace [96]. Google provides CLOUD AI [98] which is a managed PaaS environment.

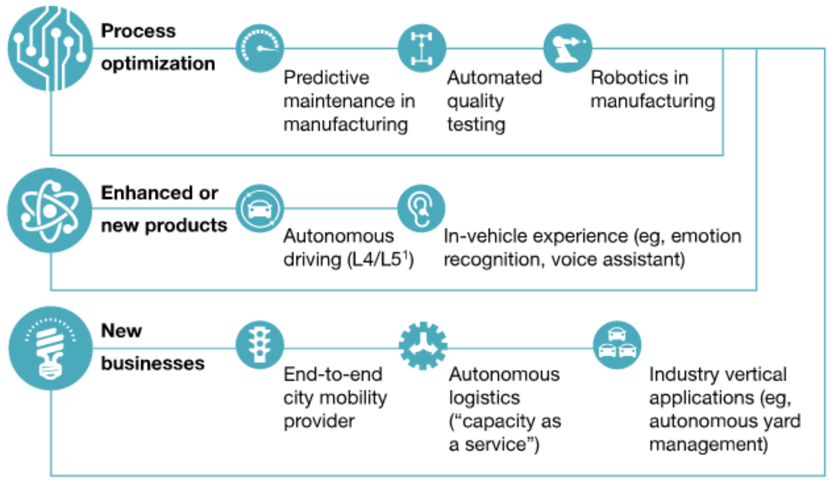
Google was one of the first companies to offer services in the cloud about Machine Learning classifications and predictions, Microsoft and Amazon offer similar services as well. Those type of services allow fast and simple access and deployment of Machine Learning, there are several models available for deployment. It is common that those companies often provide black-box models with limited calibration and modifications of the model [78].

2.5.6. Machine Learning and Deep Learning in the automotive industry

Machine Learning has the potential to unlock completely new product offerings which could improve productivity and build these products on consumer interest. Thereby, Machine Learning will create new value pools which could drive the Artificial Intelligence technology's impact on the automotive industry [74].

Despite all the technological developments, the automotive/mobility industry is still only at the beginning of the AI disruption. State-of-the-art AI applications remain in the narrow realm of the AI (performing better than humans but only in very specific tasks). Matching human ability in an even larger number of contexts is still some years out.

Deep Learning architectures have a wide application field in the automotive industry inside and outside the car. Those applications can involve topics like visual inspection in manufacturing, predictive maintenance, and topics about user experience as social analytics, customization settings, Human-Computer Interaction (i.e., conversational user interfaces), Automated Driving, among others [84].



[1]Level 4 (high automation) or level 5 (full automation), according to SAE International definitions laid out in SAE J3016.

Figure 33. - Different opportunity areas in the automotive industry [84]

A common challenge of these applications is the need for storage and processing of large volumes of data as well as the need to deal with unstructured data as videos, images or text. As an example, the images from camera-based sensors on machines in manufacturing and assembly processes.

2.5.6.1. Deep Learning in Self-Driving Vehicles

Deep Learning models are already starting to complement and, in some cases, replace rule-based systems in ADAS modules. The main goal is to design and deploy fully integrated, learning-based systems enhanced by AI algorithms. Sensor processing, data interpretation, planning, decision making, and execution are the major steps needed to fulfil the goal [72].

Eventually, the systems will learn from all new situations they encounter. Therefore the performance is continuously improving. Also, the systems will be able to share their knowledge and learnings in a centralized platform or other methods as direct interaction. Ideally, the accumulated knowledge of all the systems could be used to enhance the performance of each individual vehicle of the market.

Companies like Google and Tesla, used in their test vehicles the first hybrid systems that add self-learning elements to conventional systems. Several automotive start-ups aim at extending the usage of AI, some of these start-ups are Argo.ai [101], Drive.ai [103], nuTonomy [102], Otto [106], Preferred Networks [105] and Zoox [104].

These companies are pursuing L4 and L5 automated vehicles [100], which aim at operating vehicles without restrictions in any environment.

As humans started to hand off their decision making to machines, the interaction between machines will become more important [72]. It is important to mention that advancements in self-driving cars are closely correlated with those of machine-to-machine interactions.

“Automated Driving holds the promise of a smoother, safer, and more comfortable mobility experience” [72]

3. Concept Development Process

This Chapter is confidential due to a Non-Disclosure Agreement. The complete document is kept in the Centre of Mechanical Design and Technological Innovation of the National Autonomous University of México (UNAM). For more details, please contact Dr. Alejandro C. Ramírez Reivich.

4. Concepts

This chapter is confidential due to a Non-Disclosure Agreement. The complete document is kept in the Centre of Mechanical Design and Technological Innovation of the National Autonomous University of México (UNAM). For more details, please contact Dr. Alejandro C. Ramírez Reivich.

5. Final Concept

This chapter is confidential due to a Non-Disclosure Agreement. The complete document is kept in the Centre of Mechanical Design and Technological Innovation of the National Autonomous University of México (UNAM). For more details, please contact Dr. Alejandro C. Ramírez Reivich.

II. Second Stage: Igniting the spark with AI.

Following some ideas of the first concepts, Artificial Intelligence is an excellent way to improve the final thought. Deep Learning is a trending Machine Learning technique which has shown good performance for different tasks on several application fields. By using Deep Learning techniques, it is possible to improve some of the final concept functions and upgrade the overall performance of the system.

The objective of this stage was to test some Deep Learning models and set the foundations of how it will be the best option to increase the reliability of the system. Therefore, creating a new State of the Art model is not on the scope of this stage.

In this section, the results of the testings and deployment of some models will be shown along with the proposal of how to create a new dataset with new “hazardous” objects of interest as potholes, construction sites, and other situations which must be labelled personally by the user.

6. Data & Deep Learning Models

As introduced in section 2.5, Deep Learning is guiding the way that automotive safety systems are designed. However, it is needed that these systems are in a continuous improvement process, and one of the key-elements is that the models keep learning new elements on the driving scene. The foundations of the process how to achieve this will be described in the next sub-sections.

6.1. Datasets

There are several datasets for object detection challenges. The most common are PASCAL Visual Object Classification (Pascal VOC) [117], ImageNet [94] and Common Objects in COntext (COCO) [118,119]. These datasets include daily basis objects from bottles, dogs, birds, surfing boards, persons to cars and buses.

Another dataset is the KITTI Dataset [150,151]. This dataset includes diverse scenarios from real-world traffic situations, ranging from freeways over rural areas to inner-city scenes including several dynamic and static objects. Usually, this dataset is used in mobile robotics and Automated Driving research. In this case, the datasets selected are the KITTI and COCO Datasets, which include categories of the driving environment.

6.2. Tests & Results

The objective is not to design a new state of the art model or improve one state of the art model. In this case, the main purpose is to select and test some pre-trained Deep Learning models which can improve the main functions of the final concept, for example, the object detection task for Front Collision Warning and the creation and maintenance of the training dataset. In this stage, transfer learning was used with the Tensorflow Object Detection API [149,152].

This API provides a collection of detection models called the Tensorflow detection model zoo, pre-trained on the COCO dataset, the KITTI Dataset and the Open Images dataset [153]. Also, these models can be useful for initializing the customized models when training on the novel dataset. The selected COCO and KITTI trained models are shown in figure 88. In this API, it is important to mention that the KITTI dataset only has two classes, pedestrians, and cars.

COCO-trained models {#coco-models}

Model name	Speed (ms)	COCO mAP[^1]	Outputs
faster_rcnn_inception_resnet_v2_atrous_coco	620	37	Boxes
faster_rcnn_nas	1833	43	Boxes
faster_rcnn_resnet101_coco	106	32	Boxes
mask_rcnn_inception_resnet_v2_atrous_coco	771	36	Masks
mask_rcnn_resnet101_atrous_coco	470	33	Masks

Kitti-trained models {#kitti-models}

Model name	Speed (ms)	Pascal mAP@0.5	Outputs
faster_rcnn_resnet101_kitti	79	87	Boxes

Figure 88.- Models used for inference.

As one of the objectives is the labelling of new objects on the road that are not present in the common datasets (i.e., potholes, construction sites, among other hazardous objects), the idea is to use the trained models to infer some classes that the model already know and add the new classes in an effective way. For this reason, the used models in this stage will perform high accuracy results, and it will not matter that the inference speed is slow.

Labelling is one of the most consuming tasks in the creation of the datasets. There are several computer programs which assist to label the datasets, one of them is the software LabelIMG [155]. The normal procedure is to have one image (road image in figure 89) and label (create bounding boxes) the different classes contained in that image.

Figure 89.- Image on LabelIMG software

Figure 89 shows one image before the annotation. It is possible to add as many classes as needed. Figure 90 shows the bounding boxes of two classes in the same image, car, and truck.

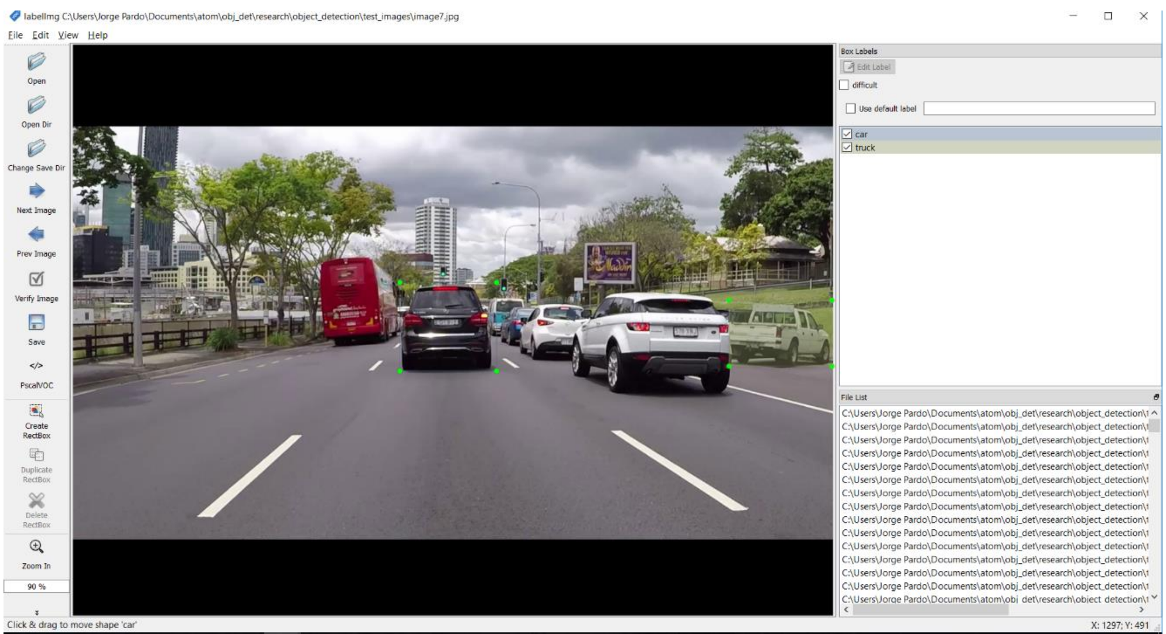


Figure 90.- Manual labelling

The problem relies on that this process is really slow. Let’s imagine that there is a dataset of ten videos with a duration of 15 minutes each video, and each video is recorded at 60 frames per second (FPS). This means that for only one video there are 54000 labelling frames (15 min * 60 secs * 60 FPS).

LabelIMG has the option to open and add to the current image an annotation file. This file must be in an XML format and it will have the coordinates of the bounding box and its class. By editing the code used in the API, it is possible to add the coordinates of the bounding boxes into an XML file and then read that file in LabelIMG.

A comparison of the pre-trained models was made to determine which is the best classifier model and fulfils the objective, label the dataset of the final concept. Elapsed time is not included in the comparison. Hence comparison between other methods only focus on the performance in object detection, regardless of run time. It was decided to use the best accuracy model regardless of the running time because it will reduce the time of labelling. Thus, the annotator only must include the desired new classes, increasing the efficiency of the labelling process.

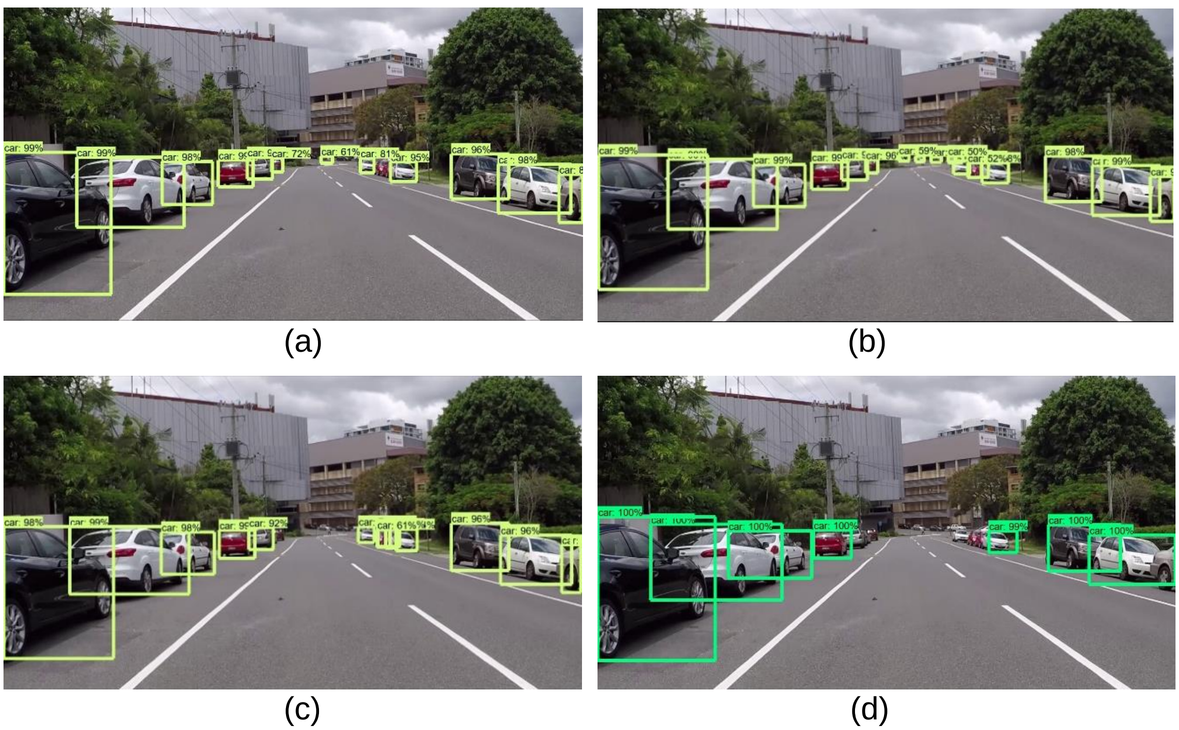
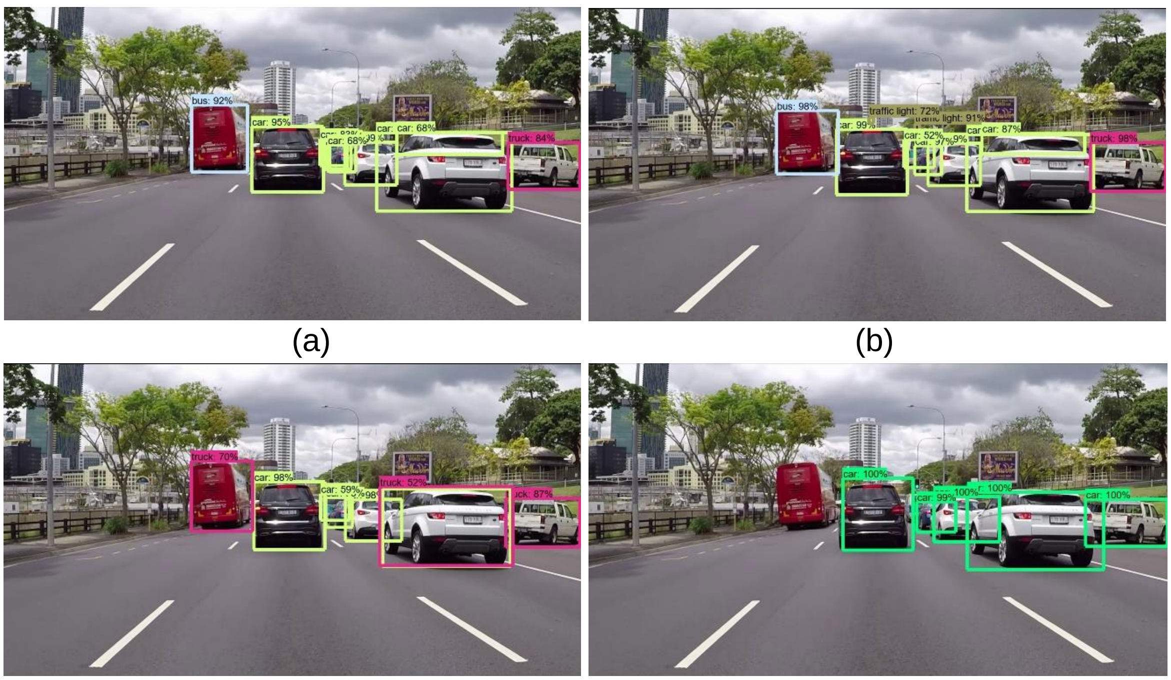


Figure 91.- Comparing Faster-RCNN models. (a) Faster-RCNN with inception-ResNet on COCO dataset, (b) Faster-RCNN with NAS on COCO dataset, (c) Faster-RCNN with ResNet101 on COCO dataset and (d) Faster-RCNN with ResNet101 on KITTI dataset

These models have a good performance at the accuracy and show high confidence levels. However, they have small differences between each model. Some show a better accuracy in some classes and show different results for the same classes. For example, in figure 92 it is possible to notice some of the differences that these models can have between each other.

In this image, the model trained with inception-ResNet is missing the traffic lights as well as the ResNet101 model. The ResNet101 model trained on the KITTI dataset is not detecting the bus because there is not a class defined as bus. Also the truck on the right is detected as car due to the same reason. The ResNet101 model trained on the COCO dataset is detecting the bus as a truck.



(a) (b) (c) (d)

Figure 92.- Comparison of Faster-RCNN models. (a) Faster-RCNN with inception-ResNet on COCO dataset, (b) Faster-RCNN with NAS on COCO dataset, (c) Faster-RCNN with ResNet101 on COCO dataset and (d) Faster-RCNN with ResNet101 on KITTI dataset

With the analysis of randomly frames from the video, around 300 randomly selected frames, it was decided that the best model to fit the purpose is the Faster-RCNN NAS [156], which achieved high confidence results in a city drive video.

As the COCO dataset has some variety of classes, there are some objects that could be detected as an “unnecessary object” in road situations. In figure 93, an example of this situation is shown. In this situation, there is an advertisement on the road from an international brewery, and the model detects the bottle in this ad.

Although the model makes the detection with a high level of confidence, this class is not necessary for the project purpose. Therefore this class can be kicked out on the next training stage.

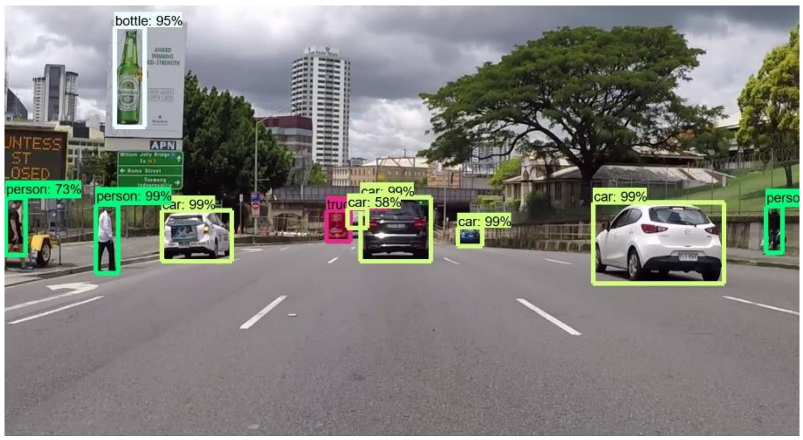


Figure 93.- Bottle detection.

In other cases, the model makes some mistakes that do not affect the purpose of the new dataset. This situation is shown in the next figure when the model detects a “potted plant” which is a tree, which is a big mistake but has no difference for the dataset purpose. As well, this class can be kicked out.

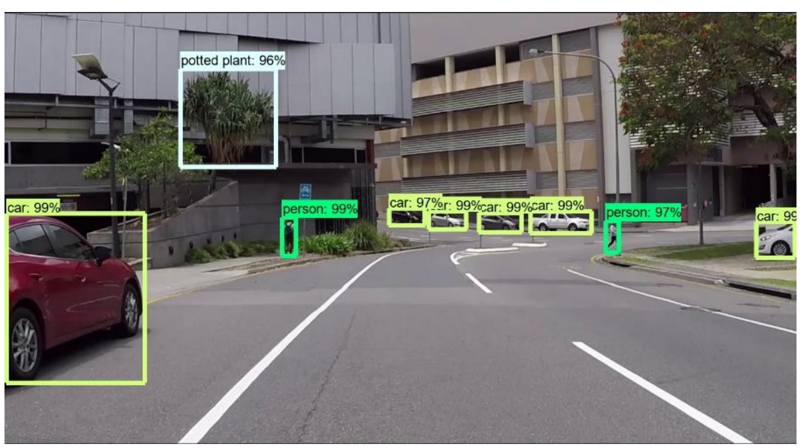


Figure 94.- Tree detected as a potted plant

These types of results, mean no danger for the driver or passengers safety, however, it may exist the possibility to gather all the types of objects that represent a hazardous situation for the driver, and create a new class called “Hazardous Object” with the condition that those objects must be on the road surface. Nevertheless, that possibility will not be addressed in this project. In other occasions, the model could miss objects, like the pedestrian in figure 95 (which is the next frame of figure 94) or have false positives when detecting cars at long distances. In those cases, the annotator will address that situation and label the missed objects.

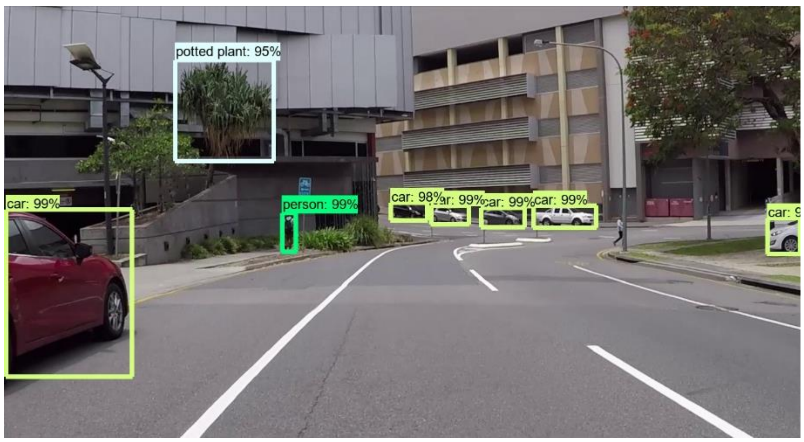


Figure 95.- In this frame, the model misses the pedestrian next to the light pole.

The selected model has good performance even with significant changes in the light conditions. Figure 96 shows the results with dark conditions and bright conditions.

Figure 96.- Results with contrasting light conditions.

After the object detection, the program will create an .xml file and will write the bounding boxes results (coordinates, class, and confidence) into the file. Consequently, every frame will have its annotation file which the annotator could add new classes if needed and/or delete bounding boxes of classes that are not part of the dataset's purpose. Therefore, this procedure will make the labelling task more efficient. The process is described in figure 97.

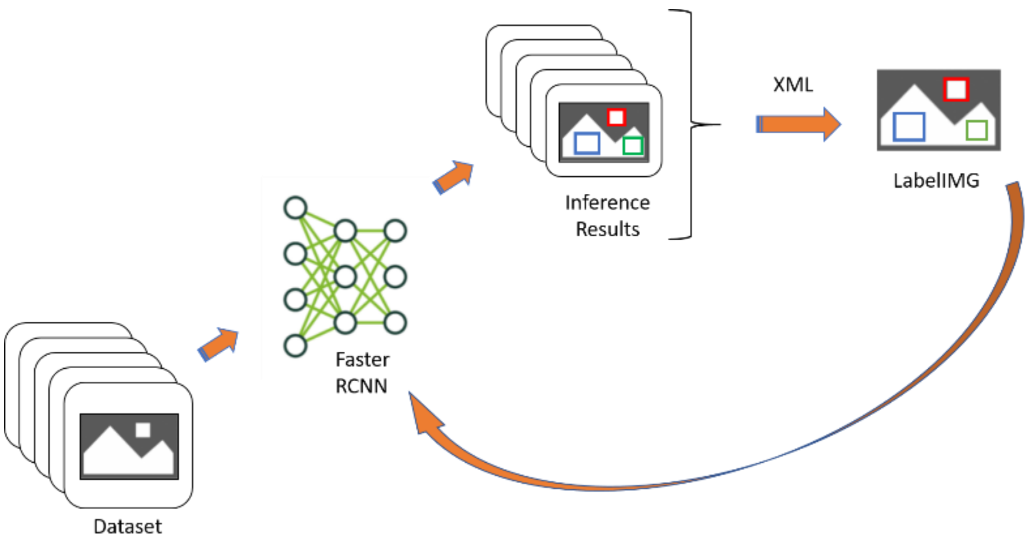


Figure 97.- Labelling workflow diagram.

The bright image in figure 96 will be used as an example of the annotation file for the LabelIMG software, due to the big number of labels that are on that image. In this case, the only excluded class was the "potted plant" and new classes were not added to the file. However, new classes as "bus stop sign" may be added to the dataset.

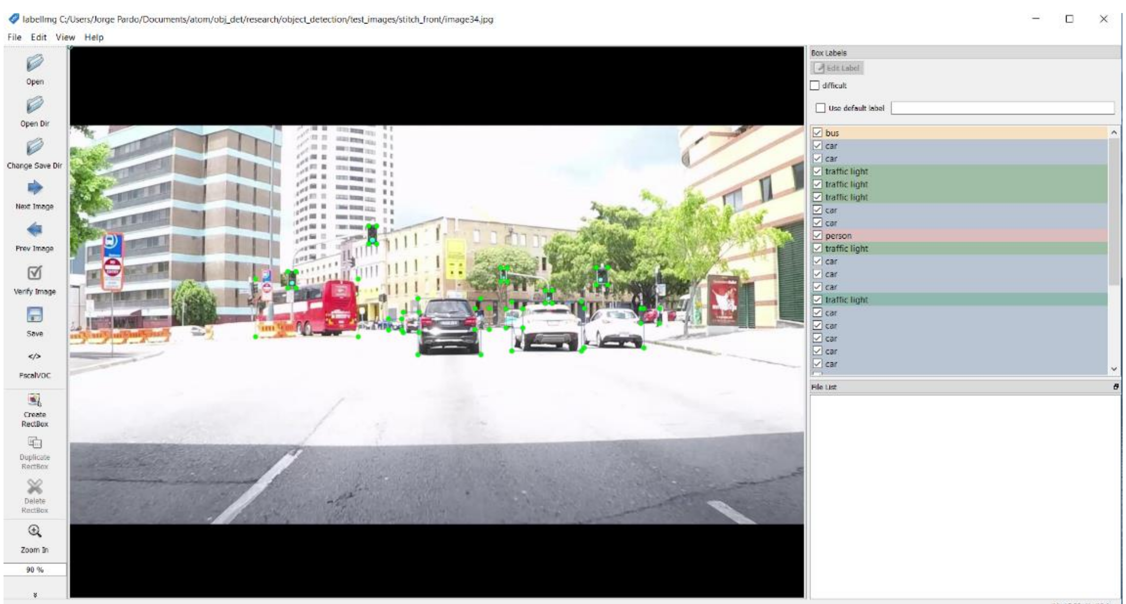


Figure 98.- Labels with bounding boxes neural network output.

Once that the labelling process has been completed, the next step is to use transfer learning. The idea is to use a pre-trained COCO model as a fixed feature extractor and then use the new dataset to fine tune that model with the new data. Two models must be selected to fulfil the improvement of the final concept functions. One of them will be chosen to keep improving the dataset when new events are sent by the concept's social community, this model must fulfil the same requirements as the model used in the labelling process: the runtime will not be important, and the accuracy is the key element. For now, the Faster-RCNN NAS has good results and could be selected again, unless a model with better accuracy it is available.

The second model must be able to run in the Jetson TX2 board with real-time inference. To this end, several light-weight ConvNets models have emerged, explicitly designed for real-time processing in embedded systems, for example, the SSD model presented in section 2.5. Those models have decreased computational complexity to make the system run in a feasible time frame. Although these models are able to perform in real-time requirements, the accuracy of those models is also decreased. The training, comparison, testing, and selection of the model for the Jetson TX2 board implementation, will be addressed in future work.

Lastly, it is important to mention that by improving the training dataset, the general performance of other main functions of the final concept, as the Front Collision Warning, are being improved as well.

Naturally, there are other ways to improve the Front Collision Warning and the other main functions of the final concept, but a different approach is needed, and consequently, the scope of the improvements will be more specific about technical details as run-time. As well, those improvements will be addressed in the future work of the concept.

III. Third Stage: UX between human and AI

User experience is a hot topic nowadays and at the same time, is one of the most important elements that need to be considered when designing a system or a product. Using critical function prototypes and simulators in the first stage of the project, not only contributed to having a better overview of the possible features for the final concept, but these tools also guided to obtain some insights about the behaviour of the driver, allowing to know about their concerns and feelings with specific situations or actions.

It is really interesting to notice that some of the issues detected on the first stage have a big relationship with issues that Artificial Intelligence and self-driving technology are confronting lately. One of those concerning issues is “trust”. The trust issue stands out in the last months in the automotive industry due to the road traffic crashes involving some companies pursuing the self-driving car. Trust, therefore constitutes an essential research area for the introduction of Automated Vehicles to road traffic and the improvements of Advanced Driver-Assistance Systems.

The purpose of this stage was to address the problem regarding the acquisition of drivers’ effective trust metrics. This stage gave as a result, the conceptual design of a study which could evaluate the relationship between the drivers’ automation trust and their physiological state, specifically the driver’s gaze behaviour.

It is important to note that this study is focused about measuring driver’s trust in self-driving technology. However, it could be used to measure the driver’s trust with ADAS technology, as the final concept showed above. The study is on the development stage and, if possible, will be conducted in collaboration with an automotive safety research centre.

7. Study Description

This chapter is confidential due to a Non-Disclosure Agreement. The complete document is kept in the Centre of Mechanical Design and Technological Innovation of the National Autonomous University of México (UNAM). For more details, please contact Dr. Alejandro C. Ramírez Reivich.

8. Future Work

As this is a conceptual design, there are a lot of opportunity areas which can be addressed with a multi-disciplinary team. However, following the work of this thesis, the future work is divided into two parts.

8.1. Future work for Stage I and II

Several tasks have been settled to keep working on, like the creation of the new dataset adding specific road hazards with the labelling process mentioned in section 7. Shortly, it is probable that new labelling tools and methods will be created, allowing to achieve better results with a reduction of the labelling time. Anyway, a significant effort in maintaining this dataset must be made to keep improving the general performance of the final concept.

Another critical task is the deployment of the training and inference systems on the cloud. It is needed to plan, and design how is going to be the procedure to receive new dataset information, how often the models will be trained, and which platform will be used to train the models and manage all the information regarding the final concept functions.

As well, real-time implementation is a critical requirement in the real-world deployment. Different models must be tested and compared before the deployment on the Jetson TX2 board. Therefore, this task opens a new whole horizon to explore and apply Deep Learning architectures with good performance in real-time.

8.2. Future work for Stage III

Stage III unlocked a new opportunity area. With the deployment of Deep Learning models, there is a possibility to improve and thrive the User Experience along with Augmented Reality. Nowadays there is a big expectation about Head-Up Displays (HUDs), and extensive research is in the process to identify the best way to deliver critical information from the driving task, without distracting or overloading the drivers. Using models as the mask-RCNN, which highlight specific information about the environment could help to the aim of delivering critical information.

Figure 101.- Instance segmentation with Mask-RCNN.

For the reason that Highly Automated Vehicles are not commercialized yet, it is important to examine drivers' behaviours and attitudes toward Highly Automated Vehicles. Deep Learning models could not be only used for developing new experiences, with those models it is possible to study the behaviour and physiological state of the drivers. Therefore, a better understanding of the drivers could be achieved. One possible application is to use physiological data along with vehicle data as an input to the neural network. By training a model to predict driver manoeuvres and intentions, the capabilities of ADAS and the driver's performance could be enhanced.

The most ambitious future work for this stage is the design of a system which can calibrate the drivers' trust in real-time. This novel system could lead to a huge and disruptive change in drivers' User Experience.

9. Discussion

9.1. Technology

For several fields, including automotive and mobility industries, Machine Learning will not be optional; it will be a technological foundation making it the source of significant competitive advantage for the next years. AI can be used to help human users make critical decisions, especially with cars that do not have ADAS. While powerful and appropriate hardware exists to run the algorithms in the back-end environment, the embedding of the whole systems requires technological hardware advancements and connection solutions [74].

Designing and implementing a system that can operate in (mostly) unrestricted environments will require huge amounts of effort, due to the virtually infinite number of use cases that must be covered and tested. For instance, the system must be able to guess which areas are appropriate for moving vehicles where the lane markings are inexistent or roads without any traffic sign. Those types of issues can become more difficult for specific areas. For example, a Computer Vision problem appears if the road surface is significantly the same as its surroundings (roads covered in snow), the system would be struggling to define where are the road and the sidewalks. Nowadays, even the biggest automotive manufacturers are struggling with these kinds of situations.

9.2. Regulations and Standards

The term “Automotive-grade”, requires higher safety standards and more accurate than many current Machine Learning application cases. In automotive and mobility, regulatory involvement is usually together with coordination of industry standards to allow scaling and integration of various systems. It is expected that the first “players” in the field might shape the standards during the early development of the technology and its applications, as some did when Antilock Braking System (ABS) was developed [74]. As the final concept is not part of the car’s systems, it might not have to get the “automotive grade,” but it might be needed to fulfil some standards and regulations about data usage and data confidentiality.

9.3. Test and Validation

Although the system is not forced to obtain the "automotive grade," it is needed to assure that the system is robust and efficient. In the automotive industry, techniques as *Brute Force, Software in the Loop (*or *Model in the Loop)* and *Hardware in the Loop* are used for test-and-validation purposes [86]. The *Brute Force* technique, which consists of exposing the system to millions of driving miles to determine statistically that the system is safe and operates as expected. If it is decided to use this technique, the main challenge is the number of miles required. Some research shows that the driving miles required to achieve, with 95% confidence, that the failure rate was 1.09 fatalities per 100 miles (as the equivalent of the 2013 US human-fatality rate) is about 275 million miles. "With a fleet of 100 Automated Vehicles being test-driven 24 hours a day, 365 days a year at an average speed of 25 miles an hour, this would take about 12.5 years to achieve 275 million miles" [99]. To demonstrate better-than-human performance, the number of miles required can quickly reach the billions. Of course, if the decision is taken to drive on highways, it is easier to achieve, but the system will not learn any new hazardous situations from city environments, making even harder to use this technique for the purposes of the final concept. The *Software in the Loop* technique combines real-world tests with simulations. Using this technique could make possible, to greatly reduce the number of testing miles required. Lastly, by using the *Hardware in the Loop* technique, pre-recorded sensor data can be fed into the system simulation, making possible to validate the operation of actual hardware.

It is crucial to choose an effective way to implement test and validation. The goal is to achieve the required confidence levels in the least amount of time for the purposes of the final concept. For this reason, it might be needed to develop or implement a hybrid approach involving the techniques previously mentioned.

9.4. Human Factors

Although the next insights are more related to the self-driving technology, they have a big relationship with ADAS. Nowadays there is a huge misunderstanding about partially automated VS. Fully automated cars. If this issue is not addressed in the right way, it will only grow over time. There are several videos on social media showing motorists putting enormous faith in the partially automated driving technology and the Advanced Driver-Assistance Systems. “There is a sense of complacency when you’re driving the same loops over and over, and you trust the vehicle” [159] says an Uber former self-driving vehicle operator.

The partially automated systems have a good performance in the functions that were designed for, then it is natural that people trust them more than they should, but these systems are a form of Driver-Assistance Systems, this means that the most critical decisions are still determined by the driver.

Unfortunately, controlled experiments as well as real-life situations, show that drivers who place too much trust in automation end up crashing into stationary vehicles or other objects [86]. The last two years, the automotive industry had some serious road traffic crashes involving partially automated driving features. In the first fatal road traffic accident, the vehicle crashed into a semi-truck in Florida while the partially automated driving feature was engaged. The vehicle’s data logs revealed that the driver had his hands on the wheel for 25 seconds of the 37 minutes the partially automated driving feature was activated during the trip, and that the driver had received seven visual and six auditory warnings to return his hands to the wheel. In the second fatal road traffic crash, the vehicle crashed into a highway barrier in California. It was confirmed that the partially automated driving feature was engaged during the collision and that the driver had received multiple warnings to put his hands on the wheel during the drive. In another road traffic crash, the driver had the partially automated driving feature activated at the time of the crash. The driver took her hands off the wheel over 12 times during the drive, including the 80 seconds before the collision, the driver was using her phone [160].

Those drivers seemed to expect more from their cars than they can actually perform. And that is understandable since the systems give the perception of being able to self-drive km after km of uneventful highway driving. The better they work, the more the drivers will trust them, to the point where it is easy to forget that they are not infallible. The current capabilities of ADAS are limited, something many early adopters fail to understand. Right now, the systems are not sophisticated enough to operate without human oversight, all these systems rely on the human paying attention, ready to intervene in case the car encounters something it cannot handle on its own.

The conceptual design of the study shown in section 8, represents a step forward in the development of more objective and continuous measurements of drivers’ trust in self-driving systems. Driver's physiological state could be used as a way to assess real-time changes in driver's trust having an important role in real-world scenarios. For instance, if the drivers over-trust or mistrust the self-driving vehicle, the system could react by modifying the driving behaviour or it may provide to the driver, more information about the road situation or the driving performance. Several studies have shown that designing systems that provide users with accurate enhanced feedback of how they operate could create appropriate trust [168]. Thereby, a safer and efficient driver-vehicle interaction could thrive.

In this way, it could be possible to avoid the fatal accidents caused by over-trust and at the same time, eliminate the mistrust of the users so they can benefit from the self-driving and Advanced Driver-Assistance systems. Trust is a sensitive topic that can affect users from minor discomforts to more safety critical situations in the driving task. Lastly, trust is among the key mindsets and attitudes of successful human-machine collaboration.

9.5. Thesis Outcome and Conclusion

This master research project was the journey into discovering a new horizon of the automotive safety systems. At the beginning of this thesis, the objective was to create (along the Safety Cloud teammates) a conceptual design for a retrofittable automotive safety system. Then, the objective changed when the approach with Artificial Intelligence (AI) surfaced.

Although the AI is nowadays a trending topic and marvellous things have been developed with AI, it is stunning to observe that some AI issues are very similar and very related to the issues that the Safety Cloud team found within the GlobalDrive project. Consequently, the initial objective changed again.

Thus, the scope added the research of human factors in the automotive safety. In summary, it is possible to say that the main objective of the human factors research is to provide an understanding of how drivers perform as a system component in the safety systems. This role recognizes that driver performance is influenced by many environmental, psychological, and system design factors.

Therefore, it is possible to think each stage as one different element. In other words, the GlobalDrive stage (a), the human factors stage (b) and the AI stage (c).

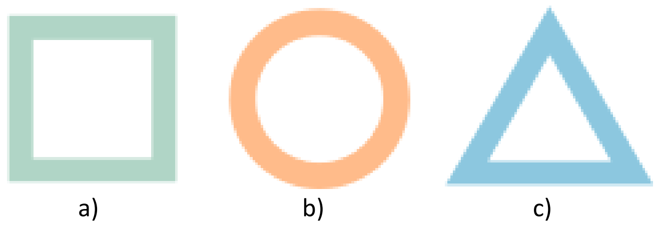


Figure 102.- Each stage figure description.

The thesis outcome, in the opinion of the author, is that even though each stage is from a totally different study field, several relevant sub-fields must be bonded for the success of using existing or new safety technologies and thus, users' benefit.

The designers, researchers and developers will need time to adjust to this shift, to find the relevant sub-fields and to implement the techniques and methodologies that will thrust the whole user experience regarding safety and comfort. This thesis was the journey of embracing that AI, human factors and concept development must be linked for a successful ideation and development of an automotive system.

Figure 103.- Sub-fields fusion.

The human factors research must be at the centre of this shift because it is the foundation of how researchers could determine which aspects of the safety systems should be modified to improve drivers' performance and reduce unsafe behaviours. Getting started early not only helps produce results quickly but also helps speed up an organization's journey toward embracing this new paradigm shift.

At the same time, AI must be embraced so the full potential of AI could be deployed for the benefit of the automotive industry. “An organizational culture open to the collaboration of humans and machines is crucial for getting the most out of AI” [72]. Now is the right time to start thinking about the best way to adjust social and economic structures to be aware of this new AI reality.

“Often people only think of AI boosting growth by substituting humans, but actually huge value is going to come from the new goods, services and innovations AI will enable [107]”. David Autor, Professor of economics MIT.

AI is much more than a new set of tools and technology, it will help reshaping the world as we know, just as it was the internet in its time. To date, the researchers, designers, developers and engineers are deploying systems that were upgraded by using AI, however it is uncertain if they are using human factors research in the design process of those systems. Fortunately, at the same time the AI will assist in the human factors research and new insights will be generated by the deployment of AI technologies.

The automotive industry still needs to overcome some challenges and shortcomings to allow that self-driving and ADAS technology can thrive in the future and generate a new era for transportation. The end-to-end solution for a disruptive change in transportation needs to implement not only better hardware and software, it needs the research and application of knowledge about human factors (human abilities, limitations and other human characteristics) and AI to achieve the goal of having a novel and disruptive driving experience. There is a long way to go, but history has demonstrated that when humans push their own limits, amazing things happen.

References

[1] 5 Disturbing Texting-While-Driving Stats.
Available: http://www.refinery29.com/2015/05/87688/texting-while-driving-stats-att.
[Accessed 17 January 2017].

[2] ACM Special Interest Group on Mobility of Systems, Users, Data and Computing. 2013. CarSafe App: Alerting Drowsy and Distracted Drivers using Dual Cameras on Smartphones. Available:
http://www.cs.dartmouth.edu/~lorenzo/Papers/mobisys13.pdf. [Accessed 7 December 2016].

[3] aCoDriver 5 – Android-Apps auf Google Play.
Available:
https://play.google.com/store/apps/details?id=com.evotegra.aCoDriver&hl=de.
[Accessed 22 January 2017].

[4] ADTSEA. Aggressive Driving Research Update 2009. Available:
http://www.adtsea.org/Resources%20PDF's/AAA%202009%20Aggressive%20Driving%20Research%20Update.pdf. [Accessed 11 August 2017]

[5] Autoliv and Neonode to showcase the future of Steering Wheel technology | Business Wire. Available:
http://www.businesswire.com/news/home/20141217006558/en/Autoliv-Neonode-showcase-future-Steering-Wheel-technology. [Accessed 16 January 2017].

[6] WayRay. Available:
https://wayray.com/industry/. [Accessed 12 January 2017].

[7] Beanland, V., Lenné, M. G., Fuessl, E., Oberlader, M., Joshi, S., Bellet, T., Banet, A., Rößger, L., Leden, L., Spyropoulou, I., Yannis, G., Roebroeck, H., Carvalhais, J., and Underwood, G. 2013. Acceptability of rider assistive systems for powered two-wheelers. Transportation Research Part F: Traffic Psychology and Behaviour 19, 63–76.

[8] Bike Lane Safety Light. Available:
http://thexfire.com/products-page/lighting-system/bike-lane-safety-light-duplicate.
[Accessed 10 January 17].

[9] Caruma – Internet of Things for Automotive. Available: https://caruma.tech/.
[Accessed 16 January 2017].

[10] Carvi Promise: The Open Road Is In Safe Hands. Available: http://getcarvi.com/. [Accessed 16 January 2017].

[11] Cloud-Computing Definition. Available: http://www.gruenderszene.de/lexikon/begriffe/cloud-computing. [Accessed 14 January 2017].

[12] DIN. Allgemeine Leitsätze für das sicherheitsgerechte Gestalten von Produkten, DIN 31000:2011-05; VDE 1000:2011-05.

[13] Facts and Stats: Drowsy Driving – Stay Alert, Arrive Alive. Available: http://drowsydriving.org/about/facts-and-stats/. [Accessed 16 January 2017].

[14] First ever holographic navigator. Available: https://wayray.com/navion. [Accessed 12 January 2017].

[15] GoPro Official Website - Capture + share your world - Homepage. Available: https://gopro.com/. [Accessed 13 January 2017].

[16] Continental Corporation. Head-up Displays. Available: http://continental-head-up-display.com/. [Accessed 10 January 2017].

[17] 2016. IBM - What is cloud computing? Available: https://www.ibm.com/cloud-computing/learn-more/what-is-cloud-computing/. [Accessed 17 January 2017].

[18] 2017. Introducing the Waylens Camera - The World's Best Data Driven Automotive Cam. Available: https://www.waylens.com/. [Accessed 13 January 2017].

[19] L. M. Bergasa, D. Almería, J. Almazán, J. J. Yebes, and R. Arroyo. 2014. DriveSafe: An app for alerting inattentive drivers and scoring driving behaviors. In 2014 IEEE Intelligent Vehicles Symposium Proceedings, 240–245. DOI=10.1109/IVS.2014.6856461.

[20] Lane Keeping Assist Systems. Available: https://www.vda.de/en/topics/safety-and-standards/lkas/lane-keeping-assist-systems. [Accessed 10 January 2017].

[21] Mell, P. M. and Grance, T. 2011. The NIST definition of cloud computing. National Institute of Standards and Technology, Gaithersburg, MD.

[22] Mercedes-Benz Intelligent Drive. Available:

https://www.mercedes-benz.com/de/mercedes-benz/innovation/mercedes-benz-intelligent-drive/ . [Accessed 16 January 2017].

[23] 2017. Nauto. Available: http://www.nauto.com/. [Accessed 16 January 2017].

[24] Regan, M. A., Hallett, C., and Gordon, C. P. 2011. Driver distraction and driver inattention: definition, relationship and taxonomy. Accident; analysis and prevention 43, 5, 1771–1781.

[25] Reinier Coetzer, G. H. 2011. Eye detection for a real-time vehicle driver fatigue monitoring system. IEEE, Piscataway, NJ.

[26] RiderEcall – Schuberths einzigartiges Notrufsystem. Available: http://www.alpentourer.eu/alpentourer-news/schuberth-rider-ecall/. [Accessed 13 January 2017].

[27] 2014. Rider-Ecall vom Markt genommen - Motorrad-Notruf vorerst gescheitert. Available: http://www.motor-talk.de/news/motorrad-notruf-vorerst-gescheitert-t4883930.html. [Accessed 13 January 2017].

[28] Statistisches Bundesamt. Verkehrsunfälle - Fachserie 8 Reihe 7 - 2015 - Aktualisierung vom 12.10.2016.

[29] Statistisches Bundesamt. Verkehrsunfälle - Zeitreihen 2015 (Stand: 12.10.2016)

[30] 2016. Stop Distracted Driving. Available: http://www.jgiani.com/StopDD.html. [Accessed 17 January 2017].

[31] The Mercedes-Benz F 015 Luxury in Motion. Available: https://www.mercedes-benz.com/en/mercedes-benz/innovation/research-vehicle-f-015-luxury-in-motion/. [Accessed 10 January 2017].

[32] THE WORLD'S FIRST CAR WITH A HEARTBEAT. Available: https://www.lexus.com.au/brand/heartbeat?category=ALL. [Accessed 13 January 2017].

[33] Under Control - Motorrad Fahrertraining. Available: http://www.motorrad-unter-kontrolle.de/informationen/eine-kritische-situation/. [Accessed 22 January 2017].

[34] Waze. 2017. Kostenlose community-basierende Mapping-, Verkehrs- & Navigations-App. Available: https://www.waze.com/de/. [Accessed 16 January 2017].

[35] What is lane departure warning, and how does it work? Available: https://www.extremetech.com/extreme/165320-what-is-lane-departure-warning-and-how-does-it-work. [Accessed 23 January 2017].

[36] Cui, Y., Sun, L., Kang, J., Jiang, Y. Interactive Safety Analysis Framework of Autonomous Intelligent Vehicles. MATEC Web of conferences 44, 2016. DOI: https://doi.org/10.1051/matecconf/20164401029

[37] World Health Organisation. GLOBAL STATUS REPORT ON ROAD SAFETY 2015.

[38] Yerkes Dodson Law. Available: http://l2bet.com/theory-of-motivation-and-its-application-in-betting/. [Accessed 23 January 2017].

[39] Ziegler, M. 2014. ESP ab 1. November 2014 Pflicht. Available: http://www.automobil-industrie.vogel.de/esp-ab-1-november-2014-pflicht-a-465145/. [Accessed 31 December 2016].

[40] Accident free vector icons designed by Freepik. Available: http://www.flaticon.com/free-icon/accident_130154. [Accessed 25 January 2017].

[41] ADVICY Drive – Security system for drivers | We prevent people from falling asleep at the wheel. Available: http://www.advicydrive.com/. [Accessed 25 January 2017].

[42] Anti Sleep Alarm StopSleep keeps you awake at the wheel. Available: http://stopsleep.biz/. [Accessed 25 January 2017].

[43] Autonomes Fahren: Chancen der Digitalisierung und des Wandels im Mobilitätsmarkt erkennen und für die Flexibilisierung des Öffentlichen Nahverkehrs nutzen.

[44] Behance. Steering wheel rendering tutorial. Available: https://www.behance.net/gallery/34859813/Steering-wheel-rendering-tutorial. [Accessed 25 January 2017].

[45] Boy with Sleeping Face free vector icons designed by Freepik. Available: http://www.flaticon.com/free-icon/boy-with-sleeping-face_75562. [Accessed 25 January 2017].

[46] Building crane free vector icons designed by Freepik. Available:

http://www.flaticon.com/free-icon/building-crane_72791. [Accessed 25 January 2017]

[47] Camera free vector icons designed by Freepik. Available: http://www.flaticon.com/free-icon/camera_112545. [Accessed 25 January 2017].

[48] Car collision free vector icons designed by Freepik. Available: http://www.flaticon.com/free-icon/car-collision_65788. [Accessed 25 January 2017]

[49] Cruise control and adaptive cruise control: the complete guide. Available: http://www.autoexpress.co.uk/car-tech/98225/cruise-control-and-adaptive-cruise-control-the-complete-guide. [Accessed 25 January 2017].

[50] Artis JHD-97 Universal Car Rear View Mirror Smartphone Mount. Ebay. Available:https://m.ebay.in/itm/Artis-JHD-97-Universal-Car-Rear-View-Mirror-Smartphone-Mount-/272644506828. [Accessed 25 January 2017].

[51] Free mobile & web prototyping (iOS, iPhone, Android) for designers – Marvel. Available: https://marvelapp.com/. [Accessed 25 January 2017].

[52] Kraftfahrt-Bundesamt - Bestand. Available: http://www.kba.de/DE/Statistik/Fahrzeuge/Bestand/bestand_node.html. [Accessed 25 January 2017].

[53] Road with broken line free vector icons designed by Freepik. Available: http://www.flaticon.com/free-icon/road-with-broken-line_62512. [Accessed 25 January 2017].

[54] Speed limit free vector icons designed by Retinaicons. Available: http://www.flaticon.com/free-icon/speed-limit_204154. [Accessed 25 January 2017].

[55] Phonearena, Tomtom 2015. tomtom-app-google-android-2015-02-450x258.jpg (JPEG-Grafik, 450 × 258 Pixel). Available:https://www.phonearena.com/news/TomTom-offers-new-premium-navigation-and-mapping-app-for-Android_id67590. [Accessed 25 January 2017].

[56] Warning free vector icons designed by Gregor Cresnar. Available: http://www.flaticon.com/free-icon/warning_159060. [Accessed 25 January 2017].

[57] Sostenedor de Taza Del Coche Palo de Teléfono móvil Soporte Universal Smartphone Telefon Accesorio para Montaje Nota Película de Succión de Parabrisas para GPS. AliExpress. Available:

https://es.aliexpress.com/item/Skymall-Car-cell-phone-Suction-Holder-Windshield-Dashboard-universal-suction-cradle-bracket-cradle-daily-watch-video/32628580084.html. [Accessed 25 January 2017].

[58] Blitzer Deutschland Österreich – Android-Apps auf Google Play. Available: https://play.google.com/store/apps/details?id=com.sygic.speedcamapp&hl=de. [Accessed 25 January 2017].

[59] Deutschland, A. Spielerisch besser Autofahren – mit der App Drive von AXA! | AXA. Available: https://www.axa.de/das-plus-von-axa/auto-kfz-unterwegs/apps/Drive-App. [Accessed 25 January 2017].

[60] Flo. Flo - Driving Insights - apps + gps tracking for trip recording. Available:https://www.driveflo.com/. [Accessed 25 January 2017].

[61] iOnRoad. Available: http://www.ionroad.com/. [Accessed 25 January 2017].

[62] Wuidi. Wuidi. Available: https://www.wuidi.de/. [Accessed 25 January 2017].

[64] Ki-Yeong Park and Sun-Young Hwang. 2014. Robust range estimation with a monocular camera for vision-based forward collision warning system. The Scientific World Journal 2014.

[65] Koch, C. and Brilakis, I. 2011. Pothole detection in asphalt pavement images. Advanced Engineering Informatics 25, 3, 507–515.

[66] Kumar, S. 2016. Lane Detection based on Contrast Analysis (Master thesis).

[67] Microsoft. 2017. Pipelines. Available: https://msdn.microsoft.com/en-us/library/ff963548.aspx. [Accessed 25 January 2017].

[68] Slimane, A. B., Khoudeir, M., Brochard, J., and Do, M.-T. 2008. Characterization of road microtexture by means of image analysis. Wear 264, 5-6, 464–468.

[69] Nvidia Jetson. Available:https://www.nvidia.com/en-us/autonomous-machines/embedded-systems-dev-kits-modules/. [Accessed 20 April 2017].

[70] Blackfly 2.3 MP Mono USB3 Vision (Sony Pregius IMX249). Available:https://www.ptgrey.com/blackfly-23-mp-mono-usb3-vision-sony-pregius-imx249. [Accessed 20 April 2017].

[71] Copeland, M. (2017, December 15) What's the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning? Nvidia Blog. Available:

https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/. [Accessed 28 December 2017].

[72] Smartening up with Artificial Intelligence (AI) – What's in it for Germany and its Industrial Sector? Mckinsey & Co. Digital / McKinsey. Available:https://www.mckinsey.de/files/170419_mckinsey_ki_final_m.pdf. [Accessed 28 December 2017].

[73] Roland Berger, Automation in financial services. Available:https://www.adetem.org/wp-content/uploads/2017/05/RolandBerger.pdf. [Accessed 28 December 2017].

[74] McKinsey Center for Future Mobility. Smart moves required. The road towards Artificial intelligence in Mobility. McKinsey & Company.

[75] Labs' Deep Learning Cookbook headlines the launch of HPE's AI platforms and services. Available: https://community.hpe.com/t5/Behind-the-scenes-Labs/Labs-Deep-Learning-Cookbook-headlines-the-launch-of-HPE-s-AI/ba-p/6981300#.Wp2veGrOWUk. [Accessed 28 December 2017].

[76] What Is Deep Learning? | How It Works, Techniques & Applications. Mathworks. Available: https://www.mathworks.com/discovery/deep-learning.html#howitworks. [Accessed 28 December 2017].

[77] Blog, G., Choudhary, A., Shaikh, F., N., & Gupta, D. (2016, October 06). The Evolution and Core Concepts of Deep Learning & Neural Networks. Available: https://www.analyticsvidhya.com/blog/2016/08/evolution-core-concepts-deep-learning-neural-networks. [Accessed 28 December 2017].

[78] A. Luckow, M. Cook, N. Ashcraft, E. Weill, E. Djerekarov and B. Vorster, "Deep learning in the automotive industry: Applications and tools," 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, 2016.

[79] Goodfellow I., Bengio Y., Courville A. Deep Learning. 2016. Available: http://www.deeplearningbook.org/contents/convnets.html. [Accessed 30 December 2017].

[80] Stanford CS231n Class notes: Convolutional Networks for Visual Recognition. Convolutional Networks (CNNs/Convnets). Available: http://cs231n.github.io/convolutional-networks/. [Accessed 30 December 2017].

[81] What is visual Recognition? Clarifai. Available:https://www.clarifai.com/technology. [Accessed 30 December 2017].

[82] Dertat, A. (2017, November 08). Applied Deep Learning - Part 4: Convolutional Neural Networks. Available: https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2. [Accessed 30 December 2017].

[83] Stanford CS231n Class notes: Convolutional Networks for Visual Recognition. Transfer Learning. Available: http://cs231n.github.io/transfer-learning/. [Accessed 30 December 2017].

[84] McKinsey Center for Future Mobility (MCFM). McKinsey & Company. Available: https://www.mckinsey.com/features/mckinsey-center-for-future-mobility/overview. [Accessed 4 January 2018].

[86] Heineke K., Kampshoff P., Mkrtchyan A., Shao E. Self-driving car technology: When will the robots hit the road? McKinsey & Company. Available: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/self-driving-car-technology-when-will-the-robots-hit-the-road. [Accessed 6 January 2018].

[87] Bouée, Charles-Edouard. The new era of artificial Intelligence. Roland Berger Available: https://www.rolandberger.com/en/Blog/Growth-and-Inclusive-Prosperity.html. [Accessed 6 January 2018].

[89] B. Marr. (Jul 4, 2018). What Is The Difference Between Deep Learning, Machine Learning and AI? Available: https://www.forbes.com/sites/bernardmarr/2016/12/08/what-is-the-difference-between-deep-learning-machine-learning-and-ai/#15e3437e26cf. [Accessed 7 July 2018].

[90] T. Instruments. (Jul 4,2018). *AI in Automotive: Practical deep learning*. Available: https://e2e.ti.com/blogs_/b/behind_the_wheel/archive/2018/02/08/ai-in-automotive-practical-deep-learning. [Accessed 7 July 2018].

[91] A. Krizhevsky, Nair, V., Hinton G. (Jul 4, 2018). *The CIFAR-10 and CIFAR-100 datasets*. Available: https://www.cs.toronto.edu/~kriz/cifar.html. [Accessed 7 July 2018].

[92] IMAGENET. Large Scale Visual Recognition Challenge (ILSVRC). Available: http://www.image-net.org/challenges/LSVRC/. [Accessed 7 July 2018].

[93] Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. Available: https://arxiv.org/abs/1602.07261?context=cs. [Accessed 7 July 2018].

[94] ImageNet. Available: http://www.image-net.org/. [Accessed 7 July 2018].

[95] Amazon Web Services. Available: https://aws.amazon.com/. [Accessed 7 July 2018].

[96] Paperspace. Available: https://www.paperspace.com/. [Accessed 7 July 2018].

[97] Microsoft Azure. Available: https://azure.microsoft.com/. [Accessed 7 July 2018].

[98] Google CLOUD AI. Available: https://cloud.google.com/products/machine-learning/. [Accessed 7 July 2018].

[99] Kalra, N., Paddock, S. How many miles of driving would it take to demonstrate Autonomous Vehicle Reliability? RAND Corporation. RR-1478-RC, 2016. Available: https://www.rand.org/pubs/research_reports/RR1478.html. [Accessed 21 July 2018].

[100] Taxonomy and Definitions of Terms Related to Driving Automation Systems for On-Road Motor Vehicles. J3016 SEP 2016. SAE International. Available: https://saemobilus.sae.org/content/j3016_201609. [Accessed 21 July 2018].

[101] Argo AI. Available: https://www.argo.ai/. [Accessed 21 July 2018].

[102] nuTonomy. Available: https://www.nutonomy.com/. [Accessed 21 July 2018].

[103] drive.ai. Available: https://www.drive.ai/. [Accessed 21 July 2018].

[104] Zoox. Available: https://zoox.com/. [Accessed 21 July 2018].

[105] Preferred Networks. Available: https://www.preferred-networks.jp/en/. [Accessed 21 July 2018].

[106] Ottomotors Co. Available: https://ottomotors.com/. [Accessed 21 July 2018].

[107] Purdy, M., Daugherty, P. Why artificial Intelligence is the future of growth. Accenture 2016. Available: https://www.accenture.com/t20170927T080049Z__w__/us-en/_acnmedia/PDF-33/Accenture-Why-AI-is-the-Future-of-Growth.PDFla=en. [Accessed 21 July 2018].

[108] Definition of Artificial Intelligence. Available: https://en.oxforddictionaries.com/definition/artificial_intelligence. [Accessed 21 July 2018].

[109] Definition of Safety. Available: https://en.oxforddictionaries.com/definition/safety. [Accessed 21 July 2018].

[110] Adaptive Cruise Control. Available: http://www.volkswagen.co.uk/technology/adaptive-cruise-control-acc. [Accessed 23 January 2017].

[111] What is adaptive cruise control, and how does it work? Available: https://www.extremetech.com/extreme/157172-what-is-adaptive-cruise-control-and-how-does-it-work. [Accessed 23 January 2017].

[112] Autoliv Inc. Available: https://www.autoliv.com/. [Accessed 21 July 2018].

[113] Embedded Vision Alliance. The evolution of Deep Learning for ADAS Applications. Available: https://www.embedded-vision.com/industry-analysis/technical-articles/evolution-deep-learning-adas-applications. [Accessed 21 July 2018].

[114] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[115] Scanlan, D., Solana, L., Deep Learning for Robust Road Object Detection. Using Convolutional Neural Networks to Detect Cars, Pedestrians, & Ciclysts in Single Colour Images. Master’s thesis. 2017.

[116] J. Uijlings, K. Van De Sande, and T. Gevers, “Selective search for object recognition,” *International journal of*, 2013.

[117] PASCAL VOC dataset. Available: http://host.robots.ox.ac.uk/pascal/VOC/. [Accessed 21 July 2018].

[118] COCO dataset. Available: http://cocodataset.org/#home. [Accessed 21 July 2018].

[119] T.-Y. Lin *et al.*, "Microsoft COCO: Common Objects in Context," *arXiv [cs.CV]*, 01-May-2014.

[120] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *arXiv [cs.CV]*, 11-Nov-2013.

[121] Deep Learning for Object Detection: A Comprehensive Review. Available: https://towardsdatascience.com/deep-learning-for-object-detection-a-comprehensive-review-73930816d8d9. [Accessed 21 July 2018].

[122] R. Girshick, "Fast R-CNN," *arXiv [cs.CV]*, 30-Apr-2015.

[123] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *arXiv [cs.CV]*, 04-Jun-2015.

[124] W. Liu *et al.*, "SSD: Single Shot MultiBox Detector," *arXiv [cs.CV]*, 08-Dec-2015.

[125] Deep Learning for Object Detection: A Comprehensive Review. Available: https://medium.com/comet-app/review-of-deep-learning-algorithms-for-object-detection-c1f3d437b852. [Accessed 21 July 2018].

[126] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *arXiv [cs.CV]*, 20-Mar-2017.

[127] NeuroNuggets: Segmentation with Mask R-CNN. Acces: https://medium.com/neuromation-io-blog/neuronuggets-segmentation-with-mask-r-cnn-c76d363b67fb.

[128] Brown, C. M., & Noy, Y. I. (2004). Behavioural adaptation to in-vehicle safety measures: Past ideas and future directions. In T. Rothengatter & R. D. Huguenin (Eds.), Traffic and transport psychology: Theory and application (Vol. 25–46, pp. 25–46). Oxford, UK: Elsevier.

[129] Wickens, C. D., & Xu, X. (2002). Automation trust, reliability and attention (Tech. Rep. No. AHFD-02-14/MAAD-02-2) (pp. 0–66). Savoy, IL: University of Illinois, Aviation Research Lab.

[130] Carlson, M. S., Desai, M., Drury, J. L., Kwak, H., & Yanco, H. A. (2014, March). Identifying factors that influence trust in automated cars and medical diagnosis systems. In AAAI Symposium on The Intersection of Robust Intelligence and Trust in Autonomous Systems (pp. 20–27). Palo Alto, CA: Association for the Advancement of Artificial Intelligence.

[131] Sheridan, T. B., & Parasuraman, R. (2005). Human-automation interaction. Reviews of Human Factors and Ergonomics, 1(1), 89–129. doi:10.1518/155723405783703082

[132] Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep Your Scanners Peeled Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving. Human Factors: The Journal of the Human Factors and Ergonomics Society, 58(3), 509-519.

[133] Jian, J., Bisantz, A., & Drury, C. (1998). Foundations for an empirically determined scale of trust in automated systems. United States Air Force Research Laboratory.

[134] Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in automation–Before and after the experience of take-over scenarios in a highly automated vehicle. Procedia Manufacturing, 3, 3025- 3032.

[135] Madhani, K., Khasawneh, M. T., Kaewkuekool, S., Gramopadhye, A. K., & Melloy, B. J. (2002). Measurement of human trust in a hybrid inspection for varying error patterns. In Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting (pp. 418–422). Santa Monica, CA: Human Factors

[136] Master, R., Jiang, X., Khasawneh, M. T., Bowling, S. R., Grimes, L., Gramopadhye, A. K., & Melloy, B. J. (2005). Measurement of trust over time in hybrid inspection systems. Human Factors and Ergonomics in Manufacturing & Service Industries, 15(2), 177–196.

[137] Brown, R., & Galster, S. (2004). Effects of reliable and unreliable automation on subjective measures of mental workload, situation awareness, trust and confidence in a dynamic flight task. In Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting (pp. 147–151). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/154193120404800132.

[138] Muir, B., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics, 39(3), 429–460. doi:10.1080/00140139608964474

[139] Korber, M., Baseler, E., Bengler, K. Introduction matters: Manipulating trust in automation and reliance in automated driving.

[140] Wickens, C., Gordon, S., Liu, Y. An introduction to Human factors engineering. 1998. Longman. Pp 14.

[141] Carsten, O., CH Lai, F., Barnard, Y., Jamson, A., Merat, N. 2012. Control task substitution in semiautomated driving: Does it matter what aspects are automated? Human factors 54 ,5(2012) ,747 – 761.DOI: http://dx.doi.org/10.1177/0018720812460246.

[142] Kemeny, A. and Panerai, F. 2003. Evaluating perception in driving next term simulation experiments. Trends in Cognitive Sciences 7, 1 (2003), 31–37. DOI: http://dx.doi.org/10.1016/ S1364-6613(02)00011-6.

[143] Korber, M. Theoretical considerations and development of a questionnaire to measure trust in automation. Proceedings 20th Triennial Congress of the IEA (2018), Springer.

[144] Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

[145] West, J., Yurdana, M., Jordan, A. A matter of Trust: How smart design can accelerate automated vehicle adoption. White paper, Intel. Available: https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/trust-autonomous-white-paper-secure.pdf. [Accessed 3 August 2018].

[146] Marinik, A., Bishop, R., Fitchett, V., Morgan, J. F., Trimble, T. E., & Blanco, M. (2014, July). Human factors evaluation of level 2 and level 3 automated driving concepts: Concepts of operation. (Report No. DOT HS 812 044). Washington, DC: National Highway Traffic Safety Administration.

[147] Tesla Autopilot Crash: Why We Should Worry About a Single Death. Available: https://spectrum.ieee.org/cars-that-think/transportation/self-driving/tesla-autopilot-crash-why-we-should-worry-about-a-single-death. [Accessed 3 August 2018].

[148] Advanced Driving Simulator. Available: https://research.qut.edu.au/carrsq/services/advanced-driving-simulator/. [Accessed 3 August 2018].

[149] Tensorflow Object Detection API. Available: https://github.com/tensorflow/models/tree/master/research/object_detection. [Accessed 3 August 2018].

[150] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Rob. Res.*, vol. 32, no. 11, pp. 1231–1237, Aug. 2013.

[151] KITTI Vision Benchmark Suite. Available: http://www.cvlibs.net/datasets/kitti/. [Accessed 3 August 2018].

[152] J. Huang *et al.*, "Speed/accuracy trade-offs for modern convolutional object detectors," *arXiv [cs.CV]*, 30-Nov-2016.

[153] Open Images Dataset V4. Available: https://storage.googleapis.com/openimages/web/factsfigures.html. [Accessed 3 August 2018].

[154] Tensorflow detection model zoo. Available: https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_model_zoo.md. [Accessed 3 August 2018].

[155] LabelIMG. Available: https://github.com/tzutalin/labelImg. [Accessed 3 August 2018].

[156] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," *arXiv [cs.CV]*, 21-Jul-2017.

[157] Discover the difference between deep learning training and inference. Available: http://blog.exxactcorp.com/discover-difference-deep-learning-training-inference/. [Accessed 3 August 2018].

[158] Firas Lethaus. Using pattern recognition to predict driver intent. International Conference on Adaptive and Natural Computing Algorithms. ICANNGA 2011.

[159] The unavoidable folly of making humans train self-driving cars. Available: https://www.wired.com/story/uber-crash-arizona-human-train-self-driving-cars/. [Accessed 3 August 2018].

[160] Tesla's Autopilot used to be a huge advantage for the company -- but now it's becoming a problem. Available: https://www.businessinsider.com.au/teslas-autopilot-faces-scrutiny-after-accidents-2018-5?r=US&IR=T. [Accessed 3 August 2018].

[161] Ford Motor Company. Active Park Assist. Available: https://www.ford.com.au/technology/active-park-assist/. [Accessed 3 August 2018].

[162] European Commission. Brake Assist. Available: https://ec.europa.eu/transport/road_safety/specialist/knowledge/esave/esafety_measures_unknown_safety_effects/brake_assist_en. [Accessed 3 August 2018].

[163] European Commission. Collision avoidance systems. Available: https://ec.europa.eu/transport/road_safety/specialist/knowledge/esave/esafety_measures_unknown_safety_effects/collision_avoidance_systems_en. [Accessed 3 August 2018].

[164] Daimler Chrysler AG. Real World Safety Benefits of Brake Assistance Systems. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.691.7824&rep=rep1&type=pdf. [Accessed 3 August 2018].

[165] Bosch AG. Lane change assist. https://www.bosch-mobility-solutions.com/en/products-and-services/passenger-cars-and-light-commercial-vehicles/driver-assistance-systems/lane-change-assist/. [Accessed 3 August 2018].

[166] AUTOMATIC Pro. Available: https://automatic.com/pro/. [Accessed 3 August 2018].

[167] dguard – your motorcycle emergency call system. Available: https://www.dguard.com/en. [Accessed 3 August 2018].

[168] Hoff, K. A., Bashir, M. Trust in automation: integrating empirical evidence on factors that influence trust. 2015

[169] Koo, J., Kwae J. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust and performance.

[170] Walker, F., Verwey, W., Martens, M. Gaze Behaviour as a Measure of Trust in Automated Vehicles.

[171] eHorizon, Continental. Available: https://www.continental-automotive.com/en-gl/Passenger-Cars/Interior/Software-Solutions-and-Services/eHorizon. [Accessed 16 August 2018].

[172] BMW Motorrad Safety 360. Available: http://www.bmw-motorrad.com/com/en/innovation-technology/safety_360/main.html. [Accessed 16 August 2018].

[173] Honda Introduces "Cooperative Mobility Ecosystem" at CES 2017. Available: http://www.multivu.com/players/English/7988331-honda-ces-cooperative-mobility-ecosystem/. [Accessed 16 August 2018].

[174] Connected Motorcycle Consortium. Available: http://www.cmc-info.net/. [Accessed 16 August 2018].

[175] Holley, P. Elon Musk: Human-driven cars may be outlawed because they're 'too dangerous'. The Washington Post. Available: https://www.washingtonpost.com/news/the-switch/wp/2015/03/18/elon-musk-human-driven-cars-may-be-outlawed-because-theyre-too-dangerous/?utm_term=.c4fa227ba214. [Accessed 16 August 2018].

[176] Faulkner, W. 1998. Extraordinary journeys around ordinary technologies in ordinary lives. Soc. Stud. Sci. 28 (3), 484-489.

[177] Biassoni, F. Ruscio, D., Ciceri, R. Limitations and automation. The role of information about device-specific features in ADAS acceptability. Safety Science, Vol. 85. https://doi.org/10.1016/j.ssci.2016.01.017.

[178] Ekman, F., Johansson M., Sochor J. 2018. Creating Appropiate Trust in Automated Vehicle Systems: A Framework for HMI Design. IEEE Transaction on

Human-Machine Systems. Vol. 48-1. Pp:95-101
https://doi.org/10.1109/THMS.2017.2776209.

[179] Volvo Cars. Drive Me Project. Available:
https://www.volvocars.com/intl/buy/explore/intellisafe/autonomous-driving/drive-me.
[Accessed 16 August 2018].

[180] Nvidia Co. NVIDIA DRIVE Developer Program. Available:
https://developer.nvidia.com/driveworks. [Accessed 16 August 2018].

[181] Waymo. Waymo 360° Experience. Available:
https://waymo.com/360experience/. [Accessed 16 August 2018].

[182] Geitner, C., et. al, A link between trust in technology and glance allocation in on-road driving.

[183] KD Nuggets, Castrounis A. Artificial Intelligence, Deep Learning and Neural Networks Explained. Available: https://www.kdnuggets.com/2016/10/artificial-intelligence-deep-learning-neural-networks-explained.html [Accessed 16 August 2018].

[184] Amazon. Adafruit 24 RGB LED Neopixel ring. Available:
https://www.amazon.com/Adafruit-RGB-LED-Neopixel-Ring/dp/B00K9M3WXG
[Accesed 16 August 2018]

Appendix

Appendix A – Benchmarking Description.

External Devices

Besides ADAS that are directly equipped in new cars, there are different retrofit-systems available that can easily be used in older vehicles regardless of manufacturer or type. These systems are using different ways to gather information about the driver, the driving state, the road or the car E.g. the OBD port, cameras or various sensors.

CarVi (Price $300+)

CarVi [10] is a stick-on unit that adds front collision warnings, lane change assistance and driver skills assessment to the driver's current ride. It incorporates a 720p camera, a three-axis accelerometer, a microphone, a speaker, and a Wi-Fi unit to allow wireless connection to the smartphone. It costs around $300 and is under development until now. The camera captures video from what's happening in front of the driver, analysing the data in real time to provide feedback to the driver and warn against dangerous situations. Amongst that, it builds up a report on the user driving skills, taking note of overly incidents like hard braking and dodgy "jackrabbit" starts, to give a score for the trip. The idea behind that is to improve the driving score by following certain driving tips.

AUTOMATIC (Price: $80+)

Another device in the market is AUTOMATIC [166]. It connects the car via the OBD port to the smartphone to provide several driving data. There are two main versions, the Pro and Lite versions. In 2013, the first AUTOMATIC device was released to the market. The most complete version is the AUTOMATIC Pro, which includes trained responders that are ready 24/7 and will contact the driver in case of an emergency. The device allows to connect the car to several apps that let the driver for example log trips to a google sheet, turn on lights and includes also engine diagnostics. In the technical aspect, the AUTOMATIC use the vehicle OBD to provide car-data, connection via Bluetooth or 3G. Depending on the version, an app or not us used for the synchronizing the driving data.

Caruma (No price yet)

Caruma [9] is an innovative connected car camera with dual high-definition cameras, a variety of sensors, and advanced wireless technology that seamlessly integrates with the driver's smartphone. It improves driver safety by monitoring subtle details about the driver to detect fatigue, attentiveness, and driver distraction.

Advicy Drive (No price yet)

This wearable device, measures the driver's heartbeat to prevent him from falling asleep. To check if driver is driving safely or not, Advicy Drive [41] detects its personal attention value (AV). At driving the heartbeat changes continuously, and even the attention value changes with it, this can be easily checked by looking the cursor movements in the app.

If the cursor is in the green zone the driver is driving safely, but if the cursor moves to the red zone, the driver's physical conditions are lower than the safety threshold and an alarm will warn the driver to prevent a possible falling asleep at the wheel.

Navion (No price yet)

Using True Holography, Navion [14] displays virtual indicators exactly where the driver needs. Using True Holography, Navion displays virtual indicators on the road ahead, exactly where the driver needs them to be. It is the first true Augmented Reality car navigation system that applies aeronautical principles to land navigation. The system responds to simple voice commands, letting the driver interact with the car without having to peer about the dashboard. Also, it works with gesture commands for determined features.

Navion allows controlling navigation using the smartphone, the driver will be able to set routes and check map for points of interest. The driver can also see all the data from previous trips. The system fits for almost any car, no headgear or eyewear required.

Xfire bike lane Safety light [8] ($40)

This device is designed to project a bike path onto roadways only. The light uses two 5mW red lasers to create bike lane markings and five red LEDs for increased visibility overall. "To enjoy the night ride Safer".

Dguard ($569 +)

Dguard [167] is a retrofit eCall system for motorcycles. It is using acceleration sensors to detect accidents and automatically send emergency call. Also, it offers an antitheft system with notification on the smartphone.

Existing Smartphone Applications

Currently there are around 2.6 million apps available on the Google Play Store and their number is growing day by day. This shows that it is impossible to give a complete overview of safety related applications available for costumers. Nevertheless, some apps that most influenced the development process of the concepts are presented in the following.

Some of them have only been developed as benchmark in university research and are not available in the app store. The three overall services provided by existing applications on the market are road monitoring, driver surveillance and providing position-based information about the environment. Due to the resource-constrained environment on the smartphone, the apps are mostly only able to provide a subset of possible safety functionalities.

ADAS functionality

CarSafe [2] is an app developed at the Dartmouth College in 2013. Both the road and the driver should be monitored continuously by rear-end and front camera respectively. Machine Learning algorithms analyse the images delivered from the inner car about whether the driver is tired or distracted. It focuses on driver features such as his head pose, eye gaze direction and blinking rate, which allow inferring on his drowsiness, periods of micro sleep and distraction. Meanwhile, the back camera tracks environmental conditions.

Combining with other smartphone sensors such as the GPS, accelerometer, and gyroscope, the app is able to alert the driver about potentially dangerous situations, such as short distance to the car in front, lane weaving/drifting or risky lane change conditions. If the app detects a risky situation in or out the car, it will alert the driver by displaying a related icon on the touchscreen accompanied by an audible alert. Similar apps available on the app market are aCoDriver [3] or IOnRoad [61]. They serve as an additional smartphone-camera-based co-pilot for the driver: the app checks the distance to the car in front to prevent from rear-end collisions; also, it warns the driver on unintentionally leaving his lane. Furthermore, it assists by tracking road signs and speed limits.

Educating the Driver

The *DriveSafe* [19] application, developed at the University of Alcalá, processing the phone's accelerometers, GPS, and rear-camera to produce a complete behaviour profile of the driver that can be checked either on real-time during the trip or be reviewed in detail when having reached the target. The app monitors and scores the driving performance as well as alerts on dangerous behaviour. Evaluating the rear-camera, microphone, inertial sensors, and GPS, seven different manoeuvres are monitored: accelerations, braking, turnings, lane-weaving, lane-drifting, speeding and car-following. Each trip will then be classified as normal, drowsy, or aggressive.

The application is designed to educate the driver by giving feedback about his driving patterns. By doing this in a game-like character that allows for comparing to others, this should be made more appealing to drivers. Apps following a similar intention are for example *Flo-Driving Insights* [60] or the *AXA Drive app* [59].

Position- based Information and Navigation Service

The app *Waze* [34] provides real-time road information to its users based on reports of its community. As millions of users report incidents as accidents, construction sides, road closure or dangerous spots together with their position, future drivers approaching the same spot may be warned beforehand. Furthermore, it provides useful information about the driver's positions' surrounding, e.g. cheap fuel costs, good restaurants. As with Google Traffic, Waze is able to provide users with the currently best route to their target based on real-time traffic information of users using their applications.

While *Waze* reports a bunch of different information categories to the driver, there also exist lots applications targeting one specific area: the app *Wuidi* [62] warns drivers about dangerous areas for animals crossing on the road for example and tells them step by step what to do in case of an accident with an animal. Warnings are displayed depending on the position, the time of day, the season and information from hunters. Another example is the application *Blitzer Deutschland Österreich* [58] that warns drivers from both static and moving radar controls, as well as warns the driver on exceeding his speed limit.