Valeo

DDS Technical Report

2 Month Progress Update

Supervisor Ibrahim Sobh, PhD.
Dec-23-2019

TABLE OF CONTENTS

TABLE OF FIGURES

1 Introduction

This document will provide the problem of Driver Distraction in the road, what are the current solutions and systems to solve the problem? and how our project/solution will solve it (using Deep Learning Computer Vision and NLP Voice Commander) for greater impact.

The task of driving requires continuous attention to road and vehicle control. Drivers may pay insufficient attention to driving because: they are occupied with other activities such as making a phone call or texting, tuning the radio or the air conditioning, talking with a passenger. etc.... and this requires a Full System that will prevent the driver from losing attention and alert him/her if they are distracted.

Figure 1-1 Distracted Driver

1.1 What is Driver Distraction?

Driver Distraction is a form of driver inattention that involves the diversion of attention away from safety critical activities within the driving task[1].

2 Problem Statement

According to the World Health Organization, the number of road traffic deaths continues to increase, reaching 1.35 million in 2016. A research[2] conducted around traffic safety indicates that about 25% of car crashes have been caused by driver distraction.

Every year around 12,000 Egyptians lose their lives as a result of road traffic crashes, 48% of those killed in motor vehicle crashes are occupants (passengers and drivers). The Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS) reported that more than 20% of road accidents occur due to driver inattention/distraction (Visual – Manual – Cognitive).

Figure 2-1 Types Of Driver Distraction

[1] (Regan et al., 2011; Young et al., 2008)
[2] Observed by Lipovac et al.,

2.1 Statistics – Driver Distraction in Numbers

For many, driving is a daily activity, not requiring much thought or consideration. However, the sad reality is that there are 3,287 deaths each day due to fatal car crashes. On average, 9 of these daily fatalities are related to distracted driving[3].

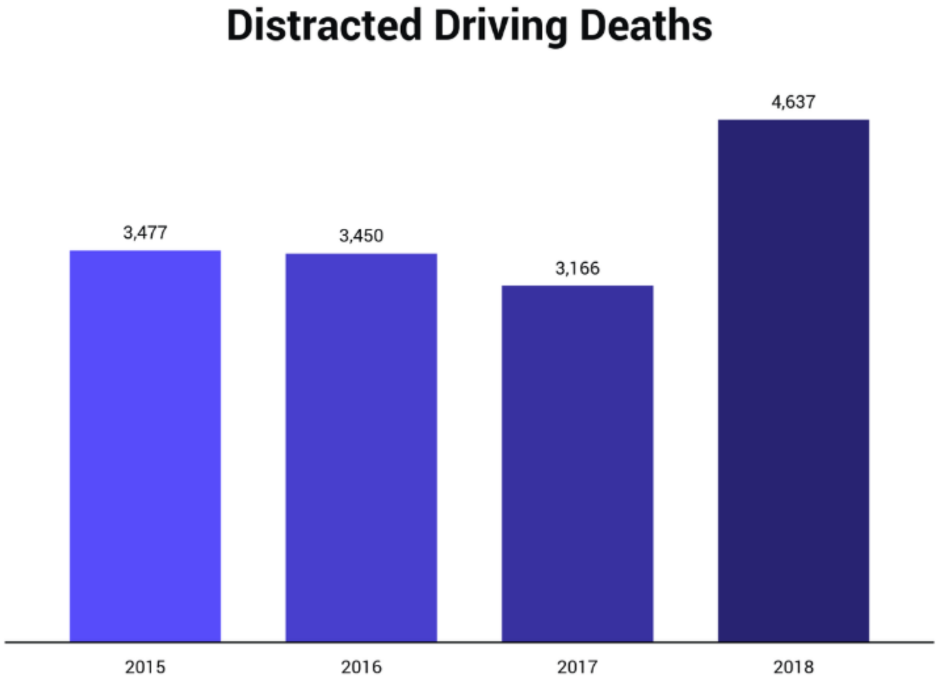


Figure 2-2 Distracted Driving Deaths in years

Teens have been the largest age group that reported being distracted while driving. Driver distraction is reported to be responsible for more than 58% of teen crashes. And that is due to the using of Mobile Phones while driving (Texting – Phone calls – Listening to Music. Etc..)[4].

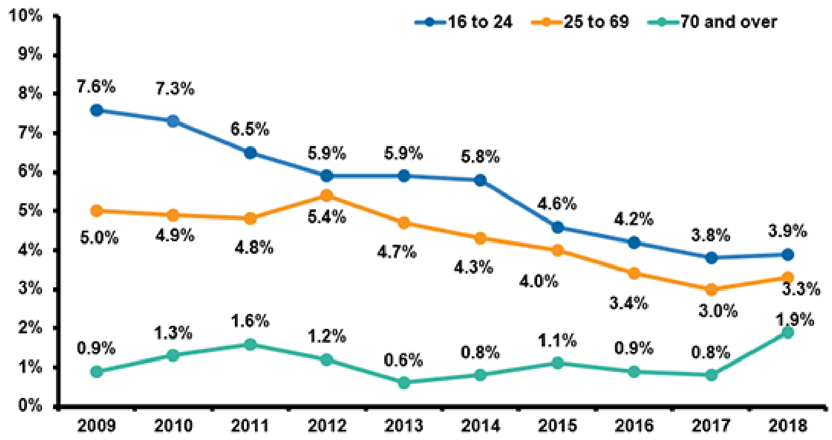


Figure 2-3 Using cellphone while driving among age groups

---

[3] U.S. Department of Motor Vehicles 2017.

[4] For more visit Teensafe.com Distracted Driving Facts and Statistics

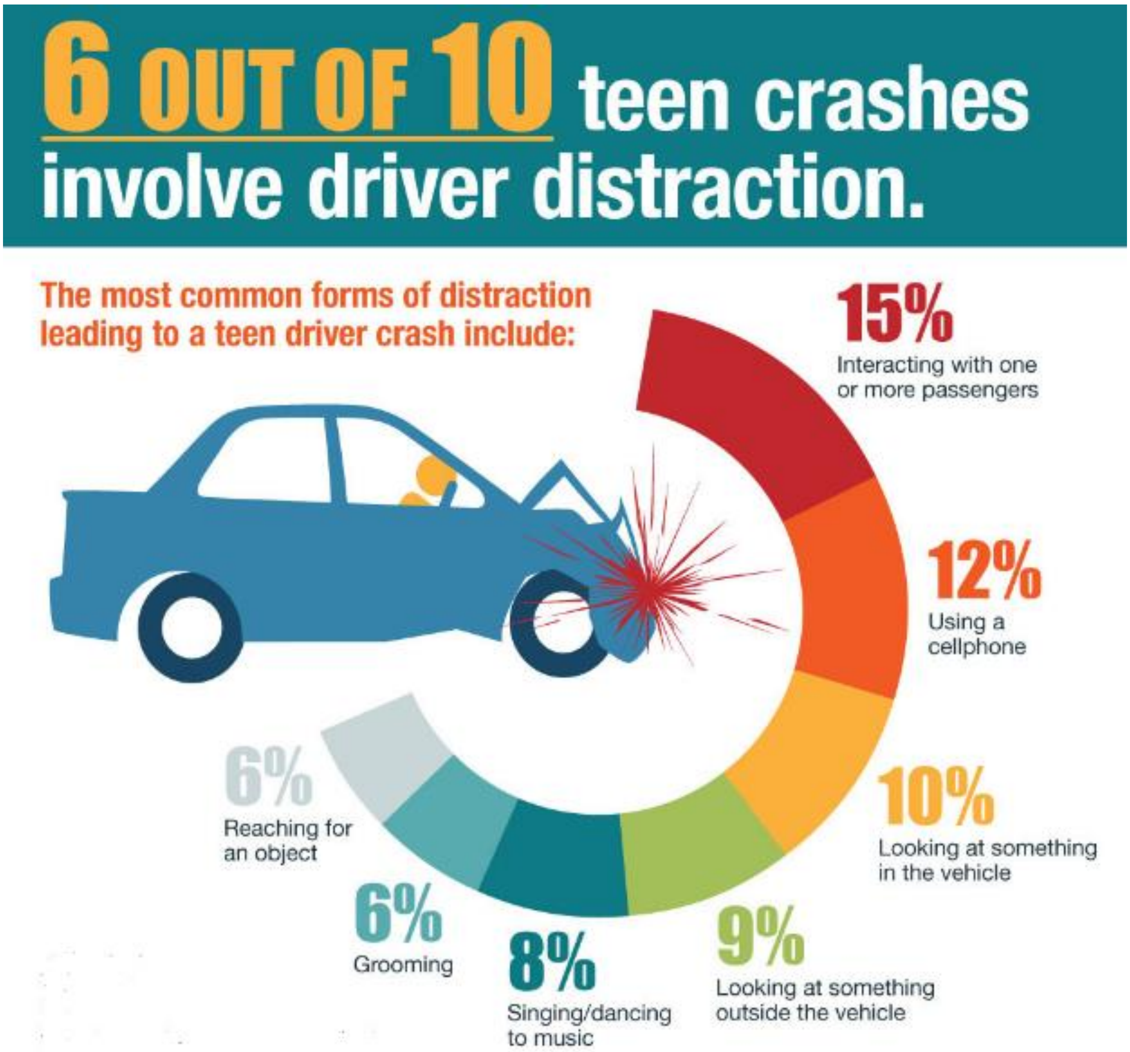


Figure 2-4 Teens and Distracted Driving (teendriving.aaa.com)

3 Current Solutions

Many solutions were developed to solve this rising problem using Embedded systems and sensors but they were not as accurate and not very cost efficient.

3.1 Forward-Collision Warning (FCW)

It provides a visual, audible, and/or tactile alert to warn drivers of an impending collision with a car. Using sensors in front of the car to only alert if the car reaches a specific distance before collision

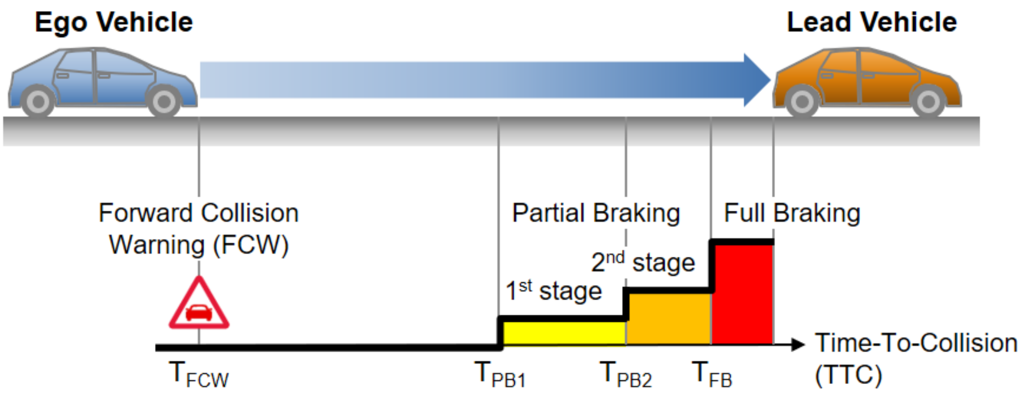


Figure 3-1 FCW and AEB Solution from https://es.mathworks.com/help/driving/examples/autonomous-emergency-braking-with-sensor-fusion.html

3.2 Automatic Emergency Braking (AEB)

If the system senses a potential collision and the driver doesn't react in time, it engages the brakes. It uses the same tech as FCW.

3.3 Apple IOS 11 and AT&T DriveMode

This latest operating system includes a Do Not Disturb While Driving mode (DND) that can block notification of incoming calls and texts when your iPhone senses driving motion or is connected to a car via Bluetooth. (It doesn't block functions that work through Apple's CarPlay system, such as music and navigation.) The DND feature can automatically send a text reply that says you're driving and will reply later. Phone calls are allowed if the iPhone is connected via Bluetooth.

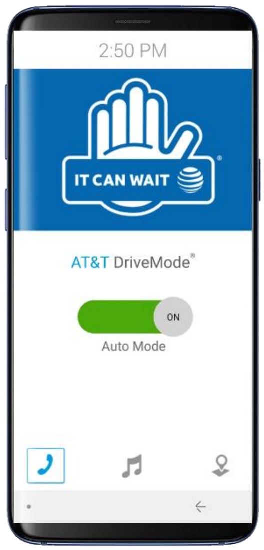


Figure 3-2 AT&T DriveMode App

3.4 Drivesafe.ly and DriveMode apps

This app is designed to read incoming text messages and email aloud for drivers so they keep their eyes on the road.

3.5 Using Machine learning

Using simple machine learning techniques to detect if the driver is distracted or not using facial recognition systems.

4 Our Solution – What do we introduce?!

Our Project will introduce deep learning into the equation to monitor the driver and prevent from being distracted as a packed solution.

The solution will run as an End Device on the car that doesn't need an internet connection.

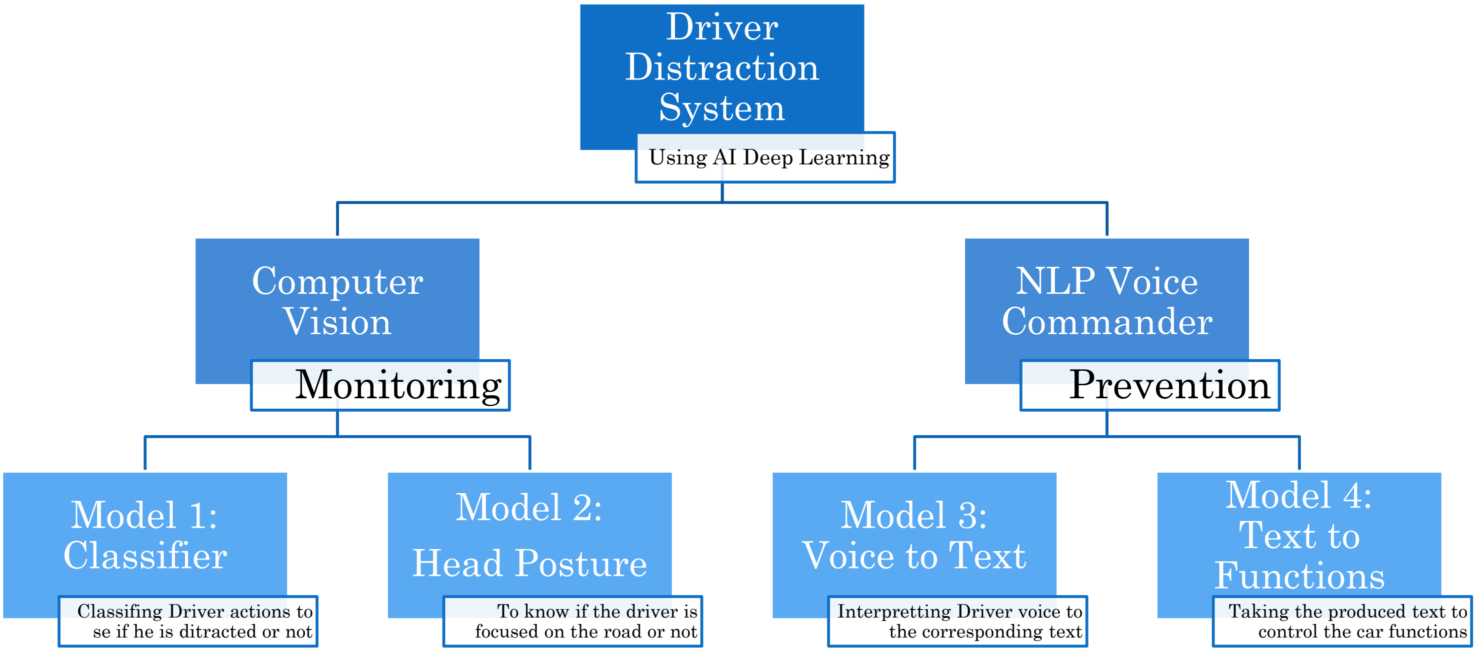


Figure 4-1 Our project in a nutshell

4.1 Monitoring

We need to monitor the driver and see if he is distracted or not (using a camera) by classifying the camera result to know if he is distracted or not, and as an added bonus we will classify what he is actually doing using classes of actions (Safe Driving – Texting – Talking on the phone – Operating the radio – Drinking – Reaching Behind – Talking to a passenger, ETC). Using a Deep Learning Neural Network. We also need to know the driver's eyes are on the road or not. This will be obtained by using a Head Posture Detection model

Figure 4-2 Driver Monitoring from our model using Kaggle Dataset

4.2 Prevention

Introducing a Car Voice Commander to help take commands from the driver and take actions upon these commands without the driver using hands and losing the grip of the steering wheel.

Figure 4-3 Car Voice Command

5 Technical Description

Technically, the system will work as follow

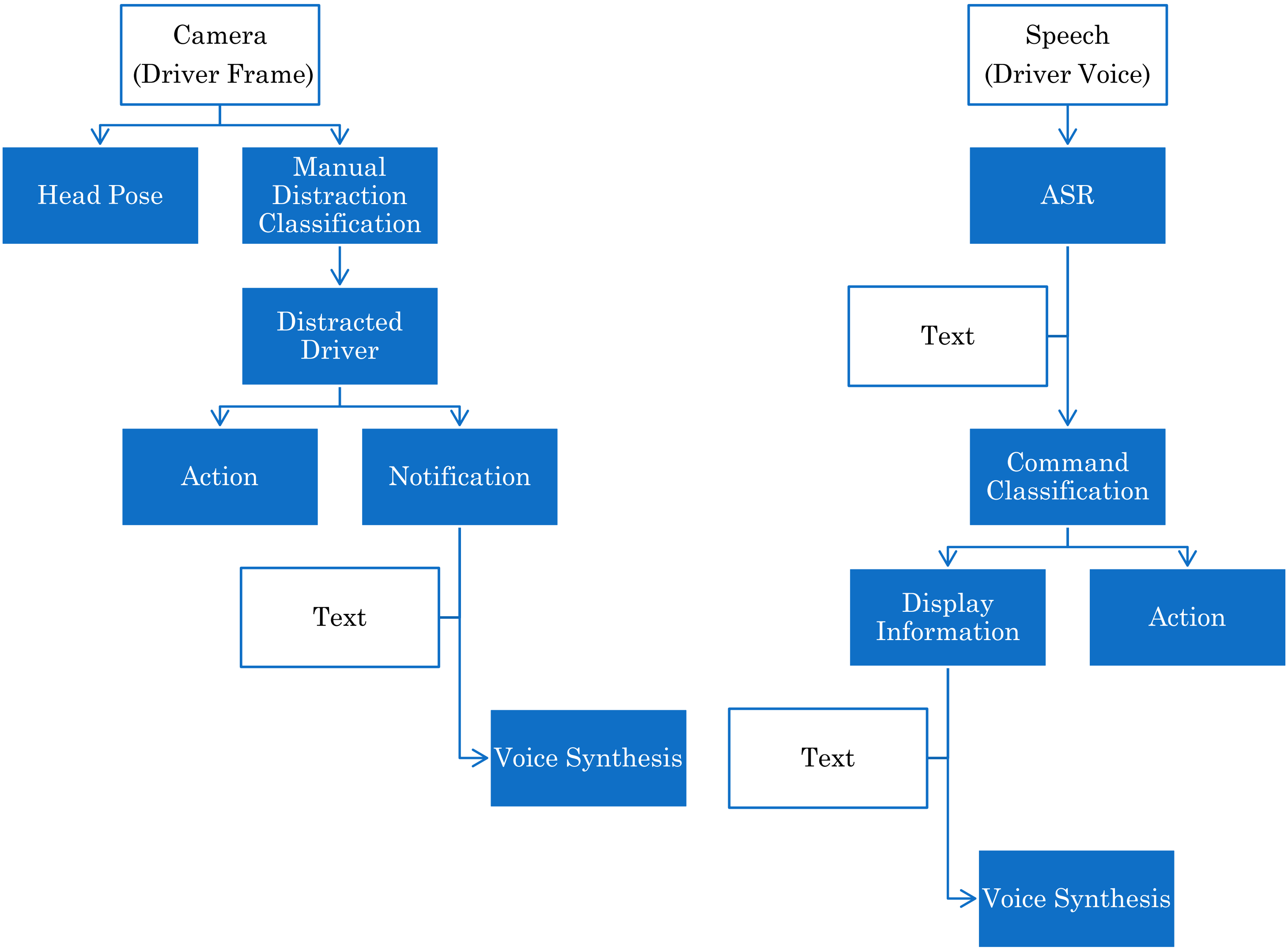


Figure 5-1 Technical Project Scheme

5.1 Datasets

5.1.1 State Farm Distracted Driver Detection

Contains images for driver inside the car with 10 different actions that the driver does in the car (10 Classes). To determine if he is distracted or not.

Figure 5-2 Dataset 10 classes

Class	Number of Images
safe driving	2489
texting - right	2267
talking on the phone - right	2317
texting - left	2346
talking on the phone - left	2326
operating the radio	2312
drinking	2325
reaching behind	2002
hair and makeup	1911
talking to passenger	2129

5.2 Model 1: Head Pose

This Model will detect the angel of the driver's head to know if he is focused on the road or not.

It uses three angles as follow:

The car will have a camera in the front of the driver, and every interval of time, we will take a photo of the driver, then we will use a YOLO Face detection arch to detect the location of the driver face in the image, then we will pass the image of the driver face to a head pose model that first will extract the features from the driver's face Using CNN Arhc and then pass a classifier that will detect the angels of the driver's face in the 3D space.

5.2.1 Architecture

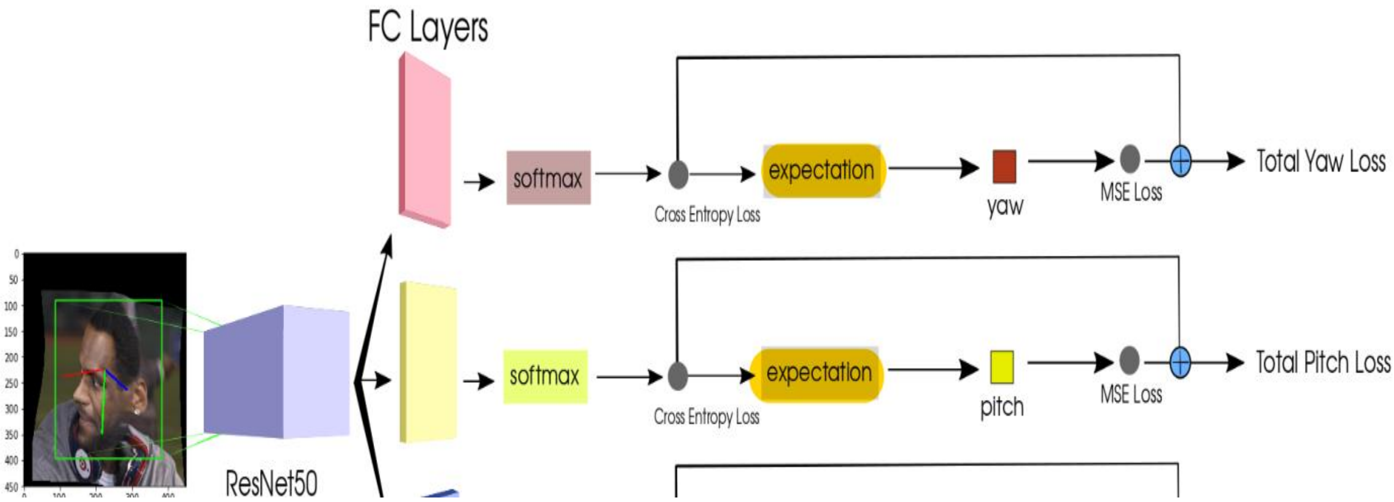


Figure 5-3 Model 1: Head Pose Arch Ref - 8.1

5.3 Model 2: Manual Distraction Classification

This model has data about the driver's actions. And it detects the action of the driver if he is safe driving or doing an action that distracts him (talking to a passenger – holding the phone – texting, etc…).

We will have a camera that will be focused on the driver's body, and every interval of time and we will pass the image to a CNN arch that will classify the actions of the driver based on a predefined 10 action that will be used if the driver is keeping attention to the driving or distracted and if the driver is classified as distracted more than a specific number consecutive frames we will notify the driver to focus on the driving or slowing the car speed and stop the car if the driver didn't take action.

We tried different CNN Models to classify the driver action and our best arch achieved 86 % accuracy, but we will try to enhance the accuracy by focusing the regions of interest in the classification like the driver's hands and head.

5.3.1 Data Augmentation

Our images dataset, each taken in a car with a driver doing something in the car (texting, eating, talking on the phone, makeup, reaching behind, etc.). Our goal is to predict the likelihood of what the driver is doing in each picture.

Choosing the right defect on your images as augmentation[5] can improve your performance we did some experiences to figure out what would lead us to better model evaluation.

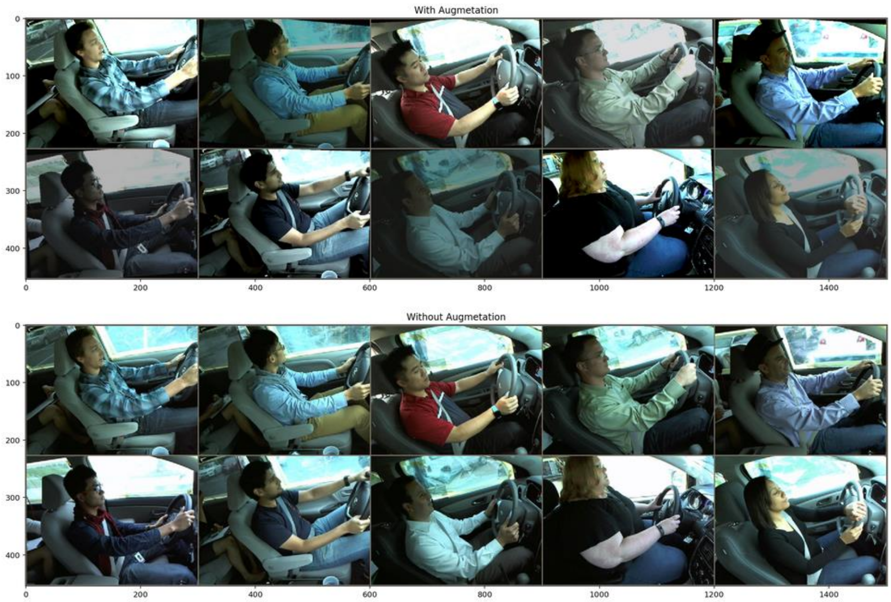


Figure 5-4 Data with (upper) and without Augmentation (lower)

[5] Data augmentation is the process of increasing the amount and diversity of data. We do not collect new data, rather we transform the already present data like resizing, color change, random image flip, etc.

5.3.2 Architecture

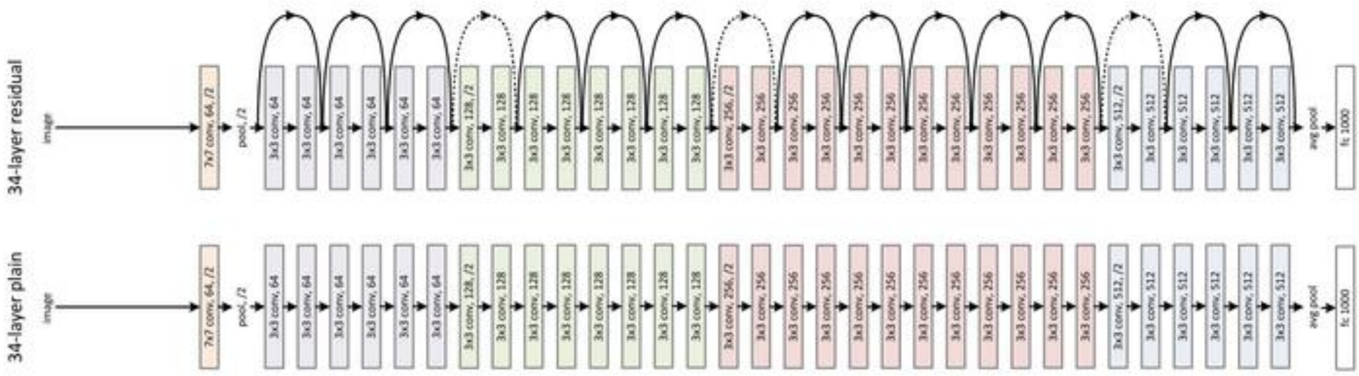


Figure 5-5 Model 3: Distraction Classifier Arch CNN Ref -9.3

5.3.3 Building and Training the Classifier

	Architecture	Accuracy	Notes
Experiment 1	ResNet 152-Layers	78%	Layers all layer freezed
Experiment 2	ResNet 18-Layers	81%	No layer freezed
Experiment 3	ResNet 50-Layers	82%	No freezing
Experiment 4	ResNet 50-Layers	76%	25 layers freezed no resizing – Very Slow
Experiment 5	ResNet 50-Layers	88%	25 layers freezed with resizing

5.3.4 Best Model Evaluation

Classes	Precision	Recall	F1 Score
	Scores		
safe driving	0.79	0.88	0.83
texting - right	0.87	0.83	0.85
talking on the phone - right	0.90	0.94	0.92
texting - left	0.85	0.935	0.89
talking on the phone - left	0.925	0.87	0.89
operating the radio	0.94	0.955	0.94
drinking	0.87	0.89	0.88
reaching behind	0.93	0.94	0.93
hair and makeup	0.82	0.77	0.79
talking to passenger	0.87	0.79	0.82
Total Scores	**0.88**	**0.88**	**0.87**

5.4 Model 3: Speech Recognition

Converts driver's voice into text that will be used in the next model to actions that drivers want without using his/her hands.

We Used an Automatic Speech Recognition model that will take the driver command as a voice then convert it to text so we can process it to execute the command. We used deep speech arch as ASR as it's an open source project and we can use it locally without needing internet connection. Deep speech takes in a voice audio then convert it to a spectrogram that will be easier to process, the it will bass the spectrogram to an RNN arch that will convert it to text at the end.

We created a script that will detect the driver active tacking then it will save the voice command and it will bass it to the model to convert it to text.

5.4.1 Architecture

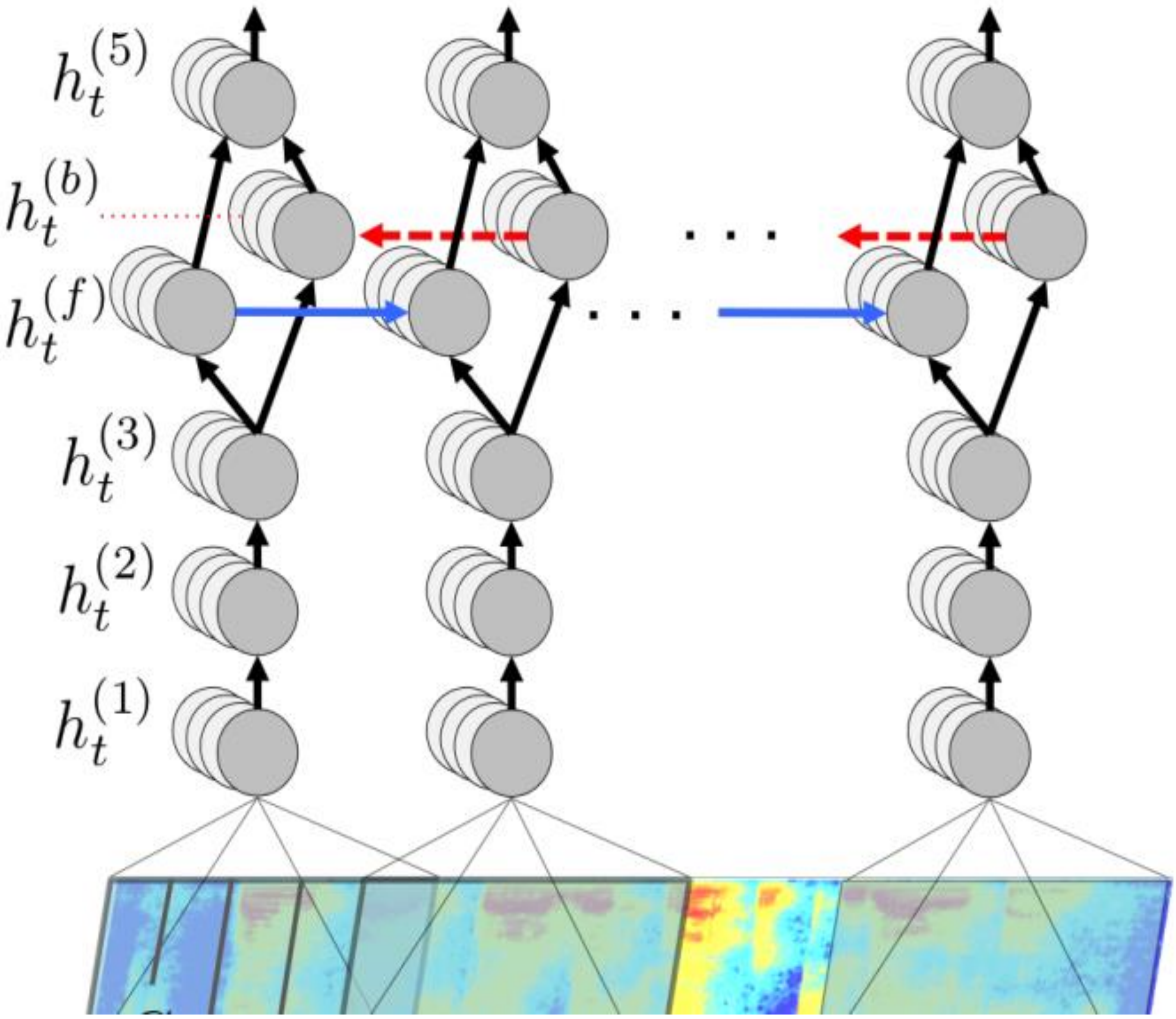
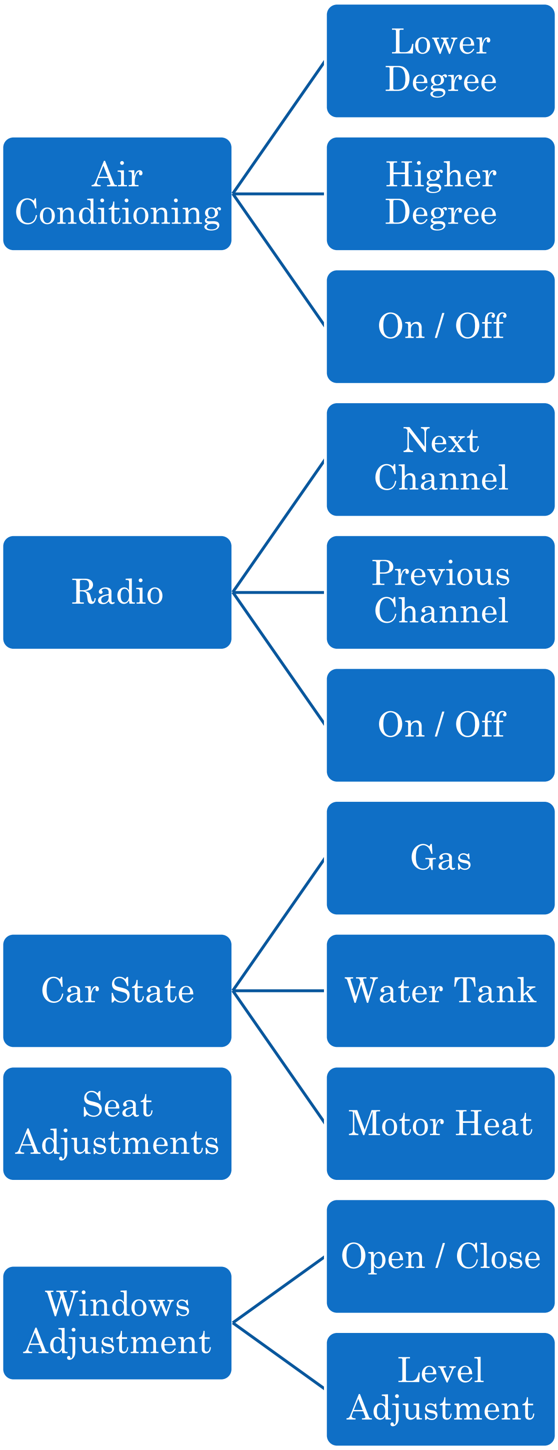


Figure 5-6 Model 3: RNN Arch Ref – 8.2

5.5 Model 4: Commands Classifications

This model will take the text produced by model 3.

After the ASR convert the voice command to text, we will then classify the text to specify the desired command from a predefined set of command and we will also extract the information from text that will be relevant to the command evacuation.

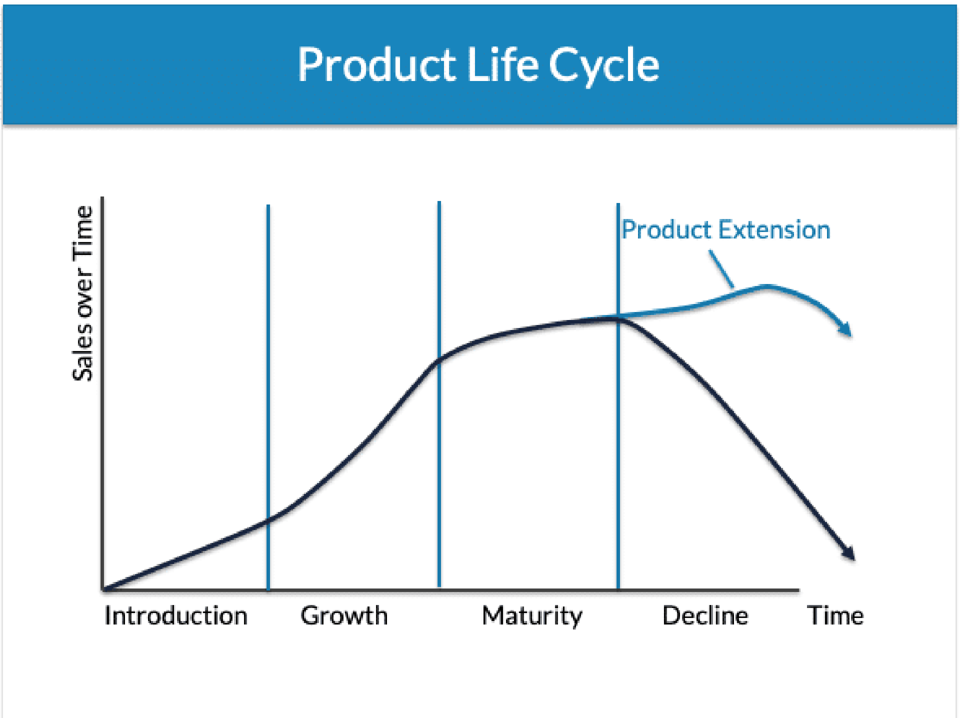


.

Figure 5-7 Voice Commands

6 Planning and Deployment

6.1 Product Development Cycle



6.1.1 Initiation – Introduction

After searching how common driving accidents are. And how driver distraction is one if not the main reason for causing terrible accidents.

And after searching and experiencing most common driver distraction solutions, we came to conclusion that deep learning model as discussed would be a best fit nowadays, by deploying the aspects of prevention using voice commands and detection using image recognition and action classification.

6.1.2 Problem Child – Growth

It is the question of how our solution will fit the market share to grow?

Most solutions in the market are focused on only just one aspect just as discussed before. And they only properly focus on distraction detection. But our solution adds prevention as a bonus.

Since our solution is a standalone solution that can be easily integrated in any car as it doesn't require internet (just a camera – mic and the chip) and it will produce signals to control the car functions.

6.1.3 Maturity

Our plan is to propose this model to the high branded cars. And with our Supervisor(Mercedes Benz – BMW, Egypt sectors) and since our project is supported by Valeo. We are trying to propose an MVP that is updated on the go (Producing fthe minimum Features – testing – adding more features) cycles.

6.1.4 Decline

To extend our project cycle we will add more features to tackle cognitive distraction. And trying to find and acquire cognitive datasets. And this will impressively extend the life span of our solution.

However, with the revolution of self-driving cars that will take our project hopefully as an initial point to go from there, our project will be declined as a standalone solution and it will be converted to a self-driving cars' feature.

6.2 Solution deployment logistics

Target audience and business process

6.3 Quality Assurance

Testing on various datasets. And actual people putting them on a distracted driving simulation.

7 Time Planning

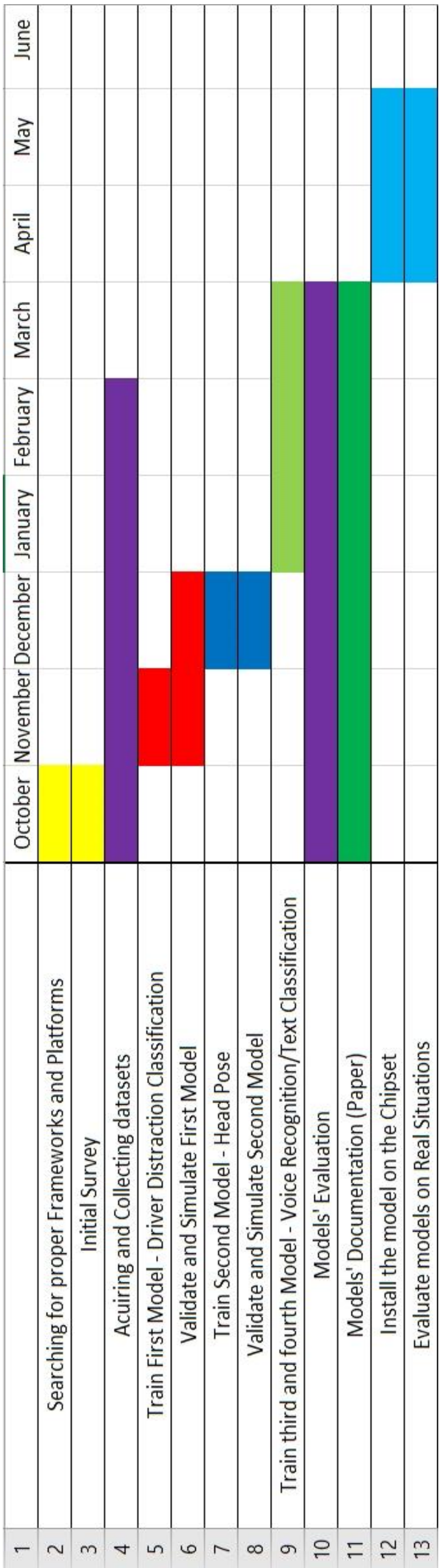


Figure 7-1 Gantt Chart

8 References

8.1 Fine-Grained Head Pose Estimation Without Keypoints, Nataniel Ruiz, Eunji Chong, James M. Rehg, V5 13 Apr 2018.

8.2 Deep Speech: Scaling up end-to-end speech recognition, Awni Hannun, Carl Case, Jared Casper, Andrew Y. Ng, V2 19 Dec 2014

8.3 Deep Residual Learning for Image Recognition, Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, 10 Dec 2015

8.4 Driver Distraction European Commission 2015

8.5 https://www.rospa.com/Road-Safety/Advice/Drivers/Distraction

8.6 https://www.sciencedirect.com/book/9780123819840/handbook-of-traffic-psychology

8.7 Driver Distraction A Sociotechnical Systems Approach, Katie J. Parnell, Neville A. Stanton and Katherine L. Plant

8.8 https://www.nhtsa.gov/risky-driving/distracted-driving

8.9 100 DISTRACTED DRIVING FACTS & STATISTICS FOR 2018

8.10 Facts + Statistics: Distracted driving

8.11 24 Distracted Driving Statistics & Facts – 2019

8.12 Technology That Can Reduce Driving Distractions and Their Dangers