Driver Assistance System

Project report submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Technology
in
Electronics and Communication Engineering

by

Aakash Negi - 15UEC001
Abhimanyu Singhal - 15UEC002
Chaitanya Maheshwari - 15UEC016

Under Guidance of
Dr. Joyeeta Singha

LNMIIT
The LNM Institute of
Information Technology

Department of Electronics and Communication Engineering
The LNM Institute of Information Technology, Jaipur

April 2018

The LNM Institute of Information Technology
Jaipur, India

CERTIFICATE

This is to certify that the project entitled Driver Assisted System , submitted by Aakash Negi (15UEC001), Abhimanyu Singhal (15UEC002) and Chaitanya Maheshwari (15UEC016) in partial fulfillment of the requirement of degree in Bachelor of Technology (B. Tech), is a bonafide record of work carried out by them at the Department of Electronics and Communication Engineering, The LNM Institute of Information Technology, Jaipur, (Rajasthan) India, during the academic session 2017-2018 under my supervision and guidance and the same has not been submitted elsewhere for award of any other degree. In my/our opinion, this thesis is of standard required for the award of the degree of Bachelor of Technology (B. Tech).

Date	Adviser: Dr. Joyeeta Singha

Dedicated to My Family, Friends and Teachers

Acknowledgments

We would like to thank The LNMIIT for providing us with ample knowledge, opportunities and facilities which are playing substantial role in helping us overcoming the challenges along the lines of developing this project. We would like to express our sincere gratitude to our Project Guide, Dr. Joyeeta Singha for the magnanimous guidance in the B.Tech. Project and also in our life in college, which helped us a lot in overcoming a lot of hurdles. We offer our sincere thanks to all the persons who knowingly or unknowingly helped us in completing this project. We would like to utilize this opportunity to thank our parents for being the centre pillars in whatever we have gained and learned so far. Thank you for the blessings and working day in and day out for carving our present and sculpting our future and making us what we are today. Thank you for being the right role models and teaching us how to learn which has been an imperative quality in developing this project.

Abstract

We may not have arrived at a Jetsons-like future of flying automated vehicles just yet, but the self-driving car is getting closer and closer to becoming a widespread, viable method of transportation. Despite this, it will take decades for commercialised autonomous vehicles to run on the roads of India. Meanwhile, drowsy driving or fatigued driving becomes a major concern for road safety. We have intended to curtail drowsy driving by designing a driver assisting system for detecting drowsiness or sleepiness of the driver, while driving. This system will be of a lot of help in places where such cases of drowsy driving related accidents are high. We have designed executable files that can be dumped on a number of hardwares and can be used thereon for practical usage. Our system detects drivers' face in real time and if it detects tired eyes, or driver to be yawning, it rings an alarm in the car to focus the drivers' attention back to the road and will help to avoid a potential accident.

Contents

Chapter *1*

Introduction

1.1 The Area of Work

Our area of work is related to computer vision, image processing and analysis. Computer Vision is the art of distilling actionable information from images. Algorithm development is central to image processing and computer vision because each situation is unique, and good solutions require multiple design iterations. Computer vision is an interdisciplinary field that deals with how computers can be made for gaining high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do. Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subcategory or field of digital signal processing, digital image processing has many advantages over analog image processing. It allows a much wider range of algorithms to be applied to the input data and can avoid problems such as the build-up of noise and signal distortion during processing. Since images are defined over two dimensions (perhaps more) digital image processing may be modeled in the form of multidimensional systems.

Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images, and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information, e.g., in the forms of decisions. Understanding in this context means the transformation of visual images (the input of the retina) into descriptions of the world that can interface with other thought processes and elicit appropriate action. This image understanding can be seen as the disentangling of symbolic information from image data using models constructed with the aid of geometry, physics, statistics, and learning theory. As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, or multi-dimensional data from a medical scanner. As a technological discipline, computer vision seeks to apply its theories and models for the construction of computer vision systems. Sub-domains of computer vision include scene reconstruction, event detection, video tracking, object recognition, 3D pose estimation, learning, indexing, motion estimation, and image restoration.

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subcategory or field of digital signal processing, digital image processing has many advantages over analog image processing. It allows a much wider range of algorithms to be applied to the input data and can avoid problems such as the build-up of noise and signal distortion during processing. Since images are defined over two dimensions (perhaps more) digital image processing may be modeled in the form of multidimensional systems.

1.2 Problem Addressed

Sleep-deprived driving (commonly known as tired driving, drowsy driving, or fatigued driving) is the operation of a motor vehicle while being cognitively impaired by a lack of sleep. Sleep deprivation is a major cause of motor vehicle accidents, and it can impair the human brain as much as alcohol can. According to a 1998 survey, 23% of adults have fallen asleep while driving. When a person does not get an adequate amount of sleep, his or her ability to function is affected. As listed below, their coordination is impaired, have longer reaction time, impairs judgment, and memory is impaired. Sleep deprivation has been proven to affect driving ability in four areas:

1. It impairs coordination.
2. It causes longer reaction times.
3. It impairs judgment.
4. It impairs memory and ability to retain information.

Sufficient sleep before driving improves memory. Researchers recorded activity in the hippocampus during learning, and recorded from the same locations during sleep. The results were patterns that occurred during sleep resembled those that occurred during learning, except they were more rapid during sleep. Also, the amount of hippocampal activity during sleep correlated highly with a subsequent improvement in performance. Signs that tell a driver of a need to stop and rest are as follows:

1. Difficulty focusing, frequent blinking, or heavy eyelids.
2. Daydreaming; wandering/disconnected thoughts.
3. Trouble remembering last few miles driven or missing exits and street signs.
4. Yawning repeatedly/rubbing eyes.
5. Trouble keeping head up. Drifting from lane to lane, tailgating, or hitting a shoulder or rumble strip.
6. Feeling restless and irritable.

It has been estimated that approximately 20% of vehicle accidents have sleep deprivation as a cause. Accidents related to sleep deprivation are most likely to happen in the early to midafternoon, and in the very early morning hours. The reason that accidents are mostly likely to happen during the early to mid afternoon may have to do with the biological time clock. Each person's body has its own. Most people run on a daily rhythm of approximately 24 hours, but this can vary from person to person. The reason night time driving is so risky is because sleep becomes an irresistible urge especially from about midnight until 6 a.m. A sleepy period is also ”programmed” for the afternoon which makes that a risky time.

Chapter *2*

Existing Algorithms

2.1 Existing Algorithms

Existing algorithms for Face Detection in Image Processing can be classified as follows:

1. Viola –Jones Object Detection Framework
2. Principal Component Analysis Method
3. Fisherfaces

2.1.1 Viola –Jones Object Detection Framework

The basic principle of the Viola-Jones face detection algorithm is to scan the detector many times through the same image - each time with a new size. Even if an image should contain one or more faces it is obvious that an excessive large amount of the evaluated sub-windows would still be negatives (non-faces). This realization leads to a different formulation of the problem: Instead of finding faces, the algorithm should discard non-faces. The thought behind this statement is that it is faster to discard a non-face than to find a face.

$$h(x) = sgn\left(\sum_{j=1}^{M} \alpha_j h_j(x)\right) \tag{2.1}$$

2.1.2 Principal Component Analysis Method

One of the simplest and most effective PCA approaches used in face recognition systems is the so-called Eigenface approach. This approach transforms faces into a small set of essential characteristics, Eigenfaces, which are the main components of the initial set of learning images (training set). The whole recognition process involves two steps:

1. Initialization process
2. Recognition process

2.1.2.1 Initialization process

1. Acquire the initial set of face images called as training set.
2. Calculate the Eigenfaces from the training set, keeping only the highest eigenvalues.
3. Calculate distribution in this M-dimensional space for each known person by projecting his or her face images onto this face-space.

2.1.2.2 Recognition process

1. Calculate a set of weights based on the input image and the M eigenfaces by projecting the input image onto each of the Eigenfaces.
2. Determine if the image is a face at all (known or unknown) by checking to see if the image is sufficiently close to a "free space ".
3. Update the eigenfaces or weights as either a known or unknown.

2.1.3 Fisherfaces

The Fisherface method is an enhancement of the Eigenface method that it uses Fishers Linear Discriminant Analysis (FLDA or LDA) for the dimensionality reduction. The LDA maximizes the ratio of between-class scatter to that of within-class scatter, therefore, it works better than PCA for purpose of discrimination. The Fisherface is especially useful when facial images have large variations in illumination and facial expression.

Chapter *3*

Litrature Survey

3.1 Introduction

During the last decades we have witnessed great advancement in the field of Image Processing. Most of them are done for the problem of automatic facial point detection. Numerous methodologies have been proposed that are shown to achieve great accuracy and efficiency.

3.2 Database

The main reason why many researchers of the field focus on the problem of face alignment is the plethora of publicly available annotated facial databases. These databases can be separated in two major categories:

1. The first category includes databases that are captured under controlled conditions, normally within special indoor laboratories/studios in which the camera position and the lighting source and intensity can be controlled. In most of these databases, each subject is asked to perform a posed facial expression, thus we find more than one images per subject. The most popular such databases are Multi-PIE , FRGC-V2, XM2VTS and AR.

2. The facial databases of the second major category consist of images that are captured under totally unconstrained conditions (in-the-wild). In most cases, these images are downloaded from the web by making face-related queries to various search engines. The most notable databases of this category are LFPW , HELEN , AFW , AFLW and IBUG. [2]

3.3 One Millisecond Face Alignment Method

This method presents an algorithm to precisely estimate the position of facial landmarks in a computationally efficient way which utilizes a cascade of regressors. [1]

$$\hat{S}^{t+1} = \hat{S}^{(t)} + r_t(I, \hat{S}^{(t)}) \tag{3.1}$$

This method starts by using:

1. A training set of labeled facial landmarks on an image. These images are manually labeled, specifying specific (x, y)-coordinates of regions surrounding each facial structure. Priors, of more specifically, the probability on distance between pairs of input pixels.

2. Priors, of more specifically, the probability on distance between pairs of input pixels.

Given this training data, an ensemble of regression trees are trained to estimate the facial landmark positions directly from the pixel intensities themselves (i.e., no feature extraction is taking place). The end result is a facial landmark detector that can be used to detect facial landmarks in real-time with high quality predictions.

The indexes of the 68 coordinates can be visualized on the image below:

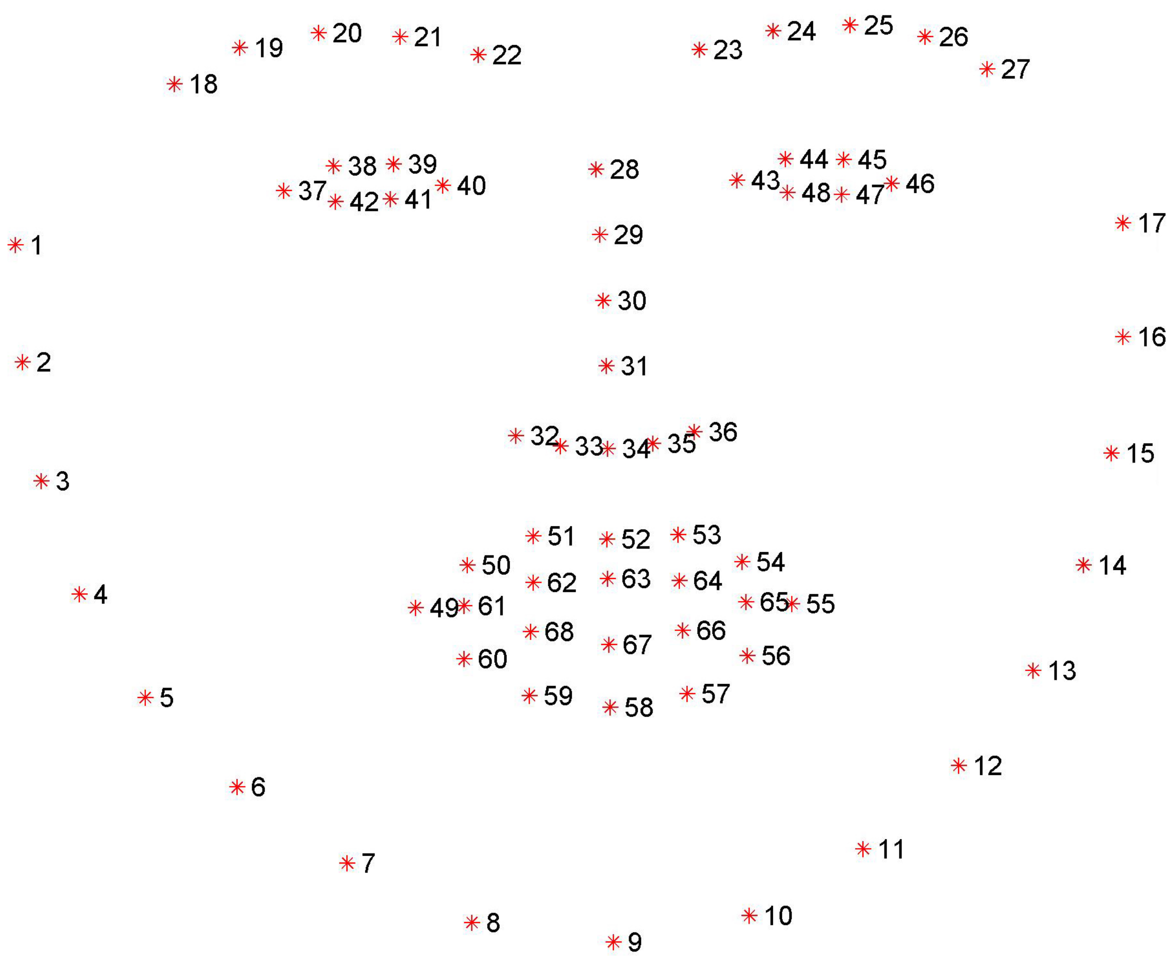


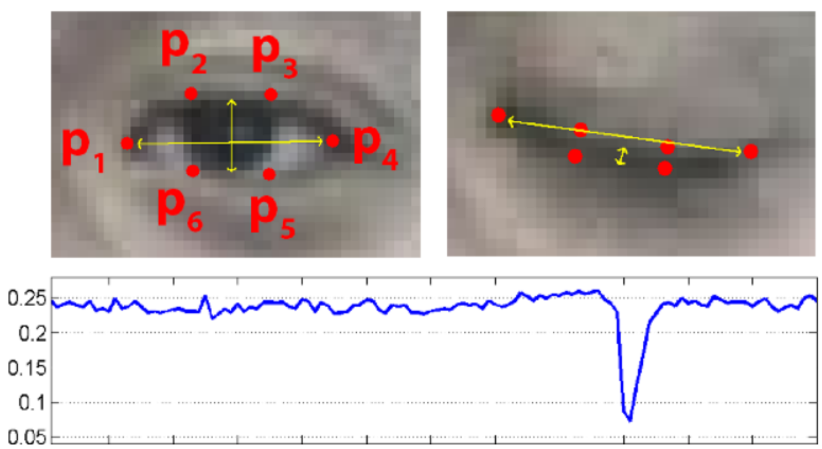
Figure 3.1: Visualizing the 68 facial landmark coordinates from the iBUG 300-W dataset. [3]

3.4 Eye Tracking and Detection

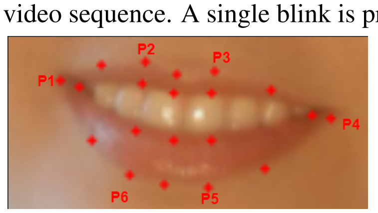
Detecting eye blinks is important for instance in systems that monitor a human operator vigilance, e.g. driver drowsiness. The eye blink is a fast closing and reopening of a human eye. Each individual has a little bit different pattern of blinks. The pattern differs in the speed of closing and opening, a degree of squeezing the eye and in a blink duration. The eye blink lasts approximately 100-400 ms. We propose to exploit state-of-the-art facial landmark detectors to localize the eyes and eyelid contours. From the landmarks detected in the image, we derive the eye aspect ratio (EAR) that is used as an estimate of the eye opening state. Since the perframe EAR may not necessarily recognize the eye blinks correctly, a classifier that takes a larger temporal window of a frame into account is trained.

The eye aspect ratio (EAR) between height and width of the eye is computed by the following equation

$$EAR = \frac{||p2 - p6|| + ||p3 - p5||}{2||p1 - p4||} \tag{3.2}$$



(a) Open and closed eyes with landmarks automatically detected. The eye aspect ratio EAR plotted for several frames of a video sequence. A single blink is present. [4]



(b) Mouth Landmarks Detected

The EAR is mostly constant when an eye is open and is getting close to zero while closing an eye. It is partially person and head pose insensitive. Aspect ratio of the open eye has a small variance among individuals and it is fully invariant to a uniform scaling of the image and in-plane rotation of the face. Since eye blinking is performed by both eyes synchronously, the EAR of both eyes is averaged.

Chapter *4*

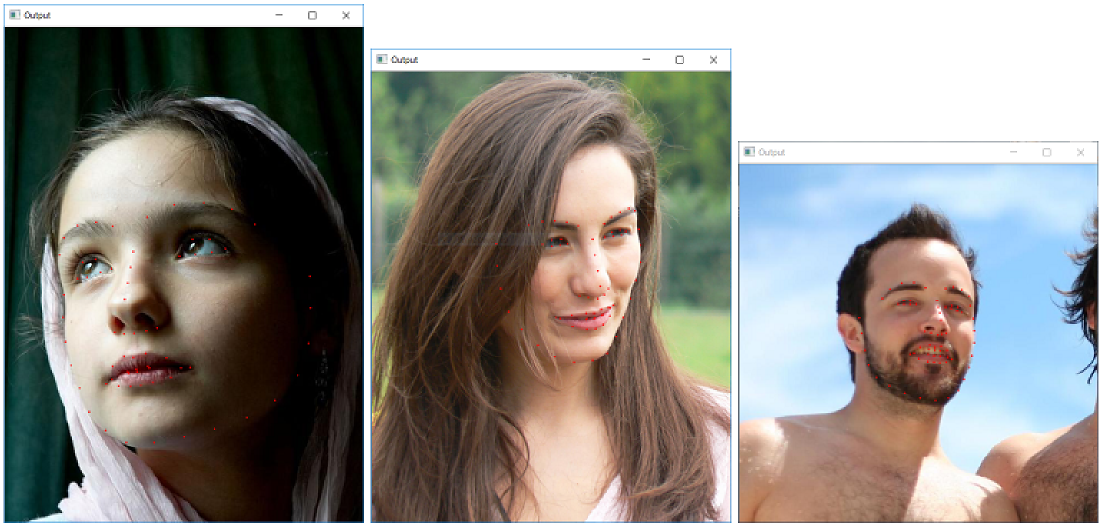
Simulation and Results

4.1 Sample Images For Helen Database

(a) Sample Image 1 (b) Sample Image 2 (c) Sample Image 3

Figure 4.1: Sample Images

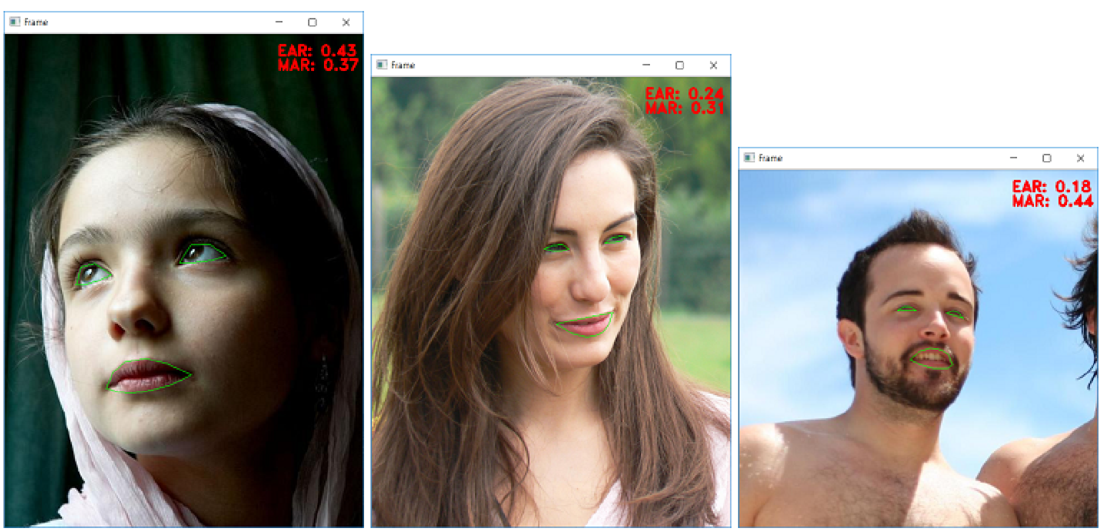
4.2 One Millisecond Face Alignment Method



(a) Sample Image 1 (b) Sample Image 2 (c) Sample Image 3

Figure 4.2: Applied On Helen Database

4.2.1 Facial Contours Detected Using Our Code



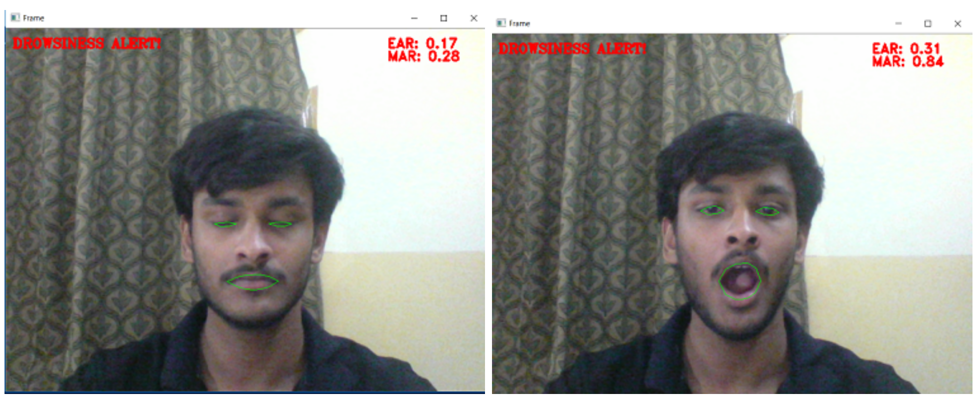
(a) Sample Image 1 (b) Sample Image 2 (c) Sample Image 3

Figure 4.3: Applied On Helen Database

4.3 Results

Here are the following results that we obtained

Figure 4.4: Normal face and eye detection.



(a) Tired eyes detected.

(b) Yawn detected

Figure 4.5: Drowsiness detection.

Chapter *5*

Conclusions and Future Work

5.1 Conclusion

We explored the problems faced by people due to driver drowsiness and chose to step in. We observed different databases and chose the most suitable one for our aim, which came out to be Helen dataset. We trained the database with variant alignments to the camera, using One Millisecond Face Alignment Method. Face detection was used to detect faces from our live video stream, frame by frame. On the event of detecting a face, eye and mouth were detected separately and then were passed on for drowsiness check. For checking a face for drowsiness, aspect ratios of eyes and face were measured respectively, and upon the aspect ratios (EAR, MAR) crossing a pre-calculated threshold value, driver was alerted.

Computer vision using image processing is a near-to-perfect way to use images in form of data and working out algorithms accordingly, but it comes with its own set of limitations. Camera quality should be good, which comes with a slightly high price. Also, if the system is to be used in night, then the camera must have night light, adding to the budget. The hardware where we dump should be really fast for real-time detection, as time is of essence in such cases, where fatal accidents are a possibility. One major problem that doesn't seem to be able to be rectified is if the driver is using sunglasses. In such a case, our system is unable to detect and driver's eyes and is rendered useless. Despite the above mentioned limitations, our system has proved time and again to produce desirable results, thus is ready to be used by anyone who plans to launch our executable file on a suitable hardware with an appropriate camera, and can be used to help the problem of driver drowsiness and resultant accidents to quite an extent. Our system can and should be used to reduce the number of road accidents till at least self driving cars on the roads of India become a reality, and perhaps even after that with certain adjustments and rigorous betterment of the same.

5.2 Proposed work for next Semester

In this semester, we have completed the driver detection part of our project. In the upcoming semester we will be working towards the detection of the roads in real time. For that, we will be taking the help of the deep learning model CNN through which we will train our machine to identify roads. Since our project is mainly focussed on the safety of the driver, we will also do car and pedestrian detection so that the driver may know if any object is coming close to his car or not. Also, we will be adding a distance sensor to our working model so that the drive could be notified whether any object is coming dangerously close to his car or not and the driver could avoid it.

Bibliography

[1] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *CVPR*, 2014.

[2] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang. Interactive facial feature localization. In *UIUC*, 2012.

[3] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: Database and results. In *Image and Vision Computing (IMAVIS), Special Issue on Facial Landmark Localisation "In-The-Wild"*., 2016.

[4] T. Soukupova and J. Cech. Real-time eye blink detection using facial landmarks. In *Computer Vision Winter Workshop*, 2014.