Machine Learning Engineer Nanodegree

Capstone Proposal

Ameen Al Baz
August 7th, 2018

Detecting drivers' attention and mood using CNN

Domain Background

The National Highway Traffic Safety Administration (NHTSA) 2016 data shows that an alarming 37,461 people were killed in 34,436 motor vehicle crashes, an average of 102 per day [1]. Most road accidents occurred due to human error. Thankfully advanced driver-assistance systems such as traction control, adaptive cruise control, lane departure warning, anti-lock brakes, and electronic stability control, have been adopted to mitigate human error and improve vehicle safety.

Despite these advanced driver-assistance systems, no current technology is implemented in vehicles that respond to the state of the driver while they are operating the vehicle. For example, collision assist will attempt to apply brakes if the driver is predicting a collision with an object or vehicle ahead, regardless if the driver is alert or not. This can frustrate some drivers who do not want to hear the warning sound of the system and will turn off or change the settings of the system to make it less strict, which can increase the risk of collision. Advanced driver-assistance systems mostly act on exterior conditions and have little to no access to information about the state of the driver.

Problem Statement

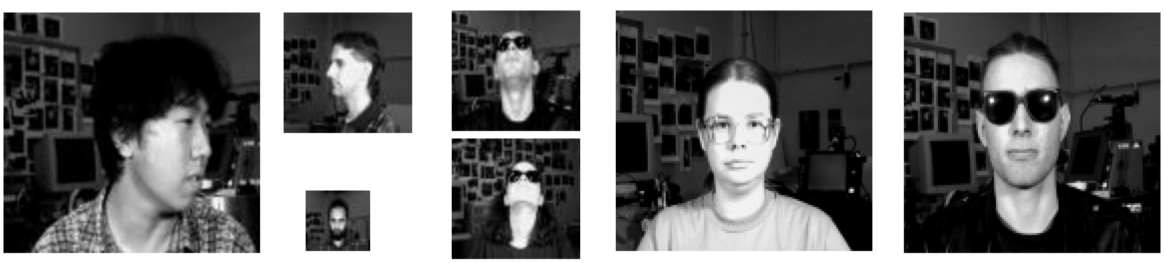
The state of the driver must be determined through a non-obstructive method and with an accuracy at least higher than random chance. The direction the driver is facing, as well as their emotional state must be determined regardless of whether the user is wearing sunglasses or not while driving. With this information, advanced driver-assistance systems will also include the driver's condition along with the road conditions while it is in operation to provide a seamless and enhanced driving experience.

Datasets and Inputs

Ideally, the dataset should include a collection of images that contain the driver operating a vehicle in real world driving conditions: with their head facing different directions, including straight ahead, while expressing different moods, both while wearing sunglasses or without, and under different lighting conditions. Unfortunately, no such dataset exists yet and it will difficult to create the dataset from scratch. Thus, the closest dataset available will be used: CMU Face Images Data Set [2].

The CMU Face Images data consists of 640 black and white face images of people taken and categorized with varying pose (straight, left, right, up), expression (neutral, happy, sad, angry), eyes (wearing sunglasses or not), and image size. Figure 1 below shows samples from the dataset.

Figure 1: Samples from the CMU Face Images data set labelled according to pose, expression, eyes/sunglasses, and image size.



For the different image sizes in the set, the original 128x120 size image is reduced to half the resolution for a medium size image and reduced again for a small size image. Only the largest copy of the images in the dataset will be used for the capstone because smaller image resolutions are not relevant to the problem. The ID label will be dropped from the image as well as facial recognition is not in the scope of the problem. The labels remaining that will be used for the capstone are head orientation, mood, and sunglasses on/off. These labels align with the problem statement which requires the detection of head orientation and driver mood regardless of whether they are wearing sunglasses or not.

Solution Statement

An onboard camera directed at the driver's head captures a few images per second of the driver. The onboard camera may have an infrared illuminator to obtain clear images of the driver at night, while remaining undetectable and nonintrusive. The images from the camera will serve as the input and an onboard computer will run an algorithm to output a multi-classification for the images. The multi-classifier will classify if the driver is distracted based on head orientation and classify the mood of the driver based on their facial expression. The algorithm must perform these classifications even if the driver is wearing sunglasses.

The advanced driver-assistance systems can then use this information to cater to the driver. For example, if the system notices that the driver is not looking straight ahead, it will employ stricter auto-braking rules to prevent a crash, and alert the driver if they are distracted for too long. Ultimately, driver's attention and mood should be provided as an additional input to the advanced driver-assistance systems to seamlessly and dynamically adjust its settings to correspond with the driver's state.

Benchmark Model

Currently, most proposed solutions to monitoring the driver state is by the use of an eye-tracker. Tobii Tech's eye tracking system is one of the newly available systems that can detect driver distraction and collaborate with the advanced driver assistance systems in the vehicles to respond accordingly [3]. However, the system does not have mood detection, and more importantly not reliable if the user is using eyewear [4]. No other data was provided by Tobii Tech regarding accuracy.

Evaluation Metrics

The evaluation metrics that will be used to evaluate the benchmark and proposed solutions are outlined in Table 1 below. The values for the benchmark evaluation metrics are known, except for the accuracy. Because the product is advertised as fully functional, it will have an assumed accuracy of 100% without sunglasses.

Table 1. Evaluation metrics for the benchmark and proposed solution. Both solutions will be evaluated by their accuracy to correctly classify the distraction state of the driver as well as the mood of the driver, with and without sunglasses.

	Distr. Acc. w/ glasses	**Distr. Acc. w/ clear**	**Mood Acc. w/ glasses**	**Mood Acc. w/ clear**	**Avg Score /10**
Benchmark	0%	100%	0%	0%	2.5
Proposed					

Project Design

The first step in this project is to examine the data and determine the data size and balance, and to augment the images. Images can be flipped horizontally as well as slight rotations and shifts can be applied to increase the number of images available as well as balance the dataset. Duplicate images that have a reduced resolution from the original image, as well as corrupted images, will not be used in the project.

Because the problem involves classifying images, a convolutional neural network will be used. CNNs do not require as much pre-processing compared to other image classification algorithms because the network can learn the necessary filters that are otherwise hand-engineered. The proposed solution is to use a pretrained convolutional neural network as a feature extractor and training a machine learning classifier to classify the extracted features. The pretrained convolutional networks that will be explored are ones that were trained on images of humans. This increases the chances that the features extracted from these networks are relevant to this project, which involves classifying the state of humans.

Once the features are extracted, three different machine learning algorithms will be explored to train on the features: a neural network, support vector machines, and XGBoost. Both neural network and support vector machines are standard algorithms that are often used to classify features extracted from a CNN, while XGBoost is an optimized, distributed gradient boosting library that has recently gained popularity for being highly efficient and flexible. It is often used on tabular data, but for exploration, will be trained on CNN extracted features.

After the models are trained and evaluated, the model that scores the highest average score on Table 1 will be compared against the benchmark model to determine if the proposed solution is more effective. In proving so, it can be concluded that the proposed solution is able to provide relevant information about the driver’s state to the advanced driver assistance systems in current and future vehicles, which can increase the safety and wellbeing of the driver.

References

[1] https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812451.

[2] https://archive.ics.uci.edu/ml/datasets/cmu+face+images

[3] https://www.tobii.com/tech/products/automotive/

[4] https://help.tobii.com/hc/en-us/articles/210249865-Glasses-lenses-and-eye-surgery