Structured Abstract

Context

1.35 million people die in car accidents and most of the accidents happen because of drivers doing multiple tasks while driving. In other words, getting distracted. State Farm collected the data of several subjects in various situations(mentioned below) in a controlled environment.

Objective

State Farm wants to make use of machine learning and deep learning techniques to predict the current state of the driver. Whether the driver is driving safely or doing something else while driving.

Method

To carry out this analysis, python as a language and some of its packages like sklearn, TensorFlow and Keras were used. The agile process(the whole process was done into multiple iterations) was used to make sure the results are reliable.

Results

The best model was CNN with transfer learning using VGG16 pre-trained model with training accuracy of 95%, validation accuracy of 80% and Test accuracy of 75-80% (tested manually by randomly selecting 20 images from the images provided in the test folder). Also tried traditional machine learning algorithm to see how it stack against CNN.

Novelty

Previous publications mainly focused on various CNN architectures. However, we also tried traditional machine learning algorithm but it was very time consuming to extract features from images hence, we used only 4000 images which were only 22% of the training set and the model gave an accuracy of only 13% which can definitely be increased by using all the images. Overall, CNN with transfer learning was the winner here.

S.No.	Model	Accuracy
1	Traditional Machine Learning Algorithm (Random Forest)	13%
2	CNN Trained From Scratch (VGG16 Architecture)	11%
3	CNN with Transfer Learning using VGG16 Pre-Trained Model	80%