Machine Learning Engineer Nanodegree

Capstone Proposal

Ayşin Taşdelen
October 11st, 2017

Using Computer Vision Detect Distracted Driver

Domain Background

State Farm is interested in alleviating the problem of car accidents caused by distracted drivers. According to AAA foundation, foundation for traffic safety, more than 80% of drivers in the annual AAA Foundation Traffic Safety Culture Index cite distraction as a serious problem and a behavior that makes them feel less safe on the road.Distracted driving is a deadly behavior. Federal estimates suggest that distraction contributes to 16% of all fatal crashes, leading to around 5,000 deaths every year. ***AAA's latest research*** has discovered that distraction "latency" lasts an average of 27 seconds, meaning that, even after drivers put down the phone or stop fiddling with the navigation system, drivers aren't fully engaged with the driving task.

In this project I created a web application which predicts the likelyhood of what the driver is doing in each picture from the State Farm Kaggle competition dataset.

Problem Statement

The goal is to implemet and train a classifier that can predict the likelihood of what actual action the driver is doing (multi-class classification)by using computer vision and machine learning methods. The tasks involved are the following:

1. Split the given training data into train and validation data.
2. Preprocessing the data with computer vision methods.
3. Train a classifier with a preprocessed training data that can predict if an image has a distracted driver.
4. Predict the likelihood using this model of what the driver is doing using the test set.
5. Computing accuracy of the model.

The final application is expected to be usefull for determining distracted users and what are they occupied with. Therefore, this project can be usefull building later

projects like installing distracted driver systems in peoples cars for everyday use to warn distracted drivers and prevent accidents.

Datasets and Inputs

In this Problem , I will use the dataset from State Farm Distracted Driver Detection competition on Kaggle. In this dataset I am given driver images, each taken in a car with a driver doing something in the car (texting, eating, talking on the phone, makeup, reaching behind, etc).

The dataset consist of imgs folder and driver_imgs_list.csv file. İmgs folder has test and train folder of 640 x 480 jpg files. Each image consist of a driver performing a task from the list below:

c0: safe driving 2489
c1: texting – right 2267
c2: talking on the phone – right 2317
c3: texting – left 2346
c4: talking on the phone – left 2326
c5: operating the radio 2312
c6: drinking 2325
c7: reaching behind 2002
c8: hair and makeup 1911
c9: talking to passenger 2129

Total training data: 22424

Maximum file number is : 2489
Minimum file number is : 1911
Mean is: 2242.4
Median is: 2314.5

Safe driving has the most examples and it is close to median so this will not effect our model dractically. Moreover, mean and median is close to each other meaning data is evenly divided around the mean and median is bigger than mean that more of our data is on the right side of the mean. Hair and makeup has the least examples. This might be because we do not expect all of our drivers to perform this task at all.

Total testing data: 79726

Total testing data is close to 3 times more than training data which can be enough to calculate the accuracy of this model.

driver_imgs_list.csv file lists training images, driver names and tasks. As we can see from the figure below, we have 26 drivers. This number seems to be small but State Farm generated this data by experimenting controlled environment with drivers, thus this number might be kept intentionally.

	subject	classname	img
count	22424	22424	22424
unique	26	10	22424
top	p021	c0	img_97080.jpg
freq	1237	2489	1

Solution Statement

We have sophisticated images with complicated patterns. Using CNN is best choice to make multi-class image classification. I wil construct my CNN with 5 convulutional layer, 5 pooling layer, 1 dropout layer to prevent overfitting. I will use python as software language, Keras and Tensorflow libraries to train my models. I will train my model using my p2.xlarge aws GPU compute enalbled instance.

Benchmark Model

I will augment the given test data and use transfer learning using the model I’ve built as solution than compare their accuracies. I will use this augmented data and input them to ResNe-50 model and process them than take the transfer values ans save them. Use them as a new input to a different neural network. Train this network with one fully connected layer and a softmax layer (10 feature in size because we have 10 different classes)

Evaluation Metrics

I will use logloss function evaluation metric for both models. State Farm Kaggle competition submissions were evaluated using the multi-class logarithmic loss. Each image has been labeled with one true class. For each image, I will calculate a set of predicted probabilities (one for every image). The formula is then,

$$logloss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\log(p_{ij}),$$

Project Design

I summarized some of the tasks briefly at problem statement section.If I expand the task list to my intentional total task list :

1. Since I don't need any color to detect an object in the driver's hand or orientation of their head, I don't need any color. Thus, I will convert my images to grayscale images.
2. Rescale the images by dividing every pixel in every image by 255.
3. From Keras with Model_CheckPoint I will split validation data from the training data.
4. Train CNN with 5 convulutional layer and 5 pooling layer and 1 droopout layer with my data that can predict if an image has a distracted driver from one of the classes .
5. Predict the likelihood using this model of what the driver is doing using the test set.
6. Compute the test accuracy.
7. Augment the test data to use in benchmark model.
8. Input this augmented data to ResNet-50 model and process them than take the transfer values and save them.
9. Since I have to remove last softmax layer of ResNet-50 model and I have 10 features to classify my images with, I will construct new neural network formed from 1 fully connected layer and a softmax layer.
10. Test my test set with this new model and compute test accuracy.
11. Compare accuracy results of two models.

References

AAA Foundation Website: https://www.aaafoundation.org/distracted-driving

State Farm Kaggle Competition Data: https://www.kaggle.com/c/state-farm-distracted-driver-detection/data