Chaofu Zhang

Individual Report

Personal Code: 20%

Introduction

The goal of this project is to anticipate how a vehicle would do autonomous driving based on the analysis of the surroundings (position of nearby automobiles). We choose this problem because we want to explore the truth of autonomous driving and find out if there is space to improve self-driving to make this technique widespread. We have discussed the model that we are interested in. To get proper data input for our model we need fu

Individual Work

After Qianhui's EDA part, I have continued to analysis the car_model data and done the further visualization. Since if we use YOLO3(we are trying to use in the very beginning), the bounding box would be built based on the car_model detail. So these works have not been usd. Besides that, my major work for this project is to do the data pre-processing and argmentation. Since the original image size is large, to save the computation power, we need to resize them to a proper size to fit in our model input. I've found all vehicle points (draw based on x, y, z) gathered at the bottom half of the image, yet some of the center points are outside of the image boundary. To get rid of these issues, I cut the top half image off and just keep the bottom half as our input images. In addition, we also need to expand the image width to include outlier points. To fit the proper size that Yuan want use as input, I resized the pictures to 2048*640 and do the standardization to keep pictures all in the same size. While resizing the image, the pose information would change if the weight-height ratio is changed, therefore, to avoid this happen, I kept the ratio of width and height of the image unchanged. The original ratio is 1:2.5,

while our resized image has a ratio of 1:3.2, so we applied two paddings to widen the images. The padding size for the left part is 1/8 of the images and for the right part is to fill up to 2048. Furthermore, I randomly flipped 10% of the images vertically as the augmentation, and divide the image RGB values by 255 to standardize the input. Then, the data was given to the Model Training part.

Results

After the 10 epochs training, the training loss for our model is more than 20, whichis really high. The main reason for the high loss comes from the BCE(Binary Cross Entropy) of the center point classification. Without BCE, the loss of the regression part is only about 0.4. The first figure below shows the heatmap with the prediction targets, it looks good but not very clearly. After doing the threshold operation, each prediction dot can be discovered distinctly in the second figure. By ensuring the position of each car in one image, the prediction of autonomous driving can be done by avoiding existing cars in sight.

Summary and conclusions

In conclusion, we implemented the EfficientDet as our network and performed prediction on the 6-DoF pose from camera captured pictures. We borrowed the idea of predicting center points heatmap and regressing on the center point related features from the CenterNet, and we combined the predictions from multiple scaled feature maps similar to U-Net, as we upsampled the outputs to the same size, stacked them together, and further fed the stacked outputs to a final output layer to predict the center point heatmap and pose features.