Final Project Individual Report

Yuan Le - Group 2

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. In this report, we will introduce our dataset, talk about how we pre-processing our data and choose parameters, discuss our deep learning network and training algorithm, and do the analysis based on the results.

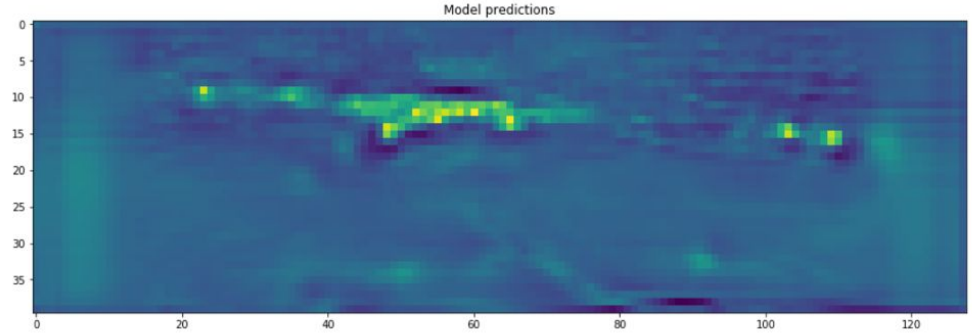
Individual works

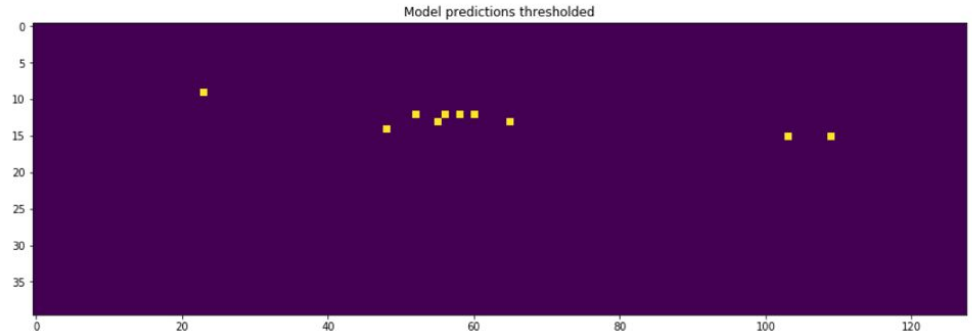
My major work in this project is to implement the BiFPN modules based on the paper EfficientDet. I also designed and implemented the prediction heads modules. Instead of using anchors as in the EfficientDet, I used an approach similar to the CenterNet, which predicts the heatmap of center points for each object, and regresses the pose features for each center point.

Results

After the 10 epochs training, the training loss for our model is more than 20, which is 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.

Due to the time and computational power limitations, our performance is not comparable to others in the Kaggle leaderboard. In our future work, we may further tune the network and modify the network to improve performance. We may also consider advanced data augmentation methods or including dropout layers in our networks to prevent overfitting.

References

Dataset retrieved from: https://www.kaggle.com/c/pku-autonomous-driving/overview

Tan, M., Pang, R., & Le, Q. V. (2019). EfficientDet: Scalable and Efficient Object Detection. *arXiv preprint arXiv:1911.09070*.

Mingxing Tan, Quoc V. Le.(2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *arXiv preprint arXiv:1905.11946.*

Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie.(2019). Feature Pyramid Networks for Object Detection.*arXiv:1612.03144*

Xingyi Zhou, Dequan Wang, Philipp Krähenbühl.(2019) Objects as Points.

arXiv:1904.07850