3D OBJECT DETECTION FOR SELF-DRIVING CARS

Burak Çolakoğlu – Furkan Han Keçeli
Advisor Dr. Öğr. Gör. Faris Serdar Taşel
Co-Advisor Dr. Öğr. Üyesi Roya Choupani

Çankaya University, Department of Computer Engineering

Abstract

Self-driving cars presents a rare opportunity to improve the quality of life in many of our communities with the combination perceive systems such as lidars, cameras and sensors technology. 3D object detection is an import task for autonomous vehicles. In this project, we combine the LiDAR and RGB data, to build an end-to-end system for 3D object detection. We research and implement the advanced models, including U-Net, Point RCNN,F-ConvNet and SVGA-Net to solve this problem. The intersection over union (IoU) is used for our model evaluation. We've finished the training and testing of U-Net based method. The Point RCNN based method is implemented.

Keywords: Object Detection, Deep Learning, Self-Driving Cars, LiDAR, RCNN

Introduction

3D models are becoming widely available and easier to capture, making available 3D information crucial for progress in object classification. The approach we are taking to address this problem is by checking various SOTA models that are available for similar kind of problems to this particular Lyft's dataset and output type that not only needs the bounding boxes and object class but also yaw, as this real time output has to be feed to self driving control system. Further, due to training computation constraints we are looking for pre-trained models that we can fine tune.

Solution

1) U-net-Based Method: UNet is proposed for semantic segmentation. In this method, after we project 3D LiDAR data into 2D Bird Eye View data, we view bounding box as the object ground truth that we'd like to segment as Figure 4 shown. Then we use a transformation matrix to transform predicted boxes back to world space.

2) PointRCNN Based Method : Only 3D LiDAR point cloud data will be used. We'll use a stage-1 sub-network to generate a small number of high-quality. Then we use a stage-2 sub-network to transform the pooled points of each proposal to canonical coordinates to learn better local spatial features.

Figure 1 – 2D Output

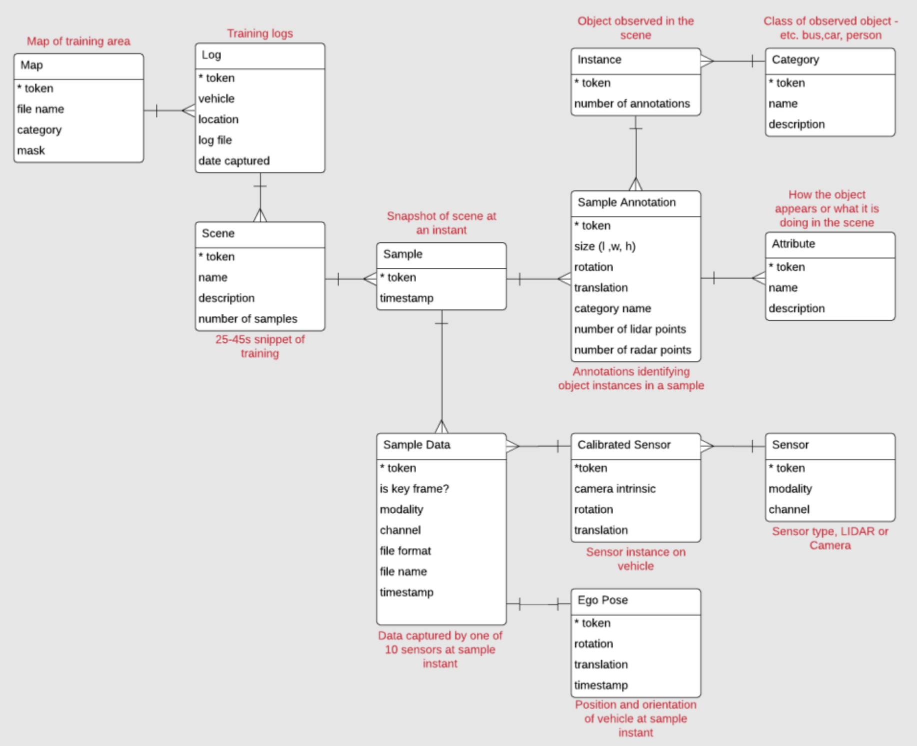

Figure 2 – Dataset Structure

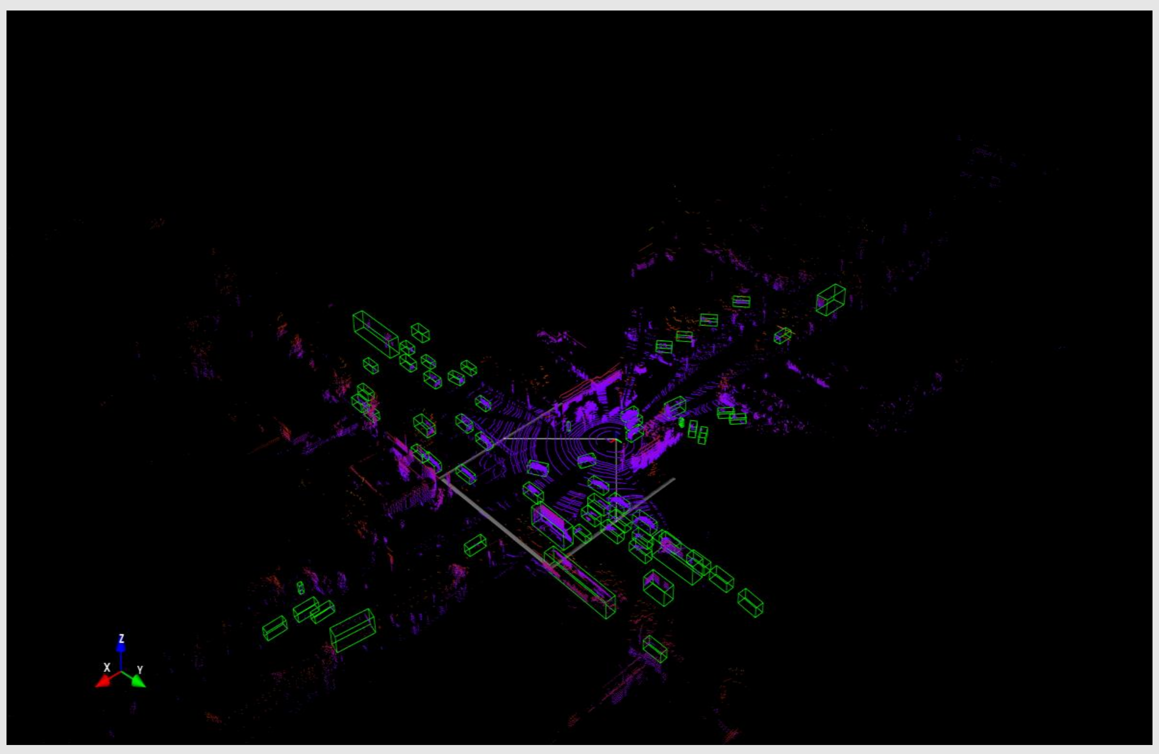

Figure 3 – LiDAR Point Cloud

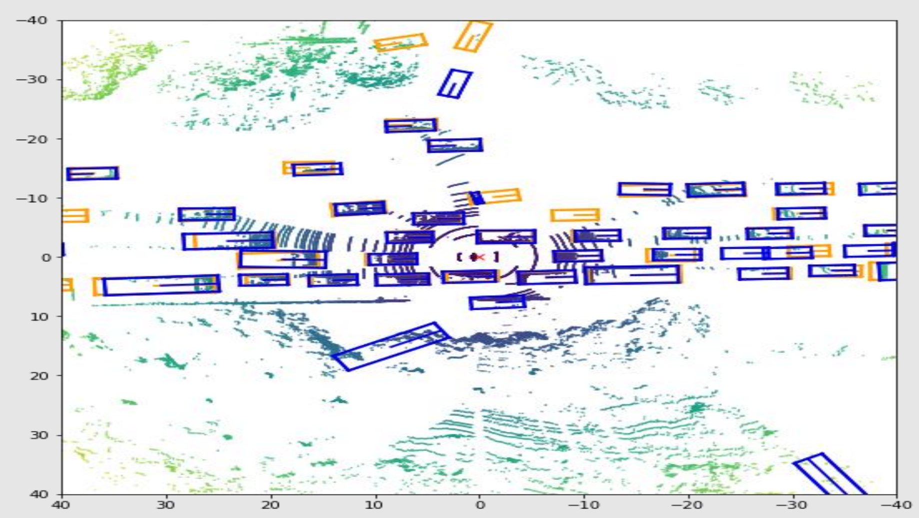

Figure 4 – 3D BEV Result

Results & Conclusion

So far we implemented the training and testing of U-Net based method. We have first transformed lyft dataset from Nuscene format into Bird-eye-view format, then we could have 2D ground truth for image segmentation. We segment objects from background with a global background pixel threshold which is a adjustable parameter. The results and its ground truth are shown in Figure 5. Then we use a transformation matrix to transform predicted bounding boxes into voxel space.

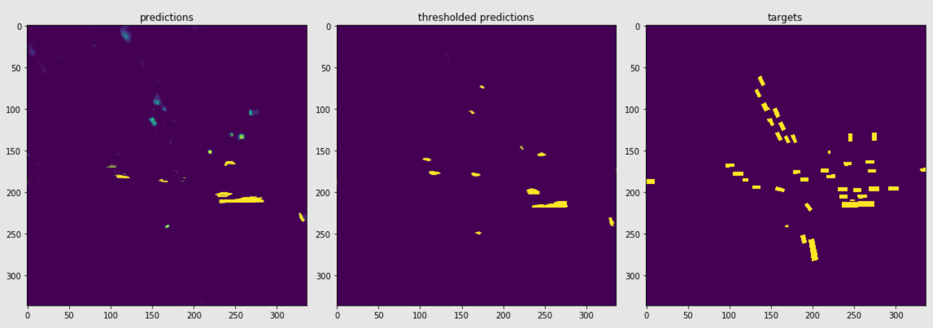

Figure 5 – U-Net Output

Acknowledgement

Special thanks to our Advisors Faris Serdar Taşel and Roya Choupani for their insightful comments and suggestions.
We are also grateful to Efe Çiftçi for his great supports.

Figure 6 – 2D BEV/Segments

Burak Çolakoğlu   Furkan Han Keçeli

Figure 7 – Team Members