Abstract

We are working on the end-to-end fashion 3D object detection project based on the KITTI dataset. The overall idea is to leverage the profoundly developed Deep Neural Network to extract features within point clouds which LiDAR collected and use the information to detect the interested objects (car, pedestrian, cyclist, etc.) on the road in real time. This project can benefit the mobile robots or Autonomous Driving technology at the visualization level. In the future, the accurate and confidential road condition information will be provided to the autonomous driving system, improve the guidance system.

Recent robotic technology has witnessed the rapidly developed deep learning method within many research field. Within which, computer vision science has gained the benefit not only from traditional machine learning method, but also the fast increasing deep neural networks. Began from AlexNet, the neural network has showed the strong power of extracting features, doing real time inference, solving complex problems in many fields. Many image datasets and benchmarks sprung up, provide us with plenty of opportunities to work on different interesting problems.

The KITTI benchmark is designed for the researchers who are interested in the autonomous driving research. We can see from the KITTI benchmark, that the 2D object detection has reached a relative optimal level, where 3D object detection still got some bottlenecks to deal with. On the other hand, the 3D object detection problem can be largely affected by the scene conditions such as light, object orientation, etc. These are all the problems that need to be handled and what we are interested in. Therefore, we decided to tackle with the 3D object detection problem using the DNN, trying to find a better solution for the problem.

We first studied recent pioneering papers on this topic, and we found that we can find a way to deal with the point clouds collected from LiDAR. RGB or RGB-D data are fragile to the environment which cannot provide us with reliable information. By contrast, the point clouds gather detected points in the 3D spaces, excluding the impact from the environment, which are more robust and stable.

Then we designed the neural network based on the VoxelNet. The main idea is to partition the 3D space into several voxels, and sample point data based on some random selection rules. With grouping and selection, we can play with the compressed data more efficiently. After the augmentation manipulation, we trained our neural network to do some unsupervised clustering, which can grasp more detailed features, and predict for “neglect” properties. We then let the convolutional neural network to do their part, get the classification map and the regression map.

The novel point of our neural network is that it can capture several “might be neglect” features quickly comparing to supervised learning method. The overall training process would not be very hard to tune the parameters, and the inference time would be quick enough for the real time testing.

Now we’re going to train our network, and we really hope that the project will benefit the autonomous driving research field.