# 3D Object detection Pipeline Report

This is the report of the project "3D Object detection Pipeline".

## 1. Pipeline

SVD Based Ground Segmentation -> DBSCAN Foreground Clustering -> PointNet++ based Foreground Objects Classification -> Target Objects BBox Fitting (in progress)

## 2. Dataset Building

All the training and test data come from KITTI 3D Object dataset. In each laser scan, the provided label and calibration files are used to crop out the objects of the target class in this detection task - Car, Cyclist and Pedestrian. The other objects on the foreground are clustered using DBSCAN and saved with label DontCare. Original data distribution:
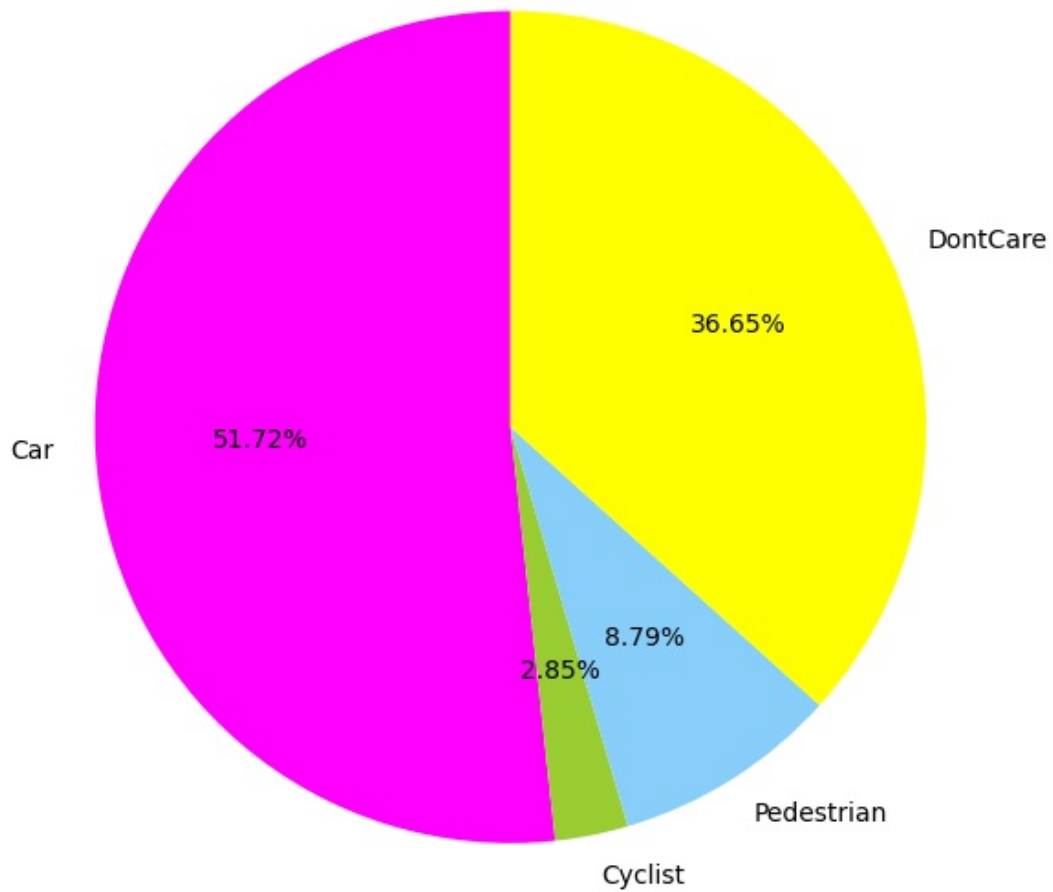
To alleviate the problem of unbalanced data, data augmentation based on random rotation, random jitter, random dropout, radom scale and combination of them are performed on class Cyclist and Pedestrian. Augmented data distribution:
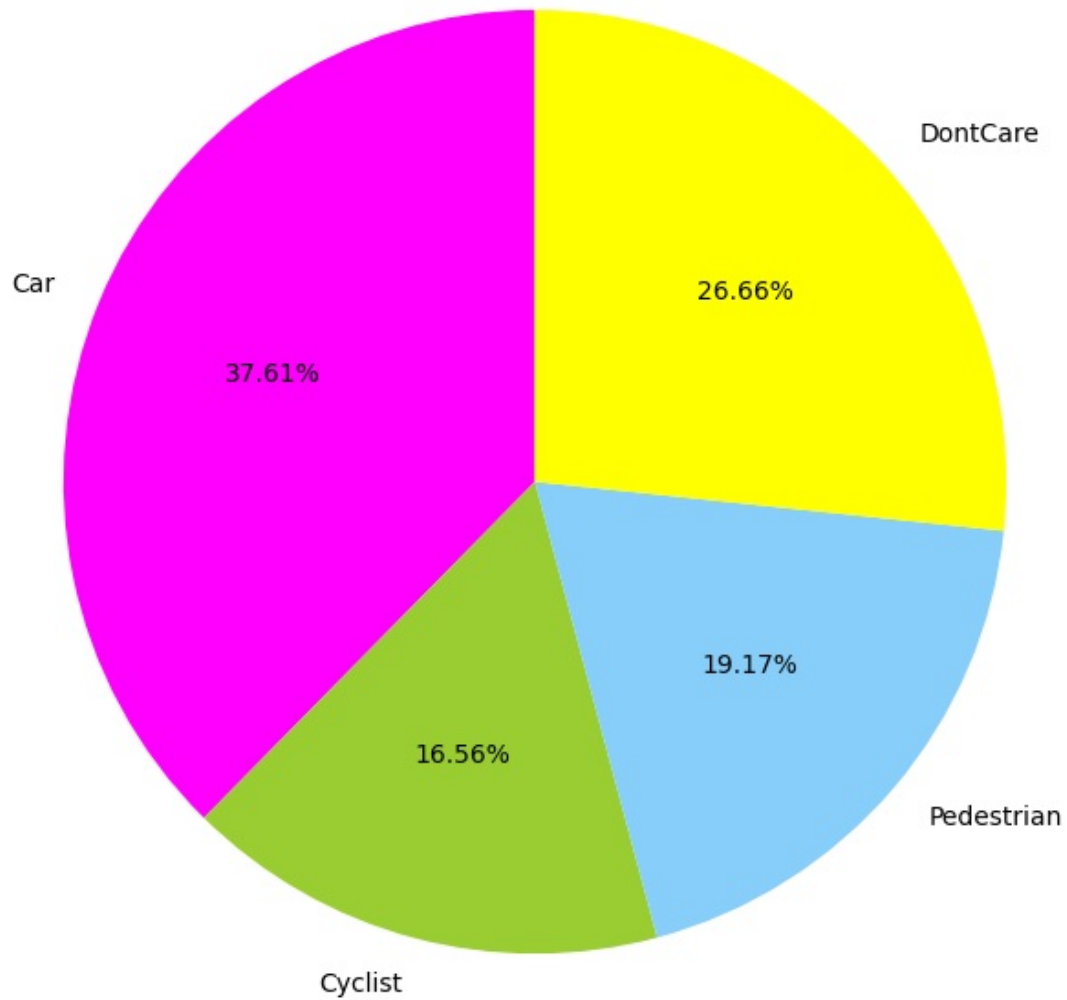
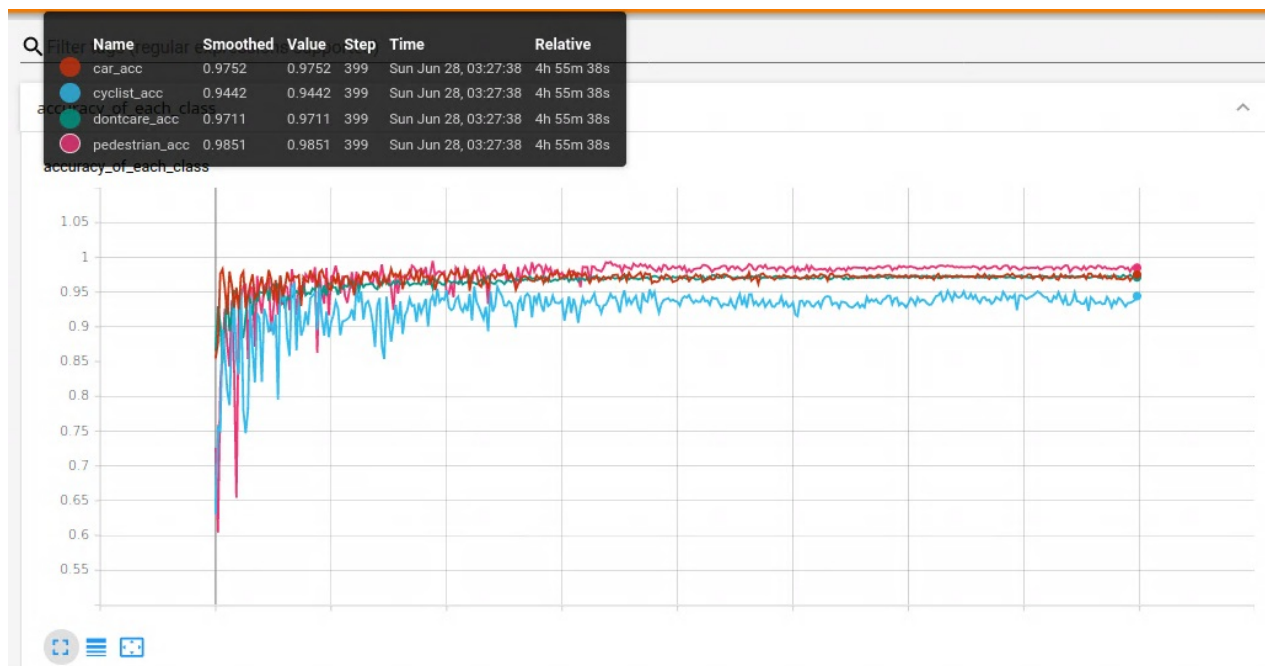| | Car | Cyclist | Pedestrian | DontCare |
|---|---|---|---|---|
| Object Number (before) | 23427 | 1289 | 3980 | 16603 |
| Object Number (after) | 23427 | 10312 | 11940 | 16603 |

Since the number of points in each object point cloud are quite different. Some objects far away from lidar has very few points. A good choice of sampling size on the input point cloud of the network is important. Histogram of cloud size of each class:
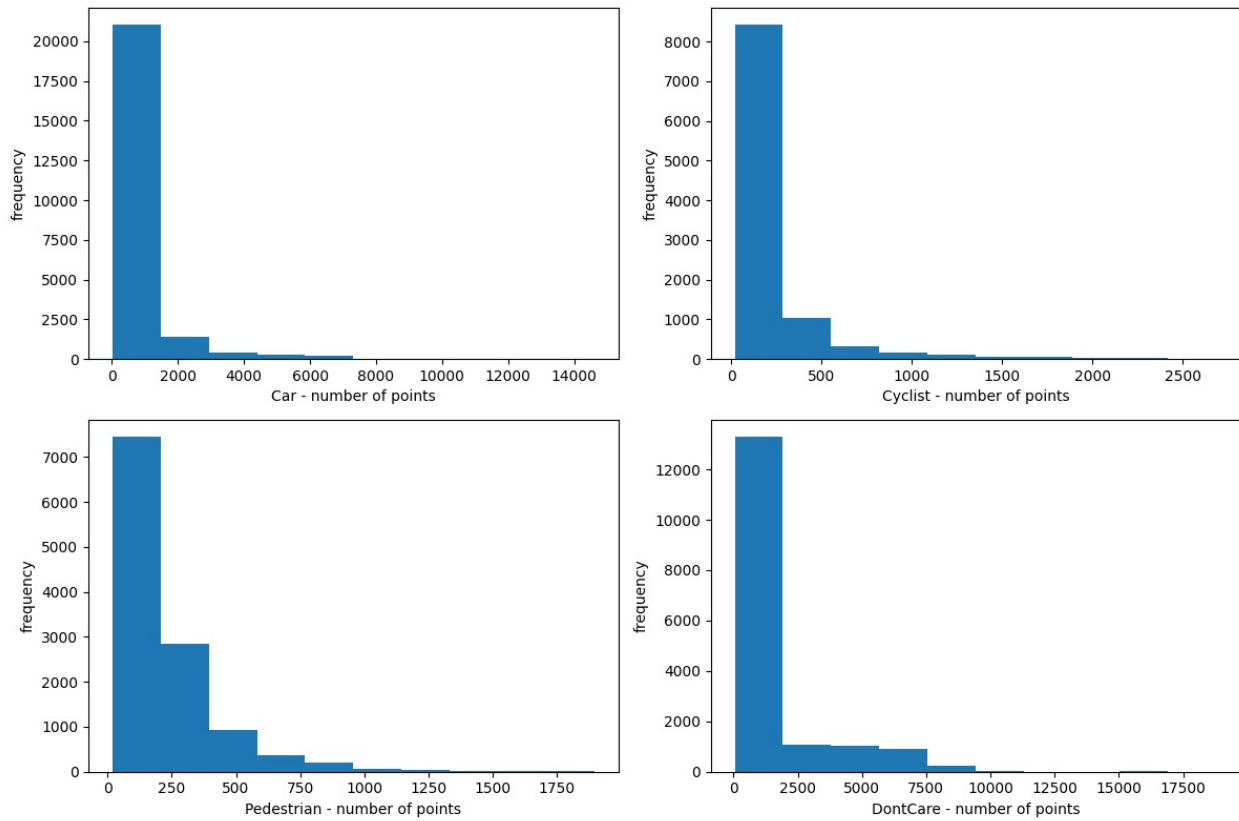
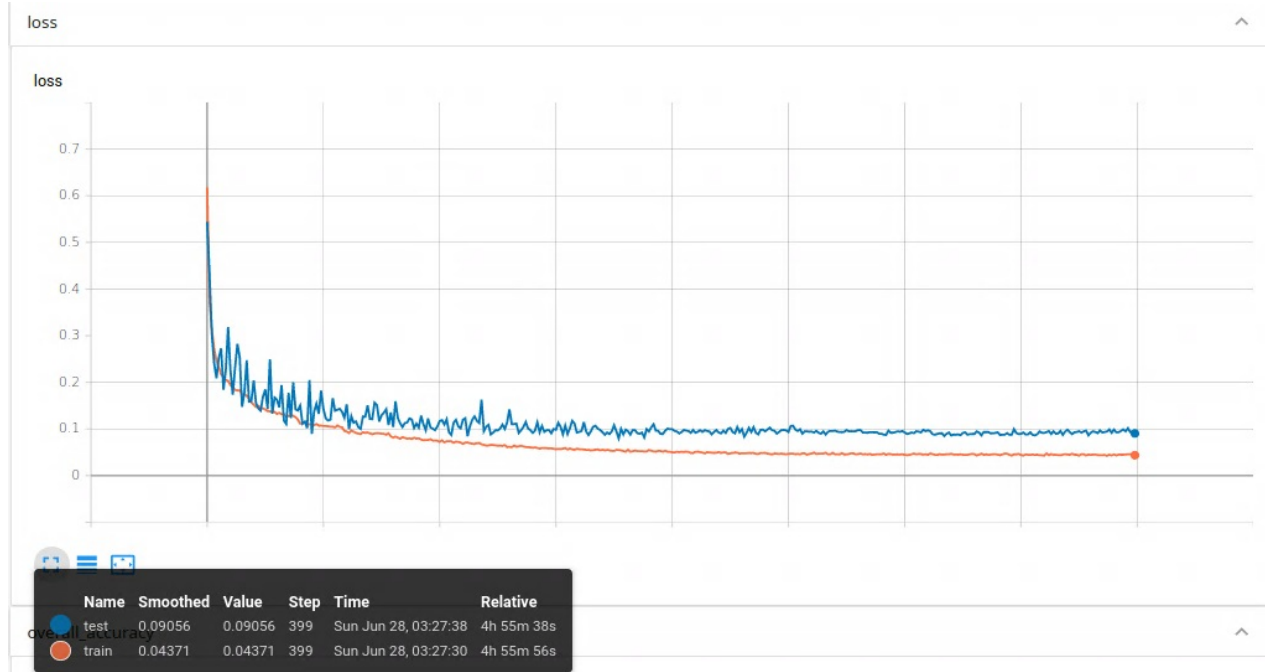To cover the most of the objects, 256 is chosen as the sampling size. And the sampling strategy is that, if the input cloud has more than 256 points, farthest point sampling is performed, otherwise the raw cloud is retained and (256 - cloud_size) points are randomly selected with replacement from the cloud to make up of a point cloud with size 256.

## 3. Training Process

Accuracy of each class

| Name | Smoothed | Value | Step | Time | Relative |
|------|----------|-------|------|------|----------|
| car_acc | 0.9752 | 0.9752 | 399 | Sun Jun 28, 03:27:38 | 4h 55m 38s |
| cyclist_acc | 0.9442 | 0.9442 | 399 | Sun Jun 28, 03:27:38 | 4h 55m 38s |
| dontcare_acc | 0.9711 | 0.9711 | 399 | Sun Jun 28, 03:27:38 | 4h 55m 38s |
| pedestrian_acc | 0.9851 | 0.9851 | 399 | Sun Jun 28, 03:27:38 | 4h 55m 38s |

Training and test loss



| Name | Smoothed | Value | Step | Time | Relative |
|------|----------|-------|------|------|----------|
| test | 0.09056 | 0.09056 | 399 | Sun Jun 28, 03:27:38 | 4h 55m 38s |
| train | 0.04371 | 0.04371 | 399 | Sun Jun 28, 03:27:30 | 4h 55m 56s |

Training and test accuracy

| Name | Smoothed | Value | Step | Time | Relative |
|------|----------|-------|------|------|----------|
| test | 0.9706 | 0.9706 | 399 | Sun Jun 28, 03:27:38 | 4h 55m 38s |
| train | 0.9765 | 0.9765 | 399 | Sun Jun 28, 03:27:30 | 4h 55m 56s |
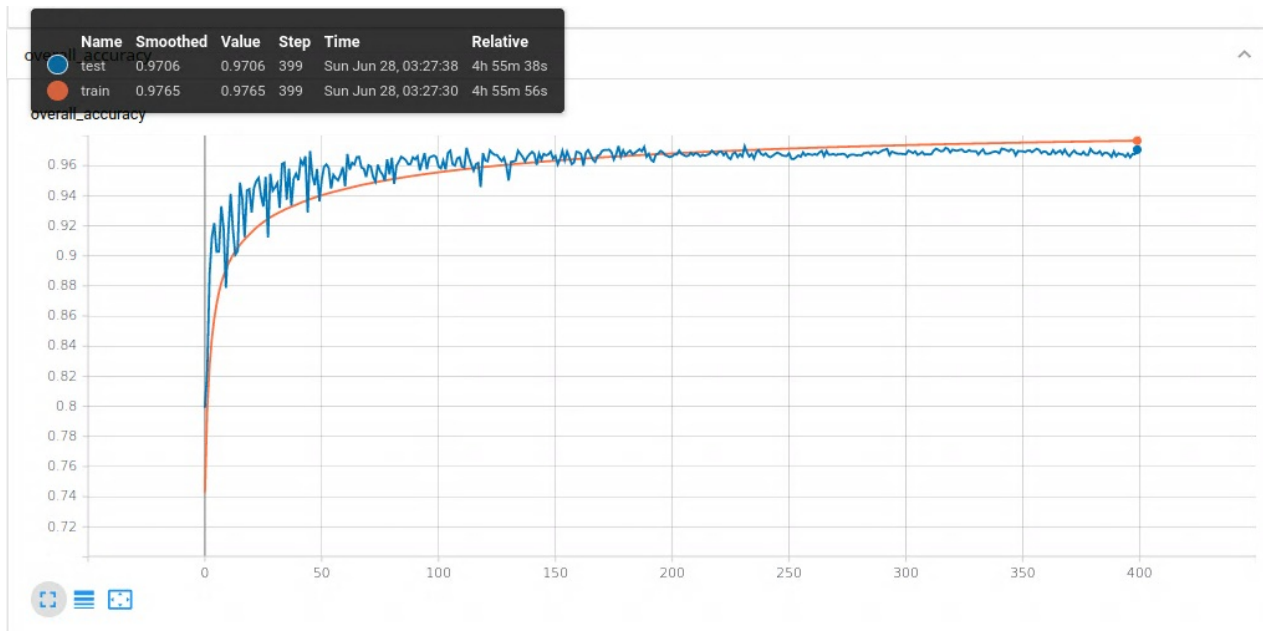
overall_accuracy

The best model is saved at epoch 178. The evaluation results:

Overall accuracy:

| Test Instance Accuracy | Class Accuracy |
|---|---|
| 0.974043 | 0.973101 |

Accuracy of each class:

| Car | Cyclist | Pedestrian | DontCare |
|---|---|---|---|
| 0.976624 | 0.953772 | 0.989494 | 0.972515 |

## 4. Some results (in progress)

The Bboxs used in the results are axis aligned bboxes, which can't indicate the real orientation of the detected object. Object oriented bbox need to be fitted for those objects.