3D Object Detection

1 Introduction

The goal of this project is to perform object detection using Lidar data from Waymo open dataset.

1.1 Task 1: ID_S1_EX1

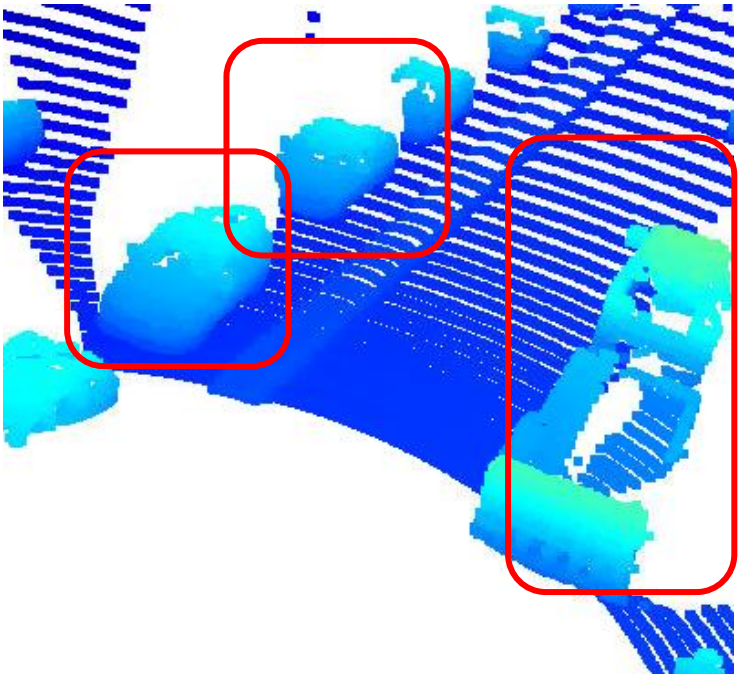
In the Waymo Open dataset, lidar data is stored as a range image. Therefore, this task is about extracting two of the data channels within the range image, which are "range" and "intensity", and convert the floating-point data to an 8-bit integer value range. Sample output is shown below.

1.2 Task 1: ID_S1_EX2

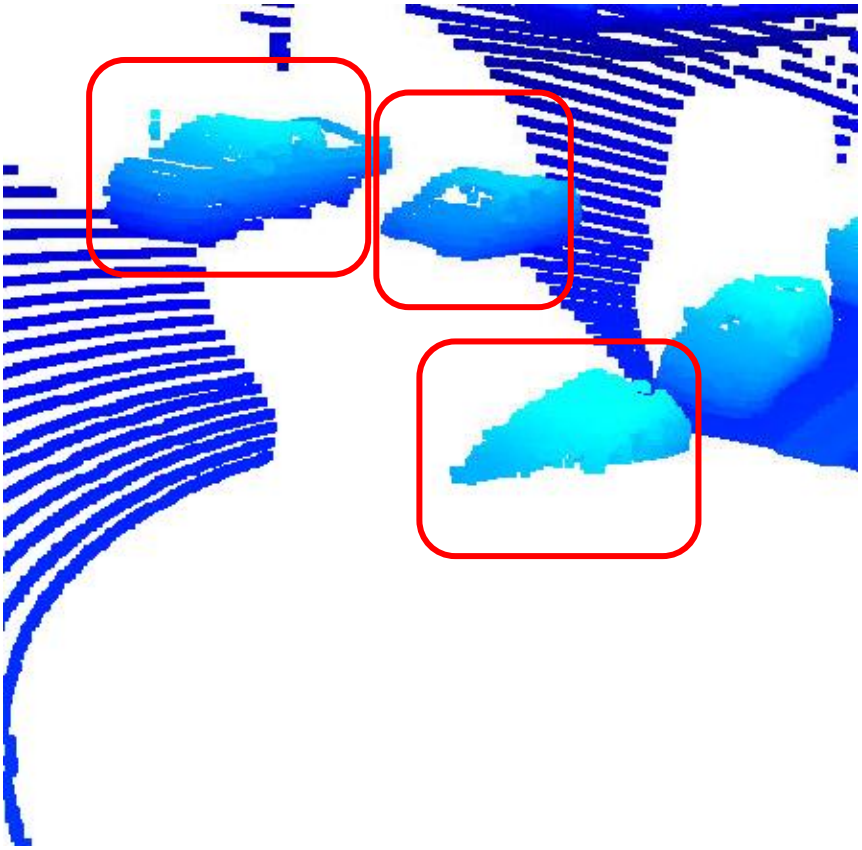
The goal of this task is to use the Open3D library to display the lidar point-cloud in a 3d viewer in order to develop a feel for the nature of lidar point-clouds.

The scenario here is that the vehicle is moving across an intersection. Thus, vehicle can be observed in multiple directions with varying degrees of visibility.

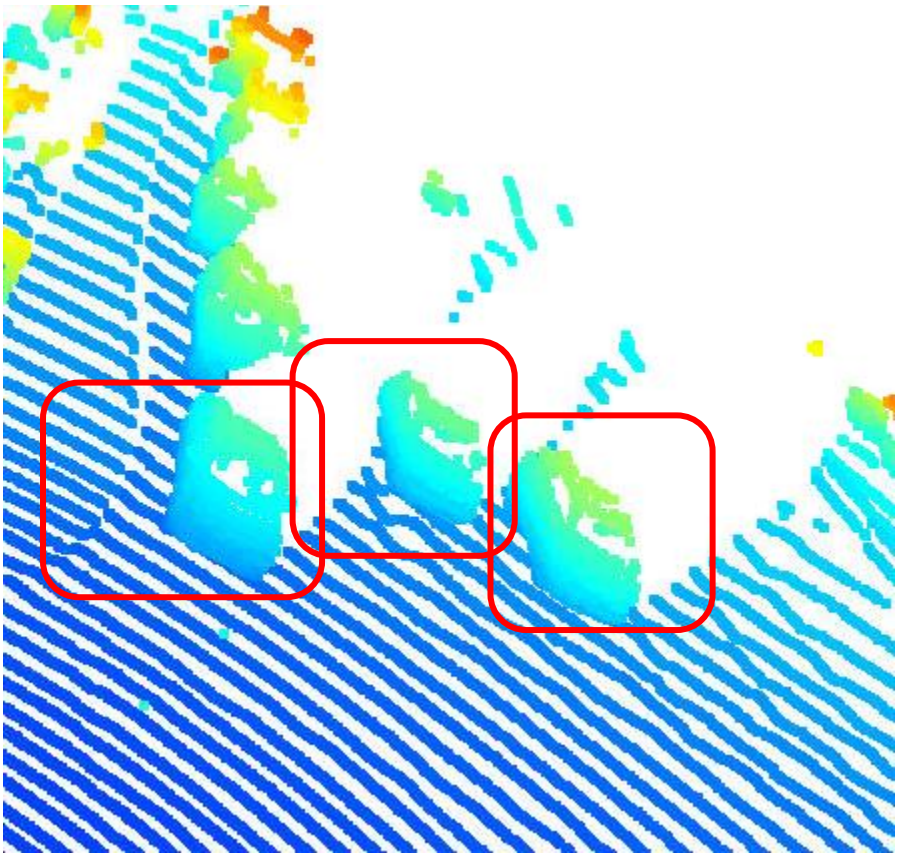
Three vehicles with high visibility



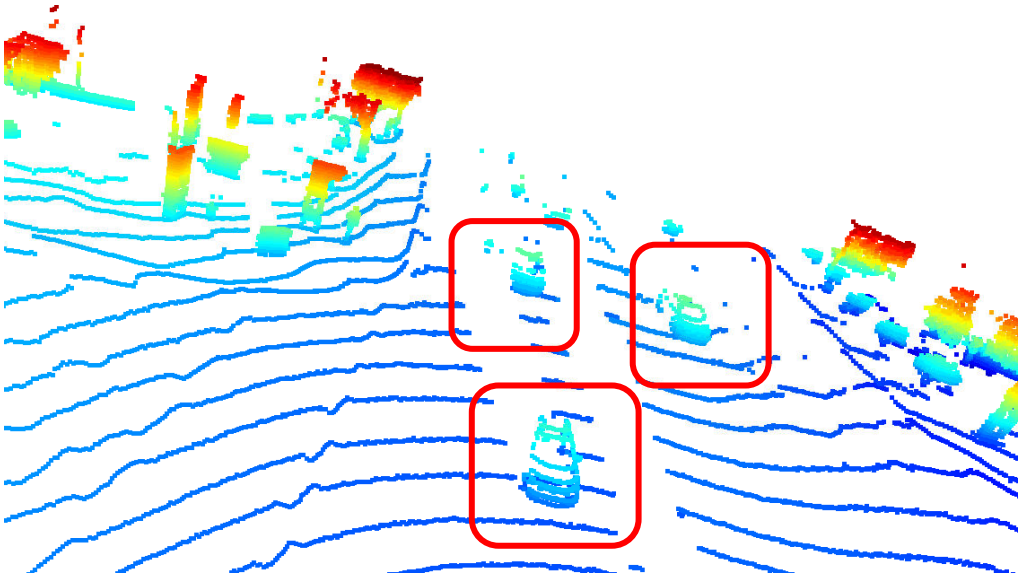
Two medium levels of visibility of vehicles created by occlusion of sensor due to mounting and vehicles.



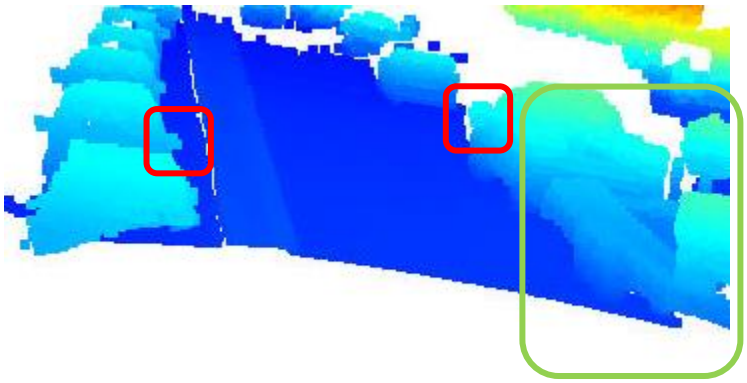
Three examples of good visibility on the other side of the intersection.



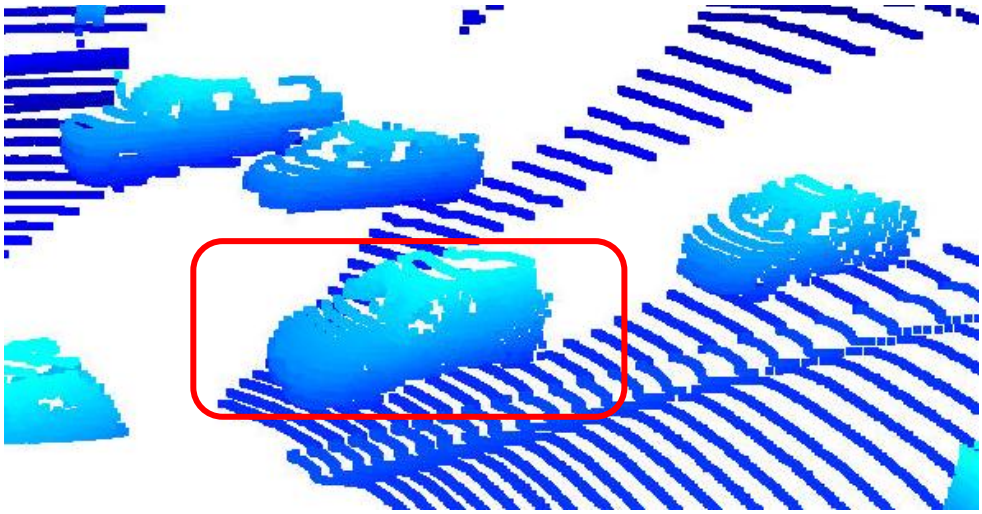
Three examples of barely visible vehicles behind the lidar.



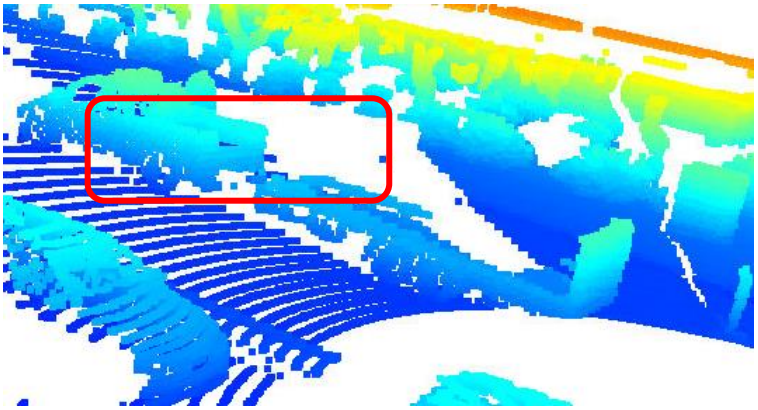
Features Identified on vehicles – Side View mirror, Trailer attached to pick up truck. So we can identify type of vehicle as well in this data.



From this image we can also get a sense of direction of direction of travel of the vehicle as the shape of the vehicle is identified by the front bumper and windshield

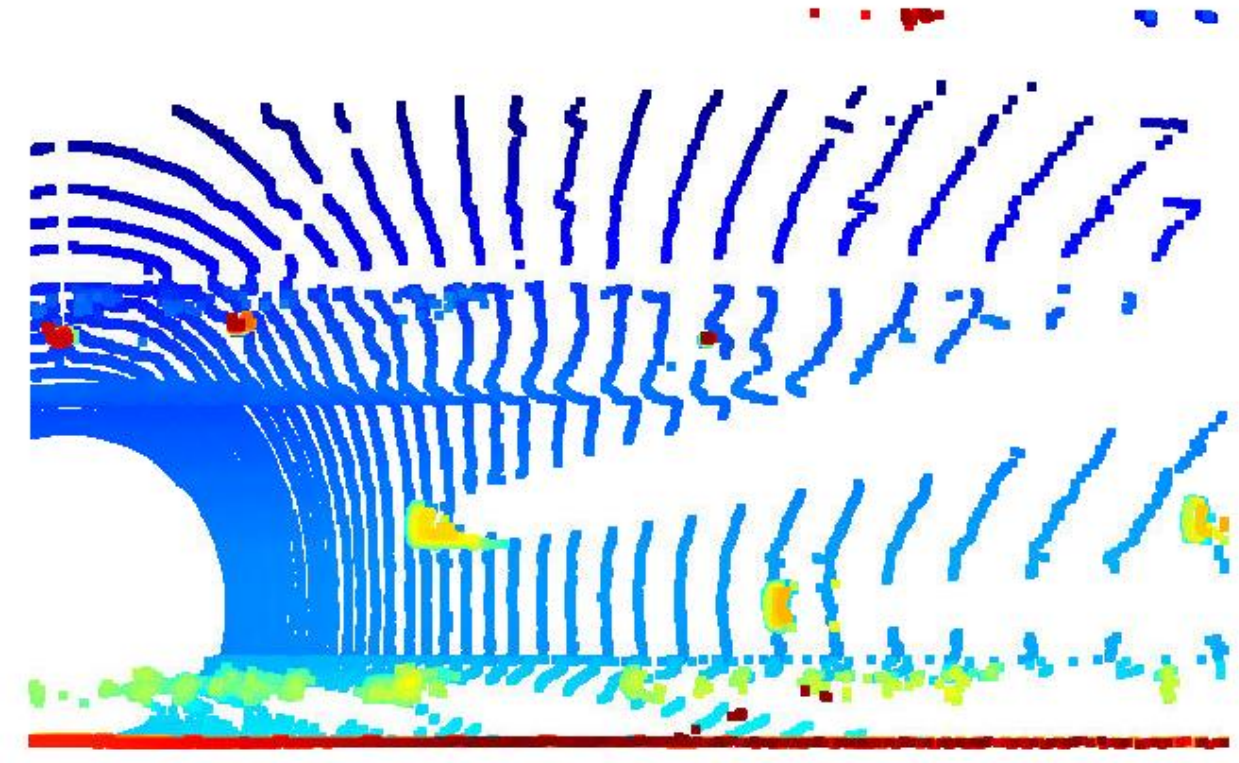


From this angle we can see the start and end of the pickup truck bed and also the rear wheel can be spotted.



1.3 Task 2: ID_S2_EX1

The goal of this task is to perform the first step in creating a birds-eye view (BEV) perspective of the lidar point-cloud.



1.4 Task 2: ID_S2_EX2

The goal of this task is to fill the "intensity" channel of the BEV map with data from the point-cloud.

1.5 Task 2: ID_S2_EX3

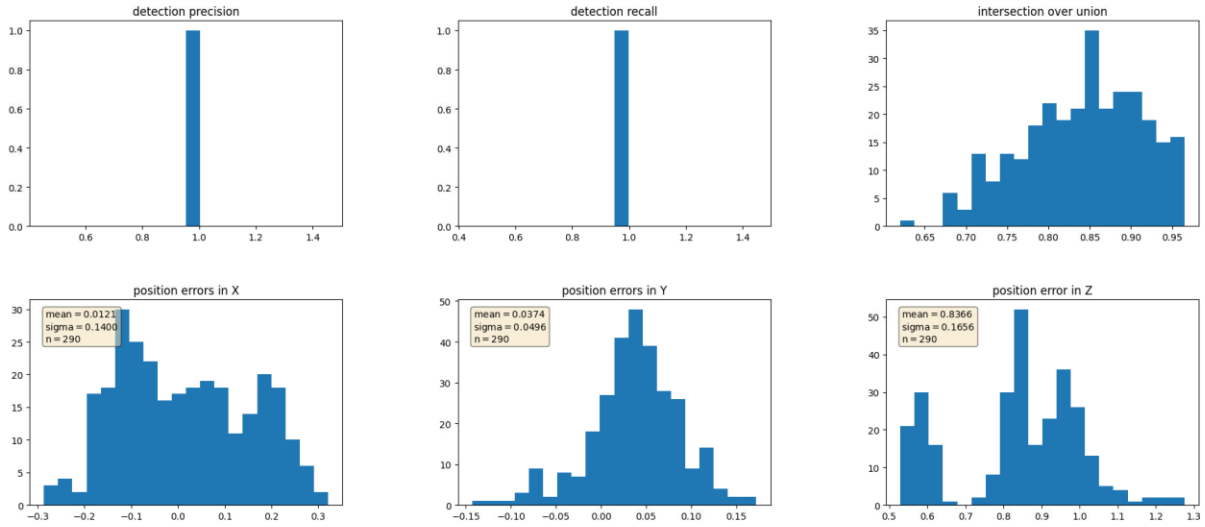
The goal of this task is to fill the "height" channel of the BEV map with data from the point-cloud.

1.6 Task 3: ID_S3_EX2

As the model input is a three-channel BEV map, the detected objects will be returned with coordinates and properties in the BEV coordinate space. Thus, before the detections can move along in the processing pipeline, they need to be converted into metric coordinates in vehicle space. This task is about performing this conversion such that all detections have the format [1, x, y, z, h, w, l, yaw], where 1 denotes the class id for the object type vehicle.

1.7 Task 4: ID_S4_EX4

After processing all the frames of a sequence, the performance of the object detection algorithm shall now be evaluated. To do so in a meaningful way, the two standard measures "precision" and "recall" will be used, which are based on the accumulated number of positives and negatives from all frames.



The Precision value is 0.9539473684210527 and the recall value is 0.9477124183006536