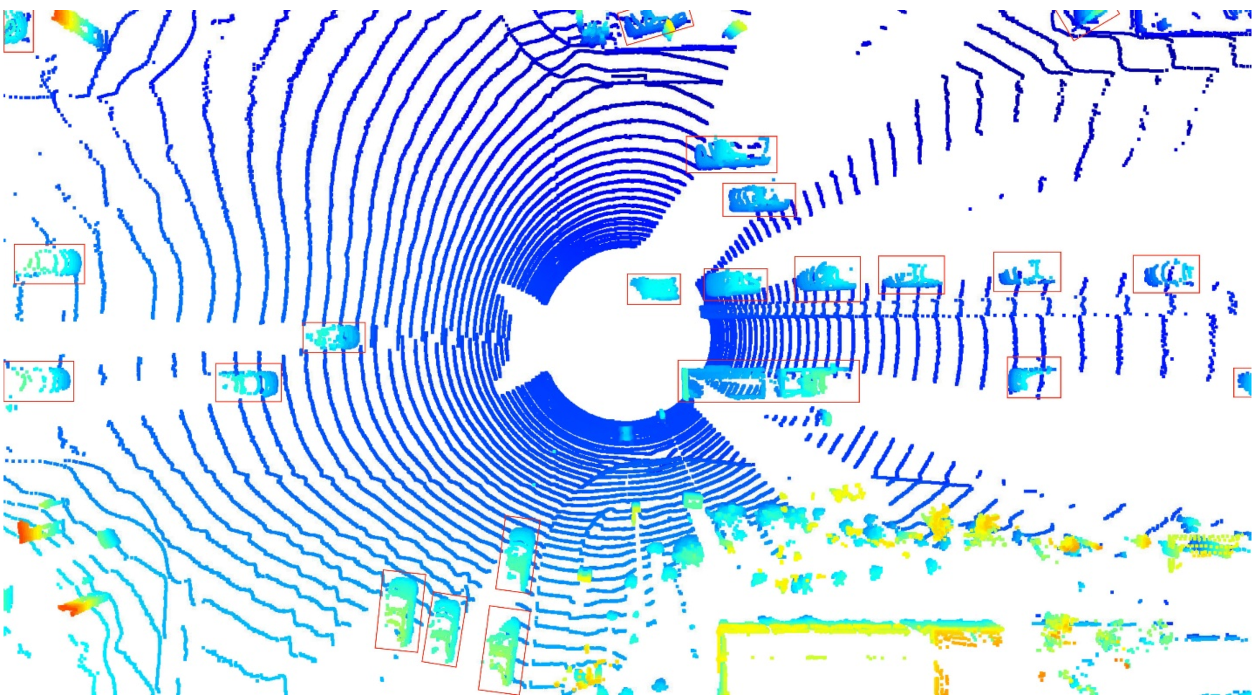
Midterm Project Writeup Sensor Fusion

Note: To view the images when executing loop_over_dataset.py, set the *display_image* variable to True in the bev_from_pcl function located in the objdet_pcl.py inside the student folder. It is set to False by default.

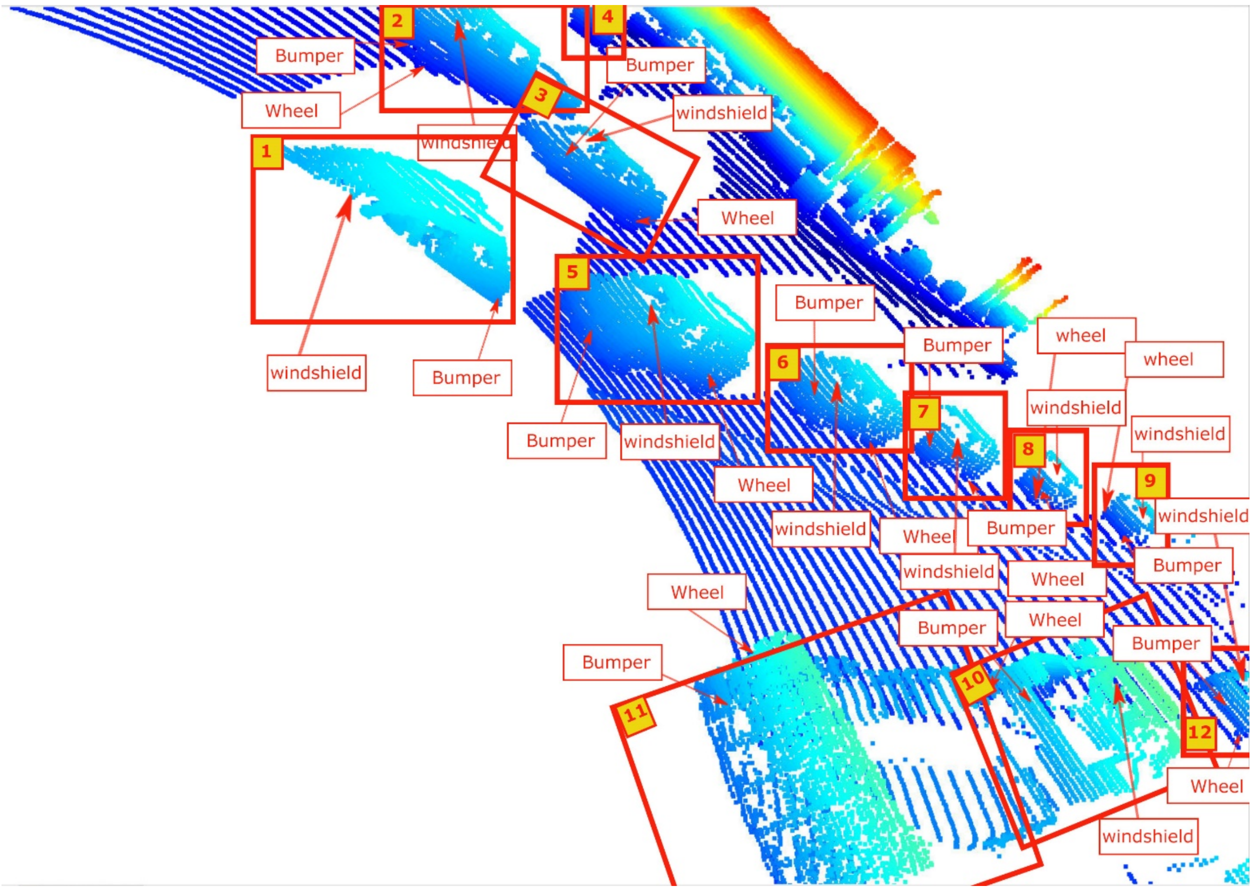
When viewing images, pressing the spacebar closes the current image and automatically opens up the next one until the all images are seen. *To exit early, press the right arrow key on your keyboard*.

The picture below comes from a point cloud obtained from the WAYMO dataset training_segment-10963653239323173269_1924_000_1944_000_with_camera_labels.tfrecord. It was taken by a car equipped with a Lidar going through what seems to be an intersection. Multiple vehicles can be observed in this picture with their bodies either showing fully or partially. Each observable vehicle is surrounded by a red bounding box.

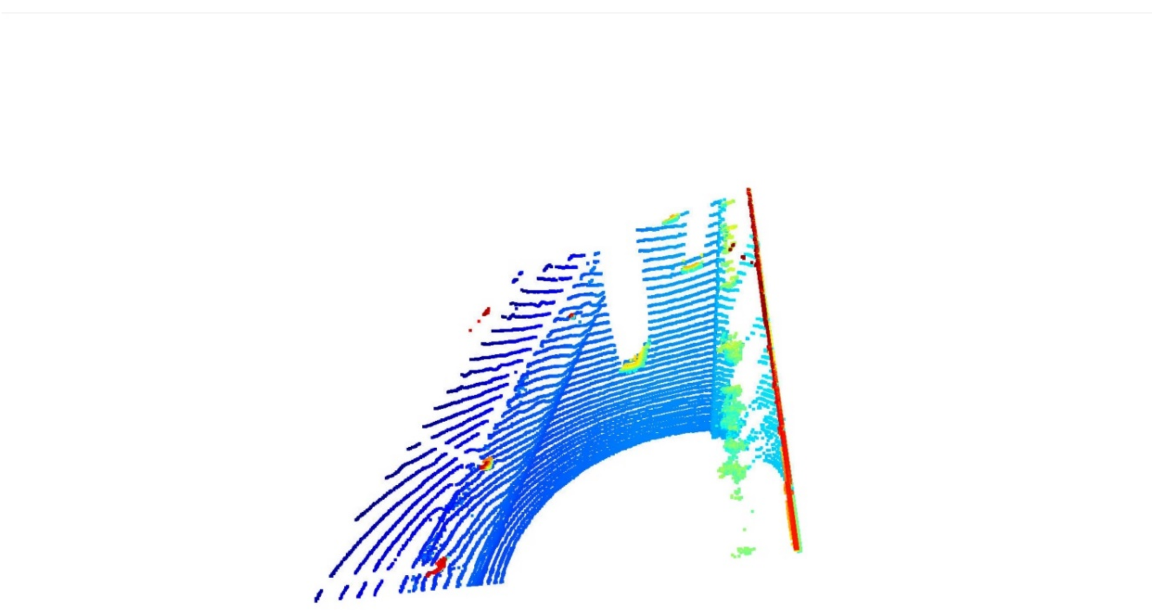
As can be seen in the same picture, the vehicles observed in this picture consistently show blank spots where the glass portions of those vehicles are located. Additionally, not all vehicles have the same number of points. It appears that the closer a vehicle is to the lidar car, the more points it will have, hence the more details it will have.



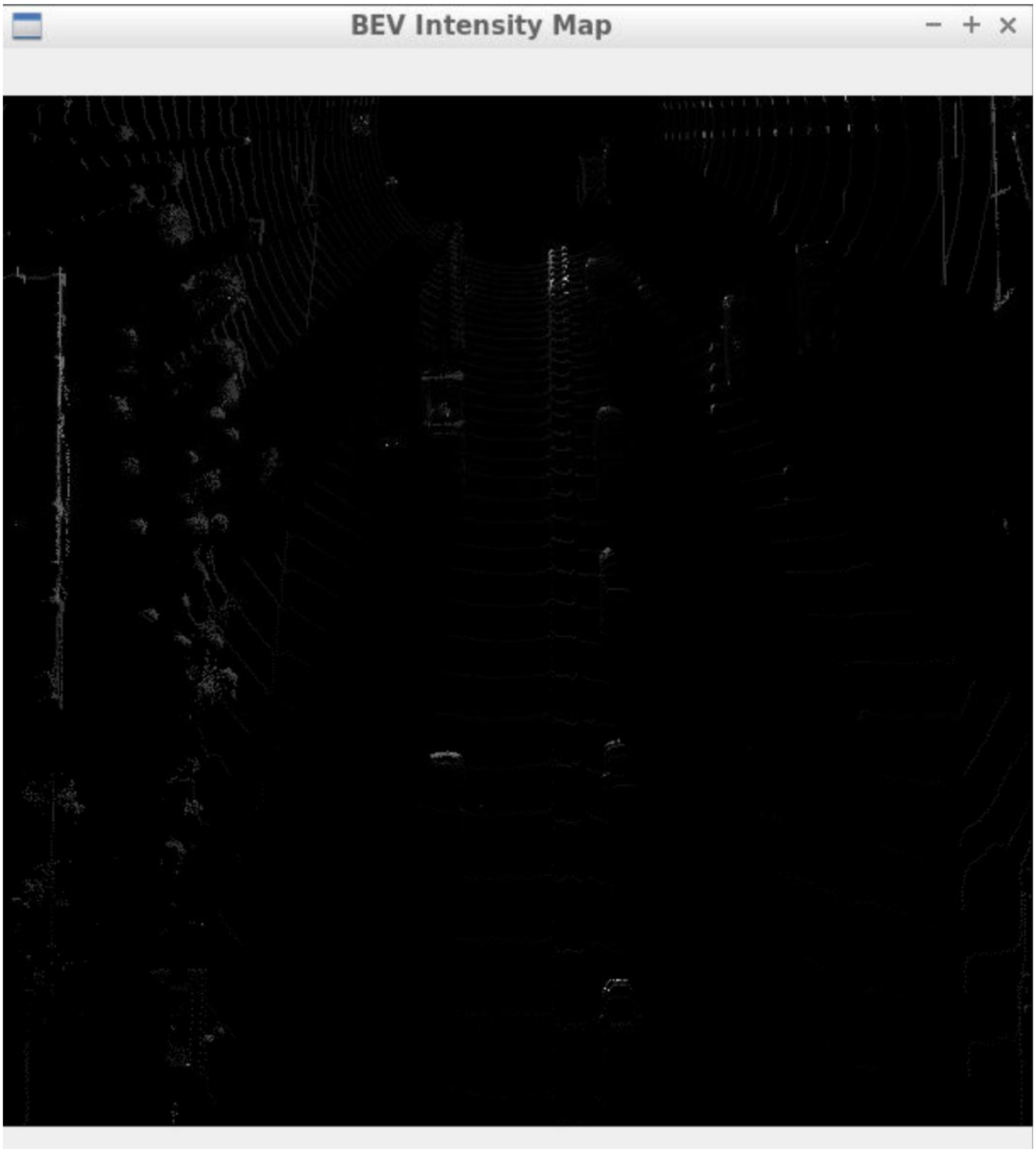
The point cloud also allows us to make a distinction among types of vehicles. For example, we can clearly see that there is a trailer being towed by a pickup truck. Continuing with our observations, we have noticed that bumpers are clearly delimited and the tires can also be seen on all vehicles. This leads us to believe that rubber has great lidar visibility as seen in the picture below.



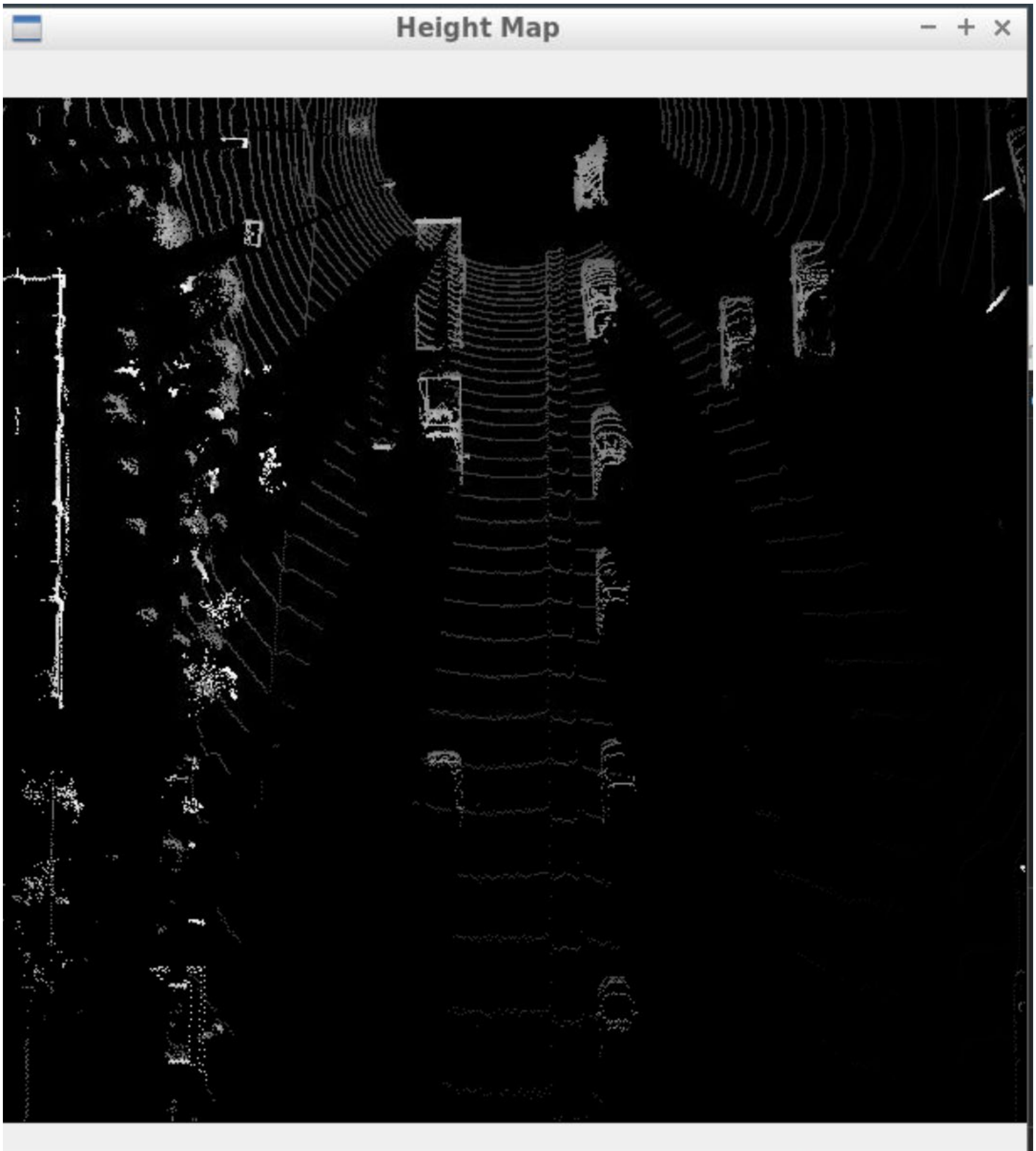
Using data from the point cloud map, we then created a Bird's Eye View (BEV) of the point cloud data based on (x,y) coordinates in sensor space translated into BEV coordinates.



The creation of the BEV map is very important. With the BEV map, we can now identify all points with the same (x,y) coordinates within the BEV map and assign the intensity value of the top-most lidar point to the respective BEV pixel. This is known as an intensity map. The resulting intensity image is normalized using percentiles, in order to make sure that the influence of outlier values (very bright and very dark regions) is sufficiently mitigated and objects of interest (e.g. vehicles) are clearly separated from the background.



We then apply the height information from the lidar channel to the BEV map thereby filling the height channel of the BEV map with data from the point cloud. This step literally pulls the vehicle out of the ground as can be seen in the picture below.



The last step was to calculate how accurate we were in our detections. To do that, we had to find pairings between ground-truth labels and detections, so that we could determine whether an object has been (a) missed (false negative), (b) successfully detected (true positive) or (c) has been falsely reported (false positive). We were able to achieve the following results: