Week 4 Report

Doğaç Güzelkokar, Niklas Bubeck

June 2022

1 Week 4 Goals

The goals of this weeks are to experiment more with the sceneflow. Introducing selective sceneflow for only dynamic classes. Removing ground classes, as well as getting first insights into self-supervised sceneflow schemes, that make use of the semantic kitti dataset rather than using RGB-D converted data.

2 SceneFlow

2.1 Experiment: Ground Removal

2.1.1 Methods

The idea behind ground removal was to get rid of a lot of unnecessary points and to emphasize the point resolution to the things classes. For this, we thought of two methods: (a) we use the ground truth label data and remove certain classes, (b) we filter the data based on a certain height. For the sake of simplicity and neglect adding noise with an additional filter decided to first go with method (a).

2.1.2 Results

We applied method (a) and removed the following classes: ("unlabeled", "outlier", "road", "parking", "sidewalk", "other-ground", "lane-marking"). The scaling paramters are: 0.01 for things classes and 1 for stuff classes. The resulting downsampled pointcloud with the sceneFlow already appied can be seen in 1

The image shows, that estimating scene flow is still hard for poles, and signs, independent of the voxelscale/sampling size. Disproving the theory that the classes are overall too underrepresented, and strengthening the theory of dataset adaption.

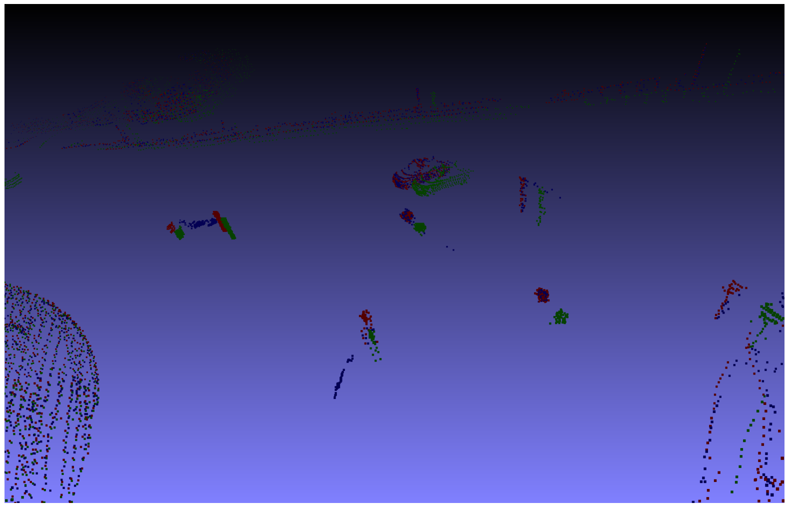

Figure 1: Sampled pointcloud with ground removal

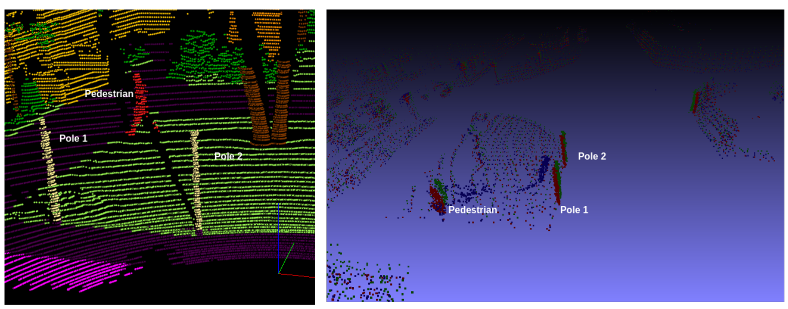

Figure 2: An example including a pedestrian.

2.2 Experiment: Evaluation on pedestrians

We searched for a good scenario where a pedestrian walks through the scanner. This was found in seq 05 scan 1-10.

From the image in 2 one can see, that also a pedestrian is not well estimated when applying the sceneFlow. It seems like the network searches for structures that represent certain geometries that are not given. Thus the network wants to combine the points of poles and the pedestrian.

3 Further Literature Review

This week, We focused on "Weakly Supervised Learning of Rigid 3D Scene Flow"[1] and "Just Go with the Flow: Self-Supervised Scene Flow Estimation"[2] papers. Most methods are reliant on Kitti Scene Flow for self supervision, which has point clouds generated by the images from a front camera of the vehicle and has no 360coverage. We think that, with adaptations such as retraining or further training on our dataset, and modifying certain preprocessing steps etc., we can adapt these works to our case.

4 Next Steps

- Setup and run the "Just Go with the Flow" model on our dataset.
- Setup " Rigid 3D Scene Flow" model, adapt it to our dataset by skipping certain preprocessing steps and run on our dataset.
- In case "Just Go with the Flow" does not perform well, retraining it on SemanticKITTI can be considered.
- We were able to do inferene with EfficientLPS[3], run EfficientLPS on Titan GPU to check if we can get better inferences.
- In the future, assuming we are able to get reasonably good inferences for single scans, and reasonably good scene flows, we will get the inferences on some sequence to have our first validation result.

References

[1] Z. Gojcic, O. Litany, A. Wieser, L. J. Guibas, and T. Birdal, "Weakly supervised learning of rigid 3d scene flow," 2021. [Online]. Available: https://arxiv.org/abs/2102.08945

[2] H. Mittal, B. Okorn, and D. Held, "Just go with the flow: Self-supervised scene flow estimation," 2019. [Online]. Available: https://arxiv.org/abs/1912.00497

[3] K. Sirohi, R. Mohan, D. Büscher, W. Burgard, and A. Valada, "Efficientlps: Efficient lidar panoptic segmentation," 2021. [Online]. Available: https://arxiv.org/abs/2102.08009