Week 5 Report

Doğaç Güzelkokar, Niklas Bubeck

July 2022

1 Week 5 Goals

The goals of this weeks are to experiment with self-supervised sceneflow methods like Just Go with The Flow[1] as well as PointPWCNet[2] and evaluate those against the previous supervised methods as FLOT[3]. Furthermore, training with self-supervised learning models on the preprocessed semanticKitti[4] data is another goal.

2 SceneFlow

2.1 Experiment: FLOT against PointPWC

2.1.1 Methods

We applied the same voxel downsampling (roughly 120k to 10k points and with 0.01m scale for things classes and 1m for stuff classes) to both methods. While the FLOT algorithm can handle different number of points for the inference step, the PointPWC method needs exactly 8192 points. Thus, we from the previously voxeld points we again randomly sample 8192 points from it.
Important is to keep in mind, that both models were trained on the same datasets (sceneFlowKitti[5], FlyingThings3D[6]). PointPWC additionally needs colorvalues. Below results use a simple random init of the color values.

2.1.2 Results

Tl;dr
PointPWC keeps local structures but lacks in precision.

Fig. 1 and Fig. 2, show a comparison of the method for a scenario with comparably large rotation and translation. It shows that though Flot exceptionally works great for car structure with good precision, it fails for vertical objects like poles and signs, ending up with cluttered points (also form multiple different objects). The PointPWC method on the other hand, does not have such high precision, but is much better at keeping local structure and identifying single objects. This can especially be seen for the pedestrian in Fig. 4.
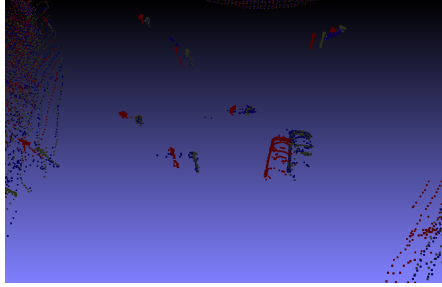
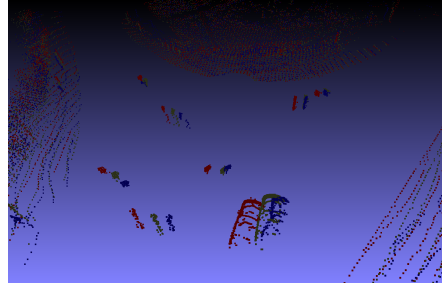Figure 1: Flot on scenario 2, scene 70/71.

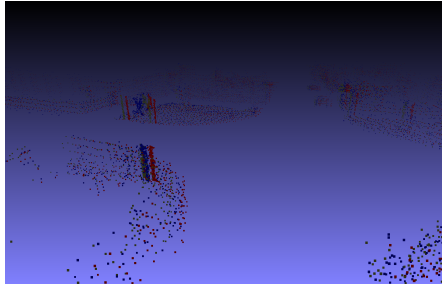Figure 2: PointPWC on scenario 2, scene 70/71.

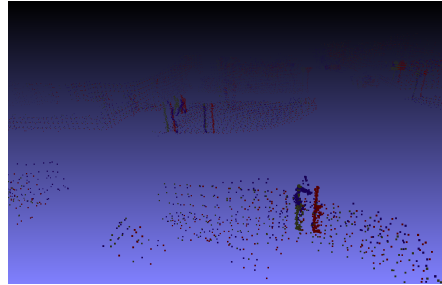Figure 3: Flot on scenario 5, scene 10/11.

Figure 4: PointPWC on scenario 5, scene 10/11.

2.1.3 Outlook

From this results we have high hopes that performance can be further boosted with training the network on data of a preprocessed semanticKitti dataset and train the color values based on the intensity. Or in general that self-supervised sceneflow can achieve great results on the semanticKitti dataset.

2.2 Just Go With The Flow

We mostly tried to run Just Go With The Flow on Titan X GPU at atcremers9. This process is blocked by the "illegal instruction" error. Therefore, currently we are adapting it to run on a GTX 1050 and trying to understand and hopefully fix the illegal instruction error, as the ability to use the Titan X would be helpful for us.

2.3 Training PointPWC

We adapted the data loader of the PointPWC method to take the downsampled semanticKitti data as input. Evaluation of the training is currently pending.

2.4 Technical Issues

2.4.1 Hardware Incompatibiliy on cremers9

Running training of both Just go With The Flow, as well as PointPWC, results in a **Illegal Instruction** error. We are not sure how to overcome this. Regarding the training, 4Gb makes it only possible to train with a batch size of 1.

References

[1] H. Mittal, B. Okorn, and D. Held, "Just go with the flow: Self-supervised scene flow estimation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11 177–11 185.

[2] W. Wu, Z. Wang, Z. Li, W. Liu, and L. Fuxin, "Pointpwc-net: A coarse-to-fine network for supervised and self-supervised scene flow estimation on 3d point clouds," *arXiv preprint arXiv:1911.12408*, 2019.

[3] G. Puy, A. Boulch, and R. Marlet, "Flot: Scene flow on point clouds guided by optimal transport," 2020.

[4] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, J. Gall, and C. Stachniss, "Towards 3D LiDAR-based semantic scene understanding of 3D point cloud sequences: The SemanticKITTI Dataset," *The International Journal on Robotics Research*, vol. 40, no. 8-9, pp. 959–967, 2021.

[5] M. Menze, C. Heipke, and A. Geiger, "Object scene flow," *ISPRS Journal of Photogrammetry and Remote Sensing (JPRS)*, 2018.

[6] N. Mayer, E. Ilg, P. Häusser, P. Fischer, D. Cremers, A. Dosovitskiy, and T. Brox, "A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation," in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, arXiv:1512.02134. [Online]. Available: http://lmb.informatik.uni-freiburg.de/Publications/2016/MIFDB16