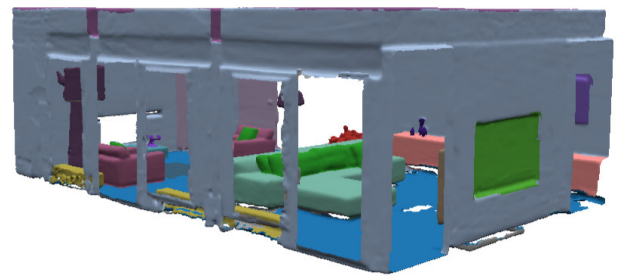
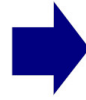




Input PTX point cloud



Semantic Segmentation result

Project Description

Semantic information of 3D scans is a fundamental step for scene understanding [1]. Given a 3D point cloud, the desired output is a set of semantic labels (e.g., chair, table, window) over all input points.

The Minkowski engine [2] is an auto-differentiation library for sparse tensors. It supports all standard neural network layers, such as convolution, pooling, unpooling, and broadcasting operations for sparse tensors. In this project, we will study methods that use the Minkowski engine as the backbone architecture for 3D semantic segmentation.

The goal of the project is to compare and analyze existing methods' strengths and weaknesses with respect to different point cloud characteristics (e.g., irregular sampling, occlusions, and low resolution).

Tasks

- Perform statistical analysis (density sampling, degree of occlusion) of indoor scene data from two different sources: 2.5D data (scan with spherical images), 3D data (photogrammetry).
- Reproduce methods results with the existing datasets and benchmarks. (S3DIS [3], Scannet [4])
- Evaluate the ability to generalize to new scan models (unseen in the training).
- The metrics considered for comparison are:
 - Training time.
 - Model complexity.
 - Accuracy evaluation (IoU, F1 score, precision, recall).

Requirements

Interest in 3D graphics and deep learning. Application development in Python.

Work Load

- 30% Theory
- 50% Implementation
- 20% Test

Project Type

Based on the scope of the topic and optional tasks, this project goals can be adjusted for Bachelor or Master thesis.

Supervision

- Prof. Dr. Renato Pajarola
- Lizeth J. Fuentes Perez (assistant)

Contact

Write an email to fuentes@ifi.uzh.ch.

References

- [1] Yulan Guo, Hanyun Wang, Qingyong Hu, Hao Liu, Li Liu, and M. Benamoun. Deep learning for 3d point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence*, PP, 2020.
- [2] Christopher Choy, JunYoung Gwak, and Silvio Savarese. 4d spatio-temporal convnets: Minkowski convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3075–3084, 2019.
- [3] Iro Armeni, Ozan Sener, Amir R. Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, and Silvio Savarese. 3d semantic parsing of large-scale indoor spaces. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1534–1543, 2016.
- [4] Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. ScanNet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 2432–2443, 2017.