TORONTO-3D

A Large-scale Mobile LiDAR Dataset for Semantic Segmentation of Urban Roadways

The dataset can be accessed at **this link**. A detailed description of the dataset and semantic segmentation benchmarks are published at **2020 CVPR Workshops**.

Semantic segmentation of large-scale outdoor point clouds is essential for urban scene understanding in various applications, especially high-definition (HD) for smart cities and autonomous cars. High-quality publicly accessible large-scale labelled datasets are essential for developing learning-based methods. The intention of releasing this new Toronto-3D dataset is to encourage developing creative deep learning models on mobile LiDAR point clouds.

The proposed dataset, Toronto-3D is a large-scale labelled urban outdoor point cloud dataset acquired by an mobile laser scanning (MLS) system. The dataset covers approximately 1 km of urban road and consists of about 78.3 million points in Toronto, Ontario, Canada. The point clouds were collected by a vehicle-mounted 32-line LiDAR sensor, having a high point density of approximately 1000 points/m² on road surfaces. The dataset was manually classified into 8 classes: road, road marking, natural, building, utility line, pole, car and fence. The point clouds have the following 10 attributes: x, y, z, R, G, B, intensity, GPS time, scan angle rank and label. The dataset is provided in four sections illustrated in Figure 1, and samples of the dataset is shown in Figure 2.

Baseline experiments for semantic segmentation were conducted and the results confirmed the capability of this dataset to train deep learning models effectively. We encourage the research community to develop and test creative methods in HD map generation, urban 3D modelling and other relevant areas.
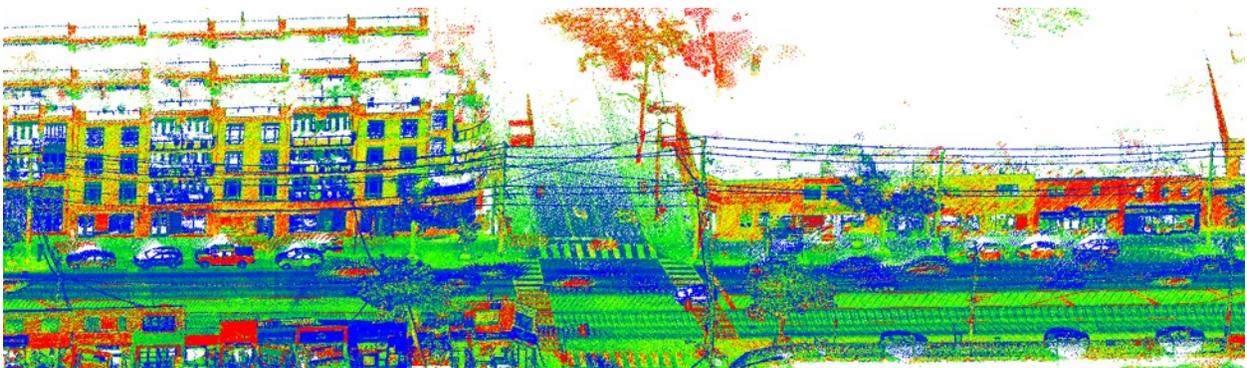

Figure 1 Overview of Toronto-3D

RGB

Intensity

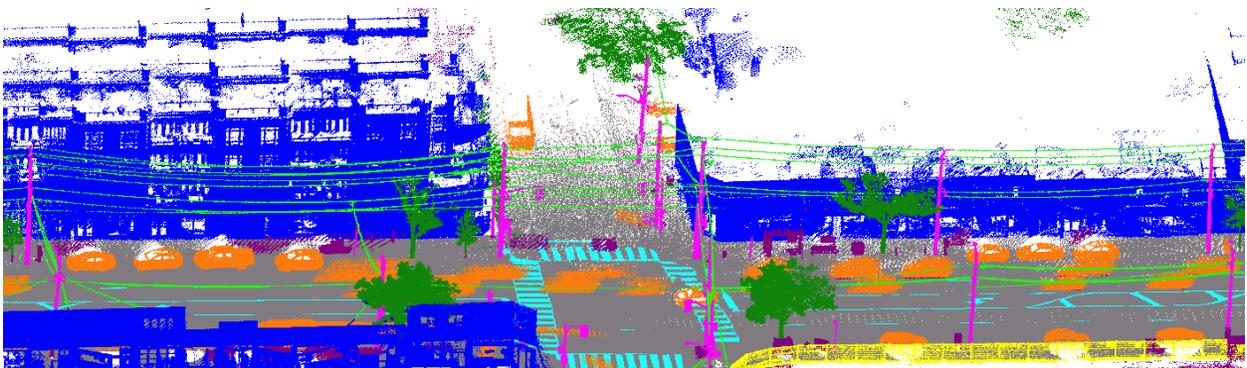Classification

| unclassified | Road | Road mrk. | Natural | Building | Util. line | Pole | Car | Fence |

Figure 2 Sample of Toronto-3D