Manual-Label Free 3D Detection via An Open-Source Simulator

Zhen Yang, Chi Zhang, Huiming Guo, and Zhaoxiang Zhang

Institute of Automation, Chinese Academy of Sciences (CASIA)
Beijing Aerospace Changfeng Co.Ltd., The 2nd Institute of CASIC

Background

Recently, the simulators are being increasingly used to remedy the shortage of labeled data.

Challenges

The synthetic data is severely distorted, and such discrepancies would cause significant performance drop.

Our Methods

LiDAR-guided Sampling:
- Virtual lidar helps reduce the distortion of the synthetic point cloud data.
- Domain adaptation: Feature alignment helps reduce the impact of the discrepancy between the synthetic data and the real data on model performance.

LiDAR-guided sampling: The high fidelity point cloud samples obtained by using LiDAR-guided sampling method can improve the detector's generalization ability on real scenes.

DA-VoxelNet: DA-VoxelNet gain a large performance improvement compared to the VoxelNet, which reveals a promising perspective of training a LiDAR-based 3D detector without any hand-tagged label.

Contribution

1) We produce high quality 3D models and embed these models into CARLA simulator to get more realistic virtual point clouds. We then propose a novel sampling algorithm, LiDAR-guided sampling, to generate high fidelity point cloud samples. By utilizing the high fidelity point clouds to augment the training set, we can implement a promising 3D detector with exponentially reduced manual labeled data.

2) We propose two novel domain adaptation components to cross the gap between the synthetic data and the real data. We further impose a consistency constraint to stabilize the training process. Combine the both themes and based on the 3D detector VoxelNet, the proposed DA-VoxelNet can get rid of the manual annotations thoroughly.

Algorithm and Objective Function

LiDAR-guided sampling:

$$f(x;\beta,D) = \frac{1}{N} \sum_{i=1}^{N} \text{log} (1 - D_i(x;\beta))$$

$$\delta(d) = \frac{1}{d \geq \delta_{\text{max}}}$$

$$f(x;\theta,P) = \int f(x;\theta,P) \mathcal{D}(x;\phi(x)) \text{d}x$$

$$P = D_i / P_i$$

Anchor-Level Adaptation:

$$L_{\text{consistency}} = -\frac{1}{N_i} \sum_{i=1}^{N_i} \text{log} (D_i(x;\beta))$$

$$L_{\text{anchor}} = L_{\text{consistency}} + L_{\text{anchor}}$$

Consistency Constraint:

$$M_i(x;\phi) = \frac{1}{N_i} \sum_{i=1}^{N_i} D_i(x;\phi(x))$$

$$M(x;\phi) = \frac{1}{N_i} \sum_{i=1}^{N_i} D_i(x;\phi(x))$$

Overall Objective:

$$L = L_{\text{anchor}} + L_{\text{consistency}}$$

Experimental Results

Conclusions

- LiDAR-guided sampling is helpful.
- The high fidelity point cloud samples obtained by using LiDAR-guided sampling method can improve the detector’s generalization ability on real scenes.
- DA-VoxelNet can relieve the domain gap.