Pointwise Convolutional Neural Networks
Supplemental Document

Binh-Son Hua   Minh-Khoi Tran   Sai-Kit Yeung

The University of Tokyo   Singapore University of Technology and Design

Here we show in Table 1 an additional performance comparison with state-of-the-art 3D deep learning techniques including PointNet [2] for the scene segmentation task. Our network is implemented in Tensorflow. For PointNet, we use an re-implementation in PyTorch. Both implementations are available in our code repository on our homepage.

We train both networks from scratch on SceneNN dataset. We re-annotate 76 scenes in SceneNN with 40 semantic classes from NYU-D v2 dataset [3]. The training set consists of 56 scenes, and 20 scenes are used for testing. As can be seen, our method has very competitive accuracy to PointNet and even outperforms at several common classes.

Class	Pointwise [1]	PointNet [2]	Class	Pointwise [1]	PointNet [2]
wall	0.868	**0.897**	table	**0.412**	0.235
floor	0.864	**0.891**	counter	**0.144**	0.052
cabinet	**0.214**	0.090	desk	**0.362**	0.310
bed	**0.513**	0.457	pillow	**0.175**	0.067
chair	**0.639**	0.596	tv	**0.178**	0.114
sofa	**0.298**	0.167	box	0.141	**0.163**

Table 1: Performance comparison to other 3D deep learning methods.

References

[1] B.-S. Hua, M.-K. Tran, and S.-K. Yeung. Pointwise convolutional neural networks. In *Computer Vision and Pattern Recognition (CVPR)*, 2018. 1

[2] C. R. Qi, H. Su, K. Mo, and L. J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. *CVPR*, 2017. 1

[3] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus. Indoor segmentation and support inference from rgbd images. In *ECCV*, 2012. 1