Supplemental Material
End-to-End CAD Model Retrieval and 9DoF Alignment in 3D Scans

Armen Avetisyan Angela Dai Matthias Nießner

Technical University of Munich

A. Online Benchmark

In this supplemental document, we provide additional results, including measurements on the hidden test set of the Scan2CAD benchmark [1]. Specifically, we provide a quantitative comparison in Tab. 1, which was submitted to official benchmark website on March 29th, 2019. In addition, we show qualitative results of our approach in Fig. 1.

B. SUNCG

We conduct experiments on the SUNCG dataset [6] to verify the effectiveness of our method. For training and evaluation, we create virtual scans of the synthetic scenes, where we simulate a large-scale indoor 3D reconstruction by using rendered depth frames similar to [3, 2] with the distinction that we add noise to the synthetic depth frames in the fusion process. The voxel resolution for the generated SDF grids is at $4.68cm$. The ground truth models are provided by the SUNCG scenes, where we discard any objects that have not been seen during the virtual scanning (no occupancy in the scanned SDF). We show a quantitative evaluation in Tab. 2, where we outperform the current state-of-the-art method Scan2CAD [1] by a significant margin. We show that our method can align CAD models robustly through all classes. Additionally, we see that our Procrustes loss notably improves overall alignment accuracy. In particular, for less frequent CAD models (e.g., those summarized in *other*), we observe a considerable improvement in alignment accuracy.

Fig. 2 shows qualitative results on scanned SUNCG scenes. Our end-to-end approach is able to handle large indoor scenes with complex furniture arrangements.

References

[1] Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang, and Matthias Nießner. Scan2cad: Learning cad model alignment in rgb-d scans. In *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 2019. 1, 2

[2] Angela Dai, Daniel Ritchie, Martin Bokeloh, Scott Reed, Jürgen Sturm, and Matthias Nießner. Scancomplete: Large-scale scene completion and semantic segmentation for 3d scans. In *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 2018. 1

[3] Ji Hou, Angela Dai, and Matthias Nießner. 3d-sis: 3d semantic instance segmentation of rgb-d scans. In *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 2019. 1

[4] Yangyan Li, Angela Dai, Leonidas Guibas, and Matthias Nießner. Database-assisted object retrieval for real-time 3D reconstruction. In *Computer Graphics Forum*, volume 34, pages 435–446. Wiley Online Library, 2015. 2

[5] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (fpfh) for 3d registration. In *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, pages 3212–3217. Citeseer, 2009. 2

[6] Shuran Song, Fisher Yu, Andy Zeng, Angel X Chang, Manolis Savva, and Thomas Funkhouser. Semantic scene completion from a single depth image. *Proceedings of 30th IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 1, 2

[7] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. In Kostas Daniilidis, Petros Maragos, and Nikos Paragios, editors, *Computer Vision – ECCV 2010*, pages 356–369, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. 2

[8] Andy Zeng, Shuran Song, Matthias Nießner, Matthew Fisher, Jianxiong Xiao, and Thomas Funkhouser. 3dmatch: Learning local geometric descriptors from rgb-d reconstructions. In *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, pages 199–208. IEEE, 2017. 2

	bath	bookshelf	cabinet	chair	display	sofa	table	trash bin	other	class avg.	avg.
SHOT (Tombari et al. [7])	0	1.8	0	8.8	0.0	1.2	0	0	2.2	1.5	2.8
FPFH (Rusu et al. [5])	0	0	1.5	10.7	0	1.2	2.1	2.9	0	2.0	3.7
Li et al. [4]	0	1.8	2.3	1.11	0	2.8	6.4	2.7	0	3.0	4.6
3DMatch (Zeng et al. [8])	0	5.3	3.8	19.5	1.7	5.2	17.0	6.0	6.5	7.2	9.2
Scan2CAD (Avetisyan et al. [1])	25.0	28.1	30.8	39.7	20.3	14.3	51.1	31.5	19.6	28.9	28.8
Ours	**40.6**	**38.6**	**36.2**	**68.1**	**25.4**	**27.0**	**63.8**	**38.0**	**40.2**	**42.0**	**44.1**

Table 1: Performance comparison (%) on the hidden test set of the Scan2CAD alignment benchmark [1]. We outperform existing methods by a significant margin on all classes; the last two rows provide class and average instance alignment accuracy, respectively.

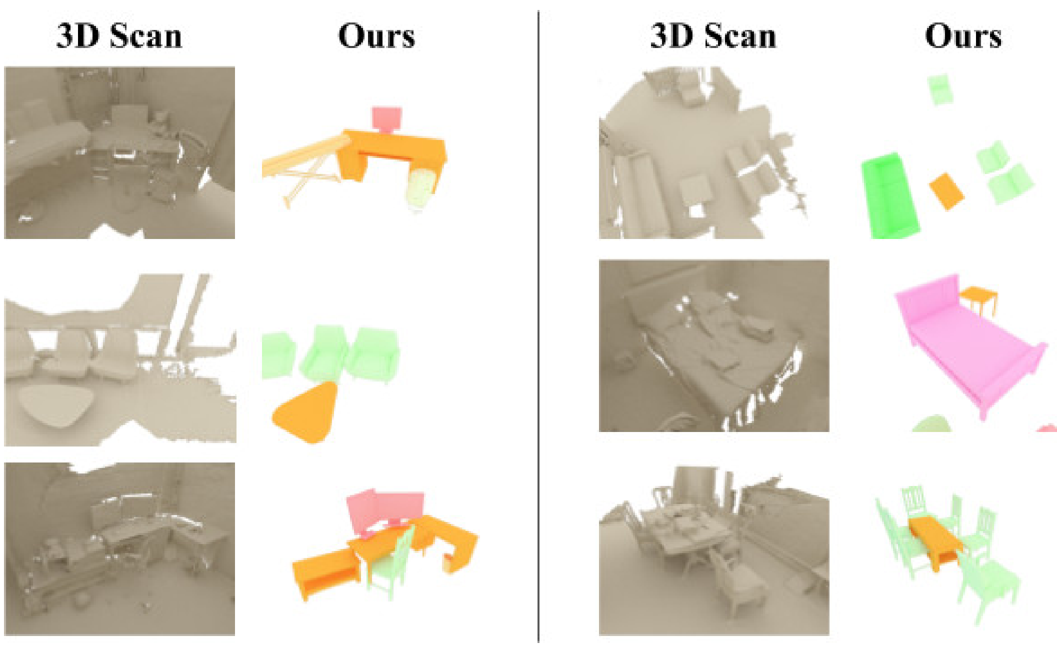

Figure 1: Qualitative results on the Scan2CAD alignment benchmark [1] (submitted to official benchmark website on March 29th, 2019)

	bed	cabinet	chair	desk	dresser	other	shelves	sofa	table	class avg.	avg.
SHOT (Tombari et al. [7])	13.43	3.23	10.18	2.78	0	0	1.75	3.61	11.93	5.21	6.3
FPFH (Rusu et al. [5])	38.81	3.23	7.64	11.11	3.85	13.21	0	21.69	11.93	12.39	9.94
Scan2CAD (Avetisyan et al. [1])	52.24	17.97	36	**30.56**	3.85	20.75	7.89	40.96	43.12	28.15	29.23
Ours (No Procrustes)	**71.64**	29.95	39.27	23.61	30.77	20.75	9.65	**69.88**	40.37	37.32	36.42
Ours (final)	**71.64**	**32.72**	**48.73**	27.78	**38.46**	**37.74**	**14.04**	67.47	**45.87**	**42.72**	**41.83**

Table 2: CAD alignment accuracy comparison (%) on SUNCG [6]. We compare to state-of-the-art handcrafted feature descriptors FPFH [5], SHOT [7] as well as a learning based method Scan2CAD [1] for CAD model alignment. Note that the Procrustes loss considerably improves overall alignment accuracy.

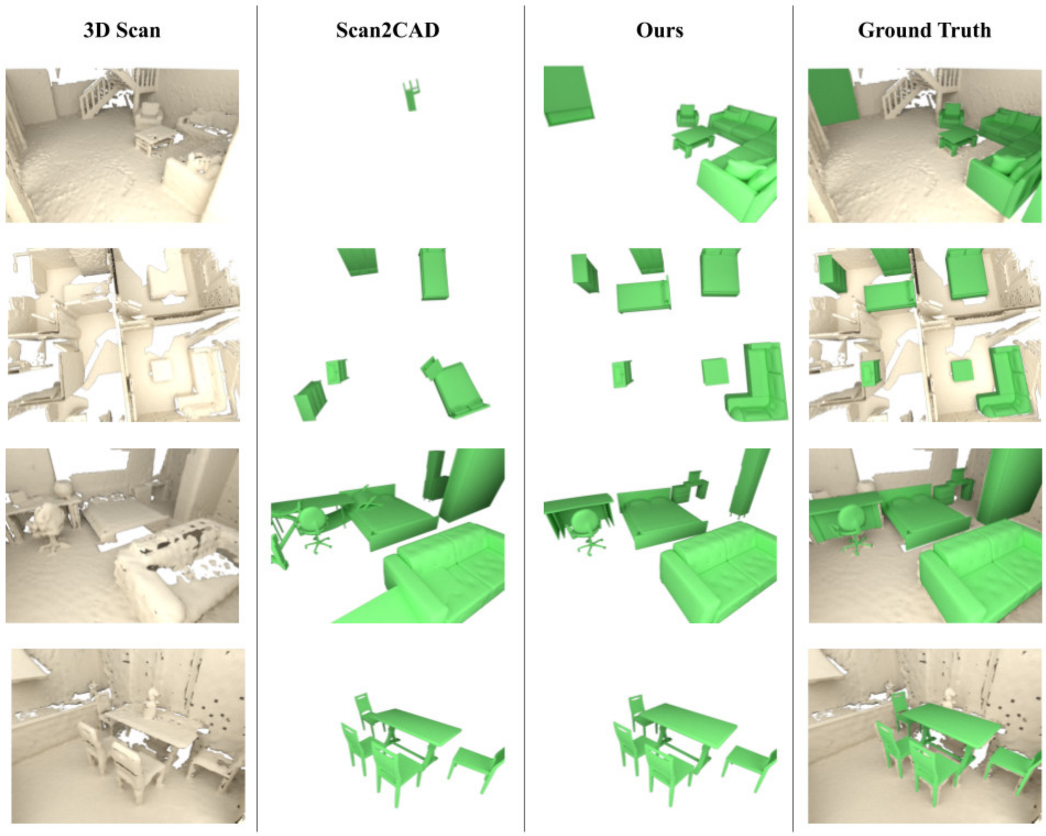

Figure 2: Qualitative results on virtual scans from SUNCG. Note that our method handles complex CAD arrangements better than Scan2CAD.