

Real2CAD: Shape Matching of Real 3D Object Data to Synthetic 3D CADs

Yue Pan
ETH Zurich
yuepan@ethz.ch

Bingxin Ke
ETH Zurich
bingke@ethz.ch

Yuanwen Yue
ETH Zurich
yuayue@ethz.ch

Yujie He
EPFL
yujie.he@epfl.ch

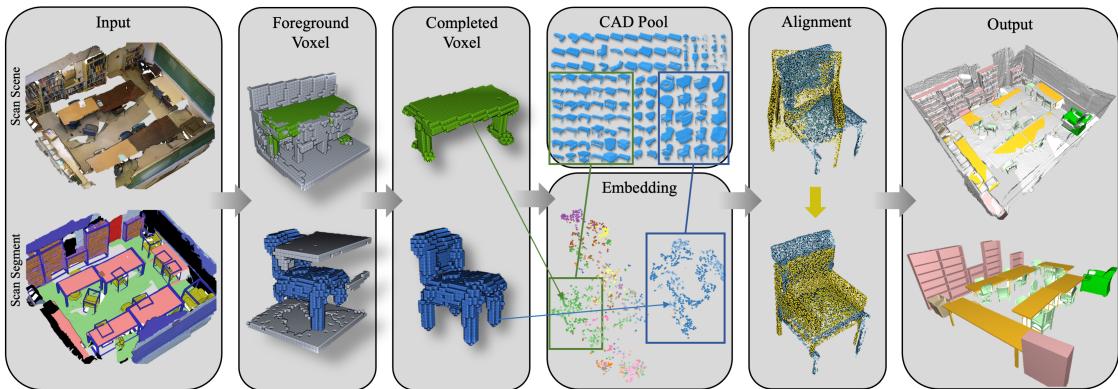


Figure 1: Overview framework of our Real2CAD system. We first generate the 3D object proposals from the 3D scan and then train a neural network to conduct foreground separation and completion using the voxel representation. The network then embeds the completed scan into a joint embedding manifold to efficiently retrieve the best match CAD model from a filtered model pool with the predicted category. Finally, each corresponding CAD and scan segments are aligned with each other in a coarse-to-fine fashion to generate the CAD-based compact map of the scene.

Abstract

Establishing a mapping between the real 3D scan and CAD model is challenging due to substantial differences between domains. Based on joint embedding [9], a novel end-to-end scan to CAD retrieval neural network is proposed. The foreground object is separated from the background cluttering and further completed into a CAD-like representation through the network. A mixed strategy of offline triplet sampling is applied to learn a rotation-aware joint embedding of scan and CAD to enable the fine-grained CAD retrieval with the approximate rotation. Besides, a classification module is connected to the network bottleneck to provide semantic supervision and realise the category-based model pool filtering. Next, based on the initial rotation guess, a scale-aware ICP is applied to align the retrieved CAD model in the scan scene. The proposed method is evaluated for both the CAD retrieval and alignment task on scan and CAD similarity [9] and Scan2CAD [1] benchmark. Extensive experiments demonstrate that our method outperforms the state-of-art algorithms by 19% in retrieval accuracy and 3% in alignment accuracy. ¹

1. Introduction

3D semantic reconstruction of indoor environments has been exhaustively studied in recent years with the increasing availability of consumer-grade RGB-D sensors. State-of-the-art 3D reconstruction approaches have driven forward many potential applications, such as indoor navigation, architectural design, augmented and virtual reality [7, 11, 20].

Semantic segmentation based approaches have achieved impressive reconstruction quality but remained limited in scenarios that contain many artefacts such as noise, missing surface parts due to occlusions and imperfections of existing sensors. Building upon the availability of synthetic CAD models, an alternative is to reconstruct the scene by retrieving CAD models from existing datasets and aligning them to the real scans, instead of reconstructing 3D geometry in a bottom-up manner. Unfortunately, matching CAD model to scan geometry is also challenging due to (1) significant difference between CAD model geometry (clean and complete) and scan geometry (noisy, incomplete) in low-level geometric features; (2) a lack of exact matches between synthetic models and real-world objects.

A pioneering work towards this problem employed a deep neural model to learn pairwise keypoint correspondences between CAD model and input scan, with a further optimization over potential matching correspondences for each candidate CAD model [1]. Avetisyan et al. [2] combine the decoupled steps in [1] and propose an end-

¹All members contribute equally.

to-end approach for scan-to-CAD alignment featuring. To tackle the lower-level geometric differences, Dahnert et al. [9] propose a model retrieval approach by learning a joint embedding space between scan and CAD geometry, where semantically similar objects from both domains lie close together. However, the above works either focus on scan to CAD retrieval in a single canonical pose or CAD to scan alignment with an exhaustive search of the best match model. In this work, we propose the Real2CAD pipeline for both model retrieval and alignment to generate the CAD-based compact map from the real scene.

Built upon the joint-embedding work [9], our approach applies a novel triplet sampling strategy to learn a rotation-aware joint embedding of scan and CAD, thus enabling the fine-grained CAD retrieval with the approximate rotation. Besides, instead of performing model retrieval from the whole CAD pool, we employ a classification module to provide semantic supervision and realize the category-based model pool filtering. A coarse-to-fine registration pipeline is proposed to align the retrieved CAD model to the scan scene. Our approach outperforms state-of-the-art methods for CAD model retrieval and alignment by 19% and 3%, respectively.

2. Related works

In recent years, the retrieval of CAD models and matching real-world RGB-D scan data has received increasing attention. Earlier studies mainly used categories for matching evaluation due to the lack of sufficient real-world data for CAD models annotations. SHREC challenges [14] for CAD model retrieval to real-world object scans have become very popular in this regard. Therefore, Avetisyan et al. proposed a large-scale Scan2CAD dataset [1] based on real-world dataset ScanNet [10] and synthetic objects dataset ShapeNet [5] for keypoint-based matching and 9DoF alignment between synthetic models. Based on Scan2CAD, the authors further developed the Scan-CAD object similarity benchmark [9] by ranking the similarities between the scanned object data with CAD model, which provide the potential for cross-domain embeddings learning.

Hand-crafted features such as Fast Point Feature Histograms (FPFH) [18], or Point-Pair Features (PPF) [12], or truncated signed distance function (TSDF) based volumetric surface representation format [8], are widely used for shape descriptors to develop methods for CAD model retrieval and alignment [16, 15]. However, these retrieval and matching methods struggle to generalize to noisy, incomplete real-world data and especially struggle to close the gap from real data to CAD model.

The recent studies on real scan to CAD alignment are mainly focused on learning-based method. Based on 3D CNN architecture for heatmap prediction, Avetisyan et al. [1] proposed a new variational optimization formulation to minimize the distance between scan keypoints heatmaps for 9DoF alignments, which further extended to end-to-end ashion [3]. Moreover, Dahner et al. [9] proposed a stacked hourglass approach with a triplet loss to learn a joint embedding space between CAD model and 3D scan to achieve

retrieval from an input scan object. [4] also applied layout elements in the scene together with segmented 3D object data to form a node in a graph neural network, thus enabling globally consistent CAD model alignments. Different from the aforementioned method, Hampali et al. [13] proposed a modified Monte Carlo Tree Search to retrieve objects and room layouts.

3. Proposed method

3.1. Scan to CAD retrieval network

The overall architecture of the proposed end-to-end CAD model retrieval network is shown in Fig. 2. The input is the 32^3 voxel representation of the 3D object’s truncated distance function (TDF). To make the scan to CAD retrieval easier, we need to transform the scan segment to a form much similar to the CAD model, which is without the cluttered background and with the complete shape. The completed object is classified into coarse categories (chair, table, etc.) and then used to learn the scan-CAD joint embedding space for rotation-aware fine-grained model retrieval.

3.1.1 Foreground separation

The object is firstly fed through a 3D-CNN based encoder-decoder with skip connection to separating the foreground x_{fg} from the background. The separation is formulated as a binary classification problem with a proxy loss $L_{fg} = BCE(x_{fg}, gt_{fg})$ using balanced weight for each class (occupied by foreground or not).

3.1.2 Scan segment completion

The foreground segment x_{fg} is then fed into another encoder-decoder to generate the complete geometry of the object x_{cmp} , again with the proxy loss $L_{cmp} = BCE(x_{cmp}, gt_{cmp})$ with balanced weight, where gt_{cmp} is the CAD model corresponding to the scan segment.

3.1.3 Coarse category classification

We add a simple 2-layer category classifier with dropout as a branch to the bottleneck of the completion module, after the last layer of the hourglass encoder. The cross-entropy classification loss $L_{cls} = CE(x_{cls}, gt_{cls})$ is used here. During the test procedure, the prediction of this classifier would be used to filter the candidate CAD pool.

3.1.4 Fine-grained model retrieval

The completed scan object is encoded together with its negative and positive CAD samples to do metric learning in the joint scan-CAD embedding space. A triplet loss is leveraged here, formulated as: $L_{rtv} = \max(d(x_{cmp}, C_p) - d(x_{cmp}, C_n) + \alpha, 0)$, which tries to minimize the distance between the completed scan object x_{cmp} and positive CAD model C_p and maximize the distance between x_{cmp} and negative CAD model C_n with a margin α . We propose a novel offline sampling strategy to constitute the semi-hard triplets for more efficient and robust metric learning. To guarantee the robustness, we apply a random tiny rotation ($< 15^\circ$) to the well-aligned ground truth CAD model to generate the positive sample. As for the negative CAD, three sample strategies are designed and compared by us:

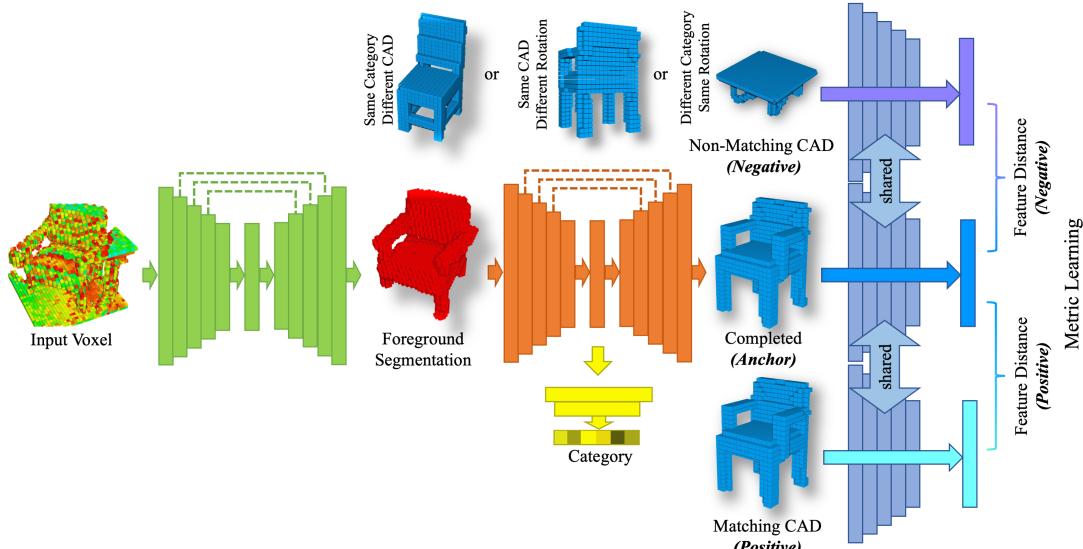


Figure 2: Architecture of the proposed rotation-aware CAD-model retrieval network

- Other category: the negative CAD models come from different categories as the completed scan object.
- Inner-category: the negative CAD models come from the same category as the completed scan object but excludes the ground-truth CAD model.
- Mixed-categories: One-third of the negative CAD models are sampled by strategy one, and the other third are sampled by strategy two. The last third is sampled from the ground-truth CAD model, but with different rotation angles around z-axis.

The experimental result in Section 4.3 reveals that the sampling with mixed-categories achieved the best performance for the downstream task in our case.

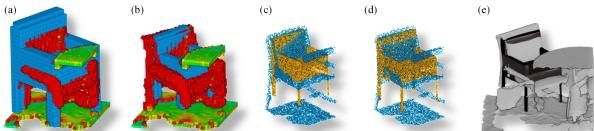


Figure 3: Coarse-to-fine registration pipeline: (a) align the center of the retrieved CAD voxel (blue) with the scan (green to red), (b) re-scale the CAD model to have the same voxel size as the scan, (c) convert both the scan and CAD to point cloud, (d) apply scale-aware registration, (e) registered CAD in the scene.

3.2. Coarse-to-fine alignment

The proposed CAD to scan alignment framework is shown in Fig. 3. Since our retrieval network is trained with the ability to distinguish rotation, we can achieve rotation-aware retrieval at inference time. By cloning each CAD model in the pool into 12 objects and applying rotation ranging from 0 to 330 degree around z-axis, we can get a rotation initial guess with the mean uncertainty of 15° theoretically. The translation and scale initial guess is achieved by aligning the centre and the size of the voxels of the scan segment and the retrieved CAD model.

Based on the initial guess, the scale-aware ICP algorithm is used to refine the relative pose. We adopt the point-to-point distance metric and continuously decrease the nearest neighbour correspondence distance threshold throughout the iterations. To compensate for the potential error of CAD model retrieval, we take the top 4 retrieval model to apply ICP and select the one with the highest fitness score.

4. Experiments

4.1. Benchmark

We train our model retrieval network on the Scan2CAD dataset [1], which provides CAD model alignments for 3049 unique ShapeNetCore [5] models to objects in 1506 ScanNet [10] scans. More details of the datasets can be found in Table 1. The training data preparation process is shown in Fig. 4. Based on ScanNet [10] annotations, we first extract scan segment, from which the scan object is separated using a foreground mask. Then, the best match CAD model is selected from ShapeNet [5] together with the relative rotation based on Scan2CAD [1] annotations and is regarded as the ground-truth CAD model corresponding to the scan object.

Our experimental evaluation focuses on eight classes: chair, table, trash bin, bed, bookshelf, cabinet, sofa, and others. Different datasets and metrics are used to evaluate the performance of model retrieval and model alignment:

- CAD retrieval: the network is evaluated on the Scan-CAD Object Similarity dataset [9], which consists of ranked scan-CAD similarity annotations. We first evaluate on a coarser level retrieval score based on whether the retrieved model's category is correct. For CAD model-based retrieval, we evaluate whether the top-1 retrieved model lies in the set of models annotated as similar to the query scan for retrieval accuracy.
- CAD alignment: we adopted the alignment benchmark from Scan2CAD [1]. A CAD alignment is considered successful only if the retrieved CAD model matches

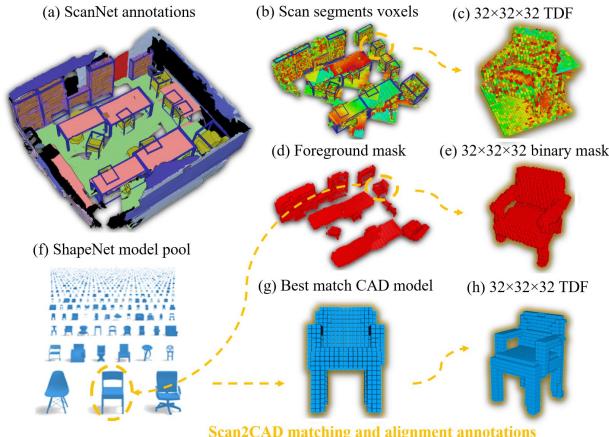


Figure 4: Training data preparation from ScanNet, ShapeNet and Scan2CAD datasets

that of the scan object and the pose error is within translation, rotational, and scale bounds.

4.2. Implementation details

Our model is trained from scratch in an end-to-end fashion on an Nvidia GTX 1080Ti. We use an Adam optimizer with a batch size of 128 and an initial learning rate of 0.001, which is decreased by 90% every 10K iterations. We train for 30K iterations (≈ 500 epochs, 1 day) with equal weight among different losses and a triplet margin $\alpha = 0.2$.

4.3. Overall evaluation

4.3.1 CAD model retrieval

The qualitative results of rotation-aware CAD model retrieval are shown in Fig. 6. The model pool is firstly filtered with the predicted category and the mixed negative sampling strategy is used here for retrieving both the CAD model and the rotation. The proposed method can correctly predict the category and for most of the cases select the correct CAD model with the closest rotation to the ground truth. Still, we can notice some failure cases for the fourth column (lamp) and the seventh column (cabinet) due to occluded scan input. The quantitative results of the proposed method are shown in Table. 2, together with its variants and the compared algorithms. With the model pool filtering and the inner-category negative CAD sampling, our method outperforms the state-of-the-art methods by 19% in instance retrieval accuracy.

Dataset	Contents
ShapeNet [5]	It covers 55 common object categories with about 51,300 unique 3D models
ScanNet [10]	It contains 2.5 million views in more than 1500 scans, annotated with 3D camera poses, surface reconstructions, and instance-level semantic segmentations.
Scan2CAD [1]	It contains 97607 keypoint correspondences between 1506 scans and 14225 CAD models.
Scan-CAD Similarity [9]	It consists of 5102 ranked scan-CAD similarity annotations covering 3979 unique scan objects and 7650 unique CAD models.

Table 1: Descriptions of the adopted datasets

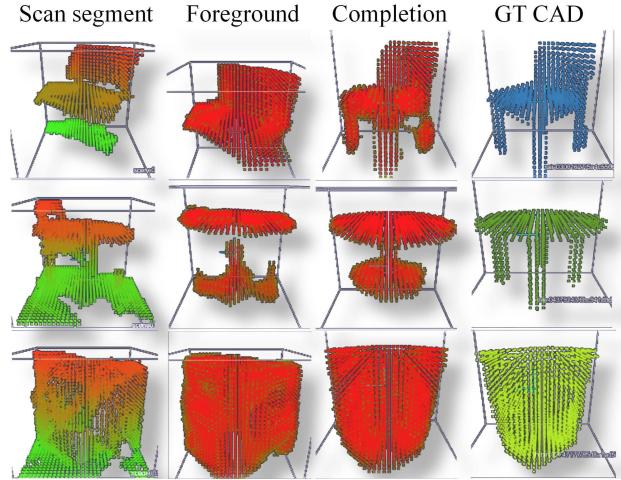


Figure 5: Foreground separation and scan completion results of the proposed method. Each point represents a occupied voxel of the 3D object. The second and third column are rendered with occupied probability.

4.3.2 Foreground separation and completion

Fig. 5 shows that as the upstream task for CAD model retrieval, the trained foreground separation and completion networks can extract the foreground with fine details and then transform itself into a clear and smooth CAD-like object. Still, there's some failure case such as the second row, where the three legs of the table are reconstructed as a single leg table since the training data is biased to the latter case. The proposed method outperforms the state-of-art algorithm JointEmbedding [9] with a large margin (19%) on model-based retrieval accuracy. The boost of accuracy is mainly due to the hierarchical retrieval strategy and the high category-based classification accuracy. However, the evaluation metric is only based on the specific CAD ID and this model cannot find the correct rotation of the CAD model at the same time. By using the mixed sampling strategy, the specific rotated CAD can also be correctly retrieved according to the input scan with small sacrifice of the rotation independent accuracy.

4.3.3 Learned joint embedding space

Our learned scan-CAD embedding space is shown in Fig. 8, visualized by t-SNE. Both the rotation-aware CAD embedding space and the scan-CAD joint embedding space are shown for both the inner-class and mixed-class negative sampling strategy. We can find that both the category (colour) and rotation (marker type) in the CAD embedding space generated by the mixed strategy are separated clearly. Besides, for the major categories, the joint embedding space is mixed together for both scan and CAD objects.

4.3.4 CAD to scan alignment

We also evaluate the CAD alignment performance of the proposed method on the Scan2CAD dataset [1]. Qualitative and quantitative results are shown in Fig. 7 and Table. 3, respectively. Our model retrieval and coarse alignment module provide a good enough initial guess for ICP to refine the registration, resulting in an accuracy better or on-par to

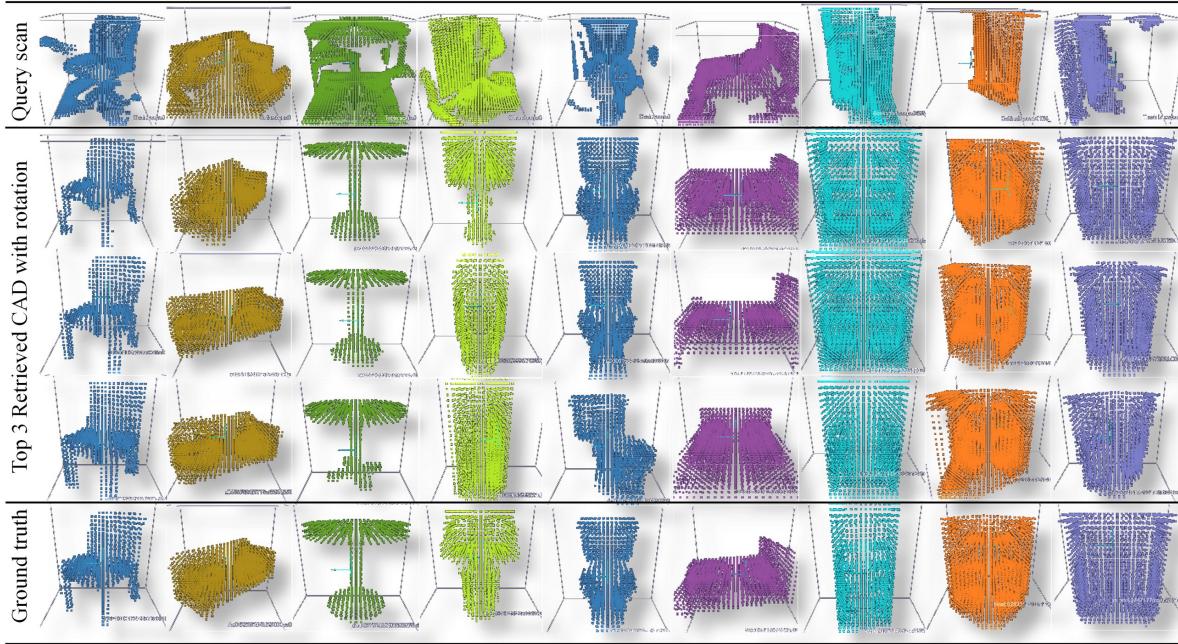


Figure 6: Scan to CAD model retrieval results of the proposed method, visualizing the top 3 retrieved models with rotation together with the ground truth. Unique CAD with rotation symmetry is only shown once.

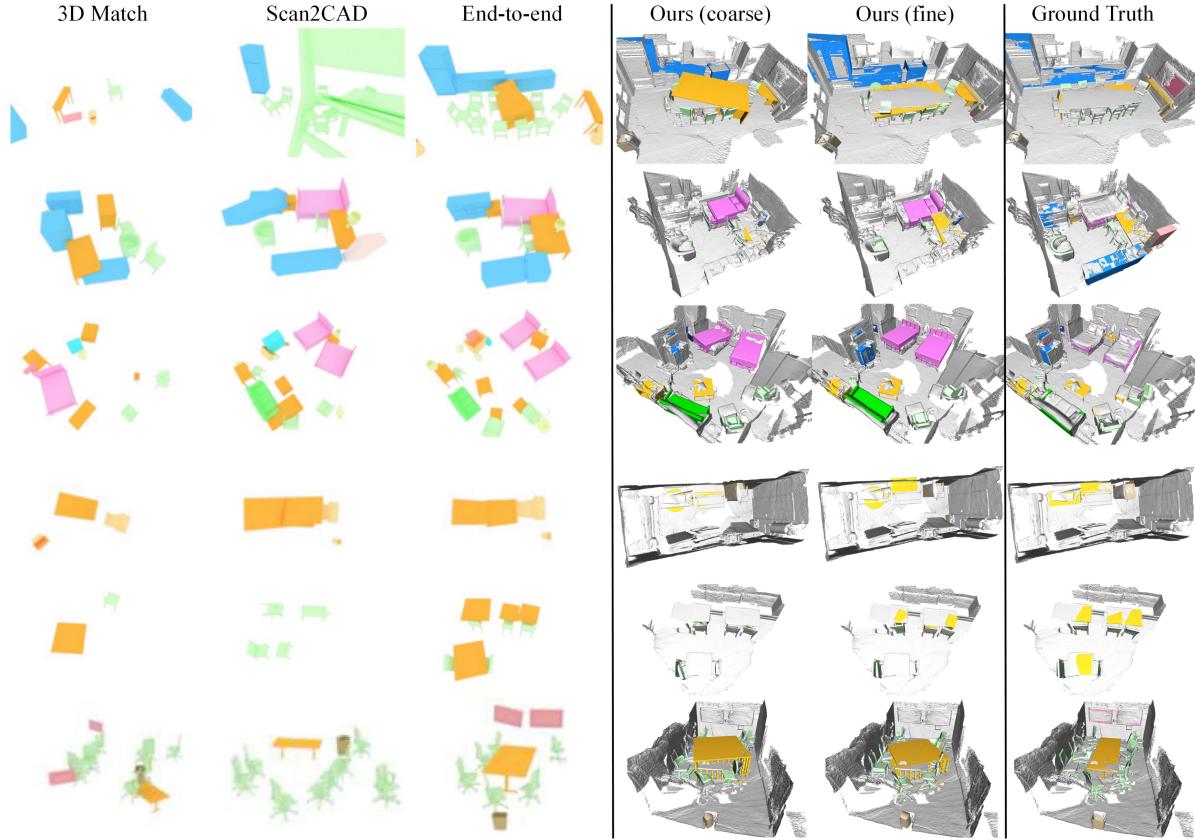


Figure 7: Qualitative comparison of CAD model alignment performance on Scan2CAD benchmark [1]. The proposed method with coarse-to-fine registration achieves more accurate and robust alignment estimation compared with 3D Match [21] and Scan2CAD [1] on various scenarios.

Method	Category based	CAD model-based								class avg.	instance avg.
		chair	table	sofa	bed	shelf	cabinet	trash bin	other		
FPFH [18]	0.14	0.18	0.02	0.07	0.00	0.00	0.00	0.02	0.03	0.04	0.08
SHOT [19]	0.07	0.06	0.02	0.07	0.09	0.00	0.01	0.00	0.03	0.05	0.04
PointNet [6]	0.49	0.43	0.13	0.09	0.61	0.23	0.04	0.38	0.07	0.23	0.29
3DCNN [17]	0.57	0.28	0.18	0.17	0.48	0.46	0.14	0.52	0.32	0.33	0.31
JointEmbedding [9]	0.68	0.55	0.32	0.33	0.42	0.19	0.26	0.50	0.43	0.39	0.43
Ours (w/o filtering)	0.53	0.51	0.20	0.20	0.17	0.13	0.11	0.31	0.10	0.22	0.34
Ours (filtered)	0.96	0.55	0.50	0.65	0.72	0.62	0.67	0.91	0.33	0.62	0.62
Ours (filtered+mix)	0.95	0.55	0.55	0.57	0.52	0.54	0.54	0.77	0.27	0.54	0.57

Table 2: Quantitative evaluation of CAD model retrieval: the proposed method with category filtering and within category negative sampling strategy performs the best on both the category-based and model-based retrieval accuracy.

Method	chair	table	sofa	shelf	cabinet	trash bin	other	class avg.	instance avg.
FPFH [18]	10.0	1.8	2.0	1.9	0.0	2.0	5.4	2.6	4.5
SHOT [19]	7.0	0.4	1.5	1.4	1.2	0.8	3.6	1.8	3.1
Li et al. [16]	14.1	3.0	6.3	1.0	1.2	1.3	1.5	3.3	6.0
3D Match [21]	21.3	1.3	7.0	5.7	2.9	4.7	10.9	6.5	10.3
Scan2CAD [1]	44.3	30.1	30.7	36.4	34.0	20.6	70.6	35.6	31.7
End2End [2]	73.0	48.2	76.9	41.5	51.5	18.2	26.8	44.6	50.7
Ours (coarse)	28.7	16.3	19.2	15.1	25.8	47.7	15.9	21.1	26.2
Ours (fine)	60.9	43.2	72.7	34.2	65.4	62.1	24.5	45.4	53.9

Table 3: CAD alignment accuracy (%) on ScanNet Scan2CAD data (Only our approach's results are achieved with the in-the-wild setup. Ground truth CAD models are not used, we instead applying retrieval from the model pool.)

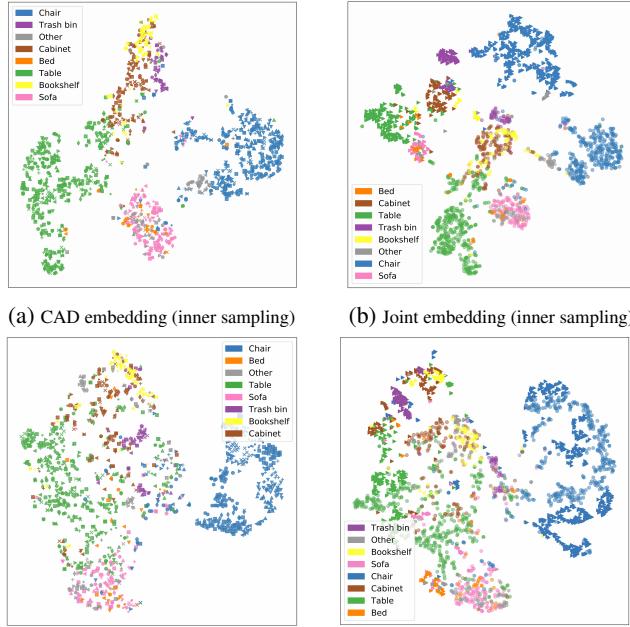


Figure 8: Comparison of the latent space of different triplet sampling strategy. For the joint embedding (b) and (d), filled triangles represent scan objects while circles represent CAD models. For the cad embedding (a) and (c), different marker represents different rotation around z-axis from the canonical pose (step=30°).

the state-of-art framework Scan2CAD [1] and End2End [2]. It should be mention that only our method is based on an in-the-wild CAD model retrieval in the scene, which casts higher difficulty for the downstream alignment. Besides, it's found that our model can work in scale on challenging large scale scenes with plentiful objects, as shown in Fig. 9.

We currently used a very lightweight classification network for model pool filtering, which only considers the ge-

ometry of scan and CAD object. For future work, more information, e.g. colour and texture can be used to enhance the semantic supervision and improve the model pool filtering performance.

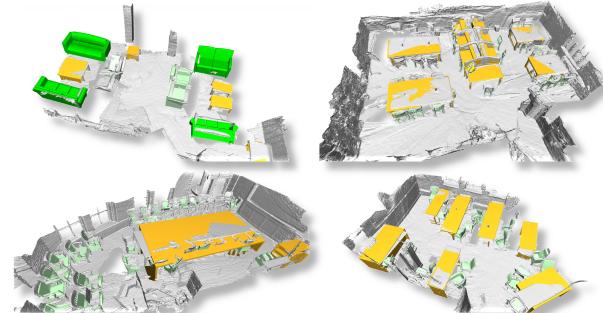


Figure 9: Examples of the proposed method on large scale challenging scenes of Scan2CAD alignment benchmark [1].

5. Conclusion

In this work, we presented Real2CAD, which retrieves and aligns a set of CAD models to 3D real scans by learning a rotation-aware joint embedding of scan and CAD object geometry. We show that by applying semantic supervision and category-based model pool filtering, our method can outperform existing state-of-the-art methods by over 19% in retrieval accuracy. In addition, our retrieval network is trained with the ability to distinguish rotation to achieve rotation-aware retrieval at inference time, thus providing helpful rotation guess for our coarse-to-fine alignment. Future works will be oriented to improve the alignment performance using semantic constraint, e.g. scene graph-based optimization. We hope our algorithm can be incorporated into an online system to work in an interactive mode and enhance the immersive experience in AR/VR environments.

References

- [1] A. Avetisyan, M. Dahnert, A. Dai, M. Savva, A. X. Chang, and M. Nießner. Scan2CAD: Learning CAD model alignment in RGB-D scans. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2614–2623, 2019. [1](#), [2](#), [3](#), [4](#), [5](#), [6](#)
- [2] A. Avetisyan, A. Dai, and M. Nießner. End-to-End CAD Model Retrieval and 9DoF Alignment in 3D Scans. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2551–2560, Oct. 2019. [1](#), [6](#)
- [3] A. Avetisyan, A. Dai, and M. Nießner. End-to-end cad model retrieval and 9dof alignment in 3d scans. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2551–2560, 2019. [2](#)
- [4] A. Avetisyan, T. Khanova, C. Choy, D. Dash, A. Dai, and M. Nießner. SceneCAD: Predicting Object Alignments and Layouts in RGB-D Scans. In *European Conference on Computer Vision*, volume 12367, pages 596–612, 2020. [2](#)
- [5] A. X. Chang, T. Funkhouser, L. Guibas, P. Hanrahan, Q. Huang, Z. Li, S. Savarese, M. Savva, S. Song, H. Su, et al. ShapeNet: An information-rich 3D model repository. *arXiv preprint arXiv:1512.03012*, pages 1–11, 2015. [2](#), [3](#), [4](#)
- [6] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 77–85, 2017. [6](#)
- [7] S. Choi, Q.-Y. Zhou, and V. Koltun. Robust reconstruction of indoor scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5556–5565, 2015. [1](#)
- [8] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 303–312, 1996. [2](#)
- [9] M. Dahnert, A. Dai, L. Guibas, and M. Nießner. Joint Embedding of 3D Scan and CAD Objects. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 8748–8757, Oct. 2019. [1](#), [2](#), [3](#), [4](#), [6](#)
- [10] A. Dai, A. X. Chang, M. Savva, M. Halber, T. Funkhouser, and M. Nießner. ScanNet: Richly-annotated 3D reconstructions of indoor scenes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5828–5839, 2017. [2](#), [3](#), [4](#)
- [11] A. Dai, M. Nießner, M. Zollhöfer, S. Izadi, and C. Theobalt. Bundlefusion: Real-time globally consistent 3d reconstruction using on-the-fly surface reintegration. *ACM Transactions on Graphics (ToG)*, 36(4):1, 2017. [1](#)
- [12] B. Drost and S. Ilic. 3d object detection and localization using multimodal point pair features. In *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, pages 9–16. IEEE, 2012. [2](#)
- [13] S. Hampali, S. Stekovic, S. D. Sarkar, C. S. Kumar, F. Fraundorfer, and V. Lepetit. Monte carlo scene search for 3d scene understanding. *arXiv preprint arXiv:2103.07969*, 2021. [2](#)
- [14] B. S. Hua, Q. T. Truong, M. K. Tran, Q. H. Pham, A. Kanezaki, T. Lee, H. Y. Chiang, W. Hsu, B. Li, Y. Lu, et al. Shrec’17: Rgb-d to cad retrieval with objectnn dataset. In *Eurographics Workshop on 3D Object Retrieval, 3DOR 2017*, pages 25–32. Eurographics Association, 2017. [2](#)
- [15] Y. M. Kim, N. J. Mitra, D.-M. Yan, and L. Guibas. Acquiring 3d indoor environments with variability and repetition. *ACM Transactions on Graphics (TOG)*, 31(6):1–11, 2012. [2](#)
- [16] Y. Li, A. Dai, L. Guibas, and M. Nießner. Database-assisted object retrieval for real-time 3d reconstruction. In *Computer Graphics Forum*, volume 34, pages 435–446. Wiley Online Library, 2015. [2](#), [6](#)
- [17] C. R. Qi, H. Su, M. Nießner, A. Dai, M. Yan, and L. Guibas. Volumetric and multi-view cnns for object classification on 3d data. In *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 2016. [6](#)
- [18] R. B. Rusu, N. Blodow, and M. Beetz. Fast point feature histograms (fpfh) for 3d registration. In *2009 IEEE International Conference on Robotics and Automation*, pages 3212–3217, 2009. [2](#), [6](#)
- [19] S. Salti, F. Tombari, and L. Di Stefano. Shot: Unique signatures of histograms for surface and texture description. *Computer Vision and Image Understanding*, 125:251–264, 2014. [6](#)
- [20] T. Whelan, S. Leutenegger, R. Salas-Moreno, B. Glocker, and A. Davison. Elasticfusion: Dense slam without a pose graph. In *Robotics: Science and Systems*. Robotics: Science and Systems, 2015. [1](#)
- [21] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, and T. Funkhouser. 3dmatch: Learning local geometric descriptors from rgb-d reconstructions. In *CVPR*, 2017. [5](#), [6](#)