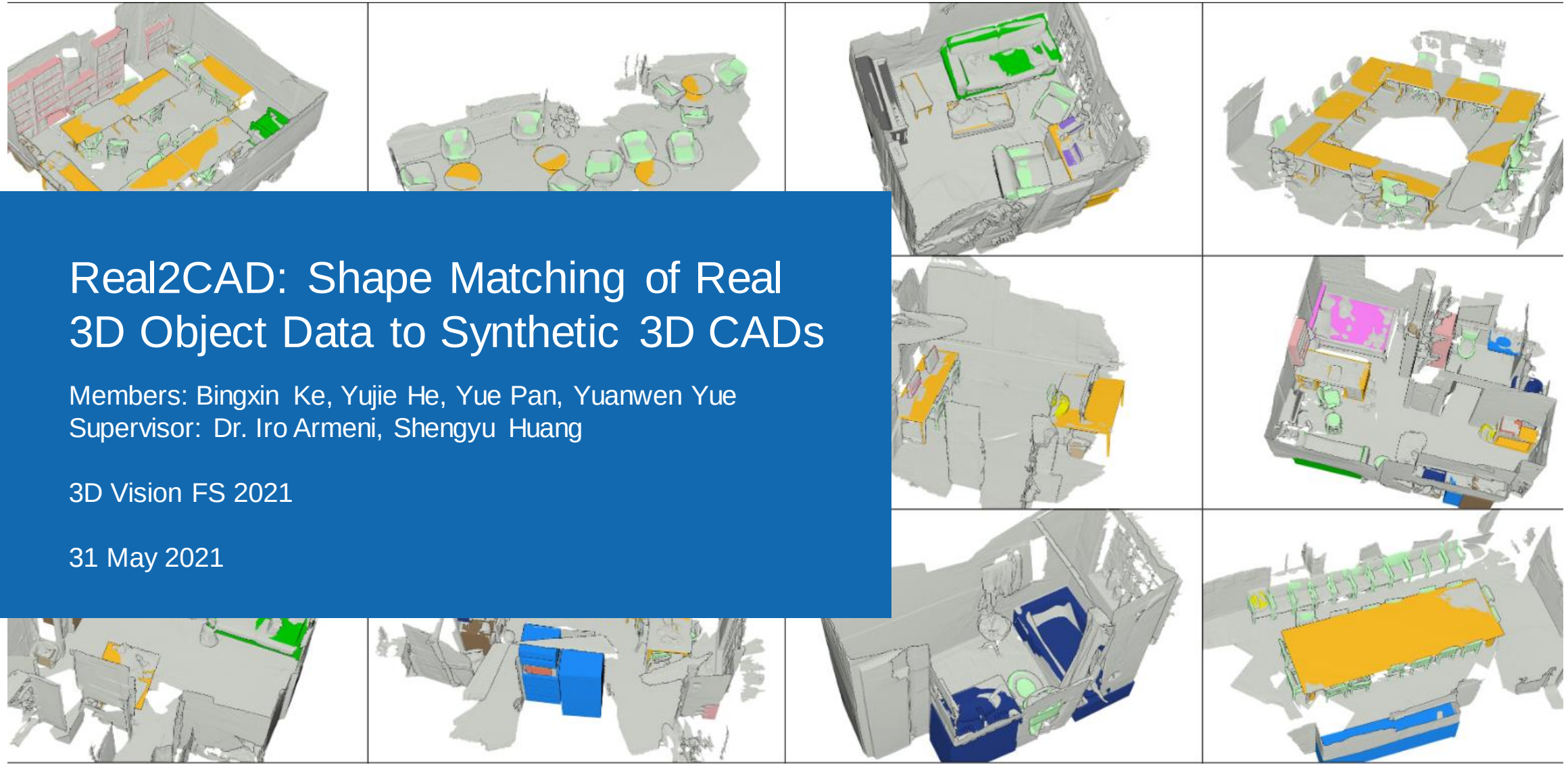
ETH zürich



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

Members: Bingxin Ke, Yujie He, Yue Pan, Yuanwen Yue
Supervisor: Dr. Iro Armeni, Shengyu Huang

3D Vision FS 2021

31 May 2021

D INFK

Introduction

Problem formation

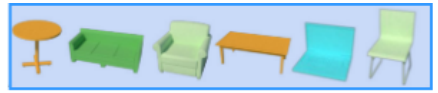
1. Segment different objects (such as chairs, tables) from 3D scans
2. Retrieve CAD models for each scan segment
3. Align CAD models to scan segments

Challenges

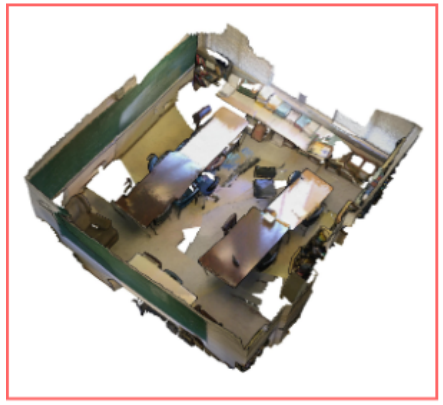
1. How to extract objects from noisy geometry in real data?
2. How to represent multi-class real object data and CAD models?
3. How to estimate 9DOF transformation from segments to CAD?

Experimental setup

- Evaluate on Scan2CAD dataset
- Compare with SOTA methods

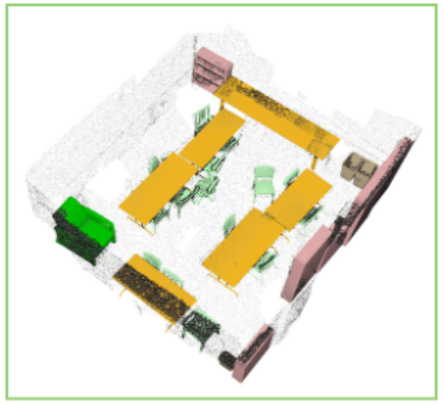


CAD pool



Input: 3D scan

Real2CAD



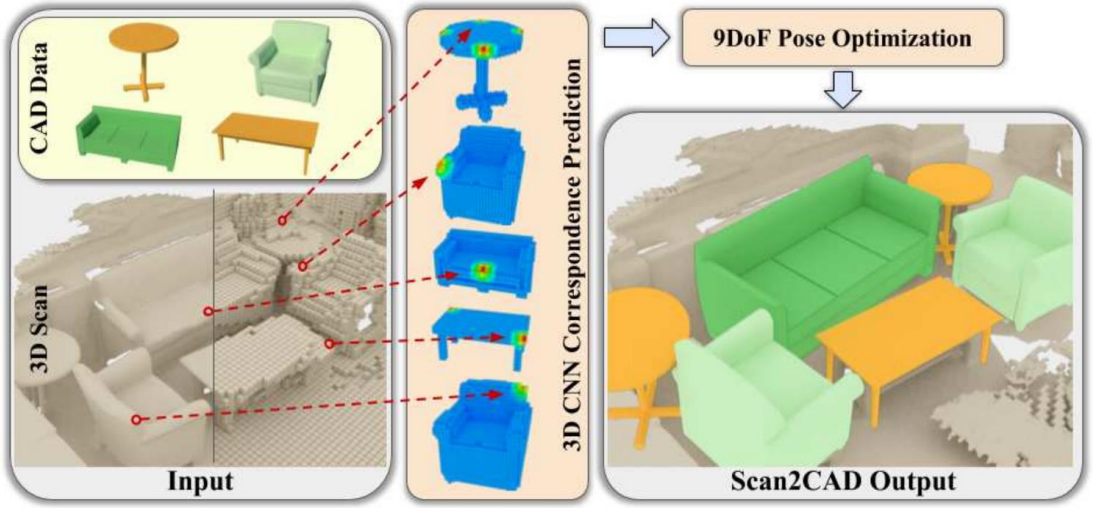
Output: CAD alignments

Progress

Month	Feb	Mar				Apr				May					Jun
Week	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Project plan proposal Literature survey on shape matching methods of real 3D object data to synthetic CADs Scan-to-CAD matching datasets exploration and preparation **Student Proposal Presentations:15 Mar**															
Re-implementation and analysis of state-of-the-art Scan-to-CAD matching pipelines Data preparation and preliminary tests on Scan2CAD datasets Experiments on one-class matching on Scan2CAD datasets **Midterm Presentations:19 April** Re-implement JointEmbedding of 3D Scan and CAD Objects Experiments and improvements of the pipelines															
Performance analysis and report writing **Project Final Presentations: 31 May**															

Related works

Scan2CAD

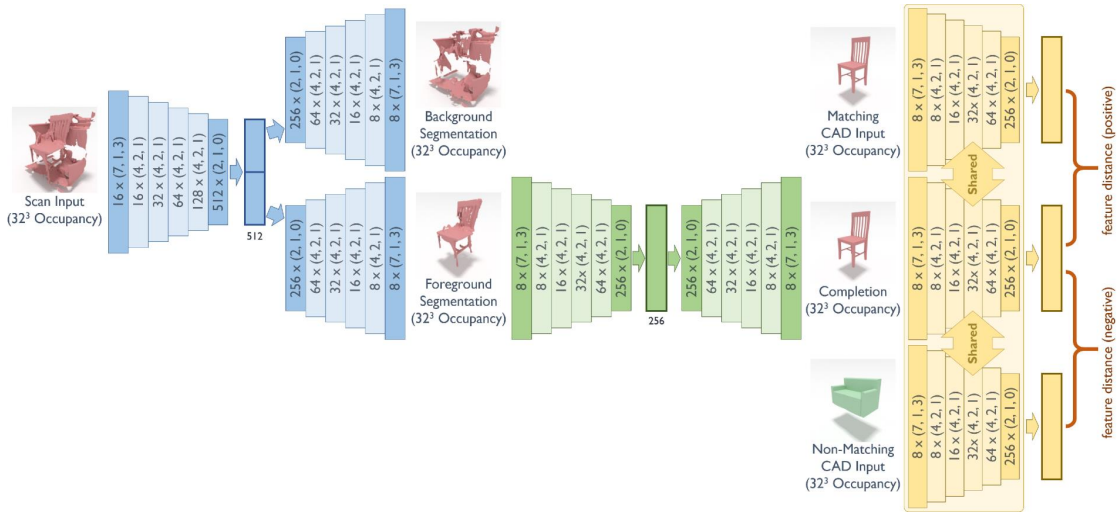


[1] Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang, and Matthias Nießner. Scan2cad: Learning cad model alignment in rgb-d scans. CVPR 2019.

+ Keypoint correspondence prediction network
+ 9DOF pose optimization (non-rigid registration)
+ Comprehensive dataset (100k keypoint correspondences between 1506 scans and 14k CAD models)

- Based on keypoints instead of scan segments
- Exhaustive searching of the corresponding CAD model (inefficient model retrieval)

Joint embedding of 3D scan and CAD objects

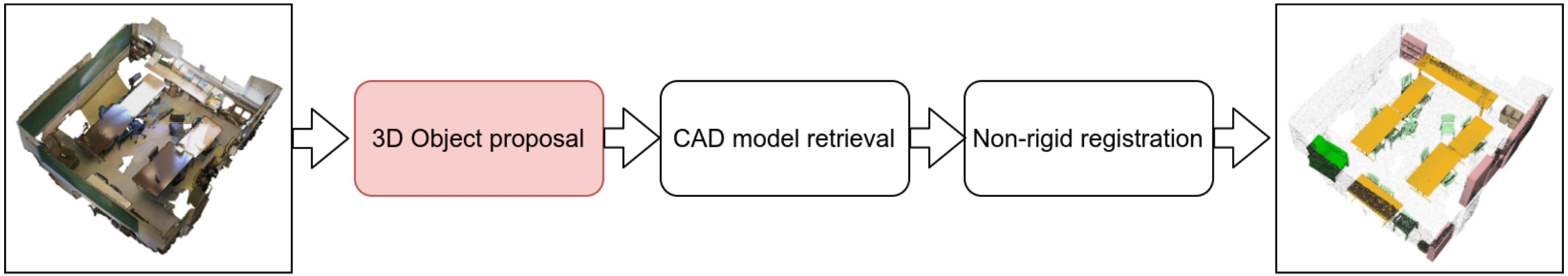


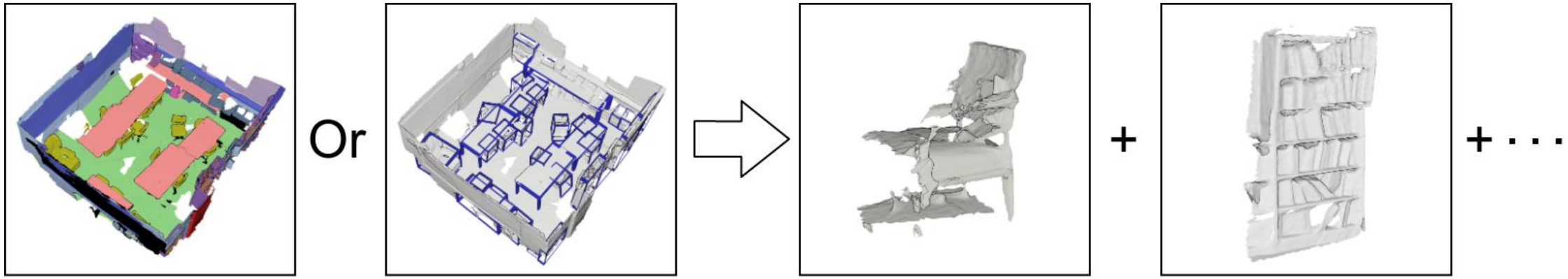
[2] Dahnert, Manuel, Angela Dai, Leonidas J. Guibas, and Matthias Niessner. Joint embedding of 3d scan and cad objects. ICCV 2019.

+ Close the gap between noisy scan data and clean CAD models
+ Enable efficient and precise model retrieval

- No CAD to scan alignment (can be only used for the model retrieval step)
- Relatively low retrieval accuracy (≈ 40%)

Real2CAD workflow

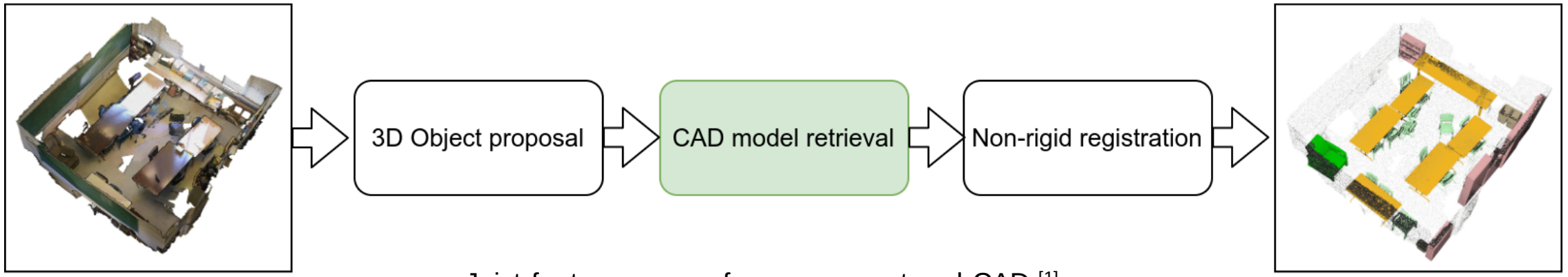




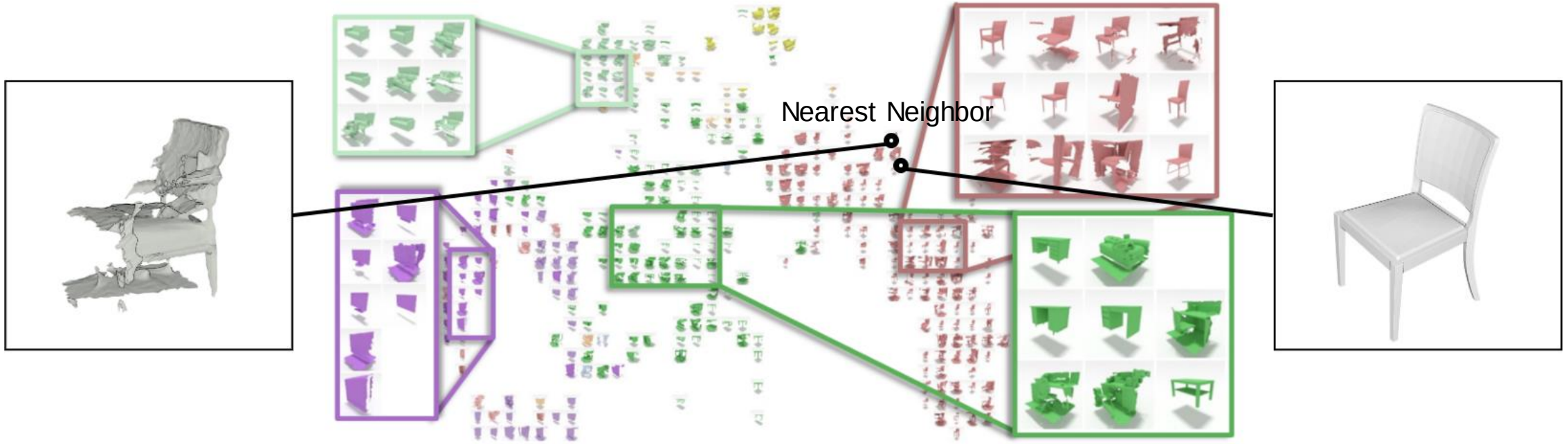
By SOTA Indoor 3D Object detection network such as 3D-SIS [1]

[1] Hou Ji and Dai Angela and Niessner Matthias. 3D-SIS: 3D Semantic Instance Segmentation of RGB-D Scans, CVPR 2019.

Real2CAD workflow

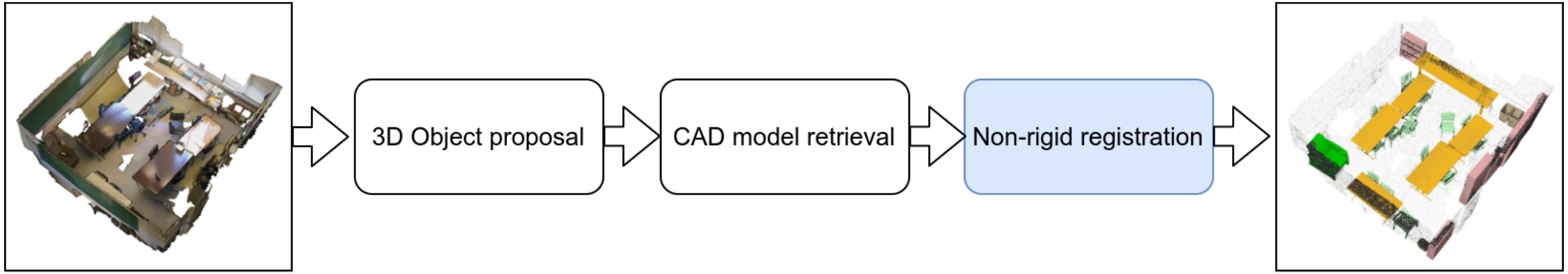


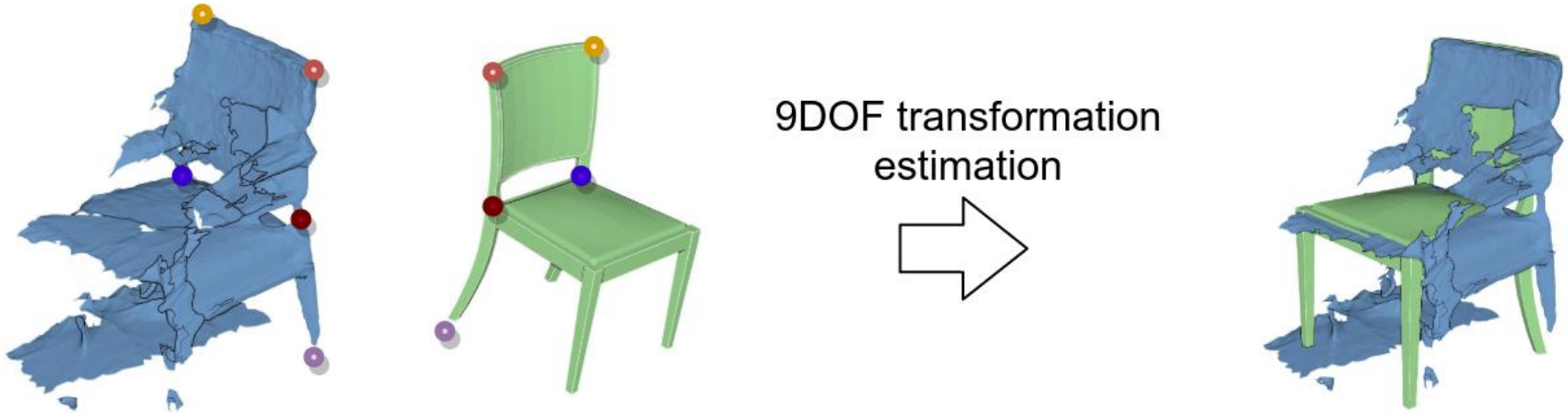
Joint feature space of scan segment and CAD [1]



[1] Dahnert, Manuel, Angela Dai, Leonidas J. Guibas, and Matthias Niessner.Joint embedding of 3d scan and cad objects. ICCV 2019.

Real2CAD workflow





Dataset

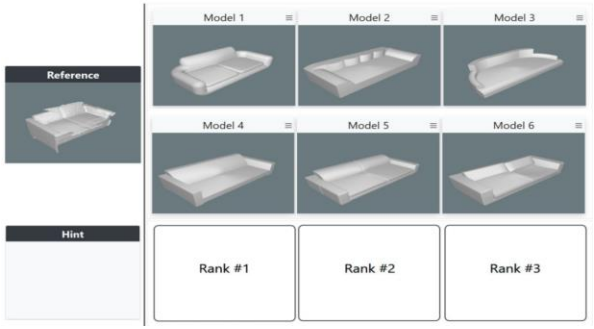
ScanNet [2]

Scan2CAD Dataset [3]

Alignment

ShapeNet [1]

Retrieval



Scan-CAD Similarity Benchmark [4]

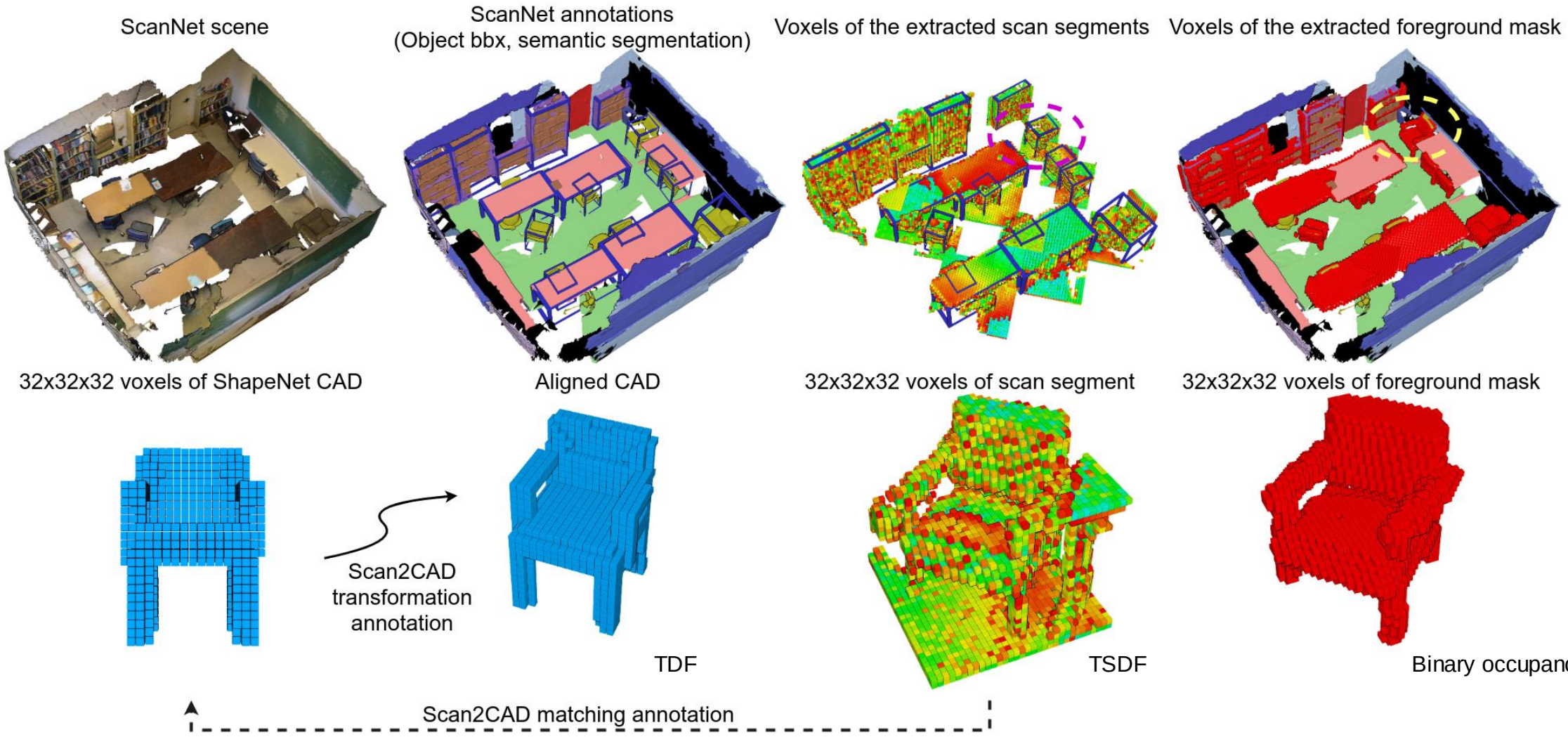
[1] Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. ShapeNet: An Information-Rich 3D Model Repository.

[2] Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. ScanNet: Richly-annotated 3D reconstructions of indoor scenes. CVPR 2017.

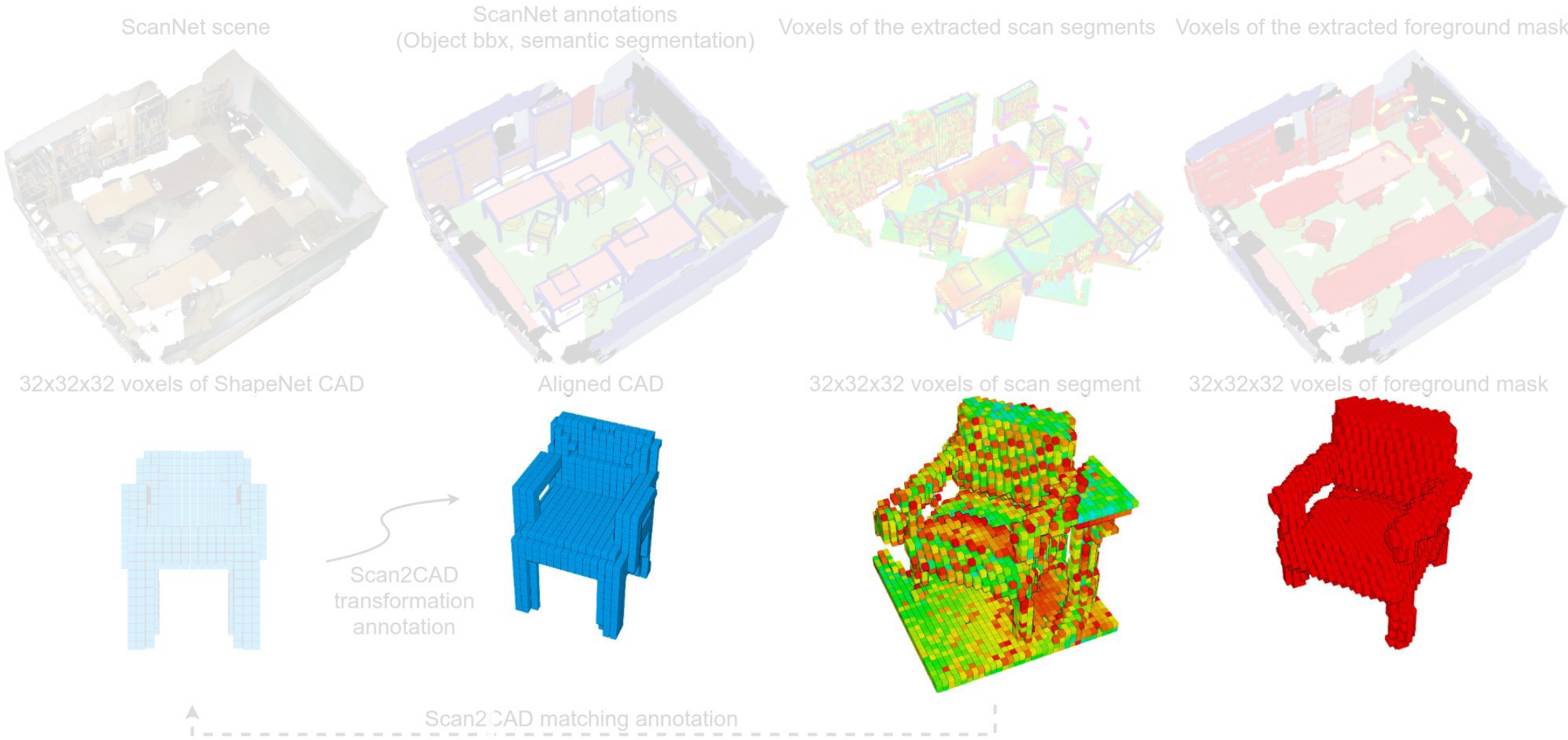
[3] Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang, and Matthias Nießner. Scan2cad: Learning cad model alignment in rgb-d scans. CVPR 2019.

[4] Dahnert, Manuel, Angela Dai, Leonidas J. Guibas, and Matthias Niessner. Joint embedding of 3d scan and cad objects. ICCV 2019.

Training data preparation



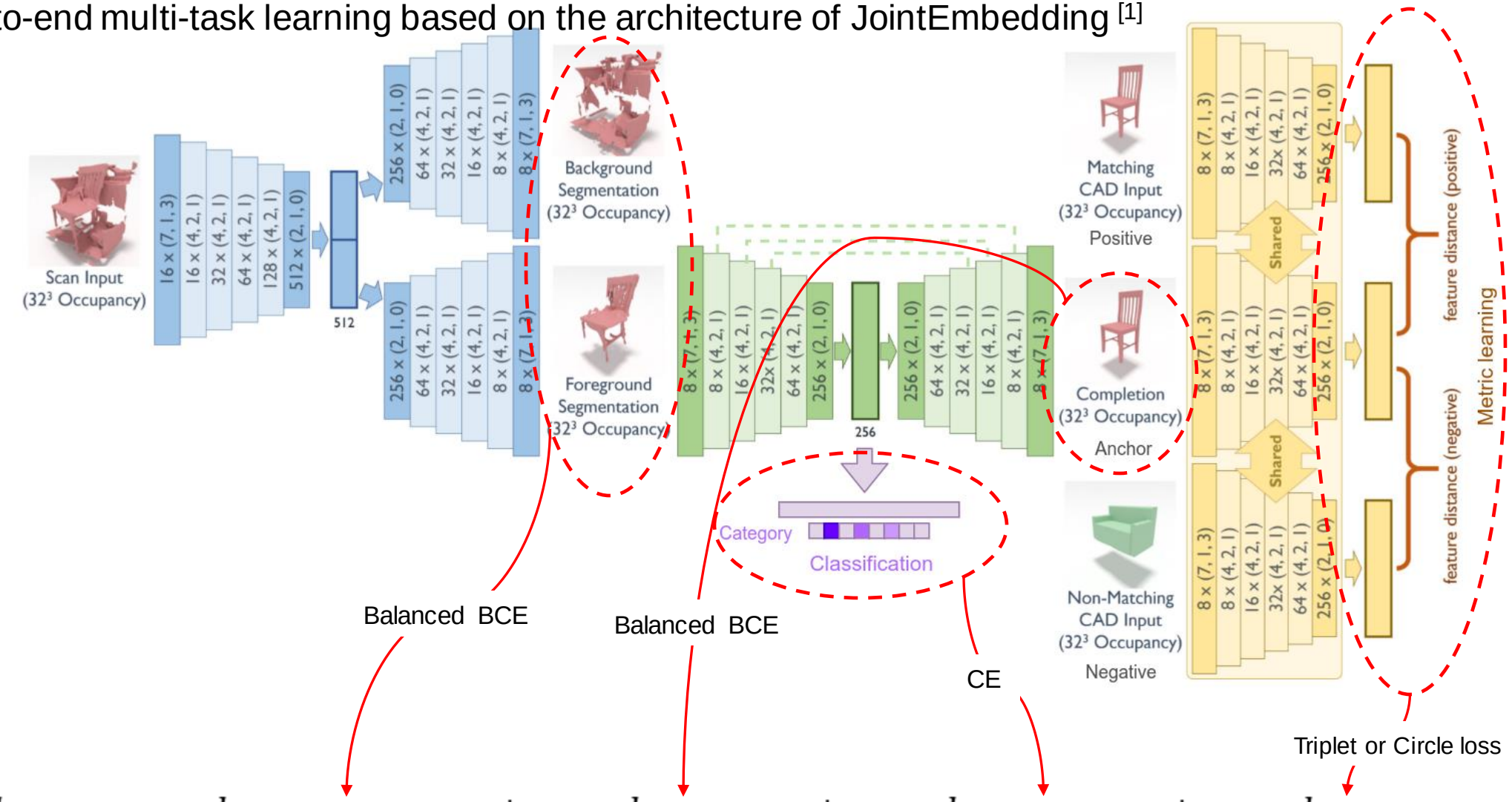
Training data preparation



In total, we get ≈ 8000 sampes for training and ≈ 2000 sampes for validation

Method: CAD Retrieval Network

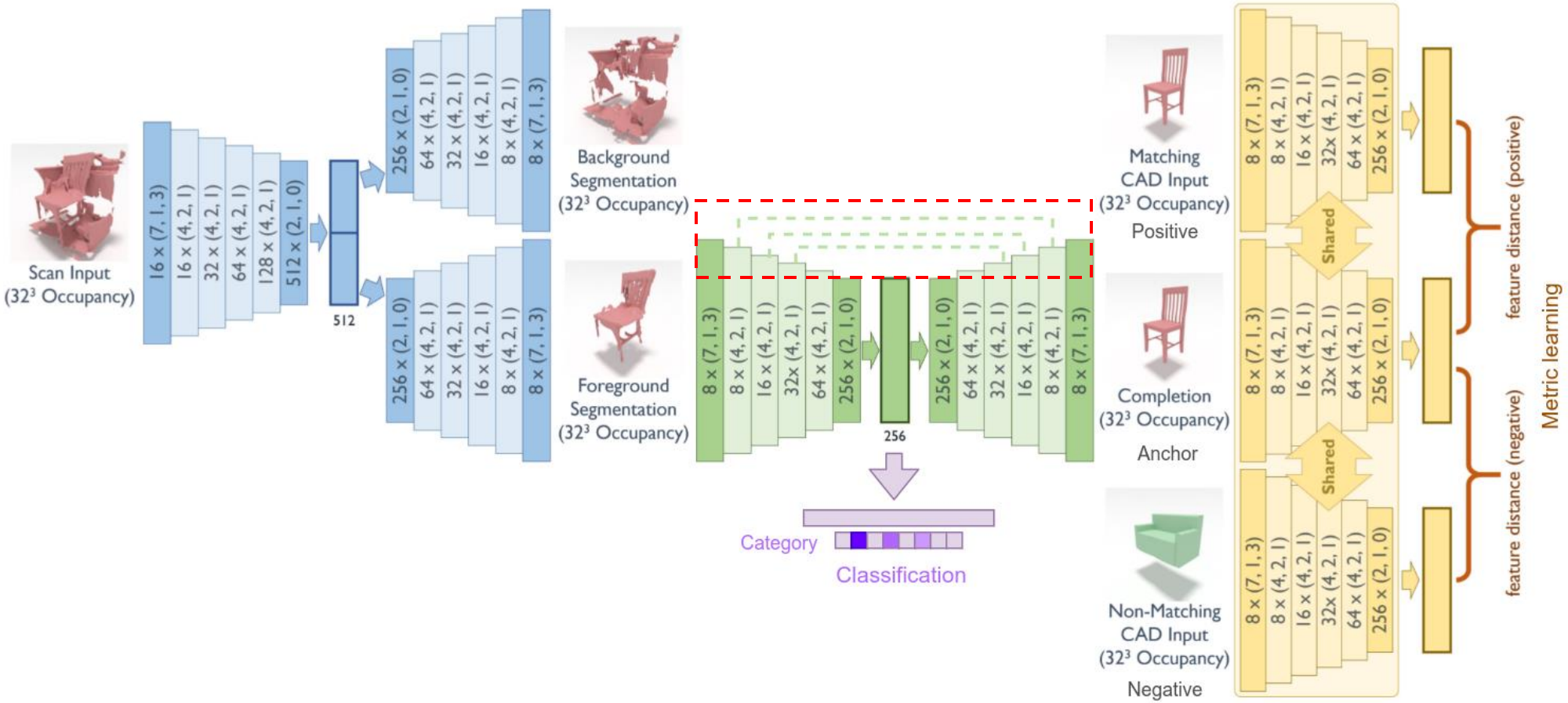
End-to-end multi-task learning based on the architecture of JointEmbedding [1]



$$loss = \omega_1 \times l_{fore/background} + \omega_2 \times l_{completion} + \omega_3 \times l_{classification} + \omega_4 \times l_{metric}$$

[1] Dahnert, Manuel, Angela Dai, Leonidas J. Guibas, and Matthias Niessner.Joint embedding of 3d scan and cad objects. ICCV 2019.

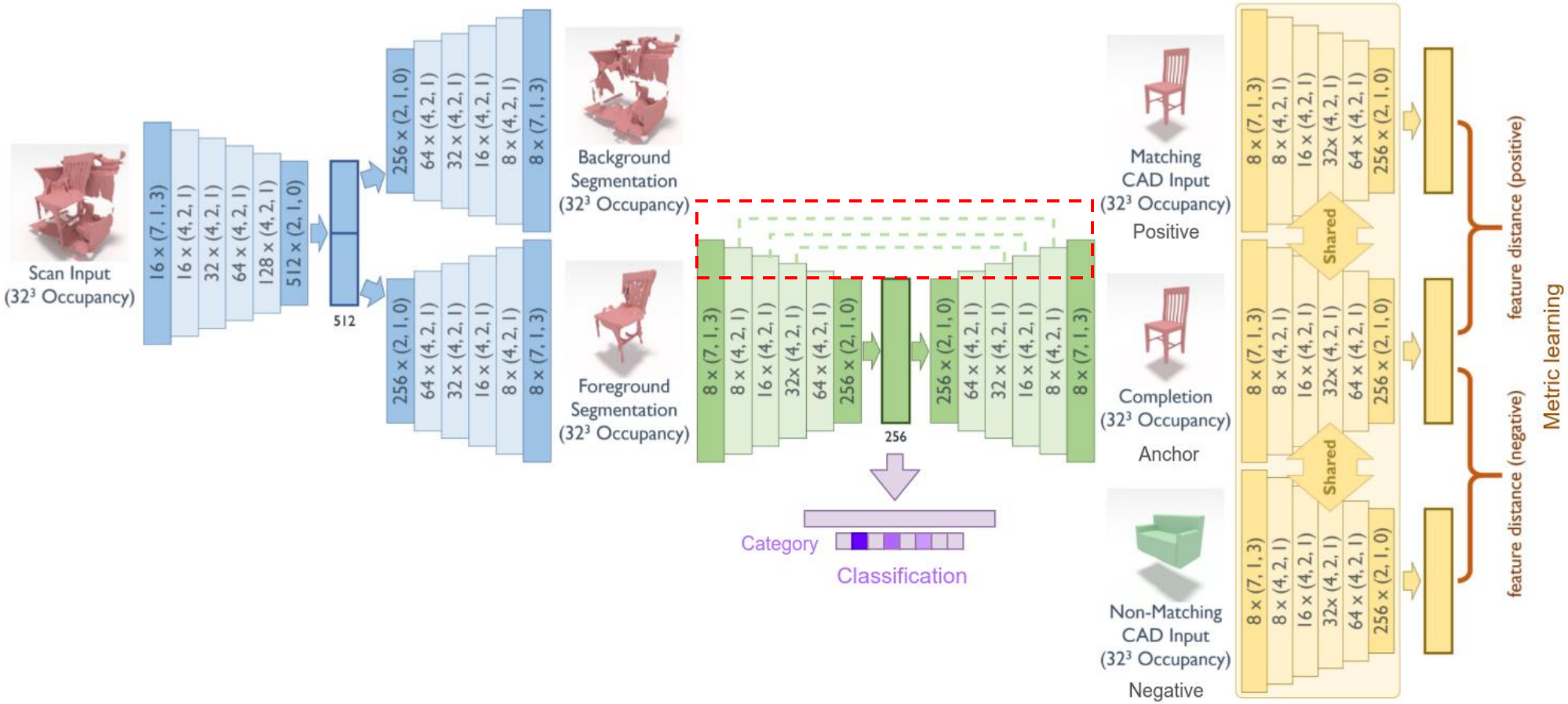
Method: CAD Retrieval Network



Our Improvements:

- Add **skip connection** to the completion encoder-decoder

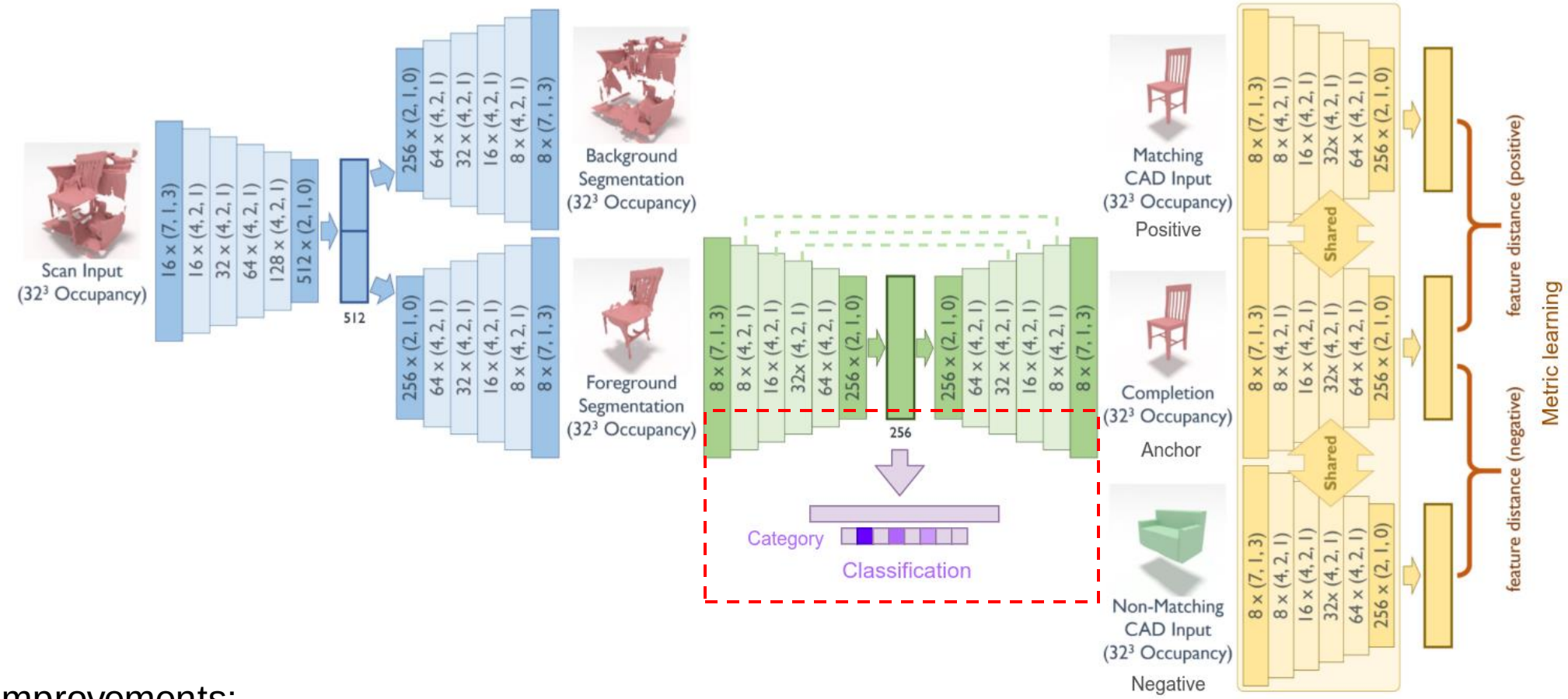
Method: CAD Retrieval Network



Our Improvements:

- Add **balanced weight** for separation and completion (binary classification problem)
- 3D object can be sparse (# occupied voxel << # non-occupied voxel)

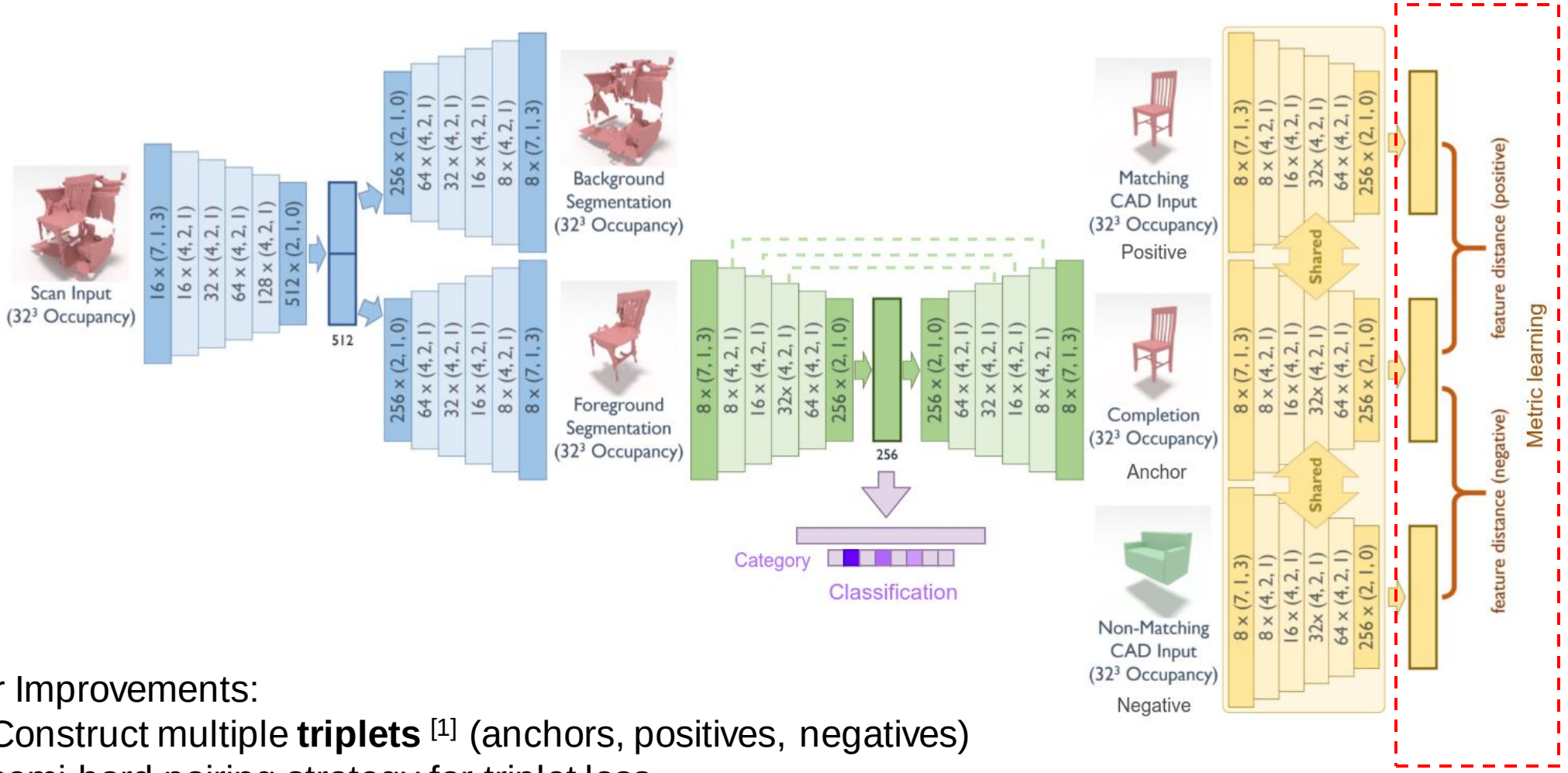
Method: CAD Retrieval Network



Our Improvements:

- Extract features from **bottleneck** of completion encoder-decoder
- Attach a **classification** module to this bottleneck.
- **Hierarchical retrieval**: filter the model pool using the classification prediction during inference

Method: CAD Retrieval Network

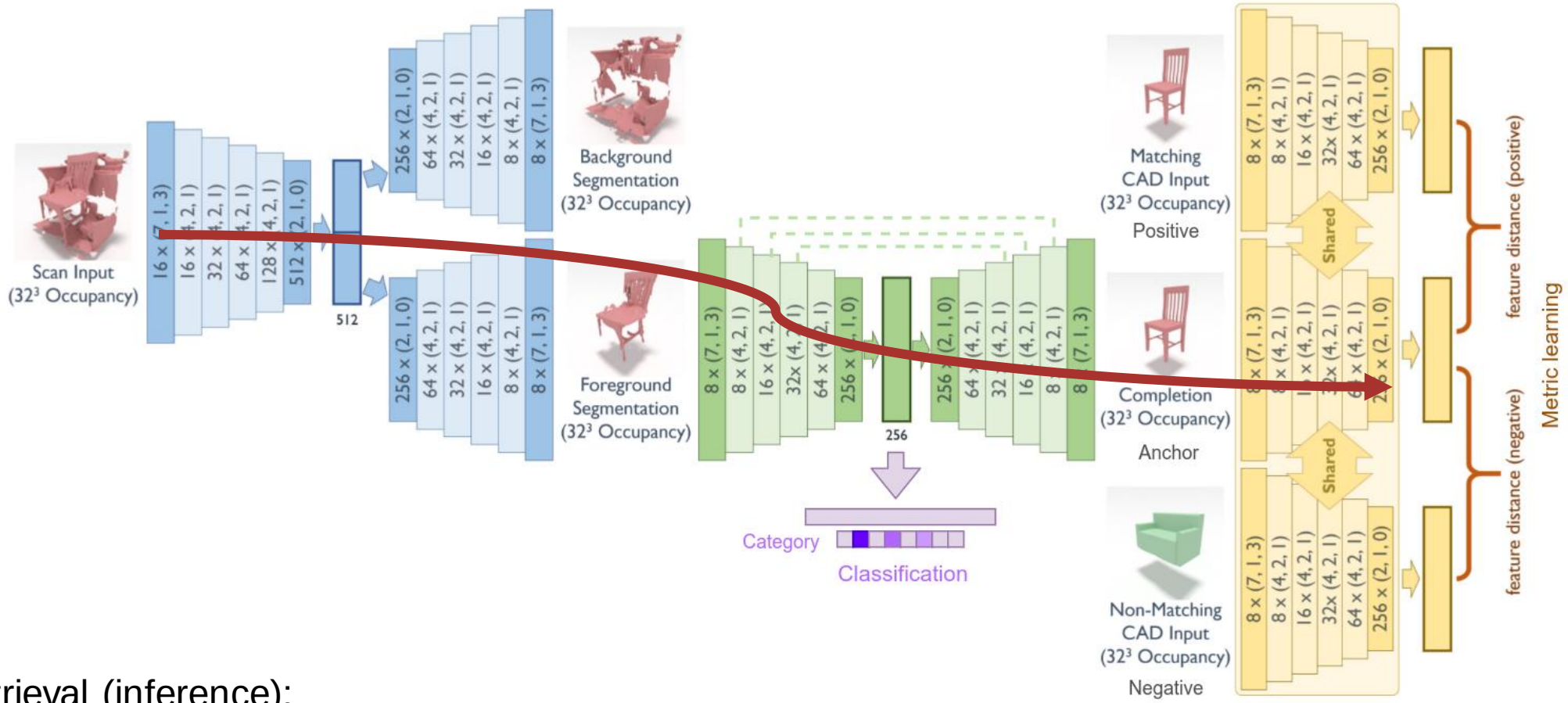


Our Improvements:

- Construct multiple **triplets** [1] (anchors, positives, negatives)
- semi-hard pairing strategy for triplet loss
- Add **circle loss** [2] as an alternative for metric learning

[1] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A unified embedding for face recognition and clustering. CVPR 2015.
[2] Yifan Sun, Changmao Cheng, Yuhan Zhang, Chi Zhang, Liang Zheng, Zhongdao Wang, and Yichen Wei. Circle loss: A unified perspective of pair similarity optimization. CVPR, 2020

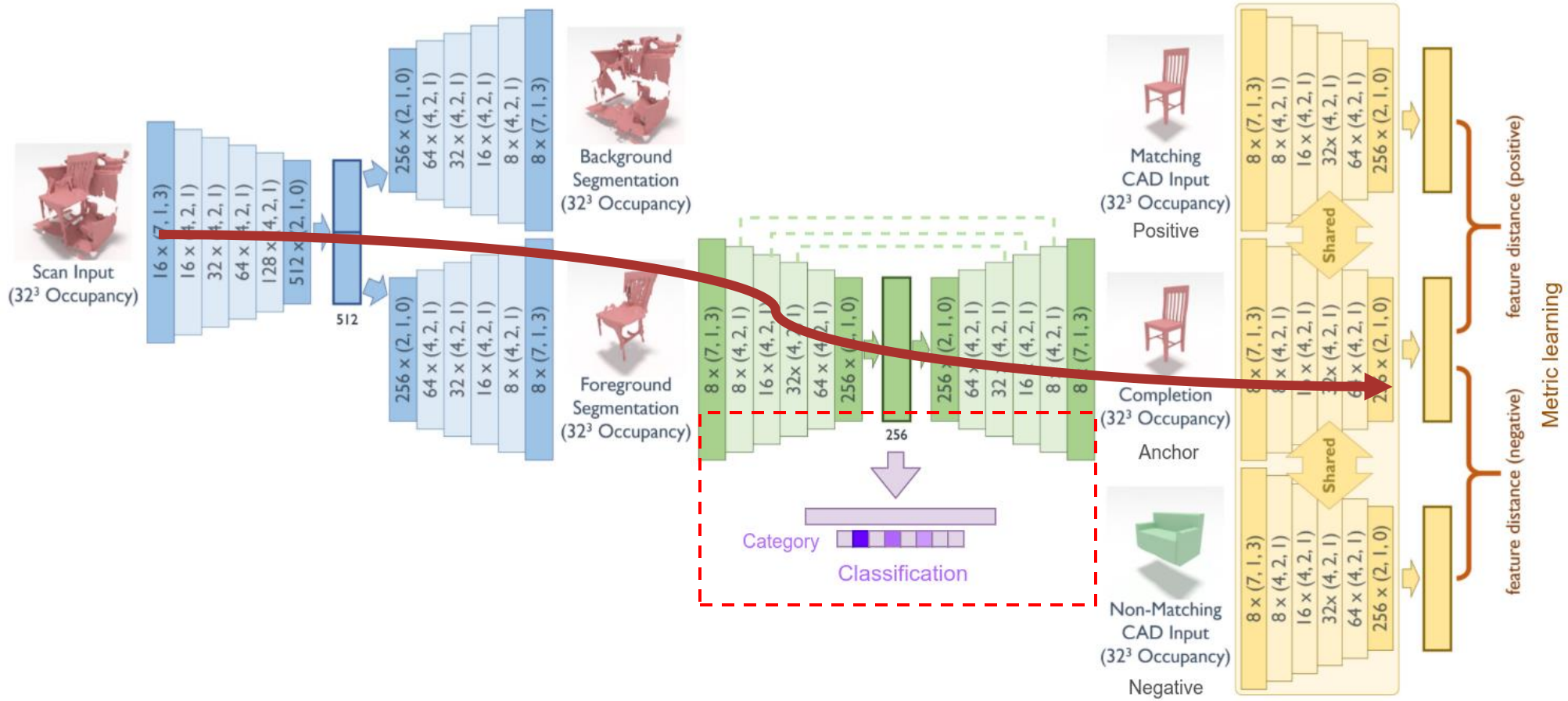
Method: CAD Retrieval Network



Retrieval (inference):

- Go through the foreground segmentation, completion and embedding network for each scan segment
- Go through the embedding network for each CAD in the model pool
- Find nearest neighbor in the joint embedding space

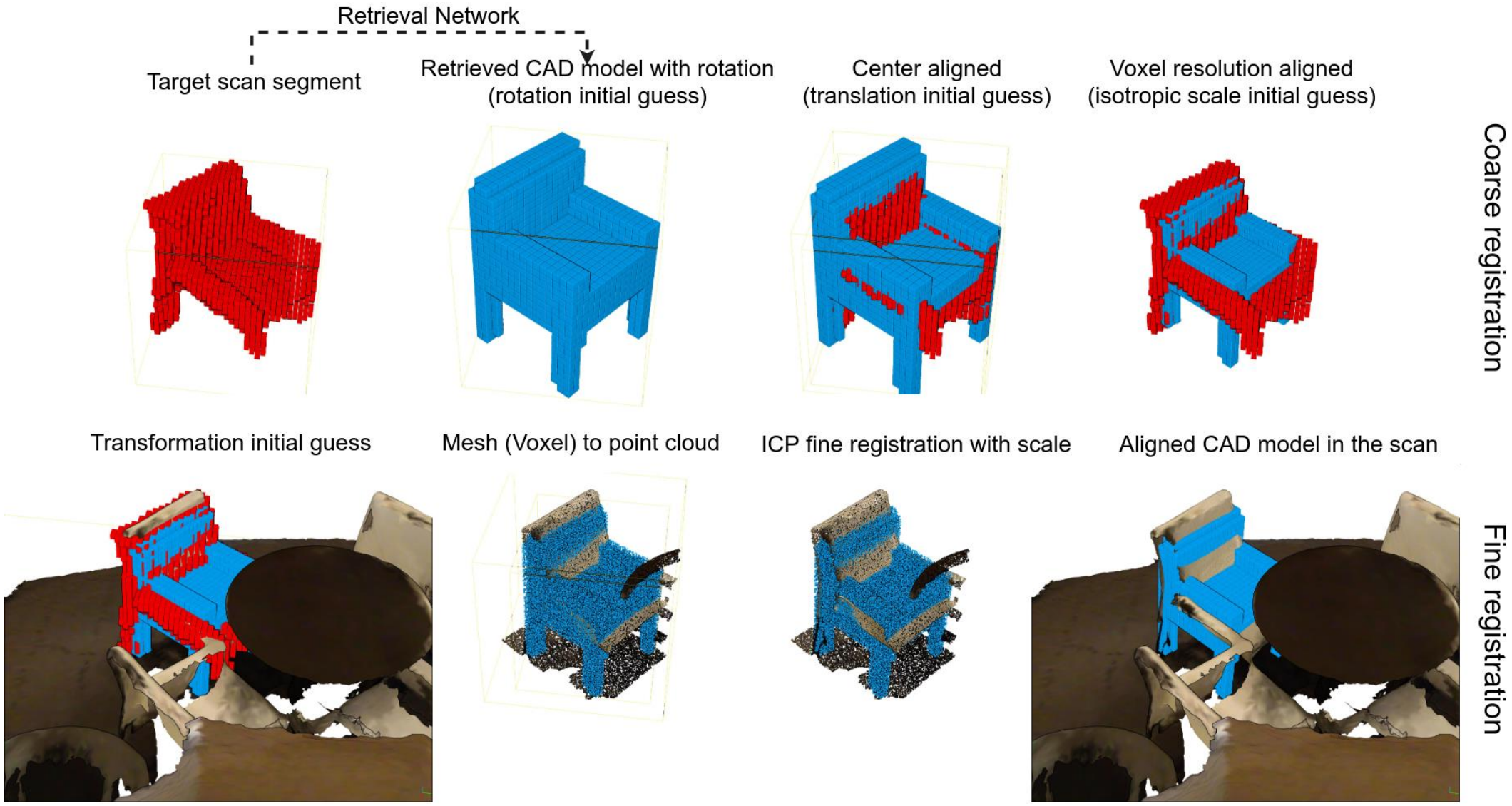
Method: CAD Retrieval Network



Hierachical retrieval:

- Filter the model pool with the predicted category of the scan segment
- Retrieve in the filtered model pool

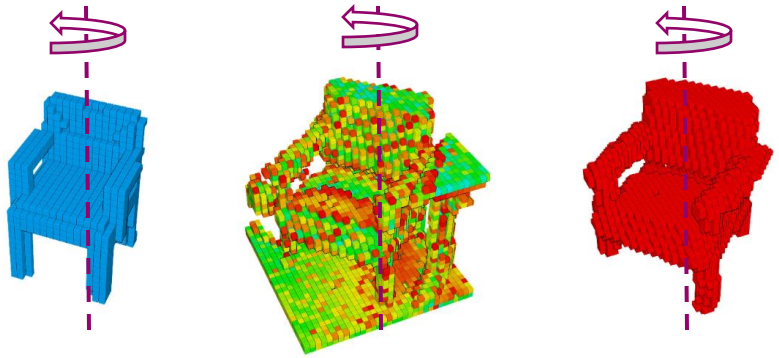
Method: Coarse-to-fine registration workflow



Implementation and Training

Weights & Biases

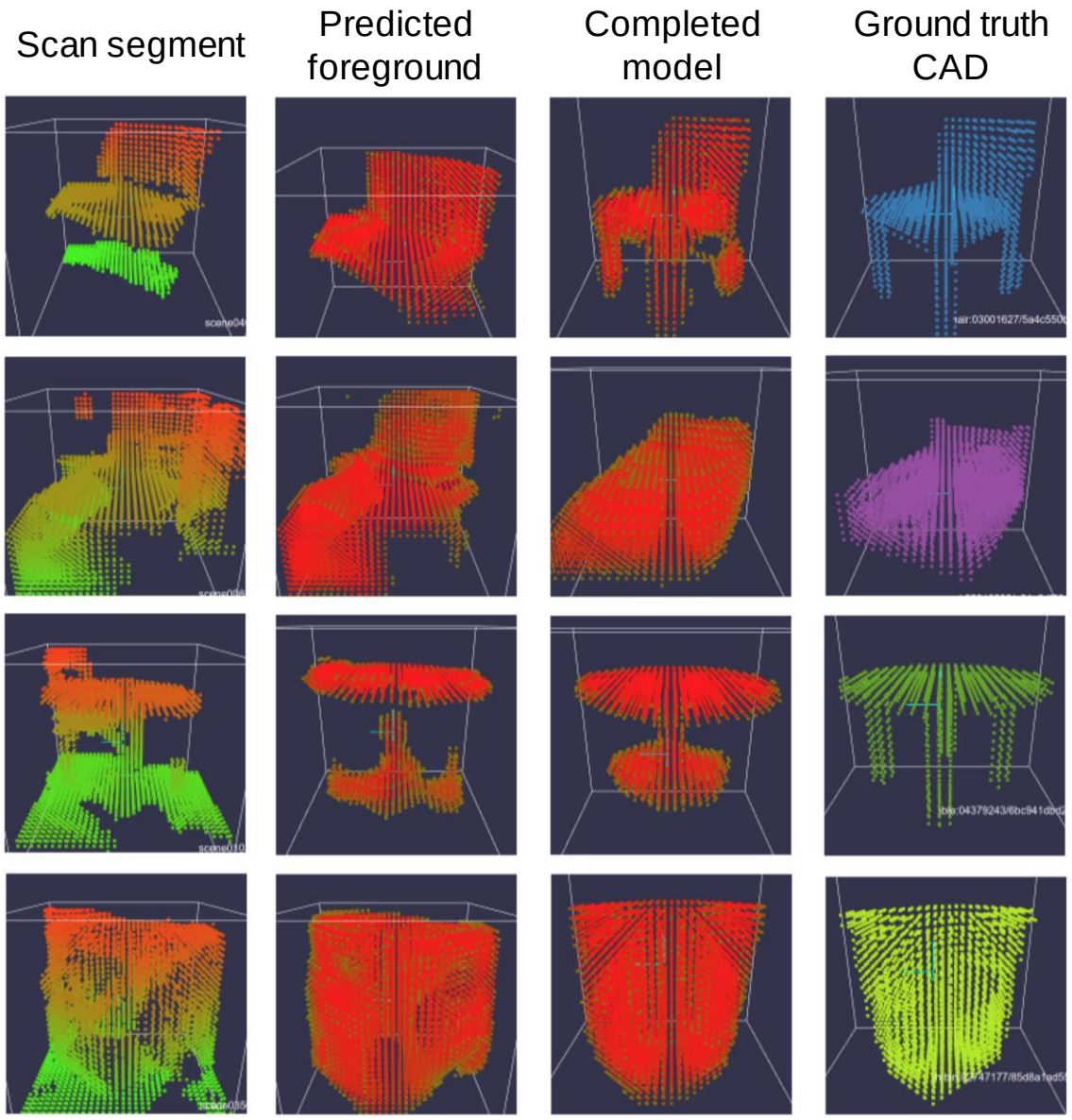
- Train with Nvidia GTX 1080Ti and 2080Ti GPU on Leonhard Cluster
- Monitoring and visualization by Weights & Biases
- Adam optimizer
- Hyperparameters Setting:
  - Learning rate (controlled by learning rate scheduler) init = 3e-4
  - Batch size = 128
  - Weight decay = 1e-4
  - Triplet margin = 0.2
- Training for 500 epochs (≈ 30k iterations, ≈ 1 day)
- Rotation augmentation around z axis (0°, 30°, 60°, …, 330°) during training [1]
- Enlarge the model pool with different rotations for rotation-aware retrieval during testing



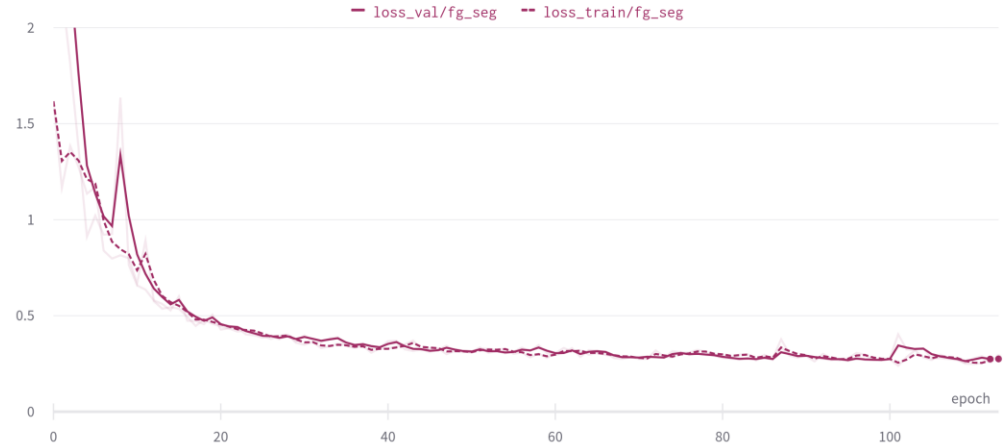
Find not only the best CAD but also the best rotation

[1] Daniel Maturana, S. Scherer, VoxNet: A 3D Convolutional Neural Network for real-time object recognition, IROS 2015.

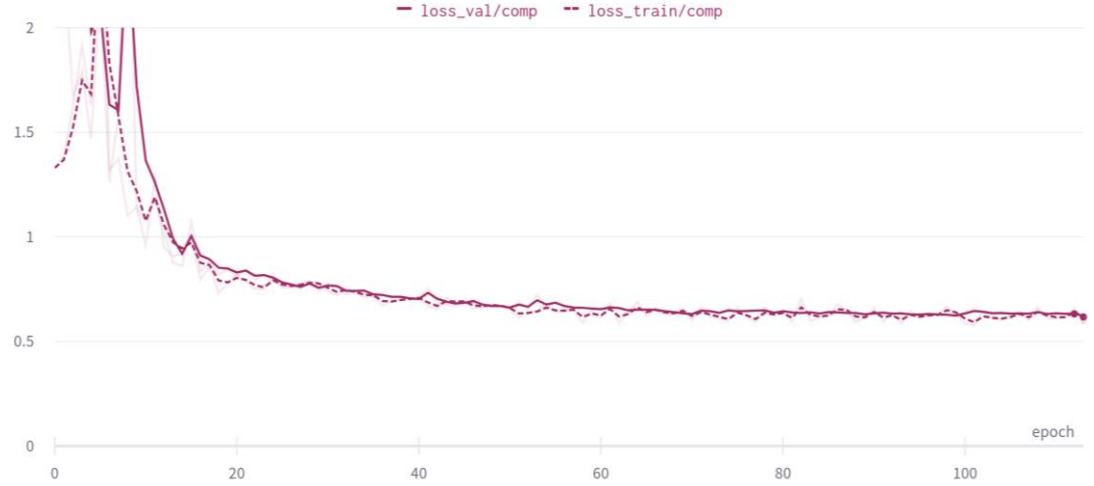
Experiment results: Sepration and completion



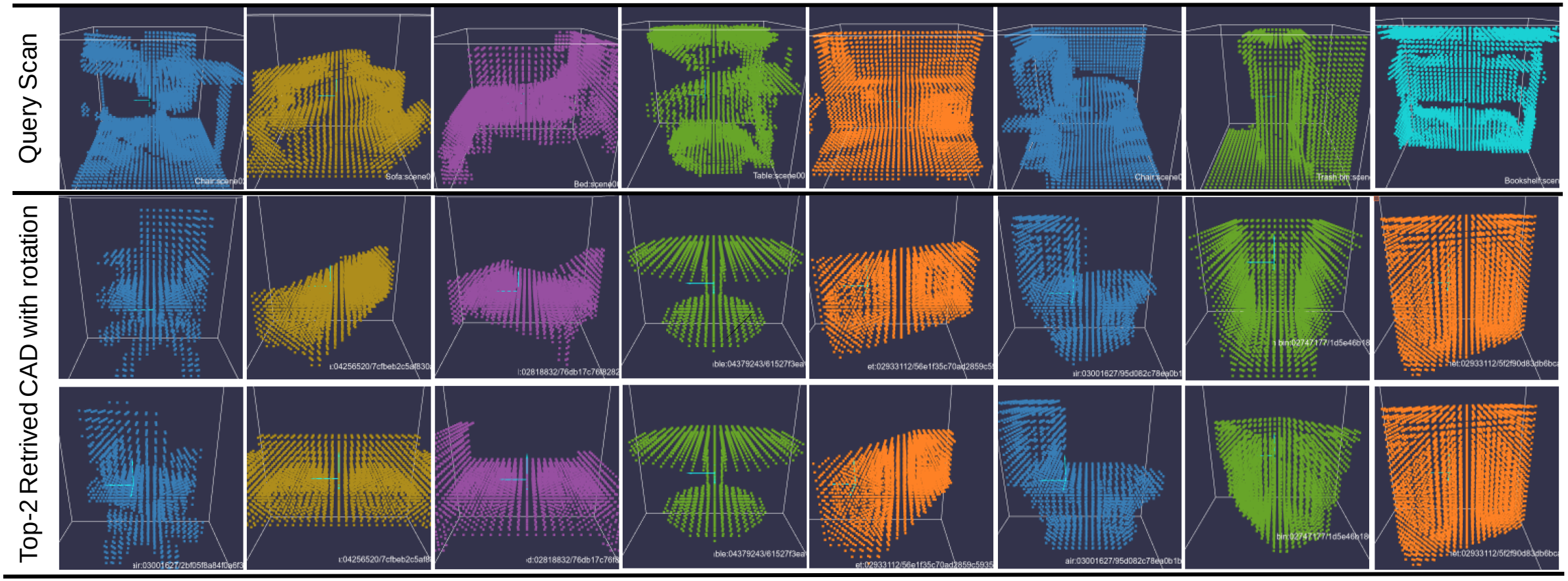
Foreground sepration loss



Completion loss

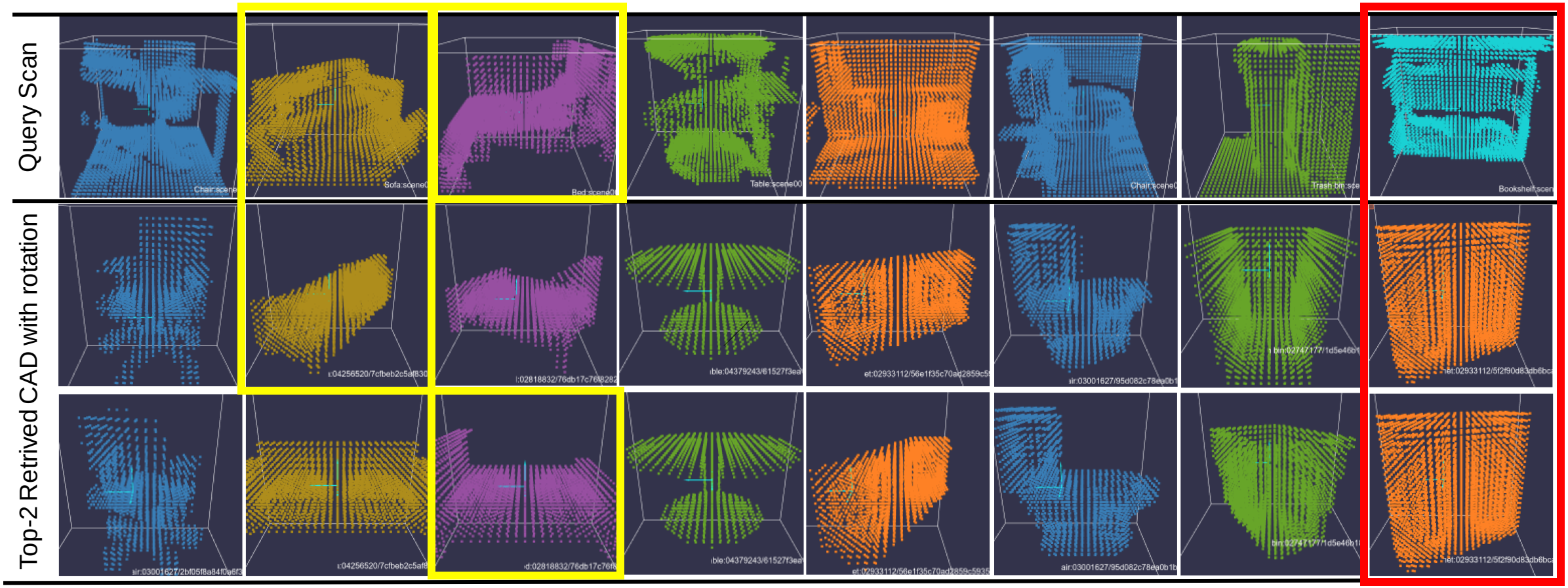


Experiment results: Fine-grained model retrieval



- High category-based retrival accuracy
- Rotation can only be retrieved roughly
- Object symmetry property: may be used to filter the model pool (redundant rotated objects)

Experiment results: Fine-grained model retrieval



Examples of the failed prediction

- Similar appearance, but different category
- Mirroring (probably due to incorrect sepration and completion)

Experiment results: Fine-grained model retrieval

Top-1 accuracy for category-based and CAD model-based retrieval on Scan-CAD similarity benchmark (# model pool = 100)

Method	category-based	CAD model-based									
		chair	table	sofa	bed	bookshelf	cabinet	trash bin	other	class avg	inst avg
FPFH [1]	0.14	0.18	0.02	0.07	0.00	0.00	0.00	0.02	0.03	0.04	0.08
SHOT [2]	0.07	0.06	0.02	0.07	0.09	0.00	0.01	0.00	0.03	0.05	0.04
PointNet [3]	0.49	0.43	0.13	0.09	0.61	0.23	0.04	0.38	0.07	0.23	0.29
3DCNN [4]	0.57	0.28	0.18	0.17	0.48	**0.46**	0.14	0.52	0.32	0.33	0.31
JointEmbedding [5]	0.68	0.55	0.32	0.33	0.42	0.19	0.26	0.50	**0.43**	0.39	0.43
Ours	0.81 +0.13	0.61	0.34	0.71	0.66	0.25	0.30	0.81	0.25	0.48	0.54 +0.11
Ours (hierachical)	**0.97** +0.29	**0.65**	**0.41**	**0.74**	**0.78**	0.42	**0.42**	**0.85**	0.32	**0.54**	**0.59** +0.16

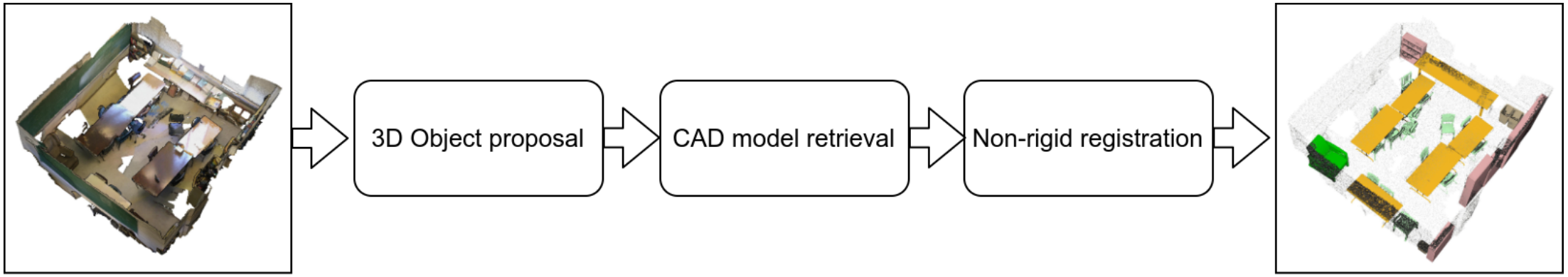
- The proposed model outperformed SOTA (JointEmbedding) by a large margin
  - By online mining the hard triplet pair
  - By using skip connection and balanced cross entropy for the sepration and completion
  - By adding a classification module to get more semantic feature involved
- With the hierarchical retrieval trick, the performance further improves
  - Our classification module has high accuracy
  - Model pool is relatively small

[1] Radu Bogdan Rusu, Nico Blodow , and Michael Beetz. Fast point feature histograms (fpfh) for 3d registration. ICRA 2009.
[2] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. ECCV 2010.
[3] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation, CVPR 2017.
[4] Charles Ruizhongtai Qi, Hao Su, Matthias Nießner, Angela Dai, Mengyuan Yan, and Leonidas Guibas. Volumetric and multi-view cnns for object classification on 3d data. CVPR 2016.
[5] Dahnert, Manuel, Angela Dai, Leonidas J. Guibas, and Matthias Niessner.Joint embedding of 3d scan and cad objects. ICCV 2019.

Experiment results: Scan2CAD alignment

- Experiments ongoing
- To be added in the final report

Conclusion

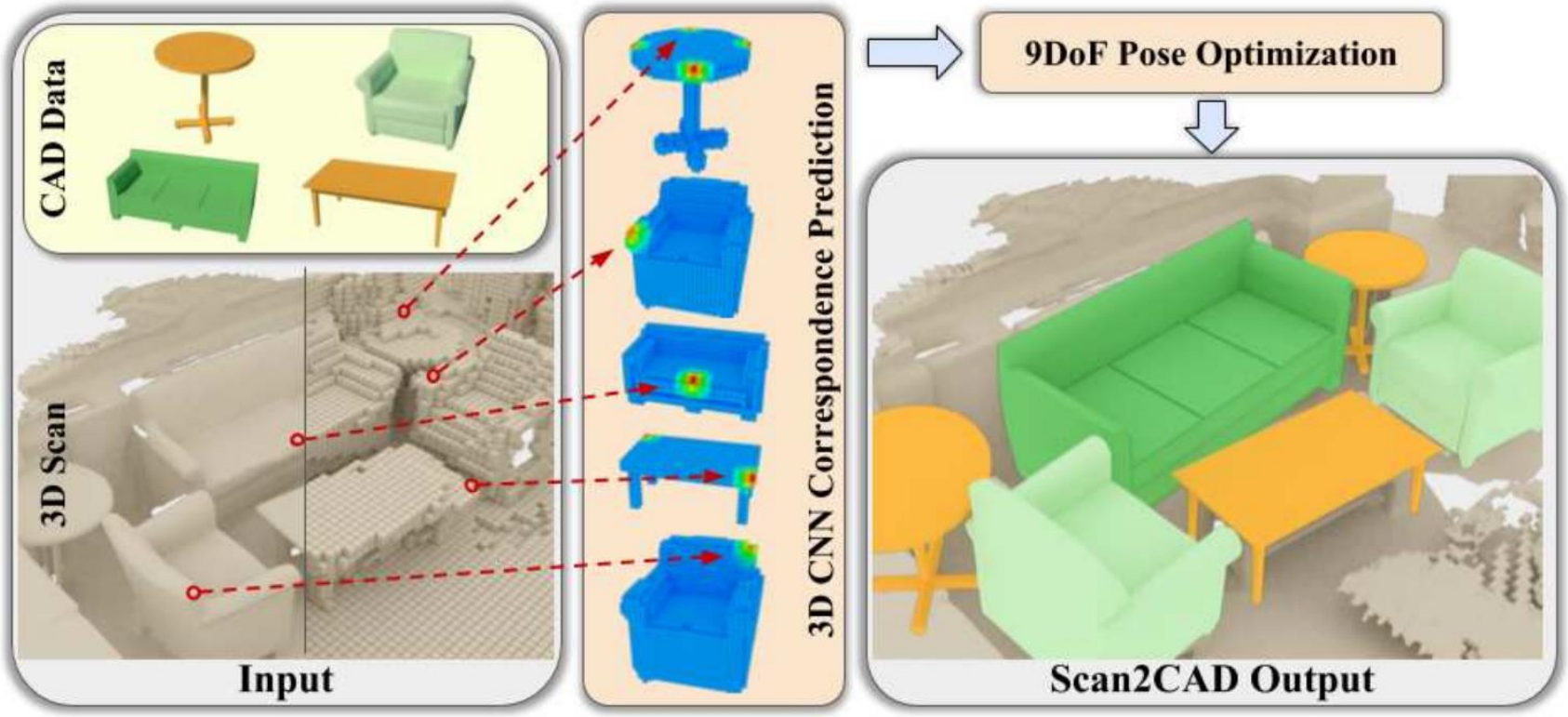


We propose the Real2CAD framework with:

- End-to-end multi-task learning for CAD model retrieval
  - Outperform SOTA by 16% on model-based retrival accuracy (29% on category-based accuracy)
  - Retrieve with relative rotation
- Coarse-to-fine registration
  - Coarse registration by rotation retrieval and voxel alignement
  - Fine registration by ICP with scale

Thanks for your attention !

Related works: Scan2CAD

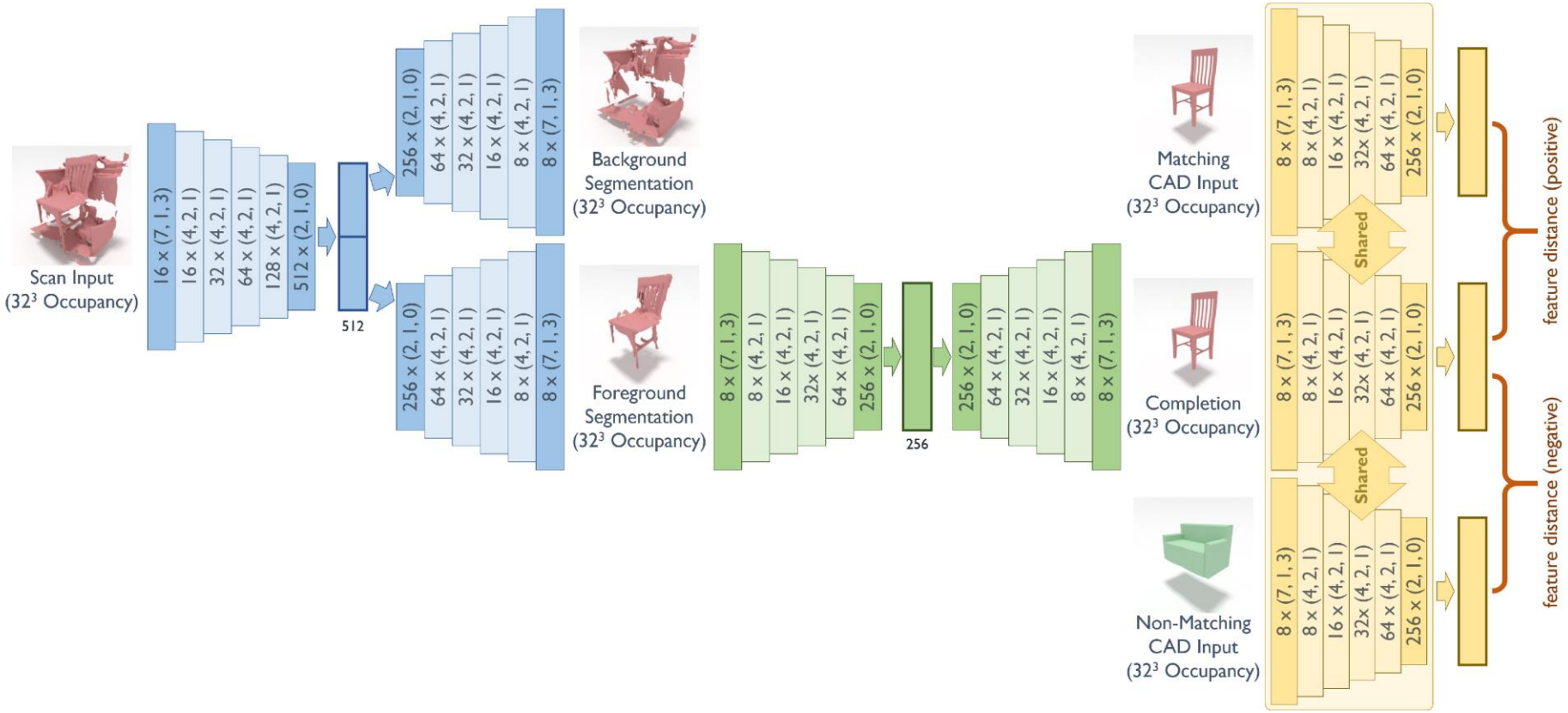


[1] Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang, and Matthias Nießner. Scan2cad: Learning cad model alignment in rgb-d scans. CVPR 2019.

+ Comprehensive dataset (100k keypoint correspondences between 1506 scans and 14k CAD models)
+ Keypoint correspondence prediction network + 9DOF pose optimization (non-rigid registration)

- Based on keypoints instead of scan segments
- Exhaustive searching of the corresponding CAD model (inefficient model retrieval)

Related works: Joint embedding of 3D scan and CAD objects



[1] Dahnert, Manuel, Angela Dai, Leonidas J. Guibas, and Matthias Niessner.Joint embedding of 3d scan and cad objects. ICCV 2019.

+ Close the gap between noisy scan data and clean CAD models

+ Enable efficient and precise model retrieval

- No CAD to scan alignment (can be only used for the model retrieval step)

- Relatively low retrieval accuracy (≈ 40%)