1 Implementation Details

Our work is implemented in Tensorflow. The batch size is set as 16. The learning rates of the reconstruction network and PCLossNet are set as 0.0001 and 0.005, respectively. They are both optimized with the Adam optimizer. The specific settings of all hyper-parameters are illustrated in Table 1, while structures of networks are presented in Table 2. All experiments are conducted on a NVIDIA 2080ti GPU with a 2.9GHZ i5-9400 CPU.

NAME	σ	ϵ	ϵ_1	N_c
CONSTANTS	0.01	0.01	0.2	128

Table 1. Illustrations of hyper-parameters

RECNET	ENCODER	DECODER
FC	MLPs(64,128,128,256,128)+MAX-POOLING	FCs(256,256,2048*3)
FOLDING	MLPs(64,128,128,256,128)+MAX-POOLING	MLPs(256,256,3) + MLPs(256,256,3)
PCLOSSNET	PARAMETER-SHARED MLPS	OUTPUT LAYERS
LNSA	MLPs(64,128)	MLPs(256,256) + MLPs(128,64,3)
LNFC	MLPs(64,128)	FCs(256,256,256,128*3)

Table 2. Illustrations of network structures. All components presented are MLPs.

2 Comparison with Pre-trained Networks

Except our proposed PCLossNet digging shapes differences with a generative-adversarial process, constraining with a pre-trained feature extraction network following [2] is also an alternative for the training of reconstruction network.

In this section, we take the encoders parts of pre-trained AE [1] and PointNet++ [3] to extract features from input and reconstructed point clouds, then training the reconstruction network by the distances between extracted features. Results are presented in Fig. 1 and Table 3. We compare Pre-trained AE (PreAE) and Pre-trained PointNet++ (PrePN+) with our proposed LNFC and LNSA.

We can see that reconstruction networks trained with PreAE and PrePN+ cannot reconstruct the input point clouds well. The reason may be that the non-linear feature spaces is still ambiguous to point clouds on 3D Euclidean space, which is similar as the reason why Point-based GAN discriminators cannot work without matching-based losses analyzed in Sec.3.1 and Sec.4.3 of the paper.

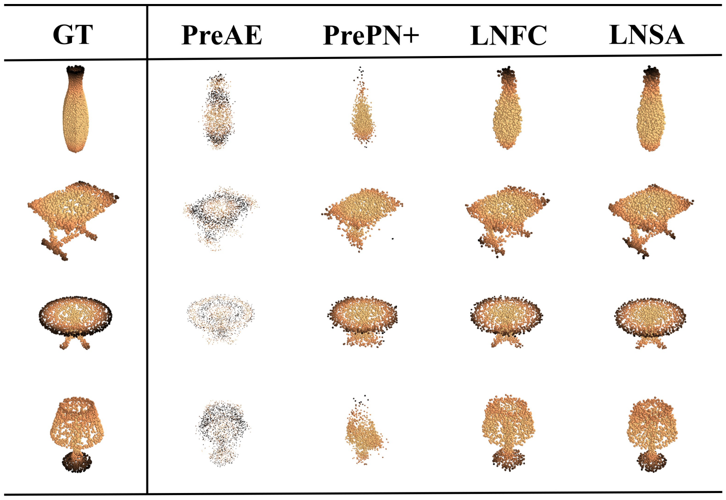

Fig. 1. Reconstruction results trained with pre-trained feature extraction networks.

An interesting thing is that the pre-trained losses perform a little better than discriminators without matching compared in Table 3, which maybe because the pre-trained processes constrain the feature spaces and reduce the ambiguity. Besides, the pre-trained method needs a separate training process before training the reconstruction network. it is very inconvenient, while the pre-training process is still limited by the matching-based losses.

Our method can fully get rid of the influence from matching processes, which can dynamically be trained together with the reconstruction network and do not need any pre-training.

DataSet	ShapeNet				ModelNet40			
Methods	PreAE	PrePN+	LNFC	LNSA	PreAE	PrePN+	LNFC	LNSA
Pre-train	✓	✓	✗	✗	✓	✓	✗	✗
MCD	67.15	2.62	0.23	**0.23**	80.43	4.53	**0.58**	0.59
HD	240.53	25.71	1.66	**1.66**	2.68	44.67	5.43	**5.30**

Table 3. Comparisons with losses based on pre-trained networks.

2.1 Comparisons on Point Cloud Completion

To make a more convincing evaluation for our method, we also conduct comparisons on point cloud completion, which is a task similar with point cloud reconstruction. Point cloud completion networks predict point clouds as identical as possible to the ground truth completed point clouds from partial input point clouds, which also needs to be trained with shape losses. In this section, we select two popular point cloud completion networks PCN [5] and CRN [4], and replace their adopted losses for completed results with our LNSA. We retrain different methods based on the 2048 points dataset provided by CRN and compared performances based on data normalized to -1∼1, where we use four metrics CD, EMD from PCN [5] and MCD, HD to evaluate the completion performances. The results are presented in Table 4. We can see that both completion networks can achieve better performances after training with LNSA, which confirms our method is quite effective. We also present some qualitative comparisons

Network	PCN				CRN			
Metrics	CD	EMD	MCD	HD	CD	EMD	MCD	HD
w/o LNSA	4.44	12.13	0.58	3.51	4.52	12.47	0.54	3.45
w/ LNSA	**4.36**	**11.52**	**0.48**	**3.20**	**4.06**	**11.19**	**0.42**	**3.07**

Table 4. Comparisons on point cloud completion. * denotes training with LNSA.

on CRN [4] trained with or without LNSA to visualize the main differences. The results are presented in Fig. 2. We can see that LNSA can help the network improve the completion performances on the bounding contours as circled.

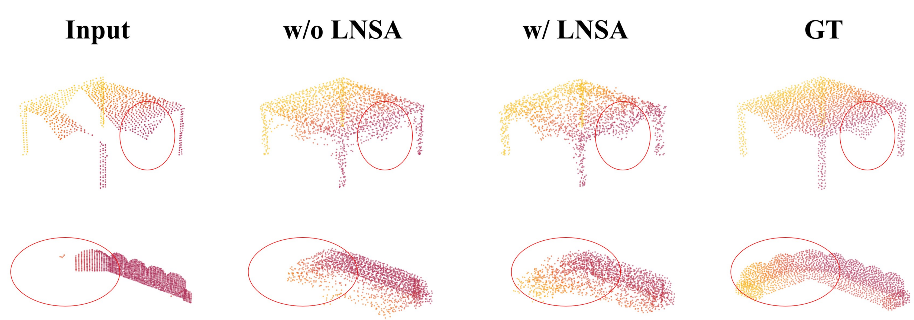

Fig. 2. Qualitative results on completion.

3 Comparison on Noised data

To further verify the effectiveness of PCLossNet, we also conduct experiments on noised objects from ShapeNet and real-world Scannet, where we add Gaussian noises $\sigma = 0.02, 0, 04, 0.06$ to ShapeNet. The quantitative results are presented in Table 5 and 6, where the qualitative results are presented in Fig.3. We can see that our method LNSA still have better performances on the noisy objects, which further confirms its robustness.

Noise	0.02		0.04		0.06	
Metric	MCD↓	HD↓	MCD	HD	MCD	HD
CD	0.33	1.92	0.41	2.32	0.52	3.07
EMD	0.25	2.23	0.29	2.44	0.37	3.01
LNSA	**0.23**	**1.70**	**0.28**	**2.14**	**0.36**	**2.99**

Table 5. Comparisons on noised objects from ShapeNet.

RecNet	AE			LAE		
Metrics	CD	EMD	LNSA	CD	EMD	LNSA
MCD ↓	0.51	0.42	**0.38**	0.30	0.34	**0.12**
HD ↓	3.72	3.88	**3.36**	0.99	4.03	**0.59**

Table 6. Comparisons on noised objects from ShapeNet.

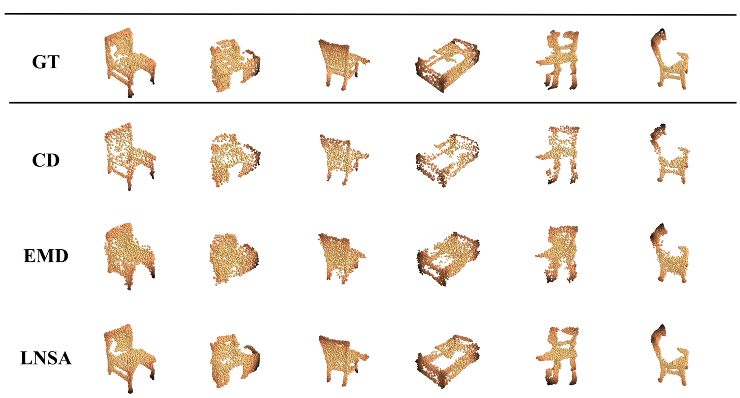

Fig. 3. Qualitative results on Scannet based on LAE reconstruction network.

4 More Reconstruction Results

To further display the differences between our methods and matching losses, we present more qualitative results on multiple reconstruction networks in Fig. 4 and Fig. 5. We can see our methods can help reconstruct more accurate details.

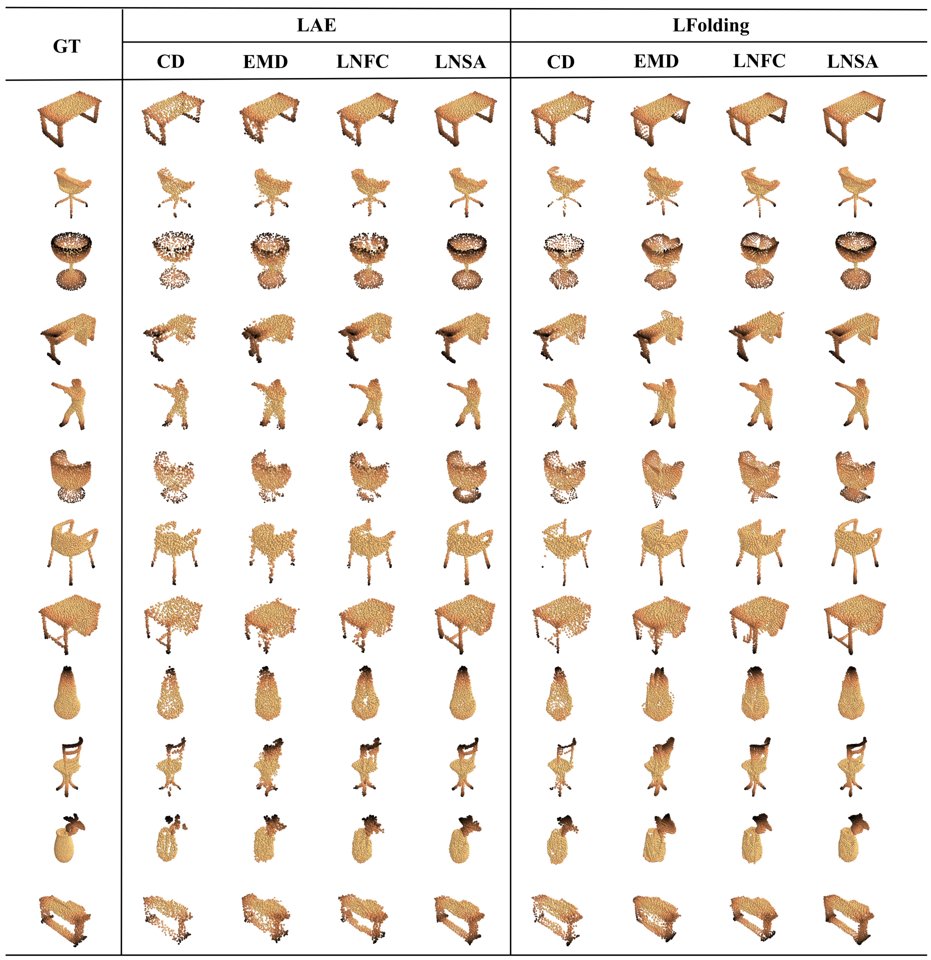

Fig. 4. More qualitative results on strong local feature-based reconstruction networks.

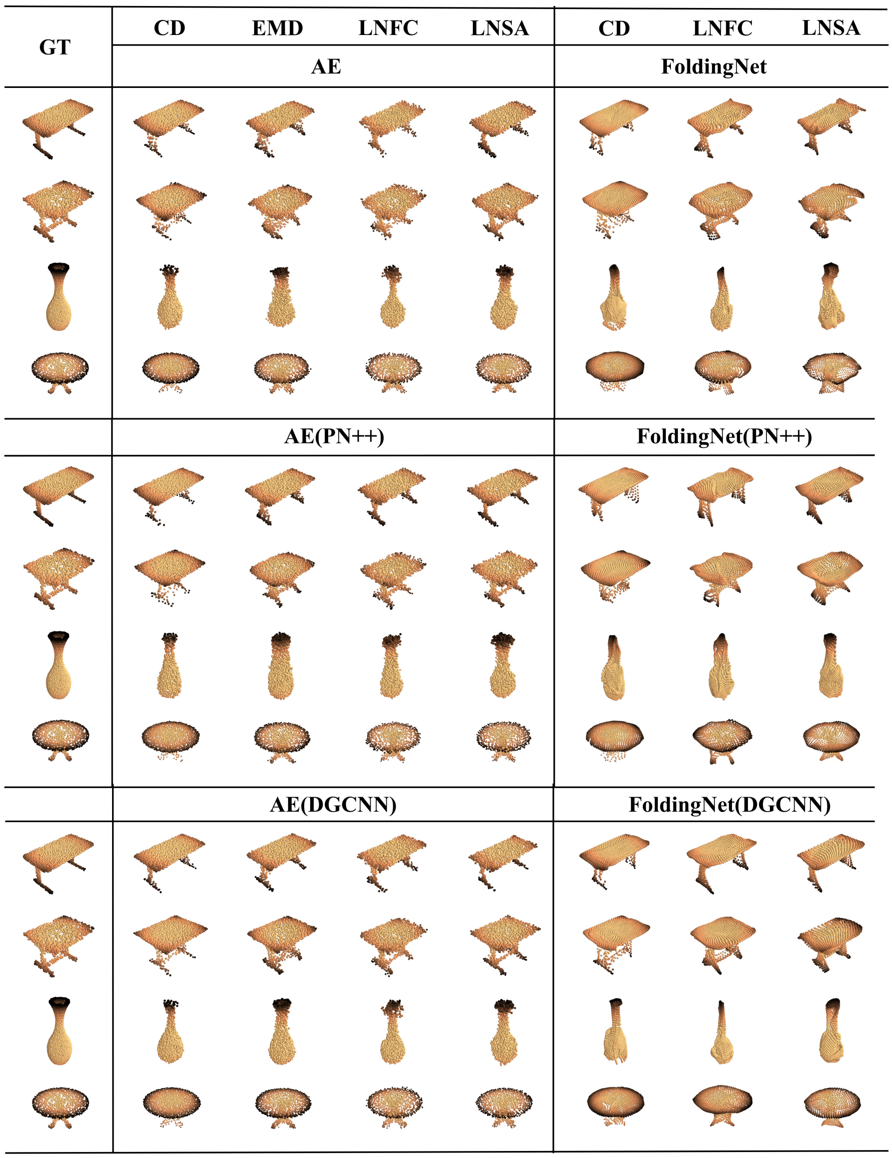

Fig. 5. Qualitative results on different global feature-based reconstruction networks.

References

1. Achlioptas, P., Diamanti, O., Mitliagkas, I., Guibas, L.: Learning representations and generative models for 3d point clouds. In: International conference on machine learning. pp. 40–49. PMLR (2018)
2. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: European conference on computer vision. pp. 694–711. Springer (2016)
3. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In: Advances in neural information processing systems. pp. 5099–5108 (2017)
4. Wang, X., Ang Jr, M.H., Lee, G.H.: Cascaded refinement network for point cloud completion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 790–799 (2020)
5. Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M.: Pcn: Point completion network. In: 2018 International Conference on 3D Vision (3DV). pp. 728–737. IEEE (2018)