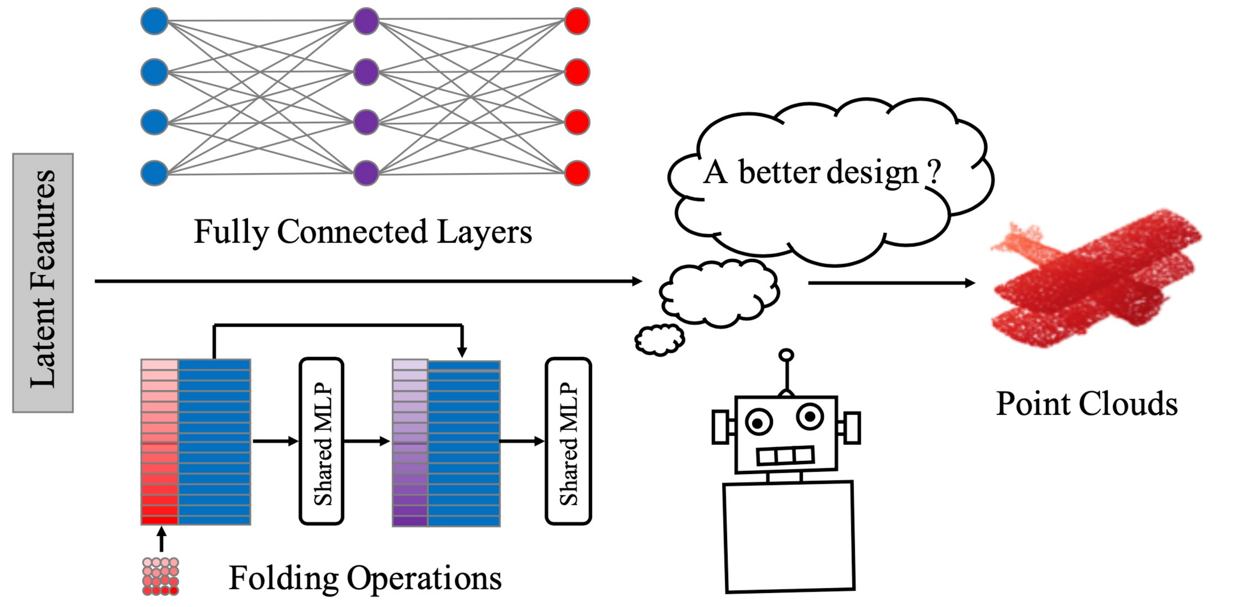
Attention-based Transformation from Latent Features to Point Clouds

Kaiyi Zhang[1], Ximing Yang[1], Yuan Wu[1], Cheng Jin[1,2]
[1]School of Computer Science, Fudan University, Shanghai, China
[2]Peng Cheng Laboratory, Shenzhen, China

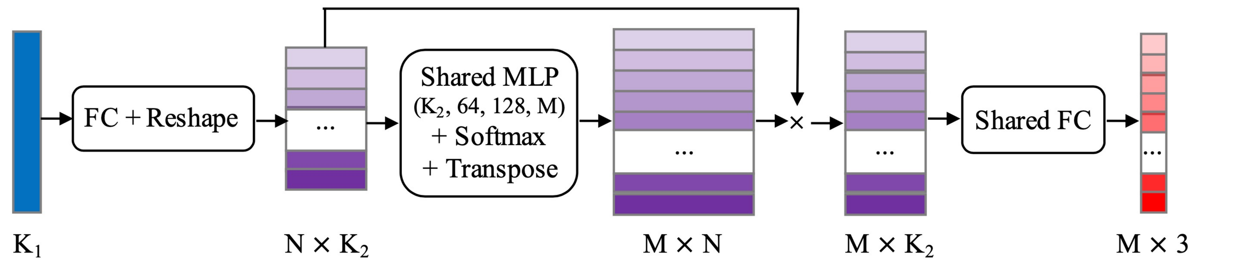
Introduction

In point cloud generation and completion, previous methods for transforming latent features to point clouds are generally based on fully connected layers (FC-based) or folding operations (Folding-based). However, point clouds generated by FC-based methods are usually troubled by outliers and rough surfaces. For folding-based methods, their data flow is large, convergence speed is slow, and they are also hard to handle the generation of non-smooth surfaces. Therefore, a better design needs to be proposed.



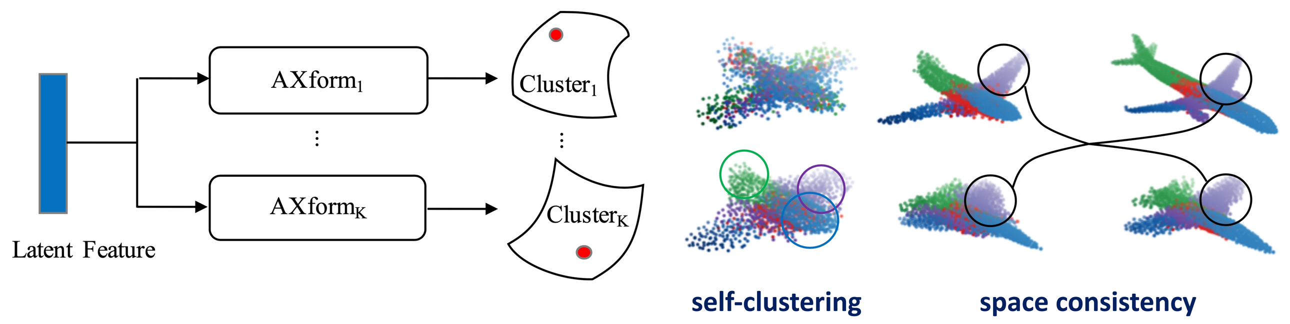
Approach

In this paper, we propose AXform, an attention-based method to transform latent features to point clouds.



It includes 3 steps:

1. Use a fully connected layer to generate points in an interim space
2. Weighting operations are performed by an attention mechanism to get centripetal constrained target points in the interim space
3. Map the target points to 3d space and get the final point cloud



Advantage 1:

When AXform is expanded to multiple branches for local generations, the centripetal constraint makes it has properties of self-clustering and space consistency.

Methods	Point constraint	#Params.	Data flow	Convergence speed
FC-based	None	Large	**Small**	Middle
Folding-based	Surface, Strong	Small	Large	Slow
Ours	Centripetal, **Soft**	**Tiny**	Middle	**Fast**

Advantage 2:

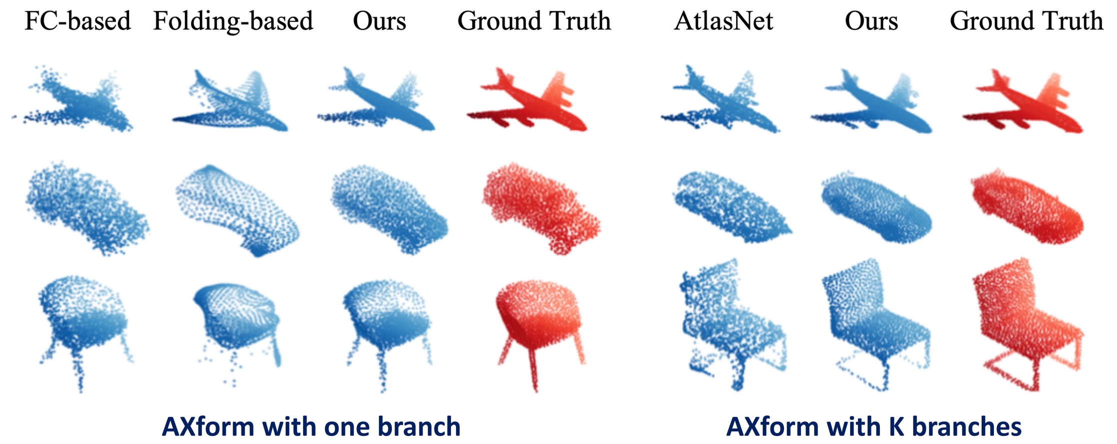
In aspects of point constraints, network parameters, data flow, and convergence speed, AXform performs better than previous two methods.

Reconstruction and Generation

AXform concatenated with a simple Shared MLP encoder can be used for point cloud reconstruction and generation.

#Branches	Category	Methods	K_1	K_2	N	CD ↓	Params. ↓
K = 1	Airplane	FC-based	1024			7.895	7.4M
		Folding-based	512			9.208	1.1M
		Ours	128	32	128	**4.386**	**0.8M**
	Car	FC-based	1024			11.523	7.4M
		Folding-based	512			20.989	1.1M
		Ours	128	32	128	**8.008**	**0.8M**
	Chair	FC-based	1024			13.861	7.4M
		Folding-based	512			23.103	1.1M
		Ours	128	32	128	**11.606**	**0.8M**
	All	FC-based	1024			8.578	7.4M
		Folding-based	512			12.980	1.1M
		Ours	128	32	128	**8.046**	**0.8M**
K = 16	Airplane	AtlasNet	1024			6.307	27.5M
		Ours	128	32	128	**3.607**	**8.9M**
	Car	AtlasNet	1024			15.269	27.5M
		Ours	128	32	128	**7.670**	**8.9M**
	Chair	AtlasNet	1024			17.154	27.5M
		Ours	128	32	128	**9.423**	**8.9M**
	All	AtlasNet	1024			11.057	27.5M
		Ours	128	32	128	**6.867**	**8.9M**

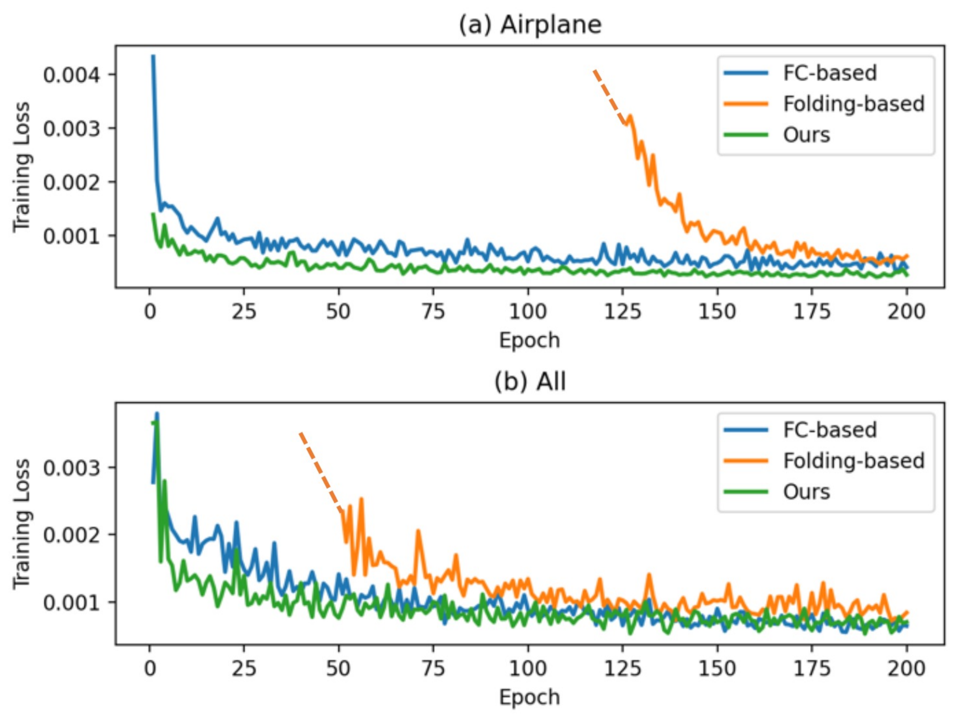
Quantitative comparison of reconstruction results on our sampled ShapeNet dataset. AXform has lower CD with fewer parameters.



Visualized comparison of reconstruction results on our sampled ShapeNet dataset. Trained on each category.

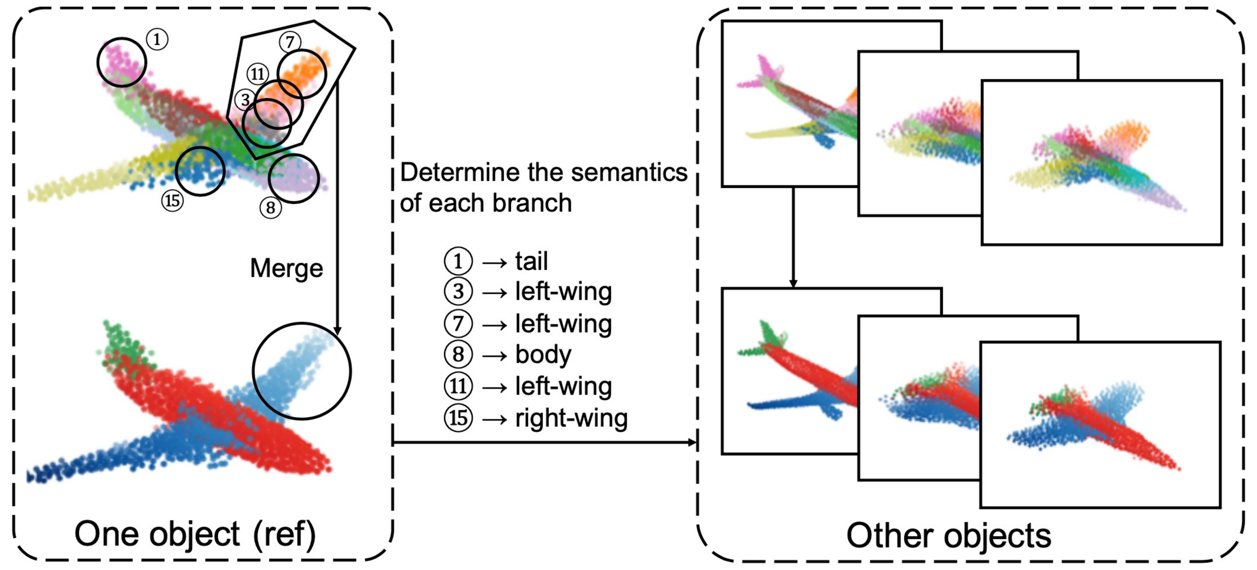
Methods	JSD ↓	MMD ↓		COV %, ↑		1-NNA %, ↓	
		CD	EMD	CD	EMD	CD	EMD
l-GAN (CD)	7.24	**0.454**	**4.43**	**33.66**	**25.70**	63.00	81.23
l-GAN-AXform	**6.27**	0.498	4.54	33.33	24.59	**61.59**	**78.55**

Quantitative comparison of replacing the MLP decoder in l-GAN (CD) with our AXform. AXform improves the most important metric 1-NNA.



Loss curve during training. The convergence speed of AXform is faster.

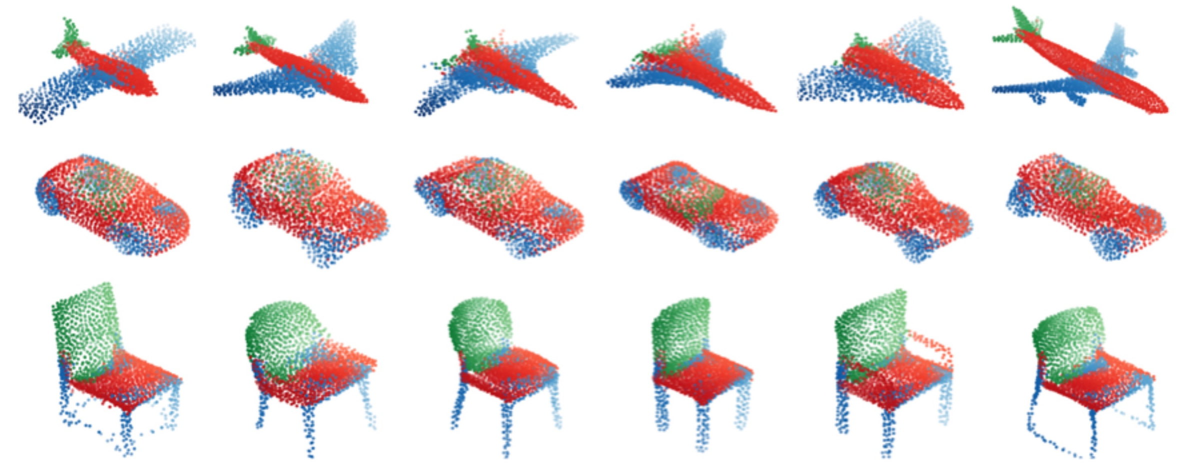
Unsupervised Semantic Segmentation



AXform with K branches can realize unsupervised semantic segmentation on the generated point clouds due to the property of space consistency.

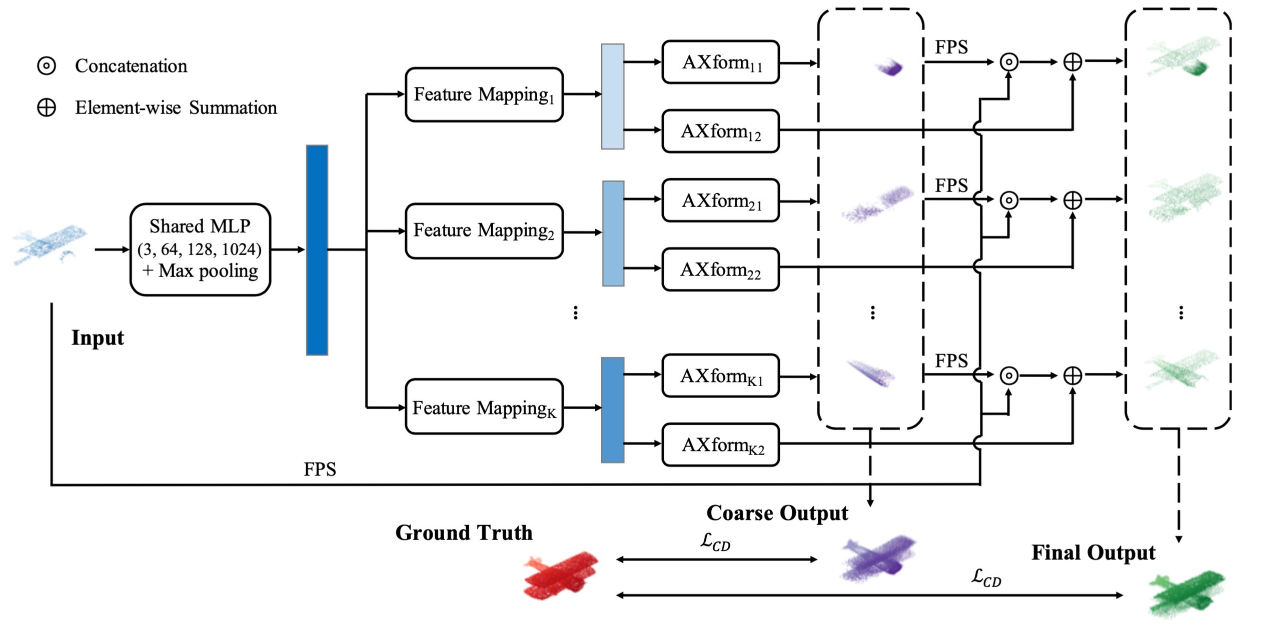
It includes 2 steps:

1. Choose one object in the training set as a reference to determine the semantics of each branch
2. For other objects in the test set. Merge the outputs of the branches with the same semantic



Unsupervised semantic segmentation results on three categories.

Point Cloud Completion



We apply AXform on point cloud completion and get AXformNet.

It includes K branches and two stages. Each branch generates a part of the target point cloud. The first stage uses AXform to generate a coarse point cloud. In the second stage, refinement is performed by combining input to generate a final point cloud.

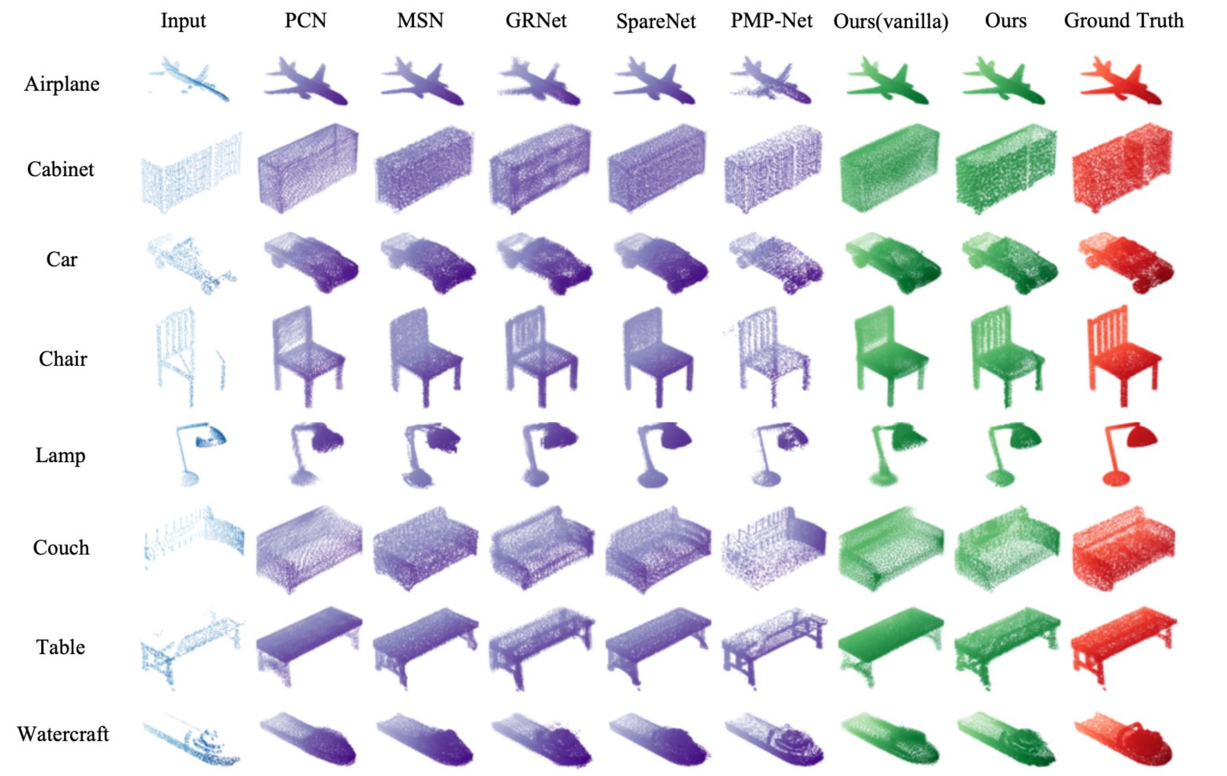
Loss Functions:

$$\mathcal{L} = \alpha\mathcal{L}_{CD}(Y_{coarse}, Y_{gt}) + \mathcal{L}_{CD}(Y_{final}, Y_{gt})$$

Methods	Airplane	Cabinet	Car	Chair	Lamp	Couch	Table	Watertcraft	Average	Methods	Airplane	Cabinet	Car	Chair	Lamp	Couch	Table	Watertcraft	Average
FoldingNet	9.491	15.796	12.611	15.545	16.413	15.969	13.649	14.987	14.308	MSN	5.596	11.963	10.776	10.620	10.712	11.895	8.704	9.485	9.969
AtlasNet	6.366	11.943	10.105	12.063	12.369	12.990	10.331	10.607	10.847	GRNet	6.450	10.373	9.447	9.408	7.955	10.512	8.444	8.039	8.828
PCN	5.502	**10.625**	8.696	10.998	11.339	**11.676**	**8.590**	9.665	9.636	SpareNet	5.956	12.567	9.956	11.931	11.105	13.388	9.950	9.589	10.555
TopNet	7.614	13.311	10.898	13.823	14.439	14.779	11.224	11.124	12.151	PMP-Net	5.650	11.240	9.640	9.510	**6.950**	10.830	8.720	**7.250**	8.730
Ours(vanilla)	**5.366**	10.677	**8.646**	**10.743**	**10.458**	11.683	8.727	**9.300**	**9.450**	Ours	**4.760**	**10.178**	**8.600**	**9.133**	8.173	**10.395**	**7.752**	7.803	**8.349**

Methods	Airplane	Cabinet	Car	Chair	Lamp	Couch	Table	Watertcraft	Average	Methods	Airplane	Cabinet	Car	Chair	Lamp	Couch	Table	Watertcraft	Average
FoldingNet	0.642	0.237	0.382	0.236	0.219	0.197	0.361	0.299	0.322	MSN	0.885	**0.644**	0.665	0.657	0.699	0.604	0.782	0.708	0.705
AtlasNet	0.845	0.552	0.630	0.552	0.565	0.500	0.660	0.624	0.616	GRNet	0.843	0.618	0.682	0.673	0.761	0.605	0.751	0.750	0.708
PCN	0.881	**0.651**	0.725	0.625	0.638	**0.581**	**0.765**	0.697	0.695	SpareNet	0.869	0.571	0.672	0.592	0.647	0.527	0.719	0.690	0.661
TopNet	0.771	0.404	0.544	0.413	0.408	0.350	0.572	0.560	0.503	PMP-Net	0.860	0.495	0.570	0.600	0.778	0.516	0.639	0.742	0.650
Ours(vanilla)	**0.893**	0.634	**0.725**	**0.632**	**0.667**	0.567	0.756	**0.703**	**0.697**	Ours	**0.920**	0.642	**0.734**	**0.704**	**0.782**	**0.605**	**0.801**	**0.778**	**0.746**

Quantitative comparison between our methods and existing methods. The metrics are CD and F-Score@1%.



Visualized completion comparison on PCN dataset.

Supported by National Natural Science Fund of China (62176064), Zhejiang Lab (2019KD0AB06)