

Supplementary Materials of Reverse2Complete: Unpaired Multimodal Point Cloud Completion via Guided Diffusion

1 CLASSIFIER ARCHITECTURE AND TRAINING DETAILS

The overview of the time-dependent classifier is displayed in Fig. 1, and is also constructed upon the point-voxel CNN [1]. Tab. 1 shows the specific classifier parameters, where SA represents the set abstraction (SA) module utilized in point-voxel CNN. We compute the cross-entropy loss between the ground truth label and the output. We set $T = 100$ instead of using the same range of the diffusion model, as the DDPM implementation results in convergence to noise in the early time steps. To optimize, we used the Adam optimizer with a fixed learning rate of 1×10^{-4} .

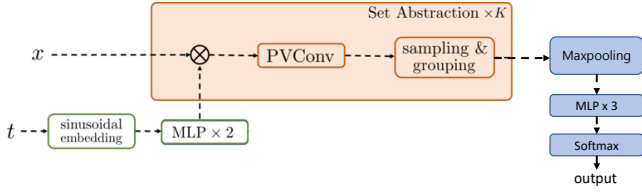


Figure 1: The details of the architecture of our proposed time-dependent classifier.

Input Feature Size: 2048×3				
Input Time Embedding Size: 64				
Output Feature Size: 2048×3				
Time Embedding				
Sinusoidal Embedding dim = 64				
MLP(64, 64)				
LeakyReLU(0.1)				
MLP(64, 64)				
	SA 1	SA 2	SA 3	SA 4
L	2	3	3	0
C_{input}	3	32	64	-
E_t	64	64	64	-
C_{output}	32	64	128	-
D	32	16	8	-
use_attn	False	True	False	-
N_{center}^-	1024	256	64	16
r	0.1	0.2	0.4	0.8
$N_{neighbor}$	32	32	32	32
MLP(512, 128, 64, 2)				
Softmax				

Table 1: Architectue details of the classifier guidance.

2 MORE VISUALIZATION RESULTS

We provide additional visualization comparisons on the 3DEPN, CRN, and real-scanned datasets in Fig. 2.

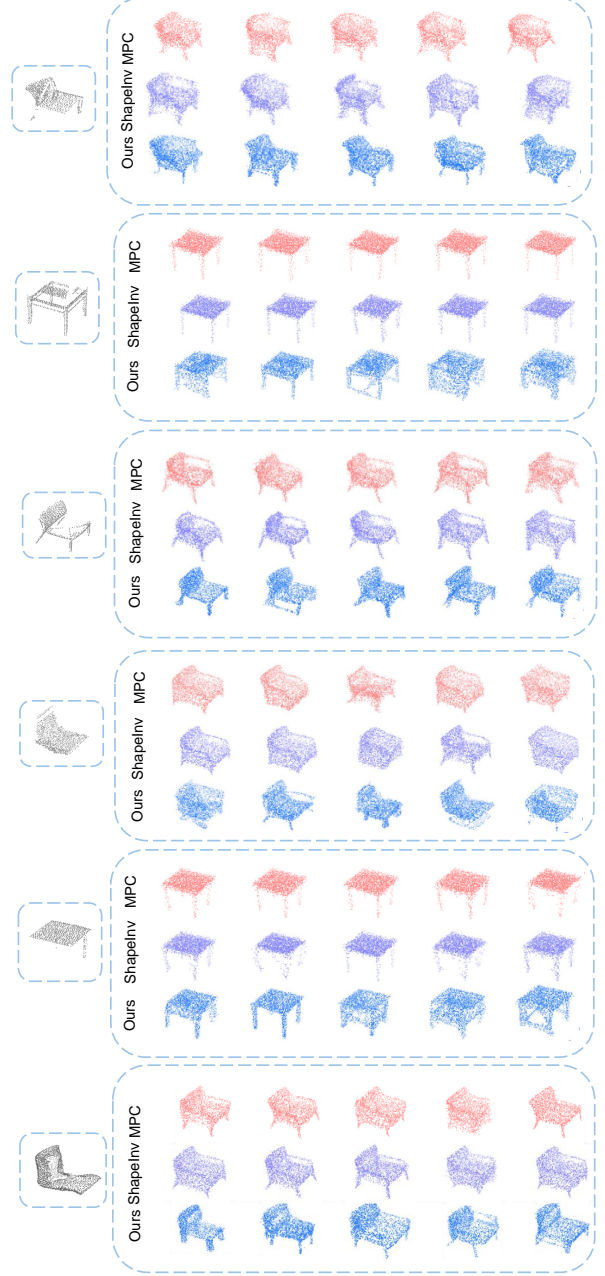


Figure 2: More visual comparison on 3DEPN, CRN and real-scanned datasets.

REFERENCES

- [1] Zhijian Liu, Haotian Tang, Yujun Lin, and Song Han. 2019. Point-Voxel CNN for Efficient 3D Deep Learning. In *Conference on Neural Information Processing Systems (NeurIPS)*.