

Supplementary Materials: GeoFormer: Learning Point Cloud Completion with Tri-Plane Integrated Transformer

Anonymous Authors

Within this supplemental material, we offer further details to enhance the manuscript. Firstly, we present more detailed experimental settings and dataset settings, and secondly, we provide more visual results and complete quantitative indicator results with other methods.

1 IMPLEMENTATION DETAILS

1.1 Implementation Settings

We implement our GeoFormer on the PyTorch platform. The number of the down-sampled point features from PointNet++ head is 256. The seed generator of GeoFormer produce a set of feature \mathcal{F} and 256 coarse complete point clouds. Two multi-scale aware upsample layers are used in the following coarse-to-fine generation procedure which output dense point clouds $\mathcal{P}_1, \mathcal{P}_2$ and \mathcal{P}_2 corresponds to the final predicted result. Our networks are trained by exploiting AdamW optimizer with the base learning rate set as 0.0002. All the models are test with a batch size of 1 on one NVIDIA A100 GPU for accurate and fair comparison.

1.2 Dataset Settings

The PCN Dataset. PCN[5] is one of the most popular benchmark in point cloud completion, it is a subset of ShapeNet[1] contains shape from 8 categories. For each shape, 16,384 points are uniformly sampled from mesh surfaces as completed ground truth and the partial input has 2,048 points.

The ShapeNet-55/34 Dataset. ShapeNet-55 and ShapeNet-34 datasets was proposed by [4], which was also generated from the ShapeNet[1] dataset while contains more object categories and incomplete patterns. All 55 categories in ShapeNet are included in ShapeNet-55 with 41,952 shapes for training and 10,518 shapes for testing. ShapeNet-34 uses a subset of 34 categories for training and leaves 21 unseen categories for testing where 46,765 object shapes are used for training, 3,400 for testing on seen categories and 2,305 for testing on novel (unseen) categories. In both datasets, 2,048 points are sampled as input and 8,192 points as ground truth. Following the same evaluation strategy with [4], 8 fixed viewpoints are selected and the number of points in the partial point cloud is set to 2,048, 4,096 or 6,144 (25%, 50% or 75% of the complete point cloud) which corresponds to three difficulty levels of simple, moderate and hard in the test stage.

The KITTI Dataset. To test the generalization ability in real-world, we further test our method on KITTI[2], which contain partial point clouds extracted from LiDAR scans from outdoor scenes where car objects are extracted in each frame according to the 3D bounding boxes, resulting in a total of 2,401 partial point clouds. Specifically, we test on these 2,401 KITTI cars extracted by [5].

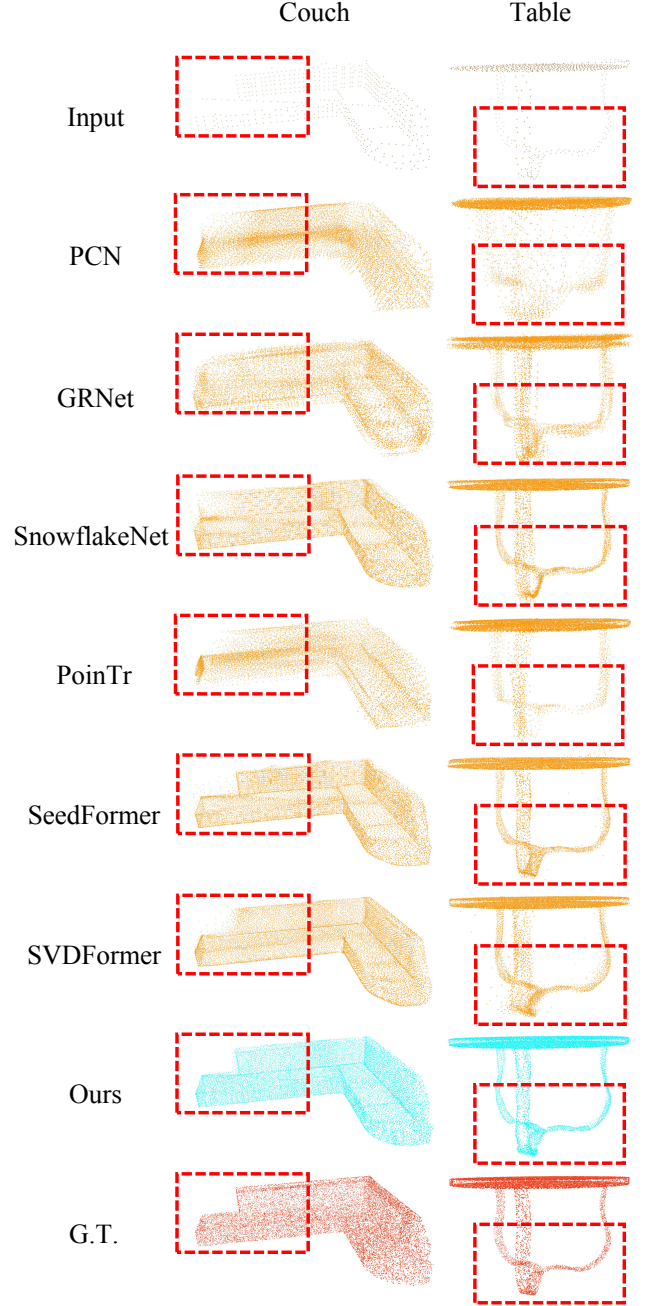


Figure 1: Visualization comparisons on PCN dataset. Results clearly show that our method can preserve better global structure, and reconstruct better local details.

Table 1: Detailed results for novel 21 categories on ShapeNet-34 dataset. *S.*, *M.* and *H.* stand for the simple, moderate and hard difficulty levels.

	PCN [5]			GRNet [3]			PoinTr [4]			SeedFormer [6]			Ours		
	S.	M.	H.	S.	M.	H.	S.	M.	H.	S.	M.	H.	S.	M.	H.
bag	2.48	2.46	3.94	1.47	1.88	3.45	0.96	1.34	2.08	0.49	0.82	1.45	0.43	0.70	1.28
basket	2.79	2.51	4.78	1.78	1.94	4.18	1.04	1.40	2.90	0.60	0.85	1.98	0.54	0.74	1.66
birdhouse	3.53	3.47	5.31	1.89	2.34	5.16	1.22	1.79	3.45	0.72	1.19	2.31	0.62	1.01	1.93
bowl	2.66	2.35	3.97	1.77	1.97	3.90	1.05	1.32	2.40	0.60	0.77	1.50	0.49	0.62	1.10
camera	4.84	5.30	8.03	2.31	3.38	7.20	1.63	2.67	4.97	0.89	1.77	3.75	0.82	1.63	3.48
can	1.95	1.89	5.21	1.53	1.80	3.08	0.80	1.17	2.85	0.56	0.89	1.57	0.47	0.72	1.42
cap	7.21	7.14	10.94	3.29	4.87	13.02	1.40	2.74	8.35	0.50	1.34	5.19	0.57	1.29	3.99
keyboard	1.07	1.00	1.23	0.73	0.77	1.11	0.43	0.45	0.63	0.32	0.41	0.60	0.28	0.35	0.46
dishwasher	2.45	2.09	3.53	1.79	1.70	3.27	0.93	1.05	2.04	0.63	0.78	1.44	0.53	0.65	1.15
earphone	7.88	6.59	16.53	4.29	4.16	10.30	2.03	5.10	10.69	1.18	2.78	6.71	1.29	2.69	7.41
helmet	6.15	6.41	9.16	3.06	4.38	10.27	1.86	3.30	6.96	1.10	2.27	4.78	1.07	2.39	5.09
mailbox	2.74	2.68	4.31	1.52	1.90	4.33	1.03	1.47	3.34	0.56	0.99	2.06	0.43	0.88	1.87
microphone	4.36	4.65	8.46	2.29	3.23	8.41	1.25	2.27	5.47	0.80	1.61	4.21	0.79	1.96	4.20
microwaves	2.59	2.35	4.47	1.74	1.81	3.82	1.01	1.18	2.14	0.64	0.83	1.69	0.56	0.69	1.36
pillow	2.09	2.16	3.54	1.43	1.69	3.43	0.92	1.24	2.39	0.43	0.66	1.45	0.37	0.57	1.16
printer	3.28	3.60	5.56	1.82	2.41	5.09	1.18	1.76	3.10	0.69	1.25	2.33	0.57	1.02	2.07
remote	0.95	1.08	1.58	0.82	1.02	1.29	0.44	0.58	0.78	0.27	0.42	0.61	0.23	0.35	0.49
rocket	1.39	1.22	2.01	0.97	0.79	1.60	0.39	0.72	1.39	0.28	0.51	1.02	0.25	0.50	0.93
skateboard	1.97	1.78	2.45	0.93	1.07	1.83	0.52	0.80	1.31	0.35	0.56	0.92	0.24	0.43	0.77
tower	2.37	2.40	4.35	1.35	1.80	3.85	0.82	1.35	2.48	0.51	0.92	1.87	0.42	0.81	1.70
washer	2.77	2.52	4.64	1.83	1.97	5.28	1.04	1.39	2.73	0.61	0.87	1.94	0.53	0.73	1.59
mean	3.22	3.13	5.43	1.84	2.23	4.95	1.05	1.67	3.45	0.61	1.07	2.35	0.55	0.99	2.15

2 ADDITIONAL EXPERIMENTAL RESULTS

2.1 Qualitive Results on PCN Dataset

As shown in Figure 1, we provide more visual results compared with other methods. Results shows that all methods are successful in generating the overall shapes in case of couch and tables. However, Our GeoFormer method significantly outperforms the other methods by producing more accurate and complete edges for detailed structures, such as couch borders and table legs.

2.2 More Results on ShapeNet-55/34

We report complete results of our method on ShapeNet-55 in Table 2 and results of novel 21 categories on ShapeNet-34 in Table 1. The models are tested under three difficulty levels: simple (S), moderate (M) and hard (H). We can see that GeoFormer achieves best scores.

REFERENCES

- [1] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al. 2015. Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012* (2015).
- [2] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. 2013. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research* 32, 11 (2013), 1231–1237.
- [3] Haozhe Xie, Hongxun Yao, Shangchen Zhou, Jiageng Mao, Shengping Zhang, and Wenxiu Sun. 2020. Grnet: Gridding residual network for dense point cloud completion. In *European Conference on Computer Vision*. Springer, 365–381.
- [4] Xumin Yu, Yongming Rao, Ziyi Wang, Zuyan Liu, Jiwen Lu, and Jie Zhou. 2021. Pointr: Diverse point cloud completion with geometry-aware transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*. 12498–12507.
- [5] Wentao Yuan, Tejas Khot, David Held, Christoph Mertz, and Martial Hebert. 2018. Pcn: Point completion network. In *2018 international conference on 3D vision (3DV)*. IEEE, 728–737.
- [6] Haoran Zhou, Yun Cao, Wenqing Chu, Junwei Zhu, Tong Lu, Ying Tai, and Chengjie Wang. 2022. Seedformer: Patch seeds based point cloud completion with upsampling transformer. In *European conference on computer vision*. Springer, 416–432.

Table 2: Detailed results on ShapeNet-55 dataset. *S.*, *M.* and *H.* stand for the simple, moderate and hard difficulty levels.

	PCN [5]			GRNet [3]			PoinTr [4]			SeedFormer [6]			Ours		
	S.	M.	H.	S.	M.	H.	S.	M.	H.	S.	M.	H.	S.	M.	H.
airplane	0.90	0.89	1.32	0.87	0.87	1.27	0.27	0.38	0.69	0.23	0.35	0.61	0.20	0.30	0.51
trash-bin	2.16	2.18	5.15	1.69	2.01	3.48	0.80	1.15	2.15	0.73	1.08	1.94	0.59	0.85	1.48
bag	2.11	2.04	4.44	1.41	1.70	2.97	0.53	0.74	1.51	0.43	0.67	1.28	0.35	0.55	1.06
basket	2.21	2.10	4.55	1.65	1.84	3.15	0.73	0.88	1.82	0.65	0.83	1.54	0.59	0.72	1.35
bathtub	2.11	2.09	3.94	1.46	1.73	2.73	0.64	0.94	1.68	0.52	0.82	1.45	0.44	0.67	1.16
bed	2.86	3.07	5.54	1.64	2.03	3.70	0.76	1.10	2.26	0.63	0.91	1.89	0.51	0.78	1.63
bench	1.31	1.24	2.14	1.03	1.09	1.71	0.38	0.52	0.94	0.32	0.42	0.84	0.26	0.33	0.61
birdhouse	3.29	3.53	6.69	1.87	2.40	4.71	0.98	1.49	3.13	0.76	1.30	2.46	0.67	1.11	2.11
bookshelf	2.70	2.70	4.61	1.42	1.71	2.78	0.71	1.06	1.93	0.57	0.84	1.57	0.51	0.74	1.33
bottle	1.25	1.43	4.61	1.05	1.44	2.67	0.37	0.74	1.50	0.31	0.63	1.21	0.26	0.53	1.02
bowl	2.05	1.83	3.66	1.60	1.77	2.99	0.68	0.78	1.44	0.56	0.65	1.18	0.46	0.51	0.86
bus	1.20	1.14	2.08	1.06	1.16	1.48	0.42	0.55	0.79	0.42	0.55	0.73	0.34	0.45	0.60
cabinet	1.60	1.49	3.47	1.27	1.41	2.09	0.55	0.66	1.16	0.57	0.69	1.05	0.47	0.54	0.83
camera	4.05	4.54	8.27	2.14	3.15	6.09	1.10	2.03	4.34	0.83	1.68	3.45	0.68	1.39	2.82
can	2.02	2.28	6.48	1.58	2.11	3.81	0.68	1.19	2.14	0.58	1.03	1.79	0.49	0.84	1.50
cap	1.82	1.76	4.20	1.17	1.37	3.05	0.46	0.62	1.64	0.33	0.45	1.18	0.28	0.36	0.64
car	1.48	1.47	2.60	1.29	1.48	2.14	0.64	0.86	1.25	0.65	0.86	1.17	0.48	0.67	0.92
cellphone	0.80	0.79	1.71	0.82	0.91	1.18	0.32	0.39	0.60	0.31	0.40	0.54	0.26	0.31	0.40
chair	1.70	1.81	3.34	1.24	1.56	2.73	0.49	0.74	1.63	0.41	0.65	1.38	0.32	0.51	1.10
clock	2.10	2.01	3.98	1.46	1.66	2.67	0.62	0.84	1.65	0.53	0.74	1.35	0.43	0.61	1.10
keyboard	0.82	0.82	1.04	0.74	0.81	1.09	0.30	0.39	0.45	0.28	0.36	0.45	0.23	0.27	0.33
dishwasher	1.93	1.66	4.39	1.43	1.59	2.53	0.55	0.69	1.42	0.56	0.69	1.30	0.48	0.55	1.07
display	1.56	1.66	3.26	1.13	1.38	2.29	0.48	0.67	1.33	0.39	0.59	1.10	0.33	0.47	0.89
earphone	3.13	2.94	7.56	1.78	2.18	5.33	0.81	1.38	3.78	0.64	1.04	2.75	0.52	0.79	3.05
faucet	3.21	3.48	7.52	1.81	2.32	4.91	0.71	1.42	3.49	0.55	1.15	2.63	0.52	1.03	2.26
filecabinet	2.02	1.97	4.14	1.46	1.71	2.89	0.63	0.84	1.69	0.63	0.84	1.49	0.55	0.71	1.27
guitar	0.42	0.38	1.23	0.44	0.48	0.76	0.14	0.21	0.42	0.13	0.19	0.32	0.10	0.17	0.28
helmet	3.76	4.18	7.53	2.33	3.18	6.03	0.99	1.93	4.22	0.79	1.52	3.61	0.62	1.18	3.09
jar	2.57	2.82	6.00	1.72	2.37	4.37	0.77	1.33	2.87	0.63	1.13	2.36	0.52	0.90	1.96
knife	0.94	0.62	1.37	0.72	0.66	0.96	0.20	0.33	0.56	0.15	0.28	0.45	0.14	0.24	0.40
lamp	3.10	3.45	7.02	1.68	2.43	5.17	0.64	1.40	3.58	0.45	1.06	2.67	0.44	1.05	2.67
laptop	0.75	0.79	1.59	0.83	0.87	1.28	0.32	0.34	0.60	0.32	0.37	0.55	0.27	0.28	0.41
loudspeaker	2.50	2.45	5.08	1.75	2.08	3.45	0.78	1.16	2.17	0.67	1.01	1.80	0.54	0.82	1.54
mailbox	1.66	1.74	5.18	1.15	1.59	3.42	0.39	0.78	2.56	0.30	0.67	2.04	0.25	0.57	1.49
microphone	3.44	3.90	8.52	2.09	2.76	5.70	0.70	1.66	4.48	0.62	1.61	3.66	0.60	1.76	3.52
microwaves	2.20	2.01	4.65	1.51	1.72	2.76	0.67	0.83	1.82	0.63	0.79	1.47	0.54	0.65	1.24
motorbike	2.03	2.01	3.13	1.38	1.52	2.26	0.75	1.10	1.92	0.68	0.96	1.44	0.51	0.78	1.22
mug	2.45	2.48	5.17	1.75	2.16	3.79	0.91	1.17	2.35	0.79	1.03	2.06	0.64	0.88	1.67
piano	2.64	2.74	4.83	1.53	1.82	3.21	0.76	1.06	2.23	0.62	0.87	1.79	0.48	0.65	1.31
pillow	1.85	1.81	3.68	1.42	1.67	3.04	0.61	0.82	1.56	0.48	0.75	1.41	0.37	0.50	0.93
pistol	1.25	1.17	2.65	1.11	1.06	1.76	0.43	0.66	1.30	0.37	0.56	0.96	0.31	0.50	0.81
flowerpot	3.32	3.39	6.04	2.02	2.48	4.19	1.01	1.51	2.77	0.93	1.30	2.32	0.72	1.07	1.96
printer	2.90	3.19	5.84	1.56	2.38	4.24	0.73	1.21	2.47	0.58	1.11	2.13	0.50	0.86	1.79
remote	0.99	0.97	2.04	0.89	1.05	1.29	0.36	0.53	0.71	0.29	0.46	0.62	0.23	0.38	0.48
rifle	0.98	0.80	1.31	0.83	0.77	1.16	0.30	0.45	0.79	0.27	0.41	0.66	0.22	0.35	0.57
rocket	1.05	1.04	1.87	0.78	0.92	1.44	0.23	0.48	0.99	0.21	0.46	0.83	0.15	0.36	0.73
skateboard	1.04	0.94	1.68	0.82	0.87	1.24	0.28	0.38	0.62	0.23	0.32	0.62	0.18	0.27	0.41
sofa	1.65	1.61	2.92	1.35	1.45	2.32	0.56	0.67	1.14	0.50	0.62	1.02	0.41	0.49	0.79
stove	2.07	2.02	4.72	1.46	1.72	3.22	0.63	0.92	1.73	0.59	0.87	1.49	0.51	0.74	1.24
table	1.56	1.50	3.36	1.15	1.33	2.33	0.46	0.64	1.31	0.41	0.58	1.18	0.35	0.47	0.91
telephone	0.80	0.80	1.67	0.81	0.89	1.18	0.31	0.38	0.59	0.31	0.39	0.55	0.26	0.31	0.43
tower	1.91	1.97	4.47	1.26	1.69	3.06	0.55	0.90	1.95	0.47	0.84	1.65	0.40	0.72	1.56
train	1.50	1.41	2.37	1.09	1.14	1.61	0.50	0.70	1.12	0.51	0.66	1.01	0.37	0.53	0.82
watercraft	1.46	1.39	2.40	1.09	1.12	1.65	0.41	0.62	1.07	0.35	0.56	0.92	0.28	0.48	0.79
washer	2.42	2.31	6.08	1.72	2.05	4.19	0.75	1.06	2.44	0.64	0.91	2.04	0.56	0.74	1.51
mean	1.96	1.98	4.09	1.35	1.63	2.86	0.58	0.88	1.80	0.50	0.77	1.49	0.41	0.64	1.25