Supplementary material

PointAugment: an Auto-Augmentation Framework for Point Cloud Classification

CVPR 2020

Overview

This supplementary document contains the following parts.

- In Section A, we provide the architecture details of our Augmentor network.
- In Section B, we show more visual comparisons of shape retrieval.
- In Section C, we show the detailed classification results of each category.

A. Details of Network Architecture

As shown in Figure 4 of the main paper, we design an augmentation network to automatically transform input training samples into augmented training samples. Figure 1 below illustrates the detailed architecture of the augmentor network. Note that we produce the shape-wise transformation $\mathcal{M}$ by first output a 4-dimension quaternion and then converting it into a 3×3 matrix.

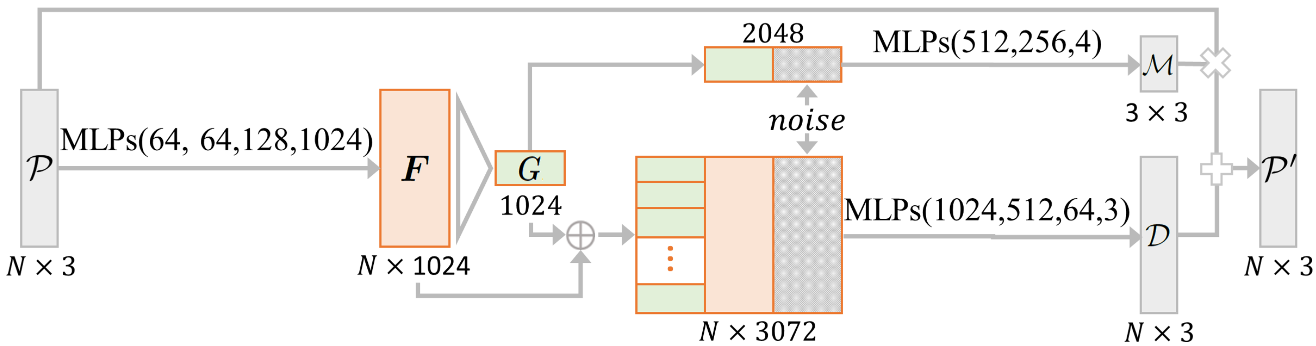

Figure 1. The detailed architecture of the augmentor network.

B. Additional Visual Comparisons of Shape Retrieval

In this section, we show more shape retrieval examples in Figures 2-5. Compared with the original PointNet [1], which is equipped with conventional data augmentation, the augmented version of PointNet with our PointAugment generates more plausible shape retrieval results.

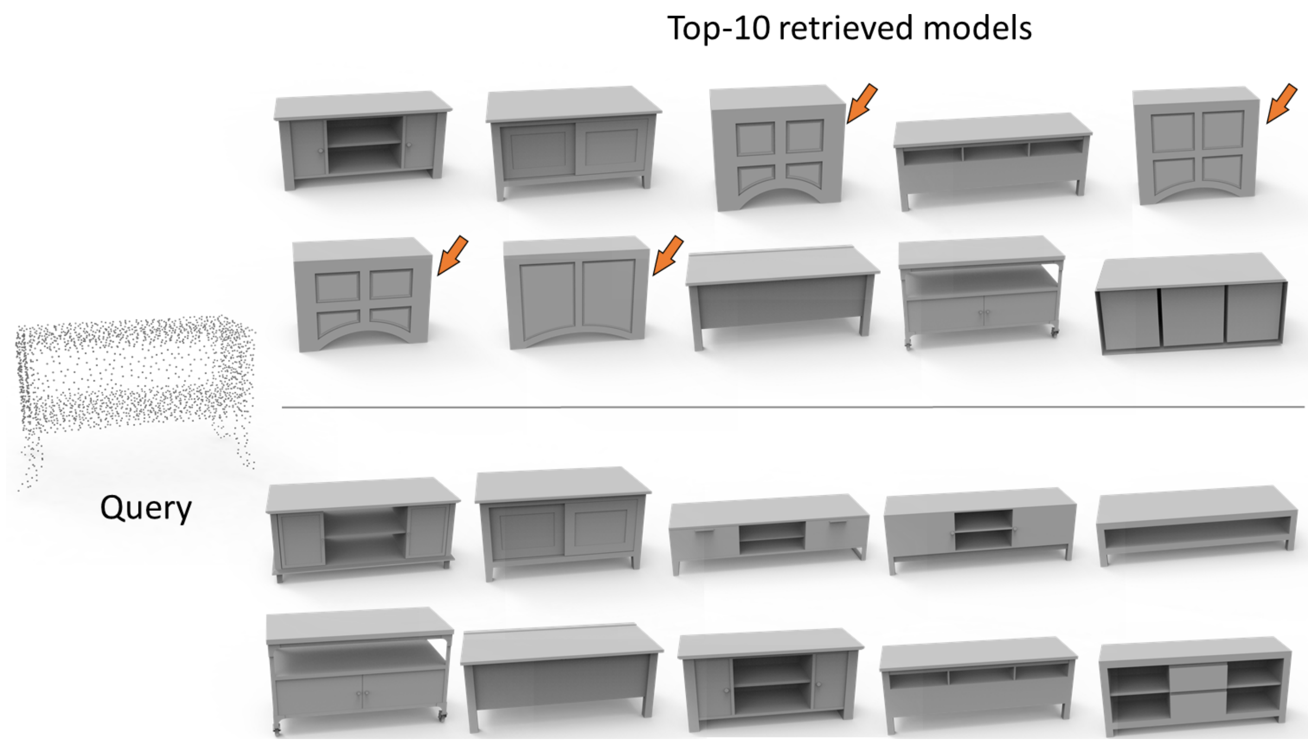

Figure 2. Shape retrieval results on the ModelNet40 dataset [5]. For each query shape shown on the left, we present two groups of Top-10 retrieval results: the top group uses PointNet [1] and the bottom group uses PointNet+PointAugment. Note that the obviously-wrong retrieval results are marked with red arrows (1/4).

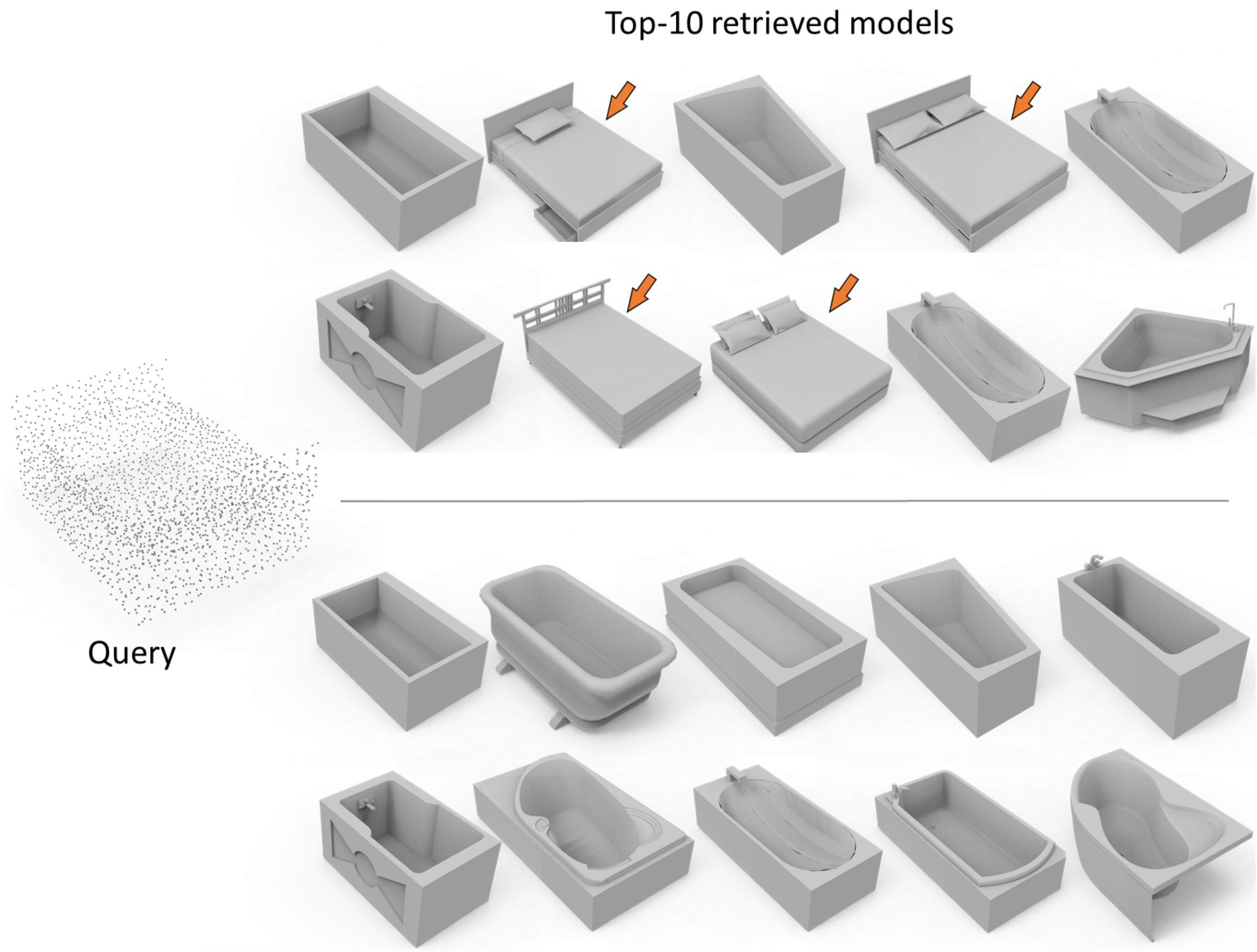

Figure 3. Shape retrieval results on the ModelNet40 dataset [5]. For each query shape shown on the left, we present two groups of Top-10 retrieval results: the top group uses PointNet [1] and the bottom group uses PointNet+PointAugment. Note that the obviously-wrong retrieval results are marked with red arrows (2/4).

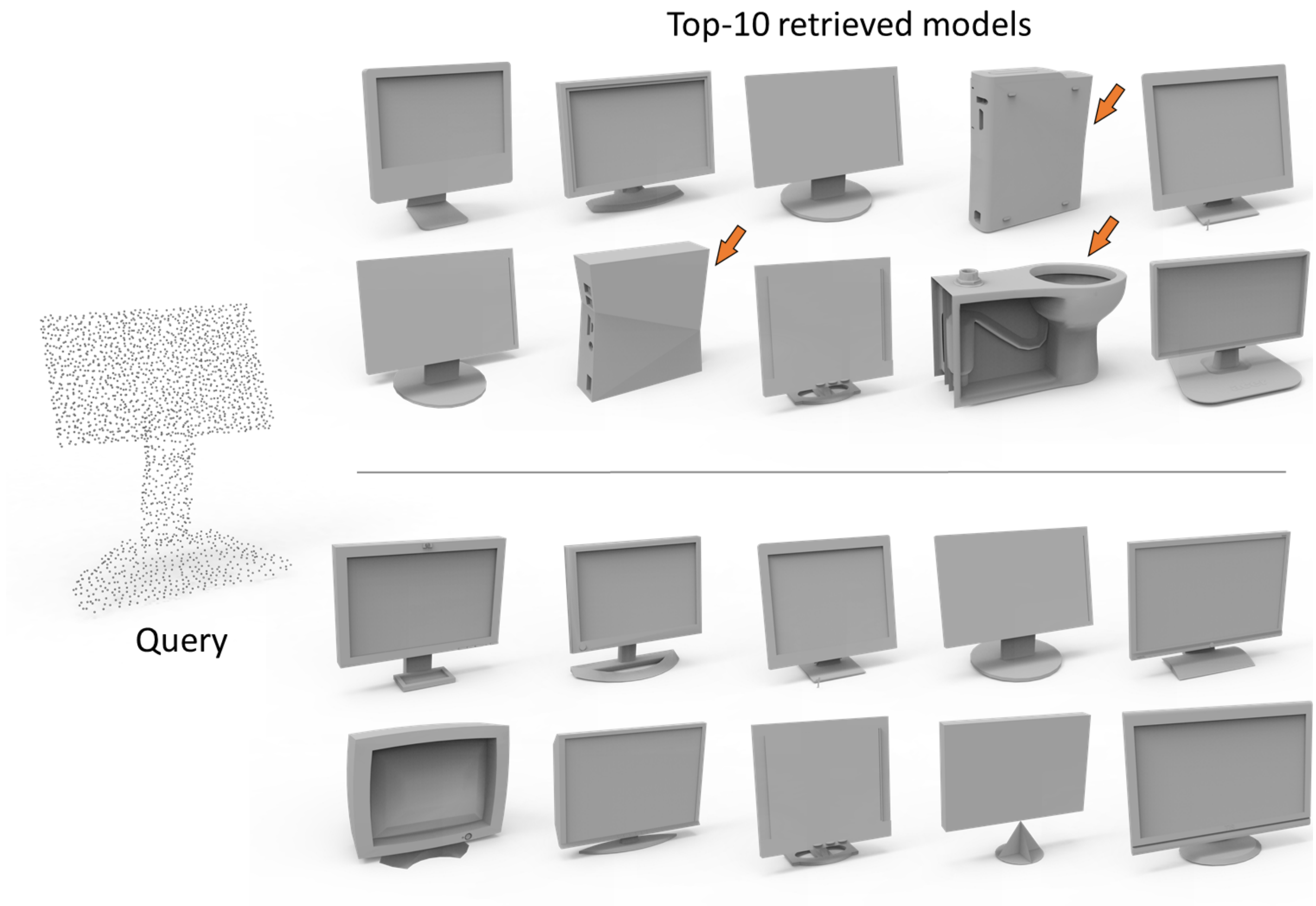

Figure 4. Shape retrieval results on the ModelNet40 dataset [5]. For each query shape shown on the left, we present two groups of Top-10 retrieval results: the top group uses PointNet [1] and the bottom group uses PointNet+PointAugment. Note that the obviously-wrong retrieval results are marked with red arrows (3/4).

Figure 5. Shape retrieval results on the ModelNet40 dataset [5]. For each query shape shown on the left, we present two groups of Top-10 retrieval results: the top group uses PointNet [1] and the bottom group uses PointNet+PointAugment. Note that the obviously-wrong retrieval results are marked with red arrows (4/4).

C. Classification Breakdowns across Categories

In this section, we provide the classification details for PointNet [1] (PN), PointNet++ [2] (PN2), and DGCNN [4] (DG), trained with conventional DA or with our PointAugment (PA) using ModelNet40 (MN40) [5] and SHREC16 (SR16) [3] datasets, respectively. The results are summarized in Tables 1 and 2 below. In each table, we present the number of training and testing samples in each category, and the classification accuracy of each category. We also provide in the last row the overall results of all categories, including the total number of training and testing samples, and the overall classification accuracy of each classifier.

Table 1. Detailed classification accuracies on MN40 [5] for each classifier with conventional DA or with our PA.

MN40	#training	#testing	PN	PN+PA	PN2	PN2+PA	DG	DG+PA
airplane	626	100	1.000	1.000	1.000	1.000	1.000	1.000
bathtub	106	50	0.900	0.900	0.900	0.920	0.880	1.000
bed	515	100	0.980	0.980	0.980	0.980	0.990	0.990
bench	173	20	0.750	0.700	0.750	0.750	0.700	0.800
bookshelf	572	100	0.960	0.980	0.970	0.980	0.970	0.950
bottle	335	100	0.950	0.960	0.970	0.970	0.990	0.970
bowl	64	20	0.800	0.900	0.850	1.000	1.000	0.950
car	197	100	0.970	0.980	0.990	0.980	1.000	1.000
chair	889	100	0.990	0.990	0.960	0.960	0.990	0.970
cone	167	20	0.900	0.950	1.000	0.950	0.950	0.950
cup	79	20	0.600	0.700	0.600	0.850	0.800	0.600
curtain	138	20	0.900	0.900	1.000	0.950	0.900	0.950
desk	200	86	0.895	0.930	0.942	0.884	0.907	0.942
door	109	20	0.900	0.900	0.700	0.750	0.800	0.950
dresser	200	86	0.826	0.919	0.826	0.872	0.849	0.895
flower_pot	149	20	0.100	0.200	0.100	0.100	0.000	0.100
glass_box	171	100	0.890	0.910	0.960	0.950	0.950	0.940
guitar	155	100	0.970	1.000	1.000	1.000	1.000	1.000
keyboard	145	20	0.800	0.950	1.000	0.950	0.950	1.000
lamp	124	20	0.800	0.950	0.800	0.850	0.850	0.900
laptop	149	20	1.000	1.000	1.000	1.000	1.000	1.000
mantel	284	100	0.980	0.960	0.950	0.940	0.980	0.980
monitor	465	100	0.970	0.990	1.000	0.990	1.000	0.990
night_stand	200	86	0.837	0.767	0.512	0.814	0.756	0.767
person	88	20	0.900	0.850	0.850	0.950	0.950	0.950
piano	231	100	0.870	0.880	0.920	0.930	0.910	0.940
plant	240	100	0.770	0.790	0.960	0.950	0.960	0.890
radio	104	20	0.500	0.650	0.450	0.900	0.700	0.750
range_hood	115	100	0.950	0.920	0.970	0.960	0.940	0.980
sink	128	20	0.700	0.800	0.850	0.950	0.850	0.950
sofa	680	100	0.980	1.000	0.990	0.990	1.000	1.000
stairs	124	20	0.850	0.850	0.950	0.950	0.850	0.950
stool	90	20	0.700	0.750	0.850	0.850	0.850	0.850
table	392	100	0.840	0.820	0.780	0.840	0.820	0.870
tent	163	20	0.900	0.950	0.900	0.950	0.950	0.950
toilet	344	100	1.000	0.980	1.000	1.000	0.990	1.000
tv_stand	267	100	0.800	0.820	0.880	0.890	0.900	0.900
vase	475	100	0.790	0.860	0.890	0.870	0.820	0.870
wardrobe	87	20	0.600	0.650	0.450	0.750	0.650	0.700
xbox	103	20	0.800	0.950	0.950	0.900	0.900	1.000
total/mean	9843	2468	0.892	0.909	0.907	0.929	0.922	0.934

Table 2. Detailed classification accuracies on SR16 [3] for each classifier with conventional DA or with our PA.

SR16	#training	#testing	PN	PN+PA	PN2	PN2+PA	DG	DG+PA
airplane	2831	405	1.000	1.000	1.000	1.000	1.000	1.000
ashcan	240	34	0.618	0.618	0.559	0.500	0.559	0.500
bag	58	8	0.625	0.750	0.000	0.750	0.750	0.750
basket	79	11	0.273	0.182	0.000	0.182	0.000	0.364
bathtub	599	85	0.824	0.918	0.894	0.906	0.929	0.906
bed	163	23	0.783	0.913	0.870	1.000	0.957	1.000
bench	1269	181	0.481	0.625	0.630	0.685	0.663	0.646
bicycle	41	6	1.000	1.000	0.000	1.000	0.883	1.000
birdhouse	51	7	0.286	0.571	0.143	0.857	0.429	0.857
bookshelf	316	45	0.578	0.778	0.800	0.778	0.800	0.644
bottle	348	50	0.700	0.720	0.800	0.740	0.780	0.720
bowl	130	18	0.944	0.944	0.944	0.889	0.889	0.944
bus	657	94	0.947	0.989	0.968	0.968	0.968	0.957
cabinet	1099	157	0.764	0.771	0.580	0.675	0.688	0.834
camera	79	11	0.000	0.000	0.000	0.182	0.000	0.091
can	75	11	0.800	0.636	0.182	0.636	0.545	0.727
cap	39	5	0.400	0.800	0.200	0.800	0.200	0.600
car	2473	353	0.992	0.994	1.000	1.000	0.994	0.997
chair	4744	678	0.948	0.950	0.953	0.959	0.944	0.962
clock	455	65	0.354	0.523	0.631	0.662	0.569	0.600
keyboard	45	7	0.714	0.714	0.857	0.857	0.857	0.857
dishwasher	65	9	0.333	0.556	0.000	0.444	0.000	0.556
display	765	109	0.844	0.908	0.807	0.899	0.899	0.954
earphone	51	7	0.571	0.714	0.286	0.714	0.857	0.714
faucet	520	75	0.667	0.867	0.853	0.907	0.933	0.933
file cabinet	208	30	0.167	0.200	0.000	0.433	0.000	0.500
guitar	557	80	0.963	0.975	0.975	0.988	0.988	0.988
helmet	113	16	0.688	0.750	1.000	1.000	0.750	1.000
jar	417	59	0.644	0.492	0.458	0.542	0.458	0.746
knife	296	43	0.953	0.977	0.953	0.953	0.953	0.977
lamp	1622	232	0.871	0.875	0.828	0.901	0.789	0.914
laptop	322	46	0.978	0.978	1.000	0.978	0.978	0.978
loudspeaker	1117	160	0.550	0.637	0.675	0.675	0.675	0.700
mailbox	65	10	0.400	0.600	0.000	0.400	0.600	0.600
microphone	46	7	0.000	0.143	0.000	0.571	0.000	0.429
microwave	106	15	0.533	0.933	0.533	1.000	0.867	0.933
motorcycle	235	34	0.912	1.000	1.000	1.000	1.000	1.000
mug	149	22	0.727	0.955	0.955	0.955	0.955	0.955
piano	167	24	0.375	0.458	0.458	0.542	0.500	0.542
pillow	67	9	0.889	0.889	0.889	1.000	0.889	1.000
pistol	214	31	1.000	0.968	1.000	1.000	1.000	0.935
pot	421	60	0.133	0.367	0.217	0.467	0.283	0.417
printer	116	16	0.250	0.875	0.562	0.812	0.625	0.812
remote	46	6	0.000	0.000	0.000	0.000	0.000	0.333
rifle	1661	237	0.954	0.966	0.954	0.954	0.962	0.987
rocket	59	9	0.222	0.556	0.000	0.778	0.111	0.667
skateboard	106	15	0.933	0.933	0.800	0.933	0.933	0.933
sofa	2221	317	0.912	0.918	0.902	0.902	0.902	0.902
stove	152	22	0.045	0.591	0.045	0.682	0.091	0.636
table	5905	843	0.980	0.967	0.943	0.972	0.976	0.983
telephone	728	105	0.790	0.914	0.886	0.933	0.933	0.943
tower	93	13	0.231	0.923	0.000	1.000	0.077	0.846
train	272	39	0.564	0.923	0.879	0.949	0.897	0.974
vessel	1357	194	0.943	0.979	0.979	0.979	0.974	1.000
washer	118	17	0.118	0.588	0.176	0.706	0.471	0.647
total/mean	36148	5165	0.844	0.884	0.851	0.895	0.870	0.906

References

[1] C. R. Qi, H. Su, K. Mo, and L. J. Guibas. PointNet: Deep learning on point sets for 3D classification and segmentation. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 652–660, 2017. 3, 4, 5, 6, 7

[2] C. R. Qi, L. Yi, H. Su, and L. J. Guibas. PointNet++: Deep hierarchical feature learning on point sets in a metric space. In *Conference and Workshop on Neural Information Processing Systems (NeurIPS)*, pages 5099–5108, 2017. 7

[3] M. Savva, F. Yu, H. Su, M. Aono, B. Chen, D. Cohen-Or, W. Deng, H. Su, S. Bai, X. Bai, et al. SHREC16' track: largescale 3D shape retrieval from ShapeNet Core55. In *Proceedings of the eurographics workshop on 3D object retrieval*, pages 89–98, 2016. 7, 8

[4] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon. Dynamic graph CNN for learning on point clouds. *ACM Trans. on Graphics*, 38(5):146:1–12, 2019. 7

[5] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao. 3D ShapeNets: A deep representation for volumetric shapes. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 1912–1920, 2015. 3, 4, 5, 6, 7

The End