Point MixSwap: Attentional Point Cloud Mixing via Swapping Matched Structural Divisions (Supplementary Material)

Ardian Umam[1*], Cheng-Kun Yang[2*], Yung-Yu Chuang[2], Jen-Hui Chuang[1], and Yen-Yu Lin[1,3]

[1]National Yang Ming Chiao Tung University, Taiwan
ardianumam.ee09@nycu.edu.tw, jchuang@cs.nctu.edu.tw, lin@cs.nycu.edu.tw
[2]National Taiwan University, Taiwan
d08922002@csie.ntu.edu.tw, cyy@csie.ntu.edu.tw
[3]Academia Sinica, Taiwan

A Experiments on Other Backbones

In addition to the backbones which are typically investigated in existing methods [1,2,3,4], PointNet [6] and DGCNN [7], we also experiment with two modern backbones, 3D-GCN [5] and GDANet [8]. Table A presents the results. There is a similar trend on performance improvement as shown in Table 1 of the main paper, which demonstrates the effectiveness of the proposed data augmentation approach in different settings.

B Performance Varies among Classes under Different Division Numbers

We provide per-class accuracy scores under different division numbers in Table B, and present the common classes shared by the ModelNet40 (M40) and ScanObjectNN (SON) datasets. Although different classes have their optimal numbers of divisions, the class-specific variances are not significant. Furthermore, the proposed method provides positive performance gains for all classes under the majority of division numbers, as compared to the baseline without the use of augmented data. Therefore, our method is not sensitive to division numbers and is effective.

C More Mixup Samples

The section provides several examples of mixups. First, we present the mixup samples in Figure A, which corresponds to the last two rows of Figure 7 in the main paper. In these examples, the source samples have varying poses and the mixup samples are generated before the alignment is performed. Second,

[*] The authors have equal contribution to this work

Table A: Accuracy scores of the proposed Point MixSwap on 20%, 50%, and 100% of the ModelNet40 (M40) and ModelNet10 (M10) datasets using other backbones.

Method	Rate 20%		Rate 50%		Rate 100%	
	M40	M10	M40	M10	M40	M10
3D-GCN	87.3	93.1	91.1	94.2	92.1	94.7
3D-GCN + Ours	**90.0** (2.7↑)	**94.6** (1.5↑)	**92.4** (1.3↑)	**94.9** (0.7↑)	**93.0** (0.9↑)	**95.4** (0.7↑)
GDANet	89.4	93.3	91.7	94.5	93.8	95.6
GDANet + Ours	**91.4** (2.0↑)	**94.6** (1.3↑)	**92.9** (1.2↑)	**95.1** (0.6↑)	**94.0** (0.2↑)	**96.1** (0.5↑)

Table B: The class-specific accuracy scores and variances, in %, of the proposed Point MixSwap on ModelNet40 (M40) and ScanObjectNN (SON) datasets, associated with Table 3(b) in the main paper, with different numbers of divisions.

Dataset	Variable	Bed	Chair	Desk	Display	Door	Shelf	Sink	Sofa	Table	Toilet
M40	Baseline	92.0	91.0	83.6	95.0	88.0	86.0	80.0	90.0	84.0	91.0
	2 divisions	**96.0**	94.0	**87.2**	98.0	91.0	89.0	82.0	94.0	**88.0**	93.0
	3 divisions	95.0	**96.0**	86.5	**100.0**	**92.0**	**91.0**	**83.0**	**96.0**	87.0	**94.0**
	4 divisions	94.0	95.0	86.5	99.0	90.0	90.0	**83.0**	95.0	87.0	**94.0**
	5 divisions	95.0	94.0	86.5	99.0	91.0	89.0	**83.0**	95.0	87.0	93.0
	Variance	0.67	0.92	0.12	0.67	0.67	0.92	0.25	0.67	0.25	0.33
SON	Baseline	68.8	93.3	61.8	72.2	93.4	76.4	55.3	81.4	77.3	55.6
	2 divisions	**72.9**	95.7	**64.6**	75.5	94.4	80.2	57.4	**83.7**	78.8	58.4
	3 divisions	72.7	**96.2**	**64.6**	**76.2**	**95.2**	**81.6**	**58.0**	83.3	**79.6**	**58.8**
	4 divisions	72.7	95.2	**64.6**	75.1	94.8	80.8	57.4	82.9	78.8	57.8
	5 divisions	71.6	95.7	64.0	75.5	94.8	81.2	57.0	83.3	78.0	58.3
	Variance	0.35	0.17	0.09	0.21	0.11	0.36	0.17	0.11	0.43	0.17

mixswap samples from other categories are shown in Figure B, from top to bottom are the categories of person, car, cup, bathtub, and toilet. For both Figures A and B, the division numbers are set to 2 and 3 for subfigures (i) and (ii), respectively. In each setting, the source sample pair/triplet is shown on the left and the resulting mixswap samples are shown on the right.

Figures A and B illustrate that Point MixSwap successfully and consistently identifies structural divisions across samples of the same category. Thus, swapping one or a few divisions can result in diverse and structure-preserved samples. With regard to categories in which there is only one major structural division, such as bathtubs and cars, Point MixSwap can still generate structure-preserved samples by utilizing the cross correspondence between structural divisions.

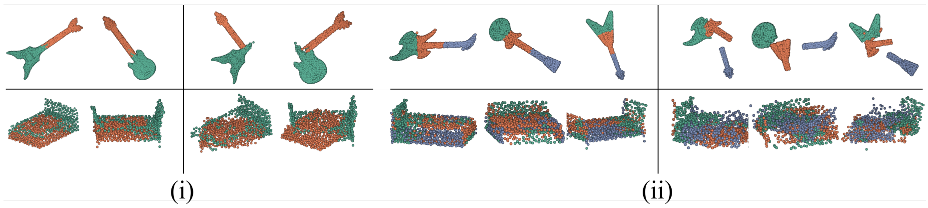

Fig. A: Mixup examples generated by Point MixSwap with (i) two divisions and (ii) three divisions. In each subfigure, the second column displays the generated mixup samples from the source samples in the first column before the alignment is applied. Note that the mixup samples produced after applying alignment are already provided in Figure 7 of the main paper. Each point is colored according to its division.

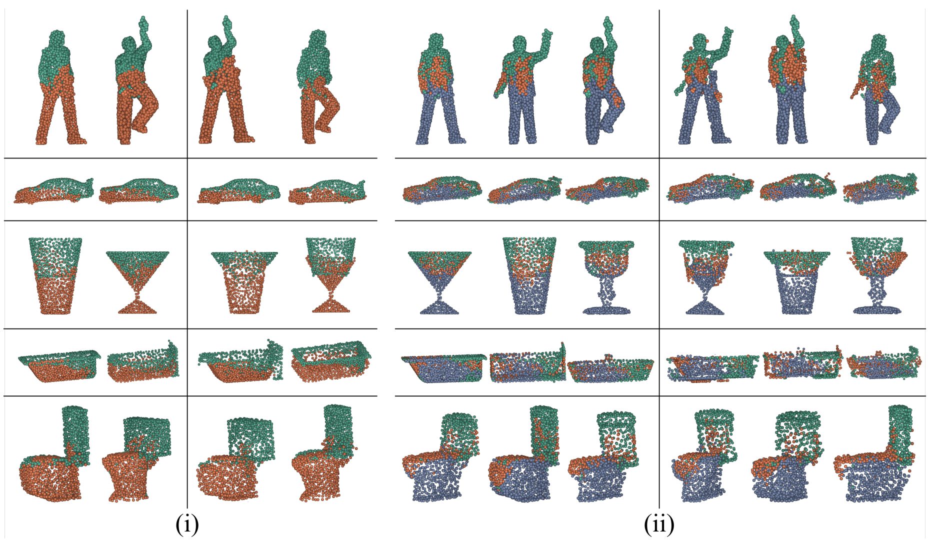

Fig. B: More examples of Point MixSwap results on the ModelNet40 dataset with (i) two divisions and (ii) three divisions. In each subfigure, the second column depicts the generated mixup samples from the original samples, which have similar poses as shown in the first column. Each point has a corresponding color indicating its division.

References

1. Chen, Y., Hu, V.T., Gavves, E., Mensink, T., Mettes, P., Yang, P., Snoek, C.G.: Pointmixup: Augmentation for point clouds. In: ECCV. pp. 330–345 (2020) 1
2. Kim, S., Lee, S., Hwang, D., Lee, J., Hwang, S.J., Kim, H.J.: Point cloud augmentation with weighted local transformations. In: ICCV (2021) 1
3. Lee, D., Lee, J., Lee, J., Lee, H., Lee, M., Woo, S., Lee, S.: Regularization strategy for point cloud via rigidly mixed sample. In: CVPR (2021) 1

4. Li, R., Li, X., Heng, P.A., Fu, C.W.: Pointaugment: an auto-augmentation framework for point cloud classification. In: CVPR (2020) 1
5. Lin, Z.H., Huang, S.Y., Wang, Y.C.F.: Convolution in the cloud: Learning deformable kernels in 3d graph convolution networks for point cloud analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1800–1809 (2020) 1
6. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: Deep learning on point sets for 3d classification and segmentation. In: CVPR (2017) 1
7. Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph cnn for learning on point clouds. TOG (2019) 1
8. Xu, M., Zhang, J., Zhou, Z., Xu, M., Qi, X., Qiao, Y.: Learning geometry-disentangled representation for complementary understanding of 3d object point cloud. arXiv preprint arXiv:2012.10921 **2** (2021) 1