Supplementary Material to: Enhancing Robustness to Noise Corruption for Point Cloud Recognition via Spatial Sorting and Set-Mixing Aggregation Module

Dingxin Zhang[1], Jianhui Yu[1], Tengfei Xue[1], Chaoyi Zhang[1], Dongnan Liu[1], and Weidong Cai[1]

School of Computer Science, University of Sydney, NSW 2006, Australia
{dzha2344, jianhui.yu, txue4133, chaoyi.zhang, dongnan.liu, tom.cai}@sydney.edu.au

1 Model Structure

The proposed Set-Mixer module is incorporated into PointNet++ structure and the detailed structure of our model is represented using the formats SA(num of sampled points, number of points sampled in a patch K, [Fully connected layer channel]) and Set-Mixer(num of input channel c, K, dropout rate):

$$SA(512, 32, [64, 64, 64], \text{Set-Mixer}(64, 32, 0.2)) \rightarrow$$
$$SA(128, 64, [128, 128, 128], \text{Set-Mixer}(128, 64, 0.2)) \rightarrow$$
$$SA(1, 128, [256, 512, 512], \text{Set-Mixer}(512, 128, 0.2)) \rightarrow$$
$$FC(512, 0.5) \rightarrow FC(256, 0.5) \rightarrow FC(40).$$

Here, the decimal between 0 and 1 inside $FC()$ indicates the dropout rate.

2 Sorting Strategy

In the proposed Set-Mixer, we perform feature sorting three times independently along the xyz-axis to enrich spatial information. We also evaluated alternative sorting strategies, including sorting once or twice, with the comprehensive results illustrated in Table S1.

As demonstrated by Table S1, all sorting methods exhibit superior performance compared to PointNet++. Notably, conducting multiple sorting operations yields a significant improvement. In general, increasing the number of sorting operations would enrich the spatial information of point sets, leading to enhanced performance and robustness.

3 Spatial Center Visualization

In Fig. S1, we visualize both the query points and the spatial centers generated by Set-Mixer in layers 1 and 2. Through this visualization, we can observe that

Table S1. Classification error rate on ModelNet40-C with different sorting strategies.

	APS			EDS	PCS			
Model	x-axis	y-axis	z-axis	distance	XY-XZ-YZ plane	ER_{clean}	ER_{noise}	RmCE
A	✓	✓	✓			**8.1**	**10.0**	0.13
B	✓	✓				8.8	10.9	0.14
C	✓		✓			9.5	11.4	0.13
D		✓	✓			8.7	10.5	0.13
E	✓					9.9	13.0	0.21
F		✓				9.6	13.5	0.27
G			✓			9.4	12.6	0.22
H				✓		8.2	20.2	0.83
I					✓	9.6	11.0	**0.10**
J				✓	✓	8.7	11.3	0.18
K	✓	✓	✓	✓		8.2	10.1	0.13
L	✓	✓	✓		✓	8.8	10.2	**0.10**
M	✓	✓	✓	✓	✓	8.3	10.3	0.14

the spatial center not only provides a more effective representation of the spatial positions of point sets but also exhibits greater stability.

4 Noisy Feature Visualization

To demonstrate the impact of noisy points on the extracted features, we employ visualizations to depict the degree of variation in the second layer features of both Set-Mixer and PointNet++. As illustrated in Fig. S2, when subjected to noise with severity 3, PointNet++ exhibits a significantly greater degree of feature variation. We further analyze the differences across five noise types on a randomly selected plane sample and a chair sample, as illustrated in Fig. S3 and Fig. S4. Moreover, we visualize samples from more categories, introducing Impulse noise at severity 3, illustrated in Fig. S5.

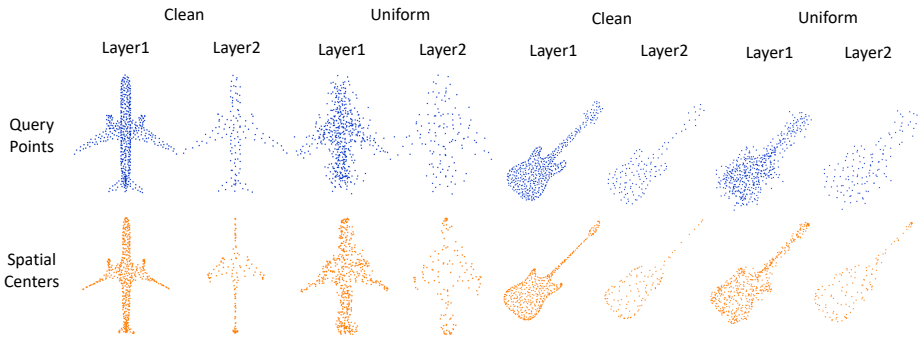

Fig. S1. Visualization of query points and spatial center points of layer 1 (512 sets, 32 neighbours) and layer 2 (128 sets, 64 neighbours) of Set-Mixer. Query points are depicted in blue, while the spatial centers are marked in orange.

Fig. S2. Visualization of feature change magnitudes. Blue indicates no change, while red indicates a high degree of change. The left graph of the top line is the result of the clean sample, while the middle and right graphs illustrate the results of PointNet++ and Set-Mixer, respectively, with the impulse of severity 3.

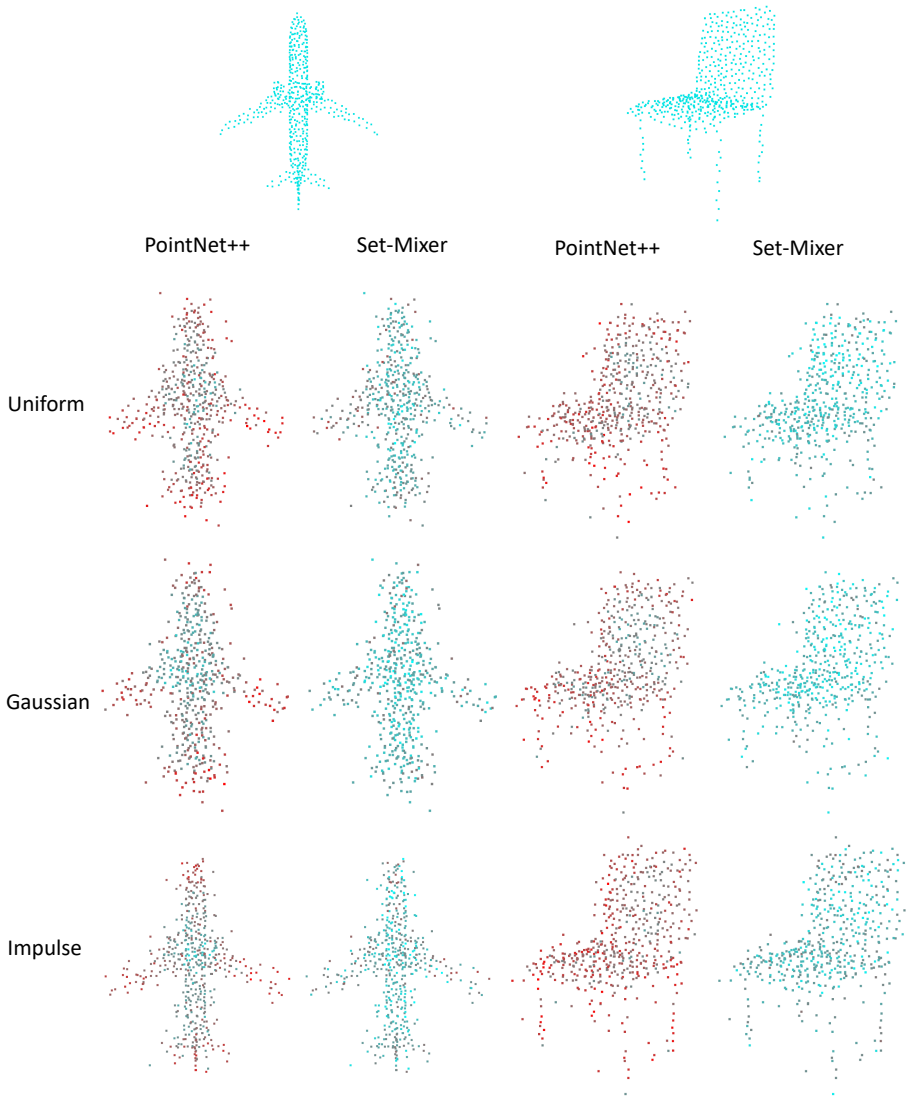

Fig. S3. Visualization of feature change magnitudes compared to clean input. Blue indicates no change, while red indicates a high degree of change. The two graphs at the top represent the results for the original clean sample. The severity is set to 3.

Fig. S4. Visualization of feature change magnitudes compared to clean input. Blue indicates no change, while red indicates a high degree of change. The two graphs at the top represent the results for the original clean sample. The severity is set to 3.

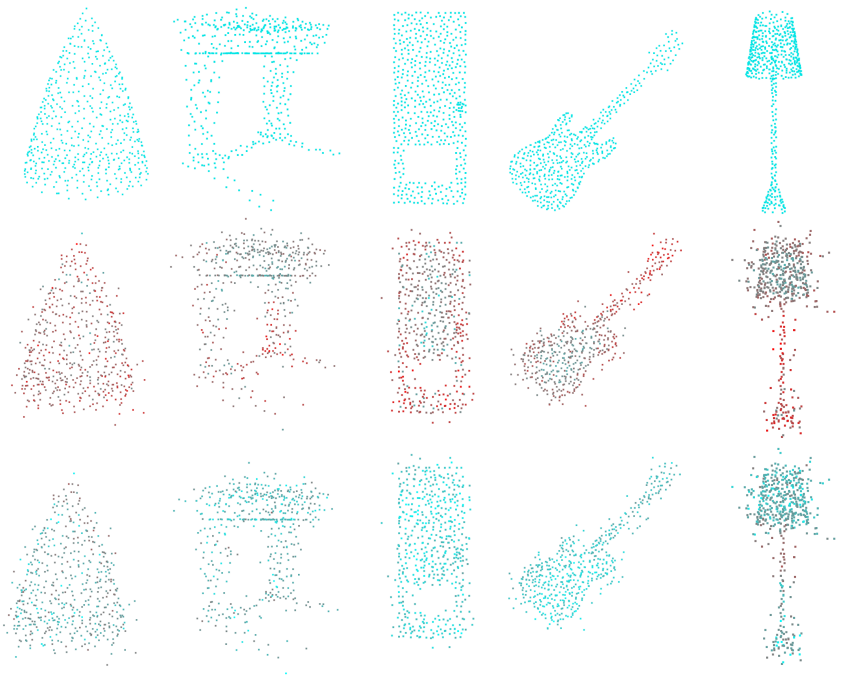

Fig. S5. Visualization of feature change magnitudes compared to clean input. Blue indicates no change, while red indicates a high degree of change. The graphs in the top row represent the results for a clean sample, the second row displays graphs from PointNet++, and the third row exhibits graphs from Set-Mixer. The noise type is Impulse, and the severity is set to 3.