Supplementary Material of Symmetric Shape-Preserving Autoencoder for Unsupervised Real Scene Point Cloud Completion

In the supplementary material, we first show more details about network, metric, and multi-category loss. Next, we conduct more experiments including the robustness of noises and the corrupted inputs, and the error bar of the training experiments. Finally, more completion on real objects and more qualitative comparisons are given to show the advantages of our USSPA.

1. More Details

1.1. Network Details

Inspired by PCN encoder [8], we design the structures of the encoders $E_R$, $E_A$, and the point discriminator $D_p$ in our USSPA as shown in Figure 1. The parameter of the weight-shared MLPs of our encoders are $[3, 128, 256]$ and $[512, 512]$. And the parameters of the weight-shared MLPs of our point discriminator are $[3, 128, 256]$ and $[512, 256]$. Besides, to get the probability, we also employ an MLP in our point discriminator after extracting the global feature. The parameter of this MLP is $[256, 64, 64, 1]$.

The decoders $D_R$, $D_A$, $D_{sl}$ and the feature discriminator $D_f$ are MLPs with parameters $[512, 512, 512, 1024, 1024, 1536]$, $[512, 512, 512, 1024, 1024, 1536]$, $[512, 128, 128, 2]$ and $[512, 64, 64, 1]$ separately. The parameters of the weight-shared MLPs of the upsampling refinement module are $[3, 256, 256, 256]$, $[512, 256, 256, 256]$, $[512, 512]$. And the parameters of the weight-shared MLPs in our USSPA as shown in Figure 1. The parameter of the MLP is $[515, 512, 512, 12]$.

For multi-category data, the channel numbers of the output layer in our discriminators are set to $11 = 10 + 1$ where 10 is the category number and 1 represents the fake label mentioned in our paper.

In practice, we alternately optimize the generator and the discriminator of our USSPA.

\[ cd_{11}^{11}(P_1, P_2) = \frac{1}{|P_1|} \sum_{x \in P_1} \min_{y \in P_2} ||x - y|| + \frac{1}{|P_2|} \sum_{x \in P_2} \min_{y \in P_1} ||x - y||. \]

As mentioned in [2], the Chamfer Distance indicator is sometimes misleading because of its sensitivity to outliers. We then further take advantage of F-score [2] to deliver more comprehensive evaluations. F-score is defined as:

\[ F_{score} = \frac{2P(d)R(d)}{P(d) + R(d)}, \]

where $P$ is precision:

\[ P(d) = \frac{1}{|P_1|} \sum_{x \in P_1} \left[ \min_{y \in P_2} ||x - y|| < d \right], \]

and $R$ is recall:

\[ R(d) = \frac{1}{|P_2|} \sum_{y \in P_2} \left[ \min_{x \in P_1} ||x - y|| < d \right]. \]

Both $P$ and $R$ take a parameter $d$ to control the strictness. Smaller $d$ means a smaller neighborhood, which brings both

---

\(^1\)For a MLP $[c_1, c_2, ..., c_n]$ means that there are $n$ layers in MLP. The input layer has $c_1$ channels, the output layers have $c_n$ channels, and $c_2, ..., c_{n-1}$ denote the channel numbers of hidden layers.
lower precision and recall. As suggested, we set $d$ at 0.01 and 0.001 separately.

1.3. Multi-Category Loss Details

Let $D_p(P_{R_0})_i$ represents the probability that the prediction $P_{R_0}$ belongs to the $i$-th category, and $g$ denotes the ground truth label of the artificial input point cloud $P_A$. As we mentioned in the paper, the label $0$ denotes fake data while the label $l > 0$ denotes real data belonging to category $l$. Thus, for multi-category data with $K$ categories, we change the loss $L_{F→R}$, $L_R$ and $L_F$ into:

$$L_{F→R} = - \left( \log[1 - D_p(P_{R_0})_0] + \log[1 - D_p(P_{R_2})_0] \right) + \log[1 - D_f(f_{R})_0],$$

(5)

$$L_F = - \left( \log D_p(P_{R_0})_0 + \log D_p(P_{R_2})_0 \right) + \log D_f(f_{R})_0,$$

(6)

and

$$L_R = - \sum_{i=0}^K (g_i \log D_p(P_A)_i + g_i \log D_f(f_A)_i),$$

(7)

respectively.

2. More Experiments

2.1. Robustness of Noises

We conduct an experiment by adding random shifts to the points of input point clouds to evaluate the robustness of the noises of our method. For an input point cloud $P_R \in \mathbb{M}_{m_0 \times 3}$, we randomly generate a shift matrix $x_{shift} \in [-s, s]_{m_0 \times 3}$ through uniform sampling, where $s$ indicates the noise scale. Then, we add the shift matrix to the input point cloud by $P_R \leftarrow P_R + x_{shift}$ to randomly shift the points. The Chamfer Distance $cd_{l1}$ and F-Score $F_{score}^{0.01\%}$ of our prediction on chair category with different noise scales are shown in Figure 2. The results show the robustness of the noises of our method when the noise scale $s$ is lower than 0.1.

2.2. Robustness of Corrupted Inputs

To evaluate the robustness of corrupted inputs and verify the necessity of our symmetry learning module, we remove the points of input point cloud along the X-axis as shown in Figure 3. The completion results are nearly the same with different keeping ratios of input point clouds. These results show the robustness of corrupted inputs of our method which benefits from our carefully designed symmetry learning module. Our symmetry learning module can generate the symmetrical point cloud of input which fully leverages the existing information.

2.3. Error Bar of Training Experiments

The error bars of our training experiments on 10 categories are shown in Figure 4, which shows that the performance of our USSPA is stable and repeatable.

3. More Completion on Real Object

More completion on real objects by Unpaired [6], ShapeInv [9], and our USSPA are shown in Figure 5. The results of the TV on the first line show that our prediction is more uniform with the vertical right boundary. And the results of the table on the second line show that our prediction is cleaner with fewer outliers. More comparisons indicate the superiority and generalization ability of our USSPA on real objects.
4. More Qualitative Comparisons

More qualitative comparisons of our USSPA and other works [3–7,9] are shown in Figures 6, 7 and 8.

References


Figure 6. Visualization of completion results by our USSPA and other works. From top to bottom: chair, table, bookshelf, and TV.

Figure 7. Visualization of completion results by our USSPA and other works. From top to bottom: bed, cabinet, lamp, and sofa.
Figure 8. Visualization of completion results by our USSPA and other works. From top to bottom: trash bin, and tub.