Supplementary Material: Maximizing Mutual Information Module for Learning on Point Clouds

Anonymous Author(s)
Affiliation
Address
email

A Overview

This document provides additional information on the main paper, including technical details, extra visualization results, and more descriptions of our experiments. Aim to present more theories and experiments to support our method, the main contents of the supplementary material are as follows.

- Section B introduces and discusses the detailed structures of the point cloud learning networks applied in our work.
- Section C provides additional descriptions of our experiments, including a brief introduction of the existing works compared in the classification task and more visualization results in the point cloud completion task.

Moreover, we also provide part of the code in the ZIP file to illustrate the authenticity and accuracy of our experimental results.

B Neural Network Structure

This section introduces the detailed network structure of PointNet*, DGCNN*, and PCN* with the MMI module in the point cloud classification and completion tasks because we slightly change their architecture to fit the MMI module in our method. Note that Figure 1, Figure 2, and Figure 3 are the flowcharts, which illustrate the detailed structures of our model and highly revert the implementation codes to provide a reference for the reader.

B.1 Point Cloud Classification

PointNet*+MMI PointNet [3] first explores deep learning on point clouds directly, and it is the basis of many studies on point cloud learning. Figure 1 shows the detailed structure of PointNet*+MMI, including the baseline models (in the green shadow) and the MMI module (in the blue shadow). The low-level features f_L (also called shallow-layer features in the paper) and the high-level features f_H (also called the deep-layer features in the paper) are extracted from two different layers from PointNet*. In our experiments, the dimension of the extracted features is all with the same dimension (channels) of 64. Then, f_L and f_H form positive sample pairs and negative sample pairs through pooling, repeating, and pairing operations. Meanwhile, these sample pairs are distinguished by local and global features. Local MI applies MLP $\{64, 64, 1\}$ in the MI neural estimator while global MI employs MLP $\{128, 64, 32, 1\}$ with the activation function - softplus. Finally, the MI-based loss in the MMI module L_{MMI} is measured by the average of the sum of L_{gMI} and L_{lMI}.

DGCNN*+MMI DGCNN [6] considers the local neighborhood information using EdgeConv, a classical and stable method for learning point clouds. Hence, we apply it as another baseline model in our work. Also, we slightly modify its structure by adding one MLP layer, shown in Figure 2, and call it DGCNN*. Note that the main working steps of DGCNN*+MMI are similar to PointNet*+MMI.

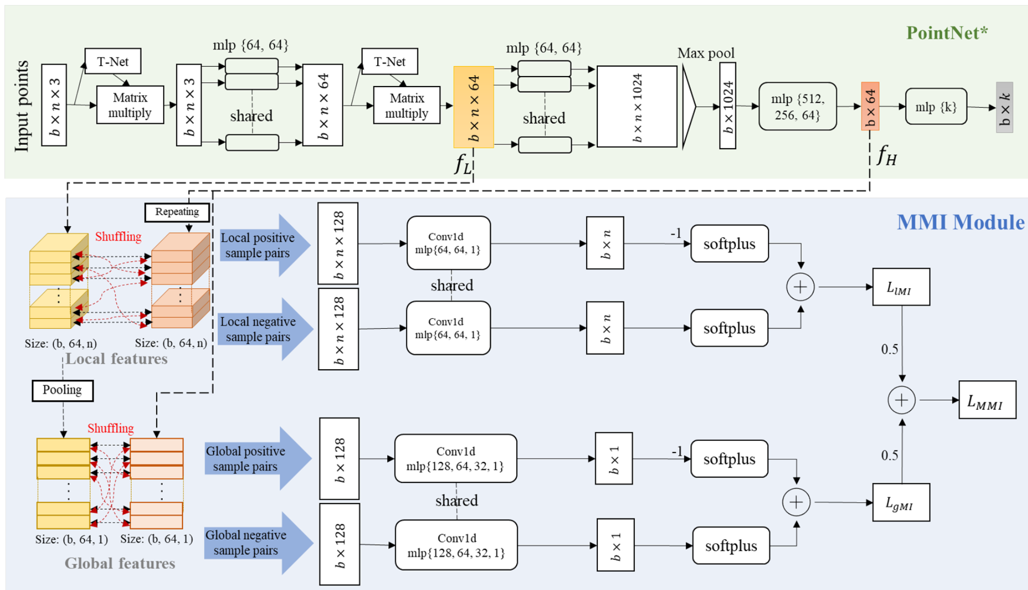

Figure 1: PointNet*+MMI

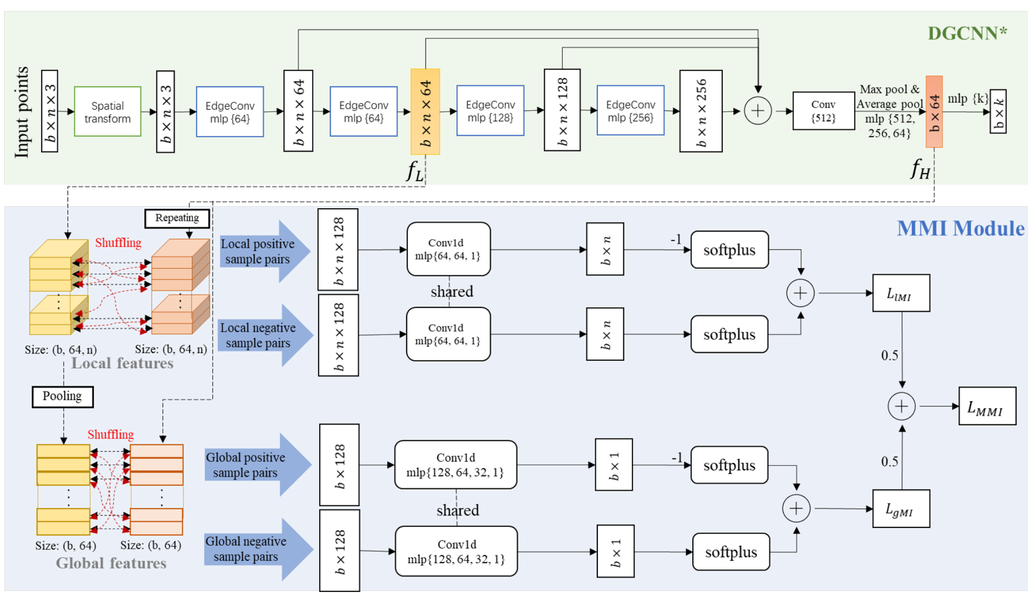

Figure 2: DGCNN*+MMI

B.2 Point Cloud Completion

PCN*+MMI PCN [8] is an encoder-decoder network to complete the partial point clouds by deep learning on the point-wise method. The encoder is an extended version of PointNet [3] and we also change its structure, naming it as PCN*. Particularly, f_L and f_H are the high-level features and the low-level features extracted from PCN* with the exact dimension of 256. Different from the above two classification models, additional layers of MLP $\{512, 256\}$ are adopted to obtain the f_H, shown in Figure 3.

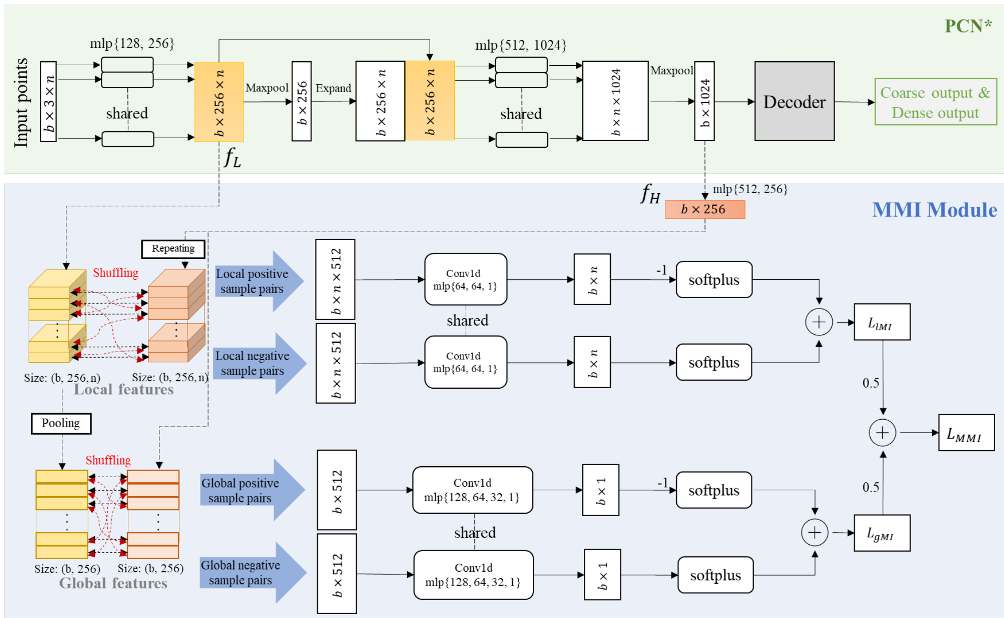

Figure 3: PCN*+MMI

B.3 Discussion

Structural Consistency By analyzing the structure of these three models, we prove that the MMI module can be used for various models with almost unchanging its structure, except slightly changing the dimension of the inputs (f_L and f_H). Structural consistency allows this module to be applied more quickly and efficiently in diverse application scenarios. We will test it on different models and tasks in future work, except the above three models (PointNet*+MMI, DGCNN*+MMI, and PCN*+MMI).

Synchronous Training The MMI module, a plug-and-play module with neural network architecture, does not require additional training. Conveniently, we can train it in parallel with the main network.

Dimensional Consistency The MMI module requires that the dimensions of the input are consistent. It can be a major drawback to this module, as it may be necessary to modify the main network to fit it.

C Additional Description on Experiments

C.1 Introduction to Contrast Models

In this subsection, we give a brief introduction to the contrast models in the classification tasks. Aim to show the superiority of our module, we compare it with some existing works (Section 4.2), including:

(1) FoldingNet [7], a point cloud auto-encoder based on deep grid deformation;

(2) PointNet++ [4], a hierarchical feature extraction network for point clouds;

(3) PCNN [1], a framework for applying convolutional neural networks to point clouds with two operators: extension and restriction;

(4) PointCNN [2], a generalization of typical CNNs to feature learning from point clouds with learning $\mathcal{X}$-transformation from the input points;

(5) DRNet [5], a network that represents point clouds from different resolutions;

(6) PointWeb [9], enhancing local neighborhood features for point cloud processing based on PointNet [3] and DGCNN [6].

C.2 More Visualization Results

To show the effect of the point cloud completion task, we provide additional visualization examples. Figure 3 illustrates four examples of point cloud completion by PCN* and PCN*+MMI. Note that the GT shapes with blue points are the ground truth with 2048 points; shapes with red points are the inputs of the whole model with missing various number of points; green point clouds represent the output point clouds with the completed shape, including coarse outputs (512 points) and dense outputs (2048 points).

The visualization results show that the performance of PCN*+MMI is better than PCN*, which implies that the MMI module can efficiently improve the capability of point cloud completion on the baseline model. Particularly, when the missing ratio is high, we can see the superior ability of PCN*+MMI to change a partial shape to a complete one. For example, in Figure 4, when the mission ratio is 75%, the outputs of PCN*+MMI are more accessible to distinguish than that of PCN* in the two shapes on the top row. The car shape in the bottom row is hard to complete in both PCN*+MMI and PCN*, but the latter's situation is obviously worse. The desk shape in the bottom row is completed well in both models, but PCN* is worse than that of PCN*+MMI.

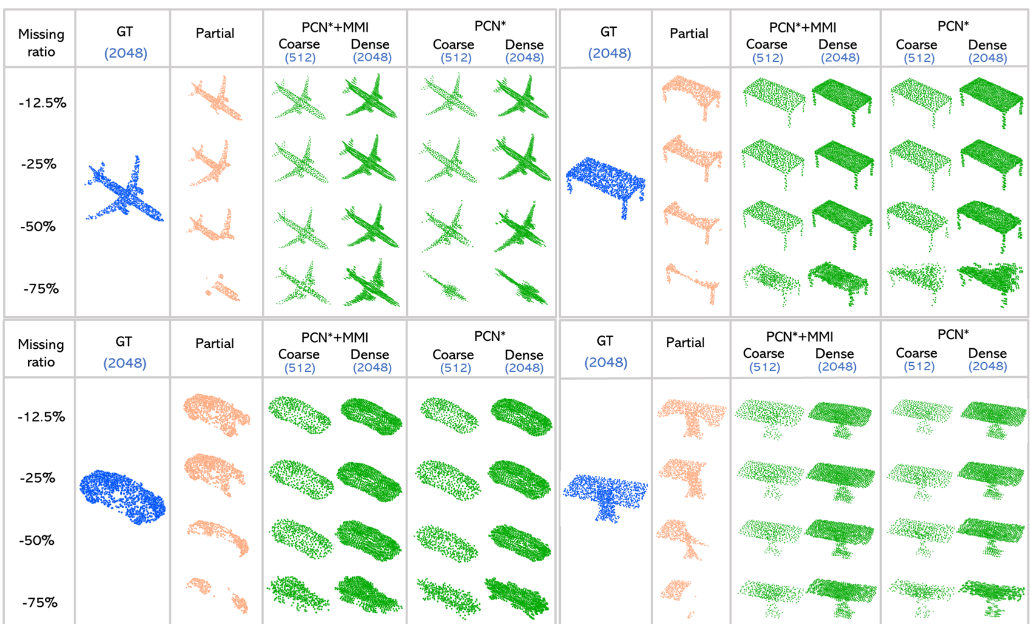

Figure 4: **Visualization Examples on Completion.**

References

[1] M. Atzmon, H. Maron, and Y. Lipman. Point convolutional neural networks by extension operators. *ACM Transactions on Graphics (TOG)*, 37(4):1–12, 2018.

[2] Y. Li, R. Bu, M. Sun, W. Wu, X. Di, and B. Chen. Pointcnn: Convolution on x-transformed points. *Advances in Neural Information Processing Systems*, 31:820–830, 2018.

[3] C. R. Qi, H. Su, K. Mo, and L. J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 652–660, 2017.

[4] C. R. Qi, L. Yi, H. Su, and L. J. Guibas. Pointnet++ deep hierarchical feature learning on point sets in a metric space. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 5105–5114, 2017.

[5] S. Qiu, S. Anwar, and N. Barnes. Dense-resolution network for point cloud classification and segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 3813–3822, 2021.

[6] 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 Transactions On Graphics (tog)*, 38(5):1–12, 2019.

[7] Y. Yang, C. Feng, Y. Shen, and D. Tian. Foldingnet: Point cloud auto-encoder via deep grid deformation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 206–215, 2018.

[8] W. Yuan, T. Khot, D. Held, C. Mertz, and M. Hebert. Pcn: Point completion network. In *2018 International Conference on 3D Vision (3DV)*, pages 728–737. IEEE, 2018.

[9] H. Zhao, L. Jiang, C.-W. Fu, and J. Jia. Pointweb: Enhancing local neighborhood features for point cloud processing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5565–5573, 2019.