PPSAN: PERCEPTUAL-AWARE 3D POINT CLOUD SEGMENTATION VIA ADVERSARIAL LEARNING

Hongyan Li, Zhengxing Sun, Yunjie Wu, Bo Li*

State Key Laboratory for Novel Software Technology, Nanjing University

Motivation

- The common used per-point loss in existing methods focuses on the average geometric similarity, but ignores the spatial consistency constraint and cannot capture the visual perceptual differences between segmentation output and ground truth.
- Two segmentation results with the same geometric similarity may have distinctly different visual perception quality.

Introduction

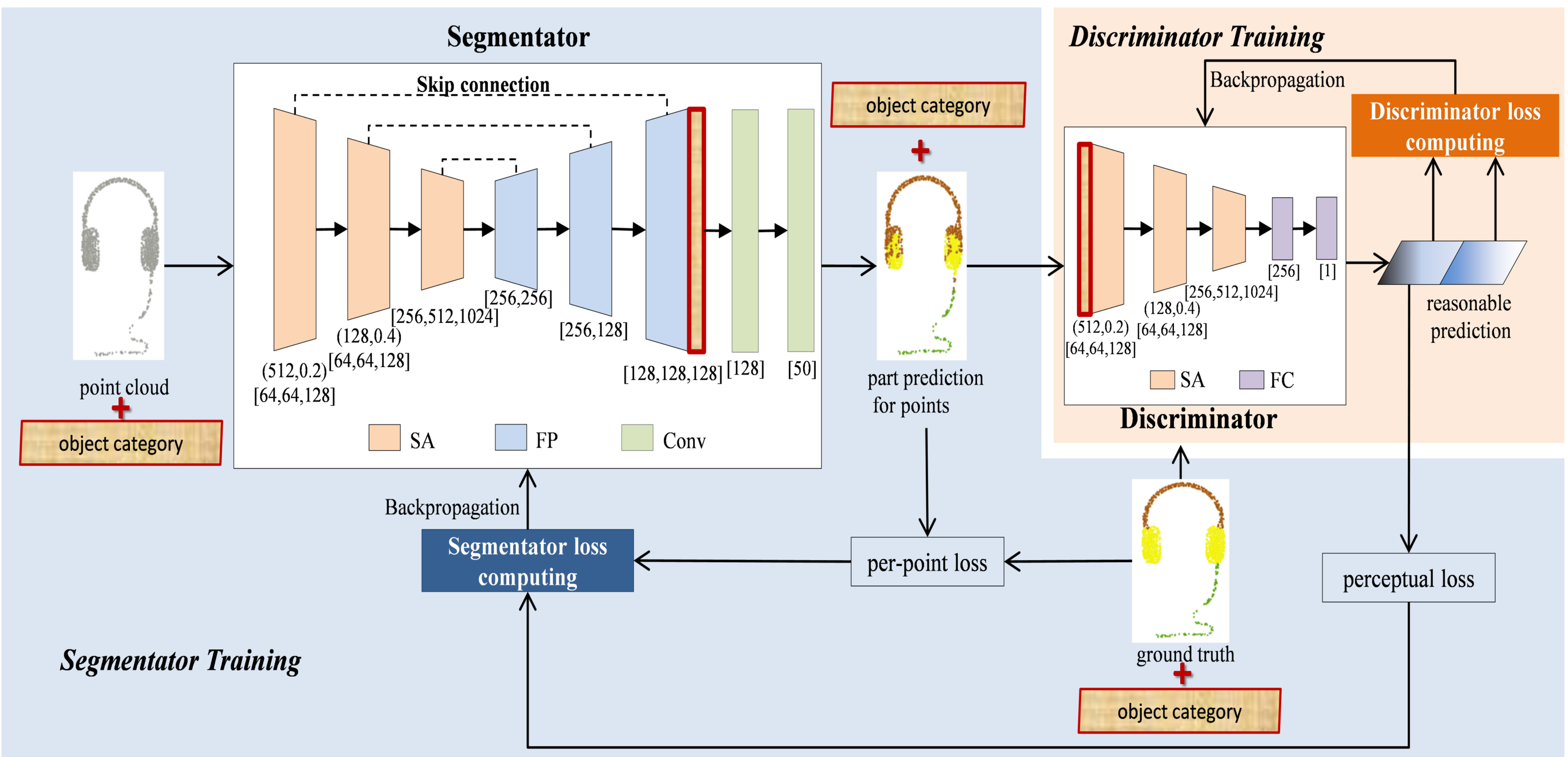
- A segmentation adversarial framework PPSAN is proposed to segmentation and labeling point cloud. With a perceptual loss provided by the discriminator, the quality of visual perception of segmentation results is concerned.
- We extend the PPSAN to the conditional setting, i.e. cPPSAN, which will further direct the segmentation and labeling process and produce more reasonable labeling results.
- Experimental results show that the proposed method can correct three common types of segmentation and labeling errors. Thus, more reasonable and better segmentation prediction results of visual perceptual quality can be obtained.

Overview

- The network architecture consists of segmentator and discriminator.

  Segmentator extracts point feature and predicts the part label for each point, which is similar to the existing point cloud segmentation networks.

  Discriminator is added, and takes the ground truth and segmentation prediction results as inputs respectively. The output of discriminator is reasonable prediction, which can determine whether the input of discriminator is a predicting result or the ground truth.



Training

- The training of discriminator network

$$L_D = \sum_{n=1}^{N}[L_{bce}(D(x_n, y_n), 1) + L_{bce}(D(x_n, S(x_n)), 0)]$$

  $S(\cdot)$ denotes the output of segmentation network and is the output of Discriminator network. $D(\cdot)$ is the output of Discriminator network. L_{bce} represents the perceptual loss of input part label maps.

- The training of segmentator network

$$L_S = \sum_{n=1}^{N}[L_{mce}(S(x_n), y_n) + \lambda L_{bce}(D(x_n, S(x_n)), 1)]$$

  where, $L_{mce}(\hat{y}, y) = -[y ln\hat{y} + (1-y)ln(1-\hat{y})]$denotes the per-point loss that computed by multi-class cross-entropy loss for segmentation prediction $\hat{y}$. L_{bce} represents the perceptual loss obtained by using segmentation prediction as the input of discriminator.

- When the PPSAN implements in condition settings, becoming cPPSAN.

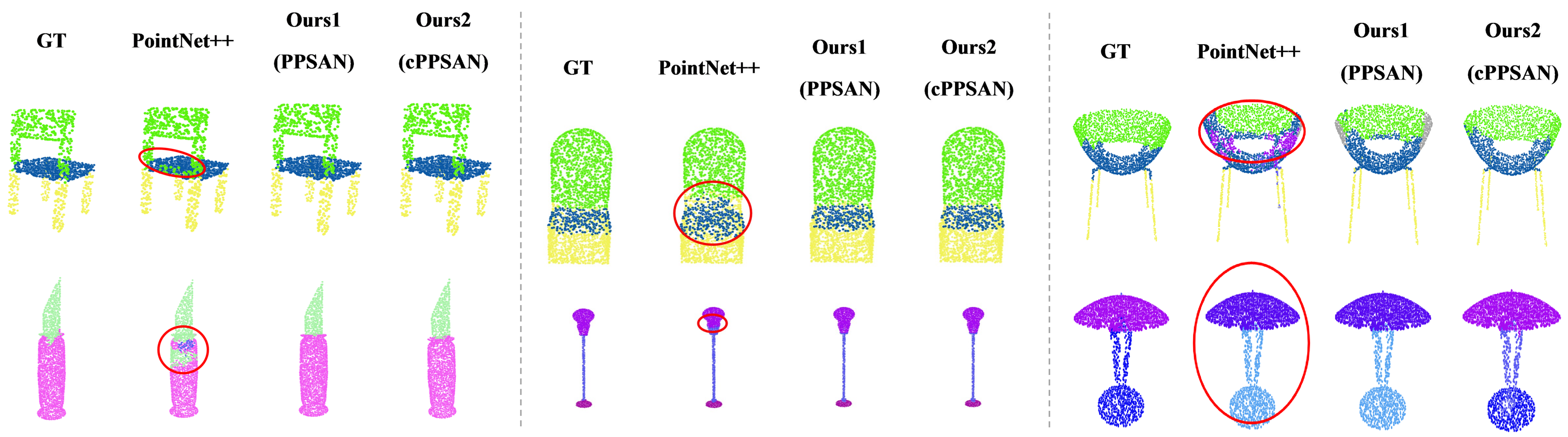
$$L_D = \sum_{n=1}^{N}[L_{bce}(D(x_n, y_n, l), 1) + L_{bce}(D(x_n, S(x_n), l), 0)]$$

$$L_S = \sum_{n=1}^{N}[L_{mce}(S(x_n), y_n, l) + \lambda L_{bce}(D(x_n, S(x_n), l), 1)]$$

  Where, l is the object category label, encoding in one-hot vector.

Experiments

We evaluate our framework in two settings, with or without global constraint, and compare the results with state-of-the-art methods in quantitative and qualitative.



References

[1] L. Yi et al., "A scalable active framework for region annotation in 3D shape collections," TOG., 2016.
[2] Charles Ruizhongtai Qi, H. Su, K. Mo, and L. J. Guibas, "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation," CVPR., 2017.
[3] Q. Huang, W. Wang, and U. Neumann, "Recurrent Slice Networks for 3D Segmentation of Point Clouds," CVPR, 2018.
[4] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space," NIPS., 2017.
[5] T. Le, G. Bui, and Y. Duan, "A multi-view recurrent neural network for 3D mesh segmentation," C&G., 2017.
[6] E. Kalogerakis, M. Averkiou, S. Maji, and S. Chaudhuri, "3D Shape Segmentation with Projective Convolutional Networks," CVPR, 2017.
[7] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," ICLR, 2016.
[8] P. Luc, C. Couprie, S. Chintala, and J. Verbeek, "Semantic Segmentation using Adversarial Networks," NIPS, 2016.

Contact us: lhynju@163.com, szx@nju.edu.cn