Experiment on Visualizing PointNet

Bo Sun
Peking University
{bosun@pku.edu.cn}

1. Introduction

PoinrNet [1] is state-of-the-art deep neural network to reason about 3D geometric data such as point clouds. In this experiment, I implemented the visualization of critical points and upper bound shape of one input point set. The method is in Sect. 2 and results are in Sect. 3.

2. Methods

2.1. Theoretical Analysis

The theoretical analysis is mainly based on the Sect. 4.2 in [1]. To test the robustness of the network, we can apply perturbation, corruption and extra noise points on the input point cloud. [1] found that the network is strongly affected by the dimension of the max pooling layer which we call K. We define $\mathbf{u} = \max_{x_i \in S}\{h(x_i)\}$ to be sub-network of f which maps a point set in $[0, 1]^m$ to a K-dimension vector. We illustrate the following theorem:

Theorem: *Suppose* $\mathbf{u} : \mathcal{X} \rightarrow \mathbb{R}$ *such that* $\mathbf{u} = \max_{x_i \in S}\{h(x_i)\}$ *and* $f = \gamma \circ \mathbf{u}$. *Then,*

(a). $\forall S, \exists \mathcal{C}_S, \mathcal{N}_S \subset \mathcal{X}, f(T) = f(S) if \mathcal{C}_S \subset T \subset \mathcal{N}_S;$

(b). $|\mathcal{C}_S| \leqslant K$

[1] explains the implications of the theorem. (a) says that $f(S)$ is unchanged up to the input corruption if all points in $\mathcal{C}_S$ are preserved; it is also unchanged with extra noise points up to $\mathcal{N}_S$. (b) says that $\mathcal{C}_S$ only contains a bounded number of points, determined by K. In other words, $f(S)$ is in fact totally determined by a nite subset $\mathcal{C}_S \subset S$ of less or equal to K elements. We therefore call $\mathcal{C}_S$ the critical point set of S and K the bottleneck dimension of f.

2.2. Implement Methods

The classification PointNet computes K (we take K = 1024 in this visualization) dimension point features for each point and aggregates all the per-point local features via a max pooling layer into a single K-dim vector, which forms the global shape descriptor.From the theoretical analysis we know that if the features extracted by max-pooling layer do not change, the output *i.e.* $f(x)$ doesn't change either.

Critical Points: We can only preserve points which contributes to the maximum of each feature rows. The input of max-pooling layer(we call $h(x)$) is of shape [points number, K] per point set and the shape of max-pooling layer is [1,K]. We can see that if we search the columns of $h(x)$ and only hold the max points, we can get a subset of the input point set which we call critical points.

Upper Bound Shape: We can add all points in input space that do not change the maximum of original $h(x)$ *i.e.* $h(x) < \mathbf{u}(S)$. Therefore, $\mathcal{T}_S$ can be obtained adding the union of all such points to $\mathcal{N}_S$.

Point Function: For each per-point function h, we calculate the values $h(p)$ for all the points p in a cube of diameter two located at the origin, which spatially covers the unit sphere to which our input shapes are normalized when training the PointNet. Here we think if $h(p) > 0.5$, p is important for the feature h extracted. So we input all points in the cube and compare them within one function h *i.e.* one column of the layer before max-pooling.

3. Results

I choose the first 5 points set in training set and visualize their critical points and upper bound shape on the classification network. I first train the network after 90 epochs and it achieves 93.1% accuracy on training set and 85.7% on test set. Some results is presented in Fig.1. We can see that the PointNet is robust to perturbation and we the max-pooling features extract the skeleton of objects.

References

[1] C. R.Qi, H. Su, and K. Mo. Pointnet: Deep learning on point sets for 3d classi?cation and segmentation. arXiv: 1612.00593.

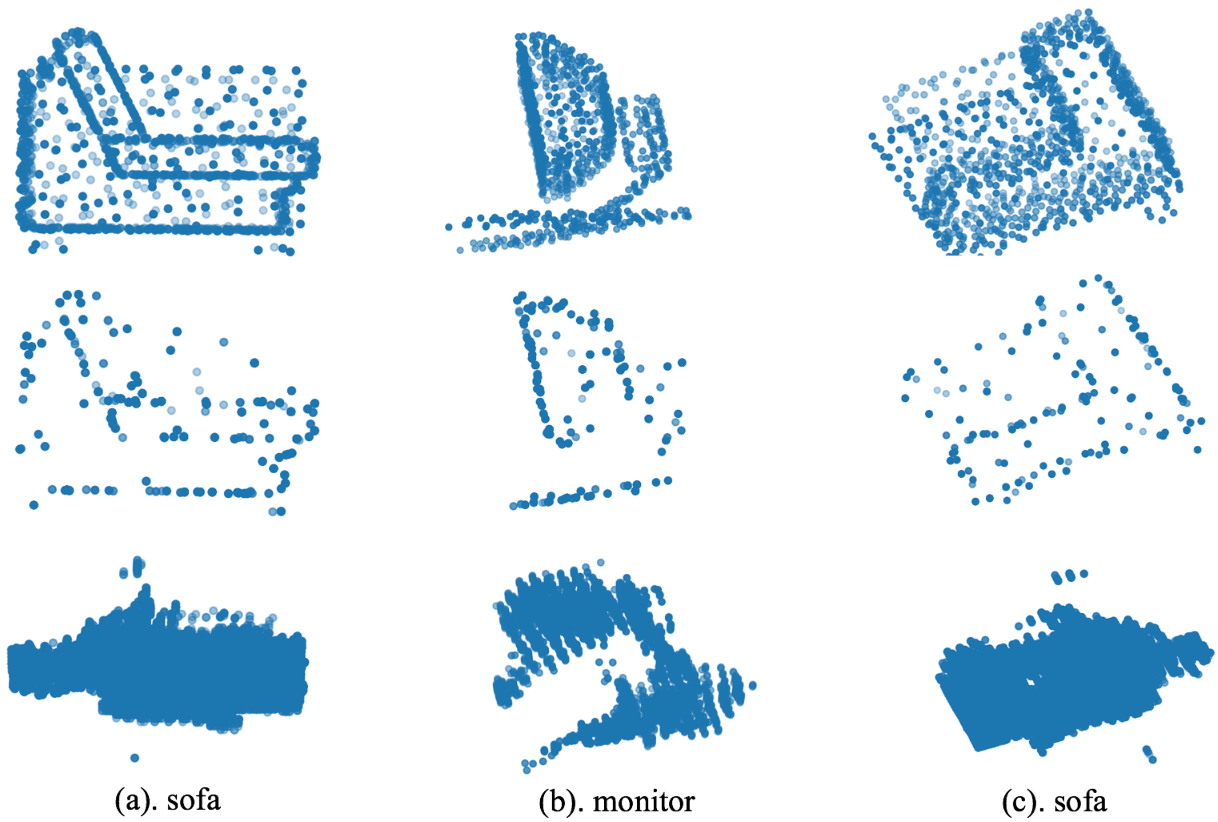

Figure 1. Critical points(medium) and upper bound shape(bottom). While critical points jointly determine the global shape feature for a given shape, any point cloud that falls between the critical points set and the upper bound shape gives exactly the same feature.