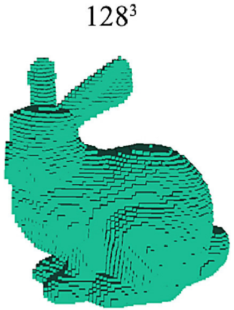
PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Presenter: Jiaxin Lu

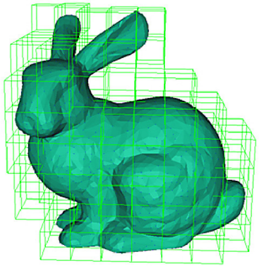
Sept. 6, 2022

3D Data vs. 2D Data

Image

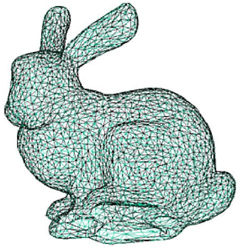


Voxel

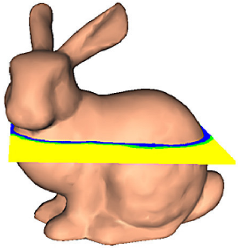


Octree

Point Cloud



Mesh



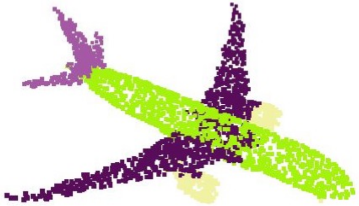
SDF

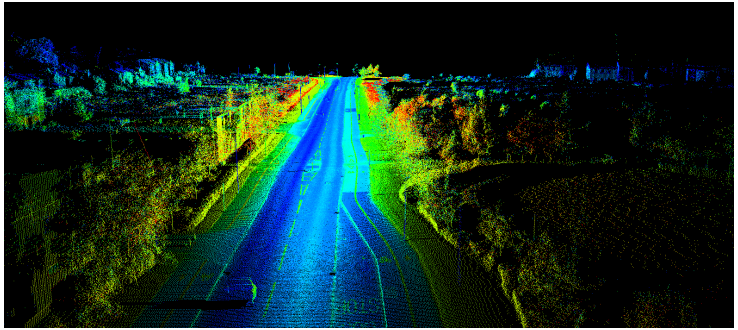
Regular Data

Irregular Data

Point Cloud

- Simplest, only points, no connectivity;
- (x, y, z) coordinates, with some features (normal, curvature, etc.)
- Easy to get directly from existing devices;
- Might be the only thing you can get directly from existing devices…





Problem Setting: Learning on Point Cloud

- **A set of 3D points** $\boldsymbol{M} = \{\boldsymbol{P_i} | \boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{n}\} \subseteq \mathbb{R}^{\boldsymbol{n}}$; for PointNet and PointNet++
- A distance metric d;
- $\mathcal{X} = (M, d)$ is a discrete metric space;

for PointNet++

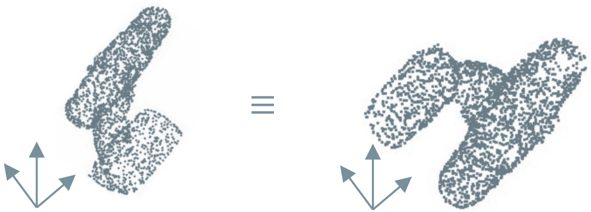
- Learn function f that take $\mathcal{X}$ as input and produce information regarding $\mathcal{X}$.

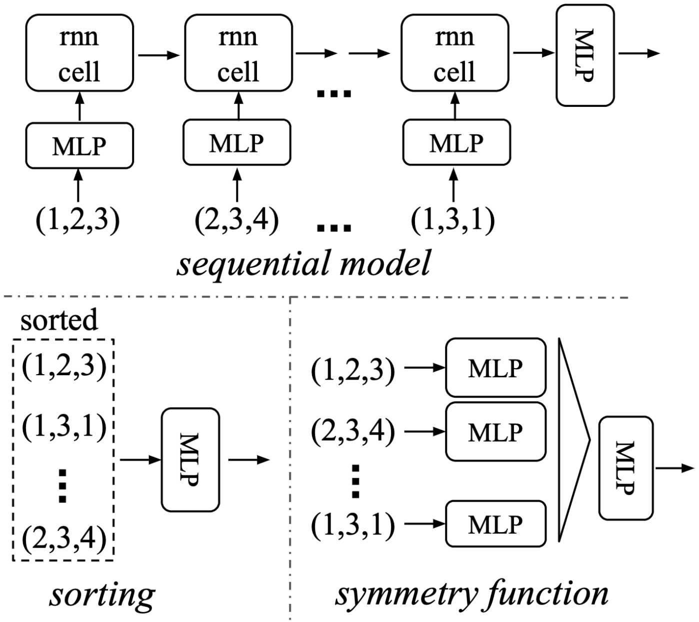
Properties of Desired Network

❖ Tackle **unordered** point set;

❖ Permutation Invariance;

❖ Transformation equivariance;





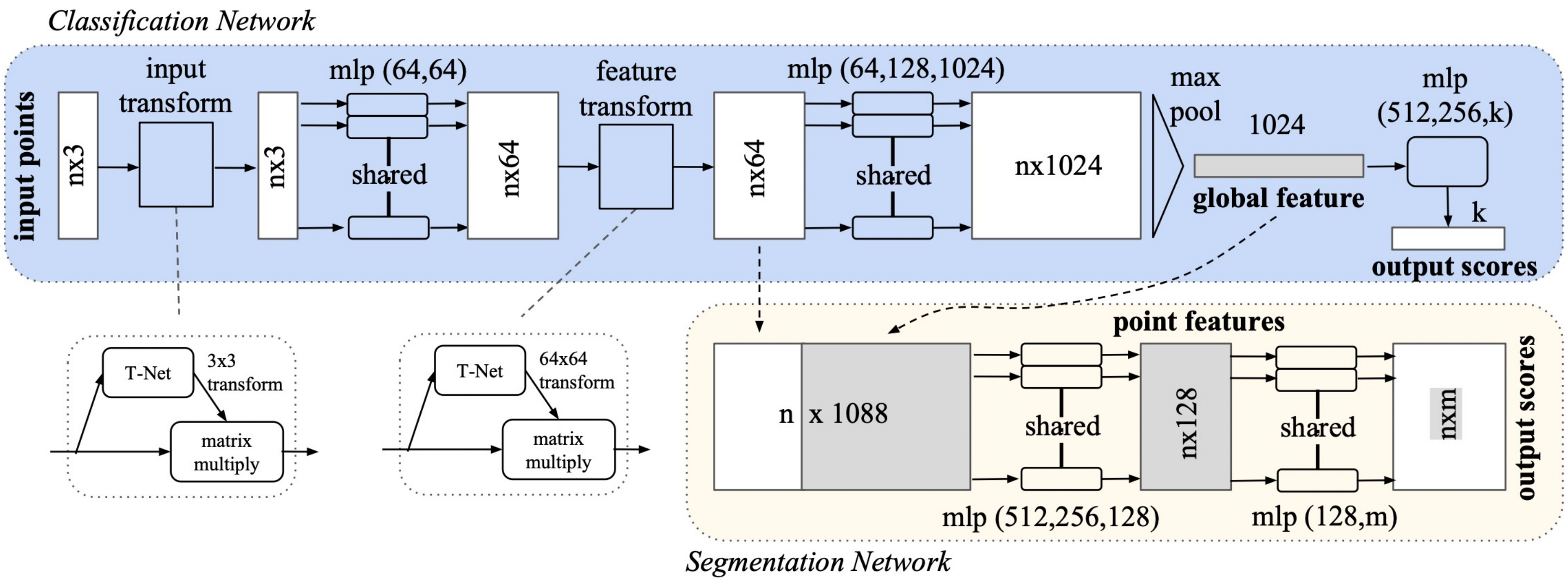
PointNet

- Fundamental idea: approximate a general function defined on a point set by applying a symmetric function on transformed elements in the set:

$$f(\{x_1, \dots, x_n\}) = \gamma\left(\underbrace{\underset{i=1,\dots,n}{\mathrm{MAX}}}_{\text{MLP}}\{h(x_i)\}\right)$$

- Symmetric function: $f(x_1, x_2, \dots, x_n) \equiv f(x_{\pi_1}, x_{\pi_2}, \dots, x_{\pi_n}), x_i \in \mathbb{R}^d$

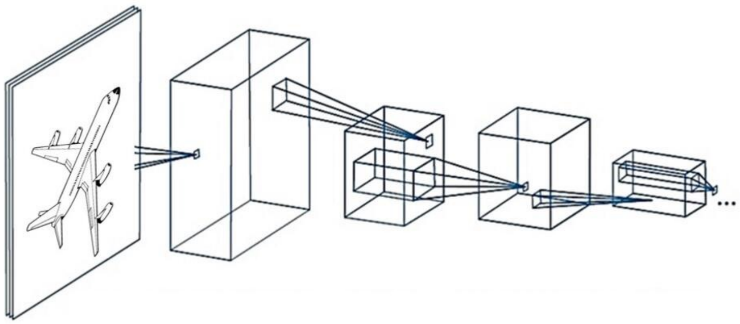
PointNet

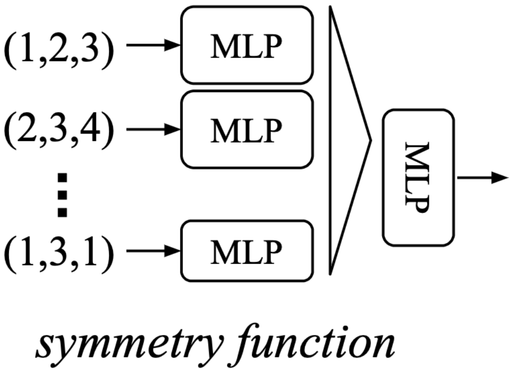


Classification Network
input points
nx3
input transform
nx3
mlp (64,64)
shared
nx64
feature transform
nx64
mlp (64,128,1024)
shared
nx1024
max pool
1024
global feature
mlp (512,256,k)
k
output scores
T-Net
3x3 transform
matrix multiply
T-Net
64x64 transform
matrix multiply
point features
n x 1088
shared
nx128
shared
nxm
output scores
mlp (512,256,128)
mlp (128,m)
Segmentation Network

Limitation of PointNet

❖ Only **global** feature, no **local** context for each point.



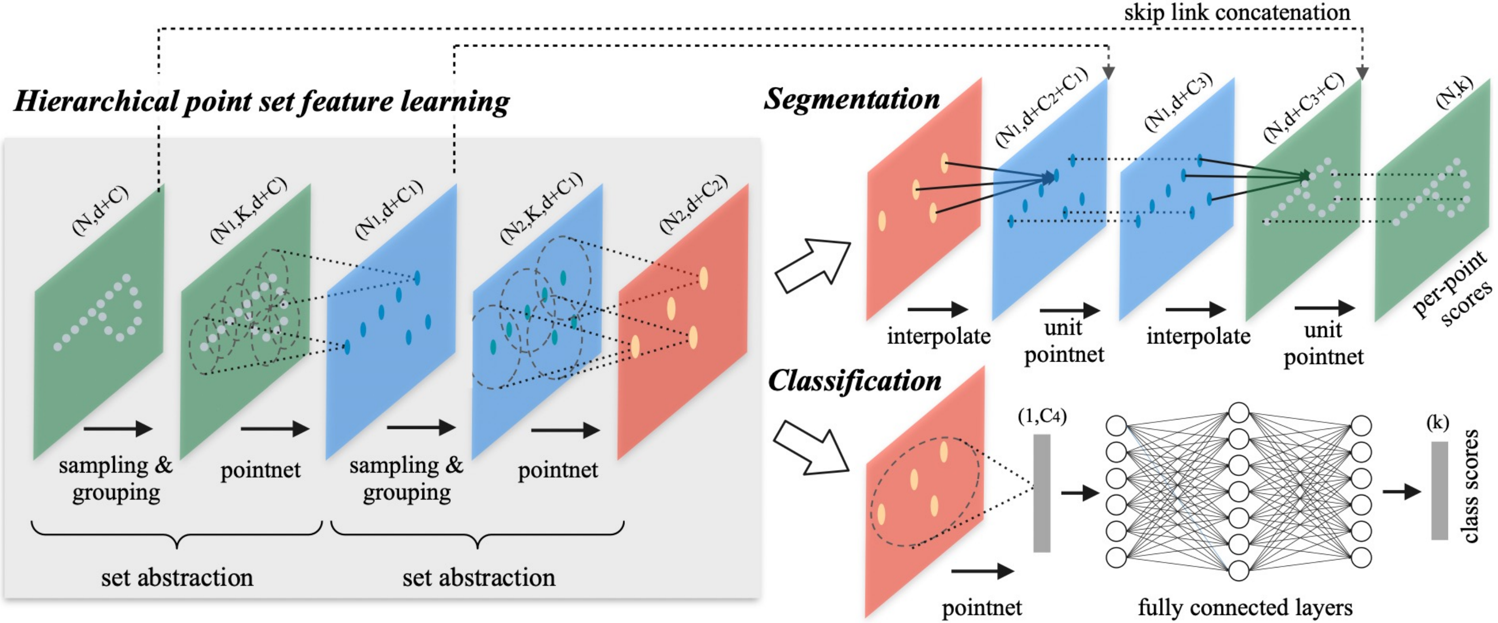


CNN

PointNet

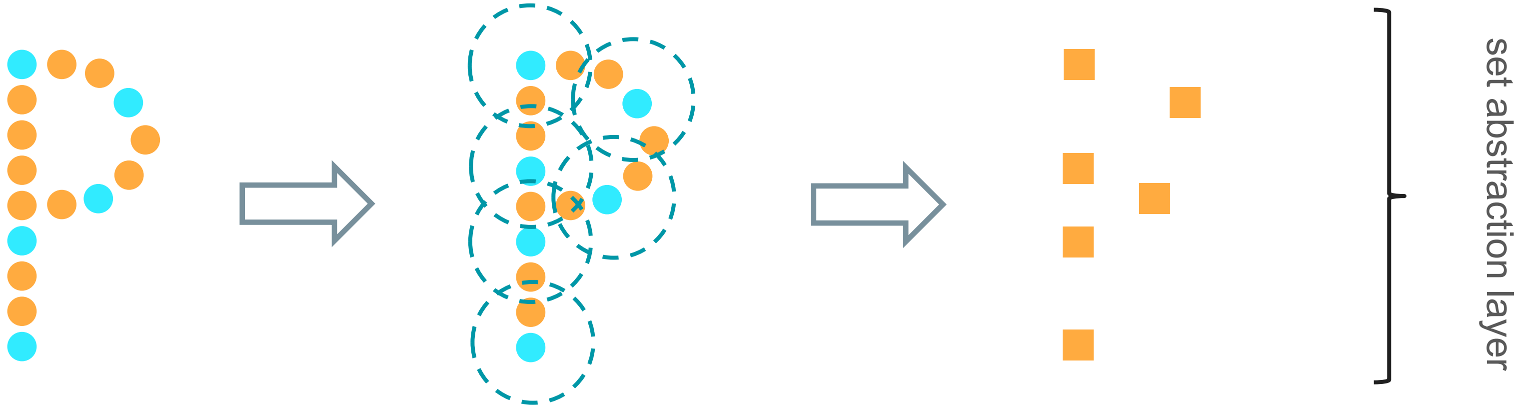
PointNet++

❖ Fundamental idea: hierarchical point set feature learning.



Hierarchical Point Set Feature Learning

- Sample anchor points (sampling layer);
- Find neighboring points of each anchor points (grouping layer);
- Use PointNet as basic building block for local pattern learning (PointNet layer).

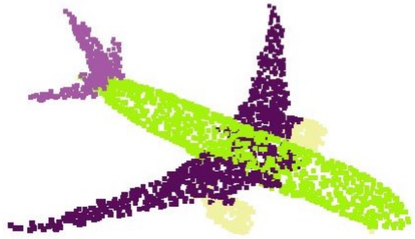


N points in $(x, y, \mathcal{F}_c)$
Farthest point sampling (FPS)

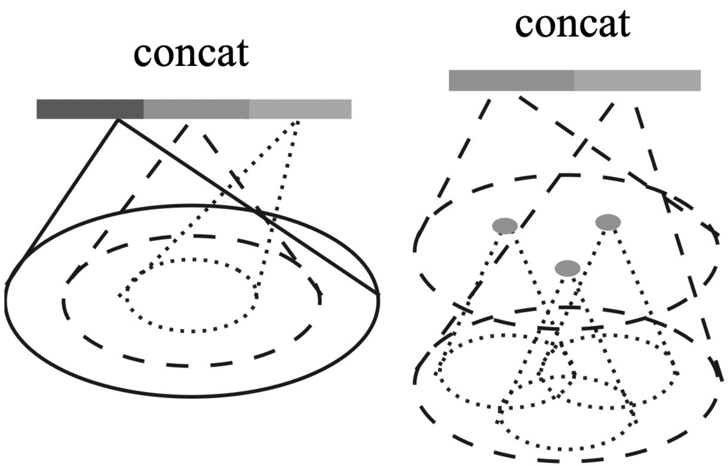
$N' \times K$ in $(x, y, \mathcal{F}_c)$
Ball query / kNN

N' points in $(x, y, \mathcal{F}_{c'})$
PointNet

Density Adaptive PointNet Layer

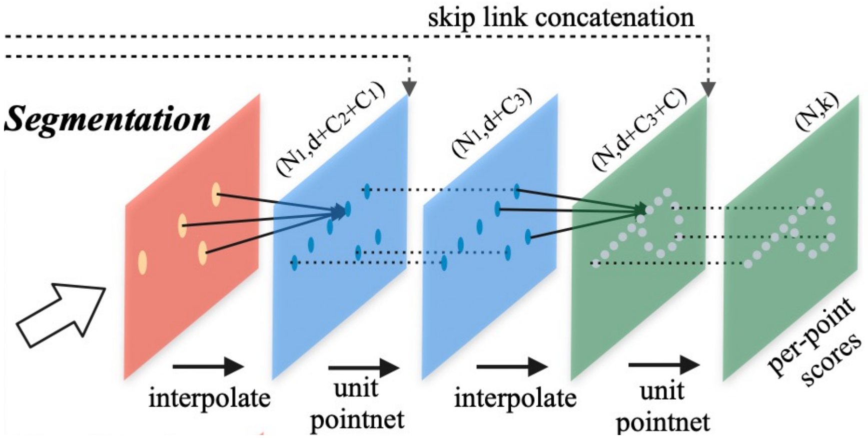


- More challenge: non-uniform density in point cloud.
- Multi-scale grouping (MSG)
  - Apply grouping layers with different scales
  - Random input dropout
- Multi-resolution grouping (MRG)
  - At layer L_i, summarizing features from lower level and process all raw points in the local region

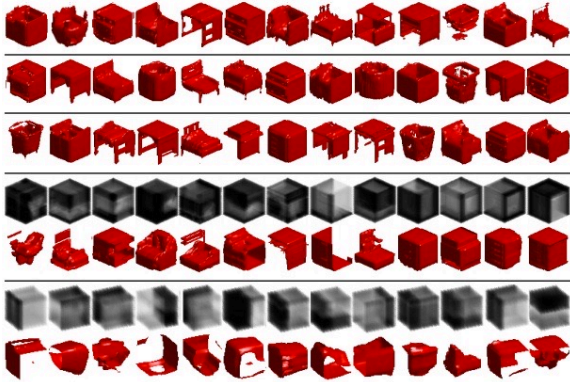


Hierarchical Feature Propagation

- Challenge: for segmentation task, we want features of **all** the original points.
- $f^{(j)}(x) = \frac{\sum w_i(x) f_i^{(j)}}{\sum w_i(x)}$, where $w_i(x) = \frac{1}{d(x,x_i)^p}, j = 1, \ldots, C$



Experimental Setup



- ❖ Datasets
  - ❖ MNIST: 2D images;
  - ❖ ModelNet40: 3D CAD models;
  - ❖ SHREC15: Shape in non-Euclidean metric space (non-rigid);
  - ❖ ScanNet: Indoor scenes;
- ❖ Metrics
  - ❖ Error rate / Accuracy

Experimental Results

- On MNIST and ModelNet40

Method	Error rate (%)
Multi-layer perceptron [24]	1.60
LeNet5 [11]	0.80
Network in Network [13]	**0.47**
PointNet (vanilla) [20]	1.30
PointNet [20]	0.78
Ours	0.51

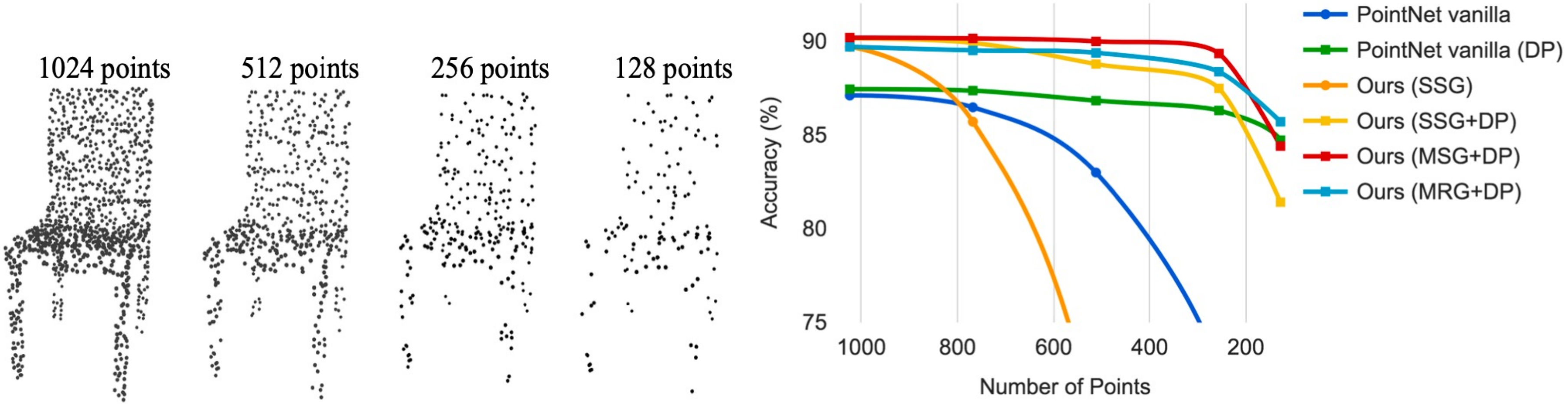
Table 1: MNIST digit classification.

Method	Input	Accuracy (%)
Subvolume [21]	vox	89.2
MVCNN [26]	img	90.1
PointNet (vanilla) [20]	pc	87.2
PointNet [20]	pc	89.2
Ours	pc	90.7
Ours (with normal)	pc	**91.9**

Table 2: ModelNet40 shape classification.

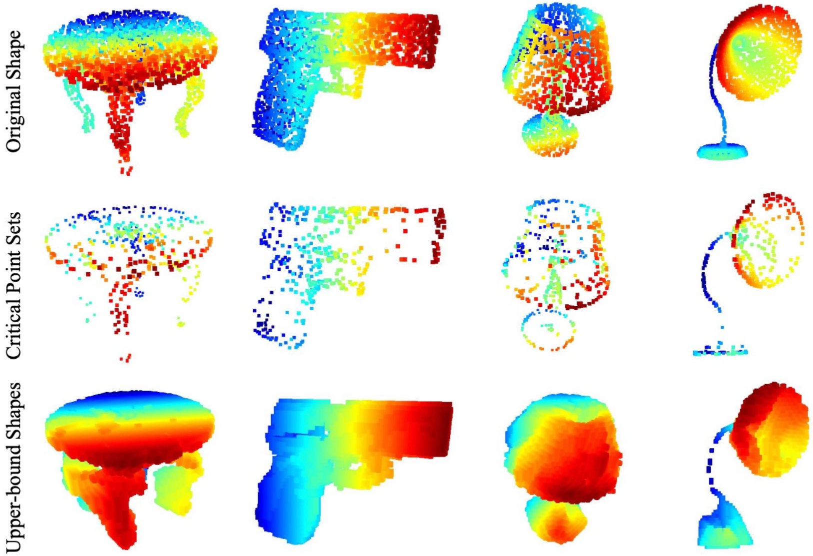
Experimental Results

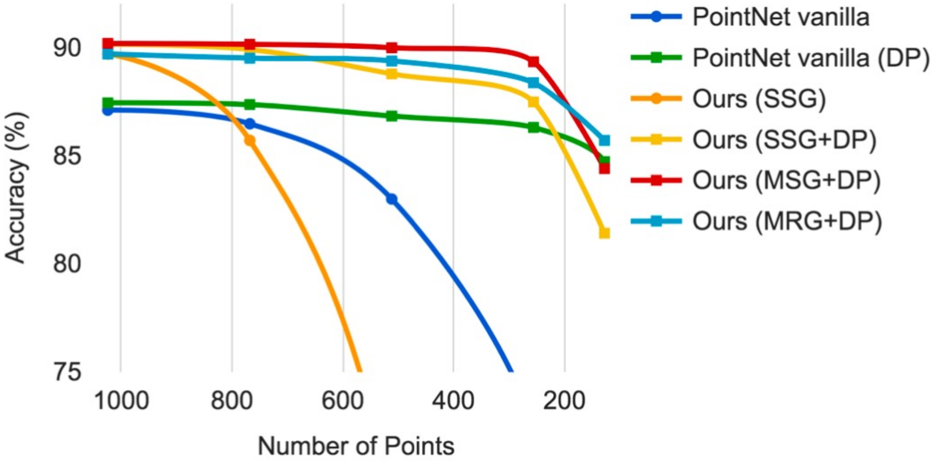
❖ On robustness to sampling density variation



Experimental Results

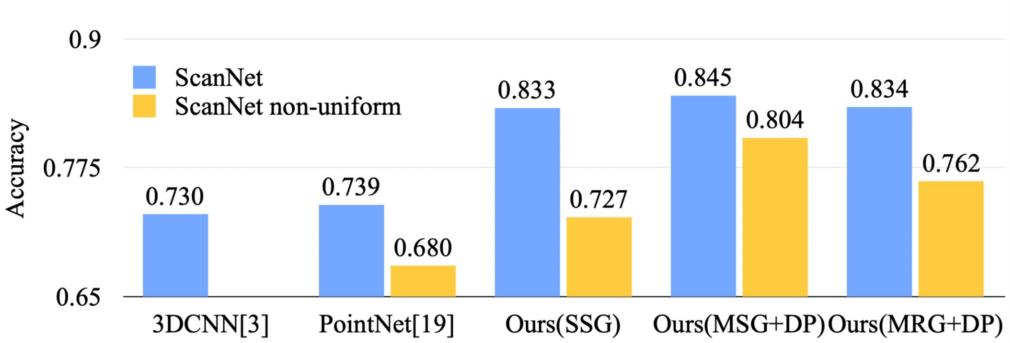
❖ On robustness to sampling density variation

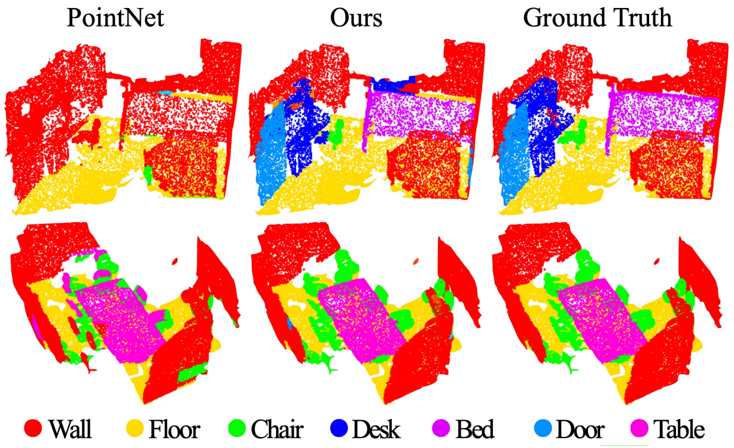




Experimental Results

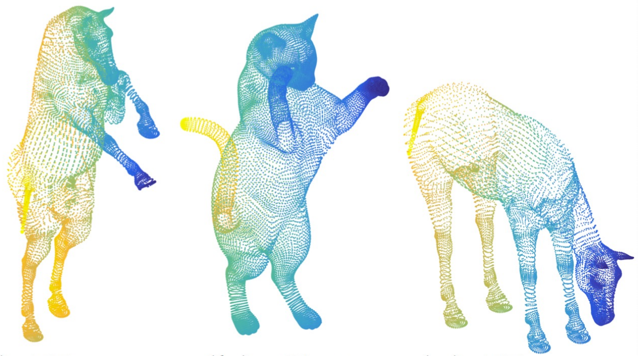
- On large scale point cloud analysis





Experimental Results

- On non-Euclidean metric space
  - distance metric d;
  - Intrinsic features: geodesic distance, WKS, HKS, Gaussian curvature …



(a) Horse (b) Cat (c) Horse

	Metric space	Input feature	Accuracy (%)
DeepGM [14]	-	Intrinsic features	93.03
Ours	Euclidean	XYZ	60.18
	Euclidean	Intrinsic features	94.49
	Non-Euclidean	Intrinsic features	**96.09**

Experimental Results

- kNN v.s. Ball Query

kNN (k=16)	kNN (k=64)	radius (r=0.1)	radius (r=0.2)
89.3	90.3	89.1	90.7

- Farthest Point Sampling

Feature difference std.	Accuracy std.
0.021	0.0017

- Time and Space Complexity

	PointNet (vanilla) [1]	PointNet [1]	Ours (SSG)	Ours (MSG)	Ours (MRG)
Model size (MB)	9.4	40	8.7	12	24
Forward time (ms)	11.6	25.3	82.4	163.2	87.0

Conclusion and Open Issues

- An effective hierarchical learning network on point cloud.
  - Sampling & Grouping Layer;
  - Non-uniform point sampling;
  - Feature propagation.
- State-of-the-art performance.

- How to sample better local regions? (VoxelNet)
- How to propagate point features? (EdgeConv, RS-CNN, etc.)
- …

Extended Readings

- Maturana, Daniel, and Sebastian Scherer. "Voxnet: A 3d convolutional neural network for real-time object recognition." 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, 2015.
- Qi, Charles R., et al. "PointNet: Deep learning on point sets for 3d classification and segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- Wang, Yue, et al. "Dynamic graph CNN for learning on point clouds." Acm Transactions On Graphics (tog) 38.5 (2019): 1-12.
- Liu, Yongcheng, et al. "Relation-shape convolutional neural network for point cloud analysis." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.
- Zhou, Yin, and Oncel Tuzel. "Voxelnet: End-to-end learning for point cloud based 3d object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

Summary

- 3D data and point cloud
- PointNet
  - Limitation: only global feature, no local context.
- PointNet ++
  - A hierarchical design using grouping and sampling;
  - Learning on non-uniform sampling density;
  - Feature propagation.
- State-of-the-art results on point cloud processing.