CVPR与ICCV 3D点云深度学习论文详解

CVer 今天

以下文章来源于中国图象图形学报 ，作者刘永成

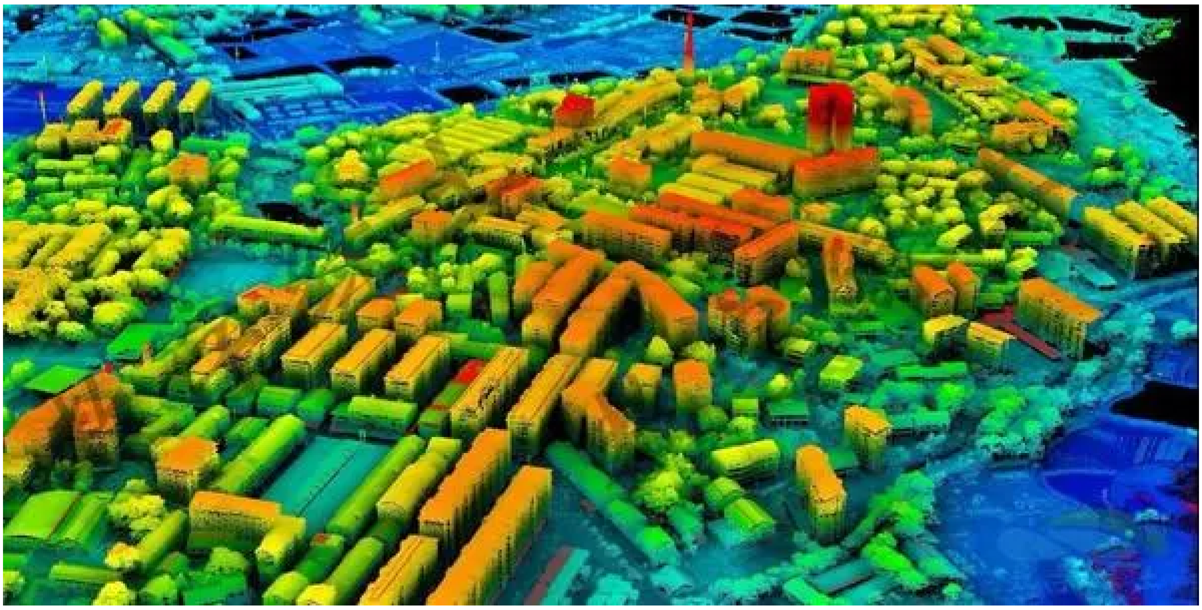
中国图象图形学报

中国图象图形学报网络宣传与推广

点击上方“**CVer**”，选择加"星标"或“置顶”

重磅干货，第一时间送达

本文转载自：中国图象图形学报



图片来源网络

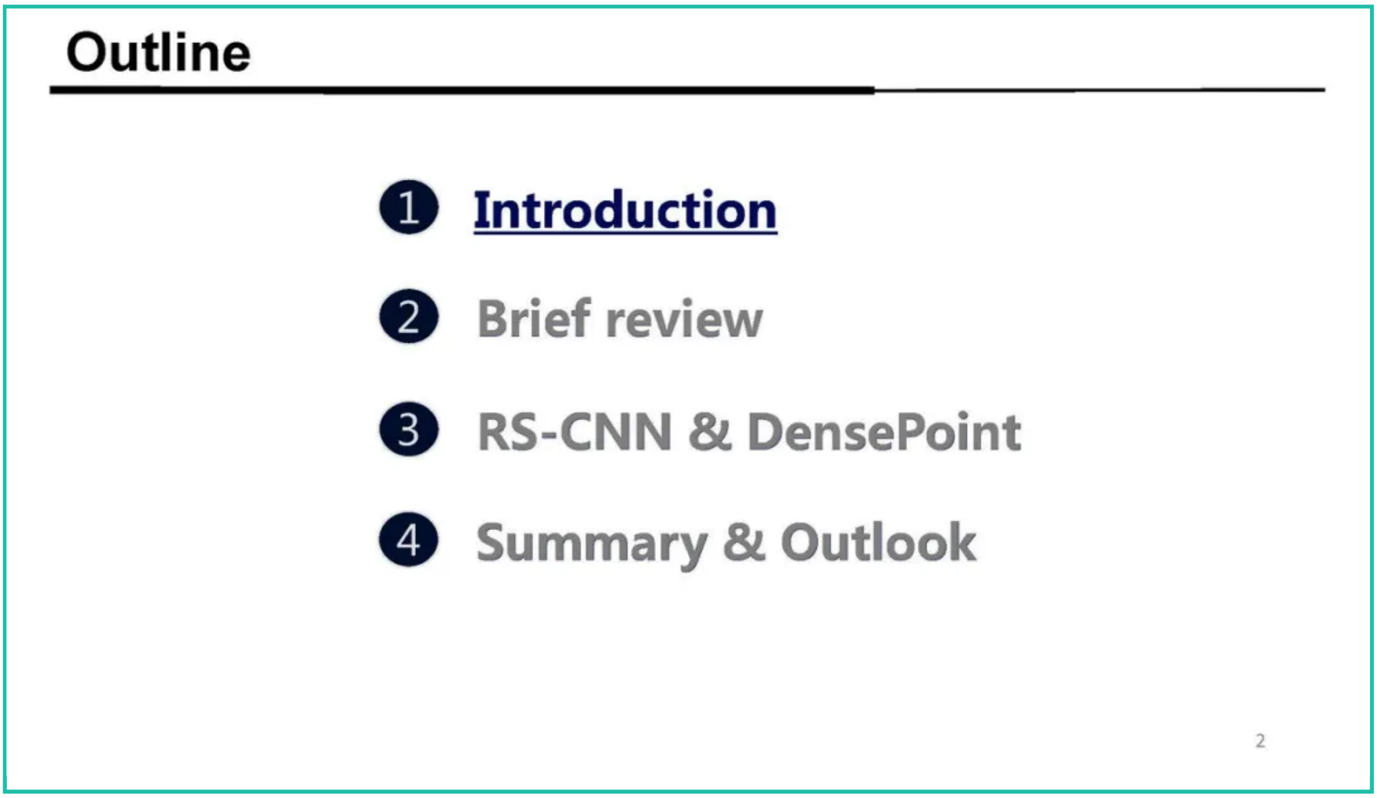
深度学习在图像识别中成效显著，然而如何进行拓展以识别不规则的3D点云，仍然是一个开放性的问题。

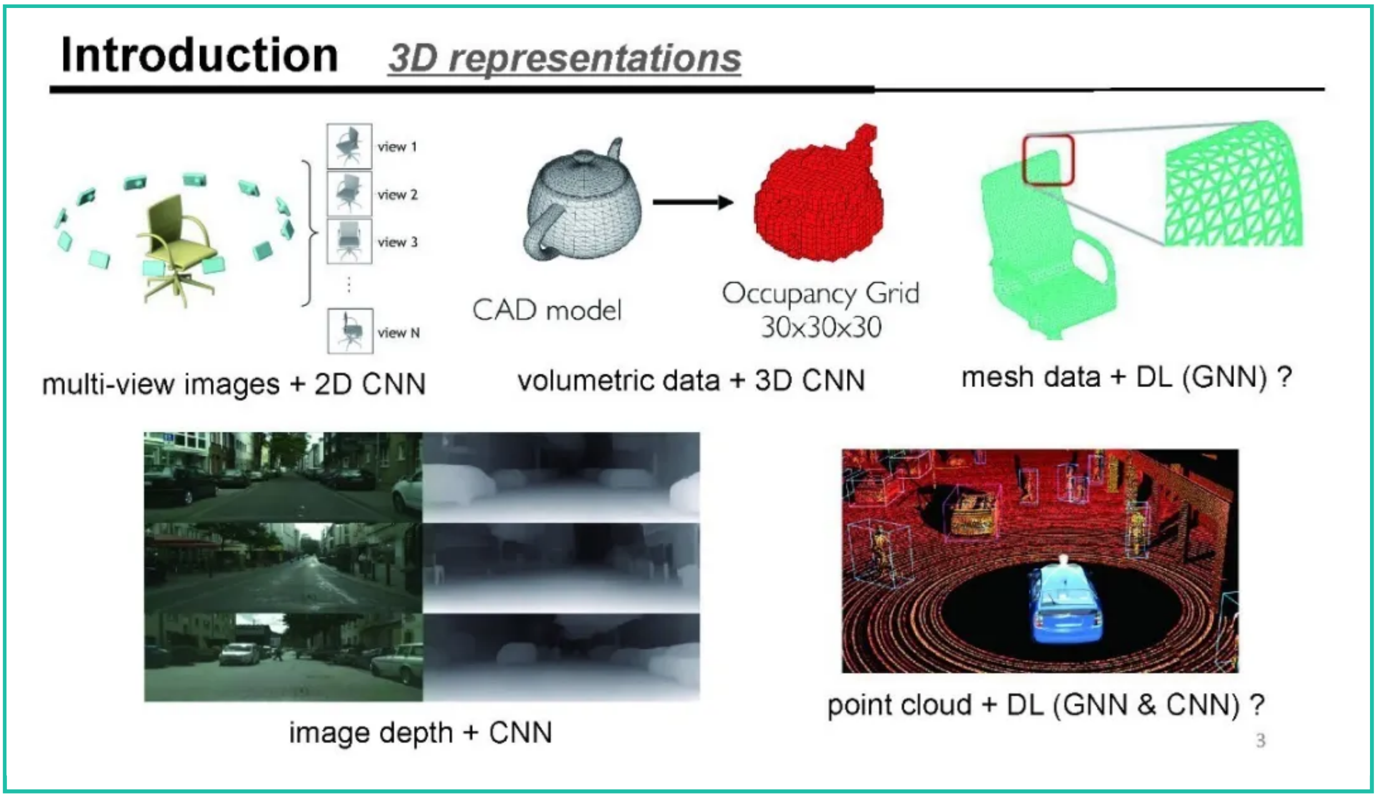
近两年3D点云识别发展迅速，研究方法上，从对称函数聚合到局部模式挖掘，再到几何图建模以及卷积核设计，该领域呈现百家争鸣的状态。

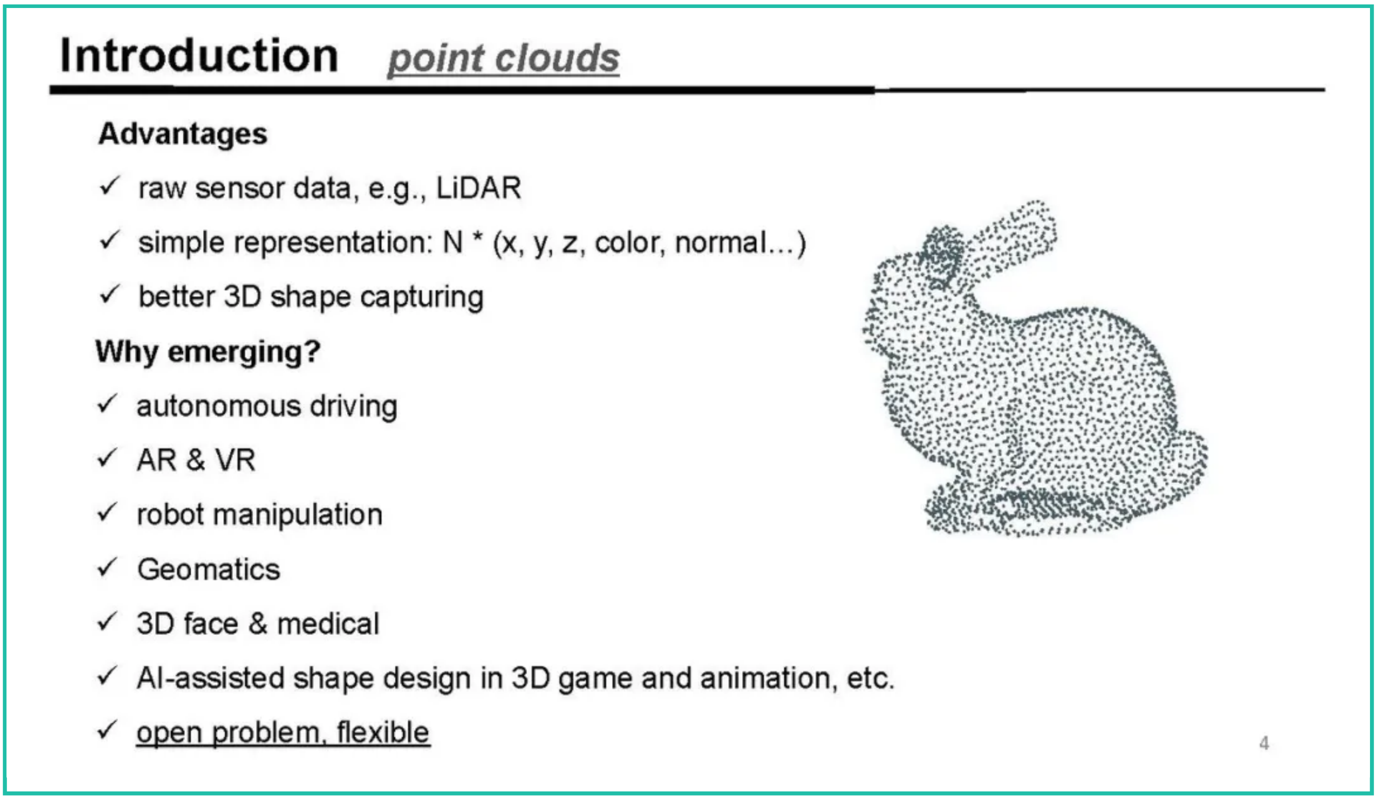
图图带来中科院自动化所**刘永成博士**的报告：《**深度学习在点云识别中的应用**》，报告对领域研究现状进行综述并介绍了其研究团队的一些探索性成果：

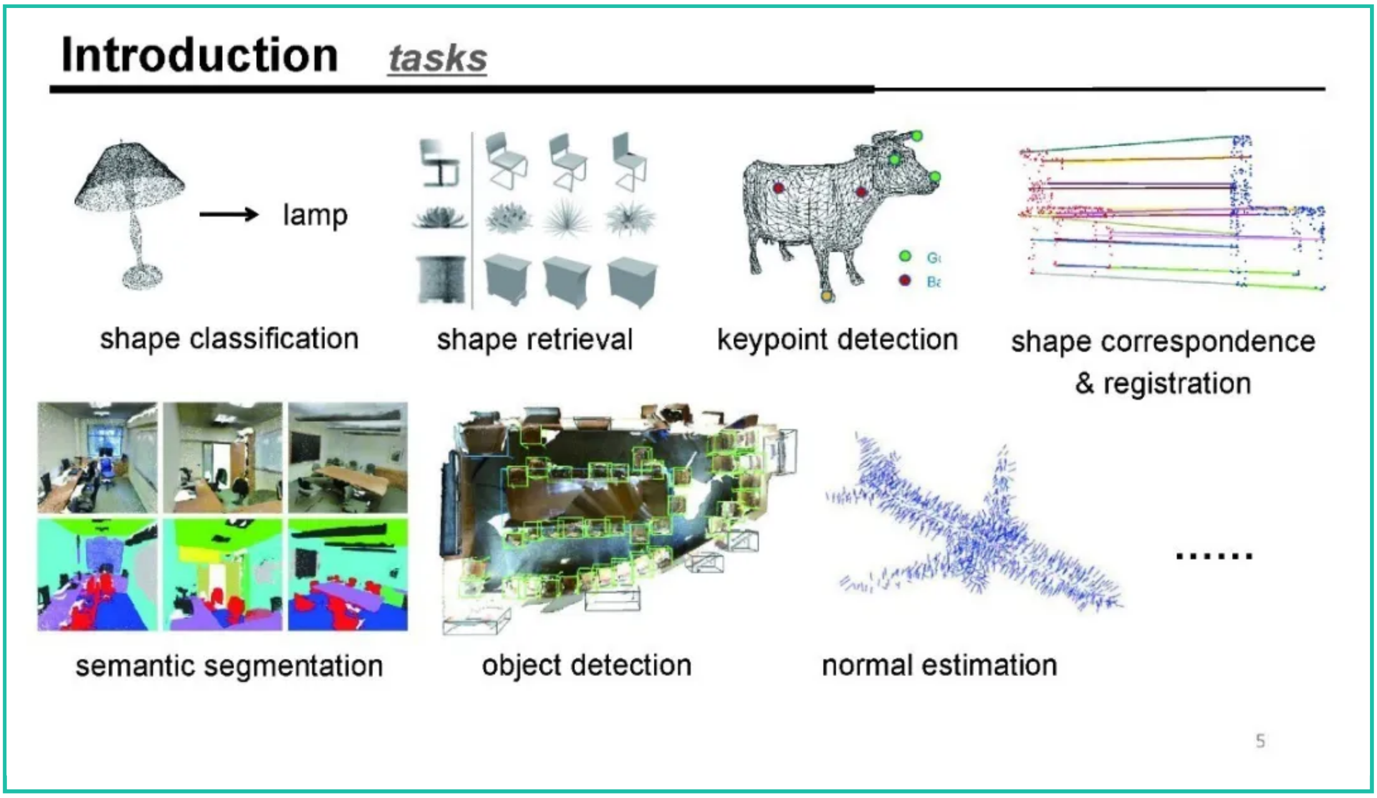
- Relation-Shape CNN(CVPR 2019)
- DensePoint(ICCV 2019)

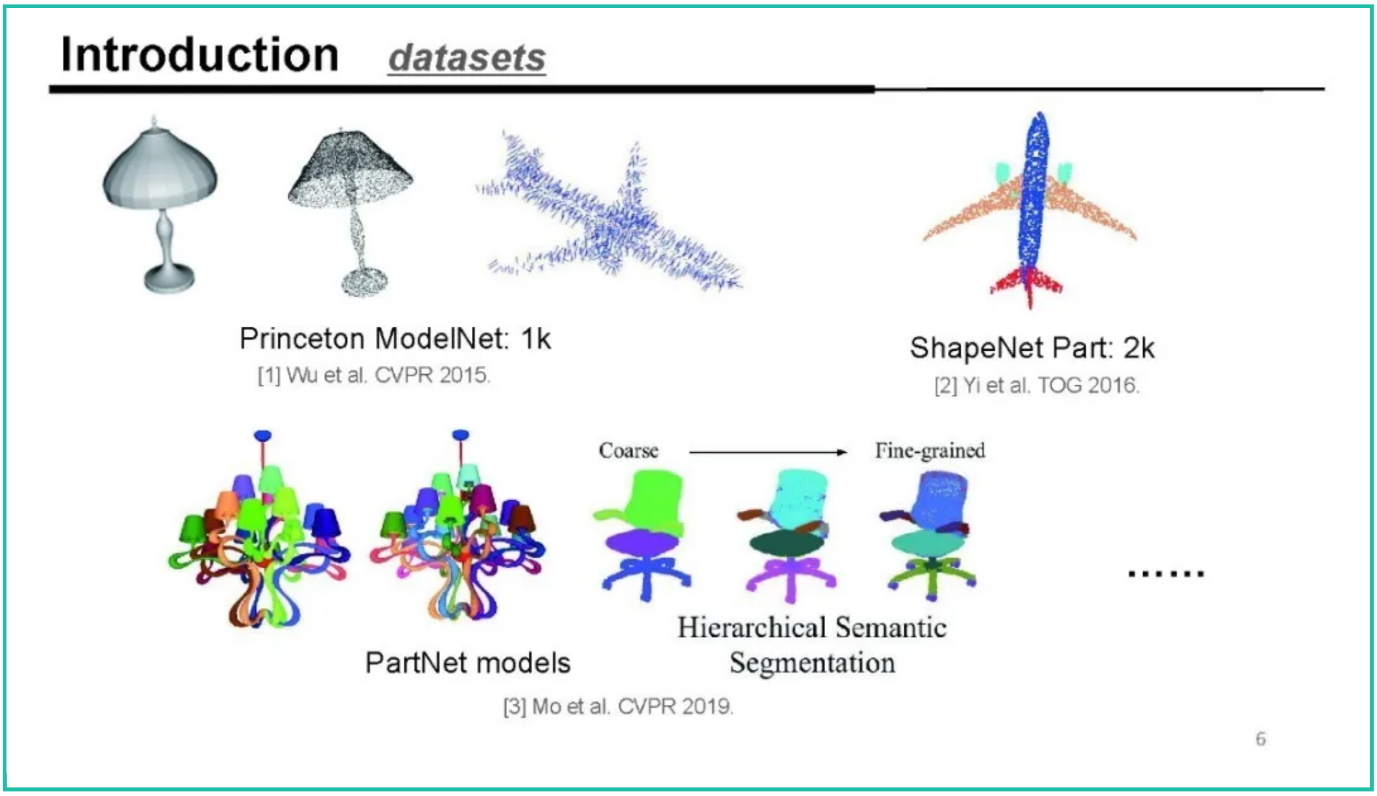
报告内容

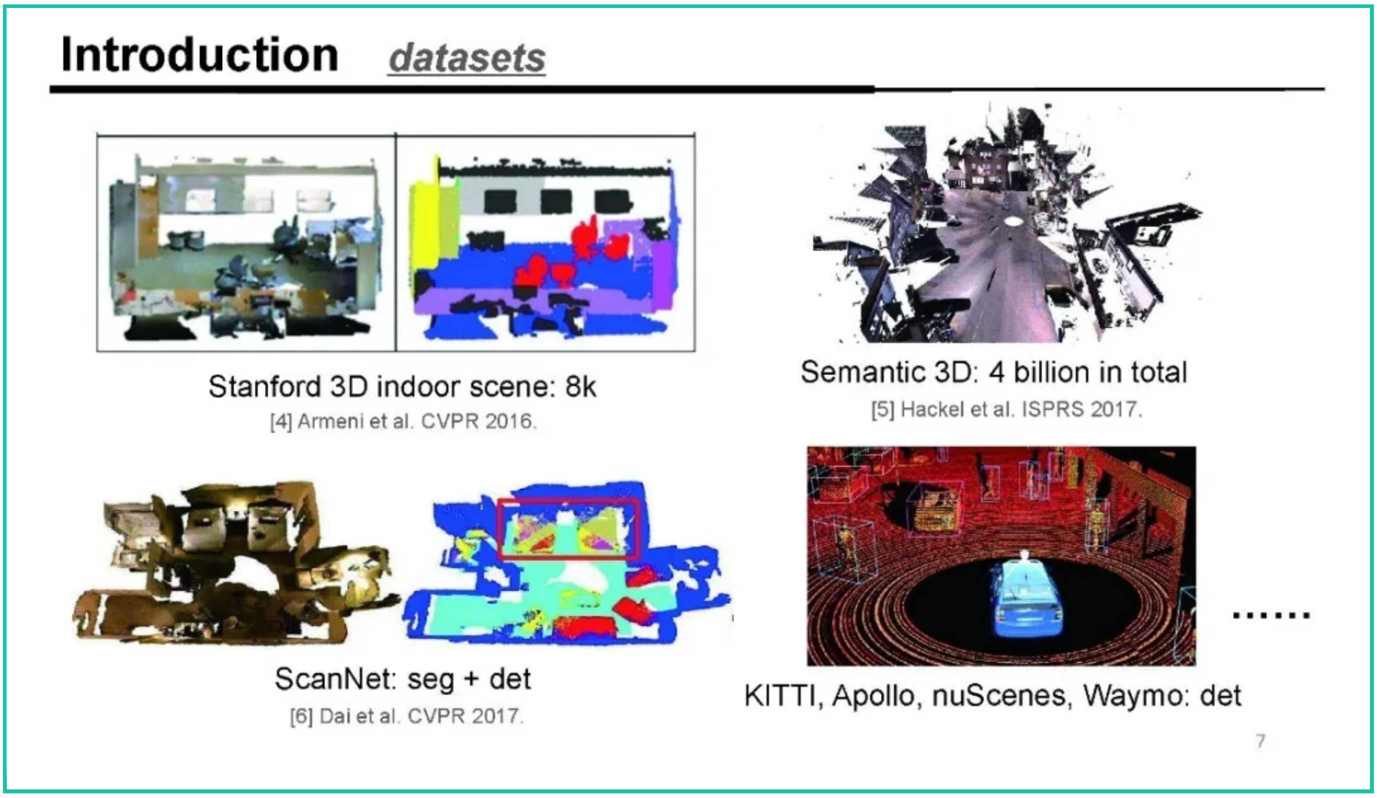


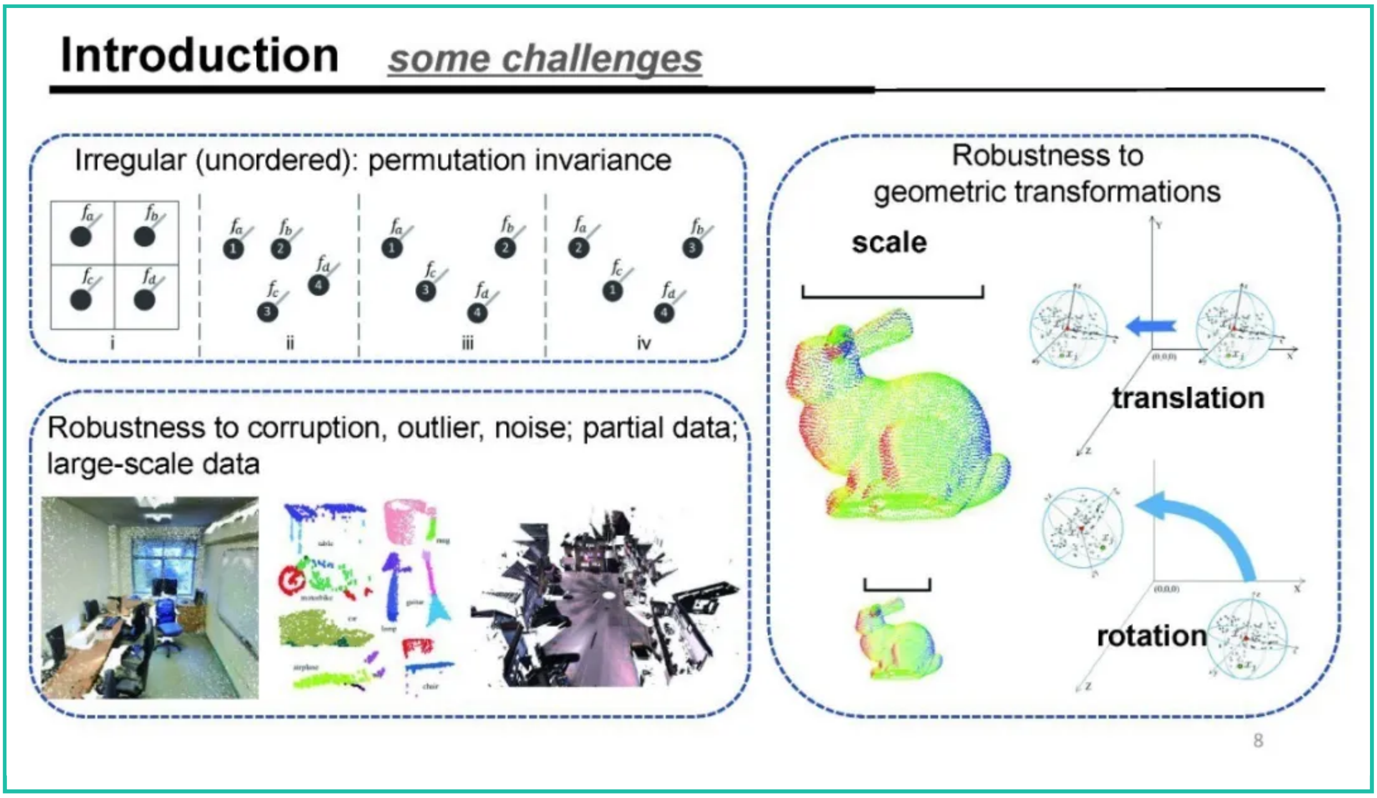

Introduction 3D representations
view 1
view 2
view 3
view N
multi-view images + 2D CNN
CAD model
Occupancy Grid 30x30x30
volumetric data + 3D CNN
mesh data + DL (GNN) ?
image depth + CNN
point cloud + DL (GNN & CNN) ?
3


Introduction point clouds
Advantages
✓ raw sensor data, e.g., LiDAR
✓ simple representation: N * (x, y, z, color, normal...)
✓ better 3D shape capturing
Why emerging?
✓ autonomous driving
✓ AR & VR
✓ robot manipulation
✓ Geomatics
✓ 3D face & medical
✓ AI-assisted shape design in 3D game and animation, etc.
✓ open problem, flexible
4


Introduction tasks
lamp
shape classification
shape retrieval
keypoint detection
shape correspondence & registration
semantic segmentation
object detection
normal estimation
......
5


Introduction datasets
Princeton ModelNet: 1k
[1] Wu et al. CVPR 2015.
ShapeNet Part: 2k
[2] Yi et al. TOG 2016.
Coarse
Fine-grained
PartNet models
Hierarchical Semantic Segmentation
[3] Mo et al. CVPR 2019.
......
6

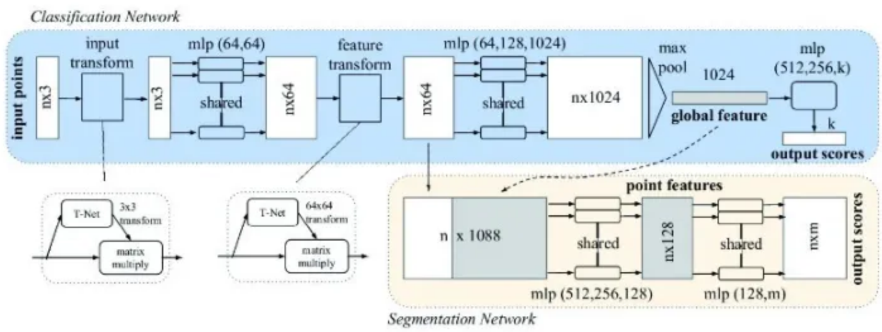

Introduction
datasets
Stanford 3D indoor scene: 8k
[4] Armeni et al. CVPR 2016.
Semantic 3D: 4 billion in total
[5] Hackel et al. ISPRS 2017.
ScanNet: seg + det
[6] Dai et al. CVPR 2017.
......
KITTI, Apollo, nuScenes, Waymo: det
7


Introduction
some challenges
Irregular (unordered): permutation invariance
i
ii
iii
iv
Robustness to corruption, outlier, noise; partial data;
large-scale data
Robustness to
geometric transformations
scale
translation
rotation
8

Outline

1. Introduction
2. **Brief review**
3. RS-CNN & DensePoint
4. Summary & Outlook

Related Work *PointNet family*



[18] Qi et al. PointNet. CVPR 2017.

Shared MLP + max pool
(symmetric function)

General function: permutation invariance

[4] Zaheer et al. Deep sets. NeurIPS 2017.

$$f(\{x_1, \ldots, x_n\}) \approx g(h(x_1), \ldots, h(x_n))$$

h: shared transform function
MLP

g: symmetric aggregation function, e.g., + or *
Max pooling + fc layers

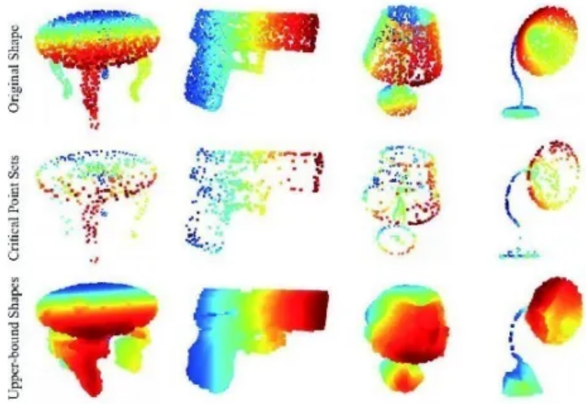
Related Work *PointNet family*

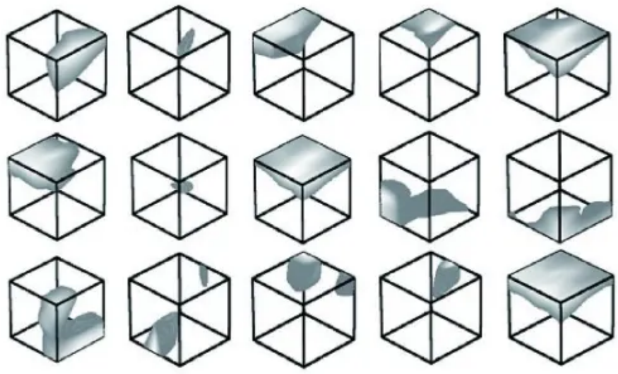
[18] Qi et al. PointNet. CVPR 2017.

- What does PointNet learn?

summarize an input point cloud by
a sparse set of key points (shape skeleton)

$$f(\{x_1,\ldots,x_n\}) \approx g(h(x_1),\ldots,h(x_n))$$



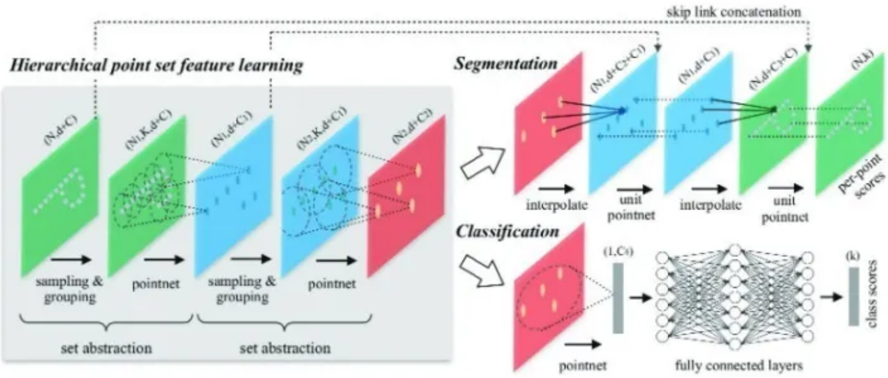


- What does PointNet learn?

$h_i(p) > 0.5$
partition/probe the space

11

Related Work *PointNet family*



[19] Qi et al. PointNet++. NeurIPS 2017.

Sampling + Grouping + PointNet

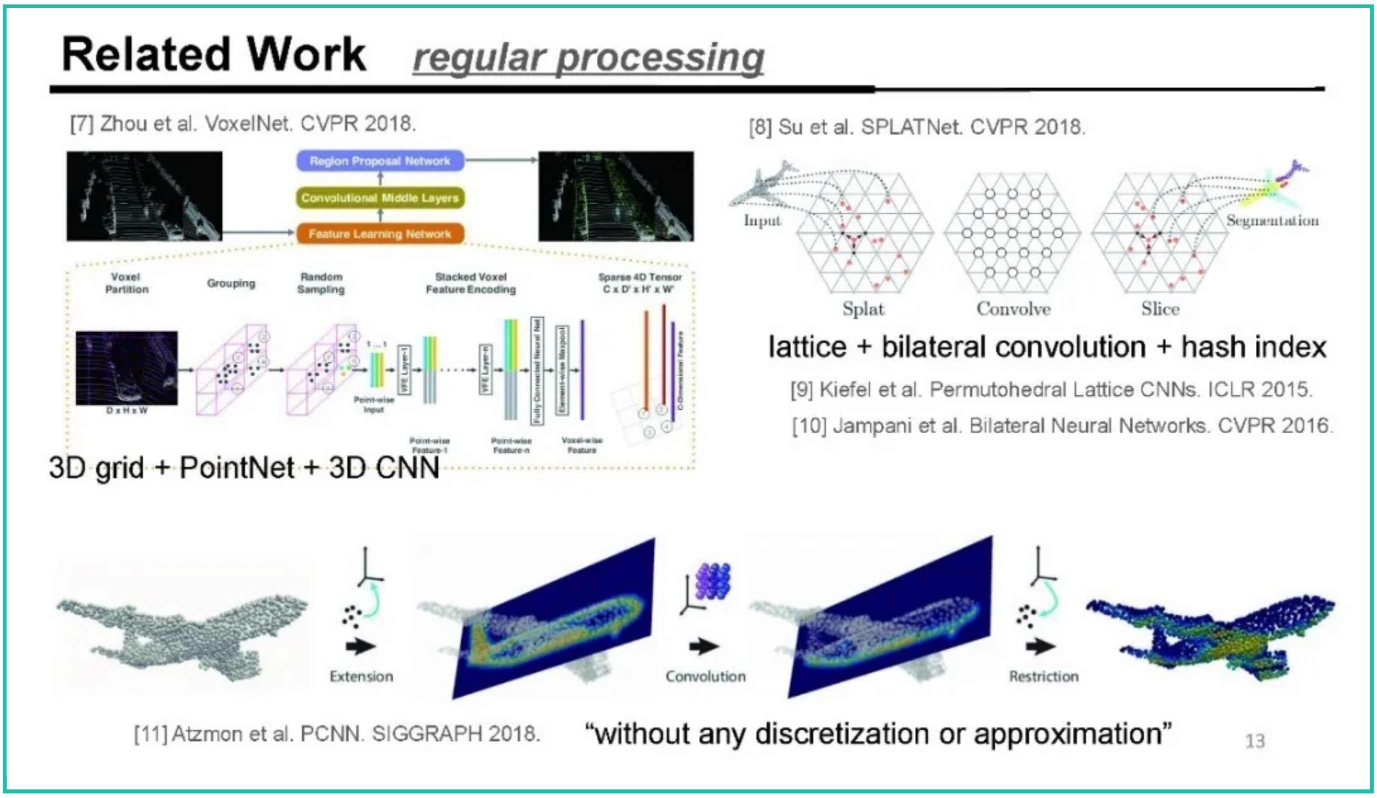
capture local patterns better

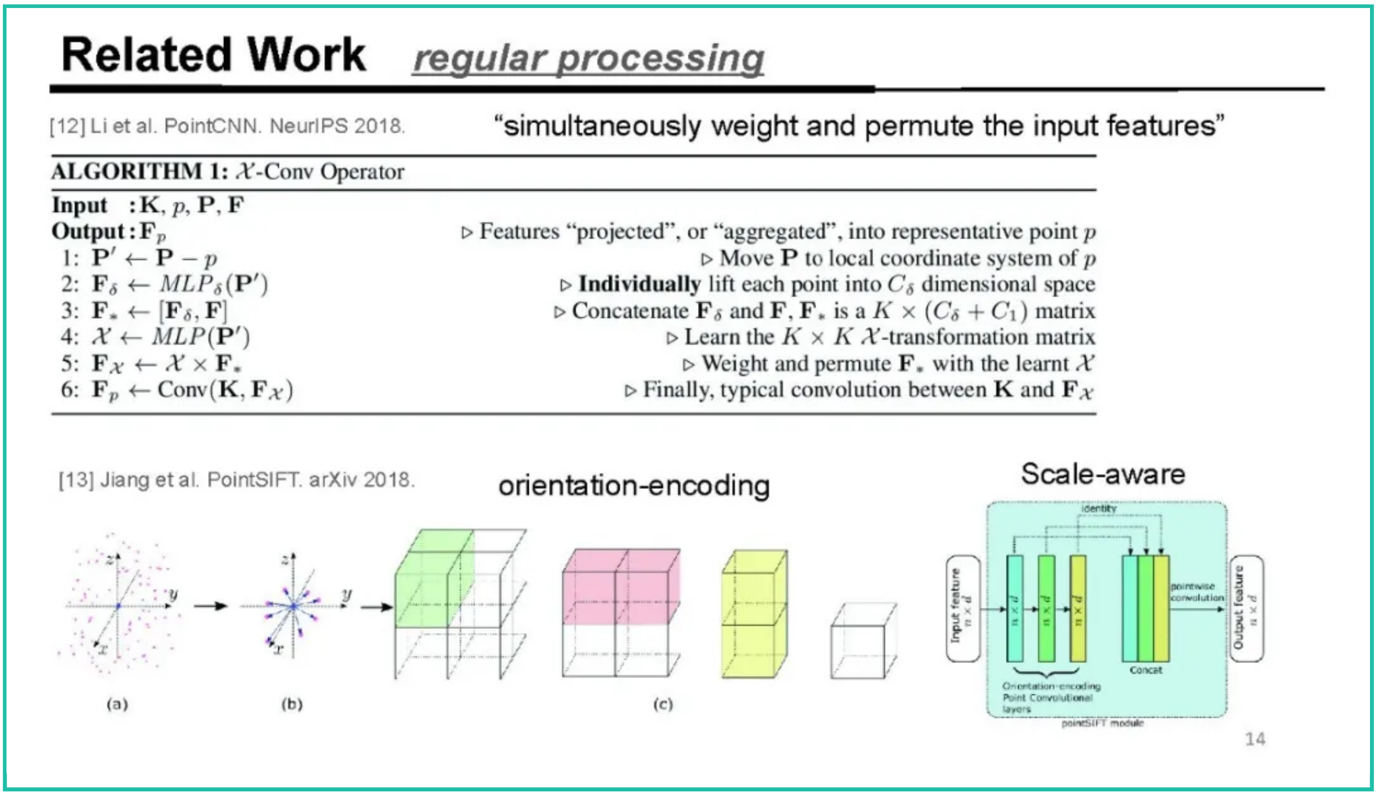
CNN like

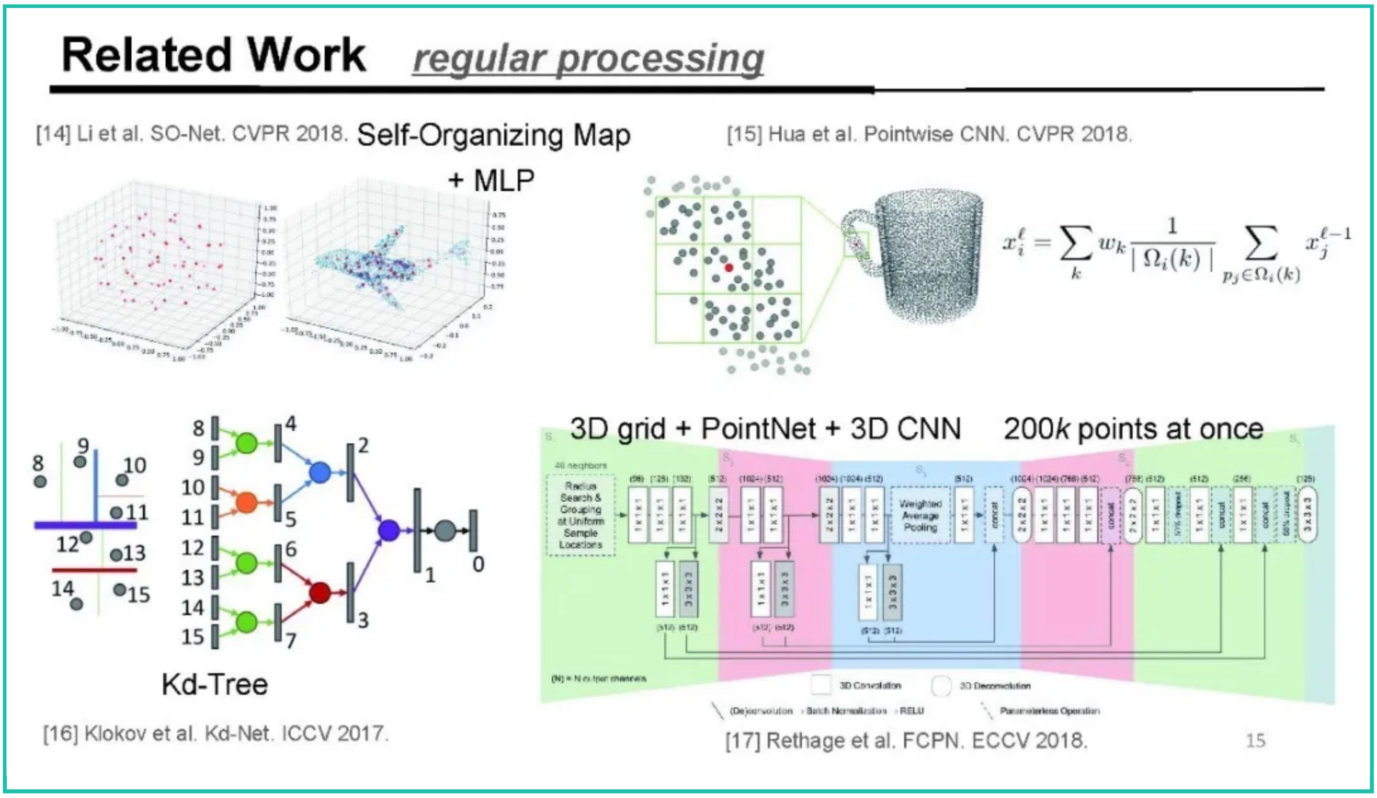
Segmentation decoder upsampling:

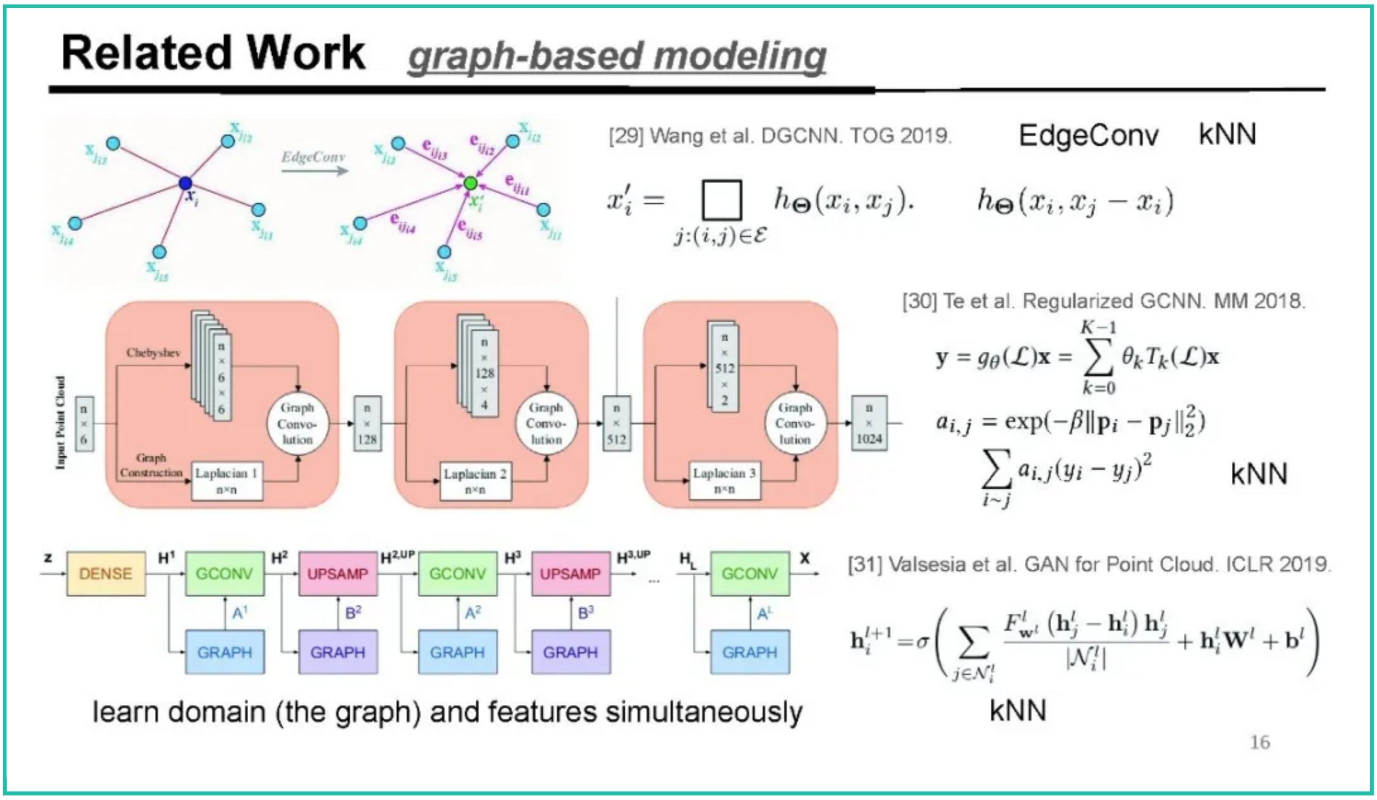
$$f^{(j)}(x) = \frac{\sum_{i=1}^{k} w_i(x) f_i^{(j)}}{\sum_{i=1}^{k} w_i(x)} \quad \text{where} \quad w_i(x) = \frac{1}{d(x,x_i)^p},\ j = 1, \ldots, C$$

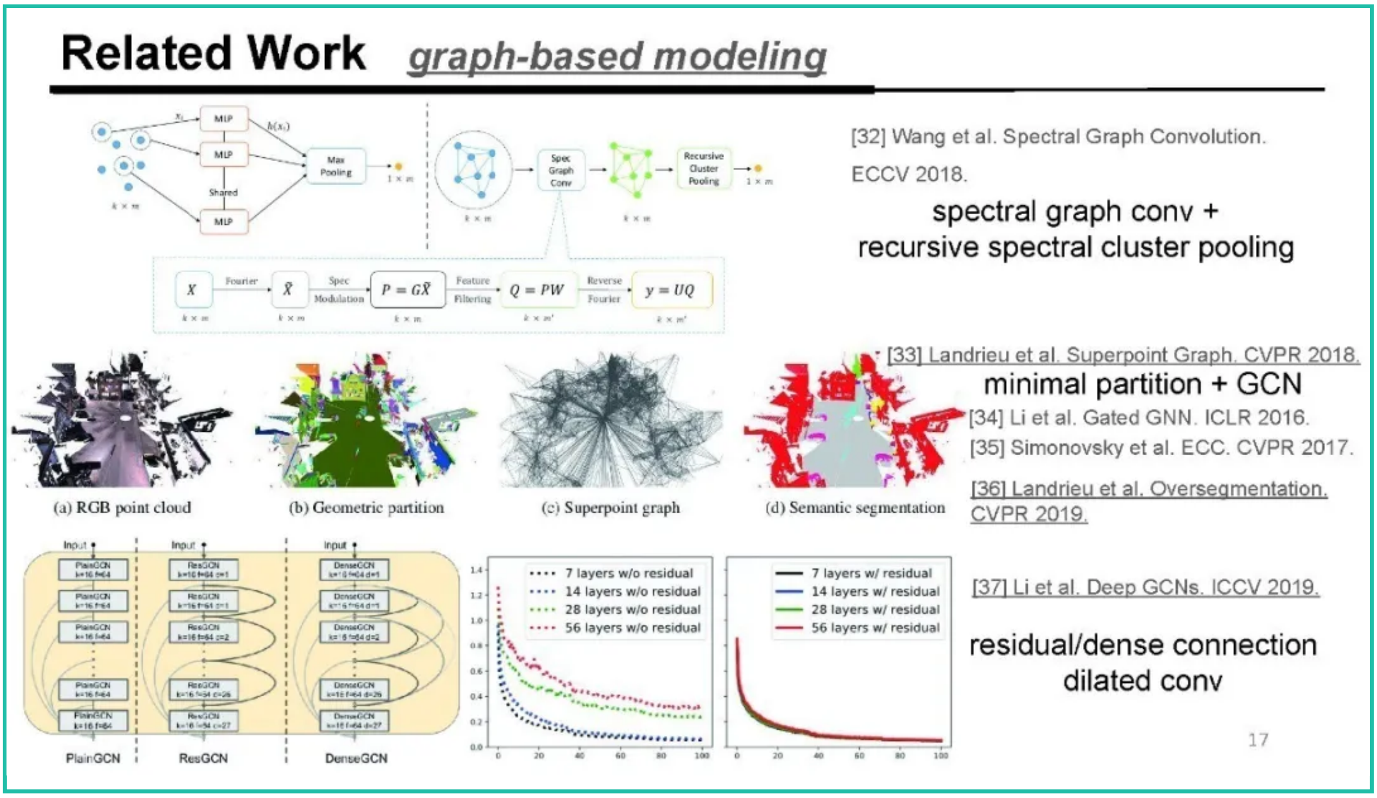
12

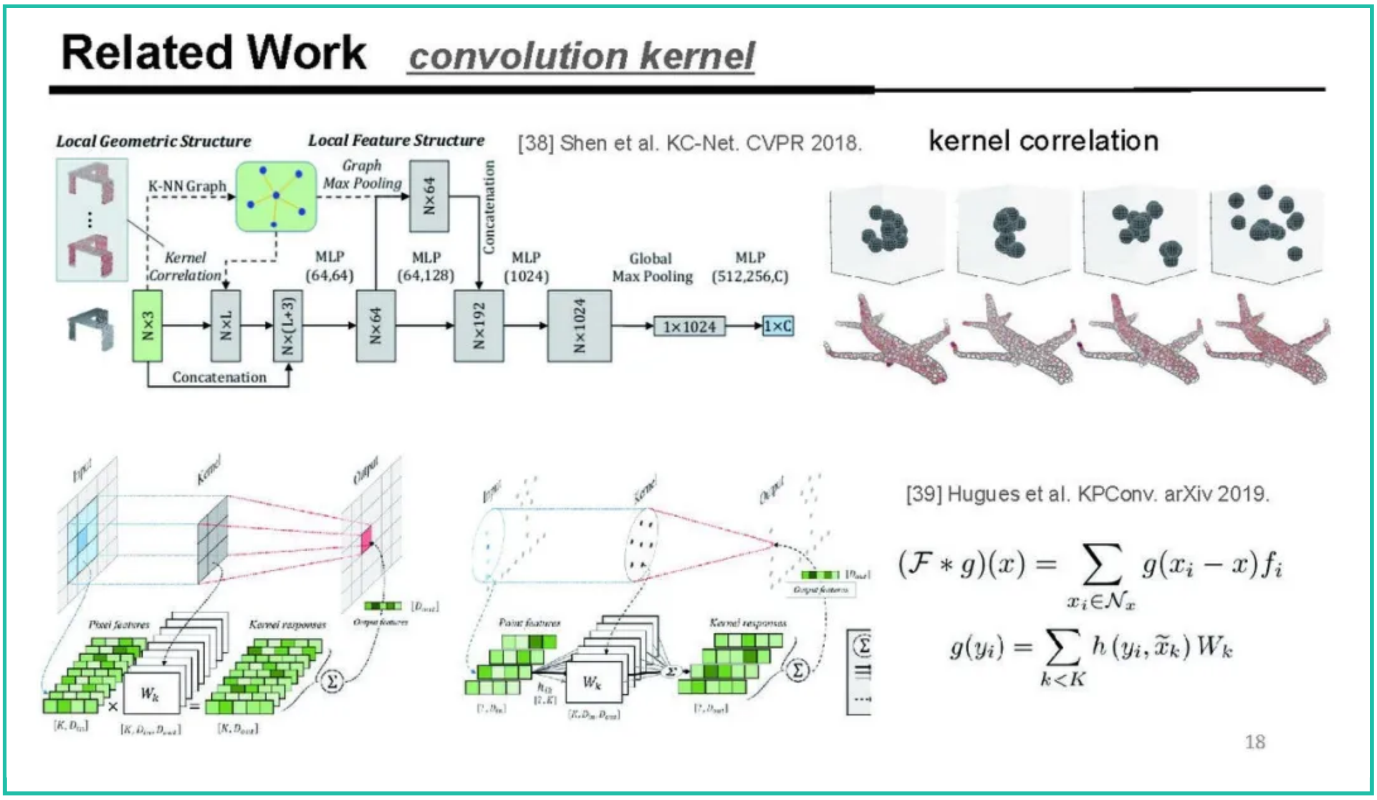

Related Work regular processing
[7] Zhou et al. VoxelNet. CVPR 2018.
Region Proposal Network
Convolutional Middle Layers
Feature Learning Network
Voxel Partition
Grouping
Random Sampling
Stacked Voxel Feature Encoding
Sparse 4D Tensor
3D grid + PointNet + 3D CNN
[8] Su et al. SPLATNet. CVPR 2018.
Input
Splat
Convolve
Slice
Segmentation
lattice + bilateral convolution + hash index
[9] Kiefel et al. Permutohedral Lattice CNNs. ICLR 2015.
[10] Jampani et al. Bilateral Neural Networks. CVPR 2016.
Extension
Convolution
Restriction
[11] Atzmon et al. PCNN. SIGGRAPH 2018.
"without any discretization or approximation"
13

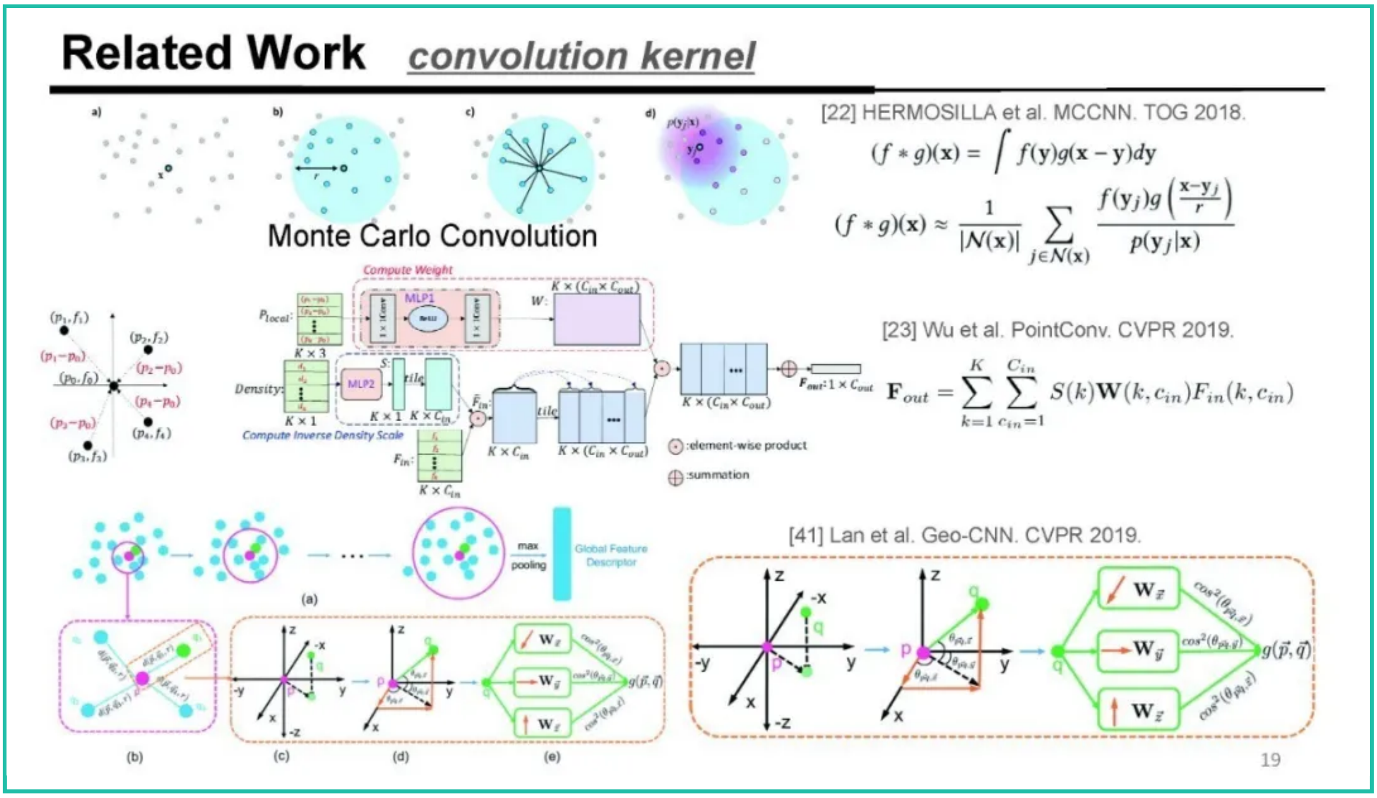

Related Work regular processing
[12] Li et al. PointCNN. NeurIPS 2018.
"simultaneously weight and permute the input features"
ALGORITHM 1: $\mathcal{X}$-Conv Operator
Input : $\mathbf{K}, p, \mathbf{P}, \mathbf{F}$
Output : $\mathbf{F}_p$ ▷ Features "projected", or "aggregated", into representative point p
1: $\mathbf{P}' \leftarrow \mathbf{P} - p$ ▷ Move $\mathbf{P}$ to local coordinate system of p
2: $\mathbf{F}_\delta \leftarrow MLP_\delta(\mathbf{P}')$ ▷ **Individually** lift each point into C_δ dimensional space
3: $\mathbf{F}_* \leftarrow [\mathbf{F}_\delta, \mathbf{F}]$ ▷ Concatenate $\mathbf{F}_\delta$ and $\mathbf{F}$, $\mathbf{F}_*$ is a $K \times (C_\delta + C_1)$ matrix
4: $\mathcal{X} \leftarrow MLP(\mathbf{P}')$ ▷ Learn the $K \times K$ $\mathcal{X}$-transformation matrix
5: $\mathbf{F}_\mathcal{X} \leftarrow \mathcal{X} \times \mathbf{F}_*$ ▷ Weight and permute $\mathbf{F}_*$ with the learnt $\mathcal{X}$
6: $\mathbf{F}_p \leftarrow \mathrm{Conv}(\mathbf{K}, \mathbf{F}_\mathcal{X})$ ▷ Finally, typical convolution between $\mathbf{K}$ and $\mathbf{F}_\mathcal{X}$
[13] Jiang et al. PointSIFT. arXiv 2018.
orientation-encoding
Scale-aware
(a)
(b)
(c)
identity
Input feature
Output feature
pointwise convolution
Concat
Orientation-encoding Point Convolutional layers
pointSIFT module
14


Related Work regular processing
[14] Li et al. SO-Net. CVPR 2018.
Self-Organizing Map + MLP
[15] Hua et al. Pointwise CNN. CVPR 2018.
$x_i^\ell = \sum_k w_k \frac{1}{|\Omega_i(k)|} \sum_{p_j \in \Omega_i(k)} x_j^{\ell-1}$
Kd-Tree
[16] Klokov et al. Kd-Net. ICCV 2017.
3D grid + PointNet + 3D CNN
200k points at once
[17] Rethage et al. FCPN. ECCV 2018.
15

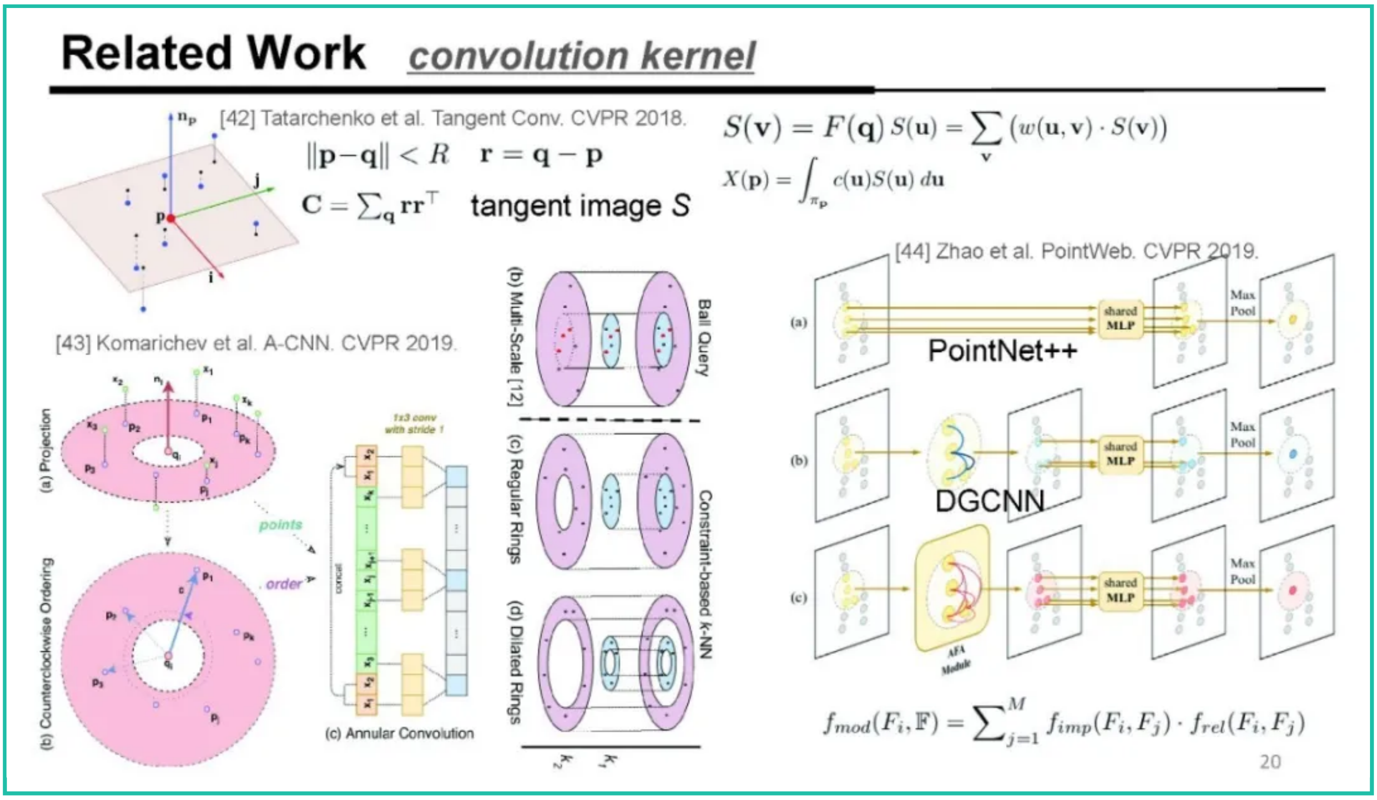

Related Work graph-based modeling
EdgeConv
[29] Wang et al. DGCNN. TOG 2019.
EdgeConv kNN
$x_i' = \square_{j:(i,j)\in\mathcal{E}} h_\Theta(x_i, x_j).$
$h_\Theta(x_i, x_j - x_i)$
[30] Te et al. Regularized GCNN. MM 2018.
$\mathbf{y} = g_\theta(\mathcal{L})\mathbf{x} = \sum_{k=0}^{K-1} \theta_k T_k(\mathcal{L})\mathbf{x}$
$a_{i,j} = \exp(-\beta\|\mathbf{p}_i - \mathbf{p}_j\|_2^2)$
$\sum_{i\sim j} a_{i,j}(y_i - y_j)^2$
kNN
Input Point Cloud
Chebyshev
Graph Construction
Graph Convolution
Laplacian 1
Laplacian 2
Laplacian 3
DENSE
GCONV
UPSAMP
GRAPH
[31] Valsesia et al. GAN for Point Cloud. ICLR 2019.
$\mathbf{h}_i^{l+1} = \sigma\left(\sum_{j\in\mathcal{N}_i^l} \frac{F_{\mathbf{w}^l}^l(\mathbf{h}_j^l - \mathbf{h}_i^l)\,\mathbf{h}_j^l}{|\mathcal{N}_i^l|} + \mathbf{h}_i^l\mathbf{W}^l + \mathbf{b}^l\right)$
learn domain (the graph) and features simultaneously
kNN
16


Related Work graph-based modeling
[32] Wang et al. Spectral Graph Convolution. ECCV 2018.
spectral graph conv + recursive spectral cluster pooling
(a) RGB point cloud
(b) Geometric partition
(c) Superpoint graph
(d) Semantic segmentation
[33] Landrieu et al. Superpoint Graph. CVPR 2018.
minimal partition + GCN
[34] Li et al. Gated GNN. ICLR 2016.
[35] Simonovsky et al. ECC. CVPR 2017.
[36] Landrieu et al. Oversegmentation. CVPR 2019.
PlainGCN
ResGCN
DenseGCN
7 layers w/o residual
14 layers w/o residual
28 layers w/o residual
56 layers w/o residual
7 layers w/ residual
14 layers w/ residual
28 layers w/ residual
56 layers w/ residual
[37] Li et al. Deep GCNs. ICCV 2019.
residual/dense connection
dilated conv
17

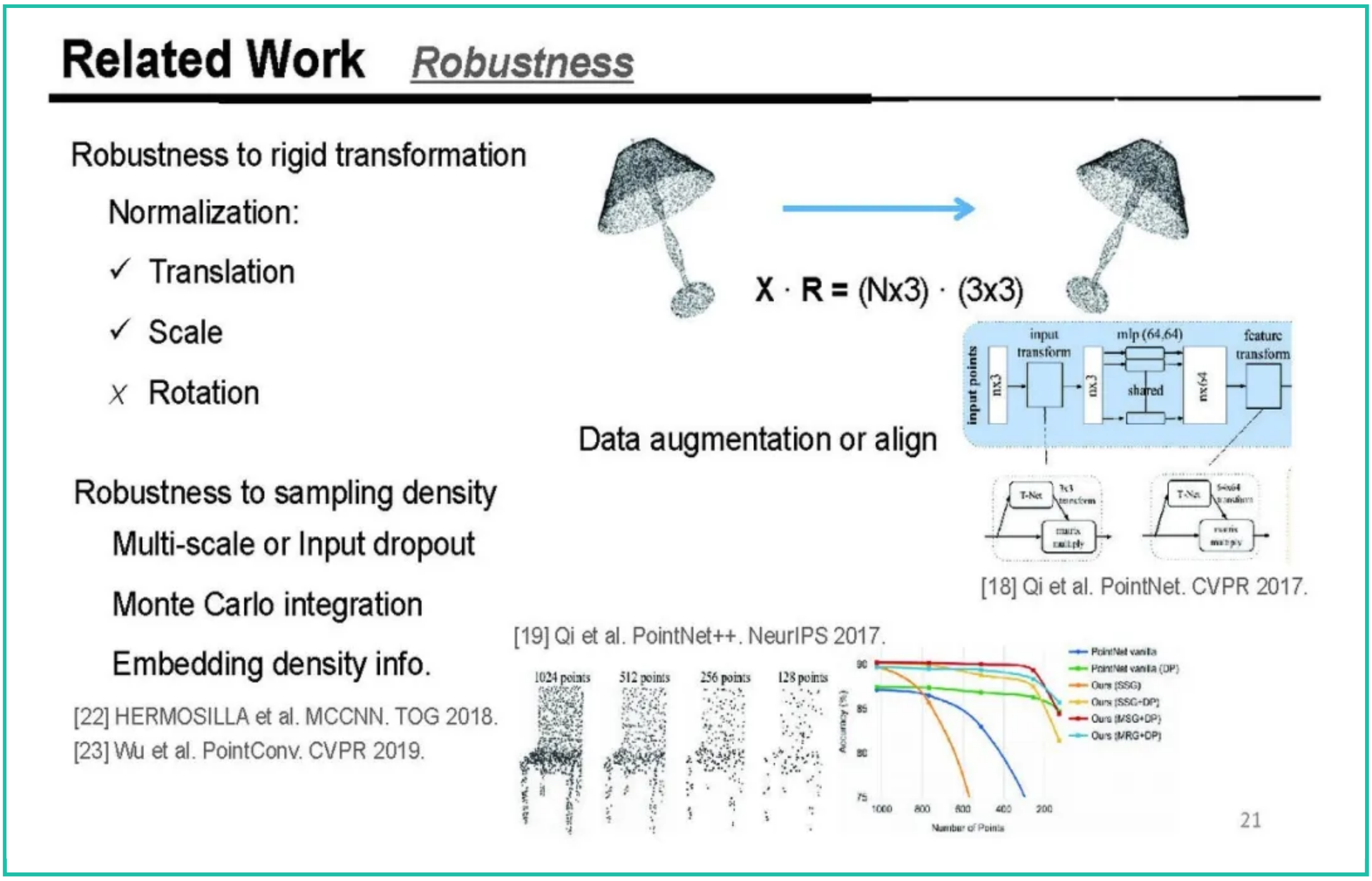

Related Work convolution kernel
Local Geometric Structure
Local Feature Structure
[38] Shen et al. KC-Net. CVPR 2018.
kernel correlation
K-NN Graph
Graph Max Pooling
Kernel Correlation
MLP (64,64)
MLP (64,128)
MLP (1024)
Global Max Pooling
MLP (512,256,C)
Concatenation
1×1024
1×C
[39] Hugues et al. KPConv. arXiv 2019.
$(\mathcal{F} * g)(x) = \sum_{x_i \in \mathcal{N}_x} g(x_i - x) f_i$
$g(y_i) = \sum_{k<K} h(y_i, \tilde{x}_k) W_k$
18

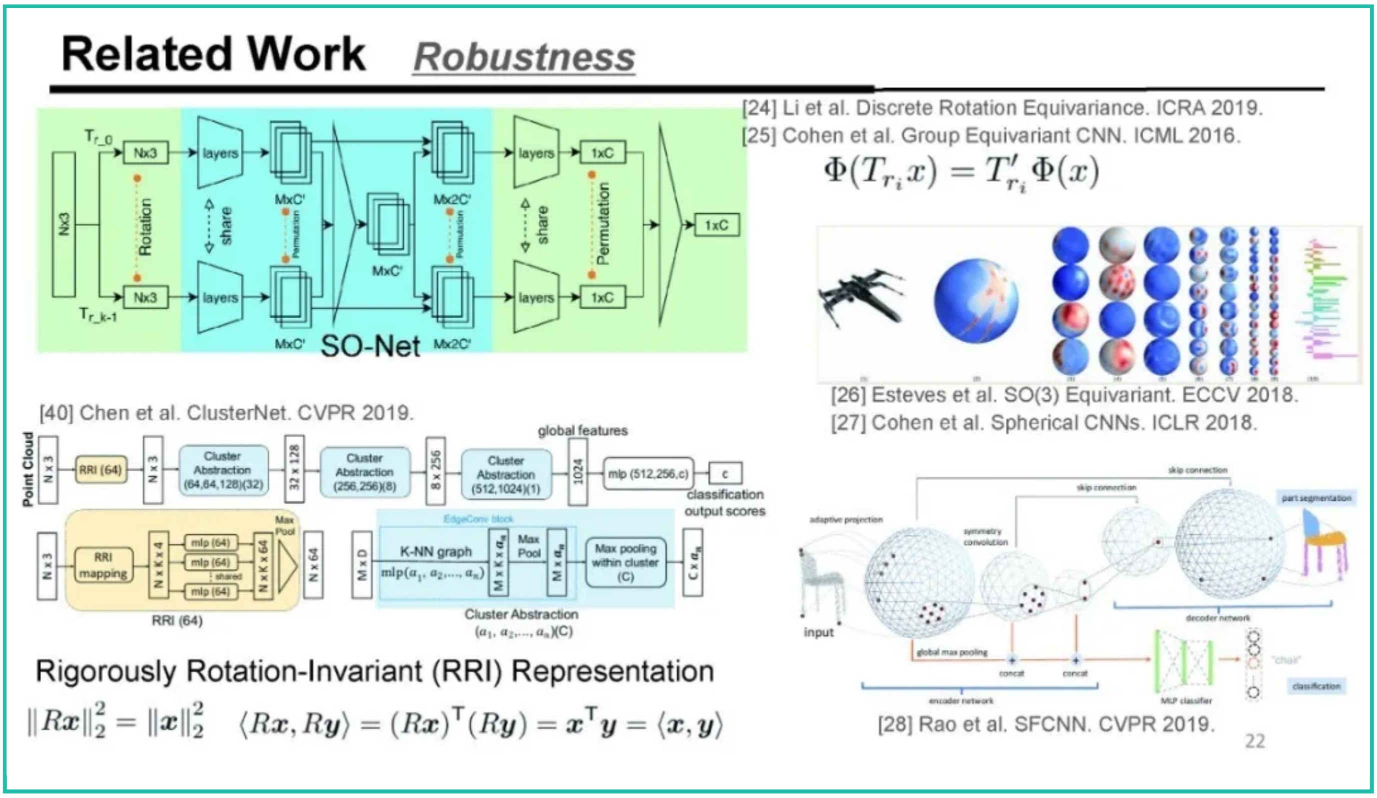


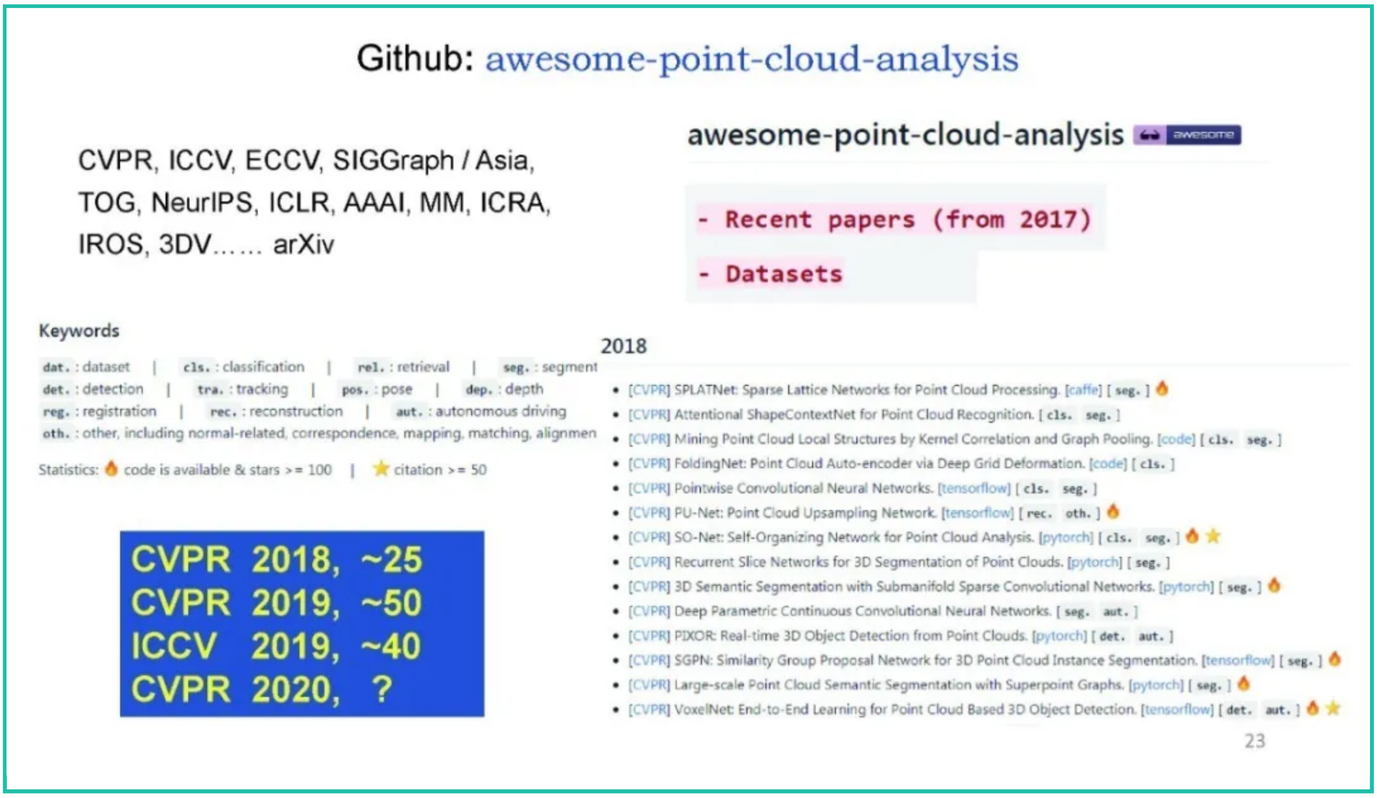
Related Work
convolution kernel
[22] HERMOSILLA et al. MCCNN. TOG 2018.
$(f * g)(\mathbf{x}) = \int f(\mathbf{y})g(\mathbf{x}-\mathbf{y})d\mathbf{y}$
$(f * g)(\mathbf{x}) \approx \frac{1}{|\mathcal{N}(\mathbf{x})|}\sum_{j\in\mathcal{N}(\mathbf{x})}\frac{f(\mathbf{y}_j)g\left(\frac{\mathbf{x}-\mathbf{y}_j}{r}\right)}{p(\mathbf{y}_j|\mathbf{x})}$
Monte Carlo Convolution
Compute Weight
Compute Inverse Density Scale
[23] Wu et al. PointConv. CVPR 2019.
$\mathbf{F}_{out} = \sum_{k=1}^{K}\sum_{c_{in}=1}^{C_{in}} S(k)\mathbf{W}(k, c_{in})F_{in}(k, c_{in})$
:element-wise product
:summation
max pooling
Global Feature Descriptor
[41] Lan et al. Geo-CNN. CVPR 2019.
19

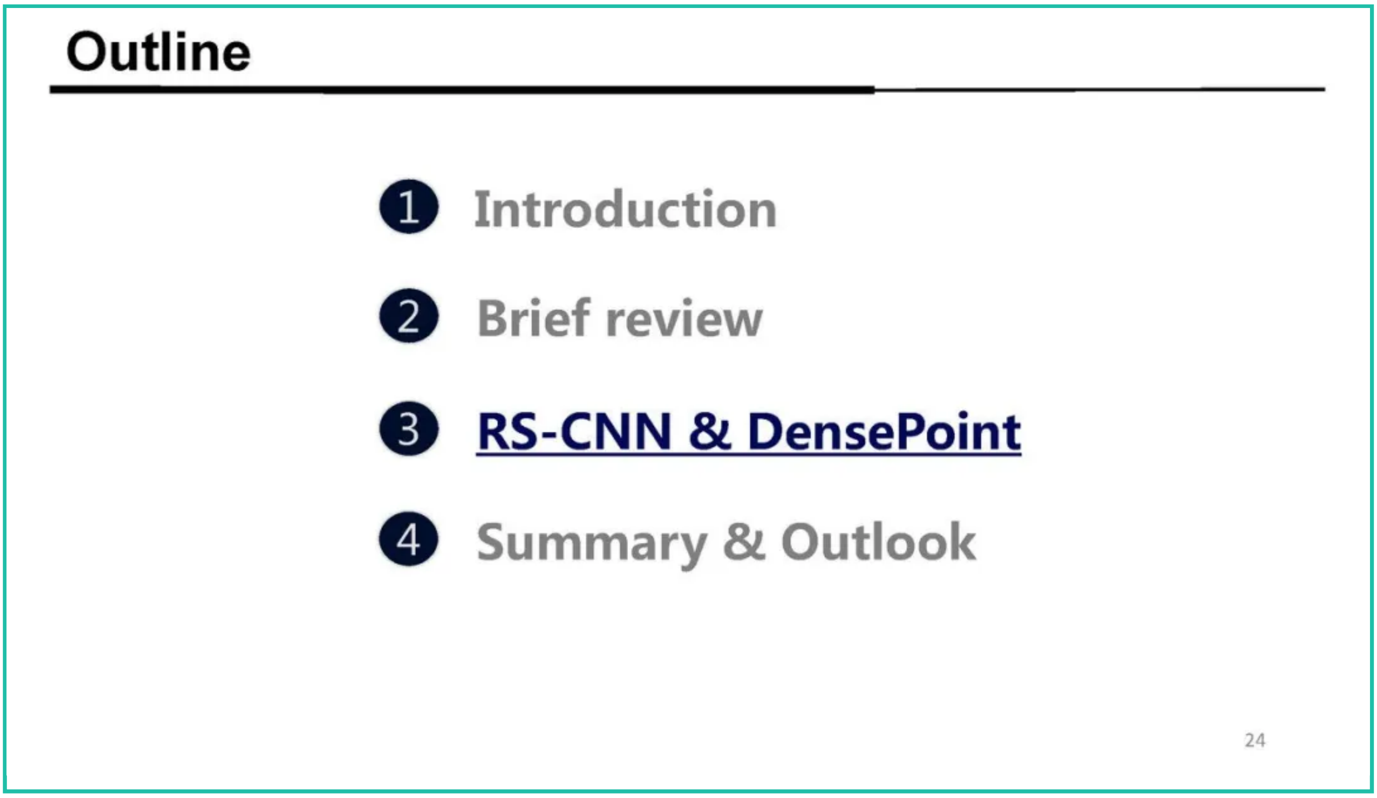


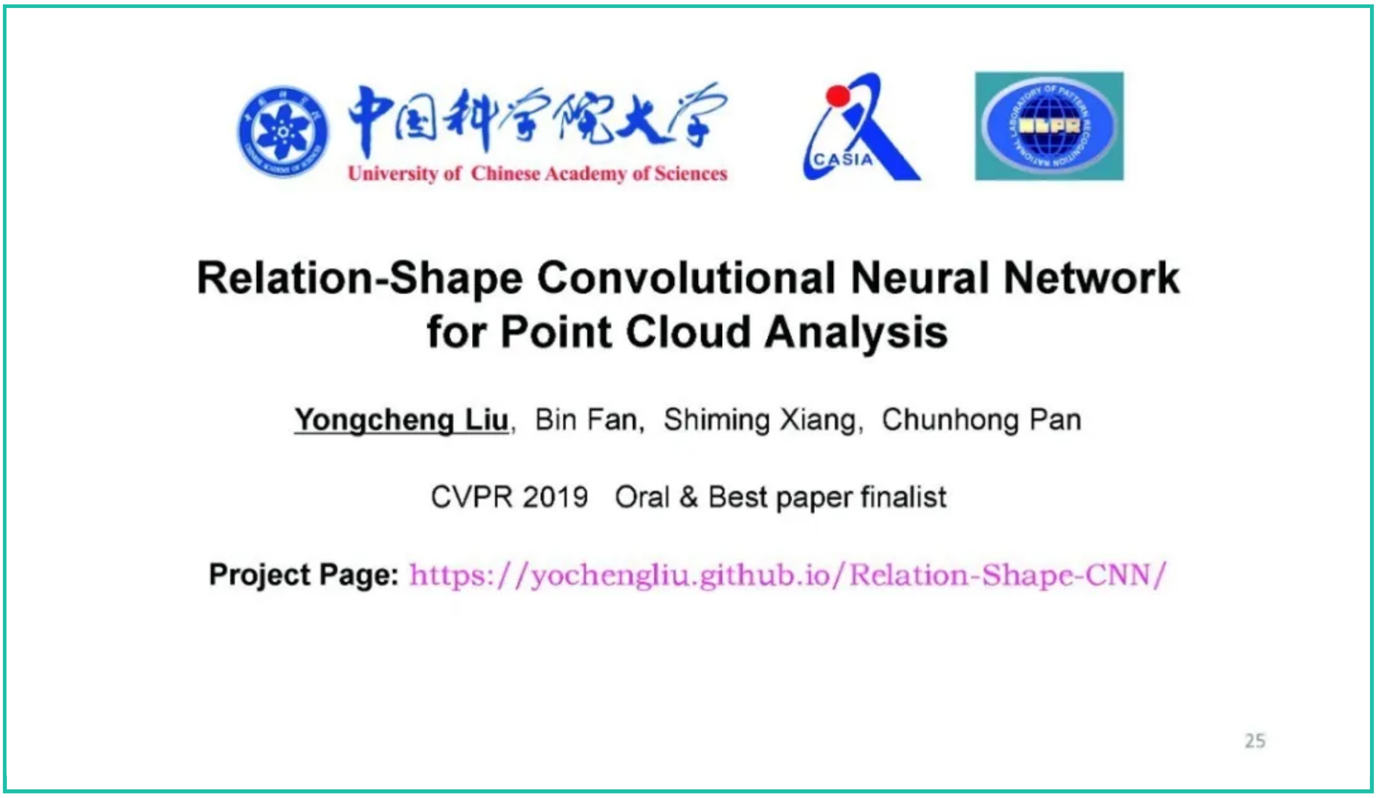
Related Work
convolution kernel
[42] Tatarchenko et al. Tangent Conv. CVPR 2018.
$\|\mathbf{p}-\mathbf{q}\| < R \quad \mathbf{r} = \mathbf{q}-\mathbf{p}$
$\mathbf{C} = \sum_{\mathbf{q}} \mathbf{r}\mathbf{r}^{\top}$ tangent image S
$S(\mathbf{v}) = F(\mathbf{q})\ S(\mathbf{u}) = \sum_{\mathbf{v}} (w(\mathbf{u},\mathbf{v}) \cdot S(\mathbf{v}))$
$X(\mathbf{p}) = \int_{\pi_{\mathbf{p}}} c(\mathbf{u})S(\mathbf{u})\, d\mathbf{u}$
[43] Komarichev et al. A-CNN. CVPR 2019.
(a) Projection
(b) Counterclockwise Ordering
points
order
1x3 conv with stride 1
concat
(c) Annular Convolution
(b) Multi-Scale [12]
Ball Query
(c) Regular Rings
Constraint-based k-NN
(d) Dilated Rings
[44] Zhao et al. PointWeb. CVPR 2019.
PointNet++
DGCNN
shared MLP
Max Pool
AFA Module
$f_{mod}(F_i, \mathbb{F}) = \sum_{j=1}^{M} f_{imp}(F_i, F_j) \cdot f_{rel}(F_i, F_j)$
20

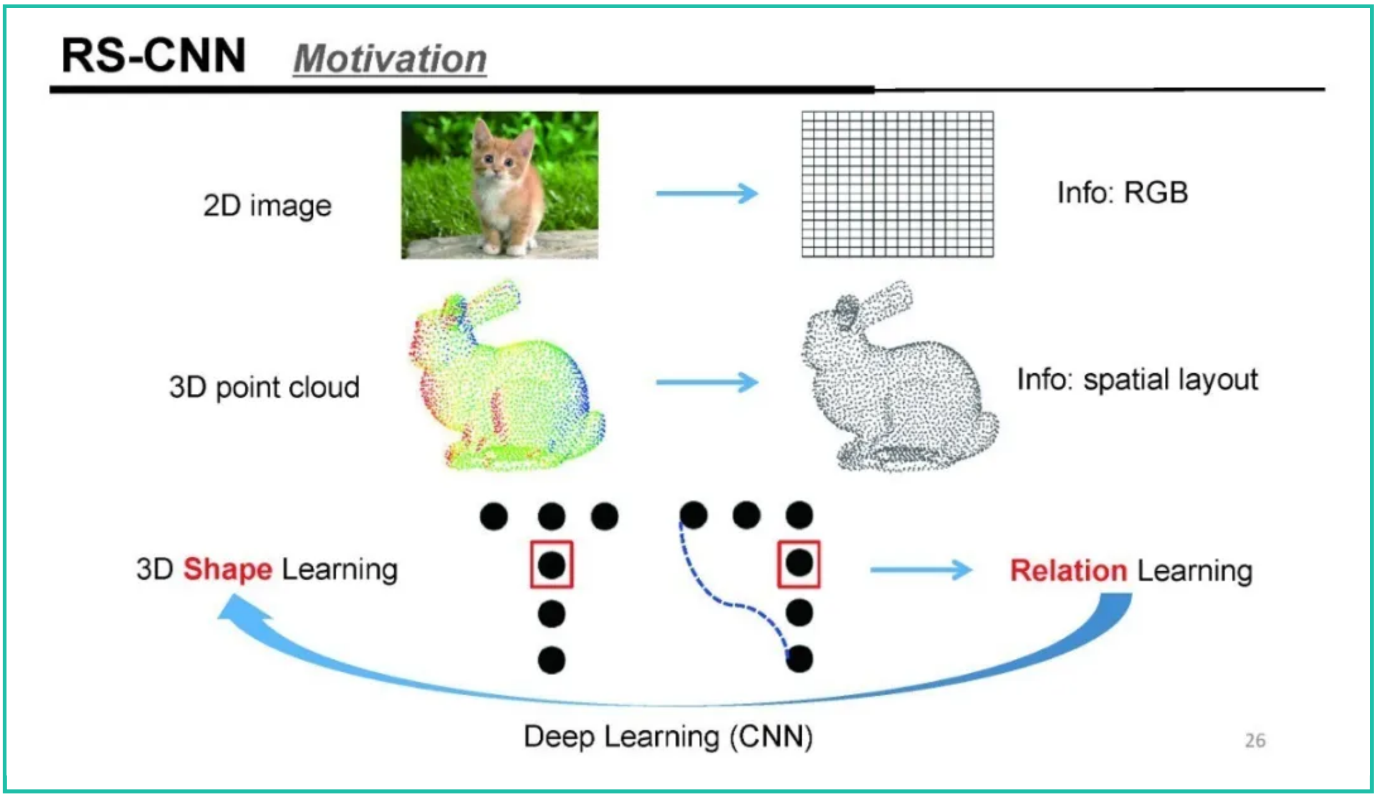

Related Work Robustness
Robustness to rigid transformation
Normalization:
✓ Translation
✓ Scale
✗ Rotation
X · R = (Nx3) · (3x3)
Data augmentation or align
input transform
mlp (64,64)
feature transform
shared
input points
T-Net
[18] Qi et al. PointNet. CVPR 2017.
Robustness to sampling density
Multi-scale or Input dropout
Monte Carlo integration
Embedding density info.
[22] HERMOSILLA et al. MCCNN. TOG 2018.
[23] Wu et al. PointConv. CVPR 2019.
[19] Qi et al. PointNet++. NeurIPS 2017.
1024 points
512 points
256 points
128 points
Number of Points
21


Related Work Robustness
SO-Net
Rotation
Permutation
share
layers
[24] Li et al. Discrete Rotation Equivariance. ICRA 2019.
[25] Cohen et al. Group Equivariant CNN. ICML 2016.
$\Phi(T_{r_i}x) = T'_{r_i}\Phi(x)$
[26] Esteves et al. SO(3) Equivariant. ECCV 2018.
[27] Cohen et al. Spherical CNNs. ICLR 2018.
[40] Chen et al. ClusterNet. CVPR 2019.
Point Cloud
RRI (64)
Cluster Abstraction
global features
classification output scores
RRI mapping
K-NN graph
Max pooling within cluster
Cluster Abstraction
Rigorously Rotation-Invariant (RRI) Representation
$\|R\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}\|_2^2 \quad \langle R\boldsymbol{x}, R\boldsymbol{y}\rangle = (R\boldsymbol{x})^\top(R\boldsymbol{y}) = \boldsymbol{x}^\top\boldsymbol{y} = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$
input
[28] Rao et al. SFCNN. CVPR 2019.
22


Github: awesome-point-cloud-analysis
CVPR, ICCV, ECCV, SIGGraph / Asia, TOG, NeurIPS, ICLR, AAAI, MM, ICRA, IROS, 3DV…… arXiv
awesome-point-cloud-analysis
- Recent papers (from 2017)
- Datasets
Keywords
dat. : dataset | cls. : classification | rel. : retrieval | seg. : segment
det. : detection | tra. : tracking | pos. : pose | dep. : depth
reg. : registration | rec. : reconstruction | aut. : autonomous driving
oth. : other, including normal-related, correspondence, mapping, matching, alignmen
Statistics: code is available & stars >= 100 | citation >= 50
CVPR 2018, ~25
CVPR 2019, ~50
ICCV 2019, ~40
CVPR 2020, ?
2018
[CVPR] SPLATNet: Sparse Lattice Networks for Point Cloud Processing. [caffe] [seg.]
[CVPR] Attentional ShapeContextNet for Point Cloud Recognition. [cls. seg.]
[CVPR] Mining Point Cloud Local Structures by Kernel Correlation and Graph Pooling. [code] [cls. seg.]
[CVPR] FoldingNet: Point Cloud Auto-encoder via Deep Grid Deformation. [code] [cls.]
[CVPR] Pointwise Convolutional Neural Networks. [tensorflow] [cls. seg.]
[CVPR] PU-Net: Point Cloud Upsampling Network. [tensorflow] [rec. oth.]
[CVPR] SO-Net: Self-Organizing Network for Point Cloud Analysis. [pytorch] [cls. seg.]
[CVPR] Recurrent Slice Networks for 3D Segmentation of Point Clouds. [pytorch] [seg.]
[CVPR] 3D Semantic Segmentation with Submanifold Sparse Convolutional Networks. [pytorch] [seg.]
[CVPR] Deep Parametric Continuous Convolutional Neural Networks. [seg. aut.]
[CVPR] PIXOR: Real-time 3D Object Detection from Point Clouds. [pytorch] [det. aut.]
[CVPR] SGPN: Similarity Group Proposal Network for 3D Point Cloud Instance Segmentation. [tensorflow] [seg.]
[CVPR] Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs. [pytorch] [seg.]
[CVPR] VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection. [tensorflow] [det. aut.]
23


Outline
1 Introduction
2 Brief review
3 RS-CNN & DensePoint
4 Summary & Outlook
24


中国科学院大学
University of Chinese Academy of Sciences
CASIA
Relation-Shape Convolutional Neural Network for Point Cloud Analysis
Yongcheng Liu, Bin Fan, Shiming Xiang, Chunhong Pan
CVPR 2019 Oral & Best paper finalist
Project Page: https://yochengliu.github.io/Relation-Shape-CNN/
25


RS-CNN Motivation
2D image
Info: RGB
3D point cloud
Info: spatial layout
3D Shape Learning
Relation Learning
Deep Learning (CNN)
26

RS-CNN Method: Relation-Shape Conv

local point subset $P_{\text{sub}} \subset \mathbb{R}^3$ → spherical neighborhood: x_i + $x_j \in \mathcal{N}(x_i)$

$$\mathbf{f}_{P_{\text{sub}}} = \sigma\big(\mathcal{A}(\{\mathcal{T}(\mathbf{f}_{x_j}),\ \forall x_j\})\big) \quad \Longleftrightarrow \quad y = \sigma(\sum \mathbf{W} * \mathbf{X})$$

$\mathcal{T}$: feature transformation $\mathcal{A}$: feature aggregation

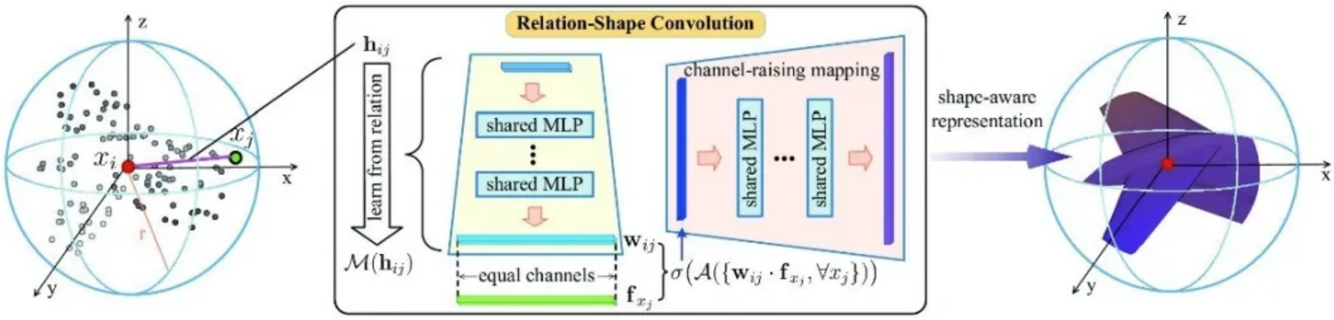
- Permutation invariance: only when A is symmetric and T is shared over each point
- Limitations of CNN: weight is not shared
  gradient only w.r.t single point - implicit $\mathcal{T}(\mathbf{f}_{x_j}) = \mathbf{w}_j \cdot \mathbf{f}_{x_j}$
- Conversion: learn from relation $\mathcal{T}(\mathbf{f}_{x_j}) = \mathbf{w}_{ij} \cdot \mathbf{f}_{x_j} = \mathcal{M}(\mathbf{h}_{ij}) \cdot \mathbf{f}_{x_j}$

$\mathbf{h}_{ij}$: predefined geometric priors → low-level relation

$\mathbf{f}_{P_{\text{sub}}} = \sigma\big(\mathcal{A}(\{\mathcal{M}(\mathbf{h}_{ij}) \cdot \mathbf{f}_{x_j},\ \forall x_j\})\big)$ $\mathcal{M}$: mapping function(shared MLP) → high-level relation

27

RS-CNN Method



$$\mathbf{f}_{P_{\text{sub}}} = \sigma\big(\mathcal{A}(\{\mathcal{M}(\mathbf{h}_{ij}) \cdot \mathbf{f}_{x_j},\ \forall x_j\})\big)$$

high-level relation encoding + channel raising mapping

low-level relation $\mathbf{h}_{ij}$: (3D Euclidean distance, $x_i - x_j$, x_i, x_j) 10 channels

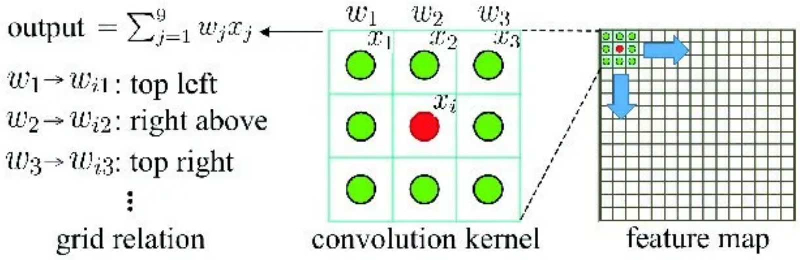
28

RS-CNN *RS-Conv: Properties*

$$\mathbf{f}_{P_{\text{sub}}} = \sigma\big(\mathcal{A}(\{\mathcal{M}(\mathbf{h}_{ij}) \cdot \mathbf{f}_{x_j},\ \forall x_j\})\big)$$

- ✓ Permutation invariance
- ✓ Robustness to rigid transformation in Relation Learning, *e.g.*, 3D Euclidean distance
- ✓ Points' interaction
- ✓ Weight sharing

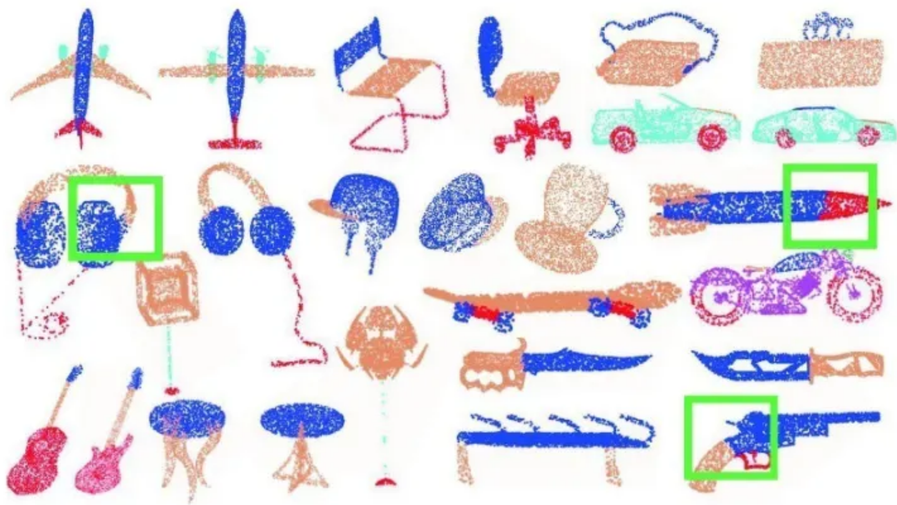
Revisiting 2D Conv:



RS-Conv with relation learning is more general and can be applied to model 2D grid spatial relationship.

29

RS-CNN *ShapePart Segmentation*



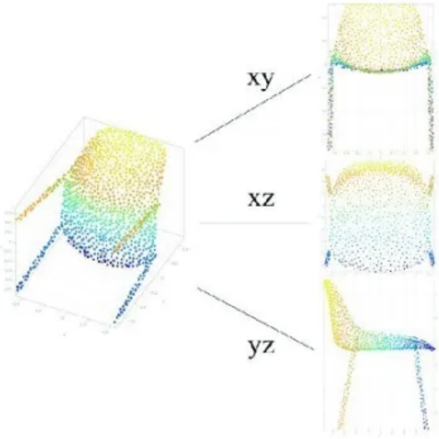
Diverse, confusing shapes

30

RS-CNN *Geometric priors*

Princeton ModelNet40

$$\mathbf{f}_{P_{\text{sub}}} = \sigma\big(\mathcal{A}(\{\mathcal{M}(\mathbf{h}_{ij}) \cdot \mathbf{f}_{x_j},\ \forall x_j\})\big)$$



model	low-level relation $\mathbf{h}$	channels	acc.
A	(3D-Ed)	1	92.5
B	(3D-Ed, $x_i - x_j$)	4	93.0
C	(3D-Ed, $x_i - x_j$, x_i, x_j)	10	**93.6**
D	(3D-cosd, x_i^{nor}, x_j^{nor})	7	92.8
E	(2D-Ed, $x_i' - x_j'$, x_i', x_j')	10	$\approx$ 92.2

low-level relation $\mathbf{h}$	channels	acc.
(XY-Ed, $x_i^{\text{xy}} - x_j^{\text{xy}}$, x_i^{xy}, x_j^{xy})	10	92.1
(XZ-Ed, $x_i^{\text{xz}} - x_j^{\text{xz}}$, x_i^{xz}, x_j^{xz})	10	92.1
(YZ-Ed, $x_i^{\text{yz}} - x_j^{\text{yz}}$, x_i^{yz}, x_j^{yz})	10	92.2
fusion of above three views		92.5

31

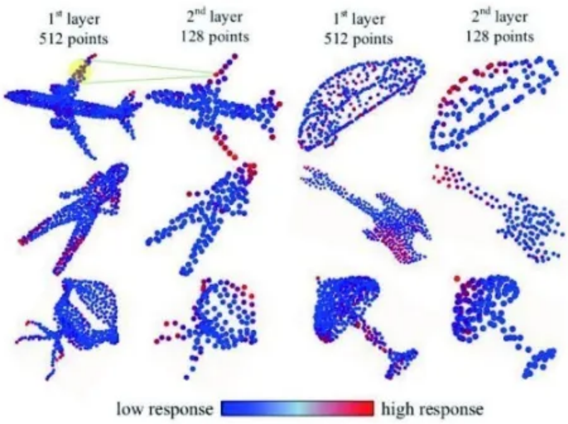
RS-CNN *Model analysis*

Robustness to point permutation and rigid transformation

relation: 3D
Euclidean distance

method	acc.	perm.	+0.2	-0.2	90°	180°
PointNet [24]	88.7	88.7	70.8	70.6	42.5	38.6
PointNet++ [26]	88.2†	88.2	88.2	88.2	47.9	39.7
Ours	**90.3**†	**90.3**	**90.3**	**90.3**	**90.3**	**90.3**

$$\mathbf{f}_{P_{\text{sub}}} = \sigma\big(\mathcal{A}(\{\mathcal{M}(\mathbf{h}_{ij}) \cdot \mathbf{f}_{x_j},\ \forall x_j\})\big)$$



Model complexity

method	#params	#FLOPs/sample
PointNet [24]	3.50M	440M
PointNet++ [21]	1.48M	1684M
PCNN [21]	8.20M	**294M**
Ours	**1.41M**	295M

32

University of Chinese Academy of Sciences

DensePoint: Learning Densely Contextual Representation for Efficient Point Cloud Processing

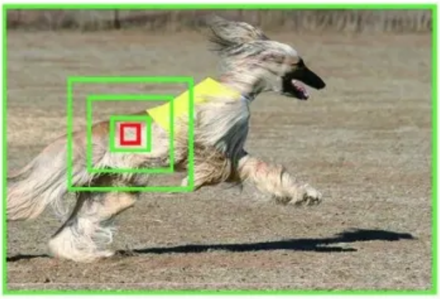
Yongcheng Liu, Bin Fan, Gaofeng Meng, Jiwen Lu, Shiming Xiang, Chunhong Pan

ICCV 2019

Code: https://github.com/Yochengliu/DensePoint

33

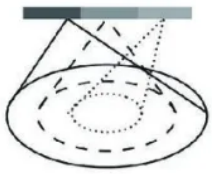
DensePoint *Motivation*



Context: potential semantic dependencies between a target pattern and its surroundings

Multi-scale learning – high complexity

- parameters
- FLOPs
- scale limitation
- unintuitive (scale ←→ semantic level)



✓ Efficient solution using deep learning?
✓ How does it perform on point cloud (shape learning)?

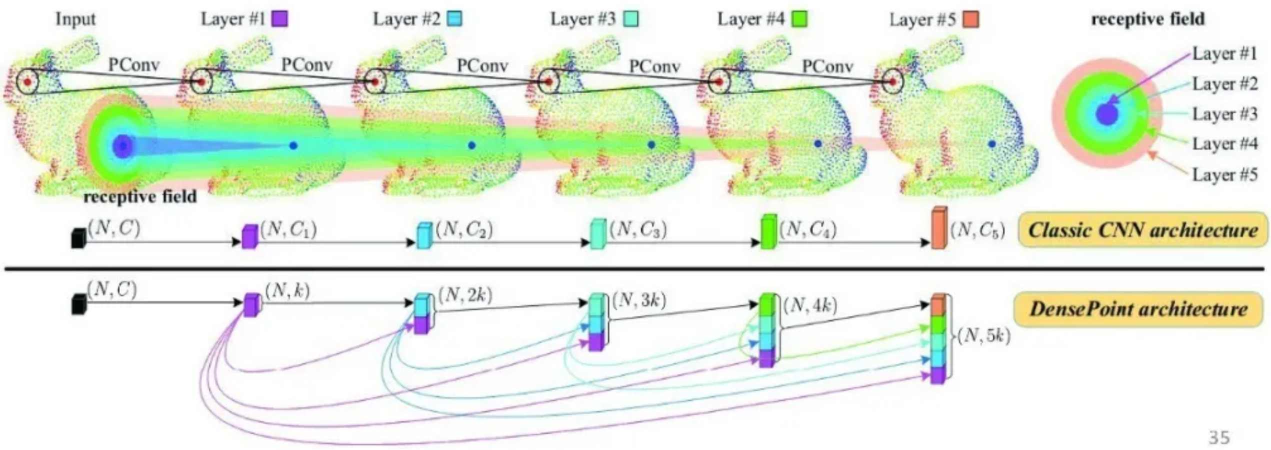
34

DensePoint *Method*

key idea: multi-level receptive fields + efficient conv on point cloud

dense connections + efficient point convolution

progressively aggregate multi-scale info. in an ***organic*** manner!



35

DensePoint *Method: efficient PConv*

$$\mathbf{f}_{\mathcal{N}(x)} = \rho\left(\{\phi(\mathbf{f}_{x_n}),\ \forall x_n \in \mathcal{N}(x)\}\right)$$

ø: single-layer perceptron

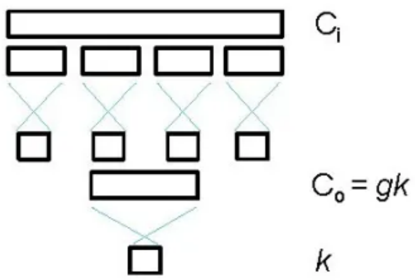
enhanced PConv: filter grouping

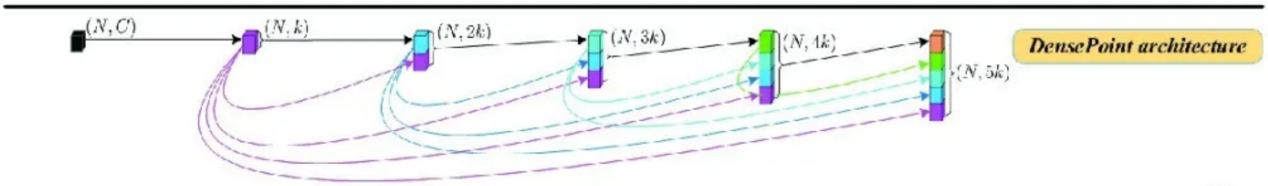
$$\mathbf{f}_{\mathcal{N}(x)} = \psi\left(\rho\left(\{\widehat{\phi}(\mathbf{f}_{x_n}),\ \forall x_n \in \mathcal{N}(x)\}\right)\right)$$

C_i*k

vs.

$C_i*gk/g + 4k^2$



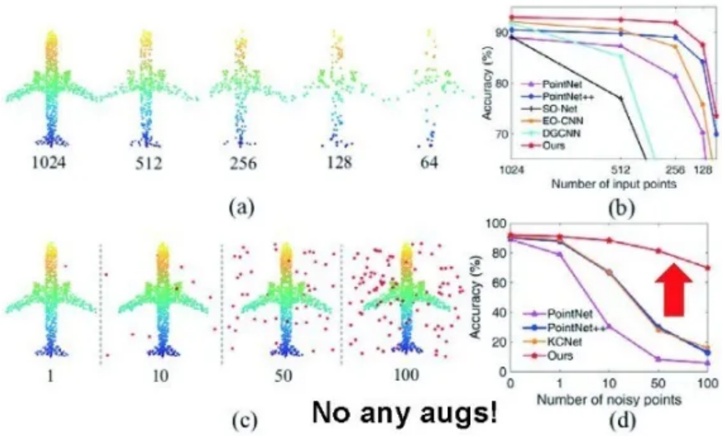


36

DensePoint *Shape classification*

ModelNet40 benchmark

Robustness to sampling density and noise



No any augs!

method	input	#points	M40	M10
Pointwise-CNN [12]	pnt	1k	86.1	-
Deep Sets [60]	pnt	1k	87.1	-
ECC [40]	pnt	1k	87.4	90.8
PointNet [31]	pnt	1k	89.2	-
SCN [55]	pnt	1k	90.0	-
Kd-Net(depth=10) [21]	pnt	1k	90.6	93.3
PointNet++ [33]	pnt	1k	90.7	-
MC-Conv [11]	pnt	1k	90.9	-
KCNet [39]	pnt	1k	91.0	94.4
MRTNet [4]	pnt	1k	91.2	-
SpecGCN [49]	pnt	1k	91.5	-
DGCNN [52]	pnt	1k	92.2	-
PointCNN [26]	pnt	1k	92.2	-
PCNN [1]	pnt	1k	92.3	94.9
Ours	**pnt**	**1k**	**93.2**	**96.6**
SO-Net [24]	pnt	2k	90.9	94.1
Kd-Net(depth=15) [21]	pnt	32k	91.8	94.0
O-CNN [50]	pnt, nor	-	90.6	-
Spec-GCN [49]	pnt, nor	1k	91.8	-
PointNet++ [33]	pnt, nor	5k	91.9	-
SpiderCNN [56]	pnt, nor	5k	92.4	-
SO-Net [24]	pnt, nor	5k	93.4	95.7

37

DensePoint *model complexity*

method	#params	#FLOPs/sample	acc.(%)
PointNet [31]	3.50M	440M	89.2
PointNet++ [26]	1.48M	1684M	90.7
DGCNN [26]	1.84M	2767M	92.2
SpecGCN [26]	2.05M	1112M	91.5
KCNet [39]	0.90M	-	91.0
PCNN [26]	8.20M	294M	92.3
PointCNN [26]	0.60M	1581M	92.2
Ours ($k=12$, $L=11$)	0.56M	294M	92.1
Ours ($k=24$, $L=11$)	0.67M	651M	93.2
Ours ($k=24$, $L=6$)	**0.53M**	**148M**	92.1

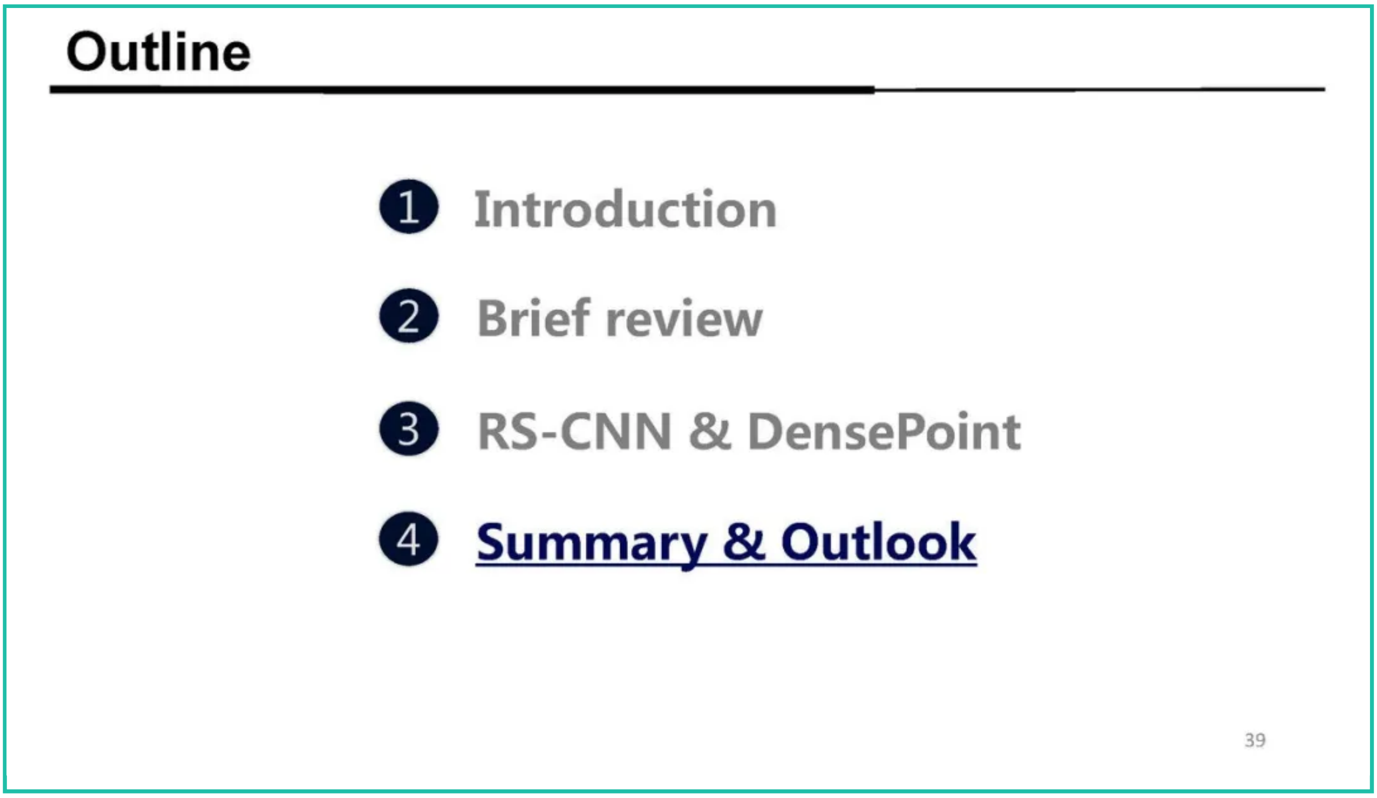
1024 points

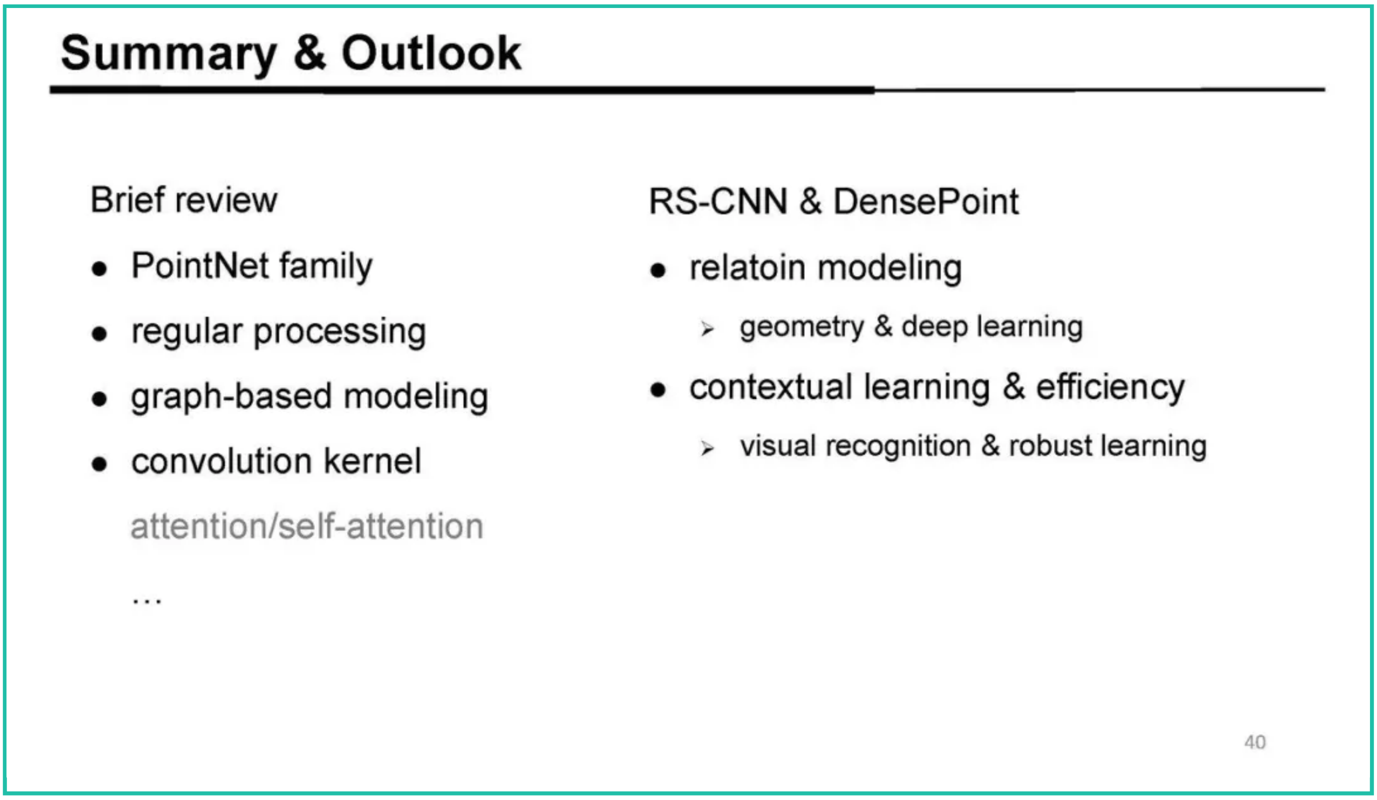
method	#points	Time (ms)		Memory (GB)	
		training	test	training	test
PointNet [31]	1024	55	22	1.318	0.469
PointNet++ [33]	1024	195	47	8.311	2.305
DGCNN [52]	1024	300	68	4.323	1.235
PointCNN [26]	1024	55	38	2.501	1.493
Ours (k=24, L=11)	1024	21	10	3.745	1.228
Ours (k=24, L=6)	1024	10	5	1.468	0.886
Ours (k=24, L=11)	4096	21	10	7.503	1.767
Ours (k=24, L=6)	4096	10	5	2.417	1.638
Ours (k=24, L=11)	8192	21	10	14.521	3.027
Ours (k=24, L=6)	8192	10	5	4.335	2.776

batchsize = 16

Titan Xp

38


Outline
1 Introduction
2 Brief review
3 RS-CNN & DensePoint
4 Summary & Outlook
39


Summary & Outlook
Brief review
● PointNet family
● regular processing
● graph-based modeling
● convolution kernel
attention/self-attention
…
RS-CNN & DensePoint
● relatoin modeling
➢ geometry & deep learning
● contextual learning & efficiency
➢ visual recognition & robust learning
40

Summary & Outlook

Advantages

- ✓ raw sensor data, e.g., Lidar
- ✓ simple representation: N * (x, y, z, color, normal…)
- ✓ better 3D shape capturing

Why emerging?

- ✓ autonomous driving
- ✓ AR & VR
- ✓ robot manipulation
- ✓ Geomatics
- ✓ 3D face & medical
- ✓ AI-assisted shape design in 3D game and animation, etc.
- ✓ open problem, flexible

- efficiency in large-scale point cloud
- multi-sensor/multi-modal
- reconstruction
- high-precision
- robustness

- ✓ geometric DL, segmentation (instance), detection, completion, registration…
- ✓ capsule, GAN, one-shot/zero-shot, meta-learning, NAS

Welcome to the world of 3D point cloud!

41

References

[1] Wu et al. 3D ShapeNets: A Deep Representation for Volumetric Shapes. CVPR 2015.

[2] Su et al. Multi-view Convolutional Neural Networks for 3D Shape Recognition. ICCV 2015.

[3] Mo et al. PartNet: A Large-scale Benchmark for Fine-grained and Hierarchical Part-level 3D Object Understanding. CVPR 2019.

[4] Armeni et al. 3d semantic parsing of large-scale indoor spaces. CVPR 2016.

[5] Hackel et al. Semantic3d. net: A new large-scale point cloud classification benchmark. ISPRS 2017.

[6] Dai et al. ScanNet: Richly-annotated 3D Reconstructions of Indoor Scenes. CVPR 2017.

[7] Zhou et al. VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection. CVPR 2018.

[8] Su et al. SPLATNet: Sparse Lattice Networks for Point Cloud Processing. CVPR 2018.

[9] Kiefel et al. Permutohedral Lattice CNNs. ICLR 2015.

[10] Jampani et al. Learning sparse high dimensional filters: Image filtering, dense CRFs and bilateral neural networks. CVPR 2016.

[11] Atzmon et al. Point Convolutional Neural Networks by Extension Operators. SIGGRAPH 2018.

[12] Li et al. PointCNN: Convolution On X-Transformed Points. NeurIPS 2018.

[13] Jiang et al. PointSIFT: A SIFT-like Network Module for 3D Point Cloud Semantic Segmentation. arXiv 2018. 42

References

[14] Li et al. SO-Net: Self-Organizing Network for Point Cloud Analysis. CVPR 2018.
[15] Hua et al. Pointwise Convolutional Neural Networks. CVPR 2018.
[16] Klokov et al. Escape from Cells: Deep Kd-Networks for the Recognition of 3D Point Cloud Models. ICCV 2017.
[17] Rethage et al. Fully-Convolutional Point Networks for Large-Scale Point Clouds. ECCV 2017.
[18] Qi et al. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. CVPR 2017.
[19] Qi et al. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. NeurIPS 2017.
[20] Qi et al. Frustum PointNets for 3D Object Detection from RGB-D Data. CVPR 2018.
[21] Liu et al. FlowNet3D: Learning Scene Flow in 3D Point Clouds. CVPR 2019.
[22] HERMOSILLA et al. Monte Carlo Convolution for Learning on Non-Uniformly Sampled Point Clouds. TOG 2018.
[23] Wu et al. PointConv: Deep Convolutional Networks on 3D Point Clouds. CVPR 2019.
[24] Li et al. Discrete Rotation Equivariance for Point Cloud Recognition. ICRA 2019.
[25] Cohen et al. Group Equivariant Convolutional Networks. ICML 2016.
[26] Esteves et al. Learning SO(3) Equivariant Representations with Spherical CNNs. ECCV 2018.
[27] Cohen et al. Spherical CNNs. ICLR 2018.
[28] Rao et al. Spherical Fractal Convolution Neural Networks for Point Cloud Recognition. CVPR 2019. 43

References

[29] Wang et al. Dynamic Graph CNN for Learning on Point Clouds. TOG 2019.
[30] Te et al. RGCNN: Regularized Graph CNN for Point Cloud Segmentation. MM 2018.
[31] Valsesia et al. Learning Localized Generative Models for 3D Point Clouds via Graph Convolution. ICLR 2019.
[32] Wang et al. Local Spectral Graph Convolution for Point Set Feature Learning. ECCV 2018.
[33] Landrieu et al. Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs. CVPR 2018.
[34] Li et al. Gated graph sequence neural networks. ICLR 2016.
[35] Simonovsky et al. Dynamic edge-conditioned filters in convolutional neural networks on graphs. CVPR 2017.
[36] Landrieu et al. Point Cloud Oversegmentation with Graph-Structured Deep Metric Learning. CVPR 2019.
[37] Li et al. Can GCNs Go as Deep as CNNs. ICCV 2019.
[38] Shen et al. Mining Point Cloud Local Structures by Kernel Correlation and Graph Pooling. CVPR 2018.
[39] Thomas et al. KPConv: Flexible and Deformable Convolution for Point Clouds. arXiv 2019.
[40] Chen et al. ClusterNet: Deep Hierarchical Cluster Network with Rigorously Rotation-Invariant Representation for Point Cloud Analysis. CVPR 2019.
[41] Lan et al. Modeling Local Geometric Structure of 3D Point Clouds using Geo-CNN. CVPR 2019.
[42] Tatarchenko et al. Tangent Convolutions for Dense Prediction in 3D. CVPR 2018. 44

References

[43] Komarichev et al. A-CNN: Annularly Convolutional Neural Networks on Point Clouds. CVPR 2019.

[44] Zhao et al. PointWeb: Enhancing Local Neighborhood Features for Point Cloud Processing. CVPR 2019.

[45] Zaheer et al. Deep Sets. NeurIPS 2017.

45

University of Chinese Academy of Sciences

CASIA

Thanks for your attention !

yongcheng.liu@nlpr.ia.ac.cn

46

作者简介

刘永成

中科院自动化所在读博士，导师为**潘春洪研究员**。研究兴趣包括3D点云处理、图像分割、多标签图像分类等。

目前已在CVPR、ICCV、ACM MM等国际会议发表学术论文5篇，其中2篇入选Oral，1篇提名CVPR 2019 Best Paper。

E-mail: yongcheng.liu@nlpr.ia.ac.cn

个人主页: https://yochengliu.github.io/

作者发表论文

Yongcheng Liu, Bin Fan, Shiming Xiang, and Chunhong Pan. Relation-Shape Convolutional Neural Network for Point Cloud Analysis. CVPR (Oral & Best paper finalist), pp. 8895-8904, 2019.

点击下载pdf

Yongcheng Liu, Bin Fan, Gaofeng Meng, Jiwen Lu, Shiming Xiang, and Chunhong Pan. DensePoint: Learning Densely Contextual Representation for Efficient Point Cloud Processing. ICCV, pp. 5239-5248, 2019.

点击下载pdf

Yongcheng Liu, Lu Sheng, Jing Shao, Junjie Yan, Shiming Xiang, and Chunhong Pan. Multi-Label Image Classification via Knowledge Distillation from Weakly-Supervised Detection. ACM MM, pp. 700-708, 2018.

点击下载pdf

Yongcheng Liu, Bin Fan, Lingfeng Wang, Jun Bai, Shiming Xiang, Chunhong Pan. Semantic Labeling in Very High Resolution Images via A Self-Cascaded Convolutional Neural Network. ISPRS Journal of Photogrammetry and Remote Sensing, 145, pp. 78-95, 2018.

点击下载pdf

Jianbo Liu, Yongcheng Liu, Ying Wang, Véronique Prinet, Shiming Xiang, and Chunhong Pan. Decoupled Representation Learning for Skeleton-Based Gesture Recognition. CVPR (to appear), pp. 1-10, 2020.

学报相关论文

苏本跃，储小玉，盛敏.2019. 同态滤波下RGB-D点云配准方法. 中国图象图形学报，24(11): 1985-1997

http://www.cjig.cn/html/jig/2019/11/20191113.htm

文威威，文贡坚，回丙伟，陈鼎新.2019. 结合全局与局部信息的点云目标识别模型库构建. 中国图象图形学报，24(2): 248-257

http://www.cjig.cn/html/jig/2019/2/20190209.htm

孙士杰，宋焕生，张朝阳，张文涛，王璇. 2018.点云下地平面检测的RGB-D相机外参自动标定. 中国图象图形学报，23(6): 866-873

http://www.cjig.cn/html/jig/2018/6/20180609.htm

吴霄，陈聪梅，王加俊.2018. 使用霍夫变换的3维点云拼合算法. 中国图象图形学报，23(4): 478-489

http://www.cjig.cn/html/jig/2018/4/20180402.htm

编辑：秀　秀　指导：梧桐君

审校：夏薇薇　总编辑：肖　亮

同时也可申请加入CVer大群和细分方向技术群，细分方向已涵盖：目标检测、图像分割、目标跟踪、人脸检测&识别、OCR、姿态估计、超分辨率、SLAM、医疗影像、Re-ID、GAN、NAS、深度估计、自动驾驶、强化学习、车道线检测、模型剪枝&压缩、去噪、去雾、去雨、风格迁移、遥感图像、行为识别、视频理解、图像融合、图像检索、论文投稿&交流、PyTorch和TensorFlow等群。

一定要备注：研究方向+地点+学校/公司+昵称（如论文写作+上海+上交+卡卡），根据格式备注，可更快被通过且邀请进群

▲长按加微信群

▲长按关注CVer公众号

麻烦给我一个在看！