藤蔓技术论坛

面向三维点云处理的深度学习研究进展

https://yochengliu.github.io/
http://vslab.ia.ac.cn/

University of Chinese Academy of Sciences

刘永成　助理研究员

中国科学院自动化研究所, 模式识别国家重点实验室

2021.06.17

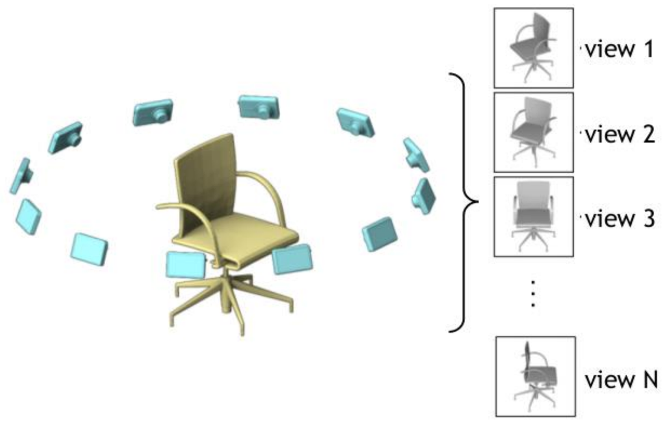
提纲

1 **背景简介**

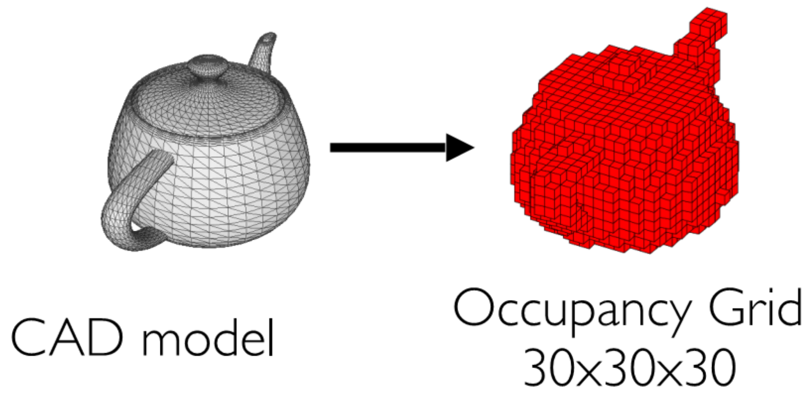
2 研究综述

3 工作介绍

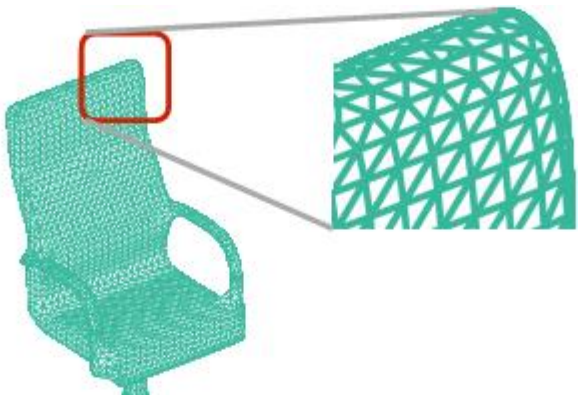
背景简介 3D数据表示



多视角图像 + 2D CNN



体素数据 + 3D CNN

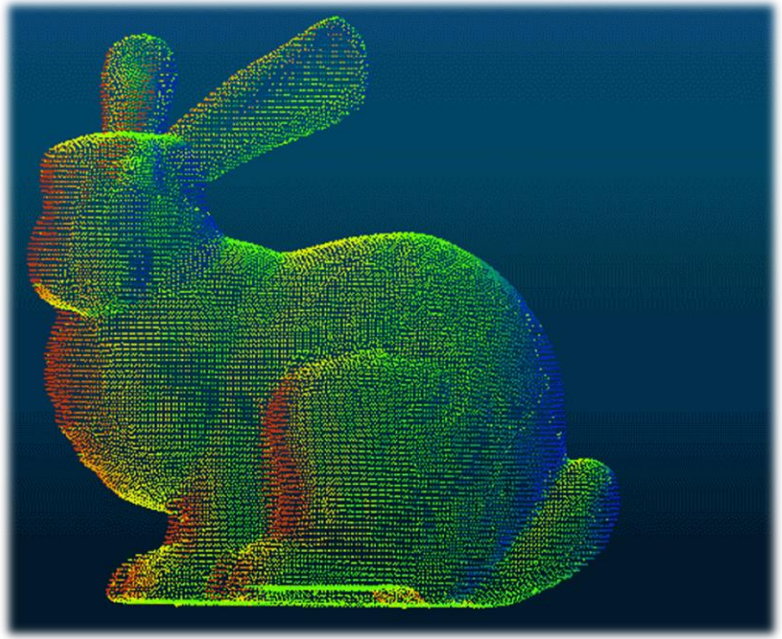


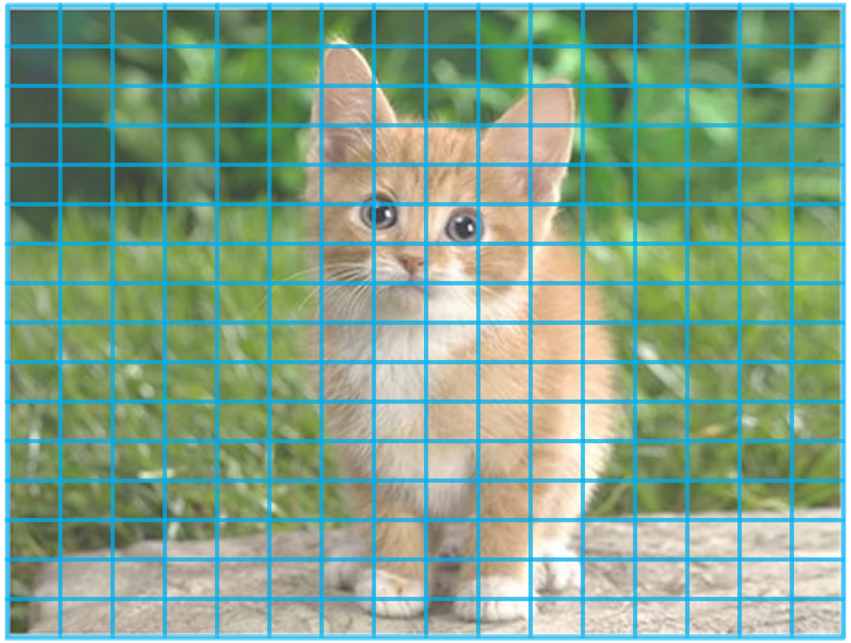
网眼（Mesh） 数据 + GNN

深度图 + 2D CNN

3D点云 + Deep Learning

背景简介 3D点云 vs. 2D图像





□ 3D点云

✓ (x, y, z)，不规则分布，无序，稀疏，不均匀

✓ 物体3D形状结构信息

□ 2D图像

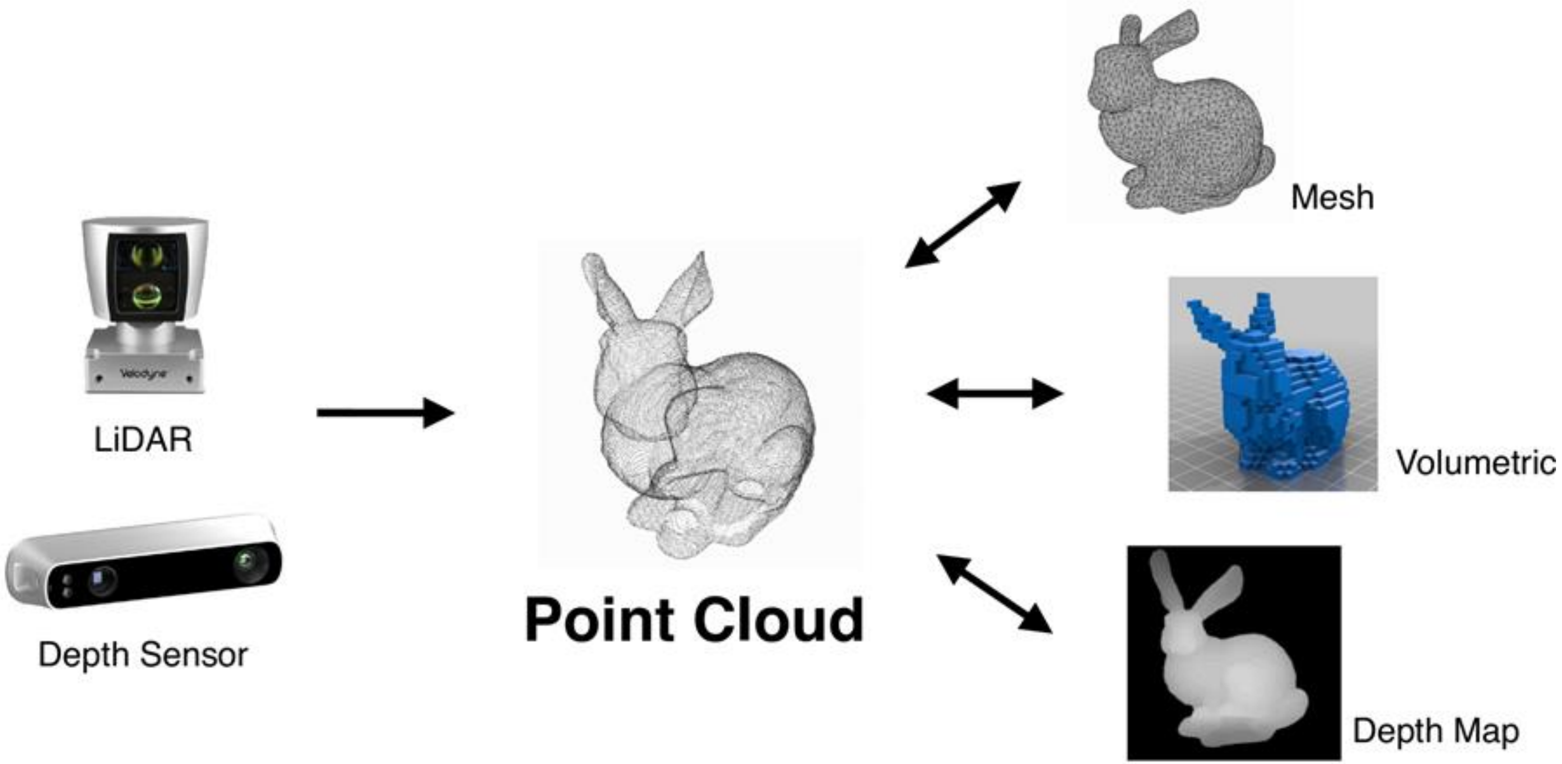
✓ (r, g, b)，规则网格分布，有序，密集，均匀

✓ 物体表观纹理信息

背景简介 3D点云优势

□ 3D点云 - 数据本身优势 vs. 其他3D表示

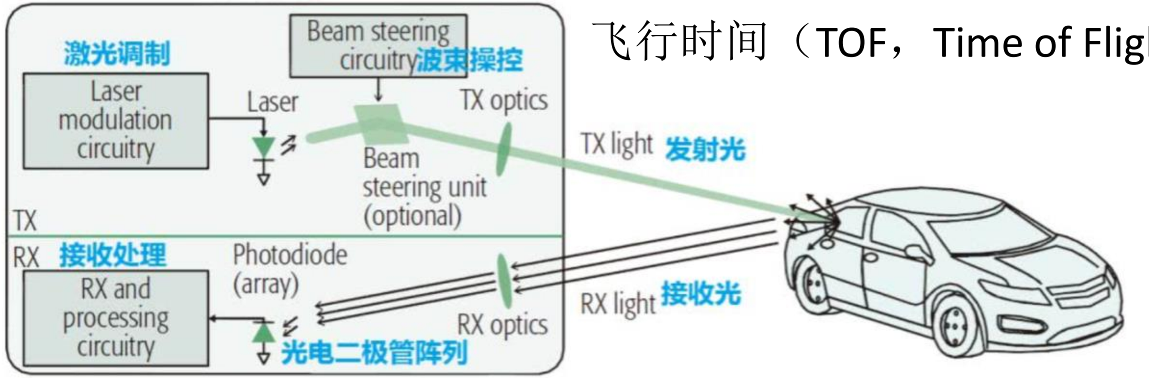
- ✓ 传感器获取的原始数据（raw data）
- ✓ 数据表示与存储简单：N * (x, y, z, r, g, b, intensity)
- ✓ 良好的3D形状表征特性，视角图像（自遮挡）、3D体素（细节缺失）、深度图（2. 5D）、Mesh(重建)



背景简介 3D点云优势

□ 3D点云 – 传感器采集优势 (LiDAR) vs. 自然图像

- ✓ 雨、雾、雪，强光、逆光等
- ✓ 全天候采集3D信息 vs. 红外2D信息
- ✓ 视野范围较大（360度环视）

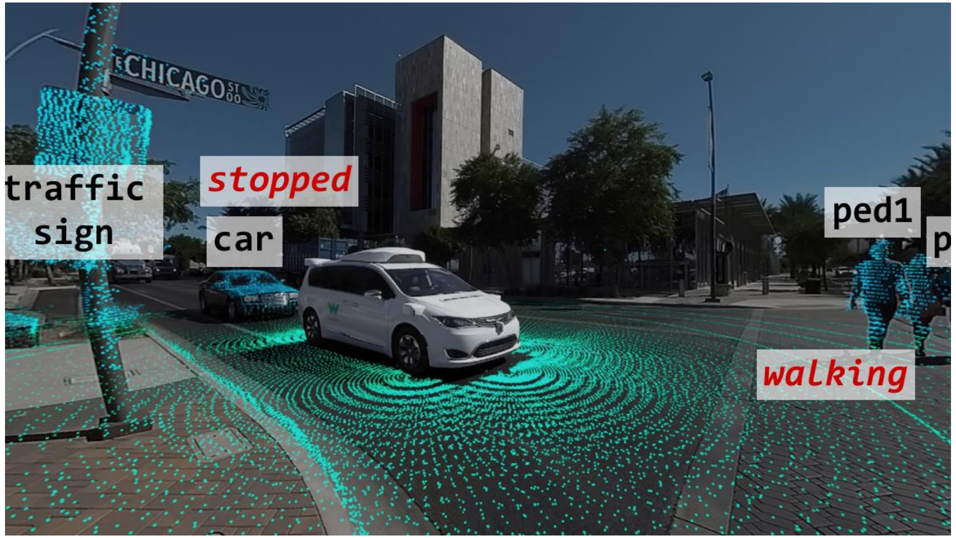


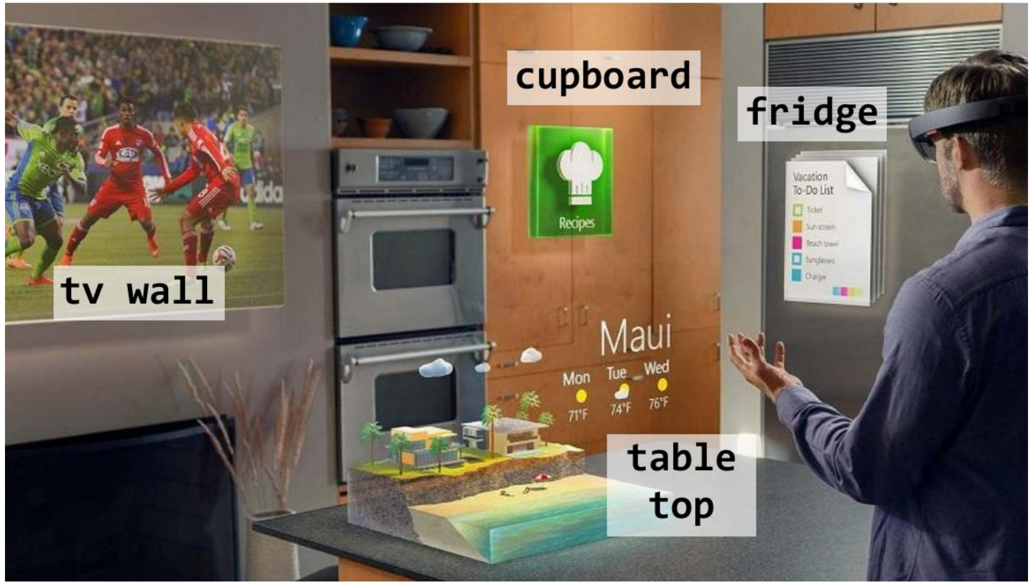
飞行时间（TOF，Time of Flight）

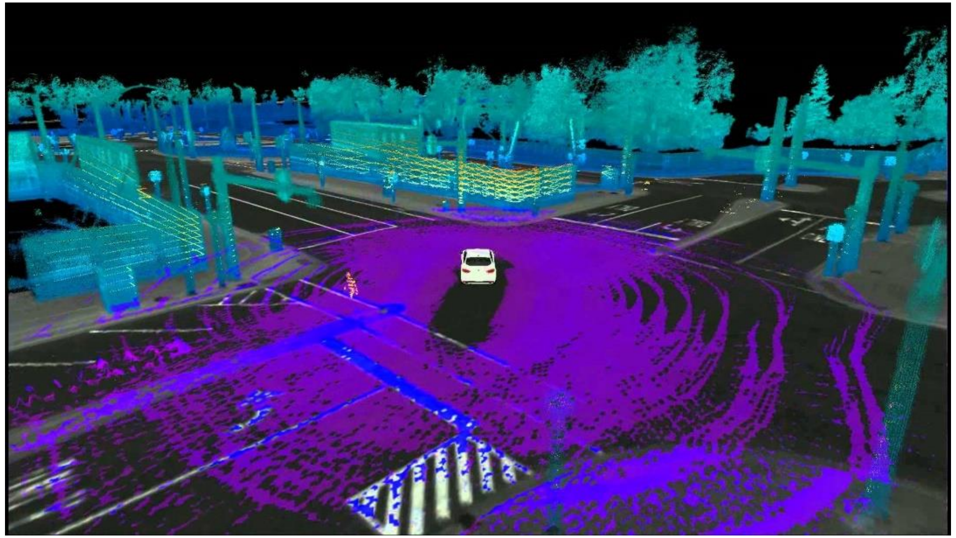
背景简介　3D点云应用场景

☐ 学术界和工业界广泛关注?

- ✓ 自动驾驶
- ✓ 高精地图
- ✓ AR & VR

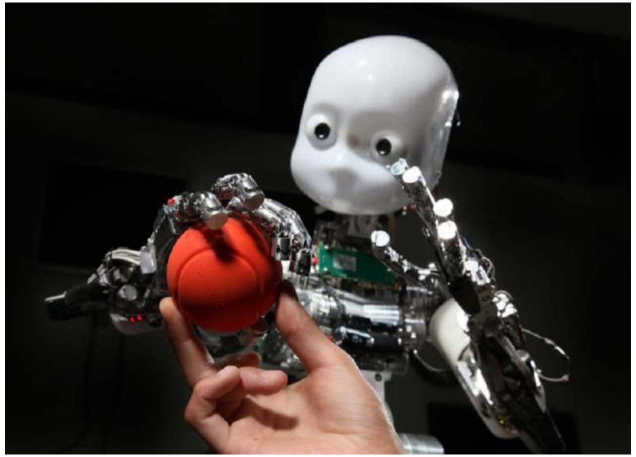


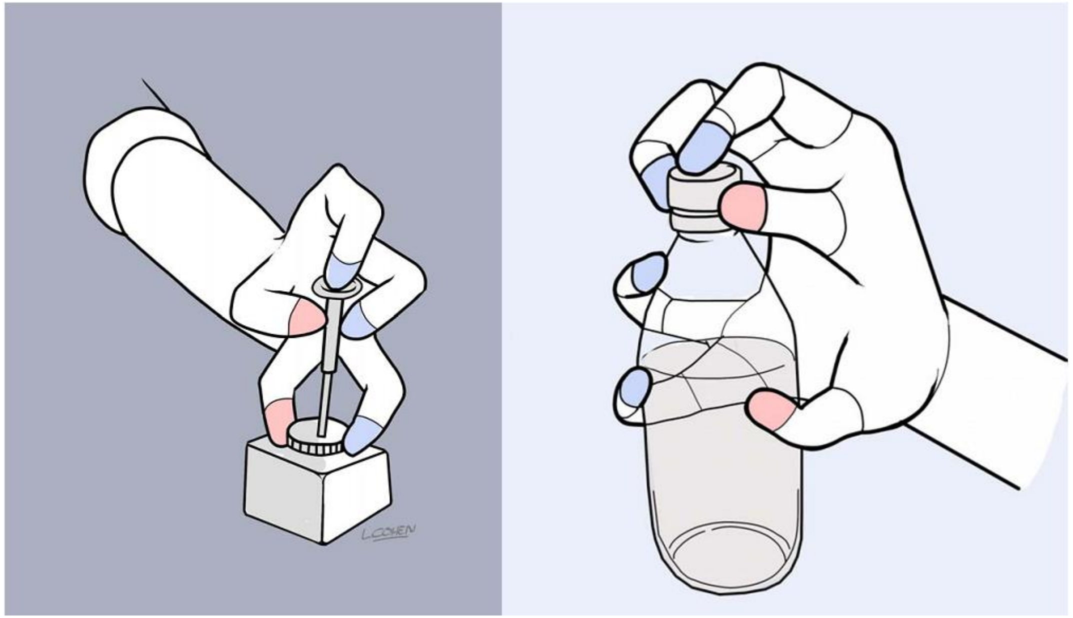


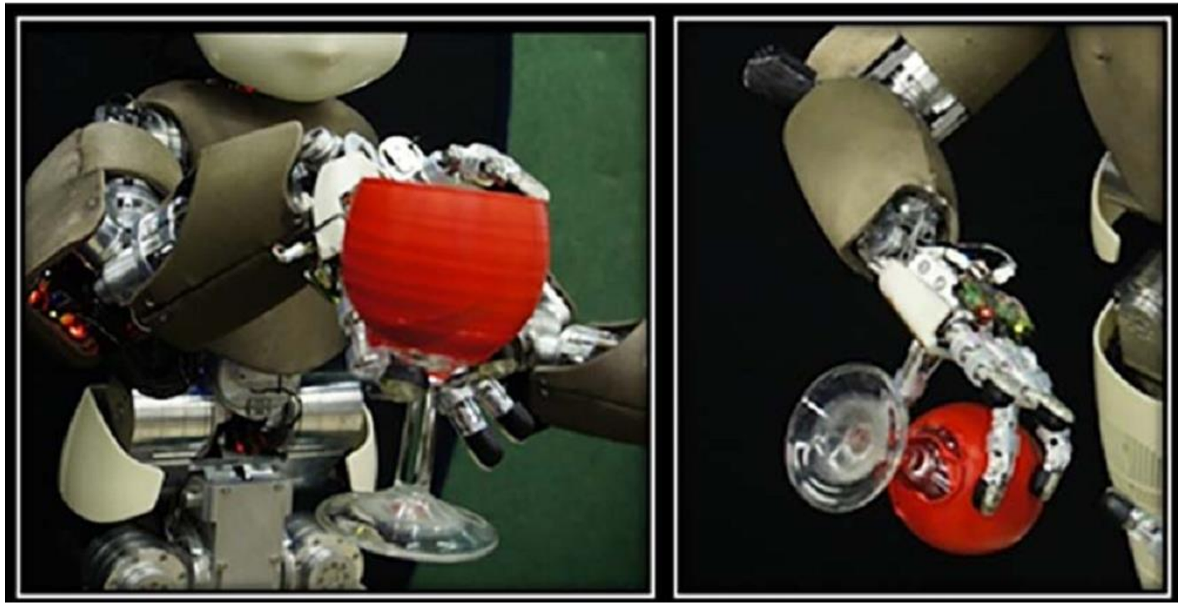


背景简介 3D点云应用场景

- ☐ **学术界和工业界广泛关注?**
- ✓ 自动驾驶、高精地图、AR & VR
- ✓ 机器人操作 (robot manipulation)：抓取



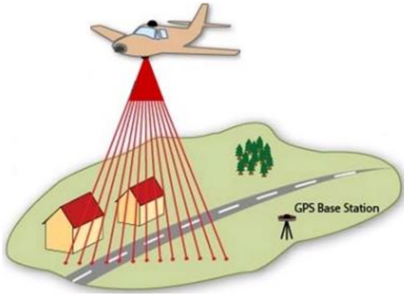


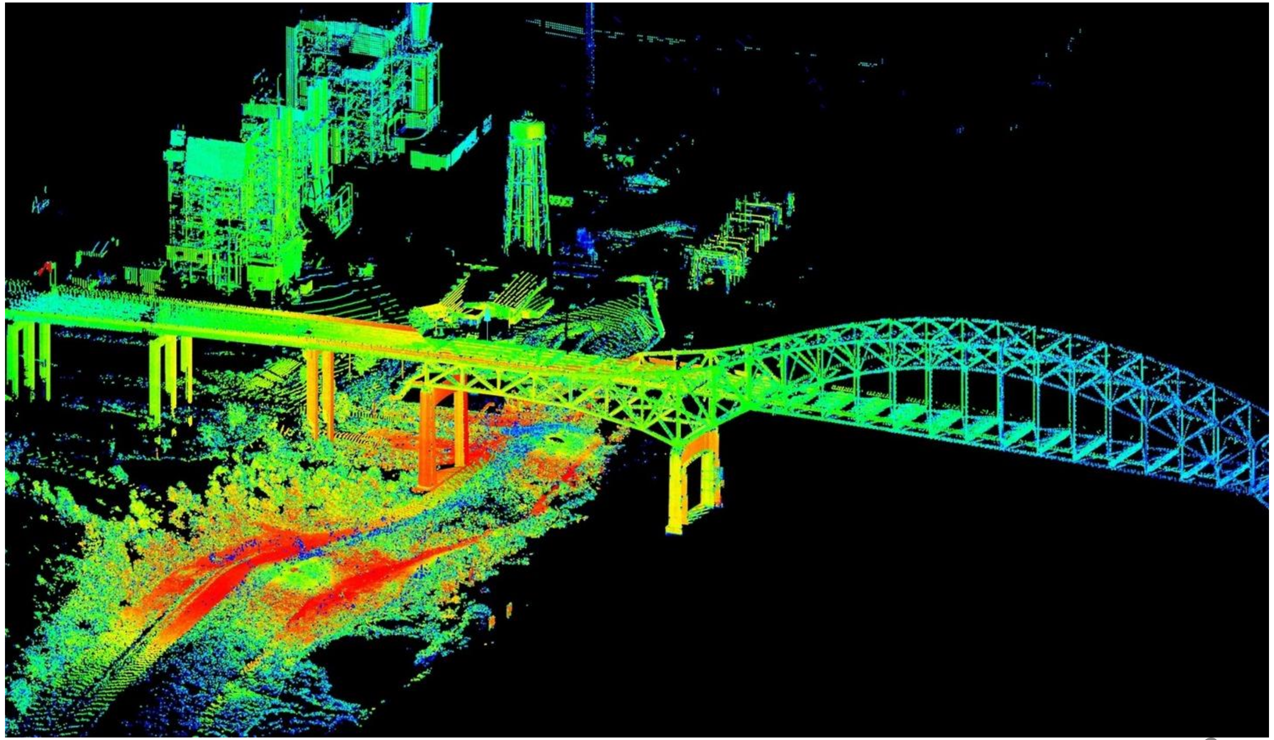


背景简介　3D点云应用场景

☐ 学术界和工业界广泛关注?

- ✓ 自动驾驶、高精地图、AR & VR
- ✓ 机器人操作：抓取
- ✓ 测绘地理、遥感 (数字城市)



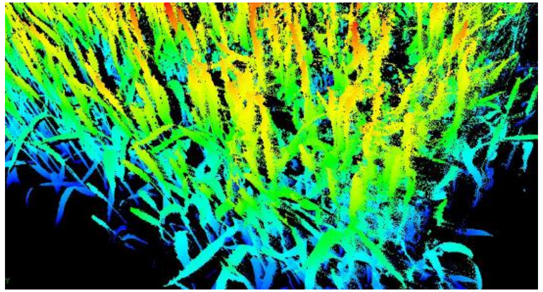


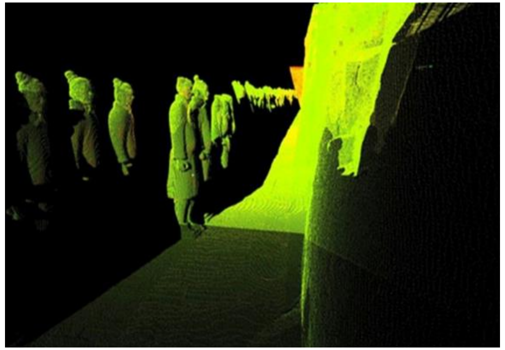
背景简介 3D点云应用场景

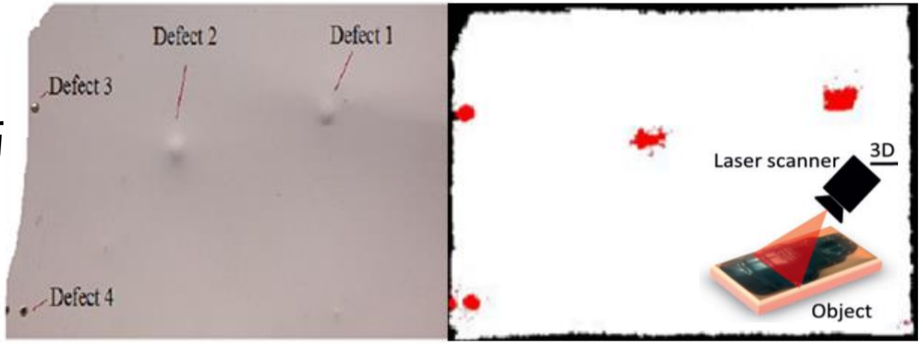
□ 学术界和工业界广泛关注?

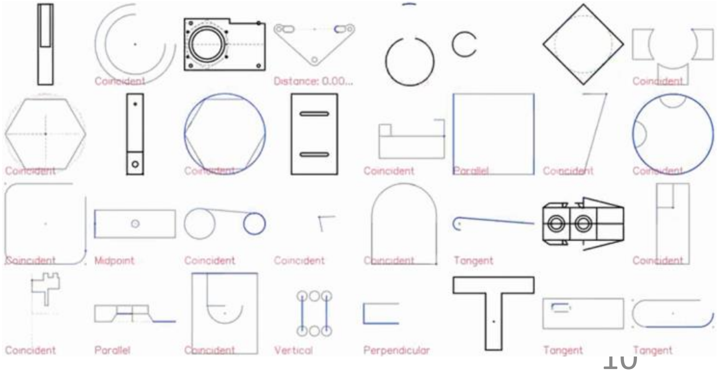
- ✓ 自动驾驶、高精地图、AR&VR
- ✓ 机器人操作：抓取
- ✓ 测绘地理、遥感 (数字城市)
- ✓ 农业、林业
- ✓ 考古与文物保护
- ✓ 3D人脸识别
- ✓ 医疗：3D精准手术
- ✓ 3D游戏、3D动漫
- ✓ 缺陷检测、工业探伤
- ✓ 计算机辅助设计

……





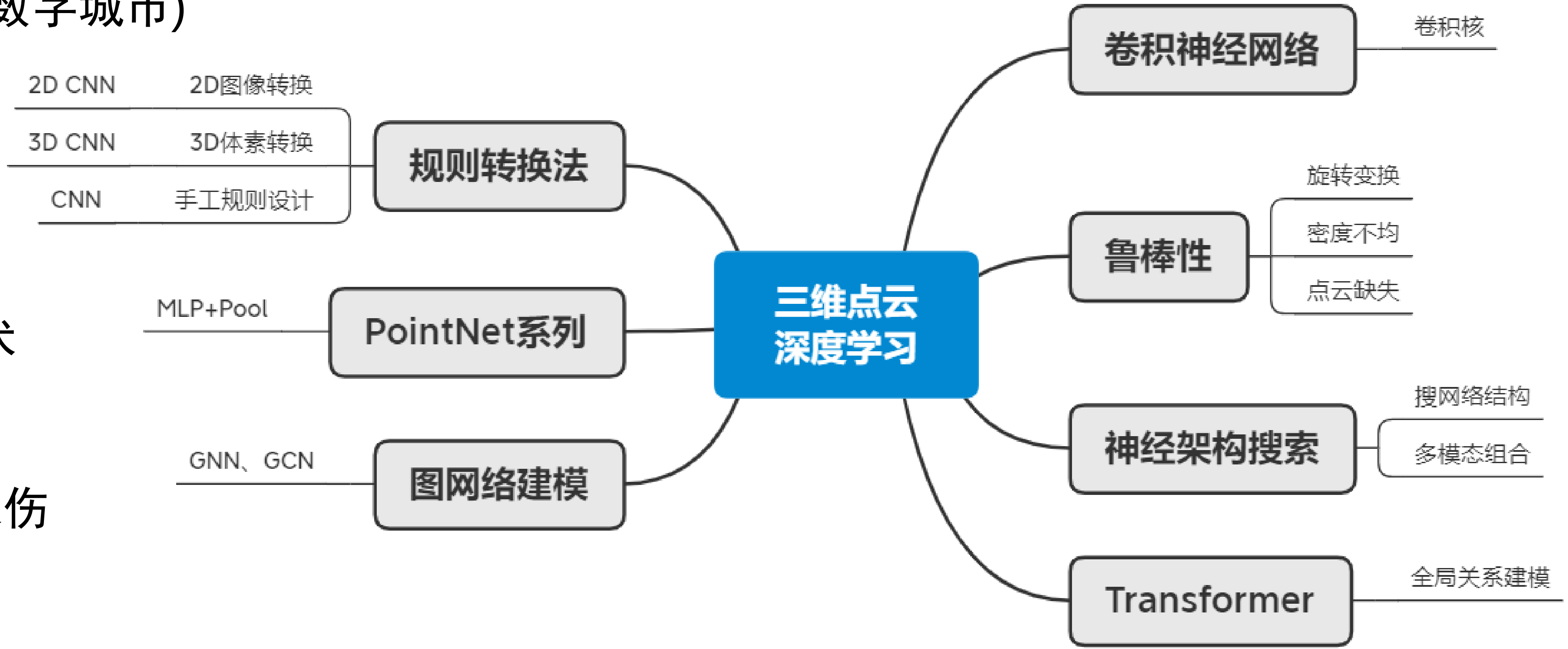




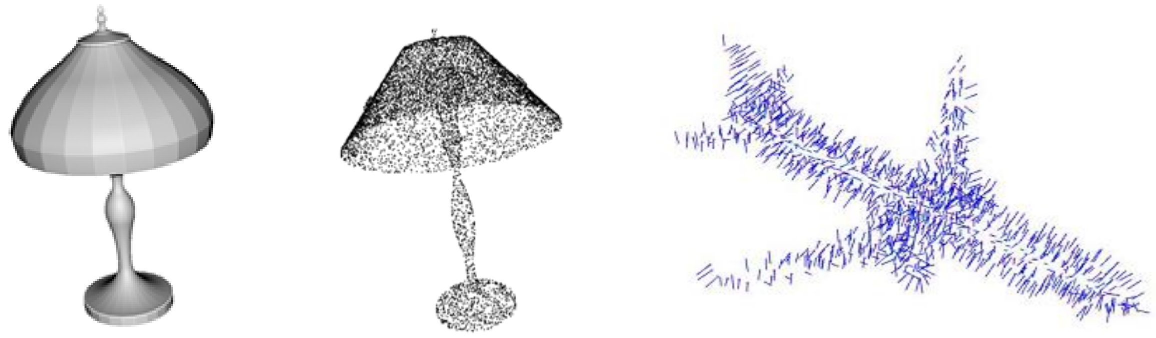
背景简介 3D点云应用场景

□ 学术界和工业界广泛关注?

- ✓ 自动驾驶、高精地图、AR&VR
- ✓ 机器人操作：抓取
- ✓ 测绘地理、遥感 (数字城市)
- ✓ 农业、林业
- ✓ 考古与文物保护
- ✓ 3D人脸识别
- ✓ 医疗：3D精准手术
- ✓ 3D游戏、3D动漫
- ✓ 缺陷检测、工业探伤
- ✓ 计算机辅助设计
- ● 学术层面：开放性研究问题

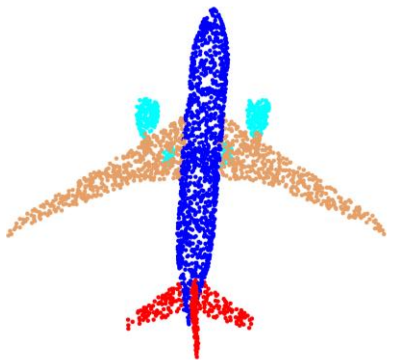


背景简介　学术研究数据集



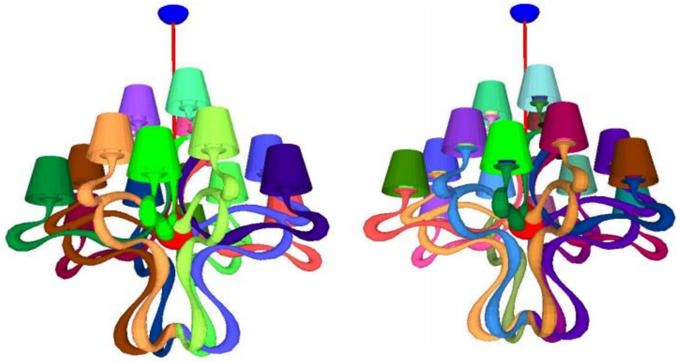
Princeton ModelNet: 1k

[1] Wu et al. CVPR 2015.

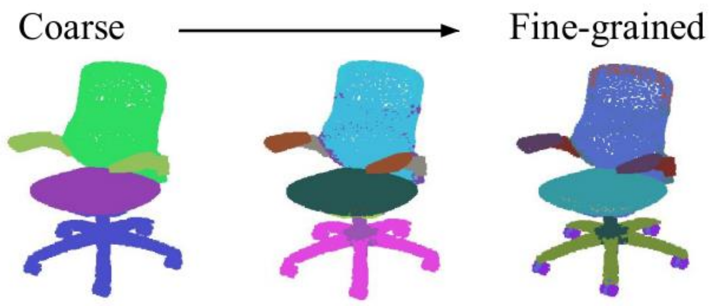


ShapeNet Part: 2k

[2] Yi et al. TOG 2016.



PartNet models

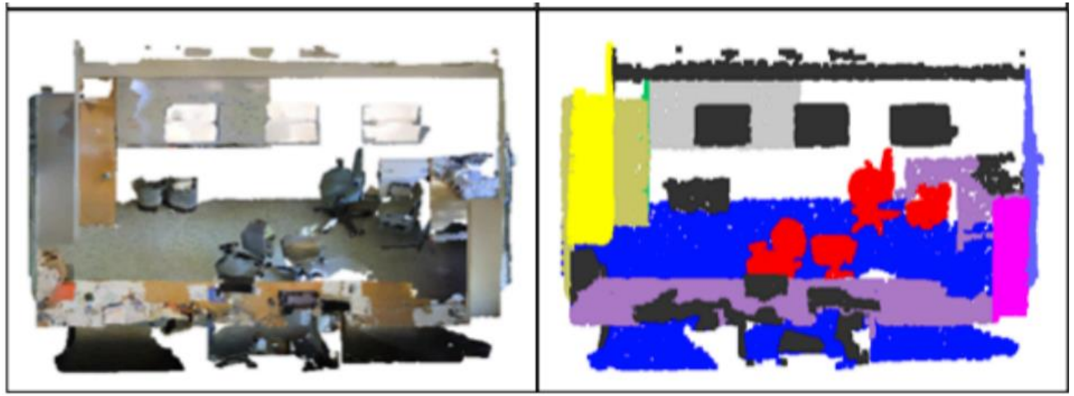


Hierarchical Semantic
Segmentation

……

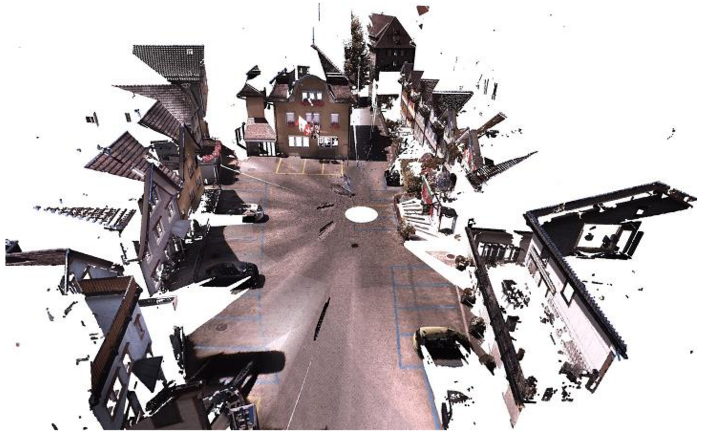
[3] Mo et al. CVPR 2019.

背景简介　学术研究数据集



Stanford 3D indoor scene: 8k

[4] Armeni et al. CVPR 2016.



Semantic 3D: 4 billion in total

[5] Hackel et al. ISPRS 2017.

ScanNet: 语义分割 + 目标检测

[6] Dai et al. CVPR 2017.

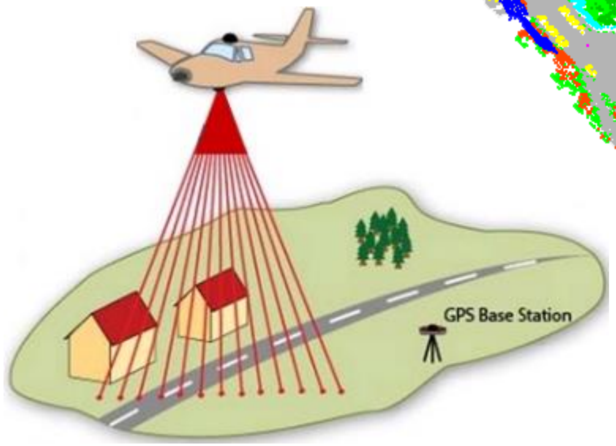
KITTI, Apollo, nuScenes, Waymo

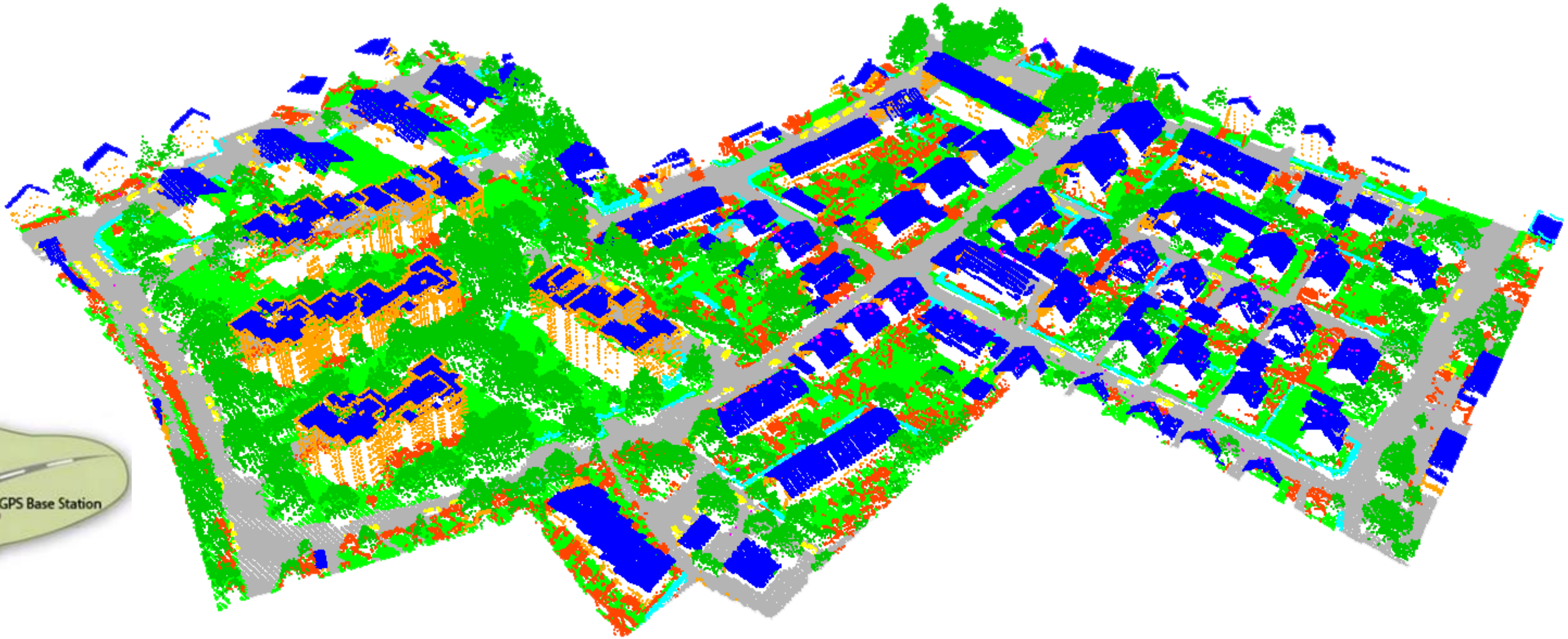
语义分割 + 目标检测

……

背景简介　学术研究数据集

- 机载激光雷达点云 **(Airborne Laser Scanning, ALS)**

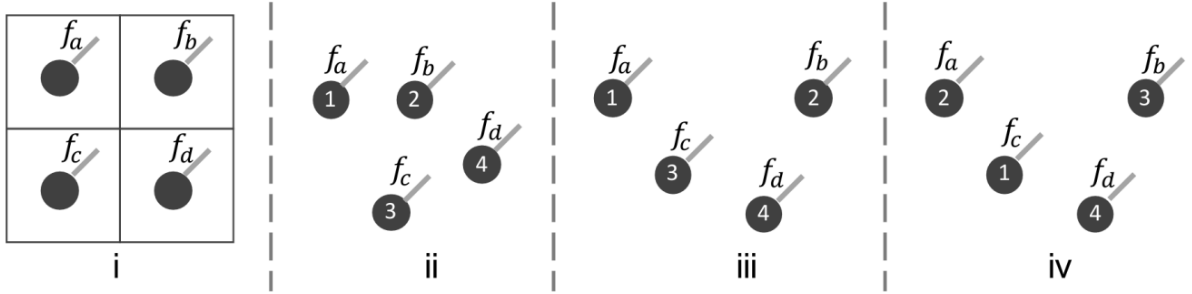




电线　矮植　树木　道路　汽车　栅栏　建筑　墙面　灌木

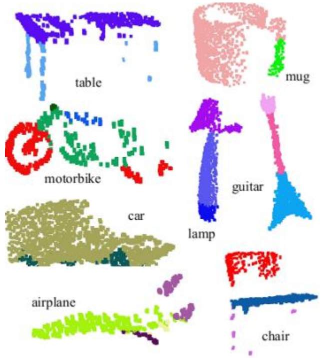
背景简介 点云深度学习：问题与挑战

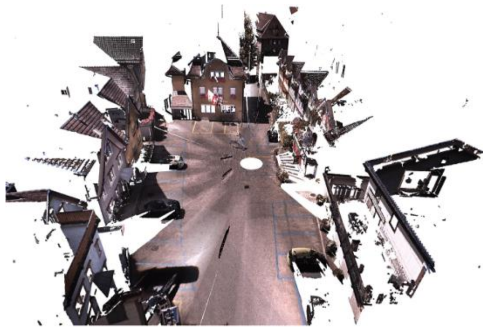
不规则、无序：结构感知 & 置换排列不变性



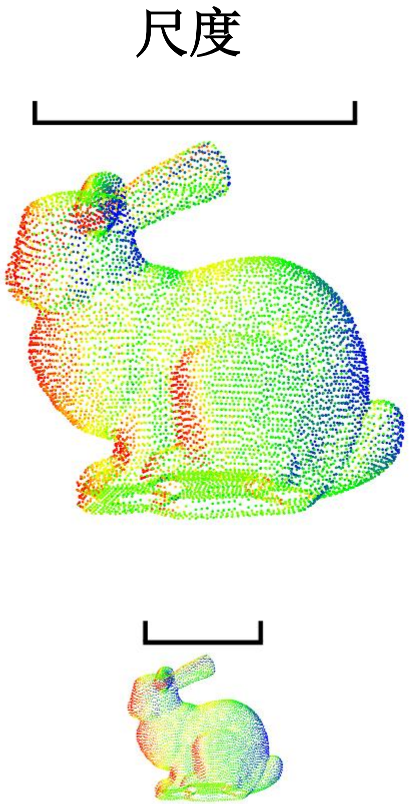
点云缺失、大量噪声、分布稀疏且不均匀：鲁棒性；

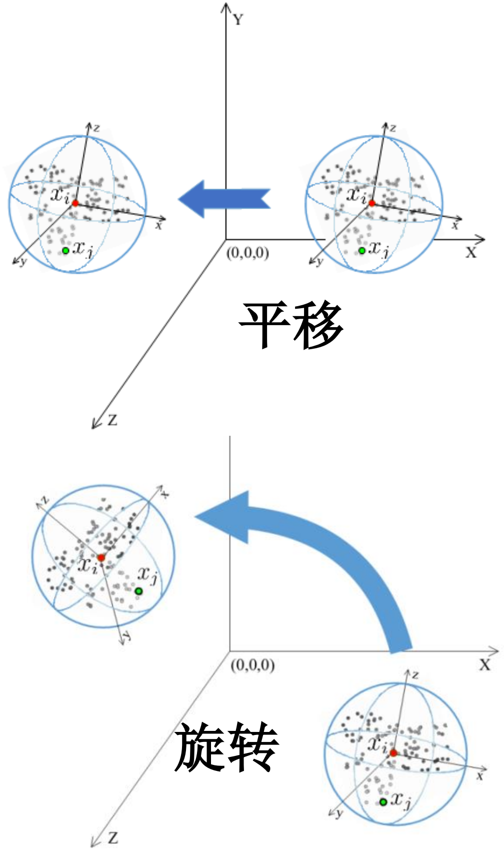
大尺度场景点云数据：如何高效处理？





几何变换：鲁棒性

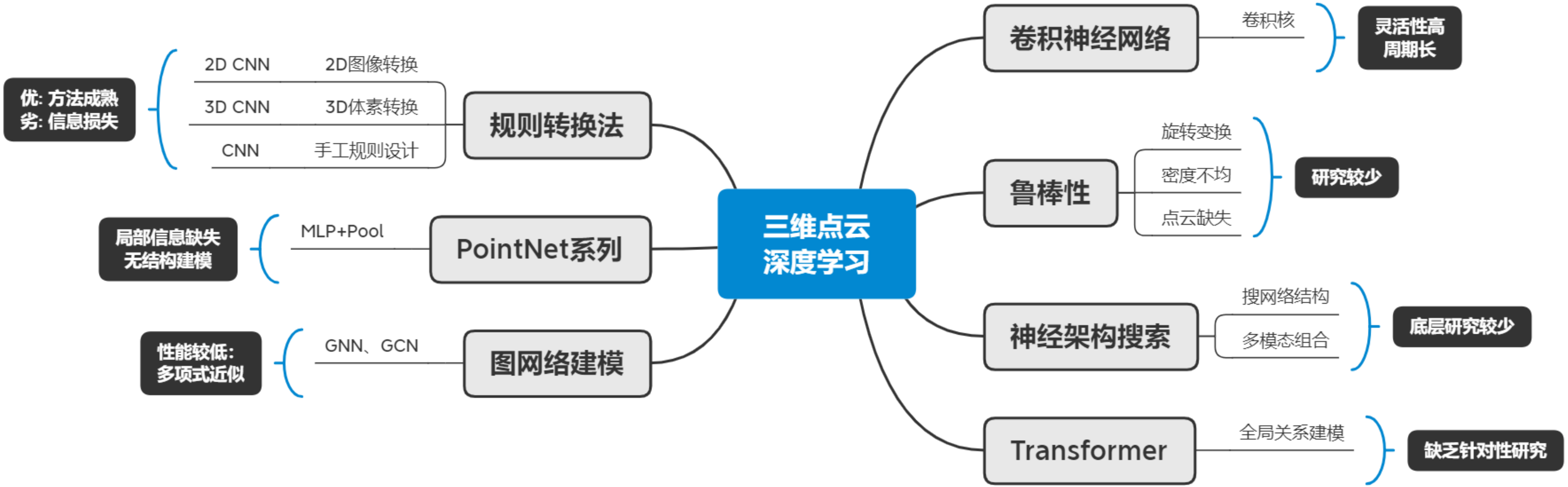




提纲

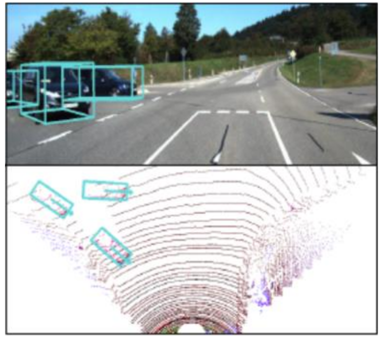
1. 背景简介
2. 研究综述
3. 工作介绍

研究综述



三维点云深度学习
规则转换法
2D CNN
2D图像转换
3D CNN
3D体素转换
CNN
手工规则设计
优: 方法成熟
劣: 信息损失
PointNet系列
MLP+Pool
局部信息缺失
无结构建模
图网络建模
GNN、GCN
性能较低:
多项式近似
卷积神经网络
卷积核
灵活性高
周期长
鲁棒性
旋转变换
密度不均
点云缺失
研究较少
神经架构搜索
搜网络结构
多模态组合
底层研究较少
Transformer
全局关系建模
缺乏针对性研究

研究综述　点云规则化处理



2DCNN

多视角图像

鸟瞰图 (bird view)

切平面图

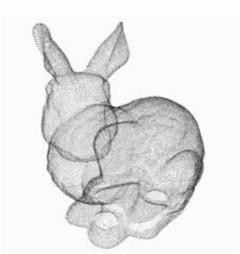
2D图像转换

规则转换法

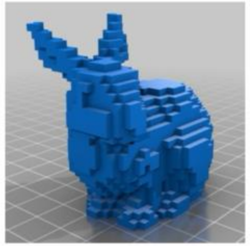
高维规则空间

3D体素转换

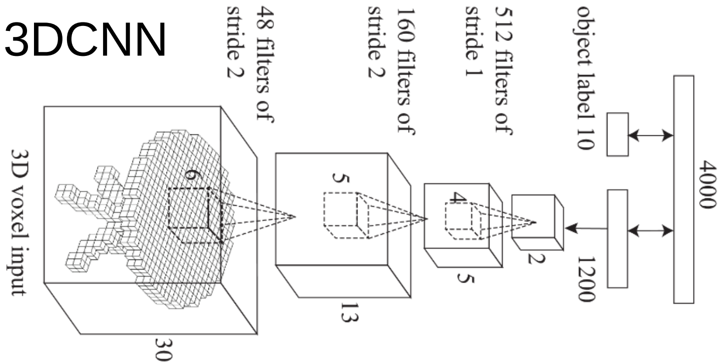
手工规则设计



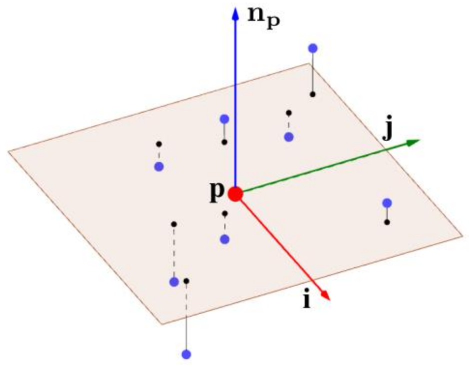
voxelization



3DCNN



研究综述 点云规则化处理 – 切平面图



$\|\mathbf{p}-\mathbf{q}\| < R \quad \mathbf{r} = \mathbf{q} - \mathbf{p}$

$\mathbf{C} = \sum_{\mathbf{q}} \mathbf{r}\mathbf{r}^{\top}$ 正切图像S

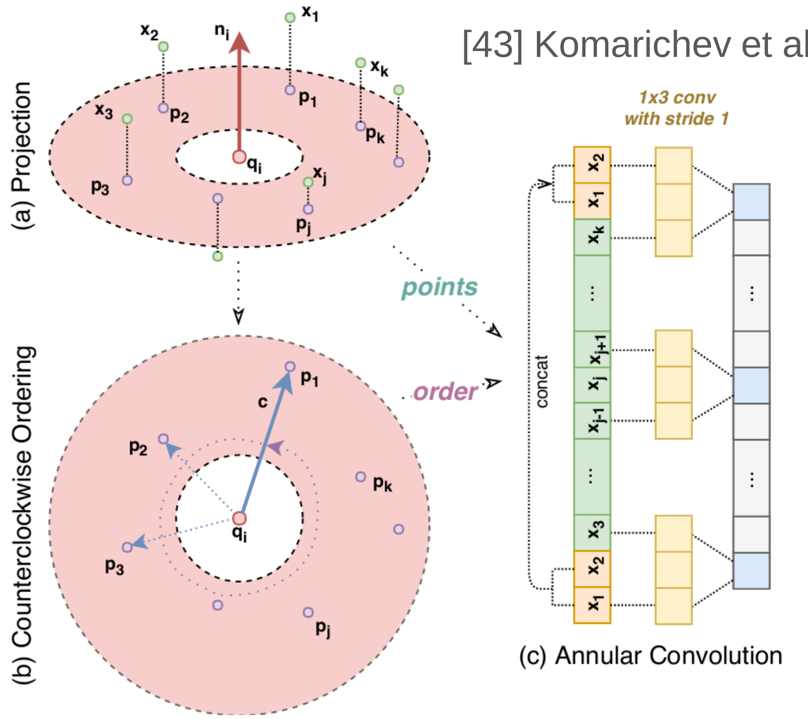
[42] Tatarchenko et al. Tangent Conv. CVPR 2018.

$$S(\mathbf{v}) = F(\mathbf{q}) \qquad S(\mathbf{u}) = \sum_{\mathbf{v}} \left(w(\mathbf{u},\mathbf{v}) \cdot S(\mathbf{v})\right)$$

$$X(\mathbf{p}) = \int_{\pi_{\mathbf{p}}} c(\mathbf{u}) S(\mathbf{u})\, d\mathbf{u}$$

计算切平面(稀疏) → 图像插值 → 完整图像 → 图像卷积

[43] Komarichev et al. A-CNN. CVPR 2019.

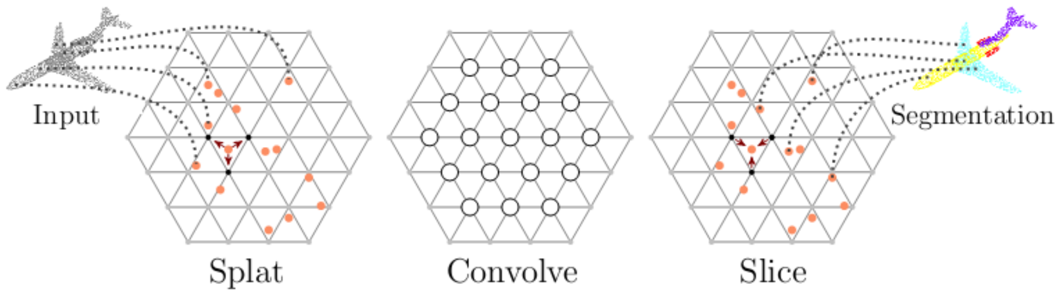


(c) Annular Convolution

协方差计算复杂度较高

3D形状信息存在缺失

多视角(12 views)：复杂度问题
间接处理方法



[8] Su et al. SPLATNet. CVPR 2018.

高维晶格 + 双边卷积 + 哈希表提速

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

[10] Jampani et al. Bilateral Neural Networks. CVPR 2016.

间接处理方法

[12] Li et al. PointCNN. NIPS 2018.

“simultaneously weight and permute the input features”

K*K 咖方变换矩阵：加权 + 置换

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

$\mathbf{P}' \leftarrow \mathbf{P} - p$ ▷ Move $\mathbf{P}$ to local coordinate system of p

$\mathbf{F}_\delta \leftarrow MLP_\delta(\mathbf{P}')$ ▷ **Individually** lift each point into C_δ dimensional space

$\mathbf{F}_* \leftarrow [\mathbf{F}_\delta, \mathbf{F}]$ ▷ Concatenate $\mathbf{F}_\delta$ and $\mathbf{F}$, $\mathbf{F}_*$ is a $K \times (C_\delta + C_1)$ matrix

$\mathcal{X} \leftarrow MLP(\mathbf{P}')$ ▷ Learn the $K \times K$ $\mathcal{X}$-transformation matrix

$\mathbf{F}_\mathcal{X} \leftarrow \mathcal{X} \times \mathbf{F}_*$ ▷ Weight and permute $\mathbf{F}_*$ with the learnt $\mathcal{X}$

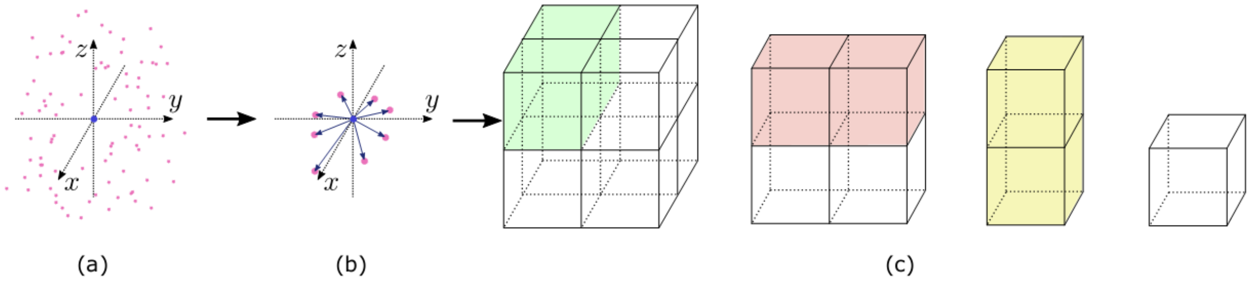
$\mathbf{F}_p \leftarrow \mathrm{Conv}(\mathbf{K}, \mathbf{F}_\mathcal{X})$ ▷ Finally, typical convolution between $\mathbf{K}$ and $\mathbf{F}_\mathcal{X}$

研究综述 点云规则化处理 – 手工规则设计

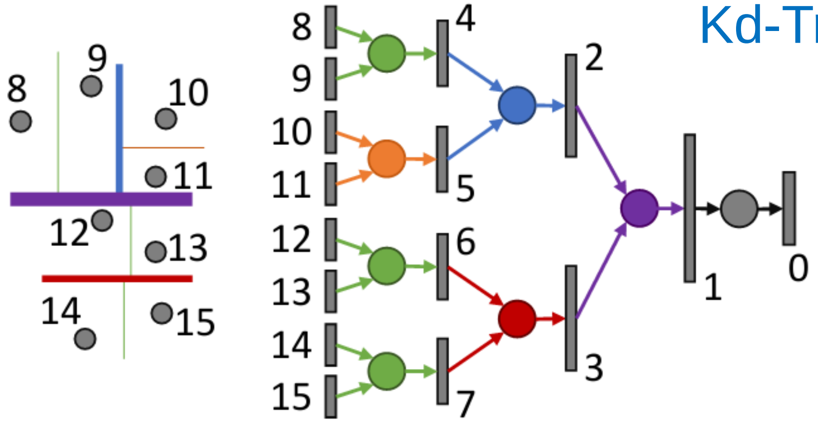
[13] Jiang et al. PointSIFT. arXiv 2018.

按方向编码

规则普适性不强

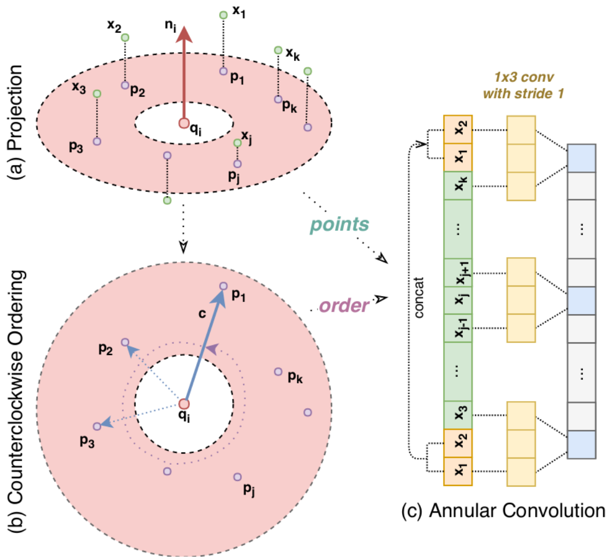


[43] Komarichev et al. A-CNN. CVPR 2019.



Kd-Tree

[16] Klokov et al. Kd-Net. ICCV 2017.



指定排列方向

顺/逆时针

PointNet 系列

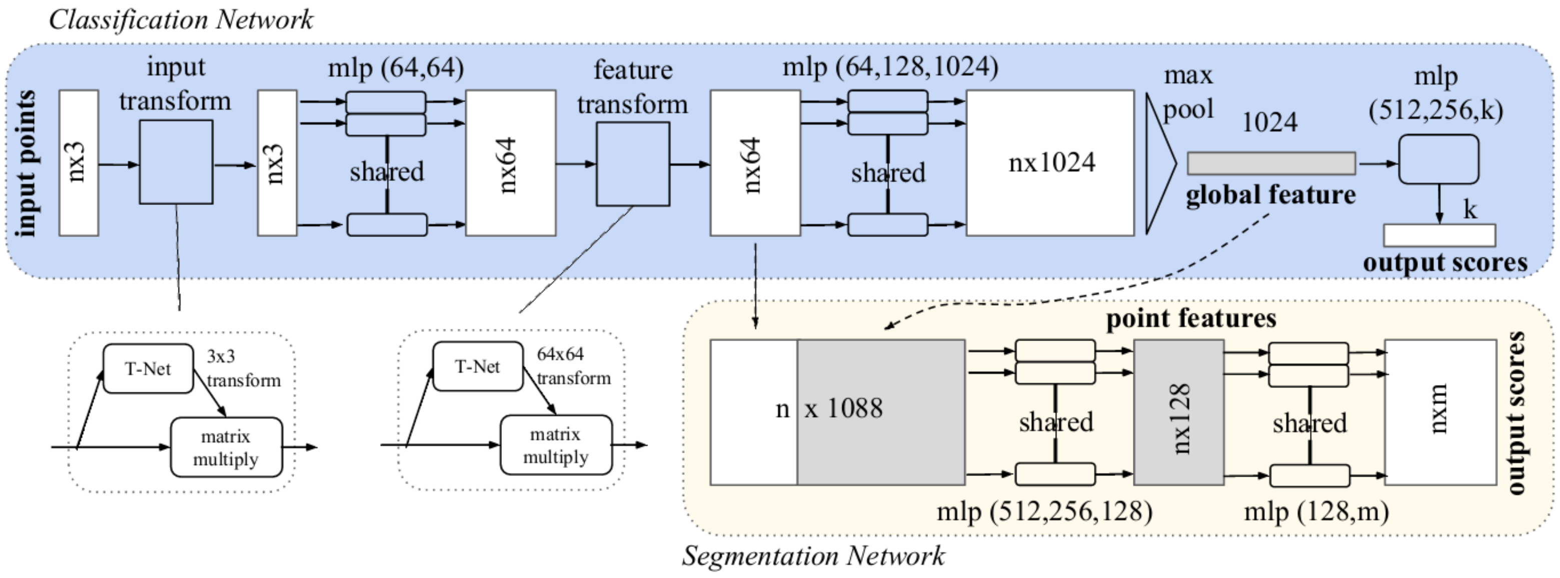
Leonidas J. Guibas

Paul Pigott Professor of Computer Science and
Electrical Engineering (courtesy)
in the School of Engineering

Research Statement

Professor Guibas heads the Geometric Computation group in the Computer Science Department of Stanford University. He is acting director of the Artificial Intelligence Laboratory and member of the Computer Graphics Laboratory, the Institute for Computational and Mathematical Engineering (iCME) and the Bio-X program. His research centers on algorithms for sensing, modeling, reasoning, rendering, and acting on the physical world. Professor Guibas' interests span computational geometry, geometric modeling, computer graphics, computer vision, sensor networks, robotics, and discrete algorithms --- all areas in which he has published and lectured extensively. Examples of current and recent activities include:

研究综述 PointNet系列 – PointNet



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

Shared MLP + max pool

(对称函数)

未建模局部结构

泛化函数：置换排列不变性 & 有效性

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

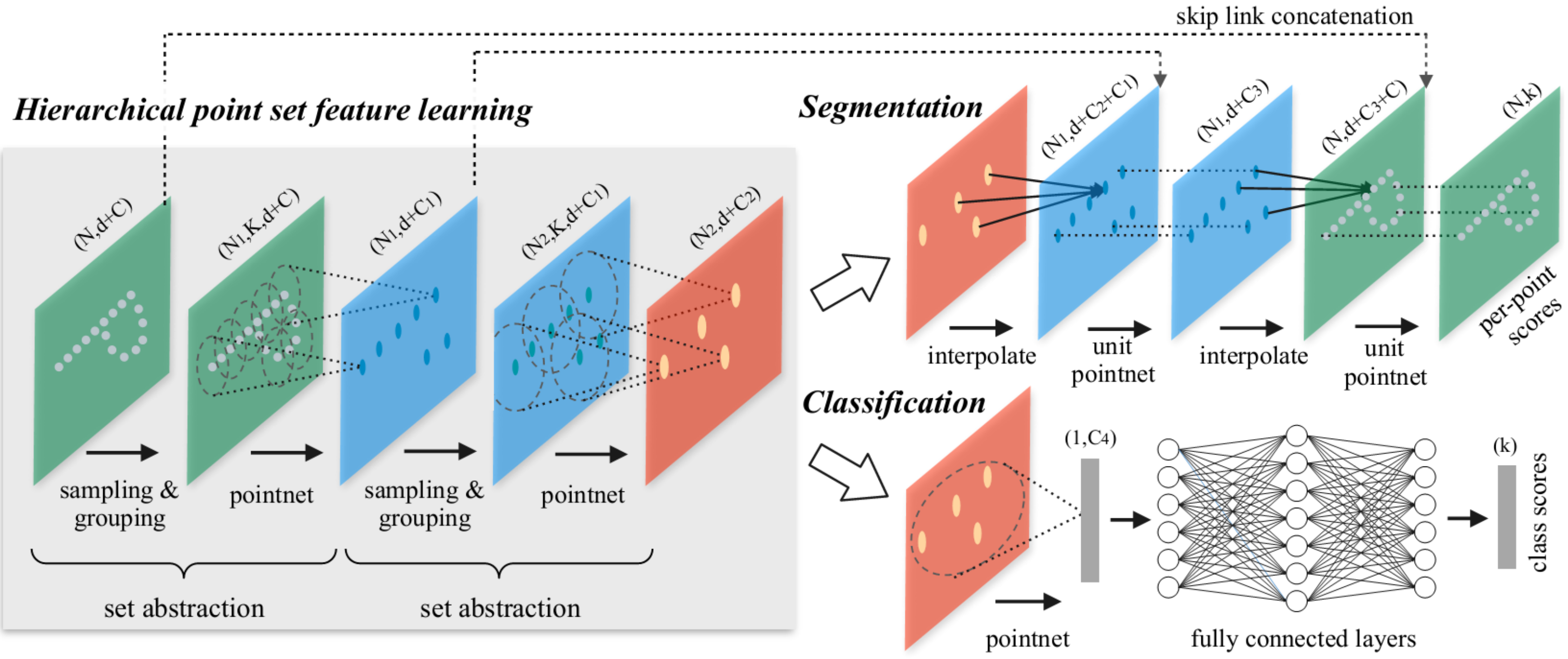
$$f(\{x_1,\ldots,x_n\}) \approx g(h(x_1),\ldots,h(x_n))$$

h(.): 共享的特征变换函数 *g(.)*: 对称的特征聚集函数，比如求和、max pool

MLP Max pool + fc layers

T-Net：学习3x3矩阵变换，将不同位姿的输入点云对齐

研究综述　PointNet系列 – PointNet++



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

Sampling + Grouping + PointNet

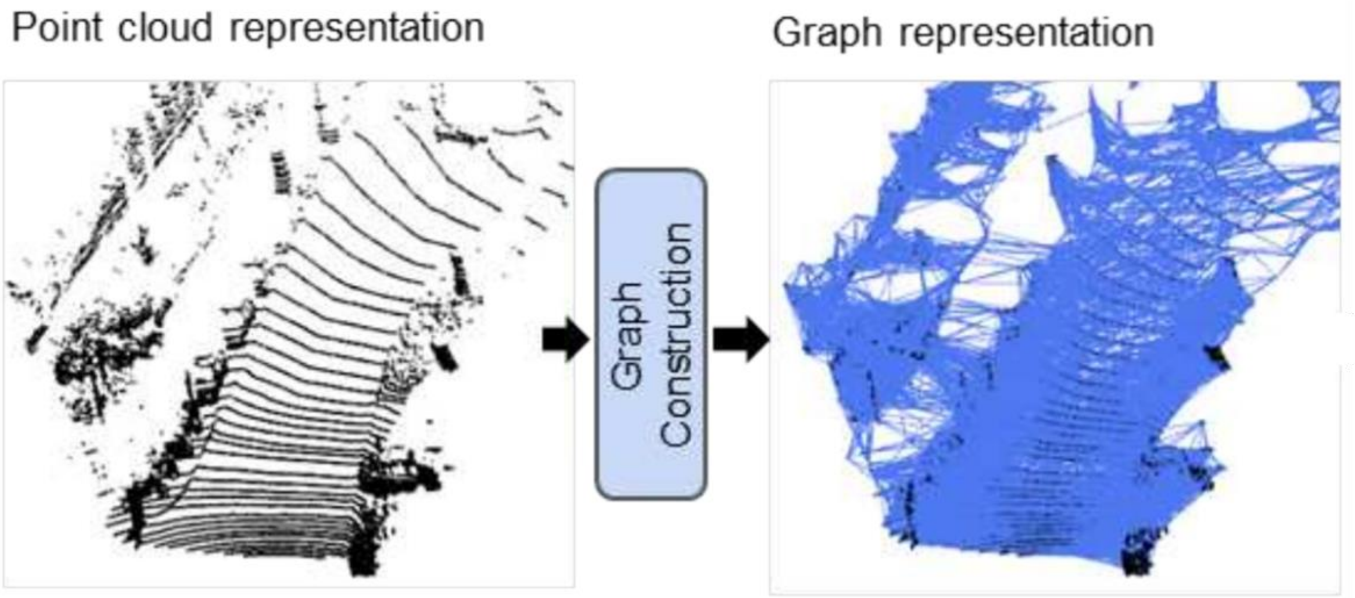
- 局部信息建模
- 类似CNN架构，由局部到全局地学习形状上下文

语义分割：解码上采样

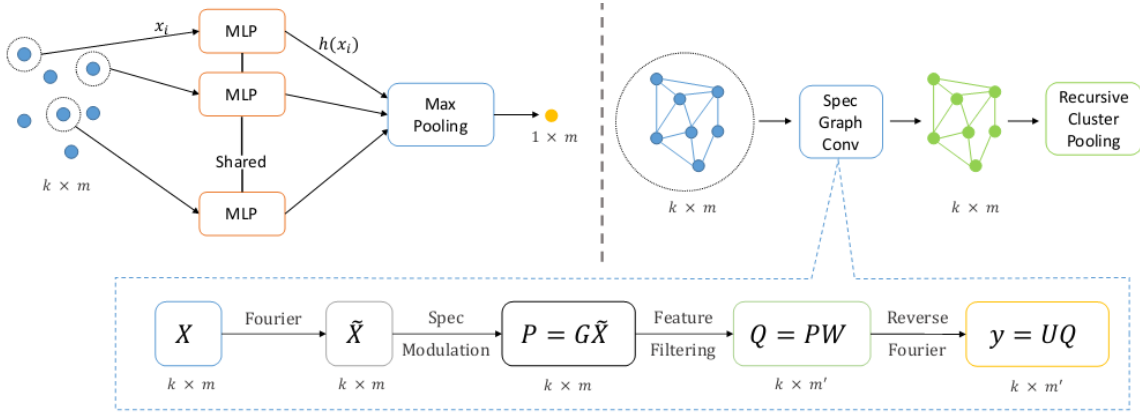
简单有效，广泛应用于检测、分割等后端任务上

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

图网络建模


Point cloud representation
Graph Construction
Graph representation

研究综述　图网络建模：谱域 & 空域



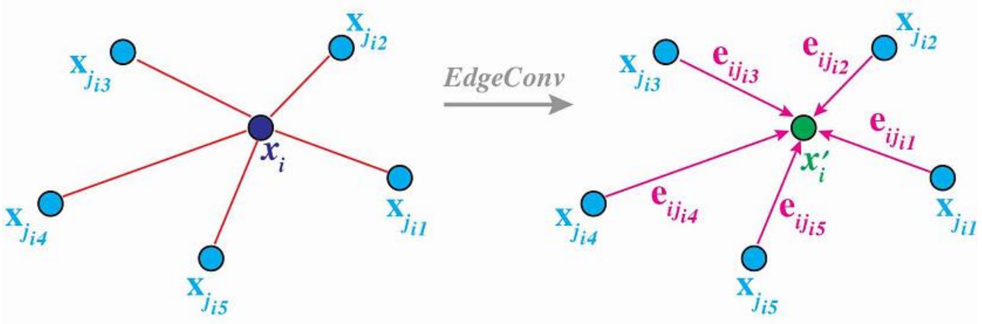
[32] Wang et al. Spectral Graph Convolution. ECCV 2018.

谱域变换 + 图卷积GCN

$$\mathbf{y} = g_\theta(\mathcal{L})\mathbf{x} = \sum_{k=0}^{K-1} \theta_k T_k(\mathcal{L})\mathbf{x}$$

复杂度较高

多项式近似 & 动态图静态核

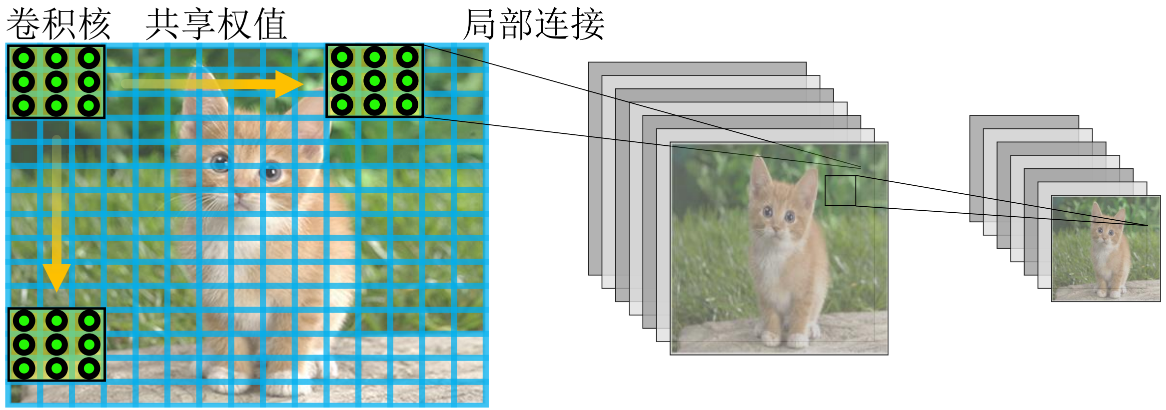


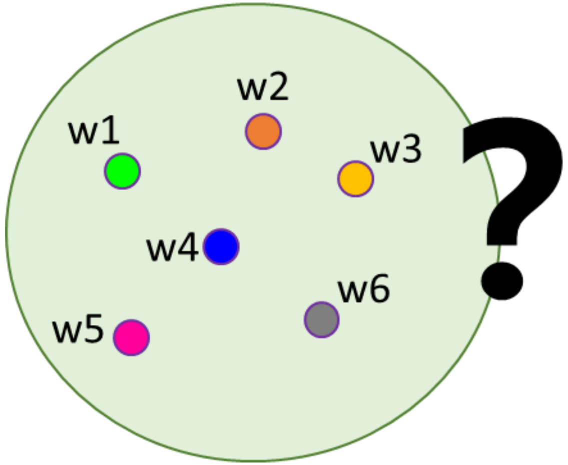
[29] Wang et al. DGCNN. TOG 2019.

空域动态建图 + EdgeConv

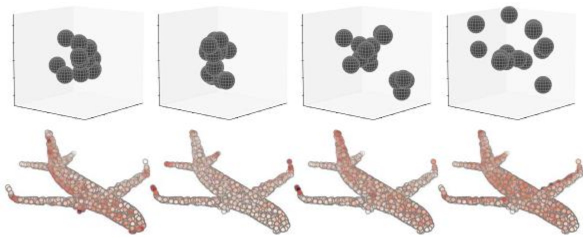
$$x'_i = \underset{j:(i,j)\in\mathcal{E}}{\square} h_{\Theta}(x_i, x_j). \qquad h_{\Theta}(x_i, x_j - x_i)$$

卷积核设计





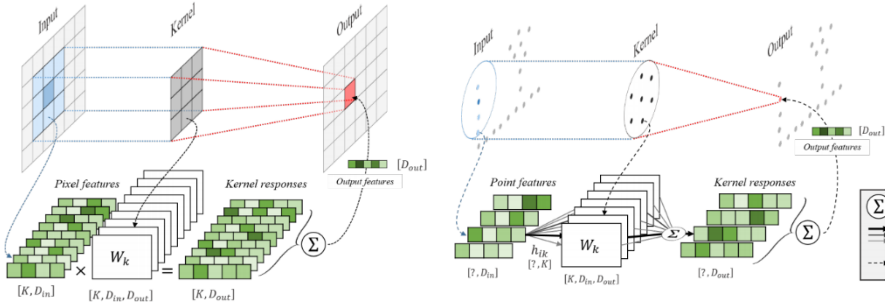
研究综述 卷积核设计



核相关滤波

(kernel correlation filtering)

$$\mathrm{KC}(\boldsymbol{\kappa}, \mathbf{x}_i) = \frac{1}{|\mathcal{N}(i)|} \sum_{m=1}^{M} \sum_{n \in \mathcal{N}(i)} \mathrm{K}_\sigma(\boldsymbol{\kappa}_m, \mathbf{x}_n - \mathbf{x}_i)$$



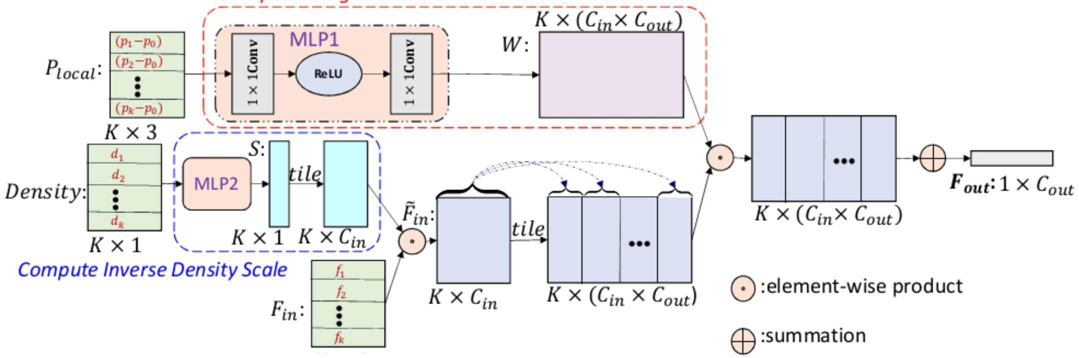
[39] Hugues et al. KPConv. ICCV 2019.

kernel points

$$(\mathcal{F} * g)(x) = \sum_{x_i \in \mathcal{N}_x} g(x_i - x) f_i \qquad g(y_i) = \sum_{k<K} h(y_i, \widetilde{x}_k) W_k$$

百家争鸣，各有优劣

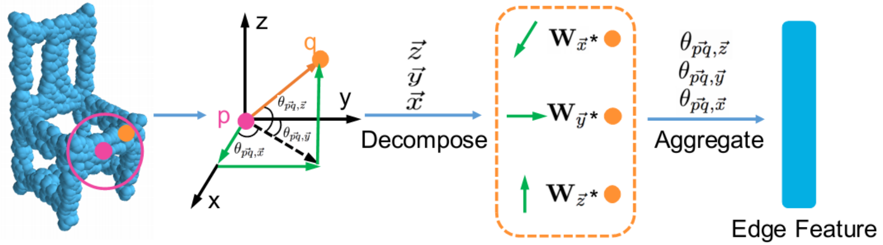
[23] Wu et al. PointConv. CVPR 2019.



$$\mathbf{F}_{out} = \sum_{k=1}^{K} \sum_{c_{in}=1}^{C_{in}} S(k) \mathbf{W}(k, c_{in}) F_{in}(k, c_{in})$$

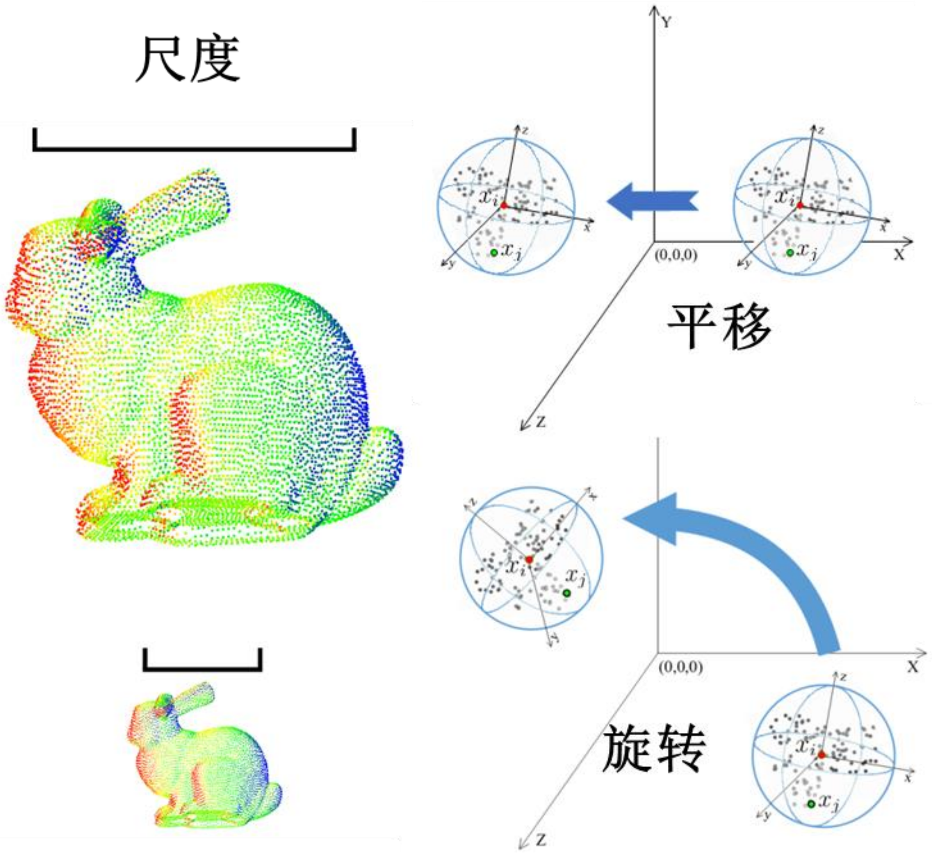
显式引入密度项

[41] Lan et al. Geo-CNN. CVPR 2019.



显式编码方向

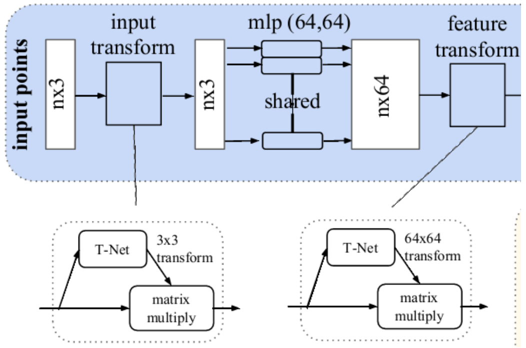
鲁棒性问题



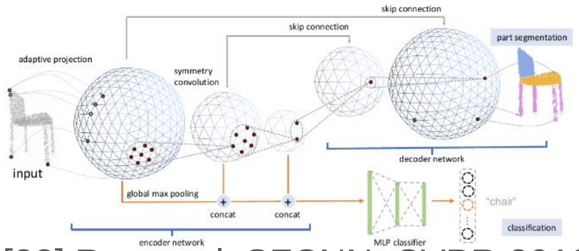
研究综述 鲁棒性

- **几何变换鲁棒性**
  - 数据预处理：
  - ✓ 平移
  - ✓ 尺度
  - x 旋转

数据增强 or 姿态归一化



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



[28] Rao et al. SFCNN. CVPR 2019.

Rotation-Invariant (RI) 表示

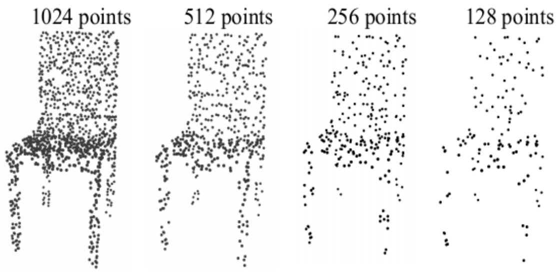
$$\|R\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}\|_2^2 \quad \langle R\boldsymbol{x}, R\boldsymbol{y}\rangle = (R\boldsymbol{x})^\mathsf{T}(R\boldsymbol{y}) = \boldsymbol{x}^\mathsf{T}\boldsymbol{y} = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$$

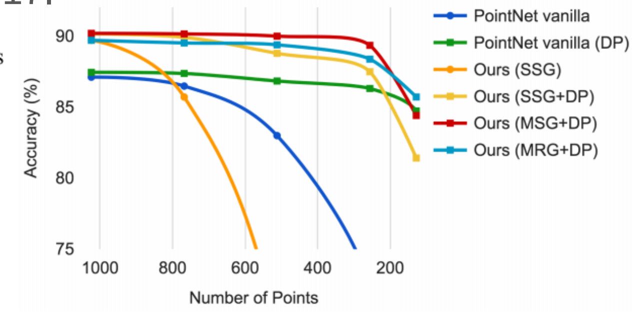
[40] Chen et al. ClusterNet. CVPR 2019.

- **密度鲁棒性**
  - 多尺度/多分辨率学习
  - Monte Carlo 积分
  - 引入密度信息

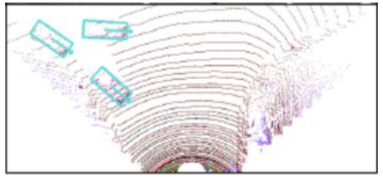
[22] HERMOSILLA et al. MCCNN. TOG 2018.
[23] Wu et al. PointConv. CVPR 2019.

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



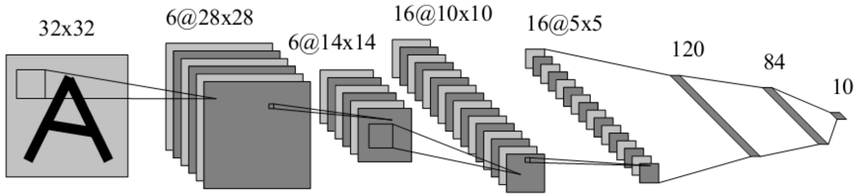


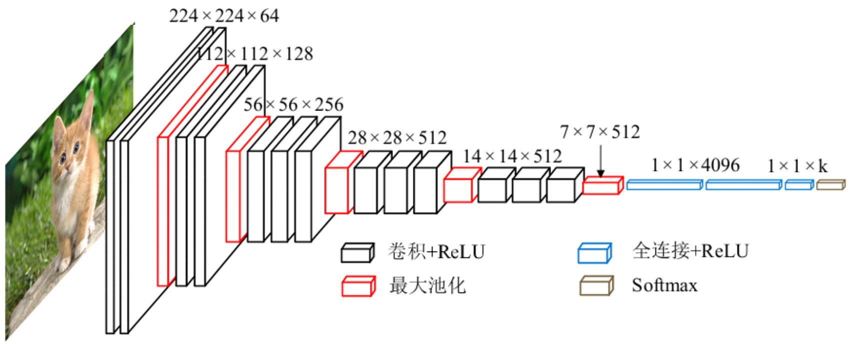
- **点云缺失鲁棒性**
  - 点云补全

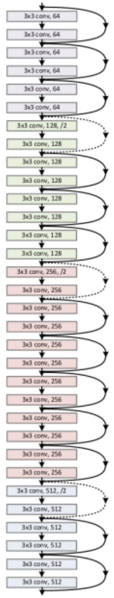


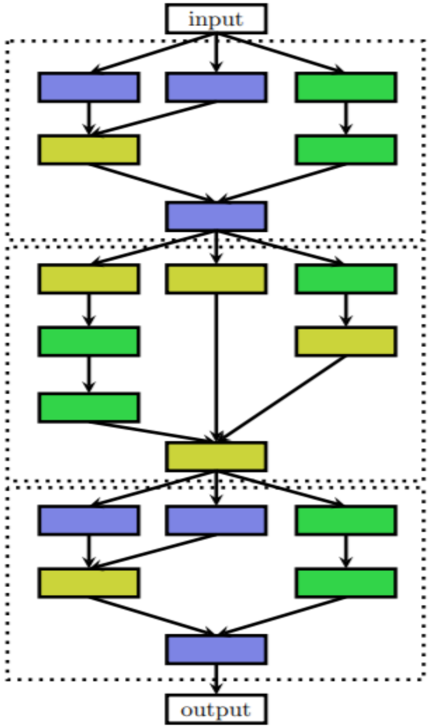
鲁棒性在方法、模型层面缺乏研究

神经架构搜索 (NAS)

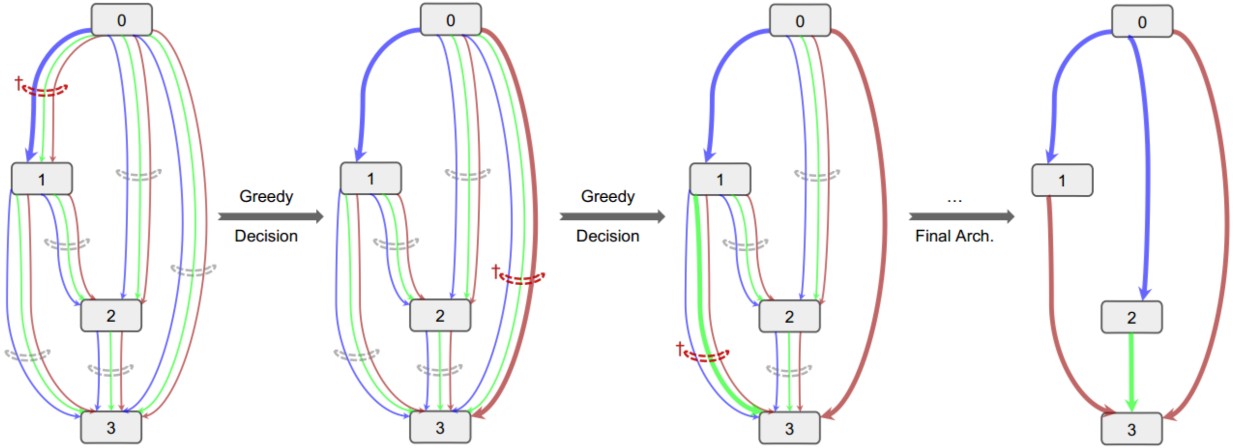

32x32
6@28x28
6@14x14
16@10x10
16@5x5
120
84
10


224×224×64
112×112×128
56×56×256
28×28×512
14×14×512
7×7×512
1×1×4096
1×1×k
卷积+ReLU
全连接+ReLU
最大池化
Softmax

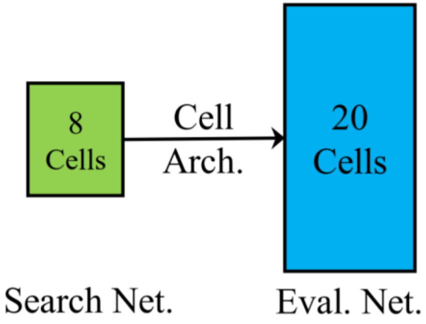



input
output

研究综述 神经架构搜索



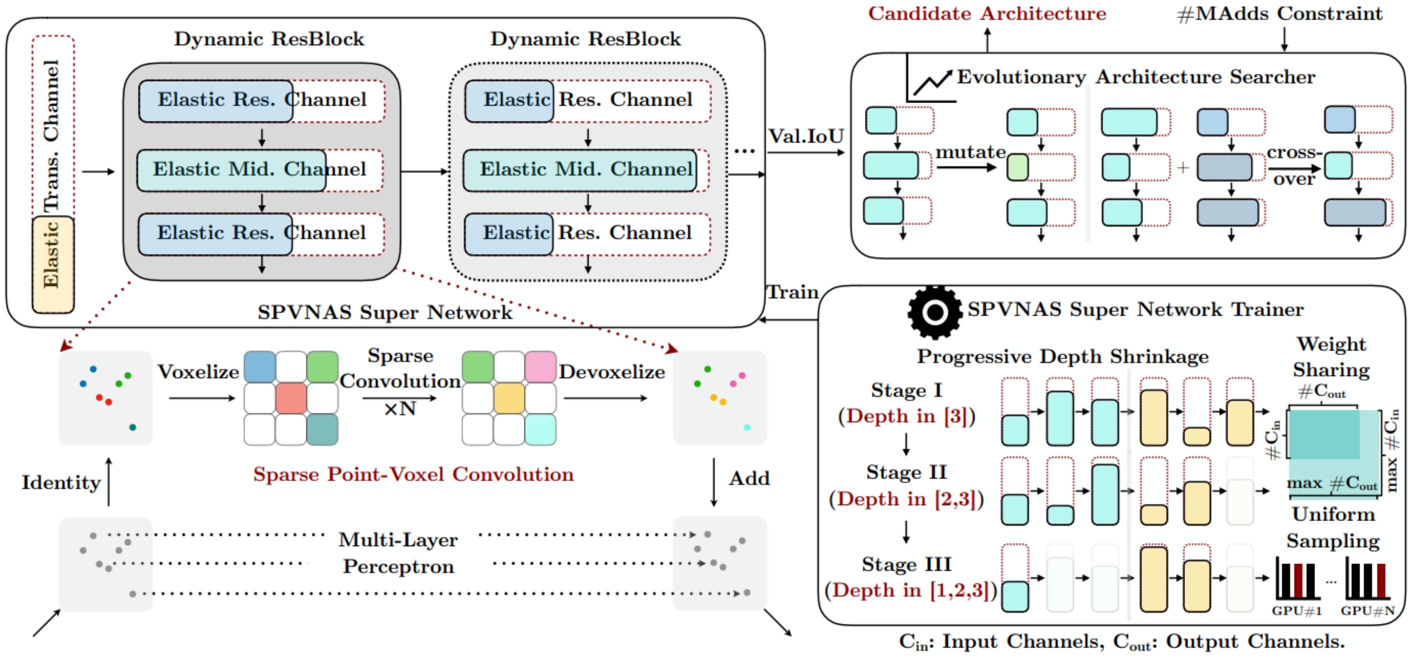
[50] Li et al. SGAS. CVPR 2020.



Gap: 搜索 -> 评估

Darts框架：贪婪策略

候选算子：卷积集成



[51] Tang et al. SPV-NAS. ECCV 2020.

卷积层面：

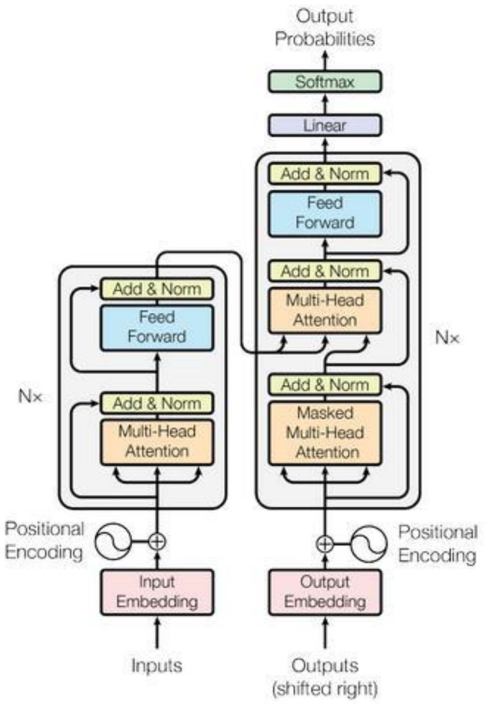
体素（粗糙） + 点云（精细）

架构层面：

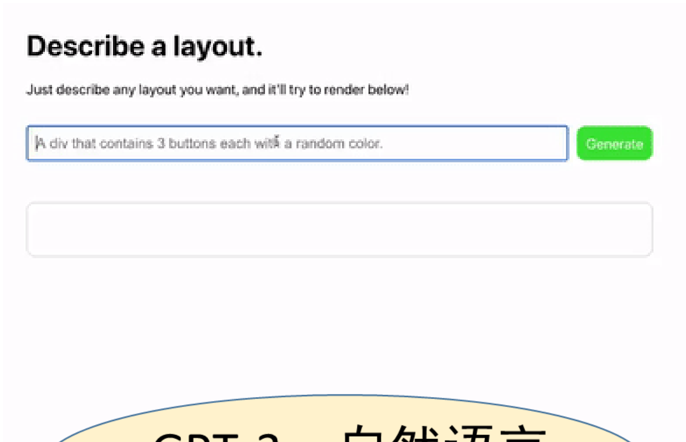
常规NAS + PDS深度衰减

缺乏针对点云结构学习的NAS研究

Transformer应用



OpenAI：GPT-3



GPT-3：自然语言

Generative Pre-Training

写文章

算数

翻译

做报表

聊天

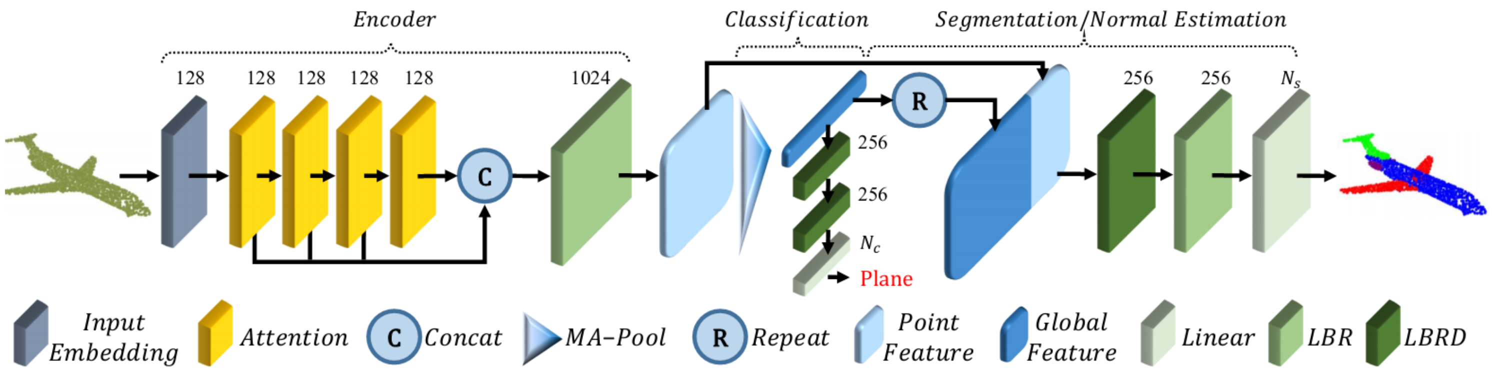
写代码

问答

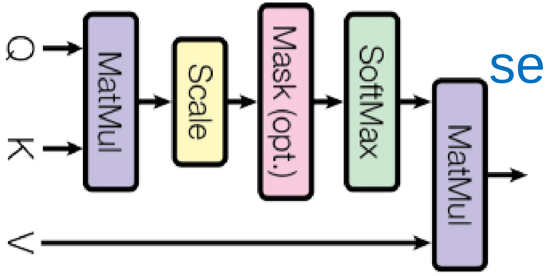
....

....

研究综述 Transformer应用



[52] Guo et al. PCT. CVMJ 2021.



标准 self-attention

应用Transformer为主

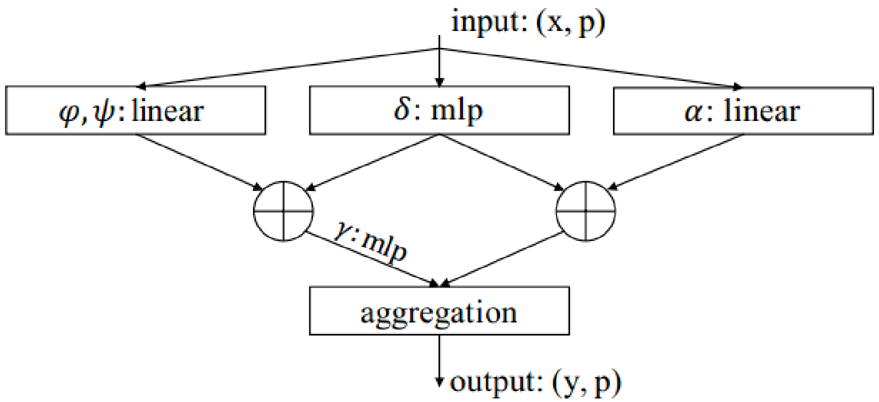
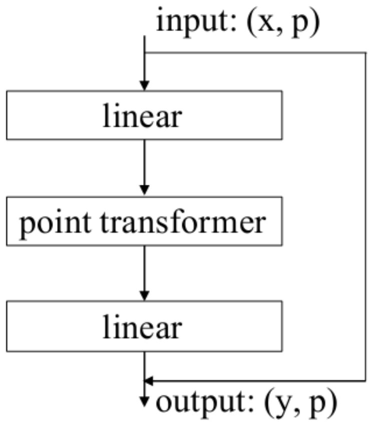


Figure 2. Point transformer layer.



(a) point transformer block

[53] Zhao et al. Point transformer. arXiv 2021.

标准 self-attention

$$\mathbf{y}_i = \sum_{\mathbf{x}_j \in \mathcal{X}} \rho\big(\varphi(\mathbf{x}_i)^\top \psi(\mathbf{x}_j) + \delta\big)\alpha(\mathbf{x}_j)$$

图像: self-attention 位置编码

$$\mathbf{y}_i = \sum_{\mathbf{x}_j \in \mathcal{X}(i)} \rho\big(\gamma(\varphi(\mathbf{x}_i) - \psi(\mathbf{x}_j) + \delta)\big) \odot \big(\alpha(\mathbf{x}_j) + \delta\big)$$

[54] Shaw et al. Relative position SA. NAACL 2018.

Github: **awesome-point-cloud-analysis**

CVPR, ICCV, ECCV, SIGGraph / Asia, TOG, NeurIPS, ICLR, AAAI, MM, ICRA, IROS, 3DV…… arXiv

awesome-point-cloud-analysis awesome

```
- 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, ~70
ECCV 2020, ~40
CVPR 2021, ?

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.`] 🔥 ⭐

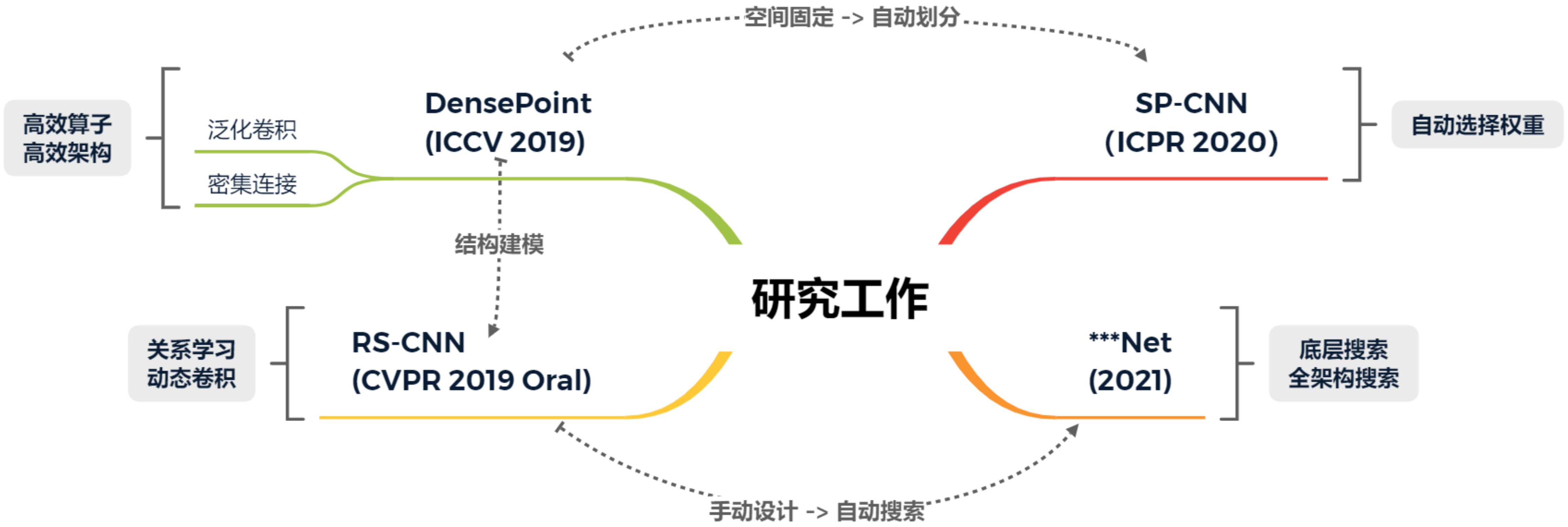
提纲

1 背景简介

2 研究综述

3 **工作介绍**

工作介绍



空间固定 -> 自动划分
高效算子
高效架构
泛化卷积
密集连接
DensePoint
(ICCV 2019)
SP-CNN
(ICPR 2020)
自动选择权重
结构建模
研究工作
关系学习
动态卷积
RS-CNN
(CVPR 2019 Oral)
***Net
(2021)
底层搜索
全架构搜索
手动设计 -> 自动搜索

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

工作一：DensePoint 研究方法

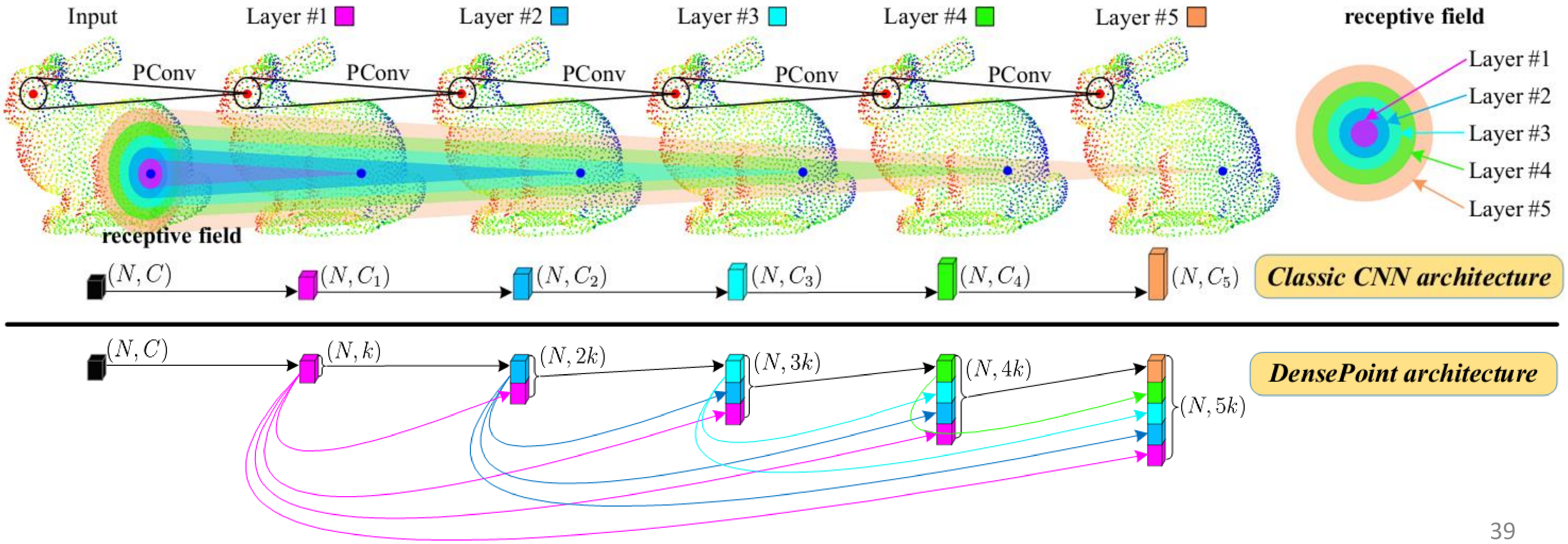
核心思想：<u>多级别感受野</u> + <u>高效点云卷积</u>

密集连接 + 高效点云卷积

以**有机协同**的方式，渐近地聚集**语义层级化的多尺度**信息

语义层级化

视觉上下文



工作一：DensePoint 研究方法

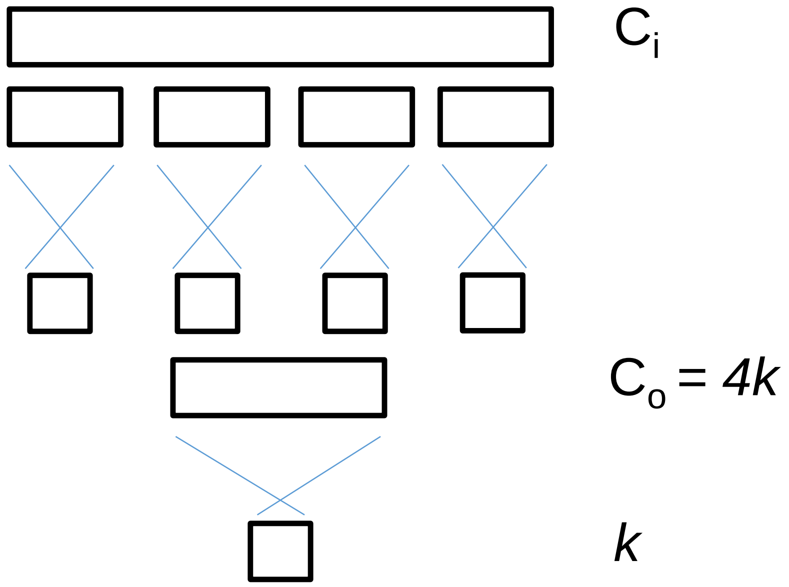
$$\mathbf{f}_{\mathcal{N}(x)} = \rho\Big(\{\phi(\mathbf{f}_{x_n}),\ \forall x_n \in \mathcal{N}(x)\}\Big)$$

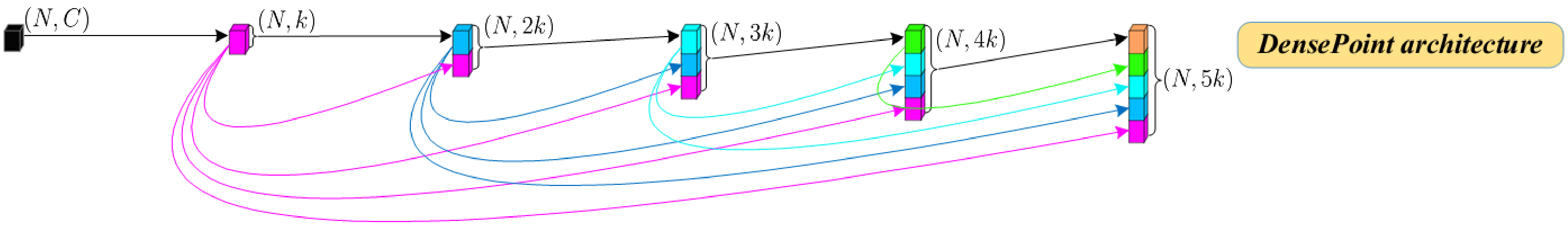
ø: 单层感知器

强化版本: 滤波器分组 (filter grouping)

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

复杂度：C_i*k vs. $C_i*4k/g + 4k^2$



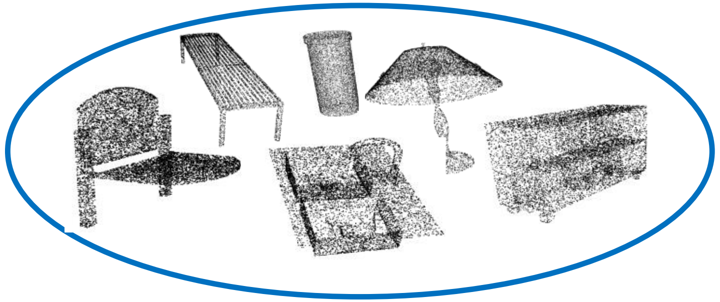


工作一：DensePoint 点云形状分类

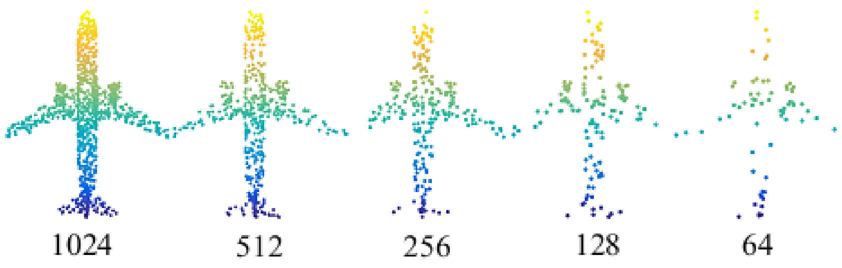
ModelNet40

ModelNet10

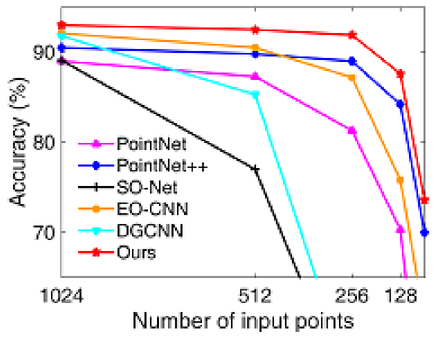
Benchmark



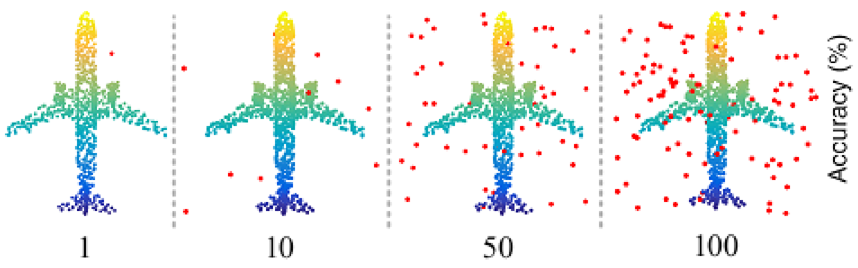
鲁棒性：采样密度 & 噪声点



(a)

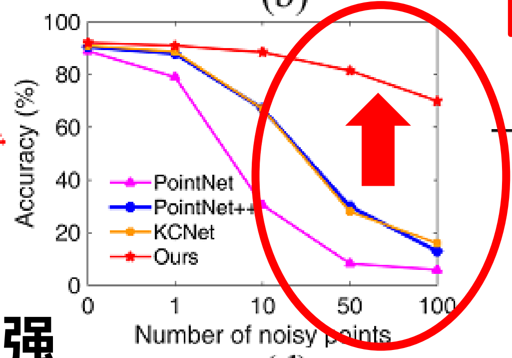


(b)



(c)

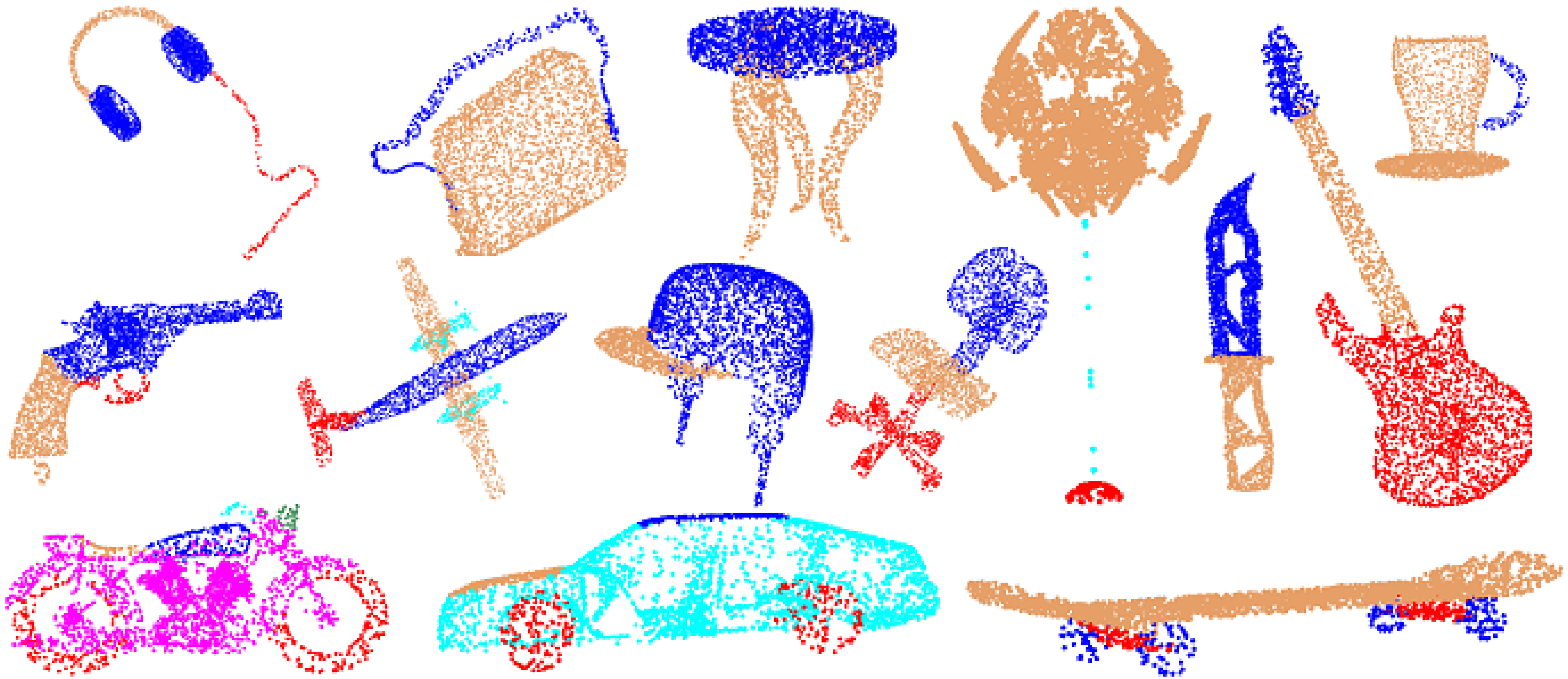
无数据增强



(d)

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

工作一：DensePoint 点云部件分割



✓ 对丰富多样的形状结构具有较强辨识能力

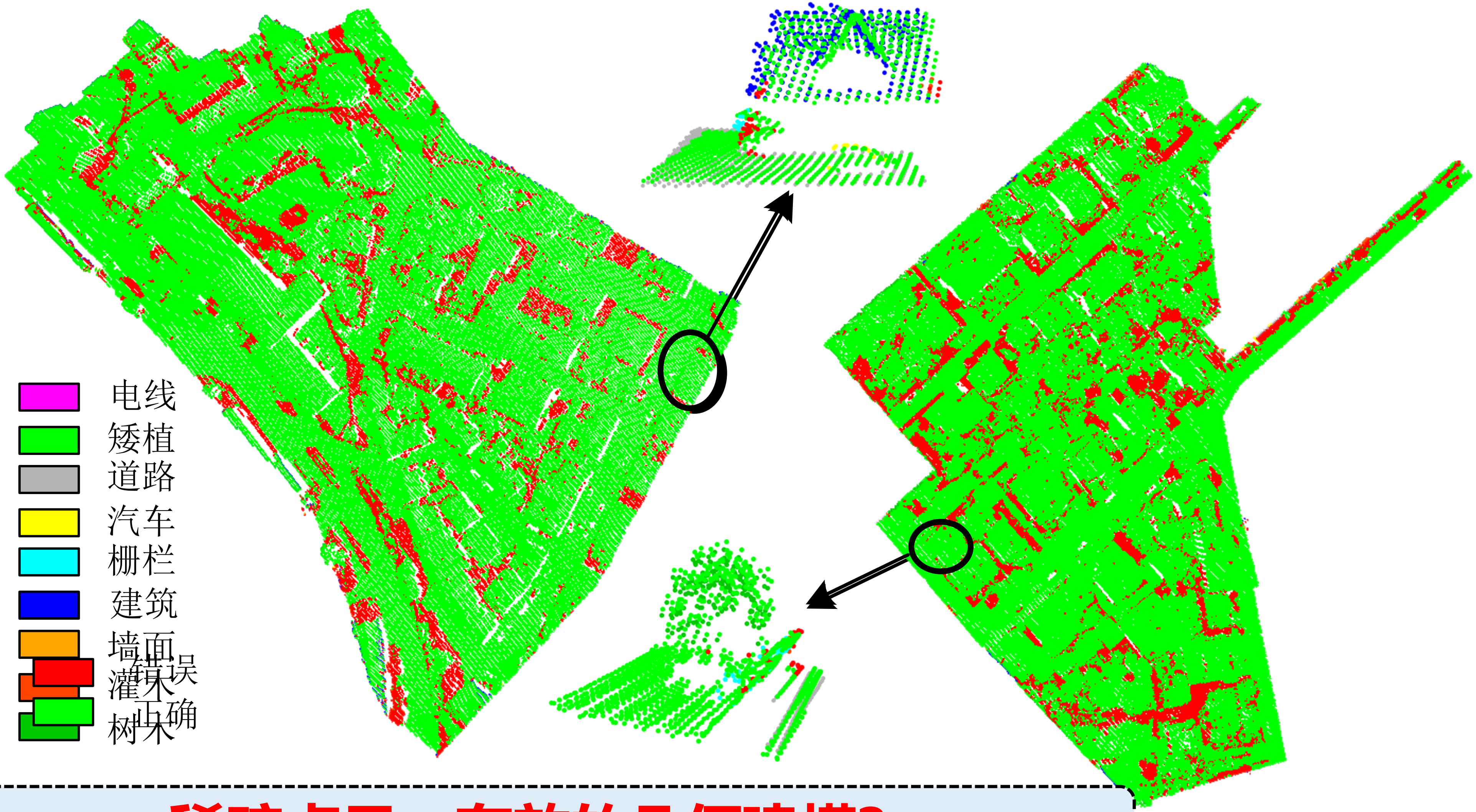
$$\mathbf{f}_{\mathcal{N}(x)} = \rho\left(\{\phi(\mathbf{f}_{x_n}),\ \forall x_n \in \mathcal{N}(x)\}\right)$$ ϕ：单层感知器

缺乏内在的几何结构建模

工作一：DensePoint　ALS点云场景分割

全局分割较好

局部有待改善



稀疏点云：有效的几何建模?

工作一：DensePoint 复杂度实测

空间

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 点

↑ 3 倍

时间

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

↑ 4 倍

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

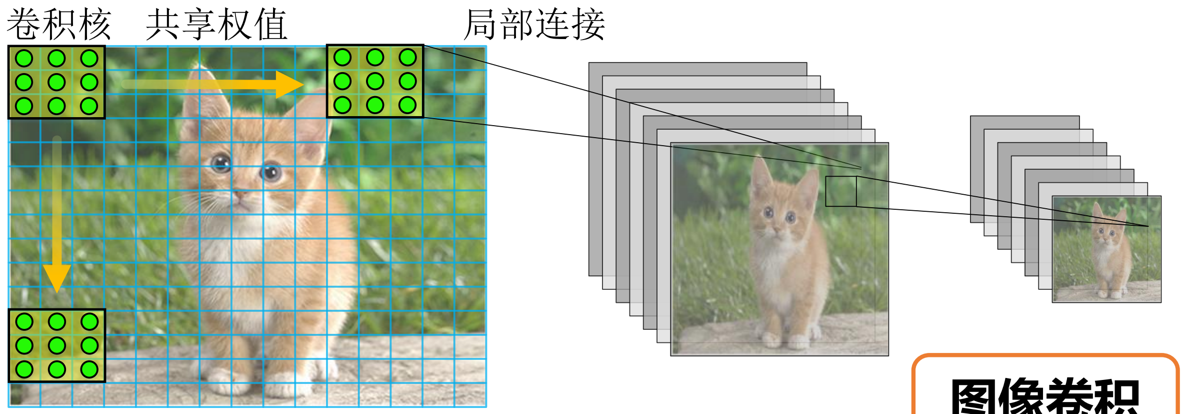
Project Page: https://yochengliu.github.io/Relation-Shape-CNN/

工作二：Relation-Shape CNN 研究方法

➢ 图像 – 像素卷积

- 规则排列，权重 w_j 与有序索引 j 对应
- 一权一点，权值不等 – **因像素信息而异**

$$y = \sigma(\textstyle\sum_j w_j x_j + b)$$



图像卷积

➢ 点云 – 关系卷积

- 无序性 – 置换排列不变性

(i) 特征变换 $\mathcal{T}$

(ii) 特征聚集 $\mathcal{A}$

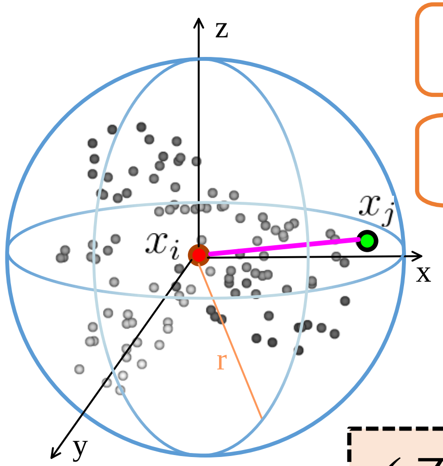
$$\mathbf{f}_{P_{\text{sub}}} = \sigma\big(\mathcal{A}(\{\mathcal{T}(\mathbf{f}_{x_j}),\ \forall x_j\})\big)$$

当且仅当 T 函数对邻域内所有点共享权重，

A 函数为对称函数，比如求和，max pool等

$$\mathcal{T}(\mathbf{f}_{x_j}) = \boxed{\mathbf{w}_j} \odot \mathbf{f}_{x_j}$$

- ■ 邻域内不共享
- ■ 未建模交互



目标转变: $\mathcal{T}(\mathbf{f}_{x_j}) = \boxed{\mathbf{w}_{ij}} \odot \mathbf{f}_{x_j}$

从几何关系中学习: $= \mathcal{M}(\mathbf{h}_{ij}) \odot \mathbf{f}_{x_j}$

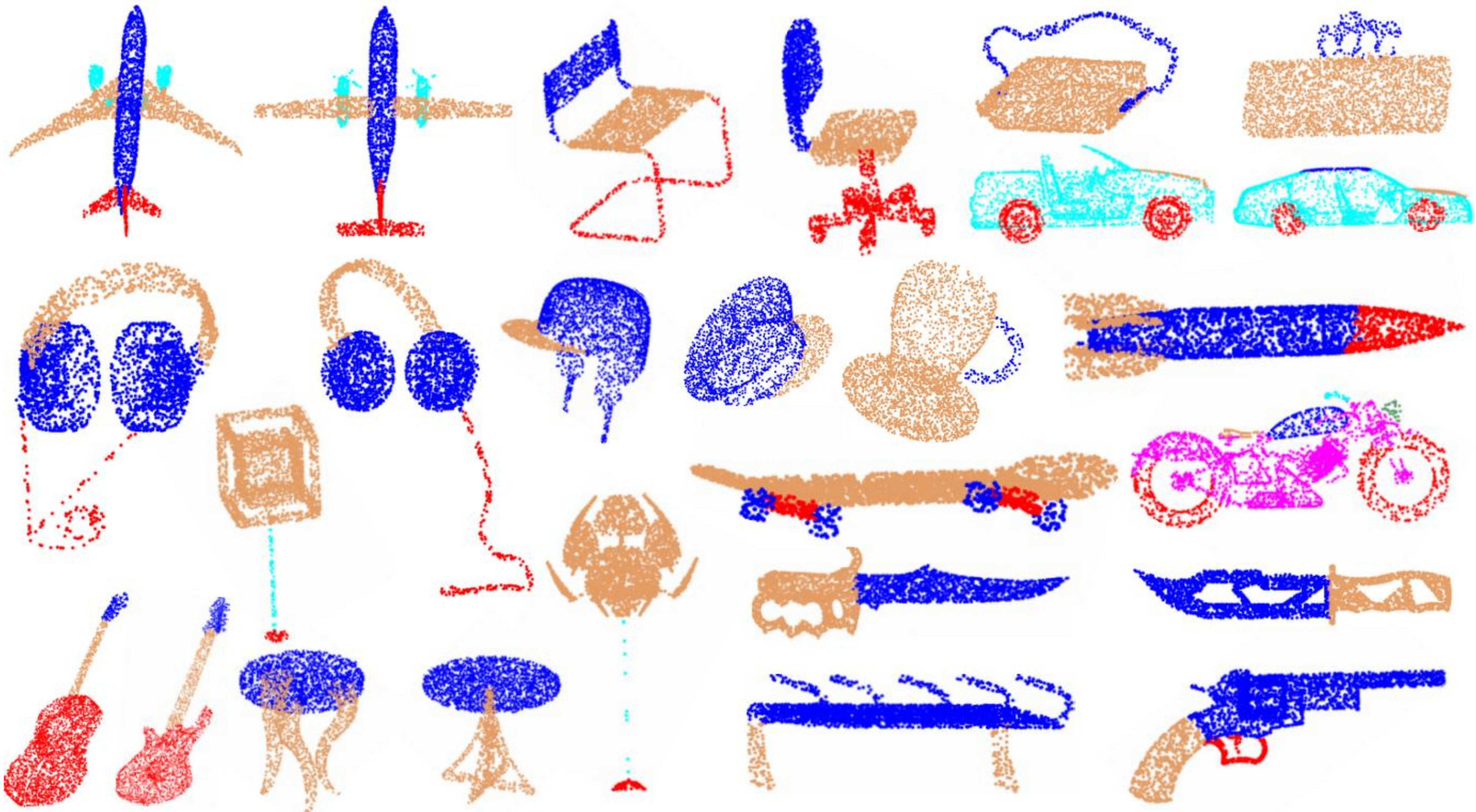
$\mathbf{h}_{ij}$ ：预定义几何先验 – 低维几何交互

$\mathcal{M}$ ：映射函数 (MLP) – 高维几何表达

✓ 不规则排列，权重 w_{ij} 与无序点对(i, j)对应

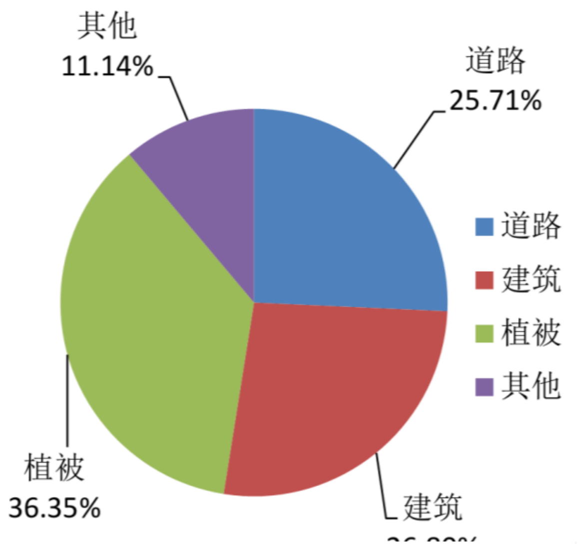
✓ 一权一点，权值不等 – **因几何关系而异**

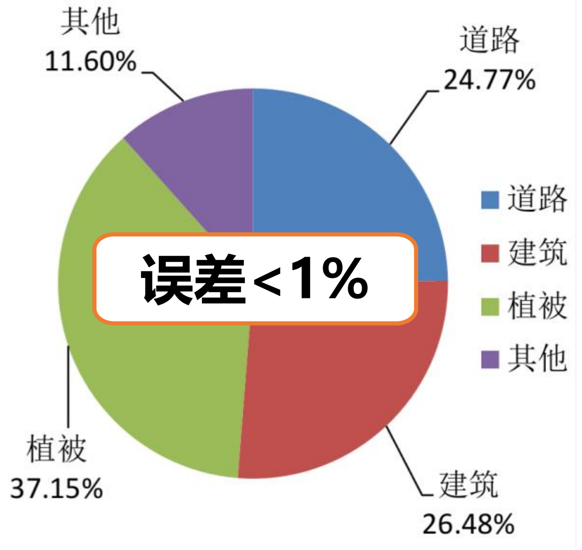
工作二：Relation-Shape CNN　点云部件分割

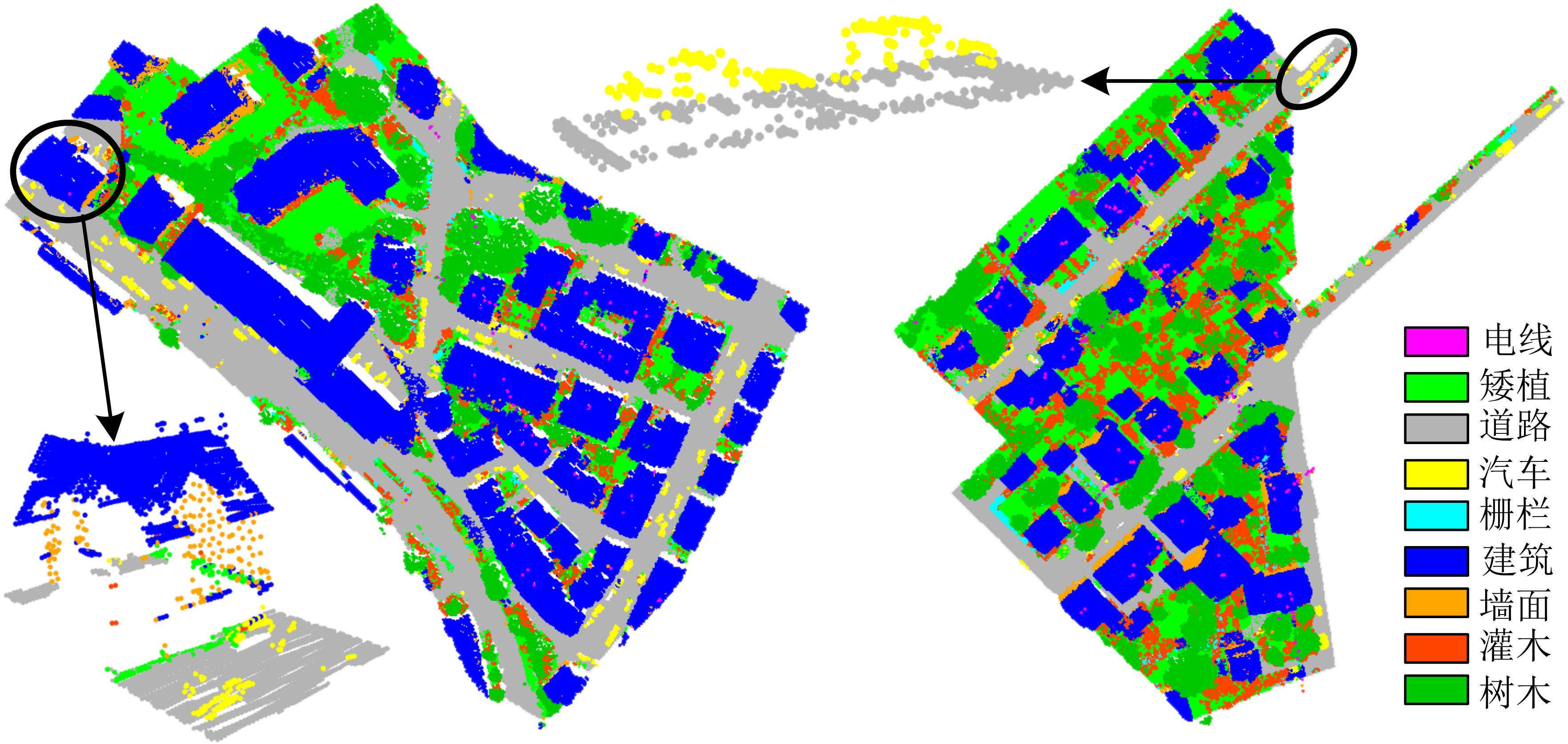


✓ 空间结构建模充分：细粒度分割效果提升

工作二：Relation-Shape CNN ALS点云场景分割







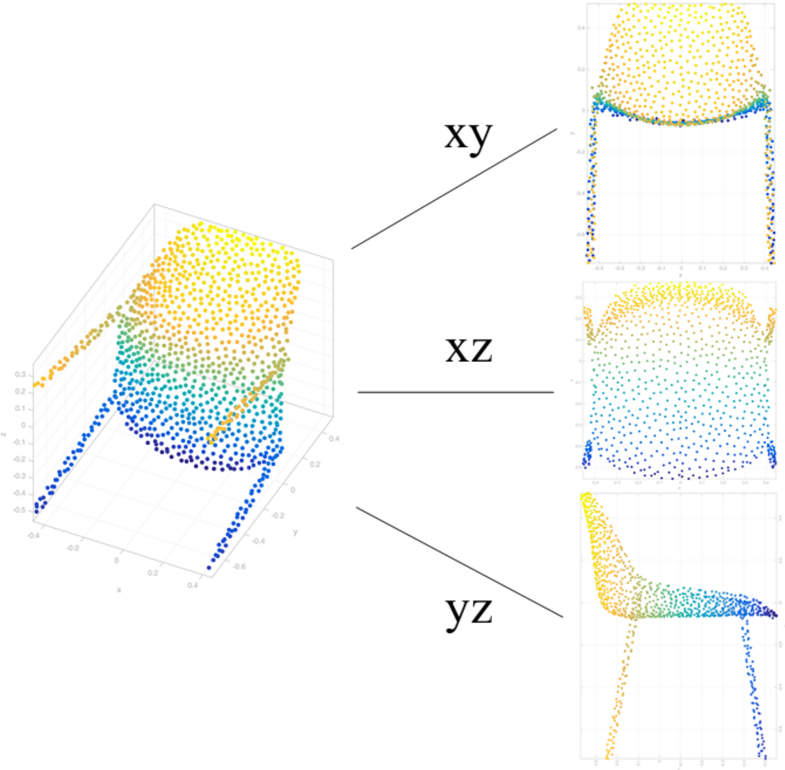
全局连贯，局部精准

工作二：Relation-Shape CNN 几何先验分析

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

几何先验可灵活定义

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

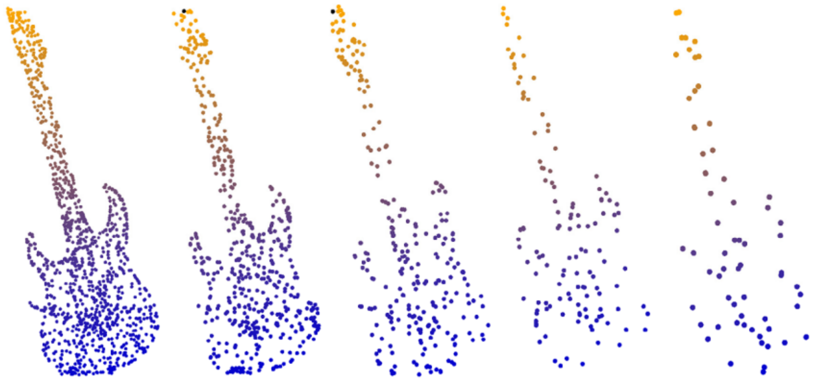


low-level relation $\mathbf{h}$	channels	acc.
(XY-Ed, $x_i^{\text{xy}} - x_j^{\text{xy}}$, x_i^{xy}, x_j^{xy})	10	92.1
(XZ-Ed, $x_i^{\text{xz}} - x_j^{\text{xz}}$, x_i^{xz}, x_j^{xz})	10	92.1
(YZ-Ed, $x_i^{\text{yz}} - x_j^{\text{yz}}$, x_i^{yz}, x_j^{yz})	10	92.2
fusion of above three views		92.5

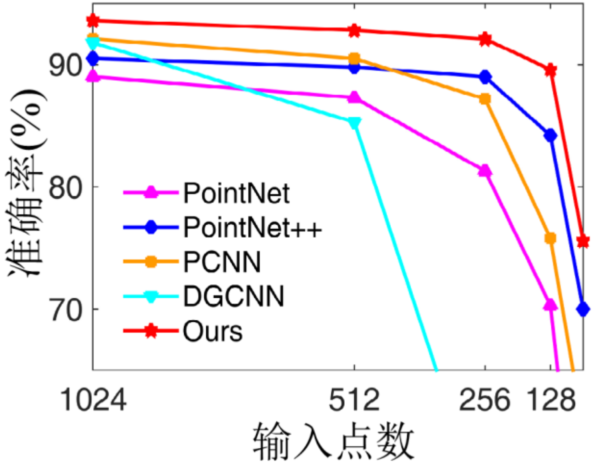
工作二：Relation-Shape CNN　复杂度与鲁棒性

鲁棒性：几何变换 & 采样密度

方法	OAcc.	点集置换	平移+0.2	平移-0.2	旋转90°	旋转180°
PointNet [100]	88.7	88.7	70.8	70.6	42.5	38.6
PointNet++ [98]	88.2	88.2	88.2	88.2	47.9	39.7
RS-CNN	**90.3**	**90.3**	**90.3**	**90.3**	**90.3**	**90.3**



1024　512　256　128　64



计算复杂度

模型	#点数	#参数量	#FLOPs/样本
PointNet [100]	1024	3.50M	440M
PointNet++ [139]	1024	1.48M	1684M
DGCNN [139]	1024	1.84M	2767M
SpecGCN [139]	1024	2.05M	1112M
KCNet [174]	1024	0.90M	-
PCNN [139]	1024	8.20M	**294M**
PointCNN [139]	1024	**0.60M**	1581M
RS-CNN	1024	1.41M	295M

↑1.5 倍

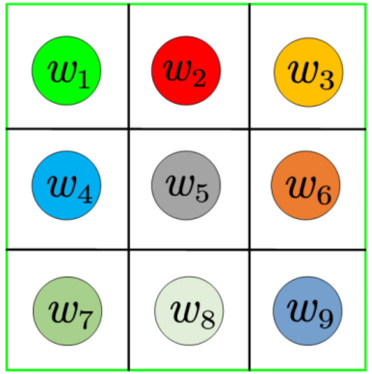
Deep Space Probing for Point Cloud Analysis

Yirong Yang, Bin Fan, **Yongcheng Liu**, Hua Lin, Jiyong Zhang,

Xin Liu, Xinyu Cai, Shiming Xiang, Chunhong Pan

ICPR 2020

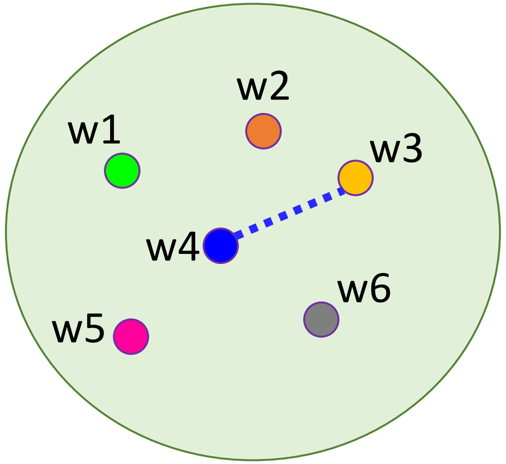
工作三： Space-Probe CNN 研究动机



卷积操作

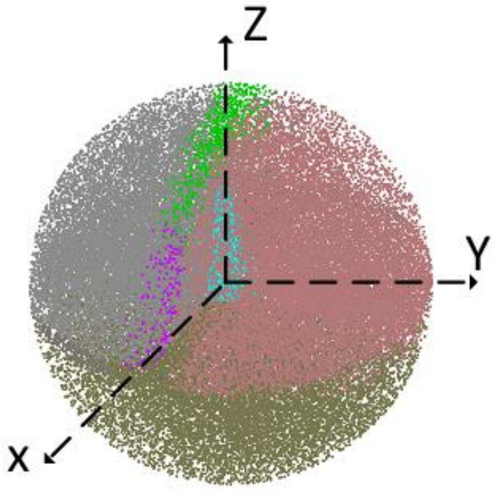
2D 图像 ← $$y = \sigma\left(\sum_j w_j x_j + b\right)$$ → 3D 点云

栅格结构一对一　　　　空间几何结构一对一

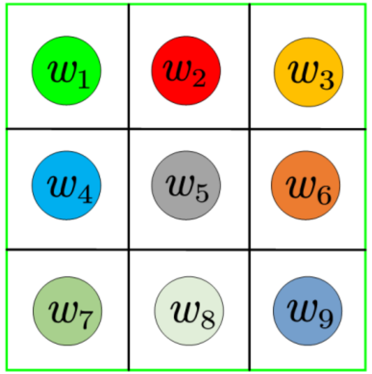


是否能将连续的3D空间划分成多个区域?

✓ 数据驱动　　✓ 可学习　　✓ 几何自适应



工作三： Space-Probe CNN 研究方法

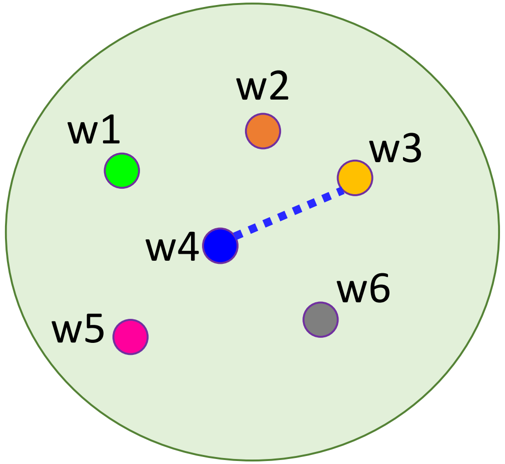


$$\hat{\mathbf{f}}_i^k = \sigma\left(\sum_{j}^{n\times n} \left(\boxed{\mathbf{w}_j^k \cdot \mathbf{f}_j} \mid \mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)\right) + b\right)$$

函数抽象化

$$\hat{\mathbf{f}}_i^k = \sigma\left(\mathcal{G}\left\{\boxed{\mathcal{W}_{\mathcal{I}_j}^k(\mathbf{f}_j)} \mid \mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)\right\} + b\right)$$

$\mathcal{I}_j$: matching函数 $\quad \mathcal{I}_j = j$



PointNet: $\mathcal{I}_j = 1$

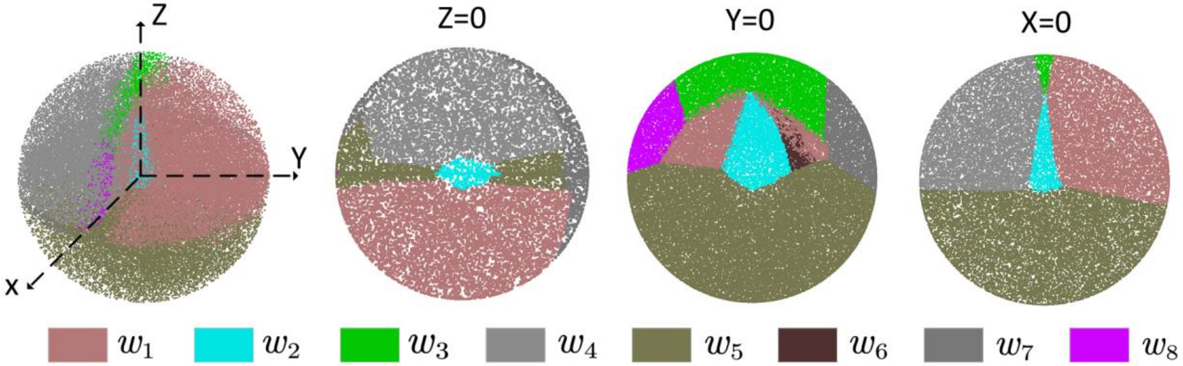
$$\hat{\mathbf{f}}_i^k = \sigma\left(\max_j \boxed{\left(\mathbf{w}^k \cdot \mathbf{f}_j\right.} \mid \mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)) + b\right)$$

index-mapping函数:

$$\mathcal{I}_j = \gamma(\mathbf{g}_{ji}) \mid \mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)$$

$$\mathbf{g}_{ji} = (\mathbf{x}_j - \mathbf{x}_i) \mid \mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)$$

$$\mathcal{I}_j = \operatorname{argmax}\left(\mathrm{MLP}(\mathbf{g}_{ji}) \mid \mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)\right)$$

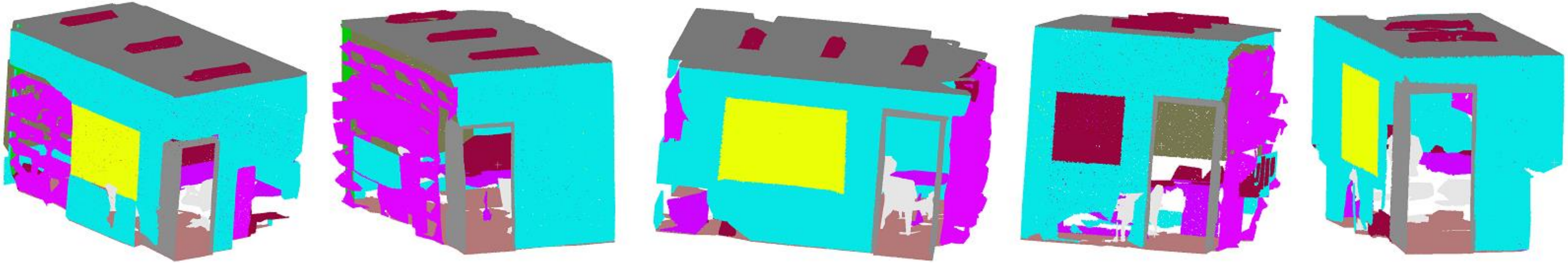


✓ 数据驱动 ✓ 可学习 ✓ 几何自适应

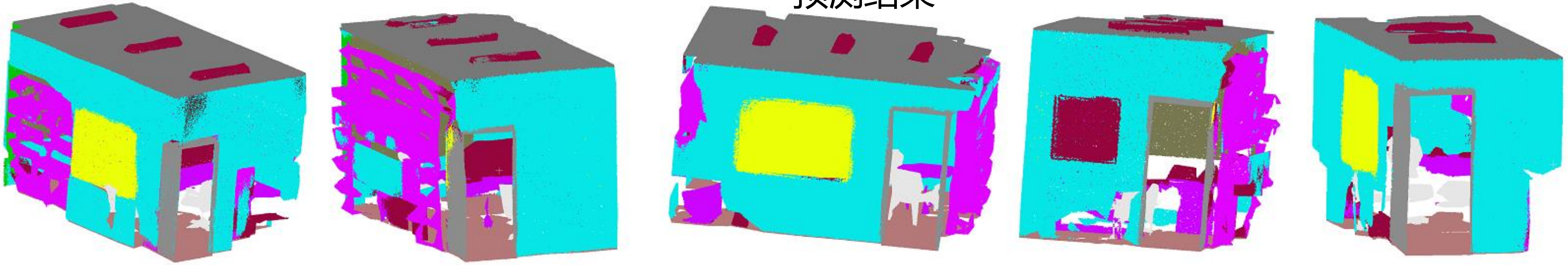
工作三： Space-Probe CNN 点云场景分割

S3DIS 大型场景分割Benchmark

Ground truth

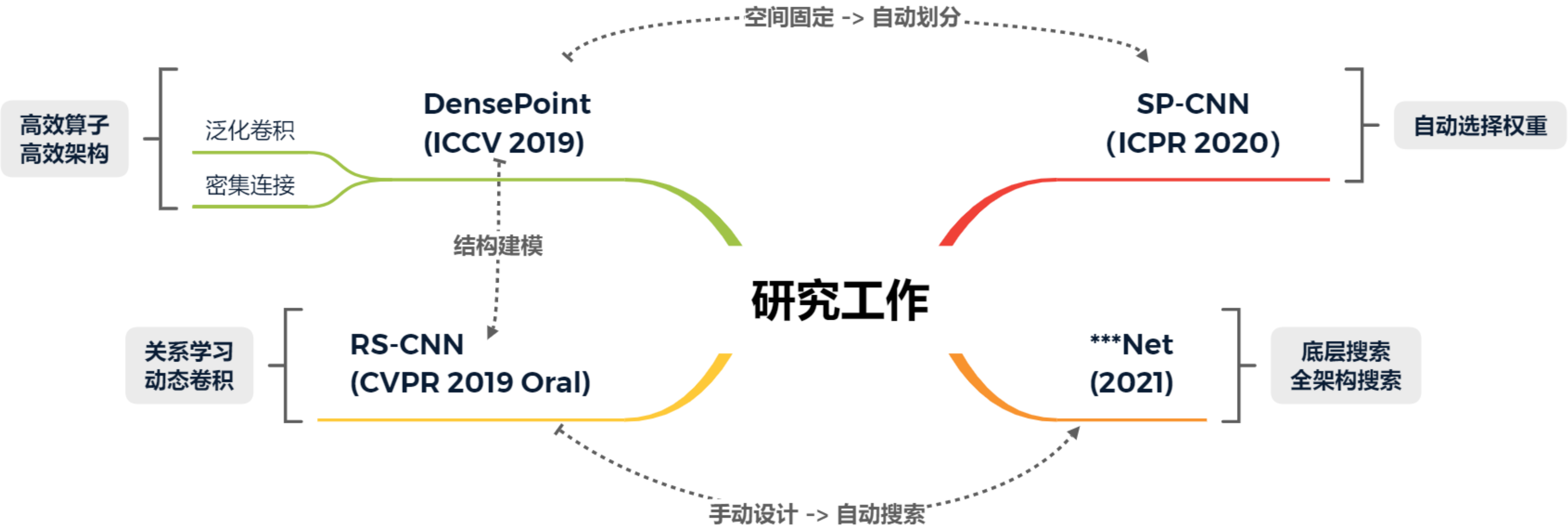


预测结果



场景分割：良好的语义连贯性

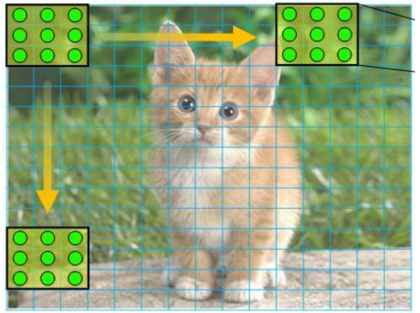
工作介绍



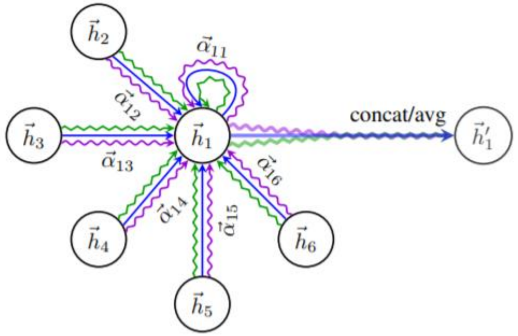
空间固定 -> 自动划分
高效算子
高效架构
泛化卷积
密集连接
DensePoint
(ICCV 2019)
SP-CNN
(ICPR 2020)
自动选择权重
结构建模
研究工作
关系学习
动态卷积
RS-CNN
(CVPR 2019 Oral)
***Net
(2021)
底层搜索
全架构搜索
手动设计 -> 自动搜索

工作四：***Net

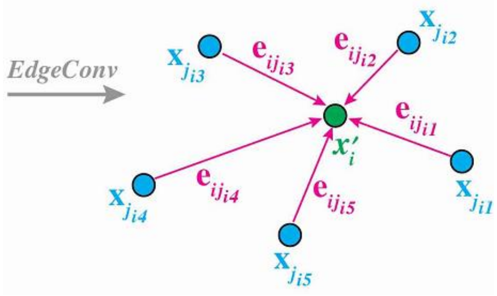
图像：1*1， 3*3， 5*5， 3*1， 1*3 各种尺寸的卷积核

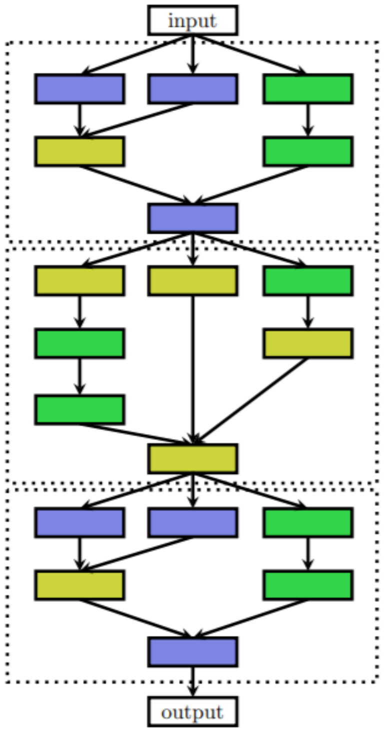


点云：GAT，SemiGCN，EdgeConv 各种现有手工卷积



……





现有NAS均基于固定的卷积算子搜索上层网络架构

能否突破到底层卷积算子内部?

实现由底层操作到顶层架构的全网络搜索

工作四：***Net

$$\boldsymbol{p} = \mathcal{G}\left(\{\psi(\boldsymbol{p}_j)\}_{j=1,\ldots,n}\right)$$

↓ **Relation-Shape CNN (CVPR 2019)**

$$\boldsymbol{p}_i' = \mathcal{G}\left(\{\psi(\mathcal{D}(\boldsymbol{p}_i, \boldsymbol{p}_j))\}_{\boldsymbol{p}_j \in \mathcal{N}(\boldsymbol{p}_i)}\right)$$ 结构 D + 参数 ψ

$$\boldsymbol{p}_i' = \mathcal{G}\left(\left\{h_\gamma\left(\mathcal{D}(\boldsymbol{p}_i, \boldsymbol{p}_j)\right)\right\}_{\boldsymbol{p}_j \in \mathcal{N}(\boldsymbol{p}_i)}\right),$$

$$\mathcal{D}(\boldsymbol{p}_i, \boldsymbol{p}_j) = f_\beta\left(\forall \text{ EA} \in \mathcal{D}(\boldsymbol{p}_i, \boldsymbol{p}_j)\right).$$

PointNet (CVPR 2017)　　PointWeb (CVPR 2019)　　DGCNN（TOG 2019)

均为特例

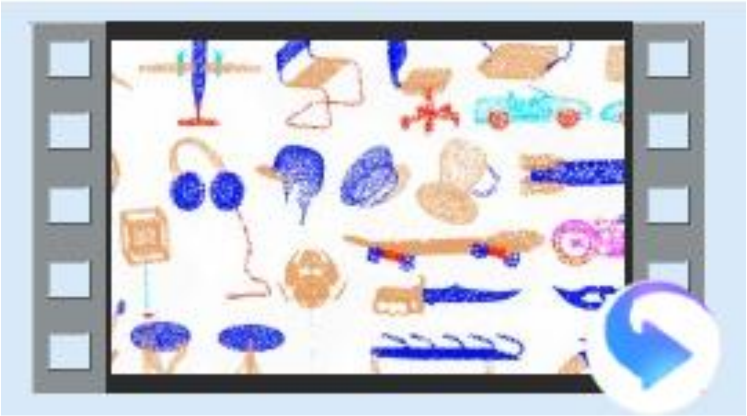
Relation-Shape-CNN（CVPR 2019)　　Pointwise（CVPR 2018)　　......

实现由底层卷积操作到顶层网络架构的全架构搜索

抛砖引玉，未完待续

点云分割效果展示

点云分割效果展示



一些经验之谈

☐ 数据层面

PointAugment. CVPR 2020

- ✓ 数据增强：scale (↑分类)、旋转（↑回归）、T-Net（↑回归）、输入Dropout（↑密度鲁棒）
- ✓ 数据稀疏：上采样up-sampling
- ✓ 分布不均匀：引入密度项　MCCNN. TOG 2018　PointConv. CVPR 2019
- ✓ 数据缺失：点云补全completion
- ✓ 超大尺度场景：重视数据切分方式　Superpoint Graph. CVPR 2018　Oversegmentation. CVPR 2019

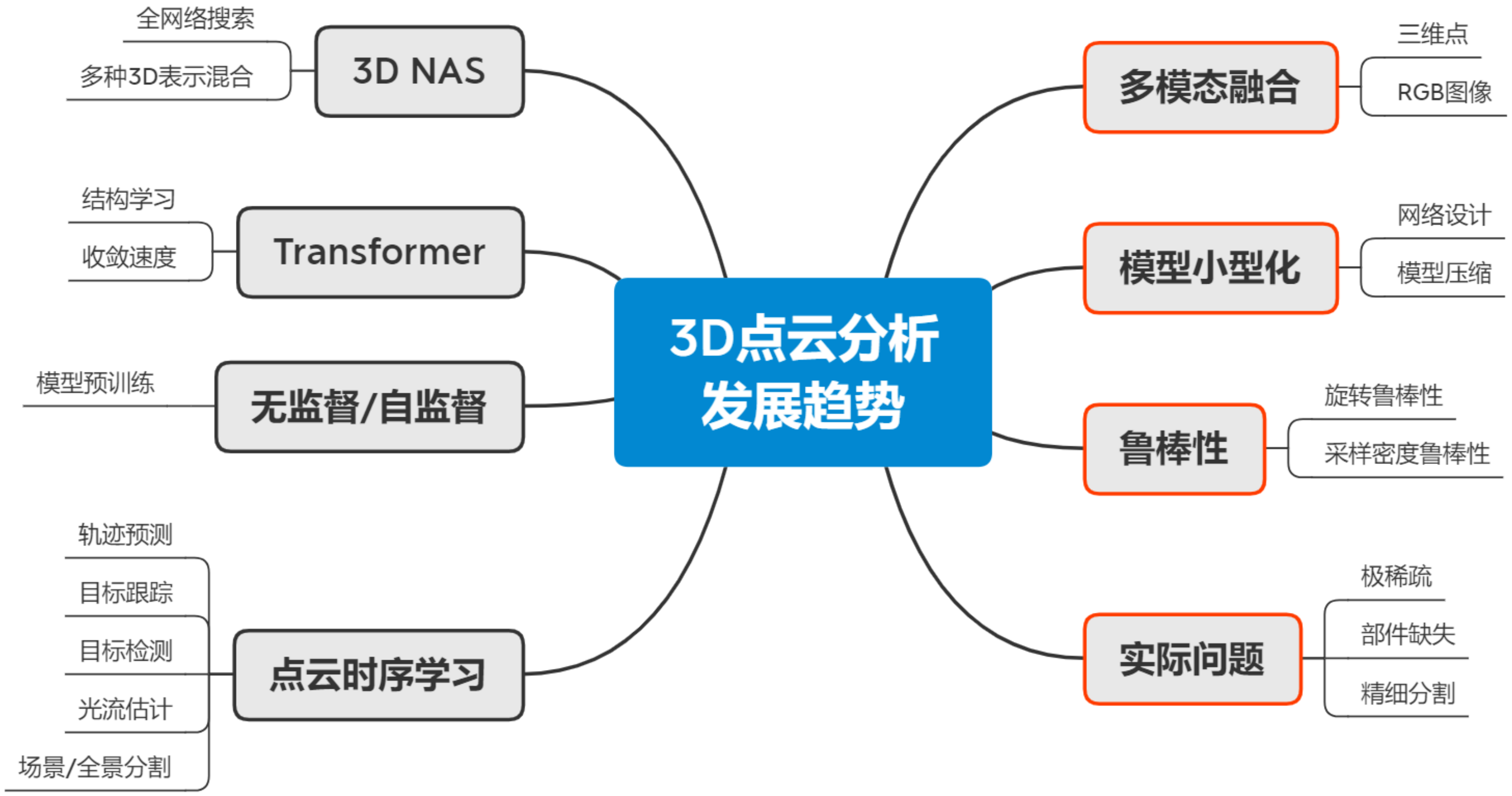
☐ 模型算法层面

- ✓ 追求精度：数据增强、高分辨率点云、NAS搜网络结构、Transformer
- ✓ 追求速度：原始点云(PointNet, DensePoint)、体素化voxelization or PointPillar
- ✓ Trade-off：体素化、原始点云均利用 (SPV-NAS, ECCV 2020)

研究发展趋势

- □ **3D NAS + Transformer**
  - 以Transformer为基础
  - Tranformer 与 CNN 混合
- □ **Transformer +** 自监督
  - 挖掘Transformer表征能力
  - 训练3D大模型
- □ 自监督 **+ 3D NAS**
  - 挖掘3D NAS潜力

……



- □ **Transformer +** 模型加速
  - Transformer 速度较快
- □ 多模态融合 —> 实际问题
  - 综合提升性能

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.

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.

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.

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.

[46] Wang et al. MSNet: Multi-Scale Convolutional Network for Point Cloud Classification. Remote Sens. 2018.

[47] Qin et al. Deep fusion of multi-view and multimodal representation of ALS point cloud for 3D terrain scene recognition. ISPRS Journal of Photogrammetry and Remote Sensing. 2018.

[48] Wang et al. A random forest classifier based on pixel comparison features for urban LiDAR data. ISPRS 2019.

[49] Mohammed et al. A multi-scale fully convolutional network for semantic labeling of 3D point clouds. ISPRS 2018.

[50] Li et al. SGAS: Sequential Greedy Architecture Search. CVPR 2020.

[51] Tang et al. Searching Efficient 3D Architectures with Sparse Point-Voxel Convolution. ECCV 2020.

[52] Guo et al. PCT: Point Cloud Transformer. CVMJ 2021.

[53] Zhao et al. Point Transformer. arXiv 2021.

[54] Shaw et al. Self-Attention with Relative Position Representations. NAACL-HLT 2018.

Thanks for your attention!

欢迎交流与合作！

http://vslab.ia.ac.cn/

yongcheng.liu@nlpr.ia.ac.cn