基于激光雷达点云的3D检测方法汇总(LiDAR only)

3D视觉工坊 2022-02-10 07:00

作者丨柒柒@知乎

来源丨https://zhuanlan.zhihu.com/p/436452723

编辑丨3D视觉工坊

前段时间比较忙，鸽了挺久，近期应该会恢复更新。

这篇文章主要是梳理一下近期3D Detection的进展，分类列举出一些我认为的比较重要的、有代表性的工作。

论文总结（一）主要讲解基于激光雷达点云的3D检测方法(LiDAR only)，欢迎补充指正。

一、论文分类汇总

1. 基于激光雷达点云的3D检测方法（LiDAR only）

方法名称	方法类别	发表年份	代码是否开源
Part-A^2	LiDAR only	TPAMI 2020	√
PointRCNN	LiDAR only	CVPR 2019	√
STD	LiDAR only	ICCV 2019	
PV-RCNN++	LiDAR only	CVPR 2020	√
PointPillar	LiDAR only	CVPR 2019	
MVP	LiDAR only	NIPS 2021	√
SE-SSD	LiDAR only	CVPR 2021	√
SA-SSD	LiDAR only	CVPR 2020	√
HVPR	LiDAR only	CVPR 2021	√
LiDAR RCNN	LiDAR only	CVPR 2021	√
SECOND	LiDAR only	Sensors 2018	√
3DIoUMatch	LiDAR only	CVPR 2021	√
CenterPoint	LiDAR only	CVPR 2021	√
3DSSD	LiDAR only	CVPR 2021	√
CIA-SSD	LiDAR only	AAAI 2021	

2. 基于多模态融合的3D检测方法（LiDAR+RGB）

方法名称	方法类别	发表年份	代码是否开源
AVOD-FPN	LiDAR+RGB	IROS 2018	√
F-PointNet	LiDAR+RGB	CVPR 2018	
F-ConvNet	LiDAR+RGB	IROS 2019	√
4D-Net	LiDAR+RGB	ICCV 2021	
MV3D	LiDAR+RGB	CVPR 2017	√
CM3D	LiDAR+RGB	WACV 2021	
H^2 3D RCNN	LiDAR+RGB	TCSVT 2021	√
ContFuse	LiDAR+RGB	ECCV 2018	

3. 基于单目图像的3D检测方法（Monocular）

方法名称	方法类别	发表年份	代码是否开源
AutoShape	Monocular	ICCV 2021	√
CaDDN	Monocular	CVPR 2021	√
MonoDLE	Monocular	CVPR 2021	√
DDMP	Monocular	CVPR 2021	√
GUPNet	Monocular	ICCV 2021	
FCOS3D	Monocular	ICCVW 2021	√
PGD	Monocular	CoRL 2021	√
MonoGRNet	Monocular	TPAMI 2021	

4. 基于双目图像的3D检测方法（Stereo）

方法名称	方法类别	发表年份	代码是否开源
SIDE	Stereo	WACV 2022	
LIGA-Stereo	Stereo	ICCV 2021	√
E2E-PL	Stereo	CVPR 2020	√

5. 基于视角特征提取的3D检测方法

方法名称	方法类别	发表年份	代码是否开源
H^2 3D RCNN	Front & Bird view	TCSVT 2021	√
PointPillar	Bird view	CVPR 2019	√
F-PointNet	Frustum	CVPR 2018	
F-ConvNet	Frustum	IROS 2019	√
TANet	Bird view	AAAI 2020	√

6. 基于特征补充/伪点云生成的3D检测方法（pseudo augment）

方法名称	方法类别	发表年份	代码是否开源
PointPainting	pseudo augment	CVPR 2020	
PointAugmenting	pseudo augment	CVPR 2021	
E2E-PL	pseudo augment	CVPR 2020	√
Pseudo-LiDAR	pseudo augment	CVPR 2019	√
Pseudo-LiDAR++	pseudo augment	ICLR 2020	√
MVP	pseudo augment	NIPS 2021	√

7. 基于transformer的3D检测方法 （Transformer）

方法名称	方法类别	发表年份	代码是否开源
VoTr	Transformer	ICCV 2021	
CT3D	Transformer	ICCV 2021	√
M3DETR	Transformer	Arxiv	
DETR3D	Transformer	CoRL 2021	√
PoinTr	Transformer	ICCV 2021	√

8. 基于半监督学习的3D检测方法（Semi supervised）

方法名称	方法类别	发表年份	代码是否开源
3DAL	Semi supervised	Arxiv	
3DIoUMatch	Semi supervised	CVPR 2021	√
WS3D	Semi supervised	TPAMI 2021	√

二、论文分类解读

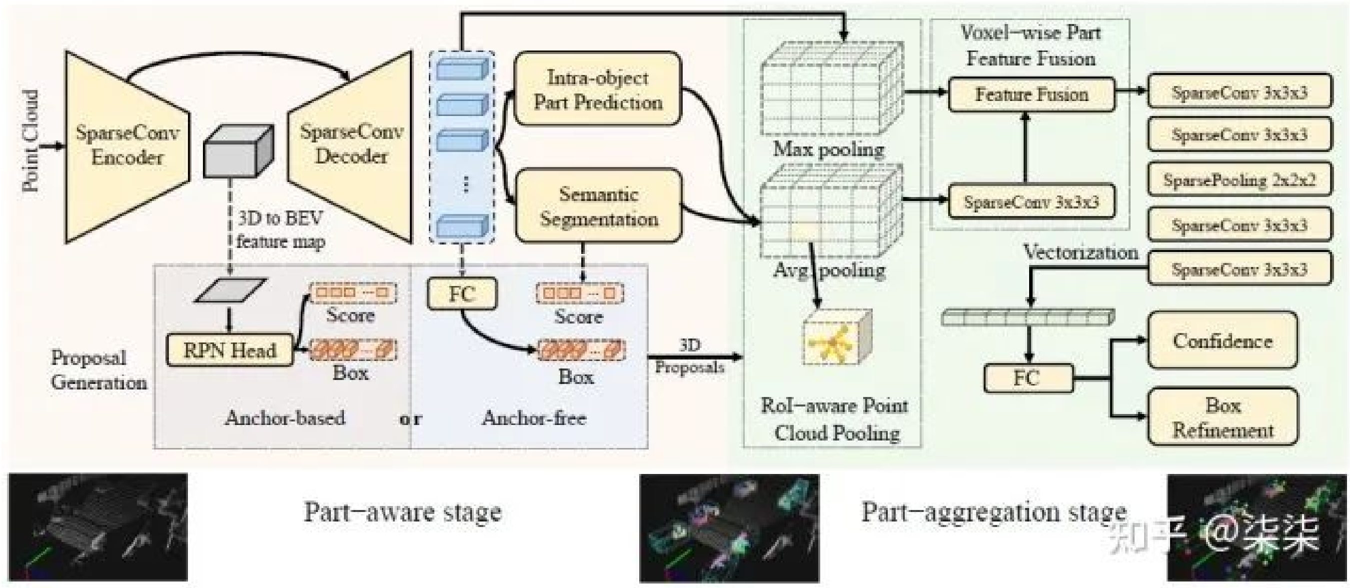
由于篇幅限制，论文总结（一）主要讲解基于激光雷达点云的3D检测方法(LiDAR only)。LiDAR only 指的是此类方法仅仅采用点云数据作为输入，方法的主要区分性在于对点云数据不同的特征提取方式。

1. Part-A^2 （TPAMI 2020）

链接：https://zhuanlan.zhihu.com/p/436797682

> 论文地址：https://arxiv.org/pdf/1907.03670.pdf
> 作者单位：The Chinese University of Hong Kong
> 代码地址：GitHub - open-mmlab/OpenPCDet: OpenPCDet Toolbox for LiDAR-based 3D Object Detection.

一句话读论文：The ground-truth boxes of 3D object detection not only automatically provide accurate segmentation mask because of the fact that 3D objects are naturally separated in 3D scenes, but also imply the relative locations for each foreground 3D point within the ground truth boxes.



网络框架图

Method	Modality	3D Detection (Car)			BEV Detection (Car)			3D Detection (Ped.)			BEV Detection (Ped.)			3D Detection (Cyc.)			BEV Detection (Cyc.)		
		Mod.	Easy	Hard	Mod.	Easy	Hard	Mod.	Easy	Hard	Mod.	Easy	Hard	Mod.	Easy	Hard	Mod.	Easy	Hard
MV3D [1]	RGB + LiDAR	62.35	71.09	55.12	76.90	86.02	68.49	-	-	-	-	-	-	-	-	-	-	-	-
ContFuse [5]	RGB + LiDAR	66.22	82.54	64.04	85.83	88.81	77.33	-	-	-	-	-	-	-	-	-	-	-	-
AVOD-FPN [4]	RGB + LiDAR	71.88	81.94	66.38	83.79	88.53	77.90	42.81	50.80	40.88	51.05	58.75	47.54	52.18	64.00	46.61	57.48	68.09	50.77
F-PointNet [6]	RGB + LiDAR	70.39	81.20	62.19	84.00	88.70	75.33	44.89	51.21	40.23	50.22	58.09	47.20	56.77	71.96	50.39	61.96	75.38	54.68
PC-CNN-V2 [8]	RGB + LiDAR	73.80	84.33	64.83	86.10	88.49	77.26	-	-	-	-	-	-	-	-	-	-	-	-
UberATG-MMF [64]	RGB + LiDAR	76.75	86.81	68.41	87.47	89.49	79.10	-	-	-	-	-	-	-	-	-	-	-	-
VoxelNet [29]	LiDAR only	65.11	77.47	57.73	79.26	89.35	77.39	33.69	39.48	31.51	40.74	46.13	38.11	48.36	61.22	44.37	54.76	66.70	50.55
SECOND [7]	LiDAR only	73.66	83.13	66.20	79.37	88.07	77.95	42.56	51.07	37.29	46.27	55.10	44.76	53.85	70.51	40.90	56.04	73.67	48.78
PointPillars [9]	LiDAR only	74.99	79.05	68.30	**86.10**	88.35	79.83	43.53	52.08	41.49	50.23	58.66	47.19	59.07	75.78	[illegible]	[illegible]	[illegible]	[illegible]
PointRCNN (Ours)	LiDAR only	75.76	**85.94**	68.32	85.68	89.47	79.10	41.78	49.43	38.63	47.53	55.92	44.67	59.60	73.93	[illegible]	[illegible]	[illegible]	[illegible]
Part-A^2-anchor (Ours)	LiDAR only	**77.86**	**85.94**	**72.00**	84.76	**89.52**	**81.47**	**44.50**	**54.49**	**42.36**	**51.12**	**59.72**	**48.04**	**62.73**	**78.58**	**57.74**	**68.12**	**81.91**	**61.92**

KITTI testset 实验结果

整体网络框架分为两个部分：Part-aware stage和Part-aggregation stage。

Part-aware stage：作者认为前景点的相对位置(intra-object part location)可以表征物体的形状信息。因此，通过估计前景点的相对位置，作者认为可以得到更具有辨别性的特征。

The part-aware network aims to extract discriminative features from the point cloud by learning to estimate the intra-object part locations of foreground points, since these part locations implicitly encode the 3D object's shapes by indicating the relative locations of surface points of 3D objects.

Part-aggregation stage：既然是一个aggregation mechanism，作者具体聚合了哪些特征呢？文中作者主要融合了两部分特征，point-wise part location以及 point-wise sementic features。利用融合后的特征，进一步预测每一个候选框的置信度和位置。

By considering the spatial distribution of the predicted intraobject part locations and the learned point-wise part features in a 3D box propsoal from stage-I, it is reasonable to aggregate all the information within a proposal for box proposal scoring and refinement.

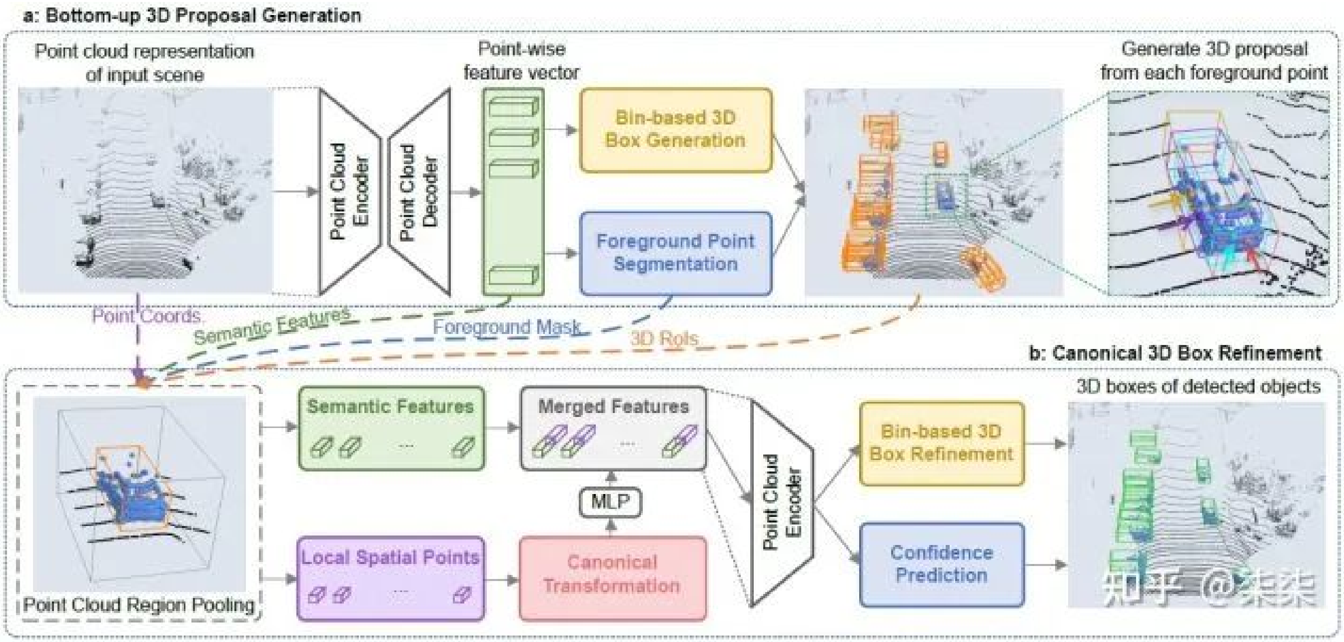
2. Point RCNN （CVPR 2019）

链接：https://zhuanlan.zhihu.com/p/390767889

链接：https://zhuanlan.zhihu.com/p/436419513

> 论文地址：https://arxiv.org/pdf/1812.04244.pdf
>
> 作者单位：The Chinese University of Hong Kong
>
> 代码地址：https://github.com/sshaoshuai/PointRCN
>
> N一句话读论文：The learned point representation from segmentation is not only good at proposal generation but is also helpful for the later box refinement.

备注：感谢微信公众号「**3D视觉工坊**」整理。



网络框架

Method	Modality	Car (IoU=0.7)			Pedestrian (IoU=0.5)			Cyclist (IoU=0.5)		
		Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
MV3D [4]	RGB + LiDAR	71.09	62.35	55.12	-	-	-	-	-	-
UberATG-ContFuse [17]	RGB + LiDAR	82.54	66.22	64.04	-	-	-	-	-	-
AVOD-FPN [14]	RGB + LiDAR	81.94	71.88	66.38	50.80	42.81	**40.88**	64.00	52.18	46.61
F-PointNet [25]	RGB + LiDAR	81.20	70.39	62.19	**51.21**	**44.89**	40.23	71.96	56.77	50.39
VoxelNet [43]	LiDAR	77.47	65.11	57.73	39.48	33.69	31.51	61.22	48.36	44.37
SECOND [40]	LiDAR	83.13	73.66	66.20	51.07	42.56	37.29	[illegible]	[illegible]	46.90
Ours	LiDAR	**85.94**	**75.76**	**68.32**	49.43	41.78	38.63	**73.93**	**59.60**	**53.59**

KITTI testset 实验结果

PointRCNN 整体为two-stage的框架，第一级生成proposal (Bottom-up 3D Proposal Generation)，第二级对proposal进行微调并得到最终的检测结果 (Canonical 3D Box Refinement)。

> **Bottom-up 3D Proposal Generation**：主要目的是做proposal的生成。对每一个point，提取point-wise feature，预测其属于前景点的概率和相应的proposal大小。对于生成的大量的proposal，利用NMS进行过滤，只保留其中的300个送入第二级进行微调。

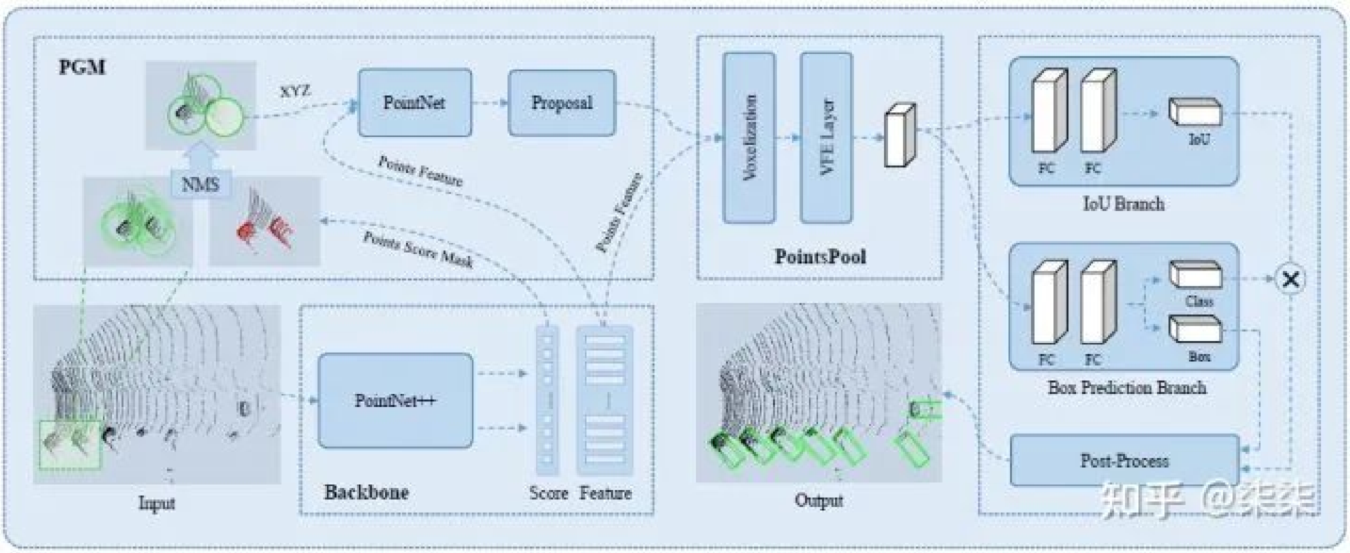
We propose a novel bottom-up point cloud-based 3D bounding box proposal generation algorithm, which generates a small number of high-quality 3D proposals via segmenting the point cloud into foreground objects and background. The learned point representation from segmentation is not only good at proposal generation but is also helpful for the later box refinement.

Canonical 3D Box Refinement：提取第一级proposal更精细的特征用于分类回归。更精细的特征包括：点特征+空间位置特征+RoI特征。

The proposed canonical 3D bounding box refinement takes advantages of our highrecall box proposals generated from stage-1 and learns to predict box coordinates refinements in the canonical coordinates with robust bin-based losses.

3. STD （ICCV 2019）

论文地址：https://arxiv.org/pdf/1907.10471v1.pdf
作者单位：Youtu Lab, Tencent 等
一句话读论文：They propose a point-based proposal generation paradigm on point cloud with spherical anchors.



网络框架

Class	Method	Modality	$AP_{BEV}(\%)$			$AP_{3D}(\%)$		
			Easy	Moderate	Hard	Easy	Moderate	Hard
Car	MV3D [4]	RGB + LiDAR	86.02	76.90	68.49	71.09	62.35	55.12
	AVOD [14]		86.80	85.44	77.73	73.59	65.78	58.38
	F-PointNet [26]		88.70	84.00	75.33	81.20	70.39	62.19
	AVOD-FPN [14]		88.53	83.79	77.90	81.94	71.88	66.38
	RoarNet [31]		88.75	86.08	78.80	83.95	75.79	67.88
	UberATG-MMF [19]		89.49	87.47	79.10	**86.81**	76.75	68.41
	VoxelNet [37]	LiDAR	89.35	79.26	77.39	77.47	65.11	57.73
	SECOND [34]		88.07	79.37	77.95	83.13	73.66	66.20
	PointPillars [15]		88.35	86.10	79.83	79.05	74.99	68.30
	PointRCNN [30]		89.47	85.68	79.10	85.94	75.76	68.32
	Ours		**89.66**	**87.76**	**86.89**	86.61	**77.63**	**76.06**
Pedestrian	AVOD [14]	RGB + LiDAR	42.51	35.24	33.97	38.28	31.51	26.98
	F-PointNet [26]		58.09	50.22	47.20	51.21	**44.89**	40.23
	AVOD-FPN [14]		58.75	51.05	**47.54**	50.80	42.81	40.88
	VoxelNet [37]	LiDAR	46.13	40.74	38.11	39.48	33.69	31.51
	SECOND [34]		55.10	46.27	44.76	51.07	42.56	37.29
	PointPillars [15]		58.66	50.23	47.19	52.08	43.53	41.49
	Ours		**60.99**	**51.39**	45.89	**53.08**	44.24	**41.97**
Cyclist	AVOD [14]	RGB + LiDAR	63.66	47.74	46.55	60.11	44.90	38.80
	F-PointNet [26]		75.38	61.96	54.68	71.96	56.77	50.39
	AVOD-FPN [14]		68.09	57.48	50.77	64.00	52.18	46.61
	VoxelNet [37]	LiDAR	66.70	54.76	50.55	61.22	48.36	44.37
	SECOND [34]		73.67	56.04	48.78	70.51	53.85	46.90
	PointPillars [15]		79.14	62.25	56.00	75.78	59.07	52.92
	Ours		**81.04**	**65.32**	**57.85**	**78.89**	**62.53**	**55.77**

KITTI testset 实验结果

整体框架依然属于two-stage的网络，第一级生成proposal，第二级提取更精细的proposal (point+voxel)特征用于微调。与其他工作相比，STD 网络最大的不同之处在于proposal generation的过程种使用了球形anchor (spherical anchors)。那么如何从球形anchor得到proposal呢？其具体步骤是：

为所有点设定球形anchor → 判断其为前景点的概率 → NMS 过滤冗余anchor → 预测余下的anchor对应的proposal。

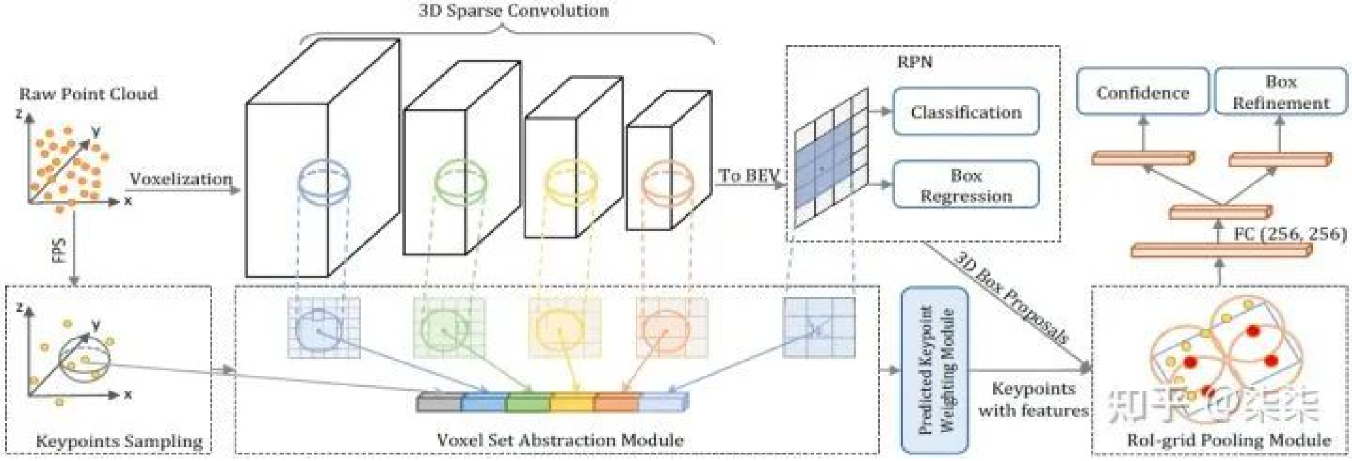
作者认为这种球形anchor的优点在于：

球形anchor可以不必考虑物体朝向问题，极大减小了计算量；

> Considering that a 3D object could be with any orientations, we design spherical anchors rather than traditional cuboid anchors. As a result, the number of spherical anchors is not proportional to the number of pre-defined reference box orientation, leading to about 50% less anchors. With computation much reduced, we surprisingly achieve a much higher recall with spherical anchors than with traditional ones.

4. PV-RCNN/PV-RCNN++（CVPR 2020）

> 论文地址：https://arxiv.org/pdf/2102.00463.pdf
> 作者单位：The Chinese University of Hong Kong
> 代码地址：https://github.com/open-mmlab/OpenPCDet
> 一句话读论文：They propose a novel two-stage detection network for accurate 3D object detection through a two-step strategy of point-voxel feature aggregation.



网络框架

Method	Reference	Modality	Car - 3D Detection Easy	Mod.	Hard	Car - BEV Detection Easy	Mod.	Hard	Cyclist - 3D Detection Easy	Mod.	Hard	Cyclist - BEV Detection Easy	Mod.	Hard
MV3D [9]	CVPR 2017	RGB + LiDAR	74.97	63.63	54.00	86.62	78.93	69.80	-	-	-	-	-	-
ContFuse [13]	ECCV 2018	RGB + LiDAR	83.68	68.78	61.67	94.07	85.35	75.88	-	-	-	-	-	-
AVOD-FPN [11]	IROS 2018	RGB + LiDAR	83.07	71.76	65.73	90.99	84.82	79.62	63.76	50.55	44.93	69.39	57.12	51.09
F-PointNet [20]	CVPR 2018	RGB + LiDAR	82.19	69.79	60.59	91.17	84.67	74.77	72.27	56.12	49.01	77.26	61.37	53.78
UberATG-MMF [17]	CVPR 2019	RGB + LiDAR	88.40	77.43	70.22	93.67	88.21	81.99	-	-	-	-	-	-
3D-CVF at SPA [57]	ECCV 2020	RGB + LiDAR	89.20	80.05	73.11	93.52	89.56	82.45	-	-	-	-	-	-
CLOCs [90]	IROS 2020	RGB + LiDAR	88.94	80.67	77.15	93.05	89.80	86.57	-	-	-	-	-	-
SECOND [16]	Sensors 2018	LiDAR only	83.34	72.55	65.82	89.39	83.77	78.59	71.33	52.08	45.83	76.50	56.05	49.45
PointPillars [15]	CVPR 2019	LiDAR only	82.58	74.31	68.99	90.07	86.56	82.81	77.10	58.65	51.92	79.90	62.73	55.58
PointRCNN [21]	CVPR 2019	LiDAR only	86.96	75.64	70.70	92.13	87.39	82.72	74.96	58.82	52.53	82.56	67.24	60.28
3D IoU Loss [91]	3DV 2019	LiDAR only	86.16	76.50	71.39	91.36	86.22	81.20	-	-	-	-	-	-
STD [23]	ICCV 2019	LiDAR only	87.95	79.71	75.09	94.74	89.19	**86.42**	78.69	61.59	55.30	81.36	67.23	59.35
Part-A2-Net [24]	TPAMI 2020	LiDAR only	87.81	78.49	73.51	91.70	87.79	84.61	-	-	-	-	-	-
3DSSD [67]	CVPR 2020	LiDAR only	88.36	79.57	74.55	92.66	89.02	85.86	**82.48**	64.10	56.90	85.04	67.62	61.14
Point-GNN [92]	CVPR 2020	LiDAR only	88.33	79.47	72.29	93.11	89.17	83.90	78.60	63.48	57.08	[illegible]	[illegible]	[illegible]
PV-RCNN-v1 (Ours)	-	LiDAR only	**90.25**	81.43	76.82	**94.98**	**90.65**	86.14	78.60	63.71	57.65	82.49	68.89	62.41
PV-RCNN-v2 (Ours)	-	LiDAR only	90.14	**81.88**	**77.15**	92.66	88.74	85.97	82.22	**67.33**	**60.04**	**84.60**	**71.86**	**63.84**

KITTI testset 实验结果

整体框架属于two-stage，有两个核心内容，第一个voxel feature → keypoints feature，第二个keypoints feature → proposal/grid feature。

voxel feature → keypoints feature。将大量的voxel feature 整合在少量的keypoints上，整合的过程包括了：原始raw points feature + multi-scale voxel feature + bev feature；

> Our proposed PV-RCNN-v1 first aggregates the voxel-wise scene features at multiple neural layers of 3D voxel CNN into a small number of keypoints, which bridge the 3D voxel CNN feature encoder and the proposal refinement network.

keypoints feature → proposal/grid feature。这一步其实就是利用之前整合的keypoints feature对每一个proposal做RoI Grid Pooling。只是需要额外注意的是，这个的grid 半径是多尺度的，作者认为这种方式可以提取更丰富的proposal feature。

> In this step, we propose keypoint-to-grid RoI feature abstraction to generate accurate proposal-aligned features from the keypoint features for fine-grained proposal refinement.

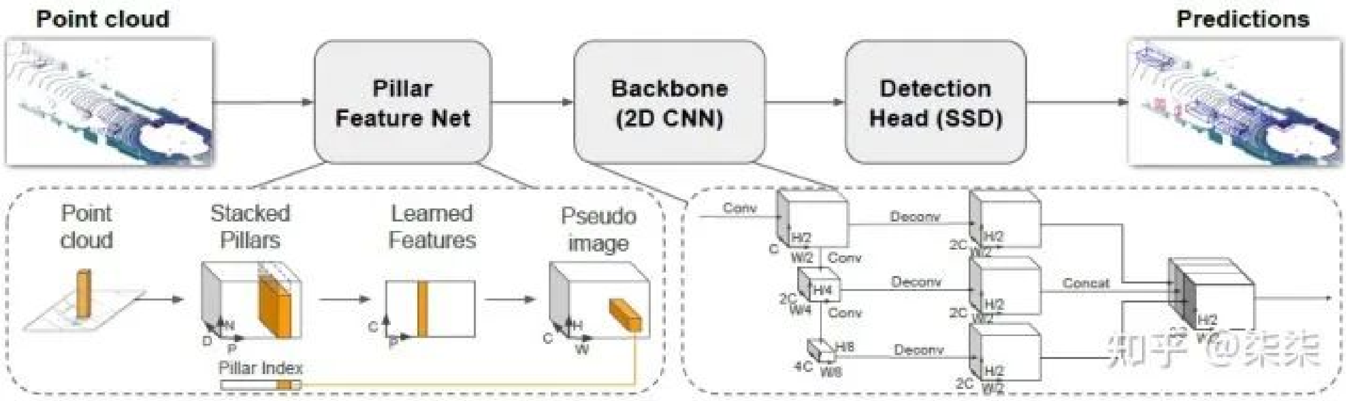
6. PointPillar（CVPR 2019）

论文地址：https://arxiv.org/pdf/1812.05784.pdf

作者单位：Oscar Beijbom and nuTonomy: an APTIV company

一句话读论文：A novel encoder which utilizes PointNets to learn a representation of point clouds organized in vertical columns (pillars).

https://zhuanlan.zhihu.com/p/389034609



网络框架

Method	Modality	Speed (Hz)	mAP	Car			Pedestrian			Cyclist		
			Mod.	Easy	Mod.	Hard	Easy	Mod.	Hard	Easy	Mod.	Hard
MV3D [2]	Lidar & Img.	2.8	N/A	71.09	62.35	55.12	N/A	N/A	N/A	N/A	N/A	N/A
Cont-Fuse [15]	Lidar & Img.	16.7	N/A	82.54	66.22	64.04	N/A	N/A	N/A	N/A	N/A	N/A
Roarnet [25]	Lidar & Img.	10	N/A	**83.71**	73.04	59.16	N/A	N/A	N/A	N/A	N/A	N/A
AVOD-FPN [11]	Lidar & Img.	10	55.62	81.94	71.88	66.38	50.80	42.81	40.88	64.00	52.18	46.61
F-PointNet [21]	Lidar & Img.	5.9	57.35	81.20	70.39	62.19	51.21	**44.89**	40.23	71.96	56.77	50.39
VoxelNet [31]	Lidar	4.4	49.05	77.47	65.11	57.73	39.48	33.69	31.5	61.22	48.36	44.37
SECOND [28]	Lidar	20	56.69	83.13	73.66	66.20	51.07	42.56	37.29	70.51	53.85	46.90
PointPillars	Lidar	**62**	**59.20**	79.05	**74.99**	**68.30**	**52.08**	43.53	**41.49**	[illegible]	**59.07**	**52.92**

Table 2. Results on the KITTI test 3D detection benchmark.

KITTI testset 实验结果

这篇文章的核心内容是**Pillar Feature Network （Pillar点云提取）**。主要作用是将点云数据处理为常见的三维数据，也就是**point clouds → C×H×W**。具体流程为：**point clouds → D×P×N → C×P×N → C×P → C×H×W，其中：**

> D=9，包括每个点坐标**(x, y, z)**，对应的pillar中心点坐标**(x_c, y_c, z_c)**，该点到中心点的偏移**(x_p, y_p)**，以及反射值r。**P=12000**，表示单场景中的pillar数目。**N=100**，表示每个pillar中采样的point数目。

> The points in each pillar are then augmented with xc, yc, zc, xp and yp where the c subscript denotes distance to the arithmetic mean of all points in the pillar and the p subscript denotes the offset from the pillar x; y center.

2. D×P×N → C×P×N 通过**卷积运算**实现。

> Next, we use a simplified version of PointNet where, for each point, a linear layer is applied followed by Batch- Norm and ReLU to generate a C×P×N sized tensor.

3. C×P×N → C×P**是对N维度做max运算**，类似max pooling。

> This is followed by a max operation over the channels to create an output tensor of size C×P.

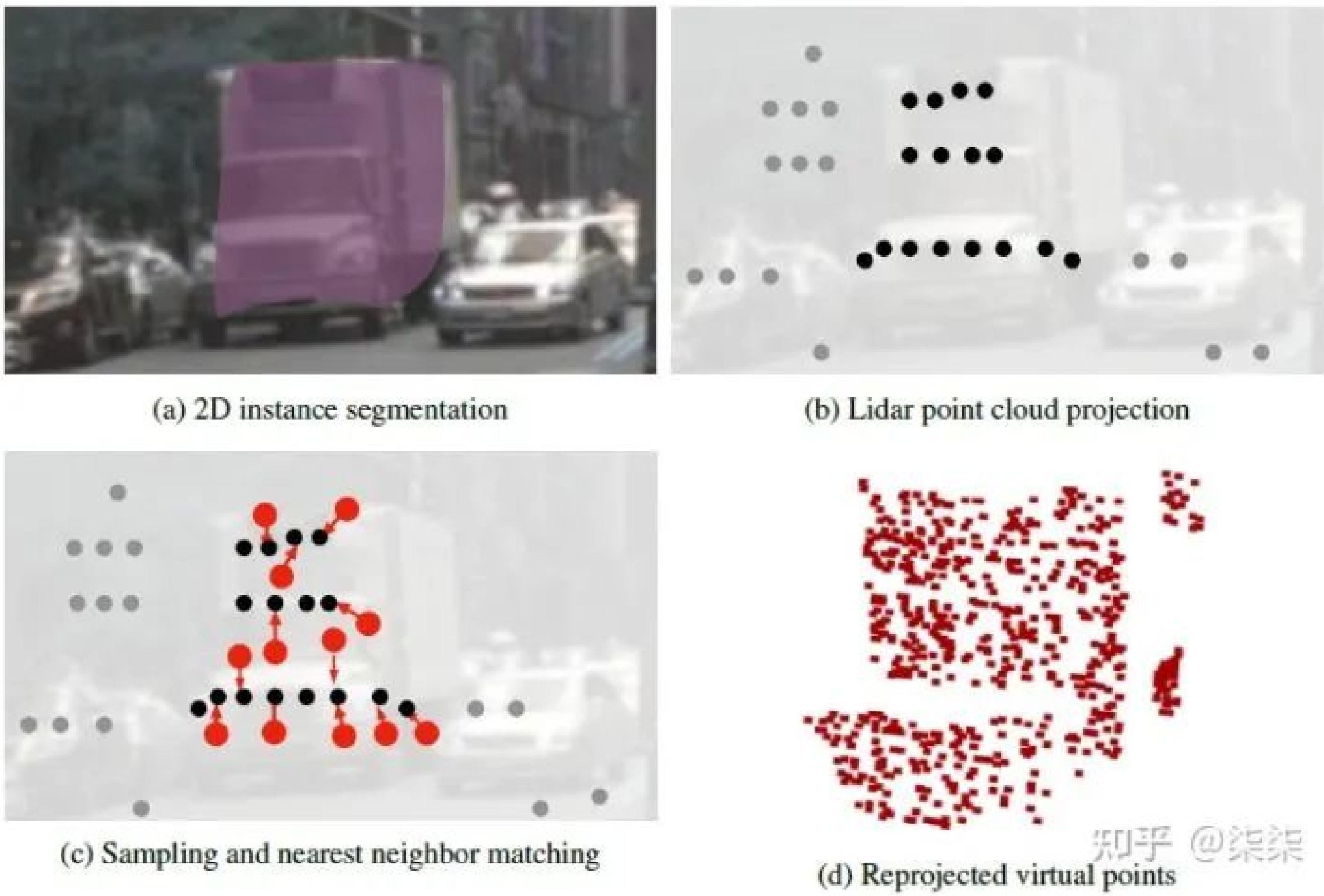
4. C×P → C×H×W是**对P维度的展开**，因为P=12000，表示该场景中的pillar数目，因此作者认为可以一定程度上描述完整场景，所以将P维度直接展开，可以得到我们熟悉的C×H×W三维数据。

> Once encoded, the features are scattered back to the original pillar locations to create a pseudo-image of size C×H×W where H and W indicate the height and width of the canvas.

至此，其实核心部分已经处理完毕，接下来可以用我们熟知的检测方法处理此点云数据。

7. MVP（NIPS 2021）

> 论文地址：https://arxiv.org/pdf/2111.06881.pdf
> 作者单位：https://www.utexas.edu/
> 代码地址：https://tianweiy.github.io/mvp/
> 一句话读论文：The approach takes a set of 2D detections to generate dense 3D virtual points to augment an otherwise sparse 3D point cloud.



网络框架

Method	mAP	NDS	Car	Truck	Bus	Trailer	CV	Ped	Motor	Bicycle	TC	Barrier
PointPillars [23]	30.5	45.3	68.4	23.0	28.2	23.4	4.1	59.7	27.4	1.1	30.8	38.9
WYSIWYG [19]	35.0	41.9	79.1	30.4	46.6	40.1	7.1	65.0	18.2	0.1	28.8	34.7
3DSSD [62]	42.6	56.4	81.2	47.2	61.4	30.5	12.6	70.2	36.0	8.6	31.1	47.9
PMPNet [65]	45.4	53.1	79.7	33.6	47.1	43.1	18.1	76.5	40.7	7.9	58.8	48.8
PointPainting [52]	46.4	58.1	77.9	35.8	36.2	37.3	15.8	73.3	41.5	24.1	62.4	60.2
CBGS [76]	52.8	63.3	81.1	48.5	54.9	42.9	10.5	80.1	51.5	22.3	70.9	65.7
CVCNet [4]	55.3	64.4	82.7	46.1	46.6	49.4	22.6	79.8	59.1	31.4	65.6	69.6
HotSpotNet [5]	59.3	66.0	83.1	50.9	56.4	53.3	23.0	81.3	63.5	36.6	73.0	71.6
CenterPoint [66]	58.0	65.5	84.6	51.0	60.2	53.2	17.5	83.4	53.7	28.7	76.7	70.9
MVP (Ours)	**66.4**	**70.5**	**86.8**	**58.5**	**67.4**	**57.3**	**26.1**	**89.1**	**70.0**	**49.3**	**85.0**	**74.8**

NuScenes testset 实验结果

这篇文章的核心内容是**如何利用2D信息对3D点云进行补全，也就是2D instance segmentation → virtual 3D points**。作者的具体做法是：3D raw points → 2D points → 3D virtual points。

> 3D raw points → 2D points。这一步是将3D物体上的所有点映射到2D图像上，但是映射之后的点仅考虑落在对应物体segmentation区域内的。

> We start by projecting the 3D Lidar point cloud onto our detection. Specifically, we transform each Lidar point into the reference frame of the RGB camera, then project it into image coordinates with associated depth using a perspective projection. The frustum only considers projected 3D points that fall within a detection mask. Any Lidar measurement outside detection masks is discarded.

2. 2D points → 3D virtual points。我们知道，3D到2D的映射属于降维，而2D到3D的映射属于升维。因此，如果没有增加额外的depth信息，从2D映射到3D有无穷多解，这显然是不合适的。作者尝试通过补全每一个2D point的depth信息，将原本“一到无穷”的映射变换为“一到一”的映射。更具体的，对于一个random 2D point，作者将其depth估计为其最邻近的映射3D raw point的depth。

> We start by randomly sampling 2D points from each instance mask. We sample points uniformly at random without repetition. For each sampled point, we retrieve a depth estimate from its nearest neighbor in the frustum.

8. SE-SSD（CVPR 2021）

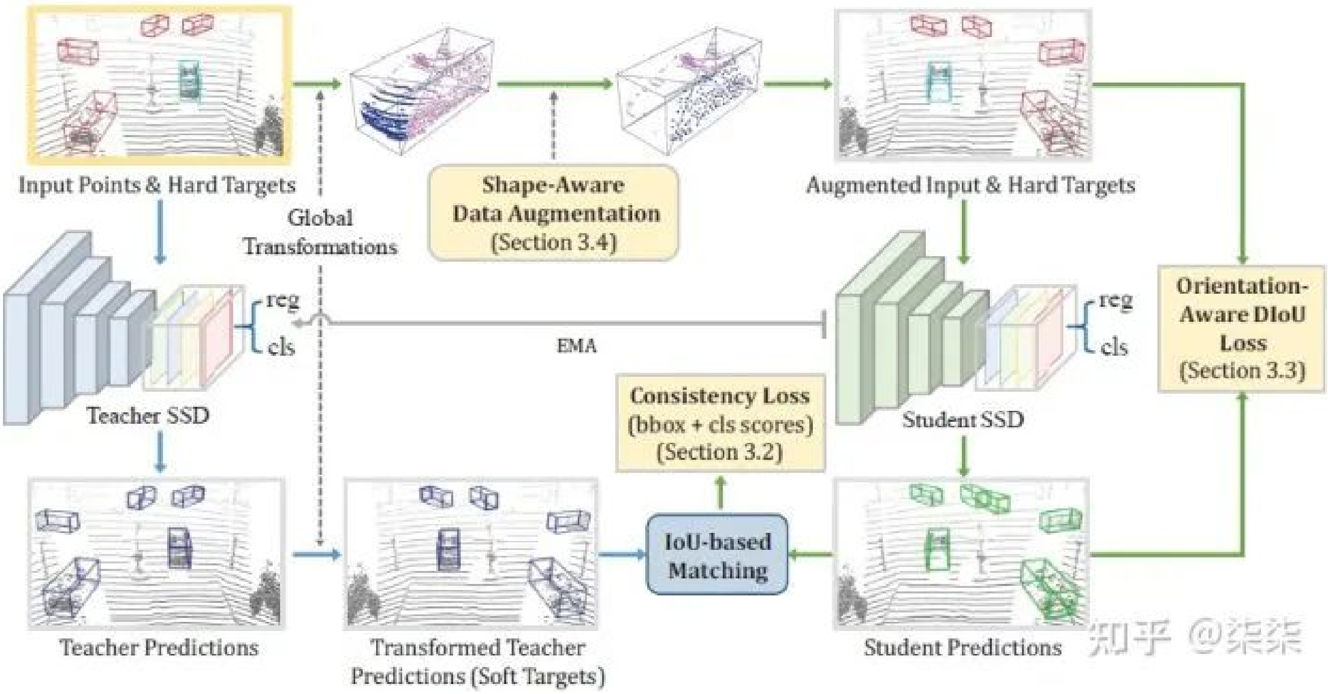
> 论文地址：arxiv.org/pdf/2104.0980
> 作者单位：The Chinese University of Hong Kong
> 代码地址：GitHub - Vegeta2020/SE-SSD: SE-SSD: Self-Ensembling Single-Stage Object Detector From Point Cloud, CVPR 2021.
> 一句话读论文：The key focus is on exploiting both soft and hard targets with formulated constraints to jointly optimize the model, without introducing extra computation in the inference.

https://zhuanlan.zhihu.com/p/371520457

https://zhuanlan.zhihu.com/p/390167950

SE-SSD: Self-Ensembling Single-Stage Object Detector From Point Cloud

Wu Zheng Weiliang Tang Li Jiang Chi-Wing Fu
The Chinese University of Hong Kong
{wuzheng, lijiang, cwfu}@cse.cuhk.edu.hk tangwl123@foxmail.com



网络框架

	Method	Reference	Modality	3D Easy	3D Mod	3D Hard	3D mAP	BEV Easy	BEV Mod	BEV Hard	BEV mAP	Time (ms)
Two-stage	MV3D [3]	CVPR 2017	LiDAR+RGB	74.97	63.63	54.00	64.2	86.62	78.93	69.80	78.45	360
	F-PointNet [20]	CVPR 2018	LiDAR+RGB	82.19	69.79	60.59	70.86	91.17	84.67	74.77	83.54	170
	AVOD [10]	IROS 2018	LiDAR+RGB	83.07	71.76	65.73	73.52	89.75	84.95	78.32	84.34	100
	PointRCNN [24]	CVPR 2019	LiDAR	86.96	75.64	70.70	77.77	92.13	87.39	82.72	87.41	100
	F-ConvNet [28]	IROS 2019	LiDAR+RGB	87.36	76.39	66.69	76.81	91.51	85.84	76.11	84.49	470*
	3D IoU Loss [37]	3DV 2019	LiDAR	86.16	76.50	71.39	78.02	91.36	86.22	81.20	86.26	80*
	Fast PointRCNN [4]	ICCV 2019	LiDAR	85.29	77.40	70.24	77.64	90.87	87.84	80.52	86.41	65
	UberATG-MMF [12]	CVPR 2019	LiDAR+RGB	88.40	77.43	70.22	78.68	93.67	88.21	81.99	87.96	80
	Part-A^2 [25]	TPAMI 2020	LiDAR	87.81	78.49	73.51	79.94	91.70	87.79	84.61	88.03	80
	STD [32]	ICCV 2019	LiDAR	87.95	79.71	75.09	80.92	94.74	89.19	86.42	90.12	80
	3D-CVF [33]	ECCV 2020	LiDAR+RGB	89.20	80.05	73.11	80.79	93.52	89.56	82.45	88.51	75
	CLOCs PVCas [19]	IROS 2020	LiDAR+RGB	88.94	80.67	77.15	82.25	93.05	89.80	86.57	89.81	100*
	PV-RCNN [23]	CVPR 2020	LiDAR	90.25	81.43	76.82	82.83	94.98	90.65	86.14	90.59	80*
	De-PV-RCNN [1]	ECCVW 2020	LiDAR	88.25	81.46	76.96	82.22	92.42	90.13	85.93	89.49	80*
One-stage	VoxelNet [38]	CVPR 2018	LiDAR	77.82	64.17	57.51	66.5	87.95	78.39	71.29	79.21	220
	ContFuse [13]	ECCV 2018	LiDAR+RGB	83.68	68.78	61.67	71.38	94.07	85.35	75.88	85.1	60
	SECOND [30]	Sensors 2018	LiDAR	83.34	72.55	65.82	73.9	89.39	83.77	78.59	83.92	50
	PointPillars [11]	CVPR 2019	LiDAR	82.58	74.31	68.99	75.29	90.07	86.56	82.81	86.48	**23.6**
	TANet [17]	AAAI 2020	LiDAR	84.39	75.94	68.82	76.38	91.58	86.54	81.19	86.44	34.75
	Associate-3Ddet [5]	CVPR 2020	LiDAR	85.99	77.40	70.53	77.97	91.40	88.09	82.96	87.48	60
	HotSpotNet [2]	ECCV 2020	LiDAR	87.60	78.31	73.34	79.75	94.06	88.09	83.24	88.46	40*
	Point-GNN [26]	CVPR 2020	LiDAR	88.33	79.47	72.29	80.03	93.11	89.17	83.90	88.73	643
	3DSSD [31]	CVPR 2020	LiDAR	88.36	79.57	74.55	80.83	92.66	89.02	85.86	89.18	38
	SA-SSD [8]	CVPR 2020	LiDAR	88.75	79.79	74.16	80.90	95.03	91.03	85.96	90.67	40.1
	CIA-SSD [35]	AAAI 2021	LiDAR	89.59	80.28	72.87	80.91	93.74	89.84	[illegible]	[illegible]	[illegible]
	SE-SSD (ours)	-	LiDAR	**91.49**	**82.54**	**77.15**	**83.73**	**95.68**	**91.84**	**86.72**	**91.41**	30.56

KITTI testset 实验结果

这篇文章的核心其实更偏向于augmentation，从本质上讲，本文中包含的两个核心模块**soft targets生成训练（label space augmentation）**和**Shape-aware data augmentation (target space augmentation)**均属

于augmentation范畴。通过ensemble的形式，可以同时利用soft targets和hard targets的信息。

a）Hence, we exploit both soft and hard targets with our formulated constraints to jointly optimize the model, while incurring no extra inference time.
b）On the other hand, to enable the student SSD to effectively explore a larger data space, we design a new augmentation scheme on top of conventional augmentation strategies to produce augmented object samples in a shape-aware manner.

更具体地，有两个注意点：

第一，Shape-aware Data Augmentation是怎么做的？首先，shape-aware尝试解决的问题是：same samples but with different point representation，用作者文中的话说就是由于遮挡距离等问题的存在，相同物体具有不同的点云表达。

Our insight comes from the observation that the point cloud patterns of ground-truth objects could vary significantly due to occlusions, changes in distance, and diversity of object shapes in practice.

2. 第二，需要注意的是，作者这里是使用了不同loss监督student SSD的学习过程。其中，consistency loss用于监督predicitons和soft targets，orientation-aware DIoU loss用于监督predictions和hard targets。

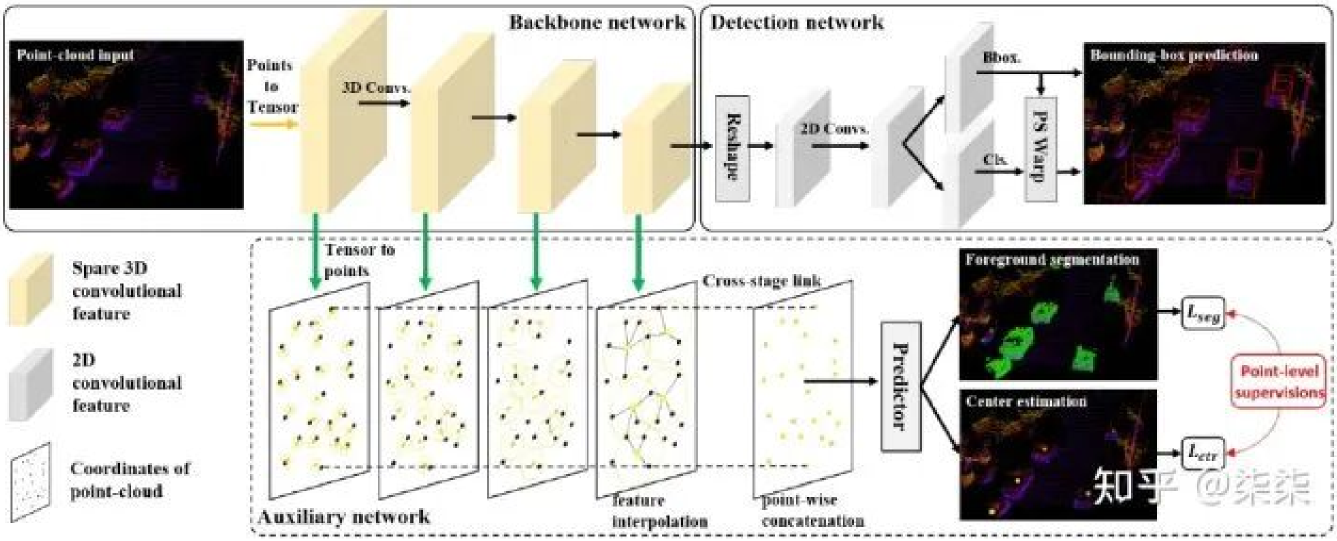
Our consistency loss is to align the student predictions with the soft targets and when we augment the input, we bring along its hard targets to supervise the student with our orientationaware distance-IoU loss.

9. SA-SSD（CVPR 2020）

论文地址：https://www4.comp.polyu.edu.hk/~cslzhang/paper/SA-SSD.pdf
作者单位：The Hong Kong Polytechnic University 等
代码地址：https://github.com/skyhehe123/SA-SSD
一句话读论文：The auxiliary network is introduced, which is jointly optimized by two point-level supervisions, to guide the convolutional features in the backbone network to be aware of the object structure.

https://zhuanlan.zhihu.com/p/378017015

https://zhuanlan.zhihu.com/p/390452834



网络框架

Method	Modality	BEV			3D			FPS
		Easy	Moderate	Hard	Easy	Moderate	Hard	
Two-stage:								
MV3D[1]	LiDAR+RGB	86.49	78.98	72.23	74.97	63.63	54.00	2.8
F-PointNet[16]	LiDAR+RGB	91.17	84.67	74.77	82.19	69.79	60.59	5.9
AVOD[7]	LiDAR+RGB	89.75	84.95	78.32	76.39	66.47	60.23	10
PointRCNN[19]	LiDAR	92.13	87.39	82.72	86.96	75.64	70.70	-
F-ConvNet[23]	LiDAR+RGB	91.51	85.84	76.11	87.36	76.39	66.69	2.1
Fast PointRCNN[2]	LiDAR	90.87	87.84	80.52	85.29	77.40	70.24	15.4
MMF[10]	LiDAR+RGB	93.67	88.21	81.99	88.40	77.43	70.22	12.5
STD[28]	LiDAR	94.74	89.19	**86.42**	87.95	79.71	**75.09**	10
One-stage:								
VoxelNet[30]	LiDAR	87.95	78.39	71.29	77.82	64.17	57.51	4.4
ContFuse[11]	LiDAR+RGB	94.07	85.35	75.88	83.68	68.78	61.67	16.7
SECOND[25]	LiDAR	89.39	83.77	78.59	83.34	72.55	65.82	20
PointPillars[8]	LiDAR	90.07	86.56	82.81	82.58	74.31	68.99	**42**[2]
SA-SSD (ours)	LiDAR	**95.03**	**91.03**	85.96	**88.75**	**79.79**	74.16	25

KITTI testset 实验结果

这篇文章的核心是利用辅助任务（auxiliary network）引导backbone学习到更丰富的结构信息。更具体地，这个辅助任务包含两个部分：前背景点分类（foreground/background classification）和中心点估计（center estimation）。

> We propose a structure-aware single-stage 3D object detector, which employs a detachable auxiliary network to learn structure information and exhibits better localization performance without extra cost.

更具体地，

1.前背景点分类（foreground/background classification），这一步为每一个point输出其前/背景分类得分。这样做的好处是，可以促使backbone学习到更清晰的边界特征。

> Specifically, we employ a sigmoid function to the segmentation branch to predict the foreground/background probability of each point. The segmentation task enables the backbone network to more precisely detect the object boundary.

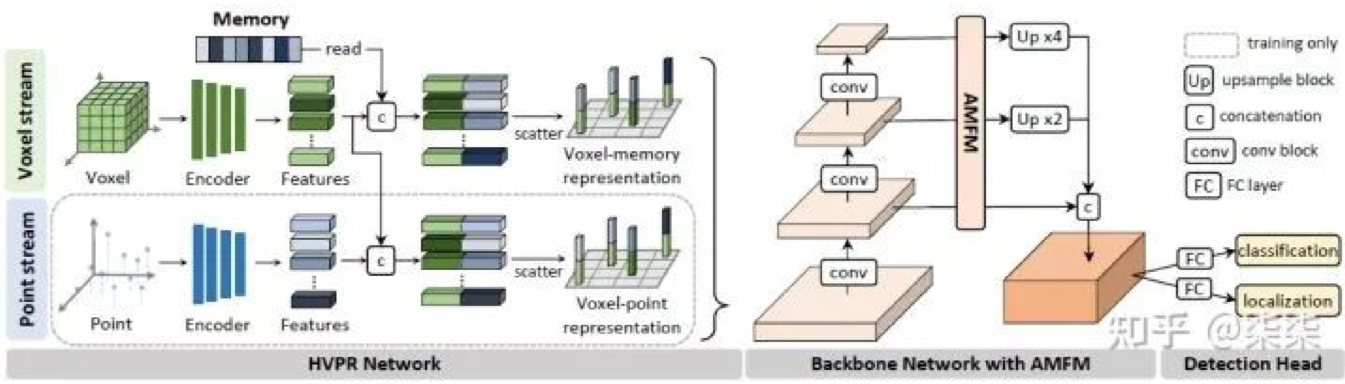
2.中心点估计（center estimation），这一步通过整合intra-object points预测每一个object的中心点。这样做的好处是，可以促使backbone学习到更精准的shape和scale信息。

> To further improve the localization accuracy, we employ another auxiliary task to learn the relative position of each object point to the object center. This intra-object relationship can help determine the scale and shape of the object, resulting in more precise localization.

10. HVPR（CVPR 2021）

> 论文地址：https://arxiv.org/pdf/2104.00902.pdf
> 作者单位：School of Electrical and Electronic Engineering, Yonsei University
> 代码地址：https://cvlab.yonsei.ac.kr/projects/HVPR/
> 一句话读论文：It proposes a novel single-stage 3D detection method having the merit of both voxel-based and point-based features.

https://zhuanlan.zhihu.com/p/373069090



网络框架

Type	Stage	Methods	Reference	Runtime (Hz)	GPU	Car Easy	Car Mod.	Car Hard	Pedestrian Easy	Pedestrian Mod.	Pedestrian Hard
Point	Two	PointRCNN [34]	CVPR 2019	10	TITAN XP	86.96	75.64	70.70	47.98	39.37	36.01
	Two	FastPointRCNN [6]	ICCV 2019	16.7	Tesla P40	85.29	77.40	70.24	-	-	-
	One	F-ConvNet [38]	IROS 2019	2.1	GTX 1080	87.36	76.39	66.69	52.16	43.38	38.80
	Two	STD [45]	ICCV 2019	12.5	TITAN V	87.95	79.71	75.09	53.29	42.47	38.35
	One	3DSSD [44]	CVPR 2020	25	TITAN V	88.36	79.57	74.55	54.64	44.27	40.23
	Two	PV-RCNN [33]	CVPR2020	12.5	GTX 1080Ti	90.25	81.43	76.82	52.17	43.29	40.29
Voxel-3D	One	VoxelNet [50]	CVPR 2018	4.5	TITAN X	77.47	65.11	57.73	39.48	33.69	31.51
	One	3DIoULoss [49]	3DV 2019	12.5	-	86.16	76.50	71.39	-	-	-
	One	SA-SSD [12]	CVPR 2020	25	GTX 2080Ti	88.75	79.79	74.16	-	-	-
	One	HotSpotNet [3]	ECCV 2020	25	-	87.60	78.31	73.34	53.10	45.37	41.47
Voxel-PI	One	SECOND [42]	Sensors 2018	20	GTX 1080Ti	84.78	75.32	68.70	45.31	35.52	33.14
	One	PointPillars [18]	CVPR 2019	**42.4**	GTX 1080Ti	82.58	74.31	68.99	51.45	41.92	38.89
	One	TANet [23]	AAAI 2020	28.5	TITAN V	84.39	75.94	68.82	**53.72**	**44.34**	40.49
	One	Associate-3D [7]	CVPR 2020	20	GTX 1080Ti	85.99	77.40	70.53			
	One	Ours		36.1	GTX 2080Ti	**86.38**	**77.92**	**73.04**	53.47	43.96	**40.64**

KITTI testset 实验结果

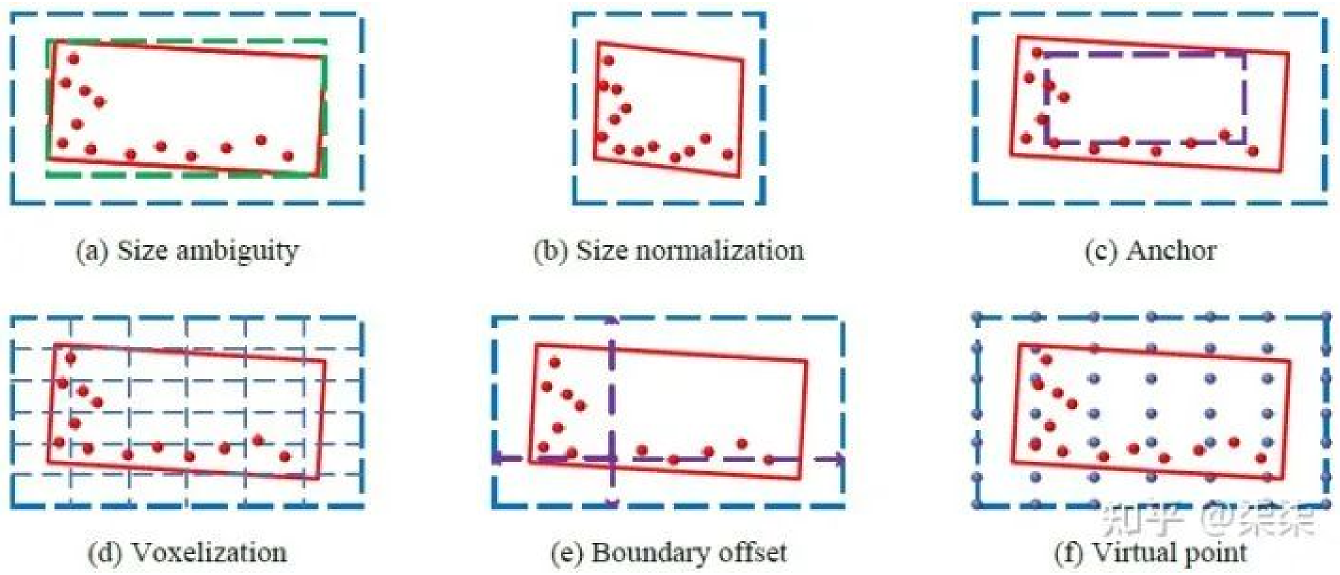
11. LiDAR RCNN（CVPR 2021）

> 论文地址：arxiv.org/pdf/2103.1529
> 作者单位：TuSimple

代码地址：GitHub - TuSimple/LiDAR_RCNN: LiDAR R-CNN: An Efficient and Universal 3D Object Detector
一句话读论文：The authors analyze the size ambiguity problem in detail and propose several methods to remedy it.

https://arxiv.org/pdf/2103.15297.pdf

https://zhuanlan.zhihu.com/p/372199358



网络框架

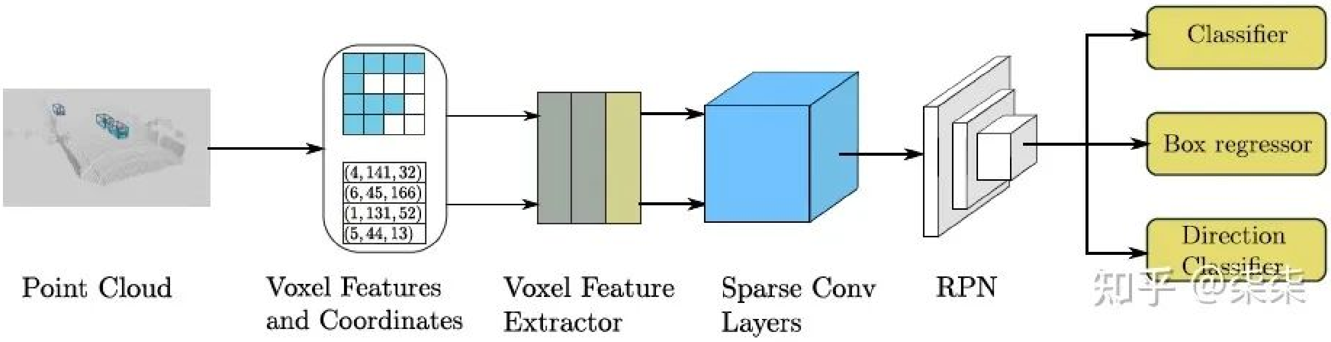
Difficulty	Method	3D AP (IoU=0.7) Overall	0-30m	30-50m	50m-Inf	3D APH (IoU=0.7) Overall	0-30m	30-50m	50m-Inf	BEV AP (IoU=0.7) Overall	0-30m	30-50m	50m-Inf	BEV APH (IoU=0.7) Overall	0-30m	30-50m	50m-Inf
LEVEL_1	StarNet [22]	53.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	PointPillar [12]	56.6	81.0	51.8	27.9	-	-	-	-	75.6	92.1	74.1	55.5	-	-	-	-
	MVF [51]	62.9	86.3	60.0	36.0	-	-	-	-	80.4	93.6	79.2	63.1	-	-	-	-
	Pillar-od [39]	69.8	88.5	66.5	42.9	-	-	-	-	87.1	95.8	84.7	72.1	-	-	-	-
	PV-RCNN [29]	70.3	91.9	69.2	42.2	69.7	91.3	68.5	41.3	83.0	**97.4**	83.0	65.0	82.1	**96.7**	82.0	63.2
	RCD [2]	66.4	86.6	65.6	40.0	66.1	86.3	65.3	39.9	82.1	93.3	80.9	67.2	81.4	92.8	80.2	66.2
	RV first stage	53.4	73.0	49.0	28.1	52.8	72.3	48.4	27.8	70.5	86.2	68.5	51.9	69.5	85.2	67.6	51.0
	LiDAR R-CNN (rv)	68.7	84.8	67.6	47.3	68.2	84.3	67.1	46.6	81.1	90.5	80.5	68.9	80.4	89.9	79.8	67.8
	PointPillars*	72.1	88.3	69.9	48.0	71.5	87.8	69.3	47.3	87.9	96.6	87.1	78.1	87.1	96.0	86.2	76.5
	LiDAR R-CNN (pp)	75.6	92.1	74.3	53.3	75.1	91.6	73.8	52.6	88.2	97.1	87.6	78.3	87.5	96.6	86.9	76.8
	LiDAR R-CNN (2x)	75.6	91.9	74.2	53.5	75.1	91.5	73.6	52.7	90.0	97.0	89.4	78.5	89.2	96.6	88.6	77.0
	LiDAR R-CNN (2xc)	**76.0**	**92.1**	**74.6**	**54.5**	**75.5**	**91.6**	**74.1**	**53.4**	**90.1**	97.0	**89.5**	**78.9**	**89.3**	96.5	**88.6**	**77.4**
LEVEL_2	PV-RCNN [29]	65.4	**91.6**	65.1	36.5	64.8	**91.0**	64.5	35.7	77.5	**94.6**	80.4	55.4	76.6	**94.0**	79.4	53.8
	RV first stage	46.0	70.6	44.1	21.0	45.5	69.8	43.6	20.7	63.5	85.5	62.7	41.1	62.7	84.5	61.9	40.3
	LiDAR R-CNN (rv)	60.1	84.1	61.8	35.7	59.7	83.6	61.3	35.2	74.2	89.8	74.7	54.8	73.5	89.2	74.0	53.9
	PointPillars*	63.6	87.4	62.9	37.2	63.1	86.9	62.3	36.7	81.3	94.0	81.7	65.5	80.4	93.5	80.8	64.1
	LiDAR R-CNN (pp)	66.5	89.1	68.3	40.8	66.1	88.7	67.7	40.3	81.2	94.2	81.8	63.6	80.5	92.7	81.9	62.3
	LiDAR R-CNN (2x)	68.0	91.1	68.1	41.1	67.6	90.7	67.6	40.5	81.6	94.3	82.2	65.6	[illegible]	[illegible]	[illegible]	[illegible]
	LiDAR R-CNN (2xc)	**68.3**	91.3	**68.5**	**42.4**	**67.9**	90.9	**68.0**	**41.8**	**81.7**	94.3	**82.3**	**65.8**	**81.0**	93.9	**81.5**	**64.5**

KITTI testset 实验结果

本文由公众号「3D视觉工坊」整理。

12. SECOND（Sensors 2018）

论文地址：https://www.mdpi.com/1424-8220/18/10/3337
作者单位：Chongqing University 等
代码地址：https://github.com/traveller59/second.pytorch
一句话读论文：It investigates an improved sparse convolution method, which significantly increases the speed of both training and inference.



网络框架

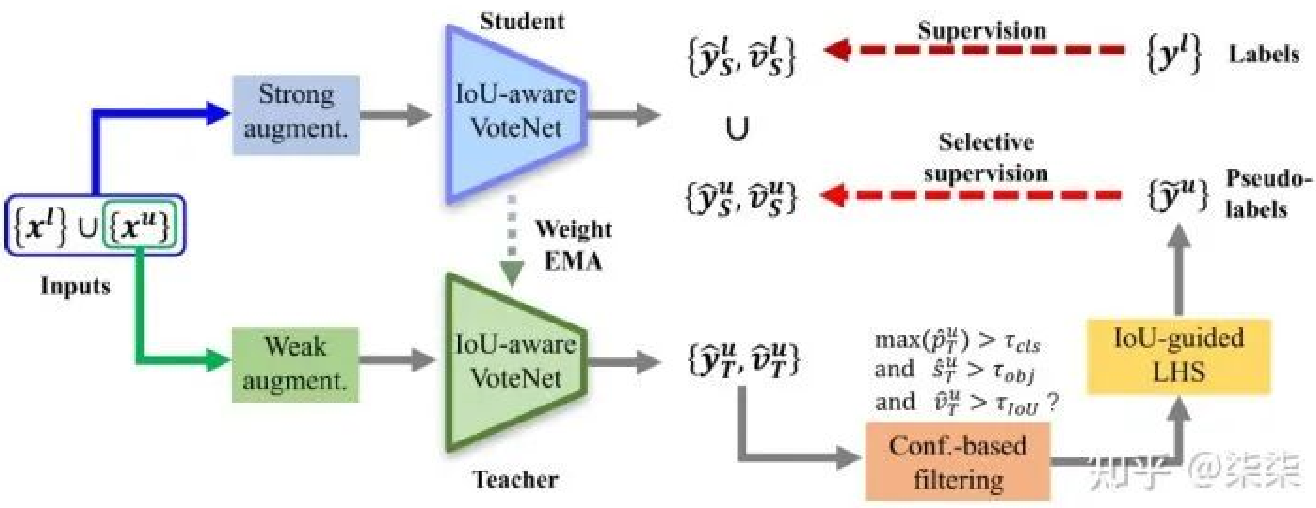
Method	Time (s)	Car			Pedestrian			Cyclist		
		Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
MV3D [8]	0.36	71.09	62.35	55.12	N/A	N/A	N/A	N/A	N/A	N/A
MV3D (LiDAR) [8]	0.24	66.77	52.73	51.31	N/A	N/A	N/A	N/A	N/A	N/A
F-PointNet [11]	0.17	81.20	70.39	62.19	51.21	44.89	40.23	71.96	56.77	50.39
AVOD [9]	0.08	73.59	65.78	58.38	38.28	31.51	26.98	60.11	44.90	38.80
AVOD-FPN [9]	0.1	81.94	71.88	66.38	46.35	39.00	36.58	59.97	46.12	42.36
VoxelNet (LiDAR) [14]	0.23	77.47	65.11	57.73	39.48	33.69	31.51	61.22	48.36	44.37
SECOND	0.05	83.13	73.66	66.20	51.07	42.56	37.29	70.51	53.85	46.90

KITTI testset 实验结果

13. 3DIoUMatch（CVPR 2021）

论文地址：https://arxiv.org/pdf/2012.04355.pdf
作者单位：Stanford University 等
代码地址：https://thu17cyz.github.io/3DIoUMatch/
一句话读论文：It proposes to use the estimated 3D IoU as a localization metric and set category-aware selfadjusted thresholds to filter poorly localized proposals.

https://zhuanlan.zhihu.com/p/390114438



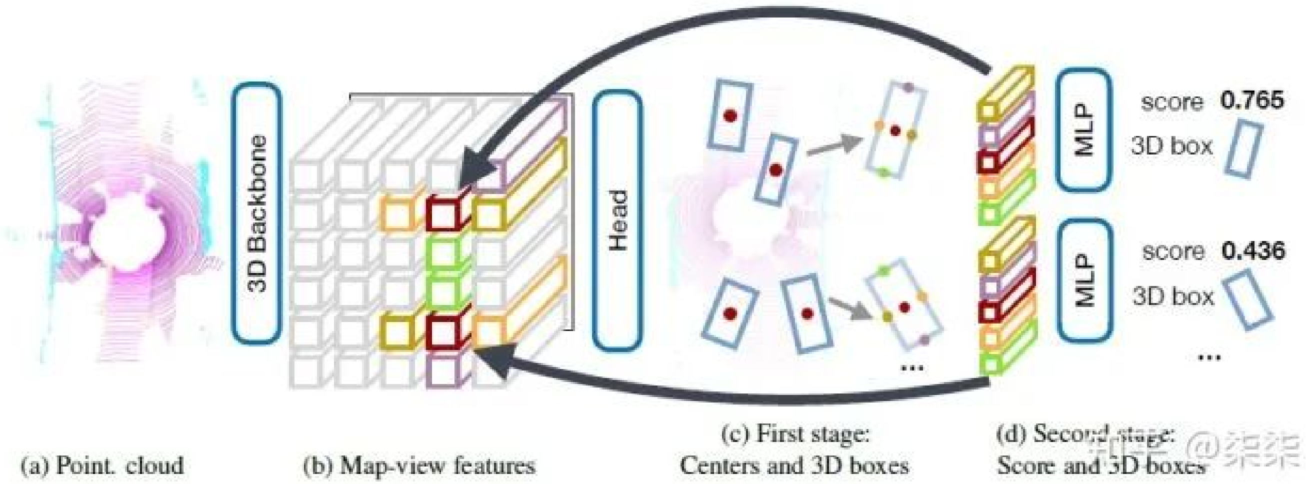
网络框架

Dataset	Model	5%		10%		20%		100%	
		mAP @0.25	mAP @0.5	mAP @0.25	mAP @0.5	mAP @0.25	mAP @0.5	mAP @0.25	mAP @0.5
ScanNet	VoteNet	27.9±0.5	10.8±0.6	36.9±1.6	18.2±1.0	46.9±1.9	27.5±1.2	57.8	36.0
	SESS reported	\	\	39.7±0.9	18.6	47.9±0.4	26.9	62.1	38.8
	SESS	32.0±0.7	14.4±0.7	39.5±1.8	19.8±1.3	49.6±1.1	29.0±1.0	61.3	39.0
	Ours	**40.0±0.9**	**22.5±0.5**	**47.2±0.4**	**28.3±1.5**	**52.8±1.2**	**35.2±1.1**	**62.9**	**42.1**
	Abs. improve.	+8.0	+8.1	+7.7	+8.5	+3.2	+6.2	+1.6	+3.1
SUN-RGBD	VoteNet	29.9±1.5	10.5±0.5	38.9±0.8	17.2±1.3	45.7±0.6	22.5±0.8	58.0	33.4
	SESS reported	\	\	42.9±1.0	14.4	47.9±0.5	20.6	61.1	37.3
	SESS	34.2±2.0	13.1±1.0	42.1±1.1	20.9±0.3	47.1±0.7	24.5±1.2	60.5	38.1
	Ours	**39.0±1.9**	**21.1±1.7**	**45.5±1.5**	**28.8±0.7**	**49.7±0.4**	**30.9±0.2**	**51.5**	**41.3**
	Abs. improve.	+4.8	+8.0	+3.4	+7.9	+2.6	+6.4	+1.0	+3.2

ScanNet & SUN-RGBD 实验结果

14. CenterPoint（CVPR 2021）

论文地址：https://arxiv.org/pdf/2006.11275.pdf
作者单位：https://www.utexas.edu/
代码地址：https://github.com/tianweiy/CenterPoint
一句话读论文：It first detects centers of objects using a keypoint detector and regresses to other attributes, including 3D size, 3D orientation, and velocity.



网络框架

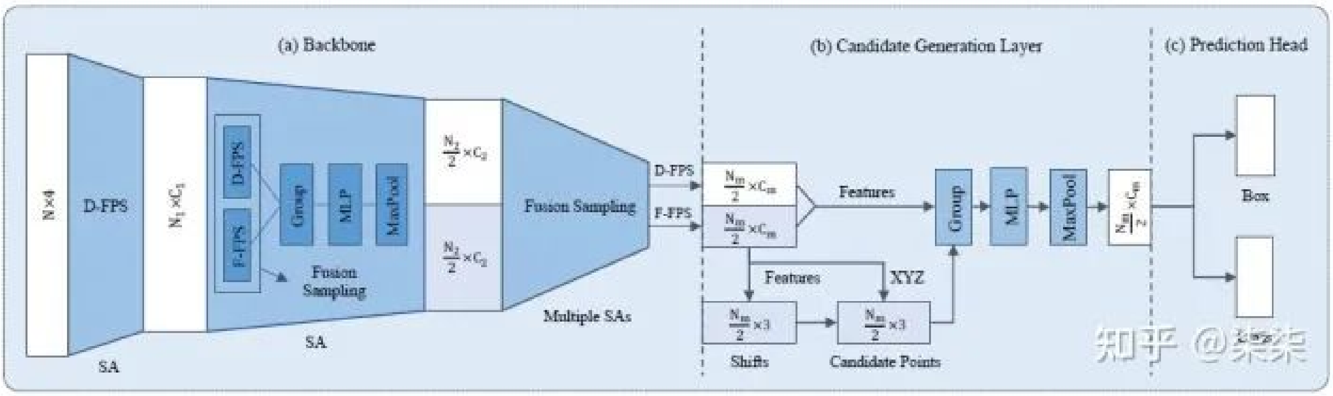
Method	mAP	NDS	Car	Truck	Bus	Trailer	CV	Ped	Motor	Bicycle	TC	Barrier
WYSIWYG [23]	35.0	41.9	79.1	30.4	46.6	40.1	7.1	65.0	18.2	0.1	28.8	34.7
PointPillars [28]	30.5	45.3	68.4	23.0	28.2	23.4	4.1	59.7	27.4	1.1	30.8	38.9
PointPainting [49]	46.4	58.1	77.9	35.8	36.2	37.3	15.8	73.3	41.5	24.1	62.4	60.2
CVCNet [7]	55.3	64.4	82.7	46.1	46.6	49.4	**22.6**	79.8	**59.1**	**31.4**	65.6	69.6
PMPNet [62]	45.4	53.1	79.7	33.6	47.1	43.1	18.1	76.5	40.7	7.9	58.8	48.8
SSN [68]	46.4	58.1	80.7	37.5	39.9	43.9	14.6	72.3	43.7	20.1	54.2	56.3
CBGS [67]	52.8	63.3	81.1	48.5	54.9	42.9	10.5	80.1	51.5	22.3	70.9	65.7
Ours	**58.0**	**65.5**	**84.6**	**51.0**	**60.2**	**53.2**	17.5	**83.4**	53.7	28.7	**76.7**	**70.9**

nuScenes testset 实验结果

15. 3DSSD（CVPR 2021）

论文地址：https://arxiv.org/pdf/2002.10187.pdf
作者单位：The Chinese University of Hong Kong 等
代码地址：https://github.com/dvlab-research/3DSSD
一句话读论文：It proposes a fusion sampling strategy in downsampling process to make detection on less representative points feasible.

https://zhuanlan.zhihu.com/p/380595350



网络框架

Type	Method	Sens.	$AP_{3D}(\%)$ Easy	Mod	Hard
2-stage	MV3D [4]	R + L	71.09	62.35	55.12
	AVOD [12]		76.39	66.47	60.23
	F-PointNet [23]		82.19	69.79	60.59
	AVOD-FPN [12]		83.07	71.76	65.73
	IPOD [34]		80.30	73.04	68.73
	UberATG-MMF [16]		88.40	77.43	70.22
	F-ConvNet [30]		87.36	76.39	66.69
	PointRCNN [26]	L	86.96	75.64	70.70
	Fast Point-RCNN [5]		85.29	77.40	70.24
	Patches [14]		88.67	77.20	71.82
	MMLab-PartA^2 [27]		87.81	78.49	73.51
	STD [35]		87.95	79.71	75.09
1-stage	ContFuse [17]	R + L	83.68	68.78	61.67
	VoxelNet [36]	L	77.82	64.17	57.51
	SECOND [31]		84.65	75.96	68.71
	PointPillars [13]		82.58	74.31	68.99
	Ours		**88.36**	**79.57**	**74.55**

KITTI testset 实验结果

本文仅做学术分享，如有侵权，请联系删文。

3D视觉精品课程推荐：

1.面向自动驾驶领域的多传感器数据融合技术

2.面向自动驾驶领域的3D点云目标检测全栈学习路线！(单模态+多模态/数据+代码)

3.彻底搞透视觉三维重建：原理剖析、代码讲解、及优化改进

4.国内首个面向工业级实战的点云处理课程

5.激光-视觉-IMU-GPS融合SLAM算法梳理和代码讲解

6.彻底搞懂视觉-惯性SLAM：基于VINS-Fusion正式开课啦

7.彻底搞懂基于LOAM框架的3D激光SLAM: 源码剖析到算法优化

8.彻底剖析室内、室外激光SLAM关键算法原理、代码和实战(cartographer+LOAM +LIO-SAM)

9.从零搭建一套结构光3D重建系统[理论+源码+实践]

10.单目深度估计方法：算法梳理与代码实现

11.自动驾驶中的深度学习模型部署实战

12.相机模型与标定(单目+双目+鱼眼)

13.重磅！四旋翼飞行器：算法与实战

14.ROS2从入门到精通：理论与实战

重磅！3DCVer-学术论文写作投稿 交流群已成立

扫码添加小助手微信，可**申请加入3D视觉工坊-学术论文写作与投稿 微信交流群，旨在交流顶会、顶刊、SCI、EI等写作与投稿事宜。**

同时也可申请加入我们的细分方向交流群，目前主要有**3D视觉、CV&深度学习、SLAM、三维重建、点云后处理、自动驾驶、多传感器融合、CV入门、三维测量、VR/AR、3D人脸识别、医疗影像、缺陷检测、行人重识别、目标跟踪、视觉产品落地、视觉竞赛、车牌识别、硬件选型、学术交流、求职交流、ORB-SLAM系列源码交流、深度估计**等微信群。

一定要备注：**研究方向+学校/公司+昵称**，例如：”3D视觉 + 上海交大 + 静静“。请按照格式备注，可快速被通过且邀请进群。**原创投稿**也请联系。

▲长按加微信群或投稿

▲长按关注公众号

3D视觉从入门到精通知识星球：针对3D视觉领域的**视频课程（三维重建系列、三维点云系列、结构光系列、手眼标定、相机标定、激光/视觉SLAM、自动驾驶等）、知识点汇总、入门进阶学习路线、最新paper分享、疑问解答**五个方面进行深耕，更有各类大厂的算法工程人员进行技术指导。与此同时，星球将联合知名企业发布3D视觉相关算法开发岗位以及项目对接信息，打造成集技术与就业为一体的铁杆粉丝聚集区，近4000星球成员为创造更好的AI世界共同进步，知识星球入口：

学习3D视觉核心技术，扫描查看介绍，3天内无条件退款

圈里有高质量教程资料、答疑解惑、助你高效解决问题

觉得有用，麻烦给个赞和在看~