



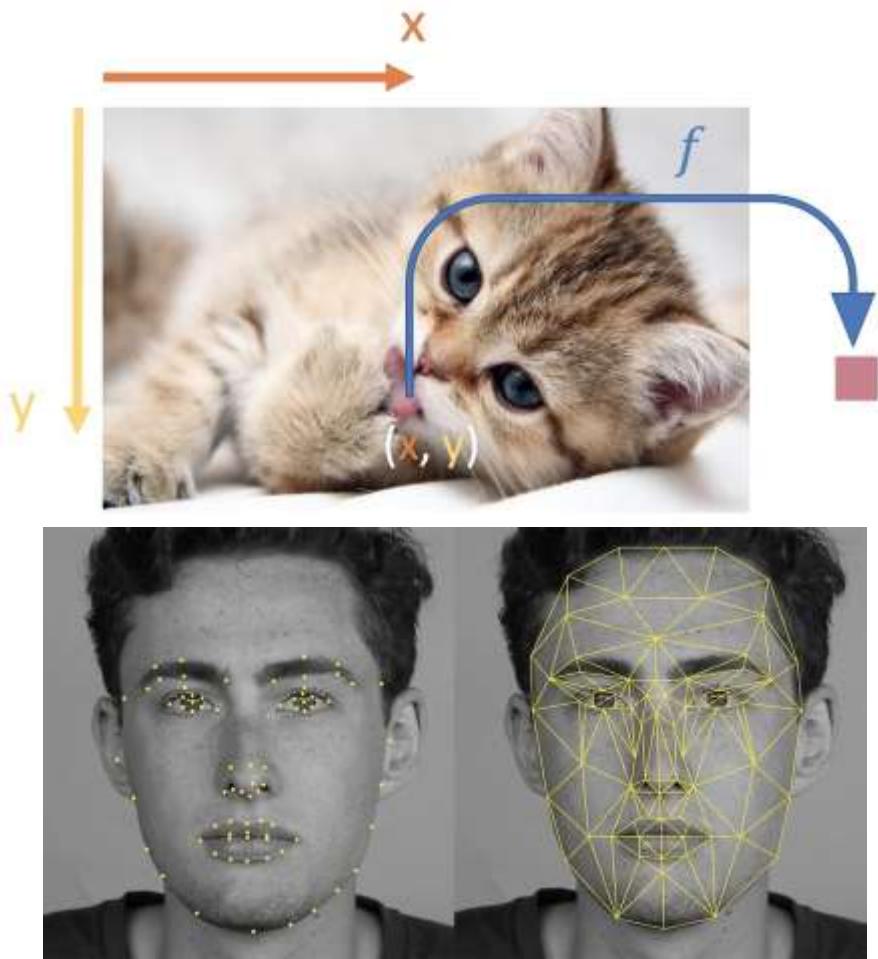
# 机器视觉：算法与应用

同济大学 王志成  
二〇二一年三月



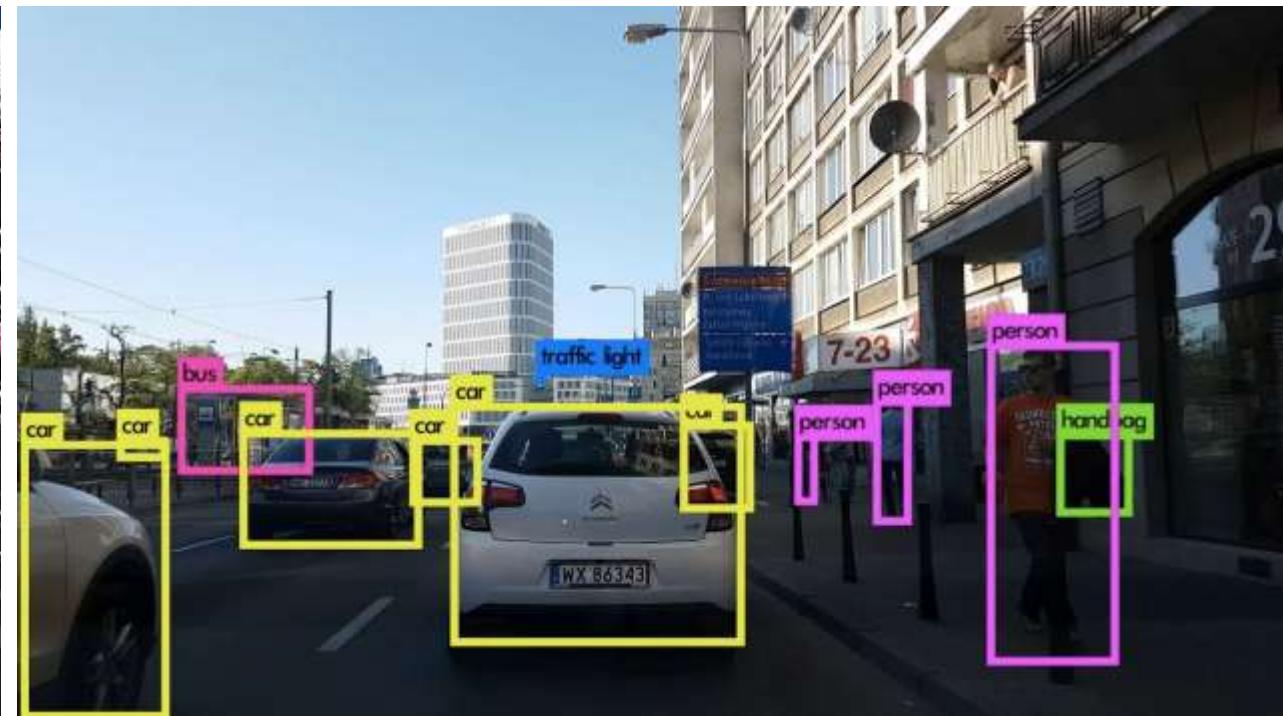
# 机器视觉的典型应用

## ◆ 图像分类识别



# 机器视觉的典型应用

## ◆ 目标检测 (Object Detection)



# 机器视觉的典型应用

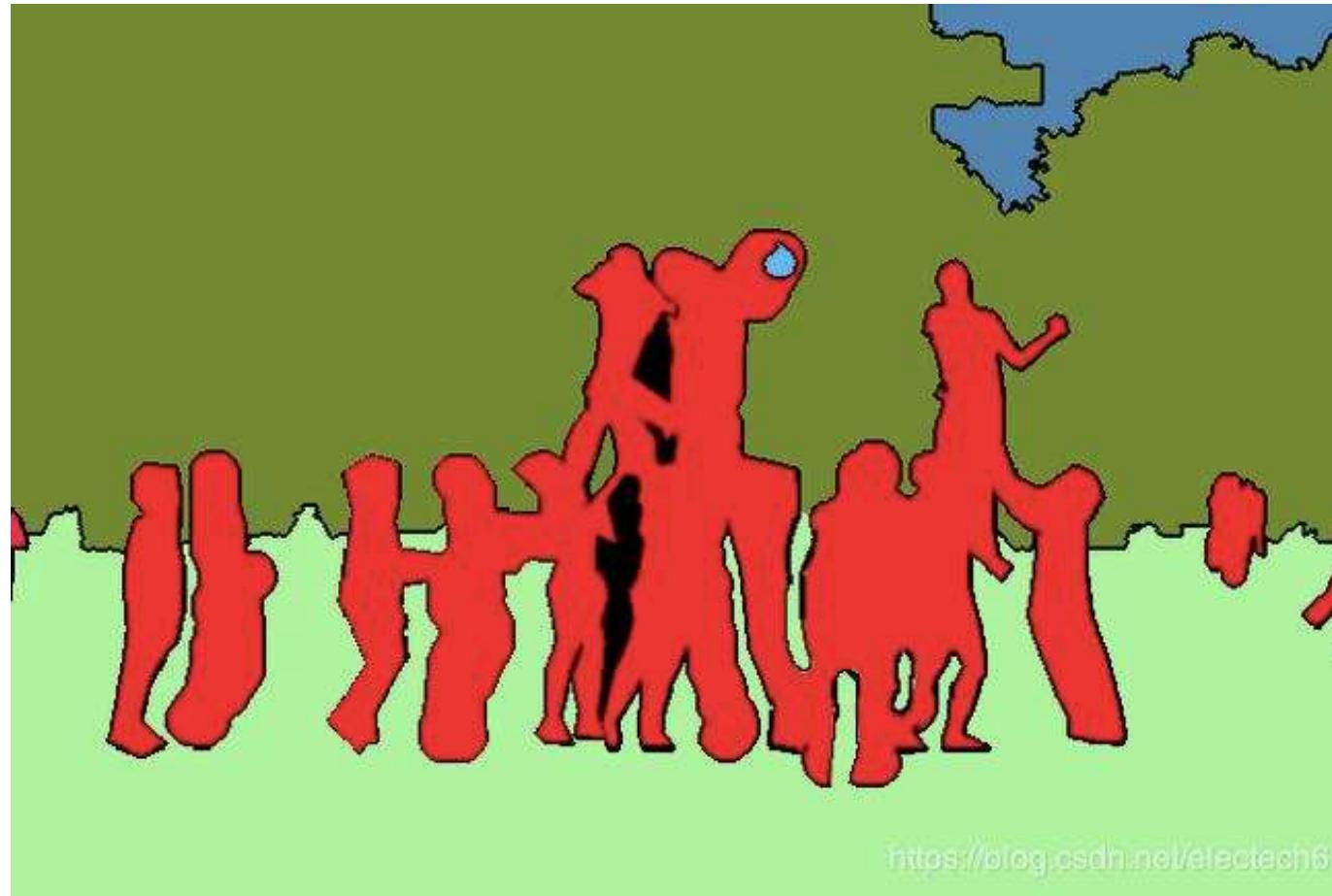
## ◆ 图像分割——超像素 (Superpixels)



- 超像素是一系列像素的集合，这些像素具有类似的颜色、纹理等特征，距离也比较近。
- 左图中每条白线所包括的区域则属于一个 Superpixels。同一个目标可以包括多个 Superpixels。

# 机器视觉的典型应用

## ◆ 图像分割——语义分割 (Semantic Segmentation)



- 给出图像中每个pixel的类别。
- 语义分割只能判断类别，无法判断个体。  
(有两个红色pixel时，无法给出它们是属于同一个人还是不同的人。)

# 机器视觉的典型应用

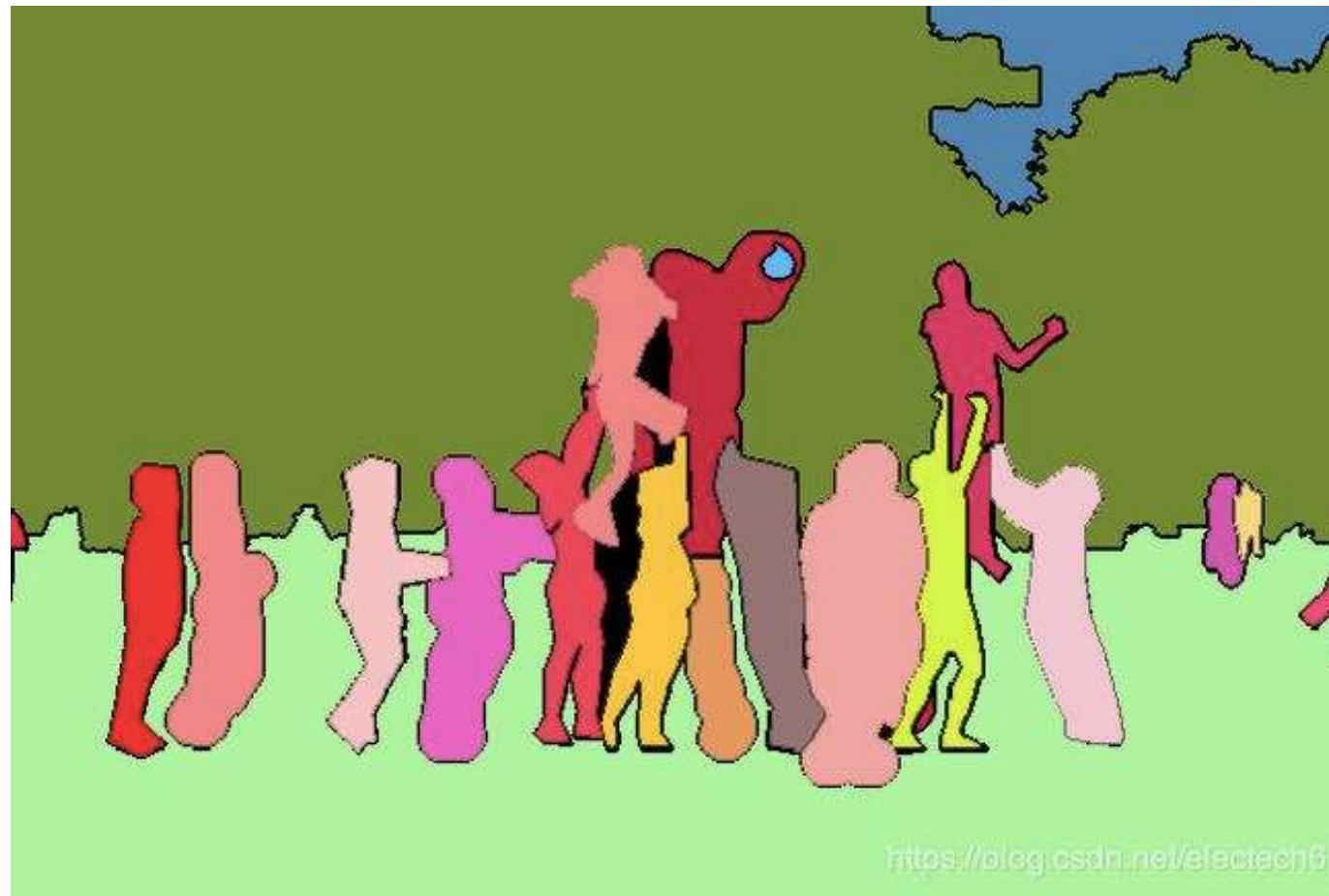
## ◆ 图像分割——实例分割 (Instance Segmentation)



- 实例分割=目标检测+语义分割。
- 相对目标检测的边界框，实例分割可精确到目标的边缘信息。
- 相对语义分割，实例分割需要用不同颜色标注出图像上同一目标的不同个体。

# 机器视觉的典型应用

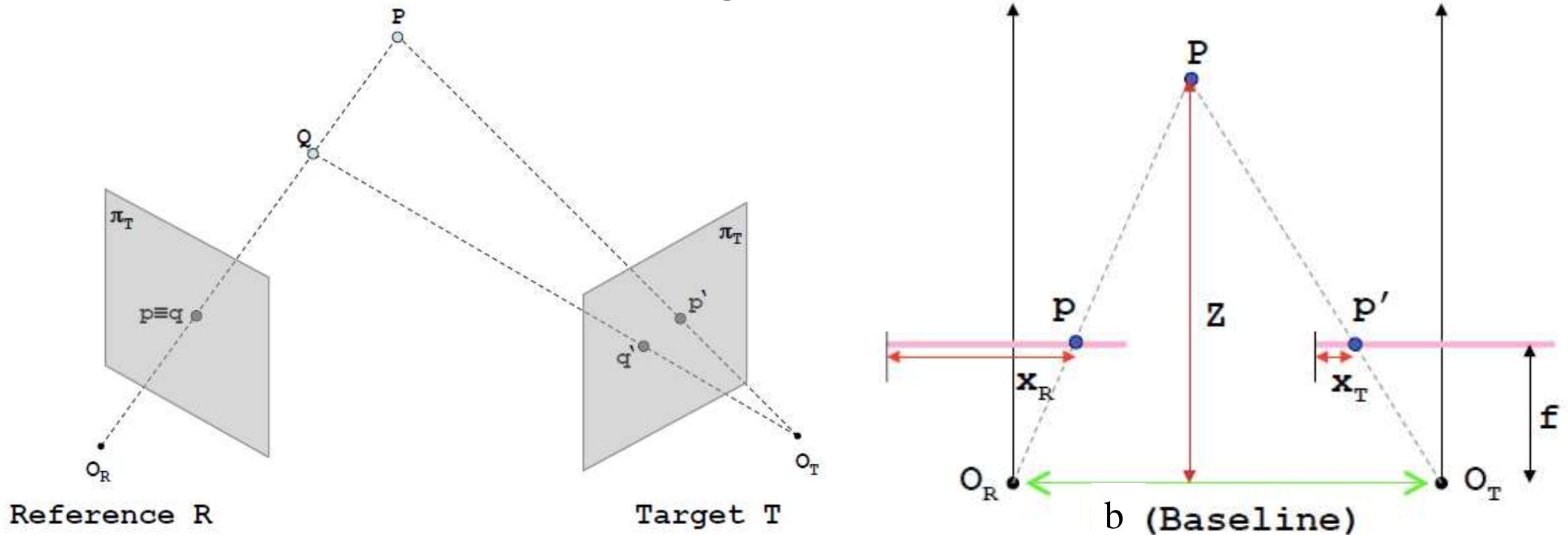
## ◆ 图像分割——全景分割 (Panorama Segmentation)



- 全景分割=语义分割+实例分割。
- 对图像中所有目标和背景都要进行检测和分割。
- 背景区域的分割属于语义分割，而目标的分割属于实例分割。

# 机器视觉的典型应用

## ◆ 立体匹配 (Stereo Matching)



$$\frac{b}{Z} = \frac{(b + x_T) - x_R}{Z - f} \rightarrow Z = \frac{b \cdot f}{x_R - x_T} = \frac{b \cdot f}{d}$$

# 交流内容

## 复杂场景目标检测

红外弱小目标检测  
遥感图像目标检测

01

02

## 单目深度估计

03

04

## 复杂场景语义分割

## 其他场景应用

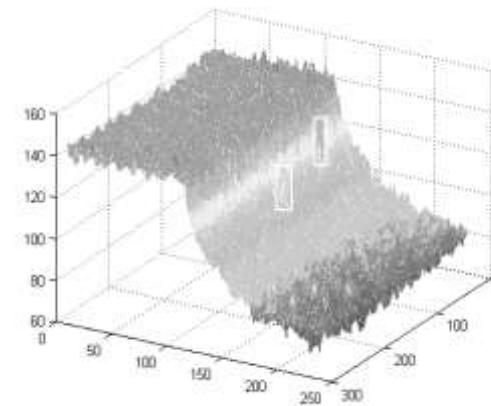
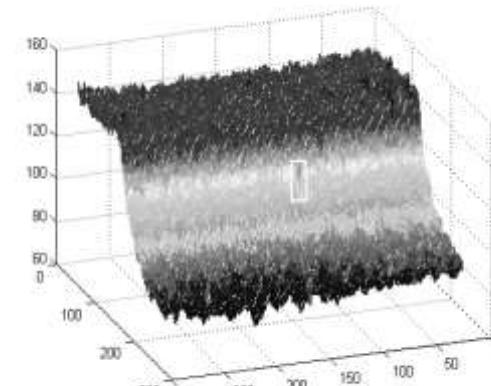
物料箱二维码自动识别  
序列图像自动拼接  
点云语义/实例分割  
.....

面向目标检测、语义分割、视差估计等传统的机器视觉任务，利用深度学习技术，在多个复杂场景中开展有针对性的算法研究与应用。

# 一、复杂场景目标检测

- ◆ 红外弱小目标检测
- ◆ 遥感图像目标检测

# 1 红外弱小目标检测



- 目标无明显形状、尺寸和纹理
- 外部环境变化，缺乏运动先验知识

难点

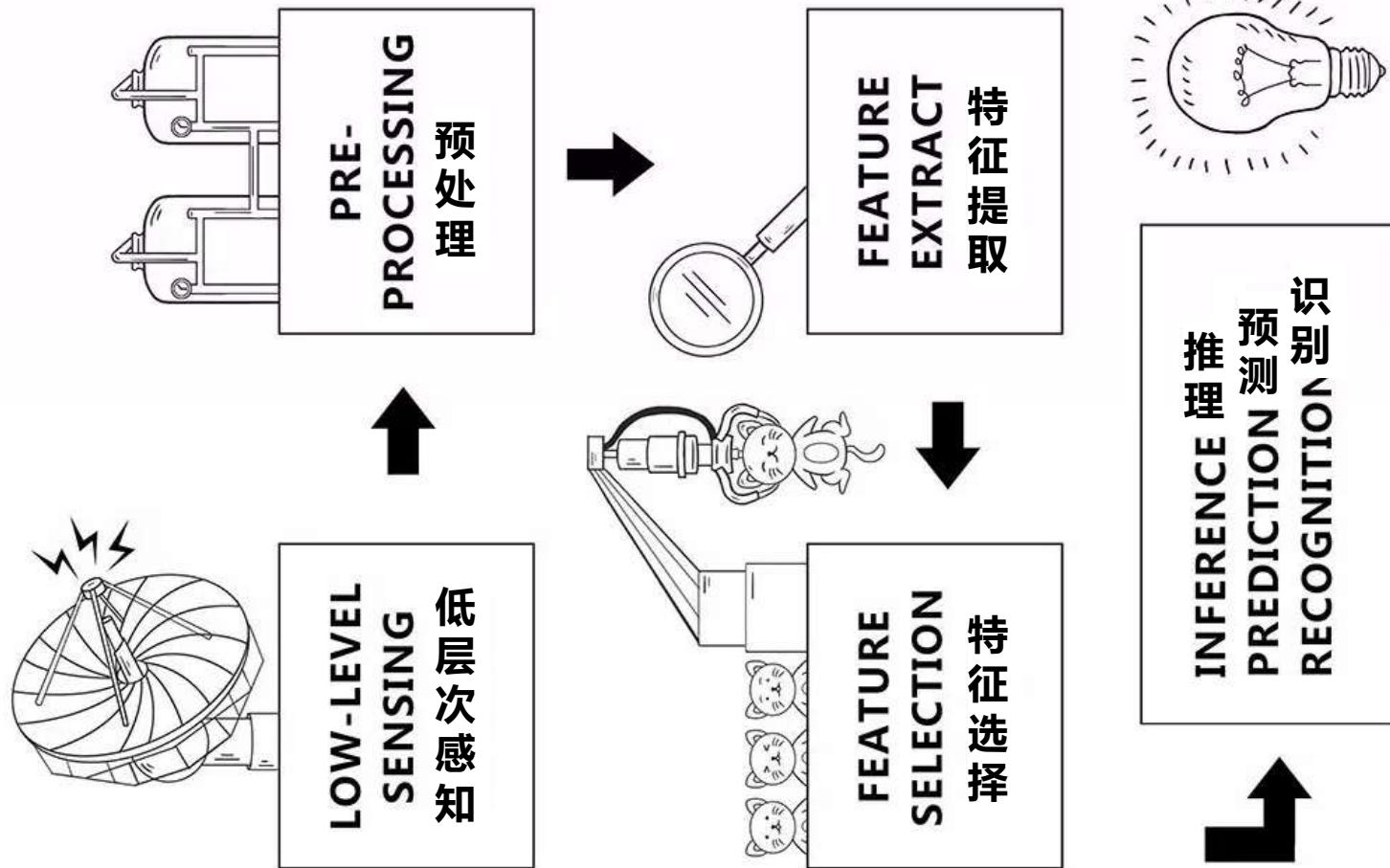
4

2

3

- 样本少且目标面积小，图像信噪比低
- 多帧图像累计能量，数据量大，难实时

# 1 红外弱小目标检测



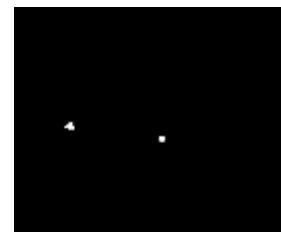
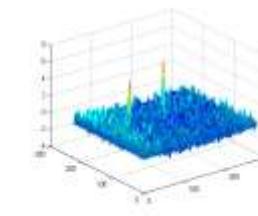
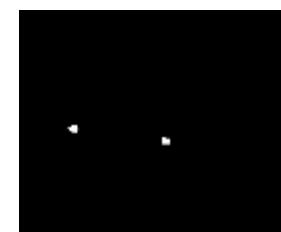
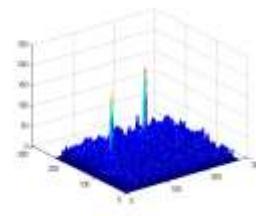
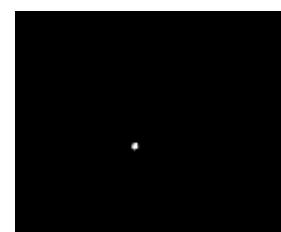
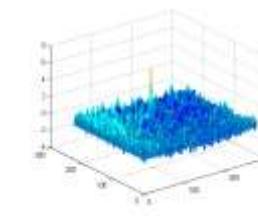
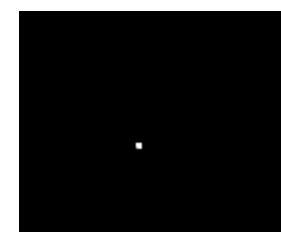
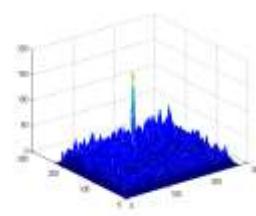
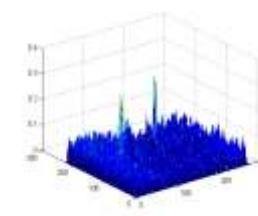
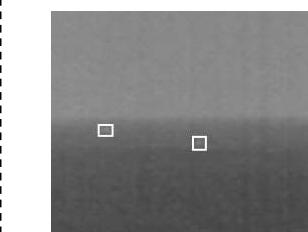
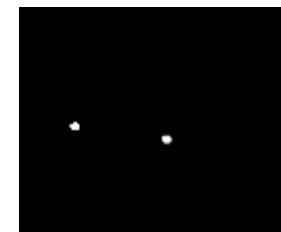
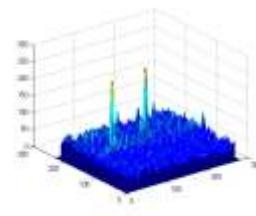
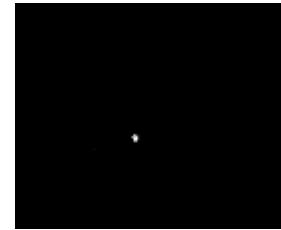
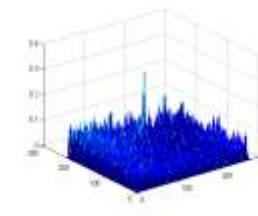
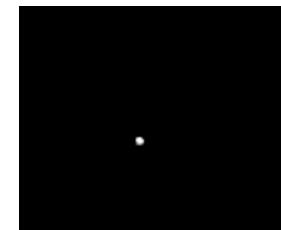
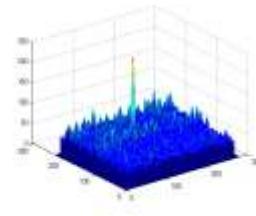
- 特征表达（三步曲）：  
预处理、特征提取、特征选择。
- 特征表达对算法的准确性起着确定性的作用。
- 特征提取：Harris（角点）、SIFT（尺度不变特征变换）、SURF（加速稳健特征）、LBF（局部二值模式）、HOG（方向梯度直方图）、DPM（可形变部件模型）。

# 1 红外弱小目标检测



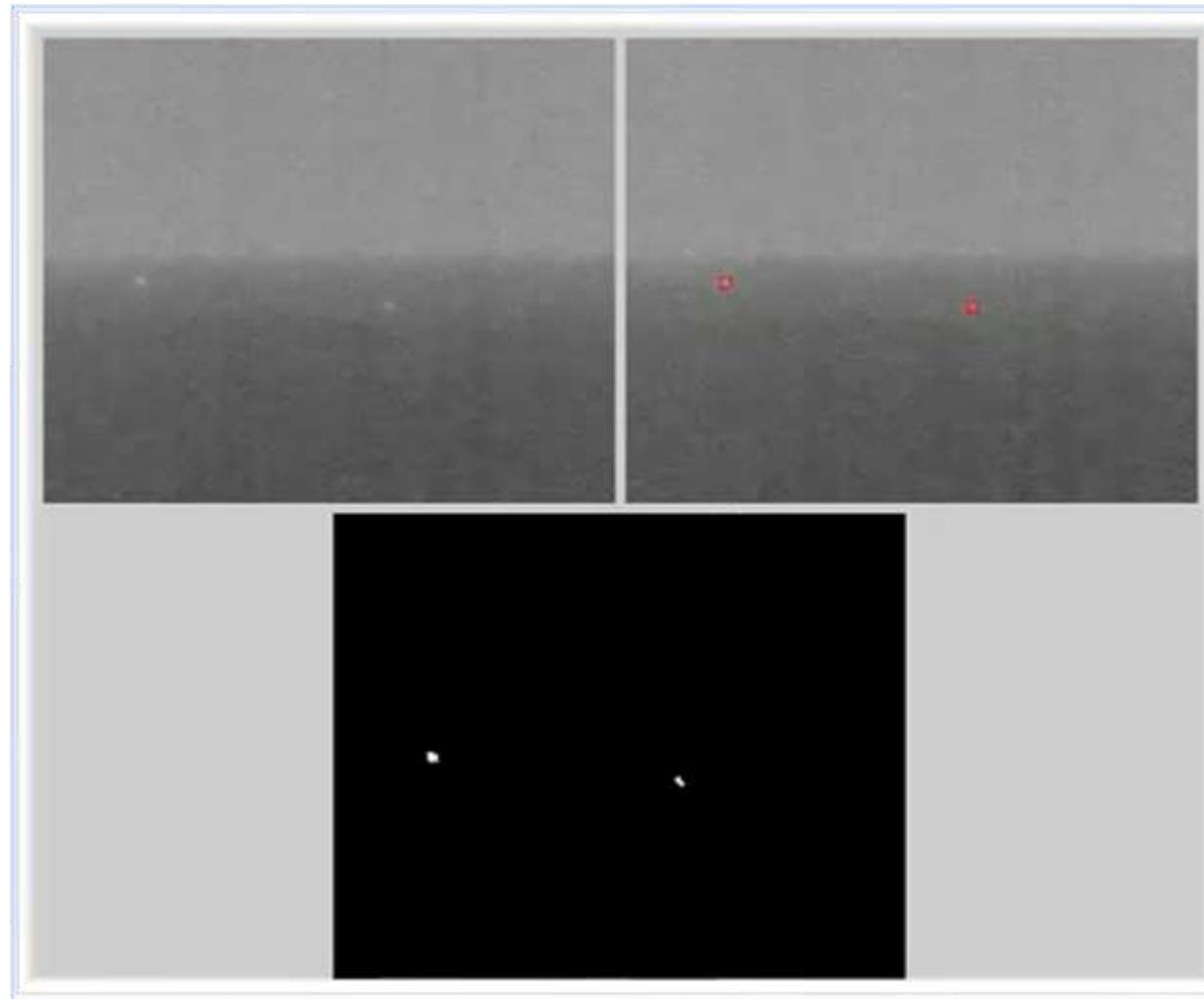
博士论文题目：复杂背景下红外弱小目标融合检测方法研究（2003-2006），研究成果应用于中船集团船舶系统工程部的原型样机。

# 1 红外弱小目标检测



Gabor 融合	小曲面 拟合
小波融 合	多特征 融合

# 1 红外弱小目标检测

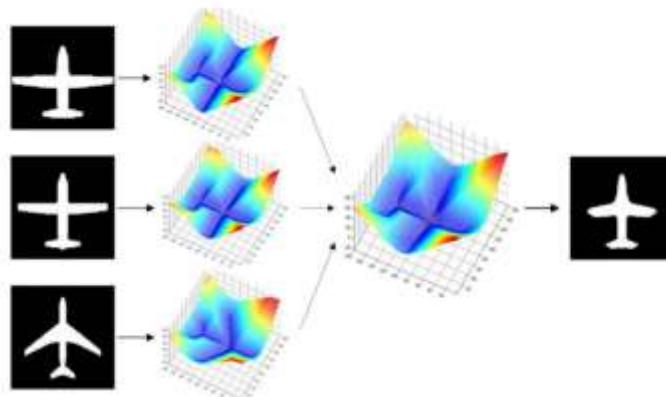
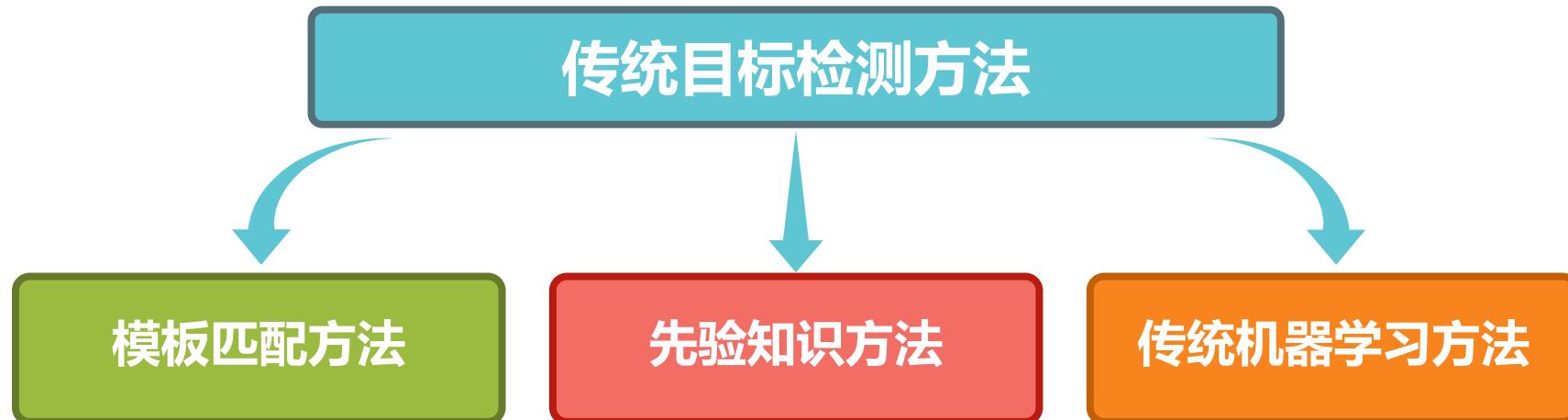


## 2 遥感图像目标检测



遥感图像的目标检测与识别，不仅有助于土地利用、城市规划、环境监测，而且对军事目标识别、战场环境仿真来说有着更为重要的意义。

# 2 遥感图像目标检测



知识库	特征内容	
大型舰船特征 先验知识库	几何特征	面积
		边长
		延伸率
		圆度
		矩形度
		形状因子
	色彩特征	色调
		饱和度
		强度

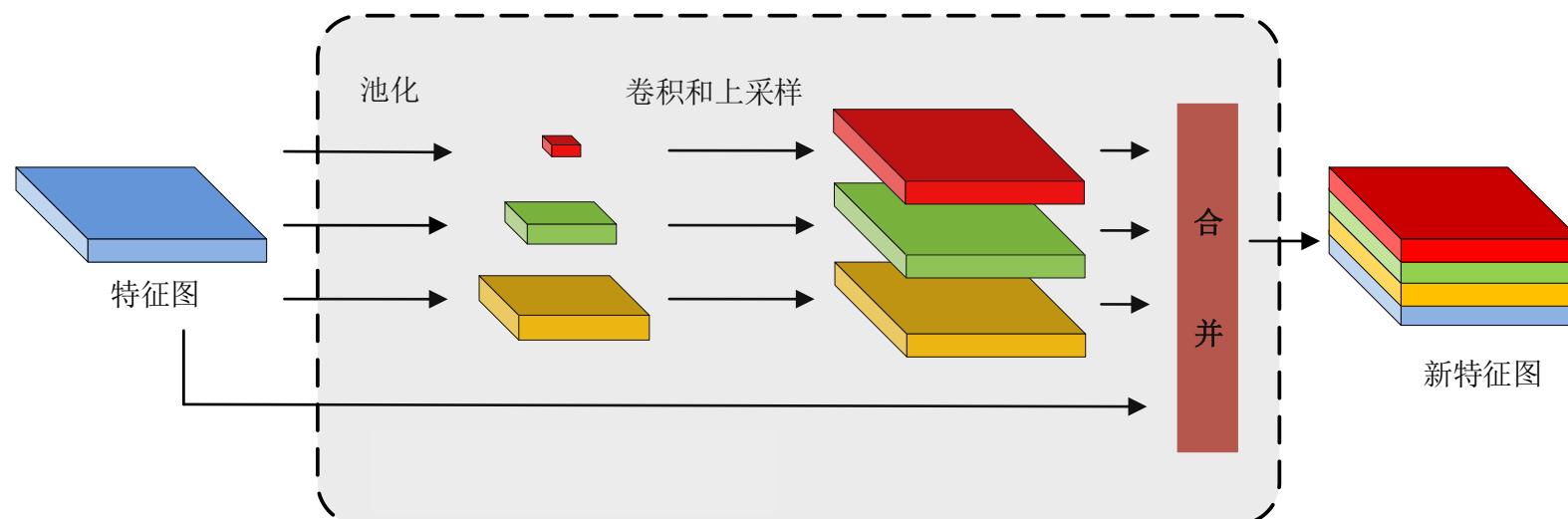
常见特征提取算法	优势	局限性
HOG	有效提取目标的边缘特征，多用于行人检测	无法挖掘深度特征
SIFT	具有尺度不变性，多用于图像匹配	不适用于彩色图片
Haar	适合几何结构边缘、线条明显的目标，多用于人脸检测	不适用于不规则物体

传统目标检测方法依赖于人的经验知识，利用的多是浅层或中层的特征，对复杂背景的适应能力弱。

## 2 遥感图像目标检测

### ● 金字塔池化结构

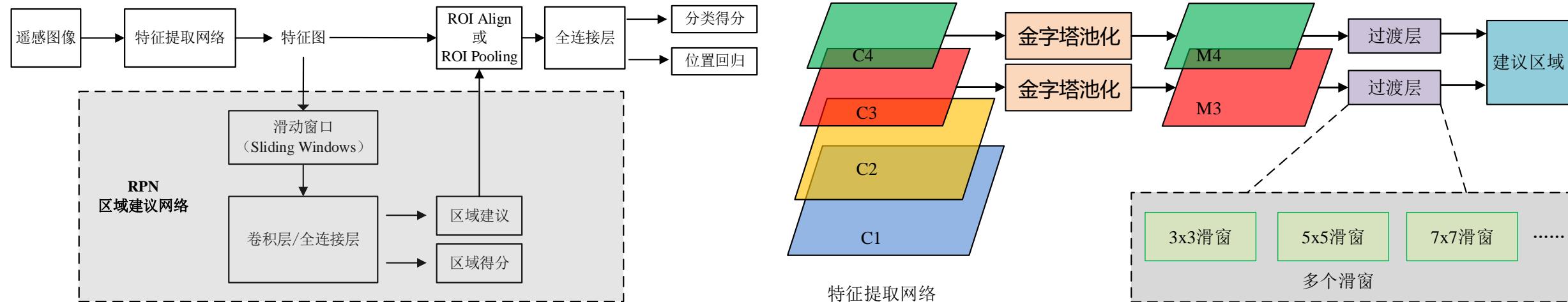
- ✓ 在深度神经网络中，感受野大小极大地影响语义信息的获取。
- ✓ 感受野和上下文信息对目标检测任务至关重要，有限的感受野容易产生虚警和错误的分类。
- ✓ 遥感图像背景复杂，目标尺寸往往较小，如可充分利用目标周围的环境信息，将有利于目标识别。
- ✓ 设计金字塔池化结构，聚合不同尺度上的特征语义信息，扩大有效感受野，获取原有位置附近的环境信息，提供更有辨识力的特征。



# 2 遥感图像目标检测

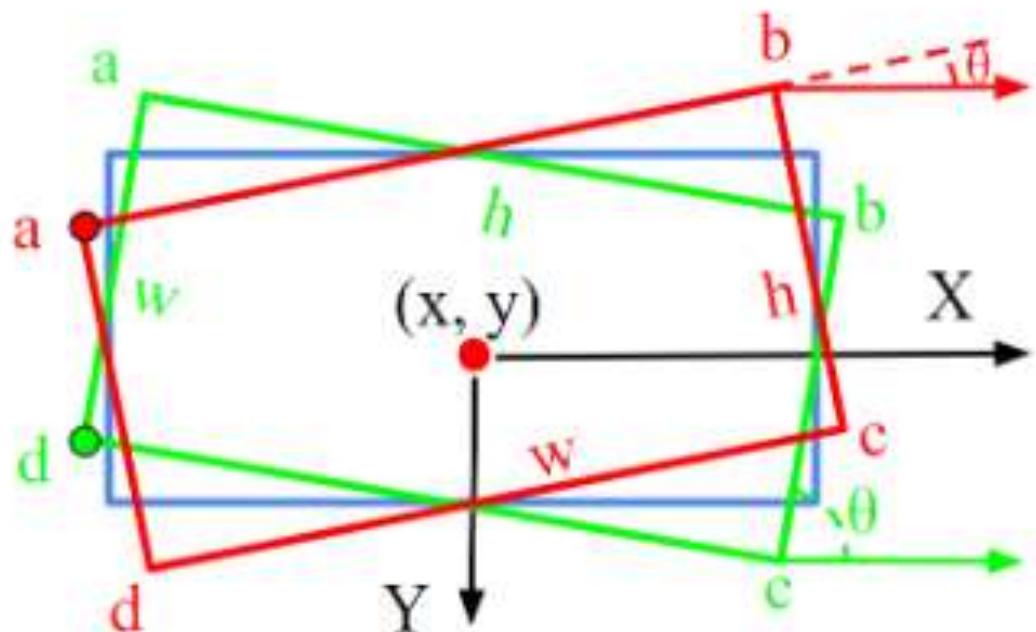
## ● 多尺度区域建议网络

- ✓ 传统的RPN（区域建议网络）作用于C4特征层上，仅用 $3 \times 3$ 卷积进行候选区域生成，有一定的局限性。
- ✓ 对C3、C4特征经过金字塔池化模块处理后，获得更多语义信息和更大感受野的M3、M4特征。
- ✓ 对M3、M4每一个特征层进行候选区域生成时，使用 $3 \times 3$ 卷积、 $5 \times 5$ 卷积和 $7 \times 7$ 卷积，更多大尺寸的卷积核能够提取到更多的语义信息和更大的感受野。
- ✓ 多个滑窗在多个特征图上生成的建议区域，一并进入后续流程，进行建议区域的置信度评价和位置回归，保留高置信度的结果。

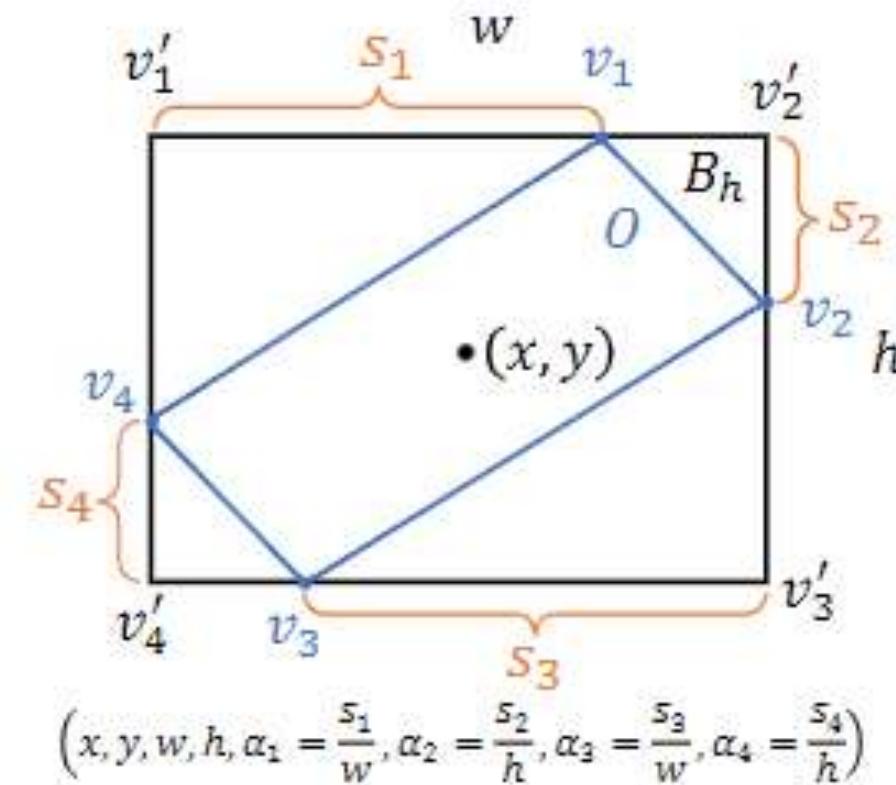


# 2 遥感图像目标检测

- 多方向（倾斜框）目标检测方法



5参数预测旋转角度:  $(x, y, w, h, \theta)$



8参数预测外接矩形内偏移 $(x, y, w, h, s_1, s_2, s_3, s_4)$

# 2 遥感图像目标检测

## ● 实验结果

- ✓ PL-飞机, BD-网球内场, BR-大桥, GTF-田径场, SV-小型车辆, LV-大型车辆, SH-船只, TC-网球场, BC-篮球场, ST-储油罐, SBF-英式足球场, RA-环行路, HA-海港, SP-游泳池, HC-直升飞机。
- ✓ 旋转Faster R-CNN的基准算法mAP为60.27, 分别加入金字塔池化结构、多尺度区域建议网络后, 整个模型性能提升显著。

表1 模型在旋转框任务的AP结果与已有方法对比

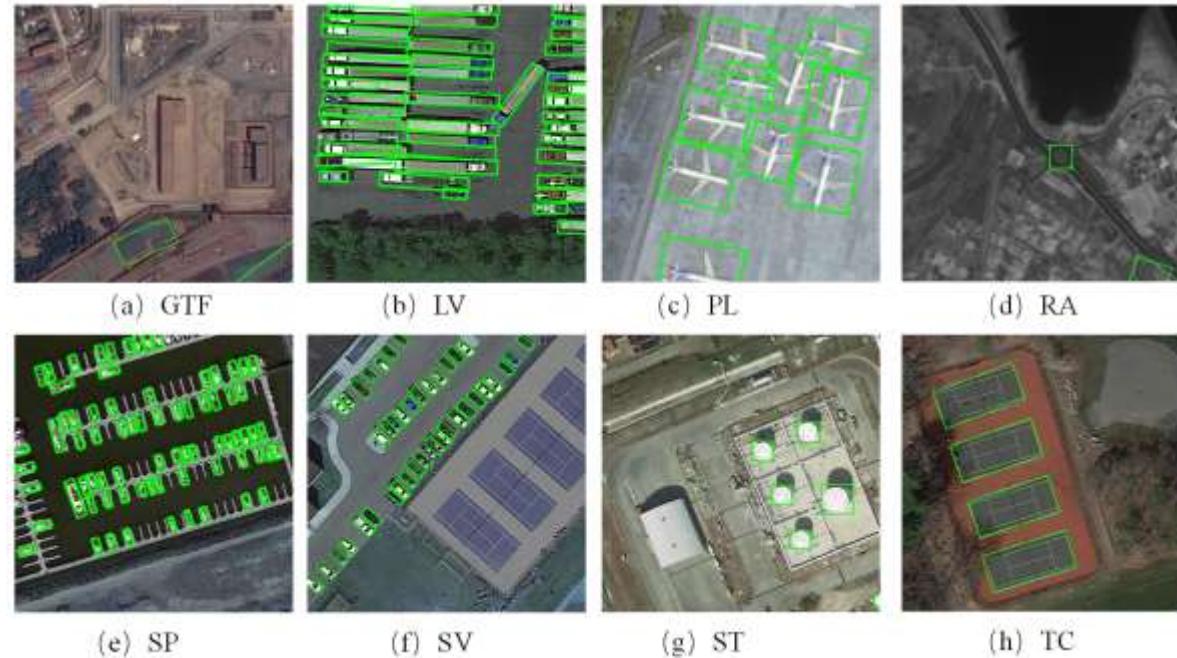
方法	mAp	PL	BD	BR	GTF	SV	LV	SH	TC
Faster R-CNN	54.13	79.42	77.13	17.7	64.05	35.3	38.02	37.16	89.41
R2CNN	60.97	80.94	65.67	35.34	67.44	59.52	50.91	55.81	90.67
Ours	<b>65.94</b>	<b>82.21</b>	<b>77.97</b>	<b>36.1</b>	<b>72.4</b>	<b>64.1</b>	<b>63.2</b>	<b>64.31</b>	<b>91.36</b>
		BC	ST	SBF	RA	HA	SP	HC	
		69.64	59.28	50.3	52.91	47.89	47.4	46.3	
		66.92	72.39	55.06	52.23	55.14	53.35	48.22	
		<b>71.2</b>	<b>74.31</b>	<b>57.29</b>	<b>56.96</b>	<b>62.54</b>	<b>62.1</b>	<b>53.1</b>	

表2 不同改进设置下的实验结果

	基准模型	不同设置下的模型		
金字塔池化模块		√		√
多尺度区域建议网络			√	√
<b>mAP</b>	60.27	62.46	63.74	65.94

Faster R-CNN: Towards real-time object detection with region proposal networks, 2015.

R2CNN: Rotational region CNN for orientation robust scene text detection, 2017.



## 2 遥感图像目标检测

### ● 实验结果

- ✓ 如左图所示的球场区域，在没有金字塔池化的模型中，框中物体被错判为船只，然而根据常识这显然是不合理的，船只是不可能出现在球场上的，在加入金字塔池化结构后，物体转而被正确识别为小汽车。
- ✓ 如右图所示，图中间橙黄色位置的小船，在原有的基准网络中无法检出，而在加入了多尺度建议网络之后，小船被成功检出。



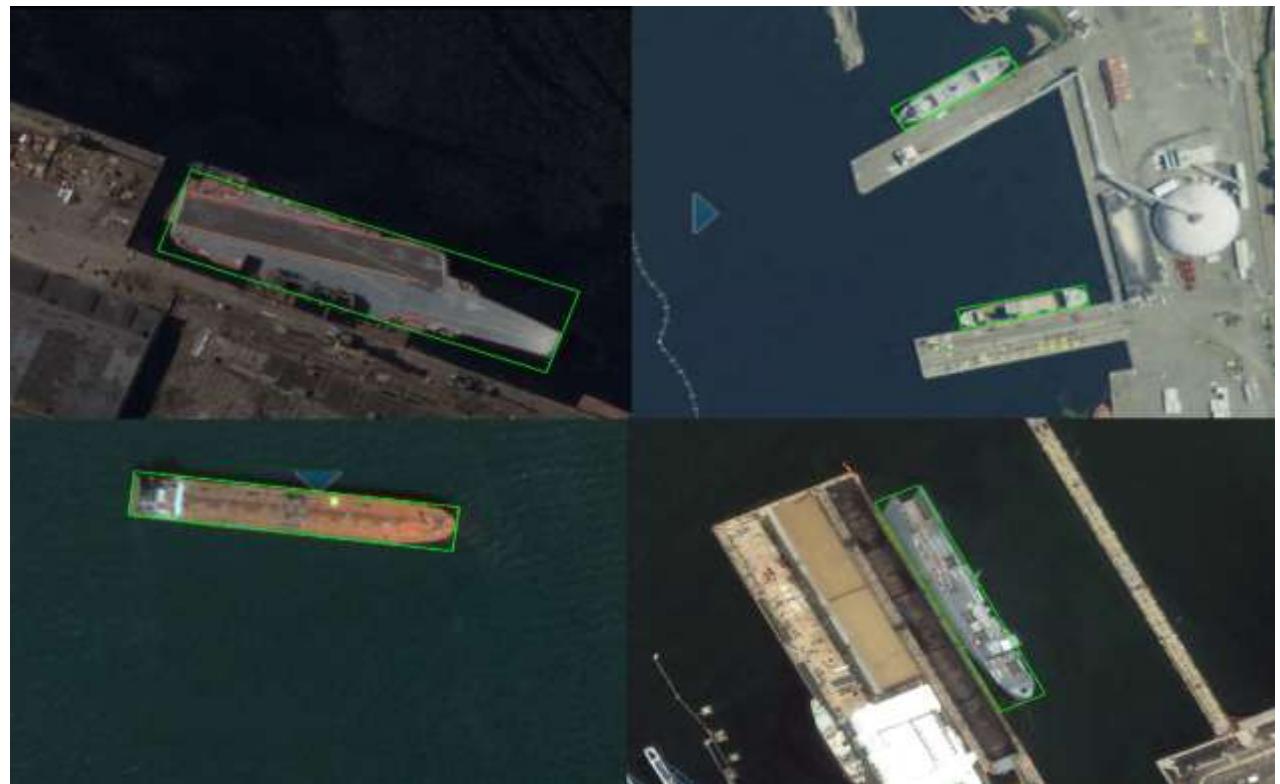
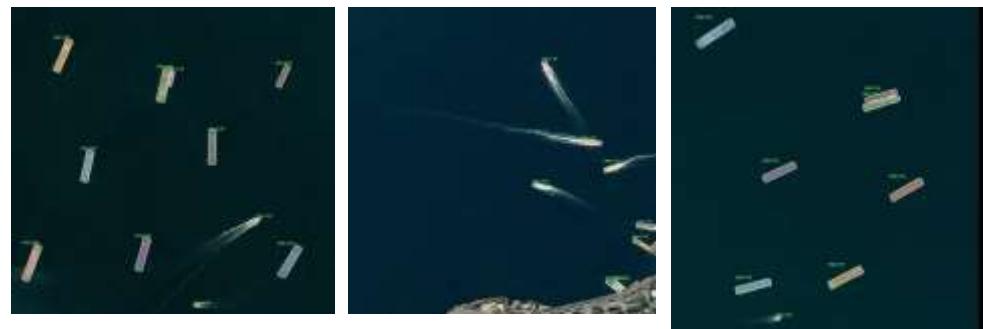
# 2 遥感图像目标检测

## ● 实验结果

✓ 在HRSC2016数据集上进行了测试，旋转Faster R-CNN的基准算法mAP为81.5，只加入金字塔结构后，模型性能提升约3.1个百分点，只加入多尺度区域建议网络后，模型提升了约0.5个百分点，二者全部加入后，整个模型提升越4.2个百分点，充分证明了方法的有效性。

表3 不同改进设置下的实验结果（HRSC206）

	基准模型	不同设置下的模型		
金字塔池化模块		√		√
多尺度区域建议网络			√	√
mAP	81.5	84.6	82.0	85.7



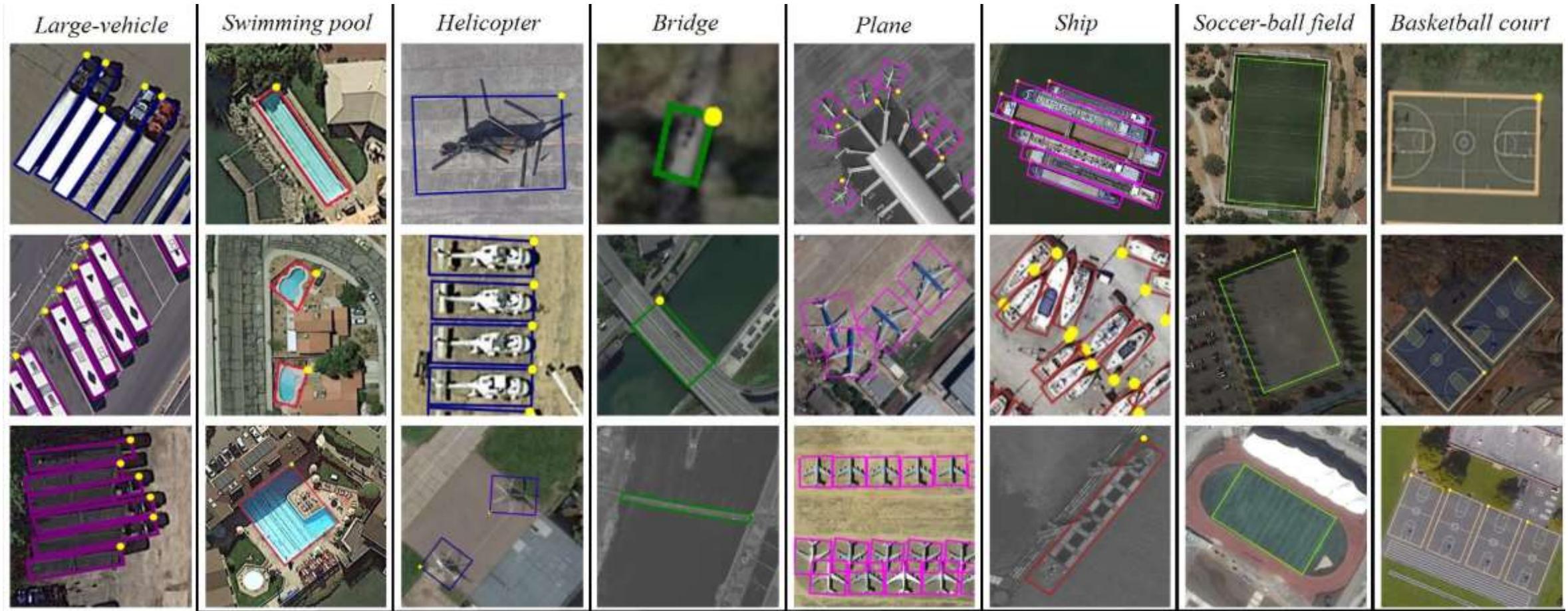
# 2 遥感图像目标检测



# 2 遥感图像目标检测



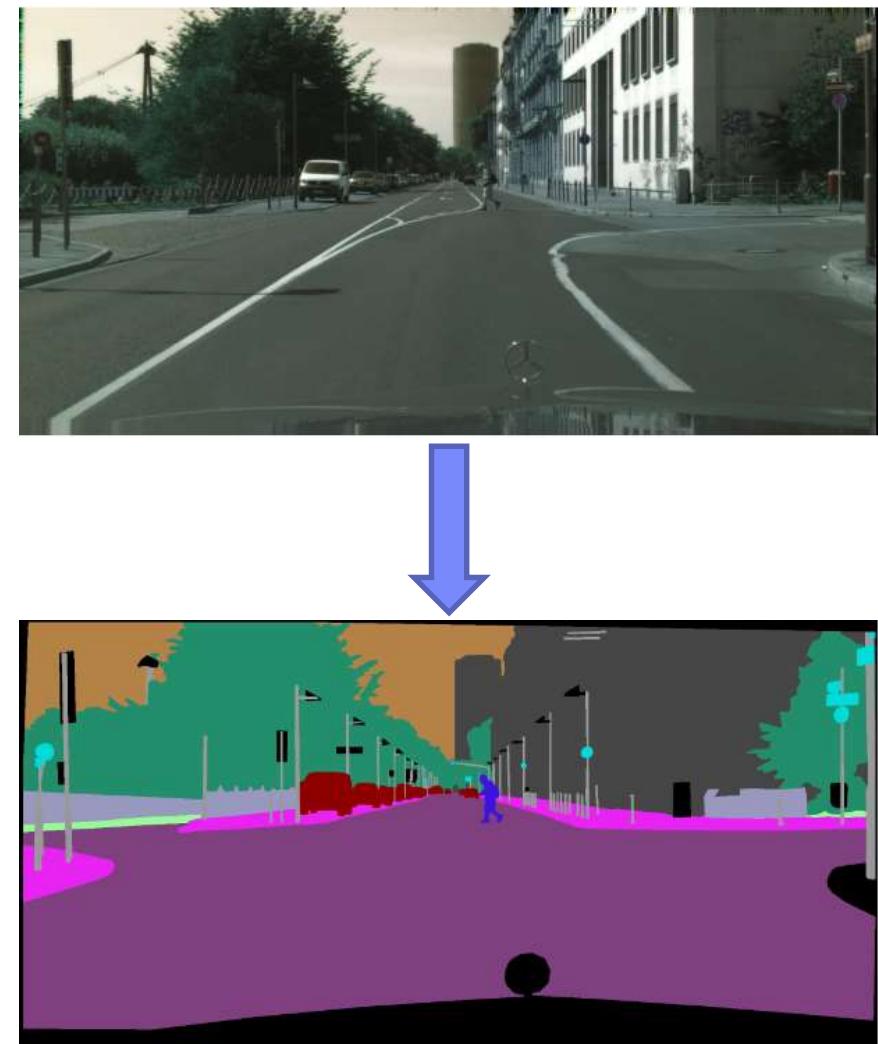
# 2 遥感图像目标检测



# 二、复杂场景语义分割

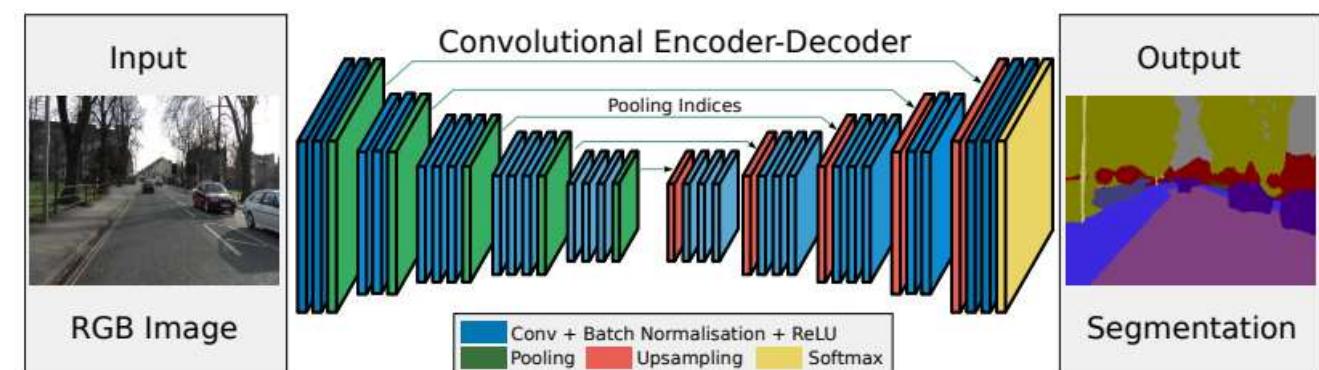
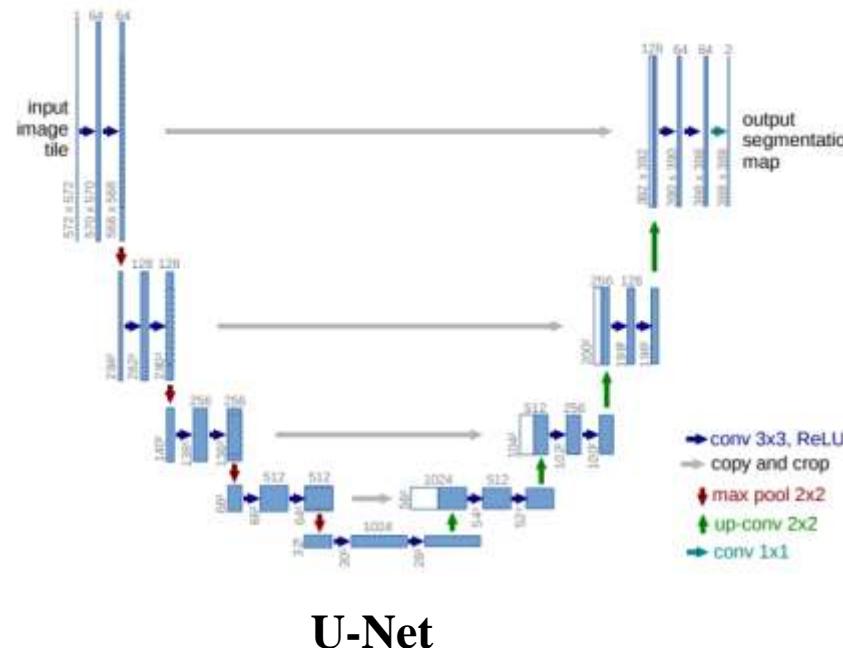
# 1 研究背景

- 图像语义分割 (Image Semantic Segmentation)  
为图像中的每一个像素分配一个预先定义好的表示其语义类别的标签。
- 引入深度神经网络 (Deep Neural Network)  
将分割过程发展为端到端 (end-to-end) 进行像素分类的形式。
- 依赖于神经网络的问题  
习惯使用更复杂的网络换取更好的效果。
- 研究目的  
针对图像语义分割任务特性，改进网络结构，提高性能，减少冗余。



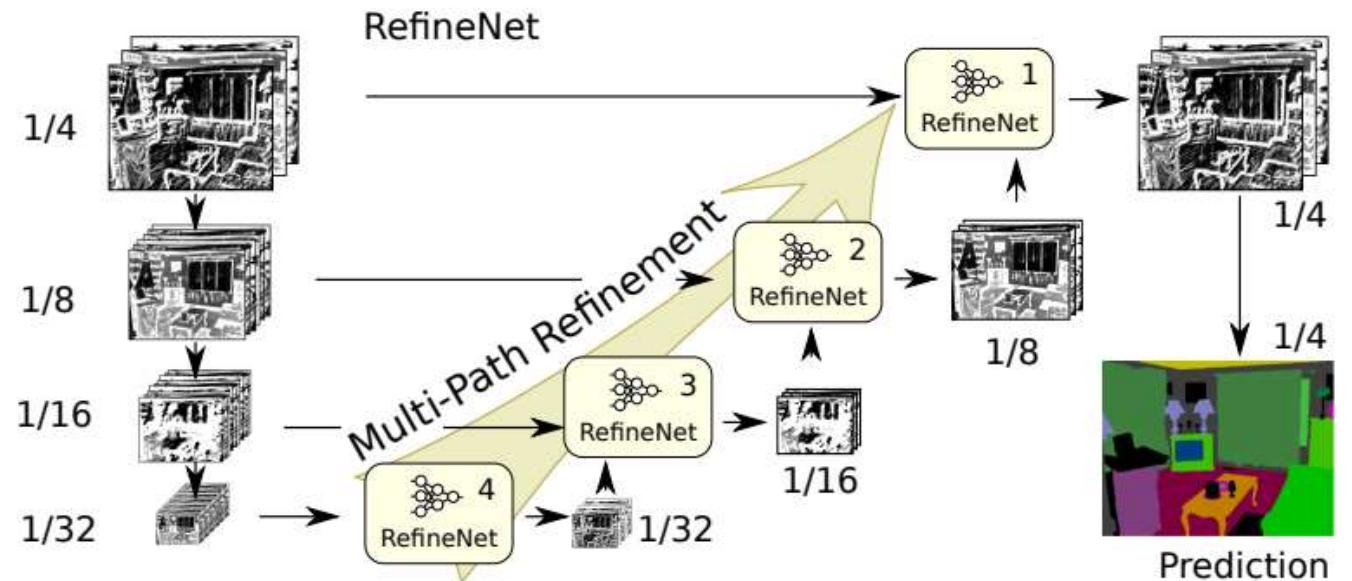
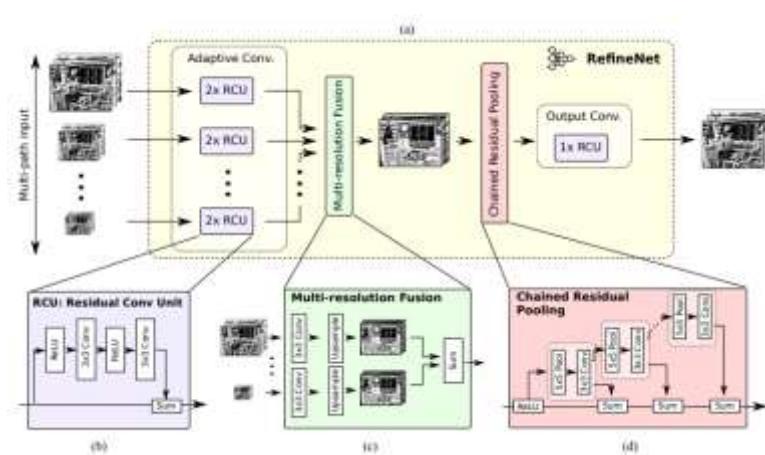
## 2 研究现状

**编码-解码模型** (Encoder-Decoder) 引出了融合多尺度特征获取丰富的语义信息的思路，以弥补连续的卷积和池化操作造成的信息丢失。



## 2 研究现状

**多尺度特征融合** (Multi-scale Feature Fusion) : 大多数方法在编码-解码模型基础上添加特征融合模块，使用相同的融合模块逐级融合各阶段特征。



# 3 算法思想

## Motivation:

- ✓ 对不同阶段的特征融合使用相同的融合结构，没有针对性融合策略；
- ✓ 重复使用同一融合结构会造成结构冗余，计算量大。

### 根据互补原则选择特征

- 特征提取网络不同阶段的特征在语义分割任务中的作用不同
- 选择相应阶段特征进行互补

### 基于双通道注意力机制 融合高层语义特征

- 避免C3阶段的错误特征对C4阶段特征造成影响
- 充分利用高阶特征的主导性

### 基于空间注意力机制 选择低层空间细节特征

- 充分利用低阶特征的细节信息
- 避免缺少准确的语义特征带来的问题

# 3 算法思想

**互补：**高层特征语义信息更强，整体性强，但是边缘细节表现不好；低层特征虽然边缘细节分割效果好，但是目标内部容易出现错误分类。



原始图像



真实标签



高层特征表现

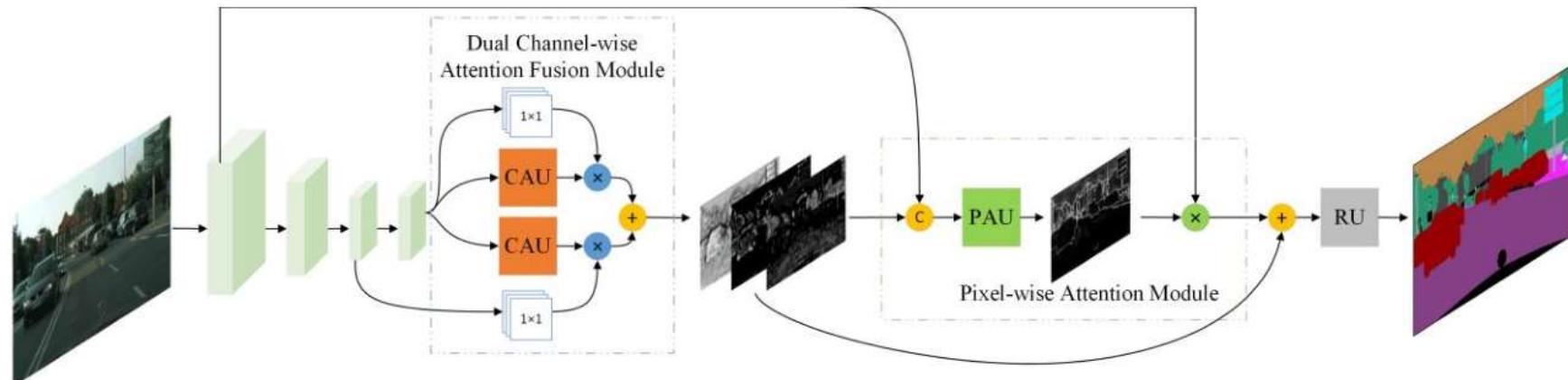


低层特征表现

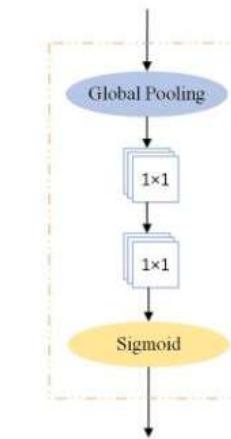
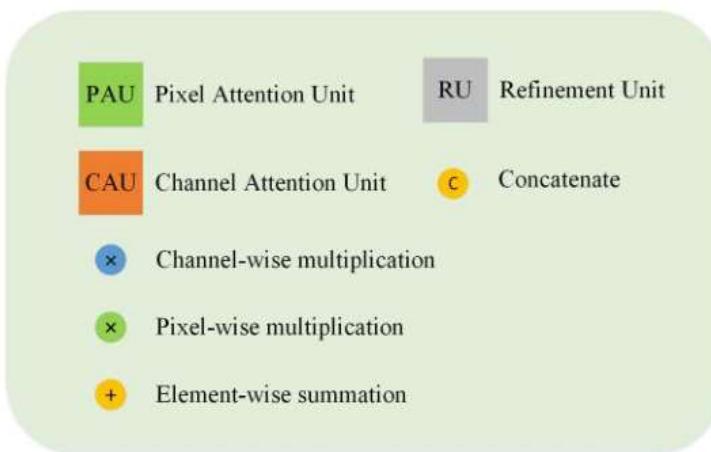
# 4 算法实现

## 网络架构：

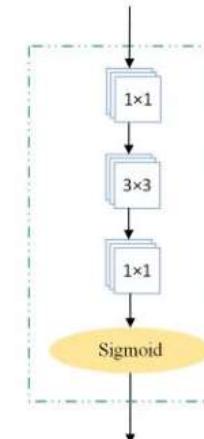
- ✓ 对Resnet的第一阶段的特征使用空间注意力机制；
- ✓ 对第3、4阶段特征使用双通道注意力机制。



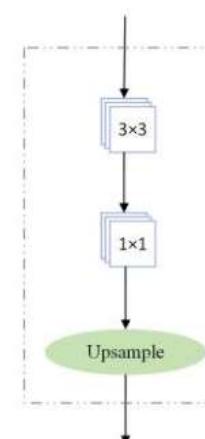
(a) Overall Network



(b) Channel Attention Unit



(c) Pixel Attention Unit



(d) Refinement Unit

# 4 算法实现

**双通道注意力机制：融合高层特征。高阶特征对所有特征进行指导，选择更具区分力的特征。**

某一位位置在各通道的值组成向量为：

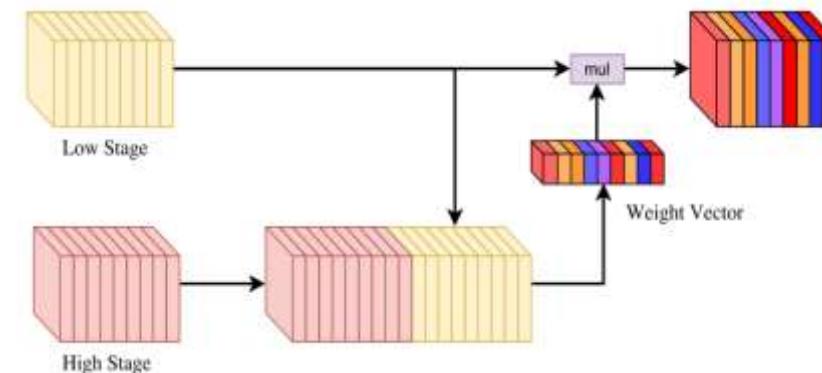
$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix}$$

其中， $k$ 为通道数。使用高阶信息学习得到的权重为：

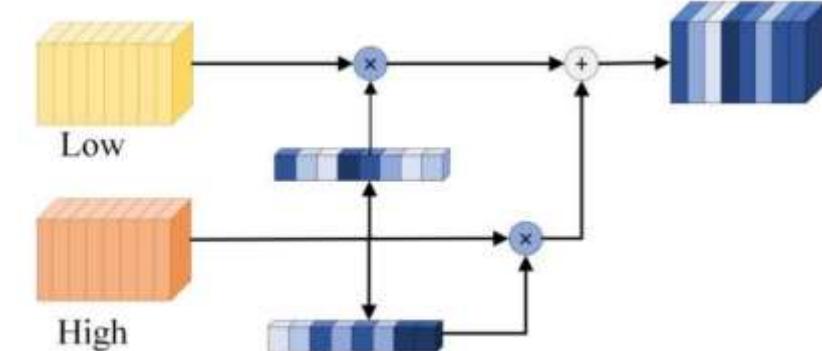
$$\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix}$$

调整后的该位置对应各通道的分布为

$$y^* = \alpha y = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} \alpha_1 y_1 \\ \vdots \\ \alpha_k y_k \end{bmatrix}$$



单通道注意力机制



双通道注意力机制

# 4 算法实现

## 空间注意力机制：选择低层特征。

空间信息的选择过程与通道的权重调整不同，是对每个通道都使用同一置信图，某一通道特征图为 $H \times W$ 大小

$$X_k = \begin{bmatrix} x_{1,1} & \cdots & x_{1,W} \\ \vdots & \ddots & \vdots \\ x_{H,1} & \cdots & x_{H,W} \end{bmatrix}$$

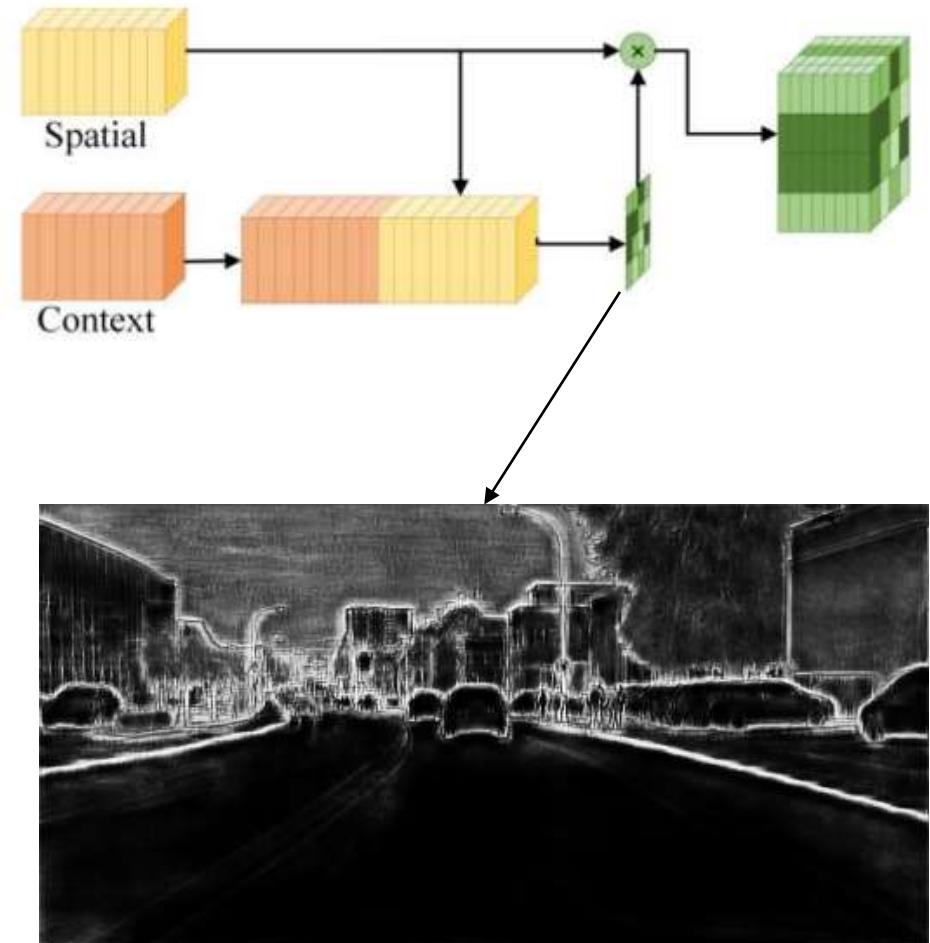
其中 $k \in \{1, 2, \dots, K\}$ ， $K$ 为通道数。置信图为

$$C = \begin{bmatrix} c_{1,1} & \cdots & c_{1,W} \\ \vdots & \ddots & \vdots \\ c_{H,1} & \cdots & c_{H,W} \end{bmatrix}$$

选择后的特征图为

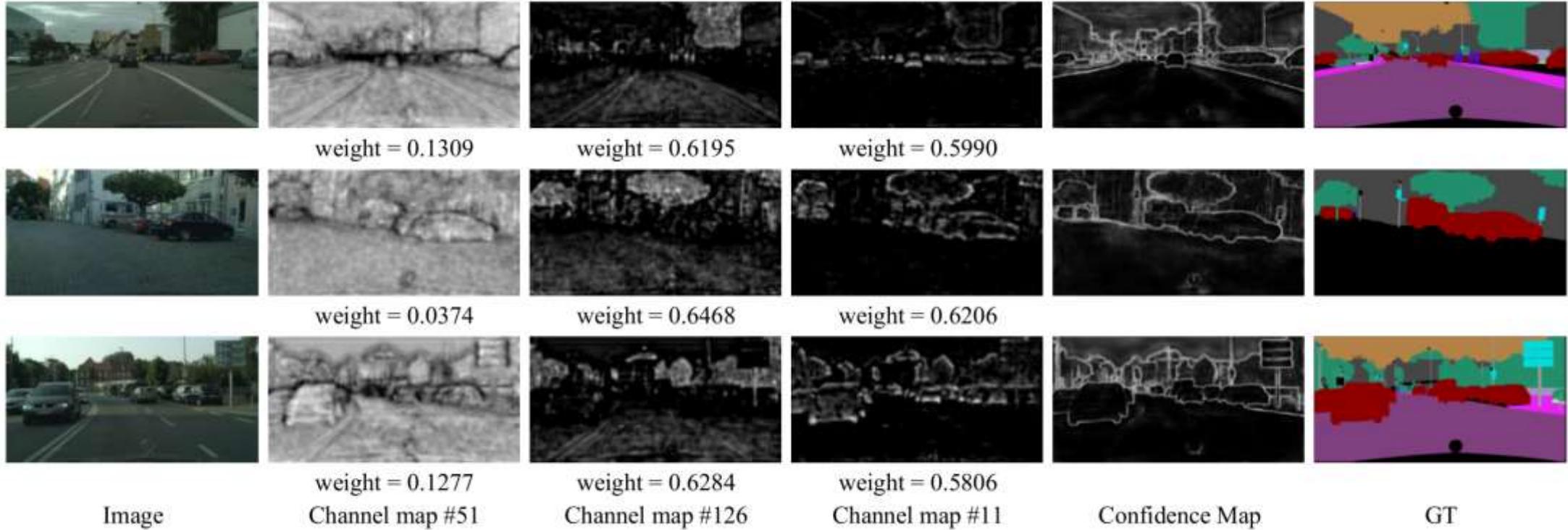
$$X_k^* = CX_k = \begin{bmatrix} c_{1,1} & \cdots & c_{1,W} \\ \vdots & \ddots & \vdots \\ c_{H,1} & \cdots & c_{H,W} \end{bmatrix} \cdot \begin{bmatrix} x_{1,1} & \cdots & x_{1,W} \\ \vdots & \ddots & \vdots \\ x_{H,1} & \cdots & x_{H,W} \end{bmatrix}$$

$$= \begin{bmatrix} c_{1,1}x_{1,1} & \cdots & c_{1,W}x_{1,W} \\ \vdots & \ddots & \vdots \\ c_{H,1}x_{H,1} & \cdots & c_{H,W}x_{H,W} \end{bmatrix}$$



# 5 实验结果

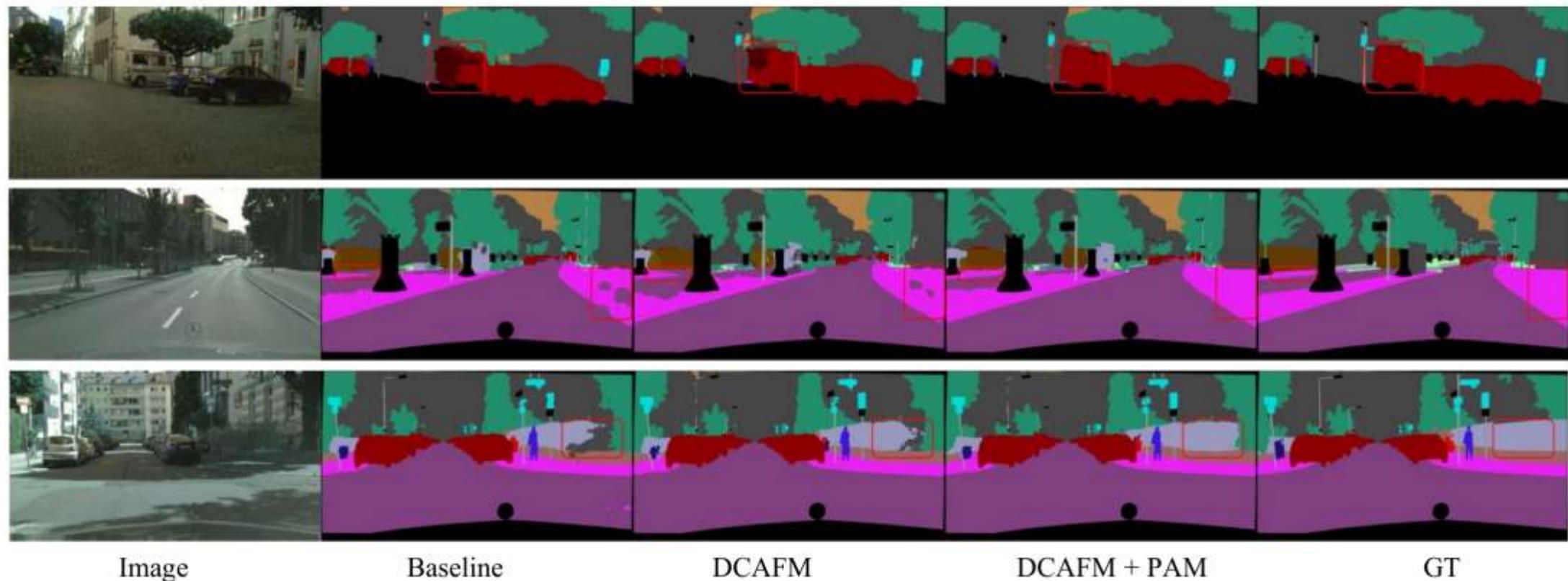
## 注意力机制可视化分析



- 双通道注意力机制：权重较高的通道（#126、#11）的特征更具区分度（象征了某一种特征，如树、车），权重较低的通道（#51）的特征值普遍较高，区分度不高。
- 空间注意力机制：（Confidence Map）权重较高的位置集中在边缘，符合预期。

# 5 实验结果

**消融实验结果：**逐步添加模块，分割效果逐步提升，验证了模块的有效性。



# 5 实验结果

## 实验结果与代表性方法对比

Cityscapes数据集

Method	Backbone	mIoU%	Params(M)	GFLOPs
FCN-8s-heavy [1]	VGG-16	65.3	134.5	1319.1
Dilation10 [15]	VGG-16	67.1	134.4	997.2
DeepLab [5]	ResNet-101	70.4	43.9	1481.4
RefineNet [3]	ResNet-101	73.6	118.0	6748.2
DUC [11]	ResNet-101	77.6	163.7	1135.7
BiSeNet[38]	ResNet-18	77.7	12.9	<b>104.1</b>
PSPNet [4]	ResNet-101	78.4	75.0	2350.9
Auto-DeepLab[39]	Searched-F20-ASPP	79.9	-	333.3
Dynamic[40]	Layer33-PSP	80.7	-	270.0
DANet[33]	ResNet-101	<b>81.5</b>	66.5	376.4
Ours	ResNet-18	73.6	<b>11.57</b>	105.2

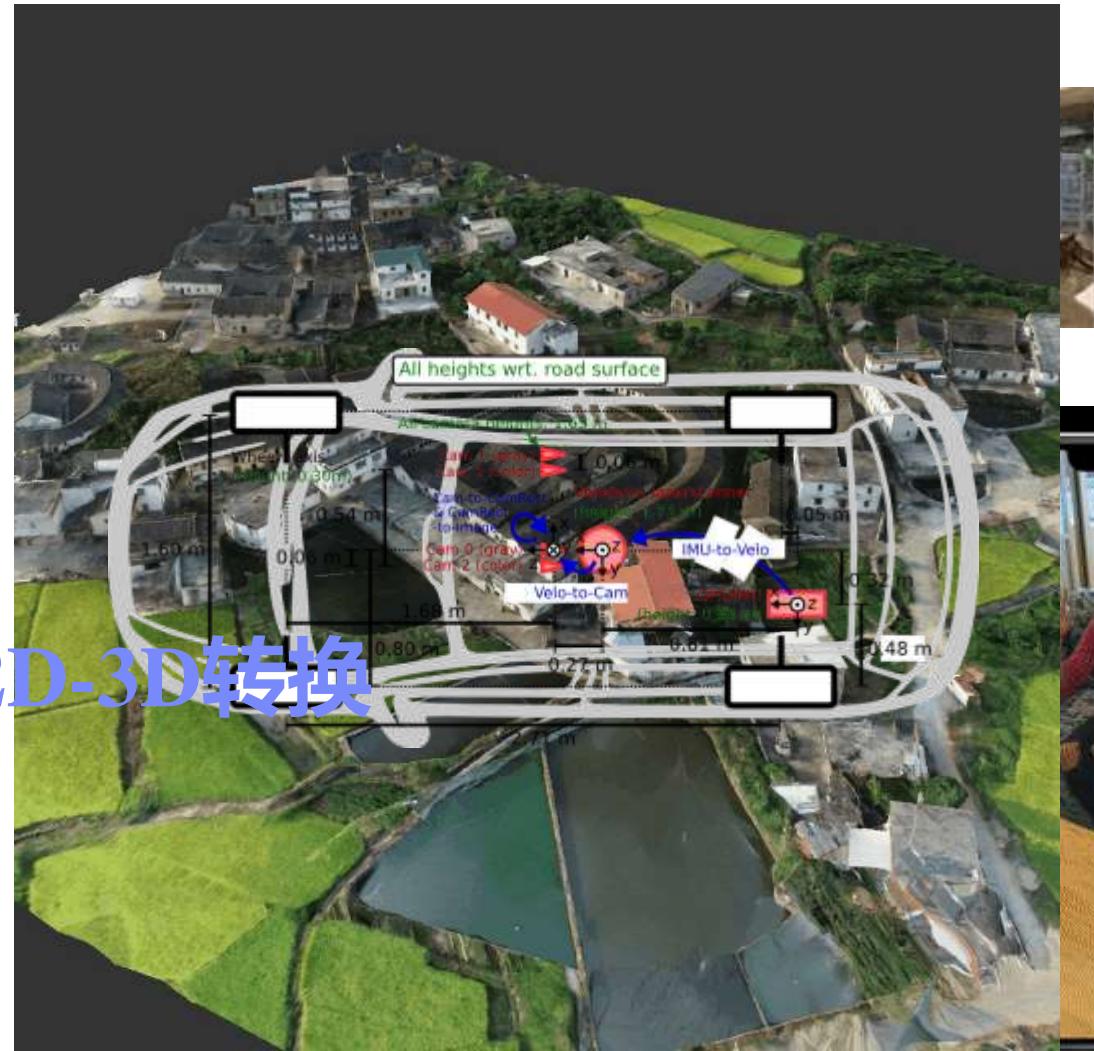
Pascal VOC 2012数据集

Method	Backbone	mIoU%	Params	GFLOPs
FCN-8s-heavy [1]	VGG-16	67.2	134.5	230.6
DeepLab [5]	ResNet-101	71.6	44.1	189.4
RefineNet <sup>†</sup> [3]	ResNet-101	82.4	118.0	843.7
Auto - Searched-F20-DeepLab <sup>†</sup> [39]	ASPP	82.5	-	41.7
DANet [33]	ResNet-101	82.6	66.5	47.1
PSPNet [4]	ResNet-101	82.6	75.0	293.9
DFN[17]	ResNet-101	82.7	90.7	280.2
DUC <sup>†</sup> [11]	ResNet-152	83.1	163.7	142.2
Dynamic <sup>†</sup> [40]	Layer33-PSP	<b>84.0</b>	-	<b>30.8</b>
Ours	ResNet-101	82.3	<b>43.2</b>	45.3

- 与已有的具有代表性方法相比，我们的方法在保证一定精度的同时，大幅度降低了模型复杂度（Params）和计算量（GFLOPs）。

# 三、单目深度估计

# 1 深度估计的应用场景

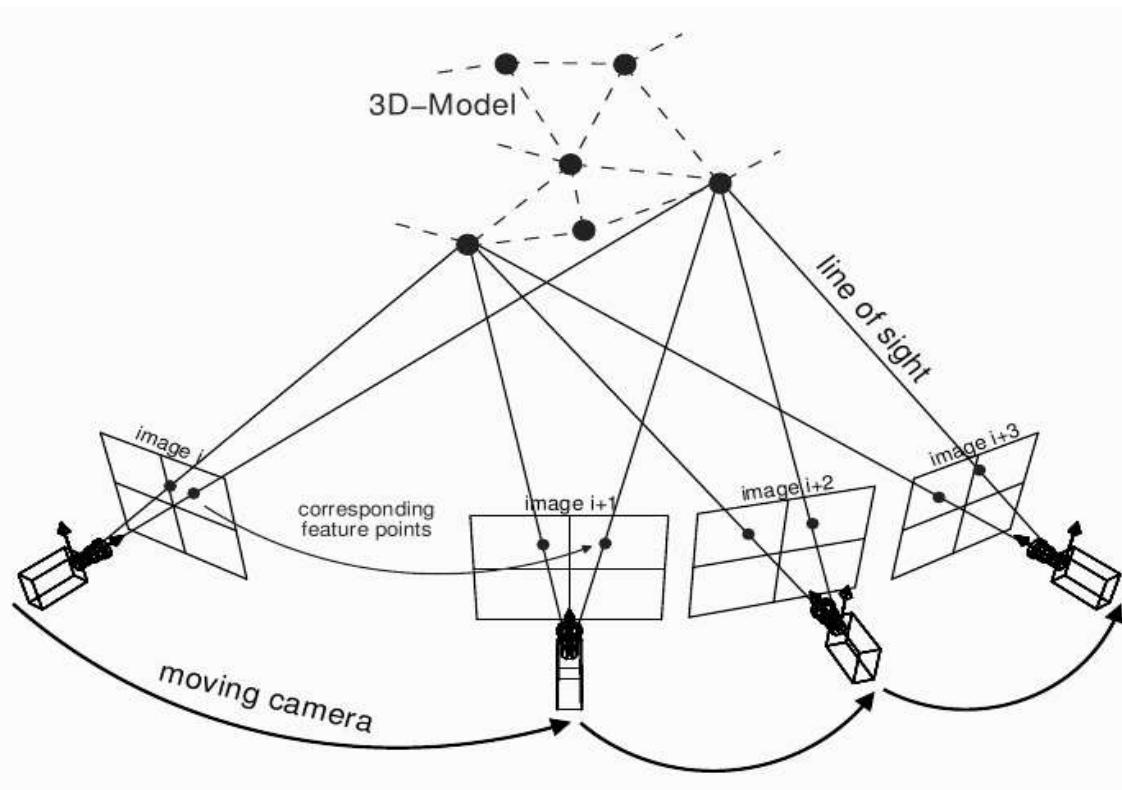


# 2D-3D转换

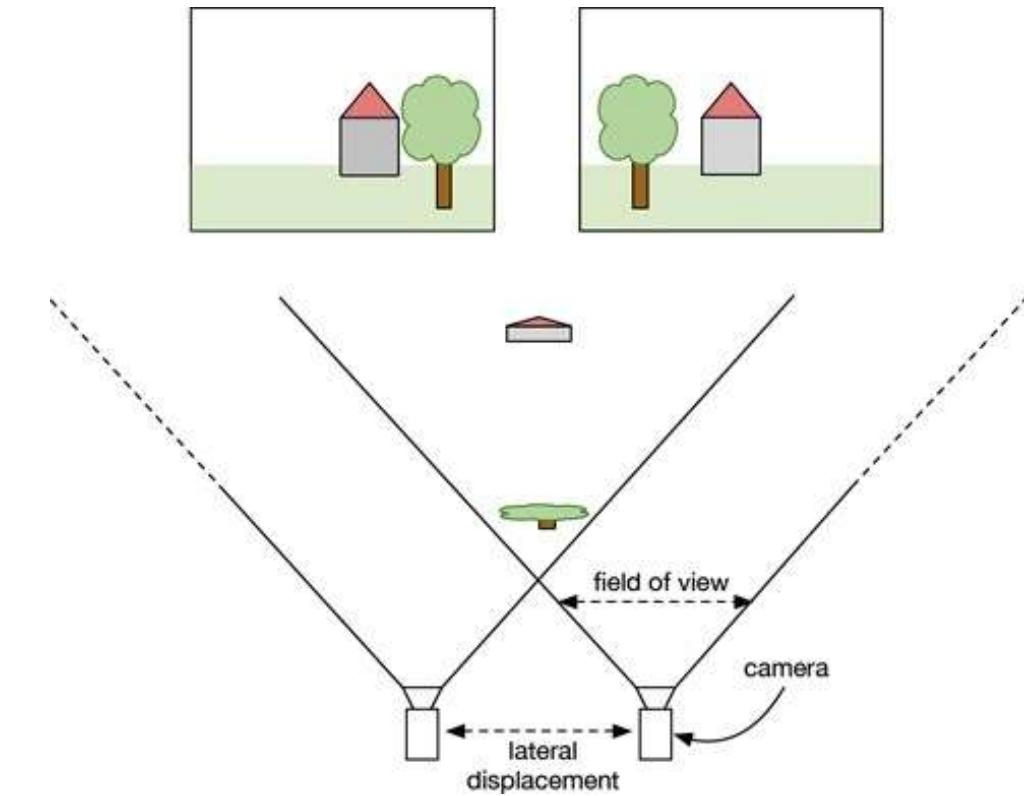
# 无人驾驶

# 2 深度估计的分支

借助SfM方法对图像序列进行处理

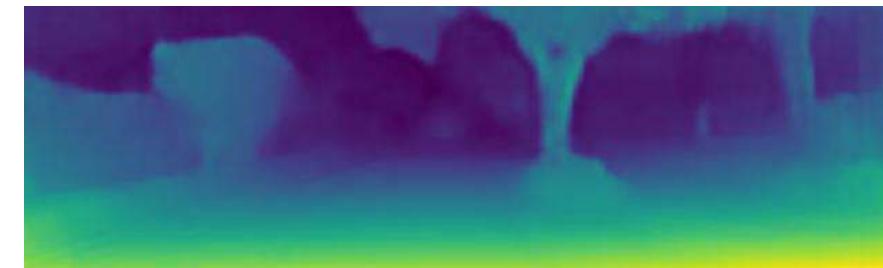
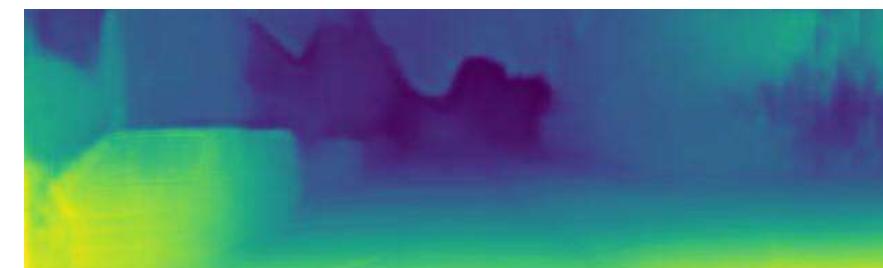


借助双目视差法对双目图像对进行处理



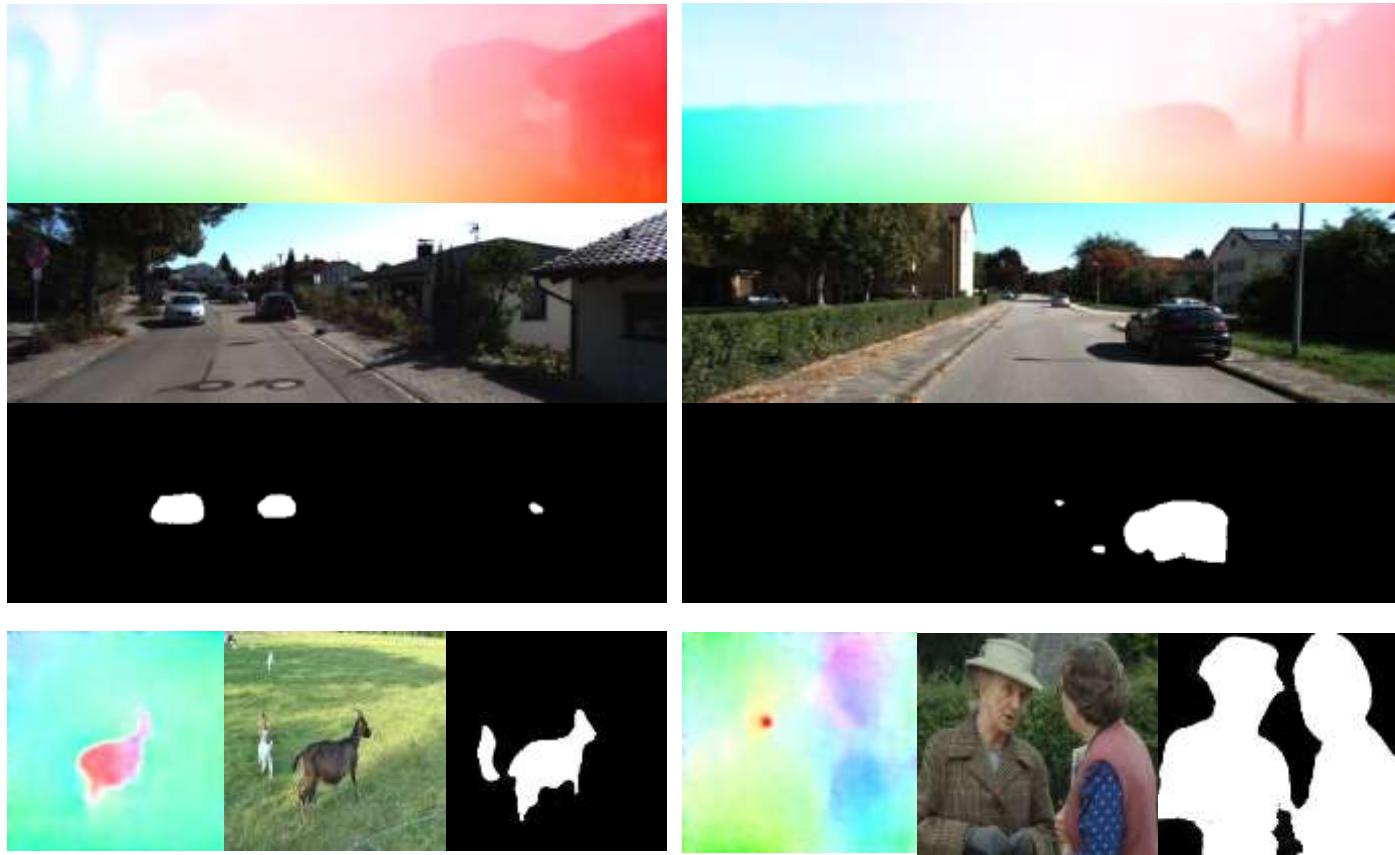
## 2.1 基于SfM的方法

- 基于SfM方法的深度估计借助相机位姿估计，通过优化重投影误差实现无监督学习。
- 可用于对无标注的视频进行单目深度重建。



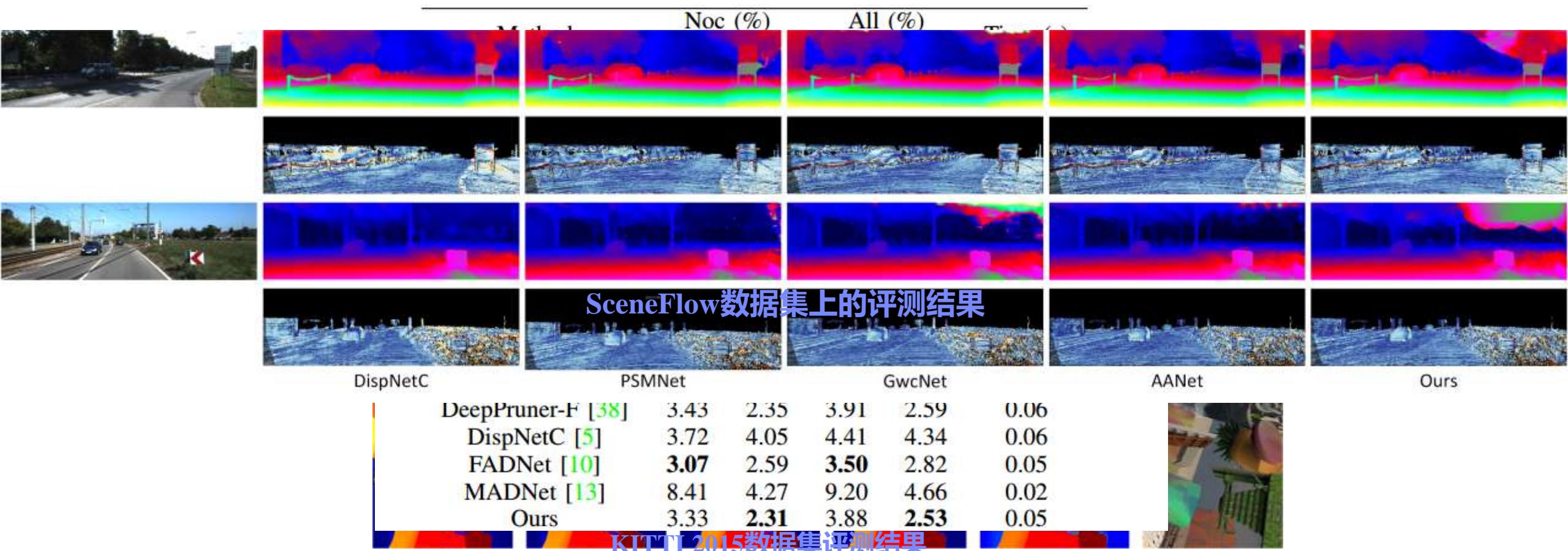
## 2.1 基于SfM的方法-运动分割

- 基于SfM方法的深度估计对于图像序列中含有运动物体的区域进行重投影时存在问题。
- 需要对视野中的运动物体进行分割提取，减小对重投影误差计算的影响。
- 针对运动分割标注难度大，数据量少的问题，可采用迁移学习的思路，在有标注数据集的训练中加入关于无标注数据的约束，使得网络可输出与数据集分布无关的特征用于分割。



## 2.2 基于双目视差的方法

**EDNet:** Efficient Disparity Estimation with Cost Volume Combination and Attention-based Spatial Residual (CVPR2021在投)



## 2.2 EDNet

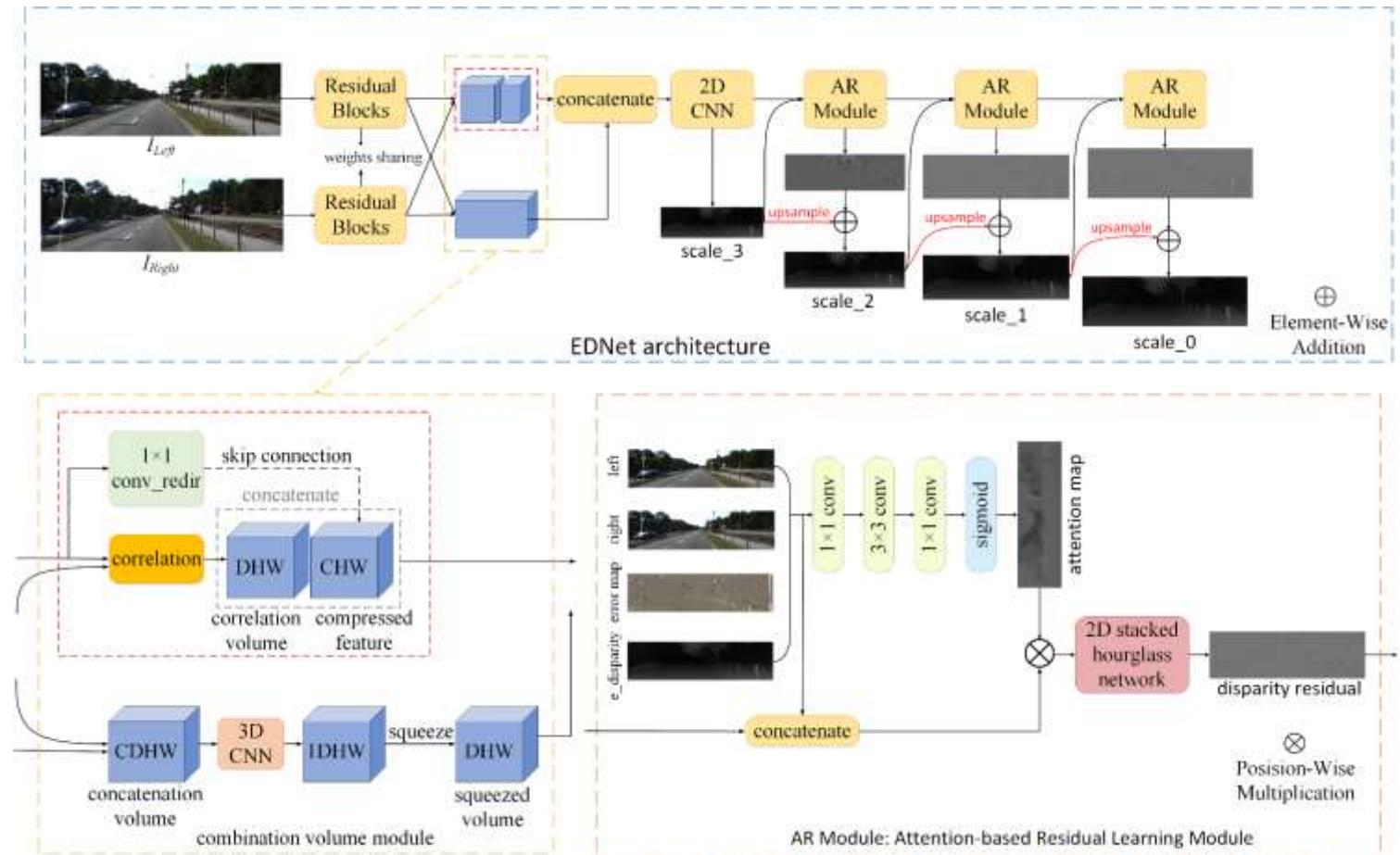
**模型设计的出发点：**希望能够实现更为快速的立体匹配，减少3D卷积的使用；充分利用现有的两种代价量计算方法，实现优势互补；利用误差图提高残差学习的效率。

### ➤ Cost Volume Combination

将当前用于视差计算的两种代价量进行融合，充分利用二者在特征相似度计算和上下文语义特征方面的优势，提取更为鲁棒的特征。

### ➤ Attention-based Spatial Residual

利用空间注意力机制根据多尺度的误差图，计算残差的空间注意力分布，对误差较大的区域进行重点的修正，有效提高了残差学习的效率。



## 2.2 EDNet

**模型设计的出发点：**希望能够实现更为快速的立体匹配，减少3D卷积的使用；充分利用现有的两种代价量计算方法，实现优势互补；利用误差图提高残差学习的效率。

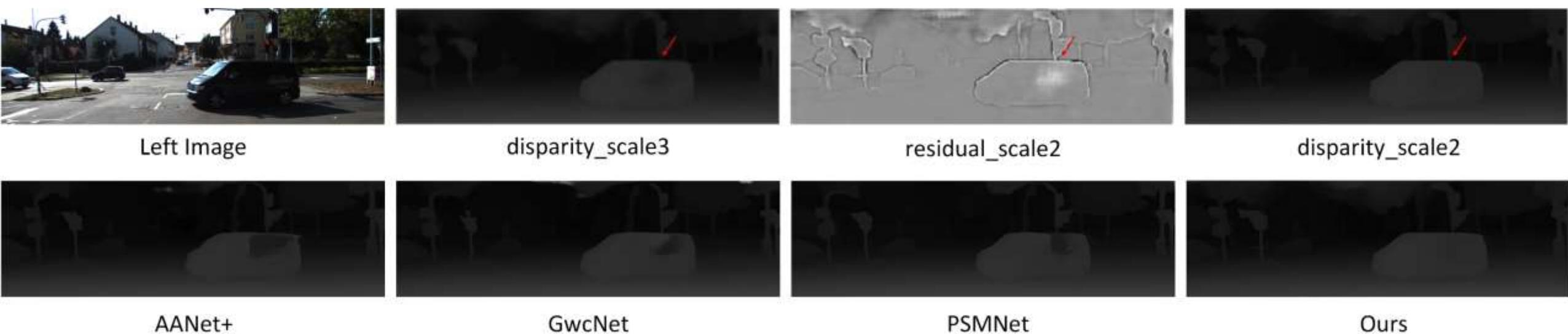
Method	Cost	Volume	Residual	Module	Results		
	correlation	s-concatenation	normal	attention	EPE	>1px(%)	>3px(%)
EDNet-NRS	✓				1.67	27.9	9.7
EDNet-NRCo		✓			1.89	29.9	10.7
EDNet-NR	✓	✓			1.63	27.4	9.8
EDNet-NA	✓	✓	✓		1.04	12.7	<b>5.4</b>
EDNet-NS	✓			✓	1.07	13.1	5.8
EDNet-F	✓	✓		✓	<b>1.00</b>	<b>12.2</b>	<b>5.4</b>

TABLE I: Evaluation of EDNet with different settings. We compute the end-point error, 1-pixel and 3-pixel error on Scene Flow dataset. We use "Co", "S", "R", "A" to denote the correlation volume, squeezed concatenation volume, normal residual, and attention-based residual respectively. "N" stands for "not applied". "EDNet-F" represents the model with all our proposed components and "s-concatenation" means the squeezed concatenation volume.

**在SceneFlow数据集上的消融实验结果，充分证明了我们提出匹配代价量融合方法以及基于空间注意力机制的效果。**

## 2.2 EDNet

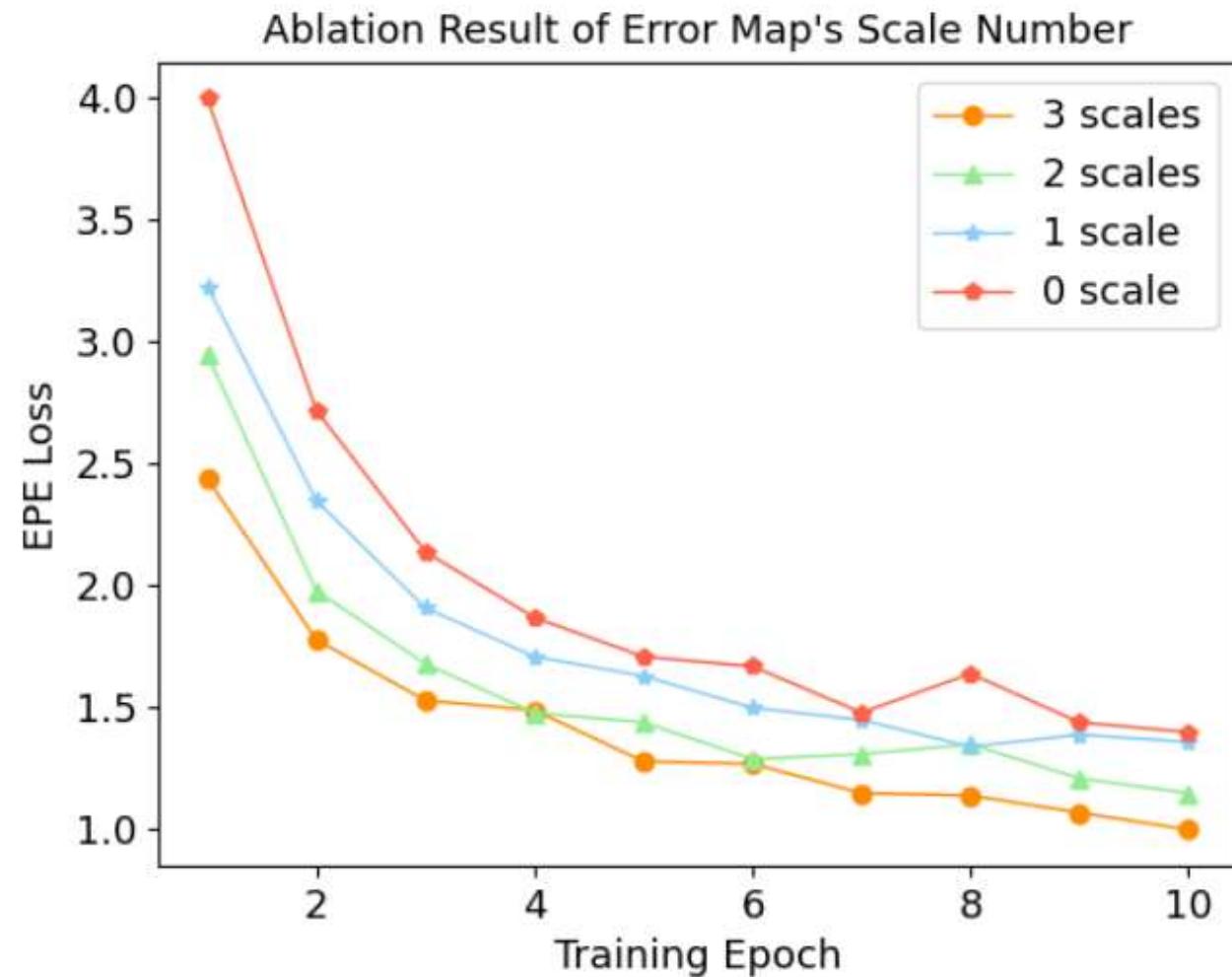
**模型设计的出发点：**希望能够实现更为快速的立体匹配，减少3D卷积的使用；充分利用现有的两种代价量计算方法，实现优势互补；利用误差图提高残差学习的效率。



残差学习过程的可视化，通过我们提出的基于空间注意力机制的残差学习模块，即使是**弱纹理区域**的视差估计结果也能够得到准确的纠正，同时对物体的边缘也进行了有效的锐化。相较于其他代表性工作，我们的模型可以获得更为连续、准确的视差估计值。

## 2.2 EDNet

通过实验，我们证明了误差图对于残差学习的影响，无论是初始的损失函数值还是最终收敛的结果，有了误差图的引导，残差学习的效率得到了有效的提升。



## 2.2 EDNet

Method	Memory (GB)	GFLOPs	Time (s)
PSMNet [2]	4.83	937.9	0.393
GANet [4]	6.53	1936.98	2.43
GwcNet [3]	4.27	899.99	0.272
Bi3D [33]	10.74	4212.05	0.899
Ours	<b>2.52</b>	<b>162.92</b>	<b>0.053</b>

TABLE V: Comparisons of the runtime, running memory and computational cost. All the results are tested on a single NVIDIA RTX 2080Ti GPU at a resolution of  $1248 \times 384$ .

与基于3D卷积的模型相比，我们提出的EDNet在显存占用与模型复杂度方面具有明显的优势。经后期进一步的修改，最新模型的参数量减少了30余倍，显存占用为1.86G，运算复杂度降低为80.45GFLOPs，运算速度为37ms，在SceneFlow上的EPE误差为0.83，在同样轻量级的模型中，依然具有出色的精度。

# 四、其它场景应用

## 4.1 物料箱二维码自动识别



# 4.1 物料箱二维码自动识别



# 4.1 物料箱二维码自动识别



**二维码坐标:**

- (1) 1983 2102 1959 1877
- (2) 2083 2190 586 482

**识别的二维码信息:**

(1)  
b'[]>|x1e06|x1dP92281577|x1d  
Q8|x1d1JUN5303170151193954  
6|x1d20LJQNPGA|x1dK851926  
097|x1d2P|x1dTLS194|x1dC|x1  
e|x04'

(2)  
b'[]>|x1e06|x1dP92281577|x1d  
Q8|x1d1JUN5303170151193954  
7|x1d20LJQNPGA|x1dK851926  
097|x1d2P|x1dTLS194|x1dC|x1  
e|x04'

# 4.1 物料箱二维码自动识别



**二维码坐标:**

(1) 2656 2755 1896 1778

**识别的二维码信息:**

(1)

b'[>|x1e06|x1dP23241633|x1d  
Q20|x1d1JUN530317015119402  
38|x1d20LJQSPGA|x1dK80259  
9157|x1d2P|x1dTLS194|x1dC|x  
1e|x04'

# 4.1 物料箱二维码自动识别



**二维码坐标:**

- (1) 2311 2456 2035 1894
- (2) 2105 2229 319 179

**识别的二维码信息:**

- (1)  
`b'[]>|x1e06|x1dP84308946|x1d  
Q16|x1d1JUN654560929842608  
75|x1d20LJQNPAG|x1dK|x1dT  
LS19003|x1dC|x1e|x04'`
- (2)  
`b'[]>|x1e06|x1dP84308946|x1d  
Q16|x1d1JUN654560929842608  
74|x1d20LJQNPAG|x1dK|x1dT  
LS19003|x1dC|x1e|x04'`

# 4.1 物料箱二维码自动识别



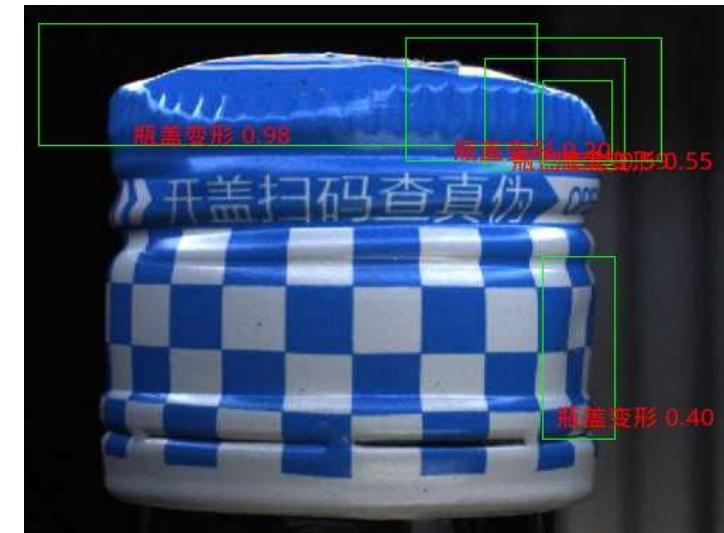
**二维码坐标:**

- (1) 2247 2409 272 133
- (2) 2218 2358 2101 1949

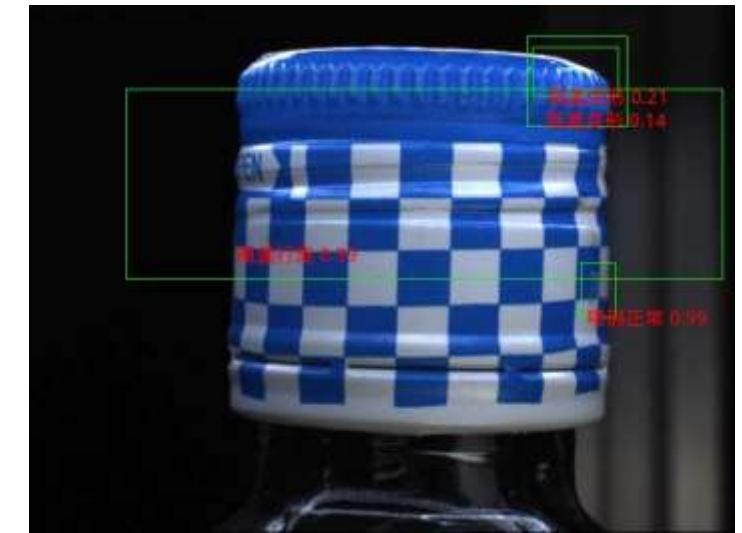
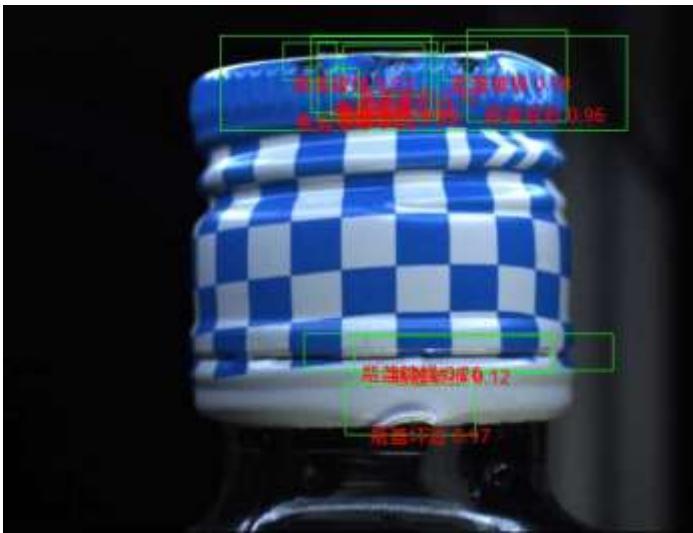
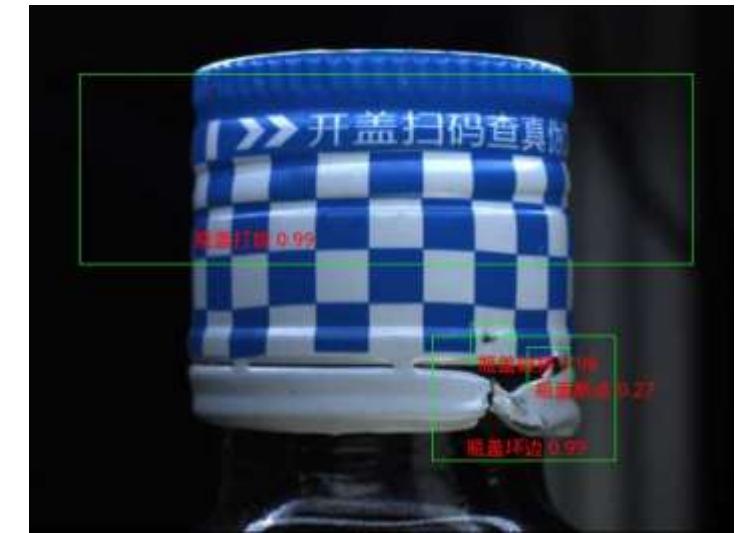
**识别的二维码信息:**

- (1)  
b'[]>|x1e06|x1dP84498219|x1d  
Q8|x1d1JUN6545609298426079  
5|x1d20LJQNPAGA|x1dK|x1dT  
LS19003|x1dC|x1e|x04'  
(2)  
b'[]>|x1e06|x1dP84498207|x1d  
Q8|x1d1JUN6545609298426074  
8|x1d20LJQNPAGA|x1dK|x1dT  
LS19003|x1dC|x1e|x04'

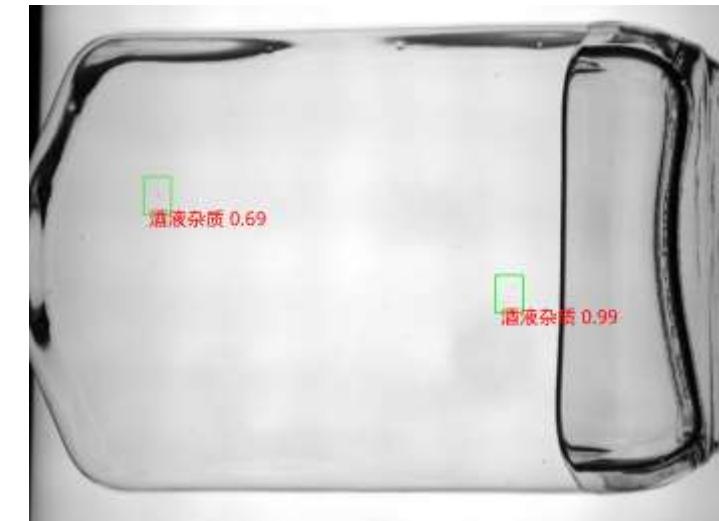
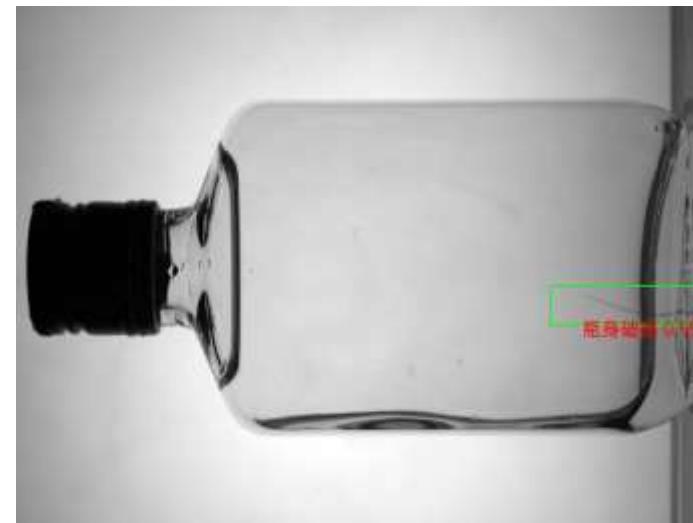
## 4.2 酒瓶缺陷检测



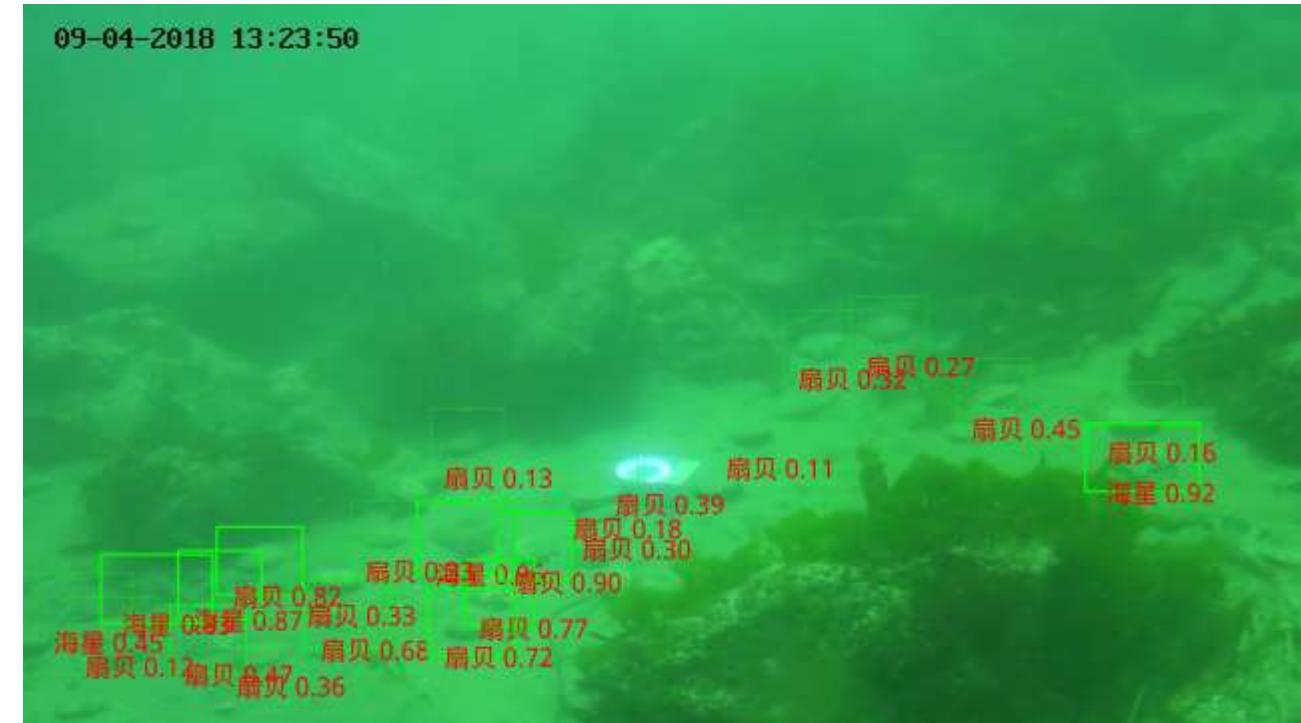
## 4.2 酒瓶缺陷检测



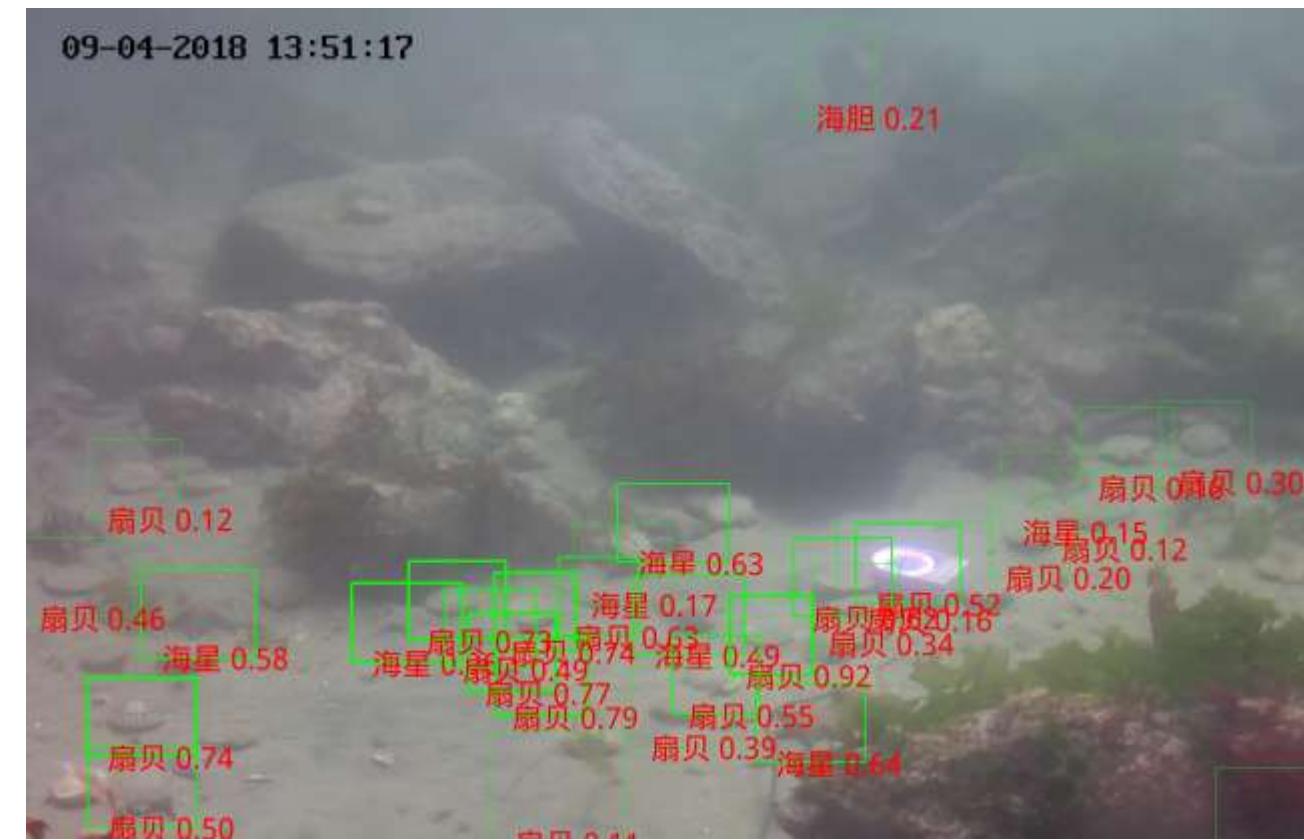
## 4.2 酒瓶缺陷检测



## 4.3 水下目标检测



## 4.3 水下目标检测



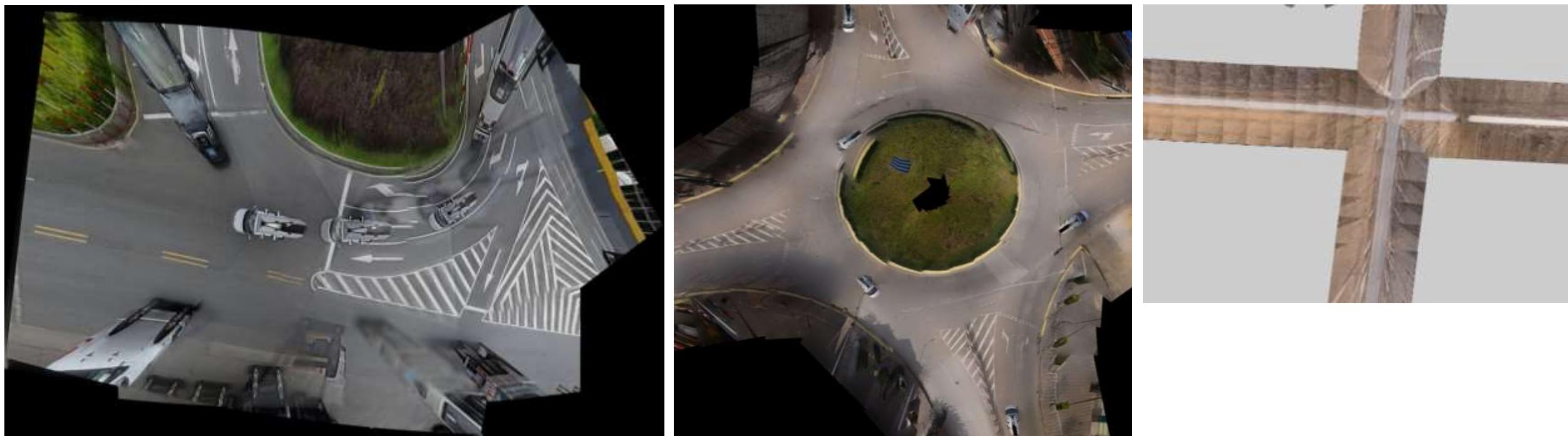
## 4.4 序列图像自动拼接



## 4.4 序列图像自动拼接



## 4.4 序列图像自动拼接



根据多张百度地图全景采集车所采集到的图像进行处理和拼接，得到较为全面的路面俯视信息，可应用于道路信息提取、检测等工作中，目前所研发的百度地图全景视图拼接系统已经上线运行。

## 4.4 序列图像自动拼接

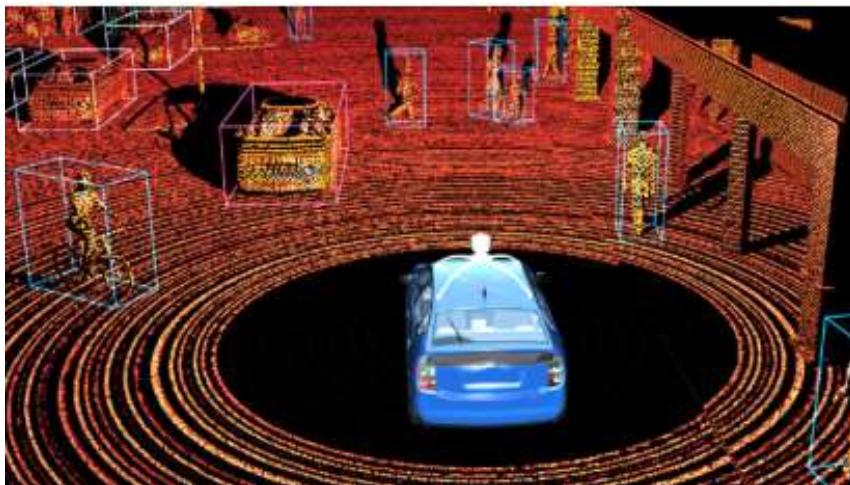


以莫高窟第61窟西壁为例，共需采集**4797**张原始图像（单幅分辨率为**5616×3744**像素），拼接后得到**51.9GB**大小的壁画全景图像，熟练工程师按平均每天拼接**20张/人**，若进展顺利，一个人则需拼接**240天**。如果期间有拼接错误则所需时间更长，费时费力，还不能保证最终的拼接质量。自动拼接系统只需要**5小时**（小型工作站）。

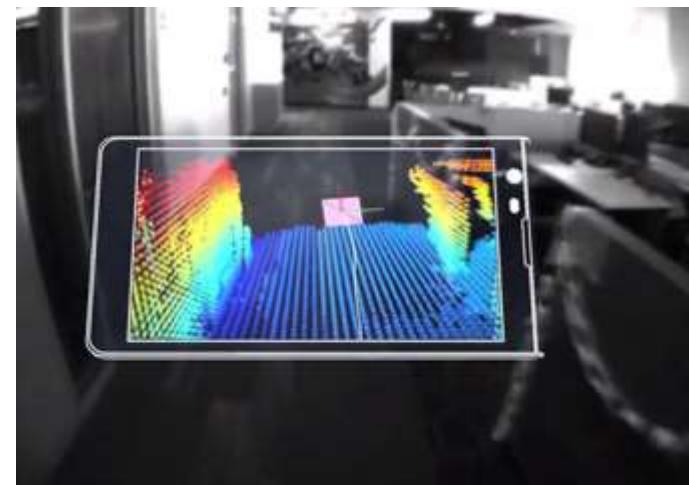
## 4.5 点云语义/实例分割

近年来，传感器的快速发展，使得三维数据的获取越来越容易。此外，随着激光雷达，RGBD相机等3D传感器在机器人、无人驾驶、AR等领域的广泛应用，深度学习在三维点云数据的研究在近几年取得了广泛关注。

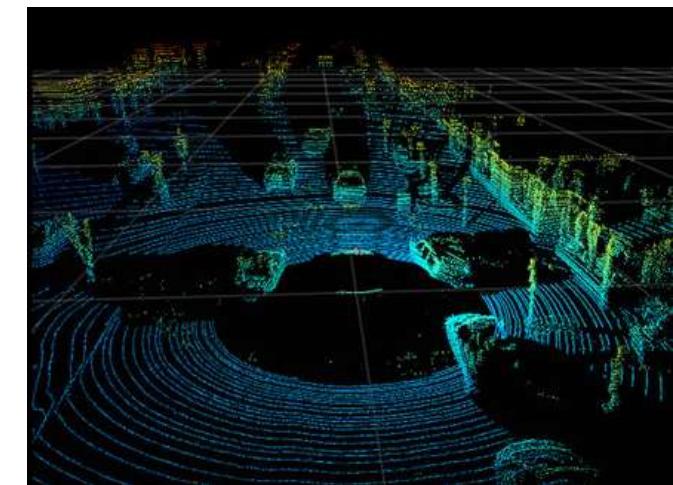
Robot Perception



Augmented Reality

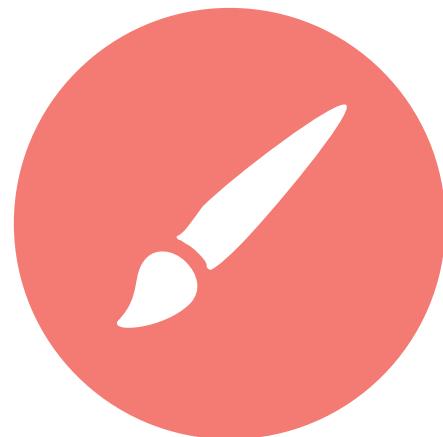


Autonomous Vehicle



Need for 3D Deep Learning!

## 4.5 点云语义/实例分割



### 非结构化

点云以及多边形网格是非结构化的，难以直接输入神经网络。



### 姿态变换的类别不变性

目标点云的任意旋转平移操作，改变空间坐标，但不改变类别属性。



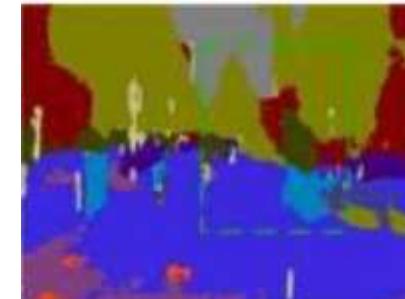
### 数据的多样性

不同目标点云大小不同，模型需要自适应处理不同尺度的点云。



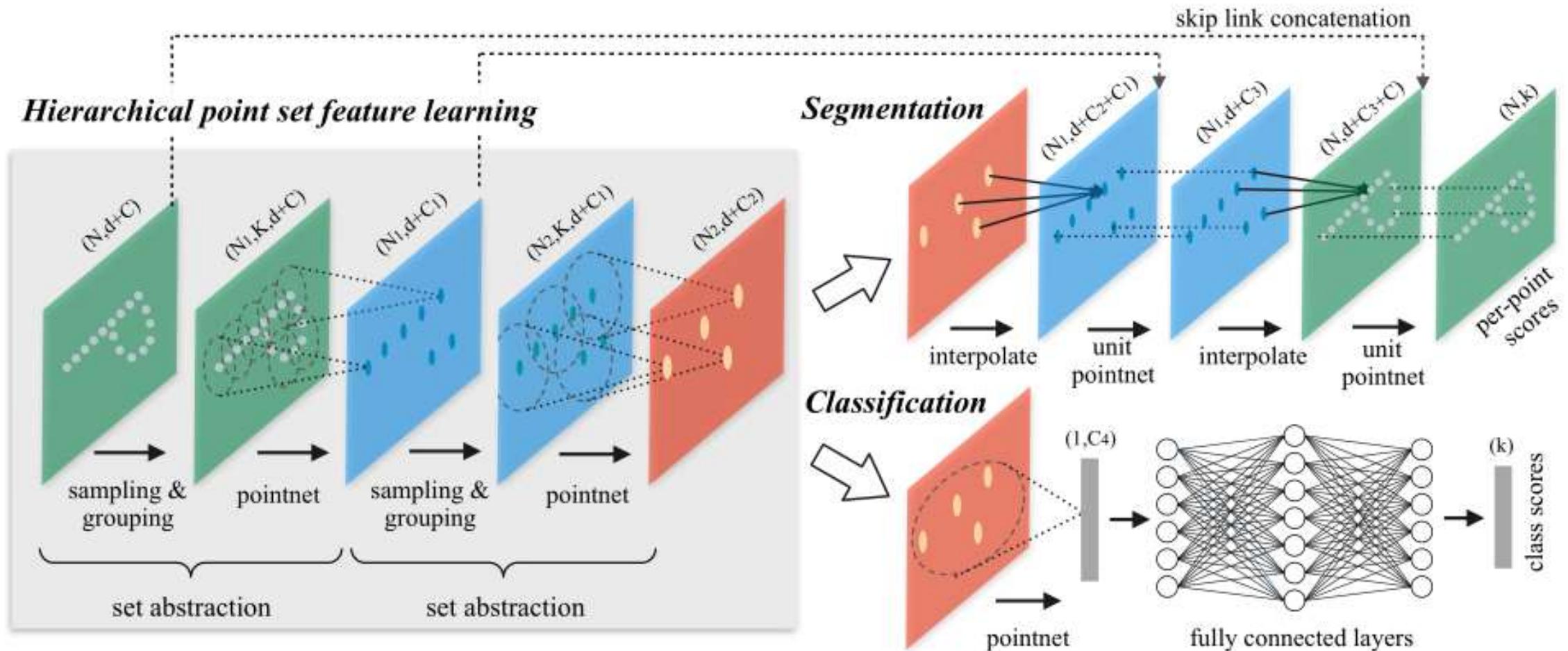
### 挑战

点云数据集规模不大，也增加了场景重构的难度。

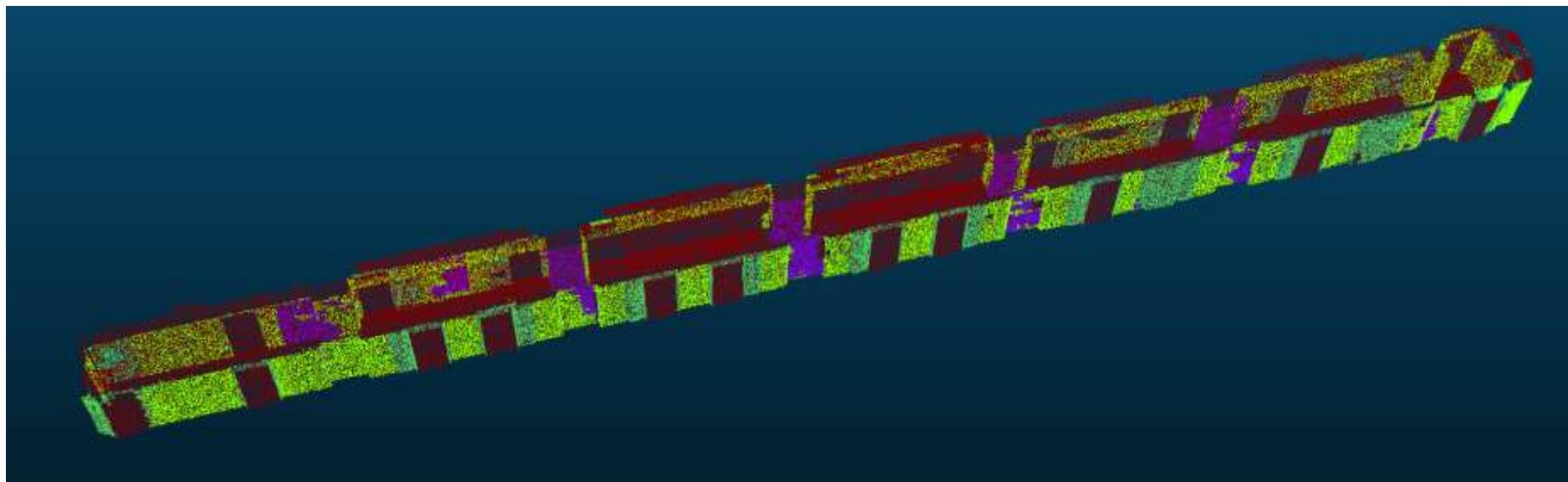
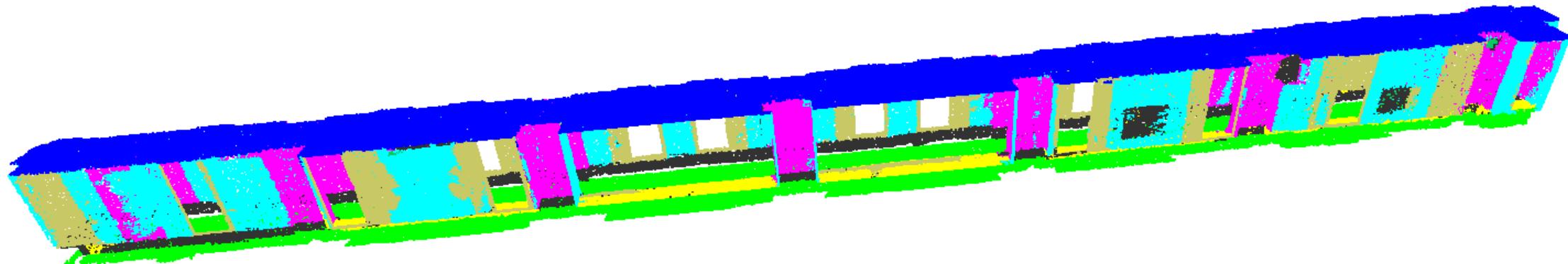


目标遮挡和光照变化会在一定程度上影响3D重构质量。

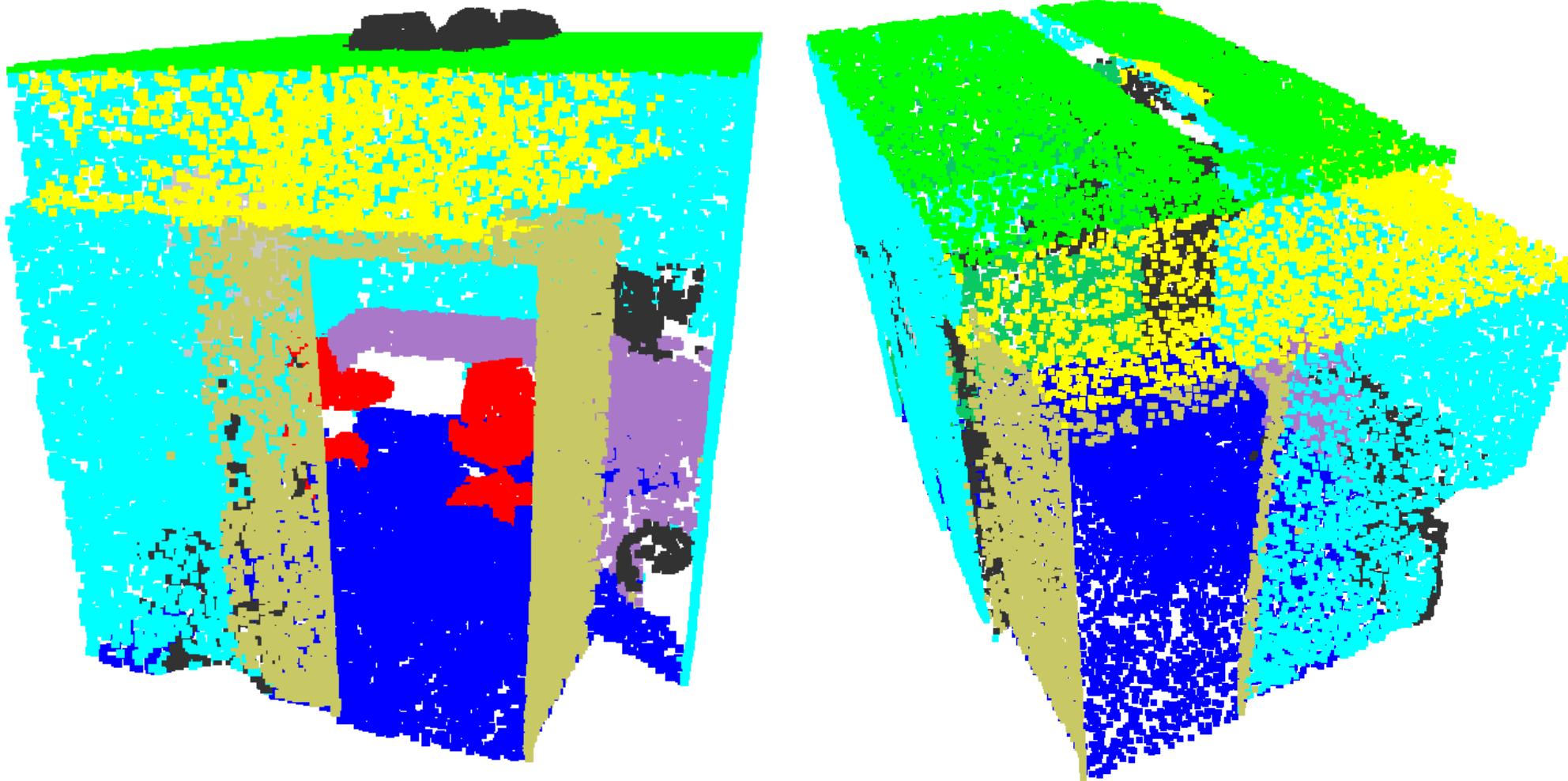
## 4.5 点云语义/实例分割



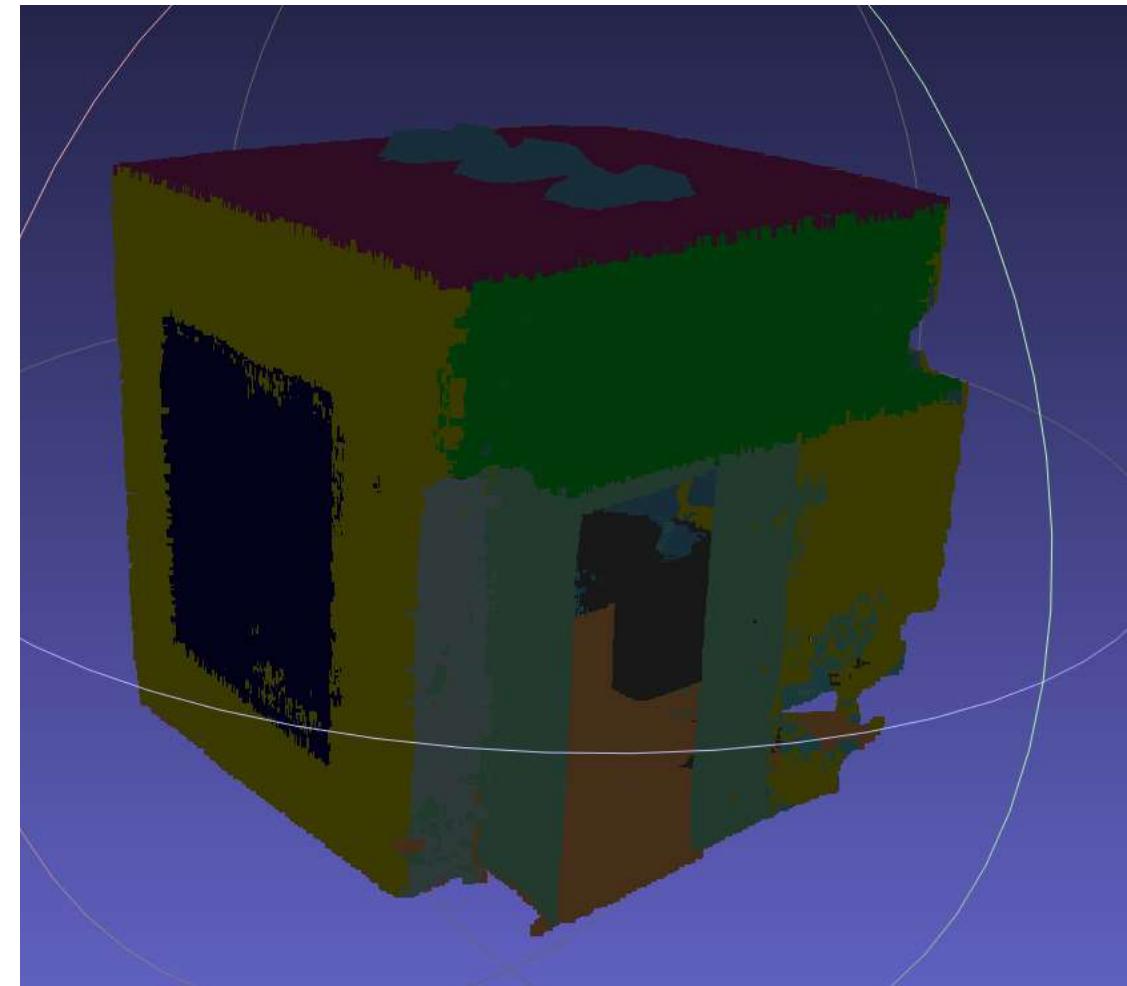
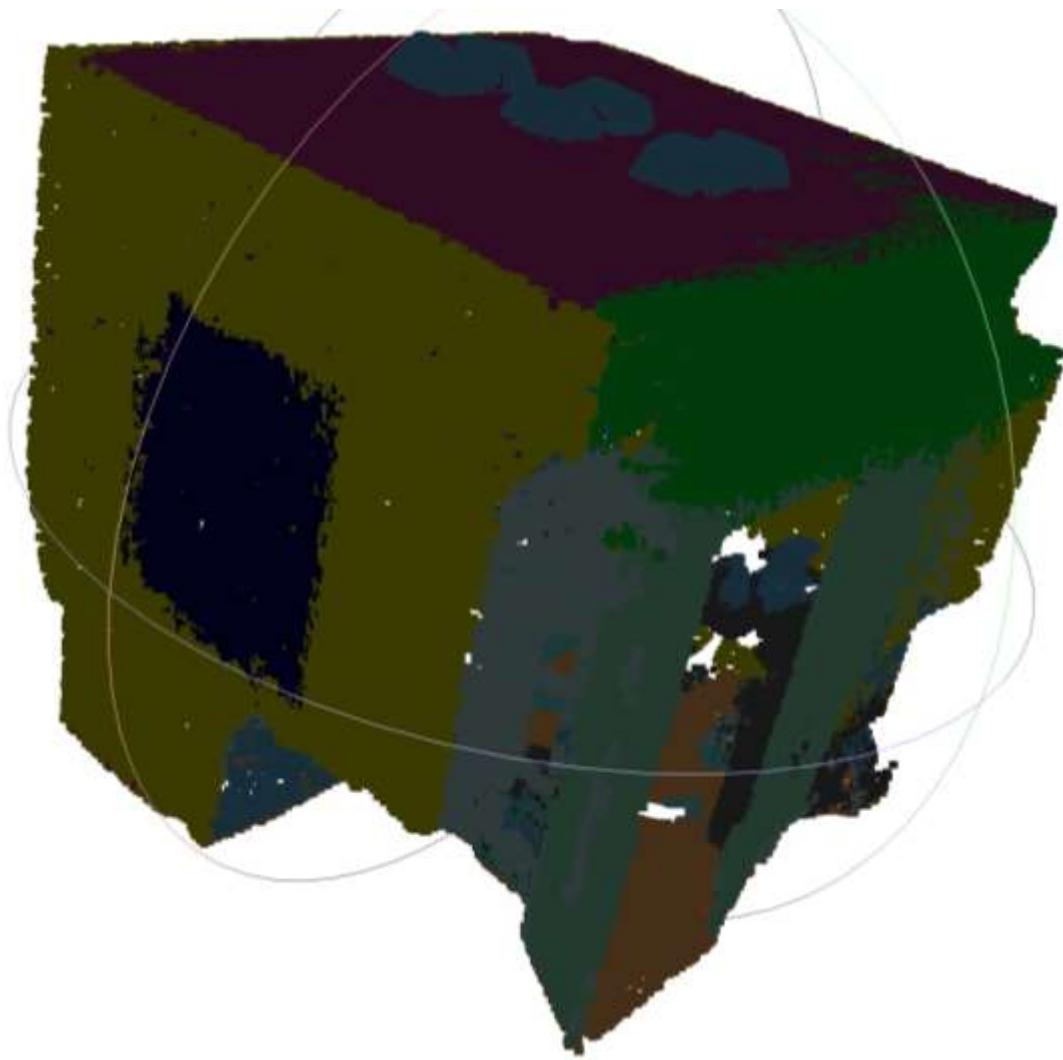
## 4.5 点云语义/实例分割



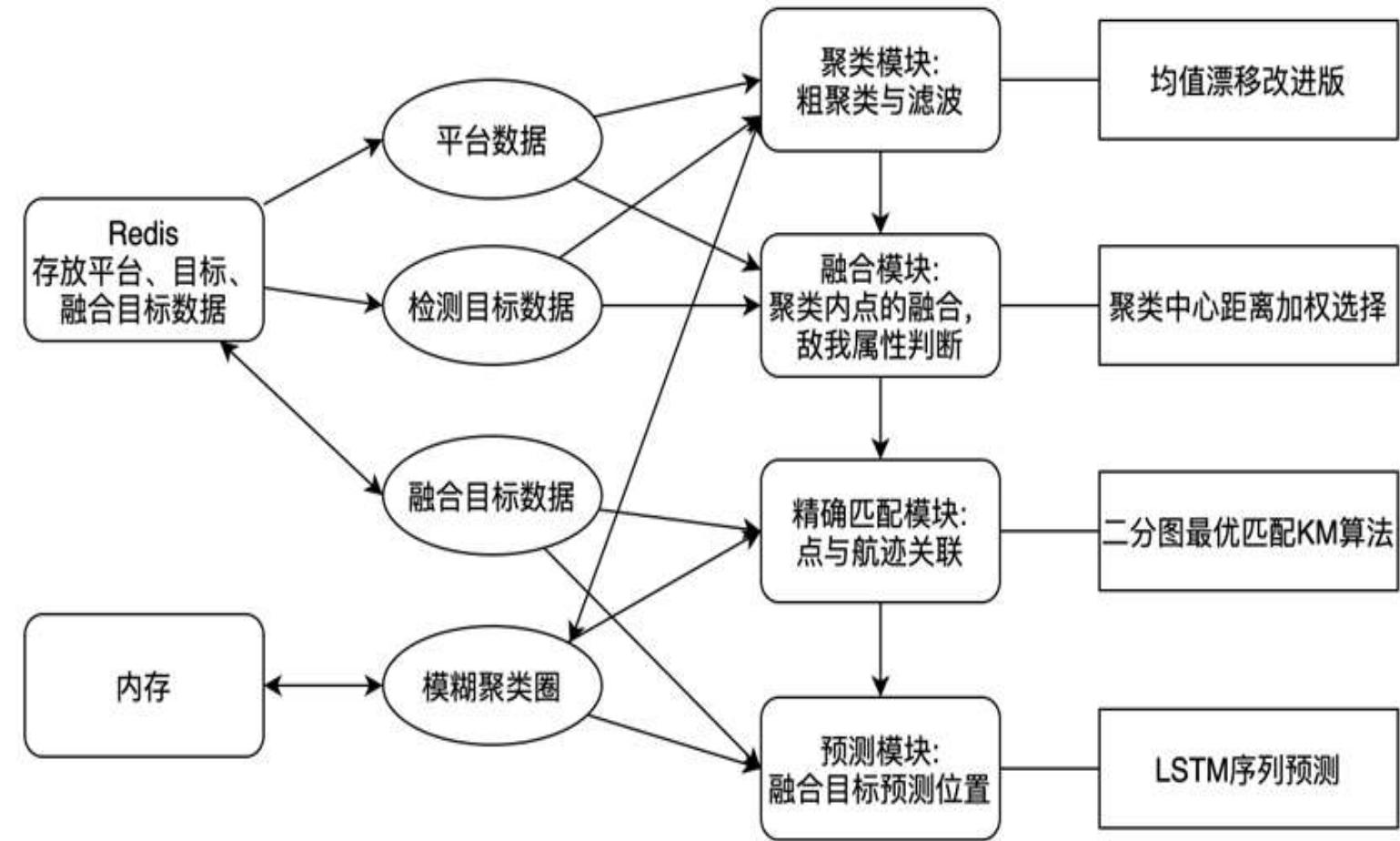
## 4.5 点云语义/实例分割



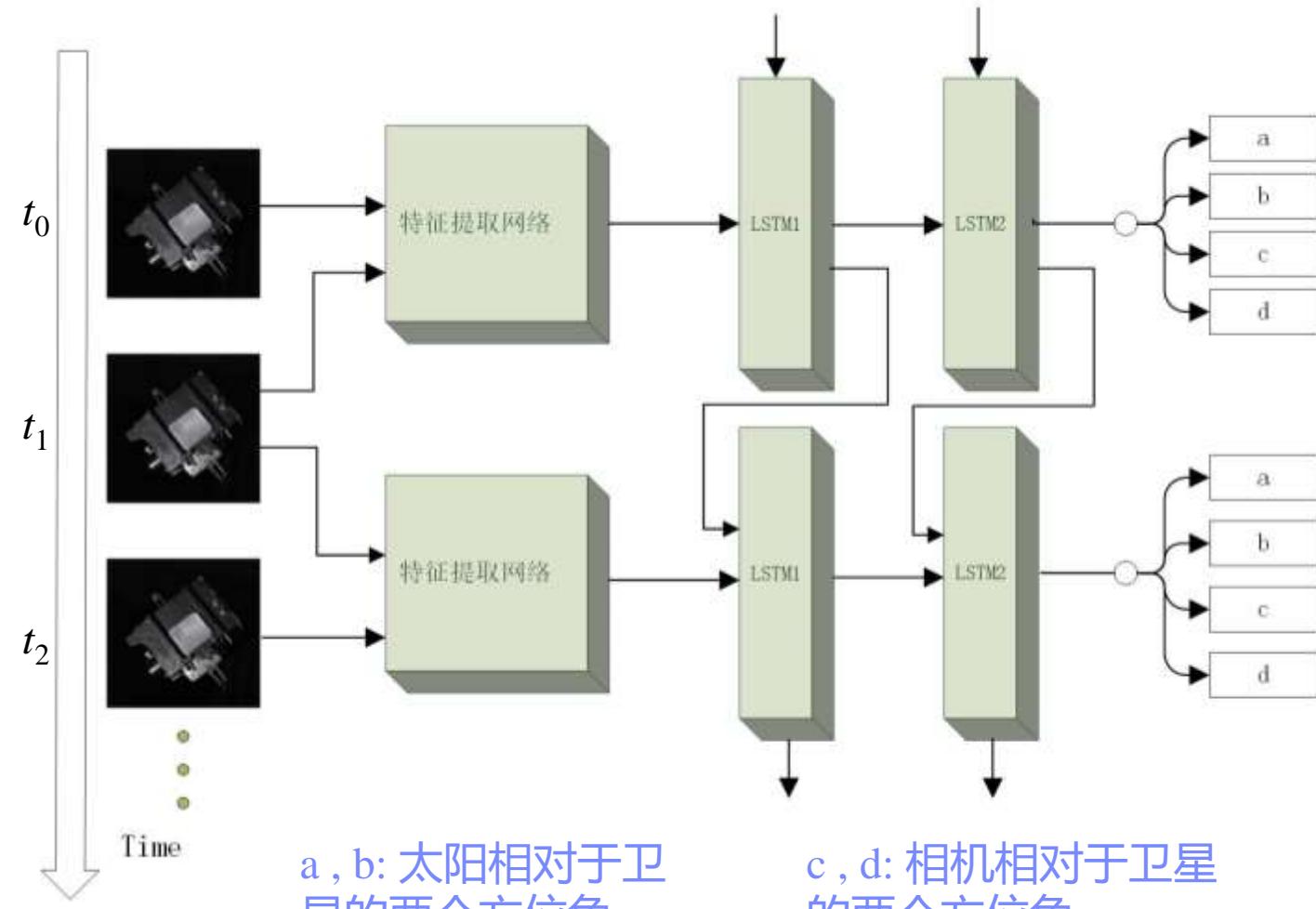
## 4.5 点云语义/实例分割



# 4.6 无人系统多域探测数据关联与融合



## 4.7 非合作目标姿态测量方法



## 4.8 \*\*概念设计方案快速生成

\*\*样本库



ZZ需求：攻击、防  
御、海区、航速、  
续航、排水量等

智能推理

No

方案可行

Yes

误差允许  
范围内？

\*\*概念设计方案集



VS (主尺度参数库)

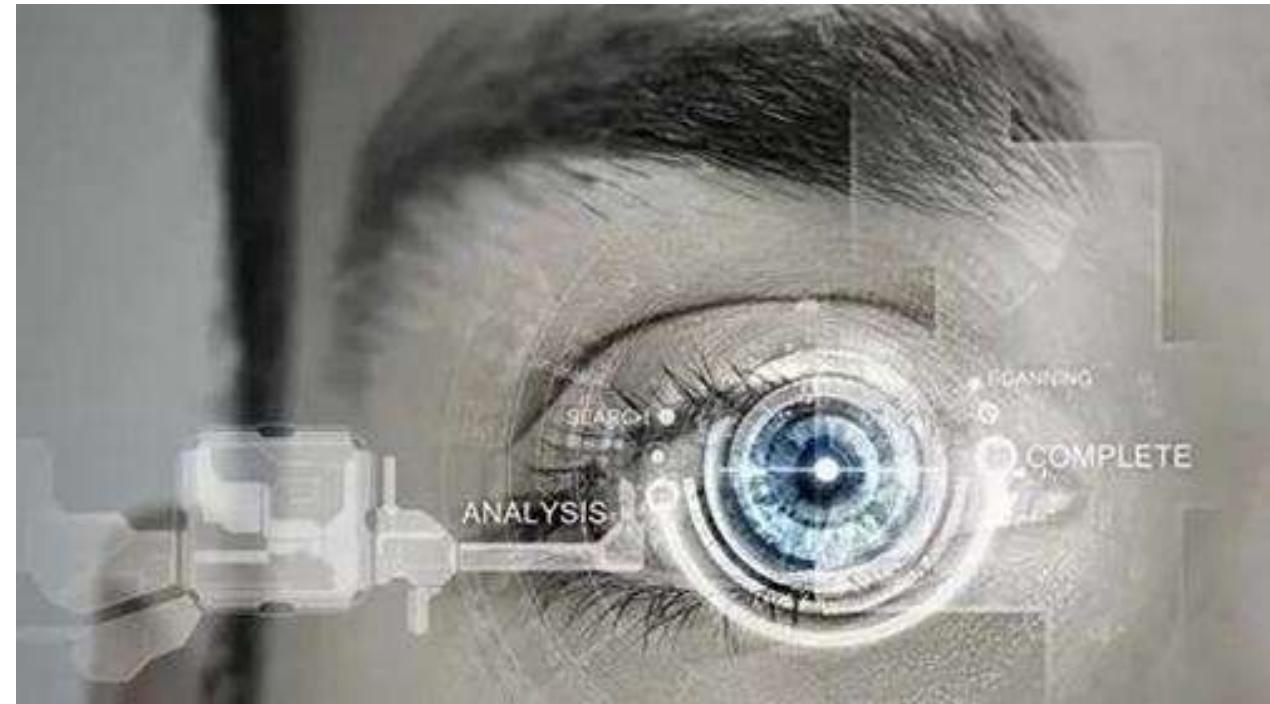


# 结语



机器视觉是人工智能正在快速发展的一个分支。机器视觉研究的是怎么用机器代替人眼来感知、理解和认知自己所处的空间与环境，让机器学会思考，进而将感知与行动有效连接起来。

机器视觉的落地应用离不开特定的应用场景，未来的发展趋势是深度（迁移、强化）学习+传统机器视觉技术。





谢谢大家

汇报人 王志成 同济博士  
单 位 同济大学CAD研究中心  
邮 手 地 址 上海曹安公路4800号  
箱 机 号 上海曹安公路4800号  
地 址 上海曹安公路4800号

汇报人 王志成 同济博士  
单 位 同济大学CAD研究中心  
邮 手 地 址 上海曹安公路4800号  
箱 机 号 上海曹安公路4800号  
地 址 上海曹安公路4800号  
汇报人 王志成 同济博士  
单 位 同济大学CAD研究中心  
邮 手 地 址 上海曹安公路4800号  
箱 机 号 上海曹安公路4800号  
地 址 上海曹安公路4800号