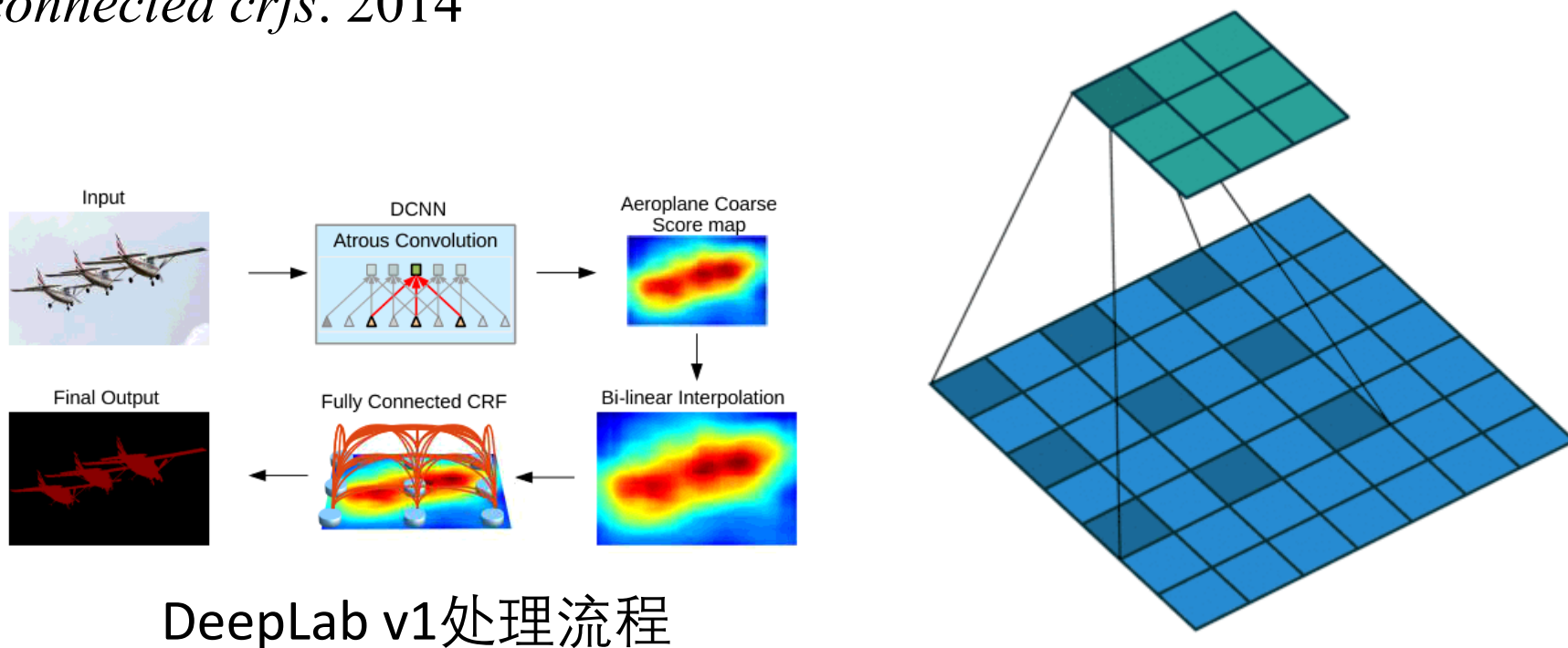


• 3.1 语义分割

DeepLab v1

Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, et al.,
Semantic image segmentation with deep convolutional nets and fully connected crfs. 2014



卷积核为3、扩张率为2和
无padding的二维空洞卷积

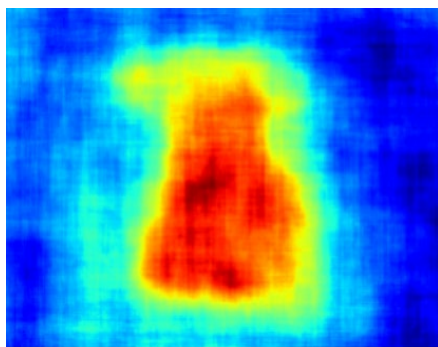
- 3.1 语义分割

DeepLab v1

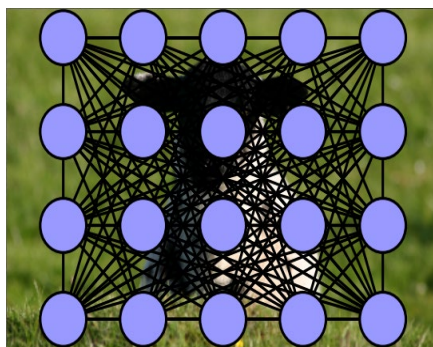
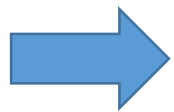
- DeepLabv1是在VGG16的基础上做了修改：
 - VGG16的全连接层转为卷积
 - 最后的两个池化层去掉了下采样
 - 后续卷积层的卷积核改为了空洞卷积
 - 在ImageNet上预训练的VGG16权重上做finetune

Conditional Random Fields(CRF)

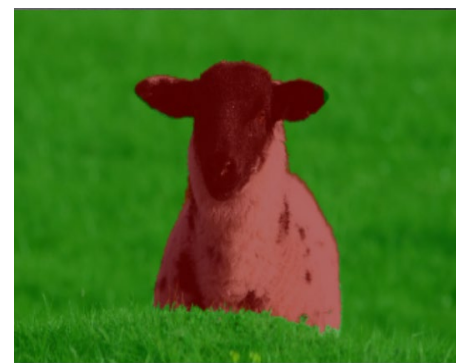
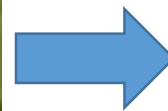
- FCNs, SegNet等基于CNN的分割算法是对每个像素进行相互独立的分类，分割时像素之间关系没有利用。
- CRF是一种概率图模型算法，此算法可以结合先验信息进行结构预测。



像素级分类概率



CRF 建模



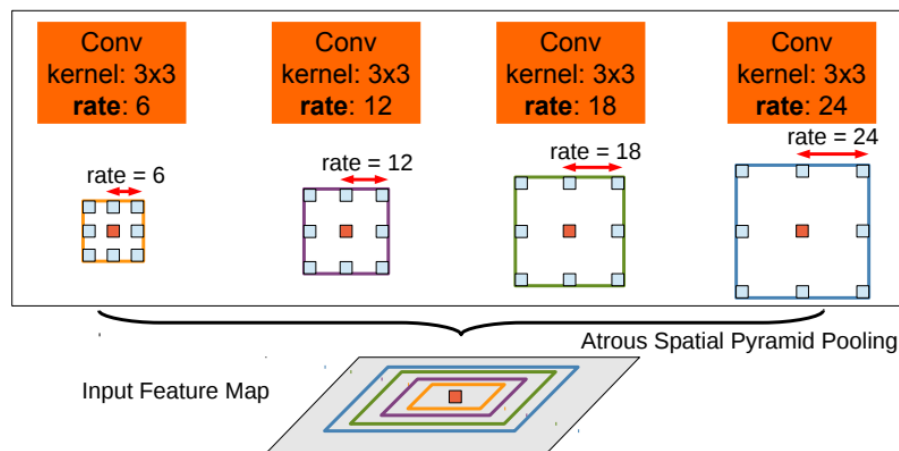
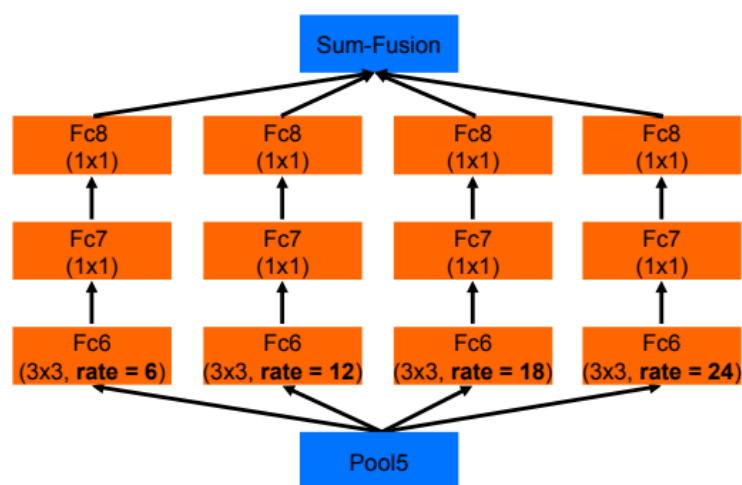
CRF输出

• 3.1 语义分割

DeepLab v2

Liang-Chieh Chen, George Papandreou, et al., *Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs*. IEEE transactions on PAMI, 2018.

多孔空间金字塔池化ASPP:



左图说明ASPP使用多个不同倍率的卷积核对池化结果进行卷积，实现从多尺度捕捉目标和上下文。右图进行一个直观的说明，图中黄色的点综合不同尺度的区域（不同颜色框）进行识别。

- 3.1 语义分割

DeepLab v2

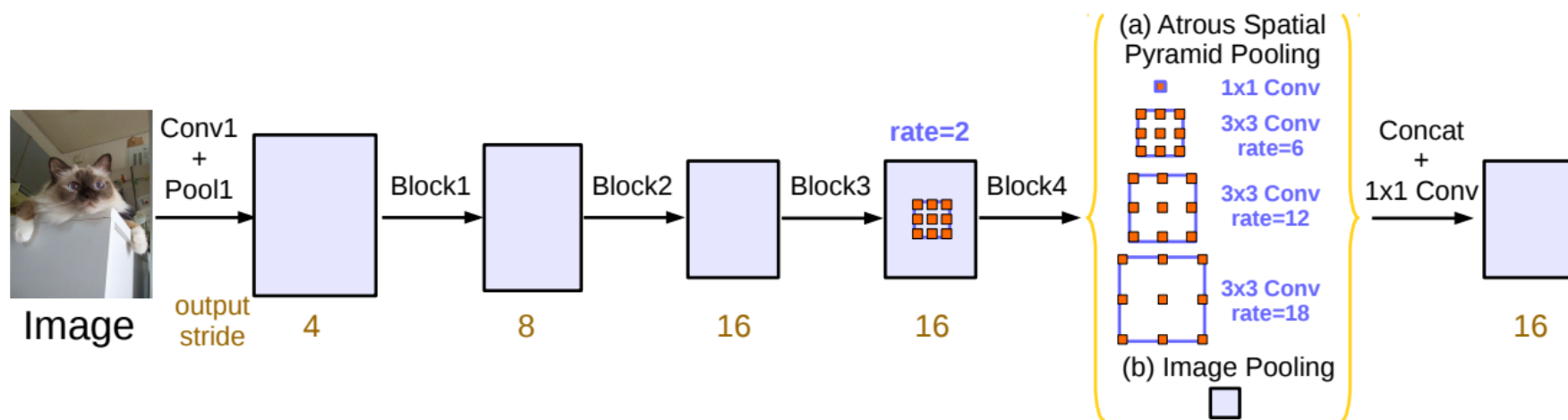
- DeepLab v2在v1的基础上进行改进

- 用多尺度获得更好的分割效果(使用ASPP)
- 基础层由VGG16转为ResNet
- 使用不同的学习策略(poly)

• 3.1 语义分割

DeepLab v3

Liang-Chieh Chen, George Papandreou, Florian Schroff, et al.,
Rethinking atrous convolution for semantic image segmentation., 2017.

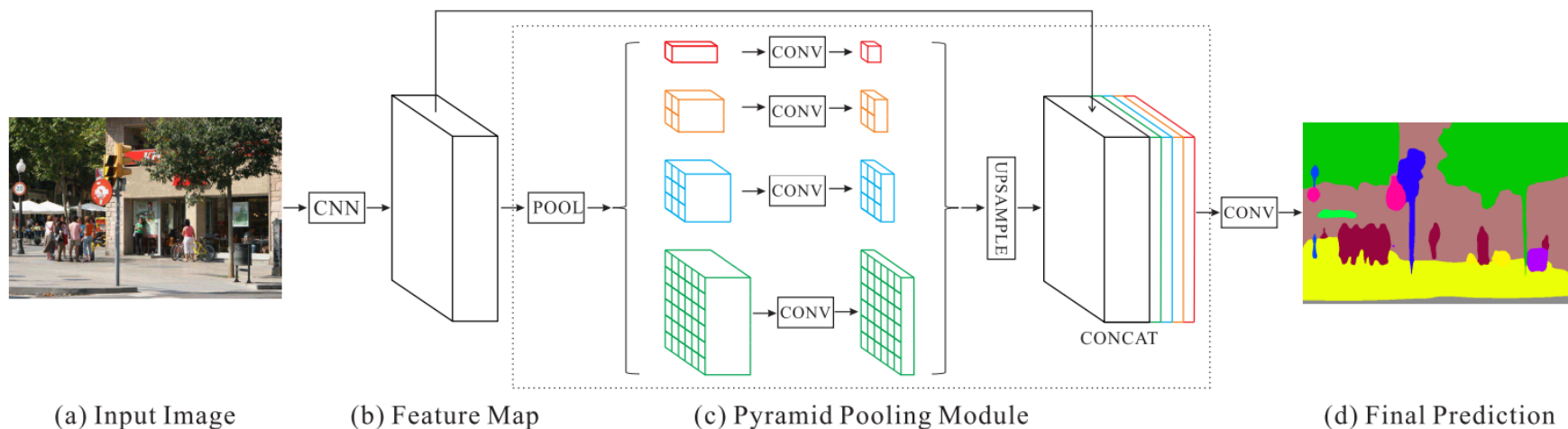


相较于v2对ASPP模块进行了改进，多了一个1x1的卷积操作和一个全局平均池化操作

• 3.1 语义分割

PSPNet

Zhao H., Shi J., Qi X., et al. *Pyramid scene parsing network*. in IEEE Conf. on CVPR. 2017.



PSPNet网络结构

主要是提出了一种**金字塔池化**（Global Pyramid Pooling），具有更好的多尺度信息

主要内容

- 1. 背景介绍
- 2. 传统图像分割
- 3. 基于深度学习的图像分割
 - 3.1 语义分割
 - 3.2 实例分割
- 4. 数据集和评价指标

• 3.2 实例分割

实例分割的定义：

图像的实例分割就是在语义分割的基础上，对同种物体的不同实例个体也用不同的类标进行标注。



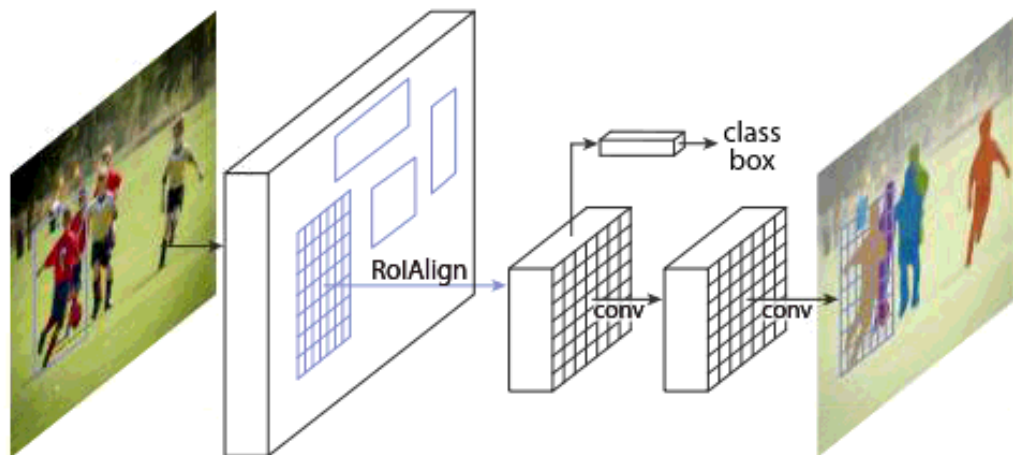
图中不仅将person和bottle进行分类，而且对同一类别（person或bottle）中的不同个体进行了区分，分别用不同颜色表示。

• 3.1 实例分割

Mask R-CNN

He K., Gkioxari G., Dollar P., et al. *Mask R-CNN*. in ICCV, 2017.

Facebook AI Research (FAIR)



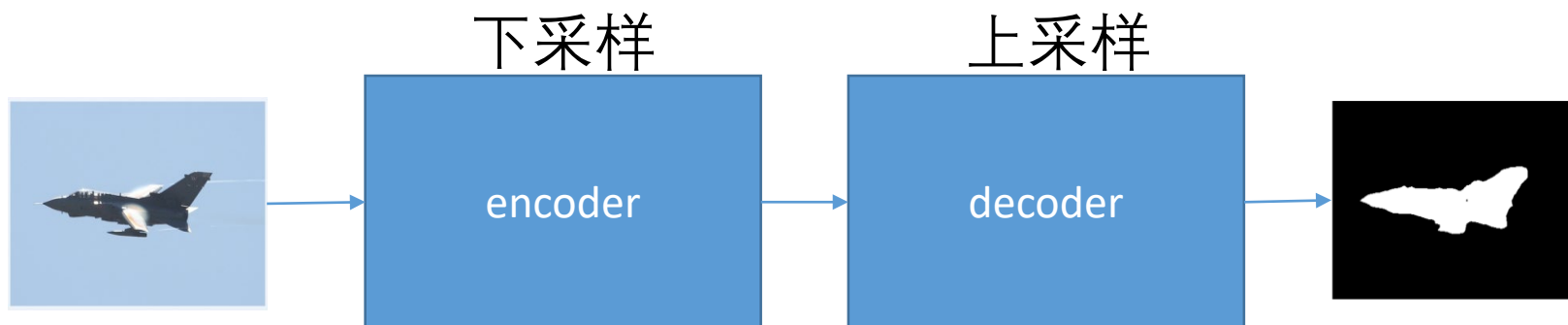
Mask R-CNN框架

使用多任务学习方法，结合了目标检测和语义分割任务，采用多分支网络框架

• 3. 基于深度学习的图像分割

图像分割网络总结

可以看出，这些网络结构都是非常相似的，都是基于encoder-decoder结构的，只不过说法不同。encoder一般是由卷积层、池化层构成的一个下采样过程，decoder则是一个上采样过程。这些网络的最大区别也就是上采样方法的不同。



主要内容

- 1. 背景介绍
- 2. 传统图像分割
- 3. 基于深度学习的图像分割
 - 3.1 语义分割
 - 3.2 实例分割
- 4. 数据集和评价指标

• 4. 数据集和评价指标

数据集

数据集中需要包括的信息：

原图



+

语义分割标签

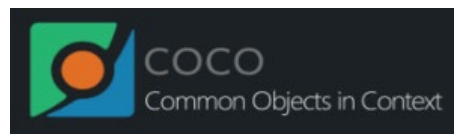


实例分割标签



常见数据集：

Visual Object Classes Challenge 2012 (VOC2012)



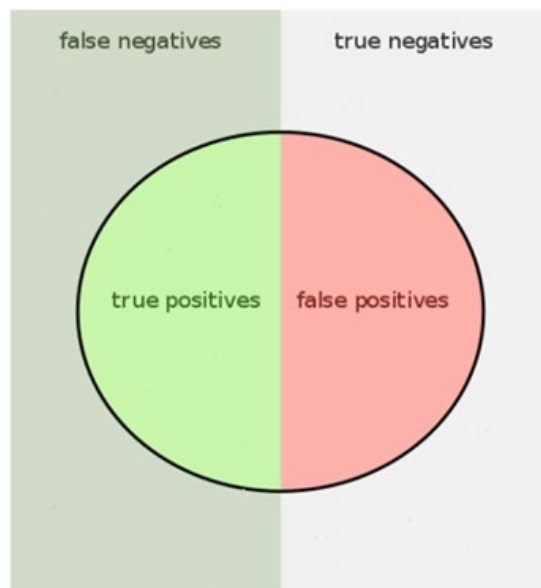
• 4. 数据集和评价指标

评价指标

如何判断图像分割预测结果的好坏，有以下几个常用的指标：

1. AP (Average Precision)

2. IU/IoU (Intersection over Union)



圆圈里表示预测的前景；
圆圈外表示预测的背景；
矩形左表示真实的前景；
矩形右表示真实的背景；

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$
$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

在图像分割中，AP则是先对一张图片中的某一类物体的预测结果计算precision，然后对所有图片求平均。对于多类别的分割任务而言，一般采用mean AP，即将AP对所有类别求平均。

• 4. 数据集和评价指标

评价指标

如何判断图像分割预测结果的好坏，有以下几个常用的指标：

1.AP (Average Precision)

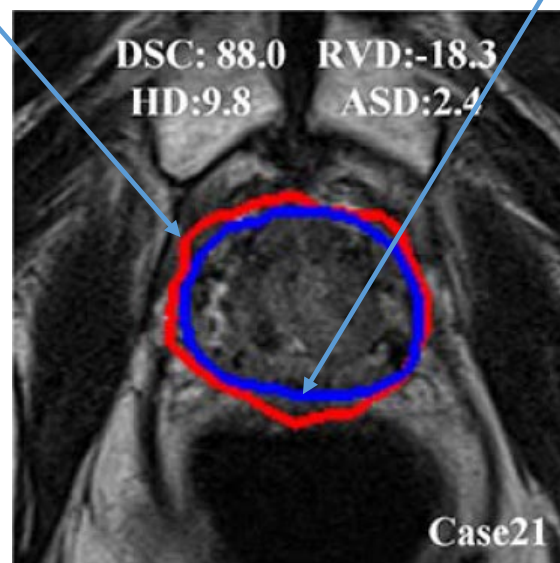
2.IU/IoU (Intersection over Union)

红色表示预测结果 蓝色表示真实结果

IU/IoU:

对于两个区域R和R'，其overlap
计算如下：

$$O(R, R') = \frac{|R \cap R'|}{|R \cup R'|}$$



二者覆盖的越好，表面结果越精确

		mean	aero plane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	dining table	dog	horse	motor bike	person	potted plant	sheep	sofa	train	tv mon
		▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	
▶	DeepLabv3+_JFT [?]	89.0	97.5	77.9	96.2	80.4	90.8	98.3	95.5	97.6	58.8	96.1	79.2	95.0	97.3	94.1	93.8	78.5	95.5	74.4	93.8	
▶	DeepLabv3+_AASPP [?]	88.5	97.4	80.3	97.1	80.1	89.3	97.4	94.1	96.9	61.9	95.1	77.2	94.2	97.5	94.4	93.0	72.4	93.8	72.6	93.3	
▶	SRC-B-MachineLearningLab [?]	88.5	97.2	78.6	97.1	80.6	89.7	97.4	93.7	96.7	59.1	95.4	81.1	93.2	97.5	94.2	92.9	73.5	93.3	74.2	91.0	
▶	MSCI [?]	88.0	96.8	76.8	97.0	80.6	89.3	97.4	93.8	97.1	56.7	94.3	78.3	93.5	97.1	94.0	92.8	72.3	92.6	73.6	90.8	
▶	ExFuse [?]	87.9	96.8	80.3	97.0	82.5	87.8	96.3	92.6	96.4	53.3	94.3	78.4	94.1	94.9	91.6	92.3	81.7	94.8	70.3	90.1	
▶	DeepLabv3+ [?]	87.8	97.0	77.1	97.1	79.3	89.3	97.4	93.2	96.6	56.9	95.0	79.2	93.1	97.0	94.0	92.8	71.3	92.9	72.4	91.0	
▶	DeepLabv3-JFT [?]	86.9	96.9	73.2	95.5	78.4	86.5	96.8	90.3	97.1	51.4	95.0	73.4	94.0	96.8	94.0	92.3	81.5	95.4	67.2	90.8	
▶	DIS [?]	86.8	94.0	73.3	93.5	79.1	84.8	95.4	89.5	93.4	53.6	94.8	79.0	93.6	95.2	91.5	89.6	78.1	93.0	79.4	94.3	
▶	** Gluon DeepLabV3 152 ** [?]	86.7	96.5	74.3	96.1	80.2	85.2	97.0	93.8	96.4	49.7	93.6	77.6	95.1	95.3	93.9	89.6	75.8	94.4	70.8	89.7	
▶	CASIA_IVA_SDN [?]	86.6	96.9	78.6	96.0	79.6	84.1	97.1	91.9	96.6	48.5	94.3	78.9	93.6	95.5	92.1	91.1	75.0	93.8	64.8	89.0	
▶	IDW-CNN [?]	86.3	94.8	67.3	93.4	74.8	84.6	95.3	89.6	93.6	54.1	94.9	79.0	93.3	95.5	91.7	89.2	77.5	93.7	79.2	94.0	
▶	DFN [?]	86.2	96.4	78.6	95.5	79.1	86.4	97.1	91.4	95.0	47.7	92.9	77.2	91.0	96.7	92.2	91.7	76.5	93.1	64.4	88.3	
▶	** GluonCV DeepLabV3 ** [?]	86.2	96.3	69.7	93.5	76.2	86.5	96.5	92.2	95.8	47.8	95.0	81.6	93.0	96.0	91.2	90.7	77.1	94.7	68.9	89.3	
▶	EncNet [?]	85.9	95.3	76.9	94.2	80.2	85.3	96.5	90.8	96.3	47.9	93.9	80.0	92.4	96.6	90.5	91.5	70.9	93.6	66.5	87.7	
▶	HPN [?]	85.8	94.1	67.0	95.2	81.9	88.3	95.5	90.4	95.9	40.0	92.7	82.5	91.7	95.3	92.6	91.6	73.6	94.1	69.4	91.1	
▶	DeepLabv3 [?]	85.7	96.4	76.6	92.7	77.8	87.6	96.7	90.2	95.4	47.5	93.4	76.3	91.4	97.2	91.0	92.1	71.3	90.9	68.9	90.8	
▶	** Auto-DeepLab-L ** [?]	85.6	96.5	77.3	94.8	74.1	84.0	97.1	88.7	94.5	53.5	91.6	79.2	88.4	94.2	90.2	91.2	75.1	90.1	70.7	89.1	
▶	** DP-CAN_decoder ** [?]	85.5	95.9	77.8	91.6	75.0	81.7	96.6	92.4	97.1	42.7	93.5	74.1	93.9	95.0	91.4	91.2	78.1	94.6	66.5	89.8	
▶	PSPNet [?]	85.4	95.8	72.7	95.0	78.9	84.4	94.7	92.0	95.7	43.1	91.0	80.3	91.3	96.3	92.3	90.1	71.5	94.4	66.9	88.8	
▶	** GluonCV PSP ** [?]	85.1	95.7	70.9	92.8	75.6	85.0	96.5	91.7	95.0	41.8	92.3	78.8	90.4	95.6	93.4	90.6	76.1	93.5	66.7	89.5	
▶	** ResNet-38_COCO ** [?]	84.9	96.2	75.2	95.4	74.4	81.7	93.7	89.9	92.5	48.2	92.0	79.9	90.1	95.5	91.8	91.2	73.0	90.5	65.4	88.7	
▶	** DP-CAN ** [?]	84.6	96.5	77.7	87.6	73.9	79.9	96.8	92.9	95.7	40.8	92.9	74.0	91.7	95.0	92.5	89.7	77.2	94.6	64.6	90.2	
▶	** resnet 101 + fast laddernet ** [?]	84.2	95.4	73.9	94.9	75.7	83.2	96.3	91.2	93.9	35.3	90.0	79.4	90.2	94.2	92.8	90.1	73.2	92.3	64.5	88.0	

<http://host.robots.ox.ac.uk:8080/leaderboard/displaylb.php?cls=mean&challengeid=11&compid=6&submid=19760>

参考资料

- 1.反卷积: <https://blog.csdn.net/itleaks/article/details/80336825>
- 2.空洞卷积: <http://www.dataguru.cn/article-11860-1.html>
- 3.深度学习语义分割发展史: https://www.sohu.com/a/155907339_610300
- 4.图像分割综述: <https://blog.csdn.net/Julialove102123/article/details/80493066>
- 5.分水岭分割算法: <https://www.cnblogs.com/mikewolf2002/p/3304118.html>
- 6.实例分割图: <https://blog.csdn.net/lanyuxuan100/article/details/70800246>
- 7.CS231n课程视频: <http://cs231n.stanford.edu/>