**A Simple Survey of Image Segmentation**

**➀Image Segmentation**

图像分割是指根据灰度、彩色、空间纹理、几何形状等特征把图像划分成若干个互不相交的区域，使得这些特征在同一区域内表现出一致性或相似性，而在不同区域间表现出明显的不同。

**➁Semantic Segmentation and Instance Segmentation**

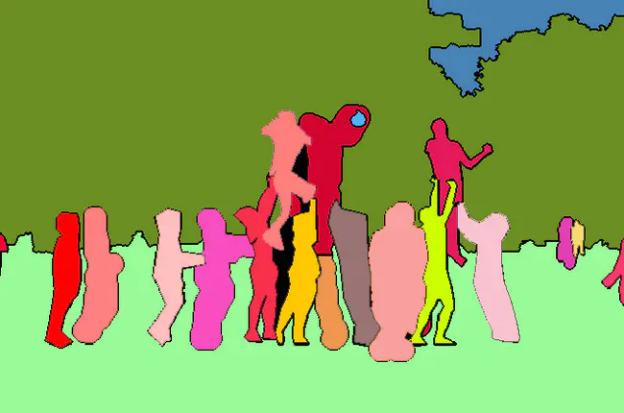
语义分割是在像素级别上的分类，属于同一类的像素都要被归为一类

实例分割是在语义分割的基础上更为细致的分割



**➂Panoramic Segmentation(全景分割)**

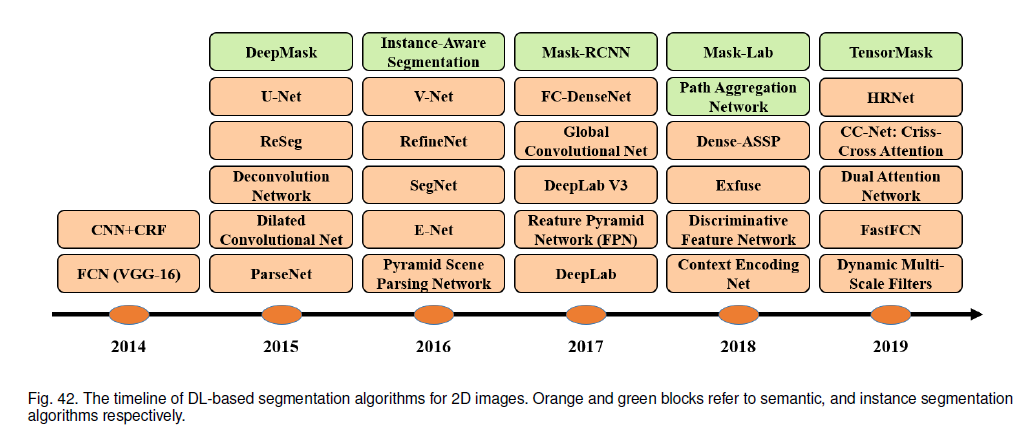
全景分割是对图中的所有物体包括背景都要进行检测和分割。



**总结：类别之间的区分、个体之间的区分**

**Based Timeline**

Figure 42 illustrates the timeline of popular DL-based works for semantic segmentation, as well as instance segmentation since 2014. Given the large number of works developed in the last few years, we only show some of the most representative ones.



**Based technical contributions**

1) Fully convolutional networks

2) Convolutional models with graphical models

3) Encoder-decoder based models

4) Multi-scale and pyramid network based models

5) R-CNN based models (for instance segmentation)

6) Dilated convolutional models and DeepLab family

7) Recurrent neural network based models

8) Attention-based models

9) Generative models and adversarial training

10) Convolutional models with active contour models

**Instance segmentation（lastest paper）**

1. PolarMask: Single Shot Instance Segmentation with Polar Representation CVPR2020
2. CenterMask : Real-Time Anchor-Free Instance Segmentation CVPR2020
3. Deep Snake for Real-Time Instance Segmentation CVPR2020

4、TensorMask: A Foundation for Dense Object Segmentation 2019 ICCV

5、YOLACT：Real-time Instance Segmentation 2019 CVPR

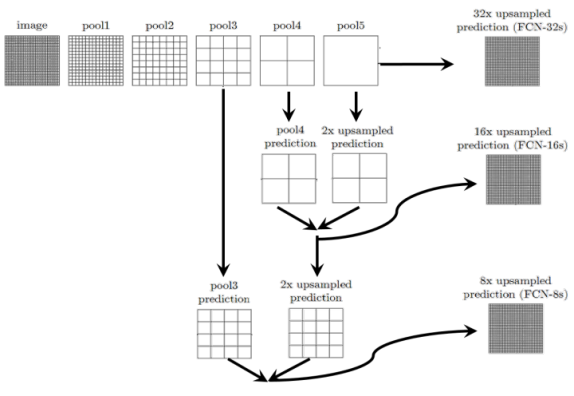
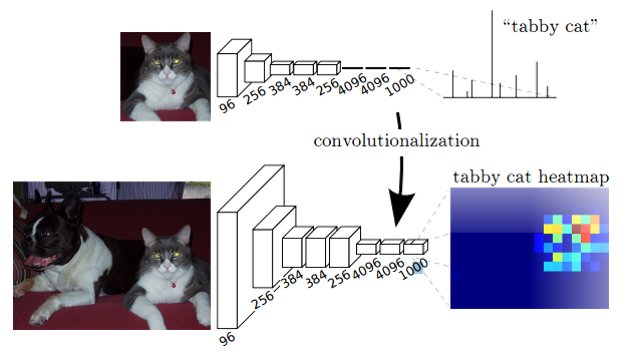
**Image segmentation’s Development**

**1、FCN(CVPR2015 best paper)**

➀卷积化（Convolutional）：常用网络架构VGG、ResNet、MobileNet等

➁上采样（Upsample）：引入反卷积Deconvolution(一对多，相乘相加运算)

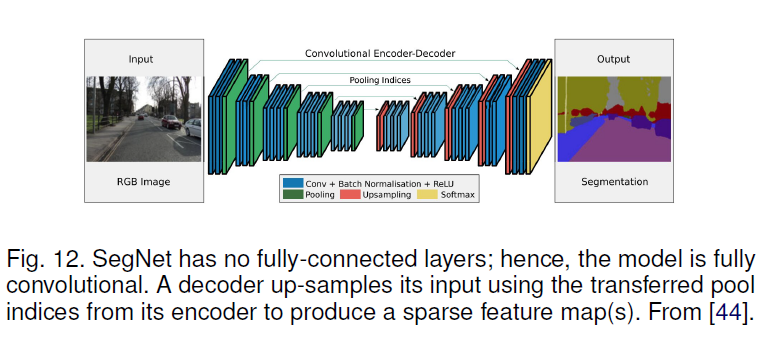
➂连接结构（Skip Layer）： 融合浅层的细粒度（fine-grain）信息与深层的粗糙（coarse）信息



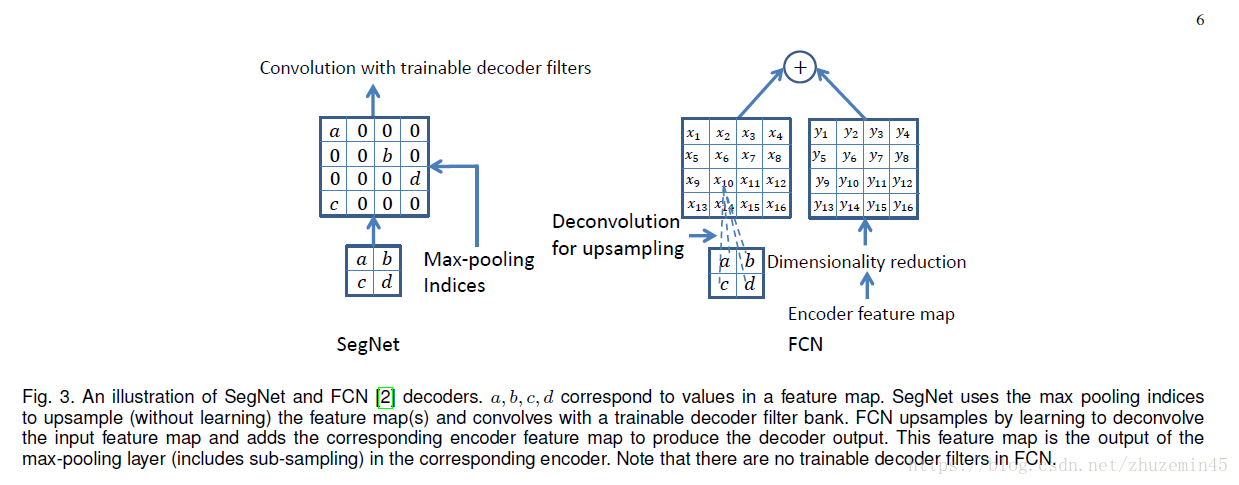
1. **SegNet（encoder-decoder based FCN）**

➀Encoder采用VGG-16的结构，记录池化窗口中的最大值

➁Decoder中的upsampling 产生稀疏特征图—>Conv 填充缺失值+ Softmax



CamVid road scenes datase：是一个小型数据集，其中包含有367训练样本，233涨测试样本，均为360X480分辨率的RGB图像



1. **Mask-RCNN（ICCV2017 best paper）（检测+分割）**

目标检测，直接在结果图上绘制了目标框(bounding box)。

目标分类，对于每一个目标，需要找到对应的类别(class)。

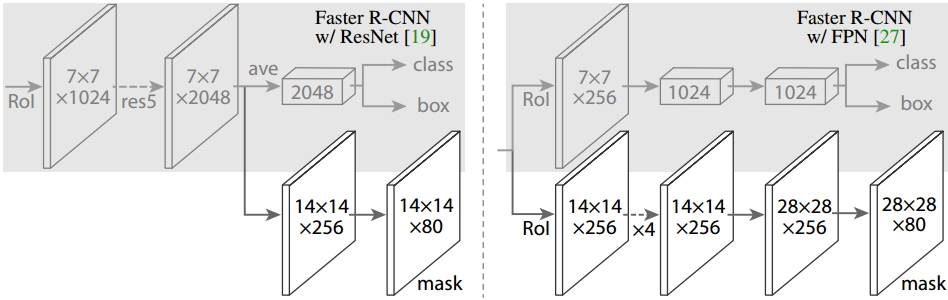
像素级目标分割，在每个目标中在像素层面区分前景和背景。

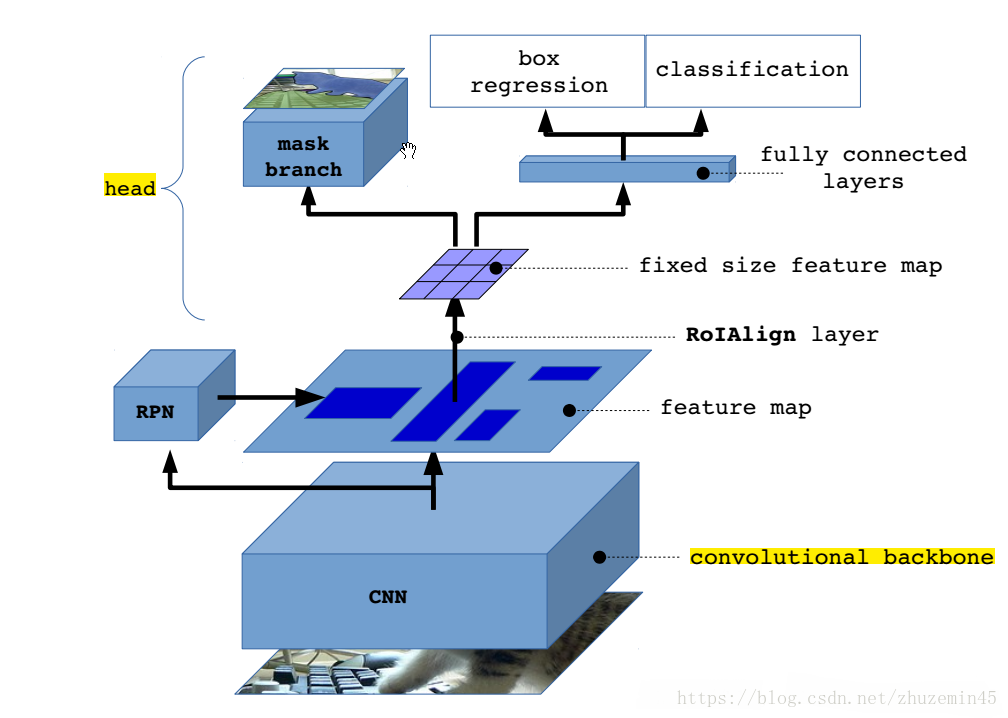
➀ResNet-101+FPN作为特征提取网络

➁ROIAlign解决Mis-alignment 的问题

➂增加了mask的Loss Function

**Mutil-tasks：object detection/classify/像素级目标分割**

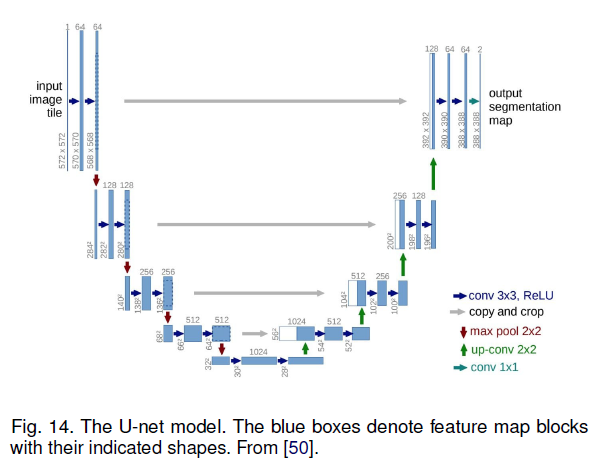




1. **U-Net（encoder-decoder）**

➀拼接形式的特征融合，而不是常规的对应点相加

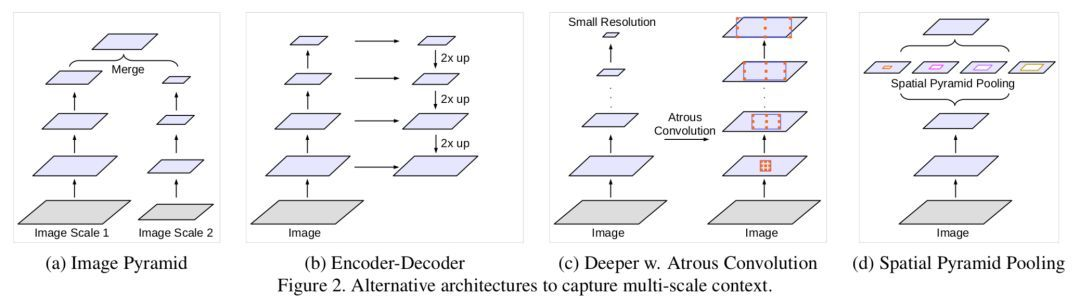
➁多尺度特征的融合



1. **DeepLab系列（空洞卷积）**

**➀**空洞卷积+全连接CRF+ASPP模块——v2

在给定的输入上以不同采样率的空洞卷积并行采样，相当于以多个比例捕捉图像的上下文，称为 ASPP (atrous spatial pyramid pooling)



**CRF：条件随机场(Conditional Random Fields, 以下简称CRF)是给定一组输入序列条件下另一组输出序列的条件概率分布模型，在自然语言处理中得到了广泛应用。**

随机场是由若干个位置组成的整体，当给每一个位置中按照某种分布随机赋予一个值之后，其全体就叫做随机场。还是举词性标注的例子：假如有一个十个词形成的句子需要做词性标注。这十个词每个词的词性可以在已知的词性集合（名词，动词...)中去选择。当我们为每个词选择完词性后，这就形成了一个随机场。

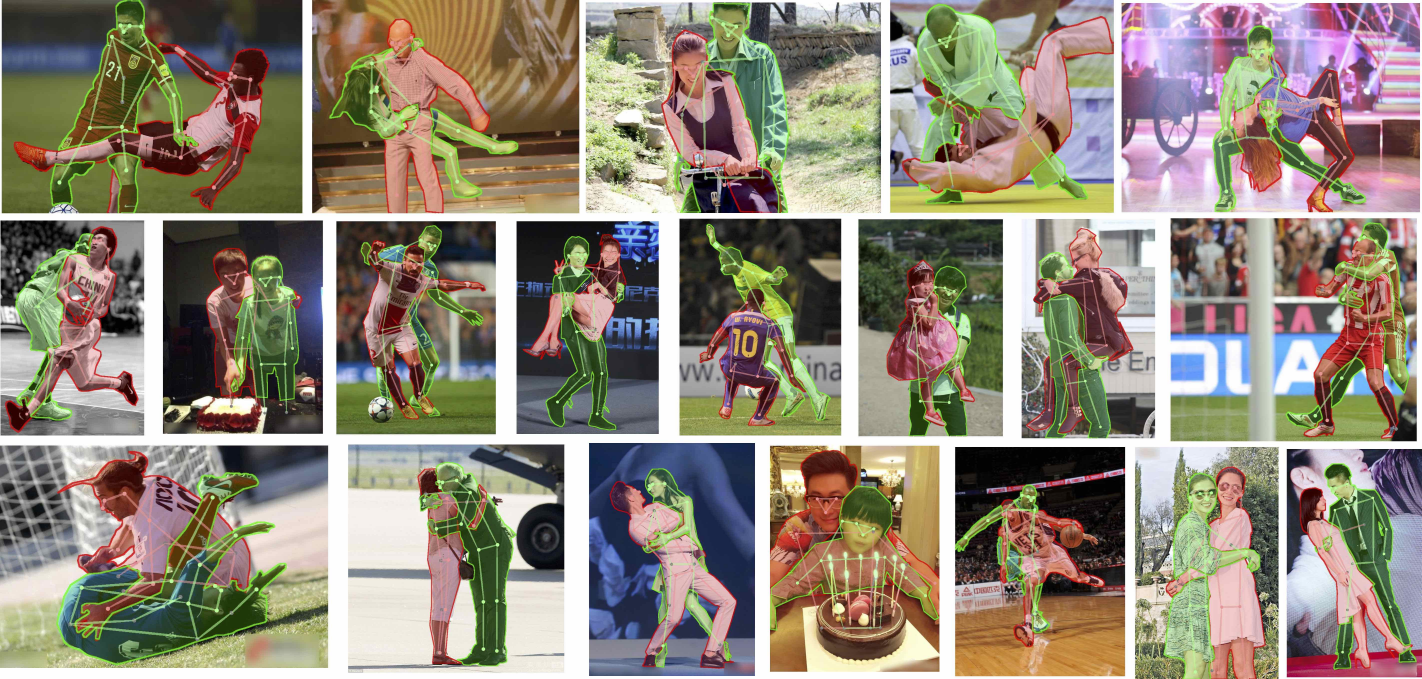
了解了随机场，我们再来看看马尔科夫随机场。马尔科夫随机场是随机场的特例，它假设随机场中某一个位置的赋值仅仅与和它相邻的位置的赋值有关，和与其不相邻的位置的赋值无关。继续举十个词的句子词性标注的例子：　如果我们假设所有词的词性只和它相邻的词的词性有关时，这个随机场就特化成一个马尔科夫随机场。比如第三个词的词性除了与自己本身的位置有关外，只与第二个词和第四个词的词性有关。

对于CRF，给出准确的数学语言描述：设X与Y是随机变量，P(Y|X)是给定X时Y的条件概率分布，若随机变量Y构成的是一个马尔科夫随机场，则称条件概率分布P(Y|X)是条件随机场。

Pose2Seg: Detection Free Human Instance Segmentation

Idea：基于人体姿势而不是提议区域检测来分离实例，处理遮挡问题

Datasets：Ochuman



**2020.06.01—2020.06.05—周报—吕锋**

**本周工作汇总**

·熟悉办公环境，配置工作软件；

·了解图像分割概念，阅读图像分割相关文献综述（30页），了解了该领域中的基本算法框

架，以及经典的网络架构；

·研读了近几年顶会文章（Mask-RCNN、UNet、SegNet、Deeplab等），加深了一些理解；

·目前在精读实时实例分割相关论文：YOLACT（ICCV2019）。

**下周工作安排**

·复现论文YOLACT

·继续精读以下论文

**Instance segmentation（lastest paper）**

1、PolarMask: Single Shot Instance Segmentation with Polar Representation CVPR2020

2、CenterMask : Real-Time Anchor-Free Instance Segmentation CVPR2020

3、Deep Snake for Real-Time Instance Segmentation CVPR2020

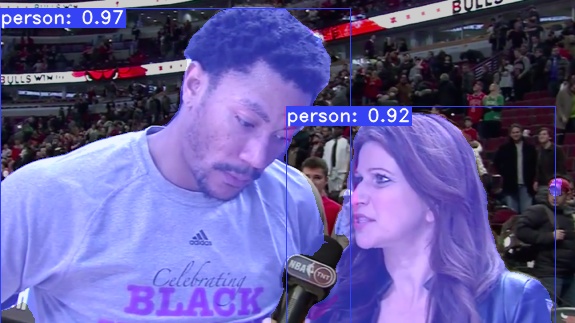
4、TensorMask: A Foundation for Dense Object Segmentation 2019 ICCV

**2020.06.08—2020.06.12—周报—吕锋**

**本周工作汇总**

·延续上周的工作，本周主要熟悉该工程的代码，在服务器上可训练。在本地进行了简单的调试后进行了测试，与论文中的观点相吻合，对于一些有遮挡的分割，效果较差；其次，对于代码中使用双线性插值来放大图片与其他方法做了一组对比实验，结果表明论文方法比较合理；最后，有部分代码还有待琢磨；

·阅读了最新论文：CenterMask（CVPR2020，美团），精读报告下周分享。



**下周工作安排**

·精读上述论文，工程复现，提出一些想法。

**2020.06.15—2020.06.19—周报—吕锋**

**本周工作汇总**

（1）精度论文CenterMask，对其结构组成Backbone所采用的的VoVNet展开学习（结合了ResNet与DenseNet，效果很好，后续工作可以将其作为基准展开），此外还学习无锚检测器FCOS（很多细节没看懂，需要看工程代码）。

（2）复现了CenterMask，但是由于缺少ImageNet预训练模型，测试效果很差，但该问题目前已可解。复现过程中的另一个问题，由于该工程基于Detectron平台实现，代码高度集成化，阅读修改较为缓慢。

（3）对现有的数据集进行简单的整理（查看标注格式精细程度等）

**下周工作安排**

（1）对数据集进行初步的评估整理

（2）争取把代码弄懂，整理好思路