分割

Segmentation

分割

- 介绍
- 发展
- 前沿

▶分割 Segmentation

分类级别

Semantic segmentation

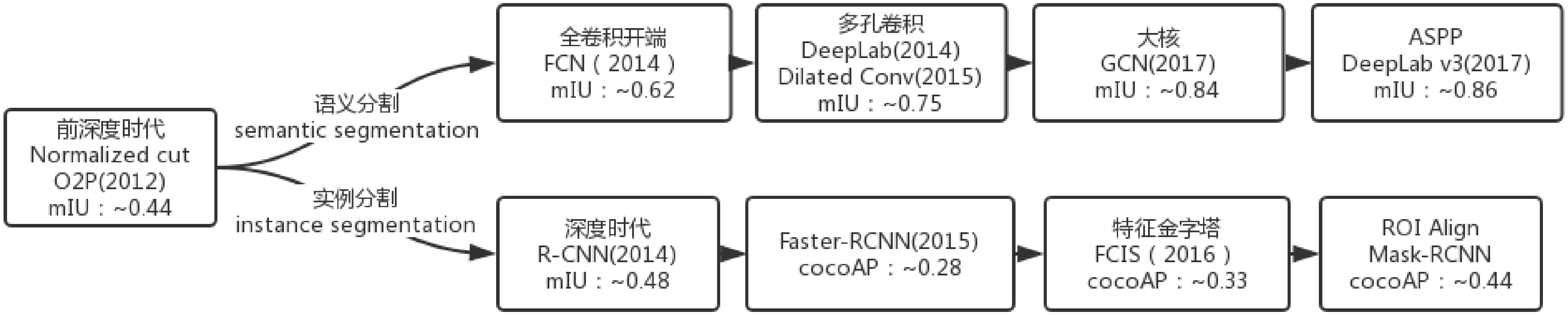
实例级别

Instance segmentation

▶ 分割

分割的发展

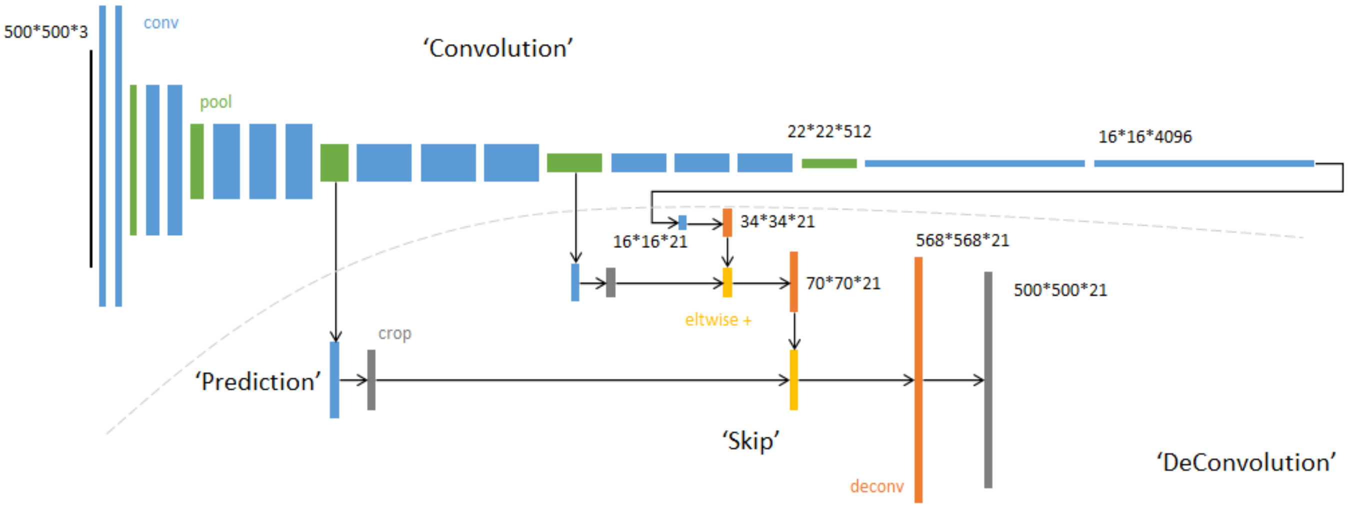
不 止 于 代 码



前深度时代
Normalized cut
O2P(2012)
mIU：~0.44
语义分割
semantic segmentation
全卷积开端
FCN（2014）
mIU：~0.62
多孔卷积
DeepLab(2014)
Dilated Conv(2015)
mIU：~0.75
大核
GCN(2017)
mIU：~0.84
ASPP
DeepLab v3(2017)
mIU：~0.86
实例分割
instance segmentation
深度时代
R-CNN(2014)
mIU：~0.48
Faster-RCNN(2015)
cocoAP：~0.28
特征金字塔
FCIS（2016）
cocoAP：~0.33
ROI Align
Mask-RCNN
cocoAP：~0.44

FCN

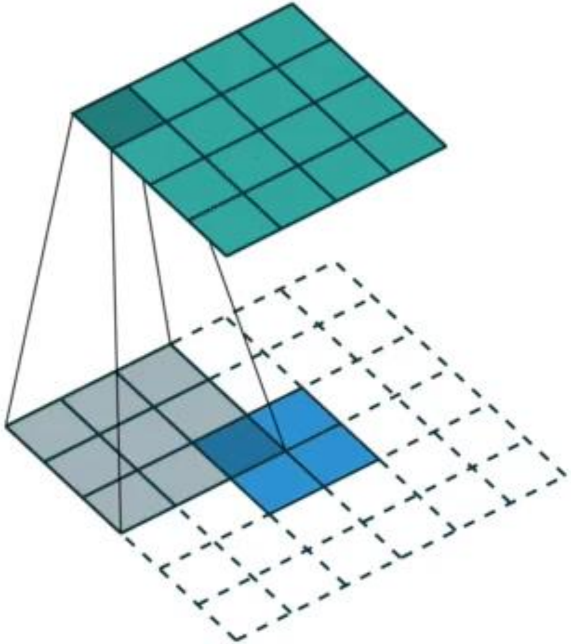
- Fully Convolutional Networks for Semantic Segmentation arXiv:1411.4038
- 全卷积
- 多层feature跳接结构（Skip Architecture）



FCN

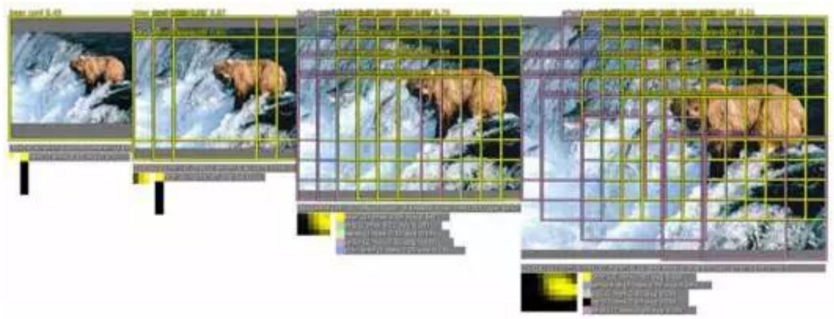
- Fully Convolutional Networks for Semantic Segmentation arXiv:1411.4038
- 全卷积
- 多层feature跳接结构（Skip Architecture）

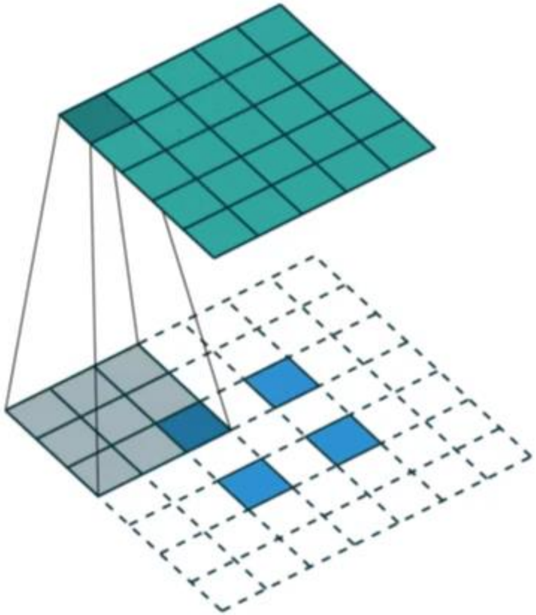
0.0625,	0.1875,	0.1875,	0.0625
0.1875,	0.5625,	0.5625,	0.1875
0.1875,	0.5625,	0.5625,	0.1875
0.0625,	0.1875,	0.1875,	0.0625



FCN

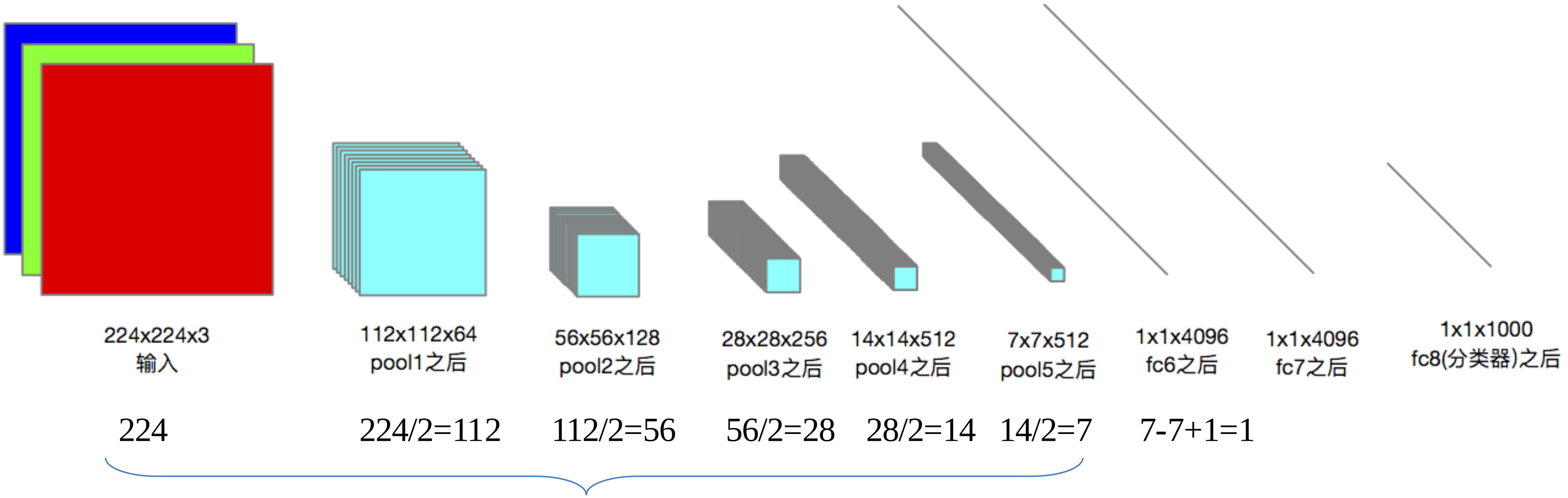
- Fully Convolutional Networks for Semantic Segmentation arXiv:1411.4038
- 全卷积
- 多层feature跳接结构（Skip Architecture）





FCN

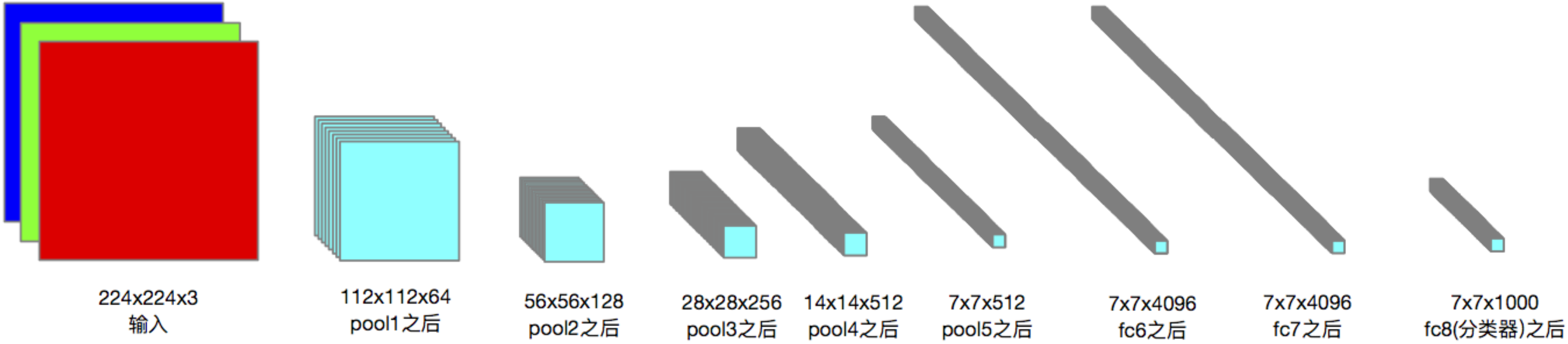
- 回忆一下经典的VGG-16



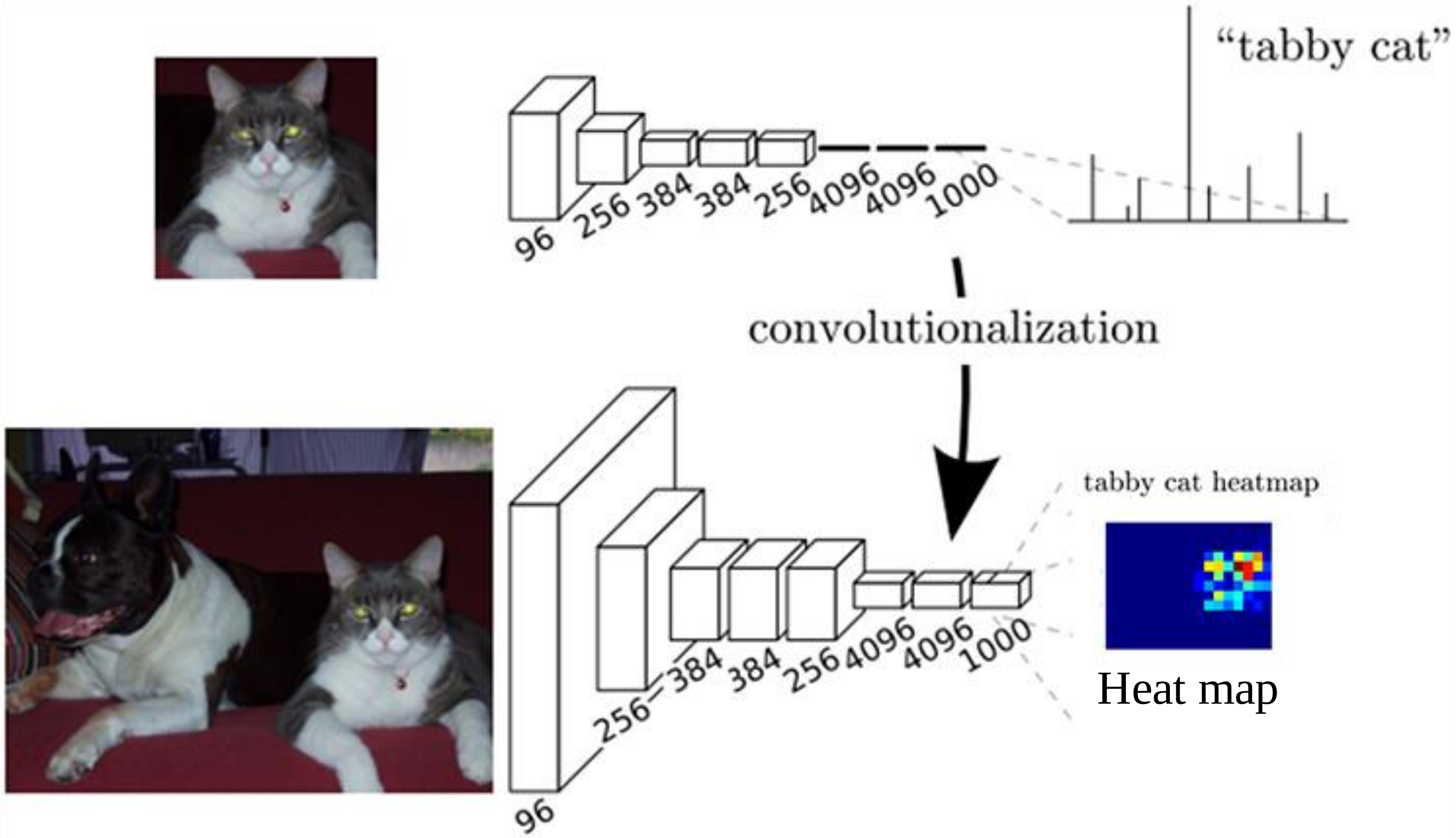
空间（spatial）尺度缩小了2^6=32倍

FCN

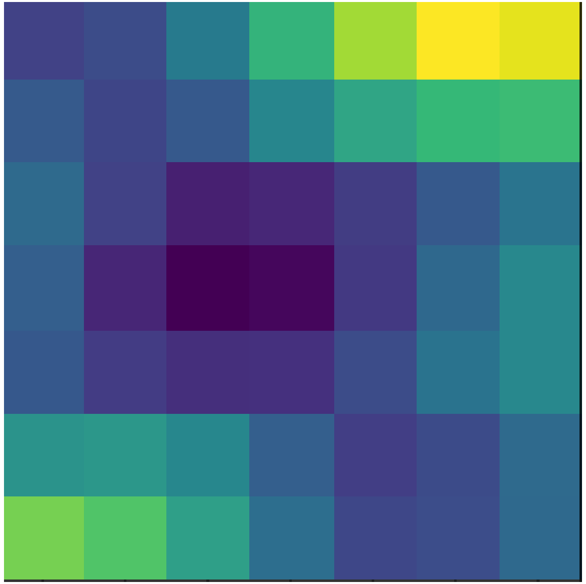
- 将最后的7x7 conv（fc6）padding改为SAME



FCN

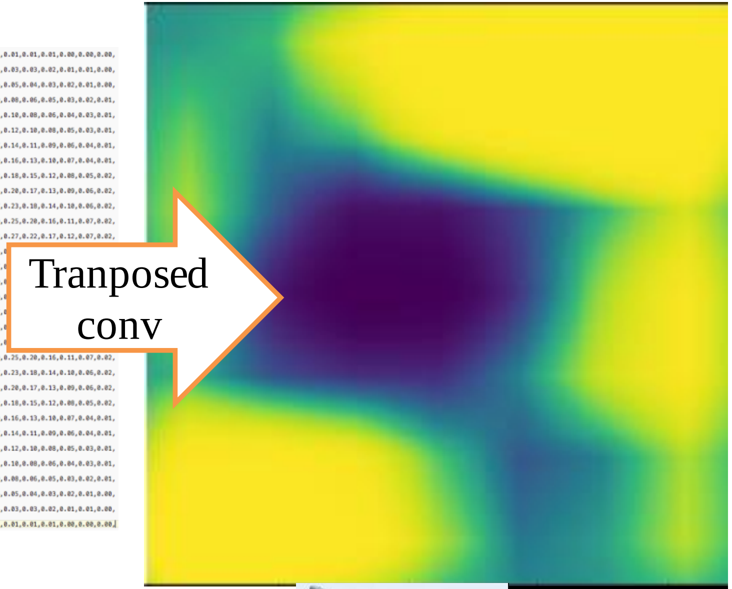

"tabby cat"
96
256
384
384
256
4096
4096
1000
convolutionalization
tabby cat heatmap
Heat map
96
256
384
384
256
4096
4096
1000

FCN



Heat map

Bilinear interpolation

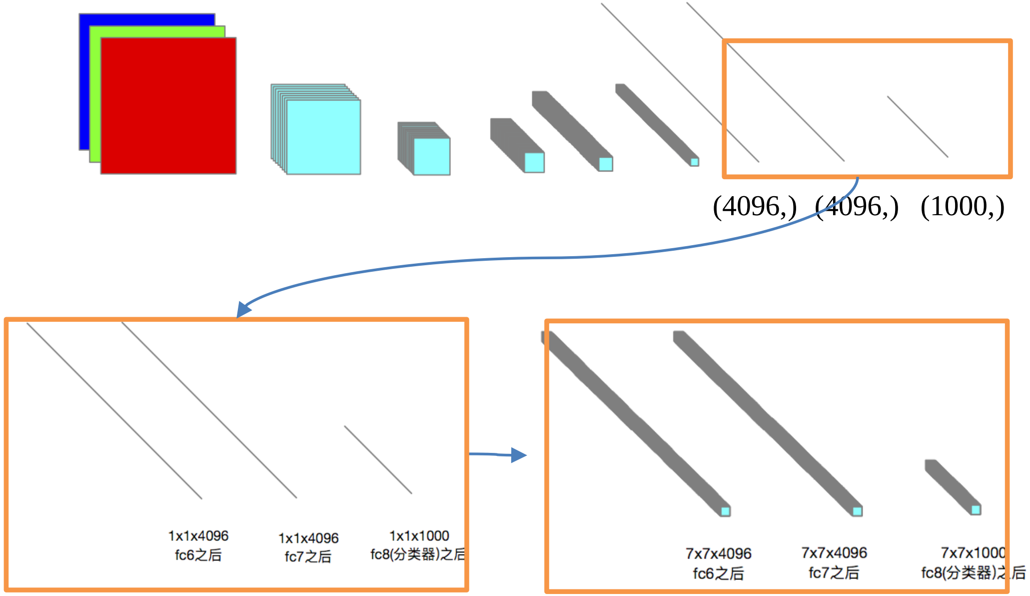


FCN-CRF（conditional random fields）

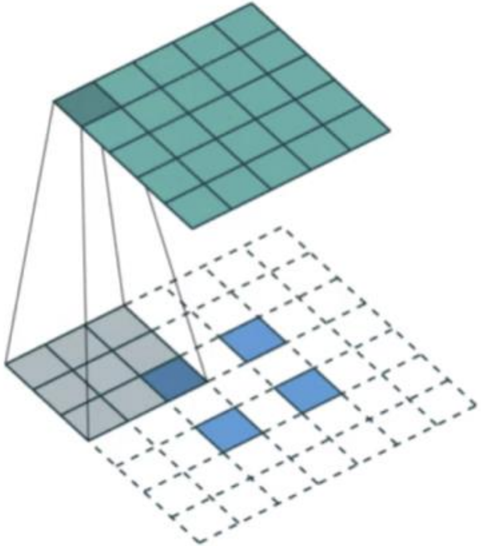
$$P(\mathbf{X}=\mathbf{x}|\mathbf{I})=\frac{1}{Z(\mathbf{I})}\exp(-E(\mathbf{x}|\mathbf{I}))$$

$$E(\mathbf{x})=\sum_{i}\Psi_u(x_i)+\sum_{i<j}\Psi_p(x_i,x_j)$$

FCN



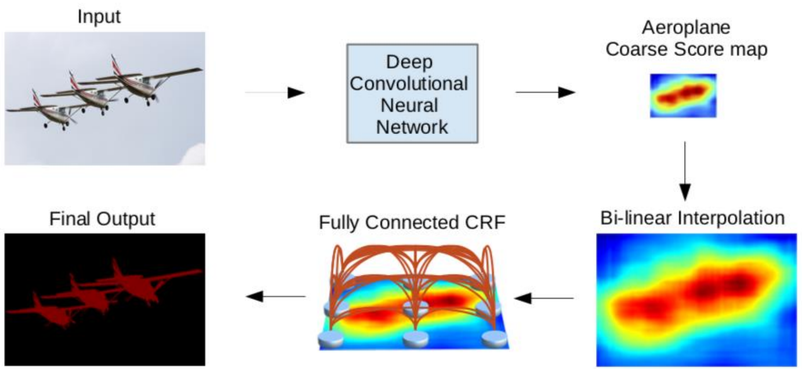
+

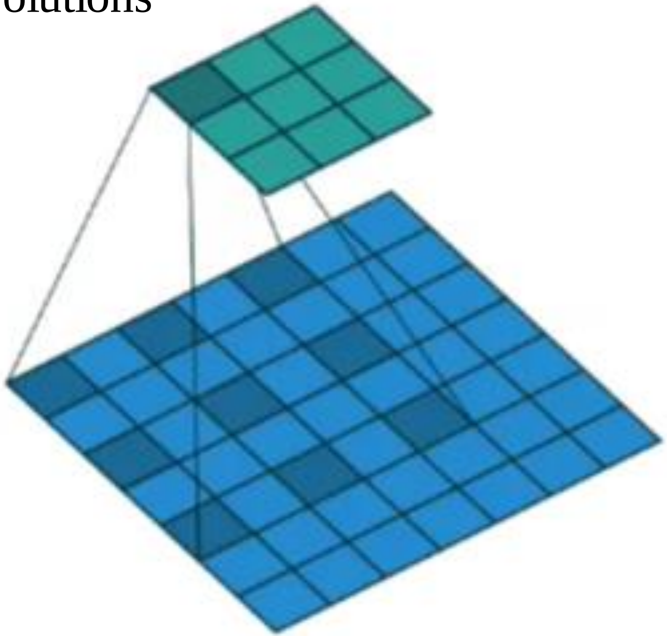


$$+ \quad P(\mathbf{X}=\mathbf{x}|\mathbf{I}) = \frac{1}{Z(\mathbf{I})}\exp(-E(\mathbf{x}|\mathbf{I}))$$

A’trous conv/Dilated conv

- Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs arXiv:1412.7062
- Multi-Scale Context Aggregation by Dilated Convolutions arXiv:1511.01722
- 利用多孔卷积替代掉原有VGG的Conv





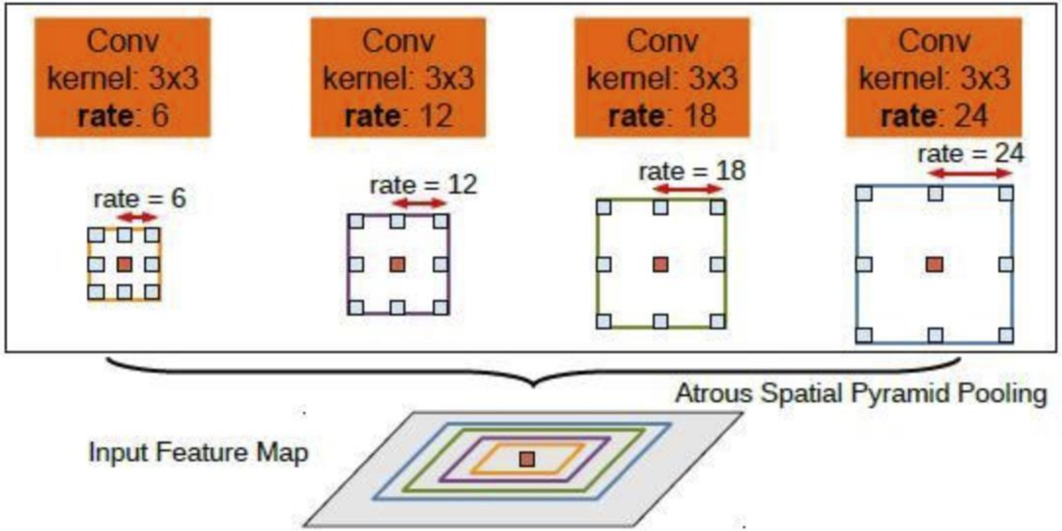
Atrous Conv

不 止 于 代 码

```
net = slim.max_pool2d(net, [2, 2], scope='pool3')
net = slim.repeat(net, 3, slim.conv2d, 512, [3, 3], scope='conv4')
net = slim.max_pool2d(net, [1, 1], stride=1, scope='pool4')
with tf.variable_scope('conv5'):
  with tf.variable_scope('conv5_1'):
    kernel = tf.Variable(tf.truncated_normal(shape=[3, 3, 512, 512]),
        name='weights')
    biases = tf.Variable(tf.zeros([512]), name='biases')
    net = tf.nn.atrous_conv2d(net, kernel,
        rate=2, padding='SAME', name='conv2d')
    net = tf.nn.relu(net+biases)
```

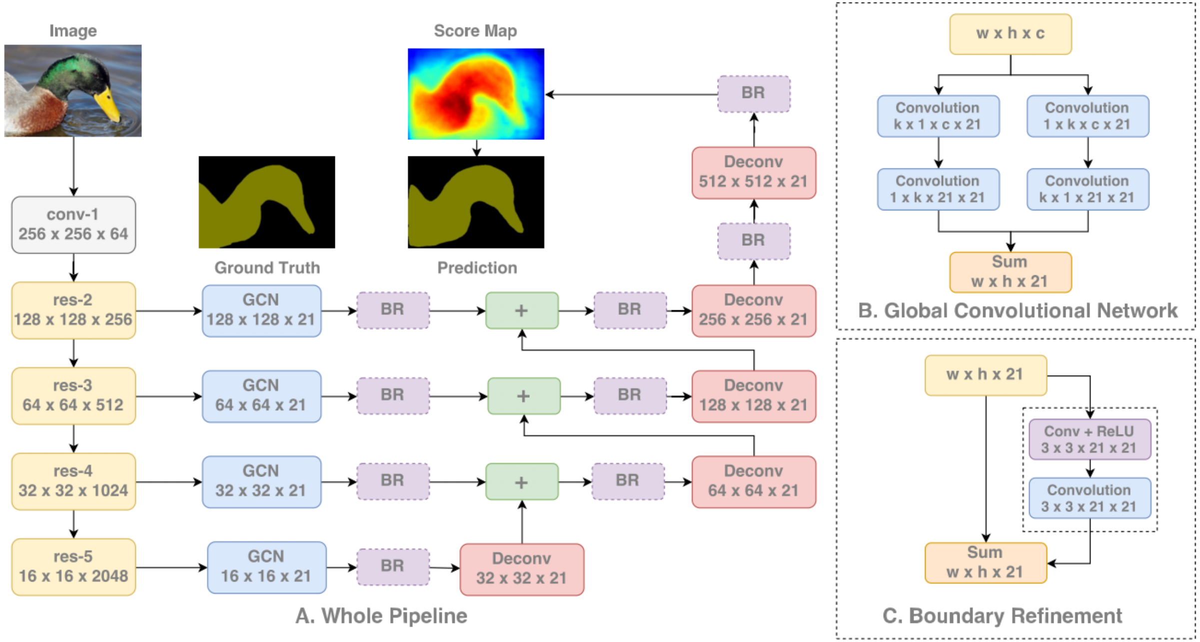
Deep lab v2

- DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs arXiv:1606.00915
- a’trous spatial pyramid pooling(ASPP)



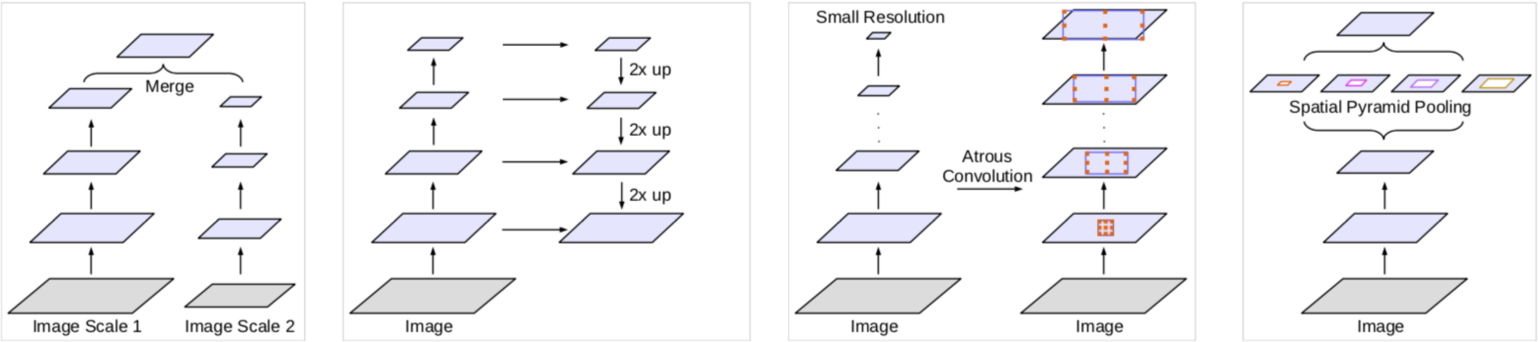
GCN

- Large Kernel Matters
  — Improve Semantic Segmentation by Global Convolutional Network
  arXiv:1703.02719
- 大内核（Large kernel）



Deep lab V3

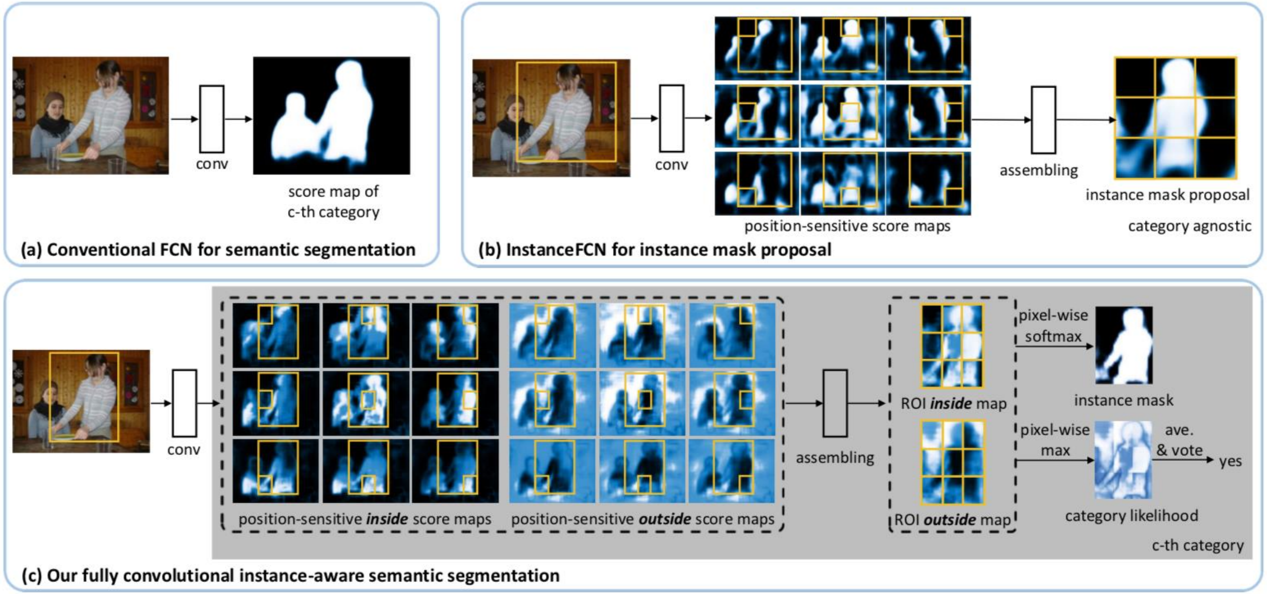
- Rethinking Atrous Convolution for Semantic Image Segmentation arXiv:1706.05587
- 多尺度信息结合方式的探究
- 工程上的胜利



(a) Image Pyramid (b) Encoder-Decoder (c) Deeper w. Atrous Convolution (d) Spatial Pyramid Pooling

FCIS

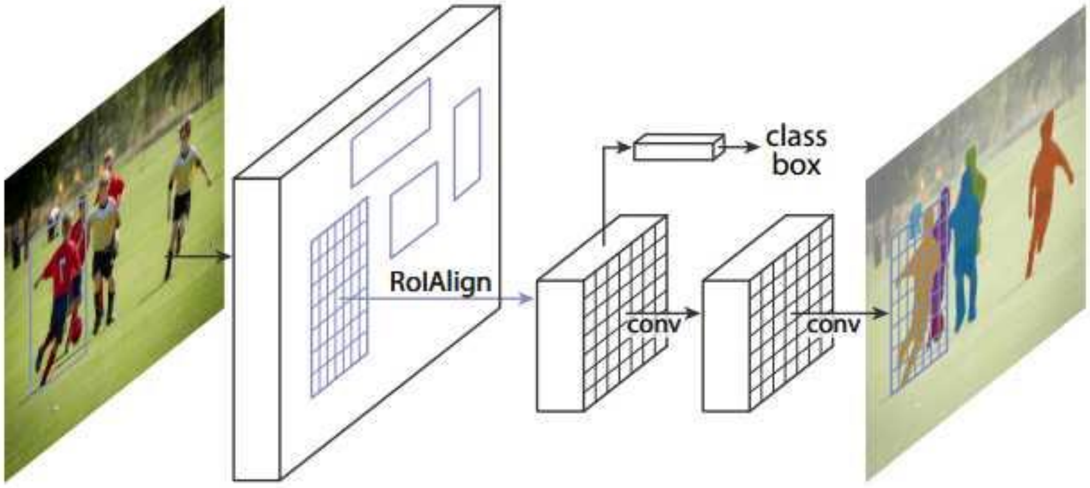
- Fully Convolutional Instance-aware Semantic Segmentation arXiv:1611.07709
- Inside/outside score



FCIS

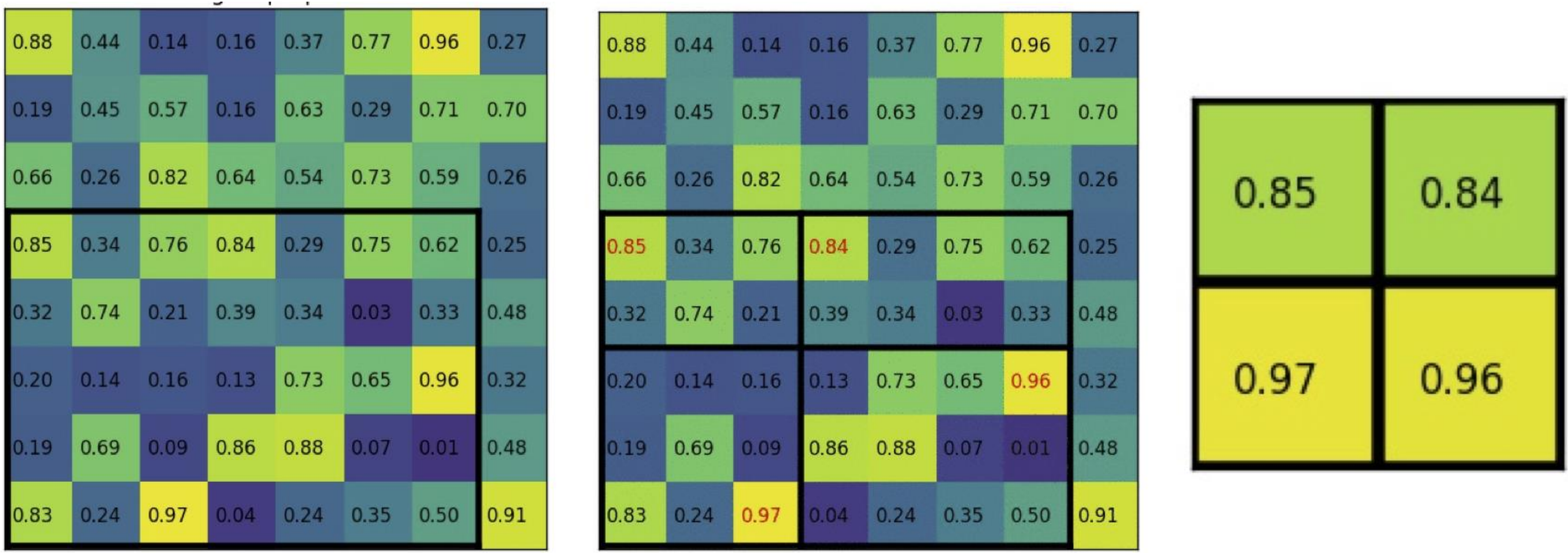
Mask R-CNN

- Mask R-CNN
  arXiv: 1703.06870
- 结构上：Faster R-CNN+FCN
- 为了得到更精确的mask位置，加入RoI Align



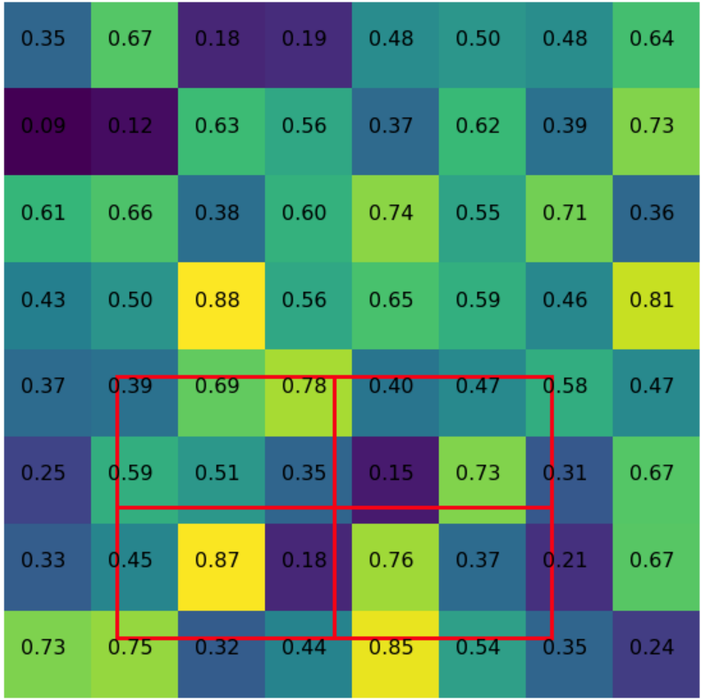
ROI Align

- 回顾一下ROI pooling



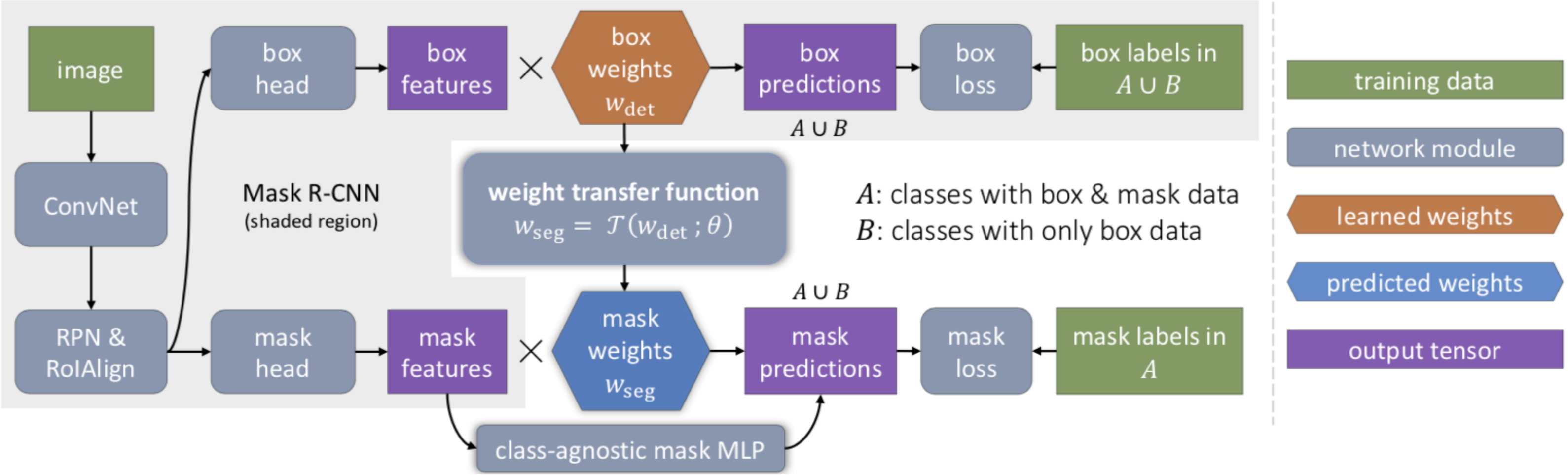
ROI Align

- ROI Align：采用双线性插值的方式，允许非整数像素的切割点



Mask ^ X R-CNN

- Learning to segment every thing
  arXiv: 1711.10370
- 采用迁移学习思路，学习一个映射
- 可能大幅度减小训练数据的成本



不 止 于 代 码

THANK YOU

AI100