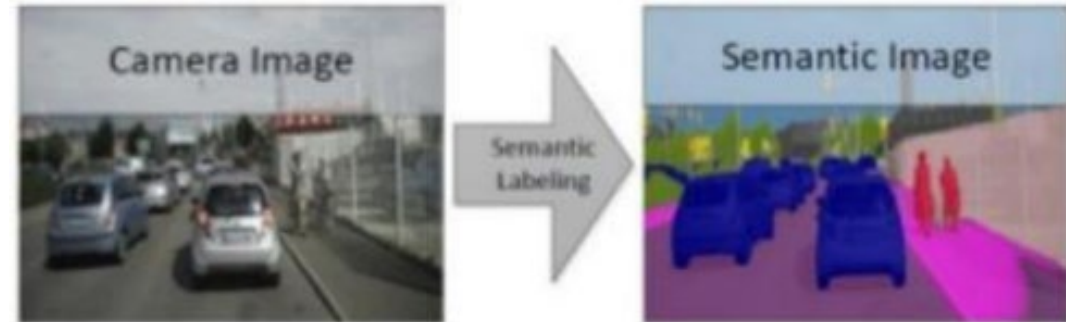


语义分割

语义分割 (Semantic Segmentation)



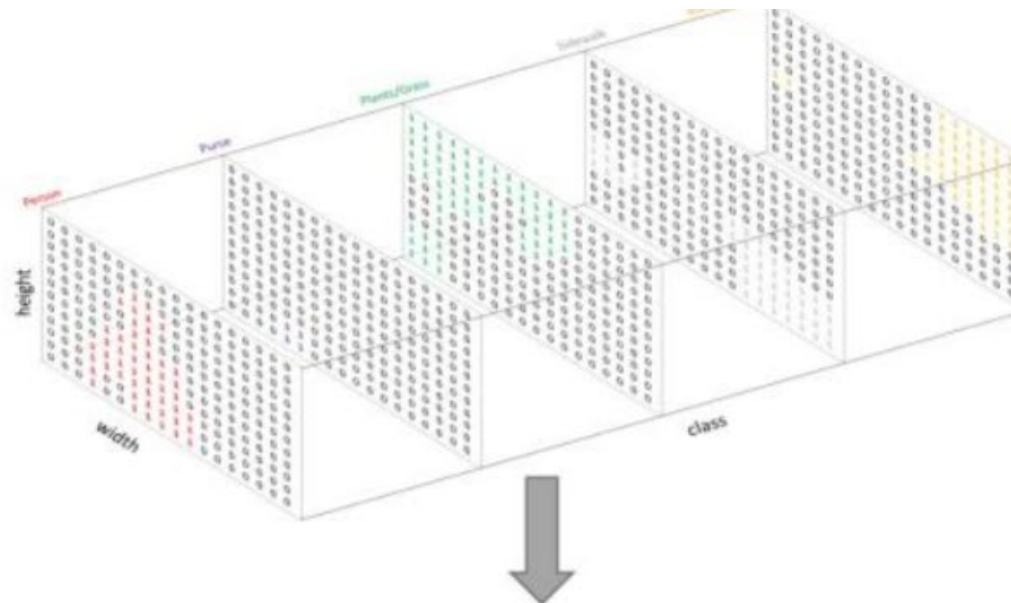
语义分割 (Semantic Segmentation)



Input



- 1: Person
- 2: Purse
- 3: Plants/Grass
- 4: Sidewalk
- 5: Building/Structures



- 0: Background/Unknown
- 1: Person
- 2: Purse
- 3: Plants/Grass
- 4: Sidewalk
- 5: Building/Structures



3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	1	1	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	1	1	1	1	3	3	3	5	5	5	5	5	5
3	3	3	3	3	1	1	1	3	3	3	5	5	5	5	5	5	5
5	5	3	3	3	3	1	1	3	3	5	5	5	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	4	4	5	5	5
4	4	4	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	1	4	4	4	4	4	4	4

Semantic Labels

语义分割 (Semantic Segmentation)

- 早期：基于像素块的语义分割



Input:
 $3 \times H \times W$

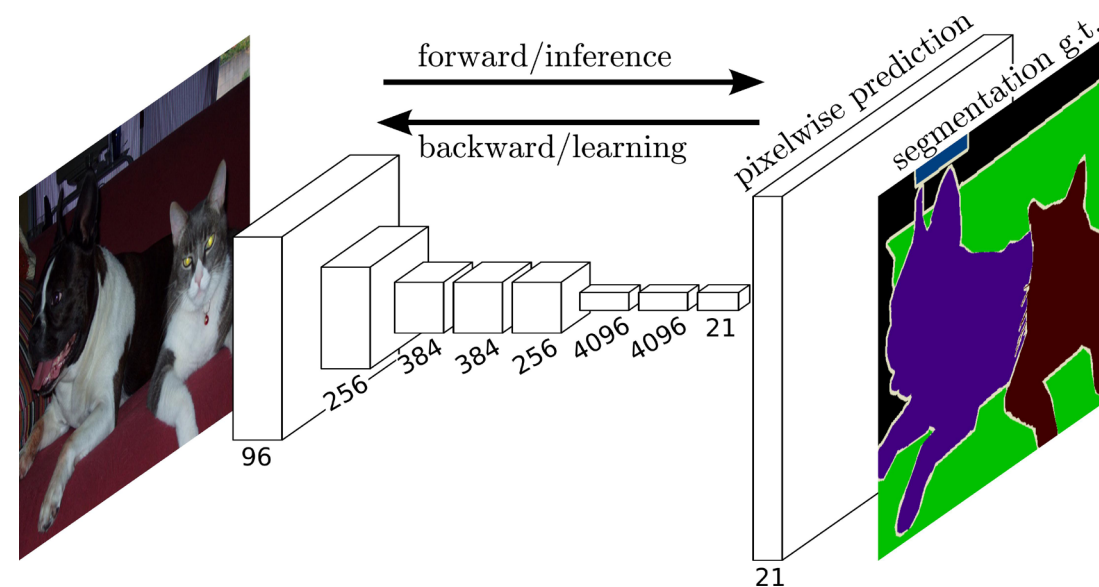


Predictions:
 $H \times W$

语义分割 (Semantic Segmentation)

● 全卷积网络 (FCN)

- Fully Convolutional Networks for Semantic Segmentation (2015 CVPR);
- 将用于图像分类的卷积神经网络中的全连接层替换成卷积层;
- 主要特点:
 - 卷积化 (Convolutional): 将分类网络的特征提取层迁移到语义分割结构中;
 - 上采样 (Upsample): 采用上采样方法将低分辨率的语义特征转化为像素级分类结果;
 - 跳跃结构 (Skip Layer): 将不同深度的特征进行融合。

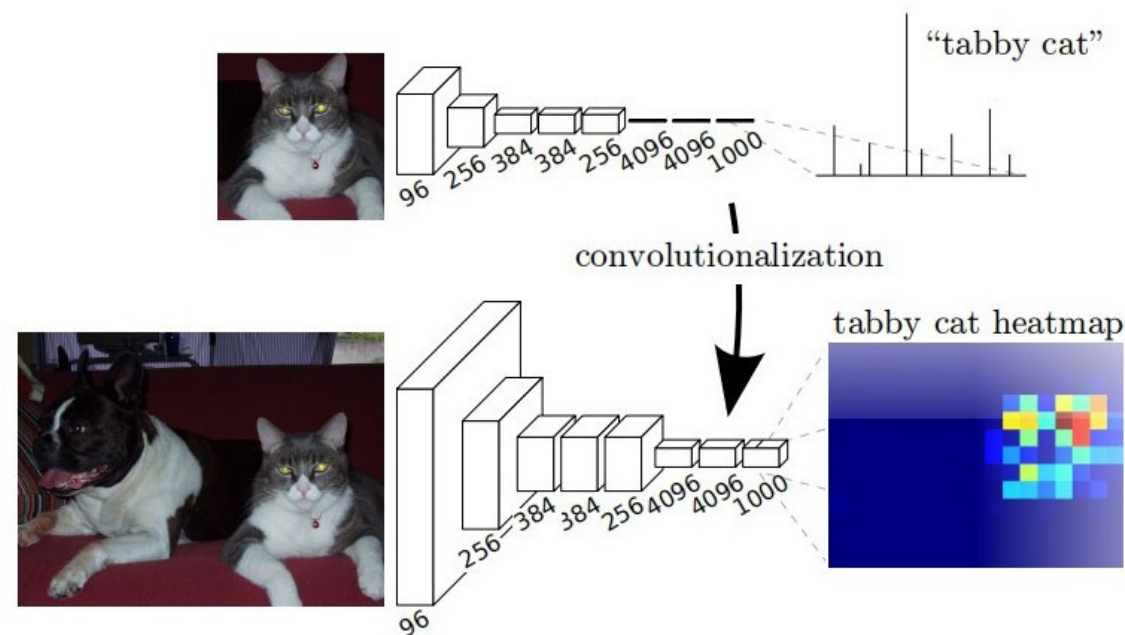


语义分割 (Semantic Segmentation)

● 全卷积网络 (FCN)

■ 卷积化:

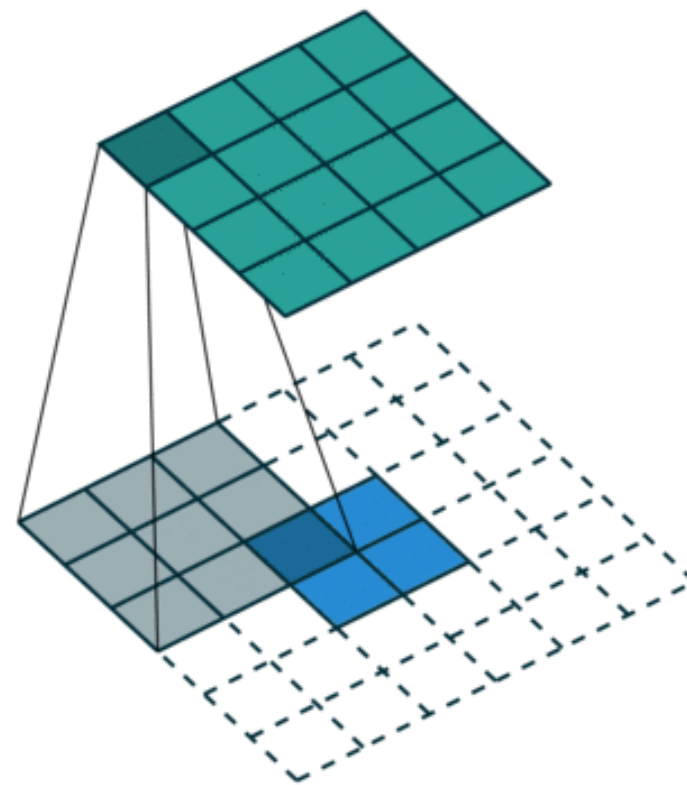
- 在一个传统的CNN结构中 (如右图), 前5层是卷积层, 第6层和第7层分别是一个长度为4096的一维向量, 第8层是长度为1000的一维向量, 分别对应1000个类别的概率;
- FCN将这3层表示为卷积层, 卷积核的大小分别为 $(4096, 1, 1)$ 、 $(4096, 1, 1)$ 、 $(1000, 1, 1)$;
- 此时, 卷积神经网络所有的层都是卷积层, 故称 “全卷积网络” 。



● 全卷积网络 (FCN)

■ 上采样:

- 一般来说, 经过CNN的多次卷积、池化操作以后, 得到的特征图 (Feature map) 越来越小, 分辨率越来越低;
- 为了恢复原图的分辨率, 使用上采样方法;
- 上采样的常见策略包括插值法和转置卷积 (Transpose convolution) (也称微步卷积、反卷积) 。

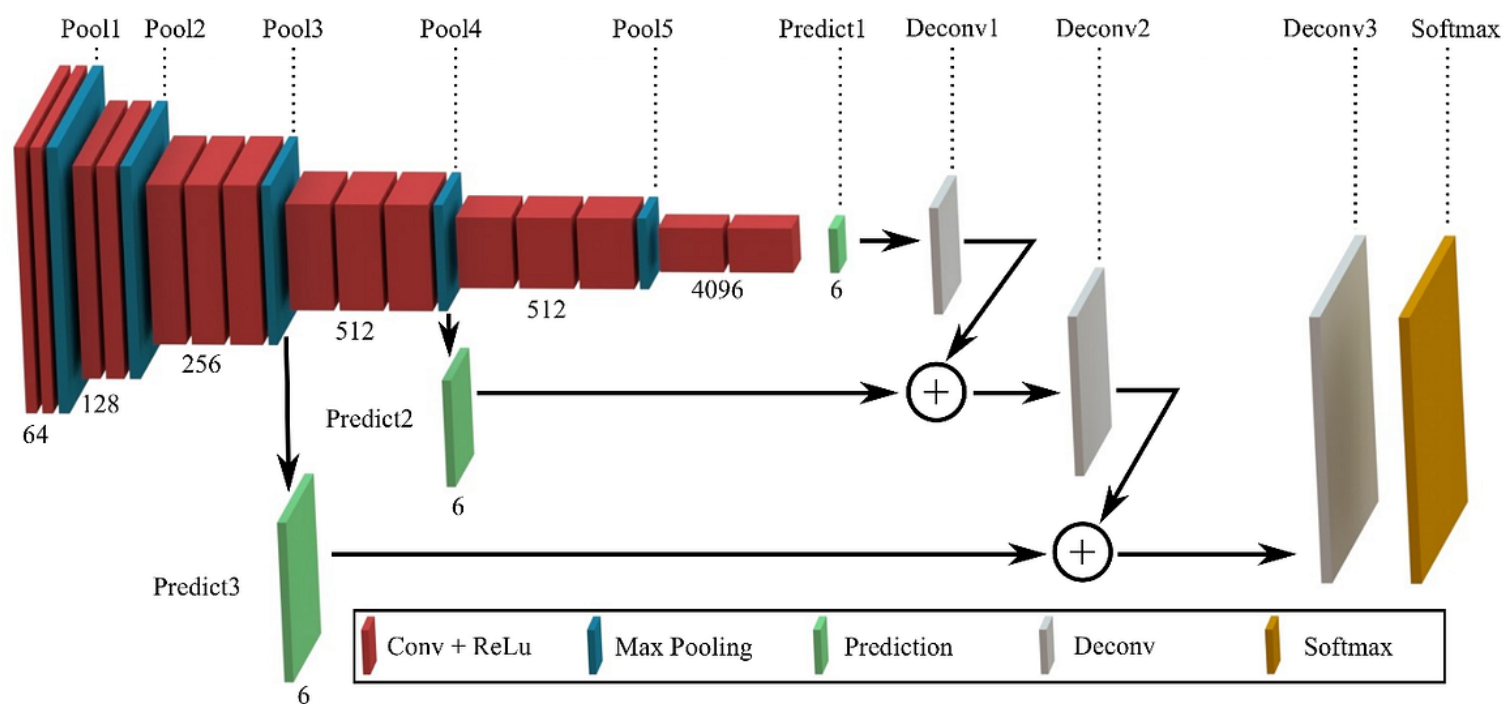


语义分割 (Semantic Segmentation)

● 全卷积网络 (FCN)

■ 跳跃结构:

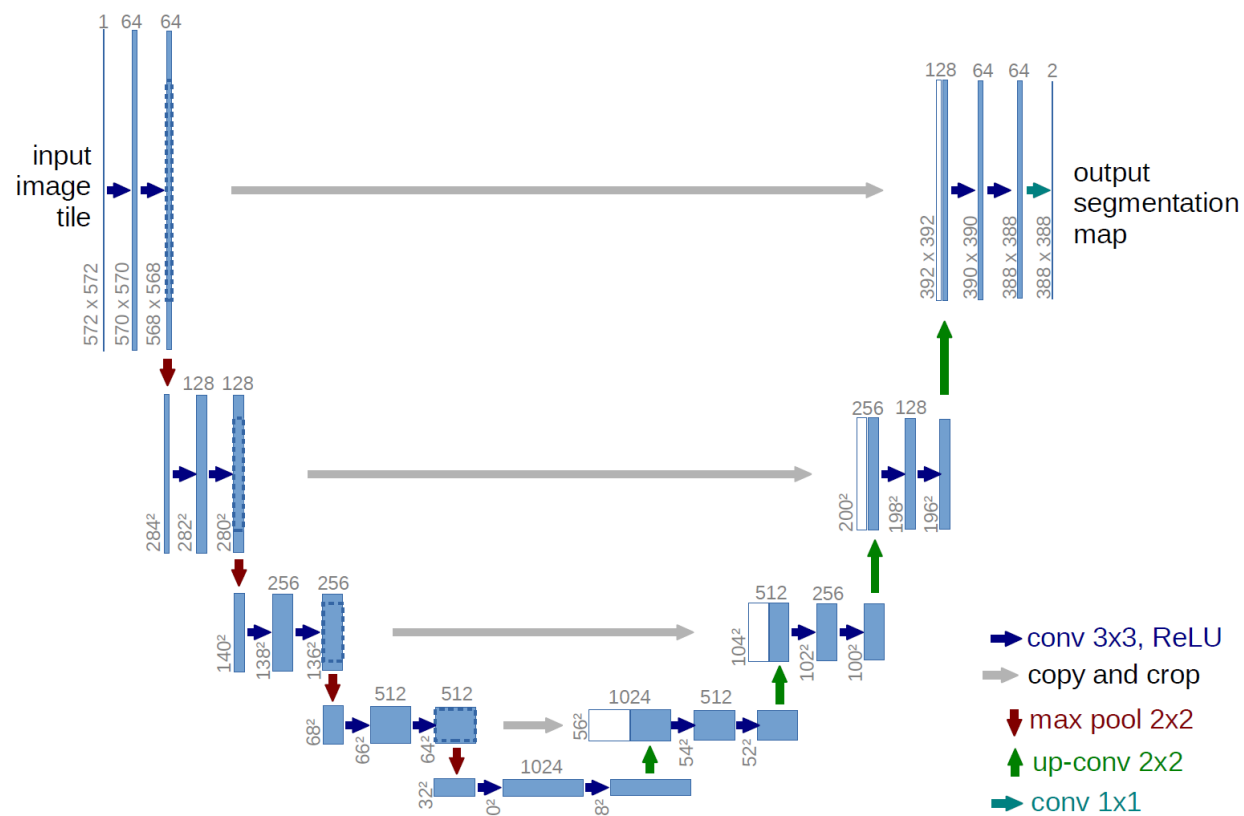
- 因为如果将全卷积之后的结果直接进行上采样，得到的结果很粗糙；
- FCN通过融合编码阶段不同分辨率的特征图来增强空间特征信息，从而加强分割效果。



语义分割 (Semantic Segmentation)

• U-Net

- U-Net: Convolutional Networks for Biomedical Image Segmentation (2015 MICCAI);
- 将Encoder的特征图与对应分辨率的Decoder特征图进行拼接操作，从而形成“阶梯结构”；
- 允许不同阶段的Decoder层能够弥补Encoder因为池化等操作丢失的信息；
- 支持很少的数据训练模型，速度快，在只包含了30幅医学图像上训练就能得到 State-of-the-art 的效果。



总结

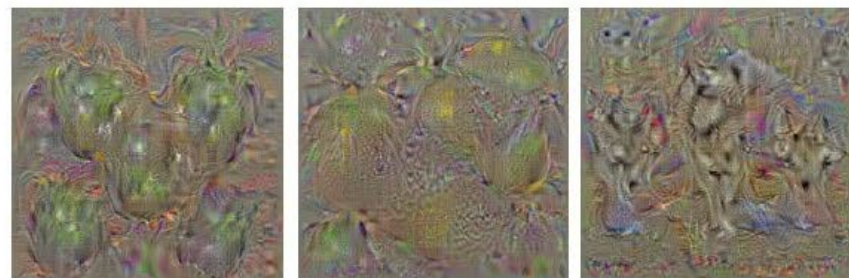
总结



dumbbell

cup

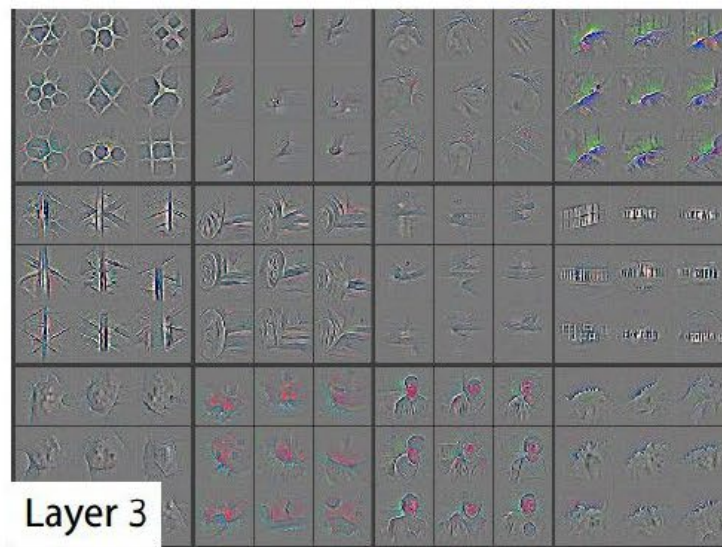
dalmatian



bell pepper

lemon

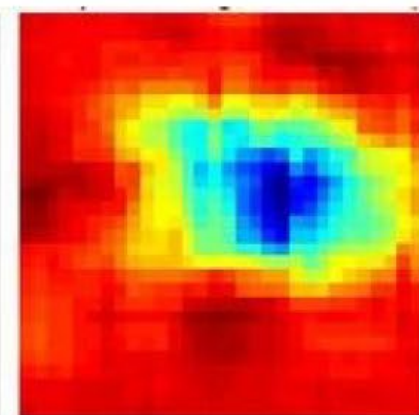
husky



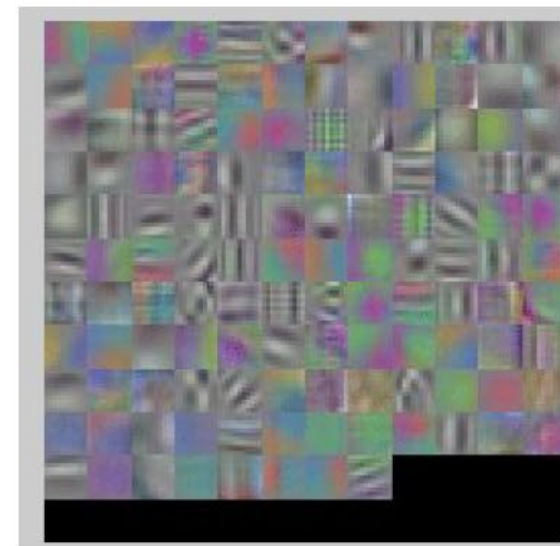
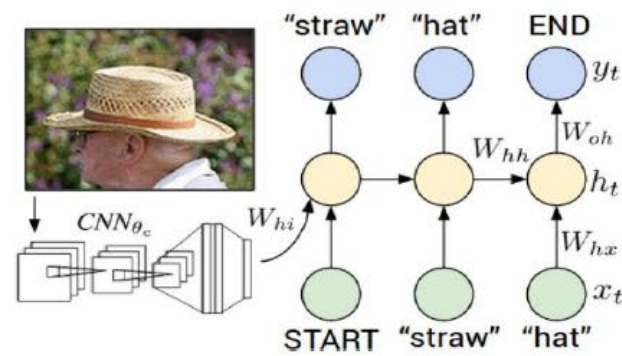
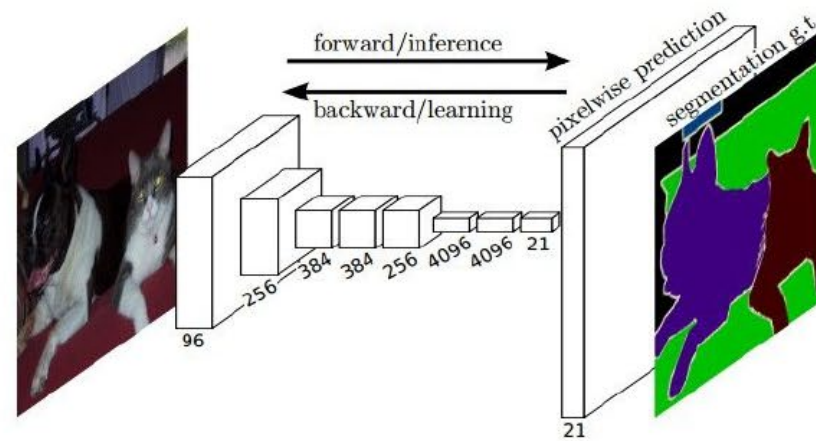
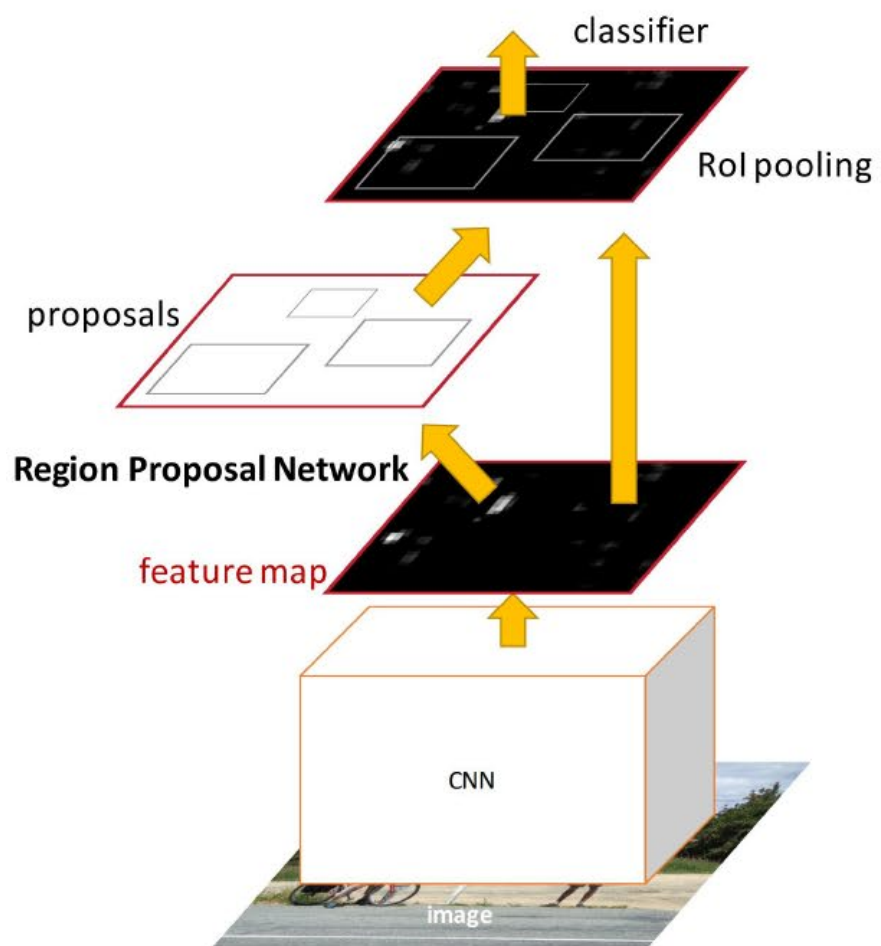
Layer 3



True Label: Pomeranian

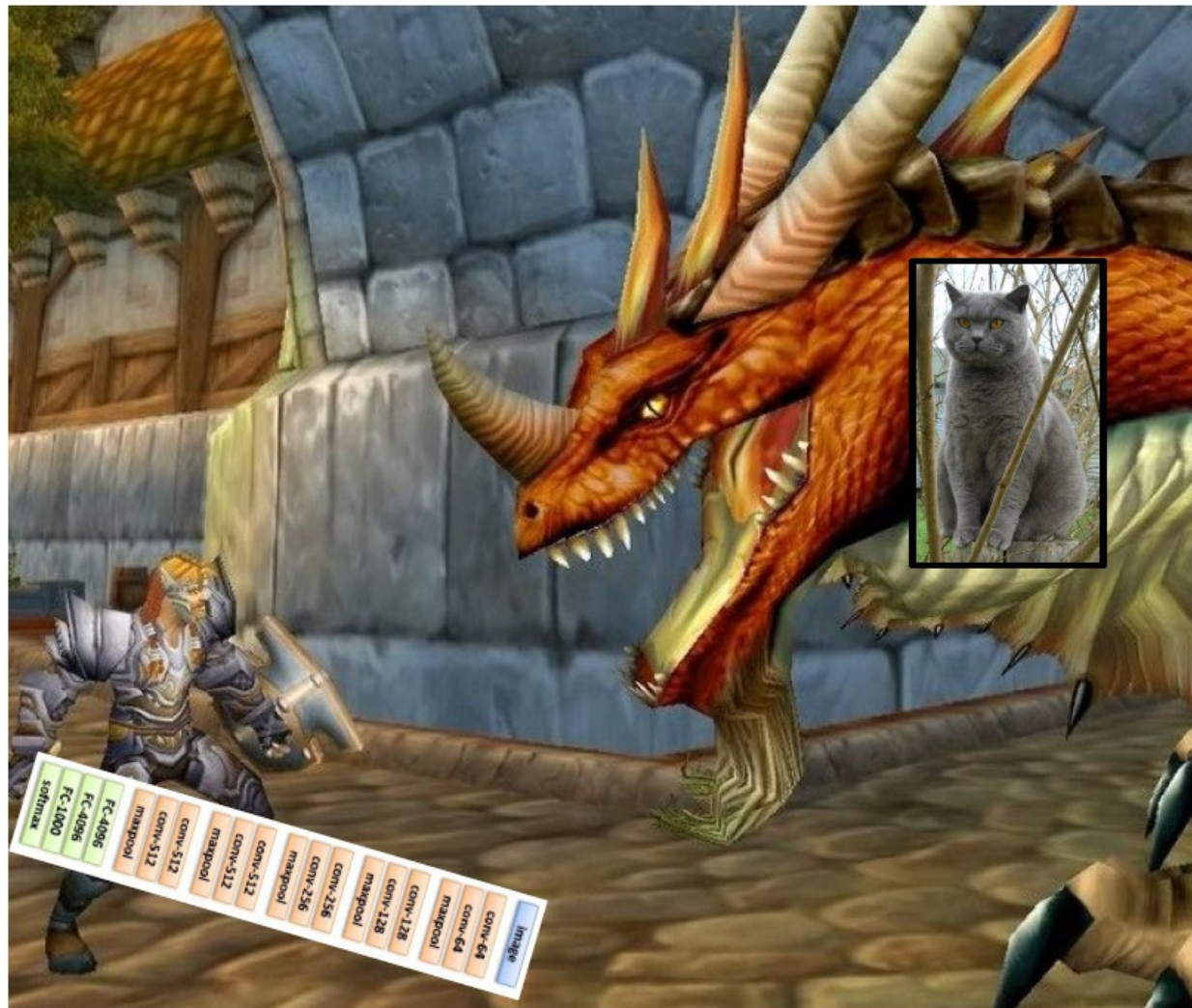


总结





总结



- 更多学习资源:

- Stanford CS231n: <https://cs231n.stanford.edu/>
- MIT 6.8300/1 Advances in Computer Vision: <https://advances-in-vision.github.io/>



谢谢！