

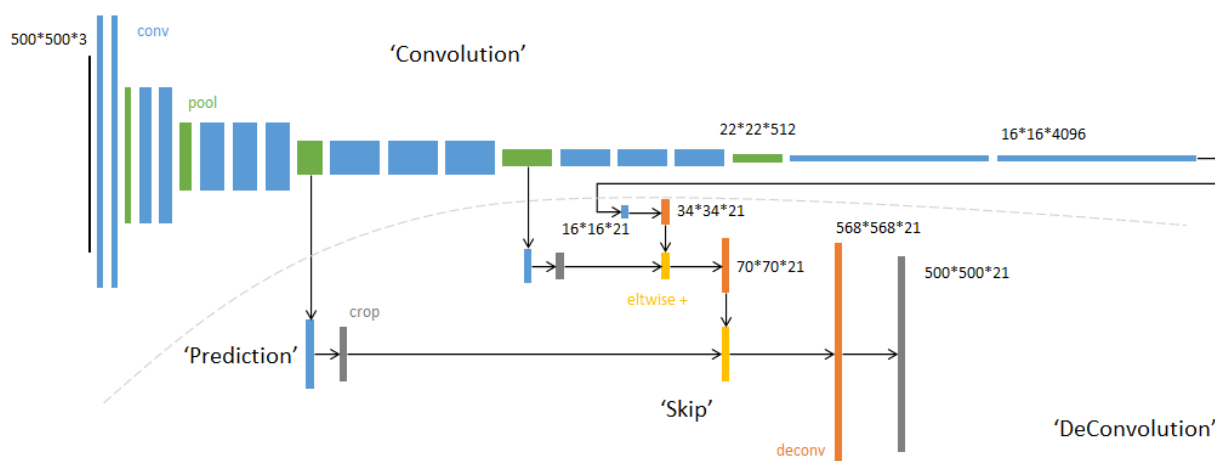
## 1.核心思想

1. 不含全连接层，最后几层都是卷积层，可以适应任意尺寸输入，并可逐像素分类
2. 最后添加反卷积层，将图像变成输入图像大小，能输出精细结果
3. 结合不同深度层结果的跳级skip结构，同时确保鲁棒性和精确性

## 2.一些重点

- 1、损失函数是在最后一层的 spatial map上的 pixel 的 loss 和，在每一个 pixel 使用 softmax loss
- 2、使用 skip 结构融合多层（3层）输出，底层网络应该可以预测更多的位置信息，因为他的感受野小可以看到小的 pixels
- 3、上采样 lower-resolution layers 时，如果采样后的图因为 padding 等原因和前面的图大小不同，使用 crop，当裁剪成大小相同的，spatially aligned，使用 concat 操作融合两个层

## 3.网络结构



论文中，达到最高精度的分类网络是VGG16，但提供的模型基于AlexNet。此处使用AlexNet便于绘图。

全连接层转换为卷积层：在两种变换中，将全连接层转化为卷积层在实际运用中更加有用。假设一个卷积神经网络的输入是  $224 \times 224 \times 3$  的图像，一系列的卷积层和下采样层将图像数据变为尺寸为  $7 \times 7 \times 512$  的激活数据体。AlexNet使用了两个尺寸为4096的全连接层，最后一个有1000个神经元的全连接层用于计算分类评分。我们可以将这3个全连接层中的任意一个转化为卷积层：

针对第一个连接区域是 $[7 \times 7 \times 512]$ 的全连接层，令其滤波器尺寸为 $F=7$ ，这样输出数据体就为 $[1 \times 1 \times 4096]$ 了。

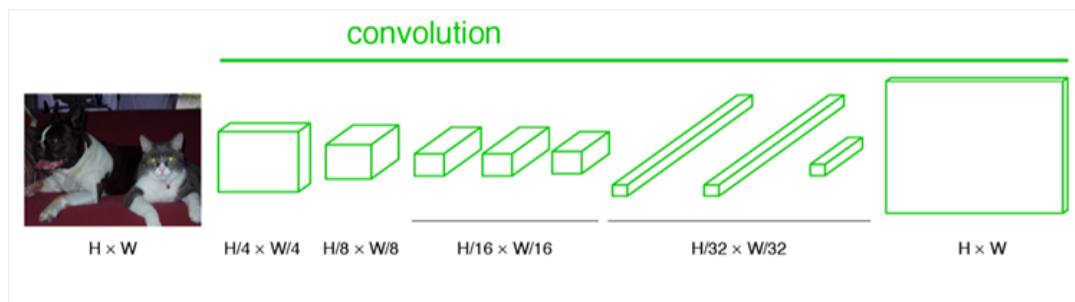
针对第二个全连接层，令其滤波器尺寸为 $F=1$ ，这样输出数据体为 $[1 \times 1 \times 4096]$ 。

对最后一个全连接层也做类似的，令其 $F=1$ ，最终输出为 $[1 \times 1 \times 1000]$

## 4.具体怎么逐像素预测

采用反卷积层对最后一个卷积层的feature map进行上采样，使他恢复到输入图像相同的尺寸，从而可以对每个像素都产生了一个预测，同时保留了原始输入图像中的空间信息，最后再上采样的特征图上进行逐像素分类。

具体过程：



最后的输出是21张heatmap经过upsampling变为原图大小的图片，为了对每个像素进行分类预测label成最后已经进行语义分割的图像，**这里有一个小trick，就是最后通过逐个像素地求其在21张图像该像素位置的最大数值描述（概率）作为该像素的分类。**因此产生了一张已经分类好的图片，如下图右侧有狗狗和猫猫的图。

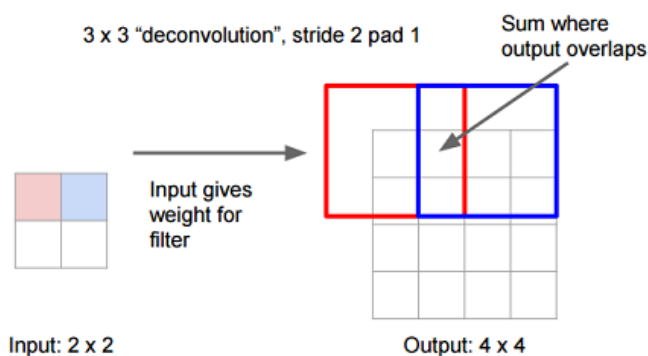
## 5.反卷积-上采样

这里会先进行上采样，即扩大像素；再进行卷积——通过学习获得权值）

下半部分，反卷积层（橙色 $\times 3$ ）可以把输入数据尺寸放大。和卷积层一样，上采样的具体参数经过训练确定。

这里图像的反卷积与下图的full卷积原理是一样的，使用了这一种反卷积手段使得图像可以变大，FCN作者使用的方法是这里所说反卷积的一种变体，这样就可以获得相应的像素值，图像可以实现end to end。

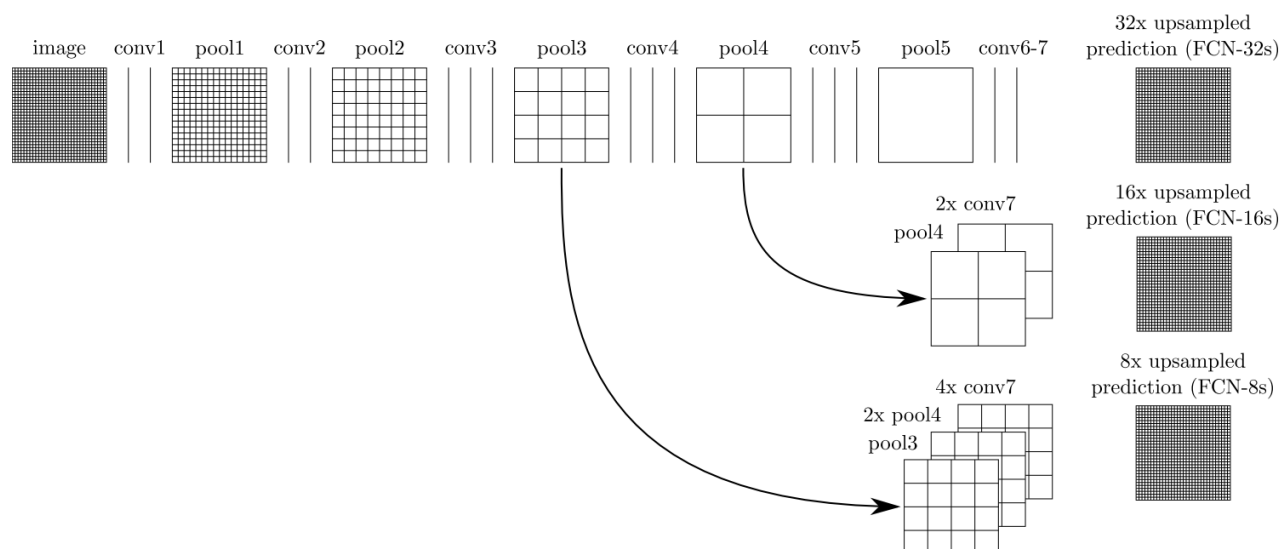
（feature map值与权重不同，生成的上采样的二值区域也是不一样的。）



## 6.skip结构

如下图所示：对原图进行卷积conv1、pool1后图像缩小为1/2；对图像进行第二次卷积conv2、pool2后图像缩小为1/4；对图像进行第三次卷积conv3、pool3后图像缩小为1/8，此时保留pool3的featuremap；对图像进行第四次卷积conv4、pool4后图像缩小为1/16，此时保留pool4的featuremap；对图像进行第五次卷积conv5、pool5后图像缩小为1/32，然后把原来CNN操作过程中的全连接变成卷积操作的conv6、conv7，图像的featuremap的大小依然为原图的1/32, 此时图像不再叫featuremap而是叫heatmap。

其实直接使用前两种结构就已经可以得到结果了，这个上采样是通过反卷积（deconvolution）实现的，对第五层的输出（32倍放大）反卷积到原图大小。但是得到的结果还上不够精确，一些细节无法恢复。于是将第四层的输出和第三层的输出也依次反卷积，分别需要16倍和8倍上采样，结果过也更精细一些了。这种做法的好处是兼顾了local和global信息。



## 结论

总体来说，本文的逻辑如下：

- 想要精确预测每个像素的分割结果
- 必须经历从大到小，再从小到大的两个过程
- 在升采样过程中，分阶段增大比一步到位效果更好
- 在升采样的每个阶段，使用降采样对应层的特征进行辅助