**Imagenet classification with deep convolutional neural networks**

1. ImageNet：15000000个 有标记的 高清图片集 22000个类别

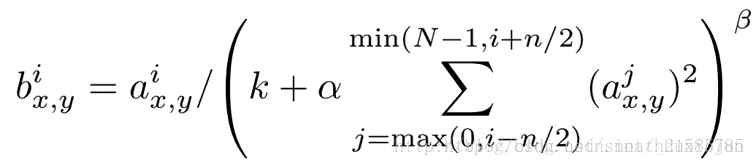
LabelMe：成百上千个 完全分割 的图像

1. CNNs的优势：（1）关联更少（2）使用卷积参数少（3）易训练
2. 使学习速度快：（1）非饱和神经元non-saturating

（2）GPU，两个GTX580 3GB GPUs ，5-6天

1. 实验证明Relus比tanh拟合的速度快
2. 选择连接模式是交叉验证一个问题，但允许我们精确调整通信量，直到它成为计算量一个可接受的部分
3. LRN（local response normalization）

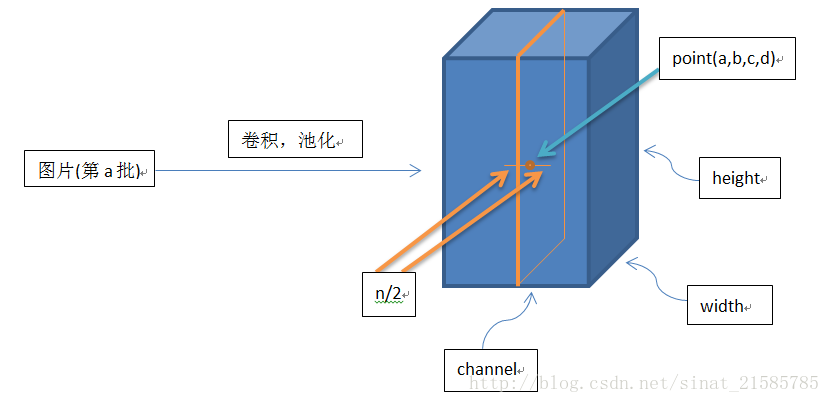
对局部神经元的活动创建竞争机制，使得其中响应比较大的值变得相对更大，并且抑制反馈比较小的神经元，增强模型的泛化能力。



K = 2 ， n =5， a = e\*(-4) , b = 0.75

ai(x,y)表示在这个输出结构中的一个位置[a,b,c,d]，可以理解成在某一张图中的某一个通道下的某个高度和某个宽度位置的点，即第a张图的第d个通道下的高度为b宽度为c的点。

a,n/2,k,α,β分别表示函数中的input,depth\_radius,bias,alpha,beta，其中n/2,k,α,β都是自定义的，特别注意一下∑叠加的方向是沿着通道方向的，即每个点值的平方和是沿着a中的第3维channel方向的，也就是一个点同方向的前面n/2个通道（最小为第0个通道）和后n/2个通道（最大为第d-1个通道）的点的平方和(共n+1个点)。



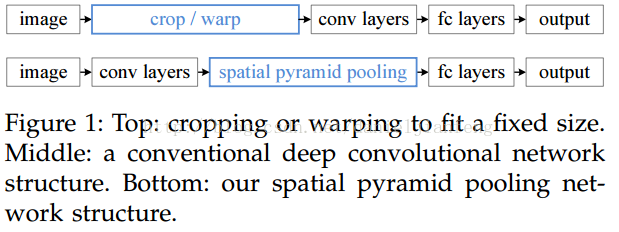
1. 池化
2. 一般池化

SizeX = stride (max pooling/mean pooling)

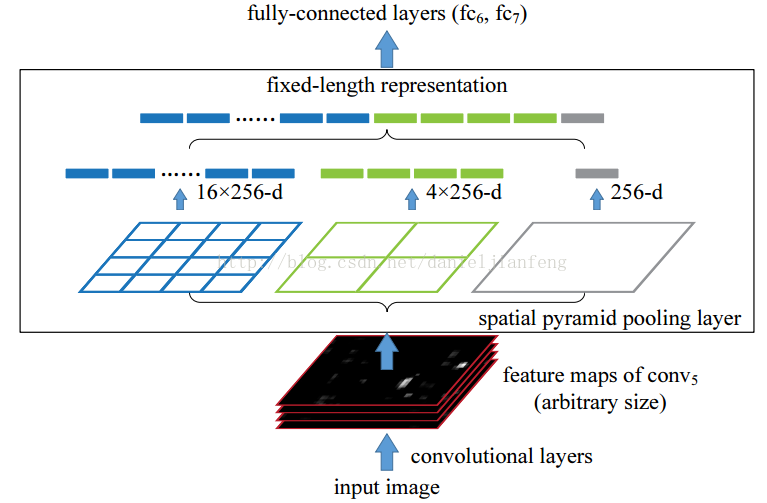
1. 重叠池（Overlapping pooling）

SizeX > stride, 不丢失信息

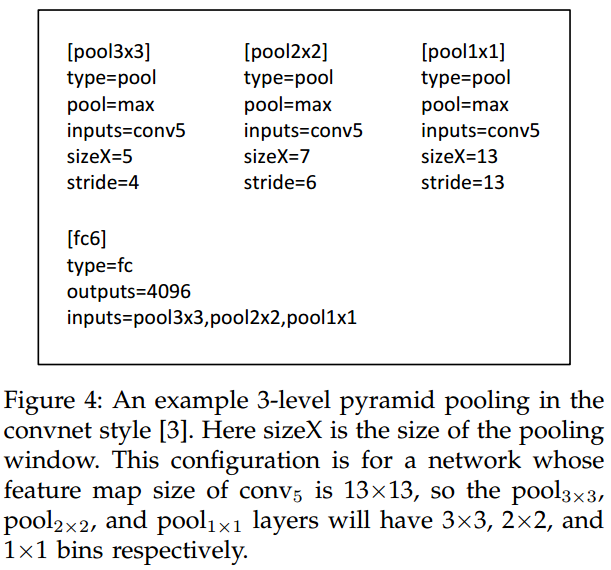
1. 空金字塔池（Spatial Pyramid Pooling）



空间金字塔池化的思想来自于Spatial Pyramid Model，它一个pooling变成了多个scale的pooling。用不同大小池化窗口作用于卷积特征，我们可以得到1X1,2X2,4X4的池化结果，由于conv5中共有256个过滤器，所以得到1个256维的特征，4个256个特征，以及16个256维的特征，然后把这21个256维特征链接起来输入全连接层，**通过这种方式把不同大小的图像转化成相同维度的特征。**



对于不同的图像要得到相同大小的pooling结果，就需要根据图像的大小动态的计算池化窗口的大小和步长。假设conv5输出的大小为a\*a，需要得到n\*n大小的池化结果，可以让窗口大小sizeX为IMG_256，步长为IMG_257 。下图以conv5输出的大小为13\*13为例。



1. 避免过拟合的方法：

(1)数据增强（data augmentation）

(2)Dropout

1. input:224\*224\*3 96kernels 11\*11\*3 stride=4

5个卷积层，3个全连接层