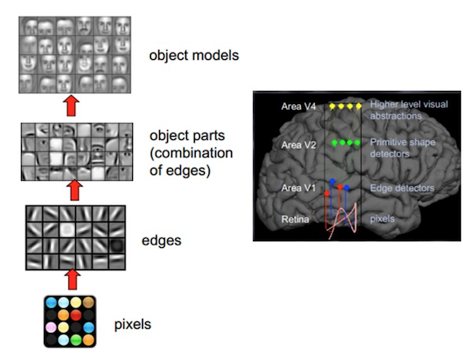
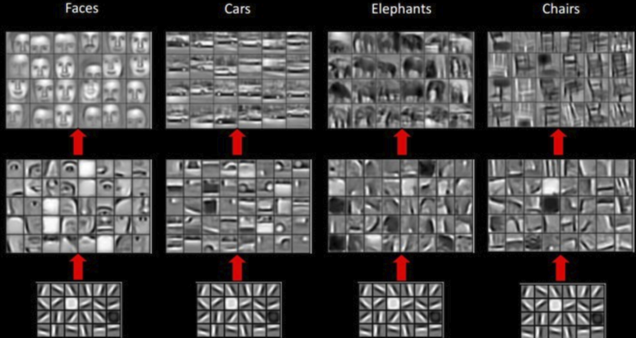
卷积神经网络

# 一、视觉神经网络

从原始信号摄入开始(瞳孔摄入像素 pixels)，接着做初步处理(大脑皮层某些细胞发现边缘和方向)，然后抽象出原始形状，然后再做更高层次的视觉抽象。下面是人脑进行人脸识别的一个示例

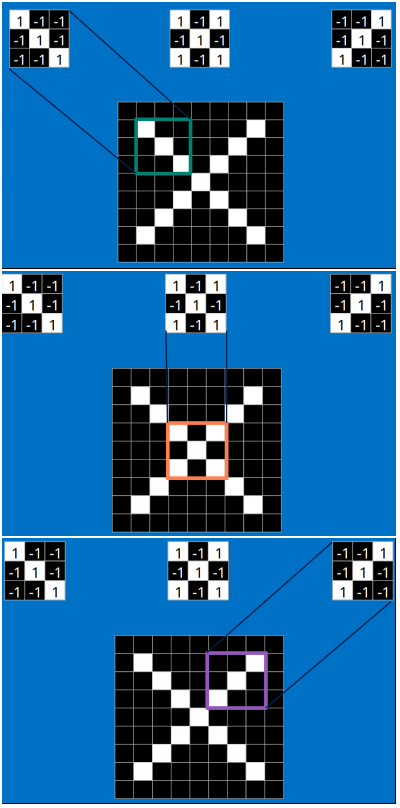
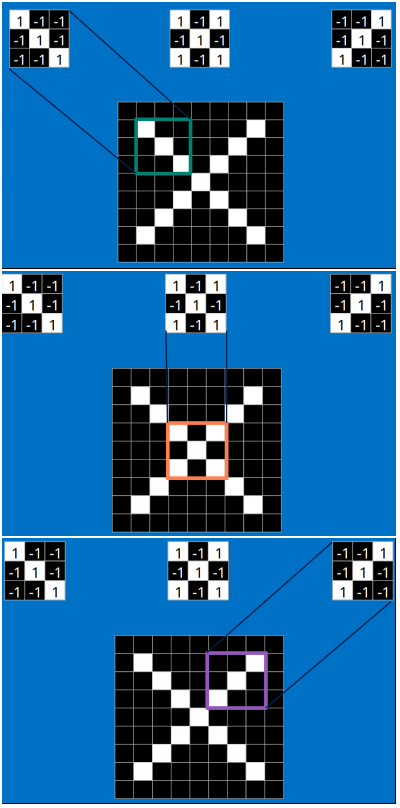
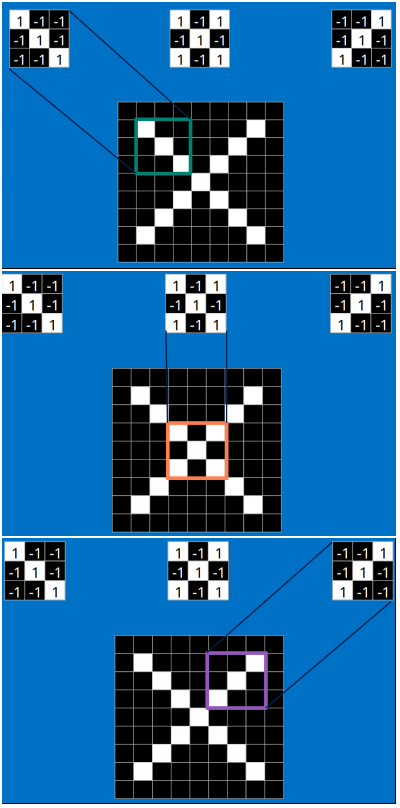
人类对外界的认知一般是从局部到全局，先对局部有感知的认识，再逐步对全体有认知。在图像中的空间联系也是类似，局部范围内的像素之间联系较为紧密，而距离较远的像素则相关性较弱。对于不同的物体，人类视觉也是通过这样逐层分级来进行认知的：



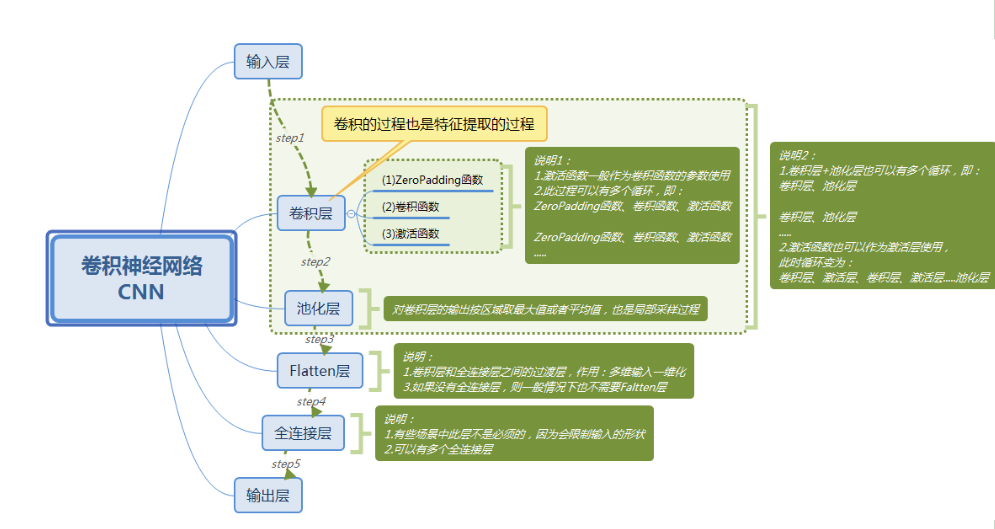
最底层的特征都是类似的，就是各种边缘等，越往上，越能提取出对应物体的一些特征(眼睛、轮子、躯干等)，到最上层，不同的高级特征最终组合成相应的图像，从而能够让人类准确地区分不同的物体。

# 二、卷积神经网络

计算机不能准确地知道这些特征到底要匹配原图的哪些部分，所以在原图中把每一个可能的位置都进行尝试，相当于把这个feature（特征）变成了一个过滤器

对于字母X，可以使用三个过滤器分别提取出以下三个重要特征两个交叉线、一个对角线

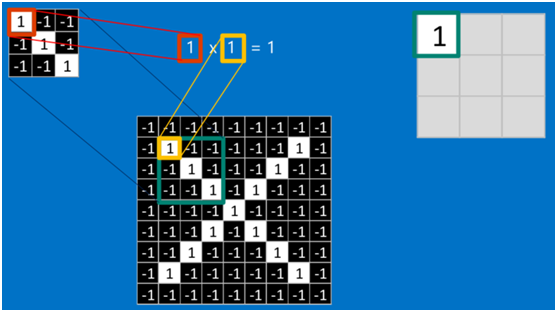
## 2.1 卷积神经网络的基本结构

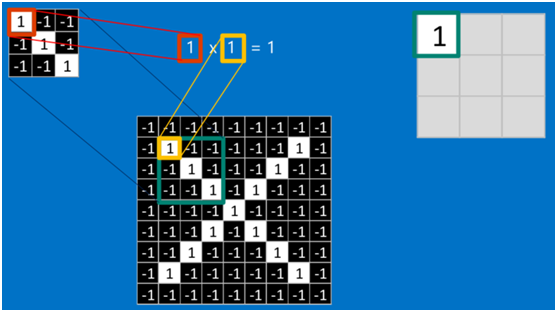


## 2.2 卷积（Convolution）

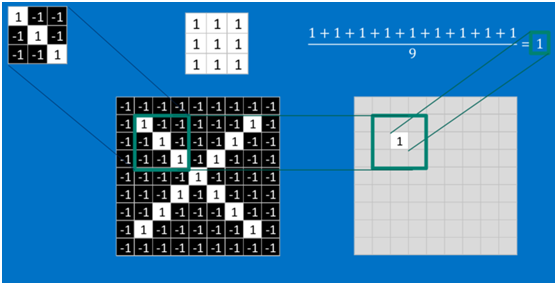
卷积层的主要作用是进行特征提取

计算一个feature（特征）和其在原图上对应的某一小块的结果

1. 将两个小块内对应位置的像素值进行乘法运算

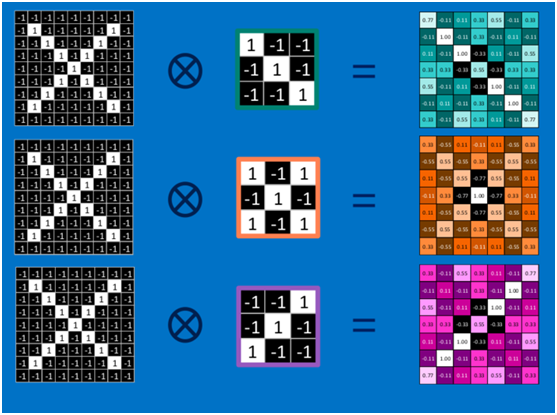


1. 将整个小块内乘法运算的结果累加起来
2. 除以小块内像素点总个数即可（注：也可不除以总个数的）



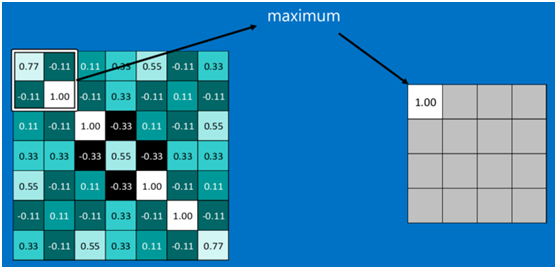
三个feature（特征）分别进行卷积操作，得到新的二维数组，称之为feature map

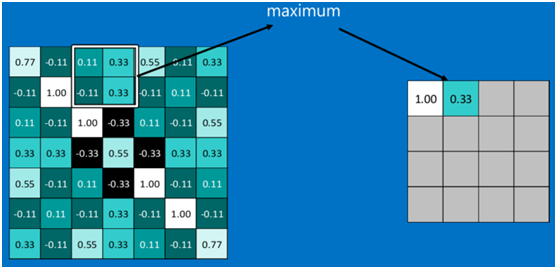
* 值越接近1表示对应位置和feature的匹配越完整
* 值越是接近-1，表示对应位置和feature的反面匹配越完整
* 值接近0的表示对应位置没有任何匹配或者说没有什么关联



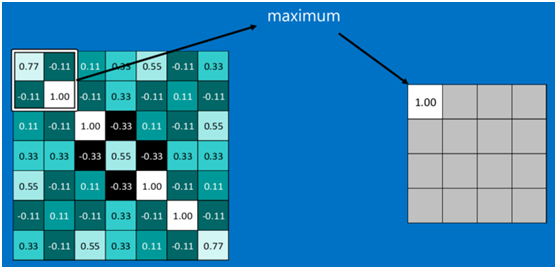
## 2.3 池化（pooling）

目前，有两种广泛使用的池化操作——平均池化（average pooling）和最大池化（max pooling），其中最大池化是两者中使用最多的一个操作，其效果一般要优于平均池化。池化层用于在卷积神经网络上减小特征空间维度，但不会减小深度。当使用最大池化层时，采用输入区域的最大数量，而当使用平均池化时，采用输入区域的平均值。

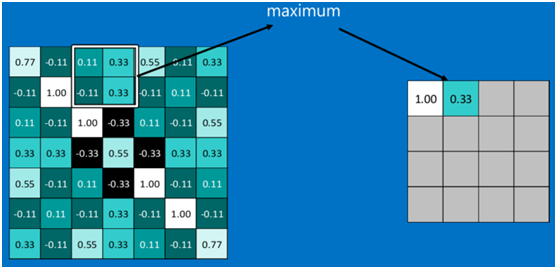
池化(Pooling)就是将输入图像进行缩小，减少像素信息，只保留重要信息

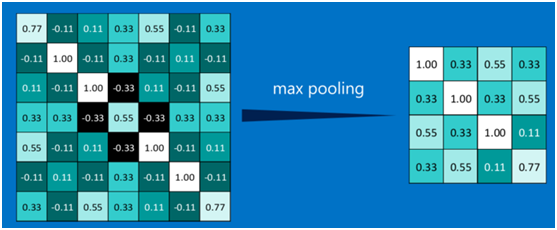


通常情况下，池化区域是2\*2大小，然后按一定规则转换成相应的值，如取池化区域内的最大值（max-pooling）、平均值（mean-pooling）等，以这个值作为结果的像素值

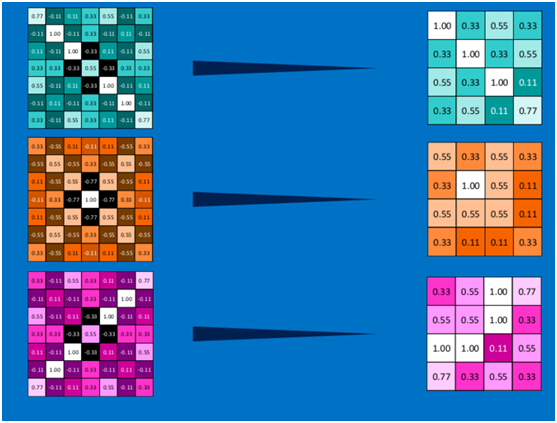


池化区域往左或者往下，依次取最大值池化区域往左，第二小块取大值max(0.11,0.33,-0.11,0.33)，作为池化后的结果





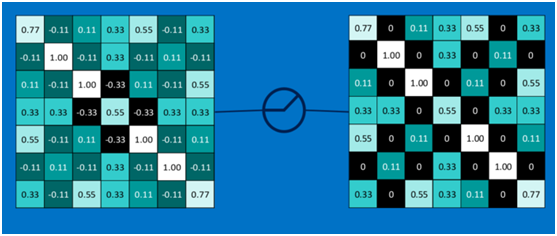
对所有的feature map执行同样的操作，结果如下：



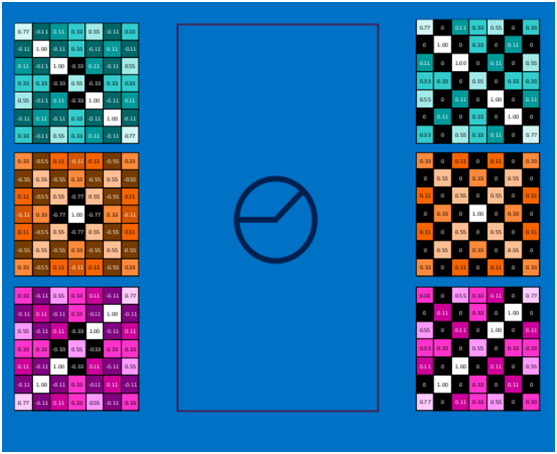
## 2.4 激活函数

常用的激活函数有sigmoid、tanh、relu等等，前两者sigmoid/tanh比较常见于全连接层，后者ReLU常见于卷积层。

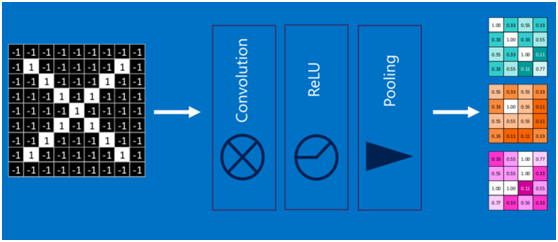
激活函数是用来加入非线性因素的，使得神经网络可以任意逼近任何非线性函数，提高神经网络对模型的表达能力，解决线性模型所不能解决的问题，这样神经网络就可以应用到众多的非线性模型中。



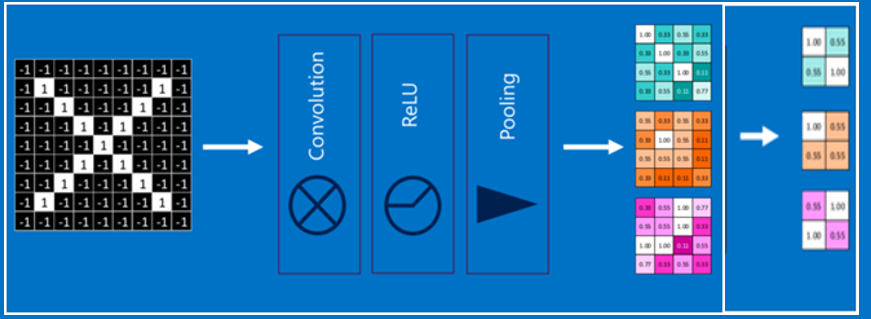
在卷积神经网络中，激活函数一般使用ReLU(The Rectified Linear Unit，修正线性单元)，它的特点是收敛快，求梯度简单。计算公式也很简单，max(0,T)，即对于输入的负值，输出全为0，对于正值，则原样输出。

对所有的feature map执行ReLU激活函数操作，结果如下：

# 深度神经网络

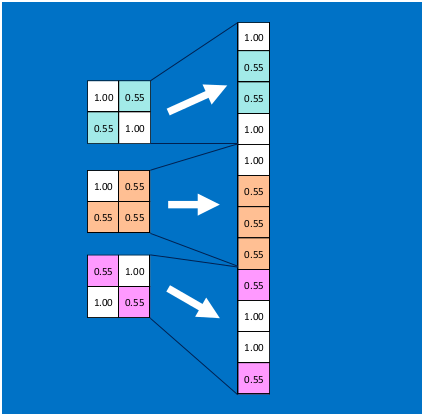
通过将上面所提到的卷积、激活函数、池化组合在一起

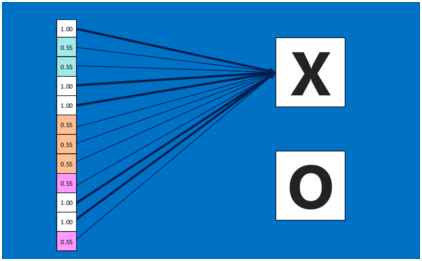
通过加大网络的深度，增加更多的层，就得到了深度神经网络，如下图

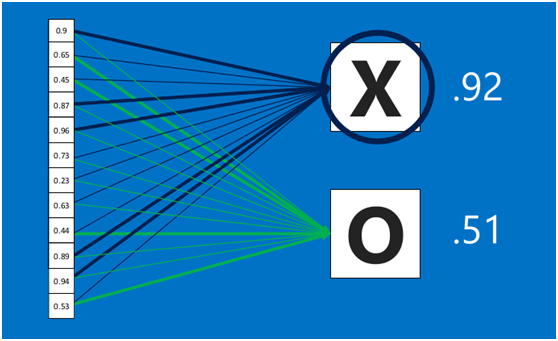


## 3.1 全连接层

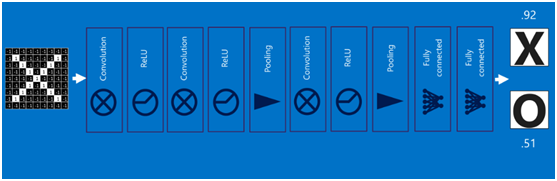
全连接层在整个卷积神经网络中起到“分类器”的作用，即通过卷积、激活函数、池化等深度网络后，再经过全连接层对结果进行识别分类。：



由于神经网络是属于监督学习，在模型训练时，根据训练样本对模型进行训练，从而得到全连接层的权重（如预测字母X的所有连接的权重)

最后计算出来字母X的识别值为0.92，字母O的识别值为0.51，则结果判定为X

“卷积神经网络”（CNN）结构，如下图所示：



## 3.2 总结

**卷积神经网络主要由两部分组成** 1、特征提取（卷积、激活函数、池化） 2、分类识别（全连接层）