3、卷积神经网络

卷积神经网络是一种特殊的人工神经网络，在图像处理方面有广泛应用。大部分应用场景下，图像对神经网络来说是非常高维的输入。如，一张1920\*1080的RGB三通道图片，组织成列向量输入给网络的话就有622万维。如果使用全连接的神经网络，每相邻层之间的神经网络均互相相连，则权重参数的个数将随神经元数量平方增长。对于图像这种高维输入，权重参数过多，巨大的计算量使训练和应用都变得不现实。

卷积神经网络利用了图像本身的局部性（相距较近像素的相关性强），提出了新的网络连接方式——每一层的神经元只和上一层的局部相连，一般格子取3x3、5x5、7x7等尺寸，【如图所示、计算公式】称为卷积核。则，前向计算相当于以卷积核为滤波器，对上一层提供的输入（称特征图，feature map）作二维离散卷积操作。

基于如下假设：如果某个滤波器接受一片二维的输入特征图区域后卷积出了有用的信息，那么这个滤波器一定也能在别的区域得到有用信息，可以在滑动过程中使用同一个滤波器。遍历过程中卷积核的权重保持不变，实现权重共享。

卷积神经网络靠局部连接和权重共享，相比全连接网络计算量大幅减少，而且因为利用了图像本身的局部性，图像内部的物体发生平移等变换不会引起网络的性能下降，故在计算机视觉领域有优秀的性能。

3.1、卷积层

卷积层是CNN的主要组成层，其原理如上述。

描述该层如何卷积需要如下几个预定的超参数：

1、深度（Depth），又称频道数（Channels），控制该层每个神经元与多少个上一层的神经元区域相连。如，CNN输入层接受RGB三通道图像时，输入层的Depth=3；

2、卷积核大小（Kernel Size），规定了连接区域的大小，即卷积滤波器的大小。常用的区域大小有3x3，5x5，7x7。AlexNet【来源请求】原论文的网络使用3x3的卷积核；

3、步长（Stride），规定了卷积核在输入特征图上每次滑动移过多少像素，常用值为1、2，AlexNet取1；

4、填充（Padding），规定卷积核在滑动到图像边缘，滤波器部分格子在图像以外时，采用什么值在图像边界外扩充一段，使卷积计算合法。AlexNet取图片边缘外为零。

显然，这几个超参数不是互相独立的。卷积层神经元个数满足关系(W-K+2P)/S+1。

由于卷积层在维数较高的空间上操作，计算量也很大。【来源请求】显示AlexNet 90%的浮点数集中在卷积层，故该层是我们的优化重点。

3.2、ReLU激活函数

ReLU函数的定义：【公式，函数图像】

AlexNet提出，使用不同的激活函数时，损失函数收敛到极小值的速度不同【对比图】。对卷积神经网络，ReLU激活函数值没有上界，收敛较快，故得到了包括AlexNet在内的许多CNN模型的应用。

3.3、池化层（Pooling Layer）

池化层对输入进行下采样操作，以达到对输入的降维，同时可以提高网络对输入图像的空间不变性。最常用的池化方法是Max Pooling（每个下采样窗口保留值最大的）和Average Pooling（每个下采样窗口取包含特征图像素的平均值）

3.4、全连接层

全连接层接受前面卷积层的输出，最后得到图像的类别，和普通全连接神经网络的分类问题一样。AlexNet【来源请求】使用5层卷积层接3层全连接层的网络。由于全连接层的输入是已经经过卷积层和池化层充分降维的数据，引入全连接层不会增加过多的计算量。