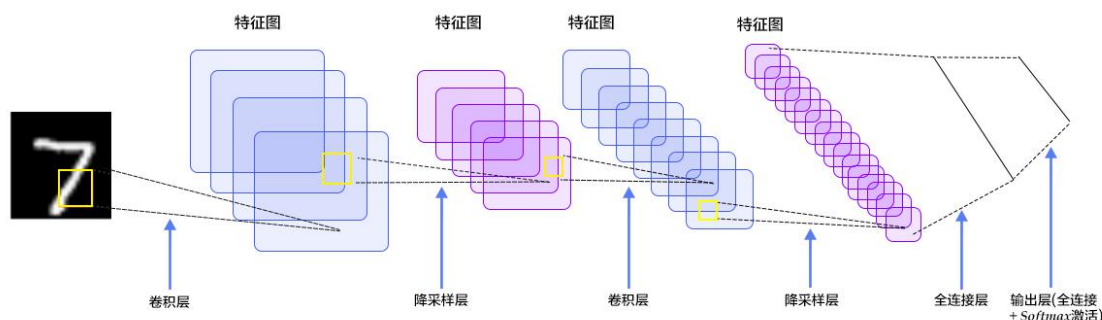


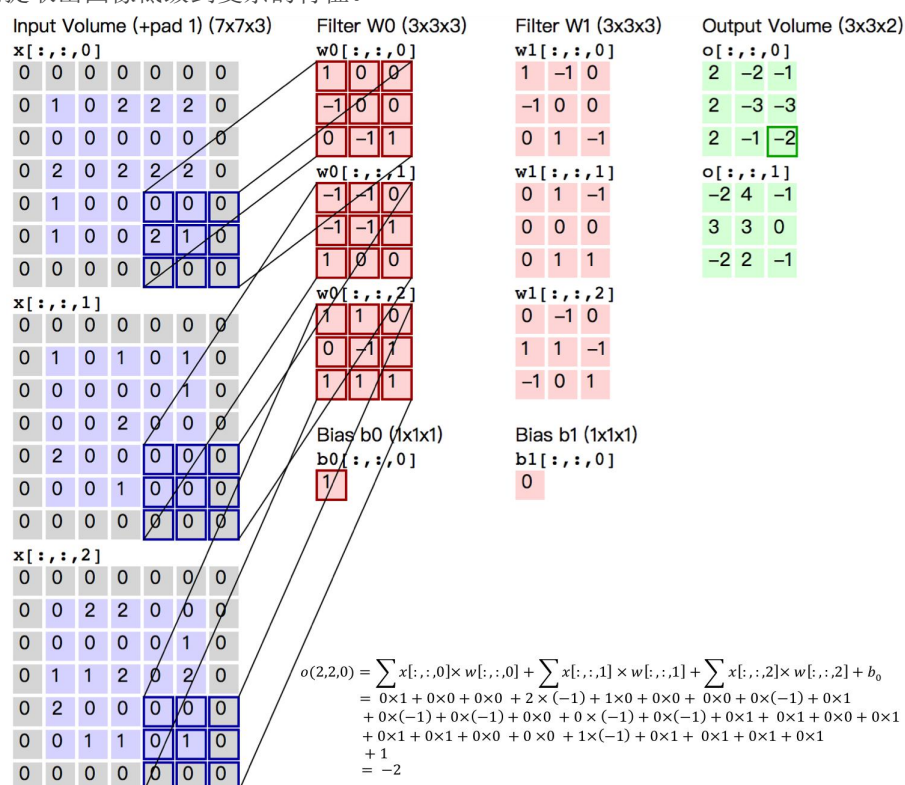
What is CNN?

卷积神经网络能够很好的利用图像的结构信息。下图是一个简单的网络结构：输入的二维图像，先经过两次卷积层到池化层，再经过全连接层，最后使用 softmax 分类作为输出层。



1. 卷积层

卷积层是卷积神经网络的核心。在图像识别中提到的卷积指二维卷积，即离散二维滤波器（也称卷积核）与二维图像做卷积操作，简单来说就是二维滤波器滑动到二维图像上的所有位置，并在每个位置与该像素点及其邻域像素点做内积。卷积操作被广泛应用于图像处理领域，不同卷积核可以提取不同的特征，例如边缘、线性、角等特征。在深层卷积神经网络中，通过卷积操作可以提取出图像低级到复杂的特征。



上图为一个卷积计算过程的示例图，输入图像大小为 H=5, W=5, D=3，即 5×5 大小的 3 通道（RGB，也称作深度）彩色图像。这个示例图中包含两（用 K 表示）组卷积核，即图中滤波器 W0 和 W1。在卷积计算中，通常对不同的输入通道采用不同的卷积核，如图示例中每组卷积核包含（D=3）个 3×3（用 F×F 表示）大小的卷积核。另外，这个示例中卷积核在图像的水平方向（W

方向)和垂直方向(H方向)的滑动步长为2(用S表示);对输入图像周围各填充1(用P表示)个0,即图中输入层原始数据为蓝色部分,灰色部分是进行了大小为1的扩展,用0来进行扩展。经过卷积操作得到输出为 $3 \times 3 \times 2$ (用 $H_o \times W_o \times K$ 表示)大小的特征图,即 3×3 大小的2通道特征图,其中 H_o 计算公式为: $H_o = (H - F + 2 \times P) / S + 1$, W_o 同理。而输出特征图中的每个像素,是每组滤波器与输入图像每个特征图的内积再求和,再加上偏置bo,偏置通常对于每个输出特征图是共享的。输出特征图 $o[:, :, 0]$ 中的最后一个-2计算如上图右下角公式所示。

H: 图片高度; W: 图片宽度; D: 原始图片通道数,也是卷积核个数; F: 卷积核高宽大小; P: 图像边扩充大小; S: 滑动步长。

在卷积操作中卷积核是可学习的参数,经过上面示例介绍,每层卷积的参数大小为 $D \times F \times F \times K$ 。卷积层的参数较少,这也是由卷积层的主要特性即局部连接和共享权重所决定。

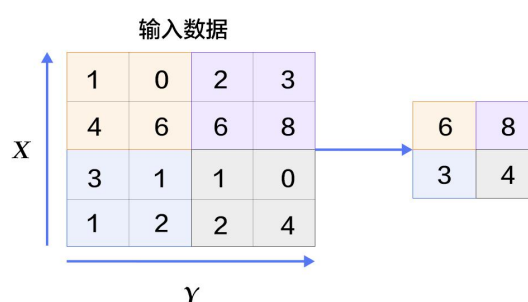
局部连接: 每个神经元仅与输入神经元的一块区域连接,这块局部区域称作感受野(receptive field)。在图像卷积操作中,即神经元在空间维度(spatial dimension,即上图示例H和W所在的平面)是局部连接,但在深度上是全部连接。对于二维图像本身而言,也是局部像素关联较强。这种局部连接保证了学习后的过滤器能够对于局部的输入特征有最强的响应。局部连接的思想,也是受启发于生物学里面的视觉系统结构,视觉皮层的神经元就是局部接受信息的。

权重共享: 计算同一个深度切片的神经元时采用的滤波器是共享的。例上图中计算 $o[:, :, 0]$ 的每个神经元的滤波器均相同,都为 W_0 ,这样可以很大程度上减少参数。共享权重在一定程度上讲是有意义的,例如图片的底层边缘特征与特征在图中的具体位置无关。但是在一些场景中是无意的,比如输入的图片是人脸,眼睛和头发位于不同的位置,希望在不同的位置学到不同的特征。请注意权重只是对于同一深度切片的神经元是共享的,在卷积层,通常采用多组卷积核提取不同特征,即对应不同深度切片的特征,不同深度切片的神经元权重是不共享。另外,偏重对同一深度切片的所有神经元也都是共享的。

通过介绍卷积计算过程及其特性,可以看出卷积是线性操作,并具有平移不变性(shift-invariant),平移不变性即在图像每个位置执行相同的操作。卷积层的局部连接和权重共享使得需要学习的参数大大减小,这样也有利于训练较大卷积神经网络。

2. 池化层

池化是非线性下采样的一种形式,主要作用是通过减少网络的参数来减小计算量,并且能够在一定程度上控制过拟合。通常在卷积层的后面会加上一个池化层。池化包括最大池化、平均池化等。其中最大池化是用不重叠的矩形框将输入层分成不同的区域,对于每个矩形框的数取最大值作为输出层,如下图所示。



@author: 专业拉网线