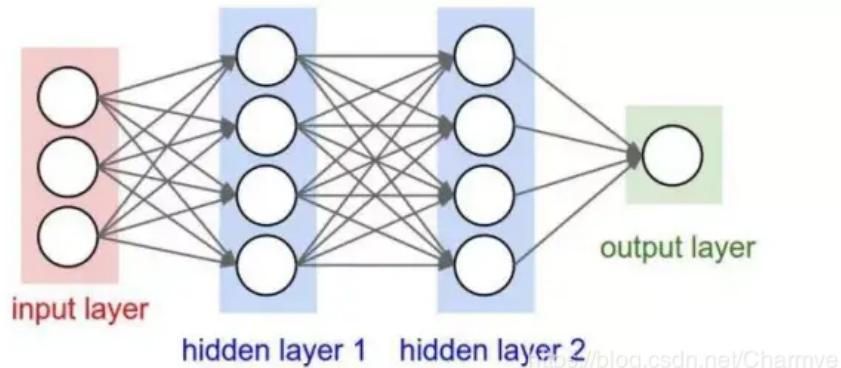


回顾和复习常见神经网络结构的底层原理：

Day2: CNN

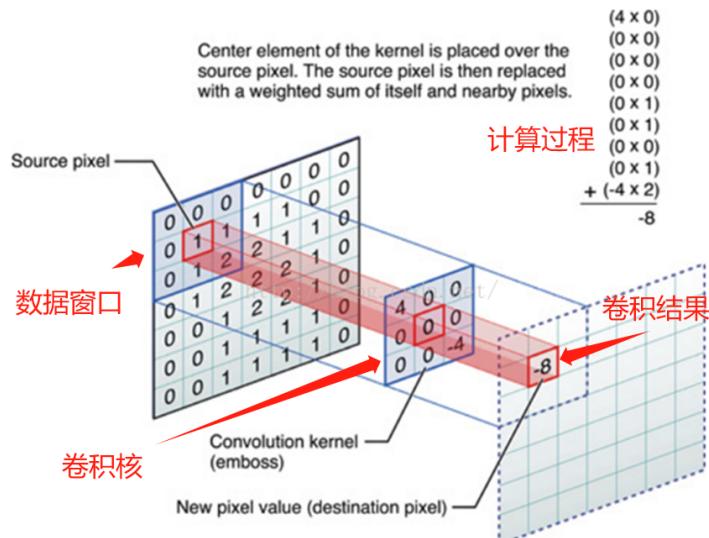
传统神经网络处理图像的本质还是展开二维图像进行权重训练，因此特征的辨别受限于具体平面的位置。



而CNN在卷积层通过使用卷积操作，从而捕捉到图像中的局部特征而不受其分布位置的影响。

1. 卷积

定义：卷积操作是指将一个可移动的小窗口与图像进行逐元素相乘然后相加的操作。这个小窗口其实是一组固定的权重，它可以被看作是一个特定的滤波器（filter）或卷积核。



如图，需要被卷积计算的部分是：

0	0	0
0	1	1
0	1	2

此时卷积核的权重是：

4	0	0
0	0	0
0	0	-4

处理后的结果计算为（数据窗口中的数据与卷积核逐个元素相乘再求和）： $4 \times 0 + 0 \times 0 + 0 \times 0 + 0 \times 0 + 0 \times 1 + 0 \times 1 + 0 \times 0 + 0 \times 1 + (-4) \times 2 = -8$

2. 卷积神经网络结构

一个卷积神经网络主要由以下5层组成：

(1) 数据输入层

该层要做的处理主要是对原始图像数据进行预处理，包括取均值、归一化、PCA降维

输入层接收原始图像数据。图像通常由三个颜色通道（红、绿、蓝）组成，形成一个二维矩阵，表示像素的强度值。

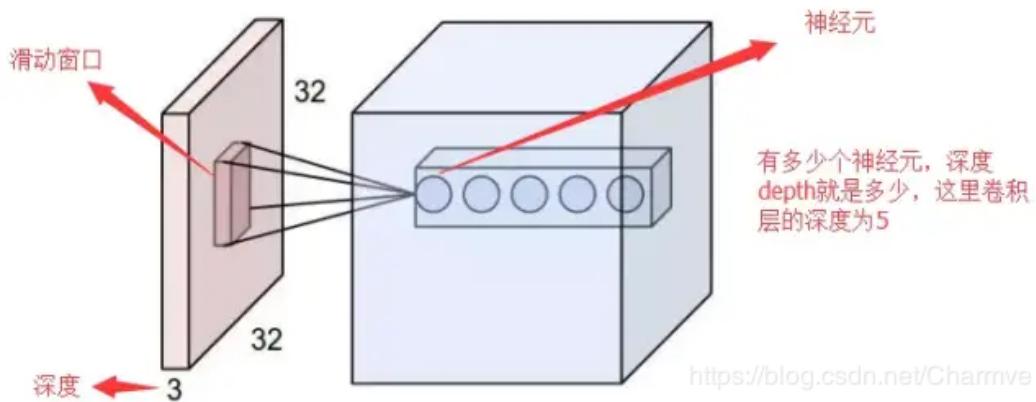
(2) 卷积计算层

卷积操作需要考虑以下参数：

步长：窗口一次滑动的长度。

步长越大，获得重叠更少的接受域和更小的空间维度

卷积核的个数：决定输出的depth厚度。同时代表卷积核的个数。



每一个卷积核都要计算最后叠在一起

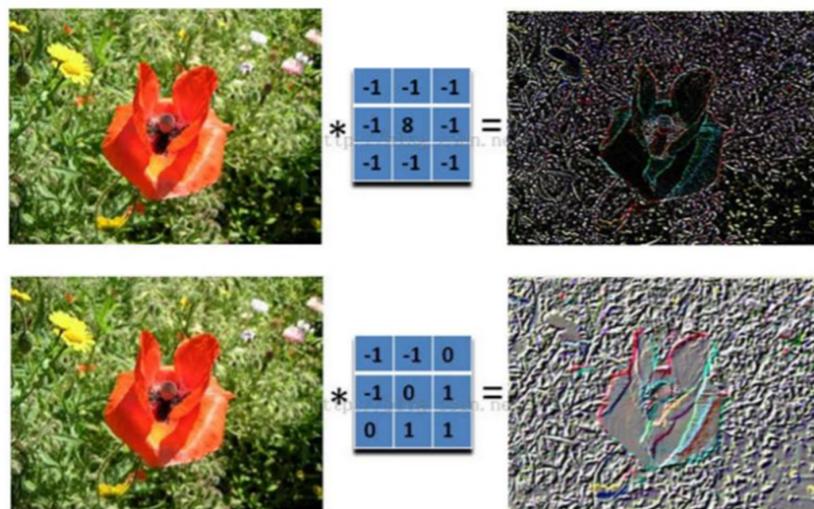
zero-padding: 在外围边缘补充若干圈0, 方便从初始位置以步长为单位可以刚好滑倒末尾位置, 通俗地讲就是为了总长能被步长整除。

$$\text{zero_padding} = (k - 1)/2$$

$$O = (W - K + 2P)/S + 1$$

其中 O 是输出尺寸, K 是卷积核尺寸, P 是填充, S 是步幅。

参数共享机制: 神经元就是图像处理中的滤波器, 在卷积层中每个神经元连接数据窗的权重是固定的, 每个神经元只关注一个特性。



卷积层的每个滤波器都会都有自己所关注一个图像特征, 比如垂直边缘, 水平边缘, 颜色, 纹理等等, 这些所有神经元加起来就好比就是整张图像的特征提取器集合。

(3) 非线性层/激活层

把卷积层输出结果做非线性映射, CNN采用的激活函数一般为ReLU(The Rectified Linear Unit), 它的特点是收敛快, 求梯度简单, 但较脆弱。

(4) 池化层

池化层夹在连续的卷积层中间, 用于压缩数据和参数的量, 减小过拟合。简而言之, 如果输入是图像的话, 那么池化层的主要作用就是压缩图像。

- 特征不变性: 去掉无关紧要的信息
- 特征降维: 把最重要的特征提取出来
- 一定程度上防止过拟合, 更方便优化

(5) 全连接层

两层之间所有神经元都有权重连接，通常全连接层在卷积神经网络尾部。也就是跟传统的神经网络神经元的连接方式是一样的。

