

# 卷积神经网络

## 一、卷积神经网络的层级结构

### 1. 数据输入层

\*对原始图像数据进行预处理

\*\*去均值

\*\*归一化：减少各维度数据取值范围的差异而带来的干扰

\*\*PCA/白化：用PCA（主成分分析）方法降维；白化是对数据各个特征轴上的幅度归一化，目的是去掉数据之间的相关关联度。

### 2. 卷积计算层

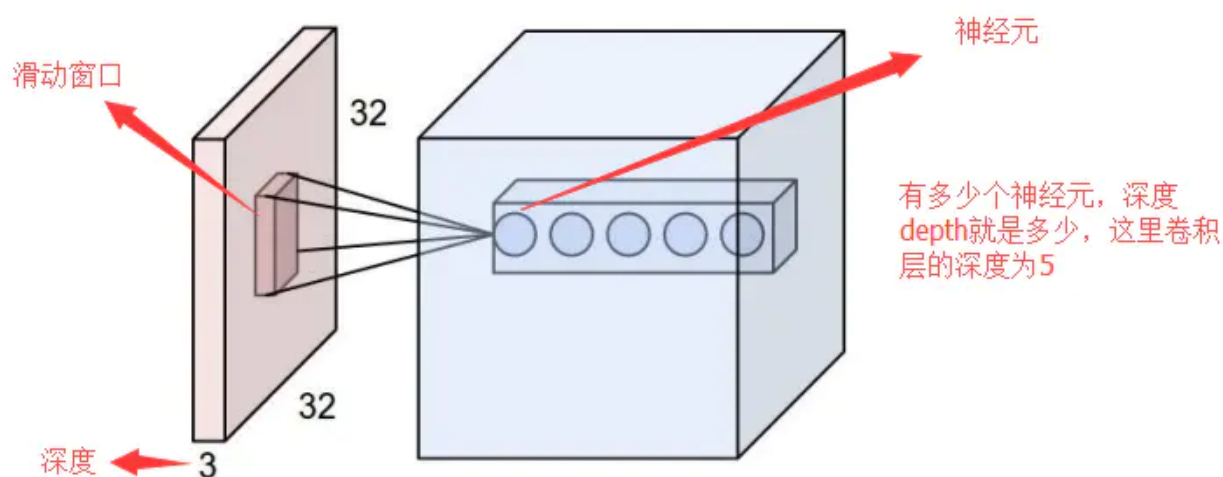
\*这一层是卷积神经网络最重要的一层，有两个关键操作：一个是局部关联，每个神经元看做一个滤波器。另一个是窗口滑动，滤波器对局部数据计算。

\*几个名词

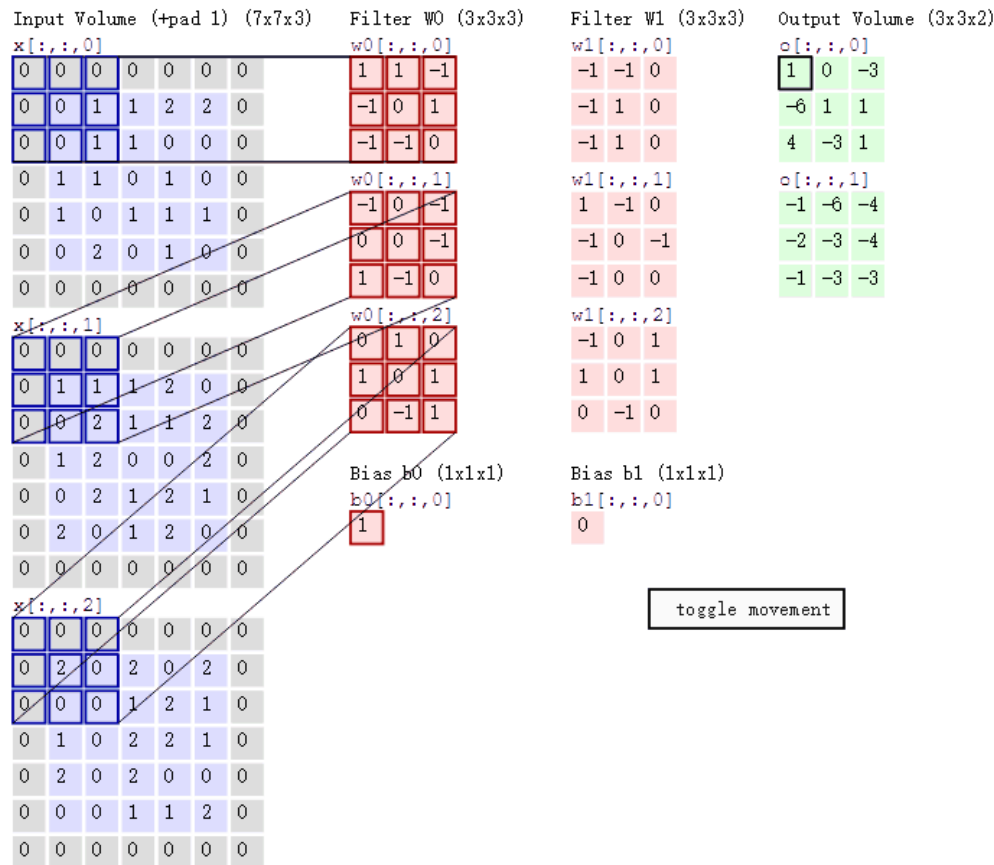
\*\*深度/depth

\*\*步长/stride（窗口一次滑动的长度）

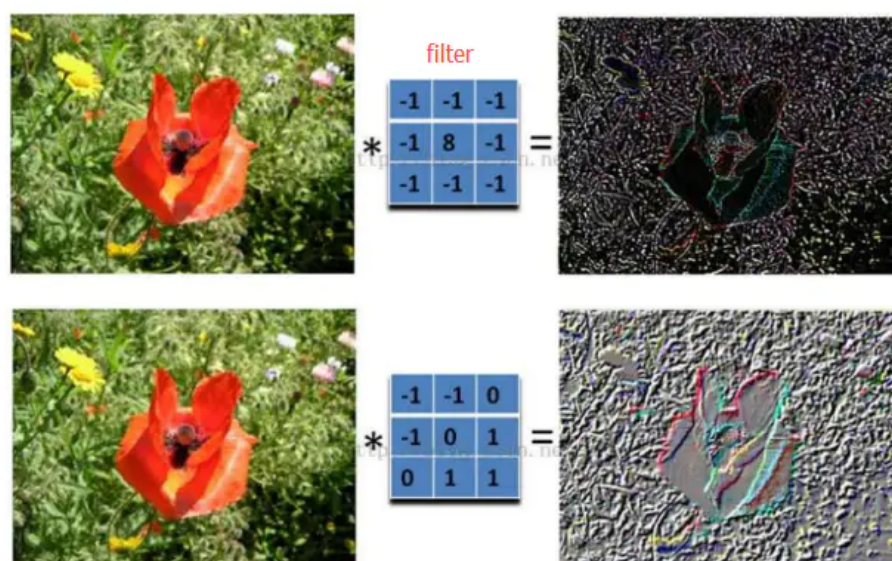
\*\*填充值/zero-padding：填充输入矩阵使得滑动能够完成。



\*\*下图两个粉色的是两个神经元。蓝色的是输入矩阵，其中灰色的是填充值。绿色矩阵是卷积层输出的结果。



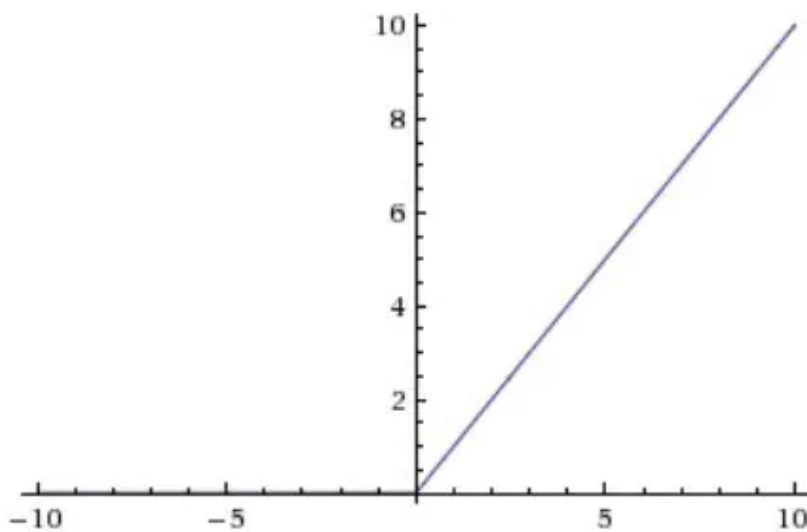
\*在卷积层中每个神经元连接数据窗的权重是固定的，每个神经元只关注一个特性。神经元就是图像处理中的滤波器，比如边缘检测专用的Sobel滤波器，即卷积层的每个滤波器都会有自己所关注一个图像特征，比如垂直边缘，水平边缘，颜色，纹理等等，这些所有神经元加起来就好比就是整张图像的特征提取器集合。



### 3. 激活s层

\*把卷积层的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/修正线性单元)，它的特点是收敛快，求梯度简单，但较脆弱。

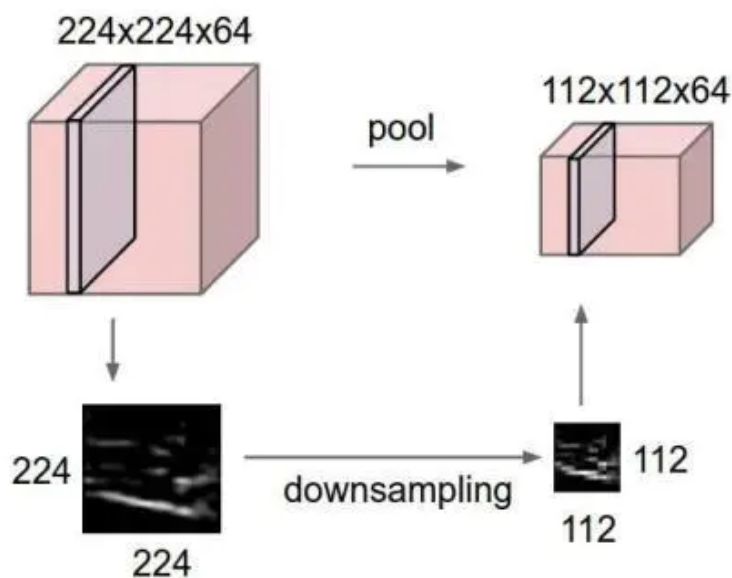


#### 4.池化层

\*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，那么池化层的最主要作用就是压缩图像。

\*池化层的具体作用如下。

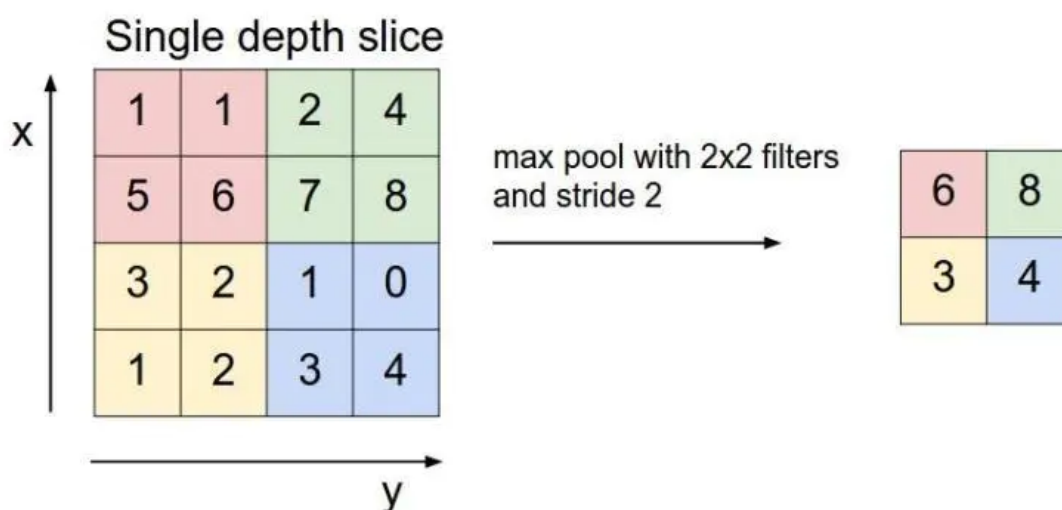
\*\*特征不变性，也就是我们在图像处理中经常提到的特征的尺度不变性，池化操作就是图像的resize，平时一张狗的图像被缩小了一倍我们还能认出这是一张狗的照片，这说明这张图像中仍保留着狗最重要的特征，我们一看就能判断图像中画的是一只狗，图像压缩时去掉的信息只是一些无关紧要的信息，而留下的信息则是具有尺度不变性的特征，是最能表达图像的特征。



**\*\*特征降维**，我们知道一幅图像含有的信息是很大的，特征也很多，但是有些信息对于我们做图像任务时没有太多用途或者有重复，我们可以把这类冗余信息去除，把最重要的特征抽取出来，这也是池化操作的一大作用。

**\*\*在一定程度上防止过拟合**，更方便优化。

\*池化的方法有Max pooling和average pooling，常用的是Max pooling，下图是Maxpooling的展示，对于每个2\*2的窗口选出最大的数作为输出矩阵的相应元素的值。



## 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

## 6. 一般CNN结构依次为

\*输入层

\*【卷积层—》激励层】\* $N$  —》池化层?\* $M$

\*【全连接层—》激励层】\* $K$

\*全连接的输出层

## 二、卷积神经网络之训练算法

\*同一般机器学习算法，先定义Loss function，衡量和实际结果之间差距。

\*找到最小化损失函数的 $W$ 和 $b$ ， CNN中用的算法是SGD（随机梯度下降）。

## 三、卷积神经网络之优缺点

\*优点

\*\*共享卷积核，对高维数据处理无压力

\*\*无需手动选取特征，训练好权重，即得特征分类效果好

\*缺点

\*\*需要调参，需要大样本量，训练最好要GPU

\*\*物理含义不明确（也就是说，我们并不知道每个卷积层到底提取到的是什么特征，而且神经网络本身就是一种难以解释的“黑箱模型”）