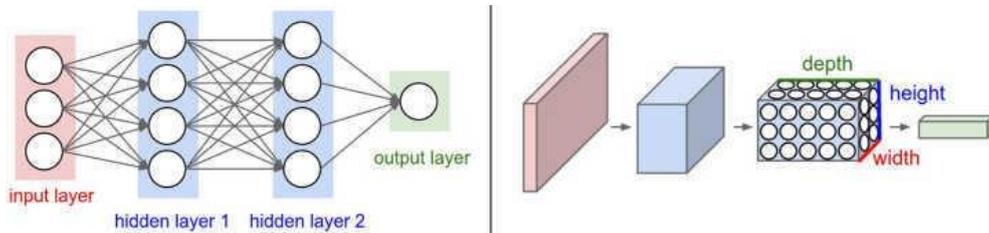


卷积神经网络简述

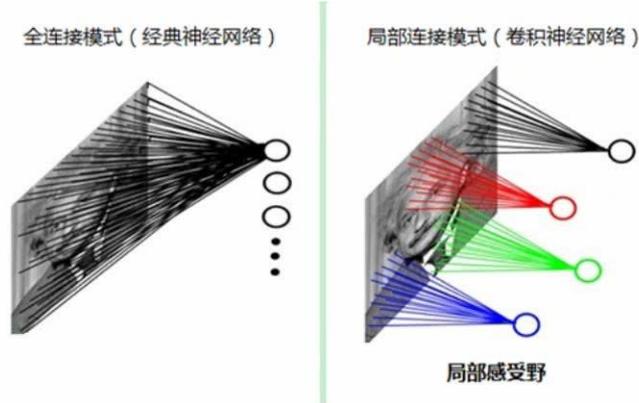
Convolutional Neural Networks, CNN



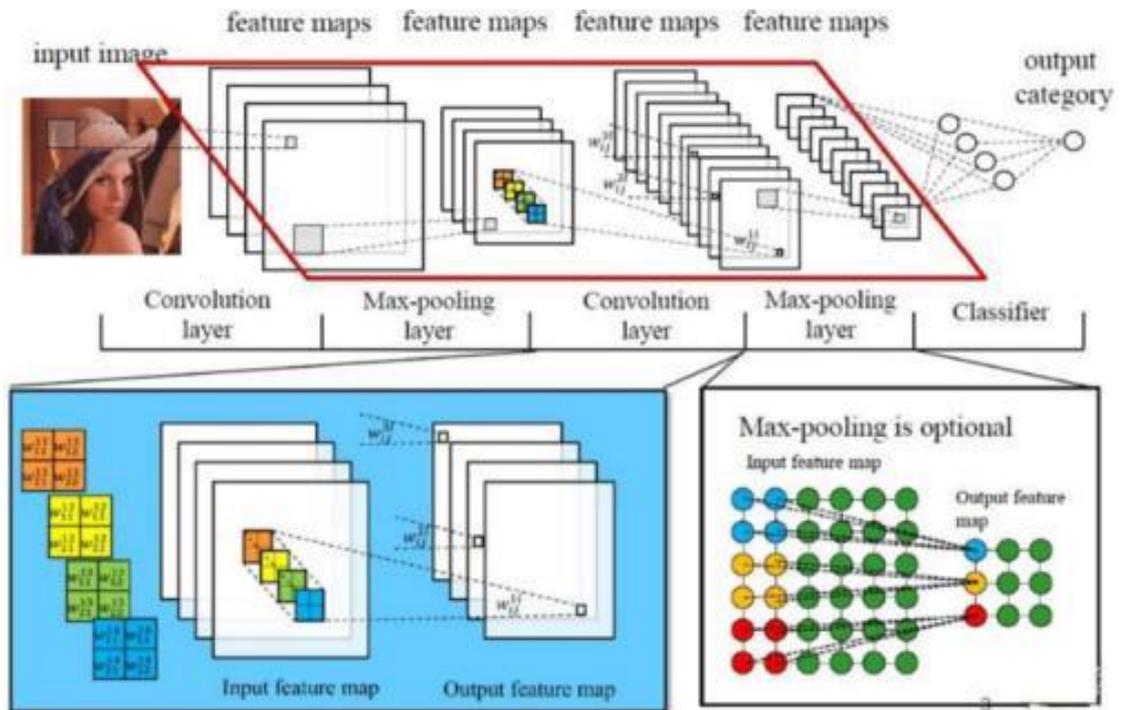
1 卷积神经网络(CNN)结构

1.1 全连接神经网络的劣势

与局部连接模式相比，全连接模式的参数太多，训练效率较低；同时大量的参数也很快会导致网络过拟合。



1.2 卷积神经网络的结构



卷积神经网络由三大块组成：输入层、隐含层和输出层。

其中隐含层是卷积神经网络的核心所在，如上图红框内所示，隐含层又由卷积层、池化层和全连接层组成。

2 输入层

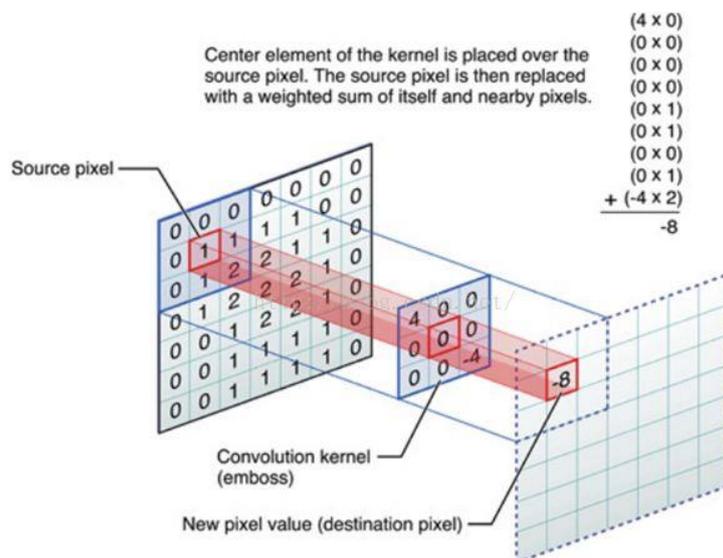
卷积神经网络的输入层可以处理多维数据，常见地，一维卷积神经网络的输入层接收一维或二维数组，其中一维数组通常为时间或频谱采样；二维数组可能包含多个通道；二维卷积神经网络的输入层接收二维或三维数组；三维卷积神经网络的输入层接收四维数组。由于卷积神经网络在计算机视觉领域应用较广，因此许多研究在介绍其结构时预先假设了三维输入数据，即平面上的二维像素点和 RGB 通道。

与其它神经网络算法类似，由于使用梯度下降算法进行学习，卷积神经网络的输入特征需要进行标准化处理。具体地，在将学习数据输入卷积神经网络前，需在通道或时间/频率维对输入数据进行归一化，若输入数据为像素，也可将分布于 $[0, 255]$ 的原始像素值归一化至区间 $[0, 1]$ 。输入特征的标准化有利于提升卷积神经网络的学习效率和表现。

3 隐含层原理

3.1 卷积层

3.1.1 卷积



如上图的操作就被称为卷积，卷积核的大小就是卷积层的感受野。

3.1.2 卷积核

卷积层的功能是对输入数据进行特征提取，其内部包含多个卷积核，组成卷积核的每个元素都对应一个权重系数和一个偏差量 (bias vector)，类似于一个前馈神经元的神经元。同时可以存在多个卷积核，即多卷积核。

3.1.3 卷积层参数

卷积层参数包括**卷积核大小**、**步长**和**填充**，三者共同决定了卷积层输出特征图的尺寸。

其中卷积核大小可以指定为小于输入图像尺寸的任意值，卷积核越大，可提取的输入特征越复杂^[1]。卷积步长定义了卷积核相邻两次扫过特征图时位置的距离，卷积步长为 1 时，卷积核会逐个扫过特征图的元素，步长为 n 时会在下一次扫描跳过 n-1 个像素^[33]。

由卷积核的交叉相关计算可知，随着卷积层的堆叠，特征图的尺寸会逐步减小，例如 16×16 的输入图像在经过单位步长、无填充的 5×5 的卷积核后，会输出 12×12 的特征图。为此，填充是在特征图通过卷积核之前人为增大其尺寸以抵消计算中尺寸收缩影响的方法。常见的填充方法为按 0 填充和重复边界值填充 (replication padding)。

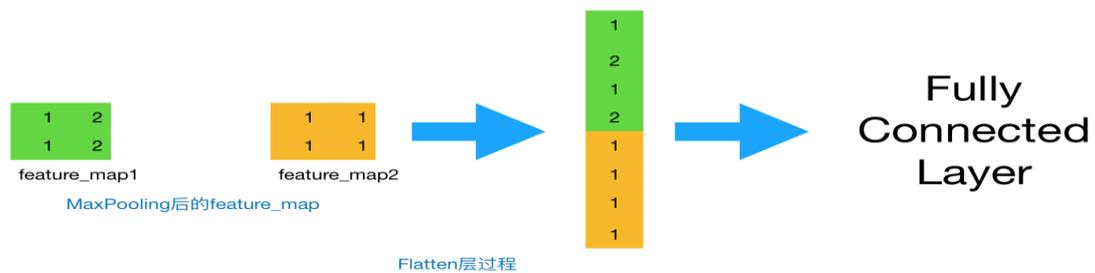
3.1.4 权值共享

在卷积层中权值共享是用来控制参数的数量。假如在一个卷积核中，每一个感受野采用的都是不同的权重值 (卷积核的值不同)，那么这样的网络中参数数量将是十分巨大的。同一个深度的神经元的权值都是相同的。

3.2 池化层

通常在连续的卷积层之间会周期性地插入一个池化层。它的作用是逐渐降低数据体的空间尺寸，这样的话就能减少网络中参数的数量，减少计算资源耗费，也能有效控制过拟合。

普通池化 (General Pooling)：除了最大池化，池化单元还可以使用其他的函数，比如平均池化 (average pooling) 或 L-2 范式池化 (L2-norm pooling)。平均池化历史上比较常用，但是现在已经很少使用了。因为实践证明，最大池化的效果比平均池化要好。



3.3 全连接层

卷积神经网络中的全连接层等价于传统前馈神经网络中的隐含层。全连接层位于卷积神经网络隐含层的最后部分，并只向其它全连接层传递信号。特征图在全连接层中会失去空间拓扑结构，被展开为向量并通过激励函数。

4 输出层

卷积神经网络中输出层的上游通常是全连接层, 因此其结构和工作原理与传统前馈神经网络中的输出层相同。对于图像分类问题, 输出层使用逻辑函数或归一化指数函数 (**softmax function**) 输出分类标签。在物体识别 (**object detection**) 问题中, 输出层可设计为输出物体的中心坐标、大小和分类。在图像语义分割中, 输出层直接输出每个像素的分类结果。