

# 卷积神经网络

xbZhong

2024-06-01

[本页PDF](#)

## Convolutional Neural Network（卷积神经网络）

局部连接：

CNN 不是每个神经元都连全图（像 MLP 那样），而是只看**小块区域**（比如  $3 \times 3$ ）

权重共享：

整个图上，这个卷积核在每个区域滑动时 **用的都是同一套参数**

- 在CNN里面，通常输入的张量维度会很大，在进行识别类别的时候，只需要让每个**Neural**去负责自己的**Receptive field**，从而识别特征，**不需要看整张图片**
- 卷积神经网络(CNN)训练的核心就是学习滤波器(filter)中的参数数值**

## Fully Connected Layer（全连接层）

- 层中的每个神经元都与前一层的所有神经元相连接
- 每个神经元有独立的权重参数
- 通常用于网络末端进行分类或回归
- 参数多，更容易过拟合

## Receptive field（感受野）

- CNN中的每个神经元只连接到输入图像的一个**局部区域**（而不是整个输入），这个局部区域就是该神经元的感受野
- 感受野大小可以不一，可以只考虑某些**channel**
- 彼此之间可以**重叠**，可以学习到图像中的**所有特征**
- 一个感受野可以由**多个**神经元学习

## Typical Setting

*Description of Receptive field:*

- kernel size**（卷积核）：用于执行卷积操作的滤波器（或称为卷积核）
  - 常见大小： $3 \times 3$ ，更大的卷积核可以学习到更多的特征，获得更大的感受野，但也会提高计算成本
- stride**（步长）：步长，包括横向步长和纵向步长
  - 步长一般设置为1或2，需要注意的是图像可以重叠学习，即当步长小于卷积核尺寸时，卷积操作会在输入上产生重叠，这有助于学习更丰富的特征表示
- padding**（补丁）：当感受野（卷积核）超出图像范围，需要对超出范围进行参数补充，称为padding
- 输出尺寸计算**：输出尺寸 =  $(\text{输入尺寸} - \text{卷积核尺寸} + 2 \times \text{padding}) / \text{stride} + 1$

## Parameter Sharing

负责不同感受野的不同神经元的参数完全一致

- 共享参数，使得计算成本低，防止**OverFitting**

## Typical Setting

- 一个感受野会由多个神经元负责学习
- 不同感受野的同组神经元**参数一致**

## Tensor（张量）

- 张量是多维矩阵
- 一张图片可以由三维张量表示
  - 两个维度是图片长和宽
  - 还有一个维度是图片的**Channels（通道）**，可以理解为RGB通道，即三原色

## Convolutional Layer（卷积层）

- 一层卷积层有多个**Filter（卷积核）**
  - 不同的卷积核会在同一个感受野上进行学习
  - 不同的卷积核学习不同的特征
  - 最后这层卷积层的输出是一个**Feature Map**，是三维张量
    - **Feature Map**是三维张量是因为不同的**Filter**学习了不同的特征，一个特征对应一个通道，即通道数直接等于Filter的数量（每个Filter贡献一个通道）

## Pooling（池化）——MaxpPooling

没有要学习的参数！！

- 池化一般在卷积之后进行，是为了让图片变得更小，易于处理，减少运算量
- 一般是几次**Convolution**之后做一次Pooling

## Flatten（展平）

- 无需要学习的参数
- 将**多维输入数据转换为一维向量**
- 常用于连接**卷积层（CNN）**和**全连接层（FC）**之间的过渡

## The Whole CNN

# The whole CNN

