

# 卷积与池化

余荫铠

2022 年 4 月 28 日

## 1 卷积

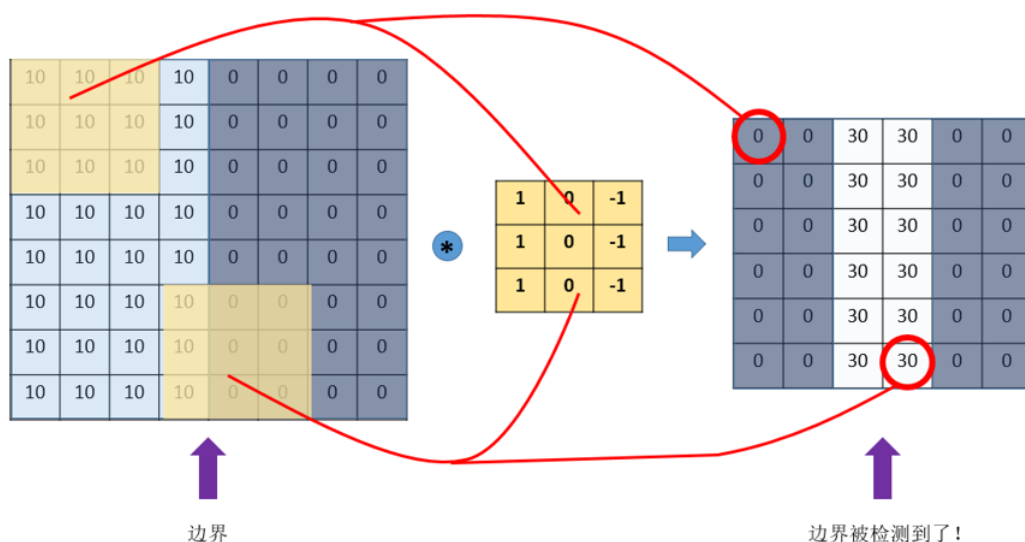


图 1: 卷积示意图

卷积是图像特征提取的一种方法。

根据想要提取的特征构建卷积核，即一个比较小的矩阵，包含了特征信息。

图1的卷积核是一个  $3 \times 3$  的矩阵，要提取的是水平方向的数值梯度特征。将卷积核矩阵与输入矩阵中大小等同于卷积核矩阵的分块按元素相乘再求和（需注意这不是内积），映射到输出矩阵的对于位置的元素上，以同样的方式将卷积核作用在输入矩阵的不同位置上（相当于用卷积核扫描输入矩阵），得到输出矩阵不同位置的元素。

这一过程的数学表达式为

$$\mathbb{M}_{m,n}^{(out)} = \sum_{a,b=1,1}^{A,B} \mathbb{C}_{a,b} \mathbb{M}_{m+a-1,n+b-1}^{(in)} \quad (1)$$

其中  $\mathbb{M}^{(in)}$  为大小为  $N \times M$  的输入矩阵， $\mathbb{M}^{(out)}$  为大小为  $(N - A + 1) \times (M - B + 1)$  的输出矩阵， $\mathbb{C}$  为大小为  $A \times B$  的卷积核。

## 2 池化

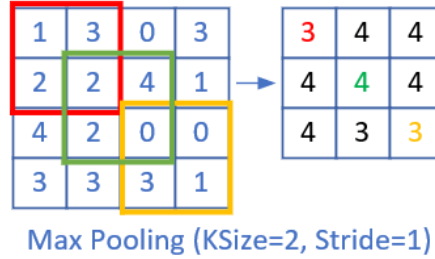


图 2: 池化示意图

池化是一种图像采样压缩的方法。

以图2为例，“池”的大小为  $3 \times 3$ ，从输入矩阵的不同位置的  $3 \times 3$  的局部中提取最大元素或者平均值，映射到输出矩阵的对应位置的元素上，这个过程类似于卷积一样分块扫描整个输入矩阵。图2标了颜色的即是对应的输入和输出。

从扫描方式上，可分为重叠池化和局部池化；从所提取的元素是最大元素还是平均值来看，可分为最大池化和平均池化。图2展示的是重叠池化/最大池化。这种池化的表达式为

$$\mathbb{M}_{m,n}^{(out)} = \max_{a \leq A, b \leq B} \mathbb{M}_{m+a-1, n+b-1}^{(in)} \quad (2)$$