



卷积神经网络

作者: Calvin

QQ: 179209347

Mail: 179209347@qq.com

介绍

笔记简介:

- 面向对象: 深度学习初学者
- 依赖课程: **线性代数, 统计概率**, 优化理论, 图论, 离散数学, 微积分, 信息论

知乎专栏:

<https://zhuanlan.zhihu.com/p/693738275>

Github & Gitee 地址:

https://github.com/mymagicpower/AIAS/tree/main/deep_learning

https://gitee.com/mymagicpower/AIAS/tree/main/deep_learning

* 版权声明:

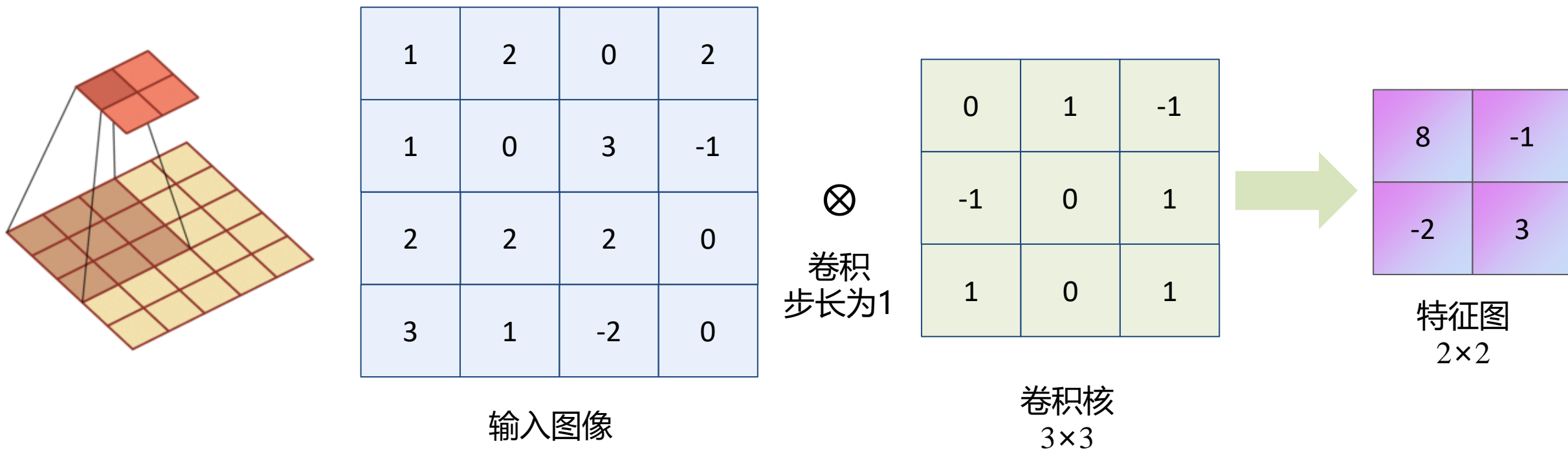
- 仅限用于个人学习
- 禁止用于任何商业用途

卷积层

卷积是一种数学运算，通常用于处理信号、图像和其他类型的数据。在深度学习中，卷积运算是指将一个函数与另一个函数经过翻转和滑动后的乘积的积分。在图像处理中，卷积运算可以用来提取图像的特征，比如边缘、纹理等。

卷积核：

卷积核是一个小的矩阵，用来提取输入数据中的特征。

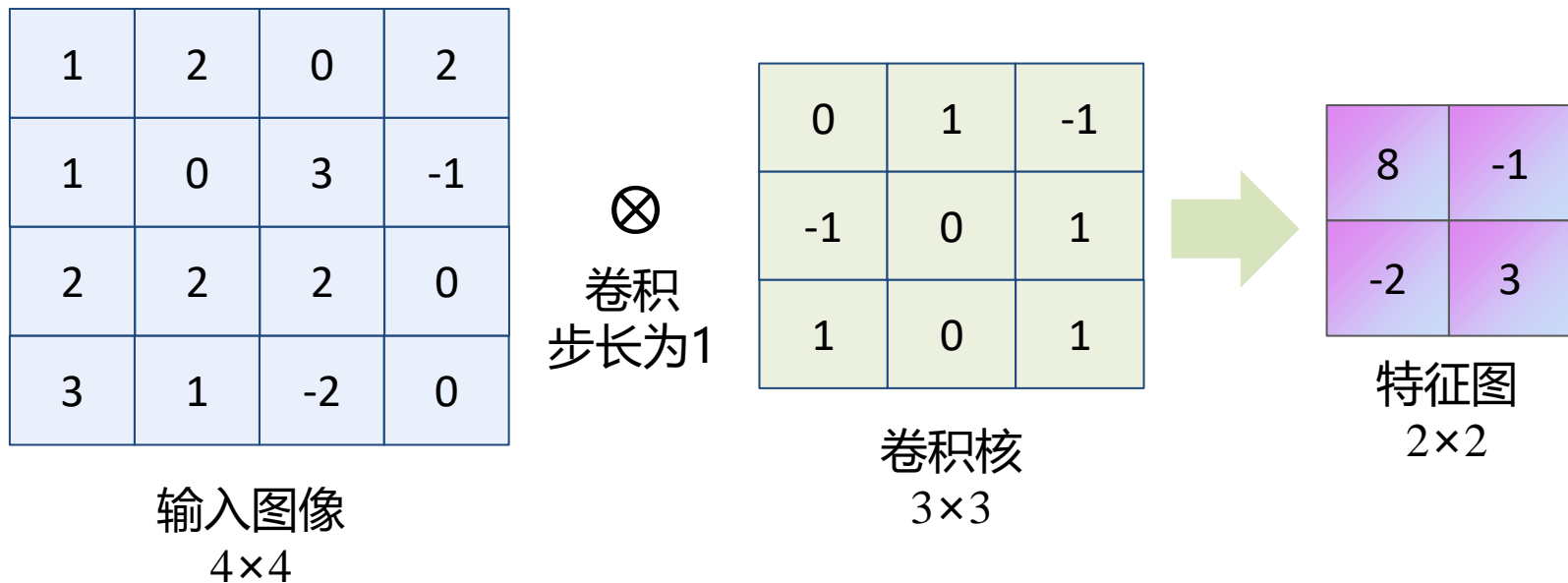


卷积层

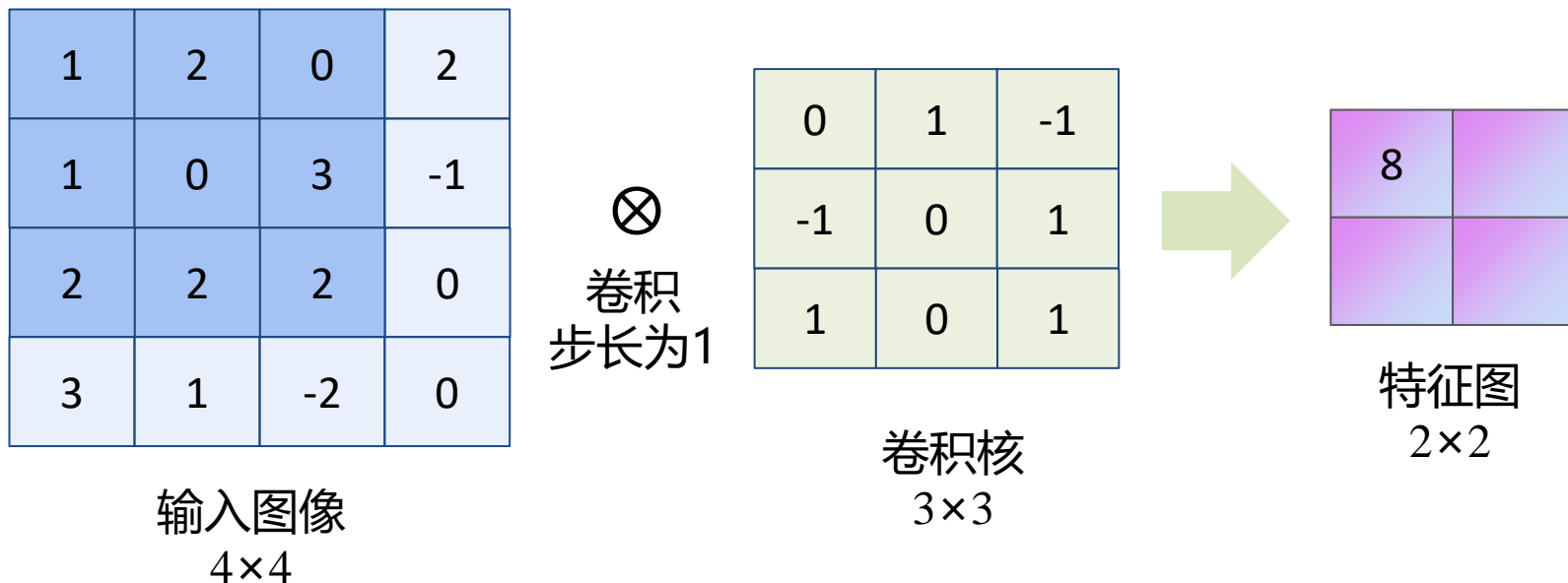
卷积核的尺寸通常设置为 3×3 、 5×5 、 7×7 等。卷积核的权重只需要初始化，然后即可在卷积运算中经梯度下降算法和反向传播算法得到更新优化。

卷积运算的简要步骤：

1. 定义卷积核 (filter/kernel)
2. 对齐卷积核
3. 元素相乘并相加
4. 填充 (padding)
5. 步长 (stride)
6. 输出结果

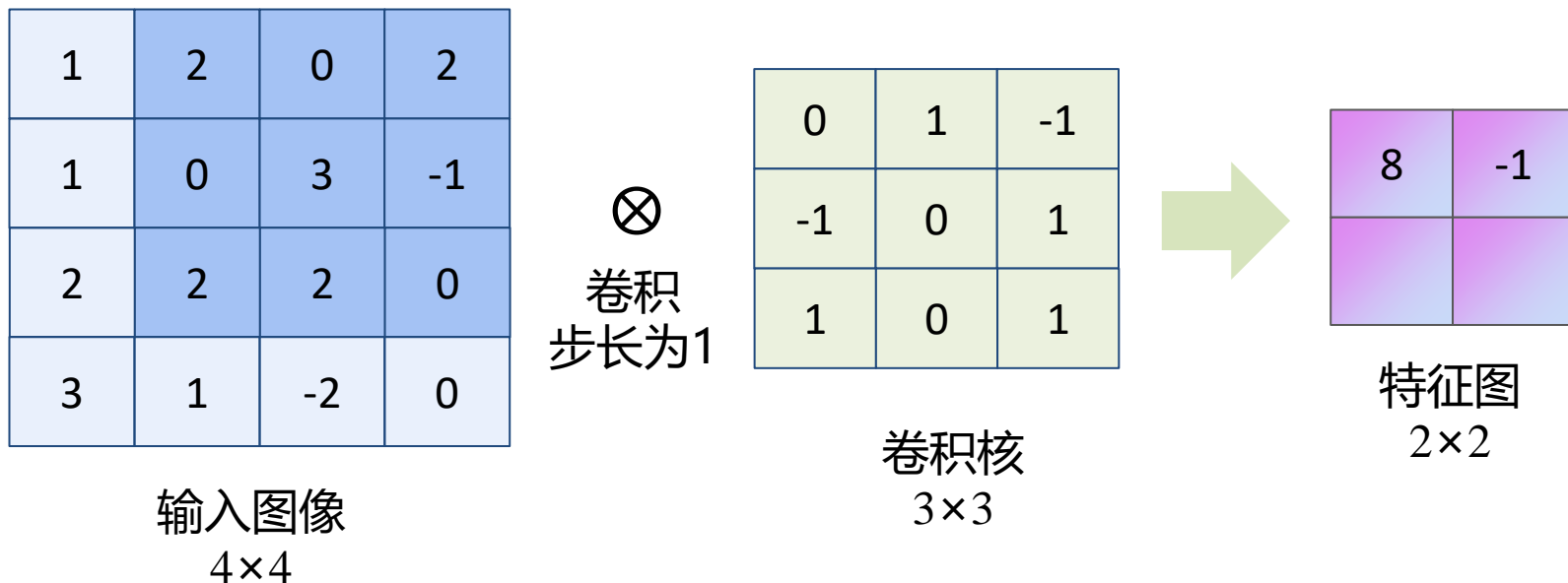


卷积层 - 卷积运算 (1/4)



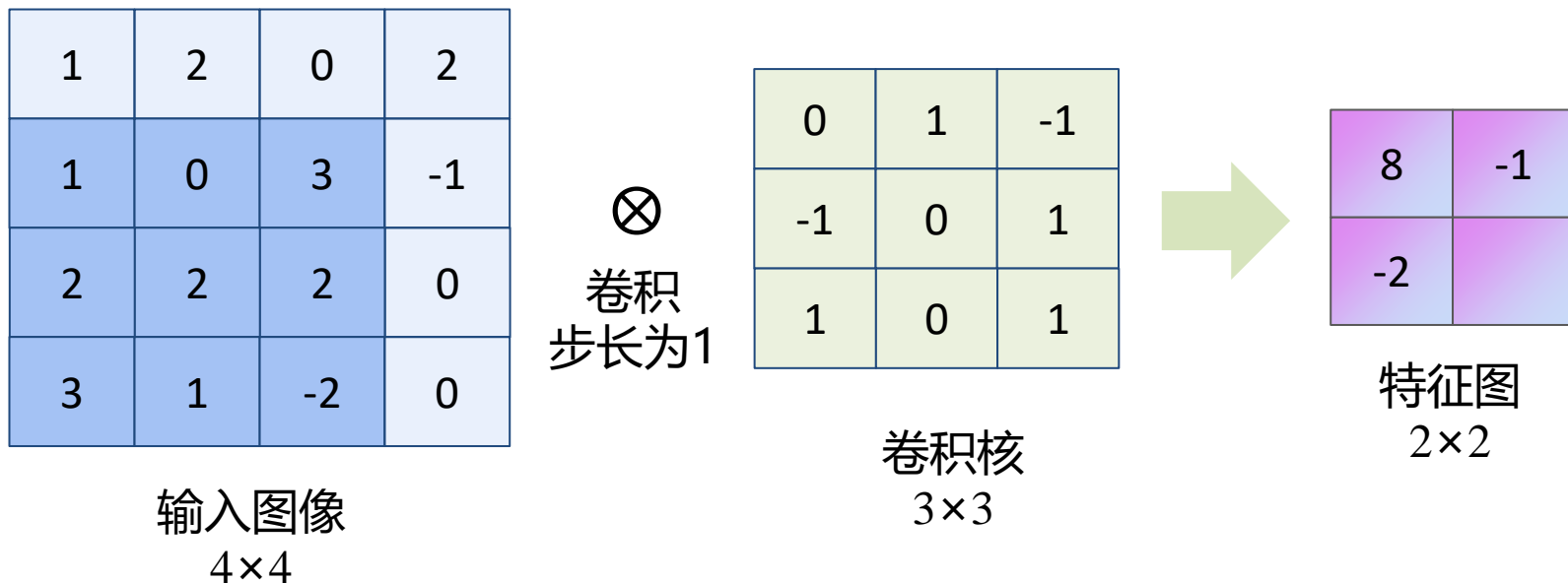
计算过程： $1 \times 0 + 2 \times 1 + 0 \times (-1) + 1 \times (-1) + 0 \times 0 + 3 \times 1 + 2 \times 1 + 2 \times 0 + 2 \times 1 = 8$

卷积层 - 卷积运算 (2/4)



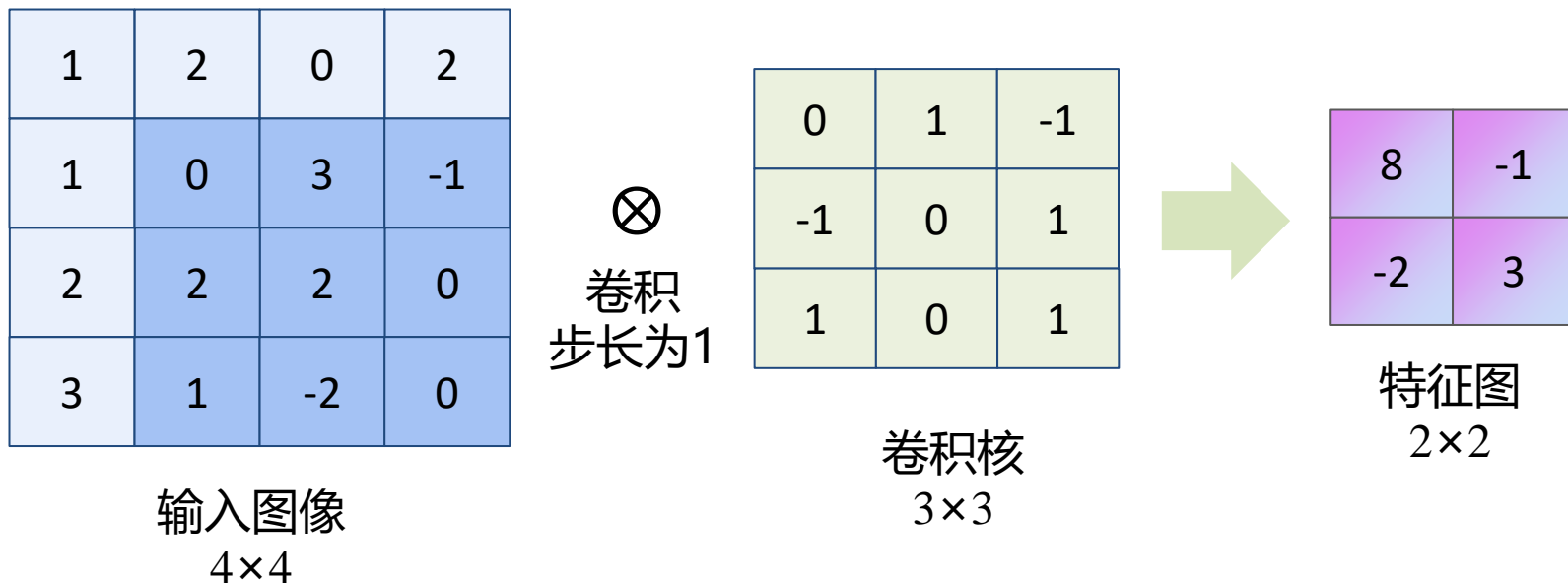
计算过程： $2 \times 0 + 0 \times 1 + 2 \times (-1) + 0 \times (-1) + 3 \times 0 + (-1) \times 1 + 2 \times 1 + 2 \times 0 + 0 \times 1 = -1$

卷积层 - 卷积运算 (3/4)



计算过程： $1 \times 0 + 0 \times 1 + 3 \times (-1) + 2 \times (-1) + 2 \times 0 + 2 \times 1 + 3 \times 1 + 1 \times 0 + (-2) \times 1 = -2$

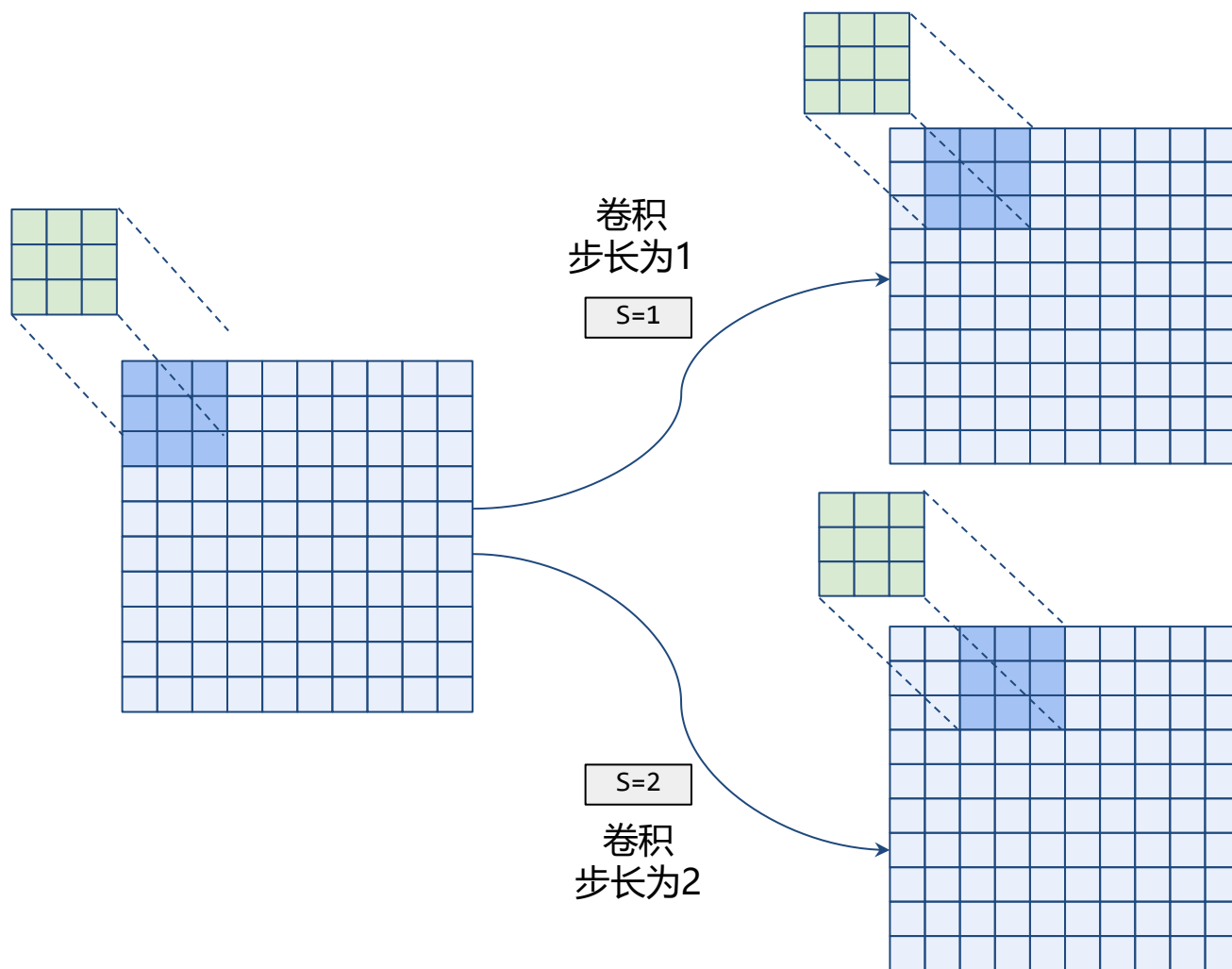
卷积层 - 卷积运算 (4/4)



计算过程： $0 \times 0 + 3 \times 1 + (-1) \times (-1) + 2 \times (-1) + 2 \times 0 + 0 \times 1 + 1 \times 1 + (-2) \times 0 + 0 \times 1 = 3$

卷积层 - 步长

步长 (stride) : 定义卷积核在输入数据上滑动的步长, 通常为1, 也可以是其他值。



卷积层 - 填充

填充 (padding)： 可以选择在输入数据周围填充一些额外的值，以保持输出的尺寸与输入相同，常用的填充方式有"valid"和"same"。

VALID 卷积

VALID卷积为无填充卷积，卷积核在输入图像的有效范围之内移动。

1	2	0	2
1	0	3	-1
2	2	2	0
3	1	-2	0

SAME 卷积

SAME卷积为有填充卷积，即保证输出特征图尺寸与输入图像一致。

0	0	0	0	0	0
0	1	2	0	2	0
0	1	0	3	-1	0
0	2	2	2	0	0
0	3	1	-2	0	0
0	0	0	0	0	0

卷积层 - 填充

可以发现，如果输入图像宽度用 w 表示，高度用 h 表示，卷积核尺寸用 k 表示，步长用 s 表示，填充的层数用 p 表示，那么输出特征图的宽度 w' 和高度 h' 可用如下公式表示。

$$w' = \frac{w-k+2p}{s} + 1 = \frac{4-3+2}{1} + 1 = 4$$

$$h' = \frac{h-k+2p}{s} + 1 = \frac{4-3+2}{1} + 1 = 4$$

1	2	0	2
1	0	3	-1
2	2	2	0
3	1	-2	0



0	0	0	0	0	0
0	1	2	0	2	0
0	1	0	3	-1	0
0	2	2	2	0	0
0	3	1	-2	0	0
0	0	0	0	0	0

⊗
卷积
步长
 $s = 1$

0	1	-1
-1	0	1
1	0	1

卷积核 k
 3×3

- 输入图像
- 宽度 $w = 4$
 - 高度 $h = 4$

卷积层 - 多通道卷积

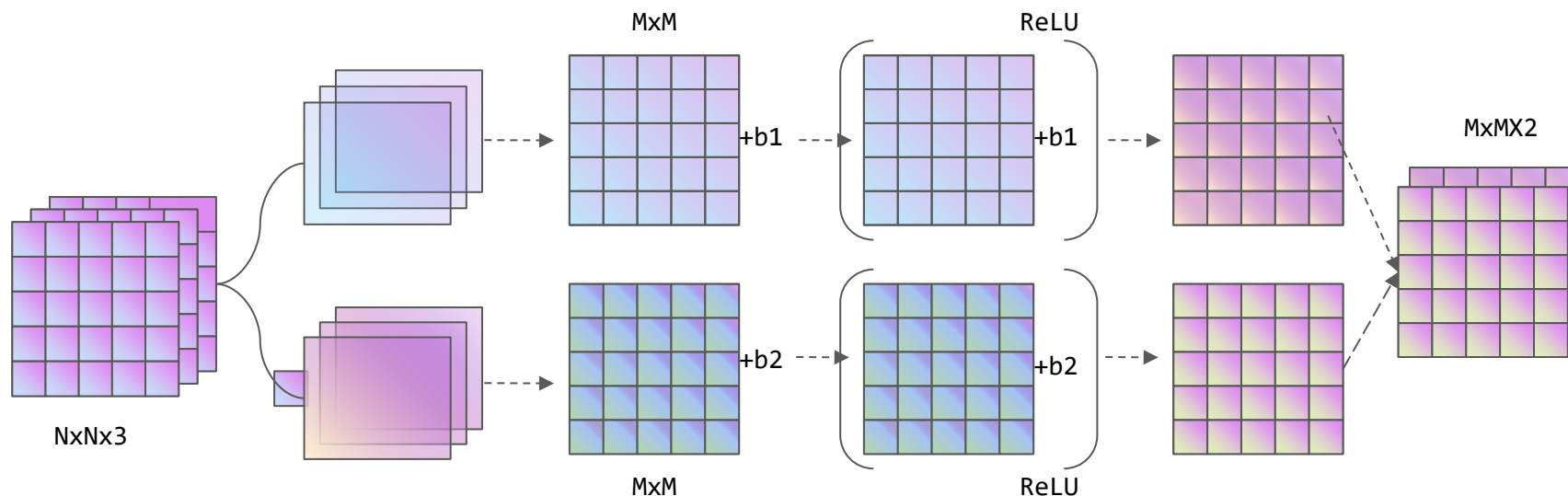
多通道卷积：对于彩色图像，图像通常由三个通道组成：RGB - 红、绿、蓝 三个通道。此时卷积核要变成相应的深度。

卷积核和输入通道数关系如下：

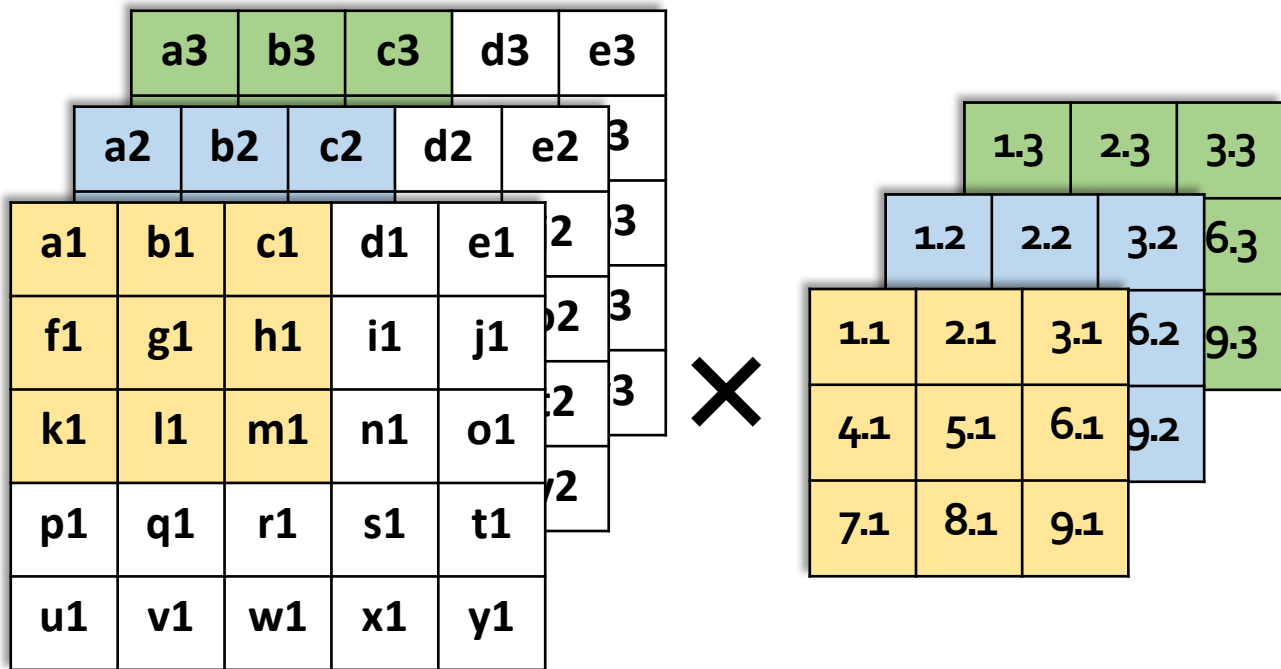
- 卷积核的通道数和卷积层的输入通道数相同。
- RGB图的通道数为3，则卷积核的通道数为3
- 灰度图的通道数为1，则卷积核的通道数为1

卷积核和输出通道数关系如下：

- 卷积核的数量决定卷积层的输出通道数
- 卷积层的输出通道数就是下一层卷积层的输入通道数

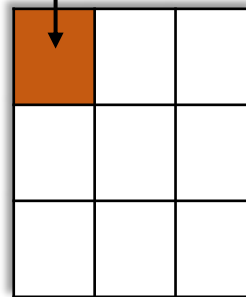


卷积层 - 多通道卷积 - 一个卷积核



$$\Sigma(\Sigma \Sigma \Sigma)$$

```
for(int h = 0; h < 5; h++) {
    for(int w = 0; w < 5; w++) {
        float sum = 0;
        for(int ci = 0; ci < 3; ci++) {
            for(int m = 0; m < 3; m++) {
                for(int n = 0; n < 3; n++) {
                    sum += A[ci][h+m][w+n] * W[ci][m][n];
                }
            }
        }
        B[h][w] = sum;
    }
}
```



卷积层 - 多通道卷积 - 一个卷积核



RGB - 红

0	1	-1	0
2	0	1	1
0	1	1	0
0	2	-1	1

RGB - 绿

1	0	0	2
2	0	1	-1
0	2	1	0
1	1	-1	0

RGB - 蓝

1	2	0	2
1	0	3	-1
0	2	0	0
3	1	-2	0

卷积核的尺寸为 $3 \times 3 \times 3$

\otimes

0	1	0
0	0	1
1	1	0

3

\otimes

-1	1	0
-1	0	1
-1	1	1

1

\otimes

0	1	-1
-1	0	1
1	0	1

4

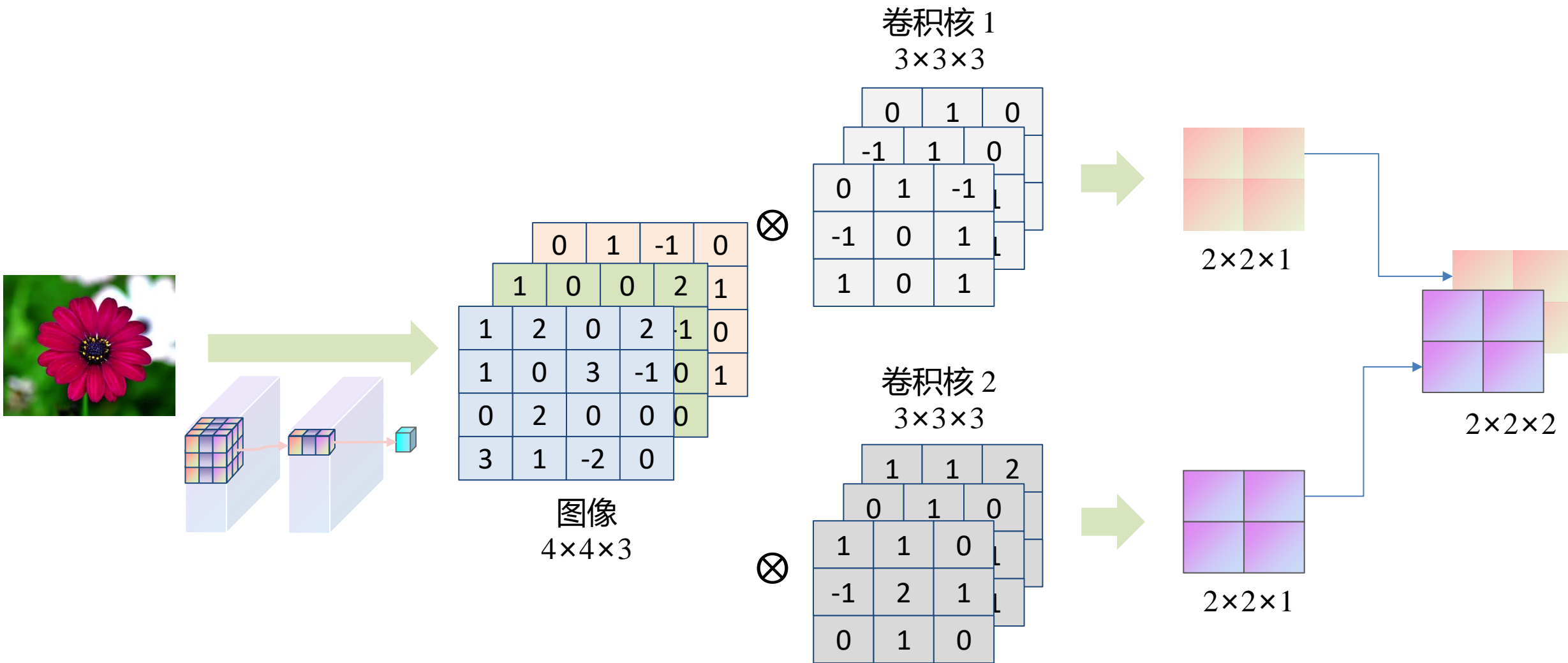
偏置 $b=1$

9	

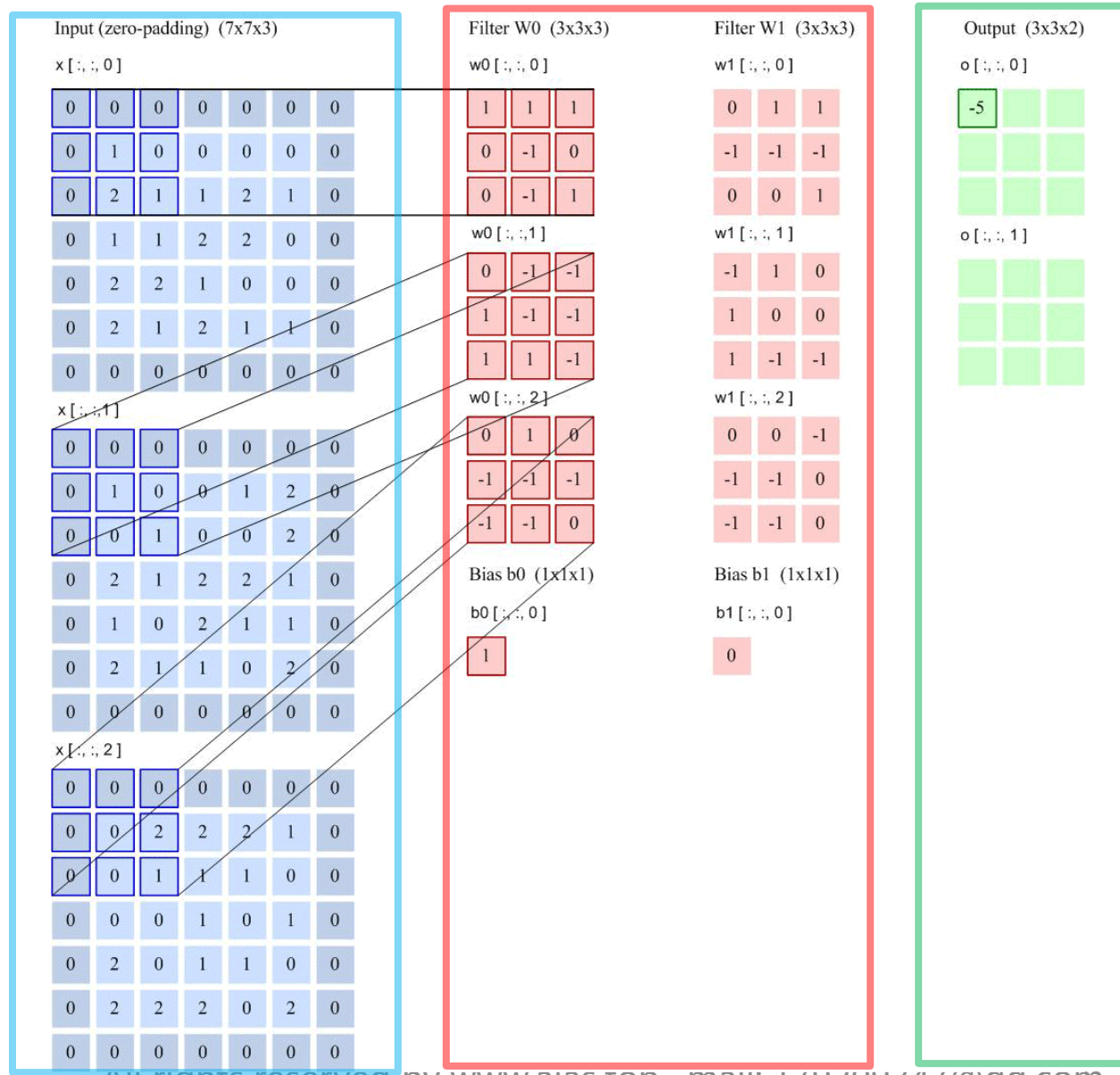
计算过程：

$$\begin{aligned}
 &0 \times 0 + 1 \times 1 + (-1) \times 0 + 2 \times 0 + 0 \times 0 + 1 \times 1 + 0 \times 1 + 1 \times 1 + 1 \times 0 \\
 &+ 1 \times (-1) + 0 \times 1 + 0 \times 0 + 2 \times (-1) + 0 \times 0 + 1 \times 1 + 0 \times (-1) + 2 \times 1 + 1 \times 1 \\
 &+ 1 \times 0 + 2 \times 1 + 0 \times (-1) + 1 \times (-1) + 0 \times 0 + 3 \times 1 + 0 \times 1 + 2 \times 0 + 0 \times 1 \\
 &+ 1 \\
 &= 9
 \end{aligned}$$

卷积层 - 多通道卷积 - 两个卷积核



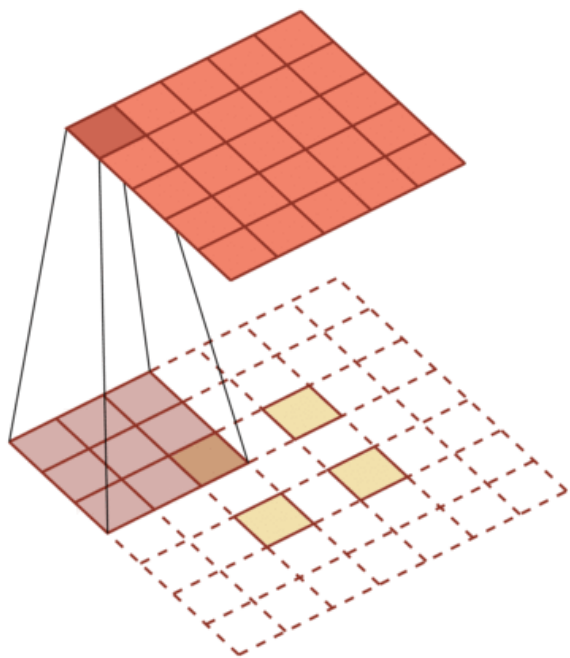
卷积层 - 多通道卷积 - 两个卷积核动态演示



其它卷积种类

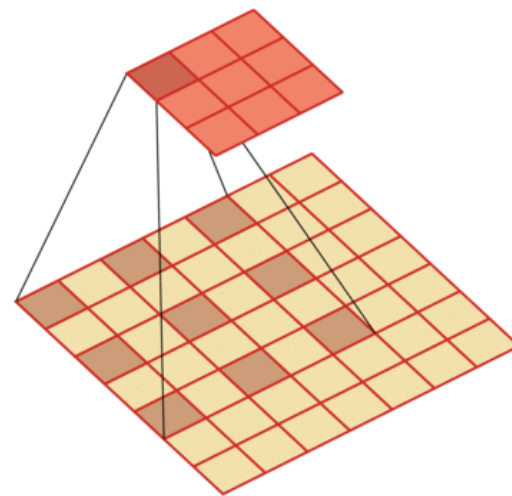
转置卷积

转置卷积 (transpose convolution), 也被称为反卷积 (deconvolution), 是一种在深度学习中常用的操作, 用于上采样或反卷积操作。转置卷积的本质是**低维特征映射到高维特征**。



空洞卷积

空洞卷积 (Dilated Convolution) 也称为扩张卷积或膨胀卷积。与传统的卷积操作不同, 空洞卷积在卷积核中引入了一个或多个空洞 (或称为孔), 使得卷积核可以跨越更大的空间范围进行卷积运算, 而不增加参数数量, 从而增加输出单元的感受野。





Thank

You