

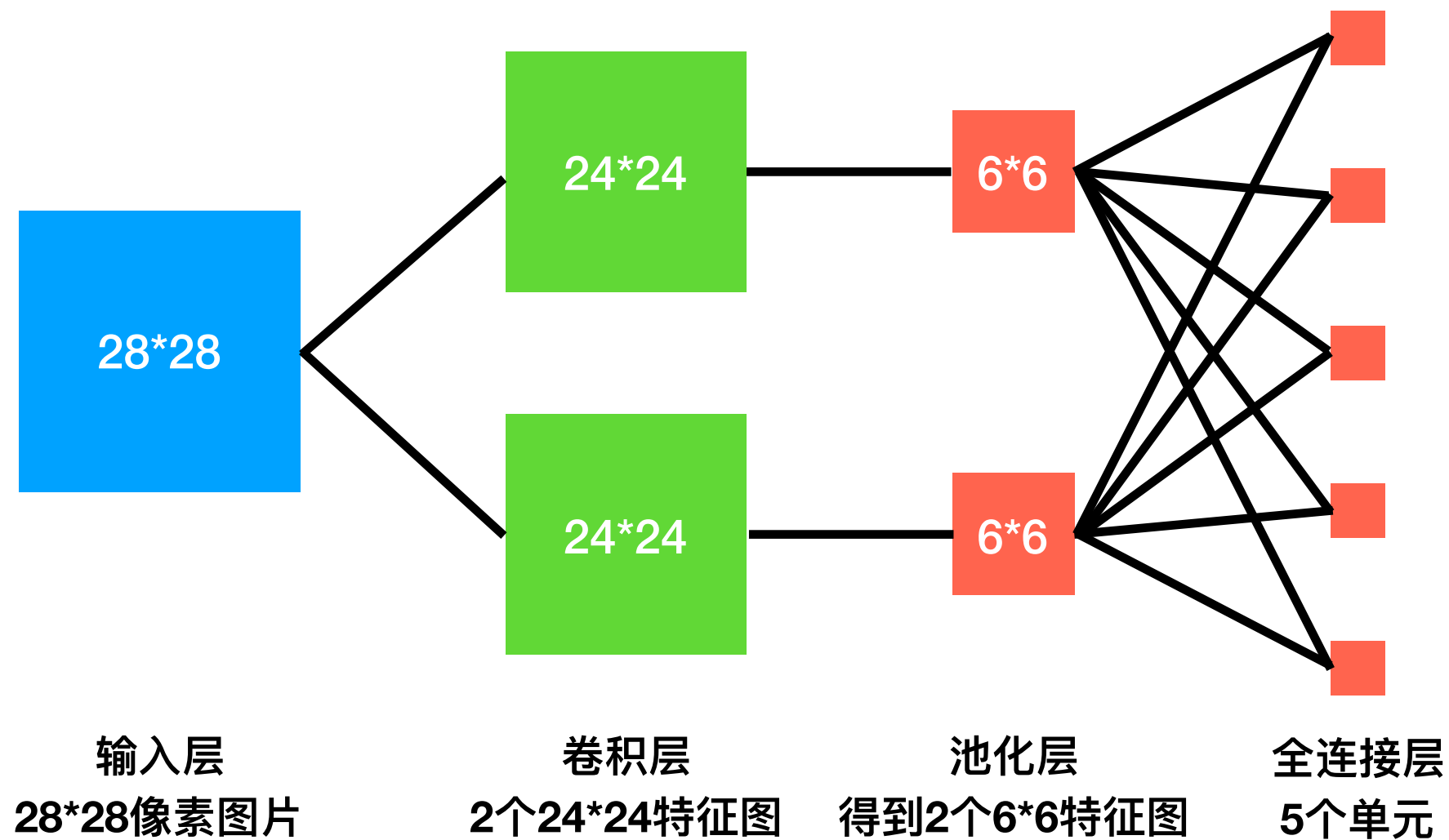
深度学习

卷积神经网络 (2)

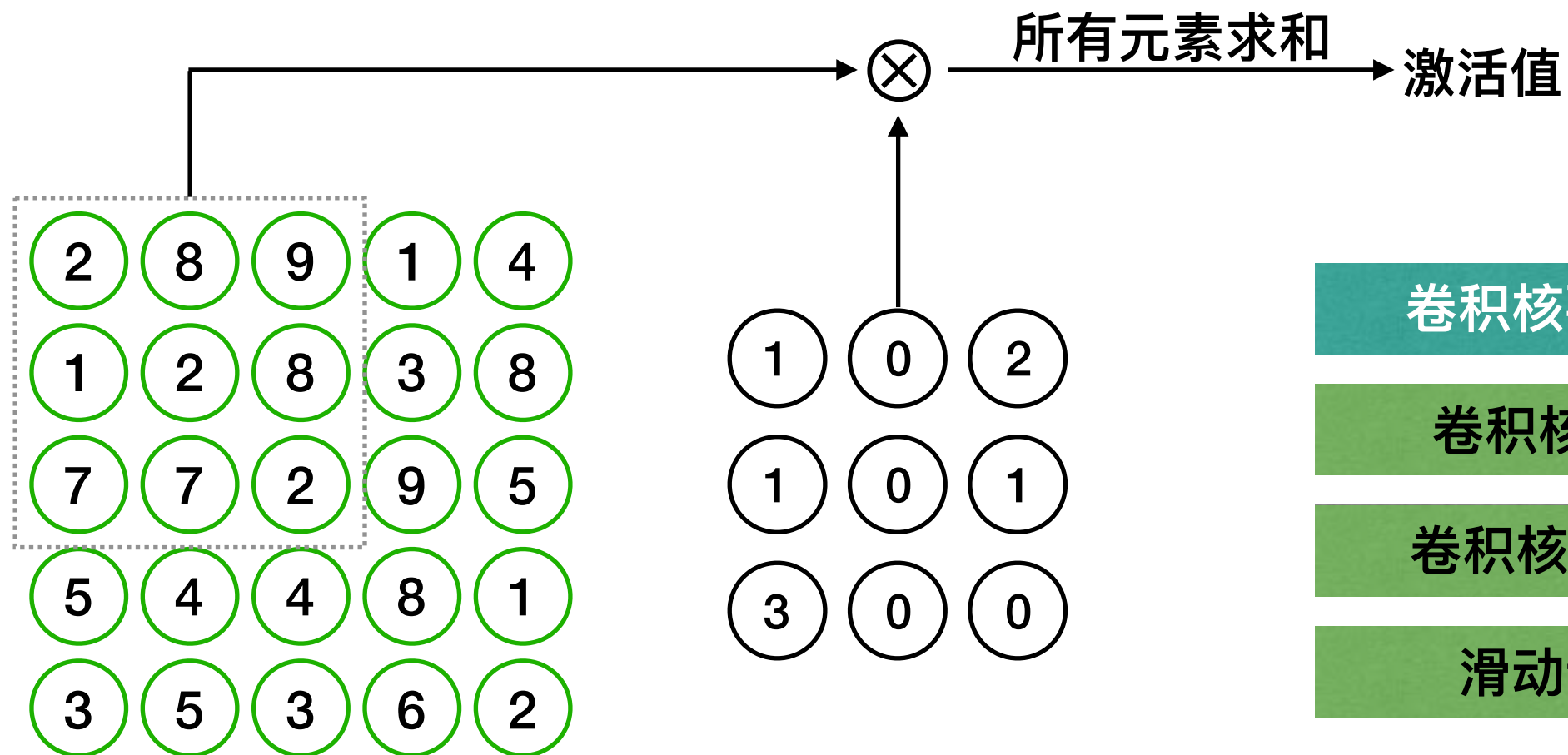
上节回顾

- 图像、语音等数据可以使用微观特征描述。
- 卷积核是特征筛选器，可以将图像转化为某个特征的分布图。
- 多个卷积核可以提取数据的完整特征。同时会得到多个特征图。
- 特征图规模较大时，使用下采样技术——池化来降维。

卷积与池化



卷积核



一张图片的像素值

一个卷积核的参数

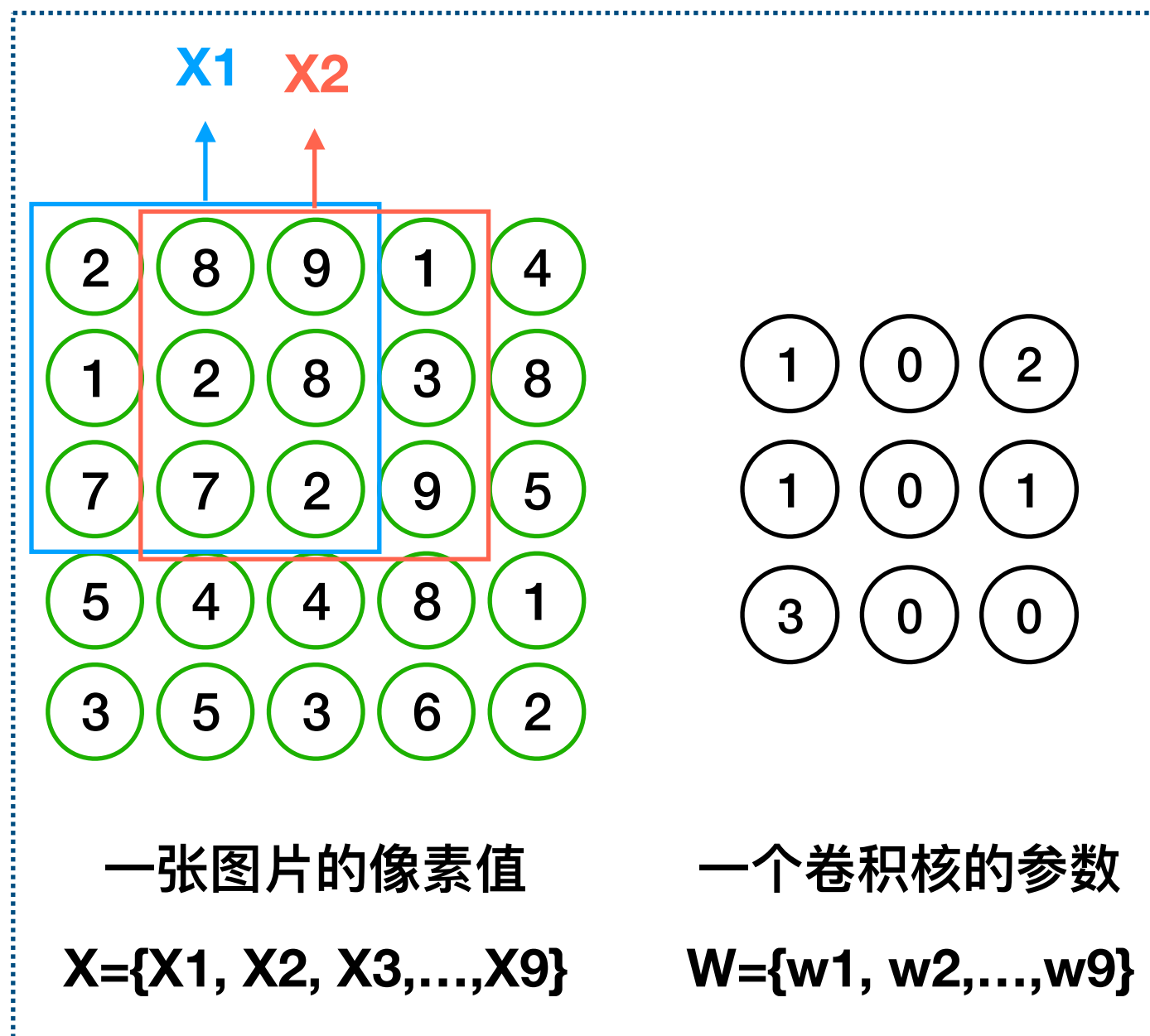
卷积核要素:

卷积核参数

卷积核shape

滑动步长

卷积核



此处我们规定 a_n 为第 n 个激活值

$$a_n = \text{sum}(X_n \otimes W)$$

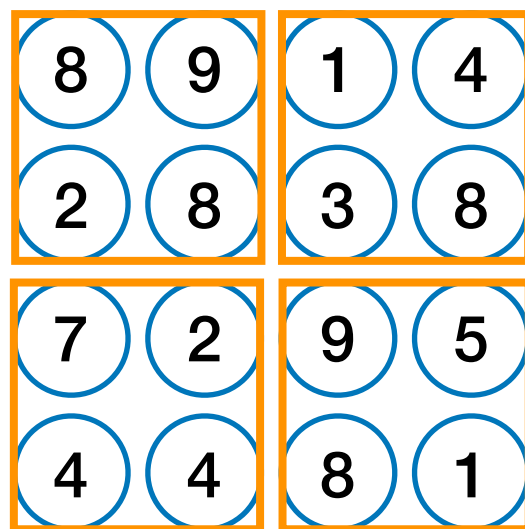
所以 a 可以看做是卷积后得到的特征图。

通常是非线性激活函数

带激活函数的卷积

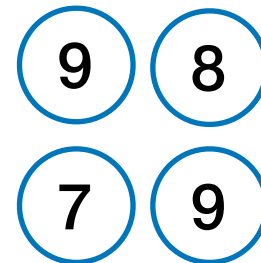
$$a_n = a_n = \text{activation_function}(\text{sum}(X_n \otimes W) + b_n)$$

池化核



特征图

最大池化



池化后的特征图

$$a_n = \max(z_n)$$

平均池化



池化后的特征图

$$a_n = avg(z_n)$$

卷积核要素：

池化核shape

滑动步长

池化方式

带激活函数的池化

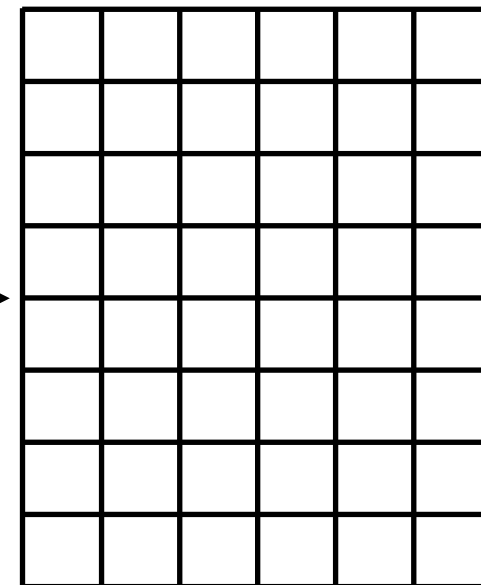
$$a_n = \text{activation_function}(\text{pooling}(z_n))$$

通常无激活函数或为线性激活函数

图片通道



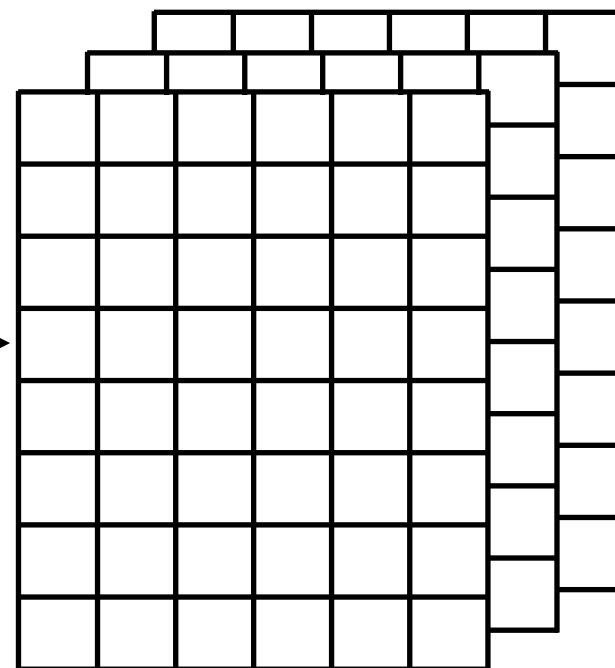
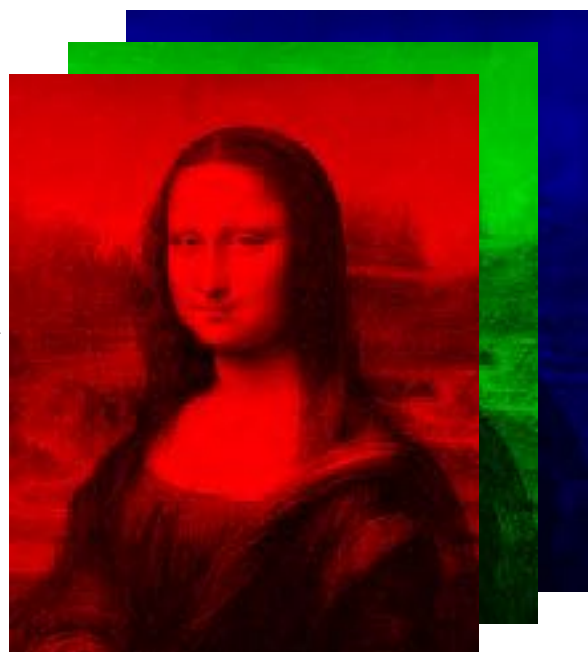
灰度图像



颜色值矩阵



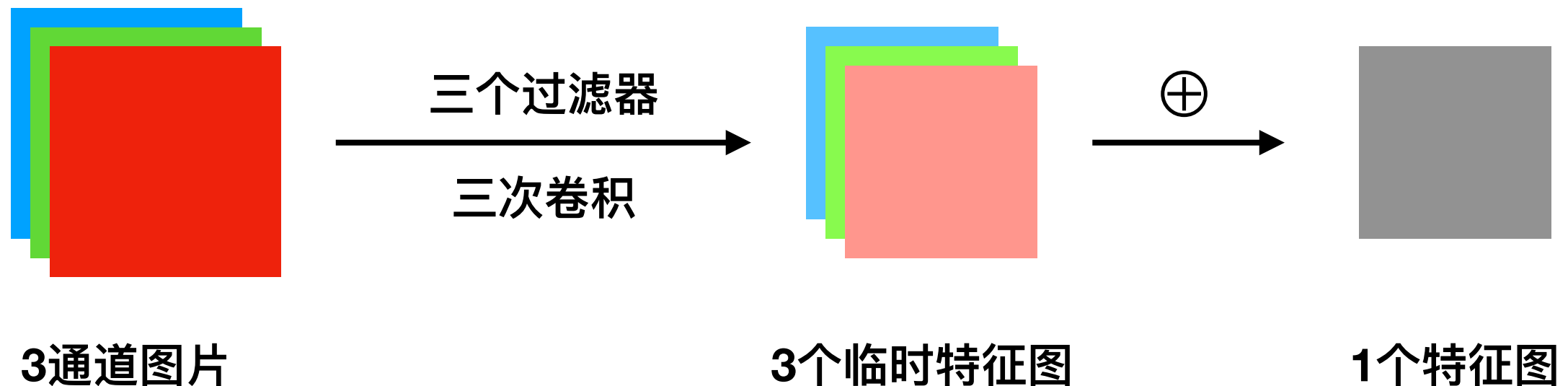
RGB图像



颜色值矩阵

多通道卷积

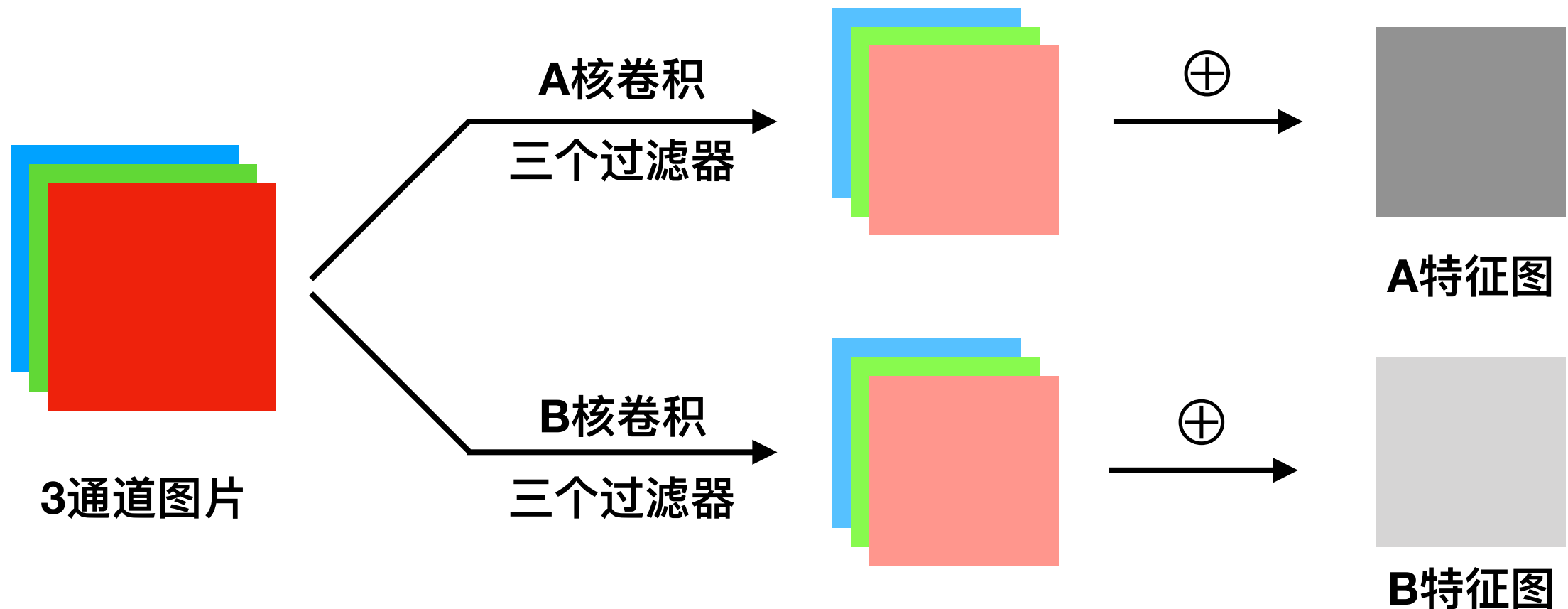
多通道“单核”卷积过程



虽然是一次卷积，但其用到了三个shape相同的过滤器（过滤器也是卷积核，但为了区分，此处我们称之为过滤器），最终可以得到3个临时特征图。最终，将三个临时特征图对应元素相加，得到特征图。卷积的激活函数对特征图使用。

多通道卷积

多通道“多核”卷积过程



我们将图片的颜色通道、同一层中生成的多个特征图的数量统称为通道。

思考？多通道单核卷积后的卷积核参数数量应该怎么算？

小练习

假如有一张大小为12px*12px大小的png图片，用2个3*3大小的卷积核处理。

问题1：得到了几个特征图？

问题2：每个特征图的大小分别是多少？

问题3：单核卷积对应多少个过滤器？

问题4：所有卷积核一共有多少个参数？

卷积核的大小与卷积运算量

图片大小	卷积核大小	特征图大小	计算量
28px*28px	3*3	26*26	6084
28px*28px	5*5	24*24	14400
28px*28px	14*14	15*15	44100
28px*28px	20*20	9*9	32400

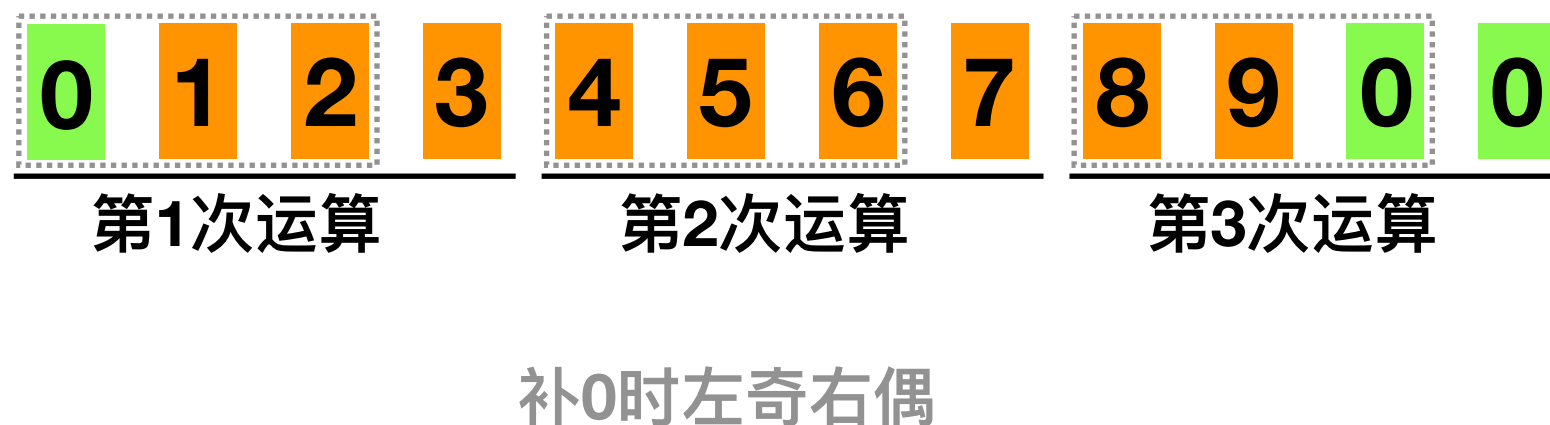
边界处理

padding="VALID"



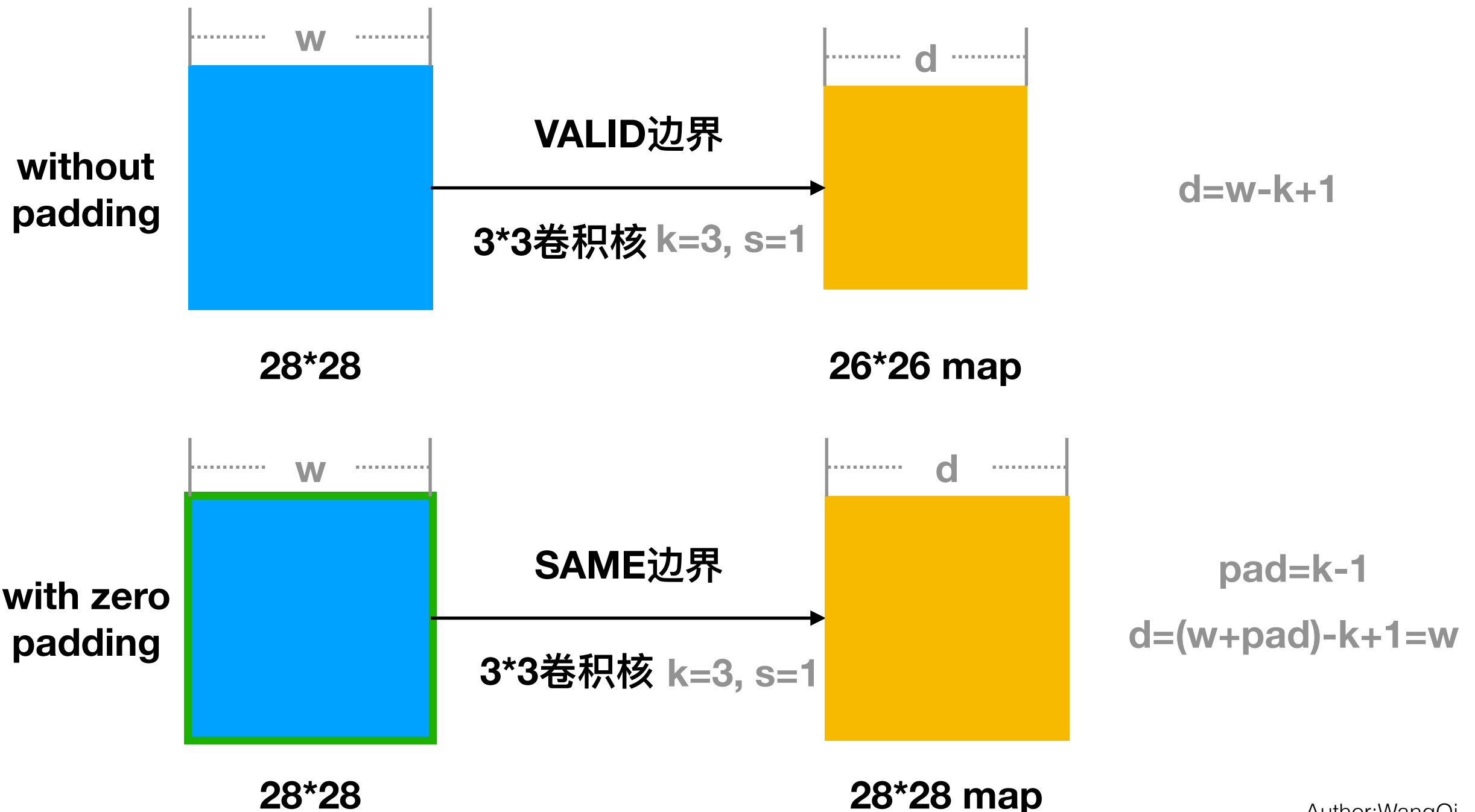
width: 9
kernel: 3
strides: 4

padding="VALID"

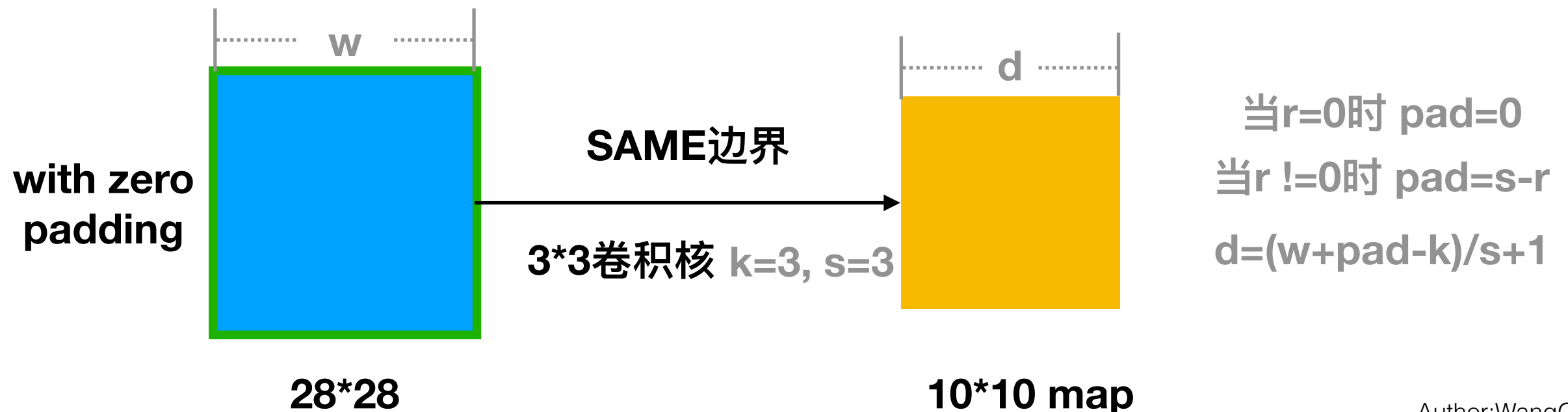
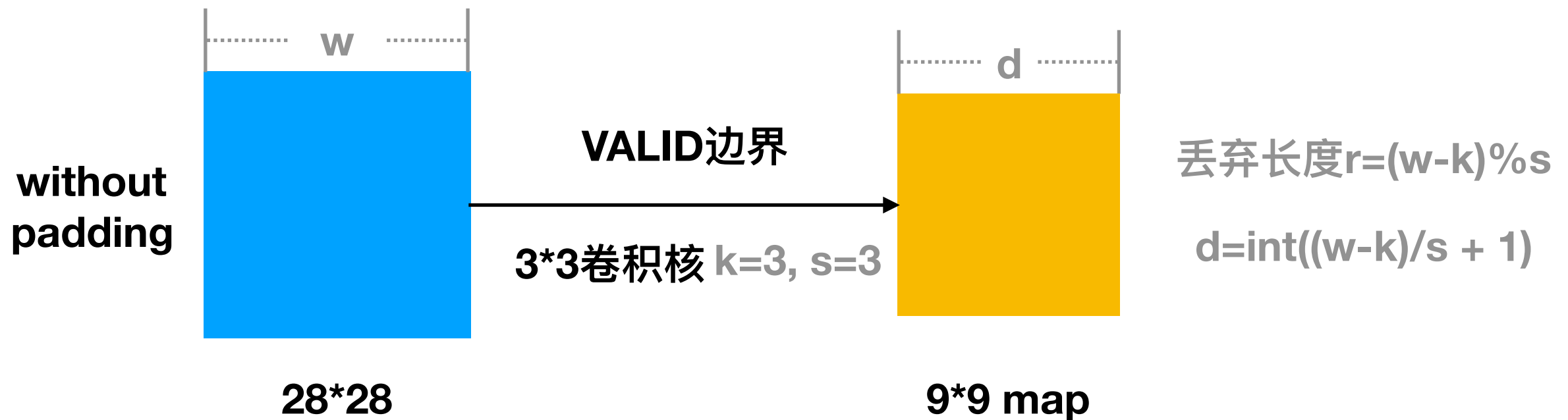


width: 9
kernel: 3
strides: 4

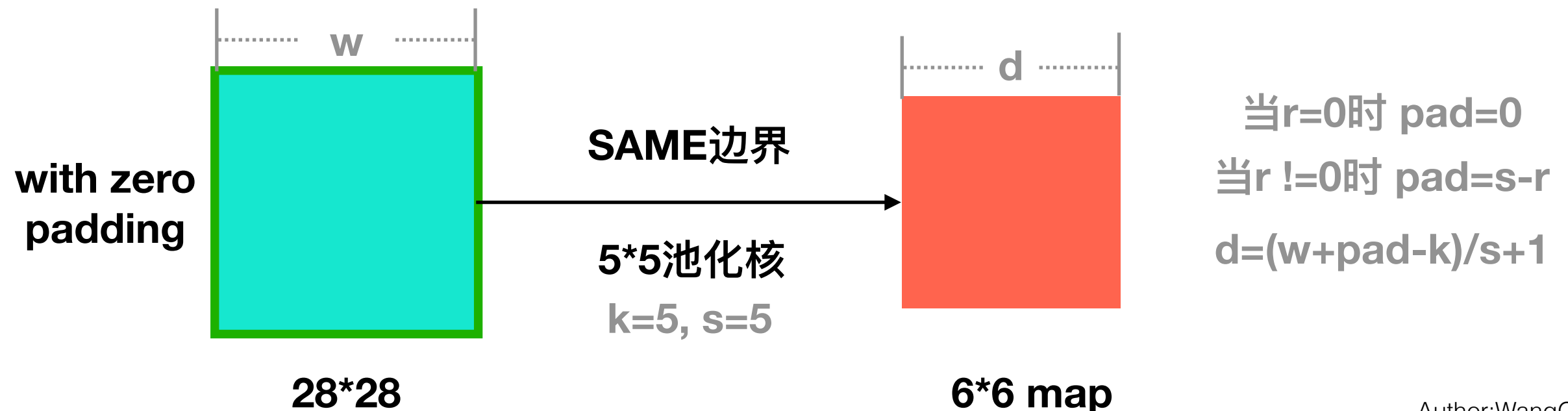
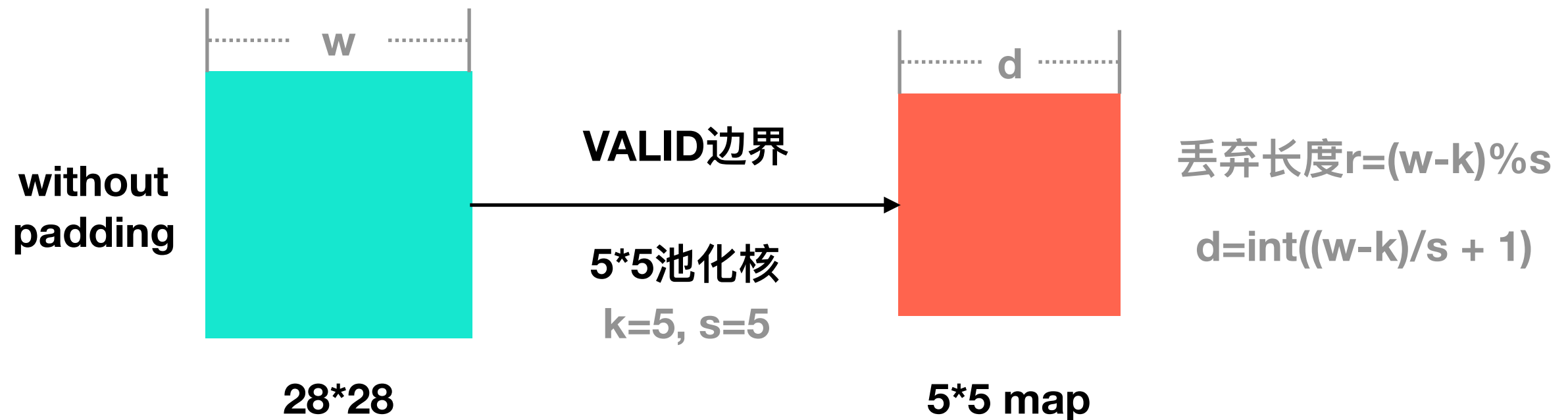
卷积边界处理



卷积边界处理



池化边界处理



小练习

技巧：任何时候都是先算冗长（VALID中的丢弃长度）。

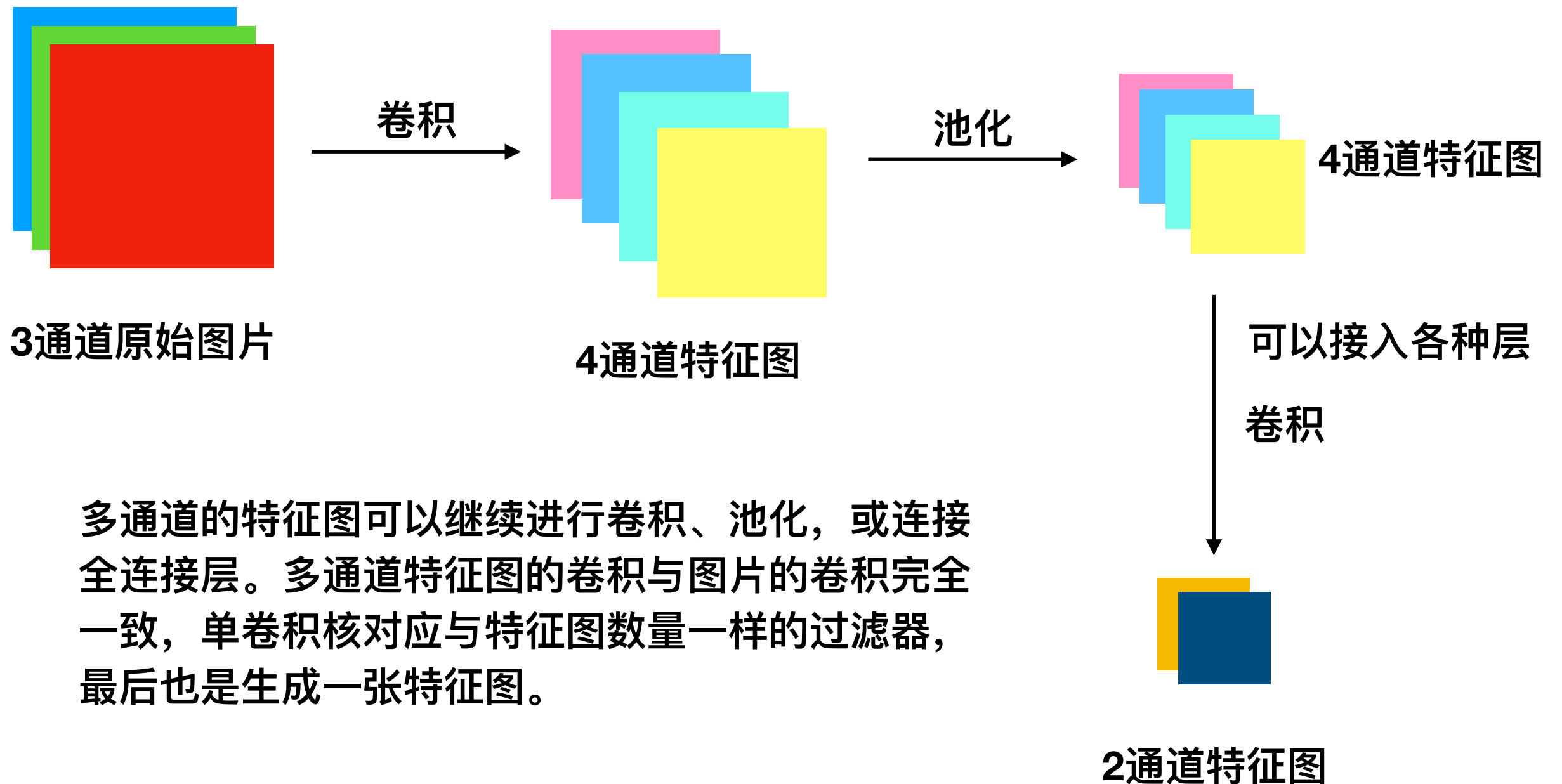
已知单通道的某图片大小为13*15，现使用6*4大小的卷积核以5*7的步长进行卷积，请问：

问题1：如果卷积使用VALID的边界处理方式，那么卷积得到的特征图大小是多少？丢弃的长度多少？

问题2：如果卷积使用SAME的边界处理方式，那么卷积得到的特征图大小是多少？补0的长度是多少？

	边长	卷积核边长	步长	冗长	VALID下的map	补0	SAME下的map
X轴	13	6	5	2	2	3	3
Y轴	15	4	7	4	2	3	3

多通道特征图



多通道的特征图可以继续进行卷积、池化，或连接全连接层。多通道特征图的卷积与图片的卷积完全一致，单卷积核对应与特征图数量一样的过滤器，最后也是生成一张特征图。

卷积层与池化层的灵活应用

1. 卷积层之后通常会接入池化层，但也可以接入新的卷积层或全连接层。
2. 池化层通常不会放在第一层，某些特殊场景可能会这样做。
3. 通常池化层之后不再接池化层，因为这样是没有意义的。
4. 卷积层之后接池化层是组合中黄金搭档。多组卷积池化连接在一起往往可以得到很好性能。

多次的卷积与池化组合，极大地降低了神经网络的参数规模，同时有效的提取了输入数据特征。

早期CNN的成功案例LeNet-5

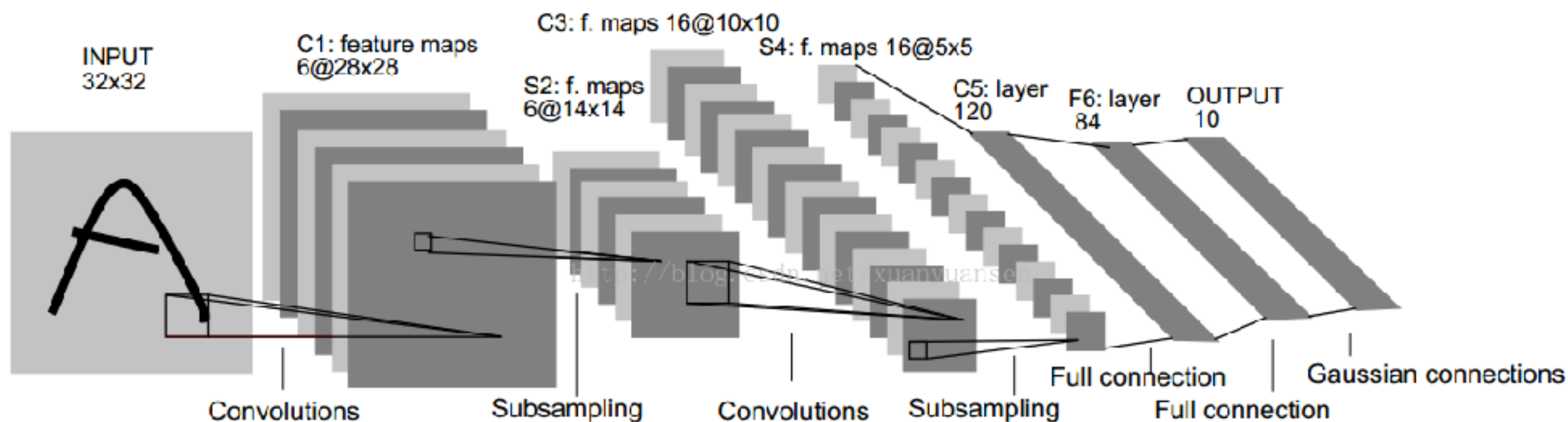


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

LeNet曾是美国各大银行等大机构最常用的识别手写体的算法。

平移不变性



训练时使用的图片



预测时使用的图片

卷积：卷积得到的map是微观特征的激活值，所以特征位置改变时，对应位置的最激活值位置也改变了，但特征之间的相对位置不会改变。而更深层的网络学习到的是微观特征的组合方式，只要特征的组合方式没有变化，**CNN**就能识别。

池化：池化操作会在一定范围内检测特征。以最大池化为例，特征位置改变了，但在池化核的范围内，最大值没有变。

CNN中的参数

- 卷积层的参数包括卷积核参数与激活函数的参数中的偏置值。偏置值数量等于生成的特征图的激活值数量。
- 池化层在没有激活函数的情况下是没有激活值的，在有激活函数的情况下，可以有偏置值。偏置值的数量等于生成的特征图的激活值数量。
- 全连接层的参数。

CNN的训练

- CNN的前向传播与普通ANN的类似。我们可以求得每一层的残差以及每一个参数的偏导数，通过梯度下降等优化算法更新参数。CNN的反向传播算法中比较复杂的部分是需要考虑卷积层与池化层的不同连接关系时不同的计算方式，在此不予以说明。

思考：存在 $1*1$ 卷积核吗？如果存在，卷积的结果是什么？有什么意义？

$1*1$ 卷积核是存在的，如果1个特征图通过 $1*1$ 卷积得到了一个新的特征图，则两个特征图的关系是线性的（卷积激活函数是线性的情况下），这时候 $1*1$ 卷积是没有意义的。但如果有多通道的特征图通过 $1*1$ 卷积可以实现通道维度的改变，同时也实现了多个通道信息的汇聚。

小结

- 卷积核与特征图或图片的运算。
- 多核卷积可以得到多个特征图，其数量与核数量相同。
- 多通道多核卷积中每一个核（多个过滤器）均进行与通道维度数相同数量的卷积。
- 卷积核通常设置比较小，运算量也比较小，容易提取到微观特征。
- 通常卷积与池化的边界处理包括VALID与SAME两种。VALID可能丢弃一部分信息，SAME可能会用0填充边界。

THANKS