CNN基础与LeNet

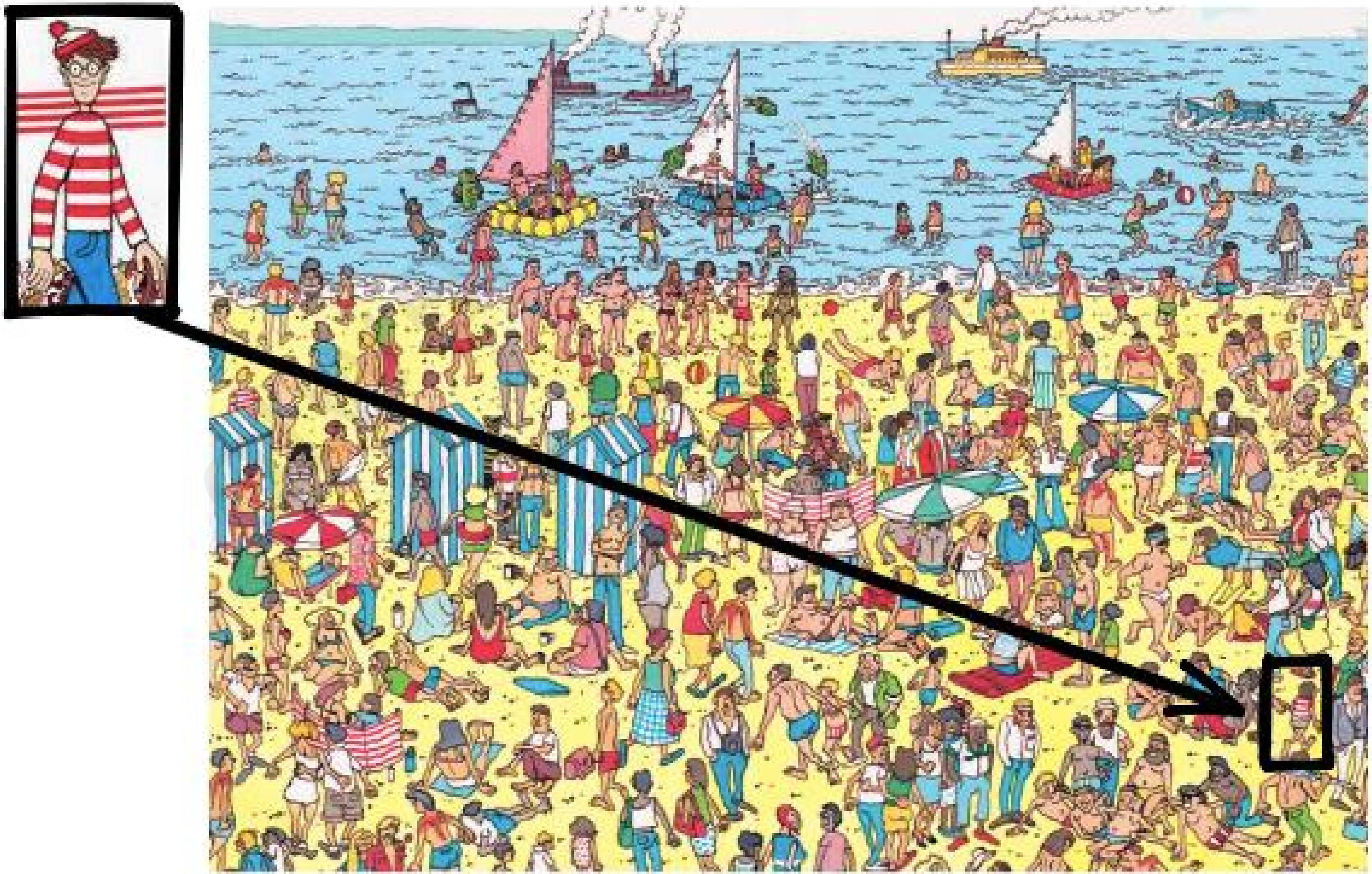
1. MLP的局限性是什么？如何从MLP联系到CNN？

MLP运用：

	特征	
样本1	M1	
样本2	M2	

局限性：

2. 为什么CNN能够解决上述MLP的局限性？ -----定性分析



3. 为什么CNN能够解决上述MLP的局限性？-----量化分析

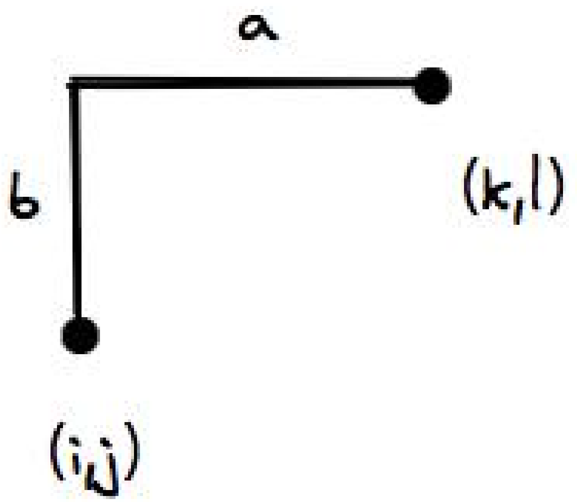
位置（i,j）处隐藏表示中的像素

偏置

权重

位置（k,l）处输入图像中的像素

$$
\begin{aligned}
[\mathbf{H}]_{i,j} &= [\mathbf{U}]_{i,j} + \sum_{k}\sum_{l}[\mathsf{W}]_{i,j,k,l}[\mathbf{X}]_{k,l} \\
&= [\mathbf{U}]_{i,j} + \sum_{a}\sum_{b}[\mathsf{V}]_{i,j,a,b}[\mathbf{X}]_{i+a,j+b}
\end{aligned}
$$



3. 为什么CNN能够解决上述MLP的局限性？ -----量化分析

位置（i,j）处隐藏表示中的像素

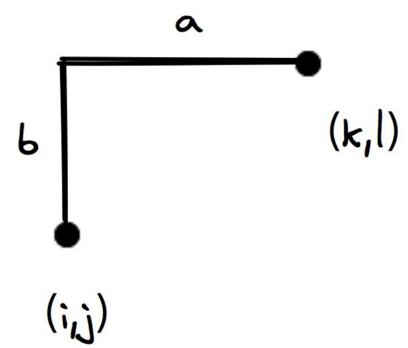
偏置

权重

位置（k,l）处输入图像中的像素

$$[\mathbf{H}]_{i,j} = [\mathbf{U}]_{i,j} + \sum_k \sum_l [\mathsf{W}]_{i,j,k,l} [\mathbf{X}]_{k,l}$$

$$= [\mathbf{U}]_{i,j} + \sum_a \sum_b [\mathsf{V}]_{i,j,a,b} [\mathbf{X}]_{i+a,j+b}$$



常数

位置不变性，则 $[\mathsf{V}]_{i,j,a,b} = [\mathbf{V}]_{a,b}$

so:

$$\boxed{[\mathbf{H}]_{i,j} = u + \sum_a \sum_b [\mathbf{V}]_{a,b} [\mathbf{X}]_{i+a,j+b}}$$

卷积核

卷积操作

局部性

$$\boxed{[\mathbf{H}]_{i,j} = u + \sum_{a=-\Delta}^{\Delta} \sum_{b=-\Delta}^{\Delta} [\mathbf{V}]_{a,b} [\mathbf{X}]_{i+a,j+b}}$$

卷积层

回顾卷积公式

两个连续函数的卷积

翻转，移位

$$(f*g)(\mathbf{x}) = \int f(\mathbf{z})g(\mathbf{x}-\mathbf{z})d\mathbf{z}$$

离散对象

$$(f*g)(i) = \sum_a f(a)g(i-a)$$

二维

$$(f*g)(i,j) = \sum_a \sum_b f(a,b)g(i-a,j-b)$$

理解

$$\boxed{[\mathbf{H}]_{i,j} = u + \sum_a \sum_b \underbrace{[\mathbf{V}]_{a,b}}_{\text{卷积核}}[\mathbf{X}]_{i+a,j+b}}$$

卷积操作

4. 二维张量？对于三通道图片怎么办？

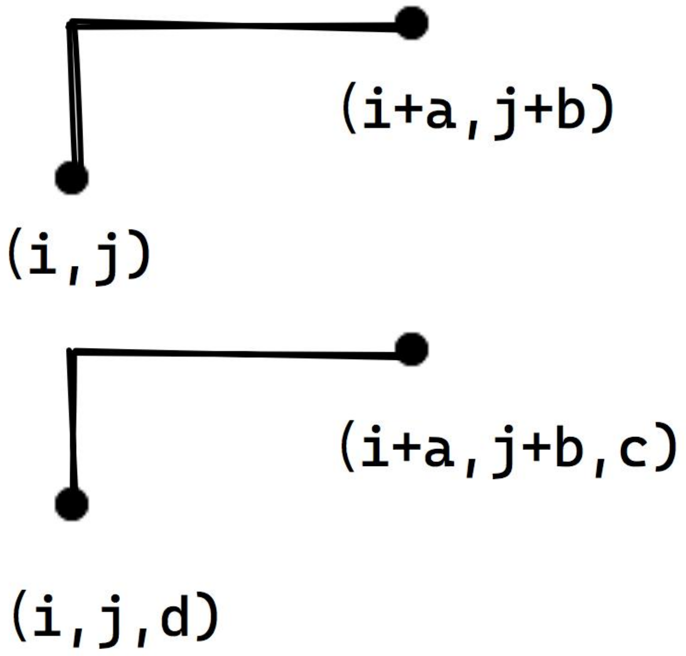
$$[\mathbf{H}]_{i,j} = u + \sum_{a=-\Delta}^{\Delta} \sum_{b=-\Delta}^{\Delta} [\mathbf{V}]_{a,b} [\mathbf{X}]_{i+a,j+b}$$

卷积层

多个通道的卷积层

$$[\mathsf{H}]_{i,j,d} = \sum_{a=-\Delta}^{\Delta} \sum_{b=-\Delta}^{\Delta} \sum_{c} [\mathsf{V}]_{a,b,c,d} [\mathsf{X}]_{i+a,j+b,c}$$

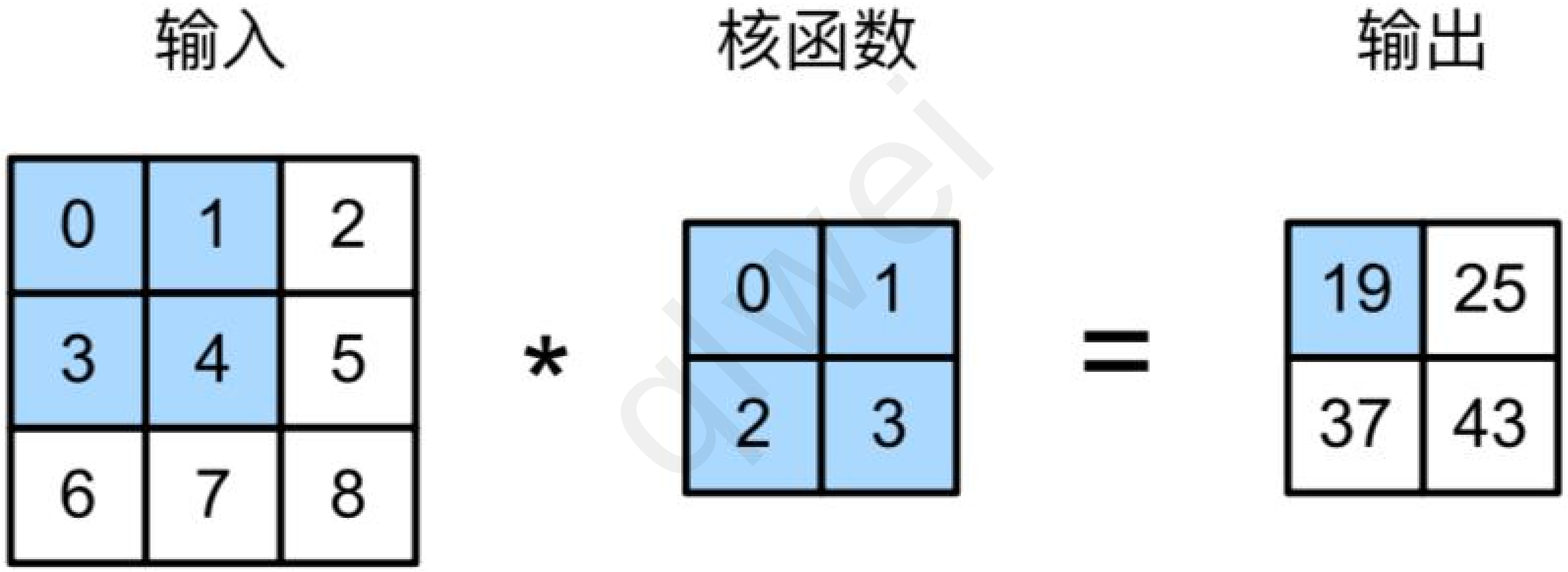
隐藏层输出维度



以上，解析了“卷积层”的原理

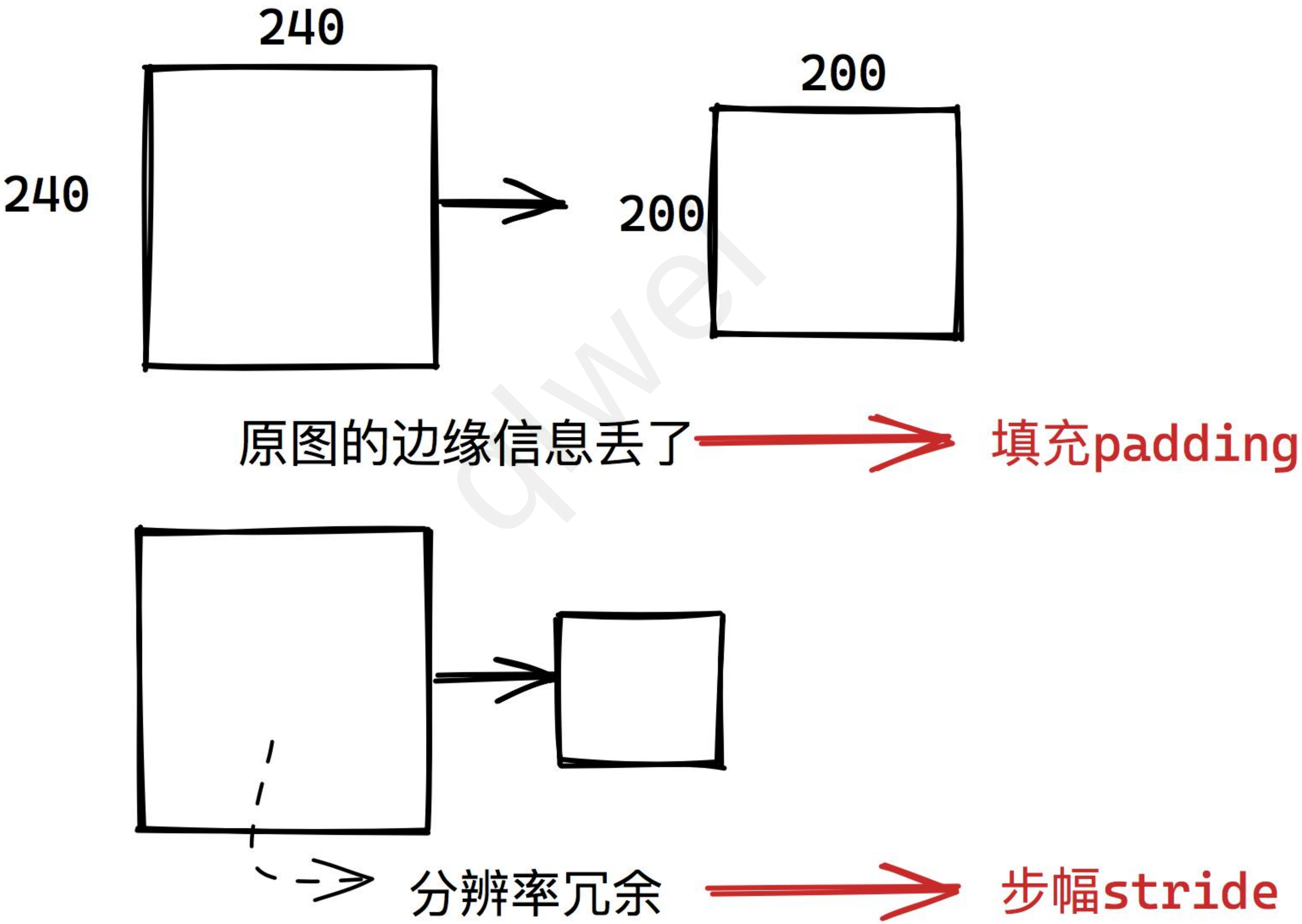
简单实践

5. 图像中卷积应用

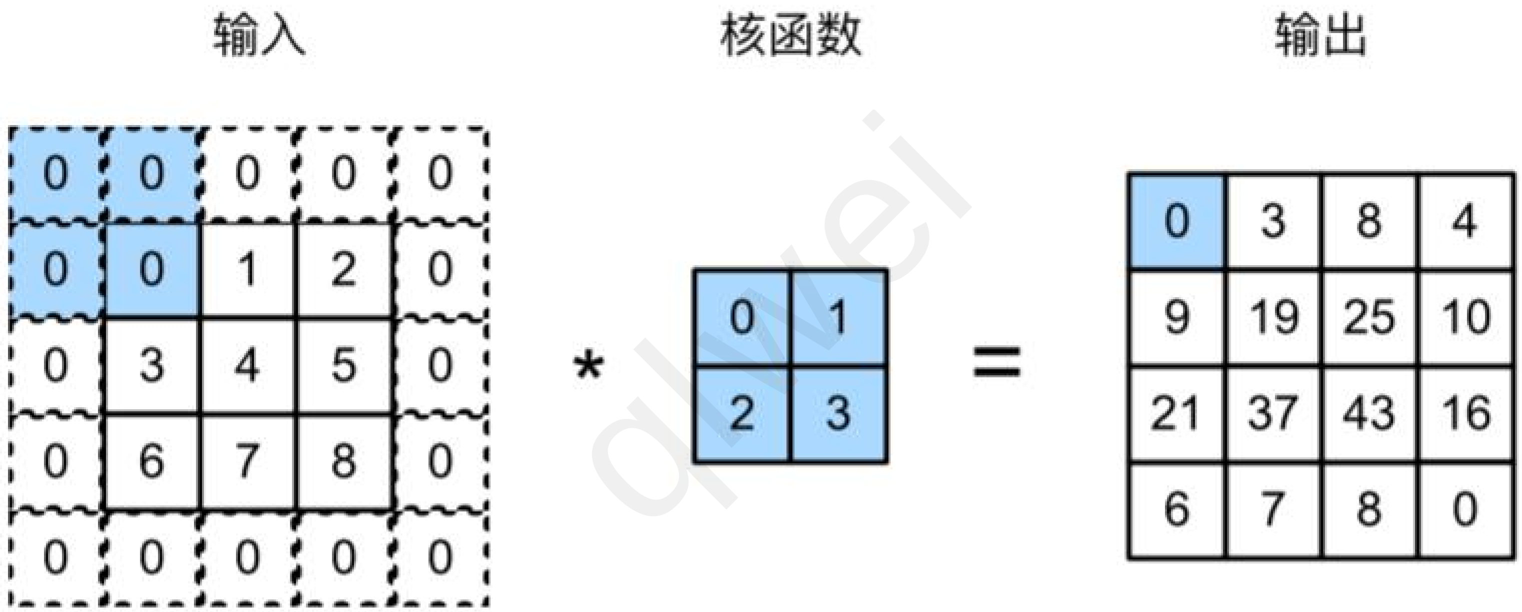


变小了。。。

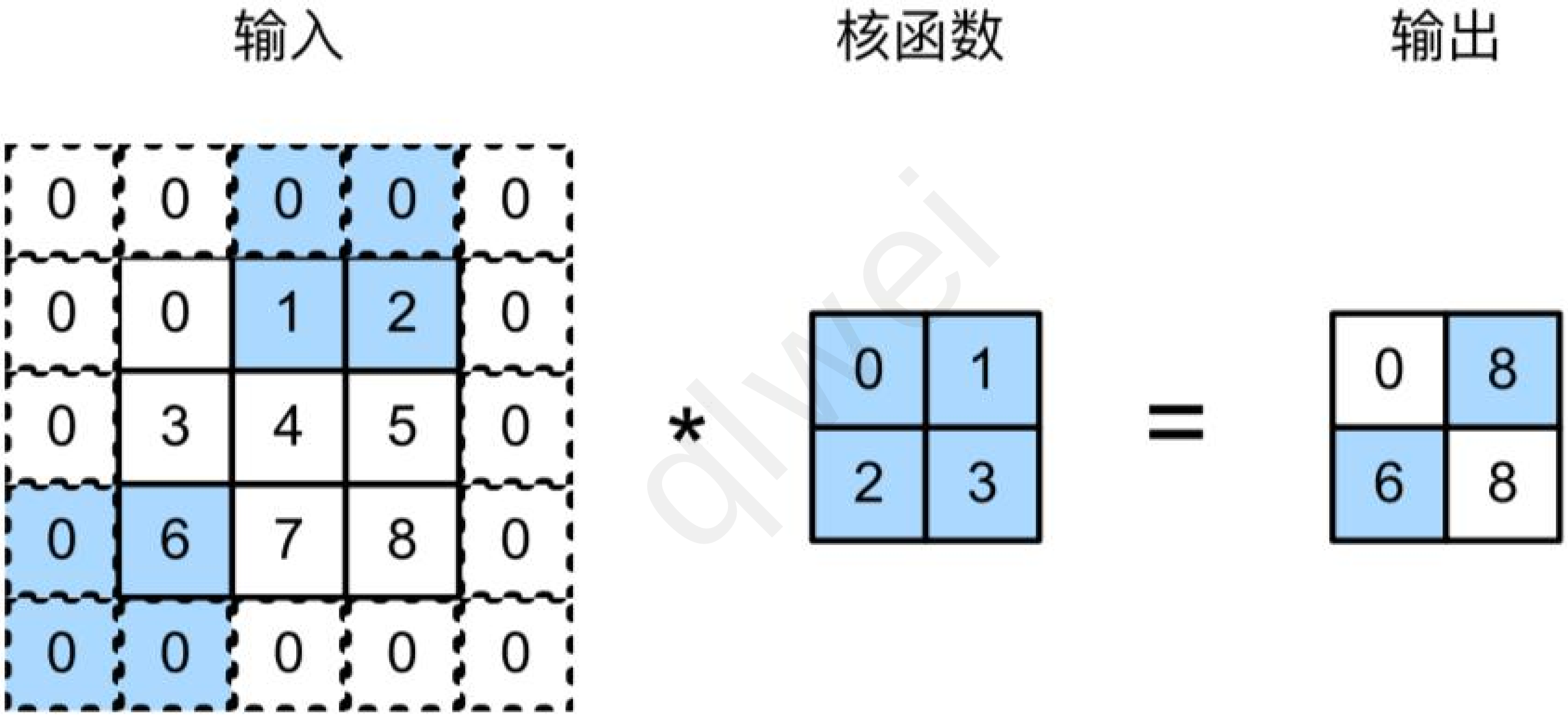
6. 还有什么因素会影响输出的大小呢?



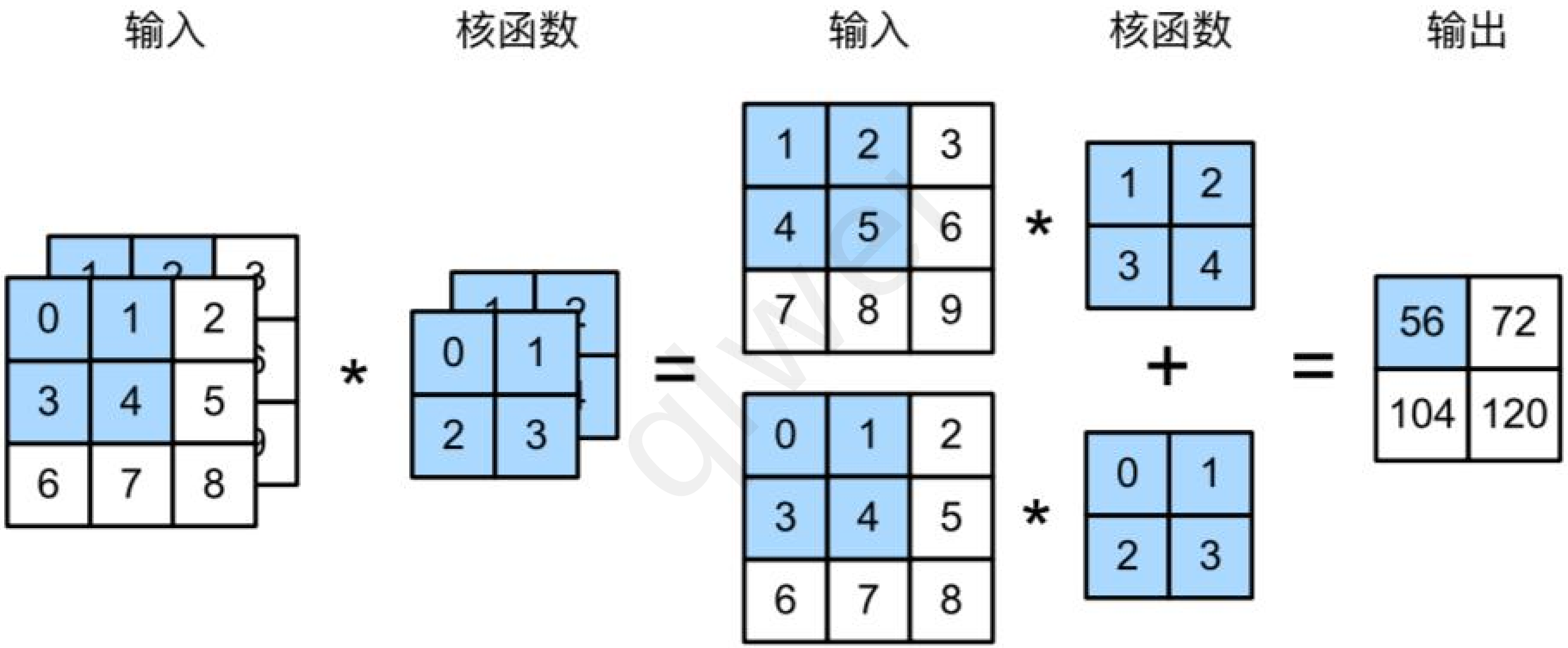
padding


输入
核函数
输出
0 0 0 0 0
0 0 1 2 0
0 3 4 5 0
0 6 7 8 0
0 0 0 0 0
*
0 1
2 3
=
0 3 8 4
9 19 25 10
21 37 43 16
6 7 8 0

stride



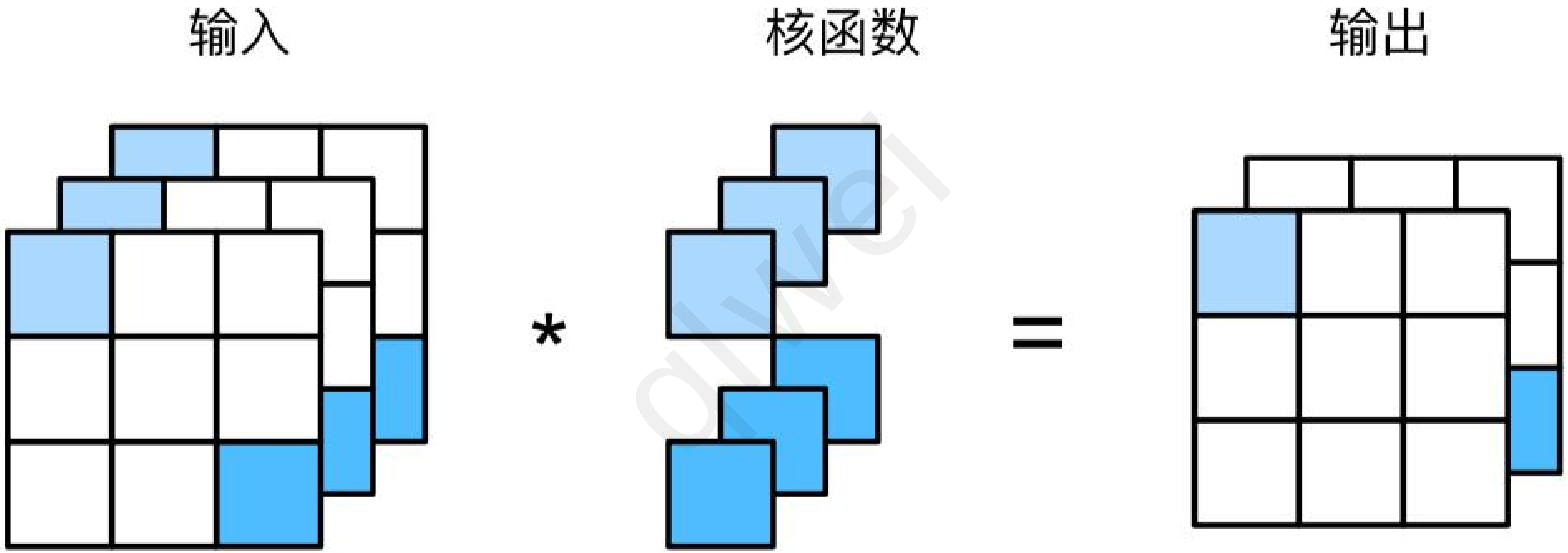
7. 那么多通道的图像卷积运算？



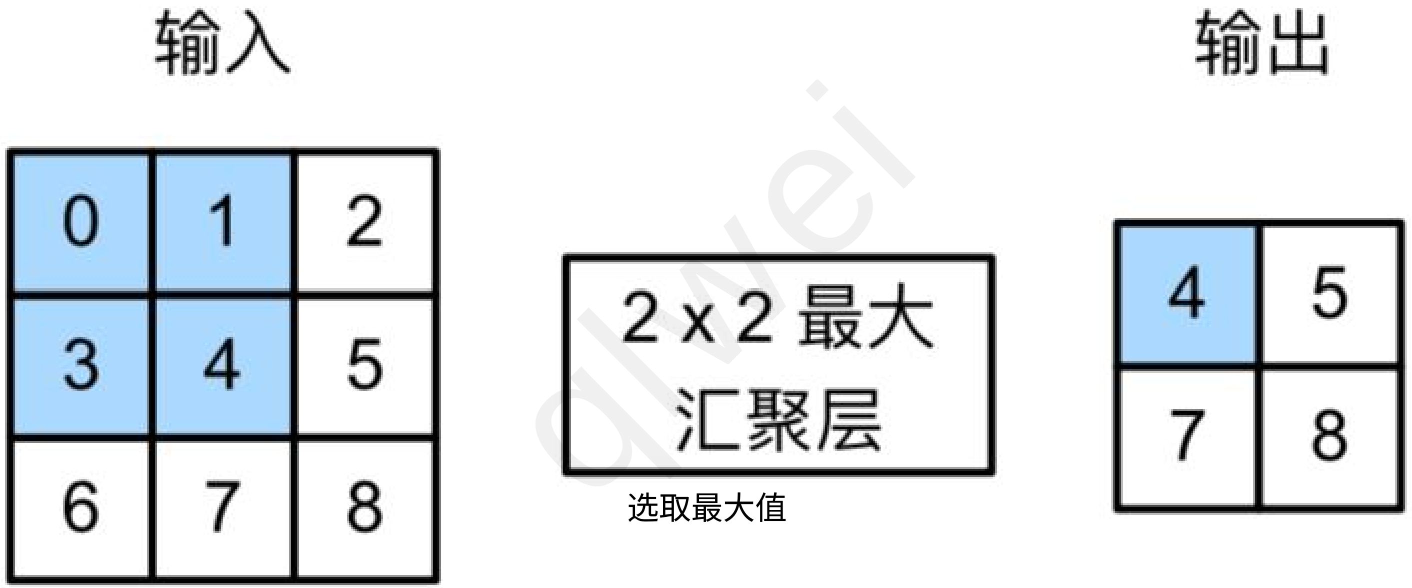
8. 有没有1*1大小的卷积核?

输入输出相同宽高，通道数改变
（同一位置的线性组合）

调整通道数目，降低模型的复杂性



9. 多层卷积后如何表达全局抽象意义？----pooling



组件们over～～～

迎来第一个CNN～～～

LeNet

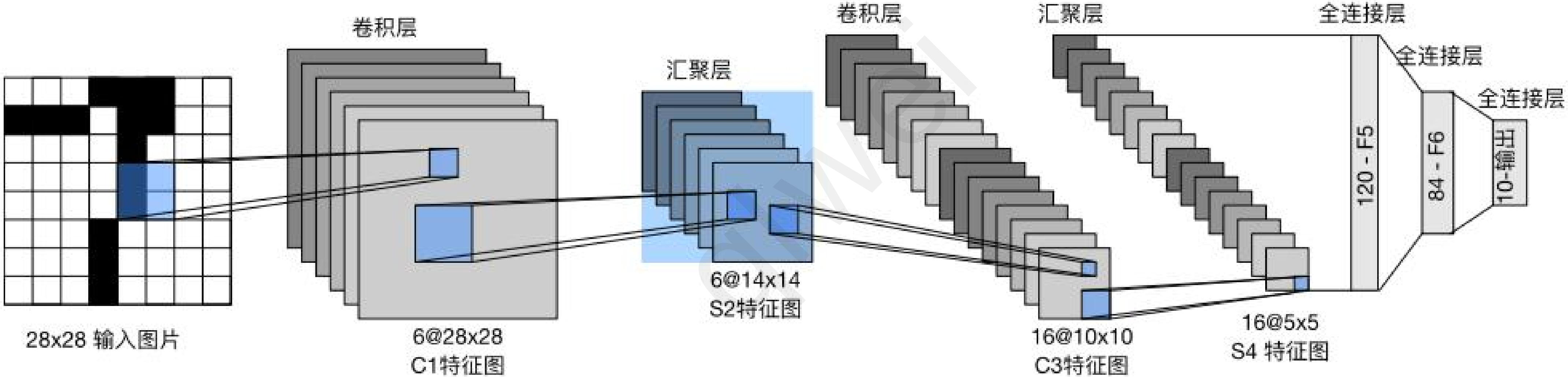


图6.6.1: LeNet中的数据流。输入是手写数字，输出为10种可能结果的概率。

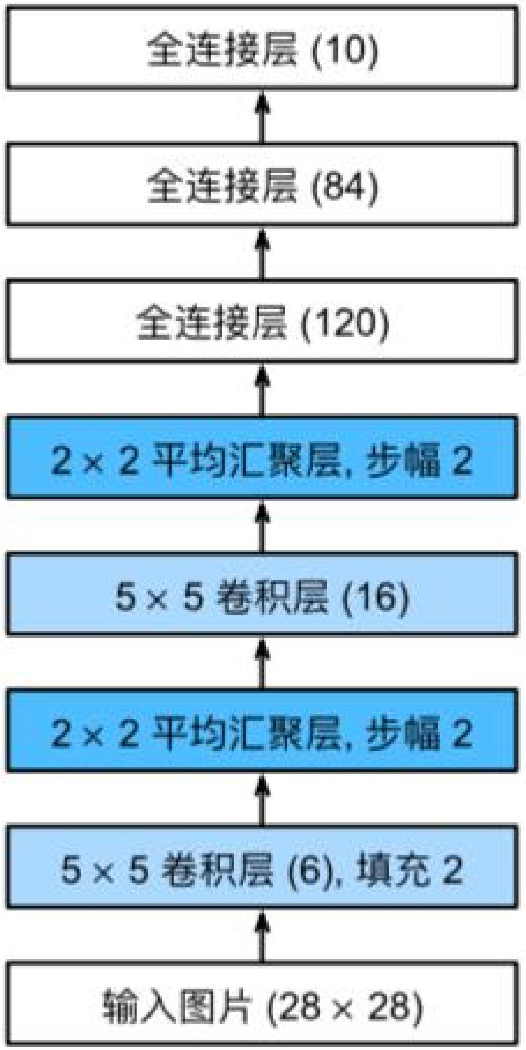


图6.6.2: LeNet 的简化版。

层与块?

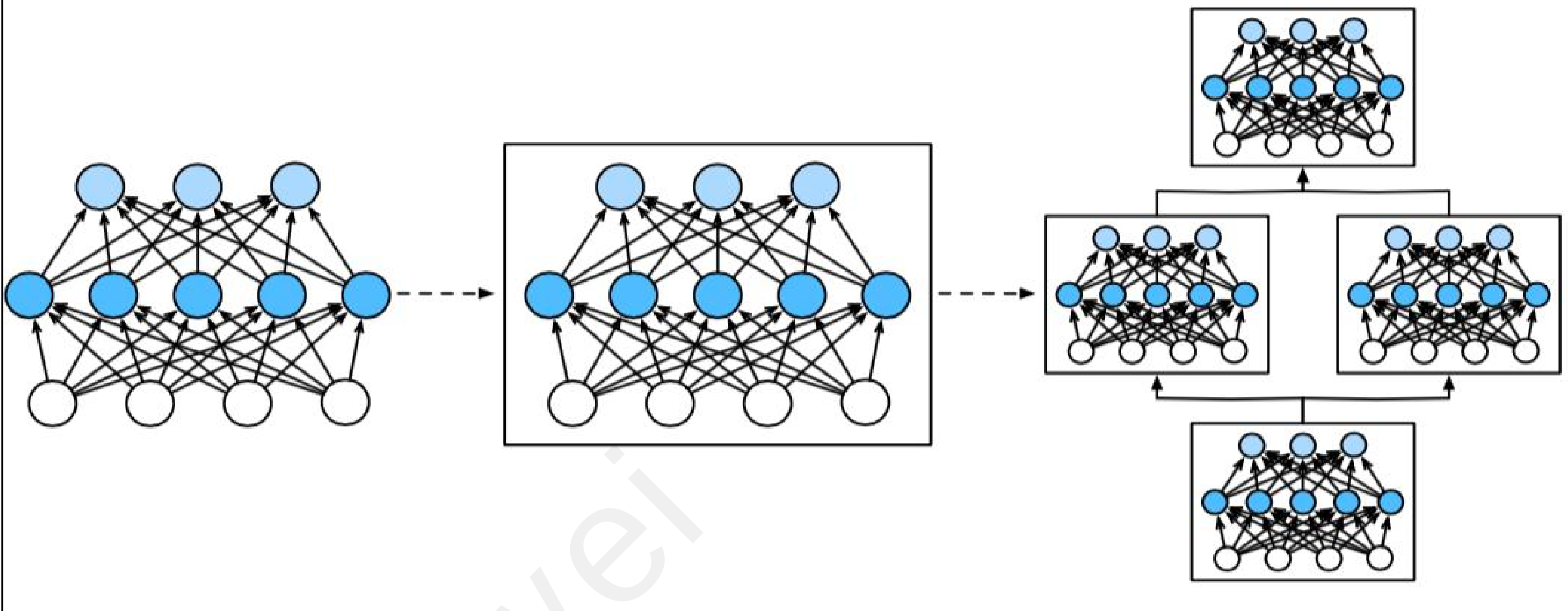


图5.1.1: 多个层被组合成块，形成更大的模型