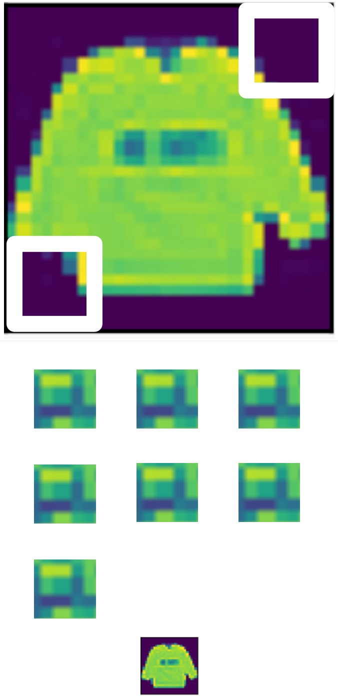
填充

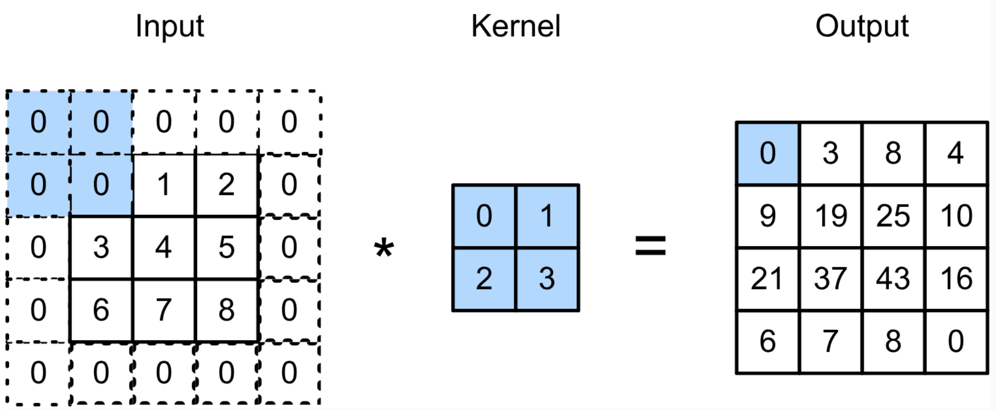
- 给定（32 x 32）输入图像
- 应用 5 x 5 大小的 卷积核
  - 第1层得到输出大小 28 x 28
  - 第7层得到输出大小4 x 4
- 更大的卷积核可以更快地减小输出大小
  - 形状从 $n_h \times n_w$ 减少到

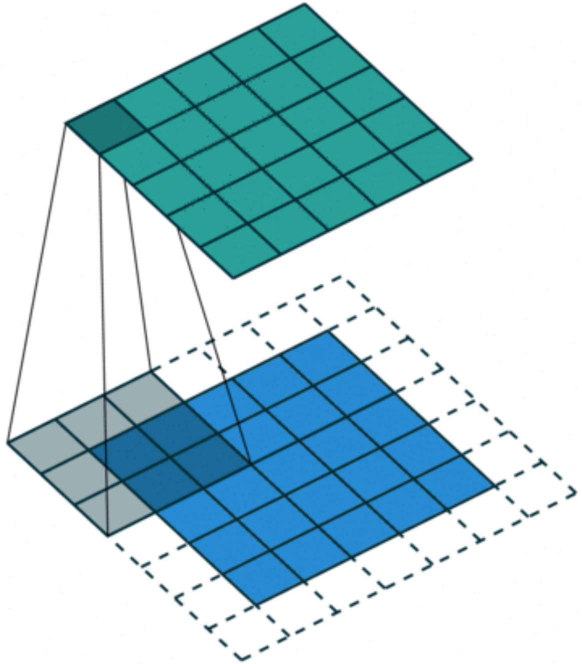
    $(n_h - k_h + 1) \times (n_w - k_w + 1)$



填充

在输入周围添加额外的行 / 列





$$0 \times 0 + 0 \times 1 + 0 \times 2 + 0 \times 3 = 0$$

填充

- 填充 p_h 行和 p_w 列，输出形状为

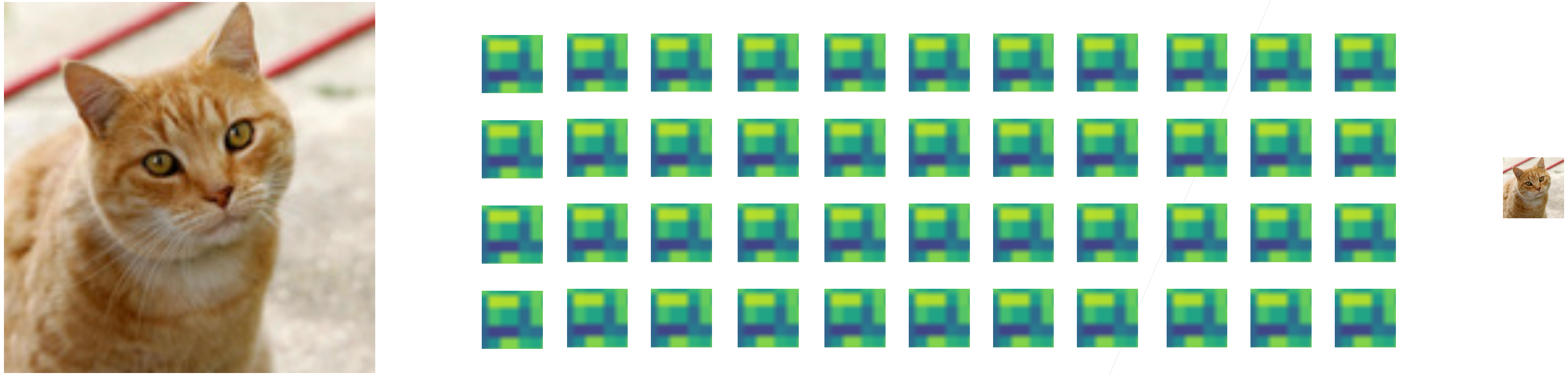
$$(n_h - k_h + p_h + 1) \times (n_w - k_w + p_w + 1)$$

- 通常取 $p_h = k_h - 1, \quad p_w = k_w - 1$
  - 当 k_h 为奇数：在上下两侧填充 $p_h/2$
  - 当 k_h 为偶数：在上侧填充 $\lceil p_h/2 \rceil$，在下侧填充 $\lfloor p_h/2 \rfloor$

步幅

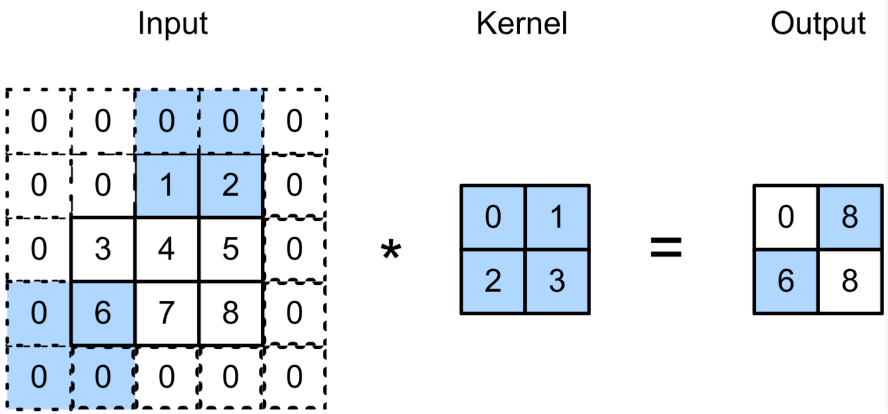
- ~~填充减小的输出大小~~与层数线性相关
  - 给定输入大小 224 x 224，在使用 5 x 5 卷积核的情况下，需要 ~~44~~ 层将输出降低到 4 x 4
  - 需要大量计算才能得到较小输出

课堂提问：需要多少层

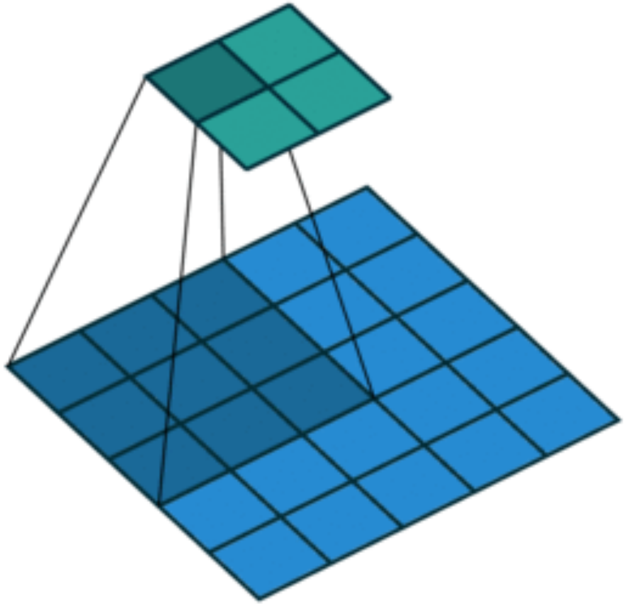


步幅

- 步幅是指行/列的滑动步长
  - 例：高度3 宽度2 的步幅



$$0 \times 0 + 0 \times 1 + 1 \times 2 + 2 \times 3 = 8$$
$$0 \times 0 + 6 \times 1 + 0 \times 2 + 0 \times 3 = 6$$



步幅

- 给定高度 s_h 和宽度 s_w 的步幅，输出形状是

$$\lfloor (n_h - k_h + p_h + s_h)/s_h \rfloor \times \lfloor (n_w - k_w + p_w + s_w)/s_w \rfloor$$

- 如果 $p_h = k_h - 1, \quad p_w = k_w - 1$

$$\lfloor (n_h + s_h - 1)/s_h \rfloor \times \lfloor (n_w + s_w - 1)/s_w \rfloor$$

- 如果输入高度和宽度可以被步幅整除

更进一步

$$(n_h/s_h) \times (n_w/s_w)$$

课堂：从零开始推导（最后一个结果）

总结

- 填充和步幅是卷积层的超参数
- 填充在输入周围添加额外的行 / 列，来控制输出形状的减少量
- 步幅是每次滑动核窗口时的行/列的步长，可以成倍的减少输出形状