



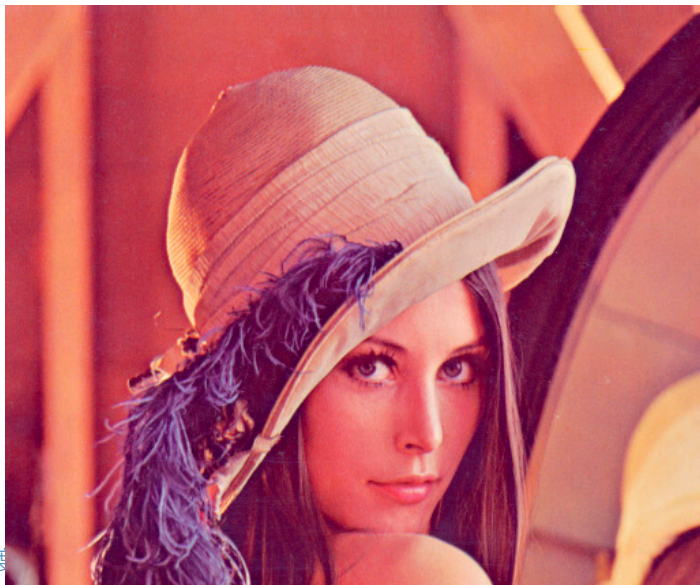
李沐 · AWS

多个输入和 输出通道



多个输入通道

- 彩色图像可能有 RGB 三个通道
- 转换为灰度会丢失信息





多个输入通道

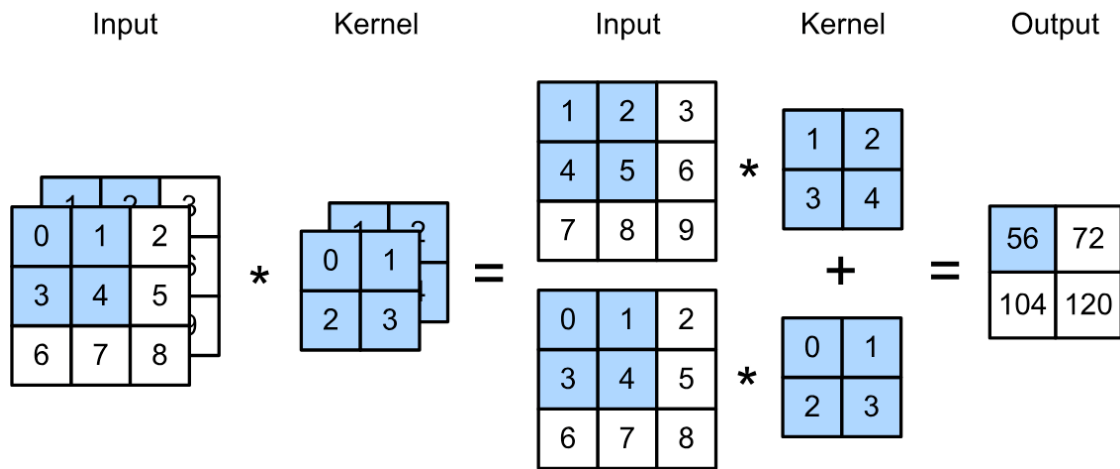
- 彩色图像可能有 RGB 三个通道
- 转换为灰度会丢失信息



多个输入通道



- 每个通道都有一个卷积核，结果是所有通道卷积结果的和



$$(1 \times 1 + 2 \times 2 + 4 \times 3 + 5 \times 4) \\ + (0 \times 0 + 1 \times 1 + 3 \times 2 + 4 \times 3) = 56$$



多个输入通道

- 输入 $\mathbf{X} : c_i \times n_h \times n_w$
- 核 $\mathbf{W} : c_i \times k_h \times k_w$
- 输出 $\mathbf{Y} : m_h \times m_w$

$$\mathbf{Y} = \sum_{i=0}^{c_i} \mathbf{X}_{i,:,:} \star \mathbf{W}_{i,:,:}$$



多个输出通道

- 无论有多少输入通道，到目前为止我们只用到单输出通道
- 我们可以有多个三维卷积核，每个核生成一个输出通道
- 输入 $\mathbf{X} : c_i \times n_h \times n_w$
- 核 $\mathbf{W} : c_o \times c_i \times k_h \times k_w$
- 输出 $\mathbf{Y} : c_o \times m_h \times m_w$

$$\mathbf{Y}_{i,:,:) = \mathbf{X} \star \mathbf{W}_{i,:,:,:) \quad \text{for } i = 1, \dots, c_o$$



多个输入和输出通道

- 每个输出通道可以识别特定模式

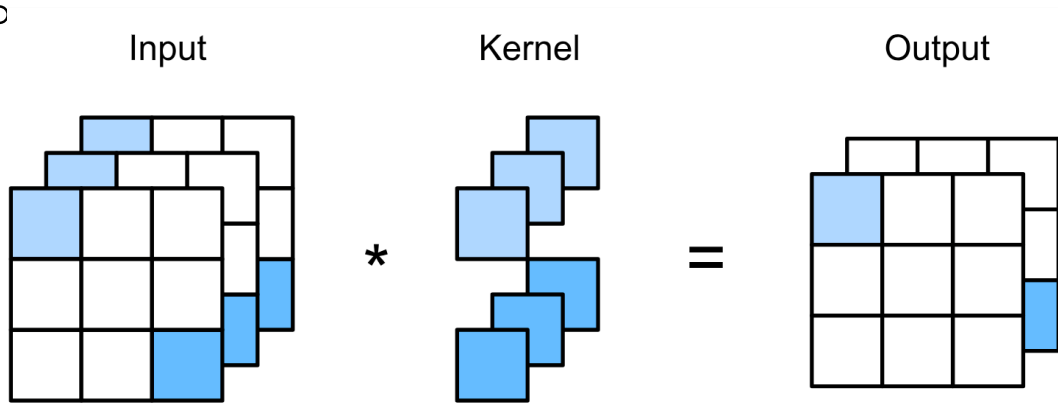


- 输入通道核识别并组合输入中的模式



1 x 1 卷积层

$k_h = k_w = 1$ 是一个受欢迎的选择。它不识别空间模式，只是融合通道。



相当于输入形状为 $n_h n_w \times c_i$ ，权重为 $c_o \times c_i$ 的全连接层



二维卷积层

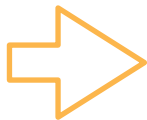
- 输入 $\mathbf{X} : c_i \times n_h \times n_w$
- 核 $\mathbf{W} : c_o \times c_i \times k_h \times k_w$
- 偏差 $\mathbf{B} : c_o \times c_i$
- 输出 $\mathbf{Y} : c_o \times m_h \times m_w$
- 计算复杂度（浮点计算数 FLOP） $O(c_i c_o k_h k_w m_h m_w)$

$$\mathbf{Y} = \mathbf{X} \star \mathbf{W} + \mathbf{B}$$

$$c_i = c_o = 100$$

$$k_h = h_w = 5$$

$$m_h = m_w = 64$$



1GFLOP

- 10 层, 1M 样本, 10 PFlops
(CPU: 0.15 TF = 18h, GPU: 12 TF = 14min)

总结



- 输出通道数是卷积层的超参数
- 每个输入通道有独立的二维卷积核，所有通道结果相加得到一个输出通道结果
- 每个输出通道有独立的三维卷积核