

日期： 2022 年 6 月 22 日



成绩： _____

学院： 智能工程学院

课程： 深度学习

周次： 第 19 周

专业： 智能科学与技术

姓名： 方桂安

学号： 20354027

1 题一

1.1 题目

使用一层深度可分离卷积将 $7 \times 7 \times 3$ 大小的 tensor 转换为 $5 \times 5 \times 64$ ，其中前两个维度为空间维度，第三个维度为特征维度。请使用图示说明计算过程，并分析使用的参数量及计算量的大小。

1.2 解答

1.2.1 标准卷积

假设我们有一个 $7 \times 7 \times 3$ 的输入图像，即图像尺寸为 7×7 ，通道数为 3，对图像进行 3×3 卷积，没有填充（padding）且步长为 1。如果我们只考虑图像的宽度和高度，使用 3×3 卷积来处理 7×7 大小的输入图像，最终可以得到一个 5×5 的输出特征图。然而，由于图像有 3 个通道，我们的卷积核也需要有 3 个通道。这就意味着，卷积核在每个位置进行计算时，实际上会执行 $3 \times 3 \times 3 = 27$ 次乘法。我们使用一个 $3 \times 3 \times 3$ 的卷积核进行卷积运算，最终可以得到 $5 \times 5 \times 1$ 的输出特征图。

如果我们想增加输出的 channel 数量让网络学习更多种特征，这时我们可以创建多个卷积核，比如 64 个卷积核来学习 64 个不同类别的特征。此时，64 个卷积核会分别进行运算，得到 64 个 $5 \times 5 \times 1$ 的输出特征图，将其堆叠在一起，最终可以得到 $5 \times 5 \times 64$ 的输出特征图。

1.2.2 深度可分离卷积

(1) 深度卷积运算

首先，我们对输入图像进行深度卷积运算，这里的深度卷积运算其实就是逐通道进行卷积运算。对于一幅 $7 \times 7 \times 3$ 的输入图像而言，我们使用大小为 3×3 的卷积核进行逐通道运算，其实就是使用 3 个 $3 \times 3 \times 1$ 的卷积核分别提取输入图像中 3 个 channel 的特征，每个卷积核计算完成后，会得到 3 个 $5 \times 5 \times 1$

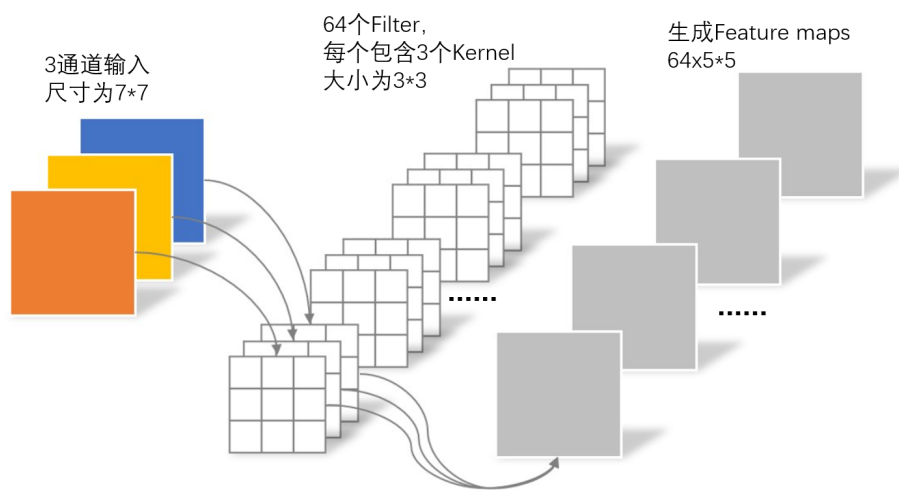


图 1: 标准卷积

的输出特征图，将这些特征图堆叠在一起就可以得到大小为 $5 \times 5 \times 3$ 的最终输出特征图。这里我们可以发现深度卷积运算的一个缺点，深度卷积运算缺少通道间的特征融合，并且运算前后通道数无法改变。

因此，接下来就需要连接一个逐点卷积来弥补它的缺点。

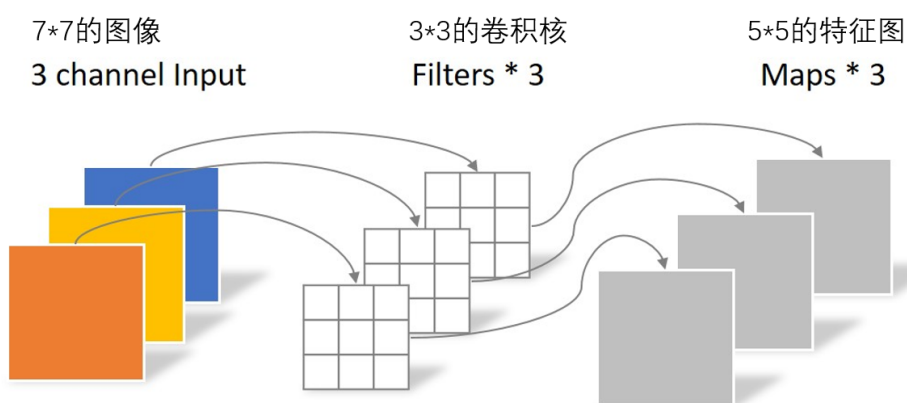


图 2: 深度卷积

(2) 逐点卷积运算

前面我们使用深度卷积运算完成了从一幅 $7 \times 7 \times 3$ 的输入图像中得到 $5 \times 5 \times 3$ 的输出特征图，并且发现仅使用深度卷积无法实现不同通道间的特征融合，而且也无法得到与标准卷积运算一致的 $5 \times 5 \times 64$ 的特征图。那么，接下来就让我们看一下如何使用逐点卷积实现这两个任务。

逐点卷积其实就是 1×1 卷积，因为会遍历每个点，所以我们称之为逐点卷积。我们使用一个 3 通道的 1×1 卷积对上文中得到的 $5 \times 5 \times 3$ 的特征图进行运算，可以得到一个 $5 \times 5 \times 1$ 的输出特征图。此时，我们就使用逐点卷积实现了融合 3 个通道间特征的功能。

此外，我们可以创建 64 个 3 通道的 1×1 卷积对上文中得到的 $5 \times 5 \times 3$ 的特征图进行运算，这样，就可以实现得到与标准卷积运算一致的 $5 \times 5 \times 64$ 的特征图的功能。

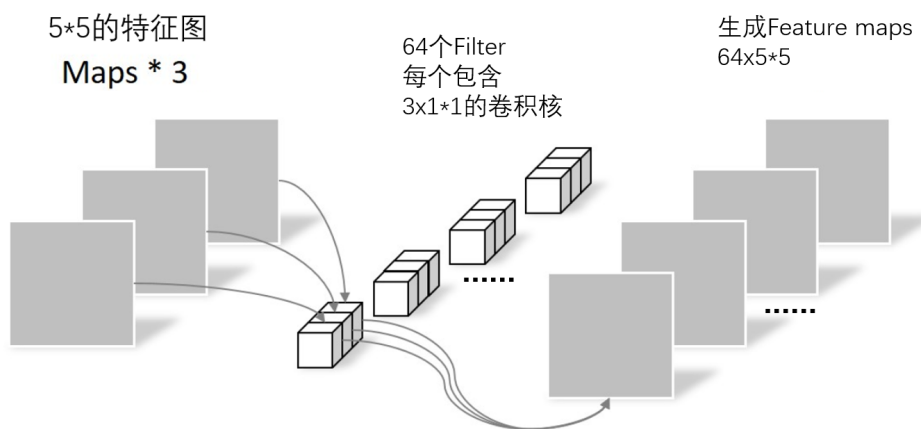


图 3: 逐点卷积

1.2.3 计算量与参数量分析

这里我们看一下上文例子中标准卷积的乘法运算个数，我们创建了 64 个 $3 \times 3 \times 3$ 的卷积核进行卷积运算，每个卷积核会在输入图片上移动 5×5 次，因此总的乘法运算个数为：

$$64 \times 3 \times 3 \times 3 \times 5 \times 5 = 43200 \quad (1)$$

参数量为：

$$3 \times 3 \times 3 \times 64 = 1728 \quad (2)$$

而换成深度可分离卷积后，在深度卷积运算时，我们使用 3 个 $3 \times 3 \times 1$ 的卷积核在输入图片上移动 5×5 次，此时乘法运算个数为：

$$3 \times 3 \times 3 \times 5 \times 5 = 675 \quad (3)$$

在逐点卷积运算时，我们使用 64 个 $1 \times 1 \times 3$ 的卷积在输入特征图上移动 5×5 次，此时乘法运算个数为：

$$64 \times 1 \times 1 \times 3 \times 5 \times 5 = 4800 \quad (4)$$

将这两步运算相加，即可得到，使用深度可分离卷积后，总的乘法运算个数变为：5475。

参数量为：

$$3 \times 3 \times 3 + 3 \times 64 = 219 \quad (5)$$

可以看到，深度可分离卷积的运算量相较标准卷积而言，计算量与参数量都少了很多。