

深度可分离卷积



Joe.Zhao

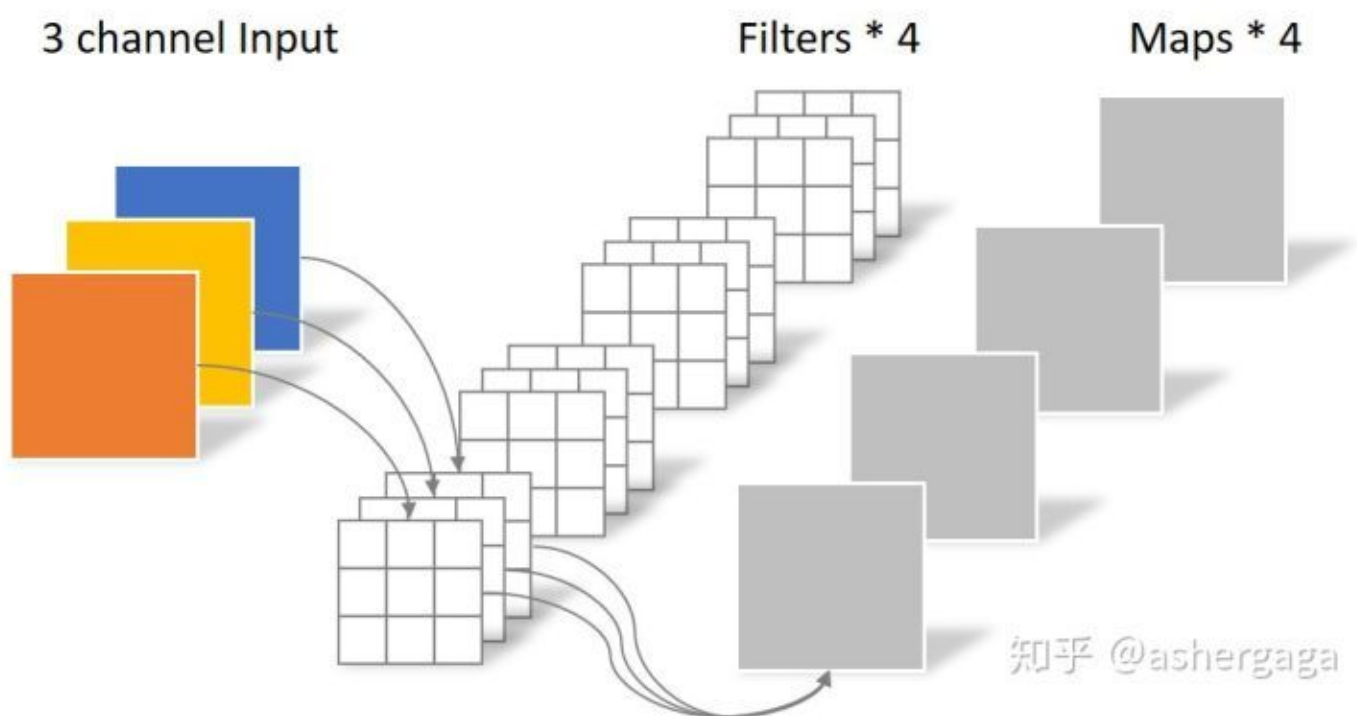
301 人赞同了该文章

一些轻量级的网络，如mobilenet中，会有深度可分离卷积depthwise separable convolution，由depthwise(DW)和pointwise(PW)两个部分结合起来，用来提取特征feature map

相比常规的卷积操作，其参数数量和运算成本比较低

常规卷积操作

对于一张 5×5 像素、三通道（shape为 $5 \times 5 \times 3$ ），经过 3×3 卷积核的卷积层（假设输出通道数为4，则卷积核shape为 $3 \times 3 \times 3 \times 4$ ，最终输出4个Feature Map，如果有same padding则尺寸与输入层相同（ 5×5 ），如果没有则为尺寸变为 3×3 ）



卷积层共4个Filter，每个Filter包含了3个Kernel，每个Kernel的大小为 3×3 。因此卷积层的参数数量可以用如下公式来计算：

$$N_{std} = 4 \times 3 \times 3 \times 3 = 108$$

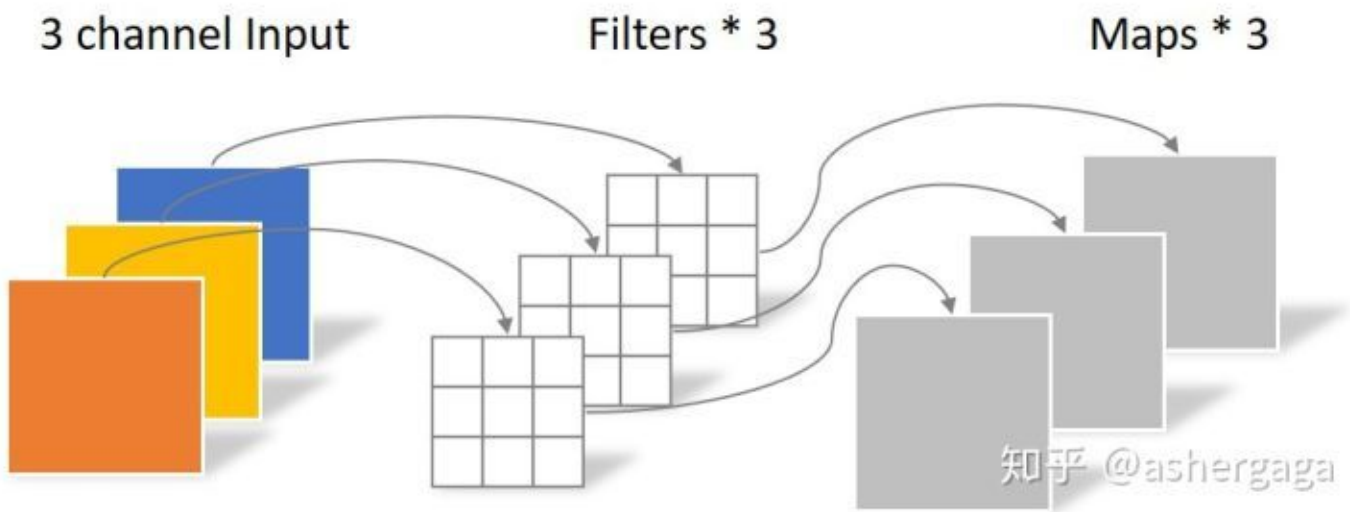
深度可分离卷积



- 逐通道卷积

Depthwise Convolution的一个卷积核负责一个通道，一个通道只被一个卷积核卷积

一张 5×5 像素、三通道彩色输入图片（shape为 $5 \times 5 \times 3$ ），Depthwise Convolution首先经过第一次卷积运算，DW完全是在二维平面内进行。卷积核的数量与上一层的通道数相同（通道和卷积核一一对应）。所以一个三通道的图像经过运算后生成了3个Feature map(如果有same padding则尺寸与输入层相同为 5×5)，如下图所示。



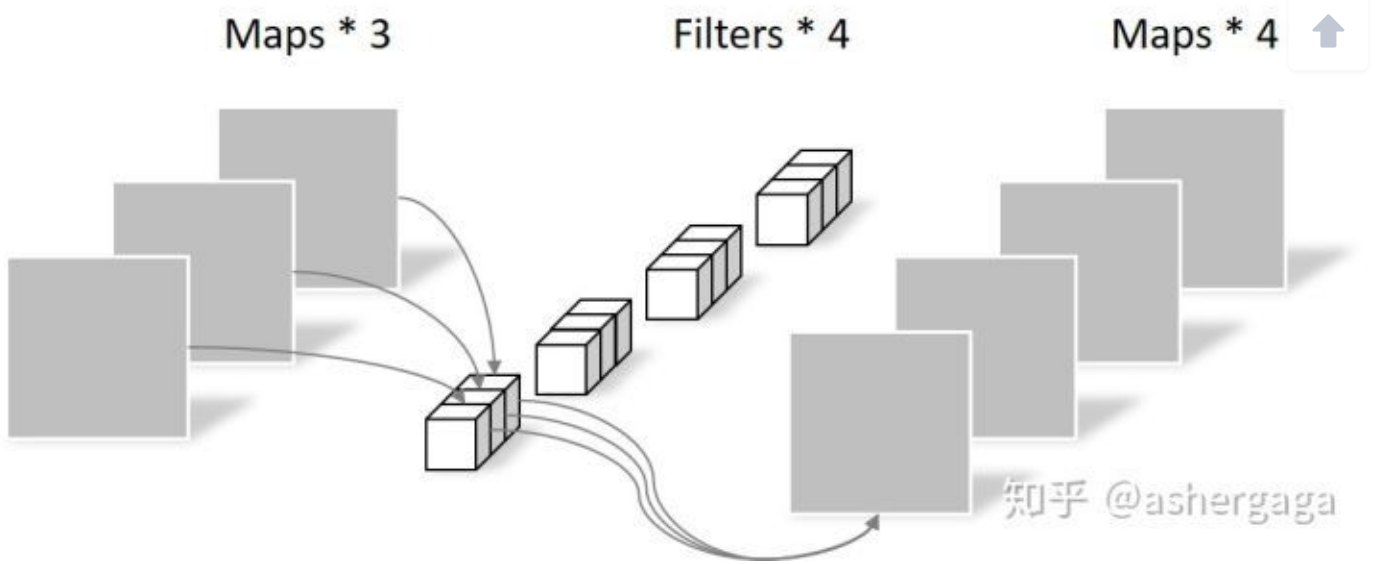
其中一个Filter只包含一个大小为 3×3 的Kernel，卷积部分的参数个数计算如下：

$$N_{\text{depthwise}} = 3 \times 3 \times 3 = 27$$

Depthwise Convolution完成后的Feature map数量与输入层的通道数相同，无法扩展Feature map。而且这种运算对输入层的每个通道独立进行卷积运算，没有有效的利用不同通道在相同空间位置上的feature信息。因此需要Pointwise Convolution来将这些Feature map进行组合生成新的Feature map

- 逐点卷积

Pointwise Convolution的运算与常规卷积运算非常相似，它的卷积核的尺寸为 $1 \times 1 \times M$ ，M为上一层的通道数。所以这里的卷积运算会将上一步的map在深度方向上进行加权组合，生成新的Feature map。有几个卷积核就有几个输出Feature map



由于采用的是 1×1 卷积的方式，此步中卷积涉及到的参数个数可以计算为：

$$N_{\text{pointwise}} = 1 \times 1 \times 3 \times 4 = 12$$

经过Pointwise Convolution之后，同样输出了4张Feature map，与常规卷积的输出维度相同

参数对比

回顾一下，常规卷积的参数个数为：

$$N_{\text{std}} = 4 \times 3 \times 3 \times 3 = 108$$

Separable Convolution的参数由两部分相加得到：

$$N_{\text{depthwise}} = 3 \times 3 \times 3 = 27$$

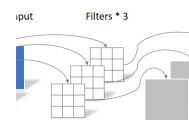
$$N_{\text{pointwise}} = 1 \times 1 \times 3 \times 4 = 12$$

$$N_{\text{separable}} = N_{\text{depthwise}} + N_{\text{pointwise}} = 39$$

相同的输入，同样是得到4张Feature map，Separable Convolution的参数个数是常规卷积的约 $1/3$ 。因此，在参数量相同的前提下，采用Separable Convolution的神经网络层数可以做的更深。

ashergaga: Depthwise卷积与
Pointwise卷积

zhuanlan.zhihu.com



<https://yinguobing.com/separable-convolution/#fn2>

yinguobing.com

