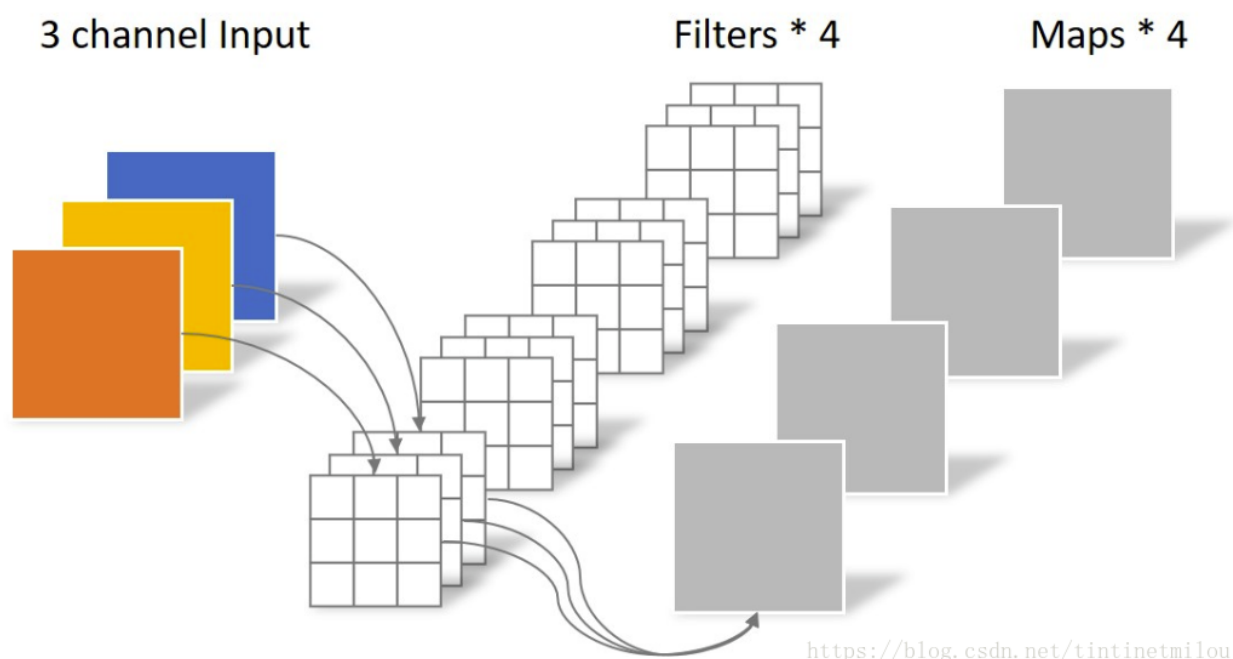


Depthwise (DW) 卷积与Pointwise (PW) 卷积，合起来被称作Depthwise Separable Convolution(参见Google的[Xception](#))，该结构和常规卷积操作类似，可用来提取特征，但相比于常规卷积操作，其参数量和运算成本较低。所以在一些轻量级网络中会碰到这种结构如[MobileNet](#)。

常规卷积操作

对于一张 5×5 像素、三通道彩色输入图片（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 。



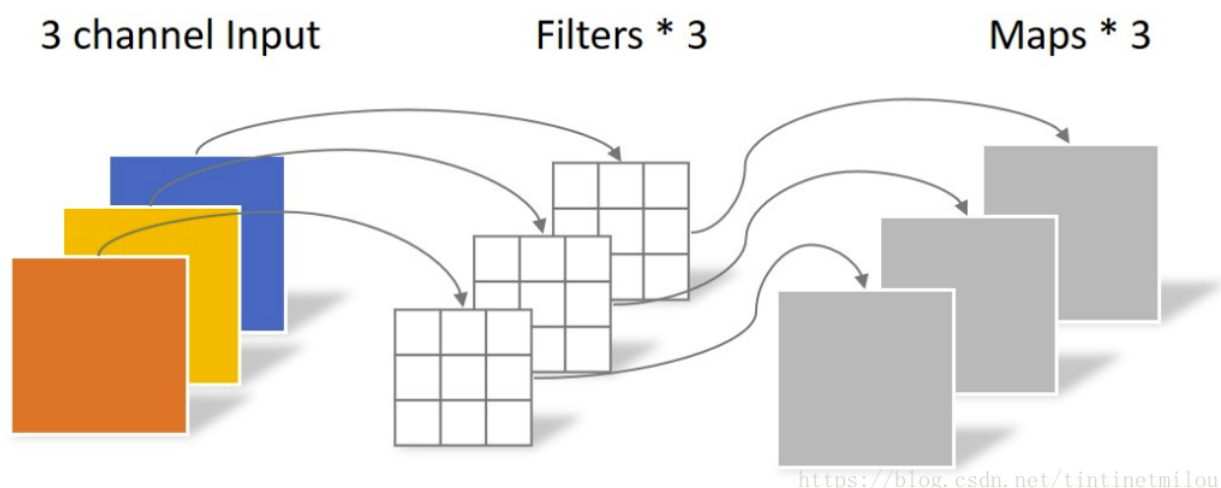
Depthwise Separable Convolution

Depthwise Separable Convolution是将一个完整的卷积运算分解为两步进行，即Depthwise Convolution与Pointwise Convolution。

Depthwise Convolution

不同于常规卷积操作，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。上面所提到的常规卷积每个卷积核是同时操作输入图片的每个通道。

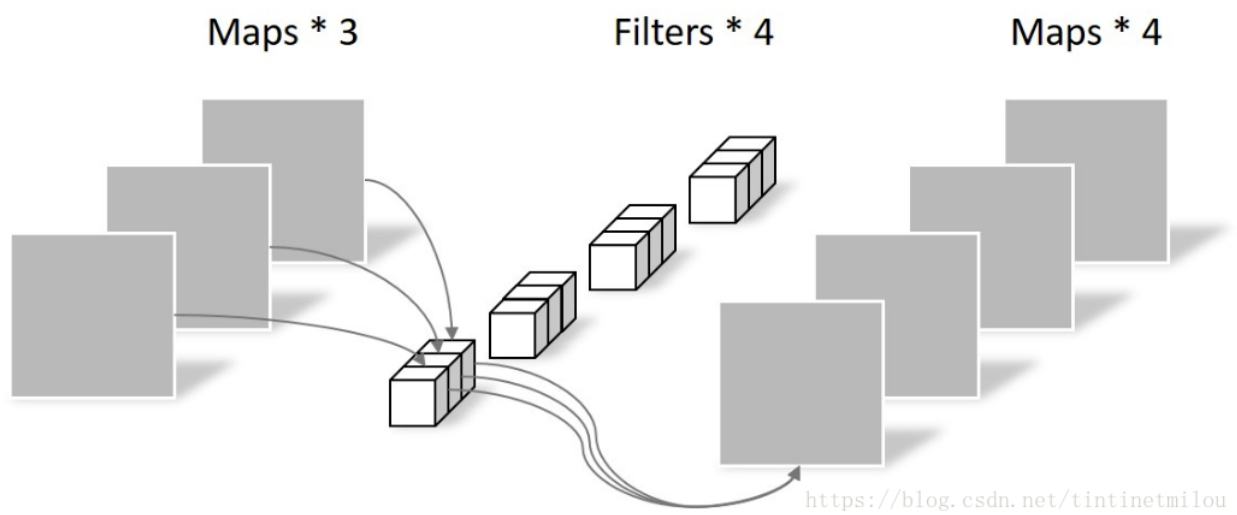
同样对于一张 5×5 像素、三通道彩色输入图片（shape为 $5 \times 5 \times 3$ ），Depthwise Convolution首先经过第一次卷积运算，不同于上面的常规卷积，DW完全是在二维平面内进行。卷积核的数量与上一层的通道数相同（通道和卷积核一一对应）。所以一个三通道的图像经过运算后生成了3个Feature map(如果有same padding则尺寸与输入层相同为 5×5)，如下图所示。



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。

Pointwise Convolution

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。如下图所示。



本文参考尹国冰的博客—[卷积神经网络中的Separable Convolution](https://blog.csdn.net/tintinetmilou)