

【完结】总结12大CNN主流模型架构设计思想



有三AI

🕒 2019-07-09 10:49:44

💬 0

👁 529

👍 0

文章

专栏《CNN模型解读》正式完结了，在这一个专栏中，我们给大家回顾了深度学习中的各类具有代表性的CNN模型，详细分析了各类模型的特点，设计思想。当然，这一个系列不可能包含所有的模型，但是我们可以从中洞见最核心的思想。如果有必要，以后我们还会进行补充的。

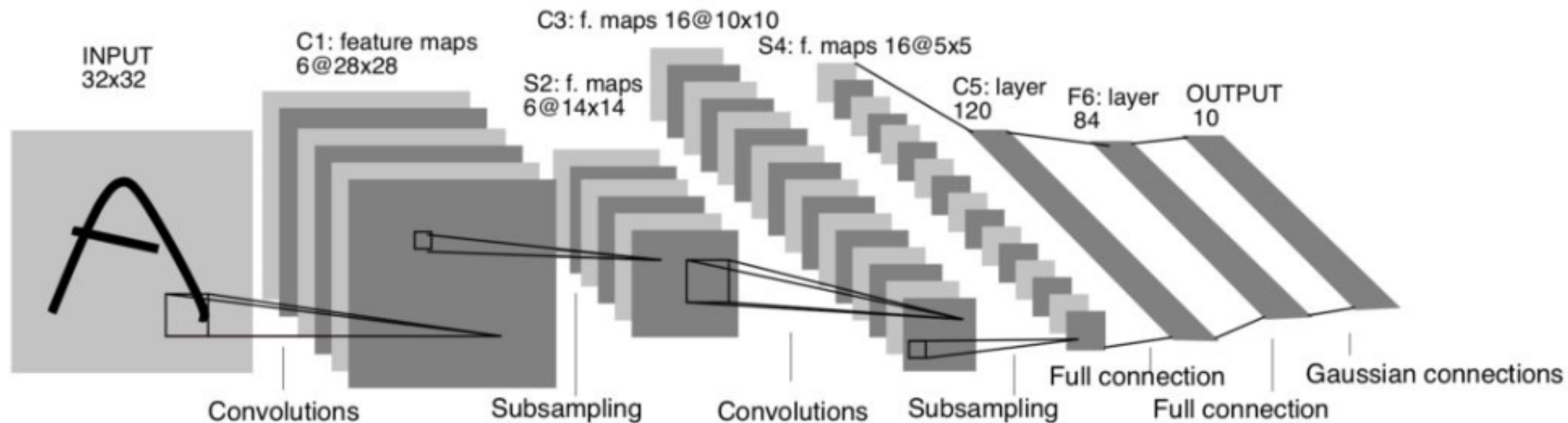
作者 | 言有三

编辑 | 言有三

来源自公众号：有三AI (yanyousan_ai)

01 从LeNet5到VGG

LeNet5不是CNN的起点，但却是它的hello world，让大家看到了卷积神经网络商用的前景。



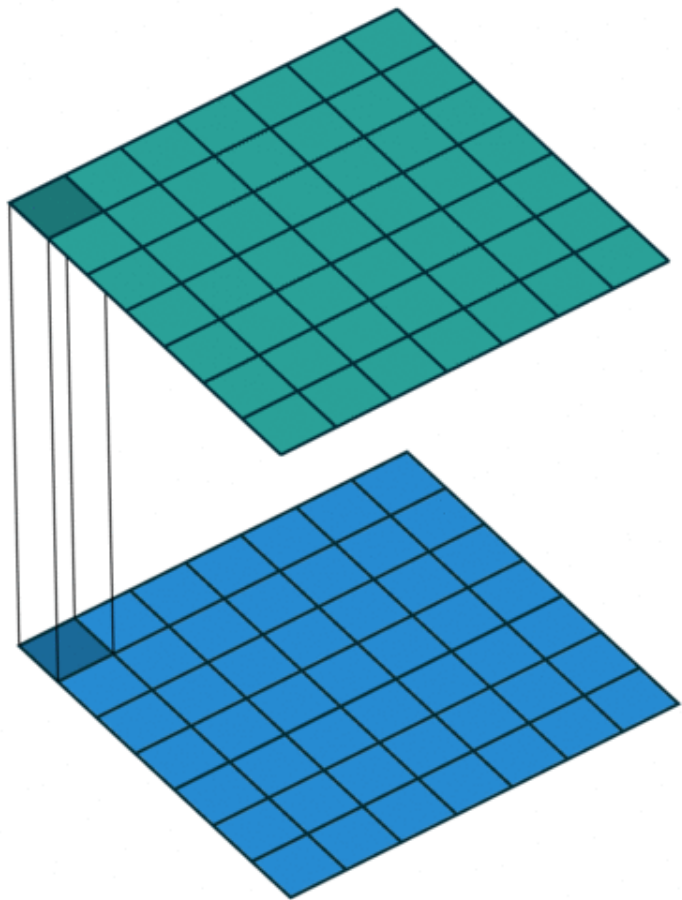
AlexNet是CNN向大规模商用打响的第一枪，夺得ImageNet 2012年分类冠军，宣告神经网络的王者归来。VGG以其简单的结构，在提出的若干年内在各大计算机视觉领域都成为了最广泛使用的benchmark。

它们都有着简单而又优雅的结构，同出一门。诠释了增加深度是如何提高了深度学习模型的性能。详细解读如下：

[【模型解读】从LeNet到VGG，看卷积+池化串联的网络结构](#)

02 1*1卷积

1*1卷积本身只是 $N \times N$ 卷积的卷积核半径大小退化为1时的特例，但是由于它以较小的计算代价增强了网络的非线性表达能力，给网络结构在横向和纵向拓展提供了非常好的工具，常用于升维和降维操作，尤其是在深层网络和对计算效率有较高要求的网络中广泛使用。

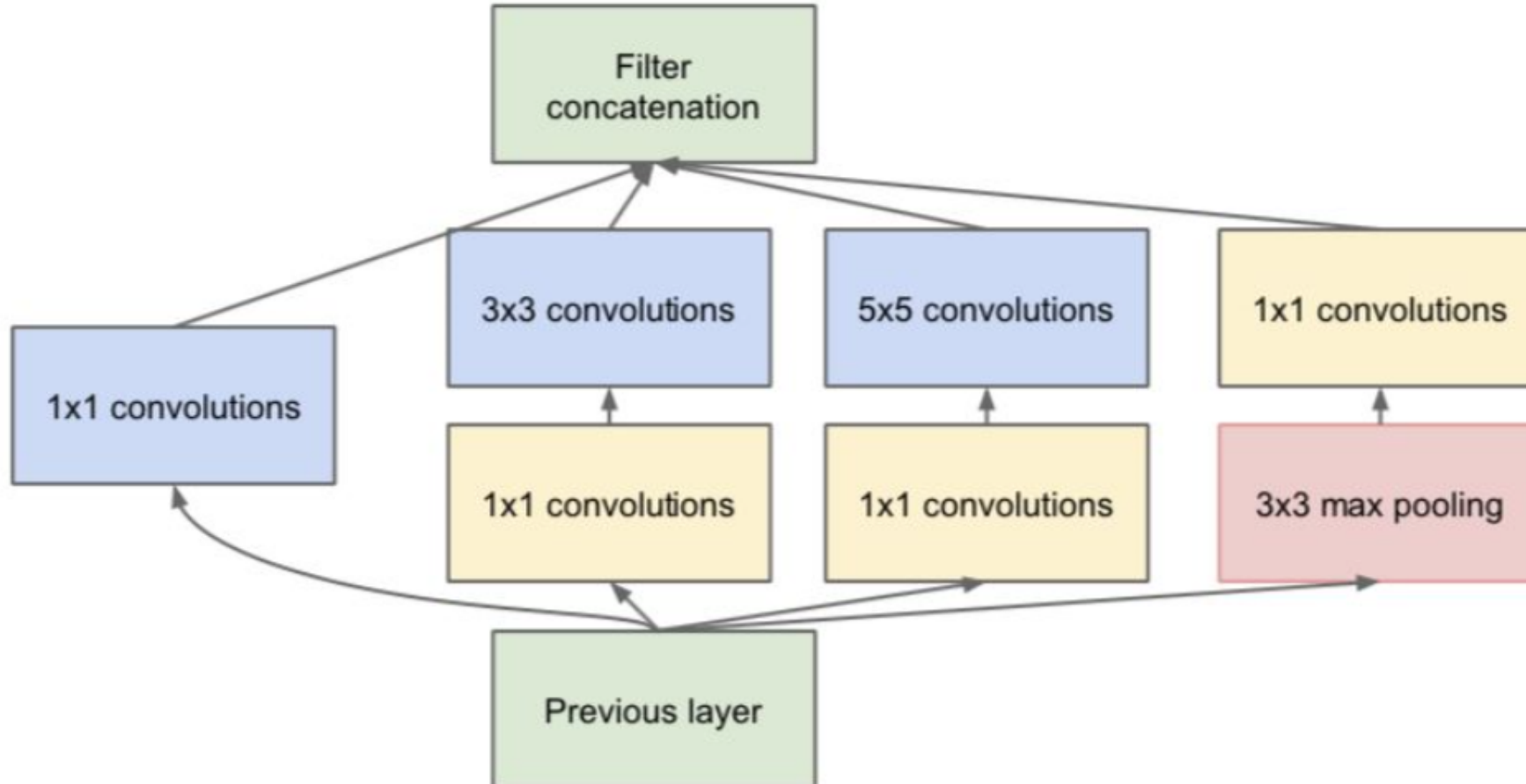


详细解读如下:

[【模型解读】network in network中的1*1卷积, 你懂了吗](#)

03 GoogLeNet

GoogLeNet夺得ImageNet2014年分类冠军, 也被称为Inception V1。Inception V1有22层深, 参数量为5M。同一时期的VGGNet性能和Inception V1差不多, 但是参数量却远大于Inception V1。Inception的优良特性得益于Inception Module, 结构如下图:



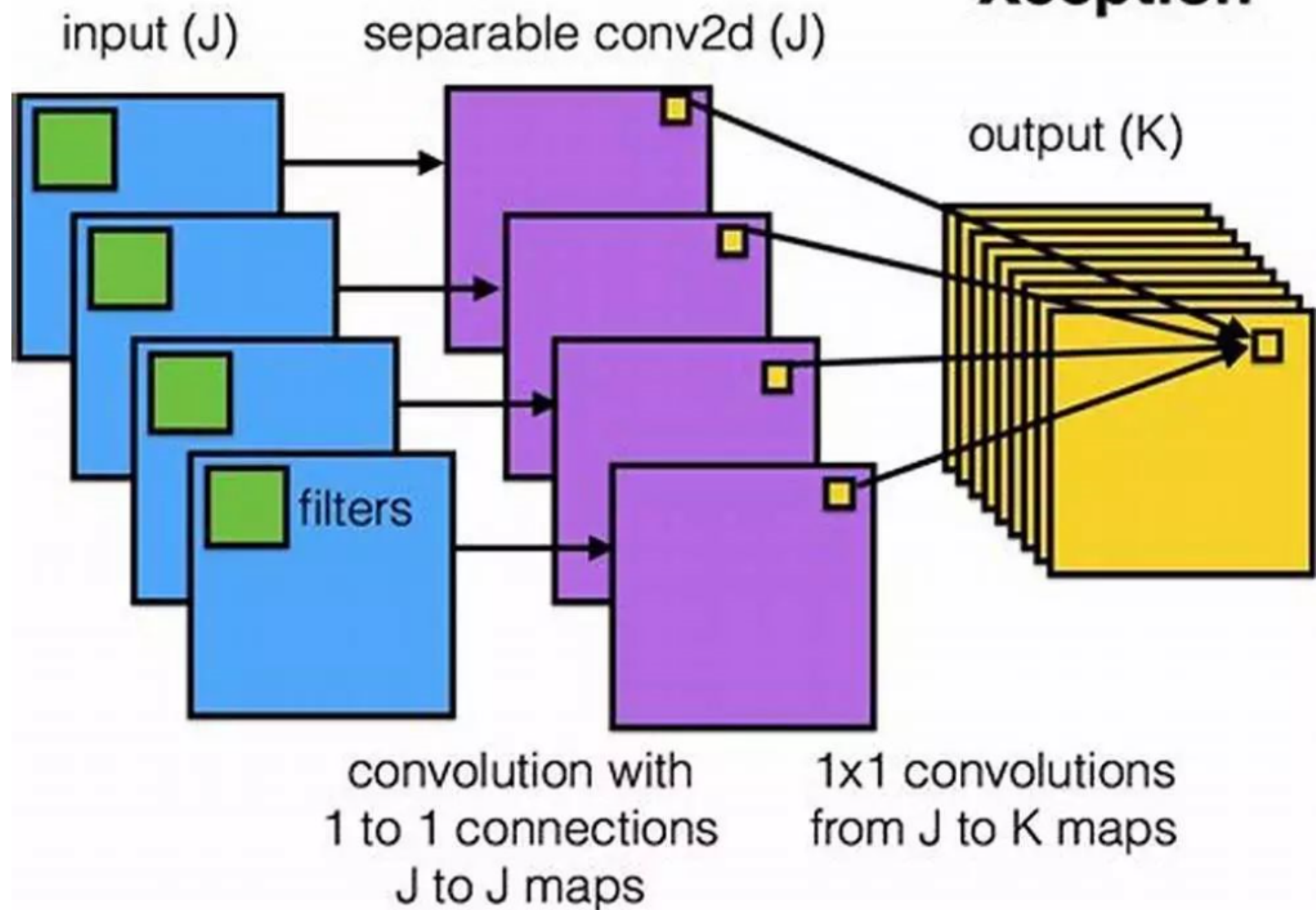
由1*1卷积，3*3卷积，5*5卷积，3*3最大池化四个并行通道运算结果进行融合，提取图像不同尺度的信息。如果说VGG是以深度取胜，那么GoogLeNet可以说是以宽度取胜，当然1*1卷积起到了很大的作用，这一点在SqueezeNet中也很关键。详细解读如下：

[【模型解读】GoogLeNet中的inception结构，你懂了吗](#)

04 MobileNets

脱胎于Xception的网络结构MobileNets使用Depthwise Separable Convolution(深度可分离卷积)构建了轻量级的28层神经网络，成为了移动端上的高性能优秀基准模型。

Xception

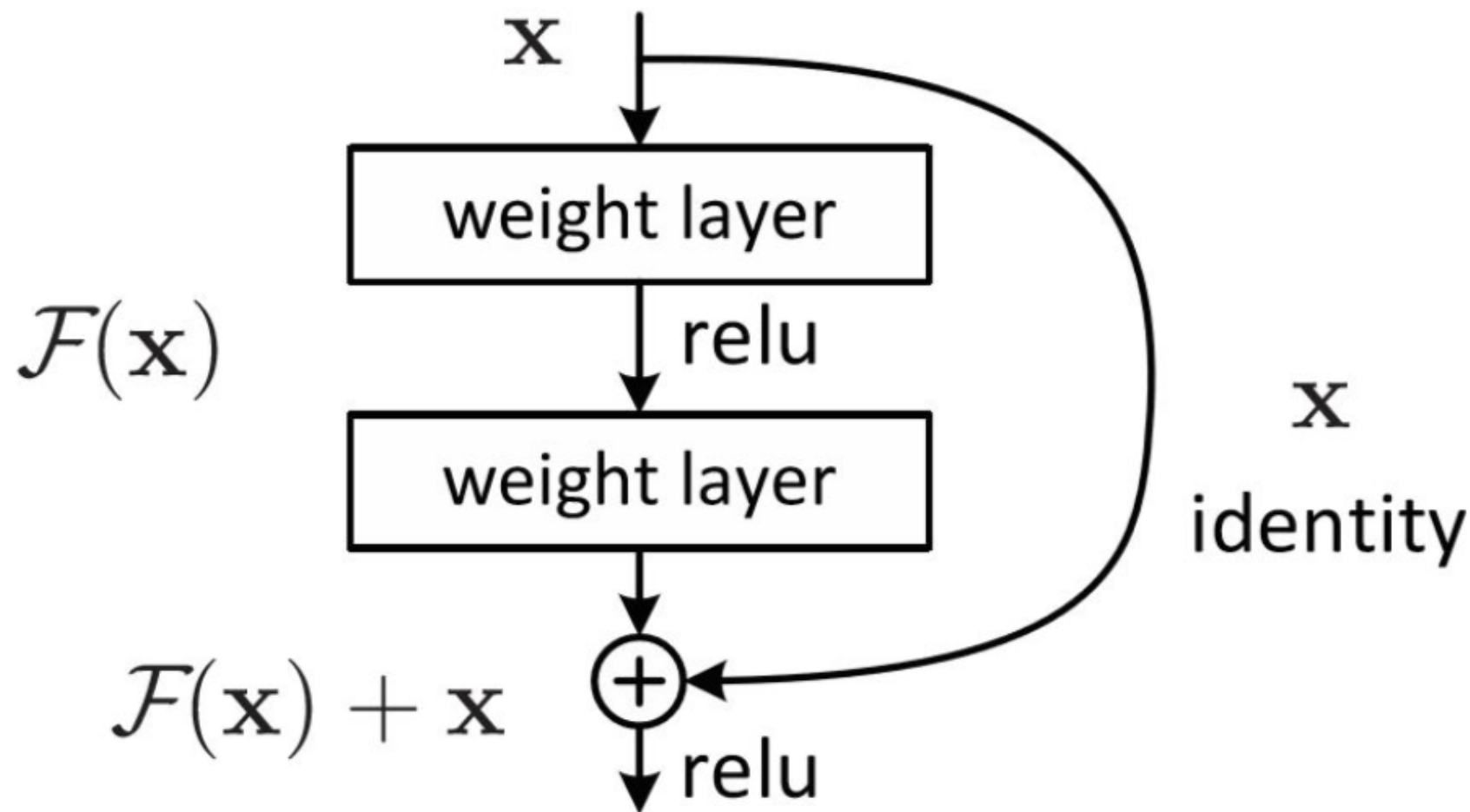


一个depthwise convolution，专注于该通道内的空间信息，一个pointwise convolution，专注于跨通道的信息融合，两者共同努力，然后强大，在此基础上的一系列模型如shufflenet等都是后话。详细解读如下：

[【模型解读】说说移动端基准模型MobileNets](#)

05 残差网络

当深层网络陷身于梯度消失等问题而导致不能很有效地训练更深的网络时，脱胎于highway network的残差网络应运而生，附带着MSRA和何凯明的学术光环，诠释了因为简单，所以有效，但你未必能想到和做到的朴素的道理。



详细解读如下：

[【模型解读】resnet中的残差连接，你确定真的看懂了？](#)

06 非正常卷积

谁说卷积一定要规规矩矩四四方方呢？MSRA总是一个出新点子的地方，在spatial transform network和active convolution的铺垫下，可变形卷积deformable convolution network如期而至。

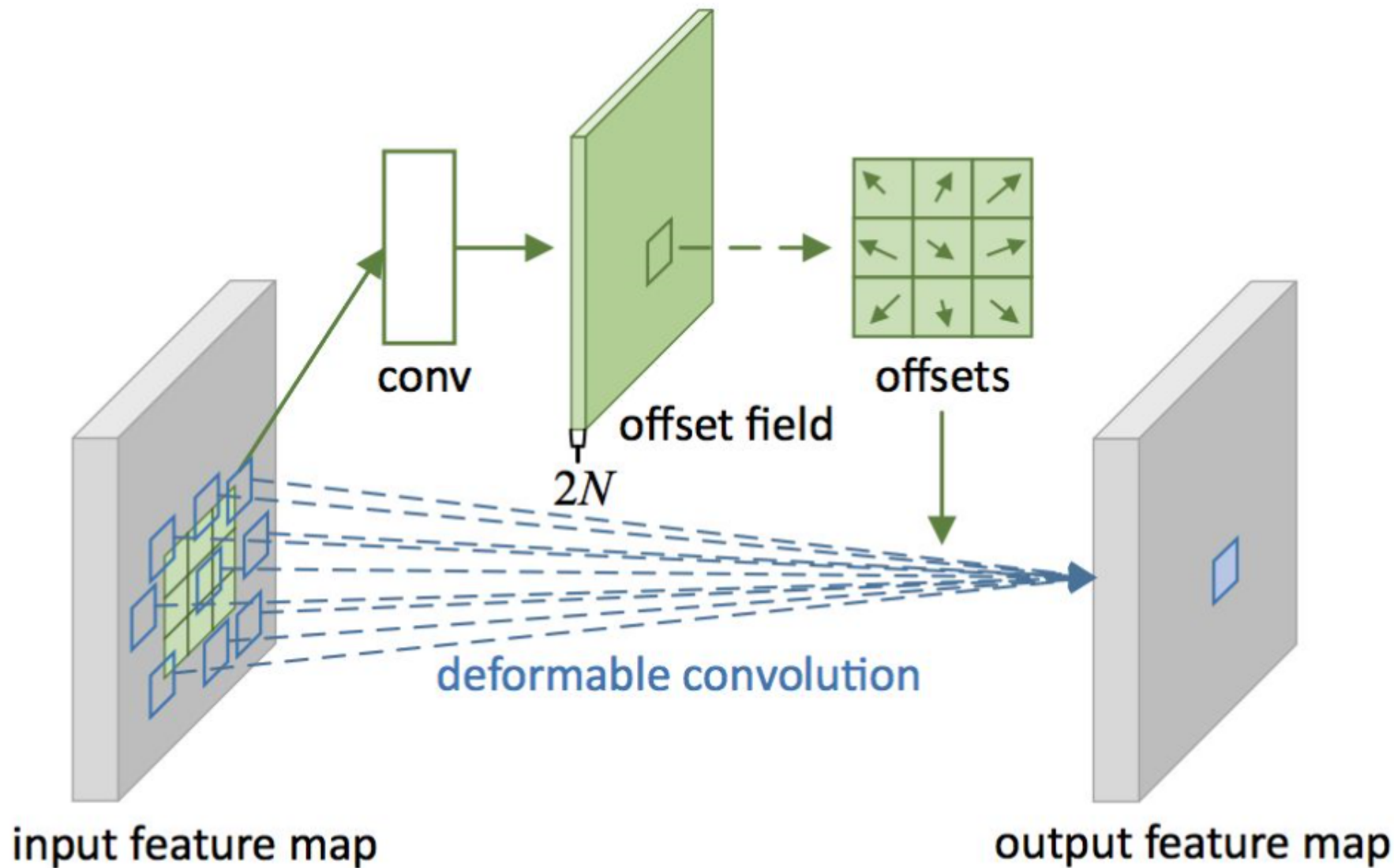
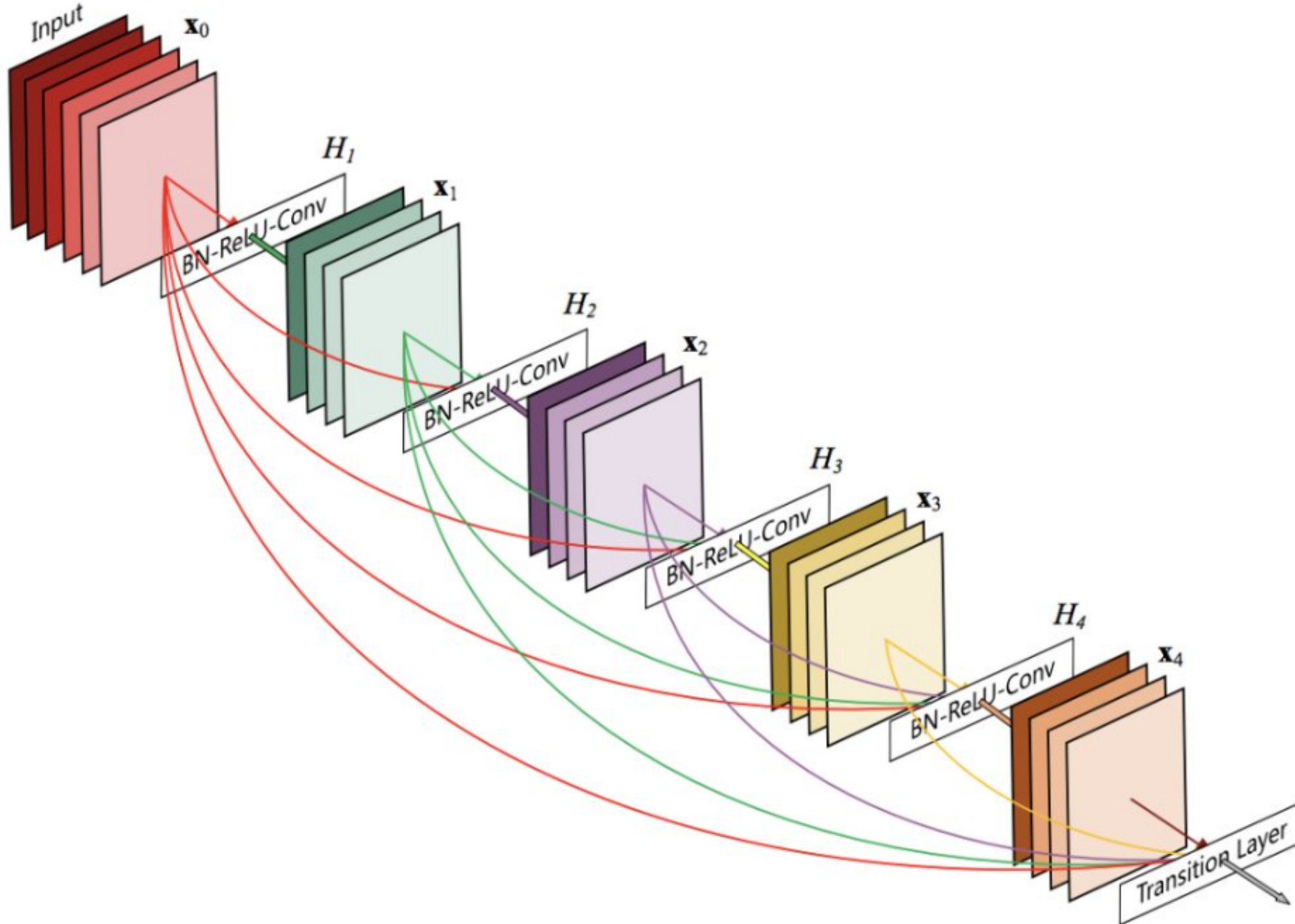


Figure 2: Illustration of 3×3 deformable convolution.

文章依旧写的很简单，这是一个致力于提升CNN对具有不同几何形变物体识别能力的模型，关键在于可变的感受野。

07 密集连接网络

说起来，DenseNet只不过是残差网络的升级版，将网络中的每一层都直接与其前面层相连，把残差做到了极致，提高了特征的利用率；因为可以把网络的每一层设计得很窄，提高计算性能。

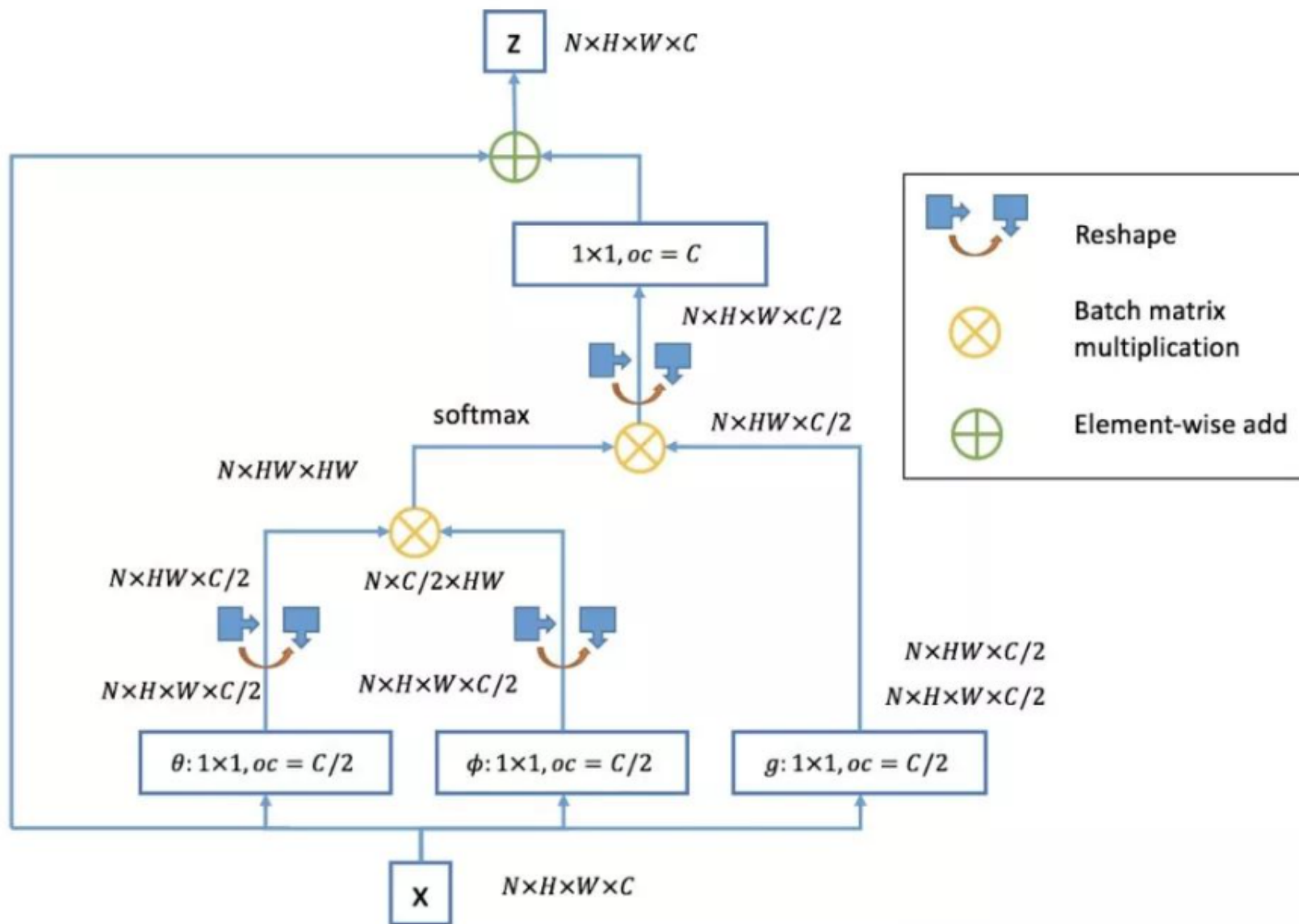


不过还是那句话，就算你能想到，也未必能做到，我们还是单独详细解读如下：

[【模型解读】全连接的卷积网络，有什么好？](#)

08 非局部神经网络

卷积神经网络因为局部连接和权重共享而成功，但是它的感受野是有限的。为了这样，我们不得不使用更深的网络，由此带来了三个问题。(1) 计算效率不高。(2) 感知效率不高。(3) 增加优化难度。这一次又是学神凯明带队出发，从传统降噪算法Non-Local中完成借鉴。

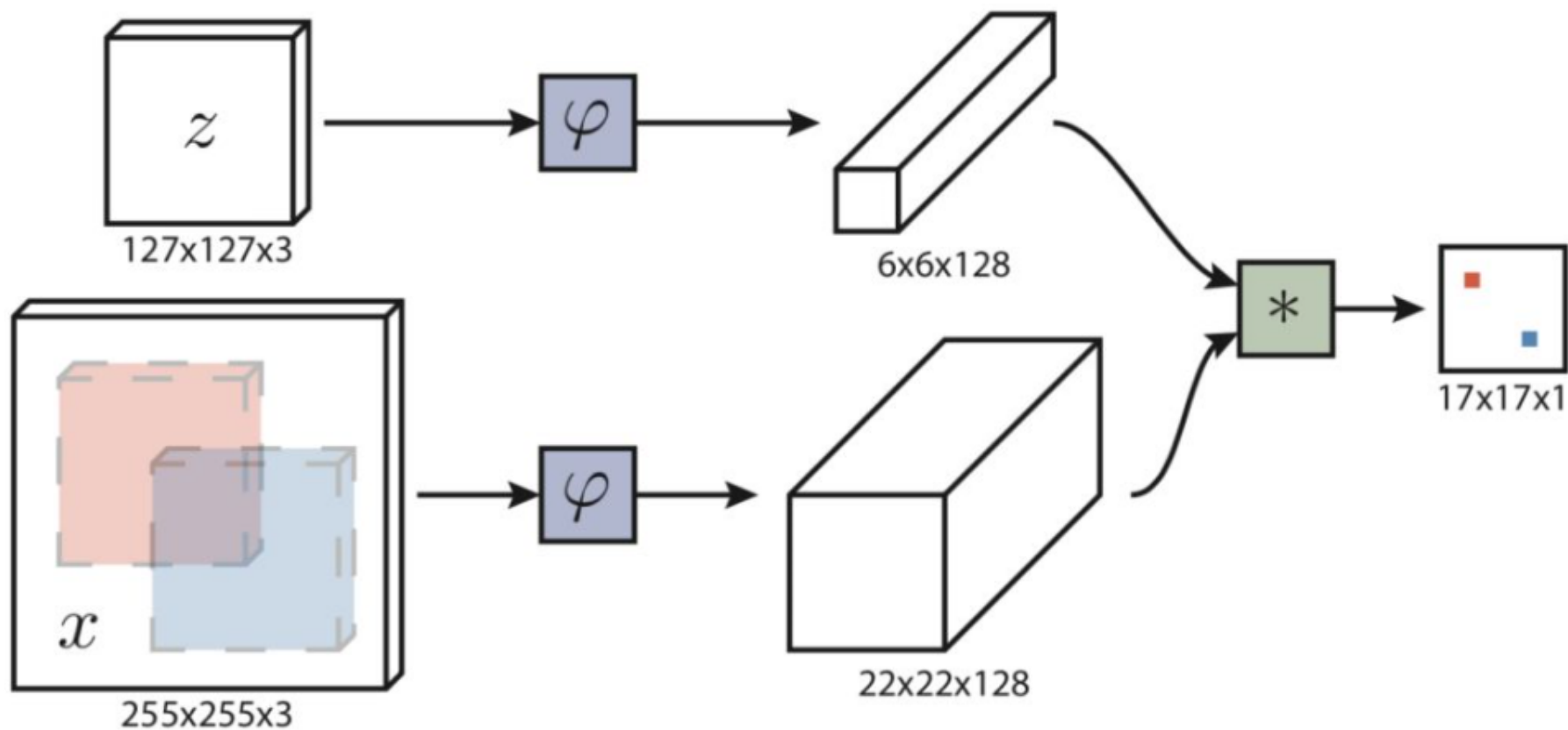


虽非真主流，了解一下也无妨。

[【模型解读】从“局部连接”回到“全连接”的神经网络](#)

09 多输入网络

见惯了输入一个图像或者视频序列，输出分类，分割，目标检测等结果的网络，是否会想起输入两张，或者多张图片来完成一些任务呢，这就是多输入网络结构。

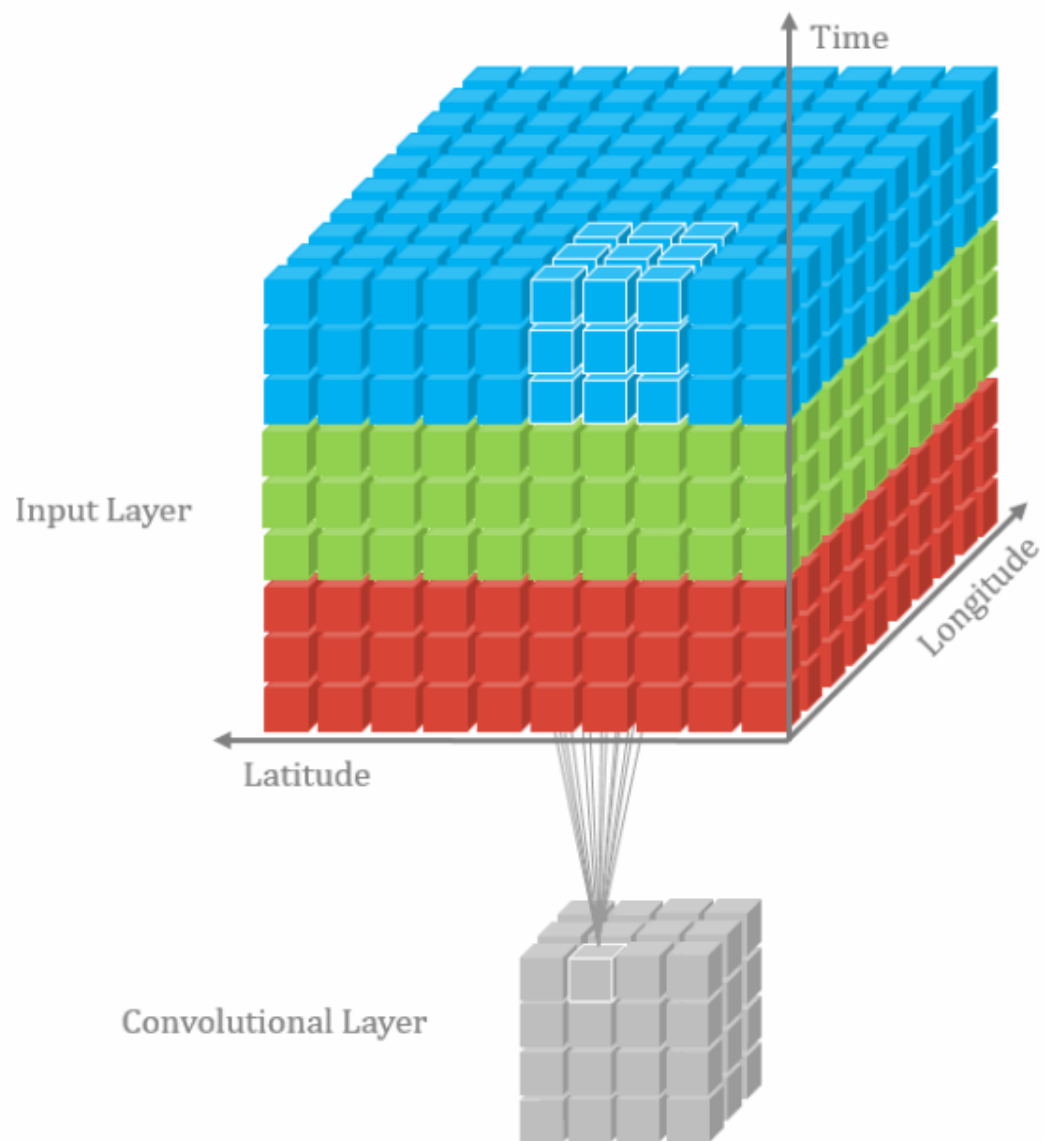


从检索，比对，到排序，跟踪，它可以做的事情有很多，你应该了解一下。

[【模型解读】深度学习网络只能有一个输入吗](#)

10 3D卷积

2D卷积玩腻了，该跳到更加高维的卷积了，常见的也就是3D卷积了。

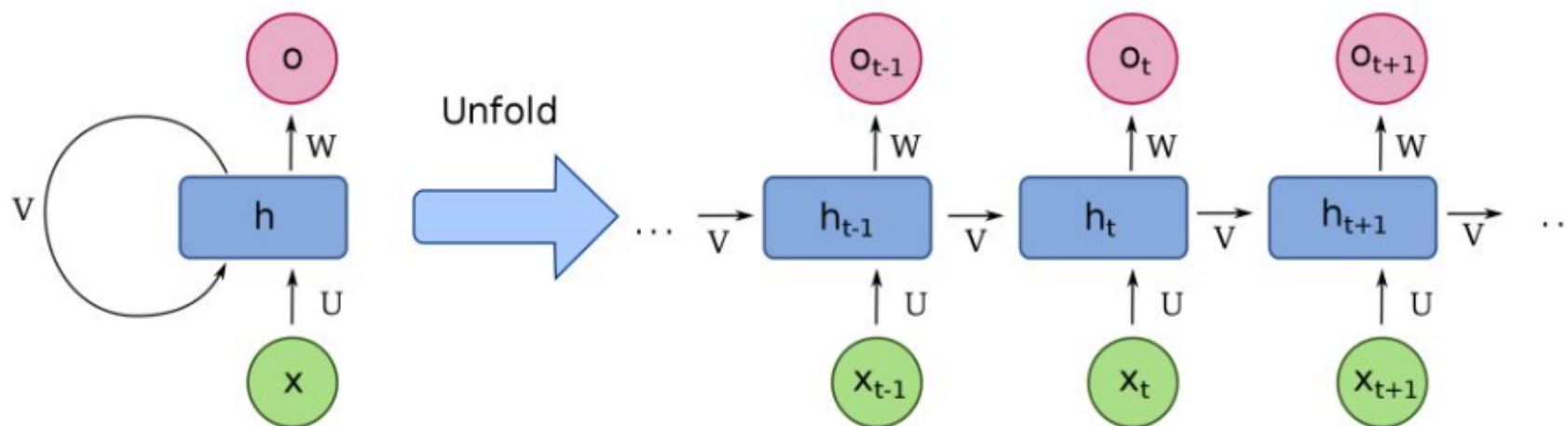


虽然3D带来了暴涨的计算量，但是想想可以用于视频分类和分割，3D点云，想想也是有些小激动呢。

[【模型解读】从2D卷积到3D卷积，都有什么不一样](#)

11 RNN和LSTM

不是所有的输入都是一张图片，有很多的信息是非固定长度或者大小的，比如视频，语音，此时就轮到RNN，LSTM出场了。

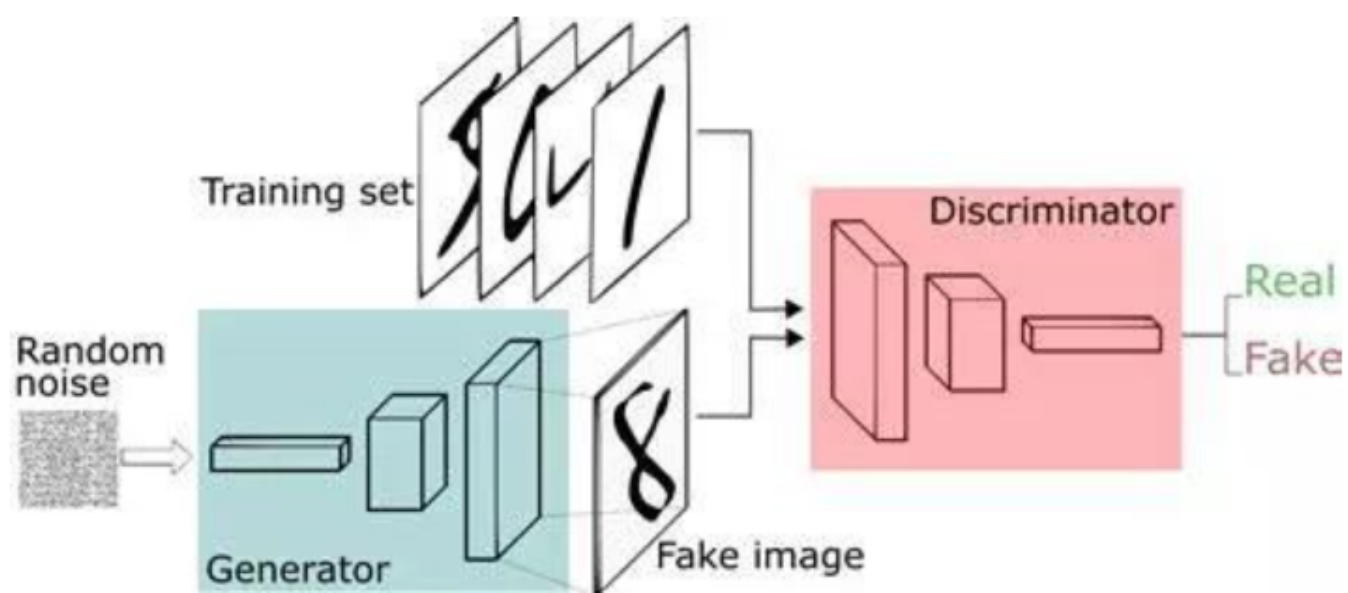


话不多说，好好学：

[【模型解读】浅析RNN到LSTM](#)

12 GAN

近几年来无监督学习领域甚至是深度学习领域里最大的进展非生成对抗网络GAN莫属，被誉为下一代深度学习，不管是研究热度还是论文数量，已经逼近甚至超越传统判别式的CNN架构。在研究者们的热情下，GAN已经从刚开始的一个生成器一个判别器发展到了多个生成器多个判别器等各种各样的结构。



快上车，因为真的快来不及了。

[【模型解读】历数GAN的5大基本结构](#)