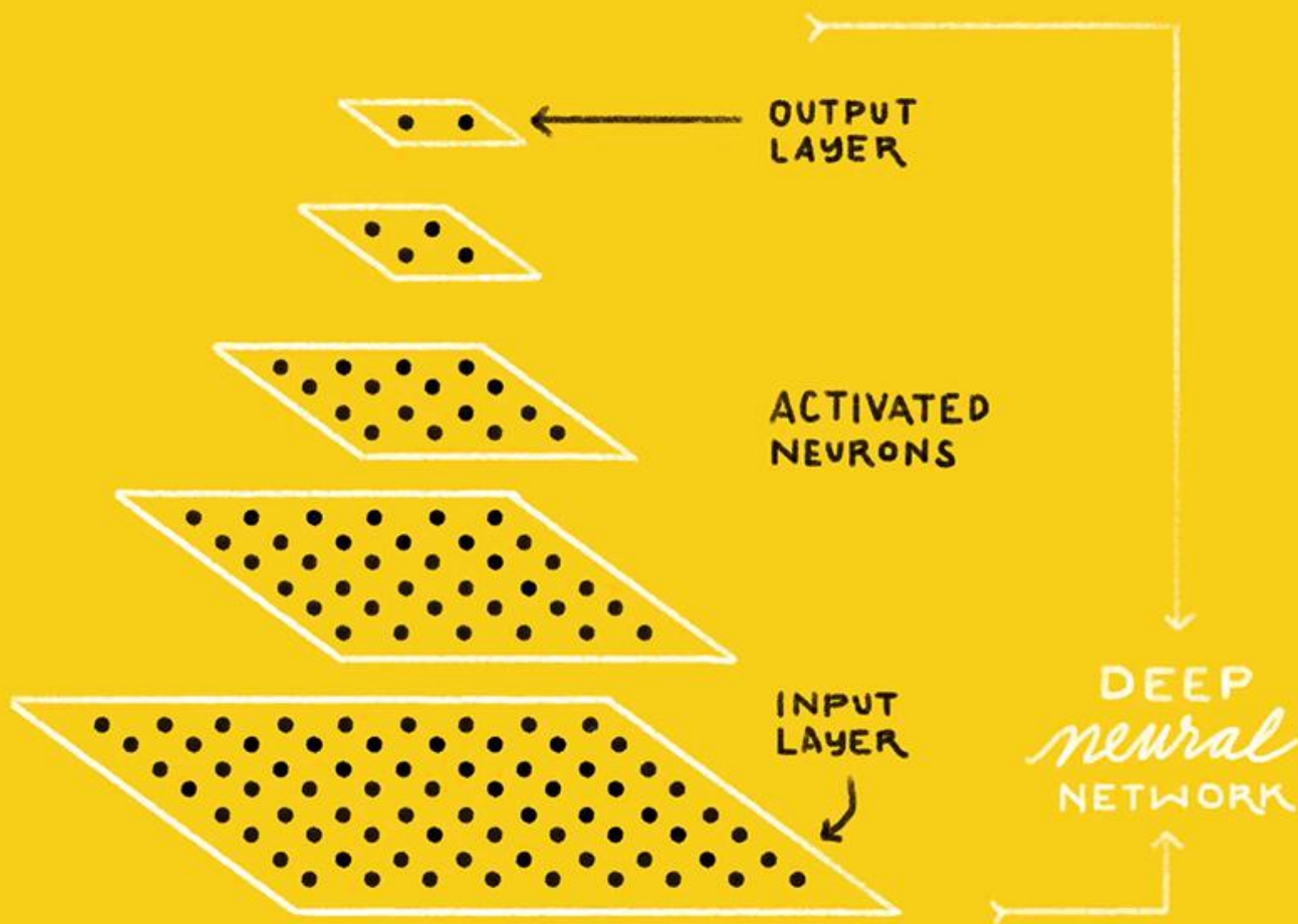


经典卷积神经网络结构

IS THIS A
CAT or **DOG**?

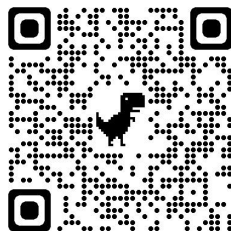
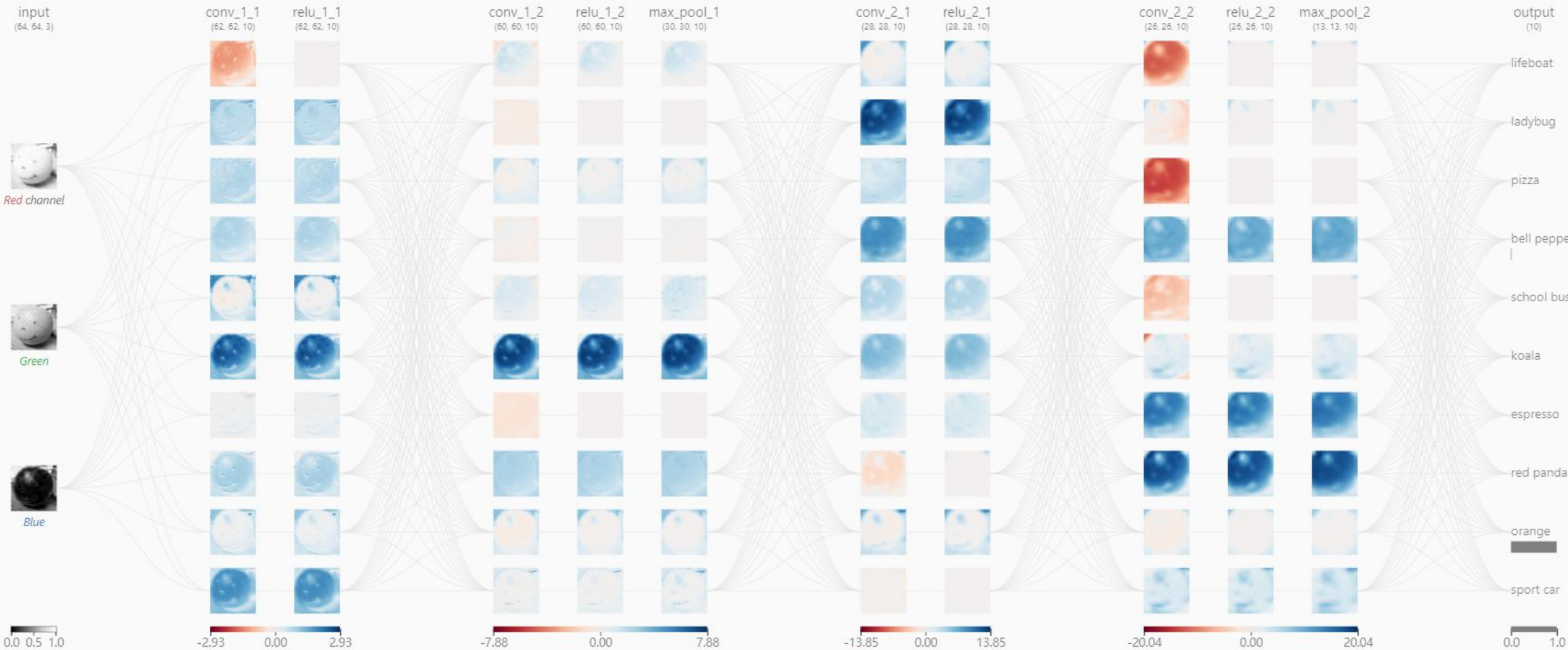


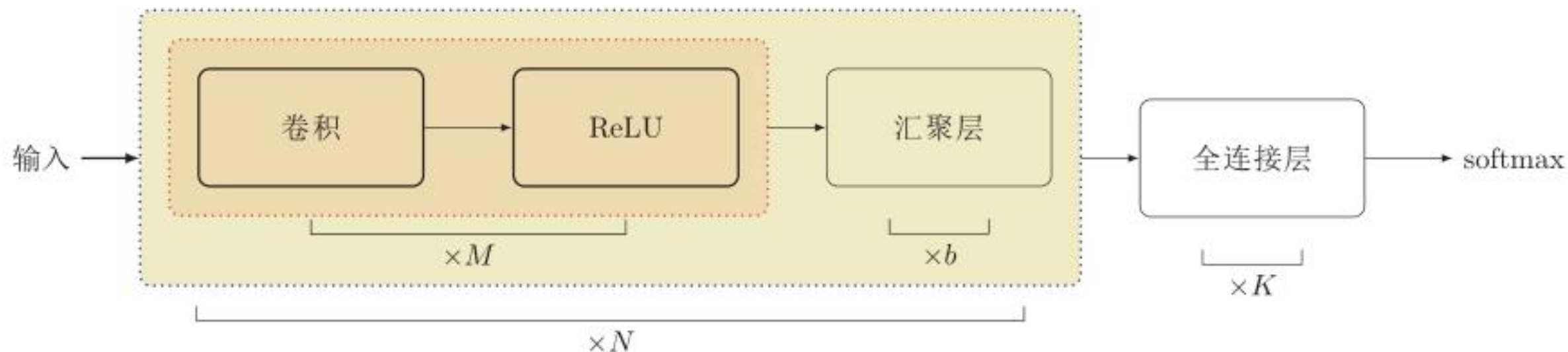
CAT **DOG**





Show detail Unit

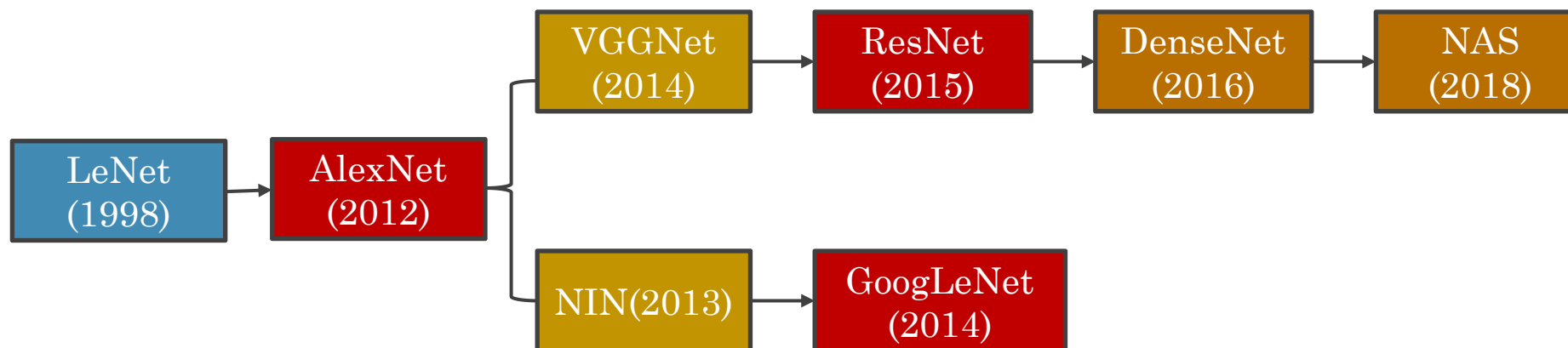




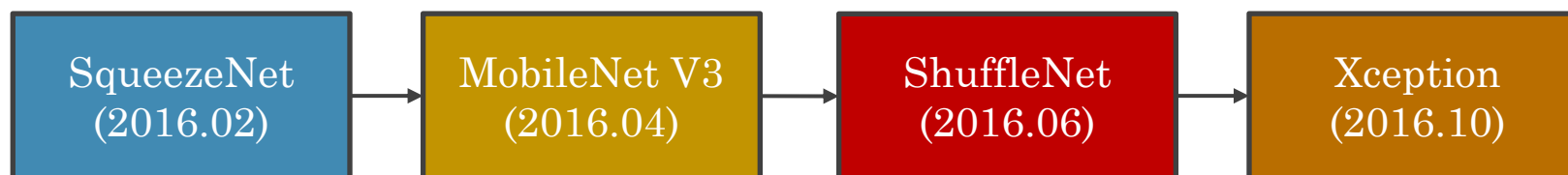
- 一个卷积块为连续 M 个卷积层和 b 个汇聚层
 - (M 通常设置为 $2 \sim 5$, b 为 0 或 1)。
- 一个卷积网络中可以堆叠 N 个连续的卷积块, 然后在接着 K 个全连接层
 - (N 的取值区间比较大, 比如 $1 \sim 100$ 或者更大; K 一般为 $0 \sim 2$)。

典型的卷积网络

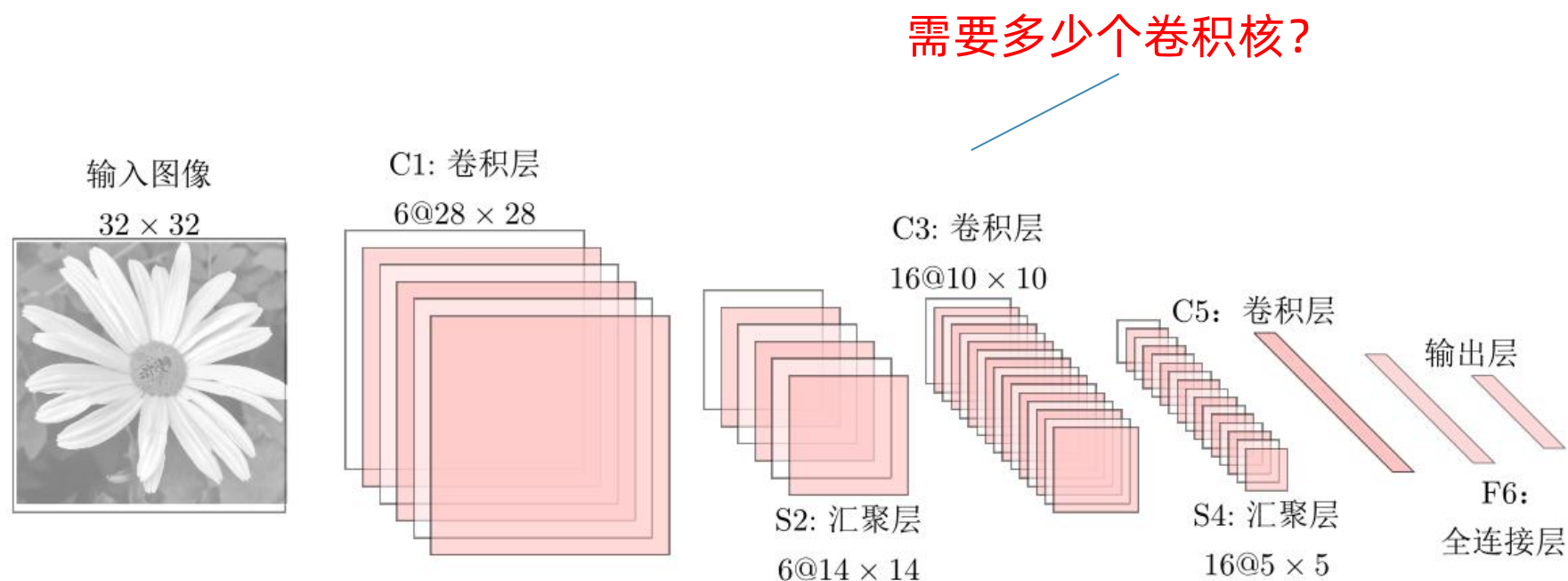
- CNN模型结构变迁



- 针对移动端的轻量型网络



- LeNet-5 是一个非常成功的神经网络模型。
 - 基于 LeNet-5 的手写数字识别系统在 90 年代被美国很多银行使用，用来识别支票上面的手写数字。
 - LeNet-5 共有 7 层。



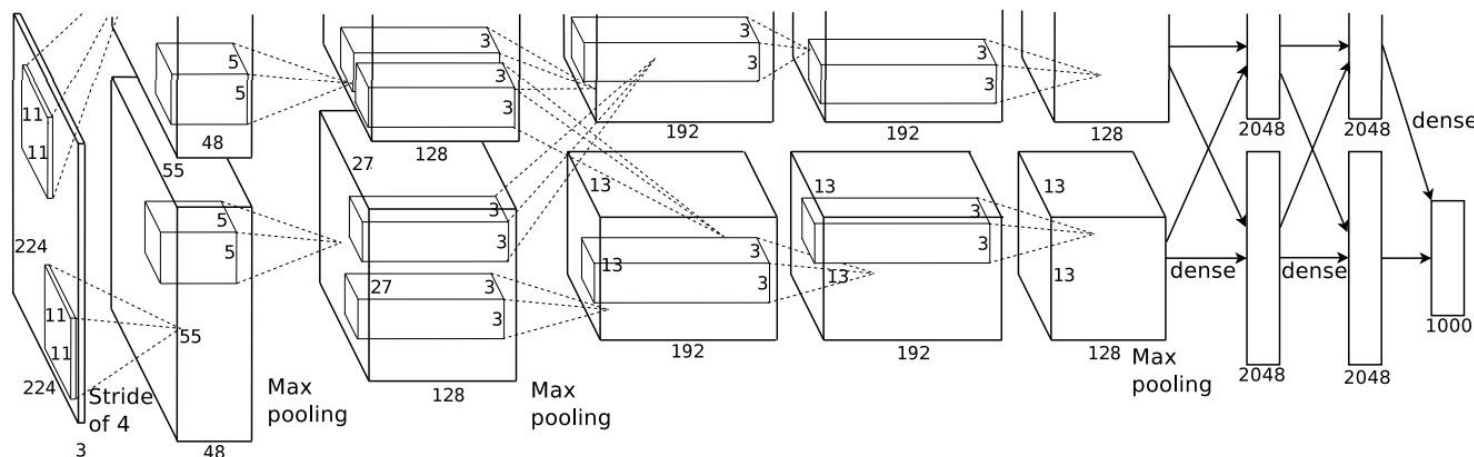


<https://atcold.github.io/pytorch-Deep-Learning/zh/week03/03-2/>



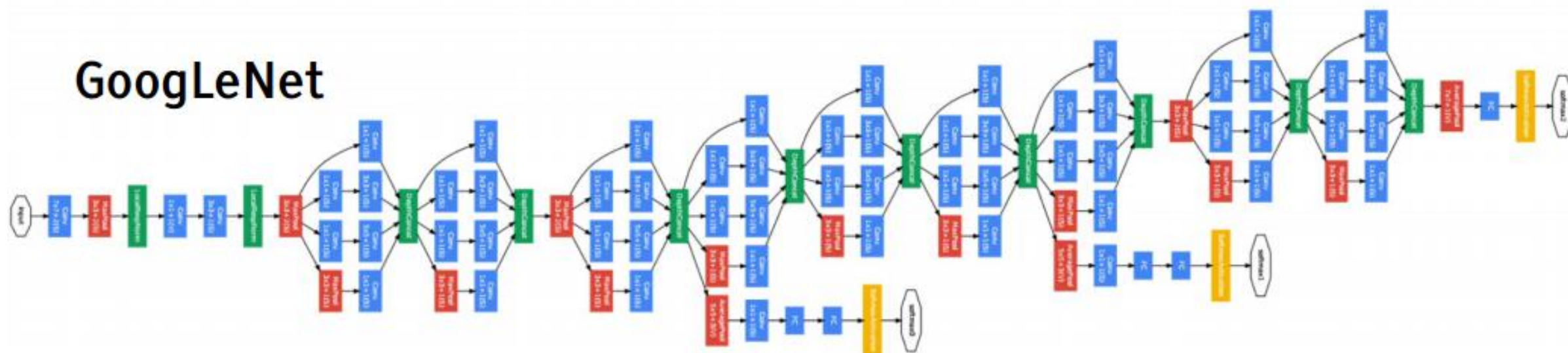
- 2012 ILSVRC winner

- (top 5 error of 16% compared to runner-up with 26% error)
- 第一个现代深度卷积网络模型，首次使用了很多现代深度卷积网络的一些技术方法，
 - 比如使用GPU进行并行训练，采用了ReLU作为非线性激活函数，使用Dropout防止过拟合，使用数据增强
- 5个卷积层、3个汇聚层和3个全连接层



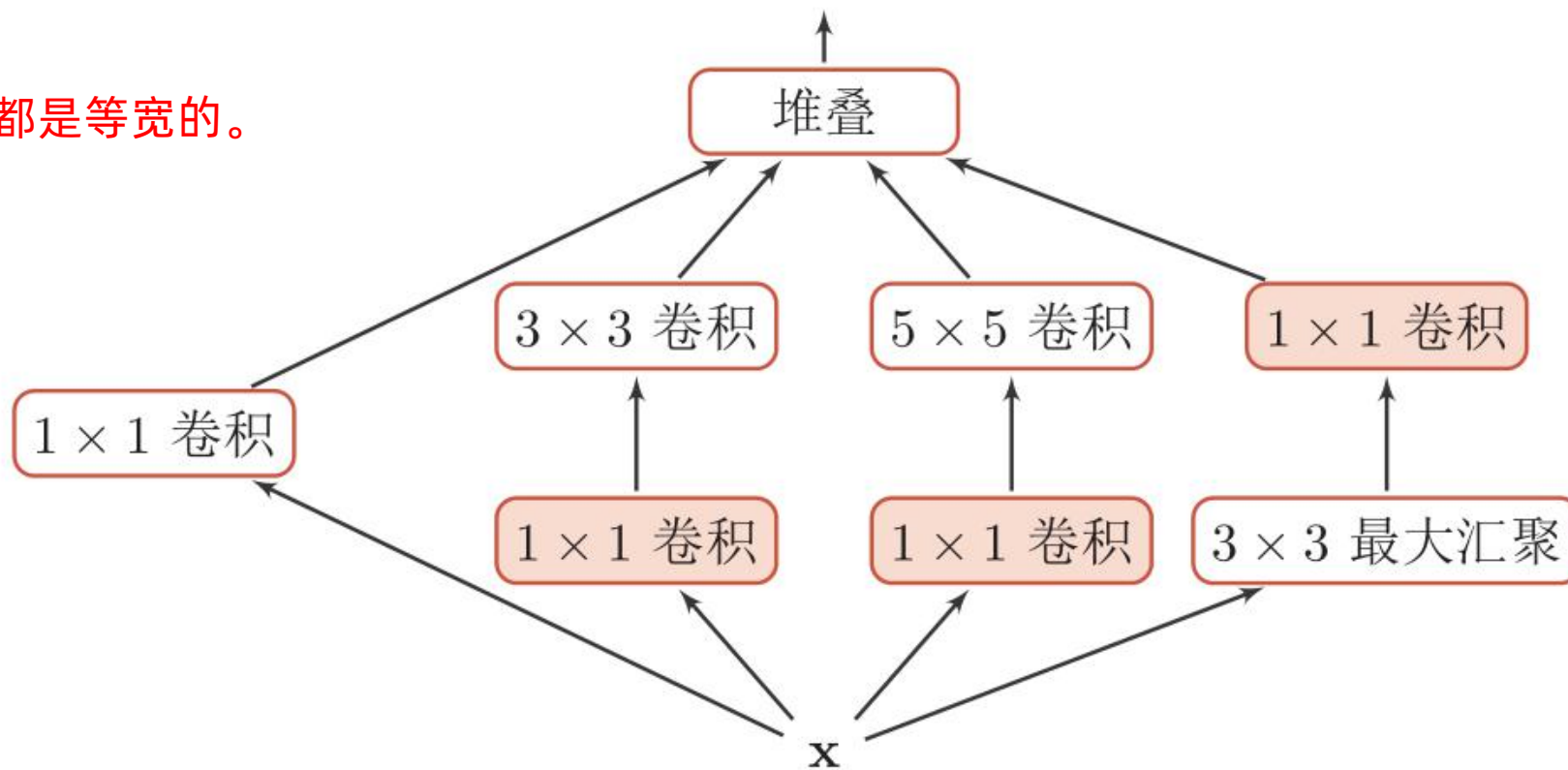
- 2014 ILSVRC winner (22层)

- 参数：GoogLeNet: 4M VS AlexNet: 60M
- 错误率：6.7%
- Inception网络是由有多个inception模块和少量的汇聚层堆叠而成。



- 在Inception网络中，一个卷积层包含多个不同大小的卷积操作，称为Inception模块。
 - Inception模块同时使用 1×1 、 3×3 、 5×5 等不同大小的卷积核，并将得到的特征映射在深度上拼接（堆叠）起来作为输出特征映射。

卷积和最大汇聚都是等宽的。



- 用多层的小卷积核来替换大的卷积核，以减少计算量和参数量。
 - 使用两层 3×3 的卷积来替换v1中的 5×5 的卷积
 - 使用连续的 $n \times 1$ 和 $1 \times n$ 来替换 $n \times n$ 的卷积。

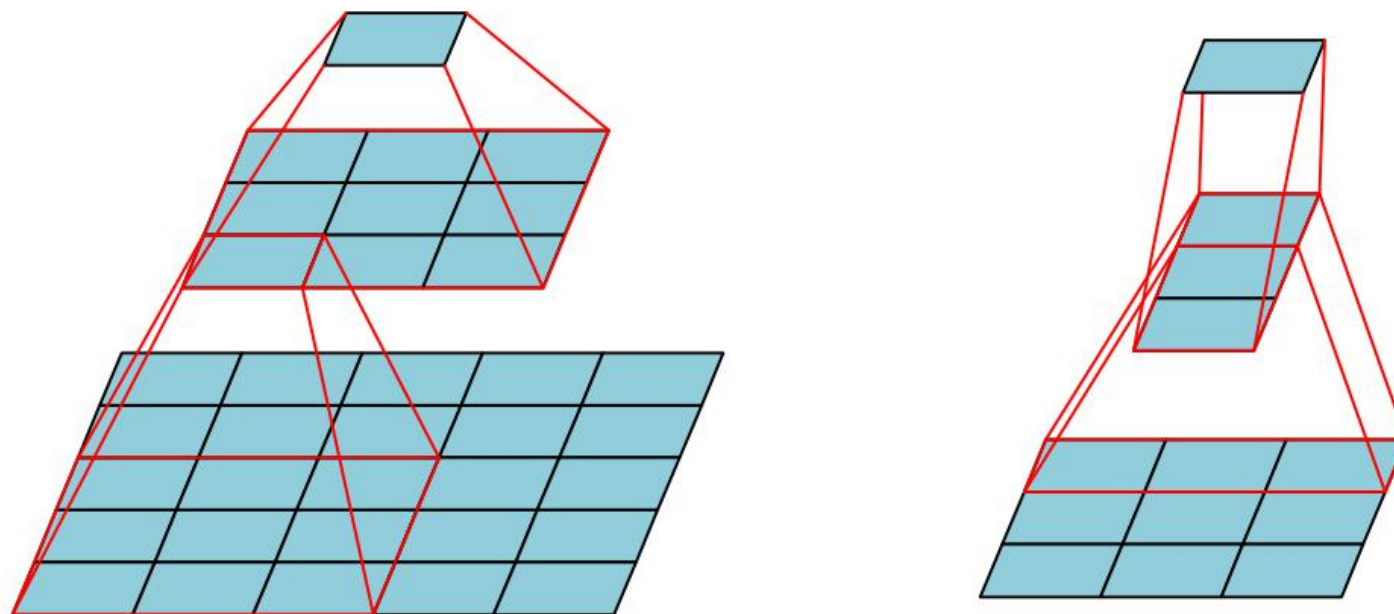


图 1 左图为 5×5 卷积分解为两层 3×3 卷积,右图为 3×3 卷积分解为连续的 1×3 和 3×1 的卷积

- 用多层的小卷积核来替换大的卷积核，以减少计算量和参数量。
 - 使用两层3x3的卷积来替换v1中的5x5的卷积
 - 使用连续的nx1和1xn来替换nxn的卷积。

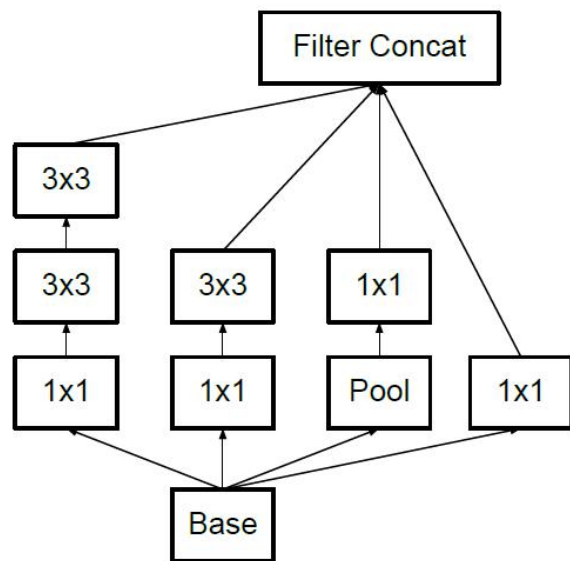


Figure 5. Inception modules where each 5×5 convolution is replaced by two 3×3 convolution, as suggested by principle [3] of Section [2].

<http://blog.csdn.net/xbinworld>

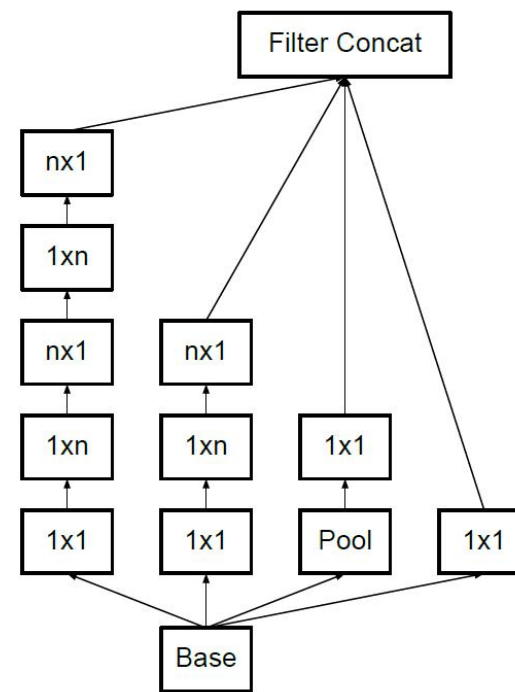


Figure 6. Inception modules after the factorization of the $n \times n$ convolutions. In our proposed architecture, we chose $n = 7$ for the 17×17 grid. (The filter sizes are picked using principle [3])

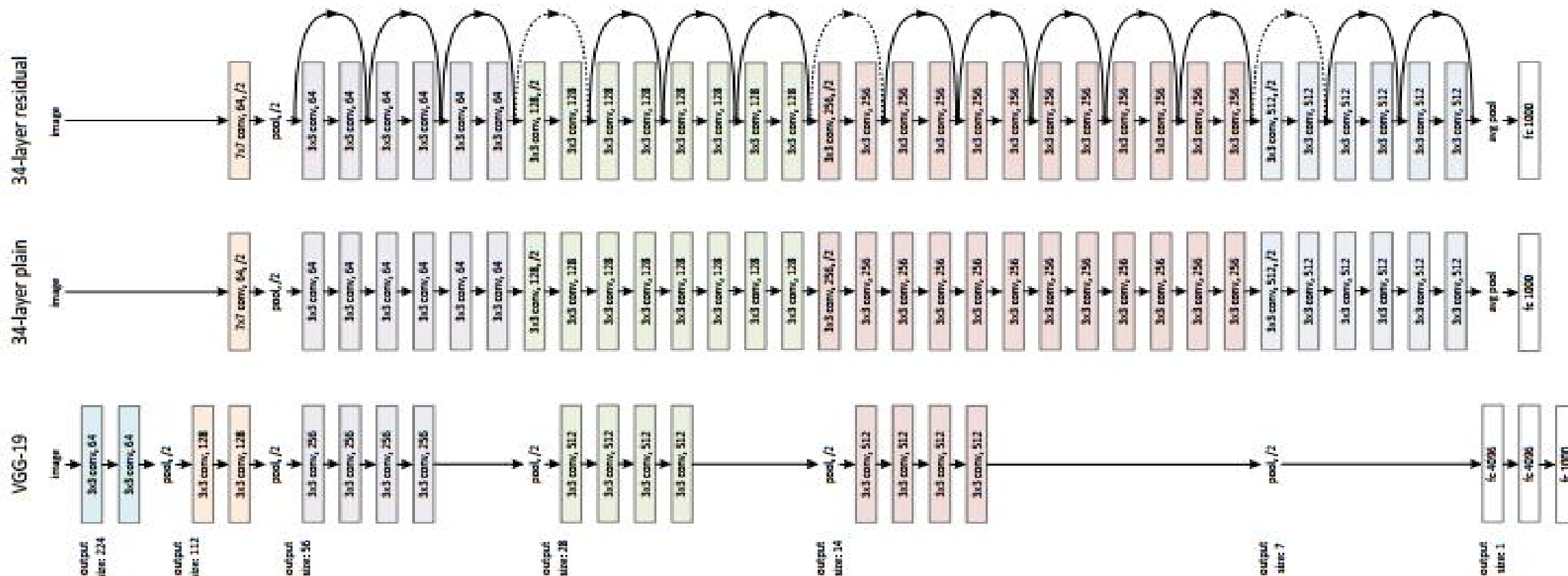
<http://blog.csdn.net/xbinworld>

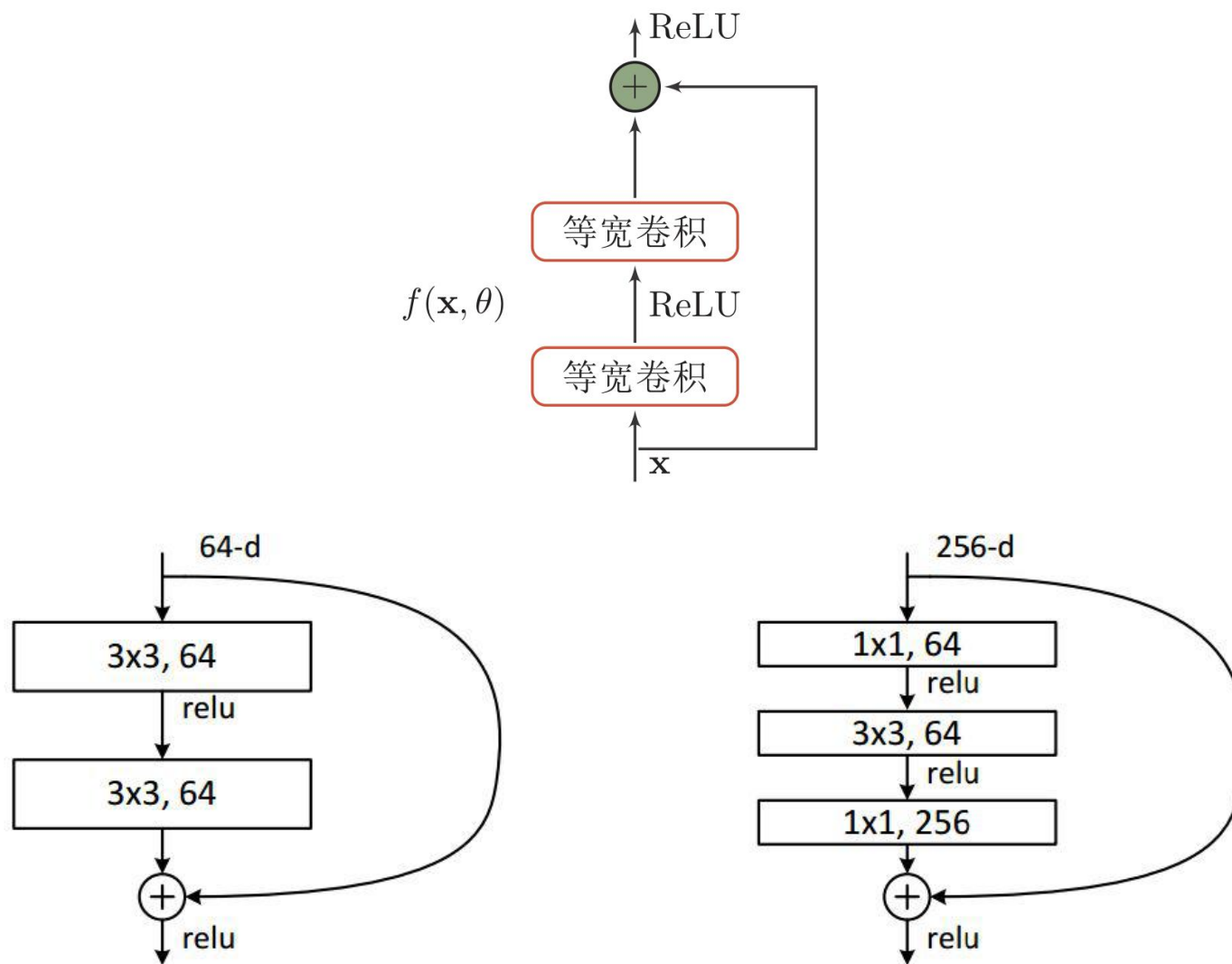
- 残差网络（Residual Network, ResNet）是通过给非线性的卷积层增加直连边的方式来提高信息的传播效率。
 - 假设在一个深度网络中，我们期望一个非线性单元（可以为一层或多层的卷积层） $f(x, \theta)$ 去逼近一个目标函数为 $h(x)$ 。
- 将目标函数拆分成两部分：恒等函数和残差函数

$$h(\mathbf{x}) = \underbrace{\mathbf{x}}_{\text{恒等函数}} + \underbrace{(h(\mathbf{x}) - \mathbf{x})}_{\text{残差函数}} \rightarrow f(\mathbf{x}, \theta)$$

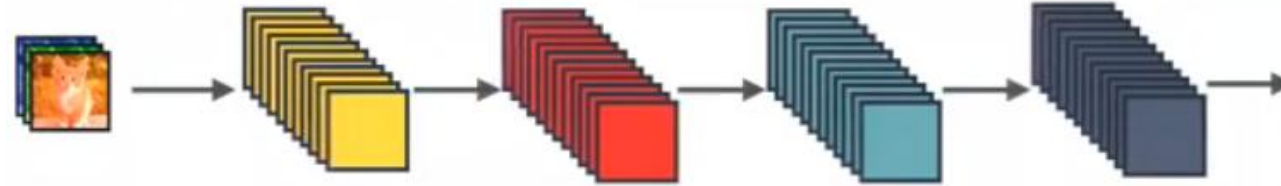
- 2015 ILSVRC winner (152层)

- 错误率: 3.57%

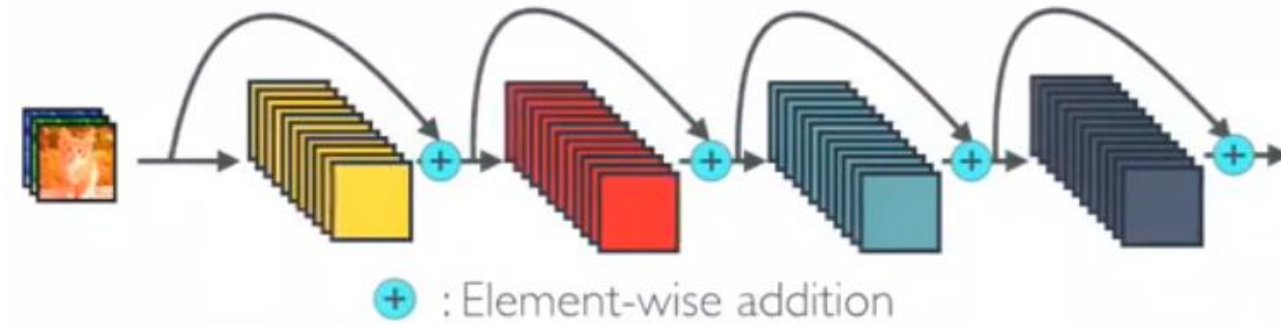




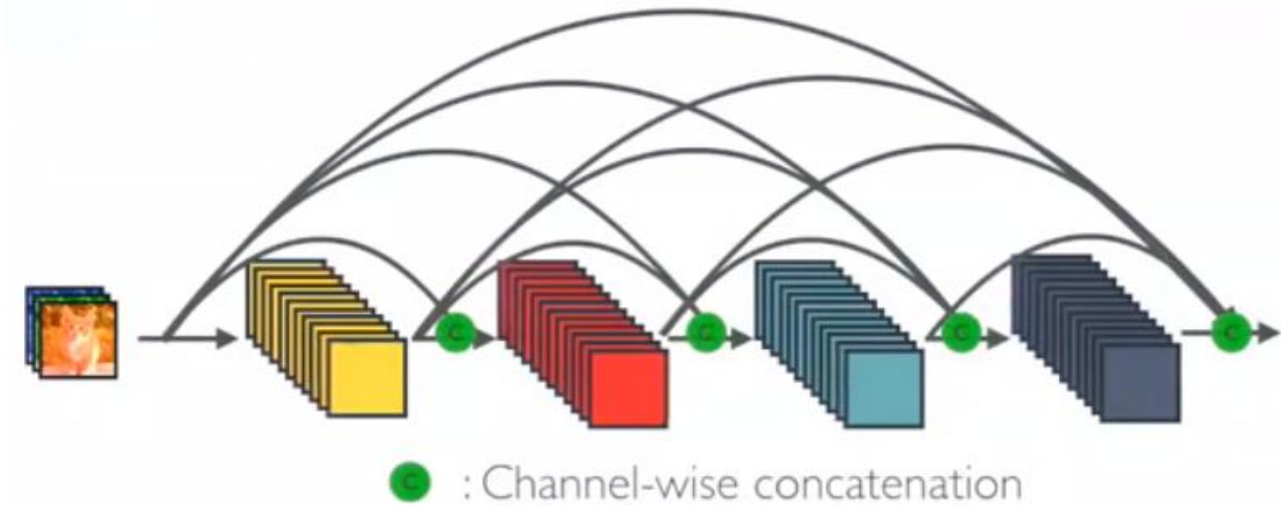
Standard CNN

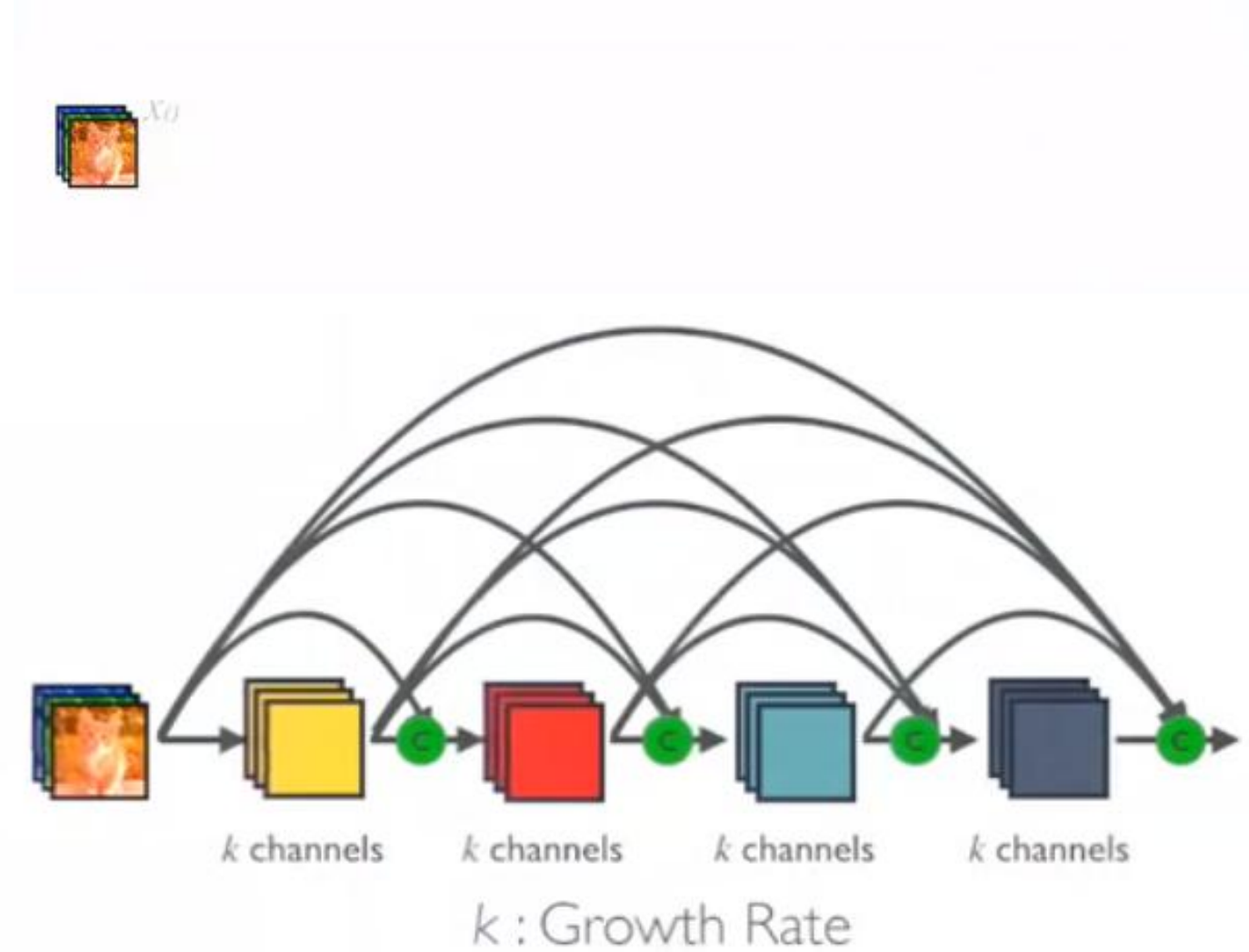


ResNet



DenseNet





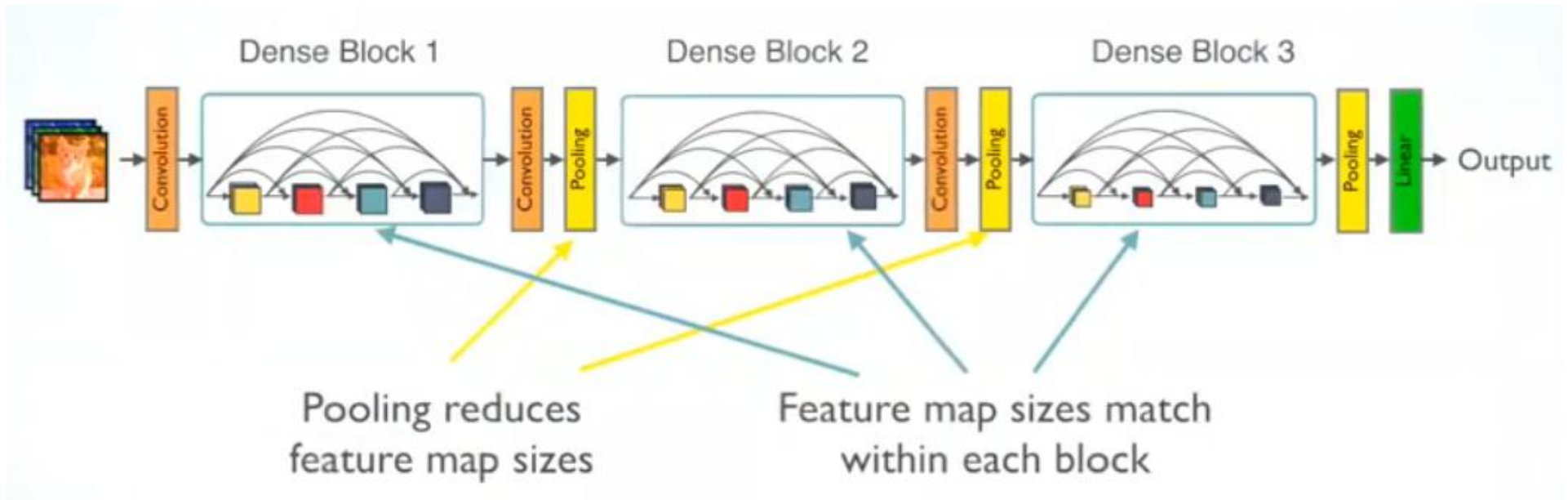


表 2 常用卷积神经网络总结
Table 2 Summary of popular DNNs

方法	top-1 精度/%	top-5 精度/%	输入分辨率/像素	卷积核大小	卷积层深度	参数量
AlexNet	56.6	80.2	224×224	3,5,11	5	6.14×10^7
VGG-16	70.3	89.4	224×224	3	13	1.38×10^8
GoogLeNet	68.9	89.1	224×224	1,3,5,7	21	7.0×10^6
ResNet-50	75.1	92.3	224×224	1,3,7	49	2.55×10^7
MobileNetV2	72.2	90.5	224×224	3	20	3.4×10^6

注：表中的精度与特定数据集相关联。

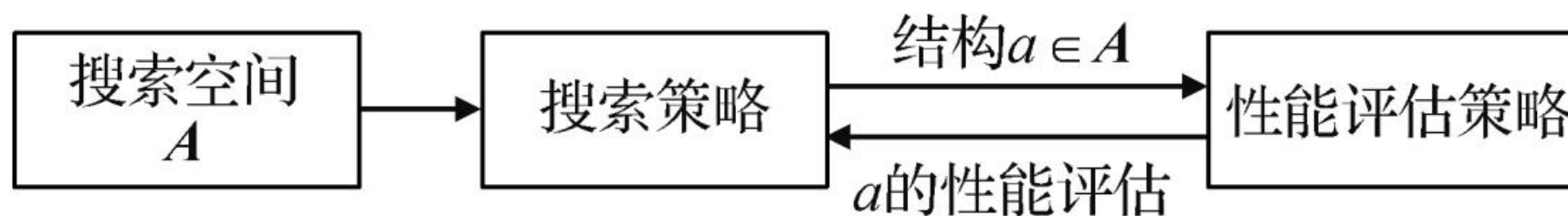


图 1 神经网络结构搜索过程

Fig. 1 Illustration of neural architecture search

论文引用格式: Tang L, Li H X, Yan C Q, Zheng X W and Ji R R. 2021. Survey on neural architecture search. Journal of Image and Graphics, 26(02): 0245-0264 (唐浪, 李慧霞, 颜晨倩, 郑侠武, 纪荣嵘. 2021. 深度神经网络结构搜索综述. 中国图象图形学报, 26(02): 0245-0264) [DOI:10.11834/jig.200202]

表 5 不同 NAS 算法在 ImageNet 数据集的性能

Table 5 Performance of different NAS methods on ImageNet dataset

方法	测试集 Top1 准确率/%	FLOPs/M	搜索时间/ GPU hours
NASNet-A	74	564	48 000
DARTS	73.1	595	96
MnasNet	74	317	40 000
ChamNet-B	73.8	323	28 000
FBNet-C	74.9	375	216
ProxylessNAS	74.6	320	200
SinglePathNAS	74.7	328	312
AutoSlim	74.2	305	180
MobileNetV3-L	75.2	219	–
OFA	76.4	238	40
FBNetV2-F4	76	238	200

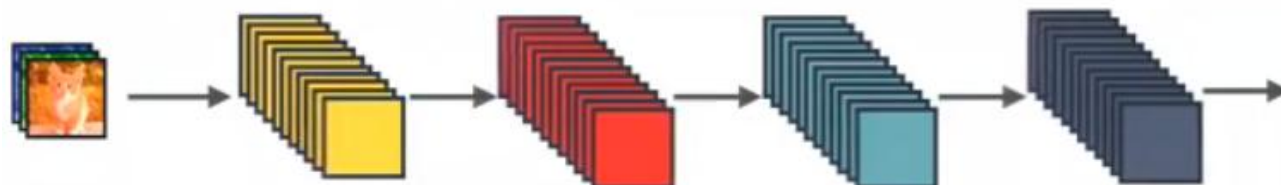
方法	top-1 精度/%
AlexNet	56.6
VGG-16	70.3
GoogLeNet	68.9
ResNet-50	75.1
MobileNetV2	72.2

注：“–”表示未知。

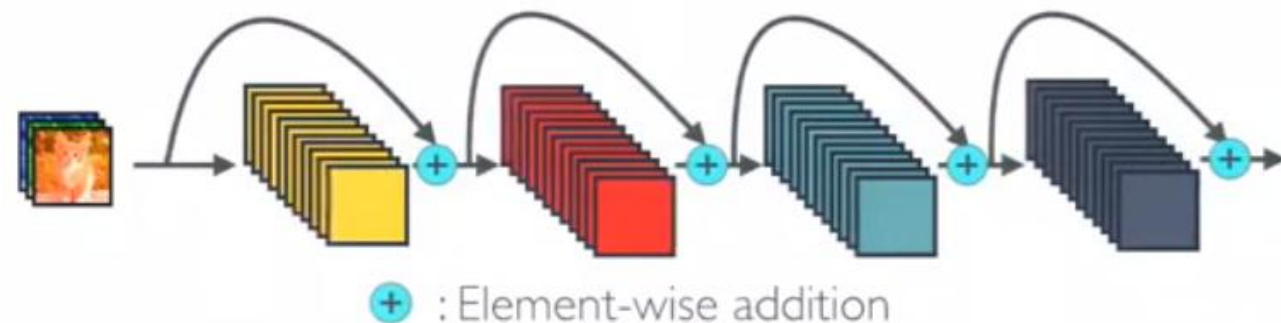
总结

- LeNet,
- AlexNet,
- VGG,
- ResNet,
- DenseNet

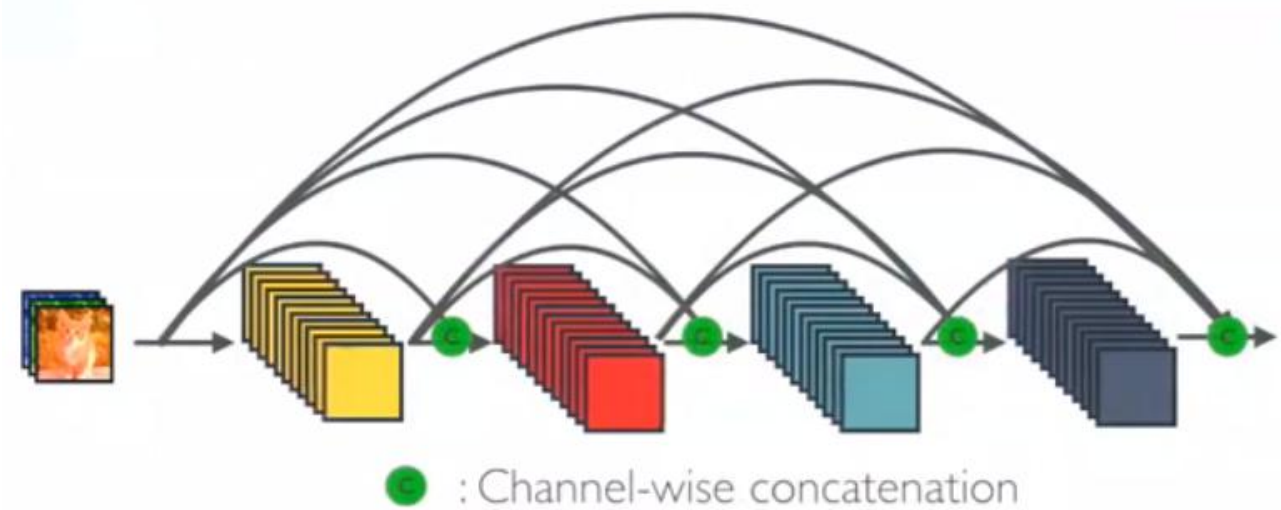
Standard CNN



ResNet



DenseNet





Thank You !

Q & A