机器学习
Machine Learning

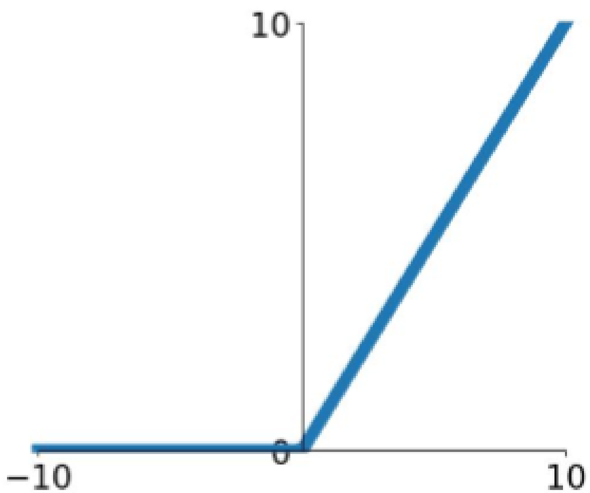
第9讲：神经网络和深度学习-II

Part1：激活函数、CNN

新的激活函数

整流线性激活函数
(Rectified Linear Unit: ReLU

$$z = h(a) = \max\{0, a\}$$



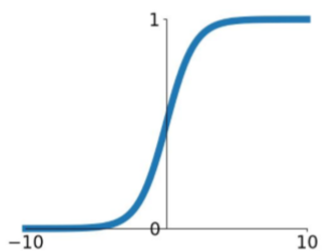
目前深度学习中
最常用的激活函数
~90%

常用激活函数比较

Activation functions

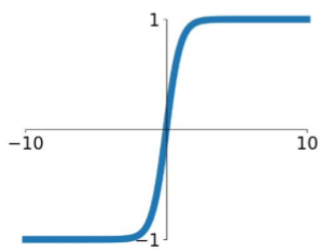
Sigmoid

$\sigma(x) = \frac{1}{1+e^{-x}}$



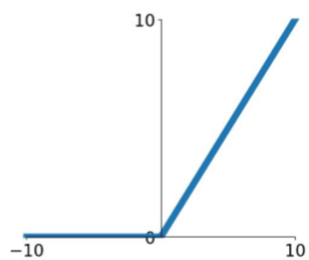
tanh

$\tanh(x)$



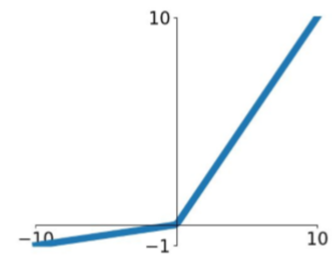
ReLU

$\max(0, x)$



Leaky ReLU

$\max(0.1x, x)$

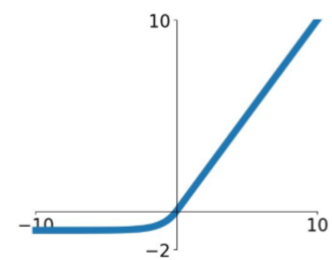


Maxout

$\max(w_1^T x + b_1, w_2^T x + b_2)$

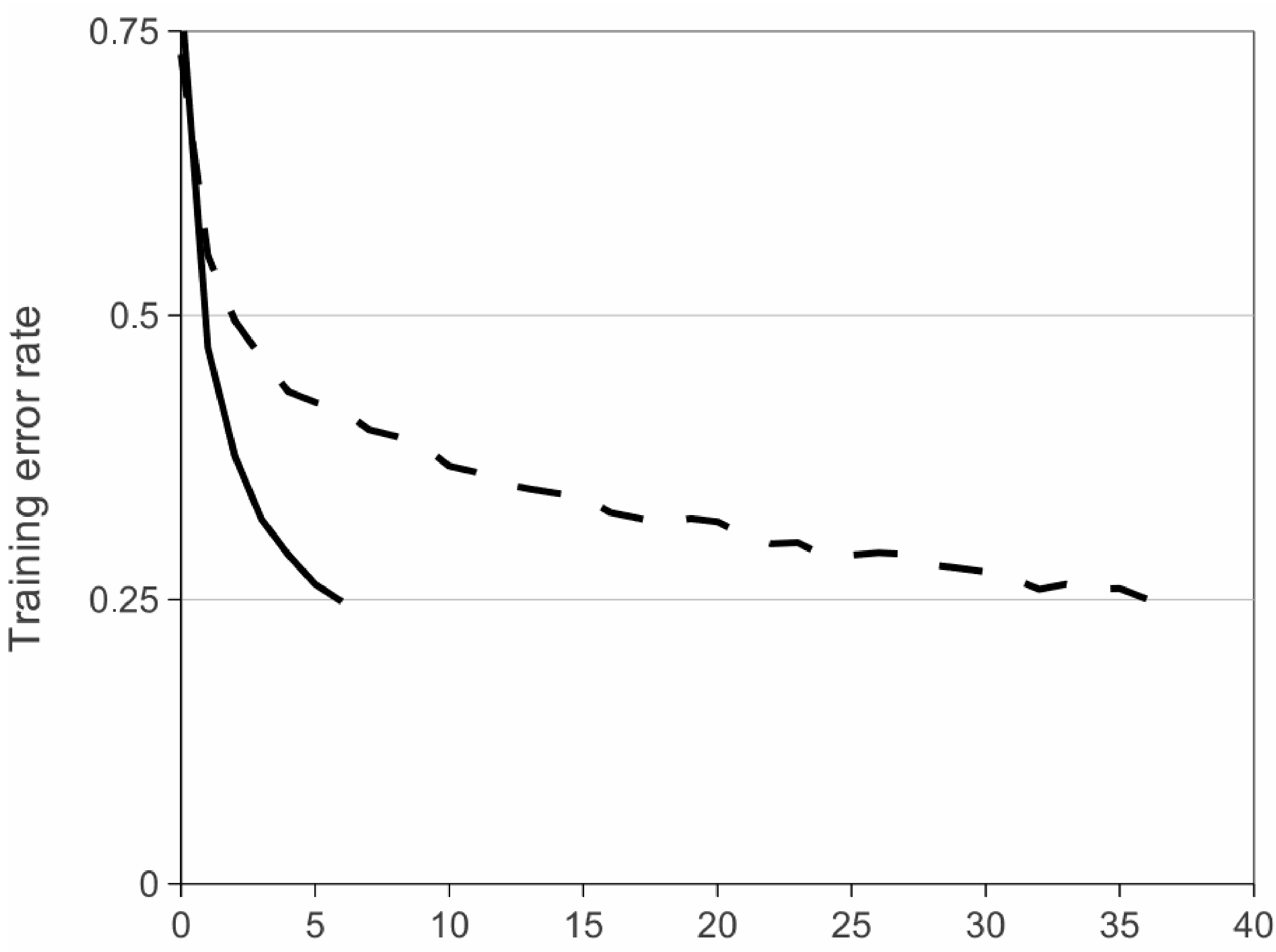
ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



激活函数收敛速度比较

本例：4层网络，ReLU和Tanh激活比较，
达到25%训练误差，ReLU激活快6倍[Krizhevsky]

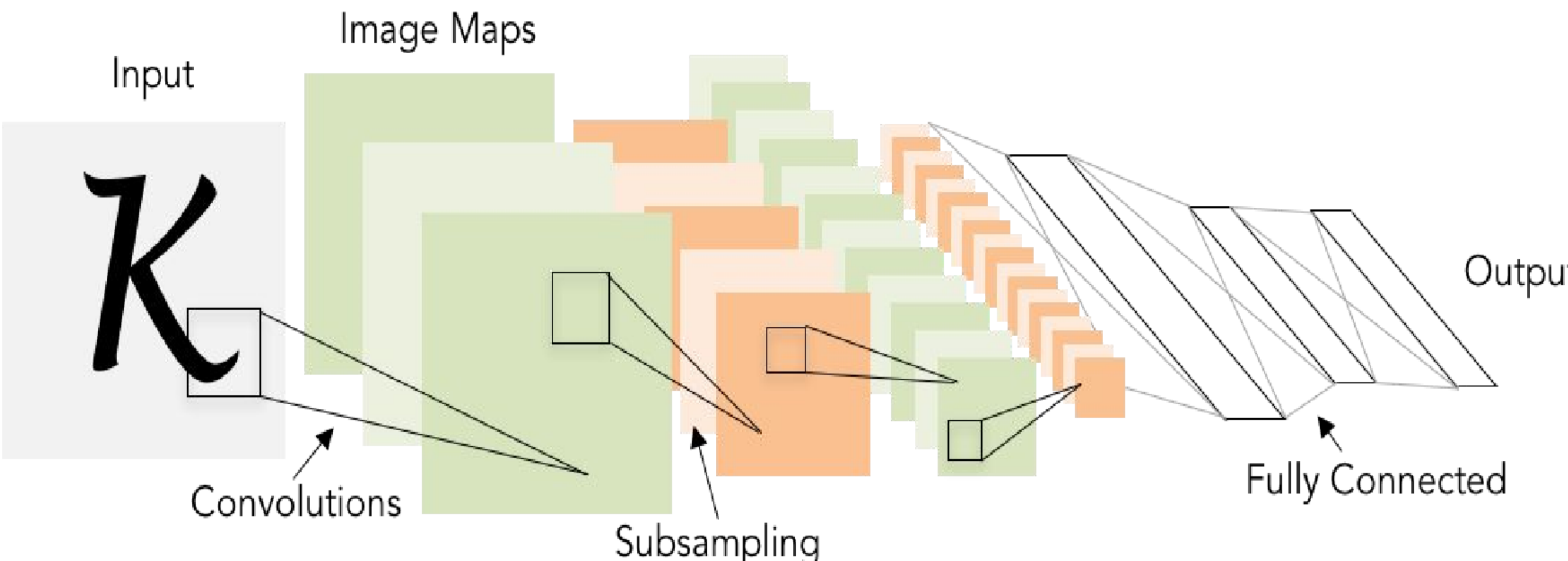


卷积网络（Convolutional Neural Network：CNN）

- 至少在网络的一层中使用卷积运算替代一般的线性全连接加权组合运算。
- CNN的特点：
  - 稀疏连接
  - 参数共享
  - 等变表示（输入平移不变性，LTI系统的基本特性）

一个完整卷积神经网络的示例

Gradient-based learning applied to document recognition [LeCun, Bottou, Bengio, Haffner 1998]

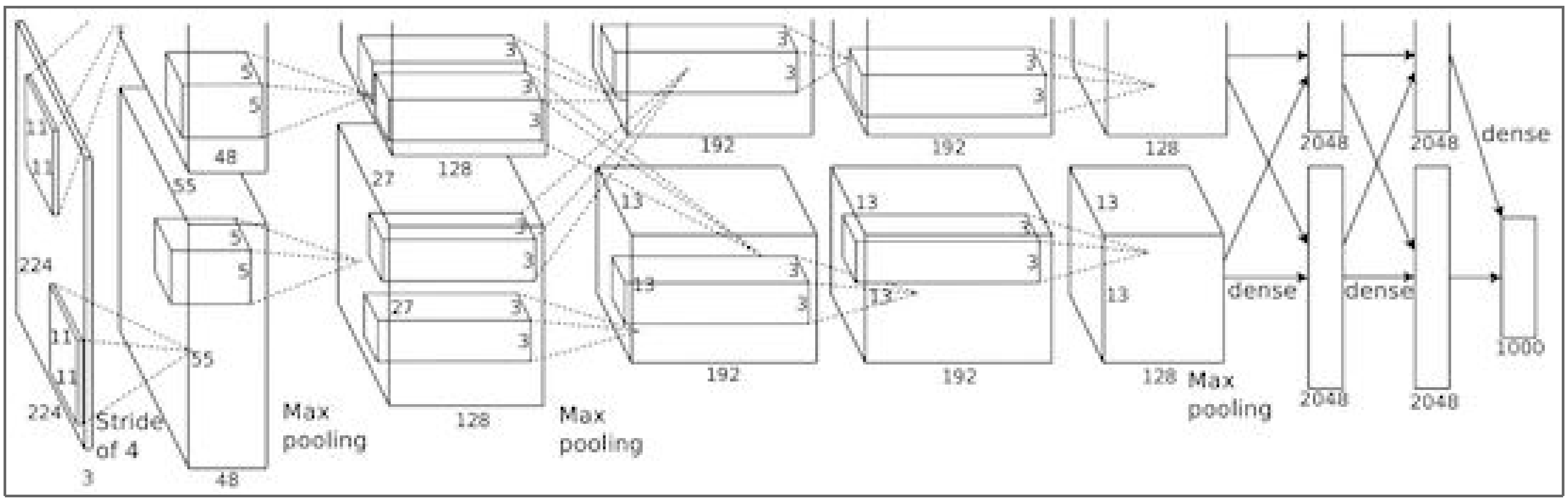


LeNet-5

可以通过多个卷积核（滤波器）将一幅图像产生多个卷积层
卷积网络可以连接全连接网络作为输出层

CNN近期重新得到极大关注的例子

ImageNet Classification with Deep Convolutional Neural Networks [Krizhevsky, Sutskever, Hinton, 2012]



"AlexNet"

卷积运算

连续情况

$$s(t)=\int_{-\infty}^{+\infty}x(\tau)w(t-\tau)d\tau=(x*w)(t)$$

离散情况

$$s(n)=\sum_{m=-\infty}^{+\infty}x(m)w(n-m)$$

$$=\sum_{m=-\infty}^{+\infty}w(m)x(n-m)$$

标准卷积满足可交换性

神经网络中卷积核 w 有限长，可写成

$$s(n)=\sum_{m=m_1}^{m_2}w(m)x(n-m)=\sum_{m}w(m)x(n-m)$$

卷积运算（续）

二维卷积（图像最常用）

$$s(i, j) = \sum_m \sum_n I(m, n) K(i-m, j-n)$$

$$= \sum_m \sum_n I(i-m, j-n) K(m, n)$$

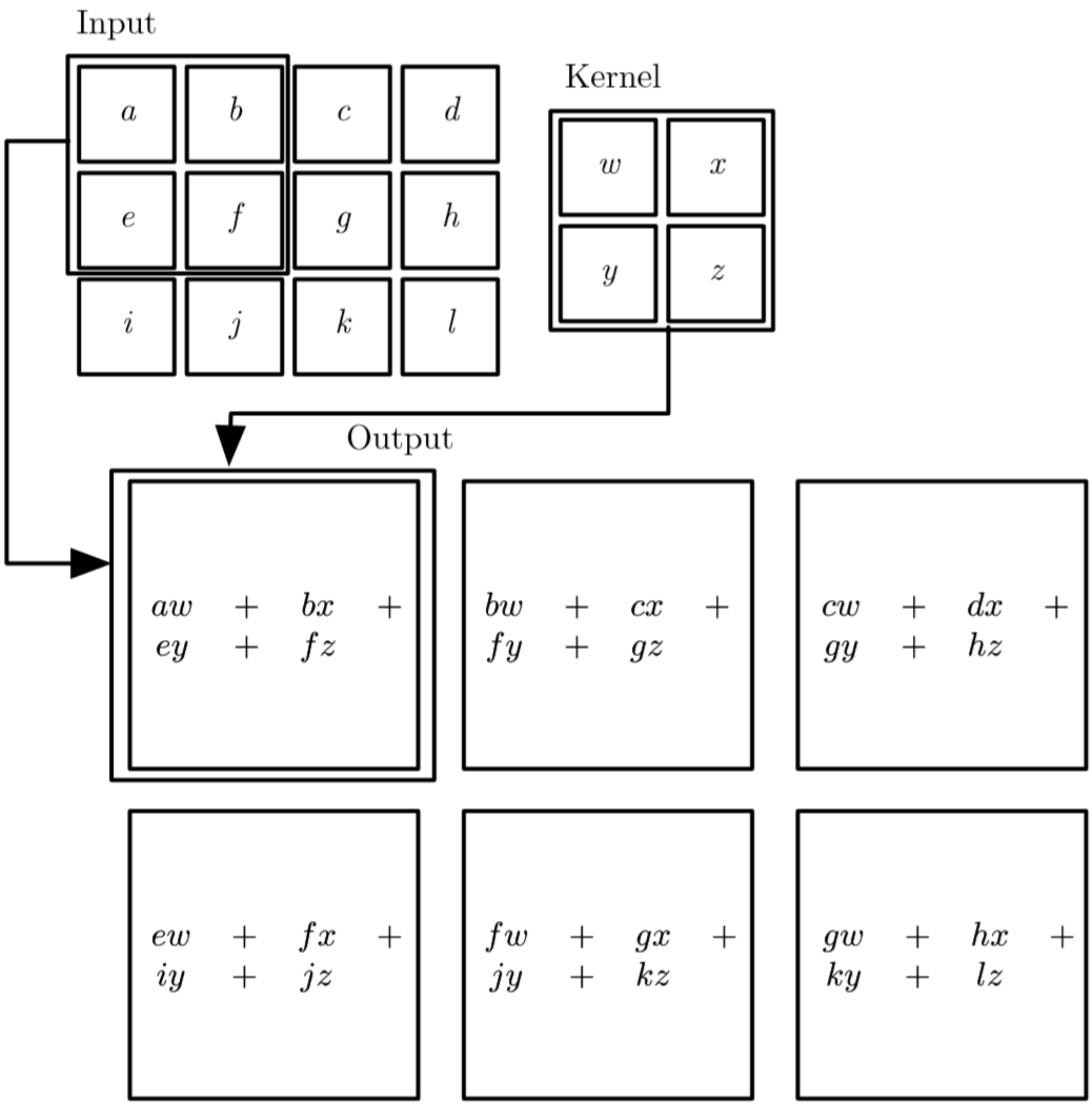
<u>神经网络中常适用一种变体的卷积，为</u>

$$s(n) = (w * s)(n) = \sum_m x(n+m) w(m)$$

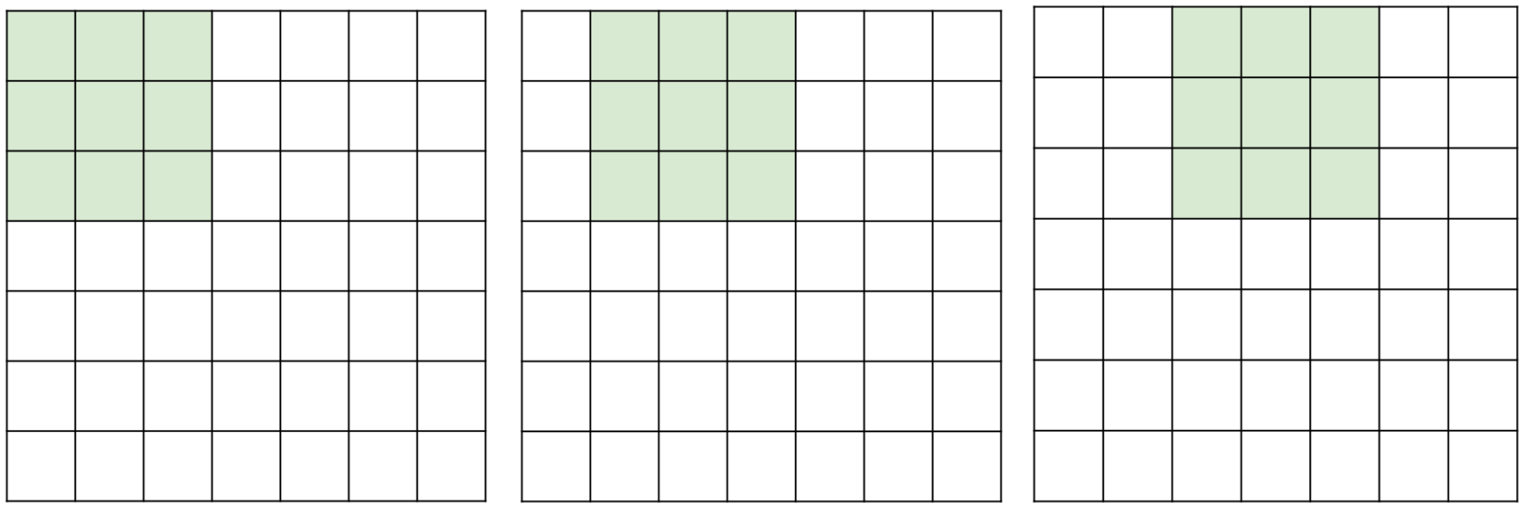
$$s(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i+m, j+n) K(m, n)$$

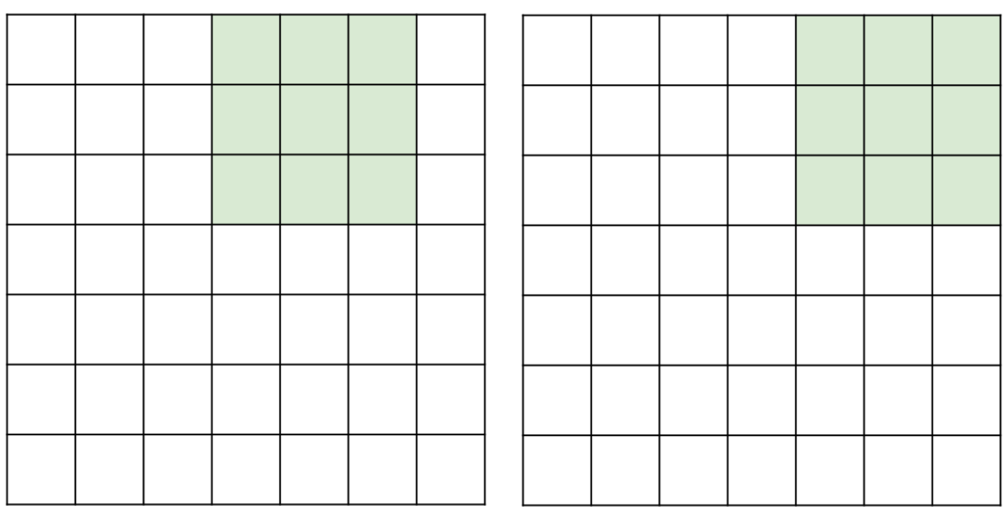
即将标准卷积一方的“翻转”运算取消，实际是一种互相关

二维卷积运算示意图


Input
a b c d
e f g h
i j k l
Kernel
w x
y z
Output
aw + bx + ey + fz
bw + cx + fy + gz
cw + dx + gy + hz
ew + fx + iy + jz
fw + gx + jy + kz
gw + hx + ky + lz

卷积运行（第一行）示例（有效卷积）



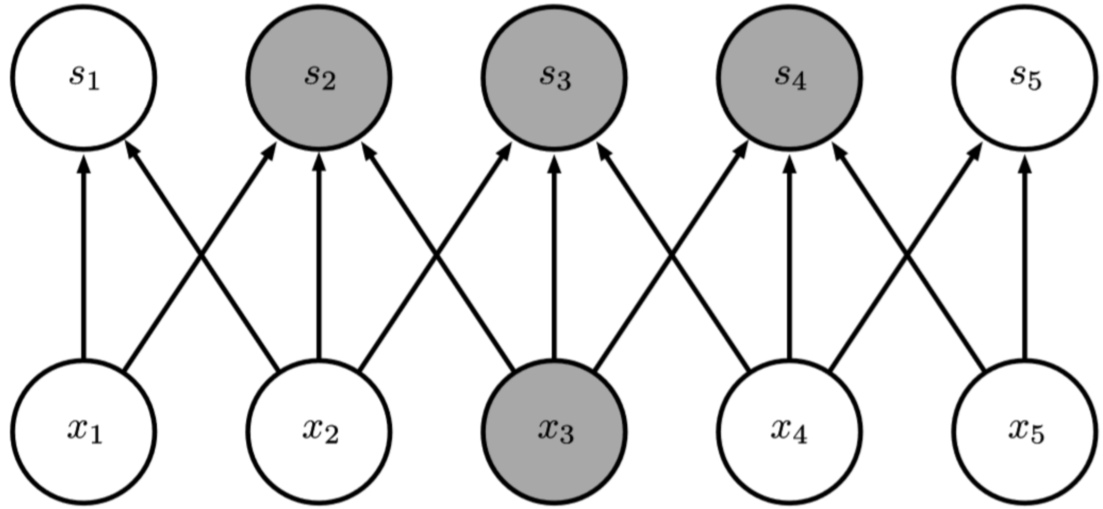


说明：
一个7x7输入图像，
一个3x3滤波器核，
进行有效卷积，
输出5x5卷积层图像。
一般MxM输入，KxK核，
卷积输出：
（M-K+1）x（M-K+1）

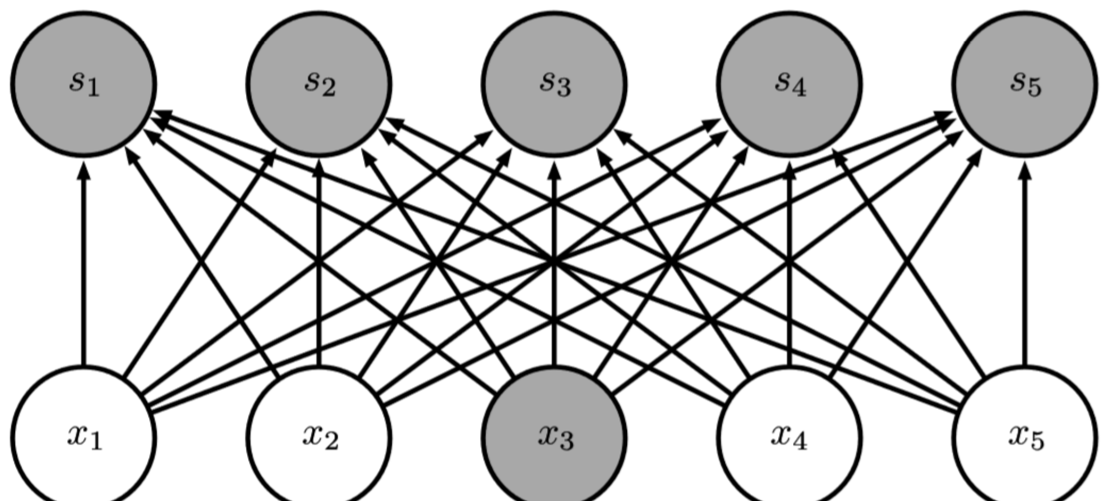
稀疏连接示意：从下向上看

CNN网络：权系数稀疏性

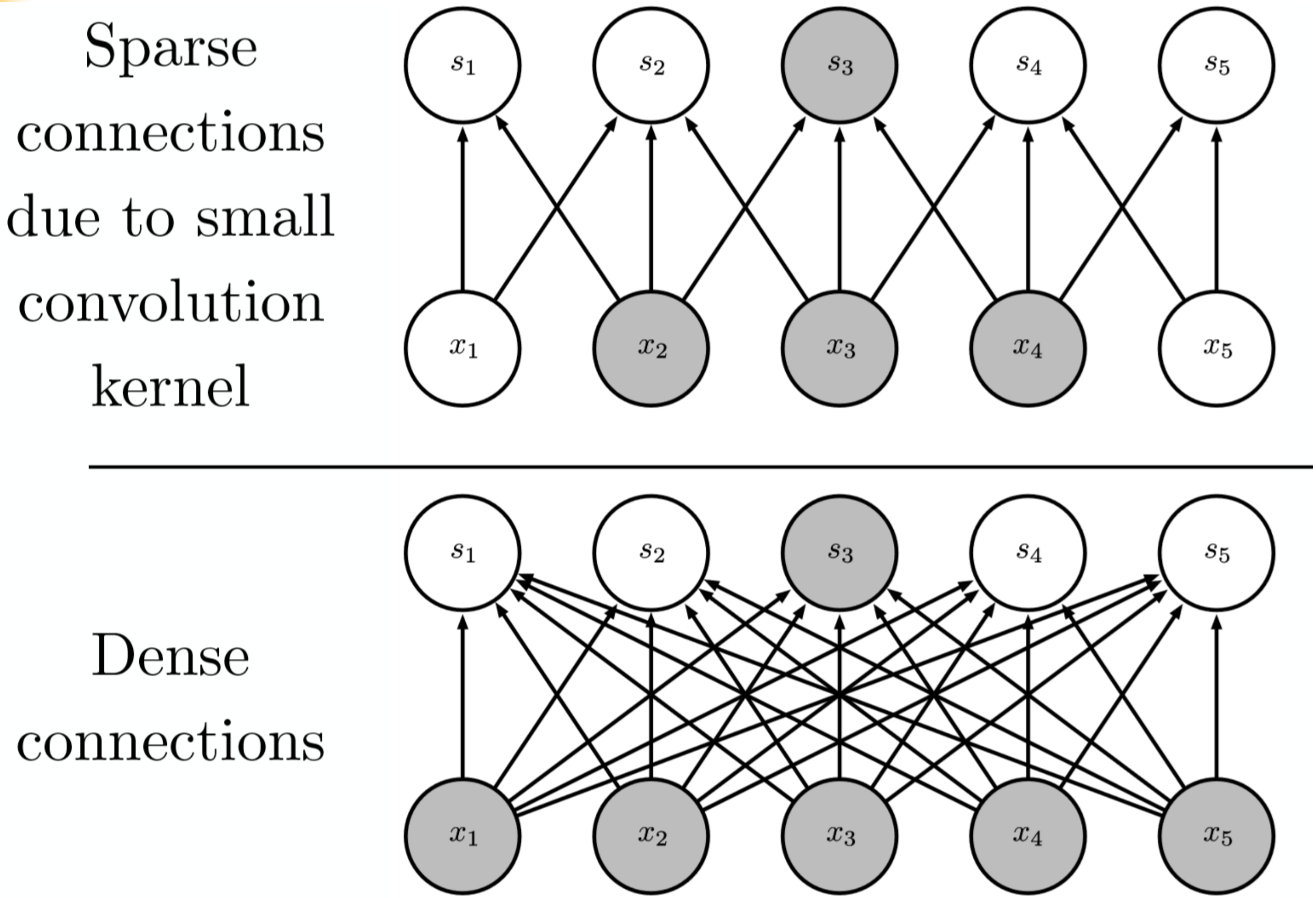
Sparse connections due to small convolution kernel



Dense connections

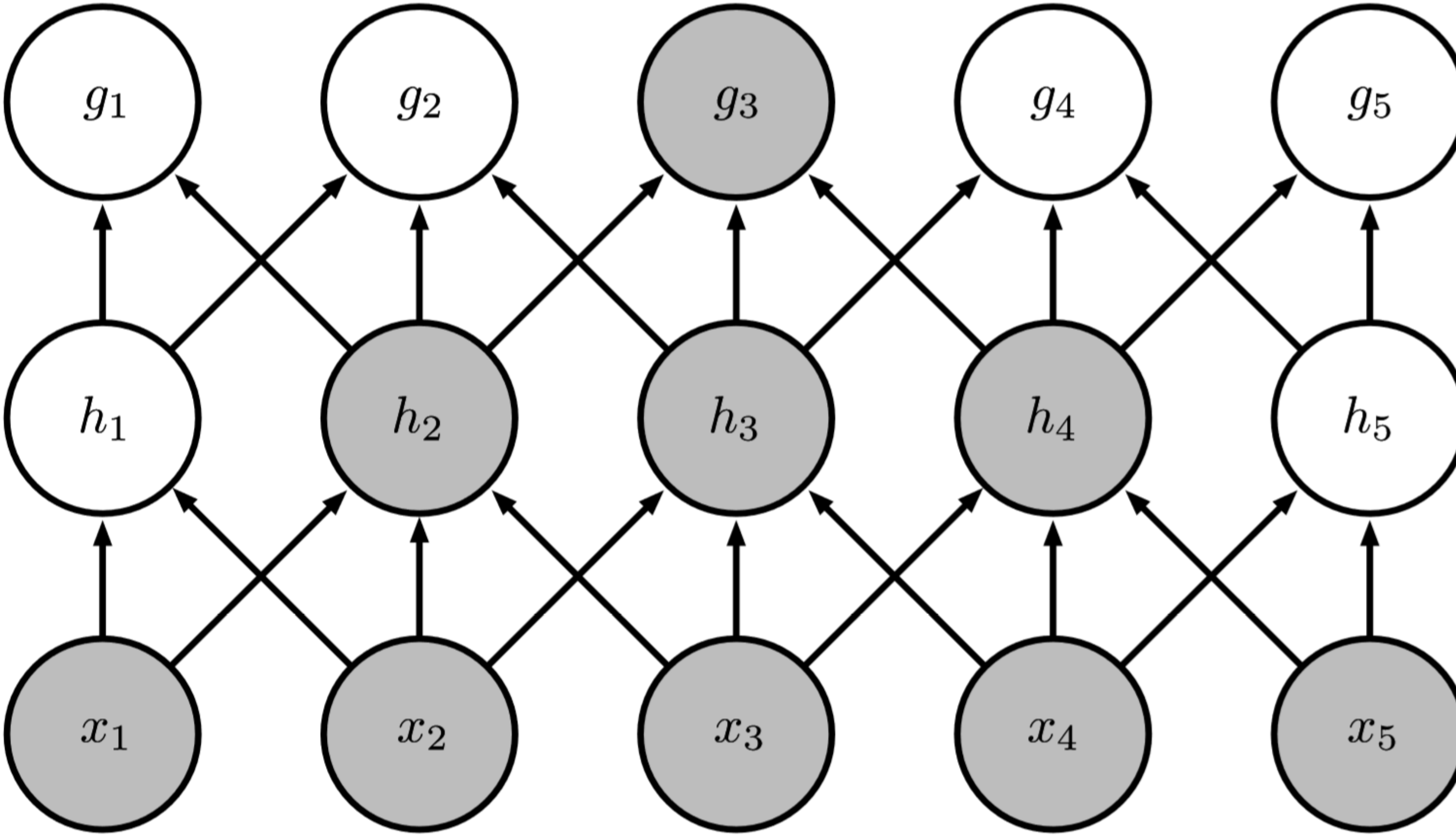


稀疏连接示意：从上向下看



随层数增加，一个输出对应的接受域增加

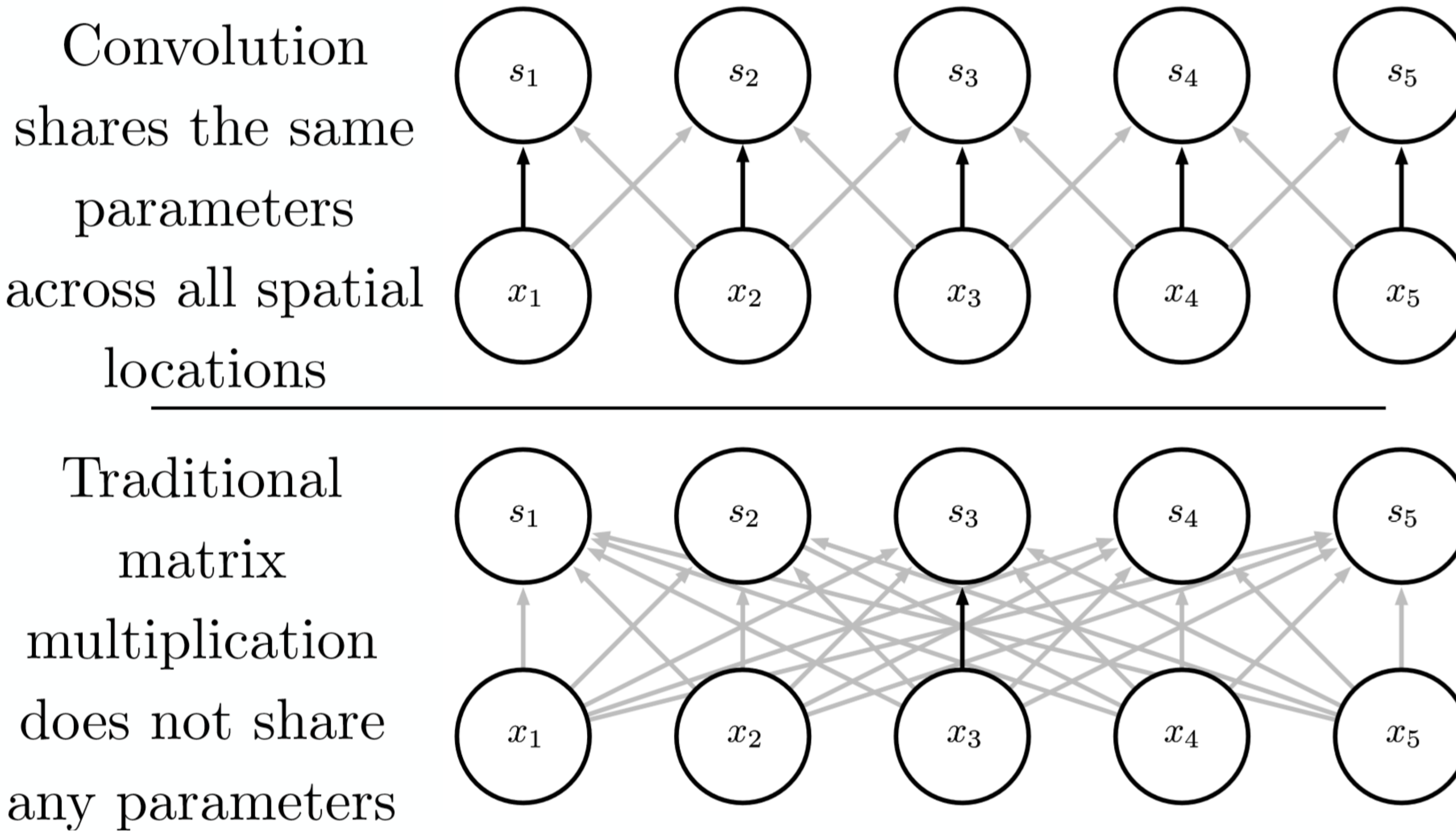
尽管单层接受域有限，但多层具有感受域传播能力



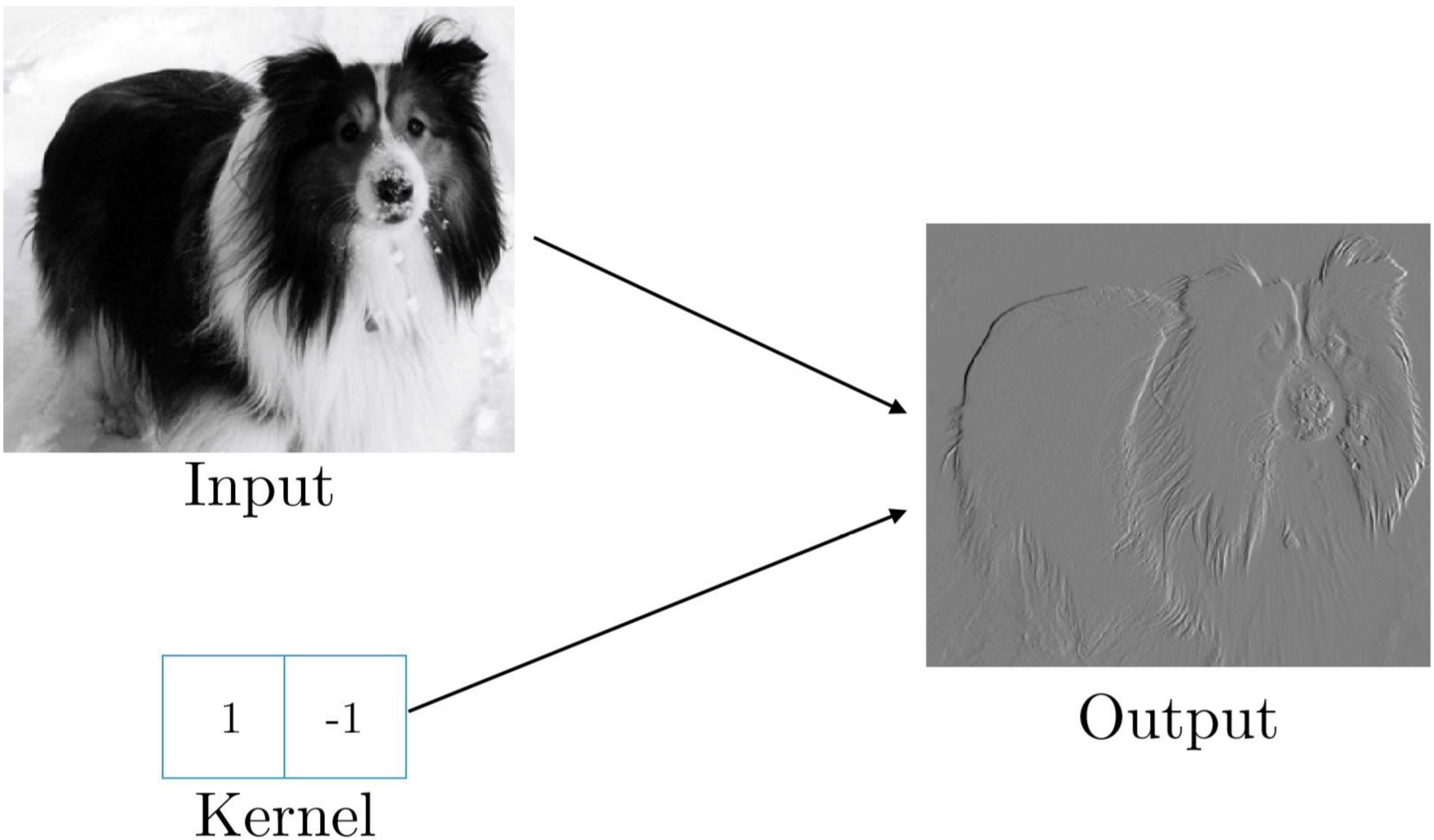
参数共享

基本CNN网络，一层使用一个卷积核共用

卷积核参数数目K，也是CNN一层共用的权系数数目

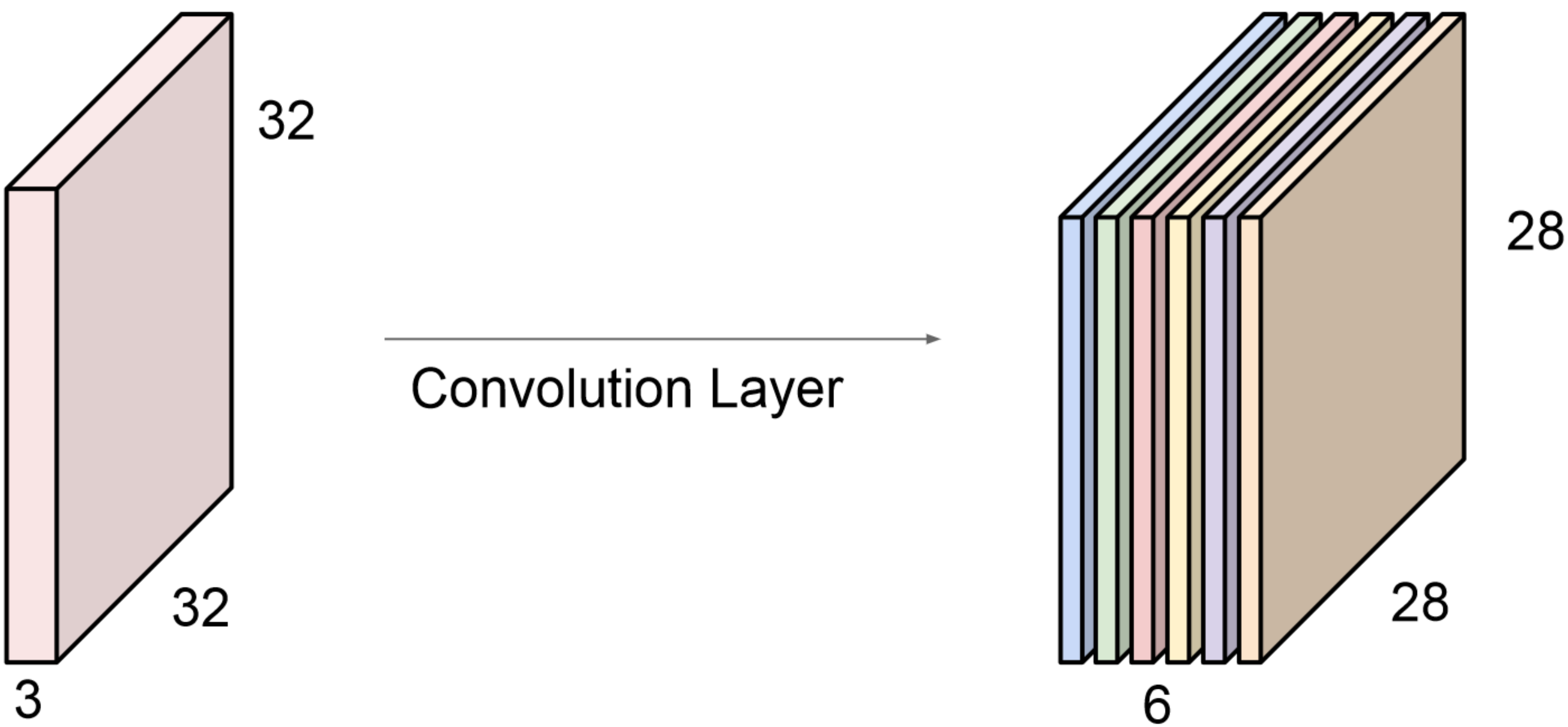


通过一层卷积做边缘检测的例子



图像产生完整卷积层的示例

例：图像是由3通道（例如RGB）组成，32x32
有6个5x5滤波核，则产生6个卷积通道，



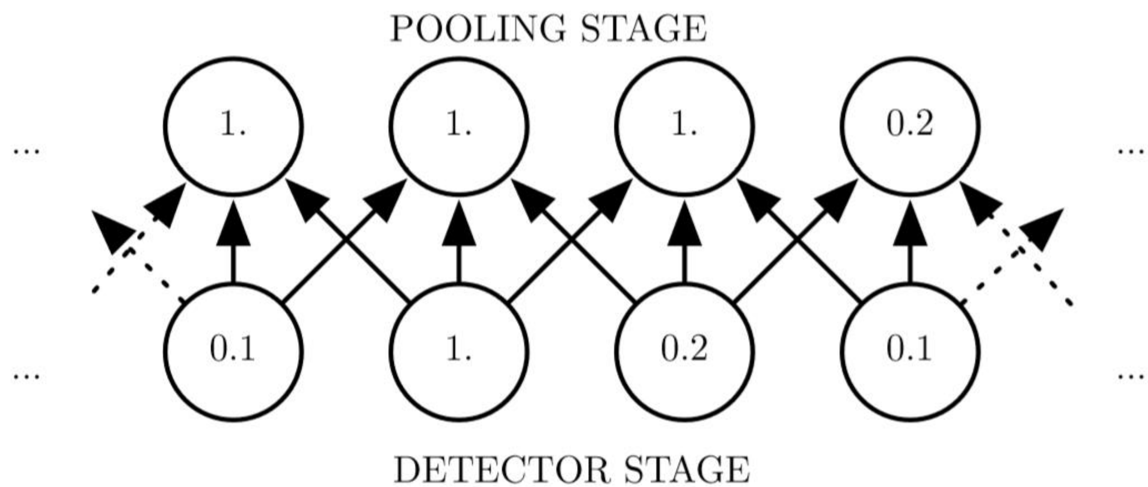
池化（Pooling）

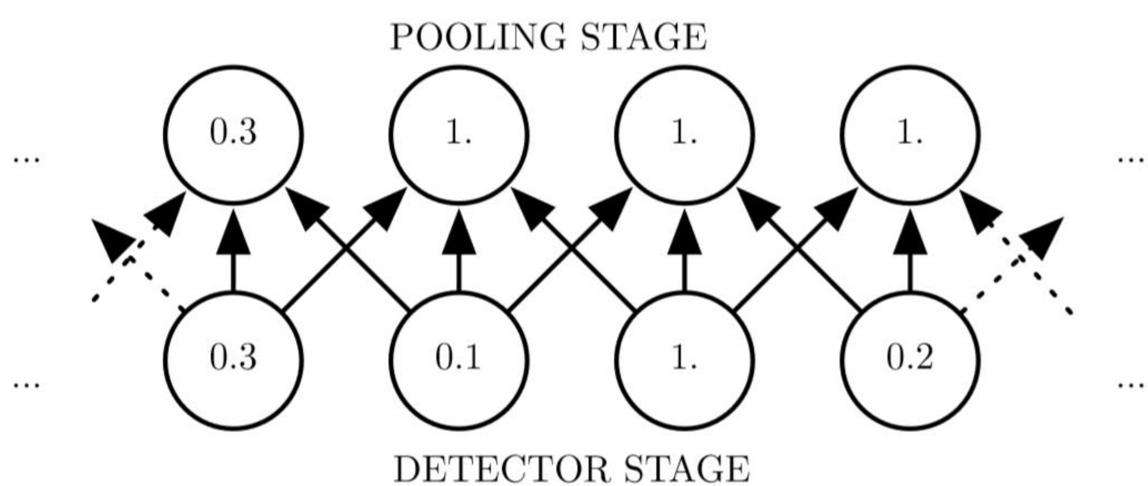
- 卷积网络一个典型层由3级组成：
  - 第一级并行地计算多个卷积层
  - 通过非线性激活函数，例如ReLU（探测级）
  - 使用池化函数
- 池化层帮助实现输入表示的近似不变性
  - 最大池化函数（Max Pooling）-输出相邻区域最大值
  - 相邻区域均值
  - 简单抽取

最大池化例子（步进为1）

从激活函数输出到池化输出（池化窗为3）

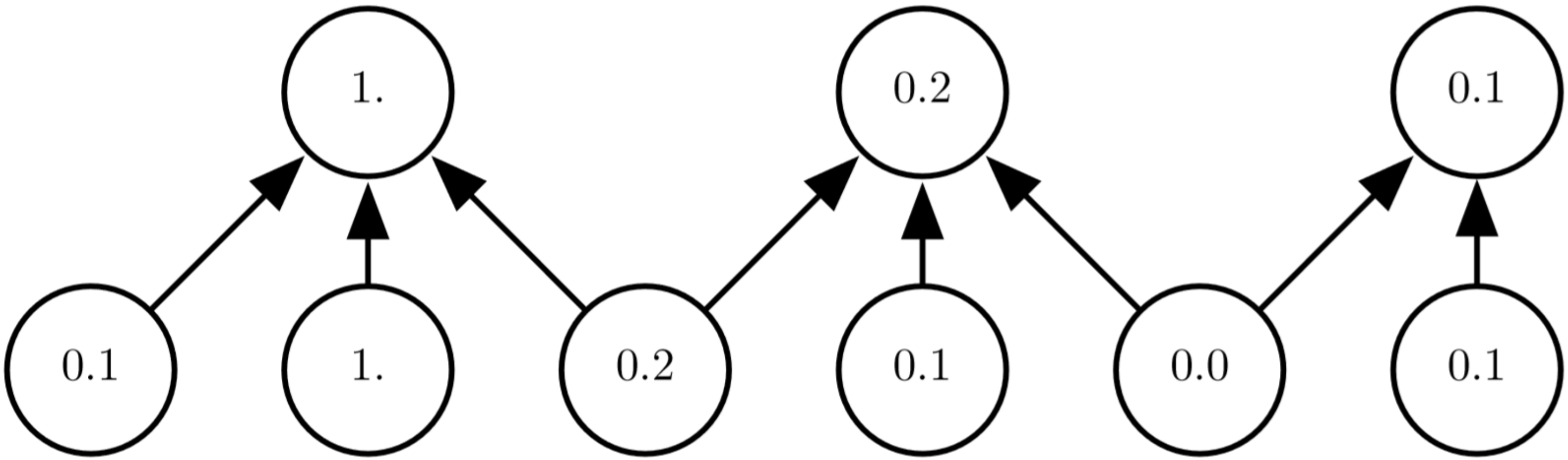
注意，上下图的激化输出有移动。



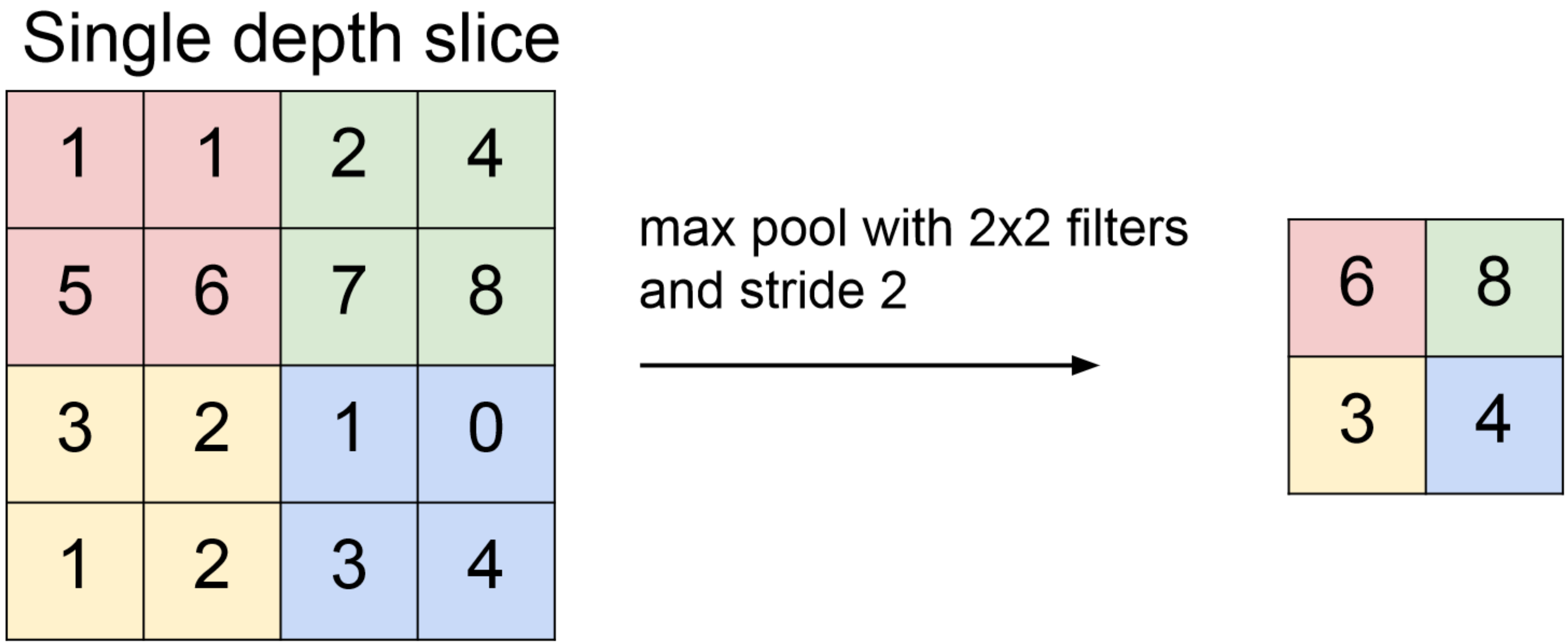


最大池化例子（步进为2）

从激活函数输出到池化输出（池化窗为3）
相当于降采样池化

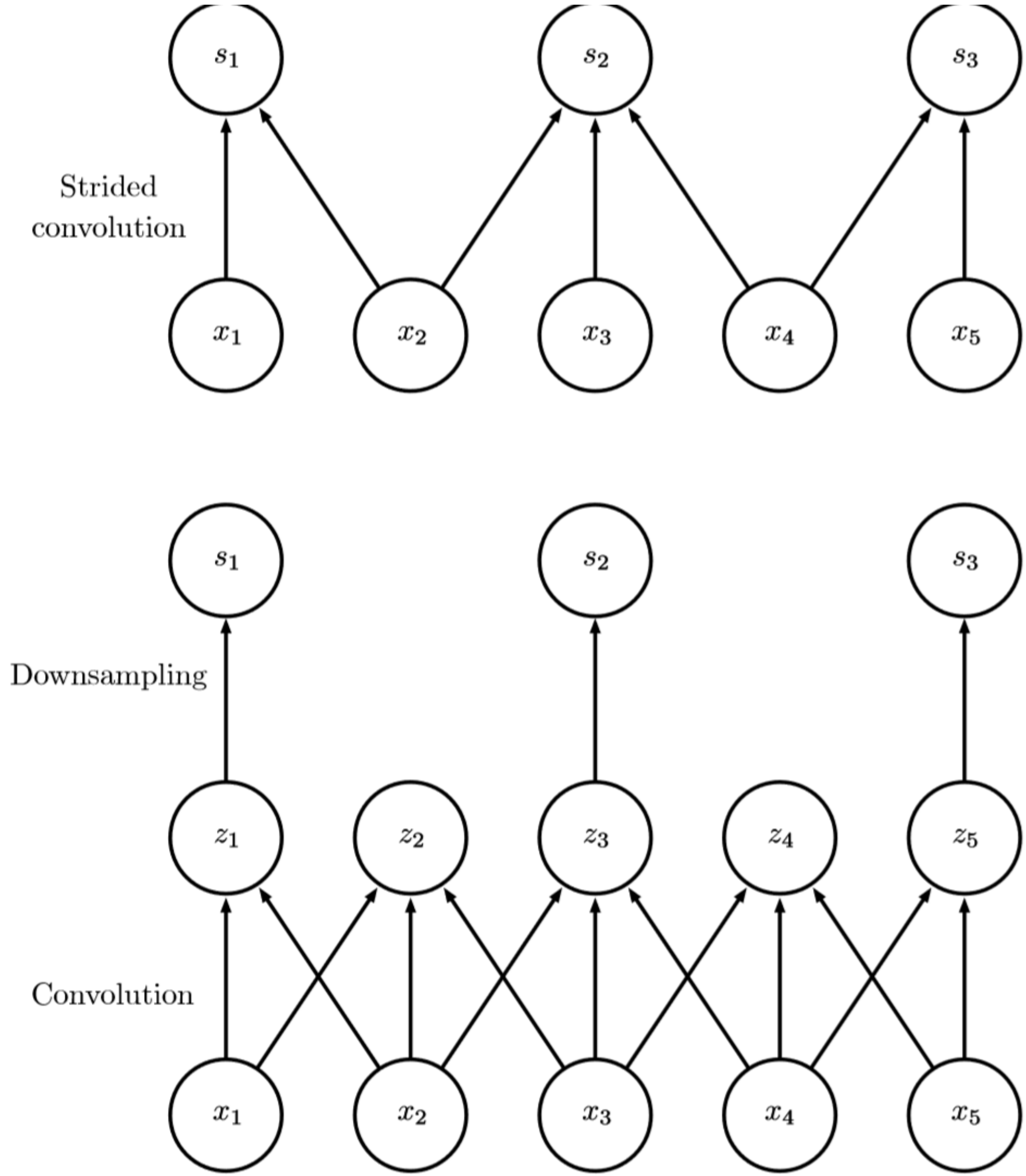


一个池化示例：
最大化池化，图像行列步进均为2



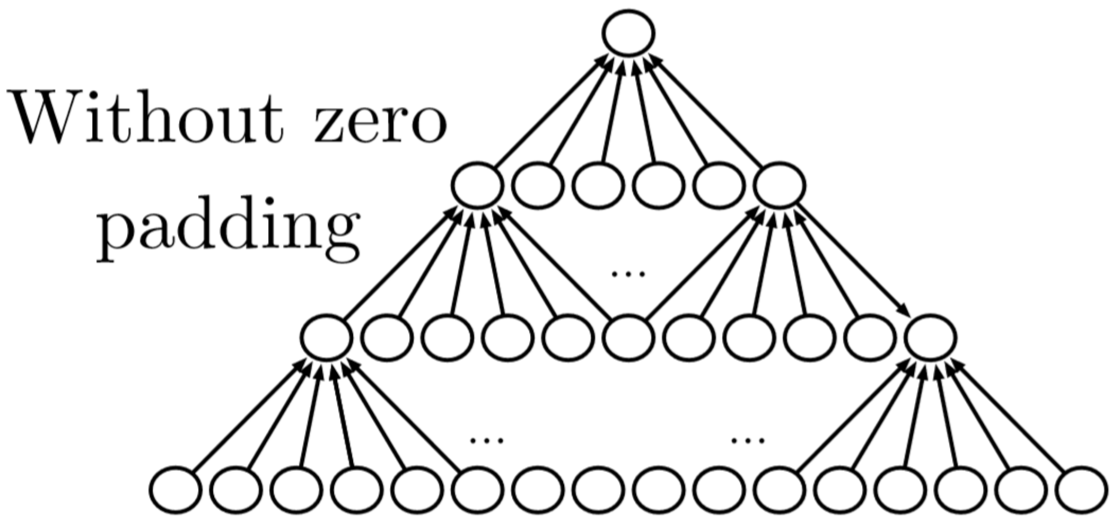
步进卷积

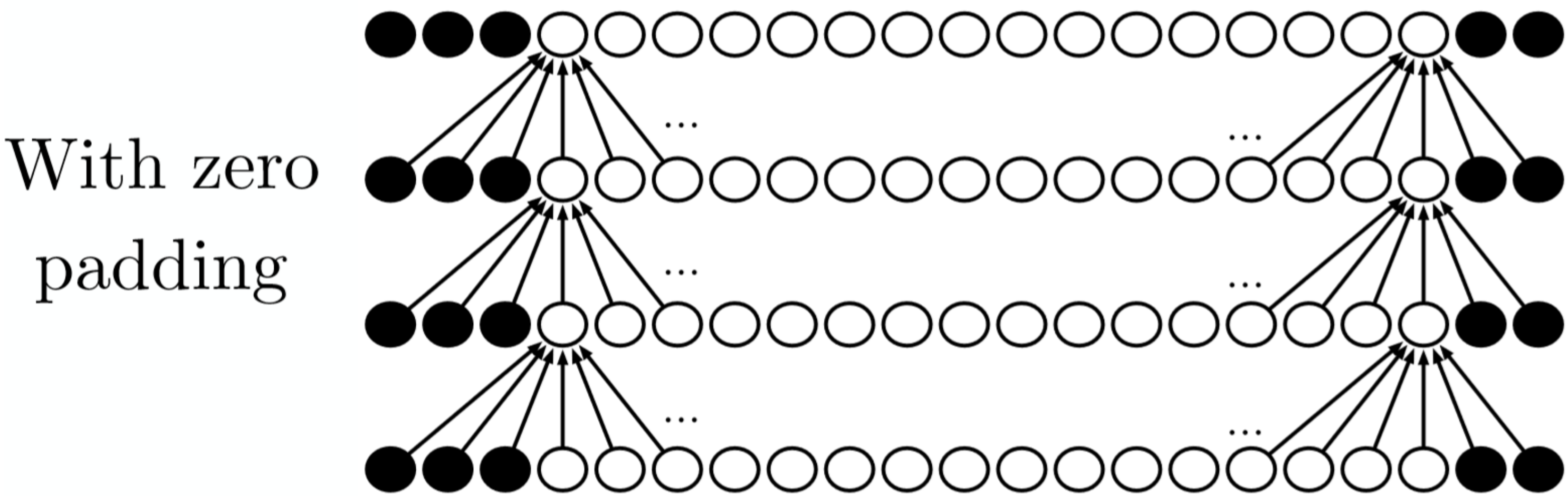
步进>1卷积
相当于1步进
卷积加下采样



填充零保证卷积相同大小

上图卷积逐步减小
下图保持卷积相同





张量卷积（体卷积）的扩展

$$a_{i,j,p} = \sum_{n,m,k} h^{(p)}_{n,m,k} X_{i+n,j+m,k}$$

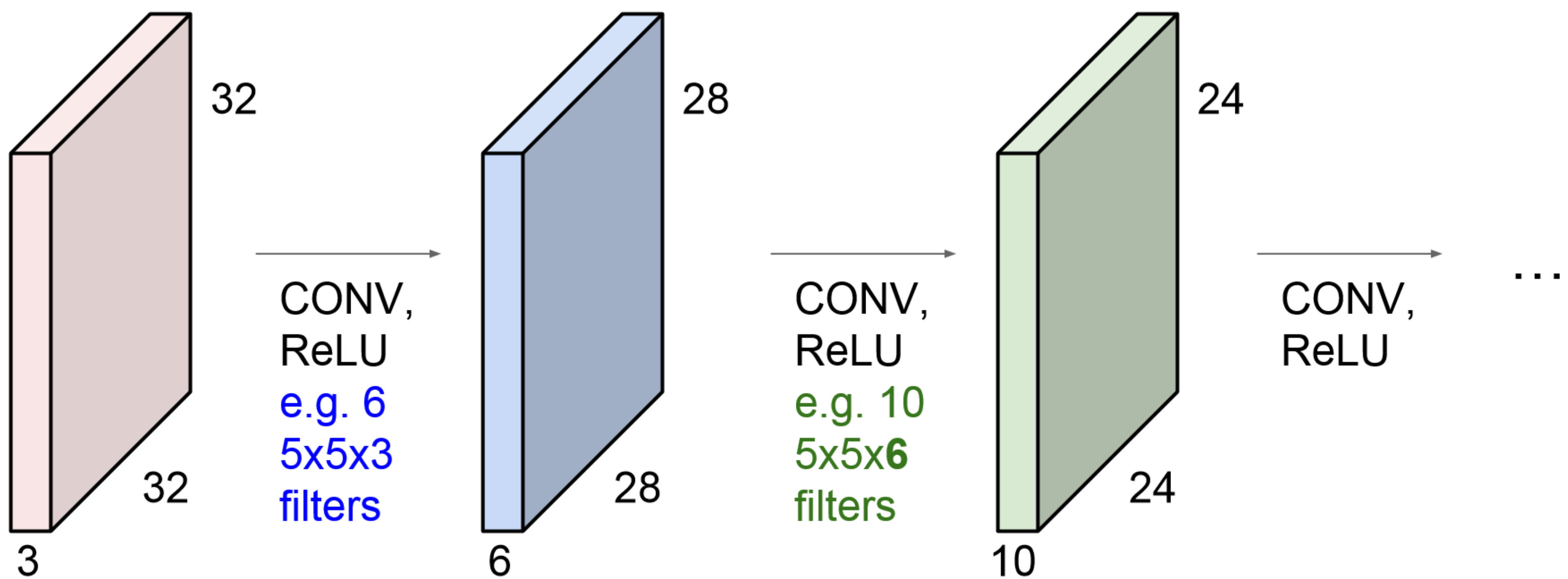
多个卷积核 $h^{(p)}_{n,m,k}$ 来产生多个卷积平面输出

也可以直接进行步幅为s的降采样卷积

$$a_{i,j,p} = \sum_{n,m,k} h^{(p)}_{n,m,k} X_{i\times S+n,j\times S+m,k}$$

一个张量卷积例子

该例主要说明体卷积，紧跟ReLU激活函数，池化步幅为1，实际中池化步幅大多>1。



1×1卷积

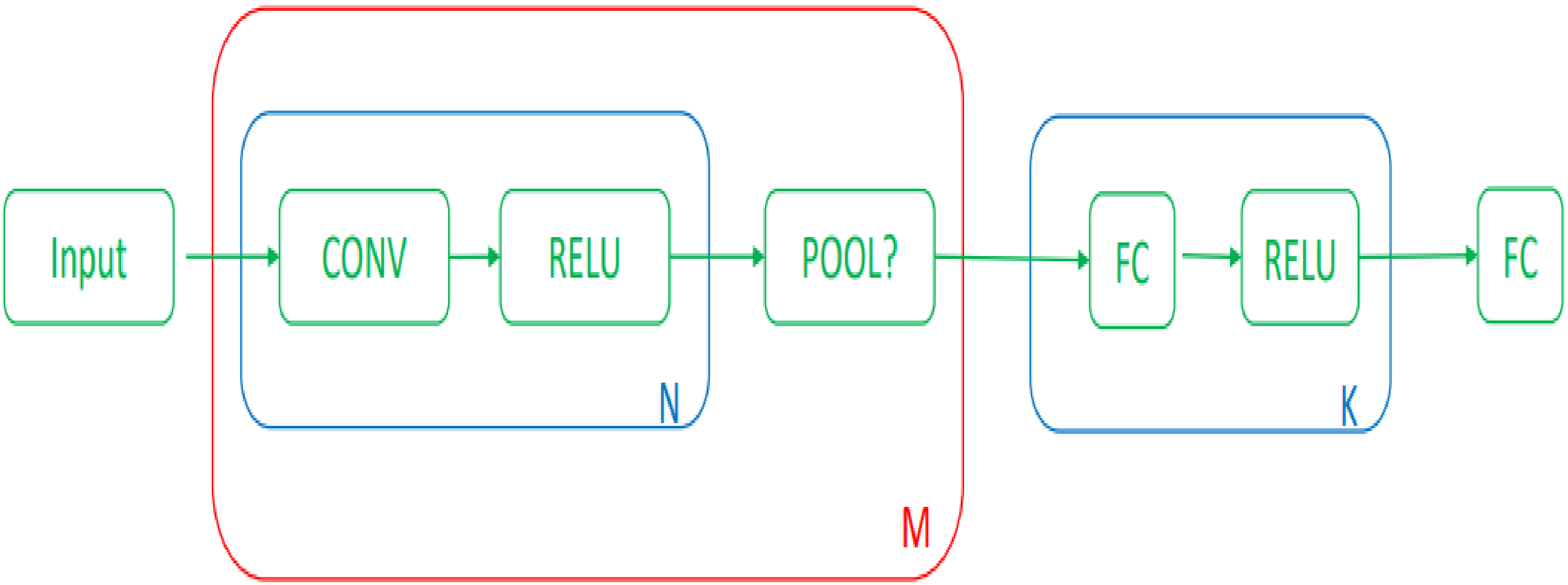
$$a_{i,j,p} = \sum_k h_k^{(p)} X_{i,j,k}$$

卷积核在行、列方向长度为1

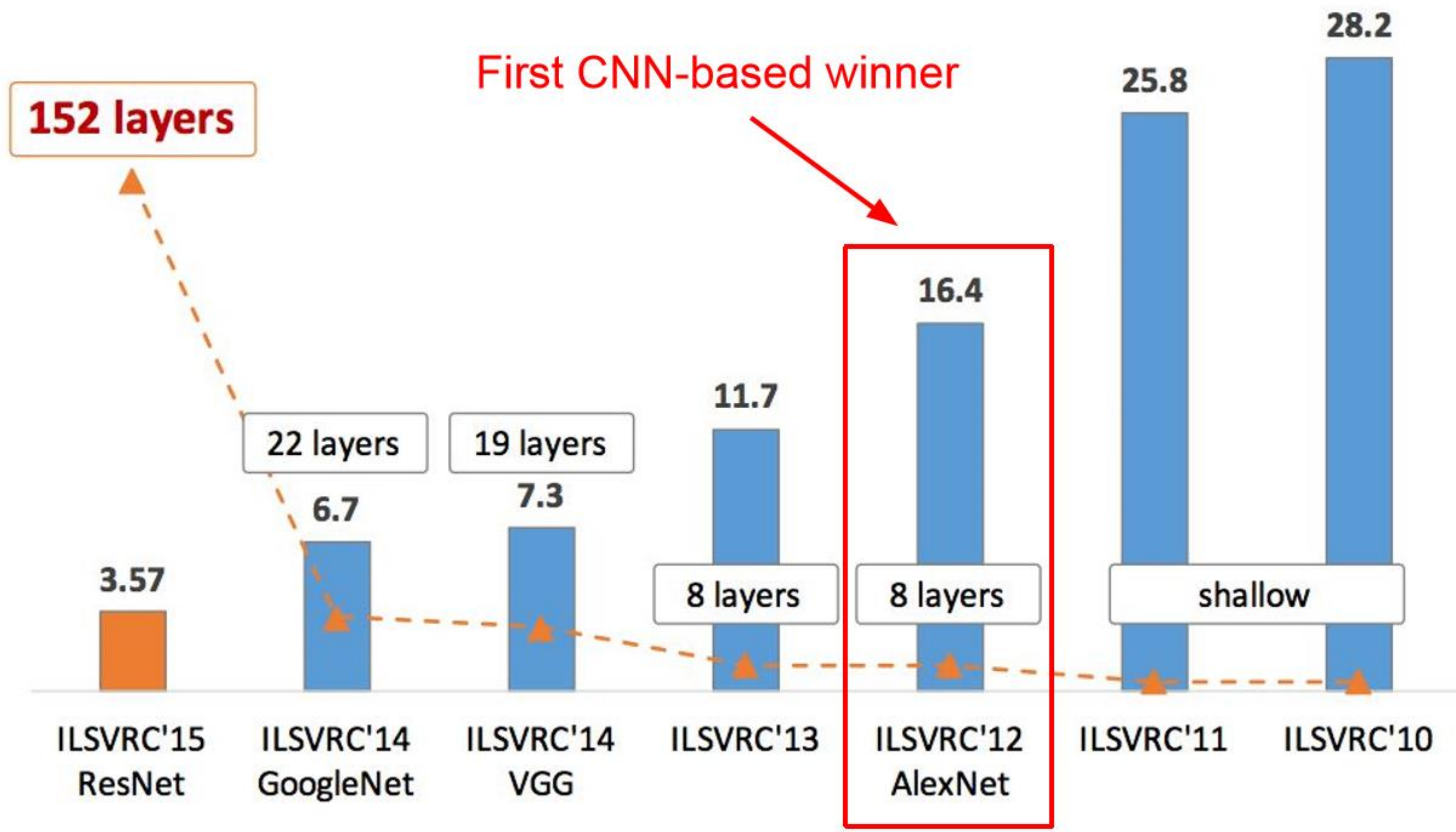
实际运算主要在通道维进行

多个1×1卷积核 $h_k^{(p)}$ 得到多个卷积平面

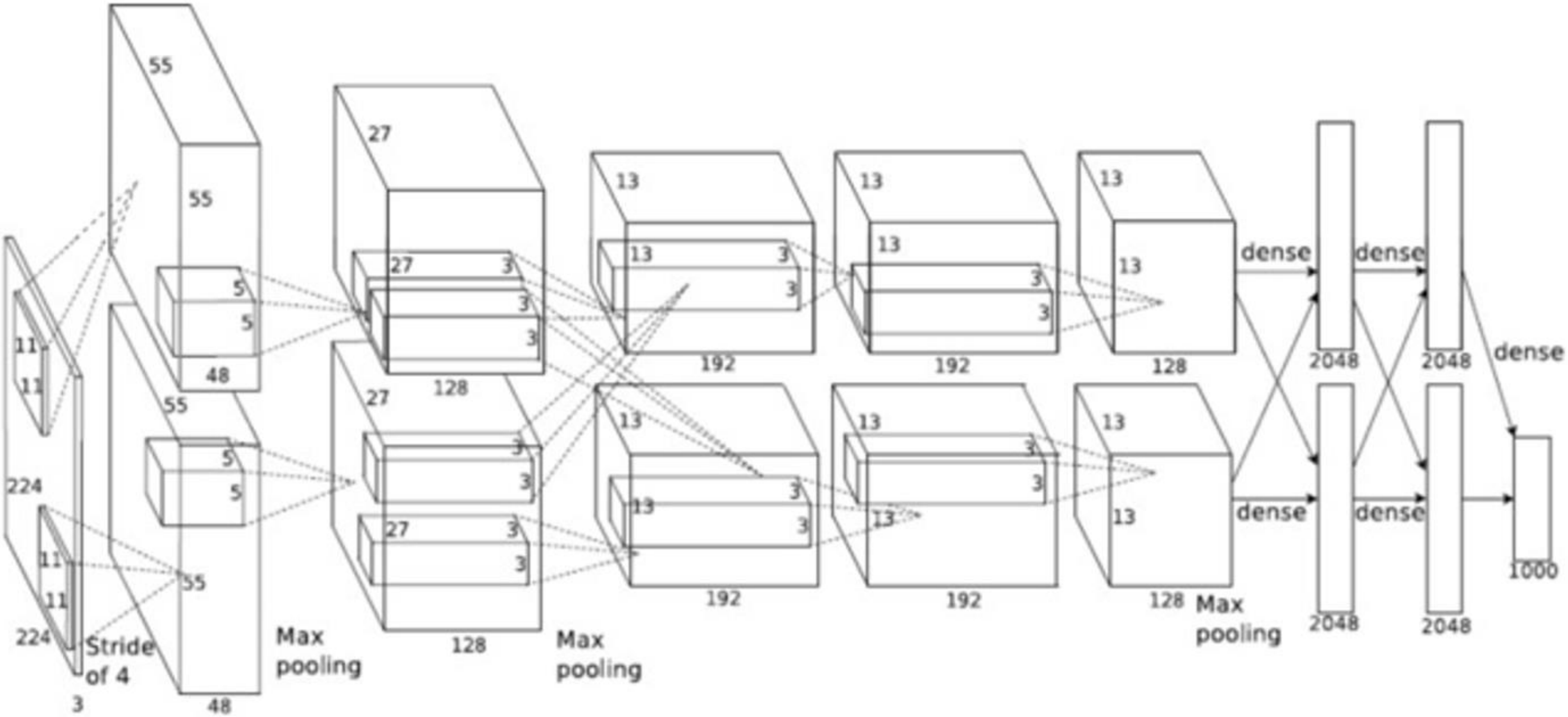
典型CNN网络结构示意图



CNN在图像识别的成果
（ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners）



AlexNet

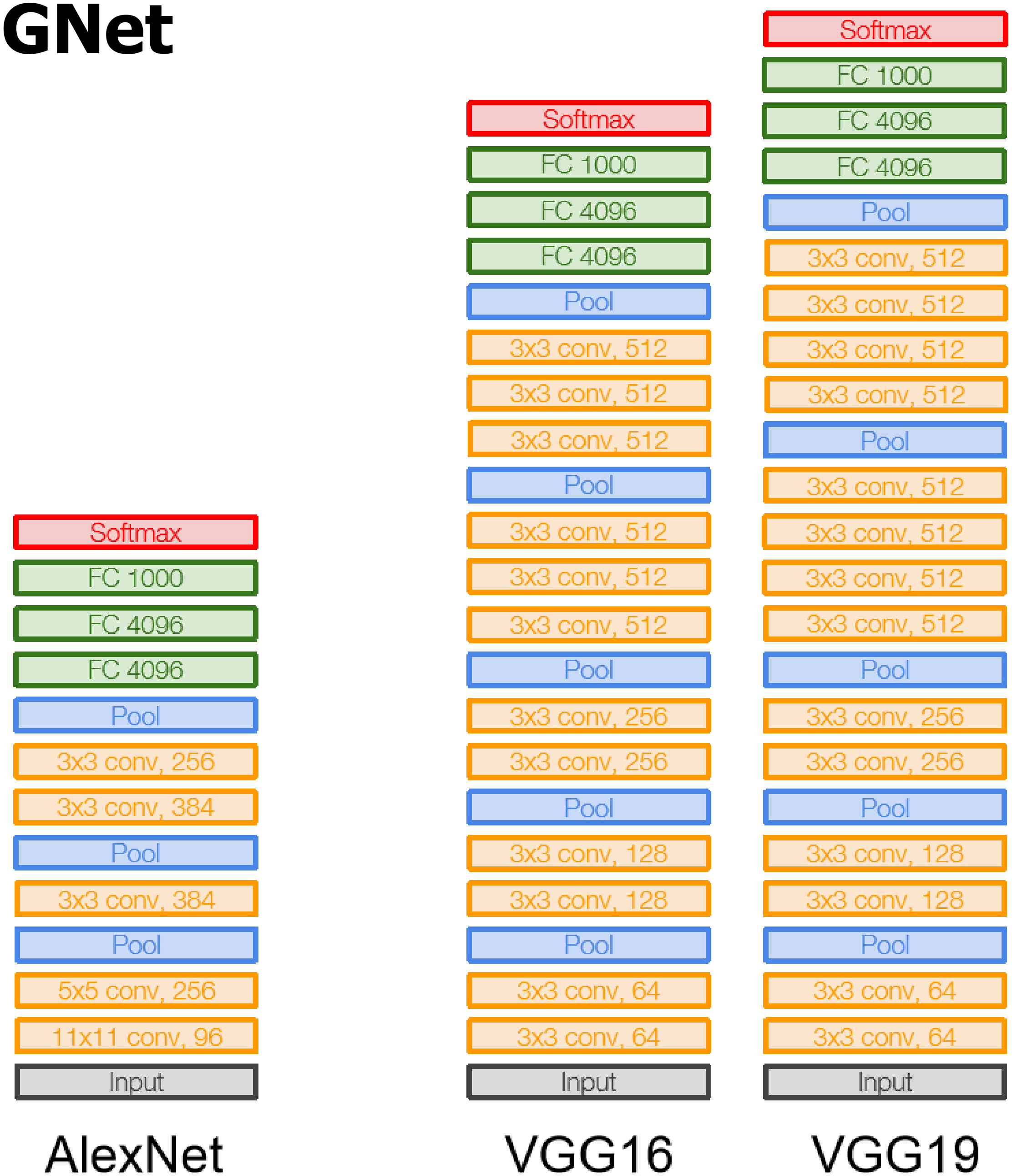


224
224
3
11
11
Stride of 4
55
55
48
5
5
Max pooling
27
27
128
3
Max pooling
13
192
128
Max pooling
dense
2048
1000

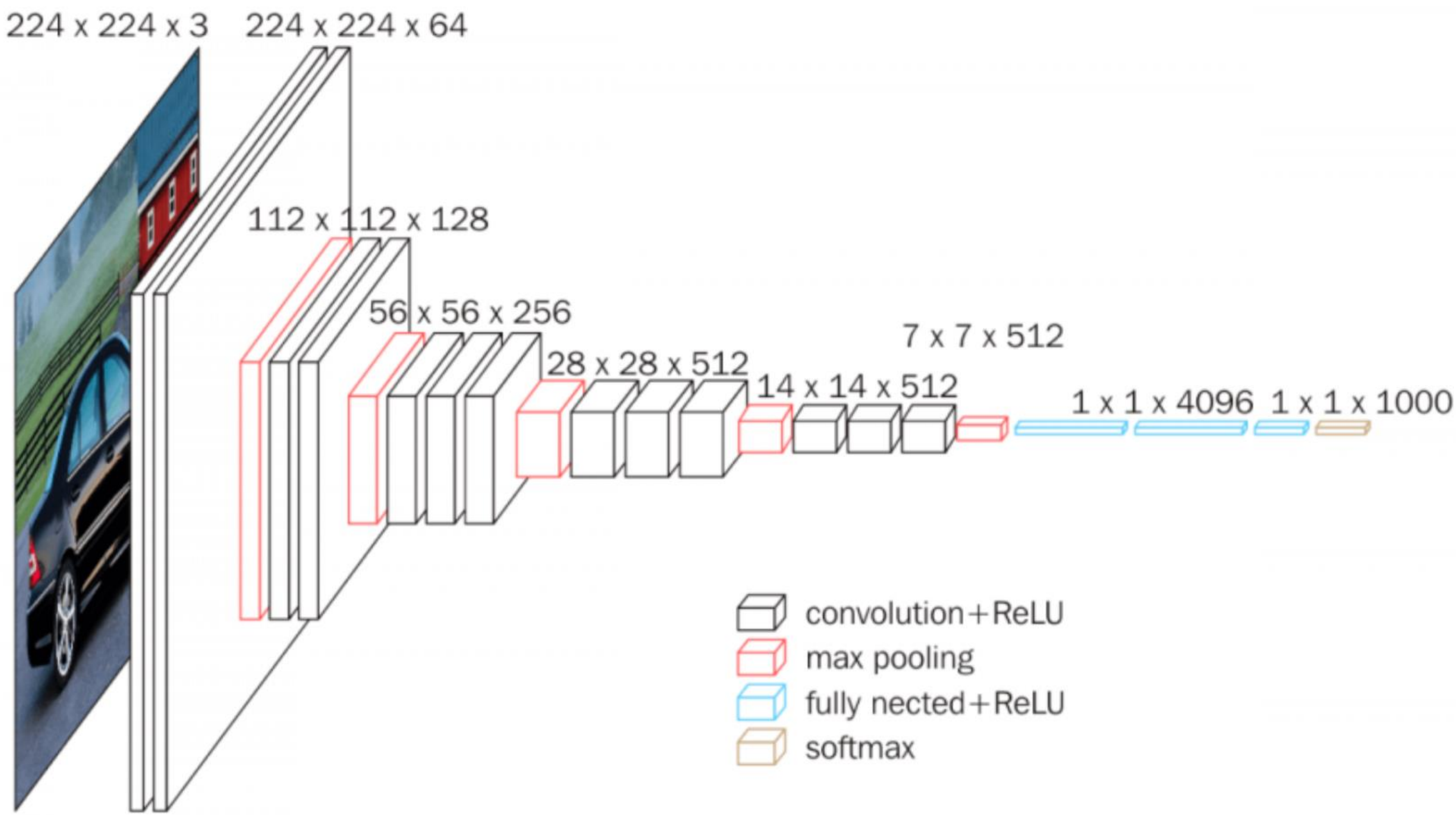
CNN网络实例：VGGNet

Small filters,
Deeper networks

8 layers (AlexNet)
-> 16 - 19 layers
(VGG16Net)
Only 3x3 CONV
stride 1, pad 1 and
2x2 MAX POOL
stride 2

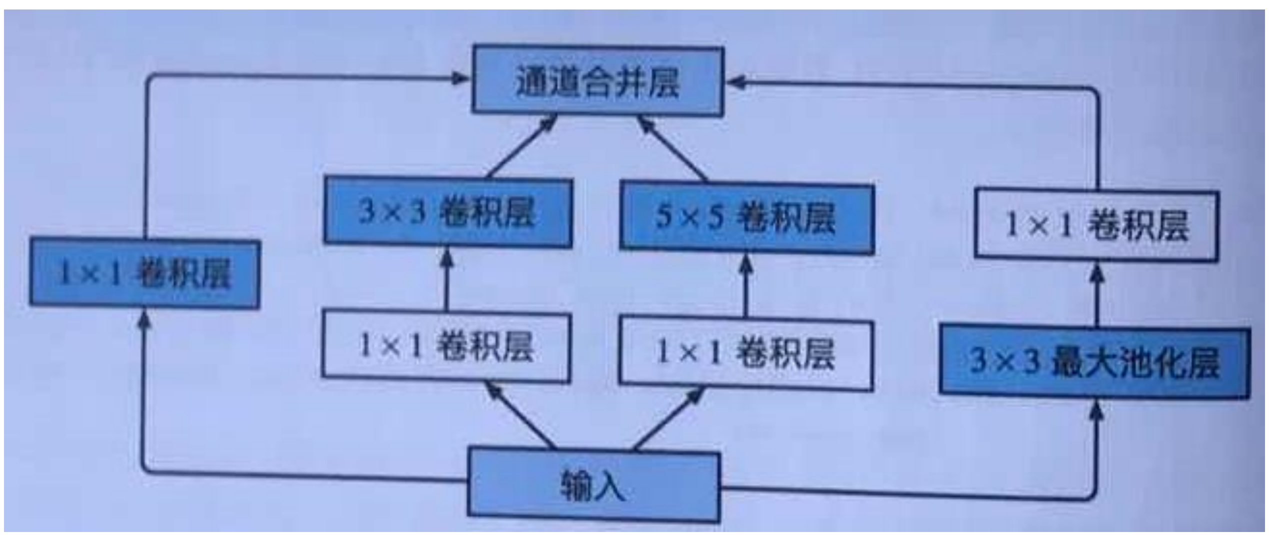


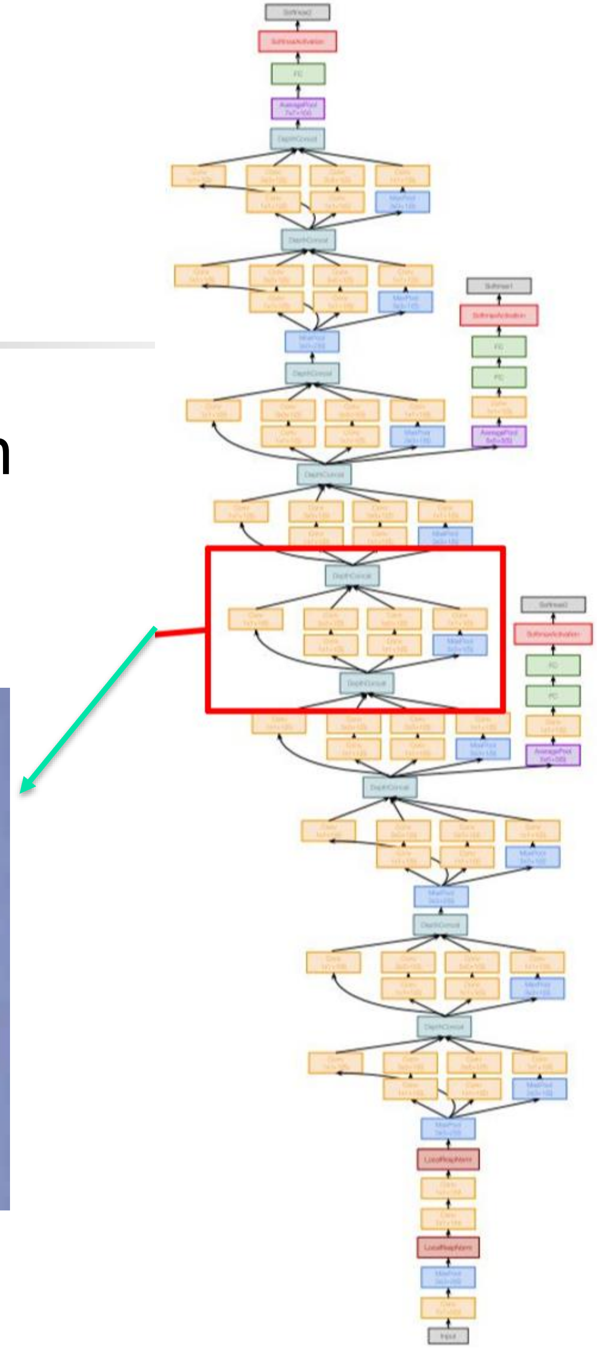
VGGNet结构示意图



CNN网络实例：GoogLeNet

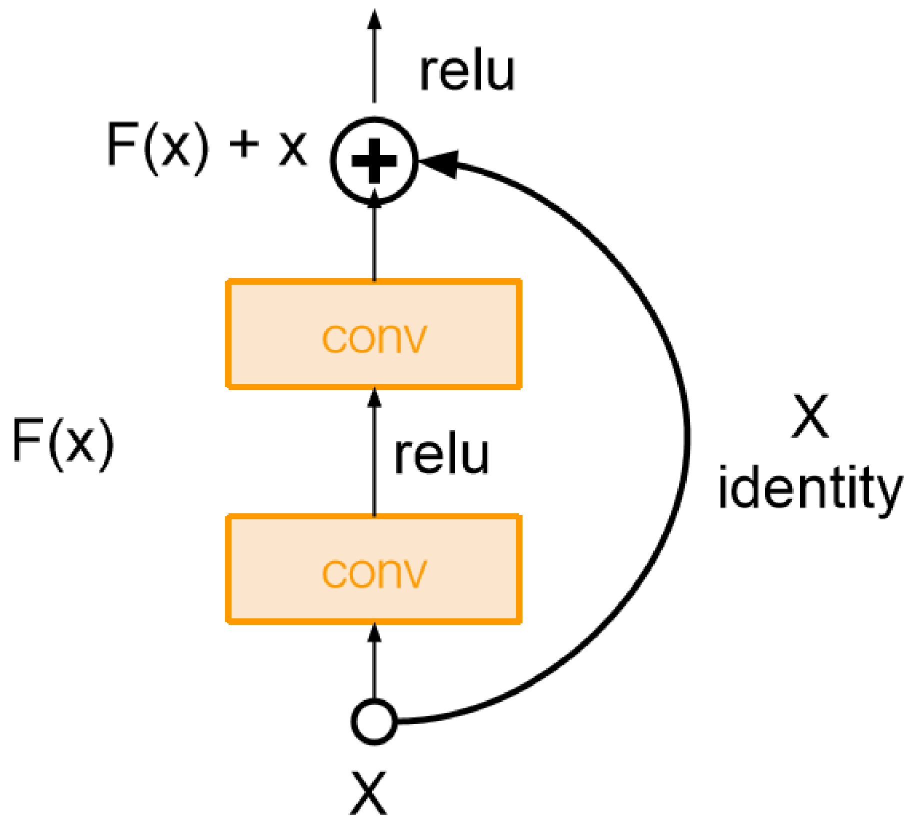
基本构造块：Inception (取自电影Inception（盗梦空间）)



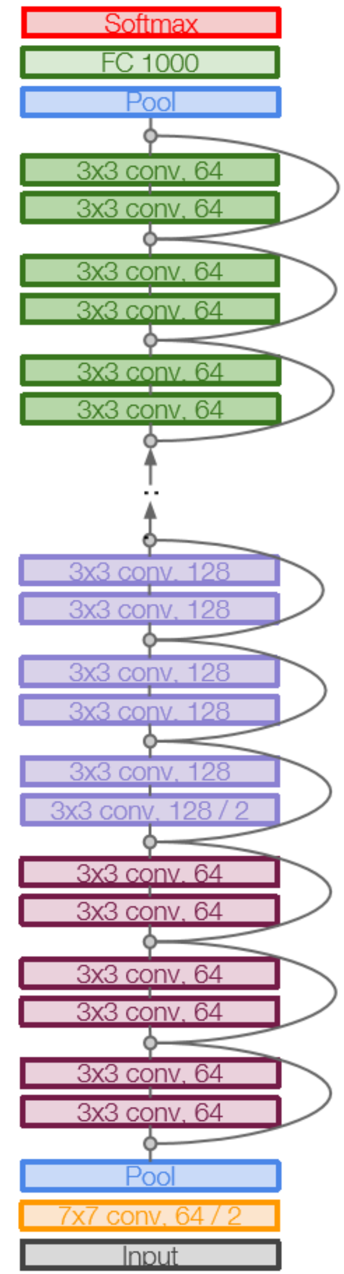


CNN网络实例：ResNet
残差网络

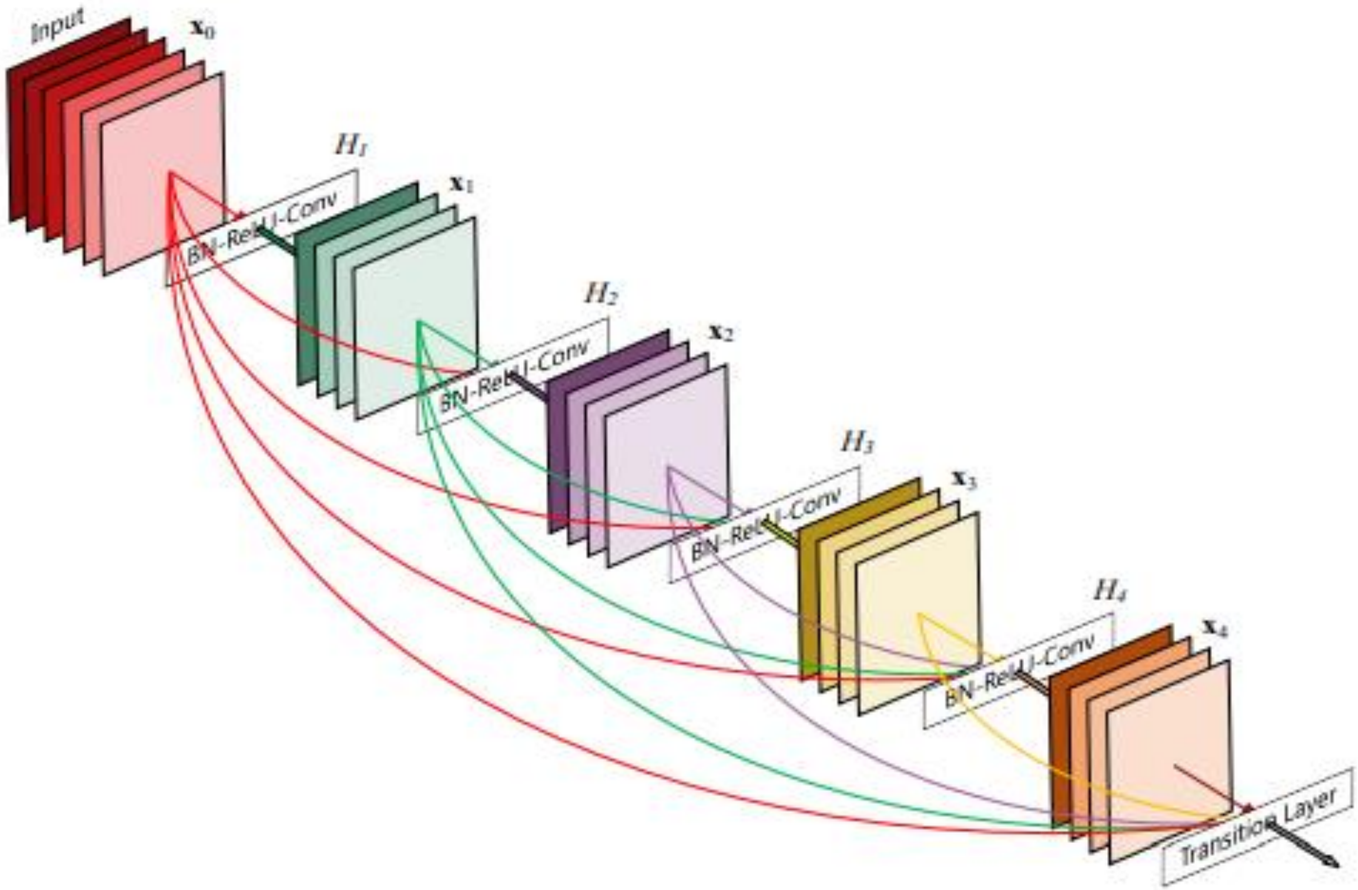
Very deep networks using residual connections
- 152-layer model for ImageNet - ILSVRC’ 15 classification winner (3.57% top 5 error) -



Residual block



DenseNet


Input
$\mathbf{x}_0$
H_1
BN-ReLU-Conv
$\mathbf{x}_1$
H_2
BN-ReLU-Conv
$\mathbf{x}_2$
H_3
BN-ReLU-Conv
$\mathbf{x}_3$
H_4
BN-ReLU-Conv
$\mathbf{x}_4$
Transition Layer

一些代表性的CNN结构示例

- AlexNet
- VGG
- GoogLeNet
- ResNet
- DenseNet
- FractalNet
- SqueezeNet

- NiN (Network in Network)
- Wide ResNet
- ResNeXT
- Stochastic Depth

CNN网络小结

- CovnNET一般有多卷积层（卷积+激活+池化）+全连接层（FC）组成；
- 目前趋势是更小的滤波核和更深的层数，甚至取消池化和全连接层，构成纯卷积网络。
- 一种典型结构：
  - [(CONV-RELU)*N-POOL?]*M-(FC-RELU)*K,SOFTMAX，这儿N可高达~5, M 是较大值， 0 <= K <= 2.
- 新的结构不断被构造和尝试。

CNN的一些趋势

- VGG, GoogLeNet, ResNet all in wide use, available in model zoos
- ResNet current best default
- Trend towards extremely deep networks
- Significant research centers around design of layer / skip connections and improving gradient flow
- Even more recent trend towards examining necessity of depth vs. width and residual connections

2012年以后，每年ICML、NIPS等会议都有大量CNN结构的论文有兴趣可参考。这里只给出几个比较经典的工作。

参考论文

- LeCun, Y. et al. Handwritten digit recognition with a back-propagation network. In Proc. Advances in Neural Information Processing Systems 396–404 (1990).
- LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. Proc. IEEE 86, 2278–2324 (1998).
- Krizhevsky, A., Sutskever, I. & Hinton, G. ImageNet classification with deep convolutional neural networks. In Proc. Advances in Neural Information Processing Systems 25 1090–1098 (2012).
- K. Simonyan and A. Zisserman "Very Deep Convolutional Networks for Large-Scale Image Recognition"，ICLR，2014
- C. Szegedy, Liu W., Jia Y. et al, Going Deeper with Convolutions, IEEE CVPR, 2015
- He K. et al. Deep Residual Learning for Image Recognition，IEEE CVPR, 2016