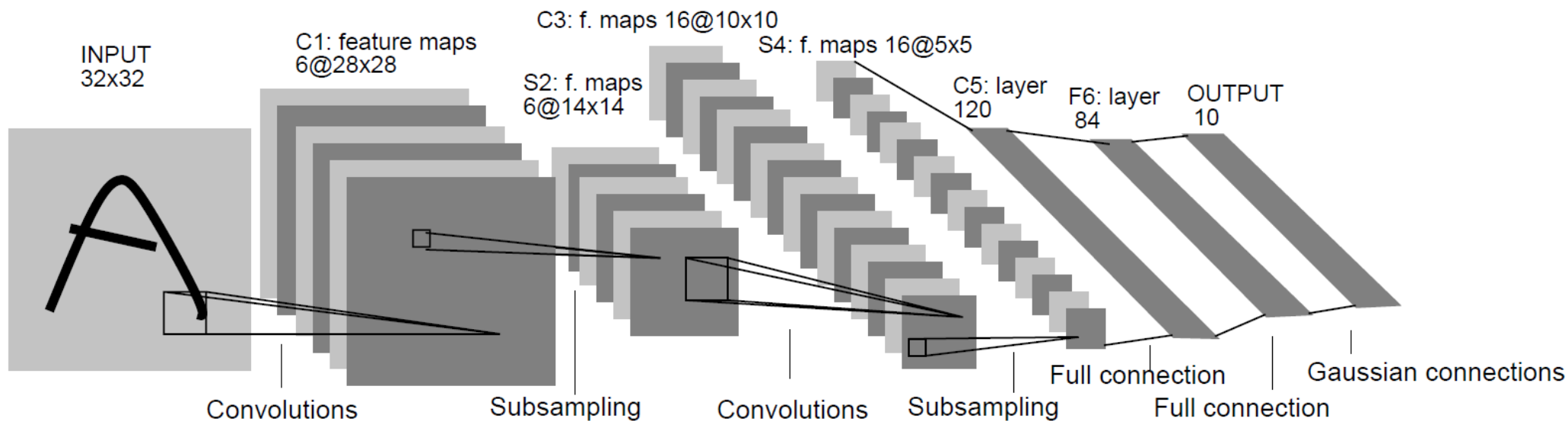


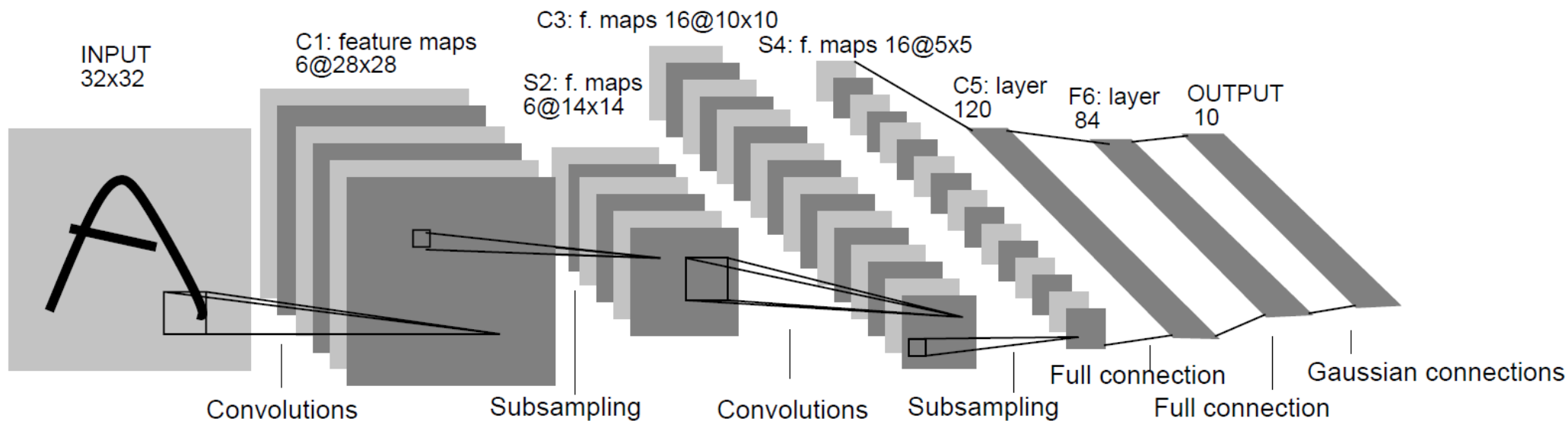
### 3. 经典CNN（分类网络）

# 1. LeNet-5



LeNet-5 出自论文 Gradient-Based Learning Applied to Document Recognition，是一种用于手写体字符识别的非常高效的卷积神经网络。

# 1. LeNet-5



输入尺寸：32\*32

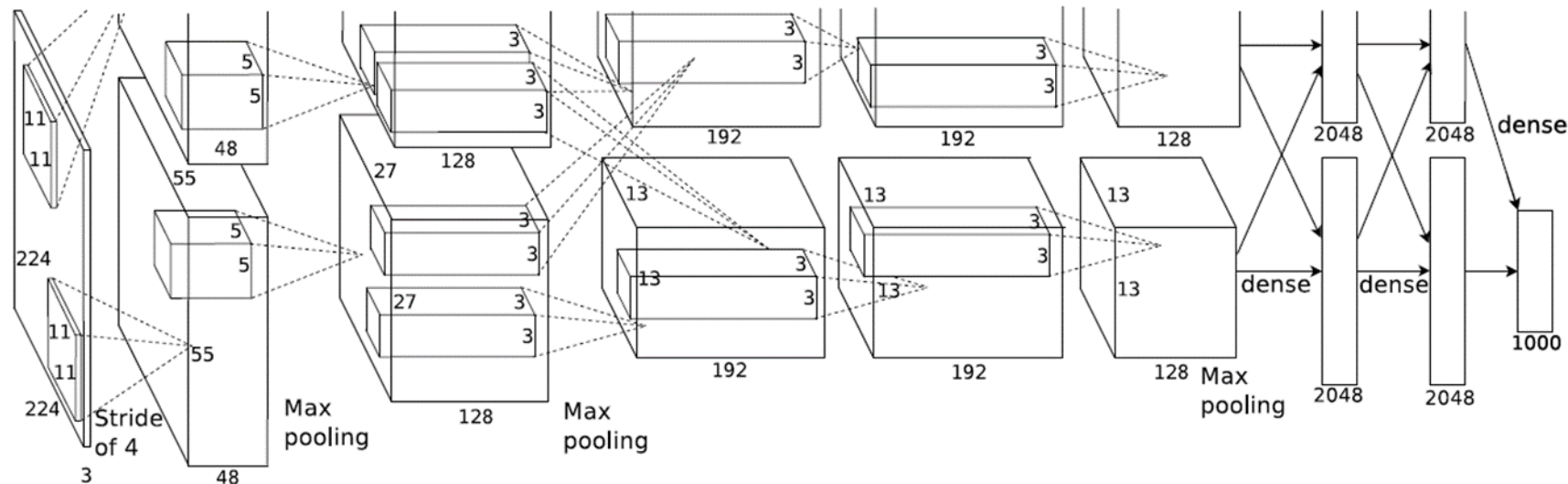
卷积层：2个

下采样层：2个

全连接层：2个

输出：10个类别（数字0-9的概率）

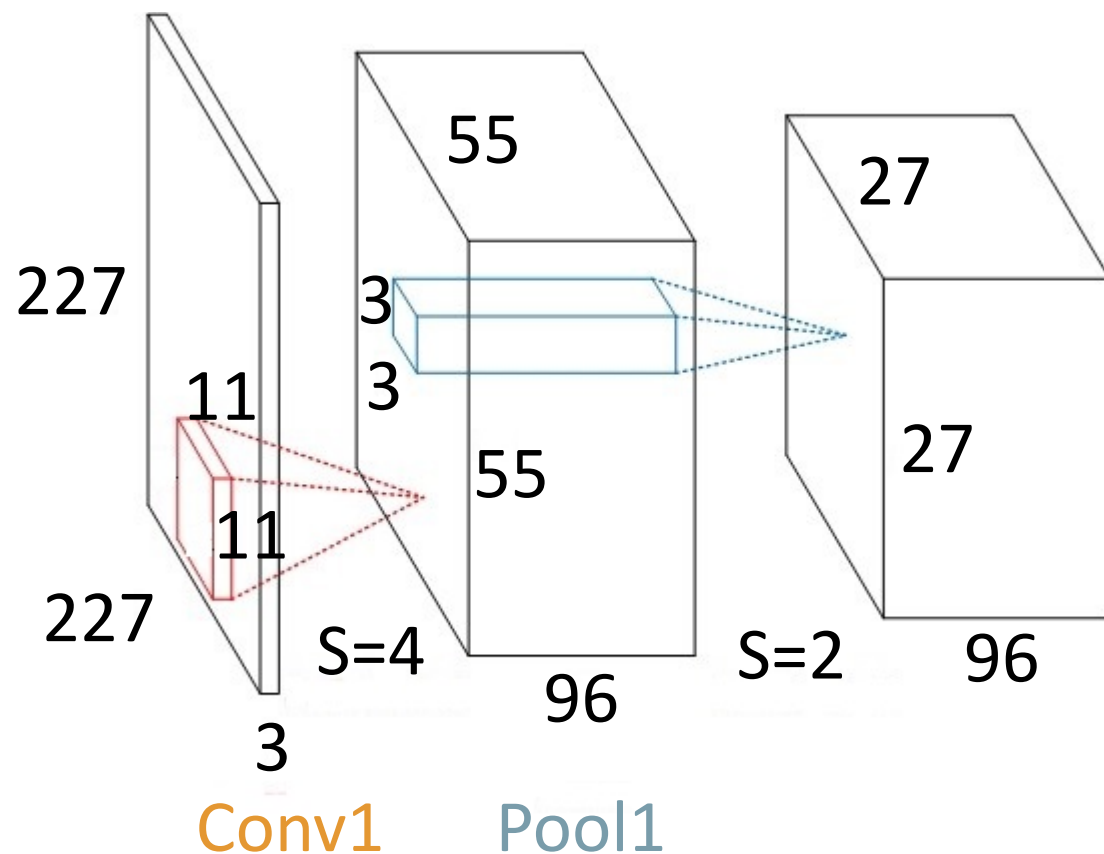
## 2. AlexNet



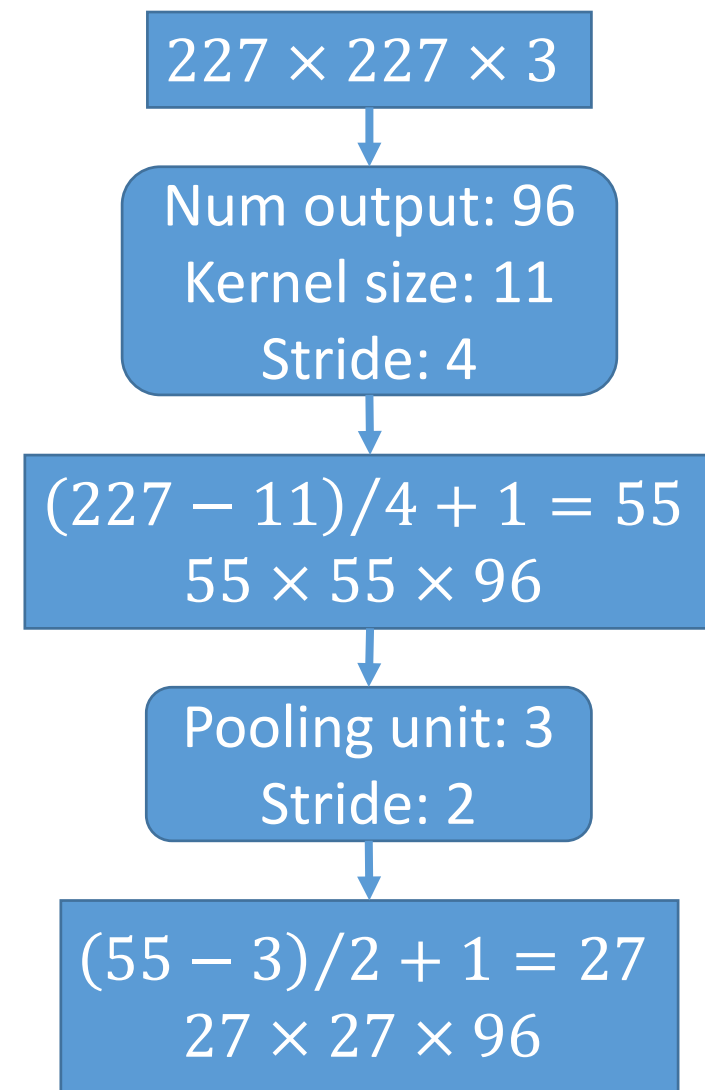
### AlexNet架构

**AlexNet**由Alex Krizhevsky于2012年提出，夺得2012年ILSVRC比赛的冠军，**top5**预测的错误率为**16.4%**，远超第一名。**AlexNet**采用**8层**的神经网络，**5个卷积层**和**3个全连接层**(3个卷积层后面加了最大池化层)，**6000万个参数**和**65万个神经元**。

## 2. AlexNet

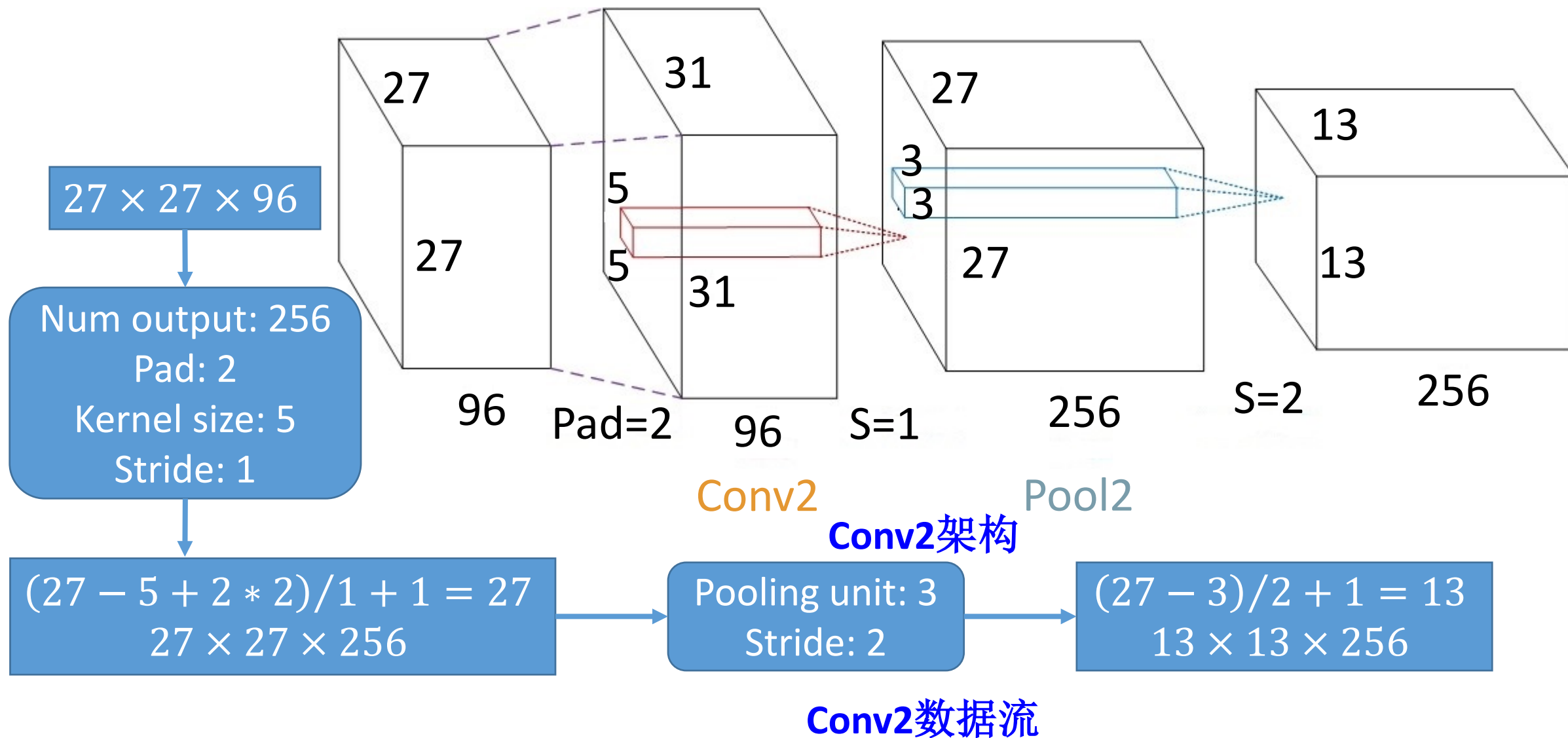


Conv1的架构

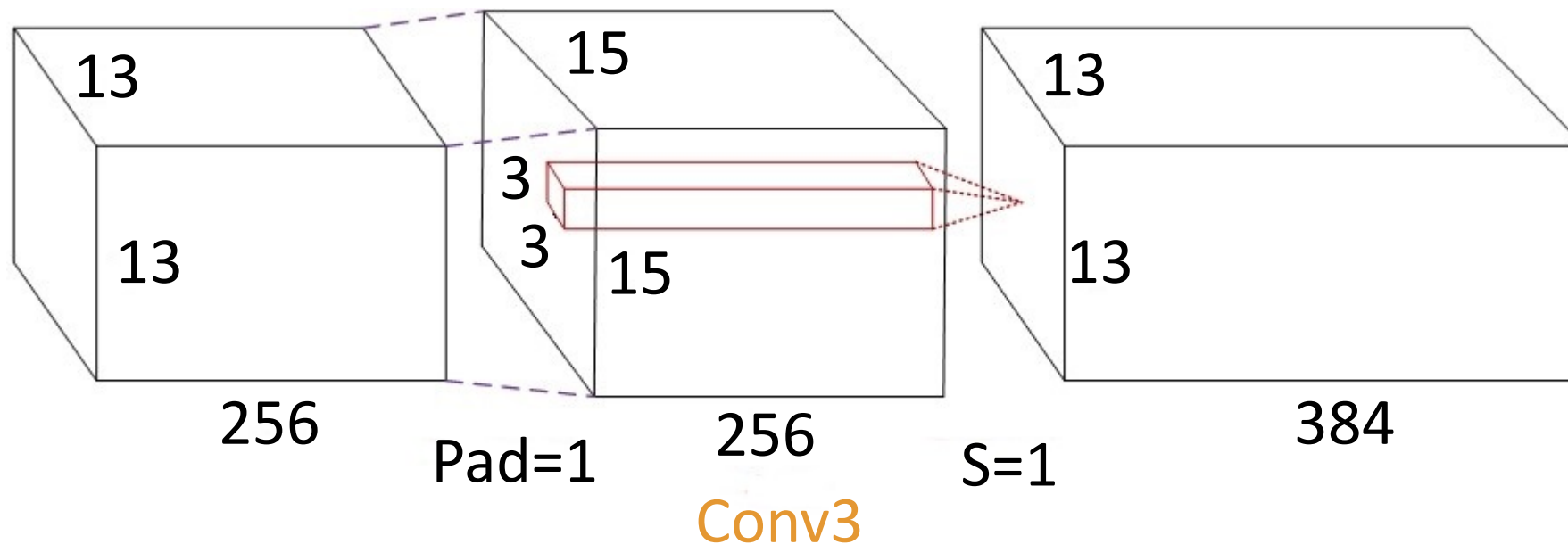


Conv1数据流

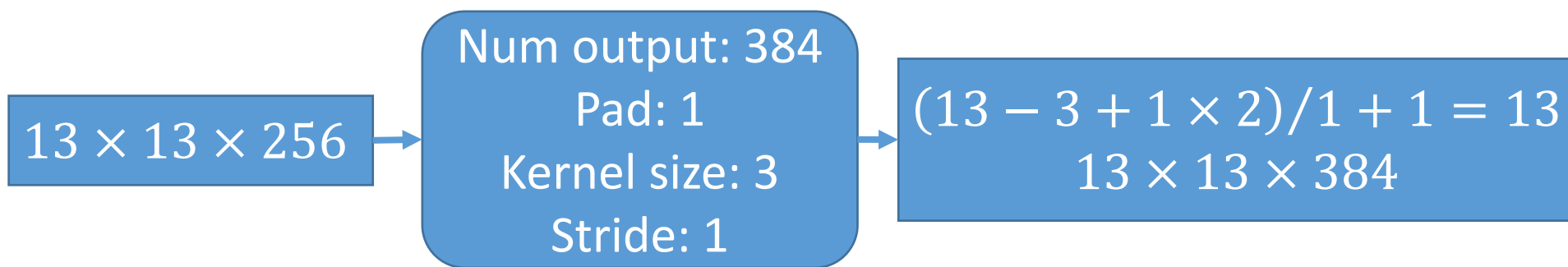
## 2. AlexNet



## 2. AlexNet

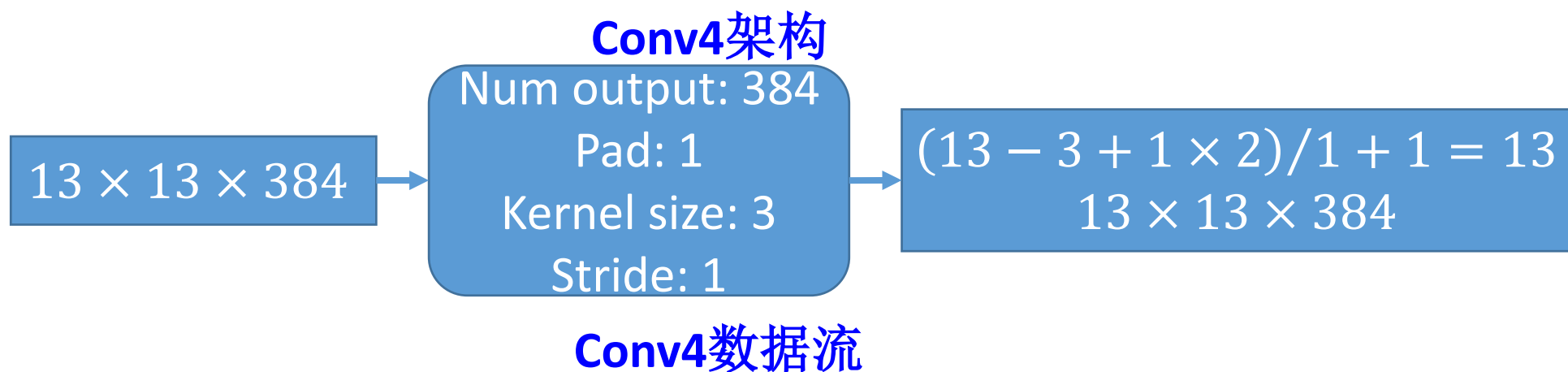
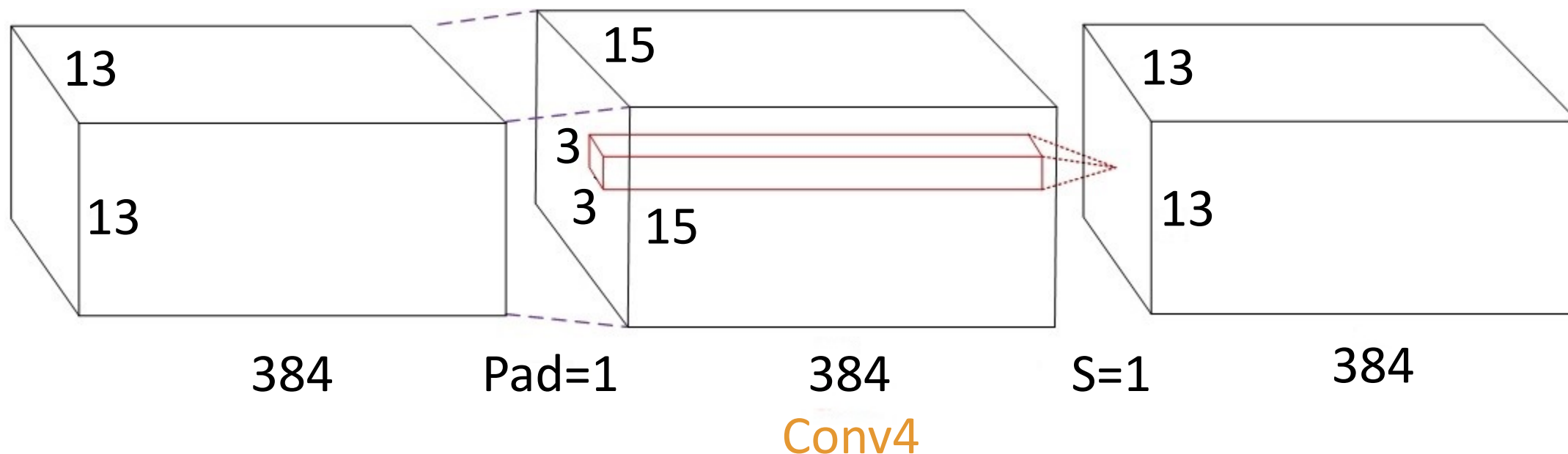


Conv3架构



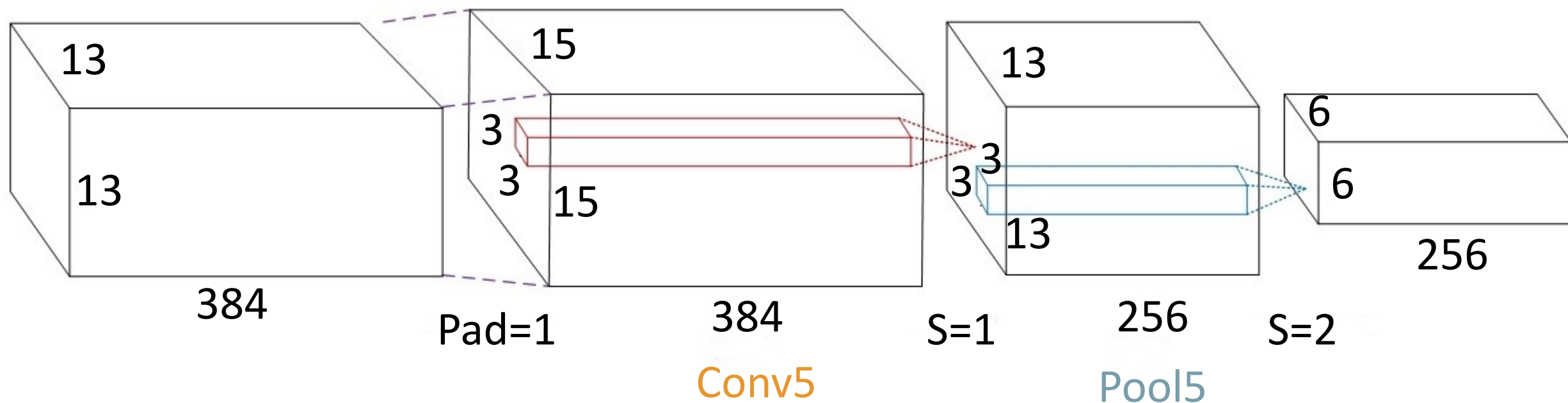
Conv3数据流

## 2. AlexNet

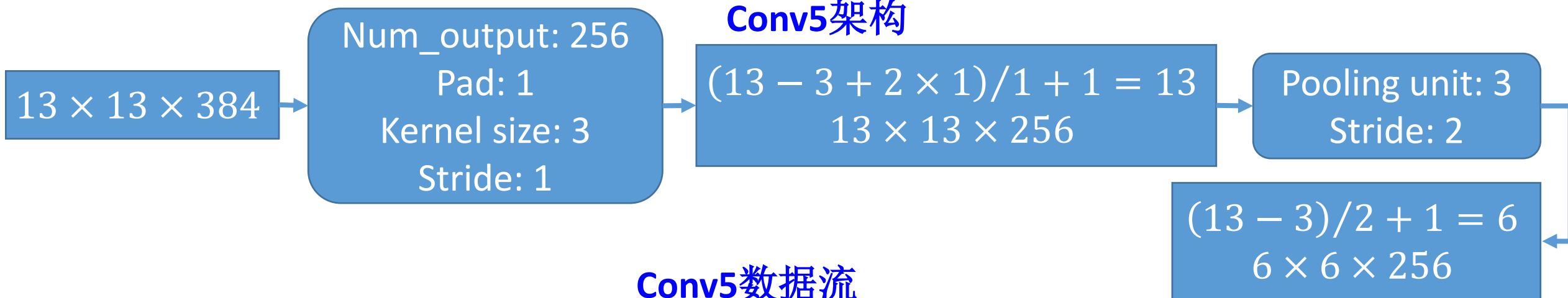




## 2. AlexNet

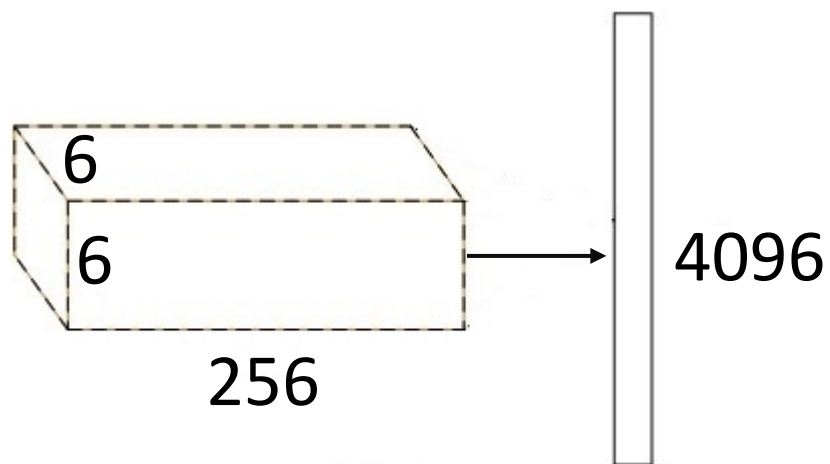


### Conv5架构



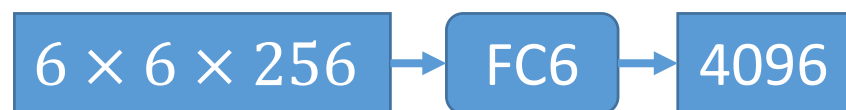
### Conv5数据流

## 2. AlexNet

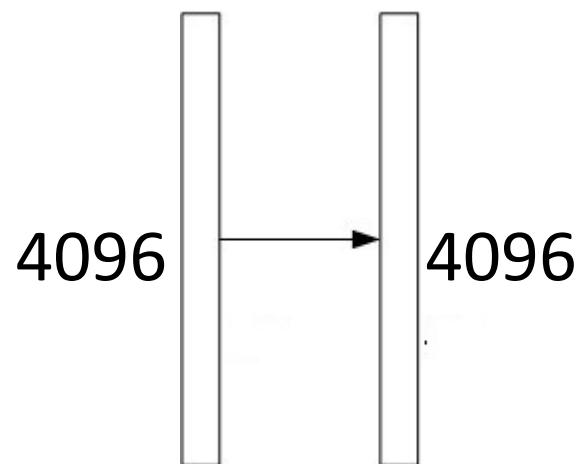


FC6

FC6架构



FC6数据流

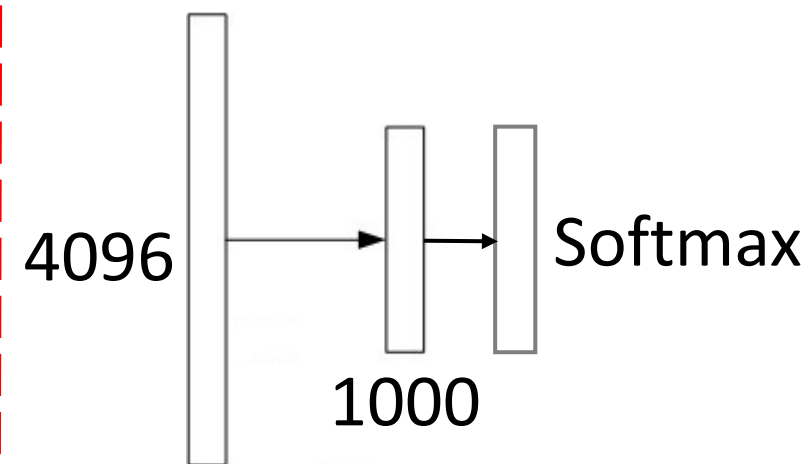


FC7

FC7架构



FC7数据流



FC8

FC8架构



FC8数据流

## 2. AlexNet

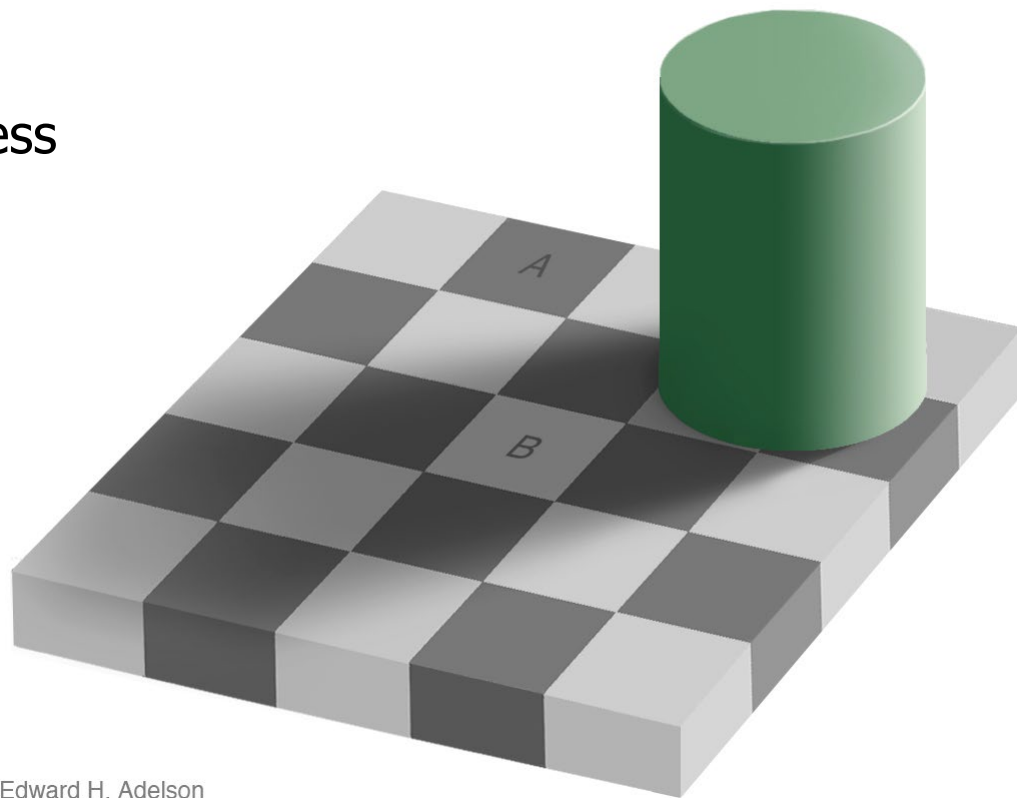
层名称	配置	节点数量	参数数量 (w和b)
Input	227*227*3	227*227*3=154587	
Conv1	11*11*3*96	55*55*96=290400	11*11*3*96+96=34944
Conv2	5*5*96*256	27*27*256=186624	5*5*96*256+256=614656
Conv3	3*3*256*384	13*13*384=64896	3*3*256*384+384=885120
Conv4	3*3*384*384	13*13*384=64896	3*3*384*384+384=1327488
Conv5	3*3*384*256	13*13*256=43264	3*3*384*256+256=884992
FC1	全连接层	4096	6*6*256*4096+4096=37752832
FC2	全连接层	4096	4096*4096+4096=16781312
FC3(Output)	全连接层和Softmax	1000	4096*1000+1000=4097000

### AlexNet参数量

# 计算机视觉中的全局和局部特征

## ■ **Reality and Illusion** / Measurement vs. Perception

Brightness

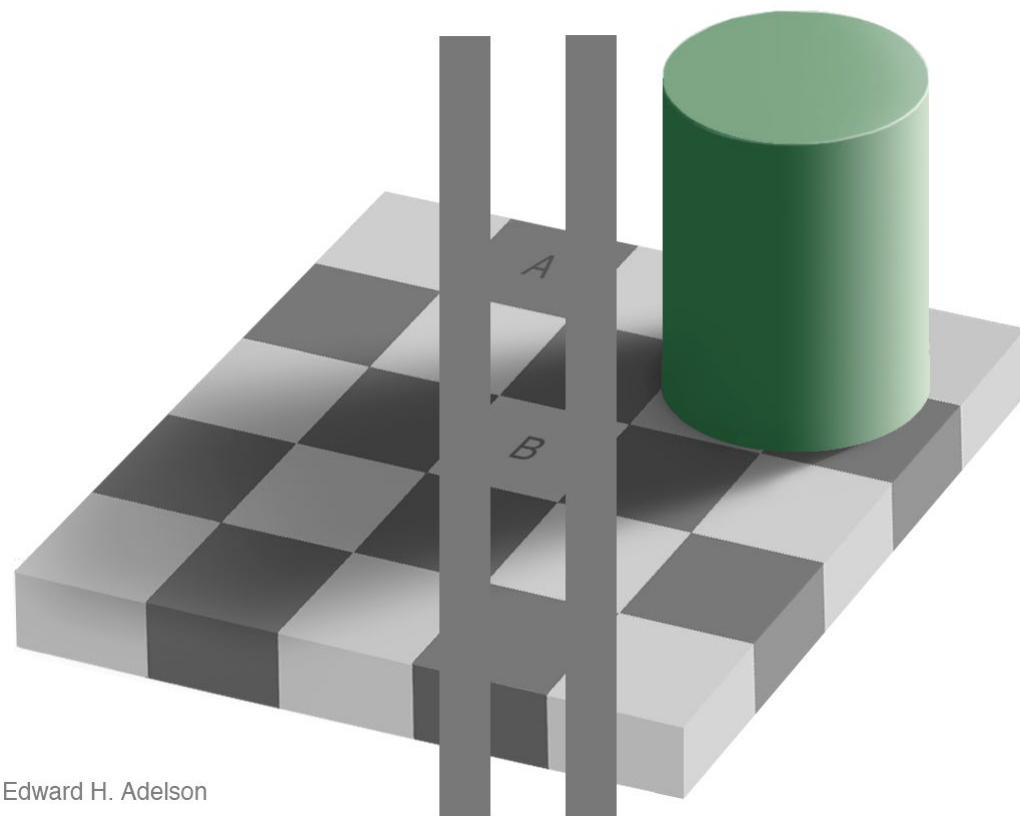


Edward H. Adelson

# 计算机视觉中的全局和局部特征

## ■ **Reality and Illusion** / Measurement vs. Perception

Brightness

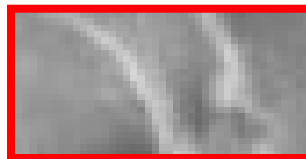
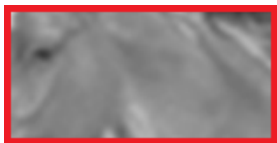


Edward H. Adelson

**Proof!**

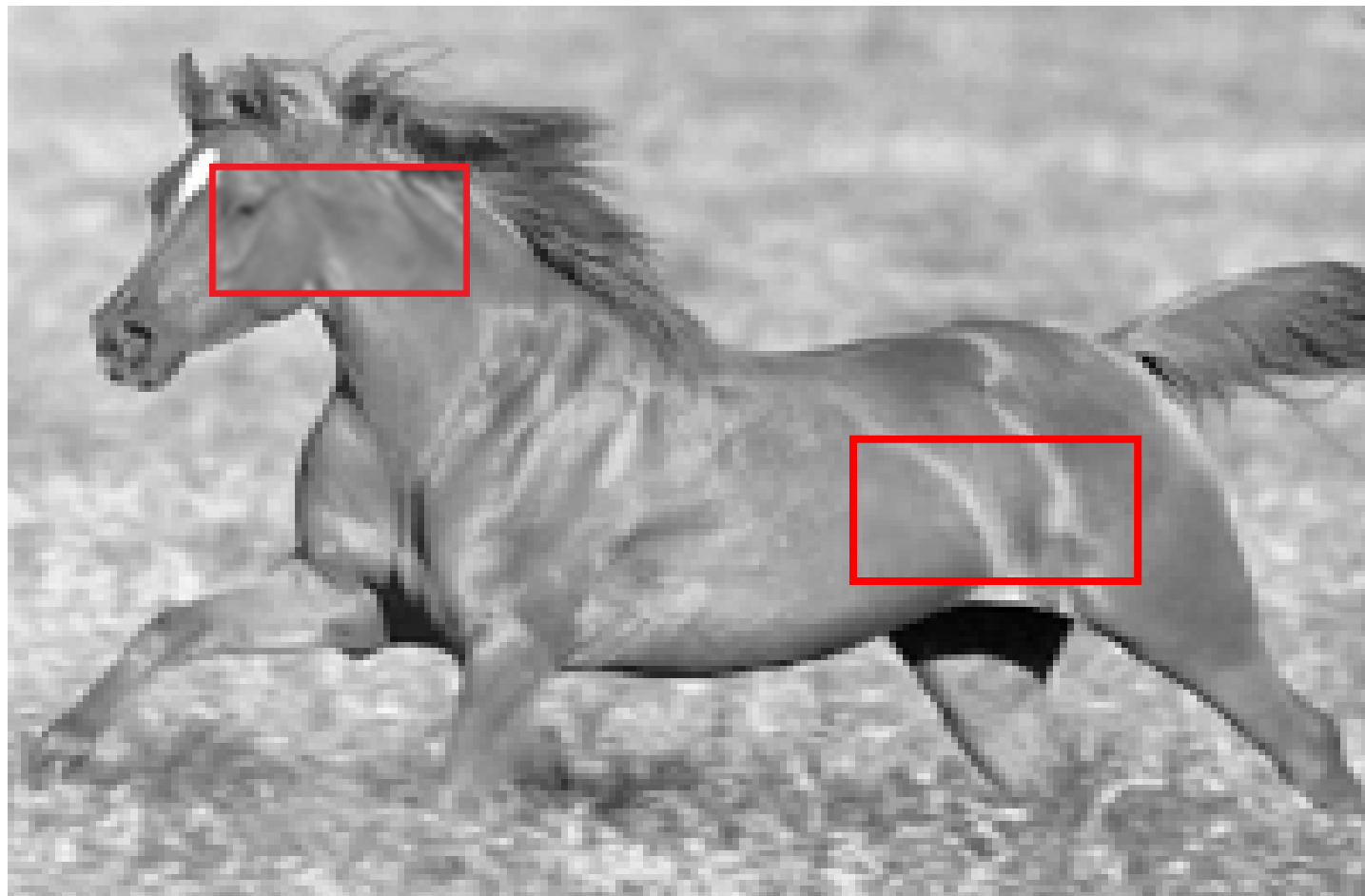
# 计算机视觉中的全局和局部特征

- **Visual Perception needs global information propagation (local & global)**



# 计算机视觉中的全局和局部特征

- **Visual Perception needs global information propagation**



### 3. Zeiler&Fergus Net (ZF-Net)

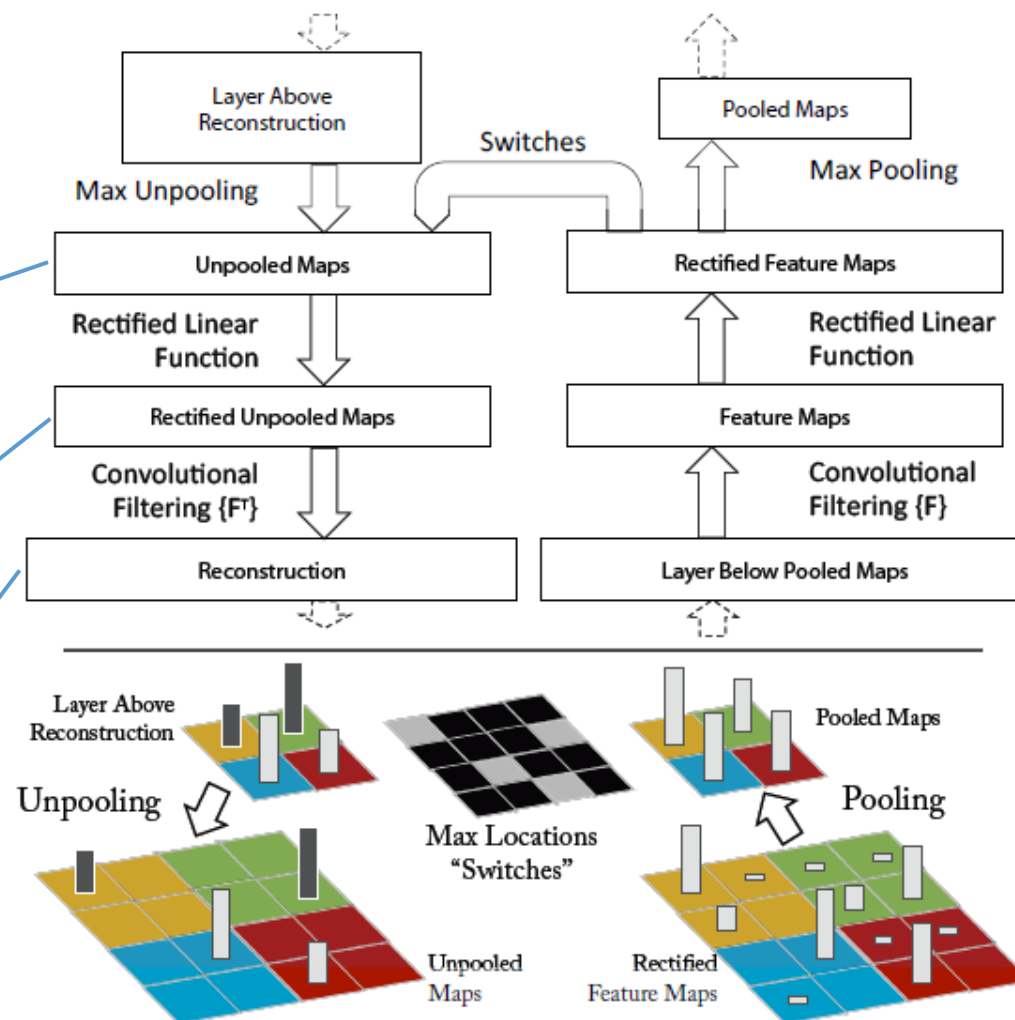
ZFNet是在AlexNet基础上进行了一些细节的改动，网络结构上并没有太大的突破。ZFNet最大的贡献在于通过使用可视化技术揭示了卷积神经网络各层到底在干什么，起到了什么作用。



### 3. Zeiler&Fergus Net

ZFNet可视化的核心是反卷积网络，由反池化层，矫正层，反卷积层组成。

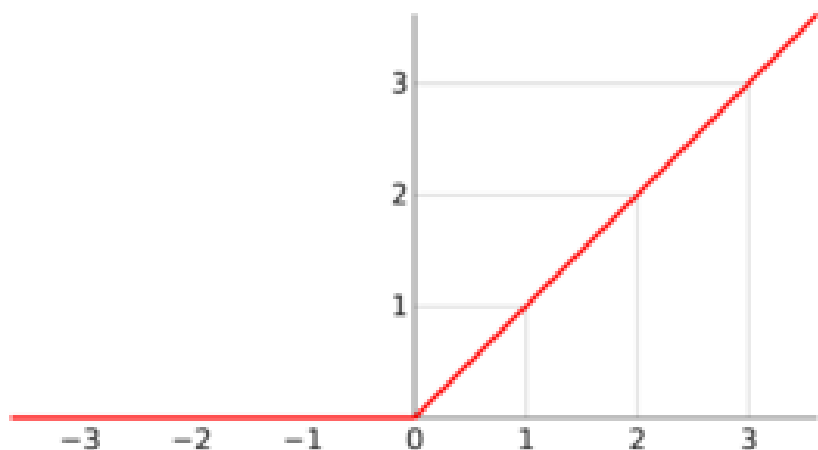
- 最大池化操作是不可逆的，在本文中用了一种近似方法来计算最大池化的逆过程：在最大池化过程中，用**Max Locations “Switches”**表格记录每一个最大值位置，在**反池化**过程中将最大值标注回记录所在位置，其余位置为0。
- 卷积网中使用**非线性的ReLU**来修改特征图来确保特征图总是正值。为了获得在每层的有效（也应该是正值）的特征重建，也在反卷积网中利用了**ReLU**。
- 卷积网使用学习到的卷积核与上层输出做卷积，得到特征。为了实现逆过程，**反卷积网**使用相同卷积核的转置作为卷积核与矫正后的特征进行卷积运算



反卷积网络的结构

# 线性整流函数（Rectified Linear Unit, ReLU）

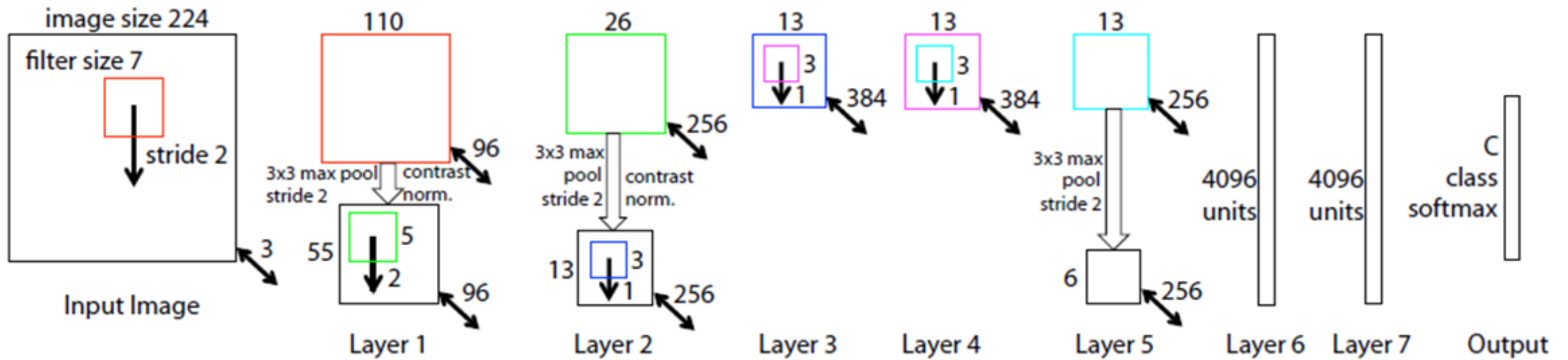
$$f(x) = \max(0, x)$$



相比Sigmoid 和 Tanh函数

- 1.更加有效率的梯度下降以及反向传播；
- 2.简化计算过程，没有了其他复杂激活函数中诸如指数函数的影响。

### 3. Zeiler&Fergus Net

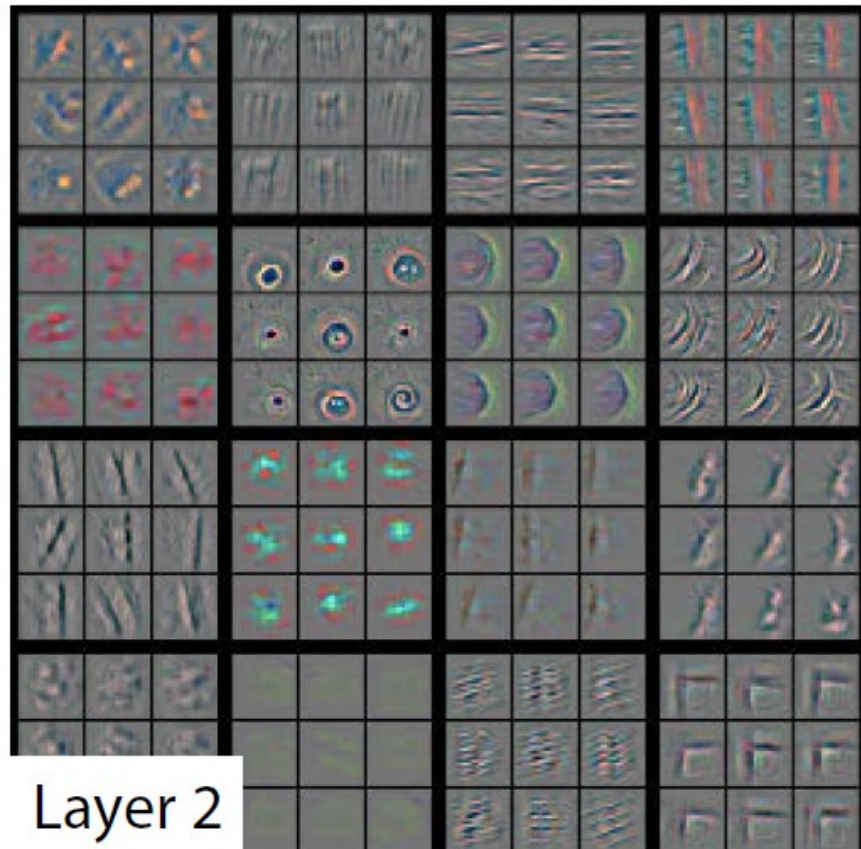
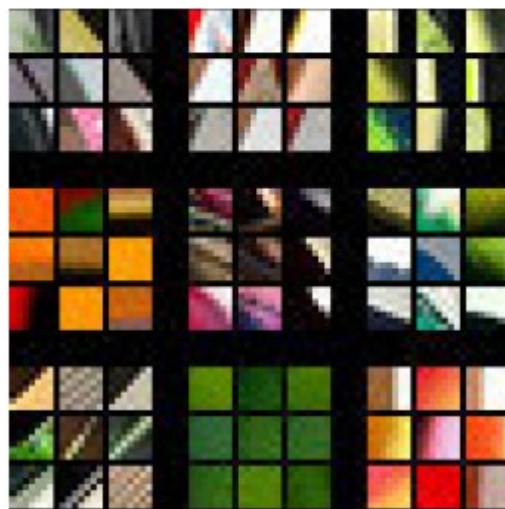


Zeiler&Fergus Net

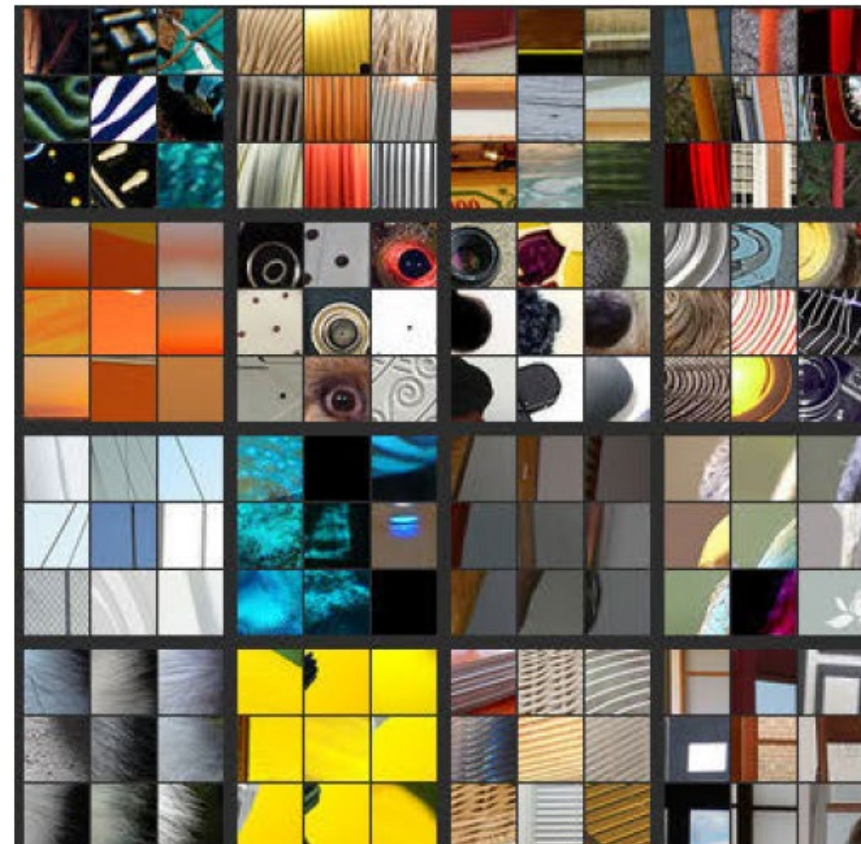
### 3. Zeiler&Fergus Net



Layer 1



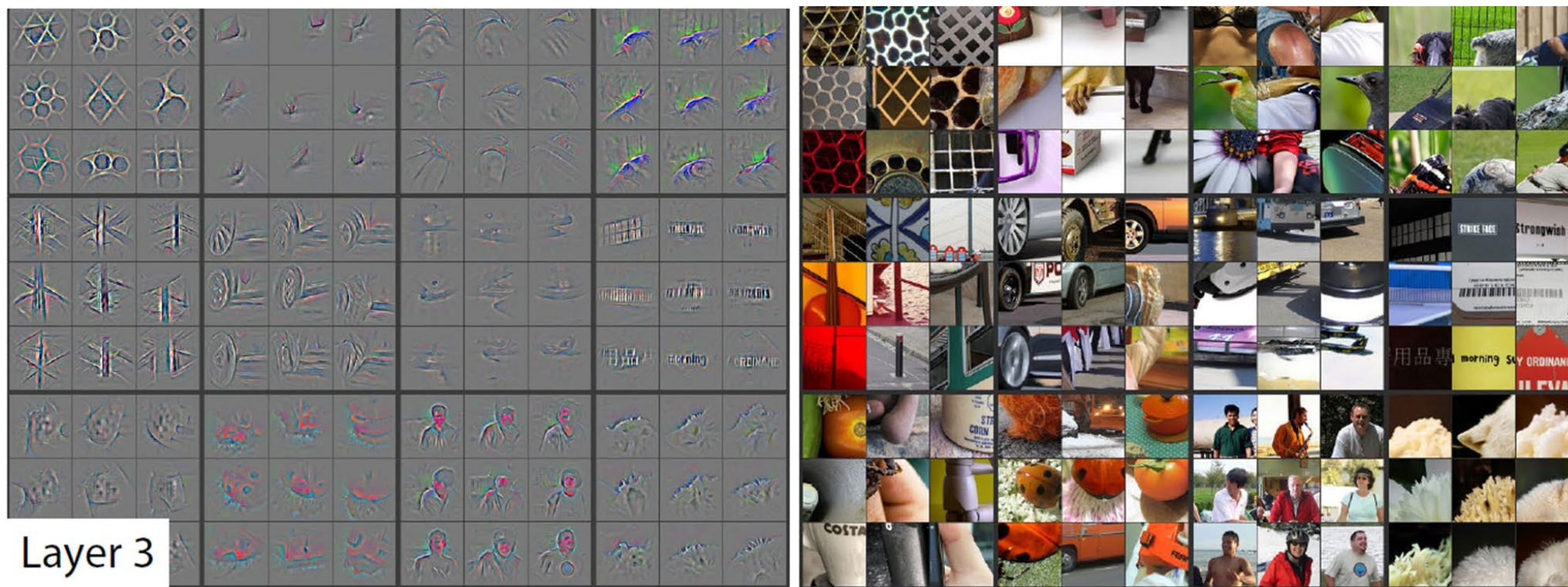
Layer 2



卷积层的可视化效果



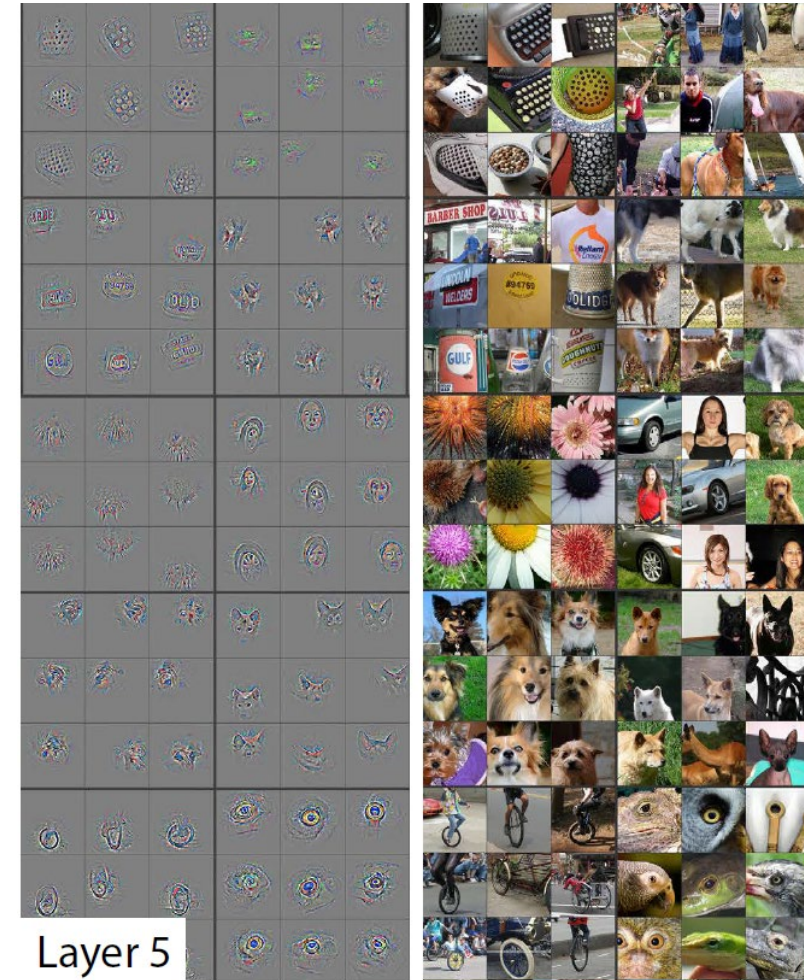
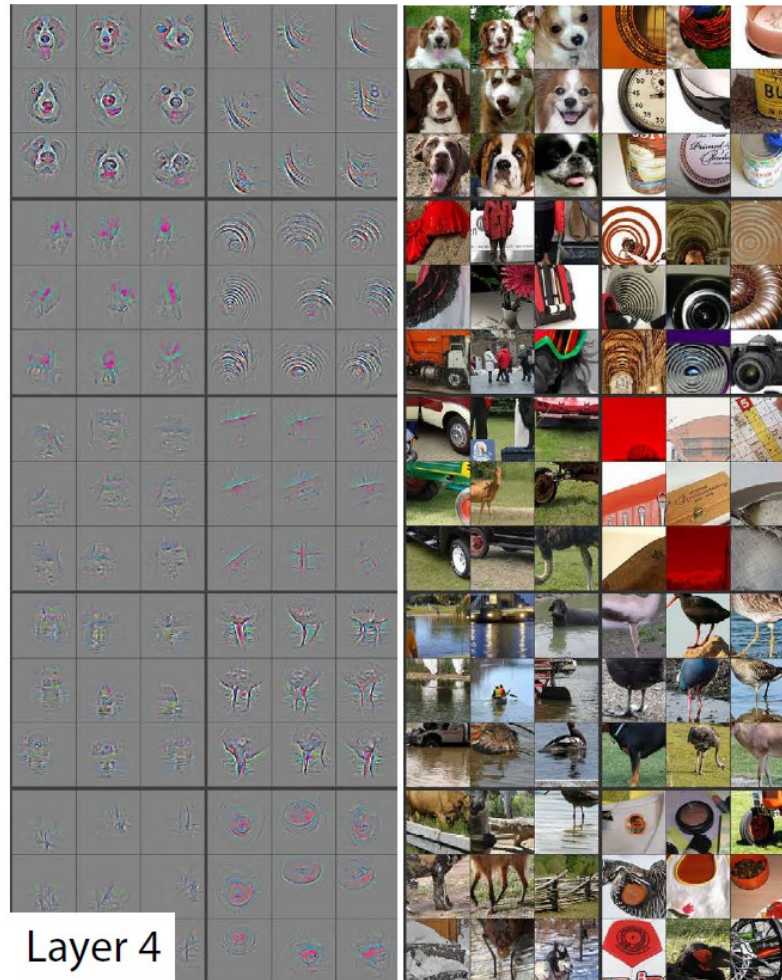
### 3. Zeiler&Fergus Net



卷积层的可视化效果



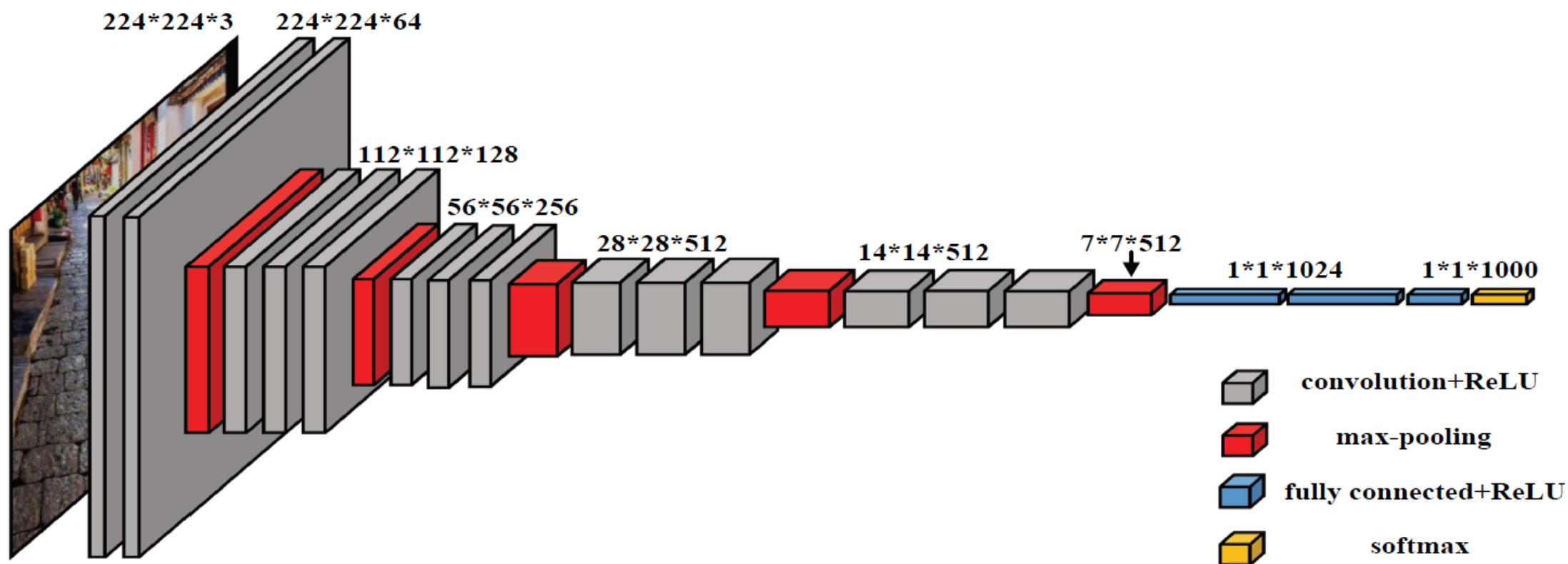
### 3. Zeiler&Fergus Net



## 4. VGGNet

VGGNet 是牛津大学计算机视觉组（Visual Geometry Group）和Google DeepMind公司一起研发出的深度卷积神经网络，VGGNet探索了卷积神经网络的深度与其性能之间的关系，成功地构筑了16~19层深的卷积神经网络，证明了增加网络的深度能够在一定程度上影响网络最终的性能，使错误率大幅下降，同时拓展性又很强。VGGNet可以看成是加深版本的AlexNet，都是由卷积层、全连接层两大部分构成。

## 4. VGGNet



VGG的网络结构



## 4. VGGNet

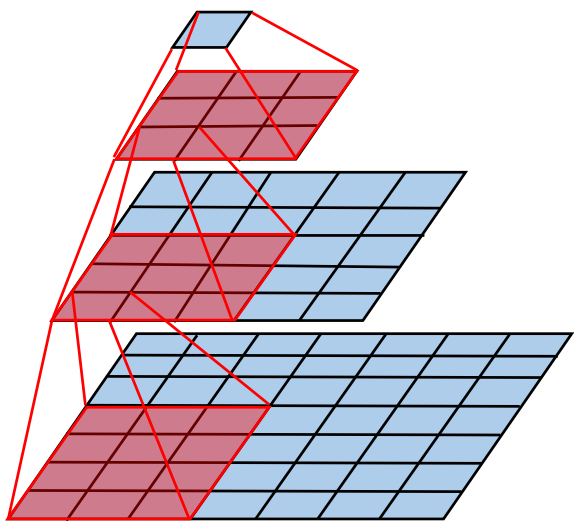
VGGNet一共有六种不同的网络结构，但是每种结构都有含有**5组卷积**，每组卷积都使用**3 x 3**的卷积核，每组卷积后进行一个**2 x 2最大池化**，接下来是三个全连接层。

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

VGG16 VGG19

## 4. VGGNet

- VGGNet的优点:**
- 结构简洁，VGG由5层卷积层、3层全连接层、softmax输出层构成，层与层之间使用max-pooling（最大化池）分开，所有隐层的激活单元都采用ReLU函数。
  - VGG使用多个较小卷积核（3x3）的卷积层代替一个卷积核较大的卷积层，一方面可以减少参数，另一方面相当于进行了更多的非线性映射，可以增加网络的拟合表达能力。



3个3\*3卷积核代替一个7\*7卷积核

例如：3个 $3 \times 3$ 的卷积核的叠加相当于一个 $7 \times 7$ 的卷积核，而且参数更少。大约是 $7 \times 7$ 卷积核卷积层的 $(3 \times 3 \times 3)/(7 \times 7) = 0.55$ 倍，而且拥有和 $7 \times 7$ 卷积核一样的感受视野，3个卷积层的叠加，经过了更多次的非线性变换，对特征的学习能力更强。

## 4. VGGNet

- 采用较小的池化核，相比AlexNet的3x3的池化核，VGG全部采用2x2的池化核。
- 通道数变多，VGG网络第一层的通道数为64，后面每层都进行了翻倍，最多到512个通道，通道数的增加，使得更多的信息可以被提取出来。
- 层数更深，特征图更深（？）。由于卷积核专注于扩大通道数、池化专注于缩小宽和高，使得模型架构上更深更宽的同时，控制了计算量的增加规模。