



# 模式识别与深度学习 (28-30)

## 卷积神经网络-1

左旺孟

综合楼712

视觉感知与认知组

哈尔滨工业大学计算机学院

[cswmzuo@gmail.com](mailto:cswmzuo@gmail.com)

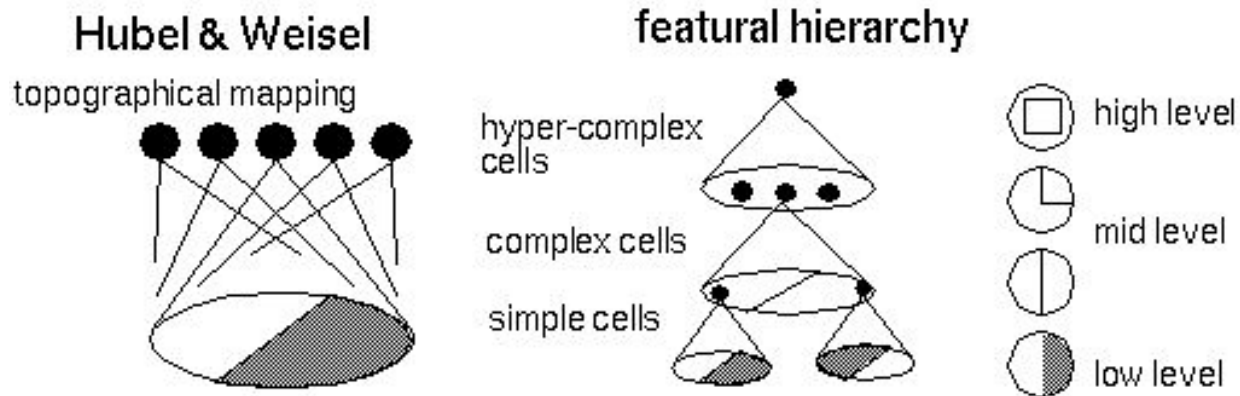
13134506692

# 卷积神经网络

- 历史和动机
- 基本操作
  - 卷积、池化、归一化、卷积神经网络
- 新进展
  - 3x3、Dilated Convolution
- 典型网络架构
  - LeNet、AlexNet、VGGNet、Inception
  - ResNet、SENet、DenseNet、Attention

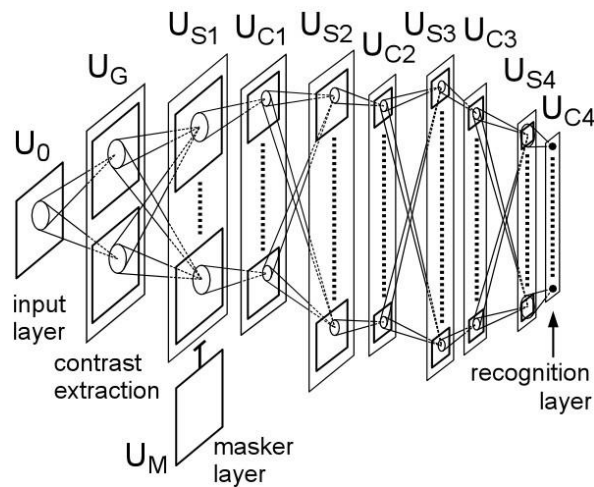
# Hubel/Wiesel架构

- D. Hubel and T. Wiesel (1959, 1962, Nobel Prize 1981)
  - 视觉皮层：包括 *simple*, *complex*, and *hyper-complex* 细胞

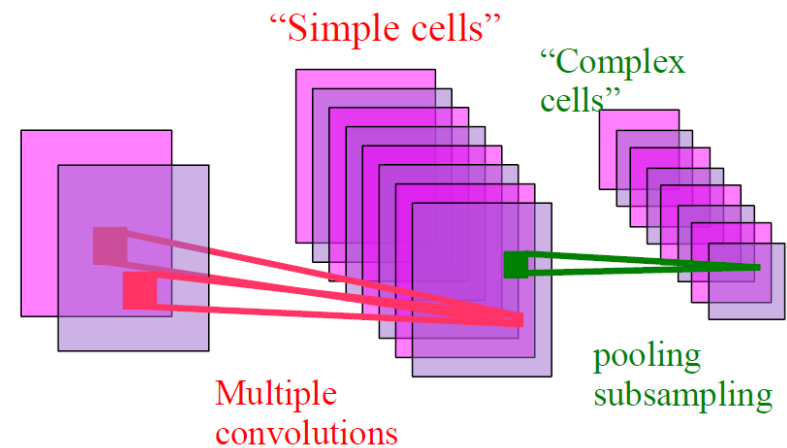


# Neocognitron

- [Hubel & Wiesel 1962]:
  - 简单细胞：局部特征检测
  - 复杂细胞：简单特征输出的聚合



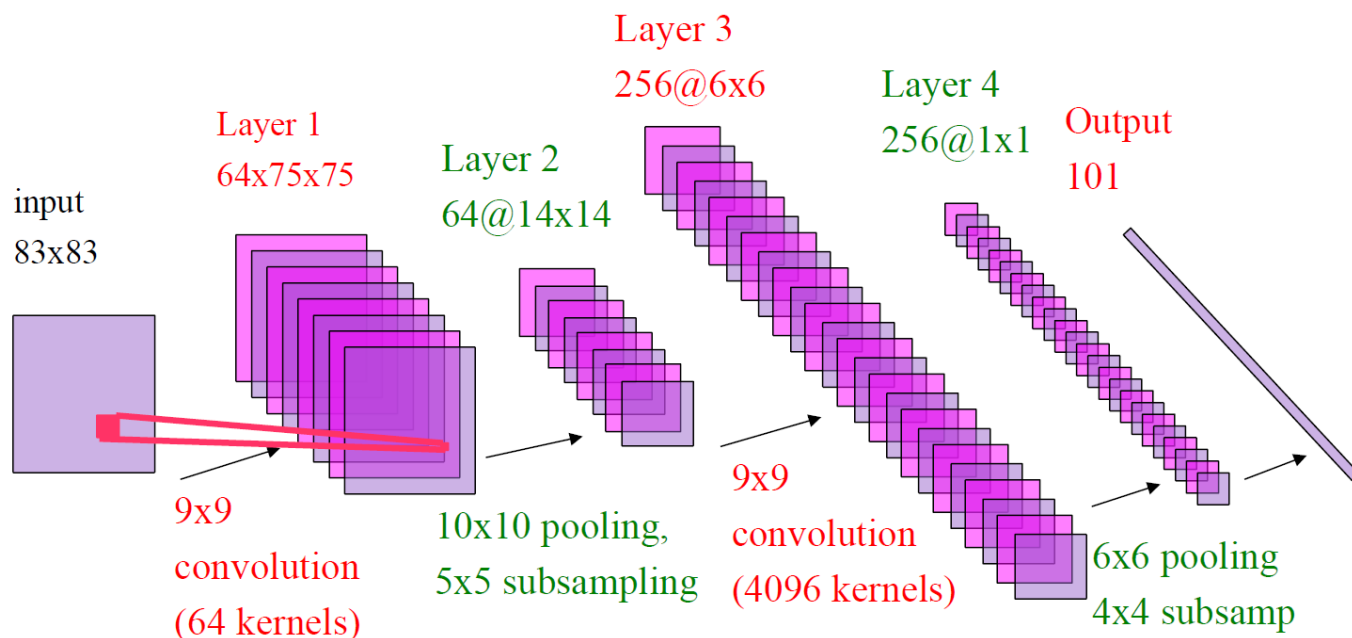
Cognitron & Neocognitron  
[Fukushima 1974-1982]



卷积神经网络

# 卷积神经网络（上半场）

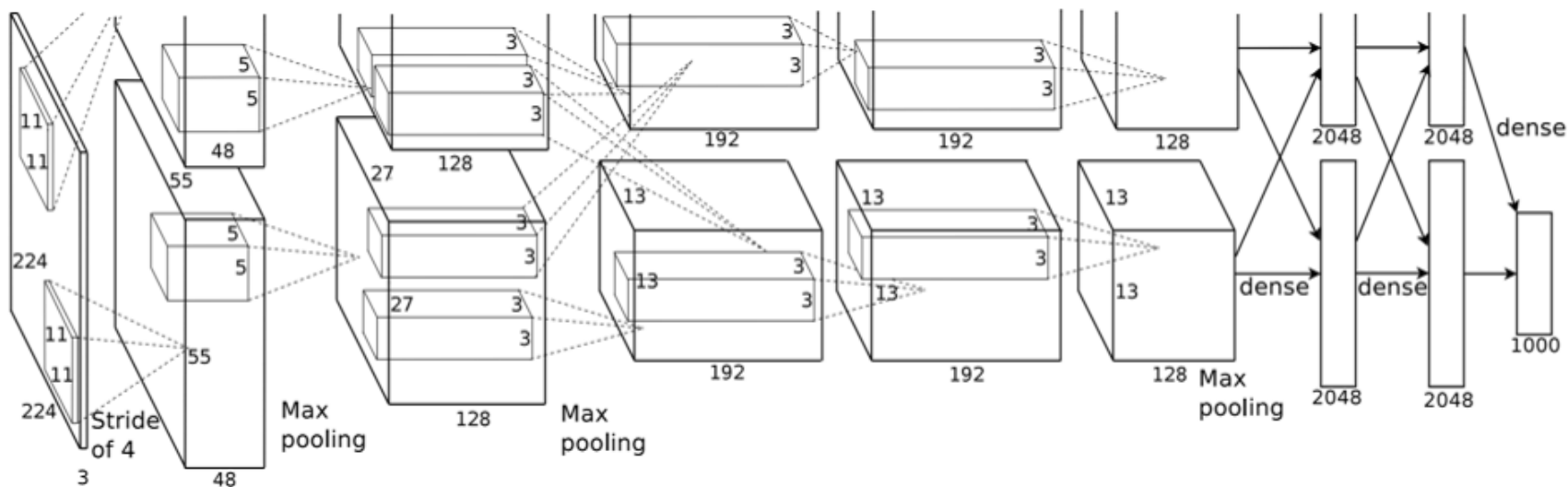
- LeCun et al., NIPS 1989



- 2个卷积层一个全连接层

# 深度卷积神经网络（下半场）

- Krizhevsky et al. NIPS 2012

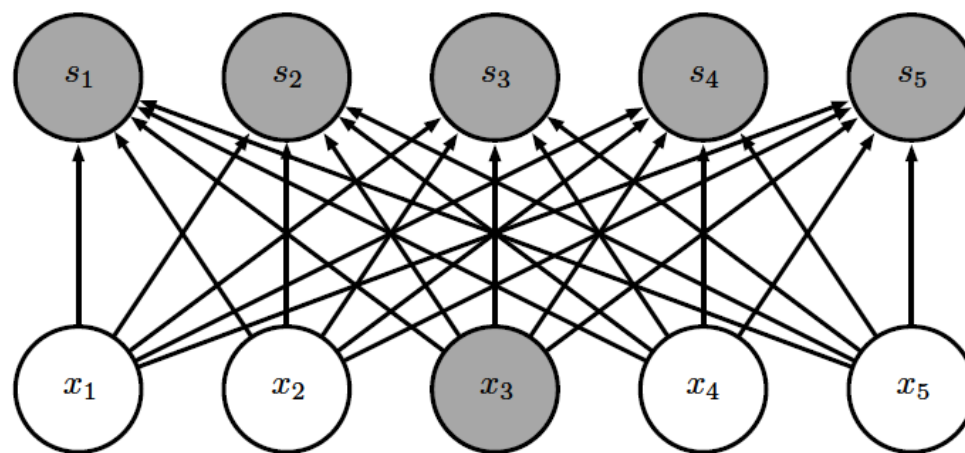
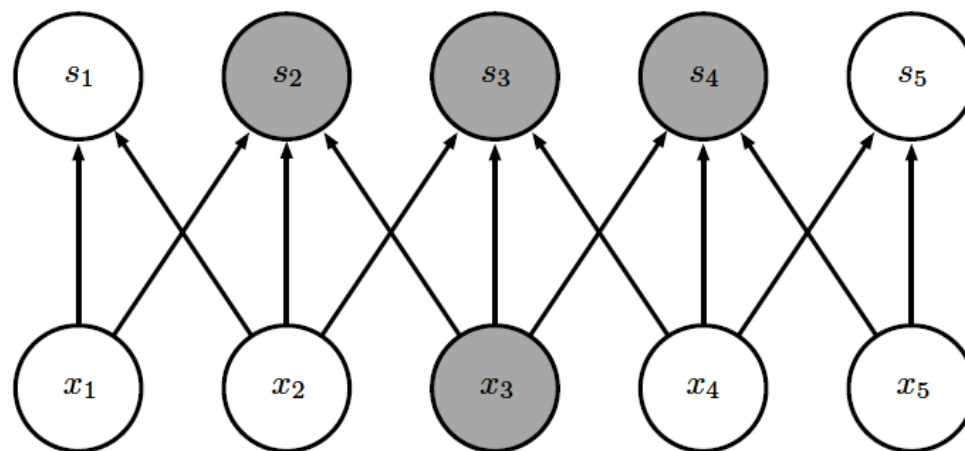


- 5个卷积层、3个全连接层

动机：

- 稀疏交互 (sparse interactions)
- 参数共享 (parameter sharing)
- 等变表示 (equivariant representation)
  - 不变表示 (invariant representation)

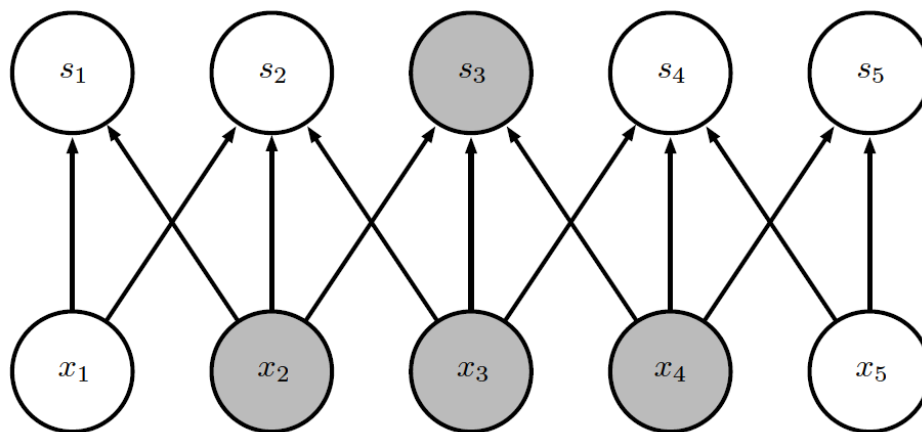
# 稀疏交互（稀疏连接）



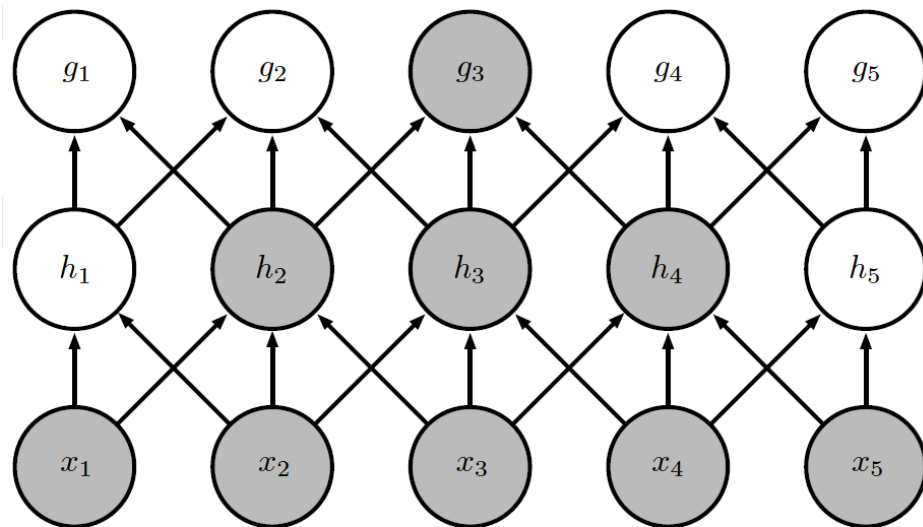


# 感受野 (Receptive Field)

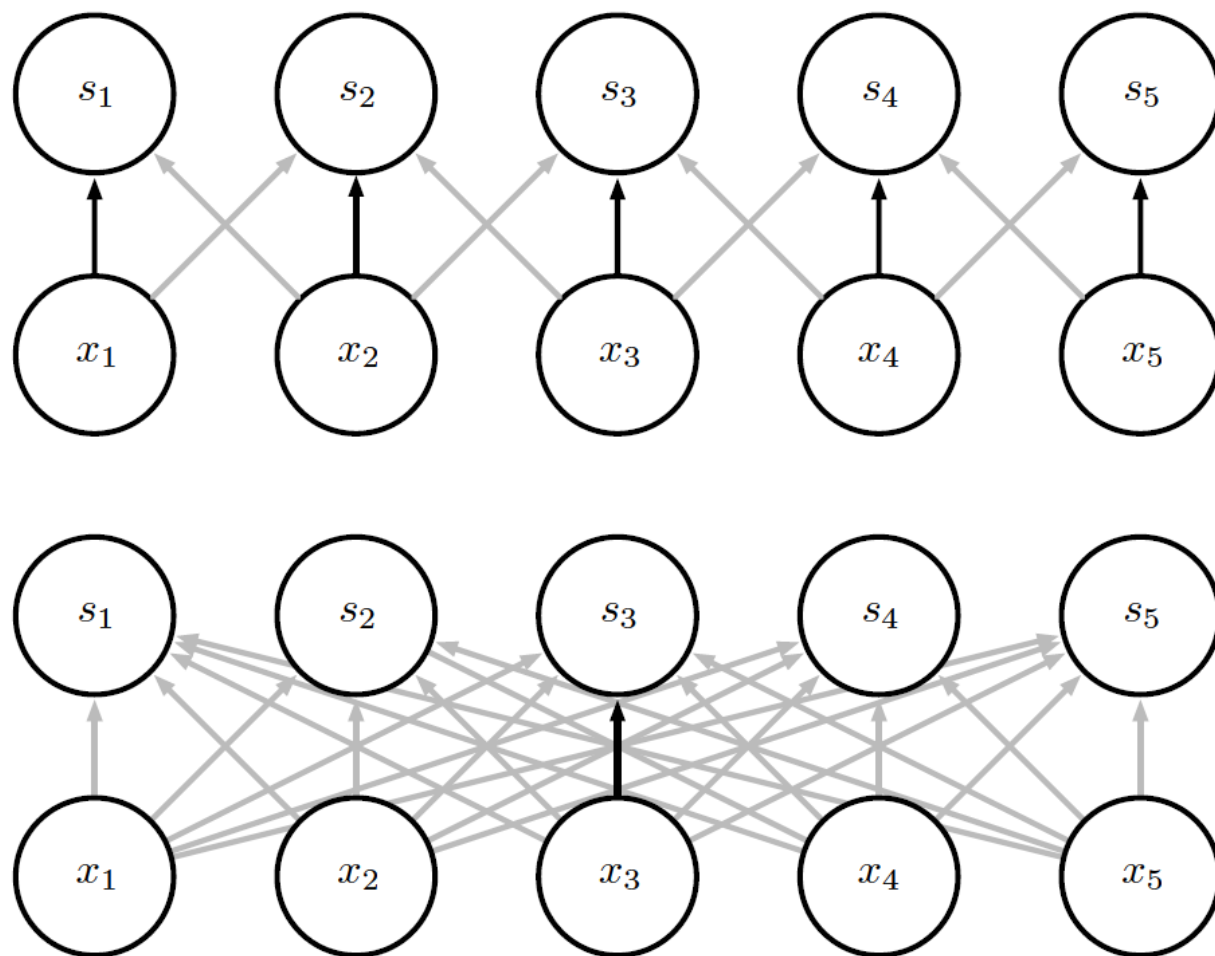
• 层1:



• 层2:



# 参数共享



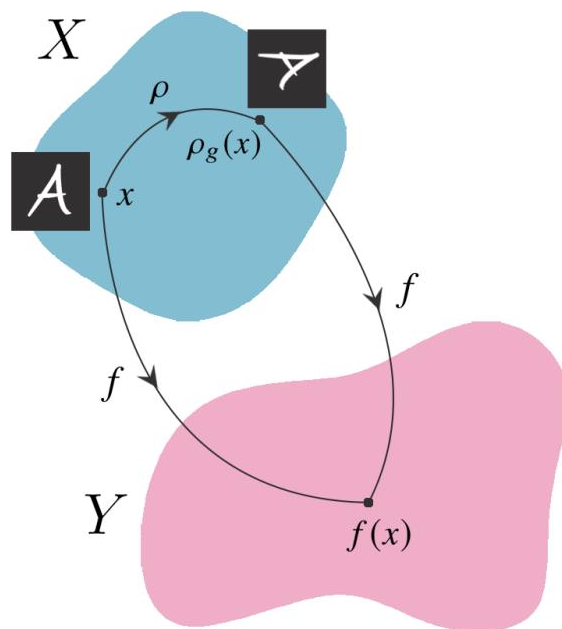
# 等变表示 (Equivariant Representation)

- 如果一个函数满足输入改变，输出也以同样方式进行改变的话，我们称它是等变的
- 卷积：平移等变
- 不变表示 (Transform-invariant)

# 等变与不变表示 (Equivariant/Invariant)

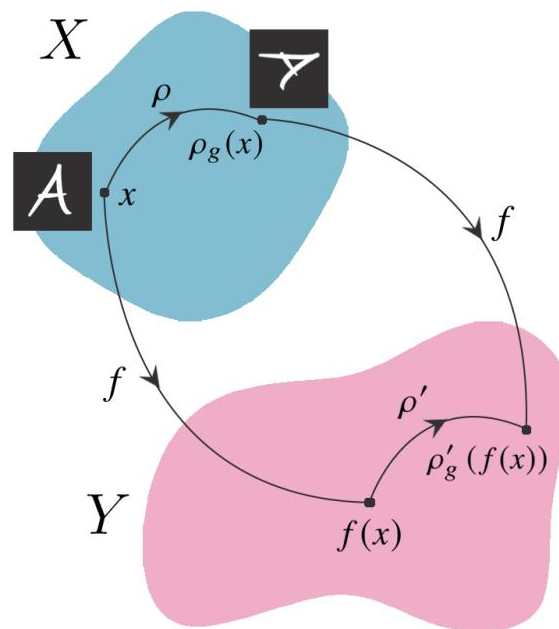
## Invariance

$$f(\rho_g(x)) = f(x)$$



## Equivariance

$$f(\rho_g(x)) = \rho'_g(f(x))$$



# 卷积神经网络

- 历史和动机
- 基本操作
  - 卷积、池化、归一化、卷积神经网络
- 新进展
  - 3x3、空洞卷积 ( dilated convolution )
- 典型网络架构
  - LeNet、AlexNet、VGGNet、Inception
  - ResNet、SENet、DenseNet、Attention

# 卷积：参数共享和稀疏连接

- 连续卷积

$$s(t) = \int x(a)w(t-a)da$$

$$s(t) = (x * w)(t).$$

- 输入、核函数

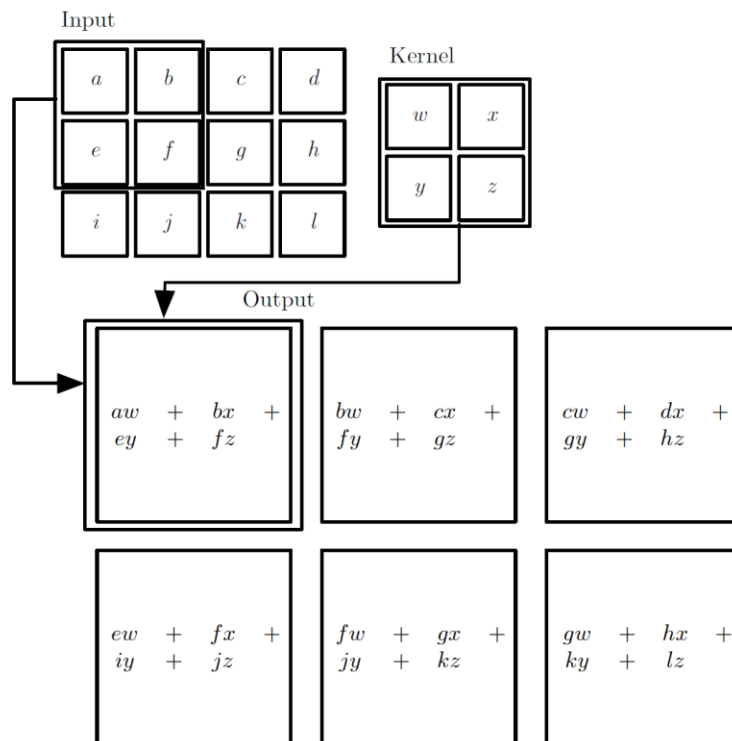
- 离散卷积

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t-a)$$

# 二维卷积

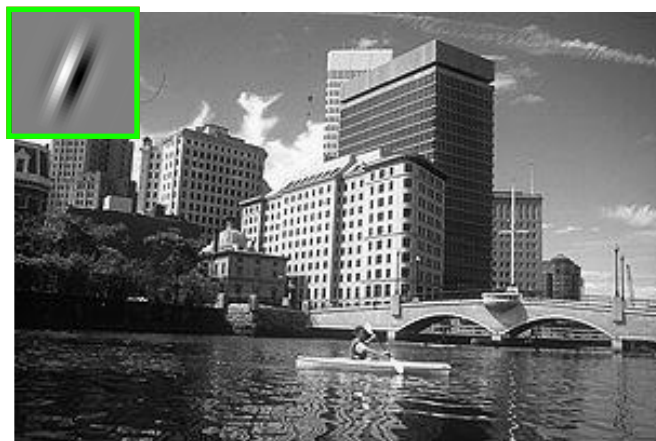
- 二维卷积

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



# 直观展示

- 卷积核(Conv. Kernel)
- 特征图(Feature Map)



Input



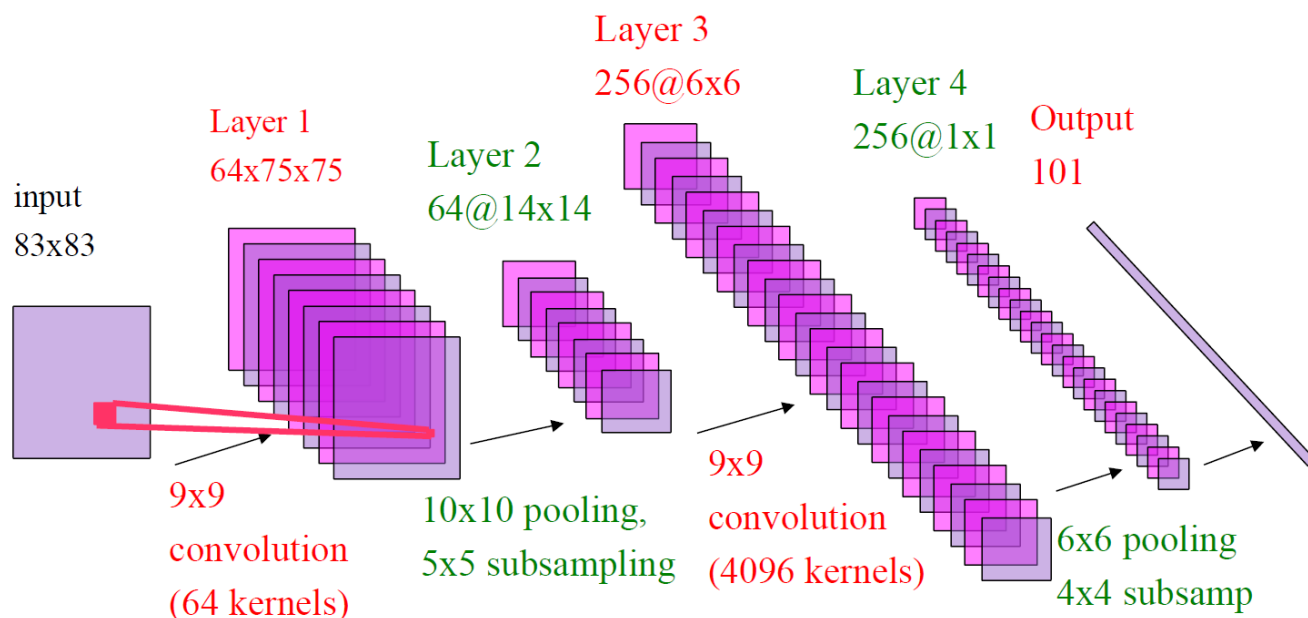
Feature Map



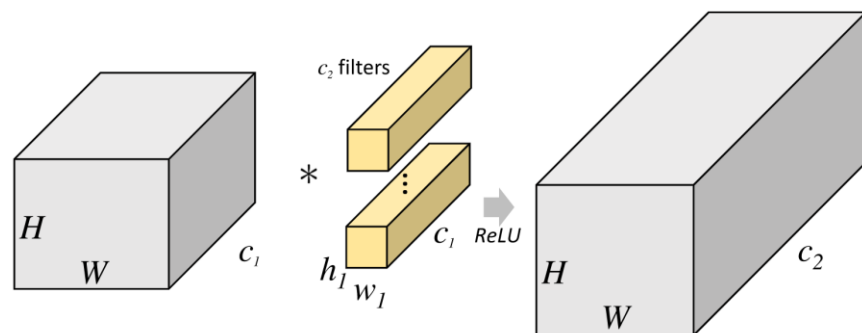
# 拓展：多通道卷积

- 多通道卷积

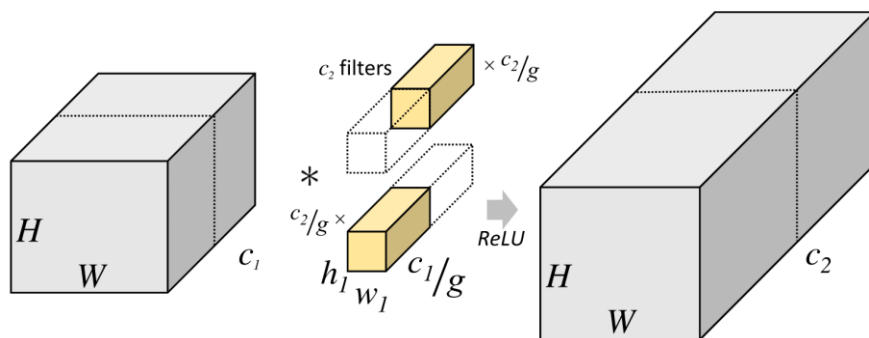
$$F_i = \sum_{j=1}^C w_{j,i} * x_j$$



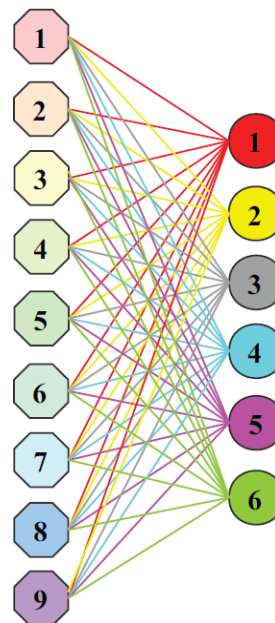
# 拓展: Grouped Convolution



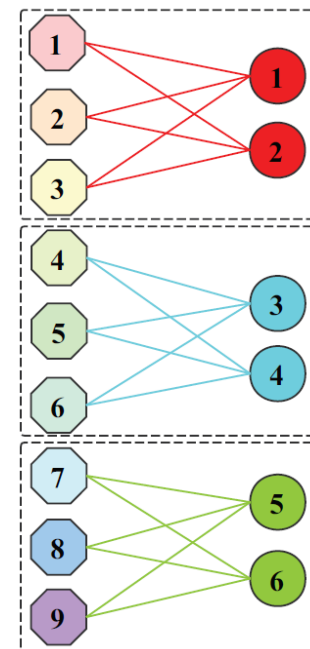
标准卷积



Grouped Convolution



标准卷积



Grouped Convolution

# 拓展：步幅（Stride）

## • 步幅为1

$$\begin{array}{|c|c|c|c|} \hline a_{11} & a_{12} & a_{13} & a_{14} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} \\ \hline \end{array} \quad * \quad \begin{array}{|c|c|c|} \hline k_{11} & k_{12} & k_{13} \\ \hline k_{21} & k_{22} & k_{23} \\ \hline k_{31} & k_{32} & k_{33} \\ \hline \end{array} \quad = \quad \begin{array}{|c|c|} \hline a_{11}k_{11}+a_{12}k_{12}+a_{13}k_{13} & a_{12}k_{11}+a_{13}k_{12}+a_{14}k_{13} \\ \hline + & + \\ a_{21}k_{21}+a_{22}k_{22}+a_{23}k_{23} & a_{22}k_{21}+a_{23}k_{22}+a_{24}k_{23} \\ \hline + & + \\ a_{31}k_{31}+a_{32}k_{32}+a_{33}k_{33} & a_{32}k_{31}+a_{33}k_{32}+a_{34}k_{33} \\ \hline \end{array} \dots$$

$I \qquad K \qquad S$

## • 步幅为2

$$\begin{array}{|c|c|c|c|c|} \hline a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} & \\ \hline \end{array} \quad * \quad \begin{array}{|c|c|c|} \hline k_{11} & k_{12} & k_{13} \\ \hline k_{21} & k_{22} & k_{23} \\ \hline k_{31} & k_{32} & k_{33} \\ \hline \end{array} \quad = \quad \begin{array}{|c|c|} \hline a_{11}k_{11}+a_{12}k_{12}+a_{13}k_{13} & a_{13}k_{11}+a_{14}k_{12}+a_{15}k_{13} \\ \hline + & + \\ a_{21}k_{21}+a_{22}k_{22}+a_{23}k_{23} & a_{23}k_{21}+a_{24}k_{22}+a_{25}k_{23} \\ \hline + & + \\ a_{31}k_{31}+a_{32}k_{32}+a_{33}k_{33} & a_{33}k_{31}+a_{34}k_{32}+a_{35}k_{33} \\ \hline \end{array} \dots$$

$I \qquad K \qquad S$

## 拓展：边界条件

- 特征图尺寸逐渐减小
- 零填充（ Zero Padding ）、镜像填充
- 其他方式： Partial Conv.

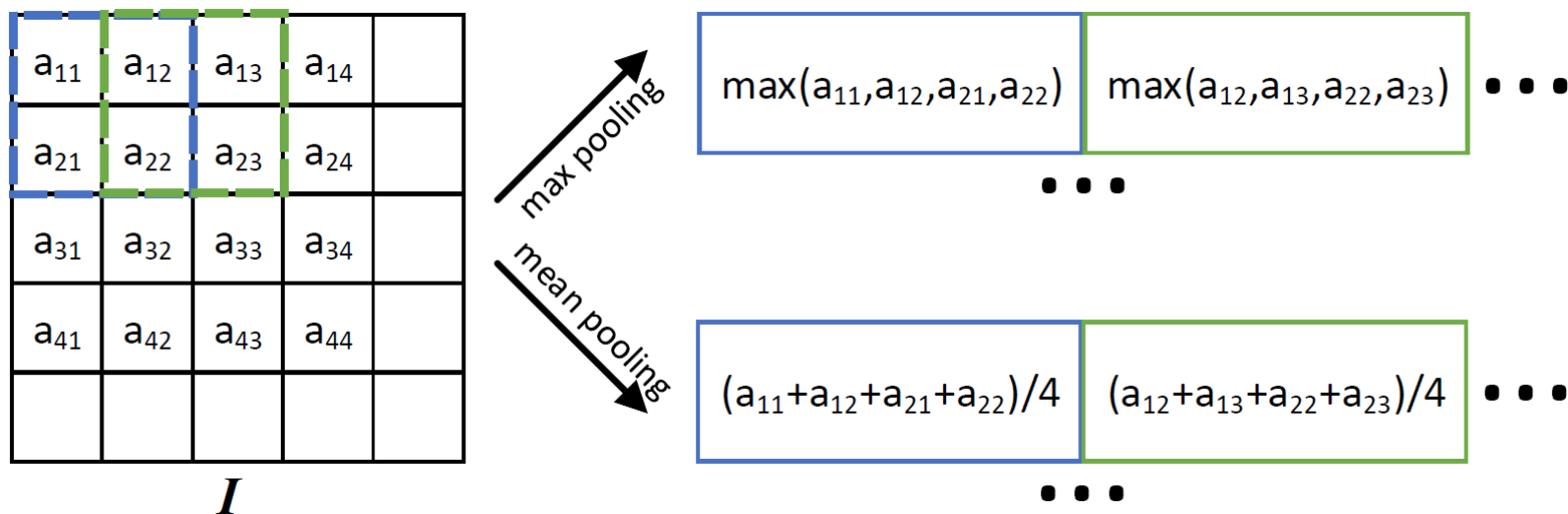
Guilin Liu, Kevin J. Shih, Ting-Chun Wang, Fitsum A. Reda, Karan Sapra, Zhiding Yu, Andrew Tao, Bryan Catanzaro, [Partial Convolution based Padding](#), arXiv:1811.11718 .

# 卷积神经网络

- 历史和动机
- 基本操作
  - 卷积、池化、归一化、卷积神经网络
- 新进展
  - 3x3、空洞卷积（dilated convolution）
- 典型网络架构
  - LeNet、AlexNet、VGGNet、Inception
  - ResNet、SENet、DenseNet、Attention

# 池化：形变不敏感

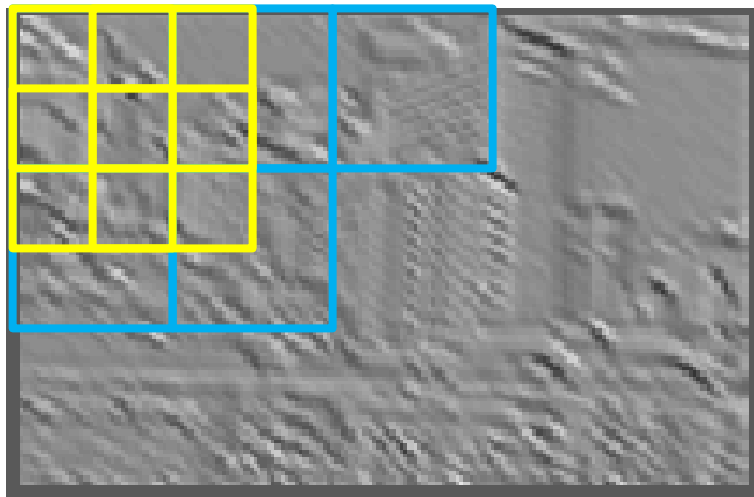
## • 池化



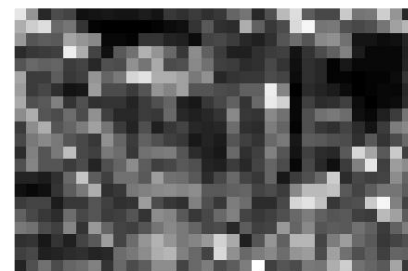
## • 下采样

# 直观展示

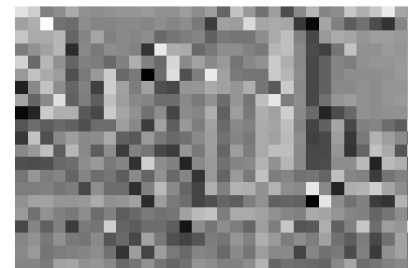
- Max Pooling
- Average Pooling



Max

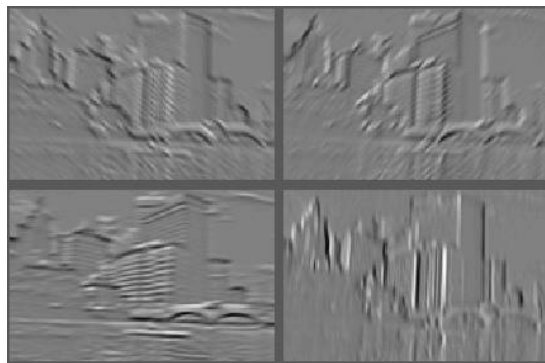


Sum

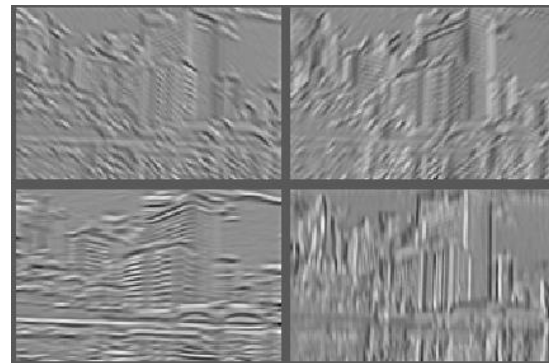


# 归一化：光照不敏感

- 每个channel或所有channel归一化
- 池化前或池化后归一化



Feature Maps



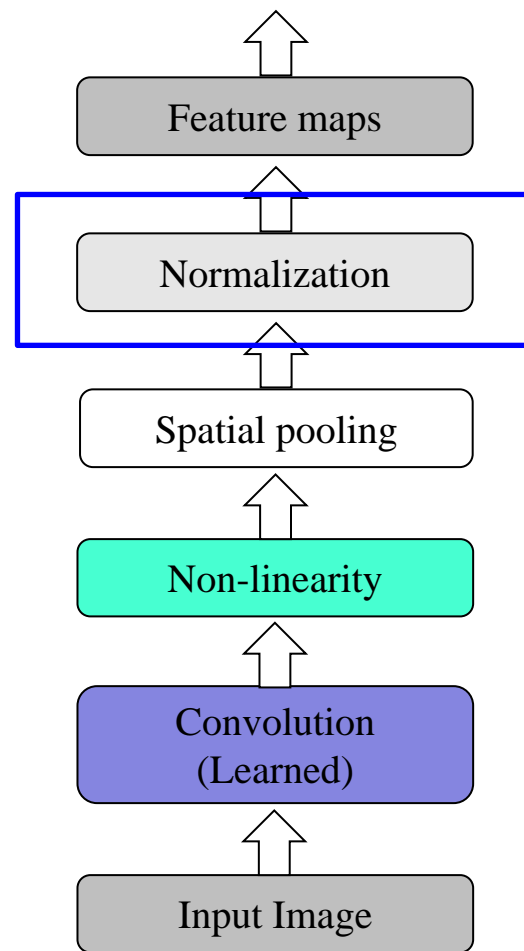
Feature Maps  
After Contrast Normalization

- 已不太常用或结合Batch Normalization



## 总结：CNN网络层

1. 卷积
2. 非线性激活函数
3. 池化
4. 归一化

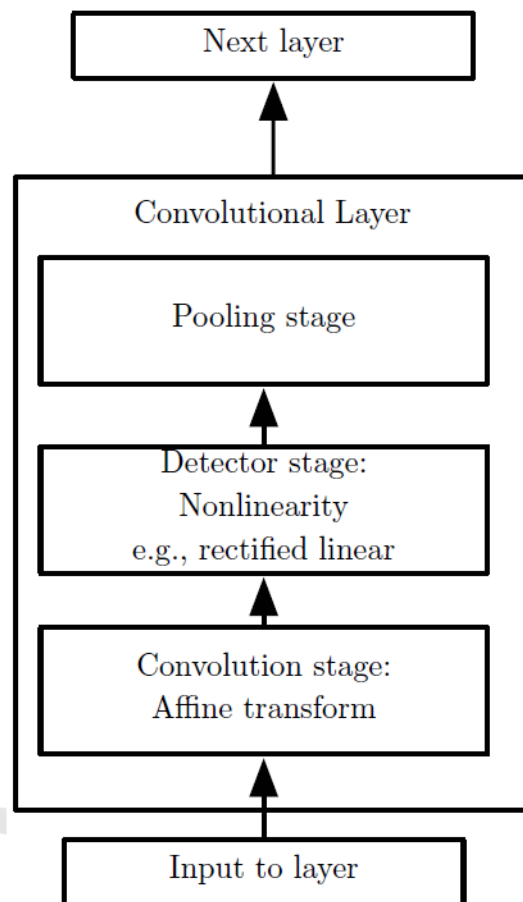


# 卷积神经网络

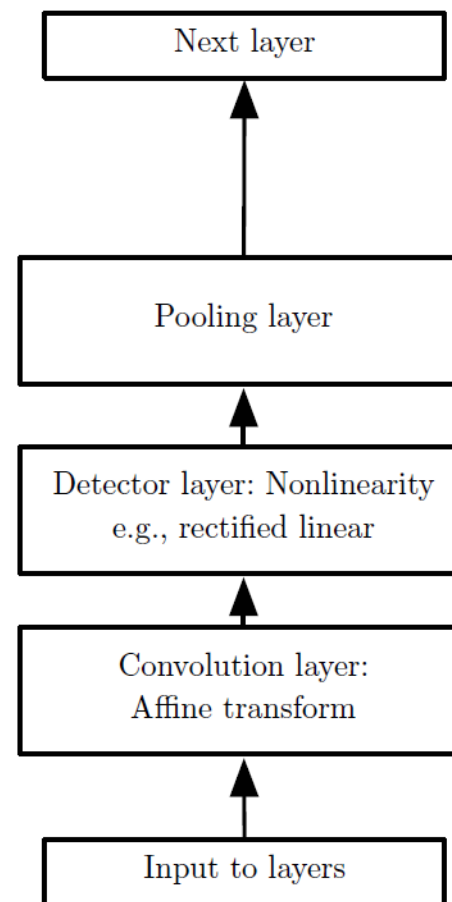
- 历史和动机
- 基本操作
  - 卷积、池化、 归一化、 卷积神经网络
- 新进展
  - 3x3、空洞卷积 ( dilated convolution )
- 典型网络架构
  - LeNet、 AlexNet、 VGGNet、 Inception
  - ResNet、 SENet、 DenseNet、 Attention

# 典型的卷积神经网络层

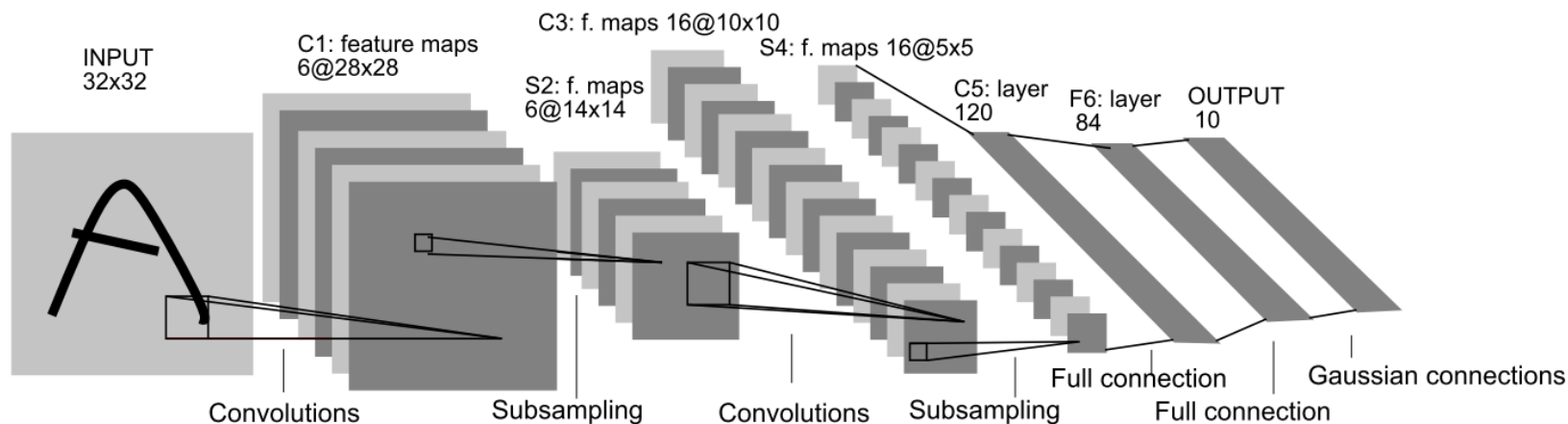
Complex layer terminology



Simple layer terminology

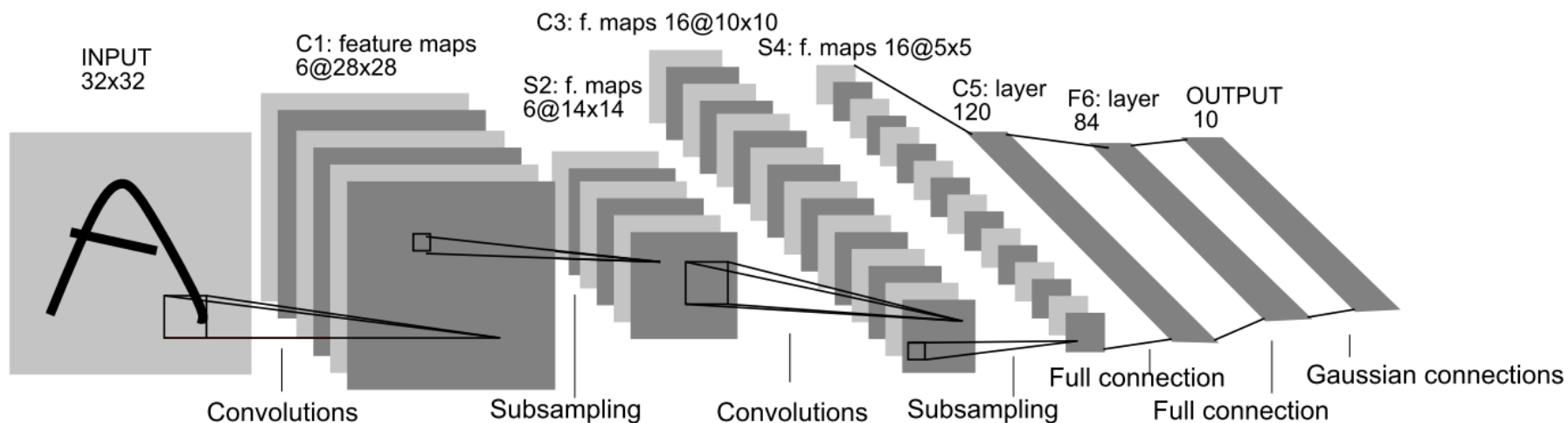


# 卷积神经网络示例：LeNet5（1998）



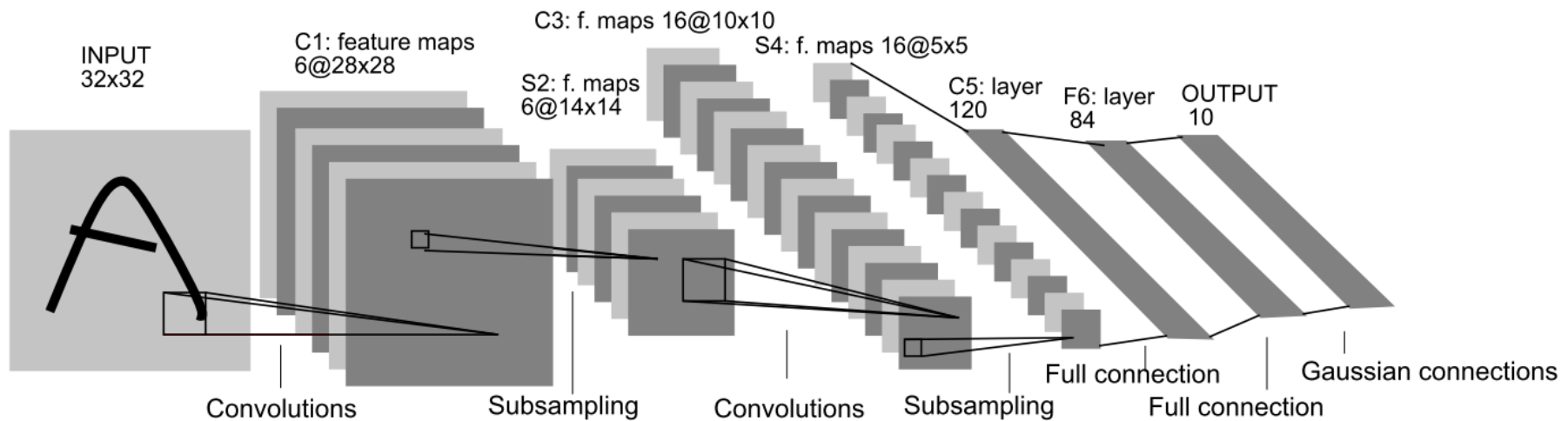
- 输入: 32x32图像
- C<sub>x</sub>: 卷积层
- S<sub>x</sub>: 下采样层
- F<sub>x</sub>: 全连接层

# LeNet 5, Layer C1



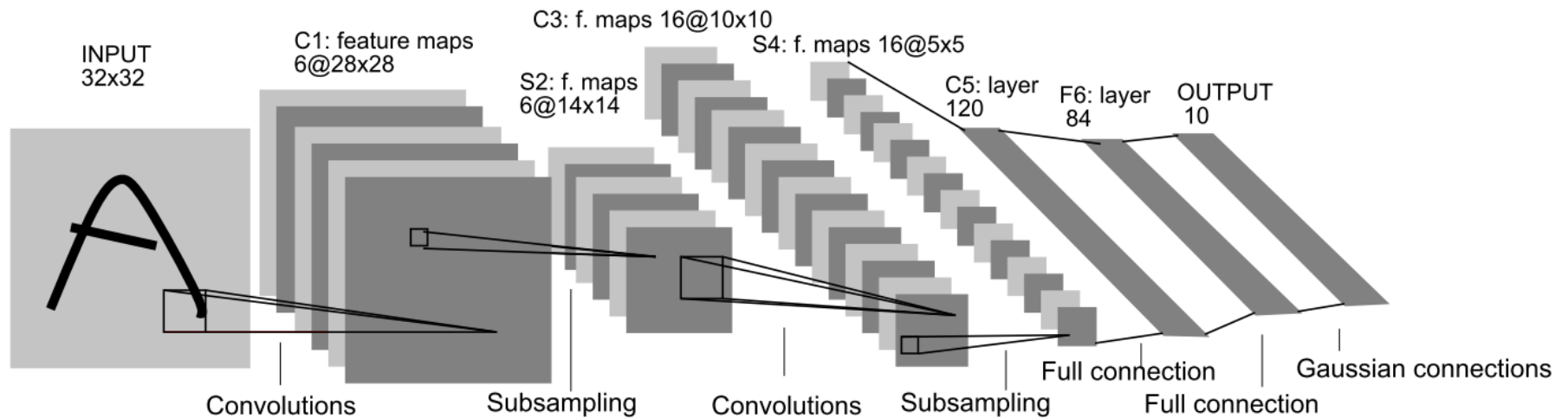
- C1: 卷积层，通道数为6，特征图大小 28x28. 卷积核 5x5.
  - 稀疏连接
  - 参数共享：参数量： $(5*5+1)*6=156$   
 非参数共享： $28*28*(5*5+1)*6=122304$   
 全连接： $(32*32+1)*(28*28)*6$  parameters

# LeNet 5, Layer S2



- S2: 下采样层，6通道，特征图大小14x14
- 2x2 感受野
- 学习参数:  $6 * 2 = 12$ .
- 全连接:  $14 * 14 * (2 * 2 + 1) * 6 = 5880$

# LeNet 5, Layer C3



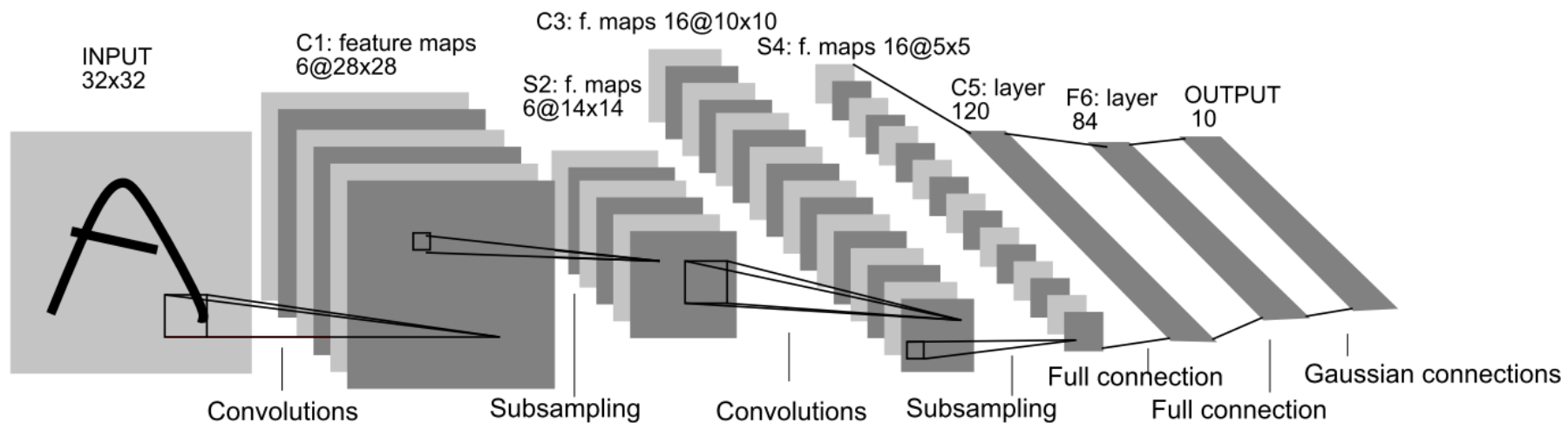
- C3: 卷积层，通道数16，特征图大小 10x10
- Each unit in C3 is connected to several! 5x5 receptive fields at identical locations in S2
- 参数量: 1516.
- 全连接: 151600

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X	X		X
2	X	X	X				X	X	X			X		X	X	X
3			X	X	X			X	X	X	X		X		X	X
4				X	X	X			X	X	X	X		X	X	X
5					X	X	X			X	X	X	X		X	X

TABLE I

EACH COLUMN INDICATES WHICH FEATURE MAP IN S2 ARE COMBINED BY THE UNITS IN A PARTICULAR FEATURE MAP OF C3.

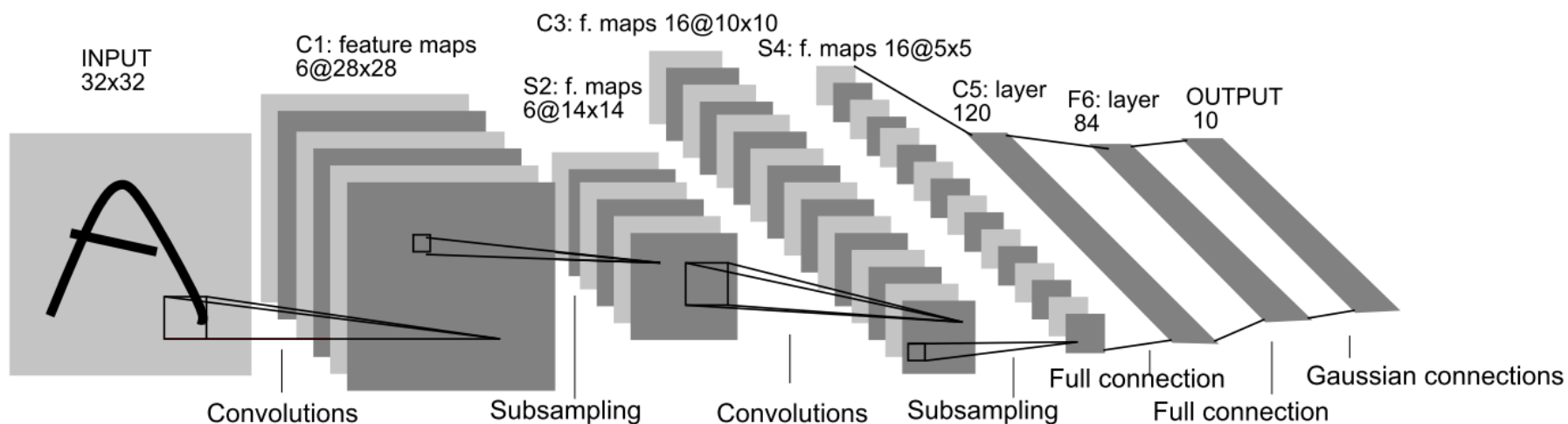
# LeNet 5, Layer S4



- S4: 下采样层, with 16 feature maps of size 5x5
- 感受野: 2x2
- 参数量:  $16 * 2 = 32$ .
- 全连接:  $5 * 5 * (2 * 2 + 1) * 16 = 2000$

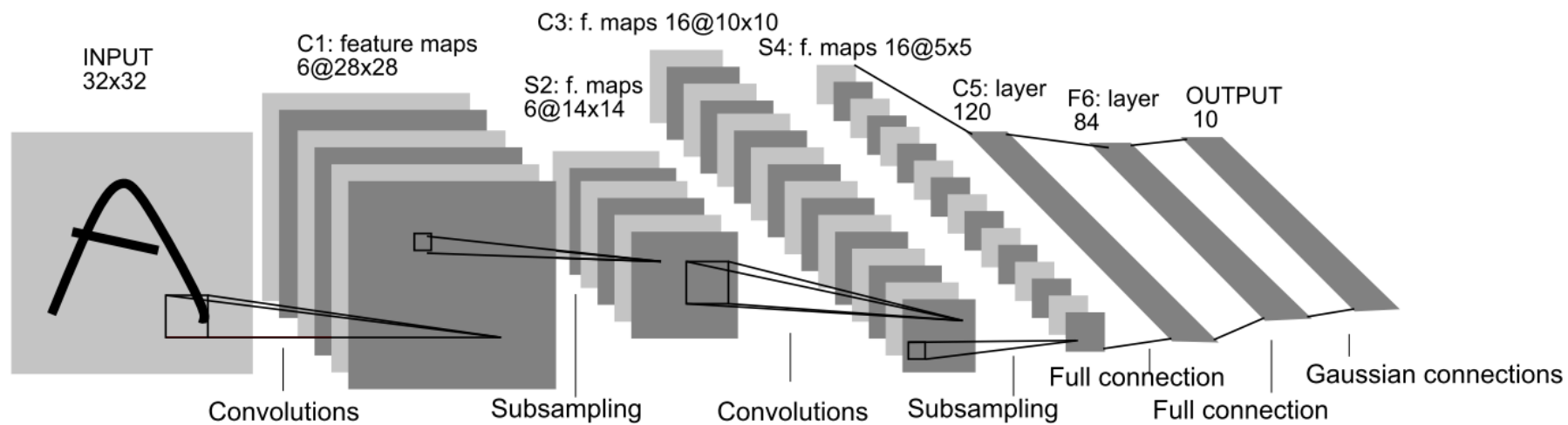


# LeNet 5, Layer C5



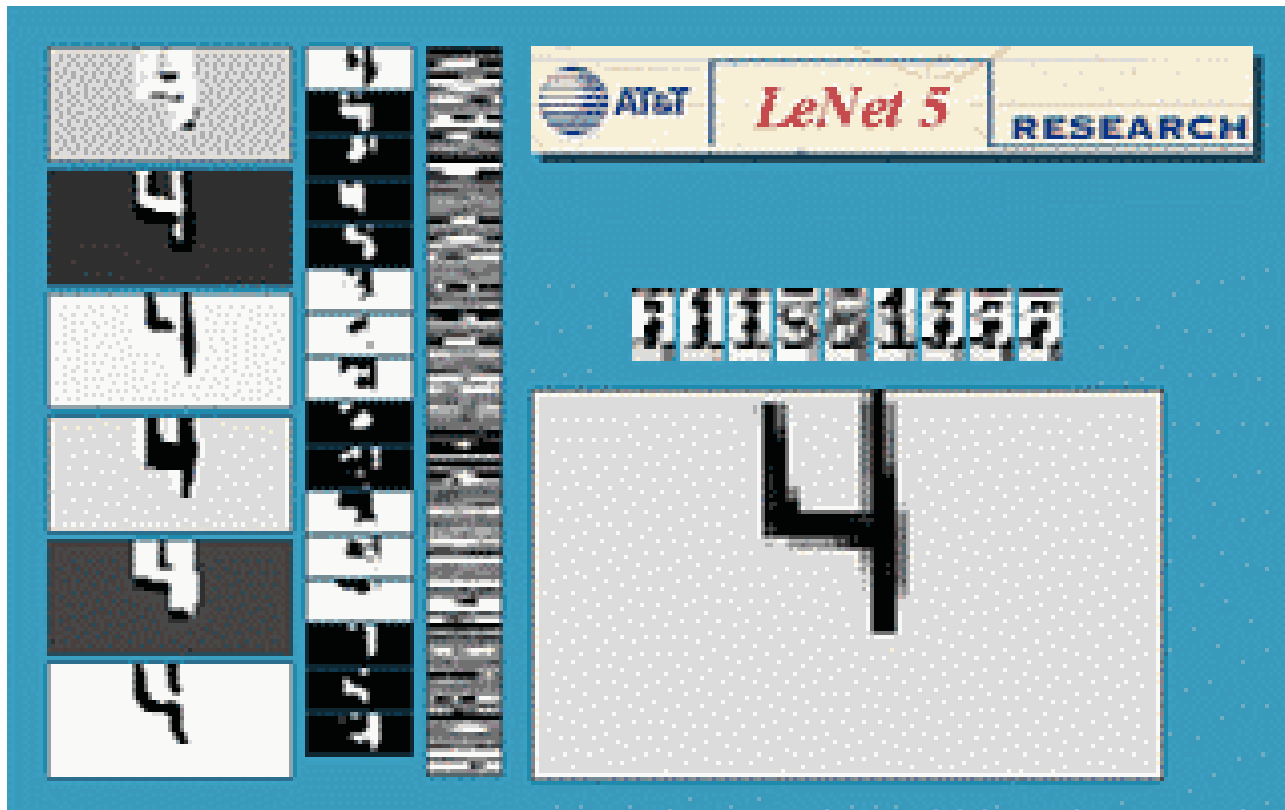
- C5: 卷积层，通道数120，特征图大小1x1
- 感受野 5x5
- 参数量:  $120 \times (16 \times 25 + 1) = 48120$

# LeNet 5, Layer F6

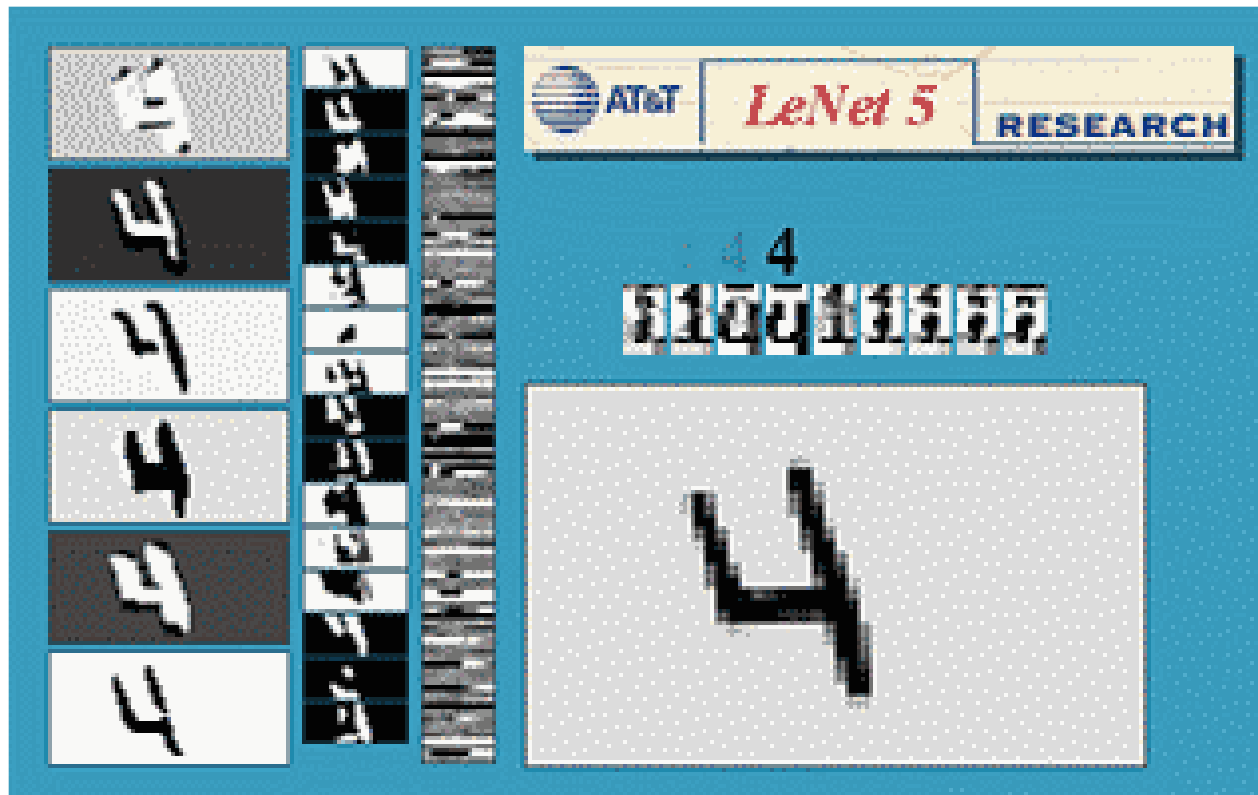


- Layer F6: 全连接层，特征位数84
- 参数量:  $84 \times (120 + 1) = 10164$ .
- 输出: 10 RBF (One for each digit)
- 学习算法: BP

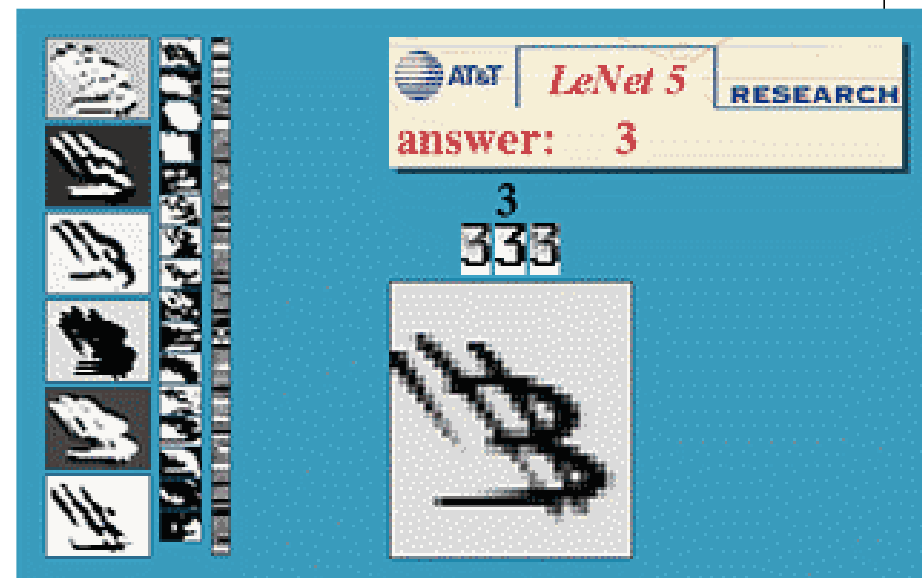
# LeNet 5, Shift invariance



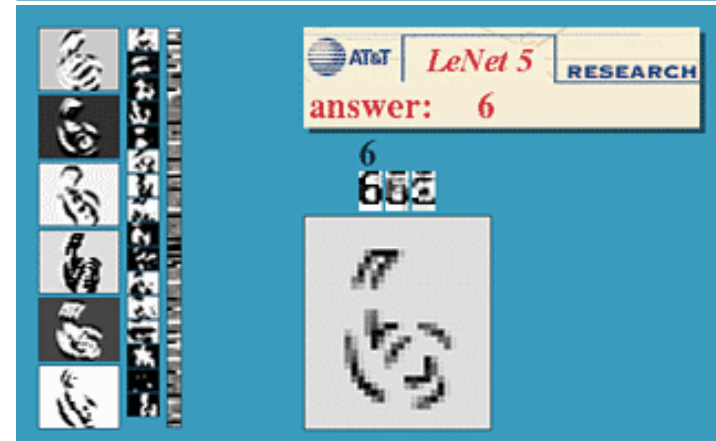
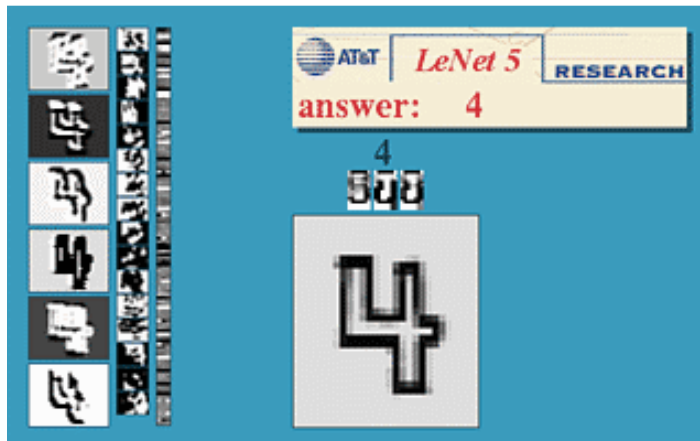
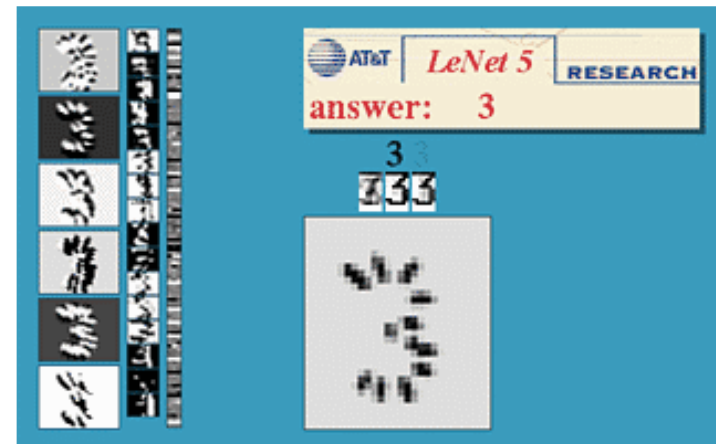
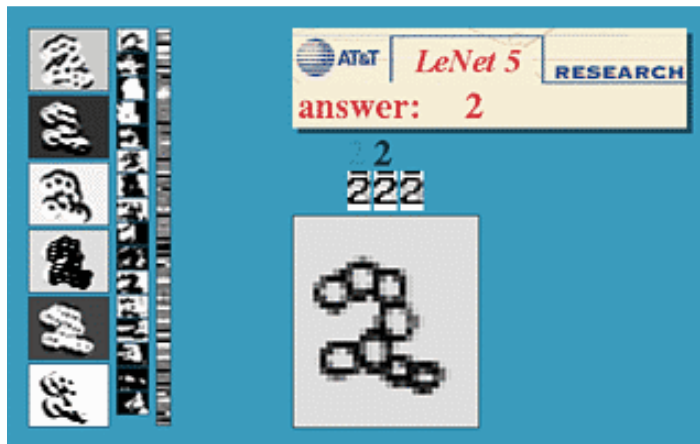
# LeNet 5, Rotation invariance



# LeNet 5, Noise resistance



# LeNet 5, Unusual Patterns



# 卷积神经网络

- 历史和动机
- 基本操作
  - 卷积、池化、归一化、卷积神经网络
- 新进展：操作
  - 3x3、dilated convolution
- 典型网络架构：思想和网络结构
  - LeNet、AlexNet、VGGNet、Inception
  - ResNet、SENet、DenseNet、Attention