



北京交通大学

图像处理与机器学习

Digital Image Processing and Machine Learning

主讲人：黄琳琳

电子信息工程学院



第七章 深度学习基础

- ◆ 深度学习引言
- ◆ 卷积神经网络
- ◆ 几种典型网络
- ◆ 问题及方向



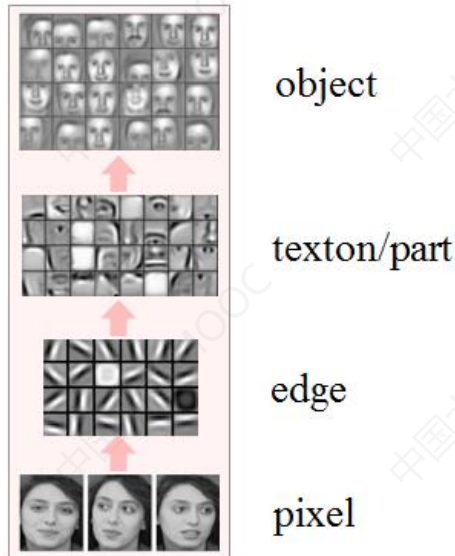
卷积神经网络

◆ 动物视觉系统对外界的感知

- 从局部到全局
- 视觉皮层的每个神经元
- 只响应某些特定区域的刺激。
 - 感受野 (Receptive fields)
 - 信息分层处理机制

卷积神经网络

(Convolutional Neural Network, CNN)



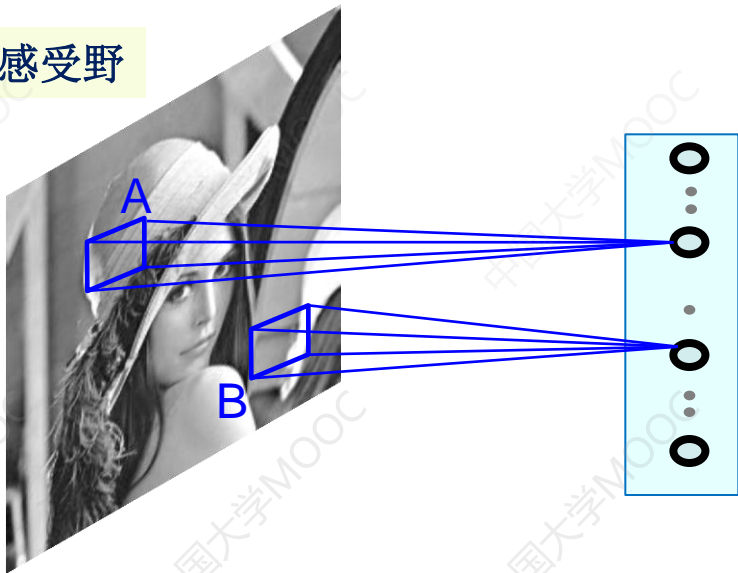


卷积神经网络

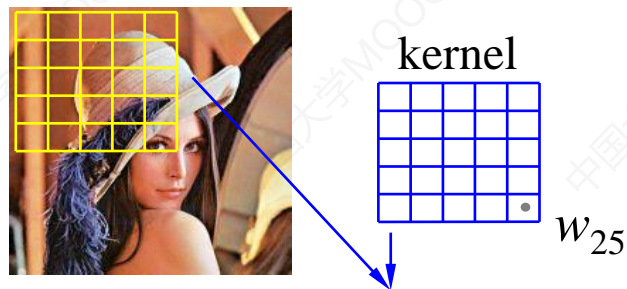
◆ 卷积神经网络 (Convolutional Neural Network, CNN)

- 每个神经元只需要对局部图像进行感知
- 在更高层将局部的信息综合起来得到全局信息

局部感受野



信息综合



$$y = p_1 w_1 + p_2 w_2 + \cdots + p_{25} w_{25}$$

加权求和

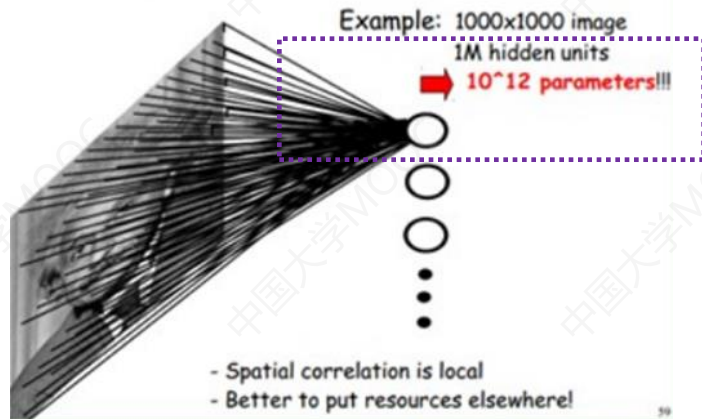


卷积神经网络

◆ 卷积神经网络 (Convolutional Neural Network, CNN)

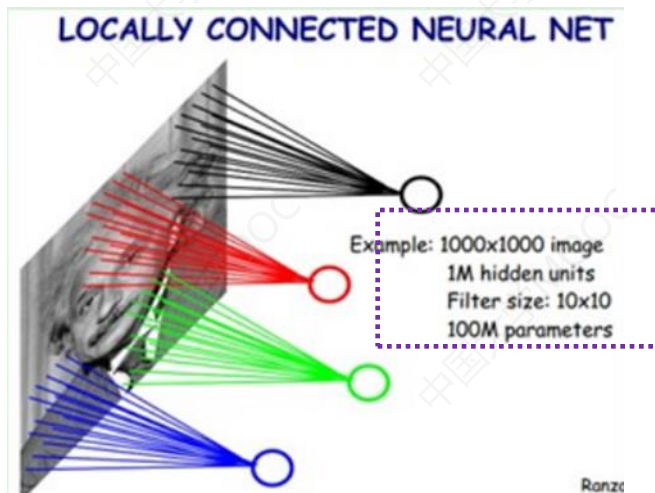
全局（整图）输入

FULLY CONNECTED NEURAL NET



局部感受野

LOCALLY CONNECTED NEURAL NET



$$10^{12} \rightarrow 10^8$$

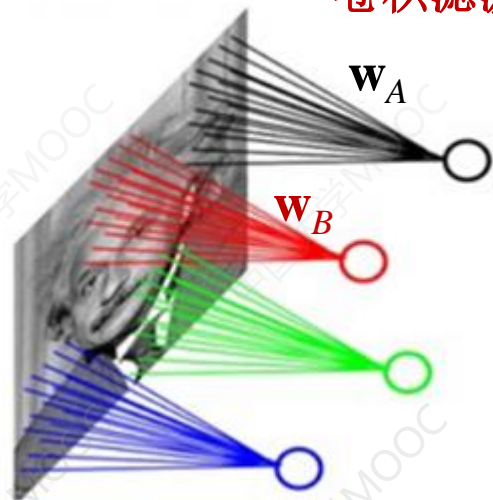


卷积神经网络

◆ 卷积神经网络 (Convolutional Neural Network, CNN)

- 连接权值 (weights) 如何确定？
- 连接权值 100M, 如何再减？

卷积滤波



object

texton/part

edge

pixel

空间域滤波

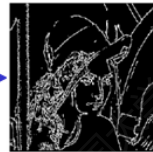
待增强图像



$x(j,i)$

图像增强
系统

增强后图像



$y(j,i)$

$$y(j,i) = \sum_m \sum_n h(m,n) x(j+m, i+n)$$

空间域滤波



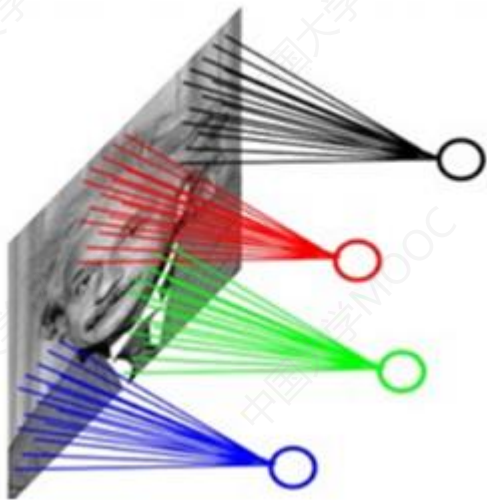
卷积神经网络

◆ 卷积神经网络 (Convolutional Neural Network, CNN)

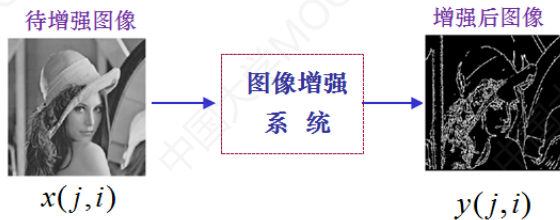
- 连接权值 (weights) 如何确定?

卷积滤波

- 连接权值 100M, 如何再减?



空间域滤波



$$y(j,i) = \sum_m \sum_n h(m,n) x(j+m,i+n)$$

空间域滤波



卷积神经网络

◆ 卷积神经网络 (Convolutional Neural Network, CNN)

- 连接权值 (weights) 如何确定?

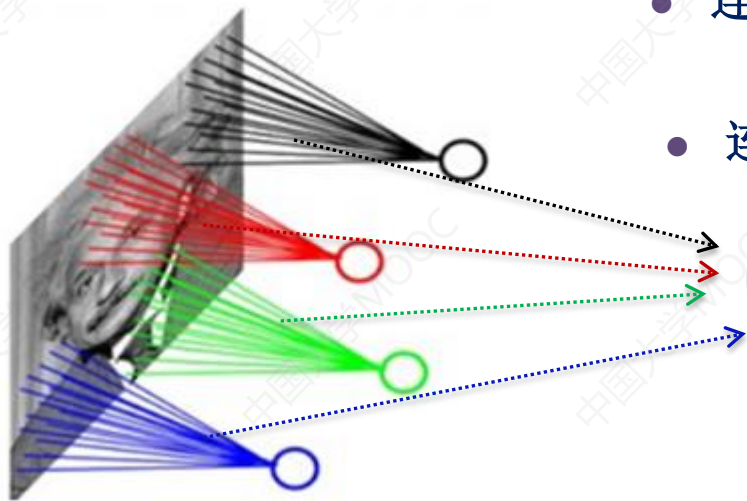
卷积滤波

- 连接权值 100M, 如何再减?

卷积核

局部连接中, 每个神经元都对应 100 个连接权值 (参数), 采用相同的卷积核, 那么总参数为 100.

权值共享





卷积神经网络

降采样 (Subsampling)

图像的统计特性具有空间不变性

- ◆ 在一个图像区域有用的特征极有可能在另一个区域同样适用
- ◆ 为描述大的图像，可以对不同位置的特征进行聚合统计

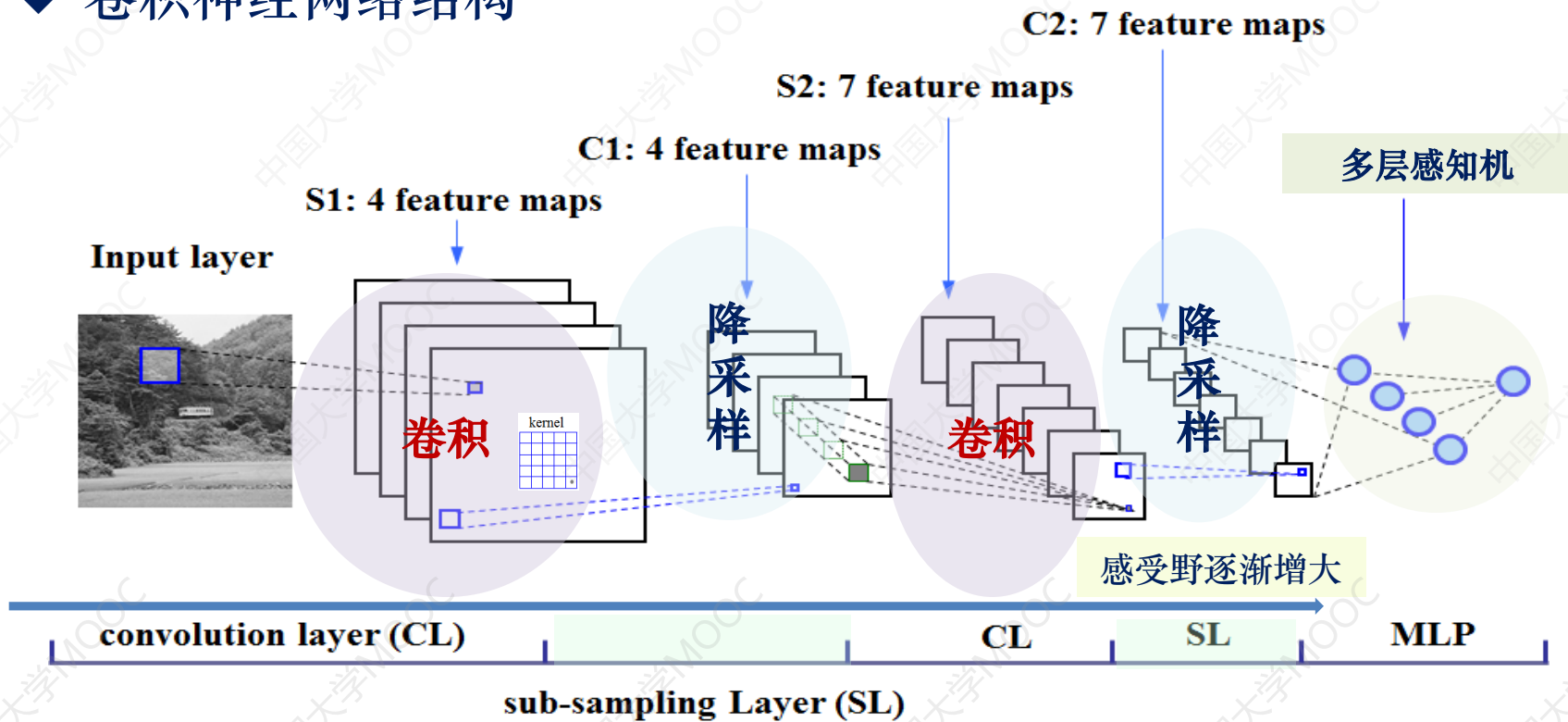
采用图像区域上某个特征的平均值 (或最大值),
维度低且有效(不容易过拟合)

池化 (Pooling)



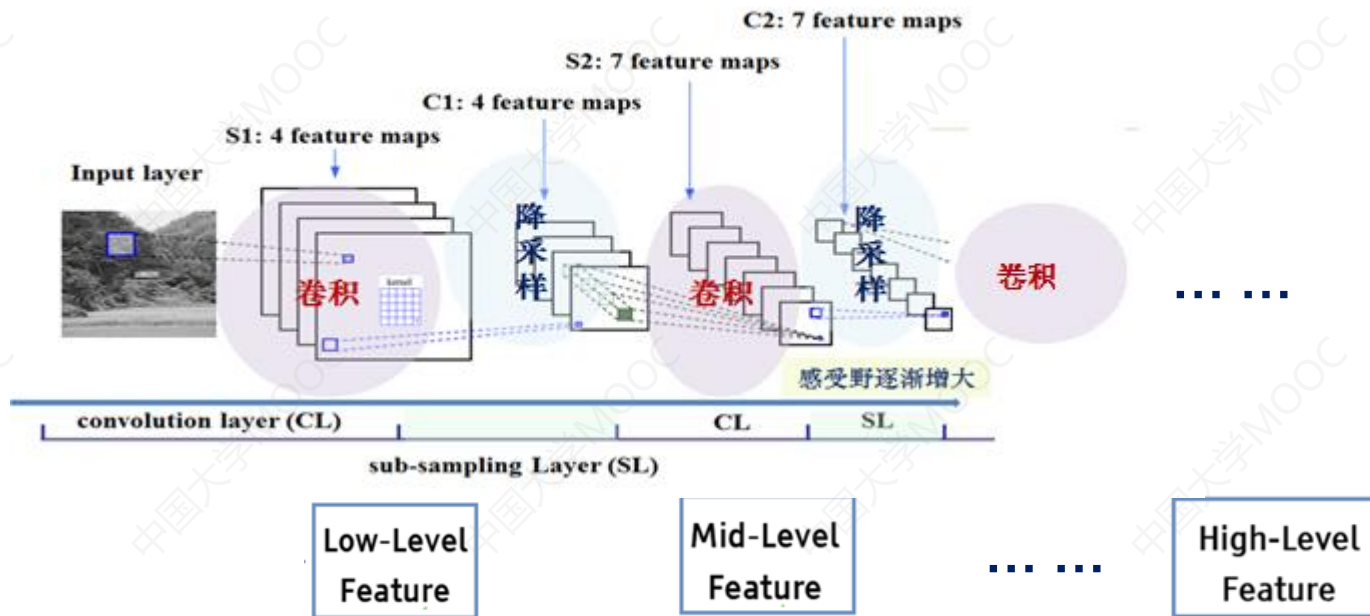
卷积神经网络

◆ 卷积神经网络结构





卷积神经网络





卷积神经网络

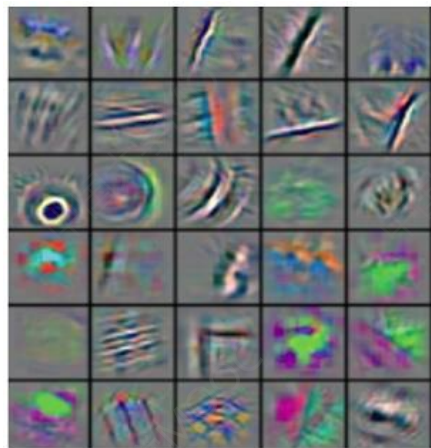
◆ 卷积神经网络 feature map



Low-Level
Feature



Mid-Level
Feature



High-Level
Feature



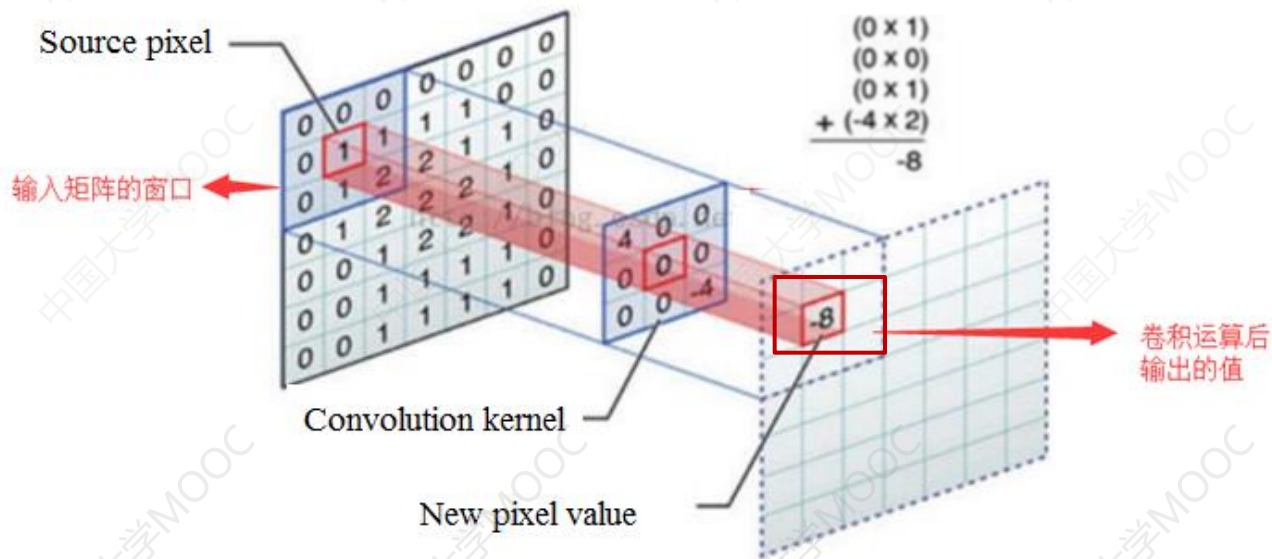


卷积神经网络

◆ 卷积神经网络 (Convolutional Neural Network, CNN)

卷积计算

$$y(j,i) = h(j,i) * x(j,i) \quad y(j,i) = \sum_m \sum_n h(m,n)x(j+m,i+n)$$

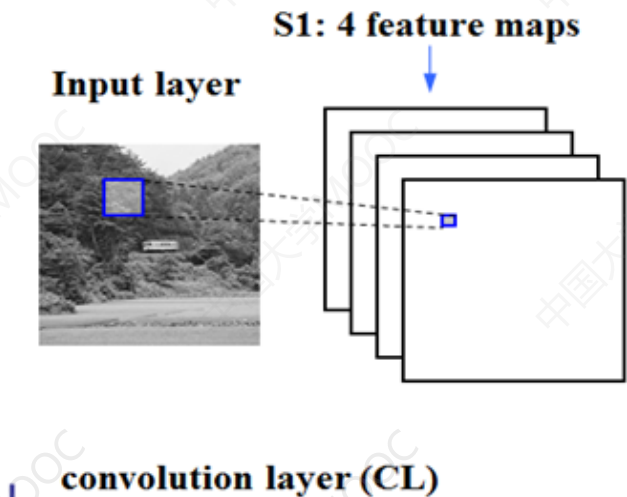




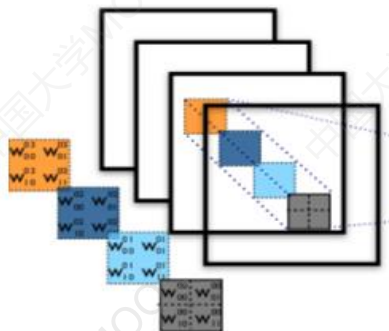
卷积神经网络

◆ 卷积神经网络 (Convolutional Neural Network, CNN)

卷积计算 $y(j,i) = h(j,i) * x(j,i)$



◆ 在卷积神经网络中每一组输出被称为一组特征映射 (Feature Map)

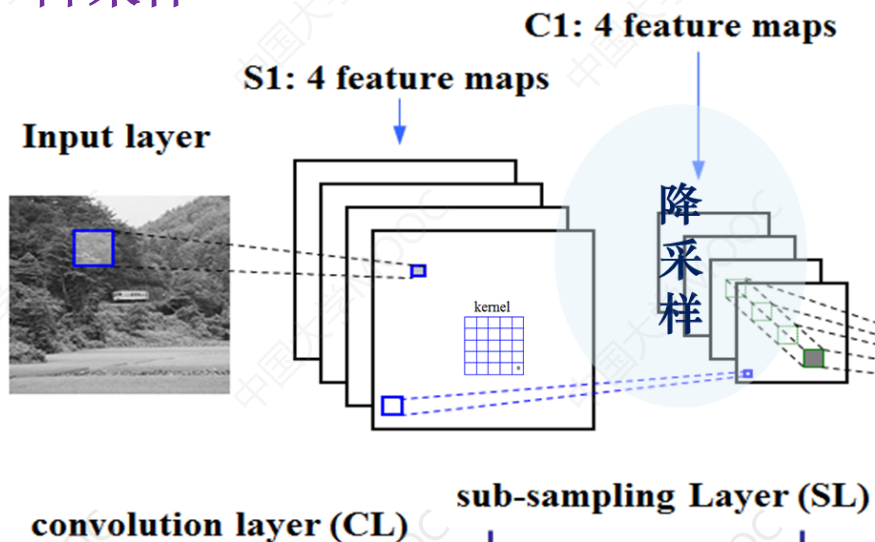




卷积神经网络

◆ 卷积神经网络 (Convolutional Neural Network, CNN)

降采样



- ◆ 卷积层虽然可以显著减少连接的个数，但是每一个特征映射的神经元个数并没有显著减少。
- ◆ 输入分类器的特征维数依然很高，很容易出现过拟合。

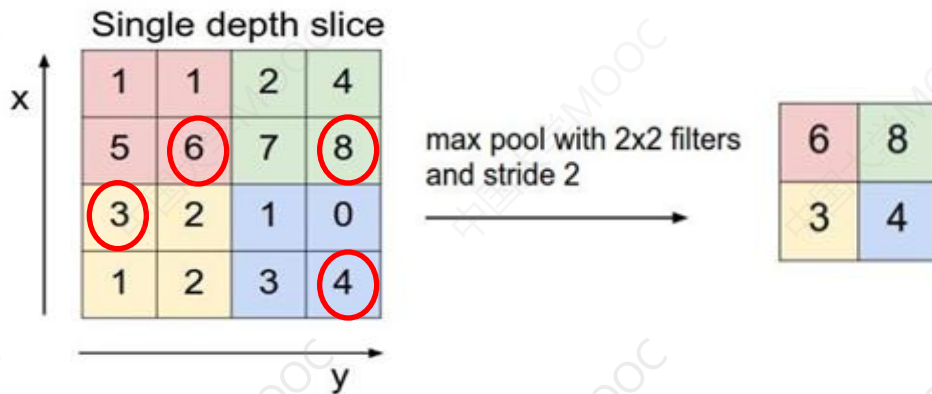
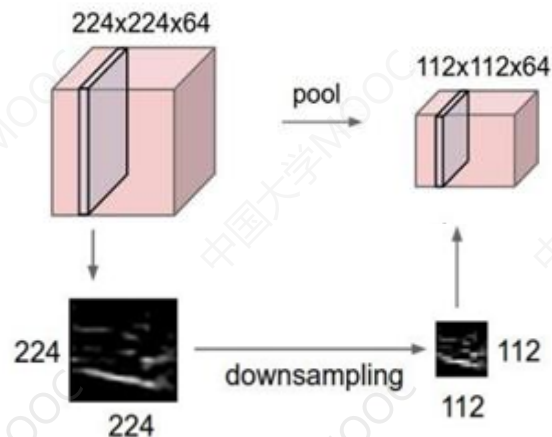
池化 (Pooling)



卷积神经网络

池化 (Pooling)

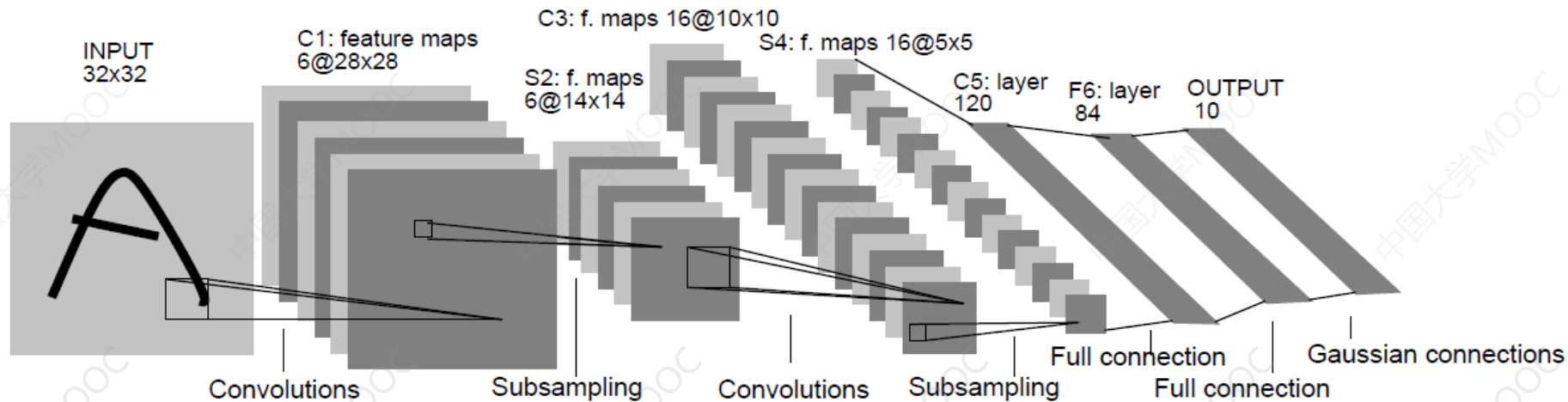
- 对于卷积层得到的一个特征映射 $X(i)$ ，可以将 $X(i)$ 划分为若干区域 R_k , $k = 1, \dots, K$
- 取区域内所有神经元输出值的**最大值**或**平均值**



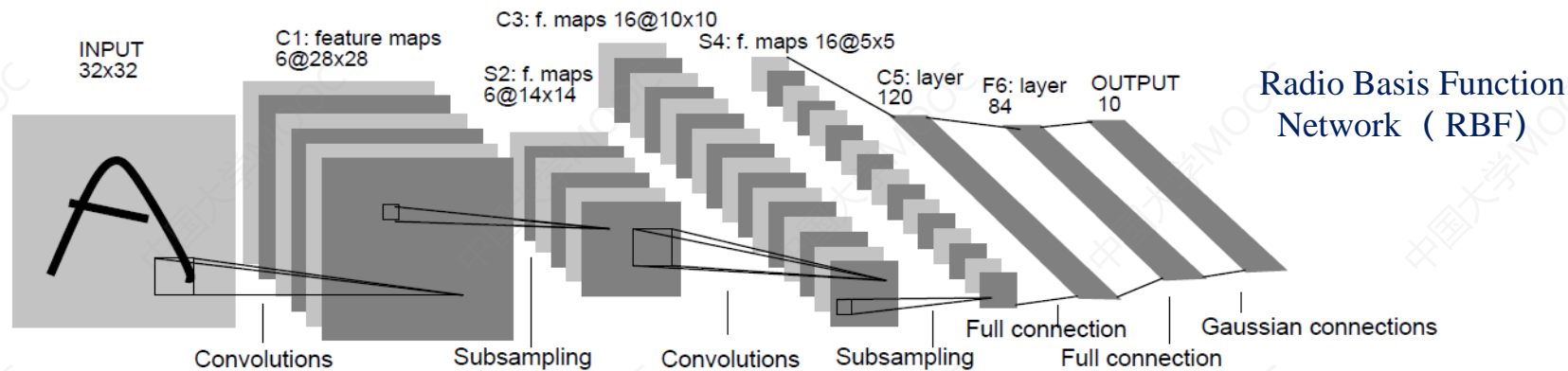
深度卷积神经网络: LeNet [LeCun et al., 1998]



. LeNet : LeCun et.al., 1998. Gradient-based learning applied to document recognition

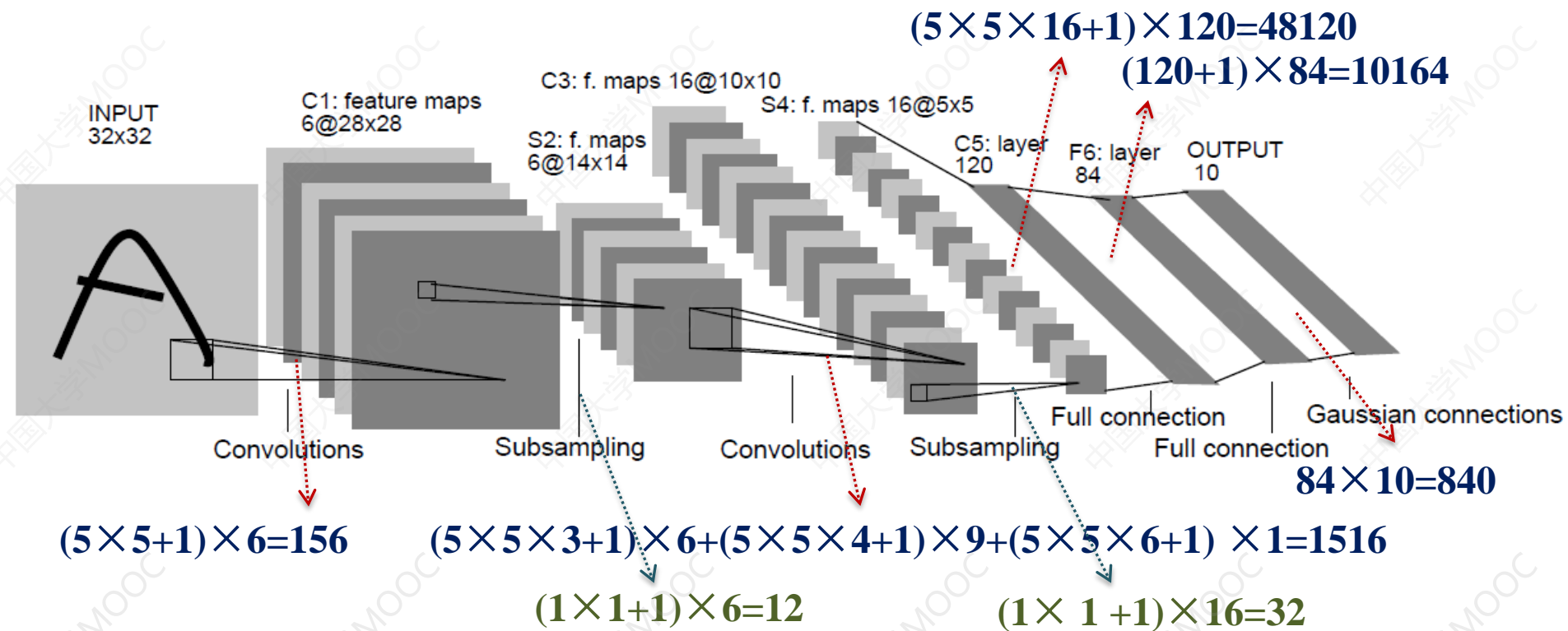


深度卷积神经网络：LeNet [LeCun et al., 1998]



- 输入层： $32 \times 32 = 1024$
- C1层：卷积层 滤波器卷积核大小： $5 \times 5 = 25$ ，共有 6 个滤波器
 $(32 - 5) = 27 \rightarrow 28 \times 28$ C1层为 6 组大小为 $28 \times 28 = 784$ feature maps
- C1层的神经元个数为 $6 \times 784 = 4,704$
- C1层训练参数个数为 $(5 \times 5 + 1) \times 6 = 156$
- S2 层：降采样层（池化），由C1层每组特征图中的 2×2 邻域点降采样为 1 个点（即求 4 个数的平均）
 $28 \times 28 \rightarrow 14 \times 14$
- S2层训练参数个数为 $(1 \times 1 + 1) \times 6 = 12$

深度卷积神经网络：LeNet



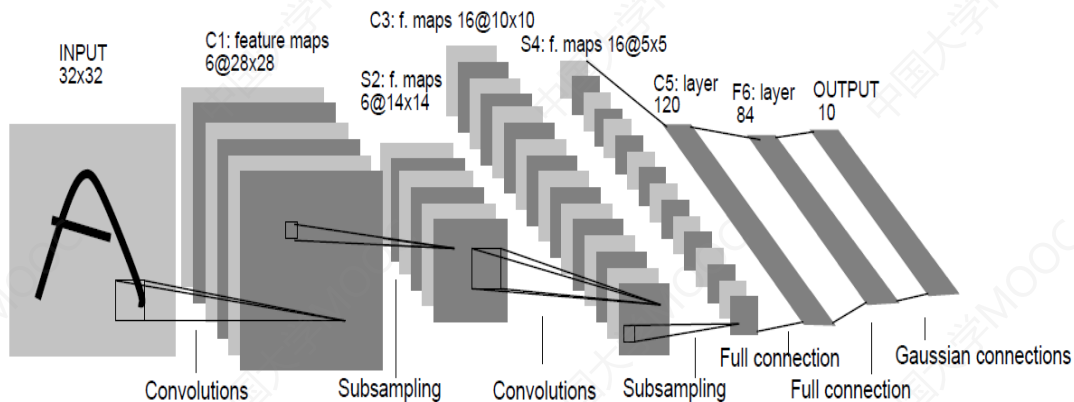
Totally 60,840 trainable free parameters

深度卷积神经网络：LeNet [LeCun et al., 1998]

► 举例：手写字符识别 (Handwritten Character Recognition)



How to make **computer** to **recognize** handwritten character?



基于LeNet-5 的**手写数字识别**系统在九十年代被美国很多银行用来识别支票上面的手写数字，取得了**很好的效果**。



谢 谢

本课程所引用的一些素材为主讲老师多年的教学积累，来源于多种媒体及同事和同行的交流，难以一一注明出处，特此说明并表示感谢！