



模式识别

主讲：崔林艳&邹征夏

单位：宇航学院

开课专业：飞行器控制与信息工程

深度学习背景简介

➤ 传统模式识别

通过传感器来获得数据，然后经过预处理、特征提取、特征选择，再到推理、预测或者识别。



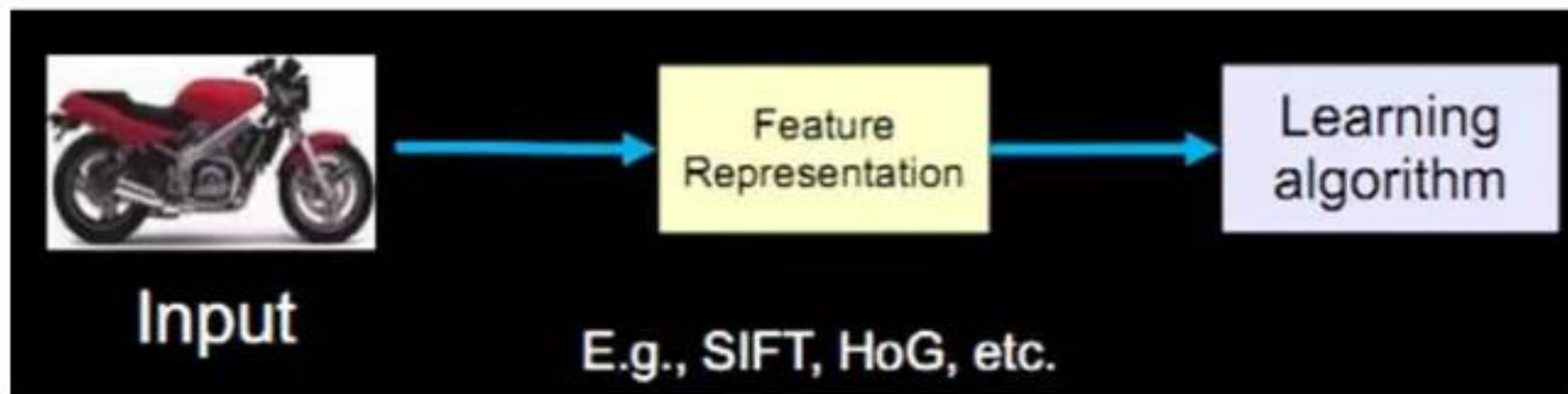
- 良好的特征表达，对最终算法的准确性非常关键。
- 传统模式识别系统计算和测试工作耗时主要集中在特征提取部分。
- 特征提取部分：靠人工提取特征。

深度学习背景简介

➤ 传统模式识别

特征提取和分类器设计分开处理，在应用时再合在一起；

举例：输入一张摩托车图像，对其进行识别。首先要有一个特征表达或者特征提取的过程，然后把表达出来的特征放到学习算法中进行分类学习，从而识别出这是辆摩托车。该识别过程，特征提取和分类器设计是分开的。



深度学习背景简介

➤ 传统模式识别

手工选取特征存在问题：

- 启发式（需要专业知识）
- 能不能选取好很大程度上靠经验和运气；
- 调节需要大量的时间。

分类问题为例：



(assume given set of discrete labels)
{dog, cat, truck, plane, ...}

—————→ cat

深度学习背景简介

➤ 传统模式识别

分类遇到的挑战：

a) 目标姿态变化：



深度学习背景简介

➤ 传统模式识别

分类遇到的挑战：

b) 遮挡问题：



深度学习背景简介

➤ 传统模式识别

分类遇到的挑战：

c) 背景：



d) 类间变化：



深度学习背景简介

➤ 传统模式识别

分类难点：

- a) 人工提取特征困难
- b) 模式识别子问题（特征提取+分类器设计）最优未必意味着全局最优。



问题引出：能否自动学习特征，并实现全局最优？

Deep Learning， 别名Unsupervised Feature Learning。Unsupervised的意思就是不要人参与特征的提取过程。

深度学习背景简介

➤ 深度学习

彻底颠覆了手工提取特征模式，开启了数据驱动的学习方式。

Example dataset: **CIFAR-10**

10 labels

50,000 training images, each image is tiny: 32x32

10,000 test images.

airplane



automobile



bird



cat



deer



dog



frog



horse



ship



truck



➤ 深度学习

彻底颠覆了手工提取特征模式，开启了数据驱动的学习方式。

- a) 所谓的经验和知识也在数据中，在数据量足够大时可直接从数据中学到；
- b) 直接从原始数据开始学习，无需人为转换到某个所谓“更好”的空间再进行学习。
- c) 无需根据经验和知识针对不同问题设计不同的处理流程，大大提高了算法的通用性，降低了解决新问题的难度。

深度学习背景简介

➤ 深度学习

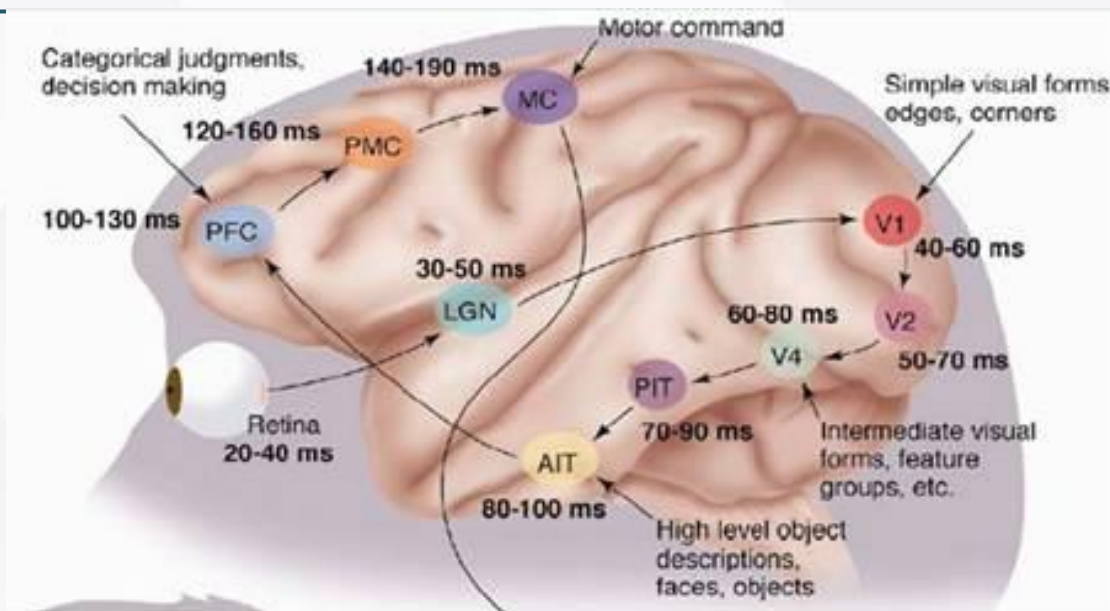
深度学习更强调端到端的学习（end-to-end learning）

- a) 不去人为的分步骤或者划分子问题，而是完全交给神经网络直接学习从原始输入到期望输出的映射。
- b) 相比传统模式识别，端到端的学习具有协同增效（synergy）（各模块协同作用，效果增加， $1+1>2$ ）的优势，更大可能获得全局上更优的解。

深度学习目标：像人脑一样深层次地思考

深度学习背景简介

➤ 人脑视觉处理

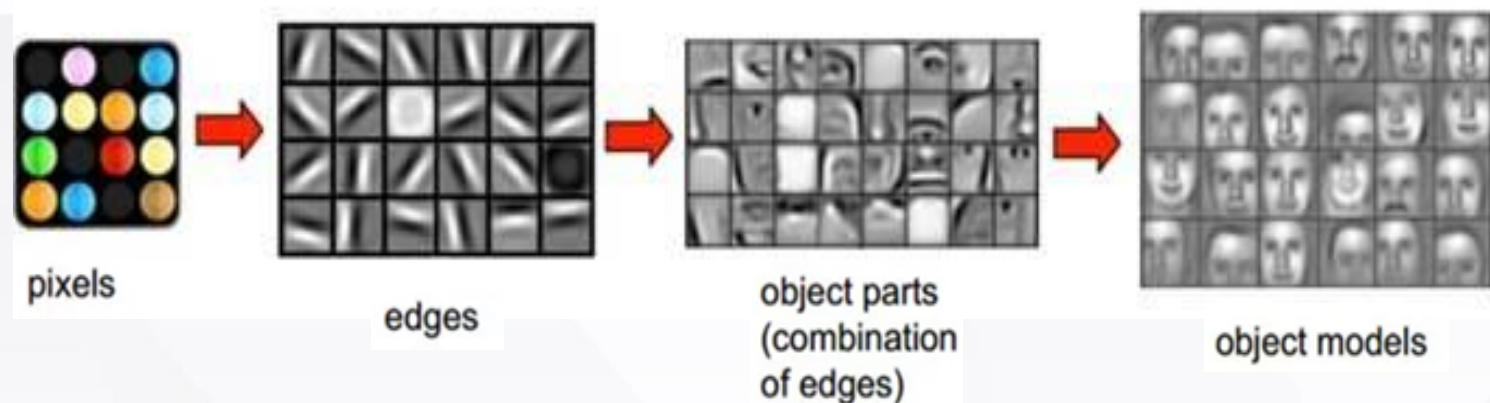


- 人类大脑如何工作：1981年诺贝尔医学奖，颁发给了David Hubel（神经生物学家）。研究成果：**发现了人的视觉系统的信息处理是分级的。**
- 如图所示，从视网膜（Retina）出发，经过低级的**V1区提取边缘特征**，到**V2区的基本形状或目标的局部**，再到**V4区高层的整个目标**，以及到更高层的PFC（前额叶皮层）进行**分类判断**等。
- 也就是说，高层的特征是低层特征的组合，从低层到高层的特征表达越来越抽象和概念化，也即越来越能表现语义。

深度学习背景简介

➤ 人脑视觉处理

大脑的工作过程是一个不断迭代、不断抽象概念化的过程。



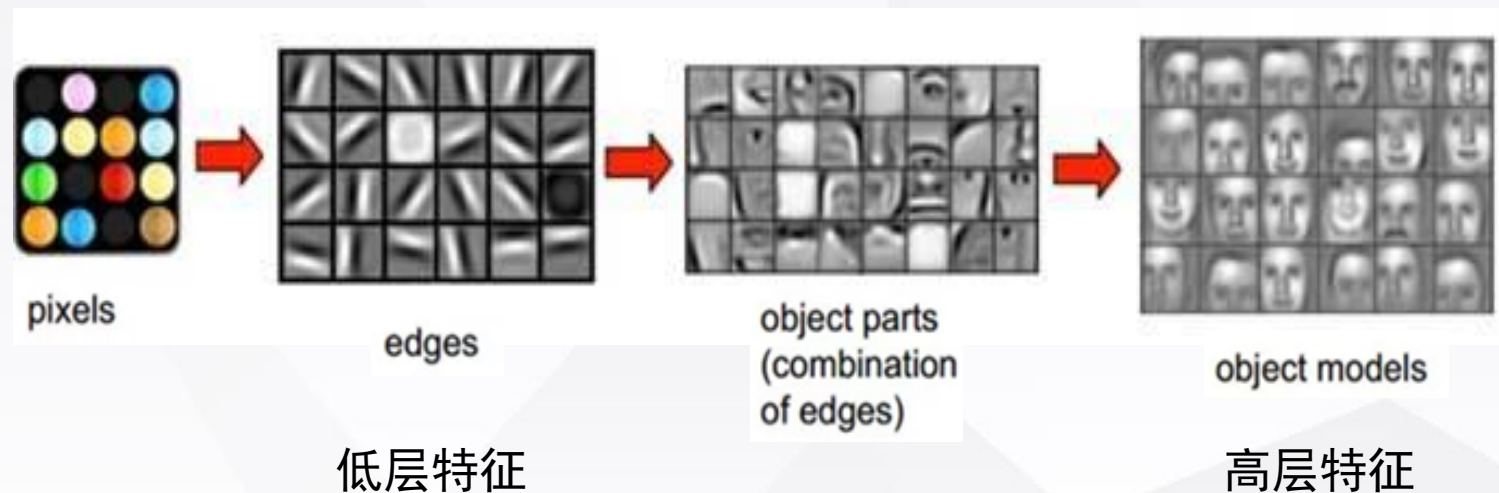
视觉的分层处理结构

大脑工作过程：从**原始信号**摄入开始（瞳孔摄入像素），接着做**初步处理**（大脑皮层某些细胞发现边缘和方向），**然后抽象**（大脑判定眼前物体的形状，比如是椭圆形的），**然后进一步抽象**（大脑进一步判定该物体是张人脸），最后识别眼前的这个人。

深度学习背景简介

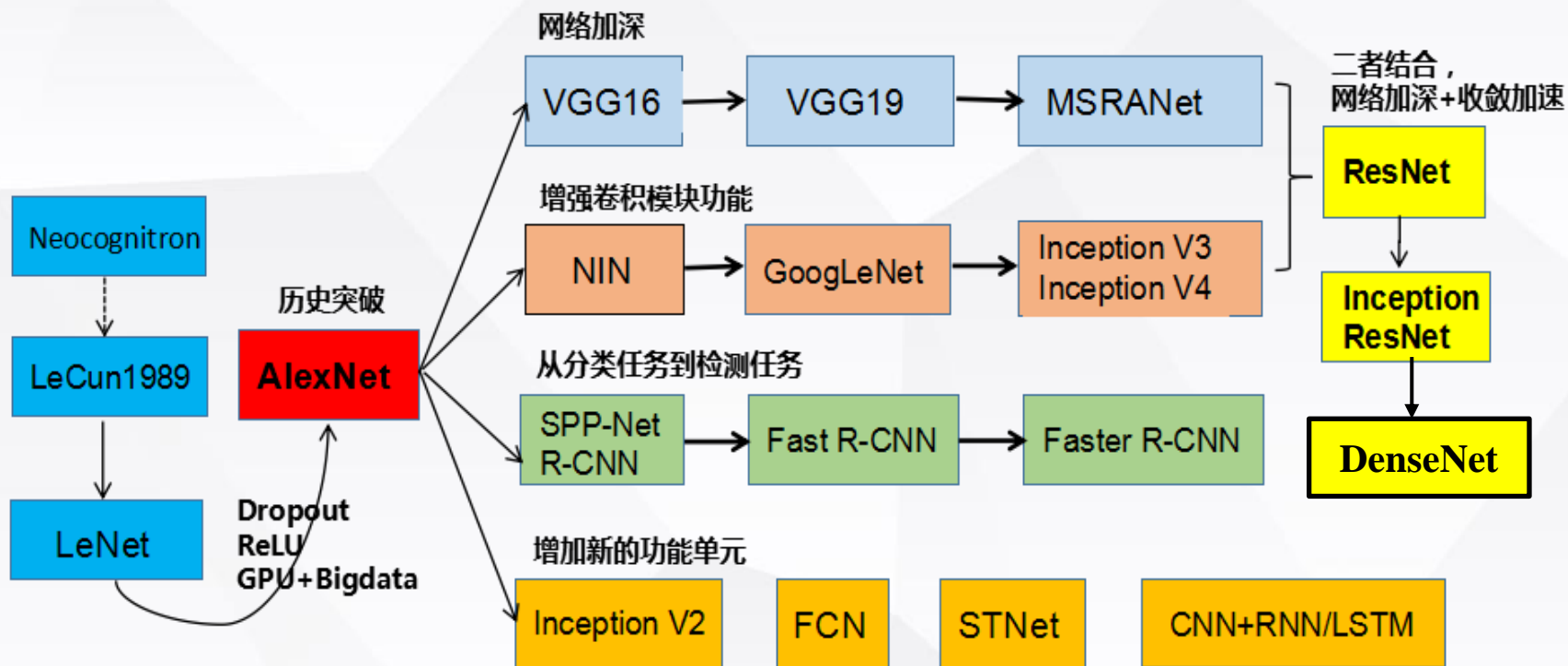
➤ 深度学习

深度学习模拟人类大脑工作机制，通过组合低层特征形成更加抽象的高层特征，进而完成整个学习过程。



深度学习背景简介

➤ CNN发展历程



经典的LeNet诞生于1998年。然而之后CNN开始被SVM等手工设计的特征盖过。随着ReLU和Dropout的提出，以及GPU和大数据带来的历史机遇，CNN在2012年迎来了历史突破—AlexNet。

第八章 卷积神经网络分类器

CONTENTS PAGE

- 8.1 卷积神经网络的基本组成
- 8.2 卷积网络结构与特性分析
- 8.3 卷积网络的训练
- 8.4 经典的卷积网络结构
- 8.5 卷积网络结构设计举例

第八章 卷积神经网络分类器

CONTENTS PAGE

8.1 卷积神经网络的基本组成

8.2 卷积网络结构与特性分析

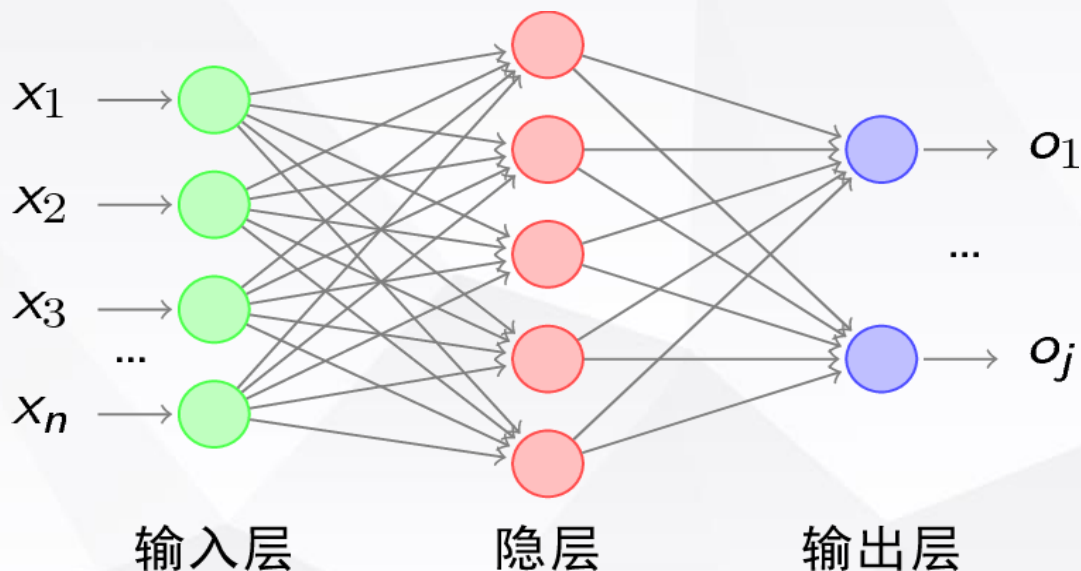
8.3 卷积网络的训练

8.4 经典的卷积网络结构

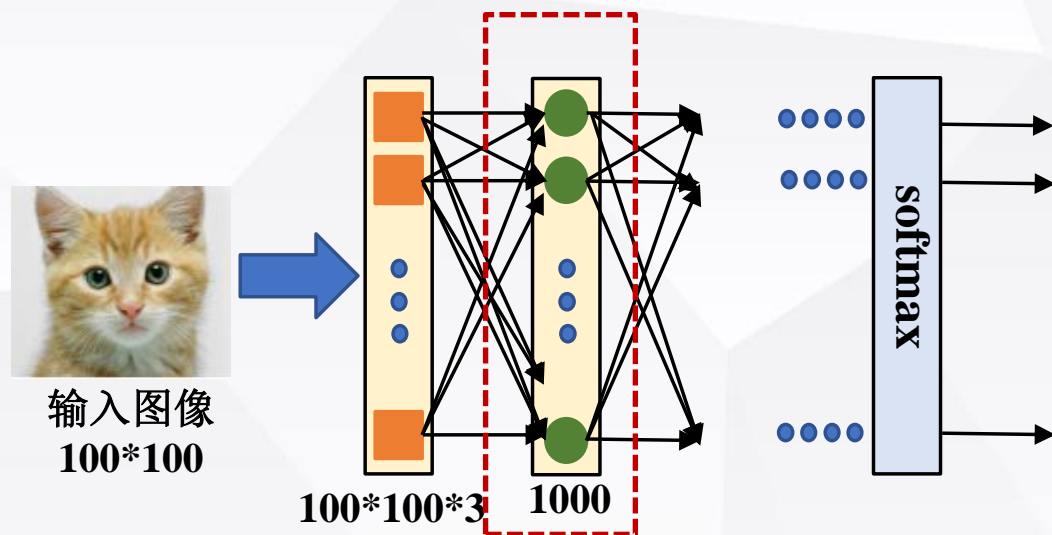
8.5 卷积网络结构设计举例

全连接网络

- 对于一个层，如果它的每一个节点（神经元）接收前一层的所有输入，则该层是**全连接层**，对应的网络为全连接网络。



全连接网络



a) **参数太多：**导致整个神经网络训练效率非常低

每个神经元有 $100*100*3=3*10^4$ 个相互独立连接

该层有 $3*10^7$ 个需要确定的权值参数

b) **难以提取图像的局部不变特征。**

图像中的物体具有局部不变特征，例如尺度缩放、平移、旋转等操作并不会影响其语义信息。如果采用全连接网络模型却很难提取这些局部不变性特征。

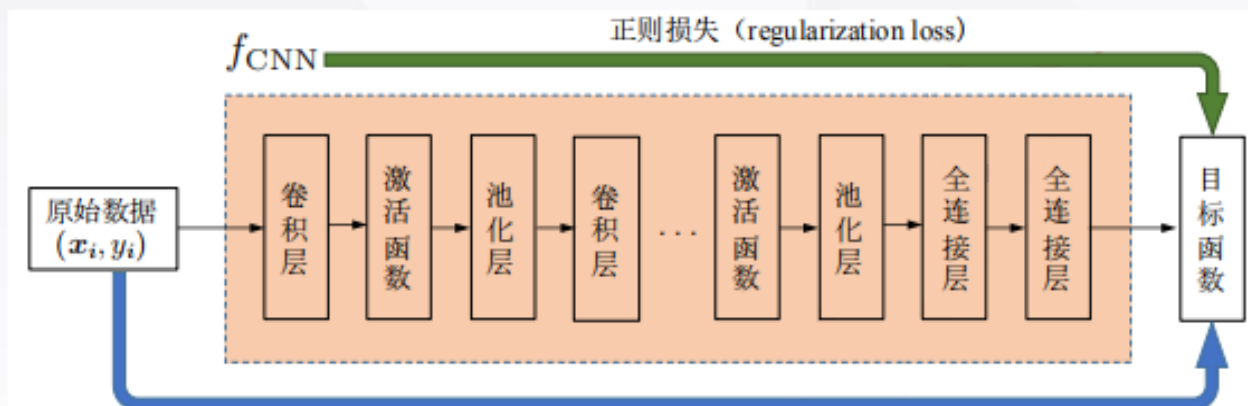
8.1 卷积神经网络的基本组成

- 卷积神经网络（convolutional neural networks, CNNs）是一种**深度前馈型人工神经网络**，可以实现数据的端到端处理，在图像视觉领域发挥了巨大的潜力。
- 卷积神经网络**是当下深度学习领域的基本模型**。

8.1 卷积神经网络的基本组成

8.1.1 卷积神经网络基本结构

一个卷积神经网络通常包括若干个卷积层、激活函数层、池化层、全连接层。输入信号通常是多通道的二维图像数据，输出则是分类或拟合结果。



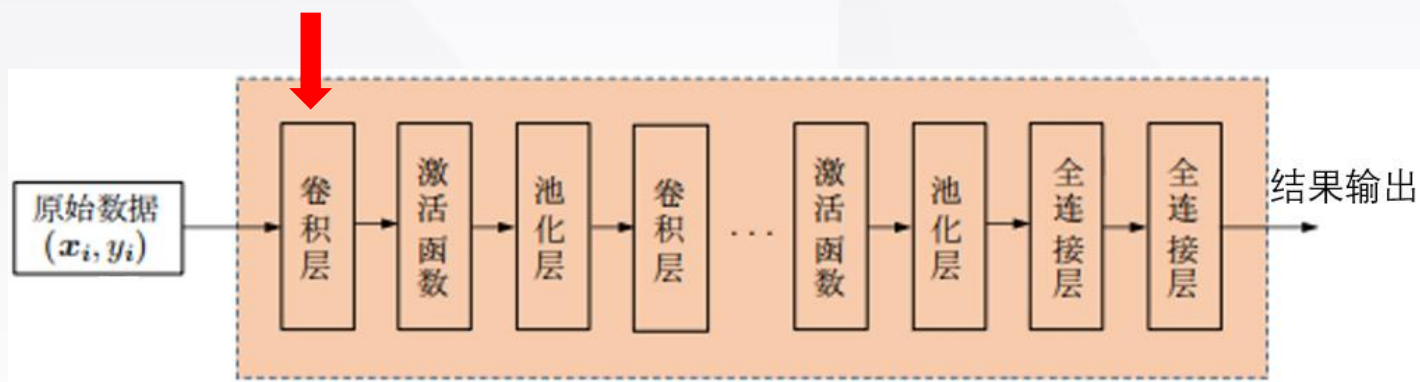
具有局部连接、权值共享特性。

- a) 使得卷积神经网络具有一定程度上的平移、缩放和旋转不变性。
- b) 参数更少。

8.1 卷积神经网络的基本组成

8.1.2 卷积层

卷积层（convolution layer）是卷积神经网络的核心模块，其通过权值共享的卷积核以滑动窗口的方式提取二维图像数据中的局部特征。



卷积神经网络的基本组成

8.1 卷积神经网络的基本组成

8.1.2 卷积层

➤ 单通道（二维）卷积操作

设输入图像数据为 I ，卷积核为 K ，卷积核大小为 $M * N$ ，则对任意一点 (i, j) ，卷积结果为

$$S_{i,j} = (I * K)(i, j) = \sum_{m=1}^M \sum_{n=1}^N I(i + m, j + n) K(m, n)$$

1	0	1
0	1	0
1	0	1

卷积核

1	2	3	4	5
6	7	8	9	0
9	8	7	6	5
4	3	2	1	0
1	2	3	4	5

输入数据



27	28	29
28	27	16
23	22	21

卷积结果（特征）

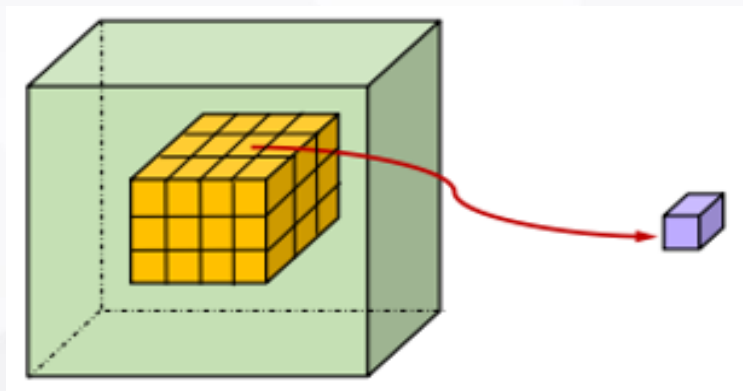
8.1 卷积神经网络的基本组成

8.1.2 卷积层

➤ 多通道（三维）卷积操作

如果有 C 个通道，则 i, j 位置的卷积输出为

$$S_{i,j} = \sum_{c=1}^C \sum_{m=1}^M \sum_{n=1}^N I_c(i+m, j+n) K_c(m, n)$$



三维场景卷积操作

卷积特征

$3*4*3$

8.1卷积神经网络的基本组成

8.1.2 卷积层

➤ 数据填充（边界扩充）

- ① 0填充，即用于填充的数据均为0
- ② 重复填充，即填充的数据为距离当前位置最近的像素原始数据
- ③ 常数填充，即直接指定填充的数据值

1	0	1
0	1	0
1	0	1

卷积核

0	0	0	0	0	0	0
0	1	2	3	4	5	0
0	6	7	8	9	0	0
0	9	8	7	6	5	0
0	4	3	2	1	0	0
0	1	2	3	4	5	0
0	0	0	0	0	0	0

输入数据



8	16	19	12	14
16	27	28	29	10
19	28	27	16	15
14	23	22	21	10
4	8	7	6	6

卷积结果（特征）

8.1 卷积神经网络的基本组成

8.1.2 卷积层

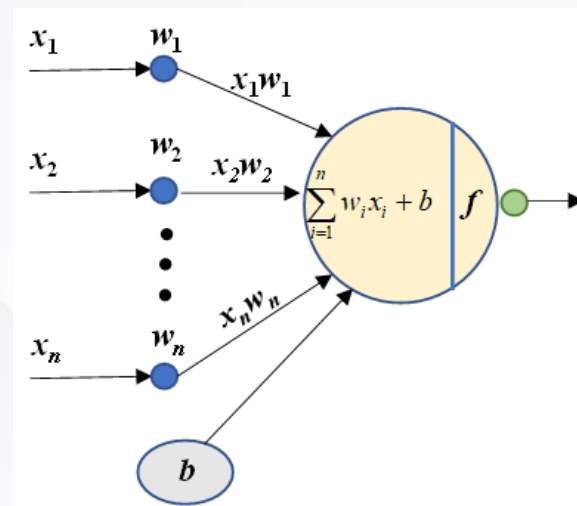
➤ 卷积核的偏置

用 $I * K$ 表示卷积核 K 对输入图像 I 的卷积，卷积核对应偏置为 b ，则卷积操作、偏置与人工神经元数学模型的对应关系如下图所示

□ 卷积网络: $(I * K) + b$

□ 人工神经元模型:

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right)$$



人工神经元模型示意图

8.1 卷积神经网络的基本组成

8.1.2 卷积层

➤ 卷积的作用

提取特征，获得图像或数据的局部信息。



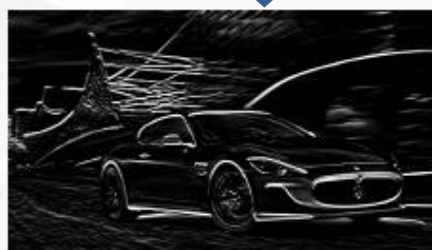
卷积核

0	-4	0
-4	16	-4
0	-4	0



整体边缘滤波

-1	-2	-1
0	0	0
1	2	1



横向边缘滤波

-1	0	1
-2	0	2
-1	0	1



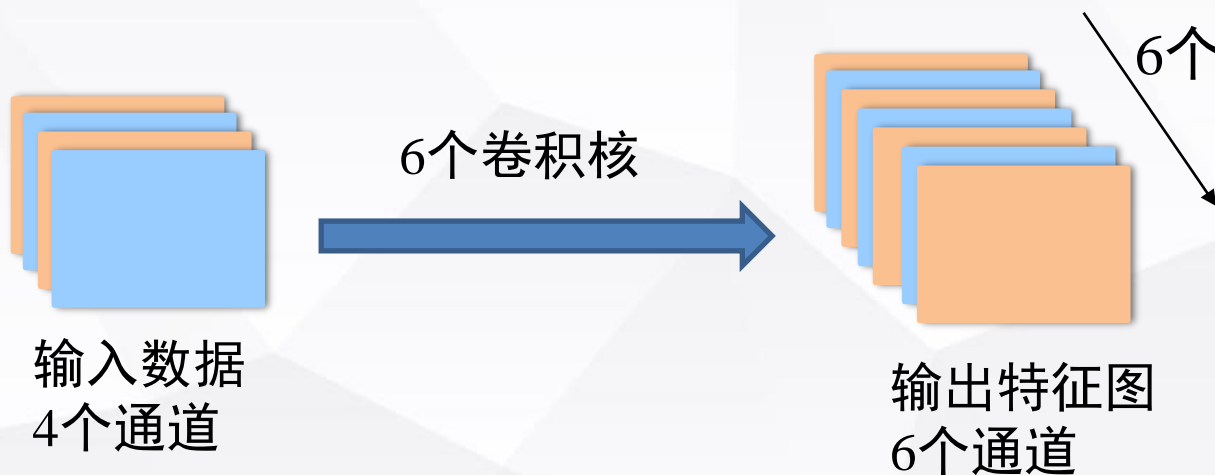
纵向边缘滤波

8.1 卷积神经网络的基本组成

8.1.2 卷积层

➤ 卷积层及参数

一个卷积核对输入数据进行卷积可以得到一个特征图， D 个卷积核可以得到一组（ D 个）特征图，特征图的大小通常用 $H*W*D$ 表示。

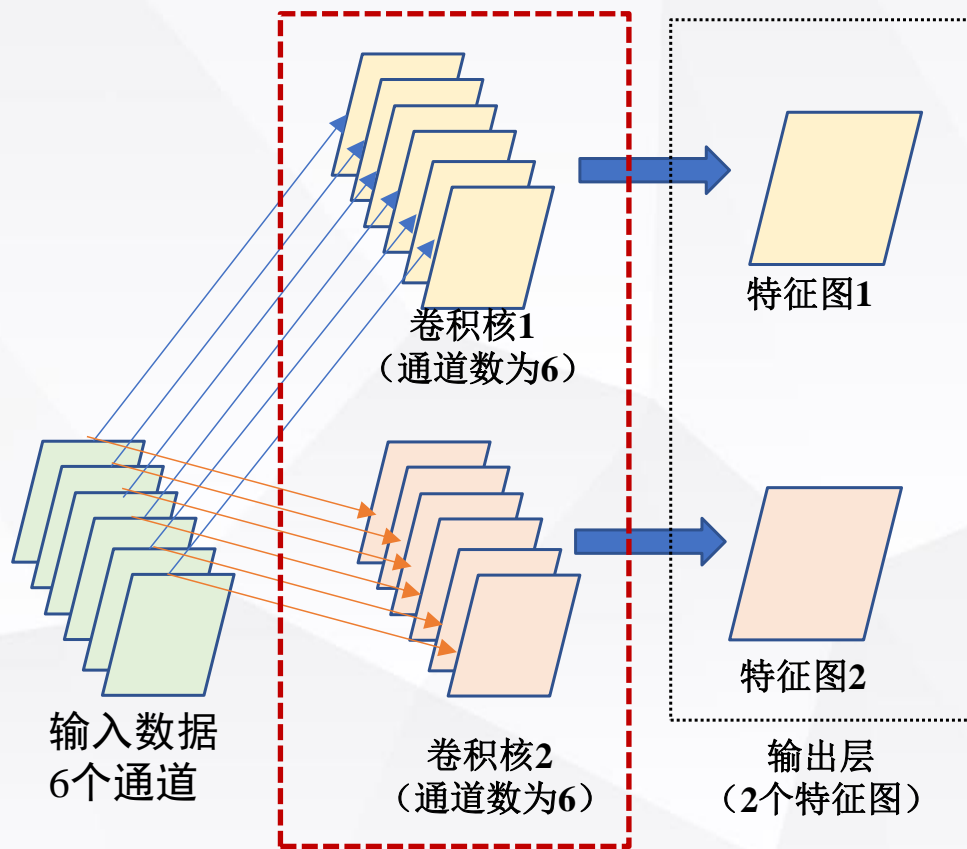


8.1 卷积神经网络的基本组成

8.1.2 卷积层

➤ 卷积核通道数计算

每个卷积核的通道数与前一层特征图的通道数一致。

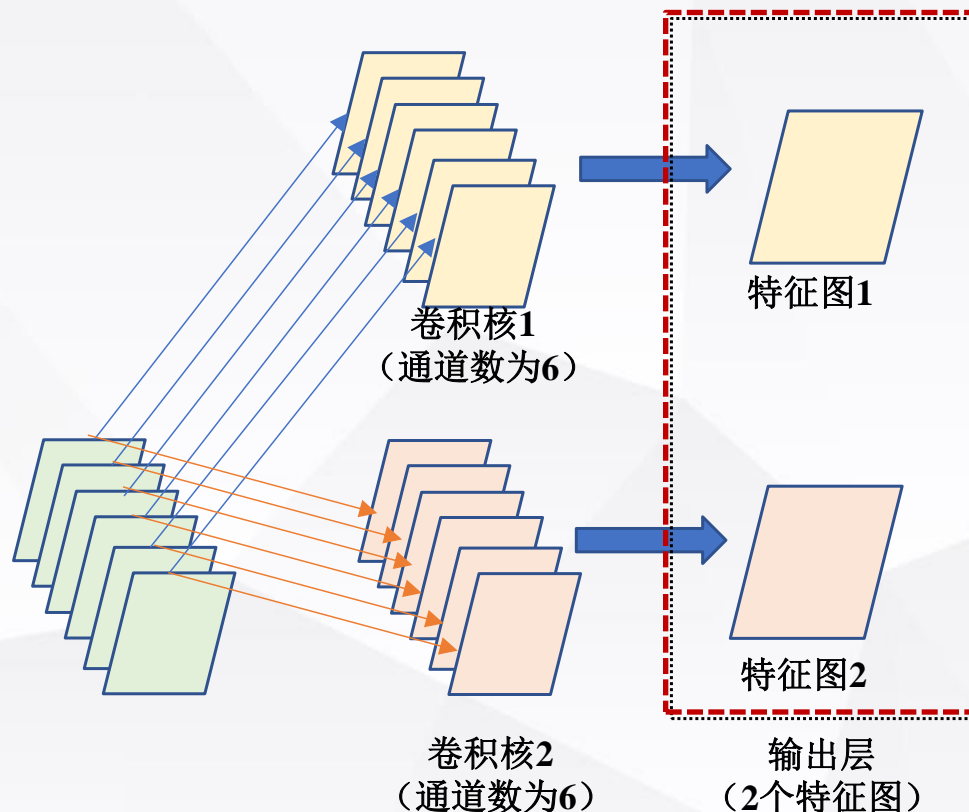


8.1 卷积神经网络的基本组成

8.1.2 卷积层

➤ 特征图数目计算

特征图的数量等于该卷积层中卷积核的数目。



8.1 卷积神经网络的基本组成

8.1.2 卷积层

➤ 特征图大小计算

对于 $N*N$ 大小的输入图像，卷积核（滤波器）大小为 $F*F$ ，移动步长（stride）为 S ，边界每侧扩充尺寸为 P ，则经过卷积操作后得到的特征图大小为：

$$\frac{(N - F + 2P)}{S} + 1$$

- **卷积核大小**（Kernel Size）：定义了卷积操作的感受野。例如设为3，表示卷积核大小为 $3*3$ 。
- **步长**（Stride）：定义了卷积核遍历图像时的步幅大小。默认值通常=1。
- **边界扩充**（Padding）：定义了网络层处理样本边界的方式，可以采用补零、补1，或者其他的方式。

8.1卷积神经网络的基本组成

8.1.2 卷积层

➤ 输入：大小为 $W_1 * H_1 * D_1$ ，其中 $W_1 * H_1$ 为输入特征图尺寸， D_1 为通道数

卷积核：一共 k 个卷积核，大小为 $F * F$ ，移动步长 S ，边界扩充为尺寸为 P

➤ 输出：大小为 $W_2 * H_2 * D_2$ ，其中 $W_2 * H_2$ 为输入特征图尺寸， D_2 为通道数

$$W_2 = \frac{W_1 - F + 2P}{S} + 1$$

$$H_2 = \frac{H_1 - F + 2P}{S} + 1$$

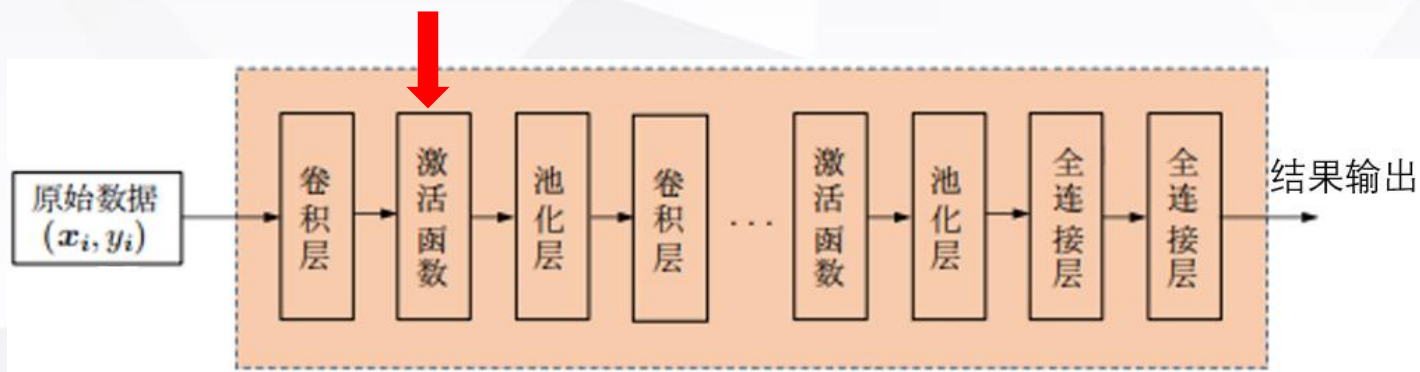
$$D_2 = k$$

每个卷积核还对应一个偏置参数，因此该卷积层参数总量为： $(F \times F \times D_1) \times k + k$

8.1 卷积神经网络的基本组成

8.1.3 激活函数层

- 卷积层的运算是线性的，线性运算的堆叠仍然是线性的。如果网络只有卷积层，那么不管其规模如何庞大，最终都只能实现线性映射。
- 激活函数（activation function）的作用是去线性化，用来实现网络的非线性表达，增强网络的映射能力。



卷积神经网络的基本组成

8.1卷积神经网络的基本组成

8.1.3 激活函数层

1) 激活函数的一般要求

一般要求激活函数具有非线性、良好的可导性、计算简单、非饱和性、单调性、合理的输出范围。

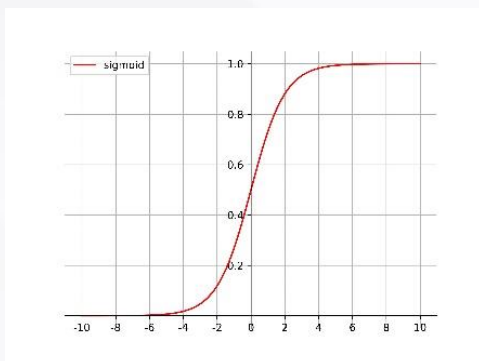
8.1卷积神经网络的基本组成

8.1.3 激活函数层

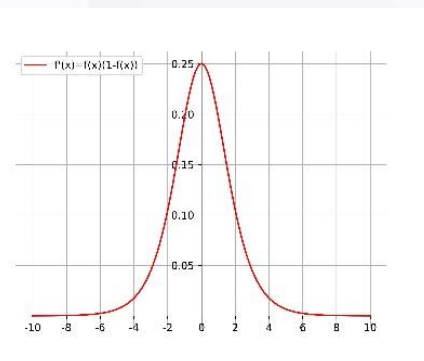
2) 常用的激活函数

➤ Sigmoid激活函数

$$f(x) = \frac{1}{1 + e^{-x}} \quad \text{具有指数形式}$$



Sigmoid函数



Sigmoid导数

Sigmoid函数导数的取值范围是(0,0.25], 数值相对较小, 在梯度反向传播中容易产生梯度消失; 函数包含幂运算, 计算量较大; 输出不是零均值, 随着网络加深, 可能会改变数据的原始分布。

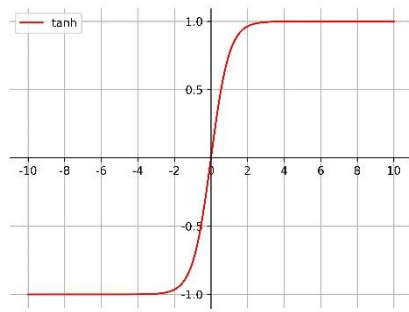
8.1卷积神经网络的基本组成

8.1.3 激活函数层

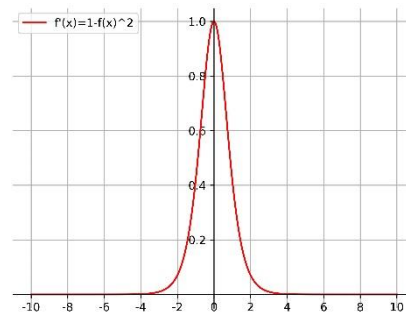
2) 常用的激活函数

➤ 双曲正切（Tanh）函数

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$



Tanh函数



Tanh导数

双曲正切函数解决了输出零均值的问题，其导数范围为(0,1]，比Sigmoid函数增大，但梯度消失问题仍然存在。

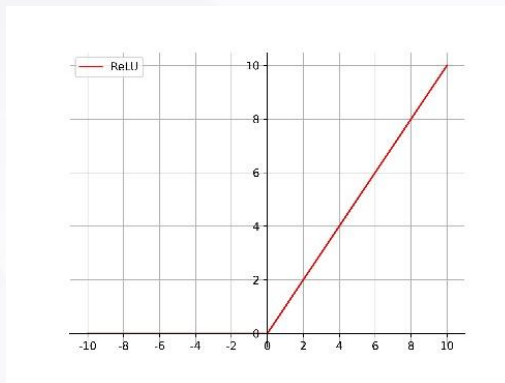
8.1卷积神经网络的基本组成

8.1.3 激活函数层

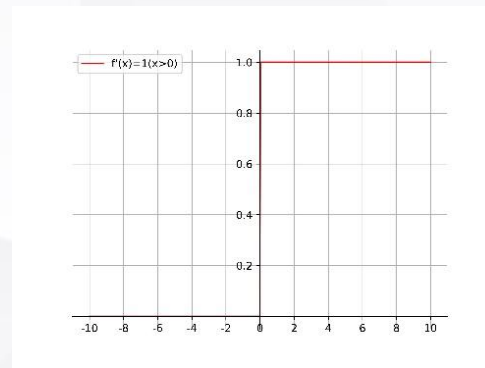
2) 常用的激活函数

➤ ReLU激活函数：

$$\text{ReLU}(x)=\max\{0,x\}=\begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$$



ReLU函数



ReLU导数

是分段函数， $x \geq 0$ 时消除了梯度饱和效应，有助于梯度下降法收敛，成为当前卷积网络的首选。但 $x < 0$ 时梯度为0，出现“死区”。

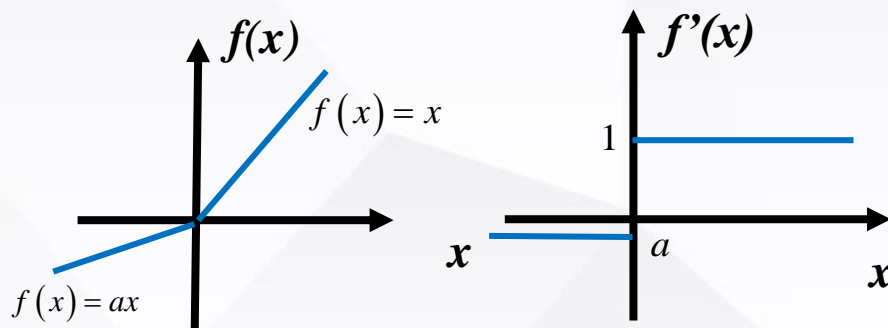
8.1卷积神经网络的基本组成

8.1.3 激活函数层

2) 常用的激活函数

➤ 参数化ReLU:

$$\text{Leaky ReLU}(x) = \begin{cases} x & x \geq 0 \\ \alpha x & x < 0 \end{cases}$$

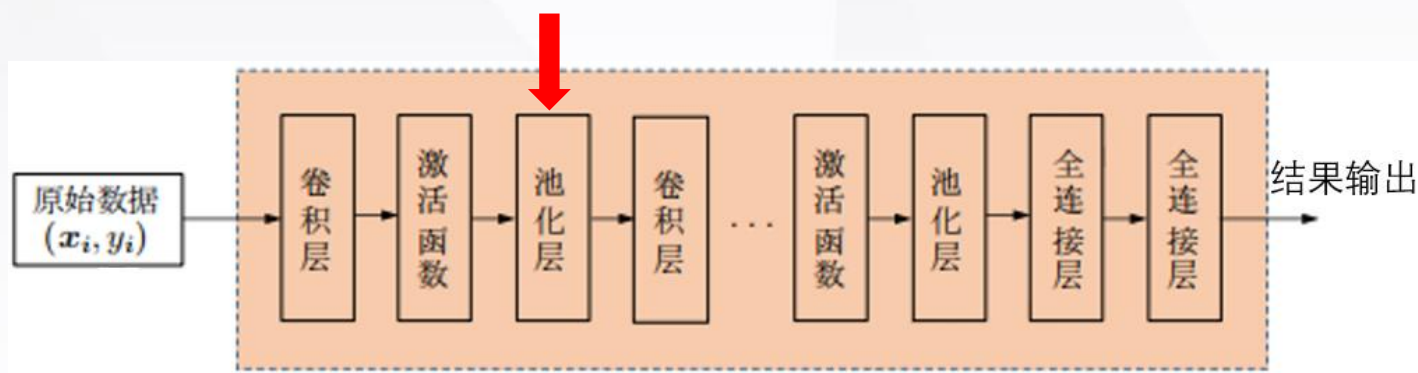


引入超参数 α ，作为一个网络中可学习的变量融入模型的整体训练过程，解决“死区”问题。

8.1 卷积神经网络的基本组成

8.1.4 池化层

- 池化层（pooling layer）的作用是对卷积层中提取的特征进行挑选。
- 使模型更关注是否存在某些特征而不是特征的具体位置。
- 该操作是降采样过程，可以提高后续特征感受野，并减少计算量。
- 池化层一般没有模型参数，反向传播过程中不需要进行权值优化。



卷积神经网络的基本组成

8.1 卷积神经网络的基本组成

8.1.4 池化层

➤ 池化层的计算

常见的池化操作：最大池化和平均池化。池化层是由矩阵窗口滑动来进行计算的，求窗口中的最大值和平均值等。

最大池化和平均池化的计算示例：滤波核的大小为 2×2 ，步长为2

1	3	0	3
2	2	4	1
4	2	0	0
3	3	3	1

3	4
4	3

最大池化

1	3	0	3
2	2	4	1
4	2	0	0
3	3	3	1

2	2
3	1

平均池化

8.1卷积神经网络的基本组成

8.1.4 池化层

➤ 池化层的特征图尺寸

输入：大小为 $W_1 * H_1 * D_1$ ，其中 $W_1 * H_1$ 为输入特征图尺寸， D_1 为通道数

池化核参数：尺寸 $F * F$ ，移动步长 S

输出：大小为 $W_2 * H_2 * D_2$ ，其中 $W_2 * H_2$ 为输出特征图尺寸， D_2 为通道数

$$W_2 = \frac{W_1 - F}{S} + 1$$

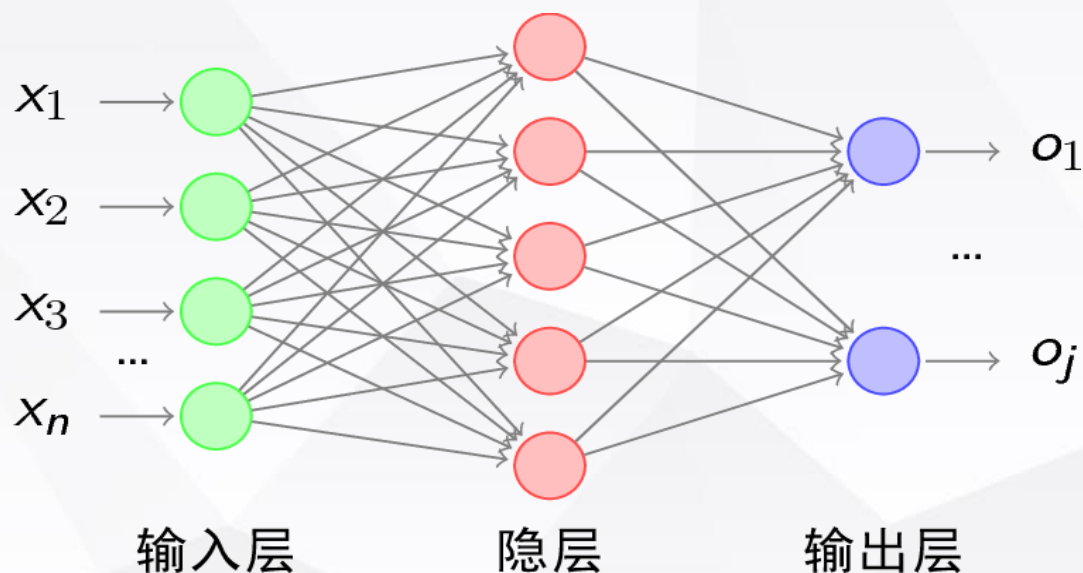
$$H_2 = \frac{H_1 - F}{S} + 1$$

$$D_2 = D_1$$

8.1 卷积神经网络的基本组成

8.1.5 全连接层

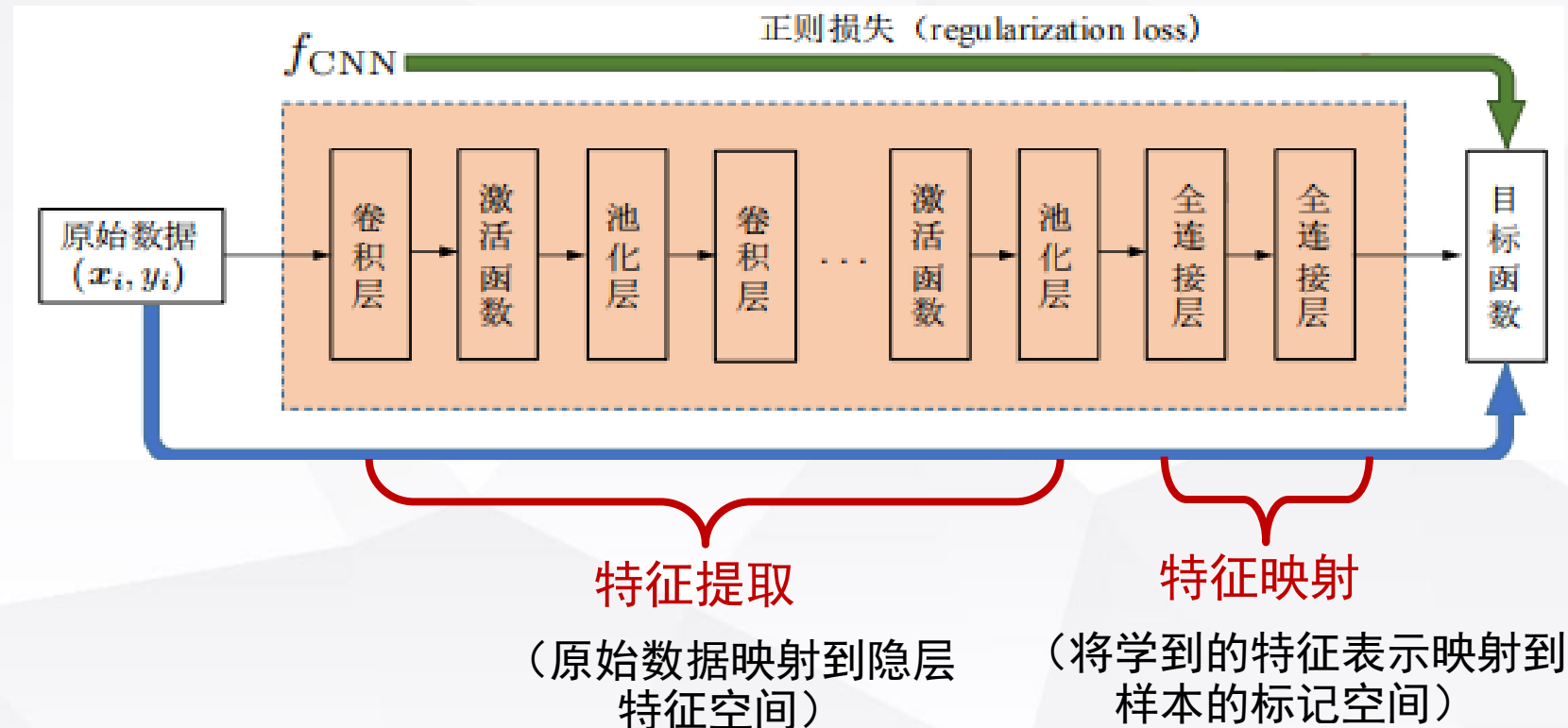
对于一个层，如果它的每一个节点（神经元）接收前一层的所有输入，则该层是**全连接层**，如左侧网络的后边两层是全连接的。



8.1 卷积神经网络的基本组成

8.1.5 全连接层

全连接层起到“分类器”的作用。卷积层、池化层和激活函数层等操作将原始数据映射到隐层特征空间，是特征提取的过程。全连接层则将学到的特征表示映射到样本的标记空间。



第八章 卷积神经网络分类器

CONTENTS PAGE

8.1 卷积神经网络的基本组成

8.2 卷积网络结构与特性分析

8.3 卷积网络的训练

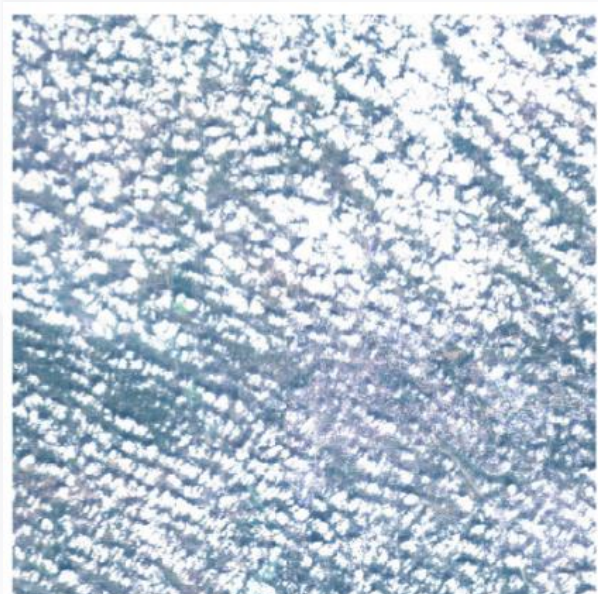
8.4 经典的卷积网络结构

8.5 卷积网络结构设计举例

8.2 卷积网络结构与特性分析

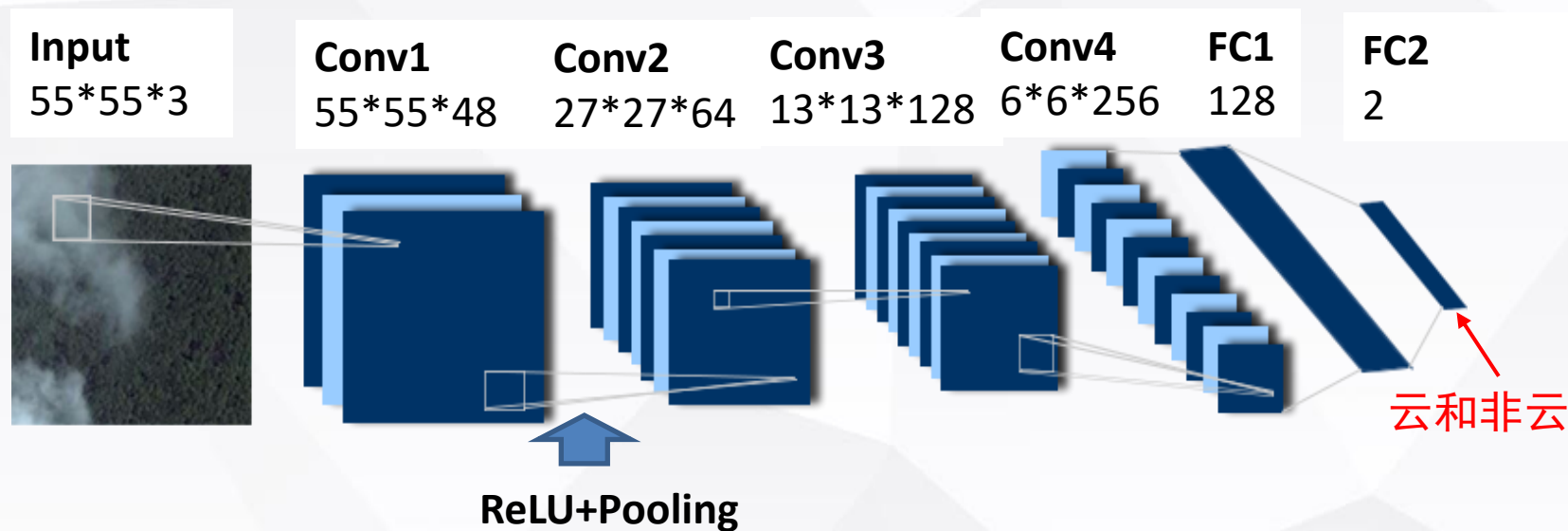
➤ 以云识别卷积网络为例：

遥感图像中经常会出现云，对遥感图像中的云区进行检测和识别，对于气象预报、生态环境监测、气象灾害的预防等方面都有积极的作用。



8.2 卷积网络结构与特性分析

- 下图是遥感图像中云区域识别网络，包括4个卷积层和2个全连接层，卷积层之间插入了激活函数层（ReLU）和池化层。

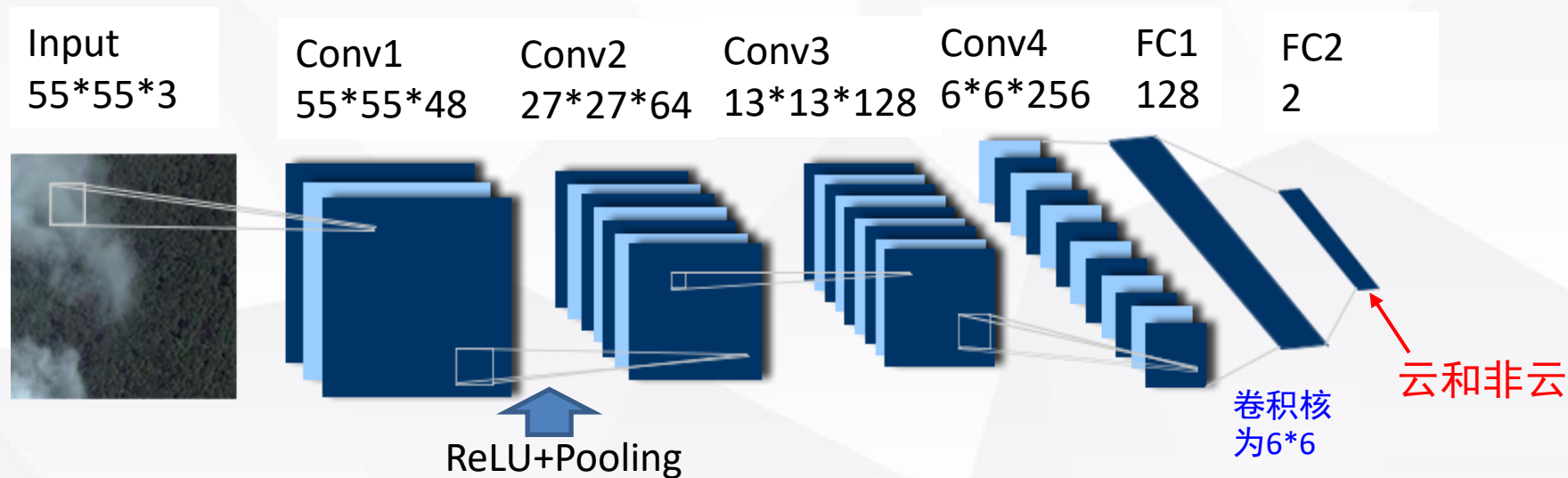


- 输入图像块大小为55*55*3，输出为该图像块是云和非云的概率。
- 卷积核尺寸为3*3，移动步长为1
- 池化层的核尺寸为2，移动步长为2

8.2 卷积网络结构与特性分析

- 表中给出了每一层的参数个数，例如卷积层1中，由于输入是3通道的，因此卷积核的通道数为3，该层有48个卷积核，每个卷积核对应一个偏置，则该层参数为 $(3*3*3)*48+48=1344$ ，其他层可以做类似的分析。

层	参数个数
Conv1	$(3*3*3)*48+48=1344$
Conv2	$(3*3*48)*64+64=27712$
Conv3	$(3*3*64)*128+128=73856$
Conv4	$(3*3*128)*256+256=295168$
FC1	$(6*6*256)*128=1179684$
FC2	$128*2=256$



8.2 卷积网络结构与特性分析

➤ 卷积网络的特点分析：

1) 权值共享

权值共享含义：一个卷积核计算出一个特征图，计算过程中参数不变，这就是卷积网络中的权值共享。

权值共享作用：经过权值共享，每个特征具有平移不变性，同一个特征可以出现在不同数据的不同位置，可以用同一个卷积核来提取这一特征。同时，通过权值共享，**减少了卷积层的参数**。

8.2 卷积网络结构与特性分析

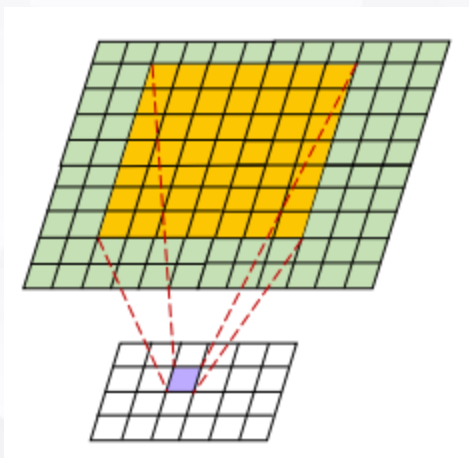
➤ 卷积网络的特点分析：

2) 感受野

如图所示，在卷积操作中，后层的一个神经元对应前层一个区域的大小，就是感受野。卷积核的大小不同，所对应前一层的感受野也不同。

前层

后层

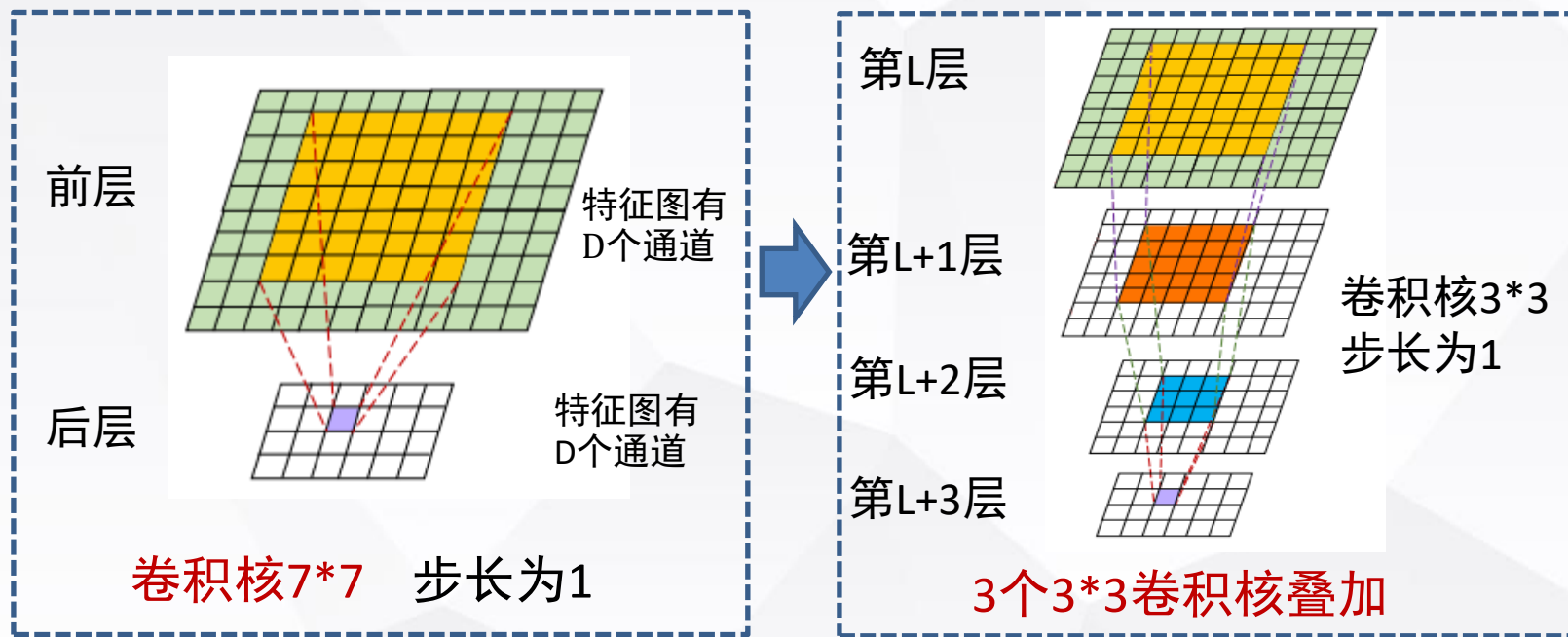


7*7感受野

8.2 卷积网络结构与特性分析

➤ 卷积网络的特点分析：

小卷积核可以通过多层叠加取得与大卷积核同等规模的感受野，同时加深了网络深度进而增大了网络容量，并能减少参数个数。



【举例】：假设当前输入的特征图有D个通道，采用7*7卷积核，卷积核个数为D，则对应下一层的参数为 $7 * 7 * D * D = 49D^2$ ，若采用3*3卷积核进行堆叠，只需要 $3(3 * 3 * D * D) = 27D^2$ 。

8.2 卷积网络结构与特性分析

➤ 卷积网络的特点分析：

3) 深度特征的层次性

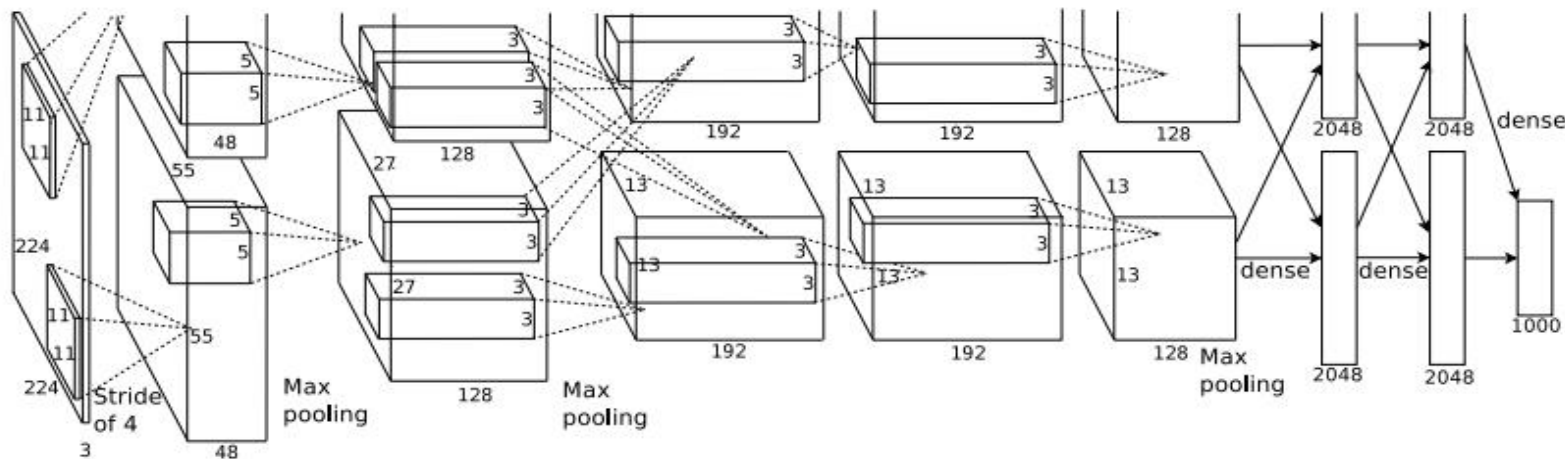
- 一个卷积网络具有多个卷积层，前层对应低层特征，后层对应高层特征。
- 低层特征一般对应边缘、纹理等细节。越到高层，一个特征点对应前端的感受野越大，越能抽象出物体的轮廓和整体，也就对应高层语义特征。

8.2 卷积网络结构与特性分析

➤ 卷积网络的特点分析：

3) 深度特征的层次性

2014年，Zeiler利用反卷积（转置卷积）技术对网络特征进行可视化，验证深度特征的层次性。

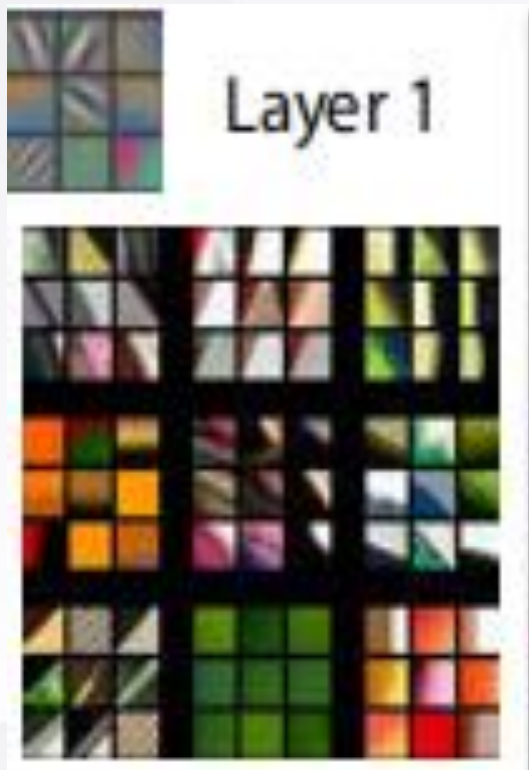


Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In proceedings of European conference on computer vision, pages 818-833,2014.

8.2 卷积网络结构与特性分析

➤ 卷积网络的特点分析：

3) 深度特征的层次性



第一层对应边缘和方向（细节信息）

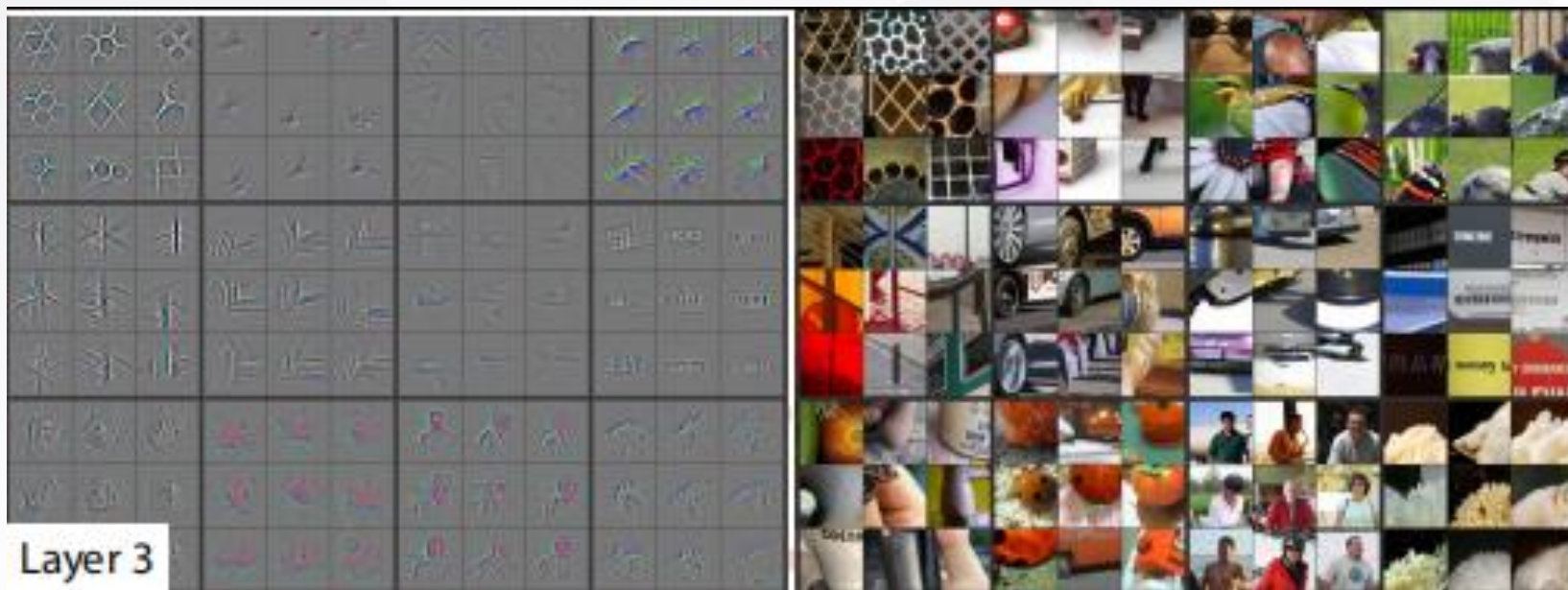


第二层对应纹理特征

8.2 卷积网络结构与特性分析

➤ 卷积网络的特点分析：

3) 深度特征的层次性

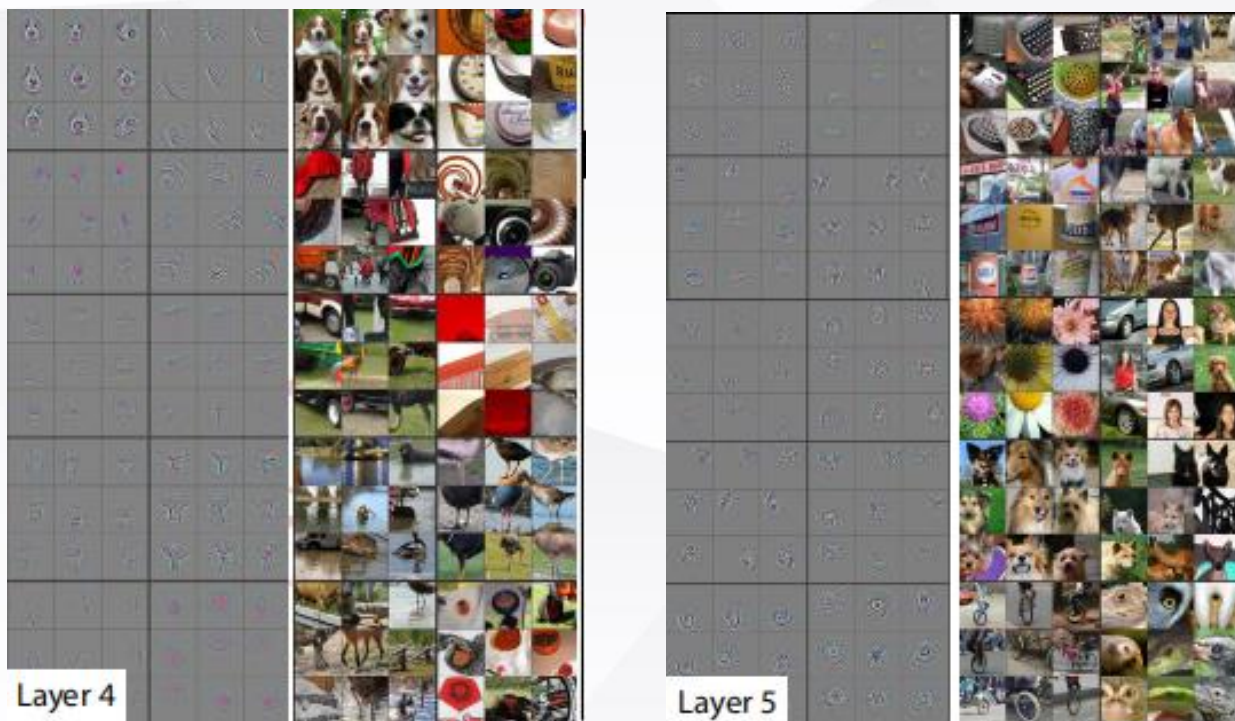


第三层开始出现一些高层语义模式，如车轮、文字和人脸形状的模式

8.2 卷积网络结构与特性分析

➤ 卷积网络的特点分析：

3) 深度特征的层次性



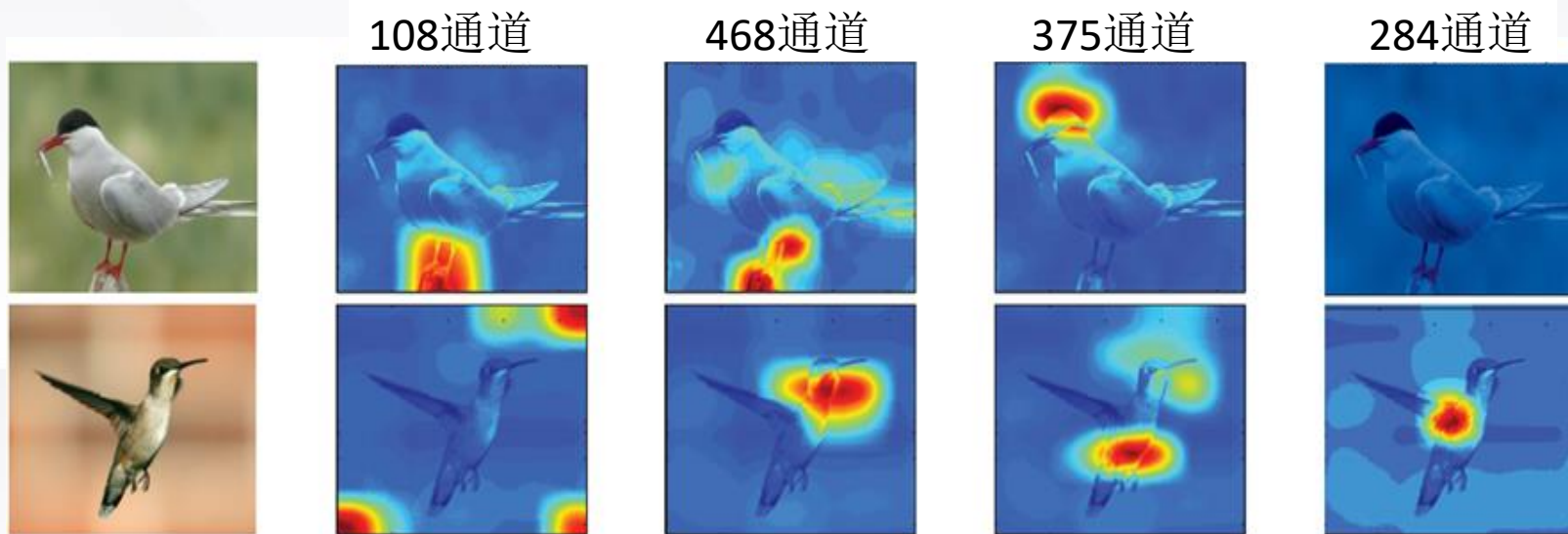
第四层和第五层，则更具分辨能力的模式被卷积网络所捕获

8.2 卷积网络结构与特性分析

➤ 卷积网络的特点分析：

4) 分布式表示

- 分布式表示是指语义概念到神经元的一种多对多映射。
- 神经网络具有分布式表示特性：每个语义概念由许多分布在不同神经元中被激活的模式表示，而每个神经元又可以参与到许多不同语义概念的表示中去。



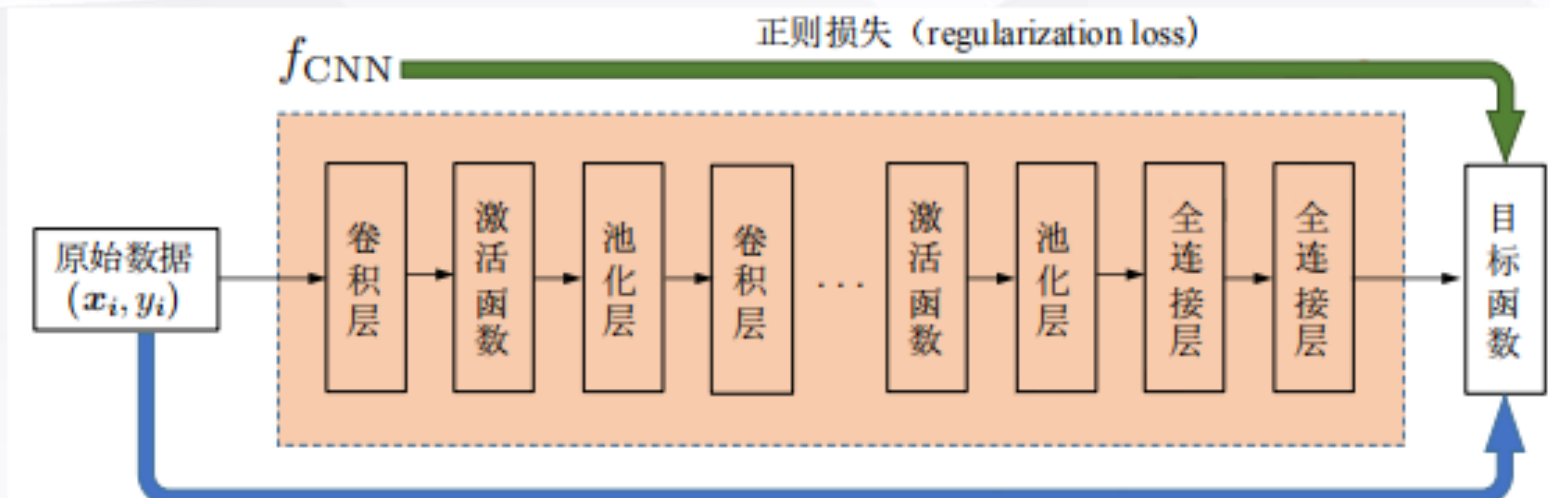
第八章 卷积神经网络分类器

CONTENTS PAGE

- 8.1 卷积神经网络的基本组成
- 8.2 卷积网络结构与特性分析
- 8.3 卷积网络的训练
- 8.4 经典的卷积网络结构
- 8.5 卷积网络结构设计举例

8.3 卷积网络的训练

- 在搭建的卷积网络基础上，需要用训练集对网络进行训练，确定网络参数，从而获得一个具体的网络模型。



- ① 损失函数设计
- ② 数据集处理
- ③ 网络训练

8.3 卷积网络的训练

① 损失函数 (loss)

- 也称为目标函数，通过样本的预测结果与真实标记之间产生的误差反向传播，指导网络学习，获得模型参数。
- 网络任务不同，损失函数的设计也不同，已经有很多种loss。如用于分类任务的loss有交叉熵loss、合页loss、中心损失函数、大间隔交叉熵loss等；用于回归任务的loss有 $l1$ loss、 $l2$ loss等。

8.3 卷积网络的训练

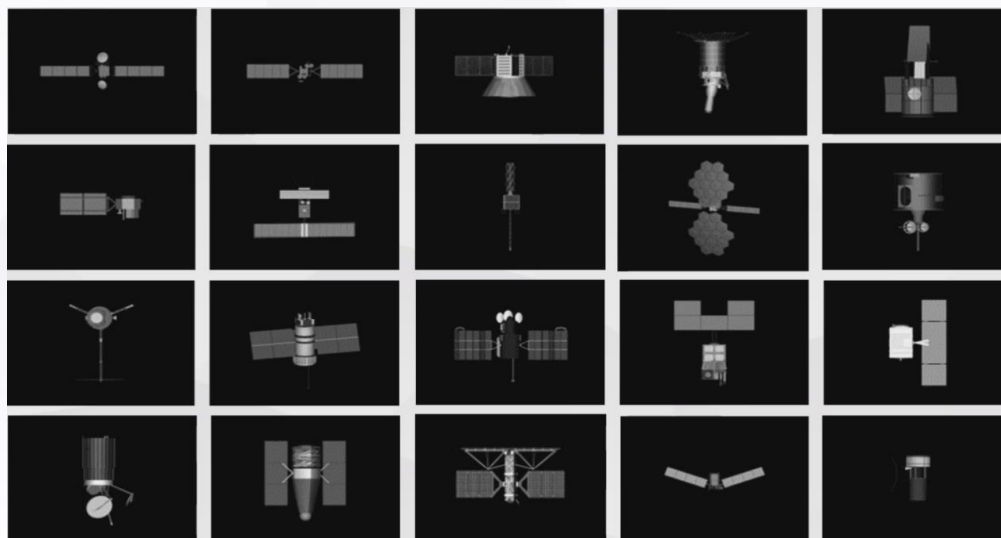
② 数据集处理

1) 数据集准备

经典的开源数据，比如ImageNet数据集和COCO数据集，当公开数据不能满足要求时，需要制作面向特殊应用的数据集。



COCO数据集



天基卫星目标数据集

8.3 卷积网络的训练

② 数据集处理

2) 数据增广

深度网络参数量大，需要大样本数据库进行训练。当训练样本不足时，可采用翻转（镜像）、旋转、随机抠取、尺度变换等方式进行扩充，扩充后数据量可达原有数据量的几十倍。



原图



水平反转



随机抠取

8.3 卷积网络的训练

② 数据集处理

3) 数据标准化

向卷积网络模型输入数据时，还需要进行数据标准化处理，得到均值为0、标准差为1的服从标准正态分布的数据，**移除共同部分，凸显个体差异。**

例如，均值方差标准化：常用的标准化方法

$$x^* = \frac{x - \mu}{\sigma}$$

其中， μ 为样本均值， σ 是样本标准差

8.3 卷积网络的训练

③ 网络训练

1) 超参数设置

进入训练之前，需要对超参数进行设置。

- 卷积层的超参数包括卷积核尺寸、数量及卷积的移动步长，实践中最常用的卷积核尺寸为 $3*3$ 、 $5*5$ 。
- 池化层的超参数主要是池化核尺寸、移动步长及池化方式。
- 其他需要设定的超参数。。。

8.3 卷积网络的训练

③ 网络训练

2) 参数初始化

卷积网络参数初始化定义了训练起点，常用的初始化方法：

➤ 均匀分布初始化

在Pytorch中，默认的参数初始化方法是 $\left[-\frac{1}{\sqrt{n_{\text{in}}}}, \frac{1}{\sqrt{n_{\text{in}}}}\right]$ 之间的均匀分布初始化，其中 n_{in} 为神经元个数

➤ Xavier初始化

是一种正态分布初始化方法，适合于卷积层参数的初始化：

$$w \sim N\left(0, \frac{2}{n_{\text{in}} + n_{\text{out}}}\right)$$

其中， n_{in} 表示前一层的输入维度， n_{out} 表示当前层的输出维度

8.3 卷积网络的训练

③ 网络训练

3) 网络优化算法

- 优化算法给出了如何根据损失函数的梯度迭代、更新网络参数，使目标函数逼近最优。
- 有多种经典的优化方法，如批量梯度下降法、随机梯度下降法、动量算法、AdaGrad法、Adadelat法、RMSProp法、Adam法等。
- 以批量梯度下降法为例，批量梯度下降法使用所有样本计算梯度，表达式为：

$$w^{(n)} = w^{(n-1)} - \eta \frac{1}{m} \sum_{i=1}^m \frac{\partial L(w, x_i)}{\partial w^{(n-1)}}$$

其中， $L(w, x_i)$ 为目标函数， m 为数据集样本总数， w 为各层参数， x_i 为第 i 个样本， η 为学习率。

8.3 卷积网络的训练

③ 网络训练

4) 批规范化 (Batch Normalization, BN)

➤ 问题引入:

对于具有多层结构的深度神经网络，训练过程中各层参数在一直变化，容易出现内部协变量移位(Internal Covariate Shift)的现象。

- 内部协变量移位概念：网络训练过程中参数不断改变导致后续每一层输入的分布也发生变化，称之为internal covariate shift（内部neuron的数据分布发生变化）。

- ##### ➤ BN层：针对卷积层后的特征图进行的操作，对每一层的输出特征图进行规范化。

8.3 卷积网络的训练

③ 网络训练

4) 批规范化 (Batch Normalization, BN)

➤ 批规范化的一般步骤：

S1：计算批处理数据均值 μ 。

S2：计算批处理数据方差 σ 。

S3：规范化。减均值除方差， $\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}}$

S4：尺度变换和偏移。 $y_i = \alpha \hat{x}_i + \beta$ ，两个参数通过学习得来。

8.3 卷积网络的训练

③ 网络训练

5) 网络正则化

- 针对损失函数进行的操作，防止参数过大而导致的过拟合。
- l_1 正则化和 l_2 正则化是两个常用的正则化方式，假设待正则化的网络层参数为 w_i ($0 < i \leq n$ ，是网络的所有参数)，它们的定义为：

$$l_1 = \sum_{i=1}^n |w_i| \quad l_2 = \frac{1}{2} \sum_{i=1}^n w_i^2$$

实际训练中参数更新用： $\text{loss} + \lambda l_2$

其中 λ 控制正则项的大小，较大的取值将较大程度地约束模型的复杂度。

8.3 卷积网络的训练

④ 从头学习和迁移学习

1) 从头学习

卷积网络需要大样本量数据进行训练，当有足够的训练样本时，可以采用从头学习方式训练，即从头训练网络模型，从0开始学习。

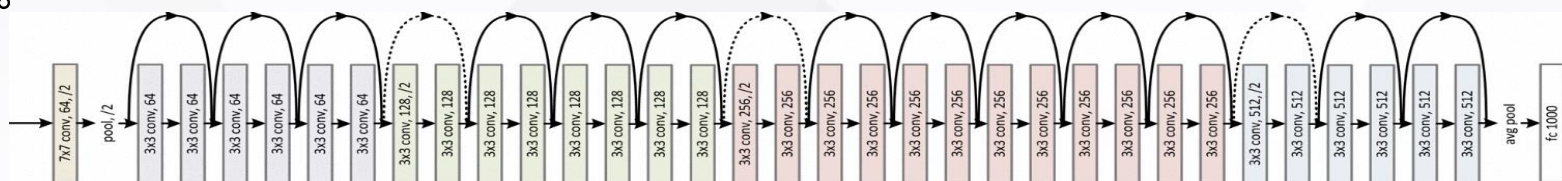
8.3 卷积网络的训练

④ 从头学习和迁移学习

2) 迁移学习

当训练样本不足时，可以采用迁移学习的方式。迁移学习是把已训练好的模型参数迁移到新的模型来帮助新模型训练，从而加快并优化模型的学习效率。

例如，采用残差网络进行飞机目标分类，可以用原来在ImageNet上的参数做初始值，然后用飞机样本继续训练网络，解决飞机样本不足问题。



残差网络模型

据研究，数据集样本足够多时，从头学习和迁移学习差别不大。

第八章 卷积神经网络分类器

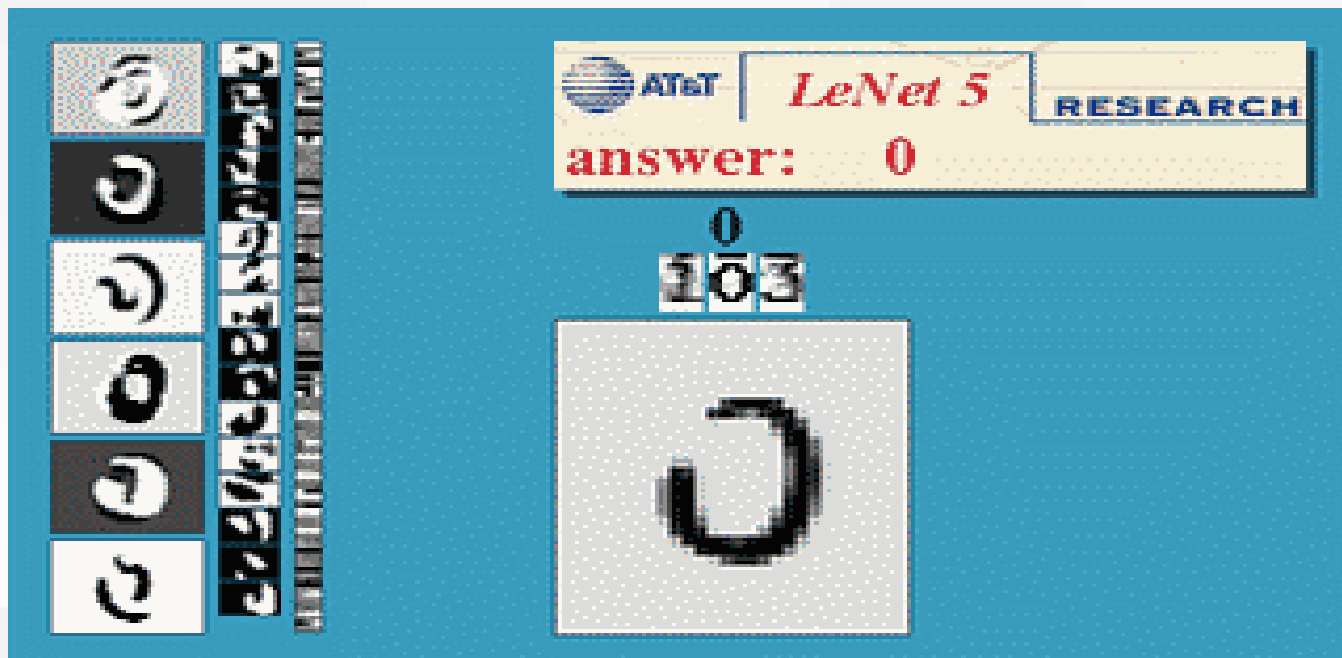
CONTENTS PAGE

- 8.1 卷积神经网络的基本组成
- 8.2 卷积网络结构与特性分析
- 8.3 卷积网络的训练
- 8.4 经典的卷积网络结构
- 8.5 卷积网络结构设计举例

8.4 经典的卷积神经网络结构

① LeNet网络

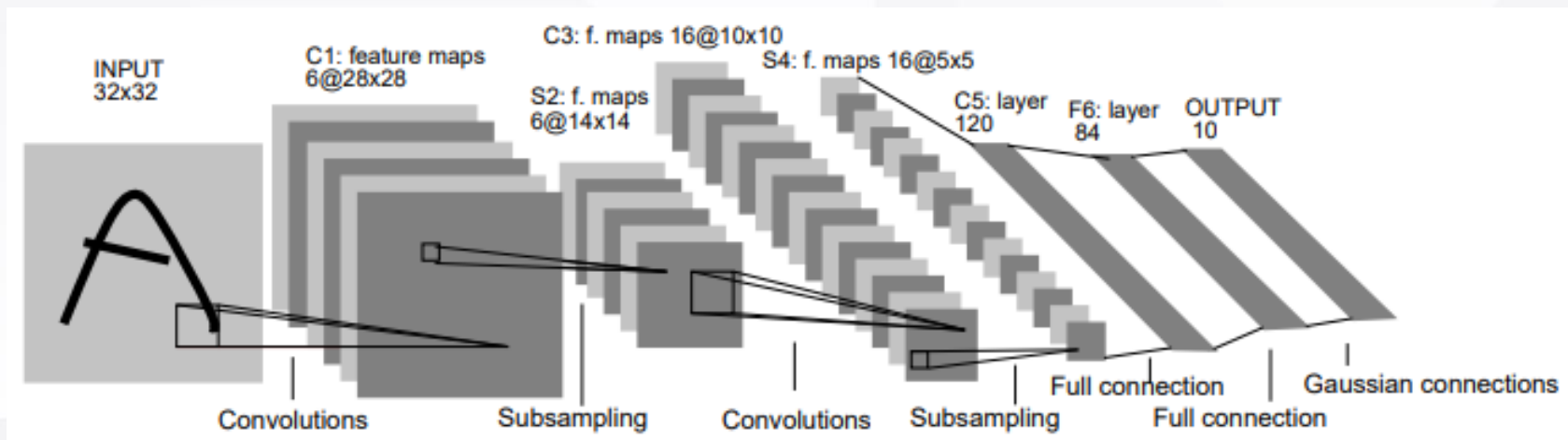
LeNet是早期卷积神经网络的代表之作，由Yann LeCun提出，用于手写数字识别。LeNet有几个版本，介绍最为熟知的LeNet5，其发表于1998年。



8.4 经典的卷积网络结构

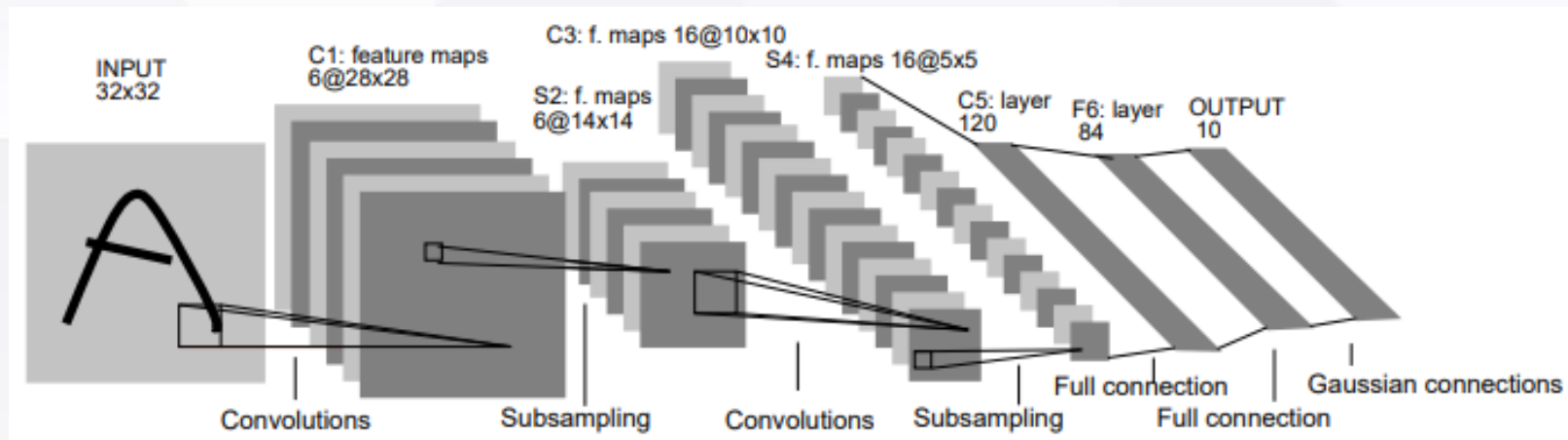
① LeNet网络

- LeNet-5网络共有7层（不包含输入层），包含了卷积层、池化层、全连接层等卷积神经网络的基本模块。
- 输入图像大小为 $32 \times 32 \times 1$ ，输出为对应10个类别（数字0-9）的概率。



8.4 经典的卷积网络结构

① LeNet网络



- **C1层**：卷积核大小为 5×5 ，通道数为1，步长为1，不采用数据填充，则6个卷积核得到6个 28×28 的特征图。
- **S2层**：池化核大小为 2×2 ，步长为2。C1层经过池化，特征图降维 $14 \times 14 \times 6$ 。
- **C3层**：卷积核大小为 5×5 ，通过对S2层输出的6个特征图进行如表所示的特殊组合计算得到C3层16个特征图。

8.4 经典的卷积网络结构

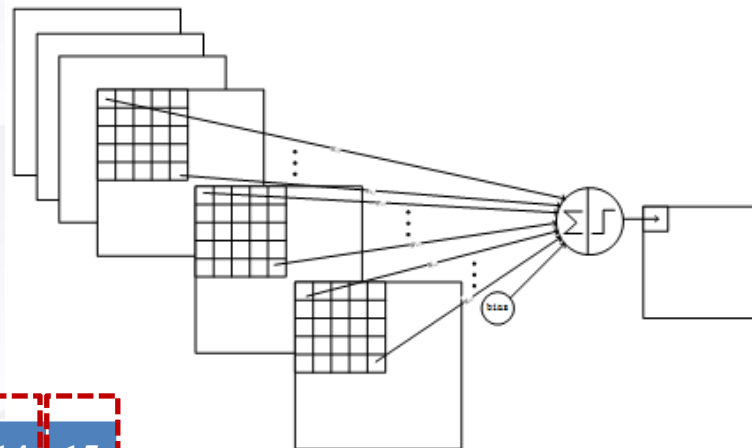
① LeNet网络

➤ C3 卷积层

S2层

C3层

	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	X

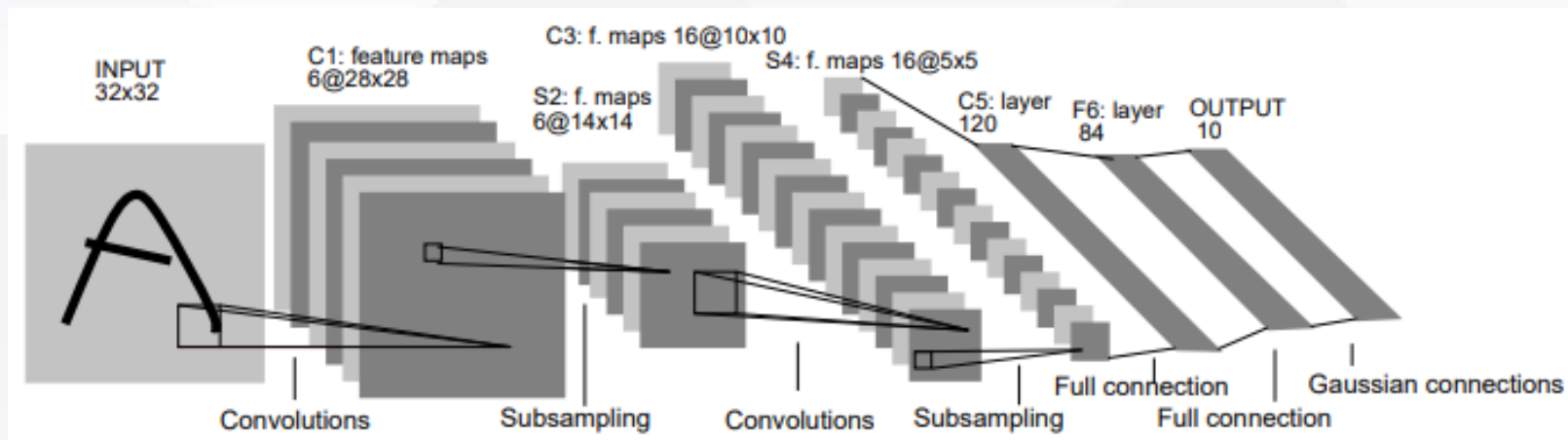


C3与S2中前3个图相连的卷积结构

- 这种不对称的组合连接方式有利于提取多种组合特征。
- 卷积核多少种？ $3*6+4*9+6*1=60$
- 输出特征图大小：10*10， $(14-5)/1+1=10$ ，卷积核大小为5*5
- 有 $(5*5*3+1)*6 + (5*5*4+1)*9 + (5*5*6+1)*1 = 1516$ 个训练参数。

8.4 经典的卷积网络结构

① LeNet网络



- **S4层**：与S2操作相同，得到16个 5×5 的特征图
- **C5层**：卷积核大小为 5×5 ，得到120个 1×1 的特征图
- **F6层**：是84个单元的全连接层
- **输出层**：全连接层，输出10个值，对输入数字为0~9的概率值，实现手写输入数字的自动识别。

8.4 经典的卷积网络结构

① LeNet网络

LeNet采用的结构、部件至今仍被广泛使用。可以说，LeNet是学习卷积神经网络的“Hello World”，是学习其他网络结构的基础。

LeNet各层操作的总结

Layer		Feature Map	Size	Kernel Size	Stride	Activation
Input	image	1	32x32	-	-	-
1	Convolution	6	28x28	5x5	1	tanh
2	Average Pooling	6	14x14	2x2	2	tanh
3	Convolution	16	10x10	5x5	1	tanh
4	Average Pooling	16	5x5	2x2	2	tanh
5	Convolution	120	1x1	5x5	1	Tanh
6	FC	-	84	-	-	tanh
Output	FC	-	10	-	-	softmax

8.4 经典的卷积网络结构

② AlexNet网络

2012年提出，第一个现代深度卷积神经网络模型。2012年ImageNet挑战赛冠军。具有历史意义的网络结构，在AlexNet之前，深度学习已经沉寂了很久。

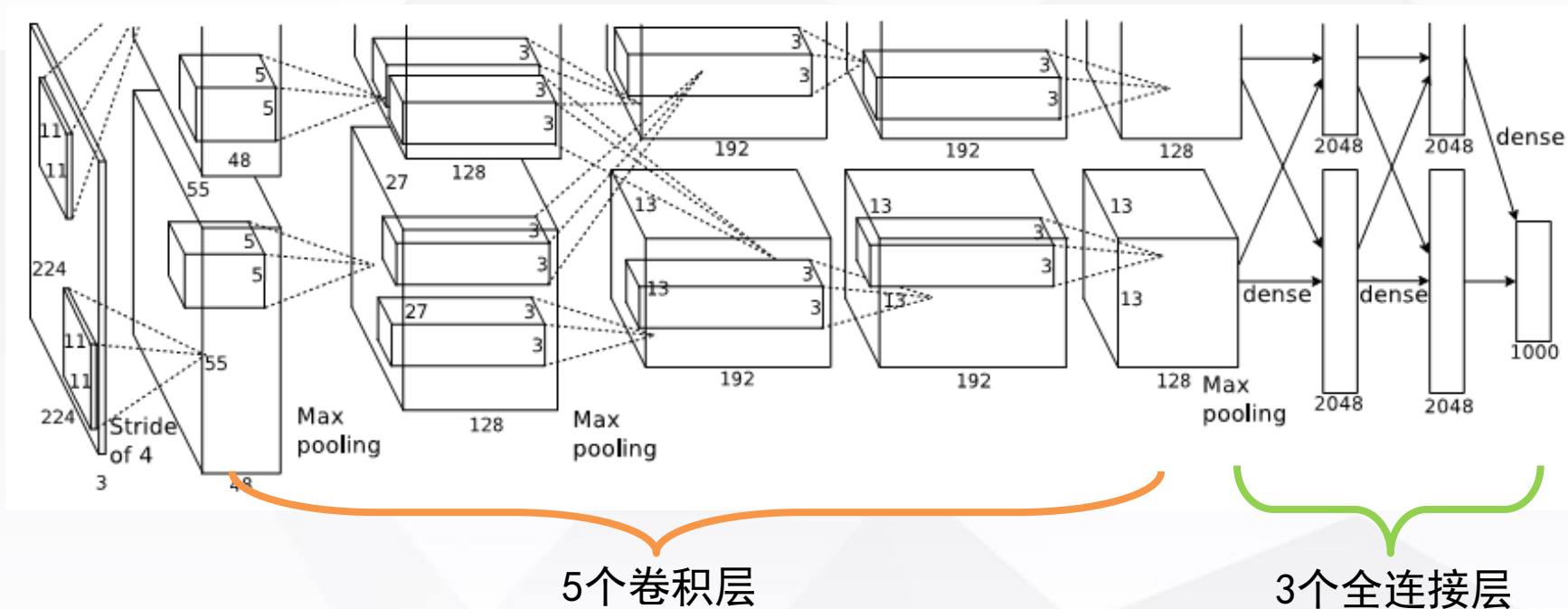
- 2012年，ImageNet图像分类挑战赛中，AlexNet模型赢得了冠军，top-5错误率（Top 5误差率是指给定一张图像，其标签不在模型认为最有可能的5个结果中的几率）15.4%；
- 错误率比上一年的冠军下降了十个百分点，而且远远超过当年的第二名（26.2%）。



A. Krizhevsky, L. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Proc.Neural Information Processing Systems,2012.

8.4 经典的卷积网络结构

② AlexNet网络



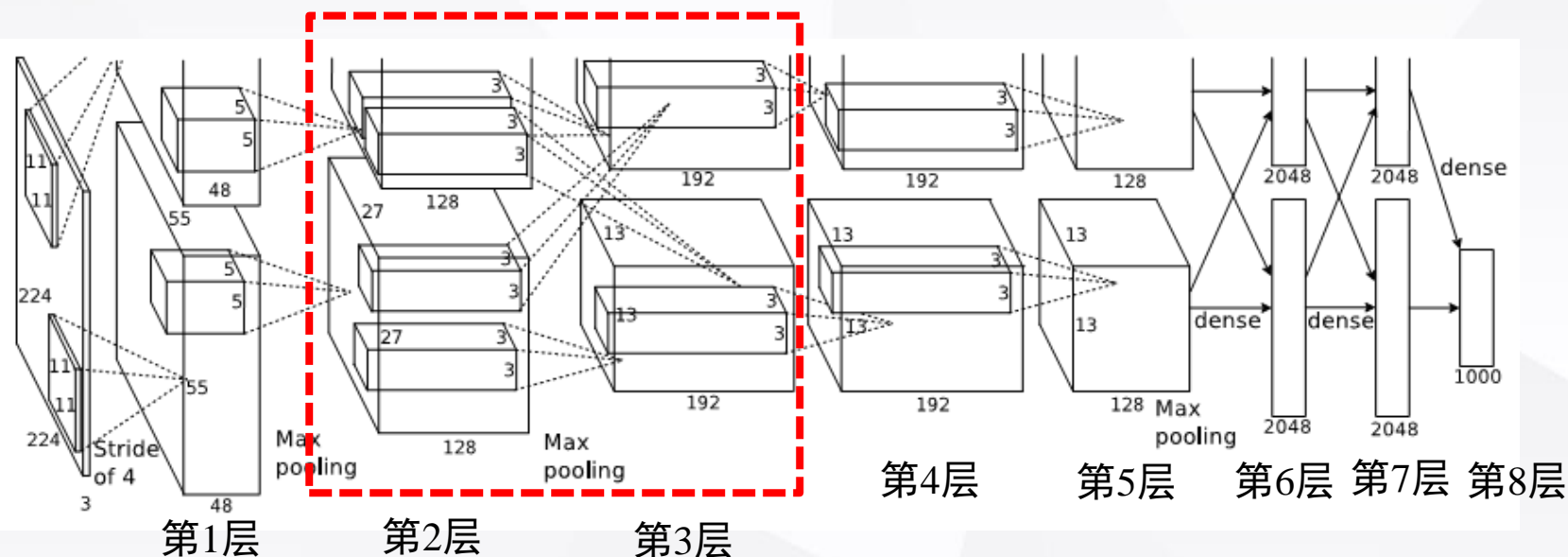
输入：227*227*3的彩色图像，输出：1000个类别的概率

8.4 经典的卷积神经网络结构

② AlexNet网络

➤ 多GPU加速训练

- 将网络分布在两个GPU上。
- 采用并行方案：在每个GPU中放置一半核（或神经元），GPU间的通讯只在某些层进行。



8.4 经典的卷积网络结构

③ VGGNet网络

2014年牛津大学计算机视觉组和DeepMind公司共同研发，2014年ILSVRC比赛上获得了分类项目的第二名和定位项目的第一名。

- 最深做到了19层；

<Very Deep Convolutional Networks for Large-Scale Image Recognition>, 2014.

VggNet一共有六种不同的网络结构，分别为11层，13层，16层，19层。在训练高级别的网络时，可以先训练低级别的网络，用前者获得的权重初始化高级别的网络，可以加速网络的收敛。

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
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

8.4 经典的卷积网络结构

③ VGGNet网络

➤ 特点分析

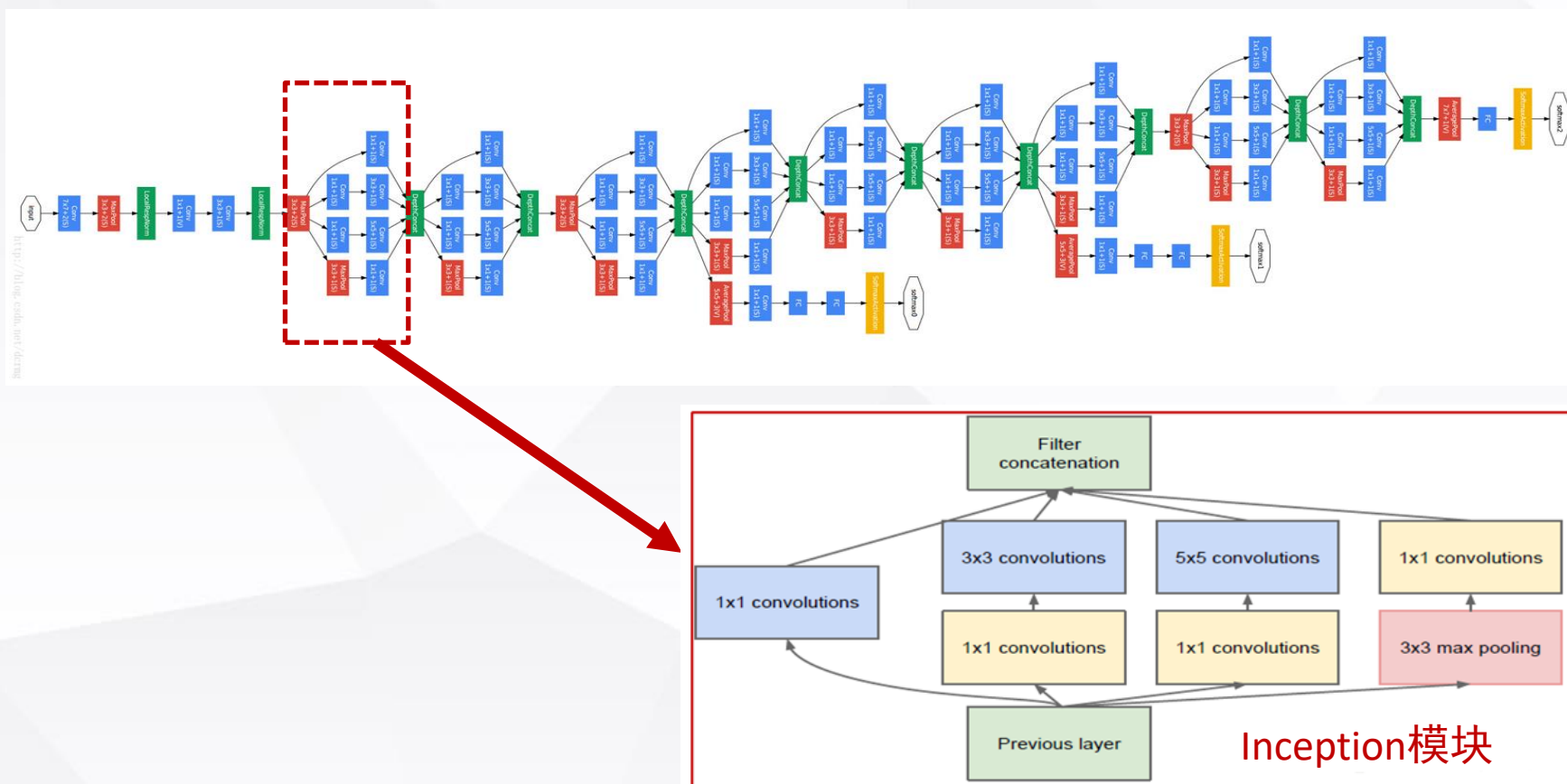
VGGNet使用多层小卷积核替代较大的卷积核。VGGNet用 3×3 的卷积核，代替了AlexNet中的大卷积核（ 5×5 、 11×11 ），加深了网络深度，增加了网络特征提取能力。

可堆叠的块状网络结构。VGGNet多次重复使用同一尺寸的卷积核，这种通过重复使用简单的基础模块来构建深度网络的思想，在VGGNet之后被广泛使用。

8.4 经典的卷积神经网络结构

④ GoogLeNet网络

谷歌团队提出，在2014的ImageNet分类任务上夺得冠军。引入一种网中网结构（Inception模块），网络达到了22层。



8.4 经典的卷积网络结构

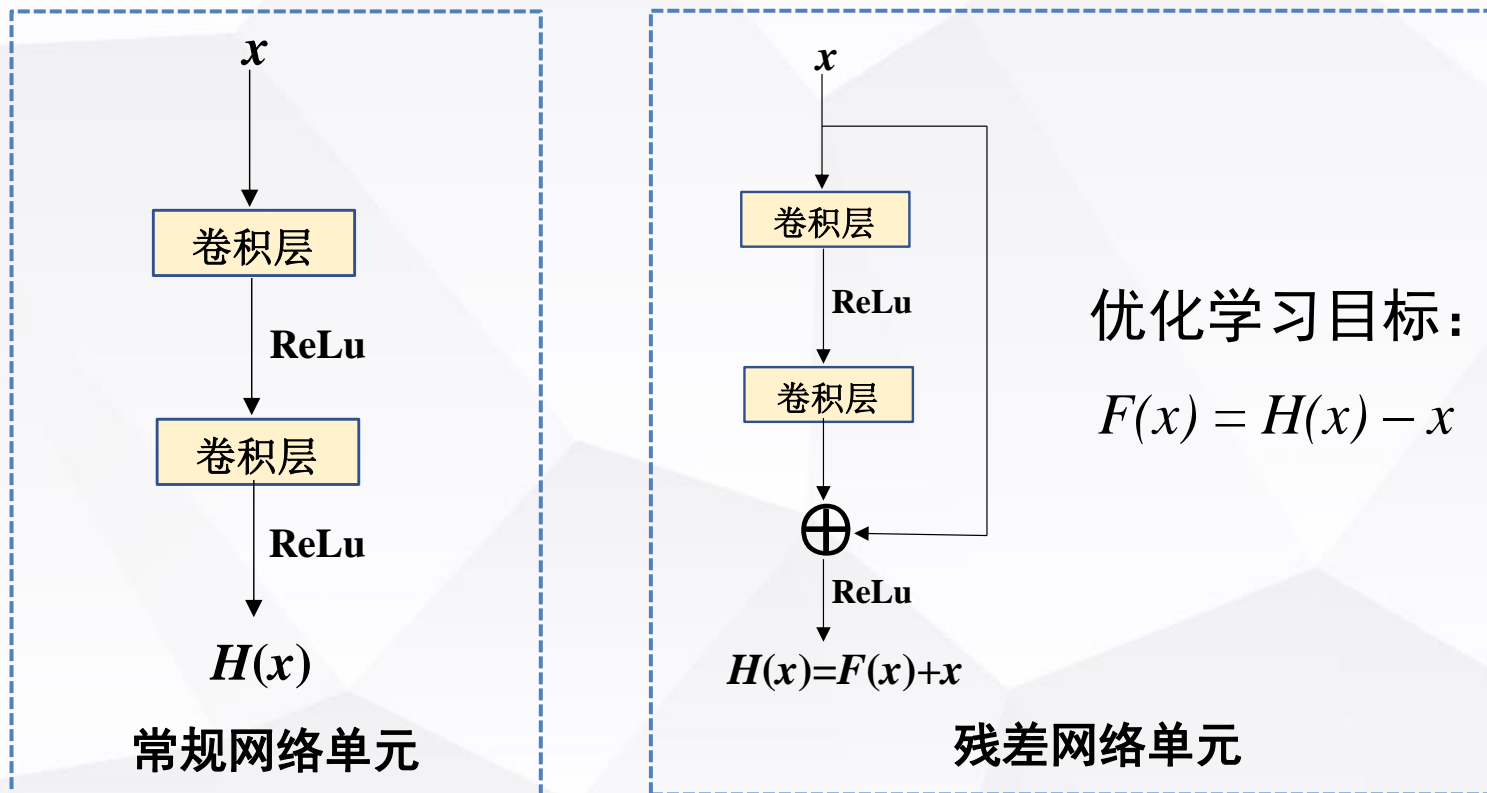
⑤ ResNet（残差网络）

- 从LeNet到AlexNet，再到VGGNet和GoogLeNet，网络层级越来越深，性能越来越好。更深的网络可以提取更加复杂的图像特征，能够更好、更全面地描述目标。
- 但是，卷积网络在信息传递的时候会存在信息丢失/失真，同时还会导致梯度消失或者梯度爆炸，从而很深的网络无法训练。
- 为了解决这一问题，何凯明等于2015年提出了残差网络结构。

8.4 经典的卷积网络结构

⑤ ResNet（残差网络）

- 重构学习过程：引入全新的“**残差学习**”原则，指导设计网络。

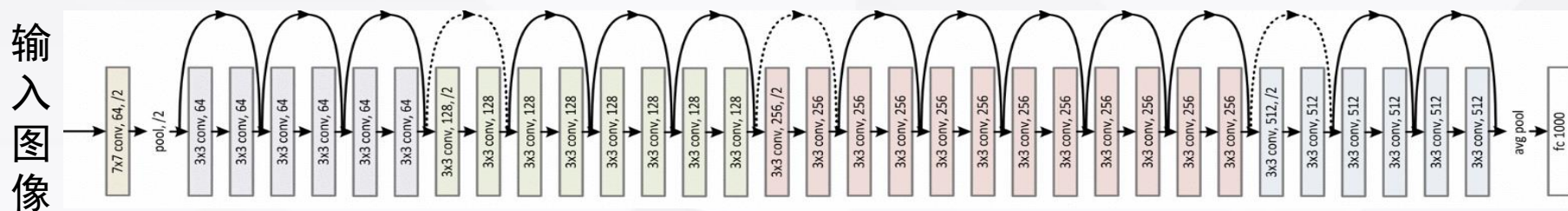


信号输入和输出之间增加了一条连线，使得网络**只需要学习输入、输出差别的那部分**，从而简化学习目标和难度。

8.4 经典的卷积神经网络结构

⑤ ResNet（残差网络）

- 将残差块堆叠在一起，组成残差网络结构，残差网络可以达上百层，下图给出的是34层的残差网络。

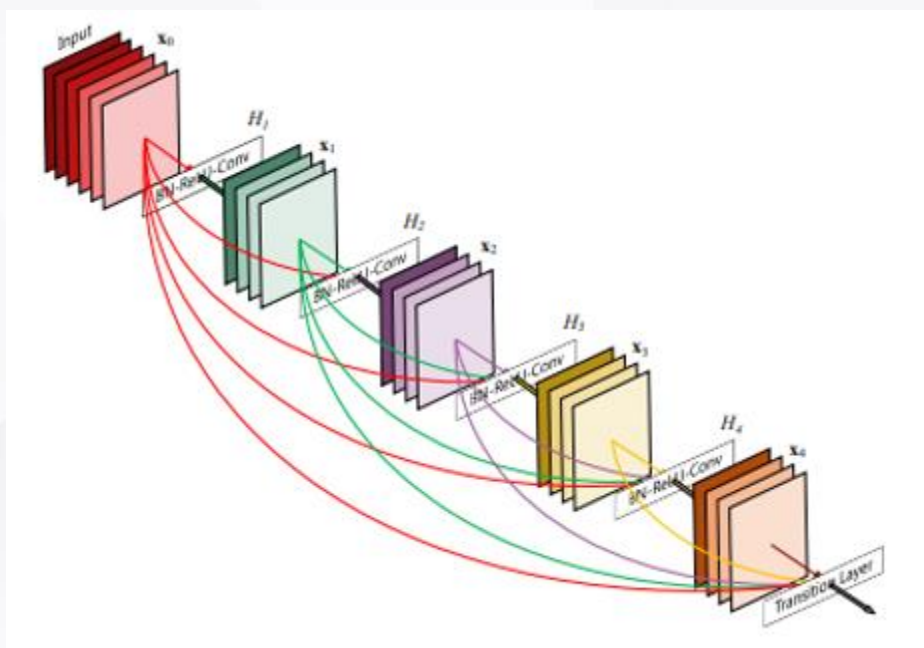


一个34层的残差网络

8.4 经典的卷积网络结构

⑥ DenseNet

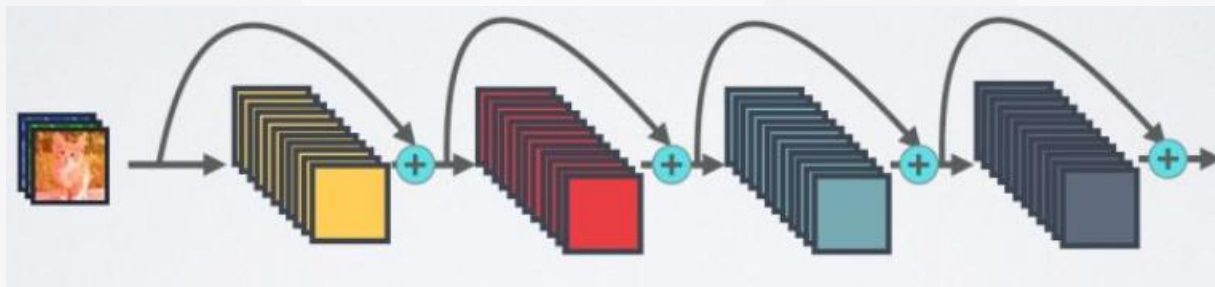
- DenseNet (Densely connected convolutional networks) 模型;
- 其基本思路与ResNet一致，但是它建立的是前面所有层与后面层的密集连接 (dense connection)，如图所示：
- DenseNet网络中的每个层都会接受其前面所有层作为其输入。



8.4 经典的卷积网络结构

⑥ DenseNet

- ResNet 网络和 DenseNet 网络的连接机制对比。相比 ResNet，DenseNet 是一种密集连接。而且 DenseNet 直接 concat 来自不同层的特征图，可以实现特征重用，提升效率。



ResNet网络的短路连接机制

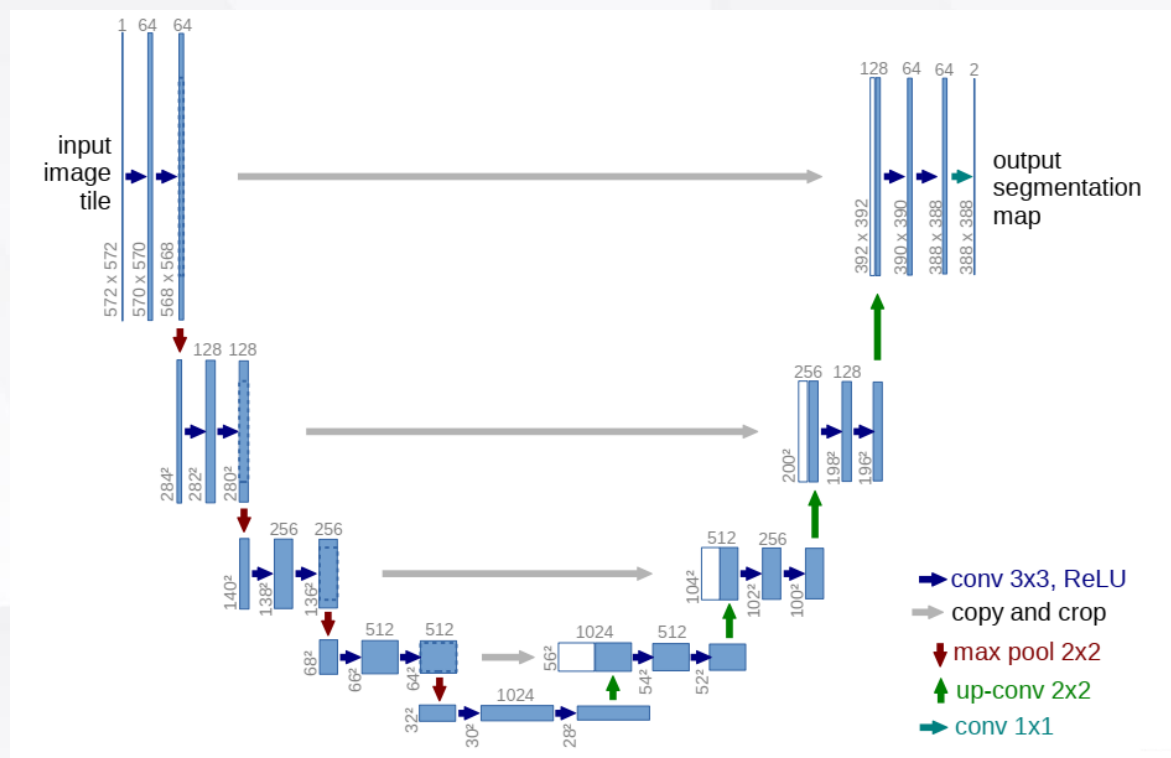


DenseNet网络的密集连接机制

8.4 经典的卷积网络结构

⑦ Unet网络

- 一种用于图像分割的网络——Unet网络。
- 基于编码器（encode）-解码器（decode）思想，通过四个下采样提取目标特征，再通过四个上采样，最后逐个对其像素点进行分类。



第八章 卷积神经网络分类器

CONTENTS PAGE

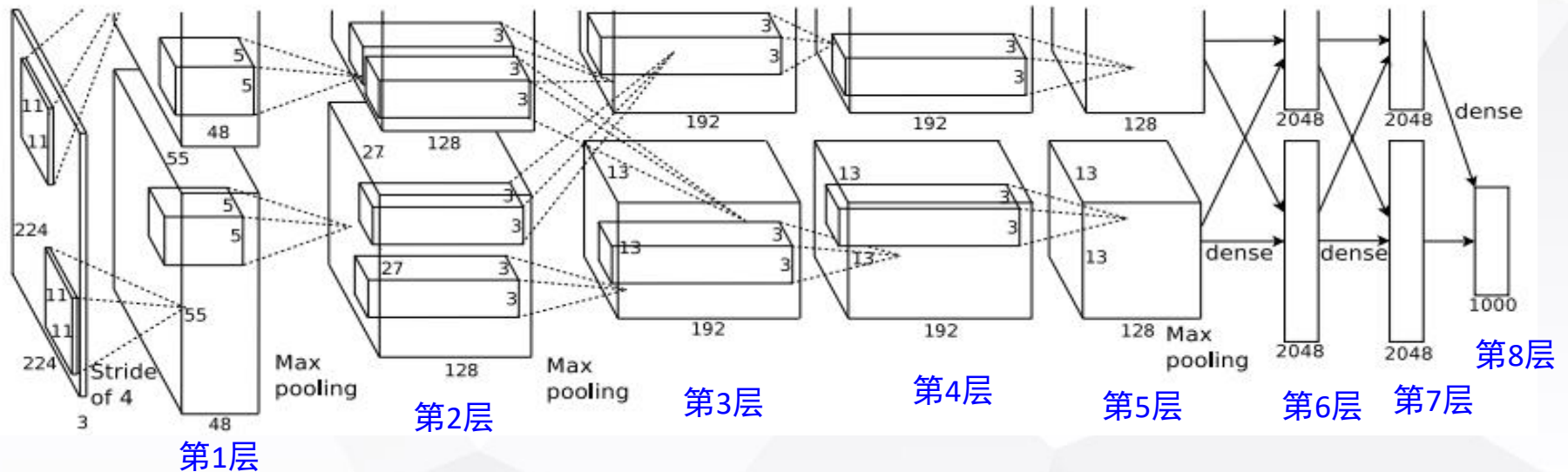
- 8.1 卷积神经网络的基本组成
- 8.2 卷积网络结构与特性分析
- 8.3 卷积网络的训练
- 8.4 经典的卷积网络结构
- 8.5 卷积网络结构设计举例

8.4 经典的卷积网络结构

参见智慧树视频讲解（自学）！

第 8 章 作业

假设所有层卷积核均为 3×3 ，计算如下网络每层的参数个数（不考虑偏置量），两个分支（两个GPU）的相同层按一层算（求和）。



计算举例：以第1层为例，该层神经元布置在2个GPU上，每个GPU上的特征图通道数为48，若采用 3×3 卷积，该层参数个数： $(3 \times 3 \times 3) \times 48 \times 2$ ；

本学期课程结束！