

卷积神经网络

Convolutional Neural Network

简称 CNN

本节学习目录

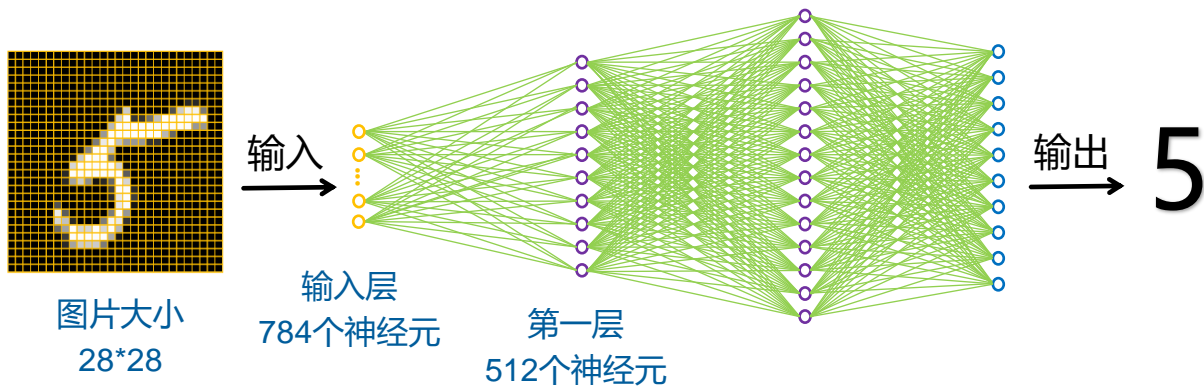
- 01 卷积神经网络由来
- 02 卷积神经网络结构组成
- 03 卷积层运算详解
- 04 池化层运算详解
- 05 常见卷积神经网络结构
- 06 经典卷积神经网络

01. 卷积神经网络的由来

01.卷积神经网络的由来

用全连接网络做图片分类，有哪些问题？

思考：第一层的权重数量有多少？计算第一层需要做多少次乘法？



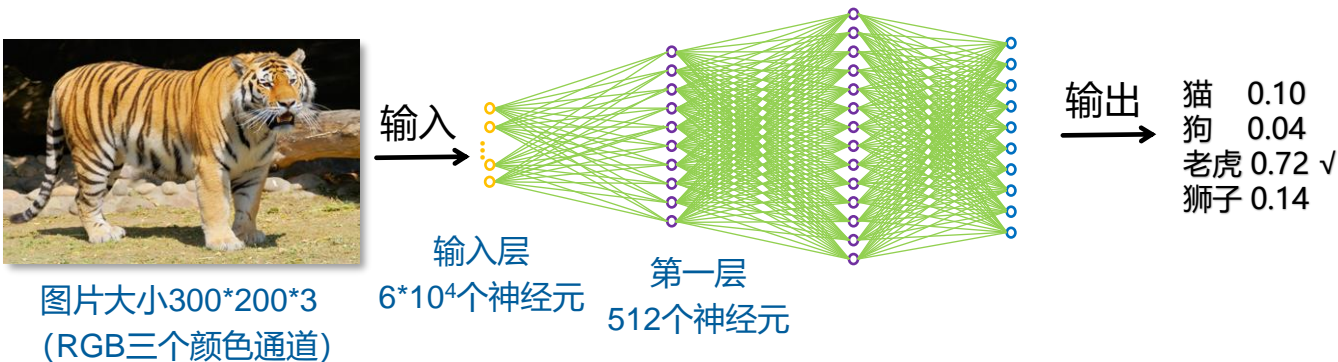
第一层权重数：784*512=401,408 (复杂度 $O(MN)$)

第一层乘法数：784*784*512=314,703,872 (复杂度 $O(M^2N)$)

01.卷积神经网络的由来

用全连接网络做图片分类，有哪些问题？

思考：第一层的权重数量有多少？计算第一层需要做多少次乘法？



第一层权重数： $1.8 \times 10^5 \times 512 \approx 9 \times 10^8$ (复杂度 $O(MN)$)

第一层乘法数： $1.8 \times 10^5 \times 1.8 \times 10^5 \times 512 \approx 1.6 \times 10^{13}$ (复杂度 $O(M^2N)$)

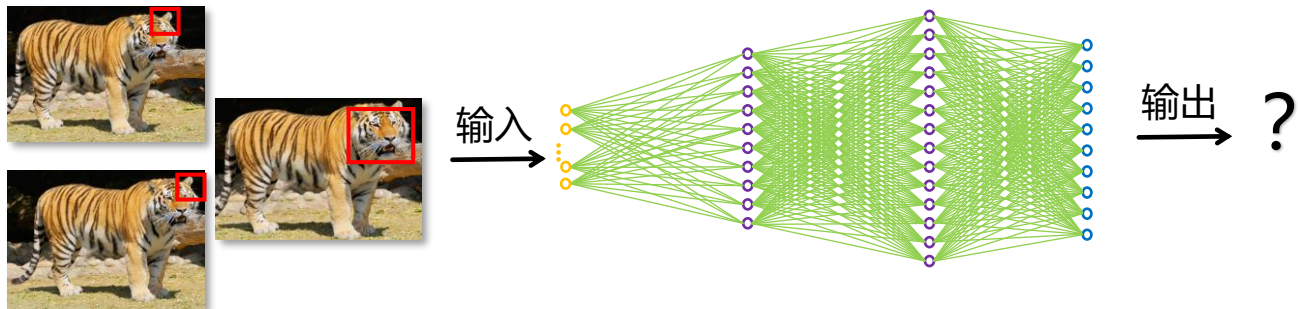
对自然图像使用全连接神经网络

问题①：参数数量过多，计算复杂度高

01.卷积神经网络的由来

用全连接网络做图片分类，有哪些问题？

思考：如果老虎向前或向后走一步，图片放大缩小了，对全连接网络有什么影响？



全连接网络把图像拉成一维向量，忽视图像中的形状，
全连接网络无法捕捉图像局部区域的关联信息（即图像的空间信息）

当图像局部区域位置变化时，会与全连接层不同的权重相乘

对自然图像使用全连接神经网络，问题②：丢失图像的空间信息

01.卷积神经网络的由来

对自然图像使用全连接神经网络的问题：

- ①参数数量过多，计算复杂度高
- ②丢失图像的空间信息

卷积的特点：

- ①局部连接
- ②参数共享



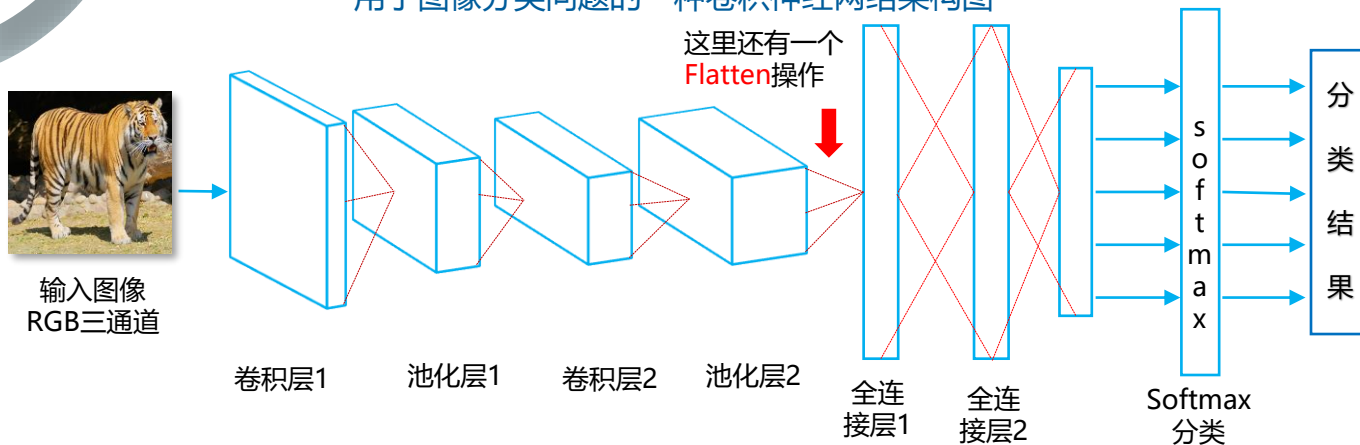
卷积的优势

- ①参数数量减少，计算量减小
- ②可以捕捉图像的空间信息

02.卷积神经网络结构组成

02. 卷积神经网络结构组成

用于图像分类问题的一种卷积神经网络架构图

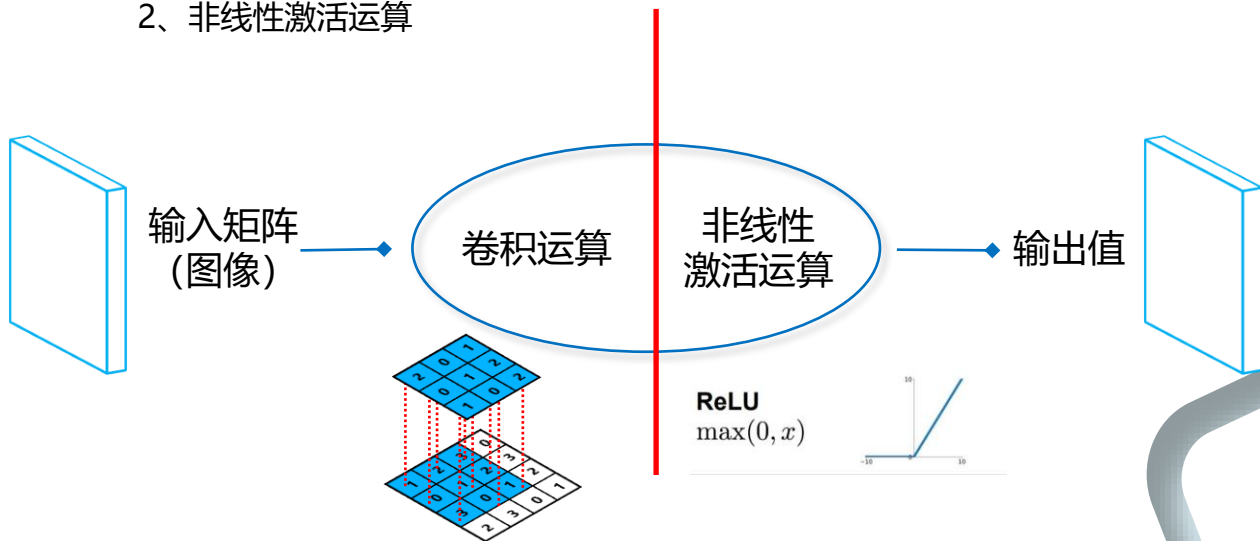


CNN的组成：输入层、卷积层、池化层、全连接层、softmax分类
其中卷积层和全连接层中包括激活函数（如ReLU）

03. 卷积层运算详解

03. 卷积层 - 运算过程

- 1、卷积运算
- 2、非线性激活运算

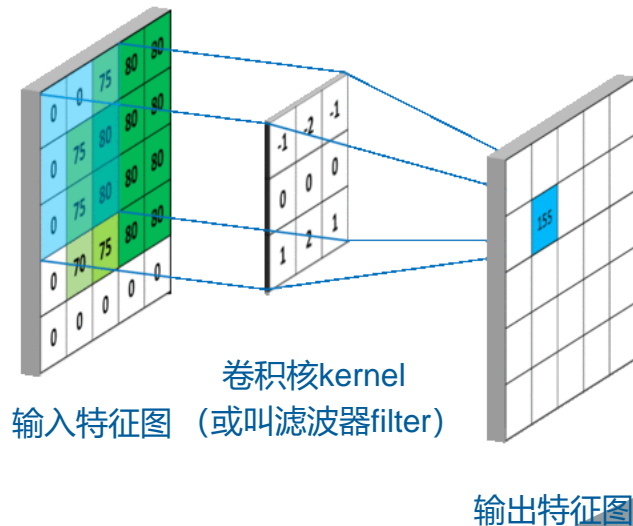


03. 卷积层 - 卷积运算演示

卷积的动态示例

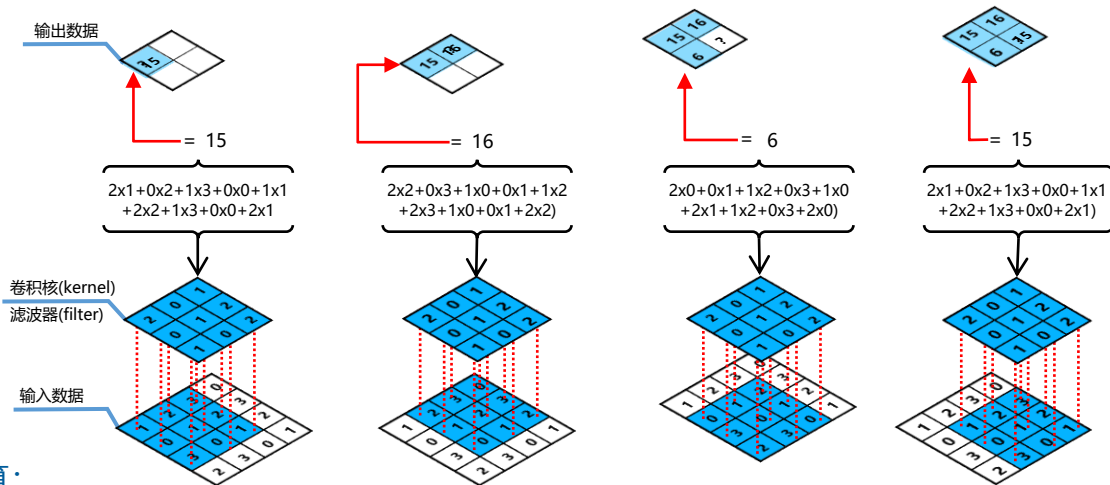
卷积层的输入输出都是三维矩阵，称为特征图
卷积核中的数值就是卷积层的权重
卷积核一般尺寸很小（如 3×3 ）

卷积的计算：
卷积核与输入特征图对应位置的像素相乘后再求和
从左到右、从上到下一行一行扫描计算相乘求和



03. 卷积层 - 卷积运算分步详解

输入4*4
卷积核3*3
输出2*2



卷积的计算：

卷积核与输入特征图对应位置的像素相乘后再求和

从左到右、从上到下一行一行扫描计算相乘求和

计算：输出大小=输入大小-卷积核大小+1

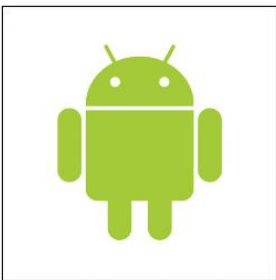
$m = n - k + 1$ ，输入大小 n ，卷积核 k ，输出大小 m

卷积覆盖图像的一个局部区域：
捕捉图像的空间信息（上下左右）

03. 卷积层 – 卷积的作用

卷积核的作用是提取图像的局部特征
不同的卷积核提取不同的局部特征

边缘检测示例



输入图像
(由不同方向的线条构成)

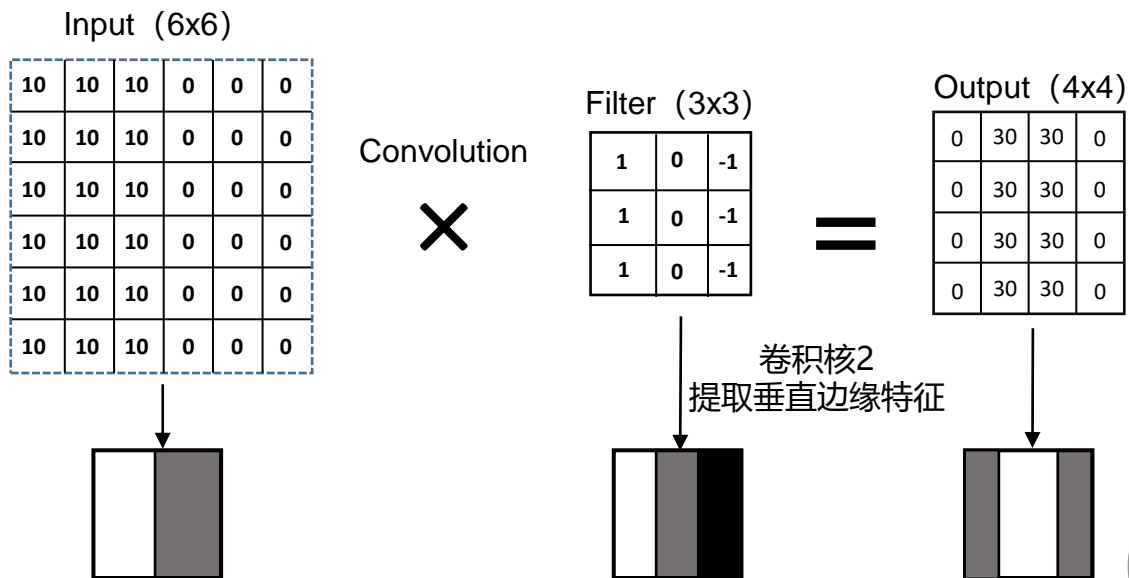
1	1	1
0	0	0
-1	-1	-1

卷积核1
提取水平边缘

1	0	-1
1	0	-1
1	0	-1

卷积核2
提取垂直边缘

03. 卷积运算—边缘检测原理

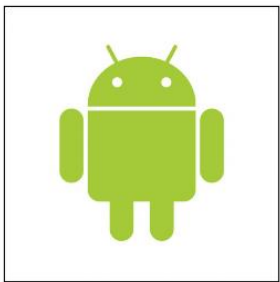


03. 卷积层 – 卷积的作用

卷积核的作用是提取图像的局部特征
不同的卷积核提取不同的局部特征

面试问题：卷积核的作用是什么？

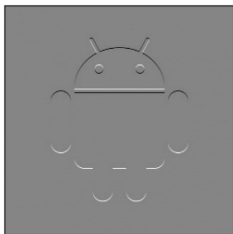
边缘检测示例



输入图像
(由线条构成)

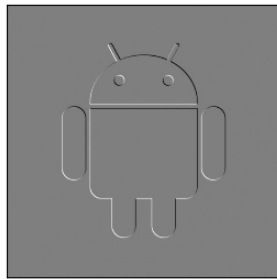
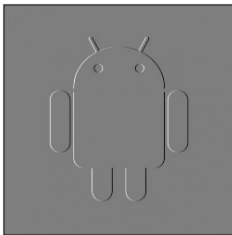
1	1	1
0	0	0
-1	-1	-1

卷积核1
提取水平边缘特征



1	0	-1
1	0	-1
1	0	-1

卷积核2
提取垂直边缘特征



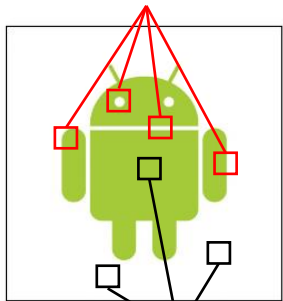
融合特征
(物体边缘)

03. 卷积层 – 卷积的作用

为什么要提取局部特征？

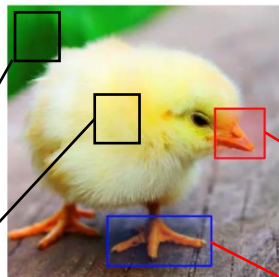
图像的属性往往由局部特征决定

有用的局部特征：
边缘，角点，形状

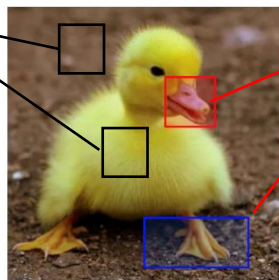


无用的局部特征：
大片相同纹理，背景

无用的局部特征：
大片相同纹理，背景



有用的局部特征：
嘴的形状（尖/扁）
脚的形状（是否有蹼）



03. 卷积层 – 卷积的特点和优势

卷积的特点:

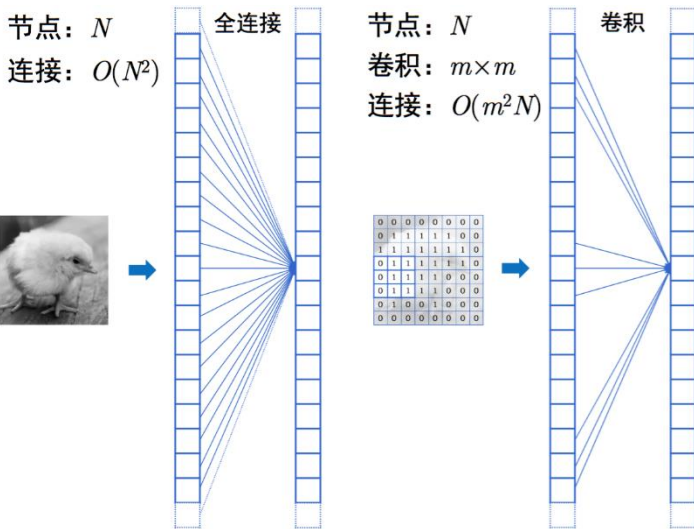
- ①局部连接
- ②参数共享

03. 卷积层 – 卷积的优势

局部连接

卷积可视为局部连接的全连接层（不参与运算的权重为0）

相比对应的全连接，卷积的参数量和计算量大大减少



面试问题1: 卷积层的局部连接特点指什么?

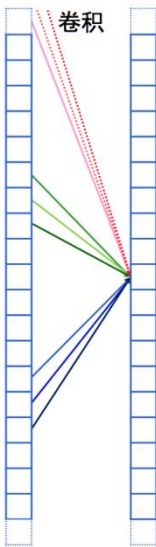
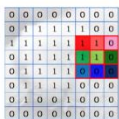
面试问题2: 卷积层的局部连接特性带来了什么好处?

03. 卷积层 – 卷积的优势

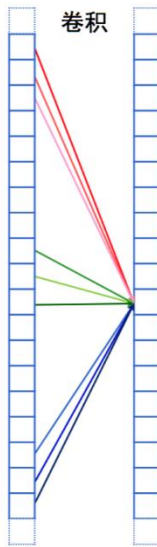
参数共享

同一个卷积核，扫描特征图不同位置时，使用相同的参数
相比对应的全连接，卷积的参数量大大减少

全连接: N^2
卷积核: m^2
 $O(N^2) \rightarrow O(m^2)$



节点: N
卷积: $m \times m$
连接: $O(m^2)$

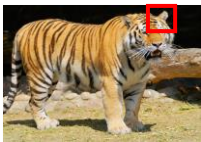


03. 卷积层 – 卷积的优势

参数共享

同一个卷积核，扫描特征图不同位置时，使用相同的参数
相比对应的全连接，卷积的参数量大大减少

等变表示：局部特征所处的位置不影响卷积的结果



面试问题1：卷积层的
参数共享特点指什么？

面试问题2：卷积层的
参数共享特性带来了
什么好处？

图像中物体的位置不会影响对物体的识别
符合人类理解图像的习惯

03. 卷积层 – 卷积的特点和优势

对自然图像使用全连接神经网络的问题：

- ①参数数量过多，计算复杂度高
- ②丢失图像的空间信息

卷积的特点：

- ①局部连接
- ②参数共享



卷积的优势

- ①参数数量减少，计算量减小
- ②可以捕捉图像的空间信息
- ③等变表示：局部特征所处的位置不影响卷积的结果

面试问题1：卷层相比全连接层，主要的特点/不同点有哪些？

面试问题2：全连接神经网络的主要问题/缺陷是什么？

面试问题3：相比之下卷积网络的优势是什么？

03. 卷积运算-填充padding

思考：

经过多次卷积运算后，输出特征图会变小，会丢失图像边界处的特征
如何保持特征图大小不变？

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

输入: 4*4

×

2	0	1
0	1	2
1	0	2

过滤器: 3*3

=

15	16
6	15

输出值: 2*2

计算：输出大小=输入大小-卷积核大小+1

$m=n-k+1$ ，输入大小 n ，卷积核 k ，输出大小 m

$k>1 \Rightarrow m<n$

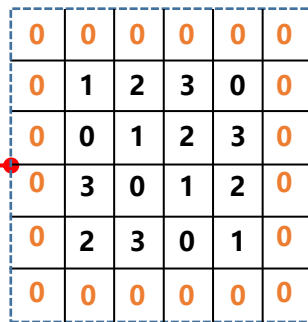
03. 卷积层 - 卷积运算填充(Padding)

Padding填充

向周围填充固定数据 (比如0)

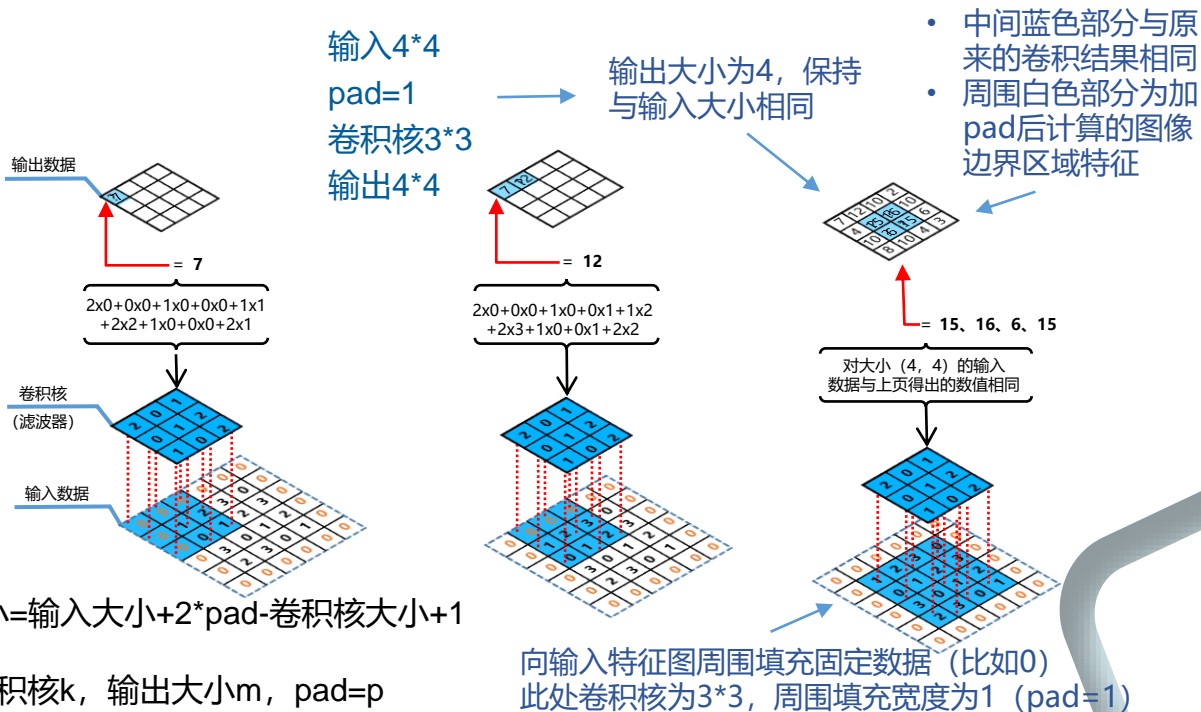
然后进行卷积运算同上:

对应的位置相乘, 乘积再求和



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

03. 卷积层 - 填充后卷积运算详解



03. 卷积层 - Valid卷积与Same卷积

Valid 卷积：不填充

$$m=n-k+1$$

输出特征图变小，丢失图像边界处的特征

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

输入: $n \times n$

*

2	0	1
0	1	2
1	0	2

卷积核: $k \times k$

=

15	16
6	15

输出: $m \times m$

面试问题1：卷层运算时，填充和不填充，分别如何计算输出大小？

Same 卷积：填充

$$m=n+2p-k+1$$

输出特征图不变，不会丢失图像边界处的特征

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

输入: $n \times n$

*

2	0	1
0	1	2
1	0	2

卷积核: $k \times k$

=

11	15	13	9
7	15	16	15
13	6	15	11
13	13	7	9

输出: $m \times m$

面试问题2：卷层运算时，不填充有什么问题？填充有什么好处？

03. 卷积层 - 卷积运算填充数量原则

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

输入: $n \times n$

*

2	0	1
0	1	2
1	0	2

卷积核: $k \times k$

=

11	15	13	9
7	15	16	15
13	6	15	11
13	13	7	9

输出: $m \times m$

要保持输出大小=输入大小 (即 $m=n$, 其中 $m=n+2p-k+1$)

Valid卷积填充宽度公式: $p = (k-1) / 2$

3*3卷积, $pad=1$

5*5卷积, $pad=2$

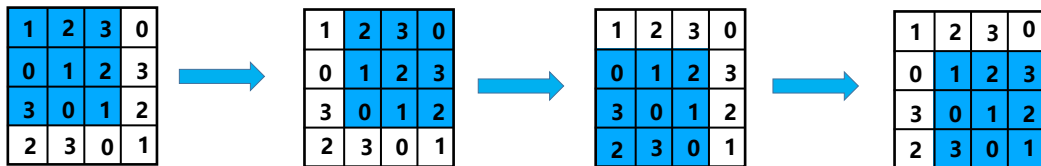
7*7卷积, $pad=3$

.....

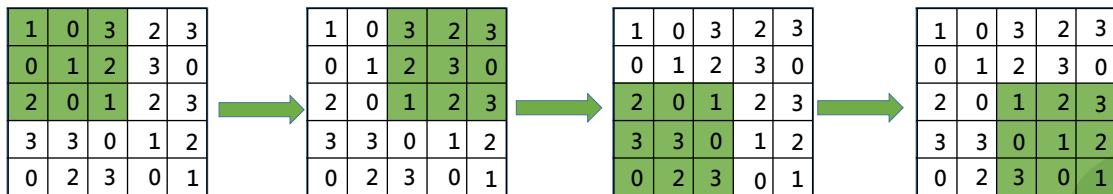
面试问题1: Valid卷积运算, 为保持输出大小=输入大小, 填充宽度如何计算?

面试问题2: 一些常用的卷积 (如3*3), $pad=?$

03. 卷积层 - 卷积运算步长

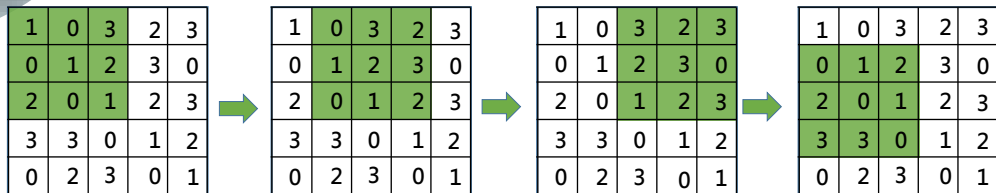


卷积输入4*4矩阵（图像）经过3*3卷积核，平移步长为1



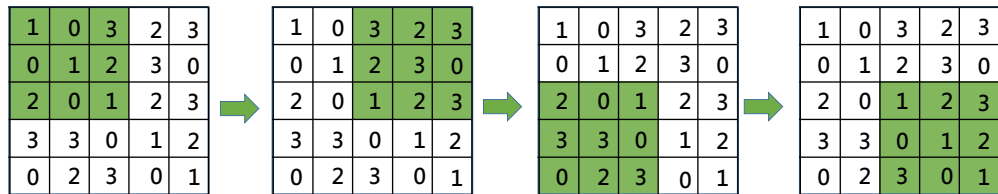
卷积输入5*5矩阵（图像），经过3*3卷积核，平移步长为2

03. 卷积层 – 卷积运算输出公式



输出3*3
步长s=1
 $m=n-k+1$

卷积输入5*5矩阵（图像），经过3*3卷积核，平移步长为1



输出2*2
步长s=2
 $m=(n-k)/s+1$

卷积输入5*5矩阵（图像），经过3*3卷积核，平移步长为2

03. 卷积层 – 卷积运算输出公式

同时考虑填充和步长的输出大小计算

输入大小 n ，卷积核 k ，填充 $\text{pad}=p$ ，步长 $\text{step}=s$ ，输出大小 m ，

$$m = (n+2p-k)/s+1,$$

例如，输入 $7*7$ ，卷积核 $3*3$ ，填充 $p=1$ ，步长 $s=2$

输出 $m=(7+2-3)/2+1 = 4$

面试问题：给定输入大小，卷积核大小，填充，步长，计算输出大小

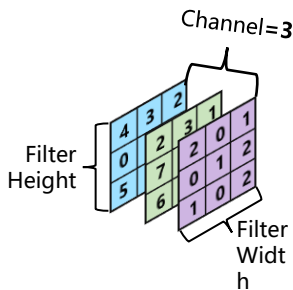
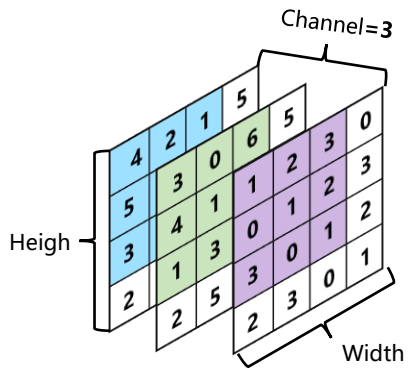
注意：如果老师没有讲填充和步长大小，可以主动提问，表示自己知道填充和步长

03. 卷积层 – 多通道卷积运算

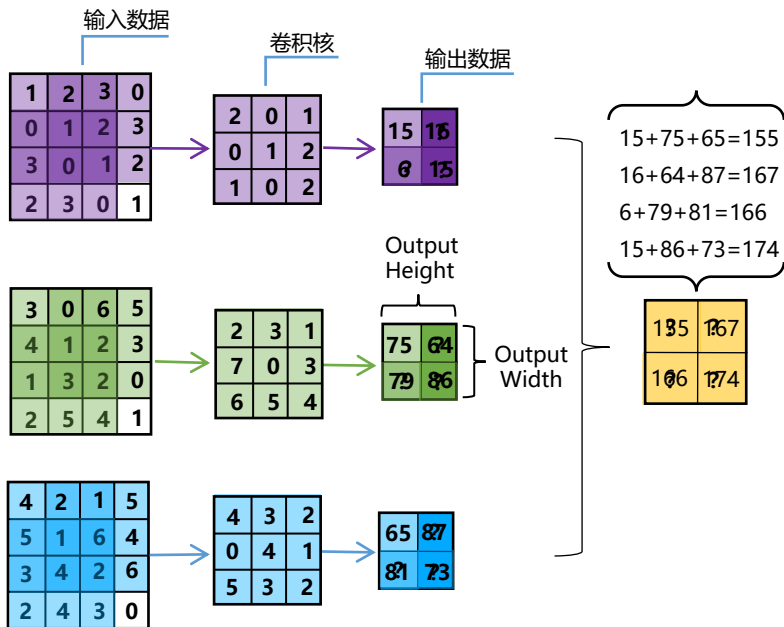
输入特征图维度: (N, Channel=3, Heigh=4, Width=4)

- 步长: 1
- 卷积核尺寸(长宽): 3x3
- 卷积核通道数: 3 (必须和输入通道数相同)
- 卷积核个数: 1

输出特征图维度: (N, Channel=1, Heigh=2, Width=2)



各通道分别进行卷积,
再把卷积的结果相加



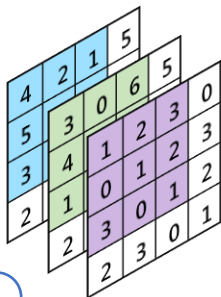
03. 卷积层 – 多通道卷积运算

输入特征图维度: (N, $C_{in}=3$, Heigh=4, Width=4)

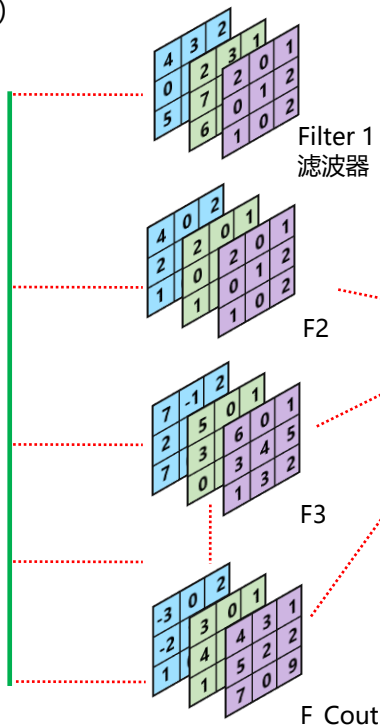
- 步长: 1
- 卷积核尺寸(长宽): 3x3
- 卷积核通道数: 3
- 卷积核个数: **Cout**

输入特征图维度: (N, **Cout**, Heigh=2, Width=2)

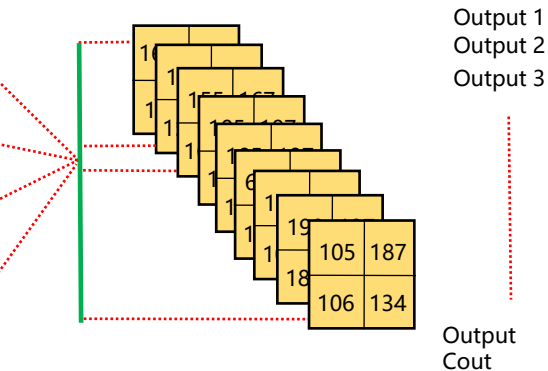
面试问题: 多通道卷积 (多个卷积核) 的作用是什么?



Input data
输入数据



一个卷积核提取一个特征,
多个卷积核产生多个特征



每个特征对应于一个通道,
卷积: 多通道输入 -> 多通道输出

03. 卷积层 – 卷积运算输出公式

多通道卷积计算

问：输入维度(N, C_{in} , H, W)，卷积核(C_{in} , C_{out} , k, k)，填充
 $pad=p$ ，步长 $step=s$ 。输出维度如何计算？

答：输出维度为(N, C_{out} , H' , W')，其中 C_{out} 为卷积核个数，

$$H' = (H+2p-k)/s+1$$

$$W' = (W+2p-k)/s+1$$

卷积核权重个数为 $C_{in} * C_{out} * k * k$

面试问题：给定输入维度，卷积核大小，填充，步长，计算输出维度

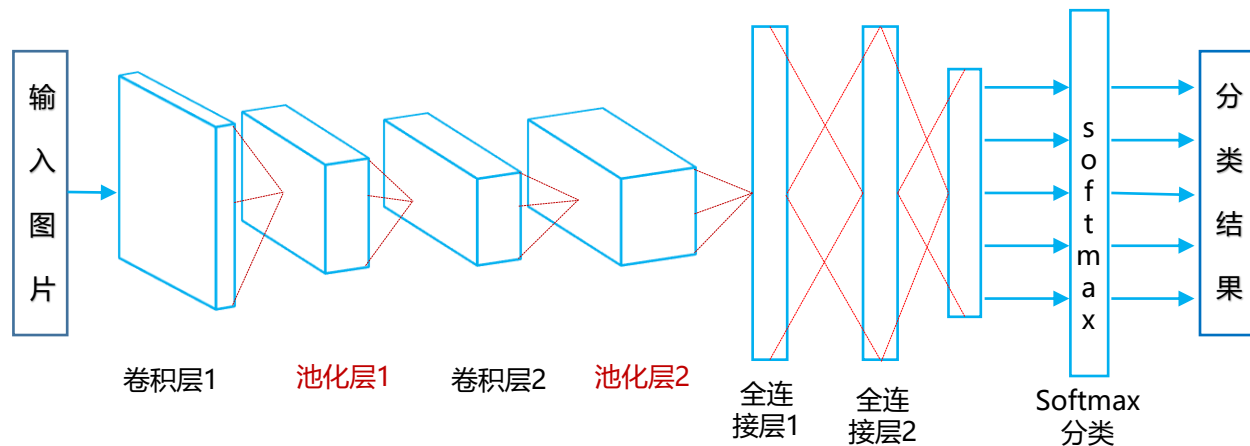
面试问题：如何计算卷积层的权重个数？

注意：如果老师只问了单通道卷积，可以主动提自己知道多通道卷积的计算

04. 池化层运算详解

04. 池化层(Pooling) - 卷积神经网络结构回顾

九曲阑干



用于图像分类问题的一种卷积神经网络架构图

04. 池化层(Pooling) - 池化分类作用

九曲阑干

池化操作分类

最大池化

均值池化

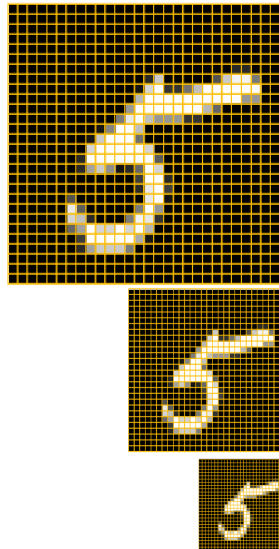
最大池化：选取值最大的特征

平均池化：计算特征的平均值

池化层的作用：

- ①减少特征维度（降维，可以减少计算复杂度，防止过拟合）
- ②保留主要特征，使特征具有不变性（平移、尺度不变性）

面试问题：池化层的作用是什么？



04. 池化层(pooling) - 最大池化操作

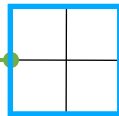
九曲阑干

Max Pooling : 获取**最大值**进行运算

池化核步长(Stride) : 2

池化核尺寸 : 2x2

Kernel Size : 2X2池化核尺寸2x2



面试问题：最大池化操作如何计算？

输出数据

输入数据

2	

最大值

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

2	3

最大值

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

2	3
4	

最大值

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

2	3
4	2

最大值

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

04. 池化层(pooling) - 均值池化操作

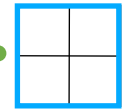
Ave Pooling : 获取平均值进行运算

池化核步长(Stride) : 2

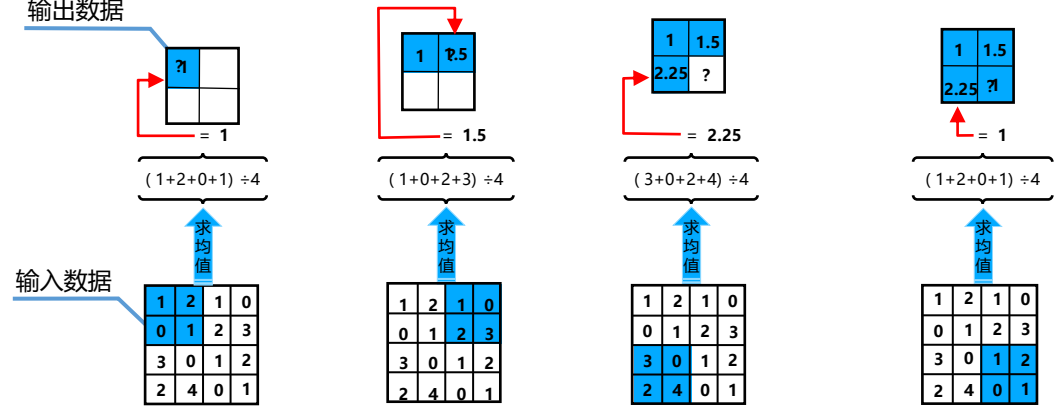
池化核尺寸 : 2x2

Kernel Size : 2X2池化核尺寸2x2

输出数据



面试问题：平均池化操作如何计算？



04. 池化层(Pooling) - 池化分类作用

九曲阑干

池化操作分类

最大池化

均值池化

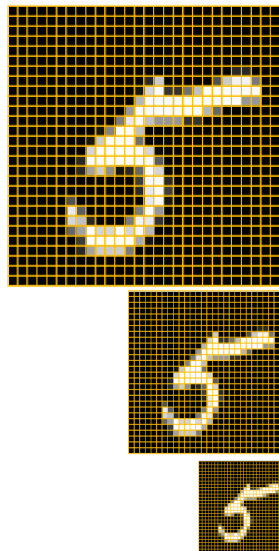
最大池化：选取值最大的特征

平均池化：计算特征的平均值

池化层的作用：

- ①减少特征维度（降维，可以减少计算复杂度，防止过拟合）
- ②保留主要特征，使特征具有不变性（平移、尺度不变性）

面试问题：池化层的作用是什么？



04. 池化层(Pooling)输出公式

池化计算

问：输入维度 $n \times n$ ，池化步长 $step=s$ 。输出维度如何计算？

答：输出维度为 $m \times m$, $m=n/s$

如果输入是三维特征图，池化后通道数不变

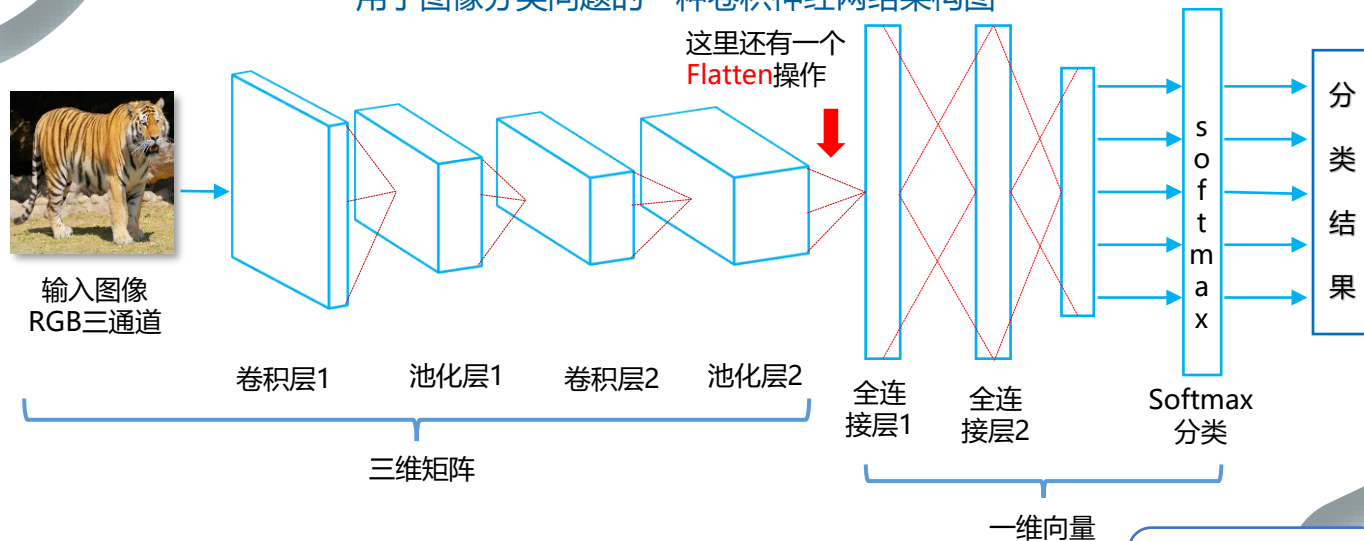
池化层没有参数

池化层后面没有激活函数

面试问题：给定输入维度，
池化步长，计算输出维度

02. 卷积神经网络结构组成

用于图像分类问题的一种卷积神经网络架构图

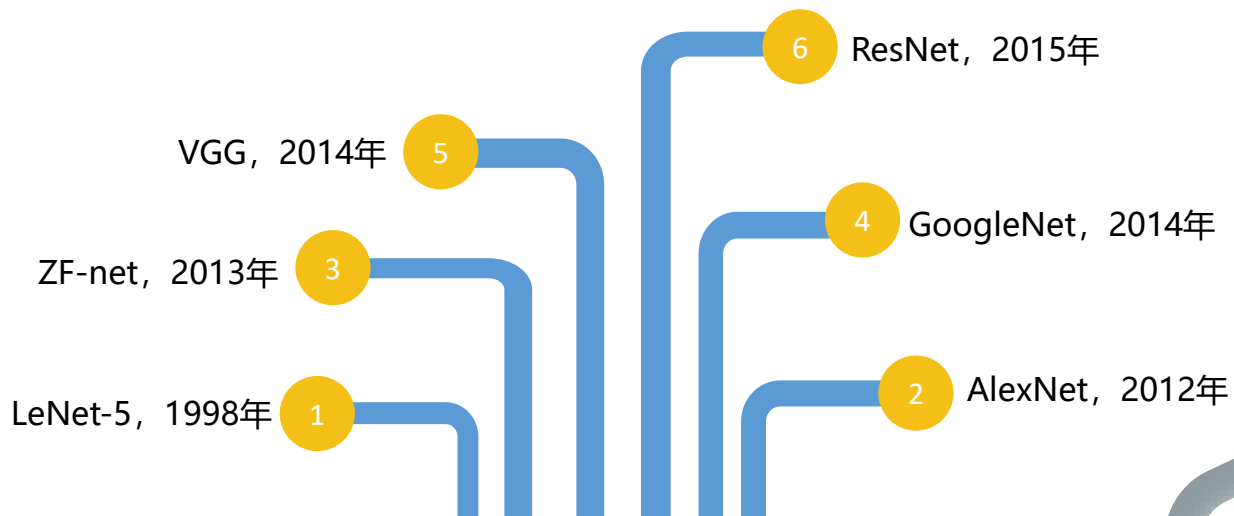


CNN的组成：输入层、卷积层、池化层、全连接层、softmax分类
其中卷积层和全连接层中包括激活函数（如ReLU）

面试问题：卷积神经网络一般由哪些层组成？

06. 经典卷积神经网络

九曲阑干

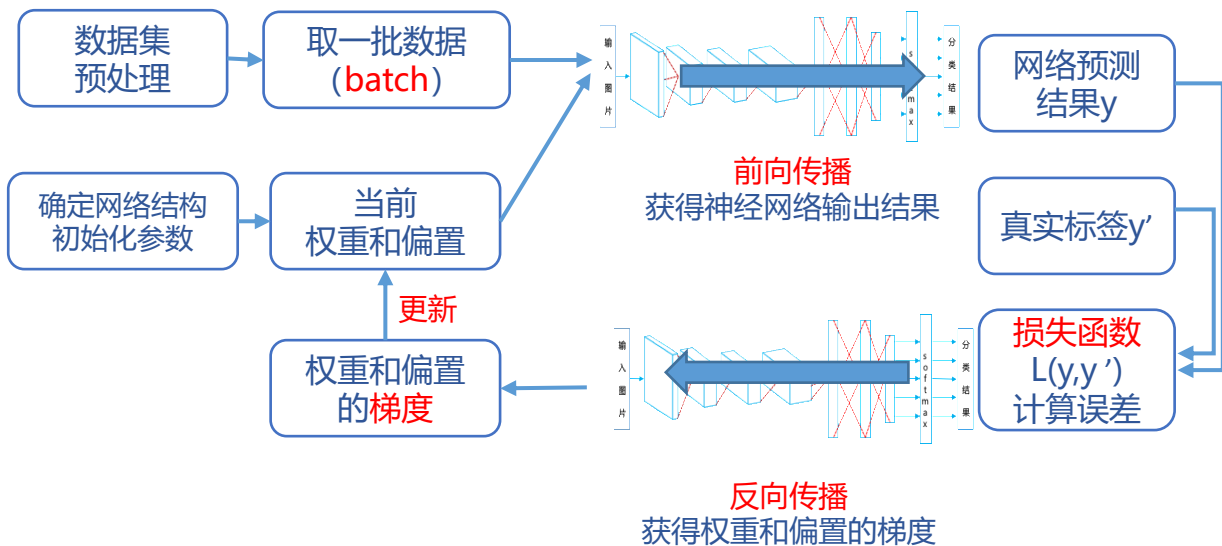


常见经典卷积神经网络主要包括

卷积神经网络训练总览



卷积神经网络的训练流程和全连接网络相同（需要计算卷积层和池化层的梯度）
使用**随机梯度下降法**进行训练，**迭代更新**神经网络的参数



九曲阑干

感谢聆听

Thank You