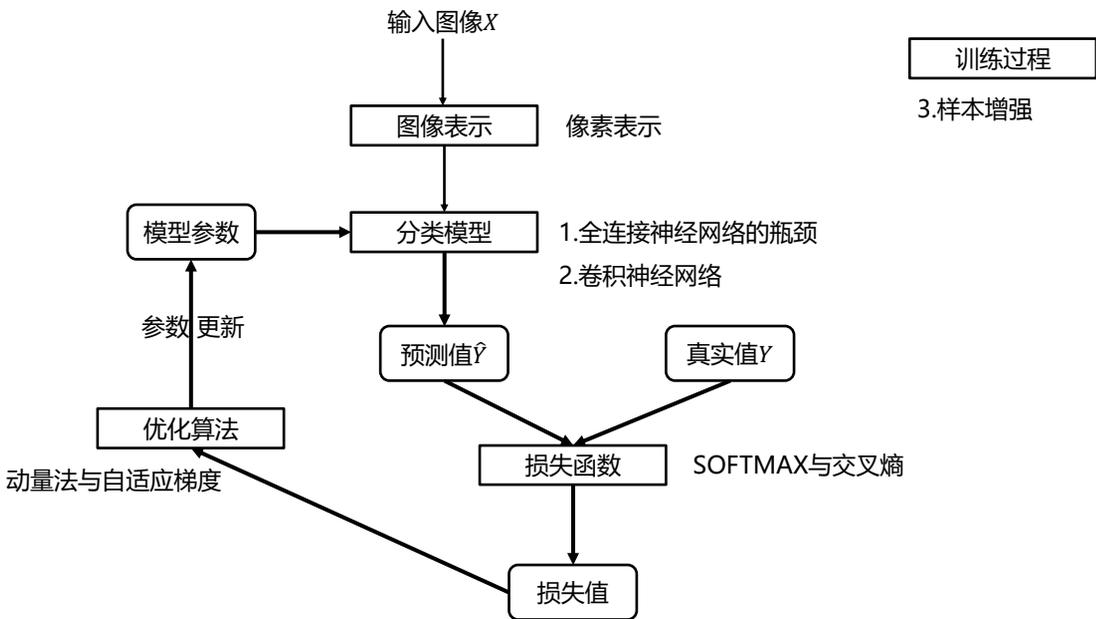
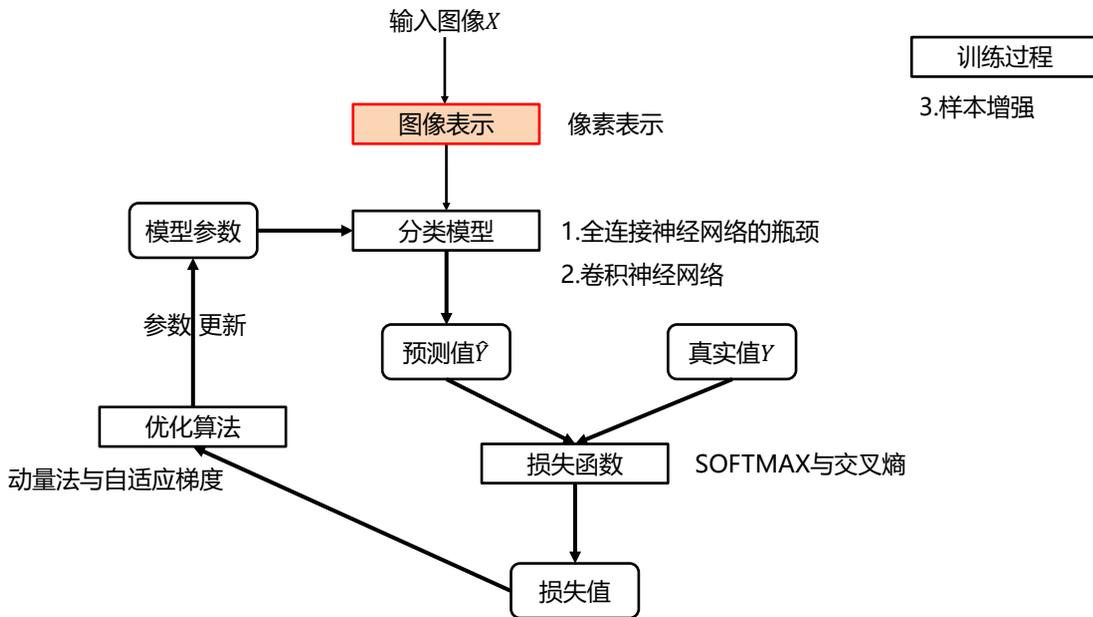


卷积神经网络

鲁鹏

北京邮电大学 计算机学院 智能科学与技术中心





2020/4/21

北京邮电大学计算机学院 鲁鹏

2

如何表示图像？

直接利用原始像素作为特征，展开为列向量。

$$\begin{bmatrix} 76 & 32 \\ 4 & 157 \end{bmatrix} \xrightarrow{\text{将矩阵转成列向量}} x = \begin{bmatrix} 76 \\ 32 \\ 4 \\ 157 \end{bmatrix}$$

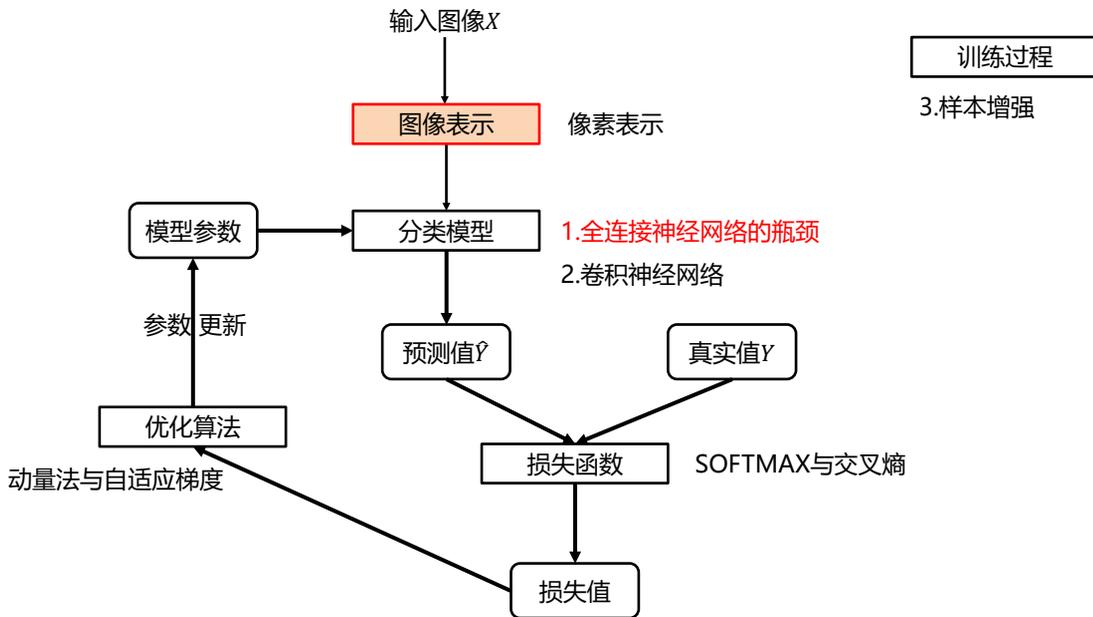
图像

cifar10中每个图像可表示为一个3072 (32*32*3) 维的向量

2020/4/21

北京邮电大学计算机学院 鲁鹏

3



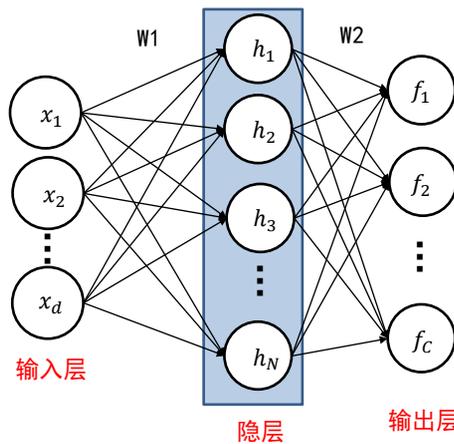
2020/4/21

北京邮电大学计算机学院 鲁鹏

4

全连接神经网络的瓶颈

两层全连接网络



问题: CIFAR10图像尺寸 $32 \times 32 \times 3$, 隐层每个神经元权值个数是多少?

回答: $32 \times 32 \times 3 = 3072$

问题: 如果图像大小 $200 \times 200 \times 3$, 隐层每个神经元权值个数是多少?

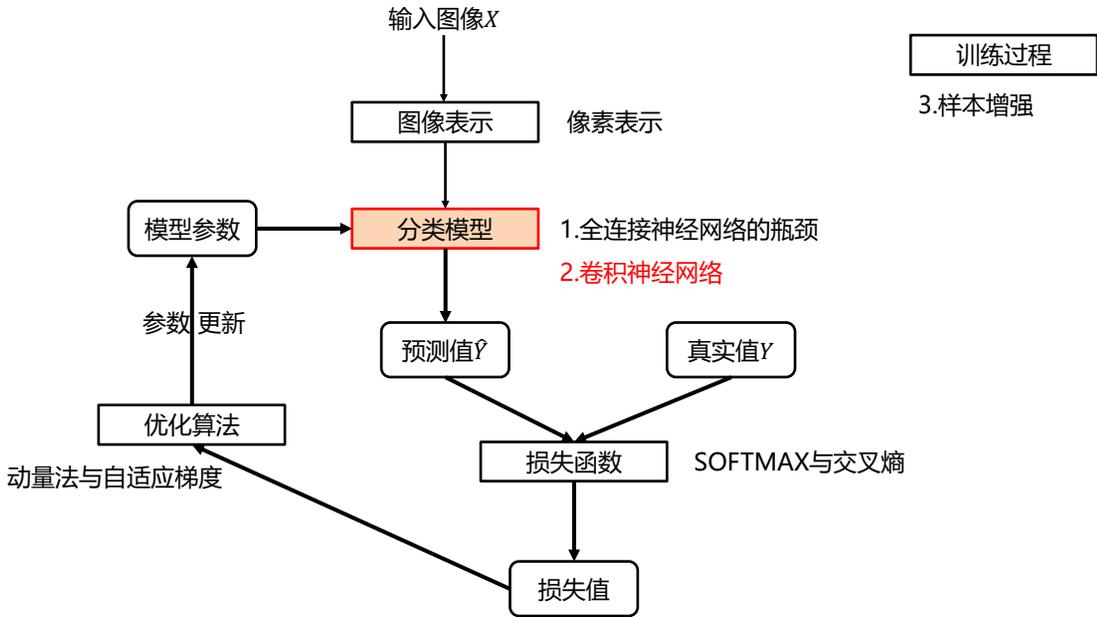
回答: $200 \times 200 \times 3 = 120,000$ 个

仅适合处理小图像!

2020/4/21

北京邮电大学计算机学院 鲁鹏

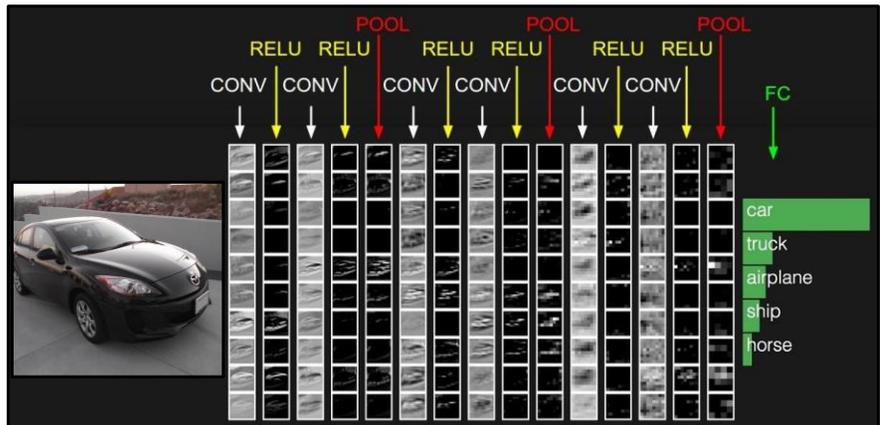
5



训练过程
3. 样本增强

卷积神经网络

- CONV——卷积层
- RELU——激活层
- POOL——池化层
- FC——全连接层



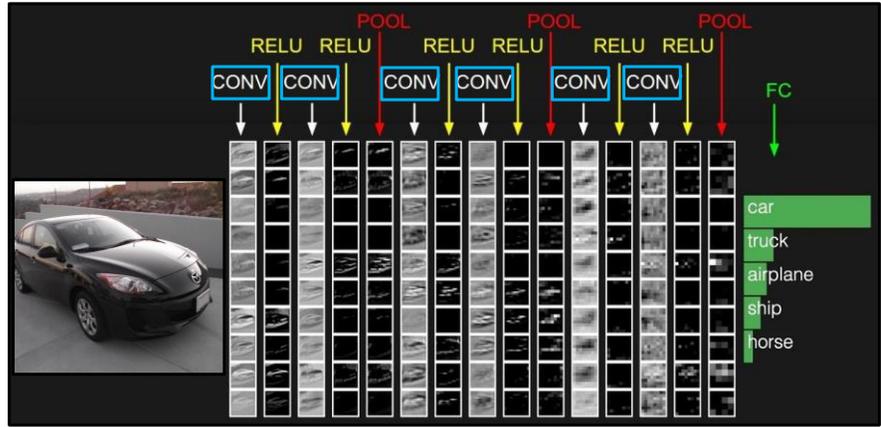
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



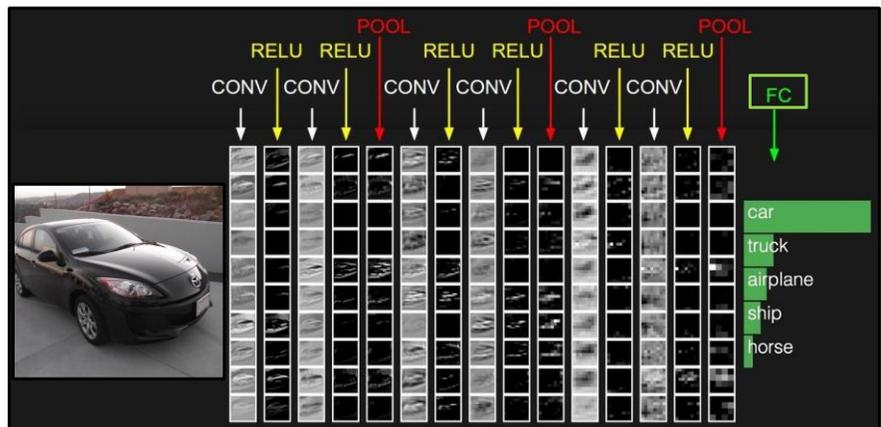
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



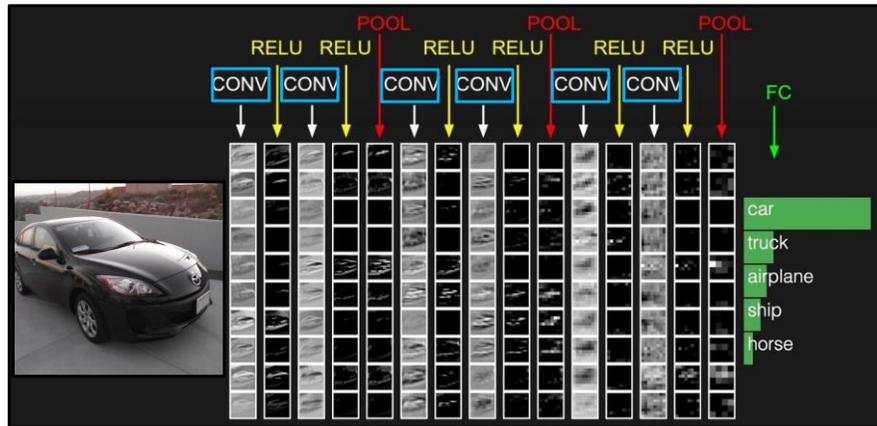
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



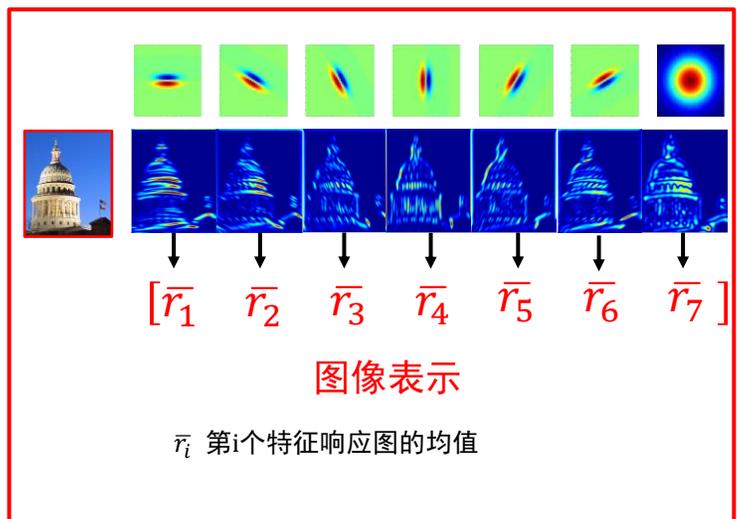
2020/4/21

北京邮电大学计算机学院 鲁鹏

10

基于卷积核组的图像表示

1. 设计卷积核组；
2. 利用卷积核组对图像进行卷积操作获得对应的特征响应图组；
3. 利用特征响应图的某种统计信息来表示图像中的纹理。

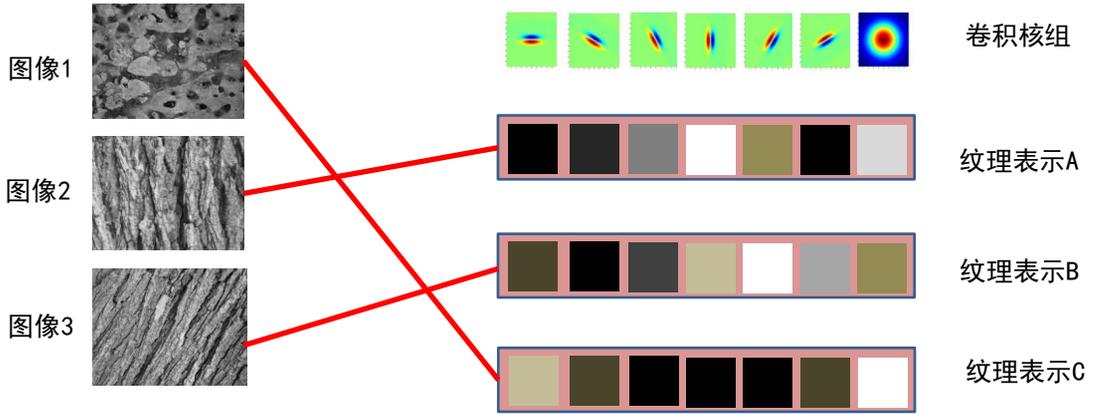


2020/4/21

北京邮电大学计算机学院 鲁鹏

11

纹理表示小游戏



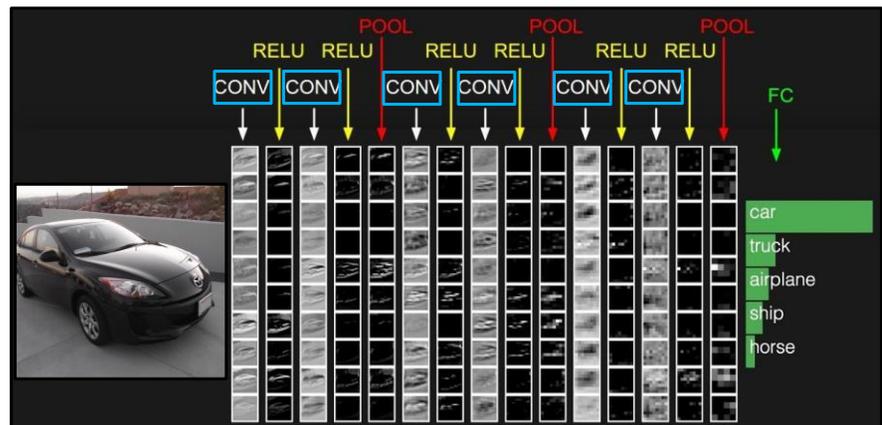
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



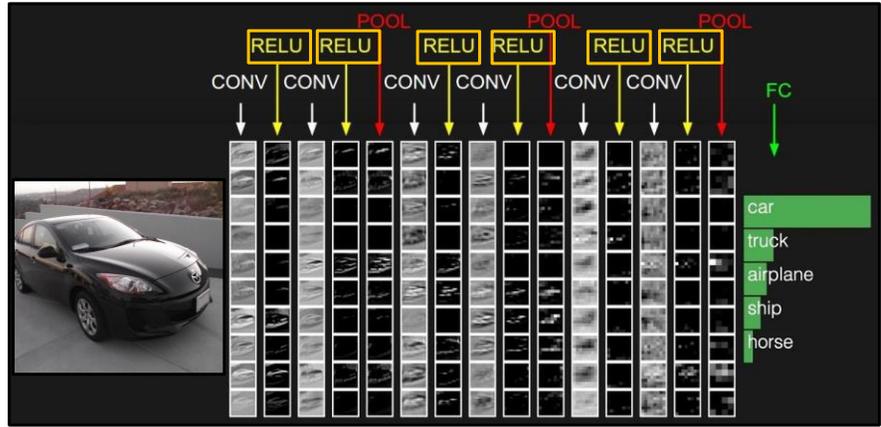
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



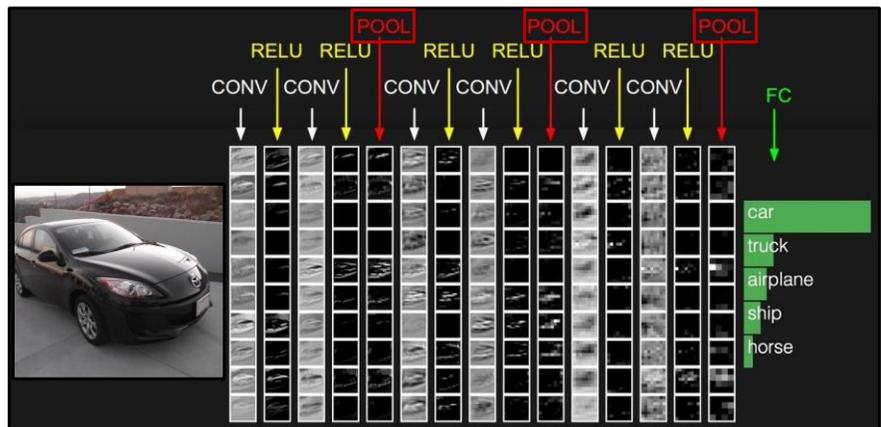
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



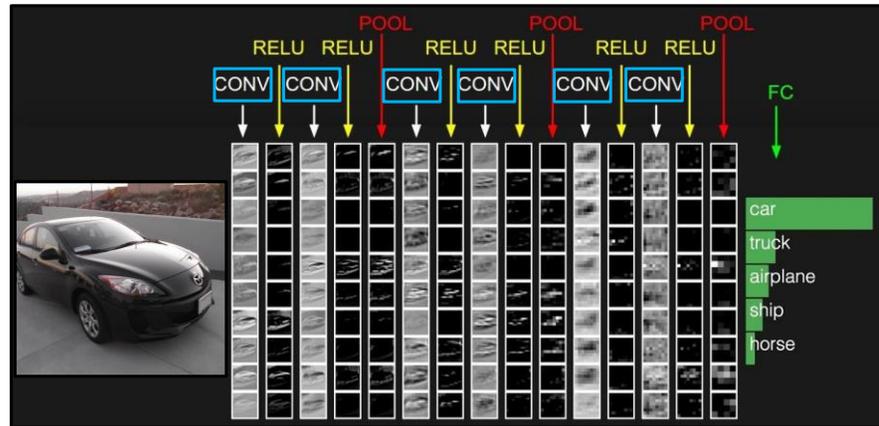
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



卷积网络中的卷积核

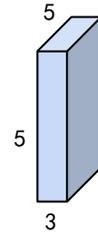
卷积核：

- 不仅具有宽和高，还具有深度，常写成如下形式：
宽度 × 高度 × 深度
- 卷积核参数不仅包括核中存储的权值，还包括一个偏置值

卷积网络中的卷积核

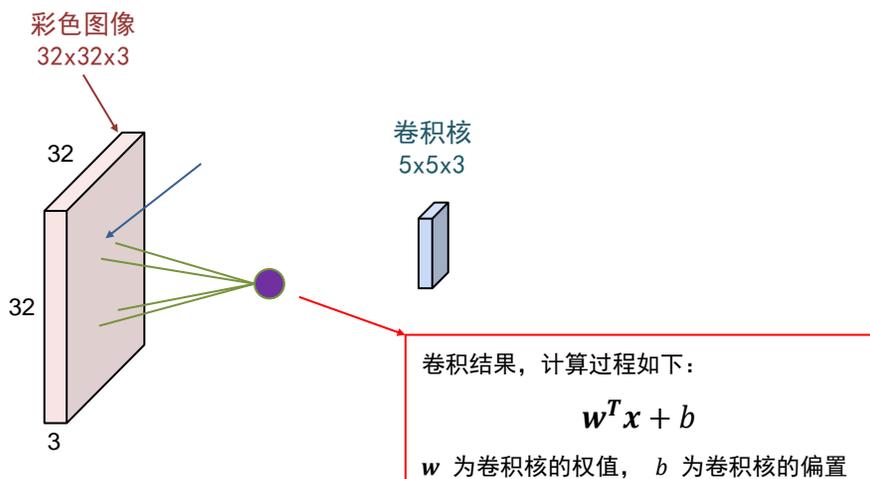
卷积核：

- 不仅具有宽和高，还具有深度，常写成如下形式：
宽度 × 高度 × 深度
- 卷积核参数不仅包括核中存储的权值，还包括一个偏置值

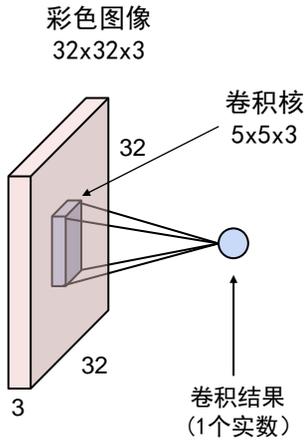


5×5×3的卷积核

卷积网络中的卷积操作



卷积网络中的卷积操作



计算过程:

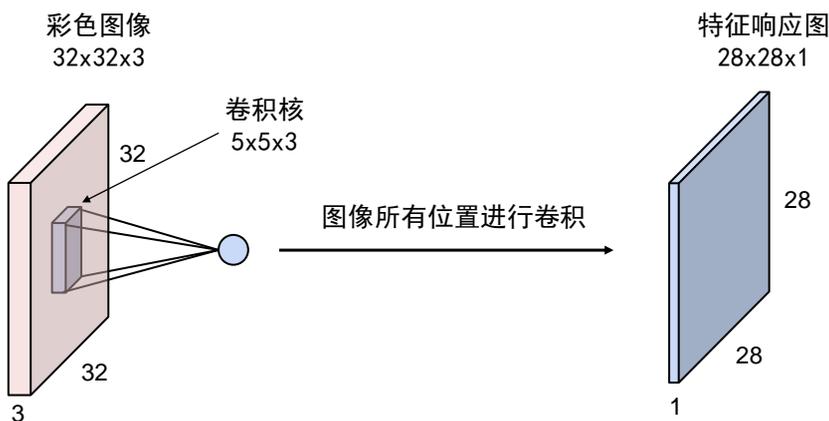
- 将卷积核展成一个5x5x3的向量，同时将其覆盖的图像区域按相同的展开方式展成5x5x3的向量
- 计算两者的点乘。
- 在点乘的结果上加上偏移量

数学公式:

$$w^T x + b$$

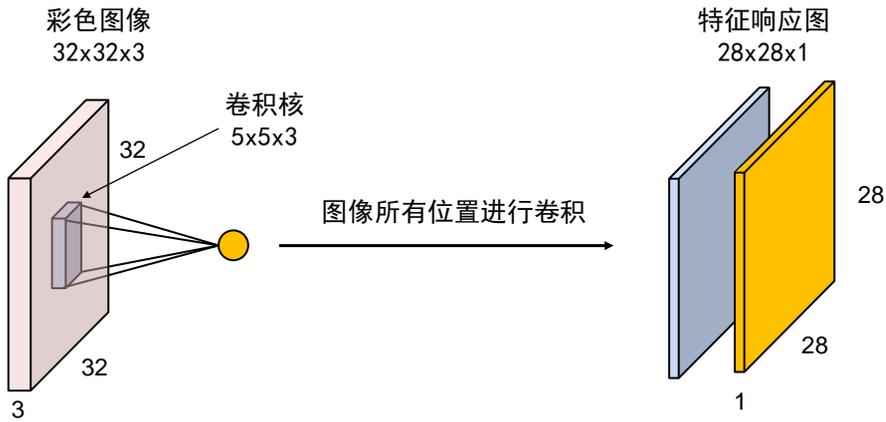
w 为卷积核的权值， b 为卷积核的偏置

卷积网络中的卷积操作



特征响应图中每个位置上的值反映了图像上对应位置是否存在卷积核所记录的基元结构信息。

卷积层

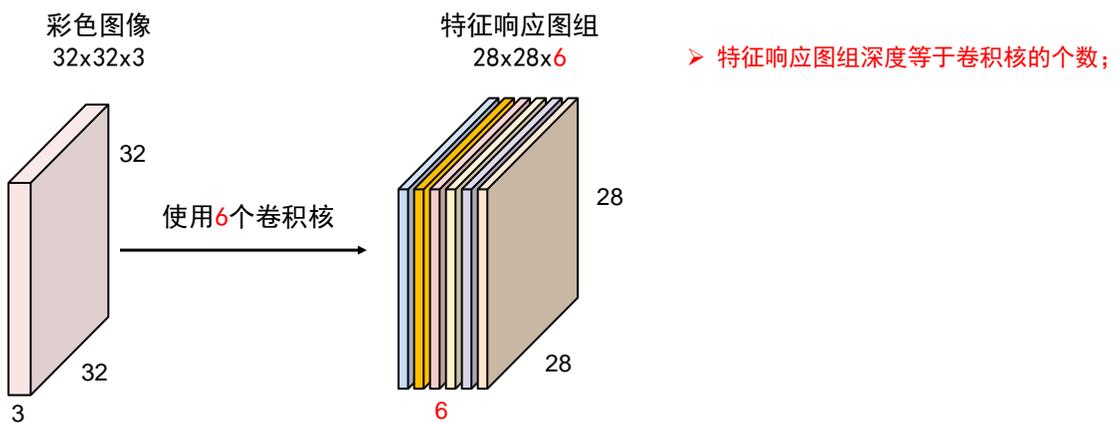


2020/4/21

北京邮电大学计算机学院 鲁鹏

22

卷积层

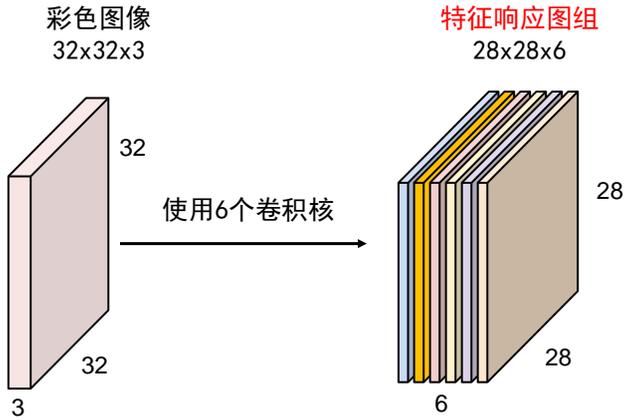


2020/4/21

北京邮电大学计算机学院 鲁鹏

23

卷积层



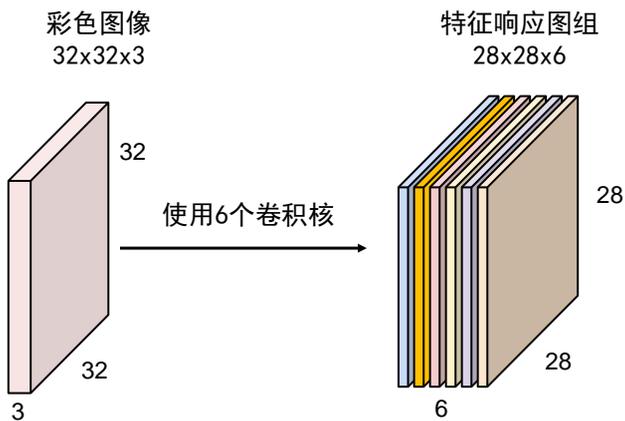
- 特征响应图组深度等于卷积核的个数；

2020/4/21

北京邮电大学计算机学院 鲁鹏

24

卷积层



- 特征响应图组深度等于卷积核的个数；

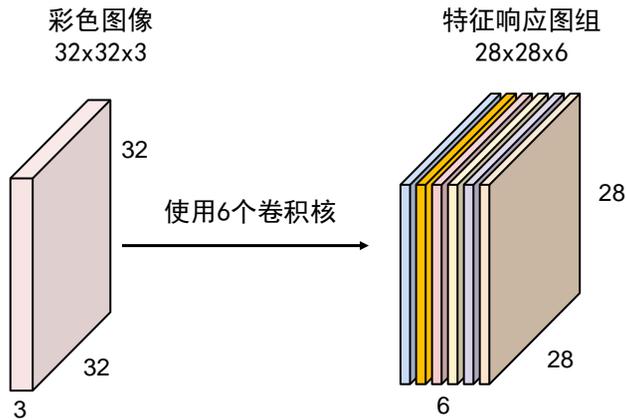
- 不同的特征响应图反映了输入图像对
不同卷积核的响应结果；

2020/4/21

北京邮电大学计算机学院 鲁鹏

25

卷积层



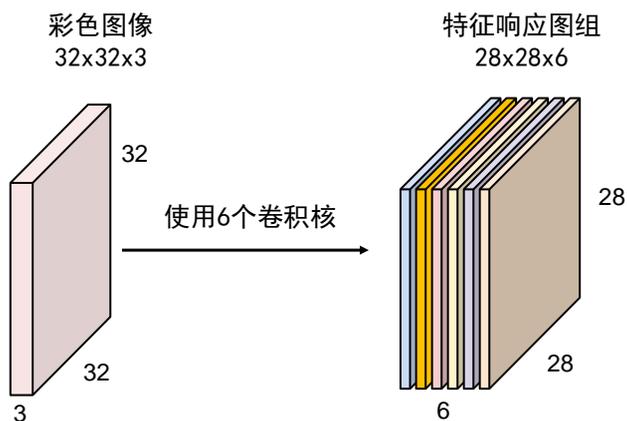
- 特征响应图组深度等于卷积核的个数；
- 不同的特征响应图反映了输入图像对不同卷积核的响应结果；
- 同一特征响应图上不同位置的值得表示输入图像上不同位置对同一卷积核的响应结果。

2020/4/21

北京邮电大学计算机学院 鲁鹏

26

卷积层



- 特征响应图组深度等于卷积核的个数；
- 不同的特征响应图反映了输入图像对不同卷积核的响应结果；
- 同一特征响应图上不同位置的值得表示输入图像上不同位置对同一卷积核的响应结果。

注意：

卷积层输入不局限于图像，可以是任意三维数据矩阵；

该层的卷积核深度要求与输入的三维矩阵的深度一致。

2020/4/21

北京邮电大学计算机学院 鲁鹏

27

卷积步长 (stride)

2020/4/21

北京邮电大学计算机学院 鲁鹏

28

卷积步长 (stride)

卷积神经网络中，卷积核可以按照指定的间隔进行卷积操作，这个间隔就是**卷积步长**。

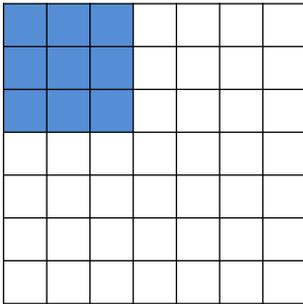
2020/4/21

北京邮电大学计算机学院 鲁鹏

29

卷积步长 (stride)

卷积神经网络中，卷积核可以按照指定的间隔进行卷积操作，这个间隔就是**卷积步长**。



图像尺寸： 7x7

卷积核尺寸： 3x3

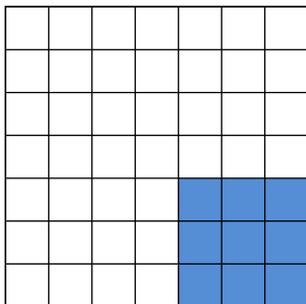
2020/4/21

北京邮电大学计算机学院 鲁鹏

30

卷积步长 (stride)

卷积神经网络中，卷积核可以按照指定的间隔进行卷积操作，这个间隔就是**卷积步长**。



图像尺寸： 7x7

卷积核尺寸： 3x3

卷积步长： 1

卷积结果： 5x5

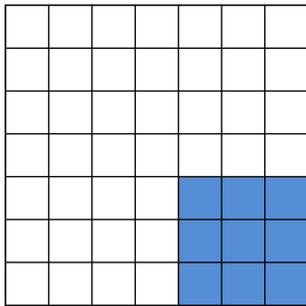
2020/4/21

北京邮电大学计算机学院 鲁鹏

31

卷积步长 (stride)

卷积神经网络中，卷积核可以按照指定的间隔进行卷积操作，这个间隔就是**卷积步长**。



卷积结果：3x3

图像尺寸： 7x7

卷积核尺寸： 3x3

卷积步长： 2

输入数据矩阵尺寸： $W1 \times H1$

输出特征图组尺寸： $W2 \times H2$

W2与W1关系如下：

$$W2 = (W1 - F) / S + 1$$

$$H2 = (H2 - F) / S + 1$$

F——卷积核尺寸

S——卷积步长

边界填充

无像素



边界填充

卷积神经网络中最常用的填充方式是**零值填充**。

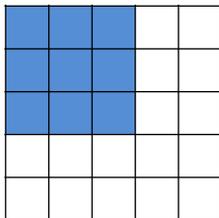
2020/4/21

北京邮电大学计算机学院 鲁鹏

34

边界填充

卷积神经网络中最常用的填充方式是**零值填充**。



图像尺寸： 5x5

卷积核尺寸： 3x3

卷积步长： 1

零值填充： 0

卷积结果：

2020/4/21

北京邮电大学计算机学院 鲁鹏

35

边界填充

卷积神经网络中最常用的填充方式是**零值填充**。

图像尺寸： 5x5

卷积核尺寸： 3x3

卷积步长： 1

零值填充： 0

卷积结果： 3x3

边界填充

卷积神经网络中最常用的填充方式是**零值填充**。

0	0	0	0	0	0	0
0						0
0						0
0						0
0						0
0						0
0	0	0	0	0	0	0

图像尺寸： 5x5

卷积核尺寸： 3x3

卷积步长： 1

零值填充： 1

卷积结果：

边界填充

卷积神经网络中最常用的填充方式是**零值填充**。

0	0	0	0	0	0	0
0						0
0						0
0						0
0						0
0						0
0	0	0	0	0	0	0

图像尺寸： 5x5

卷积核尺寸： 3x3

卷积步长： 1

零值填充： 1

卷积结果： 5x5

2020/4/21

北京邮电大学计算机学院 鲁鹏

38

边界填充

卷积神经网络中最常用的填充方式是**零值填充**。

0	0	0	0	0	0	0
0						0
0						0
0						0
0						0
0						0
0	0	0	0	0	0	0

图像尺寸： 5x5

卷积核尺寸： 3x3

卷积步长： 1

零值填充： 1

卷积结果： 5x5

F——卷积核尺寸

S——卷积步长

P——零填充数量

输入数据矩阵尺寸： $W1 \times H1$

输出特征图组尺寸： $W2 \times H2$

W2与W1关系如下：

$$W2 = (W1 - F + 2P) / S + 1$$

$$H2 = (H2 - F + 2P) / S + 1$$

作用：保持输入、输出尺寸的一致！

2020/4/21

北京邮电大学计算机学院 鲁鹏

39

特征响应图组尺寸计算

2020/4/21

北京邮电大学计算机学院 鲁鹏

40

特征响应图组尺寸计算

给定输入数据矩阵时，影响输出的特征图组尺寸大小的因素：

- 卷积核的宽、高；

2020/4/21

北京邮电大学计算机学院 鲁鹏

41

特征响应图组尺寸计算

给定输入数据矩阵时，影响输出的特征图

组尺寸大小的因素：

- 卷积核的宽、高；
- 是否采用边界填充操作；

特征响应图组尺寸计算

给定输入数据矩阵时，影响输出的特征图

组尺寸大小的因素：

- 卷积核的宽、高；
- 是否采用边界填充操作；
- 卷积步长；

特征响应图组尺寸计算

给定输入数据矩阵时，影响输出的特征图

组尺寸大小的因素：

- 卷积核的宽、高；
- 是否采用边界填充操作；
- 卷积步长；
- 该层的卷积核个数。

特征响应图组尺寸计算

给定输入数据矩阵时，影响输出的特征图

组尺寸大小的因素：

- 卷积核的宽、高；
- 是否采用边界填充操作；
- 卷积步长；
- 该层的卷积核个数。

输入数据矩阵尺寸： $W1 \times H1 \times D1$

输出特征图组尺寸： $W2 \times H2 \times D2$

特征响应图组尺寸计算

给定输入数据矩阵时，影响输出的特征图组尺寸大小的因素：

- 卷积核的宽、高；
- 是否采用边界填充操作；
- 卷积步长；
- 该层的卷积核个数。

输入数据矩阵尺寸： $W1 \times H1 \times D1$

输出特征图组尺寸： $W2 \times H2 \times D2$

F——卷积核尺寸
S——卷积步长
P——零填充数量
K——卷积核个数

卷积层的4个超参数

特征响应图组尺寸计算

给定输入数据矩阵时，影响输出的特征图组尺寸大小的因素：

- 卷积核的宽、高；
- 是否采用边界填充操作；
- 卷积步长；
- 该层的卷积核个数。

输入数据矩阵尺寸： $W1 \times H1 \times D1$

输出特征图组尺寸： $W2 \times H2 \times D2$

特征图组尺寸计算如下：

$$W2 = (W1 - F + 2P) / S + 1$$

$$H2 = (H1 - F + 2P) / S + 1$$

$$D2 = K$$

F——卷积核尺寸
S——卷积步长
P——零填充数量
K——卷积核个数

卷积层的4个超参数

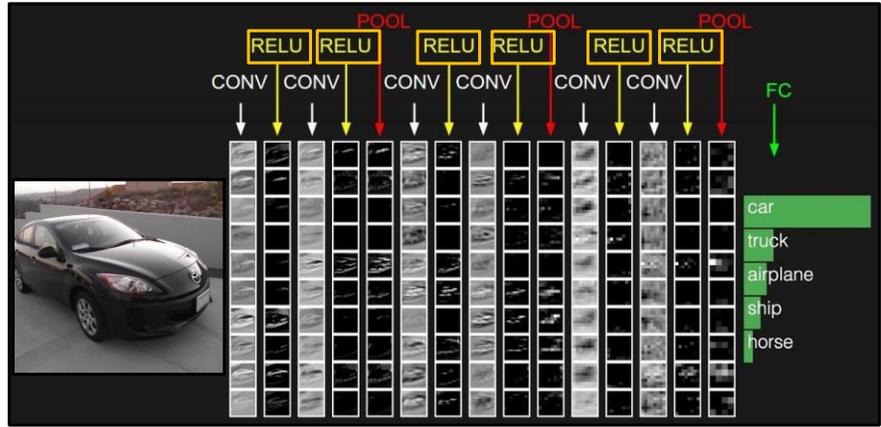
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



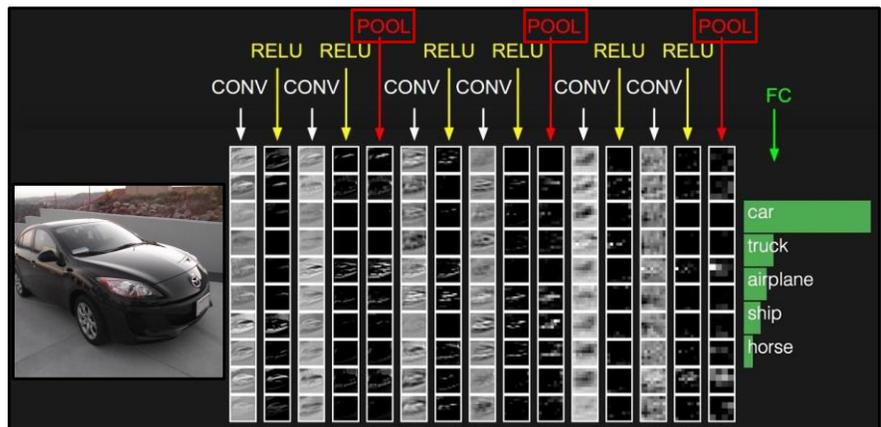
卷积神经网络

CONV——卷积层

RELU——激活层

POOL——池化层

FC——全连接层



池化操作

- 池化的作用：对每一个特征响应图独立进行，降低特征响应图组中每个特征响应图的宽度和高度，减少后续卷积层的参数的数量，降低计算资源耗费，进而控制过拟合。

池化操作

- 池化的作用：对每一个特征响应图独立进行，降低特征响应图组中每个特征响应图的宽度和高度，减少后续卷积层的参数的数量，降低计算资源耗费，进而控制过拟合。
- 池化操作：对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。

池化操作

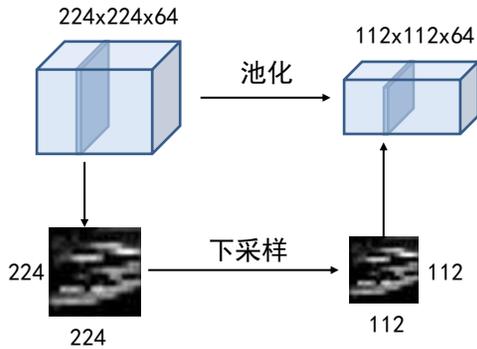
- 池化的作用：对每一个特征响应图独立进行，降低特征响应图组中每个特征响应图的宽度和高度，减少后续卷积层的参数的数量，降低计算资源耗费，进而控制过拟合。
- 池化操作：对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。
- 常见的池化操作：
 - ❑ 最大池化——使用区域内的最大值来代表这个区域；
 - ❑ 平均池化——采用区域内所有值的均值作为代表。

池化操作

- 池化的作用：对每一个特征响应图独立进行，降低特征响应图组中每个特征响应图的宽度和高度，减少后续卷积层的参数的数量，降低计算资源耗费，进而控制过拟合。
- 池化操作：对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。
- 常见的池化操作：
 - ❑ 最大池化——使用区域内的最大值来代表这个区域；
 - ❑ 平均池化——采用区域内所有值的均值作为代表。
- 池化层的超参数：池化窗口和池化步长

池化操作示例

- 池化操作对每一个特征响应图独立进行；



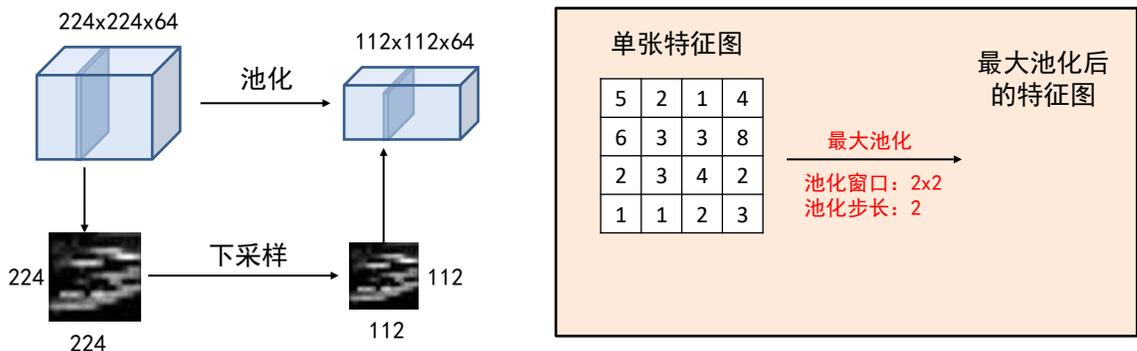
2020/4/21

北京邮电大学计算机学院 鲁鹏

54

池化操作示例

- 池化操作对每一个特征响应图独立进行；
- 对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。



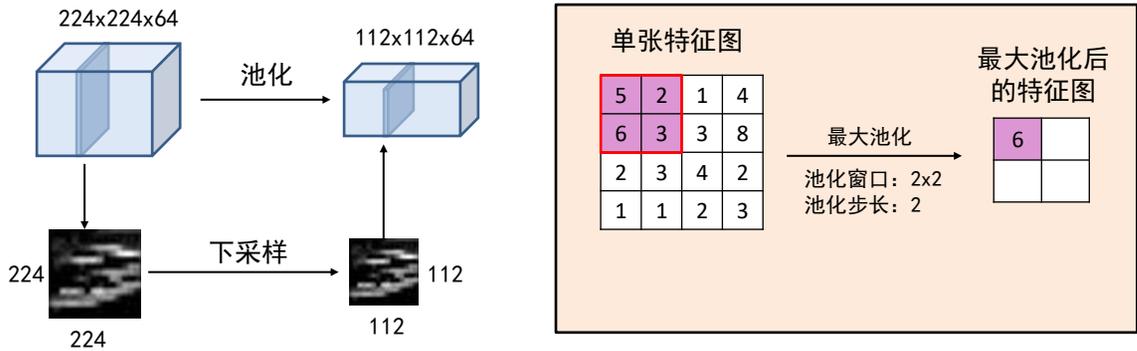
2020/4/21

北京邮电大学计算机学院 鲁鹏

55

池化操作示例

- 池化操作对每一个特征响应图独立进行；
- 对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。



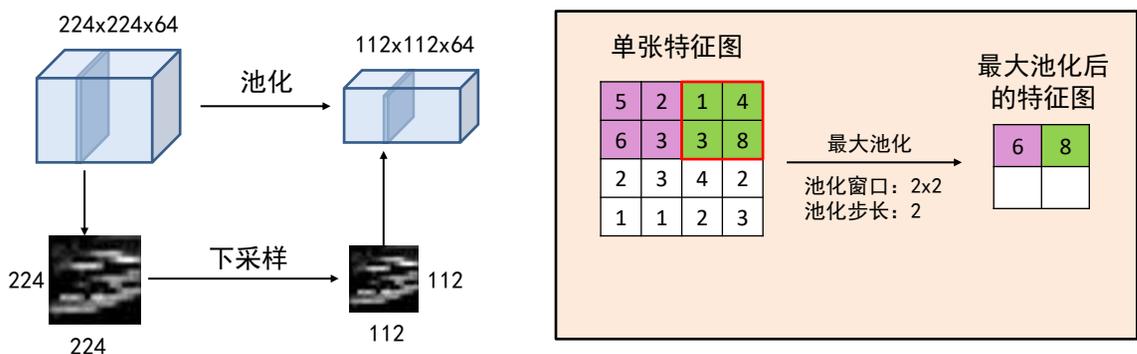
2020/4/21

北京邮电大学计算机学院 鲁鹏

56

池化操作示例

- 池化操作对每一个特征响应图独立进行；
- 对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。



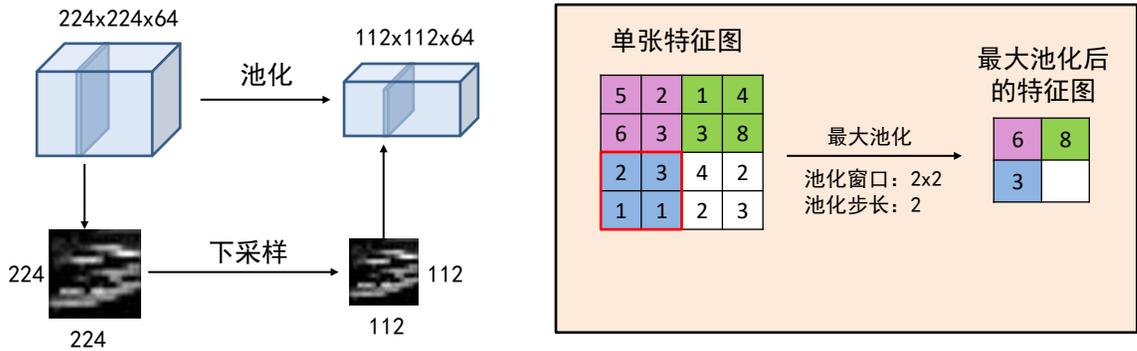
2020/4/21

北京邮电大学计算机学院 鲁鹏

57

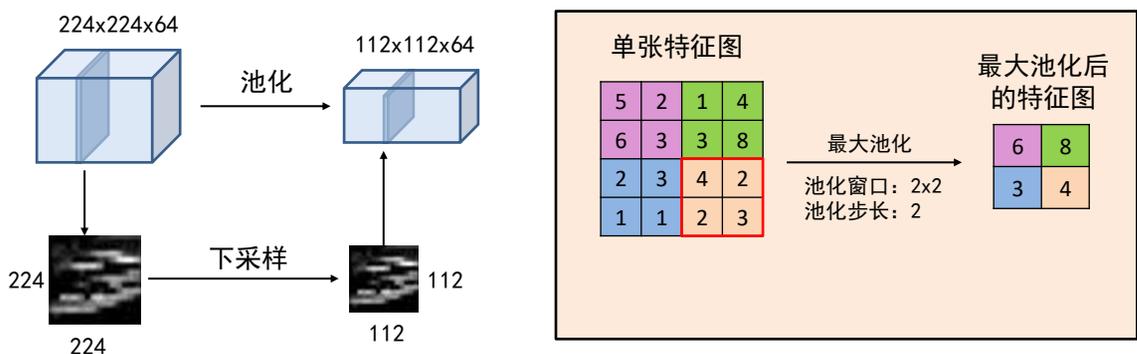
池化操作示例

- 池化操作对每一个特征响应图独立进行；
- 对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。



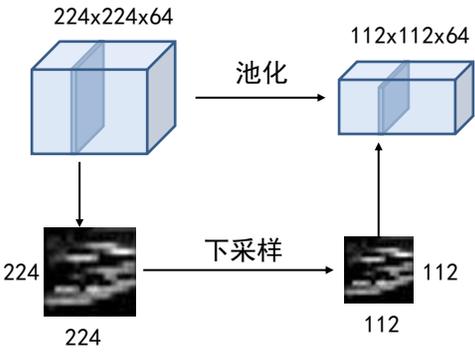
池化操作示例

- 池化操作对每一个特征响应图独立进行；
- 对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。



池化操作示例

- 池化操作对每一个特征响应图独立进行；
- 对特征响应图某个区域进行池化就是在该区域上指定一个值来代表整个区域。



单张特征图

5	2	1	4
6	3	3	8
2	3	4	2
1	1	2	3

最大池化

池化窗口: 2x2
池化步长: 2

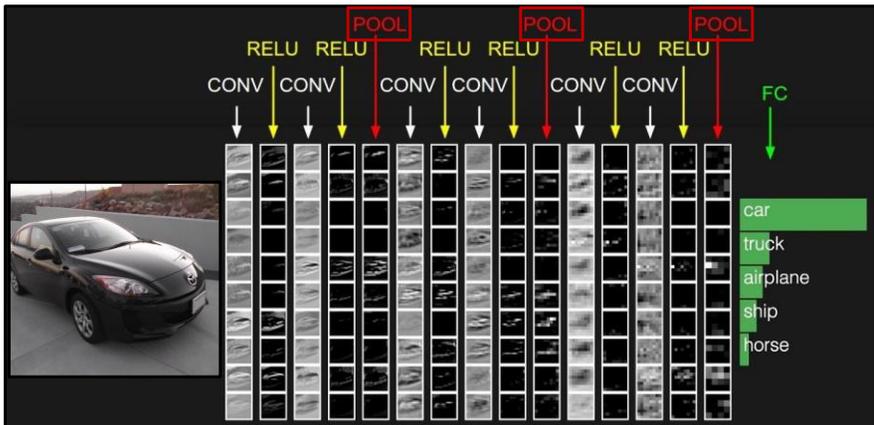
最大池化后的特征图

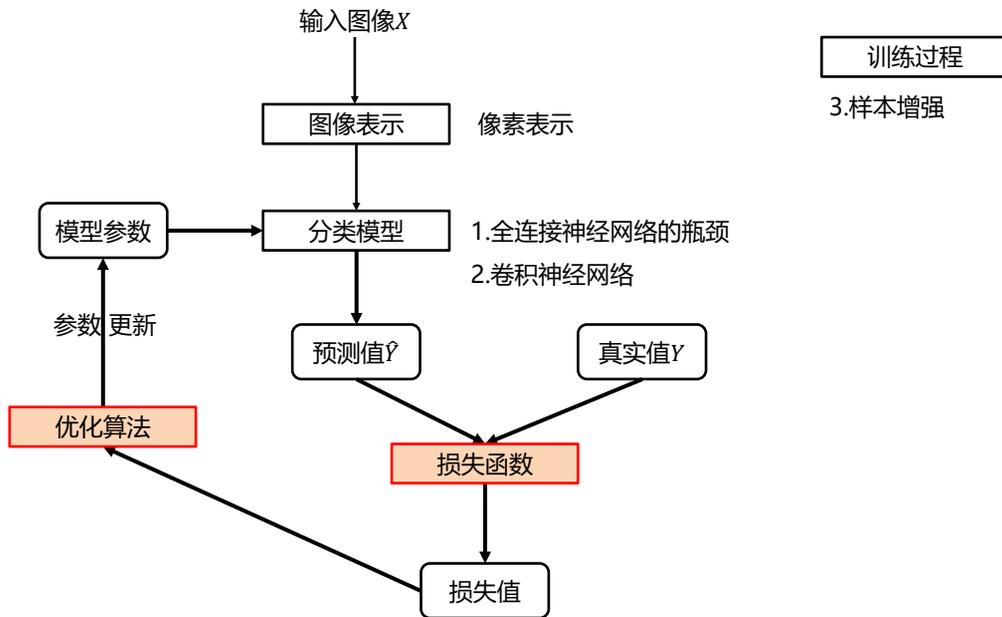
6	8
3	4

操作后，特征响应图中的75%的响应信息都丢掉，但不改变特征响应图的个数。

卷积神经网络

- CONV——卷积层
- RELU——激活层
- POOL——池化层
- FC——全连接层





2020/4/21

北京邮电大学计算机学院 鲁鹏

62

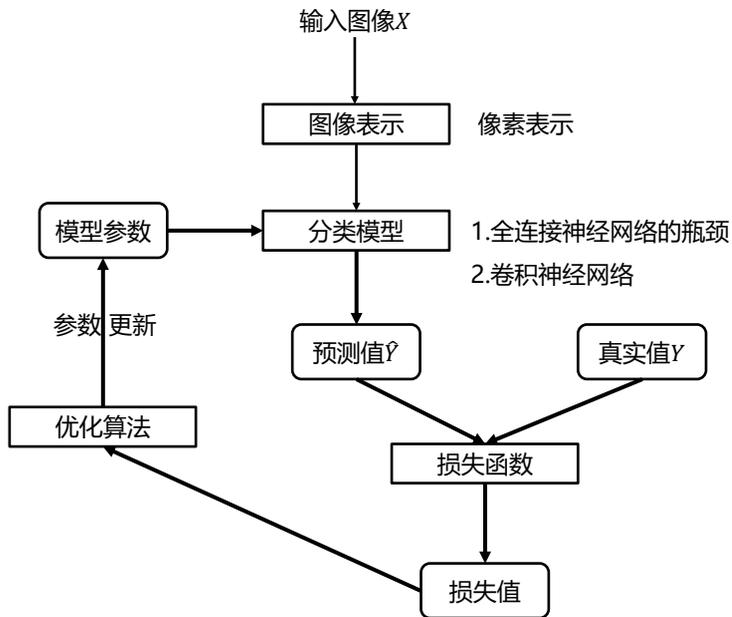
损失函数&优化算法

- 损失函数：交叉熵损失
- 优化算法：SGD、带动量的SGD以及ADAM

2020/4/21

北京邮电大学计算机学院 鲁鹏

63



训练过程

3.样本增强

2020/4/21

北京邮电大学计算机学院 鲁鹏

64

图像增强

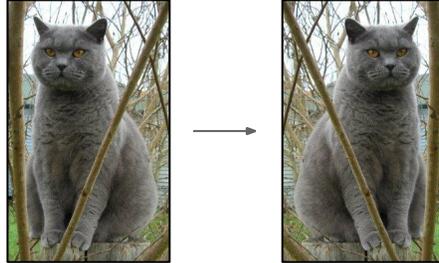
- **存在的问题**：过拟合的原因是学习样本太少，导致无法训练出能够泛化到新数据的模型。
- **数据增强**：是从现有的训练样本中生成更多的训练数据，其方法是利用多种能够生成可信图像的随机变换来增加样本。
- **数据增强的目标**：模型在训练时不会两次查看完全相同的图像。这让模型能够观察到数据的更多内容，从而具有更好的泛化能力

2020/4/21

北京邮电大学计算机学院 鲁鹏

65

样本增强——翻转



2020/4/21

北京邮电大学计算机学院 鲁鹏

66

样本增强——随机缩放&抠图

以残差网络中的样本增强方法为例

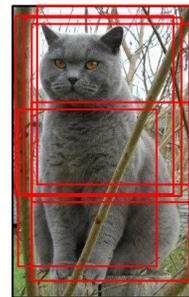
输入要求：224x224的彩色图片

训练阶段：在不同尺度、不同区域随机扣取

1. 在 [256, 480] 之间随机选择一个尺寸L
2. 将训练样本缩放至短边 = L
3. 在该样本上随机采样一个224 x 224的图像区域

测试阶段：按照一套预先定义的方式扣取

1. 将图像缩放成5种尺寸：{224, 256, 384, 480, 640}
2. 对每一个尺度的图像及其镜像图像，分别在其四个角及中间位置扣取224 x 224区域，即可获得10个图像



2020/4/21

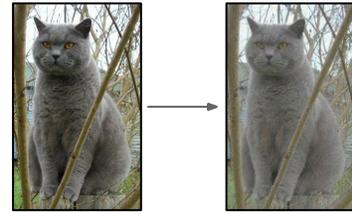
北京邮电大学计算机学院 鲁鹏

67

样本增强——色彩抖动

操作步骤：

1. 利用主成分分析方法提取当前图像的色彩数据（[R G B]）的主轴；
2. 沿着主轴方向随机采样一个偏移；
3. 将偏移量加入当前图像的每个像素。



样本增强——其他方案

随机联合下述操作

- 平移
- 旋转
- 拉伸
- 径向畸变（相关描述见摄像机几何章节）
- 裁剪
- ...