

# 深度学习2

主讲：郭春乐、刘夏雷  
南开大学计算机学院

致谢：本课件主要内容来自浙江大学吴飞教授、  
南开大学程明明教授

下面对误差反向传播 (error back propagation, BP)描述不正确的是 ( )

- ☐ A BP算法是一种将输出层误差反向传播给隐藏层进行参数更新的方法
- ☐ B BP算法将误差从后向前传递，获得各层单元所产生误差，进而依据这个误差来让各层单元修正各单元参数
- ☐ C 对前馈神经网络而言，BP算法可调整相邻层神经元之间的连接权重大小
- ☒ D 在BP算法中，每个神经元单元可包含不可偏导的映射函数

提交

以下关于批量梯度下降和随机梯度下降的说明，哪个描述是不正确的？

- ☐ A 在批量梯度下降和随机梯度下降中，为了最小化损失函数，通常使用循环迭代的方式不断更新模型参数；
- ☒ B 在每次迭代中，随机梯度下降需要计算训练集所有样本的误差和，用于更新模型参数；
- ☐ C 在每次迭代中，梯度下降使用所有数据或者部分训练数据，用于更新模型参数。

提交

关于sigmoid激活函数，下列描述正确的是？

- ☐ A 它是凸函数，凸函数无法解决非凸问题；
- ☐ B 它可以有负值；
- ☐ C 它无法配合交叉熵损失函数使用；
- ☒ D 当输入值过大或者过小时，梯度趋近于0，容易造成梯度消失问题。

提交

下面对前馈神经网络这种深度学习方法描述不正确的是（ ）

- ☐ A 是一种端到端学习的方法
- ☐ B 是一种监督学习的方法
- ☐ C 实现了非线性映射
- ☒ D 隐藏层数目大小对学习性能影响不大

提交

请问深度学习和传统的机器学习有什么区别？从数据大小、硬件要求、特征构建和解决问题方式等方面出发。

数据：深度学习适合处理大数据，机器学习算法更适用于小数据；

硬件：深度学习由于巨大的计算量，需要大量计算资源，比如GPU，机器学习算法对计算资源的需求相对较低；

特征构建：深度学习试图从数据中学习特征，机器学习中许多特征都需要由行业专家确定，并手工构造；

解决问题方式：深度学习通常利用“端到端”的方式构建模型，机器学习通常将问题分为几个步骤，每个步骤逐一解决，然后将结果组合。

作答

- (1) 请说明单层感知机为何无法模拟逻辑异或操作？
- (2) 请设计一个前馈神经网络，用于拟合异或函数。

(1) 根据图6.2.可知，在异或操作中，点 $(0,0)$ 与 $(1,1)$ 为0，点 $(0,1)$ 和 $(1,0)$ 为1，由于感知机是线性分类器，无法构建超平面可以将这两类分开。

(2) 输入层维度为2，用于接收两个操作数；至少有一层隐藏层，且隐藏层中的神经元需要包含非线性激活函数；输出层维度为1，用于输出结果。

作答

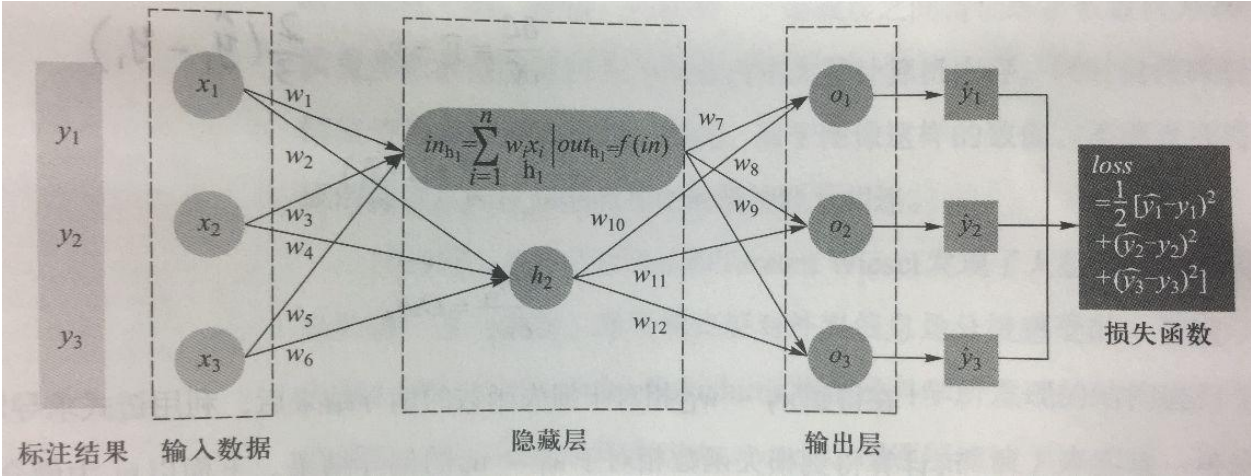
在前馈神经网络中，所有的参数能否被初始化为0？  
如果不能，能否全部被初始化为其他相同的值？  
为什么？

不能初始化为0，也不能被同时初始化为其他相同的值。  
如果参数被初始化为相同的值后，在误差反向传播过程中，同一层的神经元所接收到的误差都相同，更新后这些参数的值仍然相同。不管经过多少轮迭代，同一层神经元的参数保持相同，因此不同的神经元无法学习到不同特征的重要程度，失去了深度神经网络学习特征的能力。

作答



利用链式法则，求损失函数L对w6和w12的偏导数：（同书上推导过程）



单元	相邻前序神经元传递信息线性累加	非线性变换
$h_1$	$In_{h_1} = w_1 * x_1 + w_3 * x_2 + w_5 * x_3$	$Out_{h_1} = sigmoid(In_{h_1})$
$h_2$	$In_{h_2} = w_2 * x_1 + w_4 * x_2 + w_6 * x_3$	$Out_{h_2} = sigmoid(In_{h_2})$
$o_1$	$In_{o_1} = w_7 * Out_{h_1} + w_{10} * Out_{h_2}$	$\hat{y}_1 = Out_{o_1} = sigmoid(In_{o_1})$
$o_2$	$In_{o_2} = w_8 * Out_{h_1} + w_{11} * Out_{h_2}$	$\hat{y}_2 = Out_{o_2} = sigmoid(In_{o_2})$
$o_3$	$In_{o_3} = w_9 * Out_{h_1} + w_{12} * Out_{h_2}$	$\hat{y}_3 = Out_{o_3} = sigmoid(In_{o_3})$

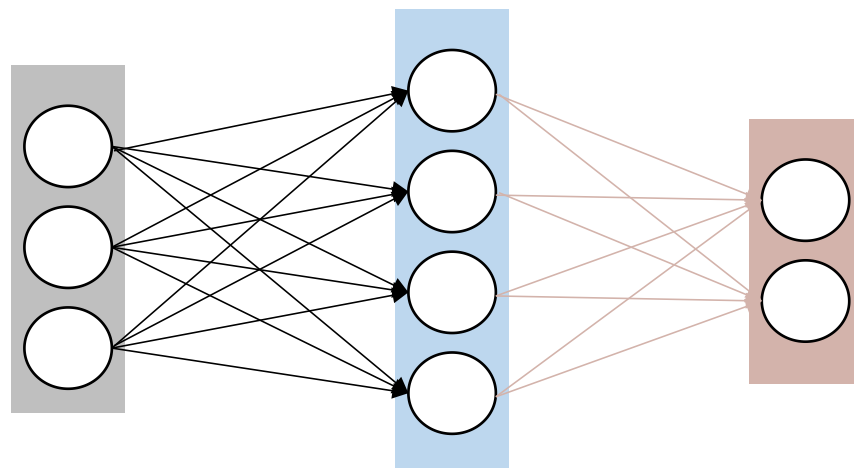
作答

# 提纲

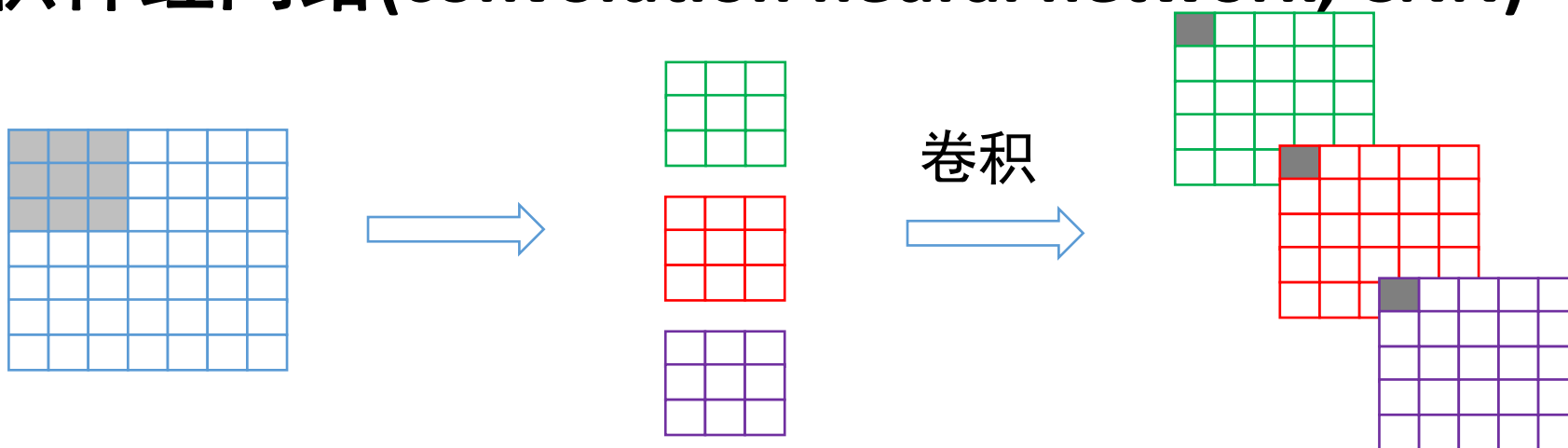
- 深度学习历史发展
- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 深度生成学习
- 深度学习应用

# 从前馈神经网络到卷积神经网络

- 三层前馈神经网络



- 卷积神经网络(convolution neural network, CNN)



# 卷积神经网络 (CNN)

- 前馈神经网络中，输入层的输入直接与第一个隐藏层中所有神经元相互连接。
  - 如果输入数据是一幅 $1000*1000$ 的图像，即 $1,000,000$ 维的向量，考虑输入数据与第一个隐藏层（与输入维度一致）中所有神经元均相连，则输入层到第一个隐藏层之间有 $10^{12}$ 个参数。
  - 模型参数数量如此巨大不仅会占用大量计算机内存，同时也使神经网络模型变得难以训练收敛。因此，对于图像这样的数据，不能直接将所构成的像素点向量与前馈神经网络神经元相连。

# 卷积神经网络 (CNN)

- 人脑 “视觉系统的信息处理” 这一机制
  - 即可视皮层对外界信息是分级感受的
  - 1959年，David Hubel和Torsten Wiesel发现
- 卷积神经网络的前身：满足平移不变性的网络Neocognitron
  - 1980年Kunihiko Fukushima，将神经科学所发现的结构进行了计算机模拟，提出级联方式（cascade，即逐层滤波）
- LeNet-5卷积神经网络用于手写体识别
  - 20世纪90年代，LeCun等人初步确立了CNN的基本结构。

# 卷积神经网络 (CNN)

- **图像中像素点具有很强的空间依赖性**
  - 卷积就是针对像素的空间依赖性来对图像进行处理的一种技术。
- **在图像卷积计算中，需要定义一个卷积核 (kernel)**
  - 卷积核是一个二维矩阵，矩阵中数值为对图像中与卷积核同样大小的子块像素点进行卷积计算时所采用的权重。
- **卷积核中的权重系数 $w_i$ 是通过数据驱动机制学习得到**
  - 用来捕获某像素点及其邻域像素点所构成的特有空间模式。
  - 一旦从数据中学习得到权重系数，这些权重系数就刻画了图像中像素点构成的空间分布不同模式。

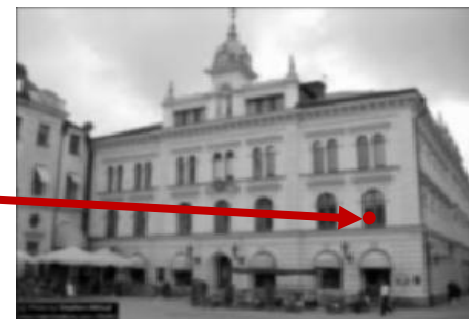


# 卷积神经网络：卷积操作

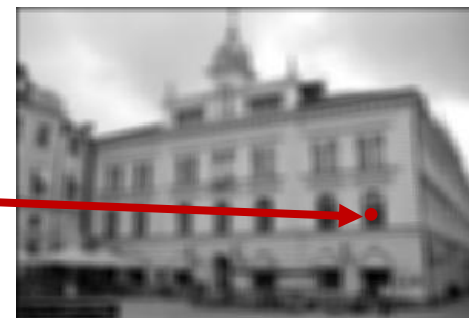
- 图像经过特定卷积矩阵滤波后，所得到的卷积结果可认为是保留了像素点所构成的特定空间分布模式
  - 如下给出了两个7\*7高斯卷积核的卷积结果



0.0000	0.0002	0.0011	0.0018	0.0011	0.0002	0.0000
0.0002	0.0029	0.0131	0.0216	0.0131	0.0029	0.0002
0.0011	0.0131	0.0586	0.0966	0.0586	0.0131	0.0011
0.0018	0.0216	0.0966	0.1592	0.0966	0.0216	0.0018
0.0011	0.0131	0.0586	0.0966	0.0586	0.0131	0.0011
0.0002	0.0029	0.0131	0.0216	0.0131	0.0029	0.0002
0.0000	0.0002	0.0011	0.0018	0.0011	0.0002	0.0000



0.0194	0.0199	0.0202	0.0203	0.0202	0.0199	0.0194
0.0199	0.0204	0.0207	0.0208	0.0207	0.0204	0.0199
0.0202	0.0207	0.0210	0.0211	0.0210	0.0207	0.0202
0.0203	0.0208	0.0211	0.0212	0.0211	0.0208	0.0203
0.0202	0.0207	0.0210	0.0211	0.0210	0.0207	0.0202
0.0199	0.0204	0.0207	0.0208	0.0207	0.0204	0.0199
0.0194	0.0199	0.0202	0.0203	0.0202	0.0199	0.0194



# 卷积神经网络：卷积操作示意



转变为灰度图



$$(77 \times 1 + 75 \times 2 + 76 \times 1 + 74 \times 2 + 65 \times 4 + 77 \times 2 + 75 \times 1 + 75 \times 2 + 74 \times 1) / (1 + 2 + 1 + 2 + 4 + 2 + 1 + 2 + 1) = 68.125 = 68$$

77	75	76
74	65	77
75	75	74

X

1	2	1
2	4	2
1	2	1

得到

68

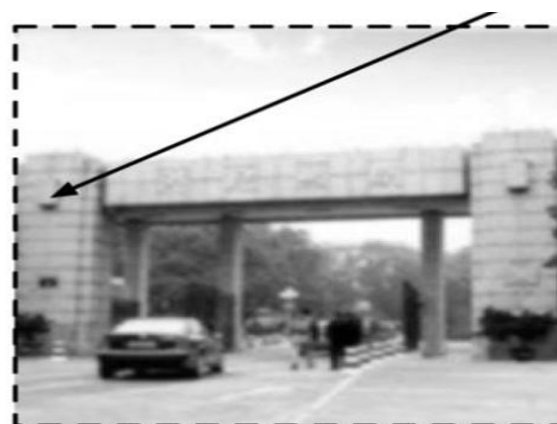
$$(77 \times 1 + 75 \times 2 + 76 \times 1 + 74 \times 2 + 65 \times 4 + 77 \times 2 + 75 \times 1 + 75 \times 2 + 74 \times 1) / (1 + 2 + 1 + 2 + 4 + 2 + 1 + 2 + 1) = 68.125 = 68$$

灰度图中的一个3 X 3 子块

3 X 3 高斯滤波器



使用了5 X 5 高斯模版卷积后的结果图

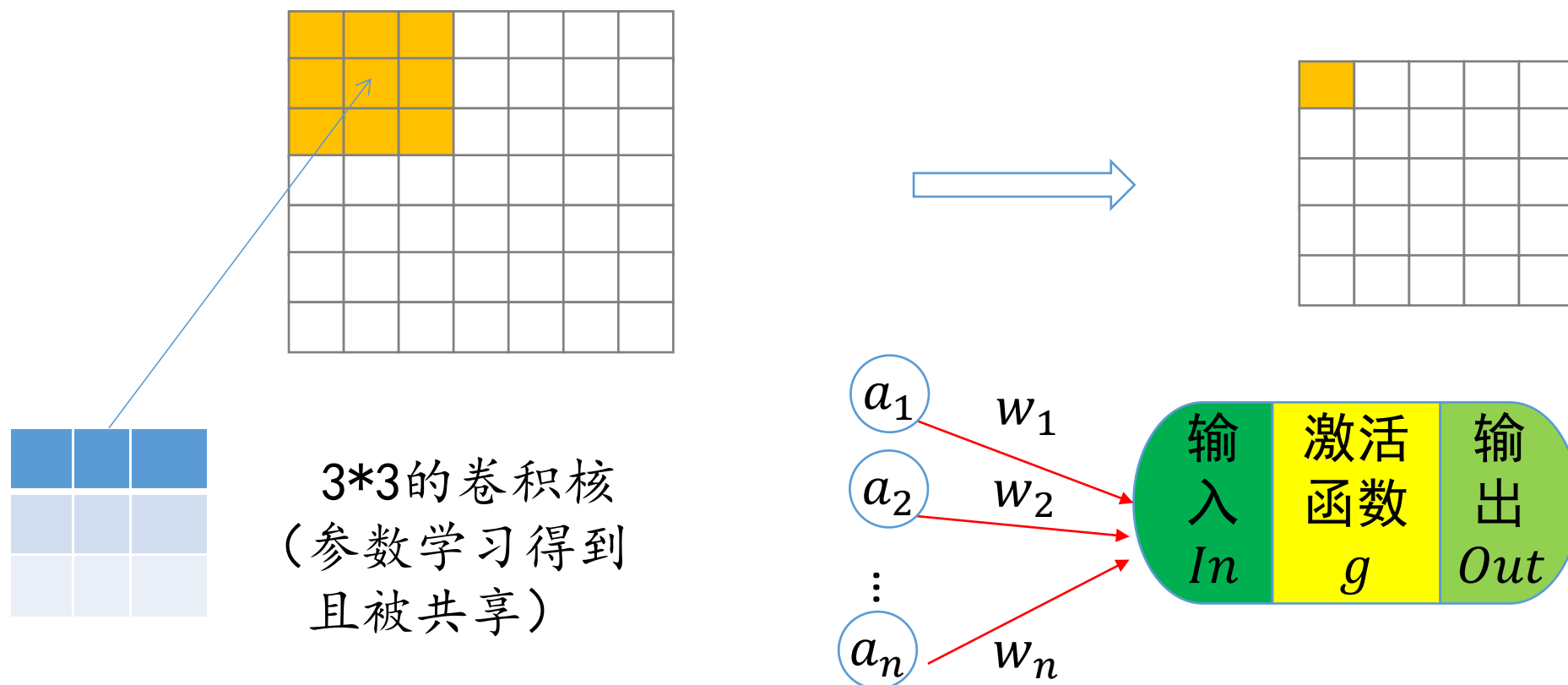


使用了3 X 3 高斯模版卷积后的结果图



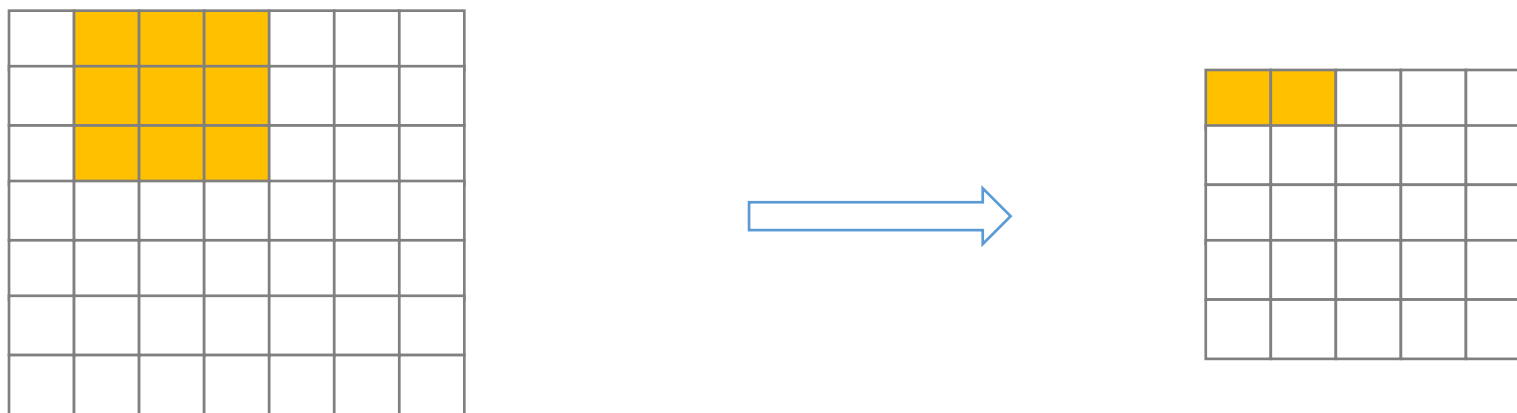
# 卷积神经网络：卷积操作

7\*7大小的图像，通过3\*3大小卷积矩阵以1的步长进行卷积操作，可得到5\*5大小的卷积结果



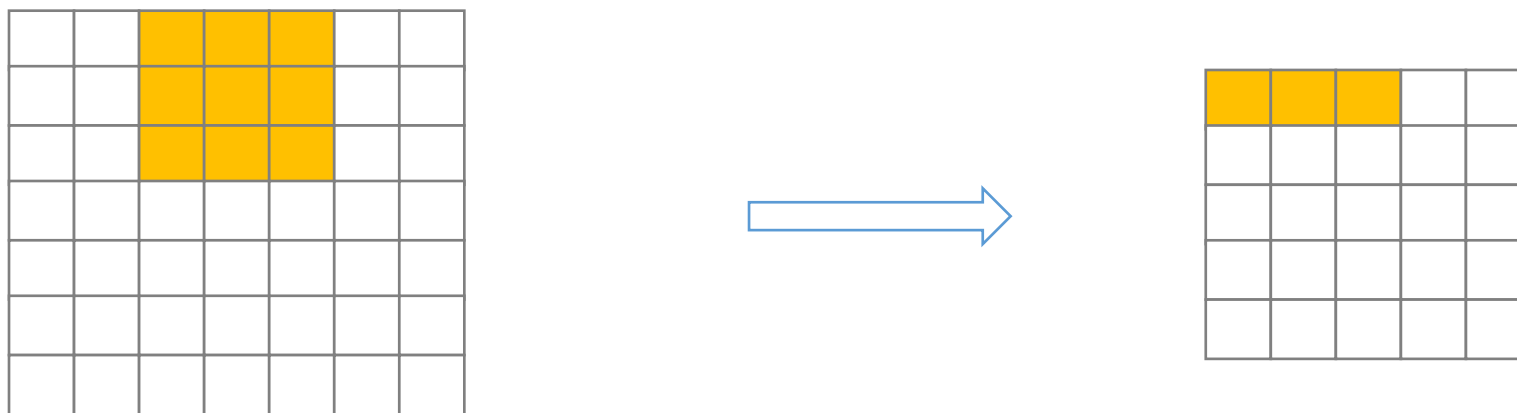
# 卷积神经网络：卷积操作

7\*7大小的图像，通过3\*3大小卷积矩阵以1的步长进行卷积操作，可得到5\*5大小的卷积结果



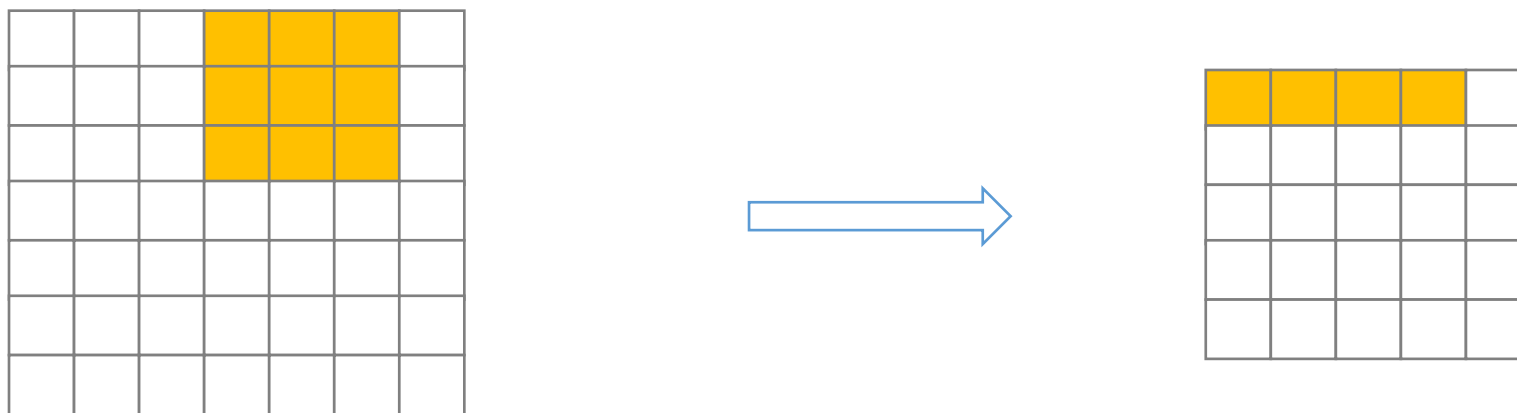
# 卷积神经网络：卷积操作

7\*7大小的图像，通过3\*3大小卷积矩阵以1的步长进行卷积操作，可得到5\*5大小的卷积结果



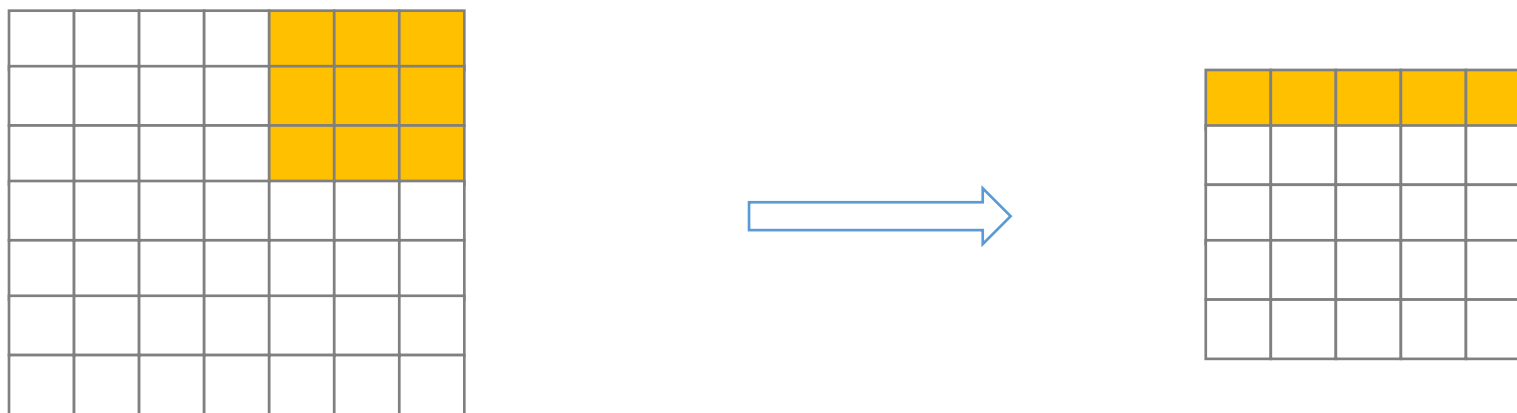
# 卷积神经网络：卷积操作

7\*7大小的图像，通过3\*3大小卷积矩阵以1的步长进行卷积操作，可得到5\*5大小的卷积结果



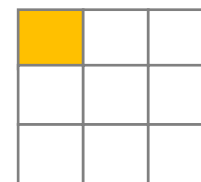
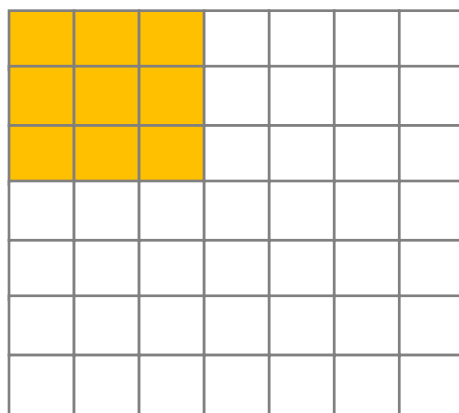
# 卷积神经网络：卷积操作

7\*7大小的图像，通过3\*3大小卷积矩阵以1的步长进行卷积操作，可得到5\*5大小的卷积结果



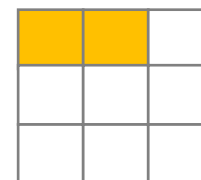
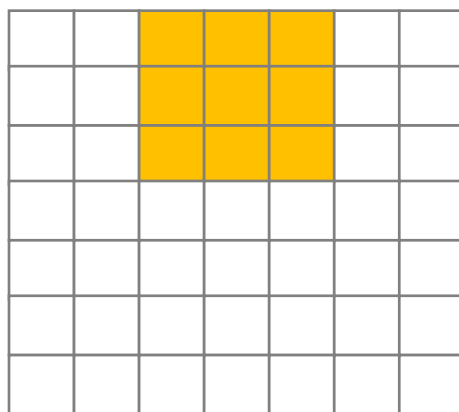
# 卷积神经网络：卷积操作

如果步长改为2



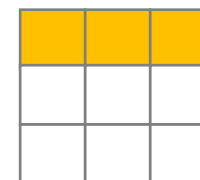
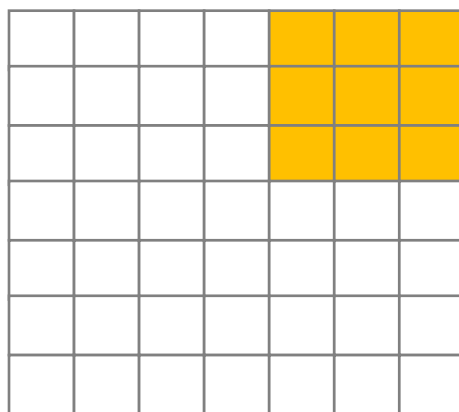
# 卷积神经网络：卷积操作

如果步长改为2



# 卷积神经网络：卷积操作

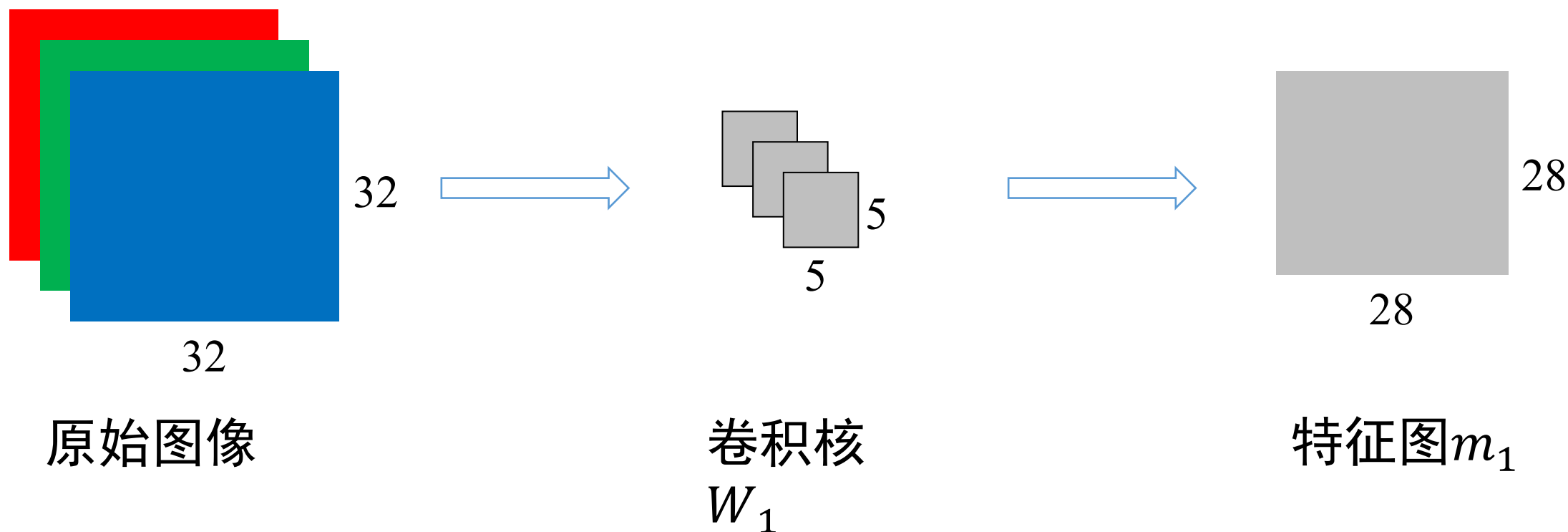
如果步长改为2





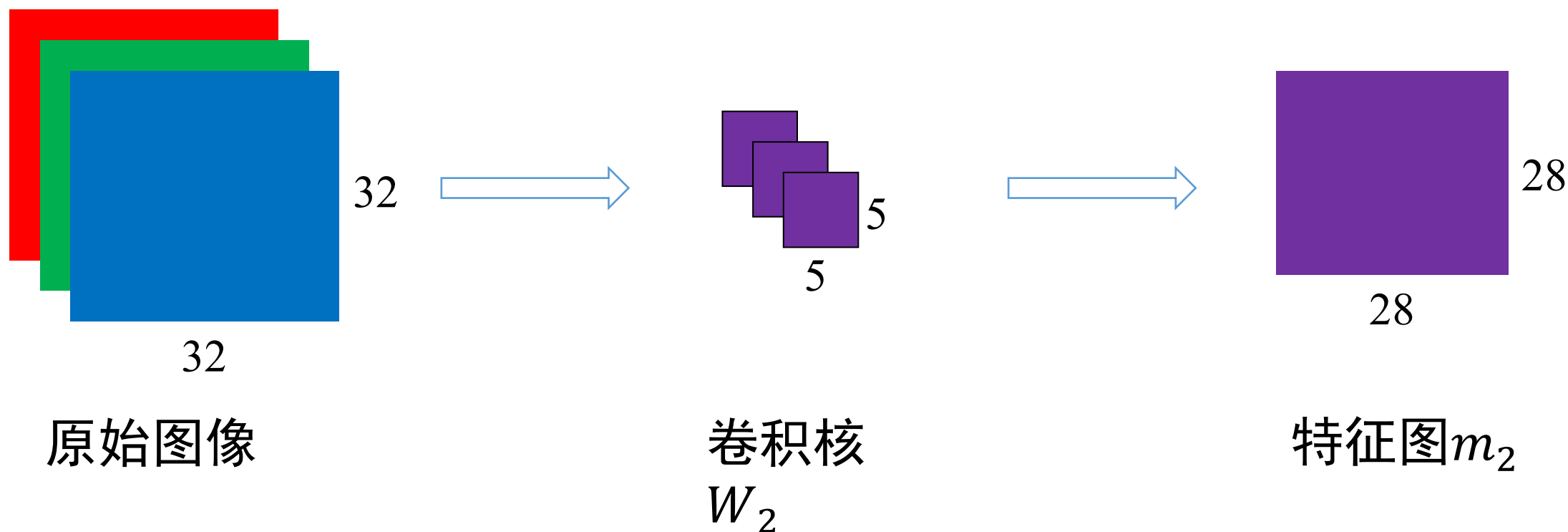
# 卷积神经网络：卷积操作

有一张 $32*32*3$  (RGB) 的图像，使用 $5*5*3$ 的卷积核 $W_1$ ，步长为1对其进行卷积操作。卷积核 $W_1$ 在原始图像上从左到右、从上到下进行计算，改变 $5*5$ 子块区域中的中心像素点值，得到 $28*28$ 的特征图 $m_1$



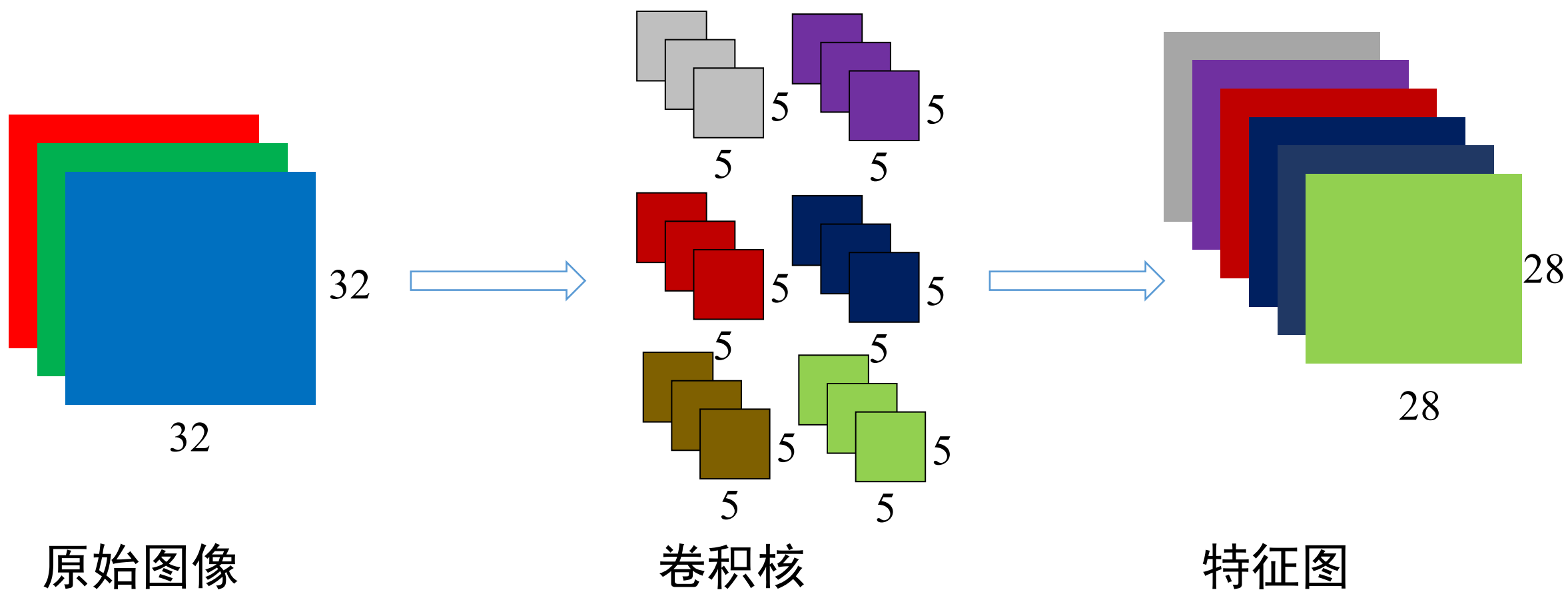
# 卷积神经网络：卷积操作

使用另一个 $5*5*3$ 的卷积核 $W_2$ 与原始图像做卷积操作，得到特征图 $m_2$



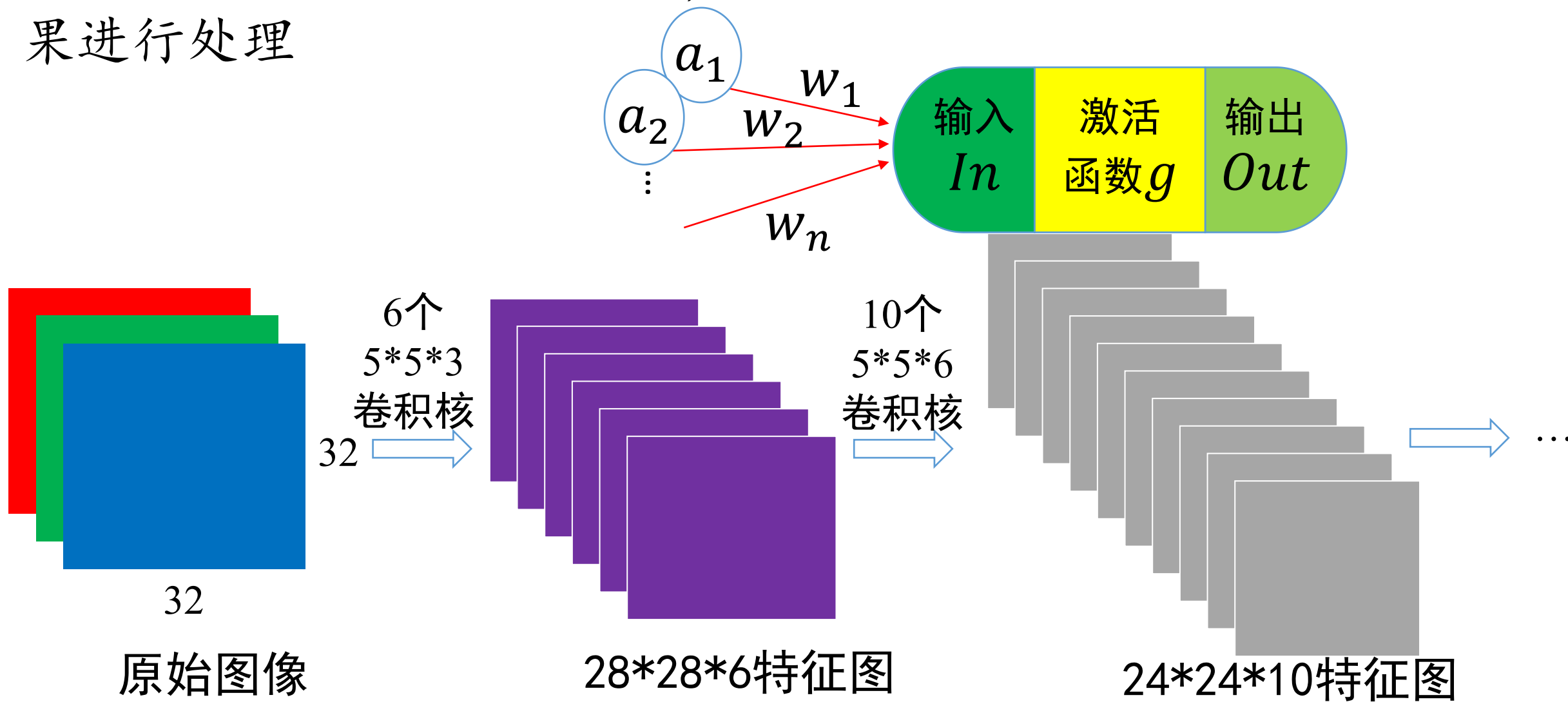
# 卷积神经网络：卷积操作

- 使用6个 $5*5*3$ 的卷积核与原始图像做卷积操作，则得到6个 $28*28$ 的特征图
- 注意:6个 $5*5*3$ 的卷积核均是数据驱动学习得到，其刻画了不同的视觉模式



# 卷积神经网络：卷积+激活函数(非线性映射)

在对原始图像做卷积操作后，可使用ReLU激活函数对卷积后结果进行处理



# 卷积神经网络：卷积操作

- 神经科学家发现，人的视觉神经细胞对不同的视觉模式具有**特征选择性**（Feature Selectivity）
  - 不同视觉神经细胞对边界、运动和颜色等不同信息具有强弱不同的选择性。不同卷积核可被用来刻画视觉神经细胞对外界信息感受时的不同选择性。
  - 卷积所得结果中，每个输出点的取值仅依赖于输入图像中该点及其邻域区域点的取值，与区域之外的其他无关，该区域被称为**感受野** (receptive field)。感受野是CNN每一层输出的特征图（feature map）上的一个点在输入图像上映射的区域。

# 卷积神经网络：池化(pooling)操作

- 对输入的特征图进行下采样，以获得最主要信息

- 常用的池化操作有：

- 最大池化

- 平均池化

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

→  
最大池化

6	4
8	9

→  
平均池化

3.5	2.75
3.5	7

# 卷积神经网络：最大池化操作

在输入特征图中每一个区域寻找最大值。

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

对2\*2大小区域，按照步长为2进行最大池化操作



6	4
8	9

# 卷积神经网络：平均池化操作

对输入特征图的每一个区域的值求平均值（取整）

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

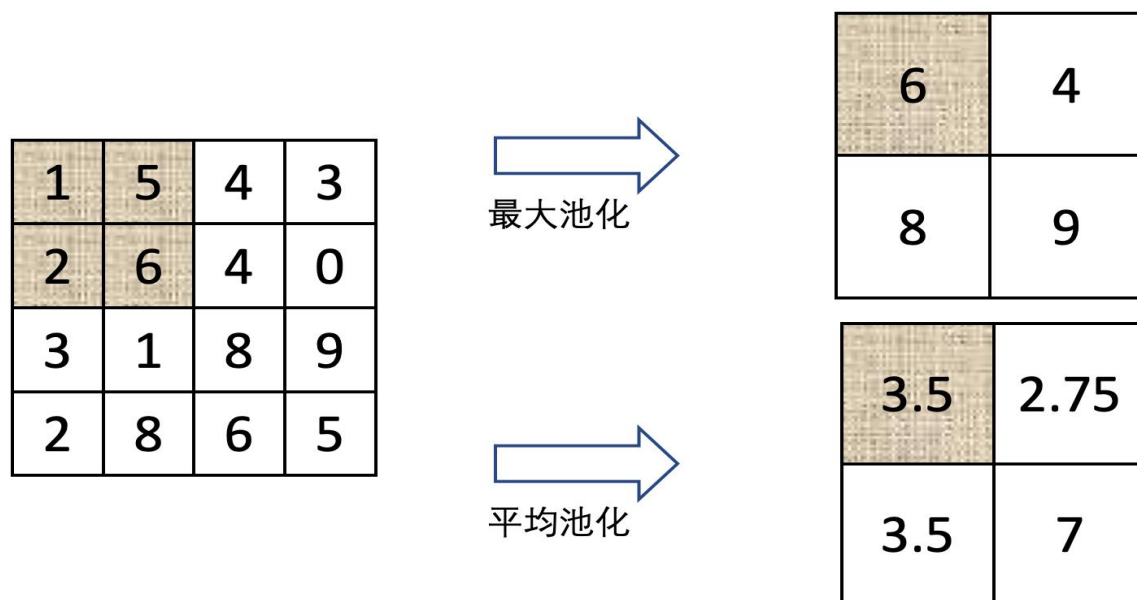
对2\*2大小区域，按照步长为2进行平均池化操作



3	2
3	6



# 卷积神经网络：池化操作



由于图像中存在较多冗余，在图像处理中，可用某一区域子块的统计信息（如最大值或均值等）来刻画该区域中所有像素点呈现的空间分布模式，以替代区域子块中所有像素点取值，这就是卷积神经网络中池化(pooling)操作。池化操作对卷积结果特征图进行约减，实现了下采样，同时保留了特征图中主要信息。

# 卷积神经网络：池化操作

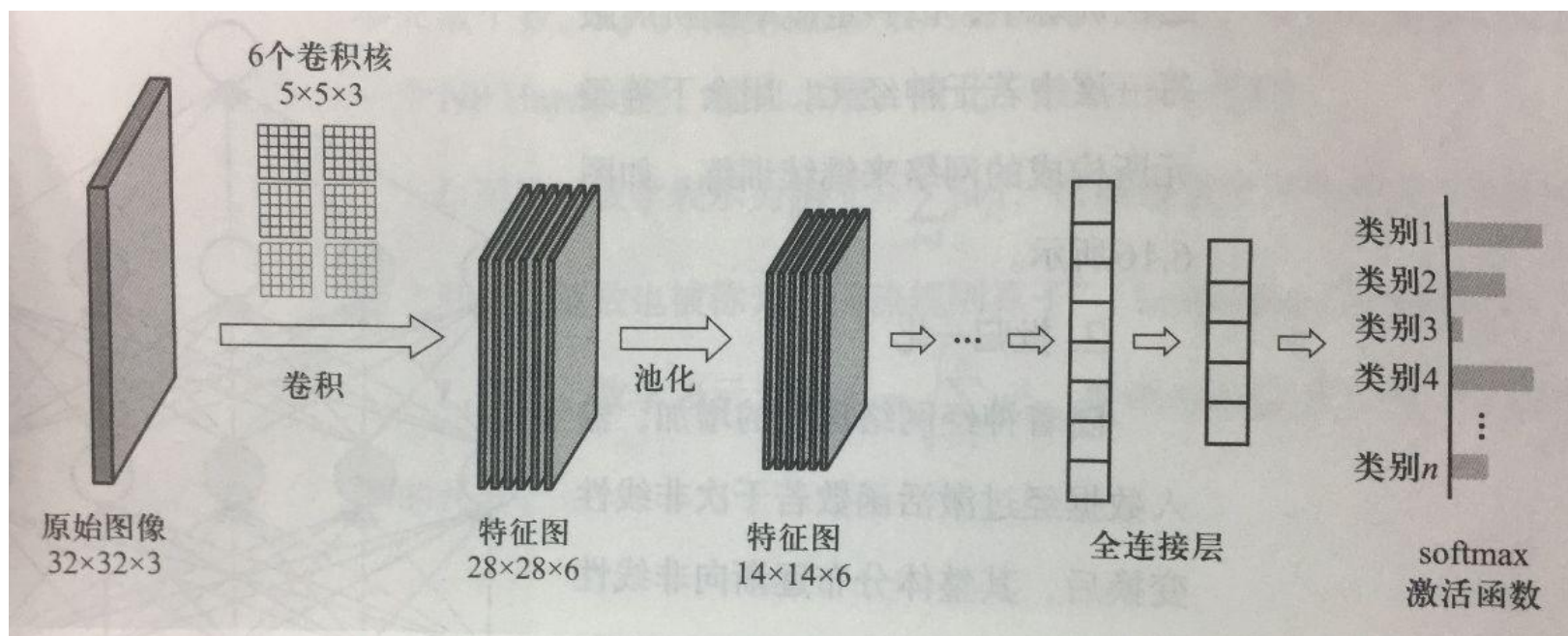
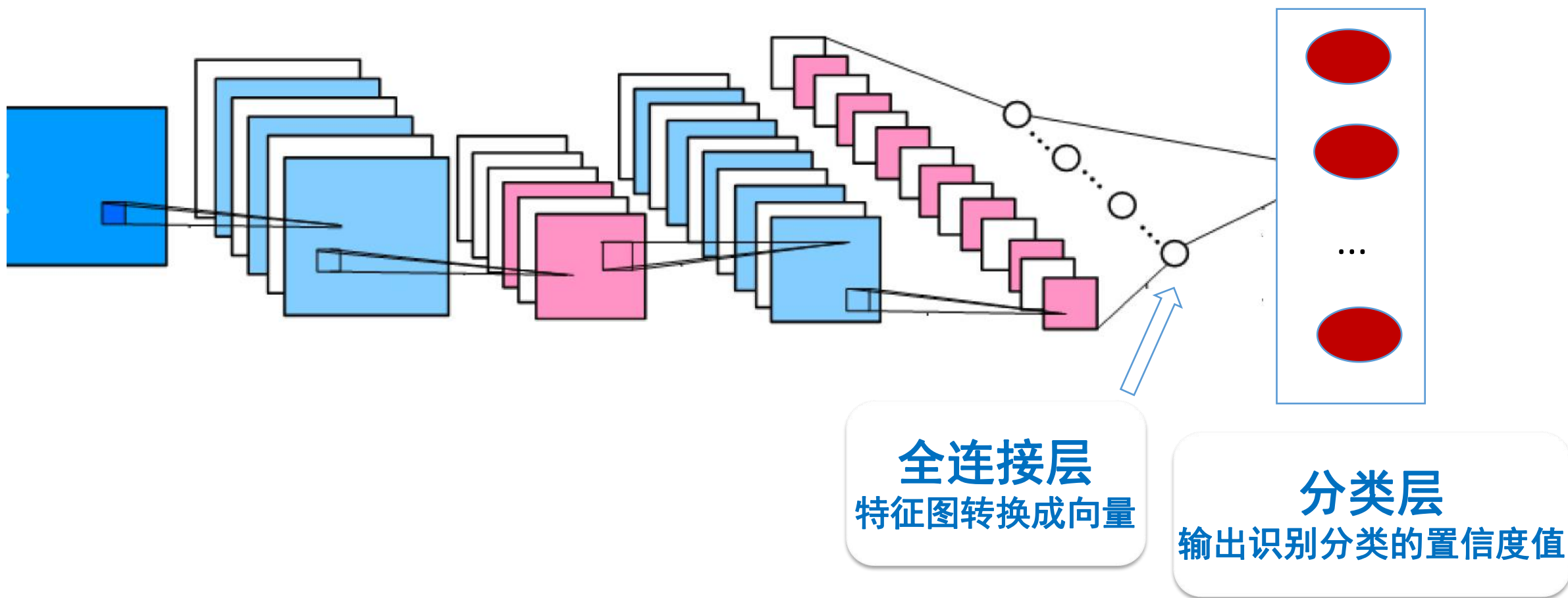


图6.15 基于卷积神经网络的图像分类示意图

对于输入的海量标注数据，通过多次迭代训练，卷积神经网络在若干次卷积操作、接着对卷积所得结果进行激活函数操作和池化操作下，最后通过全连接层来学习得到输入数据的特征表达，即**分布式向量表达(distributed vector representation)**。

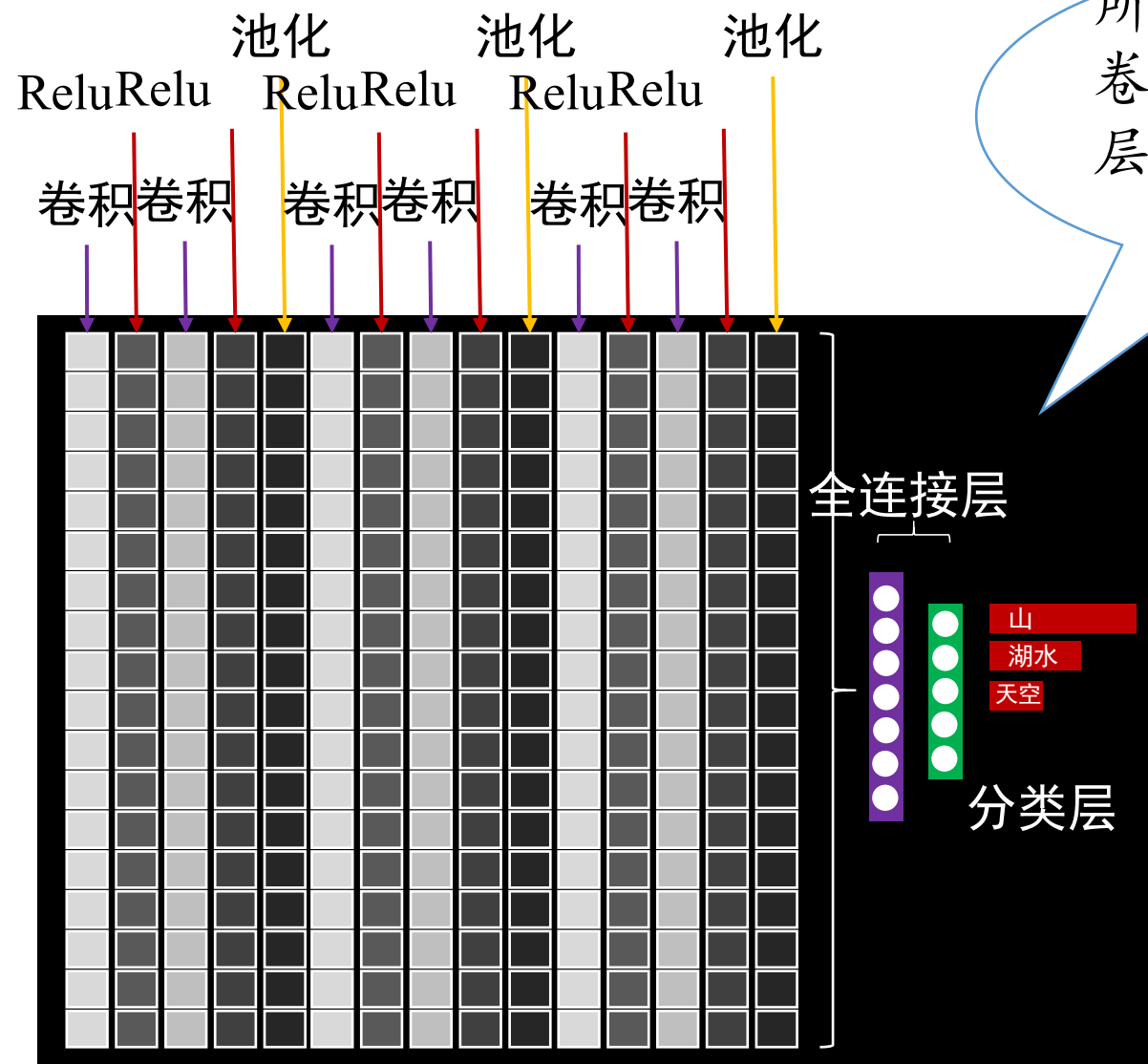
# 卷积神经网络：全连接层与分类层



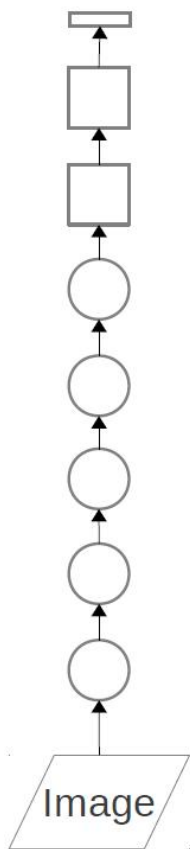
# 卷积神经网络基本架构



输入



# 卷积神经网络的参数



- Trained with stochastic gradient descent on two NVIDIA GPUs for about a week
- 650,000 neurons
- 60,000,000 parameters
- 630,000,000 connections
- **Final feature layer:** 4096-dimensional

体现了数据、模型和算力的结合



**Convolutional layer:** convolves its input with a bank of 3D filters, then applies point-wise non-linearity



**Fully-connected layer:** applies linear filters to its input, then applies point-wise non-linearity

Alex Krizhevsky, et.al., ImageNet Classification with Deep Convolutional Neural Networks



# 神经网络正则化

为了缓解神经网络在训练过程中出现的过拟合问题，需要采取一些正则化技术来提升神经网络的泛化能力 (generalization)

- Dropout
- Batch-Normalization
- L1-Norm & L2-Norm

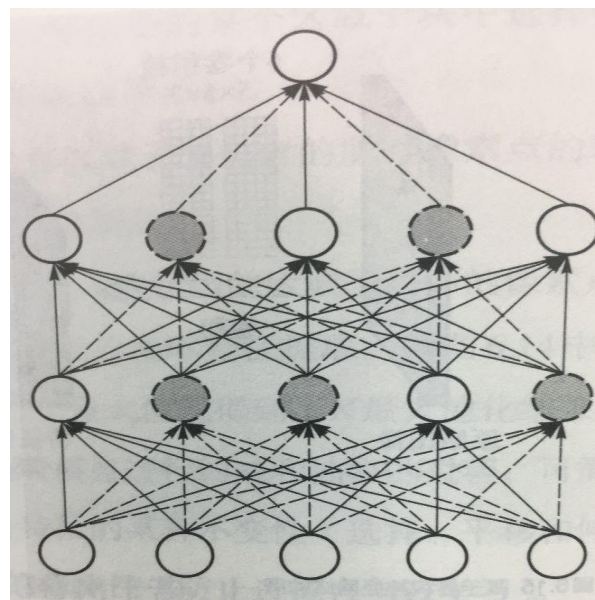


图6.16 使用Dropout的神经网络模型

$L_1$  范数：数学表示为  $\|W\|_1 = \sum_{i=1}^N |w_i|$ ，指模型参数  $W$  中各个元素的绝对值之和。 $L_1$  范数也被称为“稀疏规则算子” (Lasso regularization) 。

$L_2$  范数：数学表示为  $\|W\|_2 = \sqrt{\sum_{i=1}^N w_i^2}$ ，指模型参数  $W$  中各个元素平方和的开方。

# 提纲

- 深度学习历史发展
- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 深度生成学习
- 深度学习应用

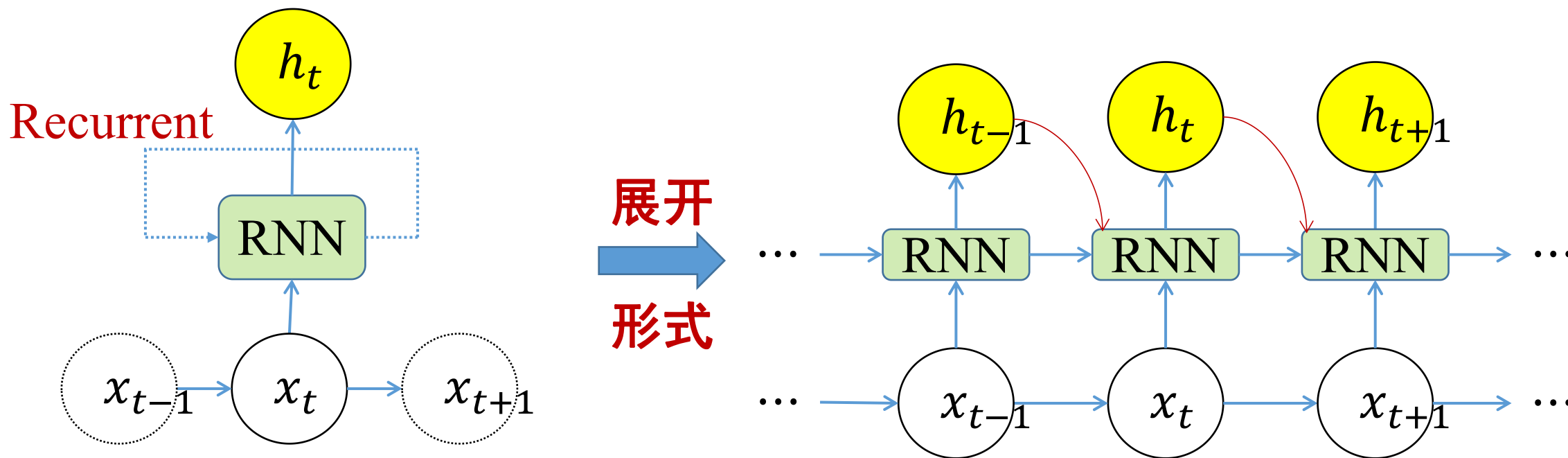
# 循环神经网络（RNN）

- 循环神经网络（Recurrent Neural Network, RNN）是一类处理序列数据（如文本句子、视频帧等）时所采用的网络结构。先前所介绍的前馈神经网络或卷积神经网络所需要处理的输入数据一次性给定，难以处理存在前后依赖关系的数据。
- 循环神经网络的本质是希望模拟人所具有的记忆能力，在学习过程中记住部分已经出现的信息，并利用所记住的信息影响后续结点输出。循环神经网络在自然语言处理，例如语音识别、情感分析、机器翻译等领域有重要应用。



# 循环神经网络（RNN）

- 循环神经网络在处理数据过程中构成了一个循环体
  - 对于序列数据，在 $t$ 时刻循环神经网络单元会读取当前输入数据 $x_t$ 和前一时刻输入数据 $x_{t-1}$ 所对应的隐式编码结果 $h_{t-1}$ ，一起生成 $t$ 时刻的隐式编码结果 $h_t$ 。接着将 $h_t$ 后传。



# 提纲

- 深度学习历史发展
- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 深度生成学习
- 深度学习应用

# 深度生成学习模型 (deep generative learning model)

在本章之前的介绍中，神经网络模型从数据中提取出高层语义在数据中所蕴含的“模式”，并利用这些模式实现对数据的分类和检测等，这种模型通常称为判别模型，判别模型不关心数据如何生成，它只关心数据蕴含哪些模式以及如何将数据进行分类。与之相对的模型类型被称为生成模型(generative model)，生成模型需要学习目标数据的分布规律，以合成属于目标数据空间的新数据。

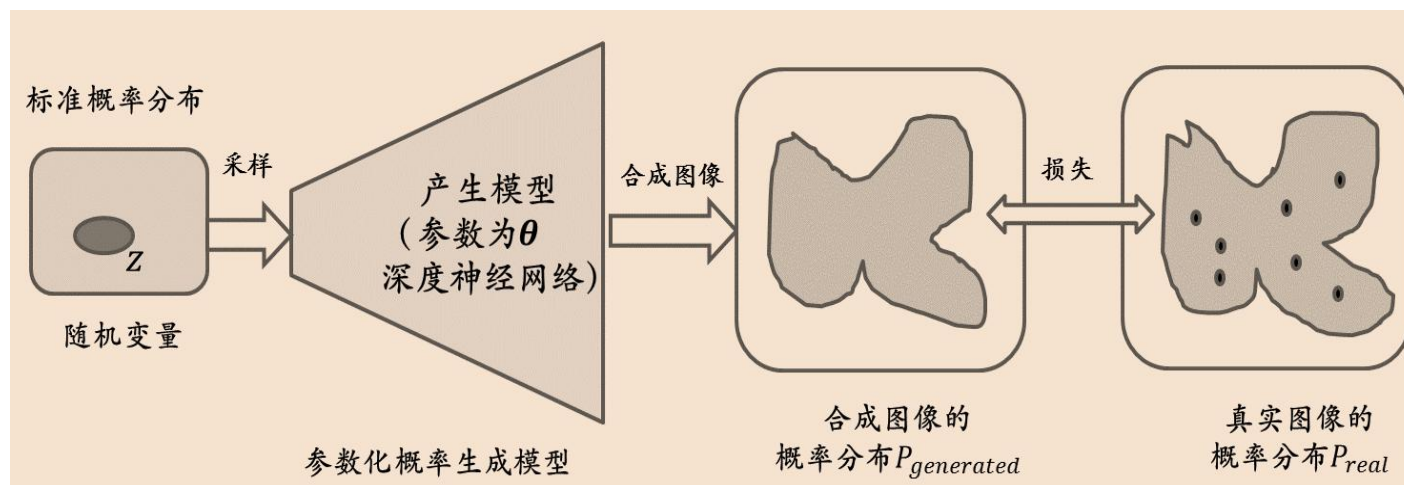


图6.23 生成模型的示意图

变分自编码器(variational auto-encoder, VAE)、自回归模型(Autoregressive models)与生成对抗网络 (generative adversarial network, GAN) 等.....

# 生成对抗网络（deep generative learning model）

生成对抗网络是由Ian Goodfellow等人于2014年提出的一种生成模型[Goodfellow2014]，该模型可视为两个神经网络相互竞争的零和游戏（zero-sum game），“以子之矛，陷子之盾”，最终达到纳什均衡。

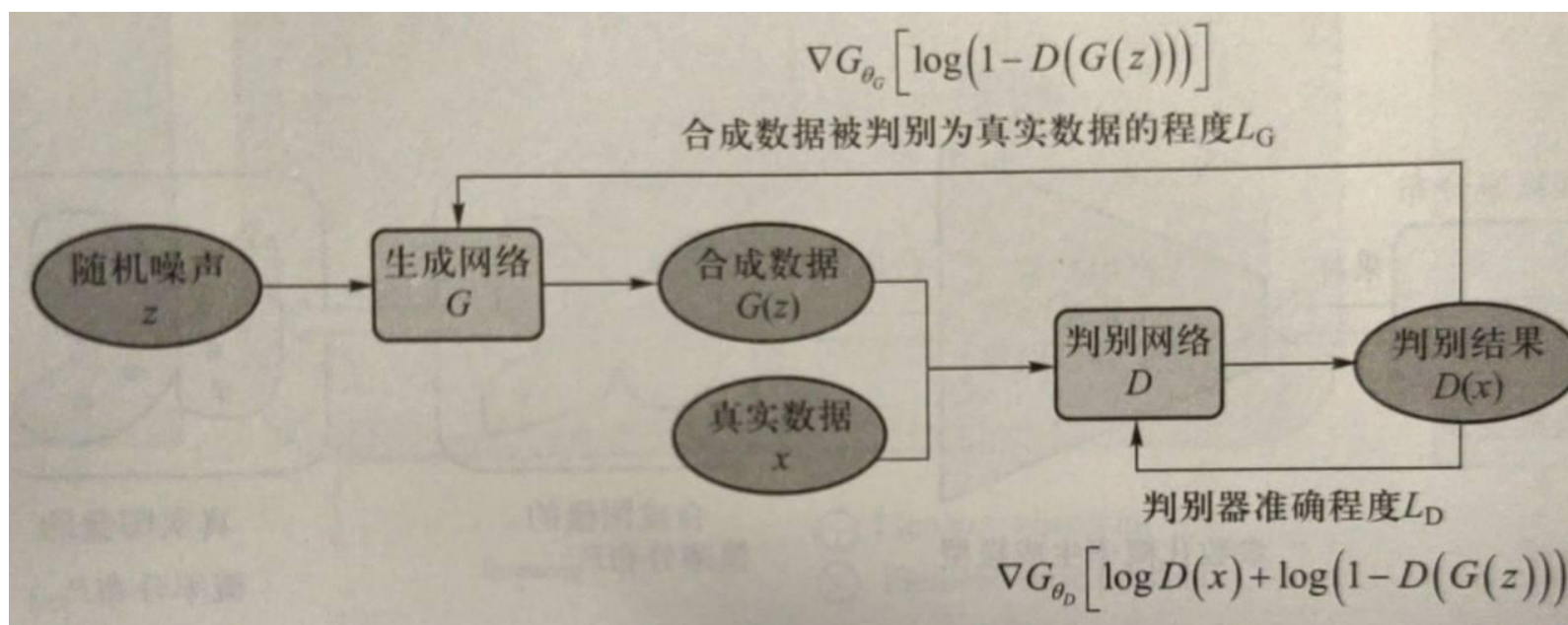


图6.24 生成对抗网络模型示意图

# 生成对抗网络 (deep generative learning model)

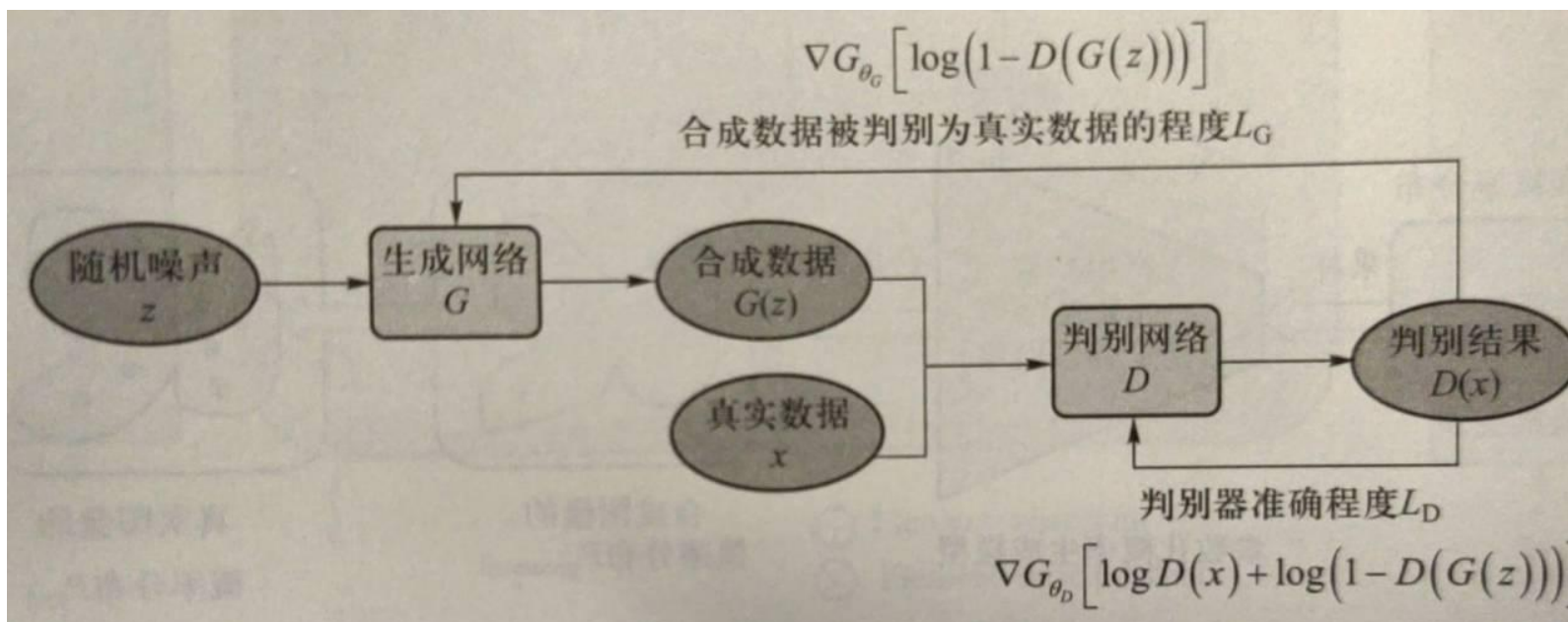


图6.24 生成对抗网络模型示意图

生成对抗网络由一个生成器（generator，简称G）和一个判别器（discriminator，简称D）组成。GAN的核心是通过生成器和判别器两个神经网络之间的竞争对抗，不断提升彼此水平以使得生成器所生成数据（人为伪造数据）与真实数据相似，判别器无法区分真实数据和生成数据。

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$



# 生成对抗网络 (deep generative learning model)

## 算法6.5.1 训练生成对抗网络

输入：神经网络  $D$ 、 $G$ ，噪声分布  $p_z(z)$ ，真实数据分布  $p_{data}(x)$

输出：神经网络参数  $\theta_d$ 、 $\theta_g$

算法步骤：

每轮训练循环执行：

    训练  $k$  轮判别器：

        从噪声分布  $p_z(z)$  中采样  $m$  个样本  $\{z^{(1)}, z^{(2)}, z^{(3)} \dots z^{(m)}\}$

        从真实数据分布  $p_{data}(x)$  中采样  $m$  个样本  $\{x^{(1)}, x^{(2)}, x^{(3)} \dots x^{(m)}\}$

        沿梯度上升方向更新判别器参数：

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m [\log D(x^{(i)}) + \log(1 - D(G(z^{(i)})))]$$

        从噪声分布  $p_z(z)$  中采样  $m$  个样本  $\{z^{(1)}, z^{(2)}, z^{(3)} \dots z^{(m)}\}$

        沿梯度下降方向更新判别器参数：

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log(1 - D(G(z^{(i)})))$$

在训练初期，因为生成数据质量较低，判别器  $D$  可以轻易将其区分，此时  $\log(1 - D(G(z)))$  无法为生成器  $G$  提供显著的梯度，因此可以使用  $-\log D(G(z))$  来代替损失函数。这么做并不会影响其能够收敛至最优解的性质，且可以在训练初期提供足够的梯度。

# 生成对抗网络 (deep generative learning model)

在后续的应用中，发现GAN具有收敛困难，且非常依赖于合适的交替训练轮次选择等不足。为此，Martin Arjovsky等人提出了Wasserstein GAN[Arjovsky 2017]，对模型结构进行了一些修改。在Wasserstein GAN中，模型输出不使用以概率值形式输出的sigmoid激活函数，而是直接根据输出值大小评价其效果。相应地，目标函数中也不需要再取对数：

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [D(x)] - \mathbb{E}_{z \sim p_z(z)} [D(G(z))]$$

# 条件生成对抗网络

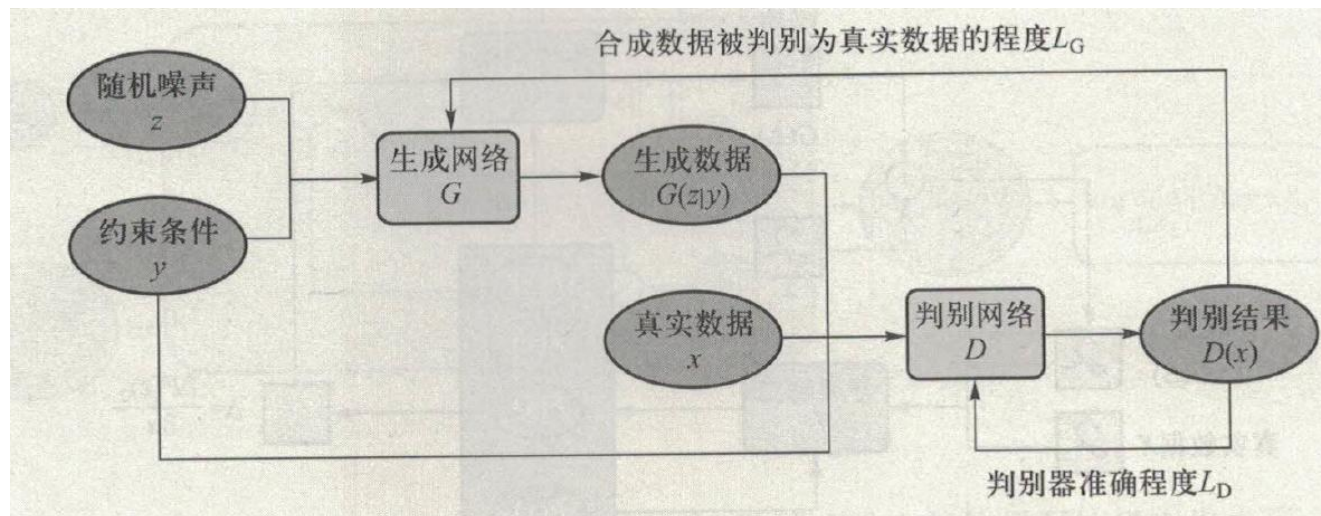


图6.25 条件生成对抗网络模型示意图

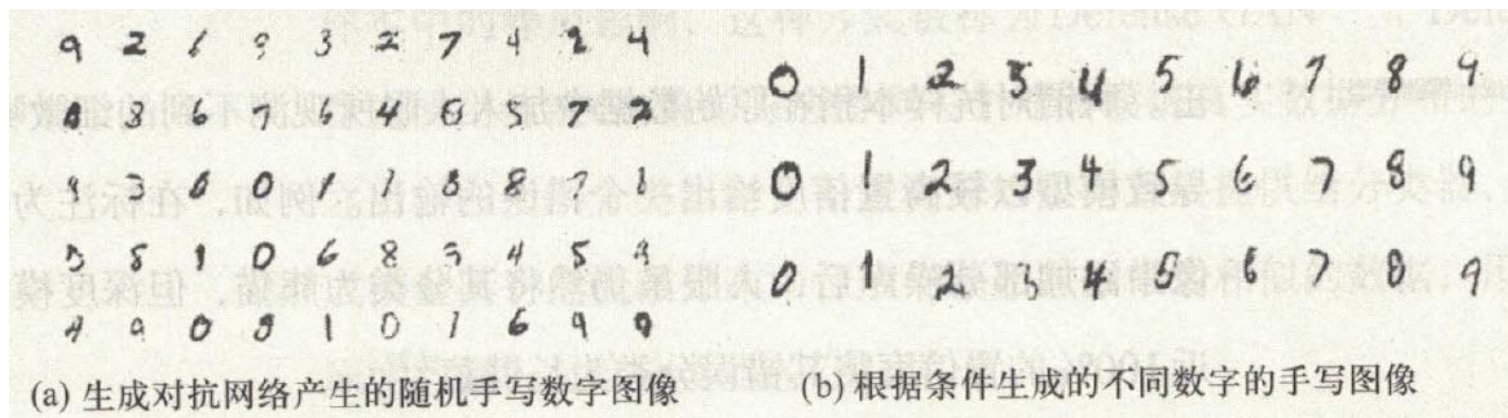


图6.26 不同生成对抗网络模型产生的手写数字图像



# 提纲

- 深度学习历史发展
- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 深度生成学习
- 深度学习应用

# 研讨内容

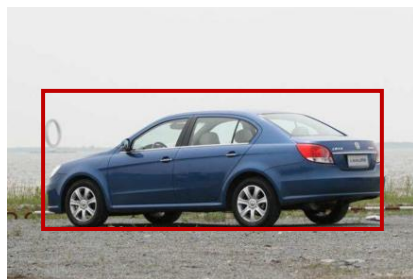
- **选题：图像超分、图像生成、图像分类、目标检测**
- **组队：3人**
- **提交报告：问题描述、每篇文章核心内容、相互联系等**
- **截止日期：2022年5月30日**

# 图像分类和目标定位

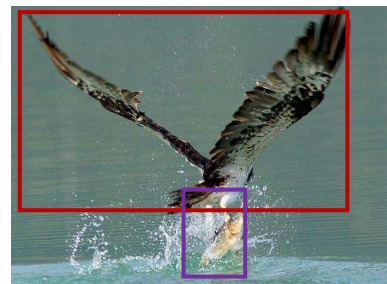
**Classification**

**Localization**

**Object Detection**



...



...

**Car**

**(x, y, w, h)**

**Eagle, Fish**

**Single Object**

**Multi Objects**

# 图像分类



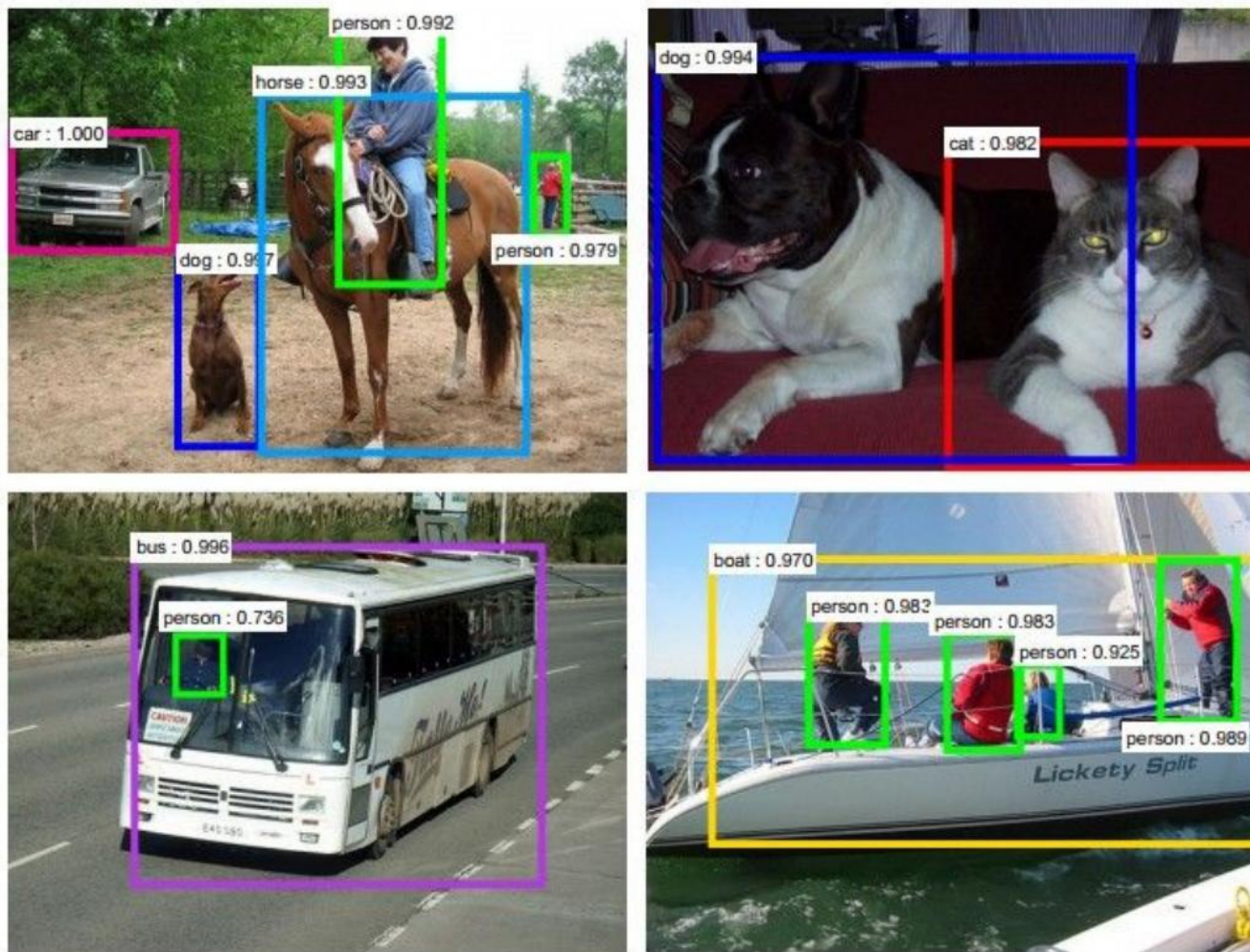
Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems* 25 (2012).

# 图像分类

- **[AlexNet]** Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems* 25 (2012).
- **[VGGNet]** Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *ICLR 2015*.
- **[ResNet]** He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 770-778. 2016.

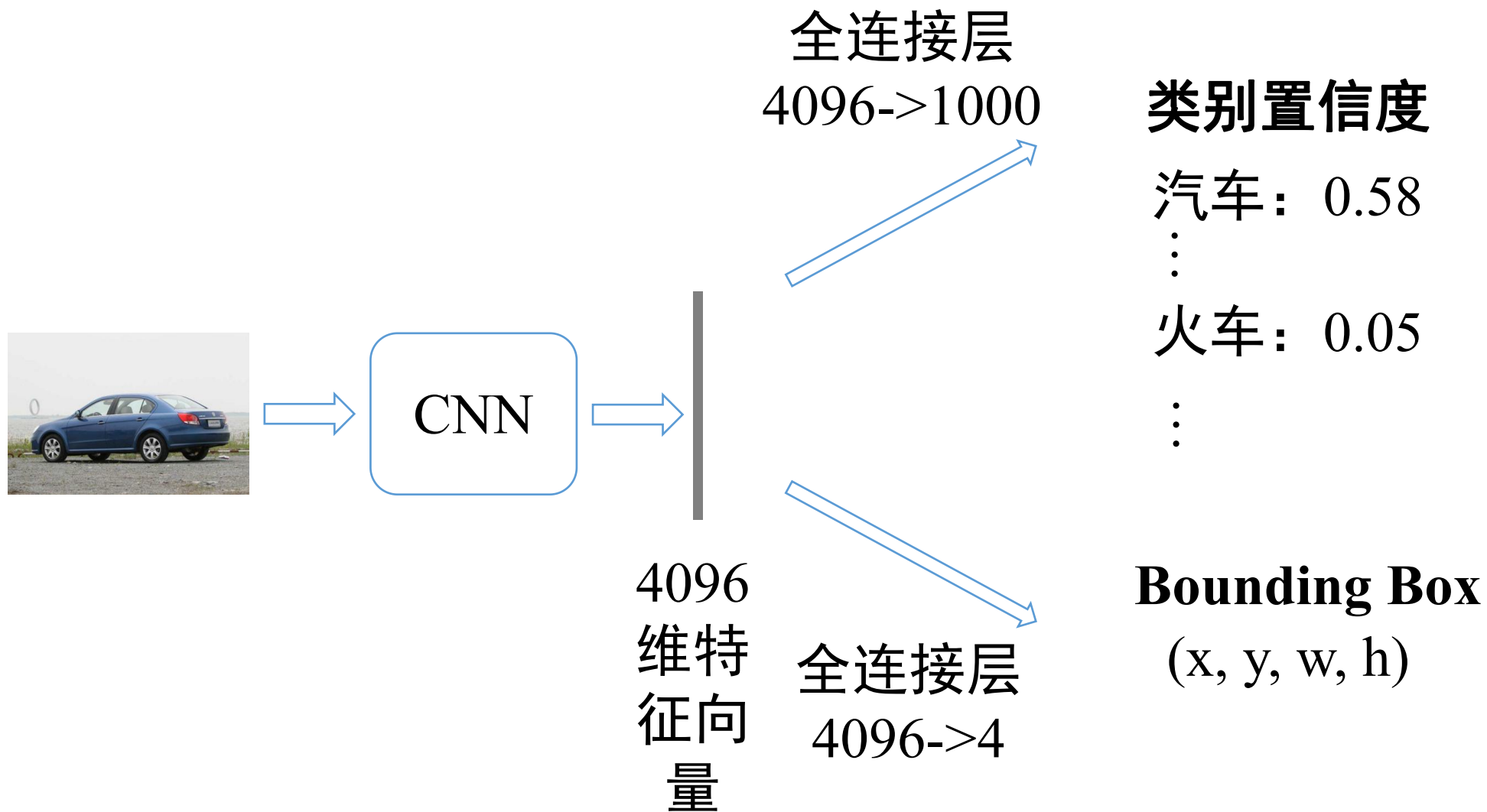


# 目标检测

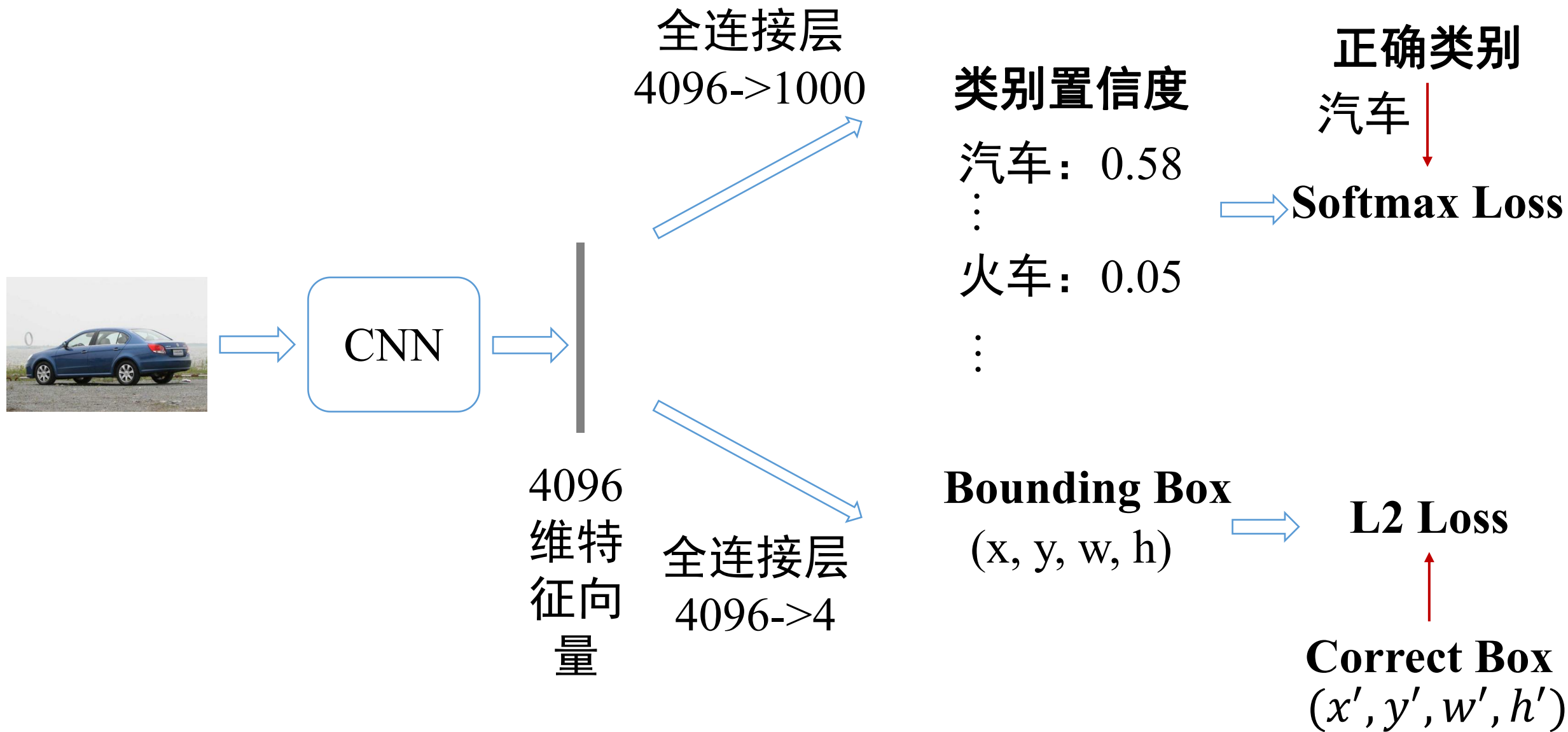


Girshick, Ross. "Fast r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2015.

# 图像分类和目标定位

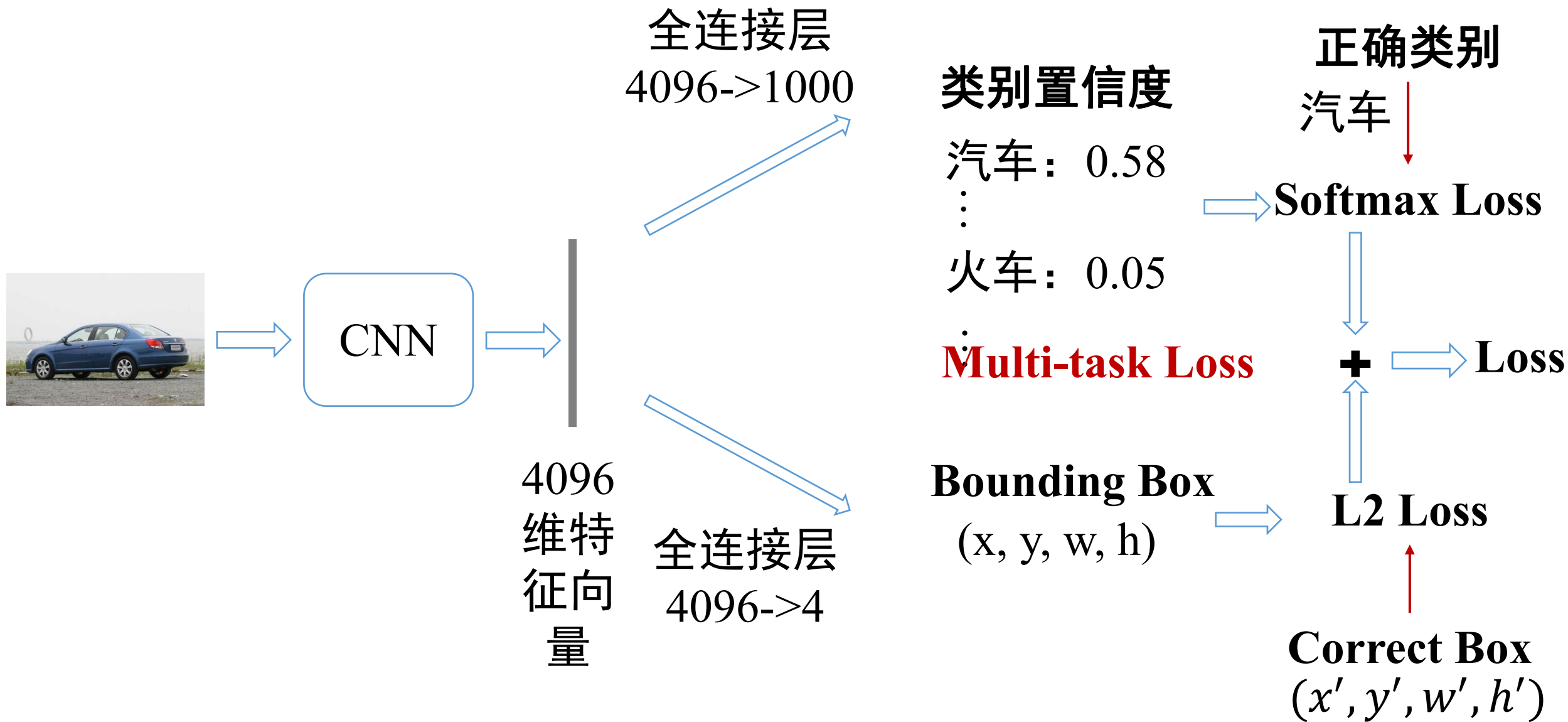


# 图像分类和目标定位





# 图像分类和目标定位



# 目标检测

- **[R-CNN]** Girshick Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580-587. 2014.
- Girshick, Ross. "**Fast r-cnn**." *Proceedings of the IEEE international conference on computer vision*. 2015.
- Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "**Faster r-cnn**: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems* 28 (2015).

# 图像超分辨率重建

图像分辨率是一组用于评估图像中蕴含细节信息丰富程度的性能参数，包括时间分辨率、空间分辨率及色阶分辨率等，体现了成像系统实际所能反映物体细节信息的能力。相较于低分辨率图像，高分辨率图像通常包含更大的像素密度、更丰富的纹理细节及更高的可信赖度。但在实际上情况中，受采集设备与环境、网络传输介质与带宽、图像退化模型本身等诸多因素的约束，我们通常并不能直接得到具有边缘锐化、无成块模糊的理想高分辨率图像。



部分图片和文字来自网络: [https://blog.51cto.com/u\\_15404184/4323223](https://blog.51cto.com/u_15404184/4323223)

# 图像超分辨率重建

图像的超分辨率重建技术指的是将给定的低分辨率图像通过特定的算法恢复成相应的高分辨率图像。具体来说，图像超分辨率重建技术指的是利用数字图像处理、计算机视觉等领域的相关知识，借由特定的算法和处理流程，从给定的低分辨率图像中重建出高分辨率图像的过程。



# 图像超分辨率重建

## 基于深度学习的超分辨率重构算法：

1. **[SRCNN]** Dong C, Loy C C, He K, et al. Learning a deep convolutional network for image super-resolution[C]//European conference on computer vision. Springer, Cham, 2014: 184-199.
2. **[VDSR]** Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1646-1654.
3. **[SRGAN]** Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.



# 图像生成

生成模型是一种训练模型进行无监督学习的模型，给模型一组数据，希望从数据中学习到信息后的模型能够生成一组和训练集尽可能相近的数据。图片生成模型就是更具体的指向说给模型一组图片作为训练集让模型进行学习，希望模型生成一组和训练集图片尽可能相近的图片。



利用DCGAN生成动漫人脸实例，分别为0，100，200次迭代结果。

部分图片和文字来自网络：[https://blog.csdn.net/like\\_red/article/details/79192374](https://blog.csdn.net/like_red/article/details/79192374)

[https://blog.csdn.net/qq\\_34739497/article/details/79902356](https://blog.csdn.net/qq_34739497/article/details/79902356)

# 图像生成

## 图像生成经典算法：

1. **[GAN]**Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.
2. **[CGAN]**Mirza M, Osindero S. Conditional generative adversarial nets[J]. arXiv preprint arXiv:1411.1784, 2014.
3. **[DCGAN]**Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.

# 报名链接

## \* 1. 选择课程报告题目

☐ 图像超分

☐ 图像生成

☐ 图像分类

☐ 目标检测

## \* 2. 组长

## \* 3. 组员

## \* 4. 组员



问卷  
外观





# 谢谢!