

人工智能原理

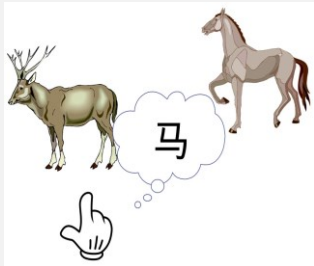


第四章 卷积神经网络

- 4.1 基于手工特征的图像分类
- 4.2 卷积神经网络的发展历程
- 4.3 卷积的本质
- 4.4 卷积神经网络的结构
- 4.5 卷积层
- 4.6 非线性激活层
- 4.7 池化层
- 4.8 全连接层
- 4.9 CNN网络的训练
- 4.10 经典的卷积神经网络
- 习题

4.1 基于手工特征的图像分类

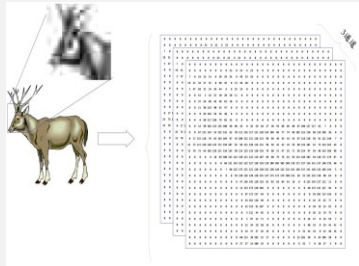
第四章 卷积神经网络



赵高的“指鹿为马”

4.1 基于手工特征的图像分类

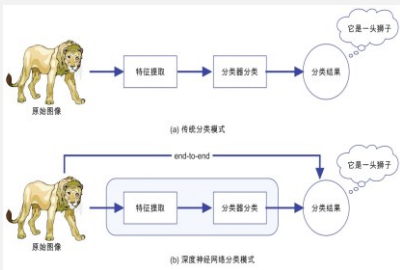
第四章 卷积神经网络



计算机内的图片表示

4.1 基于手工特征的图像分类

第四章 卷积神经网络



传统模式分类与深度神经网络的区别与联系

第四章 卷积神经网络

- 4.1 基于手工特征的图像分类
- 4.2 卷积神经网络的发展历程
- 4.3 卷积的本质
- 4.4 卷积神经网络的结构
- 4.5 卷积层
- 4.6 非线性激活层
- 4.7 池化层
- 4.8 全连接层
- 4.9 CNN网络的训练
- 4.10 经典的卷积神经网络
- 习题

4.2 卷积神经网络的发展历程

第四章 卷积神经网络

特征层次

高级特征

中级特征

低级特征

由低级特征向上抽象形成高级特征

David Hunter Hubel

4.2 卷积神经网络的发展历程

第四章 卷积神经网络

Yann LeCun

1989: CNN

LeNet 5

answer: 0

LeNet的网络架构 (7层)

第四章 卷积神经网络

4.1 基于手工特征的图像分类

4.2 卷积神经网络的发展历程

4.3 卷积的本质

4.4 卷积神经网络的结构

4.5 卷积层

4.6 非线性激活层

4.7 池化层

4.8 全连接层

4.9 CNN网络的训练

4.10 经典的卷积神经网络

习题

4.3 卷积的本质

第四章 卷积神经网络

Kernel	Image	Output																			
<table><tr><td>0</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>5</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>0</td></tr></table>	0	-1	0	-1	5	-1	0	-1	0	<table><tr><td>2</td><td>2</td><td>2</td></tr><tr><td>2</td><td>3</td><td>2</td></tr><tr><td>2</td><td>2</td><td>2</td></tr></table>	2	2	2	2	3	2	2	2	2	<table><tr><td>7</td></tr></table>	7
0	-1	0																			
-1	5	-1																			
0	-1	0																			
2	2	2																			
2	3	2																			
2	2	2																			
7																					

Kernel	Image	Output																			
<table><tr><td>0</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>5</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>0</td></tr></table>	0	-1	0	-1	5	-1	0	-1	0	<table><tr><td>2</td><td>2</td><td>2</td></tr><tr><td>2</td><td>1</td><td>2</td></tr><tr><td>2</td><td>2</td><td>2</td></tr></table>	2	2	2	2	1	2	2	2	2	<table><tr><td>-3</td></tr></table>	-3
0	-1	0																			
-1	5	-1																			
0	-1	0																			
2	2	2																			
2	1	2																			
2	2	2																			
-3																					

4.3 卷积的本质

第四章 卷积神经网络

1D卷积运算

(b)

短向量: 1 2 3

长向量: 4 5 6 7 8

卷积结果: 32 38

(c)

向量: 1 2 3

卷积结果: 32 38 44

4.3 卷积的本质

第四章 卷积神经网络

8

猫

计算机“眼中”的图像

4.3 卷积的本质

第四章 卷积神经网络

11100

01101

00111

00110

01100

卷积核

输入数据

101

010

101

卷积核

简化版本的图像矩阵和卷积核

4.3 卷积的本质

第四章 卷积神经网络

11100

01101

00111

00110

01100

Image

4

Convolved Feature

动态2D卷积

4.3 卷积的本质

第四章 卷积神经网络

原始图片

操作	卷积核 (滤波器)	卷积之后的特征图
边缘检测 (Edge detection)	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
锐化 (Sharpen)	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	

常用的图像卷积核

4.3 卷积的本质

第四章 卷积神经网络

Box blur (normalized)

$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Gaussian blur 3 x 3 (approximation)

$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$

Gaussian blur 5 x 5 (approximation)

$\frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$

常用的图像卷积核

4.3 卷积的本质

第四章 卷积神经网络

```
[9] dog_image = np.array(Image.open('./dog.jpg').convert('L'))
plt.imshow(dog_image, cmap='gray')
```

<matplotlib.image.AxesImage at 0x7f07e6bf88d0>

4.3 卷积的本质

第四章 卷积神经网络

横向 edge detection

kernel = np.array([[1, 0, 1], [-2, 0, 2], [-1, 0, 1]], np.float32)

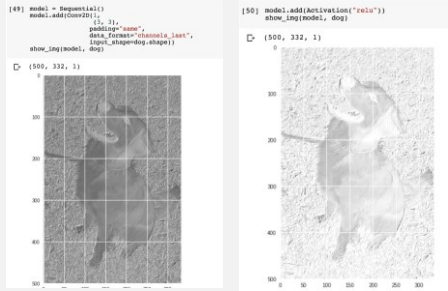
dx = show_difference(dog_image, kernel)

Original image

Convolved image

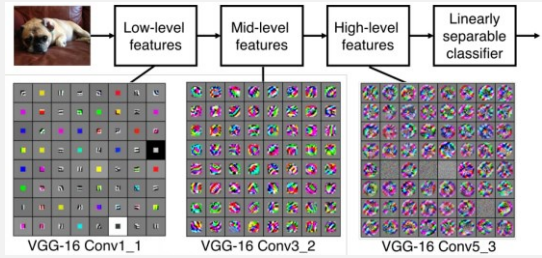
4.3 卷积的本质

第四章 卷积神经网络



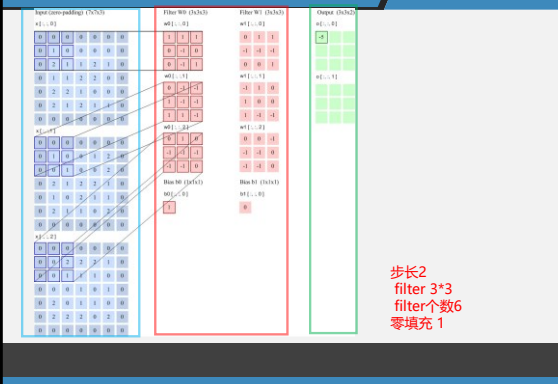
4.3 卷积的本质

第四章 卷积神经网络



4.3 卷积的本质

第四章 卷积神经网络



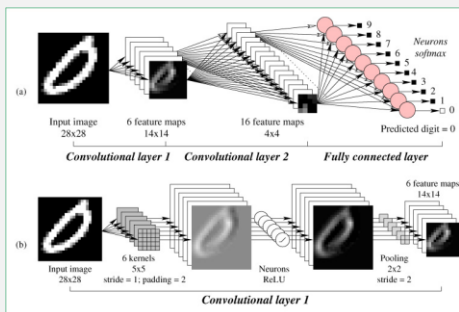
新工科信息技术基础系列规划教材

第四章 卷积神经网络

- 4.1 基于手工特征的图像分类
- 4.2 卷积神经网络的发展历程
- 4.3 卷积的本质
- 4.4 卷积神经网络的结构
- 4.5 卷积层
- 4.6 非线性激活层
- 4.7 池化层
- 4.8 全连接层
- 4.9 CNN网络的训练
- 4.10 经典的卷积神经网络
- 习题

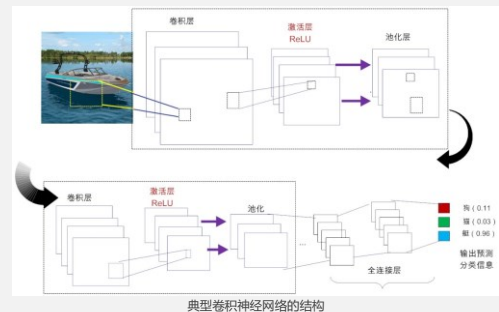
4.4 卷积神经网络的结构

第四章 卷积神经网络



4.4 卷积神经网络的结构

第四章 卷积神经网络

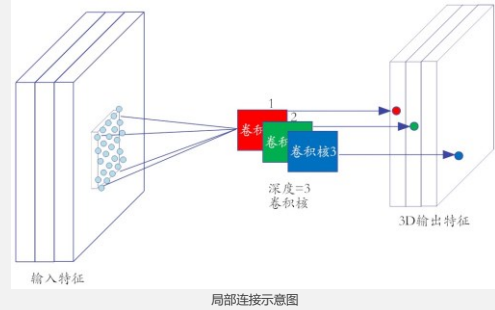


第四章 卷积神经网络

- 4.1 基于手工特征的图像分类
- 4.2 卷积神经网络的发展历程
- 4.3 卷积的本质
- 4.4 卷积神经网络的结构
- 4.5 卷积层**
- 4.6 非线性激活层
- 4.7 池化层
- 4.8 全连接层
- 4.9 CNN网络的训练
- 4.10 经典的卷积神经网络
- 习题

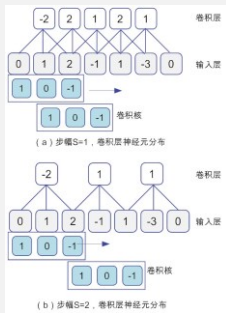
4.5 卷积层

第四章 卷积神经网络



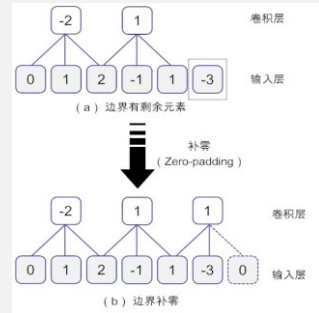
4.5 卷积层

第四章 卷积神经网络



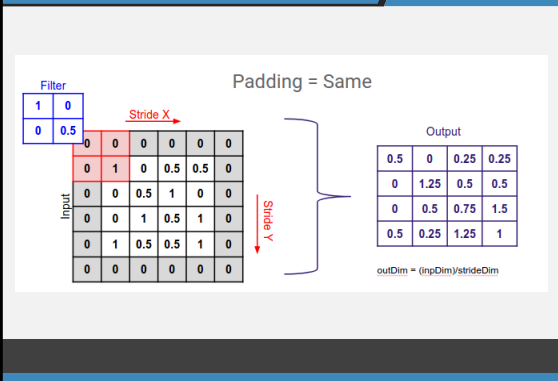
4.5 卷积层

第四章 卷积神经网络



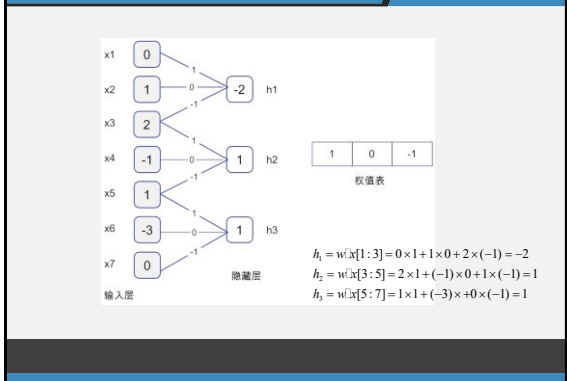
4.5 卷积层

第四章 卷积神经网络



4.5 卷积层

第四章 卷积神经网络



第四章 卷积神经网络

- 4.1 基于手工特征的图像分类
- 4.2 卷积神经网络的发展历程
- 4.3 卷积的本质
- 4.4 卷积神经网络的结构
- 4.5 卷积层
- 4.6 非线性激活层**
- 4.7 池化层
- 4.8 全连接层
- 4.9 CNN网络的训练
- 4.10 经典的卷积神经网络
- 习题

4.6 非线性激活层

第四章 卷积神经网络

Activation Functions

Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



Leaky ReLU

$$\max(0.1x, x)$$



tanh

$$\tanh(x)$$



Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ReLU

$$\max(0, x)$$



ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



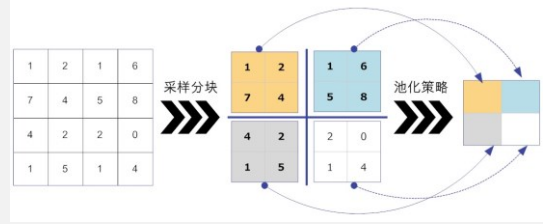
激活层：非线性变换 提升神经网络的表达能力

第四章 卷积神经网络

- 4.1 基于手工特征的图像分类
- 4.2 卷积神经网络的发展历程
- 4.3 卷积的本质
- 4.4 卷积神经网络的结构
- 4.5 卷积层
- 4.6 非线性激活层
- 4.7 池化层**
- 4.8 全连接层
- 4.9 CNN网络的训练
- 4.10 经典的卷积神经网络
- 习题

4.7 池化层

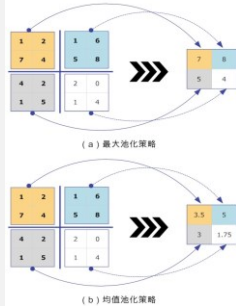
第四章 卷积神经网络



采样的本质就是力图以合理的方式“以偏概全”

4.7 池化层

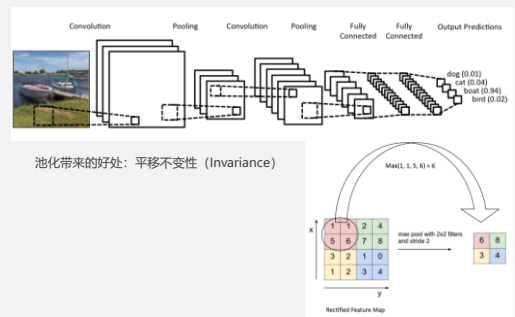
第四章 卷积神经网络



最大池化和均值池化策略对比图

4.7 池化层

第四章 卷积神经网络



池化带来的好处：平移不变性 (Invariance)

4.7 池化层

第四章 卷积神经网络

(a) 池化前的特征图谱

池化后的特征图谱变化

(b) 池化后的特征图谱

第四章 卷积神经网络

4.1 基于手工特征的图像分类

4.2 卷积神经网络的发展历程

4.3 卷积的本质

4.4 卷积神经网络的结构

4.5 卷积层

4.6 非线性激活层

4.7 池化层

4.8 全连接层

4.9 CNN网络的训练

4.10 经典的卷积神经网络

习题

4.8 全连接层

第四章 卷积神经网络

全连接层 (Dense Layer) 目的：在于分类/回归等。
前面若干层（卷积、激活、池化等）目的：数据变换-->提取特征，作为全连接层的输入，为全连接层服务。

4.8 全连接层

第四章 卷积神经网络

Softmax 输出

$$\text{softmax}(z_j) = \frac{e^{z_j}}{\sum_j e^{z_j}} \quad (j = 1, 2, 3, \dots, k)$$

第四章 卷积神经网络

4.1 基于手工特征的图像分类

4.2 卷积神经网络的发展历程

4.3 卷积的本质

4.4 卷积神经网络的结构

4.5 卷积层

4.6 非线性激活层

4.7 池化层

4.8 全连接层

4.9 CNN网络的训练

4.10 经典的卷积神经网络

习题

4.9 CNN网络的训练

第四章 卷积神经网络

CNN的反向传播 (BP) 过程

第四章 卷积神经网络

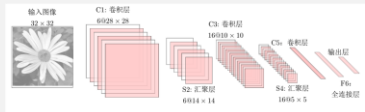
- 4.1 基于手工特征的图像分类
- 4.2 卷积神经网络的发展历程
- 4.3 卷积的本质
- 4.4 卷积神经网络的结构
- 4.5 卷积层
- 4.6 非线性激活层
- 4.7 池化层
- 4.8 全连接层
- 4.9 CNN网络的训练
- 4.10 经典的卷积神经网络

习题

4.10 经典的卷积神经网络

第四章 卷积神经网络

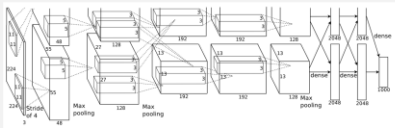
- ▶ **LeNet-5 是一个非常成功的神经网络模型。**
- ▶ 基于 LeNet-5 的手写数字识别系统在 90 年代被美国很多银行使用，用来识别支票上面的手写数字。
- ▶ LeNet-5 共有 7 层。



4.10 经典的卷积神经网络

第四章 卷积神经网络

- ▶ **2012 ILSVRC winner**
- ▶ (top 5 error of 16% compared to runner-up with 26% error)
- ▶ 第一个现代深度卷积网络模型
- ▶ 首次使用了很多现代深度卷积网络的一些技术方法
- ▶ 使用 GPU 进行并行训练，采用了 ReLU 作为非线性激活函数，使用 Dropout 防止过拟合，使用数据增强
- ▶ 5 个卷积层、3 个汇聚层和 3 个全连接层



4.10 经典的卷积神经网络

第四章 卷积神经网络

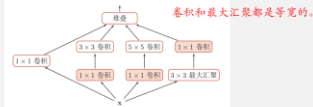
- ▶ **2014 ILSVRC winner (22 层)**
- ▶ 参数: GoogLeNet: 4M VS AlexNet: 60M
- ▶ 错误率: 6.7%
- ▶ Inception 网络是由多个 inception 模块和少量的汇聚层堆叠而成。



4.10 经典的卷积神经网络

第四章 卷积神经网络

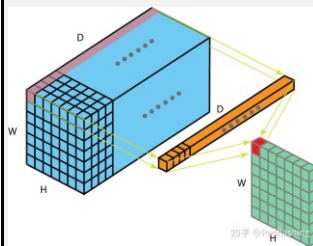
- ▶ 在卷积网络中，如何设置卷积层的卷积核大小是一个十分关键的问题。
- ▶ 在 Inception 网络中，一个卷积层包含多个不同大小的卷积操作，称为 Inception 模块。
- ▶ Inception 模块同时使用 1×1 、 3×3 、 5×5 等不同大小的卷积核，并将得到的特征映射在深度上拼接（堆叠）起来作为输出特征映射。



4.10 经典的卷积神经网络

第四章 卷积神经网络

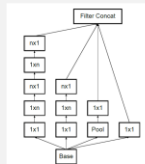
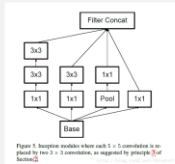
1X1 卷积核的妙用



4.10 经典的卷积神经网络

第四章 卷积神经网络

- ▶ 用多层小卷积核替换大卷积核，以减少计算量和参数量。
- ▶ 使用两层3x3的卷积来替换v1中的5x5的卷积
- ▶ 使用连续的1x1和1x1来替换nxn的卷积。



4.10 经典的卷积神经网络

第四章 卷积神经网络



4.10 经典的卷积神经网络

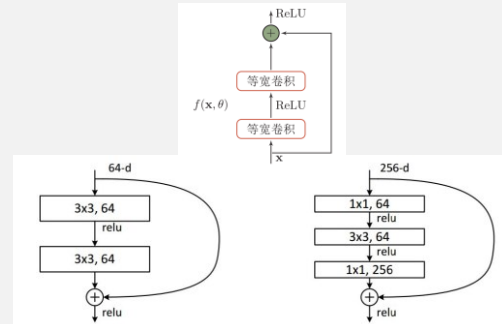
第四章 卷积神经网络

- ▶ 残差网络 (Residual Network, ResNet) 是通过给非线性的卷积层增加直连边的方式来提高信息的传播效率。
- ▶ 假设在一个深度网络中，我们期望一个非线性单元（可以为一层或多层的卷积层） $f(x, \theta)$ 去逼近一个目标函数为 $h(x)$ 。
- ▶ 将目标函数拆分成两部分：恒等函数和残差函数

$$h(x) = \underbrace{x}_{\text{恒等函数}} + \underbrace{(h(x) - x)}_{\text{残差函数}} = f(x, \theta)$$

4.10 经典的卷积神经网络

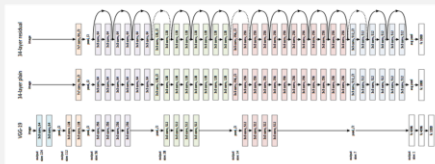
第四章 卷积神经网络



4.10 经典的卷积神经网络

第四章 卷积神经网络

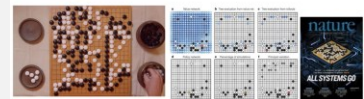
- ▶ 2015 ILSVRC winner (152层)
- ▶ 错误率: 3.57%



4.10 经典的卷积神经网络

第四章 卷积神经网络

AlphaGo



The input to the policy network is a $19 \times 19 \times 48$ image stack consisting of 48 feature planes. The first hidden layer zero pads the input into a 23×23 image, then convolves k filters of kernel size 5×5 with stride 1 with the input image and applies a rectifier nonlinearity. Each of the subsequent hidden layers 2 to 12 zero pads the respective previous hidden layer into a 21×21 image, then convolves k filters of kernel size 3×3 with stride 1, again followed by a rectifier nonlinearity. The final layer convolves 1 filter of kernel size 1×1 with stride 1, with a different bias for each position, and applies a softmax function. The match version of AlphaGo used $k = 192$ filters; Fig. 2b and Extended Data Table 3 additionally show the results of training with $k = 128, 256$ and 384 filters.

policy network:

[19x19x48] Input

CONV1: 192 5x5 filters, stride 1, pad 2 => [19x19x192]

CONV2: 12-192 3x3 filters, stride 1, pad 1 => [19x19x192]

CONV: 1 1x1 filter, stride 1, pad 0 => [19x19] (probability map of promising moves)

分布式系统: 1202 个CPU 和176 块GPU

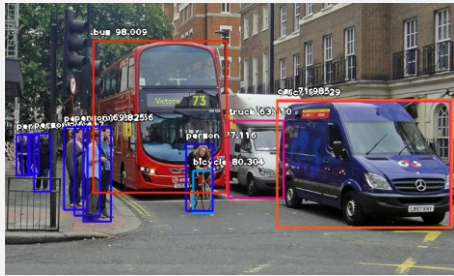
单机版: 48 个CPU 和8 块GPU

走子速度: 3 毫秒-2 微秒

4.10 经典的卷积神经网络

第四章 卷积神经网络

目标检测 (Object Detection)



4.10 经典的卷积神经网络

第四章 卷积神经网络

Mask RCNN



Figure 4. More results of Mask R-CNN on COCO test images, using ResNet-101-FPN and running at 5 fps, with 35.7 mask AP (Table 1).

4.10 经典的卷积神经网络

第四章 卷积神经网络

OCR



4.10 经典的卷积神经网络

第四章 卷积神经网络

Deep Dream



4.10 经典的卷积神经网络

第四章 卷积神经网络

画风迁移



4.10 经典的卷积神经网络

第四章 卷积神经网络

画风迁移



4.10 经典的卷积神经网络

第四章 卷积神经网络

DALL-E



4.10 经典的卷积神经网络

第四章 卷积神经网络

DALL-E

新工科信息技术基础系列规划教材

第四章 卷积神经网络

4.1 基于手工特征的图像分类

4.2 卷积神经网络的发展历程

4.3 卷积的本质

4.4 卷积神经网络的结构

4.5 卷积层

4.6 非线性激活层

4.7 池化层

4.8 全连接层

4.9 CNN网络的训练

4.10 经典的卷积神经网络

习题

习题:

1. 除了文中描述的常见卷积核, 你还知道哪些常用于图像处理的卷积核?

2. CNN的反向传播部分用的是BP算法, 而BP算法是否符合生物视觉系统的工作机制, 是存在争议的, 你知道这些争议都体现在何处?

3. 编程 (如何使用Python) 实现 (或用深度学习框架如Tensorflow、Keras等) 卷积神经网络, 并在手写字符识别数据集MNIST上进行试验测试。

感谢聆听

