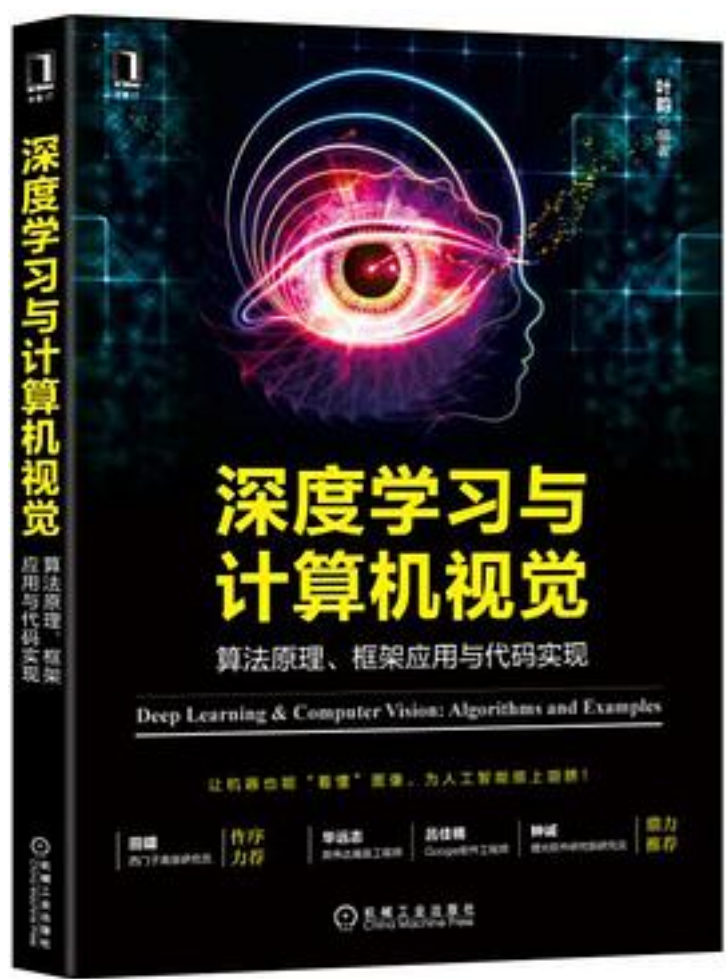


深度学习与计算机视觉



[深度学习与计算机视觉_下载链接1](#)

著者:叶韵

出版者:机械工业出版社

出版时间:2017-7-25

装帧:平装

isbn:9787111573678

全书共13章，分为2篇。第1篇基础知识，介绍了人工智能发展里程、计算机视觉概要、深度学习和计算机视觉中的基础数学知识、神经网络及其相关的机器学习基础、卷积

神经网络及其一些常见结构，最后对最前沿的趋势进行了简单探讨。第2篇实例精讲，介绍了Python基础、OpneCV基础、最简单的分类神经网络、图像识别、利用Caffe做回归、迁移学习和模型微调、目标检测、度量学习和图像风格迁移等常见的计算机视觉应用场景。从第5章开始包含了很多有趣和实用的代码示例。从第7章开始的所有实例都基于当前最流行的深度学习框架中的Caffe和MXNet，其中包含了作者原创的大量代码和搜集的数据，这些代码和作者训练好的部分模型已分享到本书github页面上供读者自行下载。

代码库地址：

https://github.com/frombeijingwithlove/dlcv_for_beginners

作者介绍：

叶韵博士，现在京东从事深度学习和计算机视觉算法研发。加入京东之前，曾在ProPlus Design Solutions硅谷和北京研发中心任职研发经理，负责统计建模算法研发，后加入Siemens Corporate Technology担任Research Scientist，专注计算影像和计算机视觉的研究。叶博士于2007年7月获得北京大学微电子学士学位，2011年4月获得Arizona State University的Electrical Engineering博士学位。

目录: 序言

前言

第1篇 基础知识

第1章 引言 2

1.1 人工智能的新焦点——深度学习 2

1.1.1 人工智能——神话传说到影视漫画 2

1.1.2 人工智能的诞生 3

1.1.3 神经科学的研究 4

1.1.4 人工神经网络的兴起 5

1.1.5 神经网络的第一次寒冬 6

1.1.6 神经网络的第一次复兴 8

1.1.7 神经网络的第二次寒冬 9

1.1.8 2006年——深度学习的起点 10

1.1.9 生活中的深度学习 11

1.1.10 常见深度学习框架简介 12

1.2 给计算机一双眼睛——计算机视觉 14

1.2.1 计算机视觉简史 14

1.2.2 2012年——计算机视觉的新起点 16

1.2.3 计算机视觉的应用 17

1.2.4 常见计算机视觉工具包 19

1.3 基于深度学习的计算机视觉 19

1.3.1 从ImageNet竞赛到AlphaGo战胜李世石——计算机视觉超越人类 19

1.3.2 GPU和并行技术——深度学习和计算视觉发展的加速器 21

1.3.3 基于卷积神经网络的计算机视觉应用 22

第2章 深度学习和计算机视觉中的基础数学知识 27

2.1 线性变换和非线性变换 27

2.1.1 线性变换的定义 27

2.1.2 高中教科书中的小例子 28

2.1.3 点积和投影 28

2.1.4 矩阵乘法的几何意义 (1) 30

- 2.1.5 本征向量和本征值 34
- 2.1.6 矩阵乘法的几何意义 (2) 37
- 2.1.7 奇异值分解 38
- 2.1.8 线性可分性和维度 39
- 2.1.9 非线性变换 42
- 2.2 概率论及相关基础知识 43
 - 2.2.1 条件概率和独立 43
 - 2.2.2 期望值、方差和协方差 44
 - 2.2.3 熵 45
 - 2.2.4 最大似然估计 (Maximum Likelihood Estimation, MLE) 47
 - 2.2.5 KL散度 (Kullback-Leibler divergence) 49
 - 2.2.6 KL散度和MLE的联系 49
- 2.3 维度的诅咒 50
 - 2.3.1 采样和维度 50
 - 2.3.2 高维空间中的体积 51
 - 2.3.3 高维空间中的距离 53
 - 2.3.4 中心极限定理和高维样本距离分布的近似 54
 - 2.3.5 数据实际的维度 56
 - 2.3.6 局部泛化 58
 - 2.3.7 函数对实际维度的影响 59
 - 2.3.8 PCA——什么是主成分 60
 - 2.3.9 PCA——通过本征向量和本征值求主成分 60
 - 2.3.10 PCA——通过主成分分析降维 61
 - 2.3.11 PCA——归一化和相关性系数 63
 - 2.3.12 PCA——什么样的数据适合PCA 64
 - 2.3.13 其他降维手段 65
- 2.4 卷积 66
 - 2.4.1 点积和卷积 66
 - 2.4.2 一维卷积 67
 - 2.4.3 卷积和互相关 68
 - 2.4.4 二维卷积和图像响应 69
 - 2.4.5 卷积的计算 70
- 2.5 数学优化基础 71
 - 2.5.1 最小值和梯度下降 72
 - 2.5.2 冲量 (Momentum) 73
 - 2.5.3 牛顿法 75
 - 2.5.4 学习率和自适应步长 77
 - 2.5.5 学习率衰减 (Learning Rate Decay) 78
 - 2.5.6 AdaGrad: 每个变量有自己的节奏 78
 - 2.5.7 AdaDelta的进一步改进 79
 - 2.5.8 其他自适应算法 80
 - 2.5.9 损失函数 81
 - 2.5.10 分类问题和负对数似然 82
 - 2.5.11 逻辑回归 83
 - 2.5.12 Softmax: 将输出转换为概率 84
 - 2.5.13 链式求导法则 84
- 第3章 神经网络和机器学习基础 87
 - 3.1 感知机 87
 - 3.1.1 基本概念 87
 - 3.1.2 感知机和线性二分类 87
 - 3.1.3 激活函数 88
 - 3.2 神经网络基础 89
 - 3.2.1 从感知机到神经网络 89
 - 3.2.2 最简单的神经网络二分类例子 90

- 3.2.3 隐层神经元数量的作用 93
- 3.2.4 更加复杂的样本和更复杂的神经网络 94
- 3.3 后向传播算法 95
 - 3.3.1 求神经网络参数的梯度 95
 - 3.3.2 计算图 (Computational Graph) 95
 - 3.3.3 利用后向传播算法计算一个神经网络参数的梯度 97
 - 3.3.4 梯度消失 99
 - 3.3.5 修正线性单元 (ReLU) 100
 - 3.3.6 梯度爆炸 101
 - 3.3.7 梯度检查 (gradient check) 102
 - 3.3.8 从信息传播的角度看后向传播算法 103
- 3.4 随机梯度下降和批量梯度下降 104
 - 3.4.1 全量数据 (full-batch) 梯度下降 104
 - 3.4.2 随机梯度下降 (SGD) 和小批量数据 (mini-batch) 104
 - 3.4.3 数据均衡和数据增加 (data augmentation) 106
- 3.5 数据、训练策略和规范化 108
 - 3.5.1 欠拟合和过拟合 108
 - 3.5.2 训练误差和测试误差 109
 - 3.5.3 奥卡姆剃刀没有免费午餐 111
 - 3.5.4 数据集划分和提前停止 112
 - 3.5.5 病态问题和约束 113
 - 3.5.6 L2规范化 (L2 Regularization) 113
 - 3.5.7 L1规范化 (L1 Regularization) 114
 - 3.5.8 集成 (Ensemble) 和随机失活 (Dropout) 115
- 3.6 监督学习、非监督学习、半监督学习和强化学习 117
 - 3.6.1 监督学习、非监督学习和半监督学习 117
 - 3.6.2 强化学习 (reinforcement learning) 118
- 第4章 深度卷积神经网络 120
 - 4.1 卷积神经网络 120
 - 4.1.1 基本概念 120
 - 4.1.2 卷积层和特征响应图 121
 - 4.1.3 参数共享 123
 - 4.1.4 稀疏连接 124
 - 4.1.5 多通道卷积 125
 - 4.1.6 激活函数 125
 - 4.1.7 池化、不变性和感受野 126
 - 4.1.8 分布式表征 (Distributed Representation) 128
 - 4.1.9 分布式表征和局部泛化 130
 - 4.1.10 分层表达 131
 - 4.1.11 卷积神经网络结构 131
 - 4.2 LeNet——第一个卷积神经网络 132
 - 4.3 新起点——AlexNet 133
 - 4.3.1 网络结构 133
 - 4.3.2 局部响应归一化 (Local Response Normalization, LRN) 136
 - 4.4 更深的网络——GoogLeNet 136
 - 4.4.1 1×1 卷积和Network In Network 136
 - 4.4.2 Inception结构 138
 - 4.4.3 网络结构 138
 - 4.4.4 批归一化 (Batch Normalization, BN) 140
 - 4.5 更深的网络——ResNet 142
 - 4.5.1 困难的深层网络训练：退化问题 142
 - 4.5.2 残差单元 142
 - 4.5.3 深度残差网络 144
 - 4.5.4 从集成的角度看待ResNet 144

- 4.5.5 结构更复杂的网络 146
- 第2篇 实例精讲
- 第5章 Python基础 148
 - 5.1 Python简介 148
 - 5.1.1 Python简史 148
 - 5.1.2 安装和使用Python 149
 - 5.2 Python基本语法 150
 - 5.2.1 基本数据类型和运算 150
 - 5.2.2 容器 153
 - 5.2.3 分支和循环 156
 - 5.2.4 函数、生成器和类 159
 - 5.2.5 map、reduce和filter 162
 - 5.2.6 列表生成（list comprehension） 163
 - 5.2.7 字符串 163
 - 5.2.8 文件操作和pickle 164
 - 5.2.9 异常 165
 - 5.2.10 多进程（multiprocessing） 165
 - 5.2.11 os模块 166
 - 5.3 Python的科学计算包——NumPy 167
 - 5.3.1 基本类型（array） 167
 - 5.3.2 线性代数模块（linalg） 172
 - 5.3.3 随机模块（random） 173
 - 5.4 Python的可视化包——matplotlib 175
 - 5.4.1 2D图表 175
 - 5.4.2 3D图表 178
 - 5.4.3 图像显示 180
- 第6章 OpenCV基础 182
 - 6.1 OpenCV简介 182
 - 6.1.1 OpenCV的结构 182
 - 6.1.2 安装和使用OpenCV 183
 - 6.2 Python-OpenCV基础 184
 - 6.2.1 图像的表达 184
 - 6.2.2 基本图像处理 185
 - 6.2.3 图像的仿射变换 188
 - 6.2.4 基本绘图 190
 - 6.2.5 视频功能 192
 - 6.3 用OpenCV实现数据增加小工具 193
 - 6.3.1 随机裁剪 194
 - 6.3.2 随机旋转 194
 - 6.3.3 随机颜色和明暗 196
 - 6.3.4 多进程调用加速处理 196
 - 6.3.5 代码：图片数据增加小工具 196
 - 6.4 用OpenCV实现物体标注小工具 203
 - 6.4.1 窗口循环 203
 - 6.4.2 鼠标和键盘事件 205
 - 6.4.3 代码：物体检测标注的小工具 206
- 第7章 Hello World! 212
 - 7.1 用MXNet实现一个神经网络 212
 - 7.1.1 基础工具、NVIDIA驱动和CUDA安装 212
 - 7.1.2 安装MXNet 213
 - 7.1.3 MXNet基本使用 214
 - 7.1.4 用MXNet实现一个两层神经网络 215
 - 7.2 用Caffe实现一个神经网络 219
 - 7.2.1 安装Caffe 219

7.2.2 Caffe的基本概念	220
7.2.3 用Caffe实现一个两层神经网络	221
第8章 最简单的图片分类——手写数字识别	227
8.1 准备数据——MNIST	227
8.1.1 下载MNIST	227
8.1.2 生成MNIST的图片	227
8.2 基于Caffe的实现	228
8.2.1 制作LMDB数据	229
8.2.2 训练LeNet-5	230
8.2.3 测试和评估	235
8.2.4 识别手写数字	239
8.2.5 增加平移和旋转扰动	240
8.3 基于MXNet的实现	242
8.3.1 制作Image Recordio数据	242
8.3.2 用Module模块训练LeNet-5	243
8.3.3 测试和评估	245
8.3.4 识别手写数字	247
第9章 利用Caffe做回归	249
9.1 回归的原理	249
9.1.1 预测值和标签值的欧式距离	249
9.1.2 EuclideanLoss层	250
9.2 预测随机噪声的频率	250
9.2.1 生成样本：随机噪声	250
9.2.2 制作多标签HDF5数据	252
9.2.3 网络结构和Solver定义	253
9.2.4 训练网络	259
9.2.5 批量装载图片并利用GPU预测	260
9.2.6 卷积核可视化	262
第10章 迁移学习和模型微调	264
10.1 吃货必备——通过Python采集美食图片	264
10.1.1 通过关键词和图片搜索引擎下载图片	264
10.1.2 数据预处理——去除无效和不相关图片	267
10.1.3 数据预处理——去除重复图片	267
10.1.4 生成训练数据	269
10.2 美食分类模型	271
10.2.1 迁移学习	271
10.2.2 模型微调法（Finetune）	272
10.2.3 混淆矩阵（Confusion Matrix）	276
10.2.4 P-R曲线和ROC曲线	278
10.2.5 全局平均池化和激活响应图	284
第11章 目标检测	288
11.1 目标检测算法简介	288
11.1.1 滑窗法	288
11.1.2 PASCAL VOC、mAP和IOU简介	289
11.1.3 Selective Search和R-CNN简介	290
11.1.4 SPP、ROI Pooling和Fast R-CNN简介	291
11.1.5 RPN和Faster R-CNN简介	293
11.1.6 YOLO和SSD简介	294
11.2 基于PASCAL VOC数据集训练SSD模型	296
11.2.1 MXNet的SSD实现	296
11.2.2 下载PASCAL VOC数据集	297
11.2.3 训练SSD模型	298
11.2.4 测试和评估模型效果	299
11.2.5 物体检测结果可视化	299

11.2.6 制作自己的标注数据 302
第12章 度量学习 304
12.1 距离和度量学习 304
12.1.1 欧氏距离和马氏距离 304
12.1.2 欧式距离和余弦距离 305
12.1.3 非线性度量学习和Siamese网络 306
12.1.4 Contrastive Loss：对比损失函数 307
12.2 用MNIST训练Siamese网络 307
12.2.1 数据准备 307
12.2.2 参数共享训练 309
12.2.3 结果和可视化 314
12.2.4 用t-SNE可视化高维特征 316
第13章 图像风格迁移 317
13.1 风格迁移算法简介 317
13.1.1 通过梯度下降法进行图像重建 317
13.1.2 图像风格重建和Gram矩阵 318
13.1.3 图像风格迁移 320
13.2 MXNet中的图像风格迁移例子 320
13.2.1 MXNet的风格迁移实现 321
13.2.2 对图片进行风格迁移 326
• • • • • ([收起](#))

[深度学习与计算机视觉 下载链接1](#)

标签

深度学习

计算机视觉

计算机

数字图像处理/计算机视觉

机器学习/深度学习/TensorFlow/Caffe

人工智能

计算机科学

机器学习

评论

印刷的小错误不少，但是确实是本好书

很好的入门书。最近又重新刷了一遍，感觉作者对理论的理解还是很好的。

三天读完，可以说前面部分理论讲的很好，但是后面实践部分就有点弱了，示例代码也不是很好

入门，详细有条理

也许这种书不该在豆瓣发，不过还是推荐一下，后面实例部分有点不好，前面理论讲解很不错

多个主题的简单提及，深度还差一些

工作需要，一周多读完。还不错，通过此书caffe算是入门了。网上竟然很难找到跑得通的caffe预测的demo，这本书上的例子虽然简单但还是挺实用。

亚马逊推荐的书籍。

前三章由浅入深，紧扣主题，并非面面俱到不冗余，有独特见解；第四章跳跃比较大，阅读难度大，还是得找相关的论文。

同事写的书，支持一下了。。。凭良心说，还是不错的

入门，一个知识框架

算是很给力的书了。现在关于深度学习的书很多，其中灌水的居多。这本书讲了很多很多有用的tricks以及一些dl基础理论，感觉很不错

理论部分讲解的挺好,电子版

我竟然发现了尼古拉斯赵四隐藏彩蛋……

非常浅显易懂的深度学习入门书，先复习了一遍高数再讲算法。可惜的是代码实例有些过时了，用的是python2和caffe。

尽管有些印刷错误，比起国内灌水的那些书要强很多。作者有自己的见解，这点弥足珍贵，计算机视觉相关技术也基本覆盖（但也缺少OCR，语意分割，等这类经典场景）。希望国内多出些此类的书，而不是毫无辨别的互相copy。

开始先大致是梳理了一遍该领域的发展史。而后从所需基础数学知识开始逐渐展开。比较偏重于背后原理的理解吧。作为入门课程的补充参考书还是很不错的。

这本书的理论部分是我读过的最言简意赅的版本，一句废话都没有，数理部分深度刚好合适，很好理解，建议作为Reference收藏。
可以说是深度学习中文书籍中的翘楚了。

讲解得很详细易懂

这是一本融入作者自己思维沉淀的书，不是那种官方教程翻译书所能比的。

[深度学习与计算机视觉_下载链接1](#)

书评

一提到深度学习，言必称花书，但花书微言大义，对于基础及概念，涉及较少，内容艰深，不适合新手理解基本概念。但是这本书很好地弥补了花书的不足。作者从数学基础讲起，对于部分概念配合图片说明深入浅出，化抽象于具体，很好。同时作者的逻辑清晰，条理性强，由浅入深循序渐...

从书中其实可以看出作者对于机器学习理解的深厚功力，比那些到处在网上拼凑内容的书强多了，但是感觉干活还可以再多点，比如python的介绍居然用了一章，我想介绍python是没必要的，还有就是排版了，代码的排版太惨不忍睹了简直没法看，其他的都还比较好吧，书的内容和里面推荐...

[深度学习与计算机视觉_下载链接1](#)