



深度学习与 计算机视觉

算法原理、框架应用与代码实现

Deep Learning & Computer Vision: Algorithms and Examples

让机器也能“看懂”图像，为人工智能插上翅膀！

田疆

西门子高级研究员

作序
力荐

华远志

英伟达高级工程师

吕佳楠

Google软件工程师

钟诚

理光软件研究院研究员

鼎力
推荐



机械工业出版社
China Machine Press

深度学习与 计算机视觉

算法原理、框架应用与代码实现

Deep Learning & Computer Vision: Algorithms and Examples

叶韵◎编著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

深度学习与计算机视觉：算法原理、框架应用与代码实现 / 叶韵编著. —北京：机械工业出版社，2017.6

ISBN 978-7-111-57367-8

I. 深… II. 叶… III. 计算机视觉—研究 IV. TP302.7

中国版本图书馆 CIP 数据核字 (2017) 第 165473 号

深度学习与计算机视觉

算法原理、框架应用与代码实现

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）

责任编辑：欧振旭

责任校对：姚志娟

印刷：中国电影出版社印刷厂

版次：2017 年 8 月第 1 版第 1 次印刷

开本：185mm×260mm 1/16

印张：21.5

书号：ISBN 978-7-111-57367-8

定价：79.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88379426 88361066

投稿热线：(010) 88379604

购书热线：(010) 68326294 88379649 68995259

读者信箱：hzit@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问：北京大成律师事务所 韩光/邹晓东

注重原理和上手实战，让读者不仅能理解算法背后的思想，还能具备独立开发基于深度学习的计算机视觉算法的能力。

原理讲解通俗易懂，能通过图文定性讲解的就尽量不用公式，不可避免要用公式的地方尽量让公式作为图文讲解的辅助手段。

结合常见的应用场景，通过大量有趣、实用的实例和原创代码，带领读者一步步亲自动手，不断提高动手能力。

内容简介

本书全面介绍了深度学习及计算机视觉中最基础的知识，并结合最常见的应用场景和大量实例，带领读者进入丰富多彩的计算机视觉领域。作为一本“原理+实践”教程，本书在讲解原理的基础上，通过有趣的实例带领读者一步步亲自动手，不断提高动手能力，而不是枯燥和深奥原理的堆砌。

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历程、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。本书从第5章开始包含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流行的深度学习框架Caffe和MXNet，其中包含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。

本书适合对人工智能、机器学习、深度学习和计算机视觉感兴趣的读者阅读。阅读本书要求读者具备一定的数学基础和基本的编程能力，并需要读者了解Linux的基本使用。

作者简介

叶韵

2007年7月毕业于北京大学信息科学技术学院，获学士学位。2011年4月获得了美国亚利桑那州立大学的电气工程博士学位。拥有超过5年的机器学习研发经验。目前在京东专注于深度学习和计算机视觉算法的研发。加入京东前，曾先后在ProPlus Design Solutions硅谷和北京研发中心任职研发经理，负责统计建模和机器学习算法的研发。后加入西门子中国研究院担任研究员，专注于计算影像和计算机视觉的研究。

欢迎IT领域的各位技术牛人洽谈出书事宜。如果有写书或投稿意向，请加QQ或者微信具体商谈。

QQ: 627173439

微信: oyzx_sp

HZBOOKS | 华章IT | Information Technology



序言

深度学习是机器学习的一个重要分支，它以简化的方式模拟人脑复杂的神经系统，从而达到对数据的高级抽象。近些年，深度学习在语音识别、计算机视觉、自然语言处理、生成网络和无监督学习等领域都有着广泛的应用，从很多方面改变着人们的日常生活。

互联网巨头谷歌、Facebook、亚马逊、微软、百度、阿里巴巴和腾讯等公司都建立了相应的深度学习部门和平台。随着近几年深度学习的快速发展，相继出现了大量的开源软件平台，如 Caffe、MXNet、TensorFlow 和 Torch 等。这些平台多数都有相应的 Python 和 C++ 接口，功能非常强大。但是对于初学者来说，还是有一定的门槛。

本书架起了一座初学者和开源深度学习软件之间的桥梁，致力于帮助初学者进入机器学习特别是深度学习在计算机视觉中的应用等领域。本书涵盖了基础的数学、机器学习和图像识别等内容，同时对两个主流的开源深度学习库 Caffe 和 MXNet 都有大量的实战例子描述分类和回归等问题。

本书作者在深度学习领域有着深入的研究，善于把复杂的问题用浅显易懂的语言描述出来，使得本书内容引人入胜。本书结构合理，内容涵盖了计算机视觉领域的一些主要问题。对于一个学习计算机视觉的新手来说，本书的数学推导浅显易懂，从一些简单例子开始，然后推广到抽象的矩阵描述方式，大大减轻了学习负担。读者可以通过前 7 章的学习，对神经网络、基础的数学和编程技巧有一个全面的了解。在此基础上可以根据具体的问题参考本书中具体的章节，例如图像识别、回归和目标检测等。在这些章节中，读者可以按照书中的步骤搭建自己的应用。

田疆

西门子高级研究员

附英文原文

Deep learning is a branch of machine learning based on a set of algorithms that attempt to model high level abstractions in data, and mimics human brain's neural networks in a simplistic manner. It is powerful and driving advances in speech recognition, computer vision, natural language processing, generative networks, and unsupervised machine learning in recent years, which are changing our daily life from different aspects.

Internet giants Google, Facebook, Amazon, Microsoft, Baidu, Alibaba, and Tencent launch their deep learning teams and platforms. With the rapidly growing related research and development community, in recent years, there appear plenty of open source frameworks for deep learning, such as: Caffe, MXNet, TensorFlow, Torch, etc. These frameworks can usually

interface with Python, Lua, and C++. They are powerful and usually well maintained, however, they are not that straightforward and "easy cooking recipes" for beginners.

This book bridges the gap with the spirit of open source. It is a resource intended to help students and practitioners enter the field of machine learning in general and deep learning (computer vision) in particular. This book not only covers basic theory on math, machine learning, image recognition, but also helps deep learning practitioner on two major libraries Caffe and MXNet by plenty of examples on classification, regression, metric learning, etc. The tutorials in this book also provide step-by-step instructions for creating models for specific types of applications.

The writing of this book is good and also interesting, which shows the author's delving very deep into this field. The structure of this book is quite reasonable. Its topics cover the major tasks in computer vision. The mathematics derivation in this book is solid and understandable for a computer vision novice, it starts from simple examples, and then to more abstract matrix counterparts, which greatly reduces the "activation threshold" for a learning curve.

Readers of this book may start the first seven chapters to get a comprehensive understanding of neural network, basic math and programming techniques. Thereafter, he may pick up the related applications such as image recognition, regression, fine-tuning, image detection, metric learning, and neural art in the following chapters, wherein, he can get step by step instructions to build the demo, which is "train by example" learning.

Jiang Tian, PHD
Lead Research Scientist, Siemens

前言

为什么要写这本书

当本书编辑找到我时，我才意识到，这两年火爆程度堪比一线城市房价的深度学习，在国内竟没有几本专门的中文书籍。目前市场上为数不多的有关深度学习的中文书，有大而全的名家之作，但其中关于深度学习的内容只有很少的一部分；有针对框架的工具书，但其偏重框架本身，缺少基础知识和理论，就好像介绍了一把剑，却没教剑法；有科普型的书，主要作用是提振读者自信，其实读者并不能从中学到任何动手能力。最重要的是，作为深度学习被应用最多，也是最有趣的领域——计算机视觉，却没有专门的书籍，这是笔者编写这本书的初衷。

我并非计算机科班出身，在转向深度学习的过程中遇到了很多问题，总结了许多经验。从写书的角度来讲，这没准是个优势。希望我在业界实际项目中的经验，加上自学时总结的教训，可以有效地帮助尚未进入深度学习和计算机视觉领域的读者。

本书面向有志于进入深度学习和计算机视觉领域的学生和技术研发人员，通过介绍深度学习及计算机视觉中最基础的知识，并结合最常见的应用场景和大量实例，带领读者进入丰富多彩的计算机视觉领域。作为一本“原理+实践”教程，本书用最基本的公式推导加上大量的配图和实例代码，给读者生动的内容，而不是枯燥的原理堆砌。

本书特色

- **注重原理和上手实战。**力求读者在阅读本书后，不仅能了解常见算法背后的思想，还能具备独立开发常见的基于深度学习的计算机视觉算法的能力。
- **原理讲解通俗易懂。**本书能通过文字定性讲解的就不用计算和公式，能用图表述的就不用文字。公式虽然不可避免，但尽量做到公式是讲解和图示的辅助，而不是需要细致理解的部分。
- **大量原创代码和图示。**本书结合作者在实际工作中的经验，将入门实战的例子和其他示例力求做到实用性和趣味性并存，并将代码分享到本书的 `github` 页面上供读者下载。本书的 `github` 页面地址为 https://frombeijingwithlove.github.io/dlcv_for_beginners/。

本书内容

本书共 13 章，分为 2 篇。

第 1 篇 基础知识

第 1 章从历史讲起，介绍了深度学习和计算机视觉的基础概念和常见应用。

第 2 章介绍了最基础的数学知识，包括线性代数、概率、高维空间、卷积和数学优化。

第 3 章以神经网络为重点，讲解了机器学习和数据的基础知识。

第 4 章在第 3 章的基础上，介绍了卷积神经网络的概念、原理和一些常见结构。

第 2 篇 实例精讲

第 5 章介绍 Python 基础，并通过小例子介绍了 NumPy 和 matplotlib 的基本使用。

第 6 章介绍了基于 Python 的 OpenCV 使用，并动手实现了数据增加小工具和物体框标注小工具。

第 7 章分别通过 MXNet 和 Caffe 实现了最简单的神经网络，以及结果的可视化。

第 8 章从头开始一步步基于 MNIST 数据集，分别实现了基于 MXNet 和 Caffe 的卷积神经网络模型用于手写数字识别，并介绍了如何测试和评估模型。

第 9 章实现了一个基于 Caffe 的用卷积神经网络做回归的例子，并介绍了如何制作 HDF5 格式数据，如何用 GPU 批量对数据进行运算，以及如何实现可视化训练的模型。

第 10 章首先实现了一个图片爬虫用于搜集图片数据，并以美食图片分类为例子一步步讲解如何基于 Caffe 实现迁移学习。然后在此基础上进一步讲解了如何对数据进行预处理，如何使用混淆矩阵、P-R 和 ROC 曲线，最后介绍了如何实现可视化卷积神经网络对输入图片的响应。

第 11 章针对 R-CNN 系和 YOLO/SSD 系这两类算法，简要介绍了基于深度学习的目标检测算法的发展史，并给出了基于 MXNet 的 SSD 检测算法实例，以及分析了结果的可视化。

第 12 章介绍了度量学习的基本概念，并从图片开始，一步步实现了基于 Caffe 的 Siamese 网络，还实现了基于 τ -SNE 的结果可视化。

第 13 章讲解了最基本的图像风格迁移算法，并给出了基于 MXNet 的图像风格迁移例子，读者可以用自己喜欢的图片做出算法生成的艺术作品。

适合阅读本书的读者

本书适合以下读者阅读：

- 对人工智能、机器学习感兴趣的读者；
- 对深度学习和计算机视觉感兴趣的读者；
- 希望用深度学习完成设计的计算机或电子信息专业的学生；
- 讲授机器学习和深度学习实践课的老师；
- 希望进一步提升编程水平的开发者；
- 机器学习与机器视觉研发人员和算法工程师；
- 人工智能产品经理。

阅读本书的读者最好具备以下要求：

- 至少具备高中以上的数学基础，本科以上最佳；
- 具备基本的编程能力；
- 了解 Linux 的基本使用；
- 拥有一台 NVIDIA 显卡的计算机，最好是 2GB 以上的显存。

本书虽然定位为入门书，但并不能保证 21 天或者 3 个月包会。学习是一件没有捷径可走的事情，希望本书能帮助读者少走弯路，也希望每一位读者翻开书前，都是带着好奇和兴趣。

纠错

由于是第一次写书，且时间仓促，错误之处估计难以避免，敬请读者朋友们发现错误后到本书的 [github](#) 页面指出，我会尽快更新在勘误表里，不胜感激！

本书作者

本书由叶韵主笔编写。其他参与编写的人员有张昆、张友、赵桂芹、晁楠、高彩琴、郭现杰、刘琳、王凯迪、王晓燕、吴金艳、尹继平、张宏霞、张晶晶、陈冠军、魏春、张燕、范陈琼、孟春燕、王晓玲、项宇峰、肖磊鑫、薛楠、杨丽娜、闫利娜、王韶、李杨坡、刘春华、黄艳娇、刘雁。

致谢

成书过程中，得到了很多人的帮助，在此向他们表示诚挚的谢意。

首先感谢负责本书的编辑，我既不是大 V，博客文章也没几篇，是他们的信任让我能通过这次写作梳理知识，并获得稿费。

感谢我的朋友们：NVIDIA（英伟达）的高级工程师华远志，谷歌的软件工程师吕佳楠，清华大学的马晨同学，NVIDIA 的高效 GPU 架构师欧阳晋博士，西门子的高级研究员田疆博士，理光软件研究院的研究员钟诚博士和亚马逊的科学家庄晓天博士。他们几位在成书过程中给予了我很多帮助和建议，特别是田疆博士、钟诚博士和欧阳晋博士，对本书内容给出了全面且独到的建议，并指出了部分错误。

感谢京东从事 AI 方向的同事们，他们各个都身怀绝技，工作中和他们的交流给了我很多技术上的帮助和灵感。

感谢插画师翟少昂为本书绘制插画，感谢星河互联高级投资经理杨森授权本书使用他于漠河的严冬中拍摄的照片。

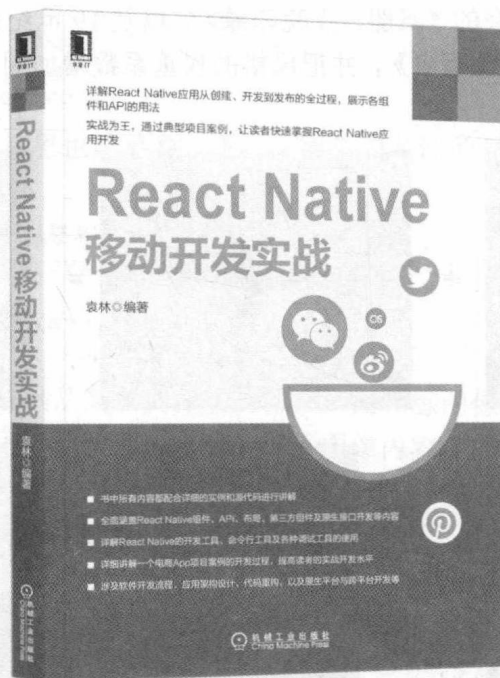
感谢我的博士生导师 Yu Cao 教授和我第一家效力公司的 CTO——Bruce McGaughey 博士。在应试教育的影响下，我浑浑噩噩求学近 20 载，未曾找到学习的意义。和他们二位接触的过程中我才渐渐明白，原来学习最重要的是兴趣。也因此，后来我才有勇气放弃了钻研了近 10 年的老本行，转向自己更感兴趣的算法和机器学习。

最后也是我最想感谢的，是我的家人！写书期间我的家人承担了一切家务劳动，是他们全方位的支持和细致照顾，让我在身体健康的状态下完成了写作。特别是我的妻子，尽管她看不懂我在写什么，却认为写得很棒。是她每天的鼓励和支持才让我没有半途而废，能在此表达我对她的谢意，是我完成这本书的最大动力。

叶韵

于北京市西郊机场老干部活动中心

推荐阅读



React Native移动开发实战

作者：袁林 书号：978-7-111-57179-7 定价：69.00元

**详解React Native应用从创建、开发到发布的全过程，展示各组件和API的用法
实战为王，通过典型项目案例，让读者快速掌握React Native应用开发**

本书以实战开发为主旨，以React Native应用开发为主线，以iOS和Android双平台开发为副线，通过完整的电商类App项目案例，详细地介绍了React Native应用开发所涉及的知识，让读者全面、深入、透彻地理解React Native的主流开发方法，从而提升实战开发水平和项目开发能力。

本书共12章，分为4篇，涵盖的主要内容有搭建开发环境、Nuclide、各种命令行工具（Git、Node.js）、布局与调试、组件、API、第三方组件、基于Node.js的服务器、fetch API、AsyncStorage/SQLite/Realm数据库存储、原生平台接口开发、redux开发框架、应用打包与发布、热更新与CodePush等。

本书适合iOS和Android原生平台应用开发者，以及有兴趣加入移动平台开发的JavaScript开发者阅读。当然，本书也适合相关院校和社会培训学校作为移动开发的教材使用。

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