

21个项目玩转 深度学习

基于TensorFlow的实践详解

何之源◎编著

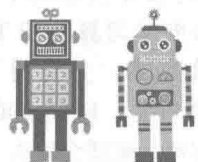


中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

内容简介



21个项目玩转 深度学习

基于TensorFlow的实践详解

何之源◎编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

《21个项目玩转深度学习——基于 TensorFlow 的实践详解》以实践为导向，深入介绍了深度学习技术和 TensorFlow 框架编程内容。

通过本书，读者可以训练自己的图像识别模型、进行目标检测和人脸识别、完成一个风格迁移应用，还可以使用神经网络生成图像和文本，进行时间序列预测、搭建机器翻译引擎，训练机器玩游戏。全书共包含 21 个项目，分为深度卷积网络、RNN 网络、深度强化学习三部分。读者可以在自己动手实践的过程中找到学习的乐趣，了解算法和编程框架的细节，让学习深度学习算法和 TensorFlow 的过程变得轻松和高效。本书代码基于 TensorFlow 1.4 及以上版本，并介绍了 TensorFlow 中的一些新特性。

本书适合有一定机器学习基础的学生、研究者或从业者阅读，尤其是希望深入研究 TensorFlow 和深度学习算法的数据工程师，也适合对人工智能、深度学习感兴趣的在校学生，以及希望进入大数据应用的研究者。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目 (CIP) 数据

21 个项目玩转深度学习：基于 TensorFlow 的实践详解 / 何之源编著. —北京：电子工业出版社，2018.3

ISBN 978-7-121-33571-6

I. ①2… II. ①何… III. ①人工智能—算法 IV. ①TP18

中国版本图书馆 CIP 数据核字(2018)第 018503 号

策划编辑：孙学瑛

责任编辑：孙学瑛

印 刷：北京季蜂印刷有限公司

装 订：北京季蜂印刷有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：720×1000 1/16 印张：23.25 字数：368 千字 彩插：4

版 次：2018 年 3 月第 1 版

印 次：2018 年 5 月第 3 次印刷

定 价：79.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888, 88258888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

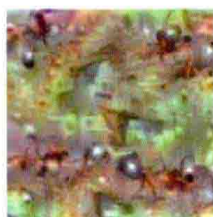
本书咨询联系方式：010-51260888-819, faq@phei.com.cn。



Hartebeest



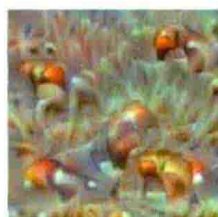
Measuring Cup



Ant



Starfish



Anemone Fish



Banana



Parachute



Screw

图 4-2 极大化神经网络各个输出类别概率得到的图片



图 4-6 拉普拉斯金字塔标准化后得到的 DeepDream 图片

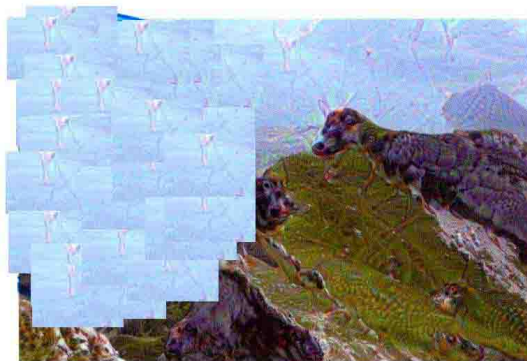


图 4-9 带有背景的 DeepDream 图片

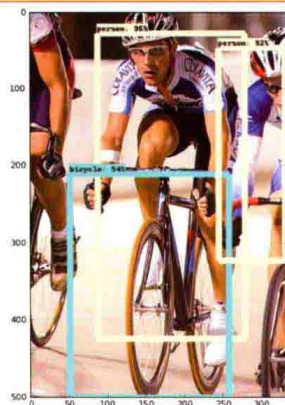


图 5-13 训练深度学习目标检测模型



风格化之前

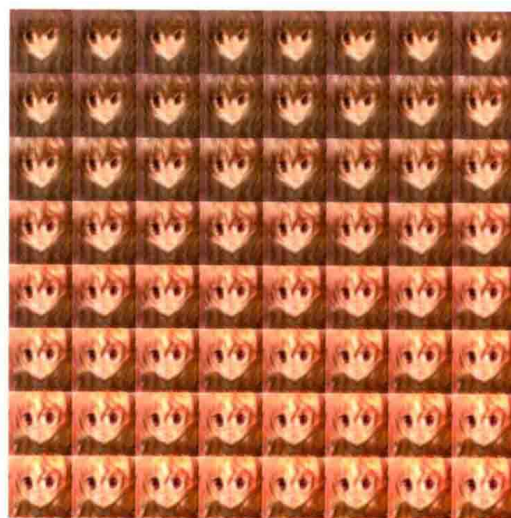


风格化之后

 图 7-6 图像的快速风格迁移



 图 8-8 GAN 模型自动生成的图像




 图 8-10 GAN 模型隐空间中的插值可视化



图 9-11 使用 pix2pix 模型为黑白食物图像自动上色

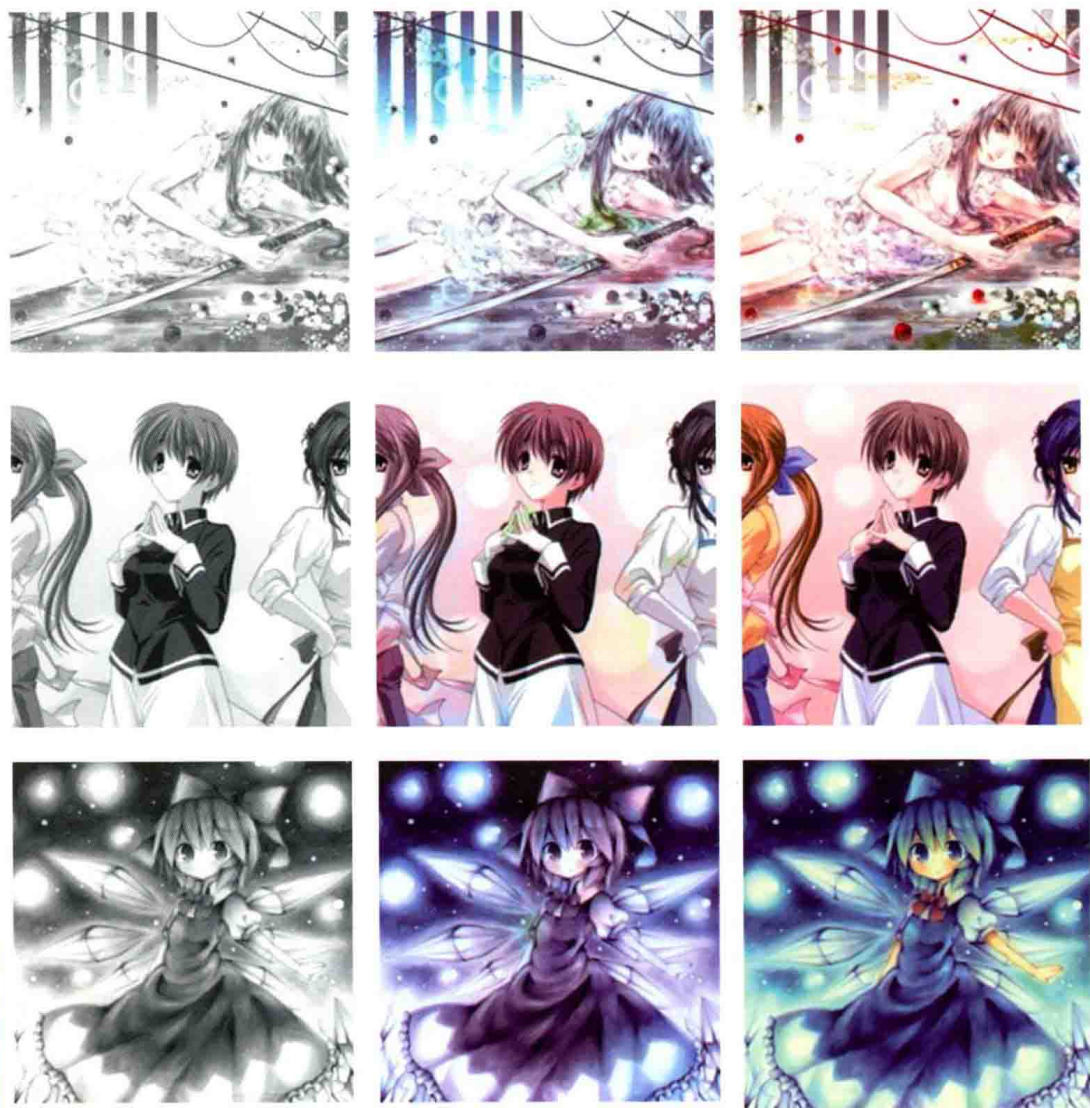


图 9-12 使用 pix2pix 模型为动漫图像自动上色



图 10-2 图像的超分辨率



图 11-6 训练 CycleGAN 模型将男性图片变为女性图片

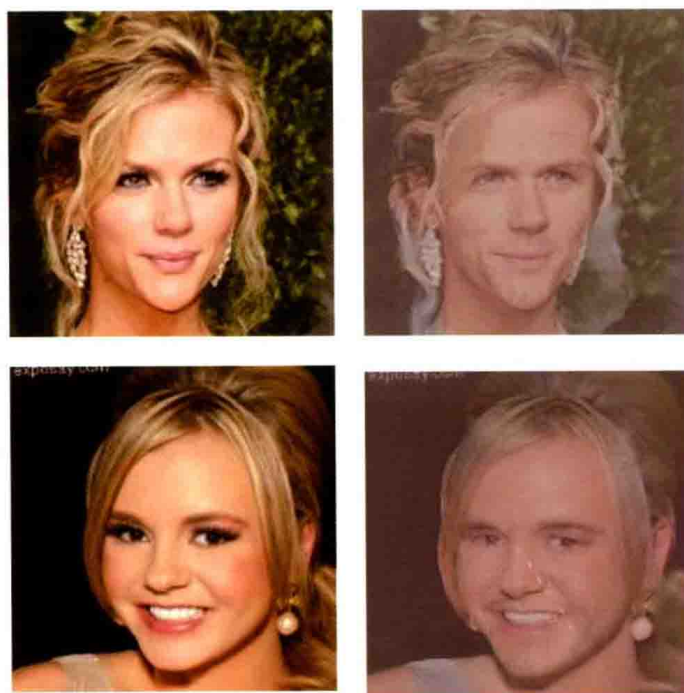


图 11-7 训练 CycleGAN 模型将女性图片变为男性图片



21 个项目玩转深度学习

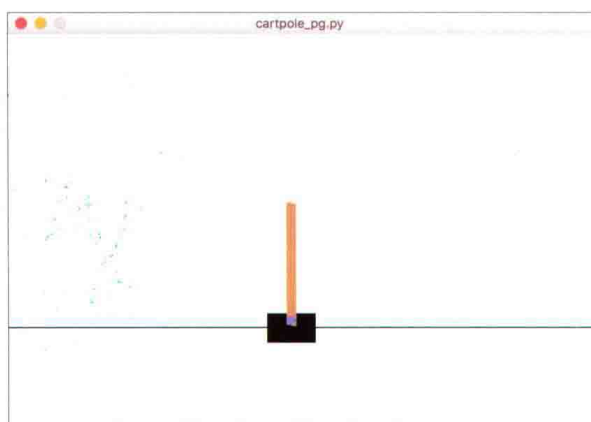
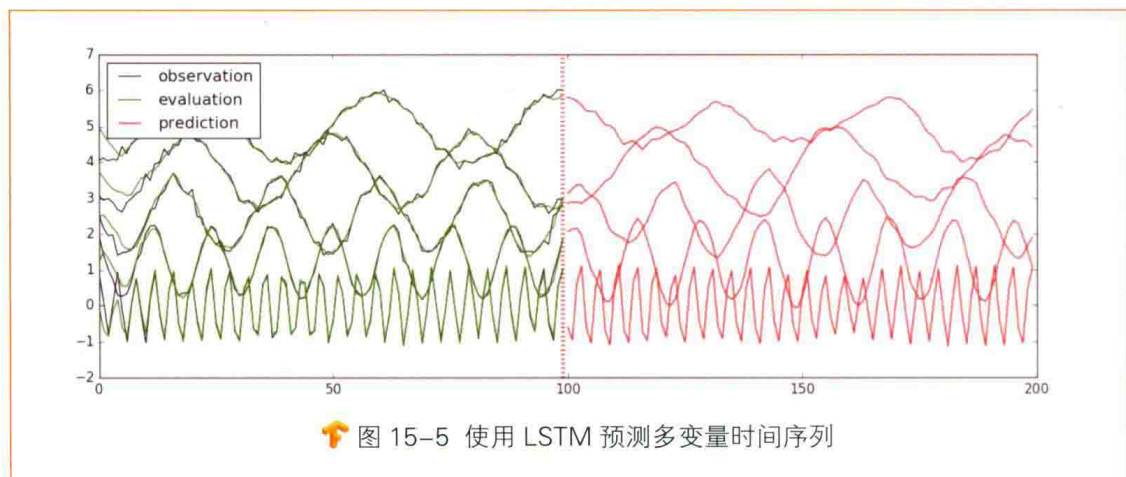


图 21-2 训练策略梯度模型解决 CartPole 问题

前言

我们正处在一个日新月异、飞速变革的时代，层出不穷的新技术每天都在冲击和改变我们的生活。人工智能无疑是其中最受关注、也是影响最深远的技术领域。它为计算机插上了翅膀，演变出许多从前根本无法想象的新技术、新应用。AlphaGo Zero——一台没有任何先验知识的人工智能机器，可以在几天内通过自我博弈成长为世界第一的围棋大师，超越人类几千年积累的经验；风格迁移应用能够自动将用户的照片转变为著名的绘画艺术风格；机器可以在零点几秒内完成翻译，把一种语言译成另一种语言。此外，有关人脸识别、自动驾驶等新技术的应用也都纷纷开始落地。在过去的几年内，人工智能技术不仅在学术上取得了巨大的突破，也开始走向寻常百姓家，真正为人们的生活提供便利。

本书主要为读者介绍这次人工智能浪潮最重要的组成部分——深度学习技术，使用的编程框架是 Google 公司的 TensorFlow。借助于 Google 公司巨大的影响力，TensorFlow 一经发布就引起了广泛的关注。目前（截至 2017 年年底），TensorFlow 在 Github 上已经有了 8 万 4000 多个 Star，是所有深度学习框架中最多的。

本书的主要特点如下：

- 以实践、应用导向。深度学习需要深厚的数理基础，对于初学者来说有一定难度。本书希望从实践出发，用具体的例子来引导读者学习深度学习技术和 TensorFlow 编程技巧。我们主要从实用性和趣味性两个方面考

量, 选择了 21 个实践项目, 其中既有 MNIST 图像识别入门项目, 也有目标检测、人脸识别、时间序列预测实用性项目, 还有一些诸如 Deep Dream 趣味性项目。读者可以在实践中找到乐趣, 逐渐进步, 让学习深度学习和 TensorFlow 的过程不再那么痛苦。

- 清晰而有深度的介绍。在编写过程中, 尽量用简单的语言描述算法的原理, 做到清晰而有条理。此外, 深度学习是一门还在快速发展的新技术, 限于篇幅, 很多内容不能完全展开, 在大部分章节的末尾列出了“拓展阅读”材料, 有兴趣的读者可以参考进一步学习。
- 基于 TensorFlow 1.4 及以上版本。TensorFlow 的发展非常迅速, 本书代码全部基于 TensorFlow 1.4 及以上版本 (包括 1.4.0、1.4.1 和 1.5.0), 并介绍了 TensorFlow 的一些新特性, 如 Time Series 模块 (1.3 版本添加)、新的 MultiRNNCell 函数 (1.2 版本更改) 等。本书的代码会在如下 GitHub 地址上提供, 并会随新的 TensorFlow 版本的发布而同步更新:
<https://github.com/hzy46/Deep-Learning-21-Examples>。

本书代码推荐的运行环境为: Ubuntu 14.04, Python 2.7、TensorFlow 1.4.0。请尽量使用类 UNIX 系统和 Python 2 运行本书的代码。

本书主要内容

本书共包括 21 章, 内容编排如下:

第 1~11 章主要介绍深度卷积神经相关的项目。其中, 第 1~3 章属于入门章节, 主要讨论深度学习中最基础的图像识别问题; 第 4~7 章讨论了其他计算机视觉相关的实践案例, 如目标识别、人脸识别、图像风格迁移等; 第 8~11 章介绍了 GAN 模型和它的几个重要变体。

第 12~17 章主要介绍 RNN、LSTM 相关的项目。RNN、LSTM 通常用来处理序列型数据, 第 12 章是一个入门章节, 会详细介绍 RNN 和 LSTM 的原理、实现方法和一个应用实例——Char RNN; 第 13~17 章讨论一些更复杂也更具体的案例, 如序列分类、词嵌入表示、时间序列预测、机器翻译等。

第 18~21 章主要介绍强化学习相关的项目。第 18、19 章分别介绍了相对简单的 QLearning 和 SARSA 算法,第 20 章和第 21 章介绍了更复杂的 DQN 和策略梯度算法。

如何阅读本书

在阅读本书前,读者应当了解 Linux 系统的基本操作,并会使用 Python 进行简单的编程,还需具备基础的机器学习知识。本书的章节安排是依据读者对深度学习的了解循序渐进设立的,建议初学者从前至后阅读。由于深度学习和 TensorFlow 不易于理解,我建议读者分几遍阅读本书:

- 第一遍先简单浏览一下,看一看书中都有哪些实践项目。当对基本的概念有初步的理解时,就可以尝试配置一下开发环境。
- 第二遍,从前至后阅读各章中算法的原理,并运行相应的实践项目。在这个过程中,希望读者能在自己动手的过程中找到学习的乐趣。读者可以对照源代码和书中的内容,深入学习各个模型的细节,此外还可以根据自己的需求对代码进行修改。本书会涉及大量 TensorFlow 中的函数,由于篇幅限制,不可能将每一个函数都介绍一遍,只介绍了比较重要的函数,读者可以参考 TensorFlow 的官方文档,查看其他函数的功能说明。
- 第三遍,根据需要对照源程序看相关章节。此外,读者还可以参阅章节最后的拓展阅读进一步学习。最后,如果你是一名深度学习和 TensorFlow 的精通者,也可以根据需要直接跳读到相关章节,查阅你需要的内容。

致谢

我首先要感谢我的父母,是他们将我养育成人,父母永远健康快乐是我最大的心愿。

感谢互联网时代,感谢网络上数不清的优秀开发者和博主,也向 Google 公司的开源精神致敬,让我们可以如此紧跟时代最前沿的技术,并为技术的进步做出自己微薄的贡献。

我还要真诚地感谢电子工业出版社对这本书的认可和兴趣。感谢电子工业出版社的孙学瑛女士，她的热情推动最终促成了我与电子工业出版社的合作。感谢宋亚东编辑，他对本书的重视和诚恳的建议，在写作过程中给了我莫大的帮助。

最后，感谢刘婧源同学给本书提出的宝贵意见。

由于本人水平有限，书中不足及错误之处在所难免，敬请专家和读者给予批评指正。如果您想和我进行技术交流，可以发送意见反馈邮件至 hzydl21@163.com，也可在知乎上找到我：<https://www.zhihu.com/people/he-zhi-yuan-16/>，此外还可以访问书友论坛 <http://forum.broadview.com.cn>。

何之源

2018 年 1 月

读者服务

轻松注册成为博文视点社区用户 (www.broadview.com.cn)，扫码直达本书页面。

- **下载资源**：本书如提供示例代码及资源文件，均可在 [下载资源](#) 处下载。
- **提交勘误**：您对书中内容的修改意见可在 [提交勘误](#) 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动**：在页面下方 [读者评论](#) 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/33571>



目 录

| | |
|---|----|
| 第 1 章 MNIST 机器学习入门..... | 1 |
| 1.1 MNIST 数据集 | 1 |
| 1.1.1 简介..... | 1 |
| 1.1.2 实验：将 MNIST 数据集保存为图片..... | 4 |
| 1.1.3 图像标签的独热表示..... | 6 |
| 1.2 利用 TensorFlow 识别 MNIST..... | 7 |
| 1.2.1 Softmax 回归..... | 7 |
| 1.2.2 两层卷积网络分类..... | 13 |
| 1.3 总结..... | 17 |
| 第 2 章 CIFAR-10 与 ImageNet 图像识别 | 19 |
| 2.1 CIFAR-10 数据集..... | 19 |
| 2.1.1 CIFAR-10 数据集简介..... | 19 |
| 2.1.2 下载 CIFAR-10 数据..... | 21 |
| 2.1.3 TensorFlow 的数据读取机制..... | 23 |
| 2.1.4 实验：将 CIFAR-10 数据集保存为图片形式..... | 29 |
| 2.2 利用 TensorFlow 训练 CIFAR-10 识别模型..... | 33 |
| 2.2.1 数据增强..... | 33 |
| 2.2.2 CIFAR-10 识别模型..... | 35 |
| 2.2.3 训练模型..... | 37 |
| 2.2.4 在 TensorFlow 中查看训练进度..... | 38 |
| 2.2.5 测试模型效果..... | 40 |

| | | |
|--------------|-------------------------------------|-----------|
| 2.3 | ImageNet 图像识别模型 | 43 |
| 2.3.1 | ImageNet 数据集简介 | 43 |
| 2.3.2 | 历代 ImageNet 图像识别模型 | 44 |
| 2.4 | 总结 | 47 |
| 第 3 章 | 打造自己的图像识别模型 | 49 |
| 3.1 | 微调的原理 | 49 |
| 3.2 | 数据准备 | 51 |
| 3.3 | 使用 TensorFlow Slim 微调模型 | 54 |
| 3.3.1 | 下载 TensorFlow Slim 的源代码 | 54 |
| 3.3.2 | 定义新的 datasets 文件 | 55 |
| 3.3.3 | 准备训练文件夹 | 57 |
| 3.3.4 | 开始训练 | 58 |
| 3.3.5 | 训练程序行为 | 60 |
| 3.3.6 | 验证模型准确率 | 61 |
| 3.3.7 | TensorBoard 可视化与超参数选择 | 62 |
| 3.3.8 | 导出模型并对单张图片进行识别 | 64 |
| 3.4 | 总结 | 67 |
| 第 4 章 | Deep Dream 模型 | 68 |
| 4.1 | Deep Dream 的技术原理 | 68 |
| 4.2 | TensorFlow 中的 Deep Dream 模型实践 | 70 |
| 4.2.1 | 导入 Inception 模型 | 70 |
| 4.2.2 | 生成原始的 Deep Dream 图像 | 73 |
| 4.2.3 | 生成更大尺寸的 Deep Dream 图像 | 76 |
| 4.2.4 | 生成更高质量的 Deep Dream 图像 | 79 |
| 4.2.5 | 最终的 Deep Dream 模型 | 84 |
| 4.3 | 总结 | 87 |
| 第 5 章 | 深度学习中的目标检测 | 88 |
| 5.1 | 深度学习中目标检测的原理 | 89 |
| 5.1.1 | R-CNN 的原理 | 89 |
| 5.1.2 | SPPNet 的原理 | 90 |