

深度学习系列

Packt


Keras

高级深度学习

Advanced Deep Learning with Keras

[菲] 罗韦尔·阿蒂恩扎 (Rowel Atienza) 著

蔡磊 潘华贤 程国建 译

 机械工业出版社
CHINA MACHINE PRESS

深度学习系列

Keras 高级深度学习

[菲] 罗韦尔·阿蒂恩扎 (Rowel Atienza) 著
蔡 磊 潘华贤 程国建 译

 **机械工业出版社**
CHINA MACHINE PRESS

本书是高级深度学习技术的综合指南，内容包括自编码器、生成对抗网络（GAN）、变分自编码器（VAE）和深度强化学习（DRL），在这些技术的推动下，AI 于近期取得了令人瞩目的成就。

本书首先对多层感知器（MLP）、卷积神经网络（CNN）和循环神经网络（RNN）进行了概述，这些是本书中介绍的更高级技术的构建模块。之后探索了包括 ResNet 和 DenseNet 在内的深度神经网络架构以及如何创建自编码器。读者将学习如何使用 Keras 和 TensorFlow 实现深度学习模型，并进一步实现其高级应用。随后，读者将会了解到有关 GAN 的所有知识，以及认识到其如何将 AI 性能提升到新的水平。在此之后，读者可快速了解 VAE 的实现方式，并将认识到 GAN 和 VAE 是如何具备生成数据的能力的，并且使所生成的数据对人类来说极具说服力。因此，该类方法已成为现代 AI 的一个巨大进步。为充分了解该系列相关先进技术，读者将会学习如何实现 DRL，例如深度 Q-Learning 和策略梯度方法，这些方法对于 AI 在现代取得很多成就至关重要。

本书适合想要深入了解深度学习高级主题的机器学习工程师，以及高等院校人工智能、数据科学、计算机科学等相关专业学生阅读。

Copyright © 2018 Packt Publishing

First published in the English language under the title “Advanced Deep Learning with Keras” / by Rowel Atienza / ISBN: 978-1-78862-941-6

Copyright in the Chinese language (simplified characters) © 2020 China Machine Press

This translation of *Advanced Deep Learning with Keras* first published in 2020 is published by arrangement with Packt Publishing Ltd.

This title is published in China by China Machine Press with license from Packt Publishing Ltd. This edition is authorized for sale in China only, excluding Hong Kong SAR, Macao SAR and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书由 Packt Publishing Ltd 授权机械工业出版社在中华人民共和国境内（不包括香港、澳门特别行政区及台湾地区）出版与发行。未经许可的出口，视为违反著作权法，将受法律制裁。

北京市版权局著作权合同登记 图字：01-2019-0589 号。

图书在版编目（CIP）数据

Keras 高级深度学习 / (菲) 罗韦尔·阿蒂恩扎 (Rowel Atienza) 著；蔡磊，潘华贤，程国建译。
—北京：机械工业出版社，2020.3

（深度学习系列）

书名原文：Advanced Deep Learning with Keras

ISBN 978-7-111-64796-6

I . ① K… II . ①罗… ②蔡… ③潘… ④程… III . ①机器学习 IV . ① TP181

中国版本图书馆 CIP 数据核字（2020）第 028952 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）

策划编辑：刘星宁 责任编辑：刘星宁 任鑫

责任校对：肖琳 封面设计：马精明

责任印制：孙炜

保定市中华美凯印刷有限公司印刷

2020 年 5 月第 1 版第 1 次印刷

184mm × 240mm · 17.75 印张 · 4 插页 · 416 千字

标准书号：ISBN 978-7-111-64796-6

定价：89.00 元

电话服务

网络服务

客服电话：010-88361066 机工官网：www.cmpbook.com

010-88379833 机工官博：weibo.com/cmp1952

010-68326294 金书网：www.golden-book.com

封底无防伪标均为盗版 机工教育服务网：www.cmpedu.com

作者简介

Rowel Atienza是菲律宾大学蒂利曼分校电气与电子工程学院副教授，并担任Dado和Maria Banatao研究所人工智能讲席教授。Rowel毕业于菲律宾大学，并一直对智能机器人十分着迷。他在新加坡国立大学从事AI增强四足机器人方面的工作，并以此获得工程学硕士学位。此后，他凭借其在人机交互主动视线跟踪领域的贡献，获得澳大利亚国立大学的博士学位。Rowel当前的研究重点是AI和计算视觉，他梦想着构建出可以感知、理解并推理的实用机器。Rowel获得了来自于菲律宾科学技术部（DOST）、菲律宾三星研究院和菲律宾高等教育委员会-菲律宾加利福尼亚高级研究所（CHED-PCARI）的多项科研资助。

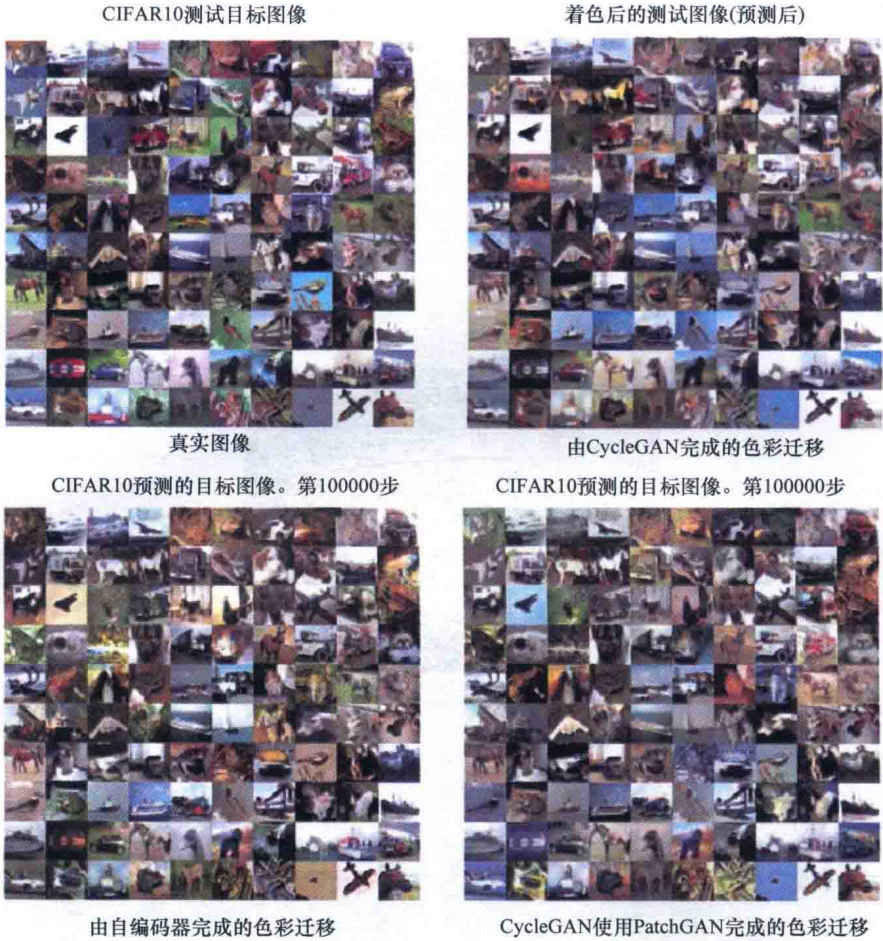
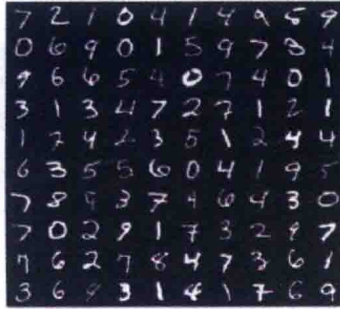


图 7.1.9 使用不同技术的色彩迁移结果。包括真实图像，使用自编码器进行着色的结果，使用 vanillaGAN 判别器进行辅助的 CycleGAN 的结果，以及使用 PatchGAN 判别器辅助的 CycleGAN 的着色结果（彩色图片可从本书 Github 代码库中找到）



图 7.1.11 两个不同的域，数据未对齐（彩色图片可从本书 Github 代码库中找到）

MNIST测试源图像



MNIST 数字

所预测的SVHN目标图像，
第100000步



SVHN域中的MNIST数字

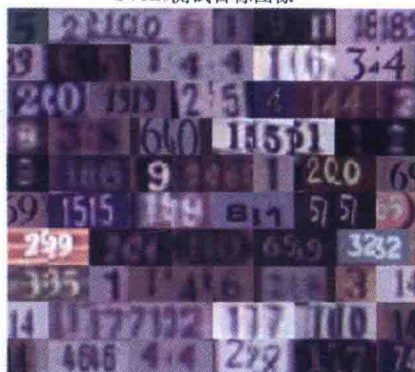
所预测的SVHN目标图像，
第100000步



使用PatchGAN的SVHN
域中的MNIST数字

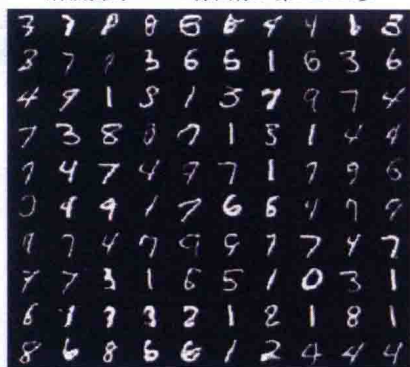
图 7.1.12 测试数据从 MNIST 域到 SVHN 的风格转换
(彩色图片可从本书 GitHub 代码库中找到)

SVHN测试目标图像



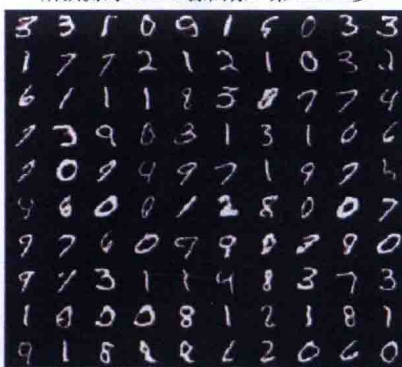
SVHN

所预测的MNIST源图像。第100000步



MNIST 域中的SVHN数字

所预测的MNIST源图像。第100000步



使用PatchGAN的SVHN域中的
MNIST数字

图 7.1.13 测试数据从 SVHN 域到 MNIST 的样式转换
(彩色图片可从本书 GitHub 代码库中找到)

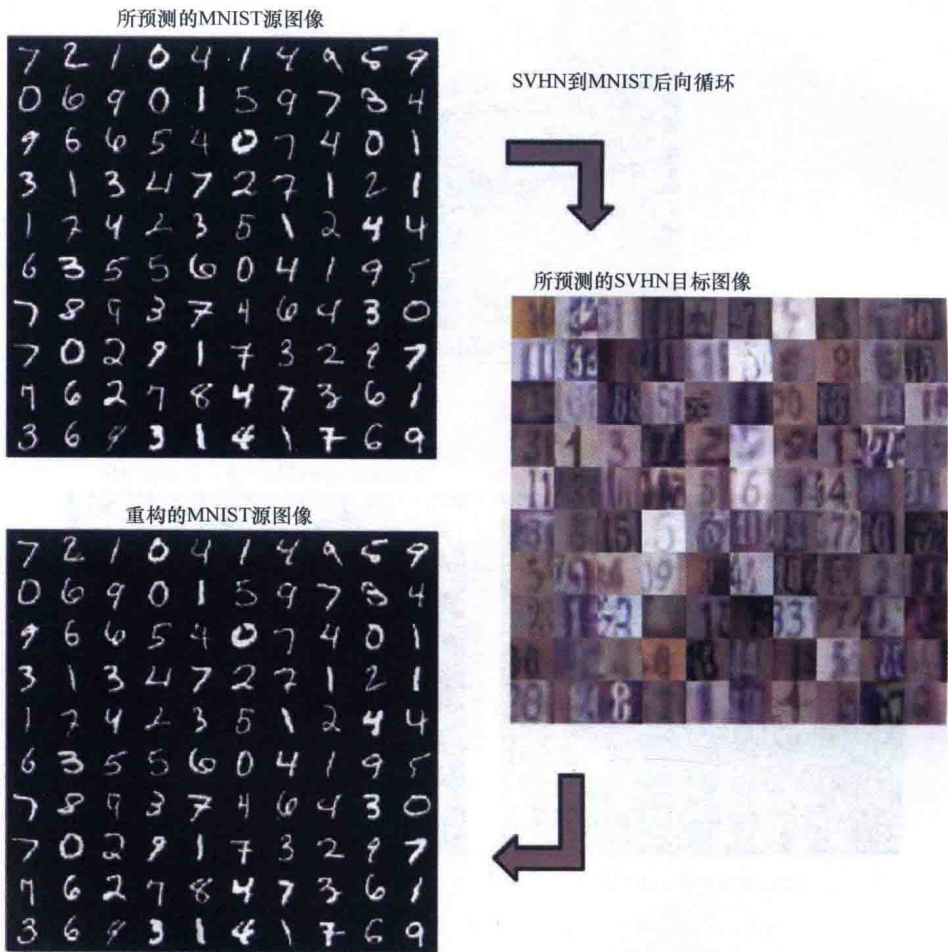
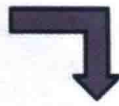


图 7.1.14 使用 PatchGAN 的前向循环 CycleGAN 从 MNIST (源) 转换到 SVHN (目标)
(彩色图片可从本书 GitHub 代码库中找到)

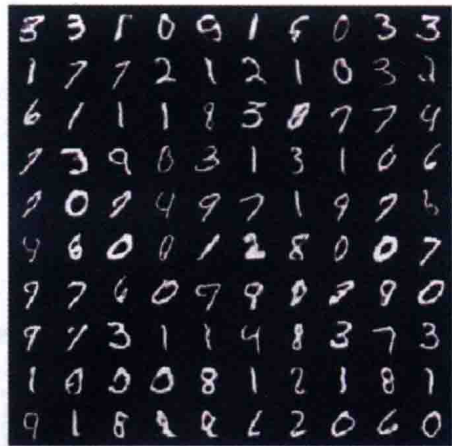
SVHN测试目标图像



SVHN到MNIST后向循环



所预测的MNIST源图像



重构的SVHN目标图像



图 7.1.15 使用 PatchGAN 的后向循环 CycleGAN 从 MNIST (源) 转换到 SVHN (目标)。重建后的目标与原始目标不完全相同 (彩色图片可从本书 GitHub 代码库中找到)

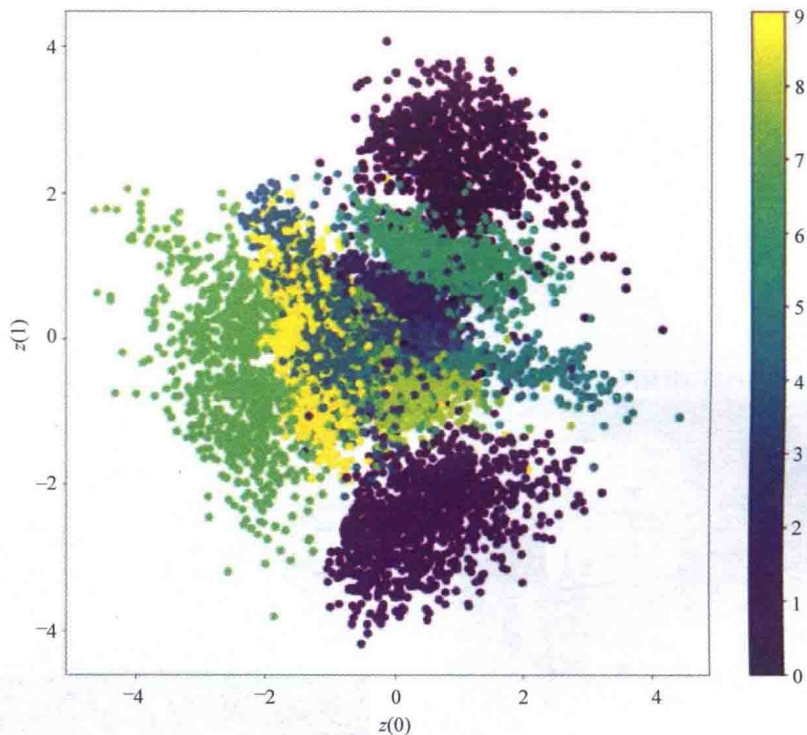


图 8.1.6 测试数据集 (VAE MLP) 上潜向量均值。色条显示随 z 的函数变化所展现出的相应 MNIST 数字 (彩色图片可从本书 GitHub 代码库中找到)

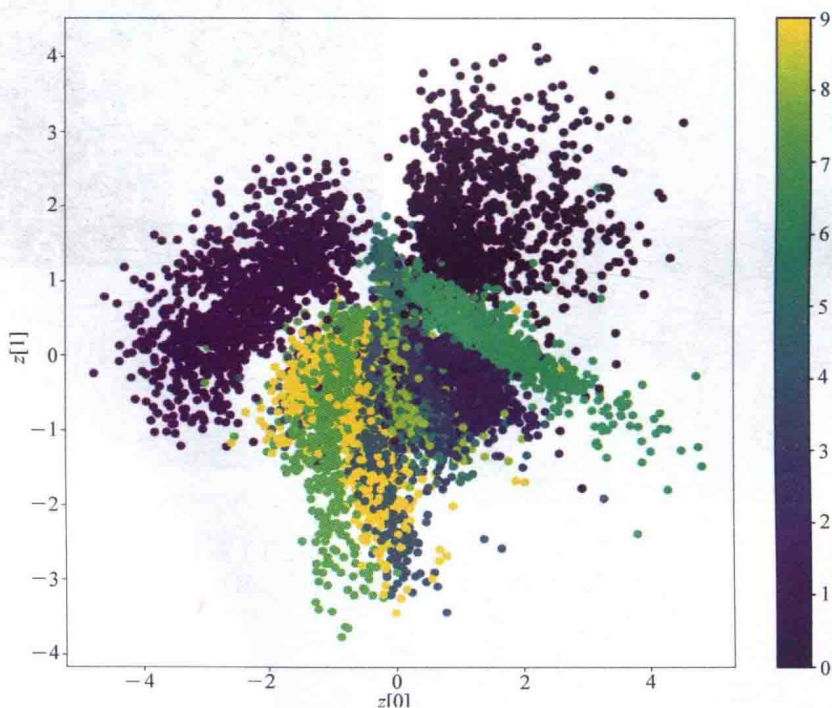


图 8.1.11 测试数据集 (VAE CNN) 上的潜向量平均值。色条显示随 z 的函数变化所展现出的相应 MNIST 数字 (彩色图片可从本书 GitHub 代码库中找到)

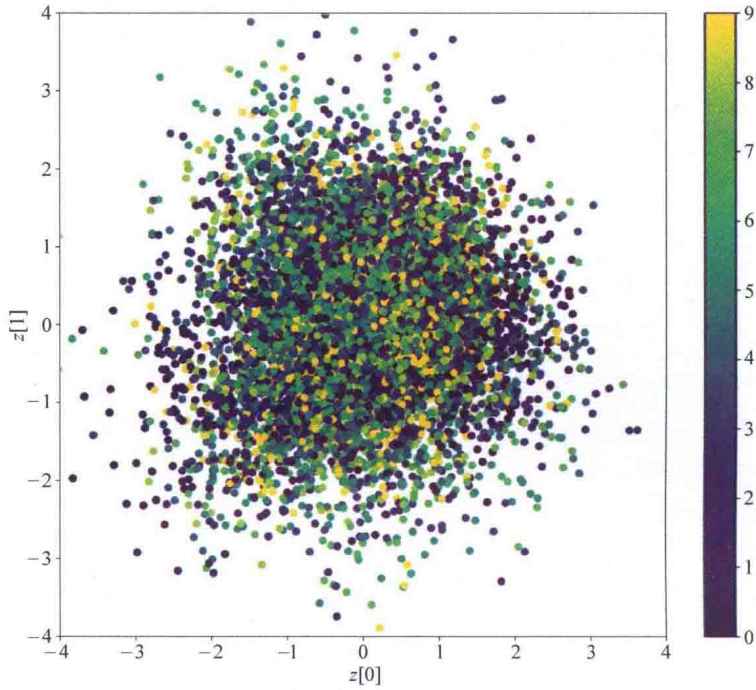


图 8.2.4 测试数据集上潜向量均值 (CVAE CNN)。
色条显示随 z 的函数变化所展现出的相应 MNIST 数字
(彩色图片可从本书 GitHub 代码库中找到)

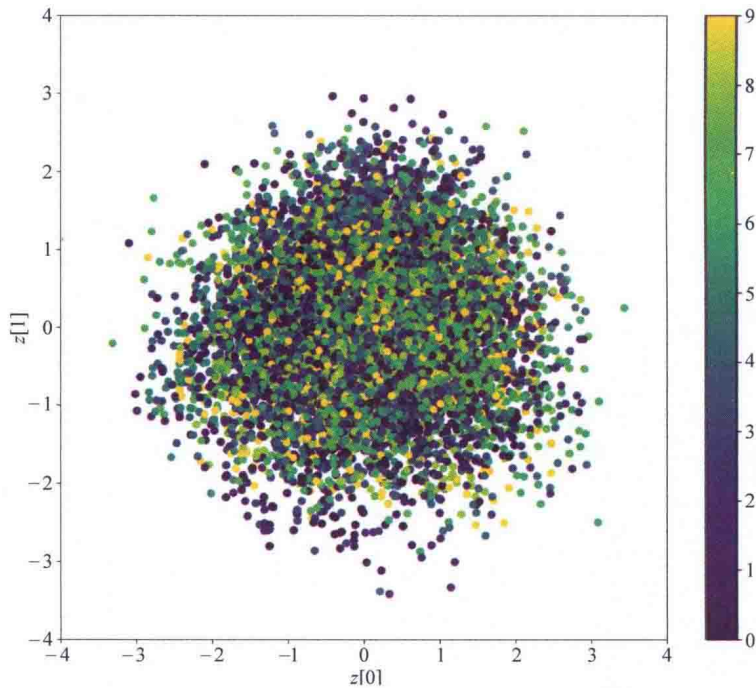


图 8.3.1 测试数据集上潜向量的均值 (β -VAE, 其中 $\beta=7$)
(彩色图片可从本书 GitHub 代码库中找到)

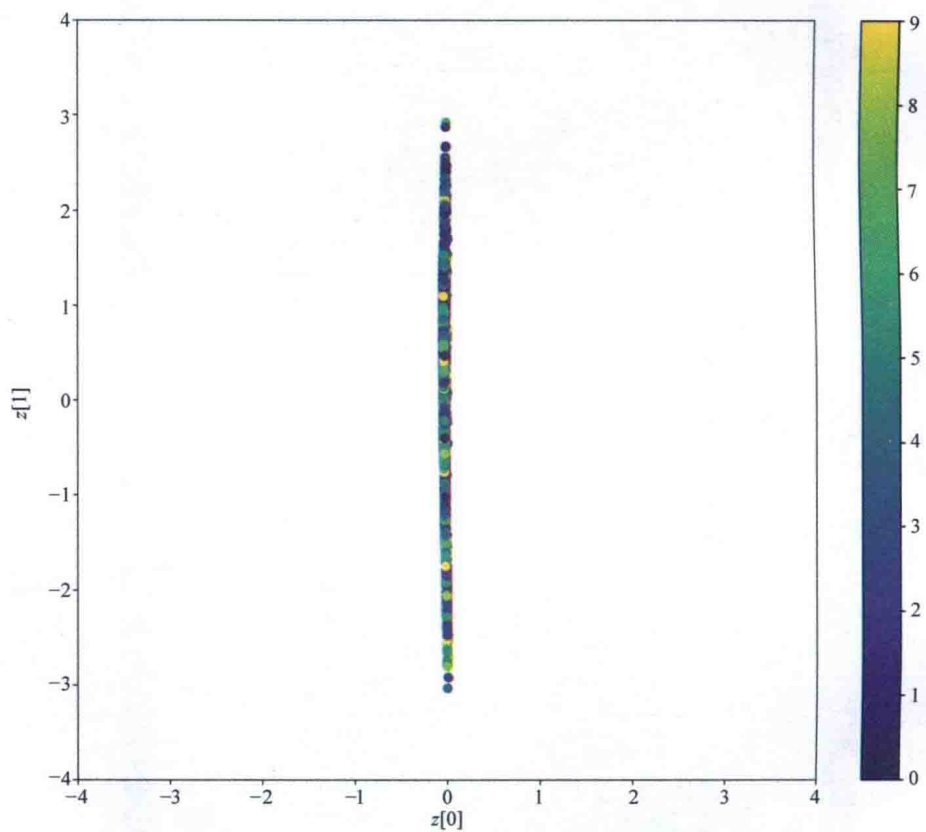


图 8.3.2 测试集合上潜向量的均值 (β -VAE, 其中 $\beta=10$)
(彩色图片可从本书 GitHub 代码库中找到)

本书特色

最近，深度学习的发展，包括生成对抗网络（GAN）、变分自编码器（VAE）和深度强化学习等，所创造的令人瞩目的AI成就不断出现在报纸的头条——Alpha Zero战胜世界象棋冠军；由于十分接近人类的绘画作品，通过生成式AI所创作的作品以超过40万美元的价格被售出。

本书是高级深度学习技术的综合指南，内容包括自编码器、GAN、VAE和深度强化学习，在这些技术的推动下AI于近期取得了令人瞩目的成就。

本书主要特色：

- 探索推动现代AI取得诸多成就的高级深度学习技术；
- 介绍了如何实现深度神经网络、自编码器、GAN、VAE和深度强化学习；
- 对GAN进行了广泛研究，包括改进的GAN、跨域GAN和分离表示GAN。

本书假定读者熟悉Python。作为一本介绍高级技术的书籍，假定读者熟悉一些机器学习方法，如对深度学习有一定实践经验将更有助于理解本书的内容。无需对Keras和TensorFlow知识有事先了解。

译者序

目前业界与学界流行的热词是 ABCDIV (A—人工智能、B—区块链、C—云计算、D—数据科学与大数据、I—物联网、V—虚拟现实)，而人工智能则包含了机器学习与深度学习。三位深度学习之父 Yoshua Bengio、Yann LeCun、Geoffrey Hinton 因他们在神经计算领域的突出贡献获得了 2019 年度计算机科学界的最高奖——图灵奖。今天，深度学习已经成为人工智能技术领域最重要的技术之一。在最近数年中，计算机视觉、语音识别、自然语言处理和机器人技术取得的巨大进展都离不开深度学习。

Keras 的命名来自古希腊语，含义为将梦境化为现实的“牛角之门”。出自《荷马史诗·奥德赛》第 19 卷涅涅洛佩与奥德修斯的对话——无法挽留的梦幻拥有两座门，一座门由牛角制成，一座门由象牙制成。象牙之门内光彩夺目，却仅是无法实现的梦境；唯有走进牛角之门，才能由梦境看见真实。Keras 是一款用 Python 编写的 AIML (人工智能与机器学习) 深度学习函数库及 API，它能够以开源平台 TensorFlow、CNTK、Theano 等作为后端运行。Keras 的特点是支持敏捷开发及快速 AIML 实验环境搭建。它允许简单而快速的原型设计，可在 CPU 和 GPU 上无缝运行，能够以最小的时延把您的想法转换为实验结果。它允许实践者以尽可能少的限制将各模块组装在一起形成需要的网络结构，诸如神经网络层数、损失函数、优化器、初始化方法、激活函数、正则化方法等都可以结合起来构造成为新的集成模型，从而轻松地创建可以提高表现力的新模块。

本书内容涵盖了 AIML 前沿技术介绍；使用 Keras 实现高级深度学习模块的集成；多层感知器 (MLP)、卷积神经网络 (CNN)、循环神经网络 (RNN) 等模块的构建与组合；自编码器和变分自编码器 (VAE)；生成对抗网络 (GAN) 及其 AI 实现技术；分离表示 GAN 和跨域 GAN；深度强化学习的方法及其实现；使用 OpenAI Gym 产生符合行业标准的應用；深度 Q-Learning 和策略梯度方法等。

本书适合于本科院校作为培训教材，对研究生课题及其实现也有极大的帮助；对自学及实战者也是不可多得的辅助参考。本书的翻译出版得益于机械工业出版社刘星宇老师的推荐与鼓励，在此特致感谢。译者的几位在读研究生在全书的初稿形成、图表编辑等诸多方面给予了帮助，在此一并致谢。

由于译者水平有限、加之 AIML 领域新兴概念繁多，难免误译或词不达意，敬请读者赐教与原谅。

译者

| 原书前言

近年来，深度学习在视觉、语音、自然语言处理与理解等存在大量数据处理的难题中取得了空前的成功，进而引起了公司、高校和政府对该领域的研究兴趣，从而更加加速了该领域的发展。本书精选了深度学习当中一些高级内容。在本书中，这些高级理论通过介绍背景原理、挖掘概念背后的直观表述、使用 Keras 实现公式和算法并对实验结果进行检验的方式来进行讲述。

纵观今日，人工智能（Artificial Intelligence, AI）仍不是一个易于理解的领域。深度学习作为 AI 的子领域，面临着相同的境遇。尽管深度学习还不是一个成熟的领域，但很多现实世界的应用，例如基于视觉的检测和识别、产品推荐、语音识别与合成、节能、药品开发、金融和市场营销，已经成功使用了深度学习算法。此外，更多相关应用还会源源不断地被挖掘并实现。本书的目的就是尝试解释相关高级概念，并给出实现示例，让专家在其领域内识别出所需的应用目标。

尚未完全成熟的领域是一把双刃剑。一方面，它为探索和利用提供了诸多机会。深度学习中尚存大量未解决的问题。这些问题可转变为率先进入市场的机会，如产品开发、发布或认可。然而，另一方面，某些关键任务环境中很难信任一个尚未完全理解的领域。当被问及时，我们可以肯定地说，仅有极少数的机器学习工程师愿意去乘坐一架由深度学习系统所控制的自动驾驶飞机。要达到这种信任级别，还有很多工作需要去做。本书中所讨论的高级概念，很有可能为达到这种信任级别而扮演了重要的角色。

深度学习的每一本书都不可能完全覆盖整个领域，本书也是如此。在给定的时间和空间内，我们会接触到一些有趣的领域，例如检测、分割和识别、视觉理解、概率推理、自然语言处理和理解、语音合成以及自动机器学习。然而，本书相信，对这些精选领域的认知将有助于读者从事其他未涵盖的领域。

作为将要阅读本书的读者，需谨记，您选择了一个令人兴奋的领域，并可以对社会产生重大的影响。我们很有幸，能从事一份早晨起来便迫不及待想为之奋斗的工作。

本书适用的读者

本书面向希望深入了解深度学习高级主题的机器学习工程师和学生。书中所探讨的内容都通过 Keras 代码进行了补充说明。本书适合那些想了解如何将理论通过 Keras 转化成为可实际运行代码的读者。除了理解理论外，代码实现通常是将机器学习应用于实际问题的艰巨任务之一。

本书所覆盖的内容

第 1 章，Keras 高级深度学习入门，介绍了深度学习的相关概念，例如优化、正则化、损失函数、基本层和网络，以及这些概念在 Keras 中的实现。本章同时介绍了深度学习和

Keras 中的 Sequential API。

第 2 章，神经网络，讨论了 Keras 函数式 API 的应用，并使用函数式 API 在 Keras 中进行了检验，实现了两种广泛使用的深度网络架构——ResNet 和 DenseNet。

第 3 章，自编码器，涵盖了自编码器的通用网络架构，该架构用于发现输入数据的潜在表示。在 Keras 中讨论并实现了自编码器的两个示例应用，即去噪和色彩迁移。

第 4 章，生成对抗网络，讨论了深度学习最新的重大进展之一。生成对抗网络 (GAN) 用于生成以假乱真的新数据。本章介绍了 GAN 的原理，并在 Keras 中检验和实现了 GAN 的两个示例——DCGAN 和 CGAN。

第 5 章，改进的 GAN 方法，介绍了对基本 GAN 进行改进的算法。该算法解决了 GAN 训练方面的困难，并提高了所合成数据的感知质量。

第 6 章，分离表示 GAN，讨论了如何控制 GAN 所生成的合成数据的属性。如果潜表示可被分离，则属性可控制。在 Keras 中介绍并实现了两种解表征的技术，即 InfoGAN 和 StackedGAN。

第 7 章，跨域 GAN，覆盖了 GAN 的一个实际应用，它可将图像从一个域转换到另一个域，该过程称为跨域转换。

第 8 章，变分自编码器，讨论了深度学习近期的另一个重大进展。与 GAN 类似，变分自编码器 (VAE) 是一个用于产生合成数据的生成模型。但与 GAN 不同的是，VAE 专注于可解码的连续潜空间，该空间适用于变分推断。在 Keras 中涵盖并实现了 VAE 及其变种算法——CVAE 和 β -VAE。

第 9 章，深度强化学习，介绍了强化学习和 Q-Learning 的原理，提出了在离散动作空间实施 Q-Learning 的两种技术——Q-Table 更新和深度 Q 网络 (DQN)。在 Keras 中使用 Python 和 DAN 对所实现的 Q-Learning 方法在 OpenAI Gym 环境下进行了演示。

第 10 章，策略梯度方法，介绍了如何使用神经网络对强化学习中的决策策略进行学习。在 Keras 和 OpenAI Gym 环境下实现了四种方法，分别是 REINFORCE、基线 REINFORCE、Actor-Critic 和优势 Actor-Critic。本章中的示例演示了在连续动作空间中的策略梯度方法。

如何充分利用本书

- 深度学习和 Python: 读者应具备深度学习及其在 Python 中进行实现的基础知识。如果有使用过 Keras 实现深度学习算法的经验更好，但非必须。第 1 章介绍了深度学习的概念及其在 Keras 中的实现。
- 数学: 本书中讨论的内容假设读者熟悉大学级别的微积分、线性代数、统计和概率论。
- GPU: 本书中绝大多数 Keras 实现都需要 GPU。没有 GPU，将会花费大量时间（数小时至数天），因此，很多代码示例的执行将不切实际。本书中的示例尽可能控制所使用数据的大小，以尽量减少使用高性能计算机的可能。推荐读者至少使用 NVIDIA GTX 1060 的显卡辅助计算。
- 编辑器: 本书中的示例代码使用了 Ubuntu Linux 16.04 LTS、Ubuntu Linux 17.04 和

macOS High Sierra 中的 vim 进行编辑。任何支持 Python 的编辑器都可行。

- TensorFlow: Keras 需要一个后端。本书中的所有示例代码采用以 TensorFlow 为后端的 Keras 完成。请确保 GPU 驱动和 TensorFlow 都正确安装。
- GitHub: 为能通过示例和相关实验进行学习, 请使用 git pull 或 fork 命令从 GitHub 代码库中获取本书的代码集。获取成功后, 可进行检验、运行、更改并再次运行。可通过调整示例代码的方式去完成所有创造性的实验。这也是领会章节中所展示理论知识的唯一方法。如果您能为本书的 GitHub 代码库加星点赞, 我们将十分感激。

下载代码文件

本书的代码集存放在 GitHub, 可通过以下网址进行访问:

<https://github.com/PacktPublishing/Advanced-Deep-Learning-with-Keras>

下载彩色图片

我们也提供本书所涉及的截屏和流程图的彩色图片 PDF 文件。可从以下网址下载:

http://www.packtpub.com/sites/default/files/downloads/9781788629416_ColorImages.pdf

使用惯例

本书的代码示例都使用 Python 编写。具体而言, 是使用 Python3 编写。代码模式基于 vim 的 syntax highlighting 插件。以下面代码为例:

```
def encoder_layer(inputs,
                  filters=16,
                  kernel_size=3,
                  strides=2,
                  activation='relu',
                  instance_norm=True):
```

""" 构建一个由 Conv2D-In-LeakyReLU 组成的一般编码器层。

IN 为可选项, LeakyReLU 可被 ReLU 所替代

"""

```
conv = Conv2D(filters=filters,
              kernel_size=kernel_size,
              strides=strides,
              padding='same')
x = inputs
if instance_norm:
    x = InstanceNormalization()(x)
if activation == 'relu':
    x = Activation('relu')(x)
else:
    x = LeakyReLU(alpha=0.2)(x)
x = conv(x)
return x
```