

青少年AI学习之路：从思维到创造

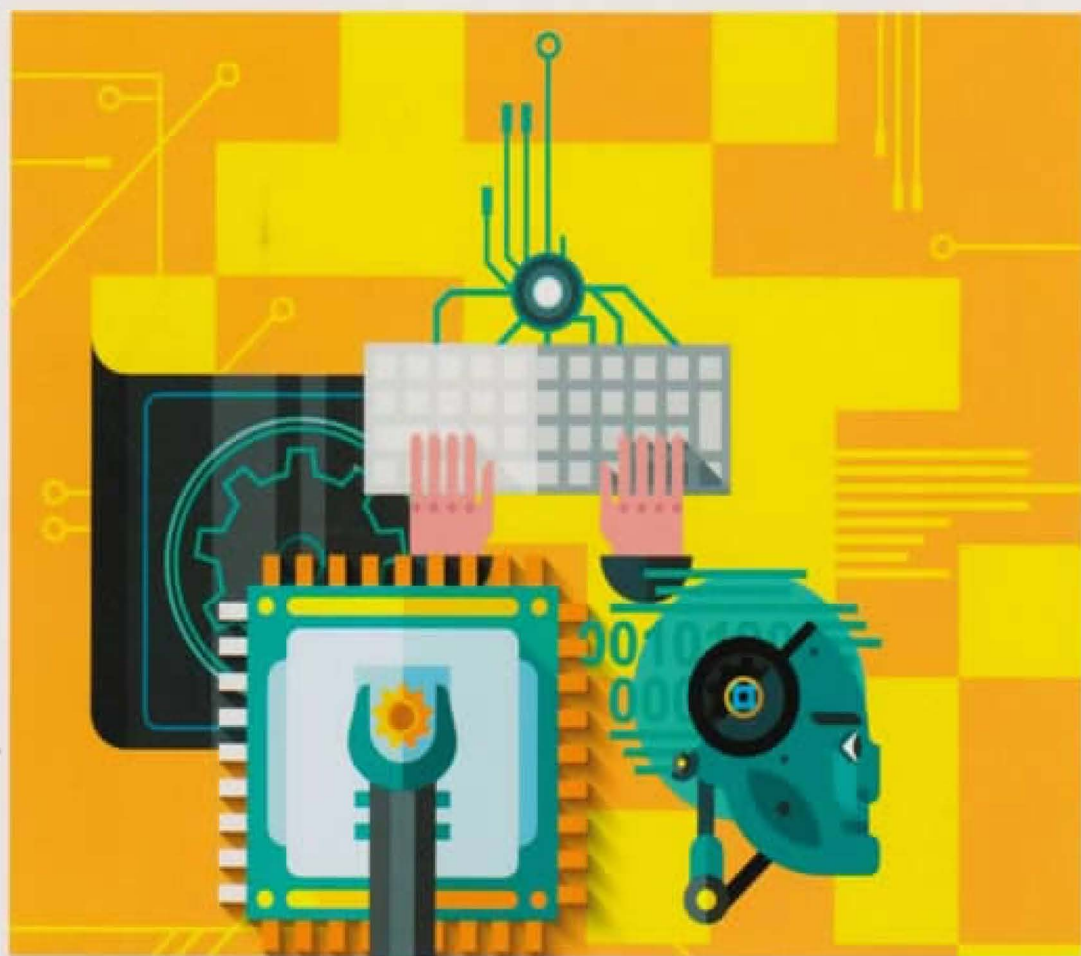
4

丛书主编：俞 勇

人工智能实践

动手做你自己的AI

编著：张惠楚 张伟楠





AI

青少年AI学习之路——从思维到创造

人工智能实践

——动手做你自己的AI

俞 勇 主编



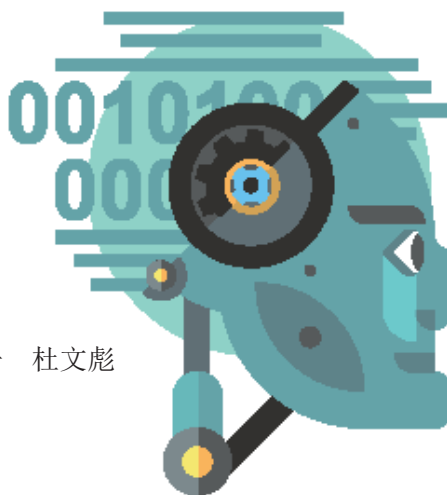
上海科技教育出版社

图书在版编目(CIP)数据

人工智能实践:动手做你自己的AI/俞勇主编. —
上海:上海科技教育出版社,2019.8
(青少年AI学习之路.从思维到创造)
ISBN 978-7-5428-7082-7

I. ①人… II. ①俞… III. ①人工智能-青少年读物
IV. ①TP18-49

中国版本图书馆CIP数据核字(2019)第167787号



责任编辑 丁 祯 王丹丹 杜文彪
装帧设计 杨 静

青少年AI学习之路:从思维到创造
人工智能实践——动手做你自己的AI
丛书主编 俞 勇

出版发行 上海科技教育出版社有限公司
(上海市柳州路218号 邮政编码200235)

网 址 www.sste.com www.ewen.co
经 销 各地新华书店
印 刷 上海昌鑫龙印务有限公司
开 本 889×1194 1/16
印 张 12
版 次 2019年8月第1版
印 次 2019年8月第1次印刷
书 号 ISBN 978-7-5428-7082-7/G·4124
定 价 96.00元



总序

清晰记得，2018年1月21日上午，我突然看到手机里有这样一则消息“【教育部出大招】人工智能进入全国高中新课标”，我预感到我可以为此做点事情。这种预感很强烈，它也许是我这辈子最后想做、也是可以做的一件事，我不想错过。

从我1986年华东师范大学计算机科学系硕士毕业来到上海交通大学从教，至今已有33年。其间做了三件引以自豪的事，一是率领上海交通大学ACM队参加ACM国际大学生程序设计竞赛，分别于2002年、2005年及2010年三次获得世界冠军，创造并保持了亚洲纪录；二是2002年创办了旨在培养计算机科学家及行业领袖上海交通大学ACM班，成为中国首个计算机特班，从此揭开了中国高校计算机拔尖人才培养的序幕；三是1996年创建了上海交通大学APEX数据与知识管理实验室（简称APEX实验室），该实验室2018年度有幸跻身全球人工智能“在4个领域出现的高引学者”世界5强（AMiner每两年评选一次全球人工智能“最有影响力的学者奖”）。出自上海交通大学的ACM队、ACM班和APEX实验室的杰出校友有：依图科技联合创始人林晨曦、第四范式创始人戴文渊、流利说联合创始人胡哲人、字节跳动AI实验室总监李磊、触宝科技联合创始人任腾、饿了么执行董事罗宇龙、森亿智能创始人张少典、亚马逊首席科学家李沐、天壤科技创始人薛贵荣、宾州州立大学终身教授黎珍辉、加州大学尔湾分校助理教授赵爽、明尼苏达大学双子城分校助理教授钱风、哈佛大学医学院助理教授李博、新加坡南洋理工大学助理教授李翼、伊利诺伊大学芝加哥分校助理教授孙晓锐和程宇、卡耐基梅隆大学助理教授陈天奇、乔治亚理工学院助理教授杨笛一、加州大学圣地亚哥分校助理教授商静波等。

我想做的第四件事是创办一所民办学校，这是我的终极梦想。几十年的从教经历，使得从教对我来说已不只是一份职业，而是一种习惯、一种生活方式。当前，人工智能再度兴起，国务院也发布了《新一代人工智能发展规划》，且中国已将人工智能上升为国家战略。于是，我创



建了伯禹教育，专注人工智能教育，希望把我多年所积累的教育教学资源分享给社会，惠及更多需要的人群。正如上海交通大学党委书记姜斯宪教授所说，“你的工作将对社会产生积极的影响，同时也是为上海交通大学承担一份社会责任”。也如上海交通大学校长林忠钦院士所说，“你要做的工作是学校工作的延伸”。我属于上海交通大学，我也属于社会。

2018年暑假，我们制订了“青少年AI实践项目”的实施计划。在设计实践项目过程中，我们遵循青少年“在玩中学习，在玩中成长”的理念，让青少年从体验中感受学习的快乐，激发其学习热情。经过近半年的开发及完善，我们完成了数字识别、图像风格迁移、文本生成、角斗士桌游及智能交通灯等实践项目的设计，取得了非常不错的效果，并编写了项目所涉及的原理、步骤及说明，准备将其编成一本实践手册给青少年使用。但是，作为人工智能的入门读物，光是一本实践手册远远满足不了读者的需要，于是本套丛书便应运而生。

本套丛书起名“青少年AI学习之路：从思维到创造”，共有四个分册。

第一册《从人脑到人工智能：带你探索AI的过去和未来》，从人脑讲起，利用大量生动活泼的案例介绍了AI的基本思维方式和基础技术，讲解了AI的起源、发展历史及对未来世界的影响。

第二册《人工智能应用：炫酷的AI让你脑洞大开》，从人们的衣食住行出发，借助生活中的各种AI应用场景讲解了数十个AI落地应用实例。

第三册《人工智能技术入门：让你也看懂的AI“内幕”》，从搜索、推理、学习等AI基础概念出发解析AI技术，帮助读者从模型和算法层面理解AI原理。

第四册《人工智能实践：动手做你自己的AI》，从玩AI出发，引导读者从零开始动手搭建自己的AI项目，通过实践深入理解AI算法，体



验解剖、改造和创造AI的乐趣。

本套丛书的特点：

■ 根据青少年的认知能力及认知发展规律，以趣味性的语言、互动性的体验、形象化的解释、故事化的表述，深入浅出地介绍了人工智能的历史发展、基础概念和基本算法，使青少年读者易学易用。

■ 通过问题来驱动思维训练，引导青少年读者学会主动思考，培养其创新意识。因为就青少年读者来说，学到AI的思维方式比获得AI的知识更重要。

■ 用科幻小说或电影作背景，并引用生活中的人工智能应用场景来诠释技术，让青少年读者不再感到AI技术神秘难懂。

■ 以丛书方式呈现人工智能的由来、应用、技术及实践，方便学校根据不同的需要组合课程，如科普性的通识课程、科技性的创新课程、实践性的体验课程等。

2019年1月15日，我们召集成立了丛书编写组；1月24日，讨论了丛书目录、人员分工和时间安排，开始分头收集相关资料；3月6日，完成了丛书1/3的文字编写工作；4月10日，完成了丛书2/3的文字编写工作；5月29日，完成了丛书的全部文字编写工作；6月1日—7月5日，进行3—4轮次交叉审阅及修改；7月6日，向出版社提交了丛书的终稿。在不到6个月的时间里，我们完成了整套丛书共4个分册的编写工作，合计100万字。

在此，特别感谢张伟楠博士，他在本套丛书编写过程中给予了很多专业指导，做出了重要的贡献。

感谢我的博士生龙婷、任侃、沈键和张惠楚，他们分别负责了4个分册的组织与编写工作。

感谢我的学生吴昕、戴心仪、周铭、粟锐、杨正宇、刘云飞、卢冠松、宋宇轩、茹栋宇、吴宪泽、钱利华、周思锦、秦佳锐、洪伟峻、陈铭城、朱耀明、杨阳、卢冠松、陈力恒、秋闻达、苏起冬、徐逸凡、侯



博涵、蔡亚星、赵寒焯、任云玮、钱苏澄及潘哲逸等，他们参与了编写工作，并在如此短的时间内，利用业余时间进行编写，表现了高度的专业素质及责任感。

感谢王思捷、冯思远全力以赴开发实验平台。

感谢陈子薇为本套丛书绘制卡通插图。

感谢所有支持编写的APEX实验室成员及给予帮助的所有人。

感谢所引用图书、论文的编者及作者。

同时，还要感谢上海科技教育出版社对本丛书给予的高度认可与重视，并为使丛书能够尽早与读者见面所给予的鼎力支持与帮助。

本套丛书的编写，由于时间仓促，其中难免出现一些小“bug”（错误），如有不当之处，恳请读者批评指正，以便再版时修改完善。

过去未去，未来已来。在互联网时代尚未结束，人工智能时代已悄然走进我们生活的当前，应该如何学习、如何应对、如何创造，是摆在青少年面前需要不断思考与探索的问题。希望本套丛书不仅能让青少年读者学到AI的知识，更能让青少年读者学到AI的思维。

愿我的梦想点燃更多人的梦想！

俞 勇

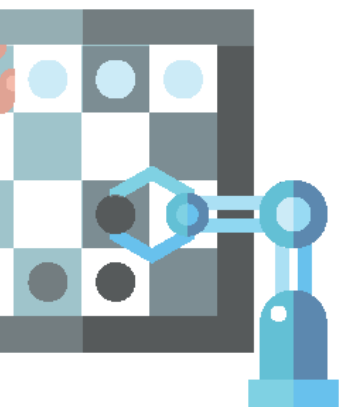
2019年8月8日于上海

目录

AI



前言	1
第1部分 图像识别	
第一章 图像识别的基础知识	6
第二章 K近邻算法	18
第三章 卷积神经网络	25
第2部分 图像风格迁移	
第四章 图像风格迁移基础知识	48
第五章 色彩的模仿——K近邻算法	49
第六章 像素级别梯度优化	54
第七章 生成对抗网络	68
第3部分 文本生成	
第八章 文本生成基础知识	92
第九章 n元语法模型——从数数开始	96
第十章 循环神经网络	101
第十一章 编码与解码——seq2seq	111





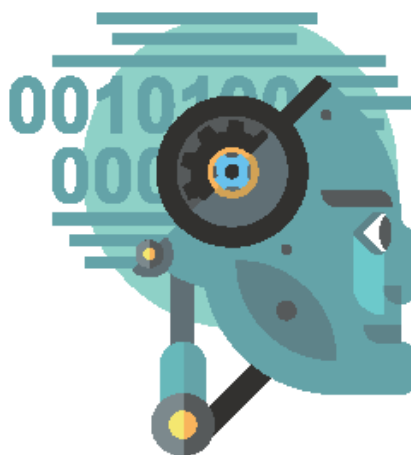
第4部分 角斗士棋

第十二章	角斗士棋基础知识	120
第十三章	常用的搜索算法	125

第5部分 红绿灯调度

第十四章	红绿灯调度基础知识	144
第十五章	马尔可夫决策过程	147
第十六章	Q-learning	152
第十七章	Deep Q-Network	157
第十八章	智能红绿灯的代码实现	162

结语		179
附录一	“思考与实践”解答参考	180
附录二	参考文献	182



前言

AI

“Talk is cheap, show me the code”, 大名鼎鼎的Linux系统创始人林纳斯·本纳第克特·托瓦兹 (Linus Benedict Torvalds) 在2000年给编写Linux系统核心的成员群发了一封邮件, 里面提到了这句话, 翻译成中文就是“多说无益, 给我看你的代码”。这句名言在计算机科学界广为流传并奉为金句。计算机作为一门实践性很强的学科, 理解理论后动手写出可运行的、能看到成果的代码才是硬道理。人工智能是计算机科学中应用场景最多的子学科之一, 这更需要学习者在理解原理后, 动手实践、应用算法解决实际问题。正所谓“纸上得来终觉浅, 绝知此事要躬行”。

所以在2018年的夏天, 俞勇教授组建了这支编写团队, 开始尝试编写针对青少年的项目实践教程。最初的内容包括图像上色、英语语法单选题与角斗士桌游。在教程中, 我们努力用最浅显的语言将复杂的人工智能算法讲明白, 并且提供了相应的代码和可在线交互的展示。一些中学生看了我们编写的这些教程之后给出了一致的好评, 这也让我们产生了将其拓展为一套给青少年看的关于人工智能的丛书的想法。2019年初, 我们给这套丛书加入了更多关于实践的内容, 引导读者一边理解原理, 一边编写代码, 理论与实践相结合地认识人工智能。在本书中, 我们精心安排了兼顾项目的有趣程度、使用技术的流行程度, 以及知识的涵盖程度的项目。在编写方式上, 我们也努力精简知识, 仅介绍与项目任务相关

的必要知识, 希望能够让零基础的读者, 仅通过阅读本书也能快速地了解用到的算法。对每一种算法, 我们都会给出Python代码实现, 配合注释与讲解, 引领读者一步一步编写人工智能算法。

五个项目各由我们编写团队中的一到两位编者负责, 每周俞勇教授与张伟楠教授





都会与编者开例会、审阅进度，并讨论可以改进的地方。在编写教材的同时，我们还与实验平台的开发人员合作，提供了线上可交互的展示，方便读者理解项目与算法。编写本套丛书前后历时近一年，俞勇教授、张伟楠教授与编者投入了大量的心血，希望能帮助青少年理解并自己实现人工智能算法，培养对人工智能的兴趣，为将来更深入地探索人工智能打下基础。

从本书中可以读到

我们将通过五个有趣的项目，带领大家在理解原理的同时，动手编写代码，实现自己的人工智能应用。本书的每一部分对应一个项目，它们分别是：

图像识别：图像识别是近几年人工智能浪潮中，发展最快、技术最成熟、落地最早的算法，如生活中常见的车牌识别、人脸识别等都是基于图像识别的技术。在这个项目中，你将了解机器处理图像数据的方法，学习并编写图像识别的两个经典算法：K近邻算法与卷积神经网络，并解决手写数字识别以及物体识别问题。

图像风格迁移：图像风格迁移是一个非常有趣的项目。利用该技术，人们可以将任意一张图片转换成其他风格的图片。一些热门的图像处理软件中就提供了类似功能，比如将照片转化成马赛克风格，甚至是梵高的画作风格。在本项目中，你将尝试用3种图像风格迁移算法进行实践。

文本生成：与文本相关的智能技术随着人工智能浪潮也得到了快速发展。在这个项目中，你将了解机器处理文本数据的方法，并学习和编写3个用于文本生成的算法。使用这些算法，可以让机器生成古诗甚至是文章。

角斗士：角斗士是一款热门的桌面游戏。在这个项目中，你将学习并使用各类搜索算法，制作角斗士游戏的AI。与上面三个项目不同的是，本项目中你将第一次接触到教机器作出决策，让机器作出正确的决策是迈向真正智能的重要一步。

红绿灯调度：大城市中如何缓解交通拥堵是一个广受关注的问题。在这个项目中，就是使用强化学习算法来优化红绿灯的控制，以达到缓解拥堵的目的。强化学习是近几年非常热门的领域，大名鼎鼎的AlphaGo就是使用强化学习算法战胜了所有人类棋手。你将了解强化学习算法的基本概念，并尝试编写强化学习算法来控制路口的红绿灯。



如何使用这本书

本书没有严格的阅读顺序要求，你可以选择自己感兴趣的项目直接阅读，若提到了其他项目中的概念均会给出提示。

除了算法原理，本书还包含了大量示例代码。如果想要理解这些代码，你需要有一些Python基础编程知识，包括变量、循环、函数、类、库的使用等。如果想要入门Python，你可以访问伯禹学堂中的Python基础课程 (www.boyu.ai/pythonbasics)。本书的所有代码都基于Anaconda3-5.2.0版本用Python3实现，Anaconda是一个免费的Python发行版，包含了大量常用的库，非常易于安装，强烈建议你在学习过程中安装使用。项目中的代码示例可以在 github.com/boyuai/textbook 下载。另外，本书配有实验平台 (www.boyu.ai/playground)，为书中涉及的项目提供一个交互式的学习体验。对每个任务，我们提供了体验、调试、理解源码、源码填空、改进源码5个阶段。不同基础背景的读者可以选择不同的阶段入门：零基础的读者可以通过体验算法了解人工智能的魅力，基础较好的读者可以自己编写代码改进算法。希望你通过阅读本书，至少可以达到理解项目内容、掌握算法原理和能够进行基本的算法实现的程度。若你对学习人工智能的积极性非常强，希望你能在实验平台上多多地尝试改进代码，并创造属于自己的人工智能算法。

本书的内容仅包含五个项目必备的知识以及代码。如果希望系统地学习人工智能算法，建议阅读本套丛书的第三本《人工智能技术入门——让你也看懂的AI“内幕”》；如果对人工智能的发展历史与未来展望感兴趣，建议阅读本套丛书的第一本《从人脑到人工智能——带你探索AI的过去和未来》；如果对人工智能在生活中的具体应用感兴趣，建议阅读本套丛书的第二本《人工智能应用——酷炫的AI让你脑洞大开》。

致谢

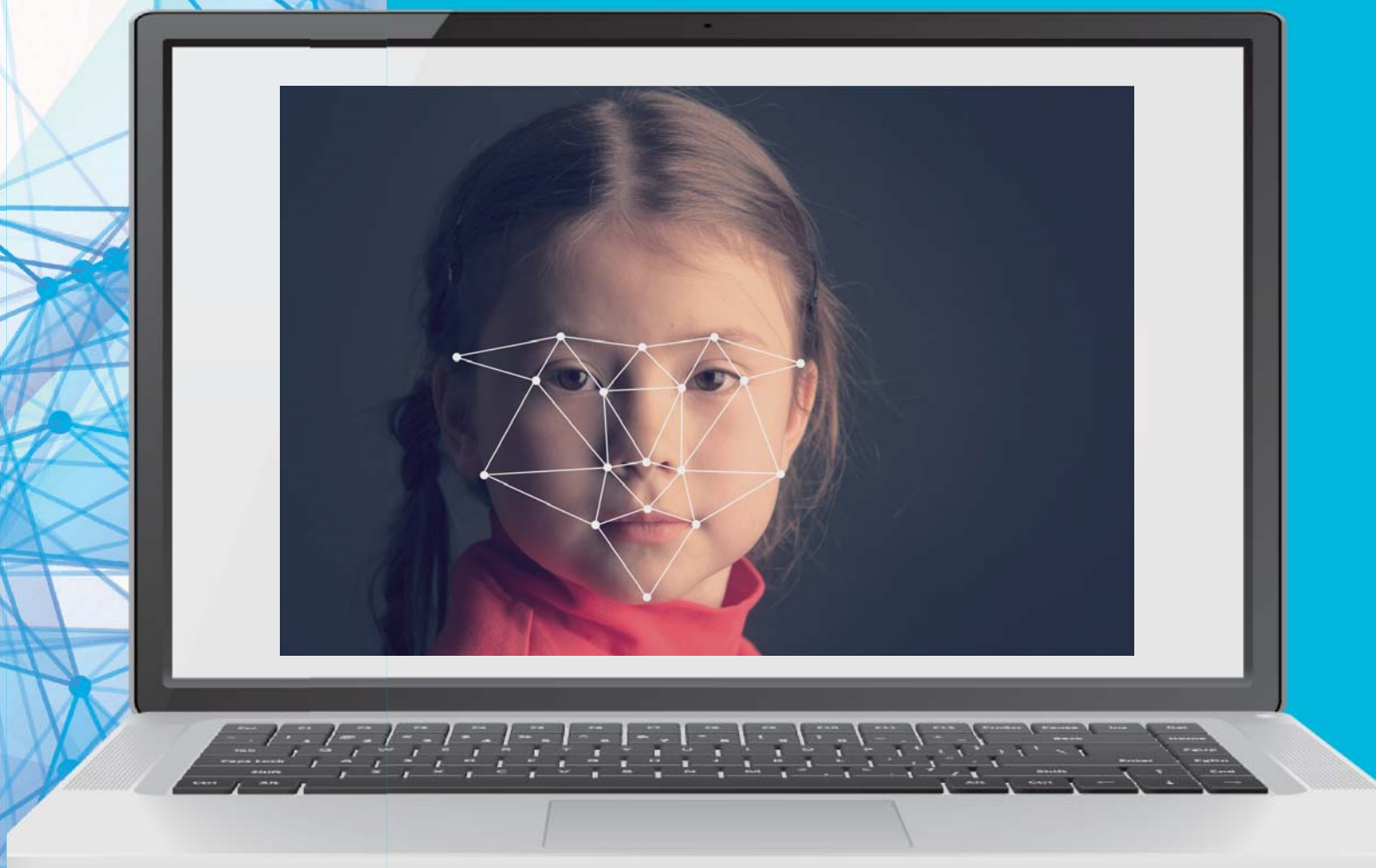
参与本书编写的主要人员有16位，俞勇教授策划并确定本书架构、内容组织及审核，张伟楠博士对全书内容进行专业指导及审核，张惠楚与秋闻达编写第一部分，苏起冬与徐逸凡编写第二部分，侯博涵与蔡亚星编写第三部分，赵寒焱与任云玮编写第四部分，钱苏澄编写第五部分。感谢陈子薇绘制本书的卡通插图，感谢王思捷、冯思远开发实验平台，并感谢任侃、沈键及潘哲逸在编写时提供的帮助。

第1部分

图像识别

人们每时每刻都在无意识地进行图像识别：识别物体，拿起它们；识别障碍物，避开它们；识别文字，进行阅读理解；识别人脸，进行身份判断。对于人类来说，图像识别是一个自然又简单的过程，从眼睛接受光信号到大脑识别出图像，几乎不需要思考。而对于计算机来说，图像识别就没那么容易了。从计算机被发明以来，科学家们一直期望让计算机能自动识别出图像内容，但是由于算法和硬件方面的限制，计算机的图像识别能力一直远低于人类的水平。直到近几年，随着深度学习的发展，图像识别的效率和准确率得到了大幅度的提升。在近

几年人工智能浪潮中，图像识别是发展最快、技术最成熟、落地最早的算法。车牌识别、人脸识别等诸多基于图像识别的技术正在改变着人们的生活。在本部分中，我们将一起了解机器处理图像数据的原理，编写图像识别的两个经典算法：K近邻算法与卷积神经网络，并解决手写数字识别以及物体识别问题。



第一章 图像识别的基础知识

在本章中，我们将了解什么是图像识别、图像在计算机中的表达、图像识别的流程以及图像识别中类别的表达方式。此外，我们将编写代码探索图像识别领域中的一些经典数据集。

一、图像识别简介

图像识别，顾名思义就是识别图像的内容。严谨地讲，就是给计算机一张图像，让其判断出图像中的内容。例如给一张手写数字的图像，判断出它是数字几；或是给一张动物的图像，判断是猫还是狗；或是给一张班级里同学的照片，判断出他/她的名字。



识别数字 / 识别动物

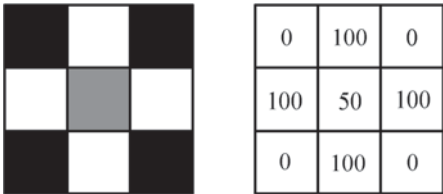
二、计算机如何表示图像

为了让计算机识别图像，我们首先得让计算机“看”到图像。所以在学习图像识别的算法之前，我们要先了解计算机“眼”中的图像是怎样的。

对于用计算机处理图像，你可能并不陌生。使用Windows操作系统自带的画图软件，可以进行一些简单的图像处理工作。此外还有Lightroom、Photoshop之类功能更为强大的图像处理软件，可见计算机处理图像的能力非常强大。你可能听说过，计算机世界里的通用语言是二进制，在计算机的“眼”中，一切都是数字。那么，我们怎样把一张图像转化为机器能处理的数字形式呢？

计算机中的图像主要有两种——位图和矢量图（矢量图在本书中不做介绍）。其中位图是使用像素矩阵来表示图像的一种方法。我们可以把图像分割成一个个小方块，每一个小方块

就是一个像素。我们常听到的广告语“2 000万像素，照亮你的美”，所表达的就是图像含有2 000万个小方块。每一个像素都可以描绘一些信息，通常像素越高，图像包含的信息就越多，图像也就越清晰。以黑白照片为例，对于其上每一个像素，都可以用一个数来描述它的灰度（或者亮度）。如果把纯白色记作100，纯黑色记作0，不同的灰色用一个1~99之间的数字表示，这样就把一张图像转化成了一个矩阵，如下图所示。通常灰度会用0~255中的一个数字来表示，这里为了方便理解所以我们用了0~100来举例。

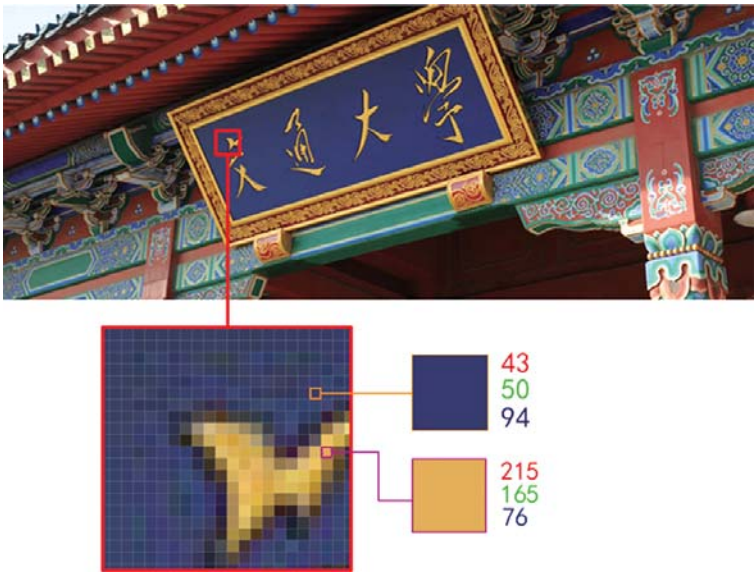


一张含有3×3=9个像素的黑白位图

根据光的三原色原理，我们知道任何颜色的光都可以通过红、绿、蓝三种原色叠加生成，因此我们可以通过描述颜色中红、绿、蓝三种成分的多少来唯一确定一种颜色。换句话说，对于图像中的每个像素，我们都可以用一个包含三个数的数组来描述它的颜色，这就是最常见的RGB色彩模型。RGB分别表示Red、Green和Blue，这三种颜色成分的多少可以分别用一个0~255的数字来表示，以下简记为(R, G, B)。例如(255, 0, 0)，表示这种颜色中红色的成分最多，没有绿色和蓝色的成分，是纯红色，同理(0, 255, 0)就是纯绿色。我们知道红光和绿光可以叠加成为黄光，那么(255, 255, 0)就应该是纯黄色。(0, 0, 0)表示什么颜色都没有，就是黑色，(255, 255, 255)表示三种颜色的“光强”都达到最大，叠加成为白色。



三原色的叠加



用RGB色彩模型来表示一张位图中的颜色



向量、矩阵、张量与通道

图像的存储方式涉及四个专业术语。

向量：用最简单的方式理解，一个向量其实就是一组数字。如下左图所示。

矩阵：矩阵也是一组数字，与向量有所不同，它由行与列构成。矩阵中的数字都可以用 x 行 y 列来定位。一个矩阵的形状由两个整数定义，如下中图的矩阵就是一个大小为 $(5, 5)$ 的矩阵。矩阵可以用来描述灰度图。

张量：张量是一种更广义的概念。如果我们希望排列方式不仅有行和列，还有更多的维度，那么就需要用到张量。向量实际上就是一阶张量，矩阵就是二阶张量，如下右图是一个三阶张量，它的大小是 $(5, 5, 3)$ 。三阶张量可以用来描述RGB彩色图。

通道：一张彩色图像通常由一个三阶张量表示，而除了行与列的维度，另外一个维度称为通道。如下右图所示，一张RGB格式的图像，通道数量为3。

1	2	1	1	0
---	---	---	---	---

向量
1阶张量

1	2	1	1	0
2	1	2	2	1
0	1	2	1	2
1	2	1	2	0
0	0	0	1	1

矩阵
2阶张量

3	1	0	1	0		
1	2	1	1	2	0	
2	1	2	2	0	1	0
2	0	0	1	1	3	1
1	1	3	2	1	1	2
0	3	1	2	1	3	
2	0	2	0	1		

通道

3阶张量

向量、矩阵、张量与通道



思考与实践

1.1 如果有一张 100×100 像素的RGB彩色图像，计算机存储这张图像需要存多少个数字呢？