

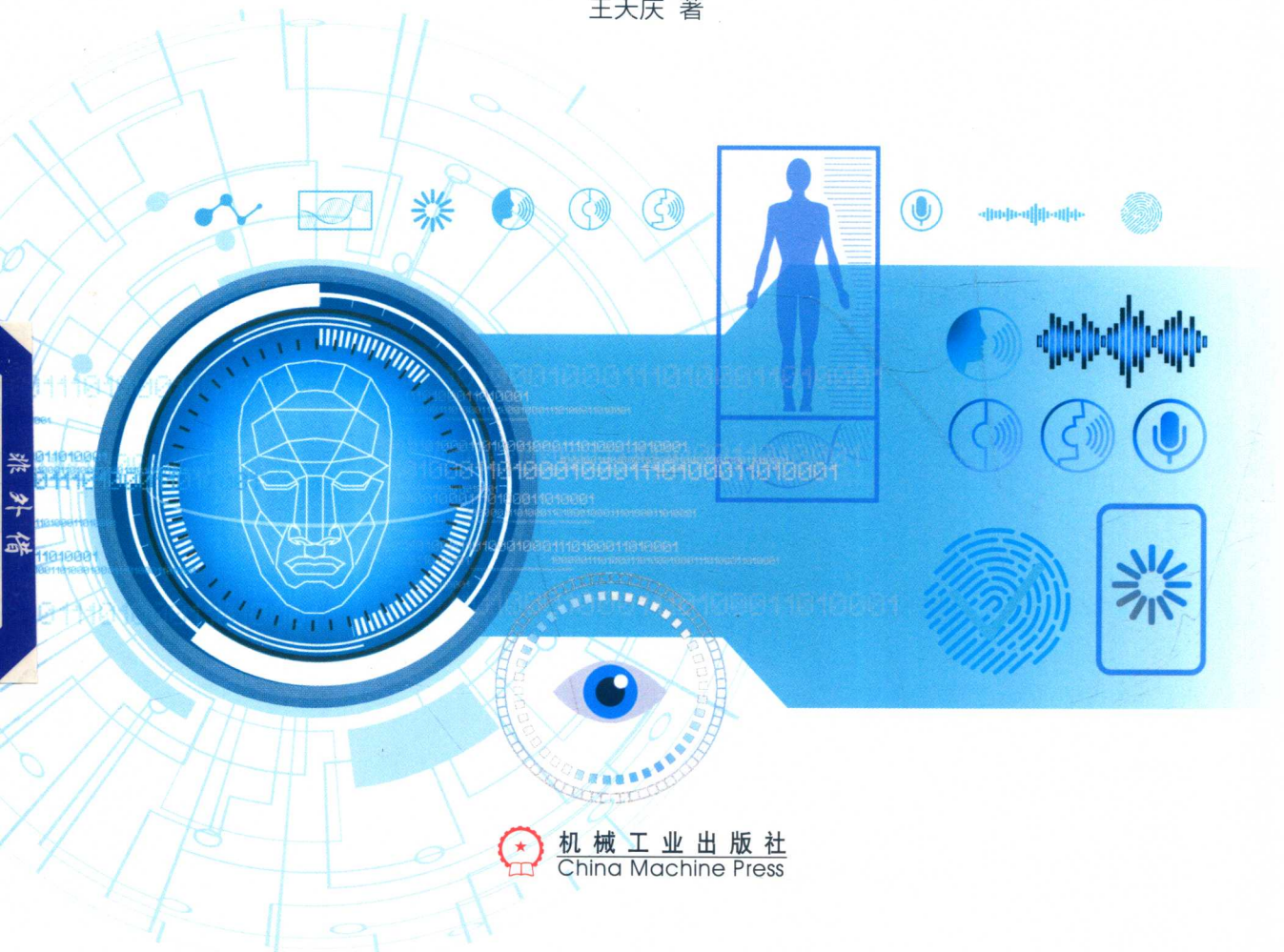
作者就职于某世界100强企业，长期从事人工智能、数据科学、分布式系统等领域的研发工作，在人脸识别领域有丰富的实践经验
涵盖机器学习、深度学习、计算机视觉、人脸识别等方面的原理、技术和算法，教读者从零实现一个工程级的人脸识别引擎

Face Recognition with Python

Python人脸识别

从入门到工程实践

王天庆 著



机械工业出版社
China Machine Press

■ ■ ■ 智能系统与技术丛书

Face Recognition with Python

Python人脸识别

从入门到工程实践

王天庆 著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

Python 人脸识别：从入门到工程实践 / 王天庆著. —北京：机械工业出版社，2019.4
(智能系统与技术丛书)

ISBN 978-7-111-62385-4

I. P… II. 王… III. ①软件工具 - 程序设计 ②人脸识别 - 研究 IV. ① TP311.561
② TP391.41

中国版本图书馆 CIP 数据核字 (2019) 第 058981 号

Python 人脸识别：从入门到工程实践

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：张锡鹏

责任校对：殷虹

印刷：北京市兆成印刷有限责任公司

版次：2019 年 4 月第 1 版第 1 次印刷

开本：186mm × 240mm 1/16

印张：16.25

书号：ISBN 978-7-111-62385-4

定价：69.00 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

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

HZBOOKS | 华章IT | Information Technology



P R E F A C E

前 言

为什么要写这本书

人类在浩瀚无垠的宇宙中只是一种渺小的存在，“科技之树”中触手可及的果实已经采摘殆尽，而在树顶的果实又是何其难以获得。在很多领域，人类当前已经难以取得像以往一样巨大的进步。在此情形下，信息科学、数学等基础科学便是驱动其他领域发展的主要动力，而人工智能则是这个领域的新星！

人工智能在近些年发展可以说是目共睹。伴随着学术发展的突飞猛进，工业界基于人工智能的应用呈现“井喷”之势，一些公司甚至提出了“All in AI”的口号，这其中比较典型的一种应用便是人脸识别。

从历史的角度来看，人们对人脸识别的探索也是比较早的，但是，人脸识别从实验室走出来，来到人们的生活中，却只是近些年才发生的事情。深度学习在计算机视觉领域应用后，使得人脸识别的精度逼近乃至超越人工水平。如果将人脸识别发展的进程用函数曲线来拟合的话，我觉得 Sigmoid 函数或许是一个不错的选择（我们会在本书的正文中介绍该函数）。

笔者在做人脸识别相关应用时，曾经面临业务压力大、无从下手、做成的模型预测质量差等一系列问题，同时有感于市面上难以找到从工程角度介绍人脸识别实现原理和方法的资料。随着自己不断地学习和尝试，模型的性能也一点点得到了改善。但是，生产环境与手写的 Demo 级代码还不一样，需要考虑的业务细节有很多。在反复的试错中，我设计的生产系统也从 0 到 1 实现了上线，并且在保证业务高可用的前提下实现了生产成本的最小

化，并逐步替换掉了原有云服务，为企业节约了大笔资金。

后来，为了帮助更多与我有类似经历的人，我将人脸识别应用开发中的琐碎细节整理出来，希望本书能够帮助到更多从事此行业的人。

读者对象

- 对机器学习、人工智能感兴趣的读者
- 对计算机视觉、深度学习感兴趣的读者
- 对人脸识别感兴趣的读者
- 希望用人脸识别技术完成课程设计的高校学生
- 人工智能或人脸识别相关产品经理
- 从事软件研发的技术工作者
- 开设相关课程的大专院校教师

本书特色

- 原理讲解通俗易懂。

本书在成文时，参考了许多专业资料，并从初学者的角度出发尽可能清楚地表达原理的含义，尝试用通俗直白的语言描述复杂深奥的数学概念，尽量少使用公式或数学语言来描述问题。

- 注重原理更注重实践。

本书虽然强调了人脸识别在工程上的重要性，但是也没有忽视对原理的讲解，对于重点内容给出了详细的文献出处，便于读者进一步加强对该内容的理解。

- 注重对实践的讲解。

本书讲解了人脸识别模型的实现原理和特点，更重要的是强调了工程上的实践能力，对工程场景中常见的问题进行了分析与探讨。

如何阅读本书

本书分为 3 部分：

第1部分为基础篇，包括第1~4章。介绍了基本知识与概念，包括与人脸识别紧密相关的机器学习与数学知识、计算机视觉以及 OpenCV 库等。

第2部分为应用篇，包括5~6章。具体讲解了深度学习的框架 Keras，以及如何通过该框架设计和实现一个人脸识别引擎。

第3部分为拓展篇，包括7~8章。人脸识别引擎做好后，无论在本地验证多么优秀，能否将其用起来才是最关键的。在这一部分中，讲解了人脸识别系统的部署与图像检索知识。

参考文献部分包含了人脸识别领域经典的研究成果和知名著作，感兴趣的读者可以根据情况有选择地阅读。

本书的第1部分相对比较基础，有经验的读者可以直接从第2部分开始阅读。在阅读的过程中，更重要的是培养一种工程实践能力。因此，希望读者一定不要只看理论（这样其实不如去看论文更加直接），而是要亲自动手实践一下。

勘误和支持

由于作者的水平有限，编写时间仓促，书中难免会出现一些错误或者不准确的地方，恳请读者批评指正。书中的全部源文件可以从 GitHub 网站上下载，网址如下：

<https://github.com/wotchin/SmooFaceEngine>

关于本书中的错误，请读者在 GitHub 中提交 issue，笔者会将从读者处汇集到的错误及时整理和修改，同样更新在该页面中。

如果读者有更多的宝贵意见，也欢迎发送邮件至邮箱 wotchin@hotmail.com，或是在 Github 上留言。期待能够得到你的真挚反馈。

致谢

本书在成文过程中，有幸得到唐文根、田瑞川、郭丹等同仁的宝贵建议，特别是唐文根博士在百忙之中完成了本书第2章中的部分内容，同时在本书初稿完成之际又校验了全书的学术性问题。没有几位同仁的支持，本书难以与读者见面。

感谢 Github 社区中开源人脸识别项目的作者们，没有你们对开源事业的支持、对技术

的热爱，我也很难坚持完成此书。这些作者包括但不限于：oarriaga、wilderrodrigues、krunal704、alexisfcote、Hrehory、xiangrufan、Joker316701882、wyliu、happyneer、bojone、hao-qiang。

感谢机械工业出版社华章公司的编辑杨福川和张锡鹏，在这段时间中他们始终支持我的写作，正是他们的鼓励和帮助使得我顺利地完成书稿。

感谢我的家人，感谢众多热爱人脸识别技术的朋友们！

王天庆

CONTENTS

目 录

前言

第 1 章 人脸识别入门 1

1.1 人脸识别概况 1

1.1.1 何为人脸识别 1

1.1.2 人脸识别的应用 2

1.1.3 人脸识别的目标 4

1.1.4 人脸识别的一般方法 ... 5

1.2 人脸识别发展状况 8

1.2.1 人脸识别历史沿革 8

1.2.2 DT 时代的呼唤 10

1.2.3 计算机视觉的新起点 ... 10

1.3 本章小结 12

第 2 章 数学与机器学习基础 ... 13

2.1 矩阵 13

2.1.1 矩阵的形式 13

2.1.2 行列式 14

2.1.3 转置 15

2.1.4 矩阵的一般运算 15

2.2 向量 17

2.2.1 向量的形式 18

2.2.2 向量的点乘 18

2.2.3 向量的范数 19

2.3 距离度量 19

2.3.1 欧式距离 19

2.3.2 曼哈顿距离 20

2.3.3 余弦距离 20

2.3.4 汉明距离 21

2.4 卷积 22

2.4.1 一维卷积 22

2.4.2 二维卷积 23

2.5 机器学习基础 25

2.5.1 机器学习类别 25

2.5.2 分类算法 26

2.6 本章小结 38

第 3 章 计算机视觉原理与 应用 39

3.1 计算机视觉介绍 39

3.2 颜色模型 40

3.2.1 彩色图像 40

3.2.2 灰度图像与二值图像 ... 42

3.3	信号与噪声	44	第5章 深度学习与 Keras 工程	
3.3.1	信号	44	实践	87
3.3.2	噪声	45	5.1 深度学习介绍	87
3.4	图像滤波	45	5.2 Keras 框架简介	89
3.4.1	均值滤波	45	5.3 Keras 的使用方法	91
3.4.2	中值滤波	47	5.3.1 深度学习的原理	91
3.5	图像的几何变换	47	5.3.2 Keras 神经网络堆叠的 两种方法	92
3.5.1	平移	48	5.4 常用的神经网络层	96
3.5.2	旋转	49	5.4.1 全连接层	96
3.5.3	缩放	50	5.4.2 二维卷积层	98
3.6	图像特征	50	5.4.3 池化层	100
3.6.1	灰度直方图	50	5.4.4 BN 层	103
3.6.2	LBP 特征	51	5.4.5 dropout 层	105
3.6.3	Haar 特征	52	5.4.6 flatten 层	106
3.6.4	HOG 特征	54	5.5 激活函数	108
3.7	本章小结	56	5.5.1 Sigmoid 激活函数	108
第4章	OpenCV 基础与 应用	58	5.5.2 Softmax 激活函数	109
4.1	OpenCV 介绍	58	5.5.3 ReLU 激活函数	110
4.2	科学计算库 Numpy	59	5.5.4 Keras 中激活函数的 使用	111
4.2.1	array 类型	60	5.6 优化器	112
4.2.2	线性代数相关	62	5.6.1 SGD 优化器	113
4.2.3	矩阵的高级函数	64	5.6.2 Adadelta 优化器	116
4.3	OpenCV 基本操作	70	5.7 损失函数	117
4.4	图像的基本变换	72	5.7.1 均方误差	117
4.4.1	颜色变换	72	5.7.2 交叉熵损失函数	118
4.4.2	几何变换	80	5.7.3 Keras 提供的损失 函数	120
4.4.3	图像噪声处理	83		
4.5	本章小结	86		

5.8	模型评估方法	122	6.6	基于深度学习的人脸识别	167
5.8.1	交叉验证	122	6.6.1	基于度量学习的方法	168
5.8.2	分类器性能评估	124	6.6.2	基于边界分类的方法	171
5.9	数据增强	127	6.7	本章小结	177
5.9.1	数据增强概述	128	第7章 人脸识别项目实战 178		
5.9.2	Keras 实现数据增强	129	7.1	人脸图片数据集	178
5.9.3	自己实现数据增强	133	7.1.1	Olivetti Faces 人脸数据集	178
5.10	Keras 的工程实践	134	7.1.2	LFW 人脸数据集	180
5.10.1	训练时的回调函数	135	7.1.3	YouTube Faces 人脸数据集	181
5.10.2	打印网络信息	137	7.1.4	IMDB WIKI 人脸数据集	181
5.10.3	输出网络结构图	139	7.1.5	Fddb 人脸数据集	182
5.10.4	获取某层的输出	140	7.2	使用 OpenCV 的人脸检测	182
5.11	本章小结	142	7.2.1	Haar 级联分类器	182
第6章 常用人脸识别算法 143			7.2.2	OpenCV 的 SSD 人脸检测器	184
6.1	特征脸法	143	7.3	使用 Dlib 的人脸检测	186
6.2	OpenCV 的方法	146	7.3.1	基于 Hog-SVM 的人脸检测	186
6.2.1	人脸检测方法	147	7.3.2	基于最大边界的对象检测器	187
6.2.2	人脸识别方法	149	7.4	深度学习实践	188
6.3	Dlib 的人脸检测方法	151	7.4.1	卷积神经网络实现	189
6.4	基于深度学习的图片特征提取	152	7.4.2	数据增强	207
6.4.1	AlexNet	152	7.4.3	自定义损失函数	211
6.4.2	VGGNet	155			
6.4.3	GoogLeNet	157			
6.4.4	ResNet	160			
6.5	基于深度学习的人脸检测	161			
6.5.1	基于深度学习的目标检测	162			
6.5.2	MTCNN	164			

7.4.4	数据预处理	213	8.1.2	云服务的形式	223
7.4.5	模型训练	214	8.1.3	云平台架构设计	224
7.4.6	实现 Web 接口	216	8.2	服务 API 设计	229
7.4.7	模型调优与总结	218	8.2.1	人脸检测	229
7.5	人脸识别的拓展应用	219	8.2.2	人脸对比	239
7.6	本章小结	220	8.3	人脸图片存储	241
第 8 章	人脸识别工程化	221	8.4	人脸图片检索	243
8.1	云平台实践	221	8.5	本章小结	244
8.1.1	云计算介绍	221	附录	参考文献	245

人脸识别入门

在本章中，我们将会接触到一个既熟悉又陌生的概念——人脸识别。之所以熟悉，是因为人脸识别技术在我们日常生活中应用极其广泛，例如火车站刷脸验票进站、手机人脸解锁等；之所以陌生，是因为我们可能并不了解人脸识别的原理，不了解人脸识别的任务目标、发展历程与趋势。

那么，在本章中，我们将会对人脸识别技术的概念、发展、目标等做简要介绍，以便读者对这项技术有一个立体的认识。

1.1 人脸识别概况

人脸识别技术是如今十分热门的一项技术，掌握人脸识别技术的优势不言而喻。下面，我们将首先介绍人脸识别的基本概况。

1.1.1 何为人脸识别

人脸识别技术由来已久，这个概念没有一个严格的定义，一般有狭义与广义之分。

狭义的表达一般是指：以分析与比较人脸视觉特征信息为手段，进行身份验证或查找的一项计算机视觉技术。从表述上看，狭义的人脸识别技术其实是一种身份验证技术，它与我们所熟知的指纹识别、声纹识别、指静脉识别、虹膜识别等均属于同一领域，即生物信息识别领域。因此，狭义上的人脸识别一般指的是通过人脸图像进行身份确认或查找的场景。

生物信息识别的认证方式与传统的身份认证方式相比，具有很多显著优势。例如传统的密钥认证、识别卡认证等存在易丢失、易被伪造、易被遗忘等特点。而生物信息则是人

类与生俱来的一种属性，并不会被丢失和遗忘。而作为生物信息识别之一的人脸识别又具有对采集设备要求不高（最简单的方式只需要能够拍照的设备即可）、采集方式简单等特点。这是虹膜识别、指纹识别等方式所不具备的优点。

人脸识别的广义表述是：在图片或视频流中识别出人脸，并对该人脸图像进行一系列相关操作的技术。例如，在进行人脸身份认证时，不可避免地会经历诸如图像采集、人脸检测、人脸定位、人脸提取、人脸预处理、人脸特征提取、人脸特征对比等步骤，这些都可以认为是人脸识别的范畴。

1.1.2 人脸识别的应用

近些年，随着人脸识别精度的提高，基于该项技术的产品也开始在我们生活中呈现“井喷”之势。例如，早在2016年2月，北京站就开启了“刷脸”进站模式。如图1-1所示是北京站“刷脸”进站的使用提示。现在，越来越多的火车站开始采用“刷脸”进站方式替代人工检票，有效地加快了检票速度。



图 1-1 北京站“刷脸”进站使用提示

人脸识别的另外一个典型应用是手机解锁。随着 iPhone X 的诞生，苹果手机家族增添了一项新的身份验证方式，即所谓的 Face ID。而苹果公司官方宣称，基于 Face ID 的识别准确率要远高于基于指纹识别的 Touch ID。实际上，通过人脸识别来解锁手机并不算什么新鲜事。早在 Android 4.0 时期，这项功能就已经集成在操作系统中了。只不过，由于种种原因，这项功能并未取得比较好的效果，因此无论是谷歌还是手机制造厂商都没有对此进行宣传，自然也不会被大众所了解。值得一提的是，据说苹果公司在这项技术上的研发时间长达 5 年之久，直至 iPhone X 才搭载了完备的人脸识别功能，可见高精度的人脸识别技术并不是一种简单的技术。

上面的两个例子只是人脸识别应用的冰山一角，人脸识别技术的典型应用场景可以总结为如下几个场景。

(1) 身份认证场景

这是人脸识别技术最典型的应用场景之一。门禁系统、手机解锁等都可以归纳为该种类别。该方法与传统的钥匙开锁、指纹识别、虹膜识别等均属于身份认证。这需要系统判断当前被检测人脸是否已经存在于系统内置的人脸数据库中。如果系统内没有该人的信息，则认证失败。

(2) 证件验证场景

证件验证与身份认证相似，也可称为人脸验证，是判断证件中的人脸图像与被识别的人脸是否相同的场景。在进行人脸与证件之间的对比时，往往会引入活体检测技术。

或许大家对活体检测技术并不陌生，就是我们在使用互联网产品时经常会出现的“眨眨眼、摇摇头、点点头、张张嘴”的人脸识别过程，这个过程我们称之为基于动作指令的活体检测。活体检测还可以借由红外线、活体虹膜、排汗等方法来实现。不难理解，引入活体检测可以有效地增加判断的准确性，防止攻击者伪造或窃取他人生物特征用于验证，例如使用照片等平面图片对人脸识别系统进行攻击。

(3) 人脸检索场景

人脸检索与身份验证类似，二者的区别在于身份验证是对人脸图片“一对一”地对比，而人脸检索是对人脸图片“一对多”地对比。例如，在获取到某人的的人脸图片后，可以通过人脸检索方法，在人脸数据库中检索出该人的其他图片，或者查询该人的姓名等相关信息。这与我们在数据库中进行查询是一样的，但人脸检索要比在数据库中查询常规数据复杂得多，例如该以何种方式才能建立高效的人脸图片检索索引呢？人脸检索的应用场景非常多，一个典型的例子是在重要的交通关卡布置人脸检索探头，将行人的的人脸图片在犯罪

嫌疑人数据库中进行检索，从而比较高效地识别出犯罪嫌疑人。

(4) 人脸分类场景

我们这里指的人脸分类主要包括判断人脸图片中的人脸是男人还是女人，所属的年龄区间是怎样的，是什么样的人种，该人的表情是什么等。当然，人脸分类能够实现的功能远不止于此，在很多场景中具有重要的应用价值。例如，社交类 App 可以通过用户上传的自拍图片来判断该用户的性别、年龄等特征，从而为用户有针对性地推荐一些可能感兴趣的人。

(5) 交互式应用场景

美颜类自拍软件大家或许都很熟悉，该类软件除能够实现常规的磨皮、美白、滤镜等功能外，还具有“大眼”“瘦脸”、添加装饰类贴图等功能。而“大眼”“瘦脸”等功能都需要使用人脸识别技术来检测出人眼或面部轮廓，然后根据检测出来的区域对图片进行加工，从而得到我们看到的最终结果。添加装饰类贴图也是在这个基础上实现的，可以认为这是一种 AR(增强现实) 应用。其实，交互式的应用场景远不止于此，还有许多游戏也属于这种交互式的应用场景。

(6) 其他应用

上面所述的内容是人脸识别中应用比较广泛的领域。其实，除这些领域外，人脸识别还有许多其他的应用。例如，人脸图片的重建技术可以应用到通信工程领域，实现低比特率的图片与视频传输；基于人脸识别技术，可以实现人脸图片的合成，甚至直接将一个视频中的人脸完全替换为另外一个人的脸。其中一个经典的项目是 DeepFake，利用该项目可以实现“视频换脸”功能，实现的效果足以以假乱真。

我们可以在日常生活中体会到人脸识别技术为我们的生活带来的便利。随着技术的进一步发展，将会有越来越多的人脸识别相关项目落地。在后面的实战内容中，我们将会围绕这些应用场景，具体介绍其原理与实现方法。

1.1.3 人脸识别的目标

我们已经介绍了人脸识别的不同应用场景。在不同的应用场景下，人脸识别的目标可能是不相同的。但是，对于绝大多数的人脸识别应用场景，人脸识别的目标是类似的。人脸识别的大致流程可以描述为：通过人脸识别模型判断图片中是否存在人脸，如果存在人脸，则定位到该人脸的区位，或者提取该人脸图像的高级特征，作为该人脸图像的特征向量，并用在后续对图片的处理中。

由于人脸识别的应用场景不同，上述步骤的选择和侧重点也不尽相同。例如，定位人脸在图片中的位置，可以用于诸如 AR 等贴图操作；通过定位人脸的关键点，可以对人脸图片进行几何变换，通过几何变换可以实现对图像中人脸的校正，与此同时，得到的人脸关键点还可以用来实现诸如“瘦脸”等操作；如果想要实现的功能并不是对人脸图片的几何变换，而是对图片中的人脸进行特定判断，如判断图片中人脸的性别、年龄等，那么此时的目标是提取出图片中人脸的高级特征，然后根据提取出来的高级特征，使用分类器进行分类，即可以实现诸如性别识别、年龄判断等功能；对于人脸对比，一个可行的思路仍然是提取图片中人脸的高级特征，然后对这两个特征进行对比，从而得出一个相似度数值，通过比较该数值与预设阈值的大小，从而判别两张图片中的人脸是否属于同一个人。

从上面的介绍中我们可以看出，不同人脸识别应用的很多步骤都是重合的，其差异仅在于操作层次的深浅。通过合理选择、组合对人脸图片的操作层次，就可以实现我们预期的目标。这个实现过程可以说是“万变不离其宗”，最核心的技术便是提取人脸图像的高级特征，我们将会在后文的例子中逐步印证这一点。

1.1.4 人脸识别的一般方法

我们首先以人脸对比场景为例，介绍一种人脸对比的可行思路。

我们在前文中提到过，虽然人脸识别的应用很广泛，而且用到的具体技术也不尽相同，但是，有很多步骤其实是类似的。以人脸对比为例，一种可行的解决方案如图 1-2 所示。

下面我们简要介绍一下其中的一些关键步骤。

1. 图像预处理

在很多计算机视觉项目中，往往需要进行图片的预处理操作。这主要是因为输入的图片常存在不合规的地方，甚至会干扰系统的后续工作。如图片带有噪声，或者图片尺寸不符合系统要求等，这些都是预处理这一步需要做的事。而对应的处理方法可以是对

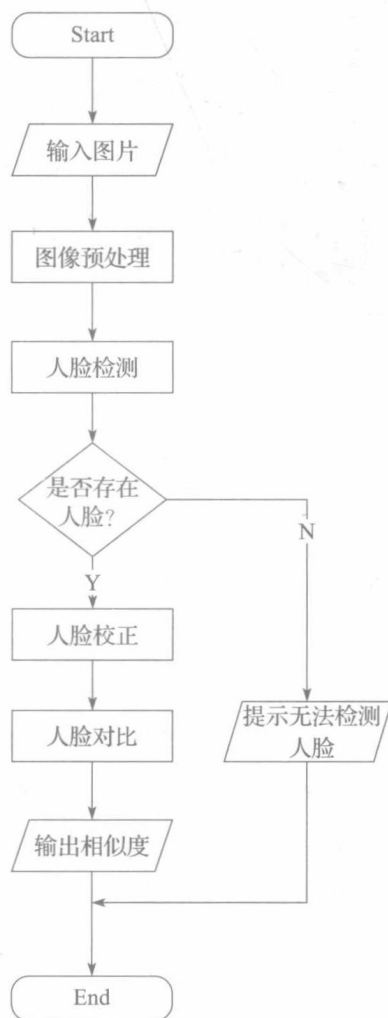


图 1-2 一种人脸对比解决方案的流程图