

深度实践OCR：基于深度学习的文字识别



[深度实践OCR：基于深度学习的文字识别_下载链接1](#)

著者:刘树春

出版者:机械工业出版社

出版时间:2020-5-20

装帧:平装

isbn:9787111654049

读者对象：

1.人工智能领域的技术工程师，尤其是从事图像识别和NLP的技术工程师；

2.专门研究OCR的专业人士或院校师生等。

这是一部融合了企业界先进工程实践经验和学术界前沿技术和思想的OCR著作。

本书由阿里巴巴本地生活研究院算法团队技术专家领衔，从组件、算法、实现、工程应用等维度系统讲解基于深度学习的OCR技术的原理和落地。书中一步步剖析了算法背后的数学原理，提供大量简洁的代码实现，帮助读者从零基础开始构建OCR算法。

全书共10章：

第1章从宏观角度介绍了OCR技术的发展历程、概念和产业应用；

第2章讲解了OCR的图像预处理方法；

第3~4章介绍了传统机器学习方法和深度学习的相关基础知识；

第5章讲解了基于传统方法和深度学习方法的OCR的数据生成；

第6章讲解了与OCR相关的一些高级深度学习方法，方便读者理解后续的检测和识别部分；

第7章讲解了文字的检测技术，从通用的目标检测到文字的检测，一步步加深读者对文字检测问题的认识；

第8章讨论了文字识别的相关技术，定位到文字的位置之后，需要对文字的内容进行进一步的解析；

第9章介绍了一些OCR后处理的方法；

第10章介绍了一些版面分析方法。

作者介绍:

刘树春

阿里巴巴本地生活研究院算法专家，前复旦七牛云联合实验室OCR算法负责人，在OCR相关技术的落地和实现方面有深入的研究和丰富的实践经验，在菜单识别、车牌检测识别、卡证识别、商业广告文字检测和识别、票据类识别等应用场景中积累了丰富的经验。曾经组队参加COCO竞赛、ICDAR刷榜等活动，并发表多篇顶级论文。

贺盼

佛罗里达大学在读博士，担任CVPR、ICCV、ECCV等10多家国际顶会审稿人或程序委员，在场景文本领域发表过多篇国际顶尖会议论文(DTRN、CTPN、SSTD等)。曾经在中科院先进技术研究院、港中文、旷世美国研究院等做过相关研究工作，对场景文本有深刻的认识，并且对场景文本学术前沿有很强的感知力。

马建奇

著名的RRPN算法的作者，发表过多篇国际顶尖会议论文，曾经在旷世研究院做过检测相关的工作，在上海高等研究院做过文字检测相关的工作，曾经组队参加过ICDAR 17

MLT比赛，获得前三名。曾获Facebook研究实习机会。目前即将开始修读博士学位，正在做文字识别相关的工作，对于OCR的问题有独到的见解。

王佳军

阿里巴巴本地生活研究院高级算法专家，负责本地生活研究院图像算法团队的管理、研发等工作，曾经在滴滴研究院做相关算法的研究。业余时间担任深度学习框架MXNet的技术专家，对实际算法的落地优化等有非常深入的耕耘，具有很强的OCR相关项目的落地能力。

谢雨飞

趣头条算法工程师，曾在七牛云做OCR、NLP相关算法的研究工作，发表过多篇国际顶尖论文。

陈明曦

纽约大学硕士研究生，曾在七牛云从事OCR相关工作，发表相关论文数篇。

目录: 推荐序

前言

第1章 绪论 1

1.1 人工智能大潮中的OCR发展史 1

1.1.1 传统OCR方法一般流程 3

1.1.2 基于深度学习OCR方法一般流程 5

1.2 文字检测 6

1.3 文字识别 8

1.4 产业应用现状 10

1.5 本章小结 11

1.6 参考文献 11

第2章 图像预处理 13

2.1 二值化 13

2.1.1 全局阈值方法 13

2.1.2 局部阈值方法 17

2.1.3 基于深度学习的方法 20

2.1.4 其他方法 22

2.2 平滑去噪 26

2.2.1 空间滤波 26

2.2.2 小波阈值去噪 28

2.2.3 非局部方法 29

2.2.4 基于神经网络的方法 33

2.3 倾斜角检测和校正 35

2.3.1 霍夫变换 36

2.3.2 Radon 变换 38

2.3.3 基于PCA的方法 38

2.4 实战 39

2.5 参考文献 43

第3章 传统机器学习方法绪论 45

3.1 特征提取方法 45

3.1.1 基于结构形态的特征提取 45

3.1.2 基于几何分布的特征提取 61

3.2 分类方法模型 63

3.2.1 支持向量机 63

3.2.2 K近邻算法	65
3.2.3 多层感知器	70
3.3 实战：身份证号码的识别	71
3.3.1 核心代码	71
3.3.2 测试结果	78
3.4 本章小结	79
3.5 参考文献	79
第4章 深度学习基础知识	80
4.1 单层神经网络	80
4.1.1 神经元	80
4.1.2 感知机	81
4.2 双层神经网络	82
4.2.1 双层神经网络简介	82
4.2.2 常用的激活函数	83
4.2.3 反向传播算法	86
4.3 深度学习	87
4.3.1 卷积神经网络	88
4.3.2 常用优化算法	90
4.4 训练网络技巧	92
4.4.1 权值初始化	92
4.4.2 L1/L2正则化	93
4.4.3 Dropout	94
4.5 实战	95
4.6 参考文献	97
第5章 数据生成	99
5.1 背景介绍	99
5.2 传统单字OCR数据生成	100
5.3 基于深度学习的OCR数据生成	101
5.3.1 文字检测数据的生成	101
5.3.2 检测图片生成	103
5.3.3 其他方法	112
5.3.4 识别数据生成	113
5.4 通过GAN的技术生成数据	114
5.4.1 GAN背景介绍	114
5.4.2 GAN的原理	116
5.4.3 GAN的变种	117
5.5 图像增广	123
5.5.1 常用的图像增强方法	123
5.5.2 深度学习方法	126
5.6 常用的开源数据集	128
5.7 ICDAR的任务和数据集	131
5.8 本章小结	138
5.9 参考文献	138
第6章 深度学习高级方法	140
6.1 图像分类模型	140
6.1.1 LeNet5	140
6.1.2 AlexNet	142
6.1.3 VGGNet	143
6.1.4 GoogLeNet	144
6.1.5 ResNet	147
6.1.6 DenseNet	151
6.1.7 SENet	153
6.1.8 轻量化网络	154
6.2 循环神经网络	154

6.2.1 RNN网络	154
6.2.2 GRU	156
6.2.3 GRU的实现	158
6.2.4 LSTM网络	160
6.3 Seq2Seq	163
6.4 CTC Loss	164
6.4.1 算法详解	166
6.4.2 前向传播	168
6.4.3 后向传播	173
6.4.4 前向/后向算法	173
6.4.5 CTC算法特性	174
6.4.6 代码解析	175
6.5 Attention	178
6.6 本章小结	181
6.7 参考文献	182
第7章 文字检测	183
7.1 研究意义	183
7.2 目标检测方法	185
7.2.1 目标检测相关术语	186
7.2.2 传统检测方法	189
7.2.3 Two-stage 方法	195
7.2.4 One-stage 方法	210
7.3 文本检测方法	217
7.3.1 传统文本检测方法	217
7.3.2 基于深度学习的文本检测方法	222
7.4 本章小结	228
7.5 参考文献	228
第8章 字符识别	232
8.1 任务概览	232
8.2 数据集说明	233
8.2.1 数据集意义	233
8.2.2 常见识别数据集介绍	234
8.3 评测指标	238
8.3.1 编辑距离	239
8.3.2 归一化编辑距离	239
8.3.3 字符准确度	239
8.3.4 词准确率	239
8.3.5 语境相关的评测方式	239
8.4 主流算法介绍	240
8.4.1 传统光学方法	240
8.4.2 完全基于深度学习的方法	244
8.5 CRNN模型实战	274
8.5.1 简介	274
8.5.2 运行环境	274
8.5.3 测试部分讲解	274
8.5.4 测试运行结果	279
8.5.5 训练部分	279
8.5.6 用ICDAR2013数据集训练CRNN模型	282
8.6 本章小结	284
8.7 参考文献	285
第9章 OCR后处理方法	288
9.1 文本纠错	288
9.1.1 BK-tree	289
9.1.2 基于语言模型的中文纠错	293

- 9.2 文本结构化 297
 - 9.2.1 模板匹配 298
 - 9.2.2 文本分类 300
- 9.3 本章小结 304
- 9.4 参考文献 304
- 第10章 版面分析 306
 - 10.1 版面分析详解 306
 - 10.2 复杂版面识别 309
 - 10.3 文档恢复 310
 - 10.4 本章小结 311
 - 10.5 参考文献 311
- • • • • [\(收起\)](#)

[深度实践OCR：基于深度学习的文字识别_下载链接1_](#)

标签

OCR

阿里巴巴

深度学习

想读的书

好书，值得一读

评论

还是非常不错的一本书，内容还是相对全面的。

[深度实践OCR：基于深度学习的文字识别_下载链接1_](#)

[深度实践OCR：基于深度学习的文字识别_下载链接1](#)