

Broadview[®]
www.broadview.com.cn

从概念到应用剖析基于深度学习的目标检测
涵盖医疗、交通、无人驾驶领域的工程实践

深度学习与 目标检测

杜鹏 谌明 苏统华◎编著

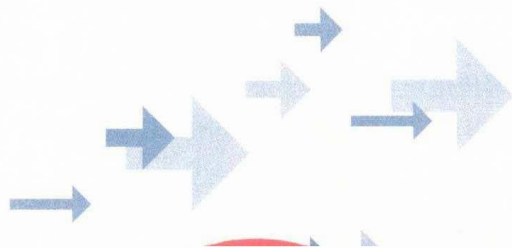
DEEP LEARNING AND OBJECT DETECTION



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>



深度学习与 目标检测

杜鹏 谌明 苏统华 编著



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书的写作初衷是，从学者的角度，用一种通俗易懂的方式，将基于深度学习的目标检测的相关论文中的理论和方法呈现给读者，同时针对作者在深度学习教学过程中遇到的难点，进行深入分析和讲解。

本书侧重对卷积神经网络的介绍，而深度学习的内容不止于此。所以，作者将深度学习分为有监督学习、无监督学习和强化学习三类，将图像分类、目标检测、人脸识别、语音识别、生成对抗网络和 AlphaGo 等应用场景归入不同的类别，并分别对其原理进行了概括性的讲解。

本书适合有一定深度学习或目标检测学习基础的学生、研究者、从业者阅读。

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

深度学习与目标检测 / 杜鹃, 谌明, 苏统华编著. —北京: 电子工业出版社, 2020.3
ISBN 978-7-121-36785-4

I. ①深… II. ①杜… ②谌… ③苏… III. ①机器学习 IV. ①TP181

中国版本图书馆 CIP 数据核字 (2019) 第 108348 号

责任编辑：潘 昕

印 刷：北京盛通印刷股份有限公司

装 订：北京盛通印刷股份有限公司

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：17 字数：307 千字

版 次：2020 年 3 月第 1 版

印 次：2020 年 3 月第 1 次印刷

定 价：89.00 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 zltts@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。

作者简介

杜鹏，博士，现任职于华为。2014年起在韩国科学技术学院和新加坡南洋理工大学从事博士后研究。回国后，曾任杭州电子科技大学副教授、浙江核新同花顺网络信息股份有限公司算法研究员。2018年被NVIDIA深度学习学院评为全球最佳校园大使，在SIGGRAPH、PG等国际著名学术会议发表论文10余篇。

谌明，博士，2004年加入美国道富集团，2011年加入浙江核新同花顺网络信息股份有限公司并任首席技术官，推动了包括计算机视觉、语音技术、自然语言处理、机器学习等在金融、医疗等领域的商业化落地。

苏统华，博士，哈尔滨工业大学副教授，自然手写中文文本识别的开拓者。曾出版手写汉字识别领域的首本英文专著，以及7本GPU计算和大数据分析方面的译作，所领导的NVIDIA GPU教育中心连续4年被NVIDIA评为中国区最佳GPU教育中心。

欢迎投稿

panxin@phei.com.cn

序

深度学习自 2006 年被正式提出后，经过 10 余年的发展，已经在很多领域取得了突破性的进展。2015 年，深度学习在著名的图像分类数据集 ImageNet 上成功超越了人类的分类准确率。2016 年，深度学习应用于强化学习领域，在围棋项目上击败了最优秀的人类棋手。2019 年，在《星际争霸 II》这样复杂的游戏项目中，深度学习同样击败了人类职业选手。

尽管深度学习让人工智能的实现变得似乎不再那么遥不可及，但是在目标检测领域，深度学习还没能让计算机超越人类。因此，系统性地整理这个领域的研究成果并成书，让更多的人参与进来，对推动深度学习在目标检测领域的研究具有积极的意义。

本书从深度学习的发展历史开始，为读者介绍了机器学习的各个流派在深度学习出现后的代表性研究成果，然后，介绍了成熟的深度学习方法和技术，以及两阶段和单阶段的基于深度学习的目标检测方法。本书内容由浅入深，适合不同层次的读者阅读，我本人读后也收获颇丰，感觉大有裨益。

本书不仅梳理了在目标检测领域应用深度学习之后出现的主流方法，还讲解了具体的训练过程，为读者学习深度学习的参数调节打下了基础，使读者能更加全面地了解相关算法，在真正进入这个领域时更容易上手。对于已经熟悉这个领域的读者，通过反复研读本书的内容，可以使自己对调参的理解更加透彻。

除了以上内容，本书还介绍了目标检测算法在一些实际问题中的应用案例。应用案例覆盖面广，并涉及当下热门的人工智能创业场景——智慧医疗和智慧交通。对每个案例，都介绍了国内外的研究现状，给出了解决方案，并详细描述了预处理、算法实现细节和实验结果。这些内容，不仅是深度学习在不同目标检测领域得以成功应用的关键，而且能帮助读者提高解决实际问题的能力。

我期待这本书能帮助更多的读者了解深度学习和目标检测领域的前沿知识，更期待这本书的读者将来能让计算机拥有超越人类的目标检测能力。

何晓飞^①

① 何晓飞，机器学习领域的世界顶级科学家，浙江大学教授、博士生导师，国家杰出青年基金获得者，首届中组部青年拔尖人才支持计划入选者，国际模式识别学会会士（IAPR Fellow），曾任滴滴研究院院长。

前 言

时光飞逝，从 2018 年 3 月开始组织素材到本书出版，一年多的时间过去了。

创作本书的初衷是希望从学者的角度，用一种通俗易懂的方式，将基于深度学习的目标检测的相关论文中的理论和方法呈现给读者。因此，本书选取了 R-CNN、SPP-Net、Fast R-CNN、Faster R-CNN、R-FCN、SSD、RetinaNet、RefineDet、YOLO 等典型的目标检测网络及 U-Net、SegNet、Mask R-CNN 等实例分割网络进行介绍，同时针对笔者在深度学习教学过程中遇到的难点进行了深入的分析 and 讲解。此外，结合三位作者的研究成果，给出了目标检测在医疗、交通等领域的深度学习应用案例。

本书内容侧重于卷积神经网络，而深度学习的内容不止于此。所以，笔者将深度学习分为有监督学习、无监督学习和强化学习三类，将图像分类、目标检测、人脸识别、语音识别、生成对抗网络、AlphaGo 等应用场景归入不同的类别，并分别对其原理进行了概括性的讲解。

本书分为三篇，第 1 章~第 3 章为基础篇，第 4 章和第 5 章为进阶篇，第 6 章~第 9 章为应用篇。

- 第 1 章介绍了深度学习的典型应用。
- 第 2 章介绍了深度神经网络的基础知识。
- 第 3 章介绍了卷积神经网络的基础知识，以及典型的卷积神经网络。
- 第 4 章介绍了两阶段目标检测方法和相应的分割算法。
- 第 5 章介绍了单阶段目标检测方法。
- 第 6 章给出了肋骨骨折检测的应用案例。
- 第 7 章给出了肺结节检测的应用案例。
- 第 8 章给出了车道线检测的应用案例。
- 第 9 章给出了交通视频分析的应用案例。

书中提到的参考链接列表，请访问 <http://www.broadview.com.cn/36785> 下载。

杜鹏、谌明负责前七章内容的编写，苏统华负责后两章内容的编写。第 6 章~第 9 章的实验部分由汪纯、许卫东、金弘晟、李松泽、孙黎完成。罗同桢、张栋、胡明玥、陈希坚等同学参与了本书的编写。金耀博士参与了本书的审校。对此，我们表示衷心的感谢。同时，感谢浙江省公益基金（No.LGF19F020014）和国家自然科学基金（No.61673140 和 No.81671771）对本书的资助。

近一年来，以 FCOS、CornerNet-Lite、CenterNet 为代表的无锚点碰撞检测成为目标检测领域新的研究热点。这类方法与基于锚点的目标检测算法相比，缩减了超参数的数量并提升了精度和性能。受时间所限，本书无法对这部分内容进行详细介绍，请感兴趣的读者参考相关文献。

本书的不足之处，恳请各位读者批评、指正。

杜鹏、谌明、苏统华

2019 年 11 月

目 录

基础篇

第1章 深度学习概述	2
1.1 深度学习发展简史	2
1.2 有监督学习	4
1.2.1 图像分类	4
1.2.2 目标检测	7
1.2.3 人脸识别	10
1.2.4 语音识别	13
1.3 无监督学习	18
1.3.1 无监督学习概述	18
1.3.2 生成对抗网络	18
1.4 强化学习	21
1.4.1 AlphaGo	21
1.4.2 AlphaGo Zero	23
1.5 小结	25
参考资料	25
第2章 深度神经网络	27
2.1 神经元	27
2.2 感知机	30
2.3 前向传递	31
2.3.1 前向传递的流程	32
2.3.2 激活函数	33
2.3.3 损失函数	37

2.4	后向传递	40
2.4.1	后向传递的流程	40
2.4.2	梯度下降	40
2.4.3	参数修正	42
2.5	防止过拟合	44
2.5.1	dropout	44
2.5.2	正则化	45
2.6	小结	46
第3章 卷积神经网络		47
3.1	卷积层	48
3.1.1	valid 卷积	48
3.1.2	full 卷积	50
3.1.3	same 卷积	51
3.2	池化层	52
3.3	反卷积	53
3.4	感受野	55
3.5	卷积神经网络实例	56
3.5.1	LeNet-5	57
3.5.2	AlexNet	59
3.5.3	VGGNet	63
3.5.4	GoogLeNet	65
3.5.5	ResNet	75
3.5.6	MobileNet	76
3.6	小结	78

进阶篇

第4章 两阶段目标检测方法		80
4.1	R-CNN	80
4.1.1	算法流程	80
4.1.2	训练过程	81

4.2	SPP-Net.....	85
4.2.1	网络结构.....	85
4.2.2	空间金字塔池化.....	86
4.3	Fast R-CNN.....	87
4.3.1	感兴趣区域池化层.....	87
4.3.2	网络结构.....	89
4.3.3	全连接层计算加速.....	90
4.3.4	目标分类.....	91
4.3.5	边界框回归.....	92
4.3.6	训练过程.....	93
4.4	Faster R-CNN.....	97
4.4.1	网络结构.....	98
4.4.2	RPN.....	99
4.4.3	训练过程.....	105
4.5	R-FCN.....	107
4.5.1	R-FCN 网络结构.....	108
4.5.2	位置敏感的分权图.....	109
4.5.3	位置敏感的 RoI 池化.....	110
4.5.4	R-FCN 损失函数.....	111
4.5.5	Caffe 网络模型解析.....	111
4.5.6	U-Net.....	115
4.5.7	SegNe.....	116
4.6	Mask R-CNN.....	117
4.6.1	实例分割简介.....	118
4.6.2	COCO 数据集的像素级标注.....	119
4.6.3	网络结构.....	120
4.7	小结.....	123
	参考资料.....	123
第 5 章 单阶段目标检测方法.....		125
5.1	SSD.....	125
5.1.1	default box.....	125

5.1.2	网络结构	126
5.1.3	Caffe 网络模型解析	127
5.1.4	训练过程	135
5.2	RetinaNet	137
5.2.1	FPN	137
5.2.2	聚焦损失函数	139
5.3	RefineDet	140
5.3.1	网络模型	141
5.3.2	Caffe 网络模型解析	143
5.3.3	训练过程	152
5.4	YOLO	153
5.4.1	YOLO v1	153
5.4.2	YOLO v2	155
5.4.3	YOLO v3	157
5.5	目标检测算法的应用场景	159
5.5.1	高速公路坑洞检测	160
5.5.2	息肉检测	161
5.6	小结	162
	参考资料	162

应用篇

第 6 章	肋骨骨折检测	166
6.1	国内外研究现状	166
6.2	解决方案	168
6.3	预处理	168
6.4	肋骨骨折检测	169
6.5	实验结果分析	170
6.6	小结	172
	参考资料	173

第 7 章 肺结节检测	174
7.1 国内外研究现状	174
7.2 总体框架	176
7.2.1 肺结节数据集	176
7.2.2 肺结节检测难点	177
7.2.3 算法框架	177
7.3 肺结节可疑位置推荐算法	178
7.3.1 CT 图像的预处理	179
7.3.2 肺结节分割算法	180
7.3.3 优化方法	182
7.3.4 推断方法	184
7.4 可疑肺结节定位算法	185
7.5 实验结果与分析 (1)	186
7.5.1 实验结果	186
7.5.2 改进点效果分析	186
7.6 假阳性肺结节抑制算法	188
7.6.1 假阳性肺结节抑制网络	188
7.6.2 优化策略	192
7.6.3 推断策略	194
7.7 实验结果与分析 (2)	194
7.7.1 实验结果	195
7.7.2 改进点效果分析	195
7.7.3 可疑位置推荐算法与假阳性抑制算法的整合	197
7.8 小结	197
参考资料	197
第 8 章 车道线检测	200
8.1 国内外研究现状	200
8.2 主要研究内容	202
8.2.1 总体解决方案	202
8.2.2 各阶段概述	203

8.3	车道线检测系统的设计与实现	206
8.3.1	车道线图像数据标注与筛选	206
8.3.2	车道线图片预处理	208
8.3.3	车道线分割模型训练	212
8.3.4	车道线检测	221
8.3.5	车道线检测结果	225
8.4	车道线检测系统性能测试	225
8.4.1	车道线检测质量测试	225
8.4.2	车道线检测时间测试	227
8.5	小结	227
	参考资料	228
第9章	交通视频分析	229
9.1	国内外研究现状	230
9.2	主要研究内容	231
9.2.1	总体设计	232
9.2.2	精度和性能要求	232
9.3	交通视频分析	233
9.3.1	车辆检测和车牌检测	233
9.3.2	车牌识别功能设计详解	235
9.3.3	车辆品牌及颜色的识别	243
9.3.4	目标跟踪设计详解	244
9.4	系统测试	247
9.4.1	车辆检测	248
9.4.2	车牌检测	251
9.4.3	车牌识别	253
9.4.4	车辆品牌识别	256
9.4.5	目标跟踪	258
9.5	小结	259
	参考资料	260

基础篇

本篇首先介绍了深度学习的基本概念、历史发展、分类和主要应用领域，然后介绍了深度神经网络和卷积神经网络的概念，最后介绍了一些具有代表性的卷积神经网络。

本篇的主要目的是帮助读者了解深度学习的基本概念和原理，以及与基于深度学习的目标检测算法有关的基础知识。

第1章 深度学习概述

建立能够模拟人类大脑进行分析和学习的神经网络，从而达到模仿大脑处理图像、声音、文本的目的——这就是深度学习的起源。

深度学习的核心是深度神经网络。深度神经网络是一种模仿神经网络进行信息分布式并行处理的数学模型。神经网络是机器学习的一个重要分支，而深度学习就是深度神经网络在近年来的重要突破。深度学习只需要使用简单的网络结构就能够实现对复杂函数的逼近，同时，由于网络层次较深，多个隐藏层能够非常好地表达数据的特征。

深度学习已经被广泛应用于语音识别、图像分类、自然语言处理等领域，尤其是谷歌旗下的深度思考（DeepMind）公司开发的人工智能围棋软件——阿尔法围棋（AlphaGo），先后击败李世石、柯洁等世界围棋名将，引起了人们对“人工智能是否能够取代人类”的大讨论。人工智能凭借深度学习的兴起，已然成为计算机领域炙手可热的研究方向。

1.1 深度学习发展简史

1. 1958年—1986年

1958年，Rosenblatt发明了用于对输入的多维数据进行二分类且能够使用梯度下降方法从训练样本中自动更新权值的感知机。但是，1969年，美国数学家及人工智能先驱马文·明斯基（Marvin Lee Minsky）在其著作中证明：感知机只能处理线性分类问题，其本质就是一种线性模型，甚至不能成功解决异或分类问题。

1986年，“深度学习之父”杰弗里·辛顿（Geoffrey Hinton）第一次打破了神经网络的非线性“诅咒”。Hinton发明了适用于多层感知机（multi-layer perceptron, MLP）的反向传播（back propagation, BP）算法，并使用sigmoid函数进行非线性映射，从而有效地解决了非线性分类和学习问题。然而，由于神经