

人工智能 前沿技术丛书

Deep Learning for

Intelligence Interpretation and Recognition of
Remote Sensing Images

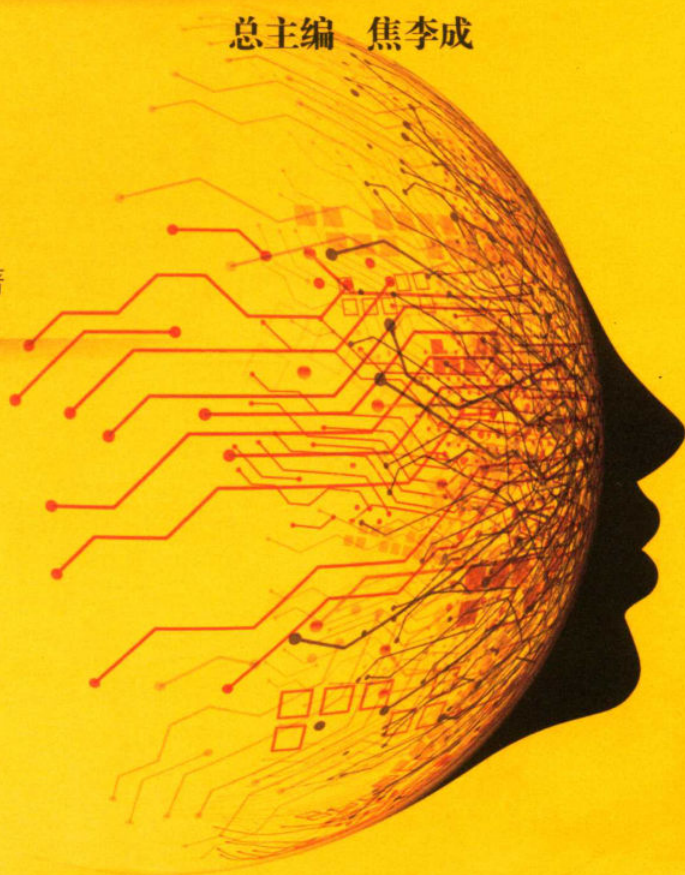
遥感影像深度学习 智能解译与识别

总主编 焦李成

焦李成 刘 芳 李玲玲 杨淑媛 著
侯 彪 杨争艳 杨 慧 孟繁荣



西安电子科技大学出版社
<http://www.xdph.com>



人工智能 前沿技术丛书

总主编 焦李成

遥感影像深度学习 智能解译与识别

焦李成 刘 芳 李玲玲 杨淑媛 著
侯 彪 杨争艳 杨 慧 孟繁荣



西安电子科技大学出版社
<http://www.xduph.com>

内 容 简 介

本书从人工智能前沿理论与技术出发,系统地论述了遥感影像深度学习智能解译与识别的基本理论、算法及应用。全书共分为四个部分,分别是 SAR 图像分类与变化检测、极化 SAR 图像分类与变化检测、高光谱影像分类、遥感影像解译描述与分类,并给出了遥感影像深度学习智能解译与识别的最新进展。每章都附有相关阅读材料,便于有兴趣的读者进一步研究。

本书为人工智能教育丛书,可为高等院校人工智能、计算机科学、电子科学与技术、信息科学、控制科学与工程等领域的研究人员提供参考,也可作为相关专业本科生及研究生的参考书。

图书在版编目(CIP)数据

遥感影像深度学习智能解译与识别/焦李成等著. —西安:西安电子科技大学出版社,2019.9
(2020.7重印)

ISBN 978-7-5606-5350-1

I. ①遥… II. ①焦… III. ①遥感图—机器学习—图像分析—研究 IV. ①TP751

中国版本图书馆 CIP 数据核字(2019)第 092600 号

策划编辑 人工智能教育丛书项目组

责任编辑 刘小莉 秦志峰

出版发行 西安电子科技大学出版社(西安市太白南路2号)

电 话 (029)88242885 88201467 邮 编 710071

网 址 www.xduph.com 电子邮箱 xdupfxb001@163.com

经 销 新华书店

印刷单位 广东虎彩云印刷有限公司

版 次 2019年9月第1版 2020年7月第2次印刷

开 本 787毫米×960毫米 1/16 印张 28.5

字 数 574千字

定 价 110.00元

ISBN 978-7-5606-5350-1/TP

XDUP 5652001-2

*** 如有印装问题可调换 ***

序言

PREFACE

随着高分辨率卫星的快速发展和国产卫星数量的不断增多,包含丰富地物信息的遥感影像数据规模日益猛增,对浩如烟海的遥感影像数据的解译成为了当前的热点、难点问题。传统的遥感影像处理手段,依赖于较强的专业知识和数据特征本身,需要消耗大量的人力和时间,但人工智能技术的快速发展为遥感图像处理技术带来了变革。本书将人工智能深度学习技术引入遥感数据的解译与识别中,有效地提升了遥感数据的自动化处理和分析能力,并成功应用于包括图像分类、图像分割、目标检测、变化检测、超高分辨率重建等多个场景中,为国土资源管理、构建智慧城市、深度军民融合等应用开拓了广阔的前景,助力空间信息行业的智能化发展。

神经网络几经沉浮,走过了艰难曲折的历程,在2006年,单隐层神经网络模型拓展到了深度神经网络模型,“深度学习”这一术语开始普及,神经网络迎来了它的第三波浪潮并开启了“深度智能”时代。近年来,依靠强大的计算设备、海量数据集以及不断完善的深度神经网络理论知识,深度学习的普及性和实用性均有了极大的发展,成为了机器学习乃至人工智能领域最热门的技术,并持续展现着强大的生命力,它将不断涌现出新的理论发展和方法实践,深刻影响人工智能、社会经济及人类生活的未来。我们团队将深度学习与稀疏认知学习、多尺度几何分析等思想相结合,并应用于遥感影像智能解译与识别,研究成果表明,此方法能够克服传统遥感影像处理中出现的复杂场景下图像处理效率和精度低、网络过拟合、鲁棒性和泛化性能差等问题。

本书基于团队工作成果,从SAR图像分类与变化检测、极化SAR图像分类与变化检测、高光谱影像分类、遥感影像解译描述与分类这四个部分论述了深度学习在遥感影像智能解译与识别中的应用。书中涉及的模型包括:DC-ResNet、脊波反卷积结构学习、改进帧间差分法与YOLO深度网络、多尺度跳跃型卷积网络、SPP Net、自步学习和对称卷积耦合网络、多层特征SENet、Task-Oriented GAN网络、阶梯网络、深度堆栈网络、复数轮廓波卷积神经网络、加权卷积神经网络与主动学习、多尺度深度Directionlet网络、局部受限卷积神经网络、Looking-Around-and-Into网络、胶囊网络、空谱解耦合双通道卷积神经网络、快速区域卷积神经网络、局部响应卷积递归神经网络等。相比于其他遥感影像解译书籍,本书从人工智能理论与技术的前沿出发,期望能为读者带来前瞻性的视角,章节安排由浅入深,在模型概述和遥感影像解译发展的基础上逐步展开,内容涵盖更为广泛,模型讨论更为深入,应用实践更为细致,希望为读者入门学习及深入钻研提供帮助。

本书的完成离不开团队多位老师和研究生的支持与帮助,感谢团队中侯彪、刘静、王

爽、杨淑媛、张向荣、侯水平、尚荣华、刘波、田小林等教授以及马晶晶、马文萍、白静、张小华、曹向海、冯捷、唐旭等副教授对本工作的关心支持与辛勤付出。同时感谢刘芳、赵进、刘旭、赵暉、朱浩、孙其功、任仲乐、宋纬、张文华等博士生，以及马丽媛、侯瑶淇、曾杰、王美玲、汶茂宁、张婷、李晰、孙莹莹、张佳琪、王继蕾、王亚明、叶维健、段丽英、张大臣、梁莹、张文豪、张娉婷、李翔等研究生的工作和劳动。最后，特别感谢杨争艳、杨慧、孟繁荣等同学的付出和辛勤劳动。本书是我们团队在该领域工作的一个小结，也汇聚了西安电子科技大学智能感知与图像理解教育部重点实验室、智能感知与计算国际联合实验室及智能感知与计算国际联合研究中心的集体智慧。在本书出版之际，特别感谢邱关源先生及保铮院士三十多年来的悉心培养与教导，特别感谢徐宗本院士、张钹院士、李衍达院士、郭爱克院士、郑南宁院士、谭铁牛院士、马远良院士、包为民院士、郝跃院士、陈国良院士、管晓宏院士及韩崇昭教授、张青富教授、张军教授、姚新教授、刘德荣教授、金耀初教授、周志华教授、李学龙教授、吴枫教授、田捷教授、贾秀萍教授、屈嵘教授、李军教授、张艳宁教授、马西奎教授、潘泉教授、高新波教授、石光明教授、李小平教授、陈莉教授、王磊教授等多年来的关怀、帮助与指导，感谢教育部创新团队和国家“111”创新引智基地的支持；同时，我们的工作也得到西安电子科技大学领导及国家“973”计划(2013CB329402)、国家自然科学基金(61836009, 61871310, U1701267, 61621005, 61573267, 61472306, 61573267, 61473215, 61571342, 61501353, 61502369)、重大专项计划(91438201, 91438103)等科研任务的支持，特此感谢。感谢书中所有被引用文献的作者。20世纪90年代初，我们出版了《神经网络系统理论》《神经网络计算》《神经网络的应用与实现》等系列专著，三十年来神经网络取得了长足的进展，我们亦出版了《深度学习、优化与识别》专著，希望能为领域的发展和普及持续做出绵薄贡献。

限于作者水平，本书难免在内容取材和结构编排上有不妥之处，希望读者不吝赐教，提出宝贵的批评和建议，我们将不胜感激。

作 者

2018年12月

西安电子科技大学

目录 CONTENTS

第一部分 SAR 图像分类与变化检测

第 1 章 基于 DC-ResNet 的 SAR 图像	
目标分类	3
1.1 引言	3
1.2 深度残差网络基础	3
1.2.1 非线性激活函数	4
1.2.2 Dropout 操作	4
1.2.3 批规范化	5
1.2.4 全局均值池化	6
1.2.5 跳跃连接	6
1.3 基于 DC-ResNet 的 SAR 图像	
目标分类	7
1.3.1 可变形卷积核	7
1.3.2 可变形卷积残差模块	8
1.3.3 DC-ResNet 模型	9
1.4 实验结果与分析	10
1.4.1 实验数据	10
1.4.2 实验环境	13
1.4.3 实验结果分析	13
本章小结	24
参考文献	24
第 2 章 脊波反卷积结构学习模型	27
2.1 引言	27
2.2 反卷积结构模型	28
2.2.1 经典的反卷积神经网络	28
2.2.2 构造反卷积结构模型的前期实验	28
2.2.3 反卷积结构模型的构造	29
2.2.4 反卷积结构模型的训练过程	32
2.3 脊波反卷积结构学习模型	35
2.3.1 模型的构造	35
2.3.2 模型的训练	37
2.4 混合聚集结构地物像素子空间的	
SAR 图像分割	39
2.4.1 算法描述	39
2.4.2 实验仿真与分析	40
本章小结	47
参考文献	48
第 3 章 基于改进帧差法与 YOLO 深度网络的	
遥感影像目标检测	51
3.1 引言	51
3.2 帧间差分法	52
3.2.1 帧间差分法原理	52
3.2.2 改进帧间差分法	52
3.3 YOLO 深度网络	53
3.3.1 YOLO 网络的思想与原理	53
3.3.2 YOLO 网络损失函数的设计	54
3.3.3 YOLO 网络的优势	55
3.4 基于改进帧差法和 YOLO 深度网络	
的遥感影像目标检测	56
3.4.1 实验数据准备及预处理	56
3.4.2 基于改进帧差法的遥感影像运动	
目标检测	58
3.4.3 基于高分辨 YOLO 深度网络的	
遥感影像目标检测	59
3.5 实验结果与分析	62
3.5.1 参数设置	62

3.5.2 结果与分析	62	5.5.3 参数化分析	102
本章小结	65	5.5.4 基于 SPP Net 的 SAR 图像变化 检测仿真实验	104
参考文献	66	本章小结	116
第 4 章 基于多尺度跳跃型卷积网络的 SAR 图像变化检测	69	参考文献	117
4.1 引言	69	第 6 章 基于自步学习和对称卷积耦合网络 的 SAR 图像变化检测	120
4.2 密集型卷积网络	70	6.1 卷积神经网络	120
4.3 多尺度跳跃型卷积网络	72	6.2 SPLPSO 算法	121
4.3.1 多尺度跳跃型卷积网络	72	6.2.1 粒子群优化算法	121
4.3.2 基于多尺度跳跃型卷积网络的 SAR 图像变化检测	73	6.2.2 年龄参数对自步学习的影响分析	122
4.4 实验结果与分析	74	6.2.3 SPLPSO 模型构造	123
4.4.1 邻域尺寸的选择	74	6.3 SCCN 网络	127
4.4.2 多尺度跳跃型卷积网络的性能	78	6.3.1 网络结构	127
本章小结	82	6.3.2 网络的学习优化	128
参考文献	82	6.4 基于自适应自步学习和 SCCN 的 SAR 图像变化检测	129
第 5 章 基于 SPP Net 的 SAR 图像变化检测 ..	86	6.4.1 算法的基本框架	129
5.1 引言	86	6.4.2 SPL-SCCN 理论分析	130
5.2 无监督方法的 SAR 图像变化检测	86	6.4.3 自步学习训练模型	131
5.2.1 深度置信网络	87	6.5 实验结果与分析	131
5.2.2 模糊 C 均值聚类算法	90	本章小结	137
5.2.3 基于 DBN 和 FCM 的 SAR 图像 变化检测	91	参考文献	137
5.3 空间金字塔池化网络	92	第 7 章 基于多层特征 SENet 的 SAR 图像 目标分类方法	140
5.4 基于 SPP Net 的 SAR 图像变化检测	94	7.1 引言	140
5.4.1 SPP Net 感兴趣区域检测网络的 构造	94	7.2 数据增强	140
5.4.2 SPP Net 感兴趣区域检测样本的 选择	96	7.3 SENet 网络	141
5.4.3 训练过程	97	7.3.1 Inception 模块	141
5.5 实验结果与参数分析	98	7.3.2 残差模块	143
5.5.1 数据集	98	7.3.3 Squeeze-and-excitation 网络	143
5.5.2 基于 DBN 和 FCM 的 SAR 图像 变化检测仿真实验	101	7.3.4 全局平均池化	144
		7.4 基于多层特征 SENet 的 SAR 图像 目标分类	145

7.4.1 训练数据集的获取及扩充	145	本章小结	166
7.4.2 分类网络模型	145	参考文献	166
7.5 实验结果与分析	147		

第二部分 极化 SAR 图像分类与变化检测

第 8 章 基于 GAN 网络的极化 SAR

影像分类	171
8.1 引言	171
8.2 GAN 最新进展	172
8.3 面向任务的生成对抗网络	173
8.3.1 模型框架	174
8.3.2 极化 SAR 数据的生成 对抗网(G-Net)	174
8.3.3 判别网络(D-Net)	176
8.3.4 任务网络(T-Net)	176
8.4 实验分析	179
本章小结	184
参考文献	184

第 9 章 基于阶梯网络模型的极化 SAR

影像分类	187
9.1 引言	187
9.2 基于图的半监督学习算法	187
9.2.1 半监督学习假设	187
9.2.2 基于图的半监督学习框架	188
9.3 Wishart 距离测度	189
9.4 梯形网络原理	189
9.4.1 降噪自动编码器	190
9.4.2 梯形网络结构	191
9.5 基于 Wishart 梯形网络的极化 SAR 地物分类	194
9.5.1 构造 Wishart 图正则	194
9.5.2 Wishart 梯形网络	196
9.5.3 基于 Wishart 梯形网络的极化 SAR 地物分类	196
9.6 实验结果与分析	197

9.6.1 荷兰 Flevoland 地区 L 波段农田 小图实验结果	198
9.6.2 美国 San Francisco 地区数据 子图实验结果	203
9.6.3 荷兰 Flevoland 地区 L 波段农田 图像实验结果	207
9.6.4 荷兰 Flevoland 地区 L 波段农田 大图	211
本章小结	213
参考文献	214

第 10 章 基于 Wishart 深度堆栈网络的

极化 SAR 影像分类	216
10.1 引言	216
10.2 从极化 SAR 数据到神经网络设计	217
10.2.1 Wishart 距离	217
10.2.2 Wishart 网络	219
10.3 Wishart 深度堆栈网络	223
10.3.1 深度堆栈网络	223
10.3.2 构建 W-DSN	224
10.3.3 W-DSN 分析	226
10.4 实验分析	226
10.4.1 传统算法与快速算法的比较	227
10.4.2 WN 的有效性	227
10.4.3 W-DSN 的有效性	231
10.4.4 W-DSN 与传统 DSN 算法比较	233
10.4.5 综合评价	234
本章小结	236
参考文献	237

第 11 章 基于复数轮廓波卷积神经网络的极化 SAR 影像分类	239	13.3.3 Germany 数据集	272
11.1 复数卷积神经网络	239	13.3.4 San Francisco 数据集	275
11.2 复数轮廓波卷积神经网络的设计及数学分析	241	本章小结	278
11.2.1 复数轮廓波卷积神经网络的框架设计	241	参考文献	278
11.2.2 复数轮廓波卷积神经网络的数学分析	242	第 14 章 基于局部受限卷积神经网络的极化 SAR 影像变化检测	280
11.3 基于复数轮廓波卷积神经网络的极化 SAR 图像分类算法	244	14.1 引言	280
11.4 实验条件以及实验结果分析	245	14.2 局部受限卷积神经网络	281
11.4.1 Flevoland 数据集	245	14.2.1 任务描述	281
11.4.2 San Francisco Bay 数据集	248	14.2.2 基于 CNN 的变化检测	283
11.4.3 Germany 数据集	249	14.2.3 构建 LRCNN	284
本章小结	250	14.3 LDIs 和 DELDIs	286
参考文献	251	14.3.1 数据分析	286
第 12 章 基于加权卷积神经网络与主动学习的极化 SAR 影像分类	253	14.3.2 构建层次差异图 LDI	287
12.1 卷积神经网络	253	14.3.3 构建 DELDI	288
12.2 基于加权卷积神经网络的主动学习算法	254	14.4 学习策略	291
12.3 极化 SAR 影像分类	256	14.4.1 CNN 训练分析	291
12.3.1 PolSAR 数据集 FN15	256	14.4.2 LRCNN 训练分析	291
12.3.2 PolSAR 数据集 SU5	261	14.4.3 训练算法	293
本章小结	264	14.5 LRCNN 与传统变化检测算法对比	294
参考文献	264	14.6 实验分析	296
第 13 章 基于多尺度深度 Directionlet 网络的极化 SAR 图像分类	267	14.6.1 评价指标	297
13.1 多尺度卷积神经网络	267	14.6.2 模拟数据集	297
13.2 基于多尺度深度方向波网络的极化 SAR 图像分类	267	14.6.3 东京数据集	306
13.3 实验结果与分析	269	14.6.4 贵州数据集	307
13.3.1 实验环境与数据	269	本章小结	312
13.3.2 Flevoland 数据集	270	参考文献	313
第 15 章 基于 Looking-Around-and-Into 网络的极化 SAR 影像变化检测	315	15.1 引言	315
15.1 引言	315	15.2 LAaI 网络	315
15.2 LAaI 网络	315	15.3 基于 LAaI 网络的变化检测	317
15.3 基于 LAaI 网络的变化检测	317	15.3.1 注意力建议卷积自编码网络	317

15.3.2 递归卷积神经网络	319	15.4.2 重点区域的多尺度变化检测结果	329
15.3.3 学习过程	321	本章小结	332
15.4 实验分析	323	参考文献	332
15.4.1 候选区域的选取	323		

第三部分 高光谱影像分类

第 16 章 基于胶囊网络的高光谱影像分类	337	高光谱影像分类	354
16.1 普通卷积层	337	17.1 卷积神经网络	354
16.2 Primary Caps 层	338	17.1.1 全卷积神经网络	354
16.3 Digit Caps 层	339	17.1.2 残差网络	355
16.3.1 层级间的传播与分配过程	339	17.1.3 ROI Align	356
16.3.2 Dynamic Routing 算法	341	17.2 空谱解耦合双通道 CNN	357
16.3.3 损失函数	341	17.2.1 光谱模块	357
16.4 重构与表示层	341	17.2.2 二维空间卷积神经网络	357
16.5 基于胶囊网络的高光谱影像分类(HSICC)	342	17.2.3 整体模型	358
16.5.1 实验设置	343	17.3 实验结果与分析	359
16.5.2 实验结果与分析	344	17.3.1 Indian Pines 数据集	359
本章小结	351	17.3.2 Pavia University 数据集	360
参考文献	352	17.3.3 Salinas-A 数据集	361
		本章小结	362
		参考文献	362
第 17 章 空谱解耦合双通道卷积神经网络的			

第四部分 遥感影像解译描述与分类

第 18 章 基于快速区域卷积神经网络的遥感语义描述	367	18.3.4 基于长短时记忆模型的文本预测	374
18.1 引言	367	18.4 实验设计与结果分析	375
18.2 快速区域卷积神经网络	367	18.4.1 实验设计	375
18.3 基于语义嵌入的快速区域卷积神经网络模型	370	18.4.2 实验结果及分析	380
18.3.1 基于快速区域卷积神经网络的特征提取	370	本章小结	385
18.3.2 基于双向循环神经网络的文本特征提取	371	参考文献	385
18.3.3 基于概率模型的图文匹配	373	第 19 章 基于局部响应卷积递归神经网络的遥感语义描述	387
		19.1 引言	387
		19.2 视觉响应机制	387

19.3 基于局部响应的卷积递归神经网络	389	20.3 结构区域图及其在 SAR 图像相干斑抑制中的应用	418
19.3.1 模型结构	389	20.3.1 高分辨 SAR 图像的初级视觉语义层	418
19.3.2 基于局部响应的卷积递归神经网络	390	20.3.2 抑制相干斑任务驱动的结构区域图的产生	419
19.4 实验设计与分析	392	20.3.3 基于几何核函数测度和匀质区域搜索的 SAR 图像相干斑抑制	420
19.4.1 实验设计	392	20.4 语义素描图及其在 SAR 图像相干斑抑制中的应用	427
19.4.2 实验结果与分析	394	20.4.1 研究动机	428
本章小结	397	20.4.2 语义素描图中方向信息的传递	428
参考文献	398	20.4.3 基于几何结构块相似性测度的非局部均值滤波方法	429
第 20 章 语义空间和像素空间信息交互联合推理框架	400	20.4.4 基于像素分类和自适应邻域搜索的 SAR 图像相干斑抑制	431
20.1 SAR 图像相干斑抑制研究现状和研究动机	400	参考文献	443
20.2 高分辨 SAR 图像的素描图	402		
20.2.1 Marr 的视觉计算理论	402		
20.2.2 光学图像的初始素描模型	403		
20.2.3 初始素描图提取方法	405		
20.2.4 高分辨 SAR 图像的素描模型	407		
20.2.5 高分辨 SAR 图像素素描图提取方法	408		

第一部分

SAR 图像分类与变化检测

1.1 引言

由于强大的非线性表达能力, 神经网络给自然图像分类领域带来了显著的提升。然而, 这些针对自然图像设计的算法并不能直接应用到 SAR 图像中。SAR 图像的处理方式与自然图像有一定的区别。SAR 图像本质上是一种微波相位相干叠加的成像手段, 具有相位相干处理的特性, 没有光学图像视觉上的直观性, 而自然图像则是可见光成像, 与人的视觉感知较为一致。另外, SAR 图像的成像角度是高空对地成像, 而自然图像一般是水平成像。将自然图像中的算法应用到 SAR 图像中时, 则需要考虑这些区别, 以解决这些差异带来的问题。

目前, 深度学习在 SAR 图像目标分类中已经得到了应用, 并且取得了一定的效果。由于自然图像和 SAR 图像在数量级上存在差异, 直接应用为自然图像设计的网络进行训练必然会带来过拟合的问题, 故需要对网络进行一定的改进。2016 年, Chen S 等人提出了一种基于深度卷积神经网络的 SAR 图像目标分类方法, 此方法所阐述的对抗过拟合的方式主要包括数据扩充以及使用卷积层代替全连接以减少参数。Furukawa H 等人提出了一种基于深度残差网络的 SAR 图像目标分类方法, 此方法使用的深度网络是残差网, 通过减少卷积核个数来降低参数量, 并获得了更好的泛化性能。

基于深度残差网和可变形卷积操作, 本章提出了一种基于可变形卷积残差网(DC-ResNet)的 SAR 图像目标分类算法。

1.2 深度残差网络基础

随着深度卷积神经网络的发展, 卷积神经网络的隐层越来越深, 结构也变得更加复杂。在理论上, 由于神经网络的非线性表达能力, 越深的隐层意味着越强的表达能力, 与此同时, 越来越深的网络层数也带来了训练上的问题。目前神经网络使用的学习算法是 BP (Back Propagation, 反向传播) 算法, 主要是通过将预测结果与真实类别之间的误差回传到每一层, 并根据回传的误差调整参数值的方式进行学习。误差反向传播的计算方式是链式法则。这意味着, 越深的网络越容易出现梯度爆炸或者梯度消失的问题。为了解决这个问

题,人们提出了很多新的网络模型设计。2015年, Kaming He 等人提出的残差网在分类、目标检测等任务中均表现出优异的性能,并于2015年在 ILSVRC 竞赛中获得了 ImageNet 数据集分类的第一名。残差网中使用的跳跃连接、批规范化等操作也成为了深度神经网络中常用的提升性能的手段。

1.2.1 非线性激活函数

激活函数能够为神经网络带来非线性的表达效果,并提升神经网络的表达能力,是神经网络的重要组件之一。Sigmoid 函数曾经是神经网络中使用最为广泛的激活函数。Sigmoid 函数如式(1-1)所示:

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (1-1)$$

其中 x 表示输入, $\text{Sigmoid}(x)$ 表示输出。Sigmoid 函数对输入值进行非线性变换,将输出值控制在 0 到 1 之间。

由于 Sigmoid 函数均值不为零等问题,使得网络的收敛速度变慢。修正线性单元 (Rectified Linear Unit, ReLU) 激活函数,也称为线性整流函数,克服了这一缺点。ReLU 函数如式(1-2)所示:

$$\text{ReLU}(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1-2)$$

ReLU 函数将输入信号中小于 0 的部分输出设置为 0,大于 0 的部分保持不变。通过这一变换处理,为神经网络带来了非线性的表达效果。目前,深度神经网络中,使用的隐层激活函数通常是 ReLU 函数及其变体。

1.2.2 Dropout 操作

Dropout 操作是一种缓解过拟合问题的有效策略。在每次的训练中,Dropout 操作随机选择一部分的结点丢弃后(这一部分结点不参与计算和更新)得到一个新的子网络,然后对这个子网络进行训练和更新。Dropout 操作的示意图如图 1.1 所示。

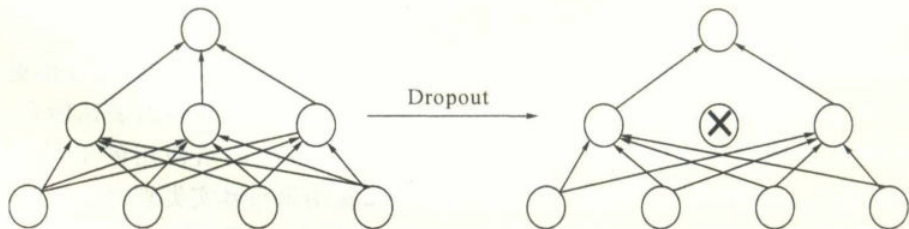


图 1.1 Dropout 操作示意图

假设一个使用了 Dropout 策略的隐层有 n 个结点, 每次随机采样选择一半的结点作为子网络, 而其余一半的结点不用更新, 则相当于有 2^n 个子网络在进行训练。在测试时, 所有的结点参与运算。这个过程相当于一个集成学习(ensemble learning)的过程, 所有的 2^n 个子网络共同决定隐层的输出, 能够有效地对抗过拟合的问题。

1.2.3 批规范化

在神经网络中, 使用数据预处理能够加快网络的收敛速度, 提高模型训练的效率, 也能够减少模型过拟合。在神经网络中, 初始的权值一般是根据标准高斯分布随机设置的, 因此, 经过权值处理操作后, 每维的均值也是与输入成一定比例关系的。如果不对数据进行归一化, 则为了学习到拟合数据的权值, 权值需要进行更多次的迭代。同时, 在这种不平衡的学习过程中, 容易使模型学习陷入局部最优值, 带来过拟合的问题。因此, 使用数据预处理, 能够加快模型收敛, 减少模型过拟合。

在深度神经网络中, 表示误差会随着层数的加深而逐步放大, 即使对输入进行了预处理, 也可能在多次的卷积处理后, 输出数据的均值和标准差产生较大的变化, 降低了模型的代表能力和训练效率。因此产生了一种新的处理方式, 即在每一层卷积中都使用批规范化处理。

批规范化(Batch Normalization, BN)是指对于每一个批次的的数据, 在隐层输出之后增加类似的归一化处理, 即对每个维度(每幅特征图)的数据, 计算出该维度的均值和标准差, 然后对该维度的数据进行减去该均值并除以该标准差的归一化操作, 得到规范化后的输出, 再送到下一层中, 其处理流程如下所述。

输入:

一个小批次上的 x 值:

$$B = \{x_1, \dots, x_m\} \quad (1-3)$$

输出:

$$\{y_i = \text{BN}_{\gamma, \beta}(x_i)\} \quad (1-4)$$

步骤 1 计算批次均值:

$$\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad (1-5)$$

步骤 2 计算批次方差:

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (1-6)$$

步骤 3 对 x 进行归一化:

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (1-7)$$

步骤4 计算输出值:

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \text{BN}_{\gamma, \beta}(x_i) \quad (1-8)$$

在每一层的激活函数之前使用批规范化操作能够对数据进行归一化, 得到更加稳定的输出结果, 防止误差的累积, 从而提高了神经网络的表达能力。

1.2.4 全局均值池化

全局均值池化是传统的卷积神经网络中全连接层的一种替代方案。传统的神经网络最后通常会有全连接层。然而, 由于在全连接的连接方式中, 全连接层之内没有参数共享机制, 参数量是输入和输出维度的乘积。在传统的卷积神经网络中, 全连接层的参数常常占到网络参数的一半以上。由于参数太多, 使用全连接层常常带来过拟合的问题。

全连接层的主要目的是去除特征的空间性, 得到全局的信息, 为最终的类别判断做准备。全局均值池化能够达到同样的目的。全局均值池化认为, 可以使用每个维度的特征图来表示一定的信息, 即信息在维度之间分离, 每个维度表示某一个方面的信息, 例如表示属于某个类别的置信度等。这种信息分布是可以通过模型的学习得到的。因此, 通过从每个维度的特征图中提取出一个值表示该方面的信息, 使得二维特征图内部的空间相关性得到了消除。

常用的池化方式分为两种: 最大池化和均值池化。最大池化摒弃了除最大值以外的信息, 而均值池化是对所有的值取均值, 因此后者能够更多地反映全局的信息。本书算法中使用的是全局均值池化。

全局均值池化的窗口大小为输出特征图的尺寸, 如图 1.2 所示。对每幅特征图, 通过全局均值池化, 整幅图像取均值, 得到一个输出值。最终得到一个和特征图维度相等的特征向量。全局均值池化操作没有任何参数, 但是却提取到了特征图中的特征信息, 去除了特征图的空间相关性, 同时带来了显著的特征约减的效果。

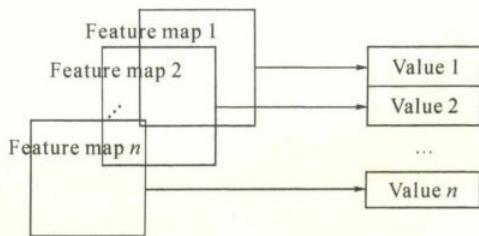


图 1.2 全局均值池化示意图

1.2.5 跳跃连接

假设在一个网络表示层中, 输入为 x , 网络映射为 $F(x)$, 输出为 $H(x)$ 。直接让表示层