



北京理工大学“双一流”建设精品出版工程

Transfer Learning: Foundations and Applications

迁移学习基础及应用

吴心筱 王晗 武玉伟 ◎ 编著



 **北京理工大学出版社**
BEIJING INSTITUTE OF TECHNOLOGY PRESS

迁移学习基础及应用

吴心筱 王 晗 武玉伟 编著

内 容 简 介

本书系统地阐述了迁移学习的解决方法和典型应用。首先,论述了迁移学习的基本概念、方法分类及发展历程,介绍了迁移学习的相关基础知识。其次,探讨了迁移学习的基本方法,包括基于样本、基于特征、基于模型和基于关系的迁移学习方法,阐述了深度迁移学习的经典方法,包括神经网络自适应迁移方法和神经网络对抗迁移方法,介绍了更加实用的部分域适应方法和开集域适应方法。最后,介绍了迁移学习在动作识别、目标检测及语义分割三个方向的应用。

本书可供从事机器学习、模式分类、人工智能以及相关领域研究和应用的技术研发人员参考,也可作为相关专业的高年级本科生和研究生的教材。

版权专有 侵权必究

图书在版编目(CIP)数据

迁移学习基础及应用 / 吴心筱, 王晗, 武玉伟编著
· -- 北京: 北京理工大学出版社, 2021. 5
ISBN 978 - 7 - 5682 - 9861 - 2

I. ①迁… II. ①吴… ②王… ③武… III. ①机器学习—研究 IV. ①TP181

中国版本图书馆 CIP 数据核字 (2021) 第 098190 号

出版发行 / 北京理工大学出版社有限责任公司

社 址 / 北京市海淀区中关村南大街 5 号

邮 编 / 100081

电 话 / (010) 68914775 (总编室)

(010) 82562903 (教材售后服务热线)

(010) 68944723 (其他图书服务热线)

网 址 / <http://www.bitpress.com.cn>

经 销 / 全国各地新华书店

印 刷 /

开 本 / 787 毫米 × 1092 毫米 1/16

印 张 / 15

彩 插 / 5

字 数 / 352 千字

版 次 / 2021 年 5 月第 1 版 2021 年 5 月第 1 次印刷

定 价 / 79.00 元

责任编辑 / 刘 派

文案编辑 / 国 珊

责任校对 / 周瑞红

责任印制 / 李志强

图书出现印装质量问题, 请拨打售后服务热线, 本社负责调换

彩 插

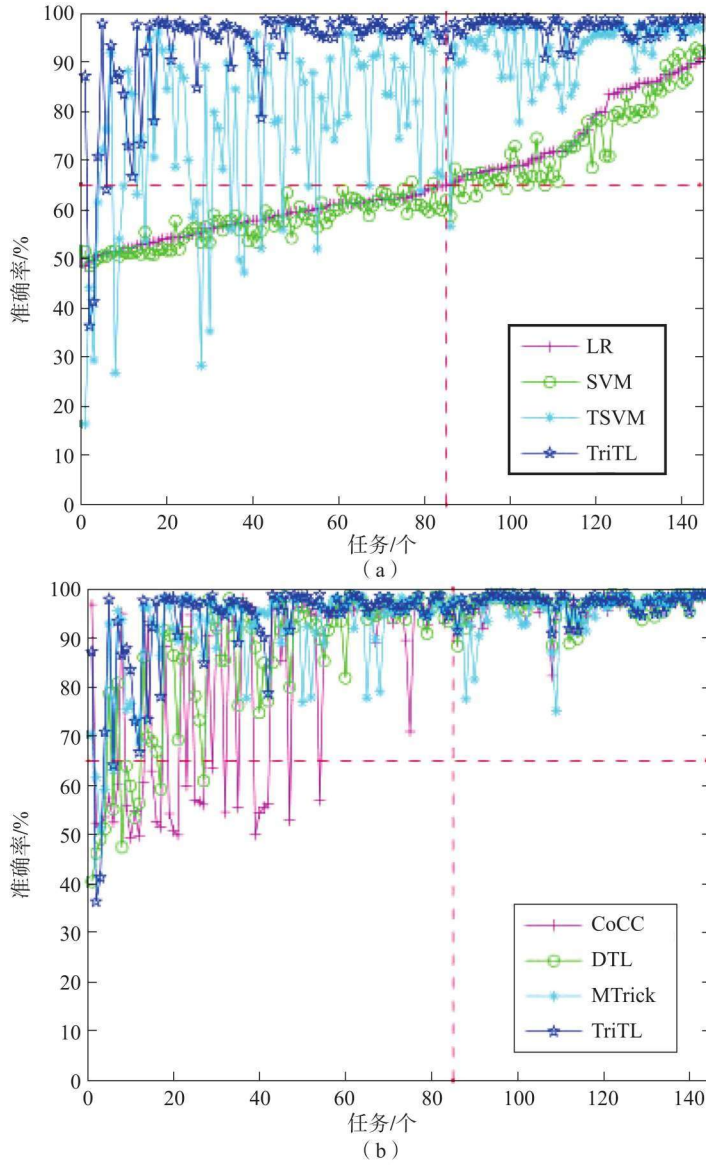
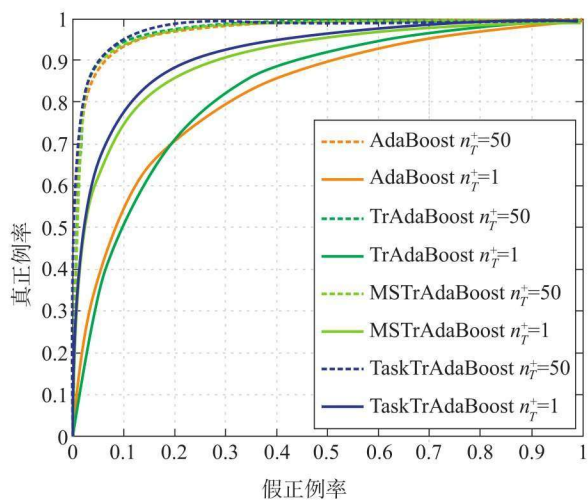
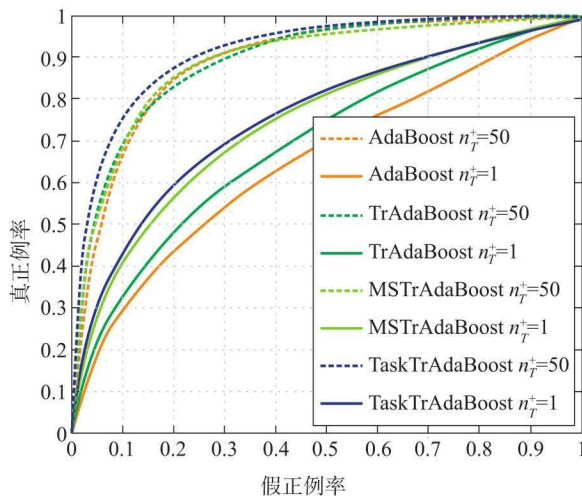


图 3.7 性能比较^[31]

(a) TriTL 和 LR、SVM、TSVM 相比；(b) TriTL 和 CoCC、DTL、MTrick 相比



(a)



(b)

图 3.8 两方法的性能比较^[7]

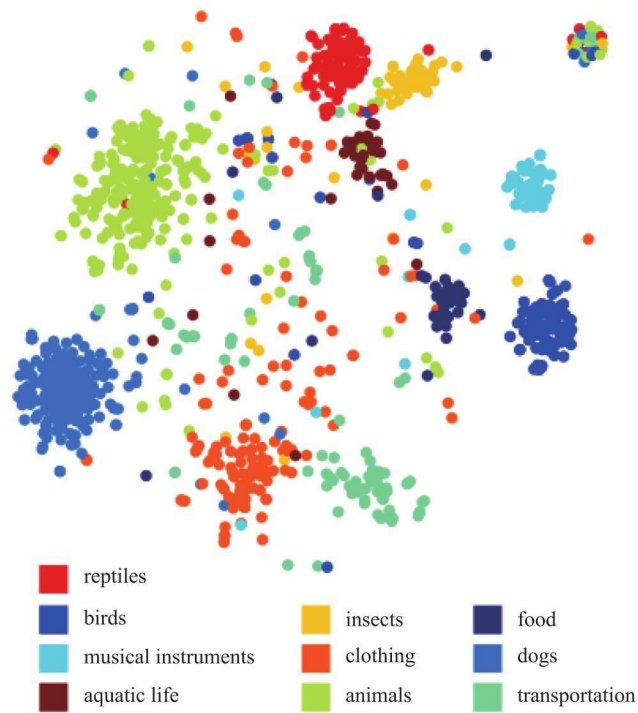


图 4.10 ILSVRC 2012 1K 中使用 skip-gram 神经语言模型的标签嵌入学习的 t-SNE 可视化子集^[18]

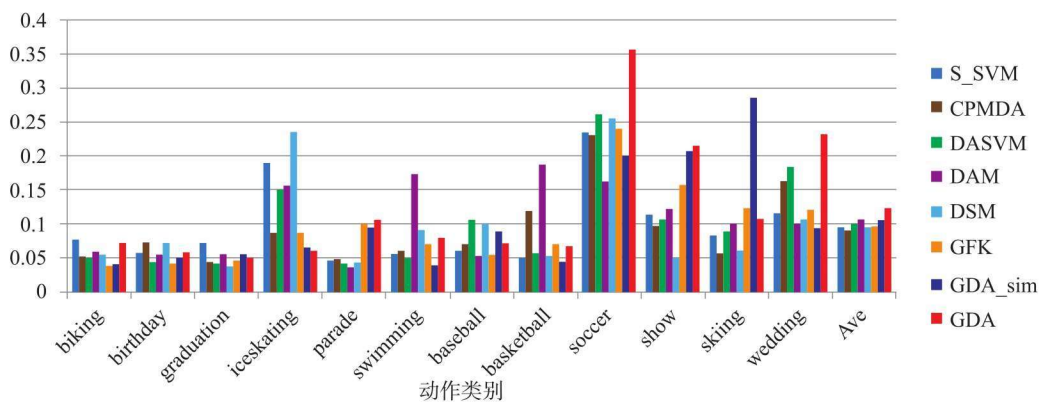


图 6.7 不同方法在 CCV 上的平均准确率^[63]

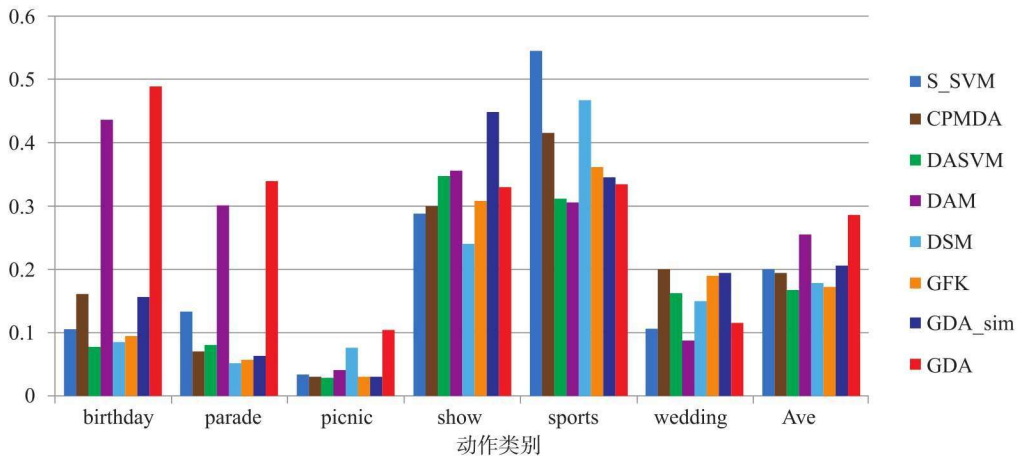


图 6.8 不同方法在 Kodak 上的平均准确率^[63]

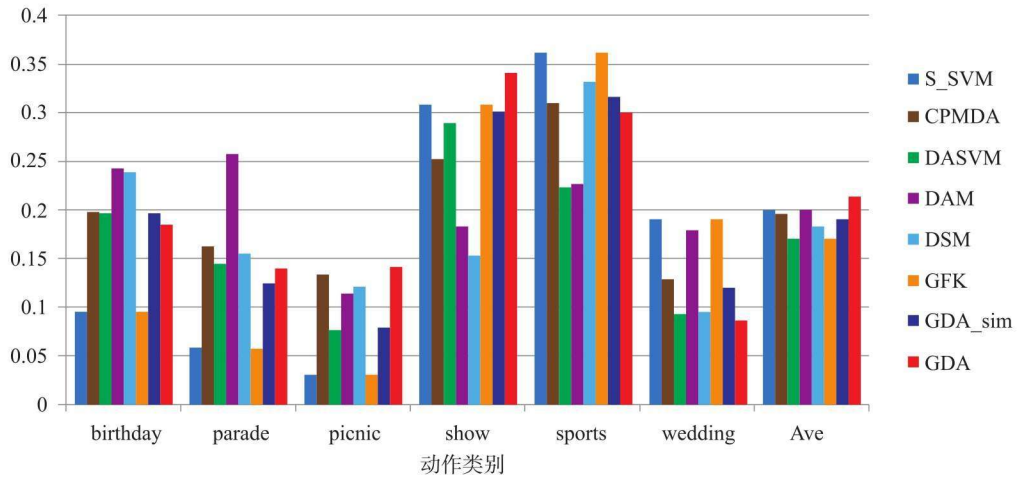


图 6.9 不同方法在 YouTube 上的平均准确率^[63]

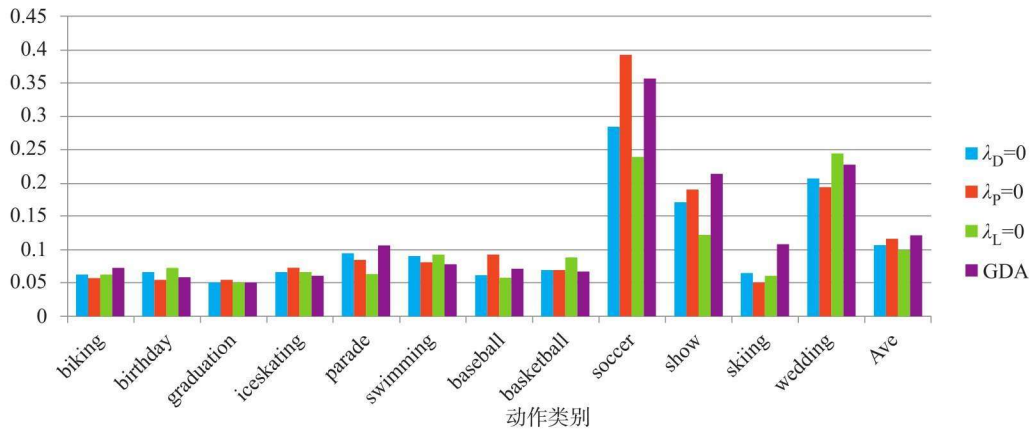


图 6.10 不同正则项对迁移性能的影响^[63]

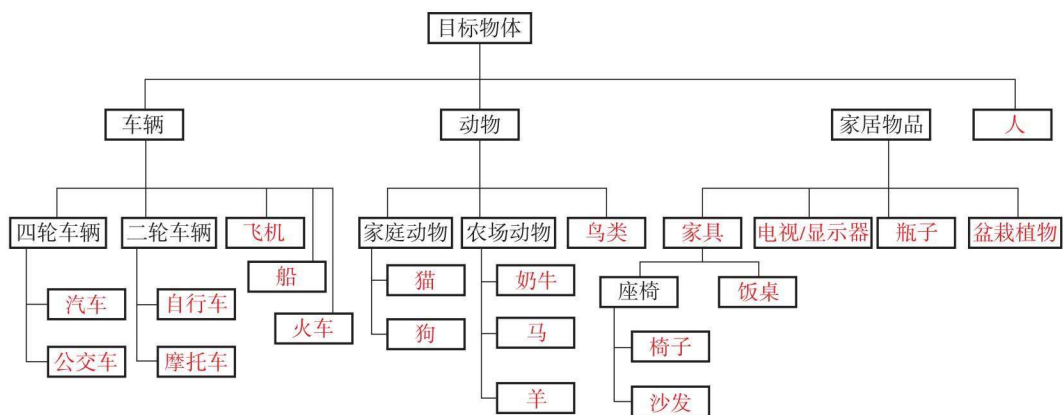


图 7.10 PASCAL VOC 数据集类别概况^[25]

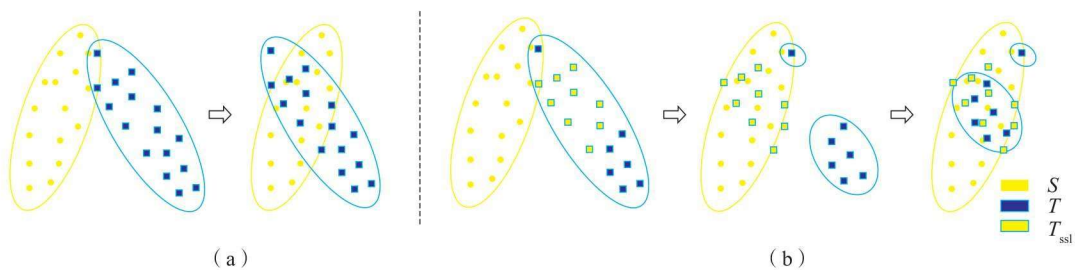


图 8.7 双向学习域适应语义分割中的自监督学习过程^[22]

(a) 步骤一；(b) 步骤二

注：图中的点表示像素点

致 谢

感谢北京理工大学研究生院对本书出版的资助。感谢北京理工大学出版社对本书的出版给予的支持和帮助。本书的撰写过程得到了北京理工大学研究生教育培养综合改革项目的资助，在此表示衷心感谢。

最后，特别感谢作者的家人在作者从事研究和撰写本书过程中给予的关心和支持。

机器学习是人工智能的一种主要实现途径，其主要研究如何让计算机模拟人类的学习行为，从数据中获取新的知识或技能，并重新组织已有的知识结构使之不断改善自身能力。传统机器学习方法依赖于大量带标签数据，且假设训练数据和测试数据来自相同的数据分布。然而，收集足够的训练数据通常是昂贵且耗时的，甚至在有些实际应用场景中是无法实现的。同时，随着时间的推移，新采集的测试数据不可能完全服从与训练数据相同的数据分布。在此背景下，迁移学习应运而生，其主要解决如何减少训练数据的标注代价，将在已标注数据域学习的模型有效迁移至新的无标注数据域，使之适应于目标数据域的任务。

自1995年机器学习领域中的迁移学习概念首次被提出，到1998年迁移学习研究领域的正式形成，再到2005年迁移学习的重新定义，再到现在基于深度神经网络的迁移学习，迁移学习引起了越来越多研究人员的关注。迁移学习的思想同样被其他机器学习方法所借鉴，如主动学习、在线学习、度量学习和强化学习。由于其出色的知识迁移和模型适应能力，迁移学习在图像识别、目标检测、情感分类、语音识别、推荐系统等诸多实际场景中具有广泛的应用。然而，迁移学习所涉及的具体任务、面临的研究难点以及采取的问题解决思路呈现多样性，导致迁移学习方法门类众多、研究成果分散，不利于初学者在短时间内系统地掌握这方面的理论、方法和技术。因此，本书对迁移学习的基本方法进行归纳、总结，同时尽量挑选具有代表性的经典方法和具有应用性的研究成果来进行介绍，以求兼顾内容的基础性和实用性。

本书分为8章。第1章为绪论，概述机器学习以及阐述迁移学习的基本概念、方法分类、发展历史及与之相关的其他学习方法。第2章简单介绍迁移学习的相关基础知识。第3章分别介绍了基于样本、基于特征、基于模型、基于关系这四类迁移学习的基本方法。第4章介绍基于深度神经网络的迁移学习方法，重点阐述神经网络自适应迁移和神经网络对抗迁移这两类方法。第5章介绍更加实用的其他迁移方法，包括部分域适应方法和开集域适应方法。第6~8章分别介绍迁移学习在动作识别、目标检测、语义分割中的应用。

本书由吴心筱、王晗和武玉伟共同编写。部分内容由研究生参与撰

写，他们是陈谨、张婷婷、刘祎程、闻子涵、伊嘉诚、滑蕊、李彤、李天宇。在此感谢学生们的辛勤劳动和付出。由于作者水平有限，书中难免存在不足，恳请同行专家和广大读者批评指正。欢迎读者来信勘误和交流。

本书可供从事机器学习、模式分类、人工智能以及相关领域研究和应用的技术研发人员参考，也可作为相关专业的高年级本科生和研究生的教材。

吴心筱 王晗 武玉伟

2020年12月2日

目 录

CONTENTS

第 1 章 绪论	001
1.1 机器学习概述	001
1.1.1 机器学习的主要步骤	001
1.1.2 机器学习的分类	003
1.2 迁移学习的基本概念	005
1.2.1 迁移学习的定义	005
1.2.2 域和任务	007
1.2.3 迁移学习的数学表示	008
1.2.4 迁移学习的基础研究问题	009
1.3 迁移学习方法的分类	010
1.3.1 根据目标域中所包含的有标签数据情况分类	010
1.3.2 根据源域与目标域的数据和任务的不同分类	011
1.3.3 根据源域与目标域的特征空间或标签是否同构分类	014
1.3.4 迁移学习方法的总结分类	014
1.4 迁移学习的发展历史	016
1.5 其他学习方法	017
1.5.1 主动学习	017
1.5.2 在线学习	018
1.5.3 度量学习	019
1.5.4 深度迁移学习	020
1.5.5 强化学习	024
参考文献	025
第 2 章 基础知识	032
2.1 矩阵分析	032

2.1.1	基本概念	032
2.1.2	矩阵分解	036
2.2	概率论	039
2.2.1	随机变量	039
2.2.2	概率分布	039
2.2.3	随机变量的数字特征	041
2.3	最优化方法	042
2.3.1	梯度下降法	043
2.3.2	牛顿法	044
2.3.3	拟牛顿法	045
2.4	神经网络	046
2.4.1	神经元模型	046
2.4.2	单层感知器	048
2.4.3	多层感知器	049
2.5	支持向量机	053
2.5.1	线性 SVM	053
2.5.2	非线性 SVM	056
	参考文献	057
第3章 迁移学习基本方法		059
3.1	基于样本迁移学习	059
3.1.1	基于样本迁移基本思想	059
3.1.2	基于样本迁移经典方法	060
3.2	基于特征迁移学习	067
3.2.1	基于特征迁移基本思想	067
3.2.2	基于特征迁移经典方法	067
3.3	基于模型迁移学习	072
3.3.1	基于模型迁移基本思想	072
3.3.2	基于模型迁移经典方法	072
3.4	基于关系迁移学习	078
3.4.1	基于关系迁移基本思想	078
3.4.2	基于关系迁移经典方法	078
3.5	异构迁移学习	080
3.5.1	同构迁移	080
3.5.2	异构迁移	080
	参考文献	082
第4章 深度迁移学习		085
4.1	深度神经网络基础	085

4.1.1	卷积神经网络	085
4.1.2	生成对抗网络	092
4.1.3	网络优化	095
4.2	深度神经网络微调迁移	099
4.2.1	网络微调基本思想	099
4.2.2	网络微调经典方法	099
4.2.3	网络微调性能分析	106
4.3	深度神经网络自适应迁移	108
4.3.1	网络自适应迁移基本思想	108
4.3.2	网络自适应迁移经典方法	108
4.3.3	网络自适应迁移性能分析	112
4.4	深度神经网络对抗迁移	114
4.4.1	网络对抗迁移基本思想	114
4.4.2	对抗迁移经典方法	115
4.4.3	网络对抗迁移性能分析	124
	参考文献	126
第 5 章	其他迁移学习	131
5.1	部分域适应	131
5.1.1	部分域适应基本思想	131
5.1.2	部分域适应经典方法	132
5.1.3	性能分析	140
5.2	开集域适应	143
5.2.1	开集域适应基本思想	144
5.2.2	开集域适应经典方法	145
5.2.3	性能分析	150
	参考文献	152
第 6 章	迁移学习在动作识别中的应用	155
6.1	动作识别介绍	155
6.2	动作识别基本方法	155
6.2.1	动作特征提取	155
6.2.2	动作分类	158
6.2.3	动作识别深度模型	159
6.3	迁移学习动作识别	162
6.3.1	自适应多核学习动作识别	163
6.3.2	多语义分组域适应动作识别	166
6.3.3	生成对抗学习动作识别	175
	参考文献	183

第 7 章 迁移学习在目标检测中的应用	190
7.1 目标检测介绍	190
7.2 目标检测经典方法	191
7.3 迁移学习目标检测方法	194
7.3.1 域适应目标检测	194
7.3.2 渐进域适应弱监督目标检测	196
7.4 常用数据集	198
7.4.1 PASCAL VOC 数据集	198
7.4.2 Cityscapes 数据集和 Foggy Cityscapes 数据集	200
7.4.3 Clipart1k 数据集	202
7.4.4 KITTI 数据集	203
7.4.5 SIM 系列数据集	205
7.4.6 目标检测评价指标	205
7.5 方法性能分析	206
7.5.1 域适应目标检测结果	206
7.5.2 渐进域适应弱监督目标检测结果	207
参考文献	208
第 8 章 迁移学习在语义分割中的应用	211
8.1 语义分割介绍	211
8.2 语义分割经典方法	212
8.3 迁移学习的语义分割方法	213
8.3.1 基于全局和局部对齐的域适应语义分割	214
8.3.2 双向学习的域适应语义分割	216
8.4 常用数据集	219
8.4.1 GTA5 数据集	219
8.4.2 SYNTHIA 数据集	220
8.4.3 语义分割评价指标	221
8.5 方法性能分析	222
8.5.1 基于全局和局部对齐的域适应语义分割方法结果	222
8.5.2 双向学习的域适应语义分割方法结果	225
参考文献	227

第 1 章

绪 论

1.1 机器学习概述

AlphaGo 作为围棋界的 AI（人工智能）代表，击败了许多人类顶尖棋手，向大众展示了人工智能的实力。而随着人工智能的进一步发展，越来越多的技术如无人驾驶汽车等也日渐成熟，开始走向大众的生活。除此之外，人工智能也已经渗透到医疗、金融、物流、保险等无数的领域，发挥着它的作用。

人工智能是计算机科学中一个比较宽泛的分支，是由机器所体现出来的智能。其应用主要有感知能力、认知能力、创造力以及更深的智能几大部分。其中感知能力主要对应人类的五官感知，即看、听、读、写等内容；认知能力主要指通过学习分析来产生判断结果的能力，如灾害预测、自动驾驶等；创造力主要是指产生新的、以前不存在的东西，如计算机作诗、计算机绘制漫画等。

机器学习是人工智能的一种方式，主要研究如何让计算机模拟人类的学习行为，从数据中获取新的知识或技能，并重新组织已有的知识结构使之不断改善自身。在数据科学中，算法是一系列统计处理步骤的流程。在机器学习的过程中，算法在大量数据中找到某种模式或特征，从而可以在新的数据上做出合适的决策或者预测结果。算法设计得越好，得到的决策和预测结果就越准确，能处理的数据也就越多。

如今，人们的生活中存在着各种各样的机器学习案例。数字语音助手根据语音指令就可以在互联网中搜索并播放出用户想要的音乐和电影；推荐系统产品通过查询用户的浏览记录、购买记录为用户播放相应的广告视频；医疗图像分析系统帮助医生定位他们可能忽略的那些肿瘤；家中无人时，扫地机器人自动为我们清洁房间的地板……对于机器学习，我们仍然有着更多的期待。随着数据的激增和计算设备算力的增强，机器学习终将使我们的生活变得越来越高效。

机器学习是一类算法的总称，这些算法企图从大量历史数据中挖掘出其中隐含的规律，并将其用于预测或者分类。更具体地说，机器学习可以看作寻找一个函数，输入是样本数据，输出是期望的结果，只是这个函数过于复杂，以至于不太方便形式化表达。需要注意的是，机器学习的目标是使学到的函数能够很好地适用于“新样本”，而不仅仅是在训练样本上表现很好。我们将学到的函数适用于新样本的能力，称为泛化（generalization）能力。

1.1.1 机器学习的主要步骤

机器学习根据数据生成预测模型，模型生成包含以下几个主要步骤（图 1.1）。

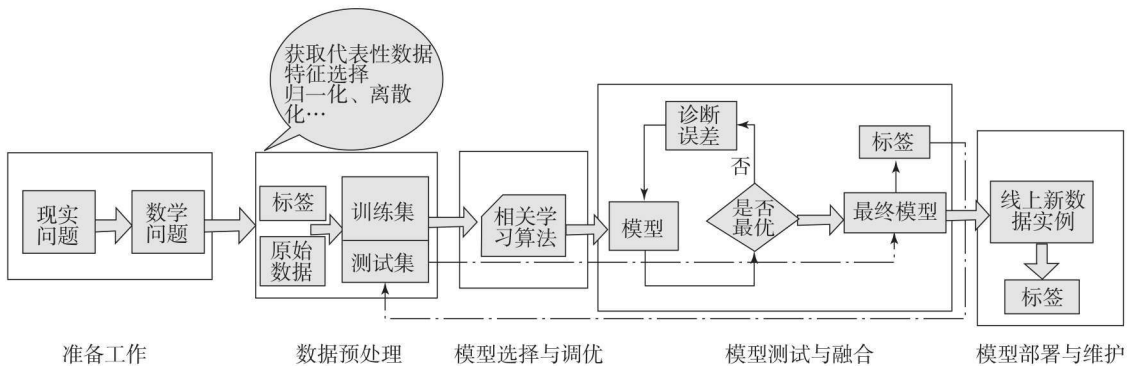


图 1.1 机器学习主要步骤

第一步，准备工作：明确问题是进行机器学习的第一步，即将现实问题抽象成数学问题。在实际应用中我们得到的并非一个明确的机器学习任务，而是一个需要解决的问题。这时我们可以将该问题抽象成数学问题，通过数学模型的方式来解决。如明确在学习过程中可以获取到的数据类型，明确学习目标为分类问题（classification problem）还是回归问题（regression problem），明确模型中的参数的作用等。根据不同的问题目标，形成不同数学表达，为后续模型选择做好准备。

第二步，数据预处理：数据预处理包括两方面，一方面是获取数据，另一方面是特征预处理与特征选择（feature selection）。数据的质量决定了机器学习结果的上限，而算法只是尽可能逼近这个上限。如在预测房屋价格时，需要通过人工获取数据，可以从和房屋相关的网站上获取数据、提取特征并进行标注。人工收集数据耗时较长且非常容易出错，只有在其他方法都无法实现时才会采用。因此，通常我们可以通过网络爬虫从相关网站收集数据，从传感器收集实测数据（如压力传感器的压力数据），从某些 API（应用程序接口）获取数据（如交易所的交易数据），从 App 或 Web 端收集数据等。对于某些领域，也可直接采用业界的公开数据集，从而节省时间和精力。此外，数据要有代表性，否则必然会过拟合（overfitting）。而且对于分类问题，数据偏移不能过于严重，即不同类别的数据所占的比例尽可能均衡。通过数据采集得到的原始数据可能并不规范，需对数据进行清洗才能满足使用需求。如去掉数据集中的重复数据、噪声数据，修正错误数据等。最后将数据转换为需要的格式，以便后续处理。对数据分布进行评估，决定使用何种特征表达。特征预处理包括标准化、归一化、信号增强等特征转换处理。这在机器学习过程中是关键步骤之一，在很大程度上影响了任何后续统计或机器学习算法的效果。通过消除数据冗余性、去除噪声，提取具有代表性的部分来提高机器学习的效率。如何获取这些特征中较为重要的部分则是特征选择所要做的工作。特征选择是模式分类、机器学习中一个关键的环节，良好的特征设计能够减少对后续机器学习算法的依赖性。特征的好坏也直接制约着整个机器学习系统的性能。因此，特征的预处理和特征选择一直是机器学习领域的一个重要的研究方向。例如在计算机视觉领域的长期研究过程中，研究人员提出各种特征提取方法用于解决具体的分类问题。这些特征包括基本的图像颜色特征、纹理特征、局部特征以及全局特征等，在基本的图像分类（image classification）和识别任务上取得了较好的应用效果。然而，这些人工设计的特征主