基于迁移学习的类别级物体识别与检测研究与进展

张雪松[1,2] 庄 严[1] 闫 飞[1] 王 伟[1]

摘 要 类别级物体识别与检测属于计算机视觉领域的一个基础性问题, 主要研究在图像或视频流中识别和定位出其中感兴趣的物体. 在基于小规模数据集的类别级物体识别与检测应用中, 模型过拟合、类不平衡和跨领域时特征分布变化等关键问题与挑战交织在一起. 本文介绍了迁移学习理论的研究现状, 对迁移学习理论解决基于小规模数据集的物体识别与检测中遇到的主要问题的研究思路和前沿技术进行了着重论述和分析. 最后对该领域的研究重点和技术发展趋势进行了探讨.

关键词 迁移学习, 物体识别, 物体检测, 小规模数据集, 类不平衡数据集

引用格式 张雪松, 庄严, 闫飞, 王伟. 基于迁移学习的类别级物体识别与检测研究与进展. 自动化学报, 2019, **45**(7): 1224−1243

DOI 10.16383/j.aas.c180093

Status and Development of Transfer Learning Based Category-Level Object Recognition and Detection

ZHANG Xue-Song[1,2] ZHUANG Yan[1] YAN Fei[1] WANG Wei[1]

Abstract Category-level object recognition and detection are the fundamental problem in computer vision, which aims to solve the challenges of identifying and localizing interested objects in a static image or dynamic video stream. For small-scale data set based category-level object recognition and detection tasks, the key issues and challenges are interwoven, such as model overfitting, class imbalance and cross-domain feature distribution shift. In this survey, we first introduce the research status of transfer learning theory, and then we focus on discussing and analyzing of the research approaches and cutting-edge technologies of how to solve the challenging problems encountered in small-scale data set based object recognition and detection applications. The research emphases and prospective technical development trends are also proposed at the end of this paper.

Key words Transfer learning, object recognition, object detection, small-scale dataset, class imbalanced dataset

Citation Zhang Xue-Song, Zhuang Yan, Yan Fei, Wang Wei. Status and development of transfer learning based category-level object recognition and detection. *Acta Automatica Sinica*, 2019, **45**(7): 1224−1243

计算机视觉是一个多学科交叉的领域, 主要研究如何从静态图像或者视频流中自动地提取、分析和理解其中的有用信息[1]. 图像数据可以使用不同类型的传感器采集, 其表现形式也是多种多样的, 例如不同分辨率的彩色图片、灰度图片、深度图片和三维激光点云等. 物体识别与检测属于计算机视觉领域的一个基础性问题, 主要任务是在静态图像或视频流中识别及定位出其中感兴趣的物体. 物体识别通常要求对于任意给定的图像能够正确地对其中包含物体的语义类别进行类标签预测, 而物体检测则还要求将图像中的物体准确地定位出来. 而且, 物体检测对于定位的精确度要求也不一样, 可以用矩形框将图像中的物体框选出来, 也可以将物体从图像中分割出来, 后者称为像素级物体检测[2].

物体识别可以分为实例级物体识别和类别级物体识别两种主要类型[3]. 对于实例级物体识别, 其主要任务是在图片或视频中找到某个特定的物体, 例如在一张含有多个行人的图片中找出航天员杨利伟. 实例级物体识别的难度主要在于感兴趣的特定物体通常被其他物体所遮挡, 而且其姿态和尺度是变化的. 目前, 实例级物体识别所采用的方法主要是从图像中提取角点、线、轮廓或者表观特征, 然后采用匹配和几何配准等方式和目标物体进行对比, 根据匹配程度找出图像中的目标物体[3]. 到目前为止, 实例级物体识别的研究已经趋于成熟, 其中比较有代表性的工作是提取 SIFT[4−5]、SURF[6] 或深度 DeCAF[7] 特征, 然后基于树或哈希算法[8] 以及几何

收稿日期 2018-02-12 录用日期 2018-05-30
Manuscript received Febrary 12, 2018; accepted May 30, 2018
国家自然科学基金 (61503056, U1508208), 辽宁省教育厅基本科研项目 (JDL2017017) 资助
Supported by National Natural Science Foundation of China (61503056, U1508208) and Fundamental Scientific Research Project of Liaoning Provincial Department of Education (JDL2017017)
本文责任编委 胡清华
Recommended by Associate Editor HU Qing-Hua
1. 大连理工大学控制科学与工程学院 大连 116024 2. 大连交通大学软件学院 大连 116028
1. School of Control Science and Engineering, Dalian University of Technology, Dalian 116024 2. Software Technology Institute, Dalian Jiaotong University, Dalian 116028

验证[9] 等方法进行特征匹配.

类别级物体识别的主要任务是判断给定的图像或视频中是否存在属于某一语义类别的物体实例, 例如判断一张图片中是否有椅子类别的物体存在. 由于在不同拍摄视角、光线条件和嘈杂场景的图像中, 属于同一类别的不同物体可能存在较大程度的类内变化, 这就给对物体类别进行数学建模带来了很大的挑战. 此外, 物体类别的定义本身属于语义层面的概念, 往往存在一词多义的现象, 比如, 多个形状和表观特征完全不同的物体都可以被人们叫做椅子. 类别级物体检测是比类别级物体识别更加严格的任务, 它不仅要求计算机能够判断出给定的图像数据中是否含有某些类别的物体, 而且还要能够准确定位出属于同一类别的物体的位置. 由于类别级物体识别与检测任务是紧密相关的, 因此其采用的方法也是相辅相成的.

为了能够对物体类别进行数学建模, 研究人员通常需要采集大量的图像数据, 然后对这些图像数据按照语义类别进行分类和手工标注. 虽然近年来不断地有各种研究机构推出自己采集和手工标注的数据集[10−11], 但是研究人员发现大量数据集都存在有偏向性[12−13] 的问题. 举例来说, 将 SUN2012 数据集[11] 训练出来的分类器在 Caltech256 数据集[14] 上进行测试, 其分类准确率将会有大幅度的下降. 在移动机器人应用中, 机器人使用异构传感器采集到的图像数据在表现形式和特征分布上差异更大, 例如用摄像头采集的彩色图像和用三维激光测距仪采集的三维点云在表现形式上就完全不同. 因此, 使用不同传感器采集的同一物体类别的图像数据的特征分布差异十分显著, 这就使得利用单一传感器采集的数据训练出来的分类器只能够适用于本领域内的物体识别与检测任务. 由于三维激光测距仪和雷达等可采集深度信息的传感器的造价比较昂贵且采集过程比较复杂, 因此三维点云数据集的规模及其所包含的物体类别数目相对于大规模 Web 图片数据集而言小很多. 当采用有监督或者半监督学习方式进行物体检测时, 分类器训练过程中使用的少量正样本通常来自训练图片中标注的矩形框或者多边形, 而大量的负样本往往直接从背景区域提取, 正负训练样本数量的绝对不平衡导致训练出的分类器模型带有偏向性且易于过拟合.

在基于小规模数据集的类别级物体识别与检测应用中, 过拟合、类不平衡和跨领域特征分布差异等关键问题与挑战交织在一起. 近年来, 随着迁移学习理论的发展与成熟, 为解决物体识别与检测中的问题与挑战提供了新的思路和解决方案. 本文围绕迁移学习理论如何应用于类别级物体识别与检测展开论述, 以问题驱动的方式对迁移学习理论在物体识别与检测中的应用研究工作进行综述. 章节安排如下: 第 1 节对迁移学习理论和类别级物体识别与检测的研究现状进行了归纳和梳理; 第 2 节对迁移学习如何解决类不平衡问题进行论述; 第 3 节对迁移学习如何解决跨领域特征分布差异进行介绍; 第 4 节对迁移学习如何解决过拟合问题进行了讨论; 第 5 节对研究重点和技术发展趋势进行说明. 最后, 对该研究方向进行总结.

1 迁移学习理论和类别级物体识别与检测的研究现状

1.1 迁移学习理论研究现状

在机器学习领域, 传统的机器学习方法通常局限于解决单一领域内的问题, 即要求训练数据和测试数据都服从相同的分布. 当训练数据和测试数据的特征分布不同时, 往往需要在新数据集上重新训练模型. 但是, 在实际应用中, 重新采集数据的成本很高甚至有时难以实现, 此时将从源领域中学习到的有用知识迁移到目标领域就变得很有必要. 自从 1995 年迁移学习研究兴起以来, 不断有研究人员将传统的机器学习方法扩展到了迁移学习中, 其中包括基于核学习的迁移学习[15−17]、基于强化学习的迁移学习[18−19]、基于流形学习的迁移学习[20] 和基于深度学习的迁移学习[21] 等.

首先, 我们给出迁移学习中一些术语的定义[22].

定义 1 (领域). 一个领域可以形式化地表示为

$$D = \{\chi, p(X)\} \tag{1}$$

其中, χ 表示特征空间, $p(X)$ 表示 $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\} \in \chi, \boldsymbol{x}_i \in \mathbf{R}^d, i = 1, 2, \cdots, n$ 的边缘分布.

定义 2 (任务). 给定某一个具体的领域 $D = \{\chi, p(X)\}$, 一个任务可以形式化地表示为

$$\Gamma = \{Y, f(\cdot)\} \tag{2}$$

其中, Y 表示类标签空间, 函数 $f(\cdot)$ 表示一个目标预测函数, 该函数可以预测一个任意实例的类标签. 一个任务的解可以从训练数据 $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^m$ 中学习获得.

Pan 等[22] 将迁移学习形式化地定义为: “给定一个源领域 D_s 和源任务 Γ_s, 一个目标领域 D_t 和目标任务 Γ_t, 迁移学习使用 D_s 和 Γ_s 中已有的知识去改进 D_t 中对目标预测函数 $f(\cdot)$ 的学习, 其中 $D_s \neq D_t$ 或者 $\Gamma_s \neq \Gamma_t$”. 因此, 迁移学习的过程就是将源领域 D_s 中解决任务 Γ_s 时学习到的 “有用” 知识迁移到目标领域 Γ_t , 结合 Γ_t 中已有的训练数据为完成任务构造出泛化能力更好的模型, 如图 1 所示.

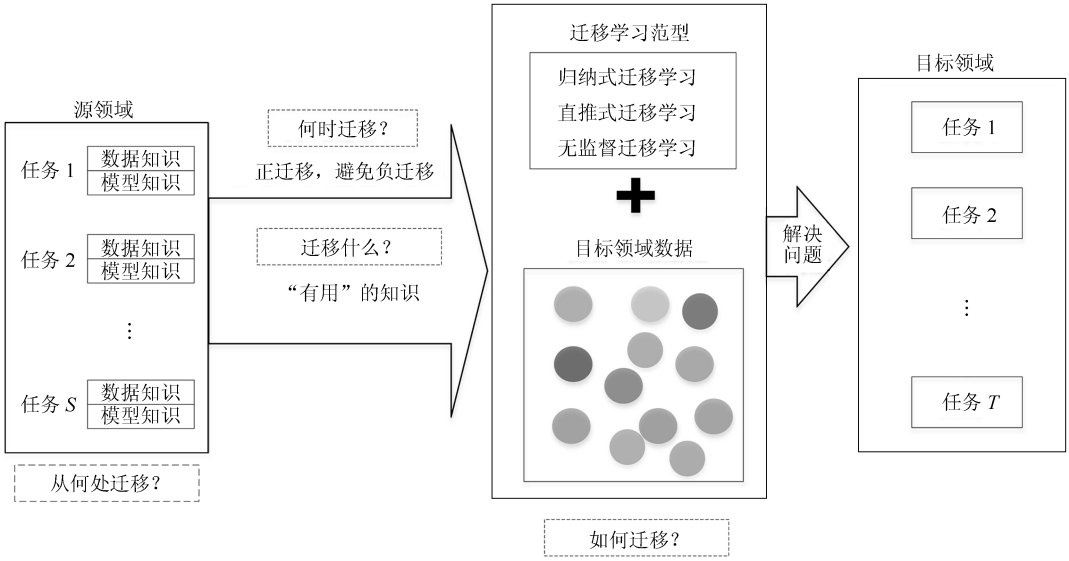


图 1 机器学习中迁移学习过程的概念表示

Fig. 1 Concept illustration of transfer learning process in machine learning

文献 [23] 中, Aytar 将迁移学习过程概括为回答下面的 4 个主要问题: 1) 从何处迁移 (Where to transfer); 2) 迁移什么 (What to transfer); 3) 如何迁移 (How to transfer); 4) 何时迁移 (When to transfer).

“从何处迁移?” 问的是被迁移到目标领域的知识的来源. 根据迁移学习的定义, 只要在源领域中存在对于完成目标领域任务有用的知识都可以将其迁移到目标领域. 值得指出, 如果迁移到目标领域的知识来自源领域或目标领域中的多个相关的任务, 则属于多任务学习 (Multi-task learning) 的范畴[24−25]. 多任务学习和迁移学习是机器学习中两个高度相关的研究领域, 多任务学习主要强调多个相关任务的共同学习, 即多个模型如何联合地进行训练. 除此之外, 还可以根据迁移学习中源领域的个数将迁移学习分为单源迁移学习 (Single source transfer learning) 和多源迁移学习 (Multi-source transfer learning)[26−29].

“迁移什么?” 问的是源领域的哪些知识可以被迁移到目标领域任务中. 由于在源领域的某些知识是局限于某个源任务的, 因此只有具有共性的知识才是可能迁移到目标领域的 “有用” 知识. 文献 [22] 将被迁移的知识类型分为数据级知识和模型级知识两种类型, 其中数据级知识主要包括实例、特征表示和关系型知识, 模型级知识指的是模型的参数. 根据从源领域迁移的知识类型, 可以将迁移学习方法分为基于实例的迁移、基于特征表示的迁移、基于参数的迁移和关系型知识的迁移 4 种主要类型, 如表 1 所示. 此外, 还可以将多种类型的知识联合迁移到目标领域中, 称为混合知识迁移 (Hybrid knowledge transfer)[30]. 例如基于实例和特征的迁移学习、基于特征和参数的迁移学习等.

“如何迁移?” 问的是使用哪种机器学习范型和方法来进行迁移. 当给定了源领域中可以迁移的知识类型之后, 只要确定了迁移学习的范型就可以根据目标领域任务的特点来设计具体的算法了. 由于表 1 中已经概括了迁移学习的常用方法, 这里重点介绍迁移学习的范型. 众所周知, 传统的机器学习理论中提出过大量的范型. 文献 [31] 根据训练数据的特点和学习目标将其分为有监督学习、无监督学习、半监督学习、归纳式学习、直推式学习和元学习等. 因为传统的机器学习范型只涉及单一领域内的学习, 所以它们不能被直接运用到迁移学习中来. 通过对传统机器学习范型定义的扩展, Pan 等[22] 最早较为系统地将迁移学习的范型分为归纳式迁移学习 (Inductive transfer learning)、直推式迁移学习 (Transductive transfer learning) 和无监督式迁移学习 (Unsupervised transfer learning) 三种. 概括地讲, 归纳式迁移学习和无监督式迁移学习允许源领域任务和目标领域任务不同但是相关, 而直推式迁移学习要求源领域任务和目标领域任务相同. 值得指出, 直推式迁移学习和领域适应 (Domain adaptation) 的研究内容相似, 但是二者存在一定差异. 在领域适应研究中, 不仅允许部分源领域训练数据有类标签, 而且也允许有少量的目标领域训练数据有类标签[32]. 文献 [33] 中指出, 对于全部目标领域训练数据有类标签的情况, 称为有监督领域适应 (Supervised domain adaptation), 对于仅少量目

标领域训练数据有类标签的情况, 称为半监督领域适应 (Semi-supervised domain adaptation), 对于全部目标领域训练数据都没有类标签的情况, 称为为无监督领域适应 (Unsupervised domain adaptation). 综上所述, 通过将迁移学习中的三种范型和表 1 中的迁移学习方法进行交叉组合, 就构成了迁移学习的主要研究内容.

表 1　迁移学习的不同方法[22]

Table 1　Different approaches to transfer learning[22]

迁移学习方法	简要描述
基于实例的迁移	通过对源领域实例加权使其用于目标领域.
基于特征表示的迁移	找到一种 “好” 的特征表示, 使其可以减少源领域和目标领域的差异以及分类和回归模型的错误率.
基于参数的迁移	发现源领域和目标领域模型中可以共享的参数或者先验知识, 使其对迁移学习有益.
关系型知识的迁移	建立源领域和目标领域关系型知识的映射. 两个领域是相关领域且每个领域中的数据非独立同分布.

“何时迁移?” 问的是迁移学习使用的时机. 显而易见, 迁移学习并不是在任何时候都是有益的, 尤其是当源领域和目标领域任务差异显著时, 进行迁移学习往往会导致负迁移的发生. 例如, 将源领域中完成骑自行车任务时学到知识迁移给目标领域中的开汽车任务必然是有害的. 虽然如何避免负迁移是一个亟待解决的问题, 但是相关的研究仍然较少. 文献 [34] 中提出了一种核方法, 可以同时用来评估不同任务的相关性和不同实例的相似性, 以此减少负迁移的可能性. Ge 等[35] 提出了一种两阶段框架, 当源领域中存在和目标领域任务完全无关知识的前提下, 仍然有效地从多个源领域中迁移 “有用” 的知识给目标领域任务, 避免负迁移的发生. 在物体识别与检测等应用中, 当源领域/任务和目标领域/任务间的差异十分显著或者目标领域的训练数据充分时, 一般不需要迁移学习.

下面以迁移学习方法为主线, 对迁移学习理论的研究现状进行概述.

1) **基于实例的迁移学习方法**

基于实例的迁移方法的主要思想是有选择性地从源领域挑选部分 “有用” 样本到目标领域, 通过扩大目标领域训练样本集的方式提高所训练出模型的泛化能力. 由于实例迁移并不进行特征层面的变换, 只是直接将源领域的样本 “借用” 到目标领域, 为了防止负迁移的发生, 一般要求源领域和目标领域数据的特征分布相似.

表 2 中归纳和总结了一些典型的基于实例的迁移学习算法, 通过对比分析可以看出集成学习、决策树和深度学习等算法已经在基于实例的迁移学习领域进行了一定的扩展性研究, 但是针对多实例迁移学习、多源和无监督方式的实例级迁移学习的理论研究仍然较为少见, 有待进一步加强.

2) **基于特征表示的迁移学习方法**

基于特征表示的迁移学习方法的主要思想是学习一种特征变换, 利用该特征变换可以将源领域的实例有效地映射到目标领域特征空间, 或者将源领域和目标领域的实例映射到一个共同的子空间. “有效映射” 的含义是特征变换应该能够减少源领域和目标领域训练数据在特征分布上的差异且不破坏数据的可分性. 因此, 利用特征变换后的源领域和目标领域样本训练模型实质上相当于通过增加训练数据来提高模型的泛化能力.

表 3 中归纳和总结了一些基于特征表示的迁移学习方法, 通过对比分析可以看出基于无监督或半监督的单源特征迁移学习算法研究较多, 多源特征迁移学习的理论研究仍然较少, 有待进一步加强.

3) **基于参数的迁移学习方法**

基于参数的迁移学习方法的主要思想是, 如果源领域任务和目标领域任务相似, 那么源领域模型中的有些参数有助于学习目标领域模型, 将从源领域学到的模型参数作为初始值训练出的目标领域模型要比用随机初始化参数值的方式更有效.

表 4 中归纳和总结了一些基于参数的迁移学习方法, 通过对比分析可以发现早期的参数迁移学习方法主要围绕支持向量机模型进行. 近几年, 针对深度卷积神经网络、贝叶斯网络和随机森林等模型的参数迁移研究已取得初步的成果. 但是, 多源的半监督和无监督方式的参数迁移学习理论研究仍然较少, 有待进一步加强.

4) **关系型知识的迁移学习方法**

关系型知识迁移学习的主要思想是将源领域内数据间的关系型知识迁移到目标领域内相似角色的数据间关系中去. 只要两个领域中属于同等级数据的角色类似, 将源领域中数据间的关系型知识迁移到目标领域同等级数据间关系的学习过程中就是 “有益” 的. 由于迁移的对象是数据间的关系型知识, 所以这种迁移方式的源领域和目标领域的数据允许非独立同分布. 举例来说, 院长在高校中所扮演的角色类似于处长在工厂中扮演的角色, 而院长和教师间的关系类似于处长和职员间的关系. 因此, 只要建立了院长和处长角色的映射以及教师和技术员角色的映射, 就可以将高校领域中的院长和教师间的关系知识迁移到工厂领域中处长和技术员间关系的学习过程中.

基于关系型知识的迁移一般适用于使用图建模的网络化数据, 例如室内场景环境中物体间的语义关系和社会网络数据间的关系等. 文献 [36] 提出了

表 2 基于实例的迁移学习方法

Table 2 Instance based transfer learning methods

序号	文献	发表年份	文献中提出的算法或方法的名称	主要思想	源领域和目标领域训练数据的形式化描述	迁移学习类型
1	[37]	2007	TrAdaboost	以目标领域少量带类标签的训练数据为指导, 在每一轮 Boosting 迭代训练弱分类器时给每个源领域样本分配一个权重, 权重越高意味着将其迁移到目标领域的价值越大.	$D_{src}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})\}_{i=1}^{m}$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$, $m>>n$	有监督单源迁移学习
2	[38]	2009	TransferBoost	提出了一种基于实例集和 Boosting 的迁移方法, 根据多个源任务中的实例集合迁移到目标领域后的分类效果, 以实例集合为单位进行整体的权重调整.	$D_{src}^{k}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})^{k}\}_{i=1}^{m_k}$, $k=1,\cdots,S$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$, $\sum_k m_k>>n$	有监督多源迁移学习
3	[39]	2012	THUNTER	通过将特征映射到子空间的方法, 将源领域中异构且不含类标签的数据迁移到目标领域的聚类任务中去.	$D_{src}=\{\boldsymbol{x}_{src}^{(i)}\}_{i=1}^{m}$, $D_{Tgt}=\{\boldsymbol{x}_{tgt}^{(i)}\}_{i=1}^{n}$ $m>>n$	无监督单源迁移学习
4	[40]	2014	DMITL	提出了一种深度多实例迁移学习算法, 将传统的单实例迁移学习扩展到了多实例迁移学习.	$\chi=\{\boldsymbol{x}_i\}_{i\in I}$, $D_{src\cup tgt}=\{\gamma_g,y_g\}_{g=1}^{G}$, $\gamma_g\in\chi$	多实例单源迁移学习
5	[41]	2014	MSD-TrAdaBoost	提出了一种多源实例迁移算法, 在 TrAdaboost 算法的基础上引入动态因子和多个源领域实例, 在每一轮分类器迭代训练时多个源领域的样本都参加, 可以有效地减少负迁移的发生.	$D_{src}^{k}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})^{k}\}_{i=1}^{m_k}$, $k=1,\cdots,S$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$, $\sum_k m_k>>n$	有监督多源迁移学习
6	[42]	2015	LSSS	提出了一种从源领域训练数据中选择和目标领域任务相关的最大实例子集的方法, 并且从理论上证明了该子集在进行实例迁移时的效果优于其超集.	$D_{src}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})\}_{i=1}^{m}$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$, $m>>n$	有监督单源迁移学习
7	[43]	2017	PreIR-DT, PreSR-DT, EmbedIR-DT, EmbedSR-DT	提出了 4 种基于决策树的实例迁移方法, 当源领域和目标领域实例的特征部分相关时, 可以有效地进行实例迁移.	$D_{src}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})\}_{i=1}^{m}$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$, $m>>n$	有监督单源迁移学习

表 3 基于特征表示的迁移学习方法

Table 3 Feature representation based transfer learning methods

序号	文献	发表年份	文献中提出的算法或方法的名称	主要思想	源领域和目标领域训练数据的形式化描述	迁移学习类型
1	[44]	2006	SCL	提出了一种结构对应的方法, 通过识别源和目标领域中频繁出现的 Pivot 特征并利用这些 Pivot 特征建立跨领域的特征对应关系, 该方法适用于各种基于特征表示的有区别模型训练.	$D_{src}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})\}_{i=1}^{t}\cup\{\boldsymbol{x}_{src}^{(i)}\}_{i=t+1}^{m}$, $D_{Tgt}=\{\boldsymbol{x}_{tgt}^{(i)}\}_{i=1}^{n}$, $m>>n$	无监督单源迁移学习
2	[45]	2007	AUGMENT	在模型开始训练之前对实例的特征空间进行扩充, 即将源领域和目标领域实例中的不同位置处填充一定数量的 0 元素. 特征扩充后的训练数据集可以使用传统的机器学习算法进行领域适应.	$D_{src}^{k}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})^{k}\}_{i=1}^{m_k}$, $k=1,\cdots,S$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$, $\sum_k m_k>>n$	有监督多源迁移学习
3	[46]	2011	TCA, SSTCA	提出了一种迁移组件分析方法, 基于最大化平均差异 (Maximum mean discrepancy, MMD) 度量在共生核希尔伯特空间中学习迁移组件. 在迁移组件张成的子空间中, 显著减少了源领域和目标领域的特征分布差异并保留了数据的可分性.	$TCA: D_{src}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})\}_{i=1}^{m}$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)})\}_{i=1}^{n}, m>>n.$ $SSTCA: D_{src}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})\}_{i=1}^{m}$, $D_{Tgt}=\{\boldsymbol{x}_{tgt}^{(i)}\}_{i=1}^{k}\cup\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=k+1}^{n}, m>>n.$	TCA: 无监督单源迁移学习 SSTCA: 半监督单源迁移学习

序号	文献	发表年份	文献中提出的算法或方法的名称	主要思想	源领域和目标领域训练数据的形式化描述	迁移学习类型
4	[47]	2015	CoMuT	提出了一种为多个源领域和目标领域寻找共同特征表示的方法, 该方法采用半监督的方式将训练数据映射到一个隐空间, 利用隐空间建立源领域和目标领域的联系.	$D_{src}^{k}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})^{k}\}_{i=1}^{m_k}$, $k=1,\cdots,S$, $D_{Tgt}=\{\boldsymbol{x}_{tgt}^{(i)}\}_{i=1}^{k}\cup\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=k+1}^{n}$, $m>>n$	半监督多源迁移学习
5	[48]	2015	TSM	提出了一个异构特征空间映射方法和任务选择方法用于无监督迁移学习.	$D_{src}^{k}=\{\boldsymbol{x}_{k}^{(i)}\}_{i=1}^{m_k}$, $k=1,\cdots,S$, $D_{Tgt}^{t}=\{\boldsymbol{x}_{tgt}^{(i)}\}_{i=1}^{n_t}$, $t=1,\cdots,T$	无监督单源迁移学习和多任务学习
6	[49]	2017	Autoencoder-Based FTL	基于降噪自动编码、共享隐层自动编码和极限学习机自动编码技术提出了三种特征迁移方法用于语音情感识别任务.	$D_{src}=\{(\boldsymbol{x}_{src}^{(i)},y_{src}^{(i)})\}_{i=1}^{m}$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)})\}_{i=1}^{n}$, $m>>n$	无监督单源迁移学习

表 4　基于参数的迁移学习方法

Table 4　Parameter based transfer learning methods

序号	文献	发表年份	文献中提出的算法或方法的名称	主要思想	源领域和目标领域训练数据的形式化描述	迁移学习类型
1	[31]	2007	Regularized SVM	通过正则化适应的方式, 将源领域中训练 SVM 分类器学习到的参数迁移到训练目标领域 SVM 分类器的过程中.	源领域模型: SVM $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$	单源有监督迁移学习
2	[50]	2011	A-SVM, PMT-SVM, DA-SVM	通过正则化源领域模型和目标领域模型间距离的方法, 将源领域 SVM 模型的参数 Ws 部分迁移到目标领域 SVM 模型 W 中.	源领域模型: SVM $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$	单源有监督迁移学习
3	[51]	2016	BNPTL	提出了一个贝叶斯网络参数迁移学习算法用于推导出网络和子图的关系, 同时解决了从源网络中找到用于迁移的最相关子图以及将源领域模型参数迁移到目标模型的方法.	源领域模型: $\{\mathrm{B}N^{i}\}_{i=1}^{k}$, $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$	多源有监督迁移学习
4	[52]	2016	CNN Fine-tuning	通过实验表明, 在目标领域训练深度卷积神经网络模型时, 微调一个预训练的源领域卷积神经网络模型要比在目标领域重新训练模型更好, 这里所说的微调实质上就是基于参数的迁移.	源领域模型: CNN $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$	单源有监督迁移学习
5	[53]	2017	SER, STRUT, MIX	首先提出了 SER 和 STRUT 算法, 可以基于源领域的随机森林模型和目标领域训练数据进行参数迁移学习. 最后, 提出了一种 MIX 方法. 将两种目标领域随机森林进行了集成.	源领域模型: Random forest $D_{Tgt}=\{(\boldsymbol{x}_{tgt}^{(i)},y_{tgt}^{(i)})\}_{i=1}^{n}$	单源有监督迁移学习

一个完备的马尔科夫逻辑网络 (Markov logic networks, MLN) 迁移系统. 在 MLN 中, 关系型领域中的实体用谓词表示, 而实体间的关系用一阶逻辑表示. 该方法首先将源 MLN 中的谓词映射到目标领域, 然后在目标领域中继续修正其映射结构以提高其性能. 文献 [54] 提出了一个方法, 可以在目标领域样本极少且存在不连通单元的特殊情况下将源 MLN 中的谓词映射到目标领域, 从而实现关系型知识的迁移. 文献 [55] 中提出了一种基于高阶语义关系型知识迁移的非参数化的语义分割方法, 可以将源领域的标注图片中的场景上下文关系型知识迁移给目标领域的未标注图片.

5) **混合迁移学习方法**

混合迁移学习的主要思想是综合地运用上面 4 种迁移学习方法, 进一步提高目标领域模型的学习效果. Xia 等[56] 提出了一种样本选择 (Sample se-

lection) 和特征集成 (Feature ensemble) 的方法, 联合使用实例迁移和特征迁移来解决源领域和目标领域的边缘概率密度与条件概率密度分布的差异. 文献 [57] 提出了一种混合迁移学习机制, 分别使用实例迁移和特征迁移来解决领域特征分布差异的问题.

虽然近年来迁移学习在理论研究上取得了很多的研究成果, 但是在物体识别与检测领域应用时仍有很多新的问题需要解决. 例如, 许多迁移学习理论研究人员提出的算法只是在低维特征空间上进行了验证, 而物体类别的特征空间通常都是高维度的. 此外, 随着深度学习等基于大数据的机器学习算法在计算机视觉领域应用中取得的巨大成功, 也给深度迁移学习理论的发展和应用研究带来了新的机遇和挑战.

1.2 类别级物体识别与检测方法研究现状

1) 类别级物体识别方法研究现状

在类别级物体识别领域, 比较有影响的方法主要包括基于词包 (Bag of words, BoW) 模型[58] 的方法和基于深度卷积神经网络 (Convolutional neural network, CNN) 模型的方法两大类. 文献 [59] 中对基于词包模型的图像分类方法进行了较为系统的综述, 归纳总结了 PASCAL VOC 竞赛从 2005 年至 2012 年的冠军团队使用的方法, 指出所有冠军团队的算法都是基于词包模型. 2012 年被视为类别级物体识别领域的一个重大转折点, 主要体现在基于深度 CNN 模型的方法以较大的优势超越了基于词包模型的方法. 近几年, 随着深度学习理论在计算机视觉领域的广泛应用, 基于大规模数据集和深度 CNN 模型的方法已成为类别级物体识别与检测研究的主流方向, 研究主要集中在 CNN 体系结构的设计与 CNN 模型训练过程的优化.

表 5 对历届 ILSVRC (ImageNet large scale visual recognition challenge, ILSVRC) 竞赛中出现的基于 CNN 模型的物体识别方法进行归纳总结. 通过对比表 5 中的数据不难看出, 2010 年和 2011 年的冠军团队使用的方法是词包模型结合 SVM 分类器, 2012 年以后 ILSVRC 的冠军团队方法都是基于深度 CNN 模型的, 性能的提高主要通过设计新的 CNN 的体系结构和提出新的模型优化技术. 虽然基于大规模数据集和深度 CNN 模型的物体识别方法取得了很高的分类准确率, 但是当训练样本不足时深度 CNN 模型易于过拟合, 而且 CNN 模型的层次过深时还会出现性能退化问题. 另外, 在源领域训练得到的深度 CNN 模型不能直接应用于特征分布不同的目标领域物体识别任务. 因此, 将迁移学习理论引入类别级物体识别领域, 降低深度 CNN 模型训练对于大规模数据集和高配置硬件资源的依赖, 进一步解决跨领域特征分布差异和类分布不平衡等实践中遇到的问题很有现实意义.

2) 类别级物体检测方法研究现状

近年来, 在类别级物体检测领域, 比较有影响的方法可以分为基于滑动窗口的方法和基于深度 CNN 模型的方法两大类. 基于滑动窗口的类别级物体检测方法以可变形部件模型[67] 为代表, 2013 年基于深度 CNN 模型和选择性搜索 (Selective search) 技术的 R-CNN[68] 的框架的出现, 使得类别级物体检测进入一个全新的时期.

表 5 ILSVRC 图像分类方法[60]

Table 5 Image classification methods in ILSVRC[60]

竞赛年度	分类模型	排名	Top-5 分类准确率	创新点概述
2010	SVM	冠军	28.2 %	LBP+SIFT 特征、局部受限线性编码 + 超向量编码、SVM 分类
2011	SVM	冠军	25.8 %	压缩的高维图像签名、Product 量化、SVM 分类
2012	AlexNet[61]	冠军	16.4 %	使用了 ReLU 非线性激活单元代替传统的 Tanh 和 Sigmoid 单元、使用了重叠的 Pooling、使用了数据集扩增和 Dropout 技术
2013	ZFNet[62]	冠军	11.7 %	使用了反卷积、前面的层使用了更小的卷积核和更小的步长、通过遮挡找出了决定图像类别的关键部位
2014	GoogLeNet[63]	冠军	6.7 %	使用了 1×1 卷积和 Inception 结构、使用 Average pooling 代替了 Fully-connect
2014	VGG[64]	亚军	7.3 %	使用了 Conv-pool-fully connect 体系结构, 其中卷积层中所有卷积核都是 3×3 大小
2015	ResNet[65]	冠军	3.57 %	引入残差结构、发现 CNN 随着深度加深 CNN 会出现性能退化问题
2016	Ensemble CNN models	冠军	2.99 %	采用深度 CNN 模型的集成和深度特征融合
2017	SENet[66]	冠军	2.25 %	提出 Squeeze-and-excitation 模块、显式地建模特征通道之间的相互依赖关系、特征重标定

近几年, 基于深度 CNN 模型的类别级物体检测方法可以分为两个主要方向:

1) 基于候选区域建议 (Region proposal) 和深度 CNN 分类的两阶段 (Two stage) 物体检测方法, 代表性工作有 R-CNN[68]、SPP-NET[69]、Fast R-CNN[70]、Faster R-CNN[71] 和 Mask R-CNN[72] 等.

2) 将物体检测转换为回归问题的一阶段 (One stage) 物体检测方法, 代表性工作有 YOLO[73−74]、SSD[75]、G-CNN[76] 和 RON[77] 等.

值得指出, 由于这两个方向采用的检测框架存在差异, 导致算法的性能也各有优势, 两阶段物体检测方法的准确率占优, 而一阶段物体检测方法的检测速度占优.

表 6 中归纳和对比了典型的物体检测方法, 对每个方法的创新点进行了简要的描述. 近年来, 物体检测领域的两阶段方法和一阶段方法普遍基于深度 CNN 模型、大规模数据集和高配置的硬件设备, 并且只是在训练和测试领域不变的前提条件下表现出了十分优异的性能. 但是, 在训练样本不足、硬件资源受限和跨领域等情况下, 基于深度 CNN 的物体检测方法的性能将出现较大幅度的退化. 因此, 基于深度 CNN 和迁移学习的类别级物体检测是一个很有意义的研究方向.

2　类不平衡问题及迁移学习解决方法

分类器训练是类别级物体识别与检测中的关键, 在类偏态分布 (Skewed class distribution) 的数据集的上训练分类器会导致容易训练出有偏向性的分类器, 因为只要投机性地将多数类标签作为预测输出就会获得很高的训练正确率. 但是, 如果少数类样本的分类错误代价远远高于多数类样本, 这种带有偏向性的分类器是毫无意义的. Weiss[78] 指出, 机器学习中的类不平衡包括类相对不平衡和类绝对不平衡两种情况. 以二元分类为例, 类相对不平衡指的是分类器训练时可以提供足够数量的正训练样本, 只是正负训练样本数量的比例失衡. 类绝对不平衡指的是正训练样本数量严重稀缺且正负训练样本数量的比例失衡. 在物体识别任务中, 由于一些物体的数据采集代价很高, 导致其训练样本数量严重匮乏. 因此, 当采用有区别的 (Discriminative) 方式训练分类器时, 分类器将偏向于始终输出训练数据中多数类的标签, 这将导致分类器无法有效地识别物体. 对于物体检测任务而言, 当采用有区别的方式训练物体的检测器时, 正训练样本通常在训练图片中标注的矩形框中提取, 而负训练样本则在训练图片的背景区域随机地大量提取. 由于训练时使用的负样本数量远远多于正样本数量, 训练出来的物体检

表 6　典型的物体检测方法

Table 6　Typical object detection methods

物体检测方法	类型	创新点概述
DPM[67]	滑动窗口	HOG 特征、特征金字塔匹配、多组件模型结合可变形部件模型实现多视角物体检测
R-CNN[68]	两阶段法	基于选择性搜索 (Selective search) 算法的区域建议、提取 CNN 深度特征、矩形框回归定位
SPP-NET[69]	两阶段法	改进了 R-CNN 方法, 在卷积层和全连接层之间加入空间金字塔池化结构、提高了产生候选框的速度和计算开销
Fast R-CNN[70]	两阶段法	改进了 SPP-NET 方法, 设计了一种 ROI pooling 的池化层结构、提出多任务损失函数思想, 将分类损失和边框回归损失结合统一学习
Faster R-CNN[71]	两阶段法	改进了 Fast R-CNN 方法, 提出了 RPN (Region proposal networks) 网络
Mask R-CNN[72]	两阶段法	改进 Faster-CNN 方法, 将 ROI_Pooling 层替换成了 ROI_Align, 添加分支 FCN 层 (Mask 层) 用于语义掩码识别, 通过 RPN 网络生成目标候选框, 再对每个目标候选框进行分类判断和矩形框回归定位
YOLO[73]	一阶段法	基于图像的全局信息进行预测、将输入图像归一化到 448×448 像素大小, 划分图像为 7×7 网格, 通过 CNN 提取特征, 直接预测每个网格内的边框坐标和每个物体类别的置信度,CNN 训练时采用 P-Relu 激活函数
YOLO 9000[74]	一阶段法	采用 Darknet-19 作为特征提取网络、批归一化 (Batch normalization) 预处理、微调两阶段预训练的深度 CNN 模型、在卷积层使用 Anchor boxes
SSD[75]	一阶段法	在不同卷积层的 Feature map 上预测物体区域, 输出离散化的 Default boxes 坐标, 利用小卷积核进行边框坐标补偿
G-CNN[76]	一阶段法	将图像划分为叠加的多尺度的规则网格, 通过网格迭代来训练模型
RON[77]	一阶段法	设计了 Reverse connection 结构、通过多任务损失函数联合优化了 Reverse connection、Objectness prior 和物体检测

测器则很容易带有偏向性且易于过拟合.

近年来, 机器学习领域围绕类不平衡学习问题进行了大量的理论研究. He 等[79] 对类不平衡条件下的模型训练和度量方法给出了一个较为全面的综述. 文献 [79] 将已有的类不平衡学习方法分为三种主要的类型: 数据级 (Data-level) 方法、算法级 (Algorithm-level) 方法和成本敏感 (Cost-sensitive) 方法. 对于 "数据级方法" 而言, 在模型学习过程中对少数类训练样本的过采样 (Over-sampling) 和对多数类训练样本的欠采样 (Under-sampling) 策略被广泛用于再平衡类的偏态分布. Zhou 等[80] 通过实验比较了随机过采样 (Random over-sampling, ROS) 和随机欠采样 (Random under-sampling, RUS) 方法在训练成本敏感 (Cost-sensitive) 神经网络时的作用. 文献 [81−83] 中提出了人工合成少数类过采样技术 (Synthetic minorities over-sampling technique, SMOTE) 及一系列扩展方法来解决类不平衡问题. 尽管 ROS、RUS 和 SMOTE 技术在解决类相对不平衡问题时取得了较好的效果, 但是在解决类绝对不平衡问题时的作用仍很有限. 对于 "算法级方法" 来说, 采用的主要策略是基于类不平衡的训练数据集将采用类平衡算法训练出来的模型适配 (Adapt) 为满足类不平衡应用. 对于 "成本敏感方法" 而言, 主要策略是在训练模型时为少数类训练数据分配一个较大的权重, 而为多数类训练数据分配一个较小的权重.

到目前为止, 在机器学习的理论研究层面, 利用迁移学习解决类不平衡学习问题的研究仍然较少. Al-Stouhi 等[84] 提出了一个 Rare-transfer 算法, 通过引入依赖于类标签的校正因子来减缓源领域样本权重的收敛速度. 文献 [84] 中特别指出, Rare-transfer 算法是第一个在实例迁移框架下解决类不平衡问题的算法, 论文中对算法的验证使用的是文本挖掘应用. 文献 [85] 在类不平衡条件下对比了主流迁移学习算法和传统机器学习算法的性能. Wang 等[86] 提出了一种基于平衡分布适应 (Balanced distribution adaptation) 的深度迁移网络模型, 通过自适应地平衡边缘分布和条件分布差异来解决类不平衡问题带来的挑战. 文献 [35] 提出了一个两阶段学习框架, 可用于解决多源迁移学习过程中的负迁移和类分布不平衡问题, 并且从理论上证明了所提出方法的错误边界. Su 等[87] 基于自适应阈值化和二次加权思想提出了 4 种方法来解决类不平衡数据对自适应正则化迁移学习 (Adaptation regularization transfer learning) 算法的影响, 并且在 UCI 数据集和脑电图分类应用中验证了所提出方法的有效性.

在物体识别与检测领域, 虽然文献 [67−68] 中通过难负例挖掘 (Hard negative mining) 结合支持向量机 (Support vector machine, SVM) 的方法在类不平衡条件下训练出了准确率较高的物体检测器, 但是难负例挖掘技术通常只适用于类相对不平衡条件下的物体识别与检测任务, 即少数类训练样本的数量不是真正的稀缺 (Scarity), 只是相对于多数类训练样本而言数量上有些比例失衡. 针对物体检测中出现的类绝对不平衡问题, 文献 [88] 中提出了一个 RUS-SMOTEBoost 算法来训练分类器并将其用于物体检测的再分类阶段, 该算法在每一轮 boosting 迭代时都对目标领域的多数类训练样本进行 RUS 采样, 对少数类训练样本进行 SMOTE 采样, 以保证在类绝对不平衡时减少 SMOTE 运算的开销.

在基于迁移学习的类不平衡物体识别与检测方面, 文献 [89] 针对类不平衡的无监督领域适应问题, 提出了一种最邻近共同空间学习 (Closest common space learning, CCSL) 算法, 通过引入隐领域变量 (Latent domain variable) 来表征领域内和领域间的类标签和结构信息, 该算法的求解过程可以视为同时进行实例加权和子空间学习. Tsai 等[90] 提出了一个领域受限迁移编码 (Domain constraint transfer coding, DcTC) 算法, 该算法通过挖掘单领域或跨领域隐子空间的方式学习一个共同的特征空间, 并将其用于领域适应和分类. Wang 等[91] 从统计学的角度提出了一个平衡分布适应 (Balanced distribution adaptation, BDA) 算法, 可以自适应地权衡源领域和目标领域间边缘分布和条件分布的差异, 然后基于 BDA 算法提出了加权平衡分布适应 (Weighted balanced distribution adaptation, W-BDA) 算法, 用于解决迁移学习中的类分布不平衡问题. 在文献 [92] 中针对物体识别中的类绝对不平衡问题, 提出了一种人工合成实例迁移学习的方法, 其基本思想是利用源领域和目标领域的训练样本人工合成出一些样本, 然后通过实例迁移的方式有选择性地从源领域实例和人工合成实例中迁移一部分 "有用" 的实例到目标领域中. 通过为每个人工合成实例分配一个合理的权重, 将其动态加入 TrAdaboost 算法的每一轮弱分类器训练过程中. 文献 [89−92] 中的每个方法均通过实验验证了其在类不平衡物体识别应用中的有效性.

综上所述, 类不平衡条件下的迁移学习理论研究积累比较薄弱, 尤其缺少各种类不平衡迁移学习算法的理论损失边界的定量化证明, 这从某种程度上阻碍了算法的应用研究. 另外, 基于迁移学习理论的类不平衡物体识别与检测研究主要集中在基于实例和特征迁移的类别级物体识别范畴, 对于类绝对不平衡条件下的物体检测方面的研究比较少见.

3 跨领域特征分布差异及迁移学习解决方法

跨领域物体识别与检测的主要挑战来自于跨领域带来的特征空间差异. 近年来, 迁移学习和领域适应中的一些理论开始被广泛应用于解决跨领域物体识别和检测任务中. 值得指出, 领域适应属于迁移学习的一个分支, 侧重于研究两个领域间的特征分布不同但任务相同的迁移学习.

3.1 同构特征迁移 (Homogeneous feature transfer)

在物体识别与检测应用中, 同构特征的含义是指源领域和目标领域的特征提取和编码方法相同. 此时, 表示源领域和目标领域物体的特征向量维度相同, 物体特征分布的差异主要由图像内在差异性导致, 例如拍摄视角变换、图像分辨率或者图像纹理特性等.

文献 [93] 中提出了一个有监督的对称特征变换学习算法用于跨领域的多类物体识别, 该算法基于 Davis 等提出的信息论度量学习 (Information-theoretic metric learning) 方法[94] 学习特征变换函数. 文献 [95] 中提出了适用于目标领域分布随时间进化的连续流形适应方法 (Continuous manifold adaptation), 该方法将每一个目标领域的样本均视为隐含采样于目标领域流形的一个不同的子空间上. 文献 [96] 提出了一种自适应结构化 SVM (Adaptive structural SVM, A-SSVM) 算法用于有监督的行人检测, 同时基于高斯过程回归等方法提出一种无监督的行人检测方法. 文献 [97] 中提出了一种深度迁移网络 (Deep transfer network) 框架, 将深度神经网络用于跨领域的特征分布匹配, 并且在跨领域多类物体识别应用中对算法的有效性进行了验证. Yang 等[98] 提出了一个 Boosting 框架下的多特征学习方法, 通过迭代地学习多特征表示来解决无监督领域适应的问题. 文献 [99] 中提出了一种深度适应网络 DAN, 将深度 CNN 应用到了领域适应中. 在 DAN 中, 每个任务的隐特征表示被嵌入共生核 Hilbert 空间中进行匹配. Yan 等[100] 提出了一个最大化独立领域适应 (Maximum independence domain adaptation) 算法, 用于从领域特征中学习出最大化独立特征, 并且在跨领域多类物体识别应用中对算法的有效性进行了验证. Long 等[101] 提出了一种无监督的领域适应方法, 利用深度神经网络从有类标签的源领域训练数据和无类标签的目标领域训练数据中同时学习分类器和可以迁移的特征. 文献 [102] 基于前向神经网络和对抗式训练方法提出了一种无监督的领域对抗神经网络 DANN, 可以使用标准后向传播和梯度下降算法进行模型训练. Long 等[103] 基于希尔伯特空间嵌入 (Hilbert space embedding, HSE) 和联合最大化平均差异准则 (Joint maximum mean discrepancy, JMMD) 提出了一种联合适应网络 (Joint adaptation networks, JAN), 通过使用对抗式训练的方法最大化源领域和目标领域间的 JMMD 度量. 文献 [104] 在 Faster R-CNN 框架下提出了领域不变候选区域生成 (Domain-invariant region proposal, DRP) 网络, 基于实例和迁移学习实现了跨领域的物体检测. 文献 [105] 提出了一种弱监督的渐进式领域适应方法, 通过领域变换 (Domain transfer, DT) 和伪标注 (Pseudo-labeling, PL) 两种方法人工生成的样本对源领域检测器进行微调实现跨领域物体检测.

表 7 对文献 [93−105] 中提出的方法进行了归纳和总结, 通过对比分析可以看出当前研究主要是基于单源无监督领域适应的物体识别应用研究, 其中深度神经网络和特征迁移学习的结合性研究较多. 不难发现, 基于同构特征迁移的物体检测研究在近三年才逐渐活跃起来, 这可能与类绝对不平衡条件下的特征迁移学习理论研究仍比较薄弱有关系.

3.2 异构特征迁移 (Heterogeneous feature transfer)

在物体识别与检测应用中, 异构特征的含义是指源领域和目标领域的特征提取和编码方法完全不同. 一般来说, 异构特征迁移算法更为通用, 它们也适用于同构特征空间的物体识别与检测任务. Kulis 等[106] 扩展了文献 [93] 中的方法并且提出了一个非对称特征变换学习算法, 该算法允许源领域和目标领域的特征维数不一样. 文献 [107] 针对异构领域适应问题, 提出了一种异构特征扩增 (Heterogeneous feature augmentation) 方法, 该方法将源领域和目标领域中维度不同的特征映射到一个共同的子空间, 以此来减少跨领域特征分布的差异性. Xiao 等[108] 基于 Hilbert-Schmidt 独立性准则将目标核矩阵 (Target kernel matrix) 和源核矩阵 (Source kernel matrix) 的子矩阵进行匹配, 以此将目标领域数据映射为源领域的相似数据, 同时利用源领域数据学习一个源领域分类器. 文献 [109] 提出了一个跨领域标志选择 (Cross-domain landmark selection) 算法用于异构领域适应, 该算法可以识别跨领域的标志性数据并将其用于生成领域不变的特征子空间.

表 8 对文献 [106−109] 中提出的方法进行了归纳和总结, 通过对比分析可以看出异构特征迁移学习理论在物体识别与检测领域的应用难度很大, 目前的研究主要集中在物体识别应用领域, 且主要采用有监督和半监督学习方式. 另外, 由于嘈杂背景的干扰和类分布不平衡等因素的影响, 异构特征迁移

表 7 基于同构特征迁移的物体识别与检测方法

Table 7 Homogeneous feature transfer based object recognition and detection methods

序号	文献	发表年份	提出的算法或方法名称	迁移学习类型	适用领域	创新点简述
1	[93]	2010	Symm	有监督单源迁移学习	物体识别	跨领域度量函数学习、KNN 多分类、采集了著名的 Office 数据集
2	[95]	2014	CMA	无监督单源迁移学习	物体识别	连续演化的目标领域、领域流形学习、KNN 和 SVM 多分类
3	[96]	2014	Supervised DA-DPM, Self-adaptive DPM	有监督单源迁学习, 无监督单源迁移学习	物体检测	自适应结构化 SVM、结构敏感–自适应结构化 SVM、基于可变形部件模型的领域适应方法
4	[97]	2015	DTN	无监督单源迁移学习	物体识别	深度迁移学习、具有线性计算复杂度适用于大规模领域适应问题
5	[98]	2015	BMFL	无监督单源迁移学习	物体识别	Boosting 集成学习、堆叠去噪自编码器 mSDA、实例加权
6	[99]	2015	DAN	无监督单源迁移学习	物体识别	核学习、基于多核最大平均差异 (MK-MMD) 的深度适应网络
7	[100]	2016	MIDA, SMIDA	无监督单源迁移学习, 半监督单源迁移学习	物体识别	核学习、Feature augment、基于 Hilbert-Schmidt 独立性准则的最大化独立适应
8	[101]	2016	RTN	无监督单源迁移学习	物体识别	端到端深度学习、最小化 MMD 度量、将 AlexNet 扩展为 RTN
9	[102]	2016	DANN	无监督单源迁移学习	物体识别	加入梯度反转层的深度 (浅) 前向神经网络、对抗式训练 (Adversarial training)
10	[103]	2017	JAN	无监督单源迁移学习	物体识别	基于深度 CNN 模型的领域适应、JMMD 准则、对抗式训练
11	[104]	2018	Domain adaptive faster R-CNN	无监督单源迁移学习	物体检测	DPR 网络、基于 H-divergence 理论设计了两个领域适应模块、对抗式训练
12	[105]	2018	DT+PL	弱监督单源迁移学习	物体检测	在人工生成数据集上微调物体检测器的两阶段渐进式领域适应

表 8 基于异构特征迁移的物体识别与检测方法

Table 8 Heterogeneous feature transfer based object recognition and detection methods

序号	文献	发表年份	提出的算法或方法名称	迁移学习类型	适用领域	创新点简述
1	[106]	2011	Arc-t	有监督单源迁移学习	物体识别	异构特征非线性变换方法、特征变换独立于分类器
2	[107]	2014	HFA, SHFA	HFA: 有监督单源迁移学习, SHFA: 半监督单源迁移学习	物体识别	跨领域核学习框架、可以同时学习核函数和分类器
3	[108]	2015	SSKMDA	半监督单源迁移学习	物体识别	基于半监督核匹配的领域适应
4	[109]	2016	CDLS	有监督单源迁移学习	物体识别	通过学习跨领域 landmarks 来生成领域不变特征子空间

学习理论在物体检测中的相关研究很少见到.

虽然已经有大量的迁移学习算法通过学习特征变换函数来解决跨领域物体识别遇到的特征分布差异问题, 但是这些方法绝大多数采用的都是离线学习的方式, 即目标领域的训练数据是以离线方式获取的. 当源领域训练数据集的规模很大或者目标领域训练数据通过在线方式获取时, 采用离线方式学习特征变换通常由于内存开销过大导致无法满足实

际应用的需求. 目前, 关于在线特征变换学习的研究较为少见, Zhao 等[110] 提出了一个在线迁移学习框架用于解决同构和异构领域适应问题, 但是他们的工作主要集中于二分类在线迁移学习和文本挖掘应用. 文献 [111] 中提出了一个在线迁移学习框架用于捕获代理 (Agent) 间的交互, 指出强化学习中的迁移学习属于在线迁移学习的一种特殊情况. 文献 [112] 针对离线方式学习特征变换时内存开销过大的缺点, 提出了一系列在线线性和非线性特征变换学习算法, 通过在线学习的方式来学习特征变换函数, 有效地减少了计算机的内存开销.

4　过拟合问题及迁移学习解决方法

近年来, 基于深度学习的物体识别与检测方法取得了极大的成功, 但是这些方法适用的前提条件是有高质量的大数据和高配置的硬件设备支持. 但是, 在移动机器人环境感知领域, 当移动机器人使用三维激光测距仪等传感器获取三维点云数据时, 由于采集成本较高, 三维点云数据集的规模和含有的物体种类相对于 Web 图片数据集而言小很多. 当目标领域训练数据不足时, 训练的模型易于过拟合, 即在训练数据上表现很好, 但在测试数据上的泛化能力很差. 因此, 解决过拟合问题对于某些特殊应用需求的物体识别与检测任务而言十分重要, 迁移学习中的多任务学习、实例迁移和参数迁移可以用来解决这个问题.

4.1　多任务学习

多任务学习通过联合利用包含在多个相关任务中的有用信息, 同时改善模型完成多个任务时的泛化能力, 如图 2 所示. 由于多任务学习涉及目标领域内多个任务间知识的共享和迁移, 本文将其视为迁移学习的范畴. 对于多类物识别与体检测任务而言, 每个类别的物体识别与检测可以视为一个任务, 多任务学习的动机是联合训练多个物体分类器要比单独地训练每一个物体分类器的效果更好.

Ghifary 等[113] 提出了一种多任务自动编码器 (Multi-task autoencoder), 可以在跨领域物体识别时具有更好的泛化能力. 文献 [114] 提出了一种层次化深度多任务学习算法, 用于大规模物体识别 (例如: 自动识别成千上万个原子物体类别) 应用, 该方法中使用视觉树 (Visual tree) 的方法将视觉上相似的原子物体 (Atom object) 类别分类为一组, 以此自动标识多个任务的内部相关性. Torralba 等[115] 基于 GentleBoost 集成学习和决策树桩弱分类器提出了一种多任务特征选择类型的算法 JointBoost, 基于共享碎片特征实现复杂场景中的多类多视角物体检测任务, 该方法通过在多物体类别间寻找共享特征的方式降低了小样本带来的模型过拟合风险并且提高了模型的健壮性. 文献 [88] 在 Torralba 等[115] 工作的基础上, 利用 BA 图表示模型和再分类等技术实现了基于三维点云的多类多视角物体检测. Li 等[116] 基于带有全局输入和输出的全连通神经网络提出了一种多任务深度显著模型. 该模型使用了多任务学习框架来发现语义图像分割和显著性检测之间的内在相关性. 文献 [117] 提出了一种端到端的多任务学习方法, 可以联合学习出物体及其行为. Chu 等[118] 提出了一种基于多任务深度卷积神经网络和感兴趣区域投票 (Region-of-interest voting) 的车辆检测框架, 该设计使得卷积神经网络模型可以共享不同的车辆属性中的视觉知识. Lu 等[119] 提出了一种多模态树结构嵌入多任务学习方法, 该方法既可以选择有区别能力的特征, 也可以选择多个相似的分类任务共享的特征. 文献 [120] 提出了

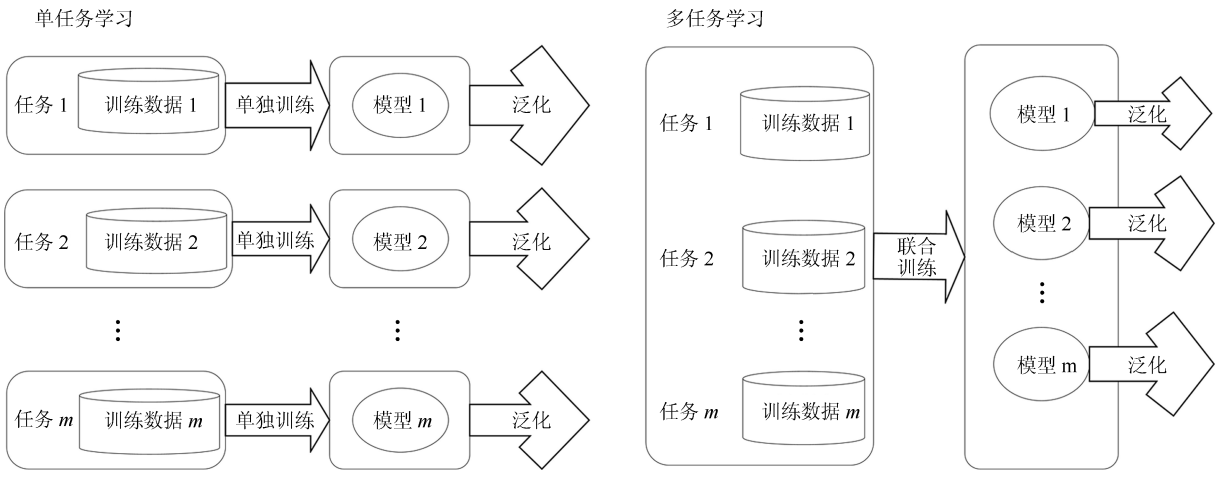


图 2　单任务学习 vs. 多任务学习过程

Fig. 2　Single task learning vs. multi-task learning process

一种联合稀疏表示和多任务学习的方法用于光谱物体检测, 利用多任务学习技术将稀疏表示模型集成到多个相关的子光谱图像中. Zhang 等[121] 提出了一种多任务相关粒子滤波 (Multi-task correlation particle filter) 方法用于物体跟踪应用中, 其中多任务相关滤波利用多个不同特征的内部相关性来联合学习相关滤波.

4.2 实例迁移

训练数据匮乏导致的分类器过拟合问题还可以通过从源领域迁移一些实例来扩大目标领域的训练数据规模的方法来解决. 由于真实场景中嘈杂背景和类分布不平衡等因素的影响, 基于实例迁移的物体检测研究较少, 主要困难在于如何有效地选择 "有用" 的实例来扩增目标领域数据集, 同时有效地避免负迁移的发生. 文献 [122] 针对多类物体数据集中部分物体类别缺少训练数据的问题, 通过从其他相关物体类别中直接迁移实例或者简单变换后再迁移实例的方式, 扩大了部分物体类别的训练数据集规模. Zhang 等[92] 提出了一种基于人工合成实例迁移学习的物体识别方法, 其基本思想是利用源领域和目标领域的训练样本人工合成出一些样本, 然后通过实例迁移的方式有选择性地从源领域实例和人工合成实例中迁移一部分 "有用" 的实例到目标领域中.

4.3 参数迁移 (也称为模型迁移)

在跨领域多类物体识别中, 有时候源领域只有若干训练好的模型却没有训练数据可以提供. 这时, 如何利用这些已有的源领域模型和少量的目标领域样本训练出泛化能力更好的目标领域模型是一个很有意义的工作. 目前, 一些基于支持向量机[23, 67] 或卷积神经网络 (Convolve neural network, CNN) 的参数迁移算法[68] 已经被用于解决这个问题. 这些算法的主要思想是以源领域模型中的部分或全部参数值作为目标领域模型训练时的参数初始值, 然后利用少量的目标领域训练数据来继续学习目标领域模型. 基于参数的迁移学习通常要求源领域和目标领域的模型使用相同的算法进行训练.

Xu 等[96] 在可变形部件模型[67] 基础上提出了两种半监督领域适应方法和一种无监督领域适应方法, 将从虚拟世界行人图片集上训练的模型知识迁移到对真实世界行人图片的物体检测任务中. Malisiewicz 等[123] 利用一个正训练标本 (Examplar) 和大量的负训练样本去训练一个标本支持向量机 (Examplar-SVM, E-SVM) 模型, 并将 E-SVM 模型集成后用于物体检测任务中. 正训练标本中包含的元数据 (Meta-data) 知识 (例如, 分割信息、几何结构和 3D 模型等) 可以直接迁移到目标领域的检测任务中. 文献 [124−125] 基于 E-SVM 的思想提出了一种基于部件的迁移正则化方法 EE-SVM, 通过从标本支持向量机模型中迁移部件级参数到新模型中构造出泛化能力更强的模型. 文献 [126] 中使用参数迁移的方法从源领域 CNN 模型中迁移参数用于训练目标领域 CNN 模型, 其中源领域模型使用的是 AlexNet 体系结构[61], 目标领域 CNN 模型的体系结构则是在 AlexNet 体系结构上针对目标领域的任务特点进行了调整修改. Hoffman 等[127] 提出了一个方法用于将 CNN 在源领域 RGB 图片集上训练的模型迁移到目标领域中基于深度图的应用中.

4.4 不同方法的对比

通过第 4.1 节 ∼ 第 4.3 节的文献综述, 可以看出多任务学习、实例迁移和参数迁移在解决训练数据匮乏所导致的模型过拟合问题上各具特点. 多任务学习适用于目标领域内有多个相关任务且每个任务的训练数据不足的情况, 此时知识的共享和迁移仅在目标领域内部进行, 不涉及到源领域中知识的迁移. 因此, 目标领域内多个任务的相关性和任务的数量对于模型的学习效果影响很大. 实例迁移在用于解决模型过拟合问题时, 主要是通过迁移单个或多个源领域的实例或者人工合成一些实例来扩增目标领域的数据集规模, 进而提高训练出的模型的泛化能力. 此时, 如何快速地挑选或人工合成出 "有用" 的辅助实例且避免负迁移很重要. 由于迁移学习中的 "负迁移" 问题尚未得到有效解决, 目前在实际应用中主要是利用图像抖动 (Image jittering) 方法 (即图像的旋转、缩放和平移等) 来扩增目标领域的数据集规模. 参数迁移在解决模型过拟合问题时, 前提条件是目标领域需要训练的模型和源领域中已有的模型体系结构一致且训练算法相同, 利用目标领域小规模的训练样本对已有的源领域模型进行参数微调. 总的来讲, 上述三种方法的适用场合各不相同, 在应用时需要根据实际情况选用.

5 研究重点和技术发展趋势

5.1 研究重点

从当前的研究现状来看, 迁移学习理论正在物体识别与检测领域进行着大量的实践应用. 所发表的研究成果主要集中在特征迁移和参数迁移的范畴内, 在利用迁移学习理论来解决类不平衡、异构特征和运动物体检测等方向的应用仍然较少. 因此, 未来的研究重点可以概括为以下 4 个方面:

1) 深度学习和迁移学习相结合

近年来, 深度学习在计算机视觉领域的应用很成功, 通过依赖大数据和 GPU 硬件设备等的支持, 将多类物体识别与检测的准确率大幅度地进行了提升. 但是, 深度学习对于数据质量和规模的要求在某

些移动机器人的应用领域是无法满足的. 因此, 如何利用迁移学习降低深度学习对大数据的依赖是个十分重要的研究工作.

2) 在线学习与迁移学习相结合

在移动机器人应用中, 在视频流中进行运动物体追踪是一个更加有挑战性的研究课题. 移动机器人通常采用多类传感器在不同环境中对感兴趣的物体进行追踪定位, 当机器人所处的外部环境不断发生变化时, 如何让机器人在进行物体检测时能够实时地适应环境的变化是个很有挑战性的研究工作.

3) 跨传感器和异构特征空间的迁移学习

由于使用不同的数据采集传感器或者特征提取和编码方法生成的物体特征分布差异显著. 例如, 将从基于 RGB 彩色图片的物体检测任务中学习到的知识迁移到基于三维点云、BA 图或者深度图的物体检测任务中是个更有挑战性的研究工作. 因此, 在跨传感器和异构特征空间下进行知识的迁移将面临来自特征分布异构、嘈杂背景干扰和小数据量等众多因素的干扰.

4) 多源知识迁移和避免负迁移问题

目前, 迁移学习领域对于多源迁移学习在物体识别与检测中的应用研究较少. 迁移学习算法在海量的多源知识中穷举搜索 "有用" 的知识会导致知识的迁移时间过长, 无法满足物体检测对实时性的需求. 另外, 源领域中 "有用" 的知识和 "有害" 的知识是并存的, 如何避免负迁移是减少迁移学习过程中嘈杂背景对于物体检测准确率影响的关键.

5.2 技术发展趋势

近年来, 随着移动机器人主动环境感知等领域对于物体识别与检测技术的需求不断变化, 对于迁移学习理论的发展和应用也提出了更高的要求, 新的基于迁移学习的物体检测方法和思路也不断涌现. 以下是对该研究领域的技术发展趋势的展望:

1) 基于深度迁移学习的物体识别与检测

基于深度学习的物体检测方法的模型训练中普遍使用了大规模数据集和深度卷积神经网络. 因此, 引入迁移学习的目的主要在于通过改进 CNN 的训练过程提高物体检测的准确率. Yosinski 等[128] 通过大量物体识别实验定量分析了卷积神经网络编码特征过程中, 哪一个卷积层编码的特征更适合进行迁移学习使用. 文献 [129] 提出了一种半监督迁移学习方法, 将多核最大平均差异 (Multi-kernel maximum mean discrepancy) 损失引入微调卷积神经网络迁移学习框架中并用于中文字符识别. Christodoulidis 等[130] 提出了一种基于多源迁移学习的方法, 使用多个源领域数据集预训练模型并将其迁移到目标领域的模式识别任务中.

2) 基于在线迁移学习的运动物体检测

近几年, 运动物体检测 (或物体跟踪) 的方法可以分为三个主要方向:

a) 基于检测的物体跟踪 (Tracking by detection), 比较有代表性的工作有 TLD[131]、Struck[132] 和 DCEM[133] 等. 该类方法与基于静态图片的物体检测方法类似, 区别在于对特征和机器学习的速度要求更高, 检测范围和尺度更小.

b) 基于相关滤波 (Correlation filter) 的物体跟踪[134], 比较代表性工作有 CSK[135]、KCF[136] 和 CN[137] 等. 该类方法的主要思想是学习一个滤波器, 然后和图像进行卷积操作得到相关信息图, 图中相关值最大的点就是物体的位置.

c) 基于深度学习的物体跟踪, 比较有代表性的工作是 DLT[138]、SO-DLT[139]、FCNT[140]、MDNet[141] 和 RTT[142] 等. 上述基于深度学习的物体跟踪方法中大量地使用了 CNN 模型的预训练和微调 (参数迁移) 的策略, 其他的迁移学习方法使用较少.

3) 基于异构特征迁移学习的物体识别与检测

近几年, 基于异构特征领域适应的跨领域多类物体识别研究的较为深入, 比较有代表性的工作有 ARC-t[106]、HFA[107]、MIDA[100] 和 CDLS[109] 等. 目前的研究主要集中在多类物体识别的应用范畴, 在跨领域物体检测方面的研究很少见, 主要原因是嘈杂背景对异构特征迁移学习的干扰较大, 影响了迁移学习的效果. 在未来, 异构特征迁移在物体检测领域的应用将是一个很有应用价值的研究工作.

6 结束语

基于迁移学习的类别级物体识别与检测的应用研究工作, 将为基于小规模数据集的物体识别与检测应用提供一系列有效的解决方案, 同时也会推动迁移学习理论的深入发展. 要在该领域取得创新性的研究成果, 必须把迁移学习理论和物体识别与检测技术有机地结合起来, 兼顾理论创新与技术创新. 此外, 迁移学习可以弥补基于大数据和深度学习的物体识别与检测方法的不足, 迁移学习的引入将为该领域的深入发展提供一个新的研究方向, 为其注入新的活力.

References

1 Szeliski R. *Computer Vision: Algorithms and Applications*. New York: Springer-Verlag, 2010. 1−28

2 Pinheiro P O, Collobert R. From image-level to pixel-level labeling with convolutional networks. In: Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 1713−1721

3 Grauman K, Leibe B. Visual object recognition. *Synthesis Lectures on Artificial Intelligence and Machine Learning*. San Rafael: Morgan & Claypool Publishers, 2011. 1−181

4 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91−110

5 Chen C C, Hsieh S L. Using binarization and hashing for efficient SIFT matching. *Journal of Visual Communication and Image Representation*, 2015, **30**: 86−93

6 Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 2008, **110**(3): 346−359

7 Donahue J, Jia Y Q, Vinyals O, Hoffman J, Zhang N, Tzeng E, et al. DeCAF: a deep convolutional activation feature for generic visual recognition. In: Proceedings of the 31th International Conference on International Conference on Machine Learning. Beijing, China: ACM, 2014. I-647−I-655

8 Friedman J H, Bentley J L, Finkel R A. An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software*, 1977, **3**(3): 209−226

9 Stewénius H, Gunderson S H, Pilet J. Size matters: exhaustive geometric verification for image retrieval. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: Springer, 2012. 674−687

10 Deng J, Dong W, Socher R, Li L J, Li K, Li F F. ImageNet: a large-scale hierarchical image database. In: Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 248−255

11 Xiao J X, Hays J, Ehinger K A, Oliva A, Torralba A. SUN database: large-scale scene recognition from abbey to zoo. In: Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 3485−3492

12 Torralba A, Efros A A. Unbiased look at dataset bias. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA: IEEE, 2011. 1521−1528

13 Tommasi T, Patricia N, Caputo B, Tuytelaars T. A deeper look at dataset bias. In: Proceedings of the 37th German Conference on Pattern Recognition. Aachen, Germany: Springer, 2015. 86−93

14 Griffin G, Holub A, Perona P. Caltech-256 object category dataset, Technical Report CNSTR-2007−001, California Institute of Technology, USA, 2007.

15 Zeng M, Ren J T. Domain transfer dimensionality reduction via discriminant kernel learning. In: Proceedings of the 16th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. Kuala Lumpur, Malaysia: Springer-Verlag, 2012. 280−291

16 Tao Jian-Wen, Chung F L, Wang Shi-Tong, Yao Qi-Fu. Sparse label propagation: a robust domain adaptation learning method. *Journal of Software*, 2015, **26**(5): 977−1000
(陶剑文, Chung F L, 王士同, 姚奇富. 稀疏标签传播: 一种鲁棒的领域适应学习方法. 软件学报, 2015, **26**(5): 977−1000)

17 Long Ming-sheng. Transfer Learning: Problems and Methods[Ph. D. dissertation], Tsinghua University, China, 2014
(龙明盛. 迁移学习问题与方法研究 [博士学位论文], 清华大学, 中国, 2014)

18 Taylor M E, Kuhlmann G, Stone P. Autonomous transfer for reinforcement learning. In: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems. Estoril, Portugal: ACM, 2008. 283−290

19 Zhu Mei-Qiang, Cheng Yu-Hu, Li Ming, Wang Xue-Song, Feng Huan-Ting. A hybrid transfer algorithm for reinforcement learning based on spectral method. *Acta Automatica Sinica*, 2012, **38**(11): 1765−1776
(朱美强, 程玉虎, 李明, 王雪松, 冯涣婷. 一类基于谱方法的强化学习混合迁移算法. 自动化学报, 2012, **38**(11): 1765−1776)

20 Cheng B, Liu M X, Suk H I, Shen D G, Zhang D Q. Multimodal manifold-regularized transfer learning for MCI conversion prediction. *Brain Imaging and Behavior*, 2015, **9**(4): 913−926

21 Long M S, Wang J M, Cao Y, Sun J G, Yu P S. Deep learning of transferable representation for scalable domain adaptation. *IEEE Transactions on Knowledge and Data Engineering*, 2016, **28**(8): 2027−2040

22 Pan S J, Yang Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 2010, **22**(10): 1345−1359

23 Aytar Y. Transfer learning for object category detection [Ph. D. thesis], University of Oxford, UK, 2014

24 Caruana R. Multitask learning. *Machine Learning*, 1997, **28**(1): 41−75

25 Zhang Y, Yang Q. A survey on multi-task learning. arXiv: 1707.08114v1, 2017

26 Ding Z M, Shao M, Fu Y. Incomplete multisource transfer learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(2): 310−323

27 Gu Xin, Wang Shi-Tong, Xu Min. A new cross-multidomain classification algorithm and its fast version for large datasets. *Acta Automatica Sinica*, 2014, **40**(3): 531−547
(顾鑫, 王士同, 许敏. 基于多源的跨领域数据分类快速新算法. 自动化学报, 2014, **40**(3): 531−547)

28 Zhang Q, Li H G, Zhang Y, Li M. Instance transfer learning with multisource dynamic TrAdaBoost. *The Scientific World Journal*, 2014, **2014**: Article No. 282747

29 Pan J, Wang X, Cheng Y, Cao G. Multi-source transfer ELM-based Q learning. *Neurocomputing*, 2014, **137**: 57−64

30 Weiss K, Khoshgoftaar T M, Wang D D. A survey of transfer learning. *Journal of Big Data*, 2016, **3**: Article No. 9

31 Li X. Regularized Adaptation: Theory, Algorithms and Applications [Ph. D. thesis], University of Washington, USA, 2007

32 Liu Jian-Wei, Sun Zheng-Kang, Luo Xiong-Lin. Review and research development on domain adaptation learning. *Acta Automatica Sinica*, 2014, **40**(8): 1576−1600
(刘建伟, 孙正康, 罗雄麟. 域自适应学习研究进展. 自动化学报, 2014, **40**(8): 1576−1600)

33 Patel V M, Gopalan R, Li R N, Chellappa R. Visual domain adaptation: a survey of recent advances. *IEEE Signal Processing Magazine*, 2015, **32**(3): 53−69

34 Shao H. Kernel methods for transfer learning to avoid negative transfer. *International Journal of Computing Science and Mathematics*, 2016, **7**(2): 190−199

35 Ge L, Gao J, Ngo H, Li K, Zhang A D. On handling negative transfer and imbalanced distributions in multiple source transfer learning. *Statistical Analysis and Data Mining*, 2014, **7**(4): 254−271

36 Mihalkova L, Huynh T, Mooney R J. Mapping and revising markov logic networks for transfer learning. In: Proceedings of the 22nd National Conference on Artificial Intelligence. Vancouver, Canada: AAAI, 2007. 608−614

37 Dai W Y, Yang Q, Xue G R, Yu Y. Boosting for transfer learning. In: Proceedings of the 24th International Conference on Machine Learning. Corvalis, USA: ACM, 2007. 193−200

38 Eaton E, DesJardins M. Set-based boosting for instance-level transfer. In: Proceedings of the 2009 IEEE International Conference on Data Mining Workshops. Miami, USA: IEEE, 2009. 422−428

39 Kong S, Wang D H. Transfer heterogeneous unlabeled data for unsupervised clustering. In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). Tsukuba, Japan: IEEE, 2012. 1193−1196

40 Kotzias D, Denil M, Blunsom P, De Freitas N. Deep multi-instance transfer learning. arXiv: 1411.3128, 2014

41 Zhang Qian, Li Ming, Wang Xue-Song, Cheng Yu-Hu, Zhu Mei-Qiang. Instance-based transfer learning for multi-source domains. *Acta Automatica Sinica*, 2014, **40**(6): 1176−1183
(张倩, 李明, 王雪松, 程玉虎, 朱美强. 一种面向多源领域的实例迁移学习. 自动化学报, 2014, **40**(6): 1176−1183)

42 Zhou S, Schoenmakers G, Smirnov E, Peeters R, Driessens K, Chen S Q. Largest source subset selection for instance transfer. In: Proceedings of the 7th Asian Conference on Machine Learning. Hong Kong, China, 2015. 423−438

43 Zhou S, Smirnov E N, Schoenmakers G, Peeters R. Conformal decision-tree approach to instance transfer. *Annals of Mathematics and Artificial Intelligence*, Switzerland: Springer, 2017, **81**(1−2): 85−104

44 Blitzer J, McDonald R, Pereira F. Domain adaptation with structural correspondence learning. In: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing. Sydney, Australia: ACM, 2006. 120−128

45 Daumé III H. Frustratingly easy domain adaptation. In: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics. Prague, Czech Republic: ACL, 2007. 256−263

46 Pan S J, Tsang I W, Kwok J T, Yang Q. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks*, 2011, **22**(2): 199−210

47 Tahmoresnezhad J, Hashemi S. Common feature extraction in multi-source domains for transfer learning. In: Proceedings of the 7th Conference on Information and Knowledge Technology (IKT). Urmia, Iran: IEEE, 2015. 1−5

48 Xue S, Lu J, Zhang G Q, Xiong L. Heterogeneous feature space based task selection machine for unsupervised transfer learning. In: Proceedings of the 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). Taipei, China: IEEE, 2015. 46−51

49 Deng J, Frühholz S, Zhang Z X, Schuller B. Recognizing emotions from whispered speech based on acoustic feature transfer learning. *IEEE Access*, 2017, **5**: 5235−5246

50 Aytar Y, Zisserman A. Tabula rasa: model transfer for object category detection. In: Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 2252−2259

51 Zhou Y, Hospedales T M, Fenton N. When and where to transfer for Bayesian network parameter learning. *Expert Systems with Applications*, 2016, **55**: 361−373

52 Tajbakhsh N, Shin J Y, Gurudu S R, Hurst R T, Kendall C B, Gotway M B, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1299−1312

53 Segev N, Harel M, Mannor S, Crammer K, El-Yaniv R. Learn on source, refine on target: a model transfer learning framework with random forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(9): 1811−1824

54 Mihalkova L, Mooney R J. Transfer learning from minimal target data by mapping across relational domains. In: Proceedings of the 21st International Jont Conference on Artificial Intelligence. Pasadena, CA: Morgan Kaufmann, 2009. 1163−1168

55 Myeong H, Lee K M. Tensor-based high-order semantic relation transfer for semantic scene segmentation. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013. 3073−3080

56 Xia R, Zong C Q, Hu X L, Cambria E. Feature ensemble plus sample selection: domain adaptation for sentiment classification. *IEEE Intelligent Systems*, 2013, **28**(3): 10−18

57 Mo Y, Zhang Z X, Wang Y H. A hybrid transfer learning mechanism for object classification across view. In: Proceedings of the 11th International Conference on Machine Learning and Applications. Boca Raton, USA: IEEE, 2012. 226−231

58 Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR′06). New York, USA: IEEE, 2006. 2169−2178

59 Huang Kai-Qi, Ren Wei-Qiang, Tan Tie-Niu. A review on image object classification and detection. *Chinese Journal of Computers*, 2014, **36**(6): 1225−1240
(黄凯奇, 任伟强, 谭铁牛. 图像物体分类与检测算法综述. 计算机学报, 2014, **36**(6): 1225−1240)

60 Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S A, et al. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 2015, **115**(3): 211−252

61 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: Curran Associates Inc., 2012. 1097−1105

62 Zeiler M D, Fergus R. Visualizing and understanding convolutional networks. In: Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014. 818−833

63 Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 1−9

64 Simonyan K, Zisserman A. Very deep convolutional networks for large-Scale image recognition. arXiv: 1409.1556v6, 2015

65 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 770−778

66 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. arXiv: 1709.01507v2, 2017

67 Felzenszwalb P F, Girshick R B, McAllester D, Ramanan D. Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(9): 1627−1645

68 Girshick R, Donahue J, Darrell T, Malik J. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(1): 142−158

69 He K M, Zhang X Y, Ren S Q, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(9): 1904−1916

70 Girshick R. Fast R-CNN. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 1440−1448

71 Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(6): 1137−1149

72 He K M, Gkioxari G, Dollár P, Girshick P. Mask R-CNN. In: Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 2980−2988

73 Redmon J, Divvala S K, Girshick R, Farhadi A. You only look once: unified, real-Time object detection. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016. 779−788

74 Redmon J, Farhadi A. YOLO9000: better, faster, stronger. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017. 6517−6525

75 Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, et al. SSD: single shot MultiBox detector. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 21−37

76 Najibi M, Rastegari M, Davis L S. G-CNN: an iterative grid based object detector. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 2369−2377

77 Kong T, Sun F C, Yao A B, Liu H P, Lu M, Chen Y R. RON: reverse connection with objectness prior networks for object detection. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017. 5244−5252

78 Weiss G M. Mining with rarity: a unifying framework. *ACM SIGKDD Explorations Newsletter*, 2004, **6**(1): 7−19

79 He H B, Garcia E A. Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 2009, **21**(9): 1263−1284

80 Zhou Z H, Liu X Y. Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Transactions on Knowledge and Data Engineering*, 2006, **18**(1): 63−77

81 Chawla N V, Bowyer K W, Hall L O, Kegelmeyer W P. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 2002, **16**(1): 321−357

82 Bunkhumpornpat C, Sinapiromsaran K, Lursinsap C. Safe-Level-SMOTE: safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. In: Proceedings of the 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining. Bangkok, Thailand: Springer, 2009. 475−482

83 Yang P Y, Yoo P D, Fernando J, Zhou B B, Zhang Z L, Zomaya A Y. Sample subset optimization techniques for imbalanced and ensemble learning problems in bioinformatics applications. *IEEE Transactions on Cybernetics*, 2014, **44**(3): 445−455

84 Al-Stouhi S, Reddy C K. Transfer learning for class imbalance problems with inadequate data. *Knowledge and Information Systems*, 2016, **48**(1): 201−228

85 Weiss K R, Khoshgoftaar T M. Comparing transfer learning and traditional learning under domain class imbalance. In: Proceedings of the 16th IEEE International Conference on Machine Learning and Applications (ICMLA). Cancun, Mexico: IEEE, 2017. 337−343

86 Wang K, Wu B. Power Equipment fault diagnosis model based on deep transfer learning with balanced distribution adaptation. In: Proceedings of the Advanced Data Mining and Applications. ADMA, Cham: Springer, 2018. 178−188

87 Su K M, Hairston W D, Robbins K A. Adaptive thresholding and reweighting to improve domain transfer learning for unbalanced data with applications to EEG imbalance. In: Proceedings of the 15th IEEE International Conference on Machine Learning and Applications (ICMLA). Anaheim, CA, USA: IEEE, 2016. 320−325

88 Zhang X S, Zhuang Y, Hu H S, Wang W. 3-D laser-based multiclass and multiview object detection in cluttered indoor scenes. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, **28**(1): 177−190

89 Hsu T M H, Chen W Y, Hou C A, Tsai Y H H, Yeh Y R, Wang Y C F. Unsupervised domain adaptation with imbalanced cross-domain data. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 4121−4129

90 Tsai Y H H, Hou C A, Chen W Y, Yeh Y R, Wang Y C F. Domain-constraint transfer coding for imbalanced unsupervised domain adaptation. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix, Arizona, USA: AAAI Press, 2016. 3597−3603

91 Wang J D, Chen Y Q, Hao S J, Feng W J, Shen Z Q. Balanced distribution adaptation for transfer learning. In: Proceedings of the 2017 IEEE International Conference on Data Mining (ICDM). New Orleans, LA, USA: IEEE, 2017. 1129−1134

92 Zhang X S, Zhuang Y, Wang W, Pedrycz W. Transfer boosting with synthetic instances for class imbalanced object recognition. *IEEE Transactions on Cybernetics*, 2018, **48**(1): 357−370

93 Saenko K, Kulis B, Fritz M, Darrell T. Adapting visual category models to new domains. In: Proceedings of the 11th European Conference on Computer Vision. Crete, Greece: Springer-Verlag, 2010. 213−226

94 Davis J V, Kulis B, Jain P, Sra S, Dhillon I S. Information-theoretic metric learning. In: Proceedings of the 24th International Conference on Machine Learning. Corvalis, Oregon, USA: ACM, 2007. 209−216

95 Hoffman J, Darrell T, Saenko K. Continuous manifold based adaptation for evolving visual domains. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014. 867−874

96 Xu J L, Ramos S, Vázquez D, López A M. Domain adaptation of deformable part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(12): 2367−2380

97 Zhang X, Yu F X, Chang S F, Wang S J. Deep transfer network: unsupervised domain adaptation. arXiv: 1503.00591, 2015.

98 Yang X S, Zhang T Z, Xu C S, Yang M H. Boosted Multifeature learning for cross-domain transfer. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2015, **11**(3): Article No. 35

99 Long M, Cao Y, Wang J, Jordan M I. Unsupervised domain adaptation with residual transfer networks. In: Proceedings of the 32th International Conference on Machine Learning. Lille, France: ACM, 2015. 97−105

100 Yan K, Kou L, Zhang D. Learning domain-invariant subspace using domain features and independence maximization. arXiv: 1603.04535, 2016.

101 Long M S, Zhu H, Wang J M, Jordan M I. Unsupervised domain adaptation with residual transfer networks. In: Proceedings of the 30th Conference on Neural Information Processing Systems. Barcelona, Spain: ACM Press 2016. 136−144

102 Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F, et al. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, 2016, **17**(1): 2096−2030

103 Long M S, Zhu H, Wang J M, Jordan M I. Deep transfer learning with joint adaptation networks. In: Proceedings of the 34th International Conference on Machine Learning. Sydney, Australia: ACM, 2017. 2208−2217

104 Chen Y H, Li W, Sakaridis C, Dai D X, Van Gool L. Domain adaptive faster R-CNN for object detection in the wild. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah, USA: IEEE, 2018. 3339−3348

105 Inoue N, Furuta R, Yamasaki T, Aizawa K. Cross-domain weakly-supervised object detection through progressive domain adaptation. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah, USA: IEEE, 2018. 5001-5009. DOI: 10.1109/CVPR.2018.00525

106 Kulis B, Saenko K, Darrell T. What you saw is not what you get: domain adaptation using asymmetric kernel transforms. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, RI, USA: IEEE, 2011. 1785−1792

107 Li W, Duan L X, Xu D, Tsang I W. Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(6): 1134−1148

108 Xiao M, Guo Y H. Feature space independent semi-supervised domain adaptation via kernel matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(1): 54−66

109 Tsai Y H H, Yeh Y R, Wang Y C F. Learning cross-domain landmarks for heterogeneous domain adaptation. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016. 5081−5090

110 Zhao P L, Hoi S C H, Wang J L, Li B. Online transfer learning. *Artificial Intelligence*, 2014, **216**: 76−102

111 Zhan Y S, Taylor M E. Online transfer learning in reinforcement learning domains. arXiv: 1507.00436, 2015.

112 Zhang X S, Zhuang Y, Wang W, Pedrycz W. Online feature transformation learning for cross-domain object category recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(7): 2857−2871

113 Ghifary M, Kleijn W B, Zhang M J, Balduzzi D. Domain generalization for object recognition with multi-task autoencoders. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 2551−2559

114 Fan J P, Zhao T Y, Kuang Z Z, Zheng Y, Zhang J, Yu J, et al. HD-MTL: hierarchical deep multi-task learning for large-scale visual recognition. *IEEE Transactions on Image Processing*, 2017, **26**(4): 1923−1938

115 Torralba A, Murphy K P, Freeman W T. Sharing visual features for multiclass and multiview object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(5): 854−869

116 Li X, Zhao L M, Wei L N, Yang M H, Wu F, Zhuang Y T, et al. Deepsaliency: multi-task deep neural network model for salient object detection. *IEEE Transactions on Image Processing*, 2016, **25**(8): 3919−3930

117 Kalogeiton V, Weinzaepfel P, Ferrari V, Schmid C. Joint learning of object and action detectors. In: Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 2001−2010

118 Chu W Q, Liu Y, Shen C, Cai D, Hua X S. Multi-task vehicle detection with region-of-interest voting. *IEEE Transactions on Image Processing*, 2018, **27**(1): 432−441

119 Lu X, Wang Y N, Zhou X Y, Zhang Z J, Ling Z G. Traffic sign recognition via multi-modal tree-structure embedded multi-task learning. *IEEE Transactions on Intelligent Transportation Systems*, 2017, **18**(4): 960−972

120 Zhang Y X, Du B, Zhang L P, Liu T L. Joint sparse representation and multitask learning for hyperspectral target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, **55**(2): 894−906

121 Zhang T Z, Xu C S, Yang M H. Multi-task correlation particle filter for robust object tracking. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017. 4819−4827

122 Lim J J, Salakhutdinov R, Torralba A. Transfer learning by borrowing examples for multiclass object detection. In: Proceedings of 24th International Conference on Neural Information Processing Systems. Granada, Spain: ACM, 2011. 118−126

123 Malisiewicz T, Gupta A, Efros A A. Ensemble of exemplar-SVMs for object detection and beyond. In: Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 89−96

124 Aytar Y, Zisserman A. Enhancing exemplar SVMs using part level transfer regularization. In: Proceedings of the 2012 British Machine Vision Conference. Guildford, UK: BMVA, 2012. 1−11

125 Aytar Y, Zisserman A. Part level transfer regularization for enhancing exemplar SVMs. *Computer Vision and Image Understanding*, 2015, **138**: 114−123

126 Oquab M, Bottou L, Laptev I, Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014. 1717−1724

127 Hoffman J, Gupta S, Leong J, Guadarrama S, Darrell T. Cross-modal adaptation for RGB-D detection. In: Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA). Stockholm, Sweden: IEEE, 2016. 5032−5039

128 Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Quebec, Canada: ACM, 2014. 3320−3328

129 Tang Y J, Wu B, Peng L R, Liu C S. Semi-supervised transfer learning for convolutional neural network based Chinese character recognition. In: Proceedings of the 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). Kyoto, Japan: IEEE, 2017. 441−447

130 Christodoulidis S, Anthimopoulos M, Ebner L, Christe A, Mougiakakou S. Multisource transfer learning with convolutional neural networks for lung pattern analysis. *IEEE Journal of Biomedical and Health Informatics*, 2017, **21**(1): 76−84

131 Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(7): 1409−1422

132 Hare S, Saffari A, Torr P H S. Struck: structured output tracking with kernels. In: Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 263−270

133 Milan A, Schindler K, Roth S. Multi-target tracking by discrete-continuous energy minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(10): 2054−2068

134 Bolme D S, Beveridge J R, Draper B A, Lui Y M. Visual object tracking using adaptive correlation filters. In: Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 2544−2550

135 Henriques J F, Caseiro R, Martins P, Batista J. Exploiting the circulant structure of tracking-by-detection with kernels. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: Springer, 2012. 702−715

136 Henriques J F, Caseiro R, Martins P, Batista J. High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(3): 583−596

137 Danelljan M, Khan F S, Felsberg M, Van De Weijer J. Adaptive color attributes for real-time visual tracking. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014. 1090−1097

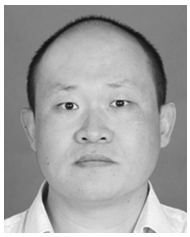
138 Wang N Y, Yeung D Y. Learning a deep compact image representation for visual tracking. In: Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: ACM, 2013. 809−817

139 Wang N Y, Li S Y, Gupta A, Yeung D Y. Transferring rich feature hierarchies for robust visual tracking. arXiv: 1501.04587v2, 2015.

140 Wang L J, Ouyang W L, Wang X G, Lu H C. Visual tracking with fully convolutional networks. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 3119−3127

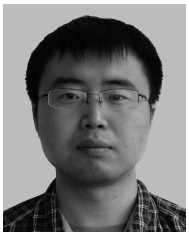
141 Nam H, Han B. Learning multi-domain convolutional neural networks for visual tracking. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016. 4293−4302

142 Cui Z, Xiao S T, Feng J S, Yan S C. Recurrently target-attending tracking. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016. 1449−1458



张雪松 博士, 大连交通大学软件学院讲师. 主要研究方向为计算机视觉, 机器学习, 室内场景理解, 物体识别与检测.
E-mail: zhangxuesong@djtu.edu.cn
(**ZHANG Xue-Song** Ph. D., lecturer at the Software Technology Institute, Dalian Jiaotong University. His research interest covers computer vision, machine learning, indoor scene understanding, object recognition and detection.)

庄 严 大连理工大学控制科学与工程学院教授. 主要研究方向为移动机器人面向复杂三维环境的自主感知、建模、地图构建、场景理解及机器学习技术在机器人领域的应用. 本文通信作者.
E-mail: zhuang@dlut.edu.cn
(**ZHUANG Yan** Professor at the School of Control Science and Engineering, Dalian University of Technology. His research interest covers mobile robot 3D environment perception, modelling, mapping, scene understanding, and machine learning in robotics. Corresponding author of this paper.)



闫 飞 大连理工大学控制科学与工程学院副教授. 主要研究方向为移动机器人地图构建、路径规划、自主导航和场景理解. E-mail: fyan@dlut.edu.cn
(**YAN Fei** Associate professor at the School of Control Science and Engineering, Dalian University of Technology. His research interest covers mobile robot map building, path planning, autonomous navigation and scene understanding.)

王 伟 大连理工大学控制科学与工程学院教授. 主要研究方向为自适应控制, 预测控制, 机器人学, 智能控制.
E-mail: wangwei@dlut.edu.cn
(**WANG Wei** Professor at the School of Control Science and Engineering, Dalian University of Technology. His research interest covers adaptive control, predictive control, robotics, and intelligent control.)