

迁移学习研究和算法综述

刘鑫鹏¹, 栾悉道^{1,*}, 谢毓湘², 黄明哲¹

(1.长沙学院计算机工程与应用数学学院, 湖南 长沙 410022; 2.国防科学技术大学系统工程学院, 湖南 长沙 410073)

摘要: 迁移学习可以将从数据域学习的知识迁移到目标数据域, 有效地解决传统机器学习中训练集和测试集经常不服从独立同分布的问题. 对迁移学习的研究价值和经典算法进行了综述, 最后对迁移学习的未来研究重点进行探讨.

关键词: 迁移学习; 领域适配; 分类; 标签

中图分类号: TP391 **文献标识码:** A **文章编号:** 1008-4681(2018)05-0028-04

我们已经进入大数据时代, 各行业对数据检索和分析能力的要求越来越高. 迁移学习是机器学习中最前沿的研究领域^[1]; 通俗来讲, 就是运用已学习到的知识来学习新知识. 斯坦福大学教授 Andrew Ng 认为: 迁移学习将成为继监督学习之后机器学习在商业领域成功应用的下一个推动力. 迁移学习在目标领域标注数据较少时可以从相关领域寻找已标注数据进行训练, 其主要目标就是将已经学会的知识很快地迁移到一个新的领域中^[2]. 相比传统的机器学习, 它的优势在于允许源域和目标域的样本、任务或者分布可以有较大的差异, 能够节省人工标注样本的时间.

本文第一部分对迁移学习进行了分类, 第二部分总结了迁移学习的研究目标与应用领域, 第三部分介绍了迁移学习经典算法, 最后是迁移学习的研究总结与展望.

1 迁移学习的分类

通常将已经学习过的领域叫做“源域”, 用 D_S 来表示; 把需要应用的新的领域叫做“目标域”^[3,4] 用 D_T 来表示. T_S 表示源任务, T_T 表示目标任务.

Pan 和 Yang^[5] 据源领域和目标领域的相似度, 将迁移学习做如下分类:

在归纳式迁移学习中, 若 D_S 包含有大量标签数据, 这时归纳式迁移学习就相当于多任务学习. 多任务学习是指知识在不同任务之间的传递, 就是将源

任务中学习到的知识迁移到目标任务中去. 若 D_S 没有可用的标签数据, 这时归纳式迁移学习就相当于自我学习. 比如区分汽车和自行车的图像, 未标注数据完全来自于一个和已标注数据不同的分布, 这种情形被称为自我学习.

表1 迁移学习的分类

学习方式	$D_S \& D_T$	$T_S \& T_T$
传统机器学习	相同	相同
归纳式迁移学习 ^[6]	相同/相关	相关
迁移学习 无监督迁移学习 ^[7,8]	相关	相关
直推式迁移学习 ^[9]	相关	相同

2 迁移学习的研究目标与应用领域

2.1 研究目标

迁移学习主要解决的是以下两个问题:

(1) 解决小数据问题. 传统机器学习存在一个严重弊端: 假设训练数据与测试数据服从相同的数据分布(但许多情况并不满足这种假设, 通常需要众包来重新标注大量数据以满足训练要求, 有时还会造成数据的浪费). 当训练数据过少时, 经典监督学习会出现严重过拟合问题, 而迁移学习可从源域的小数据中抽取并迁移知识, 用来完成新的学习任务.

(2) 解决个性化问题. 当需要专注于某个目标领域时, 源领域范围太广却不够具体. 例如专注于农作物识别时, 源领域 ImageNet 太广而不适用, 利用迁移学习可以将 ImageNet 上的预训练模型特征迁移到目标域, 实现个性化.

收稿日期: 2018-07-12

基金项目: 国家自然科学基金项目(批准号: 61571453); 湖南省教育科学规划课题(批准号: XJK17BXX010).

作者简介: 刘鑫鹏(1998—), 男, 湖南株洲人, 长沙学院计算机工程与应用数学学院学生. 研究方向: 深度学习.

* 通讯作者: 栾悉道(1976—), 男, 山东即墨人, 长沙学院计算机工程与应用数学学院副教授, 博士. 研究方向: 深度学习与模式识别.

2.2 应用领域

(1) 自然语言处理: 迁移学习应用于自然语言处理的原因是自然语言领域标注和内容数据稀缺. 可以利用源域(例如英语)中标注的样本集来对目标域(例如法语)中的样本进行处理. 迁移学习能够从长文本中迁移标注和内容知识, 帮助处理短文本语言的分析与处理.

(2) 计算机视觉: 由于图像中可能存在可变的光照、朝向等条件, 导致标注数据与未标注数据具有不同的数据属性和统计分布, 用传统机器学习显然无法满足要求. 迁移学习算法能够将领域适配, 进而达到训练效果, 提升准确率.

(3) 医疗健康和生物信息学^[10]: 在医学影像分析领域, 医学图像训练数据的标注需要先验的医学知识, 适合标注此类数据的人群稀少, 从而导致训练数据严重稀缺, 深度学习将不再适用. 可以将迁移学习应用到医学图像的语义映射中, 利用图像识别的结果帮助医生对患者进行诊断, 从而减轻医生的工作负担, 促进医疗实现转型. 例如: 胸部 X 光片的图像通常有助于检测结核、肺炎、心脏衰竭、肺癌和结节病等.

(4) 从模拟中学习: 从模拟中学习是一个风险较小的方式, 目前被用来实现很多机器学习系统. 源数据域和目标域的特征空间是一样的, 但是模拟和现实世界的边缘概率分布是不一样的, 即模拟和目标域中的物体看上去是不同的. 模拟环境和现实世界的条件概率分布可能是不一样的, 不会完全模仿现实世界中的物体交互. Udacity 已经开源了它用来无人驾驶汽车工程教学的模拟器^[11]. OpenAI 的 Universe 平台将可能允许用其他视频游戏来训练无人驾驶汽车. 另一个必须从模拟中学习的领域是机器人. 在实际的机器人上训练模型是非常缓慢和昂贵的, 训练机械臂就是一个典型案例. 从模拟中学习并且将知识迁移到现实世界的方式能缓解这个问题.

(5) 用户评价: 例如在评价用于对某服装品牌的情感分类任务中, 我们无法收集到非常全面的用户评价的数据. 因此当我们直接通过之前训练好的模型进行情感识别时, 效果必然会受到影响. 迁移学习可以将少量与测试数据相似的数据作为训练集进行训练, 能达到较好的分类效果, 并且节省大量的时间和精力.

(6) 推荐系统^[12]: 在源领域训练好一个推荐系统, 然后应用在稀疏的、新的目标领域. 比如已经成熟

完善的电影推荐系统就可以应用在书籍推荐系统中.

(7) 个性化对话^[12]: 训练一个通用型的对话系统, 该系统可能是闲聊型, 也可能是任务型的. 我们可以利用迁移学习训练特定领域的小数据集, 使这个对话系统适应不同任务.

3 迁移学习算法

3.1 归纳式迁移学习

Dai 等人^[13]基于 Boosting, 提出了基于实例的 TrAdaBoost 迁移学习算法. 当目标域中的样本被错误地分类之后, 可以认为这个样本是很难分类的, 因此增大这个样本的权重, 在下一次的训练中这个样本所占的比重变大. 如果源域中的一个样本被错误地分类了, 可以认为这个样本对于目标数据是不同的, 因此降低这个样本的权重, 降低这个样本在分类器中所占的比重. 李小璇^[14]基于 TrAdaBoost 整合两个分类器 SVM 和 NB 提出 TrSN 算法. 在 20 个 news group 数据集上的分类精度平均达到 0.94.

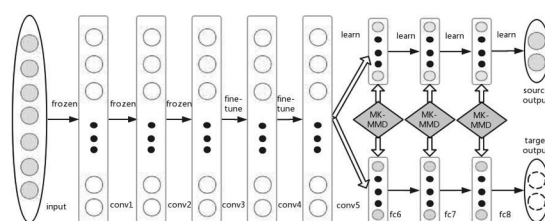


图 1 深度适配网络 (DAN) 结构图

为了解决迁移学习中的领域适配问题, Long Mingsheng 基于 DDC (Deep Domain Confusion, 深度领域适配) 提出深度适配网络 DAN (Deep Adaptation Network)^[15]. DDC 是在预训练的 AlexNet 网络的第 7 层加入了 MMD^[16] 距离来减小 D_s 和 D_t 之间的差异. DAN 是深度迁移学习方法, 它适配高层网络 DDC, 并加入了多核的 MMD (MK-MMD). MMD 是把源域和目标域用一个相同的映射, 映射到一个再生核希尔伯特空间 (RKHS) 中, 然后求映射后两部分数据的均值差异, 就当作是两部分数据的差异. 在 MMD 中核是固定的, 可以选择是高斯核或者线性核. MK-MMD 提出用多个核去构造这个总的核, 这样效果比一个核更好. 它很好地解决了 DDC 的两个问题: 一是 DDC 只适配了一层网络, 而 DAN 适配最后三层 (6~8 层) 如图 1, 网络的迁移能力在最后三层开始就会变得专化 (specific), 所以要重点适配这三层. 二是 DDC 是用了单一核的 MMD, 单一固定的核可能不是最优的核. DAN 用了多核的 MMD (MK-MMD),

效果比 DDC 更好。

Cao 等人提出了一种更一般的迁移学习“部分迁移学习”(Partial Transfer Learning)^[17],就是只迁移源域中和目标域相关的那部分样本.通过 SAN 来处理部分迁移问题.对抗网络可以很好地学习领域不变特征^[18],从而在迁移学习中能发挥很大作用。

3.2 无监督迁移学习

Pan 等人^[19]提出了迁移成分分析 TCA. TCA 通过降维来减少数据维度,首先输入两个特征矩阵,计算 L 和 H 矩阵,然后选择常用的核函数进行映射(比如线性核、高斯核)计算 K,接着求 KHK 的前 m 个特征值.然后得到源域和目标域的降维后的数据,最后就可以使用传统机器学习方法. TCA 实现简单,没有太多的限制.但是尽管它绕开了半定规划问题的求解,却需要花费很多计算时间在大矩阵伪逆的求解以及特征值分解。

Gong 等人在 2012 年基于 SGF(Sample Geodesic Flow, 采样测地线流方法)提出了 GFK^[20-21]. SGF 把源域和目标域分别看成高维空间中的两个点,在这两个点的测地线上取 n 个中间点,依次连接起来.然后由源域和目标域就构成了一条测地线的路径.找到每一步的变换,就能从源域变换到目标域. GFK 是子空间变换方面最为经典的迁移学习方法,是为了解决迁移学习中的无监督领域适配问题.它通过一个特征映射把源域和目标域变换到一个距离最小(相似度最高)的公共空间上. GFK 方法的实施步骤为:选择最优的子空间维度进行变换,构建测地线,计算测地线流式核,构建分类器。

Tahmoresnezhad 等人^[22]在 2017 年提出视觉领域自适应(Visual Domain Adaptation, VDA). VDA 利用联合转移学习和领域适应来处理分布差异较大的转换问题.特别是视觉数据集,以无监督的方式在测试集中没有可用标签的情况下减少跨域的联合边缘和条件分布.此外, VDA 构造了嵌入表示中的凝聚域不变集群,以将各个域与类转移分开,使用细化的伪目标标签来迭代收敛至最终解决方案.采用迭代过程以及新颖的优化问题为跨领域的适应创建一个稳健而有效的表示。

3.3 直推式迁移学习

联合分布适配方法(Joint Distribution Adaptation, JDA)^[23]是为了解决迁移学习中的领域适配问题,是用源数据域来标定目标数据域. JDA 假设 D_S

和 D_T 边缘分布不同, D_S 和 D_T 条件分布不同,适配联合概率恰好能解决这个问题.具体步骤是:首先用 TCA 来适配边缘分布,用 MMD 适配 D_S 和 D_T 的条件概率分布.然后通过弱分类器迭代,将上一轮得到的标签作伪标签,迭代多次以达到更高的精度. JDA 与 TCA 的区别有两点: TCA 是无监督的,即边缘分布适配不需要标签, JDA 需要源域有标签; TCA 不需要迭代, JDA 需要迭代。

Busto 和 Gall^[24]提出了开放集迁移学习(Open Set Domain Adaptation).他们利用源域和目标域的关系,给目标域的样本打上标签,并将源域转换到和目标域同一个空间中,让学习标签和学习映射进行交替,直到收敛或者目标值小于某一值即可。

3.4 其他算法

现有的域适配方法主要针对向量,这种表示所带来的问题是,当把这些数据应用于高维度表示(如卷积)时,数据首先要经过向量化.此时,无法精准完备地保留一些统计属性或者重要结构. Lu 等人^[25]基于张量的 Tucker 分解,提出了一个称为 Native Tensor Subspace Learning 的迁移学习算法.这个算法的出发点,是假设源域和目标域共享了一部分子空间,而这只在它们差异非常小时才有效.在更一般的条件下,这种共享变量要通过一个线性变换来实现. Fernando 等人提出了一个加强版的算法—Tensor - Aligned Invariant Subspace Learning (TAISL).这个算法是对 ICCV-13 的那个子空间校准的扩展版^[26]。

Feuz 和 Cook 等人^[27]提出了新颖的异构传输学习技术,特征空间重映射(FSR),它在具有不同特征空间的域之间传输知识,构建元特征来将不同特征空间中的特征相关联.这些技术利用多个源数据集来构建进一步提高性能的集成学习器, FSR 应用于活动识别问题和文档分类问题.集合技术能够胜过所有其他基线,甚至比在目标域中使用大量标记数据进行训练的分类器执行得更好。

4 总结与展望

本文对迁移学习算法以及研究进展状况进行了综述.迁移学习适合处理解决小数据和个性化问题,在计算机视觉、自然语言处理、医疗健康、模拟学习、用户评论、推荐系统和个性化对话等领域展现出其通用性.迁移学习有许多经典算法,这些算法一步步将机器学习和迁移学习的问题优化,从降维到域适

配,从完全迁移到部分迁移,从向量迁移学习到张量迁移学习,无不展现着迁移学习研究得到充分发展。然而,迁移学习还有许多具有挑战性问题有待进一步研究,具体可概括为这几个方面:(1) 负迁移表示不仅没有提升模型能力,反而降低了识别率。(2) 在非简单分类或者回归的问题上,如何更好地优化迁移算法。(3) 跨领域迁移学习。(4) 偏数据的处理,面临的问题是数据收集的时候和下一个分布是不一样的,难点在于如何利用迁移学习将偏差处理好。(5) 迁移学习与多种深度学习方法相结合的图像描述方法。

参考文献:

- [1] 解晓康. 基于迁移学习的模糊图像识别技术研究[D]. 武汉: 华中科技大学硕士学位论文, 2016.
- [2] Patel V M, Gopalan R, Li R, et al. Visual domain adaptation: A survey of recent advances [J]. Signal Processing Magazine, IEEE, 2015 (3): 53–69.
- [3] Fernando B, Habrard A, Sebban M, et al. Unsupervised visual domain adaptation using subspace alignment [A]. Proceedings of IEEE International Conference on Computer Vision [C]. IEEE, 2013: 2960–2967.
- [4] Fernando B, Habrard A, Sebban M, et al. Subspace Alignment for Domain Adaptation [EB/OL]. <https://arxiv.org/pdf/1409.5241v2.pdf> 2014.
- [5] Pan J, Yang Q. A survey on transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010 (10): 1345–1359.
- [6] Wei F M, Zhang J P, Chu Y, et al. FSFP: Transfer learning from long texts to the short [J]. Applied Mathematics & Information Sciences, 2014 (4): 2033–2044.
- [7] Dai W Y, Yang Q, Xue G R, et al. Self-taught clustering [A]. Proceedings of the 24th International Conference on Machine Learning [C]. San Francisco: Morgan Kaufmann Publishers, 2008: 200–207.
- [8] Samanta S, Tirumalai S A, Das S. Cross-Domain clustering performed by transfer of knowledge across domains [A]. Proceedings of the 2013 IEEE 4th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics [C]. IEEE, 2013: 1–4.
- [9] Dai W Y, Xue G R, Yang Q, et al. Co-Clustering based classification for out-of-domain documents [A]. Proceedings of the 13th ACM International Conference on Knowledge Discovery and Data Mining [C]. New York: ACM Press, 2007: 210–219.
- [10] 龙明盛. 迁移学习问题与方法研究[D]. 北京: 清华大学博士学位论文, 2014.
- [11] Sebastian R. Transfer Learning—Machine Learning's Next Frontier [EB/OL]. <http://sebastianruder.com/transfer-learning/index.html> 2017-04-10.
- [12] 杨强. 迁移学习的本质与实际应用[R]. KDD China 技术峰会演讲, 2016.
- [13] Dai W Y, Yang Q, Xue G R, et al. Boosting for transfer learning [A]. Proceedings of the 24th International Conference on Machine Learning [C]. San Francisco: Morgan Kaufmann Publishers, 2007: 193–200.
- [14] 李小璇. 基于 TrAdaBoost 的直推式迁移学习文本分类技术改进研究[D]. 武汉: 华中科技大学硕士学位论文, 2016.
- [15] Long M, Cao Y, Wang J, et al. Learning transferable features with deep adaptation networks [A]. Proceedings of the 32nd International Conference on International Conference on Machine Learning [C]. 2015: 97–105.
- [16] Gretton A, Borgwardt K M, Rasch M J, et al. A kernel two-sample test [J]. Journal of Machine Learning Research, 2012 (13): 723–773.
- [17] Cao Z, Long M, Wang J, et al. Partial Transfer Learning with Selective Adversarial Networks [EB/OL]. <http://cn.arxiv.org/pdf/1707.07901v1> 2017.
- [18] Pei Z, Cao Z, Long M, et al. Multi-Adversarial Domain Adaptation [EB/OL]. <http://ise.thss.tsinghua.edu.cn/~mlong/doc/multi-adversarial-domain-adaptation-aaai18.pdf> 2018.
- [19] Pan S J, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis [J]. IEEE Transactions on Neural Networks, 2010 (2): 199–210.
- [20] Gong B. Geodesic flow kernel for unsupervised domain adaptation [A]. Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition [C]. IEEE, 2012.
- [21] Gopalan R, Li R, Chellappa R. Domain adaptation for object recognition: An unsupervised approach [A]. Proceedings of 2011 IEEE International Conference on Computer Vision [C]. IEEE, 2011.
- [22] Tahmoresnezhad J, Hashemi S. Visual domain adaptation via transfer feature learning [J]. Knowledge & Information Systems, 2016, (2): 1–21.
- [23] Long M, Wang J, Ding G, et al. Transfer feature learning with joint distribution adaptation [A]. Proceedings of the IEEE International Conference on Computer Vision [C]. Venice: IEEE, 2013: 2200–2207.
- [24] Busto P P, Gall J. Open set domain adaptation [A]. Proceedings of 2017 IEEE International Conference on Computer Vision [C]. Venice: IEEE, 2017: 754–763.
- [25] Lu H, Zhang L, Cao Z, et al. When unsupervised domain adaptation meets tensor representations [A]. Proceedings of 2017 IEEE International Conference on Computer Vision [C]. Venice: IEEE, 2017: 1–16.
- [26] Fernando B, Habrard A, Sebban M, et al. Unsupervised visual domain adaptation using subspace alignment [A]. Proceedings of the IEEE International Conference on Computer Vision [C]. 2013: 2960–2967.
- [27] Feuz K D, Cook D J. Transfer learning across feature-rich heterogeneous feature spaces via feature-space remapping (FSR) [J]. ACM Transactions on Intelligent Systems and Technology, 2015, (1): 3.

(下转第 36 页)

- [18] Cevallos F, Zhao F. Minimizing transfer times in public transit networks with genetic algorithm [J]. Transportation Research Record, 2006 (1): 74-79.
- [19] Shafahi Y, Khani A. A practical model for transfer optimization in a transit network: Model formulations and solutions [J]. Transportation Research Part A, 2010 (6): 377-389.
- [20] Fujita M, Mori T. Structural stability and evolution of urban systems [J]. Regional Science and Urban Economics, 1997, (5): 399-442.
- [21] Dell'Olio L, Ibeas A, Ruisánchez F. Optimizing bus-size and headway in transit networks [J]. Transportation, 2012, (2): 449-464.

Optimization of Low Carbon Public Transportation Networks Based on Multi-objective Genetic Algorithm

TAN Qian, CHANG Zhugang

(Changsha Planning & Design Institute Co. Ltd., Changsha Hunan 410007, China)

Abstract: The multi-objective model of bi-level bus network was constructed with three objectives: the passenger travel time, bus companies operating costs and environmental pollution, and several constraint conditions: the bus line length, line nonlinear coefficient, direct ratio, line load efficiency coefficient and transfer times. Buses were classified into pure gasoline/diesel buses and hybrid buses in this model. A hybrid genetic algorithm is used to solve this model, taking the model of stochastic user equilibrium as the assignment for bus passenger distribution. MATLAB programming was ultimately applied to calculate and check the model and algorithm with numerical examples. The results show the feasibility and effectiveness of the model.

Key Words: bi-level programming; multi-objective genetic algorithm; low carbon; optimization of public transportation networks

(责任编辑: 晴川)

(上接第 31 页)

Transfer Learning Research and Algorithm Review

LIU Xinpeng¹, LUAN Xidao^{1,*}, XIE Yuxiang², HUANG Mingzhe¹

(1. College of Computer Engineering and Applied Mathematics, Changsha University,

Changsha Hunan 410022, China;

2. College of System Engineering, National University of Defense Technology, Changsha Hunan 410073, China)

Abstract: In recent years, transfer learning has been applied for many areas. Transfer learning is the transfer of knowledge learned from a data domain to a target data domain. It effectively solves the main problems in traditional machine learning: The training set and the test set often do not obey independent and identically distributed. This article summarizes the research value of transfer learning and classical algorithms, and finally points out the future progress of transfer learning.

Key Words: transfer learning; domain adaptation; classification; label

(作者本人校对)