

基于多原型重放和对齐的类增量无源域适应

田青, 康陆禄, 周亮宇

引用本文

田青, 康陆禄, 周亮宇. [基于多原型重放和对齐的类增量无源域适应](#)[J]. 计算机科学, 2025, 52(3): 206-213.

TIAN Qing, KANG Lulu, ZHOU Liangyu. [Class-incremental Source-free Domain Adaptation Based on Multi-prototype Replay and Alignment](#) [J]. Computer Science, 2025, 52(3): 206-213.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[辅助判决的案情要素关联与证据提取](#)

Case Element Association with Evidence Extraction for Adjudication Assistance

计算机科学, 2025, 52(2): 222-230. <https://doi.org/10.11896/jsjcx.240600081>

[基于源模型贡献量化的多无源域适应](#)

Multi-source-free Domain Adaptation Based on Source Model Contribution Quantization

计算机科学, 2025, 52(2): 116-124. <https://doi.org/10.11896/jsjcx.240600004>

[基于个性化PageRank和对比学习的图异常检测模型](#)

Graph Anomaly Detection Model Based on Personalized PageRank and Contrastive Learning

计算机科学, 2025, 52(2): 80-90. <https://doi.org/10.11896/jsjcx.240200005>

[融合时频特征的多粒度时间序列对比学习方法](#)

Multi-granularity Time Series Contrastive Learning Method Incorporating Time-Frequency Features

计算机科学, 2025, 52(1): 170-182. <https://doi.org/10.11896/jsjcx.231100171>

[跨城市人类移动行为预测研究综述](#)

Survey on Cross-city Human Mobility Prediction

计算机科学, 2025, 52(1): 102-119. <https://doi.org/10.11896/jsjcx.240100032>

基于多原型重放和对齐的类增量无源域适应

田青^{1,2,3} 康陆禄¹ 周亮宇¹

1 南京信息工程大学软件学院 南京 210044

2 南京信息工程大学无锡研究院 江苏 无锡 214000

3 南京大学计算机软件新技术国家重点实验室 南京 210023

摘要 传统无源域适应通常假设目标域数据全部可用,然而在实际应用中目标域数据常以流的形式出现,即未标记的目标域中的类会依次增加,这无疑带来了新的挑战。首先,在每个时间步骤中,目标域的标签空间都是源域的一个子集,盲目对齐反而会导致模型性能下降;其次,在学习新类的过程中会破坏先前学习到的知识,导致之前知识的灾难性遗忘。为了解决这些问题,提出了一种基于多原型重放和对齐(MPRA)的方法。该方法通过累积预测概率检测目标域中的共享类来应对标签空间不一致问题,并采用多原型重放来处理灾难性遗忘,提高模型的记忆能力。同时,基于多原型和源模型权重进行跨域的对比学习,从而对齐特征分布,提高模型性能。大量的实验表明,所提方法在3个基准数据集上都取得了优越的表现。

关键词: 无源域适应; 类增量学习; 多原型; 对比学习; 迁移学习

中图分类号 TP391

Class-incremental Source-free Domain Adaptation Based on Multi-prototype Replay and Alignment

TIAN Qing^{1,2,3}, KANG Lulu¹ and ZHOU Liangyu¹

1 School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China

2 Wuxi Institute of Technology, Nanjing University of Information Science and Technology, Wuxi, Jiangsu 214000, China

3 State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China

Abstract Traditional source-free domain adaptation usually assumes that all the target domain data is available, but in practice, the target domain data often appears in the form of streams, that is, the classes in the unlabeled target domain will increase sequentially, which undoubtedly brings new challenges. First, in each time step, the label space of the target domain is a subset of the source domain, and blind alignment will cause the performance of the model to deteriorate. Secondly, in the process of learning new classes, it will destroy the previously learned knowledge, resulting in the catastrophic forgetting of the previous knowledge. In order to solve these problems, this paper proposes a method based on multi-prototype replay and alignment (MPRA). In this method, the shared classes in the target domain are detected by cumulative prediction probabilities, the label space inconsistency problem is solved, and the multi-prototype replay is used to deal with catastrophic forgetting and improve the memory ability of the model. Additionally, the method incorporates cross-domain contrastive learning based on multi-prototype and source model weights to align feature distributions and improve model robustness. A large number of experiments show that the proposed method has achieved superior performance on 3 benchmark datasets.

Keywords Source-free domain adaptation, Class-incremental learning, Multi-prototype, Contrastive learning, Transfer learning

1 引言

尽管深度学习在计算机视觉领域取得了显著的成就^[1-3],但其对数据变化的敏感性带来了一系列挑战。一个突出的问题是,当模型在经过训练的源域数据上表现出色时,一旦

应用到不同的目标域,由于数据分布的变化,性能会急剧下降^[4-7]。为了解决这一问题,研究者们提出了无监督域适应(Unsupervised Domain Adaptation, UDA)的概念^[8-10],其核心目标在于将从源域数据中获得的知识泛化至目标域^[11-12]。它利用标记的源域数据和未标记的目标域数据进行学习,

到稿日期:2024-01-22 返修日期:2024-04-12

基金项目:国家自然科学基金(62176128);江苏省自然科学基金(BK20231143);南京大学计算机软件新技术国家重点实验室开放课题(KFKT2022B06);中央高校基本科研基金(NJ2022028);江苏省“青蓝工程”人才计划项目

This work was supported by the National Natural Science Foundation of China(62176128), Natural Science Foundation of Jiangsu Province, China (BK20231143), Open Projects Program of State Key Laboratory for Novel Software Technology of Nanjing University(KFKT2022B06), Fundamental Research Funds for the Central Universities(NJ2022028) and Qing Lan Project of Jiangsu Province.

通信作者:田青(tianqing@nuist.edu.cn)

从而实现了知识从源域向目标域的传递。然而,随着对数据安全和隐私保护的增强,原始的源域数据将不再可用,传统的无监督域适应方法因此不再有效。为了应对这一挑战,近年来,研究者们提出了无源域适应(Source-Free Domain Adaptation, SFDA)^[13-14],其迅速成为域适应领域的研究热点。无源域适应的关键在于,在不接触源域数据的情况下,其依然能够有效地适应到目标域。

目前,大多数无源域适应方法^[15-18]通常基于一个先验假设,即所有目标域上的数据是全部可用的。然而在实际应用中,特别是在复杂的现实环境中,目标数据往往以流的形式逐步出现^[19-20]。以图1为例,倘若考虑将一个在美术图像上训练良好的源模型迁移到现实世界的图像分类任务中,在这一情景下,目标域中现实世界的图像根据获取的难易程度逐步呈现,例如,电话、自行车等在日常生活中相对常见,因此容易获取,而野生动物等则需要花费较长的时间来逐步收集。

在这种情况下,为了提高效率,更为合适的方式是随着时间的推移不断调整源模型,而不是长时间等待直到收集完所有目标域图像。值得注意的是,为了节约存储成本,在后续的训练中,前期样本将不再可用。然而,这也带来了新的问题,因为直接使用新类别的数据来训练模型不可避免地会破坏已学习类别知识,从而导致整体性能急剧下降^[21-22]。由于目前的无源域适应方法并未考虑到此问题,因此本文进一步探索了一种更加实用的场景,称为类增量无源域适应(Class-Incremental Source-Free Domain Adaptation, CI-SFDA)。

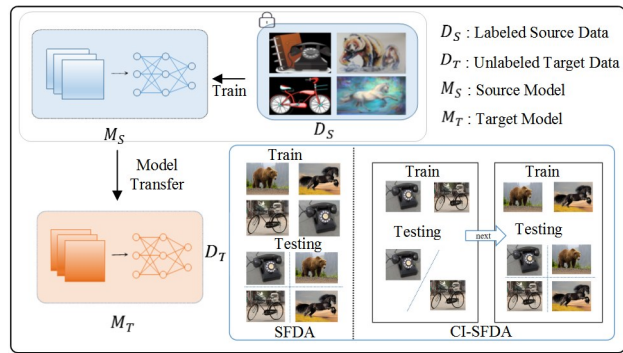


图1 无源域适应与类增量无源域适应对比

Fig.1 Comparison between source-free domain adaptation and class-incremental source-free domain adaptation

相较于传统的无源域适应设置,类增量无源域适应更加复杂且具有更高的挑战性,除了要面对无源域适应中常见域偏移和源域数据不可用^[13]等问题外,还需应对从流式数据中不断学习所带来的新挑战。在这种情况下,直接使用新类样本对网络进行调整会导致先前学习到的知识被抹去,并导致不可逆的性能下降。此外,类增量的引入使得目标域的标签空间在每个时间步骤中只占源域空间的一部分,这使得在进行域对齐之前需要先处理标签空间不一致的问题。因此,类增量无源域适应不仅需要解决传统无源域适应中的问题,还需要应对动态学习环境中的灾难性遗忘和标签空间不一致等更为复杂的挑战。这使得算法的设计变得更加关键,该设计需确保模型在不断变化的数据流中能够保持性能稳定。

为了更有效地解决这些问题,本文提出了一种基于多原型

重放和对齐(Multi-Prototype Replay and Alignment, MPRA)的方法。具体而言,本文首先通过计算目标域在当前时间步骤中所有样本的累积预测概率,来检测共享类,从而应对每个时间步骤中标签空间不一致的问题。此外,本文发现,即使在同一类别中,不同样本之间依然存在显著差异。为了最大程度地保留旧知识,缓解灾难性遗忘问题,受类增量学习中的数据重放方法^[23-25]的启发,提出了多原型重放方法以应对灾难性遗忘问题。该方法通过存储更多多样性的图像原型,在尽可能节约存储空间的前提下有效保留更多知识。这不仅有助于提高模型的记忆能力,而且在存储开销上取得了更为合理的平衡。同时,本文利用源模型提取分类器权重作为源域类中心,并基于原型通过跨域对比学习来学习域不变特征进行对齐。

综上所述,本文的贡献如下:

1)研究了一个新的实际应用中的复杂问题,探索了更加现实的类增量无源域适应场景,并充分分析了其存在的困难和挑战。

2)提出了一种基于多原型重放和对齐的类增量无源域适应方法(MPRA)。该方法通过对每个类中最具有代表性的样本进行重放和对齐来减小灾难性遗忘和域偏移的影响。

3)通过广泛的实验对所提方法进行多方面的评估和分析,验证了其在解决类增量无源域适应问题上的有效性和优越性。

本文第2章回顾了目前相关领域的工作;第3章详细介绍了本文提出的MPRA方法;第4章给出了各种实验数据相关对比和分析;最后总结全文并展望未来。

2 相关工作

2.1 无源域适应

近年来,越来越多的方法被提出用来解决无源域适应问题。具体来说,SHOT(Source Hypothesis Transfer)^[13]利用一种基于质心的标签细化技术来指导模型的自训练。GSFDA(Generalized Source-Free Domain Adaptation)^[26]和NRC(Neighborhood Reciprocity Clustering)^[27]探索了近邻结构信息,通过近邻聚类的方式提高模型性能。DIPE(Domain-Invariant Parameter Exploring)^[28]则是从模型参数的角度出发,专注于识别出源模型中的域不变参数。还有一些方法^[29]通过源模型隐藏的信息估计出源域数据的底层分布,从而实现域的对齐。最近,AdaContrast(Contrastive Test-Time Adaptation)^[30],CDCL(Cross-Domain Contrastive Learning)^[31]等方法将对对比学习的方法引入无源域适应中,通过构建正负对进行特征层面的拉远和拉近,从而提高模型的泛化性;而SFUDA-TPS(SFUDA with Trusted Pseudo Samples)^[32]则是从标签噪声角度考虑通过对样本进行划分来提高伪标签可信度。但是目前的无源域适应方法都没有考虑到更加现实的目标域数据以流的形式出现的场景,因此无法直接使用。

2.2 部分域适应

一般的域适应方法通常假设目标域与源域有着相同的标签空间,但部分域适应则假设目标域的标签空间是源域的一个子空间^[33]。因此,部分域适应与类增量域适应有着更

密切的联系,类增量域适应的每一个时间步骤都可以被看作是一个无源的部分域适应过程。目前的部分域适应方法大多通过为源域样本分配不同的权重来减轻源私有类造成的负迁移^[34],其中 PADA (Partial Adversarial Domain Adaptation)^[35]更是作为一种典型方法被广泛研究和对比。SLM (Select, Label, and Mix)^[36]等方法则是通过从源域样本中筛选出相关类所对应的样本来完成部分域适应,从而消除不相关类造成的负迁移影响。

2.3 类增量域适应

类增量域适应是将类增量学习扩展至域适应并进行研究的一个新任务。它从流式数据中学习,其中类逐渐增加,每次只有新类可用,并对目前为止所有类的样本进行预测。目前类增量域适应领域的相关研究较少。类增量域适应在 CIDA (Class-Incremental Domain Adaptation)^[37]中被首次提出,目标是在无源场景下,通过少量的标记样本,以流的形式逐步学习到源域中未知的新类别。ProCA (Prototype-Guided Continual Adaptation)^[38]则是假设一个完全未标记的目标域,通过将识别的目标原型与源域质心对齐并引入类增量学习中的样本重放方法,从而为标准的类增量无监督域适应提供了第一个可行的解决方案。但是 ProCA 需要源域样本的支持,并不适用于无源场景。因此,本文进一步研究并针对类增量无源域适应进行改进,提出了 MPRA 方法。

3 基于多原型重放和对齐的类增量无源域适应方法

本章详细介绍了基于多原型重放和对齐的方法,其主要思想是筛选出每个类中最具有代表性的样本并进行重放和对齐,以缓解先前类知识的遗忘和域偏移对模型性能的影响。首先通过计算累积预测概率的方式对当前时间步骤中的标签空间进行识别,然后通过类内聚类的方式找到每个类中最具有代表性的原型并存储到重放池中,对这些原型进行对齐并在接下来的时间步骤中进行重放约束知识的遗忘。具体思路和模型架构如图 2 所示。下文将详细阐述 MPRA 各个步骤的细节。

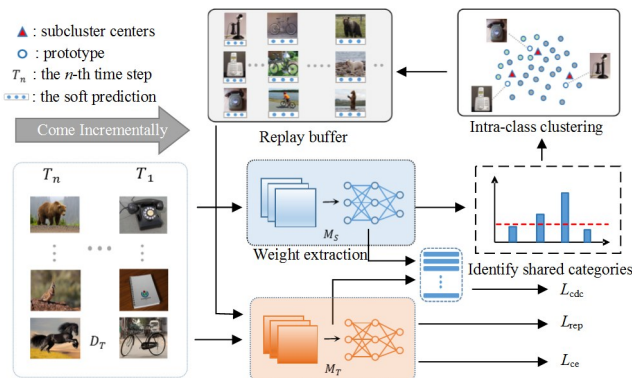


图 2 MPRA 方法的整体框架

Fig. 2 Overall architecture of MPRA

3.1 符号定义和问题描述

传统无源域适应通常假设有一个在源域数据 D_s 上训练良好的高性能源模型,表示为 M_s ,该模型包括一个特征提取器 G 和一个分类器 H 。源域和目标域具有相同的标签空间,

即 $\mathcal{Y}_t = \mathcal{Y}_s$,并且 $|\mathcal{Y}_t| = |\mathcal{Y}_s| = K$ 。然而,与传统的无源域适应任务不同的是,类增量无源域适应中目标域数据会以流的形式出现,其中的类别数量会逐渐增加。因此,本文将目标域在第 n 个时间步骤中的数据表示为 D_{T_n} ,包含 N_{T_n} 个样本。在每个时间步骤中,目标域的标签空间都只是源域标签空间的一个子集,并且在此之前的样本都不可见,即 $\mathcal{Y}_{T_n} \subset \mathcal{Y}_s, \mathcal{Y}_{T_n} \cap \mathcal{Y}_{T \neq n} = \emptyset$ 。类增量无源域适应的目标是学习一个目标模型 M_t ,在面对流动的目标域数据时依然能够保持出色的性能。为了更加详细地描述模型,表 1 列出了本文涉及的重要符号及其定义。

表 1 文中涉及的符号及其定义

Table 1 Symbols in this paper and their definitions

Symbol	Description
x_i	第 i 个目标域样本
G_s/G_t	源模型/目标模型特征提取器
H_s/H_t	源模型/目标模型分类器
P_k	第 k 个类的累积预测概率
p_i^k	第 i 个样本预测为第 k 类的概率
T_n	第 n 个时间步骤
S_{T_n}	第 n 个时间步骤时与源域的共享类
z_i	第 i 个样本的特征
\hat{y}_i	第 i 个样本的伪标签
m_k^j	第 k 个类的第 j 个子类中心
r	类原型
\mathcal{R}_{T_n}	第 n 个时间步骤时的重放池
q_i	第 i 个样本的软标签

3.2 共享类识别和伪标签生成

在类增量无源域适应中,当前时间步骤中的所有类别都是先前步骤中未曾出现的,且仅属于源域标签空间的一个子空间。有效应对不断变化的目标域的关键之一是识别当前时间步骤与源域之间共享的类别,以确保模型在源域和目标域之间建立密切联系。同时,这一过程为后续计算类别原型以构建重放池、缓解灾难性遗忘问题奠定了基础。

然而,当面临新的未标记样本时,由于目标样本未被标记,要准确检测出当前时间步骤中哪些类与源域共享是一项相对困难的任务。为解决这个问题,采用计算样本累积预测概率的方法来检测共享类^[38]。具体而言,如图 2 右侧虚线框内所示,当使用源模型对目标域样本进行预测时,当前时间步骤中所有样本属于共享类别的累积预测概率要明显大于属于源私有类别的累积预测概率,可以通过设定一个阈值对共享类进行识别。因此,首先需要计算当前时间步骤中所有目标域样本在源模型下的累积预测概率,即:

$$P_k = \sum_{i=1}^{N_{T_n}} \delta_k (H_s(G_s(x_i))) \quad (1)$$

$$P_k = \frac{P_k - \min(\mathbf{P})}{\max(\mathbf{P}) - \min(\mathbf{P})}$$

为了统一进行比较,进一步对累积预测概率 P_k 进行了 min-max 归一化。其中, δ_k 代表经过 softmax 后第 k 个类的预测概率, \mathbf{P} 表示所有 K 个类的预测概率向量。当计算完累计概率后,由于共享类的累积预测概率应该明显大于源私有类的累积预测概率,因此可以通过设定一个阈值来识别当前 T_n 时刻的共享类。具体定义如式(2)所示:

$$S_{T_n} = \{k | P_k > \text{thre}, \forall k \in K\} \quad (2)$$

其中, $thre$ 是阈值, 本文设定为 0.2。在识别出共享类之后, 就可以为当前时间步骤中的样本生成伪标签。本文借鉴文献 [13] 的伪标签生成策略, 并考虑类增量的特殊性, 加以改进。通过仅计算每个共享类的类中心的方式, 进一步剔除错误的伪标签, 提高伪标签的准确性。

$$c_k = \frac{\sum_{i=1}^{N_{T_n}} p_i^k z_i}{N_{T_n}}, k \in S_{T_n} \quad (3)$$

在类增量无源域适应中, 经过先前时间步骤中样本训练过的目标域模型对当前样本的预测是非常不准确的, 因此, 本文采用类似 warm-up 的训练方式, 先由源模型生成伪标签对目标域模型进行训练, 经过一定时间后, 再通过目标模型自己生成的伪标签进行自训练。式(3)中 p_i^k 和 z_i 的定义如式(4)所示:

$$\begin{cases} p_i^k = \delta_k(H_s(G_s(x_i))), & z_i = G_s(x_i), e < t \\ p_i^k = \delta_k(H_t(G_t(x_i))), & z_i = G_t(x_i), e > t \end{cases} \quad (4)$$

其中, e 表示当前的迭代次数, t 表示一个固定的迭代时刻。在获得共享类的中心后, 通过计算最近的类中心为第 i 个目标域样本生成伪标签。

$$\hat{y}_i = \arg \min_k Dis(\mathbf{z}_i, \mathbf{c}_k) \quad (5)$$

其中, $Dis(\cdot, \cdot)$ 表示两个向量之间的距离。为了提高伪标签的鲁棒性, 对类中心重新进行计算。 $\mathbf{c}_k = \sum_{i=1}^{N_{T_n}} \mathbb{I}(\hat{y}_i = k) \mathbf{z}_i / \sum_{i=1}^{N_{T_n}} \mathbb{I}(\hat{y}_i = k)$, 其中 $\mathbb{I}(\cdot)$ 是指示函数。得到新的类中心后, 使用式(5)对伪标签进行更新。基于伪标签, 最终可以得到标准的交叉熵损失, 即:

$$L_{ce} = -\frac{1}{N_{T_n}} \sum_{i=1}^{N_{T_n}} (\hat{y}_i)_{\text{One-Hot}} \log \delta(H_t(G_t(x_i))) \quad (6)$$

3.3 重放池的构建和更新

1) 构建

基于识别出的共享类和为每个样本生成的伪标签, 希望能够建立一个样本重放池, 以尽可能保留先前学习到的分类特征, 从而缓解灾难性遗忘问题。但是, 过量的存储是毫无意义的, 本文更希望能够在存储空间和模型记忆能力之间寻求更合理的平衡, 通过最小的样本量来存储最全面的类知识。观察图 2 可以发现, 相同的电话类别中可能包含多种款式, 不同款式之间具有显著的特征差异, 因此可以采用在类内进行聚类的方法生成多个类中心, 并以此找到每个共享类别最具代表性的原型, 在降低存储成本的同时确保涵盖各种特征差异。具体定义如式(7)所示:

$$\{\mathbf{m}_k^j\}_{j=1}^L = Kmeans(G_t(x_i))_{x_i \in D_{T_n}^k} \quad (7)$$

$$D_{T_n}^k = \{x_i | \hat{y}_i = k, \forall x_i \in D_{T_n}\}$$

其中, \mathbf{m}_k^j 表示第 k 个类内的第 j 个子类中心, L 表示子类中心的数量。然后, 计算得出与这些子类中心最相似的样本, 并将它们保存在重放池中。

$$r_k^j = \arg \min_{x_i \in D_{T_n}^k} \|G_t(x_i) - \mathbf{m}_k^j\|_2 \quad (8)$$

其中, r_k^j 代表与第 k 类的第 j 个子类中心相对应的原型。

有了这些代表性的类原型, 就可以通过对这些类原型的重放, 强制模型保留从之前类别学到的知识。具体的重放

损失定义如式(9)所示:

$$L_{rep} = -\frac{1}{|\mathcal{R}_{T_n}|} \sum_{i=1}^{|\mathcal{R}_{T_n}|} \mathbf{q}_i \log \delta(H_t(G_t(r_i))) \quad (9)$$

其中, $|\mathcal{R}_{T_n}|$ 代表当前重放池中的原型数量, \mathbf{q}_i 表示与第 i 个原型对应的软标签。

2) 更新

在识别共享类别的过程中, 可能会出现错误地识别共享类别的情况。这些被错误识别的类别可能会干扰伪标签的生成, 并在样本重放期间导致错误的特征记忆。因此, 在随后正确类别到达时, 有必要更新回放缓冲区中被错误判断的类别的原型。

为了更新这些被错误识别的类别, 一个判断依据是, 当正确的类别到达时, 这些类别的累积概率应高于在较早时期被错误判断的类别的累积概率。因此, 当在当前时间步骤计算的第 k 类的累积概率(表示为 P_k)超过重放池中先前时间步骤中获得的第 k 类的累积概率(表示为 U_k)时, 通过式(7)和式(8)对重放池中的第 k 类原型进行更新。

3.4 跨域对比对齐

尽管通过软标签进行样本重放的方式能够使模型在一定程度上保留先前样本的知识, 同时学习同类样本之间的差异以及不同类之间的相关性, 但是由于域偏移的存在, 在训练时生成的伪标签和软标签存在一定的噪声影响。因此, 本文基于多原型开发了一种新颖的跨域对比对齐方法。该方法从特征的角度进行对齐, 使模型能够在推断时做出更准确的预测。同时, 其在类原型对齐的过程中限制了特征空间的缩减, 从而缓解了在学习新的类别时可能出现的对先前类别知识的破坏, 间接地避免了灾难性遗忘问题。

由于在类增量无源域适应中, 源域数据是不可见的, 缺乏来自源域样本的正负对, 但是预训练模型的分层权重向量通常可以被看作是源域上学习到的每个类的中心特征。因此, 如图 2 下半部分所示, 可以通过将源模型分类器 H_s 的权重提取出来作为类原型的正负对来计算类对比损失, 从而解决这个问题。

$$H_s \rightarrow [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_K] \quad (10)$$

其中, \mathbf{w}_k 表示分类层中第 k 个类的权重向量。因此, 基于多原型的跨域对比损失可以表示为:

$$L_{cdc} = -\frac{1}{|\mathcal{R}_{T_n}|} \sum_{i=1}^{|\mathcal{R}_{T_n}|} \log \frac{\exp(G_t(r_i) \mathbf{w}_{\hat{y}_i} / \tau)}{\sum_{j=1}^K \exp(G_t(r_i) \mathbf{w}_j / \tau)} \quad (11)$$

其中, τ 是温度系数, 通常设定为 10。

3.5 总体优化目标

结合之前所有的损失项, 可以得出 MPRA 方法的总体优化目标, 如式(12)所示:

$$\min_{(\theta_{G_t}, \theta_{H_t}, \mathbf{x})} L_{ce}(\theta_{G_t}, \theta_{H_t}, \mathbf{x}) + \alpha L_{cdc}(\theta_{G_t}, \mathbf{r}, \mathbf{w}) + \beta L_{rep}(\theta_{G_t}, \theta_{H_t}, \mathbf{r}) \quad (12)$$

其中, α 和 β 为损失项之间的平衡超参数, θ_{G_t} 和 θ_{H_t} 分别为特征提取器和分类器的参数。可以看到, 交叉熵损失和重放损失对特征提取器和分类器的参数都进行了优化, 而跨域对比损失仅对特征提取器进行优化。详细的算法流程如算法 1 所示。

算法 1 MPRA 算法

输入: 当前时间步骤的样本 $\{x_i\}_{i=1}^{N_t}$, 重放池 $\mathcal{R}_{T_{n-1}}$, 预训练源模型 M_s ,

返回: 目标域模型 M_t ,

1. 根据式(1)识别共享类;
2. For epoch=1,2,...,E do
3. 根据式(6)计算交叉熵损失;
4. 根据式(8)计算类原型更新重放池;
5. 根据式(9)计算重放损失;
6. 根据式(11)计算跨域对比对齐损失;
7. 根据式(12)对目标模型进行优化;
8. End for
9. 返回当前时间步骤的目标模型 M_t .

4 实验

本章在 3 个公开的基准数据集上对所提的 MPRA 方法进行评估。首先介绍了本文采用的数据集和实验设置;其次通过对比、消融、参数分析等实验进行分析,验证了 MPRA 方法的有效性和优越性。

4.1 数据集

本文基于 3 个标准的域适应数据集来模拟类增量无源域适应场景,即 Office-31^[39], Office-Home^[40] 和 ImageNet-Caltech^[35]。作为域适应和迁移学习广泛使用的一个数据集, Office31 包含了办公环境中常见的 31 个对象类别,包括电脑、键盘等。该数据集由 4 652 张图像组成,分别来自 Amazon(A)、DSLR(D)和网络摄像头(W)这 3 个不同的领域。第二个数据集 Office-Home 是一个更大、更全面的数据集,由 15 500 张分布在 4 个领域、65 个类别的图像组成,即 Art (Ar), Clipart(Cl), Product(Pr) 和 Real-word(Rw)。值得注意

的是,它在领域之间呈现出显著的视觉差异和域偏移,成为域适应算法在不同视觉环境中实现鲁棒性泛化的重大障碍。第三个数据集 ImageNet-Caltech 由 ImageNet-1K 和 Caltech-256 构成,基于它们共享的 84 个类别形成了两个不同的域适应任务。

4.2 实验设置

为了充分体现本文方法针对类增量无源域适应的有效性,除了与现有的引入类增量的域适应方法 CIDA^[37] 以及 ProCA^[38] 进行对比外,还与源模型 ResNet50、一些经典的无监督域适应以及部分域适应方法进行比较,包括 DANN(Domain-Adversarial Training of Neural Networks)^[41], ETN(Example Transfer Network)^[42], PADA^[35] 和 BA³US(Balanced Adversarial Alignment and Adaptive Uncertainty Suppression)^[43]。为了方便比较,本文遵循 ProCA 中的设置,在实验过程中,每个时间步骤增加 10 个类别,并使用在 ImageNet 上预训练的 ResNet50 作为骨干网络。使用学习率为 0.001 的 SGD 的优化器来优化网络权重参数,样本批次大小设为 64,超参数 α, β, L 分别设置为 0.1, 1, 6。对于每个时间步骤中样本的训练,分别为 Office-31, Office-Home 和 ImageNet-Caltech 设定 10, 30, 15 个 epoch 进行训练并保存最优模型。

4.3 实验结果对比和分析

本节在 3 个数据集上进行了多个领域之间的适应任务,以此来评估所提方法在类增量无源域适应场景中的优越性。比较的所有结果均引自相关文献。在 3 个数据集上的实验结果如表 2 和表 3 所列,最优结果用加粗表示。从表中可以看出,MPRA 方法在 3 个数据集上的大多数任务和最终平均准确率上都取得了最优的结果。

表 2 Office-31 和 ImageNet-Caltech 上的最终准确率

Table 2 Final accuracy on Office-31 and ImageNet-Caltech

(%)

Method	DA	SF	CI	Office-31						ImageNet-Caltech			
				A→D	A→W	D→A	D→W	W→A	W→D	Avg	C→I	I→C	Avg
ResNet50	×	×	×	74.1	74.4	58.5	96.9	61.2	99.6	77.5	72.3	70.7	71.5
DANN ^[41]	√	×	×	74.9	72.5	55.7	96.6	51.4	97.7	74.8	58.8	31.4	45.1
ENT ^[42]	√	×	×	21.3	82.2	61.7	94.3	64.1	100.0	70.6	1.4	3.1	2.3
PADA ^[35]	√	×	×	56.9	61.5	12.5	82.4	46.7	84.3	57.4	37.3	45.9	41.6
BA ³ US ^[43]	√	×	×	74.1	73.3	63.3	94.8	64.0	100.0	78.3	60.8	45.0	52.9
CIDA ^[37]	√	√	√	70.4	64.5	48.1	95.1	52.7	98.8	71.6	69.3	49.2	59.2
ProCA ^[38]	√	×	√	81.8	82.5	65.2	99.1	64.1	99.6	82.1	82.9	83.1	83.0
MPRA(ours)	√	√	√	84.3	83.0	68.7	95.4	70.7	99.4	83.6	84.2	83.5	83.9

表 3 Office-Home 上的最终准确率

Table 3 Final accuracy on Office-Home

(%)

Method	Ar Cl	Ar Pr	Ar Rw	Cl Ar	Cl Pr	Cl Rw	Pr Ar	Pr Cl	Pr Rw	Rw Ar	Rw Cl	Rw Pr	Avg.
ResNet50	47.6	65.2	72.7	54.7	62.8	66.1	52.4	44.7	74.0	66.2	47.4	77.4	60.9
DANN ^[41]	33.1	40.0	45.8	36.8	36.6	44.1	32.0	29.8	49.8	42.4	40.2	55.2	40.5
ENT ^[42]	42.4	2.8	7.4	4.3	60.3	6.3	50.7	33.8	70.8	3.7	43.5	75.1	33.4
PADA ^[35]	24.8	41.4	55.1	18.3	35.0	36.3	25.9	26.2	53.7	46.8	31.4	50.0	37.1
BA ³ US ^[43]	33.7	39.7	63.2	36.6	39.1	53.7	36.5	24.9	53.4	52.2	35.9	65.9	44.6
CIDA ^[37]	32.2	45.9	49.1	36.5	38.6	46.6	51.6	33.5	59.0	64.0	38.0	65.1	47.5
ProCA ^[38]	51.9	75.2	86.1	60.8	69.7	74.7	60.1	51.0	84.2	75.8	51.2	86.4	68.9
MPRA(ours)	56.2	79.7	87.2	60.2	74.6	75.5	59.9	49.7	85.2	77.5	58.2	87.8	71.0

从表2和表3中可以观察到,在面对类增量域适应时,没有经过针对性处理的普通无监督域适应方法和部分域适应方法的分类准确度会有很大程度的下降,其表现甚至不如不作任何处理但仅在源域上训练过的 ResNet50。这是因为在对齐过程中模型会过于关注当前时间步骤中的目标域类别,尤其是部分域适应方法,这使得模型忘记了之前的类别知识。此外,从不同数据集之间的比较结果可以看出,由于 Office-31 数据集只有 31 个类,只会进行两次更新训练,并且样本数量相对较少,因此模型性能下降有限。但是,当面对需要进行 5 次更新和 7 次更新且样本数量更多、更复杂的 Office-Home 和 ImageNet-Caltech 时,模型性能的下降尤为明显。这意味着,这种灾难性遗忘现象会随着样本数量和模型更新次数的增加变得更为严重。CIDA 方法虽然设计了一个正则化项来缓解灾难性遗忘问题,但是忽略了源私有类,这可能会导致负迁移现象。ProCA 虽然充分考虑了这些问题,但是其存储的原型过于单一,没有更充分地表示整个类别知识且没有考虑无源场景。所提方法通过使用多原型的重放,能够更有效地保留类别知识;同时,针对无法接触源域数据的情况,还提出了针对性的自训练方式和跨域对比对齐方法,其性能

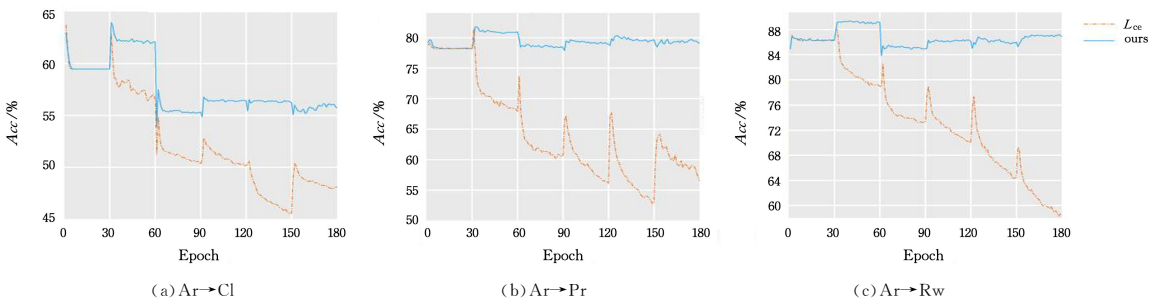


图3 分类准确率趋势对比

Fig. 3 Comparison of classification accuracy trends

从图3中可以清楚地观察到在类增量无源域适应中存在的灾难性遗忘现象。当仅使用交叉熵训练时(如虚点线所示),每当下一个时间步骤有新的类别样本到来时,整体分类准确率都会先上升,但随着不断地训练,模型在学习新类别知识的同时会对先前类别的知识造成破坏,从而造成整体分类准确率的快速下降。但是 MPRA(如实线所示)在很大程度上解决了这个问题,甚至在 Ar→Rw 任务上训练后期的性能还有所提升,证明了本文方法的有效性。此外,从表4中可以得出结论:MPRA 方法中的所有模块均发挥着至关重要的作用。从第1—第3行可以发现,单独使用交叉熵损失受到灾难性遗忘的影响较大,而重放损失仅能起到约束知识遗忘的效果,并不能从根本上提高精度。跨域对比对齐损失虽然能够通过对比样本特征进行对齐来提高分类准确性,但终究会受到类原型数量的限制,难以达到最佳性能。而从表4的第4—第7行则可以发现,当在自训练模块的基础上结合类原型重放模块和跨域对比对齐模块后,MPRA 的性能得到显著提升,这是因为跨域对比对齐能够缓解域偏移问题,提高伪标签的正确率,提升每个时间步骤的预测精度,进而也能提高类原型重放时的准确性,提高整体分类精度。

在3个数据集上都有很大的提升。

4.4 消融实验

为了评估本文提出的方法(MPRA)的各个组成模块的有效性,本节在 Office-Home 数据集的 Ar→Cl, Ar→Pr, Ar→Rw 这3个任务上系统评估了3个关键组成部分,即 L_{ce} , L_{rep} 和 L_{con} 。实验结果如表4所列。同时,为了清晰地观察到类增量学习中的灾难性遗忘现象,进一步将仅用交叉熵损失训练过程中的分类准确率变化趋势与 MPRA 方法训练过程中的分类准确率变化趋势进行对比。具体实验结果如图3所示。

表4 不同损失项的消融实验

Table 4 Ablation experiment of different loss items

L_{ce}	L_{rep}	L_{cdc}	Ar→Cl	Ar→Pr	Ar→Rw
✓			50.2	64.3	68.8
	✓		50.9	65.9	72.1
		✓	53.3	71.1	78.7
✓	✓		55.9	78.5	86.2
	✓	✓	54.7	72.8	79.1
✓		✓	55.1	77.2	83.9
✓	✓	✓	56.2	79.7	87.2

4.5 参数敏感性分析

本节分析了 MPRA 中几个重要参数对实验精度的影响,包括原型个数(L)的设置以及共享类识别的阈值($thre$)设置,依然是通过 Office-Home 数据集上的 Ar→Cl, Ar→Pr, Ar→Rw 这3个任务来进行评估,实验结果如图4所示。从图4(a)中可以发现,最终的分类准确率会随着原型个数的增加而上升,但是当达到一定数量后会变缓慢直至停止上升并稳定在最优值左右。这是因为当数量较少时,这部分原型代表的类知识有限,尚不能够很好地约束模型、记忆更全面的类知识;但是当达到一定数量后,灾难性遗忘不再是最主要的问题,而是在域适应任务中最根本的问题,即域偏移问题,这最终限制了模型的性能提升。而从图4(b)中可以观察到,当阈值较大时,模型性能会有较大程度的下降,且下降程度比取较小值时更为明显。这是因为当阈值取较小值时,当前时间步骤中会识别出本不存在的错误类别,从而导致存储错误的类原型。但是本文方法考虑了类原型更新策略,当在后续步骤中存在正确的类别时,会计算得出高于先前时刻的累积预测概率,从而进行错误类原型的更新,并学习到正确的类别知识。而阈值取较大值时,可能会导致当前时间步骤本应该被识别出

的正确类别被遗漏,且因后续不存在该类别而无法得到更新,造成整体精度的下降。

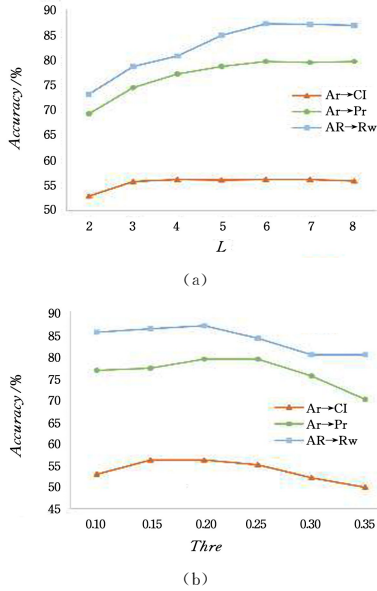


图4 参数敏感性分析

Fig. 4 Parameter sensitivity analysis

此外,本节还对每个时间步骤中到来的新样本的类别数量对整体分类准确率的影响进行研究,采用每次增加10,20,30个新类别的方式进行实验。实验结果如图5所示。可以观察到,在Ar->Pr和Ar->Rw两个任务上,当每次新类别数量增加时,整体的分类准确率有明显的下降,这说明多的类别数量会造成对先前样本知识更加严重的遗忘,从而影响整体的分类准确率。

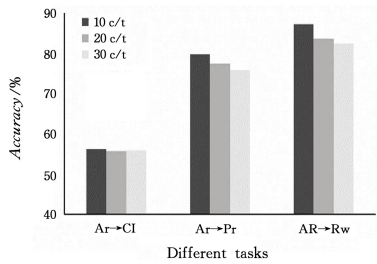


图5 每个时间步骤增加类别数量对分类精度的影响

Fig. 5 Impact of increasing the number of categories at each time step on classification accuracy

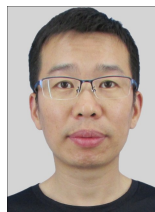
结束语 本文深入分析了类增量无源域适应相较于传统无源域适应场景面临的新挑战,包括标签空间的不一致性和学习新知识时可能导致的旧知识灾难性遗忘等。为了应对这些新挑战,本文针对性地提出了一种基于多原型重放和对齐的方法,旨在提高模型在类增量无源域适应场景中的性能,并减少旧知识的遗忘。它首先通过源模型对目标域样本的累积预测概率进行共享类的识别,然后通过类内聚类的方式找到最具有代表性的原型样本进行存储,并通过在后续训练中进行重放来记忆旧类别知识。同时,本文进一步提出基于多原型的跨域对比对齐方法,在进行目标域与源域对齐、降低域偏移带来的影响的同时,隐式地约束新类别的特征空间,减少对旧类别知识的破坏。实验结果表明,MPRA方法在类增量无源域适应问题中有着优越的表现。在后续研究中,我们将针对

域适应领域的核心问题(即域偏移问题)进行更深入的研究,寻求更有效的解决方案,以提高每个时间步骤的分类精度。同时,我们也将关注开集场景的处理,即在目标域出现源域中的未知类别时,如何确保模型的鲁棒性和性能稳定性。

参考文献

- [1] HE K,ZHANG X,REN S,et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:770-778.
- [2] CHEN G,CHEN K,ZHANG L,et al. VCANet:Vanishing-point-guided context-aware network for small road object detection[J]. Automotive Innovation,2021,4:400-412.
- [3] XU Y,ZHANG Q,ZHANG J,et al. Vitae: Vision transformer advanced by exploring intrinsic inductive bias[J]. Advances in Neural Information Processing Systems,2021,34:28522-28535.
- [4] ZHENG Z D,YANG Y. Rectifying pseudo label learning via uncertainty estimation for domain adaptive semantic segmentation [J]. International of Computer Vision, 2021, 129 (4) : 1106-1120.
- [5] TIAN Q,SUN H,MA C,et al. Heterogeneous domain adaptation with structure and classification space alignment[J]. IEEE Transactions on Cybernetics,2021,52(10):10328-10338.
- [6] HOFFMAN J,TZENG E,PARK T,et al. Cycada: Cycle-consistent adversarial domain adaptation[C]// International Conference on Machine Learning. 2018:1989-1998.
- [7] ZHOU K B,TENG L Y,ZHANG W,et al. Discriminative label semantic guidance learning for domain adaptive retrieval[J]. Journal of Chinese Computer Systems,2024,45(7):1639-1647.
- [8] DU Z,LI J,SU H,et al. Cross-domain gradient discrepancy minimization for unsupervised domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:3937-3946.
- [9] PAN S J,TSANG I W,KWOK J T,et al. Domain adaptation via transfer component analysis[J]. IEEE Transactions on Neural Networks,2010,22(2):199-210.
- [10] LONG M,CAO Z,WANG J,et al. Conditional adversarial domain adaptation[C]// Proceedings of the 32nd International Conference on Neural Information Processing System. 2018:1647-1657.
- [11] BOUSMALIS K,SILBERMAN N,DOHAN D,et al. Unsupervised pixel-level domain adaptation with generative adversarial networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:3722-3731.
- [12] LONG M,CAO Y,Wang J,et al. Learning transferable features with deep adaptation networks[C]// International Conference on Machine Learning. 2015:97-105.
- [13] LIANG J,HU D,FENG J. Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation[C]// International Conference on Machine Learning. 2020:6028-6039.
- [14] ZHANG Z,CHEN W,CHENG H,et al. Divide and contrast: Source-free domain adaptation via adaptive contrastive learning [J]. Advances in Neural Information Processing Systems,2022, 35:5137-5149.
- [15] DING N,XU Y,TANG Y,et al. Source-free domain adaptation

- via distribution estimation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 7212-7222.
- [16] LIANG J, HU D, WANG Y, et al. Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(11): 8602-8617.
- [17] CHU T, LIU Y, DENG J, et al. Denoised Maximum Classifier Discrepancy for Source-Free Unsupervised Domain Adaptation [C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2022:3472-3480.
- [18] JING M M, LI J J, LU K, et al. Visually source-free domain adaptation via adversarial style matching[J]. IEEE Transactions on Image Processing, 2024, 33: 1032-1044.
- [19] BELOUADAH E, POPESCU A, KANELLOS I. A comprehensive study of class incremental learning algorithms for visual tasks[J]. Neural Networks, 2021, 135: 38-54.
- [20] MASANA M, LIU X, TWARDOWSKI B, et al. Class-incremental learning: Survey and performance evaluation on image classification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(5): 5513-5533.
- [21] ZHAO B, XIAO X, GAN G J, et al. Maintaining discrimination and fairness in class incremental learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:13208-13217.
- [22] HU X, TANG K, MIAO C, et al. Distilling causal effect of data in class-incremental learning [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:3957-3966.
- [23] LIU Y, SCHIELE B, SUN Q. Rmm: Reinforced memory management for class-incremental learning[J]. Advances in Neural Information Processing Systems, 2021, 34: 3478-3490.
- [24] BELOUADAH E, POPESCU A. IL2m: Class incremental learning with dual memory[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019:583-592.
- [25] LIU X, WU C, MENTA M, et al. Generative feature replay for class-incremental learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020:226-227.
- [26] YANG S, WANG Y, et al. Generalized source-free domain adaptation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021:8978-8987.
- [27] YANG S, WANG Y X, et al. Exploiting the intrinsic neighborhood structure for source-free domain adaptation[J]. Advances in Neural Information Processing Systems, 2021, 34: 29393-29405.
- [28] WANG F, HAN Z, GONG Y, et al. Exploring domain-invariant parameters for source free domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022:7151-7160.
- [29] TIAN J, ZHANG J, LI W, et al. VDM-DA: Virtual domain modeling for source data-free domain adaptation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 32(6): 3749-3760.
- [30] CHEN D, WANG D, DARREEL T, et al. Contrastive test-time adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022:295-305.
- [31] WANG R, WU Z, WENG Z, et al. Cross-domain contrastive learning for unsupervised domain adaptation[J]. IEEE Transactions on Multimedia, 2022, 25: 1665-1673.
- [32] TIAN Q, PENG S, MA T. Source-free unsupervised domain adaptation with trusted pseudo samples[J]. ACM Transactions on Intelligent Systems and Technology, 2023, 14(2): 1-17.
- [33] CAO Z, LONG M, WANG J, et al. Partial transfer learning with selective adversarial networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 2724-2732.
- [34] TIAN Q, CHU Y, SUN H Y, et al. Survey on Partial Domain Adaptation[J]. Journal of Software, 2023, 34(12): 5597-5613.
- [35] CAO Z, MA L, LONG M, et al. Partial adversarial domain adaptation[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 135-150.
- [36] SAHOO A, PANDA R, FERIS R, et al. Select, label, and mix: Learning discriminative invariant feature representations for partial domain adaptation[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2023: 4210-4219.
- [37] KUNDU J N, VENKATESH R M, VENKAT N, et al. Class-incremental domain adaptation [C]//European Conference on Computer Vision. 2020:53-69.
- [38] LIN H, ZHANG Y, QIU Z, et al. Prototype-guided continual adaptation for class-incremental unsupervised domain adaptation [C]//European Conference on Computer Vision. 2022:351-368.
- [39] SAENLO K, KULIS B, FRITZ M, et al. Adapting visual category models to new domains[C]//Proceedings of European Conference on Computer Vision. Springer. 2010:213-226.
- [40] TZENG E, HOFFMAN J, SAENKO K, et al. Adversarial discriminative domain adaptation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 7167-7176.
- [41] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. The Journal of Machine Learning Research, 2016, 17(1): 2096-2030.
- [42] CAO Z, YOU K, LONG M, et al. Learning to transfer examples for partial domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019:2985-2994.
- [43] LIANG J, WANG Y, HU D, et al. A balanced and uncertainty-aware approach for partial domain adaptation[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 123-140.



TIAN Qing, born in 1984, Ph.D, professor, is a senior member of CCF (No. 33364S). His main research interests include machine learning and pattern recognition.