

文章编号: 2095-2163(2024)01-0001-08

中图分类号: TP391

文献标志码: A

基于渐进式双重对齐的无源无监督领域自适应方法

杨艳¹, 陈利娟², 唐 宋¹, 叶 茂³

(1 上海理工大学 机器智能研究院, 上海 200093; 2 上海电机学院 电气学院, 上海 201306;

3 电子科技大学 计算机科学与工程学院, 成都 611731)

摘要: 无源领域自适应的核心任务是利用无标签的目标域数据,将预训练好的源模型迁移到目标领域。基于深度聚类的方法需要在自监督学习过程中挖掘辅助信息来正则化特征分布对齐,而辅助信息中噪声常常误导该对齐过程;基于伪源域的对抗学习方法进行概率分布对齐,对所构建伪源域质量十分敏感。针对现有方法存在的不足,本文提出了一种基于渐进式双重对齐的无源无监督领域自适应方法,在进行深度聚类的同时,进行域对齐,缓解深度聚类中伪标签的噪声。首先,通过超近邻增强样本生成高质量伪源域,以克服源域不可见的问题;其次,利用对抗学习,实现两个域的概率分布初对齐;最后,引入深度特征相似,进一步强化对齐效果。在两个公开数据集上的实验结果表明了其有效性。

关键词: 领域自适应; 对抗学习; 自监督学习; 伪源域; 深度聚类

Gradual dual alignment for source-free unsupervised domain adaptation

YANG Yan¹, CHEN Lijuan², TANG Song¹, YE Mao³

(1 Institute of Machine Intelligence, University of Shanghai for Science and Technology, Shanghai 200093, China;

2 School of Electrical Engineering, Shanghai DianJi University, Shanghai 201306, China; 3 School of Computer

Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: The principle task of source-free domain adaptation is to transfer a pre-trained source model from the source domain to the target domain by only using unlabeled target domain data. The deep clustering-based methods mine the auxiliary information from the self-supervised learning process to regulate the feature distribution alignment. However, the noise in the auxiliary information always misleads this alignment. Besides, the adversarial learning-based methods, conducting a probability distribution alignment, are sensitive to the constructed pseudo source domain. Aiming at the two shortcomings, this paper proposes a new gradual dual alignment (GDA) approach, which performs deep clustering where the errors in pseudo-labels are alleviated by a domain alignment. Specifically, this hybrid framework adopts the classical paradigm to perform deep clustering. As for the domain adaptation, a dual alignment including probability distribution and feature alignment is developed: First, a pseudo-source domain is constructed by hyper-nearest neighbor sample generation to overcome the problem of the invisible source domain; followed the adversarial learning is used to achieve initial alignment of the probability distributions of the two domains; At the end, deep feature alignment is introduced to further enhance the alignment effect. Experimental results on two public datasets demonstrate its effectiveness.

Key words: domain adaptation; adversarial learning; self-supervised learning; pseudo source domain; deep clustering

0 引言

传统的无监督领域自适应 (Unsupervised Domain Adaptation, UDA),借助有标签的源域数据和无标签的目标域数据,实现源域与目标域分布对齐。解决该问题的关键是减少两个领域之间的差

异,目前主要采用量度学习和对抗学习来实现对齐。在对抗域适应中,文献[1]提出域对抗 (Domain-Adversarial Neural Network, DANN) 模型最具代表性,引入了域判别器和域对抗损失来减小域差异;文献[2]提出的条件域对抗 (Conditional Domain Adversarial Network, CDAN) 模型,使用信息熵来辅

基金项目: 国家自然科学基金(62206168)。

作者简介: 杨 艳(1998-),女,硕士研究生,主要研究方向:机器学习、计算机视觉;唐 宋(1982-),男,博士,副教授,主要研究方向:迁移学习、计算机视觉;叶 茂(1972-),男,博士,教授,主要研究方向:机器学习、计算机视觉。

通讯作者: 陈利娟(1986-),女,博士,讲师,主要研究方向:忆阻神经网络、机器学习。Email: chenlj@sdju.edu.cn

收稿日期: 2023-01-02

哈尔滨工业大学主办 ◆ 学术研究与应用

助对抗域适应。随着对信息安全与隐私保护的重视,“数据是一种关键资产”已成为一种共识,在此情景下,源域数据越来越难以获得。因此,依赖源域数据的UDA不再适用;最近的研究也发现,对于领域自适应问题,迁移过程保持所有源数据可见并不是必须的^[3-4]。而无源领域自适应(Source-Free Domain Adaptation, SFDA)因为更接近真实应用场景中预训练模型迁移的需求,具有天然数据保护的属性,越来越受到关注。与UDA不同,SFDA无需源数据,只使用源领域预训练好的模型和无标注目标域数据进行跨域迁移。

现有解决SFDA思路主要分为两种:一是在深度聚类(Deep Clustering, DC)的框架下挖掘辅助信息,通过自学习在特征空间实现从目标域向源域的特征分布对齐。文献[4]提出了源假设迁移方法(Source Hypothesis Transfer, SHOT),将深度聚类与互信息最大化(Information Maximization, IM)结合,实现目标域的特征分布和源假设对齐;文献[5]提出基于最近邻深度聚类(Nearest Neighborhood-Based Deep Clustering, N2DC)方法,利用目标数据之间的几何信息与信息熵,构建超近邻领域,实现了鲁棒的深度聚类;文献[6]提出基于最近邻几何信息的知识蒸馏的迁移方法(Gradual Knowledge Distillation, GKD)。二是构造伪源域,通过对抗学习实现伪源域和目标域的概率分布对齐。文献[7]提出模型自适应方法(Model Adaptation, MA),利用生成对抗网络生成伪源域,支撑后续基于对抗学习的域对齐;文献[8]提出基于伪源泉域的方法(Pseudo-Source Domain Based Source-Free Domain Adaptation, PS)并不生成伪源域,而是基于目标域样本,通过信息熵筛选构建伪源域;文献[9]提出基于生成对比原型的方法(Contrastive Prototype Generation and Adaptation, CPGA),通过挖掘源模型中的隐含信息来生成源域中每个类别的伪质心特征,并联合对比学习实现域对齐。

虽然以上方案取得了不错的迁移效果,但存在两方面的不足:第一种方案本质上是进行特征分布对齐,过程易受到辅助信息中噪声的误导;第二种方案本质上是进行概率分布的对齐,对伪源域构造质量有较高要求。

本文提出了一种基于渐进式双重对齐的无源无监督领域自适应方法(Gradual Dual Alignment for Domain Adaptation, GDA),该方法提出了一种混合学习框架,在深度聚类的同时,进行域概率分布对

齐,缓解深度聚类中伪标签的噪声。为了实现域对齐,本文提出了一种新的伪源域构造方法,利用对抗学习与特征相似实现概率分布与特征双重对齐,该伪源域由随机增强的目标域数据的超近邻样本构成。相比之前的方法,本文所构建的伪源域数据数量级与目标域一致。由于超近邻样本综合考虑了信息熵和数据间的几何信息,拥有高置信度的伪标签,能更准确地表达源域的分类特征。鉴于目标域样本与其超近邻样本之间的类别相似性,本文以超近邻增强样本的特征作为引导,通过施加特征相似约束加强对齐效果,进一步提升迁移性能。

1 预备知识

1.1 SFDA 问题定义

在SFDA任务中,使用 $D_s = \{\mathbf{x}_i^s, \mathbf{y}_i^s\}_{i=1}^{n_0}$ 来表示有标签的源域, $D = \{\mathbf{x}_i\}_{i=1}^{n_1}$ 表示无标签的目标域,其中 $\mathbf{x}_i^s \in X_s, \mathbf{y}_i^s \in Y_s, \mathbf{x}_i \in X, n_0$ 和 n_1 分别表示源域和目标域的样本数量。源域和目标域拥有相同的特征空间和标签空间, $X_s = X = R^d, Y_s = Y = \{1, 2, \dots, K\}$, $P_s(X_s, Y_s), P(X, Y)$ 分别表示源域和目标域样本的概率分布, $P_s(X_s, Y_s) \neq P(X, Y)$ 。由于无法直接访问源域的数据,只能获得一个由源样本训练的模型 F_s ,其中 $F_s = X_s \rightarrow Y_s$ 。SFDA的目标则是利用预训练的源模型 F_s 和无标签的目标域数据 D 来训练一个在目标域中表现良好的模型 $F, F = X \rightarrow Y$ 。

1.2 面向SFDA的深度聚类框架

面向SFDA的深度聚类是一个渐进式的学习过程,包含 T 个训练阶段(Epoch),第 t 个训练阶段记为 $E_t, t \in \{1, 2, \dots, T\}$ 。训练阶段 E_t 的目标是在已知 F_{t-1} 即上一阶段训练的条件下,学习中间模型 F_t ,所采用目标函数为式(1):

$$\begin{cases} \mathcal{L}_{SHOT} = \mathcal{L}_{im}(F_t) + \mathcal{L}_{ce}(F_t; D; \hat{\mathbf{y}}_{t-1}) \\ \mathcal{L}_{im}(F_t; D) = \mathcal{L}_{ent}(F_t; D) + \mathcal{L}_{div}(F_t; D; \hat{\mathbf{y}}_{t-1}) \end{cases} \quad (1)$$

式(1)中各损失函数分量如式(2)~式(4)所示:

$$\mathcal{L}_{ce}(F_t; D; \hat{\mathbf{y}}_{t-1}) = - \mathbb{E}_{(\mathbf{x}_i, \hat{\mathbf{y}}_i) \in D \times \hat{\mathbf{y}}_{t-1}} \sum_{k=1}^K q_k \ln \delta_k(\mathbf{a}) \quad (2)$$

$$\mathcal{L}_{div}(F_t; D) = \sum_{k=1}^K p_k \ln p_k \quad (3)$$

$$\mathcal{L}_{ent}(F_t; D) = - \mathbb{E}_{\mathbf{x}_i \in D} \sum_{k=1}^K \delta_k(\mathbf{a}) \ln \delta_k(\mathbf{a}) \quad (4)$$

其中, $\mathbf{a} = F_t(\mathbf{x}_i), \delta_k(\mathbf{a})$ 表示 K 维向量 \mathbf{a} 经过

Softmax 处理后的第 k 个元素; p_k 表示目标样本在模型 F_t 上的输出通过 Softmax 处理后的均值, 即 $p_k = \mathbb{E}_{\mathbf{x}_i \in D} [\delta_k(F_t(\mathbf{x}_i))]$; $\hat{Y}_{t-1} = \{\hat{y}_i\}_{i=1}^{n_1}$ 由基于 F_{t-1} 的加权 K-means 聚类方法得到的目标域数据伪标签集合; q 表示伪标签 \hat{y}_i 的热编码; q_k 表示在第 k 个类别上的编码值, 值为 1 代表正确类别, 其余全为 0。

面向 SFDA 的深度聚类不仅利用基于交叉熵的自监督分类损失 \mathcal{L}_{ce} 来约束模型, 还添加了信息最大化 \mathcal{L}_{im} 损失, 使目标样本既独立又均衡。本文在结合深度聚类 and 对抗学习的混合学习框架下, 进行模型优化。

2 GDA 模型

2.1 渐进双重对齐框架

GDA 采用渐进式策略进行模型训练: 以源模型为起点, 在阶段 E_1 中, 训练中间模型 F_1 , 以此类推, 阶段 E_t 对应着中间模型 F_t 的训练, 直到完成全部 T 个训练阶段。在所有训练阶段中, 模型结构一致, 如图 1 所示。在 E_t 中, 被训练的模型 F_t 包含一个特征提取器 f_t 、降维模块 b_t 、分类器 h_t 、域对抗模块 f_{t-1} 是模型 F_{t-1} 的特征提取器, 其中域对抗模块由一个梯度反向层 (Gradient Reversal Layer, GRL) 和一个域判别器 G 组成。

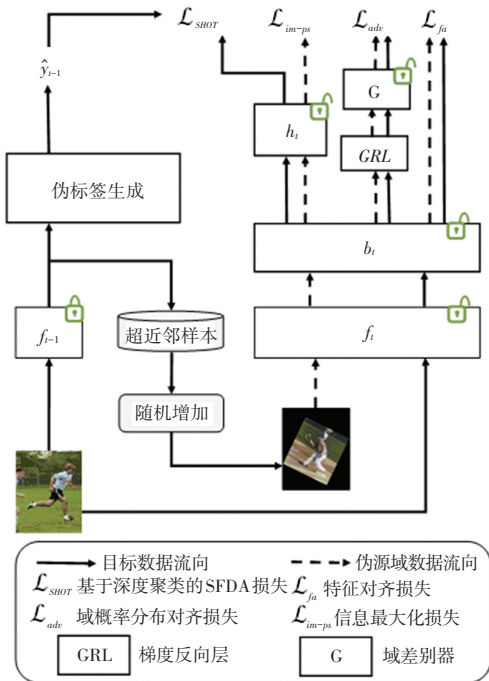


图 1 GDA 模型结构

Fig. 1 The architecture of GDA

在 E_t 的开始阶段, GDA 利用模型 F_{t-1} 为训练样

本寻找超近邻增强样本, 构建用于本阶段双重对齐的伪源域; GDA 用 F_{t-1} 中训练好的 f_{t-1} 、 b_{t-1} 和 h_{t-1} 分别初始化 f_t 、 b_t 和 h_t 而 F_1 则采用源模型 F_s 进行初始化。为了训练 F_t , GDA 采用混合学习目标函数来监督在目标域上的自学习过程, 其中 \mathcal{L}_{short} 表示面向 SFDA 的深度聚类损失函数, 而双重对齐由 \mathcal{L}_{adv} 、 \mathcal{L}_{fa} 与 \mathcal{L}_{im-ps} 联合表达。

2.2 伪源域构建方法

伪源域的构建质量直接影响最终的迁移效果, 训练数据数量均衡且兼具表征多样性的数据更有利于表达源领域的分布情况, 本文提出两阶段的构建方法。首先, 对于目标域数据 $D = \{\mathbf{x}_i\}_{i=1}^{n_1}$, 通过相似性搜索建立类别相关的初始伪源域 $\tilde{D}_{ps} = \{\tilde{\mathbf{x}}_i\}_{i=1}^{n_1}$, 获得与训练数据规模一致、语义表达关联的数据, 因为这些样本具有较高的置信度, 由其构成的伪源域更能接近真实源域分布情况; 其次, 本文使用文献 [10] 中提出的方法, 对 \tilde{D}_{ps} 进行随机增强, 通过诸如旋转、平移、翻转等数据变换, 生成具有表征多样性的伪源域 $D_{ps} = \{\hat{\mathbf{x}}_i\}_{i=1}^{n_1}$ 。构建过程可表示为式 (5):

$$D_{ps} = RA(S(D)) \quad (5)$$

其中, $RA(\cdot)$ 表示随机增强函数, $S(D) = \tilde{D}_{ps}$ 表示一种由相似性搜索确定的类别相关构造函数。

首先, 从目标域样本中划分出可信的子集 D_{con} , 即在每一个训练阶段 E_t 开始前, 基于目标域数据 D 的信息熵, 将目标域所有样本信息熵升序排列, 取前 50% 的样本为熵较小的集合 D_{con}^e ; 同时, 基于目标域数据 D 的深层特征与所属伪标签的类别的几何中心的几何距离, 将此距离值升序排列, 取前 50% 的样本为距离几何中心较近的集合 D_{con}^d , 再通过取熵较小的集合 D_{con}^e 和距离几何中心较近的集合 D_{con}^d 的交集, 得到最终的可信样本集合 $D_{con}, D_{con} = D_{con}^e \cap D_{con}^d$ 。

其次, 利用链式搜索构建初始伪源域 \tilde{D}_{ps} 。对于该集合中的任意一个样本 $\tilde{\mathbf{x}}_i \in \tilde{D}_{ps}$, $\tilde{\mathbf{x}}_i$ 被称为 \mathbf{x}_i 的超近邻样本, 其所对应的训练样本为 \mathbf{x}_i , 从 \mathbf{x}_i 到 $\tilde{\mathbf{x}}_i$ 的搜索过程在特征空间中进行。假设特征提取器 f_{t-1} 将所有目标域数据 D 映射为 $Z_{t-1} = \{\mathbf{z}_i\}_{i=1}^{n_1}$, 并将其作为搜索空间, $\mathbf{z}_i = f_{t-1}(\mathbf{x}_i)$; \mathbf{x}_i 与 $\tilde{\mathbf{x}}_i$ 在该空间的特征点分别为 \mathbf{z}_i 与 $\tilde{\mathbf{z}}_i$ 。以 \mathbf{z}_i 为初始点, 寻找离其最近的点 \mathbf{z}_{i1} , 如果 \mathbf{z}_{i1} 对应的样本 \mathbf{x}_{i1} 属于可信集合

D_{con} , 则 \mathbf{x}_{i1} 为超近邻样本, 即 $\tilde{\mathbf{x}}_i = \mathbf{x}_{i1}$; 反之, 则以 \mathbf{z}_{i1} 为新的起点, 寻找其最近邻样本, 再判断是否属于 D_{con} 。不断重复“寻找—判断”步骤, 形成链式搜索, 直到寻找到的点属于 D_{con} 。

2.3 领域概率分布对齐

得到伪源域 D_{ps} 后, 利用对抗学习对齐该域和目标域。本文添加了一个域判别器 G 用来判别特征来自于伪源域还是目标域, 该域判别器通过二进制分类损失对域进行区分, 伪源域的标签设为 1, 目标域的标签设为 0, 基于伪源域的域对抗优化目标函数如式(6)所示:

$$\mathcal{L}_{adv}(f_t; b_t; G; D; D_{ps}) = \mathbb{E}_{\hat{\mathbf{x}}_i \in D_{ps}} [\ln G(b_t(f_t(\hat{\mathbf{x}}_i)))] + \mathbb{E}_{\mathbf{x}_i \in D} [\ln(1 - G(b_t(f_t(\mathbf{x}_i))))] \quad (6)$$

模型 GDA 使用梯度反向层 GRL 实现最大化域判别器 G 和最小化特征提取器 f_t 和 b_t , 使伪源域和目标域混淆, 通过这种对抗训练的方式, 模型 GDA 学习到了域不变特征, 从而实现了伪源域和目标域的概率分布对齐。

2.4 领域特征对齐

为了加强对齐效果, 本文将目标样本的特征与对应伪源域中的特征进行进一步对齐, 提出了基于特征相似的特征对齐损失函数。文献[11]提出的平滑 $L1$ 损失在参数优化的鲁棒性上优于 $L1$ 和 $L2$ 。本文采用平滑 $L1$ 损失来实现对齐, 具体如式(7)所示:

$$\mathcal{L}_{fa}(f_t; b_t; D; D_{ps}) = \begin{cases} 0.5 (v_i - \hat{v}_i)^2, & \text{if } |v_i - \hat{v}_i| < \beta \\ |v_i - \hat{v}_i| - 0.5 * \beta, & \text{otherwise} \end{cases} \quad (7)$$

其中, $v_i = b_t(f_t(\mathbf{x}_i))$ 为目标样本经过降维模块 b_t 后的特征; $\hat{v}_i = b_t(f_t(\hat{\mathbf{x}}_i))$ 为对应的伪源域中的超近邻增强样本经过降维模块 b_t 后的特征; β 为平滑因子。

此外, 本文对伪源域进行随机增强, 虽然强化了多样性, 但随机增强会造成几何分布向外一定程度的扩展, 降低了域中样本的置信度。本文提出了针对所构造伪源域的信息最大化损失函数 \mathcal{L}_{im-ps} , 使伪源域样本分布更紧凑, 保持其高置信度, 可具体表示为式(8):

$$\mathcal{L}_{im-ps}(F_t; D_{ps}) = \mathcal{L}_{ent}(F_t; D_{ps}) + L_{div}(F_t; D_{ps}) \quad (8)$$

GDA 在深度聚类的时候, 进行域概率分布与特征双重对齐, 结合式(1)、式(6)~式(8), GDA 的最终目标函数可表示为式(9):

$$\mathcal{L}_{GDA} = \mathcal{L}_{SHOT} + \omega \mathcal{L}_{adv}(f_t; b_t; G; D; D_{ps}) + \eta \mathcal{L}_{fa}(f_t; b_t; D; D_{ps}) + \mathcal{L}_{im-ps}(F_t; D_{ps}) \quad (9)$$

其中, ω, η 为超参数。

2.5 算法伪代码

GDA 具体地实现细节如算法 1 所示。

算法 1 GDA 算法训练过程

输入 源域预训练模型 F_s (包含 f_s, b_s 和 h_s), 无标签目标域 D , 批处理大小 B , 训练轮数 T ;

输出 目标模型 F_t (包含 f_t, b_t 和 h_t)。

初始化 使用 F_s 初始化 F_t^0 ; 随机初始化判别器 G 。

for $i = 1$ to T do:

 获取自监督伪标签, 每轮训练的迭代次数 M 。

 伪源域生成: 通过公式(5)获取对应的超近邻增强样本, 构建伪源域;

 for $j = 1$ to M :

 从目标域中抽取 B 个样本; 从伪源域中抽取上述 B 个目标样本对应的超近邻增强样本;

 深度聚类损失: 通过公式(1)训练;

 双重对齐:

 域对抗对齐: 通过公式(6)训练、和域判别器 G ;

 特征对齐: 通过公式(7)和公式(8)训练 F_t ;

 end

end

return

3 实验与分析

3.1 实验数据集

本文采用两个被广泛使用的 UDA 标准测试数据集, 来验证 GDA 的有效性。

Office-31 是一个被广泛应用于领域自适应图像分类任务的小型数据集, 由 Amazon (A)、Webcam (W) 与 DSLR (D) 3 个域组成, 包含 31 个类别共计 4652 张图片, 其中, A 域图片来自亚马逊网站 2 817 张, W 域图片由网络摄像头拍摄 795 张, D 域图片由单反相机拍摄 498 张。3 个域, 共计构成 6 组迁移任务, 分别是 $A \rightarrow D$ 、 $A \rightarrow W$ 、 $D \rightarrow A$ 、 $D \rightarrow W$ 、 $W \rightarrow A$ 以及 $W \rightarrow D$ 。鉴于 Office-31 数据集轻量、迁移任务丰富的特点, 通常使用 Office-31 数据集对实验方法进行快速验证。

VisDA-C 是一个大规模领域自适应图像分类

数据集,专注于从虚拟到现实的目标识别,共计 12 个类型的迁移任务,其中源域(T)来自渲染 3D 模型生成的 152 397 张合成图像,目标域(V)为真实图像,包含 55 388 张图片,来自于微软开源的 MSCOCO 数据集。较之 Office-31 数据集,VisDA-C 中的合成图像域与真实图像域之间的差异十分巨大。本文使用 VisDA-C 数据集,不仅能够对 GDA 在大规模数据下且域差异较大时的域适应能力进行评估,而且能充分发挥深度模型的潜力。

3.2 实验设置

在源模型训练方面,本文在源领域上利用有标注数据 D_s , 采用经典的 Softmax 交叉熵目标函数实现源模型 F_s 的训练。为了提高源模型 F_s 的判别性,本文使用带标签平滑的交叉熵损失函数来优化 F_s 。

在网络结构方面,本文采用 ImageNet 上预训练的残差网络 ResNet 作为特征提取器,输出特征维度为 2 048;降维模块由一个全连接层与一个批次归一化层(Batch Normalization, BN)构成,输出维度为 256;分类输出头的分类器由一个全连接层与一个权重归一化层(Weight Normalization, WN)构成,输出维度 K 与目标分类任务相关;对齐输出头的域对抗模块采用被广泛使用的标准组件,输出一个二维的向量,用以对特征表示进行域判别。同时,为了公平比较,作为特征提取器的 ResNet 在不同数据集上采用不同的深度结构,在 Office-31 上使用 ResNet-50 作为特

征提取器,而在 VisDA-C 上使用 ResNet-101。

在训练细节方面,目标模型使用动量为 0.9 的随机梯度下降和权重衰减率为 10^{-3} 来更新网络参数,域判别器 G 的学习率为 10^{-3} 。两个数据集的 Epoch 均设置为 12, batch_size 均设置为 64, 超参数 ω 设为 1, η 设为 0.5, 平滑因子 β 设为 1。本文实验采用 Pytorch 1.6.0 深度学习框架,编程环境为 Python3.7,实验运行环境为 ubuntu18.04 系统, GPU 为 TITAN RTX, 显存为 24 GB, CUDA 版本为 11.4。

3.3 对比方法

为了证明 GDA 的有效性,本文选取 15 种典型方法作为对比,这些方法可以分为以下 3 种类型:

一是作为算法基础的模型,包括被用作特征提取器的残差网络 ResNet-50、ResNet-101 和源模型(Source only);

二是传统的 UDA 方法,包括 DANN、CDAN,这些方法的迁移过程对有标注的源域数据存在依赖;

三是不需要源域数据的 SFDA 方法,包括 SHOT、N2DC、MA、GKD、PS、CPGA,其中,SHOT 作为本文的基线,MA、PS、CPGA 是最近的基于对抗学习构造伪源域的方法,GKD 和 N2DC 是最近的基于深度聚类的框架的方法。

3.4 实验结果分析

GDA 在 Office-31 与 VisDA-C 上的多分类跨域迁移实验结果见表 1 和表 2。

表 1 在 office-31 数据集上的实验结果

Table 1 Classification accuracy (%) on Office-31

方法	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
ResNet-50	68.9	68.4	62.5	96.7	60.7	99.3	76.1
DANN	79.7	82.0	68.2	96.9	67.4	99.1	82.2
CDAN	92.9	94.1	71.0	98.6	69.3	100.0	87.7
Source-only	80.7	77.0	60.8	95.1	62.3	98.2	79.0
SDA	92.2	91.1	71.0	98.2	71.2	99.5	87.2
SHOT-IM	89.4	90.6	73.1	98.4	70.8	100.0	87.1
N2DC	93.9	89.8	74.7	98.6	74.4	100.0	88.6
SHOT	94.2	90.2	74.1	98.5	74.7	100.0	88.6
GDA(ours)	94.8	90.7	74.9	98.0	75.3	99.8	88.9

根据表 1 可知,GDA 在 Office-31 小数据集集中的 3 个迁移任务 D→A、A→D 与 W→A 上取得了最好的结果,相对于基线方法 SHOT,分别提高了 0.8、0.6、0.6 个百分点;相对于 SOTA (State of the Art) 方法,N2DC 分别提高了 0.2、0.9、0.9 个百分点。相应地,GDA 在平均分类准确率达到 88.9,较之 SHOT

与 SOTA,都提高了 0.3 个百分点。

根据表 2 可知,GDA 在 VisDA-C 大型数据集上同样取得了最好的结果,不仅在 7/12 的类别上打败了其它方法,而且在平均准确率上获得最佳。在最好的 7 个类中,4 个类别上的提升较大:较之 SHOT,“car”类提升了 12.1 个百分点,“mcycl”类提升了

10.9个百分点,“bus”和“person”类分别提升了5个百分点和4.8个百分点;从整体精度来看,GDA在准确率上平均提升了2.4个百分点,与最近的使用自监督知识蒸馏的GKD相比,提升了2.1个百分点,与最近的采用构建伪源域或伪类原型的无源无监督领域自适应的方法相比,GDA仍然超越了PS和CPGA,平均提升了1个百分点。

小数据集不能很好地支撑深度模型GDA端到

表2 在VisDA-C数据集上的实验结果

Table 2 Classification accuracy (%) on VisDA-C

Method	plane	bicycle	bus	car	horse	knife	mcycl	person	plant	sktbrd	train	truck	Avg.
ResNet-101	55.1	53.3	61.9	59.1	80.6	17.9	79.7	31.2	81.0	26.5	73.5	8.5	52.4
DANN	81.9	77.7	82.8	44.3	81.2	29.5	65.1	28.6	51.9	54.6	82.8	7.8	57.4
CDAN	85.2	66.9	83.0	50.8	84.2	74.9	88.1	74.5	83.4	76.0	81.9	38.0	73.9
Source-only	62.1	21.2	48.8	77.8	63.1	5.0	72.90	25.9	66.1	44.1	80.9	5.3	47.8
SDA	86.9	81.7	84.6	63.9	93.1	91.4	86.6	71.9	84.5	58.2	74.5	42.7	76.7
MA	94.8	73.4	68.8	74.8	93.1	95.4	88.6	84.7	89.1	84.7	83.5	48.1	81.6
SHOT	95.0	87.5	80.8	57.6	93.8	93.9	79.5	80.5	90.7	89.8	85.7	57.4	82.7
GKD	95.3	87.6	81.7	58.1	93.9	94.0	80.0	80.0	91.2	91.0	86.9	56.1	83.0
CPGA	94.8	83.6	79.7	65.1	92.5	94.7	90.1	82.4	88.8	88.0	88.9	60.1	84.1
PS	95.3	86.2	82.3	61.6	93.3	95.7	86.7	80.4	91.6	90.9	86.0	59.5	84.1
GDA (ours)	95.9	86.5	85.8	69.7	94.8	91.2	90.4	85.3	91.9	91.8	86.6	51.8	85.1

3.5 消融实验

为了分析GDA中域概率分布对齐和特征对齐对模型优化的定量影响,本文按照分量递增的方式依次实验,得到了在VisDA-C上各目标函数的消融结果,见表3。与SHOT相比,GDA模型在仅仅只添加域对抗损失函数 \mathcal{L}_{adv} 就已经提升了1.7个百分点,超越了最近的SFDA方法CPGA和PS,本文提出的利用超近邻增强样本构建伪源域进行领域概率分布对齐的有效性。在加入特征对齐损失分量 \mathcal{L}_{fa} 和伪源域信息最大化损失分量 \mathcal{L}_{im-ps} 后,GDA模型达到了最优的效果,相比SHOT提升2.4个百分点,相比SOTA提升1.0个百分点。

表3 在VisDA-C上各目标函数的消融结果

Table 3 Ablation of losses on VisDA-C

\mathcal{L}_{SHOT}	\mathcal{L}_{adv}	\mathcal{L}_{fa}	\mathcal{L}_{im-ps}	Avg. /%
✓	×	×	×	82.7 ($\omega = 0; \eta = 0$)
✓	✓	×	×	84.4 ($\omega = 1; \eta = 0$)
✓	✓	✓	×	85.0 ($\omega = 1; \eta = 0.5$)
✓	✓	✓	✓	85.1 (GDA)

3.6 可视化分析

为了进一步验证GDA的有效性,本文还利用t-SNE工具对模型迁移前后的深度特征进行可视化分

析。在迁移任务 $W \rightarrow A$ 上(属于数据集Office-31),分别采用SHOT和GDA进行跨域迁移前后目标域深度特征的分布情况如图2所示,可以看到在迁移前,目标域数据分布没有明显的聚类效果,整体较为混乱,通过SHOT方法迁移后,目标域的分布呈明显紧致分块式分布,说明这些特征类别区分度大大提高,模型更容易基于这些特征对目标域数据进行分类。通过GDA迁移后,特征分布更加具有区分度,也证实了GDA方法的有效性。

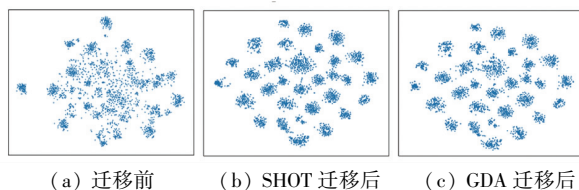
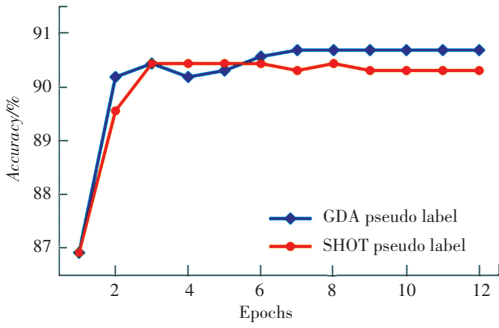


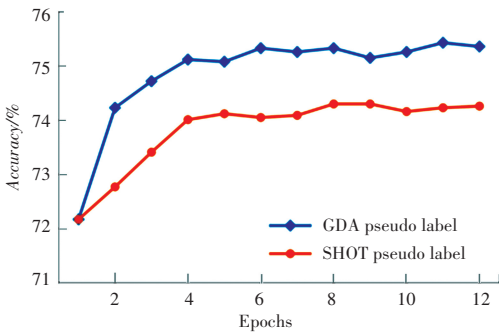
图2 $W \rightarrow A$ 任务迁移前后特征分布图

Fig. 2 The feature distribution

$A \rightarrow W$ 迁移任务和 $W \rightarrow A$ 迁移任务在训练过程中伪标签的准确率的变化过程如图3所示。GDA同SHOT一样,通过类似加权K-means策略生成伪标签,在迁移后GDA伪标签的准确率优于SHOT,说明了融合深度聚类和对抗学习的混合学习范式能够缓解深度聚类中伪标签的噪声。



(a) A→W 迁移过程中的伪标签准确率



(b) W→A 迁移过程中的伪标签准确率

图 3 伪标签准确率

Fig. 3 The pseudo label accuracy

3.7 训练稳定性

对于无监督学习任务,训练的稳定性是至关重要的。本文在每个训练阶段最后进行性能评估,绘制出 GDA 在数据集 VisDA-C 上的分类准确率的变化情况,为了增加对比,SHOT 的性能变化曲线被用作基线,结果如图 4 所示。可见,随着 Epoch 的增加,从 Epoch-1 到 Epoch-7, GDA 的曲线快速上升到 84.5%,再经过 5 个 Epoch 逐渐趋于稳定到最好的结果。同时,不论是爬升的速度还是持续能力, GDA 曲线都明显优于 SHOT 曲线,进一步验证了本文混合学习模式的合理性。

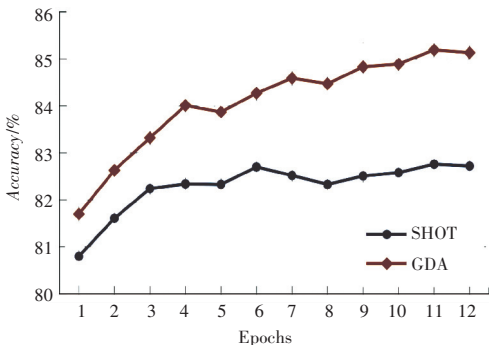


图 4 在 VisDA-C 上训练的准确率结果

Fig. 4 The accuracy of VisDA-C

4 结束语

本文提出了一种基于渐进式双重对齐的无源无监督领域自适应方法,采用域概率分布和特征双重对齐来缓解深度聚类范式中伪标签的噪声;为了实现域对齐,提出通过超近邻增强样本生成伪源域,以克服在 SFDA 中源域不可见的问题,再利用对抗学习对齐概率分布,同时通过深度特征相似来强化对齐效果。本文在两个标准测试数据集 Office-31 和 VisDA-C 进行实验,与多种近期的 SFDA 方法进行对比,验证了所提方法的有效性。但是,本文所生成的伪源域仅基于目标域,对源模型所蕴含的有用信息还有待挖掘。在未来研究中,可以结合生成模型与目标域,来构建更符合真实源域概率分布的伪源域。利用辅助信息直接在特征层进行特征对齐方法可运用到多源领域自适应、开放领域自适应任务中。

参考文献

- [1] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]// Proceedings of the International Conference on Machine Learning. New York: JMLR.org, 2015: 1180-1189.
- [2] LONG M, CAO Z, WANG J, et al. Conditional adversarial domain adaptation [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc., 2018: 31.
- [3] KIM Y, CHO D, HAN K, et al. Domain adaptation without source data [J]. IEEE Transactions on Artificial Intelligence, Piscataway, NJ:IEEE., 2021, 2(6): 508-518.
- [4] LIANG J, HU D, FENG J. Do we really need to access the source data? Source hypothesis transfer for unsupervised domain adaptation [C]// Proceedings of the International Conference on Machine Learning. New York: JMLR.org, 2020: 6028-6039.
- [5] TANG S, YANG Y, MA Z, et al. Nearest neighborhood-based deep clustering for source data-absent unsupervised domain adaptation[J]. arXiv preprint arXiv:2107.12585, 2021.
- [6] TANG S, SHI Y, MA Z, et al. Model adaptation through hypothesis transfer with gradual knowledge distillation [C]// Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and System(IROS). IEEE, 2021:5679-5685.
- [7] LI R, JIAO Q, CAO W, et al. Model adaptation: Unsupervised domain adaptation without source data [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 9641-9650.
- [8] DU Y, YANG H, CHEN M, et al. Generation, augmentation, and alignment: A pseudo-source domain based method for source-free domain adaptation [J]. arXiv preprint arXiv:2109.04015, 2021.
- [9] QIU Z, ZHANG Y, LIN H, et al. Source-free domain adaptation via avatar prototype generation and adaptation [C]//Proceedings of International Joint Conference on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann Publishers Inc, 2021: 31.