

基于风格嵌入与协同学习的多目标域自适应方法

丁美荣, 刘庆龙, 郎济聪, 卓金鑫

(华南师范大学软件学院, 广东 佛山 528225)

摘要: 多目标域自适应旨在探究如何从单个源域提取知识并将其应用到多个目标域, 包括如何确保提取到的领域不变特征具有更良好的类别区分性, 以及如何考虑单个源域与多个目标域之间的相关性等问题。为此, 提出一种基于风格嵌入和协同学习的多目标域自适应模型(SECL)。首先, 通过改变样本的特征统计量提升数据多样性; 其次, 通过双分类器的协同学习生成优化后的伪标签进行模型自训练, 采用大间隔分类调整不同类别之间的分布, 提升模型泛化能力; 最后, 采用课程学习的训练策略引导模型先适应简单领域, 再逐步过渡到更复杂的领域, 使目标域适应过程更平滑。实验表明, 该方法在Office-Home、Office31数据集、ImageCLEF-DA数据集上的平均分类准确率相较于D-CGCT分别提升2.9%、1.1%、1%。

关键词: 域自适应; 多目标; 图像分类; 伪标签; 自训练

DOI: 10.11907/rjdk.241409

中图分类号: TP18

文献标识码: A

开放科学(资源服务)标识码(OSID):

文章编号: 1672-7800(2025)006-0065-07



Multi-Target Domain Adaptation Approach Based on Style Embedding and Co-Learning

DING Meirong, LIU Qinglong, LANG Jicong, ZHUO Jinxin

(School of Software, South China Normal University, Foshan 528225, China)

Abstract: Multi objective domain adaptation aims to explore how to extract knowledge from a single source domain and apply it to multiple target domains, including how to ensure that the extracted domain invariant features have better category discriminability, and how to consider the correlation between a single source domain and multiple target domains. Therefore, a multi-objective domain adaptive model (SECL) based on style embedding and collaborative learning is proposed. Firstly, by changing the feature statistics of the samples, data diversity can be enhanced; Secondly, through collaborative learning of dual classifiers, optimized pseudo labels are generated for model self training, and large interval classification is used to adjust the distribution between different categories, thereby improving the model's generalization ability; Finally, the training strategy of course learning is adopted to guide the model to adapt to simple domains first, and then gradually transition to more complex domains, making the target domain adaptation process smoother. The experiment shows that the average classification accuracy of this method on Office-Home, Office31 dataset, and ImageCLEF-DA dataset is improved by 2.9%, 1.1%, and 1% compared to D-CGCT.

Key Words: domain adaptation; multi-target; image classification; pseudo-labels; self-training

0 引言

随着智能大数据时代的到来,海量数据以前所未有的

速度和规模不断涌现,为机器学习尤其是深度学习提供了强有力的数据支持,使其在图像分类、目标检测、语义分割等领域得到广泛关注,并取得了显著进展^[1-3]。机器学习算法通常是在拥有大量标记数据,且源域(训练集)和目标

收稿日期:2024-04-28

扫描二维码阅读全文:



作者简介: 丁美荣(1972-),女,硕士,CCF高级会员,华南师范大学软件学院副教授,研究方向为自然语言处理和智能软件开发、人工智能教育;刘庆龙(1999-),男,华南师范大学软件学院硕士研究生,研究方向为领域自适应、图像分类;郎济聪(1998-),男,华南师范大学软件学院硕士研究生,研究方向为领域自适应、语义分割;卓金鑫(1999-),男,华南师范大学软件学院硕士研究生,研究方向为迁移学习、图像分类。本文通讯作者:刘庆龙。

域(测试集)服从相同分布的前提下训练模型。然而,在实际应用场景中,源域数据和目标域数据往往展现出不同的分布特性,难以实现数据分布的一致性。例如,相同类别的图像在不同场景、光照条件或图像采集设备下可能呈现出不同的表现形式(艺术绘画、拍摄照片和剪切画等);人工收集和标注大量数据代价昂贵且耗时较大。因此,源域和目标域的数据分布差异限制了传统机器学习模型的泛化能力,影响了其在新领域的预测能力。

领域自适应(Domain Adaptation, DA)方法致力于解决数据分布不匹配的问题,其目标是对源域和目标域的数据进行建模,以最小化源域和目标域的数据分布差异^[4,5]。最终目的是即使在目标域缺乏标签信息的情况下,模型也能通过源域的分类知识进行学习,以准确分类目标域数据。然而,大多数领域自适应方法主要针对单个目标域展开研究,忽略了目标域数据来自不同场景,不同来源的目标域数据相较于源域数据在特征分布上具有不同的相似性和差异性,甚至各领域之间的类别属性也会发生变化^[6]。在这种包含多个目标域的情况下,单目标域自适应方法的作用受到限制。为此,不少学者提出了多目标域自适应(Multi-Target Domain Adaptation, MTDA)方法,旨在探究如何从单个源域中提取知识并应用多个目标域^[7]。

尽管一些多目标域自适应方法被提出,但仍存在诸多问题未得到充分解决。例如,如何保证提取到的领域不变特征具有良好的类别区分性,如何对齐单个源域与多个目标域之间的不同数据分布,如何充分考虑源域和目标域及多个目标域之间的相关性等。本文针对目前 MTDA 方法存在的问题,在探索 MTDA 方法的优化方向后提出一种基于风格嵌入和协同学习的 MTDA 方法,以提升模型对多个无标注目标域的分类准确率,强化了模型的泛化能力。

1 相关工作

MTDA 是领域自适应中极具挑战的研究方向,主要研究如何将知识从单个标记的源域迁移到多个未标记的目标域中。Roy 等^[8]提出一种课程图协同教学的 MTDA 模型,引入图卷积神经网络(Graph Convolutional Network, GCN)对伪标签进行预测与校正,通过模型的迭代自训练来增强伪标签的准确性和可靠性^[9,10]。此外,该模型还巧妙融入了领域感知的课程学习策略,按照难易程度分阶段训练,先易后难逐步深化对目标领域的理解,旨在更好地对齐特征分布。Wang 等^[11]提出一种判别式互学习的 MTDA 模型,考虑了两个关键点:①利用判别机制使模型学习到的图像特征既具有域不变性又具有区分性,减小了预测标签与真值之间的误差;②基于距离信息生成可靠的伪标签,然后用带有伪标签的目标数据进行自训练。值得注意的是,该模型在混合目标域自适应(Blending Target Domain Adaptation, BTDA)和 MTDA 设置上都表现出优越的

性能。

Nguyen-Meidine 等^[12]提出一种无监督的 MTDA 模型,依赖知识蒸馏(Knowledge Distillation, KD)迭代地将多个教师模型输入一个共同的学生模型中。其中,KD 过程以渐进方式执行,每个教师模型交替地训练学生模型如何为特定目标域执行无监督学习,从而在学习学生模型时保留每个教师模型的特异性。Zhou 等^[13]提出一种渐进式适应策略与解耦学习相结合的 MTDA 模型,通过构建跨域的解耦表示实现特征空间优化。具体而言,模型通过改进的特征解耦机制,将提取到的特征区分为类无关特征、领域不变特征和特定于领域的特征,以减少源域和目标域在特征分布上的差异,确保了特征层面的高效对齐。Wang 等^[14]提出一种注意力引导的多源域和多目标域的自适应方法,采用多种对抗策略从多个源域和目标域中提取更具区分性的特征,并提出一个域间与域内的注意力模块捕获可迁移的语义信息。Yang 等^[15]提出一种异构图注意力神经网络来探索多目标域样本节点之间的关系,通过计算源域、目标域及多个目标域之间的相关性来提升模型性能。

受以上工作启发,本文基于域间相关性和类别区分性展开探索,提出一种新的 MTDA 方法。该方法既考虑了领域间的相关性和差异性,又能聚合样本的关键特征,主要贡献为:①为了提升数据多样性并减轻因数据分布差异带来的性能下降隐患,在特征提取器中添加混合风格嵌入模块,确保生成的图像既能继承源域图像的关键语义特征,又能融合目标域图像的独特风格属性,从而达到增强数据多样性和提升模型泛化能力的目的;②引入 GCN 分类器,借助图网络特性构建领域间的互联连接,让模型捕捉和学习不同领域数据的内在关联,并在此基础上提出生成对抗网络(Generate Adversarial Networks, GAN)、多层感知器(Multilayer Perceptron, MLP)的学习机制,二者相互借鉴、共同优化,为模型自训练提供更可靠的伪标签,以增强预测准确性;③引入大间隔分类模块增强特征表示,确保类别间的界限更明晰,以进一步提升模型分类性能。

2 SECL 模型

2.1 问题定义

在 MTDA 场景中涉及一个标记的源域 $S = \{(\mathbf{x}_{s,i}, y_{s,i})\}_{i=1}^{n_s}$ 与 M 个未标记的目标域 $T = \{T_m\}_{m=1}^M$, 每个源域由 n_s 个样本构成, $y_{s,i}$ 代表源域第 i 个样本 $\mathbf{x}_{s,i}$ 的真实标签。第 m 个目标域 $T_m = \{\mathbf{x}_{t_m,j}\}_{j=1}^{n_{t_m}}$ 由 n_{t_m} 个样本组成, 目标域中所有样本无标签, 与其他领域自适应场景相同。MTDA 遵循源域与目标域的数据分布不同, 但标签空间相同这一假设, 旨在使用源域和目标域数据训练出一个能适应所有目标域的通用预测器。

2.2 域对抗网络

图1为基于风格嵌入和协同学习 (Style Embedding and Co-Learning, SECL) 的模型,采用域对抗网络 (Domain adversarial NEURAL NETWORK, DANN) 作为基础框架^[16]。

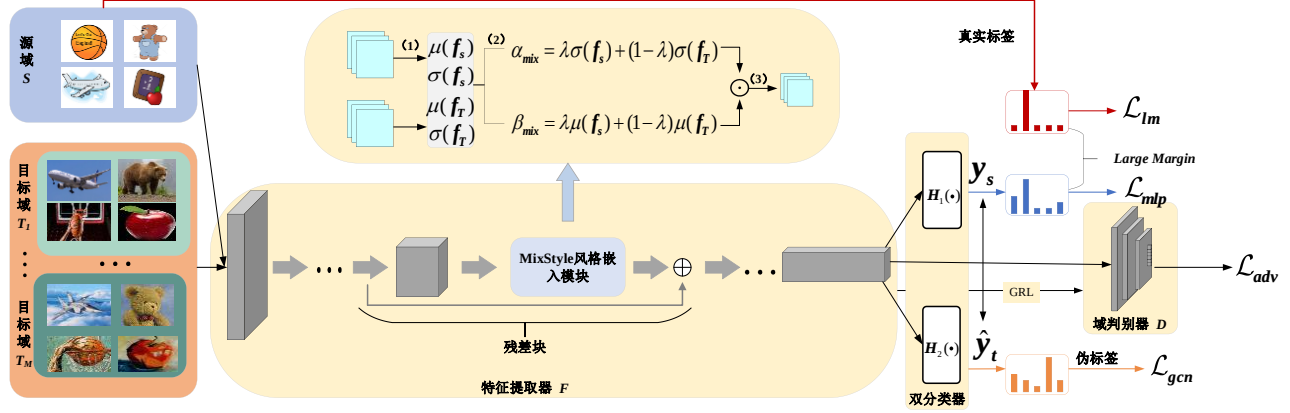


Fig. 1 SECL model

图1 SECL模型

特征提取器 F 与域判别器 D 之间通过梯度反转层 (Gradient Reversal Layer, GRL) 连接。特征提取器 F 力求从源域与目标域中提取的特征难以被域判别器 D 分辨,从而以对抗训练的方式缩小源域与目标域之间的分布差异。对齐分布所产生的域适应损失函数如式(1)~式(3)所示。SECL 分别在特征提取器 F 与分类器 H 这两个部分进行优化,以改进模型性能,提升模型在目标域上的泛化能力。

$$\mathcal{L}_{adv} = \frac{1}{n_s} \sum_{x_i \in S} \mathcal{L}_h(H(F(x_i)), y_i) - \frac{\lambda}{n_s + n_{T_m}} \sum_{x_i \in (S \cup T)} \mathcal{L}_d(D(F(x_i)), d_i) \quad (1)$$

$$\hat{\theta}_f, \hat{\theta}_h = \arg \min_{\theta_f, \theta_h} \mathcal{L}(\theta_f, \theta_h, \hat{\theta}_d), \quad (2)$$

$$\hat{\theta}_d = \arg \max_{\theta_d} \mathcal{L}(\hat{\theta}_f, \hat{\theta}_h, \theta_d), \quad (3)$$

式中: $\mathcal{L}_h()$ 为分类损失; $\mathcal{L}_d()$ 为域判别损失; $\theta_f, \theta_h, \theta_d$ 分别为特征提取器 F 、分类器 H 和域判别器 D 参数; $\hat{\theta}_f, \hat{\theta}_h, \hat{\theta}_d$ 为优化后的参数; λ 为权衡因子。

2.3 MixStyle风格嵌入模块

SECL 借鉴了领域泛化的思想,在特征提取器中添加 MixStyle 风格嵌入模块建立视觉域与图像特征统计之间的关系,以实现数据增强。由于图像特征均值 μ 、方差 σ 基本捕获了图像的风格信息,因此在特征提取器的前几层加入风格嵌入模块,通过合成具备新领域特征图像来提升数据多样性。自适应实例归一化 (adaptive instance normalization, AdaIN) 是用于风格迁移的常见技术,通过调整内容图像的特征均值与方差以匹配风格图像来实现风格迁移^[17]。该方法的核心思想是通过特征统计量 (尤其是通道均值和方差) 捕捉图像的风格信息,在调整统计量后将一

DANN 由特征提取器 F 、分类器 H 和域判别器 D 等3部分组成。其中,特征提取器 F 用于学习源域与目标域中具有类别区分性的特征表示;分类器 H 则基于这些特征表示预测样本的分类归属;域判别器 D 是标准的二分类分类器,用来确定样本领域来源。

种风格应用到另一张图像上,以保留内容图像的结构和细节。

同时,受 AdaIN 的启发提出 MixStyle 风格嵌入模块。MixStyle 不依赖附加的解码器生成图像,而是通过混合源域与目标域图像的风格信息优化特征提取器通过训练。首先,通过特征提取器的前几层卷积计算源域与目标域图像特征的均值 $\mu(f)$ 、方差 $\sigma(f)$,如式(4)、式(5)所示;其次,计算一个具备复杂风格信息的混合特征统计量 (α_{mix} 、 β_{mix}),如式(6)、式(7)所示;最后,将混合特征统计量用于实例归一化,生成具有目标风格的新特征图 $MixStyle(f_s, f_T)$,如式(8)所示。

$$\mu(f) = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W f_{c,h,w} \quad (4)$$

$$\sigma(f) = \sqrt{\frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W (f_{c,h,w} - \mu(f))^2} \quad (5)$$

$$\alpha_{mix} = \varphi \sigma(f_s) + (1 - \varphi) \sigma(f_T) \quad (6)$$

$$\beta_{mix} = \varphi \mu(f_s) + (1 - \varphi) \mu(f_T) \quad (7)$$

$$MixStyle(f_s, f_T) = \alpha_{mix} \odot \frac{f_s - \mu(f_s)}{\sigma(f_s)} + \beta_{mix} \quad (8)$$

式中: C, H, W 分别代表特征图 f 的通道数、高度和宽度; φ 为一个权重参数。

综上, MixStyle 通过融合、嵌入源域与目标域图像的特征统计,使合成图像在保留源域图像语义信息的同时融入了目标域的风格信息,实现了数据增强的目标。

2.4 双分类器协同学习机制

单个分类器预测具有一定的随机性,容易错误预测决策边界周围样本。为了防止单个分类器过度拟合自身噪声的伪标签,提出一种双分类器协同的学习机制。该机制结合了 MLP 与 GCN 分类器,对目标域图像进行预测,并将

预测结果提供给对方进行二次预测。首先,将特征输入 GCN 分类器,边网络通过构造亲和度矩阵 $\mathbf{A}$ 来表示样本间的相似度, $a_{i,j}$ 值越大代表样本 i 和样本 j 属于相同类别的可能性越大,反之越小;其次,经过节点分类器更新节点,输出样本属于某个类的概率,由于边网络训练需要带标签的样本,所以将 MLP 分类器的输出(大于某个阈值)作为节点分类器的先验知识,使节点分类器的输出向真实标签靠拢;最后,边网络使用二进制交叉熵损失进行训练,MLP 分类器和节点分类器使用标准交叉熵损失进行训练,如式(9)~式(11)所示。

$$\mathcal{L}_{edge} = \hat{a}_{i,j} \log p(a_{i,j}) + (1 - \hat{a}_{i,j}) \log (1 - p(a_{i,j})) \quad (9)$$

$$\mathcal{L}_{node} = -\frac{1}{|B|} \sum_{i=1}^{|B|} y_{x_i} \log(p(\hat{y}_{x_i})) \quad (10)$$

$$\mathcal{L}_{mlp} = -\frac{1}{|B|} \sum_{i=1}^{|B|} y_{x_i} \log(p(\tilde{y}_{x_i})) \quad (11)$$

式中: $\hat{a}_{i,j}$ 表示第 i 个节点与第 j 个节点之间边的实际标签,通常为 0 或 1; $p(a_{i,j})$ 指边网络预测第 i 个节点与第 j 个节点之间存在边的概率; y_{x_i} 代表第 i 个样本的真实标签; $\tilde{y}_{x_i}$ 、 $\hat{y}_{x_i}$ 分别代表 MLP 分类器和 GCN 分类器对第 i 个样本属于各个类别的预测概率。

GCN 分类器训练产生的损失为:

$$\mathcal{L}_{gcn} = \mathcal{L}_{edge} + \omega \mathcal{L}_{node} \quad (12)$$

式中: ω 为损失函数 $\mathcal{L}_{node}$ 的权重。

对于目标域图像,SECL 通过设定一个阈值 τ 挑选一些高置信度样本来参与双分类器的协同训练。挑选规则如式(13)所示。

$$\forall x_{T_m,j} \in T, \alpha_j = \begin{cases} 1, & \text{if } \max_c p(\hat{y}_{T_m,j} = c | x_{T_m,j}) > \tau \\ 0, & \text{其他} \end{cases} \quad (13)$$

式中: $p(\hat{y}_{T_m,j} = c | x_{T_m,j})$ 代表第 m 个目标域 T_m 所获得的标签分布; $\alpha_j = 1$ 表示样本 $x_{T_m,j}$ 被分配伪标签且参与双分类器的协同训练; $\alpha_j = 0$ 表示样本 $x_{T_m,j}$ 不参与训练。

最后,挑选的高置信度样本会被整合到数据集 $\mathbf{D}_m = \{(x_{T_m,j}, \hat{y}_{T_m,j})\}_{j=1}^{n_{T_m}}$ 中, n_{T_m} 表示 $\mathbf{D}_m$ 的样本数量。

2.5 大间隔分类模块

为了提升模型泛化能力,使其学习到更具区分性的特征。SECL 模型在基于源域数据的预训练阶段引入大间隔分类模块,旨在扩大不同类别间的差异性,提升模型分类性能,损失函数如式(14)所示。

$$\mathcal{L}_{lm} = g_{\hat{y}_{x_i} \neq y_{x_i}} \max \{0, \theta + d_{p_{x_i}, \{\hat{y}_{x_i}, y_{x_i}\}} \text{sign}(p_{\hat{y}_{x_i}}(x_i) - p_{y_{x_i}}(x_i))\} \quad (14)$$

$$d_{p_{x_i}, \{k_1, k_2\}} = \min_{\sigma} \|\sigma\|_h, s.t. p_{k_1}(x_i + \sigma) = p_{k_2}(x_i + \sigma) \quad (15)$$

式中: $d_{p_{x_i}, \{y_{x_i}, \hat{y}_{x_i}\}}$ 表示样本 x_i 在预测标签 $\hat{y}_{x_i}$ 和真实标签 y_{x_i} 之间的差异度量;为了简化公式, k_1 、 k_2 代表预测值和真实值; $\|\cdot\|_h$ 为 l_h 范数; θ 为一个阈值参数。

由此可知,当 $d_{p_{x_i}, \{y_{x_i}, \hat{y}_{x_i}\}} > \theta$ 时损失函数会惩罚预测值

与真实值差异较大的样本,当预测值与真实值足够接近时损失不会增加,从而鼓励模型更自信地预测这些样本。

2.6 课程学习训练模块

本文依序对目标域进行适应,如果先对齐与源域差异较大的目标域,可能会影响源域差异较小的目标域对齐过程,降低模型整体性能。为此,本文采用课程学习的训练策略,让模型先从简单任务开始,逐步过渡到更复杂的任务,从而使目标域对齐过程更平滑。课程学习的训练过程分为 3 个步骤:①利用有标签的源域样本对模型进行预训练;②按照与源域的相关性大小,先对最相关的目标域进行适应,再逐渐过渡到次相关的目标域;③对所有目标域完成适应后微调模型,以优化其性能。

目标域适应又可细分为 3 个阶段:①选择目标域、挑选高置信度样本、更新源域数据;②源域与目标域的相关性通过目标域预测的不确定来衡量,即采用计算目标域的熵值作为选择目标域的标准,如式(16)所示;③高置信度样本的挑选规则如式(13)所示,挑选样本会被加入源域数据中以构成新的源域。数据更新过程如式(17)、式(18)所示。

$$H(T_m) = -E_{x_{T_m,j} \in T_m} \sum_{c=1}^C p(\hat{y}_{T_m,j} = c | x_{T_m,j}) \log p(\hat{y}_{T_m,j} = c | x_{T_m,j}) \quad (16)$$

$$\hat{S}^{q+1} = \hat{S}^q \cup D_m \quad (17)$$

$$\hat{T}^{q+1} = \hat{T}^q \setminus T_m \quad (18)$$

式中: q 代表第 q 步, $q \in (1, M)$ 。

SECL 模型的总损失为:

$$\mathcal{L}_{total} = \mathcal{L}_{mlp} + \mathcal{L}_{adv} + \mathcal{L}_{gcn} + \gamma \mathcal{L}_{lm} \quad (19)$$

式中: γ 为损失函数 $\mathcal{L}_{lm}$ 的权重。

3 实验结果与分析

3.1 实验环境

本文实验基于 PyTorch 框架实现,在 GPU(NVIDIA GeForce RTX 4090)上运行,使用的 PyTorch 版本为 1.12.1, CUDA 版本为 11.3。在模型训练过程中,使用在 ImageNet 数据集上预训练得到的 ResNet 模型作为特征提取器,并采用 Adam 优化器^[18]。

3.2 数据集

本文为了验证模型的优势,选择 Office-31 数据集、Office-Home 数据集和 ImageCLEF-DA 数据集进行实验^[19-21]。每个数据集包含多个子数据集,每个子数据集对应一个不同领域。具体而言,Office-31 数据集可细分为以下 3 个领域:①Amazon(A)、Dslr(D)和 Webcam(W);②Office-Home 数据集则包含 Art(Ar)、Clipart(Cl)、Product(Pr)和 Real-world(Rw);③4ImageCLEF-DA 数据集涵盖 Bing(B)、Caltech256(C)、ImageNet ILSVRC 2012(I)和 Pascal VOC(P)。

本文依次选取数据集中的一个子数据集作为源域,其

余子数据集作为目标域进行实验。例如,Office-31 数据集包含3个迁移子任务: $A \rightarrow W + D$ 、 $W \rightarrow A + D$ 和 $D \rightarrow A + W$ 。模型性能的优劣通过计算所有子任务上的平均分类准确率进行评估,如表 1 所示。

Table 1 Statistical information of datasets
表 1 数据集统计信息

数据集	领域	特征维度	类别数	每个域的图像数量	图像总量
Office-31	A; D; W	2048(ResNet-50)	31	2 817; 498; 795	4 110
Office-Home	Ar; Cl; Pr; Rw	2048(ResNet-50)	65	2 427; 4 365; 4 439; 4 357	15 588
ImageCLEF-DA	B; C; I; P	2048(ResNet-50)	12	600; 600; 600; 600	2 400

3.3 比较实验

为了全面评估模型性能,将多种模型在 Office-31 数据集和 Office-Home 数据集上针对单目标域自适应、混合目标域自适应和多目标域自适应等 3 个场景分别进行实验,实验结果如表 2 所示。其中,单目标域自适应涉及的是从单一源域向单一目标域的知识迁移过程;混合目标域自适

应则是将知识从单一源域迁移到由多个目标域合并形成的复合域。实验表明,在混合目标域自适应场景下,SECL 相较于模型 CGCT 在 Office-Home、Office-31 数据集上分别提升 0.4%、0.3%;在多目标域自适应场景下,SECL 相较于模型 D-CGCT 在 Office-Home、Office-31 数据集上分别提升 2.9% 和 1.1%。

Table 2 Performance of various models in different scenarios
表 2 不同场景下各模型的性能

场景设置	模型	Office-Home					Office-31			
		Ar	Cl	Pr	Rw	Avg	A	W	D	Avg
单目标域 自适应	DAN ^[23]	56.1	54.2	51.7	63.0	56.3	79.5	80.3	81.2	80.4
	CDAN ^[24]	64.2	62.9	59.9	68.1	63.8	91.4	84.1	84.0	86.6
	DANN ^[16]	58.3	55.5	52.8	63.9	57.6	80.9	82.8	81.5	81.7
	SECL	68.3	69.5	62.7	71.0	67.9	89.5	85.2	86.9	87.2
混合目标 域自适应	DAN ^[23]	55.6	56.6	48.5	56.7	54.4	78.0	64.4	66.7	69.7
	JAN ^[25]	58.3	60.5	52.2	57.5	57.1	84.2	74.4	72.0	76.9
	CDAN ^[24]	59.5	61.0	54.7	62.9	59.5	93.6	80.5	81.3	85.1
	AMEAN ^[26]	64.3	65.5	59.5	66.7	64.0	90.1	77.0	73.4	80.2
	CGCT ^[8]	67.4	68.1	61.6	68.7	66.5	93.9	85.1	85.6	88.2
	SECL	67.1	68.4	62.4	69.6	66.9	94.1	85.5	86.0	88.5
	MT-MTDA ^[12]	64.6	66.4	59.2	67.1	64.3	87.9	83.7	84.0	85.2
	D-CGCT ^[8]	70.5	71.6	66.0	71.2	69.8	93.4	86.0	87.1	88.8
多目标域 自适应	AMDA ^[14]	57.8	59.4	53.9	62.4	58.4	-	-	-	-
	DCML ^[11]	72.1	74.2	69.3	72.2	72.0	94.2	86.5	87.4	89.4
	MTDA-PDA ^[13]	71.1	73.0	67.5	72.1	70.9	92.9	87.3	86.0	88.7
	SECL	73.6	73.6	69.6	74.1	72.7	93.8	88.5	87.4	89.9

综上,SECL 在处理不同场景、不同分布差异和不同任务时的性能更优,进一步证明了 SECL 具有较强的泛化能力。此外,本文将 SECL 与其他 MTDA 模型在 ImageCLEF-DA 数据集上进行试验,如表 3 所示。由此可知,所提模型相较于其他 MTDA 模型在性能上存在一定优势,相较于 D-CGCT、CDAN 分别提升 1%、2.3%,进一步验证了所提方法的有效性。

Table 3 Comparison results of various models on the ImageCLEF-DA dataset
表 3 在 ImageCLEF-DA 数据集上各模型比较结果

场景	模型	ImageCLEF-DA				
		B	C	I	P	Avg
多目标域 自适应	CDAN ^[24]	86.1	71.1	77.3	81.9	79.1
	MCC ^[27]	86.2	76.0	78.6	82.6	80.8
	DAN ^[23]	81.8	71.7	77.2	77.8	77.1
	D-CGCT ^[8]	85.7	76.1	78.8.	81.2	80.4
	SECL	87.0	76.9	78.8	83.0	81.4

3.4 参数敏感性分析

为了评估模型一些关键参数对模型性能的影响,对参数 ω 、 γ 进行敏感性分析,结果如图 2 所示。实验表明,当 $\omega = 1.2$ 、 $\gamma = 0.06$ 时,SECL 对多个目标域数据的平均分类准确率最高、性能最佳,模型不仅在单个目标域上表现出色,在面对多个目标域的自适应挑战时能有效平衡、提升在所有目标域上的平均分类准确率,性能更优。SECL 的相关超参数如表 4 所示。

Table 4 Settings of model hyperparameter
表 4 模型超参数设置

超参数名称	取值
Batch Size	32
学习率	1e-3
预训练阶段迭代次数	200(Office-31、ImageCLEF-DA)、500(Office-Home)
目标域适应阶段 迭代次数	3 000(Office-31、ImageCLEF-DA)、10 000(Office-Home)
微调阶段迭代次数	15 000

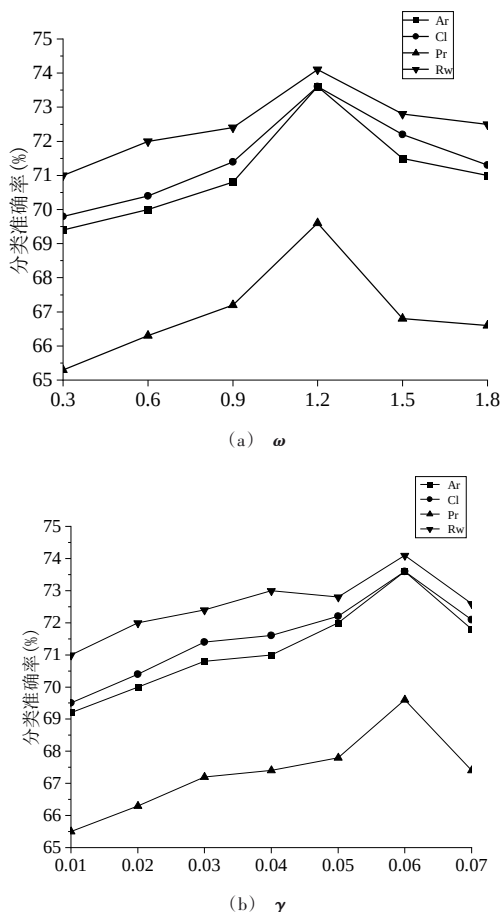


Fig. 2 Impact of different parameter values on the model on the Office-Home dataset

图2 在Office-Home数据集上不同参数取值对模型的影响

3.5 可视化分析

为了更直观地观察特征分布,利用t-SNE技术将D-CGCT和SECL生成的特征分布可视化,如图3所示(彩图扫OSID码可见,下同),每个点代表一个样本,不同颜色代表不同域^[22]。由此可见,SECL生成的t-SNE图的特征分布情况更优,特征空间中的决策边界清晰,类别之间的重叠明显减少,证明了SECL在多目标域自适应任务中能有效提取特征,这些特征表示不仅高度保留了类别信息,还显著提升了不同类别间的区分性,进一步验证了所提方法的有效性。

3.6 消融实验

为了验证MixStyle风格嵌入模块、双分类器协同学习机制及大间隔分类模块的有效性,在Office-Home数据集上分别对这3个部分进行消融实验,结果如表5所示。由此可知,采用双分类器机制相较于单一MLP分类器,模型的平均分类准确率提升6.3%;大间隔分类模块进一步提升了模型性能,平均分类准确率提升1.2%;引入MixStyle风格嵌入模块后,模型平均分类准确率再上升1.7%。

同时,为了验证课程学习训练策略的有效性,在迁移子任务Cl → Ar + Pr + Rw上比较了不同适应顺序的影响,

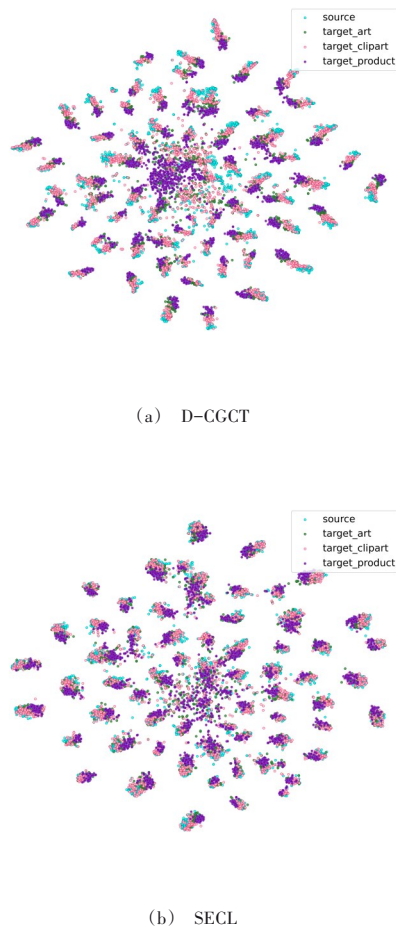


Fig. 3 t-SNE of D-CGCT and SECL under the same task

图3 D-CGCT与SECL在同一任务下的t-SNE

结果如表6所示。由此可知,由易到难的适应顺序相较于由难到易的适应顺序,模型平均分类准确率提升3.3%;由易到难的适应顺序相较于随机顺序1、2分别提升2.7%、1.8%,证明了课程学习训练策略的优越性。

Table 5 Results of ablation experiments

表5 消融实验结果

模型	$\mathcal{L}_{mlp}$	$\mathcal{L}_{gcn}$	$\mathcal{L}_{lm}$	MixStyle	Avg/%
	√	×	×	×	63.5
	√	√	×	×	69.8
	√	√	√	×	71.0
	√	√	√	√	72.7

Table 6 Experimental results of different adaptation sequences

表6 不同适应顺序的实验结果

子任务:	适应策略	适应顺序	Avg/%
Cl → Ar + Pr + Rw			
	从复杂任务开始	Ar; Pr; Rw	70.3
	随机顺序1	Ar; Rw; Pr	70.9
	随机顺序2	Rw; Ar; Pr	71.8
	从简单任务开始	Rw; Pr; Ar	73.6

4 结语

本文提出基于风格嵌入和协同学习的多目标域自适应模型SECL。首先,在特征提取器中添加Mixstyle风格嵌入模块,使新合成的图像在保留源域图像语义信息的情况下融入目标域与风格信息以增强信息;其次,采用双分类器协同学习机制降低噪声伪标签的影响,并引入大间隔分类损失优化不同类别之间的数据分布,提升了模型泛化能力;最后,采用课程学习的训练策略引导模型先适应容易的领域,再逐步适应较难的领域,使目标域适应过程更平滑。

在3个基准数据集上的实验表明,SECL在MTDA场景下具有一定的优越性和有效性,t-SNE图的特征分布情况更优,特征空间中的决策边界更清晰,类别间的重叠效果明显减少。同时,这些特征表示不仅高度保留了类别信息,还显著提升了不同类别间的区分性,进一步证实了所提方法的有效性。

参考文献:

- [1] CHEN W J, LIN L J, YANG S C, et al. Self-supervised noisy label learning for source-free unsupervised domain adaptation[C]// IEEE International Conference on Intelligent Robots and Systems, 2022: 10185–10192.
- [2] ZOU Z X, CHEN K Y, SHI Z W, et al. Object detection in 20 years: a survey[J]. *Proceedings of the IEEE*, 2023, 111(3): 257–276.
- [3] LIU Y A, ZHANG W, WANG J. Source-free domain adaptation for semantic segmentation[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2021: 1215–1224.
- [4] GUAN H, LIU M J. Domain adaptation for medical image analysis: a survey[J]. *IEEE Transactions on Biomedical Engineering*, 2021, 69(3): 1173–1185.
- [5] ZHUANG F Z, QI Z Y, DUAN K Y, et al. A comprehensive survey on transfer learning[J]. *Proceedings of the IEEE*, 2020, 109(1): 43–76.
- [6] ISOBE T, JIA X, CHEN S J, et al. Multi-target domain adaptation with collaborative consistency learning[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2021: 8187–8196.
- [7] TIAN Q, LU Z H, ZHOU J Z. Progressive adversarial learning for multi-target domain adaptation[J]. *Neural Processing Letters*, 2023, 55(9): 12239–12253.
- [8] ROY S, KRIVOSHEEV E, ZHONG Z, et al. Curriculum graph co-teaching for multi-target domain adaptation[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2021: 5351–5360.
- [9] HONG D F, GAO L R, YAO J, et al. Graph convolutional networks for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 59(7): 5966–5978.
- [10] TRANHEDEN W, OLSSON V, PINTO J, et al. DACS: domain adaptation via cross-domain mixed sampling[C]// *Proceedings of IEEE Winter Conference on Applications of Computer Vision*, 2021: 1379–1389.
- [11] WANG J, ZHONG C L, FENG C, et al. Discriminative mutual learning for multi-target domain adaptation[C]// *26th International Conference on Pattern Recognition*, 2022: 2900–2906.
- [12] NGUYEN-MEIDINE L T, BELAL A, KIRAN M, et al. Unsupervised multi-target domain adaptation through knowledge distillation[C]// *Proceedings of IEEE Winter Conference on Applications of Computer Vision*, 2021: 1339–1347.
- [13] ZHOU J Z, TIAN Q, LU Z H. Progressive decoupled target-into-source multi-target domain adaptation[J]. *Information Sciences*, 2023, 634: 140–156.
- [14] WANG Y X, ZHANG Z X, HAO W L, et al. Attention guided multiple source and target domain adaptation[J]. *IEEE Transactions on Image Processing*, 2021, 30: 892–906.
- [15] YANG X, DENG C, LIU T L, et al. Heterogeneous graph attention network for unsupervised multiple-target domain adaptation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(4): 1992–2003.
- [16] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. *Journal of Machine Learning Research*, 2016, 17(59): 1–35.
- [17] LING J, XUE H, SONG L, et al. Region-aware adaptive instance normalization for image harmonization[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2021: 9361–9370.
- [18] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770–778.
- [19] SAENKO K, KULIS B, FRITZ M, et al. Adapting visual category models to new domains[C]// *Proceedings of European Conference on Computer Vision*, 2010: 213–226.
- [20] VENKATESWARA H, EUSEBIO J, CHAKRABORTY S, et al. Deep hashing network for unsupervised domain adaptation[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 5018–5027.
- [21] CAPUTO B, MÜLLER H, MARTINEZGOMEZ J, et al. ImageCLEF 2014: overview and analysis of the results[C]// *Proceedings of International Conference of the Cross-Language Evaluation Forum for European Languages*, 2014: 192–211.
- [22] POLIČAR P G, STRAŽAR M, ZUPAN B. Embedding to reference t-SNE space addresses batch effects in single-cell classification[J]. *Machine Learning*, 2023, 112(2): 721–740.
- [23] LONG M S, CAO Y, WANG J M, et al. Learning transferable features with deep adaptation networks [DB/OL]. <https://arxiv.org/abs/1502.02791>.
- [24] LONG M S, CAO Y, WANG J M, et al. Conditional adversarial domain adaptation[DB/OL]. <https://arxiv.org/abs/1705.10667>.
- [25] LONG M S, CAO Y, WANG J M, et al. Deep transfer learning with joint adaptation networks[DB/OL]. <https://arxiv.org/abs/1605.06636>.
- [26] CHEN Z L, ZHUANG J Y, LIANG X D, et al. Blending-target domain adaptation by adversarial meta-adaptation networks[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2019: 2248–2257.
- [27] JIN Y, WANG X M, LONG M S, et al. Minimum class confusion for versatile domain adaptation[C]// *Proceedings of European Conference on Computer Vision*, 2020: 464–480.

(责任编辑:刘嘉文)