

DOI:10.19651/j.cnki.emt.2416351

# 基于类别感知与重加权的多源域自适应算法<sup>\*</sup>

谭棉<sup>1,2</sup> 李志玲<sup>1,2</sup> 陈望<sup>1,2</sup> 曾涛涛<sup>1,2</sup> 冯夫健<sup>1,2</sup>

(1. 贵州民族大学数据科学与信息工程学院 贵阳 550025;

2. 贵州民族大学贵州省模式识别与智能系统重点实验室 贵阳 550025)

**摘要:** 多源域自适应是迁移学习中的一个重要分支,类别偏移是多源域自适应领域的热点难题之一,其本质是源域和目标域类别分布不匹配的问题。针对此问题,提出了一种基于类别感知与重加权的多源域自适应算法,该算法通过类别感知策略增强相似类别间的正向迁移;同时,引入重加权矩匹配策略,减少不同层面的分布差异;此外,利用伪标签构建自适应权重,有效降低类别偏移的影响。在 Digits-five 和 Office-Caltech10 两个数据集上的任务分类准确率分别达到了 94.11% 和 97.18%,实验结果表明,所提算法相比于当前典型的多源域自适应算法在类别偏移场景下的准确性方面有显著提升。

**关键词:** 迁移学习;多源域自适应;类别感知;矩匹配;自适应加权

**中图分类号:** TP181;TN911.7 **文献标识码:** A **国家标准学科分类代码:** 520.2040

## Category-aware and reweighting-based multi-source domain adaptation algorithm

Tan Mian<sup>1,2</sup> Li Zhiling<sup>1,2</sup> Chen Wang<sup>1,2</sup> Zeng Taotao<sup>1,2</sup> Feng Fujian<sup>1,2</sup>

(1. School of Mathematical Sciences and Information Engineering, Guizhou Minzu University, Guiyang 550025, China;

2. Guizhou Minzu University, Key Laboratory of Pattern Recognition and Intelligent System, Guiyang 550025, China)

**Abstract:** Multi-source domain adaptation is an important branch of transfer learning. Category shift, a prominent challenge in this field, stems from the mismatch between category distributions in the source and target domains. To address this problem, a category-aware and reweighting-based multi-source domain adaptation algorithm is proposed. The algorithm enhances positive transfer between similar categories through a category-aware strategy and introduces a reweighting moment matching strategy to reduce distribution differences at various levels. Additionally, adaptive weights are constructed using pseudo-labels to effectively mitigate the impact of category shift. Experimental results on the Digits-Five and Office-Caltech10 datasets show that the proposed algorithm achieves classification accuracies of 94.11% and 97.18%, respectively. These results indicate that the proposed algorithm significantly improves accuracy in scenarios with category shift compared to current typical multi-source domain adaptation algorithms.

**Keywords:** transfer learning; multi-source domain adaptation; category awareness; moment matching; adaptive weighting

## 0 引言

迁移学习作为解决不同领域间知识迁移问题的重要技术,近年来受到了广泛关注。特别是在多源域环境下,由于不同源域之间存在显著的类别不一致性,类别偏移问题变得尤为突出。类别偏移指的是源域与目标域在类别分布上

的不匹配,这一问题限制了传统单源域自适应方法<sup>[1-6]</sup>在准确性方面的效果。因此,国内外学界开展了丰富的研究,通过多源域自适应技术<sup>[7-9]</sup>以更好地解决类别偏移问题,确保知识的有效迁移。

在传统域自适应方面,域自适应的核心是减小源域和目标域在分布差异,源域和目标域之间存在分布差异也归

收稿日期:2024-07-02

\* 基金项目:贵州省科技计划项目(黔科合基础-ZK[2022]一般195,黔科合基础-ZK[2023]一般143,黔科合平台人才-ZCKJ[2021]007,黔科合基础-ZK[2022]一般197),贵州省教育厅自然科学研究项目(黔教技[2023]061号,黔教技[2023]012号,黔教技[2022]015号,黔教技[2022]047号,黔教技[2024]63号),贵州省模式识别与智能系统重点实验室开放课题(GZMUKL[2022]KF01)、贵州民族大学博士科研启动项目(GZMUKL[2024]QD04)资助

由于分布不同的域之间存在的域偏移现象<sup>[10]</sup>,研究域自适应的重点在于充分挖掘并利用源域和目标域之间共同的域不变特性,使从源域中提取到的特征信息发挥更大作用,进而消除或减少域偏移的影响,并从已标记的源域学习到一个能很好地推广到不同但相关的目标域模型,从而实现目标样本的准确预测<sup>[11]</sup>。早期的研究大多属于单源域自适应,主要分为 3 类方法:基于分歧的方法:Long 等<sup>[12]</sup>提出了一种联合适应网络的深度迁移学习方法,侧重于通过对齐特定领域的数据分布和分类器,从而解决在迁移学习中常见的误分类问题。Wang 等<sup>[13]</sup>提出了一种基于流形嵌入分布对齐的视觉领域适应方法,通过在流形上嵌入源和目标域的分布,从而对齐两个领域,以减少在视觉任务中的错误分类。Sun 等<sup>[14]</sup>通过简化的适应策略,有效地对齐了源域和目标域的分布。Zhang 等<sup>[15]</sup>提出了一种再生核希尔伯特空间中看齐无限维协方差矩阵的领域适应方法,通过在核空间中嵌入源和目标域的分布,实现两个领域的对齐。对抗学习方法:Hoffman 等<sup>[16]</sup>采用极大极小博弈在特征提取器和判别器之间实现源域和目标域的对齐,但这种方法通常专注于数据分布的对齐而忽略了目标域的判别性学习。基于一致性的方法:Kumar 等<sup>[17]</sup>通过将决策边界定位在低密度区域来学习更稳定且具有区分力的特征,并单独对齐各个特征空间中的源域与目标域分布。上述研究主要应用于存在一个源域的场景下,减少域偏移从而实现域自适应。相对而言,针对域自适应中存在多个源域时,这种单一性使得在处理复杂的目标域任务时表现不佳,导致在源域样本来源的多样性和数量上存在不足。

在多源域自适应方面,多源域自适应有效地整合了多种源域,缓解了单源域自适应中源域样本来源的单一性和多样性不足的问题。然而,这也带来了更多的挑战,主要包括大量的域偏移和不同源域对目标域任务的不同贡献。针对大量域偏移的问题,Zhao 等<sup>[18]</sup>通过对抗学习在多个域间学习不变性和区分性特征表示。文献[8]则使用矩匹配来对齐源域与目标域之间的数据分布。然而,这些方法往往平等对待所有源域,忽略了它们对目标任务的不同贡献。针对不同源域对目标域任务具有不同贡献的问题,文献[7]通过计算域相关性来最小化对抗学习中的域偏移,而 Zhu 等<sup>[19]</sup>提出多特征空间自适应网络来对齐特定领域的分布和分类器,以减少错误分类。Zhang 等<sup>[20]</sup>提出了多源选择转移方法,通过不同的选择策略,如最近邻样本选择和加权选择,来选取与目标域更为相似的源域。综上所述,上述方法在解决多源域自适应中减少域偏移以及不同源域对于目标任务贡献存在差异方面取得了一定进展,但大多数研究仍基于源域与目标域之间完全共享类别的假设进行,而针对存在类别偏移的多源域自适应的研究仍存在不足。这种类别偏移导致源域与目标域标记空间不匹配和边缘分布差异,从而影响整体的域自适应性能。因此,有必要进一步研究多源域自适应方法,以更有效地解决类别偏移问题,提升

域自适应性能。

针对类别偏移的问题,Zuo 等<sup>[21]</sup>提出了一种基于注意力的多源域自适应算法,通过考虑域间的相关性,对相似度较高的源域进行加权。以约束特征提取器能够生成对齐和区分性的视觉表示。但该方法仅从域级别层面来消除类别偏移带来的负迁移。Wang 等<sup>[22]</sup>提出了类感知的样本重加权方法,结合最优传输理论,能够更好地处理多源域的异质性,减少源域与目标域之间的分布差异,实现多源和目标之间的样本级细粒度对齐。但最优传输计算通常是计算量较大的操作,特别是在多源域和高维数据情况下,计算开销较大。邱春红等<sup>[23]</sup>提出了一种基于同伴辅助学习分类器的部分域自适应方法,使得源共享域和目标域的特征分布在特征空间中是一致的,减轻了负迁移。

在上述研究的基础上,本文提出了一种基于类别感知与重加权的多源域自适应算法(category-aware and reweighting-based multi-source domain adaptation algorithm,AWMD-MSDA)。该算法旨在通过识别与目标域相关的源域类别并强化该类别的迁移,同时减少不相关类别的干扰。此外,通过一阶和二阶矩匹配策略对源域和目标域进行有效对齐,并采用自适应权重策略来优化各源域对目标域的贡献,从而提高分类的精度和适应性。该综合策略确保了算法在多源域自适应的环境下更精确地处理类别偏移问题,提高了目标域的适应性和分类性能。

总而言之,本文的学术贡献主要如下:

- 1) 提出了一种类别感知策略,选择性地加强与目标域相关的类别迁移,有效应对类别偏移;
- 2) 提出了一种重加权矩阵匹配方法,对源域和目标域的一阶和二阶统计矩进行对齐,显著减少了分布差异;
- 3) 提出了一种自适应权重机制,根据各源域的相关性动态调整其影响,优化整体适应性。

## 1 问题定义和理论分析

### 1.1 问题定义

多源域自适应中,一个重要的挑战是类别偏移问题,其中目标域与各源域只共享部分类别,而每个源域拥有其特有的类别,这种设置更贴近实际应用场景。该问题的表示形式如图 1 所示,其中左侧展示了 3 个源域的样本分布,右侧显示了目标域的样本分布。不同的形状代表不同类别的样本,其中源域 1 的特有类别用“正方形”表示;源域 2 的特有类别用“十字星”表示;源域 3 的特有类别用“五角星”表示。共享类别如“菱形”、“三角形”、“桃心”和“圆形”分别表示源域间的共享类别。颜色的变化则表示来自不同源域的样本,目标域类别是所有源域类别的一个子集。

在数据表示和采样中,假设多源域自适应由  $N$  个源域  $D_S = \{D_1, D_2, \dots, D_N\}$  和目标域  $D_T$  组成,数据分布分别是  $\{P_{s_i} = (x, y)\}_{i=1}^N$  和  $P_t = (x, y)$ 。多源域自适应从  $N$  个源分布中对训练数据  $\{(X_{s_i}, Y_{s_i})\}_{i=1}^N$  进行采样,其中  $X_{s_i} =$

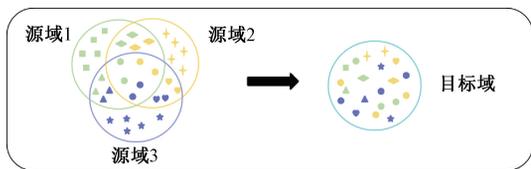


图 1 多源域自适应中的类别偏移问题示意图

Fig. 1 Illustration of the class shift problem in multi-source domain adaptation

$\{x_{si}^j\}_{j=1}^{N_{si}}$  表示第  $i$  个源域的图像,  $Y_{si} = \{y_{si}^j\}_{j=1}^{N_{si}}$  表示对应的标签。相应地, 从目标分布中对未标记的目标数据  $X_t = \{x_t^j\}_{j=1}^{N_t}$  进行采样, 其标记为  $Y_t$ 。目标域类别是所有源域类别并集的一个子集。与单源域自适应类似,  $Y_t$  在训练中不可用, 仅用于评估。因此, 多源域自适应皆在识别给定训练数据  $\{(X_{si}, Y_{si})\}_{i=1}^N$  的  $X_t$  中包含的图像。

## 1.2 理论分析

为了有效解决多源域自适应中的类别偏移问题, 本文提出了基于类别感知与重加权策略的多源域自适应算法 (AWMD-MSDA)。以下将从类别感知策略和重加权矩匹配策略两个方面进行详细的理论分析。

类别感知策略的核心思想是在多源域自适应过程中, 识别并选择性地增强与目标域相关的类别, 同时减少不相关类别的干扰。其理论依据主要基于以下几点: 首先, 类别偏移问题的本质是源域和目标域之间的类别分布不一致, 具体表现为某些源域类别在目标域中并不存在, 或者某些目标域类别在源域中没有对应的样本。传统的多源域自适应方法通常假设源域和目标域共享相同的类别空间, 这在实际应用中并不总是成立。因此, 需要一种能够区分相关和不相关类别的策略, 以避免不相关类别对迁移过程的干扰。其次, 与目标域相关的类别在迁移过程中具有更高的价值, 因为这些类别的特征对目标任务的贡献更大。通过识别这些类别并增加其权重, 可以确保这些重要特征在迁移过程中得到充分利用, 从而提高分类的精度和适应性。最后, 不相关类别的存在会对迁移过程产生干扰, 降低模型的整体性能。通过识别并减少这些不相关类别的影响, 可以减少噪声对模型训练的干扰, 从而提高模型的稳定性和泛化能力。

重加权矩匹配策略的核心思想是通过一阶和二阶统计矩的重加权处理, 对源域和目标域进行更为精确的对齐。其理论依据主要包括以下几点: 首先, 源域和目标域之间的分布差异是导致迁移学习性能下降的主要原因之一。为了实现有效的迁移, 必须尽可能减小这种分布差异。传统的分布对齐方法通常仅考虑一阶统计矩的对齐, 而忽略了高阶统计矩的差异, 这可能导致对齐效果不佳。其次, 高阶统计矩包含了更多关于数据分布的结构信息, 这对于实现更加精确的分布对齐至关重要。因此, 通过同时对齐一阶和二阶统计矩, 可以更全面地减少源域和目标域之间的分

布差异。最后, 在多源域自适应中, 不同源域对目标域的贡献可能不同。如果简单地对所有源域进行平均处理, 可能无法充分利用每个源域的特定信息。通过引入重加权处理, 可以根据源域与目标域的相关性, 动态调整各源域在分布对齐过程中的权重, 从而优化整体的适应性。基于上述理论依据, 重加权矩匹配策略通过计算源域和目标域在特征空间中的一阶和二阶矩, 并对其进行重加权处理, 以减少分布差异, 从而提高模型分类精度和适应性。

## 2 基于类别感知与重加权的多源域自适应算法

### 2.1 类别感知策略

在多源域自适应中, 面临的核心挑战是如何有效地实现目标域与多个源域之间的对齐, 尤其是在存在显著的域偏移问题时。假设通过对源域数据输出的特征进行感知能够识别各源域中不同类别的样本和源域之间共享类别的样本, 实现源域与目标域之间更紧密的对齐。因此, 本文提出了一种类别感知策略, 该策略专注于识别并加强目标域与源域中相似类别的联系, 以促进正向迁移, 并抑制不相似类别的负迁移, 从而有效减少类别偏移所造成的负面影响。

类别偏移是由不同源域中类别的差异造成的, 这些差异会影响到模型的迁移效果。具体来说, 假设有多个源域, 每个源域包含一组特定的类别。将这些源域的类别集合定义为  $Q_i$ , 并考虑将  $Q_i$  与目标域的类别集合  $Q_t$  进行对齐。因此类别偏移可以定义为源域和目标域之间类别集合的不一致性。

针对类别偏移问题, 所提出的类别感知策略旨在通过以下步骤进行优化:

1) 识别并强化源域和目标域之间共有的类别, 以利用这些类别的数据增强模型的迁移学习;

2) 通过减少源域中目标域不存在类别的依赖, 降低不相关信息的干扰, 从而减轻负迁移影响。

鉴于每个源域的数据分布可能不同, 将具有相同标识的类别视为不同的类别。基于这个假设, 重新定义源域的标签:

$$\hat{Y}_{si} = Y_{si} + Q_i \quad (1)$$

其中,  $Q$  是类别的数量, 给定所有源域图像  $X_{si} (i \in [1, N])$  沿着新标签  $\hat{Y}_{si}$ , 训练类别感知  $A$ , 训练过程中的类别的识别损失函数如下:

$$\mathcal{L}_{class} = -E [\hat{Y}_{si}^T \ln Q(X_{si})] \quad (2)$$

由于类别感知  $A$  是基于新的标签  $\hat{Y}_{si}$  推断的, 因此它不仅区分每个类别, 而且能够识别领域。因此,  $A$  用于基于一批图像计算类别相关性。一旦模型在标记的源域数据上训练完成, 就可以通过  $A$  的参数对源域的数据分布进行建模。假设每批包含  $M$  个目标图像,  $A$  用于预测属于每个类别的概率, 从而获得的类别相关性向量  $w^c = \{\omega_1^c, \omega_2^c, \dots\}$

$\dots, \omega_N^c \}$ 。

## 2.2 重加权矩匹配策略

在类别感知策略的基础上,进一步减少类别偏移对域对齐的影响是至关重要的。一个代表性的方法是多源域自适应矩匹配(M3SDA)<sup>[7]</sup>,该方法通过最小化源域与目标域之间的矩距离来学习域不变特征。但没有考虑到不同源域对目标域分类任务的贡献度差异,未能有效区分各源域对目标域任务的具体贡献差异。为了解决这一问题,本文提出了一种重加权矩匹配策略,通过计算不同源域对目标域分类任务的贡献度,以优化域对齐的精确性和效果。

具体而言,通过利用类别感知  $A$  预测的每个类别的概率,得到源域加权集成的目标域样本类别标签预测结果,并据此获得不同的贡献度。重加权矩匹配从两个角度对特征分布进行对齐:一是不同源域之间的特征分布对齐,二是源域与目标域之间的特征分布对齐。针对于不同源域之间的分布差异,构建如下形式的损失函数:

$$\begin{aligned} \mathcal{L}_{ss-awmd} &= \frac{2}{N \cdot (N-1)} \sum_{n_1=1}^{N-1} \sum_{n_2=n_1+1}^N AWMD(D_{s_{n_1}}, D_{s_{n_2}}) = \\ &= \frac{2}{N \cdot (N-1)} \sum_{n_1=1}^{N-1} \sum_{n_2=n_1+1}^N (\omega_i (\|E(G(X_{s_{n_1}})^1) - \\ &E(G(X_{s_{n_2}})^1)\|_2 + \|E(G(X_{s_{n_1}})^2) - E(G(X_{s_{n_2}})^2)\|_2)) \end{aligned} \quad (3)$$

其中,  $E$  表示各域的平均运算,  $G(X_{s_i})^1$  和  $G(X_{s_i})^2$  分别表示第  $s_i$  个源域的一阶和二阶特征,  $D_{s_{n_1}}$  和  $D_{s_{n_2}}$  表示对应源域上的处理样本,  $\omega_i$  表示贡献度  $\omega$  的第  $i$  个分量,  $\|\cdot\|_2$  表示空间的 2 范数。

此外,针对于不同源域和目标域之间的分布差异,构建如下形式的损失函数:

$$\begin{aligned} \mathcal{L}_{st-awmd} &= \frac{2}{N \cdot N} \sum_{n_1=1}^N \sum_{n_2=1}^N AWMD(D_{s_{n_1}}, D_{t_{n_2}}) = \\ &= \frac{2}{N \cdot N} \sum_{n_1=1}^N \sum_{n_2=1}^N (\omega_i (\|E(G(X_{s_{n_1}})^1) - E(G(X_{t_{n_2}})^1)\|_2 + \\ &\|E(G(X_{s_{n_1}})^2) - E(G(X_{t_{n_2}})^2)\|_2)) \end{aligned} \quad (4)$$

其中,  $E$  表示各域的平均运算,  $G(X_{t_i})^1$  和  $G(X_{t_i})^2$  分别表示目标域的一阶特征和二阶特征。  $D_{s_{n_1}}$  表示对应源域上的处理样本,  $D_{t_{n_2}}$  表示目标域上的样本,  $\omega_i$  表示贡献度  $\omega$  的第  $i$  个分量,  $\|\cdot\|_2$  表示空间的 2 范数。

综合以上特征对齐函数,可以得到如下的域级综合对齐损失函数:

$$\mathcal{L}_{awmd} = \mathcal{L}_{ss-awmd} + \mathcal{L}_{st-awmd} \quad (5)$$

第一项用于约束任意两个源域对齐,第二项用于减小每个源域和目标域之间的域偏移。与传统的矩距离的相比,所提出的重加权矩匹配具有以下优势。首先,它分别对源域间及源域与目标域间的特征分布进行对齐;其次,它通过优先考虑高贡献度的源域与目标域进行对齐,以减少不相似域带来的负面影响。

最后,通过衡量不同源域与目标域间的边缘分布差异来确定源域的贡献度,并对所得的  $N$  个贡献度进行归一化处理,由此,可以得到第  $n$  个源域对目标域的自适应贡献度权重  $\omega_i$  如下:

$$\omega_i = \frac{\omega_a^i}{\sum_{i'=1}^N \omega_a^{i'}} \quad (6)$$

$$\omega_a^i = \frac{1}{AWMD(D_{s_n}, D_t)} \quad (7)$$

其中,  $\omega_a^i$  表示对  $N$  个贡献度进行归一化的处理,  $AWMD(D_{s_n}, D_t)$  表示第  $n$  个源域和目标域在特征空间中的分布重加权矩距离。

## 2.3 基于类别感知与重加权的多源域自适应算法框架

在多源域自适应领域中,类别感知策略和重加权矩匹配策略共同构成了对源域与目标域之间的有效对齐的核心方法论。类别感知策略通过强化源域和目标域之间共有类别的联系,优化了正向迁移并减轻了由类别偏移带来的负面影响。进一步地,重加权矩匹配策略通过计算不同源域对目标域分类任务的贡献度,并在源域间及源域与目标域间进行特征分布的对齐,从而提高了域对齐的精确性和效果。这两种策略的结合,不仅优化了数据的迁移学习过程,也为多源域自适应提供了一个更为精细和高效的框架。借助这两种策略的协同作用,本文进一步提出了一个基于类别感知与重加权的多源域自适应算法,该算法在本节将详细介绍。

所提出算法的构成部分,包括特征提取、类别感知、重加权矩匹配和分类器等关键组件。该算法的模型架构如图 2 所示,其中  $S_1$  到  $S_N$  表示第 1 个到第  $N$  个源域,  $T$  表示目标域。  $G$  是对源域和目标域图像进行特征提取的共享特征提取器。  $A$  表示类别感知,  $\mathcal{L}_{class}$  表示实现类别感知的模型训练损失。  $W$  表示重加权矩匹配,  $\omega_i$  表示自适应权重机制中的自适应贡献度权重,  $\mathcal{L}_{ss-awmd}$  表示实现不同源域间的特征对齐的模型训练损失,  $\mathcal{L}_{st-awmd}$  表示实现不同源域

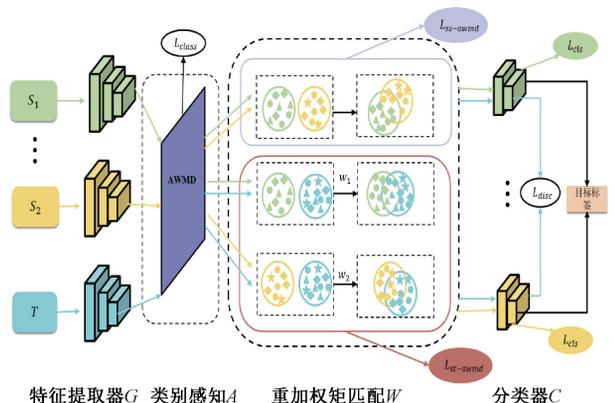


图 2 基于类别感知与重加权的多源域自适应模型结构图  
Fig. 2 Structure diagram of the multi-source domain adaptation model based on class awareness and re-weighting

与目标域之间的特征对齐的模型训练损失。 $C$  表示领域分类器,  $\mathcal{L}_{cls}$  表示域分类的交叉熵分类损失。整个多源域自适应过程通过最小化由上述损失项构成的总目标损失来实现。

为了更清晰地阐述所提 AWMD-MSDA 的完整优化过程, 算法的计算流程如算法 1:

算法 1 AWMD-MSDA 的时间复杂性主要取决于几个关键步骤: 数据采样、特征提取、损失计算及参数更新。整个算法是基于最大迭代次数  $K$  进行循环的, 每次迭代中处理  $m$  个图像。特征提取的复杂性依赖于所使用的模型细节, 假设为  $O(f)$ , 而损失计算和参数更新假设复杂性为  $O(c)$ 。因此, 每次迭代的复杂性为  $O(m \times (f + c))$ , 整个算法的总时间复杂性则为  $O(K \times m \times (f + c))$ 。这表明算法的运行时间与迭代次数、每次迭代中处理的图像数量以及特征提取与计算损失的复杂性成正比。

#### 算法 1: AWMD-MSDA 算法

输入: 有标签的源域样本, 无标签的目标域样本, 最大迭代次数  $K$ ;

输出: 目标域预测标签;

- 1: 从源域  $\{(X_{si}, Y_{si})\}_{i=1}^N$  中的一个随机采样  $m$  个图像  $\{(x_i^s, y_i^s)\}_{i=1}^m$ ;
- 2: 从目标域  $X_t$  采样  $m$  个图像  $\{X_i^t\}_{i=1}^m$ ;
- 3: 将源样本和目标样本馈送到特征提取器  $G$  以获得公共潜在表示  $G(x_i^s)$  和  $G(x_i^t)$ ;
- 4: 将源样本的公共潜在表示馈送到类别感知器以获得源样本的域特定表征;
- 5: 根据式(1)实现域特征的识别, 并计算损失  $\mathcal{L}_{class}$ ;
- 6: 根据式(2)、(6)和(7)构建并计算源域对目标域的自适应贡献权重  $\omega_i$ ;
- 7: 根据式(3)实现源域与源域之间的对齐, 并计算损失  $\mathcal{L}_{ss-awmd}$ ;
- 8: 根据式(4)实现源域与目标域之间的对齐, 并计算损失  $\mathcal{L}_{st-awmd}$ ;
- 9: 将源样本的域特定表征馈送到域特定分类器以获得  $C_j(G(X_i^s))$ , 并根据式(8)计算损失  $\mathcal{L}_{cls}$ ;
- 10: 根据式(9)计算并更新总体目标损失  $\mathcal{L}_{total}$ ;
- 11: 通过最小化式(9)来更新相应的模型参数;
- 12: 重复步骤 1~11 直至最大迭代数  $K$  及  $\mathcal{L}_{total}$  收敛。

## 2.4 目标函数

针对多源域自适应中的类别偏移问题, 本文提出的 AWMD-MSDA 算法综合考虑总体目标函数由类别感知损失函数、域级综合对齐损失函数和域分类器交叉熵损失函数组成, 其中的域分类器中的交叉熵损失函数的构建形式如下:

$$\mathcal{L}_{cls} = \sum_{j=1}^N E_{x \sim X_{sj}} J(C_j(G(X_i^s)), y_i^s) \quad (8)$$

其中,  $C$  表示是由  $N$  个域预测器  $\{C_j\}_{j=1}^N$  组成的多输出网络, 每个预测器  $C_j$  是 softmax 分类器, 并且在第  $j$  个源域的域特征提取器  $G(x)$  之后接收的不变特征,  $J(\cdot)$  表示对于每个分类器使用的交叉熵分类损失。

综合上述组成部分, AWMD-MSDA 的总体优化目标函数损失可表示为:

$$\mathcal{L}_{total} = \mathcal{L}_{class} + \mathcal{L}_{awmd} + \mathcal{L}_{cls} \quad (9)$$

其中,  $\mathcal{L}_{total}$  表示总体目标函数损失,  $\mathcal{L}_{class}$  表示类别感知损失,  $\mathcal{L}_{awmd}$  表示域级综合对齐损失,  $\mathcal{L}_{cls}$  表示域分类的交叉熵分类损失。对由上述损失项构成的总体目标损失最小优化可实现存在类别偏移场景下的多源域自适应。

## 3 实验与结果分析

### 3.1 数据集和实验设置

#### 1) 数据集

实验涉及两个广泛使用的域自适应数据集, 分别为 Digit-Five<sup>[24]</sup> 和 Office-Caltech10<sup>[25]</sup> 数据集, 具体介绍如下:

Digits-Five 是一个数字识别数据集, 包含了 5 个数字相关的数据集, 即 MNIST, MNIST-M, USPS, SVHN 和 Synthetic Digits。每个数据集由 0~9 的数字组成, 每个域中包含 10 个类别的样本, 对于每个域, 将图像大小调整为  $32 \times 32$ , 并将批量大小设置为 128 用于训练。

Office-Caltech10 是一个常用的领域适配的数据集, 它是 Caltech31 数据集扩展而来的。它来自 4 个不同领域的相同 10 个对象类别组成, 即 Webcam, DSLR, Caltech 和 Amazon。对于每个域, 将图像大小调整为  $128 \times 128$ , 并将批量大小设置为 64 用于训练。

#### 2) 实验设置

所有实验均使用 Pytorch 框架在 NVIDIA 3090 GPU 上进行, 训练过程中采用随机梯度下降(SGD)作为优化器, 权重衰减和动量分别设置为 0.0005 和 0.9, 初始学习率为 0.001。在实验过程中, 轮流选择一个域作为目标域, 其他域作为源域, 每个实验配置重复 5 次以确保结果的可靠性, 并计算平均分类精度。

本文分别设置 4 组实验:

第 1 组实验: AWMD-MSDA 与现有的域自适应方法对比分析。对比实验涵盖了单源域自适应方法和多源域自适应方法。

第 2 组实验: 类别感知策略的有效性分析。主要验证类别感知策略可以识别并加强目标域与源域之间相似类别的关联性。

第 3 组实验: 重加权矩匹配策略的有效性分析。主要验证重加权矩匹配策略可以有助于量化不同源域对目标域分类任务的贡献度差异, 以优化域对齐的精确性和效果。

第 4 组实验: AWMD-MSDA 中各组件的有效性分析。主要验证 AWMD-MSDA 算法中的每一个组成部分对最终性能有显著的正向贡献。

### 3.2 AWMD-MSDA 与现有域自适应方法对比分析

在这一节中,将 AWMD-MSDA 方法与现有的域自适应方法在 Digits-Five 和 Office-Caltech10 数据集上的表现进行了比较分析。对比方法包括了单源域自适应方法“Single Best”,它代表从所有源域中选择单一源域的最佳

性能,多源域自适应方法包括 JAN<sup>[12]</sup>、DCTN<sup>[19]</sup>、M3SDA<sup>[7]</sup>、ABMSDA<sup>[21]</sup>、CASR<sup>[22]</sup>、TSCDA<sup>[23]</sup>。在实验中,采用 ResNet-50 作为预训练的特征提取器,采用 Softmax 分类器作为初始化模型。实验结果如表 1 和 2 所示。

表 1 AWMD-MSDA 在 Digits-Five 上与现有域自适应方法对比分析

Table 1 Comparative analysis of AWMD-MSDA and existing domain adaptation methods on Digits-Five %

基准	方法	<i>mm</i>	<i>mt</i>	<i>up</i>	<i>sv</i>	<i>sy</i>	<i>Avg</i>
单源域	Single Best	63.37	90.50	88.71	63.54	82.44	77.71
	JAN	65.88	97.21	95.42	75.27	86.55	84.07
	DCTN	70.53	96.23	92.81	77.61	86.77	84.79
	M3SD A	69.76	98.58	95.23	78.56	87.56	85.94
多源域	ABMSDA	70.76	<b>98.98</b>	94.23	75.56	87.76	85.46
	CASR	89.99	98.70	<b>98.30</b>	85.95	<b>95.80</b>	93.75
	TSCDA	89.38	98.32	96.25	86.72	94.30	92.99
	AWMD-MSDA	<b>90.50</b>	98.02	97.82	<b>88.80</b>	95.40	<b>94.11</b>

表 1 显示了在 Digit-Five 数据集上进行的实验结果。其中用加粗的形式标记最优结果,“*mm*”表示以 MNIST-M 数据集作为目标域进行域自适应,“*mt*”表示 MNIST 数据集作为目标域进行域自适应,“*up*”表示 USPS 数据集作为目标域进行域自适应,“*sv*”表示 SVHN 数据集作为目标域进行域自适应以及“*sy*”表示 Synthetic Digits 数据集作为目标域进行域自适应。

实验结果显示,AWMD-MSDA 在 Digits-Five 数据集上平均精度高于其他现有域自适应方法。一方面,相比于 Single Best 方法的 77.71%,AWMD-MSDA 的效果提高了 16.40%,该结果表明,多源域自适应能够缓解单源域自适应中源域样本来源的单一性和多样性不足的问题,可以有效提高模型的分类效果。另一方面,AWMD-MSDA 的效果超过了现有的多源域自适应方法的表现。如表 1 所示,AWMD-MSDA 比现有的多源域自适应方法高 6.28%。在 Digits-Five 数据集中,SVHN(*sv*)数据集作为目标域进行域自适应是一个比较难的过程。如表 1 中结果所示,现有的域自适应方法对此任务的效果均比较差。但是,通过对特征提取后的样本进行类别感知之后再行进行目标样本的分类,结果显示本文的工作可以显著提升模型分类效果。这进一步证明了本文的方法可以缓解困难任务的域自适应问题。再如,MNIST(*mt*)数据集作为目标域进行域自适应是一个比较容易的过程。现有的域自适应方法对此任务的效果已经很好。在这种简单任务上,本文方法依然可以取得较好的结果。以上结果表明,本文方法在简单和复杂的域自适应任务上均能取得较好的分类效果,尤其是在处理结构差异较大的数据集(SVHN)时展现出优势,这一结果验证了所提方法能有效加强目标域与源域间共有的类别特征,实现更为稳定和一致的性能表现。

表 2 显示了在 Office-Caltech10 数据集上进行的实验结果。其中最优化结果用加粗的形式体现。其中,“*W*”表示 Webcam 数据集作为目标域进行域自适应,“*D*”表示 DSLR 数据集作为目标域进行域自适应,“*C*”表示 Caltech 数据集作为目标域进行域自适应,“*A*”表示 Amazon 数据集作为目标域进行域自适应。

表 2 AWMD-MSDA 在 Office-Caltech10 上与现有域自适应方法对比分析

Table 2 Comparative analysis of AWMD-MSDA and existing domain adaptation methods on Office-Caltech10 %

基准	方法	<i>W</i>	<i>D</i>	<i>C</i>	<i>A</i>	<i>Avg</i>
单源域	Single Best	99.10	98.20	85.40	88.70	92.85
	JAN	99.40	99.40	91.20	91.80	95.45
	DCTN	99.40	99.00	90.20	92.70	95.33
多源域	M3SDA	99.50	99.20	92.20	94.50	96.35
	CASR	<b>99.80</b>	<b>99.80</b>	93.20	94.80	96.90
	TSCDA	98.42	99.04	93.85	95.37	96.67
	AWMD-MSDA	99.40	99.50	<b>94.30</b>	<b>95.50</b>	<b>97.18</b>

实验结果显示,AWMD-MSDA 在 Office-Caltech10 数据集上平均精度高于其他现有域自适应方法。一方面,相比于 Single Best 方法的 92.85%,AWMD-MSDA 算法的效果提高了 4.33%。同时,AWMD-MSDA 在 Caltech(*C*)数据集作为目标域进行域自适应和 Amazon(*A*)数据集作为目标域进行域自适应这两个任务上获得了最好的分类效果。另一方面,本文的方法比其他现有的多源域自适应方法 JAN 和 DCTN 分别提高了 1.73%和 1.85%,JAN 和

DCTN 方法在进行源域样本特征提取过程中均平等地对待所有源域,对每个源域赋予相同的权重,忽略了不同源域与目标任务的贡献存在差异的问题。然而,本文的方法则在类别感知策略的基础上,采用重加权矩匹配策略,通过计算不同源域对目标域分类任务的具体贡献度,以优化域对齐的精确性和效果。所提出的 AWMD-MSDA 与现有的多源域自适应方法 M3SDA、CASR、TSCDA 平均提高了 0.54%,进一步表明了本文提出的重加权矩匹配策略,可以赋予每个源域更加细致的权重,而不是在域级别上对每个源域进行粗权分配。从而进一步有效地促进领域的对齐。

由表 1 和 2 的实验结果可知,本文的方法在 Digits-Five 数据集和 Office-Caltech10 数据集上平均分类精度分别达到了 94.11%和 97.18%,在 Digits-five 数据集相比于 ABMSDA 方法提高了 8.65%,说明了本文的方法通过从减少不相关类别的干扰上比仅考虑域级别层面来消除类别偏移带来的负迁移有着显著的提升。在 Digits-Five 数

据集上相比于 CASR 和 TSCDA 方法分别高了 0.36%和 1.12%; Office-Caltech10 数据集上相比于 CASR 和 TSCDA 方法分别高了 0.28%和 0.51%;说明了本文方法的重加权矩匹配采用从不同源域之间的特征分布对齐和源域与目标域之间的特征分布对齐两个角度比仅从源域和目标域进行有效对齐有着一定的域自适应效果提升。

### 3.3 类别感知策略的有效性分析

为了验证提出的类别感知策略在多源域自适应中存在类别偏移场景下域自适应的有效性,本节对比分析中类别感知模块的优化器参数设置与文献[23]一致。实验设置 Digit-Five 数据集上每个源域拥有 5 个类别的样本,目标域拥有 7 个类别的样本。实验过程中,轮流选择一个域作为目标域,其他域作为源域,共进行 5 轮实验,每轮实验分为 2 组,一组为未采用类别感知策略,另一组为采用类别感知策略。每轮实验重复 5 次以确保结果的可靠性,并计算平均分类精度,具体如表 3 所示,表中加粗数字表示最优值。

表 3 Digits-Five 数据集上类别感知策略的有效性分析

Table 3 Effectiveness analysis of the class-aware strategy on the Digits-Five dataset

基准	方法	<i>mm</i>	<i>mt</i>	<i>up</i>	<i>sv</i>	<i>sy</i>	<i>Avg</i>	%
多源域	未采用类别感知策略	90.00	96.39	95.70	87.50	<b>95.45</b>	93.00	
	采用类别感知策略	<b>90.50</b>	<b>98.02</b>	<b>97.82</b>	<b>88.80</b>	95.40	<b>94.11</b>	

实验结果显示,在未采用类别感知策略的情况下,平均分类精度为 93.00%。各个目标域上的具体精度分别为:*mm* (90.00%)、*mt* (96.39%)、*up* (95.70%)、*sv* (87.50%)、*sy* (95.45%)。在采用类别感知策略的情况下,平均分类精度提升至 94.11%。各个目标域上的具体精度分别为:*mm* (90.50%)、*mt* (98.02%)、*up* (97.82%)、*sv* (88.80%)、*sy* (95.40%)。从表中可以看出,采用类别感知策略后,目标域上的分类精度几乎在每个域上都得到了提升,特别是在 MNIST(*mt*)和 USPS(*up*)域,分别提高了 1.63%和 2.12%。通过引入类别感知策略,模型能够更好地识别和适应目标域的分类分布,从而提高分类精度。这是因为类别感知策略能够在源域和目标域之间更有效地进行特征对齐和知识迁移。虽然在 Synthetic Digits(*sy*)域上,采用类别感知策略后的分类精度略有下降。这是由于 Synthetic Digits 是 SVHN 数据集的合成对应物,其图像中数字的边缘特征呈较为模糊的状态,导致策略在该域上的

效果不如其他域。通过本实验发现,类别感知策略在多源域自适应中具有显著的效果,能够有效提高目标域的分类精度,验证了其在存在类别偏移场景下的有效性。尽管在个别域上略有下降,但整体而言,类别感知策略的引入对提升模型性能具有重要贡献。

### 3.4 重加权矩匹配策略的有效性分析

为了验证提出的重加权矩匹配策略在多源域自适应中存在类别偏移场景下域自适应的有效性,本节对比分析中重加权矩匹配模块的优化器参数设置与文献[21]一致。实验设置 Office-Caltech10 数据集上每个源域拥有 10 个类别的样本,目标域拥有 10 个类别的样本,尽管它们的类别数相同,但标记空间是不同的。实验过程中,轮流选择一个域作为目标域,其他域作为源域,共进行 5 轮实验,每轮实验分为 2 组,一组为未采用重加权矩匹配策略,另一组为采用重加权矩匹配策略。每轮实验重复 5 次以确保结果的可靠性,并计算平均分类精度,具体如表 4 所示,表中加粗数字表示最优值。

表 4 Office-Caltech10 数据集上重加权矩匹配策略的有效性分析

Table 4 Effectiveness analysis of the re-weighted moment matching strategy on the Office-Caltech10 dataset

基准	方法	<i>W</i>	<i>D</i>	<i>C</i>	<i>A</i>	<i>Avg</i>	%
多源域	未采用重新加权矩匹配策略	96.70	<b>99.90</b>	91.70	93.51	95.45	
	采用重新加权矩匹配策略	<b>99.40</b>	99.50	<b>94.30</b>	<b>95.50</b>	<b>97.17</b>	

实验结果显示,未采用重加权矩匹配策略的情况下,平均分类精度为 95.45%;采用重加权矩匹配策略后,平均分类精度提升至 97.17%。大多数目标域的分类精度有所提升,尤其是在 Webcam(W)和 Caltech(C)域,分别提高了 2.70%和 2.60%。尽管在 DSLR(D)域的分类精度略有下降,但总体提升的趋势明显。这种下降是由于该域的特定期类分布与其他域存在显著差异,导致策略在该域上的效果不如预期。此外,仅采用重加权矩匹配策略在某些情况下可能对少数类别的权重调整不够精准,从而影响整体效果。在类别感知策略通过精准对齐相似类别的分布,提高了源域到目标域的正向迁移效果的基础上,重加权矩匹配策略通过自适应更新的贡献度权重,量化了不同源域对目标域的贡献,减少了权重调整的误差,从而提升分类精度。通过本实验发现重加权矩匹配策略在多源域自适应中具有显著效果,能够有效提高目标域的分类精度,验证了其在存在类别偏移场景下的有效性。

### 3.5 AWMD-MSDA 中各组件的有效性分析

为了验证 AWMD-MSDA 中各组件对多源域自适应中存在类别偏移场景下域自适应的有效性,本文在 Office-Caltech10 数据集的所有域自适应任务上进行了消融实验,其中源域和目标域拥有的类别样本数与文献[22]一致。

如表 5 所示,“√”代表选用的组件, $\mathcal{L}_{cls}$ 表示仅利用域分类器交叉熵损失, $\mathcal{L}_{class}$ 表示采用的类别感知策略的识别损失, $\mathcal{L}_{awmd}$ 表示采用重加权矩匹配策略的域级综合对齐损失。结果显示,对于加入类别感知策略的模型,相比与仅利用域分类器交叉熵损失的模型域自适应效果得到了 4.08%的显著改善,这验证了通过对源域数据输出的特征进行感知能够识别各源域中不同类别的样本和源域之间共享类别的样本,得到源域与目标域之间的类别相关性,从类别级实现源域与目标域之间更紧密的对齐的假设。对于加入重加权矩匹配策略的模型,相比与仅利用域分类器交叉熵损失的模型域自适应效果得到了 5.10%的显著改善,这也验证了在域适应过程中针对不同源域对目标域任务的贡献度不同,赋予不同的权重值进行域对齐,可以抑制域适应过程中的负迁移,进而提升模型的整体适应效果的重要性。以上结果表明,本文所引入的 3 个损失在模型训练中都起着至关重要的作用。每增加一个组件,平均精度均有所提升,验证了每个组件对提升多源域自适应性能的重要性,尤其是在处理类别偏移问题时。

表 5 Office-Caltech10 数据集上 AWMD-MSDA 中各组件的有效性分析

Table 5 Effectiveness analysis of each component in AWMD-MSDA on the Office-Caltech10 dataset

$\mathcal{L}_{cls}$	$\mathcal{L}_{class}$	$\mathcal{L}_{awmd}$	Avg/%
√			91.37
√	√		95.45
√		√	96.47
√	√	√	<b>97.17</b>

## 4 结 论

本文针对多源域自适应中存在的类别偏移问题,提出了 AWMD-MSDA 算法,该算法通过引入类别感知策略,有效度量了各个源域与目标域之间的数据分布相似性,从而实现了源域知识的有效迁移。同时,引入重加权矩匹配策略,通过构建自适应更新的贡献度权重,量化了不同源域对目标域在域自适应任务中的贡献度。实验在两个广泛使用的域自适应基准数据集 Digits-Five 和 Office-Caltech10 上进行,结果表明 AWMD-MSDA 算法在处理复杂的多源域自适应问题上表现优异,尤其在克服类别偏移挑战方面,能够显著提高分类精度。在 Digits-Five 数据集上,采用 AWMD-MSDA 算法后,各个目标域的分类精度均有所提升,特别是在 MNIST(*mt*)和 USPS(*up*)域,分别提高了 1.63%和 2.12%;在 Office-Caltech10 数据集上,各个目标域的分类精度也显著提升,尤其在 Webcam(W)和 Caltech(C)域,分别提高了 2.70%和 2.60%。这些实验结果明确证明了所提算法在处理多源域自适应问题中的有效性,展示了其在实际应用中的一定潜力。

未来工作将深入探索更为精细化的特征对齐方法,以进一步提升域适应性能,解决更为复杂的类别偏移问题,推进多源域自适应领域的发展。

### 参考文献

- [1] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks [J]. Journal of Machine Learning Research, 2016, 17(1): 2096-2030.
- [2] LI K, LU J, ZUO H, et al. Multi-source contribution learning for domain adaptation [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(10): 5293-5307.
- [3] TZENG E, HOFFMAN J, SAENKO K, et al. Adversarial discriminative domain adaptation [C]. 2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2017: 7167-7176.
- [4] 赵文仓, 徐长凯, 王春鑫. 域对抗的领域自适应目标检测 [J]. 电子测量技术, 2020, 22(43): 45-49.  
ZHAO W C, XU CH K, WANG CH X. Domain-adversarial domain adaptation for object detection [J]. Electronic Measurement Technology, 2020, 22(43): 45-49.
- [5] 宗子杨, 何军, 宦海, 等. 基于欧式距离对偶的对抗性无监督域适应算法研究 [J]. 电子测量技术, 2023, 46(14): 95-101.  
ZONG Z Y, HE J, HUAN H, et al. Learning on the Euclidean discrepancy dual for unsupervised domain adaptation [J]. Electronic Measurement Technology, 2023, 46(14): 95-101.

- [6] 翟永杰, 胡哲东, 白云山, 等. 融合迁移学习的绝缘子缺陷分级检测方法[J]. 电子测量技术, 2023, 46(6): 23-30.
- ZHAI Y J, HU ZH D, BAI Y SH, et al. Integrating transfer learning for insulator defect grading detection[J]. Electronic Measurement Technology, 2023, 46(6): 23-30.
- [7] XU R J, CHEN Z L, ZUO W M, et al. Deep cocktail network: Multi-source unsupervised domain adaptation with category shift[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 3964-3973.
- [8] PENG X CH, BAI Q X, XIA X D, et al. Moment matching for multi-source domain adaptation[C]. 17th International Conference on Computer Vision, 2019: 1406-1415.
- [9] ZHU D H, CHEN N. Multi-source domain adaptation and fusion for speaker verification[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2022, 30: 2103-2116.
- [10] SUN SH L, SHI H L, WU Y B. A survey of multi-source domain adaptation[J]. Information Fusion, 2015, 24: 84-92.
- [11] YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks[J]. Advances in Neural Information Processing Systems, 2014, 2: 3320-3328.
- [12] LONG M SH, ZHU H, WANG J M, et al. Deep transfer learning with joint adaptation networks[C]. International Conference on Machine Learning, 2017: 2208-2217.
- [13] WANG J D, FENG W J, CHEN Y Q, et al. Visual domain adaptation with manifold embedded distribution alignment[C]. 26th ACM International Conference on Multimedia, 2018: 402-410.
- [14] SUN B CH, FENG J SH, SAENKO K. Return of frustratingly easy domain adaptation[C]. 13th AAAI Conference on Artificial Intelligence, 2016: 2058-2065.
- [15] ZHANG ZH, WANG M ZH, HUANG Y, et al. Aligning infinite-dimensional covariance matrices in reproducing kernel Hilbert spaces for domain adaptation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 3437-3445.
- [16] HOFFMAN J, TZENG E, PARK T, et al. CyCADA: Cycle-consistent adversarial domain adaptation[C]. 35th International Conference on Machine Learning, 2018: 1989-1998.
- [17] KUMAR A, SATTIGERI P, WADHAWAN K, et al. Co-regularized alignment for unsupervised domain adaptation[J]. Advances in Neural Information Processing Systems, 2018, 13: 9345-9356.
- [18] ZHAO H, ZHANG SH H, WU G H, et al. Adversarial multiple source domain adaptation[J]. Advances in Neural Information Processing Systems, 2018, 27: 8559-8570.
- [19] ZHU Y CH, ZHUANG F ZH, WANG D Q. Aligning domain-specific distribution and classifier for cross-domain classification from multiple sources[C]. 2019 AAAI Conference on Artificial Intelligence, 2019: 5989-5996.
- [20] ZHANG J, ZHOU W, CHEN X Q, et al. Multisource selective transfer framework in multiobjective optimization problems[J]. IEEE Transactions on Evolutionary Computation, 2019, 24(3): 424-438.
- [21] ZUO Y K, YAO H T, XU CH SH. Attention-based multi-source domain adaptation[J]. IEEE Transactions on Image Processing, 2021, 30: 3793-3803.
- [22] WANG SH SH, WANG B L, ZHANG ZH, et al. Class-aware sample reweighting optimal transport for multi-source domain adaptation[J]. Neurocomputing, 2023, 523: 213-223.
- [23] 邱春红, 邵晓根. 基于同伴辅助学习分类器的部分域自适应方法[J]. 计算机应用与软件, 2024, 41(1): 168-176.
- QIU CH H, SHAO X G. Partial domain adaptation method based on peer-assisted learning classifier[J]. Computer Applications and Software, 2024, 41(1): 168-176.
- [24] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]. 32nd International Conference on Machine Learning, 2015: 1180-1189.
- [25] GAUTHERON L, REDKO I, LARTIZIEN C. Feature selection for unsupervised domain adaptation using optimal transport[J]. Machine Learning, 2022, 111(9): 4285-4312.

## 作者简介

谭棉, 副教授, 主要研究方向为自然图像抠图、微计算。

E-mail: tanmian@gzmu.edu.cn

李志玲, 硕士研究生, 主要研究方向为迁移学习、域自适应。

E-mail: lizhiling2022@163.com

陈望, 硕士研究生, 主要研究方向为小样本图像分类。

E-mail: 2267202170@qq.com

曾涛涛, 硕士研究生, 主要研究方向为模型压缩。

E-mail: 15185762043@163.com

冯夫健(通信作者), 教授, 主要研究方向为智能计算、微计算及其应用。

E-mail: fujian\_feng@gzmu.edu.cn