

# 基于联合层级粒包络判别特征学习的 无监督领域自适应方法

李普飞, 王 品\*, 李勇明, 张锦华, 颜 芳

(重庆大学微电子与通信工程学院, 重庆 400000)

**摘 要:** 无监督领域自适应(Unsupervised Domain Adaptation, UDA)是模式识别中的重要研究领域,旨在将标注完善但分布不同的源域知识转移到未标注的目标域数据上。现有方法仅关注源域和目标域原始样本分布的对齐,因此在分布差异较大时效果不佳。近年来,基于语义的UDA方法引入了类别信息。然而,类别信息过于粗略,难以充分反映源域和目标域的分布。为了解决这一问题,本文提出了一种用于无监督领域自适应的联合层级粒包络判别特征学习方法,该方法在三个层次上聚合了原始样本对、类别和粒包络信息,从而由粗到细地反映了数据分布。具体来说,本文引入了“知识金字塔”理论,通过构造粒包络,连接原始样本与类中心,建立三层粒度样本,替代现有方法中单层粒度的原始样本。同时,基于本文方法设计了一种迭代式聚类算法,揭示了样本间的关联信息,在原始样本与类中心之间生成粒包络。通过使用Bagging集成模式整合这三种粒度层,并对不同层次的粒度进行加权,以确保令人满意的精度。在基准数据集上的实验结果表明,本文方法能有效减小域间差异并提升分类精度,且优于其它主流领域自适应方法。

**关键词:** 迁移学习;领域自适应;跨域对齐;粒包络;知识金字塔

**基金项目:** 国家自然科学基金(No.U21A20448, No.61771080)

**中图分类号:** TP391.4

**文献标识码:** A

**文章编号:** 0372-2112(XXXX)XX-0001-13

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20240793

## Unsupervised Domain Adaptation Based on Joint Hierarchical Granularity Envelope Discriminative Feature Learning

LI Pu-fei, WANG Pin\*, LI Yong-ming, ZHANG Jin-hua, YAN Fang

(School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400000, China)

**Abstract:** Unsupervised domain adaptation(UDA) is a significant research area that aims to transfer knowledge from source data, which is well-labeled but distributed differently, to unlabeled target data. Existing methods only considered aligning the distributions of the original source and target domain samples, suffering from the big difference in the distributions. In recent years, semantic-based UDA has integrated category information on this basis. However, the category information is too coarse and cannot fully reflect the distributions of source and target domain. To solve this problem, a joint hierarchical granularity envelope(JHGE) discriminative feature learning approach is proposed, which integrates information from original sample pairs, categories, and granular envelopes at three levels. This method can reflect the distribution from coarse to fine. Specifically, the "knowledge pyramid" theory is firstly incorporated into the UDA framework to realize multi-sample granularity-based semantic representation. Besides, granular envelopes are created to connect the original samples with class centers, establishing three layers of sample granularity, which replace the single layer of original samples in existing UDA methods. Secondly, an iterative clustering approach is developed to uncover associative information between samples, generating granular envelopes between original samples and class centers. This three-layer sample granularity enriches the informative content of the existing UDA methods. Thirdly, a bagging ensemble mode is implemented to integrate the three-layer granularity spaces. The different layers of granularity are weighted to ensure satisfactory accuracy. Experimental results on benchmark datasets demonstrate that this method can effectively reduce differences across domain and outper-

forms the state-of-the-art domain adaptation methods.

**Key words:** transfer learning; domain adaptation; cross-domain alignment; granularity envelopes; knowledge pyramid

**Foundation Item(s):** National Natural Science Foundation of China (No. U21A20448, No. 61771080)

## 1 引言

由于拥有大量的标注数据,传统机器学习算法在训练数据和测试数据分布一致的情况下表现出卓越的性能.然而,在许多实际情况下,由于对象图像可能在不同的角度、背景和光照条件下捕获,导致数据的边缘或条件分布存在差异<sup>[1-3]</sup>.在这种情况下,理想的方法是从训练数据(源域)中获得一个可迁移的分类器,使其能够对分布不同的测试数据(目标域)进行预测.幸运的是,领域自适应(Domain Adaptation, DA)作为一种有前景的方法,能够解决领域之间分布不同的问题.通过减小源域和目标域之间的分布差异,DA能够将源域中获得的大量知识迁移到目标域中.领域自适应方法通常可分为半监督领域自适应(Semi-Supervised Domain Adaptation, SSDA)和无监督领域自适应(Unsupervised Domain Adaptation, UDA)两种学习设置<sup>[4]</sup>.在SSDA中,目标领域具有一定数量的标签;而在UDA中,目标领域则没有可用的标签.UDA在提升机器学习模型的跨领域自适应性和泛化性方面显示出巨大的潜力.由于在许多应用中目标领域的标签通常不可用,UDA已成为近年来的研究热点<sup>[5-7]</sup>.

近年来,基于特征的UDA方法受到广泛关注<sup>[8-11]</sup>.这些方法的主要目标是解决全局域偏移问题,旨在使源域和目标域的整体数据分布相互对齐.然而,这些方法的一个明显局限是它们在进行领域自适应时忽略了类别信息.因此,即使实现了全局域适应,两个域的数据分布也只能大致对齐,导致不同类别的数据点可能过于接近,进而引发潜在的误分类问题.

为了应对这一挑战,学者们越来越专注于语义对齐方法<sup>[12-15]</sup>.这些方法在领域自适应中引入了类别信息,以更准确地对齐源域和目标域的分布.这些方法侧重于学习局部领域偏移,通过对齐类别中心来确保源域和目标域中同一类别的分布精确对齐,以最小化全局领域偏移.然而,仅仅依赖类别中心来表征局部特征分布存在风险,因为类别中心数量有限,可能无法有效减轻跨域样本类别中的偏差.此外,原始数据与类别中心之间存在显著差异,尤其是在样本数量较大的情况下.即使在结合原始数据和类别中心的语义对齐方法中,也可能忽视原始样本与类别中心之间的信息,导致模型的泛化性和准确性降低.因此,深入探索原始样本与类别中心之间的信息对于实现更好的UDA至关重要.

“知识金字塔”是指人类进行数据挖掘与知识获取

的层次结构特性,强调了挖掘数据样本间关联信息,形成信息层来获取知识的层次关系,体现了在数据处理中以层次结构考虑信息和知识对于类别分类的重要性<sup>[16,17]</sup>.在这个概念中,“数据”代表原始样本,是最小粒度,包含最多信息和冗余;而“知识”则包括数据点所属的类别.在两者之间是“信息”,信息量较少但更有助于知识的发现.因此,综合考虑这些层次结构信息对于从数据中获取类别分类知识至关重要.根据图1所示,数据集的基本单位是样本(数据点),所有原始样本被视为“数据”,而类别中心代表了原始样本所属的类别,称为“知识”.然而,数据层和信息层之间蕴含了大量有助于知识发现的信息.当前,大多数无监督领域自适应(UDA)方法主要集中在数据层,而语义对齐方法通常结合了数据和知识层.因此,有必要研究构建信息层的方法,并在这三种粒度样本层中进行跨域适应匹配.

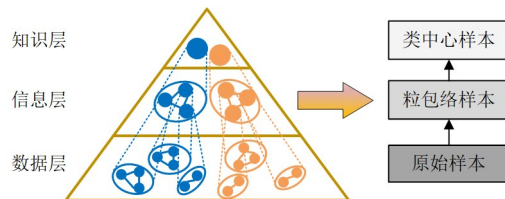


图1 知识金字塔示意图

人类的认知机制指出人类的大脑识别样本不仅仅依赖样本本身的特性,还会通过相似样本之间的关联关系增强理解<sup>[16,17]</sup>.本文为解决上述问题提出了一种用于无监督领域自适应的联合层级粒包络(Joint Hierarchical Granularity Envelope, JHGE)判别特征学习方法.粒包络是指集合一组相似样本并通过挖掘其关联信息同构出新样本,该新样本表征这组样本的内在关联信息,记为粒包络.该方法构建了数据-信息-知识三个层次的样本.其独特之处在于不仅综合考虑了原始样本和类别语义信息,同时在领域自适应过程中关注了样本之间的关联信息.与现有方法不同,本文方法包含三个层次的样本粒度:原始样本、粒包络样本和类别样本.通过同时对齐这三个层次的样本以适应跨领域分布,本方法能够有效地应对复杂的无监督领域自适应场景.具体过程如图2所示:首先对源域样本集进行粒包络生成操作以生成粒包络样本;接着,分别利用原始样本、粒包络样本和类别样本对源域和目标域的分布进行对齐,生成三个新的自适应样本集;最后,通过对这三层自适应结果进行建模和整合,得到最终结果.

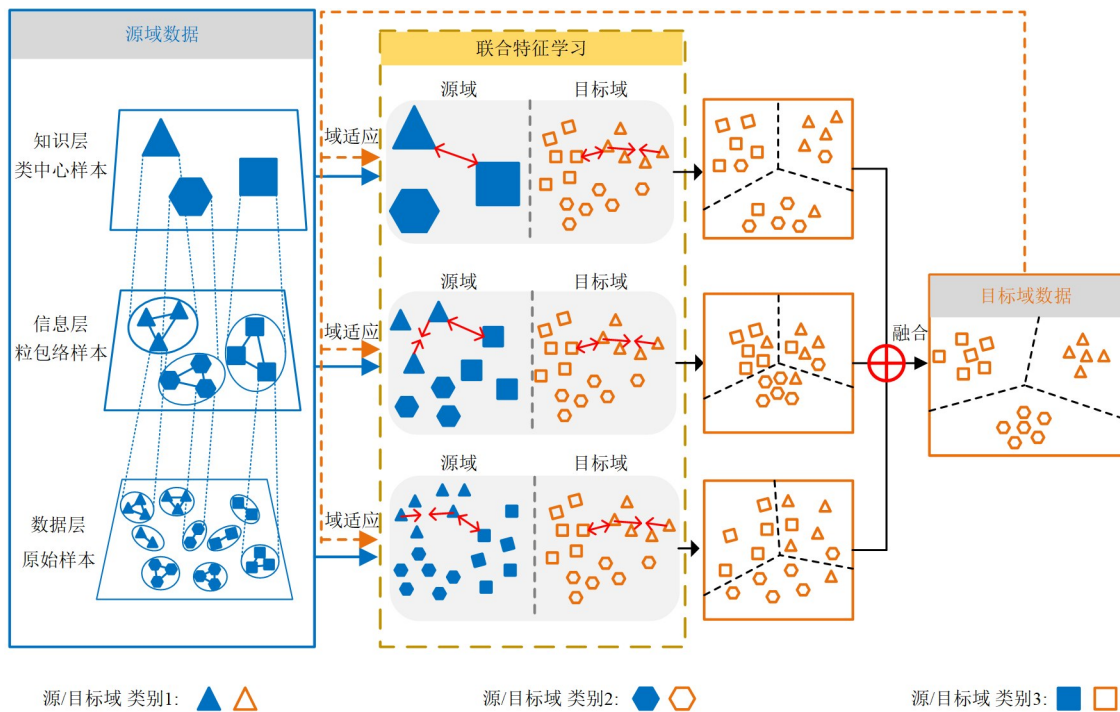


图2 基于无监督领域自适应的联合层级粒包络判别特征学习框架

我们的方法不仅涵盖了基于知识层的领域自适应,其中每个类别的中心被视为该类别的最粗粒度,还结合了基于数据层的领域自适应,重点关注最小粒度的单个数据样本,以捕捉更细节的信息.此外,受到粒计算的启发,我们在信息层中构造了可调粒度的粒包络样本作为最大和最小粒度之间的中间层次.在领域自适应的过程中,我们引入了一个直接且高效的损失项,用以惩罚同一类别中相距最远的数据对,并最小化不同类别的数据对之间的距离.这一机制的目标是识别和优化两个领域中最具挑战性的样本对,最终在学习的特征空间中增加类内紧凑性和类间可分性.

本文的主要贡献如下:(1)与现有的UDA算法不同,本文提出了一种基于三层粒度样本的算法,在数据层、信息层和知识层全面实现了UDA,这一方法能够更好地应对复杂UDA场景中的固有特征和需求.(2)不同于现有的语义对齐UDA算法,所提出的方法采用了类似Bagging集成的模式来整合三种粒度样本层,并通过调整权重灵活平衡它们的重要性,以实现最优效果.(3)设计了一种表征信息层的粒包络样本,改善了目前基于语义对齐UDA方法的性能.(4)在判别性迁移特征学习的背景下,设计了一个简单但高效的损失项,以优化最具挑战性的样本对之间的距离,旨在增加两个域中类内紧凑性和类间可分性.(5)我们在多个经典基准数据集上进行了全面实验,结果表明本文所提出的方法具有优越性.

## 2 相关工作

近年来,领域自适应取得了显著进展.本节主要回顾了先前关于无监督领域自适应的方法,重点关注两个方面:全局对齐方法和语义对齐方法.

### 2.1 全局对齐方法

这类方法通过减少跨域原始数据之间的边缘分布差异来实现分布自适应,这对应于上文所提到的基于数据层的分布匹配.例如,GONG等人<sup>[8]</sup>提出了使用测地流核(Geodesic Flow Kernel, GFK)来对格拉斯曼流形空间中的跨区域偏移进行建模.此外,一些方法利用最大均值差异<sup>[18]</sup>(Maximum Mean Difference, MMD)度量来学习共享特征表示.PAN等人<sup>[9]</sup>提出的迁移成分分析(Transfer Component Analysis, TCA)通过MMD来度量域间低维嵌入的差异,从而减少全局偏移.TZENG等人<sup>[19]</sup>提出的深度域混淆(Deep Domain Confusion, DDC)将基于MMD的域混淆损失整合到卷积神经网络架构中.在此概念的基础上,LONG等人<sup>[20]</sup>将任务特定的隐藏表示嵌入到再生核希尔伯特空间(Reproducing Kernel Hilbert Space, RKHS)中,在此空间中显式匹配源域和目标域的均值.与这些方法不同,SUN等人<sup>[21]</sup>通过线性变换方法对齐源域和目标域分布的二阶统计特征.并且他们进一步通过引入非线性变换扩展了CORAL方法<sup>[10]</sup>,并将其应用于深度网络.此外,GANIN等人<sup>[22]</sup>首次引入对抗学习来对齐源域和目标域的全局分布.而TZENG等人<sup>[11]</sup>则提出了分别训练源域和目标域特

征提取器,并使用判别器对齐两个域的特征.

尽管这些全局域对齐方法在数据层级上通过原始实例匹配边缘分布展示了其有效性,但它们未能考虑类别层和知识层维度,从而限制了其在多样化场景中的表现.

## 2.2 语义对齐方法

这些方法在数据层级适应边缘分布的同时,也在知识层级适应条件分布. 具体而言, LONG 等人<sup>[5]</sup>通过引入类别信息将 MMD 扩展到条件分布,提出了联合分布适应(Joint Distribution Adaptation, JDA),此外他们进一步对齐了多个特定层次的联合分布<sup>[23]</sup>以学习迁移网络. ZHANG 等人<sup>[24]</sup>提出了在两个耦合投影中同时减少几何偏移和两个领域的分布偏移. WANG 等人<sup>[25]</sup>提出了平衡分布自适应算法来加权边缘和条件分布的重要性,并在文献<sup>[26]</sup>中对权重因子进行了量化评估. LI 等人<sup>[27]</sup>设计了一个损失项,该项使同一类别的样本之间的距离最小化,同时使不同类别的样本间的距离最大化. 此外, LI 等人<sup>[7]</sup>进一步通过优化类内距离最大的样本对和类间距离最小的样本对,以学习判别特征. LIANG 等人<sup>[28]</sup>引入了一个与领域无关的聚类促进项,用来减小跨域语义差异和增加两个领域的类内紧凑性. JING 等人<sup>[13]</sup>提出在子空间中同时对齐一阶统计数据和二阶统计数据,在保留几何属性的同时,减少域偏移. LU 等人<sup>[14]</sup>提出一种判别不变方法,结合源域的类别信息和目标域的局部结构信息来学习域间固有特征属性. WANG 等人<sup>[12]</sup>从样本角度利用图结构对齐两个领域的分布,以表征样本与类别之间的关系. YANG 等人<sup>[15]</sup>利用类别引导的低秩约束来保持同类样本的结构一致性,在减小类内边缘的同时增加类间差异.

一些研究人员还探索了深度领域自适应方法,这些方法也表现出了优秀的性能. 例如, PEI 等人<sup>[29]</sup>提出了一种多对抗域适应方法,通过捕获每个类别的多模式结构,以基于多个域鉴别器实现不同数据分布的细粒度对齐. SAITO 等人<sup>[6]</sup>提出通过使用任务特定分类器的决策边界来对齐两个域的分布. LONG 等人<sup>[30]</sup>基于分类器预测中表征的判别信息设计了对抗适应模型提高特征的可迁移性和可分性. DENG 等人<sup>[31]</sup>结合域间判别性聚类结构信息,来对齐域间更精细的类条件分布. YU 等人<sup>[32]</sup>提出了动态对抗因子在对抗网络中平衡局部和全局分布的适配问题. CHEN 等人<sup>[33]</sup>设计的生成注意对抗分类网络利用注意力模块增强了对抗网络的可迁移性. YANG 等人<sup>[34]</sup>提出了一种跨领域几何表示的自适应几何学习方法来解决几何图形不一致的问题. HUANG 等人<sup>[35]</sup>通过学习四种特征表示形式来平衡特征的可迁移性和可区分性.

虽然大多数相关研究集中于从类中心的角度对齐

两个领域,但它们仍仅限于样本层次,目前尚未发现公开文献报道类似的分层次的无监督域适应方法. 为了解决这个局限性,本文提出了一种分层次语义对齐方法,旨在基于数据-信息-知识的三个层次的样本实现领域自适应. 本文方法从这三个层次上来对齐两个领域样本的局部和全局几何结构,有效解决了现有方法存在的类别拟合不足和样本过拟合问题. 此外,该方法能够从源域到目标域学习类别关联信息,从而增强所学特征表示的可区分性.

## 3 方法

### 3.1 问题描述

给定一个完全标注的源域  $\mathcal{D}_s = \{(x_{si}, y_{si})\}_{i=1}^{n_s}$ , 其中包含  $n_s$  个样本,  $x_{si} \in \mathbb{R}^{m \times 1}$  代表第  $i$  个源域样本, 其维度为  $m$ , 而  $y_{si}$  表示其对应的类别标签, 共有  $C$  个不同的类别. 与此同时, 给定一个未标注的目标域  $\mathcal{D}_t = \{(x_{ti})\}_{i=1}^{n_t}$ , 其中包含  $n_t$  个样本,  $x_{ti} \in \mathbb{R}^{m \times 1}$  代表第  $i$  个目标域样本, 其维度为  $m$ . 组合数据矩阵  $\mathbf{X} = [\mathbf{X}_s, \mathbf{X}_t] \in \mathbb{R}^{m \times n}$  聚合了源域和目标域的样本, 共包含  $n = n_s + n_t$  个样本. 假设  $\mathcal{D}_s$  和  $\mathcal{D}_t$  共享相同的特征空间 ( $\mathcal{X}_s = \mathcal{X}_t$ ) 和标签空间 ( $\mathcal{Y}_s = \mathcal{Y}_t$ ), 但它们的边缘分布  $P(\mathbf{x}_s) \neq P(\mathbf{x}_t)$  和条件分布  $P(y_s | \mathbf{x}_s) \neq P(y_t | \mathbf{x}_t)$  存在差异. 领域自适应的目标是学习一个特征变换矩阵  $\mathbf{A} \in \mathbb{R}^{m \times d}$ , 将源域和目标域的数据映射到一个共同的子空间, 使得跨域的边缘分布差异最小化. 具体来说, 变换后的源域和目标域数据分别表示为  $\mathbf{Z}_s = \mathbf{A}^\top \mathbf{X}_s \in \mathbb{R}^{d \times n_s}$ ,  $\mathbf{Z}_t = \mathbf{A}^\top \mathbf{X}_t \in \mathbb{R}^{d \times n_t}$ .

本文方法的目的是在三个不同粒度层次上分别学习特征变换矩阵  $\mathbf{A}$  和分类器  $F$ , 即样本层次、粒包络层次和类别层次. 然后基于这些分类器, 对目标域的分类结果进行加权决策.

### 3.2 多粒度判别特征学习方法

#### 3.2.1 基于信息层的特征学习

用于构造信息层的粒包络生成. 为了充分利用标记的源数据, 一种有效的粒包络生成的方法至关重要. 它由三个重要属性组成: 中心  $\mathbf{s} = (1/n) \sum_{i=1}^n \mathbf{x}_i$ , 半径  $r = (1/n) \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{s}\|_2$  和纯度  $P$ . 假设一个粒包络由  $k$  类的对象组成, 即  $\delta_1, \delta_2, \dots, \delta_k$ ,  $|\delta_i|$  表示第  $i$  类中的对象数量. 粒包络的纯度  $P$  可计算为

$$P = \frac{\max(|\delta_i|)}{\sum |\delta_i|}, \quad i = 1, 2, \dots, k. \quad (1)$$

粒包络的生成过程以设定的纯度阈值为条件, 如算法 1 和图 3 所示. 首先整个数据集被初始化为一个粒

包络,其初始中心是所有样本的平均值,标签是  $\arg \max(|\delta_i|)$ . 此时,初始粒包络的纯度最低,不满足所设定的阈值条件. 因此,需要对该粒包络进行分裂:随机选择  $k-1$  个剩余异构点作为新的中心,并将粒包络中的所有样本重新分配到这  $k$  个中心,生成  $k$  个新的粒包络. 计算每个子包络的纯度值,如果不满足设定的纯度阈值,则继续分裂子包络,直到满足要求. 最后对所有符合要求的粒包络进行全局划分,保证每个粒包络的分裂点更接近其中心. 至此,我们就可以获得位于信息层的粒包络样本.

#### 算法 1 粒包络生成方法

输入:有标注的数据  $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ ; 纯度阈值  $P$ .

输出:粒包络  $\mathcal{D}_g$ ; 数据矩阵  $\mathbf{X}_g$  及其标签  $\mathbf{y}_g$

1. 将整个数据集初始化为一个粒包络  $\mathcal{D}_g = \{\mathcal{D}\}$
2. 初始化包络中心  $\mathbf{a} = (1/n) \sum_{i=1}^n \mathbf{x}_i$ , 纯度  $P$
3. REPEAT
4. FOR 每个  $\mathcal{D}_i \in \mathcal{D}_g$  DO
5. 包络中心  $\mathbf{a}'_i$ , 当于粒包络的纯度  $P'_i$
6. IF  $P'_i < P$  THEN
7. 随机选择剩余  $k-1$  异类点作为新的中心
8. 计算每个样本  $\mathbf{x}_i \in \mathcal{D}_i$  与  $k$  个中心的距离;
9. 基于样本  $\mathbf{x}_i \in \mathcal{D}_i$  到  $k$  个中心的距离, 分配样本生成  $k$  个粒包络  $\{\mathcal{D}_i^{j+1}, \mathcal{D}_i^{j+2}, \dots, \mathcal{D}_i^{j+k}\}$
10. 更新旧的粒包络  $\mathcal{D}_g = \mathcal{D}_g - \mathcal{D}_i$
11. 更新当前粒包络  $\mathcal{D}_g = \mathcal{D}_g + \{\mathcal{D}_i^{j+1}, \mathcal{D}_i^{j+2}, \dots, \mathcal{D}_i^{j+k}\}$
12. END IF
13. END FOR
14. UNTIL  $|\mathcal{D}_g|$  不再增加
15. FOR 每个  $\mathcal{D}_i \in \mathcal{D}_g$  DO
16. 计算包络中心  $\mathbf{a}'_i$
17. 计算包络标签  $y'_i = \arg \max(|\delta_i|)$
18. 更新矩阵  $\mathbf{X}_g = [\mathbf{X}_g, \mathbf{a}'_i], \mathbf{y}_g = [\mathbf{y}_g; y'_i]$
19. END FOR
20. 进行全局划分

粒包络领域分布对齐:如上所述,令  $\mathbf{X}_s^g$  表示由源域生成的位于信息层中的粒包络样本,其粒度介于最大粒度和最小粒度之间中间级别. 值得注意的是,使用不同纯度设置很容易构建多个信息层,在本文中我们仅构建一个信息层. UDA 方法旨在缩小域之间的特征分布差距,这对于学习判别性迁移特征至关重要. 为了实现更有效的特征迁移,我们采用最大均值差异来同时减少边缘分布和条件分布的差异. 其数学公式可以表示为

$$\begin{aligned} \mathcal{J}_{\text{mmd}}^g &= D(P^g(\mathbf{x}_s), P(\mathbf{x}_t)) + D(P^g(y_s | \mathbf{x}_s), P(y_t | \mathbf{x}_t)) \\ &= \text{tr}(\mathbf{A}_g^T \mathbf{X}_g \mathbf{M}_g \mathbf{X}_g^T \mathbf{A}_g) \end{aligned} \quad (2)$$

其中  $\mathbf{X}_g = [\mathbf{X}_s^g \ \mathbf{X}_t]$  是粒包络层的总数据矩阵,  $\mathbf{M}_g = \mathbf{M}_0 + \sum_{c=1}^C \mathbf{M}_c$  为 MMD 矩阵,其计算公式为

$$\begin{aligned} (\mathbf{M}_0)_{ij} &= \begin{cases} \frac{1}{n_s^g}, & \mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}_s^g \\ \frac{1}{n_t^g}, & \mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}_t \\ \frac{-1}{n_s^g n_t^g}, & \text{otherwise} \end{cases} \\ (\mathbf{M}_c)_{ij} &= \begin{cases} \frac{1}{(n_s^{g,c})^2}, & \mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}_s^{g,c} \\ \frac{1}{(n_t^c)^2}, & \mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}_t^c \\ \frac{-1}{n_s^{g,c} n_t^c}, & \begin{cases} \mathbf{x}_i \in \mathbf{X}_s^{g,c}, & \mathbf{x}_j \in \mathbf{X}_t^c \\ \mathbf{x}_j \in \mathbf{X}_s^{g,c}, & \mathbf{x}_i \in \mathbf{X}_t^c \end{cases} \\ 0, & \text{otherwise} \end{cases} \end{cases} \quad (3)$$

然而简单地调整域之间的分布并不能保证学习到的特征对于分类任务具有判别性,因此需要提高类内紧凑性和类间可分性. 本文主要研究同一类别中相距最远的样本对和不同类别中相距最近的样本对. 以源域为例,它们的距离损失项可以表示为

$$\begin{aligned} \mathcal{J}_{s, \text{dist}}^g &= \sum_{c=1}^C \sum_{\mathbf{x}_{si} \in \mathcal{D}_s^{g,c}} \max_{\mathbf{x}_{sj} \in \mathcal{D}_s^{g,c}} \|\mathbf{z}_{si} - \mathbf{z}_{sj}\|^2 \\ &\quad - \sum_{c=1}^C \sum_{\mathbf{x}_{sk} \in \mathcal{D}_s^{g,c}} \min_{\mathbf{x}_{sl} \in \mathcal{D}_s^{g,c}} \|\mathbf{z}_{sl} - \mathbf{z}_{sk}\|^2 \\ &= \sum_{c=1}^C \text{Tr}(\mathbf{A}_g^T \mathbf{X}_s^g \mathbf{Q}_s^{g,c} \mathbf{X}_s^{g,c T} \mathbf{A}_g) \end{aligned} \quad (4)$$

其中  $\mathbf{Q}_s^{g,c}$  可以视为每个类别的隐式权重矩阵,其数学表达式可以写为

$$(\mathbf{Q}_s^{g,c})_{ij} = \begin{cases} \sum_{\mathbf{x}_{sq} \in \mathcal{D}_s^{g,c}} \mathbf{I} \left( \mathbf{x}_{si} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_s^{g,c}} \|\mathbf{z}_{sq} - \mathbf{z}_{sk}\|^2 \right) \\ - \sum_{\mathbf{x}_{sq} \in \mathcal{D}_s^{g,c}} \mathbf{I} \left( \mathbf{x}_{si} = \arg \min_{\mathbf{x}_{sk} \in \mathcal{D}_s^{g,c}} \|\mathbf{z}_{sq} - \mathbf{z}_{sk}\|^2 \right), & i=j \\ -\mathbf{I} \left( \mathbf{x}_{si} \in \mathcal{D}_s^{g,c}, \mathbf{x}_{sj} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_s^{g,c}} \|\mathbf{z}_{sj} - \mathbf{z}_{sk}\|^2 \right) \\ +\mathbf{I} \left( \mathbf{x}_{si} \in \mathcal{D}_s^{g,c}, \mathbf{x}_{sj} = \arg \min_{\mathbf{x}_{sk} \in \mathcal{D}_s^{g,c}} \|\mathbf{z}_{sj} - \mathbf{z}_{sk}\|^2 \right) \\ -\mathbf{I} \left( \mathbf{x}_{sj} \in \mathcal{D}_s^{g,c}, \mathbf{x}_{si} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_s^{g,c}} \|\mathbf{z}_{si} - \mathbf{z}_{sk}\|^2 \right) \\ +\mathbf{I} \left( \mathbf{x}_{sj} \in \mathcal{D}_s^{g,c}, \mathbf{x}_{si} = \arg \min_{\mathbf{x}_{sk} \in \mathcal{D}_s^{g,c}} \|\mathbf{z}_{si} - \mathbf{z}_{sk}\|^2 \right), & i \neq j \end{cases} \quad (5)$$

其中  $\mathbf{I}(\cdot)$  表示指示函数. 类似的,我们可以通过伪标

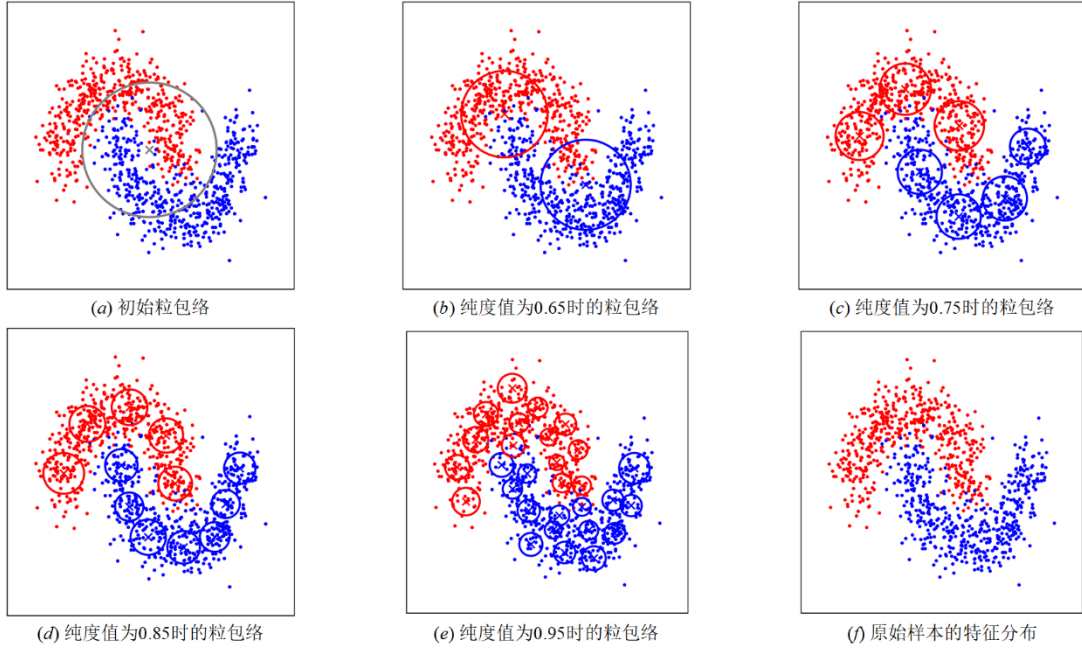


图3 粒包络生成的示意图

签获得目标域的距离损失项:

$$\begin{aligned} \mathcal{J}_{t,\text{dist}}^g &= \sum_{c=1}^C \sum_{\mathbf{x}_i \in \mathcal{D}_i^{g,c}} \max_{\mathbf{x}_j \in \mathcal{D}_j^{g,c}} \|\mathbf{z}_{ti} - \mathbf{z}_{tj}\|^2 \\ &\quad - \sum_{c=1}^C \sum_{\mathbf{x}_i \in \mathcal{D}_i^{g,c}} \min_{\mathbf{x}_k \notin \mathcal{D}_i^{g,c}} \|\mathbf{z}_{ti} - \mathbf{z}_{tk}\|^2 \\ &= \sum_{c=1}^C \text{Tr}(\mathbf{A}_g^\top \mathbf{X}_i \mathbf{Q}_i^c \mathbf{X}_i^\top \mathbf{A}_g) \end{aligned} \quad (6)$$

如果我们令  $\mathbf{Q}_s^g = \sum_{c=1}^C \mathbf{Q}_s^{g,c}$ ,  $\mathbf{Q}_t = \sum_{c=1}^C \mathbf{Q}_t^c$ , 并且  $\mathbf{Q} = \text{diag}(\mathbf{Q}_s^g, \mathbf{Q}_t)$ , 则源域样本和目标域样本的距离损失可以重写为

$$\begin{aligned} \mathcal{J}_{\text{dist}}^g &= \mathcal{J}_{s,\text{dist}}^g + \mathcal{J}_{t,\text{dist}}^g \\ &= \text{Tr}(\mathbf{A}_g^\top \mathbf{X}_g \mathbf{Q}_g \mathbf{X}_g^\top \mathbf{A}_g). \end{aligned} \quad (7)$$

通过结合(2)和(7), 迁移特征学习的整体优化项可以表示为

$$\begin{aligned} \min \mathcal{J}^g &= \mathcal{J}_{\text{mmd}}^g + \alpha \mathcal{J}_{\text{dist}}^g \\ &= \text{Tr}(\mathbf{A}_g^\top \mathbf{X}_g (\mathbf{M}_g + \alpha \mathbf{Q}_g) \mathbf{X}_g^\top \mathbf{A}_g) + \beta \|\mathbf{A}_g\|_F^2 \\ \text{s.t. } &\mathbf{A}_g^\top \mathbf{X}_g \mathbf{H} \mathbf{X}_g^\top \mathbf{A}_g = \mathbf{I}_n \end{aligned} \quad (8)$$

其中  $\mathbf{I}$  是单位矩阵,  $\mathbf{H} = \mathbf{I}_n - (1/n)\mathbf{I}_n \mathbf{I}_n^\top$  是中心矩阵. 显然, 通过引入拉格朗日技术, 令  $\mathbf{A}_g$  的导数为0, 该优化问题可以转化为广义特征值问题:

$$(\mathbf{X}_g (\mathbf{M}_g + \alpha \mathbf{Q}_g) \mathbf{X}_g^\top + \beta \mathbf{I}_m) \mathbf{A}_g = \mathbf{X}_g \mathbf{H} \mathbf{X}_g^\top \mathbf{A}_g \boldsymbol{\theta} \quad (9)$$

其中  $\boldsymbol{\theta} = \text{diag}(\theta_1, \theta_2, \dots, \theta_d) \in \mathbb{R}^{d \times d}$  是拉格朗日乘子, 最优解  $\mathbf{A}_g^*$  是其前  $d$  个最小特征向量.

分类器学习: 在信息层, 我们最终的目标是学习目标域的分类器以有效地对目标域数据进行分类, 其公式可以表示为  $f_t(\mathbf{z}) = \mathbf{w}^\top \mathbf{z}$ . 为了实现这一目标, 我们考

虑两个方面: 首先, 在所学习的投影空间中使目标域样本接近源域; 其次, 利用流形正则化来探索目标域样本之间的几何信息. 目标函数可以写为

$$\begin{aligned} f &= \min_{\mathbf{F}_t} \sum_{i=1}^{n_t} (f_s^g(\mathbf{z}_{ti}) - f_t(\mathbf{z}_{ti}))^2 + \sum_{i,j=1}^{n_t} b_{ij} (f_t(\mathbf{z}_{ti}) - f_t(\mathbf{z}_{tj}))^2 \\ &= \min_{\mathbf{F}_t} (\mathbf{F}_t - \mathbf{F}_s^g)^\top (\mathbf{F}_t - \mathbf{F}_s^g) + \mathbf{F}_t^\top \mathcal{L} \mathbf{F}_t \end{aligned} \quad (10)$$

其中  $b_{ij}$  表示样本  $i$  和  $j$  在投影空间的相似度, 拉普拉斯矩阵  $\mathcal{L} = \mathbf{D} - \mathbf{B}$  可以通过  $\mathbf{D}_{ii} = \sum_{j=1}^{n_t} b_{ij}$  计算得到. 令式(10)的导数  $\partial f / \partial \mathbf{w} = 0$ , 可得(10)的最优解为

$$\mathbf{w}^* = (\mathbf{Z}_t \mathcal{L} \mathbf{Z}_t^\top)^{-1} \mathbf{Z}_t \mathbf{F}_s^g. \quad (11)$$

### 3.2.2 基于数据层的特征学习

在数据层, 所有原始样本数据都代表最细的粒度. 与信息层类似, 我们首先对源数据和目标数据进行域不变判别特征学习. 总体损失函数可以写为

$$\mathcal{J}^d = \mathcal{J}_{\text{mmd}}^d + \alpha \mathcal{J}_{\text{dist}}^d + \beta \|\mathbf{A}_d\|_F^2 \quad (12)$$

其中, 第一项表示数据层中的 MMD 损失项, 第二项表示基于数据层的距离损失项, 第三项表示约束投影矩阵  $\mathbf{A}_d$  的正则项. 因此, 数据层中关于投影矩阵  $\mathbf{A}_d$  的迁移特征学习优化可以表示为

$$\begin{aligned} \min_{\mathbf{A}_d} &\text{Tr}(\mathbf{A}_d^\top \mathbf{X}_d (\mathbf{M}_d + \alpha \mathbf{Q}_d) \mathbf{X}_d^\top \mathbf{A}_d) + \beta \|\mathbf{A}_d\|_F^2 \\ \text{s.t. } &\mathbf{A}_d^\top \mathbf{X}_d \mathbf{H} \mathbf{X}_d^\top \mathbf{A}_d = \mathbf{I} \end{aligned} \quad (13)$$

其中  $\mathbf{X}_d = [\mathbf{X}_s \ \mathbf{X}_t]$  表示数据层的所有源样本和目标样本,  $\mathbf{M}_d$  和  $\mathbf{Q}_d$  的分别表示数据层的 MMD 矩阵和样本间距离矩阵.

在分类器学习阶段,我们同样采用实例重新加权 and 流形正则化策略. 与式(10)类似,其损失项可以写为

$$\begin{aligned} f &= \min_{f_t} \sum_{i=1}^{n_i} (f_s^d(\mathbf{z}_{ii}) - f_t(\mathbf{z}_{ii}))^2 + \sum_{i,j=1}^{n_i} b_{ij} (f_t(\mathbf{z}_{ii}) - f_t(\mathbf{z}_{ij}))^2 \\ &= \min_{F_t} (\mathbf{F}_t - \mathbf{F}_s^d)^\top (\mathbf{F}_t - \mathbf{F}_s^d) + \mathbf{F}_t^\top \mathcal{L} \mathbf{F}_t \end{aligned} \quad (14)$$

通过求解式(13)和式(14),我们可以获得数据层面的特征变换矩阵  $\mathbf{A}_d$  和目标域分类器  $\mathbf{F}_t^d$ .

### 3.2.3 基于知识层的特征学习

在知识层,源域  $\mathcal{D}_s^{kw} = \{(\mathbf{x}_{si}^{kw}, y_{si}^{kw})\}_{i=1}^C$  包含  $C$  个实例,其中  $\mathbf{x}_{si}^{kw}$  表示每个类别的中心点. 对于域间 MMD 损失  $\mathcal{J}_{\text{mmd}}^{kw}$ , 与式(2)类似,可以表示为

$$\begin{aligned} \mathcal{J}_{\text{mmd}}^{kw} &= D(P^{kw}(\mathbf{x}_s), P(\mathbf{x}_t)) + D(P^{kw}(y_s | \mathbf{x}_s), P(y_t | \mathbf{x}_t)) \\ &= \text{tr}(\mathbf{A}_{kw}^\top \mathbf{X}_{kw} \mathbf{M}_{kw} \mathbf{X}_{kw}^\top \mathbf{A}_{kw}) \end{aligned} \quad (15)$$

由于每个类别只有一个样本,因此不需要增强知识层的类内紧凑性,只需增加类间可分性,即惩罚不同类别相距最近的两个样本. 因此,距离损失项可以退化为

$$\begin{aligned} \mathcal{J}_{\text{dist}}^{kw} &= - \sum_{c=1}^C \sum_{\mathbf{x}_{ii} \in \mathcal{D}_s^{kw,c}} \min_{\mathbf{x}_{ik} \in \mathcal{D}_s^{kw,c}} \|\mathbf{z}_{si} - \mathbf{z}_{sk}\|^2 \\ &\quad - \sum_{c=1}^C \sum_{\mathbf{x}_{ii} \in \mathcal{D}_t^c} \min_{\mathbf{x}_{ik} \in \mathcal{D}_t^c} \|\mathbf{z}_{ii} - \mathbf{z}_{ik}\|^2 \\ &= \text{tr}(\mathbf{A}_{kw}^\top \mathbf{X}_{kw} \mathbf{Q}_{kw} \mathbf{X}_{kw}^\top \mathbf{A}_{kw}) \end{aligned} \quad (16)$$

通过结合式(15)和式(16),知识层中的整体特征学习损失函数可以表示为

$$\mathcal{J}^{kw} = \mathcal{J}_{\text{mmd}}^{kw} + \alpha \mathcal{J}_{\text{dist}}^{kw} + \beta \|\mathbf{A}_{kw}\|_F^2 \quad (17)$$

因此知识层对于  $\mathbf{A}_{kw}$  的整体优化可以表示为

$$\min_{\mathbf{A}} \text{Tr}(\mathbf{A}_{kw}^\top \mathbf{X}_{kw} (\mathbf{M}_{kw} + \alpha \mathbf{Q}_{kw}) \mathbf{X}_{kw}^\top \mathbf{A}_{kw}) + \beta \|\mathbf{A}_{kw}\|_F^2 \quad (18)$$

$$\text{s.t. } \mathbf{A}_{kw}^\top \mathbf{X}_{kw} \mathbf{H} \mathbf{X}_{kw}^\top \mathbf{A}_{kw} = \mathbf{I}$$

对于知识层的分类器学习阶段,类似于式(14),损失项可以写为

$$\begin{aligned} f &= \min_{f_t} \sum_{i=1}^{n_i} (f_s^{kw}(\mathbf{z}_{ii}) - f_t(\mathbf{z}_{ii}))^2 + \sum_{i,j=1}^{n_i} b_{ij} (f_t(\mathbf{z}_{ii}) - f_t(\mathbf{z}_{ij}))^2 \\ &= \min_{F_t} (\mathbf{F}_t - \mathbf{F}_s^{kw})^\top (\mathbf{F}_t - \mathbf{F}_s^{kw}) + \mathbf{F}_t^\top \mathcal{L} \mathbf{F}_t \end{aligned} \quad (19)$$

通过求解式(18)和式(19),我们可以得到知识层的特征变换矩阵  $\mathbf{A}_{kw}$  和目标域分类器  $\mathbf{F}_t^{kw}$ .

### 3.3 基于多粒度基础模型的加权融合机制

在本文的 JHGE 框架中,三个不同粒度级别的目标域分类器可以被作为基分类器. 假设定义  $\{\gamma_i, i=1,2,3\}$  为每个分类器的权重. 所提出的框架对目标域预测的

细化标签可以计算为

$$\begin{aligned} \Omega(\mathbf{X}_t) &= \gamma_1 \mathbf{F}_t^d(\mathbf{Z}_t^d) + \gamma_2 \mathbf{F}_t^g(\mathbf{Z}_t^g) + \gamma_3 \mathbf{F}_t^{kw}(\mathbf{Z}_t^{kw}) \\ \text{s.t. } \sum_{i=1}^3 \gamma_i &= 1 \end{aligned} \quad (20)$$

其中有  $\mathbf{Z}_t^d = \mathbf{A}_d^\top \mathbf{X}_t$ ,  $\mathbf{Z}_t^g = \mathbf{A}_g^\top \mathbf{X}_t$  并且  $\mathbf{Z}_t^{kw} = \mathbf{A}_{kw}^\top \mathbf{X}_t$ .

## 4 实验

### 4.1 数据集描述

我们在研究中使用了几个广泛使用的跨域基准数据集,即 CMU-PIE、Office-Caltech10 (SURF 和 DeCAF6) 和 Office-31 (DeCAF7 和 ResNet-50) 数据集. 在这里,我们对每个数据集进行简短介绍:

CMU-PIE<sup>[36]</sup>: 该数据集包含来自 68 个人,尺寸大小为  $32 \times 32$  的 40 000 多张面部图像. 与先前的研究一致,我们选取了 5 个姿势用于领域自适应,即 C05、C07、C09、C27 和 C29.

Office+Caltech-10<sup>[8]</sup>: 该数据集包含来自 10 个类别的 2 533 个实例. 为了充分评估我们的方法,我们采用了两种类型的特征: 800 维的 SURF 特征和 4096 维 DeCAF6 特征.

Office-31<sup>[37]</sup>: 该数据集由 31 个类别的 4 652 个实例组成. 我们采用了来自三个不同领域 Amazon(A)、Webcam(W)、DSLRL(D) 的特征进行实验,即 4096 维 DeCAF7 特征<sup>[38]</sup> 和 2048 维的 ResNet-50 特征<sup>[39]</sup>.

### 4.2 实验设置

为了验证本文所提算法 JHGE 的有效性,我们与大量的主流及近期无监督域自适应方法进行了对比: 1NN (1-Nearest Neighbor)<sup>[40]</sup>, GFK<sup>[8]</sup>, TCA<sup>[9]</sup>, JDA<sup>[5]</sup>, JGSA (Joint Geometrical and Statistical Alignment)<sup>[24]</sup>, MEDA (Manifold Embedded Distribution Alignment)<sup>[26]</sup>, DICD (Domain Invariant and Class Discriminative)<sup>[27]</sup>, DICE (Domain-Irrelevant Class clustering)<sup>[28]</sup>, DTLC (Discriminative Transfer feature and Label Consistency)<sup>[7]</sup>, ACE (Adaptive Component Embedding)<sup>[13]</sup>, MSL-GLR (Marginal Subspace Learning with Group Low-Rank)<sup>[15]</sup>, DIA (Discriminative Invariant Alignment)<sup>[14]</sup>, PGCD (Probability-based Graph embedding Cross-domain and class Discriminative feature learning)<sup>[12]</sup>. 这些方法均是无监督域适应方法,此外部分方法结合了语义信息. 并且为了进一步验证 JHGE 具有通用深度特征的优越性,我们还比较了几种端到端的无监督深度域自适应方法: AlexNet<sup>[41]</sup>, ResNet<sup>[39]</sup>, DDC<sup>[19]</sup>, DAN (Deep Adaptation Networks)<sup>[20]</sup>, DANN (Domain-Adversarial Neural Networks)<sup>[22]</sup>, DCORAL (Deep CORAL)<sup>[10]</sup>, JAN (Joint Adaptation Networks)<sup>[23]</sup>, DUCDA (Deep Unsupervised Convolutional Domain Adaptation)<sup>[42]</sup>, MCD (Maximum

Classifier Discrepancy)<sup>[6]</sup>, CAT(Cluster Alignment with a Teacher)<sup>[31]</sup>, DMAL(Deep Multi-representations Adversarial Learning)<sup>[35]</sup>. 这些方法是结合深度学习的无监督域适应方法.

为了保证公平性,本文参照文献[5]和文献[7],实验参数统一设置为迭代次数  $T=10$ ,  $\alpha=1$ ,对于参数  $\beta$ ,我们在 CMU-PIE 数据集上设置为  $\beta=0.1$ ,其他数据集设置为  $\beta=1$ . 对于纯度参数  $P$ ,我们在所有数据集上统一设置为  $P=0.95$ .

### 4.3 实验结果与讨论

#### 4.3.1 CMU-PIE 数据集的分类结果

为了探讨本文所介绍的 JHGE 在跨领域面部图像识别任务中的有效性,我们在 CMU-PIE 数据集上进行了实验. 我们将 JHGE 与几种近期的无监督领域自适应进行了对比,包括 ACE、MSL-GLR、DIA 和 PGCD 等方法. 根据表 1 中的比较结果显示, JHGE 在准确性方面优于其他所有方法. 具体而言, JHGE 在 20 个任务中的平均分类准确率达到 88.6%. 这一准确率明显优于最佳 Baseline 方法 PGCD 的表现. 因此,我们的研究结果表明 JHGE 在应对跨领域面部图像识别挑战方面具有独特优势.

值得注意的是, CMU-PIE 数据集包含 60 多个类别. 类别数量的增加被认为会增加迁移特征学习的复杂性,导致特征在投影空间中失真和混合增加. JHGE 不仅大幅提高了平均准确率,而且在各种跨域任务中表现出了卓越的性能. 这归因于其在多个粒度样本上对判别性迁移特征的有效学习,使得目标域特征得到充分表征,从而提高了分类性能.

#### 4.3.2 Office+Caltech10 数据集的分类结果

在 Office+Caltech10 数据集中,本文研究使用了两种不同类型的特征: SURF 和 DeCAF6. 根据表 2 的结果显示, JHGE 算法在利用 DeCAF6 特征时获得了最高的平均分类精度,并且在 12 个跨域任务中的 6 个中表现优于其他方法. 虽然 PGCD、MSL-GLR、ACE、DTLC、MEDA 和 DICE 等方法在平均准确度方面均取得了超过 90% 的结果,但相较之下, JHGE 在整体性能上保持着明显的领先地位. 对于迁移任务  $W \rightarrow C$  和  $D \rightarrow A$ ,所提算法未达到最优甚至与部分模型结果相差较大,这可能是因为在数据集上,不同域的样本数量差异较大. 尤其是在任务  $W \rightarrow C$ (295 对 1 123)和  $D \rightarrow A$ (157 对 958)上,源域和目标域样本数量的极端不均衡可能会导致模型无法有效学习到源域和目标域的共同特征.

在该数据集中, SURF 特征的分类结果见表 3. 本文方法在平均准确率方面达到 52.7%,位列第二;而 PGCD 在 SURF 特征数据集上表现最佳. 通过对比 SURF 和 DeCAF6 特征的分类结果,我们发现本文方法

在 DeCAF6 特征的平均准确率上比 PGCD 高 2.2%,而 PGCD 在 SURF 特征的平均准确率上比 JHGE 方法高出 2.2%. 由于 SURF-BoW 特征属于传统手工特征,其表达能力相对较弱,更侧重于局部特征点和描述符向量,而 DeCAF6 特征是深度学习提取的高层特征,具有一定的层次结构和抽象性. 在任务  $A \rightarrow W$ 、 $W \rightarrow C$ 、 $W \rightarrow D$ 、 $D \rightarrow C$  和  $D \rightarrow A$  中使用 SURF 特征的表现不理想,可能是因为 SURF 特征在不同领域之间的迁移性较差,无法有效捕捉跨域的语义信息. 本文算法能更好地理解 and 利用数据集中卷积神经网络提取的抽象特征表示. 尽管如此, JHGE 在这两个特征上表现出了优秀的性能,显示出其强大的泛化能力.

#### 4.3.3 Office31 数据集的分类结果

表 4 展示了所提出的 JHGE 方法与 Office-31 数据集上 DeCAF7 特征的浅层 UDA 方法相比的实验结果. 与浅层 UDA 方法相比, JHGE 在 6 个任务中取得了最高平均准确率达到 75.9%,其中有 3 个任务表现出色. 此外,我们还将其与几种主流的端到端深度域自适应方法进行了比较分析,其结果见表 5. 在  $D \rightarrow A$  和  $W \rightarrow A$  的跨域任务中, JHGE 相对于深度 UDA 方法表现出明显的性能提升. 值得注意的是,在涉及从样本数量有限的源域转移到样本量更大的目标域的任务中, JHGE 优于其他方法,尤其是在  $D \rightarrow A$  和  $W \rightarrow A$  场景中. 这突显了 JHGE 对跨域信息的有效利用及其跨域几何的稳健对齐,特别是在源域提供有限知识时. 对于具有 ResNet-50 特征的 Office-31 数据集,我们提出的 JHGE 表现出最高的平均准确率 87.2%,并在多数任务(例如  $W \rightarrow A$  和  $D \rightarrow W$ )中表现出色. 然而对于 DeCAF7 特征分类任务中的  $W \rightarrow D$ , ResNet50 特征分类任务中的  $A \rightarrow D$ ,  $W \rightarrow D$  上的表现不如预期,可能是由于模型在小样本数量的  $W \rightarrow D$  任务上无法充分适应目标域分布. 对于  $A \rightarrow D$  这样的复杂任务,即使使用 ResNet50 特征,源域和目标域之间的复杂分布差异也对模型的泛化能力提出了更高的要求. 总的来说,与 Office-31 数据集上的浅层和深层 UDA 方法相比, JHGE 均表现出了卓越的性能.

### 4.4 实验分析

在本节中,我们从几个方面分析 JHGE 的有效性: 消融研究、特征可视化和参数敏感性.

#### 4.4.1 消融实验

我们在 CMU-PIE 数据集上进行了多项消融实验,研究了基于三种粒度样本层的影响: 样本层、粒包络层和类别层. 具体来说,比较了本文方法(JHGE)与几个变体方法: “w/o C+G”(仅样本层)、“w/o G”(样本层+类别层)、“w/o C”(样本层+粒包络层)、“G+C”(粒包络层+类别层)、“only G”(仅粒包络层)和“only C”(仅类别层).

表1 CMU-PIE数据集上的分类准确率

Methods	INN	GFK	TCA	JDA	JGSA	DICD	DICE	DTLC	ACE	MSL-GLR	DIA	PGCD	Ours
C05→C07	26.1	26.2	40.8	58.6	52.9	73.0	84.1	85.1	68.8	81.4	78.2	89.3	<b>90.7</b>
C05→C09	26.6	27.3	41.8	52.0	53.1	72.0	77.9	82.7	62.4	72.7	74.1	82.8	<b>88.3</b>
C05→C27	30.7	31.2	59.6	83.7	66.0	92.2	95.9	<b>97.1</b>	84.9	95.3	93.0	95.0	96.4
C05→C29	16.7	17.6	29.4	47.7	46.1	66.9	66.5	77.2	54.4	60.6	73.1	66.0	<b>85.3</b>
C07→C05	24.5	25.2	41.8	60.6	57.5	69.9	81.3	82.8	66.8	82.5	62.3	84.4	<b>88.6</b>
C07→C09	46.6	47.4	51.5	60.2	57.2	65.9	74.0	83.9	72.7	78.1	79.8	<b>86.1</b>	82.2
C07→C27	54.1	54.3	64.7	75.4	69.2	85.3	88.6	92.1	85.5	91.9	87.4	90.8	<b>94.0</b>
C07→C29	26.5	27.1	33.7	40.9	49.8	48.7	68.8	79.7	52.0	64.2	63.5	71.1	<b>83.4</b>
C09→C05	21.4	21.8	34.7	50.9	56.1	69.4	78.8	80.0	64.5	77.0	72.9	83.8	<b>85.6</b>
C09→C07	41.0	43.2	47.7	56.1	58.7	65.4	76.7	<b>84.4</b>	69.4	76.2	74.2	83.1	83.3
C09→C27	46.5	46.4	56.2	68.0	69.5	83.4	85.2	94.3	87.9	96.2	89.7	<b>95.5</b>	94.0
C09→C29	26.2	26.7	33.2	40.3	52.2	61.4	70.8	79.9	62.3	68.6	71.1	78.2	<b>85.9</b>
C27→C05	33.0	34.2	55.6	81.0	63.2	93.1	93.3	96.7	88.2	96.9	93.7	98.1	<b>98.2</b>
C27→C07	52.7	62.9	67.8	82.8	65.7	90.1	95.0	94.8	91.7	96.1	94.7	95.5	<b>97.5</b>
C27→C09	73.2	73.4	75.9	87.2	62.6	89.0	92.3	<b>95.4</b>	92.6	93.8	88.8	93.0	93.9
C27→C29	37.2	37.4	40.3	49.9	57.0	75.6	81.1	84.4	77.3	82.2	85.6	84.5	<b>90.2</b>
C29→C05	18.5	20.4	27.0	47.0	56.3	62.9	73.8	75.4	55.3	65.3	64.7	<b>82.1</b>	79.3
C29→C07	24.2	24.6	30.0	44.8	54.7	57.0	71.2	77.8	54.5	74.7	67.8	<b>80.7</b>	78.4
C29→C09	28.3	28.5	30.0	48.1	56.4	65.9	74.1	82.4	62.1	79.2	74.3	82.8	<b>88.3</b>
C29→C27	31.2	31.3	33.6	56.5	61.7	74.8	81.8	<b>89.7</b>	70.7	88.2	89.5	84.7	88.3
Avg.	34.8	35.4	44.8	59.6	58.3	73.1	80.6	85.8	71.2	81.1	78.9	85.4	<b>88.6</b>

表2 Office+Caltech10数据集上DeCAF<sub>6</sub>特征的分类结果

Tasks	INN	GFK	TCA	JDA	JGSA	DICD	MEDA	DICE	DTLC	ACE	MSL-GLR	PGCD	Ours
C→A	87.3	87.3	89.8	90.2	91.4	90.9	93.4	92.3	92.8	92.1	93.8	92.5	<b>94.8</b>
C→W	72.5	75.9	78.3	85.4	86.8	92.2	95.6	93.6	<b>98.0</b>	89.5	95.6	91.2	96.6
C→D	79.6	83.4	85.4	86.0	93.6	93.6	91.1	93.6	93.0	88.5	90.5	92.4	<b>96.8</b>
A→C	71.7	80.3	82.6	81.9	84.9	86.0	87.4	85.9	88.2	84.5	<b>88.3</b>	86.5	87.1
A→W	68.1	77.0	74.2	80.7	81.0	81.4	88.1	86.4	93.6	87.8	93.2	84.1	<b>98.3</b>
A→D	74.5	80.9	81.5	81.5	88.5	83.4	88.1	89.8	87.3	86.0	<b>91.7</b>	90.4	90.3
W→C	55.3	67.8	80.4	81.2	84.9	84.0	<b>93.2</b>	85.3	88.1	84.5	88.4	85.3	81.7
W→A	62.6	74.3	84.1	90.7	90.7	89.7	99.4	90.7	92.0	92.3	92.6	91.6	<b>96.3</b>
W→D	98.1	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	99.4	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
D→C	42.1	69.1	82.3	80.3	86.2	86.1	87.5	87.4	<b>89.3</b>	86.3	88.7	87.6	87.5
D→A	50.0	75.8	89.1	92.0	92.0	92.1	93.2	92.5	92.9	91.9	<b>93.3</b>	92.5	91.6
D→W	91.5	98.6	99.7	99.3	99.7	99.0	97.6	99.0	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
Avg.	71.1	80.9	85.6	87.4	90.0	89.9	92.8	91.4	93.0	90.3	93.0	91.2	<b>93.4</b>

表6报告了C29→C05、C29→C07、C29→C09、C29→C27迁移任务的结果。观察发现,采用语义对齐的方法(即“w/o G”、“w/o C”、“G+C”和“JHGE”)相对于仅依赖于实例对齐的方法(“w/o C+G”)表现更佳。这凸显了为领域自适应方法引入信息和知识层的必要性。相较于“w/o G”,JHGE取得了更好的性能,这可能归因于信息层提供了样本之间的相关信息。JHGE在与“w/o C+

G”相比有显著改善(79.8% vs. 83.9%),并且与“w/o G”(81.6% vs. 83.9%)和“w/o C”(82.7% vs. 83.9%)相比也有所提升。这表明,结合三层粒度样本是影响结果的主要因素,中间层的粒包络样本在领域自适应中扮演着至关重要的角色。此外,在组合层次的对比中,“G+C”方法的平均准确率为80.2%,但当三层联合使用时(即JHGE方法),平均准确率达到83.9%,进一步

表4 Office-31数据集上DeCAF<sub>7</sub>特征的分类结果

Methods	A→W	A→D	W→A	W→D	D→A	D→W	Avg.
INN	54.0	59.6	40.8	97.8	42.4	90.9	64.3
GFK	58.6	58.4	52.4	93.6	46.1	91.0	66.7
TCA	57.8	59.0	51.6	90.2	47.3	88.2	65.7
JDA	56.6	56.9	47.0	94.1	45.0	98.0	66.3
DTSL	60.0	54.5	46.6	94.3	45.6	94.0	65.8
JGSA	63.5	56.0	<b>58.5</b>	97.2	55.4	98.2	71.5
MEDA	62.7	56.7	54.2	96.5	51.2	98.6	70.0
DICE	59.6	61.4	52.8	92.0	50.8	87.0	67.3
DTLC	66.1	68.4	56.5	<b>98.8</b>	<b>59.0</b>	97.4	74.4
Ours	<b>76.5</b>	<b>68.5</b>	57.6	97.9	55.7	<b>99.3</b>	<b>75.9</b>

表5 Office-31数据集上DeCAF<sub>7</sub>特征和ResNet50特征的分类结果

Features	Methods	A→	A→	W→	W→	D→	D→	Avg.
		W	D	A	D	A	W	
DeCAF <sub>7</sub>	AlexNet	61.6	63.8	49.8	99.0	51.1	95.4	70.1
	DDC	61.8	64.4	52.2	98.5	52.1	95.0	70.7
	DAN	68.5	67.0	53.1	99.0	54.0	96.0	72.9
	DANN	73.0	72.3	51.2	<b>99.2</b>	53.4	96.4	74.3
	DCORAL	66.4	66.8	52.8	95.7	51.5	99.2	72.1
	DUCDA	68.3	68.3	53.6	96.2	51.6	<b>99.7</b>	73.0
	Ours	<b>76.5</b>	<b>68.5</b>	<b>57.6</b>	97.9	<b>55.7</b>	99.3	<b>75.9</b>
ResNet-50	ResNet	68.9	68.4	62.5	96.7	60.7	99.3	76.1
	DAN	78.6	80.5	63.6	97.1	62.8	99.6	80.4
	JAN	84.7	85.4	68.6	97.4	70.0	99.8	84.3
	DANN	79.7	82.0	68.2	96.9	67.4	99.1	82.2
	MCD	<b>92.2</b>	88.6	69.5	98.5	69.7	<b>100.0</b>	86.5
	CAT	91.1	<b>90.6</b>	66.5	99.6	70.4	98.6	86.1
	DMAL	88.4	89.1	70.8	<b>100.0</b>	71.8	99.2	86.6
Ours	90.9	86.6	<b>73.1</b>	97.6	<b>74.8</b>	<b>100.0</b>	<b>87.2</b>	

验证了粒包络层和类别层的互补性。在C29→C05的迁移任务中，“w/o C+G”、“only G”和“only C”分别对应“仅样本层”、“仅粒包络层”和“仅类别层”，这些变体方法的准确率逐一增加，而在其他三个迁移任务上，这些方法的准确率逐渐减少。这可能是因为子域C05数据受到噪声的影响更大，通过与源域的粗粒度特征对齐，更好地保留了类别之间的结构信息，从而增强了模型的稳定性和鲁棒性。

本文方法中，粒包络层的引入能够挖掘样本之间的关联关系，反映了原样本本身和类中心所不同的信息，这在跨域任务中尤为关键。基于粒包络层的方法可以更好地对齐源域和目标域的分布，从而提升模型的域适应能力。因此，增加粒包络层不仅能够补充原样本和类别层的不足，而且能更有效地实现域对齐和跨域迁移，表明其对于提升UDA性能的必要性和互补性。

表6 在CMU-PIE数据集上消融实验的分类结果

Method	C29→ C05	C29→ C07	C29→ C09	C29→ C27	Avg.
w/o C+G	72.1	78.0	82.4	86.5	79.8
w/o G	76.9	<b>79.6</b>	84.7	85.2	81.6
w/o C	77.5	77.2	87.7	<b>88.3</b>	82.7
only G	74.2	69.4	81.2	82.3	76.8
only C	74.6	66.1	71.0	81.5	73.3
G+C	75.1	<b>79.6</b>	84.7	81.6	80.2
JHGE	<b>79.3</b>	<b>79.6</b>	<b>88.3</b>	<b>88.3</b>	<b>83.9</b>

#### 4.4.2 特征可视化

为了提升比较投影空间中目标数据分布的直观性，本节使用了t-SNE可视化技术。我们选取了Office+Caltech10的DeCAF6特征数据集，对W→A任务中的三种粒度层次的领域自适应结果进行了比较：样本、粒包络和类别。从图4(a)可以看出，目标域的数据在数据层的原始样本空间中分布均匀。然而，投影数据中同一类别的样本之间的距离较大，导致了类内紧凑性不足。图4(b)表明在信息层的粒包络样本投影空间中，类内数据的紧凑性有所改善。然而，数据类别之间的可区分性却下降了。这表明在改善类内数据分布的同时，可能会引入决策边界模糊或类别之间的交叉问题。图4(c)显示在知识层的类别投影空间中，目标域的投影数据展现出更大的类间边界。然而，由于源域中每个类别仅包含单个样本，领域自适应增加了目标域类别之间的内部差异。这可能会导致错误分类问题。因此，需要综合考虑三种粒度层级的信息，进行不同粒度级别的无监督领域自适应，以提高模型的泛化性。

#### 4.4.3 参数灵敏度

为了评估所提出模型的鲁棒性和有效性，我们对JHGE中的纯度参数进行了敏感性分析。纯度参数在决定粒包络的细化方面起着关键作用，直接影响分类性能。图5说明了在四个不同任务C→A(Office+Caltech10数据集DeCAF6特征)、D→W(Office+Caltech10数据集SURF特征)、C05→C09、W→D(Office31数据集DeCAF7特征)和A→D(Office31数据集ResNet50特征)中纯度参数和分类精度之间的关系。随着纯度参数从0.5增加到1.0，分类精度在所有任务中往往会提高。这表明更高纯度值代表更精细的包络，通常会增强模型对不同类别的分类能力。我们为所有任务选择了统一的0.95纯度值，接近最大纯度，旨在在分类准确性和模型鲁棒性之间取得平衡。尽管某些任务可能会出现波动，但总体而言，0.95的纯度水平表现出强大的性能。敏感性分析表明，纯度参数对模型性能有显著影响，尤其是在具有复杂领域迁移的任务中。然而，0.95的选定值确

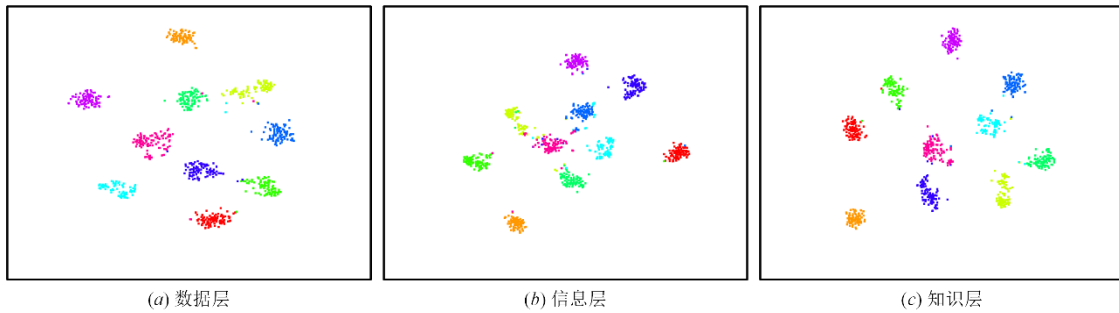


图4 目标域投影数据在三个不同粒度层次的t-SNE图

保了不同任务之间的一致性和鲁棒性,使其成为所提出方法的最佳选择。

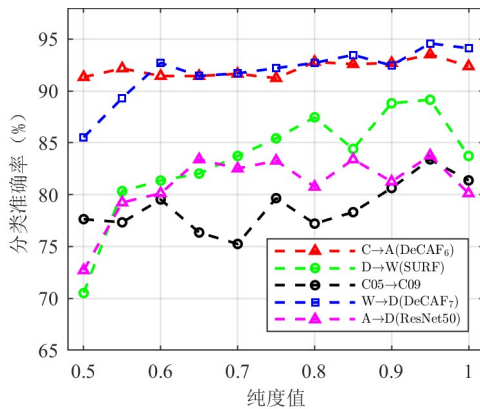


图5 纯度参数对分类准确率的影响

## 5 结论

无监督领域自适应是当前机器学习领域的研究热点.然而,现有大部分UDA算法都是基于对原始样本分布的对齐,存在着类别对齐错误、样本噪声引入以及过拟合风险较高等问题.然而在语义UDA方法中,通过在局部区域内引入样本类别信息,虽然增强了类别质心与个别样本之间的显著差异,但在大量类别样本的情况下,整合个别样本信息和类别质心可能会忽略一些重要信息,从而导致UDA的普遍性差和准确性低.为了解决这一难题,本文提出了一种用于无监督领域自适应的联合层级粒包络判别特征学习方法:首先,在信息层面,该方法采用中间层次的粒包络计算,有效地保留了大部分原始样本信息,对样本中的噪声具有良好的鲁棒性.其次,在知识层面,类别中心以最粗粒度进行计算,而原始数据在数据层面被视为最细粒度.通过在广泛的公共数据集上进行实验,本文方法实验结果均优于其它方法,在CMU-PIE数据集上达到了88.6%的平均准确率,在Office+Caltech10数据集上实现了93.4%(DeCAF6特征)和52.7%(SURF特征)的平均准确率,在Office31数据集上达到了75.9%(DeCAF7特征)和87.2%(ResNet50特征)的平均准确率,实验结

果表明了我们方法的有效性.

本文所提出的算法具有以下几个优点:(1)在每一层次,多粒度的源域样本与目标域样本分别对齐,以增强特征的类别判别性.(2)此外,该算法整合了三个层次的适应结果,以平衡全局和局部关系之间的关联性.(3)本文研究表明,在原始样本和类别之间构建新的样本信息起着重要作用,这种方法能够有效地捕捉源域和目标域之间分布差异的多级信息,实现跨域样本分布的对齐.(4)该框架有助于推动领域自适应研究,为图像处理和相关领域的实际应用提供实用的解决方案.

尽管本文方法取得了明显的效果,但仍需进一步改进:由于引入了粒包络,算法的复杂度略微增加,但可以并行化加以解决.未来我们将引入深度学习技术,将其扩展到端到端的网络中.

## 参考文献

- [1] 王格格, 郭涛, 余游, 等. 基于生成对抗网络的无监督域适应分类模型[J]. 电子学报, 2020, 48(6): 1190-1197. WANG G G, GUO T, YU Y, et al. Unsupervised domain adaptation classification model based on generative adversarial network[J]. Acta Electronica Sinica, 2020, 48(6): 1190-1197. (in Chinese)
- [2] 赵鹏, 王美玉, 纪霞, 等. 基于张量表示的域适配的迁移学习中特征表示方法[J]. 电子学报, 2020, 48(2): 359-368. ZHAO P, WANG M Y, JI X, et al. A novel feature representation based on tensor and domain adaption for transfer learning[J]. Acta Electronica Sinica, 2020, 48(2): 359-368. (in Chinese)
- [3] 蒋晓玲, 吴映波, 陈蒙, 等. 基于跨域结构保持投影的异构在线多源迁移学习方法[J]. 电子学报, 2023, 51(8): 1983-1994. JIANG X L, WU Y B, CHEN M, et al. Heterogeneous online multi-source transfer learning with cross-domain structure preserving projection[J]. Acta Electronica Sinica, 2023, 51(8): 1983-1994. (in Chinese)
- [4] ZHANG L, GAO X. Transfer adaptation learning: A de-

- cade survey[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(1): 23-44.
- [5] LONG M, WANG J, DING G, et al. Transfer feature learning with joint distribution adaptation[C]//2013 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2013: 2200-2207.
- [6] SAITO K, WATANABE K, USHIKU Y, et al. Maximum classifier discrepancy for unsupervised domain adaptation [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 3723-3732.
- [7] LI S, LIU C H, SU L, et al. Discriminative transfer feature and label consistency for cross-domain image classification [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(11): 4842-4856.
- [8] GONG B, SHI Y, SHA F, et al. Geodesic flow kernel for unsupervised domain adaptation[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2012: 2066-2073.
- [9] PAN S J, TSANG I W, KWOK J T, et al. Domain adaptation via transfer component analysis[J]. *IEEE Transactions on Neural Networks*, 2011, 22(2): 199-210.
- [10] SUN B, SAENKO K. Deep CORAL: Correlation alignment for deep domain adaptation[C]// Computer Vision - ECCV 2016 Workshops. Cham: Springer, 2016: 443-450.
- [11] TZENG E, HOFFMAN J, SAENKO K, et al. Adversarial discriminative domain adaptation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 2962-2971.
- [12] WANG W, SHEN Z, LI D, et al. Probability-based graph embedding cross-domain and class discriminative feature learning for domain adaptation[J]. *IEEE Transactions on Image Processing*, 2023, 32: 72-87.
- [13] JING M, ZHAO J, LI J, et al. Adaptive component embedding for domain adaptation[J]. *IEEE Transactions on Cybernetics*, 2021, 51(7): 3390-3403.
- [14] LU Y, LI D, WANG W, et al. Discriminative invariant alignment for unsupervised domain adaptation[J]. *IEEE Transactions on Multimedia*, 2022, 24: 1871-1882.
- [15] YANG L, ZHOU Q, LU B. Marginal subspace learning with group low-rank for unsupervised domain adaptation [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(7): 9122-9135.
- [16] CHENG D, LI Y, XIA S, et al. A fast granular-ball-based density peaks clustering algorithm for large-scale data[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 35(12): 1-14.
- [17] XIA S, WANG C, WANG G, et al. GBRs: A unified granular-ball learning model of pawlak rough set and neighborhood rough set[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 36(1): 1-15.
- [18] GRETTON A, BORGWARDT K, RASCH M, et al. A kernel method for the two-sample-problem[C]//Proceedings of the 20th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2006: 513-520.
- [19] TZENG E, HOFFMAN J, ZHANG N, et al. Deep domain confusion: Maximizing for domain invariance[EB/OL]. (2014-12-10)[2025-01-06]. <http://arxiv.org/abs/1412.3474>.
- [20] LONG M, CAO Y, WANG J, et al. Learning transferable features with deep adaptation networks[C]//Proceedings of the 32nd International Conference on Machine Learning. Lille: PMLR, 2015: 97-105.
- [21] SUN B, FENG J, SAENKO K. Return of frustratingly easy domain adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Washington: AAAI Press, 2016: 2508-2605.
- [22] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]// Proceedings of the 32nd International Conference on Machine Learning. Lille: PMLR, 2015: 1180-1189.
- [23] LONG M, ZHU H, WANG J, et al. Deep transfer learning with joint adaptation networks[C]//International Conference on Machine Learning. Sydney: JMLR, 2017: 2208-2217.
- [24] ZHANG J, LI W, OGUNBONA P. Joint geometrical and statistical alignment for visual domain adaptation[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 5150-5158.
- [25] WANG J, CHEN Y, HAO S, et al. Balanced distribution adaptation for transfer learning[C]//2017 IEEE International Conference on Data Mining (ICDM). Piscataway: IEEE, 2017: 1129-1134.
- [26] WANG J, FENG W, CHEN Y, et al. Visual domain adaptation with manifold embedded distribution alignment [C]//ACM International Conference on Multimedia. New York: ACM, 2018: 402-410.
- [27] LI S, SONG S, HUANG G, et al. Domain invariant and class discriminative feature learning for visual domain adaptation[J]. *IEEE Transactions on Image Processing*, 2018, 27(9): 4260-4273.
- [28] LIANG J, HE R, SUN Z, et al. Aggregating randomized clustering-promoting invariant projections for domain adaptation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(5): 1027-1042.
- [29] PEI Z, CAO Z, LONG M, et al. Multi-adversarial domain

- adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Washington: AAAI Press, 2018: 3934-3941.
- [30] LONG M, CAO Z, WANG J, et al. Conditional adversarial domain adaptation[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2018: 1647-1657.
- [31] DENG Z, LUO Y, ZHU J. Cluster alignment with a teacher for unsupervised domain adaptation[C]//International Conference on Computer Vision. Piscataway: IEEE, 2019: 9943-9952.
- [32] YU C, WANG J, CHEN Y, et al. Transfer learning with dynamic adversarial adaptation network[C]//International Conference on Data Mining. Piscataway: IEEE, 2019: 778-786.
- [33] CHEN W, HU H. Generative attention adversarial classification network for unsupervised domain adaptation[J]. Pattern Recognition, 2020, 107: 107440.
- [34] YANG B, YUEN P C. Learning adaptive geometry for unsupervised domain adaptation[J]. Pattern Recognition, 2021, 110: 107638.
- [35] HUANG J, XIAO N, ZHANG L. Balancing transferability and discriminability for unsupervised domain adaptation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(4): 5807-5814.
- [36] SIM T, BAKER S, BSAT M. The CMU pose, illumination, and expression database[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(12): 1615-1618.
- [37] HUTCHISON D, KANADE T, KITTLER J, et al. Adapting Visual Category Models to New Domains[M]//Computer Vision - ECCV 2010. Berlin: Springer, 2010: 213-226.
- [38] DONAHUE J, JIA Y, VINYALS O, et al. DeCAF: A deep convolutional activation feature for generic visual recognition[C]//Proceedings of the 31st International Conference on Machine Learning. New York: ICML, 2013: 647-655.
- [39] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [40] FUKUNAGA K, NARENDRA P M. A branch and bound algorithm for computing k-nearest neighbors[J]. IEEE Transactions on Computers, 1975, 24(7): 750-753.
- [41] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//Proceedings of the 26nd International Confer-

ence on Neural Information Processing Systems. Cambridge, MA, USA: MIT Press, 2012: 1106-1114.

- [42] ZHUO J B, WANG S H, ZHANG W G, et al. Deep unsupervised convolutional domain adaptation[C]//Proceedings of the 25th ACM International Conference on Multimedia. New York: ACM, 2017: 261-269.
- ZHUO J B, WANG S H, ZHANG W G, et al. Deep unsupervised convolutional domain adaptation[C]//Proceedings of the 25th ACM International Conference on Multimedia. New York: ACM, 2017: 261-269. [LinkOut]

### 作者简介



**李普飞** 男, 1995年12月出生于河北省邯郸市. 重庆大学微电子与通信工程学院博士研究生. 主要研究方向为机器学习、迁移学习.

E-mail: lipufei@cqu.edu.cn



**王品** (通讯作者) 女, 1979年11月出生于江苏省盐城市. 现为重庆大学微电子与通信工程学院副教授, 硕士生导师. 主要研究领域为图像处理与识别.

E-mail: wangpin@cqu.edu.cn



**李勇明** 男, 1976年9月出生于四川省绵阳市. 现为重庆大学微电子与通信工程学院教授, 博士生导师. 主要研究方向为医学信号处理、机器学习.

E-mail: yongmingli@cqu.edu.cn



**张锦华** 男, 1999年1月出生于山东省济宁市. 重庆大学微电子与通信工程学院硕士研究生. 主要研究方向为深度学习与图像处理.

E-mail: 202312021035@cqu.edu.cn



**颜芳** 女, 1979年10月出生于甘肃省天水市. 重庆大学微电子与通信工程学院副教授. 主要研究领域为智能信息处理.

E-mail: yanfang@cqu.edu.cn