软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software, 2022,33(4):1170−1182 [doi: 10.13328/j.cnki.jos.006478]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于自监督知识的无监督新集域适应学习*

汪云云[1], 孙顾威[1], 赵国祥[2], 薛 晖[3]

[1](南京邮电大学 计算机学院, 江苏 南京 210023)
[2](南京邮电大学 自动化学院, 江苏 南京 210023)
[3](东南大学 计算机科学与工程学院, 江苏 南京 210096)
通信作者: 汪云云, E-mail: wangyunyun@njupt.edu.cn

摘 要: 无监督域适应(unsupervised domain adaptation, UDA)旨在利用带大量标注数据的源域帮助无任何标注信息的目标域学习. 在 UDA 中, 通常假设源域和目标域间的数据分布不同, 但共享相同的类标签空间. 但在真实开放学习场景中, 域间的标签空间很可能存在差异. 在极端情形下, 域间的类别不存在交集, 即目标域中类别都为新未知类别. 此时若直接迁移源域的类判别知识, 可能会损害目标域性能, 导致负迁移问题. 为此, 提出了基于自监督知识的无监督新集域适应(unsupervised new-set domain adaptation with self-supervised knowledge, SUNDA)方法, 迁移源域的样本对比知识; 同时, 利用目标域的自监督知识指导知识迁移. 首先, 通过自监督学习源域和目标域初始特征, 并固定部分网络参数用于保存目标域信息. 再将源域的样本对比知识迁移至目标域, 辅助目标域学习类判别特征. 此外, 利用基于图的自监督分类损失, 解决域间无共享类别时目标域的分类问题. 在手写体数字的无共享类别跨域迁移和人脸数据的无共享类别跨种族迁移任务上对 SUNDA 进行评估, 实验结果表明, SUNDA 的学习性能优于无监督域适应、无监督聚类以及新类别发现方法.

关键词: 无监督域适应; 自监督学习; 标签空间; 类对比知识
中图法分类号: TP181

中文引用格式: 汪云云, 孙顾威, 赵国祥, 薛晖. 基于自监督知识的无监督新集域适应学习. 软件学报, 2022, 33(4): 1170–1182. http://www.jos.org.cn/1000-9825/6478.htm

英文引用格式: Wang YY, Sun GW, Zhao GX, Xue H. Unsupervised New-set Domain Adaptation with Self-supervised Knowledge. Ruan Jian Xue Bao/Journal of Software, 2022, 33(4): 1170–1182 (in Chinese). http://www.jos.org.cn/1000-9825/6478.htm

Unsupervised New-set Domain Adaptation with Self-supervised Knowledge

WANG Yun-Yun[1], SUN Gu-Wei[1], ZHAO Guo-Xiang[2], XUE Hui[3]

[1](School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)
[2](College of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)
[3](School of Computer Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: Unsupervised domain adaptation (UDA) adopts source domain with large amounts of labeled data to help the learning of the target domain without any label information. In UDA, the source and target domains usually have different data distribution, but share the same label space. But in real open scenarios, label spaces between domains can also be different. In extreme cases, there is no shared class between domains, i.e., all classes in target domain are new classes. In this case, directly transferring the discriminant knowledge in source domain would harm the performance of target domain, lead to negative transfer. As a result, this study proposes an unsupervised new-set domain adaptation with self-supervised knowledge (SUNDA). Firstly, self-supervised learning is adopted to learn the initial features on

* 基金项目: 国家自然科学基金(61876091, 61772284, 62006126); 中国博士后科学基金(2019M651918); 工信部模式分析与机器智能重点实验室开放基金; 江苏省高等学校自然科学研究面上项目(0KJB520004)
本文由“面向开放场景的鲁棒机器学习”专刊特约编辑陈恩红教授、李宇峰副教授、邹权教授推荐.
收稿时间: 2021-05-31; 修改时间: 2021-07-16; 采用时间: 2021-08-27; jos 在线出版时间: 2021-10-26

source and target domains, with the first few layers frozen, in order to keep the target information. Then, the class contrastive knowledge from the source domain is transferred, to help learning discriminant features for target domain. Moreover, the graph-based self-supervised classification loss is adopted to handle the classification problem in target domain without common labels. Experiments are conducted over both digit and face recognition tasks without shared classes, and the empirical results show the competitive of SUNDA compared with UDA and unsupervised clustering methods, as well as new category discovery method.

Key words: unsupervised domain adaptation; self-supervised learning; label space; class contrastive knowledge

相较于传统机器学习模型, 基于大数据的深度学习模型能让预测性能得到显著提升. 然而, 深度网络等监督模型的训练往往依赖大量的有标记数据. 在真实学习任务中, 数据标注通常需要人工参与和相关的专业知识, 耗时耗力, 代价高昂; 此外, 学习模型对新环境和新任务的适应能力弱, 当面对与训练场景不同的学习环境时, 需对新数据进行标注, 并重新训练学习模型. 域适应(domain adaptation, DA)[1]学习旨在利用相关但不同的任务域辅助当前任务域的学习, 是缓解上述数据标注稀缺问题和提升模型适应能力的有效途径之一. 目前, 目标域中无任何标注信息的无监督域适应(unsupervised DA, UDA)是域适应学习中最具挑战和备受关注的研究方向.

在 UDA 中, 当前关注的任务域为目标域(target domain), 其他相关但不同的任务域为源域(source domain). 学习目标是: 将相关的源域知识迁移至目标域中, 以提升目标域的学习性能. 由于受到如光照、噪声、角度变化以及物体自身差异的影响, 目标域与源域的样本分布往往存在差异. 因此, 当前域适应工作大多致力于减少域间的样本分布差异, 包括边缘分布和条件分布差异, 通过分布差异最小化或域混淆等策略对齐域间的特征分布. 再通过源域经验风险最小化, 将源域中的类判别知识迁移至目标域中, 指导目标域的分类学习[1]. 这些工作通常假设领域间的类标签空间一致或同构.

而在现实的开放动态环境中,样本的类别无法恒定[2]. 此时, 域间的标签空间很可能存在差异, 源域或目标域中会出现域非共享类别. 假设域间仍存在部分共享类别, 源域中包含非共享类的学习问题称为部分域适应(partial DA, PDA)[2], 而目标域中包含新类别的学习问题为开集域适应(open-set DA, OSDA)[2]. 针对上述问题, 现有工作致力于识别域间的共享类别并迁移共享类知识[2]. 但在极端条件下, 域间可能不存在任何共享类别. 如在跨种族人脸迁移和识别中, 源域和目标域中人脸数据隶属于不同种族, 除数据分布差异外, 域间的类别也是完全无重叠的, 如图 1 所示. 为方便后续介绍, 本文称该类学习问题为新集域适应(new-set DA, NSDA). 传统域适应方法在解决 NSDA 时会有如下问题: 1) 由于源域与目标域不存在共享类, 直接进行域特征对齐和知识迁移可能会引起负迁移现象; 2) 域间类别集合无交集, 基于源域学习所获分类器无法用于目标域数据分类. NSDA 的学习目标与零次学习(zero-shot learning)[3]相似, 训练学习模型以识别从未见过的类别样例. 但零次学习通常需要辅助信息, 例如属性描述、文本描述或层次结构关系等[4], 而在 NSDA 学习中不会给定任何语义或属性信息. 此外, 新类别发现(new category discovery)[5]旨在对未见类别样本进行预测, 但在学习中不考虑已知类和新类别间的分布迁移.

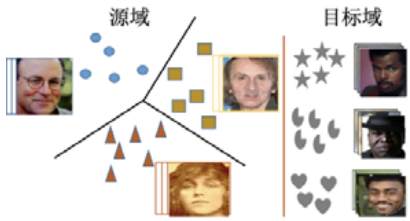


图 1 源域和目标域间无任何共享类别的域适应学习

针对 NSDA 学习场景, 本文提出了基于自监督知识的无监督新集域适应学习方法 SUNDA. 一方面, 源域和目标域间无任何共享类别, 对源域样本正确分类不一定有助于正确区分目标域中类别. 为此, 挖掘和迁移源域中的样本对比知识. 样本对比知识与具体的类别标签无关, 用于描述样本对是否源于同一类别. 将其迁

移至目标域, 帮助实现目标域样本对的有效判别, 生成强判别能力特征. 另一方面, 试图利用目标域的自监督知识指导源域知识迁移. 首先, 使用自监督预训练学习源域和目标域的底层特征, 并在后续训练中固定特征生成网络的前几层参数, 以保留目标域中有用信息. 进一步, 由于域间类别集合无交集, 域适应后源域分类器无法直接对目标域样本分类, 利用基于图的自监督学习实现目标域样本分类. 具体地, 将目标域样本多分类问题转化为判定样本对是否属于同类别的二分类问题, 并通过目标域的生成特征给样本对赋予伪标签, 实现目标域样本分类. SUNDA 方法的特点可总结如下.

1) 采用自监督预训练, 学习源域和目标域数据的底层特征, 以保存目标域信息;
2) 将源域的样本对比知识迁移至目标域, 辅助目标域生成类判别特征. 样本对比知识与具体类别标签无关, 因此源域和目标域无需包含共享类别;
3) 通过目标域中基于图的自监督知识指导知识迁移, 同时解决 NSDA 中目标域样本分类问题.

在数字和人脸数据集上, 将 SUNDA 与无监督域适应、无监督聚类以及新类发现算法进行性能对比, 实验验证了该方法的有效性.

1 相关工作

1.1 无监督域适应

域适应学习旨在将一个或多个不同但相关源域的知识迁移至目标域, 以解决目标域中样本标注稀缺等问题. 近年来, 域适应方法相继涌现, 可大致分为 3 类[1]: 基于差异的方法、基于对抗的方法以及基于重构的方法. 基于差异的域适应方法通过域间的分布差异最小化减少域间的分布距离. 常用的差异性度量包括最大均值差异(maximum mean discrepancy, MMD)[6]、KL 散度[7]和 Wasserstein 距离[8]等. 基于对抗的方法[9,10]通常利用域判别器(domain discriminator)区分域间样本, 同时, 训练特征提取器用于混淆域判别器, 生成域不变特征. Saito 等人[11]则通过两个源域分类器间的对抗学习, 实现域间的条件分布对齐. 基于重构的方法利用对抗方式生成其他域样本, 实现风格迁移, 如 cycle-GAN[12]和 CyCADA[13]等方法.

上述方式主要面向类标签空间一致的迁移学习场景. 而在现实学习任务中, 域间的标签空间很可能存在差异, 因此, 面向差异标签空间的迁移学习方法也在近期被相继提出[14–22]. 针对 PDA, Cao 等人[14]使用多个类级别和实例级别的权重机制实现各类的单独对抗; Zhang 等人[15]构造辅助域判别器用于评估源域样本和目标域间的相似性. ETN (example transfer network)方法[16]利用递进式加权策略度量源域样本的可迁移性. OSDA[17]旨在将目标域中的共享类样本正确分类, 同时拒绝新的未知类样本. 如: OSBP (open set domain adaptation by backpropagation)[17]通过对抗方式判别域共享和非共享类; SAT (separate to adapt)[18]采用由粗到细粒度的加权方式为目标域样本赋权重, 描述其属于域共享和私有类的概率. 通用域适应方法 UAN (universal domain adaptation)[21]对未知类的所属域别不作任何假设, 同时利用与源域的相似性和条件熵两种度量, 分别为源域和目标域样本赋权值, 以区分域共享和非共享类别. Saito 等人[22]基于自监督学习, 通过近邻聚类和最小化熵方式拉近域共享类, 同时分离域非共享类.

1.2 自监督学习

自监督学习(self-supervised learning, SSL)[23]是一种新的无监督学习范式, 旨在根据数据自身属性免费生成非目标标签, 以减少学习任务对大量标注数据的依赖. SSL 通过创建免人工标注的前置(pretext)任务为特征学习提供代理的监督信号, 获得有益于下游(downstream)任务的特征表示. SSL 的关键是前置任务的设计, 即如何依据数据属性自动生成免费的非目标标签. 依据属性, 自监督学习主要分为生成方法和对比方法. 生成方法主要针对图像或视频数据, 通过数据重建等方式人工构造数据标签, 并学习视觉特征, 如图像着色[23]、预测补丁的相对位置[24]、将图片裁剪后预测正确的拼图顺序[25]以及将图片随机旋转(0°, 90°, 180°, 270°)后预测旋转角度[26]等方法. 对比方法则关注样本间的对比知识, 旨在通过对比而非辅助标签学习特征, 实现样本间的有效区分. MOCO (momentum contrast)[27]将对比学习视为字典查找, 利用对比损失将编码查询与编码键字

典进行匹配, 用于训练视觉表示编码器; SimCLR[28]通过最大化同一数据示例的不同增强视图间一致性, 以及不同数据示例的增强视图间差异性, 学习特征表示.

2 方 法

为方便描述, 首先给出符号定义. 有标签源域数据和无标签目标域数据分别表示为 $(X^s,Y^s)=\{(x_i^s,y_i^s)\}_{i=1}^N$ 和 $X^t=\{x_j^t\}_{j=1}^M$, 其中, N 和 M 分别为源域和目标域样本个数. 源域数据给定类别标签 $y_i^s\in C^s=\{1,2...,r\}$, 而目标域类别标签 $y_i^t\in C^t$ 未知. 源域与目标域间无共享类别, 即 $C^s\cap C^t=\varnothing$. 上标符号 s 和 t 分别代表源域和目标域.

SUNDA 将源域中样本对比知识迁移至目标域, 同时利用目标域的自监督知识指导知识迁移, 网络结构如图 2 所示(实线箭头代表网络链接, 虚线箭头代表损失计算. 源域与目标域共享特征提取器 g 和域迁移网络 ft). 其中: g 为特征提取器; ft 为域迁移网络, 用于迁移源域中样本对比知识. 源域与目标域共享 g 和 ft. D 为域鉴别器, 用于区分源域和目标域样本. η_1 和 η_2 分别为源域和目标域分类器, 由单层全连接网络构成, 节点数分别为源域和目标域中类别数. $\mathcal{L}_{CE}$ 为源域分类损失. $\mathcal{L}_N$ 为源域 Npair 损失, 包括 $\mathcal{L}_N(g)$ 与 $\mathcal{L}_N(ft)$, 用于刻画源域中样本对比知识. $\mathcal{L}_D$ 为域判别损失. $\mathcal{L}_{BCE}$ 为目标域分类损失. 后续将一一阐述.

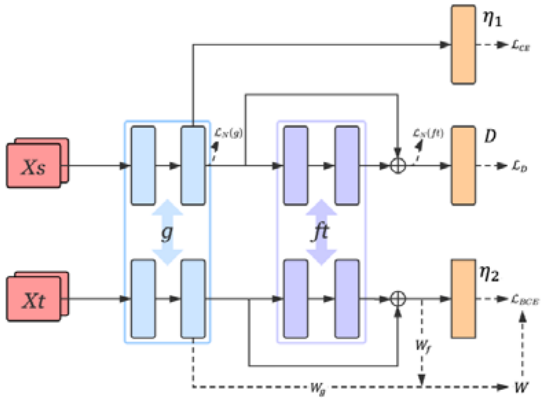


图 2 模型框架图

SUNDA 的学习流程描述如下.

1) 首先, 通过自监督学习 g 的初始参数, 并在后续学习中固定 g 的部分参数, 用于保存源域和目标域信息;
2) 使用源域数据计算类对比损失(图中 $\mathcal{L}_N$ 和 $\mathcal{L}_{CE}$), 通过特征提取器 g、域迁移网络 ft 与鉴别器 D 间的最小最大博弈, 将源域类对比知识迁移至目标域, 使目标域获得更有判别能力的特征;
3) 利用目标域特征构造伪标签矩阵 $\boldsymbol{W}$(图中 $\boldsymbol{W}^g$ 和 $\boldsymbol{W}^f$ 为不同层特征所得伪标签矩阵, 结合 $\boldsymbol{W}^g$ 和 $\boldsymbol{W}^f$ 计算 $\boldsymbol{W}$), 并通过伪标签矩阵计算基于自监督的分类损失, 实现目标域分类.

值得注意的是, 步骤 3)与步骤 2)同时进行.

2.1 自监督预训练

由于域间标签空间无交集, 源域知识迁移可能会使学习模型偏重源域学习, 导致目标域中有用信息的丢失. 自监督学习已被证明能在无类标注信息时产生有利于下游任务的底层特征. 鉴于此, SUNDA 利用自监督预训练提取源域与目标域的底层特征, 并在后续训练中固定网络中前几层参数, 用于保存目标域信息.

首先, 将源域和目标域数据合并 $X=X^s\cup X^t$, 并将样本随机旋转[26]得到 $(X,Y)=\{(x_i,\hat{y}_i)\}_i^{N+M}$, 其中, $N+M$ 为样本总数, $\hat{y}_i\in\{0,1,2,3\}$ 分别代表将 x_i 旋转 0°, 90°, 180°和 270°. 最小化交叉熵损失, 使旋转后样本与其基于旋转角度的类别标签相关联:

$$\mathcal{L}_{pre}=-\frac{1}{N+M}\sum_{i=1}^{N}\hat{y}_i\log\eta(z_i) \tag{1}$$

其中, z_i 为样本 x_i 经特征提取器 g 的生成特征, 即 $z_i=g(x_i)$; 同 η_1 和 η_2, η为包含 4 个节点的单层分类网络. 同理, 其他自监督方式, 如对比型 simCLR 和 MoCo 等方法, 也可用于提取底层特征. 这里仅以旋转方式为例, 验证 SUNDA 的有效性.

2.2 基于样本对比知识迁移的域适应

由于源域和目标域间无任何共享类别, 本文将源域与具体类别标签无关的类对比知识迁移至目标域, 以产生强判别能力特征.

- 类对比知识

源域 X^s 先后经特征提取器 g 和域迁移网络 ft 生成特征 $g(X^s)$和 $ft(g(X^s))$, 在两种特征上分别计算 Npair 损失[29], 并结合源域交叉熵损失, 得到基于源域的优化目标:

$$\mathcal{L}_{constrastive}=\frac{1}{2}(\mathcal{L}_N(g)+\mathcal{L}_N(ft))+\mathcal{L}_{CE}(X^s,Y^s) \tag{2}$$

其中,

$$\mathcal{L}_N(g)=-\mathbb{E}_{\{x_n^s,x_n^{s+}\}_{n=1}^N}\sum_{i=1}^{N}\log p_n(g),\ \mathcal{L}_N(ft)=-\mathbb{E}_{\{x_n^s,x_n^{s+}\}_{n=1}^N}\sum_{i=1}^{N}\log p_n(ft) \tag{3}$$

$$p_n(g)=\frac{\exp(g(x_n^s)^T g(x_n^{s+}))}{\sum_{k=1}^{N}\exp(g(x_n^s)^T g(x_k^{s+}))},\ p_n(ft)=\frac{\exp(ft(g(x_n^s))^T ft(g(x_n^{s+})))}{\sum_{k=1}^{N}\exp(ft(g(x_n^s))^T ft(g(x_k^{s+})))} \tag{4}$$

公式(4)中, x_n 和 x_n^+ 为属于同一类别样本. 当 $n=k$ 时, x_n 和 x_k^+ 属于同类; 反之属于不同类. $\mathcal{L}_{CE}$ 为源域交叉熵损失. Npair 损失$\mathcal{L}_N$ 能够保证类间间隔最大化, 使生成特征更具判别能力[29].

- 对抗域适应

由于域间无共享类别, SUNDA 采用不同于 DANN 的非对称域对齐方式(若 $g(x)=ft(g(x))$, 则与 DANN 相同). 具体地, 训练域判别器 D 用于区分经域迁移网络 ft 的源域数据和经特征提取器 g 的目标域数据, 同时训练特征提取器 g 和域迁移网络 ft 以混淆域判别器. 非对称域判别损失刻画如下:

$$\left.\begin{aligned}\mathcal{L}_D&=-\mathbb{E}_{x\in X^s}\log D(ft(g(x)))-\mathbb{E}_{x\in X^t}\log(1-D(g(x)))\\ \mathcal{L}_{adv}&=-\mathbb{E}_{x\in X^t}\log D(g(x))\end{aligned}\right\} \tag{5}$$

2.3 基于图的自监督分类

在获得目标域判别型特征的同时, 需对目标域数据进行分类. SUNDA 通过基于图的自监督学习方式获得目标域样本对的伪标签, 用于指导知识迁移, 同时解决 NSDA 中目标域分类问题.

- 基于图的相似矩阵

基于 cosine 距离计算目标域特征 $g(X^t)$和 $ft(g(X^t))$上样本对的相似性矩阵 $\boldsymbol{S}^g$ 和 $\boldsymbol{S}^f$, 公式如下:

$$S_{ij}^g=\frac{z_i^g\cdot z_j^g}{\|z_i^g\|_2\|z_j^g\|_2},\ S_{ij}^f=\frac{z_i^f\cdot z_j^f}{\|z_i^f\|_2\|z_j^f\|_2} \tag{6}$$

其中, $\cdot$为向量点积, z_i^g 为目标域 x_i^t 经特征提取器 g 后的特征 $z_i^g=g(x_i^t)$, z_i^f 为目标域 x_i^t 经特征提取器 g 和域迁移网络 ft 后的特征 $z_i^f=ft(g(x_i^t))$.

- 基于图的伪标签矩阵

获得相似性矩阵 $\boldsymbol{S}^g$ 和 $\boldsymbol{S}^f$ 后, 通过阈值法[28]构造样本对伪标签矩阵 $\boldsymbol{W}^g$ 和 $\boldsymbol{W}^f$:

$$W_{ij}^g=\begin{cases}1, & \text{if } S_{ij}^g\geqslant thres\\ 0, & \text{otherwise}\end{cases},\ W_{ij}^f=\begin{cases}1, & \text{if } S_{ij}^f\geqslant thres\\ 0, & \text{otherwise}\end{cases} \tag{7}$$

若两样本间相似性大于所设阈值 $thres$, 则属于同一类别($W_{ij}^g=1$或$W_{ij}^f=1$). 结合伪标签矩阵 $\boldsymbol{W}^g$ 和 $\boldsymbol{W}^f$, 目

标域伪标签矩阵可表示为

$$\boldsymbol{W}=(\boldsymbol{W}^g \geqslant thres)|(\boldsymbol{W}^f \geqslant thres) \tag{8}$$

其中, $\boldsymbol{A}|\boldsymbol{B}$ 代表 $\boldsymbol{A}$, $\boldsymbol{B}$ 矩阵内对应位置的逻辑或运算. 若 W_{ij}=1(True), 表明样本 x_i 和 x_j 属于同类; 反之, W_{ij}=0 (False), 则 x_i 和 x_j 属于不同类别. 实际上, 自监督预训练有助于提升初始伪标签矩阵的准确率.

- 基于伪标签的分类损失

获得伪标签矩阵 $\boldsymbol{W}$ 后, 希望通过伪标签实现目标域分类. 在目标域分类器 η_2 的输出空间上计算样本内积 $\eta_2(z_i^t)^T\eta_2(z_j^t)$, 其中, $z_i^t = ft(g(x_i^t))$, $\eta_2(z_i^t)$ 为经过 softmax 层归一化的概率向量. 该内积代表样本 x_i^t 和 x_j^t 在输出空间的相似性. 结合伪标签矩阵 $\boldsymbol{W}$, 可得目标域分类损失:

$$\mathcal{L}_{BCE} = -\frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}[W_{ij}\log\eta_2(z_i^t)^T\eta_2(z_j^t) + (1-W_{ij})(\log(1-\eta_2(z_i^t)^T\eta_2(z_j^t)))] \tag{9}$$

- 基于数据增强的伪标签分类损失

进一步改造公式(9), 提出了基于数据增强的 $\mathcal{L}_{BCE}$ 损失. 定义 $\hat{x}_i$ 为 x_i 经数据增强后样本, 在数据增强后的样本 $\hat{x}_i$ 和原始样本 x_j 的上计算内积 $\eta_2(\hat{z}_i^t)^T\eta_2(z_j^t)$. 最终将上述 $\mathcal{L}_{BCE}$ 改造为

$$\mathcal{L}_{BCE} = -\frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}[W_{ij}\log\eta_2(\hat{z}_i^t)^T\eta_2(z_j^t) + (1-W_{ij})(\log(1-\eta_2(\hat{z}_i^t)^T\eta_2(z_j^t)))] \tag{10}$$

2.4 一致性鲁棒损失

鲁棒的学习模型应对同一样本的不同变换保持输出不变性[30,31], 一致性损失也普遍应用于半监督学习中[32]. 给定样本 x 和变换 T, 定义变换后的样本 $\hat{x} = T\cdot x$, 学习模型应保证 x 和 $\hat{x}$ 间输出一致性. 为此, 约束学习中 x_i^t 和增强后 $\hat{x}_i^t$ 的输出一致性. 一致性鲁棒损失如下:

$$\mathcal{L}_{con} = \frac{1}{N}\sum_{i=1}^{N}(\eta_1(z_i^s)-\eta_1(\hat{z}_i^s))^2 + \frac{1}{M}\sum_{i=1}^{M}(\eta_1(z_i^t)-\eta_1(\hat{z}_i^t))^2 \tag{11}$$

其中, $z_i^s = g(x_i^s)$, $\hat{z}_i^s = g(\hat{x}_i^s)$, $z_i^t = ft(g(x_i^t))$, $\hat{z}_i^t = ft(g(\hat{x}_i^t))$.

最终, SUNDA 的优化目标可刻画如下:

$$\mathcal{L}_{g,ft,\eta_1,\eta_2} = \mathcal{L}_{constrastive} + \lambda\mathcal{L}_{adv} + \alpha\mathcal{L}_{BCE} + \omega(t)\mathcal{L}_{con} \tag{12}$$

$$\mathcal{L}_D = -\mathbb{E}_{x\in X^s}\log D(ft(g(x))) - \mathbb{E}_{x\in X^t}\log(1-D(g(x))) \tag{13}$$

最小化 $\mathcal{L}_{g,ft,\eta_1,\eta_2}$ 用于优化参数 $\boldsymbol{\theta}_g$, $\boldsymbol{\theta}_{ft}$, $\boldsymbol{\theta}_{\eta_1}$ 和 $\boldsymbol{\theta}_{\eta_2}$, 最小化 $\mathcal{L}_D$ 以优化 $\boldsymbol{\theta}_D$. 训练过程采用交替优化方式. $\omega(t)$ 为 ramp-up 函数[33], 广泛用于半监督学习中, 这里使用 sigmoid-shaped 函数 $\omega(t) = \varepsilon e^{-5\left(1-\frac{t}{T}\right)^2}$. t 是当前的时间步长, T 是 ramp-up 长度并且 $\varepsilon\in\mathbb{R}_+$. SUNDA 方法的训练流程如算法 1 所示.

算法 1. 基于自监督知识的无监督新集域适应学习.

输入: 有标签源域数据 $(X^s,Y^s)=\{(x_i^s,y_i^s)\}_{i=1}^N$ 和无标签目标域数据 $X^t=\{x_j^t\}_{j=1}^M$;

输出: 目标域数据 $X^t=\{x_j^t\}_{j=1}^M$ 的分类结果.

//自监督预训练

1. 将源域和目标域数据合并, 且旋转得到 $(X,Y)=\{(x_i,\hat{y}_i)\}_i^{N+M}$
2. **for each** epoch **do**
3. $\mathcal{L}_{pre}\leftarrow(X,Y)$ //Eq.1.
4. 更新 $\boldsymbol{\theta}_g$, $\boldsymbol{\theta}_\eta$

//源域样本对比知识迁移和目标域分类

5. **for each** epoch **do**
6. 计算伪标签矩阵 $\boldsymbol{W}$ //Eq.8.

```
    for each minibatch do
        //交替优化，实现对抗
        L_D ← (X^s, Y^t)    // Eq.13.
        更新 θ_D
        L_{g,ft,η1,η2} ← (X^s, X^t, Y^s, ω(t), λ, α)    //Eq.12
        更新 θ_g, θ_ft, θ_η1, θ_η2
    end
end
```

3 实验与结果

本文实验分为 3 个部分：首先，在数字数据集和人脸数据集上与现有方法进行性能对比；其次，基于消融实验验证 SUNDA 模型中各损失的有效性；最后，对阈值参数 *thres* 进行敏感性分析.

据笔者所知，在 NSDA 任务下，目前尚无基于域适应的目标域分类方法. 因此，将 SUNDA 与无监督域适应 DANN、无监督聚类 *K*-means、DEC、DAC、JULE-SF 和 DCC，以及发现新类 AutoNovel 方法进行性能对比. DANN 面向 closeset 问题，使用对抗学习实现域间分布对齐. 在无监督聚类方法中，*K*-means 为非深度聚类方法，其他方法均为深度学习方法. DEC 使用自编码器重构损失进行预训练，使用聚类分配强化损失对网络进行微调，同时，通过最小化软标签分布和辅助目标分布间的 KL 散度改善聚类结果. DAC 将聚类转化为成对样本的二分类问题. JULE-SF 使用 CNN 特征进行层次聚类以确定类别标签，同时根据聚类特点设计损失函数，对 CNN 特征进行调整. DCC 采用连续聚类(RCC)，无需聚类数目的先验信息. 发现新类方法 AutoNovel 采用自监督预训练和伪标签方式实现目标域分类.

由于域间无共享类别，且目标域的类别标签未知，无法将预测标签与真实标签一一对应. 如预测标签的第五类可能对应真实标签的第一类. 为此，采用聚类准确率(ACC)评估算法的学习性能，定义如下：

$$\mathrm{ACC}=\max_{m}\frac{\sum_{i=1}^{n}1\{l_i=m(c_i)\}}{n} \tag{14}$$

其中，l_i 为样本的真实标签，c_i 是预测标签，m 为所有可能的聚类和类别间映射. 通过 Hungarian 算法[34]进行优化，求解不同映射下的最大 ACC 值.

3.1 数字数据集分类任务

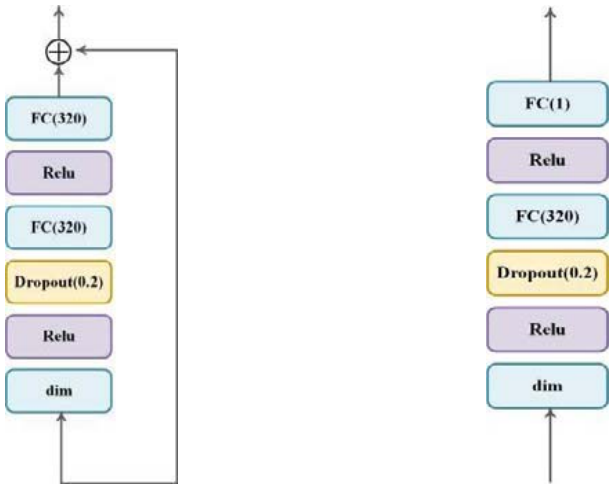
本组实验研究 SUNDA 在 MNIST，USPS 和 SVHN 数据集上，无域间共享类别下的迁移性能. 实验选择 4 组跨域分类实验来评估聚类准确率：1) MNIST 到 USPS; 2) USPS 到 MNIST; 3) SVHN 到 MNIST; 4) SVHN 到 USPS. 每组实验将数字 0~4 作为源域，5~9 作为目标域.

MNIST 由 250 个不同人的手写数字构成，包括 60 000 张训练图片和 10 000 张测试图片，每张图片的大小为 1×28×28(通道×长×宽). USPS 是从美国邮政采集的手写数字，包括 999 张训练图片和 250 张测试图片，每张图片的大小为 1×16×16. SVHN 是通过谷歌街景采集的门牌号数据，包括 4 578 张训练图片和 1 627 张测试图片，每张图片的大小为 3×32×32. 实验中，将 MNIST 和 USPS 放缩至 1×32×32，使用通道拷贝将单通道转变为三通道 3×32×32，每张的图片像素都归一化至[0,1]之间.

实验使用 ResNet-18 作为特征提取器. 域迁移网络使用两层全连接网络(dim-320-320-dim)和残差网络将输入向量映射为维度相同的向量，如图 3(a)所示，dim 为输入维度. 域分类网络包含两层全连接层(dim-320-320-1)，如图 3(b)所示. 所有参数采用 SGD 进行更新，学习速率为 3e−3，阈值 *thres* 为 0.95，超参数λ和α分别使用 0.3 和 1.

实验将 SUNDA 与无监督域适应和无监督聚类算法进行对比，结果见表 1. 表中第 1 列(METHOD)代表对比方法，最后一列(avg)为不同方法的平均性能，中间列为各方法在各迁移任务上聚类准确率. 基于表 1 可发

现: 由于域间无共享类别, DANN 实现域分布对齐和源域判别知识迁移, 并不能获得令人满意的学习性能. 而关注对比知识迁移, 同时使用非对称分布对齐, SUNDA 获得了显著的提升. 相比于非深度聚类 *K*-means, SUNDA 同样也有显著的性能提升. 但相比于其他深度聚类的方法, 如 DCC, JULE-SF 等, DANN 性能仅有小幅提升. 原因在于数字数据相对简单, 网络提取的特征容易区分, 因此无域适应的深度聚类方法已经能够获得比较好的学习效果. 接下来, 将利用更为复杂的人脸数据集, 验证 SUNDA 的有效性.



(a) 域迁移网络 *ft* 的结构示意图 (b) 域鉴别网络 *D* 的结构示意图

图 3

表 1 数字数据集目标域分类准确率(最好的准确率进行加粗)

Method	USPS to MNIST	SVHN to MNIST	SVHN to USPS	MNIST to USPS	Avg.
K-means	0.534	0.534	0.460	0.460	0.497
DCC[35]	**0.962**	0.962	–	–	0.962
DEC[36]	0.859	0.859	0.619	0.619	0.739
JULE-SF[37]	0.959	0.959	0.922	0.922	0.941
DANN[9]	0.374	0.270	0.372	0.369	0.346
SUNDA (ours)	0.955	**0.978**	**0.951**	**0.983**	**0.967**

此外, 验证了 SUNDA 在域间存在交集场景下的有效性. 在域间无交集场景中, 源域包含 0~4 数字类别, 目标域包含 5~9 数字类别. 为产生交集类别, 将源域类别依次去除{0,1,2,3,4}类, 同时增加{5,6,7,8,9}类, 分别设计 5 组实验场景, 如表 2 所示.

表 2 域间存在交集的 5 种场景

源域	目标域	重叠类别	重叠比例(%)
{1,2,3,4,5}	{5,6,7,8,9}	{5}	20
{2,3,4,5,6}	{5,6,7,8,9}	{5,6}	40
{3,4,5,6,7}	{5,6,7,8,9}	{5,6,7}	60
{4,5,6,7,8}	{5,6,7,8,9}	{5,6,7,8}	80
{5,6,7,8,9}	{5,6,7,8,9}	{5,6,7,8,9}	100

图 4 展示了 USPS 到 MINIST 任务上各对比方法在上述场景下的学习性能, 其中, 横轴为重叠比例, 纵轴为学习性能. 可以看出: 1) 因目标域类别恒定, 聚类方法的学习性能保持不变. 而随域间重叠类别逐渐增多, DANN 和 SUNDA 的学习性能都逐渐提升. 因此, 域间重叠知识越多, 源域知识迁移对目标域学习越有利; 2) 在该学习任务上, 当域间无重叠类别时, SUNDA 的性能稍差于深度聚类方法 DCC 的性能. 但随重叠类别逐渐增加, SUNDA 的性能不断提升, 超越 DCC 性能, 验证了源域相关知识的合理利用可有效辅助目标域学习; 3) 在上述 5 组实验场景下, SUNDA 的性能皆优于 DANN 的性能, 甚至包括类别完全重叠场景, 验证了 SUNDA 的有效性.

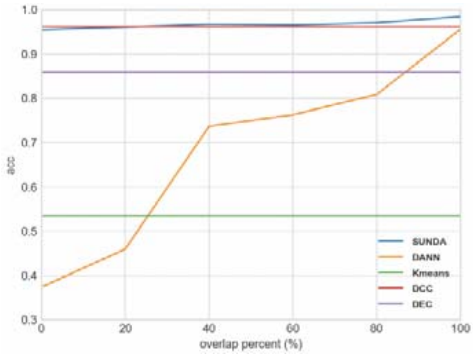


图 4 域间类别存在交集下性能对比

3.2 人脸数据集分类任务

进一步，在跨种族人脸数据上进行对比实验．人脸数据按照人种划分域，包括白人、黑人和黄种人．两两配对，共设计 6 组迁移任务评估 SUNDA 的学习性能: 1) 白人到黑人; 2) 黑人到白人; 3) 白人到黄种人; 4) 黄种人到白人; 5) 黑人到黄种人; 6) 黑人到白人．

MS-1M[38]为微软公司从网络采集的人脸照片，本组实验所使用人脸数据为从 MS-1M 中采样并对种族信息进行标注．其中，白人包含 13 类共 1 025 张图片，黑人包含 10 类共 739 张图片，黄种人包含 10 类共 825 张图片．将所有图片缩放至 3×110×110 并将像素归一化至[0,1]之间．人脸图片受到光线、角度、年龄等因素影响，相比于数字数据集场景更加复杂，更具挑战性．

特征提取器的网络结构与文献[39]中相似，如表 3 所示．域迁移网络和域分类网络与数字数据集任务中相同，如后文图 5 所示．所有参数采用 SGD 进行更新，学习率为 9e–3，阈值 *thres* 为 0.9，超参数λ和α分别使用 0.3 和 5.

表 3 人脸数据集实验使用的特征提取器 g 的结构

Operation	Kernel	Output size
Conv 1-1+ReLU	3×3	100×100×32
Conv 1-2+Maxout(2)	3×3	100×100×64
max pooling	2×2	50×50×64
ResBlock+ReLU×2	3×3, 64–64–64	50×50×64
Conv 2+Maxout(2)	3×3	50×50×128
max pooling	2×2	25×25×128
ResBlock+ReLU×4	3×3, 128–96–128	25×25×128
Conv 3+Maxout(2)	3×3	25×25×192
max pooling	2×2	13×13×192
ResBlock+ReLU×8	3×3, 192–128–192	13×13×192
Conv 4+Maxout(2)	3×3	13×13×256
max pooling	2×2	7×7×256
ResBlock+ReLU×2	3×3, 256–160–256	7×7×256
Conv 5+Maxout(2)	3×3	7×7×320
avg pooling	7×7	1×1×320
Normalize and Scale(2)	–	320

实验将 SUNDA 与无监督域适应、无监督聚类算法和发现新类方法进行对比，实验结果见表 4．从表 4 中可以看出，SUNDA 在所有迁移任务上的学习性能都显著优于对比方法的性能．与数字数据集不同，人脸场景复杂且不同类别间相似度大，因此实验中所有聚类算法均无法获得较好的学习性能．发现新类方法 AutoNovel 采用伪标签方式对目标域进行划分，而源域与目标域为跨种族人脸数据，域间差异较大，源域数据训练下的网络(无样本对比知识迁移)对目标域伪标签的获取可能并无帮助，甚至产生负面影响，最终导致学习性能差于 *K*-means 的性能．

表 4 人脸数据集目标域分类准确率

Method	Y to W	B to W	W to B	Y to B	W to Y	B to Y	Avg.
K-means	0.282	0.282	0.200	0.200	0.325	0.325	0.269
DEC[36]	0.260	0.260	0.192	0.192	0.315	0.315	0.256
DAC[40]	0.302	0.302	0.263	0.263	0.380	0.380	0.315
DANN[9]	0.218	0.207	0.264	0.227	0.225	0.294	0.239
AutoNovel[41]	0.200	0.191	0.196	0.176	0.204	0.217	0.197
SUNDA (ours)	**0.601**	**0.632**	**0.743**	**0.655**	**0.805**	**0.734**	**0.695**

综上, 在数字数据集和人脸数据集上, SUNDA 都取得了较好的学习性能. 在数字数据集上, 与 DCC 进行对比, SUNDA 的学习性能并没有很明显的优势. 而在人脸数据集上, SUNDA 相较于无监督聚类、无监督域适应以及发现新类方法均取得明显的性能提升. 此外, 在数字数据集上, 无监督聚类方法获得了较好的学习性能, 但到人脸数据集并不能很好地发挥作用. 但与此同时, SUNDA 依然可以获得较好的学习效果. 由此可见, SUNDA 在 NSDA 任务中能够充分利用源域知识辅助目标域识别, 提升目标域学习性能.

3.3 消融实验

为了进一步验证方法每一部分的有效性, 本组实验将在白人到黄种人的场景下进行消融实验. 实验包括 4 个损失: 1) 类对比损失; 2) 对抗域适应损失($\mathcal{L}_{gan}$ 包含了 $\mathcal{L}_{adv}$ 和 $\mathcal{L}_{D}$ 两个损失的对抗); 3) 目标域基于图自监督的分类损失; 4) 一致性损失. 消融实验结果见表 5. ACC 为分类准确率, NMI 为标准化互信息, ARI 为调整兰德系数. NMI 和 ARI 常用于评估聚类的好坏. ACC、NMI 和 ARI 的值越高代表性能越好. 通过表 4 可发现, 域迁移网络 *ft* 的对抗损失 $\mathcal{L}_{gan}$、类对比损失 $\mathcal{L}_{constrastive}$ 和一致性鲁棒损失 $\mathcal{L}_{con}$ 均能够带来明显的效果提升.

表 5 消融实验

Method	Metrics		
	ACC	NMI	ARI
$\mathcal{L}_{BCE}$	0.454	0.470	0.327
$\mathcal{L}_{gan}+\mathcal{L}_{BCE}$	0.625	0.574	0.481
$\mathcal{L}_{constrastive}+\mathcal{L}_{gan}+\mathcal{L}_{BCE}$	0.730	0.669	0.598
$\mathcal{L}_{constrastive}+\mathcal{L}_{gan}+\mathcal{L}_{BCE}+\mathcal{L}_{con}$	**0.805**	**0.761**	**0.716**

3.4 参数敏感性分析

本节对阈值参数 *thres* 对 SUNDA 学习性能的影响进行分析, 通过参数 *thres* 挑选正样本对, 用于基于图的自监督分类. 在人脸数据 6 组迁移任务上进行实验, 实验结果如图 5 所示. 其中, Y, W 和 B 分别代表黑人、白人和黄种人的人脸数据. 横坐标为 *thres* 的值, 纵坐标为 SUNDA 的学习性能. 从图可看出: 随着 *thres* 的增加, SUNDA 的分类效果整体逐步变好. 原因在于: 本文基于 cosine 距离, 通过 *thres* 挑选正样本对形成样本对伪标签, 随 *thres* 增大, 样本对的准确率将逐步提升, 从而使 SUNDA 的分类愈加准确.

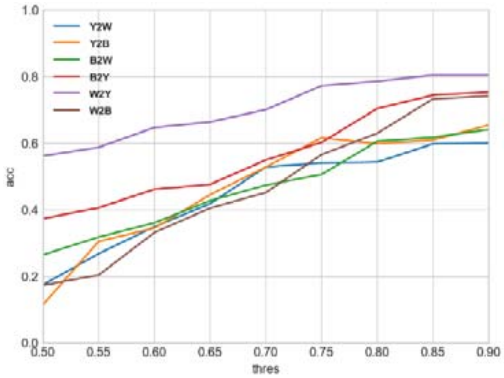
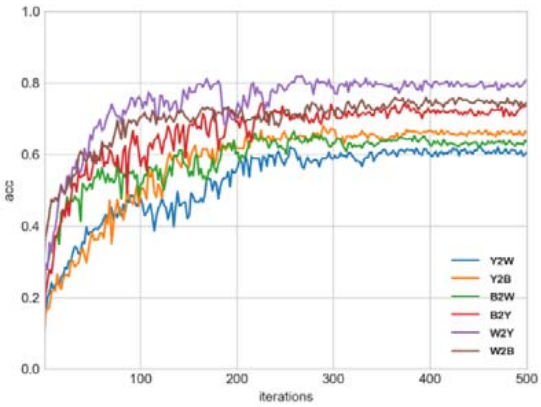


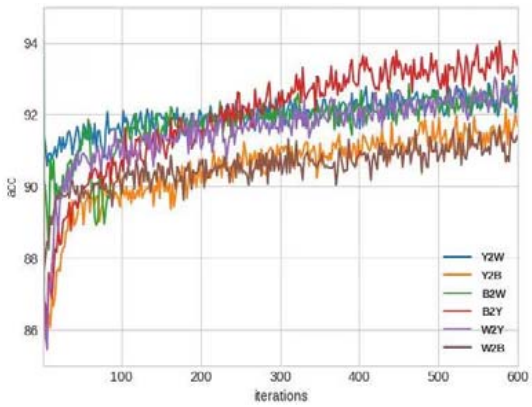
图 5 参数 *thres* 取值对准确率的影响

3.5 方法稳定性分析

为探究SUNDA的稳定性，绘制了在人脸数据集的6组迁移任务上，SUNDA学习性能随迭代的变化图，如图 6(a)所示．其中，横坐标为迭代次数，纵坐标为 SUNDA 的分类性能．由图可知: SUNDA 的学习性能随迭代次数的增加，基本呈上升趋势，并逐渐趋于稳定，验证了 SUNDA 的稳定性．此外，SUNDA 利用基于图的自监督分类，指导源域知识迁移，其中，伪标签矩阵由样本对 cosine 距离和阈值法获得．图 6(b)给出伪标签矩阵准确率随迭代的变化图，其中，横坐标为迭代次数，纵坐标为样本对伪标签准确率．由图可知: 伪标签矩阵的初始准确率基本在 80%以上，且随迭代的进行，准确率呈不断上升趋势，可对源域知识迁移进行有效指导，进一步验证了基于图的自监督分类的有效性以及方法的稳定性．



(a) SUNDA 学习性能随迭代变化图



(b) 伪标签矩阵准确率随迭代变化图

图 6

4 结束语

针对无监督域适应中，源域与目标域类别标签集合完全不同的 NSDA 学习任务，本文提出了 SUNDA 方法．首先，采用自监督学习初始特征，通过对抗迁移源域的类对比知识，帮助目标域学习判别特征．同时采用了目标域基于图的自监督来指导模型的学习，并解决了异构标签空间下目标域的分类问题．实验结果表明: 与无监督域适应和无监督聚类方法以及新类发现方法相比较，SUNDA 的学习效果得到了显著的提升．此外，在开放集域适应中，当前学习方法将目标域中的私有类都归为未知类．实际上，可以考虑将以往开放集域适应的方法与本文相结合，在对共享类实现对齐和分类的同时，实现域私有类准确分类．

References:

[1] Zhuang F, Qi Z, Duan K, *et al*. A comprehensive survey on transfer learning. Proc. of the IEEE, 2020, 109(1): 43–76.

[2] Wilson G, Cook DJ. A survey of unsupervised deep domain adaptation. ACM Trans. on Intelligent Systems and Technology (TIST), 2020, 11(5): 1–46.

[3] Wang W, Zheng VW, Yu H, *et al*. A survey of zero-shot learning: Settings, methods, and applications. ACM Trans. on Intelligent Systems and Technology (TIST), 2019, 10(2): 1–37.

[4] Pang N, Zhao X, Wang W, *et al*. Few-shot text classification by leveraging bi-directional attention and cross-class knowledge. Sciece China Information Sciences, 2021, 64(3): 29–41.

[5] Han K, Vedaldi A, Zisserman A. Learning to discover novel visual categories via deep transfer clustering. In: Proc. of the IEEE/CVF Int’l Conf. on Computer Vision. 2019. 8401–8409.

[6] Long M, Cao Y, Wang J, *et al*. Learning transferable features with deep adaptation networks. In: Proc. of the Int’l Conf. on Machine Learning. PMLR, 2015. 97–105.

[7] Kullback S, Leibler RA. On information and sufficiency. Annals of Mathematical Statistics, 1951, 22(1): 79–86.

[8] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. In: Proc. of the Int'l Conf. on Machine Learning. PMLR, 2017. 214–223.

[9] Ganin Y, Ustinova E, Ajakan H, *et al*. Domain-adversarial training of neural networks. Journal of Machine Learning Research, 2016, 17(59): 1–35.

[10] Wang YY, Gu JM, Wang C, *et al*. Discrimination-aware domain adversarial neural network. Journal of Computer Science and Technology, 2020, 35: 259–267.

[11] Saito K, Watanabe K, Ushiku Y, *et al*. Maximum classifier discrepancy for unsupervised domain adaptation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2018. 3723–3732.

[12] Zhu JY, Park T, Isola P, *et al*. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2017. 2223–2232.

[13] Hoffman J, Tzeng E, Park T, *et al*. Cycada: Cycle-consistent adversarial domain adaptation. In: Proc. of the Int'l Conf. on Machine Learning. PMLR, 2018. 1989–1998.

[14] Cao Z, Long M, Wang J, *et al*. Partial transfer learning with selective adversarial networks. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2018. 2724–2732.

[15] Zhang J, Ding Z, Li W, *et al*. Importance weighted adversarial nets for partial domain adaptation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2018. 8156–8164.

[16] Cao Z, You K, Long M, *et al*. Learning to transfer examples for partial domain adaptation. In: Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition. 2019. 2985–2994.

[17] Panareda Busto P, Gall J. Open set domain adaptation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2017. 754–763.

[18] Liu H, Cao Z, Long M, *et al*. Separate to adapt: Open set domain adaptation via progressive separation. In: Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition. 2019. 2927–2936.

[19] Saito K, Yamamoto S, Ushiku Y, *et al*. Open set domain adaptation by backpropagation. In: Proc. of the European Conf. on Computer Vision (ECCV). 2018. 153–168.

[20] Jo I, Kim J, Kang H, *et al*. Open set recognition by regularising classifier with fake data generated by generative adversarial networks. In: Proc. of the 2018 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018. 2686–2690.

[21] You K, Long M, Cao Z, *et al*. Universal domain adaptation. In: Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition. 2019. 2720–2729.

[22] Saito K, Kim D, Sclaroff S, *et al*. Universal domain adaptation through self supervision. arXiv: 2002.07953, 2020.

[23] Zhang R, Isola P, Efros AA. Colorful image colorization. In: Proc. of the European Conf. on Computer Vision. Cham: Springer, 2016. 649–666.

[24] Doersch C, Gupta A, Efros AA. Unsupervised visual representation learning by context prediction. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2015. 1422–1430.

[25] Noroozi M, Favaro P. Unsupervised learning of visual representations by solving jigsaw puzzles. In: Proc. of the European Conf. on Computer Vision. Cham: Springer, 2016. 69–84.

[26] Gidaris S, Singh P, Komodakis N. Unsupervised representation learning by predicting image rotations. arXiv: 1803.07728, 2018.

[27] He K, Fan H, Wu Y, *et al*. Momentum contrast for unsupervised visual representation learning. In: Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition. 2020. 9729–9738.

[28] Chen T, Kornblith S, Norouzi M, *et al*. A simple framework for contrastive learning of visual representations. In: Proc. of the Int'l Conf. on Machine Learning. PMLR, 2020. 1597–1607.

[29] Sohn K. Improved deep metric learning with multi-class *n*-pair loss objective. In: Proc. of the 30th Int'l Conf. on Neural Information Processing Systems. 2016. 1857–1865.

[30] Li YF, Liang DM. Safe semi-supervised learning: A brief introduction. Frontiers of Computer Science, 2019, 13(4): 669–676.

[31] Feng J, Wang X, Liu W. Deep graph cut network for weakly-supervised semantic segmentation. Science China Information Sciences, 2021, 64(3): 130105.

[32] Oliver A, Odena A, Raffel C, *et al*. Realistic evaluation of deep semi-supervised learning algorithms. arXiv: 1804.09170, 2018.

[33] Laine S, Aila T. Temporal ensembling for semi-supervised learning. arXiv: 1610.02242, 2016.

[34] Kuhn HW. The Hungarian method for the assignment problem. Naval Research Logistics, 2010, 52(1-2): 7–21.

[35] Shah SA, Koltun V. Deep continuous clustering. arXiv: 1803.01449, 2018.

[36] Xie J, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. In: Proc. of the Int'l Conf. on Machine Learning. 2016.

[37] Yang J, Parikh D, Batra D. Joint unsupervised learning of deep representations and image clusters. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 5147–5156.

[38] Guo Y, Zhang L, Hu Y, *et al*. Ms-celeb-1m: Challenge of recognizing one million celebrities in the real world. Electronic Imaging, 2016, 2016(11): 1–6.

[39] Sohn K, Liu S, Zhong G, *et al*. Unsupervised domain adaptation for face recognition in unlabeled videos. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2017. 3210–3218.

[40] Chang JL, Wang LF, Meng GF, Xiang SM, Pan CH. Deep adaptive image clustering. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2017. 5879–5887.

[41] Ehrhardt S, Zisserman A, Rebuffi SA, *et al*. Automatically discovering and learning new visual categories with ranking statistics. In: Proc. of the 8th Int'l Conf. on Learning Representations (ICLR). 2020.

汪云云(1986－), 女, 博士, 副教授, CCF 专业会员, 主要研究领域为迁移学习, 机器学习, 神经计算.

赵国祥(2000－), 男, 主要研究领域为迁移学习, 机器学习, 神经计算.

孙顾威(1996－), 男, 硕士生, 主要研究领域为迁移学习, 机器学习, 神经计算.

薛晖(1979－), 女, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为模式识别, 机器学习, 神经计算.