

# 基于改进深度子域适应网络的图像分类方法

郝海燕, 李芳

(沈阳理工大学 信息科学与工程学院, 沈阳 110159)

**摘要:** 深度子域适应网络在进行特征提取时易导致部分重要信息丢失, 且在对齐局部特征的同时会忽略源域和目标域的整体对齐, 从而影响其分类准确率, 为此提出基于改进深度子域适应网络的图像分类方法。引入卷积神经网络正则化方法提升模型在不同跨域任务中的泛化能力; 在特征提取网络中加入高效通道注意力机制, 对局部跨信道交互信息进行捕捉, 提取输入图像中的关键信息; 改进损失函数, 增加全域适应损失约束, 提升模型的全局特征对齐效果。在域适应基准数据集 Office-31 上的实验表明, 相较于原算法, 本文改进算法在一定程度上提升了分类准确率, 在跨域图像分类任务中表现更好。

**关键词:** 图像分类; 子域适应; 高效通道注意力; 全局特征对齐

**中图分类号:** TP391.41 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2024.01.011

## Image Classification Method Based on Improved Deep Subdomain Adaptation Network

HAO Haiyan, LI Fang

(Shenyang Ligong University, Shenyang 110159, China)

**Abstract:** The deep subdomain adaptation network easily leads to the loss of some important information during feature extraction, and ignores the overall alignment of the source domain and the target domain while aligning local features, thus affecting the classification accuracy. To address these issues, an improved method based on the deep subdomain adaptation network is proposed. The convolutional neural network regularization method is introduced to improve the model's generalization ability in different cross-domain tasks. Incorporating an efficient channel attention mechanism into the feature extraction network can extract key information from the input image by capturing local cross-channel interaction information. The loss function is modified by adding a global domain adaptation loss constraint which enhances the overall feature alignment effect of the model. Experiments on the domain-adapted benchmark dataset Office-31 show that the improved algorithm improves the classification accuracy to a certain extent compared with the original algorithm, and performs better in cross-domain image classification tasks.

**Key words:** image classification; subdomain adaptation; efficient channel attention; global feature alignment

以深度学习为基础的图像分类方法是计算机视觉中应用最广泛的技术之一。为成功构建图像分类系统,需预先为每个特定的分类任务提供数量足够的标注数据集,且要保证测试数据和训练数据具有相同的数据分布<sup>[1]</sup>。然而,人为采集的数据由于光照强度、拍照角度和使用场景各不相同,对其进行标注极为繁琐耗时,且对于医学影像、稀有动植物等特殊领域标准数据集的获取十分困难<sup>[2]</sup>。此外,由于不同领域的数据分布存在差异,现有数据集训练出来的模型往往难以在有别于训练环境的目标领域保持良好的泛化性能,使得源领域的知识难以迁移到目标领域,即存在领域偏移<sup>[3]</sup>。针对上述问题,人们提出了基于迁移学习的领域自适应方法,利用源领域中丰富的标记数据提高目标领域中未标记数据的分类准确率<sup>[4]</sup>,从而改善模型在跨领域任务中的性能。

领域自适应方法按特征对齐方式分为两类:以度量学习为主的域适应方法和以对抗学习为主的域适应方法。基于度量学习的域适应方法中用于衡量域与域之间分布差异的指标包括 KL 散度、最大均值差异(maximum mean discrepancy, MMD)、沃瑟斯坦(Wasserstein)距离等<sup>[5]</sup>。Pan 等<sup>[6]</sup>将 KL 散度与原型网络相结合,提出了可移植原型网络(TPN),使得源域和目标域中每个类的原型在嵌入空间中均比较接近。Pan 等<sup>[7]</sup>使用 MMD 衡量边缘分布的差异,并据此提出迁移成分分析(TCA)方法,实现了源域知识到不同但相关目标域的转移。Li 等<sup>[8]</sup>基于 TCA 提出了一种局部保留联合分布适配(LPJT)方法,将两个域的样本映射到一个低维特征空间,实现了条件概率分布和边缘概率分布的最小化。基于对抗学习的域适应算法通过在域判别器上施加一个对抗性目标,将域间分布差异的度量转化为对源域和目标域进行领域混淆,从而使两个领域的特征分布尽可能一致。Ganin 等<sup>[9]</sup>将生成对抗网络(GAN)的思想应用到域适应问题中,提出了域对抗神经网络(DANN),利用对抗学习方法对特征提取器和域判别器进行优化,实现了源域到目标域的知识迁移。Shen 等<sup>[10]</sup>提出了 Wasserstein 距离引导表示学习(WDGR)方法,通过优化特征提取器,以对抗的方式降低了源域样本和目标域样本之间的经验 Wasserstein 距离。Tzeng 等<sup>[11]</sup>将对抗性学习与判别性特征学习相结合,提出了对抗判别域适应(ADDA)方法,通过非权重共享的方式完成源和目标的独立映射,从而学习更多领域特定的

可提取特征。

与更细粒度的基于对抗学习的方法相比,基于度量学习的方法收敛更容易、收敛速度更快,但分类准确率较低。Zhu 等<sup>[12]</sup>提出的基于度量学习的深度子域自适应网络(DSAN)则在保证收敛速度的同时,达到甚至超过了目前主流的基于对抗学习算法的分类准确率。但 DSAN 算法尚存在的问题有:缺少对数据集的处理,易在多种多样的跨域任务中出现过拟合现象;仅通过卷积运算提取特征,无法区分哪些特征更重要,导致丢失许多有价值的信息;只在子领域进行对齐,导致对全局特征的关注度不够,影响特征对齐效果。

针对以上问题,本文在 DSAN 的基础上引入卷积神经网络正则化方法,加强模型在不同跨域任务中的泛化能力;加入高效通道注意力机制,使模型在提取特征时关注特征在通道间的关系,提取输入图像中的关键信息;增加全域适应损失约束,增强模型同时对齐全局特征和局部特征的能力,促使模型在训练过程中学习到更多用于域自适应的可转移特征,提升模型在无标签目标域上的分类准确率。

## 1 网络结构与原理

### 1.1 深度子域适应网络

DSAN 算法采用类别作为划分子领域的依据,其模型结构如图 1 所示。图中: $X^s$  为源域数据; $X^t$  为目标域数据; $Y^s$  表示源域数据的真实标签; $\hat{Y}^s$  和  $\hat{Y}^t$  分别表示网络模型预测的源域和目标域数据标签; $Z^s$  和  $Z^t$  分别表示  $X^s$  和  $X^t$  在  $l$  层激活( $l=1,2,\dots,L$ ), $L$  表示神经网络的层数,在深度网络的第  $l$  层进行特征适应需将  $l$  层作为激活层。 $X^s$  和  $X^t$  通过深度神经网络时被映射到同一个特征空间,经过特征提取后,两个域的数据均被划分为若干个子领域,分别在  $l$  层的同一子域内对  $Y^s$  和  $\hat{Y}^t$  使用局部最大平均差异(local maximum mean discrepancy, LMMD)计算源域和目标

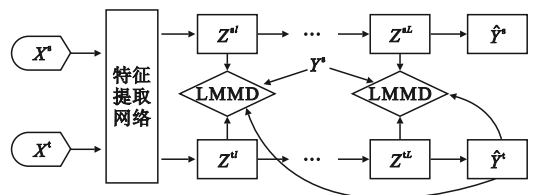


图 1 DSAN 模型结构

Fig. 1 DSAN model structure

域数据分布的距离,通过使相关域的数据分布保持一致来减少域间的特征差异。

## 1.2 改进深度子域适应网络

### 1.2.1 神经网络正则化

为使源域上训练的模型在目标域上仍有良好的表现,需要提高模型的泛化能力。DSAN 仅采用随机水平翻转的方式对训练集数据进行增强,该方式对模型在不同域适应任务中泛化能力的提升有限。

Dropout 是一种卷积神经网络正则化方法<sup>[13]</sup>,Cutout 是基于 Dropout 的扩展操作,Dropout 对经过网络提取后生成的图像特征进行遮挡,而 Cutout 直接遮挡输入的图像,对噪声的鲁棒性更好。Cutout 通过在输入图像中随机选取一固定比例的正方形区域,将这一区域的像素值设定为 0 或其他统一值,从而对训练数据集进行加噪处理,防止模型过度拟合,同时增加数据集的多样性,进而提升模型在不同数据集上的泛化能力。经 Cutout 处理后的数据集如图 2 所示。



图 2 经 Cutout 处理后的数据集

Fig. 2 Data sets after Cutout processing

### 1.2.2 高效通道注意力机制

DSAN 模型主要依赖深度神经网络提取源域和目标域的特征,然后在该特征空间中对两个领域的特征进行对齐。神经网络提取领域特征的效果对减少两个领域间的特征差异有直接影响,DSAN 采用 ResNet50 网络模型进行特征提取,对图片的所有信息给予相同的关注度,无法区分图像的哪个部位更重要。本文将高效通道注意力(ECA)模块<sup>[14]</sup>应用于 DSAN 模型的特征提取网络,改进后的特征提取网络结构如图 3 所示。

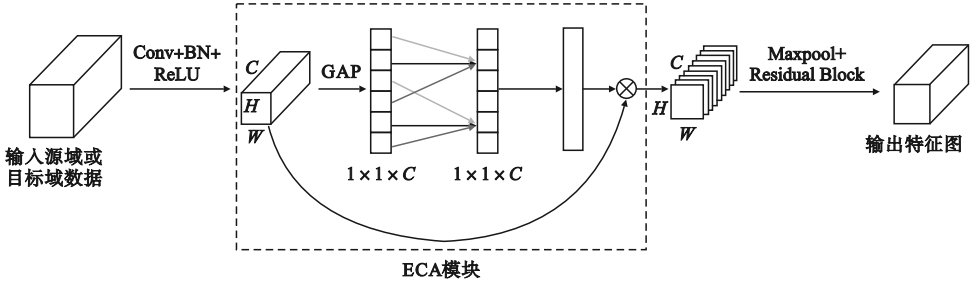


图 3 特征提取网络

Fig. 3 Feature extraction network

首先通过卷积(Conv)操作对输入神经网络的源域和目标域数据进行特征提取,同时使用批归一化(BN)和激活(ReLU)操作加速模型收敛。为提升网络模型的代表能力,在最大池化(Maxpool)操作前加入 ECA 模块,将卷积层输出的特征图进行通道注意力加权,以强化重要的特征并减少不重要的特征。然后将加权特征图输入到残差网络中进行多个残差块(Residual Block)的计算,进一步对输入特征图进行特征提取,最终输出具有更高层次语义信息的特征图。

在特征提取网络中,ECA 模块通过全局平均池化(GAP)将维度为  $H \times W \times C$  的特征图进行空间特征压缩,获取  $1 \times 1 \times C$  的特征图;对压缩后的特征图进行通道特征学习,通过一维卷积,学习不同通道之间的重要性,此时输出图像的维度仍为

$1 \times 1 \times C$ ;将经过通道特征学习的特征图(维度为  $1 \times 1 \times C$ )与原始输入特征图(维度为  $H \times W \times C$ )逐通道相乘,最终输出带通道注意力的特征图。

在执行卷积操作时,卷积核大小会影响感受野的范围,对于不同尺度的特征图,需使用不同大小的卷积核才能有效提取信息,但这样会增加模型的复杂度和计算量。为解决该问题,ECA 使用动态卷积核,通过自适应函数确定卷积核的大小,进行一维卷积,然后利用 Sigmoid 函数学习通道注意力。在通道数较大的层,使用较大的卷积核,更多地进行跨通道交互;在通道数较小的层,使用较小的卷积核,较少地进行跨通道交互。卷积核大小  $s$  和通道数  $C$  的关系定义为

$$s = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (1)$$

式中:odd表示取奇数; $\gamma$ 和 $b$ 用于改变通道数和卷积核大小之间的比例,分别设置为2和1。

将ECA引入DSAN的特征提取网络可有效提高模型捕获跨通道交互信息的能力,有助于模型学习更高效的通道注意力,从而提取输入图像中的关键信息,学习更多用于域自适应的域不变特征。

### 1.2.3 全局特征对齐

DSAN在进行特征对齐时,根据类别划分子领域,只考虑了源域和目标域相关子领域之间的关系,在对齐局部特征的同时,忽略了源域和目标域的整体对齐。

MMD是一种用于度量源域和目标域分布距离的核学习方法<sup>[15]</sup>,根据所选择的核函数不同,MMD有不同的类型,对于某一个具体任务而言,选择最适配的MMD比较困难。多核最大均值差异(multi-kernel MMD, MK-MMD)是将多个核函数综合起来构造一个总的核函数,可避免核函数选择困难的问题。MK-MMD定义为

$$d_k^2(p, q) = \| E_p[\phi(x^s)] - E_q[\phi(x^t)] \|_{H_k}^2 \quad (2)$$

式中: $p$ 和 $q$ 分别表示源域和目标域数据的概率分布; $d_k^2(p, q)$ 为 $p$ 和 $q$ 之间的MK-MMD距离; $E(\cdot)$ 表示期望值; $H_k$ 表示一个再生核希尔伯特空间; $\phi(\cdot)$ 表示映射函数,通过该映射函数可将源域数据和目标域数据映射到同一个 $H_k$ ;  $k$ 表示与特征映射 $\phi(\cdot)$ 相关的特征核。

MK-MMD中的特征核 $k$ 定义为 $m$ 个不同核的加权和,设 $\beta_u$ 表示第 $u$ 个特征核 $k_u$ 的权重系数,若满足 $\beta_u \geq 0$ ,且 $\sum_{u=1}^m \beta_u = 1$ ,则

$$k = \sum_{u=1}^m \beta_u k_u \quad (3)$$

将式(2)计算得到的 $d_k^2(p, q)$ 作为源域和目标域的全域适应损失项添加到损失函数中,重新构造的损失函数 $E$ 表示为

$$E = J(f(x^s), y^s) + \lambda \left[ \sum_{l=1}^L \hat{d}_l(p, q) + d_k^2(p, q) \right] \quad (4)$$

式中: $J(\cdot, \cdot)$ 为交叉熵损失函数,用于计算分类损失; $\sum_{l=1}^L \hat{d}_l(\cdot, \cdot)$ 表示子域自适应损失函数; $\lambda$ 为域适应损失和分类损失的权衡参数。

为使模型更好地适应训练的不同阶段,采用动态调整 $\lambda$ 的方式平衡域适应损失和分类损失之间的权重, $\lambda$ 计算式为

$$\lambda = \frac{2}{1 + \exp(-10t/T)} - 1 \quad (5)$$

式中: $t$ 为当前的训练轮数; $T$ 为总训练轮数。在训练初始阶段, $\lambda$ 取值接近于0,该阶段的主要任务是训练分类器;随着训练的深入, $\lambda$ 逐渐增大,最终趋近于1,域适应损失约束逐渐增强,该阶段的主要任务是将源域数据的特征迁移到目标域,并根据目标域的数据分布对特征进行调整和优化。

本文通过在损失函数中增加全域适应损失约束,促使深度自适应网络在训练过程中更加关注全局特征,源域和目标域数据达到更好的特征对齐效果,提高模型在跨域任务中的适应能力。

## 2 验证实验与结果分析

### 2.1 数据集

为证明改进算法的有效性,本文在公开数据集Office-31上进行验证。该数据集是域适应领域广泛使用的标准数据集,包含了31类办公室环境中常见的目标物体,共4110张图片。这些图像主要源于在线电商图片Amazon(A)、网络摄像头拍摄的低解析度图片Webcam(W)和由单反相机拍摄的高解析度图片DSLR(D),A中有2817张图片,W中有795张图片,D中有498张图片。

由于使用单一域适应任务会导致实验结果出现偶然性,因此本文选取并设定了A→W、A→D、D→W三种不同的域适应任务,其中箭头左边代表源域,右边表示目标域。

### 2.2 实验环境与参数设置

实验环境采用Ubuntu 20.04操作系统,GPU为Tesla P40,使用Python 3.8版本,深度学习框架为Pytorch 1.11.0。

为方便训练,在数据集输入模型前,首先将图片大小裁剪为 $256 \times 256$ 。实验中采用带动量的随机梯度下降算法进行参数更新,动量为0.9,批量大小为32。

### 2.3 实验结果分析

#### 2.3.1 对比实验

本文采用未经域适应算法的普通模型(Source-only)及原DSAN模型进行对比实验,各模型在A→W、A→D、D→W三种跨域任务中的分类结果如表1所示。

由表1可以看到,与Source-only相比,两种域适应算法均可显著提高跨域任务的分类准确

表 1 不同域适应任务下的分类准确率对比

Table 1 Comparisons of classification accuracy under different domain adaptation tasks %

算法	域适应任务		
	A→W	A→D	D→W
Source-only	80.63	82.73	97.61
DSAN	93.96	90.36	98.74
本文改进算法	<b>94.84</b>	<b>92.17</b>	<b>98.87</b>

率。与原 DSAN 模型相比,本文改进算法的性能在三种域适应任务中都得到了提升,在 A→D 的域适应任务中,本文改进算法的准确率达到 92.17%,比原 DSAN 算法提高了 2.00%,在 A→W 的域适应任务中,本文改进算法的准确率相较于原算法提升了 0.94%,在 D→W 的域适应任务中,本文改进算法的准确率提升了 0.13%。

为对比原算法和改进算法在不同迭代次数下的分类准确率,直观反映不同算法在整个训练进程中的性能差异和优化效果,图 4 和图 5 给出两个算法在 A→D 和 A→W 两个域适应任务中的分类准确率与迭代次数的关系。

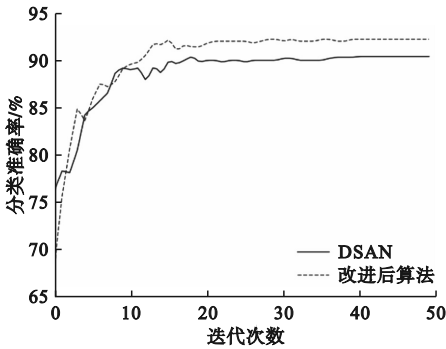


图 4 A→D 任务的分类准确率

Fig. 4 Classification accuracy in the A→D task

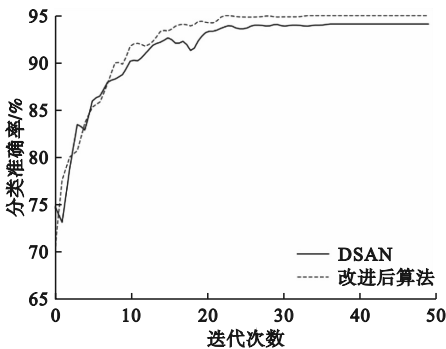


图 5 A→W 任务的分类准确率

Fig. 5 Classification accuracy in the A→W task

由图 4 和图 5 可见:在训练的前 20 轮,随着迭代次数的增加,原算法和改进算法的分类准确

率均有较大提高;随着训练的深入,两个算法的分类准确率在达到一定水平后趋于稳定;在两个域适应任务中,本文改进算法在第 10 轮到第 50 轮的训练中始终比原算法保持更高的准确率,且在 A→D 任务中分类准确率的提升优于 A→W 任务。

### 2.3.2 消融实验

为进一步评估本文方法对域适应图像分类效果的影响,在 DSAN 算法的基础上,通过依次改进损失函数,加入 ECA、引入 Cutout,探究各模块对算法性能的影响,在 A→W、A→D 两个域适应任务中进行实验,结果如表 2 所示。

表 2 各模块对域适应图像分类准确率的影响

Table 2 Impact of each module on the classification accuracy of domain adaptation image %

DSAN	改进损失函数	ECA	Cutout	域适应任务	
				A→W	A→D
✓	—	—	—	93.96	90.36
✓	✓	—	—	94.34	90.96
✓	✓	✓	—	94.59	91.57
✓	✓	✓	✓	<b>94.84</b>	<b>92.17</b>

由表 2 可见,本文方法通过改进 DSAN 算法的损失函数,增加全域适应损失约束,在 A→W 和 A→D 任务中的分类准确率分别提升了 0.40% 和 0.66%,说明全域适应损失约束能够促进特征对齐,增强对目标领域的适应性。本文方法在网络结构中引入 ECA,在 A→W 和 A→D 任务中的分类准确率分别提升了 0.26% 和 0.67%,说明注意力机制能够更好地捕捉到特征之间的关系,进一步验证了其在域适应图像分类任务中的有效性;本文应用 Cutout 神经网络正则化方法,在 A→W 和 A→D 两个域适应任务中的分类准确率分别提升了 0.26% 和 0.66%,说明 Cutout 能够有效增强模型在不同数据集上的泛化能力。

### 2.3.3 实验结果可视化

三种域适应任务下源域数据和目标域数据在域适应前后特征对齐效果的 t-SNE 图分别如图 6、图 7 和图 8 所示。图中红色表示源域样本,蓝色为目标域样本。

由图 6~8 可以看出:在域适应前,源域和目标域样本散乱地分布在一起,并没有呈现出明确的不同域间的特征关系;经过本文算法进行域间对齐后,源域和目标域中相同类的类间距离非常接近,同一类呈现出较好的聚集效果,不同类之间分散且类间距离较大。本文提出的改进算法能够

捕获每个类别的更多细粒度信息,使源域和目标域的分布尽可能保持一致,提升了特征对齐的效果。

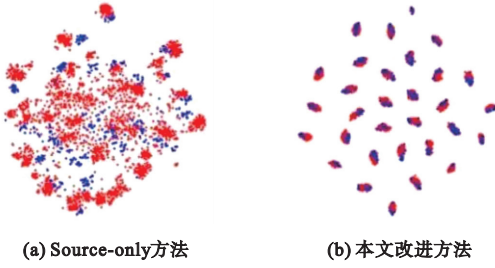


图6 A→W任务的t-SNE图

Fig. 6 t-SNE diagram of A→W task

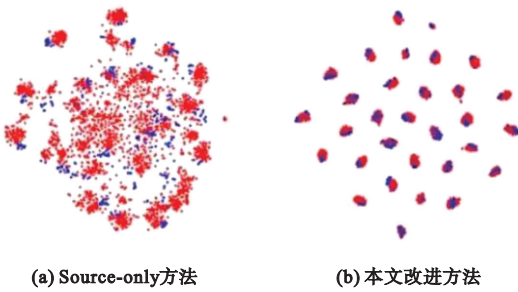


图7 A→D任务的t-SNE图

Fig. 7 t-SNE diagram of A→D task

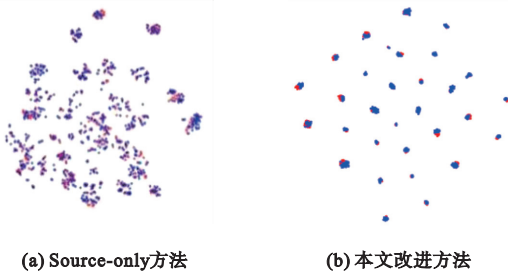


图8 D→W任务的t-SNE图

Fig. 8 t-SNE diagram of D→W task

### 3 结论

本文以DSAN算法为基础,应用Cutout缓解了过拟合现象,增强了模型的鲁棒性和泛化能力;引入ECA增强了模型通道权重的选择能力,提升了模型特征提取效果,有利于模型在后续过程中进行特征对齐;训练过程中引入全域适应损失进行约束,提高了模型的全局特征对齐能力,促使模型在训练过程中学习更多用于域自适应的域不变特征。在Office-31数据集上的实验表明,相较于原算法,改进算法在跨域图像分类任务中表现更好。

### 参考文献(References):

- [1] 田青,朱雅喃,马闯.基于深度学习的域适应方法综述[J].数据采集与处理,2022,37(3):512-541.  
TIAN Q, ZHU Y N, MA C. Review on domain adaptation methods based on deep learning[J]. Journal of Data Acquisition and Processing, 2022, 37(3): 512-541. (in Chinese)
- [2] 魏胜楠,张景异,陈亮,等.自适应局部关系网络的小样本学习方法[J].沈阳理工大学学报,2021,40(4):35-41.  
WEI S N, ZHANG J Y, CHEN L, et al. Few-shot learning method of adaptive local relational network[J]. Journal of Shenyang Ligong University, 2021, 40(4): 35-41. (in Chinese)
- [3] 贾熹滨,曾檬,米庆,等.领域对齐对抗的无监督跨领域文本情感分析算法[J].计算机研究与发展,2022,59(6):1255-1270.  
JIA X B, ZENG M, MI Q, et al. Domain alignment adversarial unsupervised cross-domain text sentiment analysis algorithm[J]. Journal of Computer Research and Development, 2022, 59(6): 1255-1270. (in Chinese)
- [4] 杨国庆,郭本华,钱淑渠,等.基于伪标签的无监督领域自适应分类方法[J].计算机应用研究,2022,39(5):1357-1361.  
YANG G Q, GUO B H, QIAN S Q, et al. Pseudo label based unsupervised domain adaptation classification method[J]. Application Research of Computers, 2022, 39(5): 1357-1361. (in Chinese)
- [5] 李晶晶,孟利超,张可,等.领域自适应研究综述[J].计算机工程,2021,47(6):1-13.  
LI J J, MENG L C, ZHANG K, et al. Review of studies on domain adaptation[J]. Computer Engineering, 2021, 47(6): 1-13. (in Chinese)
- [6] PAN Y W, YAO T, LI Y H, et al. Transferrable prototypical networks for unsupervised domain adaptation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2020: 2234-2242.
- [7] 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.
- [8] LI J J, JING M M, LU K, et al. Locality preserving joint transfer for domain adaptation[J]. IEEE Transactions on Image Processing, 2019, 28(12): 6103-6115.
- [9] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]//Proceedings of the 32nd International Conference on Machine Learning. New York, USA: ACM, 2015: 1180-1189.
- [10] SHEN J A, QU Y R, ZHANG W N, et al. Wasserstein distance guided representation learning for domain adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, California, USA: AAAI Press, 2018: 4058-4065.
- [11] TZENG E, HOFFMAN J, SAENKO K, et al. Adversarial discriminative domain adaptation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Honolulu, USA: IEEE, 2017: 2962-2971.
- [12] ZHU Y C, ZHUANG F Z, WANG J D, et al. Deep subdomain adaptation network for image classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(4): 1713-1722.