引文格式:徐仁,阿里木·赛买提,李二珠,等.面向任务对齐的遥感场景图像非监督域适应分类[J].武汉大学学报(信息科学版),2025,50(7):1359-1369.DOI:10.13203/j.whugis20230084

Citation: XU Ren, SAMAT Alim, LI Erzhu, et al. Task-Oriented Alignment for Unsupervised Domain Adaptation of Remote Sensing Scene Image Classification[J]. Geomatics and Information Science of Wuhan University, 2025, 50(7): 1359-1369. DOI: 10.13203/j.whugis20230084

面向任务对齐的遥感场景图像非监督域适应分类

徐 仁[1,2] 阿里木·赛买提[1,2] 李二珠[3] 王 伟[1,2]

1 中国科学院新疆生态与地理研究所荒漠与绿洲生态国家重点实验室,新疆 乌鲁木齐,830011

2 中国科学院大学资源与环境学院,北京,100049

3 江苏师范大学地理测绘与城乡规划学院,江苏 徐州,221116

摘 要:近年来,深度学习在端到端处理、隐含特征表示等方面的特殊优势,已成为遥感图像分类、目标识别、变化监测等任务中的主流解决方案。遥感场景图像分类是遥感图像分类领域中的研究热点之一,常规深度学习方法的性能常因图像的场景结构、空间尺度与分辨率、数据源及模型假设等因素而受到一定限制,尤其是在异源场景数据间的特征迁移、模型复用任务中。针对此问题,将计算机视觉领域中面向任务对齐的非监督域适应方法(task-oriented alignment for unsupervised domain adaptation, ToAlign UDA)引入到跨域遥感场景图像分类任务中,在解释算法原理和优化机制的基础上通过对比实验评价了其分类性能。使用ToAlign UDA对源域数据集进行训练学习,对NWPU-RESISC45、AID、PatternNet 3个目标数据集进行测试,在源域和目标域的空间分布、光谱特征、尺度等相似度较高的情况下,3个目标数据集的总体分类精度分别达到了95.16%、96.17%、99.28%。三者的分类精度均高于大多数场景分类算法,表明ToAlign UDA在遥感场景图像分类领域具有良好的算法竞争力。

关键词:遥感场景图像分类;域适应;面向任务分类;特征对齐;深度学习

中图分类号:P237 **文献标识码**:A **收稿日期**:2023-07-28

DOI:10.13203/j.whugis20230084 **文章编号**:1671-8860(2025)07-1359-11

Task-Oriented Alignment for Unsupervised Domain Adaptation of Remote Sensing Scene Image Classification

XU Ren[1,2] *SAMAT Alim*[1,2] *LI Erzhu*[3] *WANG Wei*[1,2]

1 State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China

2 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

3 School of Geography, Geomatics and Planning, Jiangsu Normal University, Xuzhou 221116, China

Abstract: **Objectives**: We primarily aim at addressing the prevailing challenges in remote sensing scene image classification, specifically those associated with the utilization of heterogeneous data and the achievement of cross-domain classification. The conventional deep learning methods, while effective, often encounter limitations due to factors such as spatial scale and resolution, data sources, model assumptions, and the inherent diversity of scene data when dealing with tasks like feature transferring and model reuse. **Methods**: In an attempt to overcome these obstacles, we introduce a novel approach called task-oriented alignment for unsupervised domain adaptation (ToAlign UDA). This algorithm, borrowed from the field of computer vision, is designed to enhance cross-domain remote sensing scene image classification. The principles and optimization mechanisms of the algorithm are explained, and its classification performance is evaluated through com-

基金项目:国家自然科学基金(42071424)。

第一作者:徐仁,硕士,主要研究方向为遥感图像分类和语义分割。xuren21@mails.ucas.ac.cn

通信作者:阿里木·赛买提,博士,研究员。alim_smt@ms.xjb.ac.cn

parative experiments. **Results:** ToAlign UDA is used in the experiment to train on the source domain dataset, while tests are conducted on three target datasets: NWPU-RESISC45, AID, and PatternNet. When the spatial distribution, spectral characteristics, scale, and other similarities between the source and target domains are high, ToAlign UDA achieves an overall classification accuracy of 95.16% on NWPU-RESISC45, 96.17% on AID, and 99.28% on PatternNet. **Conclusions:** The results clearly indicate that the ToAlign UDA approach outperforms most scene classification algorithms in terms of classification accuracy in remote sensing scene image analysis. Therefore, it holds significant potential in advancing the field of remote sensing image classification, particularly in the context of utilizing heterogeneous data and achieving cross-domain classification.

Key words: remote sensing scene image classification; domain adaptation; task-oriented classification; feature alignment; deep learning

随着对地观测技术的不断发展,高分辨率遥感影像数据的来源日益丰富,数据获取成本不断降低,其在城市环境监测、城市规划、土地利用/覆被、目标检测、精准农业等领域中发挥着重要的作用[1]。遥感图像分类作为对地观测数据处理领域中国际前沿热点研究方向之一,从分类对象的角度可分为亚像元[2]、像元[3]、超像素[4]、对象[5]及场景[6]的分类。场景分类是遥感影像分类的重要组成部分,其涵义是对每一张场景图像进行语义分析和理解,实现场景图像数据集的分类[7]。

遥感场景图像相较于自然场景图像,具有数据多源、体量大、场景结构复杂、类别多样、尺度差异大等特点,使常规的场景分类算法面临着多种问题与挑战[8]。深度学习可以自动提取目标任务特征,不需要人工设计特征,并且具有鲁棒性强、模型易迁移等优点,在土地利用/覆被分类、作物精细分类、人工用地专题制图等领域广泛应用[9]。然而基于深度学习的场景分类同样面临着来自类内多样性、类间相似性、图像差异性等方面的问题与挑战。现有的研究工作针对此类问题的处理方式包括判别特征函数训练、双流特征融合、分支特征融合、二阶统计特征学习、多尺度注意力融合、贝叶斯理论(信息融合)、数据联合等,以及主流的深度学习分类和提取,这些方法均着力于各类特征学习,依赖于大量标记样本(部分方法展开了样本受限条件下的分类研究),且模型复杂度高、部分方法没有考虑空间结构信息,这些方法也局限于同源数据训练,通用性低,难以应对深度卷积神经网络在异源遥感场景图像分类任务中的挑战[10-21]。

针对训练数据(源域)和测试数据(目标域)不同(特征异分布)而导致源域模型在目标域性能降低或失效的问题,通常采用迁移学习中域适应方法加以控制和解决[22]。域适应是一种跨领域适配模型的任务,当源域和目标域的数据分布不同时,其旨在建立可以推广到目标域的机器学习模型,并处理跨领域分布的差异,源域和目标域的数据分布差异主要表现在边缘分布差异、条件分布差异、标签空间差异、特征空间差异等方面[23-24]。本文使用闭集域适应方法来解决源域和目标域数据分布不同的问题,闭集域适应更加简单和高效,在许多实际应用中已被证明性能良好,且闭集域适应在遥感图像分类等特定场景能更好地满足需求[25]。根据目标域的数据标签也可以将域适应分为监督域适应(目标域已标记,但目标域数据规模小)、半监督域适应(目标域既有已标记的数据,也有未标记的数据)和无监督域适应(目标域无任何标记数据)[26]。

相较于监督、半监督域适应,无监督域适应不需要目标样本集有任何标注,在应用中可以避免数据标记工作,大大提高学习性能[27]。当前无监督域适应在遥感场景图像分类领域的主要挑战在于小样本和类别不平衡问题,一些方法采用辅助任务结合对抗网络、双向映射、协方差对齐、类对比迁移等手段企图解决小样本和类别不平衡对网络性能的影响,但仍然存在跨域特征映射约束不足、稀缺特征对齐不均衡、分类性能不高等情况[28-32]。且目前大多数无监督域适应采用的特征对齐手段是基于域整体特征操作的,这就导致源域中与分类任务无关的特征也参与了目标域的对齐,而一个域由多个类别组成,基于整个数据集的特征对齐不能针对性地解决分类工作中的域差异问题[33]。

面向任务对齐的非监督域适应方法(task-oriented alignment for unsupervised domain adaptation, ToAlign UDA)是计算机视觉领域的一种基于对抗学习的无监督域适应方法,该方法将源域特征划分为两种特征(与分类任务相关的特

征、与分类任务无关的特征),然后剔除任务无关特征,只将任务相关特征与目标域特征进行对齐[34]。该操作能更好地作用于分类任务,显式地减少域差异。但 ToAlign UDA 在遥感图像场景分类任务中并未得到应用,其在遥感场景图像分类中的性能尚未得到验证。

将 ToAlign UDA 应用于遥感场景图像分类时,需要考虑以下问题:

1)数据分布问题。遥感场景图像包含更多的地物信息,且与计算机视觉领域的图像在灰度值分布、空间分辨率等方面存在较大差距。

2)类别不平衡问题。遥感场景图像中存在样本类别不平衡问题,如果不进行人工均衡干预而直接使用卷积神经网络(convolutional neural network, CNN)模型,会影响整体分类性能。

3)数据集规模问题。遥感场景数据集规模往往远小于计算机视觉领域数据集,数据量不足可能会导致模型过拟合,从而影响分类性能。

本文选择成熟的数据集(NWPU-RESISC45、AID、OPTIMAL-31 和 PatternNet)来解决类别不均衡问题。这些数据集均在发布之前做了数据增强处理,可以很好地解决计算机视觉领域模型在遥感场景分类中应用的灰度值分布、空间分辨率差异、数据规模等问题。

本文将 ToAlign UDA 引入遥感场景图像分类任务,一是为了解决遥感场景图像分类任务中的数据异源问题;二是为了实现目标遥感场景数据集无标记的分类问题;三是通过剔除源域数据集的任务无关特征,以达到更高精度的遥感场景图像分类。本文采用的基本分类框架如图 1 所示。

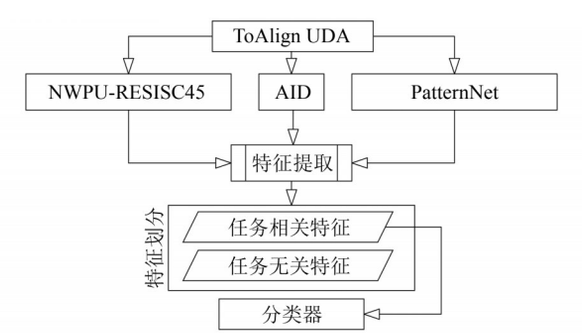


图 1　ToAlign UDA 的基本分类框架

Fig. 1　Basic Classification Framework of ToAlign UDA

1　方法原理

1.1　总体技术路线

作为一种基于对抗学习的无监督域适应分类方法,ToAlign UDA 一般采用域鉴别器来区分样本的来源(即来自源域还是目标域),并训练一个特征提取器来迷惑域鉴别器,以学习域不变特征表达[35]。特征提取器的优化目标是最大化域分类损失 L_D 并最小化图像分类损失 L_{cls},而域鉴别器则是最小化域分类损失 L_D[36]。两者的对抗需要引入梯度反转层,从而使得特征提取器既可生成正确预测的特征,又使其所提取的特征难以被判断来源[27]。

算法实施的核心过程为特征分解和域对齐。首先利用特征提取器初步提取源域和目标域的特征,然后对源域特征进行特征分解。特征分解是基于分类元知识制定面向分类任务的权重 $\boldsymbol{w}_p^{\text{cls}}$ 来筛选源域特征,即将初步提取的源域特征与分类任务权重进行哈达玛积,从而得到面向分类任务的源域特征 $\boldsymbol{f}_p^s$[37]。域对齐则将面向分类任务的源域特征 $\boldsymbol{f}_p^s$ 和目标域初步提取的特征 $\boldsymbol{f}^t$ 进行特征空间对齐,以便进行目标域的分类任务。算法流程如图 2 所示。

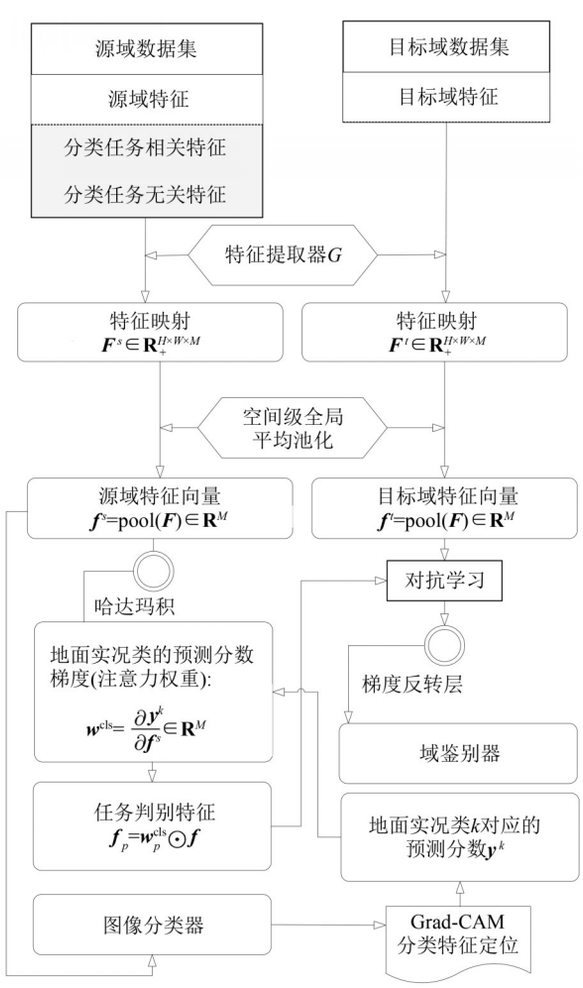


图 2　本文算法流程

Fig. 2　Flowchart of the Proposed Algorithm

1.2 特征分解

梯度加权类激活映射(gradient-weighted class activation mapping, Grad-CAM)是一种基于CNN的视觉解释技术,通过产生粗略定位映射并突出显示来预测概念的重要区域,被广泛应用于CNN中,用来定位分类的重要特征[38]。Grad-CAM梯度具有任务判别信息传递、特征识别和图像类别正确识别的能力。基于Grad-CAM的分类特征定位原理:面向任务的特征分解利用了地面实况类的预测分数梯度作为注意力权重,以获取任务判别特征。

在算法流程图中,可以从特征提取器最终的卷积块中获取特征映射$F\in\mathbf{R}_{+}^{H\times W\times M}$,即$F$为一个非负实数张量,高$H$、宽$W$,有$M$个通道。在经过空间级全局平均池化(global average pooling, GAP)之后,得到特征向量$\boldsymbol{f}=\mathrm{pool}(F)\in\mathbf{R}^{M}$。使用分类器$C$来预测所有类的对数,基于$C(\boldsymbol{f})$推导出$\boldsymbol{y}^{k}$($\boldsymbol{y}^{k}$是地面实况类$k$对应的预测分类)的梯度$\boldsymbol{w}^{\mathrm{cls}}\in\mathbf{R}^{M}$。$\boldsymbol{w}^{\mathrm{cls}}$和$\boldsymbol{f}$的关系为:

$$\boldsymbol{w}^{\mathrm{cls}}=\frac{\partial \boldsymbol{y}^{k}}{\partial \boldsymbol{f}} \tag{1}$$

梯度$\boldsymbol{w}^{\mathrm{cls}}$传递了特征$\boldsymbol{f}$中重要的优异通道信息,用于将样本数据集有效分类到地面实况类k[39]。基于Grad-CAM原理寻找分类判别特征对$\boldsymbol{w}^{\mathrm{cls}}$进行特征图信道调制,得到的任务判别特征$\boldsymbol{f}_p$为:

$$\boldsymbol{f}_p=\boldsymbol{w}_p^{\mathrm{cls}}\odot\boldsymbol{f}=s\boldsymbol{w}^{\mathrm{cls}}\odot\boldsymbol{f} \tag{2}$$

式中,$\odot$表示哈达玛积;s为自适应非负参数且$s\in\mathbf{R}_{+}$:

$$s=\sqrt{\frac{\|\boldsymbol{f}\|_2^2}{\|\boldsymbol{w}^{\mathrm{cls}}\odot\boldsymbol{f}\|_2^2}}=\sqrt{\frac{\sum_{m=1}^{M}\left\|\boldsymbol{f}_m\right\|^2}{\sum_{m=1}^{M}\left\|\boldsymbol{w}_m^{\mathrm{cls}}\boldsymbol{f}_m\right\|^2}} \tag{3}$$

与Grad-CAM找寻分类判别特征相反,任务无关的特征(即非判别特征)可以表示为$\boldsymbol{f}_n=-\boldsymbol{w}_p^{\mathrm{cls}}\odot\boldsymbol{f}$。$\left\|\boldsymbol{w}_p^{\mathrm{cls}}\right\|$和$-\left\|\boldsymbol{w}_p^{\mathrm{cls}}\right\|$的正反数关系使得任务判别通道中$\left\|\boldsymbol{w}_p^{\mathrm{cls}}\right\|$较大时,$-\left\|\boldsymbol{w}_p^{\mathrm{cls}}\right\|$反而较小。

1.3 域对齐

源域和目标域的特征对齐以最终分类任务为导向,基于分类元知识获取源域任务鉴别特征,从而利用不同的基于领域对抗学习的特征对齐方法来引导目标域和源域的任务鉴别特征进行分布对齐。源域特征$\boldsymbol{f}^{s}$经过面向分类任务的权重运算之后,最终输入到域鉴别器中的是源域的分类任务相关特征$\boldsymbol{f}_p^{s}$,此时的域分类损失函数见文献[40]。

2 遥感场景图像分类实验

2.1 实验数据

NWPU-RESISC45数据集是西北工业大学所创建的一个公开基准,隶属于遥感场景分类数据集[41]。该数据集覆盖45个场景类别,每个类别有700张图片,共31 500张图片,其部分图像如图3(a)所示。场景类别包括飞机、机场、棒球场、篮球场、海滩、桥梁、灌木丛、教堂、圆形农田、云端、商业区、密集住宅区、沙漠、森林、高速公路等。该数据覆盖面广,包含了全球100多个国家和地区。由于在空间分辨率、视角、目标姿态等方面具有较大差异,该数据集的类内差异性较大,且类间相似度较高。

AID数据集是一个大型航摄图像数据集,该数据集以提高遥感图像场景分类技术水平为初衷,数据质量高[42]。数据集由遥感解译领域的专家进行标注,并使用典型的场景分类方法和深度学习方法对其进行了性能测试。AID数据集共有10 000幅航摄影像,由30个场景类别组成,包括机场、裸地、棒球场、海滩、桥梁、中心、教堂等,每个类别有220~420张图像,部分图像如图3(b)所示。该数据集从谷歌地球上采集样本图像,谷歌地球使用的是后期红绿蓝(red green blue, RGB)渲染处理后的航空图像,已有研究表明,在像素级土地利用/覆盖映射中,经过RGB渲染的航空图像和原影像数据并无显著差异,故而AID数据集可以直接用于评价场景分类算法。

OPTIMAL-31数据集是西北工业大学在2019年发布的场景分类数据集,包含飞机、机场、棒球场、篮球场、海滩、桥、灌木、教堂等31个类别,每个类别有60张图像,共1 860张图片,其部分图像如图3(c)所示,数据集的图像来源于谷歌地球[43]。

PatternNet是一个大型高分辨率的遥感场景数据集,包含飞机、棒球场、篮球场、海滩、桥、墓地、灌木丛等38个类别,每个类别有800张图像,共30 400张图像[44]。该数据集图像展示的是美国的一些城市,数据采集于谷歌地球或使用谷歌应用程序接口收集,其部分图像如图3(d)所示。该数据集的规模较大,使其在基于深度学习的遥感场景图像分类中具有更好的适应性(相较于小规模数据集)。此外,PatternNet目标图像中的背

景和其他冗余地物占比少,目标针对性更强。

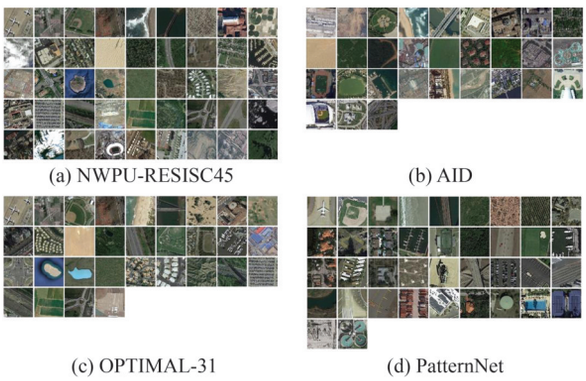


图 3　实验数据集样例图

Fig. 3　Sample Diagrams of Experimental Datasets

2.2　实验设置

2.2.1　对比算法介绍

为了客观评价 ToAlign UDA 方法在遥感场景图像域适应分类任务中的性能,本文选取了近年来遥感场景图像分类研究中的代表性方法进行对比实验分析,包括双流特征聚合深度神经网络(two-stream feature aggregation deep neural network, TFADNN)、基于分支特征融合的轻量级卷积神经网络分类方法(lightweight convolutional neural network classification method based on branch feature fusion, LCNN-BFF)、判别卷积神经网络(discriminative convolutional neural network, D-CNNs)等。实验中的源域-目标域数据集均选取公共类别部分,每对源域-目标域数据集公共类别展示在§2.3.2的混淆矩阵图以及数据集类间精度表中。

TFADNN是一种用于高空间分辨率遥感场景分类的双流特征聚合深度神经网络,该方法利用两个平行部分(判别特征流和一般特征流)获取特征表示,并对两个流的特征进行融合[11]。LCNN-BFF是一种基于分支特征融合的轻量级CNN遥感场景图像分类方法,采用双线性特征提取结构,在保证高精度前提下降低了模型的复杂性[12]。D-CNNs是一种简单高效的遥感场景图像分类方法,通过优化新的判别目标来解决分类中存在的类内多样性和类间相似性问题[10]。跳跃连接协方差(skip-connected covariance, SCCov)是一种跨连接协方差网络,该方法在传统CNN模型的基础上嵌入了跳跃连接和协方差池化,使得CNN在处理遥感场景图像分类问题中实现更具有代表性的特征学习[13]。多层特征融合网络(multilayer feature fusion network, MF^2 Net)是一种多层特征融合网络,该方法利用视觉几何组网络(visual geometry group network, VGGNet)获取多层卷积特征,对分支层特征进行处理以进行预测分类[45]。CNN-图谱卷积网络(CNN-graph convolutional network, CNN-GCN)是一种将基于全局的视觉特征和基于对象的位置特征相结合的双流结构,利用CNN进行视觉特征提取,同时采用GCN了解地物空间位置特征,是第一个在遥感场景图像分类任务中研究对象依赖关系的方法[46]。高阶GCN(high-order GCN, H-GCN)是一种基于高阶图卷积网络的遥感场景分类方法,其顾及图像内的判别分量,应用高阶图卷积网络来研究类依赖关系,并将依赖关系通过混合不同顺序节点的邻居信息来传播,因此可以获得更多的信息节点表达[47]。残差神经网络(residual network, ResNet)-152是一种常用的深度特征提取模型,是一种低复杂度且深度高达152层的残差网络[48]。注意力递归卷积网络(attention recurrent convolutional network, ARCNet)是一种注意力循环卷积网络,该方法可以针对某些关键区域或位置进行高级特征处理,从而摒弃非关键信息以提高分类性能[43]。门控双向网络(gated bidirectional network, GBNet)是一种集成了分层特征聚合和干扰信息消除的双向门网络,同时顾及了多层卷积特征和特征层次结构[49]。多级跨层双线性融合(multi-level cross-layer bilinear fusion, MLCLBF)是一种基于多级别跨层双线性融合的光学遥感场景分类算法,能较好地解决光学图像中空间复杂性、类内多样性和类间相似性问题[50]。亚历克斯·克里热夫斯基网络(Alex Krizhevsky network, AlexNet)和VGGNet是十分经典的模型,AlexNet在2012年被开发出来,该网络包含10个用于卷积和池化的层,VGGNet则相当于AlexNet的增强版,区别在于层数[51]。这两个经典模型架构简洁,但计算量大且评估成本高。计算机视觉初始架构(inception architecture for computer vision, Inception-V3)的计算成本远远低于VGGNet等经典网络,适用于大型数据的场景分类,但由于其架构的复杂性,Inception-V3难以进行网络更改[52]。ResNet-101是深度层为101的残差网络,与ResNet-152的区别在于层数。全局-局部注意力网络(global-local attention network, GLANet)是一种用于捕获全局和局部信息的注意网络,在训练中通过局部特征、全局特征和输入图像的连接特征来预测分类[53]。显著性

双重注意残差网络(saliency dual attention residual network,SDAResNet)是一种用于提取跨通道和空间显著性信息的双注意力残差网络,该方法在低层次特征中嵌入空间注意力,在高层次特征中嵌入通道注意力,达到强调位置信息并抑制背景信息的目的[54]。

2.2.2 运行环境

本文实验的运行环境为:Ubuntu 20.04系统、i7-11700k CPU、GeForce RTX 3070 GPU、64 GB RAM。实验使用的遥感场景图像分类的域适应网络ToAlign UDA基于PyTorch框架实现。为了更好地测试域适应网络ToAlign UDA在遥感场景图像中的分类性能,设计了4组正向测试数据集和4组反向测试数据集(将4组正向测试数据集源域和目标域进行调换,就得到了4组反向测试数据集),由于源域和目标域数据分布的差异,实验结果部分选取了分类表现较好的3组数据,原始设定的8组数据测试结果如表1所示。受到硬件条件的限制,设置实验批次大小为50。

2.3 实验结果

2.3.1 算法对比

为了对比ToAlign UDA算法与其他相关网络在遥感场景图像分类中的有效性,利用NWPU-RESISC45、AID和PatternNet 3个公开数据集进行实验,用于参考的其他算法的精度采用不同训练比率中分类精度的最高值。ToAlign UDA算法的参数设置采用ResNet-50作为基础网络,判别器输入维度为1 024,输出维度为1。使用域对齐损失和域混淆策略,源域和目标域学习率均为0.001。为确保实验结果的可靠性,每对实验都进行了5次。各个算法在3个数据集上的精度表现如图4所示。

表1 8组数据集的测试结果

Table 1 Test Results of 8 Datasets

源域数据集	目标域数据集	总体分类精度/%
NWPU-RESISC45	OPTIMAL-31	100.00
OPTIMAL-31	NWPU-RESISC45	95.16
NWPU-RESISC45	AID	96.17
AID	NWPU-RESISC45	90.38
AID	PatternNet	99.28
PatternNet	AID	76.16
Million-AID	NWPU-RESISC45	48.65
NWPU-RESISC45	Million-AID	64.62

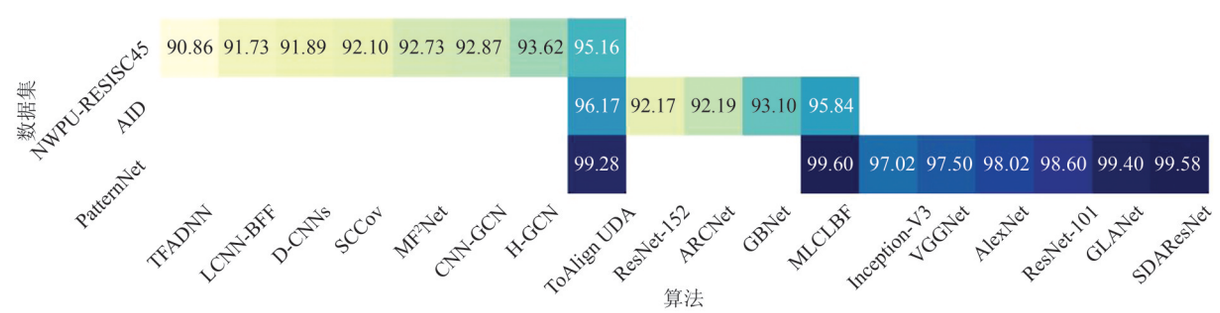


图4 不同算法在3个数据集上的精度热力图

Fig. 4 Precision Heat Maps of Different Algorithms on 3 Datasets

NWPU-RESISC45数据集的总体分类精度(与其他算法对比)如表2所示。本文所使用的ToAlign UDA算法,当源域为OPTIMAL-31数据集时,NWPU-RESISC45数据集的总体分类精度为95.16%,相较于TFADNN、LCNN-BFF、D-CNNs、SCCov等算法具有更好的分类性能。

AID数据集的场景类别多,数据规模大,像素分辨率变化范围大,这在一定程度上增加了该数据集的分类难度,其总体分类精度(与其他算法对比)如表3所示。本文所使用的ToAlign UDA算法,当源域为NWPU-RESISC45数据集时,AID数据集的总体分类精度为96.17%,具有优于ResNet-152、ARCNet、GBNet和MLCLBF算法的分类性能。

表2 不同算法在NWPU-RESISC45数据集上的精度

Table 2 Precision of Different Algorithms on NWPU-RESISC45 Dataset

方法	总体分类精度/%
TFADNN	90.86
LCNN-BFF	91.73
D-CNNs	91.89
SCCov	92.10
MF^2Net	92.73
CNN-GCN	92.87
H-GCN	93.62
ToAlign UDA	95.16

PatternNet数据集规模大、分辨率高,其总体分类精度(与其他算法对比)如表4所示。本文所

表 3　不同算法在 AID 数据集上的精度

Table 3　Precision of Different Algorithms on AID Dataset

方法	总体分类精度/%
ResNet-152	92.19
ARCNet	93.10
GBNet	94.58
MLCLBF	95.84
ToAlign UDA	96.17

使用的 ToAlign UDA 算法,当源域为 AID 数据集时,PatternNet 数据集的总体分类精度为 99.28%,分类性能优于 Inception-V3、VGGNet、AlexNet 和 ResNet-101,略逊于 GLANet、SDAResNet 和 MLCLBF。

表 4　不同算法在 PatternNet 数据集上的精度

Table 4　Precision of Different Algorithms on PatternNet Dataset

方法	总体分类精度/%
Inception-V3	97.02
VGGNet	97.50
AlexNet	98.02
ResNet-101	98.60
GLANet	99.40
SDAResNet	99.58
MLCLBF	99.60
ToAlign UDA	99.28

2.3.2　结果分析

3 个目标数据集的总体分类精度、召回率和 *F*1 分数如表 5 所示,表 5 中结果显示了 Toalign UDA 在 3 个数据集上具备优异的分类性能。

表 5　目标数据集的精度/%

Table 5　Precision of Target Datasets/%

数据集	总体分类精度	召回率	*F*1 分数
NWPU-RESISC45	95.16	95.29	95.16
AID	96.17	96.04	96.23
PatternNet	95.16	99.32	99.28

NWPU-RESISC45 数据集的混淆矩阵如图 5 所示,10 种场景类别中有 8 种实现了较高水平的分类精度。从表 6 可知,绝大部分场景类别达到了 90% 以上的分类准确率,仅有森林和跑道两个类别的分类出现了相对明显的混淆问题。森林的一部分样本被划分到了沙漠中,而跑道的部分样本被划分到了机场中。错分的两个类别之间存在一些类似的结构信息(图 6):森林的部分样本由于尺度较大,样本特征和沙漠中灌木覆盖的图像极为相似,故而导致了错分;跑道的部分大尺度样本特征与机场吻合度高,原本跑道就是机场的一部分,当出现尺度相似的问题时,跑道的样本则易被错分成机场。

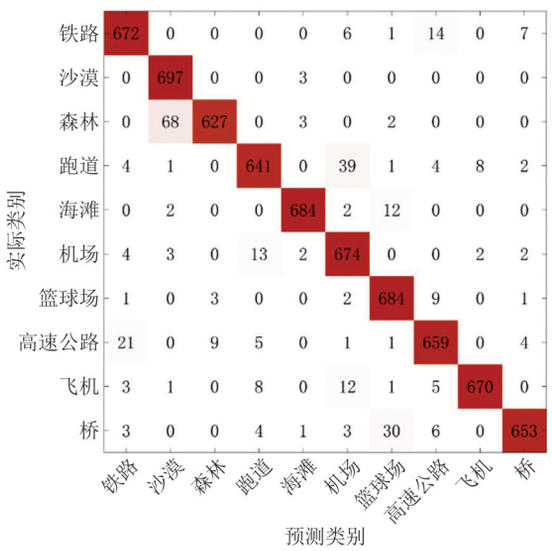


图 5　NWPU-RESISC45 数据集混淆矩阵

Fig. 5　Confusion Matrix of NWPU-RESISC45 Dataset

表 6　NWPU-RESISC45 数据集各类别分类精度/%

Table 6　Classification Accuracy of NWPU-RESISC45 Dataset/%

类别	精确率	召回率	*F*1 分数
铁路	96.00	94.92	95.45
沙漠	99.57	90.29	94.70
森林	89.57	98.12	93.65
跑道	91.57	95.53	93.51
海滩	97.71	98.70	98.21
机场	96.29	91.20	93.68
篮球场	97.71	93.44	95.53
高速公路	94.14	94.55	94.35
飞机	95.71	98.53	97.10
桥	93.29	97.61	95.40



图 6　NWPU-RESISC45 数据集的易错分样例

Fig. 6　Samples in NWPU-RESISC45 Dataset That Is Prone to Misclassification

从图 7 和表 7 可知,AID 数据集各个类别的分类精度较好,只有少量的河流和密集住宅区存在明显的错分问题(图 8)。由于部分河流和棒球

场的结构形态十分相似,导致了少量河流样本被错分成了棒球场。当密集住宅区的建筑规模较大时,其视觉表现上和商业房产区具有相似的建筑外观,从而导致少部分的密集住宅区被错分成了商业房产区。

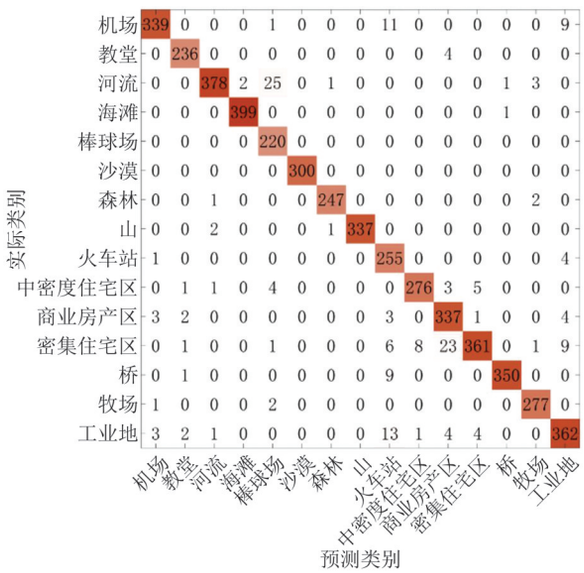


图7 AID数据集混淆矩阵

Fig. 7 Confusion Matrix of AID Dataset

表7 AID数据集各类别分类精度/%

Table 7 Classification Accuracy of AID/%

类别	精确率	召回率	*F*1分数
机场	94.17	97.69	95.90
教堂	98.33	97.12	97.72
河流	92.20	98.69	95.33
海滩	99.75	99.50	99.63
棒球场	100.00	86.96	93.02
沙漠	100.00	100.00	100.00
森林	98.80	99.20	99.00
山	99.12	100.00	99.56
火车站	98.08	85.86	91.56
中密度住宅区	95.17	96.84	96.00
商业房产区	96.29	90.84	93.48
密集住宅区	88.05	97.30	92.45
桥	97.22	99.43	98.31
牧场	98.93	97.88	98.40
工业地	92.82	93.30	93.06

PatternNet数据集生成的混淆矩阵如图9所示,9个场景类别中有8个场景类别几乎都达到了100%的分类精度(表8)。仅密集住宅区有少部分被错分成了稀疏住宅区。如图10所示,错分的密集住宅区和稀疏住宅区存在一定的分布相似性,两者的区分存在界定值问题,在高于某种界定值情况下为密集,反之为稀疏。在密集住宅区和稀疏住宅区的建筑密度相近的情况下(在界定值附近时),容易产生错分现象。

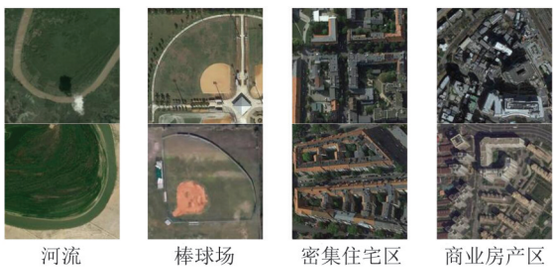


图8 AID数据集的易错分样例

Fig. 8 Samples in AID Dataset That Is Prone to Misclassification

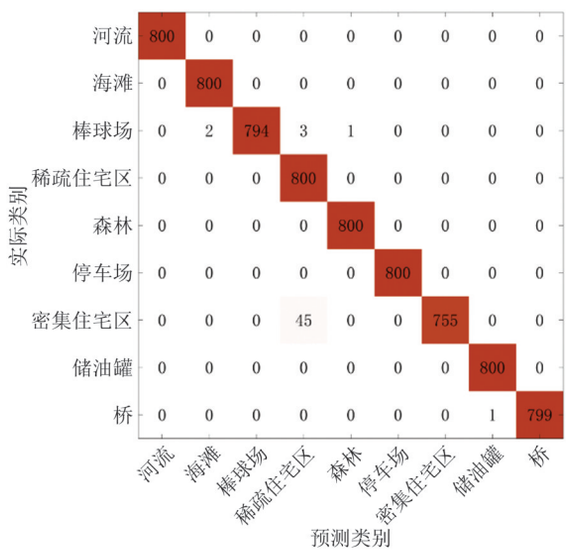


图9 PatternNet数据集混淆矩阵

Fig. 9 Confusion Matrix of PatternNet Dataset

表8 PatternNet数据集各类别分类精度/%

Table 8 Classification Accuracy of PatternNet Dataset/%

类别	精确率	召回率	*F*1分数
河流	100.00	100.00	100.00
海滩	100.00	99.75	99.88
棒球场	99.25	100.00	99.62
稀疏住宅区	100.00	94.34	97.09
森林	100.00	99.88	99.94
停车场	100.00	100.00	100.00
密集住宅区	94.38	100.00	97.11
储油罐	100.00	99.88	99.94
桥	99.88	100.00	99.94

3 结 语

传统深度学习方法采用同源数据进行训练和测试(如将NWPU-RESISC45数据集划分20%为训练数据集,剩余80%作为测试数据集),避免了数据分布差异、传感器差异和尺度问题,但对于异源数据的特征迁移却无从下手。且传



图 10　PatternNet数据集的易错分样例

Fig. 10　Samples in PatternNet Dataset That Is Prone to Misclassification

统深度学习方法受限于训练样本的数量以及场景分类算法本身性能,难以适应遥感图像处理领域的多源数据发展趋势。本文引入的 ToAlign UDA算法以分类任务为导向,实现不同数据集之间的特征提取和对齐,有效地将源域知识应用到目标域的分类任务中。相比较于经典的无监督域适应,ToAlign UDA对源域特征进行特征判别操作,剔除了与分类任务无关的特征,避免数据集总体特征对齐带来的特征冗余效应和映射偏颇问题。本文将 ToAlign UDA 在 NWPU-RESISC45、AID、PatternNet这3个目标数据集进行测试,取得的总体分类精度分别达到了95.16%、96.17%、99.28%。实验表明,在训练样本数量足够,且源域和目标域的空间分布、光谱特征、尺度等相似度较高时,ToAlign UDA在遥感场景图像分类中具有十分优异的分类性能。

由于 ToAlign UDA对于源域特征的判别操作依赖于地面实况类的预测分数梯度,使得分离出的任务判别特征存在一定的误差,以此进行的对抗学习受到源域数据集特征处理的影响较大。此外,该方法和其他无监督域适应方法一样,对源域和目标域的数据空间分布、光谱特征、尺度差异较为敏感。因此,优化分类任务特征判别以及提升无监督域适应模型的泛化能力是下一步研究中的重要方向。

参　考　文　献

[1] 秦增忍. 高分辨率影像在测绘生产中的应用及其发展潜力探讨[J]. 科技创新导报, 2015, 12(17): 75-76.
QIN Zengren. Application of High-Resolution Image in Surveying and Mapping Production and Its Development Potential[J]. *Science and Technology Innovation Herald*, 2015, 12(17): 75-76.

[2] 任武, 葛咏. 遥感影像亚像元制图方法研究进展综述[J]. 遥感技术与应用, 2011, 26(1): 33-44.
REN Wu, GE Yong. Progress on Sub-Pixel Mapping Methods for Remotely Sensed Images[J]. *Remote Sensing Technology and Application*, 2011, 26(1): 33-44.

[3] 朱述龙. 基于混合像元的遥感图像分类技术[J]. 测绘科学技术学报, 1995, 12(4): 276-278.
ZHU Shulong. Remote Sensing Image Classification Technology Based on Mixed Pixels[J]. *Journal of Geomatics Science and Technology*, 1995, 12(4): 276-278.

[4] 张思宇. 基于超像素的遥感图像分类技术研究[D]. 哈尔滨: 哈尔滨工程大学, 2018.
ZHANG Siyu. Research on Classification Technology of Remote Sensing Image Based on Superpixel[D]. Harbin: Harbin Engineering University, 2018.

[5] 曹雪, 柯长青. 基于对象级的高分辨率遥感影像分类研究[J]. 遥感信息, 2006, 21(5): 27-30.
CAO Xue, KE Changqing. Classification of High-Resolution Remote Sensing Images Using Object-Oriented Method[J]. *Remote Sensing Information*, 2006, 21(5): 27-30.

[6] 张馨月. 基于DCNN的高分辨率遥感图像场景分类[D]. 长春: 吉林大学, 2019.
ZHANG Xinyue. Scene Classification of High Resolution Remote Sensing Images Based on DCNN[D]. Changchun: Jilin University, 2019.

[7] 李学龙, 史建华, 董永生, 等. 场景图像分类技术综述[J]. 中国科学: 信息科学, 2015, 45(7): 827-848.
LI Xuelong, SHI Jianhua, DONG Yongsheng, et al. A Survey on Scene Image Classification[J]. *Scientia Sinica* (*Informationis*), 2015, 45(7): 827-848.

[8] 冯权泷, 陈泊安, 李国庆, 等. 遥感影像样本数据集研究综述[J]. 遥感学报, 2022, 26(4): 589-605.
FENG Quanlong, CHEN Bo'an, LI Guoqing, et al. A Review for Sample Datasets of Remote Sensing Imagery[J]. *National Remote Sensing Bulletin*, 2022, 26(4): 589-605.

[9] LECUN Y, BENGIO Y, HINTON G. Deep Learning[J]. *Nature*, 2015, 521(7553): 436-444.

[10] CHENG G, YANG C Y, YAO X W, et al. When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification *via* Learning Discriminative CNNS[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(5): 2811-2821.

[11] XU K J, HUANG H, DENG P F, et al. Two-Stream Feature Aggregation Deep Neural Network for Scene Classification of Remote Sensing Images

[J]. *Information Sciences*, 2020, 539: 250-268.

[12] SHI C P, WANG T, WANG L G. Branch Feature Fusion Convolution Network for Remote Sensing Scene Classification[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 5194-5210.

[13] HE N J, FANG L Y, LI S T, et al. Skip-Connected Covariance Network for Remote Sensing Scene Classification[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(5): 1461-1474.

[14] 时永欣, 周维勋, 邵振峰. 融合多尺度注意力的多视角遥感影像场景分类[J]. 武汉大学学报(信息科学版), 2024, 49(3): 366-375.
SHI Yongxin, ZHOU Weixun, SHAO Zhenfeng. Multi-view remote Sensing Image Scene Classification by Fusing Multi-scale Attention[J]. *Geomatics and Information Science of Wuhan University*, 2024, 49(3): 366-375.

[15] 虞欣, 郑肇葆, 孟令奎, 等. 遥感影像场景识别的贝叶斯共轭批次归一化方法[J]. 武汉大学学报(信息科学版), 2023, DOI: 10.13203/j.whugis20220632.
YU Xin, ZHENG Zhaobao, MENG Lingkui, et al. Scene Recognition of Remotely Sensed Images Based on Bayes Adjoint Batch Normalization[J]. *Geomatics and Information Science of Wuhan University*, 2023, DOI: 10.13203/j.whugis20220632.

[16] 毛丽君, 李明诗. GEE环境下联合Sentinel主被动遥感数据的国家公园土地覆盖分类[J]. 武汉大学学报(信息科学版), 2023, 48(5): 756-764.
MAO Lijun, LI Mingshi. Integrating Sentinel Active and Passive Remote Sensing Data to Land Cover Classification in a National Park from GEE Platform[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(5): 756-764.

[17] 卢晓燕. 面向高分辨率遥感影像大范围道路提取的深度学习方法研究[J]. 武汉大学学报(信息科学版), 2023, 48(5): 821.
LU Xiaoyan. Deep Learning Method for Large-Scale Road Extraction from High Resolution Remote Sensing Imagery[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(5): 821.

[18] 张新长, 黄健锋, 宁婷. 高分辨率遥感影像耕地提取研究进展与展望[J]. 武汉大学学报(信息科学版), 2023, 48(10): 1582-1590.
ZHANG Xinchang, HUANG Jianfeng, NING Ting. Progress and Prospect of Cultivated Land Extraction from High Resolution Remote Sensing Images[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(10): 1582-1590.

[19] 王严, 刘万军, 谭亚丽, 等. 卷积神经网络滑窗识别结合语义推理的城市功能区分类[J]. 武汉大学学报(信息科学版), 2023, 48(6): 950-959.
WANG Yan, LIU Wanjun, TAN Yali, et al. Classification of Urban Functional Areas by Convolution Neural Network Recognition Combined with Sliding Window and Semantic Reasoning[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(6): 950-959.

[20] 周亚男, 何金珂, 冯莉, 等. 深度时空卷积网络支持的地块尺度作物种植类型分类[J]. 武汉大学学报(信息科学版), 2023, DOI: 10.13203/j.whugis20230136.
ZHOU Ya'nan, HE Jinke, FENG Li, et al. Parcel-Scale Crop Type Classification Using Tile-Slice-Based Spatial-Temporal Convolutional Networks[J]. *Geomatics and Information Science of Wuhan University*, 2023, DOI: 10.13203/j.whugis20230136.

[21] 许慧琳. 样本受限条件下的高光谱遥感影像分类方法研究[J]. 武汉大学学报(信息科学版), 2023, 48(5): 822.
XU Huilin. Sample Constrained Conditions Methods for Hyperspectral Remote Sensing Image Classification[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(5): 822.

[22] WANG M, DENG W H. Deep Visual Domain Adaptation: A Survey[J]. *Neurocomputing*, 2018, 312: 135-153.

[23] FARAHANI A, VOGHOEI S, RASHEED K, et al. A Brief Review of Domain Adaptation[C]//The 16th International Conference on Data Science, Las Vegas, USA, 2021.

[24] 范苍宁, 刘鹏, 肖婷, 等. 深度域适应综述: 一般情况与复杂情况[J]. 自动化学报, 2021, 47(3): 515-548.
FAN Cangning, LIU Peng, XIAO Ting, et al. A Review of Deep Domain Adaptation: General Situation and Complex Situation[J]. *Acta Automatica Sinica*, 2021, 47(3): 515-548.

[25] PANAREDA B P, GALL J. Open Set Domain Adaptation[C]//IEEE International Conference on Computer Vision, Venice, Italy, 2017.

[26] WILSON G, COOK D J. A Survey of Unsupervised Deep Domain Adaptation[J]. *ACM Transactions on Intelligent Systems and Technology*, 2020, 11(5): 1-46.

[27] GANIN Y, LEMPITSKY V. Unsupervised Domain Adaptation by Backpropagation [C]//The

32nd International Conference on Machine Learning, Lille, 2015.

[28] 许夙晖, 慕晓冬, 张雄美, 等. 结合对抗网络与辅助任务的遥感影像无监督域适应方法[J]. 测绘学报, 2017, 46(12): 1969-1977.
XU Suhui, MU Xiaodong, ZHANG Xiongmei, et al. Unsupervised Remote Sensing Domain Adaptation Method with Adversarial Network and Auxiliary Task[J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(12): 1969-1977.

[29] SUN B C, SAENKO K. Deep CORAL: Correlation Alignment for Deep Domain Adaptation[C]// ECCV 2016 Workshop, Amsterdam, 2016.

[30] WANG X S, LI Y R, CHENG Y H. Hyperspectral Image Classification Based on Unsupervised Heterogeneous Domain Adaptation CycleGan[J]. *Chinese Journal of Electronics*, 2020, 29(4): 608-614.

[31] 周少康. 基于无监督领域适应的高光谱图像分类[D]. 徐州:中国矿业大学, 2022.
ZHOU Shaokang. Hyperspectral Image Classification Based on Unsupervised Domain Adaptation [D]. Xuzhou: China University of Mining and Technology, 2022.

[32] 陈渠川. 面向类别分布差异的无监督域适应学习研究[D]. 南京: 南京邮电大学, 2022.
CHEN Quchuan. Research on Unsupervised Domain Adaptive Learning for Category Distribution Differences [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2022.

[33] BAKTASHMOTLAGH M, HARANDI M T, LOVELL B C, et al. Unsupervised Domain Adaptation by Domain Invariant Projection[C]//IEEE International Conference on Computer Vision, Sydney, Australia, 2013.

[34] WEI G, LAN C, ZENG W, et al. ToAlign: Task-Oriented Alignment for Unsupervised Domain Adaptation[J]. *Advances in Neural Information Processing Systems*, 2021, 34: 13834-13846.

[35] 周彪. 基于对抗学习的域适应方法研究[D]. 合肥: 安徽大学, 2021.
ZHOU Biao. Research on Domain Adaptation Method Based on Adversarial Learning[D]. Hefei: Anhui University, 2021

[36] 孙语泽. 迁移图像语义分割研究[D]. 南京: 南京邮电大学, 2020.
SUN Yuze. Research on Semantic Segmentation of Migrated Images[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2020.

[37] 杨存建, 周成虎. 基于知识的遥感图像分类方法的探讨[J]. 地理学与国土研究, 2001, 17(1): 72-77.
YANG Cunjian, ZHOU Chenghu. Investigation on Classification of Remote Sensing Image on Basis of Knowledge [J]. *Geography and Territorial Research*, 2001, 17(1): 72-77.

[38] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization [C]// IEEE International Conference on Computer Vision, Venice, Italy, 2017.

[39] SIMONYAN K, VEDALDI A, ZISSERMAN A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps [C]// International Conference on Learning Representations, Banff, AB, Canada, 2014.

[40] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-Adversarial Training of Neural Networks [J]. *The Journal of Machine Learning Research*, 2016, 17(1): 2096-2030.

[41] ZHAO Z C, LI J Q, LUO Z, et al. Remote Sensing Image Scene Classification Based on an Enhanced Attention Module[J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 18(11): 1926-1930.

[42] XIA G S, HU J, HU F, et al. AID: A Benchmark Data Set for Performance Evaluation of Aerial Scene Classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(7): 3965-3981.

[43] WANG Q, LIU S T, CHANUSSOT J, et al. Scene Classification with Recurrent Attention of VHR Remote Sensing Images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(2): 1155-1167.

[44] ZHOU W X, NEWSAM S, LI C M, et al. PatternNet: A Benchmark Dataset for Performance Evaluation of Remote Sensing Image Retrieval[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 145: 197-209.

[45] XU K, HUANG H, LI Y, et al. Multilayer Feature Fusion Network for Scene Classification in Remote Sensing [J]. *IEEE Geoscience and Remote Sensing Letters*, 2020, 17(11): 1894-1898.

[46] LIANG J, DENG Y, ZENG D. A Deep Neural Network Combined CNN and GCN for Remote Sensing Scene Classification[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 4325-4338.

[47] GAO Y, SHI J, LI J, et al. Remote Sensing Scene Classification Based on High-Order Graph Convolu-

(下转第 1381 页)