

基于阴影方向先验的高分辨率遥感影像 地物阴影检测

祁昆仑¹, 马欣悦², 金准¹, 卿雅娴³, 李真强⁴, 杨超⁴, 吴华意³

1. 中国地质大学(武汉) 地理与信息工程学院, 武汉 430074;

2. 中移(杭州)信息技术有限公司, 杭州 311100;

3. 国家地理信息系统工程技术研究中心, 武汉 430074;

4. 武汉大学 测绘遥感信息工程国家重点实验室, 武汉 430079

摘要: 地物阴影的标注成本昂贵且难以全面覆盖高分辨率遥感影像蕴含的丰富信息, 其训练样本的匮乏, 严重限制了监督学习模型的性能。针对上述问题, 本文提出了一种基于阴影方向先验的高分辨率遥感影像地物阴影检测方法, 该方法探究了阴影方向先验对于遥感影像地物阴影高级语义特征表达的有效性, 并基于阴影的方向性先验构建了遥感地物阴影检测辅助任务, 实现自监督的遥感地物阴影检测方法。本文方法设计了一种方向变换无关性噪声处理机制和一种自监督阴影检测的数据增强策略, 进一步提升深度神经网络对于地物阴影关键特征的学习能力。在 AISD 数据集上的实验结果表明, 本文方法仅使用少量的标签即可显著提升地物阴影检测精度, 并且地物阴影的边界更加平滑和规整, 更接近于地面真实情况。

关键词: 阴影检测, 自监督学习, 数据增强, 阴影方向先验, 高分辨率遥感影像

中图分类号: TP75/P2

引用格式: 祁昆仑, 马欣悦, 金准, 卿雅娴, 李真强, 杨超, 吴华意. 2025. 基于阴影方向先验的高分辨率遥感影像地物阴影检测. 遥感学报, 29(3): 776-791

Qi K L, Ma X Y, Jin Z, Qing Y X, Li Z Q, Yang C and Wu H Y. 2025. Shadow detection method for high-resolution remote sensing images on the basis of shadow direction prior. National Remote Sensing Bulletin, 29(3): 776-791 [DOI: 10.11834/jrs.20243314]

1 引言

随着全球空天地一体化信息网络的进步, 兼具覆盖范围广、获取成本低、宏观可视化水平高等众多优势的遥感影像成为对地观测技术的主体研究对象(童旭东, 2016)。海量的高分影像可以用于更加精细的地物识别, 然而遥感影像中普遍存在大量阴影, 如城市中建筑物、高大树木和桥梁等地物的阴影, 严重制约遥感影像的应用(Shahtahmassebi 等, 2013; Shi 等, 2023; 廖楚江等, 2023)。地物阴影的准确识别, 是提升高分辨率遥感影像视觉效果与可辨识度的一个重要研究方向(李德仁等, 2012)。遥感影像地物阴影检测方法是一种利用特定判别条件识别地物阴影像素

的技术手段。根据不同阴影检测方法对先验条件的分析角度和依赖程度, 我们可以将遥感影像地物阴影检测的研究大致划分为两类: 基于物理模型的方法和基于特征的方法。

基于物理模型的方法通过探索光谱特性或建立照明模型等方式实现地物阴影区域的提取。李艳和宫鹏(2005)利用 DSM 阴影仿真和高度场光线跟踪技术, 提出了一种建筑物的阴影检测方法。Makarau 等(2011)根据照射光源的不同, 提出了一种黑体辐射模型, 以区分阴影区域与非阴影区域。段光耀等(2014)通过分析阴影的光谱特性, 提出一种特征分量构建与面向对象结合的阴影提取方法。Li 等(2016)实现了一种高分影像和自然图像的阴影检测方法, 提取阴影轮廓之后使用

收稿日期: 2023-07-24; 预印本: 2024-01-15

基金项目: 湖北珞珈实验室专项基金(编号: 220100034); 湖北省科技厅重大专项(编号: 2020AAA004)

第一作者简介: 祁昆仑, 研究方向为遥感影像解译和时空大数据。E-mail: qikunlun@cug.edu.cn

通信作者简介: 李真强, 研究方向为时空大数据挖掘。E-mail: lizhenqiang@cug.edu.cn

色度和亮度参数优化阴影的检测结果。Wang等(2017)通过构造成像的几何关系确定地物阴影的大致位置,利用抠图技术进一步准确提取地物阴影。然而在复杂的现实场景中,遥感影像精确的物理模型通常难以获取,并且这类方法过度依赖于数据质量和使用方式,因此不利于实际应用。基于特征的遥感影像地物阴影检测方法利用光谱、纹理、形状和上下文语义信息等特性提取遥感影像中的阴影区域。传统基于特征的方法通过探索各种手工设计特征来提取遥感影像中的地物阴影。例如,于东方等(2008)致力于拟合高分全色遥感影像的灰度统计直方图函数,采用一阶差分法计算直方图势函数,并利用其首个谷点的灰度值作为阴影分割的阈值,提升阴影检测的精度。马川等(2012)结合并优化了形态学和阈值分割两种算法,从而更加精确地获取高分辨率影像中的建筑物阴影边界。Su等(2016)运用双峰直方图阈值和抠图技术,构建了一种高分全色遥感影像阴影检测方法。然而上述传统的方法通常需要人为设定分割阈值,无法适应复杂环境下的遥感影像地物阴影检测。

随着计算机视觉技术的高速发展,深度学习方法凭借卷积神经网络结构化和层次化的高效学习能力,为基于特征的阴影检测方法提供了新思路。受限于遥感卫星重访周期和太阳光照的多样变化,遥感影像阴影检测数据的制作成本昂贵,难以获取足够的标签样本用于深度学习模型的训练,所以这类方法目前主要应用于自然图像的阴影检测(Hou等,2021;Khan等,2014;Shen等,2015)。Luo等(2020)发布了首个面向语义分割的航空遥感地物阴影检测数据集AISD(Aerial Imagery dataset for Shadow Detection),同时构建了一种用于地物阴影检测的卷积神经网络。次年,该团队基于多任务学习框架又提出了一种融合边缘信息和显著性信息的模型(Luo等,2021),在AISD数据集上进一步提升了地物阴影检测的精度。此后,学者们基于AISD数据集开展了一系列地物阴影特征提取的相关研究(Liu等,2021a;Xie等,2022;Yang等,2021;Zhu等,2022)。然而,有标注样本数据规模有限一定程度上限制了深度学习模型的泛化性能(Belkin等,2019;周培诚等,2021),已然成为制约高分辨率遥感影像地物阴影检测领域发展的主要问题之一。

自监督视觉特征表达能够有效减轻卷积神经网络对于标注样本规模的依赖,已有研究表明自监督学习方法能够灵活获取遥感影像数据内部蕴含的丰富特征(陶超等,2021)。特别是在高分影像场景分类任务中,基于对比约束的自监督方法已经取得了超过大规模有标注自然图像预训练模型的性能(Jung等,2022;Stojnic和Risojevic,2021;Tao等,2022)。对于高分影像地物阴影检测的问题,地物阴影在遥感影像中表现为信息的衰减,上述自监督方法所捕获的语义特征难以满足阴影这一非地物类别目标的特征表达,因此目前已有的方法直接用于地物阴影表征的效果存在一定的局限性。因此,如何实现基于自监督学习的高分辨遥感影像地物阴影检测,针对地物阴影的特征表达设计适合于自监督学习的辅助任务,充分利用遥感影像数据自身隐含的地物阴影语义表征,是需要关注的重点。

在高分辨率遥感影像中,地物阴影的分布展现出明显的定向规律。举例而言,在位于太阳赤道以北的地区,地物阴影的方向在一天中会随着太阳高度角的变化而变化。在上午,阴影倾向于指向西北;在正午,阴影方向趋向正北;而在下午,阴影方向则偏向东北。同样地,在太阳赤道以南的地区,上午地物阴影指向西南,正午时指向正南,下午则指向东南。地物阴影的方向可以被推测,具体取决于太阳方位角和太阳高度角。

对于地物阴影方向判别任务,模型需要同时识别地物本身以及与之相关的阴影,方能精确判断地物阴影的朝向。因此,我们提出通过设计地物阴影方向的分类任务,协助卷积神经网络学习地物和其阴影的特征表达。

然而,为了精确计算这些参数,通常需要准确的成像时间和地理位置信息,但在实际应用中,获取这些信息往往十分困难。因此,鉴于在太阳赤道以北或以南地区,阴影通常朝南或朝北,相应地,本研究将地物阴影的方向划分为两个主要类别,即 $[0^\circ, 180^\circ]$ 和 $[180^\circ, 360^\circ]$ 。如图1(a)所示,黄色区域代表地物阴影位于 $[0^\circ, 180^\circ]$,即方向一的范围内,而蓝色区域代表地物阴影位于 $[180^\circ, 360^\circ]$,即方向二的范围内。图1(b)和(c)展示了对应于遥感影像中地物阴影的这两个方向示例。

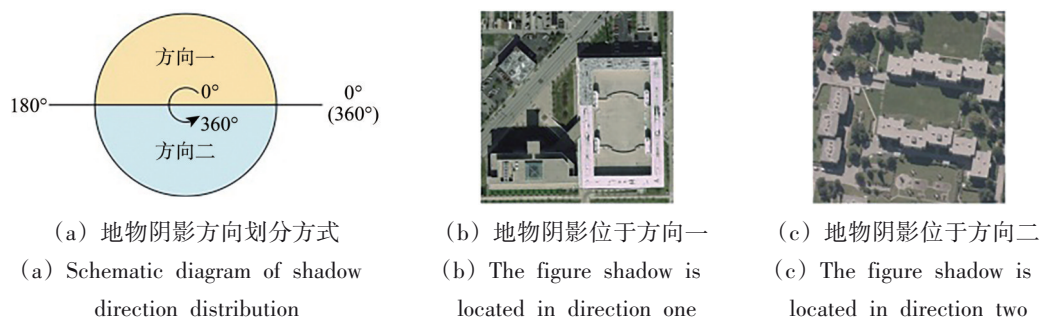


图1 地物阴影的方向划分示例图

Fig.1 Example of directional division of ground object shadows

为了解决不同阴影方向类别下的图像数量不平衡问题,针对阴影方向的特点,我们采用了一种数据增广策略,即对同一地区的原始图像进行垂直翻转,如图2所示。这一策略的目的在于构建一个数据集,其中不同阴影方向类别的样本数量保持均衡,从而降低本文方法对于不同地区影像的依赖性。最终,我们引入一个地物阴影方向判别的辅助任务,旨在从无标注的遥感图像中学习在地物阴影相关的语义信息,以提升在有限样本情况的遥感影像地物阴影检测精度。

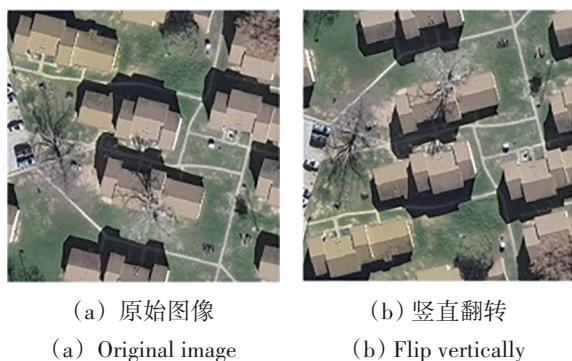


图2 遥感影像竖直翻转示例图

Fig. 2 Example of vertical flipping of remote sensing images

综上,本文提出了一种基于阴影方向先验的高分辨率遥感影像地物阴影检测方法,在AISD阴影检测数据集上的实验结果证明了本文方法的有效性与先进性。本文方法利用地物阴影的方向先验知识构建自监督辅助任务,通过竖直翻转影像构造伪标签;通过一种基于协同指导的方向变换无关性噪声处理机制DTICT (Direction Transformation Independent Co-Teaching),提高模型对于噪声样本的辨别能力,以减轻无阴影和特殊方向地物阴影的遥感影像对网络训练的干扰;同时引入融合了对比剪裁和色彩变换的数据

增强策略以帮助网络充分学习地物阴影的关键特征。

2 方法

2.1 遥感影像地物阴影方向判别模型

为了验证地物阴影的方向先验及其应用于卷积神经网络的可行性,本文通过图像的翻转构建图像二分类任务,训练卷积神经网络识别图像是否被翻转(本文中若无特殊说明外,翻转则指代竖直翻转)。本文采用在遥感影像分类中广泛使用的ResNet-18(He等,2016)网络进行实验,为了避免由图像压缩引起的图像边缘像素非对称性(Lin等,2020),需要裁剪大小为16个像素的图像边界。图3为本文建立的遥感影像地物阴影方向判别模型,具体步骤为:(1)输入的图像首先进行重采样整体放大32个像素,然后通过中心剪裁去除图像四周边界的16个像素;(2)设置原始图像的标签为Class 1(原始图像),翻转操作后的图像标签为Class 2(翻转图像),训练ResNet-18网络模型;(3)网络模型根据提取到的特征,通过分类器输出类别预测的概率。

2.2 地物阴影自监督辅助任务模型

针对高分辨遥感影像地物阴影检测任务,本文提出了一种自监督预训练方法。利用地物阴影的方向先验知识,不依赖于人工标签,而是从遥感数据自身出发构造地物阴影的监督信息,提升卷积神经网络对于地物阴影的语义表征能力。图4为本文提出的遥感影像地物阴影自监督辅助任务模型,该模型的具体流程为

(1)对输入的无标签图像进行预处理,根据地物阴影的方向性特征,将原始图像进行竖直翻转,从而构造出影像数据的类别标签(伪标签)。

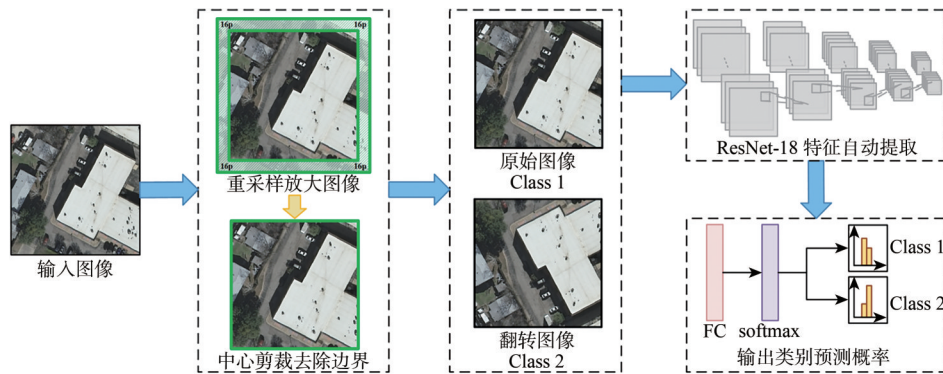


图3 遥感影像地物阴影方向判别模型

Fig.3 Shadow direction discrimination model of remote sensing image figures

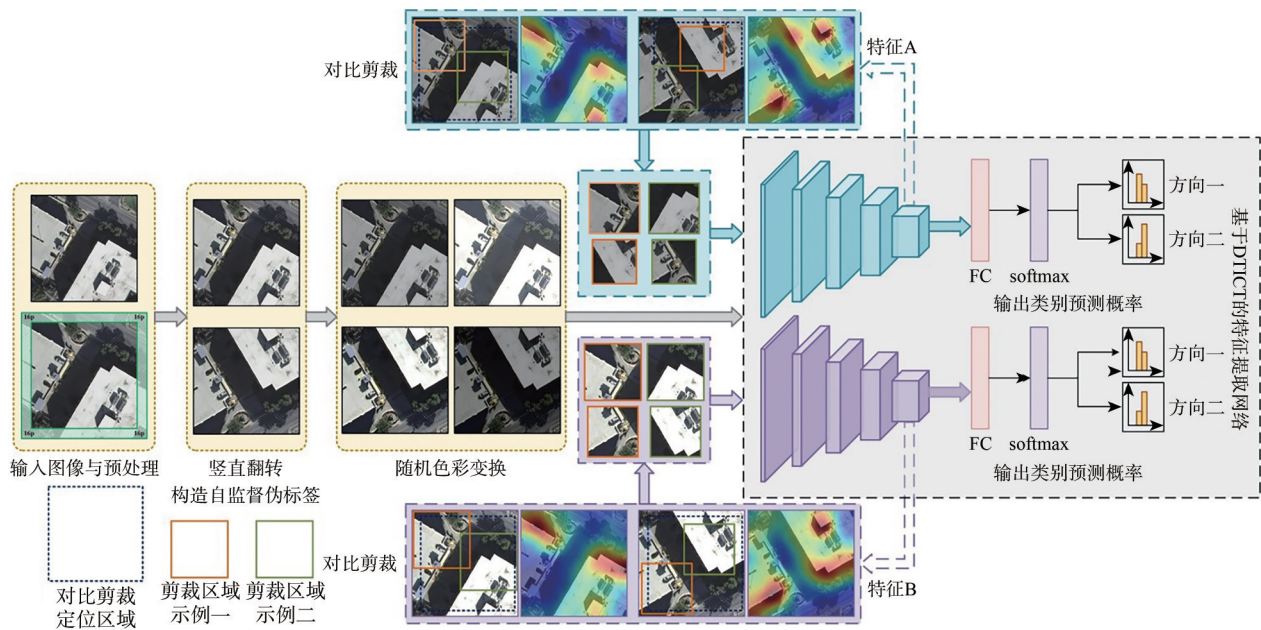


图4 遥感影像地物阴影自监督辅助任务模型

Fig.4 Self-supervised auxiliary task model for shadow detection in remote sensing imagery

(2) 面向地物阴影，建立基于协同指导的方向变换无关性噪声处理机制 DTICT (Direction Transformation Independent Co-Teaching)。针对自监督特征学习过程中的噪声样本，有效抑制噪声样本对网络特征提取的干扰，帮助自监督网络学习地物阴影正确的语义表征。

(3) 设计适合于遥感影像地物阴影自监督特征学习的数据增强策略。借助对比剪裁实现顾及语义信息的图像剪裁，通过随机改变遥感影像的亮度、对比度和饱和度进行图像色彩结构的增强，产生丰富且有价值的训练样本，保证卷积神经网络的充分训练。

2.2.1 地物阴影方向性噪声的去除

针对地物阴影的方向性特征，本文将原始的

遥感影像视为默认方向，标记为正样本，竖直（或水平）翻转变换后的影像为负样本，以此构建两个不同的阴影方向类别，从而实现自监督辅助任务模型的训练。

然而，高分辨率遥感影像中，地物阴影的分布和形态是千变万化的，必然会存在一些无阴影或者特殊方向阴影的图像（图5）。这些区域的方向性判别对于地物阴影特征的学习是毫无意义的，并且这些遥感影像在进行翻转变换后仍然处于默认方向（即自动分配的标签存在噪声）。如果直接将其输入卷积神经网络中，难免将噪声引入模型的训练。因此，针对高分辨率遥感影像地物阴影的多样性和不确定性所引起的方向性噪声数据问题，需要设计一种适合于地物阴影自监督学习网络的噪声样本过滤策略。

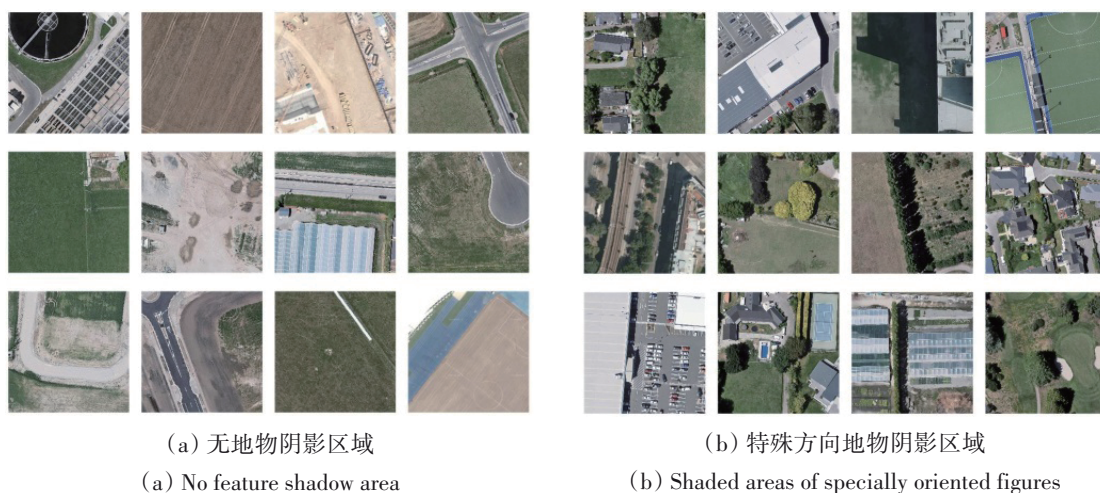


图5 无阴影和特殊方向地物阴影的遥感影像样本示例

Fig.5 Examples of remote sensing image samples without shadows and with shadows in a special orientation

本文借鉴协同指导 (Co-teaching) 的噪声样本选择方法 (Han 等, 2018), 提出了一种基于协同指导的方向变换无关性噪声处理机制 (DTICT)。本方法利用两个基础结构相同的网络分别选择各自认为可能干净的训练数据并交换这些数据以完成自身网络参数的更新, 通过引入方向变换无关性度量筛选阴影自监督学习过程中的正确样本, 进一步提升遥感影像地物阴影自监督特征的泛化性能。

DTICT 机制通过提取每个网络中原始图像和翻转变换后图像的高层特征向量, 计算其余弦距离作为方向变换无关性的度量。图 6 为 Co-teaching 方法与 DTICT 方法的数据流对比示意图。DTICT 方法通过引入方向无关性表达网络对方向变换判别的难易程度, 距离越大表示判别越容易, 样本属于方向性噪声的概率越低; 反之属于噪声样本的概率越高。最终, 结合网络模型的分损失和方向变换无关性度量, 共同筛选出干净的样本。

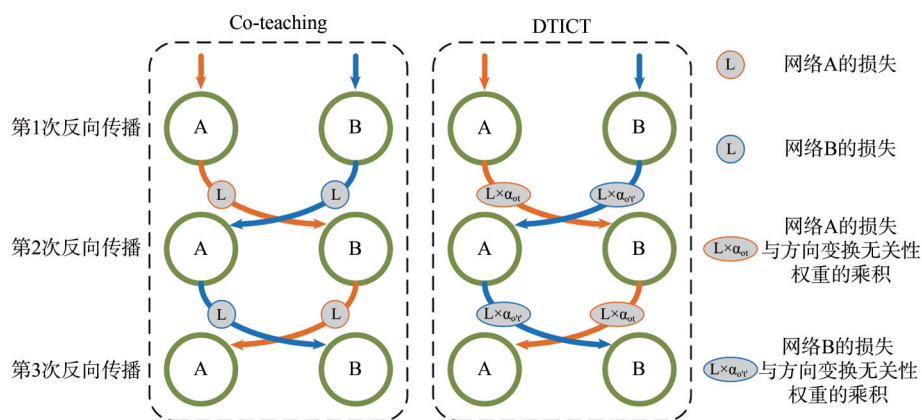


图6 Co-teaching 与 DTICT 数据流对比示意图

Fig.6 Schematic diagram of Co-teaching and DTICT data streams

基于协同指导的方向变换无关性噪声处理机制 (DTICT) 的算法流程描述如下所示。其中, f (网络参数为 w_f) 和 g (网络参数为 w_g) 表示两个网络; D 表示数据集, $\bar{D}$ 表示一个批量的样本集合; v_o 和 v_i 以及 $v_{o'}$ 和 $v_{i'}$ 分别表示不同网络提取的原始图像和方向变换后图像最后一个卷积层的特征向量, α_o 和 $\alpha_{o'}$ 表示不同网络中样本的方向变换

无关性权重; $R(e)$ 表示样本在第 e 个回合时从每个批次中选择无噪声样本的数量, 随着回合数的增多, 逐渐减少选择样本的比例; $r \in (0, 1]$ 表示样本比例常数, e_k 表示第 k 个回合样本比例的选择达到最小值; $L(f, D')$ 为样本的损失函数。

算法: DTICT 机制

1: 输入 w_f 和 w_g 、学习率 δ 、样本比例常数 r , 总共

的回合数 E 和迭代数 T ;

从 $e = 1, 2, \dots, E$ 开始:

2: 加载训练数据集 D ;

从 $t = 1, 2, \dots, T$ 开始:

3: 从 D 中取出 $\bar{D}$;

4: 计算 $\alpha_{ot} = \cos(\mathbf{v}_0, \mathbf{v}_t)$;

5: 计算 $\alpha_{o't'} = \cos(\mathbf{v}_{0'}, \mathbf{v}_{t'})$;

6: 获得 $\bar{D}_f = \operatorname{argmin}_{D': |D'| \geq R(e)|\bar{D}|} \alpha_{ot} L(f, D')$;

7: 获得 $\bar{D}_g = \operatorname{argmin}_{D': |D'| \geq R(e)|\bar{D}|} \alpha_{o't'} L(g, D')$;

8: 更新 $w_f = w_f - \delta \nabla L(f, \bar{D}_f)$;

9: 更新 $w_g = w_g - \delta \nabla L(g, \bar{D}_g)$;

结束

10: 更新 $R(e) = 1 - \min\left\{\frac{e}{e_k} r, r\right\}$;

结束

11: 输出 w_f 和 w_g 。

2.2.2 遥感影像数据增强策略的设计

自监督学习需要从大规模无监督数据中充分挖掘自身有价值的监督信息，利用构造的监督信息训练模型，使得模型学习到下游任务的正确语义表征。目前，在图像分类任务中，随机剪裁

(RandomCrop) 是以监督学习方式训练卷积神经网络最为常用的数据增强手段之一。随机剪裁的方式没有考虑图像自身的语义信息，极有可能裁剪到没有任何阴影的区域，从而将无监督数据中对下游任务造成干扰的信息引入卷积神经网络的训练。

基于上述分析，本文引入一种顾及语义信息的对比剪裁 (Peng 等, 2022) (ContrastiveCrop) 方式，协助网络充分训练从而更好地学习高分辨率遥感影像地物阴影的关键特征。对比剪裁的基本思想是，利用网络训练过程中的特征图来指导图像剪裁区域的选取，以保证语义信息的准确性，同时借助中心压制采样方法 (Center-suppressed Sampling) 增加每次剪裁的差异性，以进一步提升训练样本的丰富性。另外，对比剪裁即插即用，其产生的计算量很小，不会给网络的训练增加负担。图7为随机剪裁与对比剪裁的效果对比示例，随机剪裁可能会生成全部为背景的采样区域，对网络的特征学习产生干扰；而对比剪裁在语义信息的指导下生成了价值较高的采样区域，明显更适合用于自监督学习辅助任务。

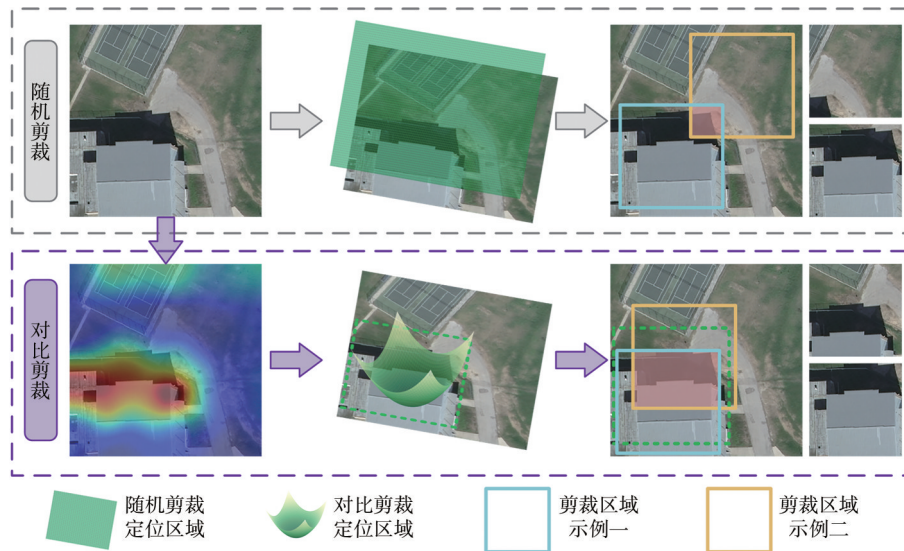


图7 随机剪裁与对比剪裁的效果对比示例

Fig.7 Example of how random clipping versus contrast clipping compares

此外，色彩特征是高分辨率遥感影像解译最为关键的特征之一 (慕晓冬 等, 2021)。本文采用颜色抖动 (Color Jitter) 的数据增强方式对高分影像的色彩结构进行建模。随机变换输入图像的亮度、对比度和饱和度，使卷积神经网络可以充分

学习图像色彩变化所带来的差异，辅助网络掌握遥感影像地物及其阴影在不同颜色下的特征表示，从而模拟不同光照条件下遥感影像的阴影检测，有利于提升模型对复杂语义情景的感知能力。

综上所述，本文设计了一种适合于高分影像

地物阴影自监督特征学习的数据增强策略。该策略能够在不改变地物阴影方向性特征的前提下,对遥感影像进行对比剪裁和随机色彩增强,使卷积神经网络更加充分地训练以学习地物阴影正确且丰富的语义信息,有助于改善自监督预训练网络在下游地物阴影检测任务中的性能表现。基于对比剪裁和色彩变换的数据增强策略如下所示:

算法: 基于对比剪裁和色彩变换的数据增强策略

1: 输入图像 G , 剪裁尺寸 s , 剪裁高宽比 r , 颜色抖动函数 T_{cj} ;

2: $G_{cj} = T_{cj}(G)$;

3: $F = \text{Forward}(G_{cj})$;

4: $M = \text{Normalize}(F)$;

5: $h = \sqrt{s \cdot r}$, $w = \sqrt{s/r}$;

6: $B = S(P[M > \delta])$;

7: $x = B_{x0} + (B_{x1} - B_{x0}) \cdot \gamma$, $\gamma \sim \beta(\alpha, \alpha)$;

8: $y = B_{y0} + (B_{y1} - B_{y0}) \cdot \varepsilon$, $\varepsilon \sim \beta(\alpha, \alpha)$;

9: 输出 $\text{Corp } G = (x, y, h, w)$ 。

2.3 基于自监督预训练的地物阴影检测流程

自监督预训练的目标是充分学习高分辨率遥感影像地物阴影的视觉表征,其性能的优劣需要根据下游任务的表现进行评估。本文提出的基于自监督预训练的遥感影像地物阴影检测方法总体流程如图8所示。该方法的具体步骤为

(1) 在模型的预训练阶段,采用本文所提出基于阴影方向先验的自监督特征学习方法完成预训练,获取预训练权重 $W1$ 。此阶段使用无标注数据,编码器 E 为去除全连接层的基础网络。

(2) 在模型的正式训练阶段,编码器 E 加载上一阶段得到的模型预训练权重 $W1$ 并进行参数微调,编码器 E 与解码器 D 组成完整的遥感影像地物阴影检测网络,模型训练完成后得到更新的权重 $W2$ 。此阶段需要使用数据标签,并且设置编码器 E 的学习率适当小于解码器 D 的学习率。

(3) 在模型的测试阶段,遥感影像地物阴影检测网络加载上一阶段得到的权重 $W2$,完成模型性能的评估。此阶段同样需要使用数据标签。

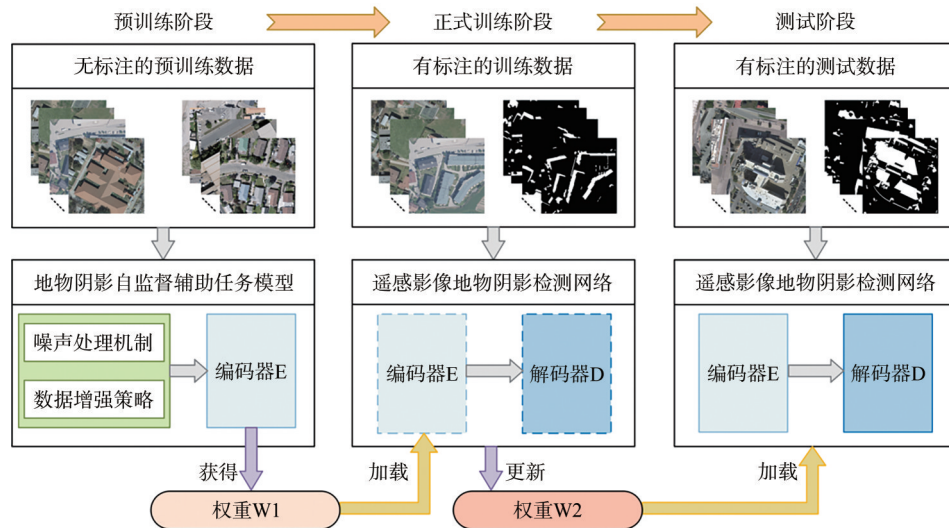


图8 基于自监督预训练的遥感影像地物阴影检测流程

Fig.8 Shadow detection of remote sensing images based on self-supervised pre-training

3 实验结果与分析

3.1 实验配置及评价指标

本实验采用分类总体准确度 OA (Overall Accuracy)、交并比 IoU (Intersection over Union) 和 F_1 分数 F_1 (F_1 -score) 作为评价指标,指标的定义如式 (1) 一式 (5) 所示。其中, TP 指正类被

判定为正类; FN 指正类被判定为负类; FP 指负类被判定为正类; TN 指负类被判定为负类。

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \quad (1)$$

$$Precious = \frac{TP}{TP + FP} \times 100\% \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (3)$$

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\% \quad (4)$$

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \times 100\% \quad (5)$$

实验环境是 Windows10 服务器，配置 Intel (R) Xeon (R) W-2133 处理器、64 G 运行内存、NVIDIA GeForce RTX 2080 Ti 显卡，基于 Pytorch 深度学习框架实现算法模型。使用交叉熵损失函数，优化器采用 Adam，学习率 (learning rate) 设置为 $1\text{E-}4$ ，权重衰减 (weight decay) 系数设置为 $1\text{E-}5$ ，批量大小 (batch size) 设置为 16。

3.2 数据集及预处理

本文选用 AISD 阴影检测数据集 (Luo 等, 2020) 进行实验，该数据集覆盖了北半球 5 个具有不同城市结构和光照特征的地区，保证了数据集的复杂性和多样性。该数据集共包含 514 张空间分辨率为 0.3 m 的影像，本文实验采用数据集的 80% 作为训练集，验证集和测试集则分别为 10%，即训练集为 412 张，验证集和测试集各 51 张。训练集采用 256×256 大小的窗口、每次滑动 64 个像素的剪裁策略，最终得到 11836 张训练图像。验证集

和测试集则按照 256×256 大小的窗口进行无重叠剪裁。

在模型的预训练阶段，无标签数据量的增加有助于提升自监督模型的性能 (Jing 和 Tian, 2021; Liu 等, 2021b)。在高分辨率遥感影像中，建筑物是产生地物阴影的典型因素之一，建筑物数据集通常存在大量广泛分布的地物阴影。因此，本文利用与 AISD 数据分辨率相当的 WHU 建筑物数据集 (Ji 等, 2019)，进行地物阴影检测网络自监督预训练阶段数据量的扩充，该数据集覆盖了南半球不同外观和用途超过 18.7 万座独立建筑物，将其包含的 8188 张大小为 512×512 的图像全部无重叠剪裁为 256×256 ，最终获得 32752 张图像。WHU 数据集和 AISD 数据集采集的地点不同，前者 WHU 阴影基本位于地物的南方，而后者阴影大多位于地物的北方。由于本实验是基于遥感影像地物阴影的方向性特征辅助构建自监督学习模型，因此需要对 WHU 数据集的图像全部进行竖直翻转，将其与 AISD 数据集中地物阴影的分布方向保持一致，图 9 为两个数据集的部分样本示例。

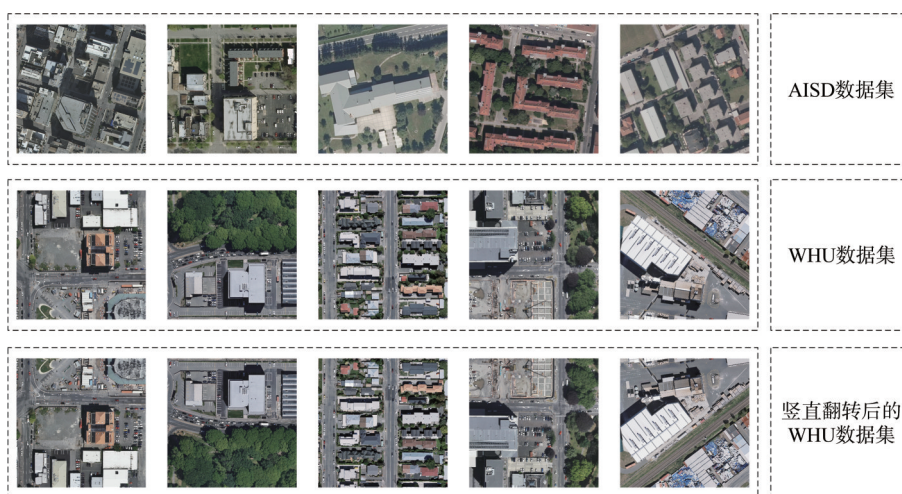


图9 AISD 和 WHU 数据集的部分样本示例

Fig.9 Examples of images in AISD and WHU datasets

基于地物阴影方向先验的自监督特征学习阶段采用竖直翻转后的 WHU 数据集和 AISD 数据集的图像数据进行训练，无需使用标签数据。为了验证本文自监督学习方法的有效性，需要评估下游阴影检测模型的性能，为此，我们使用 AISD 阴影检测数据集中的图像和标签数据对该模型进行训练。

3.3 地物阴影方向先验的分析

本文采用所提出的方法在 AISD 数据集上训练 ResNet-18 分类网络判别遥感影像中阴影的方向 (等价于是否翻转)，实验结果如表 1 所示。已知对于二分类任务，随机判别的准确率为 50%。从本文的实验结果可以看出，对于遥感影像的水平和竖直翻转，模型的分类精度均高达 90% 以上，证

明地物阴影存在明显的方向先验特性。同时，网络模型对于遥感影像竖直翻转的识别精度为99.31%左右，几乎能够正确识别全部样本，比识别水平翻转的精度高6.2%，这表明模型更擅长对竖直翻转的图像进行分类，后文将通过实验对其原因展开详细分析。

表 1 遥感影像水平与竖直翻转识别的精度
Table 1 The accuracy of the model to identify the horizontal and vertical flipping of the remote sensing image

图像翻转方式	OA/%
水平翻转	93.11±0.46
竖直翻转	99.31±0.43

我们使用梯度一类激活图（Selvaraju等，2017）对模型的特征进行可视化分析。根据梯度一类激

活图中模型关注区域的可视化结果，能够对模型精准预测遥感影像是否翻转的原因进行定性分析。本文采用Jet彩色模式绘制梯度一类激活图，红色和黄色区域代表对模型分类结果的影响较大。图10为遥感影像地物阴影方向判别模型的梯度一类激活图可视化结果，可以发现模型主要关注建筑物、树木和汽车等典型地物及其阴影的相关区域。这表明借助自监督学习的方法，通过遥感影像的翻转构造伪标签辅助前置任务，网络模型可以自动发现遥感影像地物阴影相关的高级语义特征，并且可能学习到比人工标注的数据标签更加丰富的语义信息。将辅助任务提取到的特征用于相关的下游任务，例如建筑物提取、阴影检测以及遥感场景车辆检测等，这些特征可能会提升下游分类任务的性能。

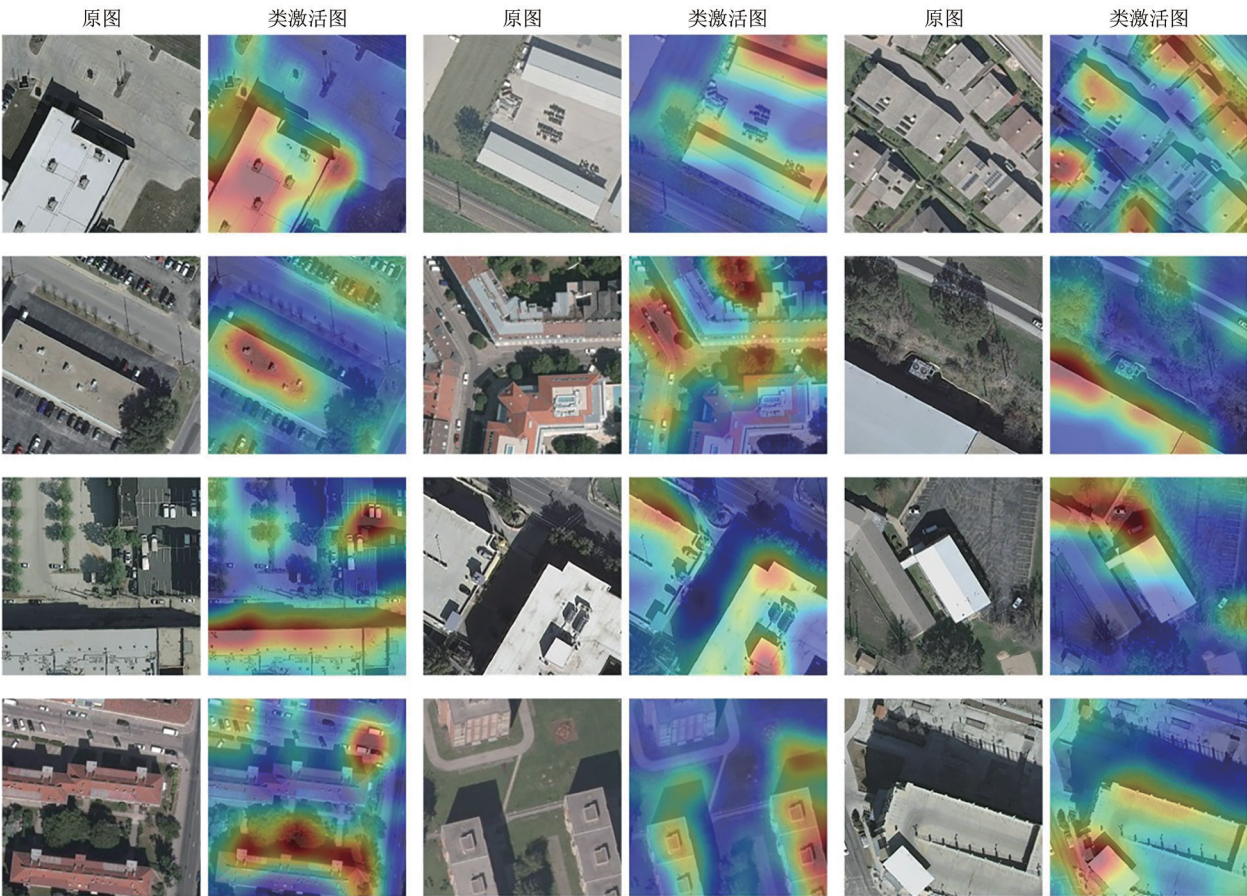


图 10 自监督阴影特征学习网络的梯度一类激活图可视化结果
Fig.10 Gradient-class activation maps of shadow images for self-supervised feature learning model

由于本实验采用的 AISD 数据集采集于北半球，阴影基本位于地物的北方，而网络模型将地物及其阴影作为识别遥感影像是否翻转的重点关注区域，因此阴影与地物的相对位置关系将会直

接影响分类结果。本文选取了一些典型的遥感影像进行说明，图11为遥感影像水平与竖直翻转的地物阴影细节对比，从上至下三行图像依次为原始影像、水平翻转影像和竖直翻转影像。图11中，原

始影像阴影位于地物的正北方，水平翻转不会改变阴影与地物的相对位置，导致网络难以辨别图像是否水平翻转。而竖直翻转将阴影转变为地物的正南方，与原始影像阴影的方向相反，易于网络的学习。在 AISD 数据集中，受成像时间影响，这种一张图像中阴影全部位于地物正北方的情况是极少的，大多数阴影向东或西偏离。原始影像

阴影分别位于地物的东北方和西北方，水平翻转会将这两种情况混淆，即原始影像与水平翻转影像同时存在东北方和西北方的阴影，导致网络无法学习到正确的类别信息。而竖直翻转后，阴影全部位于地物的东南方和西南方，不会对分类结果的辨别产生任何干扰，并且竖直翻转影像的地物阴影方向与原始影像之间的差别比较明显。

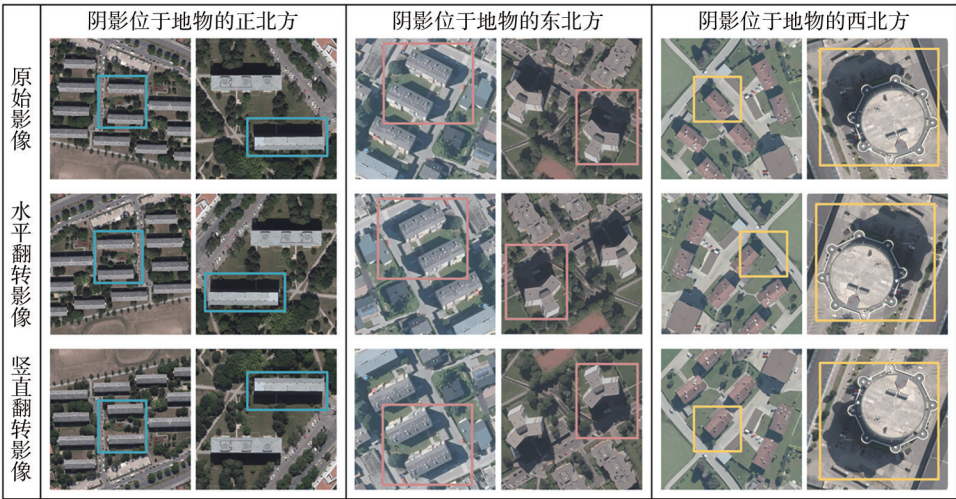


图 11 遥感影像水平与竖直翻转的地物阴影细节对比

Fig.11 The comparison of horizontal and vertical flipped for the shadow in remote sensing images

综上，对于识别遥感影像是否翻转的图像二分类任务，依据遥感影像中地物阴影分布的方向性规律，原始影像与竖直翻转影像之间阴影的方向分布差异显著，便于模型学习地物与阴影之间的位置关联，从而可以取得十分精准的分类结果。

3.4 消融实验

为研究方法中方向性噪声去除机制和数据增强策略中不同组成部分对本文所提出自监督学习算法性能的影响，本文分别以 ResNet18 和 VGG16 为主干网络的 Unet 模型作为基准，评估 DTICT 机制、对比剪裁和色彩变换 3 个模块的作用。在色彩变换模块中，我们经过比较不同参数对图像中阴影的变换效果，最终采用了以下参数范围：亮度范围为 0.6—1.8、对比度范围为 0.6—1.8、以及饱和度范围为 0.4—1.2。在 Unet 模型的正式训练阶段使用 10% 的 AISD 训练集数据（即 $N_{10\%}$ 为 1183 张图像），采用 IoU 和 F_1 作为评价指标。

3 个模块在 AISD 数据集上的消融实验结果如表 2 所示，DTICT、对比剪裁和色彩变换均对高分

辨率遥感影像地物阴影检测模型的性能有促进作用，证明本文所提出自监督辅助模型中每个组成部分均有助于提升遥感影像地物阴影检测的精度。

表 2 AISD 数据集上不同模块的消融实验($N_{10\%}$)

Table 2 Ablation experiments of different modules on the AISD dataset($N_{10\%}$)

DTICT	Contrastive-Crop	Color Jitter	ResNet18		VGG16	
			IoU/%	F_1 /%	IoU/%	F_1 /%
×	×	×	77.42	88.63	79.96	89.59
√	×	×	79.05	89.44	81.01	90.15
√	√	×	80.37	89.81	81.23	90.78
√	√	√	80.42	89.93	81.41	90.96

注：加粗表示最优结果。

其中，在只添加 DTICT 的情况下，模型性能就有明显的提升效果。以 ResNet18 为编码器基础结构时，IoU 和 F_1 分别提升了 1.63% 和 0.81%；以 VGG16 为编码器基础结构时，IoU 和 F_1 分别提升了 1.05% 和 0.56%。这表明遥感数据中存在的一些无地物阴影和特殊方向地物阴影的影像具有方向变

换无关性, 为这些遥感影像自动分配的标签存在噪声, 对自监督模型的训练产生了较多干扰, 卷积神经网络的“记忆”效应导致模型最终拟合这些噪声标签。本文基于协同指导思想建立的方向变换无关性噪声处理机制能够有效过滤噪声样本, 从而提升模型对高分辨率遥感影像复杂语义情景的感知能力, 协助网络模型学习地物阴影正确的语义表征。

结合对比剪裁和色彩变换两个数据增强机制后, 以 ResNet18 为编码器基础结构的模型相比于原始情况, IoU 和 F_1 分别提升了 3.00% 和 1.30%; 以 VGG16 为基础结构的模型对应的 IoU 和 F_1 则分别提升了 1.45% 和 1.37%。其中, 对比剪裁机制利用网络训练过程中的特征图来指导遥感影像剪裁区域的选取, 可以保证地物阴影语义信息的准确性, 并通过中心压制采样增大每次剪裁的差别以提升训练样本的丰富性。色彩变换机制则能够帮助网络学习地物及其阴影在不同颜色下的特征表示, 应对遥感影像色彩变化所带来的差异。实验结果表明, 融合了对比剪裁与色彩变换的数据增强策略能够有效改善网络对于地物阴影关键特征的提取能力。

3.5 与其他方法的对比实验

总结近年来基于 AISD 数据集发表的高分辨率遥感影像地物阴影检测方法与本文方法进行比较。表 3 为不同阴影检测方法的对比结果, 利用这些研究中通用的评价指标 F_1 进行分析, F_1 是语义分割准确率与召回率的调和平均数, 能够较为全面地衡量各方法的有效性。实验结果表明, 与现有的方法相比, 本文方法仅在 10% 的训练数据下对于地物阴影的识别准确率就能达到 90% 以上, 较其他方法在 100% 训练数据下的精度仅有 0.25%—3.12% 的差距, 有力证明了本文方法的优越性。通过自监督学习预训练, 模型能够提前捕捉地物阴影的相关特征, 因此在下游任务中仅使用少量的样本标签即可取得接近于监督学习的性能表现。

3.6 结果的讨论与分析

3.6.1 噪声样本筛选结果对比

本节旨在验证 DTICT 方法相对于 Co-teaching 方法在阴影方向识别任务中噪声去除策略方面的有效性。在第 3.2 节对 WHU 建筑物数据集的预处理

理之后, 我们最终获得了 32752 张样本数据。通过对比实验, 我们发现协同指导方法在噪声去除方面成功剔除的噪声样本占总样本数的 11.3%, 而 DTICT 方法则成功去除了 13.6% 的噪声样本。实验结果验证了本文提出的 DTICT 方法筛选阴影方向判别任务中噪声样本的有效性。图 12 展示了 Co-teaching 方法未能去除而本文方法成功剔除的噪声样本示例。

表 3 不同阴影检测方法的 F_1 结果

Table 3 F_1 -scores for different shadow detection methods

方法	训练数据	$F_1/\%$
CADNet (2021)	100%	91.21
DSSDNet (2020)		91.79
ESPFNet (2021)		92.04
CDANet (2022)		92.99 ± 0.17
本文方法 (ResNet18)	10%	89.93
本文方法 (VGG16)		90.96



图 12 DTICT 去除噪声样本示例

Fig.12 Example of removing noise samples using DTICT

3.6.2 不同训练数据量的模型性能

本文分别以编码器为 ResNet18 和 VGG16 的 Unet 模型作为基准, 不同方法在 AISD 数据集 1% (即 $N_{1\%}=118$ 张)、2% (即 $N_{2\%}=236$ 张)、5% (即 $N_{5\%}=591$ 张) 和 10% (即 $N_{10\%}=1183$ 张) 训练数据下的 IoU 和 F_1 结果对比如表 4 所示, 加粗的数字表示模型在每种训练数据比例下的最高精度。

本节分别与其他 3 种预训练方法进行了对比分析, 其中 MoCo 模型 (He 等, 2020) 是一种先进的自监督学习方法, 于 2020 年首次发表并得到了广泛的应用, 目前已在多种任务中取得了优异的性能。

表 4 各类方法在不同比例训练数据下的IoU和F₁结果

Table 4 IoU and F₁-scores of various methods under different proportions of training data

基础结构	预训练方式	带标签训练数据的比例							
		1%		2%		5%		10%	
		IoU/%	F ₁ /%	IoU/%	F ₁ /%	IoU/%	F ₁ /%	IoU/%	F ₁ /%
ResNet18	Random	72.56	85.70	75.32	87.36	76.92	88.25	77.40	88.51
	ImageNet	73.26	86.04	77.33	88.35	79.34	89.39	80.36	89.72
	MoCo	72.58	85.61	76.18	87.44	75.98	88.05	79.20	88.93
	本文方法	73.88	86.58	78.31	88.58	79.38	89.43	80.42	89.93
VGG16	Random	71.72	85.11	75.42	87.11	78.89	88.84	80.06	89.77
	ImageNet	77.98	88.39	79.40	89.16	80.19	90.09	80.91	90.63
	MoCo	73.09	86.42	76.51	88.57	79.70	89.33	79.97	89.24
	本文方法	79.53	89.26	79.86	89.49	80.34	90.21	81.41	90.96

注：加粗字体表示最优结果。

当训练数据比例为 1% 时，本文方法的 IoU 和 F₁ 分数在 ResNet18 为基础结构的情况下，较随机初始化的方式分别提升了 1.32% 和 0.88%，较 ImageNet 预训练的方式则提升了 0.62% 和 0.54%；在 VGG16 为基础结构的情况下，较随机初始化的方式分别提升了 7.81% 和 4.15%，提升效果十分显著，较 ImageNet 预训练的方式也有 1.55% 和 0.87% 的提升。当训练数据比例扩大至 2%—10% 时，本文方法在 ResNet18 上较其他方法的 IoU 和 F₁ 分别提升了 2.46%—3.02% 和 1.18%—1.42%，在 VGG16 上则分别提升了 1.44%—4.44% 和 1.19%—2.38%。上述实验结果显示，本文提出的自监督预训练方法在下游地物阴影检测任务中，相比于随机初始化、ImageNet 预训练和 MoCo 方法所得到的 IoU 和 F₁ 结果都有较为明显的提升，并且本文方法在所有情况下都取得了最高的地物阴影检测精度，达到了设计的预期效果。

如表 4 所示，本文实验中 MoCo 方法性能不足的主要原因是，本文采用的训练数据量无法满足 MoCo 充分学习特征表示的需求。这也证明了本文方法在地物阴影检测任务中的优势，由于引入了方向先验，相比于通用自监督特征学习方法，如 MoCo 对于数据量的需求较为宽松。因此，本文方法为地物阴影检测任务提供了一种有效的解决方案，特别是在面对数据规模和计算资源有限的情况下。

图 13 和图 14 为各类方法在不同比例训练数据下 IoU 和 F₁ 的精度变化曲线。可以看到，本文方法在训练数据极少的情况下具有较大优势，随着训练数据的增多，各个模型的性能逐渐接近，但本

文方法依旧保持领先水平。这是由于本文所提出的自监督特征学习辅助模型是专门针对遥感影像地物阴影检测任务设计的，充分考虑了地物阴影的方向先验特征，构建了针对地物阴影的方向性噪声处理机制和数据增强策略，有效提高了模型对于地物阴影的识别准确率。

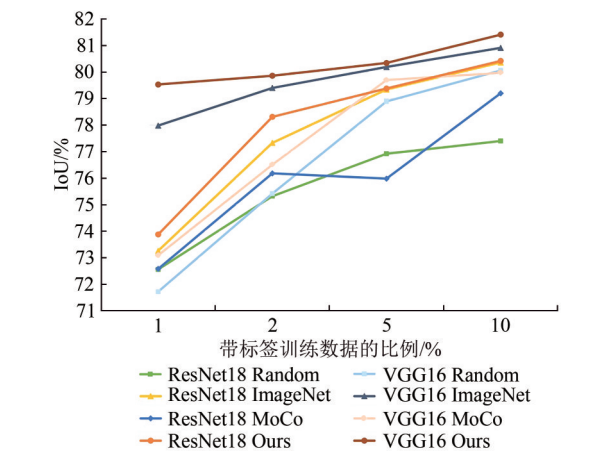


图 13 各类方法在不同比例训练数据下的 IoU 变化对比

Fig.13 IoU of various methods under different proportions of training data

除此之外，在训练数据规模有限的情况下，几乎所有的方法都能取得较高的地物阴影检测准确率。这是因为 AISD 数据集覆盖的主要为城市地区，建筑物所占比例较大，大多数建筑物所产生的阴影具有相似的纹理信息，即使存在少量建筑物的阴影细节区域或是出现一些复杂的树影很难区分，但是这些识别难度大的地物阴影所占区域较小，因此使用少量的标签数据依旧可以取得相

对较高的地物阴影总体识别精度。

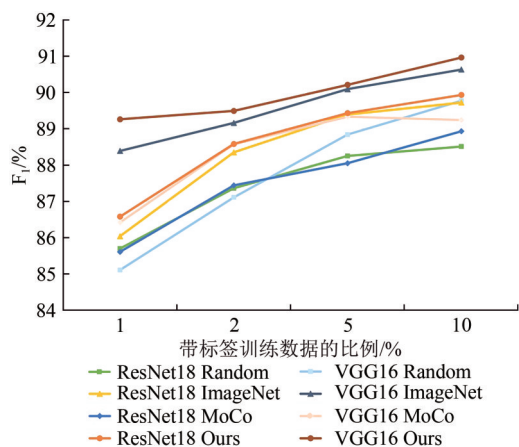


图 14 各类方法在不同比例训练数据下的 F_1 变化对比

Fig.14 F_1 -score of various methods under different proportions of training data

3.6.3 地物阴影识别结果的可视化分析

可视化图像可以清晰直观地对预测结果进行定性评价, 本文以 ResNet18 为编码器基础结构、带标签训练数据比率为 10% 的情况为例, 在 AISD 测试集中选择了 4 张具有代表性的遥感影像, 其中包含了丰富的地物及阴影类型, 通过对比不同方法所提取的地物阴影细节, 能够全面证明本文方法的有效性先进性。图 15 为地物阴影检测结果的可视化图像, 黄色矩形框区域代表不同方法对于该部分的预测结果细节差别较大。其中, 第 1 列图像为原始遥感影像, 第 2 列为真值标签, 第 3 列至第 5 列为其他对比方法的识别结果, 第 6 列为本文方法的识别结果, 可以看到本文方法所提取地物阴影的完整程度和精细程度最高。

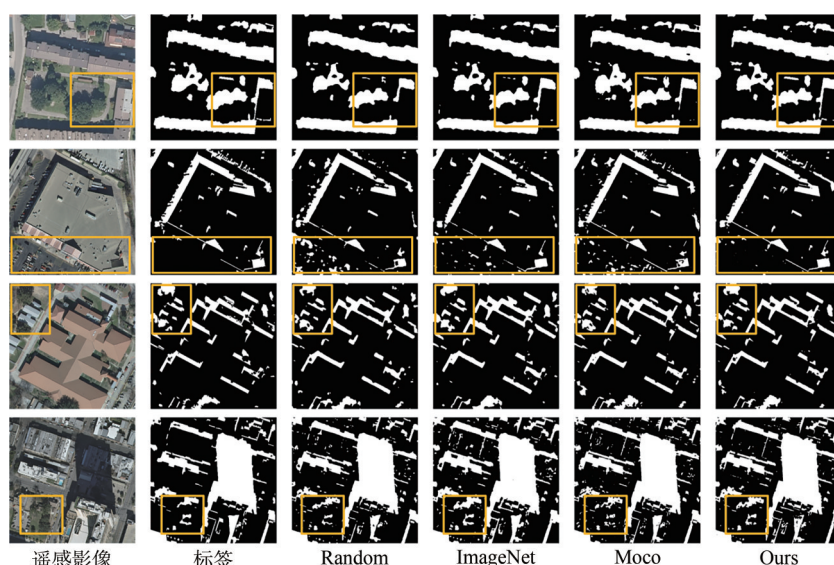


图 15 不同方法的地物阴影检测结果可视化($N_{10\%}$)

Fig.15 Shadow detection results for different methods ($N_{10\%}$)

如图 15 所示, 在第一行的图像中, 随机初始化的方式难以识别细碎的树影和建筑物边缘的阴影, MoCo 和 ImageNet 的方法能够检测到这些复杂区域的阴影, 但识别结果的细节相对较差。相比于其他方法, 本文方法能够较好地还原地物阴影的真实形态, 得到平滑和规则的阴影轮廓。在图 15 中第二行的遥感图像中, 存在许多紧密停靠的汽车阴影, 由于这些阴影非常微小, 上述 4 种方法均未能有效识别这些小尺度连续阴影, 从而导致了一些连续阴影区域的误检和漏检。这个问题一方面源于数据集本身对于这些微小阴影的忽视, 另一方面, 检测这类微小阴影本身就具有相当大

的难度, 需要设计相应的策略来加以处理。值得注意的是, 本文方法采用了对比剪裁技术, 这使得在自监督学习时难以剪裁出仅包含上述小尺度地物及其阴影的区域, 从而导致忽略了位于图像角落的微小阴影。这些因素都会对本文方法的性能产生一定的影响, 这也是本文方法在处理小尺度连续阴影时的局限性。

对于图 15 中第 3 行和第 4 行图像中的深绿色树木, 由于树木阴影的颜色与树木本身的颜色非常相似, 因此容易导致错误的检测结果。相对于其他方法的识别结果, 本文方法在辨别图像中的深色区域和阴影区域方面表现更为出色, 输出了相

对清晰的地物阴影边界。这一优势的背后在于本文采用了随机色彩变换的数据增强方法, 学习出的自监督特征能够有效协助模型学习地物阴影明暗程度变化所带来的差异。

上述实验表明, 基于自监督学习方法预训练的地物阴影检测模型不依赖于人工标签, 可以提供比传统的监督学习更加抽象和丰富的监督信息, 更好地保留高分辨率遥感影像自身蕴含的特征, 从而在下游任务中取得更好的性能表现。

4 结 论

地物阴影是高分辨率遥感影像中普遍存在的典型特征之一, 本文提出了一种基于阴影方向先验的高分辨率遥感影像地物阴影检测方法, 其贡献在于探究了阴影方向先验对于遥感影像地物阴影高级语义特征表达的有效性; 依据地物阴影的方向先验特征构造伪标签, 设计 DTICT 方向性噪声去除策略以减轻无阴影和特殊方向阴影样本对网络训练的干扰; 引入融合了对比剪裁和色彩变换的数据增强策略以帮助网络获取更丰富的训练样本, 提高预训练模型对地物阴影特征的提取能力。本文方法在 AISD 数据集上进行了实验, 结果表明本文方法相较于对比方法具有一定的优越性, 仅使用少量的样本标签训练模型即可取得较好的地物阴影检测结果。

由于数据量和计算能力的限制, 我们无法进行大规模数据的实验, 判别模型是否能够在少量标注样本下达到监督学习的性能。未来的研究将继续深入探索优化的方法, 旨在进一步提高自监督学习方法的性能。此外, 本文采用的是图像级别的阴影方向性先验知识, 难免限制其在像素级的阴影检测任务中的性能, 下一步工作将探究像素级的阴影方向性先验知识, 进一步提升阴影检测的性能。

参考文献(References)

- Belkin M, Hsu D, Ma S Y and Mandal S. 2019. Reconciling modern machine-learning practice and the classical bias-variance trade-off. *Proceedings of the National Academy of Sciences of the United States of America*, 116(32): 15849-15854 [DOI: 10.1073/pnas.1903070116]
- Duan G Y, Gong H L, Li X J and Chen B B. 2014. Shadow extraction based on characteristic components and object-oriented method for high-resolution images. *Journal of Remote Sensing (in Chinese)*, 18(4): 760-770 (段光耀, 宫辉力, 李小娟, 陈蓓蓓. 2014. 结合特征分量构建和面向对象方法提取高分辨率卫星影像阴影. *遥感学报*, 18(4): 760-770) [DOI: 10.11834/jrs.20143243]
- Han B, Yao Q M, Yu X R, Niu G, Xu M, Hu W H, Tsang I W and Sugiyama M. 2018. Co-teaching: robust training of deep neural networks with extremely noisy labels//*Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Montréal: Curran Associates Inc.: 8536-8546
- He K M, Fan H Q, Wu Y X, Xie S N and Girshick R. 2020. Momentum contrast for unsupervised visual representation learning//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE: 9726-9735 [DOI: 10.1109/CVPR42600.2020.00975]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hou L, Vicente T F Y, Hoai M and Samaras D. 2021. Large scale shadow annotation and detection using lazy annotation and stacked CNNs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(4): 1337-1351 [DOI: 10.1109/TPAMI.2019.2948011]
- Ji S P, Wei S Q and Lu M. 2019. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1): 574-586 [DOI: 10.1109/TGRS.2018.2858817]
- Jing L L and Tian Y L. 2021. Self-supervised visual feature learning with deep neural networks: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(11): 4037-4058 [DOI: 10.1109/TPAMI.2020.2992393]
- Jung H, Oh Y, Jeong S, Lee C and Jeon T. 2022. Contrastive self-supervised learning with smoothed representation for remote sensing. *IEEE Geoscience and Remote Sensing Letters*, 19: 8010105 [DOI: 10.1109/LGRS.2021.3069799]
- Khan S H, Bennamoun M, Sohel F and Togneri R. 2014. Automatic feature learning for robust shadow detection//*Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE: 1939-1946 [DOI: 10.1109/CVPR.2014.249]
- Li D R, Tong Q X, Li R X, Gong J Y and Zhang L P. 2012. Current issues in high-resolution Earth observation technology. *Science China Earth Science*, 55(7): 1043-1051 (李德仁, 童庆禧, 李荣兴, 龚健雅, 张良培. 2012. 高分辨率对地观测的若干前沿科学问题. *中国科学: 地球科学*, 42(6): 805-813) [DOI: 10.1007/s11430-012-4445-9]
- Li J Y, Hu Q W and Ai M Y. 2016. Joint model and observation cues for single-image shadow detection. *Remote Sensing*, 8(6): 484 [DOI: 10.3390/rs8060484]
- Li Y, Gong P. 2005. Integrating photogrammetry and image analysis for shadow detection. *Journal of Remote Sensing (in Chinese)*, 9(4): 357-362 (李艳, 宫鹏. 2005. 基于 DSM 阴影仿真和高度场光线跟踪的影像阴影检测. *遥感学报*, 9(4): 357-362) [DOI: 10.11834/jrs.20050452]

- Liao C J, Zhao J, Xing J, Wen J G and Liu Q H. 2023. Research on classification of GF satellite data application products and construction of common product system. *National Remote Sensing Bulletin*, 27(3): 563-572 (廖楚江, 赵坚, 邢进, 闻建光, 柳钦火. 2023. 高分卫星数据应用产品分级与共性产品体系构建研究. *遥感学报*, 27(3): 563-572) [DOI: 10.11834/jrs.20232593]
- Lin Z Q, Sun J, Davis A and Snavely N. 2020. Visual chirality//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 12292-12300 [DOI: 10.1109/CVPR42600.2020.01231]
- Liu D Y, Zhang J P, Wu Y H and Zhang Y. 2021a. A shadow detection algorithm based on multiscale spatial attention mechanism for aerial remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 19: 6003905 [DOI: 10.1109/LGRS.2021.3100294]
- Liu X, Zhang F J, Hou Z Y, Mian L, Wang Z Y, Zhang J and Tang J. 2021b. Self-supervised learning: generative or contrastive. *IEEE Transactions on Knowledge and Data Engineering*, 35(1): 857-876 [DOI: 10.1109/TKDE.2021.3090866]
- Luo S, Li H F and Shen H F. 2020. Deeply supervised convolutional neural network for shadow detection based on a novel aerial shadow imagery dataset. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167: 443-457 [DOI: 10.1016/j.isprsjprs.2020.07.016]
- Luo S, Li H F, Zhu R Z, Gong Y T and Shen H F. 2021. ESPFNet: an edge-aware spatial pyramid fusion network for salient shadow detection in aerial remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 4633-4646 [DOI: 10.1109/JSTARS.2021.3066791]
- Ma C, Feng D J and Zhang L. 2012. Detecting the building shadow with the two algorithms of threshold segmentation and mathematical morphology. *Surveying and Mapping*, 35(4): 151-154 (马川, 冯德俊, 张丽. 2012. 结合阈值分割和数学形态学的建筑物阴影检测. *测绘*, 35(4): 151-154) [DOI: 10.3969/j.issn.1674-5019.2012.04.004]
- Makarau A, Richter R, Muller R and Reinartz P. 2011. Adaptive shadow detection using a blackbody radiator model. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6): 2049-2059 [DOI: 10.1109/TGRS.2010.2096515]
- Mu X D, Bai K, You X, Zhu Y Q and Chen X B. 2021. Remote sensing image feature extraction and classification based on contrastive learning method. *Optics and Precision Engineering*, 29(9): 2222-2234 (慕晓冬, 白坤, 尤轩昂, 朱永清, 陈雪冰. 2021. 基于对比学习方法的遥感影像特征提取与分类. *光学精密工程*, 29(9): 2222-2234) [DOI: 10.37188/OPE.20212909.2222]
- Peng X Y, Wang K, Zhu Z, Wang M and You Y. 2022. Crafting better contrastive views for Siamese representation learning//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE: 16010-16019 [DOI: 10.1109/CVPR52688.2022.01556]
- Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D and Batra D. 2017. Grad-CAM: visual explanations from deep networks via gradient-based localization//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice: IEEE: 618-626 [DOI: 10.1109/ICCV.2017.74]
- Shahtahmassebi A, Yang N, Wang K, Moore N and Shen Z Q. 2013. Review of shadow detection and de-shadowing methods in remote sensing. *Chinese Geographical Science*, 23(4): 403-420 [DOI: 10.1007/s11769-013-0613-x]
- Shen L, Wee Chua T and Leman K. 2015. Shadow optimization from structured deep edge detection//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE: 2067-2074 [DOI: 10.1109/CVPR.2015.7298818]
- Shi L, Fang J and Zhao Y F. 2023. Automatic shadow detection in high-resolution multispectral remote sensing images. *Computers and Electrical Engineering*, 105: 108557 [DOI: 10.1016/j.compelceng.2022.108557]
- Stojnic V and Risojevic V. 2021. Self-supervised learning of remote sensing scene representations using contrastive multiview coding//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Nashville: IEEE: 1182-1191 [DOI: 10.1109/CVPRW53098.2021.00129]
- Su N, Zhang Y, Tian S, Yan Y M and Miao X Y. 2016. Shadow detection and removal for occluded object information recovery in urban high-resolution panchromatic satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(6): 2568-2582 [DOI: 10.1109/JSTARS.2016.2570234]
- Tao C, Qi J, Lu W P, Wang H and Li H F. 2022. Remote sensing image scene classification with self-supervised paradigm under limited labeled samples. *IEEE Geoscience and Remote Sensing Letters*, 19: 8004005 [DOI: 10.1109/LGRS.2020.3038420]
- Tao C, Yin Z W, Zhu Q and Li H F. 2021. Remote sensing image intelligent interpretation: from supervised learning to self-supervised learning. *Acta Geodaetica et Cartographica Sinica*, 50(8): 1122-1134 (陶超, 阴紫薇, 朱庆, 李海峰. 2021. 遥感影像智能解译: 从监督学习到自监督学习. *测绘学报*, 50(8): 1122-1134) [DOI: 10.11947/j.AGCS.2021.20210089]
- Tong X D. 2016. Development of China high-resolution earth observation system. *Journal of Remote Sensing (in Chinese)*, 20(5): 775-780 (童旭东. 2016. 中国高分辨率对地观测系统重大专项建设进展. *遥感学报*, 20(5): 775-780) [DOI: 10.11834/jrs.20166302]
- Wang Q J, Yan L, Yuan Q Q and Ma Z L. 2017. An automatic shadow detection method for VHR remote sensing orthoimagery. *Remote Sensing*, 9(5): 469 [DOI: 10.3390/rs9050469]
- Xie Y K, Feng D J, Chen H Y, Liao Z Y, Zhu J, Li C N and Baik S W. 2022. An omni-scale global-local aware network for shadow extraction in remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 193: 29-44 [DOI: 10.1016/j.isprsjprs.2022.09.004]
- Yang Y, Guo M Q and Zhu Q Q. 2021. CADNet: top-down contextual saliency detection network for high spatial resolution remote sensing image shadow detection//Proceedings of 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Brussels: IEEE: 4075-4078 [DOI: 10.1109/IGARSS47720.2021.9553138]
- Yu D F, Yin J P and Zhang G M. 2008. An automatic shadow detection

- method for remote sensing images based on gray histogram. Computer Engineering and Science, 30(12): 43-44, 93 (于东方, 殷建平, 张国敏. 2008. 一种基于灰度直方图的遥感影像阴影自动检测方法. 计算机工程与科学, 30(12): 43-44, 93) [DOI: 10.3969/j.issn.1007-130X.2008.12.011]
- Zhou P C, Cheng G, Yao X W and Han J W. 2021. Machine learning paradigms in high-resolution remote sensing image interpretation. National Remote Sensing Bulletin, 25(1): 182-197 (周培诚, 程杰, 姚西文, 韩军伟. 2021. 高分辨率遥感影像解译中的机器学习范式. 遥感学报, 25(1): 182-197) [DOI: 10.11834/jrs.20210164]
- Zhu Q Q, Yang Y, Sun X L and Guo M Q. 2022. CDANet: contextual detail-aware network for high-spatial-resolution remote-sensing imagery shadow detection. IEEE Transactions on Geoscience and Remote Sensing, 60: 5617415 [DOI: 10.1109/TGRS.2022.3143886]

Shadow detection method for high-resolution remote sensing images on the basis of shadow direction prior

QI Kunlun¹, MA Xinyue², JIN Zhun¹, QING Yaxian³, LI Zhenqiang⁴, YANG Chao⁴, WU Huayi³

1.School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China;

2.China Mobile (Hangzhou) Information Technology Co., Ltd., Hangzhou 311100, China;

3.National Engineering Research Center of Geographic Information System, Wuhan 430074, China;

4.State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Abstract: Given the high cost and difficulty of shadow annotation, the shadow detection performance of supervised learning models in high-resolution remote sensing images is severely limited mainly because of the lack of training samples. To solve this problem, this study proposes a shadow detection method that is based on shadow direction prior, which has unique advantages in learning the inherent shadow information contained in unlabeled high-resolution remote sensing images. Through self-supervised learning methods, the advanced semantic information of ground object shadows contained in remote sensing images can be deeply explored, thereby improving the applicability and accuracy of the model. This innovative approach provides a new way to overcome the data limitation problem in shadow detection and offers new possibilities for improving the model's ability to understand complex high-resolution images.

This study proposes an innovative shadow detection method for high-resolution remote sensing images that is based on prior knowledge of the shadow direction and explores the effectiveness of shadow direction prior in representing high-level semantic features of object shadows in remote sensing images. The key task in this method is to construct an auxiliary task through shadow direction prior so that the deep neural network can effectively understand the shadow of objects in remote sensing images. Moreover, to further enhance the deep neural network's ability to learn the key features of object shadows, this study designs a direction transformation-independent noise processing mechanism and data enhancement strategy for self-supervised shadow detection, both of which improve the generalization ability of the model.

The effectiveness of the proposed method is proven through verification on the Aerial Imagery Dataset for Shadow Detection (AISD). Experimental results show that compared with the traditional Unet benchmark model, the proposed method can successfully learn prior knowledge in unlabeled high-resolution remote sensing images. By leveraging this prior knowledge, the developed method improves the accuracy of the Unet model and demonstrates competitive efficiency and performance in small sample cases. The results of this study highlight the superiority of the proposed method in addressing the challenge of shadow identification in high-resolution remote sensing images, especially in the absence of labeled samples. By fully utilizing prior knowledge, the method successfully extends the performance of the Unet model and provides a promising solution for processing unlabeled high-resolution remote sensing images in practical applications.

The qualitative and quantitative results consistently show that using a shadow direction prior method can considerably improve the accuracy of shadow detection. The use of the DTICT strategy helps alleviate the interference caused by shadow-free samples and samples containing shadows in special directions during network training. The data enhancement strategy combined with contrast cropping and color transformation provides the network with diverse training samples, thereby enhancing the pre-trained model's ability to extract object shadow features. The proposed method achieves competitive object shadow detection accuracy by using a small number of labeled samples.

Key words: shadow detection, self-supervised learning, data augmentation, shadow direction prior, high-resolution remote sensing imagery

Supported by Special Fund of Hubei Luojia Laboratory (No.220100034); Hubei Key Research and Development Program in China (No. 2020AAA004)