

DOI:10.19651/j.cnki.emt.2517778

基于改进 UNet3+ 的露天矿无人机影像阴影提取<sup>\*</sup>

杜孙稳 宋瑞婷 高志宇 史 森 张浩然

(太原理工大学地球科学与测绘工程学院 太原 030024)

**摘要:** 利用无人机技术获取露天矿区地表正射影像图可以快速有效的实现对矿区地表形态的全面监测与分析。然而,阴影在露天矿无人机正射影像图中普遍存在,既干扰了部分地物信息的获取、还降低了无人机影像的解译和识别精度。针对目前矿区阴影提取的研究较少,现有方法不能满足露天矿阴影识别的需求,同时也未建立露天矿区阴影数据集等问题,首次采用人工标注的方法构建一组露天矿区无人机正射影像阴影数据集。模型基于 UNet3+,提出了一种结合混合注意力机制(CBAM)与深度可分离卷积层(DSC)的阴影提取方法。通过引入 ResNet 特征提取器,对原始影像进行五个尺度上的特征提取,根据提取到的特征进行全尺度跳跃连接实现特征融合,并引入 CBAM 注意力机制,以增强有用的特征,最终通过深度监督机制及解码器恢复的特征图预测每个像素的类别。对比 4 种典型的目标提取网络 FCN、UNet、UNet++、UNet3+。结果表明,所提方法与 Unet3+ 网络相比,在 mPrecision、mRecall、mF1 和 mIoU 四个指标上分别提升了 4.9%、0.44%、2.24%、4.51%。在 AISD 公开数据上对比多种已有的阴影提取方法,结果表明,所提方法与残差监督网络相比,F1 与 IoU 提高了 0.27%、2.62%,证明该方法在阴影提取方面具有准确性,适合应用于露天矿区无人机影像中阴影提取。

**关键词:** 露天矿区阴影数据集;混合注意力机制;UNet3+;深度可分离卷积;跳跃连接

**中图分类号:** TN919.8;TP751 **文献标识码:** A **国家标准学科分类代码:** 420.2030

Shadow extraction from UAV imagery of open-pit mines  
based on improved UNet3+

Du Sunwen Song Ruiting Gao Zhiyu Shi Miao Zhang Haoran

(College of Geological and Surveying Engineering, Taiyuan University of Technology, Taiyuan 030024, China)

**Abstract:** Acquiring surface orthophoto image by UAV technology can quickly and effectively realize the comprehensive monitoring and analysis of mining surface morphology. However, shadows are prevalent in the open-pit mine UAV orthophoto image, which not only interferes with the acquisition of some ground object information, but also reduces the interpretation and recognition accuracy of UAV images. There are few researches on shadow extraction in mining area at present. Existing methods can not meet the needs of shadow identification in open pit mines. And the problems of not establishing shadow data set in opencast mining area, a UAV ortho image shadow dataset is constructed by using manual annotation for the first time. Based on the UNet3+ model, a shadow extraction method combining mixed attention mechanism (CBAM) and depth separable convolution layer (DSC) is proposed. By introducing the ResNet feature extractor, feature extraction on five scales is carried out on the original image, and performs full-scale jump connection according to the extracted features to carry out feature fusion. And introduce CBAM attention mechanisms to enhance useful features. The category of each pixel is predicted by the feature map recovered by the deep monitoring mechanism and the decoder. Finally, proposed method is compared with four typical target extraction networks FCN, UNet, UNet++ and UNet3. The experimental results show that, compared with Unet3+ network, the mPrecision, mRecall, mF1 and mIoU improved by 4.9%, 0.44%, 2.24% and 4.51%, respectively. Proposed method was compared with a variety of existing shadow extraction methods on AISD public data. The experimental results showed that compared with residual supervision network, F1 and IoU improved by 0.27% and 2.62%. It is proved that this method is accurate in shadow extraction, and is suitable for shadow extraction in open pit mining area.

**Keywords:** open-pit mine shadow dataset; CBAM; UNet3+; depthwise separable convolution; skip connection

## 0 引 言

阴影是自然界中一种常见的光学现象,在现实世界中无处不在。阴影的变化可以反映植被覆盖、土壤湿度等环

境因素的变化<sup>[1]</sup>,在计算机视觉中,阴影的存在会对算法产生不利影响,阻碍视觉跟踪、语义分割等任务<sup>[2]</sup>。因此,阴影提取对后续相关的数据分析任务至关重要。

露天矿区多为阶梯状,且路面多有碎石,地表起伏情况

收稿日期:2025-01-02

<sup>\*</sup> 基金项目:国家自然科学基金(U21A20107)、山西省基础研究计划面上项目(202203021211156)资助

较为严重<sup>[3-4]</sup>。由于地表起伏、太阳方位角、拍摄角度的问题,光被物体全部或部分遮挡<sup>[5-7]</sup>,使得阴影在露天矿无人机影像中普遍存在。智慧矿山建设中,无人机获取高分辨率影像为矿山管理提供了丰富的数据支持,随着无人机影像分辨率不断提高,阴影的影响越来越明显,阴影导致被遮挡区域的信息减少,使得影像解读更加困难,智慧矿山建设的难度加大。

为提高影像质量与解译精度,促进露天矿地区高精度地表信息的获取,阴影提取是影像分析必不可少的预处理步骤。过去几十年中,阴影提取的研究多使用高分数据,范围广、面积大。提取方法可分为两类,一类是基于传统特征的方法<sup>[8-10]</sup>,主要包括:利用光谱特性,利用阴影在可见光波段(如红、绿、蓝)与近红外波段的反射率差异进行提取<sup>[11]</sup>;利用几何特性,基于阴影的几何形状和位置关系进行提取,阴影通常位于光源的相反方向,且形状与产生阴影的物体相似;利用纹理特性,利用阴影区域与周围环境的纹理差异进行提取<sup>[12]</sup>;结合多特征构造阴影指数进行提取<sup>[13]</sup>。然而,基于传统特征的方法在复杂场景下效果有限。另一类是基于深度学习的方法,随着深度学习的飞速发展,提出了众多的影像分割提取算法<sup>[14-23]</sup>,越来越多的研究者开始将深度学习应用于阴影提取领域。深度学习模型能够通过卷积获取影像中的多尺度特征,并通过大量数据进行训练,对模型中的权重进行优化,从而提高阴影提取的准确性和效率。基于 FCN、U-Net、U-Net++ 和 U-Net3+ 等经典影像语义分割模型及其改进模型被广泛应用,最新的无人机阴影提取方法包括 Luo 等<sup>[22]</sup>提出了 DSSDNet,使用了编码-解码网络,Zhu 等<sup>[24]</sup>在 UNet 网络中引入全局空间上下文注意力模块,Chen 等<sup>[23]</sup>采用 MTMT 网络,Zhang 等<sup>[15]</sup>提出了 MRPF-net 网络以提升阴影检测网络在复杂场景中的普适性,2023 年 Xie 等<sup>[25]</sup>提出了一种取水网络 SE-CU-Net 模型,以克服建筑物和其他特征的阴影效应,Yang 等<sup>[26]</sup>在 2023 年提出了一种 Rennet-Unet 模型来解决提取建筑物顶部信息时的阴影问题。这些模型通过原始模型、引入注意力机制<sup>[27]</sup>、使用轻量级网络等先进技术,进一步提高了目标提取的性能。目前,还未有人将 U-Net3+ 应用于阴影提取中,U-Net3+ 在医疗影像分割领域得到了广泛的应用,尤其是在需要高精度和细粒度分割的任务中表现突出,因此,本文将 U-Net3+ 应用于露天矿区无人机影像阴影提取中,无人机影像分辨率高,场景复杂,露天矿区地面起伏情况严重,而且地表常覆盖极易与阴影混淆的重要非阴影地物,如煤,使用 U-Net3+ 网络在无人机影像阴影提取中具有较大可行性。

过往阴影提取的研究大部分集中于城市地区<sup>[22]</sup>,针对矿区这一特殊地区研究很少,而基于阴影属性的研究在矿区适用性、准确性较差。针对这些问题,本研究基于改进 UNet3+ 网络进行露天矿区阴影提取,在编码器-解码器同尺度连接时,使用深度可分离卷积(depthwise

separable convolution,DSC)代替 U-Net3+ 中传统的二维卷积层(2D convolutional layer, Conv2D)降低网络参数量(parameter,P),提高网络训练速度,并在解码器特征融合层后嵌入混合注意力机制模块(convolutional block attention module,CBAM),抑制背景干扰,提高前景权重。最后,通过与典型语义分割方法对比、消融实验、与最新无人机阴影提取方法对比,本文方法在平均精确率(mean precision,mPrecision)、平均召回率(mean recall,mRecall)、平均 F1(mean F1,mF1)和平均 IoU(mean IoU,mIoU)值方面都得到有效提升,为露天矿区阴影提取提供了手段和方法。

## 1 基于改进的 UNet3+ 网络的阴影提取方法

### 1.1 改进 UNet3+ 网络

露天矿区阴影具有细碎、多变,且地表覆盖极易与阴影混淆的重要非阴影地物。U-Net3+ 网络是一种在影像分割具有优异性能的被广泛应用的深度学习网络架构,通过充分利用多尺度特征和深度监督机制及分类引导模块,能够更准确地分割阴影。针对含煤地带,本文在 U-Net3+ 的特征融合部分后加入 CBAM 模块,增强模型对阴影特征的提取能力,并在解码器融合同尺度特征图时使用 DSC 卷积层替换原始 Conv 卷积层,同尺度特征图特征提取已接近饱和,替换后,精度损失较少,且极大的减少参数量,从而加快模型训练速度。网络结构如图 1 所示。

UNet3+ 网络编码器通过池化下采样逐步降低输入影像的分辨率,获取多种尺度及深度的影像,同时通过卷积提取和抽象影像的高级阴影语义特征信息。解码器部分则通过上采样操作,将编码器部分提取的特征图恢复到原始尺寸,网络结构如图 1 所示。UNet3+ 在此基础上增加了全尺度跳跃连接、深度监督及分类引导模块。

#### 1) 全尺度的跳跃连接

U-Net3+ 网络中的全尺度跳跃连接是一个关键特性。旨在通过融合多尺度特征来提高图像分割的精度。

UNet3+ 编码器中,输入原图,通过两个  $3 \times 3$  的 Conv2D、批量归一化层(batch normalization layer, BN)和激活函数层(rectified linear unit, ReLU)提取特征,获得与原图同尺寸的特征图,并通过下采样,最大池化获得 1/2 原图尺寸的特征图,并再次卷积、池化,经过 4 次最大池化,获得 5 个尺度的特征图。

以  $X_{d_e}^3$  (解码器中的第 3 层)为例,首先,其融合了编码器中的小尺度  $X_{E_n}^1$ 、 $X_{E_n}^2$  和同尺度的  $X_{E_n}^3$  特征图,以及解码器的  $X_{d_e}^1$ 、 $X_{d_e}^5$  的大尺度特征图,将小尺度的特征图进行最大池化,将同尺度的特征图使用 DSC 卷积网络无重叠最大池化,将大尺度的特征图进行双线性插值,使得各层特征图尺寸相同,然后进行串联融合,实现全尺度的跳跃连接,从而所有尺度下语义全部捕获,提高分割的精度。最后,在此特征融合的基础上,增加 CBAM 模块,增强网络对

阴影特征的关注度而抑制对其他地物特征的提取,结构如图 2 所示。

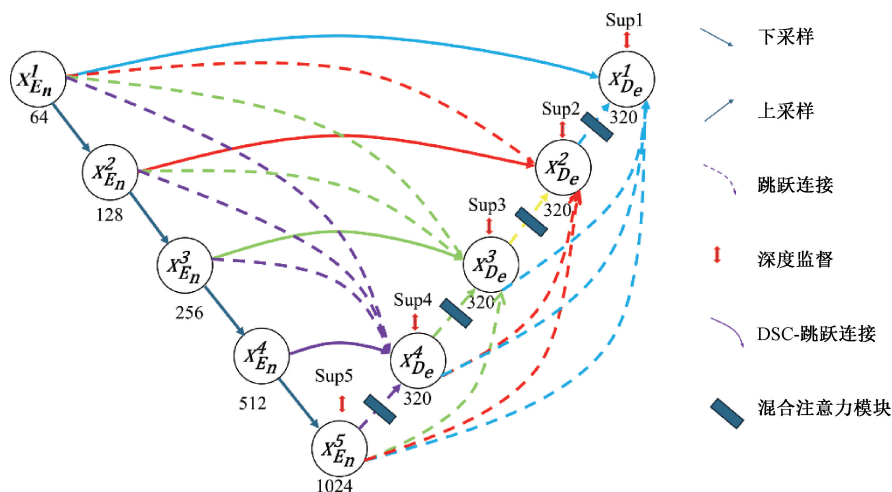


图 1 改进 UNet3+ 网络结构

Fig. 1 Improved UNet3+ network structure

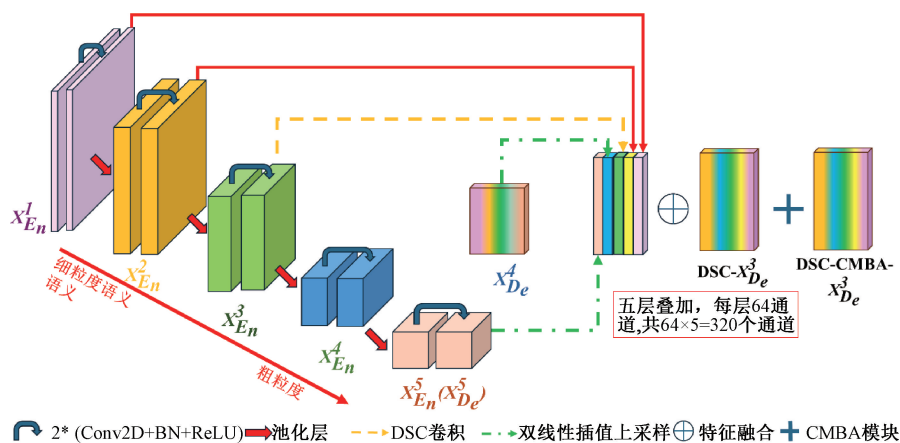


图 2 DSC-CMBA- \$X\_{De}^3\$ 层结构图

Fig. 2 The feature map of DSC-CMBA- \$X\_{De}^3\$

## 2) 深度监督

UNet3+ 采用了深度监督方法,首先,在每个解码器阶段产生一个侧输出,并使用双线性插值将标签图上采样到对应解码器输出大小,使用损失函数计算侧输出与标签图真值之间的差异,侧输出都由标签图中的真值进行监督,如图 2 中 Sup 部分。这有助于网络在训练过程中更好地学习特征表示。深度监督能够增强网络的泛化能力,并在一定程度上缓解过拟合问题。

## 3) 分类引导模块

为了减少阴影的误分,UNet3+ 引入了一个分类引导模块。该模块通过优化混合损失函数预测输入影像是否包含阴影来指导分割过程,从而提高分割的准确性。如图 3 所示,从拥有最丰富语义信息的编码层 \$x\_{En}^5\$ 中经过一系列的正则化、\$1 \times 1\$ 卷积、自适应最大池化、Sigmoid 操作

得到一个二维张量,通过预测类别结果,将二维张量转化为 0 或 1 的单一输出。其中 0 代表影像中无阴影,1 代表影像中有阴影,并将此分类结果与深度监督阶段内双线性上采样操作后的分割结果相乘,如果分类结果显示影像中无阴影,则相应地减少分割结果中原本可能被误判为阴影的区域。

## 1.2 改进的 ResNet 特征提取器

本文采用 ResNet50 作为深度特征提取的基础。在编码阶段,ResNet50 的每个卷积层逐渐提取图像的低级和高级特征,包括 \$7 \times 7\$ 卷积层、最大池化层、4 个残差层 layer1、layer2、layer3、layer4,并删除平均池化层和全连接层以适应特征提取任务,如图 4 所示,删除了 stage5 部分,设置了 strict=False 来允许权重加载时忽略因删除导致不匹配的部分,从而生成 5 个不同级别的特征图。

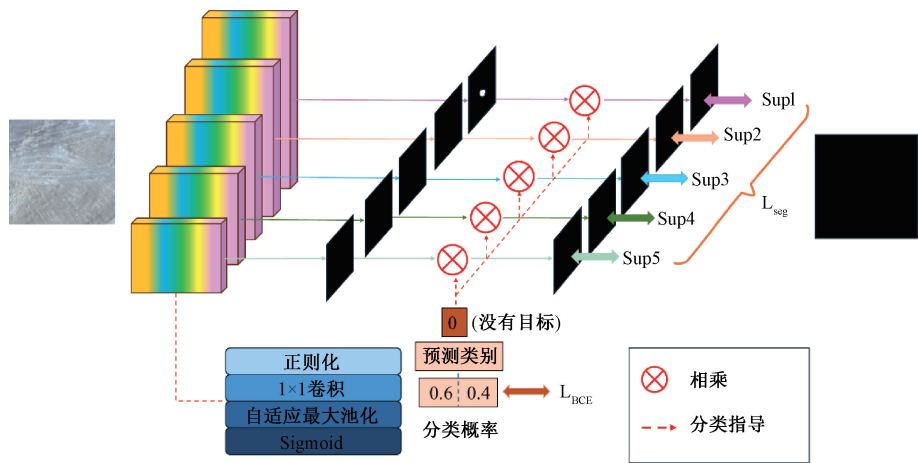


图 3 分类引导模块结构  
Fig. 3 Classification-guided module

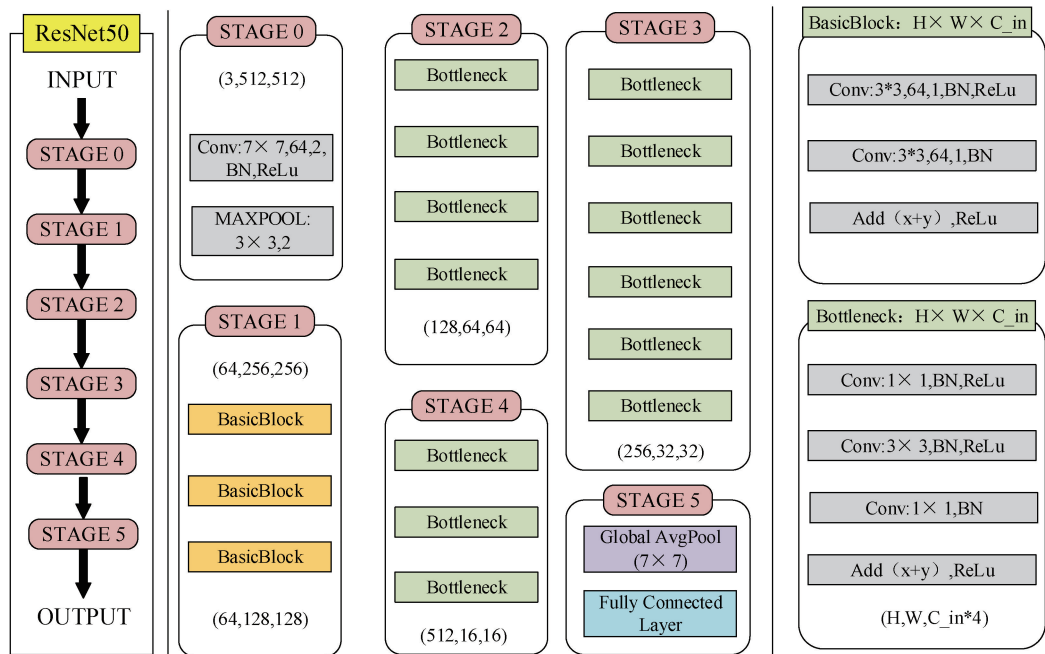


图 4 ResNet50 网络结构  
Fig. 4 ResNet50 network structure

1.3 深度可分离卷积

UNet3+网络中使用两个  $3 \times 3$  的 Conv2D、BN、ReLU 进行特征提取,这也导致网络复杂度与参数量较大。在 UNet3+网络中,在解码器融合同尺度特征图时使用深度可分离卷积层替换原始 Conv2D, DSC 是由深度卷积 (depthwise convolution, DWConv) 和逐点卷积 (pointwise convolution, PWConv) 两部分组成。两种卷积对比如图 5 所示。

DWConv 在每个通道内进行卷积操作,减少了多通道重复卷积的运算量。而 PWConv 则将 DWConv 的输出在通道间进行组合从而升维,得到与 Conv2D 相同尺寸的特

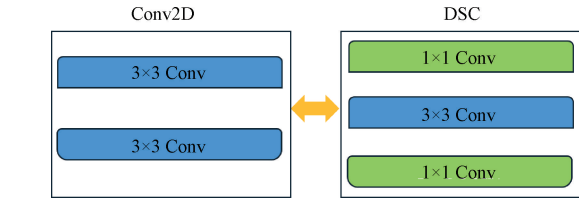


图 5 Conv2D 与 DSC 对比图  
Fig. 5 Comparison illustration of Conv2D vs. DSC

征图,可以极大减少模型的参数量和计算量。具体计算如下:

1) Conv2D 的参数量  $P_{Conv2D}$



对于标准的 Conv2D, 输入特征图为  $F \in R^{C_0 \times H_0 \times W_0}$  (其中  $C_0$  是通道数,  $H_0$  和  $W_0$  分别是特征图的高度和宽度), 输出特征图为  $F' \in R^{C_1 \times H_1 \times W_1}$ , 本文模型卷积核为  $f^{3 \times 3}$ , 则  $P_{Conv2D}$  为:

$$P_{Conv2D} = 3 \times 3 \times C_0 \times C_1 \quad (1)$$

2) DSC 的参数量  $P_{DSC}$

DSC 分为两步:

(1) DWConv: 对每个输入通道独立进行卷积, 卷积核为  $f^{3 \times 3 \times 1}$ , 每个输入通道需要一个卷积核。DWConv 的参数量  $P_{DWConv}$  为:

$$P_{DWConv} = 3 \times 3 \times C_0 \quad (2)$$

(2) PWConv: DWConv 中对每个通道单独卷积, 使得形成  $C_0$  个独立的特征图, 使用  $f^{1 \times 1}$  的卷积核对深度卷积输出的通道图进行线性组合, 以生成所需的输出通道数。PWConv 的参数量  $P_{PWConv}$  为:

$$P_{PWConv} = 1 \times 1 \times C_0 \times C_1 \quad (3)$$

因此,  $P_{DSC}$  为:

$$P_{DSC} = P_{DWConv} + P_{PWConv} = 3 \times 3 \times C_0 + 1 \times 1 \times C_0 \times C_1 \quad (4)$$

3) 减少的参数量  $\Delta P$

将  $P_{DSC}$  与  $P_{Conv2D}$  进行比较, 减少的参数量  $\Delta P$  为:

$$\Delta P = P_{Conv2D} - P_{DSC} = (8C_1 - 9)C_0 \quad (5)$$

#### 1.4 CBAM 注意力机制

UNet3+ 网络在分割含煤地表时具有一定的缺陷。而 CBAM 是一种结合了空间注意力和通道注意力的注意力机制模块, 通过对输入的特征图  $F$  进行通道维度和空间维度的注意力加权, 最终将两个注意力机制相乘得到最终的注意力特征图, 从而增强有用的特征并抑制无关的特征, 如图 6 所示。CBAM-UNet3+ 方法在解码器每个尺度特征融合后, 插入 CBAM 模块, 进一步提高对阴影特征的提取, 而抑制与煤相似特征的提取。

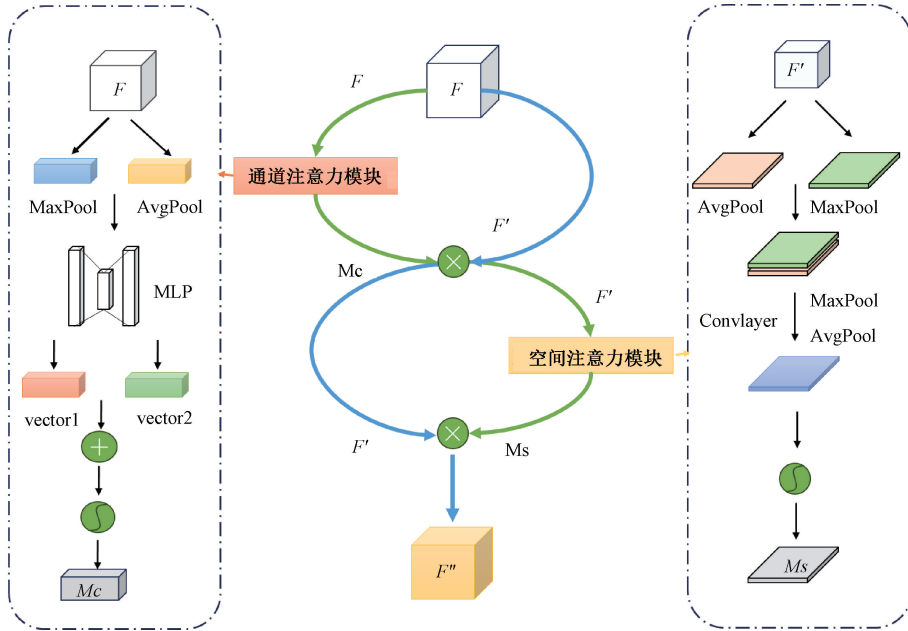


图 6 CBAM 注意力机制

Fig. 6 CBAM attention mechanism

##### 1) 通道注意力模块

首先输入特征图  $F$ , 对输入特征图分别进行全局平均池化和全局最大池化, 得到两个  $1 \times 1 \times C$  的特征图, 将两个特征图分别送入一个共享的全连接层 (由两个全连接层组成, 中间夹有 ReLU 激活函数), 将两个全连接层的输出相加, 并通过 Sigmoid 激活函数得到通道注意力图  $Mc \in R^{C \times 1 \times 1}$ , 将通道注意力图  $Mc$  与输入特征图  $F$  进行逐元素乘法, 得到加权后的特征图  $F'$ 。计算公式如式 (6)、(7) 所示。

$$Mc = \sigma\{MLP[AvgPool(F)] + MLP[MaxPool(F)]\} \quad (6)$$

$$F' = Mc \times (F) \quad (7)$$

其中,  $\sigma$  表示 Sigmoid 函数,  $AvgPool$  表示全局平均池化,  $MaxPool$  表示全局最大池化。

##### 2) 空间注意力模块

空间注意力模块, 输入经过通道注意力模块处理后的特征图  $F'$ , 对  $F'$  分别进行通道维度的平均池化和最大池化, 得到两个  $H \times W \times 1$  的特征图, 将两个特征图在通道维度上进行拼接, 得到一个新的特征图  $F_{cat} \in R^{H \times W \times 2}$ , 使用一个  $7 \times 7$  的卷积核 (通常带有适当的填充以保持空间尺寸不变) 对  $F_{cat}$  进行卷积操作, 以生成空间注意力图  $Ms \in R^{H \times W \times 1}$ , 通过 Sigmoid 激活函数将卷积结果转换为

空间注意力图  $M_s$  ,将空间注意力图  $M_s$  与经过通道注意力模块处理后的特征图  $F'$  进行逐元素乘法,得到最终的输出特征图  $F''$  。计算公式如式(8)~(11)所示。

$$P1 = MaxPool(F')$$
 (8)

$$P2 = AvgPool(F')$$
 (9)

$$M_s = \{f^{7 \times 7} \times [P1;P2]\}$$
 (10)

$$F'' = M_c \times (F')$$
 (11)

其中,  $\times$  表示两个矩阵逐元素相乘,  $f^{7 \times 7}$  表示卷积核为  $7 \times 7$  。

2 实验与分析

2.1 实验数据集

2018 年中煤平朔有限公司安家岭露天矿使用华测 P330pro 无人机搭载 HC15 正射相机进行数据采集,具体设备参数如表 1、2 所示。在数据采集阶段,依据测区范围以及高程数据,利用航线设计软件规划飞行航线,同时结合测区最低点与最高点的高程信息对航线进行优化。随后,无人机按照设计好的航线飞行,顺利获取了区域内的多幅高质量可见光影像。这些影像数据中包含完整的 POS 信息,即位置和姿态角数据,保障了较高的定位精度。最后,使用 Agisoft Photoscan Professional 软件对所采集的影像开展拼接处理工作。最终得到符合精度要求的 1:1 000 矿区正射影像图,使用该图进行阴影提取。

目前大多数公开的阴影数据集为城市航空高分辨率遥感影像数据集。因此,采用手动进行露天矿区影像的阴影掩膜标注工作,将正射影像图裁剪为 320 幅  $512 \times 512$  像

素的影像后,剔除其中不规则边界导致的低质量影像,最终得到 286 幅影像。

表 1 CTI P330pro 无人机参数  
Table 1 CTI P330pro UAV parameters

参数名称	参数值
尺寸	2 530 mm×1 210 mm
最大负载	2.0 kg
最大起飞重量	14 kg
最大电池寿命	160 min
垂直起飞高度	50~200 m
频率	100 Hz

表 2 HC15 相机参数  
Table 2 HC15 camera parameters

参数名称	参数值
传感器尺寸	35.9 mm×24 mm
像素大小	4.5 $\mu$ m
重量	330 g
像素	4.24 million
镜头焦距	36 mm

使用 python 中的库内插件 LabelMe 影像标注工具,对影像中的阴影进行标注,圈定阴影边界,获得阴影掩膜,并将其转换为与初始影像相同的 .jpg 格式,流程如图 7 所示。从数据集随机挑选 229 张为训练集、28 张为验证集、29 张为测试集,数据集命名为 DM-UAV。

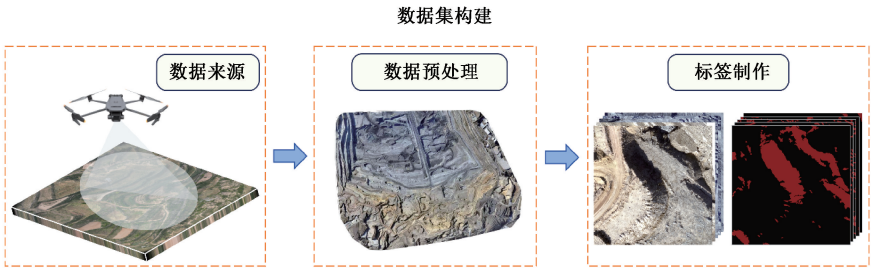


图 7 DM-UAV 数据集制作流程

Fig. 7 Production process of DM-UAV dataset

2.2 模型测试

为验证本文方法在露天矿区阴影提取中的可行性,在相同的实验环境配置条件下,在上文构建的露天矿区阴影数据集上,本文方法与 4 种典型提取模型 FCN、UNet、UNet++、UNet3+ 进行对比实验,并进行模块的消融实验。在硬件环境方面, GPU 为 NVIDIA GeForce RTX 4060 8 GB GDDR6 GPU, CPU 为 Intel(R)Core(TM) i5-14400(F)。操作系统为 Windows 11,深度学习框架为 pytorch, Python 版本为 3.8,加速平台 CUDA 版本为

12.4。设置所有模型参数为:训练迭代次数为 100,每批次训练的影像个数为乱序提供的 4 幅影像,初始学习率为 0.000 1,学习率下降(或衰减)方式'cos',使用 Adam 优化器,模型训练的损失函数使用二元交叉熵损失函数。由于数据集较小,采用多种数据增强方法扩展数据集,随机对训练影像使用 4 种增强方法,扩展训练数据集为初始 4 倍,使用数据增强方法包括:随机旋转角度、添加高斯噪声、调整亮度和高斯模糊 4 种增强方法,有效的提高模型的提取能力。

1)评价指标

为定量的评价本文模型性能,使用 4 个语义分割中广泛使用的评价指标  $mPrecision$ 、 $mRecall$ 、 $mF1$  以及  $mIoU$  对多种典型模型的预测结果进行评定。计算公式如式(12)~(19)所示。

$$Precision = \frac{TP}{TP + FP}$$
 (12)

$$mPrecision = \frac{1}{N} \sum Precision$$
 (13)

$$Recall = \frac{TP}{TP + FN}$$
 (14)

$$mRecall = \frac{1}{N} \sum mRecall$$
 (15)

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
 (16)

$$mF1 = \frac{1}{N} \sum F1$$
 (17)

$$IoU = \frac{TP}{TP + FP + FN}$$
 (18)

$$mIoU = \frac{1}{N} \sum IoU$$
 (19)

其中,  $TP$  表示正确预测为正类的样本数,  $FP$  表示错误预测为正类的样本数,  $FN$  表示错误预测为负类的样本数。

2)改进 UNet3+有效性验证

将改进的 UNet3+ 模型与 FCN、UNet、UNet++、UNet3+和改进的 UNet3+ 模型在数据集上实验进行对比,结果如图 8 所示,定量分析如表 3 所示。

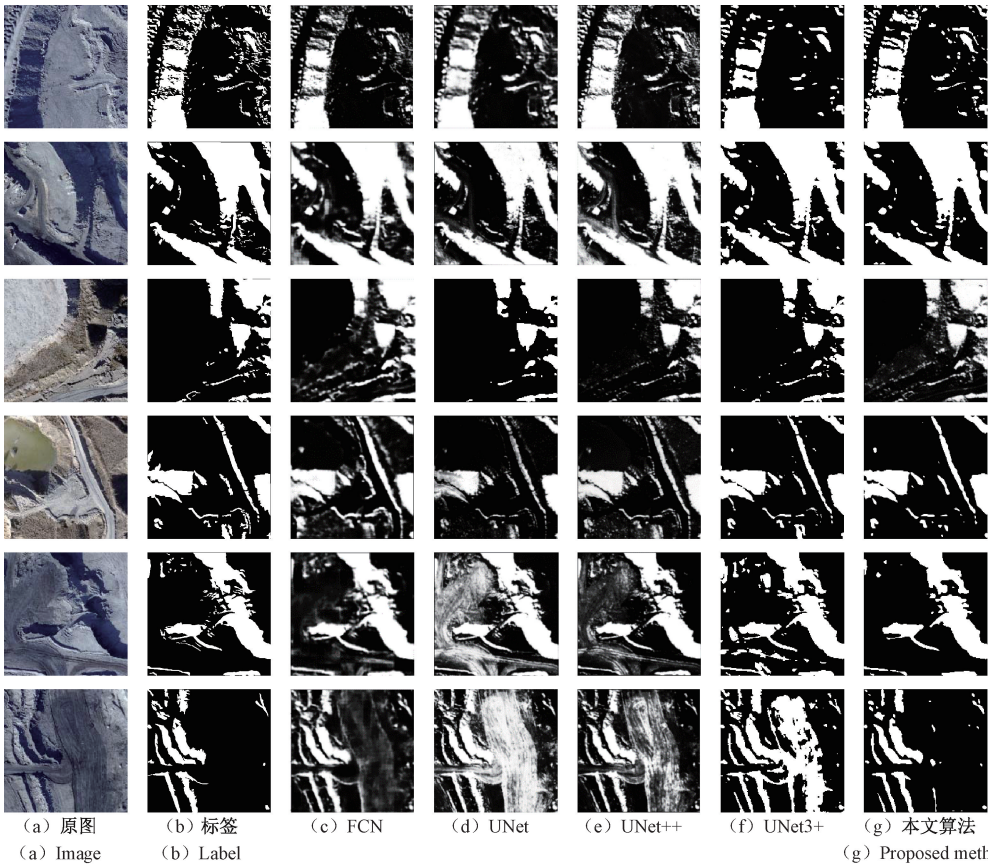


图 8 多种语义分割方法对阴影提取的结果对比

Fig. 8 Comparison of the results of shadow extraction by various semantic segmentation methods

通过图 8 分析可知,多种模型均可提取出较为明显的阴影区域,但在含煤区域,FCN、UNet、UNet++、UNet3+等方法具有明显的局限性,无法区分煤与阴影,在无煤区域,没有明显的误差,而且 FCN、UNet、UNet++边界分割较为粗略,准确性较低,UNet 相较 FCN 在边界中具有一定优势,但在分割相似类时效果比 FCN 差,UNet++相

较 UNet 网络分割效果更好,但在细碎的边界部分仍然准确度较低,UNet3+在分割边界中表现较为良好,本文方法在 UNet3+的基础上,精确率进一步提高,对应含煤区域也能准确判断从而分割提取。从表 3 中可看出相比原始 UNet3+ 网络,在  $mPrecision$ 、 $mRecall$ 、 $mF1$ 、 $mIoU$  分别提升了 4.9%、0.44%、2.24%、4.51%。



表 3 阴影提取方法结果精度对比

Table 3 Accuracy comparison of shadow extraction methods

方法	<i>mPrecision</i>	<i>mRecall</i>	<i>mF1</i>	<i>mIoU</i>
FCN	0.819 8	0.740 1	0.778 5	0.764 2
UNet	0.828 4	0.735 2	0.779 3	0.769 9
UNet++	0.852 6	0.787 3	0.815 8	0.789 4
UNet3+	0.900 3	0.915 1	0.907 4	0.835 0
本文算法	0.949 3	0.919 5	0.929 8	0.880 1

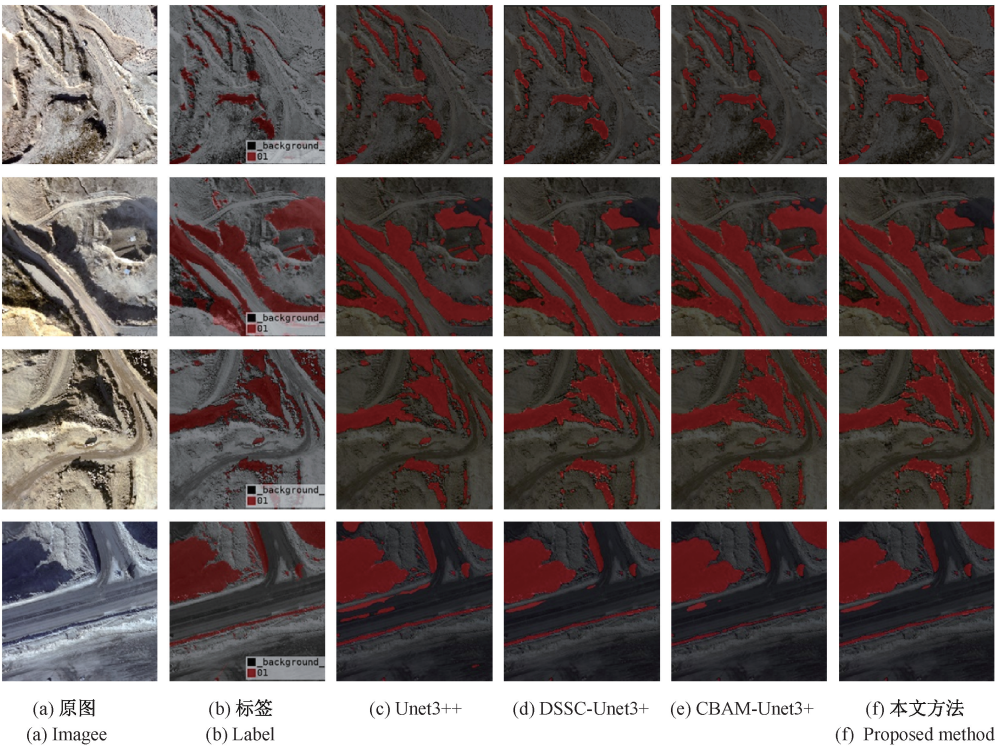


图 9 不同模块消融实验结果对比

Fig. 9 Comparisons of experiments on different modules result

表 4 不同模块的消融实验精度对比

Table 4 Accuracy comparison of ablation experiments on different modules

CBAM	DSC	<i>mPrecision</i>	<i>mRecall</i>	<i>mF1</i>	<i>mIoU</i>	<i>P</i>	<i>FPS/fps</i>
0	0	0.900 3	0.915 1	0.907 4	0.835 0	26 163 138	45
1	0	0.953 2	0.919 6	0.935 2	0.880 8	26 171 722	34
0	1	0.899 6	0.914 2	0.906 8	0.829 6	23 849 922	52
1	1	0.949 3	0.919 5	0.929 8	0.880 1	23 858 506	39

CBAM 模块的加入使得本文模型阴影分割的 *mPrecision* 提升了 5.29%，但网络模型复杂度进一步提升，网络更加复杂庞大，训练参数量增加了 8584，*FPS* 减少了 9 *fps*，而 DSC 的使用使得网络的复杂度与参数量减少，且平均精确率损失仅 0.07%，*FPS* 增加了 7 *fps*，两个模块的同时加入使得网络在降低复杂度与参数量的同时，阴影提取精度得到有效提升。

2.3 模块消融实验

为验证本文 DSC 及 CBAM 模块的有效性，在相同实验参数设置下，进行了 4 组消融实验，引入刷新率 (frames per second, *FPS*)，即模型每秒处理的影像数，评价模型处理速度。引入模型参数量，评价模型复杂程度。结果如图 9 所示，图 9(b) 中 01 代表前景，即标记出的阴影对象。定量分析结果如表 4 所示，1 代表添加该模块，0 代码未添加该模块。

2.4 与现有最佳方法的比较

为验证本文方法的有效性，选取与本文数据集相近的航空遥感图像数据集 AISD<sup>[22]</sup> 进行阴影提取对比实验，该数据集包含 514 组数据，覆盖了世界上 5 个不同的地区，即奥斯汀、芝加哥、蒂罗尔、维也纳和因斯布鲁克。图像大小从 256 pixel×256 pixel~1 688 pixel×1 688 pixel 不等，本文网络仅适用于 512 pixel×512 pixel 图像，因此，将该数据集中非 512 pixel×512 pixel 图像使用 Lanczos 重采样算法进行插值，



将图像全部统一至 512 pixel×512 pixel,并进行了训练测试。  
将基于改进 UNet3+的方法与王潇等<sup>[18]</sup>实验中涉及

的多种先进的无人机影像阴影提取方法进行了比较。结果如图 10 所示,定量分析如表 5 所示。

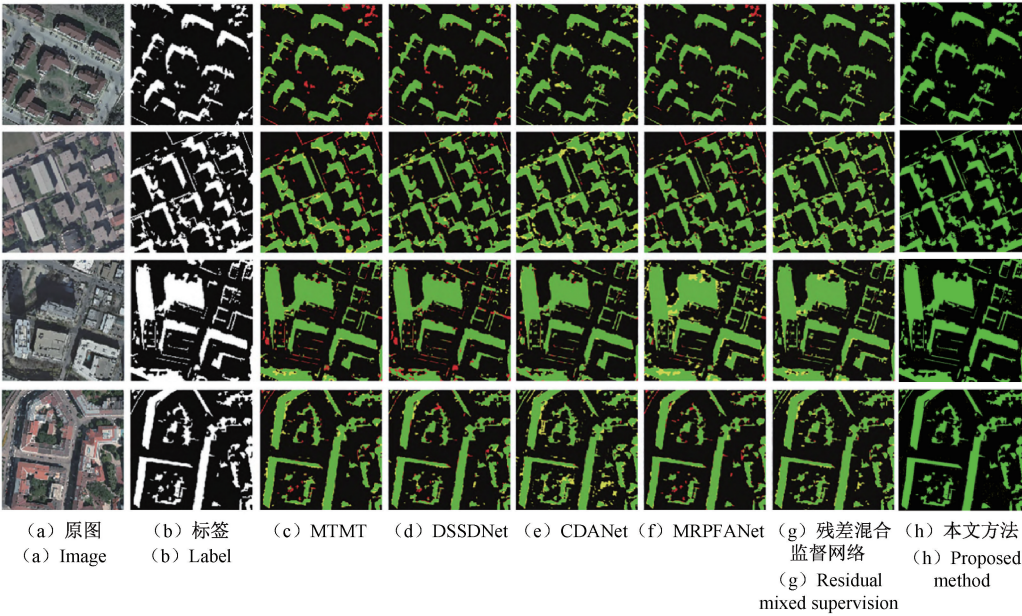


图 10 与其他方法的提取结果对比

Fig. 10 Compared with the extracted results of other methods

表 5 与现有阴影提取方法精度对比

Table 5 Compared with the accuracy of existing shadow extraction methods

方法	F1	IoU
MTMT	0.906 8	0.782 6
DSSDNet	0.917 9	0.802 9
CDANet	0.924 1	0.853 4
MRPFANet	0.926 1	0.862 4
残差混合监督网络	0.932 8	0.856 1
本文算法	0.935 5	0.880 1

如图 10 所示,绿色区域代表算法提取到正确的阴影区域,红色区域代表算法未提取到的阴影区域,黄色区域代表算法错误提取为阴影的区域。对结果分析可知,MTMT、DSSDNet、CDANet、MRPFANet 等方法的提取结果中有许多细碎的红色区域,残差混合监督网络红色区域较少,但存在较多黄色区域,本文方法的提取结果中,红色区域与黄色区域面积更小、数量更少,这意味着本文方法在提取阴影方面精度较高,能够更加精准的提取细碎阴影区域,F1 与 IoU 相比残差混合监督网络提高了 0.27%、2.62%。

3 结 论

本文制作了露天矿区的无人机影像阴影数据集,弥补了目前阴影数据集多集中于城市区域的不足。

通过在原始 UNet3+网络中同尺度跳跃时的二维卷积层替换为 DSC 卷积层、增加 CBAM 模块,解决了 UNet3+在露天矿区识别精度不高,且网络复杂度较高及参数量过多的问题,通过与其他典型语义分割网络在露天矿无人机阴影数据集上的对比实验,本文模型的 *mPrecision*、*mRecall*、*mF1* 以及 *mIoU* 分别比 UNet3+方法提升了 4.9%、0.44%、2.24%、4.51%,对露天矿区阴影具有更好的提取效果。

通过消融实验结果对比,证明了本文添加 CBAM 模块及 DSC 的有效性,CBAM 模块有效提升了提取效果,DSC 应用有效降低了参数量,提升了训练速度,提高了本文模型在露天矿区阴影提取的适用性,本文模型适合进行露天矿区无人机影像的阴影提取。

在 AISD 公开数据集上的实验表明本文所提的无人机影像阴影提取方法与现有的无人机阴影提取方法相比,F1 与 IoU 相比残差混合监督网络提高了 0.27%、2.62%,证明了该方法在阴影提取方面的准确性。

未来研究将致力于构建适用性更加广阔的露天矿区阴影数据集,并尝试阴影地带进行去除。

参考文献

[1] 吴挺. 无人机倾斜摄影在矿山生态修复中的应用研究——以长兴县李家巷鑫茂矿为例[D]. 杭州:浙江大学, 2020.  
WU T. Application of UAV tilt photography in mine ecological restoration-a case study of Xinmao mine in

- Lijiaxiang, Changxing Conty[D]. Hangzhou: Zhejiang University, 2020.
- [2] 刘晓霞, 杨风暴, 卫红, 等. 一种农田作物无人机遥感影像阴影检测算法[J]. 电子测量技术, 2022, 45(2): 135-139.
- LIU X X, YANG F B, WEI H, et al. A shadow detection algorithm for remote sensing images of farmland crops bu UAV[J]. Electronic Measurement Technology, 2022, 45(2): 135-139.
- [3] 冯占科, 钱旺, 彭建川. 基于综合权重—TOPSIS 的露天矿边坡稳定性评价[J]. 工矿自动化, 2023, 49(S1): 133-137.
- FENG ZH K, QIAN W, PENG J CH. Slope stability evaluation based on comprehensive weight and TOPSIS[J]. Journal of Mine Automation, 2023, 49(S1): 133-137.
- [4] 杨勇, 张忠政, 胡军, 等. 基于随机权重法改进 PSO-ELM 的露天矿边坡稳定性分析[J]. 有色金属工程, 2022, 12(5): 128-134.
- YANG Y, ZHANG ZH ZH, HU J, et al. Slope stability analysis of open-pit mine based on improved PSO-ELM with random weight method [J]. Nonferrous Metals Engineering, 2022, 12(5): 128-134.
- [5] IBARRA A M J, TIAHJADI T, PÉREZ O J. Shadow detection in still road images using chrominance properties of shadows and spectral power distribution of the illumination[J]. Sensors, 2020, 20(4): 1012.
- [6] LI G, WEN L, HUANG Z, et al. Data augmentation and shadow image classification for shadow detection[J]. IET Image Processing, 2022, 16(3): 717-728.
- [7] HUANG P, SU CH, WU B W, et al. Global context aware dual channel pyramid model for robust image shadow removal [J]. Engineering Applications of Artificial Intelligence, 2024, 133: 108552.
- [8] KAKOOEI M, BALEGHI Y. Shadow detection in very high resolution RGB images using a special thresholding on a new spectral-spatial index[J]. Journal of Applied Remote Sensing, 2020, 14(1): 016503.
- [9] 郭伟, 陆婉芸, 邓廷起, 等. 高分航空光学遥感影像的山体阴影去除方法[J]. 测绘通报, 2023(11): 30-35, 47.
- GUO W, LU W Y, DENG T Q, et al. Mountain shadow removal for high-resolution aerial optical remote sensing images[J]. Bulletin of Surveying and Mapping, 2023(11): 30-35, 47.
- [10] 张向阳, 罗小军, 包佳文, 等. 梯度-纹理特征超像素分割的 SAR 图像阴影识别[J]. 遥感信息, 2023, 38(4): 106-112.
- ZHANG X Y, LUO X J, BAO J W, et al. Recognition of SAR image shadow based on superpixel segmentation of gradient-texture features[J]. Remote Sensing Information, 2023, 38(4): 106-112.
- [11] ŠTAMPFL V, AHTIK J. Shadow segmentation with image thresholding for describing the harshness of light sources [J]. IEEE Transactions on Image Processing, 2024, 33: 3428-3440.
- [12] 张昊, 焦瑞莉, 乔聪聪, 等. 基于纹理特征的深度学习云和云阴影检测[J]. 计算机工程与设计, 2024, 45(5): 1580-1587.
- ZHANG H, JIAO R L, QIAO C C, et al. Deep learning based on texture features for cloud and cloud shadow detection [J]. Computer Engineering and Design, 2024, 45(5): 1580-1587.
- [13] 韩红印. 高分辨率光学遥感图像阴影检测与补偿技术研究[D]. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所), 2021.
- HAN H Y. Research on shadow detection and compensation technology for high resolution optical remote sensing images[D]. Changchun: University of Chinese Academy of Sciences (Institute of Applied Optics and Precision Mechanics, Chinese Academy of Sciences), 2021.
- [14] SHU R, ZHAO S. Multi-resolution learning and semantic edge enhancement for super-resolution semantic segmentation of urban scene images [J]. Sensors, 2024, 24(14): 4522.
- [15] ZHANG J, SHI X L, ZHENG C Y, et al. MRPFA-Net for shadow detection in remote-sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5514011.
- [16] 崔志强, 单慧琳, 张银胜, 等. 基于改进型 U-Net 的遥感云图分割方法[J]. 电子测量技术, 2022, 45(12): 127-132.
- CUI ZH Q, SHAN H L, ZHANG Y SH, et al. Method of remote sensing cloud image segmentation algorithm based on improved U-Net[J]. Electronic Measurement Technology, 2022, 45(12): 127-132.
- [17] 刘秋月, 刘雪峰, 孙绍华. 基于阴影增强和注意力机制的高光谱图像分类[J]. 电子测量技术, 2023, 46(8): 14-23.
- LIU Q Y, LIU X F, SUN SH H. Hyperspectral image classification based on shadow enhancement and attention mechanism [J]. Electronic Measurement

- Technology, 2023, 46(8): 14-23.
- [18] 王潇, 刘贞报, 史忠科. 基于残差混合监督网络的无人机目标阴影检测[J]. 航空学报, 2024, 45(17): 530062.
- WANG X, LIU ZH B, SHI ZH K. Shadow detection of UAV target based on residual mixed supervision network[J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(17): 530062.
- [19] 徐梓川, 龚晓峰. 基于语义指导和自适应卷积的遥感云检测算法[J]. 电子测量技术, 2024, 47(1): 136-143.
- XU Z CH, GONG X F. Cloud detection algorithm for remote sensing images based on semantic-guided and adaptive convolution [J]. Electronic Measurement Technology, 2024, 47(1): 136-143.
- [20] HUANG H, LIN L, TONG R, et al. UNet 3+: A full-scale connected unet for medical image segmentation[C]. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020: 1055-1059.
- [21] 杨光义, 刘泽鹏, 把硕, 等. 基于无监督学习的遥感道路阴影遮挡补全[J]. 武汉大学学报(工学版), 25(3): 482-492.
- YANG G Y, LIU Z P, BA SH, et al. Complementation of road shadowed occlusion in remote sensing images based on unsupervised learning[J]. Engineering Journal of Wuhan University, 25(3): 482-492.
- [22] LUO S, LI H, SHEN H. Deeply supervised convolutional neural network for shadow detection based on a novel aerial shadow imagery dataset[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 167: 443-457.
- [23] CHEN Z H, ZHU L, WAN L, et al. A multi-task mean teacher for semi-supervised shadow detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 5610-5619.
- [24] ZHU Q Q, YANG Y, SUN X L, et al. CDANet: Contextual detail-aware network for high-spatial-resolution remote-sensing imagery shadow detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5617415.
- [25] XIE Y, CHEN R, YU M, et al. Improvement and application of UNet network for avoiding the effect of urban dense high-rise buildings and other feature shadows on water body extraction[J]. International Journal of Remote Sensing, 44(12): 3861-3891.
- [26] YANG D, LIU Z, FENG D, et al. Resnet-Unet considering patches (RUP) network to solve the problem of patches due to shadows in extracting building top information [J]. Journal of Spatial Science, 69(1): 243-263.
- [27] HU X, FU C W, ZHU L, et al. Direction-aware spatial context features for shadow detection and removal[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(11): 2795-2808.

## 作者简介

杜孙稳(通信作者), 博士, 副教授, 主要研究方向为影像智能识别, 目标检测。

E-mail: dusunwen2006@163.com

宋瑞婷, 硕士研究生, 主要研究方向为图像处理与目标识别。

E-mail: songruiting2001@163.com