

面向梯田遥感识别的JAM-R-CNN深度网络模型

谢君洋^{1,2}, 林安琪^{1,2}, 吴浩^{1,2}, 吴紫薇^{1,2}, 吴文斌³, 余强毅³

1. 华中师范大学 城市与环境科学学院, 武汉 430079;
2. 地理过程分析与模拟湖北省重点实验室, 武汉 430079;
3. 中国农业科学院农业资源与农业区划研究所, 北京 100081

摘要: 快速准确地掌握梯田的空间分布, 不仅为水土保持提供重要的数据支撑, 也提高了山区农业的监管水平。利用深度学习方法进行梯田识别, 对于形状狭长的梯田容易因卷积运算而造成漏提现象, 并且在山区地物边界不清晰的复杂背景下, 易产生大面积粘连的识别结果, 导致最终的梯田识别精度不高。为此, 本研究提出了一种面向梯田遥感识别的JAM-R-CNN深度网络模型。该模型以Mask R-CNN为基础, 融合跳跃网络来维持高分辨率遥感影像的高语义信息, 引入卷积注意力机制模块来加强梯田的特征表达能力, 修改模型的锚框大小以适应梯田狭长的特性。并以重庆市南川区的盐井梯田研究区域, 基于国产高分二号(GF-2)卫星影像数据进行本研究模型验证实验。结果表明: JAM-R-CNN网络模型的梯田识别结果精确率为90.81%, 召回率为84.28%, F1为88.98%, IoU为83.15%。相较于经典的Mask R-CNN模型, JAM-R-CNN网络模型的梯田识别精度较高, 4个评价指标依次提升了1.96%、5.26%、3.29%和5.19%。此外, 消融实验结果验证了模型改进的3个模块均对梯田识别有明显的促进作用。综上, 本研究提出的JAM-R-CNN深度网络模型能够有效减少梯田识别结果的粘连现象, 明显提高了狭长型梯田的提取率, 实现了梯田遥感识别整体精度的显著提升, 具有较好的应用价值。

关键词: 遥感, 梯田识别, 高分辨率遥感影像, 深度学习, 跳跃网络, JAM-R-CNN

中图分类号: S127/P2

引用格式: 谢君洋, 林安琪, 吴浩, 吴紫薇, 吴文斌, 余强毅. 2024. 面向梯田遥感识别的JAM-R-CNN深度网络模型. 遥感学报, 28(12): 3136-3146

Xie J Y, Lin A Q, Wu H, Wu Z W, Wu W B and Yu Q Y. 2024. JAM-R-CNN deep learning network model for remote sensing recognition of terraced fields. National Remote Sensing Bulletin, 28(12): 3136-3146 [DOI: 10.11834/jrs.20233126]

1 引言

梯田是山区丘陵地带最为常见的基本农田, 在保障粮食安全和提升水土保持能力等方面发挥着重要作用(董世杰等, 2023)。中国梯田数量众多、分布零散、结构复杂, 给准确高效的梯田信息获取带来了极大挑战。早期的梯田识别主要通过人工实地调查的方式(刘晓燕等, 2014), 该方法成本高、时效性低, 难于实现大范围的梯田制图。随着遥感技术的迅猛发展, 遥感影像分辨率已经达到米级、甚至厘米级(陈玲玲等, 2022), 为精细化的梯田信息提取创造了有力条件。特别

是利用机器学习等智能化方法, 通过从遥感影像中挖掘梯田特性来构建自动化的识别模型, 可以大大节省人工解译的时间和成本。因此, 在高分辨率遥感影像中利用相关计算模型自动地对梯田进行识别, 能够高效地获取梯田信息, 对提升梯田智能化监测与管理水平具有重要意义。

梯田遥感识别的传统方法主要包括傅里叶变换(赵欣等, 2016; Luo等, 2020)、面向对象(Diaz-Varela等, 2014; 党恬敏等, 2017; 周珏等, 2022)以及浅层机器学习(Eckert等, 2017; Sun等, 2019; 薛牡丹等, 2019)等, 提取的梯田较为破碎、分散, 难以满足智慧农业的精细化管理

收稿日期: 2023-04-21; 预印本: 2023-06-13

基金项目: 国家重点研发计划(编号: 2022YFB3903500); 湖北省自然科学基金创新群体项目(编号: 2024AFA032); 国家自然科学基金(编号: 42071358, 42201468); 中央高校基本科研业务费资助(编号: CCNU22QN018); 湖北省自然资源厅科研计划(编号: ZRZY2023KJ03)

第一作者简介: 谢君洋, 研究方向为农业遥感。E-mail: xjy959@mails.ccnu.edu.cn

通信作者简介: 吴浩, 研究方向为地理信息科学与遥感。E-mail: haowu@ccnu.edu.cn

需求。近年来,深度学习模型逐渐得到应用,通过自动学习样本中的深层次特征来获取高分遥感影像的强语义信息(邓培芳等,2021;蔡志文等,2022;许泽宇等,2022),弥补了传统梯田提取方法只能提取样本浅层特征的不足,在精度和效率上均有较大提升。例如,Dai等(2020)基于高分辨率遥感影像和数字高程模型,利用Canny边缘检测与地形分析相结合的方法实现自动绘制梯田;刘巍等(2021)针对山区复杂的地理空间异质性特征,将耕地细分为规则耕地、梯田、坡耕地以及林间耕地,并使用RCF网络模型来提取梯田;Zhao等(2021)利用Unet对梯田进行自动提取;Yu等(2022)利用深度迁移学习方法在像素级层面提取梯田。然而,这些语义分割模型识别的梯田会产生大范围的粘连,无法准确刻画单一梯田的边界,需作进一步处理才能提升地物识别的精细化程度。例如,Waldner和Diakogiannis(2020)结合Unet和ResNet网络进行多任务学习,再通过阈值分割和分水岭分割两种方法生成封闭的边界来提取耕地地块;周楠等(2021)在语义分割结果基础上,加入边缘检测任务来进行地块级的耕地提取;Jong等(2022)在Unet模型中以残差网络为骨架来对农田边界进行识别;Long等(2022)使用多任务学习来提高农田地块边界的识别精度。而实例分割模型可以有效减少地物提取结果的粘连,例如Mei等(2022)利用Mask R-CNN模型在高分辨率图像中对规则的农田地块边界进行了准确地绘制。

不同于条田和格田,梯田具有形状狭长、轮廓不规则、面积差异大等特点,遥感识别十分困难,主要表现在两个方面:一方面,虽然Mask R-CNN模型可以在一定程度上提取高分辨率遥感影像的深层次信息,但在卷积过程中仍然会丢失大量的细节信息,导致面积较小、形状狭长的梯田无法准确识别;另一方面,对于集中连片分布的梯田来说,遥感识别结果往往容易出现粘连现象,识别精度有限。因此,为进一步实现梯田信息的准确识别,亟需解决的技术问题是如何依据梯田特性,有效维持高分遥感影像在卷积运算过程中的高语义信息,减少狭长梯田的漏提以及识别结果的粘连情况。

因此,本研究提出了一种面向梯田遥感识别的JAM-R-CNN(Jump network Anchor Mask R-CNN)

深度网络模型,该模型以Mask R-CNN为基础,针对梯田识别结果的粘连及狭长型梯田易漏提等问题,通过将跳跃网络与模型的FPN模块相融合、引入卷积注意力机制模块以及修改模型的锚框大小3方面的改进,以达到提升梯田识别精度的目的,并以重庆市南川区的盐井梯田区域为例,对本研究提出的JAM-R-CNN模型进行试验验证,最后通过消融实验来验证改进的3个模块对梯田识别的促进作用,从而为梯田地块的精确识别提供一种可靠的方法。

2 研究方法

本研究提出的面向梯田遥感识别的JAM-R-CNN深度网络模型的框架如图1所示。该方法以Mask R-CNN模型(He等,2017)为基础,采用ResNet 50作为网络骨架,通过融入跳跃网络模块和卷积注意力机制模块来增强模型对梯田的多尺度特征表达能力,并修改模型的区域建议网络模块来适应梯田本身的形状特性,以期提升梯田识别精度。

2.1 跳跃网络模块

在Mask R-CNN模型的FPN模块中,利用深度卷积神经网络中固有的多尺度特征图,加入横向连接和上采样,来以极小的附加计算量构建不同尺度的具有高级语义信息的特征金字塔网络。而在此模块的上采样过程中,由于卷积神经网络通过插值的方式从小到大对特征进行还原,再直接与对应尺寸的特征进行相加,导致高阶的语义信息不能有效传播,损失部分信息。对于面积小、形状狭长的梯田来说,就容易因此产生漏提现象。

高分辨率影像的优势是具备丰富的细节特征,有利于相关地物的提取。为了使高分辨率遥感影像中的梯田细节信息在整个运算过程中不受损失,本研究借鉴HRNet思想(Wang等,2021),在FPN原有的横向连接基础上融入跳跃网络模块,并且跳跃的层数随着相应特征层上特征图尺寸的增大而增加,具体结构如图1所示。从横向来看,在每一层特征图中通过一种重复卷积计算来保持每一层的同分辨率特征图,卷积核大小为 3×3 。由于特征图尺寸越大,包含的语义特征信息就更加丰富,通过多次重复卷积,以获取更多的高语义特征信息。所以在改进模型中,特征图尺寸越高,重复的结构就越多。从纵向来看,将浅层特征图通过

卷积的方式（卷积核大小为 3×3 ，步长为2）不断融入到深层特征图中，丰富了深层特征图中的高分辨率信息。网络中所有的连接操作均采用相加的方式，以保留特征图更完整的空间信息。在原

始FPN中融合跳跃网络模块，减少了高分遥感的语义特征损失，维持了全过程的高分辨率细节特征信息，有助于提升梯田识别的精度。

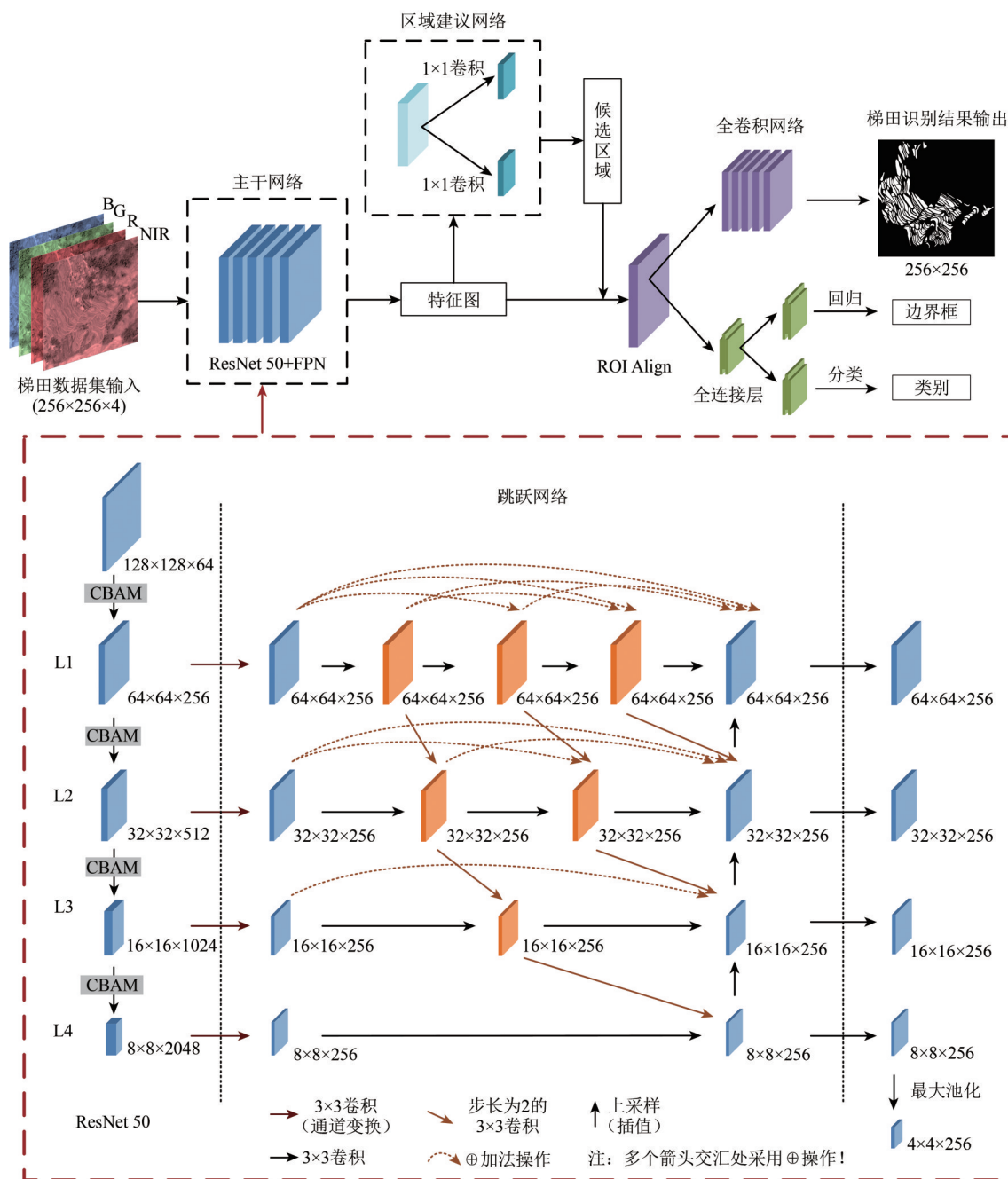


图1 面向梯田遥感识别的JAM-R-CNN深度网络模型框架图

Fig. 1 The framework of terrace recognition method based on JAM-R-CNN

2.2 卷积注意力机制模块

卷积注意力机制CBAM (Convolutional Block Attention Module) (Woo等, 2018) 模块可以让深度学习模型在通道轴和空间轴两个维度上关注重要特征, 抑制不必要特征, 其结构如图2所示。该

模块由通道注意力和空间注意力组成, 首先利用原始特征 F 计算通道注意力, 并与原始特征进行乘法运算得到 F' , 再利用 F' 计算空间注意力, 并与 F' 进行加权得到最终的卷积注意力机制结果 F'' 。

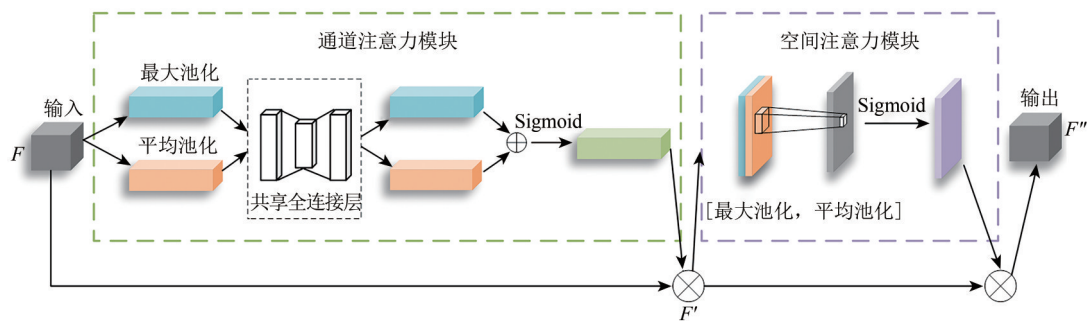


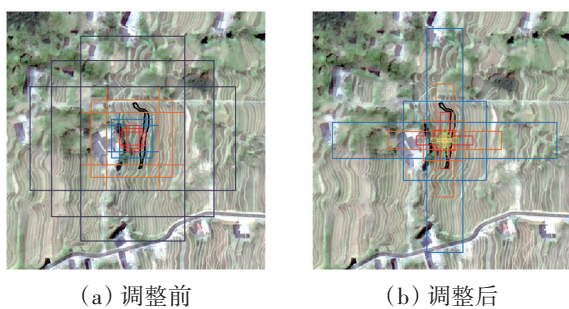
图2 卷积注意力机制结构图

Fig. 2 The structure of convolutional block attention module

原始FPN模块采用残差卷积神经网络的方式获取了不同尺度的特征图，在获取每层特征之前加入CBAM模块，可以强化梯田的多尺度特征，使输入到预测层的梯田特征更聚合，增强预测层的特征表达能力，最终提升梯田识别精度。

2.3 区域建议网络模块

区域建议网络RPN (Region Proposal Network) 用于在特征图上生成最优的候选框 (袁山等, 2022)，为模型后续的分类、检测与分割服务。由于梯田狭长的特性，原始锚框anchor的长宽比 (1:1, 1:2, 2:1) 不足以准确地对梯田进行真实表达，因此需要调整锚框的长宽比来适应梯田实际的长宽比。通过计算样本中所有梯田块的外接矩形长宽比，最终选定锚框的长宽比分别为1:1、1:5、5:1。此外，本研究修改锚框的面积分别为 (8, 16, 32, 64, 128)，图3展示了模型锚框大小调整前、后的候选区域情况。



(a) 调整前 (b) 调整后

图3 锚框大小调整前、后的候选区域展示图

Fig. 3 The candidate area before or after modifying anchor size

3 研究区域与数据

3.1 研究区域

本研究选取的研究区域为重庆市南川区石溪

镇的盐井梯田，具体位置如图4所示。盐井梯田地处渝、黔两省（直辖市）交汇点，区域内山岭纵横，河谷幽深，地形复杂。区域内耕地类型以梯田为主，属于典型的南方稻作梯田，由于区域内多数的梯田形状较为狭长，面积较小，中低分辨率影像难以准确识别梯田的边界范围，提取难度大。因此，对此区域的梯田进行识别提取较为困难，有助于深度研究基于高分辨率遥感影像的梯田识别方法，并验证其可行性。

3.2 GF-2数据及预处理

本研究使用的遥感影像数据为国产GF-2卫星影像，来源于中国资源卫星应用中心陆地观测卫星数据服务平台 (<https://data.cresda.cn/#/2dMap>)，采集时间为2021年5月9日。首先基于ENVI软件对下载的GF-2影像数据的全色波段和多光谱波段进行辐射定标、大气校正及几何校正等预处理，再使用NNDiffuse Pan Sharpening方法对全色波段和多光谱波段进行图像融合，得到分辨率为1 m的GF-2多光谱影像。最后，裁剪出10240×10240大小的影像，人工进行样本标注。

为了检验标注的梯田样本的可靠性，在研究区域影像中随机选取了600个验证点，其中梯田样本321个，非梯田样本279个，均匀分布在研究区域内，具体位置如图4所示。结合Google Earth卫星地图影像20级产品逐一检验600个验证点是否正确，以此对标注样本的精度进行验证，其中卫星地图影像分辨率为0.3 m。通过验证，本研究实验的样本数据集的总体精度为98.50%，Kappa系数为0.97，表明标注的样本可信度较高。

3.3 实验设置

本研究将原始四波段影像数据和进行标注后的矢量样本数据通过数据增强（分别旋转90°、

180°和270°)后转换为COCO数据集,共包含25600张图片,尺寸为256×256,并按照8:1:1的比例将数据集划分为20480张训练图片、2560张验证图片和2560张测试图片。在进行数据集制作过程中,采用重叠平铺策略进行影像裁剪,滑动窗口大小为256,步长为128。同时,为了使实验结果覆盖整个区域,本研究依据滑动窗口的大小

将影像数据的外围扩充了64个像元,并以0值进行填充。本研究在Ubuntu系统中,使用Pytorch框架进行实验,显卡为Nvidia GeForce RTX 3060(显存12 GB)。模型初始学习率为0.02,学习率衰减方式采用余弦衰减,迭代次数为300, batch size为4,采用SGD优化器进行实验,并选取最优迭代次数的训练结果进行最终的预测。

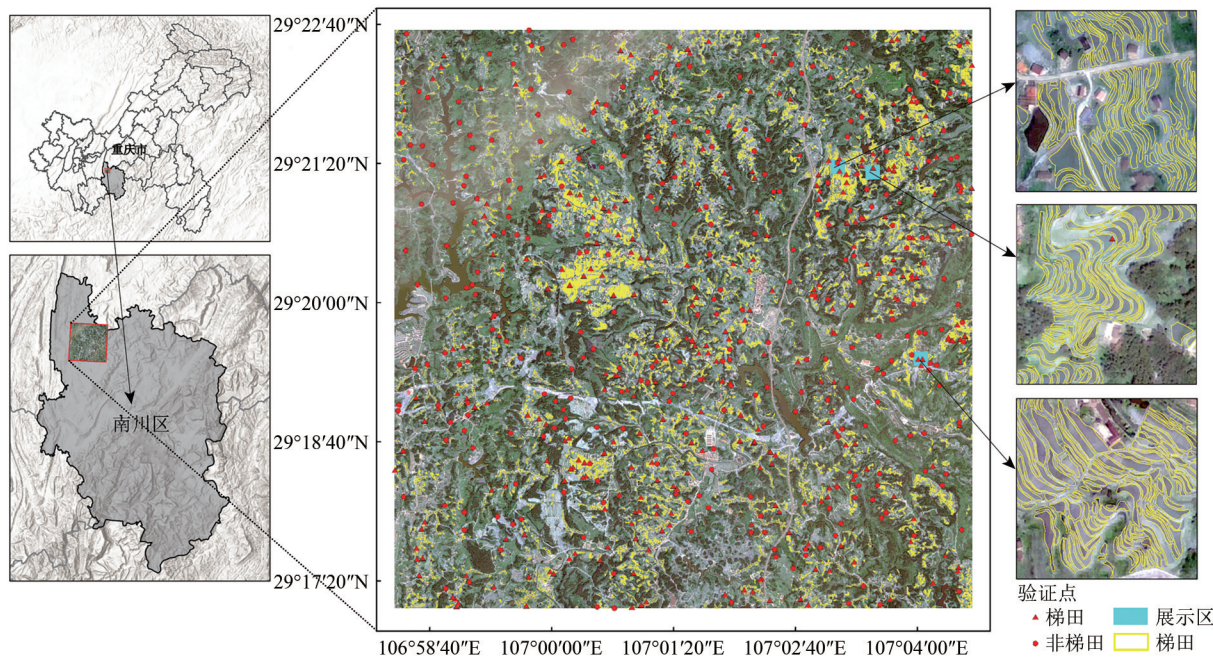


图4 实验区位置

Fig. 4 Location of the study area

3.4 精度评价指标

为了检验模型的有效性,本研究采用精确率(P)、召回率(R)、F1得分值(F1)和交并比(IoU)作为梯田提取结果的评价指标。以样本标记为真实值,模型实验的输出结果为预测值,实际和预测均为梯田的像素记为TP;实际为非梯田,预测为梯田的像素记为FP;实际为梯田,预测为非梯田的像素记为FN;实际和预测均为非梯田的像素记为TN。上述4种评价指标的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (3)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (4)$$

式(1)–(4)中, P 表示预测为梯田的像元中真正

为梯田像元的占比,即预测与实际均为梯田的像元数占预测为梯田的所有像元数的比值; R 表示实际为梯田的像元中被正确预测为梯田像元的占比,即预测与实际均为梯田的像元数占实际为梯田的所有像元数的比值;F1是对准确率和召回率的综合评价指标,其主要用来判断两者之间的差异大小;IoU表示梯田提取的结果和真实情况之间的交集与并集的比值。

4 结果与分析

4.1 梯田地块提取结果

利用本研究提出的JAM-R-CNN网络模型对样本数据进行训练后,使用最优迭代次数的训练权重数据对整个实验区进行预测,将预测结果的中心区域(128×128)进行提取拼接得到最终的结果,如图5所示。

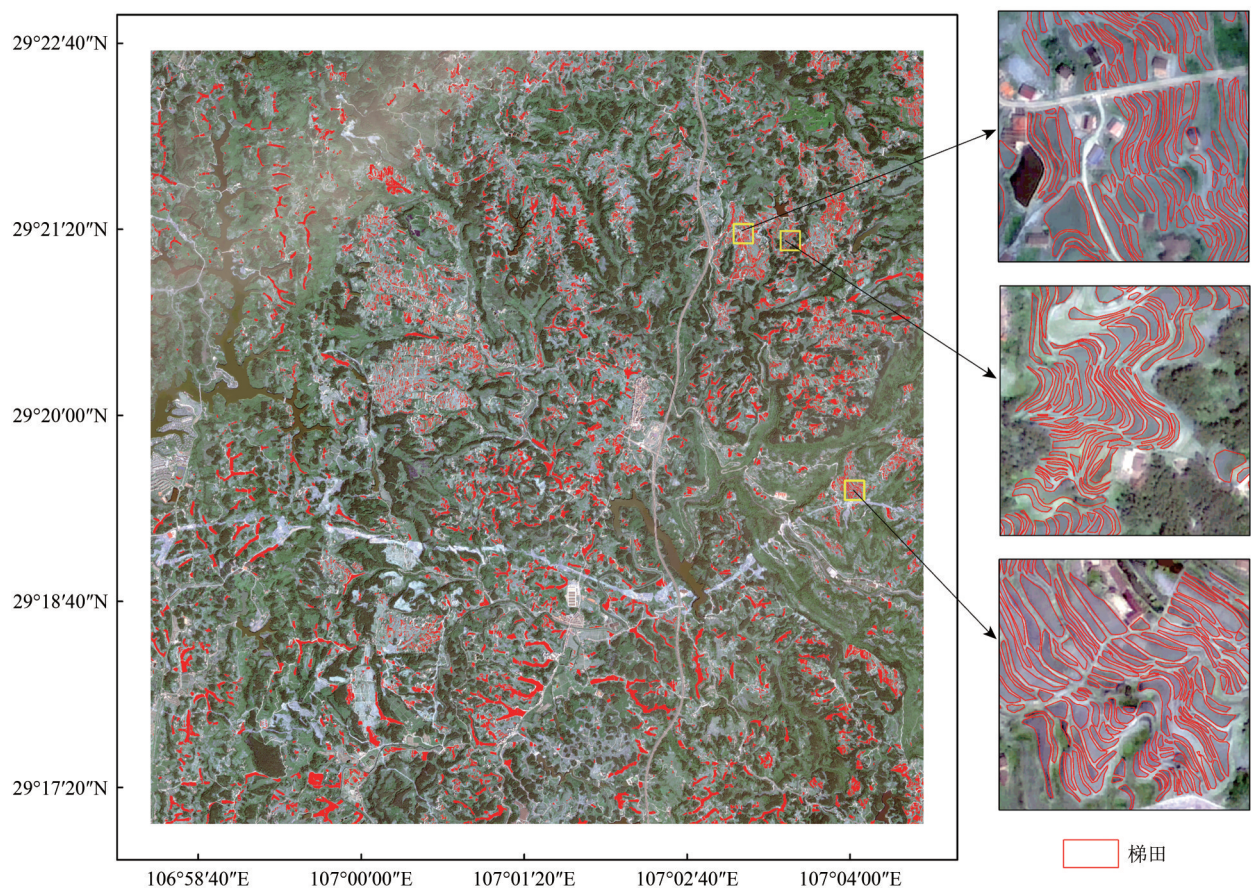


图5 JAM-R-CNN模型的梯田识别结果
Fig. 5 The results of terrace recognition based on JAM-R-CNN

为进一步的对比本研究提出的JAM-R-CNN网络模型对梯田识别的效果，选用Unet、DeepLab v3+和Mask R-CNN这3种基础网络模型进行对比实验，各模型使用相同的数据集和实验参数，具体实验结果的精度见表1。可知：本研究提出的JAM-R-CNN网络模型的梯田识别结果精确率为90.81%，召回率为84.28%，F1为88.98%，IoU为83.15%，比Mask R-CNN模型的梯田识别结果精确率、召回率、F1和IoU依次提升1.96%、5.26%、3.29%和5.19%。与其他两个语义分割模型进行对比发现，JAM-R-CNN模型的梯田识别结果精度较高。其中：精确率比Unet和DeepLab v3+模型提升了9.17%和6.31%；召回率比Unet模型提升了1.52%，比DeepLab v3+模型降低了0.89%；F1提升了7.34%和4.15%；IoU提升了11.69%和9.49%。

各个模型对梯田的识别结果如图6所示。可见：在Unet和DeepLab v3+模型的梯田识别结果中有较多的粘连情况，无法准确获取梯田信息；而在JAM-R-CNN和Mask R-CNN两种模型的梯田识别结果中，只存在极少的粘连现象，绝大部分识

别结果较好；同时，相较于这两种模型的实验结果，Unet和DeepLab v3+模型识别的梯田空洞较多，降低了梯田识别结果的精度。因此，这进一步说明了实例分割模型结果粘连现象较少，可以直接、准确的识别梯田信息，而语义分割则需进一步作地物边缘优化处理，才能达到准确识别梯田的目的。

表1 各模型的梯田识别精度
Table 1 The accuracy of terrace recognition by using different models

模型	评价精度			
	P	R	F1	IoU
Unet	81.64	82.76	81.64	71.46
DeepLab v3+	84.50	85.17	84.83	73.66
Mask R-CNN	88.85	79.02	85.69	77.96
JAM-R-CNN	90.81	84.28	88.98	83.15

为了验证重叠平铺策略对梯田的识别是否有促进作用，本研究利用JAM-R-CNN模型进行实验。在其他参数都保持一致时，对比是否使用这种策

略的梯田识别结果，识别精度详见表2。可知，采用重叠平铺策略的梯田识别结果较不采用这种策略

的识别结果精度有一定的提升，精确率、召回率、F1和IoU依次提升了0.14%、0.36%、0.21%和0.35%。

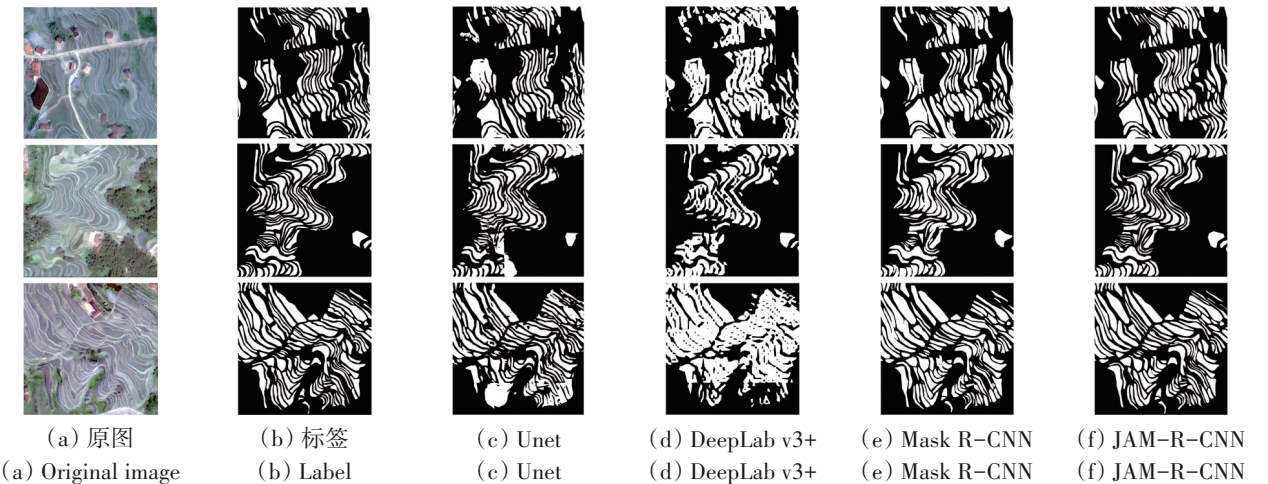


图6 各模型梯田识别结果对比

Fig. 6 Comparison of terrace recognition results of different models

表2 重叠平铺策略识别精度

Table 2 Comparison of the accuracy of terrace recognition results with overlapping tiling strategy

是否使用重叠平铺策略	评价精度			
	<i>P</i>	<i>R</i>	F1	IoU
不使用	90.67	83.92	88.77	82.80
使用	90.81	84.28	88.98	83.15

为了更直观的说明重叠平铺策略的实验效果，具体的对比如图7所示。可知，在采用重叠平铺策略的实验结果中，将影像块进行拼接后没有产生拼接痕迹，使得最终的梯田识别精度有所提升。因此，可以认为重叠平铺策略消除了由于模型运算过程中影像边缘识别精度低而在影像拼接后产生的不协调拼接痕迹，促进了梯田信息的识别。

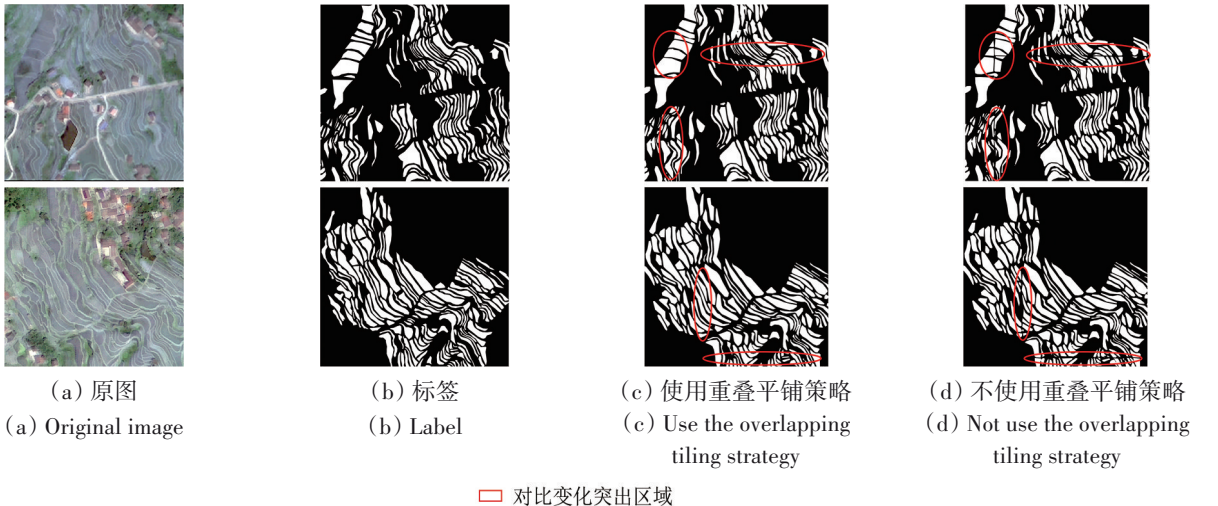


图7 重叠平铺策略结果对比

Fig. 7 Comparison overlapping tiling strategy results

4.2 消融实验

本研究对Mask R-CNN模型的改进主要在3个部分：加入跳跃网络、CBAM和修改Mask R-CNN模型的锚框大小。通过消融实验来验证改进的3个

模块对梯田识别的促进作用，具体结果见表3。可知：引入跳跃网络（编号Ⅰ）、CBAM（编号Ⅱ）和修改模型的锚框大小（编号Ⅲ）后，梯田识别结果的精确率分别提升了1.46%、1.02%和0.76%；

召回率分别提升了 3.91%、1.09% 和 0.70%；F1 提升了 2.55%、0.59% 和 0.25%；IoU 提升了 4.00%、0.91%、0.39%。整体而言，在本研究提出的 JAM-R-CNN 模型中，3 个模块全部改进的梯田识别结果优于改进某两个模块，再优于改进某 1 个模块，说明改进的 3 个模块对梯田的识别均有促进作用。

表 3 消融实验结果对比
Table 3 Comparison of the results of ablation study

编号	跳跃网络	CBAM	锚框大小	<i>P</i>	<i>R</i>	F1	IoU
I				88.85	79.02	85.69	77.96
II	✓			90.31	82.93	88.24	81.96
III		✓		89.87	80.11	86.28	78.87
IV			✓	89.61	79.72	85.94	78.35
V	✓	✓		90.62	83.50	88.66	82.63
VI	✓		✓	90.52	84.71	88.49	82.36
VII		✓	✓	90.49	81.28	86.99	79.98
VIII	✓	✓	✓	90.81	84.28	88.98	83.15

注：表中✓表示对应实验使用该模块。

图 8 直观展示了各消融实验结果的误差情况，可看出各个实验结果的错分和漏分情况。可见：加入跳跃网络后（图 8(d)），在梯田识别的误差结果中，每个梯田块外围的绿色漏分区域较不加入跳跃网络（图 8(c)）的识别结果少，跳跃网络识别的梯田较原始模型识别结果有一定的外扩，并且外扩的区域是实际的梯田区域，这是由于跳跃网络能够在模型运算过程中维持高分辨率影像的强语义信息，从而可以适当区分梯田块间的混合

像元，能够进一步捕捉梯田的边界信息；加入 CBAM 后（图 8(e)），在梯田识别的误差结果中，绿色的漏分区域整体比原始模型（图 8(c)）少，这是因为 CBAM 对梯田的多尺度特征表达能力的强化有助于更好的识别梯田信息，从而识别出部分原始模型未能识别的梯田信息；修改模型的锚框大小后（图 8(f)），对同一狭长梯田而言，识别的梯田比原始模型（图 8(c)）识别的梯田形状略长，并且这些多识别的区域在误差图中呈现白色（正确提取），在原始模型结果误差图中呈现绿色（漏分），说明修改模型的锚框大小后识别的梯田更加准确。因此，将原始模型的锚框大小依据真实标签数据中的梯田样本情况进行修改，使得模型在运算过程中产生的候选框能够更适应真实的梯田形状，进而提升识别精度。综上所述，加入跳跃网络、CBAM 和修改模型的原始锚框大小都能提升 Mask R-CNN 模型对梯田识别的精度。

此外，在本研究提出的 JAM-R-CNN 模型实验结果误差图（图 8(j)）中，有部分梯田块外围是红色的错分区域，这是由于在制作样本标签的过程中，部分标签未正确勾画出梯田的实际边界位置而造成的。但通过与实际情况进行对比，实验结果误差图中的红色区域实际是梯田，由于样本标注问题导致在计算实验结果精度时误认为是错误的提取，因而可能会造成最终计算的实验结果精度较实际精度略低的情况。因此，模型的识别结果虽然存在少部分的错分情况，但其最终识别的梯田边界位置会优于实际的样本，总体上 JAM-R-CNN 模型的识别效果较好。

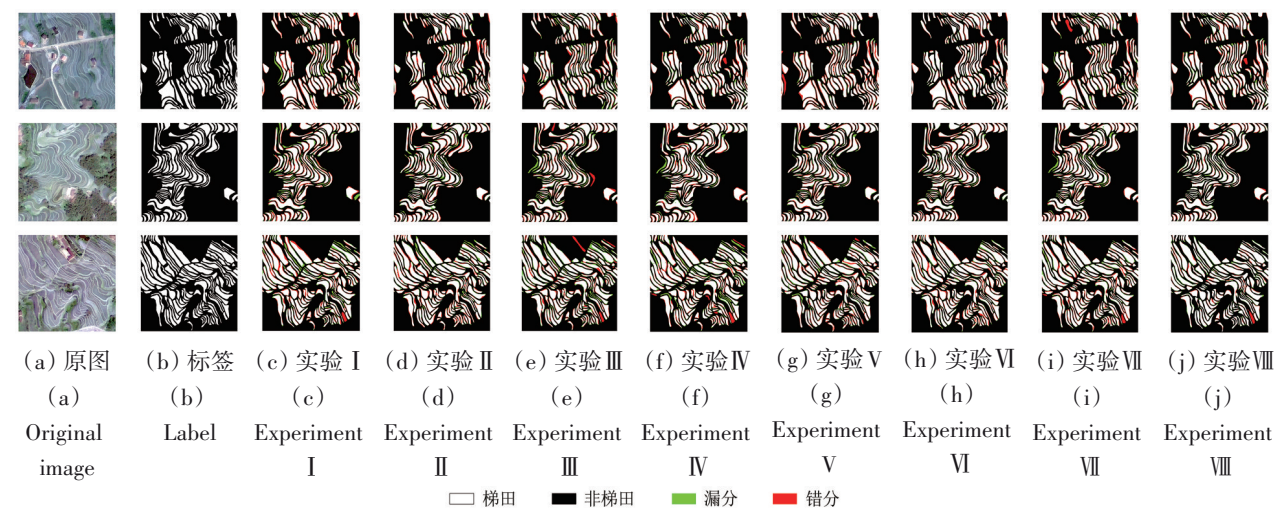


图 8 消融实验提取结果误差图
Fig. 8 Error diagram of ablation experiment

5 结 论

为实现梯田地块的高精度遥感识别, 准确地掌握梯田的空间分布, 本研究提出了一种面向梯田遥感识别的JAM-R-CNN深度网络模型, 该模型在Mask R-CNN基础上进行改进: 在模型的FPN模块中, 融入跳跃网络模块以维持高分辨率遥感影像在模型运算全过程中的高语义信息; 引入卷积注意力机制模块以加强梯田的特征表达能力; 在模型RPN模块中, 修改模型的锚框大小以适应梯田狭长的特性, 并利用GF-2遥感影像, 在中国西南丘陵地区的盐井梯田区域进行实验以验证改进模型对梯田识别的有效性, 主要结论如下: (1) 本研究提出的方法达到了较好的效果, 相较Mask R-CNN模型, JAM-R-CNN网络模型梯田识别结果精度明显提高, 精确率、召回率、F1和IoU依次提升了1.96%、5.26%、3.29%和5.19%, 具体的精确率为90.81%, 召回率为84.28%, F1为88.98%, IoU为83.15%; (2) 通过与Unet、DeepLab v3+和Mask R-CNN模型的对比, JAM-R-CNN模型比次优的Mask R-CNN模型精确率、召回率、F1和IoU依次提升1.96%、5.26%、3.29%和5.19%, 同时Unet和DeepLab v3+模型的梯田识别结果中有较多的粘连情况, 无法准确获取梯田信息, 证明了JAM-R-CNN模型不仅有效减少了提取梯田边界的粘连情况, 而且明显提高了狭长型梯田的提取率, 总体效果较好; (3) 消融实验结果表明了改进Mask R-CNN模型的3个模块均对梯田识别有促进作用, 进一步证明了本研究提出的JAM-R-CNN模型改进效果明显, 在梯田遥感识别方面展现出优越性能。

研究下一步将考虑结合梯田的一些基本浅层特征, 如地形等, 再引入“anchor free”机制提升模型锚框大小对梯田不规则轮廓的自适应能力, 从而加强模型在地物边界模糊的复杂背景下对梯田的识别能力。同时, 后续也会考虑结合梯田的多时相变化情况, 探讨时空谱特征对梯田遥感识别的作用, 以期进一步提升梯田识别精度。

参考文献(References)

Cai Z W, He Z, Wang W J, Yang J Y, Wei H D, Wang C and Xu B D. 2022. Mapping cropland at metric resolution using the spatiotem-

poral information from multi-source GF satellite data. *National Remote Sensing Bulletin*, 26(7): 1368-1382 (蔡志文, 何真, 王文静, 杨靖雅, 魏浩东, 王聪, 徐保东. 2022. 基于多源国产高分卫星时空信息的米级分辨率耕地提取. *遥感学报*, 26(7): 1368-1382) [DOI: 10.11834/jrs.20221069]

Chen L L, Shi Z, Liao K T, Song Y J and Zhang H M. 2022. Cultivated land extraction from high resolution remote sensing image based on convolutional neural network. *Transactions of the Chinese Society for Agricultural Machinery*, 53(9): 168-177 (陈玲玲, 施政, 廖凯涛, 宋月君, 张红梅. 2022. 基于卷积神经网络的高分遥感影像耕地提取研究. *农业机械学报*, 53(9): 168-177) [DOI: 10.6041/j.issn.1000-1298.2022.09.017]

Dai W, Na J M, Huang N, Hu G H, Yang X, Tang G A, Xiong L Y and Li F Y. 2020. Integrated edge detection and terrain analysis for agricultural terrace delineation from remote sensing images. *International Journal of Geographical Information Science*, 34(3): 484-503. [DOI: 10.1080/13658816.2019.1650363]

Dang T M, Mu X M, Sun W Y, Gao P and Zhao G J. 2017. Review of quickly discriminating approaches of terrace information based on high resolution remote sensing images. *Yellow River*, 39(3): 85-89, 94 (党恬敏, 穆兴民, 孙文义, 高鹏, 赵广举. 2017. 高分辨率遥感影像梯田快速提取方法研究进展. *人民黄河*, 39(3): 85-89, 94) [DOI: 10.3969/j.issn.1000-1379.2017.03.019]

Deng P F, Xu K J and Huang H. 2021. CNN-GCN-based dual-stream network for scene classification of remote sensing images. *National Remote Sensing Bulletin*, 25(11): 2270-2282 (邓培芳, 徐科杰, 黄鸿. 2021. 基于CNN-GCN双流网络的高分辨率遥感影像场景分类. *遥感学报*, 25(11): 2270-2282) [DOI: 10.11834/jrs.20210587]

Diaz-Varela R A, Zarco-Tejada P J, Angileri V and Loudjani P. 2014. Automatic identification of agricultural terraces through object-oriented analysis of very high resolution DSMs and multispectral imagery obtained from an unmanned aerial vehicle. *Journal of Environmental Management*, 134: 117-126. [DOI: 10.1016/j.jenvman.2014.01.006]

Dong S J, Xin L J, Li S F, Xie H L, Zhao Y L, Wang X, Li X B, Song H F and Lu Y H. 2023. The extent and spatial distribution of terrace abandonment in China. *Acta Geographica Sinica*, 78(1): 3-15 (董世杰, 辛良杰, 李升发, 谢花林, 赵宇鸾, 王学, 李秀彬, 宋恒飞, 卢亚哈. 2023. 中国梯田撂荒程度及空间格局分异研究. *地理学报*, 78(1): 3-15) [DOI: 10.11821/dlxb202301001]

Eckert S, Ghebremicael S T, Hurni H and Kohler T. 2017. Identification and classification of structural soil conservation measures based on very high resolution stereo satellite data. *Journal of Environmental Management*, 193: 592-606. [DOI: 10.1016/j.jenvman.2017.02.061]

He K M, Gkioxari G, Dollár P and Girshick R. 2017. Mask R-CNN// *International Conference on Computer Vision*. Venice: IEEE: 2980-2988. [DOI: 10.1109/ICCV.2017.322]

- Jong M, Guan K Y, Wang S B, Huang Y Z and Peng B. 2022. Improving field boundary delineation in ResUNets via adversarial deep learning. *International Journal of Applied Earth Observation and Geoinformation*, 112: 102877. [DOI: 10.1016/j.jag.2022.102877]
- Liu W, Wu Z F, Luo J C, Sun Y W, Wu T J, Zhou N, Hu X D, Wang L Y and Zhou Z F. 2021. A divided and stratified extraction method of high-resolution remote sensing information for cropland in hilly and mountainous areas based on deep learning. *Acta Geodae-tica et Cartographica Sinica*, 50(1): 105-116 (刘巍, 吴志峰, 骆剑承, 孙营伟, 吴田军, 周楠, 胡晓东, 王玲玉, 周忠发. 2021. 深度学习支持下的丘陵山区耕地高分辨率遥感信息分区分层提取方法. *测绘学报*, 50(1): 105-116) [DOI: 10.11947/j.AGCS.2021.20190448]
- Liu X Y, Yang S T, Wang F G, He X Z, Ma H B and Luo Y. 2014. Analysis on sediment yield reduced by current terrace and shrubs-herbs-arbor vegetation in the Loess Plateau. *Journal of Hydraulic Engineering*, 45(11): 1293-1300 (刘晓燕, 杨胜天, 王富贵, 何兴照, 马红斌, 罗娅. 2014. 黄土高原现状梯田和林草植被的减沙作用分析. *水利学报*, 45(11): 1293-1300) [DOI: 10.13243/j.cnki.slxb.2014.11.004]
- Long J, Li M M, Wang X Q and Stein A. 2022. Delineation of agricultural fields using multi-task BsiNet from high-resolution satellite images. *International Journal of Applied Earth Observation and Geoinformation*, 112: 102871. [DOI: 10.1016/j.jag.2022.102871]
- Luo L H, Li F Y, Dai Z Y, Yang X, Liu W and Fang X. 2020. Terrace extraction based on remote sensing images and digital elevation model in the loess plateau, China. *Earth Science Informatics*, 13(2): 433-446. [DOI: 10.1007/s12145-020-00444-x]
- Mei W Y, Wang H Y, Fouhey D, Zhou W Q, Hinks I, Gray J M, Van Berkel D and Jain M. 2022. Using deep learning and very-high-resolution imagery to map smallholder field boundaries. *Remote Sensing*, 14(13): 3046. [DOI: 10.3390/rs14133046]
- Sun W Y, Zhang Y Q, Mu X M, Li J Y, Gao P, Zhao G J, Dang T M and Chiew F. 2019. Identifying terraces in the hilly and gully regions of the Loess Plateau in China. *Land Degradation & Development*, 30(17): 2126-2138. [DOI: 10.1002/ldr.3405]
- Waldner F and Diakogiannis F I. 2020. Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. *Remote Sensing of Environment*, 245: 111741. [DOI: 10.1016/j.rse.2020.111741]
- Wang J D, Sun K, Cheng T H, Jiang B R, Deng C R, Zhao Y, Liu D, Mu Y D, Tan M K, Wang X G, Liu W Y and Xiao B. 2021. Deep high-resolution representation learning for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10): 3349-3364. [DOI: 10.1109/TPAMI.2020.2983686]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//15th European Conference on Computer Vision. Munich: Springer: 3-19. [DOI: 10.1007/978-3-030-01234-2_1]
- Xu Z Y, Shen Z F, Li Y, Zhao L F, Ke Y M, Li L L and Wen Q. 2022. Classification of high-resolution remote sensing images based on enhanced DeepLab algorithm and adaptive loss function. *National Remote Sensing Bulletin*, 26(2): 406-415 (许泽宇, 沈占锋, 李杨, 赵丽芳, 柯映明, 李苓苓, 温奇. 2022. 增强型DeepLab算法和自适应损失函数的高分辨率遥感影像分类. *遥感学报*, 26(2): 406-415) [DOI: 10.11834/jrs.20209200]
- Xue M D, Zhang H M, Yang J T and Li X S. 2019. Extraction of terrace information based on UAV image and topographic index. *Application Research of Computers*, 36(8): 2527-2533, 2538 (薛牡丹, 张宏鸣, 杨江涛, 李星恕. 2019. 无人机影像与地形指数结合的梯田信息提取. *计算机应用研究*, 36(8): 2527-2533, 2538) [DOI: 10.19734/j.issn.1001-3695.2018.01.0182]
- Yu M G, Rui X P, Xie W Y, Xu X J and Wei W. 2022. Research on automatic identification method of terraces on the loess plateau based on deep transfer learning. *Remote Sensing*, 14(10): 2446. [DOI: 10.3390/rs14102446]
- Yuan S, Tang H and Guo Y. 2022. Segmentation method for plant leaves using an improved Mask R-CNN model. *Transactions of the Chinese Society of Agricultural Engineering*, 38(1): 212-220 (袁山, 汤浩, 郭亚. 2022. 基于改进Mask R-CNN模型的植物叶片分割方法. *农业工程学报*, 38(1): 212-220) [DOI: 10.11975/j.issn.1002-6819.2022.01.024]
- Zhao F, Xiong L Y, Wang C, Wang H R, Wei H and Tang G A. 2021. Terraces mapping by using deep learning approach from remote sensing images and digital elevation models. *Transactions in GIS*, 25(5): 2438-2454. [DOI: 10.1111/tgis.12824]
- Zhao X, Wang X J, Zhao Y, Luo Z D, Xu Y L, Guo H and Zhang Y. 2016. Feasibility analysis on methodology of terraced extraction by using Fourier transformation based on domestic hi-resolution remote sensing images of GF-1 satellite. *Soil and Water Conservation in China*, (1): 63-65 (赵欣, 王晓晶, 赵院, 罗志东, 许永利, 郭浩, 张羿. 2016. 国产高分一号卫星数据傅里叶变换提取梯田影像可行性分析. *中国水土保持*, (1): 63-65) [DOI: 10.14123/j.cnki.swcc.2016.0050]
- Zhou N, Yang P, Wei C S, Shen Z F, Yu J J, Ma X Y and Luo J C. 2021. Accurate extraction method for cropland in mountainous areas based on field parcel. *Transactions of the Chinese Society of Agricultural Engineering*, 37(19): 260-266 (周楠, 杨鹏, 魏春山, 沈占锋, 余娟娟, 马晓宇, 骆剑承. 2021. 地块尺度的山区耕地精准提取方法. *农业工程学报*, 37(19): 260-266) [DOI: 10.11975/j.issn.1002-6819.2021.19.030]
- Zhou Y, Li M M, Wang X Q, Wu S Y and Jin S L. 2022. Extraction of farming terraces using object-based convolutional neural networks from very high resolution satellite images. *Remote Sensing Information*, 37(2): 138-144 (周珏, 李蒙蒙, 汪小钦, 吴思颖, 金时来. 2022. 面向对象卷积神经网络的耕作梯田提取. *遥感信息*, 37(2): 138-144) [DOI: 10.3969/j.issn.1000-3177.2022.02.019]

JAM-R-CNN deep learning network model for remote sensing recognition of terraced fields

XIE Junyang^{1,2}, LIN Anqi^{1,2}, WU Hao^{1,2}, WU Ziwei^{1,2}, WU Wenbin³, YU Qiangyi³

1.College of Urban and Environmental Sciences, Central China Normal University, Wuhan 430079, China;

2.Hubei Province Key Laboratory for Geographical Process Analysis and Simulation, Wuhan 430079, China;

3.Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract: Efficiently and accurately determining the spatial distribution of terraced fields provides important data support for soil and water conservation and improves the regulatory level of agriculture in mountainous areas. When deep learning methods are used for terrace recognition, narrow and elongated terraces are prone to be missed because of convolution operations, and in complex backgrounds with unclear terrain boundaries in mountainous areas, large areas of adhesive recognition results are easily generated, leading to low accuracy in the final terrace recognition. Prior to the achievement of accurate recognition of terrace information, the urgent technical problems to be solved are how to effectively maintain the high semantic information of high-resolution remote sensing images in the convolution operation process on the basis of the characteristics of terraces and how to reduce the omission of narrow and long terraces and the adhesion of recognition results. To address these problems, this study proposes the JAM-R-CNN deep learning network terrace recognition method that adopts remote sensing images with very high resolution. This network is based on the Mask Region-based Convolutional Neural Network (Mask R-CNN) model. It integrates the jumping network to maintain the high semantic information of high-resolution remote sensing images, employs the convolutional block attention module to enhance the feature expression ability of terraces, and modifies the anchor size to adapt to the narrow and long characteristics of terraces and improve terrace recognition accuracy. A part of the salt well terraces in Nanchuan District, Chongqing, China, is selected as the study area to test the proposed method, and four models in domestic GF-2 satellite image data are used for experiments. Results show that the terrace parcel map derived from the JAM-R-CNN model has a precision of 90.81%, recall of 84.28%, F1 score of 88.98%, and Intersection over Union (IoU) value of 83.15%. Compared with Mask R-CNN, JAM-R-CNN's precision, recall, F1 score, and IoU value are increased by 1.96%, 5.26%, 3.29%, and 5.19%, respectively, indicating that the JAM-R-CNN model can better identify the terraces than Mask R-CNN can. Most of the terraces identified by Unet and DeepLab v3+ are connected together, and the terraced fields with a small size are not distinguished. The JAM-R-CNN model identifies fewer missing areas on the periphery of terraces compared with Mask R-CNN model, and the number of missing narrow and long terraces is considerably reduced. This result is the effect of three improved parts and further proves that the proposed JAM-R-CNN model exerts a remarkable improvement effect and demonstrates superior performance in remote sensing recognition of terraces. The proposed JAM-R-CNN deep learning network model effectively reduces the adhesion phenomenon of terrace recognition results and considerably improves the extraction rate for narrow and long terraces, thus achieving a substantial improvement in the overall accuracy of terrace remote sensing recognition. The model has good application value.

Key words: remote sensing, terrace recognition, high resolution remote sensing image, deep learning, jump network, JAM-R-CNN

Supported by National Key Research and Development Program of China (No. 2022YFB3903500); Hubei Provincial Natural Science Foundation of China (No. 2024AFA032); National Natural Science Foundation of China (No. 42071358, 42201468); Fundamental Research Funds for the Central Universities (No. CCNU22QN018); Scientific Research Project of Hubei Natural Resources Department (No. ZRZY2023KJ03)