

doi: 10.6046/zrzyyg.2022158

引用格式: 刁明光,刘勇,郭宁博,等. 基于 Mask R - CNN 的遥感影像疏林地智能识别方法[J]. 自然资源遥感,2023,35(2):97 - 104. (Diao M G, Liu Y, Guo N B, et al. Mask R - CNN - based intelligent identification of sparse woods from remote sensing images [J]. Remote Sensing for Natural Resources, 2023, 35(2): 97 - 104.)

# 基于 Mask R - CNN 的遥感影像疏林地智能识别方法

刁明光<sup>1</sup>, 刘勇<sup>1</sup>, 郭宁博<sup>1</sup>, 李文吉<sup>2</sup>, 江继康<sup>1</sup>, 王云霄<sup>1</sup>

(1. 中国地质大学(北京)信息工程学院, 北京 100083; 2. 中国自然资源航空物探遥感中心, 北京 100083)

**摘要:** 针对遥感影像疏林地提取方法少且精度不高, 缺乏智能识别的数据集情况, 提出了一套遥感影像疏林地智能识别方法。分别使用 QGIS 插件和 Python 语言对该方法进行实现, 完成了数据集制作的环节, 为模型训练提供了数据支撑; 通过特征提取生成特征图, 在特征图中提取感兴趣区域(region of interest, ROI), 通过池化操作(ROI align)对这些感兴趣区域进行过滤操作, 减少因疏林地图像感兴趣区域过多而造成的内存消耗。实验表明, 该方法可快速进行数据集制作, 有效辅助遥感影像中疏林地的识别, 使用基于 Mask R - CNN 的遥感影像疏林地智能识别方法对疏林地目标检测的均值平均准确率可以达到 0.92。

**关键词:** 遥感影像; 疏林地; 深度学习; Mask R - CNN; QGIS

**中图法分类号:** TP 751 **文献标志码:** A **文章编号:** 2097 - 034X(2023)02 - 0097 - 08

## 0 引言

在遥感影像中, 疏林地是一种分布不均匀且背景复杂的地物, 同时也是自然资源调查中的重点识别对象<sup>[1]</sup>。疏林地具有分布连续、面积大小不一, 分布形状不规则等特点, 其周围可能是耕地、水体或建筑用地。疏林地的这些特征, 增加了在遥感影像中对其识别的难度。同时, 用于自动识别的遥感影像疏林地数据集十分缺乏, 其标注质量也难以得到保证。目前对遥感影像中疏林地进行识别存在的问题主要包括:

1) 缺乏用于自动识别的遥感影像疏林地目标检测数据集。目标检测数据集与深度学习密不可分。目前, 使用比较广泛的 3 种自然场景图像目标检测数据集有 PASCAL VOC<sup>[1]</sup>, MSCOCO<sup>[2]</sup>, ImageNet<sup>[3]</sup> 和 DOTA<sup>[4]</sup>。这些数据集的特征在于偏向自然场景下的目标识别, 并非专门针对遥感影像进行设计, 而且其中不包含疏林地识别目标, 因此不适用于遥感影像疏林地目标的识别任务。

2) 遥感影像疏林地数据集标注质量难以保证。遥感影像疏林地数据集的标注质量直接影响训练的

结果。常见数据集标注方法多采用人工标注方法, 需要专业人员具备丰富的知识储备。手工标注过程中主观因素影响较大<sup>[5-8]</sup>, 质量和效率难以得到有效保证。对于形状多样的疏林地来讲, 现有规则边缘目标识别特征提取方法<sup>[9-11]</sup>适用性不强, 这些因素导致遥感影像疏林地数据集制作和识别困难。

模型的训练需要数据集支撑, 同时算法的好坏也都需要在数据集上得到验证。本文根据遥感影像中疏林地特征, 使用 Mask R - CNN 算法模型对疏林地进行目标检测和提取。针对遥感影像疏林地数据集匮乏以及遥感影像疏林地识别中存在的问题, 研究遥感影像疏林地数据集自动生成方法<sup>[11-12]</sup>, 对遥感影像疏林地数据集特征进行分析, 提出并实现遥感影像疏林地智能识别模型及智能识别方法, 应用于遥感影像中疏林地的识别。

## 1 遥感影像疏林地数据集自动生成方法

数据是深度学习的基础, 本文提出遥感影像疏林地数据集自动生成方法并以此制作后续实验需要的数据。遥感影像疏林地数据集的制作参照 LabelMe 来进行。数据集的制作流程分为数据加载、

收稿日期: 2022 - 04 - 22; 修订日期: 2022 - 05 - 16

基金项目: 国家重点研发计划项目子课题“高亚洲和北极积雪 - 冰川与地质灾害数据可视化与算法集成研究”(编号: 2021YFE0116807)、中国地质调查局地质调查项目“全国矿山开发状况遥感地质调查与监测”(编号: DD20190511)和中国自然资源部航空物探与遥感地质重点实验室课题“遥感影像典型矿山地物(露天采场)智能识别的研究与应用”(编号: 2020YF23)共同资助。

第一作者: 刁明光(1970 - ), 男, 博士, 副教授, 主要从事地学信息工程、软件工程研究。Email: dmg@cugb.edu.cn.

影像裁剪、标注文件生成和数据集组织结构调整 4 个过程,最终生成规范的遥感影像疏林地数据集,如图 1 所示。数据加载主要利用遥感影像疏林地数据集自动生成工具,添加遥感影像数据和相应的矢量

数据,读取矢量信息。影像裁剪主要是通过遥感影像疏林地数据及自动制作工具进行栅格数据的裁剪工作。标注文件生成主要包括 JSON 数据生成和 JSON 文件解析 2 个步骤。

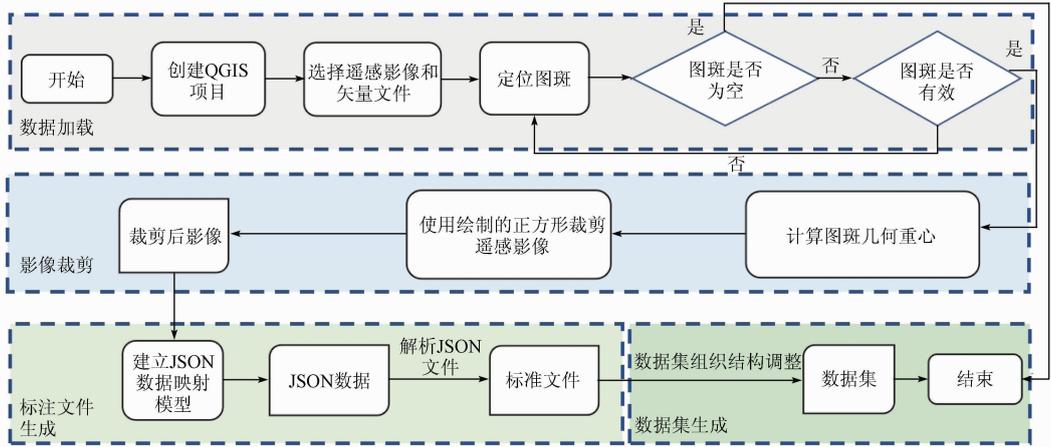


图 1 遥感影像疏林地数据集生成流程

Fig. 1 Remote sensing image sparse woods dataset generation process

### 1.1 遥感影像疏林地数据集的自动生成工具

遥感影像疏林地数据集制作工具基于 QGIS 插件<sup>[13]</sup>,采用 Python 语言实现。遥感影像疏林地数据集制作工具主要功能在于读取已有的影像和矢量文件、读取影像和矢量中的信息、完成影像裁剪和标注文件生成部分。在数据集工具的开发过程中,需要对输入数据的正确性和参数的完整性进行严格检查。根据建立的数据映射模型可以保证矢量信息与 JSON 数据转化的准确性。

### 1.2 遥感影像疏林地数据集原始图像的自动分割

遥感影像裁剪的流程主要包括数据加载、图斑信息读取、矩形绘制和图像裁剪 4 步。图斑信息读取是获取图斑的属性和地理坐标信息。从加载的矢量数据中读取图斑的属性表,之后根据图斑的编号进行遍历,在遍历的过程中,获取并保存图斑的几何重心坐标。由于疏林地类型的图斑都是不规则形状,因此采用求多边形中心的思路来计算图斑的几何重心:首先把多边形划分成多个三角形,求每个三角形的重心和面积;然后将二维多边形的面积作为权值,分别乘以对应重心坐标的 X 和 Y 值;最后将求出的 X 和 Y 值的加权平均数除以总面积,就可以得到多边形的重心坐标。

由于本文采用的遥感影像在尺寸上超出了深度学习模型能接收的图像大小,需要对原图进行裁剪。根据模型的需要,以图斑的几何重心坐标为中心点,绘制边长为 1 024 像素的正方形,以绘制的正方形为掩模边界,在像素上对正方形范围内的遥感影像文件进行裁剪。经过以上步骤,即可实现遥感影像

疏林地数据集中原始图像的生成。

### 1.3 遥感影像疏林地数据集标注文件的自动生成

标注文件的生成是制作遥感影像疏林地数据集的重要阶段。生成遥感影像疏林地标注文件的关键是建立矢量数据和 JSON 数据之间的映射关系,如表 1 所示。

表 1 JSON 数据与矢量数据映射关系

Tab. 1 Mapping relationship of JSON data and vector data

键名	值	说明
version	LabelMe 的版本号	可自动获取
flags	标志	可为空
shape_label	标签名,通常为类别名 + 编号	获取矢量信息中的类别,同一类别累加计数
shape_line_color	线条颜色	可为空
shape_fill_color	填充颜色	可为空
shape_points	掩模各个顶点的相对坐标	计算图斑上各点与矩形框左上角的坐标差
shape_type	掩模形状类型	point (点)、line (线)、polygon (多边形)
lineColor	线条颜色	可为空
fillColor	填充颜色	可为空
imagePath	原图路径	裁剪后影像的路径
imageData	64 位值	图像转换
imageHeight	原始图像的高度	图像信息中获取
imageWidth	原始图像的宽度	图像信息中获取

本文分析 LabelMe 的标注方式,提取了生成 JSON 文件中标准的几何和类别信息,在编程中通常将其理解为对象,JSON 对象是一个无序的“key - value”对集合: { "key1" : obj, "key2" : obj, "key3" : obj, ... } 其中, key 值必须为字符串类型 (string), 而 value 则可以为多种类型,且可以嵌套。

由于采用对象结构,只要建立矢量数据的属性信息和 LabelMe 生成的 JSON 文件中对应键值之间的映射模型,就可以完成根据矢量数据生成 JSON 文件的任务。

### 1.4 数据集自动生成结果

数据集生成结果自动存放在 cv2\_mask, JSON, labelme\_JSON 和 pic 等 4 个文件夹中。其中, cv2\_mask 文件夹存放掩模文件; JSON 文件夹存放 JSON 文件; labelme\_JSON 文件夹存放解析 JSON 文件后生成的文件,其中每个子文件夹内包含 5 个文件:原始影像文件 img. png、标签类别信息文件 info. yaml、掩模文件 label. png、当前影像内的对应掩模类别文件 label\_names. txt、原图与掩模文件的可视化关系图文件 label\_viz. png; pic 文件夹用来存放裁剪后的原始图像。数据集的自动生成结果验证了遥感影像疏林地数据集生成方法的可行性。

本文对提出的遥感影像疏林地数据集自动生成方法生成数据集的效率与其他生成方法进行对比,如表 2 所示,在制作相同数据集的情况下,本文遥感影像疏林地数据集自动生成方法生成数据集的速度远远高于人工制作,比矿山遥感监测实例分割数据

集自动生成方法<sup>[12]</sup>也要高出 12.8 倍,由此可以证明该遥感影像疏林地数据集自动生成方法的可行性与优越性。

表 2 不同数据集生成方法性能对比

Tab.2 Performance comparison of different dataset generation methods

数据集生成方法	数据源	数据量/ 个	耗时/ h
人工制作方法	自然资源林地遥感影像	9 156	1 008
矿山遥感监测实例分割数据集自动生成方法	自然资源林地遥感影像	9 156	96
遥感影像疏林地数据集自动生成方法	自然资源林地遥感影像	9 156	7.5

## 2 遥感影像疏林地智能识别模型

使用深度学习方法对遥感影像疏林地的智能识别,实质上是对遥感影像疏林地进行目标检测,并实现对影像的逐像素预测。因此,整个智能识别过程本质就是实现实例分割的过程<sup>[14-15]</sup>。遥感影像疏林地智能识别流程如图 2 所示。

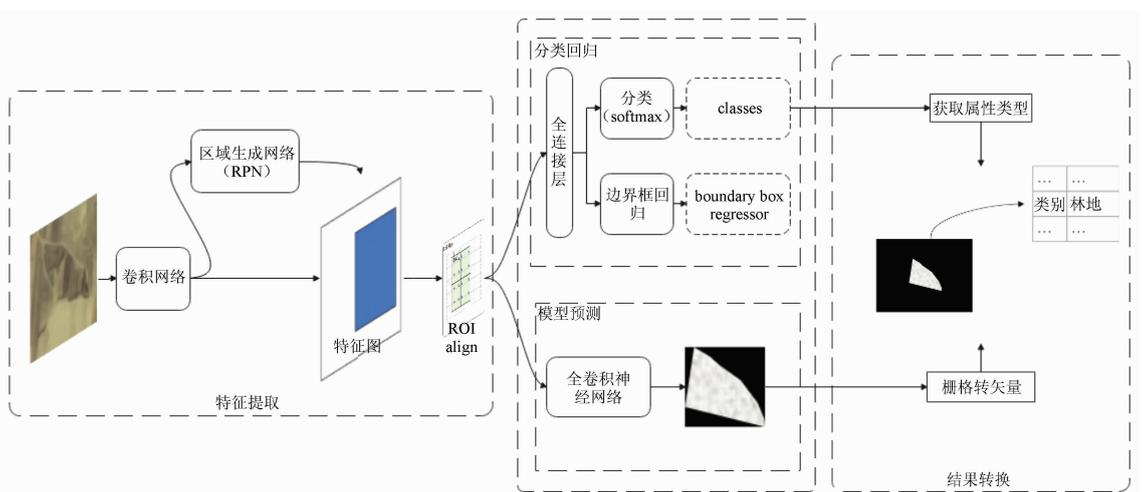


图 2 遥感影像疏林地智能识别模型框架

Fig.2 Framework for intelligent recognition model of remote sensing image sparse woods

首先通过特征提取生成特征图,然后在特征图中提取感兴趣区域(region of interest, ROI),再采用分类和回归的方法对这些候选 ROI 进行过滤操作,随后对这些 ROI 进行池化操作,最后进行分类、回归和分割掩模生成。其基本流程分为以下 5 个步骤进行: ①特征提取,首先将预处理后的图像输入到特征金字塔模型提取网络中,并将得到的特征图作为本模块的输出; ②池化操作,对上一步获得的特征图进行池化操作(ROI align); ③分类回归,在全连接层上对 ROI 进行分类和边界框回归操作,完成

定位和分类任务; ④模型预测,使用另外一支全卷积神经网络(fully convolutional network, FCN)<sup>[16]</sup>进行像素级分割,最终在每一个目标上生成对应的二值蒙版数据; ⑤结果转换,将识别的目标类别转换为矢量属性,分析并设计栅格数据转矢量数据的方法,为二值蒙版图转换为矢量图斑提供方法依据。

### 2.1 特征提取

遥感影像疏林地智能识别方法采用特征金字塔网络,将图像输入后,通过特征提取网络即可得到特征图。

1) 自下而上的路径。以 ResNet<sup>[17]</sup> 为骨干网络, 根据特征图的尺寸分为 5 个阶段, 其中由于第 1 阶段需要耗费大量内存不进行使用, 第 2—5 阶段分别为 {C2, C3, C4, C5}, 相对于原始图像的步长分别为 {4, 8, 16, 32}, 每个阶段的残差模块都将作为下一阶段的输入。

2) 自上而下的路径和从左到右的侧向连接。自上而下的路径是从最高层开始进行下采样, 这里下采样采用的是最邻近下采样。从左到右的侧向连接将下采样的结果和自下而上路径中生成的特征图进行融合, 即 {C2, C3, C4, C5} 中的每一层都使用 1×1 的卷积(Conv)操作后和下采样生成的特征图进行相加操作, 它们具有相同的尺寸; 在融合后再使用 3×3 的卷积对融合特征进行处理(消除上采样产生的混叠效应)。

3) 最终生成了特征图集合 {P2, P3, P4, P5, P6}。网络结构如图 3 所示。

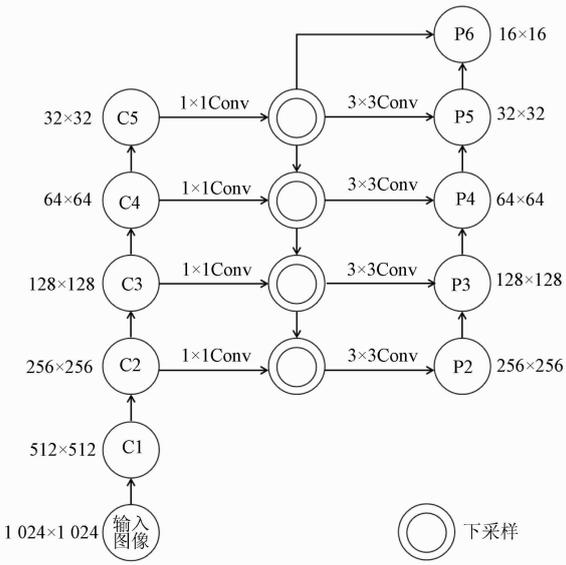


图 3 ResNet - FPN 网络结构

Fig. 3 ResNet - FPN network architecture

### 2.2 特征图的池化操作

一般的池化方法依然存在着一一些问题, 比如在计算时存在小数的情况, 此时就需要进行取整操作, 但是取整就容易造成误差。ROI align<sup>[18]</sup> 方法提出了不对池化值进行取整, 如果在计算中得到小数, 也就是没有落到真实的像素上, 那么就用最近的像素对得到的虚拟的像素进行双线性插值, 最终得到这个像素的值。

使用双线性插值法是为了降低模型在处理小目标上的误差, 提升小目标的识别精度。双线性插值计算过程见图 4。以 3×3 的池化区域 ROI align 的具体实现过程为例, 对遥感影像疏林地智能识别模

型池化的步骤进行说明: ①将边界框区域按输出要求的尺寸进行 9 等分, 等分后各顶点不会全部落到真实的像素点上; ②在每个区域内进行采样, 选择 2 位采样系数, 将每个等分区域划分出 2×2=4 个部分, 并将每个区域的中心点作为采样点, 如图 4 中的黑色小点 Q; ③对每一个黑色小点进行双线性插值计算; ④经过计算后, 在 4 个黑色小点中选用最大的值作为该区域的池化结果。

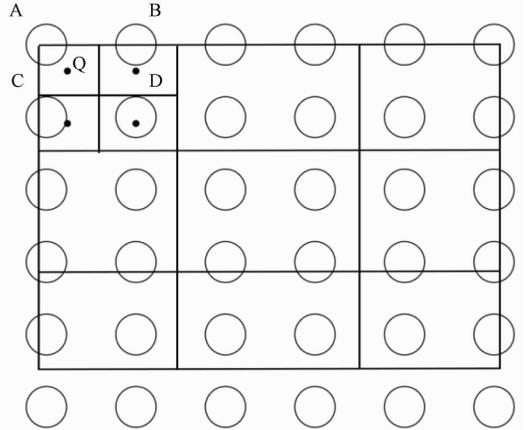


图 4 双线性插值计算过程

Fig. 4 Bilinear interpolation calculation process

### 2.3 分类回归及模型预测

遥感影像疏林地智能识别模型的预测主要包括 3 个部分: 分类、定位和分割。它主要利用 Mask R - CNN 网络定位原始图像中的 ROI(即训练结果所考虑的目标), 标记分类标签, 分割二值蒙版, 如图 5 所示, 其中“LD”是通过预测生成的分类标签; 虚线边框是用于定位的回归框; 绿色掩模部分是通过预测生成的用于分割的二值蒙版。在图像中生成二值蒙版可以有效地识别目标区域, 并为矢量图斑的生成提供数据。



图 5 遥感影像疏林地智能识别模型结果转换效果

Fig. 5 Conversion effect of remote sensing image sparse woods intelligent recognition model result

### 2.4 遥感影像疏林地识别结果转换

为了实现对遥感影像疏林地智能识别结果的编辑,结合常见的编辑格式以及遥感数据处理的方法,对识别结果的转换方法进行了分析研究。

结合常用的栅格转矢量的步骤,设计遥感影像疏林地智能识别模型识别结果转换的流程:①数据输入,即将原始影像和二值蒙版输入,读取原始影像中的地理坐标信息;②创建相同坐标系下的矢量文件,用来存储图斑数据,并对矢量创建属性“类别”,记录识别结果的标签;③读取二值蒙版图像,获取二值蒙版内各个顶点的相对坐标,使用原始图像的仿射信息进行像素坐标到地理坐标的转换,完成提取边界线的步骤;④利用转换后的坐标,生成多边

形,并添加到矢量文件中,根据识别结果显示的标签为“类别”赋值。通过以上 4 个步骤实现对所识别结果的转换。

## 3 遥感影像疏林地智能识别实验

遥感影像疏林地智能识别方法流程如图 6 所示,该流程从数据集制作开始,以输出矢量数据为结束,可以分为 4 个步骤:数据集制作、模型训练、模型测试、结果转换。其中数据集制作使用的是第 1 章介绍的遥感影像疏林地数据集自动生成方法,模型训练及测试使用的是第 2 章介绍的遥感影像疏林地智能识别模型。

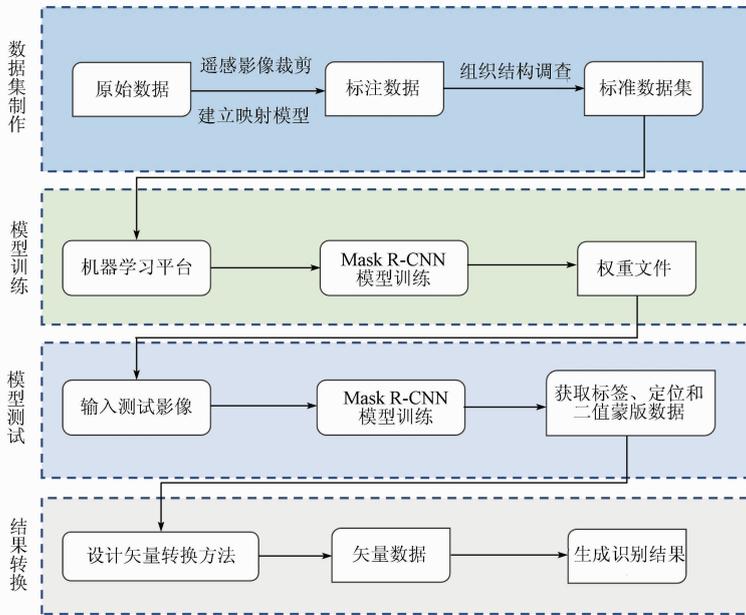


图 6 遥感影像疏林地智能识别方法框架

Fig. 6 Framework of intelligent identification method for sparse woods from remote sensing image

### 3.1 数据集介绍

本文中选用的数据来自于雄安新区高空间分辨率卫星遥感数据,如图 7 所示,为经过专业人员解译的包含疏林地的遥感影像及其矢量数据<sup>[19]</sup>,本文将利用此类遥感影像使用本文提出的遥感影像疏林地数据集自动生成方法制作数据集。由于 Mask R-CNN 实例分割模型对输入数据的尺寸有一定的要求,必须是 64 的倍数,同时考虑到选用的影像中疏林地的特征并不具有一致性,且形状大小不一,尤其是有呈带状的样本存在,该数据集在制作过程中将样本尺寸设置为 1 024 × 1 024 像素,将已经由专业人员“解译”的遥感影像数据和矢量文件放入数据集制作工具,数据集的制作依赖于遥感影像疏林地数据集自动生成工具,最终获得的数据集为 9 156 个样本数据,一个类别,即

疏林地。



图 7 疏林地遥感数据局部影像和矢量数据  
Fig. 7 Remote sensing data local image and vector data of sparse woods

### 3.2 实验环境

遥感影像疏林地智能识别模型的运行需要有一定的硬件和软件支持,本文中关于模型的相关流程都是在 Centos7 操作系统下进行的,模型的训练和测试的硬件环境为:1 个 Tesla GPU、16 G 显存、10 个 Intel CPU 和 40 G 内存。软件环境主要包括使用的编译器、语言和主要的函数库。根据对各种数据划分方法的分析结果,在结合已有的遥感影像疏林地数据集,本文中采用留出法,由于硬件的限制,将从已有的 9 156 个样本中筛选出 1 000 个样本用来划分训练集和验证集,测试集则从未参与训练和验证的剩余数据中选择;除去数据集之外,学习率和迭代次数等参数的设置也会对训练结果产生影响,在训练中设置初始的学习率为 0.01。

### 3.3 评价指标

基于 Mask R - CNN 的遥感影像疏林地智能识别模型的识别能力通过均值平均精确度(mean average precision, MAP)来进行模型总体评价。MAP 的值就是计算多个验证样本的平均精确度。计算 MAP 所需的准确率  $A$ 、精确率  $P$ 、召回率  $R$  都需要用到混淆矩阵。其中,准确率表示预测为疏林地的样本占总样本的百分比,精确率表示被预测为疏林地的样本中实际为疏林地样本的百分比,召回率表示实际为疏林地的样本被预测为疏林地的百分比,计算公式分别为:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

式中:  $TP$  为正类判定为正类,即识别中疏林地被识别为疏林地的样本数;  $FP$  为负类判定为正类,即识别中背景类判定为疏林地的样本数;  $FN$  为正类判定为负类,即应为疏林地的要素未被识别出来的样本数;  $TN$  为负类判定为负类,即背景类被识别为背景类的样本数。

## 4 结果与分析

所设计的模型性能识别能力可以由实验中得到的损失函数表示,如图 8 所示。从图 8 中可以看出其损失值( $epoch\_loss$ )随着  $epoch$  次数( $epoch\_times$ )逐渐降低,最终基本稳定在 0.5 左右,因此推断训练结果与  $epoch$  次数有一定的关系。选用  $epoch\_times$  为 10 处的训练权重文件对测试数据集进行识别测试,其结果显示如图 9 所示,图 9(a)和(b)分别为根据遥感影像疏林地数据集自动生成方法裁剪出的原始图像和掩模数据,图 9(c)为最终的识别结果,图中颜色遮掩部分为识别的二值蒙版图,外边框是目标检测结果,一个边框代表着识别出一块疏林地。虽然识别的结果存在误差,但也体现了训练结果的泛化能力。

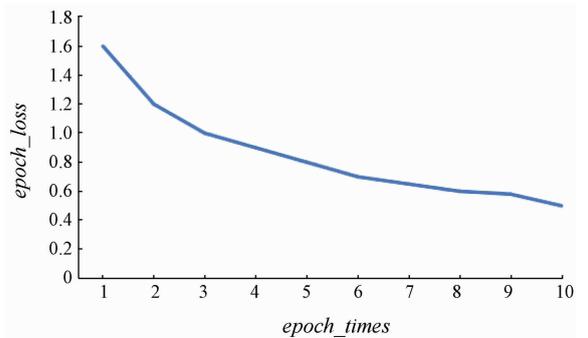
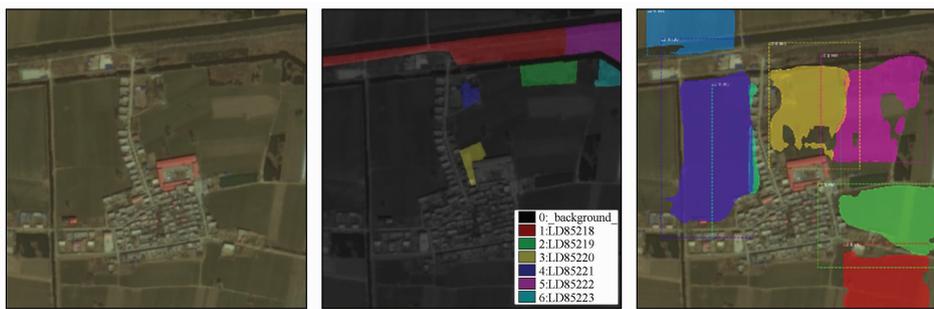


图 8 损失函数曲线  
Fig.8 Loss function graph



(a) 原始图像 (b) 掩模数据 (c) 识别结果

图 9 实验识别结果

Fig.9 Experimental recognition results

对上述实验选用  $epoch\_times$  为 10 处的训练权重文件,且在同样 10 个验证集进行验证,其验证集

精确率召回曲线如图 10 所示,验证集中的部分识别结果如图 11 所示,经过计算其 MAP 最高可达 0.92。

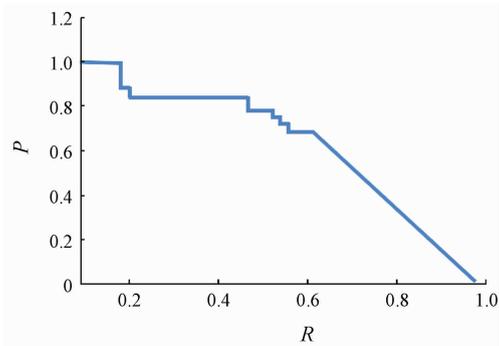


图 10 实验验证集精确率召回曲线

Fig. 10 Experimental verification set accurate recall curve



(a) 验证集原始图像

(b) 识别结果

图 11 验证集实验结果

Fig. 11 Validation set experimental results

实验中呈现的结果证明遥感影像疏林地智能识别模型具有一定的泛化能力。但是由于在实验中未进行严格的变量控制,并不能证明数据集的数量对模型性能的影响,但是可以从模型训练过程中损失函数的变化曲线可以证明损失函数随着迭代次数逐步降低,当迭代次数达到一定量时,损失函数可以达到 1.0 以下,模型的识别能力也随之增强。

## 5 结论

在已有的遥感影像疏林地数据的基础上,根据遥感影像疏林地数据集自动制作方法,生成遥感影像疏林地数据集,极大地减少了数据集制作的时间,降低了数据集制作成本。本文提出的遥感影像疏林地智能识别方法对疏林地进行了测试,通过基于 Mask R-CNN 的遥感影像疏林地智能识别模型,对遥感影像疏林地数据集进行网络智能识别,实验结果表明:针对遥感影像疏林地数据,本文构建的遥感影像疏林地智能识别方法可以进行良好的识别,证明了遥感影像疏林地智能识别方法的可行性。尽管本文方法对遥感影像疏林地识别效果较好,但仍存在以下问题:

1) 在数据预处理方面,本文采用的是第一季度的疏林地遥感数据,影像中一些农田和疏林地的边

界模糊,导致识别结果中存在边界不完整的情况,因此在数据预处理方面还需要进一步增强。

2) 相比于其他自然场景下要素的识别,遥感影像疏林地的识别精度相对较低,尤其是一些带状的疏林地影像,识别结果存在破碎的情况,如何提高遥感影像疏林地的识别精度还有待进一步的研究。

## 参考文献 (References):

- [1] Everingham M, Eslami S M A, Van Gool L, et al. The pascal visual object classes challenge: A retrospective [J]. International Journal of Computer Vision, 2015, 111(1): 98 - 136.
- [2] Lin T, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context [J]. Lecture Notes in Computer Science, 2014, 8693(1): 740 - 755.
- [3] Alex K, Ilya S, Geoffrey E H. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84 - 90.
- [4] Xia G S, Bai X, Ding J, et al. DOTA: A large - scale dataset for object detection in aerial images [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 3974 - 3983.
- [5] 刘立, 彭刚. 高分辨率遥感影像线状地物提取方法研究 [J]. 智能城市, 2021, 7(6): 63 - 64.
- [6] Liu L, Peng G. Research on the extraction method of line features from high - resolution remote sensing images [J]. Intelligent City, 2021, 7(6): 63 - 64.
- [7] 林娜, 陈宏, 李志鹏, 等. 基于 GF - 1 号影像的南方水稻种植信息提取 [J]. 地理空间信息, 2021, 19(3): 60 - 63.
- [8] Lin N, Chen H, Li Z P, et al. Information extraction of southern rice planting based on GF - 1 image [J]. Geospatial Information, 2021, 19(3): 60 - 63.
- [9] 张冠宏, 王新军, 徐晓龙, 等. 基于面向对象的无人机遥感影像荒漠植被分类 [J]. 中国农业科技导报, 2021, 23(5): 69 - 77.
- [10] Zhang G H, Wang X J, Xu X L, et al. Desert vegetation classification based on object - oriented UAV remote sensing images [J]. Journal of Agricultural Science and Technology, 2021, 23(5): 69 - 77.
- [11] 刘芸, 李雪, 廖瑶, 等. 基于面向对象分类法的贵州烤烟种植区域提取——以烤烟成熟期为例 [J]. 贵州科学, 2021, 39(2): 82 - 86.
- [12] Liu Y, Li X, Liao Y, et al. Extraction of roasted tobacco growing areas in Guizhou based on object - oriented classification method: Taking roasted tobacco maturity period as an example [J]. Guizhou Science, 2021, 39(2): 82 - 86.
- [13] Ke L, Gang W, Gong C, et al. Object detection in optical remote sensing images: A survey and a new benchmark [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 296 - 307.
- [14] 杜敬. 基于深度学习的无人机遥感影像水体识别 [J]. 江西科学, 2017, 35(1): 158 - 161, 170.
- [15] Du J. Deep learning based UAV remote sensing image water body identification [J]. Jiangxi Science, 2017, 35(1): 158 - 161, 170.
- [16] 范荣双, 陈洋, 徐启恒, 等. 基于深度学习的高分辨率遥感影像建筑物提取方法 [J]. 测绘学报, 2019, 48(1): 38 - 45.
- [17] Fan R S, Chen Y, Xu Q H, et al. A high - resolution remote sens-

- ing image building extraction method based on deep learning[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(1):38–45.
- [12] 刁明光, 于晨, 李文吉, 等. 矿山遥感监测实例分割数据集的自动生成方法[J]. *中国科技论文*, 2021, 16(3):329–335.  
Diao M G, Yu C, Li W J, et al. Automatic generation method of instance segmentation data set in mine remote sensing monitoring [J]. *China Sciencepaper*, 2021, 16(3):329–335.
- [13] Menapace A, Pisaturo G R, De Luca A, et al. EPANET in QGIS framework; The QEPANET plugin [J]. *Journal of Water Supply: Research and Technology – Aqua*, 2020, 69(1):1–5.
- [14] Zhang W W, Wang K F, Wang Y T, et al. A loss – balanced multi – task model for simultaneous detection and segmentation [J]. *Neuro-computing*, 2021, 428:65–78.
- [15] He K, Gkioxari G, Dollár P, et al. Mask R – CNN [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2):386–397.
- [16] Huang Y J, Xu H X. Fully convolutional network with attention modules for semantic segmentation [J]. *Signal, Image and Video Processing*, 2021, 15(5):1031–1039.
- [17] Miao J, Xu S W, Zou B X, et al. ResNet based on feature – inspired gating strategy [J]. *Multimedia Tools and Applications*, 2022, 81(4):19283–19300.
- [18] Bai T, Pang Y, Luo J S, et al. An optimized faster R – CNN method based on DRNet and RoI align for building detection in remote sensing images [J]. *Remote Sensing*, 2020, 12(5):762.
- [19] 孙禧勇, 许玮, 姜德才, 等. 基于遥感的雄安新区地表水时序变化与区域规划研究 [J]. *地球物理学进展*, 2021, 36(4):1443–1455.  
Sun X Y, Xu W, Jiang D C, et al. Research on time series changing trend of surface water and regional planning in Xiong’ an New Area based on remote sensing [J]. *Progress in Geophysics*, 2021, 36(4):1443–1455.

## Mask R – CNN – based intelligent identification of sparse woods from remote sensing images

DIAO Mingguang<sup>1</sup>, LIU Yong<sup>1</sup>, GUO Ningbo<sup>1</sup>, LI Wenji<sup>2</sup>, JIANG Jikang<sup>1</sup>, WANG Yunxiao<sup>1</sup>

(1. *School of Information Engineering, China University of Geosciences (Beijing), Beijing 100083, China*; 2. *China Aero Geophysical Survey and Remote Sensing Center for Natural Resources, Beijing 100083, China*)

**Abstract:** There are only a few low – accuracy methods available for the feature extraction of sparse woods from remote sensing images. Moreover, there is a lack of datasets for intelligent identification. This study proposed a method for intelligent information identification of sparse woods from remote sensing images. First, a dataset was created using QGIS and Python separately to provide data support for model training. Then, feature maps were generated through feature extraction, and then regions of interest (ROIs) were extracted from the feature maps. Subsequently, these ROIs were filtered through pooling operations (ROI align) to reduce the memory consumption caused by too many ROIs in the images. Experiments show that the method proposed in this study can create datasets quickly and facilitate the identification of sparse woods from remote sensing images. Moreover, the Mask R – CNN – based intelligent identification has a target detection mean average precision (MAP) of up to 0.92.

**Keywords:** remote sensing image; sparse woods; deep learning; Mask R – CNN; QGIS

(责任编辑: 陈理)