



武汉大学学报(信息科学版)  
*Geomatics and Information Science of Wuhan University*  
ISSN 1671-8860, CN 42-1676/TN

## 《武汉大学学报(信息科学版)》网络首发论文

题目: 模拟困难样本的 Mask R-CNN 滑坡分割识别  
作者: 姜万冬, 席江波, 李振洪, 丁明涛, 杨立功, 谢大帅  
DOI: 10.13203/j.whugis20200692  
收稿日期: 2020-12-24  
网络首发日期: 2021-08-31  
引用格式: 姜万冬, 席江波, 李振洪, 丁明涛, 杨立功, 谢大帅. 模拟困难样本的 Mask R-CNN 滑坡分割识别. 武汉大学学报(信息科学版).  
<https://doi.org/10.13203/j.whugis20200692>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字符、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20200692

引用格式：

姜万冬, 席江波, 李振洪, 等. 模拟困难样本的 Mask R-CNN 滑坡分割识别[J]. 武汉大学学报·信息科学版, 2021, DOI: 10.13203/j.whugis20200692 (Jiang Wandong, Xi Jiangbo, Li Zhenhong, et al. Landslide Detection and Segmentation Using Mask R-CNN with Simulated Hard Samples [J]. Geomatics and Information Science of Wuhan University, 2021, DOI: 10.13203/j.whugis20200692)

## 模拟困难样本的 Mask R-CNN 滑坡分割识别

姜万冬<sup>1,2</sup> 席江波<sup>1,2,3</sup> 李振洪<sup>1,2,3</sup> 丁明涛<sup>1,2,3</sup> 杨立功<sup>1</sup> 谢大帅<sup>1,2</sup>

1 长安大学地质工程与测绘学院, 陕西 西安, 710054

2 长安大学地学与卫星大数据研究中心, 陕西 西安, 710054

3 西部矿产资源与地质工程教育部重点实验室, 陕西 西安, 710054

**摘要：**随着人工智能的发展, 利用高分影像进行滑坡等地质灾害识别逐渐成为研究热点。滑坡目视解译需依赖专家经验, 传统滑坡自动识别方法又易将滑坡和裸地、道路等地物混淆。针对以上问题, 提出基于模拟困难样本的 Mask R-CNN (mask region-based convolutional network) 滑坡提取方法。在现有样本的基础上, 利用滑坡的形状、颜色、纹理等特征模拟更为复杂的滑坡背景, 进行困难样本挖掘增强, 并将得到的困难样本输入 Mask R-CNN 网络进行滑坡精细检测分割。在实际研究区域中, 由于滑坡数量有限, 因此在频率域进行小样本学习, 在减少数据需求的同时, 保证分割识别的准确度。在贵州毕节的实验结果表明, 基于模拟困难样本的 Mask R-CNN 方法检测精度为 94.0%, 像素分割平均准确率为 90.3%, 可实现低虚警率下的高性能检测分割; 采用频率域学习, 在一半数据输入量的情况下, 模型检测精度仍可得到提升; 并利用天水地区的滑坡区域进行实际验证, 进一步证明了所提方法的有效性。

**关键词：**深度学习; 滑坡提取; Mask R-CNN; 困难样本; 频率域

<sup>1</sup>项目资助: 国家自然科学基金重大项目(41941019), 国家重点研发计划(2018YFC1504805), 国家自然科学基金(61806022, 41874005), 中央高校基本科研业务费专项资金(300102260301/087, 300102260404/087, 300102269103, 300102269304 和 300102269205), 地理信息国家重点实验室开放基金(SKLGIE2018-M-3-4)。

第一作者: 姜万冬, 硕士, 主要研究方向为遥感影像地质灾害智能解译, CHD\_jwd@126.com

\*通讯作者: 席江波, 博士, 讲师, 主要研究方向包括机器学习, 遥感影像智能分析, xijiangbo@chd.edu.cn

滑坡作为一种常见的地质灾害，在世界范围内具有危害性和损毁性大、易发性和突发性强的特点。不少学者利用合成孔径雷达干涉测量（Interferometric Synthetic Aperture Radar, InSAR）<sup>[1-3]</sup>、全球导航卫星系统（global navigation satellite system, GNSS）<sup>[4-6]</sup>、激光雷达（light detection and ranging, LiDAR）<sup>[7,8]</sup>等技术对滑坡开展了早期识别<sup>[9,10]</sup>、隐患分析<sup>[11]</sup>、监测预警<sup>[12]</sup>等研究，并取得了较好的研究成果。但灾害发生后，尤其是面对重大自然灾害（如地震、海啸、洪水等）引发的大范围、分布广、数量众多的滑坡区域，如何快速精确地获取滑坡发生的位置、受灾范围等信息，对开展地应急救援工作具有重要指导意义。同时，进行滑坡风险评估、建模和制图时，获取所需滑坡数据费时费力，准确高效地识别滑坡并更新现有滑坡影像库，迄今为止仍是亟待解决的一项技术难题。

野外实地调查是滑坡勘测最为传统和直接的方法，但需要投入大量的时间、人力和物力。近年来，遥感技术不断发展，影像空间分辨率越来越高，利用遥感影像可快速获取大面积的地物信息，因而被广泛应用到地质灾害解译中。目前，利用遥感影像进行滑坡识别的主要方法有目视解译<sup>[13]</sup>、基于像元<sup>[14, 15]</sup>、面向对象<sup>[16-18]</sup>等。目视解译是遥感解译中最早使用的方法；Petschko 等<sup>[19]</sup>利用 LiDAR 和 DTM (digital terrain model) 衍生数据进行滑坡目视解译和制图，许冲等<sup>[20]</sup>针对 2008 年汶川地震引发的次生地质灾害，通过目视解译完成了 14 个重灾区，46 560 处崩塌、滑坡、泥石流等地震次生地质灾害解译工作。目视解译滑坡方法过于依赖专家经验且需投入大量的时间和精力，基于像元的方法通过分析影像数据的光谱、空间信息，选取地物特征，按特定的方法将图像中的每一个像元进行分类；基于像元的方法主要分为监督分类和非监督分类<sup>[21, 22]</sup>。面向对象方法与基于像元的方法不同，它将邻近范围区域内的像元作为最小的处理单位，可以有效地利用影像上目标物体的形状、纹理等空间特征。但面向对象方法阈值设定往往仅针对特定研究区域，其迁移性和可移植性还有待进一步提高。

机器学习在遥感影像处理中运用日渐广泛，不同的机器学习方法如支持向量机（support vector machine, SVM）<sup>[23, 24]</sup>、随机森林（Random Forests, RF）<sup>[25, 26]</sup>等，已经和基于像元、面向对象等方法结合，成为滑坡识别新的有效方法。Mezaal 等<sup>[27]</sup>利用 SVM、RF 和 K 最近邻算法分类器的结果作为概率输出，结合 DST (Dempster-Shafer theory) 规则组合对马来西亚等热带地区的滑坡进行检测；Stumpf 等<sup>[26]</sup>将面向对象方法与 RF 结合，选取重要的目标特征和分类阈值进行滑坡制图；近年来，深度学习技术在图像处理领域<sup>[28-30]</sup>（如图像分类<sup>[31]</sup>、目标检测<sup>[32]</sup>、图像分割<sup>[33]</sup>等）取得了系列成就，部分学者开始考虑将深度学习技术应用到滑坡识别上来。Shi 等<sup>[31]</sup>利用卷积神经网络（convolutional neural network, CNN）实现了中国香港地区凌峰山和大屿山地区的滑坡大面积快速检测。Omid 等<sup>[32]</sup>以尼泊尔地区为例，对比了 SVM、RF 和 CNN 等方法在滑坡识别方面的效果。针对研究区域中，滑坡样本在影像中所占的比例较小，而背景物体所占比例较大的问题，Yu 等<sup>[33]</sup>首先利用植被指数和数字高程模型（digital elevation model, DEM）剔除了大量的背景区域，筛选出可能发生滑坡的区域，再利用语义分割模型对潜在滑坡区域进行精确检测；此外，DEM、坡度、坡向等地学因子逐渐开始被运用到深度学习模型中，辅助参与滑坡的检测以提高滑坡检测的精度。Hao 等<sup>[34]</sup>先利用地形数据、地质数据和降雨等相关数据建立综合地质数据库，再利用 SVM、RF 和 CNN 等方法进行评估。Liu 等<sup>[35]</sup>利用 U-net 网络模型，并结合坡度坡向等地形因子，实现了九寨沟地区滑坡的大面积检测。整体上来看，将深度学习技术运用到滑坡检测上，可以实现大区域范围内滑坡的自动化检测，能较好地满足灾害的应急响应要求。但利用深度学习模型检测滑坡，需要利用大量的滑坡数据进行模型训练，而在实际应用中，滑坡数量往往难以达到模型训练的数据量要求。另一方面，在面对与滑坡图像特征相近的地物如裸地、旱地和某些特殊人工建筑等，如何有效地避免对这些易混淆地物误判，始终是各种滑坡检测方法面临的一个棘手问题。

针对上述问题，本文提出了基于模拟困难样本的 Mask R-CNN 滑坡提取方法，在滑坡检测的同时实现了滑坡形状像素级分割。该方法在遥感影像预处理后，建立滑坡影像数据集，根据滑坡样本的特征进行困难样本模拟，以提高模型对困难样本的识别能力，再通过 Mask R-CNN 进行滑坡的精识别分割。本文还在少量滑坡样本的情况下，利用图像的频率域信息进行模型训练。先在毕节市滑坡数据集上进行了 Mask R-CNN 模型训练和测试，随后以天水滑坡区域进一步检验模型在实际区域的适应性和有效性。

## 1 方法

### 1.1 Mask R-CNN 基本原理

1998 年，Fukushima 等<sup>[36]</sup>提出 LeNet 模型，构成了 CNN 的基础框架，其主要包含卷积、池化、全连接层三个部分。卷积部分主要通过一个  $a \times a$  的滑动窗口对图像进行卷积操作，提取图像的局部特征，再将值输入非线性激活函数。池化部分通过一个  $b \times b$  大小的滑动窗口遍历矩阵，取滑动窗口内的平均值或最大值作为该窗口的输出值，从而大幅减少了模型的参数。全连接的每一个单元都与上一层全部相连，其目的主要用于分类。

部分学者开始将 CNN 模型用于滑坡检测，但从 CNN 到 R-CNN<sup>[37]</sup>、Fast R-CNN<sup>[38]</sup>、再到 Faster R-CNN<sup>[39]</sup>，模型性能得到了不断的提升。本文采用由 Faster R-CNN 基础上改动而来的 Mask R-CNN 模型<sup>[40]</sup>进行滑坡检测，其 Mask R-CNN 的网络结构如图 1 所示。Mask R-CNN 的区域提取网络（Region of Interest, RoI）采用了双线性插值，避免了两次坐标“量化”取整引入的误差，大大提高了模型的检测性能。它的另一改动是增加了一个全连接网络预测目标的形状掩膜，故其损失函数只在 Fast R-CNN 损失函数的基础上增加了  $L_{mask}$  支路的损失值，见公式（1）。

$$\left\{ \begin{array}{l} L = \frac{1}{N_{cla}} \sum_i L_{cla}(P_i, P_i^*) + \frac{k}{N_{reg}} \sum_i [L_{cla}(P_i^* \times L_{box}(t_i, t_i^*))] + L_{mask} \\ L_{cla}(P_i, P_i^*) = -\log [P_i \times P_i^*] + (1 - P_i)(1 - P_i^*) \\ L_{box}(t_i, t_i^*) = Smooth(t_i - t_i^*) \end{array} \right. \quad (1)$$

式中  $L$  代表 Mask R-CNN 模型整体损失函数值， $N_{cla}$  代表样本总数量， $N_{reg}$  表示特征图层的尺寸； $P_i$  和  $P_i^*$  分别表示 anchor 预测为目标和背景的概率； $L_{cla}$  代表  $P_i$  和  $P_i^*$  的对数损失， $L_{box}$  代表边框回归的损失值； $t_i$  是一个向量，代表 anchor 在区域提取网络（region proposal network, RPN）训练阶段预测的偏移量， $t_i^*$  则是实际偏移量。

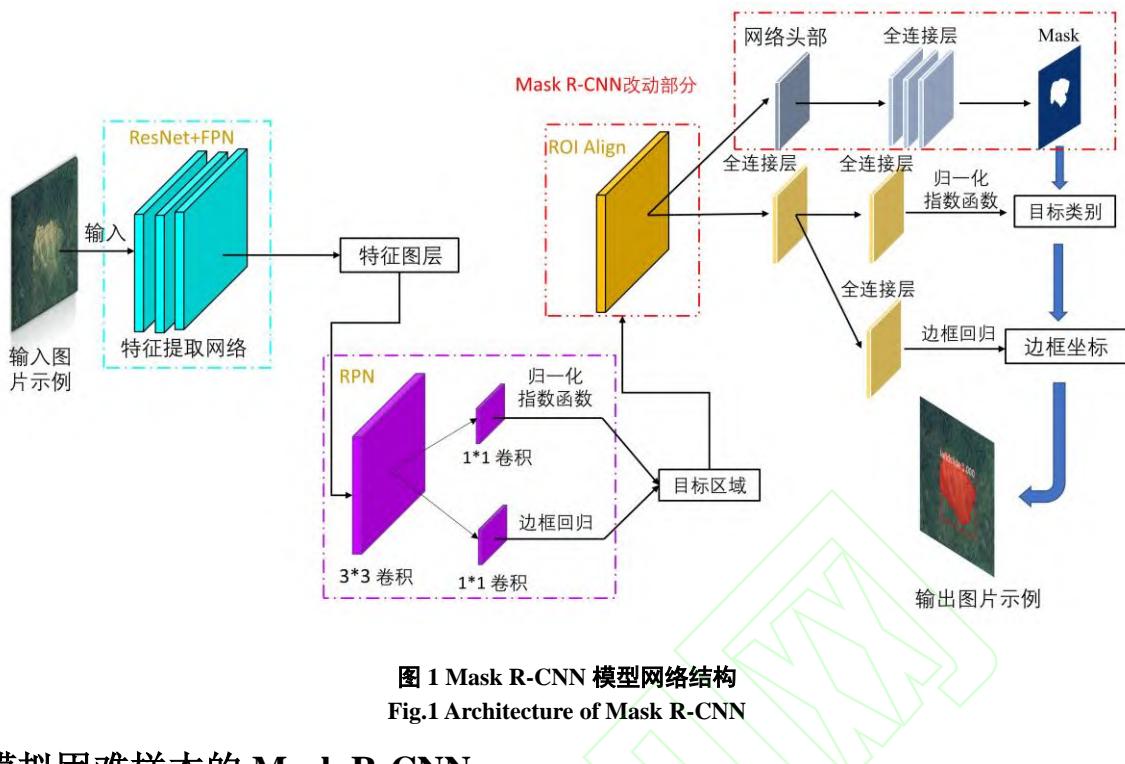


图 1 Mask R-CNN 模型网络结构  
Fig.1 Architecture of Mask R-CNN

## 1.2 模拟困难样本的 Mask R-CNN

为了提高检测模型的精度和泛化能力,一般需更多地训练数据,不断增加深度学习的网络层数以获得更好地表现效果。通常检测背景的种类、数量、复杂程度远远高于待检测目标,因而本文从增强检测背景的角度出发,引入背景增强方法<sup>[41]</sup>,该方法将三张负样本数据与正样本数据进行拼接,旨在让模型学习到背景所包含的特征,明确非检测目标,从而有效地降低模型的误检测率。由于多数背景与滑坡差异较大,若只用大量简单易分的负样本去训练模型,会主导梯度更新方向和掩盖重要信息。因此,本文根据滑坡和易混淆地物(裸地、农田、人工建筑等)的图像特征,选取了部分典型易混淆地物作为滑坡检测背景,合成困难样本,提出了基于模拟困难样本的 Mask R-CNN 方法,整体流程如图 2 所示。

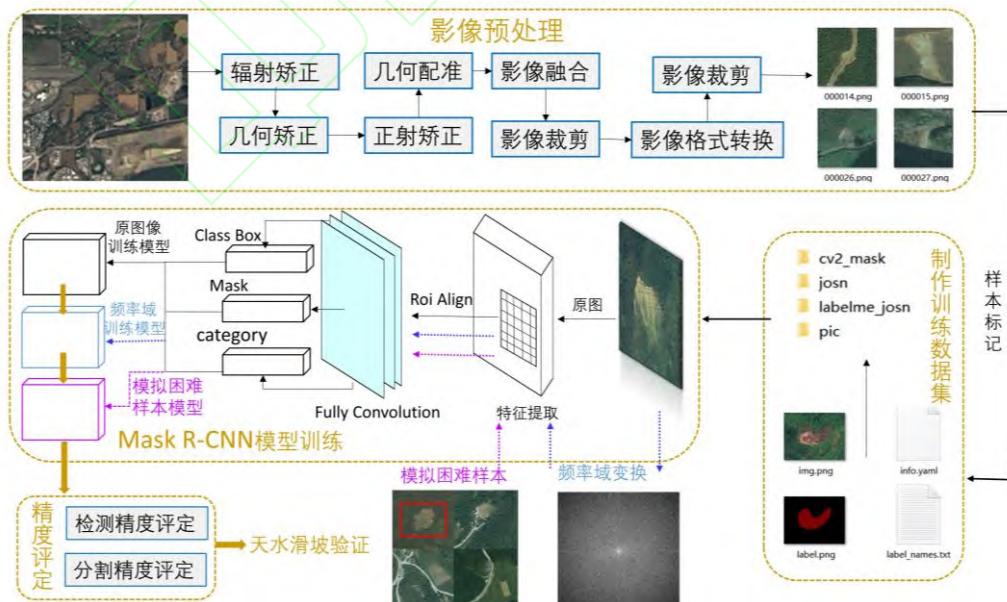
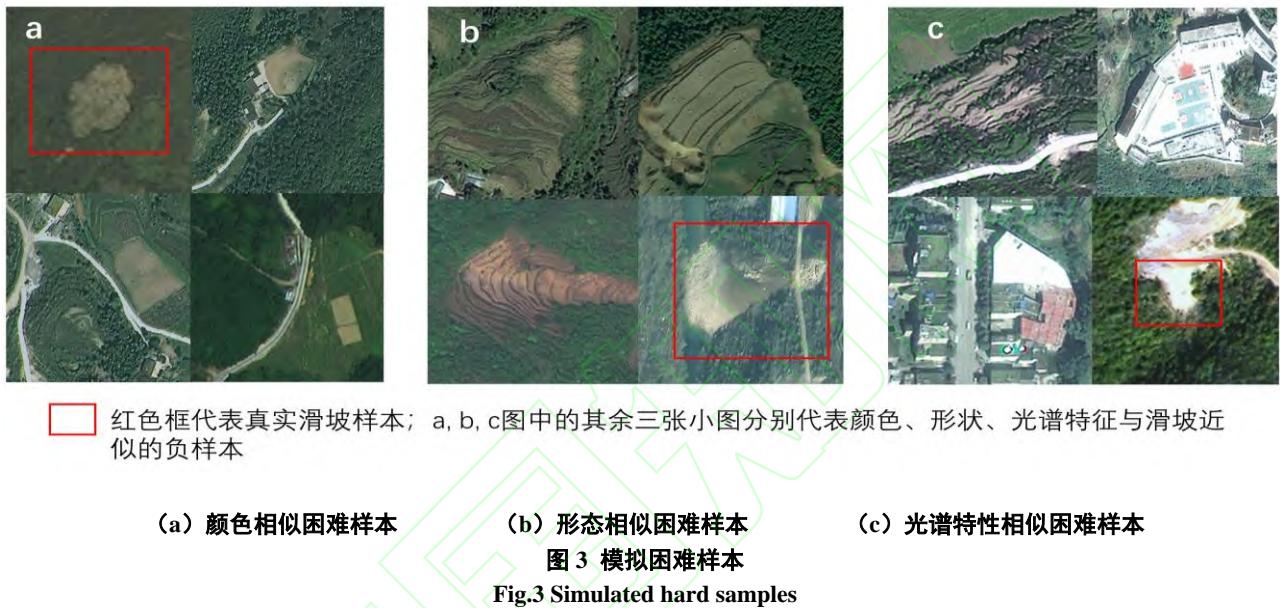


图 2 基于 Mask R-CNN 的滑坡分割识别  
Fig.2 Landslide detection and segmentation using Mask R-CNN

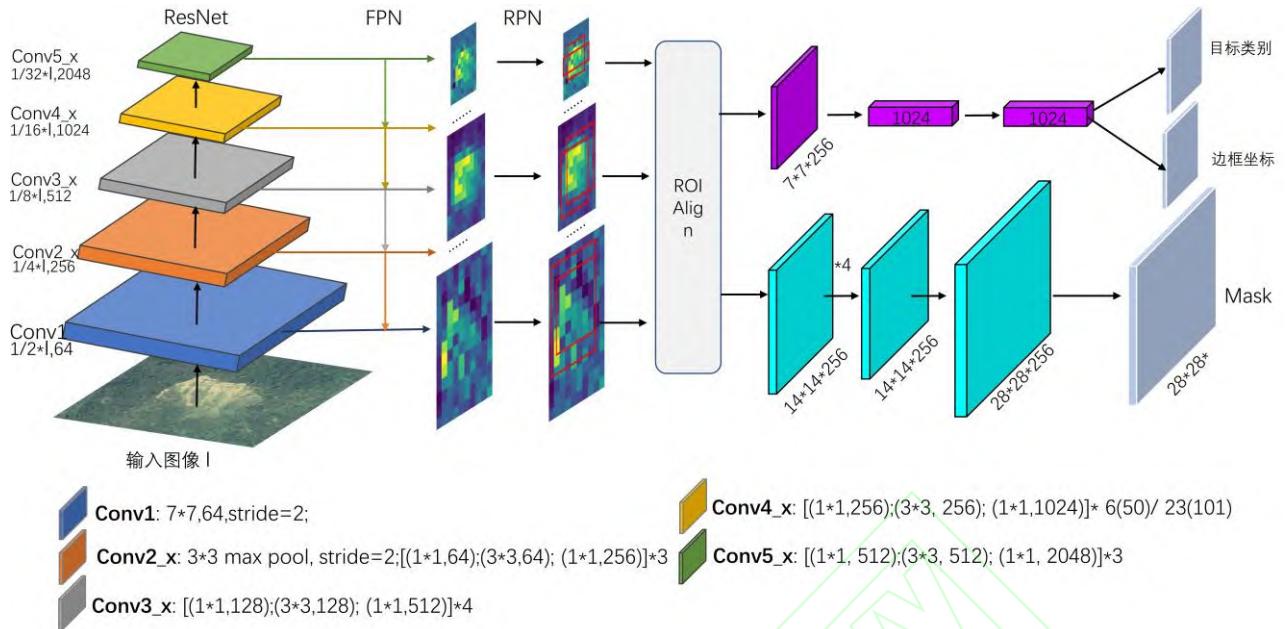
困难样本具体选取标准如下：

- 1) 根据样本的颜色特征：部分地物颜色与滑坡相近，如旱地、裸地较易与滑坡混淆（图 3 a）；
- 2) 根据样本的形态特征：部分地物呈漏斗状、舌状，与滑坡常见的形态特征较为相似（图 3 b）；
- 3) 根据样本光谱特性：部分滑坡区域土壤类型为盐碱地，其对红、绿、蓝及红外等各波段的光均有较高的反射率，在合成彩色影像时呈现白色，这些白色的滑坡区域与很多人工建筑在图像亮度均比较高，易对模型识别滑坡造成干扰（图 3 c）。

按照以上标准，在数据集样本中选取颜色、形态、光谱等特征与滑坡相近的地物，将其与真实滑坡样本进行拼接处理。模拟得到新的样本中存在多处与滑坡特征相近的地物，故统称为困难样本，输入模型训练。



将模拟的困难样本输入 Mask R-CNN 模型后，首先由 ResNet 骨干网络提取图像特征（图 4），其网络结构主要包含 5 个不同的卷积模块组成（Conv1、Conv2\_x、Conv3\_x、Conv4\_x、Conv5\_x）。Conv1 通过一个普通的卷积层 $[7 \times 7 \times 64]$ ，步长为 2，对输入图片进行特征提取，并将尺寸降采样为原来的 1/2；然后 Conv2\_x 接上 $3 \times 3$  的最大池化层，步长为 2，缩小特征图层尺寸，然后就送入残差块。残差块由 $1 \times 1$ 、 $3 \times 3$ 、 $1 \times 1$  的三个卷积层构成， $1 \times 1$  卷积将图像大小从 $256 \times 256$  降到 $64 \times 64$ ，再进行 $3 \times 3$  的卷积，最后通过 $1 \times 1$  卷积恢复图像大小。ResNet 网络常用的网络结构有 ResNet50 和 ResNet101 两种，其两者差异在于 Conv4\_x 中残差块数量不同。图像经过 ResNet 网络的 5 个不同卷积模块提取特征，得到 5 个不同尺寸的特征图层，并将其输入 FPN 网络。FPN 网络充分利用了各个阶段提取得到的特征，首先对上层特征图层升采样到与下一层相同的大小，并与下一层进行相加，从而得到一个融合好的新特征层。其次，RPN 在融合得到的特征图层上设置滑动窗口，寻找滑坡目标区域。然后，在滑坡目标区域输入 ROI Align。ROI Align 通过双线性插值方法，精确表示了滑坡目标区域坐标框的四个点坐标。最后，将 ROI Align 输出的特征图层进行分类、回归和 Mask 生成。Mask 分支将来自 ROI Align $14 \times 14$  的特征图层，通过 4 个 $3 \times 3$  的卷积核卷积，以及一个反卷积，得到 $28 \times 28$  大小的特征图层，再使用 $1 \times 1$  卷积得到 $n$ （目标类别数）个 mask。本文运用 Mask R-CNN 进行滑坡分割，故目标类别数 $n$  为 1。



### 1.3 小样本下的滑坡分割识别

具体滑坡检测案例研究中,研究地区的滑坡数量往往有限,难以满足深度学习模型训练的数据量要求。针对小样本滑坡数据的限制,利用频率域信息进行模型训练,在牺牲部分精度情况下,采用一半的原数据输入 Mask R-CNN 模型进行训练。首先对原始滑坡图像进行灰度化和二值化,然后在利用空间域信息训练的基础上,加入图像频率域信息进行 Mask R-CNN 模型训练。神经网络主要利用图像的空间域信息进行模型训练,通常需要将图像下采样到预定的神经网络输入大小。尽管下采样操作减少了模型计算量,但造成了图像有效信息的损失,导致精度下降。通过拓展输入图像的信息量,一定程度上弱化图像下采样带来的负面影响。因此,利用傅里叶函数将图像从空间域转换到频率域,公式如下:

$$F(w) = \int_{-\infty}^{+\infty} f(t) e^{-iwt} dt \quad (2)$$

式中  $w$  表示频率,  $t$  表示时间,  $e^{-iwt}$  表示复变函数。

最后,将频率域变换得到的图像和原图像一起输入 Mask R-CNN 深度模型进行训练,具体如图 5 所示。该方法可有效地利用图像的频率域信息,在适当减少样本输入量的情况下,使得模型精度受到的影响相对较小。

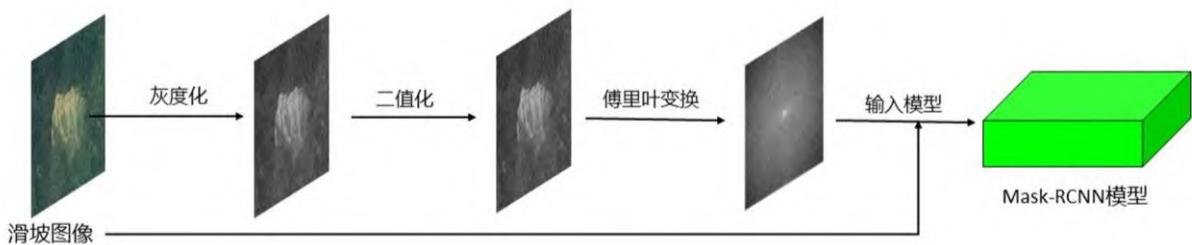


图 5 频率域的滑坡识别  
Fig.5 Landslide detection in the frequency domain

## 2 实验

### 2.1 研究区域及数据

本文实验区域位于贵州省毕节市，毕节市地处亚热带，为典型的亚热带季风气候，全年降水较为充足。同时，毕节市位于云贵高原，地势上属于二三级阶梯交汇处，区域范围内海拔相对高差大。毕节市的地形和降水条件，为山体滑坡的孕育奠定了基础。本次实验使用的模型训练数据是由 Ji 等<sup>[42]</sup>制作的贵州省毕节市滑坡数据集（[http://gpcv.whu.edu.cn/data/Bijie\\_pages.html](http://gpcv.whu.edu.cn/data/Bijie_pages.html)）。该数据集主要包括滑坡影像及对应 DEM 数据和滑坡形状掩膜，其中滑坡影像分辨率为 0.8m，DEM 分辨率为 2m，具体如图 6 所示。先在毕节市滑坡数据集上进行模型的训练和测试，随后选取天水滑坡区域对模型检测效果进一步验证。

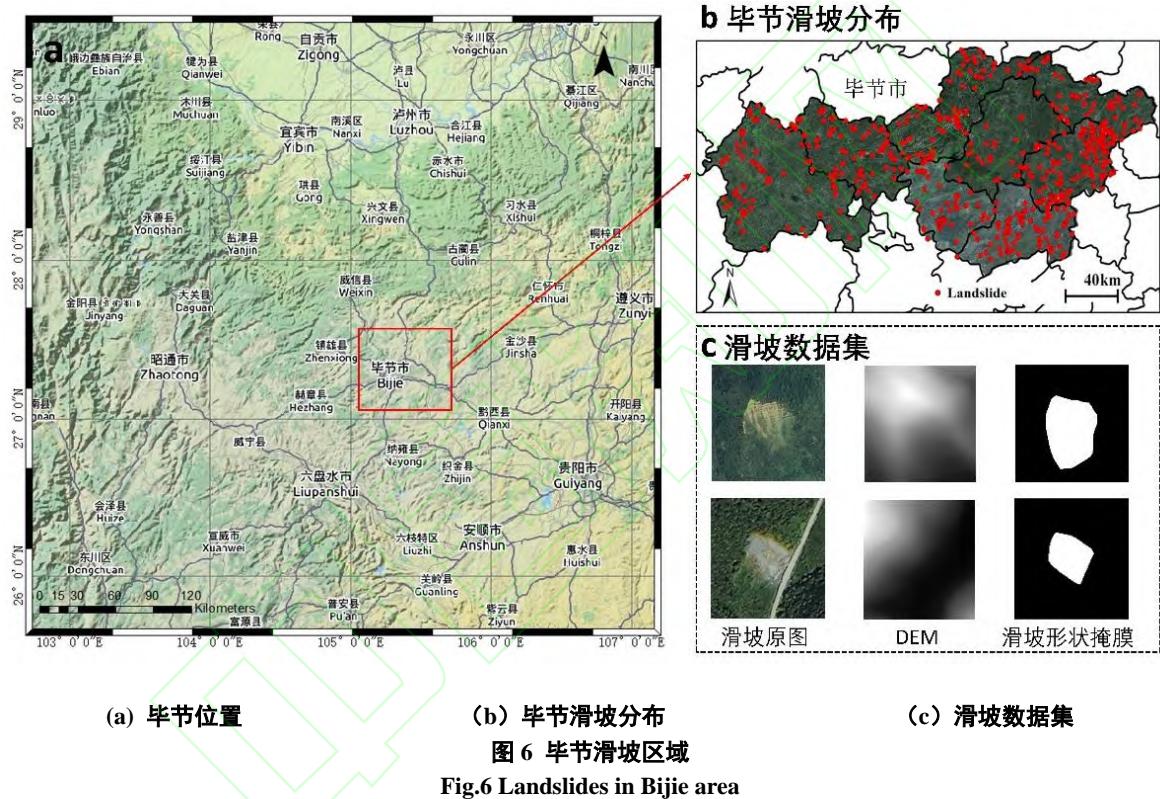


Fig.6 Landslides in Bijie area

### 2.2 数据预处理

遥感成像时，由于传感器不稳定或受到大气辐射等影响，可能导致影像辐射量失真、产生几何畸变等现象。这些现象均可严重影响遥感影像的质量，因此首先需进行图像的几何校正(粗校正、几何精校正、正射校正)、大气校正、图像镶嵌、图像融合、图像裁剪、去云去阴影等预处理操作。单景遥感影像数据量大，直接进行处理需大量计算资源。将遥感影像进行格式转换和分块处理。假设原始影像  $f(X, Y)$  大小为  $M \times N$ ，其中  $0 < X < M$ ， $0 < Y < N$ ；裁剪后影像  $f(x, y)$  大小为  $m \times n$ ，其中  $0 < x < m$ ， $0 < y < n$ ，设定裁剪时滑动窗口大小为  $m \times n$ ，移动步长为  $m$ 。先对影像边缘进行填充处理，使得  $m$  和  $n$  分别能被  $M$  和  $N$  整除，避免影像剩余边角小于滑动窗口大小。以影像左下角为原点，影像的长为  $X$  轴，宽为  $Y$  轴，滑动窗口左下角坐标记为  $[i, j]$ ，具体处理流程如图 7 所示。

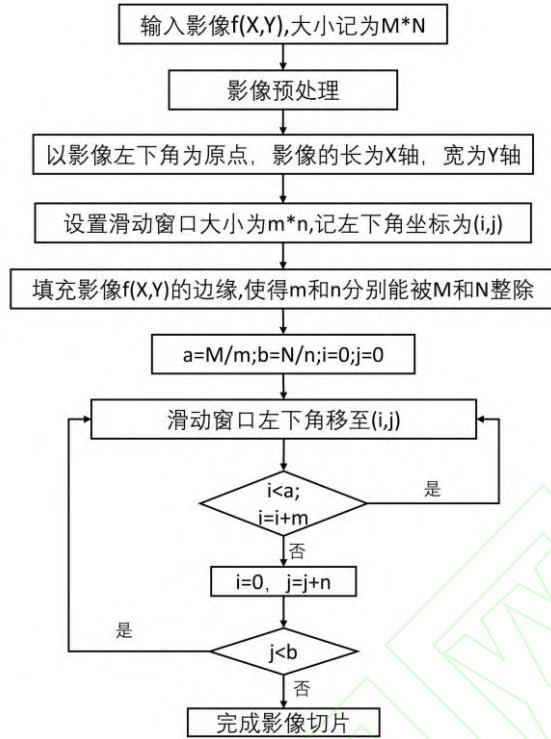


图 7 数据预处理  
Fig.7 Data preprocessing

## 2.3 实验设置

本文实验硬件配置如下：处理器 Intel (R) Core (TM) i7-8700K，运行内存 64 GB，显卡型号为 NVIDIA RTX 2080Ti。深度学习框架为 Tensorflow，其他的主要辅助软件包括 Pycharm、VC++、Anaconda 和 Python。实验预处理时，切片大小设置为  $256 \times 256$ ，标签名称设置为 landslide，实验数据集大约按照 9:1 划分训练集和测试集。Mask R-CNN 模型的训练代码可在 [https://github.com/Jiang-CHD-YunNan/Mask\\_RCNN](https://github.com/Jiang-CHD-YunNan/Mask_RCNN) 获得，实验设置了 10 个 epochs，每个 epoch 迭代次数为 1 000 次，初始学习率设置为 0.001，权重衰减系数为 0.005，动量因子为 0.9，梯度裁剪系数为 5。

利用毕节滑坡数据，对基于模拟困难样本的 Mask R-CNN 及小样本下的频率域方法进行验证，并与基于原始影像的 Mask R-CNN 模型、加入了 DEM 通道的 Mask R-CNN 模型、加入了 DEM 梯度通道的 Mask R-CNN 模型、进行背景增强得到的 100 个样本加入 Mask R-CNN 训练对比，进一步证明提出方法的有效性。本文提出的基于模拟困难样本方法结合了易混淆滑坡的样本特征，针对性地模拟了 100 个困难样本，并以原始影像训练得到模型作为初始权重，进行模型训练。

## 3 结果

### 3.1 评价指标

混淆矩阵作为评价检测结果的基础指标，是计算各种评价指标的基础。混淆矩阵虽直观地统计了检测结果的个数，但难以精确地评价模型好坏，因此，利用二级评价指标：精确度 (precision,  $P$ )、召回率(recall,  $R$ )和准确度(accuracy,  $A$ )对模型进行进一步地评价，精确度反映

的是检测为正样本的目标中正确的概率，召回率反映在所有正样本中正确识别的概率，准确度反映的是预测目标中预测正确的比例。计算公式如下：

$$\left\{ \begin{array}{l} P = \frac{TP}{TP + FP} \\ R = \frac{TP}{TP + FN} \\ Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \end{array} \right. \quad (3)$$

式中， $TP$  表示实际为正样本，预测为正样本； $FP$  表示实际为负样本，预测为正样本； $FN$  表示实际为正样本，预测为负样本； $TN$  表示实际为负样本，预测为负样本；

通常来讲，精确度高时召回率往往偏低，召回率高时则精确度往往偏低，只有在较为简单的数据集上，才可能精确度和召回率都会很高，为了综合地衡量检测模型的好坏，引进 F1 分数对模型进行评价：

$$F1 = \frac{2PR}{P + R} \quad (4)$$

本文在滑坡检测的同时实现了滑坡形状分割，为对分割结果进行合理评价，引入了平均像素准确率（mean pixel accuracy, MPA）和平均交并比指标（mean intersection over union, MIoU）。MPA 为每一预测类别正确的像素数占总像素数的比例，再累加求取平均值，反映分割模型的准确率，MIoU 每一类预测结果和真实掩膜之间交集与并集比的平均值。计算公式如下：

$$\left\{ \begin{array}{l} MPA = \frac{1}{n+1} \sum_{i=0}^n \frac{P_{ii}}{\sum_{j=0}^n P_{ij}} \\ MIoU = \frac{1}{n+1} \sum_{i=0}^n \frac{P_{ii}}{\sum_{j=0}^n P_{ij} + \sum_{j=0}^n P_{ji} - P_{ii}} \end{array} \right. \quad (5)$$

式中， $n$  为预测的类别数目； $P_{ii}$  表示原本为  $i$  类预测为  $i$  类； $P_{ij}$  表示原本为  $i$  类预测为  $j$  类； $P_{ji}$  表示原本为  $j$  类预测为  $i$  类。

### 3.2 滑坡分割识别结果

本次实验测试一共为 100 张图片，其中包含 50 个滑坡正样本，50 个滑坡负样本，其滑坡分割识别结果展示如图 8 所示，各种检测方法对比如表 1 和图 9 所示。在表 1 中， $TP$  为 50 个滑坡中正确检测个数， $FN$  为 50 个滑坡中漏检个数， $FP$  为 50 个非滑坡中的误判个数， $TN$  为 50 个非滑坡正确检测的个数。



(a) 滑坡原图

(b) 滑坡分割识别结果

图 8 滑坡分割识别  
Fig. 8 Landslide detection and segmentation

表 1 实验结果

Tab. 1 Experimental results

探测方法	TP	FN	FP	TN	P/%	R/%	A/%	F1
原图片	49	1	30	20	62.0	98.0	69.0	0.76
DEM	50	0	37	13	57.5	100.0	63.0	0.73
DEM 梯度	50	0	35	15	58.8	100.0	65.0	0.74
背景增强	44	6	13	37	77.2	88.0	81.0	0.86
频率域	50	0	33	17	60.2	100.0	67.0	0.78
本文方法	<b>48</b>	<b>2</b>	<b>4</b>	<b>46</b>	<b>92.3</b>	<b>96.0</b>	<b>94.0</b>	<b>0.94</b>

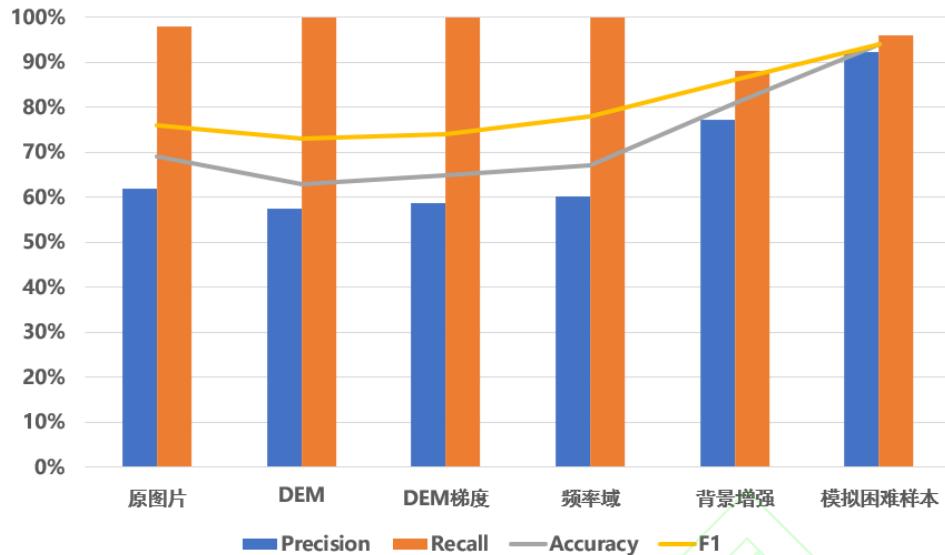


图 9 滑坡识别方法对比

Fig. 9 Comparisons between different landslide detection methods

从表 1 和图 9 可以看出，模拟困难样本的 Mask R-CNN 方法准确率、F1 分数均高于其他各种方法。利用原图片，DEM 及 DEM 梯度方法对滑坡检测准确率较高，但同时易将非滑坡样本误判为滑坡，误判率较高。经背景增强之后，召回率和检测精度有明显提升，误检率也有明显下降，模型综合评价指标 F1 分数提高 10% 左右。此外，在背景增强的基础上进行困难样本模拟，模型性能得到了进一步地提升，与原图片数据输入模型训练相比，F1 分数提高约 20%，达到 0.94，模型检测精度为 92.31%，召回率为 96%，准确率为 94%。同时，与利用原图片进行模型训练相比，利用频率域学习，在一半数据输入量的情况下，模型的综合评价指标 F1 Score 仍实现了少量的提升，这对降低数据输入量具有积极意义。

在滑坡形状分割方面，以上各种方法均取得了较好地效果。50 副滑坡图像分割的 MPA 达到了 90.3%，MIoU 达到 82.2%，整体分割效果较好。将基于模拟困难样本的 Mask R-CNN 滑坡分割方法与 边缘检测分割、迭代阈值分割、数学形态学方法、遗传算法分割等传统方法进行 比较，具体如图 10 所示。可以看到，本文方法分割滑坡具有显著优势：(1) 滑坡形状分割较为完整，有效地避免了碎屑地产生。(2) 能较好与滑坡光谱信息相似的地物区分开来。在本实验中，道路的颜色几乎是滑坡区域一致的，但利用深度学习 Mask R-CNN 模型仍能较好地将其区分开来。

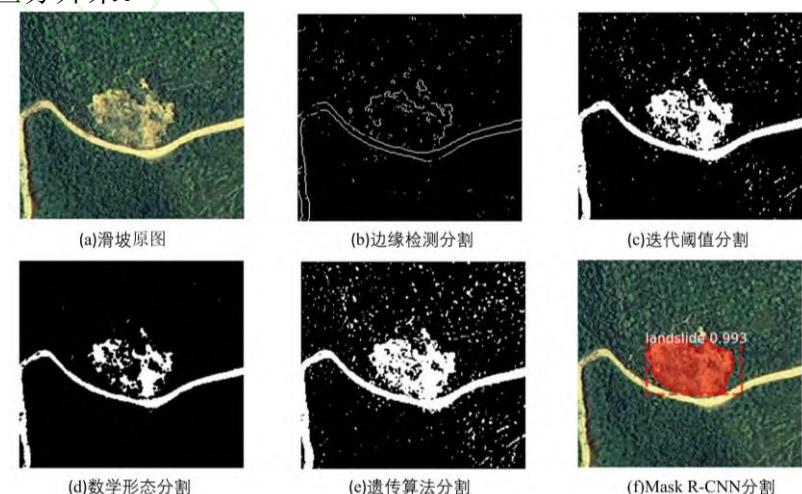


图 10 滑坡分割方法效果对比

Fig.10 Comparisons between different landslide segmentation methods

算法效率方面, Mask R-CNN 模型本身具备快速检测分割的特点, 其处理一张图像速度最快可以达到 195 ms<sup>[43]</sup>。本文测试实验部分对 100 幅不同尺寸的滑坡进行时间测试, 完成预测总时间为 657.18s, 平均单幅图片预测时间为 6.57s。通过高性能计算设备以及并行化优化处理等方式, 本文检测时间还有进一步提升的空间。

### 3.3 实际区域验证

为进一步验证模型在实际研究区域内模型对多个滑坡的检测效果, 选取甘肃省天水市秦川区娘娘坝镇柳林村的 Google Earth 影像进行验证。该区域位于 105°51'45"E、34°17'08"N, 区域面积为 1.51km<sup>2</sup>。验证区域范围内滑坡数量较多, 具有人工建筑、农田、河滩、裸地等多种地物背景, 并可通过 Google Earth 影像时间变化确认滑坡样本真值。将模拟困难样本的 Mask R-CNN 方法与变化检测、传统阈值分割方法进行对比, 由于毕节市与天水市地形地貌存在一定差异, 模拟困难样本的 Mask R-CNN 方法以毕节市滑坡数据集训练得到模型作为初始权重, 在验证区域外周围地区选取部分滑坡样本进行模型参数微调; 变化检测方法对 Google Earth 上滑坡发生前后时间最近的两幅影像, 采用 K-means 算法对两幅影像提取到的变化信息分类, 聚类中心为 2, 迭代次数为 50 次; 阈值分割方法采用了最大类间方差法自动确认阈值, 对滑坡进行分割。各种滑坡提取方法结果如图 11 所示。

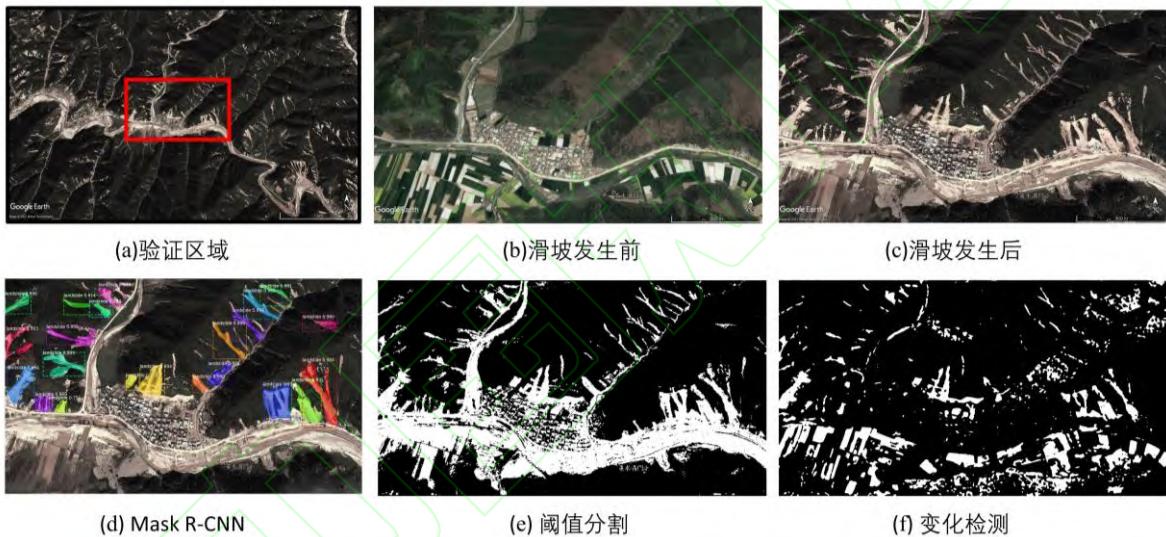


图 11 天水市滑坡分割识别

Fig. 11 Landslide detection and segmentation in Tianshui

结果显示: 模拟困难样本的 Mask R-CNN 方法较好地避免了对农田、河滩、裸地等地物误判, 实现了验证区域内大多数滑坡的识别和分割, 且无误检现象, 但针对细小滑坡存在漏检现象; 阈值分割方法对滑坡形状有较好地分割效果, 但不能较好地区分人工建筑、河滩、农田等与滑坡颜色相近的地物; 变化检测方法由于两幅影像之间的光照、地表植被、人工建筑会发生变化, 难以完全将其他发生变化的地物与滑坡区分开来; 为了精确评价各种方法对滑坡的分割效果, 本文计算利用图像分割评价指标 MPA、MIoU 对分割结果进行评价, 具体如表 2 所示。

表 2 滑坡分割结果对比

Tab.2 Comparisons between landslide segmentation results

方法	MPA/%	MIoU/%
Mask R-CNN	87.14	75.22
阈值分割	50.48	38.64

---

变化检测	50.00	38.07
------	-------	-------

---

整体上看，模拟困难样本的 Mask R-CNN 滑坡分割识别方法具备对单张影像内对多个滑坡检测和分割的能力，且该方法误检率较低，滑坡分割的 MPA、MIoU 值显著高于其他两种方法。由于毕节市滑坡数据集中细小、碎屑状的滑坡样本较少，且验证实验仅利用少量样本进行模型参数微调，故对细小滑坡的识别能力还相对较弱。该方法若能得到大量有效的样本进行模型，则验证实验中存在的这一缺陷可以得到有效改善。

## 4 结语

本文利用 Mask R-CNN 模型进行滑坡提取，针对遥感影像质量大，研究区域滑坡样本有限、易产生误判等问题，提出了基于模拟困难样本的 Mask R-CNN 方法，以毕节市滑坡数据集进行模型训练和测试，选取天水滑坡进一步验证实际区域内对多个滑坡的检测效果。毕节市滑坡数据集测试结果表明，与直接利用滑坡图像相比，利用基于模拟困难样本的 Mask R-CNN 方法滑坡误检率降低了 30.0%，F1 分数提高了 18.2%；小样本下的频率域学习方法仅在一半数据输入量的情况下，模型检测精度仍得到了提升；本文利用 Mask R-CNN 深度学习模型进行滑坡检测的同时，实现了滑坡形状像素级的分割，滑坡检测准确率达到 94.0%，像素分割平均准确率达到了 90.3%。同时，在天水实际滑坡区域验证结果显示，利用 Mask R-CNN 深度学习模型能够实现同一区域内多个滑坡的检测和分割，且较好地避免了对农田、河滩、人工建筑的误判。在未来的工作中，将结合相关地学数据对模型的适应性和迁移性进行研究，并将应用于川藏铁路廊道地质灾害隐患点的自动探测与编目更新。



## 参考文献

- [1] LI Zhenhong, SONG Chuang, YU Chen, et al. Application of Satellite Radar Remote Sensing to Landslide Detection and Monitoring: Challenges and Solutions[J]. Geomatics and Information Science of Wuhan University, 2019, 44(7): 967-979. (李振洪, 宋闯, 余琛等. 卫星雷达遥感在滑坡灾害探测和监测中的应用:挑战与对策 [J]. 武汉大学学报·信息科学版, 2019, 44(7): 967-979.)
- [2] Dai K, Li Z, Xu Q, et al. Entering the Era of Earth Observation-Based Landslide Warning Systems: A Novel and Exciting Framework [J]. IEEE Geoscience and Remote Sensing Magazine, 2020, 8(1): 136-153.
- [3] LU Huiyan, LI Weile, XU Qiang, et al. Early Detection of Landslides in the Upstream and Downstream Areas of the Baige Landslide, the Jinsha River Based on Optical Remote Sensing and InSAR Technologies[J]. Geomatics and Information Science of Wuhan University, 2019, 44(9): 1342-1354. (陆会燕, 李为乐, 许强等. 光学遥感与 InSAR 结合的金沙江白格滑坡上下游滑坡隐患早期识别 [J]. 武汉大学学报·信息科学版, 2019, 44(9): 1342-1354.)
- [4] BAI Zhengwei Z Q, HUANG Guanwen, JING Ce, WANG Jiaxing. Real-time BeiDou landslide monitoring technology of “light terminal plus industry cloud” [J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(11): 1424-1429.
- [5] Bovenga F, Nitti D, Fornaro G, et al. Using C/X-band SAR interferometry and GNSS measurements for the Assisi landslide analysis [J]. International Journal of Remote Sensing, 2013, 34(11): 4083-4104.
- [6] Qiu D, Wang L, Luo D, et al. Landslide monitoring analysis of single-frequency BDS/GPS combined positioning with constraints on deformation characteristics [J]. Survey Review, 2019, 51(367): 364-372.
- [7] Jaboyedoff M, Oppikofer T, Abellán A, et al. Use of LIDAR in landslide investigations: a review [J]. Natural Hazards, 2012, 61(1): 5-28.
- [8] McKean J, Roering J J. Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry [J]. Geomorphology, 2004, 57(3): 331-351.
- [9] XU Qiang, DONG Xiujun, LI Weile. Integrated Space-Air-Ground Early Detection, Monitoring and Warning System for Potential Catastrophic Geohazards[J]. Geomatics and Information Science of Wuhan University, 2019, 44(7): 957-966. (许强, 董秀军, 李为乐. 基于天-空-地一体化的重大地质灾害隐患早期识别与监测预警 [J]. 武汉大学学报·信息科学版, 2019, 44(7): 957-966.)
- [10] GE Daqing, DAI Keren, GUO Zhaocheng, et al. Early Identification of Serious Geological Hazards with Integrated Remote Sensing Technologies: Thoughts and Recommendations[J]. Geomatics and Information Science of Wuhan University, 2019, 44(7): 949-956. (葛大庆, 戴可人, 郭兆成等. 重大地质灾害隐患早期识别中综合遥感应用的思考与建议 [J]. 武汉大学学报·信息科学版, 2019, 44(7): 949.)
- [11] ZHU Qing, ZENG Haowei, DING Yulin, et al. A review of major potential landslide hazards analysis[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(12): 1551-1561. (朱庆, 曾浩炜, 丁雨淋等. 重大滑坡隐患分析方法综述 [J]. 测绘学报, 2019, 48(12): 1551-1561.)
- [12] ZHANG Qin, HUANG Guanwen, YANG Chengsheng. Precision Space Observation Technique for Geological Hazard Monitoring and Early Warning[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(10): 1300-1307. (张勤, 黄观文, 杨成生. 地质灾害监测预警中的精密空间对地观测技术 [J]. 测绘学报, 2017, 46(10): 1300-1307.)
- [13] Sato H P, Hasegawa H, Fujiwara S, et al. Interpretation of landslide distribution triggered by the 2005 Northern Pakistan earthquake using SPOT 5 imagery [J]. Landslides, 2007, 4(2): 113-122.

- [14] Duric U, Marjanovic M, Radic Z, et al. Machine learning based landslide assessment of the Belgrade metropolitan area: Pixel resolution effects and a cross-scaling concept [J]. *Engineering Geology*, 2019, 256:23-38.
- [15] Lu P, Qin Y, Li Z, et al. Landslide mapping from multi-sensor data through improved change detection-based Markov random field [J]. *Remote Sensing of Environment*, 2019, 231: 111235.
- [16] Pawluszek K, Marczak S, Borkowski A, et al. Multi-Aspect Analysis of Object-Oriented Landslide Detection Based on an Extended Set of LiDAR-Derived Terrain Features [J]. *Isprs International Journal of Geo-Information*, 2019, 8(8): 321.
- [17] Bacha, Werff V D, Shafique, et al. Transferability of object-based image analysis approaches for landslide detection in the Himalaya Mountains of northern Pakistan [J]. *International Journal of Remote Sensing*, 2020, 41(9): 3390-3410.
- [18] LI Qilun, ZHANG Wanchang, YI Yaning. Identification and analysis of secondary geological hazards trig-gered by a magnitude 8.0 Wenchuan Earthquake[J]. *Journal of University of Chinese Academy of Sciences*, 2020, 37(1): 93-102. (李麒崙, 张万昌, 易亚宁. 地震滑坡信息提取方法研究——以 2017 年九寨沟地震为例 [J]. *中国科学院大学学报*, 2020, 37(1): 93-102.)
- [19] Petschko H, Bell R, Glade T. Effectiveness of visually analyzing LiDAR DTM derivatives for earth and debris slide inventory mapping for statistical susceptibility modeling [J]. *Landslides*, 2016, 13(5): 857-872.
- [20] XU Chong, DAI Fuchu, CHEN Jian, et al. Identification and analysis of secondary geological hazards trig-gered by a magnitude 8.0 Wenchuan Earthquake[J]. *Journal of remote sensing*, 2009, 13(4): 754-762. (许冲, 戴福初, 陈剑等. 汶川 Ms8.0 地震重灾区次生地质灾害遥感精 细解译[J]. *遥感学报*, 2009, 13(4): 754-762.)
- [21] Danneels G, Pirard E, Havenith H B. Automatic landslide detection from remote sensing images using supervised classification methods[C], *IEEE International Geoscience & Remote Sensing Symposium*, Barcelona, Spain, 2007.
- [22] Jones J W, Desmond G B, Glover C H, et al. An approach to regional wetland digital elevation model development using a differential global positioning system and a custom-built helicopter-based surveying system [J]. *International Journal of Remote Sensing*, 2012, 33(2): 450-465.
- [23] Cao Y, Yin K, Zhou C, et al. Establishment of Landslide Groundwater Level Prediction Model Based on GA-SVM and Influencing Factor Analysis [J]. *Sensors (Basel, Switzerland)*, 2020, 20(3): 845.
- [24] Zhu C H, Hu G D. Time Series Prediction of Landslide Displacement Using SVM Model: Application to Baishuihe Landslide in Three Gorges Reservoir Area, China [J]. *Applied Mechanics & Materials*, 2013, 239-240:1413-1420.
- [25] Zhang K, Wu X, Niu R, et al. The assessment of landslide susceptibility mapping using random forest and decision tree methods in the Three Gorges Reservoir area, China [J]. *Environmental Earth Sciences*, 2017, 76(10): 405.
- [26] Stumpf A, Kerle N. Object-oriented mapping of landslides using Random Forests [J]. *Remote Sensing of Environment*, 2011, 115(10): 2564-2577.
- [27] Mezaal M R, Pradhan B, Rizeei H M. Improving Landslide Detection from Airborne Laser Scanning Data Using Optimized Dempster–Shafer [J]. *Remote Sensing*, 2018, 10(7): 1029.
- [28] Wang F, Jiang M, Qian C, et al. Residual Attention Network for Image Classification [C]. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017.
- [29] Zhao Z, Zheng P, Xu S, et al. Object Detection With Deep Learning: A Review [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 30(11): 3212-3232.

- [30] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation [M]//NAVAB N, HORNEGGER J, WELLS W M, et al. Medical Image Computing and Computer-Assisted Intervention, Pt Iii. 2015: 234-241.
- [31] Shi W, Zhang M, Ke H, et al. Landslide Recognition by Deep Convolutional Neural Network and Change Detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 1-19.
- [32] Ghorbanzadeh O, Blaschke T, Gholamnia K, et al. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection [J]. Remote Sensing, 2019, 11(2):196.
- [33] Yu B, Chen F, Xu C. Landslide detection based on contour-based deep learning framework in case of national scale of Nepal in 2015 [J]. Computers & Geoences, 2019, 135:104388.
- [34] Wang H, Zhang L, Yin K, et al. Landslide identification using machine learning [J]. Geoence Frontiers, 2020, 08(07): 9705-9726.
- [35] Liu P, Wei Y, Wang Q, et al. Research on Post-Earthquake Landslide Extraction Algorithm Based on Improved U-Net Model [J]. Remote Sensing, 2020, 12(5): 894.
- [36] Fukushima K, Miyake S, Ito T. Neocognitron: A neural network model for a mechanism of visual pattern recognition [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1983, SMC-13(5): 826-834.
- [37] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C], 2014 IEEE Conference on Computer Vision and Pattern Recognition, OH, USA, 2014.
- [38] Girshick R. Fast R-CNN [C], 2015 IEEE International Conference on Computer Vision (ICCV), Boston, MA, 2015.
- [39] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [40] 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.
- [41] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection [C]. Computer Vision and Pattern Recognition, Seattle, WA, USA, 2020.
- [42] Ji S, Yu D, Shen C, et al. Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks [J]. Landslides, 2020, 17:1337–1352.

# Landslide Detection and Segmentation Using Mask R-CNN with Simulated Hard Samples

Jiang Wandong<sup>1,2</sup>, Xi Jiangbo<sup>1,2,3</sup>, Li Zhenhong<sup>1,2,3</sup>, Ding Mingtao<sup>1,2,3</sup>, Yang Ligong<sup>1</sup>,  
Xie Dashuai<sup>1,2</sup>

1. College of Geological Engineering and Geomantics, Chang'an University, Xi'an, Shaanxi 710054, China

2. Big Data Center for Geosciences and Satellites, Chang'an University, Xi'an, Shaanxi 710054, China

3. Key Laboratory of Western China's Mineral Resources and Geological Engineering, Ministry of Education, Xi'an, Shaanxi 710054, China

**Abstract:** With the advance in artificial intelligence, using high-resolution images to detect geological hazards has gradually become a research hotspot. Visual interpretation of landslides heavily relies on expert experience, and conventional automatic landslide detection approaches are sensitive to the presence of bare land, roads and other ground objects. To address these, a Mask R-CNN with simulated hard samples is presented in this paper for landslide detection and segmentation. Based on existing landslide samples, hard landslide samples are simulated by utilizing the shapes, colors, textures, and other characteristics of landslides to make each of the samples with a more complicated background. The original imagery and simulated hard samples are then fed into the Mask R-CNN for landslide detection and segmentation. Since the number of landslides is often limited in reality, small sample learning in the frequency domain is also presented in this paper to reduce the number of input samples while ensuring the accuracy of detection and segmentation. The experimental results in Bijie, Guizhou Province, showed that the detection and the average pixel segmentation accuracies of the proposed Mask R-CNN method with simulated hard samples are 94.0% and 90.3%, respectively. It is seen that the proposed method has high performance on landslide detection and segmentation with low false alarm rates. In addition, the performance of the proposed small-sample-based learning method in frequency domain can be improved even with a half of the data input. The effectiveness of the proposed Mask R-CNN method is further proved by the successful detection of Tianshui landslides in Gansu Province.

**Key Words:** Deep learning; Landslide detection; Mask R-CNN; Hard samples; Frequency domain

Foundation support: Major Program of the National Natural Science Foundation of China (41941019), The National Key Research and Development Program of China (2018YFC1504805), The National Natural Science Foundation of China(61806022, 41874005), The Fundamental Research Funds for the Central Universities (300102260301/087, 300102260404/087, 300102269103, 300102269304, and 300102269205), the State Key Laboratory of Geographic Information (SKLGIE2018-M-3-4).

First author: Wandong Jiang, Master, focuses on intelligent interpretation of geological hazards in remote sensing images, CHD\_jwd@126.com.

Corresponding author: Jiangbo Xi, PhD, focuses on machine learning, intelligent analysis of remote sensing images, xijiangbo@chd.edu.cn.

## 网络首发：

标题：模拟困难样本的 Mask R-CNN 滑坡分割识别

作者：姜万冬，席江波，李振洪，丁明涛，杨立功，谢大帅

收稿日期：2020-12-24

DOI:10.13203/j.whugis20200692

## 引用格式：

姜万冬，席江波，李振洪，等. 模拟困难样本的 Mask R-CNN 滑坡分割识别[J]. 武汉大学学报·信息科学版, 2021, DOI: 10.13203/j.whugis20200692 (Jiang Wandong, Xi Jiangbo, Li Zhenhong, et al. Landslide Detection and Segmentation Using Mask R-CNN with Simulated Hard Samples [J]. Geomatics and Information Science of Wuhan University, 2021, DOI: 10.13203/j.whugis20200692)

网络首发文章内容和格式与正式出版会有细微差别，请以正式出版文件为准！

## 您感兴趣的其他相关论文：

线状特征约束下基于四元数描述的 LiDAR 点云配准方法

王永波，杨化超，刘燕华，等.

武汉大学学报·信息科学版, 2013, 38(9):1057-1062

<http://ch.whu.edu.cn/article/id/2751>

利用单位对偶四元数进行航空影像区域网平差解算

龚辉，姜挺，江刚武，等.

武汉大学学报·信息科学版, 2012, 37(2):154-159

<http://ch.whu.edu.cn/article/id/112>