doi:10.6041/j.issn.1000-1298.2025.09.043

基于深度学习和地理分析的淤地坝遥感识别

孙立全[1]　郭家龙[2]　苑紫岩[2]　冯　浩[3]　吴淑芳[2]

(1.新疆大学地理与遥感科学学院,乌鲁木齐 830017;
2.西北农林科技大学旱区农业水土工程教育部重点实验室,陕西杨凌 712100;
3.西北农林科技大学水土保持科学与工程学院,陕西杨凌 712100)

摘要:准确获取淤地坝数目、位置及空间分布等信息是科学分析淤地坝减蚀和拦沙作用及其未来建设规划的重要基础。本研究以黄土高原延河流域淤地坝为研究对象,以高分二号影像以及 Google Earth 影像作为数据源分别制作坝体和坝地标签,选取深度学习目标检测(Faster R-CNN、YOLO v3、Cascade R-CNN 和 YOLOX)及语义分割(FCN、U-Net、PSPNet 和 DeepLab v3+)算法分别对淤地坝坝体及坝地进行识别提取,并结合地理分析方法实现流域尺度沟道缓冲区框定,在缓冲区内进行淤地坝提取。结果表明:在淤地坝坝体识别中,YOLOX 模型表现最佳,当交并比(IoU)阈值为 0.50~0.95 时,平均精度(AP)达 69.4%。基于该模型对延河流域影像进行识别,获得 1 440 个检测框,模型准确率为 75.8%,召回率达 85.7%。在淤地坝坝地提取中,4 个模型表现均较好,平均交并比、精确率、召回率及 F1 分数均达 0.85、90%、85% 和 88% 以上。采用多数投票法结合 4 个模型预测结果分别提取了流域内 92.0 km^2生产运行期坝地和 10.6 km^2蓄水拦泥期坝地。通过引入地理分析方法,坝体识别准确率从 75.8% 提升至 84.5%,识别提取效率提升 1 倍以上。该研究为黄土高原及其他地区的淤地坝快速准确识别提供了重要依据,淤地坝监测结果为流域综合治理和水土保持规划提供了科学借鉴和理论指导。

关键词:淤地坝;深度学习;遥感;延河流域

中图分类号:S157　　**文献标识码**:A　　**文章编号**:1000-1298(2025)09-0526-10　　**OSID**:

Remote Sensing Identification of Check Dams Based on Deep Learning and Geographic Analysis

SUN Liquan[1]　GUO Jialong[2]　YUAN Ziyan[2]　FENG Hao[3]　WU Shufang[2]

(1. *College of Geography and Remote Sensing Sciences, Xinjiang University, Urumqi* 830017, *China*
2. *Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas, Ministry of Education, Northwest A&F University, Yangling, Shaanxi* 712100, *China*
3. *College of Soil and Water Conservation Science and Engineering, Northwest A&F University, Yangling, Shaanxi* 712100, *China*)

Abstract: Accurately obtaining information on the number, location, and spatial distribution of check dams is fundamental for scientifically analyzing their erosion reduction and sediment trapping effects, as well as for future construction planning. Focusing on the Yanhe River Basin in the Loess Plateau, utilizing high-resolution imagery from GF-2 and Google Earth as data sources, deep learning object detection algorithms (Faster R-CNN, YOLO v3, Cascade R-CNN, and YOLOX) and semantic segmentation algorithms (FCN, U-Net, PSPNet, and DeepLab v3+) were employed to identify and extract the check dam bodies and dam lands. Geographic analysis methods and comprehensive interpretation were used to optimize the extraction results of the check dams. The results were as follows: in the identification of check dam bodies, the YOLOX model performed the best, achieving an average precision (AP) of 69.4% at an intersection of union (IoU) threshold of 0.50~0.95. Using this model to identify images of the Yanhe River Basin, totally 1 440 detection boxes were obtained, with an accuracy of 75.8% and a recall rate of 85.7%. In the extraction of dam lands, all four models performed well, with an average IoU, precision, recall, and F1 score all exceeded 0.85, 90%, 85%,

收稿日期:2024-06-05　修回日期:2024-06-30
基金项目:国家重点研发计划项目(2023YFD1900300)
作者简介:孙立全(1993—),男,副教授,主要从事农业水土资源利用与保护研究,E-mail:liquan_sun@sina.com
通信作者:吴淑芳(1977—),女,研究员,博士生导师,主要从事土壤侵蚀过程研究,E-mail:wsfjs@163.com

and 88%, respectively. The majority voting method was used to combine the predictions of the four models, resulting in the extraction of 92.0 km^2 of dam lands in the production and operation period and 10.6 km^2 in the water storage and sediment trapping period. By incorporating geographic analysis methods, the accuracy of check dam body identification was increased from 75.8% to 84.5%, with the efficiency of identification and extraction doubling. These results indicated that the proposed method can be rapidly and accurately applied to the identification of check dams in the Loess Plateau, which was crucial for analyzing their impact on the ecological environment and for comprehensive management of small watersheds.

Key words: check dam; deep learning; remote sensing; Yanhe River Basin

0　引言

淤地坝是一项在多泥沙沟道中用于减蚀固沟、蓄洪滞洪和控制泥沙输移修建的水土保持工程措施[1-2]，对黄土丘陵沟壑区拦泥、蓄水、淤地以及发展农业生产和改善生态环境等方面发挥着至关重要的作用。根据第一次全国水利普查水土保持情况公报显示，截至 2011 年黄土高原共有淤地坝 58 446 座，淤地面积 927.6 km^2。尽管淤地坝建设取得了一定成果，但仍存在诸多问题。早期修建的淤地坝大多为群众自发建设的小型水利工程，由于缺乏合理规划，部分已淤满失去作用[3]，还有相当多的淤地坝没有泄洪设施[4]。随着全球气候变化，部分地区由于极端暴雨引起连续垮坝，对下游造成较大洪水灾害[5]。因此，对淤地坝进行动态监控十分有必要。同时，准确获取淤地坝位置及坝地面积空间分布等信息是科学分析淤地坝减蚀和拦沙作用及其未来建设规划的基础[6]。随着遥感技术、人工智能特别是计算机视觉领域深度学习、机器学习方法的快速发展，为高分辨率遥感图像分析带来了新机遇。从遥感图像中提取地物成为可能，为淤地坝空间和光谱属性提供宝贵和及时的信息。杨蕾[7]采用基于像素的方法（监督分类）从高分辨率遥感图像中提取坝地。侯群群[8]考虑坝地和水体部分的纹理特征，利用面向对象的方法对坝地进行了提取。ALFONSO-TORRENO 等[9]利用无人机（UAV）拍摄的高分辨率航空图像对淤地坝进行了识别，并估算了这些淤地坝中沉积的泥沙体积。LI 等[10]将 OBIA 和 U－Net 深度学习语义分割模型相结合，检测了淤地坝区域。ZENG 等[11]将高分辨率和易于获取的 Google Earth 图像与基于对象的分类方法相结合，在黄土高原上生成了淤地坝矢量化数据集。

综上所述，国内外学者在淤地坝提取方面取得了显著成绩，但仍存在不足之处。这些研究大多集中于淤地坝坝地的提取，且方法单一，准确率有待提高；此外，对于淤地坝具体位置的识别研究较少，空间分析也不够详尽。基于此，本研究利用遥感影像，结合深度学习和地理空间分析等方法，对黄土高原丘陵沟壑区（延河流域）的淤地坝坝体及坝地进行识别与提取，并基于提取结果对流域内的淤地坝空间分布进行分析，为流域的科学管理及淤地坝生态系统服务功能评价提供基础信息，并为黄河输沙变化和水土保持工程的未来规划提供科学借鉴。

1　研究区概况与数据处理

1.1　研究区概况

以位于黄土高原丘陵沟壑区的延河流域为例（36°22′～37°20′ N，108°39′～110°29′ E）。延河干流总长约 286.9 km，主要支流包括杏子河、西川河、蟠龙川，流域面积 7 725 km^2，平均海拔 1 216 m。该地域属典型的大陆性半干旱季风气候，多年平均气温为 8.8～10.2℃，年平均降水量为 500～550 mm。降水量随季节变化且极不均匀，其中 6—9 月降雨量占全年降雨量的 75% 以上，且降雨多呈短历时、高强度暴雨[12]。流域内水土流失严重，河流含沙量大，土壤类型主要为黄绵土，抗侵蚀能力差，极易受降雨径流分散和搬运。

自 20 世纪 50 年代以来，在延河流域开展了大量的水土保持措施建设，主要包括沟道工程措施（打淤地坝）和坡面措施（退耕还林还草）。淤地坝建设大大降低了土壤侵蚀强度，并截留了数千吨泥沙，从而显著减小甘谷驿水文站的泥沙输移量[13]。根据 2011 年第一次全国水利普查数据、2015 年延安市宝塔区及延长县淤地坝普查资料和 2018 年年底延安市水土保持措施情况数据，标记了延河流域内及周边范围内骨干坝（大型）及中型坝约 600 座（图 1）。根据文献综述及调查数据显示，目前延河流域共有 1 100 多座淤地坝，包括骨干坝 288 座和中型坝 803 座。就建坝时间而言，延河流域内的淤地坝主要建设于 20 世纪 60—70 年代，在 80—90 年代，建坝数量显著减少。然而，2003 年淤地坝工程实施期间，建坝数量有所回升。

1.2　遥感数据及其预处理

所用数据包括高分二号（GF－2）影像数据（中

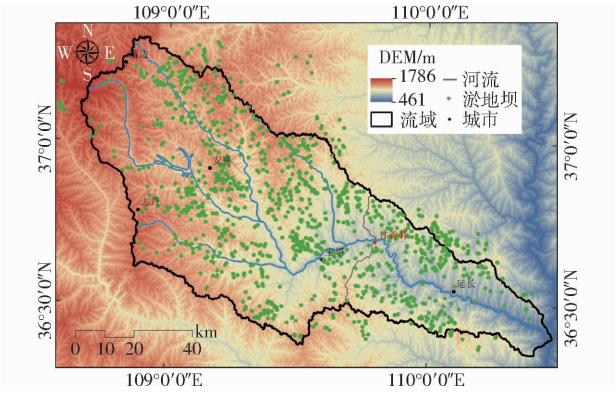


图 1 研究区概况图

Fig. 1 Study area

国资源卫星应用中心:http://www.cresda.com/CN/Satellite/3128.shtml)、Google Earth 影像数据以及 ASTER GDEMV3 数字高程模型(中国科学院计算机网络信息中心地理空间数据云平台:http://www.gscloud.cn)。

GF-2 卫星是我国自主研发的民用光学遥感卫星,搭载 2 台 PMS(多光谱/全色)可见光传感器,星下点空间分辨率为 0.8 m。全色波段(450~900 nm)空间分辨率为 1.0 m,4 个多光谱波段(蓝(450~520 nm)、绿(520~590 nm)、红(630~690 nm)、近红外(770~890 nm))空间分辨率为 4.0 m。为降低积雪对淤地坝识别的影响,同时考虑到数据质量,本试验选取 2020 年 4—11 月期间覆盖研究区不同季节的 GF-2 影像数据作为主要数据源,各景影像云量覆盖率均低于 5%。采用分辨率为 30 m 的 ASTER DEM 数据对 GF-2 数据进行正射校正,基于野外调研得到的地面控制点(Ground control point,GCP)及同期分辨率 1 m 的全色波段数据作为参考影像对 GF-2 多光谱影像进行几何精校正,然后利用 Gram-Schmidt 融合方法将全色波段和多光谱数据融合,得到分辨率 1 m 多光谱影像。此外,还选取部分空间分辨率为 0.3 m 和 1.0 m 的 Google Earth 影像数据作为补充数据,用于 GF-2 影像数据未覆盖到的区域。Google Earth 影像数据由 Google Earth 平台获取,这些影像数据包括由不同卫星和航空器拍摄的高分辨率图像,分辨率从几十厘米到几米不等,覆盖全球各地。此外,Google Earth 还提供许多地点的历史影像,允许用户查看不同时间点的图像,以进行时间序列分析。使用多源遥感数据还能克服单一数据源图像数量少、图像多样性和目标变异性不足等缺点,从而提高模型鲁棒性和泛化能力。

2 研究方法

技术路线如图 2 所示,主要是基于深度学习目标检测和语义分割算法与地理分析相结合,实现流域尺度的淤地坝精确提取。目标检测算法是计算机视觉中的一个重要分支,旨在在图像中定位并识别特定目标。在淤地坝坝体识别中,目标检测算法能在复杂的遥感图像中准确定位和识别淤地坝坝体。语义分割算法则用于对图像进行像素级分类,以提取特定区域的详细信息。在淤地坝坝地提取中,语义分割算法能精确提取出坝地的具体范围。具体过程包括:使用高分二号影像和 Google Earth 影像作为数据源,构建并评估多种目标检测和语义分割算法(如 Faster R-CNN、YOLO v3、U-Net、Deeplab v3+等),选择最佳算法进行淤地坝识别和提取;然后,将地形数据作为约束,通过 ASTER DEM 数据进行地形分析和沟道线提取,结合先验知识选择淤地坝候选区,最终综合深度学习和地理分析结果,对试验结果进行评价,以实现高精度的淤地坝识别和提取。

2.1 数据集构建

影像中淤地坝形态特征比较简单,最明显且易被识别的为坝体和坝地。坝体通常呈矩形或类矩形,少数呈类三角形;坝顶常发挥路桥作用,用于连接沟道两岸交通,呈线状特征。坝地处于坝体后方,由坝体拦截流域内泥沙淤积而成,和周围地形相比,坝地地势平缓,表面平整。

2.1.1 淤地坝坝体标签制作

在 ArcGIS Pro(版本 3.0,https://www.arcgis.com)中添加卫星影像,根据前文所述的淤地坝调查数据,对淤地坝进行标注,标注对象命名为“Check dam”,使用“导出训练数据进行深度学习”工具制作并导出淤地坝目标图像及标签文件,导出图像尺寸为 1 024 像素×1 024 像素,标注文件导出为 Pascal VOC 数据集格式,名称与图像命名一一对应,从不同数据源影像中共制作样本 1 326 个。

为减少网络过拟合现象,得到泛化能力更强的网络,更好适应应用场景,利用几何变换和色彩变换方法对样本图像进行数据增强,扩充样本量。采用样本图像旋转、翻转、平移、缩放以及模糊处理、锐化处理、添加噪声、局部扭曲和色彩抖动(包括对图像饱和度和对比度进行调整)等随机组合对制作的样本图像进行数据增强(图 3),每幅图像标注文件也同步变换。最终得到增强后的样本 12 988 个,按照训练验证集和测试集比例为 8∶2 将增强后数据集进行随机划分,得到训练验证集样本 10 392 个,测试集样本 2 596 个,又将训练验证集样本按照比例 8∶2 划分为训练集和验证集。训练集用于模型训练,验证集用于模型调参,测试集用于评估模型性能。

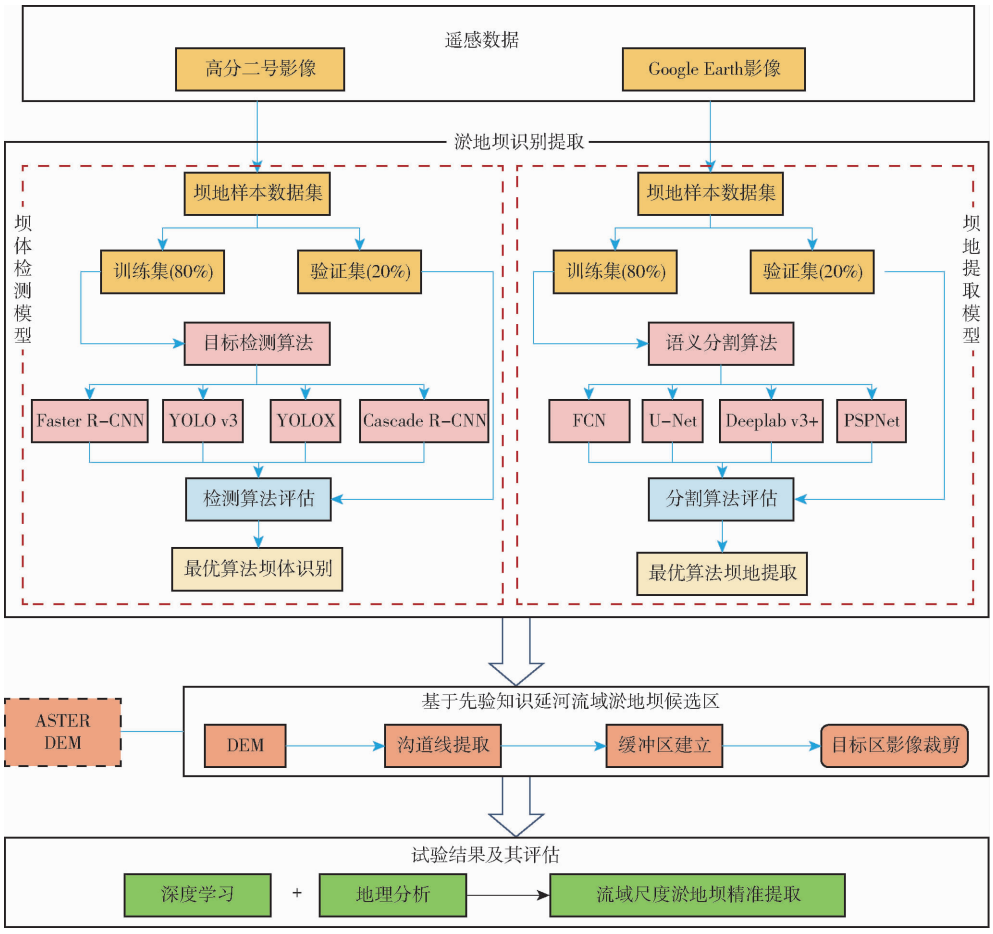


图 2　技术路线图

Fig. 2　Technology roadmap

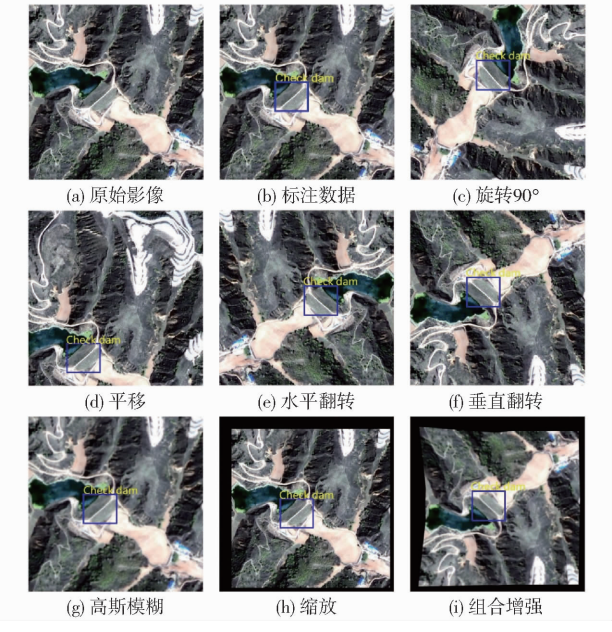


图 3　样本标注及数据增强效果示例

Fig. 3　Sample labeling and data augmentation effect examples

2.1.2　淤地坝坝地标签制作

根据淤地坝原始影像数据，选取 787 处坝地在 ArcGIS Pro 进行标注，设置蓄水拦泥期淤地坝区域值为 1，生产运行期淤地坝区域值为 2，非淤地坝区域值为 0。接着，使用 ArcGIS Pro 中的“导出训练数据进行深度学习”工具，将标注的矢量数据转换为深度学习栅格数据集，包括 TIFF 格式图像和相应的标注图像，尺寸为 512 像素 ×512 像素。为减少网络的过拟合现象，得到泛化能力更强的网络，采用旋转、翻转和平移等几何变换方法对样本图像进行数据增强，得到样本集 5 261，按照训练集和验证集比例为 8∶2 将增强后数据集进行随机划分。此外，预留了 85 处坝地用于制作测试集样本，共含影像 609 幅。淤地坝坝地及其对应标签见图 4。

2.2　深度学习模型

2.2.1　基于深度学习目标检测算法的坝体识别模型

选择深度学习典型目标检测算法 Faster R－CNN 和单阶段算法 YOLO v3、在其基础上改进的 Cascade R－CNN 和 YOLOX [14－17] 进行淤地坝坝体识别。

Faster R－CNN 通过区域生成网络（RPN）生成候选框，并通过二阶段的分类和回归进行精确定位和识别。Cascade R－CNN 是对 Faster R－CNN 的改进。它在第 1 阶段与 Faster R－CNN 相同，但在第 2 阶段通过多阶段的级联结构，提高检测准确性和鲁棒性。针对淤地坝在复杂遥感影像背景下的图像特

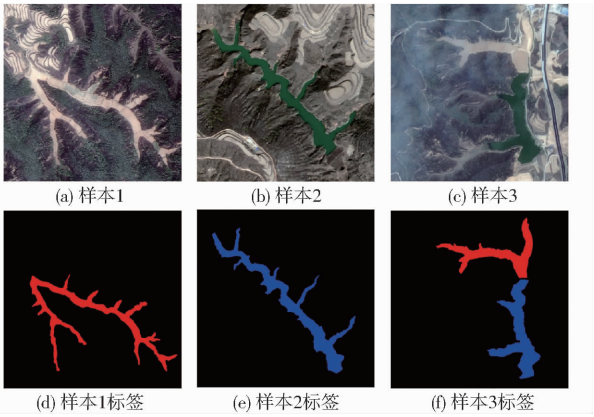


图 4 淤地坝坝地样本及其对应标签

Fig. 4 Dam area samples and their corresponding labels

征,选取 ResNeXt－101 作为骨干网络。同时为解决淤地坝检测中的多尺度问题,加入 FPN(Feature pyramid networks)网络,提升淤地坝检测性能。

YOLO 系列算法将目标检测任务转化为回归问题,通过单个卷积神经网络(CNN)处理整个图像,实现高效处理。YOLO v3 采用改进的 Darknet53 网络作为特征提取器,将图像缩放到 608 像素×608 像素,生成 19×19、38×38、76×76 共 3 个尺度特征图,并借鉴 FPN 网络的特征金字塔思想,通过 2 倍上采样传递特征图。每个网格配置 3 个不同尺度的边界框进行预测,判断边界框内是否包含物体,并调整边界框获得预测结果。YOLOX 于 2021 年由北京旷视科技有限公司开源,基于 YOLO v3 框架,集成解耦头、无锚点和 SimOTA 动态正样本匹配方法。YOLOX 采用 CSPDarknet 作为骨干网络,通过跨阶段部分连接(CSP)减少计算量,同时保持强大的特征表达能力;解耦头设计将分类和回归任务分开处理,优化每个任务的特定目标,从而提高整体检测性能;无锚框机制通过中心点检测和动态分配正样本,减少锚框设计带来的复杂性和计算开销;SimOTA 策略根据目标和预测之间的距离动态调整正样本分配,提高检测精度和鲁棒性。凭借高检测精度和速度,YOLOX 特别适用于大规模遥感图像的实时处理。

2.2.2 基于深度学习语义分割算法的坝地提取模型

考虑到 FCN、U－Net、PSPNet 和 DeepLab v3＋都是当前语义分割任务中的代表性深度学习架构[18－21],并在多个领域中表现出色,选择这 4 种模型进行进一步研究。尽管这些网络都旨在实现图像分割,但由于它们在网络架构和技术策略上的差异,在不同类型遥感数据和场景中的性能和适用性存在差异。因此,为评估这些网络在光学遥感图像与数字高程模型(DEM)数据组合上的坝地提取能力,进行了详细的试验对比。这些网络在遥感领域的其他分割任务中已经证明了其出色的性能,但在大尺度淤地坝坝地的提取任务中,应用仍然相对较少。

FCN 是第 1 个全卷积神经网络,能够处理任意尺寸图像,并实现像素级分类[22]。U－Net 专为生物医学图像分割设计,具有对称的编码器-解码器结构,通过跳跃连接融合多尺度信息,特别适用于精细的图像分割任务。

DeepLab v3＋结合前代 DeepLab 优势,由编码器和解码器组成。编码器采用空洞卷积捕获高级特征,解码器逐步恢复空间信息,描绘对象细致边界。空洞卷积扩大感受野,不降低分辨率,实现高质量分割和精确定位[23]。空间金字塔池化模块允许动态调整感受野大小,适应各种尺寸和形状的目标。但该模型复杂,需要较多训练数据和较长训练周期,计算资源受限时不一定最佳。

PSPNet(Pyramid scene parsing network)为解决语义分割中的全局上下文信息整合问题设计[24]。其金字塔池化模块(PPM)整合不同尺度的上下文信息,通过池化和上采样实现局部与全局信息融合。PSPNet 在处理多样化目标时表现出色,但其复杂性和高计算需求在资源受限场景下可能构成挑战。

2.3 模型评价指标

2.3.1 目标检测模型评价指标

使用精确率(Precision)、召回率(Recall)、准确率-召回率(P－R)曲线、平均精度(AP)和推理速度(FPS)来评估各算法性能[25－26]。在算法评估时,如果检测框和地面真值标注框之间的面积交并比(IoU)超过预定义阈值 λ,则检测框被记为真正例,否则该检测框被记为假正例。P－R 曲线显示不同 IoU 阈值下的精确率和召回率;平均精度(AP)为 P－R 曲线与坐标轴所围面积,因此 AP 越高,性能越好,本研究目标检测类别为 1,因此 AP 等于 mAP。FPS 为每秒检测到的图像数或检测每幅图像的时间。

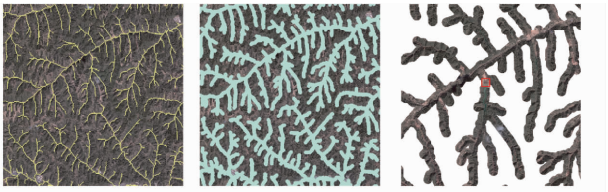
2.3.2 语义分割模型评价指标

基于混淆矩阵,分别计算精确率(Precision)、召回率(Recall)、F1 分数(F1-score)和交并比(IoU)来评估算法性能。

2.4 基于地理分析的淤地坝缓冲区提取

根据淤地坝是建设于沟道中的位置特征,通过提取沟道线及建立沟道线缓冲区框定淤地坝识别目标区范围,对深度学习算法提取的结果进一步筛选(图 5),可有效减少错分,提高淤地坝提取精度。应用 ArcGIS 软件水文分析模块基于分辨率 30 m DEM 提取研究区沟网,流程包括无洼地 DEM 生成、水流流向提取、汇流累积量计算和沟道线提取。在野外调研和内业分析的基础上,结合淤地坝分布和规模

确定能覆盖研究区淤地坝的沟道宽度，从而设定沟道线的缓冲距离，根据提取的沟道线向两侧缓冲，形成面状缓冲区。经测试分析，设定汇流累积量阈值为 200、缓冲距离为 135 m 时生成的缓冲区可使研究区内的淤地坝基本实现全覆盖。



(a) 沟道线提取　(b) 缓冲区建立　(c) 淤地坝识别目标区确定

图 5　淤地坝候选区提取

Fig. 5　Candidate area demarcation for check dam recognition

2.5　试验平台及参数设置

所有模型均在计算机上搭建的深度学习环境中运行，计算机硬件配置为 Intel(R) Core(TM) i7 - 7700 处理器，主频 3.60 GHz，运行内存(RAM) 32 GB，显卡(GPU)为 NVIDIA Tesla P100，显存 16 GB。操作系统为 Ubuntu 18.04 LTS，安装 CUDA(V11.1.105)和 CUDNN(8.0.5)以进行 GPU 加速，试验使用编译语言为 Python 3.7.12，深度学习框架为 Pytorch 1.7.1。

3　结果与分析

3.1　基于深度学习目标检测算法的淤地坝坝体识别

为验证模型识别精度，将训练好的模型应用于测试集，图 6 为 4 个网络在不同 IoU 阈值时精确率和召回率间的关系曲线，曲线下方面积为不同模型平均精度(表 1)。无论 IoU 阈值如何变化，YOLOX 曲线能完全包围住其他网络对应 P - R 曲线，表现最优，其次为 Cascade R - CNN 网络。当 IoU 阈值为 0.50 时，YOLO v3 算法明显优于 Faster R - CNN，但随着 IoU 阈值增加，其性能也明显降低，当 IoU 阈值为 0.75 时，YOLO v3 网络识别能力表现最差。

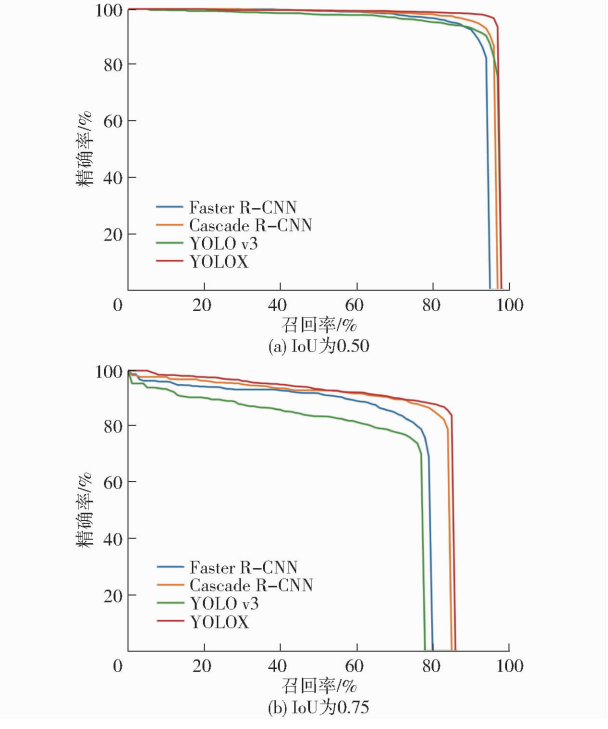


图 6　不同算法 P - R 曲线比较

Fig. 6　Comparison of P - R curves for different algorithms

表 1　不同模型淤地坝检测平均精度

Tab. 1　Average precision of check dam detection by different models

模型	主干网络	平均精度/%			检测速度/(f·s^{-1})
		IoU 为 0.50 ~ 0.95	IoU 为 0.50	IoU 为 0.75	
Faster R - CNN	ResNeXt - 101	61.5	92.5	72.3	4.2
Cascade R - CNN	ResNeXt - 101	67.0	94.9	78.5	3.9
YOLO v3	Darknet53	58.1	94.3	66.5	25.4
YOLOX	CSPDarknet	69.4	96.4	80.3	15.8

由表 1 亦可得到相似结果，YOLOX 网络 AP 达到 69.4%，检测精度最高，此外，还比较各个模型的推理速度。YOLO v3 推理速度最快，达 25.4 f/s，YOLOX 网络虽然增加了参数量，检测速度也能达到 15.8 f/s，而 Faster R - CNN 和 Cascade R - CNN 检测速度仅为 4.2、3.9 f/s。

总体而言，4 个模型在淤地坝检测上均取得了较好效果，但综合考虑检测精度和检测速度，选择 YOLOX。图 7 为择优挑选出 YOLOX 模型后，对延河流域 20 景高分二号影像进行了淤地坝识别，并保留置信度阈值大于 0.6 的结果，最终获得 1 440 个检测框(图 7)，基于淤地坝调查数据、野外调研以及 Google Earth 往期影像目视判断出 1 092 个检测框被正确识别为淤地坝，误判 200 多个，模型准确率为 75.8%，召回率达 85.7%；86 个淤地坝未被

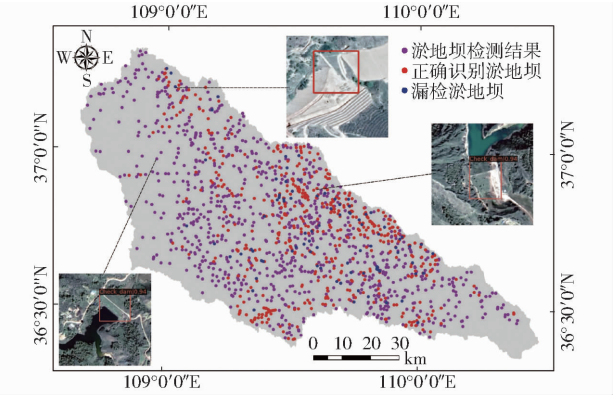


图 7　基于深度学习的延河流域淤地坝检测结果

Fig. 7　Results of check dam detection in Yan River Basin based on deep learning

识别,大多为小型淤地坝,在高分二号影像上光谱特征及纹理特征均并不明显,相关训练样本也较少。

3.2 基于深度学习语义分割算法的淤地坝坝地提取

为评价在训练过程中模型性能,详细分析了在不同波段组合下 4 种模型(FCN、DeepLab v3 +、PSPNet 和 U - Net)的损失函数 Loss 和平均交并比(mIoU)曲线(图 8)。所有训练数据迭代完一轮为一个训练周期,所有模型训练周期(即 epoch 数)均呈相似趋势,尤其在 FCN、DeepLab v3 + 和 PSPNet 3 种模型中,这可能是因为这些模型采用了相同的主干网络。在训练周期增加过程中,观察到所有模型损失值逐渐降低,同时 mIoU 逐渐提升。当训练周期达到一定数值后,各模型损失值和 mIoU 表现出稳定的趋势:①经过 300 个训练周期,迭代约 315 675 次时,U - Net 模型性能参数达到稳定状态。②在经过 350 个训练周期后,迭代约 368 288 次时,FCN、DeepLab v3 + 和 PSPNet 模型性能参数同样达到稳定。这些稳定的损失值(< 0.05)和 mIoU(>0.90)表明模型已达到收敛状态,从而证明这些模型适用于坝地提取。为进一步评估模型泛化能力,将从未用于训练或验证的测试数据集代入到训练好的模型进行评估(表 2)。PSPNet 在测试数据集上的坝地提取任务上均表现出较高的 IoU、召回率和 F1 分数。通过 IoU、召回率以及 F1 分数可以看出,深度学习模型在识别生产运行期淤地坝表现明显优于蓄水拦泥期淤地坝(表 2)。

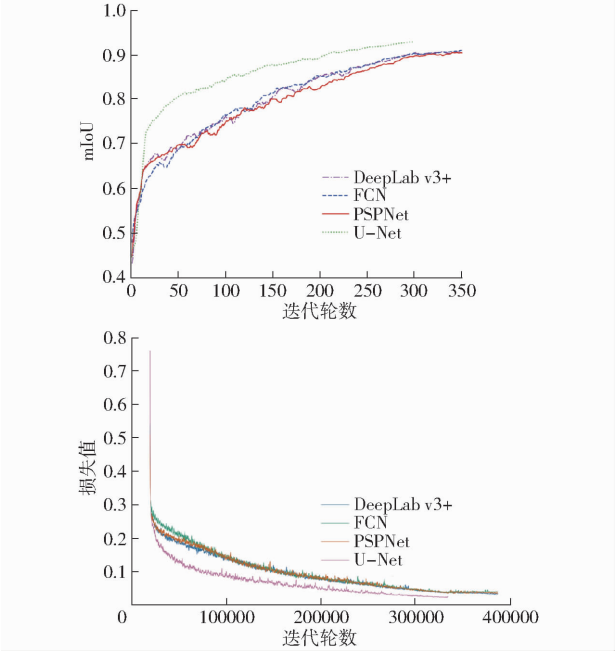


图 8 训练过程中模型平均交并比及损失值变化曲线

Fig. 8 Loss and mIoU variation in selected deep learning networks during training process

表 2 测试数据集上不同模型坝地提取精度评价

Tab. 2 Extraction accuracy of check dams using different CNN networks in test datasets

语义分割模型	交并比			精确率/%		召回率/%		F1 分数/%	
	平均值	生产运行期	蓄水拦泥期	生产运行期	蓄水拦泥期	生产运行期	蓄水拦泥期	生产运行期	蓄水拦泥期
FCN	0.856	0.820	0.771	90.4	90.5	89.8	83.9	90.1	87.1
U - Net	0.852	0.803	0.774	89.9	92.7	88.3	82.4	89.1	87.3
DeepLab v3 +	0.861	0.822	0.781	90.6	91.8	89.8	84.0	90.2	87.7
PSPNet	0.861	0.806	0.800	90.7	90.7	87.8	87.2	89.2	88.9

为定量评估各模型性能,对测试数据集中预测的淤地坝与相应的实地调查结果进行了可视化对比(图 9)。结果表明:各模型均能在测试图像中成功识别出大多数的淤地坝坝地边界与轮廓,平均 IoU 均达到 0.85 以上,为减少误识别像素并有效提高召回率,采用多数投票法[27]结合 4 个网络的预测输出。经过后处理,获得淤地坝的映射多边形。图 10 为基于本研究所构建的框架提取的延河流域淤地坝坝地结果。为进一步验证此方法的可靠性和准确性,采用无人机测量所得数据作为基准,并将其与不同季节下 3 个子区域的淤地坝提取结果进行对比分析,结果表明利用多数投票法结合 4 个模型的预测结果对大多数淤地坝的提取结果表现优异:在坝地边界方面,模型预测与实际测量值间误差小于 2.0 m;同样,在预测和测量的淤地坝面积间误差也在 5.9% 以内,这进一步印证了模型的高准确性。总体而言,淤地坝在整个研究区域内分布广泛。流域内生产运行期及蓄水拦泥期淤地坝分别淤积大约 92.0 km^2 和 10.6 km^2 坝地。

3.3 基于地理分析的淤地坝提取综合判读结果

为进一步提升淤地坝识别的准确性,引入地理分析对检测结果进一步确认。经过筛选,确定一个面积约为 3 969 km^2 候选区域,作为潜在的淤地坝所在区域。这一优化使得目标候选区相较于初始研究范围缩减约 50%,从而使得淤地坝提取更加集中于关键地区。

基于目标检测算法的坝体识别任务中,在 1 440 个检测框里最终筛选获得 1 289 个检测框作为高置信度淤地坝检测结果,其中正确识别淤地坝检测框为 1 089 个,有效筛除了约 10% 的错误检测框,淤地

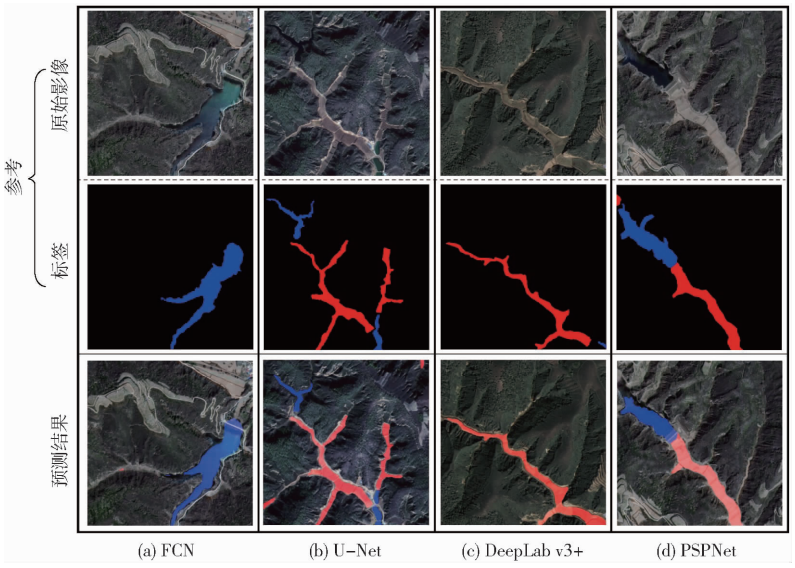


图 9　基于不同语义分割模型的淤地坝坝地提取结果

Fig. 9　Results of dam area extraction based on different semantic segmentation models

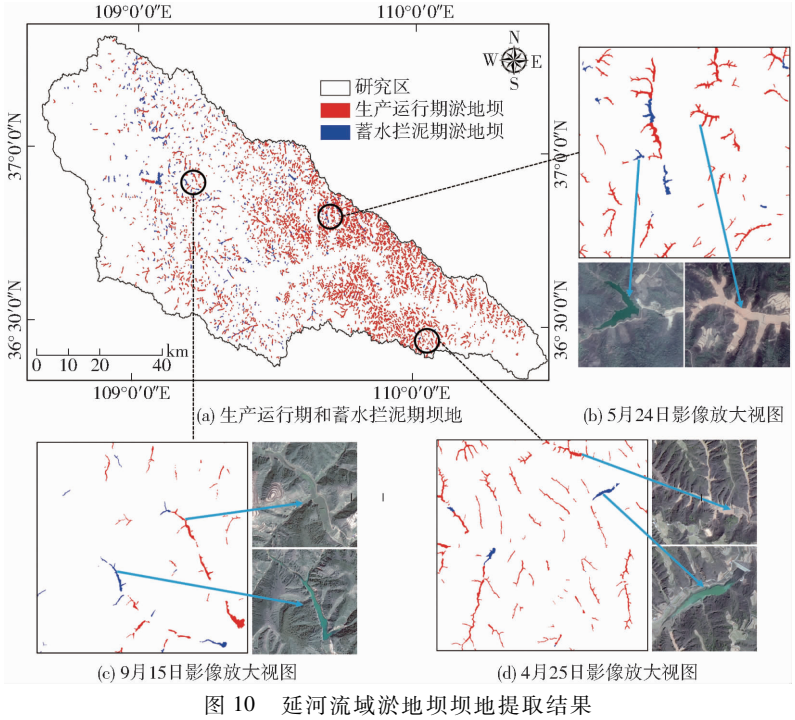


图 10　延河流域淤地坝坝地提取结果

Fig. 10　Check dam extraction results for Yan River watershed

坝识别准确率达 84.5%，召回率达 85.2%（表 3）。

表 3　地理分析后淤地坝坝体识别结果评价

Tab. 3　Evaluation of check dam bodies identification results after geographical analysis

真值数据集淤地坝数	研究方法	检测框	淤地坝正确识别数	淤地坝漏检数	召回率/%	精确率/%
602	淤地坝检测模型	1 440	1 092	86	85.7	75.8
	地理分析综合判读	1 289	1 089	89	85.2	84.5

基于语义分割算法的坝地提取任务中，通过对比试验可知空间分析提升了淤地坝地提取结果（表 4）。由表 4 可知，将候选区域作为预设的约束条件，在减少模型误分类的同时，在多个评估指标上均实现了显著提升。通过联合应用深度学习与地理空间分析，成功排除约 7.1 km^2 误检区域，并显著提高了对生产运行期和蓄水拦泥期淤地坝的识别准确率，分别提升 6.39、3.68 个百分点。更为重要的是，结合地理空间分析在保持或提高准确率的同时，计算时间节省了大约 40%，这在大规模或紧急应用场景中具有显著的实用价值。最终结果表明，共提取 106.05 km^2 坝地，其中包含生产运行期淤地坝 94.15 km^2 和蓄水拦泥期淤地坝 11.90 km^2。这与 ZENG 等[11]的提取结果类似，在延河流域内，其共

表 4 地理分析后淤地坝坝地提取结果评估

Tab. 4 Evaluation of check dam lands prediction after geospatial analysis

方法	生产运行期坝地		蓄水拦泥期坝地		汇总		时间/s
	预测面积/km²	精确率/%	预测面积/km²	精确率/%	预测面积/km²	精确率/%	
深度学习	100.74	91.22	12.41	85.40	113.15	90.58	24 535
空间分析	94.15	97.61	11.90	89.08	106.05	96.65	14 749

提取坝地 104.4 km²,但是其并未对不同坝地类型进行区分。

4 讨论

监控淤地坝最重要的问题之一是从影像等数据中将其识别。本研究提出的深度学习方法与空间分析相结合与野外调查或其他传统图像处理方法相比,具有更高的客观性、时效性和准确性,包括降低对专业人员目视解译的依赖以及高度的泛化能力[28]。在本研究中,发现 YOLOX 模型在淤地坝坝体识别中表现最佳,平均精度(AP)达到 69.4%,而 Cascade R-CNN 虽然在准确率上表现出色,但推理速度较慢,不适用于大规模实时监控。语义分割中,多数投票法结合多个模型的预测结果在坝地提取中表现优异,IoU 达 0.85 以上,表明其在不同阶段的坝地识别中具有高度稳定性和准确性。LI 等[10]也曾应用语义分割方法与面向对象相结合分析对淤地坝坝地提取的可行性,但本研究应用证实深度学习结合地理分析从大规模遥感图像中自动识别和提取淤地坝具有明显优势。通过引入地理分析方法,不仅提高了淤地坝识别准确率(提升 6.39 个百分点),还显著减少了计算时间(约 40%),使得该方法在大规模应用中具有实际价值。值得注意的是,这些深度学习网络在测试集上成功地识别了一些在标注过程中未被注意到的淤地坝,进一步证实了即便在训练数据存在不确定性的情况下,深度学习方法依然能有效地用于淤地坝提取,表明深度学习技术在补充现有淤地坝调查数据的重要性。进而,借助于卫星影像的持续更新和模型的自动化处理,几乎可以实时地进行淤地坝识别提取。相较于实地调查或其他传统方法,这种自动化策略需更少的资源和人力投入,因此能以更低的成本进行水土保持措施的动态监测。提取的淤地坝数据可用于支持与土壤侵蚀和模型构建相关研究[29],并有助于改进流域管理策略。

在实际应用中,高空间分辨率、大范围遥感影像以及来自背景信息的严重干扰将显著降低检测精度。在本研究中,结合地理分析等方法进一步提高识别结果的准确率。但所提方法在很大程度上依赖于 DEM 产品的精度,这些产品的集成会累积误差,这对淤地坝检测准确性有很大影响。未来研究可以考虑引入更高分辨率的 DEM 数据,以进一步提升算法在淤地坝识别和提取方面的精确度。同时,由于地域间地理、经济和文化差异,不同地区地表条件也不同,淤地坝建设材料以及数量等也会有很大差异。当使用所提出的地理分析方法综合判读时,应考虑地域间差异。

5 结论

(1)目标检测与语义分割算法均能有效识别提取淤地坝,YOLOX 和 PSPNet 分别在坝体识别和坝地提取中表现最佳,展示了深度学习算法在淤地坝检测中的高效性和准确性。

(2)地理空间分析显著提升了淤地坝提取精度和效率,结合地理分析方法,对检测结果进行优化,不仅提高了识别准确率和召回率,还显著减少了计算时间,展示了该方法在大规模应用中的实用价值,有助于指导淤地坝的规划和建设,更有效地防治土壤侵蚀。

参 考 文 献

[1] XU Y D, FU B J, HE C S. Assessing the hydrological effect of the check dams in the Loess Plateau, China, by model simulations[J]. Hydrology and Earth System Sciences, 2013,17(6):2185-2193.

[2] 李勉, 李平, 杨二, 等. 黄土丘陵区淤地坝建设后小流域泥沙拦蓄与输移特征[J]. 农业工程学报, 2017,33(18):80-86.
LI Mian, LI Ping, YANG Er, et al. Characteristics of sediment retention and transport in small watershed after construction of check dams in Loess Hilly Area[J]. Transactions of the CSAE, 2017, 33(18): 80-86. (in Chinese)

[3] JIN Z, CUI B, SONG Y, et al. How many check dams do we need to build on the Loess Plateau? [J]. Environmental Science & Technology, 2012,46(16):8527-8528.

[4] 高云飞, 郭玉涛, 刘晓燕, 等. 黄河潼关以上现状淤地坝拦沙作用研究[J]. 人民黄河, 2014,36(7):97-99.
GAO Yunfei, GUO Yutao, LIU Xiaoyan, et al. Amount of the intercepted sediment by the warping dams in the regions above Tongguan in the Yellow River basin[J]. Yellow River, 2014,36(7):97-99. (in Chinese)

[5] BAI L C, WANG N, JIAO J Y, et al. Soil erosion and sediment interception by check dams in a watershed for an extreme rainstorm on the Loess Plateau, China[J]. International Journal of Sediment Research, 2020,35(4):408-416.

[6] 史红艳. 黄土高原淤地坝防汛监控预警系统建设展望[J]. 中国防汛抗旱, 2019, 29(3): 16-19.

SHI Hongyan. Prospects for the construction of flood control monitoring and early warning system for warping dams on the Loess Plateau[J]. China Flood & Drought Management, 2019, 29(3): 16 – 19. (in Chinese)

[7] 杨蕾. 基于 Spot5 遥感影像提取水土保持信息的研究[D]. 西安：西北大学, 2006.
YANG Lei. A research on extraction soil and water conversation information based on Spot5 images[D]. Xi'an: Northwest University, 2006. (in Chinese)

[8] 侯群群. 面向对象的遥感影像水土保持措施信息提取研究[D]. 杨凌:西北农林科技大学, 2013.
HOU Qunqun. Soil conservation practices information extraction from remote sensing images using object-oriented method[D]. Yangling: Northwest A&F University, 2013. (in Chinese)

[9] ALFONSO-TORRENO A, GOMEZ-GUTIERREZ A, SCHNABEL S, et al. sUAS, SfM – MVS photogrammetry and a topographic algorithm method to quantify the volume of sediments retained in check-dams[J]. Science of the Total Environment, 2019, 678: 369 – 382.

[10] LI S J, XIONG L Y, HU G H, et al. Extracting check dam areas from high-resolution imagery based on the integration of object-based image analysis and deep learning[J]. Land Degradation & Development, 2021,32(7): 2303 – 2317.

[11] ZENG Y, JING T, XU B, et al. Vectorized dataset of silted land formed by check dams on the Chinese Loess Plateau[J]. Sci. Data, 2024, 11: 348.

[12] BAI M, MO X G, LIU S X, et al. Contributions of climate change and vegetation greening to evapotranspiration trend in a typical hilly-gully basin on the Loess Plateau, China[J]. Science of the Total Environment, 2019,657:325 – 339.

[13] WEI Y H, HE Z, JIAO J Y, et al. Variation in the sediment deposition behind check-dams under different soil erosion conditions on the Loess Plateau, China[J]. Earth Surface Processes and Landforms, 2018,43(9):1899 – 1912.

[14] REN S Q, HE K M, 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.

[15] REDMON J, FARHADI A. YOLOv3:an incremental improvement[EB/OL]. (2018 – 04 – 08) [2024 – 05 – 30]. https://arxiv.org/abs/1804.02767.

[16] CAI Z W, VASCONCELOS N. Cascade R – CNN:delving into high quality object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6154 – 6162.

[17] GE Z, LIU S T, WANG F, et al. YOLOX:exceeding YOLO series in 2021[J]. arXiv Preprint, arXiv:2107.08430, 2021.

[18] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640 – 651.

[19] RONNEBERGER O, FISCHER P, BROX T. U – net: convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-assisted Intervention-MICCAI 2015, 2015: 234 – 241.

[20] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6230 – 6239.

[21] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 801 – 818.

[22] PERSELLO C, STEIN A. Deep fully convolutional networks for the detection of informal settlements in VHR images[J]. IEEE Geosci. Remote Sens. Lett., 2017, 14(12): 2325 – 2329.

[23] HOESER T, KUENZER C. Object detection and image segmentation with deep learning on earth observation data: a review—Part I: evolution and recent trends[J]. Remote Sensing, 2020, 12(10): 1667.

[24] YU B, YANG L, CHEN F. Semantic segmentation for high spatial resolution remote sensing images based on convolution neural network and pyramid pooling module[J]. IEEE J. Sel. Topics. Appl. Earth Observ. Remote Sens., 2018, 11(9): 3252 – 3261.

[25] CHENG G, HAN J W. A survey on object detection in optical remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016,117:11 – 28.

[26] 郭瑞, 于翀宇, 贺红, 等. 采用改进 YOLOv4 算法的大豆单株豆荚数检测方法[J]. 农业工程学报, 2021,37(18): 179 – 187.
GUO Rui, YU Chongyu, HE Hong, et al. Detection method of soybean pod number per plant using improved YOLOv4 algorithm[J]. Transactions of the CSAE, 2021, 37(18): 179 – 187. (in Chinese)

[27] NARASIMHAMURTHY A. Theoretical bounds of majority voting performance for a binary classification problem[J]. IEEE Trans. Pattern. Anal. Mach. Intell., 2005, 27(12): 1988 – 1995.

[28] 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.

[29] BUSSI G, FRANCES F, HOREL E, et al. Modelling the impact of climate change on sediment yield in a highly erodible Mediterranean catchment[J]. J. Soils Sediments, 2014, 14(12): 1921 – 1937.