

基于卷积神经网络的遥感图像目标分类

岳启明^{1*}, 陈皖龙², 闫玉龙²

¹中国科学院文献情报中心, 中国科学院中国现代化研究中心, 北京

²太原理工大学数学学院, 山西 太原

收稿日期: 2023年12月23日; 录用日期: 2024年1月18日; 发布日期: 2024年1月23日

摘要

近年来, 大量对地观测遥感卫星被成功发射并在轨运行, 无人机等新型遥感平台也在不断发展更新, 传统的人工目视解译已不能满足遥感影像解译在效率和精度方面的需求, 深度学习在遥感影像处理方面表现出较好的可靠性和高效性。基于深度学习的遥感影像地类识别技术数据处理及特征提取能力较强, 能够有效提升识别精度, 使地类信息获取更加智能化, 因而被广泛应用于遥感影像地类处理。遥感影像应用的核心和关键是遥感影像解译, 遥感影像大数据时代智能解译提供了新的解决方案, 已经成为测绘遥感学科发展的重要驱动力量。本文就遥感影像分类开展研究, 首先介绍了本文所使用的DOTA数据集和YOLOv5算法, 其次, 建立遥感影像识别的YOLOv5项目工程, 并且设置算法的关键参数; 最后, 根据识别结果与可视化面板对模型进行分析。遥感图像信息容量大、各类地物交错复杂, 在识别方面有一定的挑战。本文的研究结果对遥感技术更好的应用提供了基础, 为人类生活提供更多帮助。

关键词

图像识别, 遥感图像, 卷积神经网络, YOLOv5

Remote Sensing Image Target Classification Based on Convolutional Neural Networks

Qiming Yue^{1*}, Wanlong Chen², Yulong Yan²

¹China Center for Modernization Research, National Science Library, Chinese Academy of Sciences, Beijing

²School of Mathematics, Taiyuan University of Technology, Taiyuan Shanxi

Received: Dec. 23rd, 2023; accepted: Jan. 18th, 2024; published: Jan. 23rd, 2024

Abstract

In recent years, Earth observation remote sensing satellites have been successfully launched and

*通讯作者。

operated in orbit, and new remote sensing platforms such as unmanned aerial vehicles are constantly developing and updating. Traditional manual visual interpretation can no longer meet the efficiency and accuracy requirements of remote sensing image interpretation. Deep learning has shown good reliability and efficiency in remote sensing image processing. The remote sensing image land class recognition technology based on deep learning has strong data processing and feature extraction capabilities, which can effectively improve recognition accuracy and make land class information acquisition more intelligent. Therefore, it is widely used in remote sensing image land class processing. The core and key to the application of remote sensing images is remote sensing image interpretation. In the era of big data in remote sensing images, intelligent interpretation provides new solutions and has become an important driving force for the development of surveying and remote sensing discipline. This article conducts research on remote sensing image classification. Firstly, relevant knowledge of the DOTA dataset and YOLOv5 algorithm used in this article were introduced. Secondly, the YOLOv5 project for remote sensing image recognition was established, and the key parameters of the algorithm were set. Finally, the model was analyzed based on the recognition results and visualization panel. Remote sensing images have a large amount of information capacity and complex intertwining of various land features, posing certain challenges in recognition. The research results of this article provide a foundation for the better application of remote sensing technology and provide more assistance for human life.

Keywords

Image Recognition, Remote Sensing Images, Convolutional Neural Networks, YOLOv5

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着科技的不断发展和进步, 遥感技术已经成为一项非常重要的技术。遥感技术通过卫星、无人机等手段获取地表信息, 并通过数据处理和分析, 为人类提供了大量的可视化信息, 为环境监测、地质勘探、城市规划、气象预测等领域的发展提供了强有力的支持。

在遥感技术中, 遥感影像的识别是其中的重要问题之一。遥感影像识别是通过将遥感影像进行图像分析和处理, 实现对遥感影像中不同地物和地物的种类进行识别和分类的过程。这个过程在很多领域都有着广泛的应用, 例如农业、林业、城市规划、环境监测等。遥感图像信息容量大、各类地物交错复杂, 对比自然图像有一定的挑战, 这给研究者带来了难题。因此, 基于深度学习的遥感影像识别成为了当前研究的热点领域[1]。深度学习算法具有自动学习和特征提取的能力, 能够自动学习出遥感影像特征, 并能够高效地分类和识别不同地物。基于深度学习的遥感影像识别已经在农业、林业、城市规划等领域取得了很好的效果, 成为了遥感影像识别的重要方法之一。

随着计算机技术的进步, 计算机硬件和软件条件的发展, 利用深度学习的方法来进行遥感影像识别, 已经获得了优异的性能。卷积神经网络(CNN)可以提取更抽象的特征, 对于复杂的场景有更好的识别效果。AlexNet [2]、VGGNet [3]、GoogLeNet [4]、ResNet [5]和 DenseNet [6]等广泛用于图像处理等领域。极大的丰富了深度学习的发展。桥婷婷[7]等提出了融合数据增广与迁移学习方法的高分辨率遥感影像场景分类方法, 大大提高了分类准确率, 孟庆祥[8]等利用卷积提取图像中的色彩和纹理等特征, 并对这些提取到的特征利用池化进行挑选, 对挑选后的特征进行提炼和综合, 利用提炼和综合后的特征进行分类,

提升了遥感影像场景分类的准确率。

本文首先对数据集的构建和项目工程的搭建,对于深度学习模型而言,数据集和需求算法是必不可少的一步,对后续模型的训练和检测更是重中之重,本文采用了YOLOv5算法开展设计;最后本文给出结果展示与模型分析,对于大部分训练的类别都有着较高的AP值,通过tensorboard可视化面板看到的结果可以得到模型在摇杆识别方面有着良好的效果。

2. 数据集与项目工程

2.1. 数据集的构建

遥感影像数据集是遥感影像目标检测的重要前提,构建出合适的数据集可以有效地提高模型的准确性。目前已经有很多研究人员在遥感数据采集开展了许多优秀的工作,本文直接选择由武汉大学制作并标注完成的DOTA数据集的遥感图片数据。DOTA数据集是一组在航空影像中应用于目标探测的数据集,它被用来探测和评价航空影像中的目标。图片来自于多种传感器和平台,其中有Google Earth, JL-1卫星,也有来自中国资源卫星资料与应用中心的GF-2卫星。该数据集共有2806张航拍图片,每一张图片的像素大小从800*800至4000*4000,包含了各种大小、方位和形状的目标。然后,专家对这些DOTA图像进行了注释,一个完整标注的DOTA图像有188,282个样本,每一个样本都用一个四边形来标注。

大部分的遥感图像的数据集中都含有角度信息,而且对象比较小。因此,常规的目标检测方法在对遥感图像进行处理时,往往会遇到很大的困难,DOTA数据集来自于多种传感器和平台的不同分辨率航拍图像,分辨率为4000*4000,并且DOTA数据集的label格式是与本文使用的YOLOV5需求格式有所出入,就需要对数据集进行预处理[9]。本文通过DOTA_devkit中的YOLO_Transform.py将分割后的数据集中的label文本文件进行格式转化,如DOTA数据格式转为YOLO数据格式。

在数据预处理过程中将数据图像打乱顺序。这是由于遥感数据中的相邻图像可能存在重叠区域,具有很强的相关性,如果不打乱顺序,他们将有可能出现在同一组样本之中,这就可能导致模型的泛化能力下降。基于先前对数据集的分割操作,已使得数据集的容量扩充了很大一部分,已然可以满足训练要求。本文的数据集划分比例经过仔细考虑后设置为:训练集:验证集:测试集 = 0.9:0.05:0.05,然后用此数据集来开展研究。

2.2. 项目搭建

本文采用的是一种单阶段目标检测算法——YOLOv5,它的总体方框图如图1所示。对于一个目标检测算法,我们一般可以将其划分为4个通用的模块:输入端、基准网络、Neck网络与Head输出端,对应于图1中的4个红色模块。

YOLOv5有YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x四个版本。文件中,这几个模型的结构基本一样,不同的是depth_multiple模型深度和width_multiple模型宽度这两个参数[10]。

通常情况下,我们会在网络中加入一个预先训练的权值,以减少网络的训练时间,以取得较高的精准度。而YOLOv5则为我们提供了多个预训练权,我们可以根据自己的需要来选择不同的预训练权。预训练权重越大,训练出来的准确度就会越高,但其检测的速度就会越慢。本文使用的是yolov5x.pt权重文件,效果较好,但是训练时间也相对较长[11]。首先需要对模型中的训练种类进行修改为数据集DOTA中所对应的数量和类别。根据自己数据集的实际情况,对训练参数进行修改。在这其中主要包括初始权重(weight)、训练模型文件(cfg)、数据集参数文件(data)、训练轮数(epochs)、批量处理文件数(batch size)、图片大小(img size)、训练设备(device)和最大工作核心数(workers)。

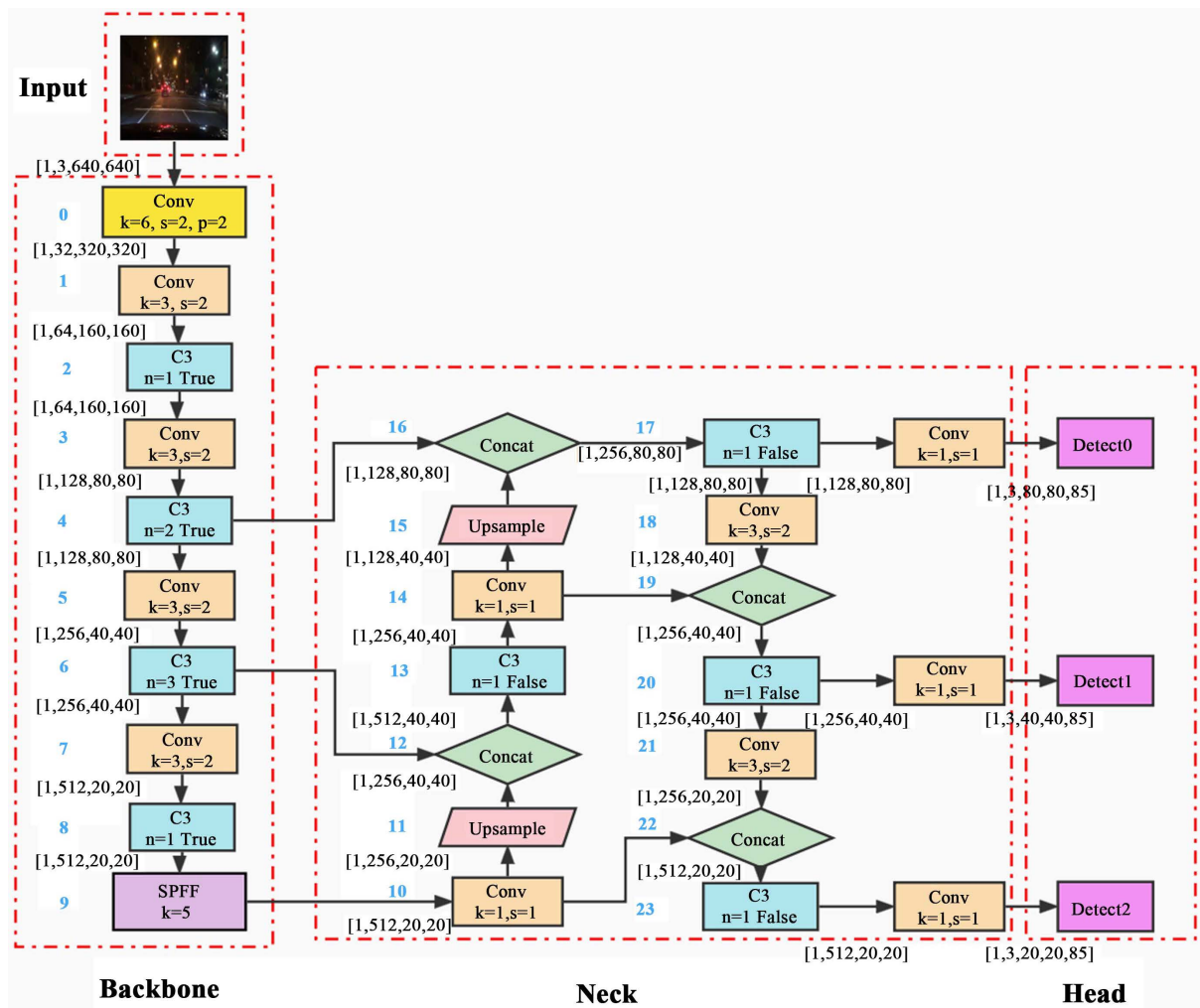


Figure 1. Diagram of the network structure

图 1. 网络结构图

YOLOv5 在对象检测方面的表现非常出色，具有以下显著的优点：首先使用 Pytorch 框架，对用户非常友好，能够方便地训练自己的数据集；不仅易于配置环境，模型训练也非常快速，并且批处理推理产生实时结果；YOLOv5 的对象识别速度高达 140FPS，用户体验较好。

3. 结果展示与模型分析

3.1. 结果展示

在对上百组样本完成测试后，发现模型的准确度和速度都能够达到预期的效果。以下是其中几组具有代表性的测试样本与测试结果。在相邻的两张图片中，左侧图片为带有标记框的测试样本，右侧图片带有预测框的测试结果。图 2(a)为对一个小型停车场的图像识别，测试样本中有两类标签共计 40 个标记框，在测试结果中有两类标签共 38 个预测框。在测试结果中，有小型车辆(small-vehicle)目标未被预测中，在其中有一个集装箱，被识别为大型车辆(large-vehicle)。图 2(b)是对一个交通环岛图像的分类识别，这一组测试对于小型车辆(small-vehicle)这一标签的预测比较准确，对于环形交叉口(roundabout)这一标签的识别，预测样本的标记框为一整块环岛区域，而预测结果则是将一个环岛的多个部分都预测为这个类。

图 2(c)为某机场的一角, 测试样本含有 4 个飞机(plane)的标记框, 在测试结果中也进行了很好的预测, 但对于测试样本中出现但未标记的小型车辆(small-vehicle), 未像图 2(a)中一般, 进行预测。

在上述几组图片中, 对模型的测试有了一个初步的了解, 本文在 YOLOv5 的基础上设计的模型对于遥感图像的目标分类有了一个较好的结果。我们可以通过预测框上的数字, 也就是置信度来判断模图像识别是否成功。本文将置信度超过 0.5 的都视为识别成功, 也可通过修改模型程序中的置信度阈值来对图像的识别结果进行筛选。

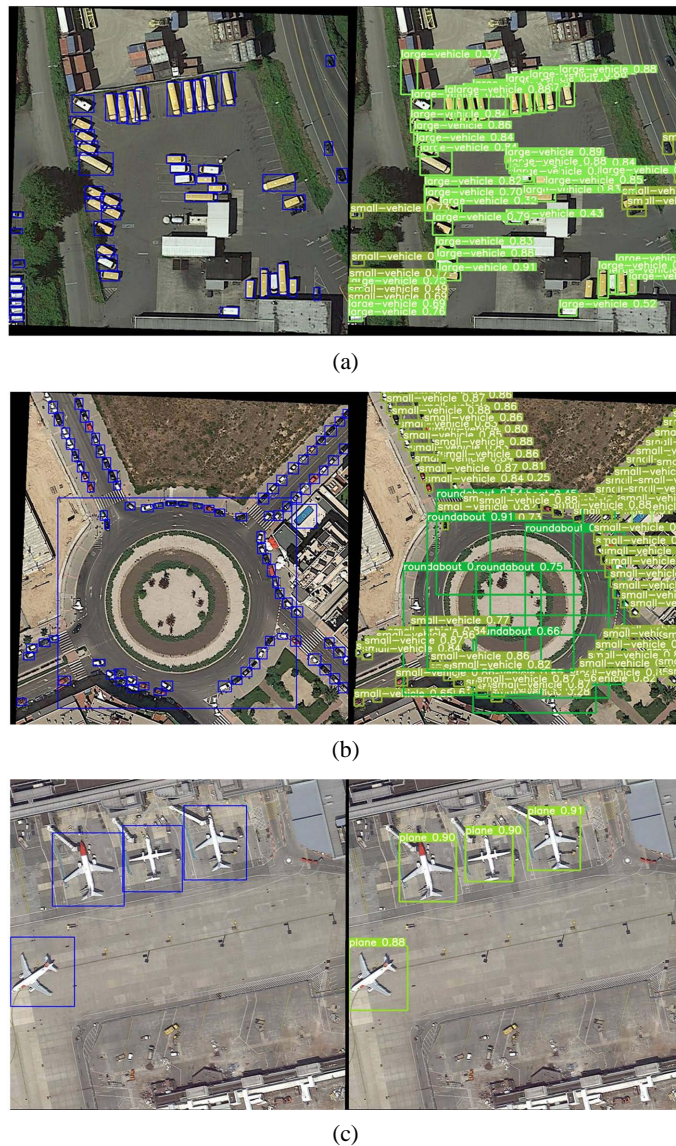


Figure 2. Diagram of image recognition
图 2. 图像识别结果

3.2. YOLO 模型评估指标

YOLO 是目前应用最为广泛的一种目标探测模式。物体的检测问题比分类问题要复杂得多, 因为物体检测不但需要准确地判断出物体的类别, 还需要准确地判断出物体的位置。本文对目标识别的评价指

标进行了归纳,分为两个部分,一个是预测框的预测指标,一个是分类预测指标。预测框的准确率用交并比(IOU)来反映[12]。交并比(IOU)是目标检测问题中的一项重要指标,它在训练阶段反映的是标注框与预测框的重合程度,用于衡量预测框的正确程度。分类预测指标采用精度和召回率进行刻画。接下来将二者结合即为评价目标检测模型的指标,即 $mAP@0.5$ 和 $mAP@[0.5:0.95]$ 。在 YOLO 模型中, $mAP@0.5$ 这种形式表示在 IOU 阈值为 0.5 的情况下, mAP 的值为多少。当预测框与标注框的 IOU 大于 0.5 时,就认为这个对象预测正确,在这个前提下再去计算 mAP 。一般来说, $mAP@0.5$ 即为评价 YOLO 模型的指标之一。YOLO 模型中还存在 $mAP@[0.5:0.95]$ 这样一种表现形式,这形式是多个 IOU 阈值下的 mAP ,会在区间[0.5, 0.95]内,以 0.05 为步长,取 10 个 IOU 阈值,分别计算这 10 个 IOU 阈值下的 mAP ,再取平均值。 $mAP@[0.5:0.95]$ 越大,表示预测框越精准,因为它去取到了更多 IOU 阈值大的情况。

受设备和环境的限制,本文对数据集的数量进行了抑制,使得标签类别在训练集中的数量减少了许多,尤其是原本在数据集中出现较少的标签类别,无法得到更好的训练。图 3 分别给出了目标检测模型的指标 $mAP@0.5$ 和 $mAP@[0.5:0.95]$ 随着训练轮数的变化图像,在谋求环境和精度的最优解的情形之下,本文的训练轮数(epochs)设计为 100 轮,尽管在如此,模型的评价价值仍达到了一个相对较好的标准。

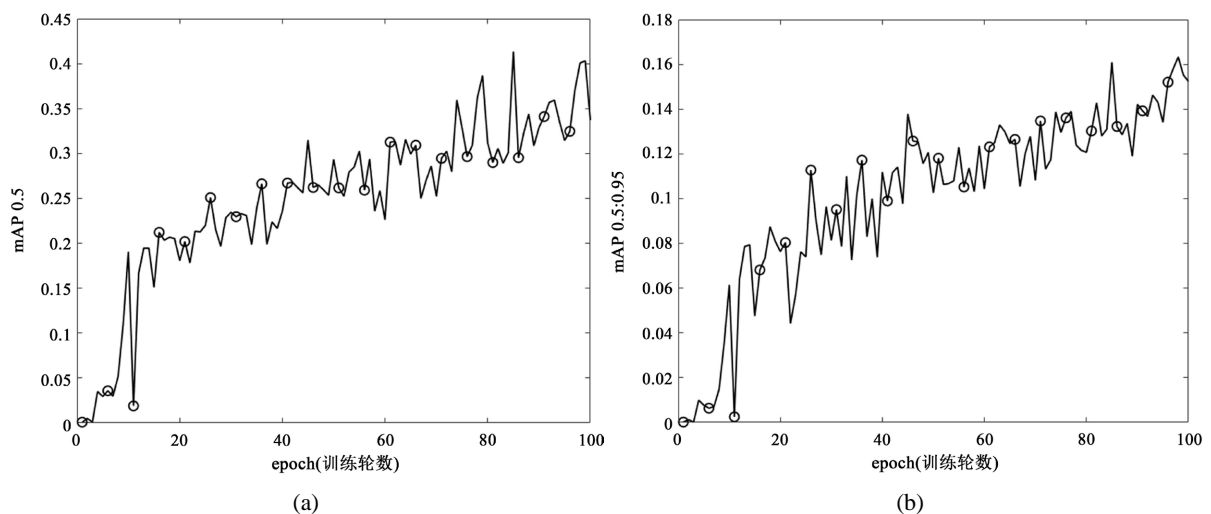


Figure 3. The relationship between the evaluation index of image recognition results and the number of training rounds
图 3. 图像识别结果的评估指标与训练轮数的关系

4. 结论

由于近年来遥感技术和卫星技术的发展,人们已经可以获取到更高分辨率的遥感影像,同时产生了大量的遥感影像。遥感影像分类是图像分割和目标检测的基础,在应用生态环境保护,城市规划等领域有重要的应用价值。本文进行了遥感影像识别算法的实现,主要内容包括了以下几个方面:一、通过对数据源进行了图像分割与格式转换,使 YOLOv5 有了合适的数据集,为后续的遥感影像识别工作的开展奠定基础。二、利用程序让标记框与预测框出现在同一张图片上,利用 IOU、 $mAP@0.5$ 和 $mAP@[0.5:0.95]$ 来对模型进行评价,对预测结果进行更好的分析。三、通过对程序运行的结果的分析,得出:所建立的工程需要靠 GPU 的算力实现对图片中的类别的检测,并且能够通过更新权重来追求理想的检测精度。遥感影像识别是一个有很大应用前景的领域,随着研究的深入,遥感影像识别算法将会更加完善和普及,为农业、城市规划、公共安全等领域提供更多有效的解决方案,具有广阔的发展空间。

基金项目

国家自然科学基金青年项目(12202301)。

参考文献

- [1] 朱寿红, 王胜利, 舒帮荣. 基于深度学习的高光谱遥感影像分类[J]. 城市勘测, 2017(4): 84-88, 92.
- [2] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. <https://doi.org/10.1145/3065386>
- [3] Simonyan, K. and Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. 2015 *International Conference on Learning Representations*, San Diego, CA, 7-9 May 2015, 1-14.
- [4] Szegedy, C., Liu, W., Jia, Y., et al. (2015) Going Deeper with Convolutions. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [5] He, K., Zhang, X., Ren, S., et al. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [6] Huang, G., Liu, Z., Laurens, V.D.M., et al. (2017) Densely Connected Convolutional Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 21-26 July 2017, 2261-2269. <https://doi.org/10.1109/CVPR.2017.243>
- [7] 乔婷婷, 李鲁群. 结合数据增广和迁移学习的高分辨率遥感影像场景分类[J]. 测绘通报, 2020(2): 37-42.
- [8] 孟庆祥, 吴玄. 基于深度卷积神经网络的高分辨率遥感影像场景分类[J]. 测绘通报, 2019(7): 17-22.
- [9] 明冬萍, 骆剑承, 沈占锋, 等. 高分辨率遥感影像信息提取与目标识别技术研究[J]. 测绘科学, 2005(3): 18-20, 23.
- [10] 邢宇驰, 李大军, 叶发茂. 基于 YOLOv5 的遥感图像目标检测[J]. 江西科学, 2021, 39(4): 725-732.
- [11] 董文轩, 梁宏涛, 刘国柱, 等. 深度卷积应用于目标检测算法综述[J]. 计算机科学与探索, 2022, 16(5): 1025-1042.
- [12] 吕石磊, 卢思华, 李震, 等. 基于改进 YOLOv3-LITE 轻量级神经网络的柑橘识别方法[J]. 农业工程学报, 2019, 35(17): 205-214.