

基于改进 EAF-YOLO 的轻量化遥感图像检测

侯林静, 李芳

(沈阳理工大学 信息科学与工程学院, 沈阳 110159)

摘要: 针对遥感图像目标检测模型难以在无人机、卫星以及移动设备等低算力场景下部署的问题, 提出一种轻量化遥感图像检测方法 EAF-YOLO。模型以 YOLOv7 框架为基础, 采用 EfficientNet 特征提取网络作为轻量级骨干网络, 降低模型参数; 在框架头部设计串行结构的金字塔池化层, 提升内部尺度交换能力, 增强对特征信息的聚合能力; 采用 WIoU 损失函数作为回归损失, 自适应地调整难易样本的权重, 提高模型的检测性能。以遥感数据集 DIOR 进行实验, EAF-YOLO 模型检测精确率达到了 90.3%, 参数量为 9.3×10^6 , 表明训练后的模型较 YOLOv7 具有较小的参数量, 适用于遥感图像计算资源有限的场景。

关键词: 遥感图像检测; 轻量化; YOLOv7; EfficientNet; 尺度交换

中图分类号: TP391.4

文献标志码: A DOI: 10.3969/j.issn.1003-1251.2025.02.002

Lightweight Remote Sensing Image Detection Based on Improved EAF-YOLO

HOU Linjing, LI Fang

(Shenyang Ligong University, Shenyang 110159, China)

Abstract: To address issues of deploying remote sensing image object detection models in low computing scenarios such as drones, satellites, and mobile devices, a novel method utilizing EAF-YOLO for lightweight remote sensing image detection is proposed. The model is based on the YOLOv7 framework. The EfficientNet feature extraction network is utilized as a lightweight backbone network to diminish the model parameters. A spatial pyramid pooling layer is crafted with a consecutive configuration within the Head section, improving the scale interactions capability and fortifying the aggregation capacity of feature information. The model employs the WIoU loss function as the regression loss, allowing for the adaptive adjustment of weights for difficult samples to enhance the detection performance. Experimenting with the remote sensing dataset DIOR reveals that, the EAF-YOLO model achieves an precision of 90.3%, while its parameter count decreases to 9.3×10^6 . Due to its low parameter count and computational intricacy, it is apt for utilization in remote sensing scenarios where there are constraints on computing resources.

Key words: remote sensing images detection; lightweight; YOLOv7; EfficientNet; scale interactions

高分辨率的卫星遥感图像中涵盖大量不同类别的地物要素信息^[1], 在智慧城市建设、应急救援以及军事侦察等领域发挥着关键作用^[2]。因此,

有必要研究如何快速准确地从高分辨率的遥感图像中获取需要的信息^[3]。

目前, 基于卷积神经网络的目标检测算法是

目标检测领域的研究热点^[4]。Ševo 等^[5]的研究结果表明,卷积神经网络在非线性特征提取、局部连接和权重共享等方面表现突出,可用于处理遥感图像目标检测任务。

基于卷积神经网络的目标检测算法通常分为单阶段目标检测和双阶段目标检测两种类型。单阶段代表算法有 YOLO 系列^[6]和 SSD^[7]等;双阶段代表算法有 Faster R-CNN 和 Mask R-CNN^[8]等。上述卷积网络模型的结构大多较为复杂,进行检测任务时涉及的计算量庞大^[9],需要依靠高性能的 GPU 保障运行,在资源受限的环境,如无人机系统、野外工作站等,复杂模型难以部署。因此,设计一个轻量化模型保证遥感图像的高效检测具有重要意义。

在轻量化的遥感图像检测领域中,刘子洋等^[10]引入 BiFormer 和 GSConv 模块减小模型规模,加强对小目标检测的精度;王燕妮等^[11]使用轻量级 VGGNet 网络作为骨干网络,并引入 FSSA 模块,改进后模型比原模型参数量缩减一半,精度达到 84.1%;郭宇阳等^[12]在 GhostNet 和 SqueezeNet 模型基础上,提出 GS-YOLO 检测模型,与 YOLOv4 相比,推理速度提高了 84%;胡根生等^[13]设计了一个轻量级网络 LiTLBNet,用于准确检测野外茶园无人机图像的茶叶枯病,该模型的内存占用量只有 2.0 MB。

从以上的研究方法可以看出,手工设计的轻量化卷积神经网络架构在特定任务和场景下可具有更好的性能,但是难以泛化到其他领域和任务上;而轻量化网络可以快速进行迭代和优化以适应不同的任务和情景,尽管可能需要权衡一定的性能损失。因此,本文提出的新方法着眼于保持检测性能的同时最大限度地减少参数量和计算量。本文在 YOLOv7 的基础上进行改进,首先使用 EfficientNet^[14]轻量级网络替换 YOLOv7 的骨干网络,通过均衡网络宽度、深度和分辨率实现模型的轻量化;其次设计串行结构的金字塔池化层,在进行特征信息融合时,该结构既能降低模型的参数数量,也能加快模型的推理过程;最后采用新的 WIoU 损失函数作为边界框回归的损失函数,降低高质量锚框的竞争力。

1 YOLOv7 模型

YOLOv7 模型主要由输入端 (Input)、主干网

络 (Backbone)、检测头 (Head) 组成, YOLOv7 算法的结构如图 1 所示。

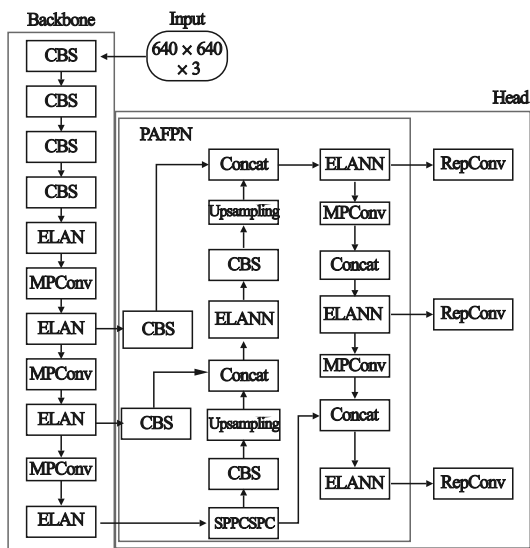


图 1 YOLOv7 结构图

Fig. 1 YOLOv7 structure diagram

第一部分为 Input, 输入图片的尺寸为 640 × 640 像素。第二部分是 Backbone, 该部分由若干卷积单元 (CBS)、高效的层级特征聚合网络 (ELAN) 以及双路径下采样卷积单元 (MPConv) 组成。其中 CBS 包含卷积层 (Conv)、批量归一化层 (BN) 和 SiLU 激活函数; ELAN 模块由多个 MPConv 模块采用双分支策略实现下采样, 一个分支执行最大池化后跟一个 1 × 1 卷积调整通道数, 另一个分支先通过 1 × 1 卷积调整通道, 再应用步长为 2 的 3 × 3 卷积进一步下采样, 最后融合两部分输出, 得出高效的下采样结果。第三部分为 Head, 首先通过特征融合网络 (PAFPN) 进行两轮特征整合, 保留更多的低层次特征信息, 然后利用空间金字塔池化跨阶段部分共享模块 (SPPCSPC) 扩展目标的感知域。上采样模块 (UpSampling) 负责调整特征图的通道数并执行上采样, 特征图通过 Concat 操作在通道维度上实现合并。ELANN 的结构与 ELAN 相近, 区别在于 ELANN 在每层的 CBS 模块之后都进行了特征拼接 (Concat) 操作。PAFPN 中的 MPCConv 模块与主干网络中的 MPCConv 模块类似, 通过增加输入特征图通道数并进行下采样, 实现与浅层特征尺寸的匹配, 便于进行特征图的拼接。最终, 三种尺度的特征通过结构重参数化网络 (RepConv) 进行优化, 并通过卷积调整通道数进行目标检测。

2 EAF-YOLO 模型

虽然 YOLOv7 具有一定的先进性,但网络结构复杂、计算量大,对硬件设备性能要求较高。另外,遥感图像中存在海量地物信息,随着网络的加深会加剧目标的特征损失,造成目标漏检和错检^[15]。本文提出的 EAF-YOLO 模型优化了原始的 YOLOv7 结构,实现了模型轻量化与检测性能之间的平衡。

2.1 基于 Efficient 的轻量化改进

本文采用 EfficientNetB0 网络替换 YOLOv7 的 CSPDarknet53 网络。EfficientNetB0 通过固定比例去缩放网络的宽度、深度、输入图像的分辨率,使模型更加轻量化。EfficientNetB0 网络的核心是移动翻转瓶颈卷积 (MBConv) 模块,其由普通卷积 (Conv)、深度可分离卷积 (Depwise Conv)、SE 注意力机制和随机失活层 (Dropout) 构成,结构如图 2 所示。

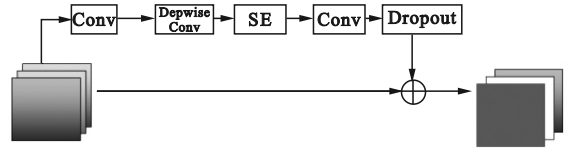


图 2 MBConv 结构图

Fig. 2 MBConv structure diagram

图 2 中前后两个 1×1 的 Conv 是为了确保在升维和降维操作后特征图尺寸一致。深度可分离卷积结合了空间上的信息处理和通道间的特征整合,使 MBConv 能够更有效地提取和传播输入图像中的特征。SE 注意力机制由一个全局平均池化、两个全连接层组成,可使模型更加关注信息量最大的通道特征,在随机失活层训练迭代中,随机地选择一些神经元并将其输出置为零,减少过拟合。

本文使用 DIOR 数据集分别将 GhostNetV2、MobileNetV3 和 EfficientNetB0 作为 YOLOv7 模型的 Backbone 进行训练,表 1 为对比实验结果,可见 EfficientNetB0 有明显优势。

表 1 不同特征提取网络结果对比

Table 1 Comparison of the results of different feature extraction networks

骨干网络	精确度/%	参数量/ 10^6	速度/ s^{-1}	mAP@0.5/%
GhostNetV2 ^[16]	89.0	33.0	46.3	88.9
MobileNetV3 ^[17]	87.2	24.5	53.8	85.2
EfficientNetB0	88.7	11.8	85.6	88.6

2.2 SPPAFC 层

针对遥感图像背景复杂的特点,在 Head 部分设计串行结构的 SPPAFC,其结构如图 3 所示。

SPPAFC 结构主要由输入 (Sinput)、两个 CBS 模块和一个基于注意力的内部尺度特征交互 (AIFI) 模块以及输出 (Soutput) 串联组成。在 Sinput 端输入特征图像,图像在经过 AIFI 处理前后,都经过一个 CBS 操作,减少通道之间的冗余信息,最后在 Soutput 端输出。其中 AIFI 采用 Transformer 编码器结构,内部包含一个多头自注意力 (MHSA) 和前馈神经网络 (FFN)。AIFI 是一个尺度内特征交互模块,可以输出具有更强语义信息的特征图,将 AIFI 模块与卷积模块串行,可使特征图经过卷积运算后具有更好的局部感知能



图 3 SPPAFC 结构图

Fig. 3 SPPAFC structure diagram

力,减少冗余信息的处理量,降低不必要的计算消耗。

EAF-YOLO 网络结构中的 Backbone 由 EfficientNetB0 网络构成,PAFPN 中原 SPPCSPC 结构被 SPPAFC 结构替代。改进后的 EAF-YOLO 网络结构如图 4 所示。

2.3 损失函数的改进

YOLOv7 网络采用的损失函数计算表达式为

$$L = L_{loc} + L_{conf} + L_{cls} \quad (1)$$

式中: L_{loc} 为定位损失; L_{conf} 为置信度损失; L_{cls} 为分类损失。采用交叉熵损失函数 CIoU 计算 L_{conf} 和 L_{cls} ,CIoU 用来评估定位损失,依据预测框与真实框宽高比的相对差异进行计算。当预测框与真实框的宽高比一致时,CIoU 损失的惩罚项达到最小值,表明仅考虑宽高比差异可能无法完全表达长宽的实际差异,导致高质量预测框较少,影响网络的训练。因此,引入动态非单调聚焦机制的 WIoU 损失函数。

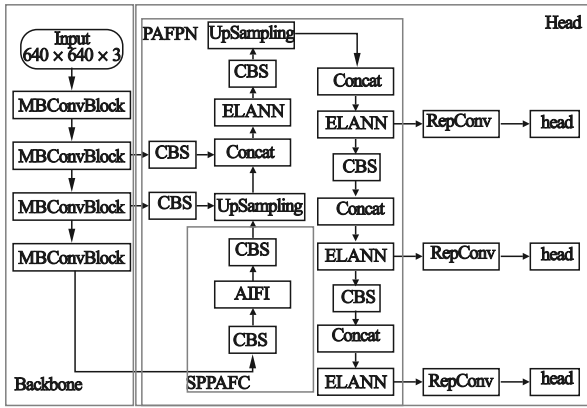


图4 EAF-YOLO 结构图

Fig. 4 EAF-YOLO structure diagram

表2 损失函数对比实验结果

Table 2 Comparison of experimental results of loss function

模型	参数量/ 10^6	GFLOPs	精确率/%	mAP@0.5/%
YOLOv7 + WIoU	37.3	105.4	89.4	89.5
YOLOv7 + CIoU	37.3	105.4	88.6	88.6

3 实验分析与讨论

3.1 实验数据集

实验采用西北工业大学于2019年发布的DIOR遥感图像数据集,该数据集包含23463张场



图5 图像样本

Fig. 5 Image samples

3.2 实验平台及网络训练参数设置

模型训练和性能评价实验在GPU服务器上完成,服务器硬件配置中CPU为Intel Xeon Platinum 8255C、显卡型号为NVIDIA GeForce RTX 3080、系统运行内存为40GB;软件环境配置中Python版本为3.8、torch版本为2.0.1、CUDA版

利用DIOR数据集对YOLOv7模型进行训练,其中损失函数分别采用CIoU和WIoU,表2为两种对比实验结果。采用参数量、计算量、精确率和mAP@0.5作为评价指标,计算量以每秒10亿次的浮点运算量(GFLOPs)来衡量。

通过动态非单调聚焦机制,模型在处理包含复杂背景和遮挡物的目标时,能够均衡处理,使模型的预测框更接近目标的真实边界框,提高模型的性能。表2的实验结果表明,在不增加参数量和计算量的前提下,WIoU损失函数有效地提高了边界框的收敛质量,并在精确率和mAP@0.5上呈现优势。

景图像和192472个目标实例组成,覆盖了20个常见目标类别,图像尺寸统一为 800×800 像素,空间分辨率在0.5~30m之间变动。本文实验按照3:1:1的比例对数据集进行训练集、验证集和测试集划分。数据集中的图像样本如图5所示。

本为11.7。实验设置统一的训练参数。

3.3 消融实验

为了验证本文提出的多种改进措施对遥感图像目标检测性能的提升效果,进行了一系列消融实验,结果如表3所示。

表 3 消融实验结果
Table 3 Ablation experiments

模型	改进措施			参数量/ 10^6	GFLOPs	精确率/%
	EfficientNetB0	SPPAFC	Wiou			
YOLOv7	×	×	×	37.3	105.4	88.6
A	√	×	×	11.8	20.3	88.3
B	√	√	×	9.6	18.9	89.4
C	√	×	√	12.0	20.9	89.7
D	×	√	√	30.3	99.7	89.8
E	√	√	√	9.3	18.3	90.3

由表 3 中的结果可以看出:A 组模型使用 EfficientNetB0 网络作为主干网络后,相比 YOLOv7 原网络参数量减少 68.4%,计算量减少 80.7%,即参数量和计算量明显下降,提高了模型的检测速度;B 组模型验证了 SAPPAC 网络的有效性,在模型精度不降低的情况下,参数量继续下降了 18.6%;C、D 组模型结果验证了 Wiou 函数改进 EAF-YOLO 模型错检漏检问题,且不消耗额外的计算资源;E 组模型为本文提出的 EAF-YOLO 模型,检测精度相较于 YOLOv7 提升了 1.7 个百分点,参数量降低了 75.1%,计算量降低了 82.6%,即 EAF-YOLO 模型在算力有限的硬件环境中能够实现对于遥感图像的精准检测。

3.4 对比实验

为体现本文改进方法的有效性,在 DIOR 数据集上进行与其他算法的对比实验,实验结果如表 4 所示,为了实验的公平性,在训练时设置统一的参数。由表 4 可以看出:相较于双阶段检测方法,单阶段检测在参数量上更小,检测速度更快;SSD 在精度方面比 EAF-YOLO 高 0.7%,但 EAF-YOLO 比 SSD 在 mAP@0.5 上提高 19.6%,说明 SSD 处理多类别检测时,由于背景噪声和类别不平衡导致目标识别率低,易出现漏检问题;与 RT-DETR 相比,在精度相差不大的情况下,EAF-YOLO 的参数量更低;本文算法的 mAP@0.5 达到了 89.7%,比 YOLOv5、YOLOv7 分别提升了 3.7% 和 1.1%。

表 4 不同算法检测结果对比

Table 4 Comparative analysis of detection effectiveness with different algorithms

算法	精确率/%	参数量/ 10^6	检测速度/ s^{-1}	mAP@0.5/%
Faster-RCNN	86.4	137.1	11.6	71.1
RT-DETR	89.5	33.0	16.5	88.9
SSD	91.0	26.3	48.9	70.1
YOLOv5	90.0	7.1	63.4	86.0
YOLOv7	88.6	37.3	67.1	88.6
EAF-YOLO	90.3	9.3	80.4	89.7

4 结论

本文提出了 EAF-YOLO 轻量化模型,该模型首先采用 EfficientNet 轻量级特征提取网络作为骨干网络进行图像特征提取;其次构建 SPPAFC 结构,加强 Head 结构的特征提取能力;最后使用 Wiou 损失函数完善原始的损失函数,提高遥感图像目标检测性能。采用公共数据集 DIOR 的实验结果表明,本文 EAF-YOLO 模型参数量降低到 9.3×10^6 ,mAP@0.5 达到了 89.7%,实现了在降

低参数的同时保持了遥感图像目标检测的精度。

参考文献 (References):

- [1] 张腾泽,李旭军,饶立明.改进 YOLOv5 的遥感图像小目标检测算法[J].计算机时代,2023(12):89-95.
ZHANG T Z,LI X J,RAO L M. Improved YOLOv5 for remote sensing image small target detection[J]. Computer Era, 2023(12):89-95. (in Chinese)
- [2] 张鹏,刘石栋,刘振军.结合高效通道注意力的轻量级遥感影像目标检测方法[J].测绘与空间地理信息,2023,46(12):53-56.
ZHANG P,LIU S D,LIU Z J. Lightweight remote sensing image target detection method combined with efficient channel attention[J]. Geomatics & Spatial Information Technology,

- 2023,46(12):53-56. (in Chinese)
- [3] HUI Y M, WANG J, LI B. STF-YOLO: a small target detection algorithm for UAV remote sensing images based on improved swintransformer and class weighted classification decoupling head[J]. *Measurement*, 2024, 224: 113936. (in Chinese)
- [4] 魏雅丽, 牛为华. 改进 YOLOv5s 的轻量化航拍小目标检测算法[J]. *计算机技术与发展*, 2024, 34(2): 53-59.
WEI Y L, NIU W H. Aerial small target detection based on improved YOLOv5s lightweight algorithm [J]. *Computer Technology and Development*, 2024, 34(2): 53-59. (in Chinese)
- [5] ŠEVO I, AVRAMOVIĆ A. Convolutional neural network based automatic object detection on aerial images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(5): 740-744.
- [6] 杨大为, 张成超. 基于改进 YOLOv7 的安全帽佩戴检测算法[J]. *沈阳理工大学学报*, 2024, 43(1): 16-21.
YANG D W, ZHANG C C. Algorithm of safety helmet wearing detection based on improved YOLOv7 [J]. *Journal of Shenyang Ligong University*, 2024, 43(1): 16-21. (in Chinese)
- [7] 康鑫英, 张德育, 王君. 压缩 SSD 低空小目标检测算法研究[J]. *沈阳理工大学学报*, 2022, 41(6): 1-5.
KANG X Y, ZHANG D Y, WANG J. Research on detection algorithm of small target at low altitude based on compressed SSD [J]. *Journal of Shenyang Ligong University*, 2022, 41(6): 1-5. (in Chinese)
- [8] LI K, CHENG G, BU S H, et al. Rotation-insensitive and context-augmented object detection in remote sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(4): 2337-2348.
- [9] 戚玲珑, 高建瓴. 基于改进 YOLOv7 的小目标检测[J]. *计算机工程*, 2023, 49(1): 41-48.
QI L L, GAO J L. Small object detection based on improved YOLOv7 [J]. *Computer Engineering*, 2023, 49(1): 41-48. (in Chinese)
- [10] 刘子洋, 徐慧英, 朱信忠, 等. Bi-YOLO: 一种基于 YOLOv8 改进的轻量化目标检测算法[J/OL]. *计算机工程与科学*, 2023: 1-15 [2024-02-08]. <https://kns.cnki.net/kcms/detail/43.1258.TP.20231107.1657.002.html>.
LIU Z Y, XU H Y, ZHU X Z, et al. Bi-YOLO: an improved lightweight object detection algorithm based on YOLOv8 [J/OL]. *Computer Engineering & Science*, 2023: 1-15 [2024-02-08]. <https://kns.cnki.net/kcms/detail/43.1258.TP.20231107.1657.002.html>. (in Chinese)
- [11] 王燕妮, 贾瑞英. 基于改进 YOLOv3 的轻量级目标检测算法[J]. *探测与控制学报*, 2023, 45(5): 98-105.
WANG Y N, JIA R Y. Lightweight object detection algorithm based on improved YOLOv3 [J]. *Journal of Detection & Control*, 2023, 45(5): 98-105. (in Chinese)
- [12] 郭宇阳, 胡伟超, 戴帅, 等. 面向路侧交通监控场景的轻量化车辆检测模型[J]. *计算机工程与应用*, 2022, 58(6): 192-199.
GUO Y Y, HU W C, DAI S, et al. Lightweight vehicle detection model for roadside traffic monitoring scenarios [J]. *Computer Engineering and Applications*, 2022, 58(6): 192-199. (in Chinese)
- [13] 胡根生, 谢一帆, 鲍文霞, 等. 面向无人机遥感图像中茶叶枯病检测的轻量型网络[J/OL]. *农业机械学报*, 2024: 1-18 [2024-03-29]. <https://kns.cnki.net/kcms/detail/11.1964.S.20240228.1722.042.html>.
HU G S, XIE Y F, BAO W X, et al. Lightweight network for tea leaf blight detection in UAV remote sensing images [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024: 1-18 [2024-03-29]. <https://kns.cnki.net/kcms/detail/11.1964.S.20240228.1722.042.html>. (in Chinese)
- [14] TAN M X, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks [C]//ICML International Conference on Machine Learning. Long Beach, California, USA: PMLR, 2019: 6105-6114.
- [15] 冯长峰, 王春平, 付强, 等. 基于深度学习的光学遥感图像目标检测综述[J]. *激光与红外*, 2023, 53(9): 1309-1319.
FENG C F, WANG C P, FU Q, et al. Survey of object detection in optical remote sensing images based on deep learning [J]. *Laser & Infrared*, 2023, 53(9): 1309-1319. (in Chinese)
- [16] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2020: 1577-1586.
- [17] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019: 1314-1324.

(责任编辑:和晓军)

(上接第6页)

- [16] LI X, ZHOU H B, ZHAO M. Transformer-based cascade networks with spatial and channel reconstruction convolution for deepfake detection [J]. *Mathematical Biosciences and Engineering*, 2024, 21(3): 4142-4164.
- [17] JIANG X, ZHUANG X T, CHEN J S, et al. YOLOv8-MU: an improved YOLOv8 underwater detector based on a large kernel block and a multi-branch reparameterization module [J]. *Sensors*, 2024, 24(9): 2905.
- [18] LIU Y Q, SONG L, LI J, et al. Enhanced tuna detection and automated counting method utilizing improved YOLOv7 and ByteTrack [J]. *Applied Sciences*, 2024, 14(12): 5321.
- [19] 吕宏泽, 李继财, 杨乔楠, 等. 基于改进 YOLOv5 车辆检测方法[J]. *计算机工程与设计*, 2024, 45(6): 1705-1712.
LYU H Z, LI J C, YANG Q N, et al. Vehicle detection method based on improved YOLOv5 [J]. *Computer Engineering and Design*, 2024, 45(6): 1705-1712. (in Chinese)
- [20] 高宏伟, 王雨桐. 基于颈颈模块与单解耦头的航拍车辆检测方法[J]. *沈阳理工大学学报*, 2024, 43(2): 1-6.
GAO H W, WANG Y T. Aerial vehicle detection method based on fast neck module and single decoupled head [J]. *Journal of Shenyang Ligong University*, 2024, 43(2): 1-6. (in Chinese)
- [21] ZHAO Y J, ZHANG Z H, ZHANG N N, et al. Nonlinear spatial decoupling strategy and deformable convnets v2 for hyperspectral image classification [J]. *Geocarto International*, 2022, 37(27): 18150-18171.

(责任编辑:和晓军)