

基于改进YOLOv5s的小目标工程车辆 定点监测识别算法*

徐世亮¹, 赖民权², 雷雨¹, 刘继忠^{2†}

(1. 江西省自然资源事业发展中心, 江西 南昌 330025; 2. 南昌大学 先进制造学院, 江西 南昌 330031)

摘要: 为有效减少并预防因违法土地开垦和矿产挖掘而造成自然环境破坏的行为, 利用部署到高塔上的摄像头, 提出了一种在复杂环境中进行各类工程车辆检测的ETS-YOLO小目标监测识别算法. 首先, 使用EfficientViT网络替换YOLOv5s的主干特征提取网络, 以提高注意力多样性, 大幅缩减模型参数量. 其次, 增加小目标检测层, 增强网络对浅层语义信息的提取, 以提高小目标检测效果. 最后, 使用软非极大值抑制算法 (soft-NMS) 替换原有NMS函数, 以有效识别遮挡、重叠目标. 实验结果表明: 改进后的模型平均准确度均值 (mAP) 为93.3%、参数量为5.90 M、检测速度为52 f/s. 相较YOLOv5s模型, mAP 提升2.6%, 参数量下降16.1%.

关键词: 目标检测; 工程车辆; EfficientViT; 改进型YOLOv5s; 软非极大值抑制算法

DOI: 10.13568/j.cnki.651094.651316.2024.04.28.0001

中图分类号: TP391 **文献标识码:** A **文章编号:** 2096-7675(2025)01-0099-08

引文格式: 徐世亮, 赖民权, 雷雨, 刘继忠. 基于改进YOLOv5s的小目标工程车辆定点监测识别算法[J]. 新疆大学学报(自然科学版中英文), 2025, 42(1): 99-106.

英文引文格式: XU Shiliang, LAI Minquan, LEI Yu, LIU Jizhong. Improved YOLOv5s algorithm for small target engineering vehicles detection with fixed-point cameras[J]. Journal of Xinjiang University(Natural Science Edition in Chinese and English), 2025, 42(1): 99-106.

Improved YOLOv5s Algorithm for Small Target Engineering Vehicles Detection with Fixed-Point Cameras

XU Shiliang¹, LAI Minquan², LEI Yu¹, LIU Jizhong²

(1. Jiangxi Natural Resources Development Center, Nanchang Jiangxi 330025, China;

2. School of Advanced Manufacturing, Nanchang University, Nanchang Jiangxi 330031, China)

Abstract: In order to effectively reduce and prevent the damage to the natural environment caused by illegal land reclamation and mineral excavation, an ETS-YOLO small target monitoring and recognition algorithm for the detection of various types of engineering vehicles in complex environments is proposed using cameras deployed to high towers. Firstly, the EfficientViT network is used to replace the backbone feature extraction network of YOLOv5s in order to improve the attention diversity and significantly reduce the number of model parameters. Secondly, a small target detection layer is added to enhance the network's extraction of shallow semantic information to improve the performance of small target detection. Finally, the original NMS function is replaced with the soft non-maximal suppression algorithm (soft-NMS) to effectively recognize occluded and overlapped targets. The experimental results show that the improved model has a mean average precision (mAP) of 93.3%, a parameter count of 5.90 M, and a detection speed of 52 f/s. Compared with the YOLOv5s model, the mAP is improved by 2.6% and the parameter count is decreased by 16.1%.

Key words: target detection; engineering vehicles; EfficientViT; improved YOLOv5s model; soft-NMS

* 收稿日期: 2024-04-28

基金项目: 江西省高层次高技能领军人才培养工程项目“基于GIS与视频融合的自然资源监管关键技术研究”(2022233); 江西省自然资源厅科技创新项目“基于智能监控与地理信息融合技术在自然资源监管中的应用研究”(202317).

作者简介: 徐世亮(1981—), 男, 硕士, 正高级工程师, 主要从事计算机软件技术、网络与信息安全的研究, E-mail: vigor2005@163.com.

† 通讯作者: 刘继忠(1974—), 男, 博士, 教授, 博士生导师, 主要从事智能机电系统与机器人、机器视觉结合的研究, E-mail: liujizhong@ncu.edu.cn.

0 引言

近年来,因违法土地开垦和矿产挖掘破坏自然环境的现象时有发生.因此,利用科技手段加强自然资源保护与合理利用尤为重要.实际执法过程中,如果只惩治违法搭建^[1],难以根除土地违法建设问题.因此,在一些关键位置(如铁塔上)安装高空瞭望摄像头对目标区域的工程车辆进行检测识别,提前作出疑似违法行为预警是很有必要的.然而,很多地区自然环境复杂,高架、高塔架设的摄像头拍摄目标小,现有目标检测算法无法达到检测要求.因此,实现小目标、遮挡目标的工程车辆检测识别具有重要的现实意义和价值.

基于深度学习的目标检测算法主要可以分为两类:单阶段算法和两阶段算法.其中,两阶段算法中典型的的目标检测算法是基于区域的卷积神经网络(R-CNN)系列中的Faster R-CNN^[2],单阶段算法中典型的的目标检测算法为YOLO(You Only Look Once)^[3-6]和单阶段多目标检测(Single Shot MultiBox Detector, SSD)^[7].随着机器视觉的快速发展,这些算法被广泛应用于建筑施工车辆的分类检测任务中,高弋淞等^[8]提出了一个基于Android平台,搭建SSD-MobileNet算法模型实时检测建筑施工车辆的方法.虽然能实现部分目标的实时检测要求,但针对一些与背景相似的小目标施工车辆的检测效果较差,漏检率较高. Guo等^[9]基于SSD和特征金字塔单阶段检测(Feature Pyramid Single Shot Detector, FSSD)^[10],提出了一种融合了方向感知边界框(Orientation-Aware Bounding Box, OABB)的方向感知特征融合单级检测(Orientation-Aware Feature Fusion Single-Stage Detection, OAFF-SSD)算法模型.该模型能够有效实现对密集、旋转目标的检测,但其轻量化程度不足,难以部署到实时检测轻量化设备中.顾晨亮等^[11]改进Faster R-CNN算法中的损失函数,提出一种基于自适应局部斥力与归一化面积损失的建筑施工车辆目标检测算法.该算法有效降低了遮挡建筑施工车辆的漏检率,但并没有很好地解决小目标检测问题.颜文杰^[12]利用可变形卷积替换Res2Net残差模块中的卷积核,提高基于Cascade R-CNN算法模型网络的空间特征表达能力.同时,设计了一个带权重的平衡化特征金字塔网络(Weighted Balanced Feature Pyramid Network, WBFPN)替换原有的特征金字塔,进一步提高了目标检测的精度.改进目标检测算法能很好地完成遮挡目标、小目标的检测任务,但存在模型参数量大、帧率低等问题,不适于配置到轻量化设备中.

针对上述问题,并结合可部署到高塔等定点位置架设摄像头进行监测的要求,本文提出一种以YOLOv5s为基础模型^[13],面向工程车辆小目标、重叠目标的实时检测需求的目标检测算法.

1 YOLOv5s网络结构

YOLOv5s算法主要包括主干网络(Backbone)、颈部网络(Neck)和头部网络(Head).YOLOv5s网络结构如图1所示.

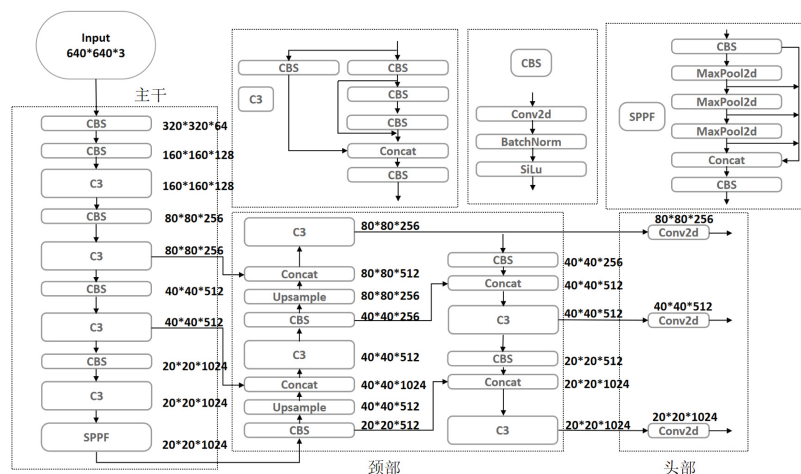


图 1 YOLOv5s网络结构

输入端由Mosaic数据增强、自适应缩放图片和自适应计算锚框等模块组成.其中,Mosaic数据增强可以通过随机缩放、随机裁剪、随机排布的方式进行拼接,进而增强训练集的多样性,提高模型的泛化能力.

主干网络主要负责提取输入图像的特征,主要部分是卷积层、C3结构和快速金字塔池化(Spatial Pyramid Pooling-Fast, SPPF)结构.卷积层主要对特征图进行压缩和下采样;C3结构用于对输入进行特征提取、融合,

丰富特征的语义信息; SPPF结构用于将不同层次和尺度的特征图进行融合, 以增强模型对目标物体的识别和定位能力.

颈部网络采用特征金字塔网络 (Feature Pyramid Network, FPN) 和路径聚合网络 (Path Aggregation Network, PANet) [14]相结合的结构. FPN采用自上而下的上采样传递语义信息, PANet采用自下而上的下采样传递定位信息, 将高级语义特征与低级细节特征进行融合, 从而在不同尺度上提取和利用目标的特征信息. 这种渐进的融合方式可以提高目标检测算法对不同尺度目标的感知能力.

头部网络主要包括三个检测头, 分别用于预测大、中、小三种目标. 输出端通过CIoU Loss[15]函数计算预测框与目标框的坐标损失.

同时, YOLOv5s中的目标检测算法一般采用非极大值抑制 (Non-Maximum Suppression, NMS) 算法进行目标检测中目标之间的重叠遮挡问题的后续处理, 即根据一定的阈值筛选出最具有代表性的检测结果, 去除冗余的重叠边界框, 保留最有可能包含物体的边界框, 以实现重叠遮挡目标的检测识别. 但采用固定阈值的传统NMS去除重叠的检测边界框方法, 容易导致过度抑制重叠边界框而产生漏检.

针对高空架设的摄像头对远距离小目标工程车辆的监测要求, 在上述YOLOv5s架构的基础上, 拟从三个角度出发, 提出一种ETS-YOLO小目标监测识别算法.

1) E (EfficientViT): 通过引入使用MBConv (MobileNetV3 Block Convolution) [16]模块和基于ReLU的轻量级自注意力机制[17]优化的EfficientViT网络[18], 并用其替换YOLOv5s的骨干网络, 增强模型对重要特征的表达能, 并降低模型参数量.

2) T (Small Target Detection Layer): 通过增加一个小目标检测层 (Small Target Detection Layer), 提高模型对浅层语义信息的捕捉能力, 增强对多尺度信息特征的融合, 提高小目标检测及整体目标的检测精度.

3) S (soft-NMS): 通过引入soft-NMS[19]代替NMS后处理方法, 提高模型对遮挡重叠目标的检测性能, 降低漏检率.

2 本文方法

2.1 ETS-YOLO模型结构

YOLOv5s算法作为单阶段目标检测算法, 具有体积小、检测精度高、实时性好等优点, 更加适用于实时检测, 部署在轻量化设备上检测效果更好. 但YOLOv5s算法对小目标以及遮挡目标识别效果不佳, 本文提出一种基于改进YOLOv5s的定点位置工程车辆检测算法, 具体结构如图2所示. 在主干网络 (Backbone) 采用EfficientViT代替原YOLOv5s的主干网络 (图2最左侧列), 在头部网络 (Head) 增加一个160×160小目标检测层, 并在后处理方法中采用soft-NMS代替NMS解决目标检测中目标间的重叠遮挡问题.

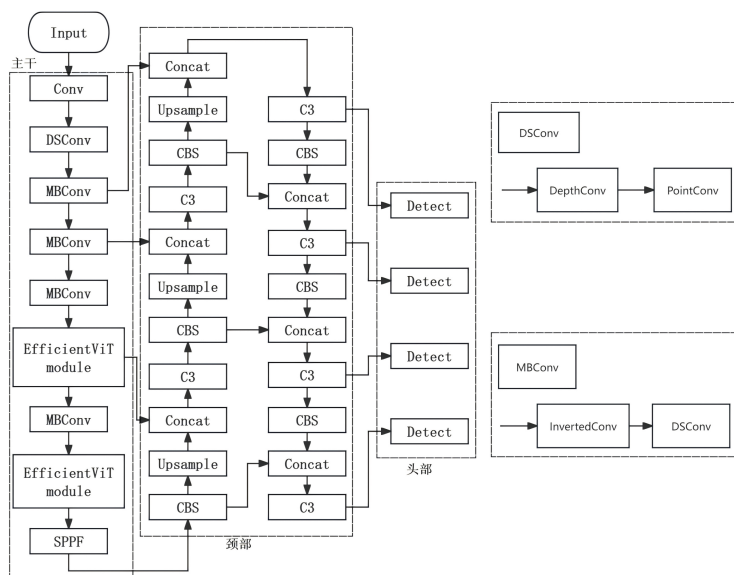


图 2 ETS-YOLO模型结构

2.2 EfficientViT网络

高效的视觉变换网络(EfficientViT)是一种为处理高分辨率图像而设计的轻量级网络结构,其优势在于它结合了视觉变换器(Vision Transformer)的优越性能和轻量级网络的高效计算能力,可以在损失小部分精度的情况下大量减少模型参数量. EfficientViT网络结构如图3所示,主要包括DSConv模块、MBConv模块和EfficientViT模块. 通过将原主干网络的CBS、C3替换为DSConv、MBConv和EfficientViT模块,达到替换主干网络的效果.

1) DSConv模块

DSConv模块是一种用于计算机视觉领域的深度可分离卷积(Depthwise Separable Convolution)模块^[20]. DSConv模块结合了深度卷积和逐点卷积两种卷积操作. 深度卷积对输入特征图的每个通道独立地进行卷积操作,有效地减少了参数数量和降低计算复杂度. 逐点卷积则将深度卷积产生的多个通道输出进行处理,将它们组合成最终的特征图. 这种模块设计旨在通过减少参数数量和计算复杂度来提高网络的效率,同时尽可能地保持、甚至提升模型的性能.

2) MBConv模块

MBConv模块是一种轻量级的网络模块,常用于构建高效的计算机视觉模型,如MobileNetV、EfficientNet^[21]等. MBConv在DSConv的基础上加入了一个InvertedConv,引入一个放大系数,对输入通道进行扩宽,最后在DSConv的逐点卷积中缩放回去,使得MBConv成为具有强大特征提取能力的轻量型卷积.

3) EfficientViT模块

EfficientViT模块结构如图4所示. EfficientViT的

核心是多尺度线性注意力模块MLA(Multi-Scale Linear Attention),它能够在不同的尺度上进行特征学习,有效地捕获图像中的细节信息. 传统的卷积网络在提取特征时可能受限于局部感受野,而EfficientViT通过多尺度线性注意力模块实现了全局感受野,有助于捕捉图像远距离依赖关系. 通过在不同尺度上进行特征提取,EfficientViT能够更好地识别不同大小的目标,特别是在处理小目标检测时,可以显著提高检测精度.



图4 EfficientViT模块结构

2.3 新增小目标检测层

在YOLOv5s网络中添加一个新的 160×160 小目标检测层以提高模型对小目标的检测能力^[22]. 具体实现如下: 首先,增加一个特征提取层,该层能够从输入图像中提取出更加丰富的特征信息,以帮助模型更准确地检测小目标. 其次,引入一个额外的检测层,该层在特征图上执行检测操作,对目标的位置和相关的分类信息进行预测. 新增小目标检测层为图2头部中尺度为 160×160 的Detect头. 最后,通过训练过程中的反向传播和优化算法,使得新增的小目标检测层能够逐步学习到更好的特征表示和检测能力.

各检测层锚框大小初始设定如表1所示,其中P2为新增检测层,新增的检测层能够充分利用输入图像的特征信息,以及通过训练过程中的优化,提高模型对小目标的识别能力.

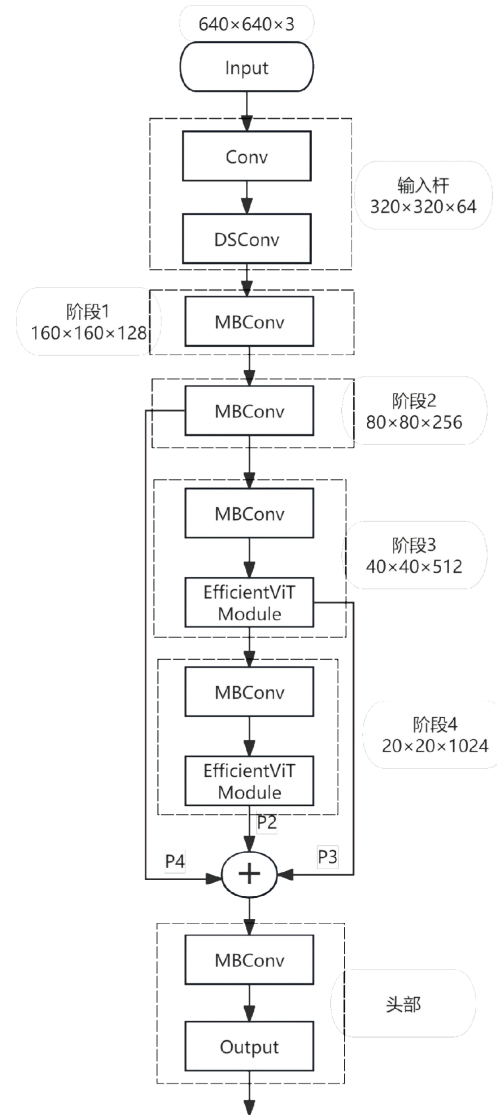


图3 EfficientViT网络结构

2.4 软非极大值抑制soft-NMS

工程车辆目标检测任务中,目标之间的重叠、遮挡现象是很常见的.当工程车辆在图像中呈现不同的尺度、角度或位置变化时,可能会导致多个检测边界框重叠.为解决这个问题,原YOLOv5s模型通过使用NMS方法进行后处理,可以根据一定的阈值筛选出最具有代表性的检测结果,去除冗余的重叠边界框,实现目标检测的准确性提升.然而,传统NMS方法在去除重叠的检测边界框时使用固定阈值,可能会导致过度抑制重叠边界框并产生漏检.

针对上述问题,引入soft-NMS算法代替原网络中的NMS算法.soft-NMS通过引入软性阈值和概率图模型来进行解决.soft-NMS将重叠边界框的概率得分进行衰减,而不是直接删除它们,从而在保留重叠边界框的同时减少它们的权重.这种衰减的方式使得重叠边界框的概率得分可以在不同的程度上降低,而不是直接丢弃,因此在检测结果中更全面地考虑了所有重叠边界框的贡献,从而提高目标检测的准确性和稳定性,避免传统NMS算法的过度抑制缺点.

3 实验与分析

3.1 数据集

本文使用的数据集为江西农村高架、高塔架设摄像头采集的分辨率为1 920×1 088的工程车辆图像.选取主要出现的三类车辆作为检测目标:挖掘机、推土机和卡车,共有590张,包含挖掘机的有246张、推土机的有128张、卡车的有222张(一张图片可能有多个检测目标).考虑到原始数据集相对不足,容易产生过拟合现象,故通过添加高斯噪声、亮度调节、旋转、裁剪、平移、镜像等多种操作方式对数据集进行增强扩充,使数据集从590张扩充至3 300张.数据集使用Labelimg进行标注,并按7:2:1的比例将数据集划分为训练集、验证集和测试集.数据集的增强扩充使数据具有更好的多样性,有助于网络对检测目标的特征学习,原始图像和增强图像如图5所示.



图5 原始图像及增强图像:(a)原始图像;(b)亮度调节;(c)旋转;(d)镜像

3.2 实验条件与评价指标

1) 实验条件

本文在Ubuntu操作系统下基于Pytorch框架进行开发训练,具体硬件实验环境配置以及软件系统如下.

操作系统:Ubuntu操作系统;处理器(CPU)Intel(R)Xeon(R)Silver 4214R;内存:90 G;显卡(GPU):RTX 3080 Ti (12 GB);Pytorch:2.1.2;Python:3.8;CUDA:11.8.

算法运行参数如下.

Epochs参数:200;学习率:0.01;图片输入尺寸:640×640;Batchsize:16;梯度下降策略:带动量的SGD梯度下降策略;动量设置:0.937;权重衰减系数:0.000 5.

2) 评价指标

表1 各检测层的锚框设定

检测层	锚框设定
P2	(5,6) (8,14) (15,11)
P3	(10,13) (16,30) (33,23)
P4	(30,61) (62,45) (59,119)
P5	(116,90) (156,198) (373,326)

为满足工程车辆目标检测模型精度、轻量化以及实时性要求,本文选用准确率 P (Precision) 和召回率 R (Recall) 共同决定的平均精度 (AP), 再对所有类别的 AP 求取平均值得到的平均准确度 mAP 、每秒传输帧数 (FPS) 以及模型参数量 (Parameter Quantity) 作为评价指标. 其中: P 表示模型预测的正样本中真正存在的样本数量占所有预测为正样本的比例, R 表示模型预测为正样本的实际正样本数量占所有实际正样本的比例,利用准确率和召回率绘制 P - R 曲线并求其积分即可得到每个种类的平均精度 AP ,计算公式如下:

$$P = \frac{TP}{TP+FP} \quad (1)$$

$$R = \frac{TP}{TP+FN} \quad (2)$$

$$AP = \int P(R)dR \quad (3)$$

其中: TP 表示真正例,即预测为正例且实际也是正例的样本数; FP 表示假正例,即预测为正例但实际为负例的样本数; FN 表示假负例,即预测为负例但实际为正例的样本数.

$mAP@0.5$ 、 $mAP@0.5 : 0.95$ 分别表示IoU为0.5以及IoU从0.5至0.95,步长为0.05下的平均准确度. FPS 表示模型每秒处理图片的数量,模型的帧率越大,处理图片的速度越快,模型的检测速度也越快.模型参数量是评估模型权重大小的指标.

3.3 实验结果与分析

为验证本文提出的各部分改进方法对模型改进的有效性,对ETS-YOLO模型进行了一系列消融实验.消融实验以YOLOv5s为基础模型,按照ETS-YOLO具体改进部分分为四组实验.首先增加一个P2检测分支,然后引入EfficientViT网络替换YOLOv5s主干网络,最后利用soft-NMS算法代替原有NMS后处理算法.具体消融及运算结果如表2所示.

表 2 消融实验

模型	$mAP@0.5/\%$	$mAP@0.5 : 0.95/\%$	参数量/M	$FPS/(f/s)$
YOLOv5s	90.7	82.9	7.03	72
YOLOv5s+P2	92.9	83.2	7.69	57
YOLOv5s+P2+EfficientViT	92.2	83.0	5.90	62
YOLOv5s+P2+EfficientViT+soft-NMS	93.3	83.6	5.90	52

由表2可知各个改进模块对模型检测性能的影响.YOLOv5s模型增加一个检测层P2后, $mAP@0.5$ 提升了2.2%, $mAP@0.5 : 0.95$ 提升了0.3%,精度得到了提升,但模型参数量增加了9.4%,帧率下降了15 f/s,运算参量有所增加,运算效率有所降低.进一步把主干网络替换成EfficientViT网络后,相比上一步, $mAP@0.5$ 和 $mAP@0.5 : 0.95$ 分别下降了0.7%和0.2%,但参数量减少了23.3%,帧率增加了5 f/s.最后,替换了后处理NMS算法为soft-NMS算法,相比上一步, $mAP@0.5$ 提升了1.1%, $mAP@0.5 : 0.95$ 提升了0.6%,帧率下降了10 f/s.

图6 (a) 为YOLOv5s模型 P - R 曲线,图6 (b) 为ETS-YOLO模型 P - R 曲线,由两个模型的 P - R 曲线可知所有工程车辆类型的 mAP 值.根据 AP 值定义可知, AP 值越高, mAP 值越高,模型识别效果越好.其中,挖掘机和推土机的 mAP 值在改进前后变化较大,提升明显,卡车的 mAP 值在改进后略有下降.

由表2及上述结论综合可以看出,本文提出的ETS-YOLO模型与原YOLOv5s相比,平均精确度 $mAP@0.5$ 和 $mAP@0.5 : 0.95$ 都有提升,分别达到了93.3%和83.6%,且 $mAP@0.5$ 提升较大,而且参数量更少,只有5.90 M.但改进算法增加了额外算法,使得帧率有所降低,但52 f/s的帧率足可以用于定点位置工程车辆目标检测任务的监测布局.

3.4 工程车辆目标检测结果与分析

图7为复杂环境下工程车辆原图像以及分别用训练后的YOLOv5s、ETS-YOLO识别原图像的结果.图7 (a) 中有两种工程车辆,分别是卡车 (Trucks) 和推土机 (Bulldozer).其中,推土机被卡车部分遮挡.图7 (b) 为使用YOLOv5s的检测结果,只检测出了3辆卡车,被部分遮挡的推土机没有被检出;图7 (c) 为使用ETS-YOLO的检测结果,3辆卡车和推土机都成功被检出.通过对比可见,YOLOv5s模型在多目标、目标相互遮挡的场景出现漏检现象,而ETS-YOLO成功实现了复杂环境下对工程车辆的检测.

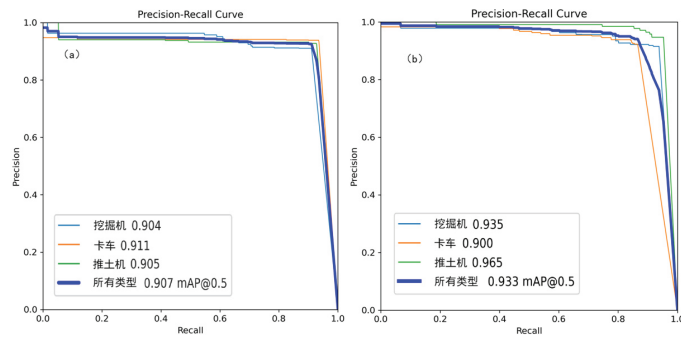


图6 P - R 曲线对比: (a) YOLOv5s模型; (b) ETS-YOLO模型



图7 原图像及模型检测结果: (a) 原图像; (b) YOLOv5s检测结果; (c) ETS-YOLO检测结果

4 结论

针对农村复杂环境中工程车辆小目标检测任务,利用EfficientViT网络的轻量化优点,增加可提高模型对浅层语义信息捕捉能力的小目标检测层,使用改善漏检的soft-NMS算法,提出了一种基于YOLOv5s的改进轻量化网络模型ETS-YOLO算法,在扩充数据集上及实验环境下的计算结果表明:改进后的ETS-YOLO模型与原模型相比,平均精确度 $mAP@0.5$ 达到了93.3%,提升了2.6%,参数量减少了16.1%,帧率降低了20 f/s,本文所提ETS-YOLO模型可以满足复杂环境背景小目标工程车辆的监测任务. ETS-YOLO模型为正在建设智能视频监控的市县提供了工程车辆目标检测识别的新方法,为有效减少土地和矿产违法行为提供了信息化技术支撑.

参考文献:

- [1] 周圣. 上海市违法建筑常态化治理面临的执行困境研究——以P区为例[D]. 上海: 华东师范大学, 2022.
ZHOU S. Study on the implementation dilemma of the normal management of illegal building in Shanghai: A case study of P district[D]. Shanghai: East China Normal University, 2022. (in Chinese)
- [2] 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.
- [3] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2023-05-20]. <https://doi.org/10.48550/arXiv.2004.10934>.
- [4] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. (2021-08-06)[2023-05-20]. <https://doi.org/10.48550/arXiv.2107.08430>.
- [5] LI C Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications[EB/OL]. (2022-09-07)[2023-06-25]. <https://doi.org/10.48550/arXiv.2209.02976>.
- [6] XU S L, WANG X X, LYU W Y, et al. PP-YOLOE: An evolved version of YOLO[EB/OL]. (2022-03-30)[2023-06-01]. <https://doi.org/10.48550/arXiv.2203.16250>.

- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [8] 高弋淞, 陈蔚卓, 王吉, 等. Android平台下的无人机输电线路施工车辆巡检[J]. 计算机系统应用, 2020, 29(2): 257-261.
GAO Y S, CHEN W Z, WANG J, et al. Detection of transmission lines construction vehicles for UAV based on Android[J]. Computer Systems & Applications, 2020, 29(2): 257-261. (in Chinese)
- [9] GUO Y P, XU Y, LI S L. Dense construction vehicle detection based on orientation-aware feature fusion convolutional neural network[J]. Automation in Construction, 2020, 112: 103124.
- [10] LI Z X, YANG L, ZHOU F Q. FSSD: Feature fusion single shot multibox detector[EB/OL]. (2017-12-04)[2023-06-01]. <https://doi.org/10.48550/arXiv.1712.00960>.
- [11] 顾晨亮, 杨恒, 刘友波, 等. 基于自适应局部斥力与归一化面积损失的工程车辆目标检测[J]. 中国安全生产科学技术, 2021, 17(11): 40-47.
GU C L, YANG H, LIU Y B, et al. Object detection of engineering vehicles based on self-adaptive local exclusion loss and normalized area loss[J]. China Production Safety Science and Technology, 2021, 17(11): 40-47. (in Chinese)
- [12] 颜文杰. 施工现场3D目标识别的关键技术研究[D]. 成都: 电子科技大学, 2022.
YAN W J. Research of key technologies of 3D object recognition on construction sites[D]. Chengdu: University of Electronic Science and Technology of China, 2022. (in Chinese)
- [13] 李建兴, 刘振宇, 马莹, 等. 改进YOLOv5在胡瓜钝绥螨品质管控系统中的设计与应用[J]. 南昌大学学报(工科版), 2023, 45(1): 85-95.
LI J X, LIU Z Y, MA Y, et al. Design and application of improved YOLOv5 in the quality control system of *Amblyseius cucumeris* Oudemans[J]. Journal of Nanchang University(Engineering & Technology), 2023, 45(1): 85-95. (in Chinese)
- [14] 蒲家鹏, 王雪梅, 高宏伟. 基于改进YOLOv4的航拍图像目标检测方法研究[J]. 沈阳理工大学学报, 2023, 42(3): 46-53.
PU J P, WANG X M, GAO H W. Research of aerial image object detection method based on improved YOLOv4[J]. Journal of Shenyang Ligong University, 2023, 42(3): 46-53. (in Chinese)
- [15] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [16] 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.
- [17] KATHAROPOULOS A, VYAS A, PAPPAS N, et al. Transformers are RNNs: Fast autoregressive transformers with linear attention[EB/OL]. (2020-08-31)[2023-10-01]. <https://doi.org/10.48550/arXiv.2006.16236>.
- [18] 徐杨, 熊举举, 李论, 等. 采用改进的YOLOv5s检测花椒簇[J]. 农业工程学报, 2023, 39(16): 283-290.
XU Y, XIONG J J, LI L, et al. Detection of pepper clusters using improved YOLOv5s[J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(16): 283-290. (in Chinese)
- [19] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS-improving object detection with one line of code[C]//2017 IEEE International Conference on Computer Vision(ICCV). Venice, Italy. IEEE, 2017: 5562-5570.
- [20] 韩文轩, 阿里甫·库尔班, 黄梓桐. 基于改进SSD算法的遥感影像小目标快速检测[J]. 新疆大学学报(自然科学版)(中英文), 2020, 37(2): 163-169.
HAN W X, ALIFU K, HUANG Z T. Small-object fast detection in optical remote sensing images based on a modified SSD[J]. Journal of Xinjiang University(Natural Science Edition in Chinese and English), 2020, 37(2): 163-169. (in Chinese)
- [21] TAN M X, LE Q V. EfficientNet: Rethinking model scaling for convolutional neural networks[EB/OL]. (2019-05-24)[2023-06-08]. <http://arxiv.org/abs/1905.11946v5>.
- [22] 王茹, 刘大明, 张健. Wear-YOLO: 变电站电力人员安全装备检测方法研究[J]. 计算机工程与应用, 2024, 60(9): 111-121.
WANG R, LIU D M, ZHANG J. Wear-YOLO: Research on detection methods of safety equipment for power personnel in substations[J]. Computer Engineering and Applications, 2024, 60(9): 111-121. (in Chinese)