

基于改进 YOLOv5 的轨道交通障碍物检测算法^{*}

赵鸿亮 郭佑民 王建鑫 杨 君
(兰州交通大学机电技术研究所 兰州 730070)

摘 要: 针对复杂的轨道交通背景下障碍物检测精度低和检测速度慢的问题,提出了一种改进 YOLOv5 的目标检测网络模型。首先,采用基于注意力机制的轻量级 Transformer 主干 EMO 来替换 YOLOv5 原有 backbone 中的部分模块,保证轻量化的同时,还能够提高模型的准确性和稳定性;其次,使用 Focal-EIoU 来替换 YOLOv5 中的 CIoU 损失函数,以解决 CIoU 引起的训练效率低、收敛速度慢等问题;最后使用轻量化上采样算子 CARAFE 来替换 YOLOv5 算法中原有的上采样层,在没有引入过多参数和计算量的情况下具有更大的感受野,提高了检测精度和检测速度。实验结果表明,该方法相较于原始的 YOLOv5 网络模型平均精确度提升了 11.1%,准确率提升了 13%,召回率提升了 11.4%,检测速度达到了 60.7 fps。所提出的方法在目标检测任务中表现出了较好的性能,有效增强了轨道交通背景下目标检测模型的检测性能。

关键词: 异物入侵;YOLOv5;注意力机制;损失函数;上采样

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.60

Rail transit obstacle detection algorithm based on improved YOLOv5

Zhao Hongliang Guo Youmin Wang Jianxin Yang Jun

(Institute of Mechanical and Electrical Technology, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: In order to solve the problems of low accuracy and slow detection speed of obstacle detection in the complex rail transit background, an improved object detection network model of YOLOv5 was proposed. Firstly, a lightweight Transformer backbone EMO based on attention mechanism was used to replace some modules in the original backbone of YOLOv5, which not only ensured the lightweight, but also improved the accuracy and stability of the model. Secondly, Focal-EIoU is used to replace the CIoU loss function in YOLOv5 to solve the problems of low training efficiency and slow convergence speed caused by CIoU. Finally, the lightweight upsampling operator CARAFE is used to replace the original upsampling layer in the YOLOv5 algorithm, which has a larger receptive field without introducing too many parameters and computational cost, and improves the detection accuracy and detection speed. Experimental results show that compared with the original YOLOv5 network model, the mean average precision of the proposed method is improved by 11.1%, the precision is improved by 13%, the recall is improved by 11.4%, and the detection speed reaches 60.7 frames per second. The proposed method shows good performance in the target detection task, and effectively enhances the detection performance of the target detection model in the context of rail transit.

Keywords: foreign body invasion;YOLOv5;attention mechanism;loss function;upsampling

0 引 言

近年来,随着社会经济水平的不断提高,轨道交通凭借着其快捷、安全、高效等突出优点,迅速成为了城市交通的重要组成部分^[1]。针对交通安全问题,目前已经出现了一些技术手段,如视频监控、智能检测^[2]等。基于机器视觉的轨道异物入侵检测作为一种重要的安全检测手段,能够有

效地提高轨道交通的安全性,减少因异物入侵引起的交通事故。

传统的检测方法存在着检测速度慢、精度低等问题,因此有很大的局限性。随着卷积神经网络的提出和深度学习的发展^[3-5],目标检测算法在实时性和精确度上都有了大幅度的提高。文献[6]改进了基于 Mask R-CNN 轨道交通障碍物检测方法,提出新的特征提取网络并综合多种多尺度

增强方法,提高了对小目标物体的检测精度。文献[7]通过对目标检测网络模型残差结构的针对性优化及训练方式的优化,提高了对于微小目标的识别准确率。文献[8]提出基于多尺度信息融合的目标检测算法和深度神经网络检测,来对铁路周边的入侵物体进行检测。文献[9]提出基于混合注意力机制来增强对于下目标物体的检测,以此应用于对铁路旁落石的检测,但是只能检测部分路段区域。

YOLOv5是一种基于单阶段目标检测的算法。通过优化网络结构、引入Anchor-free检测、多尺度训练和推理等技术,有效提升了目标检测的性能和速度。它在许多实际应用中取得了很好的效果,并成为目标检测领域的重要算法之一。文献[10]通过优化损失函数和提高生成检测框精度以达到识别目标的目的。文献[11]使用K-means聚类算法改进YOLOv5算法,提高了在雾霾天气下交通标志的识别精度。文献[12]提出的改进YOLOv5网络模型提升了对运动场景下小目标检测的精度。

为了解决轨道交通背景下远距离小目标物体和模糊目标物体检测精度较低的问题,本文提出了一种基于改进YOLOv5算法的轨道交通障碍物检测网络模型。首先使用基于注意力机制的轻量级Transformer主干高效模型(efficient model, EMO)模块来代替YOLOv5中的部分Backbone模块,以提高模型的语义信息的表达能力;其次,采用Focal-EIoU Loss来优化YOLOv5的损失函数,解决类别不平衡和目标定位精度不高的问题,提高模型的检测精度和定位精度。最后,使用轻量级上采样模块(content-aware reassembly of features, CARAFE)来替换YOLOv5中原有的上采样层,在没有引入过多参数和计算量的情况下具有更大的感受野,可以更好的利用周围的信息,以提高模型的分辨率和特征提取能力,使其能够更好地处理小目标和细节信息。改进后的YOLOv5模型在目标检测任务中表现出了较好的性能,能够更好地处理类别不平衡、小目标和细节信息等问题,具有更好的泛化性,显著提升了在轨道交通背景下目标检测模型的检测性能。

1 YOLOv5 网络模型

YOLOv5^[13-14]目标检测算法的官方代码中提供YOLOv5S、YOLOv5M、YOLOv5L和YOLOv5X四个目标检测网络模型。综合考虑到网络轻量化、检测精度和速度的要求,本文采用YOLOv5S模型作为基准进行改进。YOLOv5网络模型主要由Backbone、Neck和Head三个部分组成。

Backbone是骨干网络模块,它的作用是将输入的原始图像经过一系列卷积和下采样处理,转换为多层特征图,从而提取不同层次的图像特征。

Neck是特征融合模块,由特征金字塔网络(feature pyramid network, FPN)和路径聚合网络(path aggregation network, PAN)两部分组成。它的作用是将Backbone骨

干网络提取到的深层特征上采样,并与浅层信息逐元素相加,实现浅层图形特征和深层语义特征的融合,获得更为完整的特征信息。

Head是目标检测模块,主要由3个检测层组成,用于接收Neck模块输送的3组特征图,分别负责预测小目标、中目标和大目标。

2 改进的YOLOv5网络模型

如图1所示为改进的YOLOv5网络模型结构图。针对轨道交通障碍物检测所存在的问题,本文在YOLOv5网络模型的基础上进行了以下3点改进:首先使用基于注意力机制的轻量级Transformer主干;EMO模块来代替YOLOv5中的部分Backbone模块;其次,采用Focal-EIoU Loss来优化YOLOv5的损失函数;最后,使用轻量级上采样模块CARAFE来替换YOLOv5中原有的上采样层。

2.1 优化骨干网络

由于车载障碍物目标检测网络系统受计算资源和存储容量的限制,更加侧重于轻量级小模型网络。因此,本实验引入了一种基于注意力机制的轻量级Transformer主干EMO来替换YOLOv5中Backbone的部分模块。

本实验中的EMO是将Transformer中的多头注意力机制以及倒残差模块(inverted residual mobile block, iRMB)结合,并引入了Meta Mobile Block模型^[15],构建了一个轻量级的基于注意力机制的模型。该模块在保证高效性的同时,还能够提高模型的准确性和稳定性。

从图1中可以看出EMO在整体框架上由iRMB组成,没有其他复杂的操作符。图2所示为iRMB结构示意图,iRMB由两个子模块构成:深度可分离卷积模块(depthwise convolution, DW-Conv)和交互式注意力模块。

DW-Conv用于构建更加高效的基于注意力机制的模型。DW-Conv由改进的深度卷积(improved depthwise convolution)和逐点卷积(pointwise convolution)两个部分组成。

改进的深度卷积是一种深度可分离卷积,增加了一个可学习的分组卷积(grouped convolution)操作。这个分组卷积操作可以将通道分组,并对每个组使用不同的卷积核,从而增加了非线性性,同时保留了更多的特征信息。此外,这个深度卷积还使用了随机深度可分离卷积(random depthwise separable convolution)技巧,可以在保持相同计算量的情况下,增加模型的多样性和鲁棒性。逐点卷积与传统的逐点卷积相同,用于将改进的深度卷积的输出通道数降到所需的数量,从而进一步减少卷积层中的参数数量和计算复杂度。因为它可以将每个通道的输出与不同的权重相乘,所以增加非线性性,从而更好地捕捉特征之间的相互作用。

Attn Mat是交互式注意力模块中生成的大小为 $C \times C$ 的注意力矩阵,是交互式注意力模块中的一个输出。其中

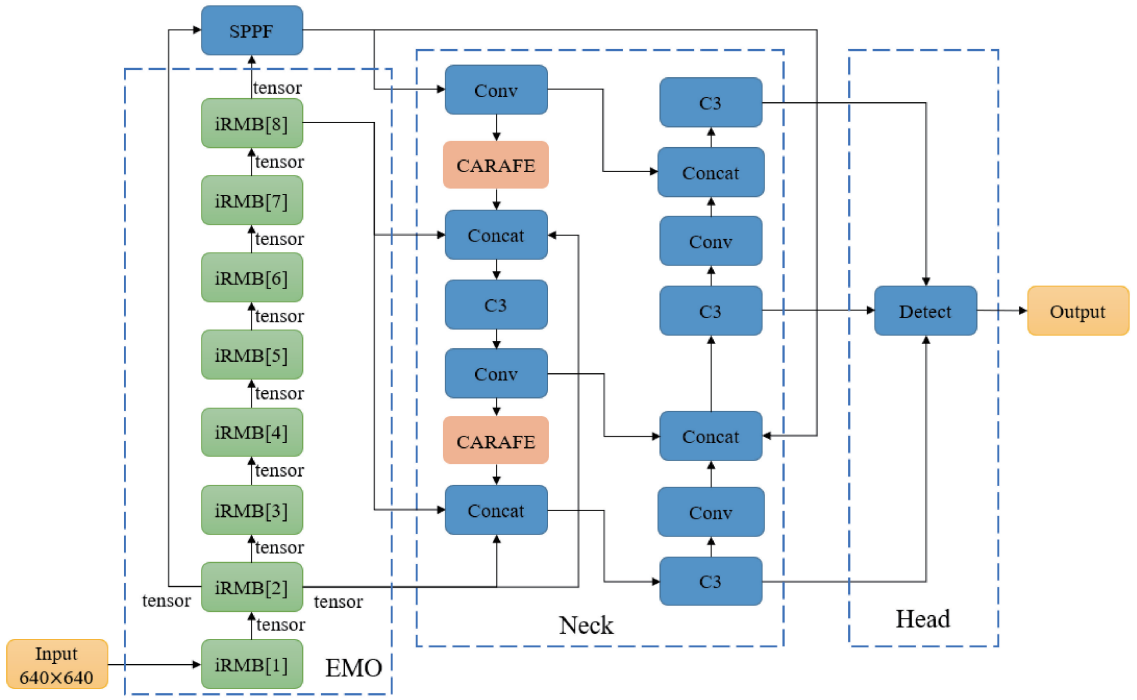


图 1 改进的 YOLOv5 网络模型结构图

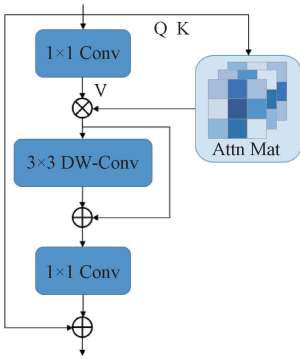


图 2 iRMB 结构示意图

C 是特征图的通道数,表示每个通道与其他通道之间的相关性和相互依赖关系,可以用于对特征进行加权聚合,从而更好地捕捉特征之间的交互和相关性,并提高模型的性能。

2.2 优化损失函数

YOLOv5 网络中的 Bounding Box 损失函数默认为 $\text{CIoU}^{[16]}$,该函数是通过计算预测框与真实框的重叠区域,中心点的距离和长宽比来计算 Loss。 CIoU 的计算公式如下:

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{C^2} + \alpha \nu \quad (1)$$

式中:

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{\text{gt}}}{h^{\text{gt}}} \arctan \frac{\omega}{h} \right)^2 \quad (2)$$

$$\alpha = \frac{\nu}{(1 - \text{IoU}) + \nu} \quad (3)$$

其中, b 表示预测框的中心点, b^{gt} 表示真实框的中心点, C 表示同时包含预测框和真实框的最小区域的对角线距离, ρ 是预测框和真实框中心点间的欧式距离, α 表示权重系数, ν 是预测框和真实框长宽比的差异性参数。

将 ν 分别对 ω 和 h 求偏导,可得到他们的梯度:

$$\frac{\partial \nu}{\partial \omega} = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{\text{gt}}}{h^{\text{gt}}} \arctan \frac{\omega}{h} \right)^2 \times \frac{h}{\omega^2 + h^2} \quad (4)$$

$$\frac{\partial \nu}{\partial h} = -\frac{4}{\pi^2} \left(\arctan \frac{\omega^{\text{gt}}}{h^{\text{gt}}} \arctan \frac{\omega}{h} \right)^2 \times \frac{\omega}{\omega^2 + h^2} \quad (5)$$

由上式可得 $\frac{\partial \nu}{\partial \omega} = -\frac{h}{\omega} \frac{\partial \nu}{\partial h}$,表明 ω 和 h 的梯度值成反比,

这也说明在训练过程中这两个参数不能同时增减,从而影响训练效果。并且 CIoU Loss 并未考虑到预测框和真实框的方向匹配问题,这也将会导致网络的效率变低、收敛速度变慢且回归结果不够精确。因此本实验引入了一种 Focal-EIoU^[17] 损失函数,该损失函数将 Focal Loss^[18] 和 EIoU Loss 两种方法进行了融合,同时引入了注意力机制,通过对不同特征层的注意力分配,提高了模型对关键区域的关注度,从而进一步提高了检测精度和效率。其公式为:

$$L_{\text{Focal-EIoU}} = \text{IoU}^\gamma L_{\text{EIoU}} \quad (6)$$

式中:

$$L_{\text{EIoU}} = L_{\text{IoU}} + L_{\text{dis}} + L_{\text{asp}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{(\omega^c)^2 + (h^c)^2} + \frac{\rho^2(\omega, \omega^{\text{gt}})}{(\omega^c)^2} + \frac{\rho^2(h, h^{\text{gt}})}{(h^c)^2} \quad (7)$$

其中, $\gamma = 0.5$, ω^c 和 h^c 是预测边界框与真实边界框的最小外接矩形的宽度和高度, ρ 是两点之间的欧式距离。

L_{IOU} 、 L_{dis} 、 L_{asp} 分别为 EIoU Loss 的 IoU 损失、距离损失、高宽损失。

由于 EIoU Loss 的计算是基于目标框的宽度和高度之比,因此对于长宽比例不一致的目标框,容易导致预测框与真实框之间的 EIoU 分数误差较大。同时 EIoU Loss 并没有考虑到正负样本之间存在的类别不平衡问题,这将会导致在训练过程中难以有效地区分出正负样本,从而影响模型的训练效果和检测精度。而 Focal-EIoU Loss 不仅考虑了基于目标框的宽度和高度的比例,还考虑了目标框的大小,从而能够更好地处理长宽比例不一致的目标框。Focal Loss 是一种针对类别不平衡问题的损失函数,它通过缩小易分类样本的权重来减轻类别不平衡的影响,Focal-EIoU Loss 也正是继承了这个优点来提升训练效果。Focal-EIoU Loss 同时解决了类别不平衡和目标定位精度不高的问题,具有较好的稳定性和适用性。

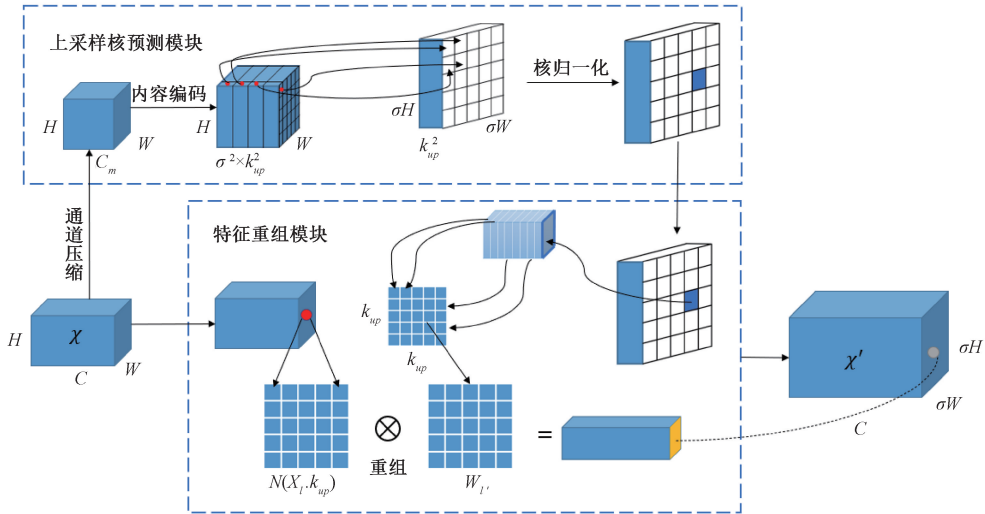


图3 CARAFE网络结构图

在上采样核预测模块,可以自适应地生成上采样核从而使得可变形卷积,根据特征图的内容自适应地调整采样位置和权重,提高特征表示的精度和准确性。如图3中所示,形状为 $H \times W \times C$ 的特征图可通过通道压缩减少可变形卷积中的通道数目,输出 $H \times W \times C_m$ 的特征图,来减少模型的参数数量和计算成本,同时保持模型的性能。内容编码由多个卷积层和池化层组成,可以对输入图像进行多次卷积和池化操作,从而逐步提取图像的特征。归一化处理的输入是预测得到的上采样核,输出则是归一化后的上采样核,使得它们的权重和为1,这样可以保证上采样核的权重分配是合理的,并且可以更好地控制采样点的位置和权重。

特征重组模块是将输入的特征图和上采样核预测后对应的特征图进行卷积和加权聚合,从而生成更准确的特征表示。与传统的上采样方法不同,CARAFE中的特征重组模块考虑了不同特征通道之间的相关性,并使用这些相

2.3 优化上采样

YOLOv5采用最邻近插值上采样算法增强特征图分辨率,算法简单且计算量小,但在边缘或线条的地方会出现锯齿状,对检测精度有一定的影响,因此本实验引一种轻量化上采样算子CARAFE^[19]来替换YOLOv5中原有的Upsample上采样层。

采用CARAFE算子替换原有的上采样算子能够有效降低图像锯齿等不良影响,提高特征表示的精度和准确性。同时,CARAFE算子的计算量相对较小,能够保持算法的高效性和实时性。相较于传统的上采样算子,CARAFE算子在没有引入过多参数和计算量的情况下具有更大的感受野,因此可以更好的利用周围的信息,在保持计算效率的同时,提高特征表示的精度和准确性。如图3所示为CARAFE网络的结构图,CARAFE算子模型主要由上采样核预测模块和特征重组模块组成。

关性来指导上采样过程中的信息重新组装。

最后,CARAFE算子将所有扩展后的小块合并为一个新的高分辨率特征图,作为整个上采样后的结果输出。

3 实验及结果分析

3.1 实验环境与数据集

本实验中使用的操作系统为Ubuntu 20.04,GPU为NVIDIA RTX3090,存为25.4 GB,CPU型号为6核Xeon Gold 6142,系统内存为60.9 GB,Python为v3.8,深度学习框架为PyTorch v2.0.0。

由于当前没有公开的轨道异物入侵数据集,本实验收集了1100张存在轨道异物的图片,并使用Labelimg对图片中的目标检测物进行标注。Pascal VOC2012数据集中包括了轨道交通中常见的行人、汽车和自行车等危险入侵物体,可以作为辅助数据集进行实验验证。因此,可以将之前标注好的文件添加到Pascal VOC 2012数据集中。最

终,本实验选取了 Pascal VOC 2012 数据集中的一些图片及其标注文件,得到了训练集图片 12 255 张,验证集图片 1 500 张,测试集图片 1 500 张,训练集和测试集的比例为 9 : 1。通过这一数据集的构建,本实验能够有效地评估所提出的改进方法在轨道异物入侵检测任务中的性能。同时,基于 Pascal VOC 2012 数据集的辅助数据集,还能够增加数据集的多样性和复杂性,提高模型的泛化能力。

3.2 评价指标

在本文实验中,使用平均精确度(mean average precision, mAP),准确率(Precision)和召回率(Recall)作为评价指标来评估模型的实时性能,上述指标的计算公式如下:

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \tag{8}$$

$$P = \frac{TP}{TP + FP} \tag{9}$$

$$R = \frac{TP}{TP + FN} \tag{10}$$

其中,TP 表示模型正确地识别出目标物体的数量,FP 表示模型将错误的非目标物体检测为目标物体的数量,FN 表示模型未能正确地识别出的目标数量, n 表示类别数。

3.3 实验结果分析:

如图 4 所示改进前后的 YOLOv5 网络模型在训练中的损失函数迭代过程对比图。训练的 epochs 设置为 300。

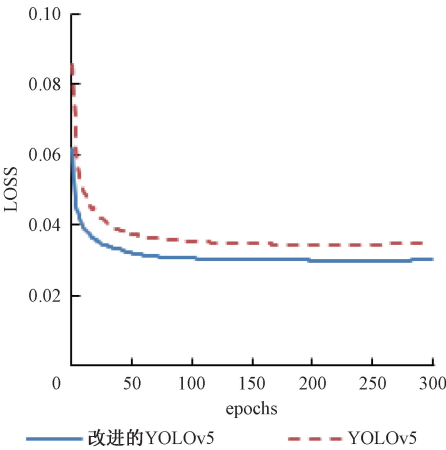


图 4 改进前后损失函数迭代过程对比图

从图 4 中可以看出,原始的 YOLOv5 网络模型的初始损失值约为 0.085,在训练 100 轮后,损失值下降到约 0.035 左右,并趋于稳定。而在使用 Focal-EIoU Loss 优化损失函数后,改进的 YOLOv5 网络模型的初始损失值约为 0.06,在训练 100 轮后,损失值下降到约 0.03 左右,并趋于稳定。这表明改进后的 YOLOv5 网络模型具有更高的检测精度,能够快速地逼近真实框。通过使用 Focal-EIoU Loss 进行损失函数优化,改进后的 YOLOv5 网络模型能够更好地处理类别不平衡和目标定位精度不高的问题,在

训练过程中收敛更快、更稳定,从而提高了模型的检测精度。

如表 1 所示,为了评估本实验引入的模块以及不同模块组合顺序对于算法性能优化的程度,设计了一系列消融实验进行分析,每个实验训练的 epochs 都为 300。其中“√”表示在 YOLOv5 网络模型的基础上加入了该方法。从表 1 中可以看出,本实验改进的 YOLOv5 网络模型相较于原始的 YOLOv5 网络模型,在平均精确度提升了 11.1%,准确率提升了 13%,召回率提升了 11.4%。特别地,在引入 EMO 后,改进后的模型在各项指标上均有显著提升。而在引入 Focal-EIoU Loss 优化损失函数后,虽然平均精确度和准确率有所下降,但是通过实验发现,相较于没有引入 Focal-EIoU Loss 的实验 6,最后的实验 8 在平均精确度提升了 5.4%,精确率提升了 4.9%,召回率上提升了 8.9%,因此引入 Focal-EIoU Loss 来优化损失函数是非常有必要的。

表 1 消融实验结果

序号	EMO	Focal-EIoU	CARAFE	mAP%	P%	R%
1				74.0	73.7	71.0
2	√			79.5	79.3	74.7
3		√		73.2	73.6	70.0
4			√	74.5	77.9	68.7
5	√	√		80.5	77.6	77.0
6	√		√	79.7	81.8	73.5
7		√	√	77.4	76.0	71.4
8	√	√	√	85.1	86.7	82.4

3.4 实验结果对比

如图 5 所示为检测效果对比图,其中图 5(a)为原始 YOLOv5 算法,图 5(b)为本文改进的 YOLOv5 算法。



图 5 检测效果对比图

在图 5(a)中可以看到原始的 YOLOv5 算法在远处小目标物体和模糊目标物体的检测方面存在一定的挑战,可能会出现漏检或误检的情况。在图 5(b)中本文改进的 YOLOv5 算法能够更好地应对这一问题,实现了对远处小目标物体和模糊物体目标的准确检测。通过实验证明,本文改进的 YOLOv5 算法在实际应用场景中具有更高的实用性和可靠性。

为了进一步验证本实验改进的 YOLOv5 网络模型的性能,在同一数据集、同一配置环境中,训练的 epochs 为 300。将本实验改进的 YOLOv5 网络模型与 Faster R-CNN、YOLOv3、YOLOv5 以及 YOLOv8 等主流目标检测算法进行了实验对比。不同检测网络模型对比实验结果如表 2 所示,从表 2 中可以看出,改进的 YOLOv5 网络模型在各个方面都优于其他算法,具有更高的平均精确度、准确率和召回率,同时每秒检测的图像帧数(frame per second, FPS)为 60.7,实时性也达到了要求。通过对比实验,进一步证明了本实验改进的 YOLOv5 网络模型在轨道异物入侵检测任务中的有效性和优越性。

表 2 不同检测网络模型对比实验

网络模型	mAP%	P%	R%
Faster R-CNN	73.5	75.3	70.2
YOLOv3	77.6	78.8	73.5
YOLOv5	74.0	73.7	71.0
YOLOv8	74.5	75.3	68.4
改进的 YOLOv5	85.1	86.7	82.4

4 结 论

本文提出了一种基于改进 YOLOv5 的目标检测网络模型,以提高轨道交通背景下障碍物检测的效率。在 YOLOv5 网络模型的基础上,针对骨干网络模块,损失函数模块和上采样模块进行了改进和优化。实验结果表明,在本实验的数据集中,改进后的网络模型相较于主流的目标检测网络模型在平均精确度、准确率和召回率方面都有大幅提升。同时,每秒检测的图像帧数达到了 60.7,符合检测实时性的要求。综上所述,本文提出的改进方法在轨道交通背景下的障碍物检测任务中取得了较好的效果,能够有效地提高了模型的检测性能和实时性。通过采用这些改进方法,不仅可以提高 YOLOv5 模型的性能,也可以为其他目标检测算法的优化提供参考。

参考文献

[1] 韩宝明,李亚为,鲁放,等. 2021 年世界城市轨道交通运营统计与分析综述[J]. 都市快轨交通,2022,35(1): 5-11.

[2] 赵朵朵,章坚武,傅剑峰. 基于深度学习的实时人流统计方法研究[J]. 传感技术学报,2020,33(8): 1161-1168.

[3] 王子冠,弋国华. 基于传统图像处理算法和深度学习的

轨道区域识别研究[J]. 电气自动化,2019,41(4): 111-114.

[4] LONG J, SHELLHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. Boston, Massachusetts, USA: IEEE, 2015, 3431-3440.

[5] 赵永强,饶元,董世鹏,等. 学习目标检测方法综述[J]. 中国图象图形学报,2020,25(4):629-654.

[6] HE D, QIU Y, MIAO J, et al. Improved mask R-CNN for obstacle detection of rail transit [J]. Measurement, 2022, 190(1):110728-110739.

[7] 初帆. 基于优化 YOLOv3 网络的在途列车障碍物检测方法[J]. 现代城市轨道交通,2021(6):19-23.

[8] 朱力强,许力之,周鑫,等. 高速铁路周界快速识别算法研究[J]. 铁道学报,2022,44(12):56-64.

[9] 胡昊,史天运,关则彬. 融合混合注意力和改进 YoloX 的铁路落石检测方法[J]. 电子测量技术, 2022, 45(20):110-116.

[10] 吕禾丰,陆华才. 基于 YOLOv5 算法的交通标志识别技术研究[J]. 电子测量与仪器学报,2021,35(10):137-144.

[11] 朱开,陈慈发. 基于 YOLOv5 的雾霾天气下交通标志识别[J]. 电子测量技术,2023,46(8):31-37.

[12] 朱瑞鑫,杨福兴. 运动场景下改进 YOLOv5 小目标检测算法[J]. 计算机工程与应用,2023,59(10):196-203.

[13] YANG G, FENG W, JIN J, et al. Face mask recognition system with YOLOV5 based on image. recognition [C]. 2020 IEEE 6th International Conference on Computer and Communications(ICC), IEEE, 2020;1398-1404.

[14] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]. IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, 2017,440(56):6517-6525.

[15] MARK S, ANDREW H, MENGLONG ZH, et al. Mobilenetv2: Inverted residuals and linear bottlenecks. In CVPR, 2018;4510-4520.

[16] REZATOFIGHI H, TSIO N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach: IEEE, 2019. 658-666.

[17] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[C]. Computer Vision and Pattern Recognition, 2021: 4321-4331.

[18] LIN T Y, COYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2):318-327.

[19] WANG J, CHEN K, XU R, et al. Carafe: content aware reassembly of features[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision,2019:3007-3016.

作者简介

赵鸿亮(通信作者),主要研究方向为计算机视觉、图形与图像处理。
E-mail:2393048967@qq.com

郭佑民,硕士,教授,主要研究方向为传感器与检测技术、计算机技术。
E-mail:15251923797@163.com