

DOI:10.19651/j.cnki.emt.2209680

融合混合注意力和改进 YoloX 的铁路落石检测方法^{*}

胡昊¹ 史天运² 关则彬³

(1. 中国铁道科学研究院研究生部 北京 100081; 2. 中国铁道科学研究院集团有限公司 北京 100081;
3. 中国铁道科学研究院集团有限公司电子计算技术研究所 北京 100081)

摘要: 铁路周边危岩落石侵入铁路限界内会严重危害旅客生命财产与铁路行车安全。针对传统检测方法对复杂动态环境识别误检高、小目标识别准确率较低等问题,提出一种基于视频的铁路落石入侵检测深度学习算法。首先,融合混合注意力模块,增强网络对与背景相似落石检测能力。其次,改进 YoloX 部分网络结构为双向特征金字塔网络,加强了不同层级特征的相互交流,提升小目标识别能力。同时采集大量不同场景模拟落石数据,构建模拟落石数据集,并在训练中使用 Mosaic 数据增强方法,增强方法的泛化能力。实验结果表明,本文方法随着改进模块的添加,识别精度不断提高。对比多种主流目标检测方法,取得了最高识别准确度,不同大小目标识别稳定,证明了本文算法在实际铁路场景的应用价值。

关键词: 铁路运输;落石检测;注意力模块;双向特征金字塔网络

中图分类号: U298 文献标识码: A 国家标准学科分类代码: 580.70

A railway rockfall detection method incorporating mixed attention and improved YoloX

Hu Hao¹ Shi Tianyun² Guan Zebin³

(1. Postgraduate Department, China Academy of Railway Sciences, Beijing 100081, China;

2. China Academy of Railway Sciences Corporation Limited, Beijing 100081, China;

3. Institute of Computing Technologies, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China)

Abstract: Dangerous rocks and falling rocks around the railway intrude into the railway boundary, which will seriously endanger the life and property of passengers and the safety of railway traffic. Aiming at the problems that traditional detection methods have high false detection in complex dynamic environment and low accuracy of small target recognition, a video-based deep learning method for railway rockfall intrusion detection is proposed. First, a hybrid attention module is incorporated into the network structure, which can enhance the network's ability to detect rockfalls similar to the background. Secondly, part of the network structure of YoloX is improved to a bidirectional feature pyramid network, which strengthens the mutual exchange of features at different levels and improves the ability to identify small targets. Simultaneously collect a large number of simulated rockfall data from different scenarios, build a simulated rockfall data set, and use the Mosaic data enhancement method in training to enhance the generalization ability of the method. The experimental results show that with the addition of improved modules, the identification accuracy of the method in this paper is continuously improved. Compared with various mainstream target detection methods, the highest identification accuracy is achieved, and the identification of different sizes targets is stable, which proves the application value of the algorithm in this paper in the actual railway scene.

Keywords: railway transportation; rockfall detection; attention module; bidirectional feature pyramid network

0 引言

目前中国已经成为世界上最高速度、最长里程、最大规

模的高速铁路运营国家。我国高铁沿线自然环境多样、气候条件复杂多变,在一些山地区域的铁路地段,当遇见强降雨、大风等恶劣天气或者地质条件不稳定时,极易发生岩石

收稿日期:2022-04-20

* 基金项目:中国国家铁路集团有限公司科技研究开发计划课题(K2020G022)项目资助

崩塌、山体滑坡等自然灾害。山体岩石落入铁路限界内时,会给铁路旅客生命财产安全以及铁路基础设施安全带来严重的危害。这些事件发生具有突发性,传统人工巡检的方式有工作量大、漏检多等问题。当前,高速铁路自然灾害及异物侵限监测系统是一种防护网阻拦方式对铁路限界进行异物监测,但是防护网缺损或者异物从防护网上进入铁路时,仍会导致落石异物侵限时发生。所以利用其他监测方式对铁路落石异物侵限进行同步监测尤为重要。

现有的铁路落石等异物监测方法,根据其原理可分为接触式与非接触式^[1]。接触式有:张力围栏、光纤围栏等,主要通过感知异物接触时产生不同信号变化,进行分析监测。非接触式有:雷达、视频、红外等,通过自身发射光波或电磁信号,或者接收图像信息,通过分析处理,进行异物监测。其中,视频分析方式因为设备安装简单、结果可视化方便等优点,是非接触式监测系统的主要检测方式。

当前很多学者针对落石等异物检测,开展许多研究工作。传统方法大都是基于背景建模,有许多开创性的研究,郭碧等^[2]提出了一种帧间差分累积的铁路异物识别方法,降低了背景噪声影响,有效对前景异物目标进行定位与提取。张怡等^[3]针对跑道异物,提出一种基于差分 GPS 定位的图像图像差分异物检测方法,通过高精度位置信息将两幅图像进行对齐操作,接着对图像进行差分异物检测,可以有效检测小尺寸目标。郭保青等^[4]针对铁路场景采集图像抖动和场景复杂多变的问题,提出快速背景差分的异物检测算法,提高速度的同时,解决了鬼影和抖动问题。伍海萍等^[5]针对动态入侵异物鬼影问题,提出了改进 ViBe 算法,抑制了和减少鬼影的帧数,同时保证了异物检测精度。史红梅等^[6]在 FPGA 和 ARM 芯片上,实现了基于支持向量机的差分图中异物目标分类,并利用卡尔曼滤波器设计了目标跟踪算法,能有效识别区域内目标,检测速度达到 13 帧/s。侯涛等^[7]提出基于改进 MOG-LRMF 异物侵限算法,建立了背景模型,并对前景目标进行检测,在光线较弱、背景复杂、多目标的场景下检测具有更好的鲁棒性。王尧等^[8]针对当前方法检测精度差、误检率高,提出一种新的基于条件随机场 CRF 的前景提取算法,实验表明,在动态背景、伪装色等问题上实现快速求解检测。

近年来,随着深度学习的广泛应用,基于神经网络的异物检测算法也取得了许多进展,于晓英等^[9]针对只能识别静态障碍物且识别速度较慢的问题,设计了基于 YOLO 的铁路侵限异物检测方法,对比 AlexNet 及 Adaboost 算法具

有更好的检测速度。王洋等^[10]提出了一种基于 CNN 的特征快速提取算法,简化其全连接网络结构,并加入稀疏性参数,实现了快速检测。张剑等^[11]提出了一个深度学习检测模型,增加了空洞卷积和改进了 FPN 结构,对于不同尺度目标检测精度均进行了提升。徐岩等^[12]针对传统检测算法识别精度低、分类不明确的等问题,提出了一种基于 Faster RCNN 的检测方法,针对铁路场景进行适应性改进,对人、车、动物的识别性能较好。Tian 等^[13]提出一种针对高速铁路场景的多尺度异物检测算法,采用双向迭代深度聚合方法融合不同层特征,在公开数据集与自建数据集上,对比其他方法取得了更高的准确率与速度。上述方法中,基于背景建模方法,在简单场景和光照条件较好的情况下,检测结果较好,但是在复杂与动态背景条件下,检测性能会有所下降。基于深度学习的方法,在环境抗干扰方面,具有优势,更便于实际应用。但是受限制于数据样本的欠缺,检测能力会受影响,同时神经网络属于统一检测,对待不同大小目标检测能力不同,提升小目标检测能力,也是需要重点解决的问题。

因此,本文针对上述异物检测当前问题,提出了一种融合混合注意力和改进 YoloX 的铁路落石检测方法。首先,基于最新的 YoloX 框架,设计铁路落石检测方法。其次,针对与背景相似落石检测难题,在网络结构中融合了混合注意力模块,提升网络相似特征的检测能力。接着,为提升不同尺度落石的检测能力,将 YoloX 的部分网络结构改为双向特征金字塔网络,加强了不同层级特征的相互交流。同时,针对真实落石样本极少的问题,本文在不同天气、不同铁路场景下采集了模拟落石数据,建立了模拟数据集。在训练过程中使用 Mosaic 数据增强方法,进一步丰富落石样本集,使网络学习到各种不同落石特征,提高落石检测场景泛化能力。最终通过多组实验,证明本文算法的有效性

1 融合混合注意力与双向特征金字塔网络的铁路落石检测网络

本文提出的铁路落石异物检测算法,是基于混合注意力与双向特征金字塔网络改进 YOLOX 目标检测框架进行实现,整体网络结构如图 1 所示。YOLOX 算法是由旷视公司于 2021 年提出了新型目标检测框架^[14],主要是基于 YOLOv3 算法的改进,网络改进主要包括主干网络结构、分类与回归解耦头、基于无锚框机制以及动态匹配正样本。

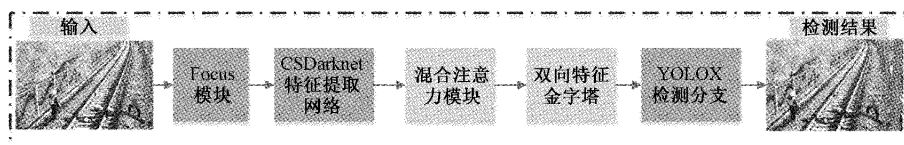


图 1 落石检测流程图

1.1 YOLOX 主干网络结构

YOLOX 模型以 YOLOv3-Darknet53 为基线进行修改,主干网络结构如图 2 所示。主干网络中结合跨阶段局部网络 (cross stage partial network, CSPNet) 构成 CSPDarknet 以提高网络学习性能。主干网络均由残差网络 (Residual) 结构构成,残差网络包括主干部分和残差边部分,主干部分由一个 1×1 和一个 3×3 卷积构成,残差边

不进行任何操作,之后将以上两部分进行拼接。由于网络中重复梯度信息会对网络计算性能和推理速度造成较大影响,CSPNet 在残差块堆叠的基础上,构成一个大的残差结构,主干部分为上述残差块,残差边经过少量处理与主干部分的输出进行拼接构成 CSPNet 的输出。CSPNet 利用跨阶段特征融合策略和截断梯度流来增强不同层内学习特征的可变性。

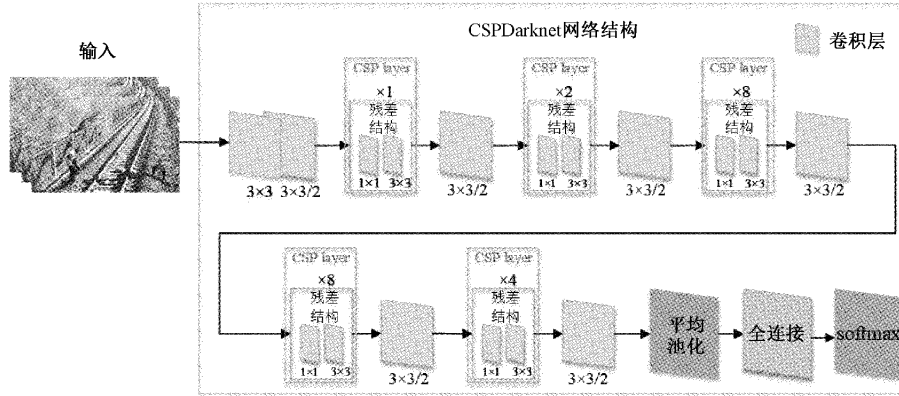


图 2 主干网络结构

YOLOX 主干部分增加了 Focus 模块,将特征图纵横方向上每间隔一个特征点提取子特征图,从而获得 4 个子特征图,对这四部分进行堆叠,最终得到的特征图长宽减半,通道数为原来的 4 倍。该模块提高特征点感受野,减少信息丢失的同时,降低了网络参数规模,提高网络训练推理速度。

YOLOX 主干网络中还使用了空间金字塔池化 (spatial pyramid pooling, SPP) 结构。SPP 将输入通过不同大小的最大池化核 ($1 \times 1, 5 \times 5, 9 \times 9, 13 \times 13$) 获取不同级别感受野信息,在将输入与各最大池化的输出进行拼接获得 SPP 的最终输出,该结构有效融合了各级特征信息,网络中加入 SPP 模块在扩大感受野的同时,对网络推理速度几乎不产生影响。

1.2 注意力模块

铁路落石异物侵限检测算法需要全天候运行,夜晚时视频能见度降低,落石与背景区别降低,且落石在画面不同位置像素面积差异较大,为了提升落石检测算法的鲁棒性。本文在 YoloX 主干网络 CSPDarknet 的最后三个 CSP 层后面加上混合注意力模块^[15],通过嵌入注意力机制,可以增强网络特征提取能力,选择目标更关键的信息^[16]。

1) 通道注意力

通道注意力模块,对于输入特征图分别从宽和高维度进行平均池化和最大池化操作,接着送入多层感知机构成的共享网络,然后输出全局平均池化和全局最大池化后的通道注意力特征,并逐元素求和等到特征向量,经过 sigmoid 的得到通道注意力学习的权重,将权重特征沿通道方向乘到输入的特征图上,得到最后的特征图。

2) 空间注意力

空间注意力模块,对于输入特征图分别从通道维度进行平均池化和最大池化操作,形成 2 个维度相同但通道为 1 的特征图,接着在通道方向进行合并,并送入输出通道为 1 的卷积,再经过 sigmoid 得到了空间维度与输入特征图相同、通道维度为 1 的空间注意力,最后将特征图和空间注意力相乘到微调后的特征图。

3) 混合通道注意力

混合通道注意力结构如图 3 所示,其是将通道注意力与空间注意力进串联,特征图首先进行通道方向的注意力操作,紧接着进行宽高方向的注意力操作,最后得到注意力特征图。混合注意力使得网络在学习的过程中,可以更突出前景落石目标部分的权重,这样相当于对原始特征图进行了加权,从而使得网络在学习过程中自主地学习到更关注于前景落石的特征。

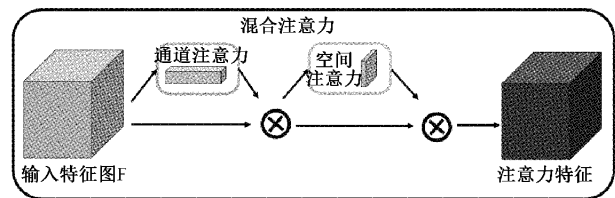


图 3 混合注意力模块结构

1.3 双向特征金字塔网络

随着网络的深度的不断增强,网络提取的特征会出现一定程度的丢失。特征金字塔网络是解决该问题的有效方式,它可以融合不同层级的特征。YoloX 的主干网络在特征提取后,将中间层、中下层与底层的特征送入特征金

字塔网络中,由其对自下而上和自上而下的特征进行融合。以往特征融合方法是对所有输入特征一视同仁,双向特征金字塔网络^[17]在同一尺度的特征间添加横向连接,并引入了加权策略,如式(1)所示。最终输出的特征图结合了当前层与上下层共三层的特征。

$$Out = \sum_i \frac{\omega_i}{\epsilon + \sum_j \omega_j} \cdot In_i \quad (1)$$

其中,Out 与 In 分别表示输出与输入特征,通过在每个 ω_i 之后施加 Relu 激活函数来保证权重 $\omega_i \geq 0$,学习率 $\epsilon = 0.0001$ 为一个小值,以避免数值不稳定。

其结构如图 4 所示。

本文针对 YoloX 特征金字塔网络部分修改为双向特征金字塔网络,增加了中间特征层的横向连接,同时为适应主干网络结构,将 BiFPN 网络简化为 3 层,不同层级的特征经过融合后分别输出到 YoloX 的 3 个检测头。经过

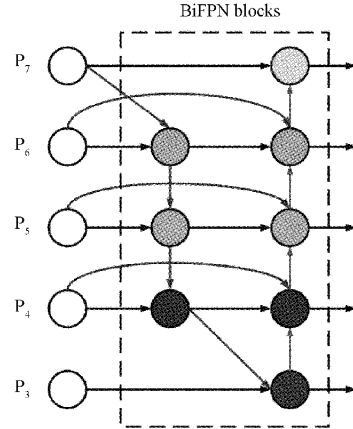


图 4 双向特征金字塔网络结构

混合注意力机制与双向特征金字塔网络改进的 YoloX 网络结构如图 5 所示。

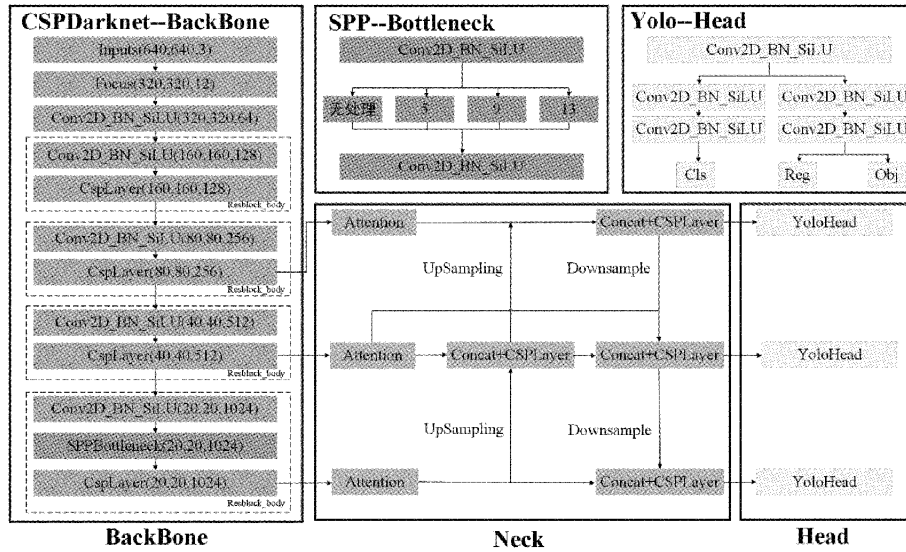


图 5 改进 YoloX 网络结构

2 网络训练

2.1 数据采集

虽然铁路沿线落石事件发生后危险性很大,但是真实铁路落石的样本相对来说仍然很少。因此,除了收集真实发生的落石样本,本文在真实铁路环境下模拟落石掉落铁路线内,并利用沿线综合视频系统进行数据录制,最后将模拟落石图像制作为模拟落石样本。根据视频监控图像模式识别的技术特点,其对入侵的监测范围一般在铁道轨面附近区域,据此设计模拟异物侵限数据集的内容及方法。如图 6 所示,沿铁轨垂直轨道方向划分为 A、B、C 三个区域。分别在不同天气状况下的白天与夜晚,模拟落石抛入 A、B、C 轨面区域,同时在离摄像头不同的距离进行模拟采集。具体的异物样本采集方案如表 1 所示,采集用例

包括单个与多个石块,分别在白天、夜间有环境光及夜间无环境光的条件下进行等间隔数据采集,采集范围为距离摄像机 0~100 m。

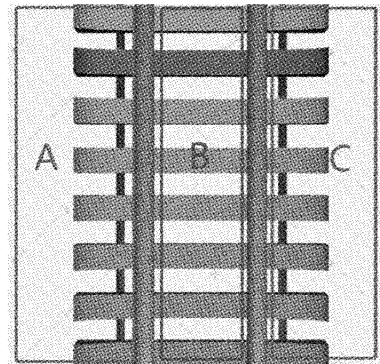


图 6 数据采集区域示意图

表 1 异物样本采集方案

采集用例	采集条件(不同天气状况下:晴天、阴天、下雨、下雪、大雾等)
1. 单个石块	白天:0~100 m,以 1 m 为间隔
	夜间有环境光:0~100 m,以 1 m 为间隔
	夜间无环境光:0~100 m,以 1 m 为间隔
2. 多个石块	白天:0~100 m,以 1 m 为间隔
	夜间有环境光:0~100 m,以 1 m 为间隔
	夜间无环境光:0~100 m,以 1 m 为间隔

在铁路沿线 32 个不同的摄像机进行落石数据集采集,以 1 m 为最小采样间隔,100 m 为最大采样距离,抛入异物样本选用泡沫模拟石头,实验购置的用来模拟落石的实物样本共 16 个,如图 7 所示,样本规格、数量如表 2 所示。最终标注模拟落石样本 2.2 万张,不同条件下的模拟落石样本如图 8 所示。

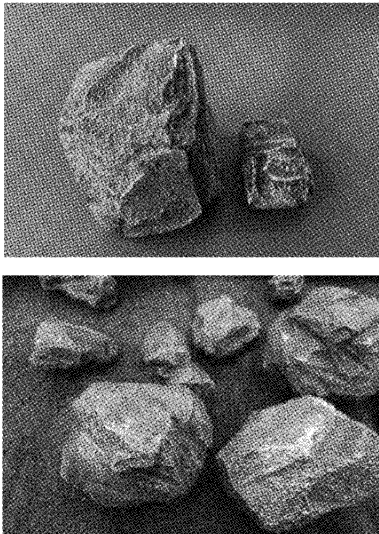


图 7 实验样本

表 2 模拟异物样本规格

体积规格/cm	物体编号	数量
10×10	0	4
15×15	1	2
20×20	2	2
30×30	3	2
40×40	4	2
50×50	5	2
60×60	6	1
70×70	7	1

2.2 数据增强

实验时发现,模型在检测 30×30 像素大小的异物时,有可能发生漏检,这是由于小目标在样本数据中占比过

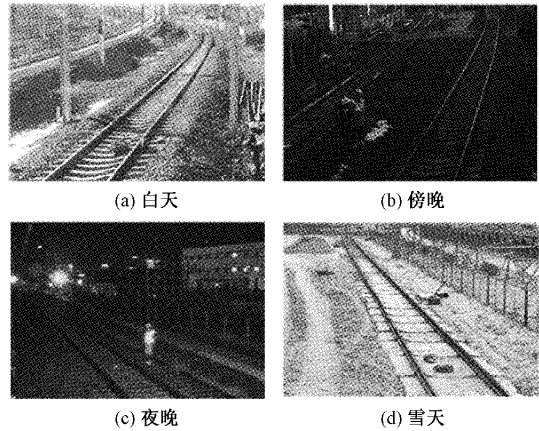


图 8 不同条件下的模拟样本

少。数据增强方法可以一定程度上解决这个问题,有效扩充不同样本。实验中使用了 Mosaic^[18] 数据增强方法,其借鉴了 CutMix 数据增强的思想,与 CutMix 混合训练集中两张图片不同,Mosaic 从训练数据中抽取四张样本,经过裁剪拼接生成一个新的样本,由于新样本融合了四张图片的信息,待检测目标的背景信息更加丰富,同时该方法使得最小 Batch_size 达到 4,显著提高了训练效率,Mosaic 数据增强后样本如图 9 所示。



图 9 Mosaic 数据增强

2.3 实验参数与训练过程

模型的好坏取决于模型训练结果。实验在 RTX 2080ti GPU 上完成,使用 Pytorch 深度学习框架。特征提取网络为 CSPDarknet,接着加了混合注意力模块与双向特征金字塔网络,最后送入检测头进行检测。网络在 2 张 GPU 中训练完成,训练数据集如 2.1 小节,YOLOX 训练迭代次数设置为 300 轮,批量大小设为 4。初始学习率为 0.001,采用 Adam 调整学习率。需要注意的是,实验中增加了 Mosaic 数据增强方法,在训练的最后 15 个迭代中,数据增强模块将被移除,实验结果证明训练过程中始终保持数据增强模块对网络性能并未有提升。

2.4 损失函数设计

YOLOX 损失函数的计算由回归损失,目标损失以及分类损失三部分构成,总损失为:

$$total_loss = l_{reg} + l_{obj} + l_{cls} \quad (2)$$

1) 回归损失(reg loss),计算真实框和预测框的交并比(IoU)损失:

$$l_{reg} = \lambda_{reg} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{(obj)} (2 - w_i - h_i) [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2] \quad (3)$$

其中, λ_{reg} 为回归损失所占的权重, S^2 为网格数, $(\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i)$ 为预测结果, (x_i, y, h_i, w_i) 为 Ground-truth 真实值。

2) 目标损失(obj loss),特征图整体由前景点与背景点组成,计算真实框与预测框的交叉熵损失作为目标损失:

$$l_{obj} = \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{(noobj)} (c_i - \hat{c}_i)^2 + \lambda_{obj} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{(obj)} (c_i - \hat{c}_i)^2 \quad (4)$$

其中, λ_{noobj} 为非目标损失权重系数, λ_{obj} 为目标损失权重系数, $1_{i,j}^{(noobj)}$ 的含义为若网格 i 的预测框 j 为负样本,则取值为 1,否则该值取 0; $1_{i,j}^{(obj)}$ 的含义为若网格 i 的预测框 j 为正样本,则取值为 1,否则该值取 0。 \hat{c}_i 为预测样本类型(正/负样本), c_i 为真实样本类型。

3) 分类损失(cls loss),计算真实框类别与预测框预测类别的交叉熵损失作为分类损失:

$$l_{cls} = \lambda_{class} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{(obj)} \sum_{c \in classes} p_i(c) \log(\hat{p}_i(c)) \quad (5)$$

其中, λ_{class} 为分类损失权重系数, $\hat{p}_i(c)$ 为预测类别概率值, $p_i(c)$ 为真实类别概率值。

3 结果分析

3.1 评价指标

为了对本文方法识别效果进行评判,目标检测算法中常选用检测精度(AP)指标,该指标由检测准确率(precision)和召回率(recall)确定。另外选用推理时间作为参考指标。

根据模型检测结果与 Ground-Truth 之间的关系,可将预测结果分为真正样本(true positive, TP)、假正样本(false positive, FP)、假负样本(false negative, FN)、真负样本(true negative, TN),据此可定义模型准确率与召回率。准确率表征了模型的查准率,召回率表征了模型的查全率,其计算方式如下:

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

单从准确率或者召回率的角度来评价模型性能具有片面性,高准确率不能确保模型能够尽可能多的预测到所有正样本,同样,高召回率也可能出现过多地负样本预测为正样本的情况。因此,需要将二者结合起来才能更全面地评价模型的检测能力,AP 即通过不同阈值下检测结果的 P-R 曲线来对模型的检测精度进行衡量。

3.2 消融实验

为了验证本文中不同改进方法对模型识别性能的影响,我们进行了多组实验对改进模块进行分析,所有实验都设置相同的训练参数,在同一数据集上进行训练。不同模型检测结果如表 3 所示,其中“√”代表在 YoloX 模型中使用了该改进方法,“×”代表未使用该改进方法。由表中结果可知,在本文的数据集上,两种改进方法都取得了检测能力的提升。首先,添加了混合注意力模块后,提升了网络对空间和通道的注意力,增强了不同大小落石的检测能力,AP 提升了 1.3%。其次,改为双向特征金字塔网络后,提高了不同层级网络交流能力,提升网络对局部和全局信息的学习能力,AP 提升了 2.1%。

表 3 不同改进方法的实验结果

方法	混合注意力	双向特征金字塔网络	AP	FPS
YOLOX	×	×	0.909	56.5
改进 1	√	×	0.922	53.1
改进 2	√	√	0.943	52.6

3.3 不同方法对比实验

使用 YOLOv5-X、YOLOv5-S、Faster RCNN 作为对比方法,与本文提出的基于混合注意力和双向特征金字塔网络改进 YOLOX 的铁路落石检测模型进行性能比较。YOLOv5-X 和 YOLOv5-S 是 YOLOv5 的不同版本,其主要区别在于网络的深度和宽度有所不同。其中 YOLOv5-S 的参数数量最少。分别使用四种模型对数据进行训练和测试,其准确率、召回率以及精度如表 4 所示。当取 IOU 阈值为 0.5 时,YOLOX 的 AP 较 Faster RCNN、YOLOv5-X、YOLOv5-S 高 13.2%、1.5%、3.4%;当取 IOU 阈值为 0.5~0.95 时,取其平均值,YOLOX 的 AP 较 Faster RCNN、YOLOv5-X、YOLOv5-S 高 6.2%、3.3%、37. %。由此可知,改进后的方法,在不同 IOU 阈值时,都取得了最好的检测精度。说明本文方法具有更好的检测性能,改进方法能有效提高落石识别能力。

表 4 指标对比

方法	P	R	AP@.5	AP@.5:.95
Faster RCNN	0.901 6	0.891 2	0.811	0.399
YOLOV5-X	0.954 4	0.893 5	0.928	0.428
YOLOV5-S	0.939 3	0.879 1	0.909	0.424
本文	0.976 7	0.893 6	0.943	0.461

同时基于检测距离对本文提出的算法进行评估。实验选择了四个不同场景的数据进行测试, 落石抛入地点据摄像机位置距离分别为 25 m、50 m、75 m 以及 100 m。场景示意图如图 10 所示。结果表明, 随着检测距离的增加, 算法检测性能损失虽然有所降低, 但是识别性能保持稳定。

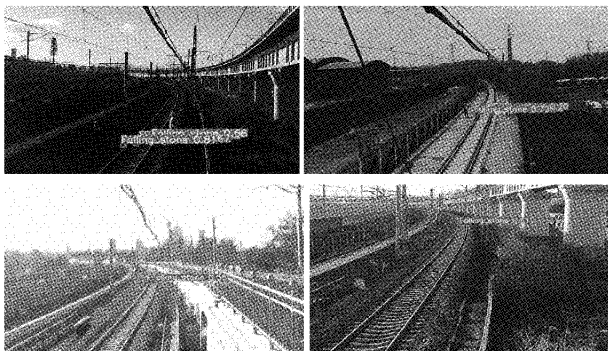


图 10 测试场景示意图

4 结 论

本文针对铁路落石检测任务, 提出了一种混合注意力和双向特征金字塔网络改进 YoloX 的落石检测算法。为解决落石与铁路背景比较相近时检测难题, 增加混合注意力模块, 加强了网络对落石特征的提取能力, 光照较低环境的落石识别能力有所提高。为了提高网络对不同大小落石识别能力, 使用了双向特征金字塔网络, 使网络对不同尺度的特征通道交流能力进行了强化, 保证检测速度的同时, 提高了目标识别准确率。同时在训练时使用了 Mosaic 方法, 增加了小目标数据样本, 进一步提高算法在复杂和远距离场景的检查性能。实验表明, 本文提出检测算法, 对比主流方法, 取得了最好的检查精度, 表明其能够更好的应用在铁路落石检测领域, 具有较好的实用价值。

参考文献

- [1] 王泉东, 杨岳, 罗意平, 等. 铁路侵限异物检测方法综述[J]. 铁道科学与工程学报, 2019, 16(12): 3152-3159.
- [2] 郭碧, 丁春平. 基于帧间差分累积的铁路限界异物检测提取算法[J]. 铁道标准设计, 2019, 63(9): 153-158.
- [3] 张怡, 孙永荣, 刘梓轩, 等. 基于 RTK 定位的图像差分跑道异物检测[J]. 电子测量与仪器学报, 2020, 34(10): 51-56.
- [4] 郭保青, 杨柳旭, 史红梅, 等. 基于快速背景差分的高速铁路异物侵入检测算法[J]. 仪器仪表学报, 2016, 37(6): 1371-1378.
- [5] 伍海萍, 侯涛, 牛宏侠. 改进 ViBe 算法的铁路动态异物入侵的实时监测[J]. 兰州交通大学学报, 2019, 38(5): 55-60, 113.

- [6] 史红梅, 柴华, 王尧, 等. 基于目标识别与跟踪的嵌入式铁路异物侵限检测算法研究[J]. 铁道学报, 2015, 37(7): 58-65.
- [7] 侯涛, 伍海萍, 牛宏侠. 改进 MOG-LRMF 的铁轨动态异物检测[J]. 交通运输系统工程与信息, 2020, 20(2): 91-100.
- [8] 王尧, 余祖俊, 朱力强, 等. 基于高阶全连接条件随机场的高速铁路异物入侵检测方法[J]. 铁道学报, 2019, 41(5): 82-92.
- [9] 于晓英, 苏宏升, 姜泽, 等. 基于 YOLO 的铁路侵限异物检测方法[J]. 兰州交通大学学报, 2020, 39(2): 37-42.
- [10] 王洋, 余祖俊, 朱力强, 等. 基于 CNN 的高速铁路侵限异物特征快速提取算法[J]. 仪器仪表学报, 2017, 38(5): 1267-1275.
- [11] 张剑, 王等准, 莫光健, 等. 基于改进 YOLOv3 的高铁异物入侵检测算法[J]. 计算机技术与发展, 2022, 32(2): 69-74.
- [12] 徐岩, 陶慧青, 虎丽丽. 基于 Faster R-CNN 网络模型的铁路异物侵限检测算法研究[J]. 铁道学报, 2020, 42(5): 91-98.
- [13] TIAN R, SHI H, GUO B, et al. Multi-scale object detection for high-speed railway clearance intrusion[J]. Applied Intelligence, 2021(8): 1-16.
- [14] GE Z, LIU S, WANG F, et al. YOLOX: Exceeding YOLO series in 2021 [J]. ArXiv Preprint, 2021, ArXiv: 2107.08430.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]. European Conference on Computer Vision, 2018: 3-19.
- [16] 崔江波, 侯兴松. 基于注意力机制的 YOLOv4 输电线路故障检测算法[J]. 国外电子测量技术, 2021, 40(7): 24-29.
- [17] TAN M, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection [C]. Conference on Computer Vision and Pattern Recognition(CVPR), 2020.
- [18] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint, 2020, ArXiv: 2004.10934.

作者简介

胡昊, 博士研究生, 主要研究方向为铁路安全智能监测、目标检测等。

E-mail: hhcars11@163.com

史天运(通信作者), 研究员, 博士生导师, 主要研究方向为铁路信息化、智能化。

E-mail: shitianyun@rails.cn

关则彬, 高级工程师, 主要研究方向为铁路视频监控。

E-mail: 76843135@qq.com