



DOI:10.12404/j.issn.1671-1815.2309259

引用格式:鲁冠宏,吕成顺,田隽,等.改进YOLOv5的公路隧道衬砌裂缝检测[J].科学技术与工程,2025,25(7):2997-3006.

Lu Guanhong, Lü Chengshun, Tian Juan, et al. Highway tunnel lining crack detection based on improved YOLOv5[J]. Science Technology and Engineering, 2025, 25(7): 2997-3006.

改进YOLOv5的公路隧道衬砌裂缝检测

鲁冠宏¹, 吕成顺¹, 田隽², 南骁聪³, 马银强², 刘健^{1,4*}, 解全一¹

(1. 山东大学齐鲁交通学院, 济南 250002; 2. 山东高速股份有限公司, 济南 250098;

3. 山东高速工程检测有限公司, 济南 250002; 4. 山东省工业技术研究院, 济南 250101)

摘要 高效、准确的衬砌裂缝检测可以为评估隧道结构安全提供依据。针对传统裂缝检测方法复杂和泛化能力弱的缺点,提出了一种基于深度学习的隧道衬砌裂缝检测网络YOLOv5-CT(YOLOv5 CBAM Transformer)。考虑到裂缝细长的形态,网络引入了Transformer模块来改善裂缝检测效果。Transformer模块较强的长距离依赖捕捉能力使得所提出的检测模型能够充分学习到裂缝区域的上下文信息。此外,该网络在特征融合部分还集成了卷积注意力机制CBAM(convolutional block attention module)。在自采集数据集上的实验结果表明:YOLOv5-CT的AP50和AP(average precision)分别可以达到85.2%和51.3%,相比于基线模型YOLOv5提高了8.9%和12.1%,在精度上优于YOLOX、YOLOv3-MobileNet等其他单阶段目标检测网络。在640×640像素条件下推理速度达到161.3 f/s(frames per second),可以实现隧道衬砌裂缝实时检测。

关键词 公路隧道; 裂缝检测; Transformer; 注意力机制; YOLOv5

中图分类号 U457+.2 TP 391.41; **文献标志码** A

Highway Tunnel Lining Crack Detection Based on Improved YOLOv5

LU Guan-hong¹, LÜ Cheng-shun¹, TIAN Juan², NAN Xiao-cong³,

MA Yin-qiang², LIU Jian^{1,4*}, XIE Quan-yi¹

(1. School of Qilu Transportation, Shandong University, Jinan 250002, China; 2. Shandong Hi-speed Company Limited,

Jinan 250098, China; 3. Shandong Hi-speed Engineering Test Co, Ltd, Jinan 250002, China;

4. Shandong Research Institute of Industrial Technology, Jinan 250101, China)

[Abstract] Efficient and accurate crack detection can provide a basis for assessing the structural safety of tunnels. Aiming at the shortcomings of traditional crack detection methods, which are complex and weak in generalization ability, an improved algorithm YOLOv5-CT(YOLOv5 CBAM Transformer) for tunnel lining crack detection was proposed. Considering the slender morphology of the cracks, the network introduced the Transformer module to improve the crack detection effect. The strong long-range dependency capture ability of the Transformer module enabled the proposed detection model to fully learn the contextual information of the crack region. In addition, the network integrated the convolutional attention mechanism CBAM(convolutional block attention module) in neck. The experiment shows that the YOLOv5-CT can achieve AP50 and AP of 85.2% and 51.3%, respectively, which is an improvement of 8.9% and 12.1% compared to the baseline model YOLOv5. It is better than other one-stage object detection networks in terms of accuracy, and the inference speed reaches 161.3 fps under 640×640 pixel conditions, which meets real-time detection of tunnel lining cracks.

[Keywords] highway tunnel; crack detection; Transformer; attention mechanism; YOLOv5

公路隧道是公路网的“咽喉”,运营管理复杂。随着使用年限增加,部分隧道出现了不同程度的病害,高效的隧道衬砌裂缝检测方法可以为道路养护部门提供有力的技术支撑,提高隧道养护质量并有效延长隧道使用寿命^[1]。目前,衬砌裂缝检测技术包括传统数字图像和基于深度学习的处理方法^[2]。

传统数字图像处理方法往往采用预设特征进行检测,其中阈值分割和边缘检测最为常见。Jiang等^[3]结合Canny算子,采用阈值自适应分割算法从隧道掌子面图像中检测并提取岩体裂缝。张振海等^[4]提出了一种改进的Otsu法首先确定隧道衬砌图像分割阈值,再进行隧道裂缝骨架提取。近年

收稿日期:2023-11-24 修订日期:2024-07-09

基金项目:山东省自然科学基金(ZR2021QE279,ZR2022DKX001);山东省泰山学者工程资助项目(tstp20221153)

第一作者:鲁冠宏(2000—),男,汉族,山东青岛人,硕士研究生。研究方向:隧道衬砌病害智能检测。E-mail:202215404@mail.sdu.edu.cn。

*通信作者:刘健(1975—),男,回族,山东梁山人,博士,教授。研究方向:交通基础设施智能感知。E-mail:liujianshanda@163.com。

投稿网址:www.stae.com.cn

来,小波变换^[5]、模糊数学^[6]等也被应用到裂缝检测领域。虽然传统检测方法在实际应用场景中取得了成功,但依赖人工先验设定特征提取参数,这使得部署工作变得复杂,并且难以平衡检测速度和准确性^[7]。

随着计算机技术的快速发展,基于深度学习的裂缝检测方法逐渐受到研究者的青睐,同时针对裂缝检测任务提出了许多改进策略如滑动窗口^[8]、图像修补策略^[9]、样本优化^[10]等。基于深度学习的裂缝检测方法可分为语义分割^[11]和目标检测,然而,Zhuang等^[12]指出大多数深度学习模型在语义分割方面的成功是以沉重的计算开销为代价的,当不需要裂缝精确的长、宽等参数时,目标检测方法是最佳选择。主流的目标检测算法有以YOLO^[13](you only look once)系列、SSD^[14](single shot detection)为代表的单阶段检测算法和以R-CNN^[15](region-based convolutional neural network)、Faster R-CNN^[16]系列为代表的两阶段检测算法两类。基于候选区域的两阶段目标检测算法虽然能够取得较高的识别精度,但是效率低,不能满足实际隧道衬砌裂缝检测需求。单阶段算法直接对边界框进行回归,以较低的精度损失换来检测速度的大幅提升。YOLO系列作为最具代表性的单阶段目标检测算法,因其精度和速度的平衡而被广泛应用^[13]。Zhang等^[17]将YOLOv3模型和自适应空间特征融合算法(adaptive spatial feature fusion, ASFF)相结合,兼顾对不同尺寸裂缝目标的检测,同时满足网络的实时性检测要求。Ma等^[18]基于YOLOv3模型和中值流追踪算法提出了YOLO-MF(median flow)裂缝识别模型,该模型能够实时检测裂缝并规避相同裂缝重复计数问题。Zhang等^[19]以YOLOv4作为基线网络,探索了不同主干网络和数据增强方法对裂缝识别模型的影响。安学刚等^[20]利用深度可分离卷积作为YOLOv4的卷积方式,同时使用Focal Loss损失函数进行训练,解决了检测模型精度低、体积大的问题。Jeong^[21]使用测试时增强(test time augmentation, TTA)方法提升YOLOv5x裂缝检测网络的预测能力。Wu等^[22]在YOLOv5的基础上进行了轻量化改进,提出了一种名为YOLO-LWNet(lightweight network)的道路病害检测模型,显著降低了模型的计算复杂度,从而实现了道路病害的高效检测。雷斯达等^[23]提出一种结合YOLOv3与MobileNet的深度学习算法,能够对桥梁混凝土结构表面的裂缝目标进行有效地识别与标定。王龙业等^[24]提出一种改进的YOLO-PCB(printed circuit board)缺陷检测算法,通过金字塔网络改进特征融合层提高网络的特征

提取能力和检测精度。周孟然等^[25]在YOLOv5s的基础上加入了注意力机制,能够有效提高网络对缺陷的检测精度。宣以国等^[26]使用可变形卷积替换原YOLOv7中特征提取网络里的卷积,能够实现道路裂缝和坑洞的检测。赵达等^[27]融合K-means++提出了一种改进YOLOv3算法,能够快速准确地识别管道内壁裂缝。然而现阶段的大部分研究并未针对隧道裂缝做较多的优化,裂缝作为一种特殊的检测目标,其长度较长、宽度较窄,且长度远大于宽度,在空间上呈现出细长的形状特征^[28],同时隧道内环境复杂,要求模型能够获取长距离依赖并显著地增强全局特征提取能力,剔除与裂缝无关的冗余特征信息,将更多的可用计算资源用于需要重点关注的目标区域。因此,现提出一种基于YOLOv5^[29]的改进模型YOLOv5-CT来检测隧道衬砌裂缝。YOLOv5-CT将Transformer^[30]模块集成到骨干网络中,改善模型特征提取能力。为了在覆盖范围较大的图像中找到注意区域,本文中采用CBAM(convolutional block attention module)^[31]注意力机制沿通道维度和空间维度依次生成注意力图。与YOLOv5相比,YOLOv5-CT可以更好地检测隧道衬砌裂缝。

1 基于深度学习算法的裂缝识别

1.1 模型架构

本研究提出的裂缝识别模型YOLOv5-CT网络结构如图1所示。该模型修改了原始的YOLOv5网络,使其适用于公路隧道裂缝检测任务。考虑到隧道衬砌裂缝通常都是大跨度延伸的,因此在特征提取网络中引入Transformer模块。首先通过卷积模块得到具有平移不变性的特征图,再利用Transformer获取特征图像素间的全局联系,结合两者之间的优势,增强复杂背景下裂缝目标的特征表达能力并强化特征图全局信息提取能力。在特征提取网络末端使用SPPF(spatial pyramid pooling fast)模块,确保不丢失任何特征信息的情况下合并不同尺度裂缝的上下文信息。

在特征融合部分引入CBAM注意力机制,通过通道注意力模块与空间注意力模块计算得出注意力权重,将注意力权重与输入特征相乘来自适应优化特征。输出端采用3个不同尺度的预测头对裂缝进行边界框回归和分类,使其适应不同尺寸的裂缝。其中,预测头的输出特征图大小为 $80 \times 80 \times 18$, $40 \times 40 \times 18$ 和 $20 \times 20 \times 18$ 。通道维度包含预测边界框信息:预测类别、边界框坐标和置信度值。此外,预测头根据置信度值进行非极大值抑制(non-

maximum suppression, NMS)^[32], 得到最终预测结果。网络选择 CIoU (cross intersection over union) 损失函数进行回归, 交叉熵损失函数进行分类。

1.2 C3Trans 特征提取模块

近年来, Transformer 在自然语言处理领域取得了巨大成功, 考虑到卷积核的大小限制了卷积神经网络全局特征捕捉能力, 研究者试图将 Transformer 移植到计算机视觉领域, 由于隧道衬砌裂缝形态各异、长宽比大, 这就要求网络具有强大的长距离依赖关系建模能力。YOLOv5 主干网络中的 C3 模块

包含 3 个卷积层, 其使用瓶颈结构和 1 × 1 的卷积层实现目标特征提取。在本文中, 把 Transformer 引入 C3 模块中构建 C3Trans 结构, 如图 2 所示。每个 Transformer 模块由多个子层构成, 主要功能由多头注意力和由全连接层构成的多层感知机实现, 子层之间采用残差结构连接。与原始的卷积相比, Transformer 模块的多头注意力能够计算整个特征图所有特征之间的相关性以获取特征图全局信息和充足的上下文信息, 因此在复杂场景下 C3Trans 模块对目标具有更好的特征提取能力。

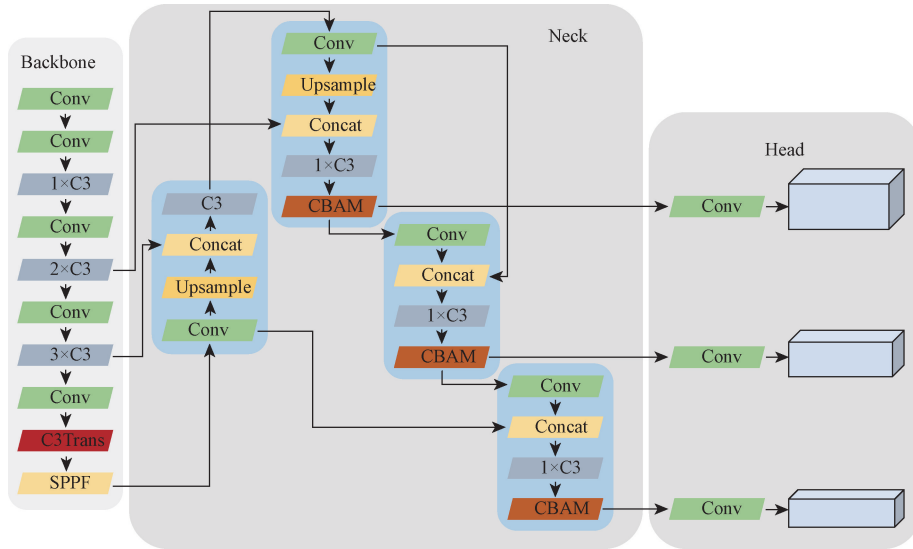
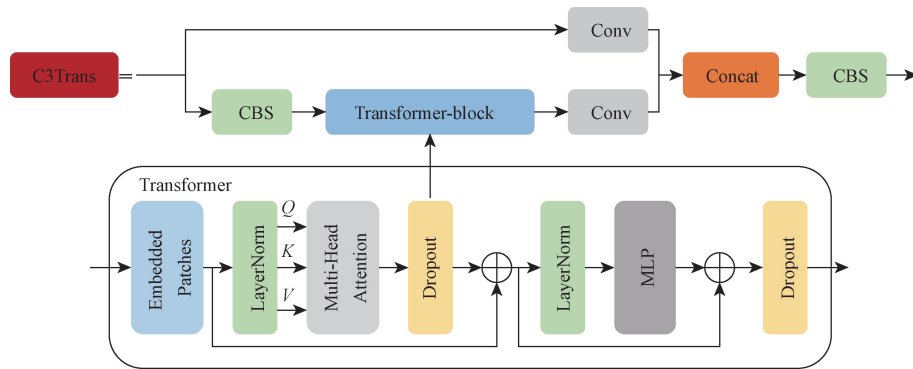


图 1 YOLOv5-CT 网络架构

Fig. 1 The overview of YOLOv5-CT



Q、K、V 分别为 query (查询)、key (键) 和 value (值)

图 2 C3Trans 模块结构

Fig. 2 C3Trans module structure

由于 Transformer 模块使用全连接层的感知机获取全局信息, 这就导致 Transformer 模块计算量与存储远超卷积模块。为了防止 Transformer 模块占用过多计算资源, 仅在特征提取网络末端引入 Transformer 构造 C3Trans 模块, 这样既能保证网络计算量在不增加很多的情况下, 又能获得 Transformer 的优点和更高的检测精度。

1.3 卷积注意力机制模块

注意力可以被解释为一种资源配置方式, 注意力机制依据特征的重要性给特征分配不同的权重, 以提高对重要裂缝特征的关注, 并抑制不必要的特征。裂缝通常只占图像的小部分区域, 大部分内容是背景。目前, 在小型深度学习网络中使用的注意力机制主要有 SE (squeeze-and-excitation networks)、

BAM(bottleneck attention module)和CBAM(convolutional block attention module)。为了剔除与裂缝无关的冗余特征信息,把更多的注意资源用于需要重点关注的目标区域,同时CBAM作为一个可以集成到CNN架构中并且能够以端到端的方式进行训练的轻量级模块,因此我们试图把CBAM注意力机制加入YOLOv5网络特征融合部分,如图3所示。对于一个给定特征图,CBAM沿着通道和空间的两个独立维度依次推断注意力图,协同学习图像中的关键局部细节信息。

将输入特征图 F 与通道权重系数 M_c 矩阵的对应元素相乘,得到能够有效体现特征关键通道信息的特征映射 F' 。在通道加权的基础上,利用串联的空间注意力机制对空间特征信息进行自适应加权,将 F' 作为空间注意力模块的输入,与空间权重系数 M_s 矩阵对应元素相乘,得到包含通道位置信息、空间位置信息的特征图 F'' 。其过程可表达为

$$\begin{cases} F' = M_c \otimes F \\ F'' = M_s \otimes F' \end{cases} \quad (1)$$

式(1)中: \otimes 表示两个矩阵对应元素相乘。

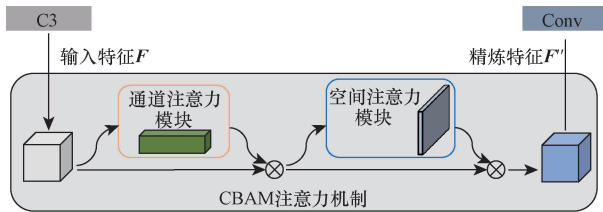


图3 CBAM注意力机制结构

Fig. 3 Convolutional block attention mechanism structure

2 实验设计

2.1 自采集数据集

将FLIR ORX-10G-310S9相机搭载于移动平台上在武汉黄龙山隧道进行公路隧道裂缝图像采集。为了避免训练过程中的过拟合,使用翻转、旋转、裁剪等操作扩充数据集。在训练裂缝识别网络之前,需要手动标注图像数据。经过注释的图像不仅可以训练裂缝识别模型,还能够与裂缝识别模型的预测结果进行比较,方便评估模型的预测性能。共有1744张图像使用流行的注释工具Labelme进行标记,数据集被随机分成两组,一组有1494张图像用于训练,另外一组有250张图像用于验证,测试集与验证集相同。Labelme的输出是JSON格式,在网络训练之前利用第三方库将标注信息转换成YOLO格式和COCO格式。

2.2 对比网络

为了验证改进后的模型YOLOv5-CT的检测能力,本文中选取了单阶段检测模型YOLOX、YOLOv5和YOLOv3-MobileNet作为对比。

YOLOv5发布于2020年,在推理速度和识别精度上有较大的提高。相较于之前的版本,YOLOv5聚焦于优化技术上,如马赛克数据增强和适应性Anchor计算。YOLOv5以CSPDarknet53作为主干网络,提升特征提取速度和精度。另外,采用SPPF、FPN(feature pyramid network)、PAN(path aggregation network)用来增强特征融合。

YOLOX^[33]于2021年由旷视科技提出。该网络在主干网络、解耦层和数据增强等部分进行了改进。YOLOX为无锚框的单阶段目标检测神经网络,不需要设置锚框的参数,极大地减少了网络的参数量,提高网络检测速度和性能。

YOLOv3-MobileNet:使用MMDetection框架中的YOLOv3-MobileNetv2模型,主干网络使用轻量级网络MobileNetV2,MobileNetV2^[34]网络相比MobileNetV1网络,准确率更高,模型更小。

在本文中,所有模型都遵循从头开始训练。环境配置为i7-11700,Nvidia T4 16 G,操作系统为Debian 11.2,开发环境为CUDA 11.1,Python 3.7.4,PyTorch 1.11.0。表1所示为所有模型的训练参数。动量是一个改善训练精度和速度的参数,使用衰减率降低权重值防止过拟合。批量大小表示在每次迭代中用于训练的样本数,训练周期指训练集的全部数据对模型进行一次完整训练。

表1 模型训练参数

Table 1 Model training parameters

模型	输入尺寸	权重衰减	动量	周期	批大小
YOLOv3-MobileNet	640 × 640	0.000 5	0.900	100	4
YOLOv5	640 × 640	0.000 5	0.937	100	4
YOLOX	640 × 640	0.000 5	0.937	100	4
YOLOv5-CT	640 × 640	0.000 5	0.937	100	4

2.3 评价指标

评估指标用于评估模型的性能,这是机器学习任务中非常重要的部分。本文中使用COCO数据集的检测评估指标来衡量不同模型的预测性能。作为最主流的目标检测评估指标之一,COCO评估指标被广泛用于各种计算机视觉任务的评估^[35]。COCO评价指标有两个主要组成部分精确率 P (precision)和召回率 R (recall),其计算公式为

$$P = \frac{TP}{TP + FP} \times 100\% \quad (2)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (3)$$

式中:TP 为被预测为正样本的正样本个数;FP 为被预测为正样本的负样本个数;FN 为被预测为负样本的负样本个数; P 为精确率,表示在预测结果中实际裂缝所占比例; R 为召回率,代表所有裂缝被正确预测的比例。

此外,目标检测任务引入了 IoU(interaction over union),这是评估指标中的一个重要概念,IoU 表示注释框和预测框交集和并集的比率,表达式为

$$IoU = \frac{A_{gt} \cap A_p}{A_{gt} \cup A_p} \quad (4)$$

式(4)中: A_{gt} 为手动标注的注释框边界范围; A_p 为预测边界框范围。

由于所用数据集只包含一种类型的对象,因此选择 3 种 COCO 评估指标来评估这些不同的裂缝检测模型,它们分别是 AP、AP50 以及 AP75。通过设置不同的 IoU 阈值(以 0.05 的步长从 0.5 到 0.95),可以获得不同的预测结果。AP(average precision)是所有 IoU 阈值下精确率的平均值,AP50 和 AP75 是当 IoU 分别为 0.5 和 0.75 时的精确率。网络模型给出的每个预测都伴随着置信水平,按照置信水平对预测结果进行排序,并使用不同的置信区间对预测结果进行分组。精确率 P 和召回率 R 随着置信水平的变化而变化,在 P - R 平面上绘制的 P 和 R 的集合就是该类的 P - R 曲线。 P - R 曲线和坐标轴所包围的区域越大,表明该模型的性能就越好。

3 实验结果与分析

3.1 消融实验

为了分析每个组件对于裂缝识别网络的影响,设置消融实验进行验证,超参数保持一致,消融实验评价指标结果如表 2 所示,其中 YOLOv5 作为基线网络,YOLOv5-T 表示仅在特征提取网络部分引入 Transformer 模块。在实验中,YOLOv5 的 AP50 和 AP 分别为 76.3% 和 39.2%;特征提取网络部分引入 Transformer 模块构建 C3Trans 后,YOLOv5-T 模型总层数从 270 层降低到 220 层,FLOPs(floating point operations per second)从 20.5 G 减少到 20.1 G,AP50 和 AP 分别增加了 0.4% 和 1.9%;在特征融合部分进一步添加 CBAM 注意力机制,相比于 YOLOv5-T,YOLOv5-CT 网络层增加 91 层,FLOPs 增加 0.1 G,AP50 和 AP 分别增加了 8.5% 和 10.2%,这表明 Transformer 模块和 CBAM 注意力机制都能够有效提升裂缝识别模型的检测精度。为验证 YOLOv5-CT 的泛化性,对原始测试集分别添加椒盐噪声和高斯噪声进行预测,结果如表 3 和表 4 所示。

表 2 消融实验结果

Table 2 Result of ablation experiments

模型	Transformer	CBAM	AP50/ %	AP/ %	层数	FLOPs/ G
YOLOv5	—	—	76.3	39.2	270	20.5
YOLOv5-T	—	✓	76.7	41.1	220	20.1
YOLOv5-CT	✓	✓	85.2	51.3	311	20.2

表 3 原始测试集添加椒盐噪声后的模型预测结果

Table 3 Prediction results after adding salt and pepper noise to the original test dataset

模型	P /%	R /%	AP50/%	AP/%
YOLOv5	69.8	47.4	61.0	34.0
YOLOv5-T	72.5	48.6	65.4	37.2
YOLOv5-CT	79.3	52.9	70.5	46.2

表 4 原始测试集添加高斯噪声后的模型预测结果

Table 4 Prediction results after adding Gaussian noise to the original test dataset

模型	P /%	R /%	AP50/%	AP/%
YOLOv5	75.6	57.7	68.8	36.9
YOLOv5-T	76.0	59.1	72.6	39.7
YOLOv5-CT	81.2	64.5	77.0	50.2

根据预测结果可以看出,无论是添加椒盐噪声还是高斯噪声,YOLOv5-CT 的各项评估指标均取得了最好的结果,并且添加高斯噪声的预测结果优于椒盐噪声。从以上分析可以得出,相比于 YOLOv5 和 YOLOv5-T 网络,YOLOv5-CT 具有更强的泛化能力和裂缝预测性能。

图 4 表示 YOLOv5 和 YOLOv5-CT 模型在特征提取过程中的部分通道特征,可以很明显地看出 YOLOv5-CT 模型提取到的裂缝区域细节信息更多,特征捕捉能力更强。此外,图 5 所示为上述 3 个模型在训练过程中的训练损失变化曲线。其中,横坐标为训练周期,纵坐标为损失值。随着训练周期的增加,3 个模型的训练损失逐渐下降并趋向于收敛状态。一般来说,损失值越低表示模型检测性能越好,在添加 Transformer 模块重构特征提取网络后,训练损失曲线下降速度加快,训练结束时损失值降低;在此基础上使用 CBAM 注意力机制增强模型特征融合能力后,损失曲线以最快的速度下降,训练结束时具有最小的损失值。从曲线波动性方面看,训练前期 3 个模型的训练曲线较为平滑,YOLOv5 模型在第 15 轮最先出现波动,YOLOv5-T 和 YOLOv5-CT 在第 22 轮开始出现波动,且曲线波动幅度减小;第 22 轮至训练结束,YOLOv5-T 和 YOLOv5-CT 损失曲线几乎平行。以上现象说明 YOLOv5-CT 在网络优化上更好,可以加快模型的收敛并提高检测性能。

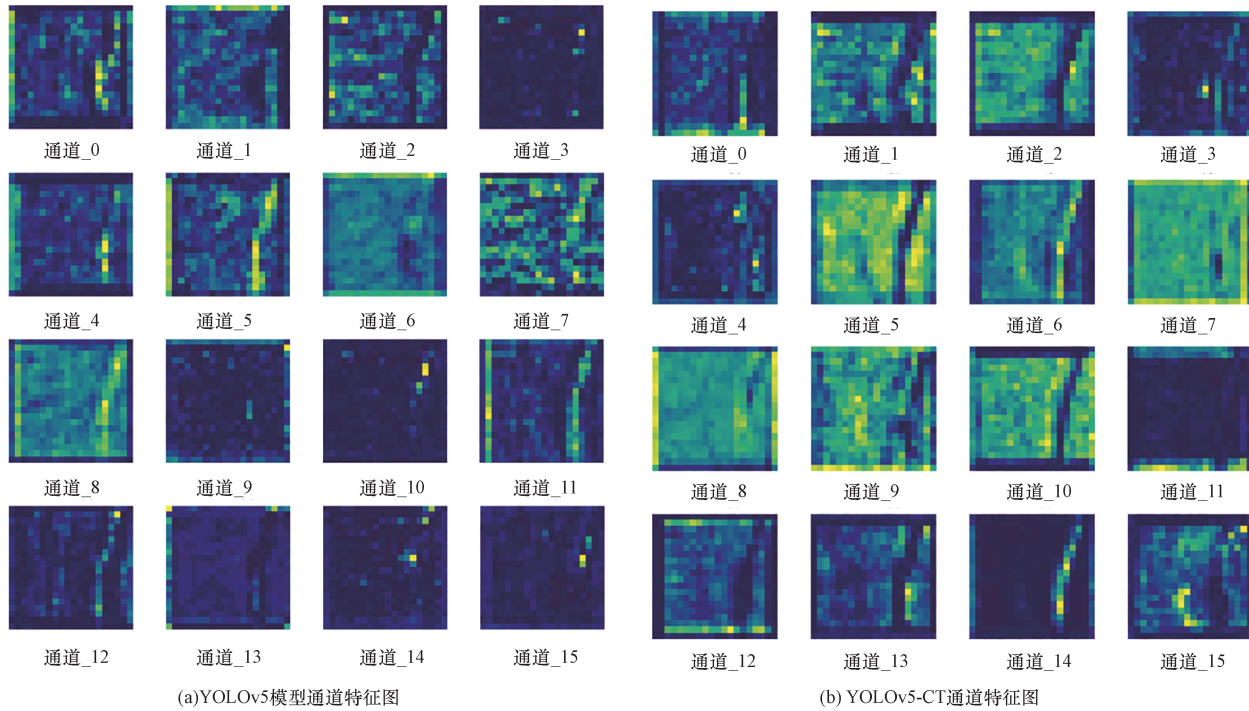


图4 通道特征图对比

Fig. 4 Comparison of channel features

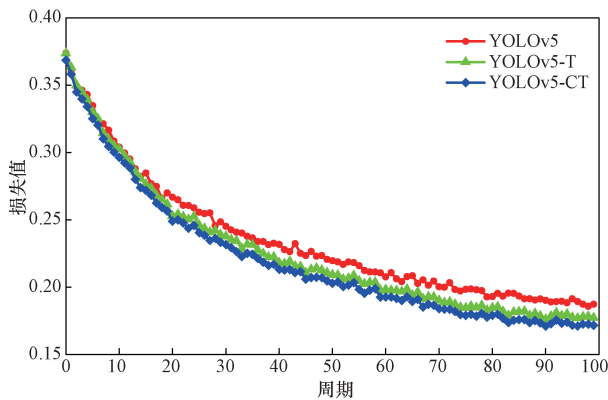


图5 3种模型在裂缝数据集上的损失曲线对比

Fig. 5 Loss curves of the three models on the crack dataset

3.2 裂缝识别网络分析

在本文中,为验证 YOLOv5-CT 的有效性和优越性,分别使用 YOLOv5、YOLOv3-MobileNet 和 YOLOX 进行了训练和测试,不同评估指标的实验结果如表 5 所示。很容易发现,对于任意一个评估指标,YOLOX 在公路隧道衬砌裂缝数据集上的表现最差,当 IOU 设置为 0.75 时,精确率仅有 12%,至于召回率也仅有 44.1%,这些指标明显低于 YOLOv5-CT 的相应值。从表 3 中可以看出,YOLOv5-CT 的 AP 分别比 YOLOv3-MobileNet、YOLOv5 和 YOLOX 高 17.2%、12.1% 和 28.4%;AP50 分别比 YOLOv3-MobileNet、YOLOv5 和 YOLOX 高 10.7%、8.9% 和 25.3%,AP75 分别比 YOLOv3-MobileNet、YOLOv5 和

表 5 不同模型在裂缝数据集上的训练指标

Table 5 Training metrics for different models on the crack dataset

模型	AP/%	AP50/%	AP75/%	R/%	帧率/fps
YOLOv3-MobileNet	34.1	74.5	28.0	45.0	84.3
YOLOX	22.9	59.9	12.0	44.1	83.8
YOLOv5	39.2	76.3	35.3	70.1	204.1
YOLOv5-CT	51.3	85.2	53.0	77.5	161.3

YOLOX 高 25%、17.7% 和 41%,模型的精确率较低意味着在更多的检测过程中产生更多的误判。相比于 YOLOv5, YOLOv5-CT 在 AP75 评估指标上的提高最为显著,表明该网络能够有效改善高 IoU 条件下裂缝预测性能,并且预测框更接近注释框。实验中使用的三个模型 YOLOX、YOLOv5 和 YOLOv5-CT 参数量分别为 8.94 M、8.84 M 和 8.75 M,三个模型在参数量上处于同一数量级,但在检测性能上却有很大差别,YOLOv3 模型由于主干网络为轻量级网络 MobileNet,在参数量仅有 3.67 M 的条件下比 YOLOX 表现好。此外,YOLOv5-CT 的检测速度达到了 161.3 f/s (frames per second),略低于 YOLOv5,但也满足实际检测工程的需求。

YOLOv5-CT 在主干网络部分引入 Transformer 模块,相比于 YOLOv5、YOLOX 和 YOLOv3-MobileNet,增强了网络特征提取能力,其中 YOLOv3-MobileNet 使用 MobileNet、YOLOv5 和 YOLOX 以

CSPDarknet 作为主干网络。同时特征融合部分的 CBAM 注意力机制使得 YOLOv5-CT 将更多的计算资源用于目标对象。

从而,本文所提出的模型在检测精度和推理速度上具有良好的平衡性能,为未来的隧道衬砌裂缝自动检测提供了可能的解决方案。

召回率是一个评估网络性能的重要指标,它能够反应模型对正样本的识别能力。YOLOv5-CT 模型的召回率分别比 YOLOv3-MobileNet、YOLOv5 和 YOLOX 高 32.5%、7.4% 和 33.4%,为了对比具有不同召回率模型的检测效果,本文随机选取 4 幅裂缝图像进行测试,预测结果如图 6 所示,其中预测框置信度设定为 0.3,图像输入尺寸为 640×640 像素。在第 1 行中,由于裂缝不明显,YOLOv5 和 YOLOX 发生漏检情况,YOLOv5-CT 和 YOLOv3-MobileNet 准确预测到了衬砌裂缝;第 2 行中,YOLOv5-CT 和 YOLOv5 都得到了预测结果,但是 YOLOv5-CT 的预测结果更接近于注释框,YOLOX 和 YOLOv3-MobileNet 未识别到裂缝;第 3 行中,仅有 YOLOX 未得到预测结果;第 4 行图像亮度值较高,4 个模型都检测到了裂缝,YOLOv5 模型的预测框和真实标注重叠度最小。根据图 6 检测结果可以看出,较高的召回率可以得到更好的检测结

果,同时也表明所提出的网络模型具有强大的裂缝捕捉能力。

一般来说,精确率和召回率都在 0~1 之间,随着召回率的增加精确率逐渐下降,每条 $P-R$ 曲线下的面积代表某一类目标在特定阈值下的 AP 值, $P-R$ 曲线与坐标轴围成的封闭图形面积越大意味着网络模型识别效果越好。在文中,IoU 阈值被设定为 0.5,阈值高于 0.5 的样本为正例,小于 0.5 的样本为负例,进而计算相应的精确率和召回率,如图 7 所示。YOLOv5-CT 的 $P-R$ 曲线完全包住了 YOLOv5、YOLOX 和 YOLOv3-MobileNet,表明 YOLOv5-CT 模型优于 YOLOv5、YOLOX 和 YOLOv3-MobileNet,预测能力 $YOLOv5-CT > YOLOv5 > YOLOv3-MobileNet > YOLOX$ 。

3.3 光照条件对模型预测性能的影响

由于隧道内光线条件差且受附属设施干扰,隧道衬砌图像中存在大量噪声。在强光条件下,裂缝特征明显有利于裂缝识别;在弱光昏暗条件下,裂缝与背景差异较小,增加了裂缝识别难度。为了研究光照对于裂缝识别模型性能的影响,将测试集图像按照 0(原始测试集)、-5%、-10% 和 -15% 降低图像亮度值来模拟不同的光照强度,预测结果如图 8 所示。

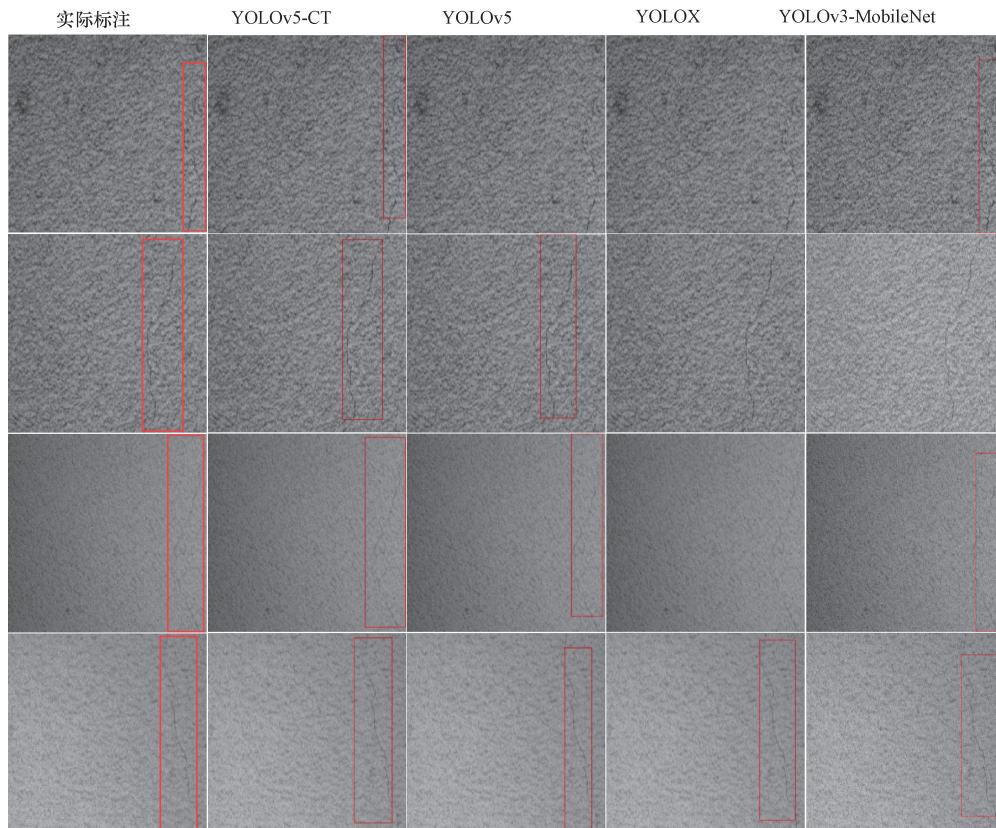


图 6 不同召回率模型的预测结果

Fig. 6 Prediction results on models with different recall values

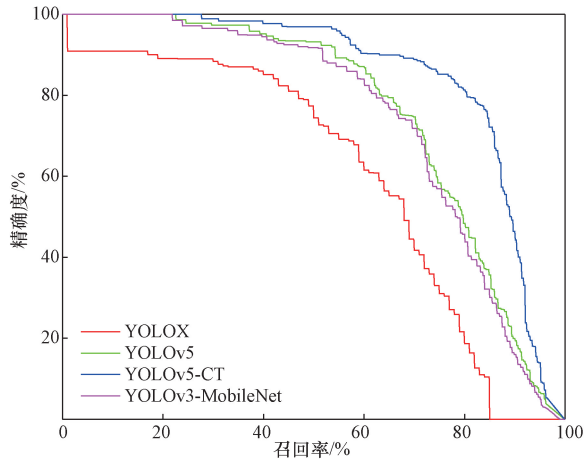


图7 不同模型的P-R曲线

Fig. 7 P-R curves of different models

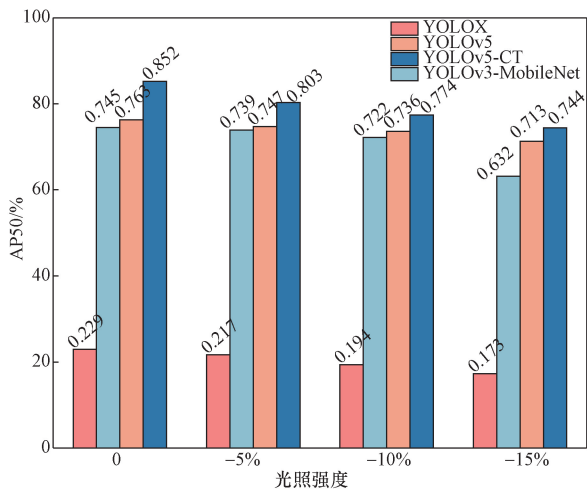


图8 不同光照强度下4种光照强度降低裂缝识别模型AP50值

Fig. 8 AP50 values of four crack detection models under different light conditions

可以看出,随着光照强度的变弱,4种不同识别模型的AP50逐渐下降,其中以YOLOv3-MobileNet从原来的74.5%下降到63.2%最为显著,YOLOv5-CT也从85.2%下降到了74.4%,光照不足时模型捕捉裂缝特征的能力受噪声干扰严重,从而造成识别结果的降低。在同一光照强度下,YOLOX检测性能最差,而YOLOv5-CT总是可以达到最好的识别效果。光照强度降低5%时,YOLOv5-CT的AP50值达到了80.3%,YOLOv5、YOLOv3-MobileNet和YOLOX以5.6%、6.4%和58.6%落后于YOLOv5-CT。进一步降低图像光照强度,YOLOv5-CT仍可以保持良好的识别效果,即使光照强度降低了15%,YOLOv5-CT的AP50值仍旧达到了74.4%,比YOLOv5、YOLOv3-MobileNet和YOLOX分别高3.1%、11.2%和57.1%。所提出的裂缝识别模型

YOLOv5-CT对不同光照强度下的裂缝图像均取得了良好的检测效果,这表明该网络可以适应不同光照环境下的隧道衬砌裂缝检测,具有较强的鲁棒性。

4 结论

在这项研究中提出了一个隧道衬砌裂缝识别网络YOLOv5-CT,旨在提升隧道裂缝检测效率和准确性。该模型基于单阶段目标检测网络YOLOv5,在构建特征提取网络时加入Transformer模块获取特征图的全局信息和上下文信息,有效增强长细裂缝目标的特征表达能力;同时采用CBAM注意力机制提高对重要裂缝特征的关注,并抑制不必要的特征,加快模型推导速度。通过构建用于模型训练和测试的隧道衬砌裂缝图像数据集,并基于该数据集完成YOLOv5-CT与其他主流方法的性能对比研究和多场景测试,可以得出以下结论。

(1)Transformer模块和CBAM注意力机制可以有效增强网络的裂缝检测能力。相较于YOLOv5,YOLOv5-CT的模型深度从270层增加到311层,FLOPs由20.5G变化到20.2G,AP50和AP分别提高了8.9%和12.1%。

(2)本文所提方法在测试集上取得了AP50为85.2%、AP75为53.0%、AP达到了51.3%、召回率为77.5%的效果。相比于基线模型YOLOv5,以及其他主流使用的YOLOX和YOLOv3-MobileNet均有大幅度提升,并且在检测精度和速度上具有良好的平衡性能。

(3)在降低光照强度时,各模型的精度均会出现不同程度的下降。光照越弱,模型精度下降越严重。实验结果表明YOLOv5-CT可以适应不同光照环境下的隧道衬砌裂缝检测,具有较强的鲁棒性。

参考文献

- [1] 周中, 闫龙宾, 张俊杰, 等. 基于深度学习的公路隧道表观病害智能识别研究现状与展望[J]. 土木工程学报, 2022, 55(S2): 38-48.
Zhou Zhong, Yan Longbin, Zhang Junjie, et al. Review and prospect of intelligent identification of apparent diseases in highway tunnels based on deep learning[J]. China Civil Engineering Journal, 2022, 55(S2): 38-48.
- [2] 翟军治, 孙朝云, 裴莉莉, 等. 多尺度特征增强的路面裂缝检测方法[J]. 交通运输工程学报, 2023, 23(1): 291-308.
Zhai Junzhi, Sun Zhaoyun, Pei Lili, et al. Pavement crack detection method based on multi-scale feature enhancement[J]. Journal of Traffic and Transportation Engineering, 2023, 23(1): 291-308.
- [3] Jiang F, Wang G, He P, et al. Application of canny operator threshold adaptive segmentation algorithm combined with digital image processing in tunnel face crevice extraction[J]. The Journal of Supercomputing, 2022, 78(9): 11601-11620.

- [4] 张振海, 贾争满, 季坤. 基于改进的 Otsu 法的地铁隧道裂缝识别方法研究[J]. 重庆交通大学学报(自然科学版), 2022, 41(1): 84-90.
Zhang Zhenhai, Jia Zhengman, Ji Kun. Crack identification method of subway tunnel based on improved Otsu method[J]. Journal of Chongqing Jiaotong University(Natural Science), 2022, 41(1): 84-90.
- [5] 苑霄哲. 基于数字图像处理的沥青路面裂缝图像检测算法研究[D]. 南京: 东南大学, 2022.
Yuan Xiaozhe. Research on image detection algorithm of asphalt pavement crack based on digital image processing [D]. Nanjing: Southeast University, 2022.
- [6] 黄梦婷, 高斐, 栾凯伦, 等. 涪陵页岩储层裂缝识别方法研究[J]. 矿产勘查, 2024, 15(5): 853-860.
Huang Mengting, Gao Fei, Luan Kailun, et al. Research on fracture identification method of Fuling shale reservoir [J]. Mineral Exploration, 2024, 15(5): 853-860.
- [7] Xiang X, Wang Z, Qiao Y. An improved YOLOv5 crack detection method combined with Transformer [J]. IEEE Sensors Journal, 2022, 22(14): 14328-14335.
- [8] Cha Y J, Choi W, Büyükköztürk O. Deep learning-based crack damage detection using convolutional neural networks: deep learning-based crack damage detection using CNNs[J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(5): 361-378.
- [9] Liu Y, Yeoh J K W. Robust pixel-wise concrete crack segmentation and properties retrieval using image patches [J]. Automation in Construction, 2021, 123: 103535.
- [10] 蓝茜茜, 张逸伦, 康志宏. 基于样本优化的神经网络方法在储层裂缝识别中的应用 [J]. 科学技术与工程, 2020, 20(21): 8530-8536.
Lan Xixi, Zhang Yilun, Kang Zhihong. Application of neural network based on sample optimization in reservoir fracture identification[J]. Science Technology and Engineering, 2020, 20(21): 8530-8536.
- [11] 张世瑶, 贺玉彬, 周新志. 基于深度学习的道路表面裂缝检测方法[J]. 科学技术与工程, 2021, 21(15): 6380-6385.
Zhang Shiyao, He Yubin, Zhou Xinzhi. Road surface crack detection method based on deep learning[J]. Science Technology and Engineering, 2021, 21(15): 6380-6385.
- [12] Zhuang J, Yang J, Gu L, et al. Shelf net for fast semantic segmentation [C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). New York: IEEE, 2019: 847-856.
- [13] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2016: 779-788.
- [14] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector [C]//Computer Vision-ECCV. Berlin: Springer International Publishing, 2016: 21-37.
- [15] 郭庆梅, 刘宁波, 王中训, 等. 基于深度学习的目标检测算法综述[J]. 探测与控制学报, 2023, 45(6): 10-20, 26.
Guo Qingmei, Liu Ningbo, Wang Zhongxun, et al. A review of object detection algorithms based on deep learning [J]. Journal of Detection and Control, 2023, 45(6): 10-20, 26.
- [16] Ren S, He K, 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.
- [17] Zhang R, Shi Y, Yu X. Pavement crack detection based on deep learning [C]//2021 33rd Chinese Control and Decision Conference (CCDC). Kunming: IEEE, 2021: 7367-7372.
- [18] Ma D, Fang H, Wang N. Automatic detection and counting system for pavement cracks based on PCGAN and YOLO-MF [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 22166-22178.
- [19] Zhang X, Xia X, Li N, et al. Exploring the tricks for road damage detection with a one-stage detector [C]//2020 IEEE International Conference on Big Data (Big Data). New York: IEEE, 2020: 5616-5621.
- [20] 安学刚, 党建武, 王阳萍, 等. 基于改进 YOLOv4 的无人机影像路面病害检测方法 [J]. 无线电工程, 2023, 53(6): 1285-1294.
An Xuegang, Dang Jianwu, Wang Yangping, et al. UAV image pavement distress detection method based on improved YOLOv4 [J]. Radio Engineering, 2023, 53(6): 1285-1294.
- [21] Jeong D. Road damage detection using YOLO with smartphone images [C]//2020 IEEE International Conference on Big Data (Big Data). New York: IEEE, 2020: 5559-5562.
- [22] Wu C G, Ye M, Zhang J L, et al. YOLO-LWNet: A lightweight road damage object detection network for mobile terminal devices [J]. Sensors, 2023, 23(6): 3268-3268.
- [23] 雷斯达, 林杰, 黄思璐, 等. 基于改进 YOLOv3 网络对混凝土表面裂缝目标识别研究 [J]. 公路, 2024, 69(1): 270-275.
Resda, Lin Jie, Huang Silu, et al. Research on target recognition of concrete surface crack based on improved YOLOv3 network [J]. Highway, 2019, 69(1): 270-275.
- [24] 王龙业, 黄黎, 曾晓莉. 基于 YOLO-PCB 的印刷电路板裸板缺陷检测 [J]. 科学技术与工程, 2024, 24(15): 6338-6345.
Wang Longye, Huang Jun, Zeng Xiaoli. Bare board defect detection of PCB based on YOLO-PCB [J]. Science Technology and Engineering, 2024, 24(15): 6338-6345.
- [25] 周孟然, 王昊男, 高立鹏, 等. 基于 YOLOv5s-FCS 的钢材表面缺陷检测 [J]. 科学技术与工程, 2024, 24(14): 5901-5910.
Zhou Mengran, Wang Haonan, Gao Lipeng, et al. YOLOv5s-FCS based steel surface defect detection study [J]. Science Technology and Engineering, 2024, 24(14): 5901-5910.
- [26] 宣以国, 余成波. 基于改进 YOLOv7 的道路裂缝和坑洞检测算法 [J]. 科学技术与工程, 2024, 24(17): 7205-7213.
Xuan Yiguo, Yu Chengbo. Improved YOLOv7 road crack and pothole detection algorithm [J]. Science Technology and Engineering, 2024, 24(17): 7205-7213.
- [27] 赵达, 王亚慧, 陈林林. 一种管道蛇形机器人的裂缝视频检测系统 [J]. 科学技术与工程, 2023, 23(6): 2492-2498.
Zhao Da, Wang Yahui, Chen Linlin. A crack video detection system of pipeline snake shaped robot [J]. Science Technology and Engineering, 2023, 23(6): 2492-2498.
- [28] 谢雄耀, 王皓正, 周彪, 等. 复杂环境公路隧道裂缝快速识别与分割算法研究 [J]. 地下空间与工程学报, 2022, 18(3): 1025-1033, 1043.
Xie Xiongyao, Wang Haozheng, Zhou Biao, et al. Research on fast identification and segmentation algorithms for cracks of highway tunnel

- nels in complex environment[J]. Chinese Journal of Underground Space and Engineering, 2022, 18(3): 1025-1033, 1043.
- [29] Jocher G, Liu C Y, Hogan A, et al. Ultralytics/YOLOv5: Initial Release[EB/OL]. 2022-11-22. <https://github.com/ultralytics/yolov5/releases>.
- [30] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]//31st Annual Conference on Neural Information Processing Systems (NIPS). California: Curran Associates, 2017: 6000-6010.
- [31] Woo S, Park J, Lee J, et al. CBAM: Convolutional block attention module [C]//Computer Vision- ECCV. Munich: Springer-Verlag, 2018: 3-19.
- [32] Neubeck A, Van Gool L. Efficient non-maximum suppression [C]//18th International Conference on Pattern Recognition (ICPR'06). Hong Kong: IEEE, 2006: 850-855.
- [33] Ge Z, Liu S, Wang F, et al. YOLOX: Exceeding YOLO series in 2021[M]. arXiv: 2107. 08430.
- [34] Sandler M, Howard A, Zhu M, et al. MobileNetV2: Inverted residuals and linear bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 4510-4520.
- [35] Lin T, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context [C]//Computer Vision-ECCV. Zurich: Springer Cham, 2014: 740-755.