



DOI:10.12404/j.issn.1671-1815.2407398

引用格式:张颖,王纪旭,曹迎康,等.基于改进YOLOv9-c的路面混合病害算法[J].科学技术与工程,2025,25(18):7793-7802.

Zhang Ying, Wang Jixu, Cao Yingkang, et al. Pavement mixed disease algorithm based on improved YOLOv9-c[J]. Science Technology and Engineering, 2025, 25(18): 7793-7802.

基于改进YOLOv9-c的路面混合病害算法

张颖^{1,2}, 王纪旭¹, 曹迎康¹, 李罡¹, 方有亮^{2*}

(1. 河北大学建筑工程学院, 保定 071002; 2. 河北大学零碳能源建筑与计量技术教育部工程研究中心, 保定 071002)

摘要 针对坑槽和裂缝两种路面病害检测实时性差、准确率低、易误检漏检等问题,提出了一种改进YOLOv9的路面混合病害算法,实现路面裂缝的自动化检测和识别。首先,在骨干网络中引入AKConv(alterable kernel convolution)替换RepNC-SPELAN4中的卷积模块,提高网络对不同病害的特征提取能力,有效解决路面病害与背景环境特征难以区分的问题;其次,在检测头中引入了SimAM注意力机制(selective image attention mechanism)和DySample上采样模块,提高网络聚焦特性并增强提取关键特征信息的能力;最后,采用inner-IoU函数优化模型的权重参数,提升对混合样本的学习能力。实验结果表明,改进后的模型与YOLOv9-c相比较,性能有了显著提升,平均精度提升40.17%、召回率提高了15.99%、mAP模型精度提高了20.95%,该优化算法能够更加精准高效的对路面混合病害进行检测,提高了路面病害检测的准确率和泛用性。

关键词 YOLOv9-c; 路面混合病害; 注意力; 特征提取; 损失函数

中图法分类号 U416.06 TP391.4; 文献标志码 A

Pavement Mixed Disease Algorithm Based on Improved YOLOv9-c

ZHANG Ying^{1,2}, WANG Ji-xu¹, CAO Ying-kang¹, LI Gang¹, FANG You-liang^{2*}

(1. Department of Civil Engineering and Architecture, Hebei University, Baoding 071002, China;

2. Engineering Research Center of Zero-Carbon Energy Buildings and Measurement Techniques, Ministry of Education, Hebei University, Baoding 071002, China)

[Abstract] Aiming at the problems of poor real-time detection, low accuracy, and false detection and omission of pavement disease detection including hole and crack, an improved algorithm based on YOLOv9 was proposed to resolve the problem. Firstly, AKConv (alterable kernel convolution) was introduced into the backbone network to replace the convolution module in RepNC-SPELAN4, which improves the feature extraction ability of the network for different diseases and effectively solve the problem that road disease is difficult to distinguish from background environment features. Secondly, selective image attention mechanism (SimAM) and DySample sampling modules were introduced to focus on the key information in the detection head, and the capability to extract information features was enhanced more efficiently. Finally, the inner-IoU function was used to optimize the weight parameters of the model to improve the learning ability of mixed samples. The experimental comparison between YOLOv9-c and our model showed that the accuracy, recall rate and MAP of the improved model are increased by 40.17%, 15.99% and 20.95% respectively. The performance has been significantly improved, and the detection effect is more accurately and efficiently, and the accuracy and generalization ability of pavement disease detection algorithm are improved.

[Keywords] YOLOv9-c; pavement mixed disease; attention mechanism; feature extraction; loss function

交通建设是国家建设的重要组成部分,而高速公路是国民出行的主要方式,因此对于路面的养护有重要的意义。公路养护的对象包括路面裂缝、车辙、坑槽等路面病害,其中路面裂缝会在雨雪等自

然因素与车辆负荷的综合作用下持续破坏路面结构,导致公路使用寿命的减少并带来行车安全隐患,是公路路面损坏的主要形式与养护的重点对象^[1]。针对道路缺陷的检测方法分为有损检测和

收稿日期:2024-10-08 修订日期:2025-04-03

基金项目:教育部春晖项目合作科研项目(HZKY20220256);国家自然科学基金青年科学基金(52108203);河北大学实验室开放项目基金(sy202236)

第一作者:张颖(1981—),女,汉族,河北石家庄人,博士,副教授。研究方向:深度学习、工程结构健康监测方法。E-mail:hbuzy@163.com。

*通信作者:方有亮(1967—),男,汉族,河北张家口人,博士后,教授,博士研究生导师。研究方向:结构健康监测与检测。E-mail:fang_youliang@163.com。

无损检测两类。有损检测主要采用钻芯取样法,以获取路面结构层的质量参数^[2]。在实际养护中,一般都是每隔一段时间,通过巡检车或者无人机采集路面的照片进行人工绘图的方式进行裂缝提取,识别效率较低,而且容易受到人为主观识别误差和工程经验的影响^[3],这样就会耗费较大的资源,因此,开发一种高效、精准、便捷的道路病害检测方法有至关重要的作用,随着深度学习在目标检测领域的应用,产生了许多基于深度学习的目标检测方法,逐渐实现了高效且精准的病害检测。基于深度学习的目标检测算法主要分为三类:单阶段算法、双阶段算法和轻量化网络。单阶段方法直接通过端到端的方式从输入图像中输出目标框和分类结果,可简化流程并减少运算量,常用模型有 SSD (single shot multibox detector) 算法与 YOLO (you only look once) 系列算法等^[4]。双阶目标检测算法分为目标定位和目标分类两步骤,先生成候选区域,再对候选区域进行分类识别。典型的双阶段检测算法有快速区域卷积神经网络 (faster region-CNN, faster R-CNN)、基于区域的全卷积网络 (region-based fully convolutional networks, R-FCN)、Mask R-CNN 等^[5]。轻量化网络主要应用于无人机算法,精度和准确率要低于单阶算法,因此,单阶算更加适用于路面病害检测。

目前关于路面病害单阶段检测的研究关注点大多聚焦于 YOLO 系列,如 Pan 等^[6]结合了 YOLOv3-tiny 检测器和传统的 KLT 跟踪算法提出了更为先进的 YOLOv3-tiny-KLT 进行实时结构振动测量。Zhang 等^[7]改进了 YOLOv4 算法,提出了 CR-YOLO,并基于 CR-YOLO、PSPNet、服务器和边缘设备,构建了人机交互桥裂纹检测分割系统。廖祥灿等^[8]基于 YOLOv5 目标检测模型,提出引用 C3-B 注意力机制模块,选取最新的 SIOU-Loss 作为边框回归的损失函数,提高训练速度和推理精度。杜磊等^[9]针对目前沥青路面裂缝目标检测技术在面对复杂路面情况(强光、积水、杂物等干扰因素)时识别精度较低的问题,通过构建沥青路面裂缝数据集,提出了一种基于改进 YOLOv5s 的沥青路面裂缝检测算法。贾晓芬等^[10]融合普通卷积、深度可分离卷积和 ECA 注意力机制设计轻量化卷积模块 ECA-Conv,再引入跳跃链接构建特征综合提取单元 E-C3 并设计特征融合模块 ECACSP,利用多组 ECACConv 和 ECACSP 模块组建细颈部特征融合模块 E-Neck,出轻量化检测模型 E-YOLOv5s。夏翔等^[11]提出了改进后的 YOLOv7 模型,mAP 指标在 VisDrone2019 数据集上提高到了 50.1%,在自制视频监控数据

集上高于现有方法 1.6 个百分点,有效提高了小目标检测的能力。王海群等^[12]在 YOLOv8 主干网络引入 CN X2f 模块的同时引入 RepConv 和 DBB 重参数化模块增强多尺度特征融合能力,改进头部采用共享参数结构,并引入 RBB 重参数模块和 SPPF_Avg 模块,解决路面病害特征丢失问题,丰富多尺度特征表达。

尽管上述方法对路面裂缝的检测有较好的效果,但实际交通场景复杂多变且存在多种病害,同时实际检测所采集的含有裂纹和坑槽的图片中往往会出现混淆一些类似裂缝或坑槽的干扰项,由于裂缝和坑槽的面积在实际图片中的占比较小,卷积过程中很容易造成信息丢失,因此寻找一种特征提取能力强的网络显得尤为重要。针对坑槽和裂缝两种主要的路面混合病害识别过程的复杂性,现采用 Yolov9-c 作为基本网络模型,进行优化改进该模型。首先,在 RepNCSPELAN4 模块中融入 AK-Conv^[13] 卷积模块,建立新的 RepNCSPELAN4-AK-Conv 模块加强网络的特征提取能力;其次引入 Si-mAM^[14] 注意力机制和 DySample^[15] 上采样算子让网络能够更好地对关键信息进行定位和采集,最后,采用 inner-IoU^[16] 函数替代 CIoU 函数改善检测边框回归效果。

1 Yolov9 算法

YOLOv9^[17]是 YOLO 系列的算法之一,主要应用对于图像进行目标检测和分类,输出每个目标框的位置和类别概率。首次引入了可编程梯度信息 PGI (programmable gradient information) 和广义高效层聚合网络 GELAN (generalized efficient layer aggregation network) 等开创性技术,PGI 是建立在辅助分支上可以自由选择适合目标任务损失函数的可逆架构,GELAN 则是结合了 CSPNet 网络和 ELAN 网络,通过整合 PGI 和多功能 GELAN 架构,YOLOv9 不仅增强了模型的学习能力,确保在整个检测过程中保留关键信息,从而实现卓越的性能。v9 和 v7、v5 构架相同两部分组成: Backbone (骨干网络) 和 Head (检测头)。Backbone 部分由 Conv, RepNCSPELAN4, CBFuse, CBLIner 等模块组成,主要进行特征提取。Head 部分包含 Conv、RepNCSPELAN4、SPPELAN、Upsample 模块,主要功能是输出之后,与真实数据标注相比较,计算出损失函数然后根据需要对数据格式进行重塑,同时对原始格点坐标做相应的激活。YOLOv9 将模型分为 t、s、m、c、e 5 种,可供下载的是 c 和 e 两个模型,上述两个模型主要差异是 params 和 FLOPs 这两项评价指标,c 模型所需

的运算要小于 e 模型,两者的精度和召回率仅相差不超过 3%,因此,本文研究采用 YOLOv9-c 模型作为基础模型。

2 改进 YOLOv9 网络结构

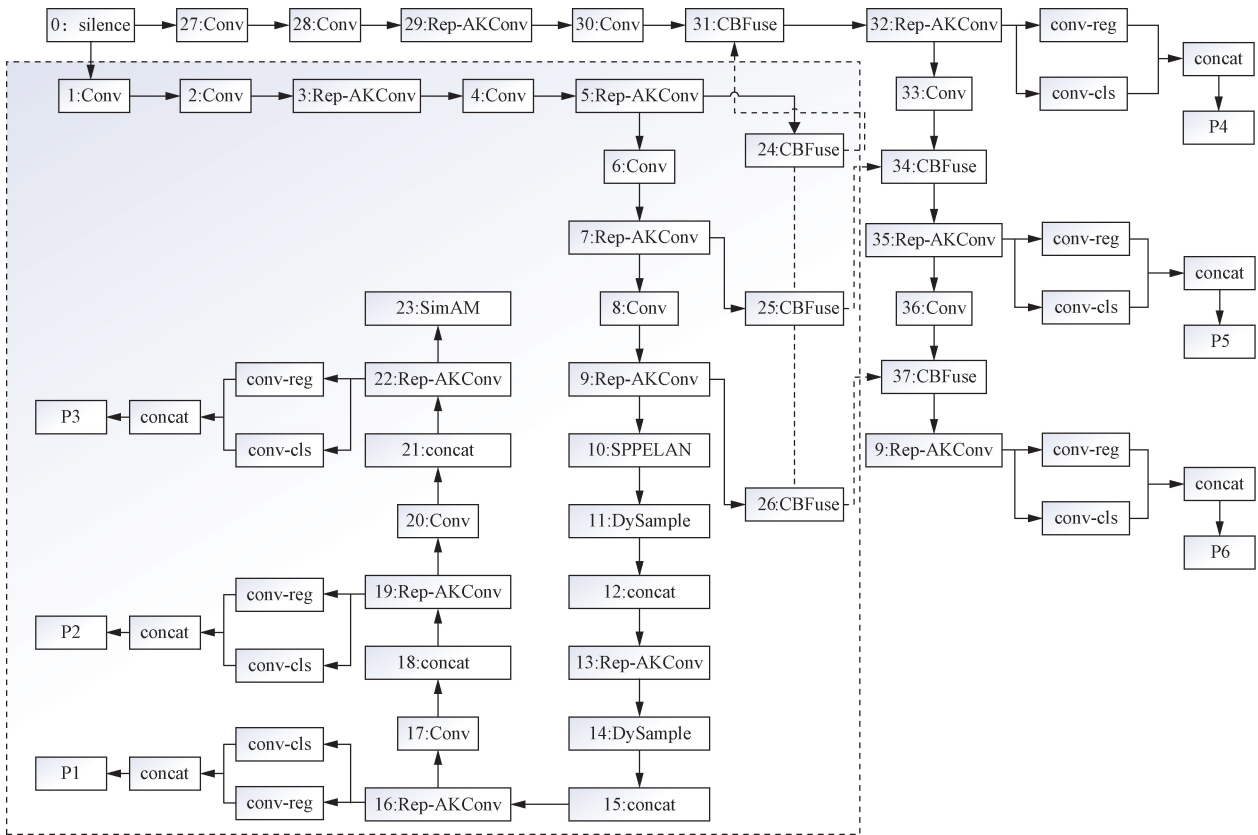
针对路面病害中常出现的坑槽和裂缝问题进行算法研究,将原始 YOLOv9 架构存在漏检误检问题进行算法改进。首先,将骨干网络中的 RepNC-SPELAN4 模块与 AKConv 模块相结合,增强了网络的特征提取能力,让算法能够更好地感知特征图的空间信息;其次,在检测头中引入了 SimAM 注意力机制,让模型能够更好地聚焦于图像关键区域的同时,增强了 convnets 模块的表达能力,让网络能够更快地提取不同病害的特征;接着将检测头部分的 nn. Upsample 模块替换为 DySample 模块,加快了模型对样本上样速度的同时,为下采样模块输入更加清晰的特征图;最后采用 inner-IoU 函数替代回归损失中的 CIoU 函数,让网络能够更好地衡量模型性能、指导参数优化,提升对差异不明显病害的学习能力。改进的 YOLOv9 网络构架图如图 1 所示。

2.1 RepNCSPELAN4-AKConv 模块

可变核卷积 (AKConv) 是一种具有任意数量的参数和任意采样形状的卷积机制,对不规则特征有更好的提取效果。RepNCSPELAN4 是 YOLOv9 中的特征提取模块,类似 YOLOv5 和 v8 中的 C2f 与 C3 模块。RepNCSPELAN4-AKConv 的主要思想是使用 AKConv 替换 RepNCSPELAN4 中的 Conv 模块,该方法突破了传统卷积局限于固定局部窗口和固定采样形状的限制,从而使得卷积操作能够更加精准地适用不同数据集和不同位置的目标,增强 YOLOv9 网络特征提取能力的同时,拥具有更高的泛用性。

传统的卷积通常具有固定的尺寸和形状,是 2x2 或 3x3 的方形卷积。本文提出 RepNCSPELAN4-AKConv 模块结构图,如图 2 所示。

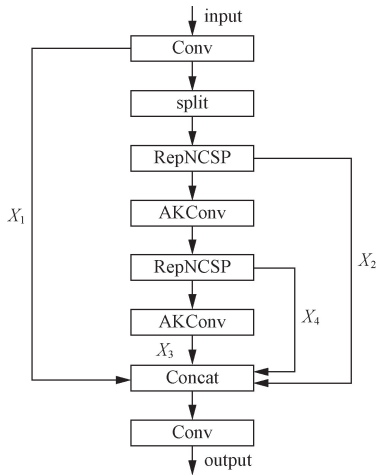
RepNCSPELAN4-AKConv 的核心原理是允许卷积核具有任意数量的参数,卷积核不再局限于标准的方形网格,根据图像特征和任务需求,采用更多样化和灵活的形状。在处理不同的图像和目标时,AKConv 的卷积核能够自动调整其采样形状,通过引入坐标生产算法,为不同大小和形状的卷积核生



Silence 为输入数据的起始占位;Conv 为卷积层;SPPELAN 为特殊池化与特征增强模块;CBFuse 为 YOLOv9 中用于特征融合的层;concat 为拼接层;conv-reg 为卷积回归层;conv-cls 为卷积分类层;P1、P2、P3、P4、P5、P6 为特征金字塔层级 1~6 层

图 1 改进的 YOLOv9 网络构架图

Fig. 1 Improved YOLOv9 network architecture diagram



X_1, X_2, X_3, X_4 为 Split 操作拆分过程;input 为输入;output 为输出;Concat 为拼接层;RepNCSP 为重参数化跨阶段局部连接

图2 RepNCSP-ELAN4-AKConv 模块结构图

Fig. 2 Structure diagram of the RepNCSP-ELAN4-AKConv module

成符合图像特征的动态初始采样坐标,这就使得卷积核可以根据图像内容的不同,而改变其采样策略,从而更有效地提取特征,同时通过学习偏移量来不断修改初始采样坐标。对于不同的目标采取相对应的采样位置,增加了 RepNCSP-ELAN4 模块的灵活性。

2.2 改进注意力模块

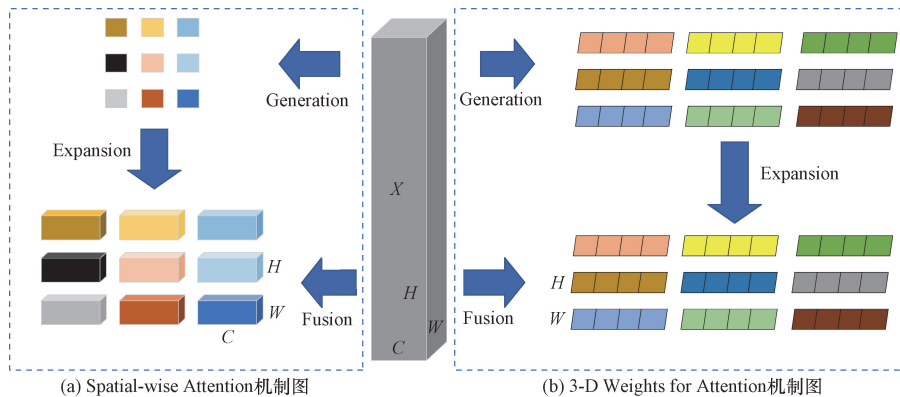
现有的注意力模块大多是根据特征 X 生成一维或二维权重,然后将生成的权重扩展为通道注意力和空间注意力,而 SimAM 注意力机制模块则直接估算三维权重。卷积神经网络注意力模块分为两种类型:通道注意力(1-D)和空间注意力(2-D)。注意力模块通常被集成到每个模块中以细化前一层的输出,而这些步骤通常会沿通道维度或空间维度进行操作。这类方法产生的一维或二维权重可能会限制它们学习更具辨别力的线索的能力,因此完整的 3-D 权重优于传统的 1-D 和 2-D 注意力。

采用基于 3-D 权重的 SimAM 注意力机制替代 2-D 的空间注意力机制,如图 3 所示。空间注意力 Spatial-wise 将 $C \times H \times W$ 的特征 X 通过 Generation 机制生成 9 个 2-D 制,结构对比如图 3 所示,接着通过 Fusion 和 Expansion 机制,将其扩展融合为 9 个与原始特征 X 相同的空间特征,但在生成和扩展特征图时,会产生信息梯度流失以及扩展错误等现象;而采用如图 3(b) 的 3-D 权重注意力时,SimAM 注意力机制将特征 X 通过 Generation 机制生成 9 个 3-D 权重的平面特征图,接着通过 Fusion 和 Expansion 机制,将其塑为完全 3-D 的平面特征图。由于全过程生成的都是平面特征图,在 Expansion 环节扩展特征图时,就能够很好地减少融入错误的信息,同时能够一定程度弥补 Generation 机制信息提取不完全的缺点。

2.3 DySample 动态上采样算子

在卷积神经网络中,由于输入图像通过卷积神经网络(CNN)提取特征后,输出的尺寸往往会变小,需要将图片恢复到原来的尺寸,以便进行进一步的计算(如图像的语义分割),需要上采样操作帮助放大图像。采用 DySample 上采样模块代替 YOLOv9 中 head 部分的 nn. Upsample 模块。DySample 上采样算子能够绕过动态卷积并从点采样的角度制定上采样,并且使用 PyTorch 中的标准内置函数实现,与以前基于内核的动态上采样器相比,DySample 不需要定制 CUDA 包,并且具有更少的参数、FLOPs、GPU 内存和延迟。除了轻量级的特点,DySample 在 5 个密集预测任务上优于其他上采样器,包括语义分割、目标检测、实例分割、全视分割和单目深度估计。

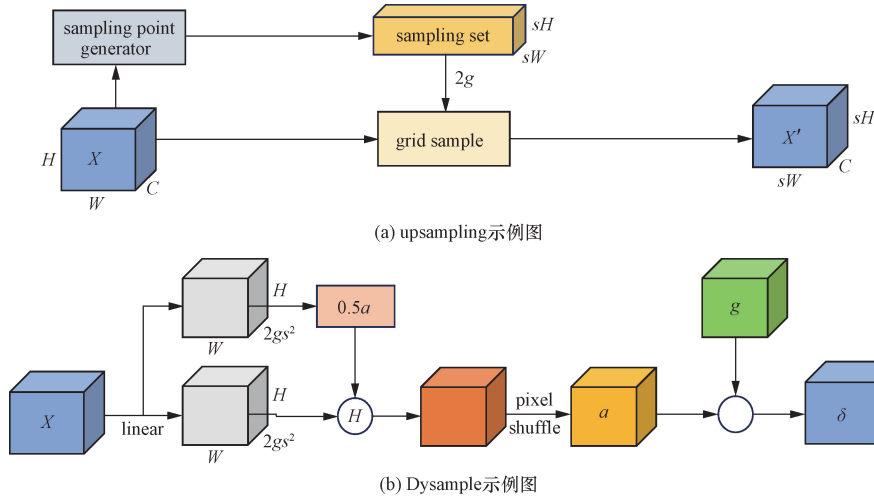
nn. Upsample 模块如图 4(a) 所示,给定一个大小为 $C \times H_1 \times W_1$ 的特征图 X ,以及一个大小为 $2g \times sH \times sW$ 的采样集 δ ,sampling point generator 为采样点生成器,sampling set 为采样集,grid sample 为网格采样



C 为通道数; H 为图像的高度; W 为图像的宽度;Generation 为生成机制;Expansion 为扩展机制;Fusion 为融合机制

图3 SimAM 机制图

Fig. 3 SimAM mechanism diagram



linear 为线性变换;pixel shuffle 为像素混洗操作

图 4 DySample 模块结构图

Fig. 4 Structure of the DySample module

操作再通过 grid sample 函数,使用 δ 中的位置对假设的双线性插值 X 进行重新采样,生成大小为 $C \times sH \times sW$ 的 X' 特征。这种局部采样位置的移动范围可能会显著重叠,这种重叠很容易影响边界附近的预测,并且此类错误会逐阶段传播并导致输出伪影图,DySample 模块如图 4(b) 所示,与 nn. Upsample 模块不同的是, DySample 额外给定一个上采样尺度因子 s ,首先使用线性层(输入和输出通道数分别为 C 和 $2s^2$)来生成大小为 $2s^2 \times H \times W$ 的偏移量 O ,其次通过 Pixel Shuffling,将其重塑为 $2 \times sH \times sW$,将偏移量 O 和原始采样网 g 相加得到采样集 δ 。通过使用 sigmoid 函数和 0.5 的静态因子,使得采样位置的移动范围在局部受到约束,不仅满足重叠与非重叠之间的理论边界条件,而且增加偏移量的灵活性。

2.4 改进损失函数

损失函数(loss function)是用来估量模型的预测值 $f(x)$ 与真实值 Y 的不一致程度,它是一个非负实值函数,通常使用 $L[Y, f(x)]$ 来表示,损失函数越小,模型的鲁棒性就越好。YOLOv9 的损失函数包括两类函数:①分类损失:BCE Loss;②回归损失:DFL Loss + CIoU Loss。

采用 inner-IoU 函数替代 CIoU 函数改进了回归损失。CIoU 计算公式为

$$\text{IoU} = \frac{|B^{\text{pred}} \cap B^{\text{gt}}|}{|B^{\text{pred}} \cup B^{\text{gt}}|} \quad (1)$$

$$\text{CIoU} = \text{IoU} - \left[\frac{\rho^2(B^{\text{pred}}, B^{\text{gt}})}{c^2} + \alpha v \right] \quad (2)$$

式中: B^{pred} 为预测框区域面积; B^{gt} 为真实框区域面积; c 为预测框与真实框的最小外接矩形的对角线长度; $\rho^2(B^{\text{pred}}, B^{\text{gt}})$ 为欧几里得距离的平方; u 用来

衡量长宽比; a 为平衡参数,表示预测框与真实框中心点之间的欧氏距离。

CIoU 在 IoU 基础上加上对中心点距离和长宽比的权衡,检测效果得到有效提升。这些改进的损失函数仍一直是通过加入新损失项来加速收敛,并没有意识到 IoU 自身的限制。Inner-IoU 提出以辅助边框来计算 IoU,提升泛化能力并且加快了收敛速度,具体计算过程如式(3)~式(5)所示。

$$\text{inter} = [\min(b_r^{\text{gt}}, b_r) - \max(b_l^{\text{gt}}, b_l)] \times [\min(b_b^{\text{gt}}, b_b) - \max(b_t^{\text{gt}}, b_t)] \quad (3)$$

$$\text{union} = w^{\text{gt}} h^{\text{gt}} (\text{ratio})^2 + wh (\text{ratio})^2 - \text{inter} \quad (4)$$

$$\text{IoU}^{\text{inner}} = \frac{\text{inter}}{\text{union}} \quad (5)$$

式中: inter 为预测框和真实框交集区域的面积; union 为预测框和真实框并集区域的面积; $\text{IoU}^{\text{inner}}$ 为预测框和真实框的交并比; b_r^{gt} 为真实右边界; b_l 为右边界坐标; b_l^{gt} 为真实左边界; b_l 为左边界; b_b^{gt} 为真实下边界; b_b 为下边界; b_t^{gt} 为真实上边界; b_t 为上边界; $\min(b_r^{\text{gt}}, b_r) - \max(b_l^{\text{gt}}, b_l)$ 为水平方向交集的长度; $\min(b_b^{\text{gt}}, b_b) - \max(b_t^{\text{gt}}, b_t)$ 为垂直方向交集的长度; w^{gt} 和 h^{gt} 分别为 GT 框的宽度和高度; w 和 h 分别为锚框的宽度和高度。

GT 框和锚框分布表示为 B^{gt} 和 B ; GT 框和 GT 框内部的中心点用 $(x_c^{\text{gt}}, y_c^{\text{gt}})$ 表示;锚框和内部锚框的中心点用 (x_c, y_c) 表示;ratio 为尺度因子,通常取值范围为 $[0.5, 1.5]$ 。为确定 ratio 值,采用天津大学机器学习与数据挖掘实验室所整理公开的 Vis-Drone2019 数据集进行消融实验^[18],实验结果如表 1 所示。

表 1 ratio 消融实验

Table 1 Ratio ablation experiment

ratio	P/%	R/%	mAP@0.5	mAP@0.5:0.9
0.5	55.4	42.3	44.9	27.2
0.7	54.2	42.4	44.1	26.8
1	56.2	42.2	45.0	27.4
1.25	54.4	43.7	45.2	27.6
1.5	54.3	43.1	44.7	27.3

当 ratio = 1 时,精度 P 为最大值 56.2%,而召回率 R 和 mAP@0.5 以及 mAP@0.5:0.9 在 ratio = 1.25 时达到最大值。对于不同的研究对象和不同的需求,ratio 应该选取不同的值,考虑到路面混合病害差异不明显的特点,ratio 取值为 1。

3 实验设计与结果分析

3.1 实验设定

实验使用的 GPU 为 NVIDIA GeForce RTX 3080 Laptop,CPU 使用了 Intel(R) Core(TM) i9-10900K CPU,pytorch 框架版本为 1.13.1+cu117,python 版本为 3.8.18,cuda 版本为 11.7。实验数据集来源于开源数据集 Concrete Crack Images for Classification 和河北省保定市保京高速和保京大桥实地考察图片。从中共计选取了 1 000 张裂缝图片,随机按比例 8:1:1 分为训练集,测试集和验证集。数据增强采用 Mosaic 图像拼接技术,Epoch 训练轮数设置为 300 轮,batch-size 为批次大小设置为 16,Patience 为早停机制,Pretrained 为预训练模型。检测目标种类为 2 种,分别为:crack(裂缝)和 hole(坑槽),运行文件采用的是 yolov9-main 下的 train_dual.py 文件,即一个辅助分支 + 一个主分支,本次实验网络训练的超参数如表 2 所示。

3.2 评估指标

选用精度 P、召回率 R、平均精度均值 mAP (mean average precision) 作为评估指标,公式为

表 2 训练超参数设置

Table 2 Training hyperparameter settings

项目设置	设置策略
数据增强	Mosaic
更新学习率	余弦退火算法
优化器	Auto
Epoch	300
Batch-size	16
Patience	100
Pretrained	False

$$P = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (7)$$

$$AP = \int_0^1 P(R) dR \quad (8)$$

$$mAP = \frac{\sum_{i=0}^n AP_i}{n} \quad (9)$$

式中:TP 为预测样本中将正样本正确预测正确的个数;FP 为预测样本中将负样本错误预测为正样本的个数;n 为检测目标类别数(本文取值 2);P(r) 为不同召回率水平下的精度函数;AP 为不同召回率水平下精度之和。

mAP@0.5:即将 IoU 设为 0.5 时,计算每一类的所有图片的 AP(AP 是对 P-R 曲线的积分,即曲线与横纵坐标所围成的面积),然后所有类别求平均。

mAP@0.5:0.95:表示在不同 IoU 阈值(从 0.5 到 0.95,步长 0.05 即 0.5、0.55、0.6、0.65、0.7、0.75、0.8、0.85、0.9、0.95)上的平均 mAP。

3.3 消融实验

为了验证所提出的改进方案对模型检测效果的提升,YOLOv9-c 模型的基础上进行了消融实验,每一组实验都设置相同的超参数,采用相同的训练策略,实验结果如表 3 所示。

表 3 消融实验结果

Table 3 Ablation test results

模型	RepNCSPELAN4-AKConv	SimAM	DySample	Inner-IOU	GFLOPs	MB	P/%	R/%	mAP@0.5	mAP@0.5:0.9
Yolov9-c	×	×	×	×	236.6	100.0	49.01	50.52	44.40	23.01
方案 1	√	×	×	√	251.2	108.5	35.90	59.30	38.60	21.70
方案 2	×	√	×	√	236.6	102.7	63.10	48.90	55.00	26.50
方案 3	×	×	√	√	236.7	102.8	51.60	57.90	54.80	27.90
方案 4	√	√	×	√	251.2	108.5	67.90	56.80	57.40	21.60
方案 5	√	×	√	√	251.2	108.5	13.60	28.20	16.00	9.160
方案 6	×	√	√	√	236.7	102.8	54.70	56.80	53.20	25.50
方案 7	√	√	√	√	251.2	108.5	64.60	57.90	53.70	19.20

注:√表示在 YOLOv9 网络基础上引入了对应的改进方案;×表示未引入该改进方案。

对比上述方案,当网络采用 RepNCSPELAN4-AKConv 模块时会增加网络的特征提取能力,但是会增加网络运算所占用的内存,引入 SimAM 注意力模块,能够有效地增加网络的精度和 mAP 指数,采用 DySample 上采样算子时能够小幅度增加精度,对于召回率和 mAP 有显著提升的同时,缩短了网络的训练时间。对比上述方案可知方案 2、方案 4、方案 6 和方案 7 均获得了较为明显的提升。方案 2 的精度提升了 28.7%, mAP@0.5 提升了 23.8%,但是回归率却有了小幅度的降低;方案 4 的精度提升了 38.5%,回归率提升了 12.4%,mAP 提升了 29.2%;方案 6 的精度提升了 11.6%,回归率提升了 12.4%,mAP 提升了 19.8%;方案 7 精确度提升了 31.8%,回归率提升了 14.6%,mAP 提升了 20.9%,方案 4 在 epoch 为 300 轮时效果最佳,但是方案 4 和方案 7 均没有提前结束训练,对于这两种

方案进行额外训练,epoch 设定为 600 轮进行二次训练,增加新的方案进行训练。

从表 4 可知,方案 8 和方案 9 各项指数几乎一致,但两者的训练时长有明显差异,方案 8 在第 534 轮时停止训练,共计耗时 8.7 h;方案九在第 503 轮时停止训练,共计耗时 7.7 h,因此认为方案 9 要比方案 8 高效。综上所述,可知方案 9 为最优方案,评估指标改进效果如图 5 所示。

3.4 检测结果可视化分析

模型训练输出结果如图 6 和图 7 所示,通过对比图 7 中 Label 和 Ours-v9 可知,在 4 种不同的单一病害路面中,提出的网络能够很好地进行病害的检测,在图 7(b)中,改进的网络不仅能够检测到 v9 网络漏检的裂缝,一定程度上克服了 YOLOv9-c 网络漏检现象同时甚至能够检测到额外的砌块裂缝,这一现象表明改进后的网络拥有更好的检测效果并

表 4 附加方案实验结果

Table 4 Experimental results of additional protocols

模型	RepNCSPELAN4-AKConv	SimAM	DySample	Inner-IOU	GFLOPs	MB	P/%	R/%	mAP@0.5	mAP@0.5:0.9
方案 8	√	√	×	√	251.2	102.8	68.9	58.4	53.2	24.6
方案 9	√	√	√	√	251.2	108.5	68.7	58.6	53.7	22.3

注:√表示在 YOLOv9 网络基础上引入了对应的改进方案;×表示未引入该改进方案。

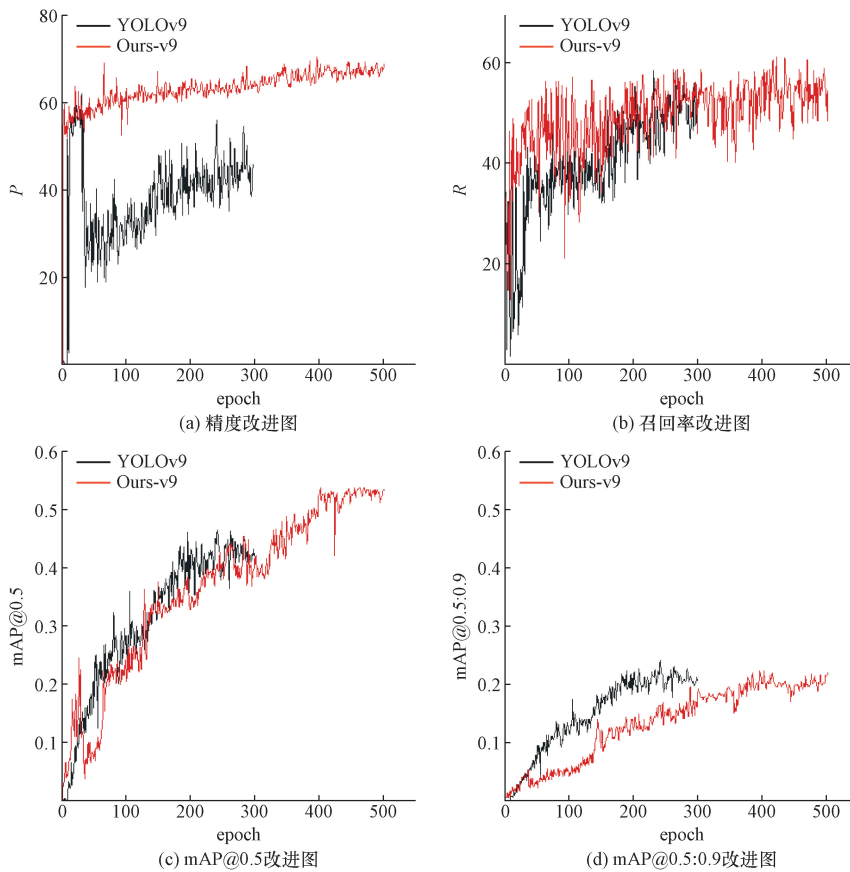


图 5 评估指标改进可视化图

Fig. 5 Evaluation indicator improvement visualization

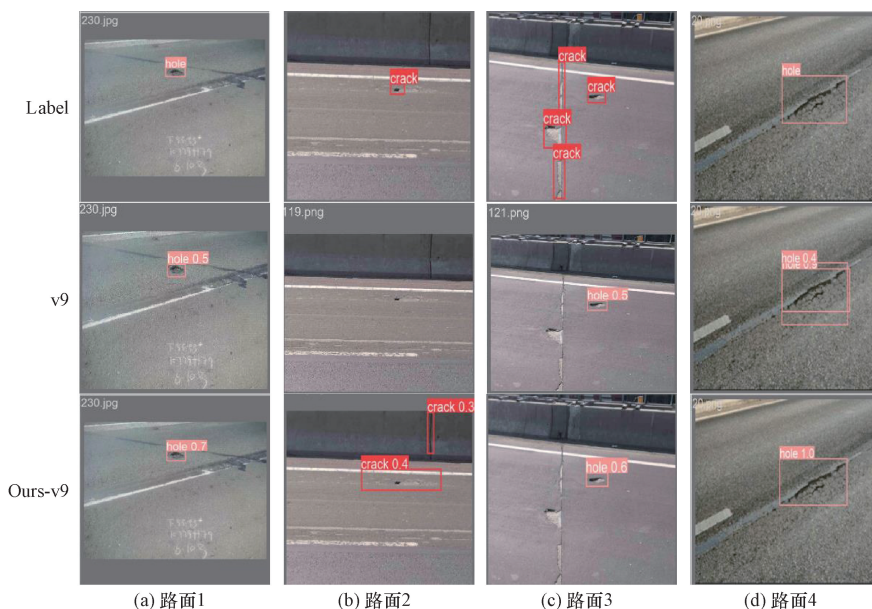


图6 单一病害检测可视化结果

Fig. 6 Visualization of a single disease detection

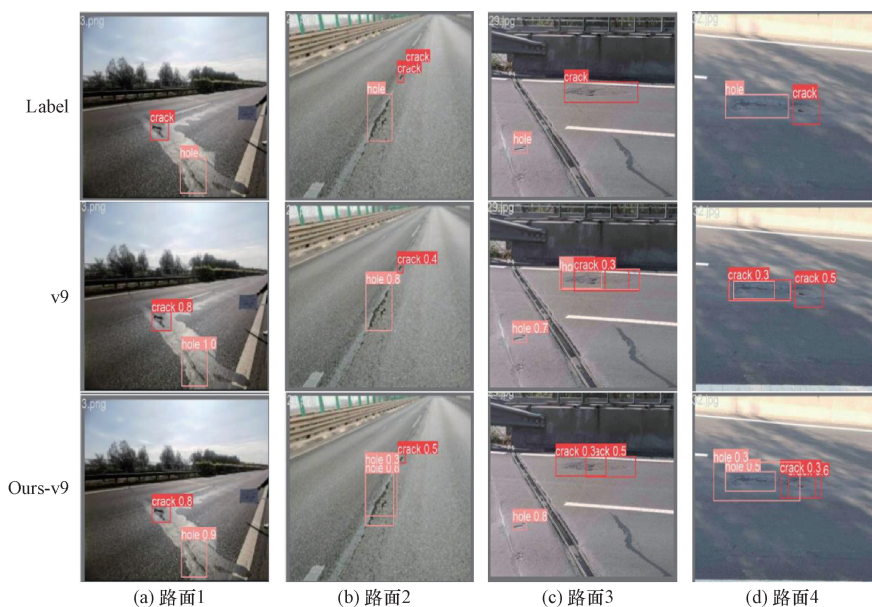


图7 混合病害检测可视化结果

Fig. 7 Visualization results of mixed disease detection

且对于裂缝这一病害能够在路面以外的背景条件下进行检测(如墙体);即使出现了如图6(c)中标签制作错误的情况,图中标签更加符合 hole,但是错误的将标签制作成 crack,但是由于大量的样本是正确的,改进前后的网络均能够对于病害进行正确的预测;在图6(c)中,改进前的网络对于同一条裂缝错误的判断为了两条叠合在一起的裂缝,而改进后的网络很好地克服了这一现象。综上可知,本文提出的改进对于单一病害的检测效果具有良好的提升。

由图7可知,当识别对象为同一背景下的混合病害时,改进的网络仍然能有较好的检测效果。当检测病害稀疏区域时如图7(a)和图7(c)所示,改进前后的网络均具有较好的检测效果,但是在图7(c)中 v9 网络错误的一个 crack 病害检测为 hole 和 crack 两个病害,检测效果与 Ours-v9 相比有明显缺陷;但当检测混合病害密集区域时如图7(b)和图7(d)时,虽能改进前后的网络均能够对图中的病害进行检测,但是对于病害的检测出现对同一病害错误的识别为多个病害,从而导致产生

图中多个锚框重叠的问题;同时对比如图 7(c) 和图 7(d) 可知,在图 7(c) 中对 crack 检测是错误的判断为了 hole,在图 7(d) 中又错误的将 hole 预测为了 crack,而提出的改进方案一定程度上克服了这两处误检现象。

在制作标签时,由于主观意识导致标签很容易产生如图 6(c) 中并不统一的效果,这就会导致最终的模型精度和召回率偏低,实际的精度和召回率要比 68.7%、58.6% 要高,但是在训练结束后的模型却能够克服这一误差,对于裂缝的定位和分割有着比人为制作更具有规律性。YOLOv9-c 网络对于裂缝的检测精度为 43.4%,召回率为 56.8%,map@0.5 为 43.4%。对于坑槽的检测精度为 54.62%,召回率为 44.24%,map@0.5 为 45.4%;改进的 YOLOv9 对于裂缝的检测精度有较大提升,达到了 62.4%,但是召回率则降低到了 49.7%,map@0.5 为 47.2% 有小幅增加,对于坑槽的检测精度为 75%,召回率为 67.5%,map@0.5 为 60.2%,各方面均有较大提升,改进的网络能够更好地检测混合病害。

3.5 对比试验

为了进一步验证改进算法对路面病害检测的优越性,将改进算法与传统的目标检测算法进行对比试验,其中二阶段检测算法有 Faster-RCNN 算法,一阶段检测算法有 YOLOv3、YOLOv4、YOLOv5、YOLOX、YOLOv7、YOLOv8n、TMDet^[19]、PPYOLOE^[20] 算法,结果如表 5 所示,由表 5 可知,改进算法提高了 YOLOv9-c 模型对混合病害的识别精度,能够对路面的损伤进行更好的检测。该算法在参数量和 GFLOPs 上并不优于大部分的其他算法,但是 P、R 和 mAP 的值均为最大值。综合考虑,所提出网络优化算法,对于检测路面混合病害具有较好的优越性和鲁棒性。

表 5 对比试验结果

Table 5 compares the test results

模型	参数量/MB	GFLOPs	P/%	R/%	mAP@0.5
Faster-RCNN	136.52	370.2	47.90	42.60	49.10
YOLOv3	207.80	282.3	58.00	43.80	45.30
YOLOv5	5.30	7.2	44.60	32.60	33.10
YOLOX	8.94	13.4	50.20	49.70	52.60
YOLOv7	6.03	13.2	51.80	44.30	53.50
YOLOv8n	3.01	8.2	57.20	50.40	56.50
TMDet	4.83	8.1	54.80	58.20	56.20
PPYOLOE	7.93	17.4	55.30	57.50	56.90
YOLOv9c	100.00	236.6	49.01	50.52	44.40
Ours-v9	108.50	251.2	68.70	58.60	53.70

4 结论

针对坑槽和裂缝两种混合路面病害,提出了一

种改进的 YOLOv9-c 模型,讨论了 YOLOv5、YOLOv8n、YOLOv9c 等多种模型的精度和计算速度、内存占用量等,在主干结构 backbone 中引入 RepNCSPeLan4-AKConv 模块和 simAM 注意力机制,在 head 部分采用 upSample 上采样,以 inner-IoU 函数替代 CIoU 函数,提升对混合病害的学习能力。通过网络训练,得到以下结论。

(1) 在路面混合病害检测中,与其他算法对比,基于 YOLOv9-c 改进的新模型中并且能够更快地处理复杂的计算任务。

(2) 与原 YOLOv9-c 训练时长相比,更换上采样子后,平均 4.5 h 训练 300epoch,模型训练速度增快。

(3) 通过改进卷积核并引入新的注意力机制模型,在一定程度上改善了 YOLOv9 网络的特征提取能力,减少了图像信息聚焦不明显而产生的漏检和误检等问题。

(4) 在 YOLOv9 网络中使用 inner-IoU 函数来指导模型的训练,提升模型对混合样本的学习能力,让新的模型能够更好地识别不同背景下的病害,使模型具有更好的泛化性和鲁棒性。

参 考 文 献

- [1] 何铁军,李华恩. 基于改进 YOLOv5 的路面病害检测模型[J]. 土木工程学报, 2024, 57(2): 96-106.
He Tiejun, Li Huaen. Pavement disease detection model based on improved YOLOv5 [J]. Chinese Journal of Civil Engineering, 2024, 57(2): 96-106.
- [2] 朱洪洲,阳绪缘. 沥青路面内部缺陷无损检测技术研究综述[J]. 科学技术与工程, 2024, 24(25): 10588-10604.
Zhu Hongzhou, Yang Xuyuan. Review of non-destructive testing techniques for internal defects in asphalt pavements [J]. Science Technology and Engineering, 2024, 24(25): 10588-10604.
- [3] 刘宇飞,樊健生,聂建国,等. 结构表面裂缝数字图像法识别研究综述与前景展望[J]. 土木工程学报, 2021, 54(6): 79-98.
Liu Yufei, Fan Jiansheng, Nie Jianguo, et al. Review and prospect of digital-image-based crack detection of structure surface [J]. China Civil Engineering Journal, 2021, 54(6): 79-98.
- [4] 宋晔,吴一全. 基于无人机航拍的风力发电机叶片表面缺陷检测综述[J]. 仪器仪表学报, 2024, 45(10): 1-25.
Song Ye, Wu Yiquan. Review on surface defect detection of wind turbine blades based on UAV aerial photography [J]. Chinese Journal of Scientific Instrument, 2024, 45(10): 1-25.
- [5] 刘传洋,吴一全,刘景景. 无人机航拍图像中绝缘子缺陷检测的深度学习研究方法研究进展[J]. 电工技术学报, 2024. DOI: 10.19595/j. enki. 1000-6753. tees. 240610.
Liu Chuanyang, Wu Yiquan, Liu Jingjing. Research progress of deep learning methods for insulator defect detection in UAV aerial images [J]. Transactions of China Electrotechnical Society, 2024. DOI: 10.19595/j. enki. 1000-6753. tees. 240610.
- [6] Pan X, Yang T Y, Xiao Y F, et al. Vision-based real-time struc-

- tural vibration measurement through deep-learning-based detection and tracking methods [J]. *Engineering Structures*, 2023, 281. DOI: 10.1016/j.engstruct.2023.115676.
- [7] Zhang J, Qian S R, Tan C. Automated bridge surface crack detection and segmentation using computer vision-based deep learning model [J]. *Engineering Applications of Artificial Intelligence*, 2022, 115. DOI: 10.1016/j.engappai.2022.105225.
- [8] 廖祥灿, 李彩林, 姚玉凯, 等. 基于改进 YOLO V5 的公路桥梁裂缝检测方法[J]. *山东理工大学学报(自然科学版)*, 2023, 37(4): 1-7.
Liao Xiangcan, Li Cailin, Yao Yukai, et al. Crack detection method of highway bridge based on improved YOLO V5 [J]. *Journal of Shandong University of Technology (Natural Science Edition)*, 2023, 37(4): 1-7.
- [9] 杜磊, 陈曦, 白朋朋, 等. 基于改进 YOLOv5 的沥青路面裂缝检测方法[J]. *科学技术与工程*, 2024, 24(31): 13579-13591.
Du Lei, Chen Xi, Bai Pengpeng, et al. Crack detection method for asphalt pavement based on the improved YOLOv5 [J]. *Science Technology and Engineering*, 2024, 24(31): 13579-13591.
- [10] 贾晓芬, 江再亮, 赵佰亭. 裂缝小目标缺陷的轻量化检测方法[J]. *湖南大学学报(自然科学版)*, 2024, 51(6): 52-62.
Jia Xiaofen, Jiang Zailiang, Zhao Baiting. Lightweight detection method for small crack target defects[J]. *Journal of Hunan University (Natural Science Edition)*, 2024, 51(6): 52-62.
- [11] 夏翔, 朱明. 改进 YOLOv7 的视频监控小目标检测[J]. *计算机系统应用*, 2024, 33(7): 52-62.
Xia Xiang, Zhu Ming. Small target detection in video surveillance based on improved YOLOv7 [J]. *Computer Systems & Applications*, 2024, 33(7): 52-62.
- [12] 王海群, 王炳楠, 葛超. 重参数化 YOLOv8 路面病害检测算法[J]. *计算机工程与应用*, 2024, 60(5): 191-199.
Wang Haiqun, Wang Bingnan, Ge Chao. Reparameterized YOLOv8 pavement disease detection algorithm [J]. *Computer Engineering and Applications*, 2024, 60(5): 191-199.
- [13] Zhang X, Song Y Z, Song T T, et al. AKConv: convolutional kernel with arbitrary sampled shapes and arbitrary number of parameters[J]. arXiv: 2311.11587, 2023.
- [14] Yang L, Zhang R Y, Li L, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks [C]//*Proceedings of the 38th International Conference on Machine Learning*. New York: PMLR, 2021: 11863-11874.
- [15] Liu W Z, Lu H, Fu H, et al. Learning to upsample by learning to sample[J]. arXiv: 2308.15805, 2023.
- [16] Zhang H, Xu C, Zhang S. Inner-IoU: more effective intersection over union loss with auxiliary bounding bBox [J]. arXiv: 2311.02877, 2023.
- [17] Wang C Y, Yeh I H, Hong Yuan M L. YOLOv9: Learning what you want to learn using programmable gradient information [J]. arXiv: 2402.13616, 2024.
- [18] Du D, Zhu P, Wen L, et al. VisDrone-DET2019: the vision meets drone object detection in image challenge results [C]// *Proceedings of the IEEE/ CVF International Conference on Computer Vision Workshops*. Seoul: IEEE, 2019: 213-226.
- [19] Lü C, Zhang W, Huang H, et al. RtmDet: an empirical study of designing real-time object detectors [J]. arXiv: 2212.07784, 2022.
- [20] Xu S, Wang X, Lü W, et al. PP-YOLOE: an evolved version of YOLO [J]. arXiv: 2203.16250, 2022.