



DOI:10.12404/j.issn.1671-1815.2403184

引用格式:原洪帅,李琦,王月明.基于YOLO-CD的路面裂缝检测[J].科学技术与工程,2025,25(9):3888-3895.

Yuan Hongshuai, Li Qi, Wang Yueming. Road crack detection based on YOLO-CD[J]. Science Technology and Engineering, 2025, 25(9): 3888-3895.

基于YOLO-CD的路面裂缝检测

原洪帅,李琦*,王月明

(内蒙古科技大学自动化与电气工程学院,包头 014010)

摘要 为解决现有路面裂缝检测算法精度低、漏检率高等问题,提出了一种基于YOLOv8n改进的路面裂缝检测算法YOLO-CD(YOLO-crack detection)。YOLO-CD算法采用ASF-YOLO架构中的尺度序列特征融合(scale sequence feature fusion,SSFF)模块和三重特征编码器(triple feature encoder,TFE)模块,提高了对多尺度裂缝的检测性能和对目标特征的感知能力。同时,在骨干网络末端和颈部网络中引入CA注意力机制(coordinate attention),将位置信息嵌入通道注意力中,强化了对裂缝特征的提取能力。此外,在YOLOv8n原有3个输出层的基础上新增了一个P2小目标检测层,增加了网络的多尺度感受野,可以同时捕获全局和局部上下文信息,提升了算法在复杂场景中对小目标裂缝的检测能力。通过将原始YOLOv8n的检测头替换为DyHead检测头,使尺度、空间和任务3种注意力机制结合统一,进一步提升了网络对裂缝的检测性能。实验结果表明,在自建数据集PD-Dataset中,改进后的算法YOLO-CD比原算法YOLOv8n的mAP50提高了4.1%。在公共数据集RDD2020中,改进后的算法YOLO-CD比原算法YOLOv8n的mAP50提高了1.5%。且算法检测速度达到了89.9帧/s,满足了路面裂缝检测实时性的要求。

关键词 路面裂缝检测;YOLOv8n;ASF-YOLO;注意力机制;小目标检测层;DyHead检测头

中图法分类号 U418.6;

文献标志码 A

Road Crack Detection Based on YOLO-CD

YUAN Hong-shuai, LI Qi*, WANG Yue-ming

(Automation and Electrical Engineering College, Inner Mongolia University of Science & Technology, Baotou 014010, China)

[Abstract] In order to address the issues of low accuracy and high missed detection rates in existing pavement crack detection algorithms, an improved pavement crack detection algorithm based on YOLOv8n, named YOLO-CD (YOLO-crack detection), has been proposed. The scale sequence feature fusion (SSFF) module and triple feature encoder (TFE) module from the ASF-YOLO architecture were utilized by the YOLO-CD algorithm to enhance the detection performance for multi-scale cracks and the perception capability of target features. Additionally, the coordinate attention(CA) mechanism was introduced at the end of the backbone network and in the neck network, with positional information embedded into channel attention, thereby strengthening the extraction capability of crack features. Furthermore, an additional P2 small object detection layer was added on top of the original three output layers of YOLOv8n, increasing the multi-scale receptive field of the network, allowing both global and local context information to be captured simultaneously, thereby improving the detection capability for small cracks in complex scenes. The original YOLOv8n detection head was replaced by the DyHead detection head, achieving the integration of scale, spatial, and task attention mechanisms, and further enhancing the network's detection performance for cracks. Experimental results show that in the self-built PD-Dataset, the mAP50 of the improved YOLO-CD algorithm is increased by 4.1% compared to the original YOLOv8n algorithm. In the public dataset RDD2020, the mAP50 of the improved YOLO-CD algorithm is increased by 1.5% compared to the original YOLOv8n algorithm. Moreover, the algorithm's detection speed is found to reach 89.9 frames/s, meeting the real-time requirements of pavement crack detection.

[Keywords] road crack detection; YOLOv8n; ASF-YOLO; attention mechanism; small object detection layer; DyHead detection head

随着中国公路建设规模的扩大,公路维护难度增加。车辆行驶、地基下沉等因素导致路面出现各种裂缝,危及行车安全且缩短公路寿命。传统的人工巡查方式效率低下、准确度有限,且工作枯燥,可能存在安全隐患。计算机视觉和深度学习为解决这一问题提

供了可能性。深度学习通过大规模训练样本获得强大的泛化能力,使路面裂缝检测成为现实。这种方法提高了检测效率,使其更准确、全面。

在目标检测领域,研究者们进行了众多经典算法的探索,其中包括单发多箱探测器(single shot multibox

收稿日期:2024-04-29 修订日期:2024-12-21

基金项目:内蒙古自治区关键技术攻关项目(2020GG0316)

第一作者:原洪帅(1996—),男,汉族,山东平度人,硕士研究生。研究方向:深度学习与人工智能。E-mail:928710929@qq.com。

*通信作者:李琦(1973—),男,汉族,陕西米脂人,硕士,教授。研究方向:图像处理。E-mail:richey@imust.edu.cn。

detector,SSD)^[1]、快速区域卷积神经网络(faster region convolutional neural networks, Faster R-CNN)^[2]、和 YOLO(you only look once)算法^[3]。而 YOLO 算法则以其出色的速度和较高的准确率著称,它使用深度神经网络来同时进行目标分类和位置检测。

随着 YOLO 算法的不断更新,包括 YOLOv3^[4]、YOLOv4^[5]、YOLOv5、YOLOX^[6]、YOLOv7、YOLOv8 和 YOLOv10 的版本也相继推出。当前,YOLO 算法已经被广泛应用于路面裂缝检测领域。Ren J^[7] 等人在 YOLOv5 中增加 CA (coordinate attention) 注意力机制,可以有效提高路面裂缝检测能力,高于其他传统和深度学习方法。林轶等^[8] 提出了一种基于改进的 YOLOv3 的 S-YOLO 交通标志算法,通过合并批归一化层、二分 K-means 聚类算法等,提高了交通标志检测的准确性和速度。徐国标等^[9] 提出了一种基于 Darknet-53 网络的改进 YOLO 算法,利用运动目标图像坐标的偏移量实现边界框长度和宽度的线性变换,解决了传统 YOLO 算法不能区分不同大小边界框的问题。Ren 等^[10] 提出了一种自动路面裂缝检测方法,通过融合注意力机制,实现对裂缝的准确检测。李松等^[11] 提出了一种改进的轻量级道路损伤检测算法 YOLOv8-RD,结合 CNN 和 Transformer 的优势,引入 BOT 模块和坐标注意力机制,成功提高了检测效率。史涛等^[12] 提出了一种改进 YOLOv8 的无人机车辆检测算法 YOLOv8-CX,通过引入 C2f-DCN 模块和 SPPF-LSKA 模块,以及采

用特征融合结构 CF-FPN,实现了对微小目标车辆的实时检测,并有效减少背景干扰。

由于 YOLO 算法未能有效解决小目标特征提取能力不足等问题,对路面裂缝等小目标的检测容易出现漏检、误检等现象。为了解决上述问题,现提出一种基于 YOLOv8n 改进的路面裂缝检测算法 YOLO-CD。采用 ASF-YOLO^[13] 架构中的尺度序列特征融合(scale sequence feature fusion, SSFF)模块和三重特征编码器(triple feature encoder, TFE)模块,提高对裂缝的检测性能和感知能力。添加 CA 注意力机制^[14],将位置信息嵌入到通道注意力中,增强对裂缝相关特征的提取。通过添加 P2 小目标检测层,网络在处理小尺寸裂缝时提高检测准确性和鲁棒性。并增强对小目标裂缝的感知能力,有效减少漏检和误检。将检测头替换为 DyHead (Dynamic Head)^[15], DyHead 动态检测头相对于普通的 Detect 头具有动态可调整卷积核大小的优势。虽参数量稍有增加,但提高裂缝特征提取能力和适应性。

1 算法改进

为了增强路面裂缝检测的精确性,提出了改进 YOLOv8n 的网络算法 YOLO-CD。YOLO-CD 网络结构如图 1 所示。

在 YOLO-CD 网络中,采用 ASF-YOLO 架构中的尺度序列特征融合(SSFF)模块和三重特征编码器(TFE)模块。SSFF 模块能获取到路面裂缝在不同

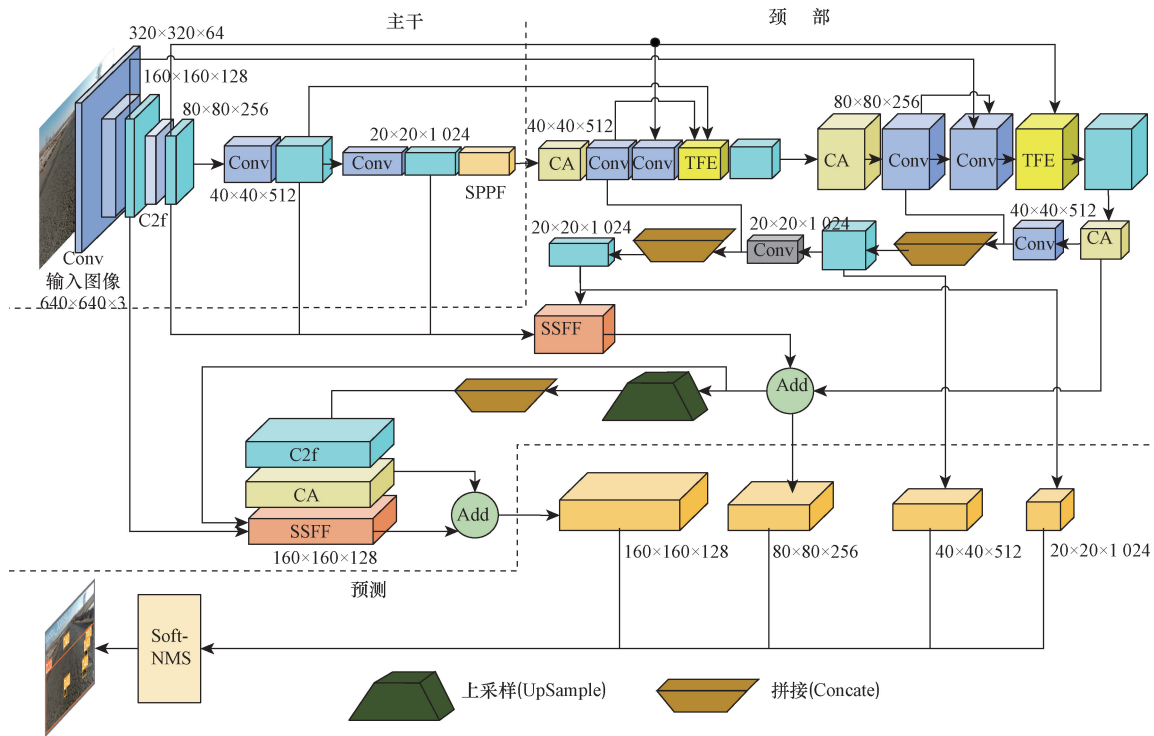


图 1 YOLO-CD 网络结构
Fig. 1 YOLO-CD network architecture

尺度上的信息,增强对裂缝特征的提取能力,而 TFE 模块则增强对裂缝形态和纹理的捕捉能力。此外,将 CA 模块嵌入 YOLOv8n 的骨干网络 SPPF 模块之后以及颈部的 C2f 之后,突出路面裂缝检测的关键信息,增强了网络对路面裂缝的特征提取能力。同时,引入了小目标检测层,进一步提高对小目标裂缝的检测精度。为了更准确地识别复杂背景下的裂缝区域,检测头更换为 DyHead。这些改进使得 YOLO-CD 算法能够更准确地检测路面裂缝,从而提升了检测的精度和可靠性。

1.1 ASF-YOLO 架构

ASF-YOLO,即基于注意力尺度序列融合的YOLO 框架,将空间和尺度特征结合在一起。该框架在传统 YOLO 框架的基础上,引入了尺度序列特征融合(SSFF)模块,以增强网络对多尺度信息的提取能力,并采用了三重特征编码器(TFE)模块来融合不同尺度的特征图,提供更加详细的信息。

尺度序列特征融合(SSFF)模块,旨在结合深度特征图的高维信息与浅层特征图的详细信息。尺度空间是沿着图像的尺度轴构建的,代表物体可以具有的各种尺度范围。模糊后的图像可能会失去细节,但图像的结构特征得以保留。作为 SSFF 输入的缩放后的图像可以通过其中, $f(\omega, h)$ 表示宽度为 ω 、高度为 h 的二维输入图像。 $F_{\sigma}(\omega, h)$ 是利用二维高斯滤波器 $G_{\sigma}(\omega, h)$ 对一系列卷积“ \times ”进行平滑得到的, σ 是用于卷积的二维高斯滤波器的标准差的缩放参数。

$$F_{\sigma}(\omega, h) = G_{\sigma}(\omega, h) \times f(\omega, h) \quad (1)$$

$$G_{\sigma}(\omega, h) = \frac{1}{2\pi\sigma^2} e^{-\frac{(\omega^2+h^2)}{2\sigma^2}} \quad (2)$$

SSFF 模块利用缩放后的图像进行处理,通过高斯滤波器对一系列卷积进行平滑处理,得到具有相同分辨率但不同尺度的结果图像。这种方法在小目标检测和分割中表现出良好效果,特别适用于基于 P3 级别设计的 SSFF 模块。

SSFF 模块使用 1×1 卷积将 P4 和 P5 特征图的通道数改为 256,采用最近邻插值法将其大小调整为 P3 级的大小,利用 unsqueeze 方法增加每个特征层的维数,将其从一个三维张量[高度、宽度、通道]变为一个四维张量[深度、高度、宽度、通道],将四维特征图沿深度维度进行拼接,形成三维特征图,用于后续的卷积,最后利用三维卷积、三维批处理归一化和 SiLU 激活函数完成尺度序列特征提取。

为了解决密集重叠小物体的识别问题,引入了 TFE 模块,旨在拆分并增加大、中、小尺度的特征图,并对其进行特征放大,以提升详细特征信息的质量。

每个三特征编码器模块均采用 3 个不同大小的特征映射作为输入,并通过最近邻插值法进行上采样。通过这些不同分辨率的特征映射,TFE 模块能够更好地捕捉和融合裂缝的细节信息和语义信息。高分辨率特征映射能够捕捉到裂缝的细微特征,确保裂缝的精确定位和识别;低分辨率特征映射提供全局语义信息,帮助模型理解裂缝的整体结构和背景;中等分辨率特征映射在细节和全局信息之间取得平衡,增强模型的综合理解能力。

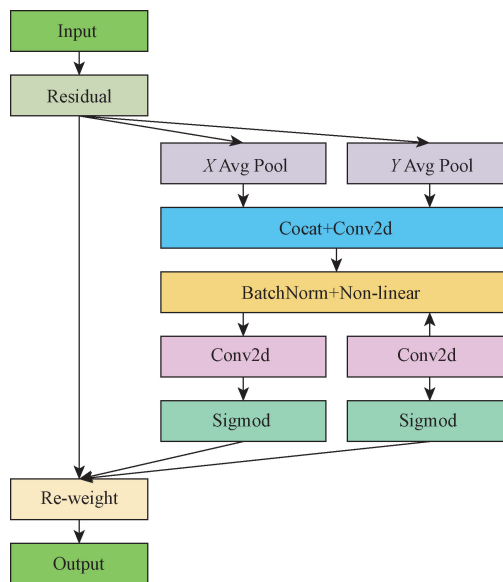
这种策略有助于保持低分辨率图像的局部特征丰富性,以防止小目标特征信息的丢失。最后,对具有相同维度的大、中、小 3 个尺寸的特征图进行一次卷积,然后在通道维度上进行拼接,实现了多尺度特征的有效融合和整合。

$$F_{\text{TFE}} = \text{Concat}(F_1, F_m, F_s) \quad (3)$$

式(3)中: F_{TFE} 为 TFE 模块输出的特征图; F_1 、 F_m 和 F_s 分别为大、中和小尺寸的特征图。 F_{TFE} 由 F_1 、 F_m 和 F_s 拼接得到,具有与 F_m 相同的分辨率,并且通道数是 F_m 的 3 倍。

1.2 CA 注意力机制

由于路面裂缝所处环境的复杂,为了增强模型对路面裂缝的泛化能力和特征表达能力,在原网络 YOLOv8n 的主干网络 SPPF 模块之后以及颈部的 C2f 之后添加 CA 注意力机制,能够有效地提取特征,通过对特征图进行通道加权,增强了有用信息的权重,减少了无关信息的干扰,提高了特征的表达能力和判别性。CA 注意力机制结构如图 2 所示。



Residual 为残差块; X Avg Pool 为 X 维平均池化; Y Avg Pool 为 Y 维平均池化; Concat + Conv2d 为拼接 + 2D 卷积; BatchNorm + Non-linear 为批归一化 + 非线性激活; Sigmoid 为激活函数; Re-weight 为重加权

图 2 CA 注意力机制

Fig. 2 Coordinate attention Mechanism

CA 注意力机制通过两个步骤对通道关系和长距离依赖关系进行编码:坐标信息嵌入和坐标注意力生成。该注意力块捕获沿一个空间方向的长距离依赖关系,并保留另一个空间方向上的精确位置信息。CA 注意力机制生成涉及聚合特征图、应用共享的 1×1 卷积变换函数、分割输出,并利用单独的变换生成注意力权重。CA 注意力机制最终输出是一个经过转换的张量,具有与输入相同大小的增强表示。

此外,CA 注意力机制有助于网络学习特征之间的相关性和重要性,提升网络的性能,包括准确率和鲁棒性,适应不同的输入数据和任务。相对于全局注意力机制,CA 注意力机制在通道维度上操作,适用于资源受限的场景。

1.3 小目标检测层

在实际路面裂缝场景中,裂缝常分布于路面各个角落,由于相机距离裂缝的距离差异较大,在图像中裂缝的尺度大小不一,尤其是图像包含多个裂缝目标时更为明显。为了进一步提升算法对小目标裂缝的检测性能,增加了一个额外的专为小目标设计的检测层 P2。

特征层的不同尺寸携带着不同的信息。尺寸较大的特征层适合捕捉小目标的信息,而尺寸较小的特征层则更适合于大目标的检测。在实际应用中,路面裂缝可视为小目标,因此增加尺寸较大的特征层以进行信息融合是必要的。考虑到原网络的三种尺寸检测头在小目标检测方面表现不佳,因此在增强特征提取网络时引入一个尺寸为 160×160 的特征层 P2 来更好地检测小目标裂缝。

小目标通常在图像中占据较少的像素,因此更容易被忽略或误判。通过专门的 P2 层,YOLO-CD 能

够更敏锐地检测和定位小目标裂缝,提高了小目标裂缝检测的准确性。

1.4 DyHead 动态检测头

YOLOv8 原有的检测头存在几个限制:首先,采用了传统的单尺度预测结构,导致难以处理多尺度目标,忽略了其他尺度特征的贡献。其次,预测头部缺乏考虑上下文信息,导致每个预测位置的预测都是独立进行的,无法形成全局视野。最后,预测头部的参数量较少,限制了其表达能力,难以深度挖掘特征中的空间结构信息。

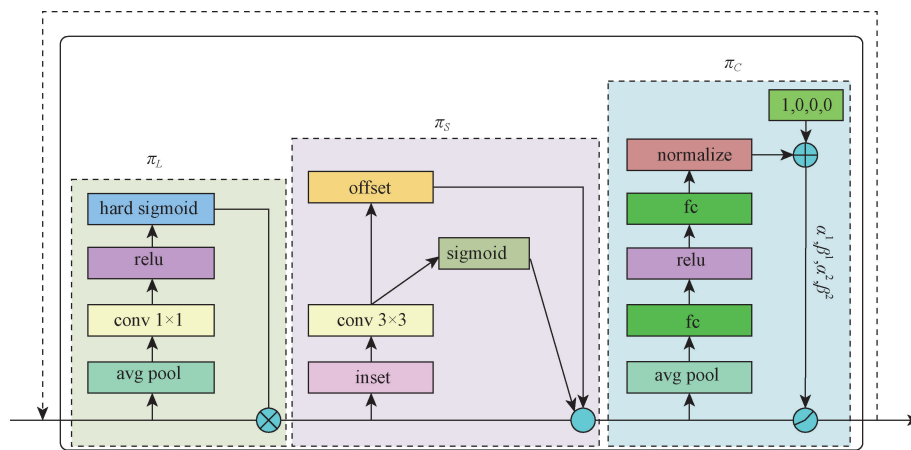
DyHead 动态检测头采用动态路由机制,能够更好地融合上下文信息,实现对多尺度目标的识别。其在特征的不同维度上分别部署注意力机制:逐级、空间和通道。尺度感知注意力模块聚焦于语义层次,空间感知注意力模块增强了空间位置的区分性表示,任务感知注意力模块将不同特征通道指向不同任务。

动态检测头利用了自我注意力机制,实现了尺度感知、空间感知和任务感知的统一。其中,在检测层上给定三维特征张量 $F \in \mathbf{R}^{L \times S \times C}$,该注意力计算公式表达式为

$$W(F) = \pi_c \{ \pi_s [\pi_L(F) \cdot F] \cdot F \} \cdot F \quad (4)$$

式(4)中: F 为一个输入的三维张量 $L \times S \times C$; \cdot 表示矩阵点乘。通过尺度感知、空间感知和任务感知这3个注意力模块 $\pi_L(\cdot)$ 、 $\pi_s(\cdot)$ 和 $\pi_c(\cdot)$,分别作用于给定的三维特征张量 $F \in \mathbf{R}^{L \times S \times C}$ 的 L 、 S 和 C 维度。 $\pi_L(\cdot)$ 、 $\pi_s(\cdot)$ 和 $\pi_c(\cdot)$ 注意力模块以及 DyHead 结构如图3所示。

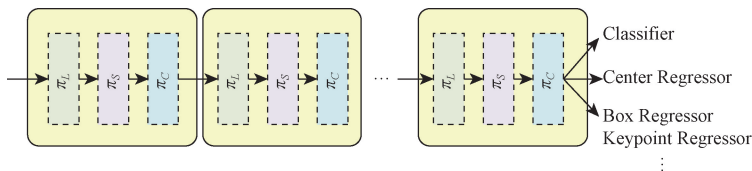
3种注意力模块的内部结构以及它们串联形成的单个 DyHead 结构如图4所示。



hard sigmoid 为激活函数;offset 为偏移操作;inset 为插入操作;fc 为全连接层; normalize 为归一化操作; $[\alpha^1, \beta^1, \alpha^2, \beta^2]^T$ 为学习控制激活阈值的超函数

图3 DyHead 动态检测头

Fig. 3 Dynamic head detection head



classifier 为分类; center regressor 为中心回归器; box regressor 为框回归器; keypoint regressor 为关键点回归器

图4 单个 DyHead

Fig. 4 Single Dynamic head

2 实验与结果

2.1 路面裂缝数据集

在研究中,安装了一台高分辨率摄像头在汽车上,用于采集内蒙古自治区包头市各种路面裂缝的数据。将采集的图像采用十帧取一帧作为路面裂缝数据经过人工挑选,去除没有裂缝和过于模糊的裂缝图像。利用 LabelImg 图像标注工具对初始路面裂缝数据集图像进行标注,制作数据集,将数据集放入算法中进行训练。根据形态特征,将裂缝分为6种类型,分别为纵向裂缝(D00)、横向裂缝(D10)、网状裂缝(D20)、坑洞(D40)、井盖(D50)和补丁缝(D51)。井盖本身并不属于路面裂缝,为了避免井盖和坑洞相互干扰,所以将井盖单独归为一类。

收集了21 597幅路面裂缝图像,构成了目标检测数据集 PD-Dataset。标注了训练集、验证集和测试集的比例设置为8:1:1。使用了17 277张裂缝图像进行训练,其中有2 159张用于验证所提出的裂缝检测方法的性能,并对该方法进行了性能测试。

公共数据集采用 RDD2020,该图像数据集包含从印度、日本和捷克共和国收集的路面裂缝图像,4种裂缝类别的注释分别为:纵向裂缝(D00)、横向裂缝(D10)、网状裂缝(D20)和坑洞(D40)。

2.2 训练参数

实验基于 PyTorch 深度学习框架构建了路面裂缝检测网络。服务器的配置包括 RAM:64GB, GPU: 2 × Nvidia GeForce RTX2080Ti, 22GB, CPU: Intel i9-9900k @ 3.6GHz。训练批次(batchsize)设置为32,并持续进行了长达300个epoch的训练。初始学习率设定为0.01,使用SGD作为优化函数对模型进行训练。

2.3 评价指标

为了全面评估所提出方法的性能,采用了重要的学术评价指标,包括 F_1 、平均精准度(mean average precision, mAP)、画面每秒帧数(frames per second, FPS)、模型参数量(Params)、每秒10亿(10^9)次的浮点运算数(giga floating-point operations per

second, GFLOPs)、权重文件大小。准确率(precision, P)是指在所有被模型预测为正类别的样本中,实际为正类别的比例。召回率(recall, R)是指在所有实际正类别样本中,模型成功预测为正类别的比例。 F_1 综合考虑了准确率和召回率,较全面地反映网络的整体性能。mAP是一种综合评价指标,它结合了不同类别的精确度和召回率,计算方法涉及每个类别的平均精度(average precision, AP),最终得到的是所有类别AP的均值。FPS表示每秒检测到的帧数,Params表示模型的大小,GFLOPs用于度量算法的复杂性。以根据准确率和召回率计算的 F_1 和 mAP 作为检测效果指标。采用 FPS、Params、GFLOPs、权重文件大小,衡量本文算法性能。相关公式如式(5)~式(9)所示。

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$AP = \frac{\sum P}{M} \quad (7)$$

$$mAP = \frac{\sum AP}{N} \quad (8)$$

$$F_1 = 2 \frac{PR}{P + R} \quad (9)$$

式中:TP为检测结果中正确目标的个数;FP为检测结果中错误目标的个数;FN为正确目标中缺失目标的个数; M 为该类别目标总数; N 为类别数;AP为类别平均精度。

2.4 消融实验对比

为了更好地验证改进模块对原模型性能提升的影响,设计了消融实验。实验结果如表1所示。

实验结果表明,随着各个改进模块的加入,如添加小目标检测层 P2、ASF-YOLO 框架、CA 注意力机制模块和 DyHead 动态检测头模块, F_1 分别提升了 0.5%、0.1%、0.3% 和 1.5%, mAP50 分别提升了 1.3%、0.8%、0.4% 和 2.5%。组合使用多个改进模块(P2、ASF-YOLO、CA、DyHead)可以进一步提升 F_1 和 mAP50。

表 1 消融实验结果

Table 1 Ablation experiment results

P2	ASF-YOLO	CA	DyHead	F_1	mAP 50/%	FPS/(帧·s ⁻¹)	Params/10 ⁵	GFLOPs	权重文件/MB
—	—	—	—	0.503	46.1	145.5	30.0	8.1	5.96
√	—	—	—	0.508	47.4	127.4	29.2	12.2	5.93
—	√	—	—	0.504	46.9	135.7	30.4	8.5	6.07
—	—	√	—	0.506	46.5	138.7	30.2	8.2	6.03
—	—	—	√	0.518	48.6	118.0	34.8	9.6	6.90
√	√	—	—	0.519	48.1	118.2	24.8	12.0	5.14
√	—	√	—	0.514	47.6	129.0	29.3	12.2	5.97
√	—	—	√	0.520	49.1	113.5	35.0	19.6	7.07
√	√	√	—	0.520	48.3	150.7	25.0	12.0	5.19
√	√	—	√	0.521	49.2	102.5	31.6	19.4	6.44
√	√	√	√	0.525	50.2	89.9	31.8	19.4	6.49

注:√表示加入该模块;—表示未加入该模块;mAP50表示将交并比(IoU)设为0.5时,计算每一个类别下所有图片的mAP。

改进后的算法 YOLO-CD 模型参数量仅为 31.8×10^5 , 较 YOLOv8n 增加了 6%, 而 F_1 、mAP50 分别为 0.525、50.2%, 较 YOLOv8n 分别增加了 2.2% 和 4.1%。虽然改进后模型相比原模型的 GFLOPs 增加较多, 但 FPS 为 89.9 帧/s, 仍然满足算法检测速度要求。上述改进使得算法在满足实时性检测需求的同时, 精度有了较大提升。

2.5 泛化能力对比

在公共数据集 RDD2020 上, 分别对原算法 YOLOv8n 与改进算法 YOLO-CD 进行训练和测试, 来评估改进算法在实际应用中的性能。将数据集按照 9:1 划分训练集和验证集。如表 2 所示, YOLO-CD 在 4 个类别上的 mAP50、mAP50-95 均优于 YOLOv8n。在 mAP50 方面, 横向裂缝(D10)的检测效果提升最高, 增加了 2.4%; 在 mAP50-95 方面, 坑洞(D40)的检测效果提升最高, 增加了 0.8%。此外, YOLO-CD 整体检测效果比 YOLOv8n 更好, mAP50 和 mAP50-90 分别提高了 1.5% 和 0.5%。这些实验结果验证了改进算法 YOLO-CD 在路面裂缝检测方面的显著提升。

2.6 不同算法对比

为了验证 YOLO-CD 算法在 YOLO 系列中的检测性能, 将其与 Faster RCNN、CenterNet、RetinaNet、YOLOv5n、YOLOv7-tiny、YOLOv5s、YOLOv10n 算法

表 2 RDD2020 数据集检测结果对比

Table 2 The comparison of detection results on the RDD2020 dataset

算法	类别	mAP50/%	mAP50 ~ 95/%
YOLOv8n	D00	47.1	21.2
	D10	46.8	19.3
	D20	69.1	36.3
	D40	53.5	23.9
	平均值	54.1	25.2
YOLO-CD	D00	47.7	21.4
	D10	49.2	20.0
	D20	70.4	37.0
	D40	55.0	24.7
	平均值	55.6	25.7

注:mAP50 ~ 95 表示在 IoU 阈值为 0.5 ~ 0.95 的范围内计算的 mAP。

在公共数据集 RDD2020 进行了训练和测试比较。从表 3 中可以看出, 与 Faster RCNN、CenterNet、RetinaNet、YOLOv5n、YOLOv7-tiny、YOLOv5s、YOLOv10n 这些算法相比, YOLO-CD 在保证运行速度的前提下, F_1 分别提高 18.5%、10.2%、13.4%、3.1%、3.5%、0.9% 和 3.9%, mAP50 分别提高了 18.1%、11.8%、14.3%、5.8%、4.7%、2.2% 和 5.2%。实验结果表明, 与其他目标检测算法相比, YOLO-CD 算法以较小的模型体积实现了极高的检测精度, 具有检测效率高、使用成本低、部署更加简单等优点。

表 3 不同算法对比

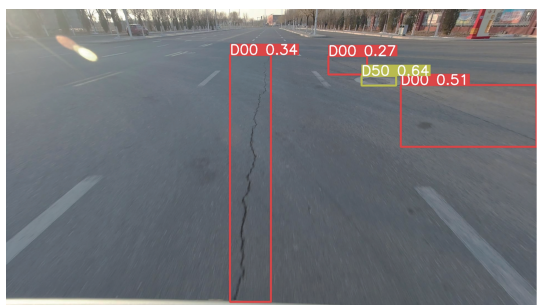
Table 3 The comparison of different algorithms

算法	F_1	mAP50/%	FPS/(帧·s ⁻¹)	Params/10 ⁵	GFLOPs	权重文件/MB
Faster RCNN	0.367	37.5	25.8	1 367.3	369.8	108.0
CenterNet	0.450	43.8	38.2	326.7	69.9	124.0
RetinaNet	0.418	41.3	29.4	379.7	170.1	139.0
YOLOv5n	0.521	49.8	160.2	17.6	4.2	3.75
YOLOv7-tiny	0.517	50.9	145.1	60.1	13.0	11.73
YOLOv5s	0.543	53.4	82.1	70.2	15.8	13.8
YOLOv10n	0.513	50.4	117.6	26.9	8.2	5.53
YOLO-CD	0.552	55.6	89.9	31.8	19.4	6.58

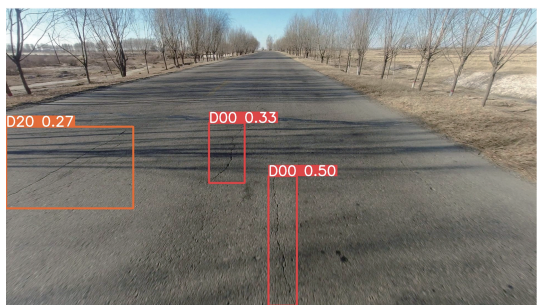
2.7 检测结果与分析

图5、图6为YOLOv8n算法与YOLO-CD算法在自建数据集PD-Dataset上的检测结果。图5(a)未能检测出D10裂缝,但图6(a)能够检测出D10裂缝,没有漏检;图5(b)未能检测出D00裂缝,但图6(b)能够正确检测出所有裂缝,没有漏检;图5(c)出现一处D00裂缝的误检测,但图6(c)没有出现误检,正确检测出了D10裂缝。并且,YOLO-CD检测图均比YOLOv8n检测图置信度高。

图7、图8为YOLOv8n算法与YOLO-CD算法在公共数据集RDD2020上的检测结果。图7(a)和图8(a)对比,YOLOv8n检测不完整,YOLO-CD比YOLOv8n多检测出了一处D10裂缝;图7(b)和图8(b)对比,YOLO-CD比YOLOv8n检测出了更多的D10裂缝;图7(c)和图8(c)对比,两种算法均识别出了D20裂缝,但YOLO-CD比YOLOv8n检测裂缝的置信度高,两个裂缝置信度分别提升0.03和0.04;图7(d)和图8(d)对比,YOLOv8n错误地检测出了D10裂缝,但YOLO-CD没有出现误检,正确检测出D20裂缝。



(a)漏检1

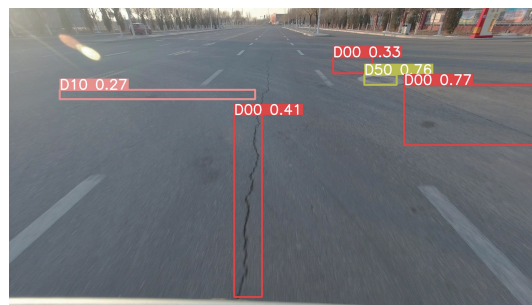


(b)漏检2

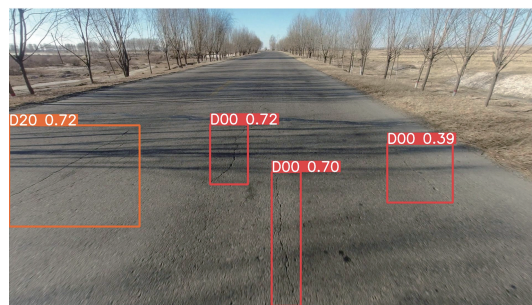


(c)误检

图5 PD-Dataset YOLOv8n 检测
Fig. 5 YOLOv8n detection on PD-Dataset



(a)无漏检1



(b)无漏检2



(c)无误检

图6 PD-Dataset YOLO-CD 检测
Fig. 6 YOLO-CD detection on PD-Dataset



(a)漏检1

(b)漏检2



(c)置信度低

(d)误检

图7 RDD2020 YOLOv8n 检测
Fig. 7 YOLOv8n detection on RDD2020

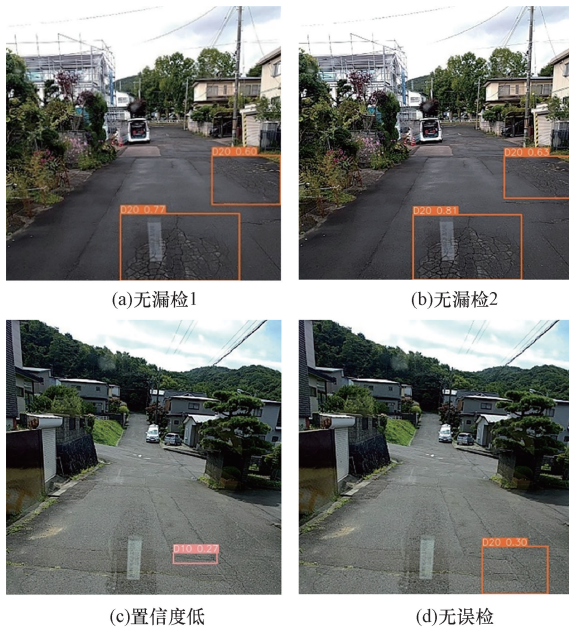


图 8 RDD2020 YOLO-CD 检测
Fig. 8 YOLO-CD Detection on RDD2020

图 5 ~ 图 8 表明, YOLO-CD 算法在自建数据集 PD-Dataset 和公共数据集 RDD2020 上的裂缝检测准确性和置信度均优于 YOLOv8n 算法, 显著减少了漏检和误检的情况。改进后的算法在检测效果上具有更强的鲁棒性和可靠性。这些结果验证了 YOLO-CD 算法在路面裂缝检测中的应用价值和良好的泛化能力。

3 结论

提出了一种改进 YOLOv8n 的路面裂缝检测算法 YOLO-CD。该算法采用 ASF-YOLO 架构中的尺度序列特征融合 (SSFF) 模块和三重特征编码器 (TFE) 模块, 通过多尺度特征融合和多层次特征编码, 提高了对不同尺度和不同层次裂缝的检测能力。在骨干网络末端和颈部网络中添加 CA 注意力机制, 通过坐标信息嵌入和坐标注意力生成, 实现了对输入裂缝图像的特征增强。此外, 网络中添加了小目标检测层, 提高了算法对小目标裂缝的检测能力。将检测头替换为 DyHead, 通过动态卷积核调整和多尺度特征融合, 提高了裂缝检测精度。在自建数据集 PD-Dataset 中, YOLO-CD 相较于原算法 YOLOv8n, mAP50 提高了 4.1%。在公共数据集 RDD2020 中, YOLO-CD 相较于原算法 YOLOv8n, mAP50 提高了 1.5%。且算法检测速度达到 89.9 帧/s, 满足路面裂缝检测实时性的要求。此外, YOLO-CD 算法在检测精度上也优于其他同类算法, 解决了现有路面裂缝检测算法精度低、误检率和漏检率高等问题, 进一步验证了其在路面裂缝检测中的实际应用价值。

所提算法还有诸多方面有待改进, 模型检测速度上还有提升的空间。在实时检测情景下, 如何更进一步地提高模型的检测精度将是下一步的研究方向。

参考文献

- [1] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [C]//Computer Vision-ECCV 2016: 14th European Conference. The Netherlands; ECCV, 2016: 21-37.
- [2] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. Advances in Neural Information Processing Systems, 2015, 28: 1137-1149.
- [3] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas; IEEE, 2016: 779-788.
- [4] Redmon J, Farhadi A. Yolov3: an incremental improvement [J]. ArXiv Preprint ArXiv, 2018: 1804.02767.
- [5] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: optimal speed and accuracy of object detection [J]. ArXiv Preprint ArXiv, 2020: 2004.10934.
- [6] Ge Z, Liu S, Wang F, et al. Yolox: exceeding YOLO series in 2021 [J]. ArXiv Preprint ArXiv, 2021: 2107.08430.
- [7] Ren J, Zhao G, Ma Y, et al. Automatic pavement crack detection fusing attention mechanism [J]. Electronics, 2022, 11(21): 3622.
- [8] 林轶, 陈琳, 王国鹏, 等. 改进的 YOLOv3 交通标志识别算法 [J]. 科学技术与工程, 2022, 22(27): 12030-12037.
Lin Yi, Chen Lin, Wang Guopeng, et al. Improved YOLOv3 algorithm for traffic sign recognition [J]. Science Technology and Engineering, 2022, 22(27): 12030-12037.
- [9] 徐国标, 侯明利, 熊辉. 基于 YOLO 改进算法的远程塔台运动目标检测 [J]. 科学技术与工程, 2019, 19(14): 377-383.
Xu Guobiao, Hou Mingli, Xiong Hui. Detection of moving targets in remote tower based on improved YOLO algorithm [J]. Science Technology and Engineering, 2019, 19(14): 377-383.
- [10] Ren J H, Zhao G W, Ma Y D, et al. Automatic pavement crack detection fusing attention mechanism [J]. Electronics, 2022, 11(21): 3662.
- [11] 李松, 史涛, 井方科. 改进 YOLOv8 的道路损伤检测算法 [J]. 计算机工程与应用, 2023, 59(23): 165-174.
Li Song, Shi Tao, Jing Fangke. Road damage detection algorithm based on improved YOLOv8 [J]. Computer Engineering and Applications, 2023, 59(23): 165-174.
- [12] 史涛, 崔杰, 李松. 优化改进 YOLOv8 实现实时无人机车辆检测的算法 [J]. 计算机工程与应用, 2024, 60(9): 79-89.
Shi Tao, Cui Jie, Li Song. Algorithm optimization of improved YOLOv8 for real-time UAV vehicle detection [J]. Computer Engineering and Applications, 2024, 60(9): 79-89.
- [13] Kang M, Ting C M, Ting F F, et al. ASF-YOLO: a novel YOLO model with attentional scale sequence fusion for cell instance segmentation [J]. ArXiv Preprint ArXiv, 2023: 2312.06458.
- [14] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 13713-13722.
- [15] Dai X, Chen Y, Xiao B, et al. Dynamic head: unifying object detection heads with attentions [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 7373-7382.