

基于边缘特征增强的 YOLOv8n 道路缺陷检测算法^{*}

曹鸿博 安维胜 梁海鹏 林 强

(西南交通大学机械工程学院 成都 610031)

摘 要: 道路病害对交通安全、道路使用寿命及行车的舒适度都有一定影响。针对现有的检测算法对于复杂特征的裂缝、坑洞识别精度较低的问题,本文构建了一种基于增强边缘特征的 YOLOv8n-Edge 道路缺陷检测算法。首先,在骨干网络中引入 RFACnv,以避免卷积核参数共享同时扩大感受野。然后设计具有强化边缘特征的 Edge Enhance Cnv,将高频信息与输入图像融合,强化输入特征。其次,融合 Manet 和 Starnet 提出 Manet-Star 替换部分 C2f 结构,加强模型的特征提取能力。最后,在浅层网络设计一个旁路分支模块—Sub-GEIM,生成具有不同尺度的边缘特征图,与对应尺度的检测头进行融合以优化目标框定位。结果表明,YOLOv8n-Edge 算法显著提高了路面病害的检测效果。虽然带来少量参数量和计算量,但是在预处理的 RDD2022 数据集上关键指标 mAP@50 达到了 72.1%,较原算法提高 3.3%。此外通过泛化实验、对比试验均验证了本文算法的有效性。

关键词: 强化边缘特征;路面病害检测;YOLOv8n;旁路分支

中图分类号: TP391.4;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.6040

YOLOv8n road defect detection algorithm based on
edge feature enhancement

Cao Hongbo An Weisheng Liang Haipeng Lin Qiang

(School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: Road surface defects impact traffic safety, road durability, and driving comfort. To address the low detection accuracy of complex features such as cracks and potholes in existing methods, this paper proposes YOLOv8n-Edge—a road defect detection algorithm based on YOLOv8n with enhanced edge features. RFACnv is integrated into the backbone to enlarge the receptive field while avoiding kernel parameter sharing. An Edge Enhance Cnv module is introduced to fuse high-frequency details with the input, reinforcing feature representation. Additionally, the Manet-Star structure, combining Manet and Starnet, replaces parts of the C2f module to boost feature extraction. A shallow-layer auxiliary branch, Sub-GEIM, generates multi-scale edge features that are fused with corresponding detection heads to improve localization accuracy. Experimental results show that YOLOv8n-Edge achieves a mAP@50 of 72.1% on the preprocessed RDD2022 dataset—an improvement of 3.3% over the baseline—while only slightly increasing model complexity. Its effectiveness is further validated through generalization and comparative experiments.

Keywords: enhanced edge features;road defect detection;YOLOv8n;auxiliary branch

0 引 言

近年来,经济的快速增长引起公路覆盖面积和车流量都大大增加。道路病害问题也日益严重,影响到驾驶的体感且具有潜在的危险性^[1]。随着深度学习的快速发展,极大程度的推动了各个不同领域的智能化发展,其在目标检测领域的意义十分重大,在道路病害检测工作中得到了广泛的运用。

与基于图像阈值的分割方法^[2]和基于边缘检测理论^[3]的传统道路缺陷检测方法不同。基于深度学习的道路病害检测主要有两种不同的实现方式,一种是两阶段的目标检测算法,如 Masked R-CNN^[4]算法。另一种则是单阶段目标检测算法,如 YOLOv8^[5]等目标检测算法。魏陈浩等^[6]为提高模型在道路障碍检测的抗干扰能力,基于 YOLOv8n 网络引入了基于可变卷积的 D_C2f 和双层路由注意力模块,强化特征提取能力,适应多变的真实场景。周

建新^[7]等使用动态蛇形卷积(dynamic snake convolution, dysnake conv)改进 C2f 结构,同时引入多路径聚合的坐标注意力机制模块(multipath coordinate attention, MPCA)在主干网络(backbone)中,以提高模型对细长类型裂缝的捕捉和提取能力。李松等^[8]结合了 CNN 和 Transformer 的优势,提出了 BOT 模块,以适应裂纹对象的大跨度与细长特征,同时引入坐标注意力机制(coordinate-attention, CA)和 C2fGhost 模块,实现模型轻量化和目标准确定位。赵佰亭等^[9]基于 YOLOv8n 网络提出 EDG-YOLO 网络,设计了 C2f-EIR 结构以增强主干网络对井壁裂缝特征的提取能力,在颈部网络(neck)中融合 CSP_EDRAN(CSP efficient dilated reparam aggregation network),促进浅层与深层特征之间的信息交互。叶发茂等^[10]为了充分利用重要的边缘特征信息提出了双骨干的 DB-YOLO_v8n 模型,在原模型基础上增加一个边缘特征骨干网络(edge feature backbone, EFB)最后有机融合不同主干之间同尺度的边缘特征和图像特征。

综上的研究工作中虽然在道路缺陷检测任务中均取得了一定的成果,例如可变卷积结构、注意力机制融合、主干网络增强等多种有效改进策略。然而,这些方法大多聚焦于深层语义特征的建模,着重改进网络结构以加强特征提取能力,从根本上来说还是缺乏对裂缝、坑洞的边缘特征信息的充分学习,导致对浅层结构中关键的边缘与纹理特征的提取与利用关注不足,因此,从此层面着手改进 YOLOv8 模型以提高检测的准确性和速度仍具有很大的空间。

实际的道路环境具有不确定性和复杂性,在实际的路面缺陷的检测任务中,裂缝、坑洞缺陷特征的边缘信息比较复杂、细碎,使得检测模型对于多缺陷类别的准确性、泛化性不高。因此强化边缘特征信息的学习、提升浅层结构对裂纹目标的表达能力,同时加强网络的特征提取能力,最终进行融合,以综合提高网络的综合性能显得十分重要。

为了强化目标特征的边缘信息以学习到更加全面的缺陷特征,从而使得检测的准确性和鲁棒性更高,本文提出了 YOLOv8n-Edge 网络,一种基于边缘特征增强改进的 YOLOv8n 网络结构,与其他主流模型相比,本网络在带来较大的精度提升同时不会造成太大的参数量。主要贡献献如下:

1)在 backbone 网络中引入感受场注意力卷积^[11](receptive-field attention convolution, RFACConv),带来更大的感受野,更好地关注目标区域特征,避免标准卷积提取特征模糊化。

2)设计边缘特征增强卷积(Edge Enhance Conv),通过突出裂缝和坑洞的边缘特征信息,使得浅层网络输出的目标特征边缘信息更加丰富、突出,便于后续网络的特征提取,也更好的促进浅层与深层特征之间的信息交互。

3)设计边缘特征信息旁路分支模块(subordinate

generate edge information model, Sub-GEIM),使用聚焦复杂边缘和高频信息的注意力机制、自适应差分卷积提取梯度和多级池化处理,从浅层的增强特征图中提取出具有裂缝和坑洞丰富轮廓信息的多尺度特征图,与网络处理的不同尺度信息融合,提高目标区域定位。

4)结合 Manet^[12]和 Starnet^[13],提出 Manet-star 改进 backbone 中的部分 C2f 结构,增强网络对增强特征的提取能力,丰富上下文中对于路面缺陷整体的理解。

与 YOLOv8n 相比,预期 YOLOv8n-Edge 网络在道路缺陷检测任务中在各项评价指标上将会有不同程度的提升,对复杂特征的裂缝、坑洞均能达到更好的检测准确性。

1 YOLOv8n-Edge 网络

YOLOv8 网络是由骨干网络、颈部网络和检测头三部分组成。为了提高网络学习路面缺陷特征的边缘信息,提出 YOLOv8n-Edge 算法,设计 Edge Enhance Conv,以强化输入特征的高频信息;同时引入 RFACConv,扩大感受野,更好的关注目标区域特征;在骨干网络中提出 Manet-Star 替代部分 C2f 结构,加强特征提取的能力;最后在浅层网络中设计分支网络 Sub-GEIM,输出不同尺度的边缘信息以辅助检测头进行精准定位,提高模型的准确性,减少漏检、误检率。YOLOv8n-Edge 结构如图 1 所示。

2 基于 YOLOv8n 的改进方法

2.1 感受场注意力卷积—RFACConv

在 YOLOv8 网络中使用标准卷积提取特征信息,由于卷积核的固定性,使得无法捕捉图像不同位置的特征信息差异,导致对于路面缺陷的边缘特征容易模糊化,而影响了网络的性能。因此在主干网络中引入 RFACConv,这是一种旨在增强卷积神经网络感受野的卷积运算,其中通过分组卷积为每个感受野生成独立的注意力权重,实现参数不共享的卷积操作,提高对不同尺寸目标的适应能力。RFACConv 流程图如图 2 所示。

RFACConv 采用双分支设计,将输入特征分别通过注意力分支和感受野特征分支,得到对应的空间注意力权重和感受野特征图,然后使用乘积操作根据空间注意力图权重提取感受野空间特征生成空间特征图。各分支流程表示如下:

$$X_1 = \text{Softmax}(\text{GroupConv}_1(\text{AvgPool}(X))) \quad (1)$$

$$X_{rf} = \text{ReLU}(\text{Norm}(\text{GroupConv}_3(X))) \quad (2)$$

$$X_{out} = \text{Conv}_{k=3, s=3}(X_1 \otimes X_{rf}) \quad (3)$$

其中, X 表示输入特征图, GroupConv_n 表示卷积核为 n 的组卷积, X_{rf} 表示感受野特征分支输出特征图, X_1 表示注意力分支输出特征图, X_{out} 表示 RFACConv 输出特征图, \otimes 表示逐元素乘法操作, $\text{Conv}_{k=3, s=3}$ 表示卷积核为 3、步长为 3 的标准卷积操作。

在注意力分支中,首先使用平均池化操作聚合全局特

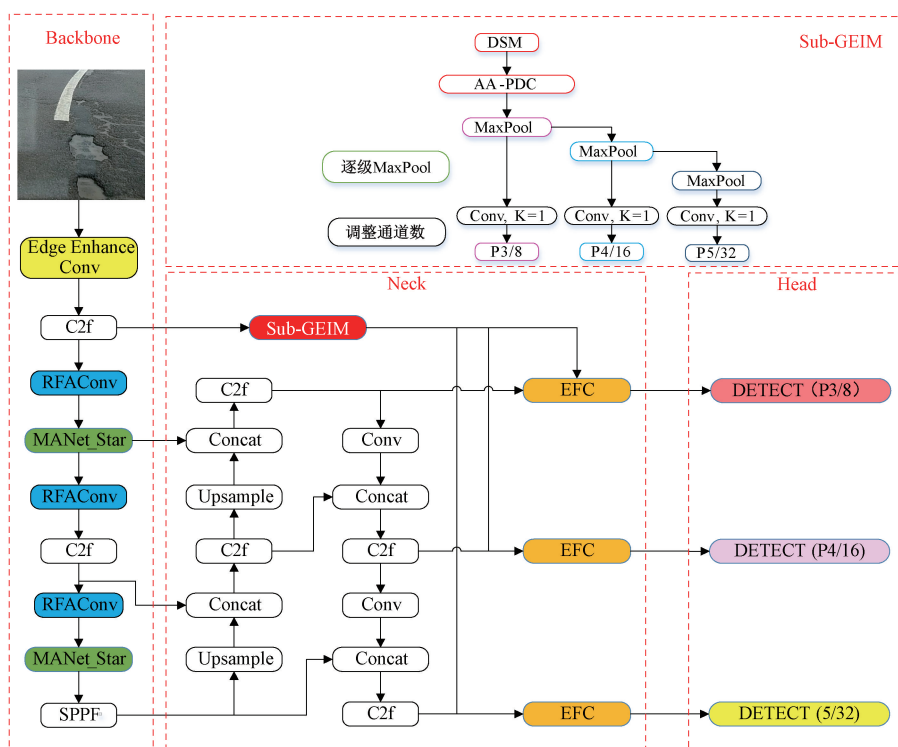


图 1 YOLOv8n-Edge 网络

Fig. 1 YOLOv8n-Edge network

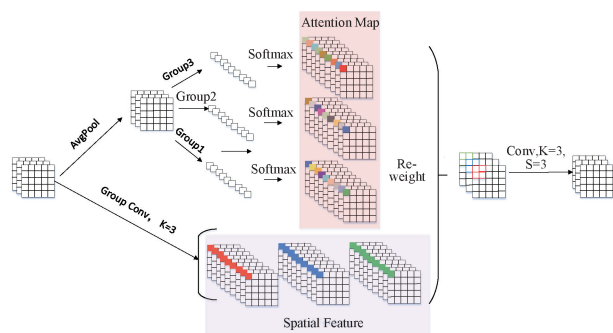


图 2 RFACnv 结构

Fig. 2 Structure of RFACnv

征,再使用多个组卷积进行信息交互,且使用 Softmax 操作,以生成不同权重的空间注意力矩阵图,并且按照权重大小进行优先排序,特征的权重越大表示更重要,以确保后续特征提取具有包含更多裂缝、坑洞病害信息的图像。

在感受野空间分支中,使用 $k=3$ 的分组卷积提取特征,然后使用归一化和 RELU 激活函数生成感受野空间特征图,输出特征图尺度和维度特征与注意力图一致。最后根据空间注意力图权重提取感受野空间特征,尺度调整输出特征更突出的结果。因此使用 RFACnv 代替标准卷积,有助于提高全局特征表达能力。

2.2 Edge Enhance Conv

由于原始图像含有大量的背景信息和干扰信息,对此

YOLOv8 网络中使用了两层下采样结构筛选了部分背景和干扰信息,但是会使得目标区域边缘信息模糊化,丢失细节信息。对此,设计了边缘信息增强模块—Edge Enhance Conv 替换原网络结构中的浅层下采样卷积结构,以增强目标区域的边缘特征,避免单纯的降采样导致特征信息损失。

首先引入 RFACnv 注意力卷积扩大感受野,之后通过设计的 Edge-Enhancer 模块增强边缘对比度,以强化特征的边缘信息,然后通过 SobelConv 分支和 MaxPool 池化分支,分别进行边缘梯度提取和特征提取,最后将各分支信息进行融合、调整通道后输出。Edge Enhance Conv 结构如图 3 所示。

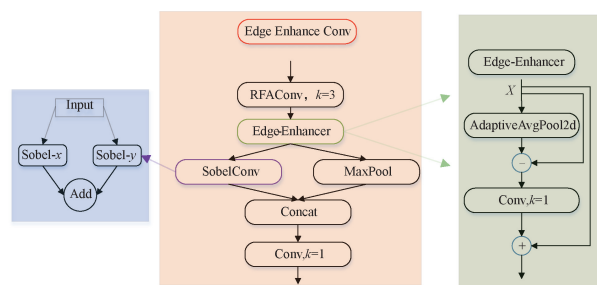


图 3 Edge Enhance Conv 结构

Fig. 3 Structure of Edge Enhance Conv

1) RFACnv 注意力卷积:通过设计感受野分支和注意力分支,以增强神经网络的感受野,筛选干扰信息强化

主要特征。

2) Edge-Enhancer: RFAConv 筛选部分干扰背景, 得到输出特征 X , 通过平均池化处理以平滑特征得到低频信息, 之后原始特征图与池化结果比对相减, 得到具有边缘特征的高频信息, 然后使用卷积对高频边缘信息进行强化突出, 最后与输入特征进行相加输出, 从而使得输出特征具有更丰富的高频信息, Edge-Enhancer 结构如图 3 所示, 具体流程如下:

$$F_0 = X + (\text{Conv}(X - \text{AvgPool}(X))) \quad (4)$$

3) SobelConv 分支与 MaxPool 分支: 通过 SobelConv 对高频信息计算水平梯度和垂直梯度, 梯度变化较大的区域则是可能的物体边界, 对其进行捕捉, 使用 MaxPool 分支提取特征, 更好的保留细节特征。梯度计算如下:

$$G_x = I \otimes W_x, G_y = I \otimes W_y \quad (5)$$

其中, W_x, W_y 为 Sobel 卷积核, I 为输入图像, G_x 表示水平方向边缘响应图, G_y 表示垂直方向边缘响应图。

$$\text{计算梯度幅值: } G = \sqrt{G_x^2 + G_y^2}.$$

4) 对分支输出特征拼接和通道、尺度的变换调整。

因此通过输入图像通过 Edge Enhance Conv 模块的处理, 不仅能够有效的减少背景干扰、调整图像尺寸, 通过其高频信息的突出以及梯度计算, 输出具有更明显的裂缝、坑洞缺陷边缘特征图, 从而为后续的深层网络提取特征提供更具有区分度的输入信息。

2.3 Sub-GEIM

在目标检测任务中, 目标框的定位十分依赖目标边缘特征, 对此 YOLOv8 采用卷积神经网络逐层提取图像特征, 然后通过不同模块完成目标分类和边界框回归, 偏向于获取整体的轮廓和纹理信息, 对于边缘信息关注度较少。

为了增加边缘特征的关注度, 设计了 Sub-GEIM 模块, 由于浅层网络使用了卷积层过滤了较多不相关的背景信息, 更加关注目标区域的边缘和纹理。因此从浅层网络中提取出边缘信息的分支模块, 获取轮廓信息, 且生成多尺度的边缘特征, 将提取的边缘信息与颈部网络进行深度融合, 提高对目标边缘的感知能力, 使目标定位更加精准。

Sub-GEIM 模块的原理如下: 首先引入聚焦复杂边缘和高频信息的注意力机制 (dual domain selection mechanism, DSM)^[14] 处理输入的浅层特征, 然后设计自适应像素差分卷积 (angle adaptive pixel difference convolution, AA-PDC), 通过重排卷积核中的元素, 以具有计算不同角度、方向梯度和差值, 便于区域的特征进行捕捉, 而突出边缘特征; 最后使用最大池化处理层进行逐级下采样, 生成具有不同尺度的边缘信息, 采用 1×1 卷积逐层调整通道数以及特征信息的映射, 同时结合 Batch Normalization (BN) 和 SiLU 激活函数, 便于生成的边缘信息与深层网络通过边缘融合卷积 (edge fusion conv, EFC) 融合。Sub-GEIM 结构如图 4 所示。

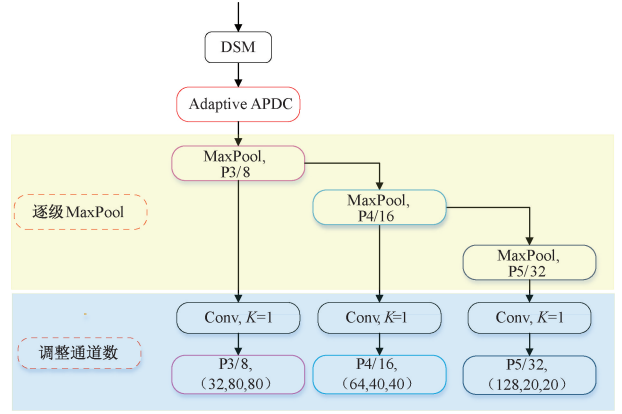


图 4 Sub-GEIM 结构

Fig. 4 Structure of Sub-GEIM

1) DSM 注意力机制: 由空间选择模块 (spatial selection module, SSM) 和频率选择模块 (frequency selection module, FSM) 组成。SSM 模块用于关注输入特征 X 的重要区域, 采用三支设计, 主分支将输入特征依次使用 MaxPool、AvgPool 进行特征压缩, 使用标准卷积生成特征图 X' 后再拓展成与原特征图尺度一致, 主分支提取结果分别与其他分支使用深度卷积提取的特征通过逐元素乘法和加法融合, 处理结果输入到 FSM 模块中。FSM 模块强调包含输入清晰图像对之间真正差异的区域, 通过均值滤波器来生成输入特征低频特征, 然后使用输入特征减去低频特征即可得到高频特征, 联合输入特征运用主元素乘法和残差链接得到注重边缘区域重要信息的输出特征图 \hat{X} 。DSM 结构如图 5 所示, 具体表示如式 (6)、(7)、(8):

DSM 阶段:

$$X' = \text{Conv}_3([\text{AvgPool}(X), \text{MaxPool}(X)]) \quad (6)$$

$$X_1 = D\text{Conv}_{5,7}(X) \otimes T(X', C) + D\text{Conv}_3(X) \quad (7)$$

FSM 阶段:

$$\hat{X} = (X_1 - \text{mean}(X_1)) \otimes X_1 + X_1 \quad (8)$$

其中 X 表示输入特征图, \hat{X} 为输出特征图, X_1 为中间特征图, $\text{Conv}_n, D\text{Conv}_n$ 分别表示卷积核为 n 的标准卷积、深度卷积, $T(X', C)$ 表示 \hat{X} 特征沿着 C 维度进行 Tile 操作, AvgPool、MaxPool 为池化处理。

2) 角度自适应像素差分卷积 (AA-PDC): 由于待检测的裂缝、坑洞边缘均具有复杂、不规则的边缘特征, 所以为了能够提取到不同角度的梯度信息, 故采用对卷积核内的像素点位置随机进行打乱的策略对卷积核进行变换以模拟角度差分, 通过随机产生的不同方向进行梯度计算。

如式 (9)、(10), 卷积核为 3 的 ad_weight 像素点随机打乱之后得到 $disturb_ad_weights$, 逐像素点相减得到新卷积核, 可以发现新卷积核中 $X_1 - X_3$ 像素点位置等可以计算水平梯度, $X_2 - X_5$ 像素点位置等可以计算垂直梯度,

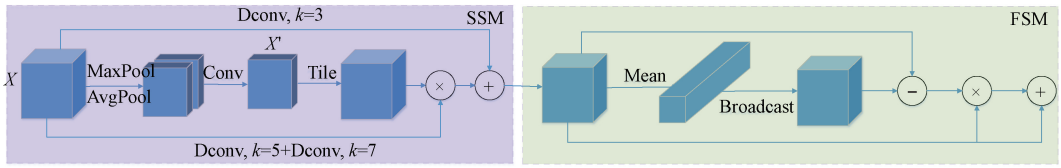


图5 DSM结构

Fig. 5 Structure of DSM

$X_6 - X_2$ 则可以计算不同角度的梯度,生成每个位置边缘响应强度,值越大表示边缘特征越明显.并且每个通道都进行权重学习,与具有方向、角度的边缘强度图逐像素乘积,确保不同通道的梯度信息能被适当放大或抑制。

$ad_weight(k=3)$ disturb_ad_weights($k=3$)

$$\begin{bmatrix} X_1 & X_2 & X_3 \\ X_4 & X_5 & X_6 \\ X_7 & X_8 & X_9 \end{bmatrix}$$

$ad_weights[:, :, [3, 5, 1, 6, 4, 2, 9, 7, 8]]$

$$\begin{bmatrix} X_3 & X_5 & X_1 \\ X_6 & X_4 & X_2 \\ X_9 & X_7 & X_8 \end{bmatrix}$$

(9)

$ad_weight(k=3) - disturb_ad_weight(k=3) =$

$$\begin{bmatrix} X_1 - X_3 & X_2 - X_5 & X_3 - X_1 \\ X_4 - X_6 & X_5 - X_4 & X_6 - X_2 \\ X_7 - X_9 & X_8 - X_7 & X_9 - X_8 \end{bmatrix}$$

(10)

因此角度自适应像素差分卷积综合各种不同的方向和角度计算的梯度,相较于 Sobel、Laplacian 算子等更能适应坑洞、裂缝缺陷特征的边缘特征检测。

3) MaxPool 逐级下采样及通道调整:通过两次的逐级下采样,逐步降低分辨率,形成 80×80 、 40×20 、 20×20 三种不同尺度的边缘特征图,对应 P3/8、P4/16、P5/32 三中不同尺度的特征层;之后调整不同尺度的特征层的通道数,结合 Batch Normalization(BN)和 SiLU 激活函数,以便于生成的边缘信息与深层网络进行融合。

4) 边缘融合卷积 EFC:将具有相同特征的边缘信息和网络输出特征信息逐像素相加,再使用卷积变换通道与特征提取,输出与连接网络匹配的通道。EFC 结构如图 6 所示。

通过 Sub-GEIM 模块,有利于强化且突出重要区域的高频边缘信息,同时生成具有不同尺度的边缘特征,针对坑洞、裂缝的复杂检测边缘特征进行细化检测,提高对小目标和细节的检测效果,对后续与网络特征进行融合,提高目标检测任务中边界框的定位精度,从而提升模型的检测性能。

2.4 Manet-Star

在浅层网络中 Edge Enhance Conv 模块增强了输入特征的边缘信息,为了加强网络对增强特征的提取,结合 Manet 与 Starblocks,提出 Manet-Star,以代替骨干网络中

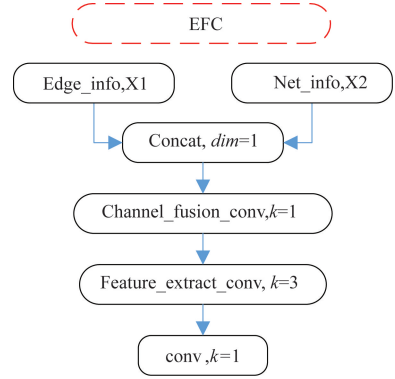


图6 EFC结构

Fig. 6 Structure of EFC

的部分 C2f 结构,以丰富梯度流,加强特征提取能力。Manet-Star 采用混合聚合机制,融合了 1×1 旁路卷积分支、DSConv(深度可分离卷积分支)和 C2f-Star 模块,充分利用它们各自的优势,提升基础特征在关键阶段的语义深度,并优化计算效率。其结构如图 7 所示。

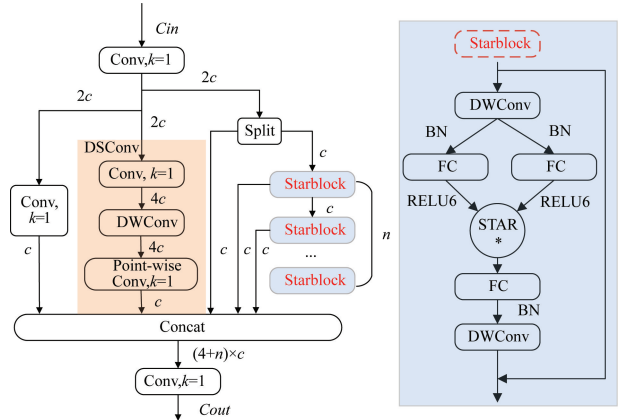


图7 Manet-Star结构

Fig. 7 Structure of Manet-Star

在 Manet-Star 模块中,输入特征图 X 首先通过 1×1 卷积扩展通道数至 $2c$,然后使用 3 条不同的分支路径进行特征处理:

1) 1×1 旁路卷积分支:直接映射输入特征,使用 1×1 卷积调整通道数,为特征融合提供丰富的梯度流信息。

2) DSConv 分支:使用 1×1 卷积调整通道数,使用 $k=3$ 的深度可分离卷积提取特征信息,最后通过 1×1 逐

点卷积压缩通道为 c 。流程如式(11):

$$X_2 = Conv_1(Conv_3(Conv_1(X))) \tag{11}$$

3)C2f-Star 模块:使用 Starblocks 代替 C2f 中堆叠的 Bottleneck,由于 Starblocks 中使用的 Star Operation,一种类似于在不同通道上进行成对特征乘法的核函数,元素乘法操作使得在神经网络中仅需对 Starblocks 进行少量的几次堆叠,即可实现隐式维度复杂性的指数式增长,从而减少了 C2f 结构的参数量,降低成本需求。Manet-Star 结构如图 7 所示。

在骨干网络中通过使用有效的结合了梯度流丰富的 Manet 结构和有助于较少参数数量的 Starblocks 形成 Manet-Star,以加强网络的对于裂缝和坑洞的特征提取能力,便于后续深层网络的特征学习。

3 实验结果与分析

3.1 实验平台及环境配置

本次实验训练的环境采用的 win10 操作系统,Python=3.8, PyTorch = 1.3.1, TorchVision = 0.14.1, CPU: GeForceRTX 3060;CUDA:11.6;image_size:640×640;学习率:0.001;批量大小:8;SGD 优化器;epoch:120。

3.2 数据集

本实验选取了 RDD2022^[15]和 RDD2018^[16]道路缺陷检测公开数据集,由于原始数据集中不同国家采集的数据集存在较大差异,故对其进行筛选,排除部分质量较差、路况不符、较为模糊的图片。为了增强模型对坑洞类别识别的抗干扰能力,引入 RDD2018 中的日本道路数据集代替 RDD2022 数据集中的日本数据集。最后为了补充坑洞缺陷数据集的样本数量不足,自建数据集进行补充,最终得到 11 000 余张图像作为本文训练的数据及内容。聚焦于五种路面缺陷类型:横向裂缝、纵向裂缝、龟裂、坑洞、井盖。

王宇航^[17]指出含有较少天空背景的数据集训练效果比含有较多天空背景的数据集进行训练的效果要更好,因此对于整理的 11 000 余张数据集进行预处理:首先读取图片及其对应的标签文件,以确定框选区域的范围,通过比较各个不同框选区域的坐标值,确定出图像中涵盖缺陷特征区域的框选区域,通过适当余量偏移以确保边缘信息的保存,得到扩充框,对其进行裁剪,得到含有较少天空背景的数据集图像,最后按照 8:1:1 的比例进行数据集的划分以训练集、验证集和测试集的划分。

3.3 评价指标

为了对模型进行评价,本实验以本次实验以精确率(precision,P)、召回率(recall,R)、平均精度均值(mAP@50、mAP@50-95)、总浮点运算量(GFLOPs)和参数量作为评价 YOLO 系列模型性能的参考指标。

3.4 Manet-Star 不同位置对网络性能的影响

使用 Manet-Star 用于代替 YOLOv8 网络中的 C2f 结

构以丰富梯度流信息,能够有效的增强特征提取能力,但是会带来一定参数量和计算量,因此为了验证并选择最适合的网络位置进行替换,对 backbone 和 neck 网络结构中的 C2f 结构都进行替换实验,从而选出最有效的替换位置,这有助于提升网络的综合性能,且不带来太多额外参数量和计算量。具体对比试验如表 1 所示,其中 B 表示骨干网络,N 表示颈部网络,对应数字表示不同通道维度的对应 C2f 结构替换,分别进行了 8 个不同位置的 C2f 结构替换实验。

表 1 Manet-Star 不同位置实验对比
Table 1 Comparison of Manet-Star at different positions

模型	mAP50/ %	mAP50-95/ %	Params/ 10 ⁶	GFLOPs
基线	68.8	40.5	3.07	8.2
B-128	69.4	40.8	3.08	8.4
B-256	70.0	41.2	3.08	8.3
B-512	69.4	40.7	3.09	8.2
B-1024	70.4	41.7	3.23	8.2
N-1-512	68.5	40.2	3.13	8.4
N-256	68.6	40.4	3.08	8.2
N-2-512	68.6	40.3	3.13	8.4
N-1024	68.9	40.7	3.30	8.4

从表 1 中可以得出:在不同位置使用 Manet-Star 替换 C2f 结构,综合的评价指标结果均会有不同的改变,其中在骨干网络替换的效果明显强于在颈部网络中的替换效果,故不考虑使用 Manet-Star 结构在颈部网络中的替换使用。

在骨干网络中,替换 B-256 处的 C2f,不带来参数量和计算量增加,而且模型的 mAP@50 提高 1.5%,mAP@50-95 提高 1%;替换 B-1024 处的 C2f,带来了部分参数量,但是整体的精度上升较多 mAP@50 提高 1.9%,mAP@50-95 提高 1.5%。相较于骨干网络中其他 C2f 结构的替换具有更好的效果,因此 Manet-Star 替换骨干网络中通道维度为 256 和 1024C2f 结构最合适,不仅加强网络的特征提取能力,同时不会增加太多的参数量和计算量。

3.5 消融实验

为了验证提出的模块对网络改进的有效性,基于 YOLOv8 的模型框架,设计了消融实验。所有实验均在相同的训练环境、训练策略和超参数下进行。主要构建了以下模型,包括 YOLOv8n 基础模型,以及使用不同改进策略的改进模型:YOLOv8n-E(使用 Edge Enhance Conv 模块替换浅层卷积层)、YOLOv8n-R(使用 RFACConv 替换骨干网络中标准卷积结构)、YOLOv8n-M(在 backbone 网络中处理通道维度为 128 的输出端添加 Sub-GEIM 分支模块与 neck 网络融合)、YOLOv8n-MS(使用 Manet-Star 替换 backbone 中通道数为 256 和 1 024 的 C2f 模块),以及组合

这些模块形成的各种算法 (YOLOv8n-EM、YOLOv8n-EMS、YOLOv8n-RM、YOLOv8n-RMS、YOLOv8n-MMS、YOLOv8n-EMMS、YOLOv8n-RMMS、YOLOv8n-Edge), 具体实验结果如表 2 所示。

表 2 消融实验对比
Table 2 Ablation experiment comparison

模型	mAP50/ %	mAP50-95/ %	Params/10 ⁶	GFLOPs	P/ %	R/ %
YOLOv8n	68.8	40.5	3.07	8.2	70.1	61.7
YOLOv8n-E	69.9	41.4	3.05	8.4	72.8	62.3
YOLOv8n-R	69.7	41.0	3.10	8.3	70.8	62.1
YOLOv8n-M	70.3	41.3	3.34	8.8	72.0	63.0
YOLOv8n-MS	70.7	42.0	3.20	8.4	73.1	63.4
YOLOv8n-EM	70.5	41.9	3.30	8.9	72.8	62.7
YOLOv8n-EMS	71.0	41.5	3.20	8.4	72.5	62.9
YOLOv8n-RM	69.9	41.6	3.35	8.8	72.9	62.6
YOLOv8n-RMS	70.5	41.7	3.22	8.4	72.7	64.0
YOLOv8n-MMS	71.3	42.0	3.48	8.9	72.3	65.3
YOLOv8n-EMMS	71.7	42.4	3.46	9.0	72.4	65.5
YOLOv8n-RMMS	71.5	42.6	3.50	9.0	72.2	65.1
YOLOv8n-Edge	72.1	43.0	3.50	9.1	72.0	65.7

表 2 模型实验结果的对比中,各项评价指标中使用加深标记标注最优项,实验结果显示出了本文基于 YOLOv8n 提出的 YOLOv8n-Edge 模型具有优秀的综合性能,消融实验中所展示出的各个模块的改进均对模型的性能有一定的提升,使用 Edge Enhance Conv 和 RFACnv 改进模型不会带来额外的计算量和参数量,而且在 mAP@50 指标上都有一定的提高。分别达到 1.1%、0.9%;使用 Manet-Star 替换 backbone 中通道数为 256 和 1024 的 C2f 模块,虽然会带来少量的参数量,但是在 mAP@50、mAP@50-95、精确率、召回率上都能达到非常优秀的指标,分别提升 1.9%、1.5%、3.0%、1.7%。同样,本文中创新性提出的 Sub-GEIM 分支模块设计也帮助网络在上述指标中分别提升了 1.5%、0.8%、1.9%、1.3%;

然后组合不同的两个模块以改进网络,可以发现结合 Manet-Star 和 Sub-GEIM 分支设计的 YOLOv8n-MMS 网络也在这些评价指标上的表现最为优异,分别提升了 2.5%、1.5%、2.2%、3.6%。然后基于 YOLOv8n-MMS 网络的框架,加以 RFACnv 和 Edge Enhance Conv 的改进,都对模型的综合性能具有一定的提升作用。最后组合各个有效的方案对 YOLOv8n 全面改进提出的 YOLOv8n-Edge 分别在各项评价指标提升了 3.3%、2.5%、1.9%、4%,较 YOLOv8n 原模型的检测结果具有较大提升。

上表的消融实验结果极有利的验证了本文中提出的 YOLOv8n-Edge 算法在增强模型检测方面的综合性能显著增加,说明各模块是有效的,且各模块之间有较好的相互联系和影响,虽然带来了一定的参数量和计算量,但是在精度、精确率、召回率上的提升巨大,依然能够满足实时

性的需求。

3.6 图像可视化

为了直观的体现 YOLOv8n-Edge 与 YOLOv8n 算法在道路缺陷检测任务中的检测结果的差异,从验证集中选取了 3 张不同的图像用于可视化结果的展示。其中通过 GradCAMPlusPlus 对改进前后的模型在路面缺陷检测任务中的表现进行了热力图可视化,并将改进模型与原模型在相同图像的检测结果进行了检测结果对比,具体可视化结果如图 8 所示。

图 8 中第 1、2 行的可视化结果显示出,在目标数量单一的检测任务中 YOLOv8n-Edge 模型更集中于捕获目标缺陷的特征信息;第 3 行的可视化结果则显示出,在多目标数量多元的检测任务中,YOLOv8n-Edge 注重全局检测效果,有效的减少了原模型存在的漏检,误检问题,检测结果更准确。综合的可视化结果反映出,YOLOv8n-Edge 能够更加准确的捕捉目标的大致位置,映证了旁路分支以及主干特征网络信息提取的有效性,较 YOLOv8n 模型有更强的检测能力,更加适用于真实场景的检测。

3.7 不同模型性能对比

为了验证本文提出的 YOLOv8n-Edge 算法的优越性,对比了主流算法和基于 YOLOv8 改进的代表性模型。主流的目标检测算法主要有:YOLOv5n^[18]、YOLOv10n、YOLOv10s^[19]、YOLOv11n^[20]。代表性改进模型:赵佰亭等^[9]提出的 EDG-YOLO、叶发茂等^[10]提出的 DB-YOLO_v8n,评价指标、参数设置等与 YOLOv8n 实验保持一致。实验结果如表 3 所示。

由表中的实验结果可见,本文中提出的 YOLOv8n-Edge

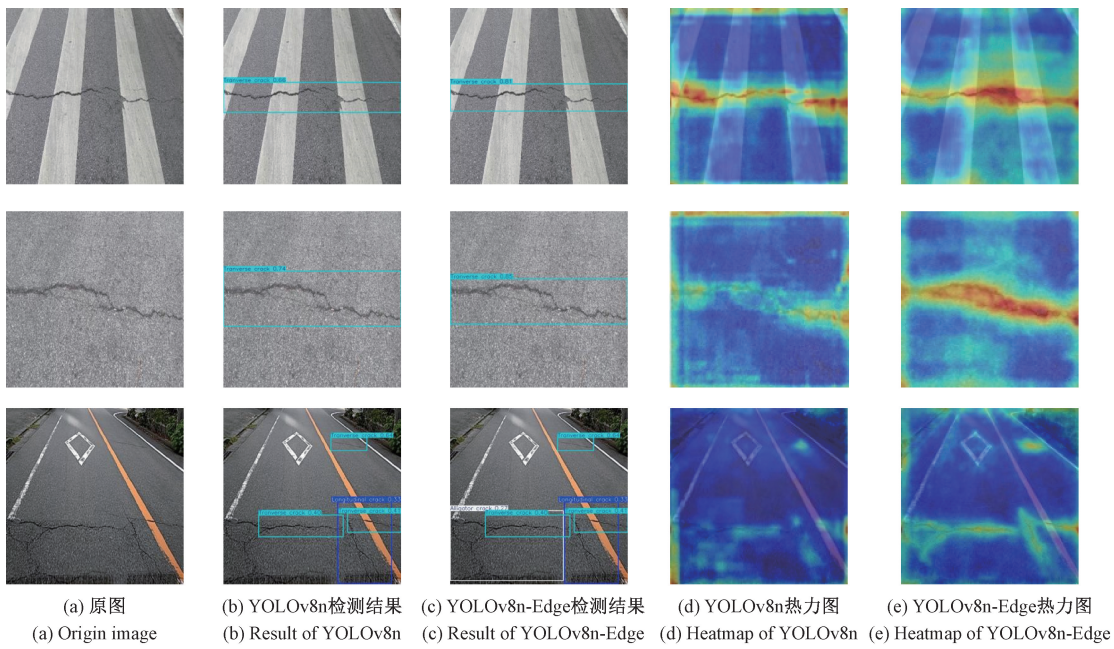


图 8 可视化结果

Fig. 8 Visualization results

表 3 不同模型对比实验

Table 3 Comparison experiments with different models

模型	mAP50/%	mAP50-95/%	Params/ 10^6	GFLOPs	P/%	R/%
YOLOv5n	66.8	38.5	2.50	7.1	69.0	61.0
YOLOv8n	68.8	40.5	3.07	8.2	70.1	61.7
YOLOv10n	67.5	39.7	2.26	6.5	69.0	60.1
YOLOv10s	71.0	42.6	7.22	21.4	72.0	63.3
YOLOv11n	68.7	40.5	2.60	6.3	70.6	62.5
DB-YOLO_v8n	71.7	41.6	3.43	10.8	71.0	64.1
EDG-YOLO	69.5	41.0	2.10	4.3	70.5	62.1
YOLOv8n-Edge	72.1	43.0	3.50	9.1	72.0	65.7

算法的表现明显优于主流的 YOLO 检测模型和代表性改进模型。虽然旁路分支 Sub-GEIM 的设计会使得参数量、计算量略有增加,但是其充分提取边缘特征信息的作用十分明显,因此在各个评价指标上显示出来的优秀检测性能是远高于其他模型的。而对比参数量较大的 YOLOv10s 模型,本文提出的算法具有较轻量效果,同时具有较好的检测性能,因此具有一定的随机性的模型优越性。

与 DB-YOLO_v8n 模型相比,YOLOv8n-Edge 的各方面的性能更优异,这就说明相较于采取双骨干网络提取边缘特征信息的方法,采用旁路分支的设计以提取边缘特征信息的方式不仅避免了更加复杂的模型计算,而且检测的结果也更加的准确、高效。

此外,相较于基于 YOLOv8n 轻量化设计的 EDG-YOLO 模型,本文提出的算法在各项 mAP、Precision、Recall 指标上均显示出明显优势。这进一步验证了将 Manet 与

Starblocks 结合构建的 Manet-Star 结构替换 C2f 模块,能有效提升特征提取的准确性。同时,也体现 Sub-GEIM 旁路边缘分支在多尺度特征融合中的作用显著。尽管整体参数量与计算量有所增加,但模型仍展现出较强的综合性能。

3.8 泛化性实验

为了检验算法的泛化能力,对 SVRDD^[21]数据集按照处理 RDD2022 数据集的方法同样的进行裁剪操作等预处理,以减少数据集背景的占比。使用 YOLOv8n 和 YOLOv8n-Edge 对上述数据集进行对比试验,具体结果如表 4 所示。

从表中的实验结果得出,在整理的 SVRDD 数据集上本文提出的算法也具有较好的表现,较 YOLOv8n 算法在 mAP50 和 mAP50-95 两个关键指标上分别提升 1.1% 和 1.6%,充分的证明了 YOLOv8n-Edge 的有效性及其较好的泛化能力。

表 4 泛化性实验

Table 4 Generalization experiment

模型	mAP50/%	mAP50-95/%	Params/ 10^6	GFLOPs	P/%	R/%
YOLOv8n	75.0	47.2	3.07	8.2	76.8	66.6
YOLOv8n-Edge	76.1	48.8	3.51	9.1	77.1	67.1

4 结 论

为了提高检测网络对边缘信息的重视,强化边缘特征,减少干扰因素的影响,提高模型的检测精度,本文提出了 YOLOv8n-Edge 检测算法,并在进行预处理的 RDD2022 和 SVRDD 数据集上训练和验证。首先,在骨干网络中引入 RFACnv,以避免卷积核参数共享同时扩大感受野。其次,通过将高频信息与输入图像融合,设计具有强化边缘特征的 EdgeEnhanceCnv 下采样模块,替换浅层网络中的浅层标准下采样模块,强化采样层输入特征。为了深层网络更好学习增强特征,融合 MANet 和 StarNet,提出 Manet-Star 替换部分 C2f 结构,加强模型的特征提取能力,建立上下文的联系。最后,在浅层网络设计一个旁路分支模块 Sub-GEIM,通过特征增强、差分卷积多方向梯度计算、逐级池化处理的操作生成具有不同尺度的边缘特征图,与对应尺度的检测头进行融合以优化目标框的定位。实验数据充分的显示出了本算法的优良性能,在各项评价指标上均有较大的提升,在 mAP@50、mAP@50-95、精确率、召回率上分别提升了 3.3%、2.5%、1.9%、4%。同时也验证了本模型具有较好的泛化性。实验结果也与预期的结果基本一致,对模型的准确性以及各项指标均有一定的提升。下一步研究可以在基于强化边缘信息的网络设计的基础上,进一步对模型进行轻量化的处理,以提高模型的实用性价值。

参考文献

[1] 郭一丹. 基于改进深度学习模型的分阶段路面裂缝检测方法研究[D]. 西安:长安大学,2024.

GUO Y D. A study on stage-wise pavement crack detection method based on improved deep learning models[D]. Xi'an:Chang'an University, 2024.

[2] 伏金浩,王剑平,闻路红,等. 基于改进阈值与边缘梯度的亮场干细胞图像分割方法[J]. 电子测量技术,2020, 43(20):109-114.

FU J H, WANG J P, WEN L H, et al. Brightfield stem cell image segmentation method based on improved threshold and edge gradient[J]. Electronic Measurement Technology, 2020, 43(20): 109-114.

[3] CHEN B, ZHANG X, WANG R T, et al. Detect concrete cracks based on OTSU algorithm with differential image[J]. The Journal of Engineering, 2019(23):9088-9091.

[4] 刘旭,陈里里,范国栋,等. 基于改进 Cascade R-CNN 算法的道路表面缺陷检测[J]. 自动化与仪器仪表, 2023,(10):10-14.

LIU X, CHEN L L, FAN G D, et al. Road surface defect detection based on an improved cascade R-CNN algorithm [J]. Automation and Instrumentation, 2023,(10): 10-14.

[5] 胡凤阔,叶兰,谭显峰,等. 一种基于改进 YOLOv8 的轻量化路面病害检测算法[J]. 图学学报,2024,45(5): 892-900.

HU F K, YE L, TAN X F, et al. A lightweight road surface defect detection algorithm based on improved YOLOv8 [J]. Journal of Graphical Science, 2024, 45(5): 892-900.

[6] 魏陈浩,杨睿,刘振丙,等. 具有双层路由注意力的 YOLOv8 道路场景目标检测方法[J]. 图学学报,2023, 44(6):1104-1111.

WEI CH H, YANG R, LIU ZH B, et al. A YOLOv8 road scene object detection method with dual-layer routing attention [J]. Journal of Graphic Science, 2023, 44(6): 1104-1111.

[7] 周建新,张媛,贾梓涵,等. 改进 YOLOv8 的路面缺陷检测算法[J]. 电子测量技术,2024,47(19):146-154.

ZHOU J X, ZHANG Y, JIA Z H, et al. The improved YOLOv8 road surface defect detection algorithm[J]. Electronic Measurement Technology, 2024, 47(19): 146-154.

[8] 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 S, SHI T, JIN F K. Improved YOLOv8 road damage detection algorithm [J]. Computer Engineering and Applications, 2023, 59 (23): 165-174.

[9] 赵佰亭,程瑞丰,贾晓芬. 融合多尺度特征的 YOLOv8 裂缝缺陷检测算法[J]. 计算机工程与应用,2024, 60(22):261-270.

ZHAO B T, CHENG R F, JIA X F. A YOLOv8-based crack defect detection algorithm integrating multi-scale features [J]. Computer Engineering and Applications, 2024, 60(22), 261-270.

[10] 叶发茂,张立,袁燎,等. DB-YOLO:特征增强融合的双骨干 YOLOv8 道路缺陷检测模型[J]. 计算机工程

与应用,2024,60(24):260-269.

YE F M, ZHANG L, YUAN L, et al. DB-YOLO: A dual-backbone YOLOv8 road defect detection model with feature enhancement and fusion[J]. Computer Engineering and Applications, 2024, 60 (24): 260-269.

[11] ZHANG X, LIU C, YANG D G, et al. RFACnv: Innovating spatial attention and standard convolutional operation [J]. ArXiv preprint arXiv: 2304. 03198,2023.

[12] FENG Y, HUANG J, DU S, et al. Hyper-YOLO: When visual object detection meets hypergraph computation[J]. 2025, 47(4): 2388-2401.

[13] MA X, DAI X Y, BAI Y, et al. Rewrite the stars[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2024, Seattle, Washington, USA, 2024: 5694-5703.

[14] YANG J W, LI CH Y, DAI X Y, et al. Focal modulation networks [J]. ArXiv preprint arXiv: 2203. 11926, 2022.

[15] ARYA D, MAEDA H, GHOSH S K, et al. RDD2022: A multi-national image dataset for automatic road damage detection[J]. Geoscience Data Journal, 2024, 11(4):846-862.

[16] MAEDA H, SEKIMOTO Y, SETO T, et al. Road damage detection and classification using deep neural networks with smartphone images [J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(12): 1127-1141.

[17] 王宇航. 基于深度学习的路面裂缝检测研究与应用[D]. 西安:西安理工大学, 2024.

WANG Y H. Research and application of pavement crack detection based on deep learning[D]. Xi'an: Xi'an University of Technology, 2024.

[18] LUO X D, WU Y Q, WANG F Y. Target detection method of UAV aerial image based on improved YOLOv5[J]. Remote Sensing, 2022, 14(19):5063.

[19] DU Q, ZHANG S, YANG S. BLP-YOLOv10: efficient safety helmet detection for low-light mining[J]. Journal of Real-Time Image Processing, 2024, 22(1):10.

[20] 李彬,李生林.改进 YOLOv11n 的无人机小目标检测算法[J]. 计算机工程与应用, 2025, 61(7): 96-104.

LI B, LI SH L. Improved YOLOv11n drone small target detection algorithm[J]. Computer Engineering and Applications, 2025, 61(7): 96-104.

[21] REN M, ZHENG X F, ZHI X B, et al. An annotated street view image dataset for automated road damage detection[J]. Scientific Data, 2024, 11(1):407.

作者简介

曹鸿博,硕士研究生,主要研究方向为深度学习、图像处理、目标检测。
E-mail:18773791923@163.com

安维胜(通信作者),副教授,主要研究方向为深度学习、控制与优化。
E-mail:anweisheng@swjtu.edu.cn

梁海鹏,硕士研究生,主要研究方向为深度学习、图像处理。

林强,副教授,主要研究方向为深度学习、图像处理。