



DOI:10.12404/j.issn.1671-1815.2403279

引用格式:姚庆安,王友刚,冯云丛,等.融合动态蛇卷积的YOLOv8道路裂缝检测[J].科学技术与工程,2025,25(12):5083-5092.

Yao Qing'an, Wang Yougang, Feng Yuncong, et al. YOLOv8 road crack target detection method integrating dynamic snake convolution[J]. Science Technology and Engineering, 2025, 25(12): 5083-5092.

## 融合动态蛇卷积的 YOLOv8 道路裂缝检测

姚庆安,王友刚,冯云丛,王雪潇

(长春工业大学计算机科学与工程学院,长春 130102)

**摘要** 针对当前道路损伤检测技术中存在的效率不高、成本过高以及不易于在移动设备上部署等挑战,提出一种基于改进 YOLOv8 算法的新型道路裂缝检测方法,命名为 YOLOv8-RC (YOLOv8-road crack)。所提方法在 YOLOv8n 的架构基础上,对 C2f 模块进行改良,引入动态蛇形卷积技术以更精确地识别管状结构特征,同时能够自适应地关注于纤细和弯曲的局部结构。所提算法中新增一种效率高的多尺度注意力机制 (efficient multi-scale attention, EMA),有效提升了识别精度。在模型的颈部结构中,加入加权双向金字塔网络 (bi-directional feature pyramid network, BiFPN),实现了特征的多尺度融合,优化了算法的精度和效率。在 RDD2022\_China\_MotorBike 和 RDD2022\_Japan 两个数据集上的实验结果显示,改进后的 YOLOv8n-RC 模型的 mAP50 分别为 78.8% 和 43.8%,较原 YOLOv8n 模型分别提高了 3.9% 和 3%。所提算法的模型参数总量仅为 2.84 M,计算复杂度为 7.8 G,从而证明了所提方法的实用性和优越性。

**关键词** 动态蛇卷积; YOLOv8; 道路裂缝; 道路安全; 目标检测; 注意力

中图分类号 TP391.4;

文献标志码 A

## YOLOv8 Road Crack Target Detection Method Integrating Dynamic Snake Convolution

YAO Qing-an, WANG You-gang, FENG Yun-cong, WANG Xue-xiao

(College of Computer Science and Engineering, Changchun University of Technology, Changchun 130102, China)

**[Abstract]** In response to the challenges of low efficiency, high cost, and difficulty in deployment on mobile devices in current road damage detection technology, a novel road crack detection method based on the improved YOLOv8 algorithm, named YOLOv8 road crack (YOLOv8-RC), was proposed. The C2f module, based on the YOLOv8n architecture, was enhanced through the introduction of dynamic snake convolution technology, which more accurately identified tubular structural features and adaptively focuses on fine and curved local structures. Furthermore, a highly efficient multi-scale attention (EMA) mechanism was incorporated into the algorithm, effectively enhancing recognition accuracy. In the neck structure of the model, a weighted bidirectional pyramid network (BiFPN) was added to achieve multi-scale fusion of features, thereby optimizing both the accuracy and efficiency of the algorithm. Experimental results on the RDD2022-China-MotorBike and RDD2022-Japan datasets demonstrate that the improved YOLOv8n-RC model achieves mAP50 scores of 78.8% and 43.8%, respectively, representing improvements of 3.9% and 3% over the original YOLOv8n model. The total number of model parameters for the proposed algorithm is only 2.84 M, and the computational complexity is 7.8 G, underscoring the practicality and superiority of this method.

**[Keywords]** dynamic snake convolution; YOLOv8; road damage; road safety; object detection; attention

路面裂缝是常见的道路病害,减少路面的服务年限,损害行驶安全性。适时、正确地开展路面裂缝检查工作是路面巡检系统重点任务,其相关技术的研发也是工作重点。

在传统的裂缝检测领域,方法通常基于数字图像处理技术,通过分析图像中的颜色、形状和边缘

等特征来识别和分类裂缝。在此背景下, Li 等<sup>[1]</sup>开发了一种结合拐角检测与分割网络的矩阵网络方法,专门用于检测道路裂缝。Teng 等<sup>[2]</sup>提出了一种基于形状特征的裂缝检测技术,该技术考虑了潜在裂缝在空间上的分布,便于进行路面健康监控。

在深度学习领域,卷积神经网络 (convolutional

收稿日期: 2024-05-06; 修订日期: 2025-01-23

基金项目: 吉林省发展改革委员会技术研发专项(2023C042-6); 吉林省自然科学基金-自由探索一般项目(YDZJ202201ZYTS422); 吉林省科技厅青年成长科技计划项目(20210508039RQ)

第一作者: 姚庆安(1975—),男,汉族,吉林长春人,硕士,副教授。研究方向:图像处理、智能数据处理、深度学习。E-mail: yao@ccut.edu.cn.

投稿网址: www.stae.com.cn

neural networks, CNN)<sup>[3]</sup>被广泛认为具有卓越的特征提取能力,特别是在识别多变道路环境中的裂缝方面显示出显著优势。这一优势使得基于 CNN 的道路裂缝检测成为近年来的研究焦点。Chen 等<sup>[4]</sup>设计了 PCSN (pavement crack segnet) 模型,这是一种全卷积网络的编解码结构,能够处理任意尺寸的图像,并支持像素级的训练方法。王森等<sup>[5]</sup>开发的 CrackFCN (creak fully convolutional networks) 模型在处理复杂背景中裂缝检测的问题上取得了实质性进展。Liu 等<sup>[6]</sup>提出的 FPCNet (fast pavement crack network) 利用不同膨胀率的空洞卷积来深入提取裂缝特征,并采用 SE 通道注意力机制优化特征的上采样过程,从而加速裂缝检测。Thi 等<sup>[7]</sup>开发的两阶段卷积神经网络先通过一阶段消除图像中的噪声或伪影,然后在第二阶段学习从背景中检测区域裂缝。Sun 等<sup>[8]</sup>则创建了一个包含多个噪声裂纹图像的数据集 NCD,以提高在噪声环境下的裂缝检测性能。张志华等<sup>[9]</sup>则通过结合空洞卷积与双注意力机制,进一步提高了裂缝检测的准确性。上述研究对推动道路裂缝检测技术的发展做出了关键贡献。

现有的深度学习方法在应对道路裂缝的复杂形状和多变环境时,仍存在特征提取能力欠缺,检测精度不足,模型参数较大等问题。为解决这些问题,提出一种基于改进 YOLOv8 算法的新型道路裂缝检测方法 YOLOv8-RC。在 C2f 中融入动态蛇形卷积技术;通过自适应地延伸和收缩卷积核,更好地捕捉道路裂缝的细长和弯曲特征。引入多尺度注意力机制 (efficient multi-scale attention, EMA), 结合多尺度信息,提高特征图的辨识能力和检测精度。采用加权双向特征金字塔网络 (bi-directional feature pyramid network, BiFPN), 实现高效的多尺度特征融合,增强了对不同尺度目标的检测能力,显著降低模型的参数量和计算复杂度。改进后模型更为轻量、高效,对道路裂缝的特征识别更为敏感,更适用于在移动设备上实时部署,在道路巡查,维护道路安全方面具有重要意义。

## 1 YOLOv8 目标检测算法

YOLO (you only look once)<sup>[10-14]</sup> 是一种广泛研究的单阶段目标检测算法,自从首次推出以来一直是学术界和业界的热点。特别是在 2023 年, Ultralytics 团队发布了 YOLOv8 版本,这一版本在保持实时性能的同时,实现了检测精度的显著提升和网络结构的轻量化。这些特点使得 YOLOv8 特别适合于对道路裂缝目标进行高效准确的检测<sup>[15-17]</sup>。鉴于移动环境下场景变化迅速,检测系统需要具备

高准确性。

在 YOLOv8n 模型的基础上,在骨干网络 C2f 中引入动态蛇卷积准确地捕捉管状结构的特征,对路面裂缝的检验测试准确度大大提高。将 EMA 注意力机制融入 YOLOv8n 的颈部和尾部网络中,并将从骨干网络获取到的更丰富的特征信息细化,特别在图像中的目标区域。为该网络引入可选择的权重,以调整不同输入信息的重要性,并利用层层自顶向下和自底向上的多尺度信息融合和迭代过程,以实现目标检测。改进后的整体框架结构如图 1 所示。

## 2 改进 YOLOv8 道路裂缝检测算法

### 2.1 C2f-模块 (动态蛇卷积)

准确分割各种拓扑管状结构,如血管和道路等,在许多领域中,这对于确保下游任务的准确性和效率至关重要。然而,由于局部结构薄弱和全局形态多变等因素,这项任务变得复杂。

动态蛇卷积 DSConv 方法用于捕捉细长且曲折的局部管状结构的特征,并引入多视角特征融合策略和基于持久同源性的连续性约束损失函数来增强特征融合和拓扑连续性约束。使用动态蛇卷积来提取管状结构的局部特征。

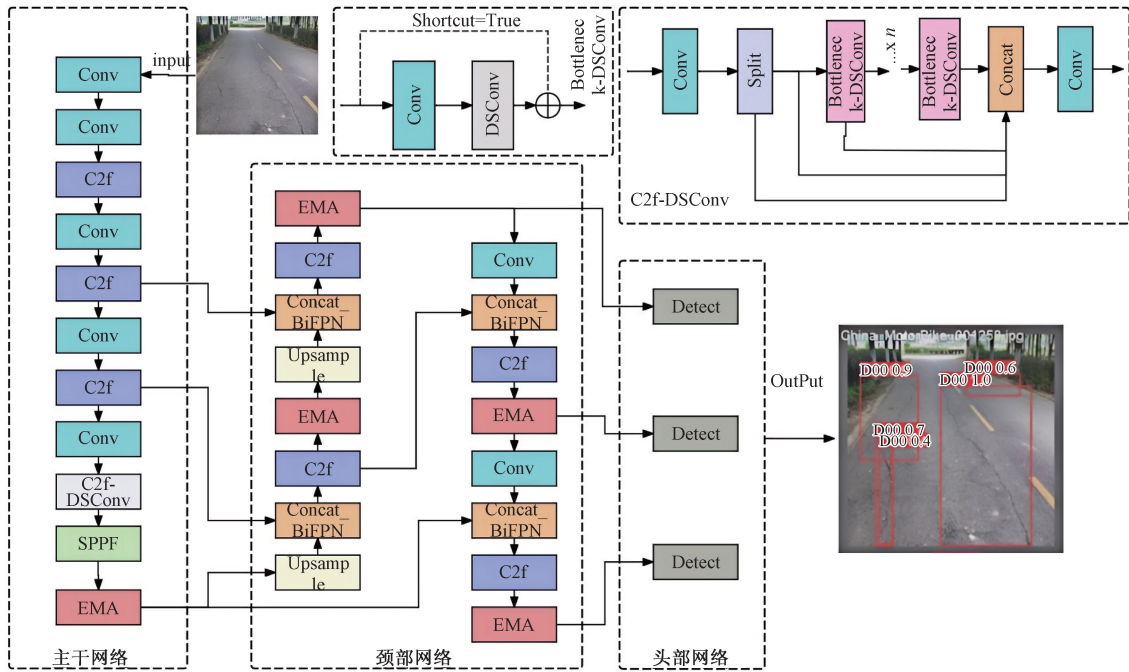
给定标准 2D 卷积坐标为  $K$ , 中心坐标为  $K_i = (x_i, y_i)$ 。具有伸缩的  $3 \times 3$  核  $K$  可表示为

$$K = \{(x-1, y-1), (x-1, y), \dots, (x+1, y+1)\} \quad (1)$$

为了使卷积核更灵活地关注目标的复杂几何特征,将引入变形偏移  $\Delta$ 。然而,如果让模型自由地学习变形偏移,但是在薄管状结构的情况下,感知场容易偏离目标。因此,采用迭代策略,如图 2 所示,对每个待处理的目标依次选择以下待观察的位置,从而保证注意的连续性,同时不会因变形偏移过大而使感知场传播太远。

在 DSConv 中,伸直标准卷积核,在  $x$  轴和  $y$  轴方向上。考虑一个大小为 9 的卷积核,以  $x$  轴方向为例,  $K$  中每个网格的具体位置表示为:  $K_{i \pm c} = (x_{i \pm c}, y_{i \pm c})$ , 其中  $c = \{0, 1, 2, 3, 4\}$  为到中心网格的水平距离。每个网格位置  $K$  的选择  $i \pm c$  在卷积核中,  $K$  是一个累积过程。从中心位置  $K$  开始  $i$  距离中心网格的位置取决于前一网格的位置:  $K_{i+1}$  与  $K_i$ 。因此,通过这种递进式偏移机制,最终使所有偏移量的累积和满足  $\Sigma$  条件,从而确保卷积核符合线性形态结构。图 3 在  $x$  轴方向上可表示为

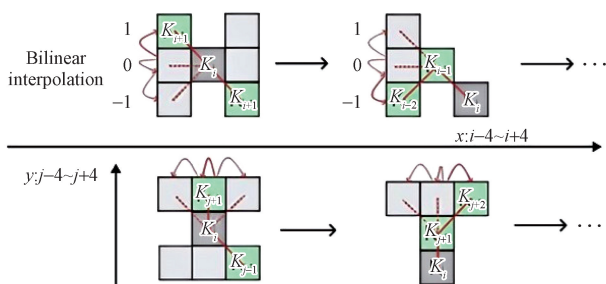
$$K_{i \pm c} = \begin{cases} (x_{i+c}, y_{i+c}) = (x_i + c, y_i + \sum_{i-c}^{i+c} \Delta y) \\ (x_{i-c}, y_{i-c}) = (x_i - c, y_i + \sum_{i-c}^i \Delta y) \end{cases} \quad (2)$$



input 为输入; Conv 为卷积块; C2f 为融合两卷积 Bottleneck 的 CSP 快速实现; C2f-DSCConv 为融入动态蛇卷积的模块; SPPF 为空间金字塔池化; Concat\_BiFPN 为加权双向特征金字塔网络; Upsample 为上采样层; EMA 为注意力机制层; Detect 为检测层; Output 为输出

图 1 优化的 YOLOv8n 网络结构

Fig. 1 Optimized YOLOv8n network structure



Bilinear interpolation 为双线性差值;  $x: i-4 \sim i+4$  为  $x$  轴取值范围;  $y: j-4 \sim j+4$  为  $y$  轴取值范围

图 2 DSCConv 的坐标计算示意图

Fig. 2 Illustration of the coordinates calculation of the DSCConv

在  $y$  轴方向上可表示为

$$K_{j \pm c} = \begin{cases} (x_{j+c}, y_{j+c}) = (x_j + \sum_j^{+c} \Delta y, y_j + c) \\ (x_{j-c}, y_{j-c}) = (x_j + \sum_j^{-c} \Delta y, y_j - c) \end{cases} \quad (3)$$

沿  $y$  轴(或  $x$  轴)正方向移动  $c$  个单位时,  $x$  坐标为  $x_j$ (或  $x_i$ )加上从位置  $j$  到  $j+c$ (或从位置  $i$  到  $i+c$ )的所有  $\Delta x$  的累加和(即  $\sum_j^{j+c} \Delta x$  或  $\sum_i^{i+c}$ )。

由于偏移量  $\Delta$  通常为分数阶, 双线性插值实现可表示为

$$K = \sum_{K'} B(K', K) K' \quad (4)$$

式(4)中:  $K$  为式(2)、式(3)在  $x$  轴与  $y$  轴分数位置;  $K'$  为枚举了在  $x$  轴与  $y$  轴所有的积分空间位

置;  $B$  为双线性插值核, 将其分离为两个一维核, 可表示为

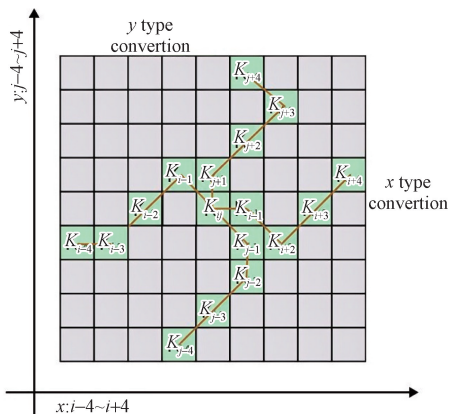
$$B(K, K') = b(K_x, K'_x) b(K_y, K'_y) \quad (5)$$

如图 3 所示, 由于二维 ( $x$  轴,  $y$  轴) 的变化, DSCConv 在变形过程中覆盖  $9 \times 9$  的范围。DSCConv 的设计是为了更好地适应基于动态结构的细长管状结构, 如道路裂缝之类的, 从而更好地感知关键特征。

## 2.2 EMA 模块

对于不同的计算机视觉任务时, 通道或空间注意力机制对产生更可确定的特征所表示的显著效果进行了描述。但是, 使用多通道降维方式对跨通道关系的模拟, 也会对深度视觉表示的信息提取方式产生副作用。

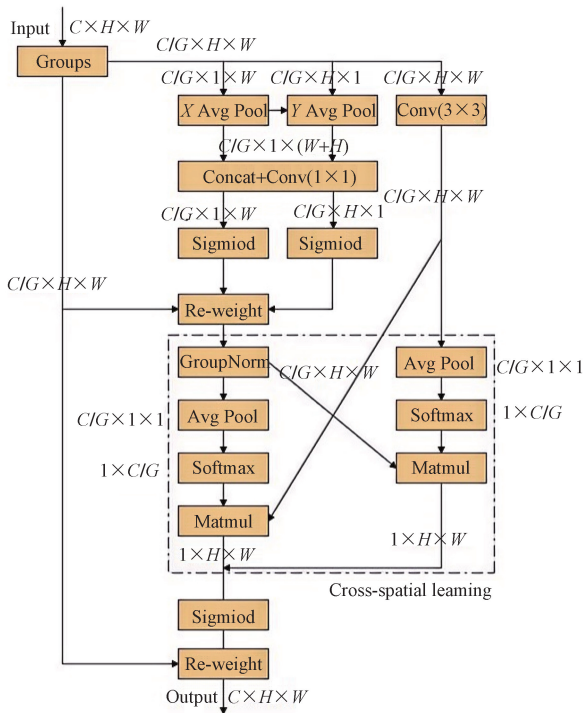
网络中的并行子结构有助于避免过多的顺序处理和深度。将并行处理策略合并到 EMA 模块中, EMA 的整体结构如图 4 所示。将讨论 EMA 如何在道路裂缝检测中学习有效的通道描述, 而不会降低通道维度, 并为高级特征图产生更好的像素级注意力。具体来说, 从 CA 模块中选择  $1 \times 1$  卷积的共享分量, 并将其命名为 EMA 中的  $1 \times 1$  分支。为了聚合多尺度空间结构信息并实现快速响应, 将一个  $3 \times 3$  内核平行于  $1 \times 1$  分支, 将其命名为  $3 \times 3$  分支。考虑特征分组和多尺度结构, 有效地建立短程和长程依赖关系有助于提高性能。



$x$  type conversion 为  $x$  型转换;  $y$  type conversion 为  $y$  型转换;  
 $x:i-4 \sim i+4$  为  $x$  轴取值范围;  $y:j-4 \sim j+4$  为  $y$  轴取值范围

图3 DSConv 的接收域

Fig. 3 The receptive field of the DSConv



Input 为输入; Cross-spatial learning 为跨空间学习组块  
Re-weight 为权重; GroupNorm 为数据归一化操作;  
Softmax 为自然非线性函数; Matmul 为矩阵点积相乘运算;  
 $X$ 、 $Y$ 、Conv 为 3 条并行子网;  $X$  Avg Pool 为 1D 水平全局池;  
 $Y$  Avg Pool 为 1D 垂直全局池; Concat + Conv(1 x 1) 为连接提取特征;  
Sigmoid 为捕捉像素级对关系的函数; Output 为输出

图4 EMA 注意力机制结构

Fig. 4 Structure of EMA attention mechanism

对于任何给定的输入特征映射,EMA 将在通道维度方向上将  $X$  划分为  $G$  个子特征,以学习不同的语义,  $C$  表示输入通道的数量,  $H$  和  $W$  分别为输入特征的空间维度。其中组样式  $X = [X_0, X_1, \dots, X_{C-1}]$ ,  $X_i \in \mathbf{R}^{C//G \times H \times W}$ 。将  $G$  个分组重塑并排列到批

量维度中,并将输入张量重新定义为形状为  $C//G \times H \times W$ ,其中  $C//G$  表示  $C$  对  $G$  整除。在不失一般性的情况下,让  $G \ll C$ ,并假设学习的注意力权重描述符将被用来加强每个子特征中的感兴趣区域的特征表示。

为捕获多尺度的空间信息,神经元的大的局部感受野能够收集广泛的信息。为了描述分组特征图的注意力权重,EMA 引导利用 3 条并行路线。其中,两条并行路由位于  $1 \times 1$  分支中,另一条并行路由位于  $3 \times 3$  分支中。为了在保持计算预算的同时捕获通道之间的依赖关系,在通道方向上建模了跨通道的信息交互。具体而言,使用两个 1D 全局平均池化操作,分别在  $1 \times 1$  分支中沿着两个空间方向对通道进行编码,并且在  $3 \times 3$  分支中只使用一个  $3 \times 3$  内核来捕获多尺度特征表示。

EMA 提出了一种不同空间维度方向的跨空间信息聚合方法,以实现更丰富的特征聚合。仍然引入了两个张量,其中一个为  $1 \times 1$  分支的输出,另一个是  $3 \times 3$  分支的输出。然后,利用二维全局平均池化将全局空间信息编码在  $1 \times 1$  分支的输出中,并且最小分支的输出将直接在通道特征的联合激活机制之前被转换为相应的维度形状,如  $\mathbf{R}_3^{1 \times C//G} \times \mathbf{R}_1^{C//G \times H \times W}$ 。2D 全局池化在通道  $c$  的操作  $z_c$  可表示为

$$z_c = \frac{1}{H \times W} \sum_{j=1}^H \sum_{i=1}^W x_c(i, j) \quad (6)$$

式(6)中:  $x_c$  为输入特征图在通道  $c$  上位置  $(i, j)$  处的特征值,其被设计用于对全局信息进行编码并对长范围依赖性进行建模。

注意力因子仅由每个组内全局和局部特征描述符之间的相似性引导。考虑跨空间信息的聚合,对长程依赖关系进行建模,将精确的位置信息嵌入 EMA。卷积神经网络能够通过融合来自多个尺度的上下文信息,实现对高级特征图的精细像素级注意力建模。此外,卷积核的并行化似乎是一种更强大的结构,可以通过采用跨空间学习方法来处理短期和长期依赖关系。与逐渐减少的感受野形成相反,  $3 \times 3$  和  $1 \times 1$  卷积的并行利用中间特征中的更多上下文信息。

### 2.3 BiFPN 模块

在计算机视觉中,特征金字塔网络(FPN)已成为多尺度特征融合的重要工具,并在不同视觉任务中显示出广泛的适用性。为了提升跨尺度特征融合的效率,新近的架构 PANet<sup>[18]</sup> (path aggregation network) 和 NAS-FPN<sup>[19]</sup> (learning scalable feature pyramid architecture for object detection) 旨在减少计算资源和模型尺寸的同时,增强特征的语义表达能力。

基于这些进展, BiFPN 采用可学习的权重来针对性地处理不同分辨率的特征, 使得网络可以自适应地识别各个输入特征对融合结果的重要性。BiFPN通过反复应用自顶向下与自底向上的融合策略, 显著提升了目标检测性能。与传统方法相比, 常规做法是将所有特征统一到同一尺度后进行简单的相加, 但这种方法未能考虑到特征间的异质性。为了解决这一问题, 引入一种创新的金字塔注意力网络, 该网络利用全局自注意力机制进行上采样, 以提高像素级的定位精度, 更好地适应特征融合的需求。BiFPN 通过实现双向特征流的集成及其快速归一化处理, 有效地捕获了细粒度信息并保持了高语义丰富度, 从而极大地增强了目标检测任务的整体性能。

在考虑双向跨尺度连接的实现中, 一种直接的方法是对各个连接权重应用双向机制, 以便对这些权重进行归一化处理, 使其范围介于  $0 \sim 1$ , 这样可以定量地表示每个输入特征的相对重要性。然而, 这种增加的双向跨尺度连接可能会对 GPU 性能产生显著影响, 从而降低计算效率。为应对这一挑战, 提出一种快速的特征融合策略, 旨在减少因增加双向连接而带来的额外计算负担。

对于快速归一化融合方法, 其计算公式为

$$O = \sum_i \frac{\omega_i}{\varepsilon + \sum_j \omega_j} I_i \quad (7)$$

式(7)中:  $O$  为输出;  $i$  为当前正在处理的输入特征的索引;  $j$  为所有输入特征的索引(包括当前  $i$  对应的特征);  $I_i$  为第  $i$  个输入特征图;  $\sum_j \omega_j$  为所有输入特征的权重之和;  $\omega_i$  为第  $i$  个输入特征的可学习权重,  $\omega_i \geq 0$  是通过在每个  $\omega_i$  之后应用 ReLU 函数来确保的;  $\varepsilon = 0.0001$  是一个小值, 以避免数值不稳定性。

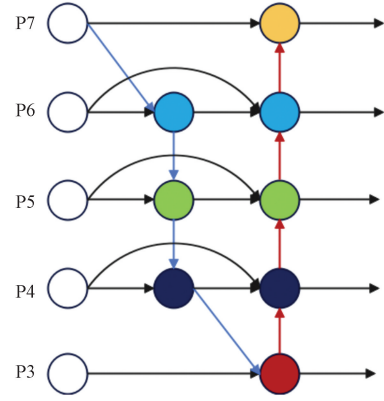
类似地, 每个归一化权重的值也落在  $0 \sim 1$ , 但由于这里没有双向跨尺度连接操作, 因此效率更高。

最终的 BiFPN 结合双向跨尺度连接和快速归一化融合。例如, 图 5 中 BiFPN 的第 6 级的两个融合特征可分别表示为

$$P_6^{\text{ld}} = \text{Conv} \left[ \frac{\omega_1 P_6^{\text{in}} + \omega_2 \text{Resize}(P_7^{\text{in}})}{\omega_1 + \omega_2 + \varepsilon} \right] \quad (8)$$

$$P_6^{\text{out}} = \text{Conv} \left[ \frac{\omega'_1 P_6^{\text{in}} + \omega'_2 P_6^{\text{ld}} + \omega'_3 \text{Resize}(P_5^{\text{out}})}{\omega'_1 + \omega'_2 + \omega'_3 + \varepsilon} \right] \quad (9)$$

式中:  $P_6^{\text{ld}}$  为自顶向下路径中第 6 级的中间特征;  $P_6^{\text{out}}$  为自底向上路径中第 6 级的输出特征;  $P_6^{\text{in}}$  为第 6 级输入特征;  $\omega_1, \omega_2, \omega'_1, \omega'_2, \omega'_3$  分别为  $i$  取不同值时的权重;  $\text{Resize}$  通常是用于分辨率匹配的上采样或下采样运算;  $\text{Conv}$  通常是用于特征处理的卷积运算。



P3 ~ P7 为 3 ~ 7 级的多尺度特征

图 5 BiFPN 特征网络设计

Fig. 5 BiFPN feature network design

该构造原则适用于其他特征层的构建。值得注意的是, 为增强模型的效率, 特征融合阶段采用了深度可分离卷积, 并在每次卷积操作后继而进行批量归一化以及激活函数处理。

### 3 实验结果与分析

#### 3.1 实验环境

实验运行环境如表 1 所示。

表 1 实验运行环境

Table 1 Experimental operating environment

类别	版本号
操作系统	Ubuntu20.04
CPU	Intel(R) Xeon(R) Silver 4210R CPU @ 2.40 GHz
GPU	NVIDIA Quadro RTX 4000
Pytorch 版本	Pytorch 1.12.1
CUDA 版本	CUDA 11.3

#### 3.2 评价指标

采用 4 个关键性能评估指标, 包括精确率  $P$ 、召回率  $R$ 、平均精度 (average precision, AP) 及均值平均精度 (mean average precision, mAP)。其中, 精确率是为分类器正确认定为正类的样本与分类器正确判断的正类样本数量的百分比。召回率, 反映的是分类机正确标识的正类样本数量与所有实际正类样本之间的百分比。平均精度则是通过计算召回率为横轴、精确率为纵轴的精确率-召回率 ( $P$ - $R$ ) 曲线下的面积来确定的。均值平均精度是对所有类别的平均精度进行算术平均得到的, 用以评价模型对整个数据集类别的总体性能。

$$AP = \int_0^1 P(R) dR \quad (10)$$

$$mAP = \frac{1}{m} \sum_{i=1}^m AP_i \quad (11)$$

式(11)中:  $m$  为数据集中的类别数;  $AP_i$  为第  $i$  个类别的平均精度值。

### 3.3 数据集

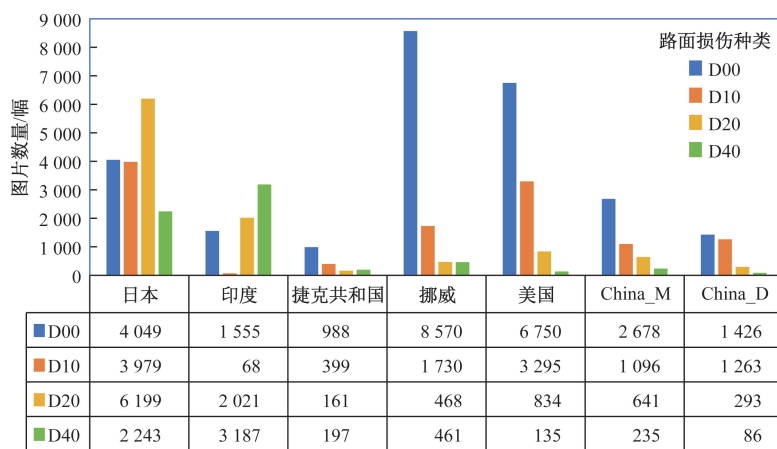
采用 RDD2022 数据收集,包括来自日本、印度、捷克共和国、挪威、美国和中国 6 个国家的 47 420 幅路面图片。这些图像共说明 55 000 个路面的裂缝案例。4 个种类的路面损伤,即 D00(纵向裂缝),D10(横向裂缝),D20(鳄鱼裂纹)和 D40(坑洞),都捕获到了数据集。数据集包含 6 个国家的道路裂缝数据,分别为中国(China)、日本(Japan)、印度(India)、捷克(Czech)、挪威(Norway)、美国(US)。中国数据分为 China\_M 和 China\_D。其中,China\_M 由摩托车上装置智能手机拍摄得到,China\_D 中的图片由无人机搭载摄像头拍摄得到。具体情况将如图 6 所示。其中,选取 China\_MotorBike 数据集,训练集共 1 977 张图片,测试集共 500 张图片,用于本次实验。

### 3.4 消融实验与分析

本次消融实验是在 YOLOv8n 网络模型的基础上实现,在实验过程中均采用相同的初始化参数,均训练 120epoch,在 RDD 2022-China\_MotorBike 数据集上进行消融实验得到对比结果。使用 50% 的 IoU(intersection over union) 阈值下的 mAP 值(mAP50)、50-95% 的 IoU 阈值范围内的 mAP 值(mAP50-95)、浮点运算次数(FLOPs)及参数量(Params)为定量评价指标。

为了验证算法改进的有效性,设计一系列的消融实验,其中 YOLOv8n-D 是在 YOLOv8n 的骨干网络中引入动态蛇卷积,从而准确地捕捉到管状结构的特征。YOLOv8n-E 是通过在 YOLOv8n 的骨干网络尾部和颈部添加了 EMA 注意力块实现。YOLOv8n-B 模型通过引入可选择的权重来判断各种输入条件的重要性,同时也通过将自上而下与自下而上的多尺度数据融合。YOLOv8n-RC 为本文算法,包括上述所有算法。

表 2 展示了针对道路裂缝检测的消融实验结果,证明了改进后的算法在各个阶段均实现了显著的性能提升。具体而言,集成了动态蛇形卷积的 YOLOv8n-D 模型在 mAP50 和 mAP50-95 指标上相较于原始 YOLOv8n 分别有着不同程度的提升。这表明该模型能够自适应地关注于细长和曲折的局部结构,有效地识别管状特征。此外,引入 EMA 注意力机制的 YOLOv8n-E 模型在 mAP 指标上相对于原始模型也表现出了明显的提升,这说明 EMA 注意力机制对于捕获多尺度空间信息非常有效。同时,通过加入可学习权重的 YOLOv8n-B 模型也在 mAP50 上实现了性能提升,该模型通过评估不同特征的重要性,并采用多次自顶向下与自底向上的特征融合策略,有效地增强了目标检测的准确性。



中国数据分为 China\_M 和 China\_D,其中,China\_M 由摩托车上装置智能手机拍摄得到,China\_D 中的图片由无人机搭载摄像头拍摄得到

图 6 RDD 2022 裂缝类别的数据统计

Fig. 6 Damage category-based data statistics for RDD2022

表 2 在 RDD 2022-China\_MotorBike 数据集上的消融实验结果

Table 2 Results of ablation experiments on the RDD 2022-China\_MotorBike dataset

模型	DSCConv	EMA	BiFPN	P/%	R/%	mAP50/%	mAP50-95/%	Params/M	FLOPs/G
YOLOv8n	×	×	×	77.9	66.7	74.9	42.8	3.01	8.1
YOLOv8n-D	√	×	×	79.3	68.3	75.4	44.4	3.70	8.5
YOLOv8n-E	×	√	×	79.6	69.1	75.6	42.6	3.03	8.4
YOLOv8n-B	×	×	√	76.2	69.1	75.7	43.9	2.76	7.3
YOLOv8n-DE	√	√	×	80.4	70.9	76.6	43.8	3.78	8.8
YOLOv8n-DB	√	×	√	80.2	67.9	76.5	42.7	3.17	7.7
YOLOv8n-RC	√	√	√	81.2	69.8	78.8	45.6	2.84	7.8

注:“√”表示在原模型中加入此改进点;“×”表示在原模型中未加入此改进点。

图7为敏感性相关矩阵热力图,直观展示不同模块和性能指标之间的相关性,并带有指示相关系数的注释。颜色渐变有助于快速识别强相关性和弱相关性。每一行对应于不同的YOLOv8模型变体,展示每个模型对C2f-DSCConv、EMA和BiFPN技术的结合如何影响其性能和计算需求。模块参数之间的相关性如图7所示。

### 3.5 对比实验与分析

为评估EMA注意力模块在改进后的算法中的效能,将EMA注意力模块与两种其他注意力机制:CBAM和CA,分别集成到YOLOv8n算法中,并进行了一系列融合实验。实验结果对比了这3种注意力模块的性能,如表3所示。

在表3中展示的数据基于对YOLOv8n网络的扩展,通过分别集成CBAM、CA及EMA3种注意力模块,并对每种配置进行120个训练周期的实验,使用RDD2022-China\_MotorBike数据集进行性能评估。实验结果显示,在相同的训练条件下,YOLOv8n结合CBAM注意力模块的mAP50为74.2%。集成CA注意力模块的YOLOv8n达到74.8%的mAP50。

表3 注意力机制对比

Table 3 Comparison of attention mechanisms

名称	mAP50/%	mAP50-95/%	Params/M
YOLOv8n + CBAM	74.2	42.2	3.95
YOLOv8n + CA	74.8	42.2	3.12
YOLOv8n + EMA	75.6	42.6	3.03

而融合EMA注意力模块的YOLOv8n表现最佳,实现了75.6%的mAP50,此外,该模型也表现出较低的训练权重和参数量,具体检测效果如图8所示。仿真实验结果清晰地表明,相较于CBAM和CA注意力模块,EMA注意力模块显著提升了网络的检测精度。

为了验证本文算法的有效性,将优化的YOLOv8n与YOLOv3、YOLOv4-tiny、YOLOv5s、YOLOv7在RDD2022\_China\_MotorBike数据集上进行对比,得到的对比结果如表4所示。

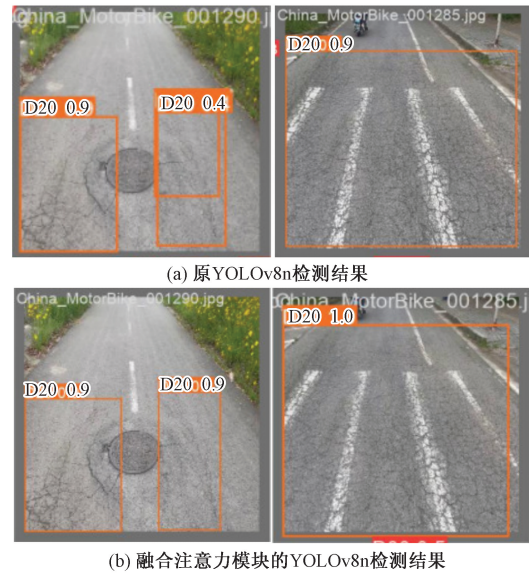


图8 检测结果对比

Fig. 8 Comparison of detection results

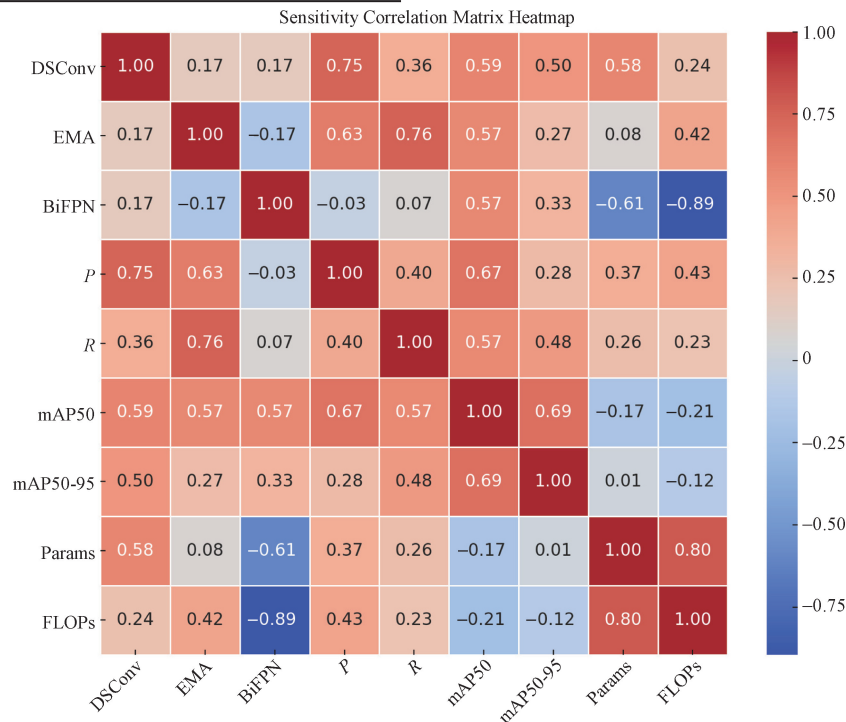


图7 模块敏感性相关矩阵热力图

Fig. 7 Heat map of the module sensitivity correlation matrix

由表 4 可知,虽然优化的 YOLOv8n 的 mAP50 略低于 YOLOv3,但是模型的复杂度、参数量、训练权重都远远小于 YOLOv3;优化的 YOLOv8n 和轻量化模型 YOLOv4-tiny 相比精度有很大的提升;优化的 YOLOv8n 虽然比 YOLOv5s 的 mAP50 高 5%;优化的 YOLOv8n 虽然比 YOLOv7 s 的 mAP50 也有不错的提升;优化的 YOLOv8n 和原 YOLOv8n 相比 mAP50 提升 4.1%, 并且参数量也少于原 YOLOv8n。所以在 RDD 2022 数据集进行实验对比,优化后的 YOLOv8n 在综合性能上和其他目标检测算法相比有一定的优势。具体检测结果效果如图 9 所示。

为了进一步验证所优化的 YOLOv8n 算法对道路裂缝有良好的检测效果,所以将优化算法在 RDD2022\_Japan 数据集上进行验证并将改进算法与 YOLOv3、YOLOv4-tiny、YOLOv5s、YOLOv7 进行对比,结果如表 5 所示。

由表 5 的数据可知,所优化的 YOLOv8n-RC 算法在 RDD2022\_Japan 也能有效的提升检测性能。所优化算法的 mAP 为 43.8,与 YOLOv3、YOLOv4-tiny、YOLOv5s 与 YOLOv7 算法相比,mAP50 分别有着不同程度的提升,与原 YOLOv8n 算法相比,mAP50 提升了 3%,模型复杂度有所下降,参数量和训练权重也有所减少。进一步验证所优化算法的有效性。具体检测结果效果如图 10 所示。

表 4 不同算法在 RDD2022\_China\_MotorBike 数据集上的实验对比结果

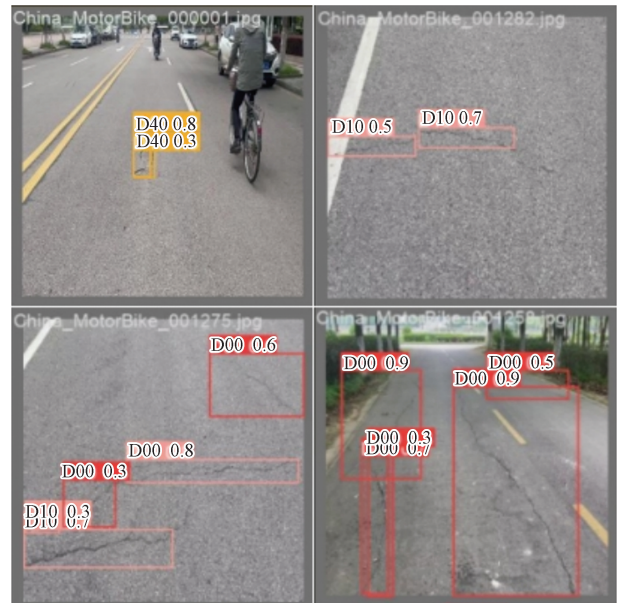
Table 4 Experimental comparison results of different algorithms on RDD2022\_China\_MotorBike dataset

模型	mAP50/%	Params/M	FLOPs/G
YOLOv3	79.6	295.60	81.4
YOLOv4-tiny	58.5	5.69	16.3
YOLOv5s	73.8	3.33	11.2
YOLOv7	73.6	3.38	11.6
YOLOv8n	74.9	3.01	8.1
本文算法	78.8	2.84	7.8

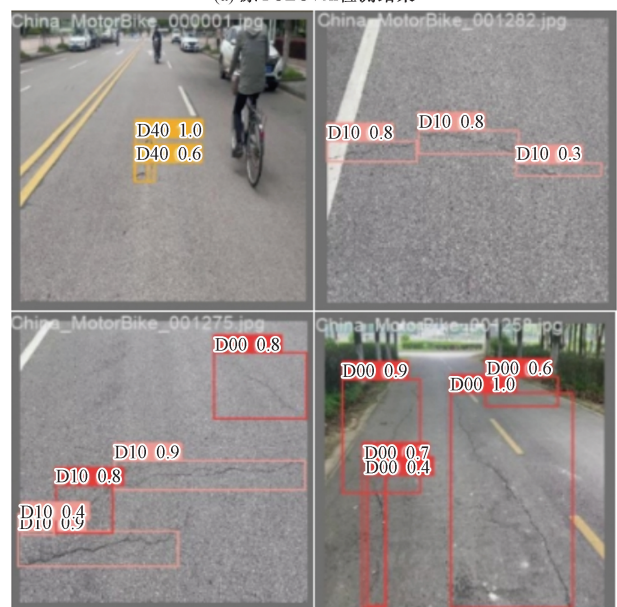
表 5 不同算法在 RDD2022\_Japan 数据集上的实验对比结果

Table 5 Experimental comparison results of different algorithms on RDD2022\_Japan data sets

模型	mAP50/%	Params/M	FLOPs/G
YOLOv3	41.5	295.60	81.4
YOLOv4-tiny	30.7	5.69	16.3
YOLOv5s	38.9	3.33	11.2
YOLOv7	38.8	3.38	11.6
YOLOv8n	40.8	3.00	8.2
本文算法	43.8	2.84	7.8



(a) 原 YOLOv8n 检测结果



(b) 改进后的 YOLOv8n 检测结果

图 9 检测结果对比

Fig. 9 Comparison of detection results

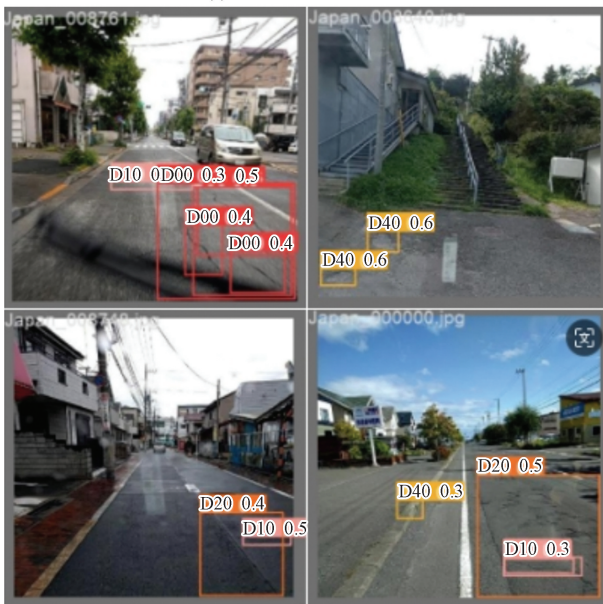
## 4 结论

为了在移动设备上实现高效的道路损伤检测,基于 YOLOv8n 模型提出一种优化的检测方案。通过对该方案进行一系列实验。得出以下结论。

(1) 在 C2f 引入动态蛇卷积,其在道路裂缝图像中的应用能够更好地捕获拓扑管状结构,从而提供了更好的准确性和连续性。传统的卷积操作主要关注图像的局部信息,而动态蛇卷积通过引入蛇形变形机制,使得卷积核能够自适应地延伸和收缩,从而更好地适应道路裂缝的形状和拓扑结构。



(a) 原YOLOv8n检测结果



(b) 改进后的YOLOv8n检测结果

图10 检测结果对比

Fig. 10 Comparison of detection results

这种自适应性能够有效地捕捉到裂缝的连续性,提高了算法在裂缝检测中的表现。

(2) 高效的多尺度EMA注意力模块通过引入通道和空间的注意机制,产生更具辨别力的特征表示,从而提高了算法的检测精度。该模块能够自适应地调整特征图的权重,将更多的注意力集中在重要的特征通道或空间位置上,以捕捉不同尺度下的关键信息。通过这种方式,EMA模块能够提升算法对于不同尺度目标的检测性能,并增强模型对于复杂场景的适应能力。

(3) BiFPN通过引入加权双向金字塔结构,实

现了简单快速的多尺度特征融合。BiFPN的设计与YOLOv8本身的特征提取与预测结构相融合得较好,它能够在不同层级的特征金字塔之间进行双向的跨尺度连接和特征融合。这种结构可以有效地降低计算成本和参数量,并在准确性和效率之间取得平衡。BiFPN通过多层金字塔结构的信息交互和特征融合,提升了算法对于不同尺度目标的检测性能,并增强了模型对于尺度变化和语义信息的建模能力。

综上所述,这3个优化思路在改进的YOLOv8n算法中发挥了关键作用,分别通过动态蛇卷积、高效的多尺度注意力和BiFPN特征融合等方法,提高了道路裂缝检测算法的准确性、连续性和效率。这些优化措施与YOLOv8本身的结构相融合,并经过实验证明了它们的有效性和优越性。

### 参 考 文 献

- [1] Li G, Chen Y, Zhou J, et al. Road crack detection and quantification based on segmentation network using architecture of matrix[J]. *Engineering Computations*, 2021, 39(2): 693-721.
- [2] Teng W, Kasthurirangan G, Omar S, et al. Automated shape-based pavement crack detection approach[J]. *Transport*, 2018, 33(3): 598-608.
- [3] 张鑫, 姚庆安, 赵健, 等. 全卷积神经网络图像语义分割方法综述[J]. *计算机工程与应用*, 2022, 58(8): 45-57.  
Zhang Xin, Yao Qing'an, Zhao Jian, et al. Review of image semantic segmentation methods of fully convolutional neural networks[J]. *Computer Engineering and Applications*, 2022, 58(8): 45-57.
- [4] Chen T, Cai Z, Zhao X, et al. Pavement crack detection and recognition using the architecture of segNet[J]. *Journal of Industrial Information Integration*, 2020, 18: 100144.
- [5] 王森, 伍星, 张印辉, 等. 基于深度学习的全卷积神经网络图像裂纹检测[J]. *计算机辅助设计与图形学学报*, 2018, 30(5): 859-867.  
Wang Sen, Wu Xing, Zhang Yin-hui, et al. Deep learning based fully convolutional network image crack detection[J]. *Journal of Computer Aided Design and Graphics*, 2018, 30(5): 859-867.
- [6] Liu W J, Huang Y C, Li Y, et al. FPCNet: fast pavement crack detection network based on encoder-decoder architecture[J]. *arXiv Preprint*, 2019; arXiv: 1907. 02248.
- [7] Thi H N N, Stuart P, Don B, et al. Two-stage convolutional neural network for road crack detection and segmentation[J]. *Expert Systems with Applications*, 2021, 186: DOI: 10. 1016/J. ESWA. 2021. 115718.
- [8] Sun M S, Zhao H W, Li J A. Road crack detection network under noise based on feature pyramid structure with feature enhancement (road crack detection under noise) [J]. *IET Image Processing*, 2021, 16(3): 809-822.
- [9] 张志华, 温亚楠, 慕号伟, 等. 结合双注意力机制的道路裂缝检测[J]. *中国图象图形学报*, 2022, 27(7): 2240-2250.  
Zhang Zhihua, Wen Yanan, Mu Haowei, et al. Road crack detection using dual attention mechanism[J]. *Chinese Journal of Image and Graphics*, 2022, 27(7): 2240-2250.

- [10] Caie X, Zhe D, Shengyun Z, et al. Fusion network for small target detection based on YOLO and attention mechanism[J]. Optoelectronics Letters, 2024, 20(6): 372-378.
- [11] Vu S P H, Van K N T, Huy L Q, et al. Road surface damages allocation with RTI-IMS software based on YOLOv5 model[J]. Sustainable and Resilient Infrastructure, 2024, 9(3): 242-261.
- [12] Mohd O, Pradeep K. PD-ITS: pothole detection using YOLO variants for intelligent transport system[J]. SN Computer Science, 2024, 5(5): DOI: 10.1007/s42979-024-02887-1.
- [13] 覃紫馨, 姜彦南, 徐立, 等. 基于 YOLO 算法的探地雷达道路图像异常自动检测[J]. 科学技术与工程, 2023, 23(27): 11505-11512.  
Qin Zixin, Jiang Yannan, Xu Li, et al. Automatic detection of anomalies in GPR images based on YOLO algorithm[J]. Science Technology and Engineering, 2023, 23(27): 11505-11512.
- [14] 赵一航, 宋淑彩, 张博. 采用 YOLOv5 模型的路面裂缝检测研究[J]. 福建电脑, 2022, 38(7): 23-26.  
Zhao Yihang, Song Shucui, Zhang Bo. Research on road crack detection using YOLOv5 model[J]. Fujian Computer, 2022, 38(7): 23-26.
- [15] 任安虎, 姜子渊, 马晨浩. 基于改进 YOLOv5s 的道路裂缝检测算法[J]. 激光杂志, 2024, 45(4): 88-94.  
Ren Anhu, Jiang Ziyuan, Ma Chenhao. Road crack detection algorithm based on improved YOLOv5s[J]. Laser Journal, 2024, 45(4): 88-94.
- [16] 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 Chenhao, Yang Rui, Liu Zhenbing, et al. YOLOv8 road scene object detection method with double layer routing attention[J]. Journal of Graphics, 2023, 44(6): 1104-1111.
- [17] 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. Improving YOLOv8 road damage detection algorithm[J]. Computer Engineering and Applications, 2023, 59(23): 165-174.
- [18] Wu Y, Yao Q, Fan X, et al. Panet: a point-attention based multi-scale feature fusion network for point cloud registration[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 17: DOI: 10.1007/s12145-024-01265-y.
- [19] Wang H, Han D, Cui M, et al. NAS-YOLOX: a SAR ship detection using neural architecture search and multi-scale attention[J]. Connection Science, 2023, 35(1): DOI: 10.1080/09540091.2023.2257399.