

文章编号 1004-924X(2026)02-0336-16

自适应特征的轻量化路面裂缝检测方法

刘媛媛, 朱 凯, 顾志辉, 岳 猛, 王靖智, 朱 路*
(华东交通大学 信息与软件工程学院, 江西 南昌 330013)

摘要:针对路面裂缝形态复杂、易受环境干扰,且检测存在精度与轻量化不平衡等问题,本文提出一种自适应特征的轻量化路面裂缝检测方法。首先,根据裂缝狭长且跨度大的特点,设计了裂缝高效注意力机制,通过压缩特征维度,以捕捉其长距离空间依赖。其次,构建动态采样金字塔进行自适应采样和提取目标特征,以增强对异构裂缝特征的表示能力。然后,改进HGNet_GS轻量化主干网络,并提出了轻量化检测头,显著降低了计算冗余;采用Powerful IoU损失函数解决框锚膨胀问题并提升小模型的收敛速度。此外为验证模型泛化性,自建了民用路面缺陷数据集,其中包含不同光照条件下路面缺陷共计2 985张。实验结果表明,与基准模型YOLOv8n相比,本文模型参数量和计算量分别减少了50%和52%。在自建数据集上,mAP50和mAP95分别提升了5.4%和4.1%;在公开的RDD2022数据集上,mAP50和mAP95分别提升了2.1%和1.5%。该模型已应用于边缘设备并完成工程作业测试,验证了其能够满足轻量化路面裂缝检测的工程应用需求,为自动化道路维护提供了技术方案。

关键词:路面裂缝检测;注意力机制;轻量化;动态采样金字塔;YOLOv8

中图分类号:TP394.1 **文献标识码:**A

doi:10.37188/OPE.20263402.0336

CSTR:32169.14.OPE.20263402.0336

Lightweight road crack detection method with adaptive features

LIU Yuanyuan, ZHU Kai, GU Zhihui, YUE Meng, WANG Jingzhi, ZHU Lu*

(School of Information and Software Engineering, East China Jiaotong University, Nanchang 330013, China)

* Corresponding author, E-mail: luyuanwanwan@163.com

Abstract: To address the issues of complex crack morphologies, environmental interference, and the imbalance between detection accuracy and model lightweight requirements in road surface inspection, this paper proposed a lightweight road crack detection method with adaptive feature extraction. First, a Crack Efficient Attention (CEA) module was designed based on the slender shape and large span characteristics of cracks, compressing feature dimensions to capture long-distance spatial dependencies. Second, a Dynamic Sampling Feature Pyramid Network (DSFPN) was constructed for adaptive sampling and target feature extraction, enhancing representation capability for heterogeneous crack features. Third, the HGNet_GS lightweight backbone network was improved, and a CEA Group Head (CGHead) was proposed, significantly reducing computational redundancy; the PIoU (Powerful IoU) loss function was adopted to solve anchor box expansion problems and improve convergence speed for small models. Additionally, a civilian road defect dataset containing 2 985 images under various lighting conditions was established to validate model generalization. Experimental results show that compared with the baseline YOLOv8n model, the

收稿日期:2025-09-25;修订日期:2025-11-06.

基金项目:国家自然科学基金项目(No. 62366015, No. 61963016, No. 61967007)

proposed method reduces parameters and computational cost by 50% and 52%, respectively; on the self-built dataset, mAP50 and mAP95 increase by 5.4% and 4.1%; on the public RDD2022 dataset, these metrics improve by 2.1% and 1.5%. The model has been deployed on edge devices and verified through engineering tests, demonstrating its capability to meet practical requirements for lightweight road crack detection and providing a technical solution for automated road maintenance systems.

Key words: pavement crack detection; attention mechanism; lightweight; dynamic pyramid; YOLOv8

1 引言

道路作为交通系统的基石,是推动经济发展与城市化进程不可或缺的关键要素。在众多国家,道路设施的病害已成为亟待解决的重大难题。裂缝作为道路病害中最为常见的类型,对公路的使用年限和车辆行驶安全构成严重威胁。因此,迅速且准确地识别路面损伤显得尤为重要,这对于道路修复、道路养护、以及延长道路使用寿命具有至关重要的意义。

传统的路面裂缝检测方式主要依靠专业人员对路面进行观察判断,这种方式不仅效率低下,而且可靠性不高。近年来,随着深度学习技术的发展,人们将深度学习的方法用在路面裂缝检测上。王安政等人^[1]设计了一种基于注意力门控的解码模块,抑制非裂缝区域并加强对裂缝区域的关注,提高检测精度。蓝章礼等人^[2]提出了一种基于Sobel算子的路面裂缝检测网络,通过融合轮廓信息实现特征提取。Yu等人^[3]基于Vision Transformer提取裂缝特征并融合语义信息提高精度,虽能捕捉裂缝的长距离依赖,但其巨大的计算成本造成了工程部署困难。李禹萱等人^[4]提出了一种基于Swin-Transformer的路面裂缝检测算法,精度有所提升但模型参数量大幅度上升。高明星等人^[5]提出了一种双金字塔网络路面裂缝检测方法,基于MobileNetv3轻量化改进并使用双金字塔特征融合传递更多编码特征。Zhipeng等人^[6]基于YOLOv7进行改进,设计了一种基于SimAM注意力的轻量化路面裂缝检测方法,并测试了注意力在模型不同位置的效果。邢仁琦等人^[7]基于EfficientNet和CBAM注意力设计了一种小目标裂缝检测算法,将参数量降低至4M。Su等人^[8]基于YOLOX设计了一种保持原始维度信息的路面裂缝检测算法,但在复杂的公共数据集中表现出性能不足。刘光辉等人^[9]通过自适应池化与网络

优化,增强了对细微裂缝的感知能力并提升了精度。夏晓华等人^[10]通过融合语义与三维信息,提出动态逆透视变换模型,提升了路面裂缝检测图像的质量。朱世松等人^[11]基于YOLOv8引入MLAP注意力与DWR卷积,实现了裂缝多尺度特征提取提升了检测精度。贾翔宇等人^[12]基于RT-DERT提出了融合动态网络和动态重排列注意力的轻量化算法,精度得到了提升。胡鹏等人^[13]提出了一种面向嵌入式的航拍路面裂缝检测方法,通过剪枝将YOLOv5模型参数量降低至1.53 M,但模型本体仍然有轻量化的改进空间。

尽管近年来各种模型取得了较为显著进展,但在精度与轻量化之间仍未取得平衡。一方面,裂缝在形态上呈现出异构性,不同环境下的裂缝特征差异很大,且裂缝具有跨度大、不明显的特征,这对模型的精度和泛化能力提出了较高要求;另一方面,模型在轻量化方面还有着进一步提升空间,以便于在边缘设备中实现部署。因此,本文提出一种平衡精度与轻量化的路面裂缝检测算法。首先,设计了CEA模块,通过压缩并融合水平与垂直方向的背景信息,增强模型对大跨度裂缝细节的关注能力。其次,构建了动态采样金字塔DSFPN,其中核心模块Dysample和SDMod用于动态提取目标特征以解决目标异构性的问题。此外,引入HGNet并进行轻量化改进为HGNet_GS。进一步,优化了原有检测头,结合CEA和组卷积,在精度和轻量化方面达到了平衡。最后将损失函数替换为PIoU解决框锚膨胀并加快小模型收敛速度。最终将模型部署于边缘设备中,并应用于工程作业,测试结果表明能够满足路面裂缝检测需求,具备良好的工程应用前景。

2 YOLOv8网络结构

YOLOv8是较为主流的目标检测模型之一,

其优秀的检测性能和泛化性使其在科研和工业领域都有广泛地应用。如图 1 所示, YOLOv8 网络架构由骨干网络、颈部和头部三部分构成。骨干网络采用 CSPDarknet 构架, 其中 C2f 模块负责的特征提取, 快速空间金字塔池化模块 SPPF, 捕获多尺度感受野下的上下文信息。颈部采用 PAFPN 结构, 通过残差连接来获取多尺度特征以增强特征表示能力。在检测头部分, 采用了无框锚的解耦头, 将分类与回归任务分离, 避免了

冗余计算提升了检测速度和精度。YOLOv8 提供 n, s, m, l, x 五个规模版本, 模型精度随规模递增, 但体量随之变大速度下降。为在模型轻量化与检测性能之间取得平衡, 本文选取最小规模的 YOLOv8n 作为基准模型。

3 改进的 YOLOv8 算法

为实现路面裂缝检测的精度与轻量化之间平衡, 解决裂缝检测中存在的异构性强、跨度大、不明显等问题, 本文对 YOLOv8n 模型结构进行改进。首先, 提出了 CEA 模块, 该模块针对裂缝而设计, 通过压缩并融合水平与垂直方向的语义, 增强模型对大跨度目标的关注能力; 其次将改进后的 HGNet_GS 结构作为 YOLOv8n 的骨干网络, 并将 CEA 模块嵌入到骨干中, 改进后的 HGNet_GS 实现了精度与轻量化的平衡; 在 Neck 部分, 构建了动态采样金字塔 DSFPN, 其引入了 Dysample 进行动态特征下采样, 并设计了 SDMod 将提取到的信息进一步细分再加工, 同步展开主要特征与边缘特征的双路径提取和整合, 以解决目标异构性的问题。在检测头中, 本文改进了原检测头存在的不足, 以组卷积和嵌入 CEA 的方式提升了精度和效率。最后用 PIoU 替换原有的 CIoU 模块, 解决框锚放大问题并加快小模型的收敛速度。改进后的 CHDP-YOLO 网络结构如图 2 所示。

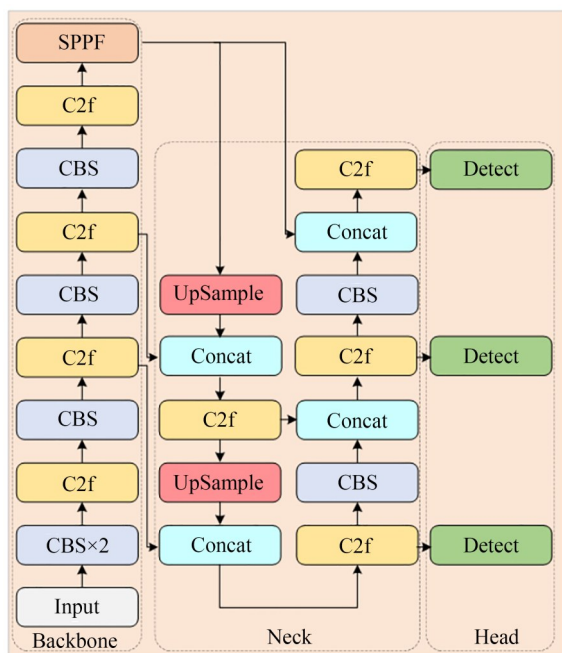


图 1 YOLOv8 结构图

Fig. 1 YOLOv8 Structure

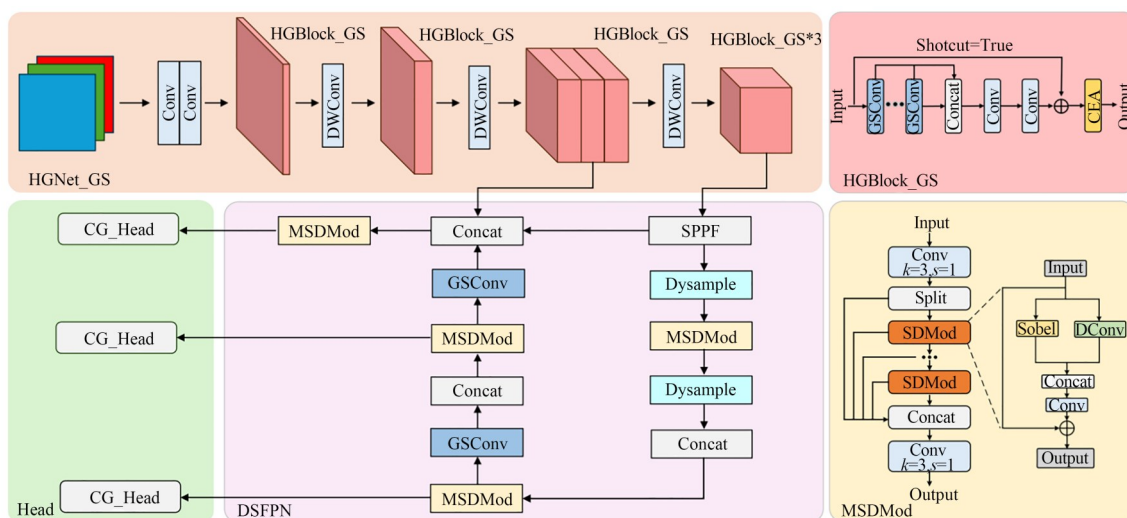


图 2 CHDP-YOLO 网络结构图

Fig. 2 CHDP-YOLO Model Structure

3.1 CEA注意力机制

注意力机制源于人类视觉系统的选择性关注的特性,通过动态分配特征权重使神经网络聚焦于关键区域。ELA^[14](Efficient Local Attention, ELA)是一种轻量级注意力机制,能够捕获感兴趣区域的位置并保持输入特征图通道的维度。通过分离空间维度的水平与垂直特征向量,有效捕捉了长距离空间依赖。其核心在于将通道注意力分解为两个独立方向的一维特征编码过程,分别沿水平和垂直方向聚合特征,随后解析这些特征向量以预测注意力权重,最终通过空间乘积运算整合这些信息,确保了对感兴趣区域位置信息的精确捕捉。

然而,ELA的局限性在于其对两个方向权重的同时依赖性。在进行特征信息融合时,如果某一方向的特征较为显著,而另一方向权重较低,目标位置的特征会被抑制。这一特性对于裂缝检测尤为不利,因为裂缝形状特征往往仅在宽度或高度方向延伸,导致重要的裂缝特征可能被忽略。此外,路面裂缝的识别还需考虑背景信息的融合,不同背景下的相似特征可能代表不同类型的裂缝。

针对上述信息抑制问题并增强模型对复杂路面环境的感知能力,本文提出了一种专为裂缝检测优化的裂缝高效注意力机制(Crack Efficient Attention, CEA)。该机制的核心在于通过解耦方向性注意力来独立计算水平和垂直方向的特征重要性,以适配裂缝的各向异性形态。在此基础上,CEA引入了上下文感知融合机制,通过在注意力权重生成前聚合各方向的全局信息,

从而将背景语义融入其中。且关键的是,CEA采用动态加法聚合策略取代了传统的乘法融合,允许单一方向的显著特征被完整保留,以此动态适应裂缝的多变形态。

具体而言,如图3所示,给定输入特征图 $X \in H \times W \times C$,CEA机制将并行处理两个独立的分支。在垂直方向分支中,首先对输入特征图 X 执行高度维度的全局平均池化(Average Pooling, AvgPool),将全局背景信息压缩为 $1 \times W \times C$ 的特征向量。随后,通过一维卷积(Conv1d, Conv)与组归一化(GroupNorm, GN)处理该向量,最终经由Sigmoid激活函数生成垂直方向的注意力权重 H_{info} 。

$$H_{info} = \delta \left(GN \left(Conv1d \left(AvgPool_h(x) \right) \right) \right). \quad (1)$$

对称地,在水平方向分支中执行相同操作,生成包含水平上下文信息的注意力权重 W_{info} 。

$$W_{info} = \delta \left(GN \left(Conv1d \left(AvgPool_w(x) \right) \right) \right). \quad (2)$$

最终,CEA将原始输入特征图 X 分别与两个方向的动态注意力权重相乘,并采用逐元素相加的方式完成聚合,公式如下:

$$Output = X * W_{info} + X * H_{info}. \quad (3)$$

这种动态加法聚合策略确保了任意方向的显著特征都得以保留,解决了特征抑制问题。通过这一系列设计,CEA模块不仅以轻量化的方式精确地提取了裂缝的各向异性特征,更通过融入背景上下文和采用动态加法聚合,有效规避了信息抑制,增强了模型在复杂场景下对多样化裂缝的识别与定位能力。

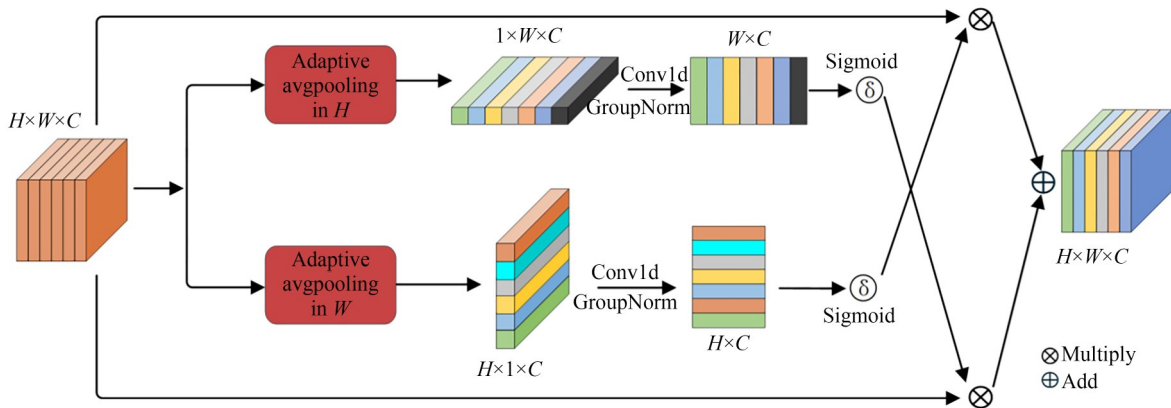


图3 Crack Efficient Attention结构图

Fig. 3 Crack Efficient Attention structure

3.2 HGNet_GS 骨干网络

HGNet 作为 RT-DETR^[15] 的骨干网络, 因其在参数量、计算开销和推理效率上的出色平衡, 在目标检测任务中展现了不错的性能。其结构主要由 HGStem 下采样模块, 基于深度可分离卷积^[16] (Depthwise Separable Convolutions, DW-Conv) 的下采样路径以及多层 HBlock 堆叠构成。然而, 其结构在特征保持能力与轻量化程度上仍存在进一步优化的潜力。为此, 本文在 HGNet 基础上提出一种改进的轻量化主干网络 HGNet_GS, 以进一步提升模型在边缘计算环境下的实用性。

HGNet_GS 的首要改进是优化网络中的前端下采样结构, 本文移除了 HGStem 模块, 并采用两个标准卷积来执行下采样。HGStem 中的多层级联压缩虽能快速降低特征图尺寸, 但也带来了严重的浅层语义信息损失, 且浅层堆叠的特征信息不利于推理速度。相比之下, 标准卷积在实现下采样功能的同时, 能够更完整地保留裂缝的浅层结构特征, 避免了冗余压缩造成的细节丢失。另一项改进在于重构其特征提取单元。本文引入了兼顾效率与表达能力的轻量化卷积单元 GSConv^[17] (Ghost Shuffle Convolution, GSConv), 并以此为基础构建了新型的 HBlock_GS 模块, 其结构如图 4 所示。GSConv 的核心思想在于高效地融合标准卷积、深度可分离卷积与通道混洗 Shuffle 机制。它首先通过标准卷积将输入通道数 C_1 压缩为 $C_1/2$, 再通过深度可分离卷积对特征进行高效提取, 随后拼接两部分特征, 并引入通道混洗操作以打破通道间的固定模式, 增强特征的可表达性与泛化能力。从计算复杂度角度分析, 假设输入特征尺寸为 $W \times H \times C_1$, 输出特征为 $W \times H \times C_2$, 卷积核尺寸为 $K_1 \times K_2$, 则使用标准卷积的计算量为:

$$F = W \times H \times K_1 \times K_2 \times C_1 \times C_2. \quad (4)$$

而使用 GSConv 时的计算量为:

$$F = \frac{1}{2} \cdot W \times H \times K_1 \times K_2 \times C_2 \times (C_1 + 1). \quad (5)$$

可见, GSConv 在保证表达能力的前提下, 降低了近 50% 的计算量, 这为模型的轻量化部署提供了有力支持。在此基础上, 本文进一步将前述的 CEA 注意力机制集成至 HBlock_GS 中。通过 CEA 强化了网络在特征通道间的语义选择能力

与对裂缝关键区域的聚焦能力。最终构建的主干网络 HGNet_GS, 在参数控制、特征提取效率以及对裂缝特征的针对性上, 均相较原始 HGNet 实现了全面优化, 为后续的裂缝检测网络提供了高效且可迁移的特征提取基础。

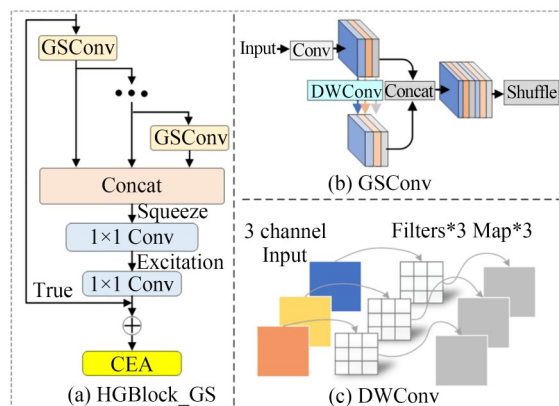


图 4 HBlock_GS 结构图

Fig. 4 HBlock_GS Structure

3.3 DSFPN 动态采样金字塔

在 YOLOv8 中, 路径聚合网络特征金字塔网络 (Path Aggregation Network-Feature Pyramid Network, PAN-FPN) 通过融合自顶向下与自底向上的双向路径, 实现了跨尺度特征的有效交互。然而, 该架构在应用于裂缝检测时暴露出固有局限: 其一, 采用固定的双线性插值法进行上采样, 易导致狭长裂缝的精细纹理信息丢失; 其二, 标准的卷积操作受限于固定的几何结构, 难以灵活适配裂缝多变的形态与空间走向。

为克服 PAFPN 架构的不足, 本文提出一种动态采样金字塔网络 (Dynamic Sampling Feature Pyramid Network, DSFPN)。该模块通过对上采样与特征提取模块动态化改造, 构建了一个灵活的采样金字塔。具体而言, DSFPN 引入了轻量化动态上采样器 Dysample 以取代固定的插值操作, 并设计了以 Sobel 滤波器为基础的 SobelConv, 结合可变形卷积 DConv, 同步展开边缘特征与主要特征的双路径提取和整合, 以解决目标异构性的问题。从而实现对异构裂缝特征的自适应提取与保留。

3.3.1 DySample 动态采样器

在特征金字塔的上采样阶段, 为避免传统插值导致的细节损失, 同时规避其他动态采样方法

高昂的计算开销,本文采用了轻量化的动态上采样器^[18](Dynamic Upsampling, Dysample)。Dysample基于点采样理论,如图5所示,给定输入特征映射,通过一个微型的采样点生成器处理自适应地预测采样点在空间上的最佳偏移量,采样点的动态生成机制,让DySample能够根据不同的输入特征,自适应地调整采样策略,显著提升了特征采样的灵活性与准确度,样点生成器将特征重塑为 $2g \times sh \times sw$ 的形式,最终通过坐标网格采样生成 $c \times sh \times sw$ 的上采样特征图,从而能够以极低的参数量和计算成本,灵活、精确地恢复高分辨率特征图中的裂缝细节。通过该设计,DSFPN有效克服了传统双线性插值在裂缝检测中的不足,在保持网络整体轻量化的同时,显著提升了特征采样的灵活性与准确度。

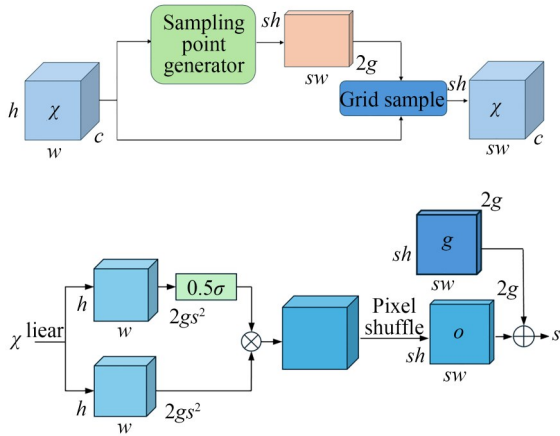


图5 Dysample 结构图

Fig. 5 Dysample structure

3.3.2 MSDMod

在特征提取层面,为解决传统卷积核形状固化的问题,本文设计了一种动态特征提取模块(Sobel Deformable Model, SDMod)。模块结构如图6所示,该模块可以提取目标的边缘特征与主体特征并进行融合,获得一个更加完整的特征图。首先,对于传入的特征图 $H \times W \times C_{in}$,首先经过一个 1×1 的卷积进行维度调整并降低计算量,随后将特征图沿通道维度拆分为两条分支,一条分支留存以用于后续特征拼接,另一条分支则经过 n 次SDMod迭代处理。在SDMod内部,同步开展动态特征与边缘特征的双路径提

取,并对两类特征进行整合,以获取更加完整的裂缝特征信息。最终,将整合后的特征与预留存分支的特征沿通道维度拼接融合,结合一个 1×1 的卷积将特征图维度调整为 $H \times W \times C_{out}$,实现完整特征信息的输出。

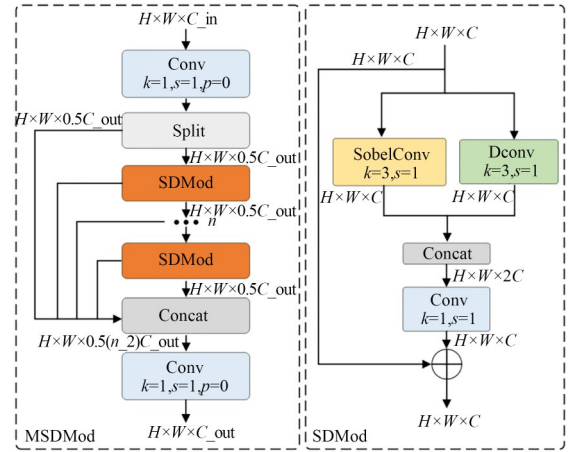


图6 MSDMod和SDMod 结构图

Fig. 6 MSDMod and SDMod Structure

具体而言:一方面,它利用可变形卷积^[19](Deformable Conv, DConv)来动态捕捉裂缝的主体形状。DConv通过引入可学习的空间偏移量,使卷积核的感受野能够根据裂缝的实际走向与形状进行自适应调整,从而灵活地提取非规则的目标特征,其公式表示为:

$$y(p_0) = \sum_{k=1}^K \omega_k \cdot \chi(p_0 + p_k + \Delta p_k) \cdot \Delta m_k, \quad (6)$$

其中: ω_k 代表卷积核的权重, p_0 代表输出特征图的位置, p_k 是卷积核大小, Δp_k 和 Δm_k 分别代表了学习到的采样点偏移量与调制权重,当 $\Delta m_k = 0$ 时,该采样点的贡献为0,当 $\Delta m_k = 1$ 时,该采样点的贡献最高,这一操作使得可变形卷积操作能聚焦于目标区。

另一方面,SDMod集成了Sobel滤波器^[20](Sobel Operator, Sobel),Sobel滤波器是一种边缘信息计算算子,该算子结合了高斯平滑和微分求导运算,利用局部差分寻找边缘,作为边缘检测算子,Sobel具有X方向和Y方向的边缘梯度计算能力。通过Sobel算子核心思想,本文设计了SobelConv,强化了对裂缝边缘高频信息的响应。Sobel算子通过计算像素邻域的梯度来精确标定边缘位置,其水平与垂直方向的梯度核定

义为:

$$\text{Output} = \text{Input} * \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} + \text{Input} * \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 1 & -2 & -1 \end{bmatrix}. \quad (7)$$

通过以 Sobel 算子为核心的 SobelConv 卷积,模块能够高效地提取出富含边缘细节的特征图。最终,SDMod 将可变形卷积提取的自适应形态特征与 SobelConv 提取的精确边缘特征进行有效融合,并通过堆叠多个模块形成 MSDMod (Multiply Sobel Deformable Model, MSDMod)进一步增强表征能力。这一设计使得 DSFPN 能够对裂缝进行更全面、更精准的特征描述,显著提升了检测性能。

3.4 CGHead 轻量化检测头

在 YOLO 检测头中,为获取更加丰富的语义特征使用了大量的卷积堆叠,但在轻量化应用场景下常因其密集的卷积层而导致计算冗余与特征提取效率不高的问题。为解决这一局限,本文基于 YOLO 检测头改进了一种轻量化检测头 CGHead (CEA GroupConv Head, CGHead),其结构如图 7 所示。该架构通过使用组卷积结构与 CEA 注意力机制,实现了计算效率与检测精度的双重提升。

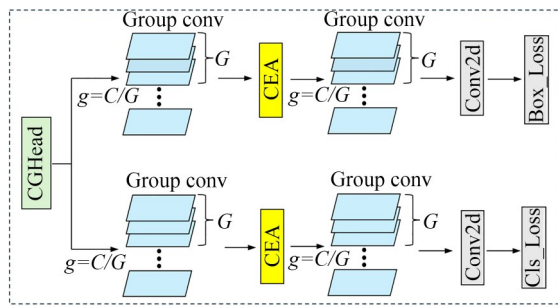


图 7 CGHead 结构图

Fig. 7 CGHead Structure

CGHead 的首要优化在于其卷积结构的重塑。它以组卷积^[21] (Group Convolution, Group-Conv)取代了标准的卷积。组卷积通过将输入特征图的通道划分为若干独立的组,并于各组内分别执行卷积运算,从而打破了通道间的全连接模式。在分组数为 G 的条件下,该策略能将卷积层

的理论计算量与参数量降低至原来的 $1/G$,因而在维持有效特征提取能力的前提下压缩了计算开销。在此基础上,CGHead 进一步集成了本文提出的 CEA 注意力机制,以强化对裂缝目标的精准感知。CEA 以其高效的方向性特征建模能力,引导检测头将计算资源聚焦于最重要的裂缝区域,并抑制背景噪声的干扰。相较于其他注意力机制,CEA 在提供精确特征选择的同时,且几乎不引入额外的计算负担。最终,通过组卷积轻量化特性与 CEA 带来的特征性精准聚焦,两者协同作用,使 CGHead 在实现轻量化的同时,提升了对裂缝目标的检测性能,为构建高效的端侧检测模型提供了坚实的基础。

3.5 PIoU 损失函数

边界框回归的损失函数对于目标检测模型的收敛速度与定位精度起着决定性作用。当前 CIoU 存在以下问题:锚框膨胀导致的梯度消失以及回归复杂引发的收敛缓慢。为解决这些局限,本文引入了 PIoU^[22] (Powerful IoU, PIoU) 作为边界框回归损失函数,其结构如图 8 所示。PIoU 的核心优势在于其引入了一个惩罚因子 P ,该因子同时考量了预测框中心点与四个顶点相对于真实框的归一化距离。这种多点约束的惩罚项不仅有效缓解了传统损失函数中的锚框膨胀问题,还能为边界框的回归提供更稳定、更高效的优化路径,从而加速模型收敛。PIoU 的损失函数定义为:

$$L_{\text{PIoU}} = L_{\text{IoU}} + 1 - e^{-P^2}, 0 \leq L_{\text{PIoU}} \leq 2. \quad (8)$$

惩罚因子 P 综合了预测框在宽度和高度方

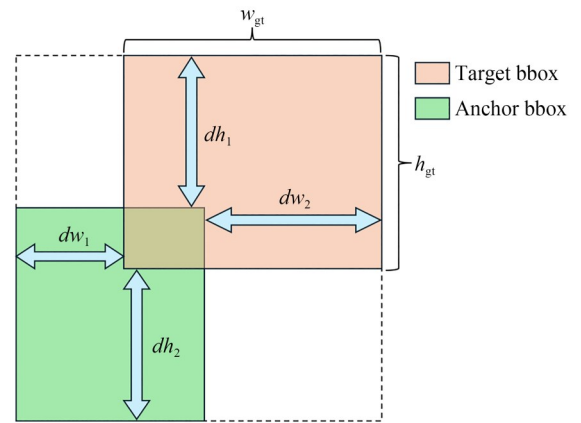


图 8 PIoU 结构图

Fig. 8 PIoU Structure.

向上相对于真实框的偏差,可进一步优化训练动态,定义为:

$$P = \left(\frac{dw_1}{w_{gt}} + \frac{dw_2}{w_{gt}} + \frac{dh_1}{h_{gt}} + \frac{dh_2}{h_{gt}} \right) / 4, \quad (9)$$

其中: dw 和 dh 表示预测框与目标框边界之间的绝对距离, w_{gt} 和 h_{gt} 表示目标框的宽度和高度。

本文采用了PIoU的改进版本PIoUv2。PIoUv2通过引入一个由单一超参数 λ 控制的非单调注意力函数, λ 通过系统性超参数敏感性分析确定,以本文模型为基准测试了 λ 在 $[1.1, 1.7]$ 范围内的多个值,多方面综合考量选取 $\lambda=1.3$ 。PIoUv2损失函数如下,实现了对不同质量锚框梯度的动态重加权。其损失函数定义如下:

$$u(x) = 3x \cdot e^{-x^2}, \quad q = e^{-p}, \quad q \in (0, 1], \quad (10)$$

$$L_{PIoUv2} = u(\lambda q) \cdot L_{PIoU}, \quad (11)$$

其中: q 是锚框质量的度量指标, $u(x)$ 注意力函数会对中等质量的锚框赋予更高的梯度权重,同时适度抑制高质量和低质量锚框的梯度。这种梯度调节策略,使得模型能够更专注于优化那些最具潜力的锚框,进一步提升了回归的效率与鲁棒性,最终实现更高的检测精度。

表1 不同 λ 取值下的性能

Tab. 1 Performance of PIoUv2 with different λ values

λ	Precision	mAP50	mAP95
PIoUv2($\lambda=1.1$)	79.2	72.5	42.3
PIoUv2($\lambda=1.2$)	78.1	72.6	42.4
PIoUv2($\lambda=1.3$)	78.6	73.0	42.9
PIoUv2($\lambda=1.4$)	78.7	72.1	42.3
PIoUv2($\lambda=1.5$)	77.2	73.1	42.8
PIoUv2($\lambda=1.6$)	77.8	72.5	42.5
PIoUv2($\lambda=1.7$)	78.2	72.9	42.7

4 实验数据与评价指标

4.1 实验数据集

为了进行实地测试,本文团队制作了一个民用路面裂缝数据集,由团队实地拍摄并使用Labeling软件对其进行标注。其中包括的路面病害共分为4个类别:纵向裂缝(D00)、横向裂缝(D10)、网状裂缝(D20)、坑洞(D40)共计1 717张,并通过数据增强的方式将其扩充到2 958张。

该数据集包含了光照条件良好/一般/恶劣场景下的水泥路面,同时,数据集还选取了多个时间段,如正午、傍晚等时刻,并采用阴影增强的方式模拟真实路面环境可能存在的阴影遮挡问题(如树荫、建筑物影子遮挡),数据集样本展示如图9所示,从左至右分别是光照条件良好,光照条件一般,光照条件恶劣的数据样本。



图9 自制路面裂缝数据集样本展示

Fig. 9 Dataset Sample Display

同时为了证明本文提出的模型在各种场景下的泛用性,本文还选用了全球道路损伤检测挑战赛提供的RDD2022^[23]公开数据集,该数据集由中国、美国、日本、印度等多个国家的道路图像,具有更加多样性的裂缝和更加复杂的路况环境。从中选取了包含各种国家不同环境下的路面状况与裂缝共20 593张作为本实验数据集,其中路面病害共分为4个类别:纵向裂缝(D00)、横向裂缝(D10)、网状裂缝(D20)、坑洞(D40)。

4.2 评价指标

为了客观地评价模型的具体检测性能,本实验采用以下评价指标。模型参数量(Params)用于衡量模型的复杂度和存储需求。浮点运算量(GFLOPs)表示模型在推理过程中执行的运算总量。当Params和GFLOPs越小代表模型对计算设备的性能要求更低。精度(Precision, P)、召回率(Recall, R)是用于评价正确预测的目标的比例的指标,平均精度(Average Precision, AP)和均值平均精度(mean Average Precision, mAP)作为模型精度的评价指标。其公式:

$$Precision = \frac{TP}{TP + FP}, \quad (12)$$

$$Recall = \frac{TP}{TP + FN}, \quad (13)$$

$$AP = \int_0^1 PdR, \tag{14}$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N}, \tag{15}$$

其中: TP 为检测正确的样本数; FP 为检测错误的样本数; FN 为每次结果中漏检的样本数; AP 是 P-R 曲线与坐标轴围成的面积; mAP 是将各个类别的 AP_i 求均值得到的平均精度; N 为检测类别数, 本文中 N 取值为 4。

5 实验参数与结果分析

5.1 实验配置与超参数

本文提出的算法基于 Python 3.8, 使用 Py-Torch 深度学习框架。操作系统为 Windows10, Cuda11.7, 在 PyCharm Community 上实现, CPU

为 Intel(R) Xeon(R) Silver 4210R 2.40 GHz 内存 128 GB, 显卡为 NVIDIA GeForce RTX 3080, 显存 10 GB。初始学习率设置为 0.01, 学习率下降方式为余弦退火算法, 输入图像的分辨率设置为 640×640 。采用 SGD 作为优化器, 其中动量设置为 0.937, 训练轮数设置为 250epoch, 批次大小设置为 32。

5.2 消融实验

本小节通过消融实验来评估各模块对网络性能的贡献, 并验证其在路面裂缝检测任务中的有效性, 本文以 YOLOv8n 为基准模型进行消融实验。通过对比 mAP、参数量和计算量, 本文验证了每种改进策略对模型的具体影响。具体实验结果如表 2 所示。“√”表示运用了该改进方法。

表 2 消融实验结果

Tab. 2 Ablation experimental results

YOLOv8n	HGNet	HGNet_GS	DSFPN	CGHead	CEA	PIOUv2	参数量/M	GFLOPS/G	mAP50/%	mAP95/%
√							3.01	8.2	67.3	38.3
√	√						2.3	7.0	66.5	37.5
√		√					2.32	7.2	68.2	38.7
√			√				2.83	7.5	70.4	40.4
√				√			2.42	5.7	69.2	38.8
√					√		3.15	8.2	70.2	39.9
√						√	3.01	8.2	68.3	39.1
√		√	√				2.14	6.6	70.9	41.1
√		√	√	√			1.55	4.0	71.1	41.9
√		√	√	√	√		1.56	4.0	72.4	42.2
√		√		√	√	√	1.74	4.8	70.9	41.1
√		√	√		√	√	2.13	6.6	72.5	42.4
√		√	√	√		√	1.54	4.0	71.1	41.3
OurModel		√	√	√	√	√	1.56	4.0	73.0	42.9

实验结果表明, 当主干使用 HGNet 时, 对比 YOLOv8n 模型, 参数量和计算量分别减少了 23% 和 15%, 但其检测精度有所降低, mAP50 降低了 0.8%。而使用本文重新设计改良的 HGNet_GS 时, 参数量和计算量对比 YOLOv8n 分别减少了 22% 和 12%, 检测精度 mAP50 提升了 0.9%。当使用 DSFPN 时, 参数计算量较 YO-

LOv8n 模型减少了 5.7% 和 8.6%, 精度 mAP50 显著上升了 3.1%。当使用 CGHead 检测头时, 模型的参数量与计算量明显降低, 相比于 YOLOv8n 模型参数量计算量分别各减少了 20% 和 31%, 检测精度 mAP50 提升了 1.9%, 证明 CGHead 有着良好的轻量性, 且不会对精度产生影响。当使用本文提出的 CEA 注意力机制时, 在

参数量仅增加 5% 的情况下,检测精度 mAP50 提升了 2.9%。此外,通过引入损失函数 PIOUSv2,在参数和计算量不变的情况下,检测精度 mAP95 提升了 0.7%。当仅移除 DSFPN 或 CEA 模块时,对模型的影响较大,mAP50% 分别降低了 2.1% 和 1.9%,mAP95% 分别降低了 1.8% 和 1.6%,说明这两个模块对精度起到了重要的作用。将上述模块引入,并使用 PIOUSv2 损失函数时,对比 YOLOv8n 模型,参数量和计算量各降低了 48.2% 和 50%,检测精度 mAP50 提升了 5.4%,检测精度 mAP95 提升了 4.1%,可见本文模型在轻量化和精度方面有着良好的平衡。

5.3 不同场景下量化分析实验

为了量化不同场景下模型的性能和指标,对数据集进行了子集分割,分别以光照优良、光照恶劣、复杂裂缝三项指标划分了数据集,其中分割出光照条件优良的裂缝样本 360 张,光照条件恶劣的裂缝样本 91 张,复杂裂缝 224 张,用于验证改进后的模型性能。从表 3 可以看到本文提出的模型在各个方面指标均领先于基准模型,特别是在光照条件恶劣和复杂裂缝的情况下,领先较为明显。基准模型出现了 Recall 较低既漏检的情况,而本文提出的模型有较好性能。

表 3 不同场景下量化分析实验
Tab. 3 Scenario-based quantitative analysis experiments

模型	光照优良			光照恶劣			复杂裂缝		
	Recall	mAP50	mAP95	Recall	mAP50	mAP95	Recall	mAP50	mAP95
YOLOv8n	61.2	70.2	40.7	43.3	45.3	24.9	58.3	68.9	39.9
Our model	70.6	75.2	44.3	69.3	51.3	29.9	66.6	74.5	44.1

5.4 不同模型对比实验

为了验证本文提出的模型在路面裂缝检测任务中的性能,在本小节,将其与其他主流的轻量级目标检测算法进行了比较。具体选择的对比模型包括 YOLOv5, YOLOv7^[24], YOLOv8, YOLOv9^[25], YOLOv10^[26] 和 YOLOv11,并且均选用了这些模型的代表性轻量版本。为了证明模型的泛用性,本文采用了更加复杂的路面数据集 RDD2022 公开数据集进行对比实验。通过表 4 可以看到,本文提出的模型在模型参数量、计算

量上,相比于其他模型均有优势。在 RDD2022 数据集上,对比基准模型 YOLOv8n,参数量降低了 50%,计算量降低了 52%,而精度提升了 1.3%,mAP50 提升了 2.1%,mAP95 提升了 1.5%。此外,YOLOv9tiny 模型虽然表现出不错性能,但 GFLOPs 较大且收敛慢。相比之下本文的模型相比其计算量降低了 63%,参数量降低了 35%,Precision, mAP50 和 mAP95 均有提升,并且收敛较快。图 10 更加直观地展示了各个模型的 mAP 曲线和收敛过程。

表 4 不同模型对比实验
Tab. 4 Comparison of different models experimental results

模型	Input size	Para/M	GFLOPS/G	Precision/%	mAP50/%	mAP95/%
YOLOv5n	640×640	2.5	7.1	63.0	60.8	30.9
YOLOv7tiny	640×640	6.0	13.2	59.5	60.0	28.2
YOLOv8n	640×640	3.0	8.2	65.9	61.8	32.2
YOLOv9tiny	640×640	2.6	10.7	66.0	63.1	33.4
YOLOv10	640×640	2.3	8.2	62.8	59.1	30.7
YOLOv11n	640×640	2.59	6.4	65.2	61.9	32.2
Our model	640×640	1.56	4.0	67.2	63.9	33.7

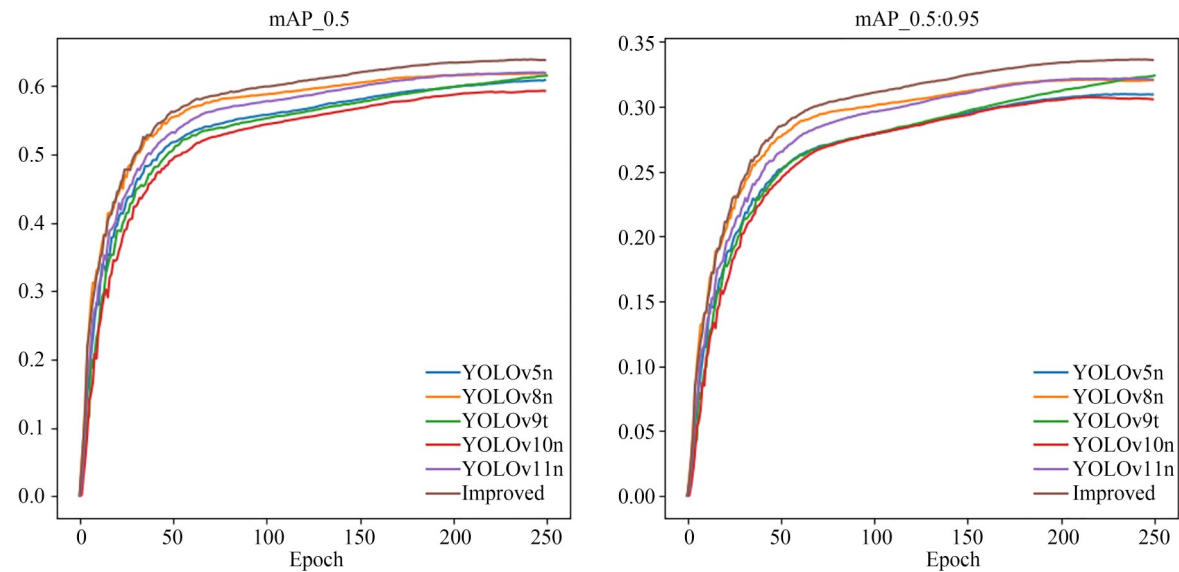


图 10 不同模型之间的 mAP 曲线对比
Fig. 10 Detection result comparison

5.5 注意力对比实验

为了验证 CEA 注意力的性能,本文将 SE^[27] (Squeeze-and-excitation, SE), CBAM^[28] (Convolutional Block Attention Module, CBAM), CA^[29] (Coordinate Attention, CA), ELA 四种主流的注意力机制以及本文提出的 CEA 注意力机制分别添加到本文提出的模型中,并进行对比实验。通过表 5 可以看到,使用 CBAM 注意力时模型参数量较高,达到了 1.7 M,但其检测效果未达到预期,表明高参数量并不直接等同于高性能。同时在表中可以看到,当 CA 注意力添加到的模型中

时,在检测 D40 类型裂缝中有着不错的效果 AP50 达到了 50.2%,但在 D00 和 D20 中表现一般。此外,当使用 ELA 注意力时,表现出了一定的性能提升,但是对于 D40 类型裂缝的检测效果较差。而当使用本文提出的 CEA 注意力机制时,可以看到在保持较低参数量的同时,对各种类型裂缝的检测效果均为最优,特别是对于纵向裂缝 D00 和横向裂缝 D10,检测效果有了大幅提升。综合考虑模型检测精度、模型复杂度,本文选择了精度和参数量间权衡较好的 CEA 模块与模型进行融合以提高综合性能。

表 5 注意力对比实验

Tab. 5 Comparison of attention results

模型	D00	D10	D20	D40	参数量
	AP50/%	AP50/%	AP50/%	AP50/%	
om+SE	69.2	62.9	68.7	49.4	1.57
om+CBAM	69.6	63.3	68.1	49.8	1.70
om+CA	69.2	62.5	68.0	50.2	1.57
om+ELA	70.3	64.2	69.0	49.7	1.55
om+CEA	70.6	64.5	69.3	51.3	1.56

5.6 检测头对比实验

如表 6 所示,对 CGHead 模块进行了对比实验,来检验 CGHead 检测头的性能。实验结果表

明,当使用 YOLOv8n 模型的检测头时,参数和计算量达到最大,却在各项数据上取得了最低的分。而在 YOLOv8n 模型中使用 CGHead 时,模

型在参数量和计算量减少的情况下,各项数据的 AP50 精度均有所上升,其中 D00 纵向裂缝和 D10 横向裂缝提升最为明显,分别上升了 0.5% 和 2.1%,这得益于 CEA 注意力的效果。同时,本文提出的模型在使用 CGHead 时,也表现出了更优的性能,D00,D20,D40 裂缝类型的 AP50 结果均有大幅提升,且参数量降低了 27%,计算量降低了 38%。通过实验结果显示,CGHead 检测头在轻量化和精度的平衡上有着更明显的优势。

表 6 检测头对比实验
Tab. 6 Comparison of detect head results

模型	D00	D10	D20	D40	参数量/ M	计算量/ G
	AP50/%	AP50/%	AP50/%	AP50/%		
YOLOv8n	68.7	61.1	67.9	49.7	3.01	8.1
YOLOv8n+CGHead	69.2	63.2	67.9	50.3	2.42	5.7
Our model+YOLOHead	69.8	64.5	69.5	50.9	2.14	6.5
Our model+CGHead	70.6	64.5	69.3	51.3	1.56	4.0

5.7 可视化分析

如图 11 所示展示了多种场景下的检测结果,可以看到,在多种场景如光照条件差、裂缝特征复杂、以及裂缝特征不明显的情况下,YOLOv8n 模型出现了漏检、误检等问题。相比之下经过改进后的本文模型体现出了较好的性能,能适应多个场景下的路面裂缝检测任务。而从热力图可以看到,本文提出的模型对于裂缝的分布更加聚焦,能够自适应特征并扩展注意力至裂缝形状,加强了检测效果。

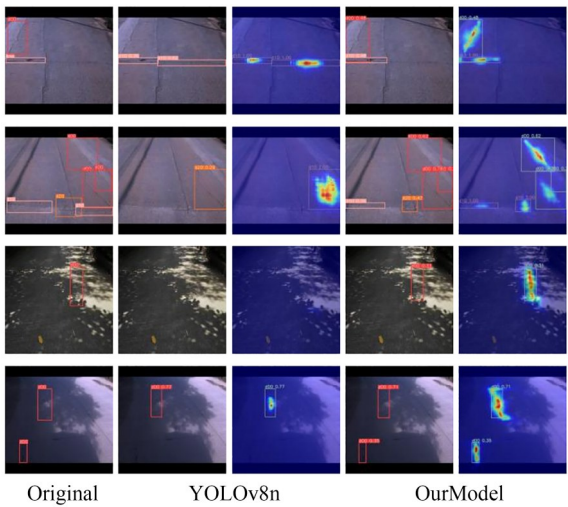


图 11 检测结果对比
Fig. 11 Detection result comparison

6 模型部署与测试

为评估本文所提算法在真实环境下的工程适用性与实时性能,团队构建了一套车载移动检测平台,并开展了现场部署与应用测试。

6.1 边缘计算设备与检测平台

该实验平台集成了图像采集、边缘计算与传感模块,其整体架构如图 12 所示。数据采集核心为一台高速矩阵相机,其分辨率设定为 640×640。数据处理端采用了一台低功耗边缘计算单元如图 12(b)所示,其硬件规格为: Intel i3-10105F CPU, 功耗 65 W; DDR4 8 GB RAM; NVIDIA GeForce RTX A2000 GPU(6 GB 显存), 功耗 70 W。操作系统为 Ubuntu 22.04.3 LTS, 集成了 CUDA 11.7, Python 3.9, PyTorch 1.13.0, 同时为了方便操作,使用 Python 构建了可视化界面如图 12(c)所示。为记录检测结果的地理空间定位,检测平台还同步集成了里程计与 GPS 模块。以上所有车载设备均由车载电源经稳压转换后统一供电。将本文提出的模型部署于该平台上,经过测试,RTX A2000 GPU 在工程作业中理论最大功耗为 70 W; 最大显存 6 GB。当 batch 设置为 1 时,显存仅占用 0.9 G, FPS 达到 101 FPS, 当 batch 设置为 4 时,显存占用 2.5 GB, FPS 达到 155 FPS, 当 batch 设置为 16 时,显存占用 5 GB, 此时显卡充分发挥并行计算性能,达到功耗满载 70 W, FPS

达到 191 FPS。均满足路面裂缝检测中 >60 FPS 的需求。因此,本文提出的模型在精度和轻量化

方面有着良好的平衡,易部署于边缘设备,可以满足路面裂缝检测的工程需求。

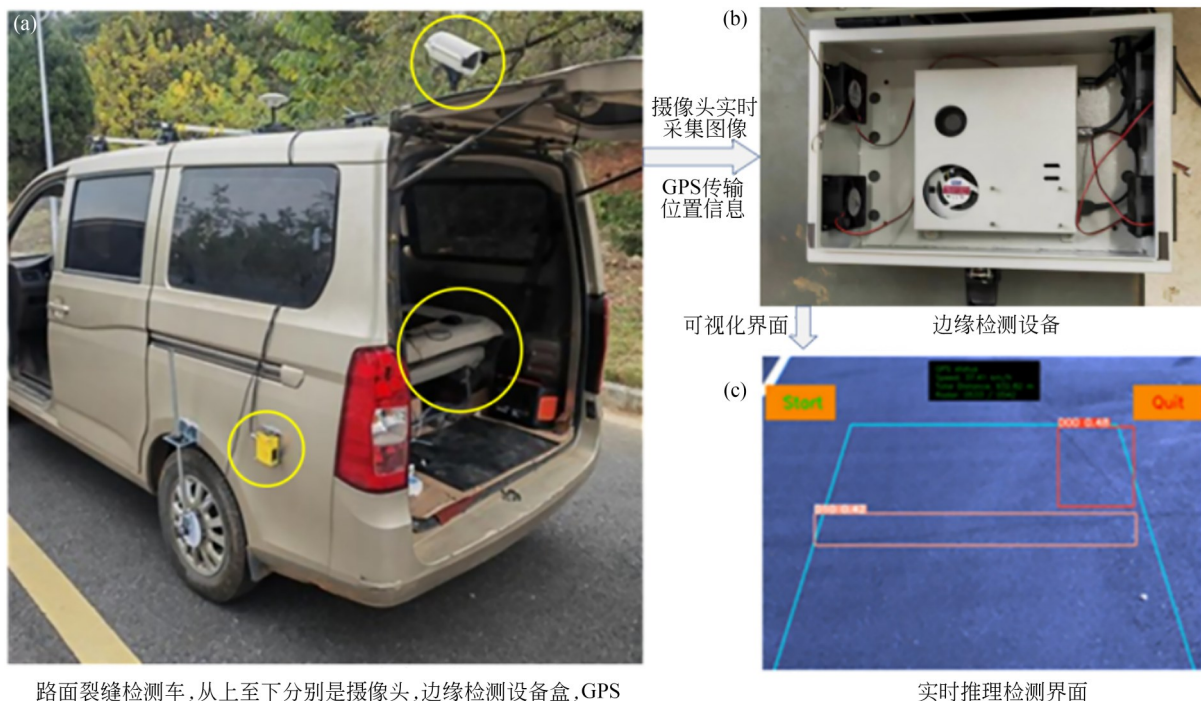


图 12 路面裂缝检测平台部署方案

Fig. 12 Deployment plan for road crack detection

7 结 论

针对路面裂缝形态复杂、易受环境干扰,且检测存在精度与轻量化不平衡等问题,本文提出一种自适应特征的轻量化路面裂缝检测方法。首先,针对路面裂缝的特征提出了CEA模块,该模块通过压缩并融合水平与垂直方向的语义,增强模型对大跨度细微目标的关注能力;其次引入HGNet并使用GSConv进行轻量化改进,改进后的HGNet_GS结构作为YOLOv8n的骨干网络,并将CEA模块嵌入到骨干中,实现了精度与轻量化的平衡;在Neck部分,构建了动态采样金字塔DSFPN,其引入了Dysamle进行动态特征下采样,并设计了SDMod将提取到的信息进一步细分再加工,同步展开主要特征与边缘特征的双路径提取和整合,以解决目标异构性的问题。在检测头中,优化了原检测头存在的不足,通过使用组卷积并嵌入本文提出的CEA模块提升了模型精度和效率。最后用PI-

OU替换原有的CIU模块,解决框锚放大问题并加快小模型的收敛速度。实验结果表明,本文改进的算法相比其他算法有着更好的轻量化与精度之间平衡。与基准模型YOLOv8n相比,本文模型参数量和计算量分别减少了50%和52%。在自建数据集上,mAP50和mAP95分别提升了5.4%和4.1%;在公开的RDD2022数据集上,mAP50和mAP95分别提升了2.1%和1.5%。改进后的算法能够适应多个复杂场景下的路面裂缝检测需求,且对比各个模型参数量和计算量都有显著的下降,更适用于边缘设备的路面裂缝检测使用场景。

作者贡献声明:

刘媛媛:论文审核与编辑;
朱凯:研究方法的提出,论文构思和撰写;
顾志辉、岳猛、王靖智:数据初步整理,数据集制作,算法对比分析;
朱路:技术支持和学术指导。

参考文献:

- [1] 王安政, 党建武, 岳彪, 等. 基于位置信息和注意力机制的路面裂缝检测[J]. 计算机工程, 2024, 50(4): 303-312.
- WANG A Z, DANG J W, YUE B, *et al.* Road crack detection based on position information and attention mechanism [J]. *Computer Engineering*, 2024, 50(4): 303-312. (in Chinese)
- [2] 蓝章礼, 徐元通, 赵胜薇, 等. 基于Sobel算子桥接的双编码器路面裂缝检测网络[J]. 重庆交通大学学报(自然科学版), 2024, 43(9): 18-24, 33.
- LAN Z L, XU Y T, ZHAO S W, *et al.* Dual encoder pavement crack detection network based on sobel operator bridging[J]. *Journal of Chongqing Jiaotong University (Natural Sciences)*, 2024, 43(9): 18-24, 33. (in Chinese)
- [3] YU M, WU D C, RAO W Y, *et al.* Automated road crack detection method based on visual transformer with multi-head cross-attention [C]. 2022 IEEE International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC). August 5-7, 2022, Chongqing, China. IEEE, 2022: 328-332.
- [4] 李禹莹, 宋伟东, 孙尚宇, 等. 基于改进Swin-Transformer的农村路面裂缝检测算法[J]. 北京交通大学学报, 2024, 48(5): 88-97.
- LI Y X, SONG W D, SUN S Y, *et al.* Rural pavement crack detection algorithm based on improved Swin-Transformer[J]. *Journal of Beijing Jiaotong University*, 2024, 48(5): 88-97. (in Chinese)
- [5] 高明星, 蒋正发, 张林, 等. 基于双金字塔网络的航拍图像路面裂缝检测方法[J]. 激光与光电子学进展, 2024, 61(22): 2212002.
- GAO M X, JIANG Z F, ZHANG L, *et al.* Detection method of pavement cracks in aerial images based on double pyramid network[J]. *Laser & Optoelectronics Progress*, 2024, 61(22): 2212002. (in Chinese)
- [6] NING Z P, WANG H, LI S L, *et al.* YOLOv7-RDD: a lightweight efficient pavement distress detection model[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2024, 25(7): 6994-7003.
- [7] 邢仁琦, 杨怀志, 薄一军, 等. 基于轻量化EfficientNet的小目标裂缝检测算法[J]. 北京交通大学学报, 2023, 47(2): 114-121.
- XING R Q, YANG H Z, BO Y J, *et al.* Small target crack detection based on lightweighted EfficientNet [J]. *Journal of Beijing Jiaotong University*, 2023, 47(2): 114-121. (in Chinese)
- [8] SU P, HAN H Z, LIU M, *et al.* MOD-YOLO: Rethinking the YOLO architecture at the level of feature information and applying it to crack detection [J]. *Expert Systems with Applications*, 2024, 237: 121346.
- [9] 刘光辉, 陈健, 孟月波, 等. 联合线性引导与网格优化的混凝土裂缝分割[J]. 光学精密工程, 2024, 32(2): 286-300.
- LIU G H, CHEN J, MENG Y B, *et al.* Concrete crack segmentation combined with linear guidance and mesh optimization [J]. *Opt. Precision Eng.*, 2024, 32(2): 286-300. (in Chinese)
- [10] 夏晓华, 高霄飞, 段智威, 等. 融合语义与三维信息的路面图像逆透视变换[J]. 光学精密工程, 2025, 33(12): 1929-1939.
- XIA X H, GAO X F, DUAN Z W, *et al.* Inverse perspective mapping of pavement image combining semantic and 3D information [J]. *Opt. Precision Eng.*, 2025, 33(12): 1929-1939. (in Chinese)
- [11] 朱世松, 高虹, 芦碧波, 等. 基于改进YOLOv8n的城市道路病害检测算法[J]. 液晶与显示, 2025, 40(9): 1356-1368.
- ZHU S S, GAO H, LU B B, *et al.* Urban road defect detection algorithm based on improved YOLOv8n [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(9): 1356-1368. (in Chinese)
- [12] 贾翔宇, 张永宏, 阙希, 等. D3F-DET: 轻量化多尺度融合的路面缺陷检测算法[J]. 计算机工程与应用, 2025, 61(17): 159-170.
- JIA X Y, ZHANG Y H, KAN X, *et al.* D3F-DET: lightweight and multiscale fusion algorithm for pavement defect detection [J]. *Computer Engineering and Applications*, 2025, 61(17): 159-170. (in Chinese)
- [13] 胡鹏, 夏晓华, 钟预全. 面向嵌入式应用的路面裂

- 缝检测方法[J/OL]. 计算机科学, 1-17 [2025-09-21]. <https://link.cnki.net/urlid/50.1075.TP.20250510.1923.004>.
- HU P, XIA X H, ZHONG Y Q. Road Surface Crack Detection Method for Embedded Applications [J/OL]. *Computer Science*, 1-17 [2025-09-21]. <https://link.cnki.net/urlid/50.1075.TP.20250510.1923.004>. (in Chinese)
- [14] XU W, WAN Y. ELA: Efficient Local Attention for Deep Convolutional Neural Networks [EB/OL]. 2024; *arXiv*: 2403.01123. <https://arxiv.org/abs/2403.01123>
- [15] ZHAO Y A, LV W Y, XU S L, *et al.* DETRs beat YOLOs on real-time object detection [C]. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 16-22, 2024, Seattle, WA, USA. IEEE, 2024: 16965-16974.
- [16] CHOLLET F. Xception: deep learning with depthwise separable convolutions [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 1800-1807.
- [17] LI H L, LI J, WEI H B, *et al.* Slim-neck by GSConv: A Lightweight-Design for Real-Time Detector Architectures [EB/OL]. 2022; *arXiv*: 2206.02424. <https://arxiv.org/abs/2206.02424>
- [18] LIU W Z, LU H, FU H T, *et al.* Learning to upsample by learning to sample [C]. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 1-6, 2023, Paris, France. IEEE, 2024: 6004-6014.
- [19] ZHU X Z, HU H, LIN S, *et al.* Deformable ConvNets V2: more deformable, better results [C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 9300-9308.
- [20] KANOPOULOS N, VASANTHAVADA N, BAKER R L. Design of an image edge detection filter using the Sobel operator[J]. *IEEE Journal of Solid-State Circuits*, 1988, 23(2): 358-367.
- [21] XIE S N, GIRSHICK R, DOLLÁR P, *et al.* Aggregated residual transformations for deep neural networks [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 5987-5995.
- [22] LIU C, WANG K G, LI Q, *et al.* Powerful-IoU: More straightforward and faster bounding box regression loss with a nonmonotonic focusing mechanism[J]. *Neural Networks*, 2024, 170: 276-284.
- [23] 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: gdj3.260.
- [24] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023: 7464-7475.
- [25] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9: learning what you want to learn using programmable gradient information [C]. *Computer Vision-ECCV 2024*. Cham: Springer, 2025: 1-21.
- [26] CHEN H, CHEN K, DING G G, *et al.* YOLOv10: real-time end-to-end object detection [C]. *Advances in Neural Information Processing Systems 37*. December 10-15, 2024. Vancouver, BC, Canada. *Neural Information Processing Systems Foundation, Inc. (NeurIPS)*, 2024: 107984-108011.
- [27] HU J, SHEN L, SUN G. Squeeze-and-excitation Networks [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [28] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module [C]. *Computer Vision-ECCV 2018*. Cham: Springer, 2018: 3-19.
- [29] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C].

2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 20-25,

2021, *Nashville, TN, USA*. IEEE, 2021: 13708-13717.

作者简介:



刘媛媛(1978—),女,江西永新人,教授,硕士生导师,2007年于华东交通大学获得硕士学位,加拿大 Dalhousie 大学访问学者(2014级),主要从事机器视觉、图像处理、模式识别等方面的研究。Email:lyy.78@163.com



朱路(1976—),男,江西永新人,博士,教授,硕士生导师,2009年于华中科技大学获得博士学位,主要从事机器视觉、物联网技术、图像处理、超构表面及优化设计等方面的研究。E-mail:luyuanwanwan@163.com