

文章编号:1007-2780(XXXX)XX-0001-16

WADW-YOLO:面向航拍的轻量化道路缺陷检测

孙瑞, 燕芳*, 李嘉辉, 向天明

(内蒙古科技大学自动化与电气工程学院, 内蒙古包头 014000)

摘要:针对无人机航拍道路检测中缺陷尺度差异大、背景干扰强且检测模型复杂度高等问题,本文提出一种基于YOLOv11n的改进检测模型WADW-YOLO。首先,在主干网络中提出C3k2-WC加权卷积模块,通过预定义的中心增强权重矩阵实现卷积核的空间位置加权,在降低模型参数数量的同时聚焦缺陷区域特征提取;其次,采用轻量化ADown下采样模块,提升对缺陷细节的保留能力。在主干末端引入DAttention可变形注意力机制,通过自适应调节感受野来捕捉全局与局部信息;最后,设计Wise-ShapeIoU损失函数,将Wise-IoU的非单调聚焦机制与ShapeIoU的几何形状约束相结合,提升检测性能。WADW-YOLO在公开数据集China_Drone上的mAP@0.5达到74.6%,相较YOLOv11n提升3.8%,且参数量降低17.8%,计算量减少23.8%,实现检测精度和模型轻量化的有效平衡。WADW-YOLO在China_Motor-Bike数据集上的mAP@0.5较YOLOv11n提升了1.8%,展现了WADW-YOLO突出的泛化性能。本文为无人机道路缺陷检测提供了一种高精度、低功耗部署方案的技术参考。

关键词:无人机;道路缺陷;YOLOv11n;WADW-YOLO;轻量化

中图分类号:TP391.4 **文献标识码:**A **doi:**10.37188/CJLCD.2026-0065 **CSTR:**32172.14.CJLCD.2026-0065

WADW-YOLO: lightweight road defect detection for aerial photography

SUN Rui, YAN Fang*, LI Jiahui, XIANG Tianming

(School of Automation and Electrical Engineering, Inner Mongolia University of Science and Technology, Baotou 014000, China)

Abstract: To address the problems of large scale variation of defects, strong background interference and high complexity of detection models in UAV aerial road inspection, this paper proposes an improved detection model WADW-YOLO based on YOLOv11n. Firstly, a C3k2-WC weighted convolution module is proposed in the backbone network, which realizes spatial position weighting of convolution kernels through a predefined center-enhanced weight matrix, reducing the number of model parameters while focusing on feature extraction of defect regions. Secondly, a lightweight ADown downsampling module is adopted to improve the ability of retaining defect details, and a DAttention deformable attention

收稿日期:2026-04-14;**修订日期:**2026-05-08.

基金项目:内蒙古自然科学基金项目(No.2025MS06002);内蒙古自治区重点研发和成果转化计划项目(No.2025SYF-HH0062)

Supported by Natural Science Foundation of Inner Mongolia Autonomous Region (No.2025MS06002); Key Research and Development and Achievement Transformation Program Project of Inner Mongolia Autonomous Region (No.2025SYFH0062)

*通信联系人, E-mail:0472yanfang@163.com

mechanism is introduced at the end of the backbone to capture both global and local information by adaptively adjusting the receptive field. Finally, a Wise-ShapeIoU loss function is designed by combining the non-monotonic focusing mechanism of Wise-IoU and the geometric shape constraint of ShapeIoU to improve detection performance. WADW-YOLO achieves 74.6% mAP@0.5 on the public China_Drone dataset, which is 3.8% higher than that of YOLOv11n, with 17.8% fewer parameters and 23.8% less computation, achieving an effective balance between detection accuracy and model lightness. WADW-YOLO also obtains a 1.8% improvement in mAP@0.5 over YOLOv11n on the China_MotorBike dataset, demonstrating its outstanding generalization performance. This paper provides a technical reference for high-precision and low-power deployment schemes in UAV road defect detection.

Key words: UAV; road defect; YOLOv11n; WADW-YOLO; lightness

1 引 言

随着城市化进程的加速推进,城市道路因长期承受车辆负荷累积而出现的裂缝、坑洞和凹坑等损伤,不仅缩短了道路使用寿命、大幅增加养护成本,且威胁行车安全。不良路况往往是导致交通事故的潜在因素,因此快速而准确地检测道路损伤对保障行车安全和社会经济发展具有重要意义。

目前的道路缺陷检测方法可分为传统目标检测方法和基于深度学习的目标检测方法。传统道路缺陷检测方法,如人工目视巡检、机械化检测等,依赖人工设计特征,难以应对复杂多变的实际道路场景,检测精度和鲁棒性较差,难以满足规模化养护的实际需求^[1]。

以深度学习为基础的目标检测算法为影像识别道路缺陷检测提供了有效解决方案^[2],在目标检测领域,学者提出多种目标检测算法,其中包括单次多框检测器(single shot multibox detector, SSD)^[3]、快速区域卷积神经网络 Faster R-CNN^[4]、和 YOLO(you only look once)^[5]算法。其中 YOLO 系列算法因兼具检测精度与实时性,成为该领域的研究热点。诸多研究通过改进 YOLO 系列算法提升道路缺陷检测性能:周晴^[2]等采用轻量化网络模型 Mobile Net-V3 为主干网络加强了对缺陷特征提取能力;徐光宪等^[6]在 YOLOv8 模型基础上引入全维度动态卷积与 Powerful-IoU 损失函数,提升了无人机影像中小目标缺陷的检测精度,对于设备算力需求为 8.1GFLOPs;龙伍丹^[7]等设计了特征融合模块 Res-RFB,以提升道路缺陷的表征能力,但导致

模型参数量较大(6.93M);张颖^[8]等针对混合路面的坑槽和裂缝两种混合路面缺陷,通过引入 AKConv 卷积模块提升了模型对不同道路缺陷的特征提取能力,但改进后模型仍需算力为 20.2GFLOPs;李勇^[9]等设计 LFDS-YOLO 模型,实现了多尺度特征融合与注意力机制改进的轻量化路面缺陷检测,并提升了检测精度;朱世松^[10]等通过在主干网络中引入多层级感知注意力机制来增强缺陷特征表达能力,使模型专注于缺陷区域。Li^[11]等通过将多核 ConvFFN(MCFN)与 C3K2 模块融合构成 C3K2-MCFN 结构,以捕捉不同感受野间的局部依赖关系,进一步提升了特征表达能力。

上述大部分改进模型中虽提升了识别检测精度,同时模型参数量与算力需求也随之增大。近年来,无人机航拍技术凭借其安全高效、可大范围巡检且无需接触道路表面的独特优势。航拍图像与深度学习技术相结合应用于道路损伤检测领域的研究层出不穷,然而该技术受限于航拍图像背景复杂、道路缺陷尺度差异大、小目标占比高、背景干扰强等因素,加之无人机平台嵌入式硬件的算力约束,导致其检测效果仍难以满足工程实际需求。本文针对上述挑战,提出一种基于 YOLOv11 的改进轻量化模型 WADW-YOLO,以解决无人机航拍道路缺陷检测任务中的目标检测精度较低、漏检率高等问题。

该模型通过以下四个方面改进实现了航拍场景下道路缺陷的检测性能的提升:

(1)提出轻量化 C3k2-WC 模块,通过引入加权卷积构成 C3k2-WC 代替 C3k2 特征提取模块,使模型在降低参数量的同时增强对缺陷特征提

取能力。

(2)采用轻量化 ADown 替换原 Conv 下采样,在降低模型复杂度的同时并减少下采样过程中信息损失。

(3)在主干网络末端引入可变形注意力 DAttention 替换 C2PSA,通过自适应调节感受野捕捉全局与局部特征。

(4)设计 Wise-ShapeIoU 损失函数,将 Wise-IoU 聚焦机制与 ShapeIoU 形状约束结合,动态调节样本梯度权重,提升模型对不规则缺陷的定位精度。

2 YOLOv11 算法原理

YOLOv11^[12]由 Ultralytics 团队于 2024 年 9 月发布,作为 YOLO 系列的最新迭代版本,其以“高精度、轻量化、易部署”为核心设计理念,相较于前代模型优势显著。该模型在保证检测精度的前提下,大幅缩减了参数量,降低对硬件计算资源的依赖,无需高性能 GPU 即可稳定运行,更适配边缘设备等资源受限场景,同时在小目标检测精度、多尺度特征建模能力和推理延迟上均有优化,兼顾实时性与检测效果,具备突出的工程

落地价值。

YOLOv11 延续了 YOLO 系列“骨干网络-颈部网络-检测头”的三级架构,核心改进集中在三大模块:骨干网络中用 C3k2 模块替代 C2f 模块,通过双分支结构设计及与重参数化技术的协同,提升特征提取的灵活性、效率并降低参数量;颈部网络提出 C2PSA 模块,融合 CSP 结构与 PSA 自注意力机制,增强多尺度特征建模能力;检测头分类分支引入深度可分离卷积替代传统卷积,结合逐通道卷积与 1×1 点卷积,在提升性能的同时显著降低计算复杂度,所有改进均经过工程验证,兼顾创新性与部署可行性。相对比前代模型,YOLOv11 在速度、精度与效率之间取得了更优的平衡。

3 WADW-YOLO 目标检测算法

图 1 为本文提出的一种基于 YOLOv11n 基础架构改进的道路缺陷检测算法(WADW-YOLO)的网络结构。

该模型在原始模型中的特征提取、下采样、注意力机制及损失函数四个层面均进行了改进。在主干网络中用 C3k2-WC 模块替代 C3k2,通过

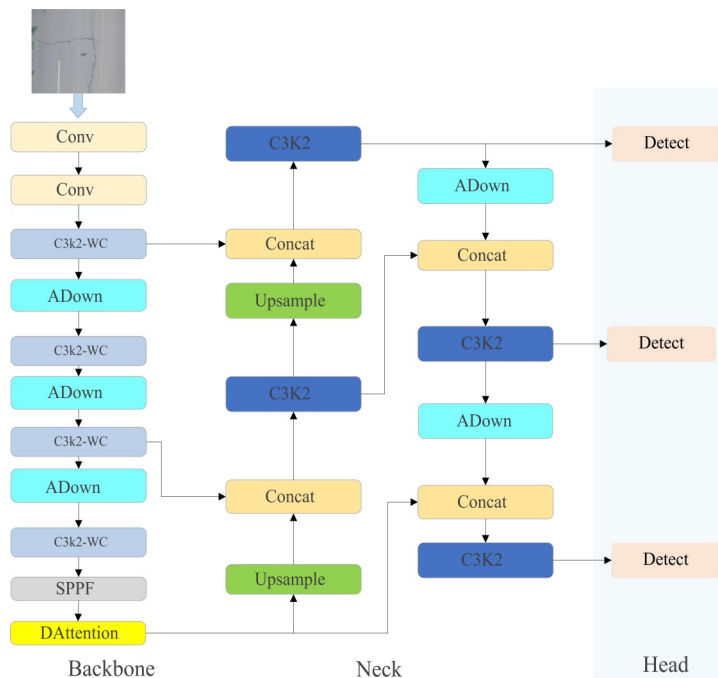


图 1 WADW-YOLO 网络结构

Fig. 1 Network architecture of WADW-YOLO model

中心增强的空间加权策略机制,模型可以聚焦于缺陷区域中心;下采样环节采用ADown模块,利用非对称卷积与通道重排降低信息损耗,保留更多边缘细节;网络末端嵌入DAttention可变形注意力,动态选择关键区域并调节感受野,以较低开销兼顾全局与局部纹理;边界框回归通过设计Wise-ShapeIoU损失,融合非单调权重与形状约束,优化困难样本的梯度分配。上述改进模块在保持算法轻量化的同时增强对裂缝、坑洼等不规则缺陷的特征提取能力,通过上述模块的协同作用增强了该模型在复杂道路场景下的适应性。

3.1 C3k2-WC 模块

无人机航拍获取的道路图像中,路面纹理与裂缝、坑槽等缺陷区域在灰度特征上具有高度相似性,导致缺陷边界模糊、辨识度不足。同时,道路标线、修补痕迹等干扰物与真实缺陷在视觉表现上难以区分,进一步加大了检测难度。此外,航拍过程中受日照角度、天气状况、飞行高度及视角变化等因素影响,图像易出现光照不均、阴影遮挡、尺度变换等问题,均对缺陷特征的稳定

提取构成挑战。YOLOv11n网络中的C3k2模块使用传统卷积,权重共享,缺乏空间差异化能力,难以聚焦缺陷区域和抑制背景干扰。这种卷积方式处理道路缺陷时,特征表达针对性和判别性不足。

为此,本文提出C3k2-WC模块,将wConv2d加权卷积嵌入C3k2特征提取单元。wConv2d采用固定空间权重矩阵对卷积核进行调制,无需额外学习参数,计算开销极低。该设计在保持了模型的轻量特性的同时,增强中心区域特征响应。适用于无人机平台的实时检测需求。通过空间维度上的权重调制,C3k2-WC能够提升卷积核对缺陷中心的敏感度,从而增强裂缝、坑槽等细微缺陷的辨识能力。C3k2-WC的结构如图2所示。

其中wConv2d加权卷积^[13]是一种空间位置加权机制,通过在标准卷积核上引入预定义的中心增强权重矩阵,基于空间位置的重要性对卷积核进行差异化加权,使模型能够增强中心区域的特征响应,从而在保持较低计算开销的同时提升对关键特征的提取能力。

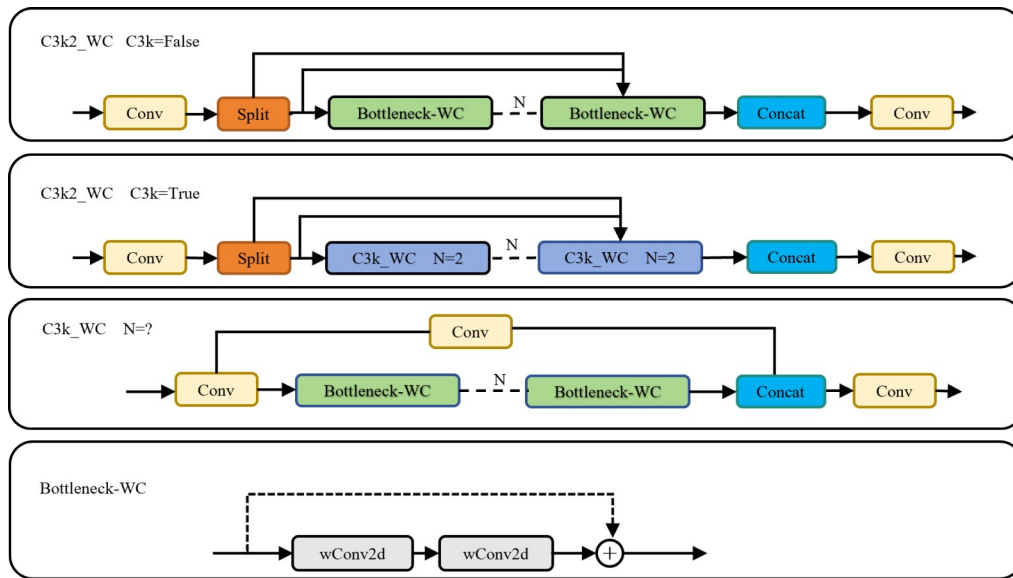


图2 C3k2_WC模块结构

Fig. 2 The structure of C3k2_WC model

如图3所示,wConv2d加权卷积通过密度参数den生成空间权重矩阵 Φ ,实现对卷积核的空间位置加权调制。首先设密度参数为 $den = [d_1, d_2, \dots, d_k]$,构造对称衰减向量 α 如式(1)所示:

$$\alpha = [d_1, d_2, \dots, d_k, 1.0, d_k, \dots, d_2, d_1] \in \mathbb{R}^{2k+1}. \quad (1)$$

如式(1)中所示,其中心位置取值为1.0,两侧通过对称翻转den实现中心聚焦的权重分布。den=[0.9]为例,生成的向量为: $\alpha = [0.9, 1.0, 0.9]$ 。然后构建空间权重矩阵。基

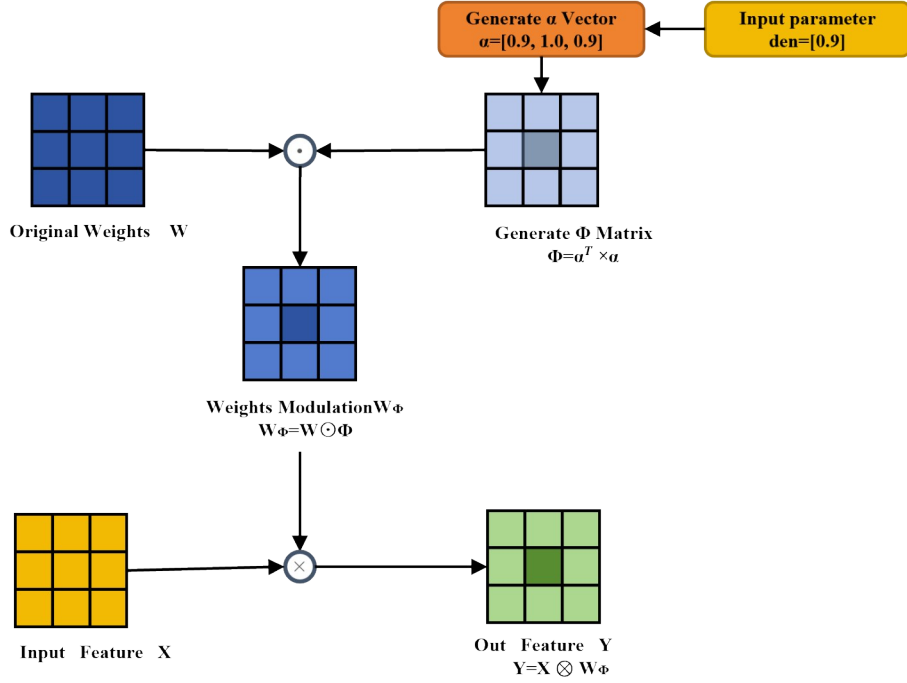


图3 wConv2d 模块结构

Fig. 3 The structure of wConv2d model

于向量 α , 通过外积运算构建空间权重矩阵 Φ 如式(2)所示:

$$\Phi = \alpha^T \times \alpha \in \mathbb{R}^{(2k+1) \times (2k+1)}. \quad (2)$$

其对应矩阵元素计算如式(3)所示:

$$\Phi_{i,j} = \alpha_i \cdot \alpha_j, \quad i, j \in \{1, 2, \dots, 2k+1\}. \quad (3)$$

当 $\text{den}=[0.9]$ 时, 生成的 3×3 空间权重矩阵:

$$\Phi = \begin{bmatrix} 0.9 \\ 1.0 \\ 0.9 \end{bmatrix} \begin{bmatrix} 0.9, 1.0, 0.9 \end{bmatrix} = \begin{bmatrix} 0.81 & 0.90 & 0.81 \\ 0.90 & 1.00 & 0.90 \\ 0.81 & 0.90 & 0.81 \end{bmatrix}$$

该矩阵呈现中心增强、四周衰减的分布模

式, 中心权重为 1.00, 边缘权重为 0.81, 形成中心聚焦的空间权重分布。为了更为直观了解这一矩阵特性, 当密度函数为 0.9, 0.8, 0.7 时其生成空间权重矩阵热力图分别为下图 4 所示。

对卷积核进行空间权重调制, 设原始卷积核权重为 $W \in \mathbb{R}^{C_{out} \times C_{in} \times (2k+1) \times (2k+1)}$, 通过 Hadamard 积(逐元素乘法)将空间权重矩阵 Φ 调制到卷积核上, 如式(4)所示:

$$W_{\phi} = W \odot \Phi, \quad (4)$$

式中, \odot 为 Hadamard 积。调制后的卷积核 W_{ϕ} 具

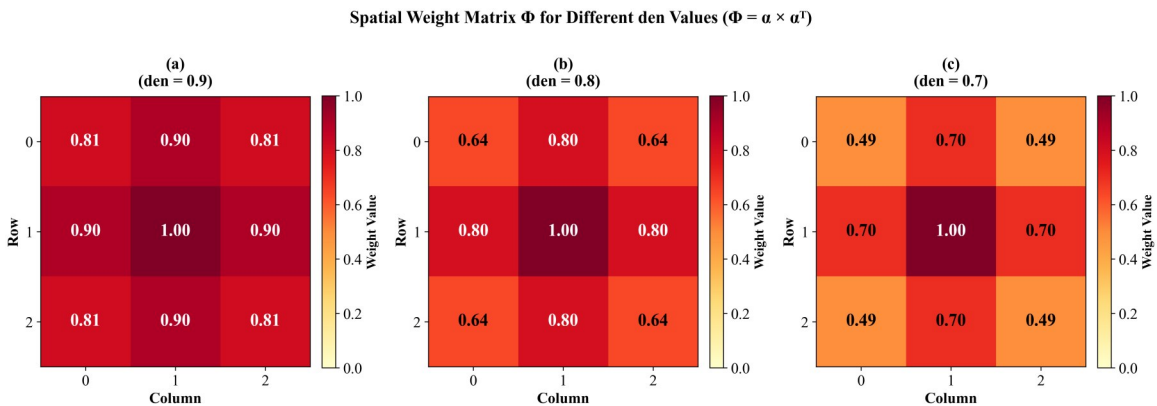


图4 不同密度函数生成的空间权重矩阵:(a)den=0.9 (b)den=0.8 (c)den=0.7。

Fig. 4 Spatial weight matrices generated by different density functions: (a) den=0.9 (b) den=0.8 (c) den=0.7.

备了空间位置相关的权重分布,使得中心区域获得更高响应,边缘区域响应相对抑制。

最后进行卷积运算,对于输入特征图 $x \in \mathbb{R}^{N \times C_{in} \times H \times W}$,使用调制后的权重 W_{\oplus} 执行标准卷积运算公式如式(5)所示:

$$Y = X \otimes W_{\oplus} + b, \quad (5)$$

其中: \otimes 为卷积运算, b 为偏置项,输出特征图 $Y \in \mathbb{R}^{N \times C_{out} \times H' \times W'}$ 。

3.2 ADown下采样模块

在无人机道路巡检的应用场景中,路面缺陷通常表现为细微裂缝或局部凹陷等形态,具有目标尺度较小、空间分布不均等特点,对检测算法的精度与实时性均构成了显著挑战。传统的空间降维方法多采用步长较大的标准卷积或池化操作,虽然可以降低特征图分辨率,但不可避免地会造成部分像素信息的丢失,导致关键特征难以充分捕捉,尤其在识别尺度较小或对比度较低的缺陷区域时,会显著降低检测性能。针对上述问题,本

文引入了一种轻量级下采样模块 ADown^[14],以替代 YOLOv11n 架构中原有的标准卷积下采样组件,其结构如图 5 所示。ADown 单元构建了一种基于平均池化与最大池化并行处理的特征提取框架。该框架首先对输入特征图执行平均池化操作,以实现空间尺度压缩,从而保留关键的语义信息。随后,特征数据被分流至两条独立的处理分支:第一条分支采用 3×3 卷积核进行特征变换,以降低计算复杂度;第二条分支则通过 3×3 最大池化操作提取更精细的局部特征,以弥补卷积操作可能造成的信息损失,并利用 1×1 卷积进行特征融合。最终,通过特征拼接机制将双路输出集成,在达成降维目标的同时最大限度地保留特征表达能力,该模块有效抑制了微小缺陷目标在特征编码过程中出现的细节信息退化现象,显著提升了模型在缺陷检测任务中的识别准确率与鲁棒性。即便在缺陷区域与背景高度混杂的复杂场景下,该结构仍能保持稳定的检测性能。

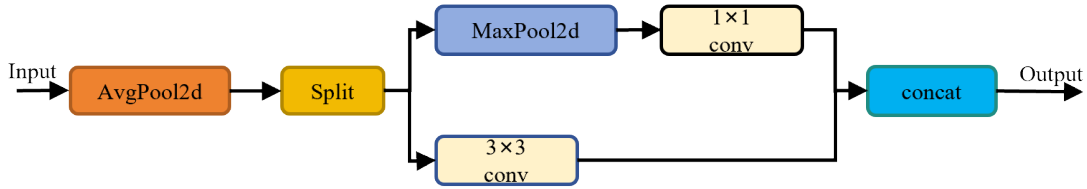


图5 ADown下采样模块结构图

Fig. 5 ADown downsampling module structure diagram

3.3 DAttention注意力机制

在无人机道路缺陷检测任务中,航拍视角下的缺陷目标呈现出显著的多尺度特性与不规则形态。裂缝、坑洞等缺陷往往表现为细长的线性结构或离散的块状区域,且与路面纹理、阴影及遮挡物存在高度相似的视觉特征。传统卷积神经网络受限于局部感受野,难以有效建模全局上下文关系;而标准自注意力机制虽能捕获长距离依赖,但其计算复杂度与特征图尺寸呈二次增长,在处理高分辨率无人机影像时面临严重的计算瓶颈。此外,规则网格采样策略难以自适应地覆盖形态各异的不规则缺陷区域,导致细粒度特征提取能力不足。

为应对上述挑战,本文引入可变形注意力机制(Deformable Attention)^[15]作为特征增强模块,通过稀疏自适应采样策略在降低计算开销的同

时提升对不规则目标的建模能力。如图 6 所示,给定输入特征图 $x \in \mathbb{R}^{H \times W \times C}$,针对输入特征图 $x \in \mathbb{R}^{H \times W \times C}$,构建一个与原始空间维度相关联的离散采样网格 $p \in \mathbb{R}^{H_G \times W_G \times 2}$,该网格通过缩放系数 r 对原图进行下采样得到:

$$H_G = \frac{H}{r}, \quad W_G = \frac{W}{r}. \quad (6)$$

随后将网格坐标映射至归一化区间 $[-1, 1]$,为获取空间变形量,首先利用投影矩阵 W_q 将输入特征转换为查询向量 $q = xW_q$,继而将其馈入轻量化的偏移预测网络 $\theta_{\text{offset}}(\cdot)$ 以估计位置偏移:

$$\Delta p = \theta_{\text{offset}}(q). \quad (7)$$

基于变形后的坐标对特征图进行重采样,得到键和值的特征表示,相应的线性变换定义为:

$$q = xW_q, \quad \tilde{k} = \tilde{x}W_k, \quad \tilde{v} = \tilde{x}W_v, \quad (8)$$

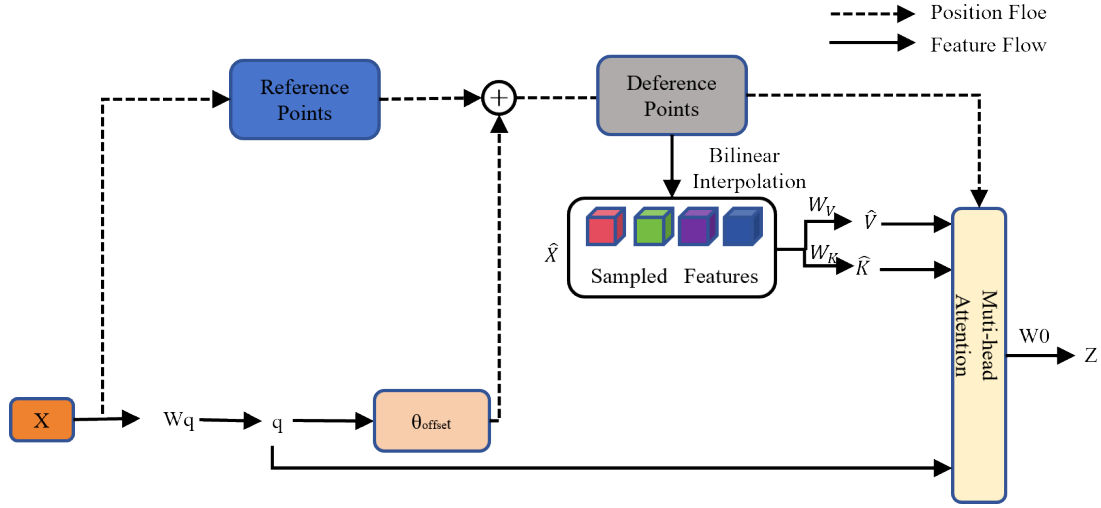


图6 可变注意力机制模块

Fig. 6 Module of deformable attention mechanism

$$\Delta p = \theta_{\text{offset}}(q), \quad \tilde{a} = \phi(x; p + \Delta p). \quad (9)$$

此处 \tilde{k} 与 \tilde{v} 分别表征经几何变形后的键嵌入与值嵌入。采样算子 $\phi(\cdot)$ 采用可微分的双线性插值实现,其数学表达如式(10)所示:

$$\begin{aligned} \phi(z; p_x, p_y) = & \sum_{(r_x, r_y)} g(p_x - r_x)g(p_y - r_y) \\ & z[y + r_x, x + r_y, :], \end{aligned} \quad (10)$$

式中插值核函数 $g(a, b) = \max(0, 1 - |a - b|)$ 用来得出两个位置的关联性, (r_x, r_y) 为特征图 $z \in \mathbb{R}^{H \times W \times C}$ 上的所有位置建立索引。鉴于核函数的紧支集特性,仅当 g 在最接近位于 (p_x, p_y) 的四个积分点内为非零,故上述求和简化为四点加权平均。对 q, \tilde{k}, \tilde{v} 实施多头注意力计算,并融入相对位置编码 R 以捕获空间结构信息。第 m 个注意力头的输出形式化如式(11)所示:

$$z^{(m)} = \sigma \left(\frac{q^{(m)} \cdot \tilde{k}^{(m)T}}{\sqrt{d}} + \phi(\hat{B}; R) \right) \tilde{v}^{(m)}, \quad (11)$$

式中 $\phi(\hat{B}; R) \in \mathbb{R}^{H \times W \times H_c \times W_c}$ 。最终,将各头输出沿通道维度拼接,并经投影矩阵 W 。线性变换得到聚合特征 z 。该机制在降低计算复杂度的同时,显著增强了对道路缺陷这类不规则目标的特征提取能力。

3.4 Wise-ShapeIoU 损失函数

在目标检测任务中,损失函数主要用于衡量模型预测值与真实值之间的偏差,通过最小化损失函数,模型能够更高效地完成目标检测任务的

学习。YOLO11在分类任务中采用的损失函数是 CIoU^[16],其原理是根据预测框和真实框之间的中心点距离和宽高比差异,来提供有效的梯度信息,进而提高边界框回归的精度。然而 CIoU 无法针对不同目标进行自适应调整,难以适配尺度差异大和形态多变的目标,YOLOv11在道路缺陷检测任务中,针对多尺度及复杂形态的小目标缺陷时其回归精度存在不足,并且目标框的定位也存在一定偏差。因此本文将 Wise-IoU V3^[17] 和 ShapeIoU^[18] 两者思想相结合并设计 Wise-ShapeIoU 作为边界框损失函数,以提升回归精度和增强定位准确性。形状损失函数通过引入长宽比惩罚项,解决边界框回归中长宽比不一致的问题。首先计算权重系数 w 和 h 其计算公式分别为式(12)与式(13)所示:

$$w = \frac{2 \times (w_{\text{gt}})^{\text{scale}}}{(w_{\text{gt}})^{\text{scale}} + (h_{\text{gt}})^{\text{scale}}}, \quad (12)$$

$$h = \frac{2 \times (h_{\text{gt}})^{\text{scale}}}{(w_{\text{gt}})^{\text{scale}} + (h_{\text{gt}})^{\text{scale}}}, \quad (13)$$

式中 w_{gt} 和 h_{gt} 分别表示真实框的宽度和高度, scale 为尺度控制参数。

基于上述权重,计算形状距离如式(14)所示:

$$\text{distance}_{\text{shape}} = \frac{hh \times (x_c - x_c^{\text{gt}})^2 + ww \times (y_c - y_c^{\text{gt}})^2}{c^2}, \quad (14)$$

式中, (x_c, y_c) 和 $(x_c^{\text{gt}}, y_c^{\text{gt}})$ 分别表示预测框和真实

框的中心点坐标, c 为两框中心点之间的距离。

最终形状损失函数定义如式(15)为:

$$\Omega_{\text{shape}} = \sum_{l=w,h} (1 - e^{-w_l})^\theta, \quad \theta = 4, \quad (15)$$

其中 w_w 和 w_h 分别是宽度和高度方向上的惩罚权重, 如式(16)与式(17)所示:

$$w_w = hh \times \frac{|w - w_{\text{gt}}|}{\max(w, w_{\text{gt}})}, \quad (16)$$

$$w_h = ww \times \frac{|h - h_{\text{gt}}|}{\max(h, h_{\text{gt}})}, \quad (17)$$

式中 w 和 h 分别表示预测框的宽度和高度。ShapeIoU 定义图如图 7 所示。其对应的边界框回归损失表达式如式(18)所示:

$$L_{\text{ShapeIoU}} = 1 - \text{IoU} + \text{distance}_{\text{shape}} + 0.5 \times \Omega_{\text{shape}}. \quad (18)$$

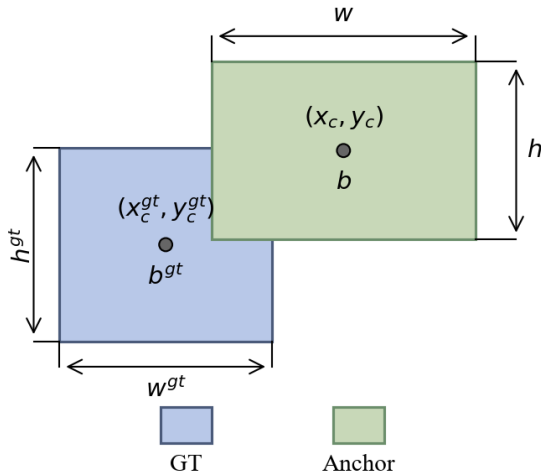


图 7 ShapeIoU 计算示意图

Fig. 7 Schematic diagram of ShapeIoU calculation

引入将 WiseIoU-V3 的动态权重策略与 ShapeIoU 的几何约束相结合, 构建为 Wise-ShapeIoU 损失函数。该损失通过引入 Wise 中的非单调聚焦机制, 自适应地调节不同质量样本的梯度贡献, 从而抑制低质量样本对模型训练的干扰。Wise-ShapeIoU 的数学表达如式(19)~(21)所示:

$$L_{\text{Wise-ShapeIoU}} = L_{\text{WiseIoU}} + \text{distance}_{\text{shape}} + 0.5 \times \Omega_{\text{shape}}, \quad (19)$$

$$L_{\text{WiseIoU}} = L_{\text{IoU}} \times r, \quad r = \frac{\beta}{\delta \alpha^{\beta - \delta}}, \quad (20)$$

$$\beta = \frac{L_{\text{IoU}}^*}{\bar{L}_{\text{IoU}}} \in [0, +\infty), \quad (21)$$

式中: r 为非单调聚焦系数; β 为锚框异常程度; α 、 δ 是超参数; L_{IoU}^* 为当前锚框的 IoU 损失; \bar{L}_{IoU} 为所有锚框 IoU 损失的均值。

ShapeIoU 部分通过引入长宽比相关的权重系数, 使模型在优化过程中对边界形状更加敏感; 而 WiseIoU 部分则依据锚框异常程度 β 计算聚焦系数 r 。检测到离群样本 (β 较大) 时自动降低其权重, 避免异常梯度主导优化方向, 反之则增强困难样本的学习信号。Wise-ShapeIoU 不仅具有 ShapeIoU 对边界框形状感知能力, 还通过 WiseIoU-V3 的聚焦机制实现具有对样本难度的动态适配能力。

4 实验结果与分析

4.1 实验数据集

本文使用 Road Defect Detection2022 (RDD2022)^[19] 开源数据集进行训练。该数据集是专门用于自动道路缺陷检测的标注图像数据集, 该数据集涵盖中国、印度、日本、捷克、美国等多国采集的道路缺陷图像总计 47 420 张。为反映中国本土道路的缺陷特征, 本文选取该数据集的中国区域数据集开展研究, 中国区域数据集包含 China_Drone 与 China_MotorBike。China_Drone 数据集由无人机航拍采集, 清洗后得到 2 396 张道路缺陷图像, 用于模型的训练与核心实验分析; China_MotorBike 数据集由车载移动设备采集, 清洗后共 1 977 张道路缺陷图像, 用于泛化性实验, 以验证模型在检测不同采集方式获取的道路场景下的泛化能力。

数据集标注的道路缺陷类型共分为 5 类, 分别为纵向裂缝 (D00)、横向裂缝 (D10)、网状裂缝 (D20)、坑洞 (D40) 与修补区域 (Repair), 各类别典型样本如图 8 所示。实验所用数据集均按照 8:1:1 的比例随机划分为训练集、验证集与测试集。

4.2 实验环境及评价指标

实验环境为 Windows 11, CPU 为 AMD Ryzen 7 5700X, GPU 为 NVIDIA GeForce RTX 5060, 显存为 8GB。编程语言为 Python 3.11, 训练框架为 PyTorch 2.11.0, 采用 CUDA 12.8 进行加速。训练参数设置如下: 训练轮次 (epochs) 为 200, 训练批次大小 (batchsize) 为 16, 输入图像尺寸为像素 640×640 像素, 优化器 (optimizer) 使

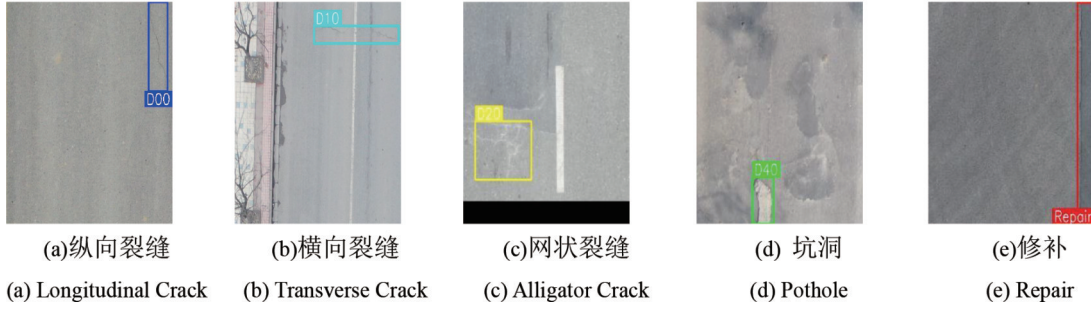


图 8 道路损伤类型样例

Fig. 8 Examples of road damage types

用随机梯度下降(SGD)算法,初始学习率为0.01,动量为0.937,权重衰减因子为0.0005。

为客观评价实验结果,本文选用以下评价指标评估模型的整体性能:精确度(Precision,P)、召回率(Recall,R)、平均精度(mean Average Precision,mAP),公式如式(22)~(25)所示:

$$P = \frac{TP}{TP + FP}, \quad (22)$$

$$R = \frac{TP}{TP + FN}, \quad (23)$$

$$AP = \int_0^1 P(R) dR, \quad (24)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i). \quad (25)$$

式中 TP 表示正确识别的正类样本数量, TN 表示正确识别的负类样本数量, FN 表示被错误识别为正类的负类样本数量, FP 表示被错误识别为负类的正类样本数量。n 为道路缺陷的类别总数。精确度 P 表示模型预测为正例的样本中真实为正例的比例。召回率 R 定义为模型正确识别为正例的样本数占所有真实正例的比例。平均准确率 mAP 是通过对所有类别的 AP 值取平均得到的,其中 AP 反映了每一个类别的预测准

确率, mAP 反应整个模型的整体性能。mAP@0.5 定义为 IOU 阈值为 0.5 时的平均精度, mAP@0.5:0.95 定义为 IOU 阈值从 0.5~0.95 计算得到的平均精度。此外,为了综合更加全面模型的检测精度、计算效率及部署可行性,增加模型参数量(Parameters)、计算量(GFLOPs)评价指标构建多维评估体系。

4.3 C3k2-WC 模块密度参数选择

为确定 C3k2-WC 模块中密度参数 den 的最优取值,本文在 YOLOv11n 基线模型上,替换主干网络中的 C3k2 为 C3k2-WC,对密度参数 den 分别在 {0.70, 0.75, 0.80, 0.85, 0.90, 0.95} 范围内进行对比实验,结果如表 1 所示。

由表 1 可见,密度参数 den 的取值对模型性能有显著影响。当 den 由 0.70 逐步增大至 0.90 时, mAP@0.5 和 mAP@0.5:0.95 指标整体呈上升趋势,并在 den=0.90 时分别达到最优值 71.4% 和 46.1%。当 den 继续增大至 0.95 时,两项指标分别降至 68.4% 和 43.9%,精度出现明显下降。可分析当 den 取值过低时,中心聚焦效果不足,难以突出缺陷区域特征;den 取值过高时,边缘区域的权重被过度压制,不利于细长

表 1 不同密度参数的 C3k2-WC 的实验结果

Tab. 1 Experimental results of C3k2-WC with different density parameters

Model	P/%	R/%	mAP@0.5/%	mAP@0.5:95/%	GFLOPs/G	Params /M
C3k2-WC(0.70)	69.9	65.1	71.1	42.7	5.8	2.58
C3k2-WC(0.75)	73.7	58.7	68.2	45.4	5.8	2.58
C3k2-WC(0.80)	65.1	63.7	69.1	44.9	5.8	2.58
C3k2-WC(0.85)	69.1	58.6	66.5	42.3	5.8	2.58
C3k2-WC(0.90)	76.1	63.1	71.4	46.1	5.8	2.58
C3k2-WC(0.95)	72.4	55.9	68.4	43.9	5.8	2.58

注:粗体数字代表最优值

裂缝等缺陷的完整提取。综合精度与稳定性, 本文后续实验均采用 $\text{den}=0.90$ 作为 C3k2-WC 模块的默认配置。

4.4 不同改进 C3k2 模块对比实验

为对比不同卷积设计在 C3k2 模块中的效果, 本文在 YOLOv11n 基础模型中对比了五种改进方案实验结果如表 2 所示。结果表明, C3k2-WC 以 5.8G 的 GFLOPs 和 2.58M 的参

数量取得了 71.4% 的 $\text{mAP}@0.5$, 在精度和效率之间达到最佳平衡; C3k2-PConv 和 C3k2-PFDCConv 虽然参数量最低 (2.40M), 但 $\text{mAP}@0.5$ 分别降至 68.5% 和 66.2% 效果不佳; C3k2-RFACConv 的所需算力较大, 精度提升有限。综合来看, wConv2d 的引入在减小计算负担的同时有效增强了 C3k2 模块的特征提取能力。

表 2 不同改进的 C3k2 的实验结果

Tab. 2 Experimental results of different improved C3k2 modules

Model	P/%	R/%	mAP@0.5/%	mAP@0.5:95/%	GFLOPs/G	Params /M
C3k2(Baseline)	76.7	62.0	70.8	45.2	6.3	2.58
C3k2-RFACon ^[20]	66.9	70.6	70.5	45.2	6.4	2.61
C3k2-PFDCConv ^[14]	71.4	55.6	66.2	41.0	5.8	2.40
C3k2-PConv ^[21]	58.5	65.3	68.5	44.1	5.9	2.40
C3k2-ODConv ^[22]	73.1	60.2	69.0	41.9	5.8	2.61
C3k2-WC(ours)	76.1	63.1	71.4	46.1	5.8	2.58

注:粗体数字代表最优值

4.5 不同损失函数对比实验

为分析损失函数改进后对模型性能影响, 在基础模型框架下引入了五种边界框回归损失函数进行对比, 实验结果如表 3 所示。由表 3 可见, Clou 作为基线方法, 取得了 70.8% 的 $\text{mAP}@0.5$, 但其召回率仅为 62.0%。WIoU-V3 通过引入动态非单调聚焦机制, 将召回率提升至 64.7%, 然而精确率却下降至 69.7%。ShapeIoU 与 FocalEIoU 的 $\text{mAP}@0.5$ 分别为 68.9% 与 66.9%, 均未超越基线性能。EIoU 表现最为不佳, $\text{mAP}@0.5$ 仅

为 66.5%。相比之下, 由于 Wise-ShapeIoU 在形状感知的基础上进一步引入了动态加权机制, 使得 $\text{mAP}@0.5$ 及 $\text{mAP}@0.5:0.95$ 分别达到 73.8% 与 47.6%, 相较于 Clou 基线分别提升了 3.0% 与 2.4%。该结果验证了引入形状感知与动态加权机制的 Wise-ShapeIoU 损失函数在道路缺陷检测任务中的有效性。 $\text{mAP}@0.5:0.95$ 涵盖 IoU 阈值 0.5 至 0.95, 高阈值段对边界框回归误差尤为敏感, Wise-ShapeIoU 在该指标上的提升间接反映了其边界框定位精度的改善。

表 3 不同损失函数的实验结果

Tab. 3 Experimental results of different loss functions

Model	P/%	R/%	mAP@0.5/%	mAP@0.5:95/%	GFLOPs/G	Params /M
Clou(Baseline)	76.7	62.0	70.8	45.2	6.3	2.58
WIoU-V3 ^[17]	69.7	64.7	70.5	43.4	6.3	2.58
ShapeloU ^[18]	71.5	60.5	68.9	45.0	6.3	2.58
FocalEIoU ^[23]	72.3	56.6	66.9	43.8	6.3	2.58
EIoU ^[24]	63.4	63.8	66.5	42.2	6.3	2.58
Wise-ShapeloU(ours)	72.3	66.5	73.8	47.6	6.3	2.58

注:粗体数字代表最优值

4.6 消融实验

本文在 China_Drone 道路缺陷数据集上,以 YOLOv11n 为基线模型,在相同的实验环境下进

行消融实验,以评估本文中改进模块对基线模型的影响,实验结果如表 4 所示,其中“√”表示引入相应模块。

表 4 消融实验

Tab. 4 Ablationexperiment

C3k2-WC	ADown	DAttention	Wise-ShapeIoU	P/%	R/%	mAP@ 0.5/%	mAP@ 0.5:95/%	GFLOPs/G	Params/M
—	—	—	—	76.7	62.0	70.8	45.2	6.3	2.58
√	—	—	—	76.1	63.1	71.4	46.1	5.8	2.58
—	√	—	—	74.3	64.1	72.5	47.5	5.4	2.14
—	—	√	—	74.2	61.1	70.9	46.2	6.3	2.60
—	—	—	√	72.3	66.5	73.8	47.6	6.3	2.58
√	√	—	—	68.8	68.3	72.4	46.3	4.8	2.10
√	√	√	—	72.9	70.5	73.4	46.7	4.8	2.12
√	√	√	√	80.6	64.4	74.6	47.9	4.8	2.12

注:粗体数字代表最优值

消融实验结果表明。基线 YOLOv11n 的 mAP@0.5 为 70.8%, 参数量 2.58M, GFLOPs 6.3G。引入 C3k2-WC 后, mAP@0.5 提升至 71.4%, GFLOPs 降至 5.8G, 可知加权卷积的权重调制使得模型能够聚焦缺陷特征。引入 ADown 替换下采样层卷积模块后, 模型的 GFLOPs 降为 5.4G, 显著降低了模型的计算量。引入 DAttention 注意力机制后, 参数量由 2.58M 微增至 2.60M, 但 mAP@0.5:0.95 由 45.2% 提升至 46.2%。这表明 DAttention 以极小的参数量代价有效增强了模型对不规则道路缺陷的多尺度特征提取能力。采用 Wise-ShapeIoU 损失函数后, mAP@0.5 达到 73.8%, 这也说明 Wise-ShapeIoU 损失函数在处理缺陷样本质量不平衡和多尺度缺陷问题具有一定优势。在模块组合时其中 C3k2-WC+ADown+DAttention 的 mAP@0.5 为 73.4%, 相较于基线模型, 改进模型 WADW-YOLO 在 China_Drone 数据集上的性能均有显著提升, 其中 mAP@0.5 提升 3.8%, 不仅增强了模型对目标的识别能力, 还大幅提升了整体检测性能。本文的改进算法通过高效的网络架构设计, 将参数量减少了 17.8%, 计算量减少 23.8%, 在精度与效率的权衡上具有明显优势, 适合部署于无人机等计算资源受限的边缘设备。

4.7 不同主流算法的对比实验

为全面系统地评估 WADW-YOLO 模型的性能, 本文选取了其他经典模型, 包括相关领域具有代表性的目标检测算法以及 YOLO 系列网络进行对比分析, 对比算法包括单阶段 YOLO 系列的 YOLOv10n^[25]、YOLOv11n^[12]、YOLOv12n^[26]、YOLOv13n^[27] 与 Hyper-YOLOn^[28], 以及基于 Transformer 架构的单阶段检测器 RT-DETR-R50^[29], 同时涵盖两阶段检测器 Faster-RCNN^[4]。所有对比模型均在相同环境下基于官方开源代码复现, 实验结果如表 5 所示。

由表 5 可见, WADW-YOLO 模型在保证较优检测精度的同时显著降低了参数量与计算量。具体而言, 其 mAP@0.5 达到 74.6%, Precision 为 80.6%, 相比表现次优的方法 Hyper-YOLO, mAP@0.5 提升了 3.3 个百分点。与基础模型 YOLOv11n 相比, WADW-YOLO 的 mAP@0.5 上提高了 3.8%, 同时参数量从 2.58M 压缩至 2.12M, 计算量由 6.3 GFLOPs 下降至 4.8 GFLOPs, 在保持检测性能的基础上进一步推进了模型的轻量化设计。相比之下, 传统两阶段检测器 Faster-RCNN 由于采用区域提议网络 (RPN) 机制, 引入了明显的计算冗余, 其计算量高达 134.0 GFLOPs, 而 mAP@0.5 仅为 51.5%, 故难以适用于对实时性要求较高的道路损伤检

测任务。基于 Transformer 架构的 RT-DETR-R50 虽借助自注意力机制实现了全局上下文建模,检测精度有一定保障,但其参数量高达 41.94M,内存与计算资源需求较大,严重限制了在计算能力有限的边缘设备上的实际部署可行性。

综上所述,WADW-YOLO 模型通过引入可变形注意力机制,有效提升了模型对道路损伤中不规则区域的表征能力,模型能够自适应地关注

损伤的关键几何结构,从而更精准地捕捉复杂形态。同时,结合特征融合等优化策略,模型实现了从不同网络层级提取和整合特征信息,增强了多尺度损伤的检测效果,避免了信息丢失或冗余。这些改进使得 WADW-YOLO 在检测精度与模型轻量化之间取得了较为理想的平衡,不仅显著降低了计算复杂度与存储需求,还保持了较高的检测准确率。

表 5 不同模型在 China_Drone 数据集下检测结果

Tab. 5 Detection results of different models on the China_Drone dataset

Model	P/%	R/%	mAP@0.5/%	mAP@0.5:95/%	GFLOPs/G	Params /M
Faster-RCNN	52.2	55.6	51.5	36.1	134.0	41.32
Hyper-YOLOn	71.8	68.2	71.3	46.7	9.5	3.62
RT-DETR-R50	77.3	63.8	70.1	39.9	125.6	41.94
YOLOv10n	67.1	65.6	67.1	43.4	6.5	2.26
YOLOv11n	76.7	62.0	70.8	45.2	6.3	2.58
YOLOv12n	77.0	54.8	64.4	39.4	6.3	2.54
YOLOv13n	76.6	58.1	69.1	45.6	6.1	2.45
WADW-YOLO(our)	80.6	64.4	74.6	47.9	4.8	2.12

注:粗体数字代表最优值

4.8 道路缺陷检测可视化及分析

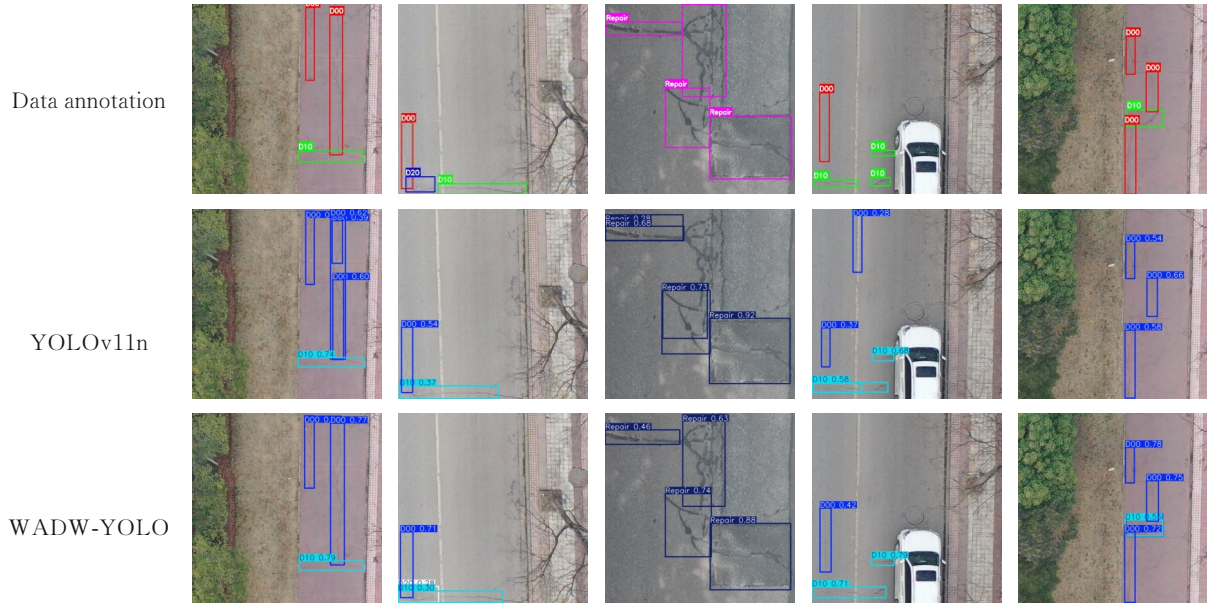
为验证 WADW-YOLO 算法在城市道路缺陷巡检场景下的泛化性能与检测精度,选取了 China_Drone 数据集 5 张强背景干扰的样本图像进行可视化分析实验,将 YOLOv11n 模型与 WADW-YOLO 模型检测效果进行对比,如图 9 所示。其中 Data annotation 为用来评估模型检测准确性的数据的真实标签。由图 9(a)组对比结果可知,针对密集分布的纵向裂缝检测时,YOLOv11n 模型存在检测框合并、轮廓贴合度差的问题。从(b)组,(c)组和(e)组对比结果可知 YOLOv11n 模型对密集裂缝和修补的边界拟合偏差较大甚至出现部分漏检情况。(d)组中 YOLOv11n 模型在识别过程中将马路分割线识别为纵向裂缝,即存在误检,故 YOLOv11n 模型存在对缺陷与背景的特征区分及弱特征缺陷识别能力不足的问题。较之 YOLOv11n 模型,WADW-YOLO 算法能够有效识别与背景高度融合的目标,且基本无漏检与误检情况;在边界回归精度上,检测框与缺陷真实轮廓的贴合度接近数据的

真实标签,密集目标的区分度与边界拟合精度大幅优化;上述可视化结果证明了本文所提改进策略的有效性。

此外,为了直观展示改进模型对缺陷区域的关注程度,选取了 China_Drone 数据集中的 5 组常见缺陷形态背景,并引入特征热力图对模型关注区域进行可视化分析,结果如图 10 所示。这些场景的选择旨在全面评估模型在多样化环境下的注意力分布。从热力图分布来看,在图 10(a)组中展示了纵向裂缝场景下,改进模型较基础模型激活更集中。(b)组展示了大面积网状裂缝场景下,改进模型响应覆盖更完整,内部空洞现象得到明显改善。(c)组展示了细小纵向裂缝检测时,改进模型实现了从头到尾的连续激活,基础模型则存在轻微断裂。(d)组展示了不规则坑洞场景下,改进模型响应紧贴实际轮廓。边缘轮廓特征提取明显。(e)组中裂缝与阴影共存时,改进模型能够有效抑制阴影干扰,裂缝响应更为清晰。相比之下,基础模型出现将人影误检为修补。从可视化实验结果可看出

WADW-YOLO模型的特征响应更加明显且集中,边缘轮廓更加清晰,凸显了模型对细微缺陷的捕捉能力。这表明WADW-YOLO模型通过

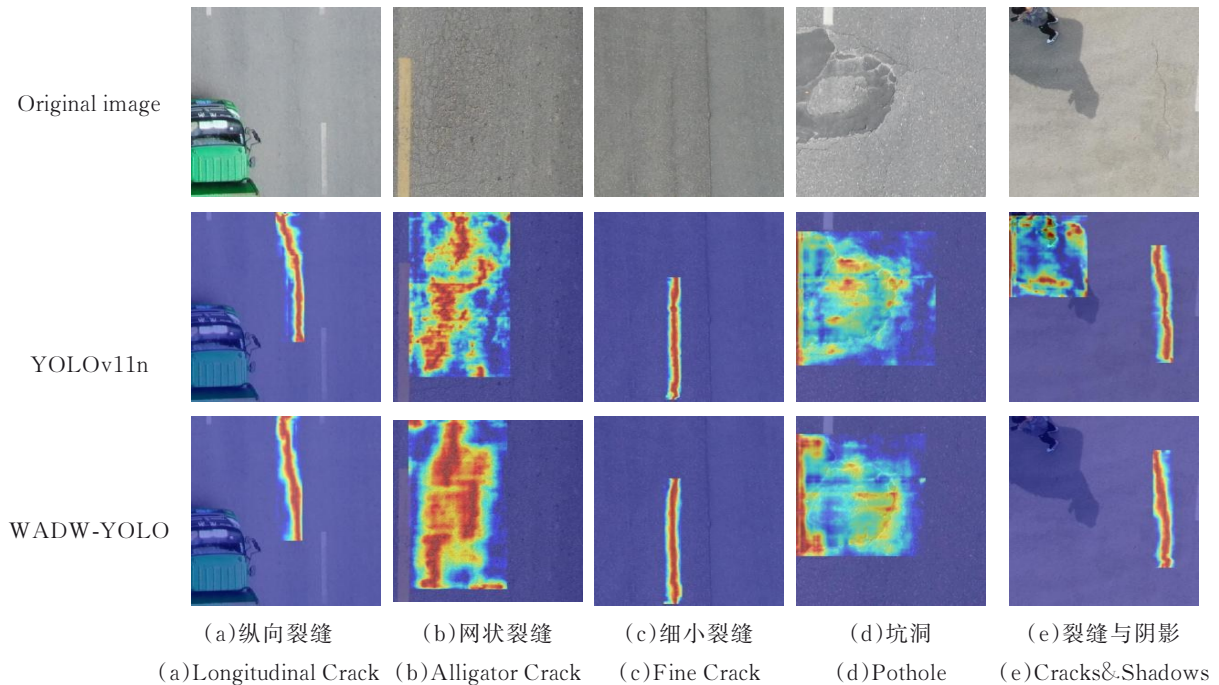
增强特征选择与融合机制,有效抑制了背景干扰,使网络更聚焦于缺陷检测目标,从而提升了识别精度与鲁棒性。



(a)多种类型裂缝 (b)多种类型裂缝 (c)修补 (d)多种类型裂缝 (e)多类型裂缝
(a)Multiple Cracks (b)Multiple Cracks (c)Repair (d)Multiple Cracks (e)Multiple Cracks

图9 在China_Drone数据集上的检测效果对比图

Fig.9 Detection performance comparison on the China Drone dataset



(a)纵向裂缝 (b)网状裂缝 (c)细小裂缝 (d)坑洞 (e)裂缝与阴影
(a)Longitudinal Crack (b)Alligator Crack (c)Fine Crack (d)Pothole (e)Cracks&Shadows

图10 YOLOv11n模型与WADW-YOLO模型的热力图可视化结果

Fig.10 Heatmap visualization results of YOLOv11n and WADW-YOLO model

4.9 泛化实验

为了检验 WADW-YOLO 模型的泛化能力,本研究选取 RDD2020 数据集中车载拍摄的 China_MotorBike 数据集进行泛化实验。实验结果如表 6 所示。WADW-YOLO 在 Precision、Recall 和 mAP@0.5 及 mAP@0.5:95 指标上分别达到 90.8%、88.5%、92.8% 和 68.5%, 优于 Faster-RCNN^[4]、RT-DETR-R50^[29] 等两阶段检测器,以及 YOLOv10n^[25]、YOLOv11n^[12]、YOLOv12n^[26]、YOLOv13n^[27] 等单阶段轻量化模型。与当前较

优的 Hyper-YOLOn^[28] 相比,本文方法的计算量相对比降低了 49.5%,参数量减少了 41.4%。实验结果表明,本文提出的加权卷积模块通过调整卷积核的权重分配,有效减少了模型的参数数量和计算开销,而特征增强策略则利用多尺度特征融合技术,显著提升了模型对道路缺陷关键信息的捕获能力,使得模型在保持高精度检测性能的同时大幅降低了模型的计算复杂度和内存占用。此外,由于模块设计轻量且适配性强,该方案也同样适合车载平台的检测需求。

表 6 不同模型在 China_MotorBike 数据集下检测结果

Tab. 6 Detection results of different models on the China_MotorBike dataset

Model	P/%	R/%	mAP@0.5/%	mAP@0.5:95/%	GFLOPs/G	Params /M
Faster-RCNN	78.4	87.4	81.0	48.0	134.0	41.32
Hyper-YOLOn	90.6	87.1	92.8	69.5	9.5	3.62
RT-DETR-R50	93.0	84.5	90.0	63.0	125.6	41.94
YOLOv10n	88.2	79.6	88.3	65.4	6.5	2.26
YOLOv11n	88.4	86.2	91.0	68.7	6.3	2.58
YOLOv12n	84.5	82.4	88.2	64.0	6.3	2.54
YOLOv13n	90.1	84.5	91.0	67.6	6.1	2.45
WADW-YOLO(our)	90.8	88.5	92.8	68.5	4.8	2.12

注:粗体数字代表最优值

5 结 论

针对无人机航拍视角下道路缺陷检测任务中存在的道路缺陷检测算法识别准确率低,误检和漏检率高以及模型计算量过高的问题,提出一种基于 YOLOv11n 的改进轻量化航拍道路缺陷检测模型 WADW-YOLO,该模型在提升检测精度的同时实现模型轻量化,从而实现高效地识别和定位道路缺陷。通过引入 DAttention 可变形注意力机制,实现感受野的自适应调节,可根据道路缺陷的形态、尺度与空间分布特征,动态调整采样点的空间偏移量与覆盖范围,从而适配多尺度、不规则分布道路缺陷。通过在模型中加入 ADown 下采样模块和 C3k2-WC 模块,优化了模型特征融合机制,同时也显著降低了模型计算量。在此基础上加入 Wise-ShapeIoU 损失函数,实现动态调节梯度权重,进一步提升道路缺陷定位检测性能。

在 China_Drone 数据集上实验结果表明, WADW-YOLO 模型相较于 YOLOv11 基础模型,其参数量降低 17.8%,计算量减少 23.8%, mAP@0.5 提升 3.8%;相较于其他主流目标检测算法, WADW-YOLO 模型在检测精度和计算效率上均表现出显著优势,降低了误检率和漏检率。在可视化及分析实验中验证了 WADW-YOLO 模型在复杂道路场景下的鲁棒性。泛化实验结果显示 WADW-YOLO 模型在 China_MotorBike 数据集上的 mAP@0.5 指标较基础模型提升 1.8%, WADW-YOLO 模型也具有更强的泛化能力。本文的研究成果可为道路养护的工程应用提供有力的技术支持,后续研究将把算法模型实际嵌入到无人机移动平台,进行进一步验证。同时,我们将扩展对道路突起、交叉裂纹等更多类别的道路缺陷检测研究,进一步提升模型的通用性。

参 考 文 献:

- [1] 《中国公路学报》编辑部. 中国路基工程学术研究综述·2025[J]. 中国公路学报, 2025, 38(12): 1-153. DOI: 10.19721/j.cnki.1001-7372.2025.12.001.
Editorial Department of China Highway Journal. Review of Academic Research on China Subgrade Engineering 2025 [J]. *China Highway Journal*, 2025, 38(12): 1-153. DOI: 10.19721/j.cnki.1001-7372.2025.12.001. (in Chinese)
- [2] 周晴, 谭功全, 尹宋麟, 等. 改进YOLOv5s的道路目标检测算法[J]. 液晶与显示, 2023, 38(05): 680-690.
Zhou Q, Tan G Q, Y S L, et al. Improvement of road target detection algorithm for YOLOv5s [J]. *Liquid Crystal and Display*, 2023, 38(05): 680-690. (in Chinese)
- [3] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector [C]// *Proceedings of the European Conference on Computer Vision*. Amsterdam: Springer, 2016: 21-37.
- [4] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks [C]// *Advances in Neural Information Processing Systems*. Montreal: Curran Associates, Inc., 2015: 91-99.
- [5] 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.
- [6] 徐光宪, 唐桂芳, 马飞. 改进YOLOv8的无人机航拍道路损伤检测方法[J]. 测绘科学, 2024, 49(09): 104-114. DOI: 10.16251/j.cnki.1009-2307.2024.09.011.
Xu G X, Tang G F, Ma F. Improved YOLOv8-based UAV aerial road damage detection method [J]. *Surveying and Mapping Science*, 2024, 49(09): 104-114. DOI: 10.16251/j.cnki.1009-2307.2024.09.011. (in Chinese)
- [7] 龙伍丹, 彭博, 胡节, 等. 基于加强特征提取的道路缺陷检测算法[J]. 计算机应用, 2024, 44(7): 2264-2270.
Long W D, Peng B, Hu J, et al. Road damage detection algorithm based on enhanced feature extraction [J]. *Computer Applications*, 2024, 44(7): 2264-2270. (in Chinese)
- [8] 张颖, 王纪旭, 曹迎康, 等. 基于改进YOLOv9-c的路面混合缺陷算法[J]. 科学技术与工程, 2025, 25(18): 7793-7802.
Zhang Y, Wang J X, Cao Y K, et al. Road surface mixed pavement damage algorithm based on improved YOLOv9-c [J]. *Science and Technology in Engineering*, 2025, 25(18): 7793-7802. (in Chinese)
- [9] 李勇, 沈坚. LFDS-YOLO: 多尺度特征融合的轻量化航拍路面缺陷检测算法[J]. 计算机工程与应用, 2025, 61(21): 81-93.
Li Y, Shen J. LFDS-YOLO: A Lightweight Aerial Road Surface Damage Detection Algorithm Based on Multi-scale Feature Fusion [J]. *Computer Engineering and Applications*, 2025, 61(21): 81-93. (in Chinese)
- [10] 朱世松, 高虹, 芦碧波, 等. 基于改进YOLOv8n的城市道路缺陷检测算法[J]. 液晶与显示, 2025, 40(09): 1356-1368.
Zhu S S, Gao H, Lu B B, et al. Improved YOLOv8n-based urban road damage detection algorithm [J]. *Liquid Crystal and Display*, 2025, 40(09): 1356-1368. (in Chinese)
- [11] Li B H, Wang K P, Liu Z H, et al. DMC-YOLOv11: an improved pavement damage detection model based on YOLOv11 [J]. *Journal of Asian Concrete Federation*, 2025, 11 (2) : 145-160. DOI: 10.1080/21642583.2025.2579987.
- [12] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements [J]. arXiv: 2410.17725, 2024.
- [13] Cammarasana S, Patanè G. Optimal Weighted Convolution for Classification and Denoising [J]. arXiv preprint, arXiv: 2505.24558, 2025.
- [14] Wang C Y, Yeh I H, Mark Liao H Y. YOLOv9: learning what you want to learn using programmable gradient information [C]// *European Conference on Computer Vision*. Berlin: Springer, 2025: 1-21.
- [15] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks; proceedings of the Proceedings of the IEEE international conference on computer vision, F, 2017 [C].
- [16] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]// *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*, New York, 2020: 12993-13000.

- [17] Tong Z J, Chen Y H, Xu Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [Z]. arXiv: 2301.10051, 2023.
- [18] [18]Zhang H, Zhang S J. Shape-IoU: more accurate metric considering bounding box shape and scale[Z]. arXiv: 2312.17663, 2023.
- [19] Arya D. , Maeda H. , Ghosh S. K. , Toshniwal D. , & Sekimoto Y. (2024). RDD2022: A multi-national image dataset for automatic road damage detection. *Geoscience Data Journal*, *11*(4), 846-862.
- [20] 张晶,苑先磊,李磊,等. 基于YOLOv11n的风电叶片表面损伤检测[J/OL]. 太阳能学报,1-9[2026-03-25]. <https://doi.org/10.19912/j.0254-0096.tynxb.2025-0829>.
Zhang J, Yuan X L, Li L, et al. Surface damage detection of wind turbine blades based on YOLOv11n [J/OL]. *Journal of Solar Energy*, 1-9 [2026-03-25]. <https://doi.org/10.19912/j.0254-0096.tynxb.2025-0829>. (in Chinese)
- [21] 郭文豪,郝斌,张飞. 改进YOLOv11的轻量化水下生物目标检测算法[J/OL]. 电子测量与仪器学报,1-17[2026-03-25]
Guo W H, Hao B, Zhang F. Lightweight underwater biological target detection algorithm improved from YOLOv11 [J/OL]. *Journal of Electronic Measurement and Instrumentation*, 1-17[2026-03-25](in Chinese)
- [22] Chao Li, ZhouAojun, YaoAnbang. Omni-Dimensional Dynamic Convolution [J/OL]. <https://arxiv.org/abs/2209.07947>.
- [23] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. *Neuro computing*, 2022, 506:146-157.
- [24] TIAN Y, SU D, LAURIA S, et al. Recent advances on loss functions in deep learning for computer vision[J]. *Neuro computing*, 2022, 497:129-158.
- [25] WANG A, CHEN H, LIU L, et al. Yolov10: Real-time end-to-end object detection[J]. *Advances in Neural Information Processing Systems*, 2024, 37: 107984-108011.
- [26] TIAN Y, YANG Q, ZHANG S, et al. YOLOv12: Attention-centric real-time object detectors[EB/OL]. (2025-02-18). <https://arxiv.org/abs/2502.12524>.
- [27] Ultralytics. YOLOv13: Object detection model[EB/OL]. (2025). <https://github.com/ultralytics/ultralytics>.
- [28] WANG Y, LIU C, ZHANG Z, et al. Hyper-YOLO: When visual object detection meets hypergraph computation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025, 47(4): 2388-2401.
- [29] LV W, XU S, ZHAO Y, et al. DETRs beat YOLOs on real-time object detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2024: 16965-16974.

作者简介:



孙 瑞,男,硕士研究生,2025年于桂林信息科技学院获得学士学位,主要从事目标检测方面的研究。E-mail: 13463948462@163.com



燕 芳,女,博士,教授,2016年于北京科技大学获得博士学位,主要从事太赫兹时域光谱技术与应用及目标检测等方面的研究。E-mail: 0472yanfang@163.com