

doi: 10.16104/j.issn.1673-1891.2026.01.007

基于 YOLO11-AKAD 的道路缺陷检测方法

阮进军

(安徽商贸职业技术学院信息与人工智能学院, 安徽 芜湖 241002)

摘要:针对现有道路缺陷检测方法在复杂道路背景下存在准确度低、易漏检误检、难以满足实时性要求等问题,提出一种轻量化 YOLO11-AKAD 道路缺陷检测算法。首先,在 YOLOv11n 的主干网络中设计了基于自适应核特征提取 C3k2_AKConv 模块,增强模型的局部特征提取;其次,使用 ADown 模块优化了 downsampling 过程,优化多尺度上下文信息传递,减少参数量;最后,在输入测试时采用自适应图片缩放的方式解决原始缩放方法可能产生大量冗余信息的问题。实验结果表明:YOLO11-AKAD 在 China_RDD 数据集上的 mAP@0.5 指标达到了 81.6%,参数量和浮点运算量分别降低至 2.09 M 和 5.1 G FLOPs;相较于基准的 YOLOv11n, mAP@0.5 指标提高了 1.8%,模型参数规模减少了 19%。该方法能在一定程度上有效提高道路缺陷检测性能。

关键词:YOLOv11n;道路缺陷检测;自适应核卷积;下采样算子

中图分类号:TP183;TP391.41;U418 **文献标志码:**A **文章编号:**1673-1891(2026)01-0055-12

Road Defect Detection Method Based on YOLO11-AKAD

RUAN Jinjun

(School of Information and Artificial Intelligence, Anhui Business College, Wuhu 241000, Anhui, China)

Abstract:To address the problems of low accuracy, high rates of missed and false detections, and poor real-time performance of existing road defect detection methods in complex road backgrounds, this paper proposes a lightweight road defect detection algorithm named YOLO11-AKAD based on YOLOv11n. Firstly, a C3k2_AKConv module based on adaptive kernel feature extraction is designed in the backbone network of YOLOv11n to enhance the model's local feature extraction capability. Secondly, the ADown module is introduced to optimize the downsampling process, improve the transmission of multi-scale contextual information, and reduce the number of model parameters. Finally, an adaptive image scaling method is adopted in the input preprocessing stage to solve the problem of excessive redundant information caused by the original scaling method. Experimental results on the China_RDD dataset show that the mAP@0.5 of YOLO11-AKAD reaches 81.6%, with the number of parameters and floating-point operations reduced to 2.09 M and 5.1 G FLOPs, respectively. Compared with the baseline model YOLOv11n, the mAP@0.5 is increased by 1.8 percentage points, and the model parameter scale is reduced by 19%. The results demonstrate that the proposed method can effectively improve the performance of road defect detection in complex scenarios.

Keywords:YOLOv11n; road defect detection; adaptive kernel convolution; downsampling operator

收稿日期:2025-11-13

基金项目:安徽省高校自然科学研究重点项目(2023AH052296、2024AH050531)。

作者简介:阮进军(1976—),男,安徽桐城人,教授,博士,主要研究方向为计算机技术、人工智能。E-mail:ruanjinjun@abc.edu.cn。

0 引言

中国作为道路基础设施发展的领先国家之一,近年来公路网络规模得到了大幅度发展。然而,道路的长期使用产生了各种缺陷,包括裂缝、龟裂和坑洼^[1]。这些损坏不仅扰乱交通流,还带来严重的安全隐患。传统的人工检测方法主要依靠工程师和道路养护人员进行人工勘察,这些方法虽然能检测到可见的缺陷,但耗时、费力且容易出现人为错误。如果使用配备摄像头和激光扫描仪专用车辆的检测技术,则需要大量的资金投入和人力资源,这限制了大规模道路监测的可行性^[2]。

计算机视觉技术的发展为道路缺陷检测提供了创新解决方案,成为该领域技术进步的主要驱动力^[3]。早期基于机器学习的计算机视觉道路缺陷检测方法依赖于人工设计的特征或特征提取器,在条件高度可变的道路缺陷检测任务中,手动调整和优化特征提取器和分类器可能很烦琐,需要耗费大量计算资源,最终导致效率低下。

随着人工智能和深度学习的快速发展,自动化道路缺陷检测系统已成为道路维护和提升的有效解决方案。深度学习模型,特别是卷积神经网络,在监控摄像头、无人机和移动设备收集的图像和视频数据中识别和分类路面缺陷方面展现出了卓越的准确性^[4]。Li 等^[5]用 Fast R-CNN 算法精准地辨识并定位坑洞、裂缝等 6 种常见的道路缺陷。虽然模型考虑了在复杂环境背景下干扰的问题,但是在光照和天气等因素的影响下,模型的鲁棒性和准确性仍旧不高。单次多框检测器(single shot multibox detector, SSD)利用多个不同尺度的特征图作为单个神经网络的输入,通过处理不同层级的特征图,在检测速度和多尺度精度之间取得了平衡^[6],近年来广泛应用于道路缺陷检测。韦正璐等^[7]提出利用改进 Inception 模块,替代 SSD 网络部分结构,在不丢失分辨率的情况下扩大感受野,提高浅层特征图

对目标物体的特征提取能力。黄凯枫等^[8]在 SSD 网络结构的基础上将其基础网络替换为 Dense-net 网络,同时引入注意力机制模块提升了道路病害识别的准确性。但在复杂的路面场景中,模型易受背景干扰,且对路面微小裂缝、远处小物体容易产生漏检。

YOLO 系列算法的核心思想是将目标检测视为一个回归问题,能够直接预测输入图像中目标的类别和边界框的位置。由于它实现了目标定位和识别的并行处理,使其在自动驾驶、机器人等应用中检测小型或密集排列的物体时特别有效。YOLO 历代版本不断对原始的 YOLO 架构进行优化和增强,众多科研人员在此基础上进行了大量的研究工作。陈飞宇等^[9]提出基于 YOLOv5s 改进模型,用于解决车载影像数据集中道路病害尺寸差异大、小尺度病害多导致的检测精度低问题,其通过通道混合滑动 Transformer 模块、广义特征金字塔、动态检测头等方法提升了 YOLOv5s 的 mAP@0.5 和检测速度。胥铁峰等^[10]设计了面向复杂自然场景的轻量级道路病害检测模型,其通过 C2f_DCNv3 模块、GS-Decoupled head 检测模块、E-Slide Loss 权重函数及通道剪枝优化性能,相较于 YOLOv8n 模型 mAP 提高 2.4%。王杰等^[11]通过将 YOLOv8n 主干网络的 C2f_EMA 模块和颈部网络的 CARAFE 上采样模块优化特征提取与细节保留, mAP@0.5 较 YOLOv8n 提高 3.4%。罗臻等^[12]基于 YOLO11n 模型,通过引入 CGBlock 上下文引导模块、DRCGN 跨尺度特征融合模块及融入尺度因子学习机制的 MPDIoU 损失函数优化性能,在 RDD2022 数据集上 mAP@0.5 较原模型分别提升了 3%。

由此可见,基于计算机视觉道路缺陷检测方法虽然在检测准确率、检测速度等方面取得了一些进步,但是在面对复杂道路检测环境下仍然存在检测准确率偏低,小缺陷目标容易误检漏检的情况。本文在 YOLO11n 框架的基础上设计了自适应核的

特征提取 C3k2_AKConv 模块,加强了局部特征提取;嵌入 Adown 下采样模块,优化多尺度上下文信息传递;最后检测输入时采用自适应图片缩放的方法,将改进后的 YOLOv11-AKAD 网络模型应用于道路缺陷检测以提升目标检测性能。

1 YOLOv11 算法改进

YOLOv11 是由众多研究学者基于 Ultralytics 框架于 2024 年提出的当前最先进 (state of the art, SOTA) 一阶段目标检测算法^[13]。算法延续了 YOLO 系列实时高效的特点,并在性能、精度和应用范围上实现突破^[14]。YOLOv11 基于传统 CSP Bottleneck,采用并行卷积分支设计了 C3k2 模块,浅层特征和深层特征都得到了关注,在 C2f 模块上设计了金字塔切片注意力机制 (pyramid split attention, PSA),增强了模型的特征选择能力;模型的 Head 部分则采用 3 任务解耦架构,将分类、回归和分割分支与共享中间特征分离,实现跨检测、分割任务的灵活部署。根据模型规模的不同,YOLOv11 可分为 5 种不同版本的模型,其中 YOLOv11n 参数量最小,在减少冗余计算的同时提升了检测速度,可以更好地满足轻量化设计需要。

在道路缺陷检测应用场景中,数据采集环境复杂,容易受到日照、天气等多变因素干扰;此外,采集设备的多样化也导致目标不可避免地出现多尺度分布特征。对 YOLOv11n 的具体研究表明,在使用相对较小的参数前提下,模型在面对环境比较复杂的道路缺陷检测应用场景中效果不理想,尤其是在小目标检测任务中常常出现漏检、误检的问题。综合考虑模型轻量化和性能要求,本文在 YOLOv11n 模型的基础上进行了改进。

1.1 融入 AKConv 模块改进网络

标准卷积采样位置在网格上是固定不变的,比较常见的是正方形卷积核。在道路缺陷检测应用场景中道路环境复杂多变,采集到的检测目标数据

差异化巨大,标准卷积核难以自适应地捕捉目标的关键特征。当面对不规则形状的物体时,固定形状的卷积核可能无法完全覆盖目标区域,导致局部信息的丢失;同时,传统正方形卷积核的参数数量会随着尺寸的增大呈平方增长。为了克服这一缺陷,研究者们提出了多种改进方法,如可变形卷积 (deformable convolution)^[15]和动态卷积^[16]等。可变形卷积通过学习偏移量来动态调整采样位置,在一定程度上提高了卷积核对目标形状的适应性。然而,可变形卷积仍受限于规则的卷积核尺寸,如基于 3×3 网格的偏移,且参数数量依然呈平方增长,未能从根本上解决问题。动态卷积则是根据输入特征动态生成卷积核权重,但通常在固定的采样网格上进行采样,同样存在局限性。

为了克服这些问题,Zhang 等^[17]提出了任意核卷积 (arbitrary kernel convolution, AKConv)。该模块通过动态偏移调整采样位置,并结合初始坐标生成算法,可以根据输入特征的变化自适应地调整卷积核的形状和参数数量,使卷积核更好地贴合目标形状,提高特征提取的准确性。此外 AKConv 的参数数量仅随尺寸线性增长,大幅度降低了模型的参数数量,模型轻量化和资源利用效率得到了显著提升。AKConv 结构如图 1 所示。

AKConv 中,假设在定义围绕中心锚点 P_0 的不规则采样点 P_n 后, P_0 对应的原始坐标集合为 Q ,可有效解决不规则卷积无法匹配相应大小的标准卷积问题。集合 Q 如式(1)所示。

$$Q = \sum [\omega \times (P_0 + P_n)] \quad (1)$$

式中: ω 为卷积参数。

AKConv 通过二维卷积操作,生成一个大小为 $2N \times H \times W$ 的偏移量张量 θ 。其中 N 为卷积核中采样点的数量, $2N$ 为每个采样点的 X 和 Y 方向的偏移量, H 和 W 分别为输入图像的高和宽。接着通过添加偏移量来动态调整采样坐标,相加得到修改后的采样坐标集合 P ,如式(2)所示。

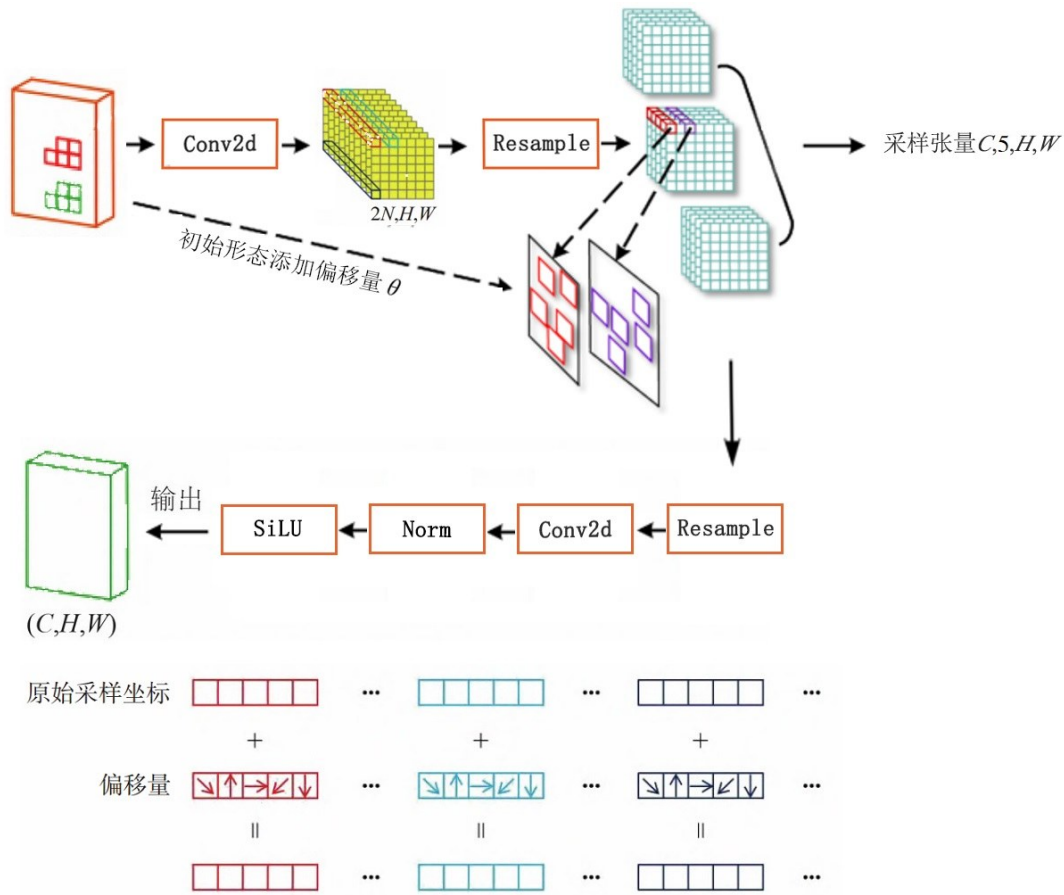


图1 AKConv 结构

$$P = Q + \theta \tag{2}$$

再使用双线性插值式(3)计算新的特征值,使得卷积能够作用在连续空间上。

$$y(p) = \sum_{q \in \tilde{q}} [\max(0, 1 - |q_x - p_x|) \times \max(0, 1 - |q_y - p_y|) \times y(q)], q \in \tilde{q} \tag{3}$$

式中: p 为修改后的采样坐标集合中的任意一点, q 为相应相邻点的坐标集合中任意一点, p_x, p_y 和 q_x, q_y 分别为 p 点和 q 点的 x, y 坐标值。

使用调整后的采样点对输入特征图进行卷积计算,提取更具适应性的特征,从而得到增强后的特征图。

最后,引入归一化和激活函数来进一步增强特征并确保数值稳定性

1.2 在 C3k2 模块中引入 AKConv

考虑到 AKConv 在特征提取中的优势,本文将 AKConv 引入 YOLOv11 的 C3k2 模块中,设计 C3k2-

A 模块以替代原始的标准卷积结构。在特征提取阶段,AKConv 能够动态调整卷积核大小,从而能够灵活地选择最优感受野,更高效地捕捉局部细节和全局上下文信息。图2展示了引入 AKConv 后改进的 C3k2-AKConv 模块的网络结构。

从图2可以看出,当 C3k 为 False 时,输入特征

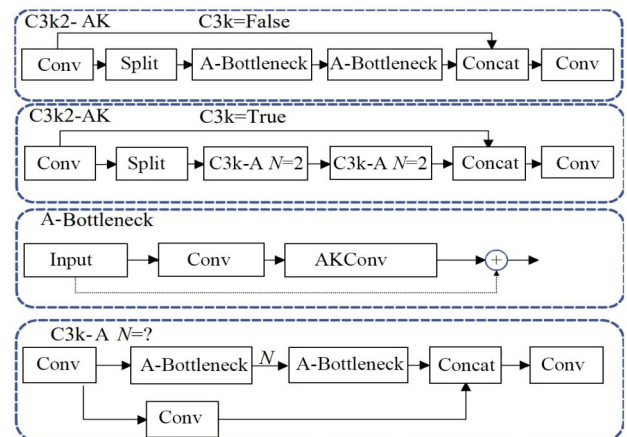


图2 C3k2-AKConv 模型结构

首先通过一个普通的卷积层生成初始特征,然后通过分割被分为2个并行分支。每个分支的核心是A-Bottleneck结构。2个分支的输出沿通道维度进行拼接,随后通过最后一个卷积层融合通道信息,完成多尺度特征聚合。在这种模式下,该模块通过A-Bottleneck内部的动态核调整和残差连接,实现局部细节与全局上下文的初步融合。当C3k为True时,输入特征同样经过初始卷积层和分割操作,但每个分支升级为C3k-A(N=2)结构。单个A-Bottleneck的输出如式(4)所示。

$$Y_{AB} = F_{Conv}(X) + F_{AKConv}(X) \quad (4)$$

每个C3k-A,包括残差连接的计算如式(5)所示。

$$F_{C3k-A}^{(i)}(X) = \sum_{j=1}^{N-2} F_{AB}^{(j)}(X) + X \quad (5)$$

式中: $N=2$ 表示每个C3k-A包含2个A-Bottleneck, X 是该分支的输入特征。

AKConv通过动态调整感受野,自适应地融合不同分支中的多尺度特征。A-Bottleneck和C3k-A均引入残差连接,以缓解梯度消失并增强特征重用。这种结构在C3k2模块中实现了高效的多尺度特征提取,同时通过AKConv的动态核机制平衡了模型复杂度与性能。

1.3 嵌入ADown下采样模块

下采样在基于深度学习的目标识别领域,能够有效减少特征图空间维度,降低模型计算复杂度与显存占用,同时扩大特征感受野,为后续特征融合与目标识别提供支撑。传统降采样方法主要采用步幅卷积、池化等方法,虽然可以实现维度压缩,但是由于基本都是采用固定的采样规则,无法根据输入特征的内容差异动态调整策略,导致采样过程中不可避免地出现特征丢失,尤其会弱化道路缺陷中数量众多的小目标的关键信息,影响模型对复杂道路场景的适应能力。Fang等^[18]提出了高效可学习下采样结构模块ADown(adaptive downsampling module)。ADown模块的结构如图3所示。

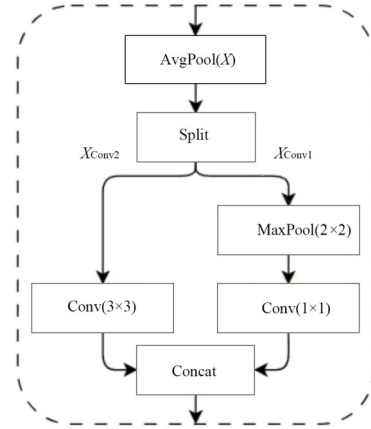


图3 ADown模块结构

ADown模块通过动态调整采样策略减少特征信息丢失,突破传统固定模式下采样方法的局限,在实现维度压缩的同时最大化保留输入特征的关键语义与边缘细节。该模块采用双分支路径架构与卷积-池化联合建模机制,在保留了特征表达能力的同时兼顾了模型轻量化,其输入特征图如式(6)所示。

$$X \in \mathbf{R}^{C \times H \times W} \quad (6)$$

式中: C 为特征通道数。

ADown模块的具体工作流程可分为全局压缩、双分支处理及特征融合3个关键阶段。

首先,对输入特征图 X 执行二维平均池化操作,通过全局压缩获得紧凑的中间特征表示 X_{avg} ,如式(7)所示。

$$X_{avg} = \text{AvgPool}(X) \quad (7)$$

式中: $\text{AvgPool}(X)$ 表示核大小与步长适配降采样需求的平均池化操作生成维度压缩后的中间特征图。

其次,将平均池化输出的中间特征图 X_{avg} 沿通道维度进行均分处理,得到2组结构一致的双分支特征子图 X_{avg1} 与 X_{avg2} ,分别对应后续分支输入 X_1 与 X_2 ,确保了2组子图能够分别聚焦不同维度的特征信息。接下来的操作将围绕上述2个结构异构的分支实现特征提取与边缘信息保留的协同优化,先将第一组特征子图 X_1 作为主分支通过带参数的卷积运算操作完成降采样与局部上下文信息建模。该

操作通过步长设置实现特征图空间维度的压缩,同时利用局部感受野聚合上下文信息,强化特征的语义表达能力。其输出如式(8)所示。

$$X_{Conv1} = \text{Conv}(X_1; k, s, p) \quad (8)$$

式中:Conv 卷积操作定义为核大小 $k \times k$ 、步长 s 、填充系数 p 的标准卷积。

在第 2 组特征子图 X_2 输入辅助分支后,首先通过 2×2 最大池化操作提取边缘轮廓与强响应区域特征,该过程保留了特征图中的关键结构信息,由此避免了传统采样过程中边缘细节消失的问题。随后对最大池化输出结果执行卷积处理,得到特征图 X_{Conv2} 实现了边缘特征的精细化建模与维度适配。其输出如式(9)所示。

$$X_{Conv2} = \text{Conv}(\text{MaxPool}_{2 \times 2}(X_2); k, s, p) \quad (9)$$

最后特征融合阶段将主分支输出 X_1 与辅助分支输出 X_2 沿通道维度进行拼接融合,生成最终的降采样特征图,其输出如式(10)所示。

$$Y = \text{Concat}(X_{Conv1}, X_{Conv2}) \quad (10)$$

ADown 模块由于整合双分支分别提取的语义特征与边缘细节,形成兼具全局上下文与局部结构信息的高维特征表示,在特征下采样过程中,它能够在保持高精度的同时,提高计算效率并降低资源消耗。

本文在 YOLOv11n 原始网络基础上改进了 YOLO11-AKAD 网络,如图 4 所示。YOLO11-AKAD 通过引入 AKConv 对 YOLOv11 的 C3k2 模块进行改进,动态调整卷积核大小,从而灵活捕捉局部细节和全局上下文信息。对模型图中 P4 头 C3k2 模块进行替换,以增强对小目标和细节特征的提取效果。在 YOLOv11n 的 Backbone 和 Neck 部分部署了 Adown 模块,以优化下采样过程,提升小目标检测性能,并增强模型对复杂背景中小尺度目标的适应性。改进后的 YOLO11-AKAD 模型在提升道路缺陷准确度的基础上大幅度减少了模型参数规模,降低了计算成本。

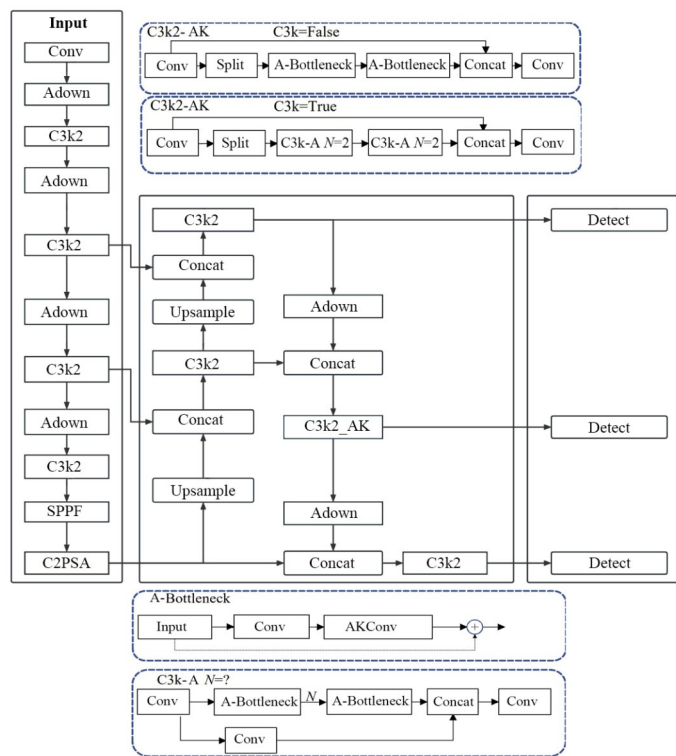


图 4 YOLO11-AKAD 网络结构

1.4 自适应图片缩放

在道路缺陷目标检测任务中,图像预处理阶段的统一尺寸缩放策略在实际应用中易出现缩放填充后黑边不对称的问题,若填充过量会引入大量冗余信息,严重拖累检测推理效率。为解决这一痛点,拟借鉴自适应图像缩放的创新思路,通过动态添加少量黑边的方式优化图像缩放流程,有效削减冗余信息。针对 YOLOv11 网络的输入特性,可以将缩放后长边尺寸设定为 YOLOv11 的标准输入尺寸 768。设原图长为 c , 宽为 k 。

首先,计算放缩比例 p , 计算公式如式(11)~(12)所示。

$$x = \begin{cases} c, & \text{if } c > k \\ k, & \text{else} \end{cases} \quad (11)$$

$$p = \frac{768}{x} \quad (12)$$

其次,根据原图尺寸与缩放比例计算缩放后图像尺寸,计算公式如式(13)所示。

$$\begin{cases} \hat{c} = c \times p \\ \hat{k} = k \times p \end{cases} \quad (13)$$

最后,计算黑边填充参数,计算公式如式(14)所示。

$$\begin{cases} w = \hat{c} - \hat{k} \\ m = w \bmod (w, 32) \\ y = \frac{m}{2} \end{cases} \quad (14)$$

式(12)~(14)中: \hat{c} 、 \hat{k} 分别为缩放后长和宽, w 为需要填充的黑边宽度, m 为模值, y 为填充到短边的2边的值。选取 32 作为模值,是因为 YOLOv11 经多次下采样后特征层的感受野与 32 强适配,可保障特征提取的完整性。该自适应缩放方法在维持 YOLOv11 轻量化推理特性的同时,通过精准控制黑边填充量,避免了冗余信息的过度引入,为后续道路缺陷目标的高效检测奠定了预处理基础。

2 实验验证

2.1 实验环境与数据准备

实验环境配置如表 1 所示,网络模型训练参数如表 2 所示。

表 1 实验环境配置

平台配置	版本
操作系统	Windows 11
编程语言	Python 3.9
深度学习框架	PyTorch 2.3.1+cu118
CPU	i9-12900KS@3.40 GHz
GPU	NVIDIA GeForce RTX 4090
Memory	DDR5 32 GB@6000 MHZ

表 2 实验训练参数

参数	配置
轮数	200 epochs
批处理大小	32
输入分辨率	640×640
数据加载线程数量	8 workers
优化器	SGD
初始学习率	0.01
动量	0.937
预热训练	warmup_epochs=3

本文基于 2022 年 IEEE 国际大数据会议上举办的全球道路损坏检测挑战赛的全球道路损坏数据集 RDD2022, 创建了中国道路缺陷实验数据集 China_RDD。选用了原数据中通过无人机和车载移动摄像头收集的中国道路数据。该数据集共有 4 378 张图像, 分属 D00(纵向裂缝)、D10(横向裂缝)、D20(网状裂缝)、D40(坑洞)4 个类别, 相应的数量分别为 4 104、2 382、934 和 321。其中, D00、D10、D20、D40 分别占总数的 93.7%、54.4%、21.3%、7.3%。D00 占绝对多数, 而数量最少, 仅约占 7%。这种严重的不平衡可能导致模型在训练过程中偏

向多数类,而忽略少数类,从而影响整体检测性能。因此,本文采用了多维度数据增强策略优化数据集。一方面,利用中国天池数据集中的坑洼数据扩充弱势样本。另一方面,采用自适应图片缩放、几何变换、色彩调整与动态模糊等数据增强手段,通过缩放、旋转、裁剪等几何变换增加数据集的多样性,降低模型过拟合的风险;利用亮度调整、色调变换等操作引入随机的颜色变化来丰富训练数据的多样性,从而增强模型对不同光照和色彩环境的适应性;同时借助动态模糊技术模拟现实场景,丰富训练数据的动态多样性,使模型能够适应不同程度的模糊干扰,有效提升对复杂场景的识别与处理能力,进一步增强模型的鲁棒性与泛化能力。经过上述处理,数据集规模由 4 378 张图像扩充至 5 119 张,其中 D40 的样本数由原先的 321 增长到 1 521 个。实验结果表明,使用数据增强后的数据集进行模型训练,其性能相较于原数据集训练的模型均有明显提升,其中 YOLOv11n 网络 D40 的检测 mAP@0.5 指标值增长了 2.06%,改进后的 YOLO11-AKAD 网络 D40 的检测 mAP@0.5 指标提升了 2.01%。

2.2 评估标准

为了全面评估该模型的性能,实验采用平均精度均值(mAP)、精确率、召回率、模型参数数量(Parameters)、模型的浮点运算量(GFLOPs)等多个指标来评估性能。

精确率计算公式如式(15)所示。

$$N_{\text{Precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \quad (15)$$

式中: $N_{\text{Precision}}$ 为精确率值, N_{TP} 为真正例的数量, N_{FP} 为假正例的数量。

召回率计算公式如式(16)所示。

$$N_{\text{Recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (16)$$

式中: N_{FN} 为假反例的数量。

N_{AP} 是召回率和精确度为横纵坐标围成的面

积,计算公式如式(17)所示。

$$N_{\text{AP}} = \int_0^1 P(R) dR \quad (17)$$

通过计算所有类别的平均精度(AP)的平均值得到 mAP,反映了模型对不同缺陷类型的整体检测能力,其计算公式如式(18)所示。

$$N_{\text{mAP}} = \frac{1}{N_{\text{class}}} \sum_{i=1}^{N_{\text{class}}} N_{\text{AP}i} \quad (18)$$

式中: N_{class} 为缺陷类别的数量, $N_{\text{AP}i}$ 为第*i*个类别的 N_{AP} 值。

2.3 实验结果与分析

使用 China_RDD 数据集训练优化的基于 YOLOv11 的道路缺陷检测算法。数据集按 7:2:1 的比例分为训练集、验证集和测试集。在网络收敛后,得到成熟的网络模型。

YOLO11-AKAD 网络损失曲线如图 5 所示。由图 5 可知,YOLO11-AKAD 网络在训练集上的训练框、训练分类、训练 DFL 损失均随迭代快速下降后趋于平稳,且平滑曲线与结果曲线趋势一致,训练过程稳定;验证集的验证框、验证分类、验证 DFL 损失也呈现持续下降并收敛的趋势,在损失下降过程中,无大幅度波动,整体表明该网络训练有效且稳定。这表明网络模型设计合理,其规模深度与数据集构建均较为完善,在边界框回归、类别分类、分布焦点任务上的学习与泛化能力良好。

改进的 YOLO11-AKAD 网络与 YOLOv11n 在多类型道路缺陷检测上的精确率-召回率 PR 曲线(图 6)对比显示,YOLO11-AKAD 的平均 mAP@0.5 达到 0.816,优于 YOLOv11n 的 0.798,且在各类缺陷检测中表现更均衡。其中 D10 检测精度达到 0.867,检测精度最高,主要由于横向裂缝特征相对清晰且纹理规律易于提取;D00 的检测精度达到 0.833,在标注纵向裂缝时,其线性特征明显,减少了背景干扰,提升了网络对该类别缺陷的识别鲁棒性;网络对 D40 的检测精度为 0.710,较 YOLOv11n 的 0.692 有明显提升,体现了该网络能够更准确地获取像素占比较

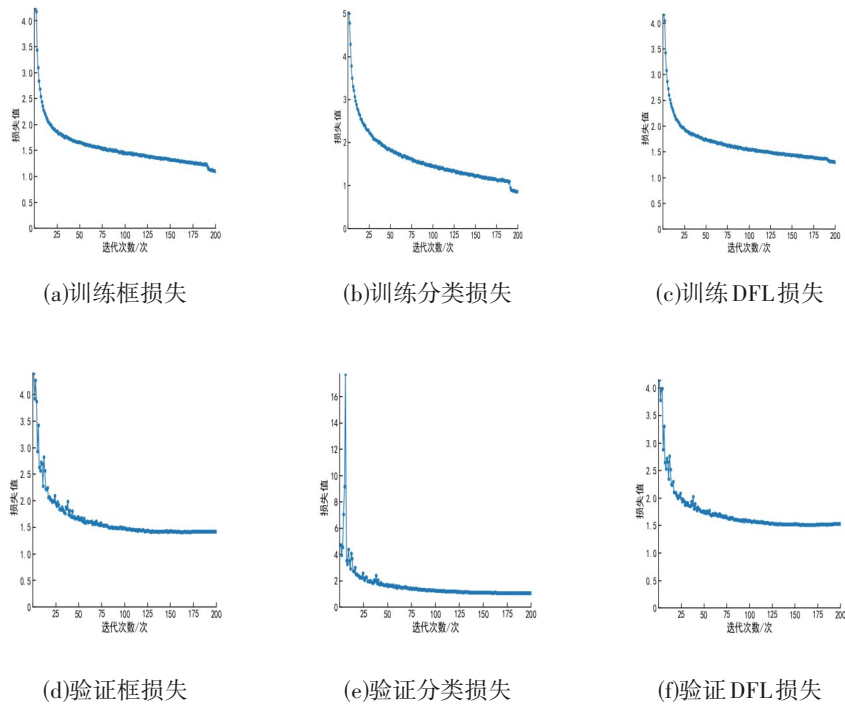


图5 YOLO11-AKAD网络损失曲线

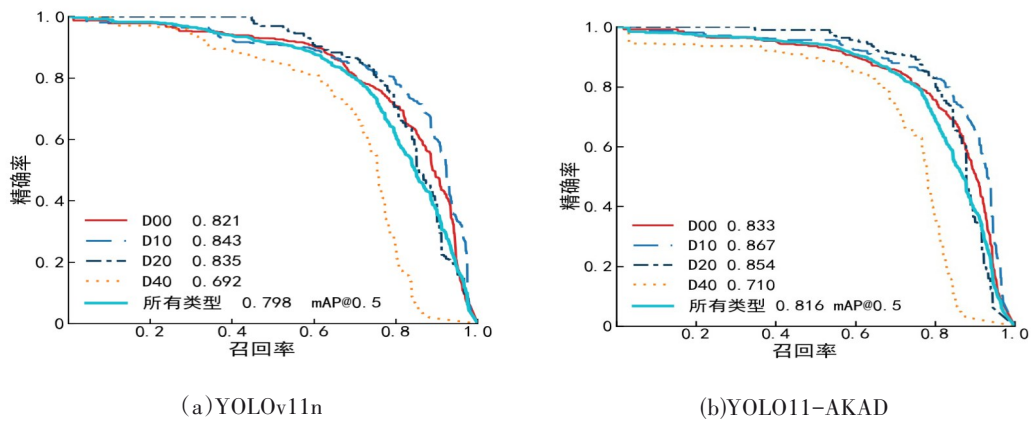


图6 网络精确率-召回率曲线

小且形态不规则的缺陷信息。从PR曲线可见，YOLO11-AKAD在各召回率区间的精确率均高于YOLOv11n，尤其是在高召回率段优势更显著，说明其在道路缺陷检测的精度和泛化能力上均有优化。

为了深入评估YOLO11-AKAD在中国道路缺陷检测数据集上的性能优势，在相同实验环境及参数设置中分别对Faster RCNN、SSD300、YOLOv5n、YOLOv8n、YOLO11n网络进行训练，训练过程中使

用相同的训练方式，网络性能对比测试结果如表3所示。

YOLO11-AKAD道路缺陷检测算法精确率为81.7%，对比FRCNN、SSD300、YOLOv5n、YOLOv8n、YOLO11n分别提升41.0%、3.5%、4.9%、2.4%、3.4%，算法检测精确率有明显提升。其mAP@0.5为81.6%，对比FRCNN、SSD300、YOLOv5n、YOLOv8n、YOLO11n分别提升5.3%、25.85%、2.9%、1.4%、

表 3 不同算法模型在 China_RDD 道路缺陷数据集中的性能比较

模型	精确率/%	召回率/%	mAP@0.5 /%	参数量 /M	浮点运算量/G
FRCNN	40.7	65.5	76.33	137.1	370.2
SSD300	85.1	22.1	55.75	26.3	85.1
YOLOv5n	76.8	72.1	78.7	2.50	7.1
YOLOv8n	79.3	75.0	80.2	3.01	8.1
YOLO11n	78.3	73.2	79.8	2.58	6.3
YOLO11-AKAD	81.7	75.6	81.6	2.09	5.1

1.8%, 检测性能更优。模型参数量为 2.09 M, 对比 FRCNN、SSD300、YOLOv5n、YOLOv8n、YOLO11n 分别减少 135.01 M、24.21 M、0.41 M、0.92 M、0.49 M。浮点运算量为 5.1 G, 对比上述模型分别减少 365.1 G、80 G、2 G、3 G、1.2 G, 参数量和计算量优势显著, 能更好适应硬件平台资源受限的应用场景。

YOLO11-AKAD 在多数场景下均能较稳定地完成道路缺陷目标检测。模型相对原始 YOLOv11n 网络, 在细裂缝与纵向缺陷特征提取方面整体更为准确, 漏检率有所降低。但在斑驳路面、标线重合、多重道路缺陷相互重叠的特殊场景中, 模型的特征表达容易受到噪声干扰, 存在漏检或误检的现象, 说明模型对复杂纹理背景的特征分离能力有待进一步增强。在小目标检测方面, YOLO11-AKAD 对细小裂缝的检出数明显多于 YOLOv11n, 但对极端微小的裂纹仍不够敏感, 模型的小目标特征保持与强

化方面需进一步优化。在光照变化方面, 模型对常规光照波动具有良好的鲁棒性, 但在逆光、强光直射和明暗交界区域, YOLO11-AKAD 模型检测框置信度出现下降, 尤其是白色标线附近的细裂纹更容易被忽略, 这表明, 模型的特征提取模块对低对比度细节的鲁棒性还需进一步提升。总体而言, 改进后的模型在一定程度上有效提升了道路缺陷检测性能, 但在复杂纹理背景、极小目标细节提取以及极端光照条件下仍需进一步优化特征提取与多尺度融合机制。

2.4 消融实验

针对实现道路缺陷检测 YOLOv11n 提出了多个模块的改进, 共设计了 4 组实验。通过逐步添加自定义模块, 观察模型的平均精度均值 mAP、精度、召回率等性能指标的变化, 验证各改进方案对检测结果的影响, 对比性能如表 4 所示。

表 4 消融实验结果对比

Baseline	AKConv	Adown	精确率/%	召回率/%	mAP@0.5/%	参数量 /M	浮点运算量/G
YOLOv11n	无	无	78.3	73.2	79.8	2.58	6.3
	有	无	78.0	75.8	80.7	2.57	6.3
	无	有	78.2	76.3	80.1	2.23	5.3
	有	有	81.7	75.6	81.6	2.09	5.1

由表 4 可知, 未添加 AKConv 和 Adown 的 YOLOv11n 精确率为 78.3%, 召回率为 73.2%, mAP@

0.5 为 79.8%, 参数量为 2.58 M, 浮点运算量为 6.3 G; 仅添加 AKConv 时, 精确率为 78.0%, 召回

率提升至 75.8%, mAP@0.5 提升至 80.7%, 参数量略微减少至 2.57 M, 浮点运算量保持 6.3 G; 仅添加 A_{down} 时, 精确率为 78.2%, 召回率提升至 76.3%, mAP@0.5 提升至 80.1%, 参数量减少至 2.23 M, 浮点运算量减少至 5.3 G; 同时添加 AK-Conv 和 A_{down} 后, 即 YOLO11-AKAD, 精确率提升至 81.7%, 召回率为 75.6%, mAP@0.5 提升至 81.6%, 参数量进一步减少至 2.09 M, 浮点运算量减少至 5.1 G。这是因为 AKConv 能够增强网络对道路缺陷特征的提取能力, 提升检测精度; A_{down} 在降低特征图分辨率的同时能有效

保留缺陷特征信息, 减少参数量和计算量。二者结合后, 在提升检测性能的同时, 显著降低了模型的参数量和计算量, 使算法能更好适应硬件平台资源受限的应用场景。

2.5 泛化实验

为了全面验证 YOLO11-AKAD 模型在不同数据域下的泛化能力、鲁棒性、迁移学习能力、结构普适性和计算效率, 本文对 2 组跨域样本进行了比较实验: 一组是 CrackForest 数据集, 另一组是由 RDD2022 中来自 6 个国家的道路图像组成的跨国数据集。实验结果如表 5 和表 6 所示。

表 5 CrackForest 数据集的实验结果

模型	精确率/%	召回率/%	mAP@0.5 /%	参数量 /M	浮点运算量/G
YOLOv11n	95.0	75.5	83.3	2.58	6.3
YOLO11-AKAD	97.7	80.4	87.3	2.09	5.1

表 6 RDD2022 数据集的实验结果

模型	精确率/%	召回率/%	mAP@0.5 /%	参数量 /M	浮点运算量/G
YOLOv11n	52.3	43.1	43.6	2.59	6.3
YOLO11-AKAD	57.8	41.8	43.8	2.09	5.1

由表 5 和表 6 可知, 在 CrackForest 数据集上, YOLO11-AKAD 精确率为 97.7%, 对比基准的 YOLOv11n 提升 2.7%, 召回率为 80.4%, 提升 4.9%, mAP@0.5 为 87.3%, 提升 4.0%; 在 RDD2022 数据集上, YOLO11-AKAD 精确率为 57.8%, 对比 YOLOv11n 提升 5.5%, mAP@0.5 为 43.8%, 提升 0.2%。同时, YOLO11-AKAD 在 2 个数据集上参数量均为 2.09 M, 对比 YOLOv11n 分别减少 0.49 M、0.5 M。这表明, YOLO11-AKAD 在不同数据域中能够在显著降低参数量与计算量的同时提升检测性能, 展现出良好的效率、轻量化特性以及跨场景的稳定表现, 具备一定的实际部署价值。然而, 在环境更复杂、采样方式更为多样的跨国家/区域数据集 RDD2022 上, 其性能提升相对有限, 说明模型的跨域泛化能

力仍需进一步提升。

3 结束语

本文针对现有道路缺陷检测方法在复杂道路背景下存在准确度低、难以满足实时性要求等问题, 以 YOLOv11n 为基准模型进行了改进。通过引入自适应核的特征提取的 C3k2_AKConv 模块, 增强模型的局部特征提取; 使用 A_{Down} 模块优化了下采样过程, 优化多尺度上下文信息传递, 减少参数量。实验结果表明, 改进的 YOLO11-AKAD 算法更适应道路缺陷目标检测特点, 算法的 mAP@0.5 达到了 81.6%, 参数量和浮点运算量分别降低至 2.09 M 和 5.1 G, 模型参数规模减少了 19%, 有一定的推广性。同时, 模型在面对复杂道路环境的泛化能力还需进一步加强, 下一步将研究在尽量保证轻量化的同

时,提升模型的泛化性能。

参考文献:

- [1] 胡璠,凌天清.基于数字图像的高速公路路面病害识别系统研究[J].黑龙江科技信息,2012(33):275-276.
- [2] MIHALJ T, LI H X, BABIĆ D, et al. Road infrastructure challenges faced by automated driving: a review[J]. Applied Sciences, 2022, 12(7):3477.
- [3] 王丹.基于机器视觉的路面缺陷检测算法研究[D].天津:河北工业大学,2022.
- [4] GUAN S Z, LIU H L, POURREZA H R, et al. Deep learning approaches in pavement distress identification: a review[EB/OL]. (2023-08-01)[2025-12-10].<https://doi.org/10.48550/arXiv.2308.00828>.
- [5] LI J, ZHAO X, LI H. Method for detecting road pavement damage based on deep learning[C]//SPIE Smart Structures + Nondestructive Evaluation. Bellingham: SPIE, 2019:517-526.
- [6] HUANG L, CHEN C, YUN J T, et al. Multiscale feature fusion convolutional neural network for indoor small target detection[J]. Frontiers in Neurorobotics, 2022, 16:881021.
- [7] 韦正璐,王家晨,刘庆华.基于改进SSD算法的路面破损检测[J].电子设计工程,2023,31(3):63-68.
- [8] 黄凯枫,张博熠,王梦,等.基于改进SSD模型的路面病害识别算法研究[J].江苏科技大学学报(自然科学版),2023,37(2):53-60.
- [9] 陈飞宇,张应迁,吴嘉懿,等.基于GOLD-YOLO改进YOLOv5模型道路病害检测研究[J].现代计算机,2024,30(19):7-12.
- [10] 胥铁峰,黄河,张红民,等.基于改进YOLOv8的轻量化道路病害检测方法[J].计算机工程与应用,2024,60(14):175-186.
- [11] 王杰,翟亚红,徐龙艳,等.基于EC-YOLO的道路缺陷检测[J].湖北汽车工业学院学报,2024,38(4):64-68.
- [12] 罗臻,蒋瑜,李维诚.基于改进YOLO11n的道路缺陷检测模型[J].微电子学与计算机,2025,42(11):25-36.
- [13] 贺智轩,陈里里,王翔,等. DMF-YOLOv11:基于改进YOLO11n的无人机航拍图像目标检测算法[J].计算机工程与应用,2025,61(14):88-100.
- [14] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. (2024-10-23)[2025-12-10]. <https://doi.org/10.48550/arXiv.2410.17725>.
- [15] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017:764-773.
- [16] CHEN Y P, DAI X Y, LIU M C, et al. Dynamic convolution: attention over convolution kernels[C]//Proc of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020:11030-11039.
- [17] ZHANG X, SONG Y Z, SONG T T, et al. AKConv: convolutional kernel with arbitrary sampled shapes and arbitrary number of Parameters[EB/OL]. (2023-11-20)[2025-11-6]. <https://arxiv.org/pdf/2311.11587>.
- [18] FANG S M, CHEN C M, LI Z J, et al. YOLO-ADual: a lightweight traffic sign detection model for a mobile driving system[J]. World Electric Vehicle Journal, 2024, 15(7):323.

责任编辑:李远辉