

融合注意力机制的轻量化路面裂缝检测方法研究

罗瑜¹, 樊赫²

(1. 陕西工业职业技术大学 电气工程学院, 陕西 咸阳 712000;

2. 西北机电工程研究所, 陕西 咸阳 712099)

摘要: 针对现有路面裂缝检测精度低、定位能力差和计算量较大等问题, 对基于 YOLO v8n 的轻量化检测模型进行研究; 采用 GhostNet v2 作为主干网络以减少参数量并提升特征提取性能, 在特征融合模块引入 CBAM 增强对路面缺陷特征的关注, 将边界框回归损失函数替换为 Focal-SIoU 以优化边界框重叠度计算; 实验结果显示, 改进后模型的精确率、召回率和全类平均正确率分别提升 4.76%、1.89% 和 2.77%; 模型大小为 YOLO v8n 的 33.33%, 推理速度达 271.5 frames · s⁻¹, 检测 1 000 张图片仅需 3.68 s; 该模型满足实时性要求, 可减少路面裂缝漏检和误检, 为路面移动端设备自动检测提供技术支持。

关键词: YOLOv8n; 路面裂缝; 目标检测; 注意力机制; 轻量化模型

Lightweight Pavement Crack Detection Algorithm Incorporating Attention Mechanism

LUO Yu¹, FAN He²

(1. School of Electrical Engineering, Shaanxi Polytechnic University, Xianyang 712000, China;

2. Northwest Institute of Mechanical and Electrical Engineering, Xianyang 712099, China)

Abstract: Aiming at the problems of low accuracy, poor positioning ability, and large computational load in existing pavement crack detection, this study investigates a lightweight detection model based on YOLO v8n. GhostNet v2 is adopted as the backbone network to reduce the number of parameters and improve feature extraction performance. The CBAM (Convolutional Block Attention Module) is introduced into the feature fusion module to enhance attention to pavement defect features, and the bounding box regression loss function is replaced with Focal-SIoU to optimize the calculation of bounding box overlap. Experimental results show that the precision, recall, and mean average precision of the improved model are increased by 4.76%, 1.89%, and 2.77%, respectively. The model size is 33.33% of that of YOLO v8n, with an inference speed of 271.5 frames · s⁻¹, and it takes only 3.68 seconds to detect 1000 images. The model meets real-time requirements, reduces missed and false detections of pavement cracks, and provides technical support for automatic detection on mobile pavement devices.

Keywords: pavement cracks; object detection; YOLOv8n; attention mechanism; lightweight algorithm

0 引言

路面裂缝的形成主要由环境因素和机械应力的持续交互作用导致^[1]。长期温度变化、降水侵蚀等自然力, 加上交通工具的反复动态压力, 通常会引起铺装层出现各种应力损伤^[2]。这一问题不仅影响路面的美观, 还对

路面行车安全产生严重威胁, 进一步加速路面的损坏, 增加养护成本^[3]。路面损伤检测作为应对路面老化问题及维护优良路况的核心环节, 快速识别并修补裂缝有助于提升行车安全系数, 进而有效遏制交通意外隐患^[4]。

现阶段, 针对路面裂缝开展检测工作主要依靠人工检测方式完成, 虽然直观、易操作, 但存在工作量大、

收稿日期:2025-06-16; 修回日期:2025-07-29。

基金项目:陕西工业职业技术大学自然科学类科研项目(2024YKYB-001)。

作者简介:罗瑜(1992-),女,硕士,讲师。

引用格式:罗瑜,樊赫.融合注意力机制的轻量化路面裂缝检测方法研究[J].计算机测量与控制,2026,34(2):65-72,86.

效率低和工作环境复杂的问题^[5]。近年来,随着深度学习技术快速发展,目标检测技术在计算机领域迈出重要一步^[6]。传统检测网络在识别过程中普遍存在响应延迟、无法实现整体自优化机制、依赖人工干预进行超参数配置以及应对海量信息时存在性能瓶颈等挑战。运算效率不足导致工程实践中难以即时解析高频率采集路面损伤样本。

深度学习目标检测模型在表面缺陷检测领域已取得一定成效,其中,单阶段检测模型如你只看一次(YOLO, you only look once)系列^[7],可实现实时检测且对光照等自然环境的适应性优于传统模型,部署简便,因此在公路工程缺陷检测处理领域得到了广泛应用^[8]。文献[9]提出基于YOLO v5模型改进的路面裂缝检测模型,全类平均正确率达87.29%,比原始模型高5.42%,模型参数量减至5.7 MB。但该模型检测的裂缝类型单一,在逆光或阴影条件下,裂缝与背景对比度降低,致使召回率下降;文献[10]通过引入改进轻量化卷积以及上采样算子改进YOLOv5s模型,以达到轻量化目的,但其实验结果表明,检测1 129张图片所用时间为39.89 s,未达到实时要求;文献[11]提出了一种基于YOLOv8s的路面损坏检测模型,实验结果表明,改进后模型全类平均正确率提高1.1%,模型计算量减少46.2%,但其应遮挡情况时,出现“漏检—误检”矛盾;文献[12]利用具有注意力机制的新型类条件生成对抗网络增强具有各种罕见路面缺陷的数据集,最终在YOLO v8模型上进行测试,验证了提出方法的可行性,但其并未考虑轻量化,不便于移动端部署。

针对上述问题,为提高网络对于不同类别、自然环境影响下路面裂缝检测能力,本研究基于YOLO v8n模型,针对路面裂缝检测的特点和难点,对模型进行了相应的改进和优化。

1) 对数据集采用多种数据增强方法,在训练样本层面构建差异化特征空间,从而增强模型对未知分布的适应性与环境扰动的抵御强度。

2) 将YOLO v8n主干网络替换为GhostNet v2^[13],其高效的特征提取和较强的鲁棒性,有助于在复杂环境中精确识别裂缝,提升模型的泛化能力,其轻量级特性可显著提高计算效率和推理速度,适应实时处理需求。

3) 为增强网络特征表达能力,使其聚焦关键特征,在特征融合模块中整合卷积块注意模块(CBAM, convolutional block attention module)^[14]。

4) 将边界框回归损失函数替换为焦点平滑交并比(Focal-SIoU, focal-smoothed intersection over union)^[15],以更精确地考量预估边界框与真实框之间的重叠程度,通过调制网络聚焦于图像关键区域,提高模型对这些区域的识别能力。

1 数据集准备

1.1 图像数据集采集

本研究图像数据集主要来自CrackForest、混凝土表面裂缝检测、路面裂缝检测等公开数据集,其余源于ImageNet、Open Images Dataset等开放数据集网站,广泛覆盖我国西北干旱地区常见路面类型。共1 245张JPG格式图像,统一裁剪至 320×320 像素分辨率,以保留裂缝缺陷区域并平衡检测速度。数据集涵盖横向、纵向、接缝、倾斜等多种裂缝类型。为适用不同场景裂缝检测,参考YOLO官方数据增强建议,对挑选图像随机分批采用随机挖空边长图像面积5%~15%像素正方形区域,模拟路面修补痕迹,且保证裂缝特征保留率 $\geq 80\%$,避免挖空面积过大导致特征丢失;采用均值0,方差0.01~0.03高斯噪声和密度0.01~0.02椒盐噪声,模拟相机传感器噪声;采用概率50%水平翻转和概率30%垂直翻转,扩展样本方向多样性;采用-30%~+30%HSV色彩空间亮度调整,以覆盖95%自然光照场景;所有参数均以“特征保留优先、训练稳定可控、场景真实匹配”为核心原则。共得到2 185张图像,各类别裂缝及其增强效果如图1所示。

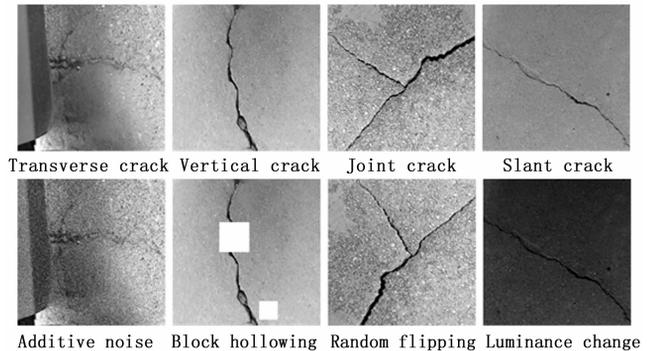


图1 各类别裂缝及其增强效果

1.2 数据集构建

利用LabelImg工具,对2 185张图像样本中不同类别路面裂缝及其位置信息进行人工标注工作。标注结果被保存为YOLO格式txt文件。随后,按照7:1:2比例进行数据集随机划分,得到训练集(1 530张)、验证集(218张)、测试集(437张)。图2展示了各类裂缝标注示例。

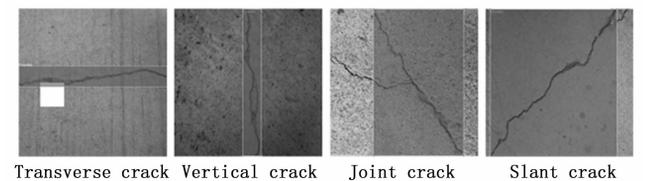


图2 各类别裂缝标注示例

2 融合注意力机制的轻量级路面裂缝检测模型

2.1 YOLO v8n 卷积神经网络模型

作为目标识别领域的前沿模型, 由 Ultralytics 公司在 2023 年提出的 YOLO v8 模型开发了 n、s、l、m 和 x 五种不同规格, 以适应多样化部署环境^[16]。各版本的参数量及存储占用如表 1 所示。

表 1 各版本 YOLO v8 模型参数量

模型	浮点运算量/G	模型体积/MB
YOLO v8n	8.2	6.3
YOLO v8s	28.6	22.1
YOLO v8m	78.9	50.9
YOLO v8l	165.3	85.7
YOLO v8x	257.8	133.7

由表 1 可知, YOLOv8n 参数量最少, 针对路面检测场景对模型推理性能的要求以及模型轻量化要求, 本研究决定采用 YOLOv8n 作为研究与应用的基础。YOLOv8n 结构包含主干 (Backbone)、颈部 (Neck) 与头部 (Head) 网络模块, 分别负责特征提取、多尺度特征融合以及检测结果输出任务。

主干网络采用 Darknet-53 作为基础框架, 通过串联跨阶段特征复用组件 (C2f, CSPDarknet53 to 2-Stage FPN) 与标准卷积单元 (CBS, convolution-batch normalization-sigmoid linear unit), 构建多层特征交互机制。在特征编码末端集成空间金字塔池化融合 (SPPF, spatial pyramid pooling fusion), 该模块通过并行池化操作实现不同粒度特征的上下文信息保留。颈部网络将路径聚合网络 (PAN, path aggregation networks) 和特征金字塔网络 (FPN, feature pyramid networks) 融合, 建立跨分辨率特征的双向交互通道, 纵横双向的特征传播体系显著提升了目标空间定位的效率, 大大提高预估边界框的产生速度, 节约计算成本。头部网络实施无预设锚点机制, 通过构建双分支独立推理单元, 分别处理目标中心概率分布与边界偏移量计算, 有效避免了传统锚框机制存在的样本分配失衡问题。

2.2 网络模型的改进

本研究中, 主要改进为将主干网络替换为 GhostNet v2, 以有效实现轻量化; 在多尺度特征融合模块, 引入 CBAM, 以显著提升模型在多尺度特征融合过程中的感知能力; 将综合交并比 (CIoU, complete IoU) 替换为 Focal-SIoU, 以改进模型的收敛速度和稳定性, 特别适合处理不规则或细小的裂

缝; 最终有助于提高路面裂缝检测任务整体性能, 获得更高检测精确率和鲁棒性。改进后的 GCFS-YOLO v8n 结构如图 3 所示, 其中, “GCV2” 代表 GhostNet v2 网络模块, “C2fGV2” 代表融合 GhostNet v2 的 C2f 模块, 其他部分为 YOLO v8n 原始网络模块。

本研究着重从以下 3 方面针对路面裂缝检测任务对 YOLO v8n 模型进行优化提升:

1) 特征提取模块替换为 GhostNet v2。本研究主干网络采用 GhostNet 系列的第二代轻量级模型 GhostNet v2, 其在第一代 GhostNet 基础上进一步提出解耦全连接 (DFC, decoupled fully connected) 模块^[17]。用于捕获长距离空间信息, 且维持轻量级卷积神经网络计算效率。

将 GhostNet v2 应用于路面裂缝检测, 能够充分捕捉长距离空间信息, 与获取图像信息的视觉系统相结合, 可实现自然环境下路面裂缝信息的实时识别。GhostNet v2 作为轻量级卷积神经网络的特性, 有效减少模型规模, 于移动端设备部署有利。GhostNet v2 由 GhostNet 的核心 Ghost 模块以及 DFC 注意力模块共同构建而成, 如式 (1) 所示:

$$O = \text{Sig}(A) \odot v(X) \quad (1)$$

其中: $\text{Sig}()$ 为 Sigmoid 函数, A 为注意力图, \odot 为按元素进行乘法运算, $v()$ 为 Ghost 模块, X 为输入特征图。

两路特征进行按元素乘法运算, 得到最终输出, 如图 4 所示。

在标准卷积运算中, 设定输入数据为 $X \in R^{h \times w \times c}$, h 、 w 、 c 分别为输入特征图高度、宽度、通道数。Ghost 模块首先对 m 个输出特征图 $Y \in R^{h' \times w' \times m}$, 完成标准卷积, 其中, h' 、 w' 分别对应输出特征图高度和宽度, 且 m 小于输出通道数 c 。随后, 借助线性操作来生

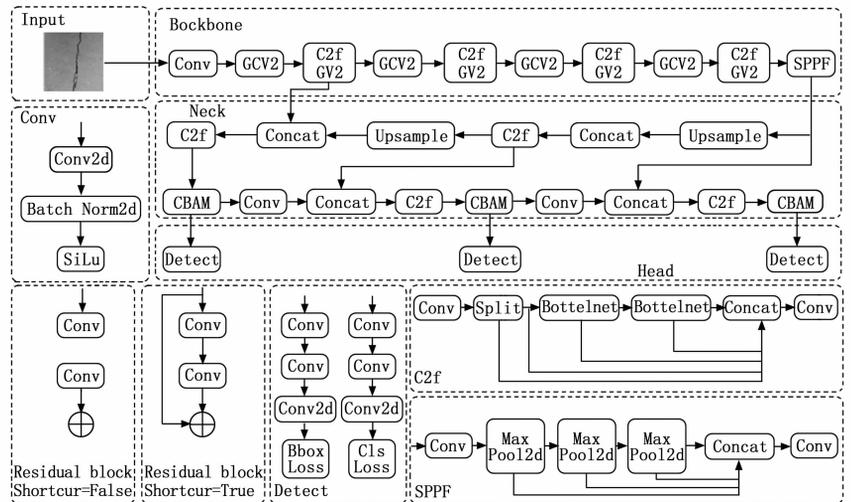


图 3 GCFS-YOLO v8n 网络结构

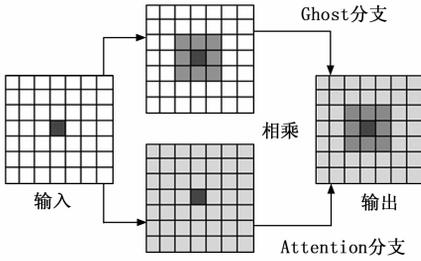


图 4 GhostNet v2 模块

成丰富的 Ghost 特征图，具体描述如公式 (2)、(3) 所示。最终，拼接生成最终特征图：

$$Y = X * f \tag{2}$$

$$y_{ij} = \Phi_{i,j}(y_i), \forall i = 1, \dots, m, j = 1, \dots, s \tag{3}$$

式中， f 为普通卷积滤波器， $*$ 表示卷积操作； y_{ij} 为最终 Ghost 特征图； $\Phi_{i,j}$ 为生成第 j 个镜像特征图 y_{ij} 的线性操作； y_i 为 Y 中第 i 个原始特征图。

DFC 注意力具备长距离建模、结构简易等优势，最终生成具备全局感受野的注意力图。首先，完成水平和垂直方向全连接层解耦操作，随后，采用两个单独全连接层，对两个方向长距离信息进行聚合。如此，计算复杂度显著下降，如式 (4)、(5) 所示：

$$a'_{hw} = \sum_{h'=1}^H F_{h,h'\omega}^H \odot z_{h'\omega}, h = 1, 2, \dots, H, \omega = 1, 2, \dots, W \tag{4}$$

$$a_{hw} = \sum_{w'=1}^W F_{w,hw'}^W \odot a'_{hw'}, h = 1, 2, \dots, H, \omega = 1, 2, \dots, W \tag{5}$$

其中： a 为生成的注意力图； F^H 和 F^W 为全连接层变换权值； z 表示特征。DFC 注意力模块结构如图 5 所示。

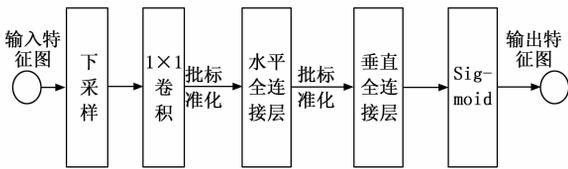


图 5 DFC 注意力模块

GhostNet v2 核心在于 GhostBottleneckV2 模块，其结果如图 6 所示。DFC 注意力模块与 Ghost_1 模块呈并行关系，扩展性增强。随后，输入到 Ghost_2 模块，进而生成输出特征。

2) CBAM 注意力机制。在实际检测工作中，由于光照条件变化、车辆等物体遮挡等因素影响，图片中干扰信息与路面裂缝特征信息相互混淆，对识别精准度造成不利影响，检测过程中易出现漏检或错检。故为进一步应对环境信息给路面裂缝特征提取带来的干扰问题，本研究在 YOLO v8n 颈部网络中引入 CBAM 模块。

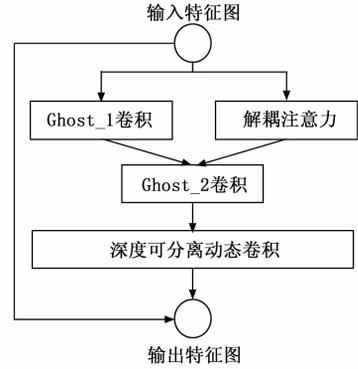


图 6 GhostBottleneckV2 模块

CBAM 模块包含通道注意力、空间注意力。在完成提取特征后，经由通道注意力和空间注意力模块对路面裂缝特征图进行信息提炼，可提升模型颈部网络对特征准确提取能力^[18]。CBAM 注意力模块结构如图 7 所示。

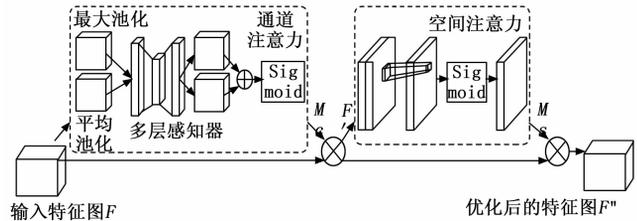


图 7 CBAM 注意力模块结构示意图

通道注意力模块先执行全局平均和最大池化操作，进而得到中间特征图。之后，借助 Sigmoid 函数算得权重系数，并与中间特征图做乘法运算，最终得到输出特征图。具体计算方法如式 (6) 所示：

$$M_c(F) = \text{Sig}\{MLP[\text{AvgPool}(F)] + MLP[\text{MaxPool}(F)]\} \tag{6}$$

式中，通道注意力图用 M_c 表示；输入特征图用 F 表示；多层感知神经网络则为 $MLP()$ 。

空间注意力模块以 $M_c(F)$ 作为输入，同样先执行两种池化操作，随后，在通道维度对池化结果进行拼接处理，再利用卷积操作将通道数量缩减为 1，经由 Sigmoid 函数生成权重系数，并与 $M_c(F)$ 做乘法运算，得到输出特征图。其计算方法如式 (7) 所示：

$$M_s(F) = \text{Sig}\{f^{7 \times 7}\{[\text{AvgPool}(F'); \text{MaxPool}(F')]\}\} \tag{7}$$

式中， M_s 为空间注意力图； $f^{7 \times 7}$ 为 7×7 卷积； F' 为输出特征图。

CBAM 模块通过构建“通道注意力—空间注意力”双分支结构以强化特征表达。在通道注意力模块中，对输入特征图执行全局平均池化和最大池化操作，随后利用多层感知机生成相应通道权重，从而增强裂缝边缘特

征通道权重。而在空间注意力模块中, 对经过通道权重加权的特征图应用 7×7 卷积操作, 以生成空间权重, 进而抑制非裂缝区域响应。

3) Focal-SIoU 损失函数。YOLO v8n 默认采用 CIoU 损失函数, 完成对边界框位置进行预测, 将预估边界框的纵横比损失纳入考量, 然而未充分考虑数据集样本平衡以及预估边界框方向问题, 尤其是在纵横比模糊的情况下。故此, 本研究引入带有角度惩罚机制的 Siou 进行边界框回归预测, 同时, 引入 Focal 机制, 以缓解正负样本数量差异^[15]; IoU 为预估与真实边界框交集与并集面积比值, 用于衡量目标检测的精准度, 如公式 (8) 所示:

$$IoU = \frac{|B \cap B_T|}{|B \cup B_T|} \quad (8)$$

式中, B 表示预估边界框区域, B_T 表示真实边界框区域。

Siou 计算主要涵盖 4 个代价函数, 分别是角度、距离、形状以及 IoU , 计算原理如图 8 所示。

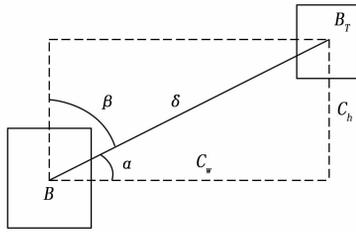


图 8 Siou 参数示意图

角度代价函数用于感知角度信息, 如公式 (9) 所示:

$$\Delta = 1 - 2\sin^2(\arcsin(x) - \frac{\pi}{4}) \quad (9)$$

其中:

$$x = c_h / \sigma = \sin(\alpha) \quad (10)$$

$$\sigma = \sqrt{(b_{cx}^{gt} - b_{cx})^2 - (b_{cy}^{gt} - b_{cy})^2} \quad (11)$$

$$c_h = \max(b_{cy}^{gt}, b_{cy}) - \min(b_{cy}^{gt}, b_{cy}) \quad (12)$$

位置代价会随着角度代价而改变, 如公式 (13) 所示:

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho_t}) \quad (13)$$

其中:

$$\rho_x = \left(\frac{b_{cx}^{gt} - b_{cx}}{c_w} \right), \rho_y = \left(\frac{b_{cy}^{gt} - b_{cy}}{c_h} \right), \gamma = 2 - \Delta \quad (14)$$

式中, $(b_{cx}^{gt}, b_{cy}^{gt})$ 为真实边界框中心点, (b_{cx}, b_{cy}) 则是预估边界框中心点。

依据距离代价计算形状代价, 如公式 (15) 所示:

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta \quad (15)$$

其中:

$$\bar{\omega}_w = \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})}, \bar{\omega}_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (16)$$

式中, ω^{gt} 、 h^{gt} 为真实边界框宽和高; $\theta=1$ 。

综上, Siou 整体损失函数计算公式如公式 (17) 所示:

$$L_{Siou} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (17)$$

为平衡对高低样本重要性评估, 本研究引入 Focal 机制并将其与 Siou 融合, Focal 机制通过引入平衡因子 τ 对样本权重调整, 让模型更聚焦于难区分困难样本。 τ 可根据实际情况进行灵活调整, 从而以实现更优效果, 具体如公式 (18) 所示:

$$L_{Focal-Siou} = -(1 - IoU)^\tau \log(IoU) L_{Siou} \quad (18)$$

由两种机制形成的 $L_{Focal-Siou}$ 函数, 能避免模型对易于区分的样本过度自信, 同时增强对难区分困难样本检测性能, 如此, 有效减弱路面复杂环境的干扰影响, 提升模型的检测性能。

2.3 训练环境与方法

2.3.1 试验环境

本研究训练所使用的硬件和软件环境如表 2 所示。

表 2 软硬件平台配置

名称	版本信息
系统	Linux
中央处理器	Intel® Xeon® Silver 4214 CPU @ 2.20 GHz
显卡	NVIDIA GeForce RTX 3090 24 G 内存
Python	3.11
深度学习框架和库	Pytorch 2.2.1, Cuda 12.1, cuDNN 8.9.6

训练超参数设置如下: lr0 设置为 0.012, 动量设置为 0.937, 批量大小设置为 8, 多线程设置为 4, 优化器为随机梯度下降^[19], 训练完整迭代 250 epoch, 早停机制为 30 次, 采用余弦退火衰减法^[20]调整学习率, 预热阶段的动量参数和权值衰减参数分别设定为 0.8 和 0.00045。

2.3.2 评价指标

为对各模型在路面裂缝检测任务中性能进行衡量, 本研究选用以下评价指标: 精确率 (P, Precision)、参数量、召回率 (R, recall)、权重尺寸、全类平均正确率 (mAP, mean average precision) 和帧率 (FPS, frames per second)^[21]。各指标计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (19)$$

$$R = \frac{TP}{TP + FN} \quad (20)$$

$$mAP = \frac{\sum_{t=1}^N P(R) dR}{N} \quad (21)$$

式中, TP 、 FP 分别指被预测为正类的正负样本数;

FN 指被预测为负类的正样本数；用 AP 表示模型 PR 曲线与坐标轴围成面积，主要用于衡量模型针对单一目标的检测性能， AP 值越高，模型性能越好； mAP 是所有类别 AP 值平均数。 N 代表检测任务总类别数，本研究类别 N 为 1。

3 结果与分析

3.1 消融试验

本研究通过消融试验验证对 YOLO v8n 应用 GhostNet v2、CBAM 以及 Focal-SIoU 改进策略的可行性。试验结果见表 3。

表 3 消融试验

编号	GhostNet v2	CBAM	Focal-SIoU	$P/\%$	$R/\%$	$mAP/\%$	参数量/ 10^6	权重大小/MB
1	/	/	/	86.72	83.64	85.71	3.01	6.3
2	✓	/	/	85.50	82.23	85.10	0.86	2.0
3	/	✓	/	90.81	83.77	86.82	3.10	6.4
4	/	/	✓	89.14	83.29	86.38	3.01	6.3
5	✓	✓	/	90.85	83.65	87.08	0.89	2.1
6	✓	✓	✓	91.27	84.71	88.04	0.89	2.1

由表 3 可知，引入 GhostNet v2 后，模型权重大小仅占 YOLO v8n 的 31.75%。此时，模型 P 、 R 和 mAP 略微下降，其原因在于 GhostNet v2 借助生成成本较低 Ghost 特征图，降低模型计算量和参数量，同时导致模型在识别目标过程时，对正样本判断精确度下降，最终对 mAP 产生影响；在颈部网络引入 CBAM 后，模型参数量增加了 0.09×10^6 ，权重大小增加了 0.1 MB，相应的模型整体检测性能也随之提升，表明 CBAM 的通道注意力模块自动学习通道重要性权重，增强关键通道特征表征，空间注意力模块聚焦目标空间位置，减少背景干扰，二者共同作用，实现特征优化；采用 Focal-SIoU 后，模型整体检测性能得到优化，其原因在于 Focal 机制提升了模型对难辨别样本关注度，Siou 则优化边界框回归损失，进而提升定位精度。两者结合产生协同效应，增强复杂场景适应性，确保目标定位精确性。

当同时引入 GhostNet v2 和 CBAM 时，与试验 2 和 3 相比，模型权重在缩减至原始模型的 33.33%，并且 P 、 R 和 mAP 相较于原始模型亦有所提升，表明 GhostNet v2 和 CBAM 模块能够有效与 YOLO v8n 模型结合，充分发挥各自优势；最后，将 GhostNet v2、CBAM 和 Focal-SIoU 全部引入后，模型 P 、 R 和 mAP 分别达到 91.27%、84.71% 和 88.04%，相对于其他组试验，模型性能最优，且模型参数量和权重仅为原始 YOLO v8n 模型的 29.57% 和 33.33%，表明采用本研究提出的 3 种改进策略后，模型同时满足高精度与轻量化需求。

3.2 轻量化网络对比试验

本试验以 YOLO v8n 为基础模型，采用 Mobile-

NetV2^[22]、MobileNetV3^[23]、ShuffleNetV2^[24] 以及 GhostNet v2 等主流轻量化特征提取网络，仅对模型主干网络替换，开展对比试验。具体结果如表 4 所示。

表 4 轻量化特征提取网络的对比试验

轻量化特征提取网络	$P/\%$	$R/\%$	$mAP/\%$	参数量/ 10^6	权重大小/MB
MobileNetV2	79.25	78.02	78.11	0.75	1.8
MobileNetV3	85.15	81.62	84.72	1.19	2.7
ShuffleNetV2	80.95	77.29	79.28	0.97	2.2
GhostNet v2	85.50	82.23	85.10	0.86	2.0

根据表 4 数据，将主干网络替换为 MobileNetV2，模型性能最差；当 GhostNet v2 作为主干，参数量和权重大小分别仅为 0.86×10^6 、2.0 MB，且模型 P 、 R 和 mAP 方面均展现出显著的优势。其中， mAP 较其他网络分别提升 6.99%、0.38% 和 5.82%。

3.3 损失函数对比试验

原始 YOLO v8n 边界框回归损失函数选用 Ciou，对传统 IOU 指标做出改进。然而，Ciou 对路面裂缝形状变化不敏感，当目标检测框的长宽比例存在较大差异时，将导致梯度更新不稳定，进而对检测模型整体性能造成负面影响，即 mAP 值产生影响。为此，本研究开展了对比试验，对 YOLO v8n 模型在分别采用 Ciou、Siou 和 Focal-SIoU 时的 mAP 变化进行评估。实验结果如图 9 所示。

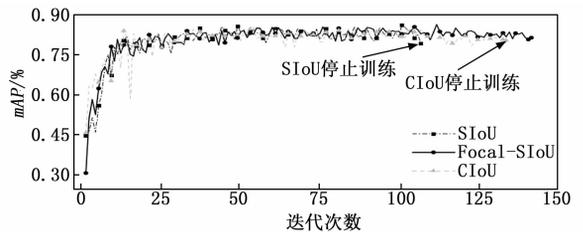


图 9 多种损失函数对比试验

由图 9 可知，由于早停机制，三组试验皆在 150 轮内完成训练，说明模型结构与数据集特征高度匹配，能够迅速学习到数据中的规律。在使用 Focal-SIoU 后，模型在训练过程中 mAP 值变化较为平稳，且趋于稳定后， mAP 值高于 Ciou 和 Siou，结合表 3 数据，在路面裂缝检测任务中，Focal-SIoU 可有效提高模型性能。

3.4 注意力机制对比试验

以 YOLO v8n 为基础模型，采用压缩-激励注意力 (SE, squeeze-and-excitation)^[25]、高效通道注意力 (ECA, efficient channel attention)^[26]、卷积神经网络注意力 (simAM, simple, parameter-free attention module for convolutional neural networks)^[27] 以及 CBAM 进行颈部注意力模块替换，展开对比试验。每组训练 5 次并记录 mAP 值，具体结果如图 10 所示。

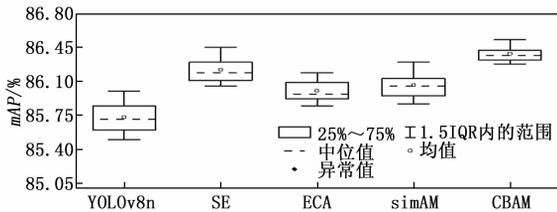


图 10 *mAP* 箱体图对比试验

由图 10 可知, 5 组试验都没有出现异常值, 说明原始 YOLO v8n 模型能够较好地学习给定数据集中学习到裂缝特征, 注意力机制也能够基于现有数据分布的基础上合理地调整模型的关注重点。试验结果表明, 相较于仅通道加权的 SE、计算效率高但无空间建模的 ECA 以及仅特征增强的 simAM, 引入 CBAM 后, 训练得到的 *mAP* 数值普遍较高, 且更趋于稳定。

3.5 主流检测模型对比试验

本研究将改进后的 GCFS-YOLO v8n 模型与 Improved-YOLOv5n^[9]、YOLOv6n^[28]、YOLOv8n、MED-YOLOv8n^[11]、YOLO v10n^[29]、YOLOv11n^[30]、Deformable-DETR^[31] and RT-DETR^[32] 等主流单阶段目标检测模型进行对比, 使用同一权重文件进行训练, 试验结果见表 5。

表 5 不同模型路面裂缝检测结果对比

模型	<i>P</i> /%	<i>R</i> /%	<i>mAP</i> /%	参数量/ 10 ⁶	权重尺寸 /MB	<i>FPS</i> / frames · s ⁻¹
Improved-YOLOv5n	86.51	82.82	85.27	2.08	4.42	200.2
YOLOv6n	90.09	82.99	87.18	4.23	8.7	116.9
YOLOv8n	86.72	83.64	85.71	3.01	6.3	148.36
MED-YOLOv8n	87.22	83.55	86.34	1.66	3.5	241.2
YOLOv10n	86.04	81.97	85.38	2.70	5.8	145.4
YOLOv11n	90.81	84.19	87.98	2.59	5.5	156.5
Deformable-DETR	89.51	84.39	87.69	40.0	457.5	35.1
RT-DETR	91.54	84.86	88.26	19.9	38.7	60.7
GCFS-YOLOv8n	91.27	84.71	88.04	0.89	2.1	271.5

表 5 数据显示, GCFS-YOLOv8n 模型的精确率、召回率和 *mAP* 均优于实验 1~7 中的模型。虽其性能略低于 RT-DETR 模型, 但 RT-DETR 模型的参数量高达 19.9×10^6 , 模型体积达 38.7 MB, 无法满足轻量化需求。此外, GCFS-YOLOv8n 模型的推理速度可达 $271.5 \text{ frames} \cdot \text{s}^{-1}$, 检测 1 000 张图片仅需 3.68 s, 显著优于其他模型, 完全符合实时性要求。

综上, 实验结果表明, 本研究提出的 GCFS-YOLOv8n 模型在路面裂缝检测任务中, 有效平衡了轻量化特性与性能表现。因此, 比其他主流单阶段网络模型具有更显著的优势。

为进一步验证改进后的 GCFS-YOLO v8n 模型在复

杂环境中对路面裂缝的检测性能, 本研究于测试集 437 张图像中随机抽取 5 张未参与训练的复杂场景图像, 利用已完成训练得到保存模型的状态字典和结构 .pt 文件以及 detect.py 进行图像检测。图 11 展示了部分 YOLO v8n 模型与 GCFS-YOLO v8n 模型的检测结果。

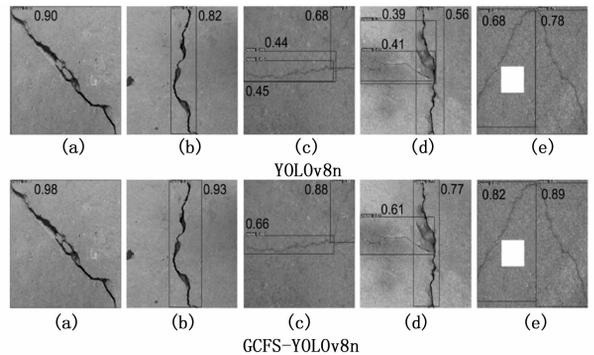


图 11 原始 YOLO v8n 与 GCFS-YOLO v8n 检测效果对比

图 11 (a)、(b) 展示原始 YOLO v8n 与 GCFS-YOLO v8n 在正常光照无遮挡条件下的检测结果, GCFS-YOLO v8n 置信度 0.98, 远高于 YOLO v8n 的 0.90。图 11 (c)、(d) 对比暗光照和亮光照场景, YOLO v8n 误检, GCFS-YOLO v8n 通过 CBAM 模块抑制背景干扰, 无误检且置信度普遍较高。图 11 (e) 显示, 在雨后水渍下, GCFS-YOLO v8n 的 Focal-SIoU 损失函数优化边界框回归, 置信框更贴合且置信度普遍较高。综上所述, 复杂环境下本研究所提模型检测性能比原始 YOLO v8n 更具优势。

5 结束语

本研究设计轻量化路面裂缝检测模型 GCFS-YOLO v8n。该模型以 YOLO v8n 单阶段目标检测模型为基础, 在主干网络引入 GhostNet v2, 参数量降低 68.25%, 检测性能超越 MobileNet 和 ShuffleNet 轻量化主干网络; 在颈部网络引入 CBAM 网络, 有效融合多尺度特征信息, 提升路面裂缝识别能力; 采用 Focal - Siou 后, *mAP* 提升 0.67%, 训练中优于 CIoU 和 Siou。融合 3 种改进策略后, GCFS-YOLO v8n 在测试集上 *P*、*R* 和 *mAP* 分别达 91.27%、84.71%、80.04%。训练完成后, 模型参数量为 0.89×10^6 , 权重大小 2.1 MB, 较原始 YOLO v8n 权重缩减 66.67%, *mAP* 提升 2.33%, 与主流单阶段目标检测模型相比优势明显。通过复杂环境下路面裂缝检测性能对比试验, 证实 GCFS-YOLO v8n 更适用于户外自然环境下路面裂缝精准检测, 为后续路面自动检测和管理提供技术支持。

GCFS-YOLO v8n 模型可通过轻量化特性直接部署于车载嵌入式设备或无人机巡检系统, 与现有路面养护管理平台通过 API 接口实现检测数据实时传输, 支持裂缝位置坐标、尺寸及严重程度的结构化输出, 形成

“采集—分析—养护”闭环管理；成本效益方面，相比人工检测，人均日检 0.5 km，搭载该模型的智能检测车可实现日均 30 km 作业效率提升 60 倍，且单张图片检测耗时 3.68 ms，可满足车载实时性要求。

本研究模型提升了路面裂缝识别精度并实现轻量化设计，但数据集未完全覆盖所有缺陷类型和复杂背景，可能导致模型在边缘设备运行时面临挑战。未来研究中，为应对实际应用中面临的雨、雪、雾覆盖挑战，可通过引入多光谱图像融合技术与迁移学习策略优化模型泛化能力，针对边缘设备算力限制问题，可采用 INT8 精度模型量化压缩与 NVIDIA Jetson 系列边缘计算模块硬件加速进一步降低部署门槛，最终实现路面病害检测从“事后修复”向“预防性养护”的模式转变。

参考文献：

- [1] 《中国公路学报》编辑部. 中国路面工程学术研究综述 · 2024 [J]. 中国公路学报, 2024, 37 (3): 1-81.
- [2] RUSSEL N S, SELVARAJ A. MultiScaleCrackNet: a parallel multiscale deep CNN architecture for concrete crack classification [J]. Expert Systems with Applications, 2024, 249 (PB): 123658.
- [3] 许正森, 雷相达, 管海燕. 多尺度局部特征增强 Transformer 道路裂缝检测模型 [J]. 中国图象图形学报, 2023, 28 (4): 1019-1028.
- [4] 冉蓉, 徐兴华, 邱少华, 等. 基于深度卷积神经网络的裂纹检测方法综述 [J]. 计算机工程与应用, 2021, 57 (9): 23-35.
- [5] CAO W M, LIU Q F, HE Z Q. Review of pavement defect detection methods [J]. IEEE Access, 2020, 8: 14531-14544.
- [6] TIAN D, HAN Y, WANG S, et al. Absolute size IoU loss for the bounding box regression of the object detection [J]. Neurocomputing, 2022, 500: 1029-1040.
- [7] 冷月香, 王健. 基于改进的 YOLOv5 绝缘子检测与识别模型 [J]. 计算机测量与控制, 2025, 33 (4): 95-101.
- [8] TAFIDA A, ALALOUL S W, ZAWAWI W B A N, et al. Advancing smart transportation: a review of computer vision and photogrammetry in learning-based dimensional road pavement defect detection [J]. Computer Science Review, 2025, 56: 100729.
- [9] ZHANG P J, LI W W, WENG Y. Detection of road crack defects based on improved YOLOv5s model [J]. Journal of Physics: Conference Series, 2022, 2303 (1): 012019.
- [10] 刘鹏杰, 姚金杰, 高晶, 等. 一种轻量化网络的道路病害检测方法研究 [J]. 计算机测量与控制, 2025, 33 (4): 40-47.
- [11] ZHAO M H, SU Y H, WANG J X, et al. MED-YOLOv8s: a new real-time road crack, pothole, and patch detection model [J]. Journal of Real-Time Image Processing, 2024, 21 (2): 26.
- [12] ORTIZ C S, ORTIZ S E, IGLESIAS L L, et al. Leveraging a deep learning generative model to enhance recognition of minor asphalt defects [J]. Scientific Reports, 2024, 14 (1): 28904.
- [13] ZHANG J, TIAN M, YANG Z, et al. An improved target detection method based on YOLOv5 in natural orchard environments [J]. Computers and Electronics in Agriculture, 2024, 219: 108780.
- [14] 李良福, 蒲应丹, 黎光耀, 等. 基于 MD-CBAM 的多样性裂缝图像修复方法 [J]. 光电子·激光, 2024, 35 (4): 351-359.
- [15] LI C, DU Y, ZHANG X, et al. YOLOGX: an improved forest fire detection algorithm based on YOLOv8 [J]. Frontiers in Environmental Science, 2025, 12: 1486212.
- [16] BAKIRCI M. Advanced aerial monitoring and vehicle classification for intelligent transportation systems with YOLOv8 variants [J]. Journal of Network and Computer Applications, 2025, 237: 104134.
- [17] ZHAI X Y, WEI H L, WU H D, et al. Multi-target tracking algorithm in aquaculture monitoring based on deep learning [J]. Ocean Engineering, 2023, 289 (P1): 116005.
- [18] SHUAI L Y, MU J, JIANG X Q, et al. An improved YOLOv5-based method for multi-species tea shoot detection and picking point location in complex backgrounds [J]. Biosystems Engineering, 2023, 231: 117-132.
- [19] LANCONELLI A, LAURIA S C. A note on diffusion limits for stochastic gradient descent [J]. Journal of Approximation Theory, 2025, 309: 106160.
- [20] 闫向彤, 罗嘉伟, 曹现刚. WGAN-GP 结合 CBAM-VGG16 轻量化网络滚动轴承故障诊断 [J]. 噪声与振动控制, 2024, 44 (5): 120-127.
- [21] 梁忠超, 黄苗, 胡兴, 等. 基于 YOLO 网络的车辆四点标定测距研究 [J]. 机械工程学报, 2023, 59 (10): 226-235.
- [22] AL-GAASHANI M S A M, XU W B, OBSIE E Y. MobileNetV2-based deep learning architecture with progressive transfer learning for accurate monkeypox detection [J]. Applied Soft Computing, 2025, 169: 112553.
- [23] SUN J D, YOU J L, LI F M, et al. Research on improvement strategies for a lightweight multi-object weed detection network based on YOLOv5 [J]. Crop Protection, 2024, 186: 106912.

(下转第 86 页)