基于改进 YOLOv8n 的道路裂缝检测轻量化模型

朱佳慧　刘艺　张登银

南京邮电大学物联网学院　南京　210003

摘 要：针对道路裂缝外观特征易受环境干扰、细小裂缝漏检率高、检测设备计算资源受限的问题，提出了轻量级检测模型 MCA-YOLO-A。该模型基于 YOLOv8n，用更轻量的 MobileNetV3 特征提取网络来代替原主干网络，并融合了精确空间信息的 CA 注意力模块，提高了特征提取能力。同时，引入了适用于轻量级网络的 Alpha-IOU 损失函数，使得网络整体性能提升。此外，增加了小目标检测层，提升细小裂缝的识别精度。MCA-YOLO-A 模型在道路裂缝数据集上平均精度均值 mAP_0.5 和 F1 分数分别达到 0.930 和 0.893，相较于原 YOLOv8n 模型提升了 7.0%和 9.7%，参数量仅为 6.0M，减少了 4.8%，检测速度达到 95 帧/s。实验结果证明，该模型具备高精度、轻量化以及出色的泛化能力，更适合应用于计算资源受限的嵌入式系统和移动终端等场景。

关键词：道路裂缝；图像检测；深度可分离卷积；YOLOv8；注意力模块；轻量化

中图分类号：TP391

基金项目：国家自然科学基金（62471241）；江苏省研究生科研与实践创新计划项目（No.KYCX23_1051）

A Lightweight Road Crack Detection Model Based on improved YOLOv8n

ZHU Jiahui, LIU Yi, ZHANG Dengyin

School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Abstract: To address the challenges of road crack appearance characteristics being susceptible to environmental interference, high miss detection rate of fine cracks, and limited computational resources of inspection equipment, a lightweight detection model, MCA-YOLO-A, is proposed. The model is based on YOLOv8n, replacing the original backbone with a lighter MobileNetV3 feature extraction network, and integrating a CA attention module that accurately captures spatial information, thereby enhancing the capability of feature extraction. Meanwhile, the Alpha-IOU loss function suitable for lightweight networks is introduced, which makes the overall performance of the network improved. In addition, a small target detection layer is added to improve the recognition accuracy of fine cracks. The average precision of MCA-YOLO-A model on road crack data sets is 0.930 and 0.893, respectively, which is 7.0% and 9.7% higher than that of the original YOLOv8n model, and the parameter quantity is only 6.0M, which is 4.8% lower, and the detection speed reaches 95 frames/s. The experimental results demonstrate that the model is highly accurate, lightweight, and capable of generalization, which makes it more suitable for deployment in scenarios with limited computational resources such as embedded systems and mobile devices.

Key Words: Road cracks, Image detection, Depthwise separable convolution, YOLOv8, Attention module, Lightweighting

引　言

公共道路及交通系统的运行效率、可靠性与安全性是社会发展的重要支撑，而当前我国公路正面临着人口增长、基础设施退化、建设与维护成本快速上涨等严峻挑战。裂缝是一种常见的路面病害，若不加注意，细小的裂缝将演变为较大的裂缝，给行车安全带来很大的危害。中国公路里程数已逾 543 万公里，若仍使用人工检测裂缝，测试过程耗时耗力，且对小裂缝的漏检率高，不及时的检测会带来安全隐患和高昂的修复成本。因此，探索能够快速准确地识别道路裂缝的自动检测技术，对提高公路养护水平，保障行车安全具有重要意义。

目前，国内外对道路裂缝自动检测技术的研究已取得一定成果。例如，配备红外或传感器设备的坑洞检测车辆[1-2]，但这类设备硬件成本较高，因此相应的检测费用也较高。许多研究人员开始使用高效、低成本的图像处理技术来检测道路裂缝。传统的图像处理方法主要有数学形态学处理[3-4]、边缘检测[5]、多特征提取与融合[6]等设计网络结构实现目标。传统的图像处理方法复杂度较低，适合于裂缝明显和背景简单的图像，当裂缝图像背景复杂、噪声大、裂缝和背景亮度相近时，传统方法的检测精度就会大幅度下降[7]，并且通常需要手动标记特征，不能提供快速和全自动的分析。

与传统图像处理技术相比，基于深度学习理论的图像处理技术具有更高的精度、更快的速度、更好的抗噪声能力和可嵌入性[8]。陈泽斌等[9]使用改进的 U-Net 网络对路面裂缝进行提取，该网络需要对每个像素进行分类，这使得它在处理速度上相对较慢，同时需要较高的计算资源和内存。Tong 等[10]基于 Faster RCNN 模型对管道裂缝进行检测，引入数据增强和上下文池化模块，模型检测精度达到 0.817，但是检测速度有待提高。近年来，由于嵌入式道路裂缝检测设备如无人机、行车记录仪等存储空间有限，且在自动驾驶和实时监控系统中，需要实时处理，许多研究人员致力于对单阶段检测模型进行改进实现目标检测，例如 YOLO（You Only Look Once）[11-14]、SSD（Single Shot Multibox Detector）[15]等。这种模型仅需一次类别预测与位置回归，因此卷积运算可以更高效地共享，实现更快的处理速度和更低的内存消耗，但不能很好地提取小目标敏感的浅层高分辨率特

征[16]。李生辉等[17]在 YOLOv5 的颈部网络引入深度可分离卷积 DSC 和通道 MLP 构建的 EC-MLP 模块来提供更充分的目标上下文特征。然而 EC-MLP 在每个通道内执行全连接操作，带来了更多的参数和计算量，检测效率 FPS 仅为 76。Li 等[18]提出了一种利用 SimAM 注意力机制和 GHostConv 改进 YOLOv8 的道路缺陷检测算法。但 SimAM 注意力机制涉及大量矩阵运算和优化，计算代价较高，模型收敛速度慢，且 GHostConv 的简化卷积操作导致特征提取的稀疏性，容易丢失小目标对象。Su 等[19]考虑到 YOLO 系列算法中可能存在的信道信息丢失和感受野不足的问题，在 YOLOX 的基础上设计了一种名为 MOD-YOLO 的道路缺陷检测模型，但将其部署到边缘设备仍然存在一定挑战。

上述方法虽然都可以较为精准地识别研究目标，但是很少考虑到所提方法是否能够兼顾精度、参数量、计算量及模型大小。基于此，本研究基于 YOLOv8n 提出一种轻量级裂缝检测模型 MCA-YOLO-A，主要设计思想是在提取裂缝特征时尽量获取全局感受野和完备的空间位置信息，在尽可能保证模型轻量化的同时，进一步优化算法，以促进模型的加速收敛和性能提升。本研究的主要创新有 3 点：（1）构建了嵌入精确空间信息的轻量级主干网络，将 YOLOv8n 的主干网络替换为融合坐标注意力模块的改进 MobileNetv3 网络，促进特征提取中裂缝和空间特征的有效结合，同时极大地减少了模型大小、参数量和计算复杂度，进一步提升检测精度，实现了模型检测精度与轻量化的有效平衡；（2）创新性地引入一种适用于轻量级网络的边界框回归损失函数 Alpha-IOU，通过自适应重新加权与改善损失函数的正态分布、对称性以及方差均衡，显著提高了模型的收敛速度和边界框回归精度，使模型具有更强的鲁棒性，有效解决现有方法中引入大量网络层和模块，使得加速模型收敛变得更加困难的问题；（3）重新设计了检测头结构，在重构模型的颈部网络中增加 160×160 像素的检测层，弥补了现有方法经过下采样和卷积操作后，易丢失重要的浅层位置信息，对于小目标尤其是小于 8×8 像素裂缝识别时的不足。

1 MCA-YOLO-A 模型设计

YOLOv8n 的核心结构由主干网络（Backbone）、颈部网络（Neck）和头部

网络（Head）组成。主干网络以 C2f 模块为基本单元，相较于 YOLOv5 采用的 C3 模块，C2f 模块具备更优秀的特征提取能力，然而堆叠了更多的瓶颈结构，不可避免地带来过多的通道信息冗余和较大的计算量。头部网络使用解耦头结构，将分类和检测头分离，有效地提升检测精度，同时边界框回归损失采用 IOU Loss 和 DFL 的组合方式，尽管 IOU Loss 有 CIOU，DIOU，GIOU 三种 Loss 可选择，但都有其局限性。颈部网络使用 PAN-FPN 结构来改善多尺度特征的融合，但小目标的特征可能在不同层次之间的融合中丢失，导致检测精度不足。由于道路裂缝形态多样、与路面纹理区分度较低且目前检测算法难以适用于计算资源受限的边缘设备，本文提出一种基于 YOLOv8n 的道路裂缝检测模型 MCA-YOLO-A，用更轻量的 MobileNetV3 特征提取网络来代替原主干，并在轻量级卷积神经网络 MobileNetV3 中嵌入坐标注意力（Coordinate Attention，CA）模块，减少了模型所需的计算量与网络参数量，同时融合了精确的空间信息。在头部预测结构中，引入 Alpha-IOU 损失函数替换边界框回归损失函数通过优化 IOU 值减小了边界框预测偏差，提升了模型收敛速度和鲁棒性。此外，为了提高模型对于细小裂缝特征的识别能力，在颈部网络中增加小目标检测层，有效地提升了模型的检测精度。改进后的 MCA-YOLO-A 模型网络结构如图 1 所示。

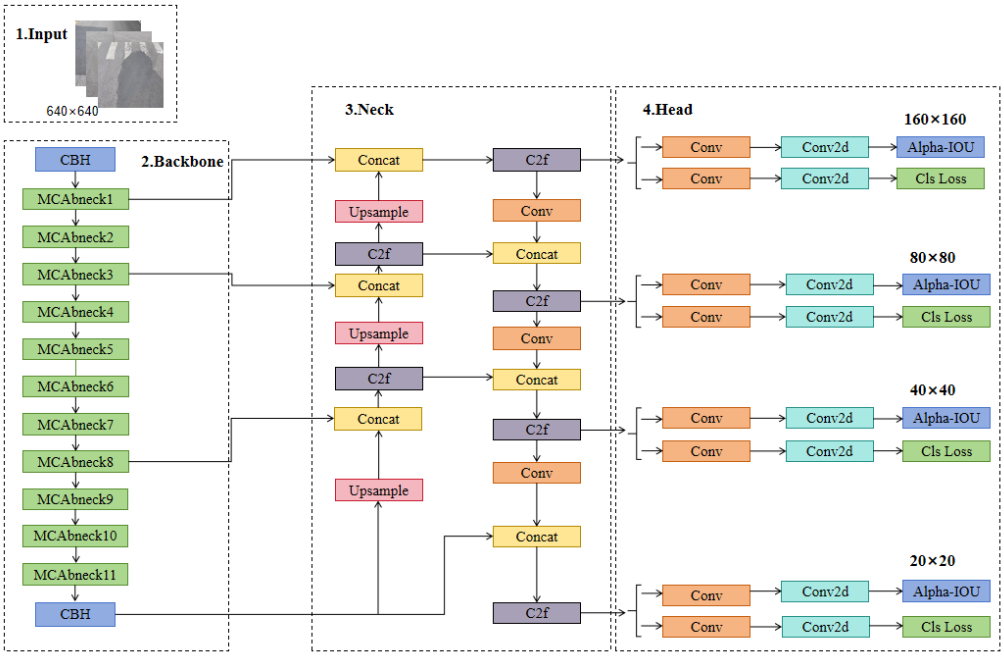


图 1 MCA-YOLO-A 网络架构图

Fig. 1 Network architecture of MCA-YOLO-A

1.1 重构的主干特征提取网络

MobileNetV3[20]是一种轻量级卷积神经网络，在保持高精度的同时，具备更低的计算开销和更快的推理速度，非常适合在移动设备和嵌入式设备上进行高效的图像处理和目标检测任务。因此，本研究选用 MobileNetV3 特征提取结构作为主干网络进行特征提取，图 2 显示了 MobileNetV3 的特征提取网络结构，其核心是采用深度可分离卷积将传统的卷积操作分解为深度卷积（depthwise convolution，DW）和逐点卷积（pointwise convolution，PW）两个更简单的步骤，以减少计算量与参数量。同时采用倒置残差结构与线性瓶颈设计，以有效提取特征并降低低维特征信息的损失，另外集成了压缩与激励(Squeeze-and-excitation，SE)注意力机制，有助于增强不同通道间的特征选择能力。

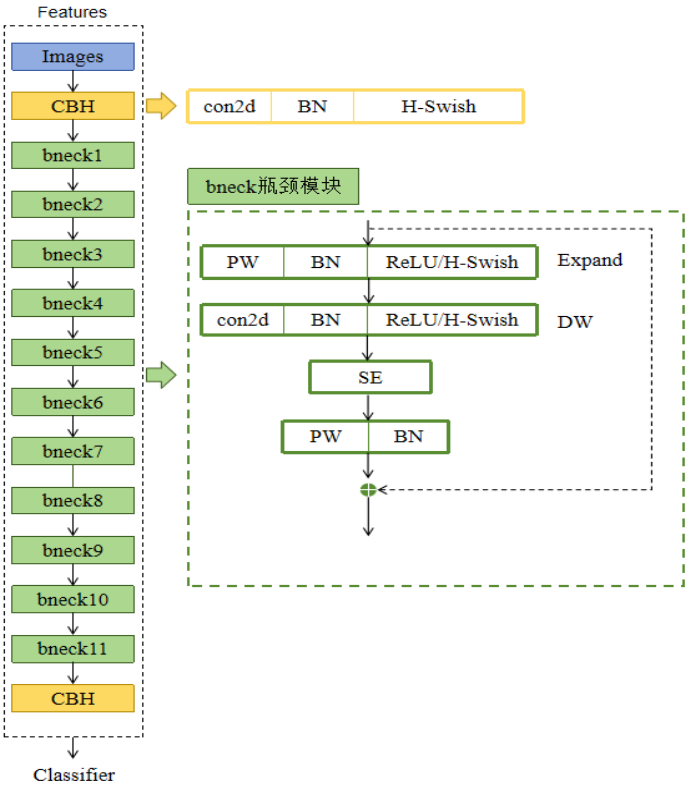


图 2 MobileNetV3 特征提取结构图

Fig. 2 Structure diagram of MobileNetV3 feature extraction

具体的，设输入特征量为 $D_f \times D_f \times M$，卷积核尺寸为 $D_k \times D_k$，输出特征量为 $D_f \times D_f \times N$，M 与 N 分别表示输入通道数和输出通道数，D_f 表示输入和输出特征图的空间维度，f 表示特征图高度值和特征图宽度值，D_k 表示卷积核的大小，k 表示卷积核高度值和宽度值。

深度可分离卷积和标准卷积的计算量之比，可表示为：

$$\frac{C_2}{C_1}=\frac{\alpha D_k\times D_k\times M\times \beta D_f+\alpha^2 M\times N\times \beta D_f\times \beta D_f}{D_k\times D_k\times M\times N\times D_f\times D_f}=\frac{\alpha\beta^2}{N}+\frac{\alpha^2\beta^2}{{D_k}^2} \qquad (1)$$

其中，C_1表示标准卷积的计算量，C_2表示深度可分离卷积的计算量，α表示调节特征通道数量的宽度系数，β表示调节特征分辨率的分辨率系数。

由式（1）可知，深度可分离卷积一次操作就可减少$\frac{\alpha\beta^2}{N}+\frac{\alpha^2\beta^2}{D^2}$的计算量，另外通过控制宽度因子和分辨率因子，从而更好地减少基于深度可分离卷积神经网络 MobileNetV3 的参数量与计算量。

原 MobileNetV3 特征提取网络瓶颈模块（Bottleneck，bneck）使用压缩与激励(Squeeze-and-excitation，SE) 模块，通过跨通道间的相互作用来提升特征的表达力。然而对于裂缝图像这类特定场景，由于其内在的通道间相关性较弱，传统的通道注意力模块在捕捉和强化这些通道特征关系可能会受限。同时，研究中通过 RDD2022（2022 年道路损坏数据集）观察到的是，大多数的道路裂缝都发生在图片的下半部分，而不同位置的裂缝可能会由于相机的视角而有不同的外观，而 SE 模块的压缩操作会损失特征图中关键的空间位置信息。因此，本文在 MobileNetV3 的瓶颈模块中引入融合精确空间信息的轻量级 CA[21]模块。该模块对 SE 模块中的全局平均池化进行了水平和垂直方向上的分解，从而在保留通道间关系的同时，捕获了更为精细的空间位置信息，提高了轻量级模块对特征的表达能力。改进后的 MobileNetv3 瓶颈模块 bneck 结构如图 3 所示。

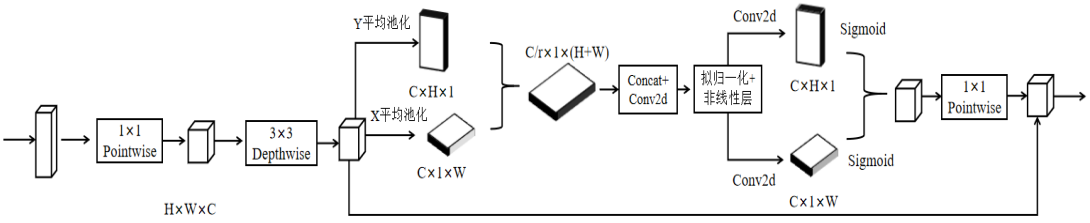


图 3 改进的 MobileNetV3 瓶颈模块结构图

Fig. 3 Structure of the improved MobileNetV3 bottleneck module

设输入特征图的通道数、高度和宽度分别为C、H、W。首先，沿 X、Y 轴，将 SE 模块的全局平均池化（式（2））分解成一维感知注意力特征图：

$$Z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i, j) \tag{2}$$

$$Z_c^h = \frac{1}{W} \sum_{0 \le i \le W} x_c(h, i) \tag{3}$$

$$Z_c^w = \frac{1}{H} \sum_{0 \le j \le H} x_c(j, w) \tag{4}$$

其中，Z 为输出，x_c 为第 c 个通道的二维特征，i 、j 为输出特征的坐标位置，w 和 h 为第 c 个通道对应的卷积核权重。

进一步地，对两个一维特征图进行拼接，并作为卷积变换函数 $F1$ 的输入，生成中间特征图 f ，包含水平和垂直方向的空间信息：

$$f = \partial(F1([z^h, z^w])) \tag{5}$$

其中，[.,.]为空间维度上的拼接操作， $f \in R^{c/r \times (H+W)}$ ， ∂ 为非线性激活函数，r 为一个调节模块尺寸的超参数。

然后，将 f 按空间维度分解成 2 个单独的张量 $f^h \in R^{c/r \times (H+W)}$ 和 $f^w \in R^{c/r \times (H+W)}$ 。再使用卷积变换函数 F_h 与 F_w 分别将 f^h 与 f^w 转换为与输入特征具有相同通道数的张量：

$$g^h = \sigma(F_h(f^h)) \tag{6}$$

$$g^w = \sigma(F_w(f^w)) \tag{7}$$

其中，σ 是 *Sigmoid* 函数。g^h 表示对张量 f^h 进行卷积 F_h 变换得到的注意力向量，$g^h \in R^{C \times H \times 1}$ ，g^w 表示对张量 f^w 进行卷积 F_w 变换得到的注意力向量，$g^w \in R^{C \times 1 \times W}$ 。

最终，对所述输入特征图扩展为坐标注意力模块的输出特征：

$$y_{c(i,j)} = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \tag{8}$$

其中，$y_{c(i,j)}$ 、$x_{c(i,j)}$ 表示的是输出和输入特征图在第 c 通道、位置为（i，j）的值，$g_c^h(i)$ 表示注意力向量 g^h 在通道 c 上水平位置 i 的注意力权重， $g_c^w(i)$ 表示注意力向量 g^w 在通道 c 上竖直位置 j 的注意力权重。

1.2 Alpha-IOU 边界框回归损失函数优化

边界框（Bounding Box，bbox）回归通过预测目标的 bbox 在图像/视频中的位置实现定位。利用损失函数求出预测框与真实框之间的误差，逐步增强模型对目标位置预测的准确性。YOLOv8 采用基于交并比 IOU 的损失，并给出了三种改进 IOU 的选择：完全 IOU(CIOU)、广义 IOU(GIOU)和距离 IOU(DIOU)，然而它们在处理非重叠边界框、小目标检测、以及边界框形状和大小差异较大的场景中仍然存在局限性。

Alpha-IOU 提供了现有基于 IOU 损失的统一幂泛化，通过自适应地重新加权高 IoU 和低 IoU 对象的损失和梯度，优化了基于 IOU 的边界框回归的准确性[22]。这种动态调整机制相对于现有损失，通过调整 α 值，可以针对不同的训练阶段或数据集特性进行优化，提供了更多的灵活性，同时文献[22]在多个基准目标检测数据集和模型上实验证明，Alpha-IOU 损失可以一致地超越现有的基于 IOU 的损失，并在小数据集和轻量模型，以及噪声较大的边界框中提供更好的鲁棒性。

Alpha-IOU 采用 Box-Cox 变换[23]，显著优化了损失函数的正态分布与方差均衡，大大提升了边界框回归的准确性。

Box-Cox 变换表达式为：

$$O(\alpha)=\begin{cases}\dfrac{O^{\alpha}-1}{\alpha}, \alpha \neq 0 \\ \log(O), \alpha = 0\end{cases} \tag{9}$$

其中，O 为连续变量，且要求取值为正，α 是用来调整数据正态分布的参数。

应用 Box-Cox 变换，将 IOU 损失（$Loss_{IOU}=1-IOU$）推广为 Alpha-IOU 损失 $Loss_{\alpha-IOU}$，有：

$$Loss_{\alpha-IOU}=\frac{1-IOU^{\alpha}}{\alpha}, \alpha>0 \tag{10}$$

由于基于 IOU 的主流损失函数中 α 值大于 0，化简得到：

$$Loss_{\alpha-IOU}=1-IOU^{\alpha} \tag{11}$$

引入惩罚项 $R(b,\ b^{gt})$，将上述 $Loss_{\alpha-IOU}$ 扩展为一个更通用的形式，b 和 b^{gt} 分别表示预测框与真实框的中心点坐标，有：

$$Loss_{\alpha-IOU}=1-IOU^{\alpha_1}+R(b,b^{gt})^{\alpha_2} \quad (12)$$

通过实验，发现 $Loss_{\alpha-IOU}$ 对 α_2 不敏感，为了简化计算，有：

$$Loss_{\alpha-IOU}=1-IOU^{\alpha}+R(b,b^{gt})^{\alpha} \quad (13)$$

于是可以将现有的基于 IOU 的损失（CIOU、DIOU 和 GIOU）推广到一个新的 Alpha-IoU 损失家族。

$$Loss_{CIOU}=1-IOU+\frac{\rho^2(b,b^{gt})}{l^2}+\beta v \rightarrow Loss_{\alpha-CIOU}=1-IOU^{\alpha}+\frac{\rho^{2\alpha}(b,b^{gt})}{l^{2\alpha}}+(\beta v)^{\alpha}$$

$$Loss_{DIOU}=1-IOU+\frac{\rho^2(b,b^{gt})}{l^2} \rightarrow Loss_{\alpha-DIOU}=1-IOU^{\alpha}+\frac{\rho^{2\alpha}(b,b^{gt})}{l^{2\alpha}} \quad (14)$$

$$Loss_{GIOU}=1-IOU+\frac{a-u}{a} \rightarrow Loss_{\alpha-GIOU}=1-IOU^{\alpha}+(\frac{a-u}{a})^{\alpha}$$

其中，IOU 代表预测边界框和真实边界框相交与相并面积之比，l 代表两边界框的最小外接矩形的对角线长度，$\rho^2(b,b^{gt})$ 代表两个边界框中心点的欧式距离，α 代表真实和预测边界框的最小外接矩形的面积，u 代表真实和预测边界框的相并面积。β 是权重函数，v 衡量宽高比的相似度，

1.3 颈部网络增加小目标检测层

原 YOLOv8 模型在特征融合后，检测层会输出 20×20、40×40 和 80×80 三种尺寸的特征图，以实现对不同大小目标的检测。较小尺寸的特征图感受野较大，同时语义信息丰富，但在局部细节上较为模糊，因此更适用于检测大目标。而最大尺寸的 80×80 像素特征图，仅能用于检测大于 8×8 像素的目标，那么在实际道路裂缝检测中，如果裂缝的高度和宽度均小于 8 像素，加之裂缝在经过下采样和卷积操作后，会造成大量浅层位置信息丢失，原始模型可能无法准确捕捉目标的特征信息从而导致漏检。为了解决这个问题，本研究增设专门针对小目标的检测层，在模型中纳入 160×160 像素的检测特征图，使网络更关注细小裂缝的检测，从而提升检测效果。

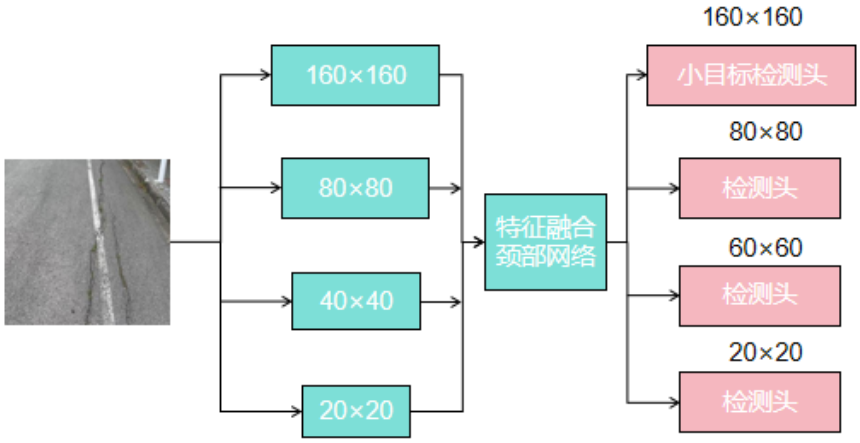


图 4 添加小目标检测层

Fig. 4 adds a small target detection layer

首先，通过 80×80 特征层与颈部网络中第三层上采样特征层拼接，再通过 C2f 特征融合，构建具有小目标特征的深层语义特征层。随后，将所述深层语义特征层和其他浅层位置特征层做拼接，最终提升 160×160 图像特征层对裂缝语义、位置等信息的表达能力。最后，将该融合特征层的结果送至头部网络中新增的解耦头进行处理，具体如图 1 所示。优化后新增的 160×160 像素的检测特征图可以有效保留浅层有效位置信息，进而使模型对道路裂缝的检测更加全面和准确。

2 模型的训练与试验

2.1 数据集构建

本文构建的道路裂缝图像数据集来源于 RDD2022 道路损伤公开数据集，RDD2022 道路损伤数据集收集了来自日本、印度、捷克共和国、中国等六个国家/地区的道路裂缝图像，具有 34,702 个地面真实标签，所述地面真实标签包括边界框和损伤类型。考虑各国国家道路类型存在差异，如果直接在这个数据集上训练深度学习模型，它在中国的道路上将无法取得较好的效果。因此本文选取其中来自中国的使用智能手机拍摄的 2477 张道路裂缝数据，如图 5 所示，其中包括四种不同道路裂缝类型的带标注图片：纵向裂缝（D00）、横向裂缝（D10）、鳄鱼裂缝（D20）和路面坑洞（D40）。

(a)D00 (b) D10 (c) D20 (d) D40

图 5 不同类型的道路裂缝图像

Fig. 5 Images of different types of road cracks

在构建数据集时，为更有效地评估模型的泛化能力，将 2477 张道路裂缝图像按照 7∶1.5∶1.5 的比例随机划分为训练集（1677 张）、验证集（400 张）、测试集（400 张）。如图 6 所示，本文数据集中分别有 2124 个纵向裂缝（D00）、860 个横向裂缝（D10）、524 个鳄鱼裂缝（D20）和 186 个路面坑洞（D40）。

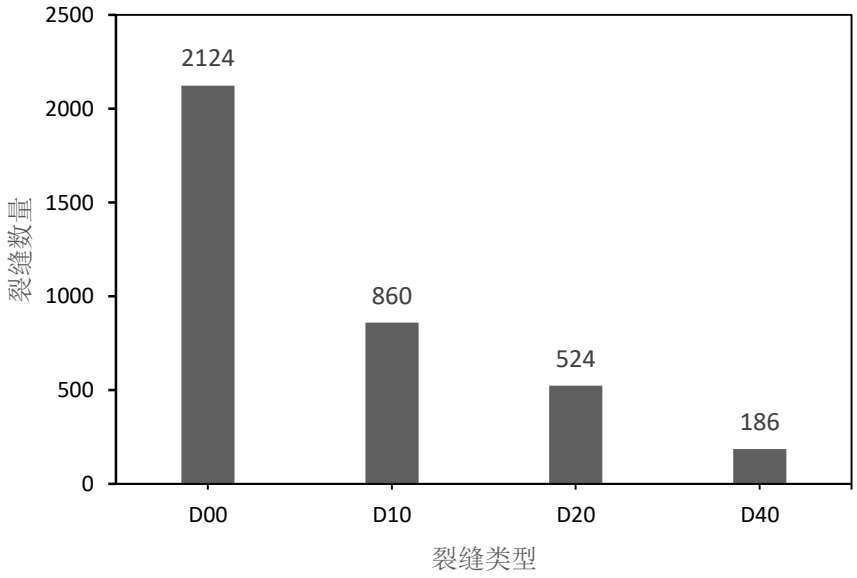


图 6 本文数据集中裂缝类型的数量分布

Fig. 6 Number distribution of crack types in the dataset in this paper

2.2 训练环境

本研究使用的算法模型的训练环境如表 1 所示。

表 1 道路裂缝检测研究的训练环境

Table 1 Training environment for road crack detection

配置	配置名称	详细信息
硬件配置	CPU	Intel(R) Core(TM) i9-12900H
	运行内存大小	16GB
	GPU	NVIDIA RTX 3060

软件配置	显存大小	6GB
	操作系统	Windows 11
	Python 版本	3.8
	深度学习框架	2.3.1
	CUDA	11.8

2.3 评价指标

为了客观评估本文所提模型的性能，本研究采用 *F1* 分数（*F1 Score*）、平均精度均值 mAP_0.5、模型大小、参数量（Params）、FPS、每秒 10 亿次浮点运算数（giga floatingpoint operations per second，GFLOPs）等参数作为模型的评价指标。其中 mAP_0.5 和 F1 分数是衡量检测精度的关键指标，它们的值越高，说明模型的检测能力越强；FPS 是衡量模型实时性的指标，其数值越高，表明模型在单位时间内能处理的图像数量越多，实时性能越好。GFLOPs 是衡量模型计算复杂度的指标，它与模型的大小和参数量一起，反映了模型的轻量化程度这些指标的值越低，意味着模型对硬件性能的依赖性越小，轻量化程度越高。

2.4 实验参数设置

实验中模型输入图像尺寸为 640×640，并配合使用 YOLOv8 默认数据增强策略。在模型训练的过程中，配合优化器随机梯度下降（Stochastic Gradient Descent，SGD）调整学习率，详细实验参数如表 2 所示。

表 2 详细实验参数

Table 2 Detailed experimental parameters

参数	值	参数	值
hsv_h	0.015	epochs	100
hsv s	0.7	batch	16
hsv v	0.4	imgsz	640
degree	0.0	workers	8
translate	0.1	seed	0
scale	0.5	close_mosaic	10
shear	0.0	fraction	1.0
perspective	0.0	mask_ratio	4
flipud	0.0	dropout	0.0

fliplr	0.5	lrf	0.01
mosaic	1.0	momentum	0.937
mixup	0.0	label_smoothing	0.00

3 实验结果与分析

3.1 不同主干特征提取网络对比实验

以YOLOv8n模型为基础，本文选取主流轻量级卷积神经网络EfficientNet-B0、ShuffleNetV2、MobileNetV2、MobileNetV3 代替原主干 DarkNet-53 网络进行对比实验，验证 MobileNetV3 作为 YOLOv8n 模型主干特征提取网络的优点，实验结果如表 3 所示。从表 3 中可知，MobileNetV2 精度表现最差，相较原 YOLOv8n 模型 mAP_0.5 下降了 32.5%，同时在实验设置的迭代次数 100 轮内未能收敛，但其轻量级特性和高帧率在资源受限的环境中可能更有优势。ShuffleNetV2 在精度和效率之间表现比较均衡，但总体来看不如 MobileNetV3 和 EfficientNet-B0。特殊的，EfficientNet-B0 的精度表现也很优秀，mAP_0.5、F1 分数分别比原 YOLOv8n 模型提升了 3.7%、6.7%，然而由于 EfficientNet 使用大量高数据读写操作的网络特性，GPU 算力发挥不充分，最终导致检测速度表现仅为 83 帧率，比原模型下降了 47.5%。综合来看，MobileNetV3 表现最佳，在检测精度 mAP_0.5 和 F1 分数方面均高于其他主流轻量级卷积神经网络，mAP_0.5、F1 分数分别比原 YOLOv8n 模型提升了 4.6%、7.4%。同时，参数量、模型大小、GFLOPs 分别降低了 7.9%、8.6%、20.9%，仅次于 MobileNetV2，并且保持了较高的帧率。综上，证明了 MobileNetV3 特征提取网络替换原 YOLOv8n 模型主干网络的有效性和高效性。

表 3 YOLOv8n 不同主干特征提取网络对比实验

Table 3 Comparison experiments of different feature extraction networks for YOLOv8n

网络模型	主干网络	mAP_0.5	*F1* 分数	Params/10^6	模型大小/MB	GFLOPs	FPS
YOLOv8n	DarkNet-53	0.869	0.814	3.01	6.3	8.1	158
YOLOv8n	EfficientNet-B0	0.901	0.869	7.21	15.8	13.4	**83**
YOLOv8n	ShuffleNetV2	0.803	0.755	6.38	13.0	16.4	137

YOLOv8n	MobileNetV2	0.586	0.574	0.71	1.7	2.6	134
YOLOv8n	MobileNetV3	**0.909**	**0.875**	**2.77**	**5.8**	**6.4**	**122**

3.2 不同注意力模块嵌入 MobileNetv3 对比实验

为进一步验证坐标注意力 CA 模块嵌入 MobileNetV3 相较于其他注意力模块的优势，本文进行不同注意力模块的对比试验，包括坐标注意力 CA 模块，结合通道注意力和空间注意力的双重注意力 CBAM（Convolutional Block Attention）模块，高效通道注意力 ECA（Efficient Channel Attention）模块。本节实验均在本文重构的主干特征提取网络为 MobileNetV3 的 YOLOv8n 模型上进行。具体实验结果如表 4 所示。

由表 4 可知，本研究嵌入 ECA 模块后的精度最低，但由于 ECA 模块轻量级的设计，计算开销较低，因此在检测速度 FPS 上比原模型提升了 6。嵌入 CBAM 模块时，网络的检测精度相比原主干网络 MobileNetv3 精度进一步提升，提升可达 1.1%，然而由于既需进行通道维度又需进行空间维度的计算，模型的检测速度下降最多，相较于添加之前 FPS 下降了 26.2%。而嵌入 CA 注意力模块时，模型的综合性能最好，模型的参数量为 2.77 M，平均检测精度 mAP_0.5 为 0.921，与原重构模型相比，模型的检测精度提高了 1.3%，同时模型参数量、GFLOPs 几乎不变，在检测速度方面也保持了 117 的 FPS 值。

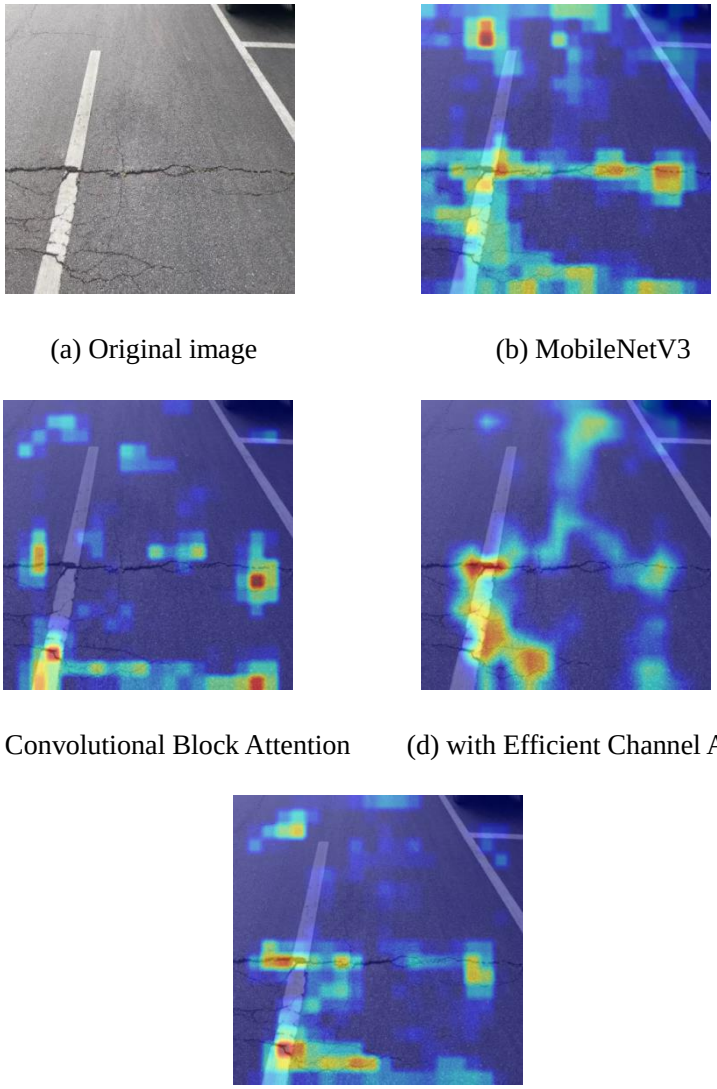
表 4 不同注意力模块对比实验

Table 4 Comparative experiments with different attention modules

主干网络	mAP_0.5	Params/10^6	GFLOPs	FPS
MobileNetV3	0.909	2.77	6.4	122
MobileNetV3+CA	**0.921**	**2.80**	**6.5**	**117**
MobileNetV3+CBAM	0.919	2.78	6.5	97
MobileNetV3+ECA	0.906	2.77	6.4	128

为直观地分析模型的注意力区域，本文通过热力图对比分析主干网络中不同注意力机制目标捕获能力的差异。可视化结果在图 7 中。深红色区域为焦点区域，表明模型更关注此区域的目标。由图 7 可知，主干网络为 MobileNetv3 的基准模型易受到背景或车辆等因素的影响，关注范围在非裂缝区域中占据了相当的比例。嵌入 CBAM 注意力模块时，网络在捕获连续密集裂缝信息方面能力较弱，嵌入

ECA 注意力模块时，网络对于远处的细小裂缝关注度低，从而影响了检测精度，而 CA 模块能更精准地识别裂缝的空间位置，从而迅速定位到关注目标，所以本文采用 CA 注意力模块嵌入主干网络更有助于道路裂缝目标的定位。



(a) Original image　　(b) MobileNetV3

(c) with Convolutional Block Attention　　(d) with Efficient Channel Attention

(e) with Coordinate Attention

图 7 主干网络嵌入不同注意力模块可视化结果

Fig. 7 Backbone network embedded with different attention modules visualization results

3.3 不同损失函数对比实验

为了进一步验证 Alpha-IOU 损失函数的有效性，对比重构的 MobileNet 模型采用 CIOU、GIOU、DIOU、Alpha-IOU 作为边界框损失函数训练时模型的平均精度 mAP_0.5、*F1* 分数和收敛情况，实验结果如图 8、图 9 所示。本节实验均在将 CA 模块嵌入 MobileNetV3 的重构主干网络的 YOLOv8n 模型上进行。

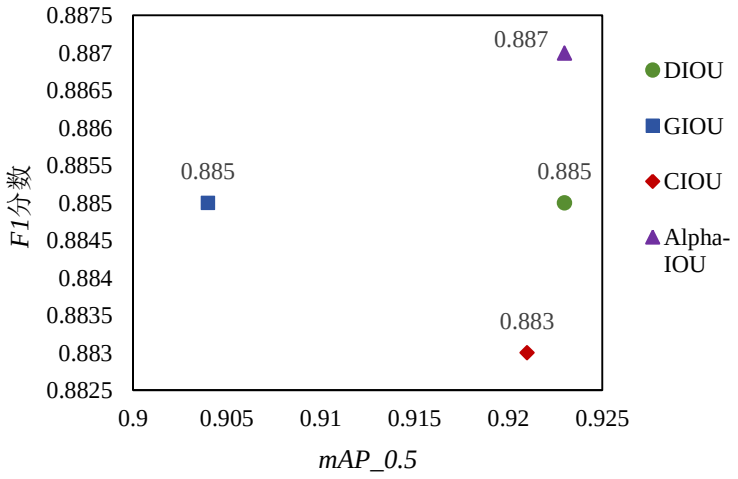


图 8 不同损失函数模型平均精度与 *F1* 分数对比图

Fig. 8 Plot of average accuracy vs. F1 score for different loss function models

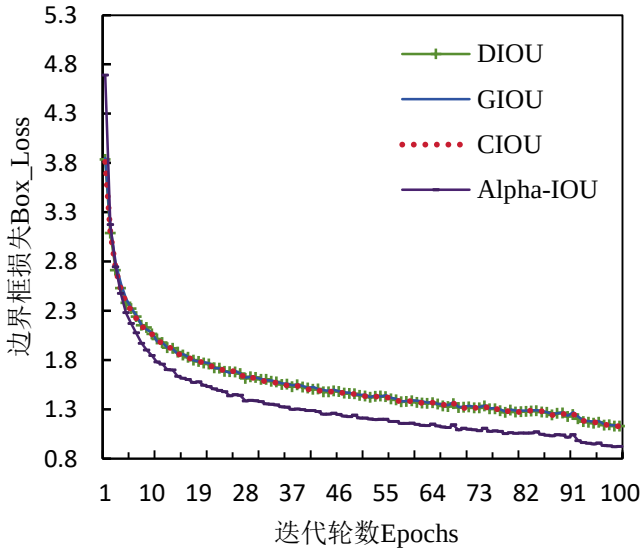


图 9 不同边界框损失函数收敛情况

Fig. 9 Convergence of loss functions for different bounding boxes

由图 8 可知，GIOU 和 DIOU 的 *F1* 分数有相同程度的提升，但 GIOU 对于不同裂缝类别检测的平均性能却略显不足，mAP_0.5 下降了 1.8%，而 DIOU 的整体性能仅略强于 CIOU。Alpha-IOU 的 mAP_0.5 和 *F1* 分数相较于 CIOU 分别提高了 0.2%，拥有最高的 *F1* 分数。并且由图 9 可知，其收敛速度同样明显优于其他损失函数，表明采用 Alpha-IOU 损失函数能够显著提高模型的收敛速度和回归精度。

图 10 给出了使用 Alpha-IOU（上行）和 CIOU（下行）损失函数在相同测试

集上的预测结果。Alpha-IOU 比 CIOU 表现得更好，因为它可以更准确地定位目标，因此可以检测到更多的真阳性目标（图像 1、图像 3）和更少的假阳性目标（图像 2）。

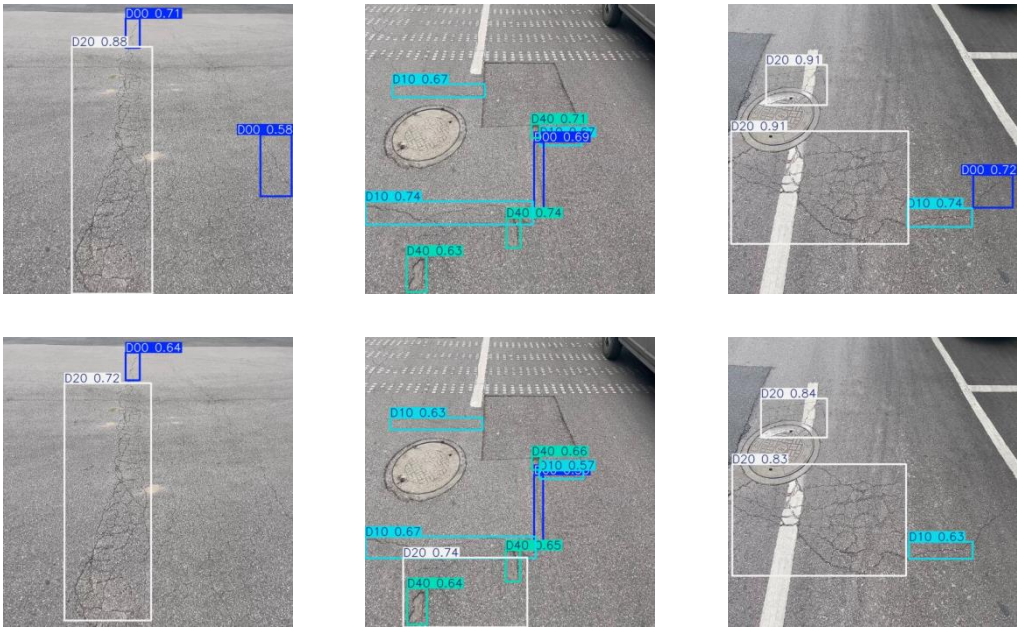


图 10 使用不同损失函数的检测情况

Fig. 10 Detection using different loss functions

3.4 消融实验

本研究在原 YOLOv8n 基础上，以 MobileNetv3 轻量型网络结构作为主干网络，降低了模型参数量，并嵌入坐标注意力 CA 模块，提升对特征图关键区域的聚焦能力；在计算边界框回归损失时使用 Alpha-IOU 损失函数，加速网络的收敛；在颈部网络中增加 160×160 小目标检测层，提高模型对于细小裂缝特征的识别能力。 为验证本研究提出的 3 种改进策略在道路裂缝检测中的性能优势，以 YOLOv8n 模型作为基础网络，进行了消融实验，具体实验结果如表 5 所示。

表 5 消融实验结果对比

Table 5 Comparison of results of ablation experiments

序号	重构轻量级主干网络	Alpha-IOU Loss	小目标检测层	mAP_0.5	*F1* 分数	Params/10^6	FPS
1	×	×	×	0.869	0.814	3.01	158
2	√	×	×	0.921	0.883	2.81	117

3	×	√	×	0.876	0.810	3.01	132
4	×	×	√	0.872	0.820	3.00	128
5	√	√	×	0.923	0.887	2.81	104
6	√	×	√	0.911	0.877	2.72	93
7	√	√	√	**0.930**	**0.893**	**2.72**	95

注：“√”表示运用了该改进策略，“×”表示未使用该改进策略。

如表 5 所示，本文在 YOLOv8n 基础模型之上进行改进，采用轻量级的 MobileNetv3 网络来提取特征，以此减少模型参数量，嵌入坐标注意力 CA 模块，有利于网络精准把握裂缝目标的空间位置，相较于原 YOLOv8n 模型，mAP_0.5 提升 5.9 个百分点，F1 分数提升 8.4 个百分点，在少量降低检测速度的同时模型参数量减少 6.6%；接着通过使用 Alpha-IOU 损失函数作为边界框回归损失函数，在保持模型参数量和帧率大小几乎不变的同时，mAP_0.5 和 *F1* 分数分别又提升了 0.002 和 0.005；最后再添加了小目标检测层，使网络更关注细小裂缝的检测，从而提升精度，以牺牲极少的 FPS 换来了 mAP_0.5 达到 0.930，*F1* 分数达到 0.893，相较于原始 YOLOv8n 模型，分别提升 7.0 个百分点和 9.7 个百分点，同时参数量减少了 9.6%，FPS 达 95 满足实时检测的要求。通过消融实验表明本文的改进策略均具有积极的意义，同时兼顾了精度与轻量化之间的平衡。

3.5 不同检测模型对比实验

为评估本文提出的 MCA-YOLO-A 模型优越的性能，将本文模型与几种近些年性能最优、应用最广泛的目标检测模型进行了对比实验。实验模型包括 Faster R-CNN[24]、YOLOv5s、YOLOv6s[25]、YOLO8s、YOLOv9-tiny[26]以及 RT-DETR[27]模型，本研究在相同的数据集和实验条件下进行了试验，且模型均已收敛。具有实验结果如表 6 所示。

表 6 不同模型性能比较

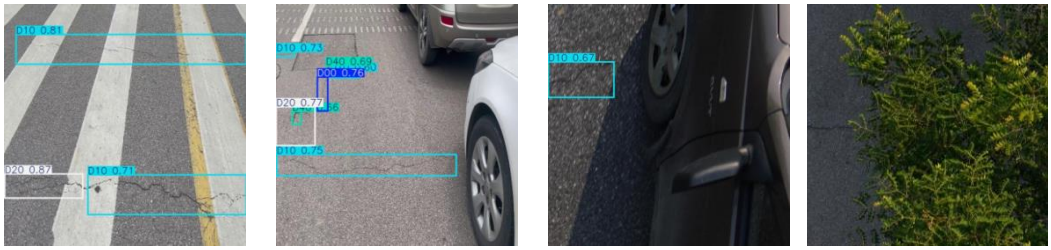
Table 6 Comparison of performance of different models

模型	mAP_0.5	*F1* 分数	Params/10^6	模型大小/MB	GFLOPs	FPS
Faster R-CNN	0.791	*0.657*	137	108.0	251.4	10
YOLOv5s	0.881	0.831	9.12	14.1	10.6	44

YOLOv6s	0.841	0.812	16.3	32.8	44.2	102
YOLO8s	0.909	0.878	11.12	22.5	28.7	110
YOLOv9-tiny	0.880	0.838	2.01	4.7	7.7	79
RT-DETR	0.866	0.834	19.9	34.3	110	43
MCA-YOLO-A	**0.930**	**0.893**	**2.72**	**6.0**	**8.1**	**95**

可以看出本文 MCA-YOLO-A 模型相比于其他主流的目标检测网络拥有最高的 mAP_0.5 和 *F1* 分数。二阶段目标检测算法 Faster RCNN 检测精度低，且算法的参数量巨大，导致算法的检测速度较低，FPS 仅有 10。本文模型与目前较新的目标检测网络 YOLOv8s、YOLOv5s、YOLOv6s、YOLOv9t 和 RT-DETR 相比，平均检测精度 mAP_0.5 分别提升 2.3%、5.5%、10.5%、5.7%和 7.3%，F1 分数分别提升 1.7%、7.4%、9.9%、6.5%和 7.1%。同时本文模型的参数量仅为 2.72，模型大小仅为 6.0，计算复杂度 GFLOPs 仅为 8.4，仅次于 YOLOv9-tiny 模型，但在检测速度上达到了 95FPS，比 YOLOv9 提升了 20.1%。综上，MCA-YOLO-A 模型相较于其他模型具有更高检测精度和稳定性，同时 FPS 达 95 满足实时检测的要求，具有良好的实际应用价值。

为了更直观地呈现改进后模型的性能，特别挑选其中表现较好的对比模型 YOLO8s、YOLOv5s 和 RT-DETR 与本文 MCA-YOLO-A 模型进行了四组可视化效果的比较，第一组选取具有连续不同裂缝特征的图像，第二组选取具有多种裂缝类型且有复杂背景干扰的图像，第三组选取有阴影遮挡的图像，第四组选取雨天环境且遮挡极端严重的图像，检测结果对比展示在图 11 中。在第一组，相比于其他模型在此类图像中容易漏检误检，MCA-YOLO-A 能有效捕捉细微裂缝特征，可以较好地将连续的网状裂缝和横向裂缝区分开来，同时拥有较高的置信度。在第二组，即使道路上存在车辆、减速带等多种背景干扰，裂缝目标小且多，只有 MCA-YOLO-A 不存在漏检误检情况。在第三组，裂缝被汽车阴影遮挡，光照强度不高，仅有 MCA-YOLO-A 顺利检出。在第四组，由于雨天光照极低，且有大量树木遮挡，所有模型均未检出。再次验证了 MCA-YOLO-A 模型在裂缝检测方面的可靠性，也突显了其在不同条件下的稳定性，不过也存在一些局限，例如由于裂缝特征的微妙差异，对于形状相似的不同类型裂缝可能存在错误检测，光线严重不足时也可能出现漏检情况。



（a）MCA-YOLO-A

（b）YOLOv8s

（c）YOLOv5s

（d）RT-DETR

（e）Original label

图 11 检测效果对比

Fig.11 Comparison of detection effect

3.6 泛化性实验

为了验证 MCA-YOLO-A 模型的实际泛化能力，本文选取公开数据集

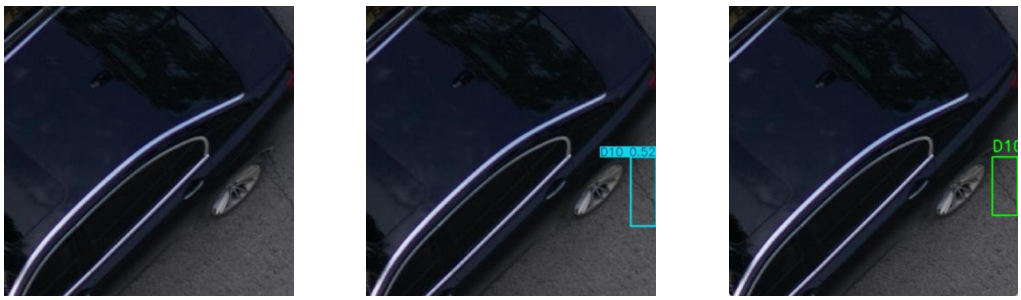
RDD2022 中来自中国的使用无人机拍摄的道路裂缝数据进行测试，相比于智能手机采集的图像数据，该数据图像包含大量行人、树木、汽车等环境干扰，总共包含 2401 张带标注道路裂缝图像，选取相同道路裂缝类别作为泛化实验中的检测任务，分别在 YOLOv8n 和 MCA-YOLO-A 模型上训练至收敛。实验结果如表 7 所示。

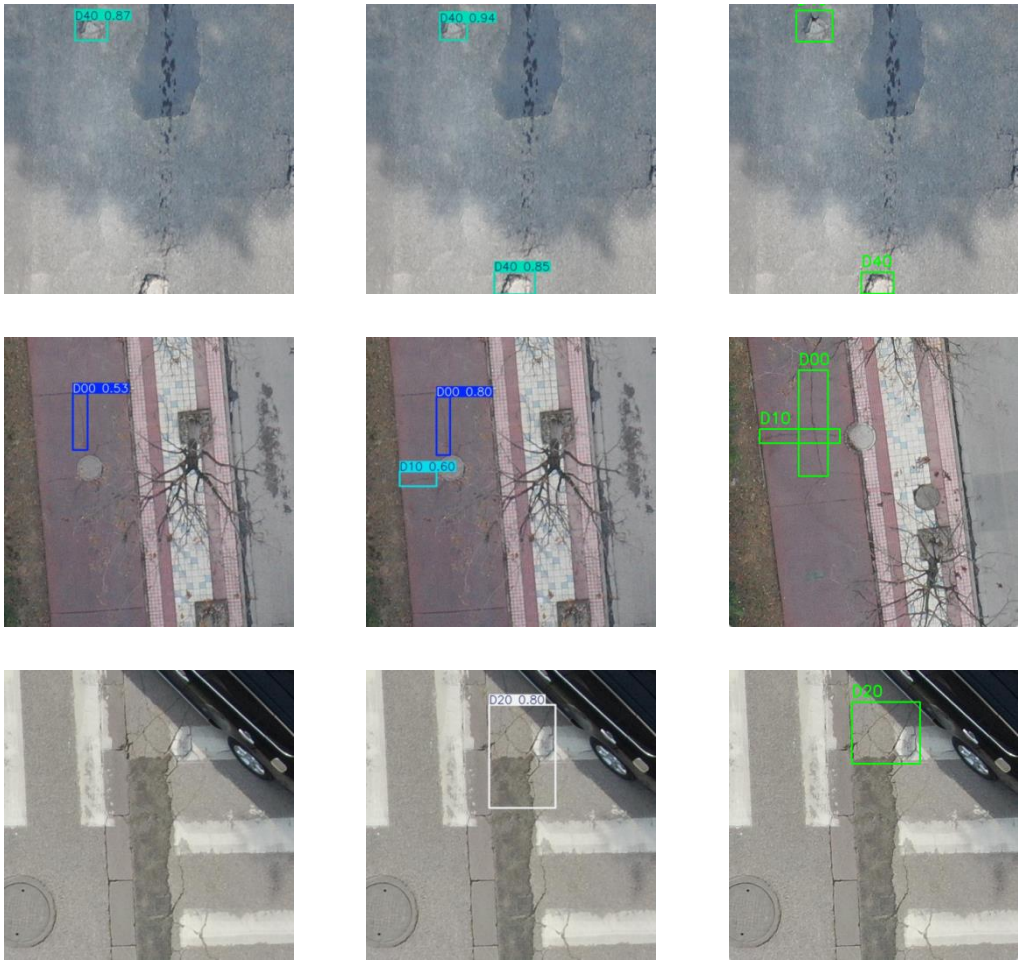
表 7 RDD2022 无人机中国数据集泛化能力测试

Table 7 RDD2022 UAV China dataset generalisation ability test

类别	YOLOv8n		MCA-YOLO-A	
	mAP_0.5	*F1* 分数	mAP_0.5	*F1* 分数
D00	0.589	0.585	0.639(+8.3%)	0.654(+11.8%)
D10	0.758	0.756	0.765(+1.0%)	0.771(+3.2%)
D20	0.369	0.378	0.536(+43.9%)	0.536(+41.8%)
D40	0.769	0.772	0.826(+7.3%)	0.838(+8.5%)
All	0.636	0.632	0.702(+10.2%)	0.701(+11.2%)

由表 7 可知，MCA-YOLO-A 在四个类别上的精度明显优于 YOLOv8n，鳄鱼裂缝提升效果最明显，相较于 YOLOv8n 模型提升 43.9 个百分点和 41.8 个百分点；在总体检测性能方面，mAP_0.5 与 *F1* 分数分别提升 10.2 和 11.2 个百分点。实验结果证明了本文所提出的模型拥有较强的泛化能力。泛化实验的部分检测结果对比如图 12 所示，由图可知，原模型对于易受环境干扰，且对于细小裂缝容易漏检，而本文所提出的 MCA-YOLO-A 模型对复杂环境下的道路裂缝以及小目标道路裂缝做出了精确检测，并且置信度更高。





（a）YOLOv8n　　（b）MCA-YOLO-A　　（c）Original label

图 12 检测结果对比

Fig.12 Comparison of detection effect

此外，为了评估 MCA-YOLO-A 模型在不同数据集上的泛化能力，选取公开数据集 UAPD[28]进一步测试。该数据集为无人机对中国南京东济大道上采集的道路裂缝图像，包含大量环境干扰，共计 2390 张多种不同类型的道路裂缝图像。实验结果见表 8。由表 8 可知，模型再次展现出同样优秀的精度与 F1 分数，再次证明了本文模型在处理更多复杂道路环境时的适应性和鲁棒性。

表 8 UAPD 数据集泛化能力测试

Table 8 UAPD dataset generalisation ability test

类别	YOLOv8n		MCA-YOLO-A	
	mAP_0.5	F1 分数	mAP_0.5	F1 分数
D00	0.834	0.823	0.873（+4.6%）	0. 839（+1.9%）

D10	0.863	0.836	0.899（+4.2%）	0. 878（+5.0%）
D20	0.862	0.852	0.887（+2.9%）	0. 876（+2.8%）
D40	0.951	0.913	0.970（+2.0%）	0. 938（+2.7%）
all	0.881	0.867	0.908（+3.1%）	0. 889（+2.5%）

4 结束语

为克服目前道路裂缝形态多样、与路面纹理区分度较低、且目前检测算法难以适用于计算资源受限的边缘设备的问题，本文基于 YOLOv8n 模型提出了轻量化检测算法 MCA-YOLO-A，并在开源数据集 RDD2022 中国图像数据上进行训练、验证和测试。首先，使用轻量型的 MobileNetv3 网络作为主干特征提取网络减少模型参数量，同时嵌入坐标注意力 CA 模块，加强对特征图裂缝区域的感兴趣程度；其次，在计算边界框回归损失时使用 Alpha-IOU 损失函数，加速网络收敛；最后，在颈部网络中增加 160×160 小目标检测层，提高模型对于细小裂缝特征的识别能力。实验结果表明，MCA-YOLO-A 模型对于背景复杂、类型多样的道路裂缝检测具有更低的参数量、更高的精度以及更快的收敛速度，并在 RDD2022 无人机拍摄的中国数据集上验证了该方法有较强泛化能力和鲁棒性。

本文工作的不足之处在于在处理具有多种环境干扰的情况，尤其是阴影遮挡和光照不足的场景时，缺乏有效的数据增强策略，导致模型在这些情况下容易出现漏检情况。因此，下一步的研究重点是：针对阴影和光照问题，我们将通过扩展数据集中的多样性来提升模型的适应能力。例如，应用基于亮度、对比度和颜色变换的技术来模拟不同光照条件，以及使用轻量级无监督学习网络来增强图像；探讨引入更具抗干扰能力的网络结构，例如使用多尺度特征融合来提升模型在不同环境下的表现；在实际环境中的部署和测试也是我们的重要任务，包括使用减枝和量化技术减少延迟。通过这些部署优化，我们希望提升模型的实用性，并使其能够适应各种实际应用场景。

参考文献

[1] ANANDHI R J, BASWARAJU S, NANDAGOPALAN S P, et al. Survey on IOT Based Pothole Detection[C]//2022 IEEE Delhi Section Conference (DELCON). New Delhi, India: IEEE, 2022: 1-6.

[2] SURESH S, PRASATH M R, CHAKARAVARTHI G. In-situ surface crack detection on metal using a passive wireless RFID-based NDT sensor[C]//2023 IEEE Microwaves, Antennas, and Propagation Conference (MAPCON). Ahmedabad, India: IEEE 2023:1-5.

[3] LI J, CHEN L, ZHANG H, et al. Research on Underwater Surface Crack Detection Technology of Concrete Dam Based on Image Processing[C]//2023 2nd International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI). Zakopane, Poland: IEEE, 2023: 99-102.

[4] 赵旭辉, 谢梦洁, 杨飚, 等. 低秩表示与深度学习结合的裂缝检测与样本生成方法[J]. 测绘学报, 2023, 52(11): 1917-1928.

ZHAO Xuhui, XIE Mengjie, YANG Biao, et al. A method for crack detection and sample generation based on low rank representation and deep learning[J]. Acta Geodaetica et Cartographica Sinica, 2023, 52(11): 1917-1928.

[5] PRASETYO A E, YUNIARTO M P, SUPROBO, et al. Application of Edge Detection Technique for Concrete Surface Crack Detection[C]//2022 International Seminar on Intelligent Technology and Its Applications (ISITIA). Surabaya, Indonesia: IEEE, 2022: 209-213.

[6] YANG F, ZHANG L, YU S J, et al. Feature pyramid and hierarchical boosting network for pavement crack detection[J]. IEEE Transactions on Intelligent Tansportation Systems, 2020, 21(4): 1525-1535.

[7] 赵凡, 李琳芸, 魏仁杰, 等. 基于通用目标检测器的大坝裂缝检测方法[J]. 数据采集与处理, 2022, 37(02): 405-414.

ZHAO Fan, LI Linyun, WEI Renjie, et al. Dam Crack Detection Method Based on Universal Target Detector[J]. Journal of Data Acquisition and Processing, 2022, 37(02): 405-414.

[8] MADASAMY K, SHANMUGANATHAN V, KANDASAMY V, et al. OSDDY: embedded system-based object surveillance detection system with small drone using deep YOLO[J]. EURASIP Journal on Image and Video Processing, 2021, 2021(1):1-14.

[9] 陈泽斌, 罗文婷, 李林. 基于改进 U-net 模型的路面裂缝智能识别[J]. 数据采集与处理, 2020, 35(02): 260-269.

CHEN Zebin, LUO Wenting, LI LIn. Automatic Identification of Pavement Crack Using Improved U-net Model[J]. Journal of Data Acquisition and Processing, 2020, 35(02): 260-269.

[10] TONG X, HUANG Y, XIAO L, et al. Surface Defect Detection Method Based on Improved Faster-RCNN[C]//2021 4th International Conference on Information Communication and Signal Processing (ICICSP). Shanghai, China: IEEE, 2021: 357-362.

[11] 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. [S.l.]: IEEE, 2016: 779-788.

[12] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 7263-7271.

[13] REDMMON J, FARHADI A. YOLOv3: An incremental improvement[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 1125-1131.

[14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[EB/OL]. (2020- 04-23)[2023-04-03]. https://arxiv.org/abs/2004.10934.

[15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Proceedings of the European Conference on Computer Vision. Amsterdam, Netherlands: Springer, 2016: 21-37.

[16] 卢宏涛, 罗沐昆. 基于深度学习的计算机视觉研究新进展[J]. 数据采集与处理, 2022, 37(02): 247-278.

LU Hongtao, LUO Mukun. Survey on New Progresses of Deep Learning Based Computer Vision[J]. Journal of Data Acquisition and Processing, 2022, 37(02): 247-278.

[17]李生辉, 李晓飞, 宋璋晗, 等. 基于改进 YOLOv5 的船舶多尺度 SAR 图像检测算法[J]. 数据采集与处理, 2024, 39(01): 120-131.

LI Shenghui, LI Xiaofei, SONG Zhanghan, et al. Multi-scale SAR Image Detection Algorithm for Ships Based on Improved YOLOv5[J]. Journal of Data Acquisition and Processing, 2024, 39(01): 120-131.

[18]LI Y, YIN C, LEI Y, et al. RDD-YOLO: road damage detection algorithm based on improved you only look once version 8[J]. Applied Sciences, 2024, 14(8): 3360.

[19]SU P, HAN H, 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.

[20]HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3[C] //Proceedings of the IEEE/CVF international conference on computer vision. Seoul, Korea (South): IEEE, 2019: 1314-1324.

[21]HOU Q , ZHOU D , FENG J .Coordinate Attention for Efficient Mobile Network Design[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. [S.l.]: IEEE, 2021: 13713-13722.

[22]HE J, ERFANI S, MA X, et al. Alpha-IoU: A Family of Power Intersection over Union Losses for Bounding Box Regression[J]. Advances in Neural Information Processing Systems, 2021, 34: 20230-20242.

[23]BOXG E P, COX D R. An analysis of transformations[J]. Journal of the Royal Statistical Society: Series B (Methodological), 1964, 26(2): 211-243.

[24]REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[25]LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. arXiv preprint arXiv: 2209. 02976, 2022.

[26]WANG C Y, YEH I H, LIAO H Y M. Yolov9: Learning what you want to learn using programmable gradient information[J]. arXiv preprint arXiv: 2402. 13616, 2024.

[27] ZHAO Y, LU W, XU S, et al. Detrs beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2024: 16965-16974.

[28] ZHU Junqing, ZHONG Jingtao, MA Tao, et al. Pavement distress detection using convolutional neural networks with images captured via UAV[J]. Automation in Construction, 2022, 133: 103991.

作者简介：

朱佳慧（2000-），女，硕士研究生，研究方向：深度学习，目标检测，E-mail: 1022072013@njupt.edu.cn。

刘艺（1999-），女，硕士研究生，研究方向：深度学习，图像增强，目标检测。

张登银（1964-），**通信作者**，男，博士，研究员，博士生导师。研究方向：信号与信息处理、信息安全。E-mail: zhangdy@njupt.edu.cn。