DOI：10.20271/j.cnki.1673−9833.2026.6005

一种可变形卷积的道路路面缺陷检测算法

刘振桢[1]，肖满生[1]，高梦兴[1]，许雅婷[1]，肖泽宇[2]

（1. 湖南工业大学 计算机与人工智能学院，湖南 株洲 412007；2. 湖南工商大学 计算机学院，湖南 长沙 410000）

摘　要：针对道路场景中缺陷尺度多变、复杂背景干扰导致检测精度低、计算量大的问题，在YOLOv8n 的基础上，提出一种可变形卷积的道路路面缺陷检测算法。首先，引入可变形卷积 DCNv4 模块构建一个 C2f_DCNv4 模块，代替原算法中的 C2f 模块，能够更加有效地处理图像中的形变和动态变化，以适应裂纹的宽范围分布和细长特性；其次，在颈部网络中设计具有层次尺度的特征金字塔 HS-FPN，以捕获更全面的裂缝特征信息；最后，采用自注意力的检测头（Detect_SA）算子对模型的头部进行改进，在保持记忆效率的同时实现特征提取。在开源道路损害数据集 RDD2022 和 UAV-PDD 2023 上的实验结果表明：所提方法的平均检测精度 *mAP*@50 为 86.3% 和 61.1%，与原算法相比提高了 6.5% 和 5.0%，并且参数量和计算量分别降低了 42.86% 和 28.40%。

关键词：深度学习；路面缺陷检测；复杂背景；可变形卷积

中图分类号：TP273　　**文献标志码**：A　　**文章编号**：1673−9833(2026)06−0031−09

引文格式：刘振桢，肖满生，高梦兴，等 . 一种可变形卷积的道路路面缺陷检测算法 [J]. 湖南工业大学学报，2026，40(6)：31−39.

A Deformable Convolution Algorithm for Road Surface Defect Detection

LIU Zhenzhen[1]，XIAO Mansheng[1]，GAO Mengxing[1]，XU Yating[1]，XIAO Zeyu[2]

（1. School of Computer Science and Artificial Intelligence，Hunan University of Technology，Zhuzhou Hunan 412007，China；
2. School of Computer，Hunan University of Technology and Business，Changsha 410000，China）

Abstract：In view of variable defect scales and complex background interference in road scenes, which lead to low detection accuracy and large computational complexity, a deformable convolution algorithm for road surface defect detection has thus been proposed based on YOLOv8n. Firstly, a deformable convolutional DCNv4 module is introduced to construct a C2f_DCNv4 module, replacing the C2f module in the original algorithm, which can more effectively handle deformation and dynamic changes in images to adapt to the wide range distribution and slender characteristics of cracks. Secondly, a hierarchical feature pyramid HS-FPN (High Level Screening Feature Fusion Pyramid) is designed in the neck network to capture more comprehensive crack feature information. Finally, the self attention detection head (Detect_SA) operator is used to improve the head of the model, thus achieving feature extraction while maintaining memory efficiency. The experimental results on the open-source road damage datasets RDD2022 and UAV-PDD 2023 demonstrate the average detection accuracy of the proposed method *mAP*@50. Compared with the original algorithm, it has increased by 6.5% and 5.0%, with 86.3% and 61.1% respectively, reducing parameter and computational complexity by 42.86% and 28.40%, respectively.

Keywords：deep learning；road defect detection；complex background；deformable convolution

收稿日期：2025−05−12

基金项目：湖南省自然科学基金资助项目（2024JJ7154）；湖南衣朵云商品智能化基金资助项目（h2024-003）

作者简介：刘振桢，男，湖南工业大学硕士生，主要研究方向为深度学习，计算机视觉，E-mail：qq131515@126.com

通信作者：肖满生，男，湖南工业大学教授，硕士生导师，主要研究方向为智能计算，智能信息处理与图像处理，E-mail：xiaomansheng@hut.edu.cn

1 研究背景

常见的道路缺陷主要包括裂缝和坑洼，这些缺陷如果不能被及时发现并修复，不仅会缩短道路的使用寿命，还会对行车安全构成严重威胁。传统的道路缺陷人工检测方法效率低下、成本高昂，且受到人为因素的影响较大，难以满足大规模、高频率的道路巡检需求。因此，发展高效、精准且自动化的道路缺陷检测技术，已成为保障交通参与者人身安全、提高道路维护效率的核心需求。

道路缺陷检测技术主要包括传统的目标检测方法和基于深度学习的目标检测方法。传统的路面缺陷检测方法通常依赖装载检测设备的巡检车辆，常用设备包括探地雷达等。通过搭载探地雷达等传感器的巡检车辆进行数据采集，探地雷达利用电磁波传播特性检测路基病害，并结合信号处理算法和图像分析技术实现缺陷识别。郭士礼等 [1] 探讨了探地雷达的应用，总结了基于特定波形特征识别裂缝宽度、深度和类型的量化方法。该类检测方法在自动识别路面问题方面具有较好的性能，但存在设备组成复杂、人力和资金投入较高、对检测环境的适应性有限等问题，因而在更大范围内的推广应用中仍面临挑战，尤其在村庄和小乡镇等地区更为明显。

神经网络架构的迭代优化推动了目标检测领域的突破性进展，目标检测领域的主流方法可分为两阶段法和一阶段法。两阶段算法体系主要涵盖 Fast R-CNN[2]、Faster R-CNN[3]、Mask R-CNN[4]、U-Net[5] 等。如肖创柏等 [6] 将 Faster R-CNN 与调整特征提取网络和数学形态学算法相结合以适应小尺寸缝隙目标检测，并在铁轨和道路裂缝数据集上进行了实验验证，取得了较高的准确率；赵志宏等 [7] 提出了一种变尺度的 U-Net 模型 VS-UNet，通过引入多尺度特征提取融合模块 Extend-LG Block，提高了道路缺陷检测的精度和分割效果。尽管这些两阶段检测算法在准确性方面通常优于单阶段算法，但它们在处理速度上的不足限制了其应用范围，相比之下，在一阶段检测算法中，YOLO（You Only Look Once）系列 [8-9] 和 SSD（Single Shot Multibox Detector）[10] 构成了主流技术路线，因其较快的检测速度，在实际道路损伤检测中更受青睐。王雪秋等 [11] 基于 YOLOv8 架构，融合可变形卷积与多路径坐标注意力机制，设计了 C2f-MPDC 模块与轻量共享检测头，实现了模型精度提升与结构轻量化的双重优化。王子龙等 [12] 在 YOLOv8n 中引入重参数化模块与注意力机制，提出轻量化的 RD-YOLO 模型，有效降低了模型复杂度并提升对微小裂缝的检测能力，但过度轻量化可能削弱对复杂背景下缺陷特征的表达能力。高泽岭等 [13] 基于 YOLOv11 结构设计裂缝方向注意力模块，加强对线性特征的感知能力，在复杂背景下提升了裂缝识别精度与稳定性。罗臻等 [14] 为解决小目标检测难题，在 YOLOv11n 中融合了 CGBlock 与 DRCGN 网络，虽有效提升了系统检测性能，但当缺陷尺寸细微、外观易于混淆时，模型仍会发生漏检。

上述方法虽然在道路缺陷检测领域取得了一定的成果，然而目前还存在一些难题，比如：道路损伤的形态多样性及尺寸差异增加了特征提取算法的设计难度；交通参与者与环境场景因素会对检测算法产生干扰，从而导致模型检测精度降低；此外，在提升模型检测精度时无法兼顾其计算成本，这也是当前面临的一个重要问题。基于此，本文在 YOLOv8n 架构基础上进行算法优化，提出一种可变形卷积的缺陷检测算法 DHS-YOLO（deformable hierarchical self-attention-YOLO）。具体包含以下几点：1）采用 DCNv4[15]（deformable convolution v4）卷积重构网络基础模块，将原始 C2f 模块与 DCNv4 进行深度整合，构建 C2f_DCNv4 模块，使网络能够自适应调整卷积核采样位置，以加强对道路缺陷特征的提取。2）为了获得更加轻量化的网络，降低计算成本，使用高效的 HS-FPN（high-level screening-feature fusion pyramid）[16] 特征融合金字塔结构，提高了模型对小尺寸道路损伤的检测能力。3）使用自注意力模块和卷积融合的自注意力检测头 Detect_SA[17]（detect_self-attention），可以捕获更广泛的上下文信息，也能保留局部信息。

2 YOLOv8n 网络结构

作为一阶段目标检测的代表算法，YOLOv8 构建了兼顾实时性的轻量化检测架构。该框架通过调整网络深度与宽度，形成 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 和 YOLOv8x 5 个版本。其中，YOLOv8n 参数量最小，更适用于计算资源受限的场景。结合工业检测对轻量化与精度平衡的需求，本文选用 YOLOv8n 作为改进基础。YOLOv8n 网络结构由 3 个主要部分组成：主干网络（Backbone）、颈部（Neck）和头部（Head）。输入模块通过多种数据增强方法提高模型的泛化能力。主干网络采用 CSPDarknet53，由 Conv、C2f 和 SPPF 模块组成，负责提取图像特征；颈部采用 FPN（feature pyramid network）和 PAN（path aggregation network）结构，

增强不同尺度的特征信息；头部利用多尺度特征图预测目标类别和边界框，并通过 NMS 筛选结果。

3　基于可变形卷积网络的改进算法

基于可变形卷积网络的改进算法 DHS-YOLO 结构见图 1，DHS-YOLO 主干部分使用可变形卷积 DCNv4 和 C2f 相结合的 C2f_DCNv4 模块，DCNv4 通过增加偏移量，更灵活地适应各种裂缝所在位置，增强了特征提取效率；在颈部，使用了一种名为 HS-FPN 的特征融合金字塔结构，这种结构能处理多尺度的特征图，有助于识别道路缺陷的细微特征，而且有效减少了冗余参数的数量；最后，在头部使用了自注意力机制的检测头（Detect_SA），由高效的上下文自注意力模块和卷积结合，对提高模型性能和效率有重要作用。

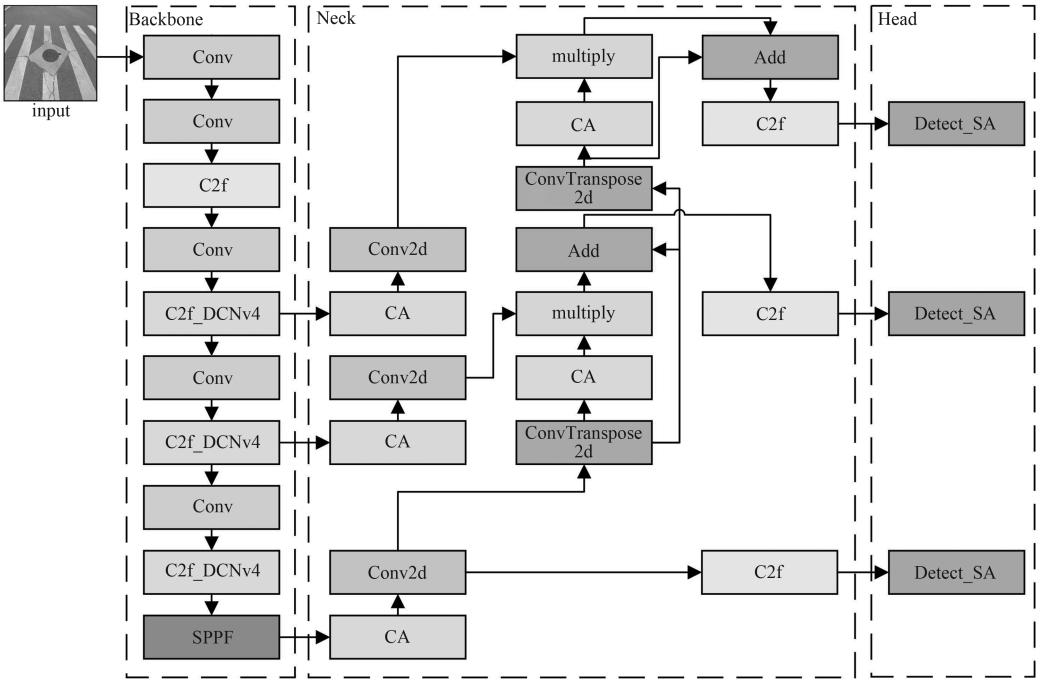


图 1　DHS-YOLO 网络结构

Fig. 1　DHS-YOLO network architecture

3.1　特征提取模块 C2f_DCNv4

针对网状裂缝这一检测难点，经典卷积检测框架受固定感受野限制，在路面缺陷边缘特征捕获方面存在不足，从而影响检测精度。为了解决这一问题，本研究引入了 DCNv4 模块，以提升对道路缺陷特征的提取能力。

给定输入特征图，常规卷积和可变形卷积的公式分别如式（1）（2）所示：

$$y(P_0)=\sum_{P_n\in\Omega}\omega(P_n)\cdot x(P_0+P_n)\,, \tag{1}$$

$$y(P_0)=\sum_{P_n\in\Omega}\omega(P_n)\cdot x(P_0+P_n+\Delta P_n)\,。 \tag{2}$$

式（1）（2）中：$x\in\mathbf{R}^{C_{\text{in}}\times H\times W}$ 为输入特征图，其中 C_{in}、H 和 W 分别为输入特征图的通道数、高度和宽度；P_0 为输入特征图中的每个位置；P_n 为卷积核采样区域中的第 n 个采样点；Ω 为常规卷积的采样网格；$\omega(P_n)$ 为位置 P_n 处对应的卷积权重；$y(P_0)$ 为位置 P_0 处的输出特征值；ΔP_n 为可学习的偏移量，用于调整卷积核的实际采样位置。

可变形卷积中使用偏移量 ΔP_n，突破传统固定网格采样方法。基于裂缝区域的形变特征，算法可自主优化特征采样点的空间分布模式，能够更加贴合地捕捉网状裂缝的边缘特征，与常规卷积相比表现出更强的鲁棒性，同时有效减少模型参数量和计算复杂度。其中，偏移量 ΔP_n 是通过一个并行的卷积层根据输入特征学习得到的，它允许采样点在常规网格位置 P_n 的基础上进行浮点数级别的移动，从而使感受野的形状能够动态适应目标的几何形态。DCNv4 模块针对输入的特征 $x_g\in\mathbf{R}^{H\times W\times C}$，其具有高度 H、宽度 W 和通道 C 时，公式描述如式（3）（4）：

$$y_g=\sum_{k=1}^{K}m_{gk}x_g\left(p_0+p_k+\Delta p_{gk}\right), \tag{3}$$

$$y=\text{Concat}\left(\left[y_1,y_2,\cdots,y_g\right],axis=-1\right)。 \tag{4}$$

式（3）（4）中：g 为沿通道维度进行空间分组的组数；x_g 和 y_g 分别为第 g 组的特征图，对于第 g 组，特征通道的维数由 $c^*=c/g$ 表示，其中 c 为 DCNv4 模块的完整特征图总通道数；m_{gk} 为第 g 组中第 k 个采样点的空间聚合权重；p_k 为网格采样的第 k 个位置，类似于标准卷积 {(−1, −1), (−1, 0), ⋯, (0, +1), ⋯, (+1,

+1)}；Δp_{gk} 为第 g 组中网格采样位置 p_k 的偏移量；Concat 为增加图像特征描述数量（即通道数）操作。

如图 2 所示，DCNv4 模块对特征图按通道进行分组，并为每个分组计算各位置的偏移量和权重，用偏移量和权重执行卷积运算，最后将所有分组的卷积结果合并以生成最终输出。DCNv4 模块能够根据不同道路病害的形状和大小自适应调整其采样位置与形状，从而更精准地提取特征。

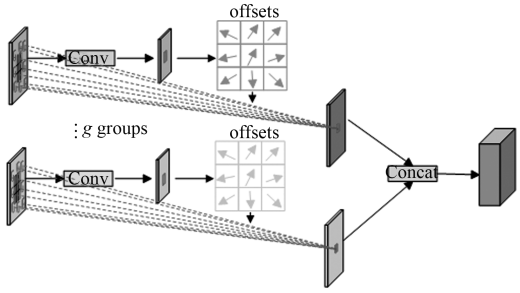


图 2 DCNv4 模块

Fig. 2 DCNv4 module

如图 3 所示，为增强对网状裂缝和坑洼特征的提取能力，采用 DCNv4 模块对 YOLOv8n 的 C2f 模块进行深度重构，构建具有形变感知能力的 C2f_DCNv4 模块。相较于标准卷积，可变形卷积的动态采样机制可自适应调整感受野范围，从而提升对不规则裂缝和坑洼的几何特征敏感度。在 C2f_DCNv4 模块里，首先采用 1×1 卷积实现输入特征通道降维，降低通道冗余显著减少参数规模。随后，通过 Split 操作分割多尺度特征，并引入跨层跳跃连接结构，该设计不仅增强了梯度反向传播效率，还通过参数共享策略进一步压缩模块参数量。DCNv4 的优化提升了特征提取的速度和效率，同时在降低计算复杂度的前提下，保持甚至增强了特征提取的质量。

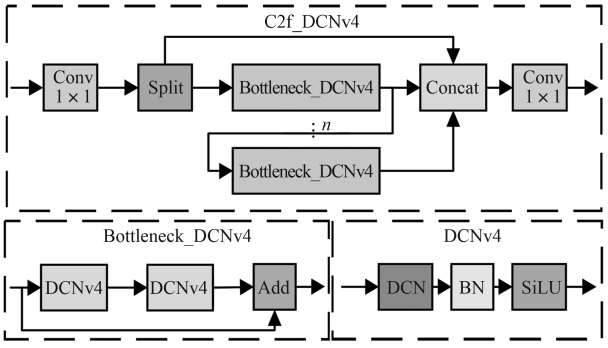


图 3 C2f_DCNv4 结构

Fig. 3 C2f_DCNv4 structure

3.2 HS-FPN 模块

YOLOv8n 的颈部主要采用特征金字塔网络（FPN[18]）和路径聚合网络（PAN[19]）进行特征融合。FPN 通过自上而下的单向路径融合特征，虽能增强语义信息，但易忽略定位信息；PAN 则在此基础上增加自下而上的路径，以补充定位特征。尽管该结构能够实现语义与定位信息的协同融合，但仍存在一定局限，如对浅层细节特征利用不足，影响微小目标的识别效果。

为了解决上述问题，本文参考 HS-FPN（high-level screening-feature fusion pyramid）的特征融合金字塔结构，重新设计一种面向 YOLOv8 的改进 HS-FPN 特征融合架构，整体结构如图 4 所示。该改进 HS-FPN 特征融合架构主要由特征选择模块与特征融合模块两部分组成。

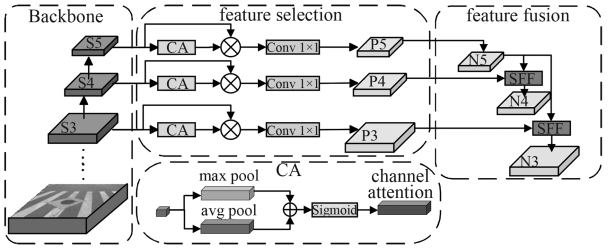


图 4 面向 YOLOv8 的改进 HS-FPN 特征融合架构

Fig. 4 Improved HS-FPN feature fusion architecture for YOLOv8

特征选择模块采用了通道注意力（channel attention，CA）和维度匹配（dimensional matching，DM）机制，以更有效地提取特征的有用信息。CA 模块对输入特征图 $f_{\text{in}} \in \mathbf{R}^{C\times H\times W}$ 进行处理。特征图依次通过全局最大池化（max pool）和全局平均池化（avg pool）两个池化层，然后将得到的特征进行融合。接着，利用 Sigmoid 激活函数计算每个通道的权重值，生成各通道的权重 $f_{\text{c}} \in \mathbf{R}^{C\times 1\times 1}$。在进行特征融合之前，它们的通道数各不相同，DM 模块使用 1×1 卷积来减少通道数，使其统一为 256。

在特征融合模块阶段，HS-FPN 采用选择性特征融合（selective feature fusion，SFF）机制，其原理如图 5 所示。

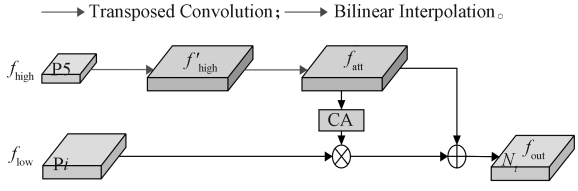


图 5 SFF 原理示意图

Fig. 5 SFF principle schematic diagram

为实现多层次特征有效融合，针对高级的特征输入 $f_{\text{high}} \in \mathbf{R}^{C\times H\times W}$，通过步长为 2 的 3×3 转置卷积（Transposed Convolution）进行上采样，将其空间维度扩展至 $f'_{\text{high}} \in \mathbf{R}^{C\times 2H\times 2W}$。为实现特征维度兼容，进一步引入双线性插值算法（Bilinear Interpolation）对 f_{high} 执行自适应缩放，输出 $f_{\text{att}} \in \mathbf{R}^{C\times H_1\times W_1}$，经 CA 模块生成的注意力权重对 $f_{\text{low}} \in \mathbf{R}^{C\times H_1\times W_1}$ 实施特征优

选，最终通过残差连接融合语义与细节特征，输出 $f_{\text{out}} \in \mathbf{R}^{C\times H_1\times W_1}$。式（5）和（6）具体展示了特征选择的融合过程。

$$f_{\text{att}} = BL\left(T - \text{Conv}\left(f_{\text{high}}\right)\right), \tag{5}$$

$$f_{\text{out}} = f_{\text{low}} \times \text{CA}\left(f_{\text{att}}\right) + f_{\text{att}}。 \tag{6}$$

3.3　自注意力检测头

在道路缺陷检测中，受裂缝宽度差异影响，细小裂缝易被忽视。为增强检测头间的信息交互并提升检测性能，课题组提出了一种融合自注意力机制的检测头（Detect_SA）。传统检测头通常独立处理各自特征图，容易导致信息割裂，从而影响检测效果。

如图 6 所示，自注意力检测头包括自注意力模块（self-attention module），自注意力模块旨在在保持记忆效率的前提下优化特征提取过程。自注意力模块由多个关键组件构成，旨在有效捕获特征间的长距离依赖关系。首先，通过一个 1 × 1 的卷积层将特征维度降低，以减少后续计算负荷。紧接着是多头自注意力层（MHSA），它能够有效捕获特征间的复杂关系。随后，另一个 1 × 1 的卷积层将特征维度恢复至原本的特征维度。为了保证训练的稳定性，引入了批量归一化层（BN）。最后，将自注意力模块的输出与初始输入相加，并应用 ReLU 激活函数，以增强模型的非线性表达能力。

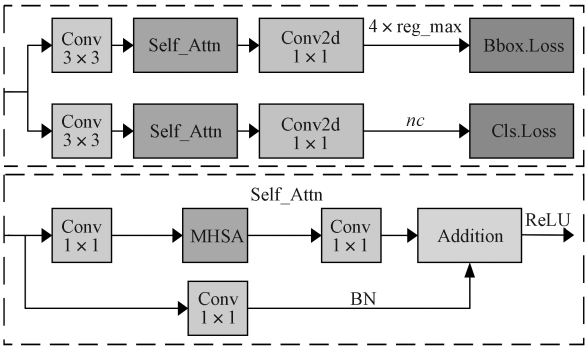


图 6　自注意力检测头结构

Fig. 6　Self-attention detection head structure

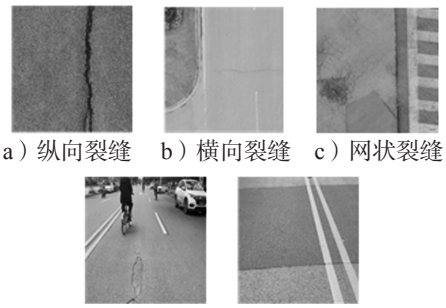
4　实验分析

4.1　实验环境

为客观评估算法性能，搭建了标准实验环境。本文基于高性能计算平台开展实验验证，硬件系统搭载 NVIDIA RTX 3080Ti 显卡（显存容量为 12 GB）及 Intel Xeon Silver 4214R 处理器（基础频率为 2.40 GHz）。软件环境部署 Ubuntu 22.04 作为操作系统，采用 Python 3.8 解释器与 CUDA 12.1 加速库，依托 PyTorch 2.1.0 框架构建深度学习模型。主要的训练超参数配置如下：输入层接收 640 × 640 像素的标准化图像，通过随机梯度下降算法（stochastic gradient descent，SGD）进行参数优化，设置总迭代周期为 200 epoch，单次训练批量样本数为 32，数据加载线程数（workers）为 8，其余参数保持框架默认配置。

4.2　数据集介绍与处理及评价指标选取

实验数据源自国际道路损伤检测挑战赛数据集 RDD2022[20]，该数据集由多国协同构建，覆盖了中国、日本、美国、印度和捷克等国家的道路表面数据，针对国内智能巡检系统对本土化特征适配性的需求，筛选出中国道路图像子集作为基准数据集。该数据图像涵盖了 5 种道路缺陷类别：D00（纵向裂缝）、D10（横向裂缝）、D20（网状裂缝）、D40（坑洼）和 Repair（道路修补），如图 7 所示，每类缺陷均提供高分辨率实例图像，总共 4 378 张道路缺陷数据图像。经过数据清洗之后，数据集共有 3 853 张包含数据带标注的图像。将数据集随机划分为训练集、验证集、测试集，比例为 8:1:1。



a）纵向裂缝　b）横向裂缝　c）网状裂缝

d）坑洼　e）道路修补

图 7　道路缺陷类别示例图

Fig. 7　Road defect category illustration

为系统评估算法性能，采用的评估指标如下：所有类别的平均精度（mean average precision，mAP）和 F1 分数（F1-score），其数值提升表明模型的判别性能增强；精确度（Precision）代表预测为正样本中正确的比例，反映正类预测的可靠性；召回率（Recall）代表真实为正样本中被正确预测为正的比例；每秒检测帧数（frames per second，FPS）为每秒能处理的图像数量，对于实现实时检测非常重要；每秒十亿次浮点运算数（giga floating point operations per second，GFLOPs）和模型参数量（Params/M）代表模型的复杂度，其值越小则模型轻量化程度越高。

4.3　模块化消融实验

为了研究改进模块对网络性能的影响，本文以 YOLOv8n 作为基础模型设计了模块化消融实验，以验证每种改进策略对模型的具体影响。D-YOLO 表示对主干网络实施可变形卷积升级，将 C2f 模块置换为 C2f-DCNv4 结构；H-YOLO 表示采用特征融合金

字塔结构 HS-FPN 重构颈部网络；将检测头升级为自注意力检测头的模型称为 S-YOLO，而 DHS-YOLO 为集成上述 3 项优化策略，专为道路缺陷数据特性设计的算法。

消融实验结果如表 1 所示，该结果展示了各改进模块的性能效果。由表中数据可得：通过引入 C2f-DCNv4，裂缝形变特征的提取能力显著增强，*mAP*@50 提升了 5.1%，即参数量约减少 3.3% 的同时保持 93 帧 /s 的推理速度；将重构的 HS-FPN 特征金字塔通过分层筛选机制融合多尺度语义信息，在参数量压缩 34.6%、计算量降低 14.8% 的轻量化优势下，仍实现 3.2% 的精度提升；通过全新设计的检测头 Detect_SA，在保持记忆效率的前提下优化特征提取过程，*mAP*@50 提升了 4.8%。多模块联合部署时，可变形卷积模块 DCNv4 和 HS-FPN 协同作用使模型参数量减少 37.9%，检测精度较基准模型提升 4.9%；最终集成的 DHS-YOLO 算法融合 3 项核心技术，在 YOLOv8n 基础上平均精度可达 86.3%，参数量与计算量分别减少 42.86% 和 28.40%，推理速度达 79 帧 /s。可见所提算法在保持高召回率的前提下，解决了路面缺陷检测中精度与实时性的平衡难题，为嵌入式设备部署提供了高效解决方案。

表 1 消融实验结果对比

Table 1 Comparison of ablation experiment results

算法	C2f-DCNv4	HS-FPN	Detect_SA	*mAP*@50/%	*Precision*/%	*Recall*/%	*GFLOPs*	*Params*/10^6	*FPS*/(帧·s^{-1})
YOLOv8n				79.8	84.2	70.5	8.1	3.01	97
D-YOLO	√			84.9	83.8	79.8	8.1	2.91	93
H-YOLO		√		83.0	82.8	76.8	6.9	1.97	88
S-YOLO			√	84.6	81.7	79.2	7.1	2.87	81
DH-YOLO	√	√		84.7	85.7	75.2	6.9	1.87	85
DS-YOLO	√		√	85.9	83.2	80.8	7.0	2.77	84
HS-YOLO		√	√	85.8	88.6	74.7	5.9	1.83	81
DHS-YOLO	√	√	√	86.3	85.9	77.7	5.8	1.72	79

4.4 与不同算法的检测结果比较

通过对比实验验证 DHS-YOLO 算法的性能优势，结果如表 2 所示。

表 2 不同算法对比实验结果

Table 2 Comparative experiment results of different algorithms

算法	*mAP*@50/%	*mAP*@50-95/%	*Params*/10^6	*Precision*/%	*Recall*/%	*FPS*/(帧·s^{-1})	$F_{1\text{-score}}$/%
Faster R-CNN	73.2	42.5	28.33	47.8	79.8	11	60
SSD	72.0	41.8	26.19	87.9	49.2	70	61
YOLOv5s	81.0	50.5	9.12	87.0	73.6	75	76
YOLOv6s	80.7	50.2	16.3	85.9	75.4	92	79
YOLOv7-tiny	78.6	48.4	6.01	84.3	71.1	90	80
YOLOv8n	79.8	51.0	3.01	84.2	70.5	97	81
YOLOv9t	82.7	52.3	1.97	82.5	74.8	58	77
YOLOv10n	83.0	54.4	2.27	83.6	76.7	110	74
YOLOv12n	82.6	53.1	2.51	74.9	82.6	62	80
RT-DETR-L	79.9	51.6	29.27	74.2	77.6	26	77
RT-DETR-ResNet50	83.9	55.6	41.94	90.8	74.3	28	81
DHS-YOLO	86.3	57.7	1.72	85.9	77.7	79	82

本研究中，选取了当前主流的目标检测算法作为对比基准。基准实验模型包括 Faster R-CNN[3]、SSD[10]、YOLOv5s[21]、YOLOv6s[22]、YOLOv7-tiny[23]、YOLOv8n[12]、YOLOv9t[24]、YOLOv10n[25]、YOLOv12n[26]，以及基于 Transformer 模型的 RT-DETR-L[27] 和 RT-DETR-ResNet50。所有对比实验均采用统一的数据集和训练策略，确保模型达到充分收敛状态，从而尽可能减小由训练设置差异带来的结果偏差，保证实验结果具有较高的公平性、客观性与可信度。

从表 2 可以看出，本文提出的 DHS-YOLO 算法在检测精度、模型轻量化与实时性方面均表现出较强的综合优势。在精度指标方面，DHS-YOLO 的 *mAP*@50 和 *mAP*@50-95 均达到全体对比模型中的

最优水平，较现有对比模型最高值分别提升 2.4% 和 2.1%，表明该方法在不同 IoU 阈值下均具有更优的目标定位与类别判别能力。同时，算法的 $F_{1\text{-score}}$ 达到 82%，在 Precision 与 Recall 之间取得了较为均衡的表现，说明所提出方法不仅能够有效提升检测准确率，还能在一定程度上减少漏检与误检现象，体现出较好的综合检测性能。在轻量化与实时性方面，DHS-YOLO 同样表现突出，其参数量仅为 1.72×10^6，较参数量次优的 YOLOv9t 进一步降低 12.69%，体现出较高的参数利用效率；与此同时，模型推理速度达到 79 帧 /s，满足实时检测需求。综合来看，DHS-YOLO 在精度、复杂度和推理速度之间取得了较好的平衡，验证了所提出改进策略在道路缺陷检测任务中的有效性与实用价值，也说明该方法具有较好的工程应用潜力。

4.5　RDD2022 数据集实验结果

为量化评估模型对不同损伤形态的适应性，表 3 对比了基准模型 YOLOv8n、YOLOv12n 与改进算法 DHS-YOLO 在 RDD2022 数据集不同损坏类别下的性能。实验结果表明，尽管 YOLOv12n 较 YOLOv8n 在总体精度上有所提升，但本文提出的 DHS-YOLO 在 5 类典型道路缺陷检测中均表现出更全面的精度优势。相比基准模型 YOLOv8n，其在类别 D10（横向裂缝）上的提升最为显著，*mAP*@50 和 *mAP*@50-95 分别提高了 7.7% 和 8.0%。

表 3　RDD2022 数据集各类型损伤的检测结果对比

Table 3　Comparison of detection results for various types of damage in RDD2022 dataset

算法	类别	*mAP*@50/%	*mAP*@50-95/%
YOLOv8n	D00	81.6	49.2
	D10	79.3	43.6
	D20	74.6	50.1
	D40	80.6	46.9
	Repair	82.9	65.3
YOLOv12n	D00	82.1	51.2
	D10	86.0	52.8
	D20	73.6	42.6
	D40	85.0	51.0
	Repair	86.2	67.9
DHS-YOLO	D00	89.0	56.2
	D10	87.0	51.6
	D20	81.5	54.0
	D40	85.3	53.2
	Repair	88.9	73.6

图 8 展示了 DHS-YOLO 算法在 RDD2022 路面缺陷检测数据集上的可视化结果。图 8 中可看出，在组别 1 中，受车辆遮挡影响，YOLOv8n 和 YOLOv12n 虽能检测出缺陷目标，但整体置信度偏低，RT-DETR-L 则出现多个目标框，存在定位不准问题；相比之下，DHS-YOLO 能够更准确地识别道路裂缝目标。在组别 2 中，图像存在树枝干扰且缺陷目标较小，YOLOv8n、YOLOv12n 和 RT-DETR-L 均出现误检，而 DHS-YOLO 能够有效抑制背景干扰，检测结果更稳定。在组别 3 中，针对明显的光照不均与明暗变化，DHS-YOLO 表现出更强的环境适应能力。在组别 4 中，面对路牌、路灯等复杂背景干扰，YOLOv8n 和 YOLOv12n 均出现 D10 类缺陷漏检，RT-DETR-L 虽提高了检测置信度，但仍存在局部误检。综合来看，DHS-YOLO 在复杂道路场景下表现出更强的鲁棒性和更高的检测准确率。

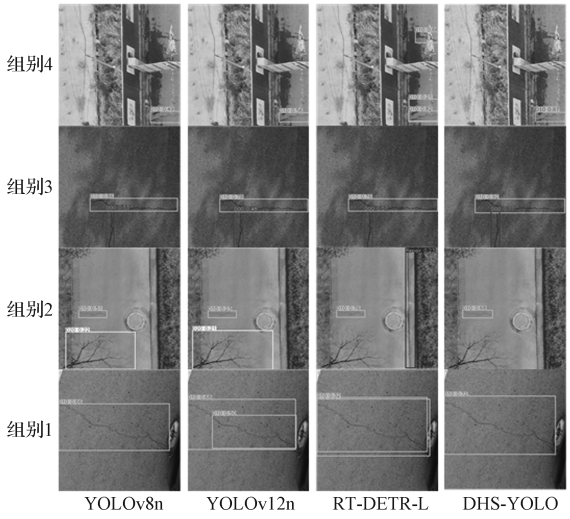


图 8　RDD 数据集检测效果对比图

Fig. 8　Detection performance comparison on the RDD dataset

4.6　实际检测能力实验

为评估本研究的泛化性能，选取 UAV-PDD 2023[28] 公开数据集展开验证。该数据集基于无人机航拍技术构建，在天津 30 m 高度拍摄，并以 PASCAL VOC 格式进行标注。数据集中共包含 2 440 张图像和 11 158 个实例，训练集、验证集和测试集按 8∶1∶1 划分。数据图像包含较多环境干扰，采集场景覆盖晴朗天气及降雨后 1 h 内的道路状况，道路类型包括高速公路、省道等。该数据集包含 6 类典型裂缝，分别为纵向裂缝（longitudinal crack）、横向裂缝（transverse crack）、网状裂缝（alligator crack）、斜裂缝（oblique crack）、修补（repair）和坑洞（pothole）。由于 UAV-PDD 2023 较 RDD2022 多出斜裂缝（oblique crack）类别，为消除类别偏差，本文选取两数据集共有的 5 类道路损伤，即 D00、D10、D20、D40 和 Repair。实验结果表明，在该数据集上原始 YOLOv8n 的 *mAP*@50 和 *mAP*@50-95

分别为 56.1% 和 26.1%，而 DHS-YOLO 分别达到了 61.1% 和 29.3%，较原算法分别提升了 5.0% 和 3.2%，验证了改进模型在复杂真实场景下的泛化能力。

为了更直观地展示算法改进前后的关注差异，本文采用 Gradient-weighted Class Activation Mapping++（Grad-CAM++）方法对 YOLOv8n 与 DHS-YOLO 模型的输出特征进行可视化处理，热力图结果如图 9 所示，颜色越接近深红色表示模型越关注该区域。从图中可以观察到，YOLOv8n 算法在处理具有背景干扰（如车辆、植被）或复杂纹理的图像时，其注意力区域分布较为分散，存在较多非缺陷区域的高响应，容易造成误检或漏检。而本文提出的 DHS-YOLO 算法则展现出更强的目标聚焦能力，其高响应区域更集中于实际缺陷位置，有效抑制了背景与车辆等干扰因素，说明改进后的模型能够更准确地提取与缺陷相关的特征信息，从而提升了检测性能与鲁棒性。

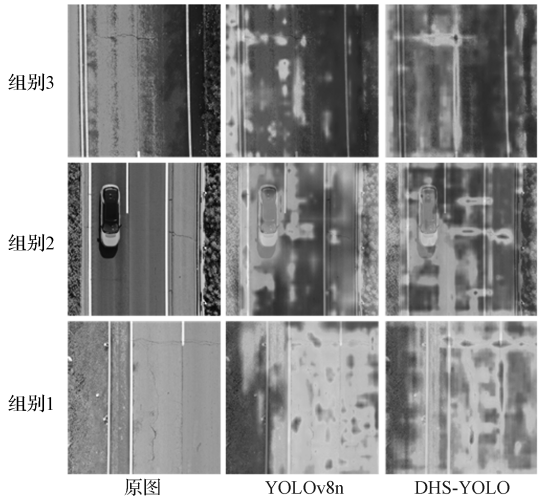


图 9 UAV-PDD 数据集实际检测结果对比图

Fig. 9 Comparison of actual detection results on the UAV-PDD dataset

5 结语

本文提出了一种可变形卷积的道路路面缺陷检测算法 DHS-YOLO，通过深度改进 YOLOv8n 基准架构实现性能提升，具体通过引入 DCNv4 模块、HS-FPN 结构以及 Detect_SA 自注意力检测头，有效提升了模型对目标位置的聚焦能力、对小目标的检测能力以及在复杂场景下的表现。实验结果表明，该算法在 RDD2022 数据集上实现了较高的检测精度和运算效率。与 YOLOv5s、YOLOv6s、YOLOv7-tiny、YOLOv8n、YOLOv9t、YOLOv10n、YOLOv12 以及 RT-DETR 等先进算法相比，本研究提出的改进 YOLOv8n 算法在检测精度上均有显著提升，尤其在复杂背景和多样化缺陷场景下，表现出更强的鲁棒性。尽管如此，相较于基准 YOLOv8n 模型，当前算法存在推理速度变慢导致 FPS 有所下降。因此，后续研究将重点探索开发基于图像预处理的动态计算简化机制或设计轻量化网络改进模块，通过平衡模型精度与推理效率，着力压缩网络参数量级，最终实现算法在嵌入式平台的轻量级部署。

参考文献：

[1] 郭士礼，梁 栋，张鸿彦，等 . 道路裂缝属性特征探地雷达检测方法综述 [J]. 地球物理学进展，2025，40(5)：2172−2186.
GUO Shili，LIANG Dong，ZHANG Hongyan，et al. Review of Ground Penetrating Radar Detection Technology for Attribute Characteristics of Cracks in Road Structural Layers[J]. Progress in Geophysics，2025，40(5)：2172−2186.

[2] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago，Chile：IEEE，2015：1440−1448.

[3] 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.

[4] HE K M，GKIOXARI G，DOLLÁR P，et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2020，42(2)：386−397.

[5] RONNEBERGER O，FISCHER P，BROX T. U-Net：Convolutional Networks for Biomedical Image Segmentation[C]//Medical Image Computing and Computer-Assisted Intervention 2015. Cham：Springer，2015：234−241.

[6] 肖创柏，柏鳗晏，禹 晶 . 基于 Faster R-CNN 的缝隙检测与提取算法 [J]. 北京工业大学学报，2021，47(2)：135−146.
XIAO Chuangbai，BAI Manyan，YU Jing. Crack Detection and Extraction Based on Faster R-CNN[J]. Journal of Beijing University of Technology，2021，47(2)：135−146.

[7] 赵志宏，何 朋，郝子晔 . 一种道路裂缝检测的变尺度 VS-UNet 模型 [J]. 湖南大学学报 (自然科学版)，2024，51(6)：63−72.
ZHAO Zhihong，HE Peng，HAO Ziye. A Variable-Scale VS-UNet Model for Road Crack Detection[J]. Journal of Hunan University (Natural Sciences)，2024，51(6)：63−72.

[8] REDMON J，FARHADI A. YOLOv3：An Incremental Improvement[EB/OL]. (2018−04−18). 2018：arXiv：

1804.02767. https://arxiv.org/abs/1804.02767.

[9] BOCHKOVSKIY A，WANG C Y，LIAO H M. YOLOv4：Optimal Speed and Accuracy of Object Detection[EB/OL]. (2020−04−23). 2020：arXiv：2004.10934. https://arxiv.org/abs/2004.10934.

[10] LIU W，ANGUELOV D，ERHAN D，et al. SSD：Single Shot MultiBox Detector[C]//Computer Vision – ECCV 2016. Cham：Springer，2016：21−37.

[11] 王雪秋，高焕兵，郑泽萌 . 改进 YOLOv8 的道路缺陷检测算法 [J]. 计算机工程与应用，2024，60(17)：179−190.
WANG Xueqiu，GAO Huanbing，JIA Zemeng. Improved Road Defect Detection Algorithm Based on YOLOv8[J]. Computer Engineering and Applications，2024，60(17)：179−190.

[12] 王子龙，陈慈发，董方敏 . 基于改进 YOLOv8n 的道路裂缝检测算法 [J]. 现代电子技术，2025，48(21)：113−119.
WANG Zilong，CHEN Cifa，DONG Fangmin. Road Crack Detection Algorithm Based on Improved YOLOv8n[J]. Modern Electronic Technique，2025，48(21)：113−119.

[13] 高泽岭 . 基于改进 YOLOv11 的道路裂缝检测系统研究 [J]. 信息与电脑，2025，37(6)：136−138.
GAO Zeling. Research on Road Crack Detection System Based on Improved YOLOv11[J]. Information & Computer，2025，37(6)：136−138.

[14] 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.
LUO Zhen，JIANG Yu，LI Weicheng. Road Defect Detection Model Based on Improved YOLO11n[J]. Microelectronics & Computer，2025，42(11)：25−36.

[15] XIONG Y W，LI Z Q，CHEN Y T，et al. Efficient Deformable ConvNets：Rethinking Dynamic and Sparse Operator for Vision Applications[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle：IEEE，2024：5652−5661.

[16] CHEN Y F，ZHANG C Y，CHEN B，et al. Accurate Leukocyte Detection Based on Deformable-DETR and Multi-Level Feature Fusion for Aiding Diagnosis of Blood Diseases[J]. Computers in Biology and Medicine，2024，170：107917.

[17] YU H Y，WAN C，DAI X Y，et al. HyCTAS：Multi-Objective Hybrid Convolution-Transformer Architecture Search for Real-Time Image Segmentation[EB/OL]. (2024−03−15). 2024：arXiv：2403.10413. https://arxiv.org/abs/2403.10413.

[18] LIN T Y，DOLLÁR P，GIRSHICK R，et al. Feature Pyramid Networks for Object Detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu：IEEE，2017：936−944.

[19] LIU S，QI L，QIN H F，et al. Path Aggregation Network for Instance Segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City：IEEE，2018：8759−8768.

[20] ARYA D，MAEDA H，GHOSH S K，et al. RDD2022：A Multi-National Image Dataset for Automatic Road Damage Detection[EB/OL]. (2022−11−18). 2022：arXiv：2209.08538. https://arxiv.org/abs/2209.08538.

[21] 卫　策，吕　进，曲晨阳 . 改进 YOLOv5s 的复杂交通场景下目标检测算法 [J]. 电子测量技术，2024，47(2)：121−130.
WEI Ce，LÜ Jin，QU Chenyang. Improved Object Detection Algorithm for Complex Traffic Scenes in YOLOv5s[J]. Electronic Measurement Technology，2024，47(2)：121−130.

[22] LI C Y，LI L L，JIANG H L，et al. YOLOv6：A Single-Stage Object Detection Framework for Industrial Applications[EB/OL]. (2022−11−07). 2022：arXiv：2209.02976. https://arxiv.org/abs/2209.02976.

[23] WANG C Y，BOCHKOVSKIY A，LIAO H M. YOLOv7：Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors[EB/OL]. (2022−07−06). 2022：arXiv：2207.02696. https://arxiv.org/abs/2207.02696.

[24] WANG C Y，YEH I H，MARK LIAO H Y. YOLOv9：Learning What You Want to Learn Using Programmable Gradient Information[C]//Computer Vision – ECCV 2024. Cham：Springer，2025：1−21.

[25] CHEN H，CHEN K，DING G G，et al. YOLOv10：Real-Time End-to-End Object Detection[C]//Advances in Neural Information Processing Systems 37. Vancouver：NeurIPS，2024：107984−108011.

[26] TIAN Y J，YE Q X，DOERMANN D. YOLOv12：Attention-Centric Real-Time Object Detectors[EB/OL]. (2025−02−18). 2025：arXiv：2502.12524. https://arxiv.org/abs/2502. 12524.

[27] ZHAO Y A，LÜ W Y，XU S L，et al. DETRs Beat YOLOs on Real-Time Object Detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle：IEEE，2024：16965−16974.

[28] YAN H H，ZHANG J F. UAV-PDD2023：A Benchmark Dataset for Pavement Distress Detection Based on UAV Images[J]. Data in Brief，2023，51：109692.

（责任编辑：申　剑）