

基于频率增强和形类协同的道路病害检测^①



李志华¹, 李佳良^{1,2}, 王臻¹, 许灿辉², 史操²

¹(浪潮企业云科技(山东)有限公司, 济南 250101)

²(青岛科技大学 信息科学技术学院, 青岛 266061)

通信作者: 李志华, E-mail: 476021201@qq.com

摘要: 道路病害检测作为衡量路面损坏状况、维护道路养护的重要方式, 存在着病害极端长宽比、大小尺度不一以及难易病害数量分布不均匀等问题. 当前基于卷积的各种方法虽然获得了更大的感受野以增强感知, 但是牺牲了蕴含细小病害的高频分量, 不适合有细小病害的道路病害检测任务. 为此, 本文提出了一种基于频率增强和形类协同的道路病害检测算法 FS-YOLO. 首先, 为实现感受野与高频信息的平衡, 我们引入了基于频率的增强卷积策略, 根据局部频率分量动态调整空间膨胀率, 为不同尺寸的病害分配合适的卷积参数. 其次, 面对不同类别病害几何形状位置各有特点, 我们引入了基于注意力机制的三维显式协同的动态检测头来实现空间几何信息与类别信息的显式协同, 使模型能够发挥病害类别与空间位置的内在潜能. 最后, 我们引入 Slide loss 损失函数来解决实际道路中难识别与容易识别的病害分布不平衡问题, 特别提升模型对难区分样本的处理能力. 实验结果表明, FS-YOLO 在自建数据集和公开道路病害检测数据集 RDD 2022、UAV-PDD 上的精确率和召回率均显著优于基线模型, 且在高速公路和国省道的实际应用中也得到了有效验证, 显著提高了病害检测的准确率和效率.

关键词: 道路病害检测; 深度学习; 机器视觉; YOLO; 膨胀卷积; 显式协同; 注意力机制

引用格式: 李志华, 李佳良, 王臻, 许灿辉, 史操. 基于频率增强和形类协同的道路病害检测. 计算机系统应用, 2026, 35(1): 64-75. <http://www.c-s-a.org.cn/1003-3254/10080.html>

Road Defect Detection Based on Frequency Enhancement and Synergy of Geometric Shape and Category

LI Zhi-Hua¹, LI Jia-Liang^{1,2}, WANG Zhen¹, XU Can-Hui², SHI Cao²

¹(Inspur Enterprise Cloud Technology (Shandong) Co. Ltd., Jinan 250101, China)

²(School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: Road defect detection, as an important method for measuring pavement damage and maintaining road maintenance, faces challenges, including extreme length-to-width ratios, varying defect sizes, and uneven distributions of easy versus difficult defects. Current convolution-based methods have achieved larger receptive fields to enhance perception, but at the expense of high-frequency components that contain small defects, making them unsuitable for road defect detection tasks. To address this, a road defect detection algorithm, FS-YOLO, based on frequency enhancement and synergy of geometric shape and category, is proposed. First, to balance the receptive field and high-frequency information, a frequency-adaptive dilation strategy is introduced, dynamically adjusting the spatial expansion rate according to local frequency components, and assigning appropriate convolutional kernels to defects of different sizes. Second, given that different types of defects have distinct geometric shapes and positions, an attention-based three-dimensional explicit synergy dynamic detection head is introduced to achieve explicit synergy between spatial geometric information and category information, enabling the model to leverage the inherent potential of defect categories and spatial locations. Finally, the Slide loss function is introduced to address the imbalance in the distribution of difficult and

① 基金项目: 国家自然科学基金 (62471272, 61806107, 62201314)

收稿时间: 2025-06-29; 修改时间: 2025-09-05; 采用时间: 2025-09-22; csa 在线出版时间: 2025-11-26

CNKI 网络首发时间: 2025-11-27

easy defects in real-world roads, particularly enhancing the model's ability to handle difficult-to-distinguish samples. Experimental results show that FS-YOLO significantly outperforms the baseline model in terms of precision and recall on both the self-built dataset and the public road defect detection datasets RDD 2022 and UAV-PDD. It has also been effectively validated in practical applications on expressways and national and provincial roads, significantly improving the accuracy and efficiency of defect detection.

Key words: road defect detection; deep learning; computer vision; YOLO; dilated convolution; explicit synergy; attention mechanism

道路病害是指道路在使用过程中由于自然因素、交通荷载以及施工质量等原因导致的各种损坏现象。这些病害不仅影响道路的使用寿命,还会对行车安全造成严重威胁^[1,2]。例如,裂缝可能导致雨水渗入地下,加速路面的结构损坏;坑槽则可能引发车辆颠簸甚至交通事故。因此,及时、准确地检测和修复道路病害对于保障交通安全、提高道路使用寿命具有重要意义^[3]。传统的道路病害检测方法主要依赖人工巡检和人工标注的方式,这种方法难以应对全国干线公路养护管理评价期间的巨大检测量,不但效率低下,人工成本高,而且人工检测面对复杂病害时的主观性强,检测结果强依赖于标注人员的经验,不同检测人员对病害的判断标准可能存在差异,难以保证检测结果的准确性和一致性。

尽管有许多机器学习、深度学习等自动化的方法在自动道路病害检测方面做出努力,但是面对道路病害极端长宽比、尺寸变化大、数量分布不均等特有问題,道路病害检测算法仍有进一步探索的空间。基于以上问題,本文提出了基于频率增强和形类协同的 YOLO-v8^[4]道路表面病害检测算法 FS-YOLO,实现了道路表面病害的检测与分类。面对道路病害尺寸变化大、卷积膨胀率与高频信息难以平衡的问题,本文引入基于频率的增强卷积模块,该模块能够根据局部频率分量动态调整卷积参数,从而为不同尺寸的病害分配合适的卷积核,有效解决了传统卷积难以适应多尺寸病害的问题。面对不同类别道路病害几何特征各有不同的情况,引入三维显式协同的动态检测头模块 (explicit synergy head, ESHead), 其在尺度感知、空间感知、任务感知这 3 个方面添加自注意力模块,显式地建模病害的几何特征与类别信息,增强了模型的信息感知能力,提升模型对复杂病害的检测效果。面对实际道路中不同类别病害数量分布不平衡的问题,引入了 Slide loss 损失函数^[5],该函数通过 IoU 阈值划分样本,并对

边界附近的样本采用滑动加权,平衡简单与困难样本的学习,特别提升了模型对困难样本的处理能力。

在自建数据集和 RDD 2022 数据集^[6]上进行的大量实验表明,FS-YOLO 在各类道路病害检测场景中都取得了较好的成绩。这不但证实了基于频率特点对特征进行增强是有效的,还说明本文方法为复杂道路病害检测这一挑战提供了有效、可实用的解决方案。

1 相关工作

1.1 膨胀率与频率响应映射的理论阐述

当卷积核的膨胀率从 1 提升至 D 时,其本质是在时域对原始核函数进行 D 倍零插入扩展。设原始卷积核的时域表达式为 $h(n)$, 扩展后的核函数 $h_D(n)$ 可定义为:

$$h_D(n) = \begin{cases} h\left(\frac{n}{D}\right), & n \equiv 0 \pmod{D} \\ 0, & \text{其他} \end{cases} \quad (1)$$

该操作在时域上表现为信号采样间隔的 D 倍扩展,即采样率降低至原来的 $1/D$, 而根据傅立叶变换的时域-频域对偶性,时域采样率的变化会直接引发频域的尺度变换,这为膨胀率与频率响应的映射提供了核心理论依据。

傅立叶变换的尺度特性^[7]明确指出:若离散时间信号 $x(n)$ 的傅立叶变换为 $\mathcal{F}\{x(n)\} = X(e^{j\omega})$, 其中 ω 为数字角频率,对信号进行时域尺寸变换后的信号 $x(an)$, 其傅立叶变换满足:

$$\mathcal{F}\{x(an)\} = \frac{1}{a} X(e^{j\omega/a}) \quad (2)$$

在膨胀率扩展场景中,时域尺度因子 $a = 1/D$ 。将其代入尺度特征公式可得到扩展后核函数 $h_D(n)$ 的傅立叶变换:

$$\mathcal{F}\{h_D(n)\} = \mathcal{F}\left\{h\left(\frac{n}{D}\right)\right\} = D \cdot H(e^{j\omega D}) \quad (3)$$

数字角频率 ω 的物理意义与信号带宽直接相关。原

始核 $h(n)$ 的频率响应带宽为 $\Omega_{\text{orig}} = [-\omega_{\text{max}}, \omega_{\text{max}}]$,也就是频率响应 $H(e^{j\omega})$ 非零的 ω 范围.相应地,对于扩展后的核函数,其频率响应 $H_D(e^{j\omega}) = D \cdot H(e^{j\omega D})$ 仅在 $\omega D \in [-\omega_{\text{max}}, \omega_{\text{max}}]$ 时非零,从而得到扩展后的带宽为:

$$\Omega_D = \left[-\frac{\omega_{\text{max}}}{D}, \frac{\omega_{\text{max}}}{D} \right] = \frac{1}{D} \Omega_{\text{orig}} \quad (4)$$

这表明,膨胀率 D 的提升会使卷积核的频率响应曲线在频域上压缩至原来的 $1/D$,带宽同步缩放至 $1/D$.这样的结论与直觉上对膨胀率的认知一致.

1.2 早期的道路病害检测算法

为了解决人工检测的低效、病害判断主观等问题,各类自动化的道路表面病害检测方法应运而生,这类研究主要解决道路图片中病害的检测和分类问题,其不但能快速、准确、客观地完成病害检测任务,而且能够有效减少人为因素的干扰,提高检测结果的准确性和一致性.

早期的道路病害检测研究主要集中在基于特征工程的机器学习方法.阈值分割法^[8,9]、边缘检测法^[3]、区域生长算法等方法最早被用来提取道路图像的视觉特征,在自动化病害检测中做出了重要的探索.基于监督学习的算法,如支持向量机(support vector machine, SVM)、随机森林(random forest)等^[10]首先利用传统方法进行视觉特征提取,随后将提取的视觉特征送入机器学习算法进行图像的分类,这些方法虽然在一定程度上提高了检测效率,但由于基于规则的方法难以应对复杂病害特征,捕捉能力有限,逐渐难以满足现代道路病害检测的需求.

1.3 基于深度学习的道路病害检测算法

随着计算机计算能力的提升以及数据量的积累,深度学习(deep learning)方法在目标检测领域取得了巨大的成功.例如,卷积神经网络(convolutional neural network, CNN)通过卷积层、池化层和全连接层的组合,自动地提取图像中的局部视觉特征,不需要人为的干预.这种自动特征提取能力使得深度学习模型在处理大规模、复杂的数据时具有显著优势.近年来,深度学习方法被广泛应用于目标检测任务中,典型的检测网络有R-CNN、SSD、YOLO、DETR等.R-CNN系列方法提取视觉特征后,利用启发式方法(如selective search)或区域提议网络(region proposal network, RPN)等生成候选框,用来定位目标物体.SSD方法在多尺度的特征图上依照规则生成不同大小和长宽比的先验框,

直接预测先验框的类别和偏移量,其检测速度快,能检测不同大小目标,但在小目标检测精度上略差.在道路病害检测领域,周秋红^[11]提出一种基于不规则四边形标注框的改进SSD道路裂缝检测方法,根据裂缝形态进行目标检测,在检测精度方面有了较高的提升. DETR架构结合Transformer和CNN,将目标检测转化为集合预测问题,通过编码器解码器结构预测目标类别和位置,其通过自注意力机制处理长距离依赖关系,不过推理速度相对较慢.朴恒剑等人^[12]在RT-DETR^[13]的基础上提出MSaE注意力机制,同时在编码器部分使用GhostConv卷积^[14]和DySample模块^[15]优化了上采样,有效提升了道路缺陷检测能力.YOLO(you only look once)作为典型的单阶段检测方法,在不同的小特征区域上进行多个预测,其检测速度快,适合实时应用,对小目标检测性能较好,且模型结构简单,易于部署和优化.

在道路病害检测任务中,YOLO凭借其便捷易用、推理速度快等特点受到了广泛的应用.Wang等人^[16]重构YOLOv8颈部网络并引入SimSPPF模块,减少模型的参数量的同时提高了模型的病害检测速度.高雪豪等人^[17]利用C2iAFF多尺度特征融合模块和通道注意力机制SE,提高了模型对于裂缝等细小病害特征的代表能力.Wang等人^[18]利用部分卷积设计了EMA Faster block结构取代YOLOv8 C2f模块中的瓶颈结构,并将增强后的C2f模块称为C2f-Faster-EMA,提升了模型的性能.Zhong等人^[19]引入C3Ghost模块增强模型的特征提取能力,并引入显式视觉中心EVC模块对特征金字塔层进行优化,从而提高模型的检测性能.

1.4 YOLOv8的目标检测算法

YOLOv8基于YOLO系列算法的核心思想,采用了一种端到端的检测框架.它将目标检测任务视为一个单一的回归问题,直接从输入图像预测目标的边界框和类别概率.与传统的两阶段检测方法相比,YOLOv8的单阶段检测机制大大提高了检测速度,使其在实时应用中表现出色.在架构设计上,YOLOv8引入了更高效的特征提取网络.它采用了改进的CSPDarknet作为骨干网络,通过引入跨阶段部分连接(cross stage partial connection)和深度可分离卷积(depth-wise separable convolution),在保持模型精度的同时显著降低了计算复杂度.这种设计使得YOLOv8能够在不同硬件平台上高效运行.YOLOv8在损失函数的设计上进行了优化.它采用了分层的损失函数,将边界框的定位损失、

类别预测损失和目标置信度损失进行了有机结合. 对于边界框的定位, YOLOv8 使用了更精确的 IoU 损失函数变体 $CIoU$ (complete IoU) 损失, 这种损失函数在考虑边界框的重叠比例的同时, 还引入了形状和尺度的惩罚项, 使得模型能够更准确地预测目标的边界框. 在训练策略方面, YOLOv8 引入了混合精度训练技术. 通过在训练过程中动态切换浮点精度和半精度计算, YOLOv8 能够在保持训练精度的同时大幅减少内存占用和计算时间. 此外, YOLOv8 还采用了数据增强技术, 如随机裁剪、颜色抖动和多尺度训练等, 这些技术能够增强模型对不同场景和光照条件的适应能力, 进一步提升检测性能.

在基于 YOLO 的研究中, 许多研究都集中于卷积操作的改进. 例如, 李思思等人^[20]利用膨胀卷积与多池化操作捕捉并融合多尺度特征信息, 提高模型识别能力. 刘美辰等人^[21]提出膨胀卷积融合膨胀注意力模块, 增强了模型的小目标检测能力. 曾孟佳等人^[22]引入 RFB 多分支空洞卷积层, 增大模型的感受野, 提升模型的特征提取能力. 尽管这些模型在各自领域可以发挥出色

的检测性能, 但是面对道路病害的特有问题, 如何平衡大感受野与高频信息仍然需要进一步讨论.

2 基于频率增强和形类协同的道路病害算法

为了应对道路病害特有的难点, 我们的模型做出对应的改进. 模型的结构如图 1 所示. 为了在扩大感受野的同时避免损失过多高频细节信息, 我们在 FADC^[23] 的基础上提出了基于频率的增强卷积模块 (frequency-based enhanced convolution, FEConv). 相较于传统卷积, FEConv 对卷积操作的 3 个重要部分: 特征图、膨胀率、卷积核分别做了适应性改进, 使得卷积操作在避免高频信息损耗的同时保持了较大的感受野, 兼顾了裂缝类细小病害和修补类大病害; 其次, 受 Dynamic head^[24] 的启发, 三维显式协同动态检测头 (explicit synergy head, ESHead) 被用来实现病害类别信息与空间几何特征的显式交互, 充分发挥二者的内在协同作用; 最后, 使用 Slide loss 损失函数进行反向传播, 通过交并比与阈值的接近程度来判别难易样本, 从而让模型更多关注难以分类的样本.

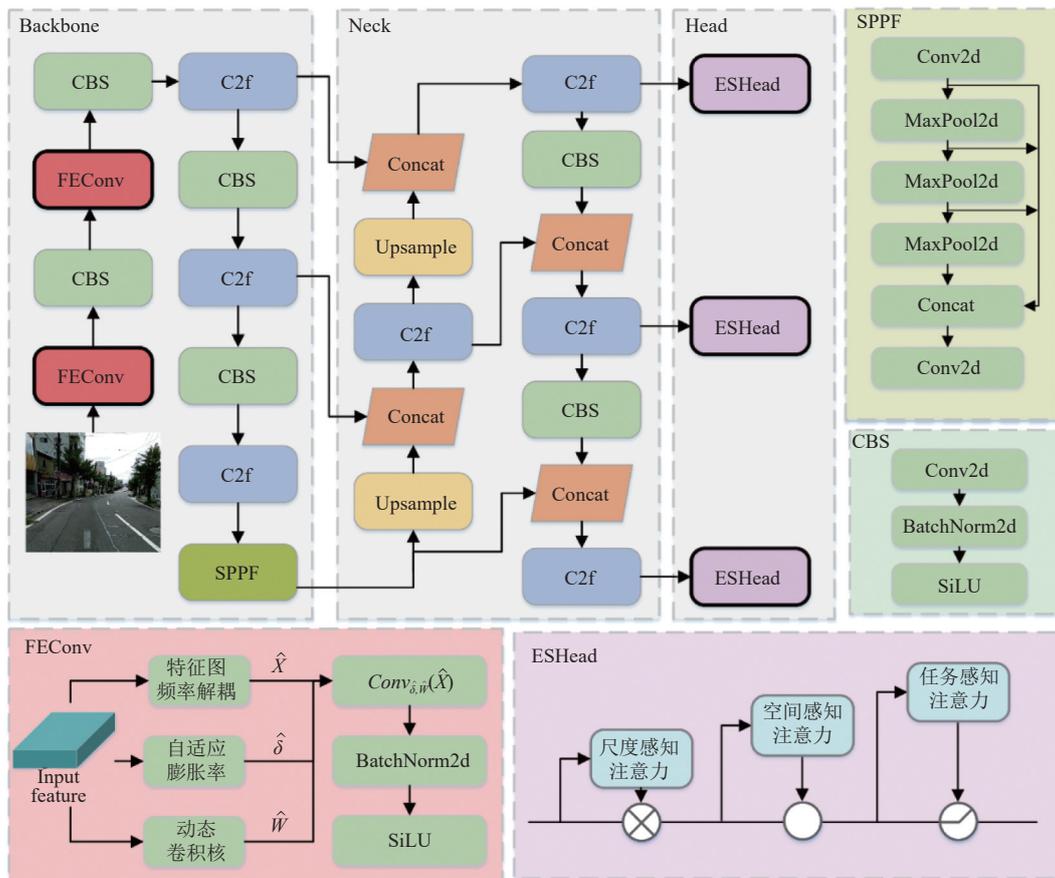


图 1 基于频率增强和形类协同的道路病害算法结构图

2.1 基于频率的增强卷积模块

膨胀卷积 (dilated convolution)^[25]是一种在卷积操作中引入空洞的技术, 在增加模型的感受野的同时不增加参数数量, 广泛应用于计算机视觉领域。

膨胀卷积可以描述为:

$$X'(p) = \sum_{i=1}^{K \times K} W_i X(p + \delta \times \Delta p_i) \quad (5)$$

其中, $X'(p)$ 代表 p 位置的特征经过扩张卷积后的输出, K 为卷积核尺寸大小, δ 是当前卷积层约定的膨胀率, $X(p + \delta \times \Delta p_i)$ 表示 p 位置经过偏移后采样的输入特征, 偏移量由 Δp_i 经过扩大膨胀率 δ 倍后得到. 特别地, 膨胀率 δ 在普通卷积中取值为1. W_i 是可学习的权重. 与普通卷积相比, 膨胀卷积引入的膨胀率 (dilation rate) 参数 δ , 使得卷积核可以在输入特征图上以更大的间隔进行操作, 从而保持输出特征图的尺寸不变, 同时有效提取更广感受野的特征, 但要承受高频分量响应损失的代价. 具体来说将卷积核的膨胀率从1增加到 δ 时, 相当于通过零插入的方式将卷积核扩展 δ 倍. 在图2(a)中, 常规卷积的膨胀率为1, 一个 3×3 的卷积核的感受野为 3×3 , 然而在图2(b)中, 膨胀率由1变为2, 使得感受野也变大为 5×5 . 然而, 根据傅立叶变换的缩放特性, 卷积核的频率响应曲线和带宽都会按 $1/\delta$ 的比例进行缩放. 带宽的减少极大地限制了该层处理高频成分的能力, 损害了模型捕捉图像细节的能力.

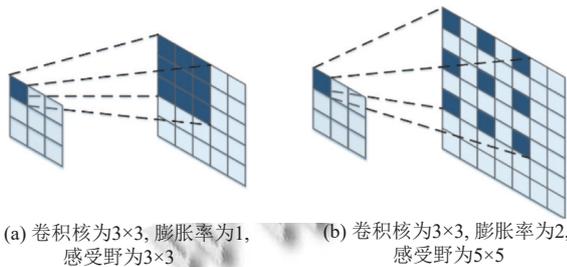


图2 不同膨胀率卷积核感受野的影响

为了在不损害高频分量响应的前提下扩大特征提取感受野, 本文从频谱分析的角度引入基于频率的增强卷积模块 (FEConv), 如图1左下角所示, FEConv通过3种策略来改进扩张卷积的各个阶段: 特征图频率解耦模块 (FFD) 将输入视觉特征解耦到不同的频域内, 通过为不同频域带加权的方式重新分配高低频信息, 在特征层面实现了频率和感受野的平衡; 自适应膨胀率模块 (ADR) 对 FFD 输出的特征图进行频率分析, 从

而为每个特征局部分配不同的膨胀率; 动态卷积核模块 (DK) 根据特征将传统静态卷积核权重解耦为高频和低频两部分, 从而进一步提高有效带宽的获取能力.

2.1.1 特征图频率解耦模块

为了让整体可学习的卷积操作充分发挥潜能, 我们首先使用特征图频率解耦模块 (feature map frequency decoupling, FFD), 通过将特征图依据阈值解耦到不同的频率带, 为各频率带施加不同的权重来动态调整不同频率信息, 具体操作如图3所示.

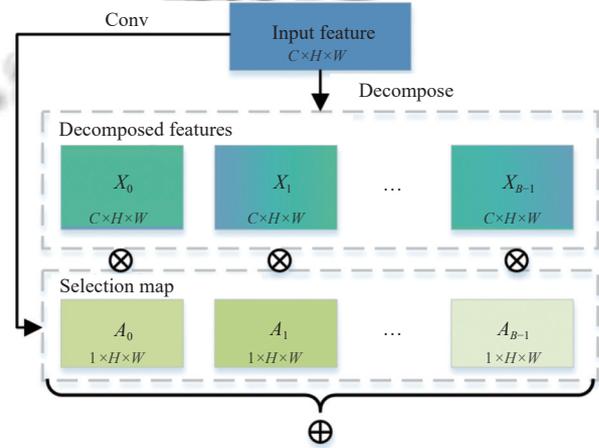


图3 特征图频率解耦模块

为了合理地进行特征解耦, 我们首先需要对输入的特征进行频率分析, 将特征从空间域变换到频率域中. 对于特征图 $X \in \mathbb{R}^{H \times W}$, 使用离散傅立叶变换 (discrete Fourier transform, DFT) 进行处理, 其可以表示为:

$$F(u, v) = \frac{1}{HW} \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} X(h, w) e^{-2\pi j(uh+vw)} \quad (6)$$

其中, $F \in \mathbb{R}^{H \times W}$ 为 DFT 输出的频率域特征, H 和 W 为其高宽, h 和 w 为特征图 X 的索引. 高度和宽度维度中的归一化频率分别由 $|u|$ 和 $|v|$ 给出.

对于频率域特征图, 我们利用不同掩码分解到不同频率带中, 再通过空间变化获得不同频率带中的视觉特征:

$$X_b = \mathcal{F}^{-1}(\mathcal{M}_b F) \quad (7)$$

其中, X_b 代表 b 频率带的空间特征. \mathcal{F}^{-1} 为逆快速傅立叶变换 (inverse fast Fourier transform, IFFT), 用于将频域信号转换回空间域. \mathcal{M}_b 是一个二值掩码 (binary mask), 用于提取特定频率范围内的成分:

$$\mathcal{M}_b(u, v) = \begin{cases} 1, & \text{if } \phi_b \leq \max(|u|, |v|) < \phi_{b+1} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

其中, ϕ_b 和 ϕ_{b+1} 是预定义的频率阈值, 用于划分不同的频率带, 其中 ϕ 取自预定义频率阈值集 $\{\phi_k | k = 0, 1, 2, \dots, B-1 \text{ and } \phi_k \in [0, 0.5]\}$. 在具体实现中, 我们按照八度卷积的做法将频域划分为 4 个频率带. 随后, FFD 通过重新加权的方法重新组合, 来平衡不同频率成分的权重:

$$\hat{X}(i, j) = \sum_{b=0}^{B-1} A_b(i, j) X_b(i, j) \quad (9)$$

其中, \hat{X} 则为增强后空间域的特征图. A_b 是第 b 个频率带的权重, 负责对频率分量选择性映射.

2.1.2 动态卷积核模块

传统的卷积操作中, 卷积核的权重是静态的, 也就是一旦训练完成就固定不变. 这种静态的卷积核虽然能够捕捉不同频率的信息, 但在处理复杂的视觉问题时仍有增强的空间. 为此, 动态卷积核模块 (DK) 根据不同的输入视觉特征来动态生成卷积核, 从而更好捕捉和处理低频与高频信息.

如图 4 所示, 我们首先将卷积核参数分解为低频与高频两部分, 这一过程会增加少量的参数数量和计算负担. 对于静态卷积核, 其权重解耦可以被描述为:

$$W = W_l + W_h \quad (10)$$

其中, W_l 是卷积核的低频部分, 充当一个 $K \times K$ 的低频均值滤波器, 其具体实现为卷积核权重的均值:

$$W_l = \frac{1}{K \times K} \sum_{i=1}^{K \times K} W_i \quad (11)$$

为增强卷积核对不同频率信息的捕捉能力, 动态卷积核 DK 模块自适应地调整低频和低频部分的权重:

$$\hat{W} = \lambda_l W_l + \lambda_h W_h = \lambda_l W_l + \lambda_h (W - W_l) \quad (12)$$

其中, \hat{W} 是调整后的动态卷积核权重, W_h 是卷积核的高频部分, 可以被定义为卷积核与低频部分的差值. λ_l 和 λ_h 作为卷积核的低频部分和高频部分的权重, 通过全局池化层和卷积层预测而来:

$$\begin{cases} \lambda_l^{(c)} = \text{GlobalPooling}(\text{Conv}_l(X^{(c)})) \\ \lambda_h^{(c)} = \text{GlobalPooling}(\text{Conv}_h(X^{(c)})) \end{cases} \quad (13)$$

其中, $X^{(c)}$ 表示在 c 通道的特征. 通过动态调整二者比重可以使得网络可以根据输入特征的复杂性, 灵活地关注低频背景信息或高频细节信息.

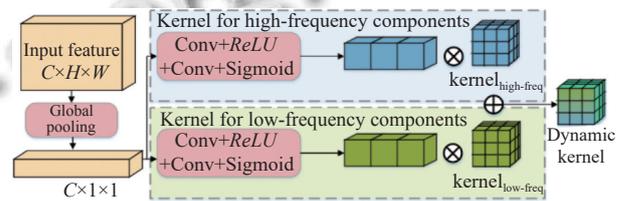


图 4 动态卷积核模块

2.1.3 自适应膨胀率模块

自适应膨胀率 (ADR) 模块基于频率信息为每个局部分配不同膨胀率. 如图 5(a) 所示, 图片的右上角有两条纵向裂缝, 左上角有一处白色污染物, 因此在图 5(b) 的特征图中, 污染物和裂缝的位置频率信息更高, 就意味着该局部中含有更多的有效视觉信息, 其余位置的沥青路面呈现出较低的频率, 也就是拥有更少的细节信息. 因此, 像图 5(c) 一样, 我们期待为低频部分如图 5(c1) 分配较大的膨胀率, 使其获得较大的感受野, 而图 5(c2) 中的纵向裂缝部分呈现出大量高频成分, 自适应膨胀率模块为该局部分配较小的膨胀率 ($\delta = 1$) 以避免局部高频细节的遗漏和丢失, 保持特征较宽的有效带宽.

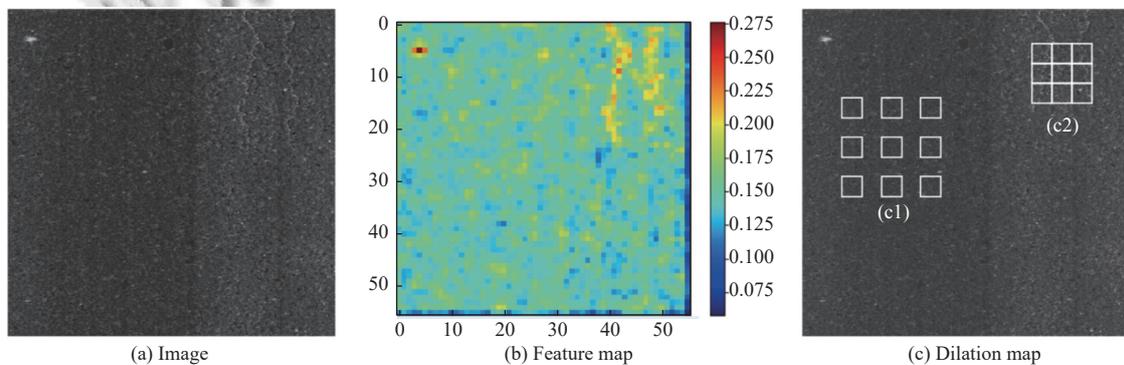


图 5 自适应膨胀率

在具体实现中,对于离散傅立叶变换得到的频率域特征图 F ,膨胀率 $\hat{\delta}$ 由参数为 θ 的卷积操作 $Conv_{\theta}$ 所预测:

$$\hat{\delta} = ReLU(Conv_{\theta}(F)) \quad (14)$$

其中, $ReLU$ 用来保证膨胀率非负, θ 作为优化目标旨在平衡感受野与高频信息,即最大化感受野与丢失的高频信息之间的差值.关于感受野的描述,我们称窗口大小为 s ,位置为 p 的特征为 $F^{(p,s)}$,它的感受野 $\mathcal{RF}(p)$ 与 $\hat{\delta}(p)$ 变化趋势相同,可以被描述为:

$$\mathcal{RF}(p) = \hat{\delta}(p) \times (K - 1) + 1 \quad (15)$$

同时,为了衡量扩张率增加导致的高频信息丢失,我们计算高频频率:

$$\mathcal{HP}(p) = \sum_{(u,v) \in \mathcal{H}_{\hat{\delta}(p)}^+} |F^{(p,s)}(u,v)|^2 \quad (16)$$

至此,优化目标可以被描述为:

$$\theta = \max_{\theta} \left(\sum \mathcal{RF}(p) - \sum \mathcal{HP}(p) \right) \quad (17)$$

最终,我们将以上模块所提到的增强特征图、动态卷积核、自适应膨胀率应用到卷积操作中:

$$X'(p) = \sum_{i=1}^{K \times K} \hat{X}_i \hat{W}(p + \hat{\delta}(p) \times \Delta p_i) \quad (18)$$

2.2 三维显式协同动态检测头

在传统目标检测任务中,建模物体之间的位置和形状关系可以更好地帮助物体识别,道路表观病害也是如此.不同类别病害所具有的形状和位置特点各有不同,尽管 YOLOv8 的目标检测头已经隐式地建模了空间与类别的关系,但是这种建模难以充分发挥两优化目标的协同作用.我们认为将病害的类别判断与几何信息进行显式交互,才能充分发挥两优化目标之间的内在的协同潜能.

为此,本文引入了三维感知显式协同的动态检测头 (ESHead),如图 6 所示,其利用多个自注意力机制在尺度感知、空间感知、任务感知之间进行显式建模,显著提升目标检测头的多目标协同和表达能力.具体来说,动态检测头利用 3 种不同的注意力机制实现不同方面的感知,其可以表示为:

$$x' = \mathcal{A}_C(\mathcal{A}_S(\mathcal{A}_L(x), x), x) \quad (19)$$

其中,尺度感知注意力 \mathcal{A}_L 应用在水平尺度上;空间感知注意力 \mathcal{A}_S 关注特征的空间维度;任务感知注意力 \mathcal{A}_C 则部署在通道维度上,其具体更新流程如图 6 所示.

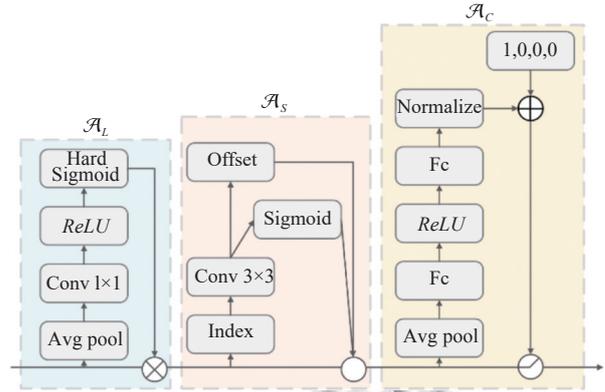


图 6 三维显式协同动态检测头

2.3 改进的损失函数

YOLOv8 模型采用了复合损失函数 BCEWithLogitsLoss 来判定当前预测框的类别置信度,其内部首先运用 Sigmoid 激活函数,然后计算二元交叉熵损失 (BCE) 损失.但由于在道路缺陷检测场景下,数据分布存在难易样本不均衡的问题.数据集中简单样本数量较多,难样本数量少,导致模型倾向于关注简单样本而无法准确分类困难样本.针对这一问题,我们将 Slide loss 权重函数引入道路病害检测任务中来优化函数,通过衡量检测框与目标框的相近程度来应对样本不平衡的挑战.在训练阶段计算损失过程时,该函数为难分类的样本分配较大膨胀因子 α ,引导模型更多地关注难以分类的样本,从而提升模型难样本的辨别能力并提高检测精度.具体而言,膨胀因子的确定依赖预测框与真实框之间的交并比 (IoU),对于 IoU 阈值 μ ,将 IoU 接近阈值 μ 的样本视为难样本, IoU 与 μ 相距较大的样本划分为容易样本.对于难分类样本,膨胀因子 α 将会更大.权重函数由函数 $f(x)$ 计算得到:

$$\alpha = f(x) = \begin{cases} 1, & x \leq \mu - 0.1 \\ e^{1-\mu}, & \mu - 0.1 < x < \mu \\ e^{1-x}, & x \geq \mu \end{cases} \quad (20)$$

在具体实现中,我们使用所有边界框 IoU 的均值作为 μ 的取值.依据样本 IoU 与阈值 μ 的关系,利用不同的权重函数对原 BCEWithLogitsLoss 损失进行加权处理,再进行反向传播:

$$\mathcal{L}_{Slide} = \alpha \cdot \mathcal{L}_{BCEWL} \quad (21)$$

在图 7 中我们取 $\mu = 0.5$,可以发现滑动损失在 0.5 附近的权重更大.对于负样本, IoU 小于 $\mu - 0.1$ 的部分不进行放大, IoU 在 $[\mu - 0.1, \mu]$ 的负样本有最高的权重,意味着需要重点学习优化.对于正样本,随着 IoU

距离阈值 μ 的距离 $|x-\mu|$ 越大,意味着样本可信度越高,从而优化的权重逐渐变小。

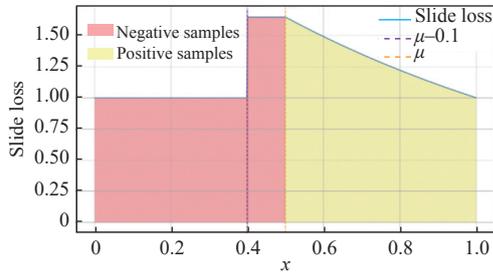


图7 Slide loss 损失函数膨胀权重

3 实验与分析

为了验证本文方法在道路病害检测任务中的有效性,首先我们进行消融对比各模块改进相较于基线模型 YOLOv8n 的差距,来验证模块的有效性。随后,在公开数据集 RDD 2022 和自建数据集上进行了对比实验,证明模型整体改进对于检测结果的提升。

3.1 数据集

我们团队在发展过程中收集了沥青路面道路病害图片 94 187 张,包含中国多个城市的高速公路、国省道等,图片均自顶而下拍摄采集。数据集标签包含纵向裂缝、横向裂缝、龟裂、块状裂缝、坑槽、松散、条状修补、块状修补等主要检测类别,主要病害的数量如表 1 所示。在本文模型实验过程中,按照 8:1:1 的比例划分训练集、验证集和测试集。

表 1 自建数据集道路病害标签情况

病害名称	训练集数量	验证集数量	测试集数量	数据集总数
纵向裂缝	24169	3071	3790	31030
横向裂缝	14579	1921	2104	18604
条状修补	79791	10192	9817	99800
块状修补	9300	1177	1271	11748
龟裂	1367	191	162	1720
块状裂缝	1137	134	163	1434
坑槽	1729	218	261	2208

在将前沿算法应用到实际行业的过程中,数据标注标准的制定格外重要。数据标注质量的高低直接限制了后续模型学习的上限,尤其在数据量较小的情况下。相较于当前的常规矩形框标注,使用小框密集标注的方式虽然能够更好地拟合任意形状病害,但是其小框捕捉特征不明显,在真实道路污染物多的情况下,容易出现误检的情况。如在国、省道路面中,道路上的

水渍、油滴、泥土等各类干扰物极易被识别为病害,造成误检问题,降低模型准确率。为此我们额外标注了许多非病害类别以增强模型的辨别能力。此外,道路病害类别的判别也至关重要,例如施工缝与裂缝的区分、修补与大面积修补的划分、裂缝与龟裂块状裂缝的辨别等都需要科研算法人员与业务人员依据国标行业标准进行研讨共同制定。让标注结果在分类正确的前提下,尽可能多地蕴含明显的视觉特征,降低模型的学习难度,做到业务正确与算法可实现的平衡。在自建数据集标注过程中,我们使用了项目历史标注数据对模型进行训练。依靠包含脏数据的约 14 万张图片对模型进行预训练可以得到基本可信的结果,随后我们通过脚本对比标注结果与预测结果,依次筛选出误标注、漏标注的数据,分批进行调整,大大减小了标注与工作量。

同时,在两个不同角度的公开数据集上进行了实验。RDD 2022 是一个多国图像数据集,用于自动道路损伤检测,由印度理工学院罗凯里分校交通系统中心等机构创建。该数据集包含来自 6 个国家的 47 420 张道路图像,标注了超过 55 000 个道路损伤实例。为保持数据的一致性,选择了其中 14 884 张从车顶向前采集的数据作为实验数据,数据集按照 8:1:1 进行划分训练集、验证集和测试集。除了以上道路巡检数据集外,我们还在无人机采集的数据集 UAV-PDD 上进行了验证。该数据集共包含 2 440 张三通道 JPG 格式图像,同时为了增强实用性,作者特意在下雨过后 1 h 拍摄了部分路段的照片。图像中分别标记了 6 种类型的道路损坏,包括纵向裂缝、横向裂缝、鳄鱼裂缝、斜裂缝、修补和坑槽。

3.2 实验设备配置及评价指标

实验服务器使用 Ubuntu 20.04.6 LTS 系统,训练使用显存为 80 GB 的 NVIDIA A100 显卡,CPU 为 Intel(R) Xeon(R) Gold 6342 CPU@2.80 GHz,运行内存 512 GB。本模型基于 Python 3.9、PyTorch 2.6.0 深度学习框架进行训练、验证,配置 CUDA 12.4 版本, batch_size 设置为 64, epoch 最大被设置为 200,网络学习率为 0.01,优化器使用 SGD。将图片缩放到 1024×1024 时占用显存约为 26 GB。模型在训练时使用了 mosaic、图像翻转、光照条件变化、运动模糊等 YOLO 默认图像增强设置。在推理时,本文模型可以在单张 GeForce RTX 2080 上完成推理。

为了衡量网络模型性能,采用目标检测常用的评

价指标: 召回率 (R)、准确率 (P)、F1 分数 (F1)、平均正确率 (mAP@0.5)。其中, 召回率 (R) 指检测到真实病害数量与全部真实病害数量的比值, 衡量模型查全能力。准确率 (P) 是检测真实病害数量与检测到全部目标数量之比, 反映模型检测准确性。F1 分数 (F1) 是召回率和准确率的调和平均数, 综合衡量两者平衡。平均正确率 (mAP@0.5) 在目标检测等任务中, 以 *IoU* 阈值 0.5 为标准, 计算模型对不同类别病害检测准确性的平均值, 评估模型整体性能。

模型效率也是模型能否工程应用的一个重要考虑因素, 在本文中每秒处理的图片数量 (FPS) 来衡量。模型的规模大小使用计算量和参数量两个指标来衡量。计算量反映了模型的计算复杂度, 通常以浮点运算次数 FLOPs 来计数, 本文中使用 GFLOPs (10^9 FLOPs) 来表示。参数量则是模型所有参数的总数, 通常以百万 (M) 计数。

3.3 消融实验

对于各模块的消融实验。为了验证各模块的有效性和对模型性能的影响, 我们在 RDD 2022 数据集上进行消融实验, 实验结果如表 2 所示。F1 和 mAP@0.5 被作为衡量模型的实验指标, baseline 模型为 YOLOv8n, “√”表明当次实验中使用了该模块。基线模型的 YOLOv8n 对图片全局使用相同的膨胀率, 难以兼顾高频信息与更大的感受野。添加 FEConv 模块的方法在 F1 和 mAP@0.5 的指标上实现了 3.6% 和 1.4% 的提升。这得益于频率的增强卷积模块更好地为不同的局部特征分配不同的膨胀率。在基线模型添加 ESHead 显示协同预测头后, 模型的检测 F1 提升至 80.9%, 证明了挖掘空间特征与病害类别的内在关系能够增强模型的判别能力。添加了 Slide loss 后, 模型也在 F1 和 mAP@0.5 的指标上获得了 0.9% 和 0.6% 的增益, 这表明增强难区分样本有益于提高模型的能力。

表 2 各模块有效性验证的消融实验

baseline	FEConv	ESHead	Slide loss	F1 (%)	$\Delta F1$ (%)	mAP@0.5 (%)	$\Delta mAP@0.5$ (%)
√	—	—	—	77.1	+0	73.5	+0
√	√	—	—	80.7	+3.6	74.9	+1.4
√	—	√	—	80.9	+3.8	74.3	+0.8
√	√	√	—	81.3	+4.2	76.9	+3.4
√	—	—	√	78.0	+0.9	74.1	+0.6
√	√	—	√	81.7	+4.6	77.6	+4.1
√	—	√	√	81.9	+4.8	78.3	+4.8
√	√	√	√	83.6	+6.5	79.1	+5.6

• 对特征图频率解耦模块的分析。如表 3 所示, 我们对频率域分为 4 段, 同时对模型为不同频率域所分配的权重的均值进行了统计。结果表明, 模型会为高频率域的特征分配更低的平均权重。这可能是由于在道路病害图片中, 模型更倾向于将高权重分配给裂缝类病害, 或面积较大病害的边缘区域, 而这一部分在图像中只占很小的面积。

表 3 不同频率域的平均权重统计

频率域	[0, 1/16]	(1/16, 1/8]	(1/8, 1/4]	(1/4, 1/2]
平均权重	1.0	0.72	0.41	0.28

除了按照八度卷积的方法进行频率划分之外, 我们还使用平均划分的方法对频率域进行划分。如表 4 所示, 可以看到基于八度卷积划分方法的 F1 指标超出平均划分方法 6.6%。这表明八度卷积这种类似于人类视觉模式的划分方式也适用于深度学习领域。

表 4 不同频率带划分方法的性能比较 (%)

频率域划分方法	R	P	F1	mAP@0.5
平均划分	71.9	82.9	77.0	75.3
八度卷积	77.2	91.2	83.6	79.1

• 对自适应膨胀率模块的分析。为了比较不同膨胀率对模型特征感知及预测能力的影响, 我们对不同膨胀率以及自适应膨胀率的方法进行了消融实验。实验结果如表 5 所示, 可以看到自适应膨胀率的方法相较于固定膨胀率的方法取得了更好的效果。

表 5 不同膨胀率的性能比较 (%)

膨胀率	R	P	F1	mAP@0.5
1	75.2	80.9	77.9	73.1
2	77.1	86.6	81.6	77.2
4	71.2	83.9	77.0	69.2
ADR (ours)	77.2	91.2	83.6	79.1

• 对动态卷积核模块的分析。对于分离后的卷积核, 通过动态调整低频部分和高频部分的比重, 可以有效提高模型的有效带宽, 增强模型的检测性能。为了量化这一模块的效果, 我们对两模块的权重比例进行了消融实验。如表 6 所示, 相较于固定比值的融合方式, 使用自适应的方式可以让模型依据当前输入图像自主地选择卷积核组成比例, 从而更好地适应不同的场景。

表 6 卷积核不同比例的性能比较 (%)

λ_h/λ_l	R	P	F1	mAP@0.5
0.5	74.6	83.1	78.6	74.9
1.0	75.3	89.7	81.9	77.8
2.0	72.9	90.5	80.8	76.7
自适应	77.2	91.2	83.6	79.1

3.4 实验结果分析

为了验证本文方法在道路病害检测中的优势,我们和其他主流的目标检测方法进行了对比实验.首先,在自建数据集上进行实验,实验结果如表7所示.通过实验比较了FS-YOLO模型与其他经典模型在病害检测方面的性能.FS-YOLO在各项指标上均取得了优异的成绩.具体来看,FS-YOLO模型的F1分数达到了73.8%,超过Faster R-CNN、YOLOv8n分别12.8%、4.9%,这表明相较于两阶段的目标检测算法,单阶段的方法在提高检测速度的同时还可以通过避免次优化的

方式提高检测精度.相较于RT-DETR方法,我们不但在精度上实现了超越,也表明了本文改进的有效性,而且基于YOLO的方法使得所提方法在参数量上有较为显著的优势.虽然YOLO不可避免的后处理会增加图片的处理时间,但是NMS这种可控的后处理操作仍有应用价值.此外,在工程应用中,DETR系列模型训练时间长,其能否满足项目推进中的时间要求是一个必须考虑的问题.具体来看,召回率(R)为66.1%,精确率(P)为83.6%,较高的精确率也贴合了用户在道路评价工作中的实际要求.

表7 不同模型在自建数据集上的性能比较

Model	R (%)	P (%)	F1 (%)	mAP@0.5 (%)	Δ mAP@0.5 (%)	Params (M)	GFLOPs	FPS (f/s)
YOLOv3	59.19	61.4	60.3	63.2	0	61.9	19.0	83.0
YOLOv4	57.9	59.3	58.6	61.7	-1.5	61.3	16.5	67.0
YOLOv5	60.3	63.9	62.0	64.1	-0.9	2.5	7.2	78.7
YOLOv8n	63.7	<u>77.3</u>	69.8	70.3	+7.1	3.0	8.2	84.8
Faster R-CNN	58.3	63.9	61.0	59.1	-4.1	60.1	108.6	29.0
RT-DETR	67.1	75.9	<u>71.2</u>	<u>71.3</u>	+8.1	42.0	136.0	108.0
FS-YOLO (ours)	<u>66.1</u>	83.6	73.8	75.2	+12.0	3.7	8.6	83.3

注:加粗字体表示最优数值,下划线表示次优数值

表8和图8中,我们详细列举了主要病害类别的检测效果指标以及混淆矩阵,可以观察到模型对于工程关注的类别实现了较高的准确率.在混淆矩阵中,我们可以直观地观察到容易分类错误的病害类别,如脱离上下文的情况下条状修补和块状修补难以区分,松散与白色土渍在黑白图像中难以辨别.

表8 FS-YOLO在自建数据集主要病害类型上的性能(%)

类别	P	R	F1	mAP@0.5
坑槽	83.0	73.1	77.7	65.3
条状修补	93.8	96.4	95.1	96.4
块状修补	62.6	71.9	66.9	69.5
块状裂缝	78.5	54.7	64.5	61.2
龟裂	94.8	67.9	79.1	80.4
横向裂缝	80.3	72.4	76.1	72.2
纵向裂缝	76.0	76.6	76.3	75.7

将YOLOv8n与本文FS-YOLO模型的检测结果进行可视化,结果如图9所示.由图9(d)、(e)可以看到,FS-YOLO可以更好地识别裂缝这类信息,这归功于基于频率的增强膨胀卷积模块更好地捕捉了高频细节信息.

为了进一步验证FS-YOLO的有效性,我们将FS-YOLO与其他模型在道路巡检数据集RDD 2022及无人机数据集UAV-PDD上进行了进一步比较.表9展

示了FS-YOLO在前景摄像头采集的数据集RDD 2022上的性能,可见其各项指标均高于其他网络.我们将这归因于FEConv的自适应设计,这使得模型应对不同场景的图像都可以展现出充分的鲁棒性.表10展示了模型在航拍数据集UAV-PDD上的结果.FS-YOLO在F1指标上高出基线模型1.2%,在mAP@0.5上高出基线模型0.9%.结果表明,我们的方法可以广泛应用于道路病害检测任务中.

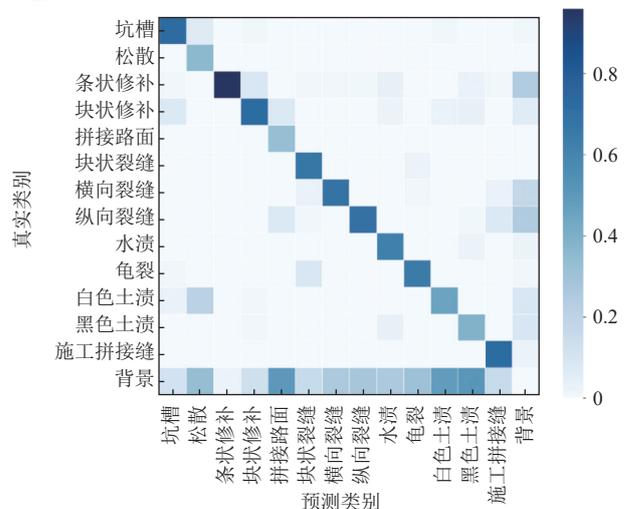


图8 模型在自建数据集上的混淆矩阵

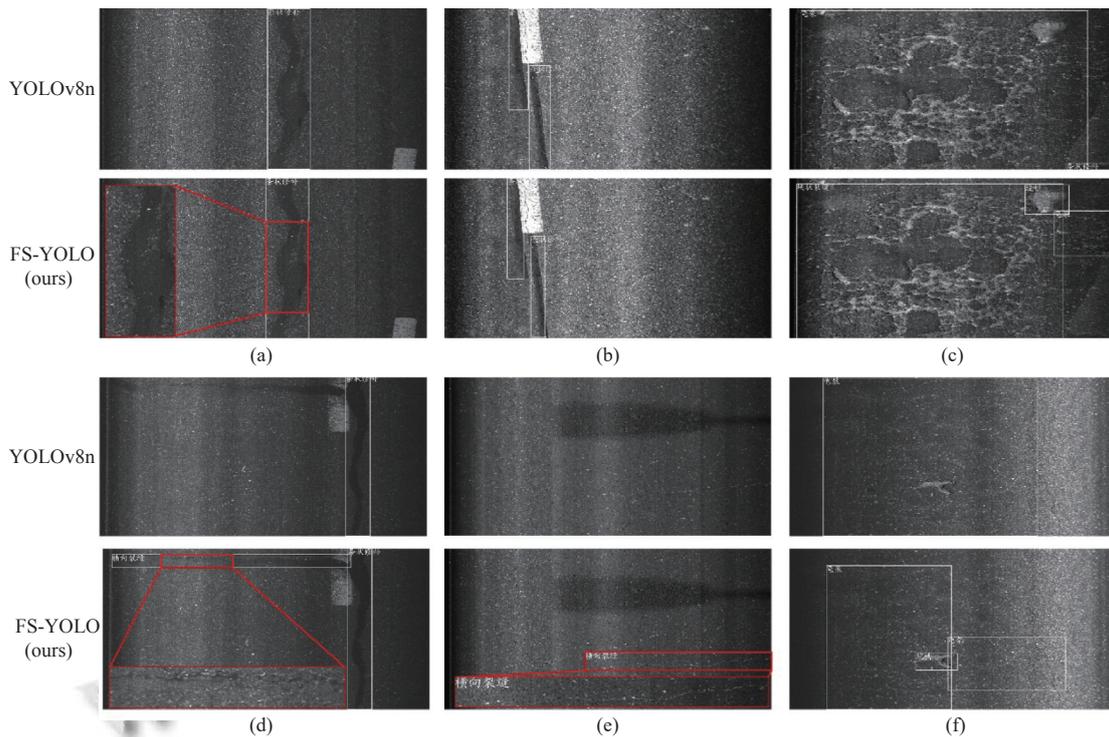


图9 YOLOv8n与FS-YOLO在自建数据集上的识别结果

表9 不同模型在RDD 2022上的性能比较(%)

Model	R	P	F1	mAP@0.5
YOLOv5	71.8	80.4	75.9	72.0
YOLOv7	66.6	76.4	71.2	69.5
YOLOv8n	<u>73.4</u>	<u>81.1</u>	<u>77.1</u>	<u>73.5</u>
FS-YOLO (ours)	77.2	91.2	83.6	79.1

注:加粗字体表示最优数值,下划线表示次优数值

表10 不同模型在UAV-PDD上的性能比较(%)

Model	R	P	F1	mAP@0.5
YOLOv5	85.1	83.9	84.5	91.7
YOLOv8n	89.1	88.9	89.0	93.0
FS-YOLO (ours)	91.3	89.1	90.2	93.9

注:加粗字体表示最优数值

4 结论与展望

本文针对道路病害检测中存在的病害尺寸多样性、类别与位置差异以及难易病害分布不均等问题,提出了一种基于频率增强和形类协同的检测算法FS-YOLO.通过引入基于频率的增强卷积模块,实现了大感受野与高频信息的平衡,使模型能够更好地处理不同尺寸的病害;借助三维显式协同的动态检测头,充分发挥了病害类别与空间位置的内在关联,提升了检测的准确性;采用Slide loss损失函数则有效解决了病害

分布不平衡问题,增强了模型对困难样本的处理能力.实验结果不仅充分证明了改进模型在公开数据集及实际道路场景中的优越性能,精确率和召回率均显著优于基线模型,而且展现出了模型在实际应用中的卓越表现.接下来,应对极端复杂路况和有水路面时,模型的鲁棒性可能还有待进一步提升.未来的工作将致力于进一步丰富雨水路面、省道复杂路面数据量,优化模型结构,提高模型在各种复杂条件下的鲁棒性和泛化能力,为道路养护和管理提供更加可靠、高效的技术支持.

参考文献

- 1 Zhou Y, Guo XM, Hou FJ, et al. Review of intelligent road defects detection technology. Sustainability, 2022, 14(10): 6306. [doi: 10.3390/su14106306]
- 2 柳雨豪, 罗浩原, 黄晓明. 智能公路系统下的路面病害图像检测技术综述. 现代交通与冶金材料, 2023, 3(1): 9-20.
- 3 Cao WM, Liu QF, He ZQ. Review of pavement defect detection methods. IEEE Access, 2020, 8: 14531-14544. [doi: 10.1109/ACCESS.2020.2966881]
- 4 Varghese R, M S. YOLOv8: A novel object detection algorithm with enhanced performance and robustness. Proceedings of the 2024 International Conference on

- Advances in Data Engineering and Intelligent Computing Systems (ADICS). Chennai: IEEE, 2024. 1–6.
- 5 Yu ZP, Huang HB, Chen WJ, *et al.* YOLO-FaceV2: A scale and occlusion aware face detector. *Pattern Recognition*, 2024, 155: 110714. [doi: [10.1016/j.patcog.2024.110714](https://doi.org/10.1016/j.patcog.2024.110714)]
 - 6 Arya D, Maeda H, Ghosh SK, *et al.* RDD2022: A multi-national image dataset for automatic road damage detection. *Geoscience Data Journal*, 2024, 11(4): 846–862.
 - 7 Proakis JG. *Digital Signal Processing: Pearson New International Edition*. Pearson, 2013.
 - 8 Otsu N. A threshold selection method from gray-level histograms. *Automatica*, 1975, 11(285–296): 23–27.
 - 9 Talab AMA, Huang ZC, Xi F, *et al.* Detection crack in image using Otsu method and multiple filtering in image processing techniques. *Optik*, 2016, 127(3): 1030–1033. [doi: [10.1016/j.ijleo.2015.09.147](https://doi.org/10.1016/j.ijleo.2015.09.147)]
 - 10 王一帆. 基于支持向量机和探地雷达技术的公路浅层病害检测研究 [硕士学位论文]. 郑州: 华北水利水电大学, 2019.
 - 11 周秋红. 基于改进 SSD 模型的道路病害检测研究. *黑龙江交通科技*, 2023, 46(4): 30–32.
 - 12 朴恒剑, 朱明. 基于改进 RT-DETR 的道路缺陷检测. *计算机系统应用*, 2025, 34(7): 107–116. [doi: [10.15888/j.cnki.csa.009877](https://doi.org/10.15888/j.cnki.csa.009877)]
 - 13 Zhao YA, Lv WY, Xu SL, *et al.* DETRs beat YOLOs on real-time object detection. *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2024. 16965–16974.
 - 14 Han K, Wang YH, Tian Q, *et al.* GhostNet: More features from cheap operations. *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 1577–1586.
 - 15 Liu WZ, Lu H, Fu HT, *et al.* Learning to upsample by learning to sample. *Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision*. Paris: IEEE, 2023. 6004–6014.
 - 16 Wang XQ, Gao HB, Jia ZM, *et al.* BL-YOLOv8: An improved road defect detection model based on YOLOv8. *Sensors*, 2023, 23(20): 8361. [doi: [10.3390/s23208361](https://doi.org/10.3390/s23208361)]
 - 17 高雪豪, 吴建平, 韦杰, 等. 基于增强多尺度融合 YOLOv8 的道路病害检测算法. *计算机技术与发展*, 2024, 34(11): 140–147.
 - 18 Wang JL, Meng RF, Huang YH, *et al.* Road defect detection based on improved YOLOv8s model. *Scientific Reports*, 2024, 14(1): 16758. [doi: [10.1038/s41598-024-67953-3](https://doi.org/10.1038/s41598-024-67953-3)]
 - 19 Zhong SH, Chen CL, Luo WS, *et al.* DDVC-YOLOv5: An improved YOLOv5 model for road defect detection. *IEEE Access*, 2024, 12: 134008–134019. [doi: [10.1109/ACCESS.2024.3453914](https://doi.org/10.1109/ACCESS.2024.3453914)]
 - 20 李思思, 葛华勇, 孟亚群, 等. 基于 AGD-YOLO 的钢材表面缺陷检测算法. *计算机测量与控制*, 2025, 33(11): 142–148. [doi: [10.16526/j.cnki.11-4762/tp.2025.11.017](https://doi.org/10.16526/j.cnki.11-4762/tp.2025.11.017)]
 - 21 刘美辰, 李杰, 陈廷伟. DTI-YOLO: 改进 YOLOv10s 的交通标志检测模型. *计算机工程与应用*, 2025, 61(17): 112–122.
 - 22 曾孟佳, 徐欢, 黄旭. 一种基于改进 YOLO v5n 的黄桃虫害检测方法. *安徽农业科学*, 2025, 53(3): 236–242.
 - 23 Chen LW, Gu L, Zheng DZ, *et al.* Frequency-adaptive dilated convolution for semantic segmentation. *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2024. 3414–3425.
 - 24 Dai XY, Chen YP, Xiao B, *et al.* Dynamic head: Unifying object detection heads with attentions. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 7369–7378.
 - 25 Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. *Proceedings of the 4th International Conference on Learning Representations*. San Juan: ICLR, 2016.

(校对责编: 李慧鑫)