

基于 YOLOv11n 的输水管道缺陷检测方法及效果分析^{*}

张 杨 秦昌浩 孔 硕 曲娅珺 赵 辉

(东北林业大学机电工程学院 哈尔滨 150040)

摘 要: 针对输水管道内壁背景纹理复杂、微小裂纹与不规则脏污难以区分等问题,提出了一种基于改进 YOLOv11n 的高精度检测算法。首先,利用双向特征金字塔网络(BiFPN)重构颈部网络,通过快速归一化融合机制实现多尺度特征的动态聚合,提升对微小及变尺度目标的表征能力;其次,在特征融合末端嵌入 GCNet 模块,建立像素级长距离依赖以增强全局背景理解,有效抑制水垢等噪声干扰;最后,构造 Focaler-MPDIoU 损失函数,结合几何优化与动态聚焦机制,提升定位精度并均衡难易样本训练。实验表明,改进算法在自建数据集上的平均精度均值(mAP) mAP50 达到 91.40%,较基线提升约 9.5%,在满足实时性的同时显著增强了鲁棒性,具有较高的工程应用价值。

关键词: 输水管道;缺陷检测;YOLOv11n

中图分类号: TN919.5;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Defect detection methodology for water conveyance pipelines
based on YOLOv11n and its performance analysis

Zhang Yang Qin Changhao Kong Shuo Qu Yajun Zhao Hui

(College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, China)

Abstract: To address the challenges associated with complex background textures on the inner walls of water pipelines and the difficulty in distinguishing minute cracks from irregular fouling, this paper proposes a high-precision detection algorithm based on an improved YOLOv11n. First, the bidirectional feature pyramid network (BiFPN) is utilized to reconstruct the neck network. By employing a fast normalized fusion mechanism, this approach achieves dynamic aggregation of multi-scale features, thereby enhancing the feature representation capability for minute and variable-scale targets. Second, a global context network (GCNet) module is embedded at the end of the feature fusion stage to establish pixel-level long-range dependencies. This strengthens global background understanding and effectively suppresses noise interference such as limescale. Finally, a Focaler-MPDIoU loss function is constructed by integrating geometric optimization with a dynamic focusing mechanism to improve localization accuracy and balance the training of hard and easy samples. Experimental results demonstrate that the improved algorithm achieves an mAP50 of 91.40% on a self-constructed dataset, representing an improvement of approximately 9.5% over the baseline. The proposed method significantly enhances robustness while satisfying real-time requirements, offering substantial value for engineering applications.

Keywords: water pipeline; defect detection; YOLOv11n

0 引 言

输水管道作为城市供水系统的“生命线”,其健康状况直接关系到城市的供水安全与社会运行的稳定性^[1]。由于长期埋设于地下,管道内壁易受多种因素影响而产生各类病害^[2]。常见的典型病害主要包括两类:1)结构性缺陷^[3],

如管壁的破裂(裂缝、破洞)和接口处的错位或脱节,这类缺陷直接威胁管道的结构完整性,易引发渗漏甚至灾难性的爆管事故;2)功能性缺陷,如长期运行形成的水垢沉积、树根侵入及异物滞留等,这类病害会大幅缩减管道的有效过流断面,降低输水效率,严重时导致管网堵塞瘫痪。

近年来,基于深度学习的计算机视觉技术为管道缺陷自

收稿日期:2025-12-02

^{*} 基金项目:黑龙江省应用技术与开发计划项目(GA21A403)、2025 年高教强省专项-产教融合研究生联合培养基地建设项目(41502123)资助

动化检测提供了新的解决方案。针对复杂场景下的目标识别,林哲等^[4]提出了融合改进 YOLO 与语义分割的方法以解决遮挡目标抓取难题,李春晖等^[5]则成功将 YOLOv8 应用于油气管道漏磁检测的缺陷智能识别,验证了该系列算法的有效性。YOLO (you only look once) 算法由 Redmon 等^[6]首次提出,因其卓越的速度与精度平衡,在工业缺陷检测领域得到了广泛应用。然而,尽管 YOLOv8、YOLOv10 及实时检测 Transformer (real-time detection Transformer, RT-DETR) 等模型在通用任务上表现优异,但面对输水管道极其特殊的背景纹理与跨尺度病害时,仍存在局限性。近年来,先进的目标检测算法已被成功应用于复杂管道环境下的缺陷识别任务中^[7]。因此,本文选取 2024 年发布的最新一代检测模型 YOLOv11n 作为基线(Baseline)。

选型依据主要基于其在轻量化架构上的突破性演进:YOLO11 作为该系列的最新版本,通过重新设计的骨干网络(Backbone)与颈部架构(Neck),在减少参数数量的同时显著增强了特征提取能力。其官方版本通过更精简的参数规模实现了超越前代中型模型的精度表现。YOLO11 引入了更加高效的 C3k2 模块替代了 YOLOv8 中的 C2f 模块,这种改进通过动态调整卷积核大小与层数,不仅显著降低了计算冗余,更提升了模型对非结构化特征的抽象能力。同时,其在快速空间金字塔池(spatial pyramid pooling-fast, SPPF)层后的特征聚合机制也进行了微调,使其在处理低分辨率图像时仍能保持较高的特征保真度,这对于受限于传输带宽的管道机器人实时回传图像尤为关键。

网络结构差异与选型合理性:相比于采用 C2f 结构的 YOLOv8n, YOLO11n 优化了训练流程与特征加权策略,提高了参数利用率;相较于侧重消除 NMS 延迟但可能牺牲部分复杂场景表达力的 YOLOv10n, YOLO11n 更加注重模型生命全周期的多功能性与鲁棒性,特别是在处理管道内壁这种高噪声环境时具有更强的特征捕获潜力。尤其是面对输水管道内壁这种光照不均、反光强烈且纹理重复度高的环境, YOLO11n 展现出了比 RT-DETR 更优秀的端侧部署适应性,避免了 Transformer 架构带来的巨大算力开销,确保了嵌入式设备上的实时推理帧率。

然而,管道内部环境复杂,现有技术仍面临严峻挑战。首先,长期运行的管壁常附着大面积水垢,形成了极为复杂的背景纹理。杨明霞等^[8-9]在研究中发现,城市管道检测机器人在作业时常受此类环境干扰,且传统的卷积神经网络受限于局部感受野,缺乏对整体管壁环境的上下文理解,难以从全局角度区分相似纹理。其次,病害特征跨度大且形态各异,细微裂缝极易被背景噪声掩盖,不同类型的缺陷在尺度上差异巨大,现有的检测模型在处理跨尺度目标时往往顾此失彼。此外,传统的边界框回归损失函数,如完全交并比(complete intersection over union, CIoU),也难以对极端不规则形状进行高精度定位。针对多变的缺陷信号,张祎等^[10-11]的研究均表明单一模态或传统阈值函数存在局

限。尽管周梦颖等^[12]尝试改进 YOLOv8 进行检测,但在复杂背景下的误检率仍有待降低。郭传斌^[13]的研究也进一步证实,传统图像处理技术难以满足市政管道高精度自动化检测的需求。

针对上述挑战,本文提出了一种名为 GFF-YOLO (global context, feature fusion and focaler-loss YOLO) 的改进算法。该算法的核心思想是针对上述四大挑战进行正交化的多维优化。首先,为应对多尺度融合不充分的挑战,本文参考文献[14]使用双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)替换基线模型中的路径聚合网络(path aggregation network, PANet)颈部。BiFPN 引入了高效的双向跨尺度连接和加权特征融合机制,通过对不同分辨率的特征图进行通道对齐与权重分配,使模型能够动态地平衡高层语义信息与底层细节信息。其次,为解决全局上下文理解不足的问题,算法在 BiFPN 特征融合网络的末端、检测头(Head)的入口处嵌入了 Cao 等^[15]提出的全局上下文网络(global context network, GCNet)模块。与仅作用于主干网络不同,该模块直接对 BiFPN 输出的 P3、P4、P5 三层多尺度特征进行全局建模。GCNet 结合了 Non-local(non-local neural networks)的长距离感知能力与 SE-Net (squeeze-and-excitation network)的轻量化优势,在特征进入检测头之前建立全局依赖关系,从而有效抑制背景干扰,增强模型区分局部相似特征(如脏污与裂纹)的能力。最后,为同时解决几何缺陷与样本不平衡两大难题,本文设计了一种新颖的 Focaler-MPDIoU 混合损失函数。该损失函数“正交”地集成了两种机制,它以 Ma 等^[16]提出的多点距离交并比(minimum point distance intersection over union, MPDIoU)为几何优化基础,通过最小化预测框与真实框的角点距离来解决完全交并比(complete intersection over union, CIoU)的几何退化问题;同时结合 Zhang 等^[17] Focaler 机制作为元包装器,通过线性区间映射动态调整难易样本的损失权重,迫使模型在训练后期专注于难检缺陷。本文旨在通过这些改进,在保证实时性的前提下,显著提升输水管道缺陷检测的精度与鲁棒性。

1 改进 YOLOv11n 算法

GFF-YOLO 的改进算法整体架构如图 1 所示。

1.1 BiFPN 加权双向特征金字塔

输水管道中的缺陷尺度差异极大:细微的裂纹往往需要高分辨率的浅层特征(保留纹理细节)才能被精准捕捉,而障碍物或大面积的错位脱节则更依赖低分辨率的深层特征(包含语义信息)进行识别。虽然原始 YOLOv11 的类 PANet 结构通过双向路径增强了特征融合,但它采用简单的拼接(Concat)方式,平等地对待所有输入特征,忽略了不同分辨率特征图对特定缺陷检测贡献度的差异。

为了解决这一问题,GFF-YOLO 引入了 BiFPN 结构。BiFPN 在 PANet 的基础上进行了关键改进:通过删除只

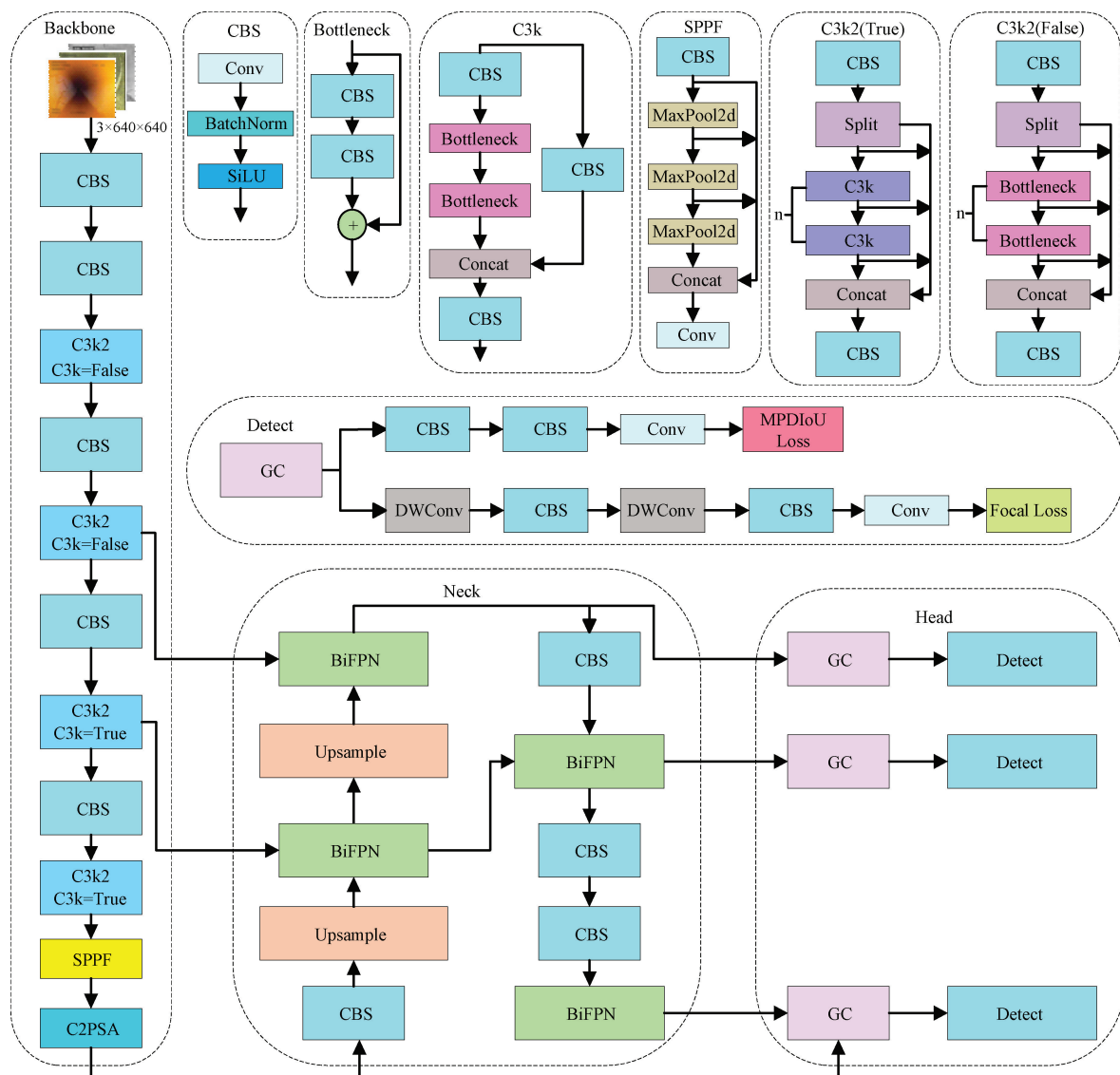


图 1 GFF-YOLO 整体架构

Fig. 1 GFF-YOLO overall architecture

有单输入节点的低效连接,并在同一层级的输入和输出节点之间增加跳跃连接,构建了更复杂的双向跨尺度融合路径。针对管道缺陷的多尺度特性,BiFPN 引入了快速归一化融合(fast normalized fusion)机制。与传统的 Softmax 融合相比,该机制在保证数值稳定性的同时,去除了昂贵的指数运算,从而提升了模型的推理速度。BiFPN 结构如图 2 所示。

BiFPN 的核心在于为每条输入边引入一个可学习的权重,让网络能够根据训练数据自动学习不同尺度特征的重要性。假设某一特征层级的输入特征列表为 $I = \{I_1, I_2, \dots, I_n\}$,对应的可学习权重为 $w = \{w_1, w_2, \dots, w_n\}$,则融合后的输出特征 O 计算公式为:

$$O = \sum_i \frac{\text{ReLU}(w_i)}{\epsilon + \sum_j \text{ReLU}(w_j)} \cdot I_i \quad (1)$$

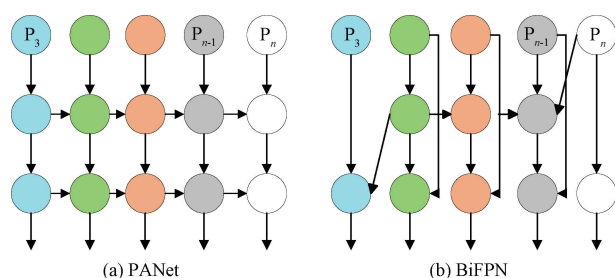


图 2 BiFPN 网络结构

Fig. 2 BiFPN network structure

式中: I_i 是不同层级的输入特征图; w_i 是学习的权重;通过 ReLU 保证了权重非负; ϵ 是一个极小值以保证数值稳定性,在 GFF-YOLO 中,将这一机制应用于所有 P3、P4、P5 节点的特征聚合处。对于微小裂纹检测,模型在训练过

程中会自动学习到 P3 层(高分辨率)对应的权重 w_{p3} 较大,从而在最终特征图中保留更多的细粒度纹理信息;而对于大面积的脏污或树根,模型则会侧重于 P5 层的深层语义信息。这种针对不同缺陷特征进行动态调整的能力,是传统 PANet 简单拼接结构所无法实现的。

1.2 GCNet 全局上下文建模

在输水管道缺陷检测任务中,准确区分“破裂”与“脏污”是一项关键挑战。由于这两类目标在局部纹理上极具相似性,仅依赖局部特征难以有效辨别。两者的核心区别在于全局上下文关系:水渍通常呈现连续的大面积分布,而裂纹则多表现为孤立的线条形态。传统的卷积神经网络(convolutional neural networks, CNN)受限于局部感受野,难以捕捉这种长距离的上下文依赖,极易因局部特征的相似而被误导。为了克服这一局限, GFF-YOLO 在 BiFPN 的末端、检测头(Head)的入口处嵌入了 GCNet 模块。本文选用的 GCNet 模块巧妙地结合了 Non-local 的长距离感知能力和 SE-Net 的轻量化优势。这使得模型在进入最终检测阶段前,能够利用全局信息有效抑制背景噪声(如大面积水渍),显著提升对裂纹等细微缺陷的识别准确率。GCNet 结构如图 3 所示。

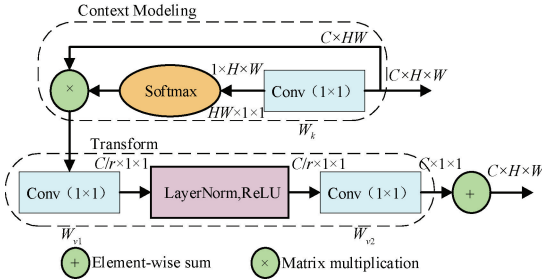


图 3 GCNet 网络结构

Fig. 3 Network structure of the GCNet

首先通过一个 1×1 卷积和 Softmax 操作生成空间注意力图,将所有位置的特征进行加权聚合,从而获得一个包含全局上下文信息的特征向量 z , 然后进行特征变换(feature transform),这是为了捕获通道间的依赖关系,利用类似 SE-Block 的瓶颈结构对 z 进行变换。由于管道缺陷特征通常较为抽象,这里引入层归一化 LayerNorm (LN)来加速收敛,变换过程包含降维和升维两个 1×1 卷积操作,并将缩放比设为 $r = 16$ 以显著降低参数量。最后进行特征融合(Fusion),采用广播(Broadcast)机制将变换后的全局上下文向量扩展至与输入特征图相同的空间尺寸。通过残差连接的方式,将其逐元素加回至原始特征中, GCNet 的数学表达为:

$$z_i = x_i + W_{v2} \cdot \text{ReLU}(\text{LN}(W_{v1} \cdot \sum_{j=1}^{N_p} \frac{e^{W_k x_j}}{\sum_{m=1}^{N_p} e^{W_k x_m}} x_j)) \quad (2)$$

式中: x_i 表示第 i 个位置的输入特征向量; z_i 为融合后的输出特征; N_p 为特征图的像素总数; W_k 、 W_{v1} 和 W_{v2} 分别为对应的卷积权重参数。通过上述机制,输出特征图中的每一个像素点都成功融合了整个管道内壁的全局背景信息。这使得模型在判断局部纹理时能够参考全局环境,从而构建出一种查询无关(query-independent)的全局注意力。该方法特别适用于管道内部这种背景纹理相对固定的场景,能够有效抑制因局部相似性导致的水垢误检,显著提升检测的鲁棒性。

1.3 Focaler-MPDIoU 损失函数

输水管道内部环境的特殊性给目标检测的边界框回归带来了双重挑战。首先,缺陷形态极其不规则,特别是裂纹通常呈现大长宽比的细长形态,而水垢等脏污则边缘模糊。传统的损失函数(如 CIoU Loss)主要依赖重叠面积和中心点距离进行优化,在面对这种高宽比极端或重叠度高但形状不匹配的情况时,容易出现定位漂移,难以实现精确定位。其次,管道图像中存在严重的样本不平衡问题,大量易于识别的简单背景样本(如正常管壁接口)在训练中占据主导地位,导致模型对少量困难样本(如隐蔽的微小裂纹、复杂异物)的关注度不足,梯度下降方向易被简单样本淹没。

为了克服上述局限性,本文提出一种融合动态加权与几何优化的 Focaler-MPDIoU 损失函数。它结合了 MPDIoU 在复杂几何形状回归上的优势和 Focaler 机制在样本难易均衡上的能力。

1) MPDIoU 几何优化

针对管道裂纹和不规则缺陷的定位难题, GFF-YOLO 引入了 MPDIoU。与传统方法不同, MPDIoU 创新性地引入了基于最小点距离的度量方式,直接最小化预测框与真实框左上角及右下角点之间的欧氏距离。这不仅简化了计算过程,还能更准确地衡量边界框的匹配程度,特别是在处理具有相同高宽比但位置微偏的缺陷目标时表现更佳。

设真实框面积为 A , 预测框面积为 B , 输入图像宽、高为 w, h 。真实框左上、右下角点坐标为 $(x_1^{\text{gt}}, y_1^{\text{gt}}), (x_2^{\text{gt}}, y_2^{\text{gt}})$, 预测框对应坐标为 $(x_1^{\text{prd}}, y_1^{\text{prd}}), (x_2^{\text{prd}}, y_2^{\text{prd}})$, 则左上角点欧氏距离 d_1^2 和右下角点欧氏距离 d_2^2 定义为:

$$d_1^2 = (x_1^{\text{prd}} - x_1^{\text{gt}})^2 + (y_1^{\text{prd}} - y_1^{\text{gt}})^2 \quad (3)$$

$$d_2^2 = (x_2^{\text{prd}} - x_2^{\text{gt}})^2 + (y_2^{\text{prd}} - y_2^{\text{gt}})^2 \quad (4)$$

MPDIoU 的计算公式定义为:

$$\text{MPDIoU} = \text{IoU} - \frac{d_1^2}{w^2 + h^2} - \frac{d_2^2}{w^2 + h^2} \quad (5)$$

对应的 MPDIoU 损失函数为:

$$L_{\text{MPDIoU}} = 1 - \text{MPDIoU} = 1 - \text{IoU} + \frac{d_1^2}{w^2 + h^2} + \frac{d_2^2}{w^2 + h^2} \quad (6)$$

这种直接优化角点的策略对于管道检测至关重要:迫

使预测框的两个关键角点向真实边界收缩。在回归细长、弯曲的管道裂纹时,能够避免预测框虽然包住了裂纹但包含了大量无效背景的问题,实现比 CIoU 更紧密的边界贴合。

2)Focaler 动态加权

针对管道缺陷检测中严重的样本不平衡问题,引入 Focaler 机制。该机制通过线性区间映射重构 IoU,提升难易样本的区分度。根据样本的回归难度(以 IoU 值衡量),自适应地调整损失权重,Focaler-IoU 的核心逻辑在于利用两个超参数;困难样本下界阈值 l 和简单样本上界阈值 u ,对原始 IoU 进行非线性截断与线性拉伸的复合重构。其重构公式为:

$$\text{IoU}^{\text{focaler}} = \text{clip}\left(\frac{\text{IoU} - l}{u - l}, 0, 1\right) \tag{7}$$

在噪声截断区 $\text{IoU} < l$ 时,当预测框与真实框的重叠度极低时,映射值被强制归零。这有效过滤了管道内壁复杂纹理产生的无效噪声梯度,避免了极难样本对训练稳定性的干扰。在梯度聚焦拉伸区 $l \leq \text{IoU} \leq u$ 时,在该关键重构区间内,原本细微的 IoU 差异被线性放大至张 $[0, 1]$ 范围。这种“拉伸”增加了损失函数对区间内样本的敏感度,迫使模型在训练后期精准捕捉那些已有初步定位但精度尚缺的微小缺陷。

而当样本回归精度达到 $\text{IoU} > u$ 梯度饱和区时,当回归精度达到阈值 u 后, $\text{IoU}^{\text{focaler}}$ 进入饱和区并稳定为 1。此时该样本产生的梯度趋于消失,从而解决了训练后期海量简单样本淹没难检样本梯度的问题。经过网格搜索实验确定,当 $l = 0.2, u = 0.8$ 时,模型能最大程度地将训练重心锁定在 $[0.2, 0.8]$ 这一关键重构区间内,通过对该区间内 IoU 差异的线性放大,迫使梯度下降方向在训练后期向难以定位的微小缺陷倾斜,从而提升整体召回率。

3)Focaler-MPDIoU 融合

将 MPDIoU 的几何优化特性与 Focaler 的动态加权机制相结合,得到最终的 Focaler-MPDIoU 损失函数:

$$L_{\text{Focaler-MPDIoU}} = L_{\text{MPDIoU}} + \text{IoU} - \text{IoU}^{\text{Focaler}} \tag{8}$$

即利用经过 Focaler 重构后的 $\text{IoU}^{\text{Focaler}}$ 替换 MPDIoU 公式中的标准 IoU 项进行计算。如图 4 所示。该改进策略实现了双重增益:一方面,MPDIoU 利用角点距离约束提升了对不规则缺陷边界的回归精度;另一方面,Focaler 机制根据样本难易程度动态调整梯度下降的权重。两者的结合显著增强了 GFF-YOLO 在复杂管道环境下对小目标(困难样本)和密集目标的检测鲁棒性。

2 实验环境与结果分析

2.1 实验环境和参数设置

本研究的所有实验均在统一的软硬件环境下进行,以确保公平对比。检测训练平台设置的实验环境如表 1 所示。

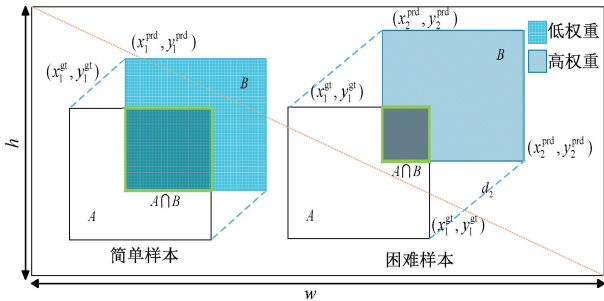


图 4 Focaler-MPDIoU 边框处理过程

Fig. 4 Focaler-MPDIoU bounding box processing process

表 1 实验环境

Table 1 Experimental environment	
类别	配置详情
处理器 (CPU)	Intel(R) Xeon(R) CPU E5-2686 v4
图形处理器 (GPU)	NVIDIA RTX A4000 (16 GB)
内存 (RAM)	60 GB
操作系统	Ubuntu20.04
编程语言	Python 3.8
深度学习框架	PyTorch 2.4
CUDA 版本	CUDA 11.6

为保证实验的可靠性与可复现性,所有模型(包括基线模型和改进模型)均采用相同的超参数设置进行训练。所有模型均加载 yolov11n.pt 预训练权重进行迁移学习,以加快收敛速度并提升模型性能。关键训练参数如表 2 所示。

表 2 超参数设置

Table 2 Hyperparameter settings	
超参数	数值
输入图像尺寸	640×640
训练周期	300 次
批量大小	16 批
优化器	SGD (随机梯度下降)
初始学习率	0.01
学习率动量	0.937
权重衰减	0.000 5
数据增强	mosaic: 1.0

2.2 数据集

本文数据集的构建结合了两个来源:自建市政管网数据集和公开数据集 Sewer-ML。自建数据集:来源于某市 CCTV 检测视频,经关键帧提取、去模糊后,参照 PACP 标准进行标注。包含 1 909 张原始图像。Sewer-ML 扩充:引入部分 Sewer-ML 数据以增加样本多样性,特别是不同管材(混凝土、镀锌钢管、球铸墨铁)的背景纹理。数据增强:

考虑到管道内部光照条件极差,采用 CLAHE(限制对比度自适应直方图均衡化)增强图像局部对比度,突出裂纹特征;同时使用 Mosaic 增强,模拟遮挡情况,提升模型对部分可见缺陷的检出率。最终数据集扩充至 4 024 张,各类别的具体样本数量分布为破裂(1 001 张)、错位(912 张)、脏污(1 196 张)、障碍物(915 张),最终按 8:1:1 的比例随机划分为训练集(3 220 张)、验证集(402 张)和测试集(402 张)。为了避免时序数据泄露,数据集划分严格按照视频片段段进行隔离,确保训练集和测试集来自不同的管道视频段。部分数据集示例如图 5 所示。

2.3 消融实验分析

为验证本文所提出的各项改进策略的有效性,进行了详细的逐级消融实验。实验以原始 YOLOv11n 为基线模型(Baseline, M1),在完全相同的超参数配置和训练环境下,逐步引入 GCNet(全局上下文模块)、BiFPN(加权双向特征金字塔)和 Focaler-MPDIoU(混合损失函数)。实验旨在定量分析各模块对模型复杂度(Params, FLOPs)及检测性能(Precision, Recall, mAP)的边际贡献。结果如表 3 所示。

从实验结果来看,单独引入 GCNet(M2)后,mAP₅₀ 提升了 3.34%,证明了全局上下文建模能有效抑制水垢等背景噪声,单独引入 Focaler-MPDIoU(M4)带来了超过 5% 的 mAP₅₀ 提升。这在一般目标检测任务中是罕见的,但在

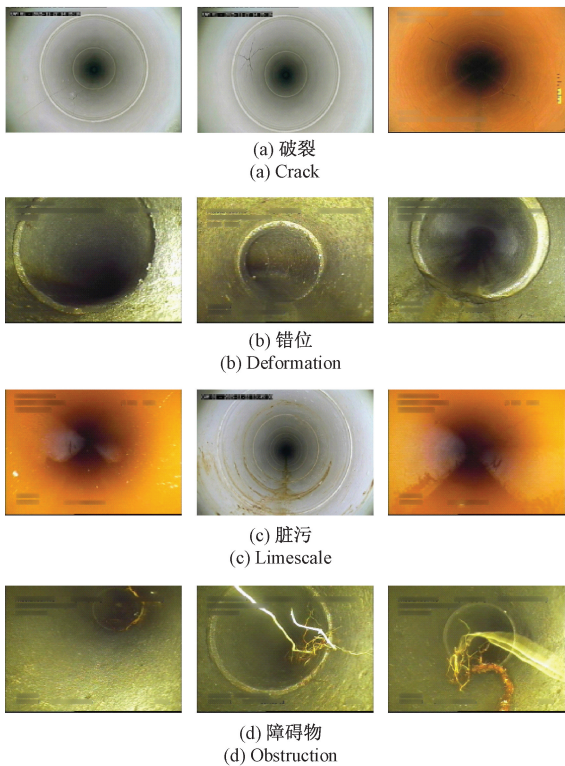


图 5 样本示例
Fig. 5 Sample examples

表 3 消融实验结果
Table 3 Ablation study results

模型	YOLOv11n	GCNet	BiFPN	Focaler-MPDIoU	Param/ ($\times 10^6$)	GFLOPs	P/%	R/%	mAP ₅₀ /%	mAP ₅₀₋₉₅ /%
M1	✓	—	—	—	2.582	6.3	83.07	72.71	81.89	54.16
M2	✓	✓	—	—	2.610	6.4	85.84	81.6	85.23	63.25
M3	✓	—	✓	—	2.670	7.1	77.73	73.03	82.46	51.06
M4	✓	—	—	✓	2.582	6.3	89.00	76.34	86.95	60.52
M5	✓	✓	—	✓	2.610	6.4	84.55	79.97	87.11	59.37
M6	✓	—	✓	✓	2.670	7.1	87.65	77.64	87.55	60.68
M7	✓	✓	✓	✓	2.510	7.6	91.47	83.86	91.40	66.99

注:“✓”表示使用了该模块,“—”表示未执行该操作

管道检测中却合乎逻辑。因为管道缺陷(如裂纹)的几何形态与 COCO 数据集中的物体(如人、车)完全不同,传统的基于矩形假设的 CIoU 在此处失效严重,而 MPDIoU 的角点约束恰好弥补了这一短板,当组合使用这些模块时,模型性能呈现出叠加效应。最终集成了所有改进策略的 GFF-YOLO(M7)取得了 91.40% 的最佳 mAP₅₀ 和 66.99% 的最佳 mAP₅₀₋₉₅,相较于基线模型,mAP₅₀ 和 mAP₅₀₋₉₅ 分别显著提升了 9.51% 和 12.83%。

值得注意的是,GFF-YOLO(M7)在大幅提升精度的同时。其模型的参数量略低于基线,这主要归因于 BiFPN 结构移除了 PANet 中冗余的单输入节点,抵消了 GCNet 带来的参数增加,但计算量 7.6 GFLOPs 相较于基线模型

6.3 GFLOPs 有所增加,但这种“以适度计算资源换取高精度”的策略在工业检测中是极具价值的,特别是考虑到其对微小裂纹和复杂脏污的检出率实现了质的飞跃。

2.4 SOTA(State-of-the-Art) 对比实验

为进一步评估 GFF-YOLO 的综合竞争力,本节将其与当前主流的目标检测模型进行横向对比,包括轻量级的 YOLOv3-Tiny, YOLOv5n, YOLOv8n, YOLOv10n 以及经典的二阶段检测器 Faster R-CNN。所有模型均在同一数据集和软硬件平台上进行重新训练与测试,结果如表 4 所示。

实验数据表明,GFF-YOLO 在检测精度上实现了全面的 SOTA 表现,GFF-YOLO 的 mAP₅₀ 高达 91.40%,大

表 4 对比实验结果

Table 4 Comparative experiment results

模型	Params/($\times 10^6$)	GFLOPs	mAP ₅₀ /%	mAP ₅₀₋₉₅ /%	Inference time/ms
YOLOv3-Tiny	9.52	14.40	70.80	43.40	10.40
YOLOv5n	2.18	5.90	63.00	35.80	13.50
YOLOv8n	2.68	6.90	71.20	43.60	12.30
YOLOv9t	2.56	7.10	74.60	45.10	15.20
YOLOv10n	2.30	6.70	78.80	50.90	20.00
RT-DETR-N	4.10	10.50	76.50	48.30	22.30
Faster R-CNN	28.31	454.78	63.52	45.71	77.20
GFF-YOLO	2.51	7.60	91.40	66.99	17.40

幅领先于所有对比模型。即便是与最新的 YOLOv10n 相比，其 mAP₅₀ 和 mAP₅₀₋₉₅ 也分别高出了 12.60% 和 16.09%。这充分说明通用的轻量化模型在面对管道内部这种极端纹理和不规则缺陷时存在瓶颈，而 GFF-YOLO 的针对性改进有效突破了这一限制。在满足工业级实时检测需求(30 fps)的前提下，大幅提升了检测精度。与基于 Transformer 架构的 RT-DETR-N 相比，GFF-YOLO 在参数量减少约 38.7% 的情况下，mAP₅₀ 领先其 14.9%，且推理速度提升了 21.9%。在与传统的二阶段大模型 Faster R-CNN 相比时，GFF-YOLO 展现出压倒性优势：仅消耗其 8.86% 的计算量，便实现了 4.4 倍的推理速度，且 mAP₅₀ 反超 27.88%。这得益于 BiFPN 的快速归一化融合机制，在极大地降低内存访问成本的同时，实现了对

复杂几何缺陷的高效捕获。

综上所述，GFF-YOLO 在输水管道缺陷检测任务中，实现了精度与速度的最佳平衡。

2.5 模型检测效果

从图 6(a)和(d)可以看出，随着迭代轮次增加，训练集和验证集的损失值均呈现平稳下降趋势，可以看出，改进算法在第 50 轮左右 mAP 提升速率明显优于基线，说明 Focaler 机制有效加速了模型对难例的学习，且未出现震荡或发散。从图 6(e)可以看出 mAP 指标在第 180 轮左右趋于稳定，最终从图 6(f)看出 mAP₅₀₋₉₅ 稳定在 66.99%。这表明模型具有良好的收敛特性和泛化能力，未出现明显的过拟合现象。

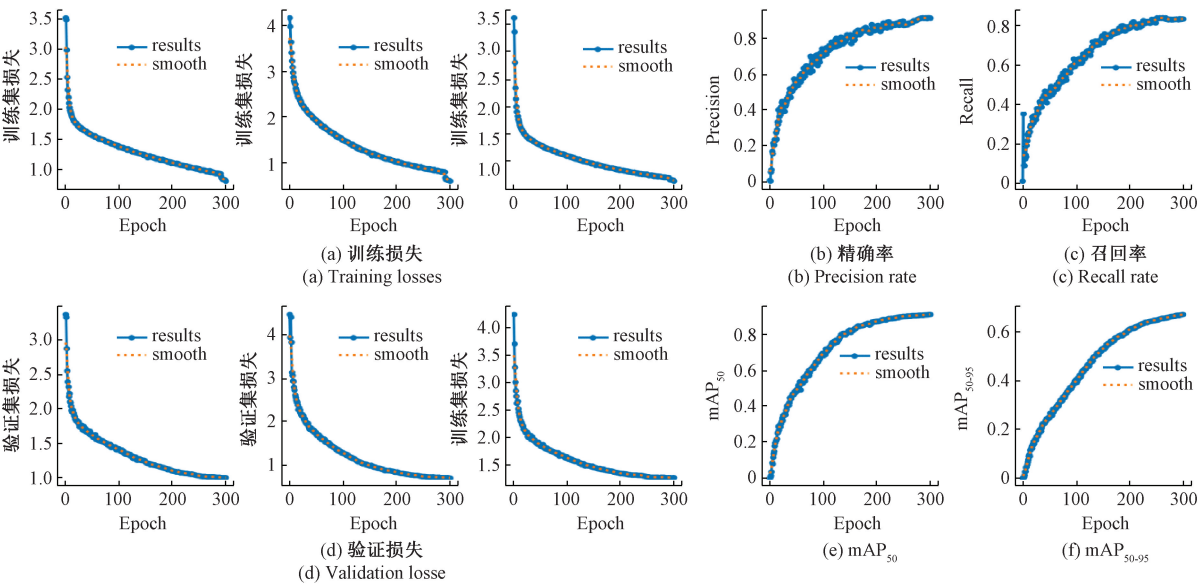


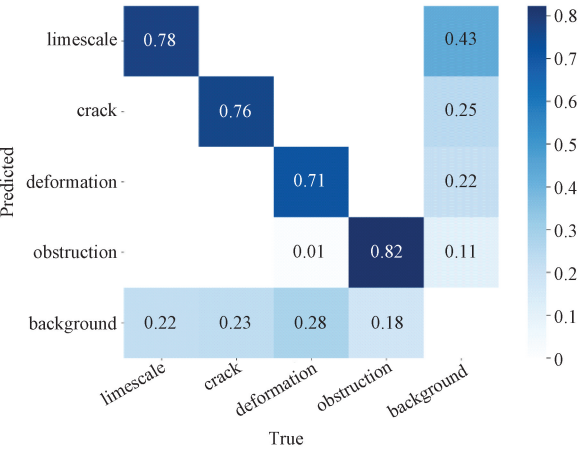
图 6 GFF-YOLO 训练结果

Fig. 6 GFF-YOLO training results

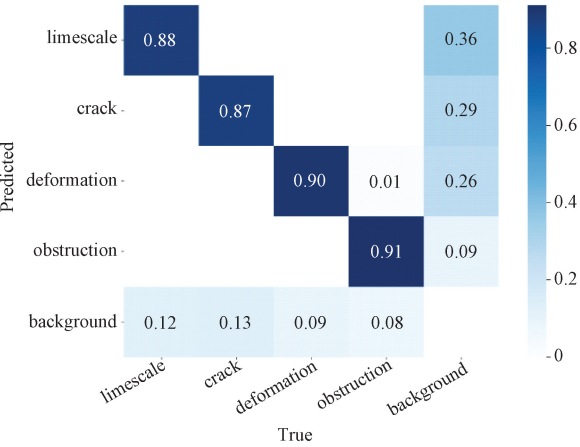
为进一步验证改进后 GFF-YOLO 模型的有效性，即能对复杂管道环境下的 4 类典型缺陷进行良好的区分和检测，在自建的管道缺陷数据集上进行了相关实验。对基线模型和改进后的 GFF-YOLO 模型进行训练，得到对应

的归一化混淆矩阵图，如图 7 所示。进行可视化描述对比分析。其中，对角线上的值表示的是模型预测正确的比例即召回，数值越高表明模型训练结果越好。混淆矩阵的最后 1 行表示训练时各类别的漏检率，最后 1 列表示错检率，

这两项的数值越小表明模型训练越好。如图 7(a)所示,基线模型在破裂与背景之间的混淆度较高,说明模型难以区分细微裂纹与相似的水垢纹理。而在图 7(b)中,改进后的 GFF-YOLO 模型成功将这一类别的误检率显著降低。这直接佐证了 GCNet 模块的有效性,通过建立全局长距离依赖,模型不再仅依靠局部纹理特征,而是结合全局环境判断目标的真实属性,从而有效解开了裂纹与水垢的混淆。同时,对于形态较长的错位缺陷,得益于 MPDIoU 的角点优化,其召回率也得到了明显提升。以上结论均表明,本文所引入的改进模块对不同类型的管道缺陷检测均有着较好的效果,能做到对 4 类缺陷的良好区分,达到了预想的效果。



(a) 基线模型效果
(a) Baseline model effect



(b) GFF-YOLO模型效果
(b) GFF-YOLO Model performance

图 7 归一化混淆矩阵

Fig. 7 Normalized confusion matrix

2.6 可视化与定性分析

为了更直观地展现 GFF-YOLO 算法相较于基线模型 YOLOv11n 的改进效果,本文选取了典型检测场景进行可视化定性分析,随机选取了未经训练的个别数据集图片作

为数据集,使用 GFF-YOLO 算法数据集进行缺陷检测如图 8 所示。

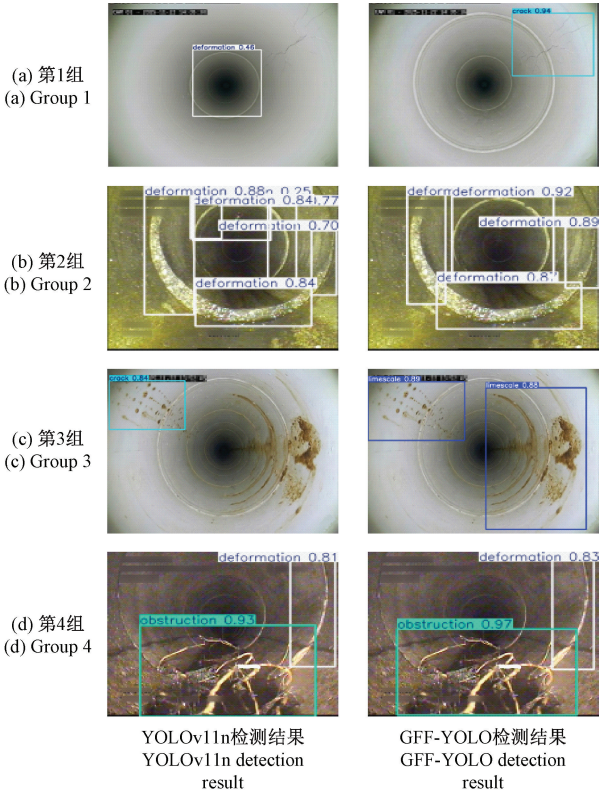


图 8 不同算法可视化对比

Fig. 8 Visual comparison of different algorithms

针对复杂背景的误检分析,如图 8(c)所示,管道内壁的水垢(limescale)纹理与裂纹(crack)高度相似。基线模型由于缺乏全局上下文理解,易将大面积水垢误检为“破裂”;而 GFF-YOLO 依靠 GCNet 模块的全局注意力机制,成功抑制了背景噪声干扰,避免了误检。针对微小尺度目标的漏检分析,如图 8(a)所示,基线模型的 PANet 结构对微小裂纹的特征提取能力不足,导致了目标漏检;GFF-YOLO 凭借 BiFPN 高效的双向特征融合能力,成功捕获并检测到了该微小缺陷。针对不规则形态的定位精度分析,如图 8(b)所示,面对形态不规则的障碍物,基线模型的边界框回归(CIoU)定位松散;GFF-YOLO 在 Focaler-MPDIoU 损失的优化下,边界框更紧密地贴合了缺陷边缘,IoU 质量更高,定位更精准。

综上所述,可视化结果与前文的定量分析(mAP)高度一致,有力地证明了 GFF-YOLO 在复杂管道环境下的检测鲁棒性(Robustness)和高精度。

3 结 论

针对传统检测效率低及现有模型在复杂管道环境下精度受限的问题,本文提出基于 YOLOv11n 的 GFF-YOLO 算法。首先,针对水垢与裂纹极易混淆的难点,引

入 GCNet 建立全局依赖,赋予模型全局视野,显著增强了对背景噪声的抑制能力。其次,利用 BiFPN 重构颈部网络,通过加权双向融合解决信息传递效率低的问题,大幅提升了微小及巨型目标的综合表征。此外,设计 Focaler-MPDIoU 混合损失函数,融合角点优化与动态聚焦机制,有效解决了异形缺陷定位不准及样本失衡问题。实验表明,该模型 mAP₅₀ 达 91.40%,较基线提升 9.5%,在保证实时性的同时实现了高精度检测,该研究成果不仅验证了各项改进策略的有效性,也为城市地下管网的智能化、自动化运维提供了一种具备高工程应用价值的可靠技术方案。

尽管 GFF-YOLO 模型取得了良好效果,但仍存在局限性。未来的研究工作可从以下方面展开:如易混淆类别,混淆矩阵分析显示,模型对破裂和脏污等视觉特征相似的类别仍存在少量混淆。未来可尝试引入更细粒度的注意力机制或特定的数据增强,以提升模型的辨识能力。

参考文献

- [1] 袁海伟, 胡露, 钟涛, 等. 铸铁供水管道沉积物污染富集特征[J]. 中国给水排水, 2025, 41(11): 86.
Yuan Haiwei, Hu Lu, Zhong Tao, et al. Pollution enrichment characteristics of sediments in cast iron water supply pipelines[J]. China Water & Wastewater, 2025, 41(11): 86.
- [2] 郭育硕, 韩小蒙, 舒诗湖. 金属供水管道内壁腐蚀行为与多因素预测模型[J]. 中国给水排水, 2025, 41(11): 25.
Guo Yushuo, Han Xiaomeng, Shu Shihu. Corrosion behavior of inner wall of metal water supply pipeline and multi-factor prediction model[J]. China Water & Wastewater, 2025, 41(11): 25.
- [3] 杨超, 孙虎, 唐超. 电磁脉冲涡流检测下金属管道缺陷检测研究[J]. 电子测量与仪器学报, 2025, 39(4): 133.
Yang Chao, Sun Hu, Tang Chao. Research on defect detection of metal pipelines based on electromagnetic pulse eddy current testing[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(4): 133.
- [4] 林哲, 潘慧琳, 陈丹. 融合改进 YOLO 和语义分割的遮挡目标抓取方法[J]. 电子测量与仪器学报, 2024, 38(12): 191-194.
Lin Zhe, Pan Huilin, Chen Dan. Occluded target grasping method fusing improved YOLO and semantic segmentation[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(12): 191-194.
- [5] 李春晖, 吕岩, 孟祥来, 等. 基于 YOLOv8 的油气管道漏磁检测缺陷智能识别技术研究[J]. 仪器仪表学报, 2025, 46(2): 252-253.
Li Chunhui, Lyu Yan, Meng Xianglai, et al. Research on intelligent recognition technology for magnetic flux leakage detection defects of oil and gas pipelines based on YOLOv8 [J]. Chinese Journal of Scientific Instrument, 2025, 46(2): 252-253.
- [6] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [7] Wang N N, Duan Y H, Li S T, et al. Advanced YOLO-DeepSort-based system for drainage pipeline defects intelligent detection[J]. IEEE Access, 2024, 12: 145857-145870.
- [8] 杨明霞. 城市排水管道检测机器人研究进展[J]. 广东建材, 2023, 39(8): 133.
Yang Mingxia. Research progress of urban drainage pipeline inspection robots [J]. Guangdong Building Materials, 2023, 39(8): 133.
- [9] 罗来鸣, 王婷. 基于优化 YOLO 算法的管道裂缝识别研究[J]. 给水排水, 2023, 59(S2): 354-356.
Luo Laiming, Wang Ting. Research on pipeline crack identification based on optimized YOLO algorithm[J]. Water & Wastewater Engineering, 2023, 59 (S2): 354-356.
- [10] 张祎, 李素贞. 基于声-压信号融合的城市供水管道泄漏监测[J]. 仪器仪表学报, 2024, 45(8): 69-70.
Zhang Yi, Li Suzhen. Leakage monitoring of urban water supply pipelines based on acoustic-pressure signal fusion [J]. Chinese Journal of Scientific Instrument, 2024, 45(8): 69-70.
- [11] 徐自强, 李成, 穆连波, 等. 基于改进阈值函数的直埋供热管网检漏研究[J]. 电子测量技术, 2024, 47(6): 149.
Xu Ziqiang, Li Cheng, Mu Lianbo, et al. Research on leak detection of directly buried heating pipe networks based on improved threshold function [J]. Electronic Measurement Technology, 2024, 47(6): 149.
- [12] 周梦颖, 张学武, 曾鹏源, 等. 基于改进 YOLOv8 的排水管网缺陷检测[J]. 电子测量技术, 2025, 48(3): 129-132.
Zhou Mengying, Zhang Xuewu, Zeng Pengyuan, et al. Defect detection of drainage pipe network based on improved YOLOv8 [J]. Electronic Measurement Technology, 2025, 48(3): 129-132.
- [13] 郭传斌. 基于图像处理技术的市政管道质量缺陷检测[D]. 武汉: 华中科技大学, 2024: 71-72.
Guo Chuanbin. Quality defect detection of municipal pipelines based on image processing technology [D]. Wuhan: Huazhong University of Science and Technology, 2024: 71-72.

[14] Tan M X, Pang R M, Le Q V. EfficientDet: Scalable and efficient object detection [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 10782-10785.

[15] Cao Y, Xu J R, Lin S, et al. GCNet: Non-local networks meet squeeze-and-excitation networks and beyond [C]//IEEE International Conference on Computer Vision (ICCV), 2019:1975-1978.

[16] Ma S L, Xu Y. MPDIoU: A loss for efficient and accurate bounding box regression [PP/OL]. V1. (2023-07-14) [2026-05-24]. <https://doi.org/10.48550/arXiv.2307.07662>

[17] Zhang H, Zhao S J. Focaler-IoU: More focused intersection over union loss[PP/OL]. V1. (2024-01-19) [2026-04-20]. <https://doi.org/10.48550/arXiv.2401.1052>.

作者简介

张杨, 硕士研究生, 主要研究方向为输水管道内壁缺陷检测与脏污清洗机器人。

E-mail: 1141975648@qq.com

秦昌浩, 硕士研究生, 主要研究方向为计算机视觉、辣椒播种机器人。

E-mail: 1264085396@qq.com

孔硕, 硕士研究生, 主要研究方向为计算机视觉、小径木砍伐机器人。

E-mail: 2864553125@qq.com

曲娅琨, 硕士研究生, 主要研究方向为计算机视觉、机器人技术。

E-mail: 2509677757@qq.com

赵辉(通信作者), 博士, 副教授, 硕士生导师, 主要研究方向为计算机视觉、智能制造。

E-mail: zhaohv@163.com