

基于 YOLO 轻量化的水下管桩裂缝检测方法^{*}陈潇威¹ 宋瑞银^{1,2} 吴瑞明¹ 李凤姝² 王天恒²

(1. 浙江科技大学机械与能源工程学院 杭州 310018; 2. 浙大宁波理工学院机电与能源工程学院 宁波 315100)

摘 要: 基于机器视觉的水下管桩裂缝检测在工作时往往面临着嵌入式设备计算资源有限、实时检测速度慢的问题,为了解决这些问题本文提出了基于管桩清洗机器人的管桩裂缝自动识别方法。设计了一种轻量化网络检测算法 YOLOv8-MLLA-Mobilenetv4-WIoU(MWM-YOLO)。在水下环境拍摄获取低质缺陷图像并进行数据增强,扩充数据集;针对浑水下低质图像,针对图像增强与目标检测不匹配所导致的抑制作用,利用 MLLA 自注意力精准聚焦关键特征区域,在保持高分辨率输出的同时,有效抑制背景干扰,从而增强图像增强与目标检测的协同作用。同时采用最新的 Mobilenetv4 主干网络,降低特征网络的参数量和计算量。在此基础上,考虑低质图像数据标注难免包含低质量示例,使用 WIoU 损失函数替换原 YOLOv8 网络模型中的损失函数,提高模型泛化性能。实验结果表明:MWM-YOLO 模型权重大小为 14.9 MB,较原模型减少了 30.3%。平均精度达到了 89.1%,推理速度为 137.54 fps,优于其他模型。改进后的网络模型相比原网络,在保持缺陷识别精度的同时,可以轻量化部署到边缘计算设备,为水下管桩清洗机器人提供技术支持。

关键词: 管桩清洗机器人;管桩裂缝检测;YOLOv8s;缺陷识别;注意力机制

中图分类号: TV36;TN0 **文献标识码:** A **国家标准学科分类代码:** 510.4

Crack detection method of underwater pipe pile based on YOLO lightweight

Chen Xiaowei¹ Song Ruiyin^{1,2} Wu Ruiming¹ Li Fengshen² Wang Tianheng²

(1. School of Mechanical and Energy Engineering, Zhejiang University of Science and Technology, Hangzhou 310018, China;

2. School of Mechatronics and Energy Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, China)

Abstract: In order to solve these problems, this paper proposes an automatic identification method for pipe pile cracks based on pipe pile cleaning robots. A lightweight network detection algorithm YOLOv8-MLLA-Mobilenetv4-WIoU (MWM-YOLO) was designed. Capture low-quality defect images in a muddy water environment and augment the data to expand the dataset. For low-quality images under muddy water, in view of the suppression effect caused by the mismatch between image enhancement and object detection, MLLA is used to accurately focus on key feature areas, which can effectively suppress background interference while maintaining high-resolution output, so as to enhance the synergy between image enhancement and object detection. At the same time, the latest Mobilenetv4 backbone network is used to reduce the number of parameters and calculations of the characteristic network. On this basis, considering that low-quality image data annotation inevitably contains low-quality examples, the WIoU loss function is used to replace the loss function in the original YOLOv8 network model to improve the generalization performance of the model. The experimental results show that the weight of the MWM-YOLO model is 14.9 MB, which is 30.3% less than that of the original model. The average accuracy reached 89.1%, and the inference speed was 137.54 fps, which was better than other models. Compared with the original network, the improved network model can be lightweight deployed to edge computing devices while maintaining the accuracy of defect identification, providing technical support for underwater pipe pile cleaning robots.

Keywords: pipe pile cleaning robot; crack detection of pipe piles; YOLOv8s; defect identification; attention mechanisms

0 引 言

裂缝是管桩常见的质量问题之一,特别是在承载力较大

的情况下,裂缝可能会导致管桩的结构强度下降^[1-3]。然而,水下管桩裂缝检测同样面临着许多挑战,当前常用的检测手段主要以潜水员水下肉眼观测或携带器材探摸为主,潜水员

往往因为海洋复杂环境干扰,检测危险系数较大,检测效率低,检测结果可靠性较差,检测成本较高。此外,因此研究一种智能化,快捷方便,低成本且能够克服水下苛刻环境影响的管桩水下结构缺陷识别技术有着十分重要意义^[4-5]。

近年来,无人机与水下机器人的发展给管桩裂缝检测带来了新的选择。其中,基于机器人的管桩裂缝陆上检测和识别方法的已经较为成熟。在深度学习快速发展的背景下,李培坤等^[6]优化YOLOv8n算法用于水下目标检测算法,为水下检测提供新的思路。李泽荣等^[7]改进YOLO算法实现了桥梁水下结构缺陷自动识别。Burguera等^[8]使用Siamese卷积神经网络的轻量级实现水下视觉环检测和分类。张云泽^[9]通过分析桥梁墩桩所处水下环境特征与裂缝特征,提出了一种新的基于深度学习的桥梁水下墩桩裂缝检测方法为工程师提供更加准确、详细的桥梁墩桩裂缝损害信息。Antoni等^[10]提出了一种针对水下环境中鲁棒、快速视觉环路检测的神经网络。

水下操控机器人实现裂缝识别相比于陆上识别具有更大的挑战性。目前,水下图像检测可以根据不同类型的感知技术进行分类,常见的分类包括光学成像、声学成像,以及其他新兴的水下探测技术。其中声学成像通过声呐设备发射声波,声波遇到物体后反射回来,通过分析反射波来构建水下环境或目标的图像,但是声学成像具有局限性,分辨率较低,生成的图像可能模糊,细节较少,难以提供如光学成像那样丰富的视觉信息。声波的反射也容易受到水体中的物质或其他干扰物影响。光学成像通过水下摄像机或相机采集图像,依赖自然光或人工照明。但是光线在水中衰减较快,受限于浑浊度、散射、吸收等因素。深水区域光线不足,颜色失真严重,图像质量下降。针对水下低质图像的处理方法主要为图像增强与图像复原两大类。田青等^[11]提出了一种Enhanced Zero-DCE++算法,改善水下选通图像光圈部分过曝光的问题,在小波域中区分噪声和真实特征,增强特征区分能力,提高成像清晰度。袁国铭等^[12]提出了纹理感知联合颜色直方图特征的水下图像增强网络,包含纹理感知网络,颜色直方图特征提取网络以及颜色纹理融合的水下图像增强网络。王宁等^[13]提出一种基于深度估计和梯度下降策略的水下图像恢复与增强框架(UIRENet)。Ziad等^[14]引入了一种新的非局部模型在噪声干扰下使得图像复原。

综上所述,水下结构检测相比于陆上结构检测还存在许多的局限性。清晰图像难以采集,设备受动态环境变化的限制,缺陷样本数据有限,识别缺陷能力有限,同时部署到边缘计算设备困难。因此本文改进YOLOv8算法,引入MLLA自注意力机制,聚焦裂缝的局部特征、抑制噪声干扰、适应复杂形状和光照变化。同时基于MobileNetV4进行主干网络改进,适合在移动端和水下设备部署,在提高精度的同时,保持模型的高效、轻量化特性。最后通过改进损失函数,减少低质量数据标注示例的干扰,提高模型泛化性

能,为水下结构智能检测提供便利。

1 图像采集与预处理

1.1 实验数据来源

本文通过团队自主研发的管桩清洗机器人如图1(a)所示,在实验室水池中的水下管桩如图1(b)所示进行浑水下拍摄获取缺陷图像,图1(c)所示为数据采集结构。机器人攀附在管桩上,电机驱动上下爬行的过程中,底层的清洗装置开始围绕管桩旋转清洗,在清洗装置上的摄像头开始拍摄。针对水下环境黑暗的问题,加装照明灯。同时通过数据传输线将图像传输到移动设备端和计算机上进行图像识别任务。在通过拍摄视频后逐帧截取照片,得到1200张图像。

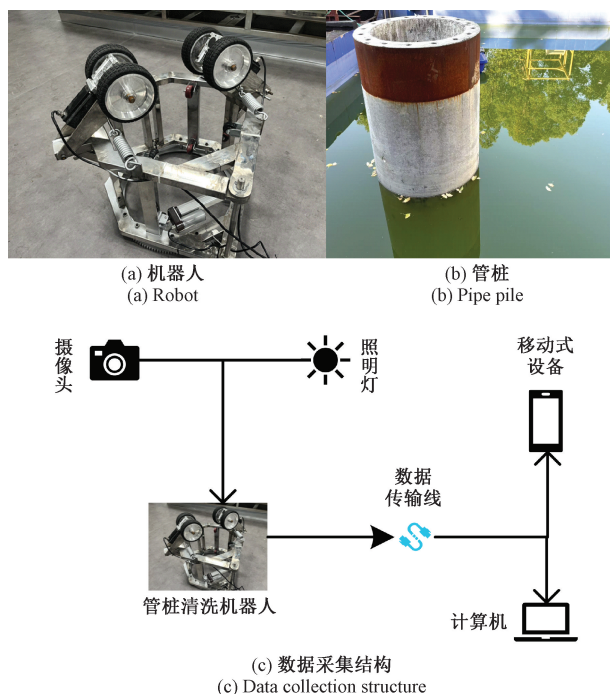


图1 数据采集场景与结构

Fig. 1 Data collection scenario and structure

管桩结构常见缺陷为裂缝。最后对采集图像数据进行处理,制作成目标检测数据集,如图2所示。



图2 管桩裂缝数据集

Fig. 2 Pipe pile crack dataset

1.2 数据预处理

YOLO 深度卷积神经网络模型依赖于特征学习方法，其依靠于大量多样的数据进行训练。由于海洋环境的复杂性，以及机器人在爬行过程中受力不平衡等不可控因素，为了模型能够适应现实中复杂的海洋场景，增强管桩裂缝数据集的多样性和质量。本研究对图像进行预处理，并且使用数据增强技术，具体包括平移、旋转、缩放、亮度调整、随机遮挡以及添加椒盐噪声等操作，增加了数据集的复杂度和覆盖范围。预处理后，最终获得 3 000 张图像，如图 3 所示。

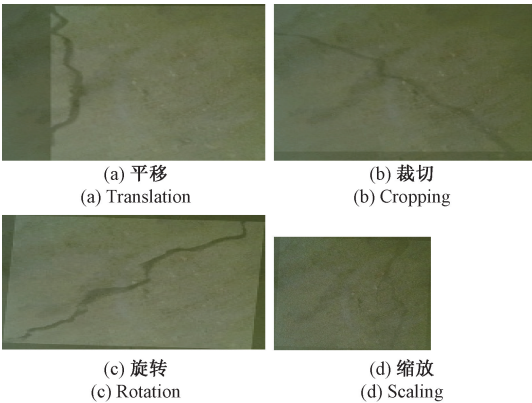


图 3 数据增强
Fig. 3 Data augmentation

使用“Labelimg”标注工具进行管桩裂缝标注。本研究按的比例随机筛选出 1 800 个训练集,600 个测试集,600

个验证集。训练集:验证集:测试集为 6 : 2 : 2 包括水下 2 562,水上 438(包含近距离,远距离,俯拍视角),数据分布如图 4 所示。数据集类别为 fissure。

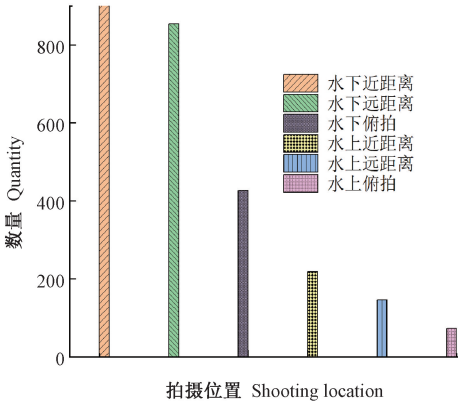


图 4 数据集分类
Fig. 4 Dataset classification

2 算法设计

2.1 YOLOv8 目标检测网络

YOLOv8 由 Ultralytics 开发,延续了 YOLO 家族核心理念,即通过单次神经网络前向传播(单阶段检测)实现目标检测的高效率,同时在精度和速度之间找到平衡。YOLOv8 网络模型的构成主要由 3 个关键部分组成: Backbone(主干网络)、Neck(颈部网络)和 Head(头部网络),如图 5 所示。

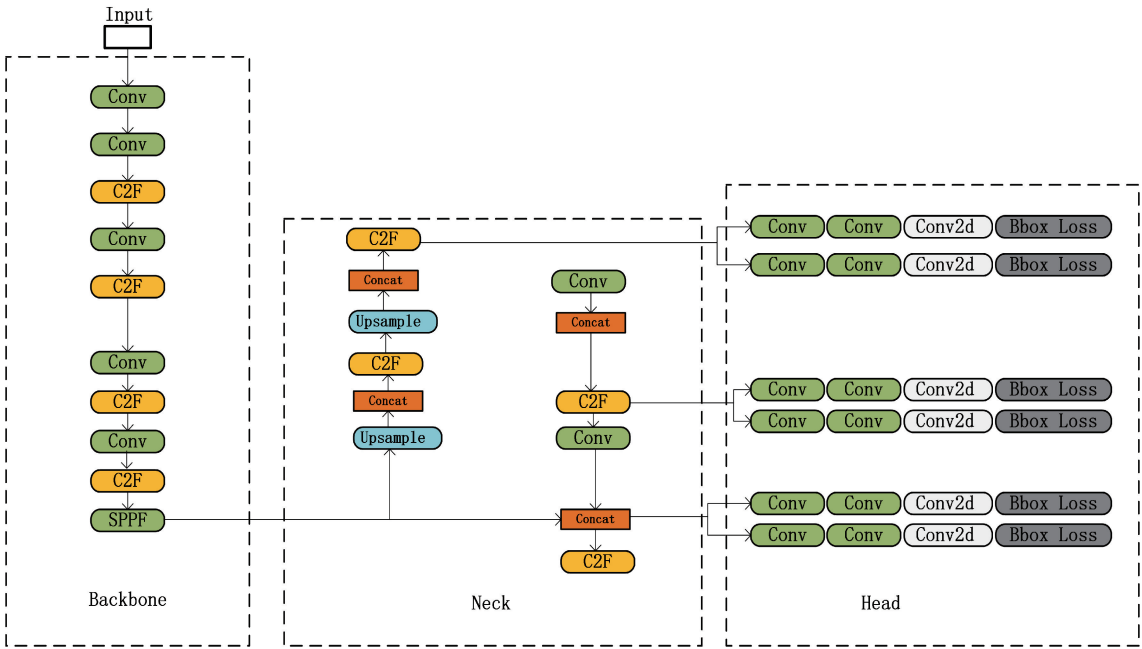


图 5 YOLOv8s 结构
Fig. 5 Structure of YOLOv8s

首先,输入图像通常会经过预处理后经过 Backbone (主干网络)。Backbone 是卷积神经网络的一部分,负责逐层提取图像的底层、中层和高层特征。通过多次卷积、池化操作,提取出不同分辨率的特征图。提取的多尺度特征接下来通过 Neck(颈部网络)进行特征融合。Neck 部分使用了 FPN (feature pyramid network) 和 PANet (path aggregation network),以结合不同层次的特征信息。FPN 通过自顶向下的路径进行高分辨率和低分辨率特征的融合,而 PANet 则进一步提升了特征信息在多尺度上的传递能力。经过特征融合的多尺度特征图被传递到 Head (头部网络),负责目标检测的最终输出。YOLOv8 的 Head 会根据每个特征图位置生成候选边界框 (bounding boxes),并对每个候选框预测目标的类别概率和边界框的坐标。结构如图 5 所示。

本文采用 YOLOv8s 作为基础模型,YOLOv8s 在水下管桩裂缝检测中具有显著优势。它不仅轻量高效,适合在移动设备上部署,实现实时检测,而且在复杂的水下环境中表现出色。YOLOv8s 在保持较高检测精度的同时,能适应光照不均、水下干扰多等条件,具备很强的泛化能力。其多尺度特征学习能力尤其适合识别微小、不规则的裂缝,加之便捷的迁移学习特性,使得它在不同的水体环境下都能稳定准确地检测裂缝,是水下裂缝检测任务中的理想选择。因此,本研究基于 YOLOv8s 提出了 MWM-YOLO 轻量化算法。

2.2 网络改进

1) MLLA 注意力机制

管桩长期浸泡在海水环境中容易积累生物附着(贝类和藻类等),当附着物吸附于管桩损伤部位,会增加损伤识别难度,同时在黑暗浑浊水下环境中,模型难以应对颜色失真和模糊特征,使得重要细节被忽略。此外,复杂的背景可能使模型难以聚焦于目标,光照变化适应性差会进一步影响检测准确性,同时小型裂缝可能被噪声掩盖,导致检测能力下降。所以引入 MLLA 注意力机制^[15]。Mamba-like linear attention (MLLA) 是通过将 Mamba 模型的一些核心设计融入线性注意力机制,从而提升模型的性能。具体来说,MLLA 主要整合了 Mamba 中的“忘记门”(forget gate)和模块设计(block design)这两个关键因素,同时 MLLA 通过使用位置编码(RoPE)来替代忘记门,从而在保持并行计算和快速推理速度的同时,提供必要的位置信息。如图 6 所示,它通过在通道和空间维度上保持较高的分辨率,有效降低了参数量,同时减少了因降维带来的信息损失。引入 MLLA 机制后,模型能够更好地关注与目标相关的图像区域,从而提升目标识别的准确性。面对水下环境中复杂的背景,MLLA 机制帮助模型捕捉重要的上下文信息,提高对目标的敏感度,并将注意力集中在目标区域,减少背景干扰。这使得模型在处理复杂背景和多个目标的图像时,展现出更强的鲁棒性和有效性。

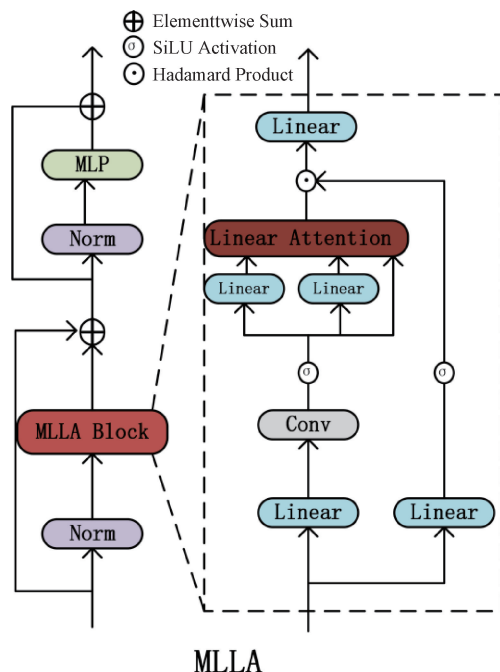


图 6 MLLA 注意力机制

Fig. 6 Attention mechanism of MLLA

2) 损失函数改进

水下管桩的裂缝通常呈现细长、不规则的形状,这些目标在常规 IoU 损失的计算中难以被准确捕捉。常规 IoU 在预测框与真实框形状差异大时表现不佳。同时水下环境中,由于光照不足或水体干扰,模型容易出现位置误差。易受噪声干扰或发生裂缝遮挡。改进为 WIoU 损失函数^[16],Wiou 在 IoU 损失中引入了形状权重,可以更好地拟合细长型目标,提高裂缝的检测准确性,同时 WIoU 在计算 IoU 时考虑了框的位置、中心点偏移,通过额外的距离权重来减小位置偏差对损失的影响,从而优化模型对裂缝位置的识别。Wiou 通过更精细的 IoU 计算来优化框的拟合效果,可以提高检测的准确率和召回率,从而确保裂缝的全面检测。

本研究采用了改进后的 WIoU V3 损失函数。该函数在锚框与目标框高度重叠时减小了几何度量的惩罚^[17],能有效减少高质量锚框之间的竞争,同时降低低质量样本产生的负面梯度影响,从而提升模型的泛化能力。图 7 展示了标注框、预测框及其最小外接矩形的参数示意图。

起初 WIoU 损失函数根据距离度量构建得到了 WIoU V1,公式为:

$$L_{WIoUv1} = R_{WIoU} L_{IoU} \quad (1)$$

$$R_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (2)$$

式中: W_g 和 H_g 表示最小包围框的宽和高; x, y 表示预测框坐标中心; x_{gt}, y_{gt} 表示真实框坐标中心; $R_{WIoU} \in [1, e)$ 会在一定程度上扩大普通质量锚框的 IoU 损失函数值

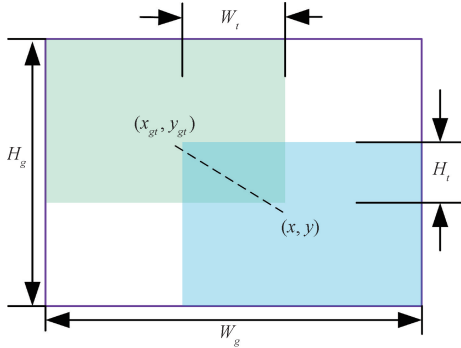


图 7 WIoU 参数示意图

Fig. 7 WIoU parameters

$L_{IoU}, L_{IoU} \in [1, e]$ 会明显降低高质量锚框的 R_{WIoU} , * 表示是将 W_g 和 H_g 从计算图中分离,可有效提高收敛效率。

WIoU V2 损失函数的计算公式 L_{WIoUv2} , 如式 (3) 所示^[18]。

$$L_{WIoUv2} = \left(\frac{L_{IoU}^*}{L_{IoU}} \right)^\gamma L_{WIoUv1} \quad (3)$$

式中: L_{IoU}^* 为单调聚焦系数, $\overline{L_{IoU}}$ 为平均值, 在公式中进行归一化, 使梯度增益保持在较高水平。

WIoU V3 损失函数通过引入离群度 β 来表示锚框的质量, L_{WIoUv3} 计算公式如式 (4)、(5) 所示。

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \quad (4)$$

$$L_{WIoUv3} = r L_{WIoUv1}, r = \frac{\beta}{\delta \alpha^{\beta-\alpha}} \quad (5)$$

式中: α, δ 表示超参数; β 为非单调聚焦系数。

3) 基于 MobileNet V4 主干网络改进

水下环境光线较暗, 裂缝与背景的对比较低, 检测难度大。同时水下检测设备往往内存和算力有限, 传统深度网络在移动端部署会因高计算需求而延迟严重, 并且水下检测任务中可能会使用不同的移动设备或平台。所以基于 MobileNetV4 的主干网络改进在深度卷积结构保持精度的同时降低了模型复杂度, 使得移动设备能有效处理低对比度图像, 提升裂缝的识别能力, 同时 MobileNetV4 的轻量化设计有效减少了计算资源消耗, 适合移动端实时运行, 同时能降低功耗和延长设备续航。

MobileNet V4^[19] 是 Google 提出的轻量级卷积神经网络架构, 可在硬件配置较低的情况下进行部署, MobileNetV4 引入了一种名为通用反转瓶颈(UIB)的新架构组件。UIB 是一个灵活的架构单元, 融合了反转瓶颈(IB)、ConvNext、前馈网络(FFN), 以及新颖的额外深度(ExtraDW)变体。如图 8 所示, 通过引入改进的 UIB 模块, C2f 主干能够更高效地提取特征, 从而提升模型的检测精度。这种高效特征提取对复杂场景中的目标检测尤其重要。

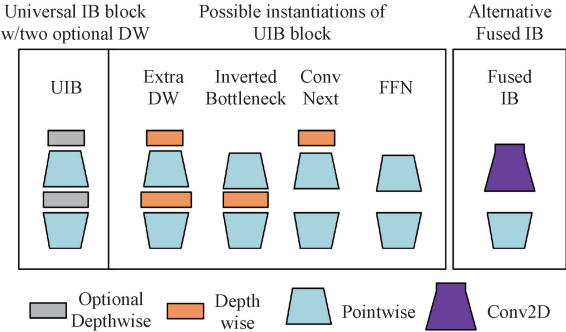


图 8 MobileNet V4-UIB

Fig. 8 MobileNet V4-UIB

2.3 融合图像增强与改进 YOLOv8 识别网络

YOLOv8 针对不同设备与开发需求提供了多个基本模型, YOLOv8s 模型最准确, 更有利于实际部署到管桩清洗机器人, 因此 本文选择采用 YOLOv8s 模型。为减少浑水环境的干扰, 提升检测视频效果和对缺陷的检测精度, 在主干部分加入 MobileNet V4, 更加高效地提取特征, 并且在主干特征提取网络各连接部位后引入 MLLA 注意力机制(Mamba-like linear attention)机制模块, 解决引入图像增强处理后过度平滑低质图像边缘对目标检测带来的抑制作用, 并将 IoU 损失函数改进为 WIoU 损失函数, 减少数据标注中低质量示例对模型性能的影响。改进 YOLOv8 的水下管桩裂缝检测方法 MWM-YOLO 模型结构如图 9 所示。

3 实验与分析

3.1 实验设备与训练

文本所用的操作系统为 windows11, CPU 型号为 12th Gen Intel(R) Core(TM) i5-12400F, GPU 型号为 NVIDIA GeForce RTX 2060。

为了验证本文的水下管桩裂缝检测的改进方法的有效性, 本实验包括以下环节: 1) 图像增强技术对浑水环境下的裂缝图像进行预处理, 使数据集扩充; 2) 对 YOLOv8 进行改进对比实验, 再进行消融实验, 验证改进的有效性。

所有实验训练模型都不采用预训练权重, batch size 设置为 32, 图片大小为 256×256 , 为单类别检测, 使用 SGD 作为优化器, 训练 200 轮。训练参数如表 1 所示。

3.2 评价指标

本次实验主要采用准确率 (precision, P)、召回率 (recall, R) 和平均精度均值 (mean average precision, mAP) 作为评价指标。如式 (6)~(8) 所示。

$$P = \frac{T_{TP}}{T_{TP} + F_{FP}} \times 100\% \quad (6)$$

$$R = \frac{T_{TP}}{T_{TP} + F_{FN}} \times 100\% \quad (7)$$

$$A_{AP} = \int_0^1 P(R) dR \quad (8)$$

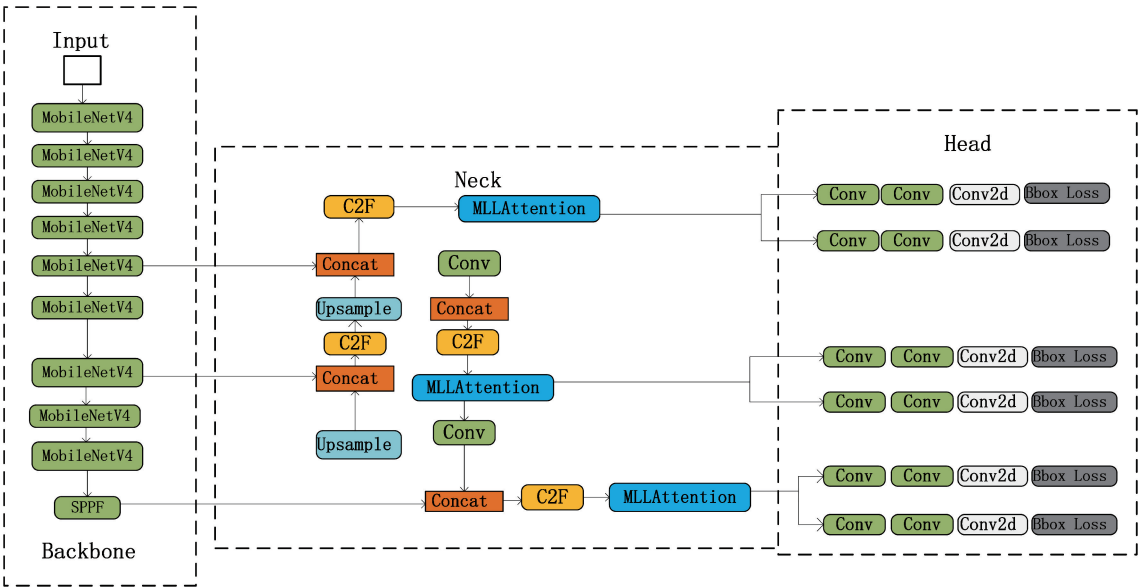


图 9 MWM-YOLO 模型结构
Fig. 9 Structure of MWM-YOLO model

表 1 训练参数

Table 1 Training parameters

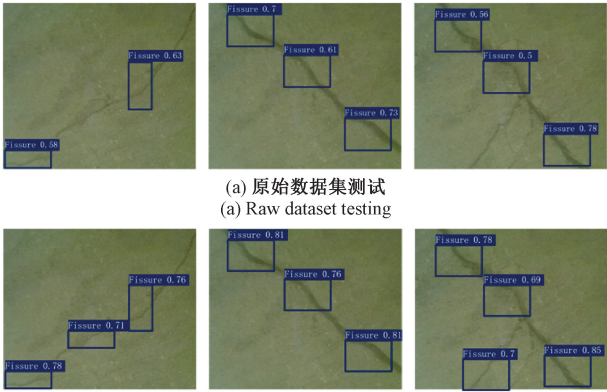
参数	参数值
batch size	32
图片大小	256×256
优化器	SGD
迭代次数	200

其中, T_{TP} 表示被正确分类为正样本的数量, F_{FP} 代表被错误分类为正样本的数量, F_{FN} 则是被错误分类为负样本的数量。准确率 P 和召回率 R 用于衡量模型性能, 而 AP 值是 P - R 曲线下的面积, 通过计算所有类别的 AP 值的平均, 得出 mAP 。

3.3 结果与分析

1) 数据集预处理有效性分析

通过数据增强技术对数据集进行亮度调整、平移处理和模糊等操作, 可以有效模拟不同光照条件、拍摄角度变化, 以及水流湍急的场景。为验证数据增强效果, 本研究使用 YOLOv8s 原模型对测试集中的平移、模糊和俯视图像进行了检测。图 10 展示了两组检测结果, 其中图 10(a) 为原始数据集的检测效果, 图 10(b) 为经过数据增强后数据集的检测效果。实验结果表明, 使用原始数据集训练的模型在模糊条件下存在漏检现象, 且在清晰度不足时预测框位置不准确, 置信度偏低(最高为 0.78, 最低为 0.5)。相比之下, 使用数据增强后数据集的模型在模糊条件下成功消除了漏检现象, 预测框位置更加准确, 置信度显著提高(最高达 0.85, 最低为 0.7)。本研究结果表明, 数据集的预处理显著提升了水下管桩裂缝检测模型的性能。



(a) 原始数据集测试
(a) Raw dataset testing
(b) 数据集预处理测试
(b) Dataset preprocessing test

图 10 图像预处理测试
Fig. 10 Image preprocessing test

2) 消融实验结果与分析

为了验证 MWM-YOLO 模型的可行性, 本文进行消融实验。结果如表 2 所示; 实验 1 为 YOLOv8s 原始模型; 实验 2 为使用 Mobilenetv4 作为网络骨干; 实验 3 表示在 Mobilenetv4 作为网络骨干加入 MLLA 注意力机制; 实验 4 表示在 Mobilenetv4 作为网络骨干的情况下采用 WIoU 损失函数; 实验 5 在加入损失函数后加入 MLLA 注意力机制。表 2 为消融实验结果, 图 11 为模型精度曲线。

如表 2 所表示, 对比实验 1 和 2, 在使用 Mobilenetv4 作为网络骨干后对比原模型, 模型大小从 21.4 MB 减小至 14.7 MB, FPS 从 100.21 提升至 183.92, 而精确率下降了 5.82%, 说明略微下降的精确率作为代价, 有效提升了模型的效率和速度。

对比实验 2 和 3, 可以看出实验 3 的精确率增加了

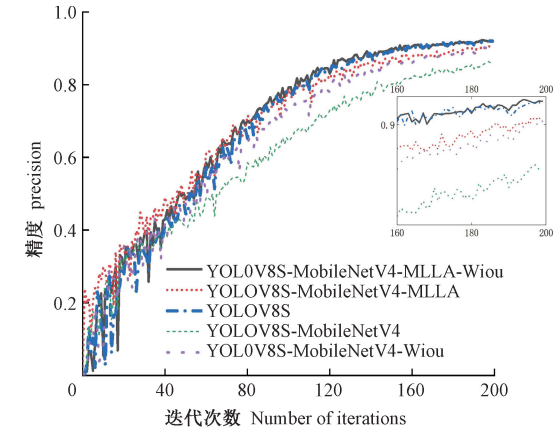


图 11 模型精度曲线

Fig. 11 Model accuracy curve

4.21%,同时召回率达到了 93.61%,增加了 4.81%,说明 MLLA 注意力机制能在识别管桩裂缝方面更为有效,可以有效抑制不重要特征,能够更好地减少假阳性和假阴性。

对比实验 2 和 4,精确率提升了 4.19%,召回率也提升了 4.72%,说明模型在处理目标检测任务时,特别是在定位和识别裂缝方面更加准确。WIoU 损失函数能够更好地衡量目标的重叠程度,特别是在处理不平衡样本或难以识别的边界时,增强了模型对重要区域的关注,从而提高了整体检测性能。

对比实验 1 和 5 模型,本研究提出的 MWM-YOLO 模型,精度提升,权重大小减少了 30.3%,速度提高了 37.25%,能够有效改善水下管桩裂缝的检测。

3)注意力机制对比实验

本研究在 YOLOv8s 模型中引入了 SEA^[20]、NAM^[21]和 MLLA^[22]等多种注意力机制,并在实验数据集上对它们的表现进行了对比。表 3 的结果显示,加入注意力机制后,模型的精度均有所提升,尤其是应用了 MLLA 机制时,平均精度均值(mAP)达到了 88.53%。在参数量略微增加的情况下,牺牲推理速度较少,显示出 MLLA 机制在性能优化上的优势。

表 2 消融结果

Table 2 Ablation results

实验	模型	Precision/ %	Recall/ %	mAP/ %	权重大小/MB	FPS/fps
1	YOLOv8s(原始)	91.93	96.41	87.85	21.4	100.21
2	YOLOv8s-Mobilenetv4	86.11	88.80	85.50	14.7	183.92
3	YOLOv8s-Mobilenetv4-MLLA	90.32	93.61	88.03	14.7	158.43
4	YOLOv8s-Mobilenetv4-WIoU	90.30	93.52	88.00	14.7	142.72
5	YOLOv8s-Mobilenetv4-WIoU-MLLA(本文)	92.09	96.03	89.10	14.9	137.54

表 3 注意力机制对比

Table 3 Comparison of attention mechanisms

模型	mAP@0.95	FPS	Params
YOLOv8s	87.85	100.21	11 135 987
YOLOv8s-SEA	86.56	85.27	11 834 867
YOLOv8s-NAM	86.86	85.56	11 137 011
YOLOv8s-MLLA	88.53	92.17	11 027 011

本研究通过热力图(Grad-CAM)可视化对比注意力机制对检测效果的影响(如图 12 所示)。结果表明,原始模型在训练过程中未能有效聚焦管桩裂缝的特征。而在引入 SEA、NAM 和 MLLA 等注意力机制后,模型对裂缝特征的关注度明显提高,背景干扰显著减少,关键信息丢失情况得到有效改善。尤其是 MLLA 模块,将浅层和深层特征进行融合,以增强对局部和全局信息的捕捉。同时,通过动态调整每层特征的重要性权重,确保模型能够在多尺度特征下聚焦于管桩裂缝的关键区域,从而显著提高识别的准确性和稳健性。提高了模型在识别管桩裂缝时的精确度。

4)损失函数对比实验

为获取高精度的管桩裂缝位置坐标,以支持后续修

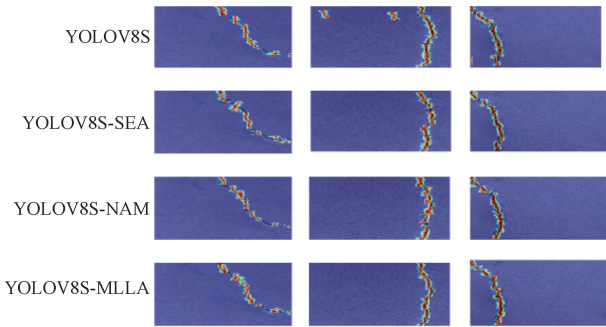


图 12 Grad-CAM 热力图

Fig. 12 Grad-CAM heat map

复,本研究对 4 种损失函数(CIoU、GIoU^[23]、WIoU 和 ShapeIoU^[24])进行了对比实验,结果如图 13 所示。实验显示,当迭代次数低于 30 时,CIoU 的边界框回归速度较慢;迭代次数超过 30 后,GIoU 的回归速度最慢。在 CIoU 和 ShapeIoU 之间,二者的收敛速度和损失值接近,且均稍优于 GIoU。WIoU 的动态非单调评估机制使模型更专注于一般质量的数据,在 120 次迭代后快速收敛,损失值最低,对裂缝位置的定位也更为精准。

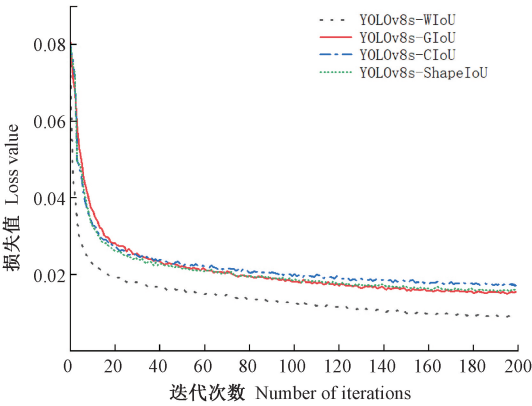


图 13 损失训练值对比
Fig. 13 Comparison of lost training values

5)不同算法检测对比

为了验证本文提出的算法相较于其他算法的优越性,本文进行了与 EDG-YOLO^[25]、YOLOv8s、YOLOv8n-4SCDP^[26]、TWRD-Net^[27]、BMS-YOLOv8^[28]、FMS-YOLOv5s^[29]以及其他研究人员提出的水下裂缝检测模型的对比实验。对比结果如表 4 所示。

表 4 算法检测对比
Table 4 Comparison of algorithms

模型	Paramater/ 10^6	mAP@0.95/%	Precision
EDG-YOLO	2.01	51.00	56.00
YOLOv8s	11.13	87.85	91.93
YOLOv8n-4SCDP	2.70	83.60	—
TWRD-Net	—	81.30	83.60
BMS-YOLOv8	11.12	84.30	83.30
FMS-YOLOv5s	5.78	52.10	—
本文	11.02	89.10	92.09

根据表所示 EDG-YOLO、YOLOv8n-4SCDP 与 FMS-YOLOv5s 参数量分别达到了 2.01×10^6 、 2.7×10^6 、 5.78×10^6 。但是精度却只有 51%、83.6%、52.1%。说明这 3 个算法对水下复杂目标的处理不够敏感,从而影响了算法的精度,与其他算法相比存在明显差距。无法满足水下机器人检测管桩裂缝的巡检要求。YOLOv8s 为 YOLO 系列的主流模型,BMS-YOLOv8 与 TWRD-Net 网络模型为最新提出的模型,他们参数量对比本文模型分别增加了 1.1×10^5 、 1×10^5 。而精度却分别减少了 1.25%、4.8%、7.8%。精确率也相比本模型较低。无法满足水下机器人巡检后的图像在移动端高效处理的需求。本文提出的模型,在参数量为 11.02×10^6 时平均精度能够达到 89.1%,能够满足水下管桩清洗机器人的巡检要求。

3.4 模型部署与验证

1)模型部署

为提升管桩清洗机器人在低算力边缘设备上的识别

能力,本研究将优化后的 YOLOv8 模型权重文件(best.pt)转为 ONNX 格式,增强模型的移植性和跨平台性能。随后,借助专为移动端和嵌入式设备优化的高效推理引擎 NCNN,将模型进一步转换为 Android 兼容的 param 和 bin 格式文件。应用 Autojs 开发软件,自行设计用户界面(UI),实现图片和实时视频检测模式,并最终生成 APK 文件,完成优化后 YOLOv8 模型在安卓平台的部署。图 14 所示为部署流程。

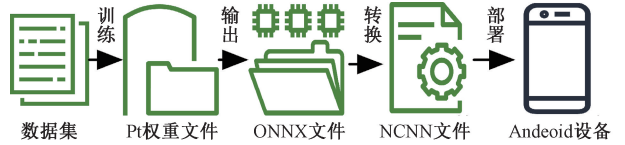


图 14 模型部署流程
Fig. 14 Model deployment process

2)部署验证

在部署验证阶段,采用 OnePLUS7 安卓设备,具体设备参数如表 5 所示,分别对 3 种模型(YOLOv8s、YOLOv8n 和改进后的 YOLOv8)在 Android 应用程序上进行了对比测试实验。

表 5 OnePLUS7 设备参数
Table 5 OnePLUS7 device parameters

配置	参数
操作系统	H2OS
CPU 型号	高通骁龙 855
GPU 型号	高通 Adreno 640
RAM 容量	12 GB
ROM 容量	256 GB

将图像输入尺寸设定为 256×256 ,以 GPU 和 CPU 的帧率(FPS)及平均识别精度作为模型部署后的性能评估指标。为了全面测试模型在多种场景下的检测准确度,实验中选用了不同角度、光照条件和距离的图像样本。安卓设备上的测试结果汇总于表 6。

表 6 安卓设备测试结果
Table 6 Test results of Android devices

输入尺寸	模型	GPU-FPS	CPU-FPS	平均精度/%
256×256	YOLOv8s	25	8	78.59
256×256	YOLOv8n	28	15	68.31
256×256	本文	31	24	85.02

根据测试结果,优化后的模型在安卓设备上展现出了比其他模型更优异的整体部署性能。应用程序的检测结果如图 15 所示。

从检测结果图像中可以明显看出,即使在远距离或侧角度的情况下,改进后的 YOLOv8 模型也能保持高水平的

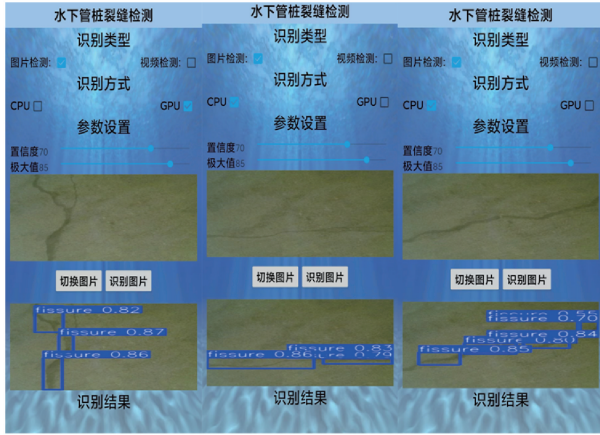


图 15 安卓应用程序效果图

Fig. 15 Renderings of an Android application

准确性。这证明了模型的鲁棒性和适应性,使其能够处理复杂的实际场景。此外,该模型使用 GPU 运行时实现了大约 25.2 fps 的平均帧率,展示了其在边缘计算设备上的高效性能,并为水下实际应用提供了强大的支持。

4 结 论

为实现水下管桩裂缝的高效检测,本研究提出了轻量化模型 MWM-YOLO,旨在提升边缘计算设备上目标识别的速度与精度。通过采用 MobileNetV4 作为主干网络,MWM-YOLO 在检测速度和计算复杂度方面表现出色,在安卓设备 GPU 模式下达到了 31 fps,显著降低了嵌入式设备的资源消耗,满足了水下机器人实时识别的需求。此外,模型结合 MLLA 注意力机制和 WiIoU 损失函数,增强了对细微裂缝特征的捕捉能力,并有效提升了检测的鲁棒性和精度。在复杂环境下,MWM-YOLO 较其他检测模型的精度提升至 89.1%,提高了 1.25%,充分证明了其在水下管桩裂缝检测中的应用潜力,为后续实际管桩裂缝的位置定位与修复工作提供了强有力的技术支持。同时下一步针对水下管桩多种缺陷与海洋复杂环境带来的干扰进行研究。

参考文献

[1] 王放,金樱,张应辉,等.国内外码头钢管桩的腐蚀状况调查及防护对策[J].材料保护,2022,55(4):172-181.
WANG F, JIN Y, ZHANG Y H, et al. Corrosion investigation and protective measures of wharf steel pipe piles at home and abroad [J]. Materials Protection, 2022, 55(4):172-181.

[2] XIA R L, JIA CH, LIU CH Y, et al. Non-uniform corrosion characteristics of the steel pipe pile exposed to marine environments[J]. Ocean Engineering, 2023, 272:113873.

[3] ALBITAR H, DANDAN K, ANANIEV A, et al.

Underwater robotics: Surface cleaning technics, adhesion and locomotion systems [J]. International Journal of Advanced Robotic Systems, 2016, 13(1): 7-7.

[4] 刘亢,刘均利,余文成. 2007~2015 年洪水导致垮塌桥梁的统计分析[J]. 城市道桥与防洪, 2017(1):90-92, 12-13.
LIU K, LIU J L, YU W CH. Statistical analysis of collapsed bridges caused by floods in 2007~2015[J]. Urban Roads and Bridges and Flood Control, 2017(1):90-92,12-13.

[5] 吉伯海,傅中秋. 近年国内桥梁倒塌事故原因分析[J]. 土木工程学报, 2010, 43(S1):495-498.
JI B H, FU ZH Q. Analysis on the causes of bridge collapse accidents in China in recent years[J]. Chinese Journal of Civil Engineering, 2010, 43(S1):495-498.

[6] 李培坤,李锋,葛忠显,等. 基于改进 YOLOv8n 的水下目标检测算法[J]. 电子测量技术, 2025, 48(3):172-179.
LI P K, LI F, GE ZH X, et al. Underwater target detection algorithm based on improved YOLOv8n[J]. Electronic Measurement Technology, 2025, 48(3): 172-179.

[7] 李泽荣,刘爱荣,陈炳聪,等. 基于融合图像增强与改进 YOLOv7 算法的桥梁水下结构缺陷识别[J]. 工程力学, 2024, 41(S1):245-252.
LI Z R, LIU AI R, CHEN B C, et al. Bridge underwater structure defect recognition based on fusion image enhancement and improved YOLOv7 Algorithm[J]. Engineering Mechanics, 2024, 41(S1): 245-252.

[8] BURGUERA A, BONIN-FONT F. An unsupervised neural network for loop detection in underwater visual SLAM[J]. Journal of Intelligent & Robotic Systems, 2020, 100(3):1-21.

[9] 张云泽. 水下桥梁墩桩裂缝视觉检测方法研究[D]. 沈阳:沈阳大学, 2021.
ZHANG Y Z. Research on visual detection method of pier pile crack of underwater bridge [D]. Shenyang: Shenyang University, 2021.

[10] ANTONI B. Lightweight underwater visual loop detection and classification using a siamese convolutional neural network[J]. IFAC-PapersOnLine, 2021, 54(16): 410-415.

[11] 田青,赵宇,张正,等. 基于 Enhanced Zero-DCE++ 的水下选通图像增强技术研究[J]. 电子测量技术, 2025, 48(1):166-174.
TIAN Q, ZHAO Y, ZHANG ZH, et al. Research on underwater gated image enhancement technology based on Enhanced Zero-DCE++ [J]. Electronic Measurement Technology, 2025, 48(1):166-174.

[12] 袁国铭,刘海军,李晓丽,等. 纹理感知联合颜色直方图特征的水下图像增强[J]. 光学精密工程, 2024,

- 32(13):2112-2127.
- YUAN G M, LIU H J, LI X L, et al. Underwater image enhancement of texture perception combined with color histogram features[J]. Optics and Precision Engineering, 2024, 32(13):2112-2127.
- [13] 王宁,贾薇,陈延政,等. 基于深度估计和梯度下降策略的水下图像恢复与增强[J]. 大连海事大学学报, 2025, 50(3):1-12.
- WANG N, JIA W, CHEN Y ZH, et al. Underwater Image Recovery and Enhancement Based on Depth Estimation and Gradient Descent Strategy[J]. Journal of Dalian Maritime University, 2025, 50(3):1-12.
- [14] ZIAD L, OUBBIH O, KARAMI F, et al. A nonlocal model for image restoration corrupted by multiplicative noise[J]. Signal, Image and Video Processing, 2024, 18:5701-5718.
- [15] HAN D CH, WANG Z Y, XIA ZH F, et al. Demystify Mamba in Vision: A linear attention perspective [J]. ArXiv preprint arXiv: 2405.16605, 2024.
- [16] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism[J]. ArXiv preprint arXiv:2301.10051, 2023.
- [17] 宁峰,赵良军,郑莉萍,等. 一种针对 SAR 图像的船舶目标检测算法[J]. 电光与控制, 2024, 31(5):60-65.
- NING F, ZHAO L J, ZHENG L P, et al. A ship target detection algorithm for SAR images [J] Electronics Optics & Control, 2024, 31(5):60-65.
- [18] 赵庆,魏鸿磊,杨伟宁,等. 基于改进 YOLOv5 网络的轮胎规格字符识别[J]. 组合机床与自动化加工技术, 2024(2):164-168.
- ZHAO Q, WEI H L, YANG Y N, et al. Tire specification character recognition based on improved YOLOv5 network[J]. Combined Machine Tool and Automatic Processing Technology, 2024(2):164-168.
- [19] QIN D F, LEICHNER C, DELAKIS M, et al. MobileNetV4-universal models for the mobile ecosystem[J]. ArXiv preprint arXiv:2404.10518, 2024.
- [20] WAN Q, HUANG Z L, LU J CH, et al. Seaformer: Squeeze-enhanced axial transformer for mobile semantic segmentation[C]. The Eleventh International Conference on Learning Representations, 2023.
- [21] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [22] ZHENG ZH H, WANG P, REN D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2021, 52(8): 8574-8586.
- [23] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.
- [24] ZHANG H, ZHUANG SH J. Shape-IoU: More accurate metric considering bounding box shape and scale[J]. Arxiv preprint arXiv:2312.17663, 2023.
- [25] 赵佰亭,程瑞丰,贾晓芬. 融合多尺度特征的 YOLOv8 裂缝缺陷检测算法[J]. 计算机工程与应用, 2024, 60(22):261-270.
- ZHAO B T, CHENG R F, JIA X F. YOLOv8 crack defect detection algorithm fused with multi-scale features[J]. Computer Engineering and Applications, 2024, 60(22):261-270.
- [26] 姜源,付波,权轶,等. 基于改进 YOLOv8n 的 PCB 缺陷检测算法[J]. 国外电子测量技术, 2024, 43(6): 22-32.
- JIANG Y, FU B, QUAN Y, et al. PCB defect detection algorithm based on improved YOLOv8n[J]. Foreign Electronic Measurement Technology, 2024, 43(6):22-32.
- [27] 高嘉,刘涛,王显峰,等. TWRD-Net: 一种用于曳引钢丝绳缺陷的实时检测网络算法[J]. 仪器仪表学报, 2023, 44(6):223-235.
- GAO J, LIU T, WANG X F, et al. TWRD-Net: A real-time detection network algorithm for traction wire rope defects [J]. Chinese Journal of Scientific Instrument, 2023, 44(6):223-235.
- [28] 庄盛. 基于深度学习的水下图像目标识别研究[D]. 吉林:吉林化工学院, 2024.
- ZHUANG SH. Research on underwater image target recognition based on deep learning [D]. Jilin: Jilin Institute of Chemical Technology, 2024.
- [29] 曹立,康少波. 基于 FMS-YOLOv5s 的轻量化交通标志识别算法[J]. 国外电子测量技术, 2024, 43(5):179-189.
- CAO L, KANG SH B. Lightweight traffic sign recognition algorithm based on FMS-YOLOv5s[J]. Foreign Electronic Measurement Technology, 2024, 43(5):179-189.

作者简介

陈潇威,硕士研究生,主要研究方向为海洋机电装备。

E-mail:1060619412@qq.com

宋瑞银(通信作者),博士,教授,博士生导师,主要研究方向为海洋机电装备。

E-mail:ruiyinsong@163.com

吴瑞明,博士,副教授,研究生导师,主要研究方向为智能装备。

E-mail:wuruiming@163.com

李凤姪,博士,讲师,主要研究方向为水轮机海洋能发电以及海洋机电设备技术。

E-mail:lfs@nbt.edu.cn

王天恒,硕士研究生,主要研究方向为渔业自动化装备。

E-mail:1440907679@qq.com