

城市排水管道缺陷智能检测方法研究

雷江浩¹, 刘光宇², 张温², 程静¹

(1. 西安建筑科技大学 建筑设备科学与工程学院, 西安 710055;

2. 西安建筑科技大学 管理学院, 西安 710055)

摘要: 针对城市排水管道缺陷自动化检测准确率低、目标定位不准确的问题, 提出一种改进 YOLOv8 的排水管道缺陷检测模型; 该模型在基线模型中引入感受野注意力卷积, 并构建了 C2F_RFAConv 模块, 通过空间感受野与卷积交互自适应学习的方式, 增强模型对缺陷特征的提取能力; 同时, 提出一种混合注意力高低阶特征融合网络, 将 Backbone 和 Neck 输出的 3 种不同尺度的低阶特征和高阶特征进行有效融合, 增强模型学习图像全局上下文信息的能力; 此外, 对影响边界框回归的重叠情况、中心点距离、宽高偏差等因素进行了综合分析, 设计 Inner-MPDIoU 损失函数, 使模型适应不同尺寸的缺陷检测任务, 提高缺陷目标边界框的定位准确率; 经过实验验证发现, 改进后的模型取得了 93.9% 的平均检测准确率, 较改进之前提升 3.7%, 漏检率和误检率仅为 9.1% 和 17.6%, 较改进之前分别降低 3.2% 和 2.7%。

关键词: 城市排水管道; 缺陷检测; YOLOv8; 感受野注意力卷积; 特征融合; Inner-MPDIoU

Research on Intelligent Defect Detection Method for Urban Drainage Pipelines

LEI Jianghao¹, LIU Guangyu², ZHANG Wen², CHENG Jing¹

(1. School of Building Services and Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China;

2. School of Management, Xi'an University of Architecture and Technology, Xi'an 710055, China)

Abstract: To address the problems of low automatic detection accuracy and inaccurate target positioning for urban drainage pipeline defects, an improved YOLOv8 drainage pipeline defect detection model is proposed. This model introduces receptive field attention convolution to the baseline model and constructs the C2F receptive field attention convolution (RFAConv) module, enhancing the model's ability to extract defect features through interactive adaptive learning of spatial receptive fields and convolutions. Additionally, a hybrid attention high-order and low-order feature fusion network is proposed, which effectively fuses the low-order and high-order features of three different scales by the backbone and neck outputs, enhancing the model's ability to learn the global contextual information of the image. The Inner-MPDIoU loss function is designed by comprehensively analyzing the factors affecting the bounding box regression, such as overlap, distance between the center points, and deviation in width and height. The model is suitable for the defect detection of different sizes, and improves the positioning accuracy of the defect target boundary box. Experimental results show that compared with the baseline model, the improved model achieves an average detection accuracy of 93.9%, with an increase of 3.7%, the missed detection rate and false detection rate of the improved model are reduced to 9.1% and 17.6%, with decreases of 3.2% and 2.7%, respectively.

Keywords: urban drainage pipes; defect detection; YOLOv8; RFAConv; feature fusion; inner-MPDIoU

0 引言

排水管道系统作为城市地下水系统正常运行的重要基础设施, 承担着运输、排放城市雨污水的任务^[1]。然而, 管道系统长时间运行, 不可避免地出现诸如沉积堵塞、破裂、脱节等缺陷, 若未能及时检修, 将会导致城市内涝、道路塌陷和地下水污染等问题。工程中常用 CCTV (Closed Circuit Television) 方法采集管道内部视频^[2], 再由专业人员根据管道普查规范对视频中的缺陷进行辨识和评定, 但人工识别效率低、主观性强、误检率高^[3], 不能适应现有大规模排水管道普查任务, 因此, 有必要进行智能化、自动

化的排水管道缺陷检测研究。

城市地下排水管道内部环境复杂, CCTV 采集的视频图像质量参差不齐, 会影响缺陷检测模型的性能。文献 [4] 使用动态直方图均衡化方法和双边滤波器解决了图像亮度不均和噪点过高问题, 提出一种利用多层全局特征提高检测性能和细粒度分类表示的自动化缺陷检测网络, 检测 5 种常见管道缺陷, 平均准确率为 72.5%。文献 [5] 通过使用一种基于注意力机制的去雾算法和一种特征金字塔结合生成对抗网络 GAN 的去模糊网络框架, 解决图像有雾和模糊问题, 提升了管道缺陷图像的质量, 并对管道中的

收稿日期: 2024-05-30; 修回日期: 2024-07-01。

基金项目: 陕西省自然科学基金面上项目 (2024JC-YBMS-286)。

作者简介: 雷江浩 (1997-), 男, 硕士研究生。

通讯作者: 张温 (1994-), 男, 博士研究生。

引用格式: 雷江浩, 刘光宇, 张温, 等. 城市排水管道缺陷智能检测方法研究[J]. 计算机测量与控制, 2024, 32(12): 81-87.

错口、破裂和障碍物等 3 种缺陷进行分割研究，最终获得了较高的分割精度。可见，高质量的图像数据对于研究是至关重要的。

文献 [6] 对比了 SSD、Faster-R-CNN 和 YOLOv3 等 3 种主流检测模型，仅检测树根、沉积物、树根和沉积物 3 种情况，实验研究发现，平均检测准确率的排名为 Faster-R-CNN>YOLOv3>SSD，而检测速度排名则是相反的，YOLOv3^[7]在平衡检测速度和准确率的方面更好，在排水管道缺陷检测方面潜力巨大。文献 [8] 在 YOLOv5^[9]中加入微尺度检测机制和注意力机制，改善了模型的缺陷检测机制，并增强了神经网络对空间和通道信息的获取能力，用于检测错口、脱节、沉积物、破裂等 12 种常见的排水管道缺陷，但平均准确率仅为 75.9%。文献 [10] 为了提高小尺寸缺陷的检测精度，在 YOLOv5 中加入自适应 CA 注意力机制，以便充分学习缺陷的边缘特征，采用 SiOU 损失函数加速模型收敛，检测变形、障碍物、浮渣、树根和脱落等 5 种常见缺陷，最终获得 71.1% 的检测准确率。文献 [11] 通过 StyleGAN2 方法对小样本数据集进行扩增，并在 YOLOX 的特征融合层中引入 SE 注意力机制，解决了不同尺度特征融合后造成的特征重叠问题，使用 CIoU 损失函数，提高了目标检测框的回归效率，但检测准确率仅有 68.76%。现有基于 YOLO 算法的排水管道缺陷检测研究虽取得不错的效果，但对于复杂背景中的多类缺陷，其检测准确率仍有提升空间。

因此，本文提出一种改进 YOLOv8^[12]的排水管道缺陷检测模型，能够进一步提高缺陷检测的准确率和定位精度。主要工作如下：1) 使用直方图均衡化方法和暗通道先验去雾算法^[13]处理图像，改善图像质量，构建一个缺陷特征丰富和样本均衡的管道图像数据集；2) 引入感受野空间注意力卷积^[14]，提出混合注意力高低阶特征融合网络，对模型 Backbone 和 Neck 输出同一尺度的低阶特征和高阶特征进行

有效融合，增强模型对图像全局上下文信息的学习能力；3) 设计 Inner-MPDIoU^[15-16]损失函数，综合分析影响边界框回归的重叠情况、中心点距离和宽高偏差等因素，使模型适应不同尺寸的缺陷，提升缺陷目标定位准确性。

1 排水管道缺陷检测模型

排水管道缺陷检测模型如图 1 所示，由 Backbone、Neck 和 Head 组成。Backbone 中的感受野注意力卷积用于提取图像中不同尺度的缺陷特征信息，空间金字塔池化层通过池化操作将不同尺度的缺陷固定尺寸并输出；Neck 通过构建自下而上和自上而下的特征金字塔，对 Backbone 输出不同尺度的缺陷特征进行深度融合；Head 解耦头输出 3 个不同尺度的特征图，每个尺度包含两个分支，分别执行缺陷的分类任务和定位任务。首先，对原始图像进行对比度增强和去雾处理，然后将其输入到检测模型中，网络通过学习缺陷图像特征，最终输出检测结果。

1.1 感受野空间注意力卷积

YOLOv8 在提取缺陷特征时，不能充分学习图像的全局上下文信息，会出现缺陷漏检、误检等问题，导致缺陷检测准确率低。卷积神经网络以滑块的形式遍历学习缺陷图像特征，可以通过插入空间注意力机制来增强网络的特征学习能力，当把空间注意力机制插入到 1×1 卷积前面时，网络学习所有滑块内的特征，不会造成冗余；当把空间注意力机制插入到 3×3 卷积前面时，网络在遍历每个感受野滑块内的特征时会产生特征冗余，导致特征学习效率低。

因此，引入感受野注意力卷积 (RFACConv, receptive-field attention convolution)^[14]，构建 C2F_RFACConv 模块，通过感受野特征与卷积进行交互自适应学习，增强模型学习图像全局上下文信息特征的能力。感受野空间注意力卷积学习的感受野空间特征是根据卷积核大小动态生成的，以 3×3 卷积核为例，输入原始特征图，输出新的特征图由多个不重叠的 3×3 感受野空间特征滑块组成，整个变换过

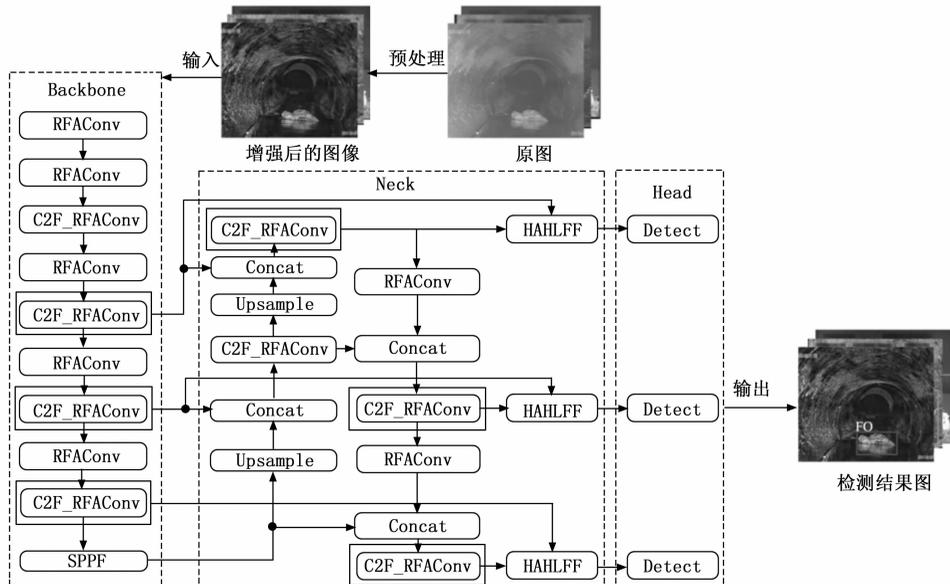


图 1 排水管道缺陷检测模型

计算如下：

$$IoU = \frac{|B_{gt} \cap B_{prd}|}{|B_{gt} \cup B_{prd}|} \quad (8)$$

$$CIoU = IoU - \frac{\rho^2(b^{gt}, b^{prd})}{C^2} - \alpha \quad (9)$$

式中， IoU 为预测框与真实框的交并比， B_{gt} 为真实框的面积， B_{prd} 为预测框的面积。 b_{gt} 和 b_{prd} 分别为 B_{gt} 和 B_{prd} 的中心点， $\rho(\cdot)$ 为欧氏距离， C 为包围真实框和预测框的最小边界框的对角线距离， α 是正权衡超参数， ν 用来衡量纵横比的一致性。

$CIoU$ 损失函数虽然考虑了边界框重叠的情况和中心点距离，但由于沉积物、障碍物之间存在遮挡、重叠情况，对于多尺寸缺陷，两边界框的宽高比相同而数值不同时，也会限制边界框回归准确性。因此，本文设计了 Inner-MPDIoU 边界框损失函数，采用辅助 (Inner) 边界框计算损失并加速边界框回归，引入 R_a 比例因子控制辅助边界框比例大小来检测不同尺寸缺陷，并利用两边界框左上角两点的距离和右下角两点的距离进行边界框相似度对比，进一步提高边界框回归准确率，损失函数计算如图 5 所示。

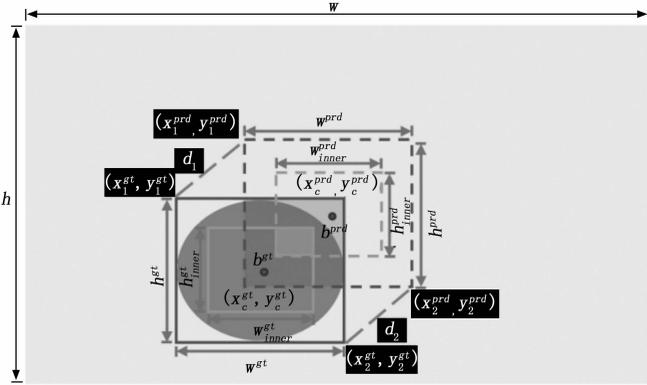


图 5 Inner-MPDIoU 损失函数计算示意图

R_a 比例因子的取值范围在 $[0.5, 1.5]$ ，当 R_a 的取值在 $[0.5, 1.0]$ 时，如图 5 所示，将生成一个比真实框 (左下较大实线框，宽为 w^{gt} ，高为 h^{gt}) 小的辅助边界框 (左下较小实线框，宽为 w^{inner} ，高为 h^{inner})，此时，真实框在外，辅助边界框在内，两边界框的中心点 b^{gt} 一致，唯一区别在于尺度不同，那么对于检测大尺寸缺陷，如沉积物、脱节和横向切口等，网络能够实现更为精准的回归定位；当 R_a 的取值在 $[1.0, 1.5]$ 时，将生成一个比真实框大的辅助边界框，辅助边界框在外，真实框在内，两个边界框仍然具有相同的中心点 b^{gt} ，此时对于裂缝、障碍物等小尺寸缺陷来说，相当于将小目标的边界框进行了放大，使得网络能够更好地学习小尺寸缺陷的特征信息。

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} * R_a}{2} \quad (10)$$

$$b_r^{gt} = x_c^{gt} + \frac{w^{gt} * R_a}{2} \quad (11)$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} * R_a}{2} \quad (12)$$

$$b_b^{gt} = y_c^{gt} + \frac{h^{gt} * R_a}{2} \quad (13)$$

$$b_l^{prd} = x_c^{prd} - \frac{w^{prd} * R_a}{2} \quad (14)$$

$$b_r^{prd} = x_c^{prd} + \frac{w^{prd} * R_a}{2} \quad (15)$$

$$b_t^{prd} = y_c^{prd} - \frac{h^{prd} * R_a}{2} \quad (16)$$

$$b_b^{prd} = y_c^{prd} + \frac{h^{prd} * R_a}{2} \quad (17)$$

$$inter = [\min(b_l^{gt}, b_l^{prd}) - \max(b_l^{gt}, b_l^{prd})] * [\min(b_t^{gt}, b_t^{prd}) - \max(b_t^{gt}, b_t^{prd})] \quad (18)$$

$$union = (w^{gt} * h^{gt}) * (R_a)^2 + (w^{prd} * h^{prd}) * (R_a)^2 - inter \quad (19)$$

$$IoU_{inner} = \frac{inter}{union} \quad (20)$$

真实框和内真实框 (辅助边界框) 的中心点 b^{gt} 用 (x_c^{gt}, y_c^{gt}) 表示，式 (10) ~ (13) 表示的是辅助内真实框。预测框 (右上较大虚线框，宽为 w^{prd} ，高为 h^{prd}) 和内预测框 (右上较小虚线框，宽为 w_{inner}^{prd} ， h_{inner}^{prd} 高为) 的中心点 b^{prd} 用 (x_c^{prd}, y_c^{prd}) 表示，式 (14) ~ (16) 表示的是辅助内预测框，真实框内的灰色椭圆代表缺陷目标。可以看到，Inner-IoU 函数的边界框损失计算原理与 $CIoU$ 一致，都是通过计算预测框与真实框的交并比来计算损失的，如式 (18) ~ (20)。在图 5 中， w^{gt} 和 h^{gt} 分别表示真实框的宽度和高度， w^{prd} 和 h^{prd} 分别表示预测框的宽度和高度。 (x_1^{gt}, y_1^{gt}) 和 (x_2^{gt}, y_2^{gt}) 分别表示真实框和预测框的左上角点与右下角点的坐标。

$$d_1^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \quad (21)$$

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \quad (22)$$

$$MPDIoU = IoU - \frac{d_1^2}{h^2 + w^2} - \frac{d_2^2}{h^2 + w^2} \quad (23)$$

$$L_{inner-MPDIoU} = L_{MPDIoU} + IoU - IoU_{inner} \quad (24)$$

式中， w 和 h 分别为图像的宽度和高度， d_1 和 d_2 分别为真实框和预测框的左上角点的距离和右下角点的距离，用来进行边界框相似度对比，解决两边界框因宽高比相同但宽高数值不同而导致的边界框回归不准确的问题。综上，引入 Inner-MPDIoU 边界框损失函数，综合分析影响边界框回归准确性的重叠情况、中心点距离、宽高偏差等因素，使模型能够适应不同尺寸的缺陷，提高目标边界框回归准确性。

2 数据集构建及图像预处理

2.1 数据集构建

本文使用的所有图像数据均来自丹麦奥尔堡大学提供的 Sewer-ML^[21] 和 Codalab 竞赛平台提供的 ICP2022^[22] 两个数据集，前一个数据集共有 130 万张管道图像，包含 18 种缺陷，用于排水管道的缺陷分类任务，但缺陷类别不平衡，图像质量不佳，如图 6 中的 (a)、(b) 和 (d)；后一个数据集是我国广州市某管道工程公司采集的管道视频数据，缺陷特征较为丰富，图像质量较高，但缺陷类别同样

不平衡, 如图 6 中的 (c) 和 (e)。从两个数据集中总共选取 12 266 张图像, 其中包含沉积物、裂缝、脱节、横向切口和障碍物等 5 种缺陷, 这样做既可以丰富数据特征, 又可以平衡缺陷类别。按照 8 : 1 : 1 将所有图像分为训练集、验证集和测试集进行实验。部分原始排水管道缺陷图像存在水雾干扰和图像不清晰的问题, 会影响模型的检测性能, 需要先对原始图像进行去雾和增强对比度操作, 再将特征增强后的排水管道缺陷图像输入到模型中进行训练和检测。

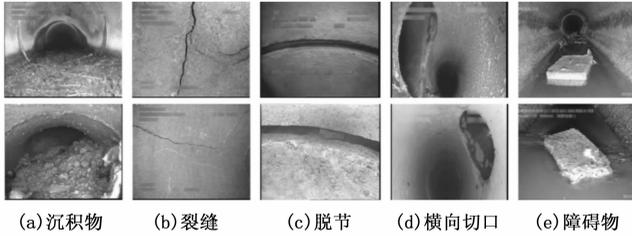


图 6 5 种典型排水管道缺陷图像

2.2 图像预处理

数据集中有一部分图像存在过亮或过暗和水雾干扰等问题, 使得缺陷图像的原始特征不清晰, 如图 7 所示。如果直接将其输入到检测模型中, 会导致模型学习不到缺陷的有效特征, 使得模型检测准确率过低, 为了提升图像质量, 本文采用直方图均衡化方法和去雾算法处理图像, 调节图像的亮度和对比度, 去除图像中的水雾和模糊, 提升图像的清晰度。

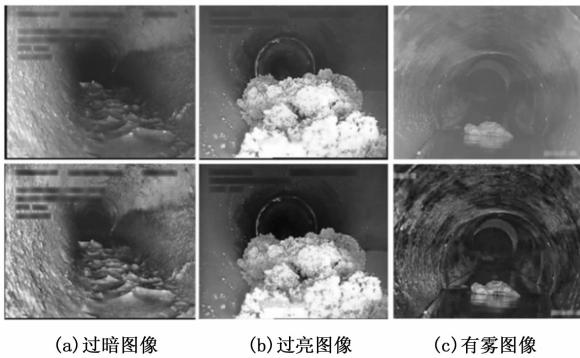


图 7 图像预处理前后对比

3 实验与分析

3.1 实验配置及评价指标

操作系统: Ubuntu18.04 操作系统; 中央处理器: Inter (R) Corei5-13400FCPU@2.50 GHz, 10 核 16 线; 显卡: NVIDIA RTX2080Ti (11.0 GB) 显卡; 内存: 16 GB; Python: 3.8.5; 深度学习框架: Pytorch1.13.1; CUDA: 11.0。

模型输入尺寸为 640×640 像素, 批量大小 (batch-size) 为 32, 模型总共迭代 150 个周期, 进程数 (works) 为 16 其他参数采用原始 YOLOv8 网络默认设置。在模型训练过程中, 为减小模型陷入局部最优的可能性, 使用随机

梯度下降 (SGD, stochastic gradient descent) 优化器。

实验以查准率 *Precision*、查全率 *Recall*、各类缺陷平均精度 *mAP*、漏检率 *Missed Detection Rate* 和误检率 *False Detection Rate* 作为评估指标。各指标定义如下:

查准率代表在被预测为正例的样本中, 被预测正确的样本所占的比例。*Precision* 表示为:

$$Precision = \frac{TP}{TP + FP} \quad (25)$$

查全率代表在实际为正例的样本中, 被预测正确的样本所占的比例。*Recall* 可以表示为:

$$Recall = \frac{TP}{TP + FN} \quad (26)$$

模型 *P-R* 曲线以查准率为纵轴, 查全率为横轴。曲线与坐标轴所围成的面积为平均精度 *AP*:

$$AP = \int_0^1 Precision_{(t)} dt \quad (27)$$

对于 *N* 个类别的检测目标, 各类平均精度均值表示为式:

$$mAP = \frac{\sum_{n=1}^N AP_n}{N} \quad (28)$$

漏检率代表在实际为正例的样本中, 被预测为负样本的正例所占的比例。*Miss rate* 则表示为:

$$FNR = \frac{FN}{TP + FN} \quad (29)$$

误检率代表在实际为负例的样本中, 被预测为正样本的负例所占的比例。*False alarm* 则表示为:

$$FPR = \frac{FP}{TN + FP} \quad (30)$$

3.2 消融实验

数据集中包含少量的小尺度障碍物和裂缝, 其余大部分缺陷的尺度相差不大, 因而将 Inner-MPDIoU 中的取值范围定在 $[0.5, 0.9]$, 以 0.1 为间隔进行实验, 确定其取值, 结果如表 1 所示。

表 1 消融实验 1

<i>IoU</i>	<i>R_s</i>	<i>P</i> /%	<i>R</i> /%	<i>mAP</i> @0.5/%	<i>FNR</i> /%	<i>FPR</i> /%
MPDIoU	—	90.9	87.7	90.4	12.3	20.3
Inner-MPDIoU	0.5	91.2	88.9	90.9	11.1	19.4
	0.6	91.8	90.1	91.3	9.9	19.0
	0.64	91.9	90.4	91.6	9.6	18.6
	0.7	91.7	89.7	91.1	10.3	18.9
	0.8	91.4	89.5	90.9	10.5	19.1
	0.9	91.2	88.1	90.6	11.9	19.8

通过实验, 由表 1 可知, 当比例因子的取值在 0.6~0.7 之间时, 模型的检测准确率高于其他取值情况, 研究将比例因子取值范围进一步缩小在 0.6~0.7 之间, 以 0.01 为间隔进行实验, 最终发现比例因子取值为 0.64 时, 平均准确率最高为 91.6%。为了验证完全改进后的模型性能, 本文使用同一批数据集进行消融实验 2, 方法 1、2、3、4 分

别是 Baseline (YOLOv8)、Baseline + RFACConv + C2F_RFACConv+ HAHLFF、Baseline + Inner-MPDIoU 和本文模型, 实验结果如表 2 所示。

表 2 消融实验 2 %

方法	<i>P</i>	<i>R</i>	<i>mAP@0.5</i>	<i>FNR</i>	<i>FPR</i>
1	91.5	87.7	90.2	12.3	20.3
2	92.1	89.8	93.1	10.2	19.5
3	91.9	90.4	91.6	9.6	18.6
4	92.3	90.9	93.9	9.1	17.6

由表 2 可知, 方法 2 在 Baseline 中加入感受野空间注意力卷积和混合注意力高低阶特征融合网络后, 模型学习缺陷图像全局上下文信息的能力得到增强; 方法 3 在 Baseline 加入 Inner-MPDIoU 后, 克服了边界框重叠情况、中心点距离和宽高偏差等因素的影响, 对于不同尺寸的缺陷, 模型的边界框回归定位更加准确; 方法 4 为完全改进的本文模型, 综合了方法 2 和方法 3 的优势, 获得 93.9% 的检测准确率, 以及更低的漏检率和误检率, 检测准确率较改进前提升了 3.7%, 证明了本文改进模型的有效性。

3.3 对比实验

为了进一步测试本文模型的性能, 选取 YOLOv5、Baseline 以及文献 [8, 10, 11, 23] 的方法与本文模型作对比实验, 设置相同的实验环境, 并采用同一数据集进行训练和测试, 评价指标同消融实验, 结果如表 3 所示。为了更加直观的展示本文模型的优势, 本文选取 Baseline 和本文模型对同一测试集数据进行检测, 结果如图 8 所示。

表 3 对比实验 %

模型	<i>P</i>	<i>R</i>	<i>mAP@0.5</i>	<i>FNR</i>	<i>FPR</i>
YOLOv5	89.8	85.8	88.7	14.2	22.5
Improved-YOLOv5 ^[8]	90.3	86.6	89.3	13.4	21.9
Improved-YOLOv5 ^[10]	89.9	86.1	89.0	13.9	22.1
Improved-YOLOX ^[11]	89.2	87.0	88.9	13.0	21.6
YOLOv7 ^[23]	90.1	87.4	89.1	12.6	22.0
Baseline	91.5	87.7	90.2	12.3	20.3
本文模型	92.3	90.9	93.9	9.1	17.6

如图 8, 图 (a) 中使用 1 号五角星标记的是沉积物缺陷, 通过对比图 (d) 和图 (g) 可以发现, 本文模型的沉积物检测框明显比 Baseline 的检测框的定位更加准确; 图 (b) 中使用 2 号和 3 号五角星标记的都是裂缝缺陷, 4 号五角星标记的则是脱节缺陷, 通过对比图 (e) 和图 (h) 可以发现, 本文模型检测出了图中的全部缺陷, 然而 Baseline 则漏检了 3 号五角星标记的裂缝缺陷; 图 (c) 中使用 5 号五角星标记的是障碍物缺陷, 通过对比图 (f) 和图 (i) 可以发现, 本文模型能够准确的检测出障碍物缺陷, 然而 Baseline 则是将障碍物误检为沉积物缺陷。综上可知, 本文模型在缺陷检测准确率、漏检和误检方面都要优于 Baseline。

结合图 8 和表 3 的实验结果可知, 本文模型的检测性能

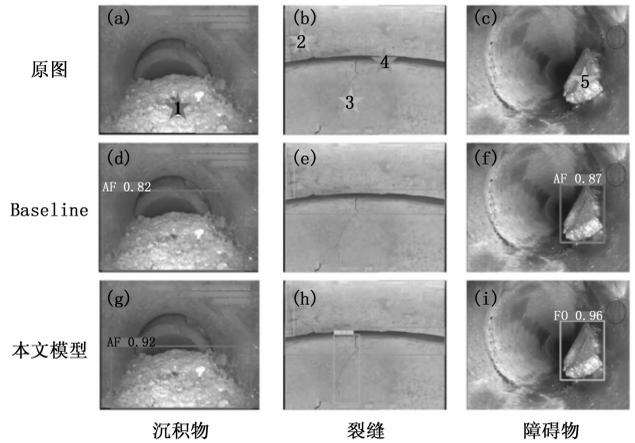


图 8 本文模型与 YOLOv8 的检测效果对比图

优于其他对比模型, 分析原因如下: YOLOv5 在训练过程中使用 Anchor-Base 的方式, 使用的耦合头虽然具有较少的参数量和计算量, 但却削弱了模型提取特征的能力; 文献 [8, 10, 11] 所提方法在 YOLOv5 或 YOLO X 的基础上进行了相应的改进, 在检测准确率、漏检率和误检率方面均优于原始模型, 文献 [23] 方法同样是使用 Anchor-Base, 同时使用带有辅助头的训练来提升检测精度, 但冗余明显增加, 其综合检测性能略好于 YOLOv5; Baseline 虽然拥有更丰富的梯度流, 性能优势明显, 但未能充分学习图像全局上下文信息, 面对不同尺寸的缺陷, 边界框回归准确率; 本文模型针对 Baseline 存在的问题进行了优化, 相比之下, 平均准确率提升 3.7%, 最终达到 93.9%, 漏检率和误检率分别降低 3.7% 和 2.7%, 综合性能表现最好, 进一步证明了本文所提模型的有效性。

4 结束语

本文针对 YOLOv8 检测排水管道缺陷准确率低、目标定位不准确的问题, 通过引入感受野空间注意力卷积, 构建 C2F_RFACConv, 并提出混合注意力高低阶特征融合网络, 进一步增强了模型学习缺陷图像全局上下文信息的能力; 综合分析了边界框重叠情况、中心点距离、宽高偏差等因素对边界框回归的影响, 设计 Inner-MPDIoU 损失函数, 适应不同尺寸的缺陷, 提高目标边界框回归准确率。实验结果表明, 本文模型的平均准确率达到 93.9%, 漏检率和误检率分别为 9.1% 和 17.6%, 综合性能优于其他对比模型。未来, 将对检测模型的轻量化部署和缺陷实时检测作进一步研究, 并增加缺陷检测种类。

参考文献:

[1] JOAKIM B H, MEYSAM M, et al. Multi-task classification of sewer pipe defects and properties using a cross-task graph neural network decoder [C] //2022 IEEE/CVF Winter Conference on Applications of Computer Vision, 2022: 1441 - 1452.
 [2] 中华人民共和国住房和城乡建设部. 城镇排水管道检测与评估技术规程: GB/CJJ181-2012 [S]. 北京: 中国建筑工业出版社, 2012.

- [3] SUN L P. Deep learning-assisted automated sewage pipe defect detection for urban water environment management [J]. *Science of The Total Environment*, 2023, 882: 163562.
- [4] LI D W, XIE Q, YU Z H, et al. Sewer pipe defect detection via deep learning with local and global feature fusion [J]. *Automation in Construction*, 2021, 129: 103823.
- [5] MA D, FANG H Y, et al. Automatic defogging, deblurring, and real-time segmentation system for sewer pipeline defects [J]. *Automation in Construction*, 2022, 144: 104595.
- [6] KUMAR S S, WANG M Z, et al. Deep learning-based automated detection of sewer defects in CCTV videos [J]. *Journal of Computing in Civil Engineering*, 2020, 34: 1943–5487.
- [7] REDMON J, FARHADI A. Yolov3: an incremental improvement [J/OL]. *ArXiv Preprint ArXiv*: 1804.02767, 2018. <https://arxiv.org/abs/1804.02767>.
- [8] CHANMI O, DANG L M, et al. Robust sewer defect detection with text analysis based on deep learning [J]. *IEEE Access*, 2022, 10: 46224–46237.
- [9] SITU Z X, et al. A transfer learning-based YOLO network for sewer defect detection in comparison to classic object detection methods [J]. *Developments in the Built Environment*, 2023, 15: 100191.
- [10] 完颜健飞, 江雅馨, 徐晓龙, 等. 基于改进 YOLOv5 的城市地下管网缺陷识别算法 [J]. *计算机测量与控制*, 2024, 32 (11): 258–264.
- [11] 陆绮荣, 丁 昕, 梁雅雯. 基于改进 YOLOX 的地下排水管道缺陷识别算法 [J]. *电子测量技术*, 2022, 45 (21): 161–168.
- [12] ZHOUF Y, CHAO Y S, et al. A small sample nonstandard gear surface defect detection method [J]. *Measurement*, 2023, 221: 113472.
- [13] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior [C] //2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009: 1956–1963.
- [14] ZHANG X, LIU C, et al. RFACConv: innovating spatital attention and standard convolutional operation [J/OL]. *ArXiv Preprint ArXiv*: 2304.03198v1, 2023. <https://arxiv.org/abs/2304.03198>.
- [15] ZHANG H, XU C, ZHANG S J. Inner-IoU: more effective intersection over union loss with auxiliary bounding box [J/OL]. *ArXiv Preprint ArXiv*: 2311.02877, 2023. <https://arxiv.org/abs/2311.02877>.
- [16] MA S L, XU Y. MPDIoU: a loss for efficient and accurate bounding box regression [J/OL]. *ArXiv Preprint ArXiv*: 2307.07662, 2023. <https://arxiv.org/abs/2307.07662>.
- [17] TANG L F, ZHANG H, et al. Rethinking the necessity of image fusion in high-level vision tasks: practical infrared and visible image fusion network based on progressive semantic injection and scene fidelity [J]. *Information Fusion*, 2023, 99: 101870.
- [18] LI J F, WEN Y, HE L H. ScConv: spatial and channel reconstruction convolution for feature redundancy [C] //2023IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 6153–6162.
- [19] LIU J X, LIU W, et al. Attention-based adaptive structured continuous sparse network pruning [J]. *Neurocomputing*, 2024, 590: 127698.
- [20] HUANG Z B, LIU J Y, et al. ReCoNet: recurrent correction network for fast and efficient multi-modality image fusion [C] //European Conference on Computer Vision, 2022: 539–555.
- [21] HAURUM J B, MOESLUND T B. Sewer-ML: a multi-label sewer defect classification dataset and benchmark [C] //2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13451–13462.
- [22] ADRIEN P, ISABELLE G, ANNE C L, et al. CodaLab competitions: an opensource platform to organize scientific challenges [J]. *Journal of Machine Learning Research*, 2023, 24 (198): 1–6.
- [23] WANG C Y, ALEXEY B, MARK L H Y. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [J/OL]. *ArXiv Preprint ArXiv*: 2207.02696, 2022. <https://arxiv.org/abs/2207.02696>.
- [22] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [C] //Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3–19.
- [23] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132–7141.
- [24] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713–13722.
- [25] CHARFI I, MITERAN J, DUBOIS J, et al. Optimized spatio-temporal descriptors for real-time fall detection: comparison of support vector machine and Adaboost-based classification [J]. *Journal of Electronic Imaging*, 2013, 22 (4): 041106.
- [18] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1–9.
- [19] ZHANG R. Making convolutional networks shift-invariant again [C] //International Conference on Machine Learning. PMLR, 2019: 7324–7334.
- [20] LU W, CHEN S B, TANG J, et al. A robust feature down-sampling module for remote sensing visual tasks [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1–12.
- [21] ESCARDINO A, MESTRE S, BARBA A, et al. Synthesis mechanism of an iron-chromium ceramic pigment [J]. *Journal of the American Ceramic Society*, 2000, 83 (1): 29–32.