

基于改进 EfficientDet 模型的城市地下管道缺陷检测

马小骞¹ 李波^{1*} 罗理¹
MA Xiaoqian LI Bo LUO Li

摘要

地下管道是城市基础设施中不可或缺的一部分,及时排除地下管道缺陷,可以有效保障城市地下管道系统的正常运转。地下管道缺陷样本背景复杂、缺陷尺度变化较大、形态各异、噪声多,给有效检测带来了很大挑战。目前目标检测方法在地下管道缺陷检测方面存在不能兼顾速度和精度的问题,且模型参数量大。文章在 EfficientDet-D0 模型基础上,提出了一种改进模型,引入 DySample 动态上采样,更好地适应复杂场景和重建特征图,并对数据集采取实时数据增强。通过改进模型比原始模型获得了更高的检测精度,同时保证了实时推理的速度,参数量小,平均精度提升了 2.75%,且未额外引入参数量,和主流检测模型相比也有一定优势。

关键词

城市地下管道;缺陷检测;DySample 动态上采样;EfficientDet 模型

doi: 10.3969/j.issn.1672-9528.2025.10.005

0 引言

地下管道系统是城市基础设施中不可或缺的一部分,保障着城市经济建设和居民日常生活的正常运转^[1]。随着使用年限的不断增长、自然老化和外力损坏等原因,可能会导致地下管道产生各种缺陷,需要及时排除和修复。

由于人工检测会存在较大的安全风险,随着技术的发展,地下管道缺陷检测方式有闭路电视(closed circuit television, CCTV)^[2]、激光^[3]、声呐^[4]等。Dirksen 等人^[5]研究指出,在识别地下管道缺陷过程中,有 25% 的缺陷无法通过视觉手段进行正确识别,表明传统人工方法在复杂情况下具有局限性。

近年来,越来越多的人利用深度学习算法进行地下管道缺陷的检测,并取得了不错的成果。Ran 等人^[6]提出的改进 U-Net 的具有 CBAM 注意力机制的智能检测算法在检测任务中表现良好。Zhou 等人^[7]基于 FCOS 模型引入轻量的 MobileOne 网络,相比于基线模型,平均精度提升了 1.83%,模型参数量也有所下降。Wang 等人^[8]将对合模块(Involution Module)和 GSConv 简化模型分别引入骨干网络和特征融合网络,改进后的 YOLOv5 精度得到有效提升。杨帆等人^[9]提出了改进的 YOLOv8 排水管道缺陷检测算法,添加 Coordinate Attention 注意力机制,增强了算法对排水管道细小缺陷的识别能力。冯丽丹等人^[10]将轻量化网络 MobileNetV3 和 YOLOv7 网络框架相结合,增加 BiFPN 特征金字塔结构以提高精确度,减少了参数计算量,且平均准确

率有所提高。Zhao 等人^[11]将倒置残差移动模块(iRMB)集成到骨干网络中,引入移位运算符来提高性能,为识别腐蚀类别的缺陷提供了一种准确的方法。

以上研究表明了利用深度学习模型进行地下管道缺陷检测的有效性与可行性。然而,随着卷积层的增加,深度学习模型在地下管道缺陷检测方面存在难以兼顾速度和精度的问题,且模型参数量往往较大。

本文的主要贡献包括:(1)选用轻量化 EfficientDet 模型作为基线模型,引入 DySample 动态上采样模块,以更好地适应复杂场景和重建特征图,提高检测精度;(2)为了保证检测算法的准确性和鲁棒性,在训练中采用了实时数据增强方法;(3)提出了一种检测精度高的轻量级的改进 EfficientDet 模型,满足实时要求。

1 相关工作

1.1 构建数据集

本文实验数据构建基于 CVPR 会议公开的 Sewer-ML 基准数据集^[12]。由于该数据集图片中未标注管道缺陷的具体位置区域,因此无法直接用来进行管道缺陷目标检测的训练建模。在本实验中,为了训练目标检测模型还需人工标注缺陷区域。本文使用 LabelImg 完成标注工作。

本文选取该数据集中最为常见的 6 类缺陷进行研究,分别是暗接、沉积、错口、结垢、树根和障碍物。这些缺陷在实际地下管道运行中出现的可能性较高,从一定程度上影响了地下管道的实际使用,因此,地下管道日常维护应该优先对这些地下管道缺陷进行及时有效检测。

1. 中南民族大学计算机科学学院 湖北武汉 430074

本文实验数据集共构建有标注缺陷区域位置的图片样本 1 685 张。训练集和验证集的总占比为 90%，测试集占 10%；训练集包含 1 364 张样本，验证集包含 152 张样本，训练集和验证集的比例为 9:1，测试集包含 169 张样本。

1.2 数据增强

由于地下管道背景复杂、光线昏暗且有的缺陷严重程度不同，有时在图像中仅占据较小区域，所以为了保证检测算法的准确性和鲁棒性，本文在训练时，对数据集进行了实时数据增强。数据增强操作包括：随机缩放、随机翻转、随机色域变化，在训练时动态增强，不需要额外存储空间，且随之更加灵活，可以有效防止模型过拟合，增加模型的泛化性。

2 EfficientDet 模型架构

EfficientDet 是一个轻量且高效的模型，可以在更少的参数下取得更高的精度。EfficientDet-D0 是 EfficientDet 参数最少的版本，其整体架构由三部分组成：基于 EfficientNet 的主干网络（Backbone）、双向特征金字塔（BiFPN）以及检测头（Detection Head）。

主干网络主要用于对输入图片中的特征进行提取，EfficientDet-B0 由 16 个 MBConvBlock 堆叠构成，MBConvBlock 的设计思路是先利用 1×1 卷积升维，在 3×3 或者 5×5 深度卷积后增加通道的注意力机制，最后利用 1×1 卷积降维后增加一个大残差边，深度卷积的使用大大降低了计算量和参数。

经过 EfficientNet 后可以提取到 5 个有效特征图 P3、P4、P5、P6、P7，随后传输入到 BIFPN 双向特征金字塔中。BIFPN 双向特征金字塔用于对提取的特征进行特征融合，每个 BIFPN 块有两条路径，首先自上而下，从高层特征图上采样后和低层特征图融合；其次自下而上，从低层特征图下采样后和高层特征图融合。多尺度特征图的加权融合使得模型自动学习不同尺度特征层的重要性，融合不同尺度的上下文信息，更好地捕获小物体的细节特征，提高检测的精度。

3 本文提出的改进模型

本文提出的整体模型改进网络结构如图 1 所示。常见的上采样器有最近邻插值和双线性插值，它们都有特定的规则，对低分辨率特征进行插值，忽略了特征图中的语义信息。针对上述问题，本文在 BIFPN 模块中用动态上采样 DySample^[13] 对最近邻插值方法进行替换，根据特征图动态学习采样点，可以更好地适应复杂场景和重建特征图，提高检测精度，添加 DySample 模块后 BiFPN 结构图如图 2 所示。

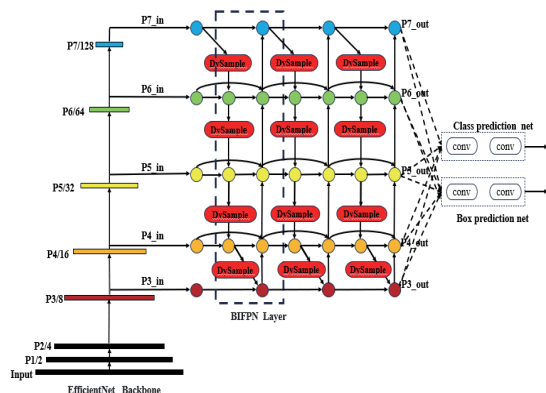


图 1 改进的 EfficientDet 总体网络结构

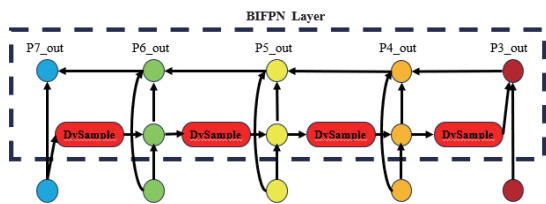


图 2 改进后的 BIFPN 结构图

DySample 是一个轻量级、高效率的动态上采样方法，受可变形卷积网络以及 Deformable DETR^[14] 的启发，根据不同的特征图，其会自适应调整采样点的位置，能够生成内容感知的采样位置，每个采样点会有 4 个偏移量。过程是特征图先经过线性层产生偏移量，然后根据初始化采样点产生采样点集合，最后根据这个采样点集合对输入特征进行重新采样，得到高分辨率特征图，用公式分别表示为：

$$O = \text{linear}(X) \quad (1)$$

$$S = G + O \quad (2)$$

$$X' = \text{grid_sample}(X, S) \quad (3)$$

式中： X 为特征图； O 为偏移量； G 为原始采样网格； S 为采样点集合； X' 为重新采样的结果；linear 为线性变换；grid_sample 为采样函数，根据 S 中指定的坐标位置，从输入特征图 X 中采样。

但如果偏移范围太大，采样点在边界处重叠，重复的特征值会造成采样混乱，特征无法正确表达。因此，文献 [13] 通过使用通过线性投影输入特征来生成动态范围因子，配合 sigmoid 函数和 0.5 的静态因子，增加偏移量的灵活性，运算过程为：

$$O = 0.5 \text{sigmoid}(\text{linear}_1(X)) \cdot \text{linear}_2(X) \quad (4)$$

4 实验结果及讨论

4.1 实验环境与超参数设置

本文实验在 AutoDL 官网租用的分布式计算集群中部署深度学习训练环境，实验环境为 Ubuntu20.04，以 PyTorch 为框架，以 Python 编写代码。

此外，权值因随机取值，特征提取效果不明显，为保证训练效果，在训练时使用预训练权重。为减少计算资源和训练时间，采用冻结主干网络训练策略，冻结 50 个 epoch，此时 batch_size 为 8，50 轮后 batch_size 为 4，输入的图片大小为 512 px×512 px，总共训练 200 个 epoch，因为在主干网络新增的模块由于没有预训练所以不进行冻结，初始学习率设置为 3e-4，采用余弦退火（Cosine Annealing）学习率调度，逐渐降低学习率，采用 adam 优化器，动量参数设置为 0.9。

4.2 评价指标

本文采用评价指标包括：精确率（Precision）、召回率（Recall）、单类别平均精度（average precision, AP）以及所有类别的平均精度均值（mean average precision, mAP），为了衡量模型的性能，本文采用 mAP、参数量以及 FPS 来对比不同模型的性能。

$$\text{Precision}=\frac{\text{TP}}{\text{TP}+\text{FP}}$$

（5）

$$\text{Recall}=\frac{\text{TP}}{\text{TP}+\text{FN}}$$

（6）

$$\text{AP}=\int_0^1 P(R)\text{d}R$$

（7）

$$\text{mAP}=\frac{\sum_i^C \text{AP}}{C}$$

（8）

$$\text{FPS}=\frac{1}{\text{FrameTime}}$$

（9）

式中：TP 为检测正确的地下管道缺陷目标数；FP 为模型检测为管道缺陷而实际为背景的目标数；FN 为漏检的地下管道缺陷目标数；精确度（Precision）为正确判定的阳性（本文中为管道缺陷）样本数量与全部阳性预测结果的比值；召回率（Recall）为计算时取正确检出的阳性样本数量与实际阳性样本总量的比例；平均准确率（AP）为精确率和召回率曲线下的面积；C 为类别的个数；FrameTime 为表示处理单帧画面所耗费的时间。

4.3 对比实验

通过分析不同模型的检测指标来验证不同模型在地下管道缺陷检测的优势，如表 1 和表 2 所示。

表 1 不同模型检测指标结果对比

模型	AP_CK	AP_CJ	AP_JG	AP_ZW	AP_AJ	AP_SG
FCOS	73.94	92.57	84.36	84.73	90.65	73.30
CenterNet	64.62	88.03	74.37	79.95	89.88	61.35
FasterRCNN	72.63	90.34	71.38	80.90	90.21	65.39
RT-DETR	84.20	89.80	69.90	83.20	81.70	75.10
EfficientDet	83.36	95.04	76.43	81.12	76.89	78.32
本文改进模型	82.79	95.55	85.97	88.06	75.70	79.57

注：表头中 AP_CK、AP_CJ、AP_JG、AP_ZW、AP_AJ 和 AP_SG 分别表示错口缺陷、沉积缺陷、结垢缺陷、障碍物缺陷、暗接缺陷和树根缺陷检测的平均精度。

表 2 不同模型检测指标结果对比

模型	mAP/%	参数 /10 ⁶	FPS/(帧·s ⁻¹)
FCOS	83.26	32.13	33.5
CenterNet	76.37	32.66	72.2
FasterRCNN	78.47	23.50	13.9
RT-DETR	80.60	42.00	79.1
EfficientDet	81.86	3.90	47.3
本文改进模型	84.61	3.90	42.9

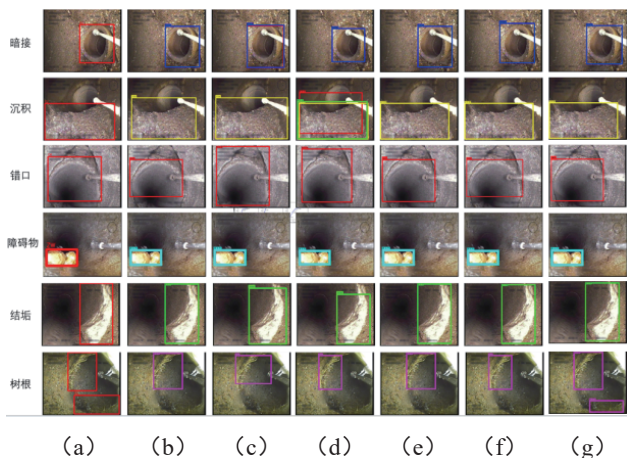
在实时性敏感场景的模型评估中，两阶段检测模型 Faster R-CNN 因级联式推理流程导致计算延迟，其处理速率不足 20 帧 /s。相较之下，单阶段无锚框检测模型 FCOS 与轻量级模型 EfficientDet 达到了 30 帧 /s 以上的实时处理能力。值得注意的是，基于关键点估计的 CenterNet 为 72.2 帧 /s 与 Transformer 驱动的 RT-DETR 为 79.1 帧 /s，表现更优，其并行化计算特性可满足工业级高帧率检测需求。

从实验结果来看，CenterNet 和 FasterRCNN 检测指标中有个别只有 60%。且模型的 mAP 低于 80%，说明检测效果并不好。而 FCOS、EfficientDet 和 RT-DETR 的 mAP 都高于 80%，相比前面两个模型检测效果较好，且检测指标远远超过，说明检测效果比前两个模型好。EfficientDet 虽然有个别指标比 FCOS 低，但树根和沉积缺陷检测效果比 FCOS 高。相比于原始模型，本文的改进模型除了错口缺陷和暗接缺陷略低，其他缺陷指标都大幅提升，在所有比较模型中 mAP 最高，达到了 84.61%。

对于模型的参数，RT-DETR 最高，其次是 CenterNet 和 FCOS，其较大的参数量使其在资源受限的环境中不太适合使用，其他模型参数适中，EfficientDet 和本文改进模型参数最低，仅有 3.9×10⁶，说明模型轻量化，适合资源受限的应用场景。综上所述，改进后的模型在不引入额外参数的同时相较于原始模型有很大的提升，检测效果最好，同时还保证了实时推理的速度。验证引入 DySample 动态上采样加强了缺陷检测的能力。

为直观评估不同模型在地下管道缺陷检测中的能力，采用可视化对比方法如图 3 所示。图 3(a) 为缺陷区域的真实框，从上到下依次为暗接、沉积、错口、障碍物、结垢及树根缺陷的检测框；RT-DETR（图 3（b））展现出较优的类别识别能力，缺陷类型均被正确检测，但检测框的定位稍有偏差；Faster RCNN（图 3（c））虽能定位缺陷区域，但存在预测框尺寸普遍偏大且在预测错口缺陷时误检的可能。检测效果低于 RT-DETR 效果；CenterNet（图 3（d））在预测框生成质量方面表现欠佳，主要缺陷为框体尺寸过小，导致定位不完整，特别是对沉积缺陷存在严重混淆和误检现象；FCOS（图 3（e））实现了全类别缺陷识别，检测框定位相对较好；

EfficientDet 模型 (图 3 (f)) 在多数缺陷类型中表现出均衡的检测能力。相较之下, 本文的改进模型 (图 3 (g)) 其预测框质量整体优于其他对比模型, 检测出了其他所有模型均未检测出的细小树枝缺陷。



注: (a) 为真实框; (b) 为 RT-DETR; (c) 为 FasterRCNN; (d) 为 CenterNet; (e) 为 FCOS; (f) 为 EfficientDet; (g) 为本文改进模型。

图 3 不同模型检测效果对比图

5 结论

本文提出了一种改进的 EfficientDet 模型, 适应地下管道复杂场景, 更好地重建特征图, 提高了检测精度。通过对比实验验证, 本文提出的方案相比原始模型获得了更高的检测精度, 保证了地下管道缺陷检测的实时性, 平均精度提升了 2.75%, 具有较好的实际应用前景。

参考文献:

- [1]SHI J Y, DI D Y, YU H R. A siltation diagnosis method for underground drainage pipelines based on multi-source data fusion and MLSTM-SVM[C]// 2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT).Piscataway:IEEE, 2024:2271-2275.
- [2]HEO G, JEON J, SON B. Crack automatic detection of CCTV video of sewer inspection with low resolution[J]. KSCE journal of civil engineering, 2019, 23: 1219-1227.
- [3]YUE Z Y, SUN H L, ZHONG R F, et al. Measurement of tunnel clearance convergence using mobile laser detection technology[J]. Canadian journal of remote sensing, 2021, 47(1): 100-118.
- [4]ALI H, CHOI J-H. A review of underground pipeline leakage and sinkhole monitoring methods based on wireless sensor networking[J]. Sustainability, 2019, 11(15): 4007.
- [5]DIRKSEN J, CLEMENS F, KORVING H, et al. The consistency of visual sewer inspection data[J]. Structure and infrastructure engineering, 2013, 9(3): 214-228.

- [6]RAN Q Y, WANG N N, WANG X Y, et al. Improved U-Net drainage pipe defects detection algorithm based on CBMA attention mechanism[C]// 2024 9th International Conference on Electronic Technology and Information Science (ICETIS). Piscataway:IEEE,2024:342-347.
- [7]ZHOU B, LI B, LAN W F, et al. SDH-FCOS: an efficient neural network for defect detection in urban underground pipelines[J]. Computers, materials & continua, 2024, 78(1): 633-652.
- [8]WANG T, LI Y H, ZHAI Y D, et al. A sewer pipeline defect detection method based on improved YOLOv5[J]. Processes, 2023, 11(8): 2508.
- [9]杨帆, 刘如飞, 刘扬胜, 等. 基于改进 YOLOv8 的城市排水管道缺陷检测算法研究 [J]. 给水排水, 2024,60(8): 120-125.
- [10]冯丽丹, 王闯, 祁军, 等. 基于改进 YOLOv7 的管道缺陷检测技术研究 [J]. 辽宁石油化工大学学报, 2024, 44(4): 82-90.
- [11]ZHAO Q, WANG G J. Multi-architecture optimization of pipeline inner wall defect detection algorithm based on YOLOv8[J]. Measurement, 2025, 242: 116305.
- [12]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 (CVPR).Piscataway: IEEE,2021:13451-13462.
- [13]LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision.Piscataway: IEEE, 2023: 6027-6037.
- [14]ZHU X Z, SU W J, LU L W, et al. Deformable DETR: deformable transformers for end-to-end object detection[EB/OL].(2023-03-18)[2025-05-25].https://doi.org/10.48550/arXiv.2010.04159.

【作者简介】

马小骞 (2000—), 男, 云南弥勒人, 硕士研究生, 研究方向: 图像处理和深度学习。

李波 (1975—), 通信作者 (email:libo@mail.scuec.edu.cn), 男, 湖北宜都人, 博士, 教授, 研究方向: 机器学习与模式识别。

罗理 (2001—), 女, 陕西西安人, 硕士研究生, 研究方向: 图像处理和深度学习。

(收稿日期: 2025-04-21 修回日期: 2025-10-13)