

一种基于生成对抗网络和改进U-Net的排水管道缺陷识别模型

刘 勇¹, 李有淋²

(1. 四川省自然资源测绘地理信息有限责任公司, 四川 成都 610212; 2. 西南石油大学 地球科学与技术学院, 四川 成都 610500)

摘 要: 通过集成生成对抗网络图像生成及增强技术, 利用卷积注意力模块 CBAM (Convolutional Block Attention Module) 改进 U-Net 网络, 提出了一种排水管道缺陷高精度识别模型, 以准确自动对排水管道影像缺陷进行分割识别。该模型取得了满意的识别结果, 在验证集中平均像素精度 mPA (Mean Pixel Accuracy)、平均交联度 mIoU (Mean Intersection over Union) 分别为 91.40%、85.02%。结果表明, 利用生成对抗网络对采集数据中较少的缺陷类别图像生成及增强, 能够大大提高少样本缺陷识别的准确度, 数据增强后对 PSPNet、U-Net 以及不同主干网络的 DeepLabv3+ 网络得到大幅度提升, mPA、mIoU 平均增长 20%~30%。加入 CBAM 注意力模块能够提升 U-Net 的分割能力, 使模型具有更好的像素级分割能力。

关键词: 排水管道缺陷; 生成对抗网络; U-Net; CBAM

中图分类号: TU992

文献标志码: A

文章编号: 1009-7716(2026)01-0033-06

A Defect Recognition Model for Drainage Pipes Based on Generative Adversarial Networks and Improved U-Net

LIU Yong¹, LI Youlin²

(1. Sichuan Natural Resources Mapping and Geographic Information Co., Ltd., Chengdu 610212, China; 2. School of Geosciences and Technology, Southwest Petroleum University, Chengdu 610500, China)

Abstract: By integrating GAN (Generative Adversarial Network) for image generation and enhancement, and improving the U-Net network with the Convolutional Block Attention Module (CBAM), a high-precision recognition model for drainage pipeline defects is proposed to accurately and automatically segment and identify the image defects of drainage pipes. The model achieves satisfactory recognition results. The mPA (Mean Pixel Accuracy) is 91.40% and the mIoU (Mean Intersection over Union) is 85.02% on the validation set. The results indicate that the generation and enhancement of images with fewer defect categories in the collected data by using GAN can greatly improve the accuracy of defect recognition with fewer samples. After data augmentation, the DeepLabv3+ networks of PSPNet, U-Net and different backbone networks have been significantly improved, with an average increase of 20%~30% in mPA and mIoU. Adding the CBAM attention module can enhance the segmentation ability of U-Net and enable the model to have better pixel-level segmentation ability.

Keywords: defects of drainage pipe; generate adversarial networks; U-Net; CBAM

0 引 言

排水管道系统作为城市重要的基础设施, 在保障防洪减灾、维持城市正常运行中起到至关重要的作用^[1]。然而, 由于污水的侵蚀、排水管道的不均匀沉降、施工问题、排水管道老化等因素, 往往会出现不同程度的缺陷^[2]。常见排水管道检测技术有闭环

电视检测(CCTV)、激光探测、潜望镜探测、探地雷达监测、红外热成像等^[3]。随着人工智能的发展, 涌现出许多优秀的深度学习算法, 并越来越多地用于更准确和自动的排水管道缺陷检测^[4-5]。图像分割方面, 在提供缺陷的形状、位置和密度等像素级描述方面更为准确和详细^[6-7]。其中在排水管道缺陷识别方面已经取得较为显著的成果, 吴哲豪等人提出一种基于 EfficientViT 改进 YOLOv8 网络的排水管道缺陷智能检测模型, 有效提高了检测速度^[8]。凌同华等人构建了轻量化 YOLO-Pipe 并与 ByteTrack 追踪算

收稿日期: 2025-09-20

作者简介: 刘勇(1977—), 男, 硕士, 正高级工程师, 从事测绘工程、地下管网探测及供排水管网修复等工作。

法融合,同时兼顾了精度和检测速率^[9]。

然而,对于数据集本身的细化增强工作目前少有研究。由于各类缺陷数据分布不均匀导致某些少样本缺陷识别效果较差,识别总体精度不高,从而很难满足现场应用的需求。因此,本文使用了生成对抗网络StyleGAN3^[10]以及传统图像变换混合增强手段。弥补了采集数据的不均匀分布,降低模型过拟合风险,并且使用了双重注意力机制模块CBAM^[11],改进了U-Net网络,获得更高的识别精度。

1 数据采集及处理

1.1 原始数据采集

本文使用CCTV管道机器人对四川省遂宁市部分街道排水管道缺陷图像进行采集,对采集后的图像进行清洗工作,筛选清晰度良好,具有排水缺陷的图像进行提取,将提取后的图像去除文字水印,并按照图像中心512×512像素大小进行裁剪,将裁剪后的图像进行排水管道缺陷的分类确定。按照中国住房和城乡建设部制定的行业标准《城镇排水管道检测与评估技术规程》(CJJ 181—2012)以及结合采集图像实际缺陷易发情况确定排水管道缺陷种类,分别确定5种排水道缺陷:(1)破裂,排水管道的外部压力超过自身的承受力致使排水管道发生破裂,如图1(b)所示;(2)障碍物,排水管道内影响过流的阻挡物,如图1(c)所示;(3)沉积,杂质在排水管道底部沉淀淤积,如图1(d)所示;(4)变形,排水管道挤压造成形状变异,如图1(e)所示;(5)树根,单根树根或是树根群自然生长进入排水管道,如图1(f)所示,以及正常排水管道图像如图1(a)所示。这5类排水管道缺陷具有不同的特征,需要一种高效、鲁棒的检测模型进行识别。根据现场数据采集情况,共选取1 400张排水管道缺陷图像数据作为基础数据,其中利用开源Python库Labelme对排水管道缺陷影像进行标注工作,随机按照8:2划分作为原始数据集的训练集和验证集。其中对各类标签的像素点出现频率进行统计,如表2所示,其中沉积出现频率最高,像素占比38%,随后依次是障碍物(像素占比21%)、变形(像素占比20%)、树根(像素占比11%)、破裂(像素占比最少,为10%)。当一张图像中存在多类缺陷时,每种缺陷像素数量单独计数。

1.2 增强数据处理

使用StyleGAN3进行数据生成并增强,其中对像素数量较低的缺陷类型进行选择性图像生成,

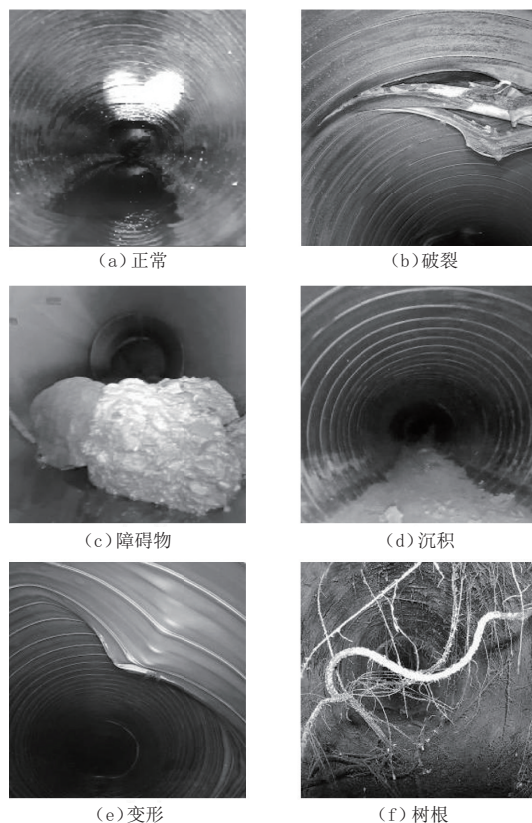


图1 排水管道缺陷分类示例

StyleGAN3相关训练配置如表1所示, cfg表示StyleGAN3-T(等变平移), batch-gpu表示图像总批次大小, gamma表示R1正则化权重, kimg表示一个总的迭代次数, snap表示间隔多久打印一次训模信息。将经过预处理的1 400张图像进行StyleGAN3训练,生成的排水管道扩充图像如图2所示,具有较高的真实性。随后利用随机数种子1-500放入训练模型中生成500张扩充图像,并选择生成图像中的380张质量较高并且属于缺陷像素数量较少类别的图像进行缺陷标注,并放入真实样本中,共计获得1 780张排水管道图像。其中各类缺陷像素点分布情况,如表2所示,各类缺陷的像素数量均有提升,由于沉积像素数量原本较高,占据一定优势,因此主要提升在于树根、变形、破裂、障碍物,其中变形占比提升到25%,树根和破裂分别提升至12%、11%,沉积占比降低至33%。随后将1 780张图像以及标签图像,按照色彩扰动、中心裁剪、水平翻转、随机噪声、垂直翻转五类进行扩充增强,如图3所示,共计获得10 680张排水管道缺陷图像以及对应的标签图像,按照8:2划分训练集(即8 544张)和验证集(2 136张)。本研究使用pytorch作为深度学习训练框架,搭建U-Net网络结构,另外采用随机动量梯度下降优化器

(SGDM)作为梯度下降函数,采用交叉熵损失函数作为损失函数。此外,采用 Windows11 操作平台, Intel (R) Core (TM) i9-14900KF 3.20 GHz CPU 和 NVIDIA 4070 GPU (24Gb 内存)。

表 1 StyleGAN3 相关参数配置

配置项	参数
cfg	StyleGAN3-t
batch-gpu	4
gamma	3.2
mirror	1
king	1 000
snap	10
tick	4



图 2 生成的扩充图像

表 2 数据扩充前后各缺陷像素及占比

类别	树根	变形	沉积	障碍物	破裂
扩充前像素数量	2 435 331	4 251 319	8 302 008	4 486 401	2 128 903
占比/%	11	20	38	21	10
扩充后像素数量	3 128 860	6 407 025	8 358 716	4 944 197	2 760 071
占比/%	12	25	33	19	11

2 模型改进及评价

2.1 改进的 U-Net 结构

针对排水管道缺陷识别场景,本文提出一种基于 CBAM 模块改进的 U-Net 网络模型,CBAM 模块结构如图 4 所示,由通道注意模块和空间注意模块组成,通道注意模块关注图像中哪些特征通道更关键,空间注意模块关注图像中哪些区域特征更关键,该模块可以根据输入图像的特征自适应优化。此外,CBAM 的一个主要优点是轻量化,可以无缝集成到任何即插即用的神经网络中。改进模型沿用 U-Net 经典的编码器-解码器对称结构,如图 5 所示,在特征提取过程中融入 CBAM 注意力机制,以增强对管道缺陷特征的捕捉能力。编码器部分通过多层级联的 3×3 卷积,搭配 ReLU 激活函数与 2×2 最大池化操作,逐步缩减特征图尺寸,从 512×512 降至 32×32 ,同时加深特征维度,从 3 通道增至 1 024 通道,每个层级

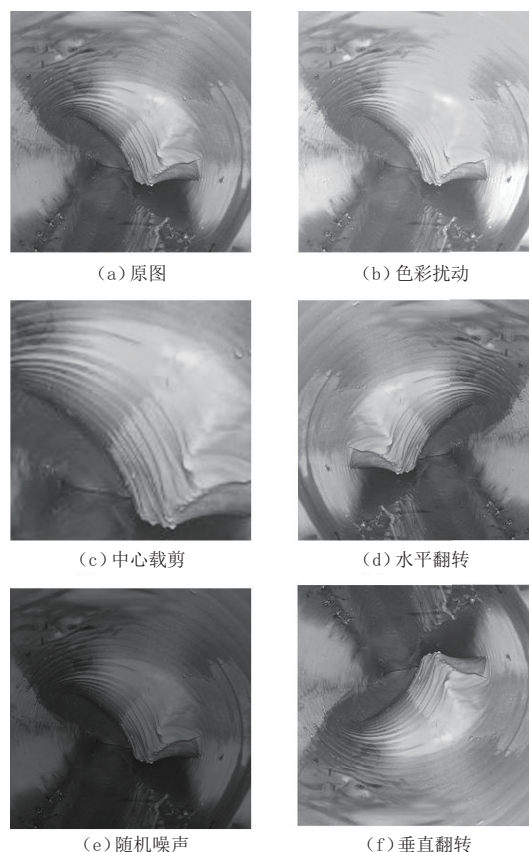


图 3 图像基础增强

后嵌入 CBAM 模块,实现通道与空间维度的注意力加权,强化缺陷特征的显著性。解码器通过 2×2 上采样操作恢复特征图尺寸,结合编码器对应层级的特征拼接,融合浅层细节与深层语义信息,再经 3×3 卷积与 1×1 卷积调整特征维度,最终输出 6 个通道结果,对应上述 5 种排水管道缺陷,加上背景共 6 个类别,从而实现对排水管道缺陷的精准分割。该改进网络通过注意力机制与特征融合策略,有效应对排水管道缺陷复杂纹理、光照不均等挑战,提升缺陷识别的准确性与鲁棒性。

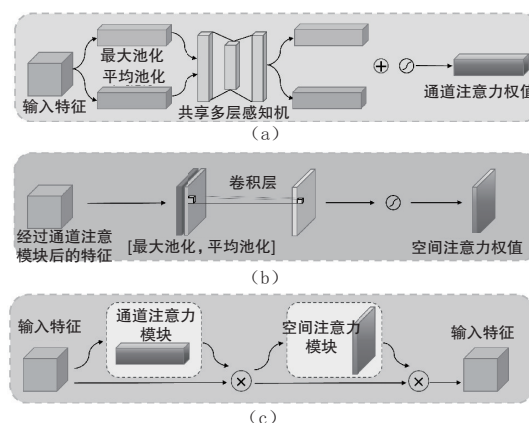


图 4 CBAM 的结构

2.2 模型评价指标

为了验证改进的 U-Net 模型的识别性能,采用交

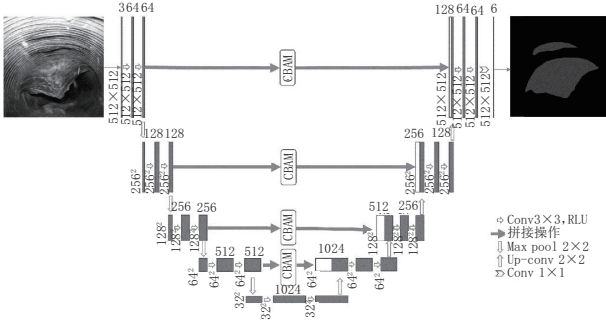


图5 改进后U-Net网络

并比 (Intersection over Union, IoU)、像素精度 (PA)、精确度 (Precision)、平均交并比 (mean Intersection over Union, mIoU)、和平均像素精度 (mean PA, mPA)、平均精确度 (mean Precision, mPrecision)、总体准确率 (overall accuracy, OA) 共 7 个精度评价指标来评价模型的识别精度。

$$IoU = \frac{TP}{FN + FP + TP} \quad (1)$$

$$PA = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN + FP + TP} \quad (4)$$

$$mPA = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{TP + FN} \quad (5)$$

$$mPrecision = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{TP + FP} \quad (6)$$

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \quad (7)$$

式中: k 为排水管道缺陷的类别数量加背景, 本研究为 6; TP 是真阳性; TN 为真负; FP 为假阳性; FN 为假阴性。对于某一类缺陷, TP 和 TN 分别表示被正确预测为“该缺陷”和“非该缺陷”的像素数, FP 表示被错误判断为“该缺陷”的“非该缺陷”像素数, FN 表示被错误判断为“非该缺陷”的“该缺陷”像素数。解释如图 6 所示。

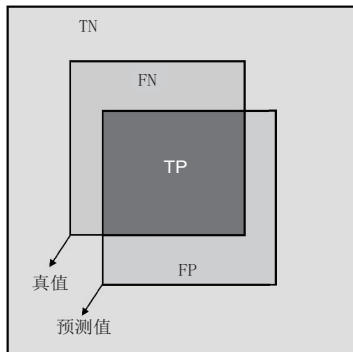


图6 评价指标解释

3 实验结果

3.1 模型评价

使用 2.1 和 2.2 章节中的数据作为原始数据集和增强数据集, 搭建多种语义分割模型进行对比, 图 7 展示了不同模型在不同数据条件下训练集和验证集的损失曲线。在训练集中, 无论是使用原始数据还是增强后的数据, 各个模型都能有较好的收敛, 其中 loss 最低的为使用增强数据训练的改进型 U-Net, 损失值为 0.034 1。PSPNet 和以 XceptionNet 作为 Deeplabv3+ 主干特征网络并使用原始数据的 Loss 高于其他模型分别为 0.121 4、0.132 8。在验证集中使用原始数据训练的 U-Net 出现了过拟合现象, 因此各项分割指标在验证集中表现大多数低于其他模型组合。而在验证集中, 使用本文数据增强方法后的各个模型收敛均较好, 实现了较好的收敛效果。表 3 总结了各研究网络的在验证集上的分割指标 (即 mIoU、mPA、mPrecision、OA), 结果表明, 各个分割网络在数据增强后 mIoU、mPA、mPrecision、OA 均有大幅提升, 基于数据增强以及改进后的 U-Net 更是分别高达 85.02%、91.40%、92.21%、96.51%。从分割指标上看, 表明大部分排水管道缺陷能够被准确检测、定位, 并且能对排水管道缺陷进行精确像素级分割, 由数据增强带来的增益明显大于模型改进后的增益, 增强数据后的分割模型 mIoU 都在 80% 以上。相比之下, U-Net 网络指标提高最多, 在数据增强后 mIoU 提高 34.29%, mPA 提高 31.03%, mPrecision 提高 20.83%, OA 提高 10.07%。改进后的 U-Net 在相对于改进前使用的 mIoU 提高 1.6%, mPA 提高 1.08%, mPrecision 提高 0.78%, OA 提高 0.53%。对于不同模型以及使用原始数据或增强数据共 9 种组合情况对各个缺陷类别的 IoU、PA、Precision 比较如图 8 所示, 数据增强有效提高了原始数据集中五类缺陷的三种分割指标。其中对于增强后的数据集, 本文改进的 U-Net 在各类缺陷中表现几乎最好, 其中 U-Net 体系网络对管道破裂和树根的识别更好, 本文改进的 U-Net 对破裂缺陷的 IoU 达 0.81, PA 为 0.88, Precision 为 0.90; 对树根 IoU 为 0.77, PA 为 0.88, Precision 为 0.86。分析其原因, U-Net 是一种编码器解码器结构的网络, 主要由卷积层和上采样层组成, 而其特点是能够同时获取全局和局部信息, 因此在处理小目标时表现较好, 适合于检测小目标破裂和树根两类缺陷。

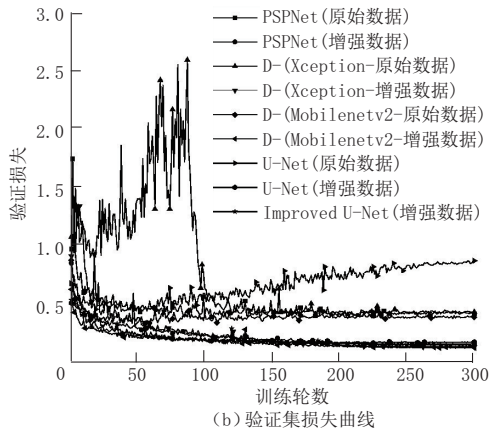
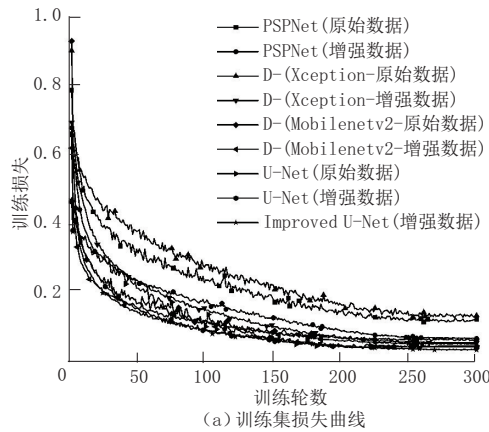


图 7 不同模型训练集验证集损失

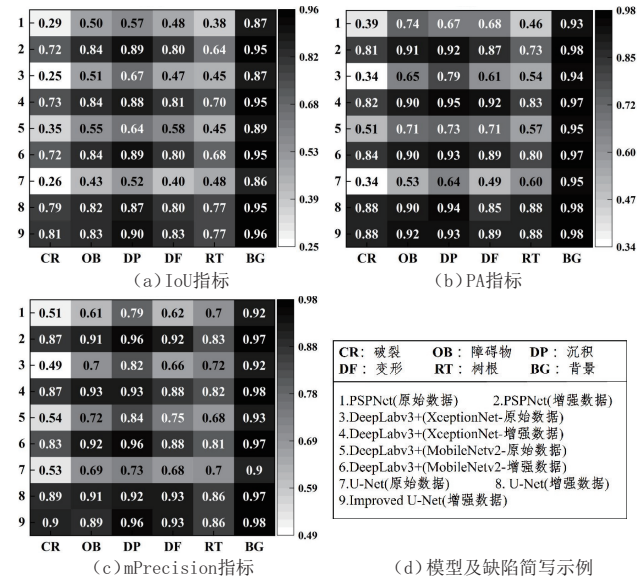


图 8 不同网络 5 类缺陷类型及背景指标

3.2 排水管道缺陷语义分割

如表 4 所示,绝大部分模型在数据增强之后,都取得不错的分割效果,都能准确的进行多类排水管道缺陷像素点分割,对于表 4 破裂缺陷中,在数据增强之前,PSPNet 网络对此破裂效果识别较差,但在数据增强之后能有效准确的识别出缺陷并能以较高精确度进行分割,取得了不错的分割效果。表 4 障碍物缺陷中,由于光线较暗,对于分割网络具有较大考

表 3 不同模型数据增强前后指标对比

网络模型	主干网络	数据类型	mIoU	mPA	mPrecision	OA
PSPNet	Resnet50	原始数据	51.29%	64.55%	69.02%	86.44%
PSPNet	Resnet50	增强数据	80.59%	86.98%	90.99%	95.51%
DeepLabv3+	Xception	原始数据	53.59%	64.67%	71.72%	87.51%
DeepLabv3+	Xception	增强数据	82.03%	89.73%	90.09%	95.71%
DeepLabv3+	Mobilenetv2	原始数据	57.68%	69.64%	74.25%	89.07%
DeepLabv3+	Mobilenetv2	增强数据	81.14%	89.15%	89.43%	95.52%
U-Net	VGG16	原始数据	49.13%	59.29%	70.60%	85.91%
U-Net	VGG16	增强数据	83.42%	90.32%	91.43%	95.98%
Improved U-Net	VGG16	增强数据	85.02%	91.40%	92.21%	96.51%

验,但在数据增强之后都能较好的改善之前的分割效果。表 4 中沉积缺陷是由于管道长时间使用造成,数据增强后的各模型取得不错的识别效果,均能对沉积范围分割较为准确。表 4 中存在有排水管道大面积变形缺陷,其中表现较为优秀的有数据增强后的 U-Net 以及改进的 U-Net 网络,其余网络在管道变形与分变形之间产生混淆。表 4 树根缺陷中,其中纵横交错,无规则分散分布,识别难度较大,但对于数据增强后的模型,基本具有分割出大部分树根的能力,其中表现较为优秀的使用增强数据训练的 U-Net 与本文提出的增强 U-Net。特别注意的是,在树根缺陷图中左下角光线十分黑暗的情况下,U-Net 具有良好的分割效果,这也再次印证了前面所述 U-Net 对树根分割效果更好。表 4 原始图像、标注样本和不同网络对样本图像的预测结果

4 结 论

本研究探讨了利用生成对抗网络用于排水管道图像增强的可行性,并基于排水管道图像数据中较少像素数量缺陷类别进行扩充生成。比较了“PSPNet”、“U-Net”以及 DeepLabv3+利用“XceptionNet”、“MobileNetv2”三种主干特征提取网络对数据增强前后的识别效果对比,各项评价指标在数据增强后得到大幅度提升。本文还提出了一种利用 CBAM 双重注意力机制改进的 U-Net 的排水管道识别模型,能够有效提高模型对密闭空间场内对排水管道缺陷的注意力。本文考虑了 5 种类型的排水管道缺陷,包括破裂、障碍物、沉积、变形和树根。验证了 StyleGAN3 数据增强后的优越性,以及 CBAM 模块的有效性。基于改进的 U-Net 模型 mIoU 达 85.02%,mPA 为 91.40%,mPrecision 达 92.21%,OA 为 96.51%。从分割指标上看,如图 8 所示,发现 3 种分割指标与原始数据中各类缺陷的像素点数量呈一定的正相

预测项	排水管道缺陷				
	破裂	障碍物	沉积	变形	树根
原始图像					
分割标签					
PSPNet (原始数据)					
PSPNet (增强数据)					
DeepLabv3+ (XceptionNet) (原始数据)					
DeepLabv3+ (XceptionNet) (增强数据)					
DeepLabv3+ (MobileNetv2) (原始数据)					
DeepLabv3+ (MobileNetv2) (增强数据)					
U-Net (原始数据)					
U-Net (增强数据)					
Improved U-Net (增强数据)					

关,在原始数据中,沉积缺陷所占像素数量最多,在使用原始数据的各个模型中均对沉积识别效果好,而破裂、变形、树根所占像素数量相对较少,分割指标相对较低,在利用StyleGAN3进行选择生成扩充数据后,即对原始数据集中较少像素数量例如破裂、

变形以及树根缺陷进行图像生成,其中在变形中扩充的像素数量最多,因此在增强数据集上各个模型对变形提升也最为明显,使得原本像素量较少的种类识别准确率得到有效提高,这有利证明了利用StyleGAN3进行下水道缺陷图像生成来增强数据集的有效性。

参考文献:

[1] Rojas-Gómez K L, Binder M, Walther M, *et al.* A parsimonious approach to predict regions affected by sewer-borne contaminants in urban aquifers[J]. Environmental Monitoring and Assessment, 2023, 195 (12): 1517.

[2] Snider B, McBean E A. Assessing the impact of pipe rehabilitation on decreasing watermain break rates using random survival forest models [J]. Journal of Water Resources Planning and Management, 2022, 148 (8): 04022045.

[3] LI M, LI M, REN Q, *et al.* Intelligent identification and classification of sewer pipeline network defects based on improved RegNetY network[J]. Journal of Civil Structural Health Monitoring, 2023, 13(2): 547-560.

[4] CHENG J C P, WANG M. Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques[J]. Automation in Construction, 2018, 95: 155-171.

[5] WANG M, LUO H, CHENG J C P. Towards an automated condition assessment framework of underground sewer pipes based on closed-circuit television (CCTV) images[J]. Tunnelling and Underground Space Technology, 2021, 110: 103840.

[6] CHEN J, ZHANG D, HUANG H, *et al.* Image-based segmentation and quantification of weak interlayers in rock tunnel face via deep learning [J]. Automation in Construction, 2020, 120: 103371.

[7] JI A, XUE X, WANG Y, *et al.* An integrated approach to automatic pixel-level crack detection and quantification of asphalt pavement[J]. Automation in Construction, 2020, 114: 103176.

[8] 吴哲豪,郭源,黄胜,等.基于EfficientViT改进YOLOv8排水管道缺陷检测模型[J].给水排水,2025,61(1):125-130.

[9] 凌同华,贝政豪,张胜,等.基于YOLO-Pipe和ByteTrack的排水管道缺陷检测[J].中国给水排水,2025,41(3):125-130.

[10] Karras T, Aittala M, Laine S, *et al.* Alias-free generative adversarial networks[J]. Advances in neural information processing systems, 2021, 34: 852-863.

[11] 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.