

DOI:10.19853/j.zgjsps.1000-4602.2022.13.004

基于 Deeplabv3+ 的排水管道缺陷检测与语义分割

周倩倩, 刘汉林, 陈维锋, 司徒祖祥, 腾 帅, 陈贡发
(广东工业大学 土木与交通工程学院, 广东 广州 510006)

摘 要: 传统计算机视觉技术应用于排水管道缺陷检测和评估, 存在识别类型单一、时效性差、判断准确率低等问题, 无法满足现代排水管道多缺陷共存、实时性强和精准检测的需求。近年来兴起的深度学习神经网络技术, 具有强大的数据特征学习和处理能力。为此, 提出了基于 Deeplabv3+ 卷积神经网络的管道缺陷检测及语义分割方法, 实现对排水管道缺陷的多类型检测、空间定位和几何属性分割。分别比较了 ResNet-18、ResNet-50、Mobilenet_v2、Xception 和 InceptionResnet_v2 这 5 类骨架特征提取网络对缺陷检测和语义分割的影响作用。结果表明, ResNet-50 的识别分割性能优于其他网络, 准确率达到 89.8%, 平均交并比和加权交并比分别为 53.2% 和 83.9%, 分割速率为 12.50 帧/s。这为排水管道缺陷的智能检测与分割提供了新的技术支撑和手段。

关键词: 管道缺陷; 卷积神经网络; 检测和定位; 语义分割

中图分类号: TU992 **文献标识码:** A **文章编号:** 1000-4602(2022)13-0022-06

Drainage Pipeline Defects Detection and Semantic Segmentation Based on Deeplabv3+

ZHOU Qian-qian, LIU Han-lin, CHEN Wei-feng, SITU Zu-xiang, TENG Shuai, CHEN Gong-fang

(School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: The application of traditional computer vision technology in inspection and evaluation of CCTV sewer defects has many problems, such as incapable of multiple defects identification, poor detection effects and low detection accuracy, which cannot solve the tasks on multiple defects, real-time and accurate detection of drainage defects. In recent years, the emerging deep learning neural network technology has powerful data feature learning and processing capabilities. This paper proposes a pipeline defect detection and semantic segmentation method based on Deeplabv3+ convolutional neural network, which is used for multi-type detection, locating and geometric attribute segmentation of sewer defects. The impacts and mechanism of five types of backbone feature extraction networks, including ResNet-18, ResNet-50, Mobilenet_v2, Xception and InceptionResnet_v2, on defect detection and segmentation were analyzed respectively. The experimental results showed that the recognition and segmentation performance of ResNet-50 outperformed other networks. The accuracy rate was 89.8%, the mean and weighted mean intersection over union are 53.2% and 83.9%, respectively, and the segmentation rate is

基金项目: 国家自然科学基金青年基金资助项目(51809049); 国家级大学生创新训练项目(202111845038)

12.50 frames/s. This method proposed can provide technical support and means for intelligent detection and segmentation of sewer defects.

Key words: pipeline defect; convolutional neural network; detection and location; semantic segmentation

高效率、自动化、大规模的管道检测已成为城市排水设施建设和管理的迫切需求。在内外多种因素的长期作用下,许多排水管网存在堵塞、破损、渗漏等问题,造成管道功能障碍、污水渗漏和排水不畅,严重的还会引起地面塌陷和环境污染等安全问题。对现有管道进行精准检测,找出安全隐患,对保证城市排水能力、延长管道使用寿命等都具有重要意义。尽管闭路电视(CCTV)检测技术为排水管道缺陷检测提供了数据支撑^[1],但依靠人工视觉和传统计算机检测技术^[2-4],无法解决多缺陷共存识别^[5-6]、时效性差^[7]、检测准确率低^[8]等问题。

当前管道缺陷检测的前沿技术,包含目标检测^[9]和语义分割^[10]等方法。其中,语义分割不仅可以对多种缺陷进行同时检测和定位,还可以对各类缺陷的几何特征和属性进行区域提取,为后期缺陷等级评估提供支撑。当前主流的语义分割算法包括FCN、SegNet、U-Net和Deeplabv3+等^[11],它们将图像级的分类诊断升级为像素级的分类检测和区域定位,检测功能远优于传统方法。FCN^[12]是语义分割的开山之作,但其在下采样时会使特征图的感受视野变小,无法对检测对象的结构或边界进行精准提取。SegNet^[13]在其基础上引入编码-解码结构,恢复了检测目标的细节和原始的空间维度,提升了对物体边缘部分的提取准确度,但在处理小尺寸目标的上采样时,仍会丢失部分边界信息。U-Net^[14]网络结构复杂,对物体边界的分割效果不够理想。

Deeplabv3+网络^[15]在编码结构中采用了金字塔池化模块(ASPP),大大改善了因感受视野的大小受到限制而导致像素被误分为其他具有相似结构物体类别的问题。同时,利用解码结构恢复目标边界的细节信息,提高了网络运行速率和分割准确度。目前,Deeplabv3+网络在路面和桥梁裂缝等结构性检测^[16]和无人驾驶领域^[17]中已有应用,并取得了很好的效果,但还未应用到排水管网缺陷的检测和特征分割,本研究是在该方向上的全新尝试。

1 方法论及数据

1.1 Deeplabv3+网络结构

目前最新且性能最优的Deeplabv3+模型于2018年提出,改善了Deeplabv3初版网络存在的生成预测图稀疏和边界信息丢失等问题。图1展示了该网络包含的编码和解码两大结构,编码结构包括基础网络和ASPP模块。

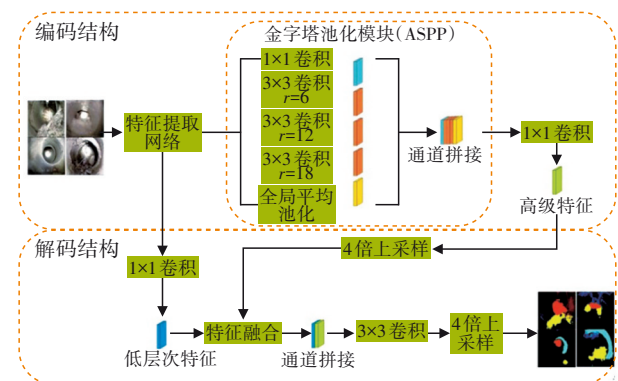


图1 Deeplabv3+网络编码-解码结构示意图

Fig.1 Deeplabv3+ network encoder-decoder structure diagram

基础网络用于对图像的数据特征进行提取。ASPP模块包括5个不同尺度的结构。其中,一个 1×1 卷积层和三个不同卷积速率 r 的 3×3 卷积层,扩大了感受视野,提取不同尺度的图像特征,再通过全局平均池化获取图像的全局特征。经拼接处理后,通过 1×1 卷积实现通道压缩^[18]。总结来看,编码结构从输入图像中提取空间特征和语义特征。而在解码结构中,用 1×1 卷积处理来自骨干网络的低层次特征图,以减少特征图的通道数,再与来自ASPP结构并经过4倍上采样的高级特征图进行融合。由于池化层的存在,特征图与原来图像相比尺寸较小(即分辨率比较低),因此进一步采用4倍上采样技术以恢复丢失的位置边界信息,得到对应的分割图,完成整体语义分割的过程。

值得注意的是,编码中的卷积层用于识别并提取原始数据的特征,在卷积层之后周期性设置池化

来缩小图像矩阵尺寸,减少参数量进而有效减少网络的训练数量。在相邻层之间的函数关系则通过激活函数来实现。特征图的上采样,能够将卷积池化后的特征图恢复到原来的尺寸大小。

1.2 数据和实验设置

本次实验采用来源于 CCTV 视频录像的 600 张排水管道图像作为数据库,在前期处理中,将所有图像的分辨率均调整为 512×512 像素,以提高网络训练时算法的收敛速度。下一步,借助 Matlab 平台的 Image Label 工具,采用 5 种颜色,分别对 5 种缺陷即裂缝(CR)、错口(DJ)、断壁(RW)、障碍物(OB)、树根(TR)以及背景进行像素级人工标注,从而获取语义分割的参照/验证标签图(即 Ground Truth,标注样本见图 2)。



图2 Image Label的像素级人工标注

Fig.2 Pixel-level manual labeling of Image Label

在 Deeplabv3+ 模型中,骨架特征提取网络至关重要,选取合适的网络结构能提高模型的检测精确度和运行效率。实验采用了 5 类主流网络进行了对比分析,即 ResNet-18、ResNet-50、Mobilenet_v2、Xception 和 InceptionResnet_v2(见图 3),进而提出性能最优的网络结构。

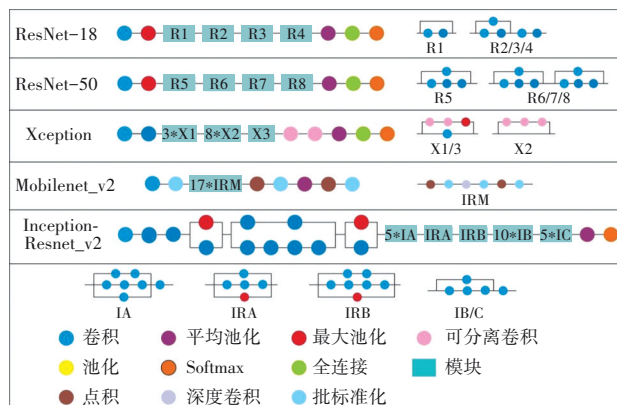


图3 5类骨架特征提取网络结构示意图

Fig.3 Schematic diagram of five types of skeleton feature extraction network structure

数据库设置按照 80% 的图像进行训练、20% 图像用于验证的方式。在训练过程中,对训练集进行了随机翻转和平移来实现数据增强,以提升训练数据量。用于训练的优化算法是随机梯度下降与动量(SGDM),学习率为每 10 轮降低 0.3,批处理为 8,从而减少训练时内存使用。

1.3 模型性能评价指标

采用所有主流的性能评价指标对模型性能进行多维度的评价,包括:①像素精度(Pixel Accuracy,记作 PA),体现了模型正确分割预测图像的像素与全体像素的比例;②平均像素精度(Mean Pixel Accuracy,记作 MPA),为各类别缺陷准确率的平均值;③交并比(Intersection over Union,记作 IoU),为图像分割预测结果与原始图像真值的重合程度;④平均交并比(Mean Intersection over Union,记作 MIoU),为所有缺陷类别的平均值;⑤加权交并比(Frequency Weighted Intersection over Union,记作 FWIoU),为 MIoU 的进阶,根据每个类别出现的频率为其设置权重,可改善图像分割中某类别的像素占全体像素的比例较小而导致预测结果较差的问题;⑥模型运行时间。以上指标中 PA、MPA、IoU、MIoU 及 FWIoU 的具体计算方法见公式(1)~(5)。

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

$$MPA = \frac{1}{2} \left(\frac{TP}{TP + FP} + \frac{TN}{FN + TN} \right) \quad (2)$$

$$IoU = \frac{\text{area}(A_p \cap A_t)}{\text{area}(A_p \cup A_t)} \quad (3)$$

$$MIoU = \frac{IoU}{N} \quad (4)$$

$$FWIoU = \frac{1}{TP + FP + FN + TN} \times \frac{IoU}{N} \quad (5)$$

式中:真正例(TP)表示模型正确预测出实际缺陷的像素;假正例(FP)表示模型错误预测出实际缺陷的像素;假反例(FN)表示模型错误预测出实际非缺陷的像素;真反例(TN)表示模型正确预测出实际非缺陷的像素; A_p 表示模型预测的缺陷像素面积; A_t 为实际的缺陷像素面积; N 为缺陷总类别数。

2 结果与讨论

研究的 5 类骨架特征提取网络对每种类型缺陷的分割精度的影响具体如表 1 所示。总体而言,除 InceptionResnet_v2 对各类缺陷的分割效果普遍较

弱外,其他网络对断壁实现了很好的分割效果,精度达到了 90% 以上。其次,对错口和树根的分割精度达到了 80% 左右。所有网络对障碍物与裂缝的像素分割相对较弱,基本低于 80%。ResNet-50 网络对断壁和背景的分割效果最好,精度分别达到了 93.3% 和 91.4%,但对裂缝、错口和障碍物这三类缺陷的分割精度略低于 Mobilenet_v2 和 ResNet-18。Mobilenet_v2 对各类缺陷的检测和分割效果均较理想,但对背景的分割能力稍弱。

表 1 各类缺陷及背景的像素分割精度

Tab.1 Prediction accuracy of various defects and background pixels %

项 目	裂缝	错口	障碍物	断壁	树根	背景
ResNet-18	75.8	80.8	73.6	92.9	77.9	89.1
ResNet-50	69.3	78.9	70.3	93.3	77.8	91.4
Mobilenet_v2	79.4	85.1	79.6	92.6	83.1	84.7
Xception	77.1	79.0	80.1	91.5	83.5	85.7
InceptionResnet_v2	77.1	67.6	75.3	61.6	80.2	89.9

5 类网络的其他性能评价指标汇总见表 2。ResNet-50 网络的预测准确率 PA 值最高,达到 89.8%,但其 MPA 值为 80.2%,低于除 Inception-Resnet_v2 外的其他网络。同时,参考 MIoU 和 FWIoU,从整体上更好地评估模型的分割效果,最高的是 ResNet-50,分别为 53.2% 和 83.9%,分割速率为 12.50 帧/s。因此,综合 5 种评价指标来看,ResNet-50 展现的性能最佳。其中的一个重要原因是由于 ResNet 网络创新性地引入了残差网络的概念,使靠前若干层的某一层数据输出直接跳过多层引入到后面数据层的输入部分,后面的特征层的内容会有一部分由其前面的某一层线性贡献。残差网络的采用克服了由于网络深度加深而产生的学习效率变低与准确率无法有效提升的问题。

表 2 各类性能评价指标

Tab.2 Various performance evaluation indicators

项 目	PA/%	MPA/%	MIoU/%	FWIoU/%	分割速率/ (帧·s ⁻¹)
ResNet-18	88.0	81.7	49.9	81.8	14.93
ResNet-50	89.8	80.2	53.2	83.9	12.50
Mobilenet_v2	84.7	84.1	45.8	78.1	12.20
Xception	85.5	82.8	46.7	79.9	10.99
Inception-Resnet_v2	87.2	75.3	45.9	80.5	8.55

图 4 选取了 5 张样本图像进行比较,从视觉上展示了 5 类网络对各类缺陷的分割效果。由 b 列图可知,所有网络对断壁均能实现理想的识别效果,尤其是 ResNet-18 和 ResNet-50,能进行完整和细致的预测标注。对于错口而言(c 列及 e 列图像),网络的分割能力较强,其中 ResNet-50 分割的错口区域更加接近原图及标签图,而其他网络对错口边界的识别和分割不够细致和平滑。在识别树根方面(a 列及 c 列图像),所有网络的预测结果也都较为准确,但由于树根细而密集的性质,预测面积往往略大于其实际面积。从 a、b、d、e 列图可以看出,所有网络对障碍物和裂缝均存在过度预测的情况,表现为分割的面积明显大于原图中的面积,甚至在 b 列图中,Xception 和 Mobilenet_v2 网络对障碍物出现了细部预测错误的情况,表现为预测了原图中未出现障碍物的区域。整体而言,ResNet-50 对各类缺陷及其边缘的分割较为精确,分割效果最优。

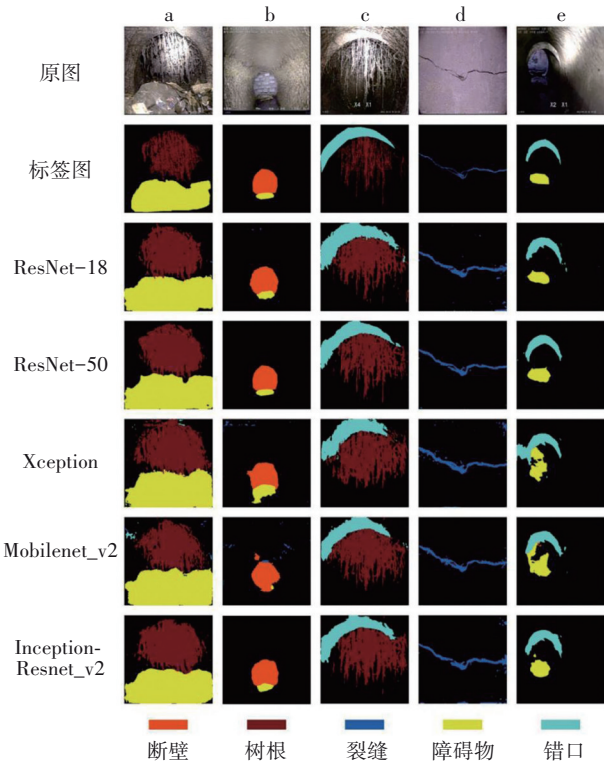


图 4 5 类骨架特征提取网络对缺陷的分割效果

Fig.4 Segmentation effect of five types of skeleton feature extraction networks on defects

针对 ResNet-50 网络进行了深入分析,图 5 展示了分割效果较好及不佳两种情况下的模型表现。由 A1 列可知,ResNet-50 识别性能优于人工标记,

其分割出的障碍物区域大于人工标记的区域,更加符合原图中障碍物的覆盖区域。同时,网络还识别出了人工未标记的障碍物和裂缝区域(A2和A3列),指出人工标记存在不准确性和不确定性的问题,这也是影响网络分割效果的重要原因之一。由此可见,Deeplabv3+的实际识别精度可能更高。

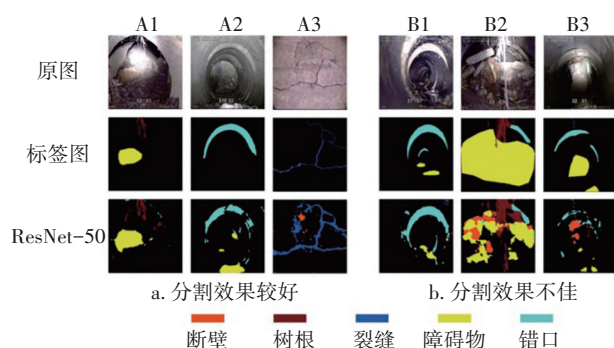


图5 基于ResNet-50网络识别各类缺陷的优劣对比

Fig.5 Comparison of the advantages and disadvantages of identifying various defects based on the ResNet-50 network

图5(b)表明,在某些特定条件下,ResNet-50的模型性能表现不佳:①当缺陷区域互不相连但相距较近时,区域间可能会出现多余的相连像素点,导致边界分割不准确,如图5(b)中B1列的障碍物区域。②当缺陷区域之间有重合部分时,会出现边界混淆和预测准确度降低的情况,如图5(b)中B2列障碍物和断壁的识别较为混乱。③当拍摄图片存在光线较暗的问题时,缺陷不易被识别出来,容易被错误预测,如图5(b)中B3列的错口识别效果较差,障碍物被错误识别为断壁。在将来的工作中,仍可针对这些问题对模型进行进一步完善。

3 结论

采用Deeplabv3+网络,对排水管道缺陷进行智能检测和语义分割,采用了5类骨架特征提取网络,结果表明,Deeplabv3+对排水管道的各类缺陷均能实现准确和高效的检测和分割。其中,整体性能较优的ResNet-50准确率达到89.8%,分割速率为12.50帧/s。同时,在人工标注训练样本的过程中,可能会存在少标、漏标的问题,后期需进一步加强人工标注的精细度,提升数据质量,有助于提高模型的预测精度。在未来的工作中,计划进一步改善和提高模型的准确率,并对各类缺陷的几何属性特征进行提取,建立缺陷综合评估机制,以完善管道缺陷的评估体系。

参考文献:

- [1] 郭翔. CCTV管道检测在扬州污水提质增效行动中的应用[J]. 中国给水排水, 2020, 36(20): 67-70.
GUO Xiang. Application of CCTV pipeline inspection in Yangzhou sewage quality improvement and efficiency enhancement action [J]. China Water & Wastewater, 2020, 36(20): 67-70 (in Chinese).
- [2] 刘玉贤, 叶绍泽, 吕兵, 等. 排水管网缺陷智能检测的信息化解决方案[J]. 中国给水排水, 2021, 37(8): 32-36.
LIU Yuxian, YE Shaoze, LÜ Bing, et al. Information solution for intelligent detection of drainage pipe network defects [J]. China Water & Wastewater, 2021, 37(8): 32-36 (in Chinese).
- [3] 王和平, 关安峰, 谢广永. 《城镇排水管道检测与评估技术规程》(CJJ 181—2012)解读[J]. 给水排水, 2014, 40(2): 124-127.
WANG Heping, GUAN Anfeng, XIE Guangyong. Interpretation of Technical Regulations for Inspection and Evaluation of Urban Drainage Pipelines (CJJ 181-2012) [J]. Water & Wastewater Engineering, 2014, 40(2): 124-127 (in Chinese).
- [4] XU K, LUXMOORE A R, DAVIES T. Sewer pipe deformation assessment by image analysis of video surveys[J]. Pattern Recognition, 1998, 31(2): 169-180.
- [5] KHALIFA I, ABOUTABL A E, BARAKAT A A. A new image-based model for predicting cracks in sewer pipes [J]. International Journal of Advanced Computer Science and Applications, 2013, 4(12): 65-71.
- [6] HALFAWY M R, HENGMECHAI J. Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine [J]. Automation in Construction, 2014, 38: 1-13.
- [7] SINHA S K. Automated Underground Pipe Inspection Using a Unified Image Processing and Artificial Intelligence Methodology [D]. Waterloo: University of Waterloo, 2000.
- [8] HAWARI A, ALAMIN M, ALKADOUR F, et al. Automated defect detection tool for closed circuit television (CCTV) inspected sewer pipelines [J]. Automation in Construction, 2018, 89(3): 99-109.
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//IEEE Computer Society. 2014 IEEE Conference on Computer Vision and Pattern

- Recognition. Columbus: IEEE Computer Society, 2014: 580-587.
- [10] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (4): 640-651.
- [11] 吕兵, 刘玉贤, 叶绍泽, 等. 基于卷积神经网络的 CCTV 视频中排水管道缺陷的智能检测[J]. 测绘通报, 2019 (11): 103-108.
- LÜ Bing, LIU Yuxian, YE Shaoze, *et al.* Intelligent detection of drainage pipe defects in CCTV video based on convolutional neural network [J]. Bulletin of Surveying and Mapping, 2019 (11): 103-108 (in Chinese).
- [12] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C]// IEEE. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2015: 3431-3440.
- [13] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [14] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]// HORNEGGER N, WELLS J, FRANGI W. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Berlin: Springer, 2015: 234-241.
- [15] CHEN L C, ZHU Y, PAPANDREOU G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//FERRARI V, HEBERT M, SMINCHISCU C. Computer Vision-ECCV 2018. Berlin: Springer, 2018: 2-5.
- [16] JI A K, XUE X L, WANG Y N, *et al.* An integrated approach to automatic pixel-level crack detection and quantification of asphalt pavement [J]. Automation in Construction, 2020, 114(6): 1-15.
- [17] 孙斌艳, 曹馨窃, 张连勇, 等. 基于语义分割的无人驾驶车道线检测算法研究[J]. 计算机时代, 2022 (5): 39-41.
- SUN Binyan, CAO Xinyao, ZHANG Lianrong, *et al.* Research on the detection algorithm of driverless lane line based on semantic segmentation [J]. Computer Times, 2022 (5): 39-41 (in Chinese).
- [18] 刘志赢, 谢春思, 李进军, 等. 基于改进 Deeplabv3+ 的烟雾区域分割识别算法[J]. 系统工程与电子技术, 2021, 43: 328-335.
- LIU Zhiying, XIE Chunsi, LI Jinjun, *et al.* Smoke region segmentation and recognition algorithm based on improved Deeplabv3+ [J]. System Engineering and Electronic Technology, 2021, 43: 328-335 (in Chinese).

作者简介:周倩倩(1984-),女,浙江乐清人,博士,副教授,研究方向为市政排水系统诊断。

E-mail:qiaz@foxmail.com

收稿日期:2021-08-04

修回日期:2021-09-20

(编辑:李德强)

大力推进水利薄弱环节建设,
提高防灾减灾能力