Jurnal Kejuruteraan 37(5) 2025: 2205-2214
https://doi.org/10.17576/jkukm-2025-37(5)-14

Sewer Pipeline Defect Detection based on YOLOv8-CPA

Jia Chaoyu* & Wan Nurshazwani Wan Zakaria

School of Mechanical Engineering,
College of Engineering, Universiti Teknologi MARA, Malaysia

**Corresponding author: jiachaoyu@xmhnphds.cn*

Received 28 October 2024, Received in revised form 18 February 2025
Accepted 18 March 2025, Available online 30 August 2025

ABSTRACT

As the service life of sewer pipelines increases, various types of defects inevitably occur, posing significant risks to urban infrastructure and public safety. Timely detection and assessment of these defects are crucial to ensuring the structural integrity and functionality of sewer systems. Traditional manual inspections are time-consuming, labor-intensive, and prone to human errors. Object detection technology based on deep learning provides an efficient and automated alternative by training models to accurately identify the type and location of pipeline defects. However, the complex internal environment of pipelines, including low-light conditions, noise interference, and varying defect appearances, presents significant challenges for detection accuracy. To address these issues, we conducted an in-depth study on common sewer pipeline defects and applied image preprocessing techniques such as grayscale conversion and denoising to enhance dataset quality. Furthermore, we improved the YOLOv8 model by integrating the CPA Enhancer module into its Backbone structure, optimizing feature extraction and defect recognition. Based on this enhancement, we developed a deep active learning framework, YOLOv8-CPA, which leverages a chain-thinking prompt mechanism to refine detection performance iteratively. Experimental results demonstrate that the YOLOv8-CPA pipeline inspection system achieves high accuracy in detecting and classifying pipeline defects. By improving detection efficiency, ensuring consistency, and accelerating validation processes, our system significantly enhances the sewer inspection workflow. The proposed method contributes to effective defect management, aiding in the timely implementation of appropriate maintenance and rehabilitation strategies for sewer infrastructure.

Keywords: Sewer pipeline detection; YOLOv8; deep learning; CPA- Enhancer

INTRODUCTION

The municipal underground sewer network is the infrastructure of a city, playing an irreplaceable role in sewer, flood prevention, and other aspects. As urban sewer networks grow, some pipelines have been built for a long time. Due to unclear standards for early pipeline construction, the problem of urban waterlogging caused by sedimentation, blockage, and damage to underground pipelines has become increasingly serious in recent years, seriously affecting the daily lives of residents and the economic development of the city. Therefore, the demand for detecting defects in sewer pipelines is rapidly increasing. Recent technological advances have led to machine learning (ML) becoming a multifaceted research method that incorporates optical technology, image processing, and pattern recognition. The introduction of machine learning technology to the task of detecting defects in sewer pipelines will provide ideas for further development in this area. Compared to machine learning, deep learning (DL) provides deeper learning models, simpler training processes, and higher detection accuracy.

Deep learning is becoming more widely used in industrial inspection as image quality and quantity improve. Unlike traditional machine learning algorithms, deep learning omits the time-consuming and crucial steps of feature engineering and connects the input-output layers end-to-end. Researchers have also applied deep learning to detect sewer pipeline defects.

In recent years, researchers have proposed a variety of deep learning-based methods in the field of sewage pipe defect detection to improve the accuracy and efficiency of detection. Xu Fang used an unsupervised machine learning algorithm to use the video data collected by the new sewage pipe detection equipment, converted it into feature vectors, and applied anomaly detection algorithms for defect identification. Dawei Li proposed a learning method based on multi-scale and global feature fusion, which enables the target detection network to simultaneously locate defects and finely classify them, thereby improving detection performance. Xianfei Yin used an advanced target detection algorithm based on YOLOv3 to train on a dataset of 4056 images containing six types of defects (breakage, holes, deposits, cracks, fractures, and tree roots) and one structural feature (pipeline interface), and achieved an average precision (mAP) of 85.37%. Duo Ma combined StyleGAN-SDM with CNN to propose a pipeline multi-defect detection system, in which StyleGAN-SDM preprocessed the original image to cope with the difficulty of data collection and small sample problems, and the multi-defect classification model (MDCM) fused with CNN combined with the Inception network and the residual network to improve detection accuracy. These studies show that the accuracy and generalization ability of sewage pipe defect detection can be effectively improved through the optimization of deep learning technology.

METHODOLOGY

We analyzed the current research status of defect detection in sewer pipelines and introduced deep learning methods to improve detection efficiency and reduce labor costs. However, there are still some problems: a complex internal environment and a poor dataset are present within the pipeline. Most deep learning methods for the detection of defects in sewer pipelines use simple applications and modifications of mature algorithms, without constructing detection algorithm models based on the characteristics of defect images. From the above perspective, this article selects the advanced detection framework YOLOv8 algorithm as the benchmark model, analyzes the characteristics of sewer pipeline defect images, and selects targeted improvements to YOLOv8, proposing a deep learning framework for detecting sewer pipeline defects called YOLOv8-CPA.

SEWER PIPELINE DATASET

In this research, we determine the defect category by considering the different types of defects, the degree of harm they can cause, and the frequency of their occurrence in CCTV shooting scenes. On this basis, the original video data is preprocessed, keyframes are extracted, image noise is removed through various machine learning methods, and problems such as overexposure and underexposure are alleviated to construct a sewer pipeline defect dataset.

There are various types of defects in sewer pipelines, with diverse causes and varying degrees of damage. Water Research Centre (WRc) was developed by the UK Water Research Centre and is widely used for the assessment of sewer networks. According to the WRc standard, defects are divided into two categories: structural defects caused by damage to the pipeline itself or abnormal connections at the pipeline, and functional defects caused by blockage by external objects inside the pipeline. The types of pipelines studied in this project include rainwater and sewage pipelines, and pipeline materials include reinforced cement pipelines and plastic pipelines. The specific types of defects are cracks, damage, sedimentation, and obstacles.

Breakage refers to significant damage to a pipeline that manifests as holes, ruptures, peeling, etc. Breakage typically results in pipeline leaks or loss of load-bearing capacity.

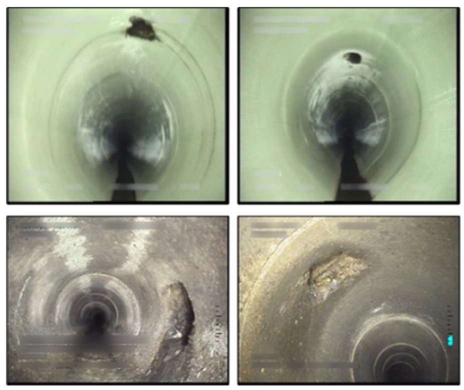

FIGURE 1. Pipeline Breakage Defect

Crack is a small gap on the surface of a pipeline, which may be caused by factors such as material aging, external forces, corrosion, etc. In pipelines, cracks may appear at any point, including the pipe walls, interfaces, etc.

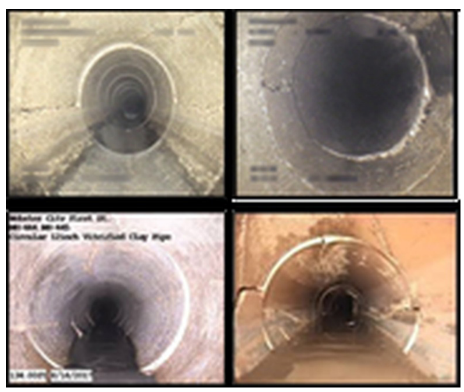

FIGURE 2. Pipeline Crack Defect

Obstacles refer to large objects such as tree roots, stones, and construction waste that affect the flow of water through pipelines. Obstacles can obstruct water flow and even cause blockages in pipelines.

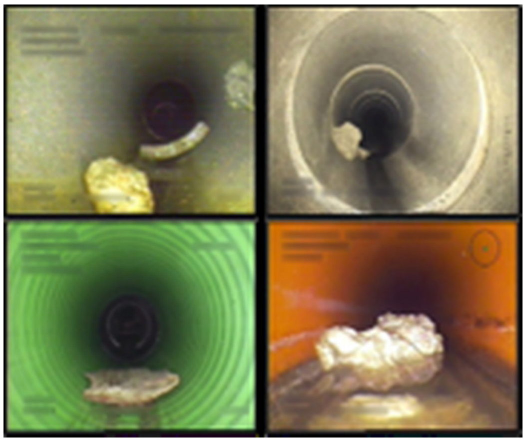

FIGURE 3. Pipeline Obstacle Defect

Sedimentation refers to the accumulation of debris such as garbage and silt in pipelines over a long period of time. Sedimentation will reduce the cross-sectional area of the pipeline and decrease its drainage capacity.

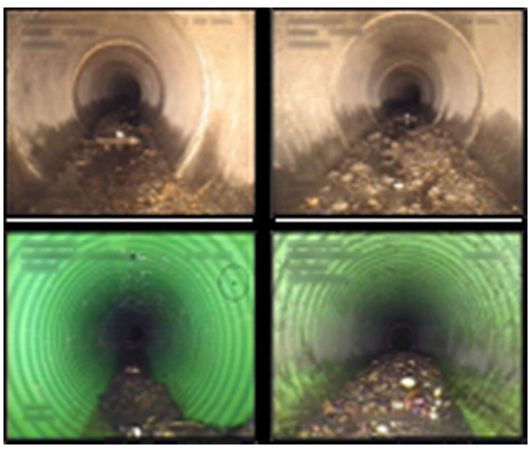

FIGURE 4. Pipeline Sedimentation Defect

DATA PREPROCESSING

Each pixel value in an image collected by CCTV equipment is composed of three different channels: R, G, and B. Compared to grayscale image data; color images store more color information internally. A large amount of color information helps technicians directly determine whether pipelines have defects and the type of defects. However, this type of color information is not helpful for machine learning models. On the contrary, this type of information greatly increases the calculation parameters of the model. To remove redundant color information, true color image data can be converted into grayscale images, simplifying matrices, reducing model calculation parameters, and improving computational speed.

There are three commonly used methods for image grayscale processing, all of which can be defined in Equation (1).

$$Gray(x,y)=\alpha R(x,y)+\beta G(x,y)+\theta B(x,y) \quad (1)$$

In the above equation, *Gray*, *R*, *G*, and *B* are different channels, and α, β, and θ are different weight values. The first method is the maximum value method, which takes the maximum pixel value of the *R*, *G*, and *B* channels and assigns it to *Gray*. At this time, the weight parameter of the channel with the highest pixel value is 1, and the other parameters are 0. The second method is the average method, where the weight parameters are all set to 0.33. The third method is the weighted average method, where each weight parameter is set to 0.299, 0.587, and 0.114. Figure 5 illustrates the images obtained by the three different grayscale processing methods described above.

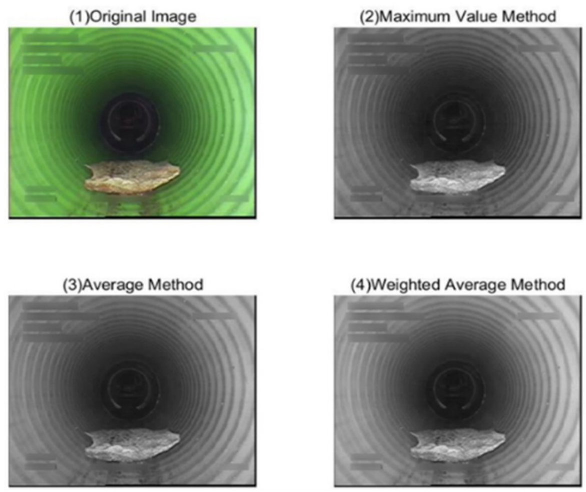

FIGURE 5. Grayscale processing

After verification, the images obtained by the weighted average method have the highest experimental effect. Therefore, we selected the weighted average method to process the grayscale values of the original data.

The presence of noise in the image is often the result of insufficient or excessive lighting when CCTV equipment is used to collect image samples from sewer pipelines. Noise adheres to the image and affects the training of the model and the identification of defects. Therefore, before annotating the data, the original samples need to be denoised. Image denoising generally involves adding filtering to the original image, preserving as much detail as possible, and suppressing noise. There are four common methods:

1. Mean filtering method. Mean filtering is a typical linear filtering method, which first fixes the pixel point, selects a filter of a certain size around the pixel point, and calculates the average pixel value inside the filter to replace the pixel value of the current fixed point. Traverse all pixels in the image and complete the mean filtering operation. Mean filtering can effectively remove Gaussian noise and smooth images, but it also has certain drawbacks: while removing image noise, it destroys the detailed features of the image, making it blurry, reducing the contrast between different pixels, and cannot effectively remove salt and pepper noise points.
   Mean filtering can be defined in Equation (2).

$$A'(x,y) = \frac{1}{(2L+1)^2}\sum_{-L}^{L}\sum_{-L}^{L} A(i+k, j+l) \tag{2}$$

2. Median filtering method. The median filtering method is based on sorting theory and is a nonlinear signal processing technique. Like the mean filtering method, median filtering also selects a filter of a certain size around a fixed pixel point. Firstly, multiple different pixel values inside the filter are sorted, and the middle value is selected to replace the fixed pixel value. Traverse the pixels of the image and complete the median filtering operation. Due to the nonlinear nature of median filtering, this method is more effective in dealing with smooth salt and pepper noise. By selecting appropriate points instead of contaminated points, this method is prone to losing image details such as edge information when removing noise from structurally complex images, thereby damaging the geometric structure of the image.

$$A'(x,y) = medf(x-k, y-l), k, l \in L \tag{3}$$

In the above equation, A represents the specified image, *(x, y)* is the pixel value at the specific position and L is the size value of the selected filter.

3. Wavelet transform denoising. The principle of wavelet transform threshold denoising is to decompose and transform noisy images in the spatial domain into the wavelet domain. At this time, the image itself has larger wavelet coefficients, while the wavelet coefficients of the noise are significantly smaller than those of the image itself. Therefore, by selecting an appropriate threshold and retaining content with larger wavelet coefficients, setting wavelet coefficients smaller than the threshold to zero, the goal of denoising can be achieved. Compared with median filtering and mean filtering methods, wavelet transform preserves the frequency information in the image while maximizing the spatial information of the image.

4. Gaussian filtering is a commonly used image processing technique, commonly used in applications such as denoising, smoothing, and edge detection. It is based on the concept of Gaussian function. Due to the properties of Gaussian function, pixels farther away from the center pixel contribute less to the new value. Weighted average processing is applied to pixels in the image, so that the influence of surrounding pixels is greater, while the influence of pixels farther away from the center is smaller. Therefore, Gaussian filtering can smooth the image and remove some noise while preserving the edges and details in the image. Gaussian filtering is achieved through a matrix (convolution kernel). The value of this matrix is calculated by a Gaussian function, which reaches its maximum at the center point and gradually decreases with increasing distance.

Peak Signal to Noise Ratio (PSNR) measures the quality of an image by calculating the ratio of its maximum signal to background noise, thus indicating the degree of distortion in the image. Its unit is dB, and the larger the value, the better the image quality and the lower the degree of distortion.

$$PSNR = 10 \cdot \log_{10}\left(\frac{max_I^2}{mse}\right) \tag{4}$$

In Equation (4), *max* is the maximum possible pixel value of the image, and *mse* is the mean square error of each pixel in the given size image and the noisy image. Select 5 different images and add noise.

TABLE 1. Comparison of Peak Signal to Noise Ratios of Various Algorithms

NO.	PSNR/dB			
	Mean Filter	Median Filter	Wavelet Filter	Gaussian Filter
1	26.842	28.473	22.636	27.357
2	24.395	24.942	22.312	25.324
3	25.258	25.822	22.604	24.865
4	25.749	26.343	22.528	26.094
5	27.459	29.527	22.689	28.241

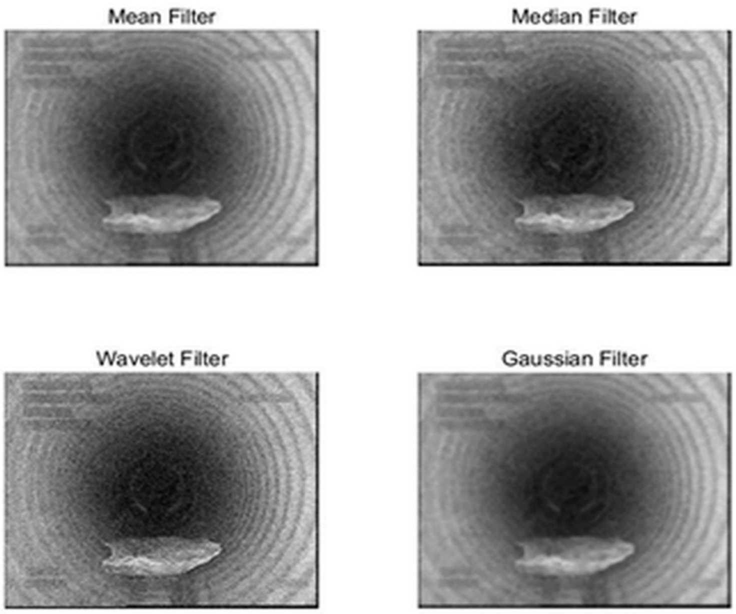

FIGURE 6. Comparison for Denoising Image 1

As shown in Table 1, all four methods were able to denoise 5 different test images. Compared with the other 3 methods, median filtering performed the best in all 5 images. The denoising effects of different methods on image 1 are shown in Figure 6. The median filtering, wavelet filtering, and Gaussian filtering have all reduced some noise, but the image appears blurry to the naked eye. Median filtering has achieved effective results in denoising, retaining more image details while removing noise. The degree of blur is smaller than the above three methods, and it is closer to the original image without added noise. Overall, this project used median filtering to reduce the amount of noise in the original image.

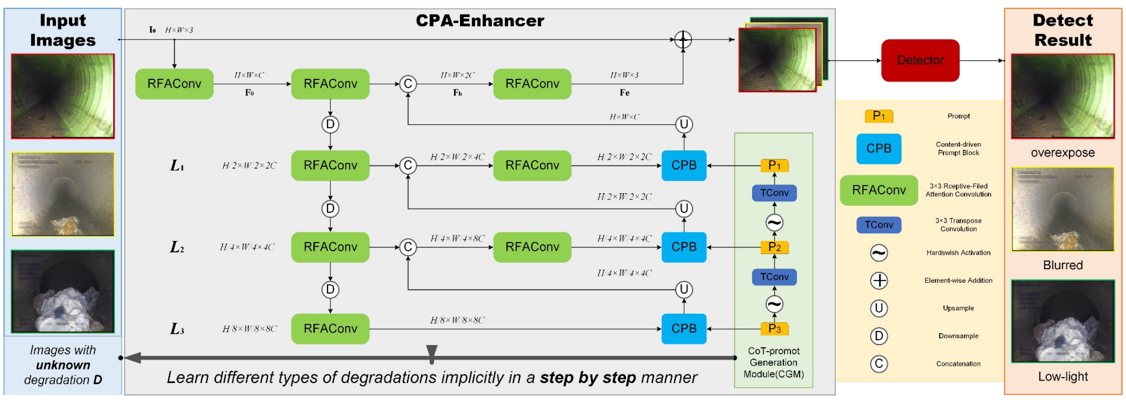

FIGURE 7. Overview of the CPA-Enhancer

YOLOV8-CPA

The present state of defect detection models for sewer pipelines have achieved certain results, but most of these models are little more than simple transfers of mature algorithms. To maximize the detection model's performance and improve its robustness, it is necessary to improve the various modules within the algorithm and design a detection framework that meets the specific tasks. In this study, we propose a deep learning framework called YOLOv8-CPA that can be used for defect detection in sewer pipelines by improving the YOLOv8 algorithm.

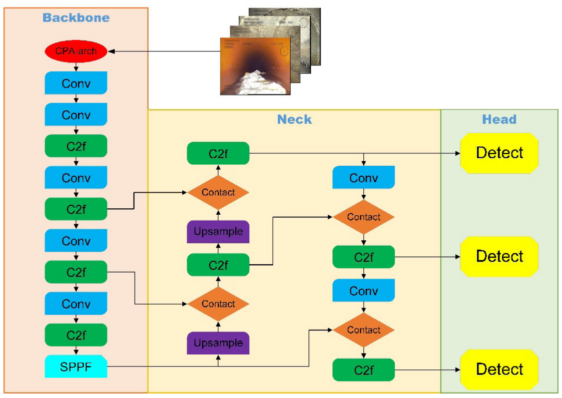

FIGURE 8. YOLOv8-CPA detection framework

The detection framework proposed in this article innovatively incorporates the CPA-Enhancer module in the Backbone section. CPA-Enhancer increases the model's ability to extract features for irregular defects (such as breakage, crack, sedimentation, etc.) through channel and position attention mechanisms, ensuring accurate localization of defect areas regardless of background complexity. The enhanced capability of feature extraction provides a more robust and accurate foundation for the detection of defects in the future. The improved YOLOv8 detection framework has demonstrated excellent performance in pipeline defect detection. It provides strong technical support for pipeline safety assessment and maintenance by not only identifying various common pipeline defects, but also accurately distinguishing them from normal areas in complex backgrounds.

Figure 7 shows the CPA Enhancer's structure. Given an input image $\boldsymbol{I_0} \in \mathbb{R}^{H\times W\times 3}$ with degradation D, the CPA Enhancer first applies receptive field attention convolution (RFAConv) to extract base features $\boldsymbol{F_0} \in \mathbb{R}^{H\times W\times C}$. This feature set $\boldsymbol{F_0}$ passes through a 4-level encoder, where each level uses a 3×3 RFAConv. The encoder reduces spatial dimensions and produces low-resolution features $\boldsymbol{F_r} \in \mathbb{R}^{\frac{H}{8}\times\frac{W}{8}\times 8C}$. The decoder then reconstructs high-resolution features from $\boldsymbol{F_r}$, and a final RFAConv generates enhanced features $\boldsymbol{F_e} \in \mathbb{R}^{H\times W\times 3}$. These combine with $\boldsymbol{I_0}$ to create an enhanced image $\boldsymbol{I_e} = \boldsymbol{I_0} + \boldsymbol{F_e}$, which is input into the YOLOv8 detector.

A chain-of-thought (CoT) module generates prompts to guide the model's adaptation to degradation. Content Driven Prompt Blocks (CPBs) further enrich features with degradation-specific details, enhancing robustness.

The CoT-Prompt Generation Module (CGM) encodes degradation-specific context, using Hardswish activation to filter relevant information:

$$P_i = \mathrm{Hardswish}\left(\mathcal{TC}_{3\times3}\,(P_{i+1})\right), i \in \{1,2\} \tag{5}$$

Content-driven Prompt Blocks (CPBs) allow input features F_i and prompts P_i to interact, refining enhancement based on degradation. This design reduces computation and improves efficiency through parallel processing.

The application of CPA-Enhancer in sewer defect detection is particularly important. Due to the complex internal environment of the pipeline, there are problems such as uneven lighting, stain interference, and complex background textures. Traditional detection methods are easily disturbed, resulting in false detection or missed detection. CPA-Enhancer effectively enhances the model's feature extraction capabilities for different types of defects (such as breakage, cracks, deposits, and obstacles) by introducing channel attention and position attention mechanisms. The channel attention mechanism can adaptively adjust the weights of feature channels so that the network pays more attention to important information in the defect area, while the position attention can accurately capture the spatial distribution of the target, ensuring that defects can still be accurately located in complex backgrounds. This improvement not only improves the detection performance of the model but also provides more reliable technical support for pipeline safety assessment and maintenance.

RESULT

This study used the Python 3.9.7 programming language. The CPU model is Intel Xeon Platinum 8375C and the GPU model is RTX 3090(24GB). The resolution of input crop image is set to 640 × 640 by cropping the original image. For different object detection networks, the training number is uniformly set to 100 epochs. The initial and end learning rates are 1×10^{-3} and 1×10^{-4}, respectively. In the training process, the optimal parameters of different networks are continuously adjusted and determined according to the pre-training results.

Overfitting is a common occurrence in the latter stages of deep learning, and the performance of the model is largely dependent on its ability to generalize. To comprehensively demonstrate the effectiveness of improving the object detection algorithm, this section sets up a control experiment to compare it with other mainstream algorithms.

TABLE 2. Comparison of the effectiveness of different detection models

Model	mAP	FPS
Faster R-CNN	78.9	23
YOLOv5s	85.7	51
YOLOv8s	92.1	60
YOLOX	89.2	45
YOLOv8-CPA	93.7	73

The mAP index is used to evaluate the detection accuracy of the algorithm, FPS is used to evaluate the detection speed of the algorithm, and the detection performance of YOLOv8-CPA is comprehensively analyzed from multiple aspects. The comparative experimental results are shown in Table 2.

Compared with other detection models, YOLOv8-CPA shows significant advantages in the task of sewer defect detection. From the perspective of detection accuracy, the mAP of YOLOv8-CPA reaches 93.7%, which is 14.8 percentage points higher than the traditional Faster R-CNN (78.9%), 8.0 and 4.5 percentage points higher than YOLOv5s (85.7%) and YOLOX (89.2%), respectively. Even compared with YOLOv8s (92.1%), it still has an advantage of 1.6 percentage points. This shows that the CPA-Enhancer module can more effectively extract defect features and improve detection accuracy, especially in the identification of pipeline defects in complex backgrounds. From the perspective of inference speed, YOLOv8-CPA runs at 73 FPS, far exceeding Faster R-CNN (23 FPS), YOLOX (45 FPS), and YOLOv5s (51 FPS), and 21.7% faster than the original YOLOv8s (60 FPS). This shows that CPA-Enhancer not only improves detection accuracy but also optimizes computational efficiency so that the model can still have real-time detection capabilities while ensuring high accuracy, providing a more efficient solution for practical engineering applications.

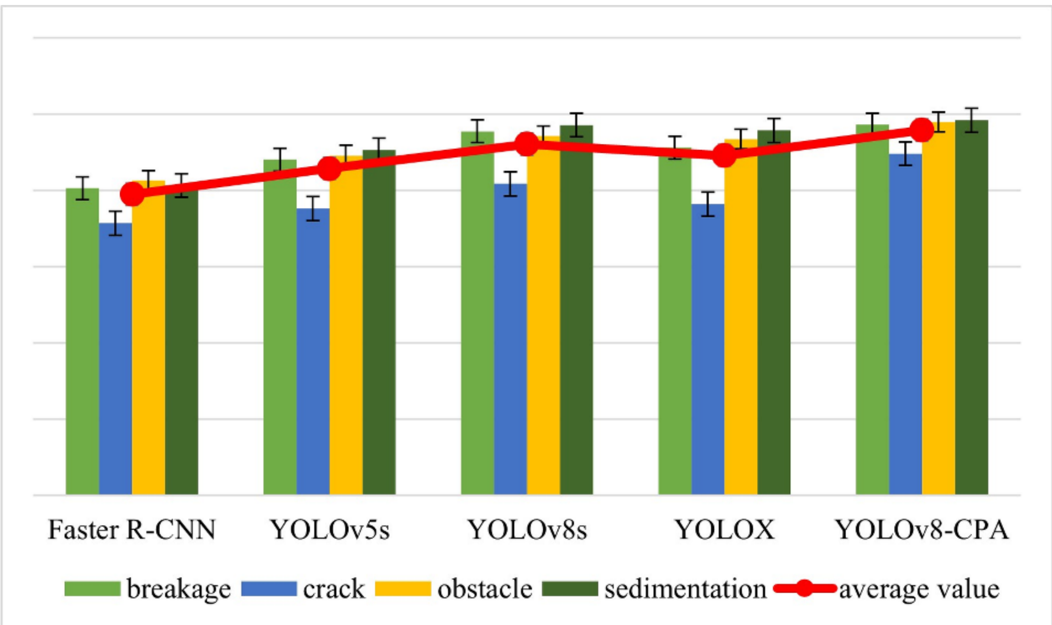

FIGURE 9. Comparison of YOLOv8-CPA and other algorithms for four types of defect detection accuracy

Figure 9 illustrates that the improved YOLOv8-CPA algorithm not only has the highest average accuracy, but also significantly improves the detection accuracy of various defects. Crack is difficult for mainstream algorithms to detect, mainly because compared to other defects, crack has fewer representational features in blurry image and has less difference from the internal environment of pipelines. Due to the stronger feature extraction and utilization ability of YOLOv8 CPA proposed in this article, it alleviates the poor detection accuracy caused by unclear defects in blurry, low light and overexposed image. YOLOv8-CPA is superior to mainstream object detection algorithms in terms of accuracy and speed while balancing different defects.

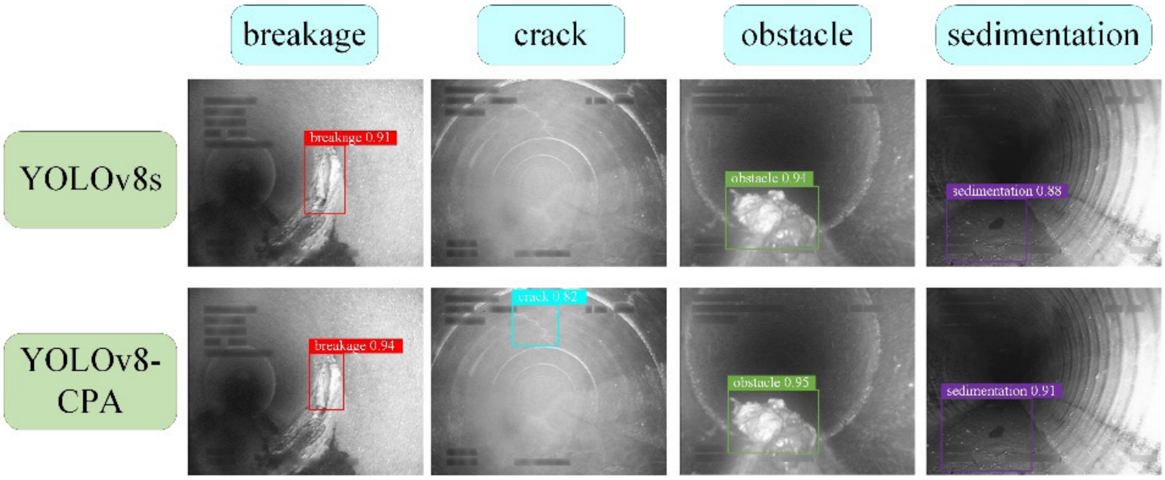

FIGURE 10. Comparison of detection effects

Figure 10 shows the partial defect detection results of YOLOv8 and YOLOv8-CPA detection algorithms. Firstly, compared to YOLOv8, YOLOv8-CPA has achieved more significant improvements in detecting four types of pipeline defects: breakage, crack, obstacle, and sedimentation. Secondly, detecting the location of defects is more accurate, and the size of the detection frame is more in line with the size of the object being detected. Last but not least, even when faced with partially blurred images, the improved model can still detect defect targets effectively, while YOLOv8 cannot. Consequently, YOLOv8-CPA is capable of detecting defects in blurry images in addition to providing high accuracy for general defect locations.

CONCLUSION

The rapid development of computer vision and deep learning technology not only promotes the progress of the computer industry but also provides creative ideas for solving problems in other industries. It is important to investigate and repair internal defects in the sewer pipeline as soon as possible to maintain the normal operation of the city. This article proposes a deep learning framework YOLOv8-CPA for defect detection in sewer pipelines.

The rapid advancement of computer vision and deep learning technologies has not only propelled innovation in computational systems but also inspired transformative solutions for cross-industry challenges. In urban infrastructure management, the timely identification and remediation of sewer pipeline defects are critical to ensuring municipal operational integrity. This study introduces YOLOv8-CPA, a novel deep learning framework tailored for sewer pipeline defect detection.

Focusing on the challenges of drainage pipeline defect identification, this work systematically reviews the current research landscape, highlighting the limitations of existing deep learning approaches in terms of generalization and robustness in complex environments. To address these gaps, we propose a comprehensive methodology: (1) Preprocessing techniques, including contrast enhancement and image denoising, are applied to improve the quality of defect images under low-visibility conditions; (2) The CPA-Enhancer module, integrated into the YOLOv8 backbone network, significantly amplifies feature extraction capabilities by leveraging multi-scale attention mechanisms, thereby enabling precise defect localization even in cluttered sewer environments. Experimental results demonstrate that the enhanced YOLOv8-CPA framework achieves superior performance in detection accuracy and computational efficiency compared to baseline models.

Beyond sewer systems, the proposed methodology holds potential for broader industrial applications, such as surface defect detection in steel strips, integrity monitoring of oil and gas pipelines, and infrastructure inspection in energy distribution networks. Furthermore, this study provides a foundational framework for advancing computer vision tasks through adaptive deep learning architectures, particularly in scenarios requiring robustness to environmental variability. Future work will focus on optimizing real-time deployment and expanding the framework's applicability to multi-modal sensor data.

ACKNOWLEDGEMENT

This research does not received any external fundings.

DECLARATION OF COMPETING INTEREST

None.

REFERENCES

Alharam, A., Almansoori, E., Elmadeny, W. and Alnoiami, H. 2020. December. Real time AI-based pipeline inspection using drone for oil and gas industries in Bahrain. In 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT) (pp. 1-5). IEEE. https://doi.org/10.1109/3ICT51146.2020.9312021

Athriyah, A. M. N., Amir, A. K. M., Zaki, H. F., Zulkifli, Z. A., & Rahman, A. 2022. Incremental learning of deep neural network for robust vehicle classification. *Jurnal Kejuruteraan* 34(5): 843-850. https://doi.org/10.17576/jkukm-2022-34(5)-11

Cao, H., Zhang, G., Zhao, A., Wang, Q., Zou, X. and Wang, H. 2024. YOLOv8n-CSE: A model for detecting litchi in nighttime environments. *Agronomy* 14(9): p.1924. https://doi.org/10.3390/agronomy14091924

Daher, S., Zayed, T. and Hawari, A., 2021. Defect-based condition assessment model for sewer pipelines using fuzzy hierarchical evidential reasoning. *Journal of Performance of Constructed Facilities* 35(1): p.04020142. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001554

Fang, X., Guo, W., Li, Q., Zhu, J., Chen, Z., Yu, J., Zhou, B. and Yang, H., 2020. Sewer pipeline fault identification using anomaly detection algorithms on video sequences. *IEEE Access* 8: 39574-39586. https://doi.org/10.1109/ACCESS.2020.2975887

Huang, F., Wang, N., Fang, H., Liu, H. and Pang, G., 2022. Research on 3D defect information management of drainage pipeline based on BIM. *Buildings* 12(2): 228. https://doi.org/10.3390/buildings12020228

Ismail, N. D., Ramli, R., & Ab Rahman, M. N. 2025. A systematic literature review of vision-based fire detection, prediction and forecasting. *Jurnal Kejuruteraan* 37(1): 191-218. https://doi.org/10.17576/jkukm-2025-37(1)-13

Kaddoura, K. and Atherton, J. 2021. Smart and automated sewer pipeline defect detection and classification. In Pipelines 2021 (pp. 135-143). https://doi.org/10.1061/9780784483602.016

Lalinia, M. and Sahafi, A. 2024. Colorectal polyp detection in colonoscopy images using YOLO-V8 network. *Signal, Image and Video Processing* 18(3): 2047-2058. https://doi.org/10.1007/s11760-023-02835-1

Li, D., Xie, Q., Yu, Z., Wu, Q., Zhou, J. and Wang, J., 2021. Sewer pipe defect detection via deep learning with local and global feature fusion. *Automation in Construction* 129: 103823. https://doi.org/10.1016/j.autcon.2021.103823

Li, Y., Wang, H., Dang, L.M., Piran, M.J. and Moon, H. 2022. A robust instance segmentation framework for underground sewer defect detection. *Measurement* 190: 110727. https://doi.org/10.1016/j.measurement.2022.110727

Li, Y., Wang, H., Dang, L.M., Song, H.K. and Moon, H. 2022. Vision-based defect inspection and condition assessment for sewer pipes: A comprehensive survey. *Sensors* 22(7): 2722. https://doi.org/10.3390/s22072722

Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X. and Pietikäinen, M. 2020. Deep learning for generic object detection: A survey. *International Journal of Computer Vision* 128: 261-318. https://doi.org/10.1007/s11263-019-01247-4

Ma, D., Liu, J., Fang, H., Wang, N., Zhang, C., Li, Z. and Dong, J. 2021. A multi-defect detection system for sewer pipelines based on StyleGAN-SDM and fusion CNN. *Construction and Building Materials* 312: 125385. https://doi.org/10.1016/j.conbuildmat.2021.125385

Ma, Q., Tian, G., Zeng, Y., Li, R., Song, H., Wang, Z., Gao, B. and Zeng, K. 2021. Pipeline in-line inspection method, instrumentation and data management. *Sensors* 21(11): 3862. https://doi.org/10.3390/s21113862

Palanisamy, P., Mohan, R.E., Semwal, A., Jun Melivin, L.M., Félix Gómez, B., Balakrishnan, S., Elangovan, K., Ramalingam, B. and Terntzer, D.N. 2021. Drain structural defect detection and mapping using AI-enabled reconfigurable robot raptor and IoRT framework. *Sensors* 21(21): 7287. https://doi.org/10.3390/s21217287

Pan, G., Zheng, Y., Guo, S. and Lv, Y. 2020. Automatic sewer pipe defect semantic segmentation based on improved U-Net. *Automation in Construction* 119: 103383. https://doi.org/10.1016/j.autcon.2020.103383

Rayhana, R., Jiao, Y., Zaji, A. and Liu, Z. 2020. Automated vision systems for condition assessment of sewer and water pipelines. *IEEE Transactions on Automation Science and Engineering* 18(4): 1861-1878. https://doi.org/10.1109/TASE.2020.3022402

Sharma, N., Sharma, R. and Jindal, N. 2021. Machine learning and deep learning applications-a vision. *Global Transitions Proceedings* 2(1): pp.24-28. https://doi.org/10.1016/j.gltp.2021.01.004

Shen, D., Liu, X., Shang, Y. and Tang, X. 2023. Deep learning-based automatic defect detection method for sewer pipelines. *Sustainability* 15(12): 9164. https://doi.org/10.3390/su15129164

Siu, C., Wang, M. and Cheng, J.C. 2022. A framework for synthetic image generation and augmentation for improving automatic sewer pipe defect detection. *Automation in Construction* 137: p.104213. https://doi.org/10.1016/j.autcon.2022.104213

Strubell, E., Ganesh, A. and McCallum, A. 2020, April. Energy and policy considerations for modern deep learning research. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 09, pp. 13693-13696). https://doi.org/10.1609/aaai.v34i09.7123

Sun, L., Zhu, J., Tan, J., Li, X., Li, R., Deng, H., Zhang, X., Liu, B. and Zhu, X., 2023. Deep learning-assisted automated sewage pipe defect detection for urban water environment management. *Science of The Total Environment* 882: 163562. https://doi.org/10.1016/j.scitotenv.2023.163562

von Chamier, L., Laine, R.F., Jukkala, J., Spahn, C., Krentzel, D., Nehme, E., Lerche, M., Hernández-Pérez, S., Mattila, P.K., Karinou, E. and Holden, S., 2021. Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nature Communications* 12(1): 2276. https://doi.org/10.1038/s41467-021-22518-0

Wang, M., Kumar, S.S. and Cheng, J.C., 2021. Automated sewer pipe defect tracking in CCTV videos based on defect detection and metric learning. *Automation in Construction* 121: 103438. https://doi.org/10.1016/j.autcon.2020.103438

Wang, Y., Li, P. and Li, J. 2022. The monitoring approaches and non-destructive testing technologies for sewer pipelines. *Water Science and Technology* 85(10): 3107-3121. https://doi.org/10.2166/wst.2022.120

Wu, Z., Yu, X., Zhu, D., Pang, Q., Shen, S., Ma, T. and Zheng, J. 2022. SR–DSFF and FENet-ReID: A two-stage approach for cross resolution person re-identification. *Computational Intelligence and Neuroscience* 2022(1): 4398727. https://doi.org/10.1155/2022/4398727

Yin, S., Li, H. and Teng, L. 2020. Airport detection based on improved faster RCNN in large scale remote sensing images. *Sensing and Imaging* 21(1): 49. https://doi.org/10.1007/s11220-020-00314-2

Yin, X., Chen, Y., Bouferguene, A., Zaman, H., Al-Hussein, M. and Kurach, L. 2020. A deep learning-based framework for an automated defect detection system for sewer pipes. *Automation in Construction* 109: 102967. https://doi.org/10.1016/j.autcon.2019.102967

Zhang, J., Liu, X., Zhang, X., Xi, Z. and Wang, S., 2023. Automatic detection method of sewer pipe defects using deep learning techniques. *Applied Sciences* 13(7): 4589. https://doi.org/10.3390/app13074589

Zhang, Y., Wu, Y., Liu, Y. and Peng, X., 2024. CPA-Enhancer: Chain-of-thought prompted adaptive enhancer for object detection under unknown degradations. arXiv preprint arXiv:2403.11220.

Zhou, Q., Situ, Z., Teng, S., Liu, H., Chen, W. and Chen, G., 2022. Automatic sewer defect detection and severity quantification based on pixel-level semantic segmentation. *Tunnelling and Underground Space Technology* 123: 104403. https://doi.org/10.1016/j.tust.2022.104403