A METHOD FOR FOREIGN OBJECT DETECTION IN UNDERGROUND CABLE PIPES BASED ON YOLO-MSAN MODEL

Yan FENG[1], Zongsheng XU[1], Yufang LIN[1], Linfeng DAN[1], Xinyu PEI[2], Xiongfeng HUANG[2, *]

Foreign objects in underground cable pipe, if undetected and unremoved, can impede cable laying and even damage cable structures. Conventional object detection methods are inefficient in poorly illuminated pipe environments and struggle to identify small foreign objects. To address these issues, this paper proposes a pipe foreign object detection method based on image preprocessing and an optimized YOLO model, enabling accurate detection of foreign objects in pipes. The proposed method tackles key challenges through three core improvements.First, gamma correction is applied to preprocess low-visibility images, resolving the low detection accuracy issue in dark environments. Second, multi-core branches and an attention mechanism are integrated into the neck structure of the YOLOv9 model to reduce missed detection of small foreign objects. Third, loss function optimization enhances model training efficiency.In practice, after collecting in-pipe images, the method first preprocesses the images and then uses the optimized YOLOv9 model to process them, identifying existing foreign objects. Experimental results show that compared with the original YOLOv9 model, the optimized model increases the mean Average Precision (mAP) for pipe foreign object detection from 86.9% to 95.2%, with significantly improved precision and recall stability.

Keywords: YOLOv9-MSAN, gamma correction, small target detection, foreign object detection

1. Introduction

The laying and maintenance of underground pipeline cables are indispensable and critical links in power grid construction. In practical engineering, however, foreign objects such as stones, metal burrs, and various metal debris are frequently present in cable pipes. If these impurities are not detected and cleaned in a timely manner, they will not only impede the cable laying process but may even cause damage to the cable structure. Considering the working environment of underground cable pipes, which is characterized by low visibility and the presence of small-sized foreign objects, there is an urgent need to develop a highly adaptable object detection algorithm. This algorithm should have stable detection capabilities

[1] Wuhan Huayuan Electric Power Design Institute Co., Ltd, Wuhan, 430050, China
[2] School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, 230009, China
[*] E-mail: hfut_huangxf@hfut.edu.cn

under low-light conditions and be able to achieve accurate identification and automatic classification of foreign objects inside the pipes.

Currently, the most commonly used object detection algorithms include YOLO [1], Faster R-CNN [2], and SSD [3]. The YOLO algorithm achieves end-to-end, efficient detection by dividing input images into fixed grids—each predicting target categories and positions, with detection results directly output via convolution [4]. The Faster R-CNN algorithm generates candidate regions using a Region Proposal Network (RPN), followed by classification and position regression for target determination [5]. The SSD algorithm performs detection using multi-scale feature maps, each responsible for targets of corresponding sizes and predicting categories/positions via convolution [6]. The above three methods all tend to have problems of missed detection and false detection in low-light environments.

The structure of this paper is as follows. Section 2 briefly elaborates on the overall process of the underground cable pipe foreign object detection scheme based on the optimized YOLO model. Section 3 investigates the low-visibility image preprocessing scheme based on gamma correction. Section 4 optimizes the neck structure and loss function of YOLOv9 respectively. Section 5 constructs a pipeline foreign object dataset and conducts comparative experiments using the original and optimized YOLOv9 object detection schemes.

2. Foreign Object Detection Scheme for Underground Cable Pipes

The flow chart of foreign object detection in underground cable pipes based on the optimized YOLO model is presented in Fig. 1.

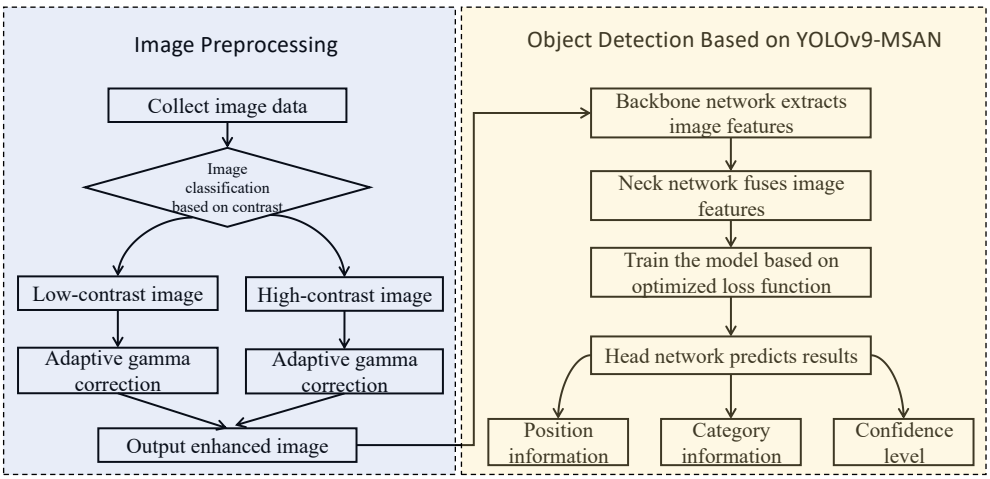


Fig. 1. Workflow of the foreign object detection

The overall process is divided into two stages: image preprocessing and object detection using the improved YOLOv9 model.

Initially, image data from the interior of cable pipes is collected by a camera. These images are then categorized into low-contrast and high-contrast images according to the statistical features of image contrast. Following this, adaptive gamma correction is applied to each category of images for enhancement.
Subsequently, the preprocessed images are fed into the optimized YOLOv9 model. The backbone network of the model extracts core feature layers of different scales from the input images. The neck structure integrated with multi-core branches and an attention mechanism performs multi-scale feature fusion to strengthen the features of small-sized foreign objects. Finally, the head network conducts bounding box regression and category prediction based on the optimized loss function, and outputs the position, category, and confidence of foreign objects in the pipes, thus completing the foreign object detection process.

3. Design of Low-Visibility Pipeline Image Enhancement Based on Gamma Correction

In the field of image processing, a variety of color models are available. Images are commonly represented in the RGB color space, where the three channels (Red, Green, and Blue) show a strong correlation. Conducting intensity transformation directly in the RGB space is prone to altering the original color of the image. For this reason, this method employs the HSV color space [7], which separates the color and brightness information of an image into three distinct channels: Hue (H), Saturation (S), and Value (V). The adaptive gamma correction operation is only applied to the V channel to enhance image brightness without altering the color information of the original image, while the H and S channels remain unchanged.

The adaptive gamma correction proposed in this paper is developed based on traditional gamma correction [8]. The formula for traditional gamma correction is given as follows:

$$I_{out} = cI_{in}^{\gamma} \tag{1}$$

In formula (1), $I_{out}$and I_{in}denote the intensity of the output image and the input image respectively; γ and c are two parameters that control the shape of the transformation curve. The adaptive gamma correction method proposed in this paper, after classifying images, dynamically sets and adjusts the parametersγ and c in the conventional gamma correction formula based on image - specific information. Through this dynamic adjustment, adaptive enhancement of low - visibility images are achieved.

3.1 Image Classification of Pipeline Interiors Based on Contrast Statistical Features

To implement targeted processing for images with distinct characteristics, it is essential to classify them using their contrast statistical features [9]. The specific classification method is based on the formulas presented below.

$$g(I) = \begin{cases} \varrho_1, D \leq 1/\tau \\ \varrho_2, \text{others} \end{cases} \tag{2}$$

where $D = 4\sigma$, τ is a parameter used to define image contrast; μ and σ denote the mean of image intensity and the standard deviation, respectively. When $D \leq 1/\tau$, the image is classified as a low-contrast category (ϱ_1), indicating that the pixel intensities of the image are mostly clustered within a small range; otherwise, it is classified as a medium-to-high contrast category (ϱ_2). From practical experience, this formula can effectively distinguish and characterize the contrast of different images when $\tau = 3$.

In addition, images whose average intensityμ is less than 0.5 are regarded as low-visibility images, and the adaptive gamma correction algorithm will process such low-visibility images.

3.2 Enhancement Algorithm for Low-Contrast Pipeline Interior Images

Most pixels in low-contrast images exhibit similar intensity values. To enhance image contrast, pixel values need to be spread over a wider range. In gamma correction, γ controls the slope of the output image intensity curve; a steeper slope corresponds to a more dispersed distribution of image intensities and thus higher contrast.

$$\Upsilon = -\log_2 \sigma \tag{3}$$

In traditional gamma correction, the parameterc serves to adjust the brightness level of the output image intensity. In the proposed method, the value of cis dynamically adjusted based on the inherent characteristics of the image, and its calculation formula is given as follows:

$$c = \frac{1}{1+\text{Heaviside}(0.5-\mu)\times(k-1)} \tag{4}$$

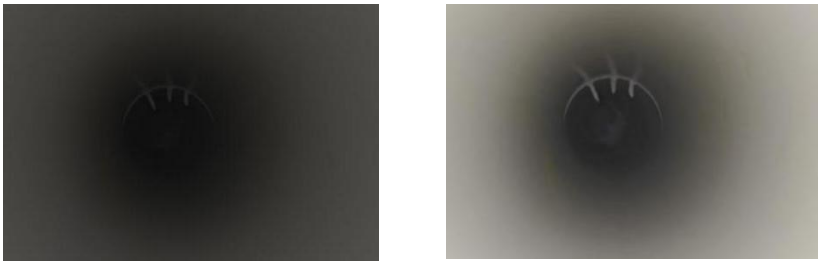
$$k = I_{\text{in}}^{\gamma} + \left(1 - I_{\text{in}}^{\gamma}\right) \times \mu^{\gamma} \tag{5}$$

The above setting correlates the value of c with the average image intensity μ. When $\mu \geq 0.5$, Heaviside(0.5-μ) = 0, and thus c=1; when $\mu < 0.5$, Heaviside(0.5-μ) = 1, and the value of c is determined by k. This enables targeted intensity transformation for images with different brightness characteristics, better adapting to the requirements of image enhancement.

Combining formulation (3), (4), and (5), the final transformation function for low-visibility images is obtained as follows:

$$I_{\text{out}} = \frac{I_{\text{in}}^{\gamma}}{I_{\text{in}}^{\gamma} + (1 - I_{\text{in}}^{\gamma}) \times \mu^{\gamma}} \tag{6}$$

By applying this method to enhance the collected images of burrs inside underground cable pipes, the resulting outcomes are presented in Fig. 2.



(a) Unenhanced pipeline burr image (b) Enhanced pipeline burr image

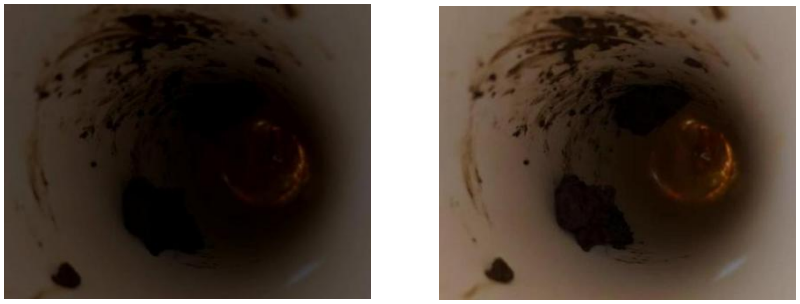
Fig.2. Enhancement effect of low-contrast pipeline images

3.3 Enhancement Algorithm for High-Contrast Images

The intensity distribution of pixels in high-contrast images is relatively scattered. The calculation method for the relevant parameters of such images $I_{out}$and c is similar to that described above. However, to avoid over-stretching of the image, adjustments need to be made to γ. The specific formula is as follows.

$$\gamma = \exp[(1 - (\mu + \sigma))/2] \quad (7)$$

By applying this method to enhance the collected images of soil inside underground cable pipes, the resulting outcomes are presented in Fig. 3.



(a) Unenhanced pipeline mud image (b) Enhanced pipeline mud image

Fig.3. Enhancement effect of high-contrast pipeline images

4. Optimization of the YOLOv9 Model for Small-Volume Object

4.1. YOLOv9 Object Detection Algorithm

YOLOv9 is a real-time object detection model [10]. It comprises five main components: The input end is responsible for image preprocessing; the backbone network, consisting of convolution modules, GELAN modules, and Adown modules [11], is used for feature extraction; the neck structure adopts an improved Path Aggregation Network (PAN) for multi-scale feature fusion [12]; the head designs independent detectors at each fused scale to predict bounding boxes, category probabilities, and confidence scores.

4.2. Observation equation and state equation of the Kalman filter

The traditional YOLOv9 model shows significant missed detection when identifying small-sized foreign objects. To resolve the issue, this paper designs a Multi-Scale Attention Neck (MSAN) structure to replace the original neck component of the model. The network structure of the improved YOLOv9-MSAN model is presented in Fig. 5. Compared with the traditional structure, the MSAN neck structure incorporates an SPDconv module and a CSP-OK module.

The structure of the CSP-OK module is depicted in Fig. 4

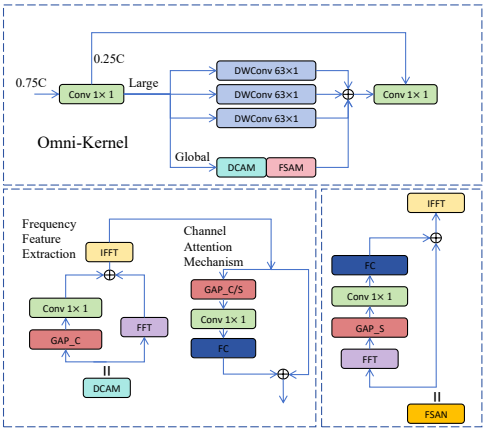


Fig.4. CSP-OK network structure

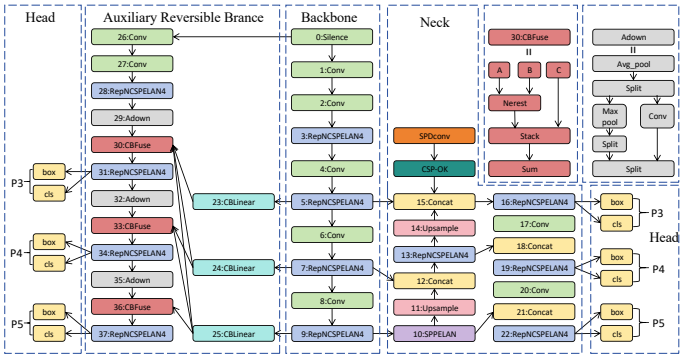


Fig.5. YOLOv9-MSAN network structure

This module consists of a multi-core (OK) branch and a large residual branch. In the OK branch, features are first processed via a 1×1 convolution, and then a Large branch, a Global branch, and a residual branch are used to handle information at different scales. The Large branch adopts 63×63 depthwise convolution to obtain receptive fields and combines 1×63 and 63×1 depthwise convolutions to capture contextual features. This design performs well for training images with similar scales, but its fixed-size convolution kernels fail to effectively cover when the input image size increases significantly. To address this, the Global

branch integrates a Dual-domain Channel Attention Module (DCAM) and a Frequency-domain Spatial Attention Module (FSAM): DCAM realizes the conversion of feature maps between the spatial and frequency domains through Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT) to extract frequency-domain features, and then assigns weights to each channel via a channel attention mechanism; FSAM applies a spatial attention mechanism in the frequency domain to generate weights for each spatial position.

Compared with directly adding a P2 detection layer to the network, the Small Object Enhancement Pyramid (SOEP) improvement scheme proposed in this paper not only avoids a substantial increase in model computation but also effectively learns global features, thereby enhancing the model's detection performance for small-sized pipeline foreign objects.

4.3 Optimization of the YOLOv9 Model Loss Function

The loss function of the original YOLOv9 model comprises Binary Cross-Entropy (BCE) Loss, Distribution Focal Loss (DFL), and Complete Intersection over Union (CIoU)[13].

Training data for object detection may contain low-quality samples, and explicitly enhancing bounding box regression for such samples will significantly degrade the performance of the detection model. To address this issue, this paper proposes a dynamic non-monotonic focusing mechanism and adopts the Wise-Intersection over Union (WIoU) loss function [14]. Instead of using traditional IoU to measure anchor box quality, this dynamic non-monotonic focusing mechanism evaluates anchor boxes through outliers and incorporates an intelligent gradient gain allocation strategy. This enables WIoU to reduce the importance of high-quality anchor boxes, mitigate harmful gradients generated by low-quality samples, and thereby focus on optimizing low-quality anchor boxes—ultimately improvi-quality anchor boxes—ultimately improving the detection accuracy of the model.

Intersection over Union (IoU) is used to evaluate the overlap between predicted bounding boxes and ground-truth bounding boxes, with its calculation formula given as follows:

$$\mathcal{L}_{\mathrm{IoU}} = 1 - \mathrm{IoU} = 1 - \frac{\mathrm{W_i H_i}}{\mathrm{S_u}} \tag{8}$$

$$\mathrm{S_u} = \mathrm{WH} + \mathrm{W_{gt} H_{gt}} - \mathrm{W_i H_i} \tag{9}$$

where W_i and H_i denote the width and height of the overlapping region, respectively; S_u is the area of the union region; W and H represent the width and height of the anchor box; and W_{gt} and H_{gt} are the width and height of the target box.

A dynamic non-monotonic focusing mechanism is designed based on Intersection over Union (IoU), and the quality of anchor boxes is measured by outliers. The outlier is defined as follows:

$$\beta = \frac{\mathcal{L}^*_{\mathrm{IoU}}}{\mathcal{L}_{\mathrm{IoU}}} \in [0, +\infty] \tag{10}$$

where $\mathcal{L}_{IoU}^{*}$ denotes the monotonic focusing mechanism. The hyperparameters used to construct the WIoU non-monotonic focusing coefficient are given by the following formula:

$$\mathcal{L}_{WIoU} = r\left(\exp\left(\frac{(x-x_{gt})^2+(y-y_{gt})^2}{W_g^2+H_g^2}\right)\right) \tag{11}$$

$$r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{12}$$

where w_g and H_g denote the width and height of the minimum bounding box, respectively; x and y arethe center coordinates of the anchor box; $x_{gt}$and y_{gt}representthe center coordinates of the target box. r is the dynamic non-monotonic focusing coefficient; β is the outlier; and α and δ are hyperparameters. By combining the outlier with hyperparameters, this coefficient enables dynamic allocation of gradient gains for anchor boxes of different qualities. It reduces the importance of high-quality anchor boxes and mitigates the harmful gradients from low-quality samples, allowing the model to focus more on low-quality anchor boxes and thereby improving detection accuracy.

The total loss function of the improved YOLOv9 is shown in Equation (13):

$$\mathcal{L} = \lambda_{box}\cdot\frac{1}{N}\sum_{i=1}^{N}\left(1-\mathcal{L}_{WIoU}(B_i,\widehat{B}_i)\right)+\lambda_{cls}\cdot\frac{1}{N}\sum_{i=1}^{N}\mathcal{L}_{BCE}(p_{cls}^{i},t_{cls}^{i})+\lambda_{dfl}\cdot \frac{1}{N}\sum_{i=1}^{N}\mathcal{L}_{DFL}(p_{dfl}^{i},t_{dfl}^{i}) \tag{13}$$

The total loss consists of three components: WIoU, BCE, and DFL, which are adjusted by weight hyperparameters λ_{box}, λ_{cls}, and λ_{dfl} respectively, followed by averaging the loss over all samples.

5. Experiment and analysis

5.1 Dataset Construction

Targets commonly present inside underground cable pipelines that may hinder the movement of cable pulling heads can generally be categorized into two types: foreign objects within pipelines and burrs at pipeline joints. Foreign objects in pipelines include stone, soil and metal. Burrs at pipeline joints are defined as small, sharp protrusions or edges formed at pipeline interfaces.

To address the above scenarios, we constructed a corresponding dataset, which consists of 2,850 images collected from actual underground cable pipeline detection scenes. The image resolution is uniformly adjusted to 640×640 pixels to match the input size of the detection model. The dataset is divided into 4 categories: stone, soil, metal debris, and pipeline joint burrs, with 720, 680, 750, and 700 images for each category, respectively. In order to avoid overfitting of the model, we adopted data augmentation strategies such as random flipping, rotation, and brightness adjustment for the training set images. The dataset is randomly split into training set, validation set, and test set according to the ratio of 7:1:2, with 1,995,

285, and 570 images respectively. Sample examples of the dataset are presented in Fig. 6.

(a) Stone (b) Mud (c) metal

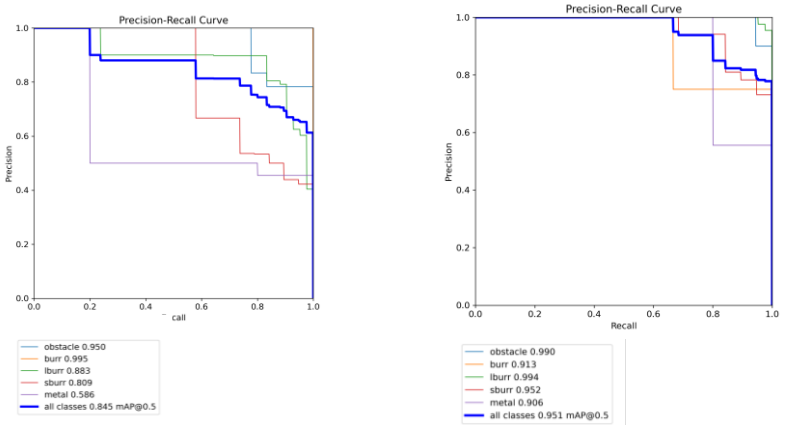
(d) Burr (e) Burr

Fig.6. Foreign objects inside the pipeline

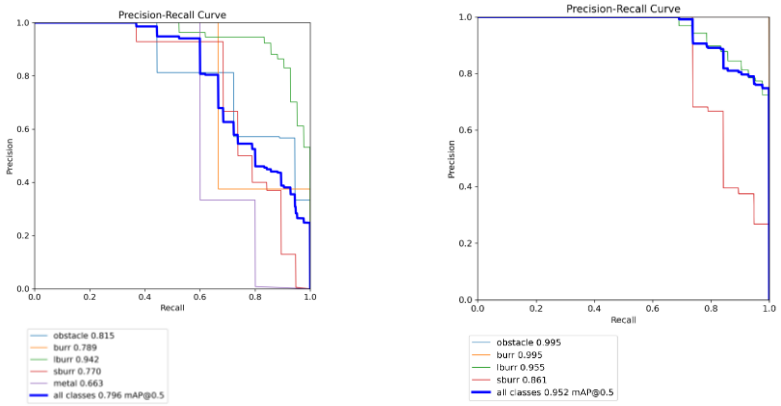
5.2 Comparison Experiment and Result Analysis

To verify the effectiveness of the proposed method, both the YOLOv9-MSAN model, original YOLOv9 model, YOLOv8 model and R-CNN model were trained using the constructed dataset in this study. The training parameters were set as follows: the number of training epochs was 200, the training batch size was 32, and the image input size was 640×640.

Subsequently, the four trained models were used to perform detection experiments on the test set separately. The mean Average Precision (mAP) is compared, and the corresponding results are presented in Fig. 7.



(a) Detection result of the YOLOv9 model (b) Detection result of the optimized model

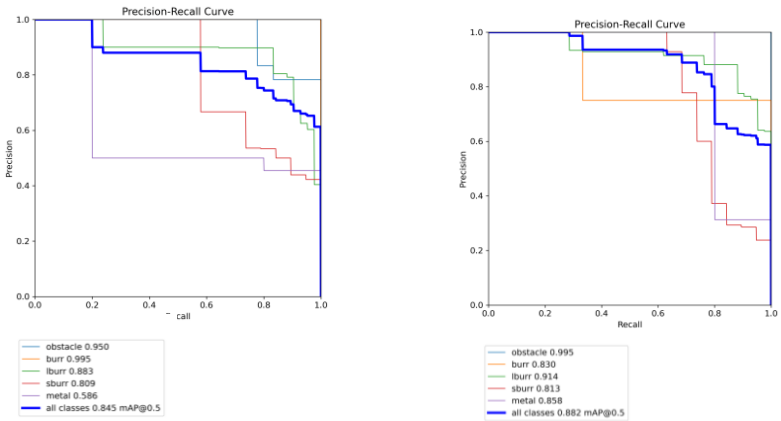


(c) Detection result of the YOLOv8 model (d) Detection result of the R-CNN model

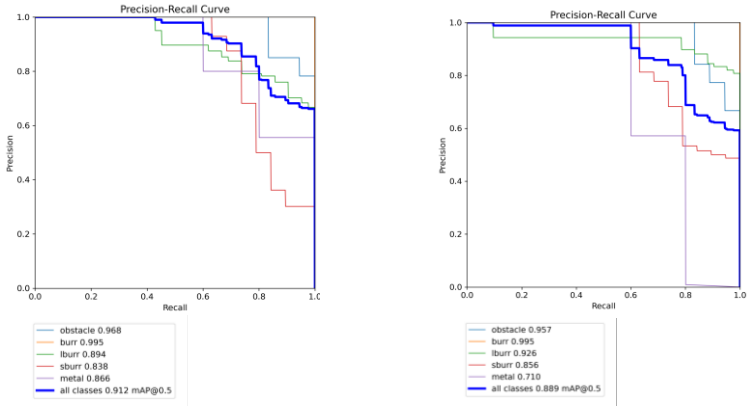
Fig.7. Comparison of detection performance among different models

Fig. 7 presents a comparison of the mean Average Precision (mAP) of the detection results from the two models. The mean Average Precision at an IoU threshold of 0.5 (mAP@0.5) serves as a core metric for evaluating model performance in object detection tasks. Specifically, a detection result is deemed valid when the Intersection over Union (IoU) between the predicted bounding box and the ground-truth box is ≥ 0.5; under this condition, the average precision for each category is calculated. From the results, it can be observed that the detection accuracy of the model improved using this method has been enhanced compared with that of YOLOv8, R-CNN, and YOLOv9.

To clearly demonstrate the improvement effects of the three modifications—gamma correction, MSAN, and WIoU—on the original YOLOv9 model, we also conducted ablation experiments. The experimental results are presented in Fig. 8. According to the experimental results, the three methods proposed in this paper can all significantly improve the detection accuracy.



(a) Detection result of the YOLOv9 model (b) YOLOv9 applying gamma correction



(c) YOLOv9 applying MSAN model (d) YOLOv9 applying WIoU

Fig.8. Comparison of improvement effects

6. Conclusion

The underground cable pipeline foreign object detection method based on the YOLO-MSAN model effectively addresses the problems of low detection accuracy in dark environments and easy missed detection of small-volume foreign objects. It improves the quality of low-light images through gamma correction, introduces multi-core branches and attention mechanisms in the neck of YOLOv9 to strengthen small target feature extraction, and optimizes the loss function to enhance training efficiency and precision.

Future optimizations can be carried out from three aspects: First, enrich the dataset scenarios and samples by supplementing samples such as water accumulation and rust, and refine data augmentation; second, explore model lightweighting by compressing module structures or combining pruning technologies to meet the requirements of real-time detection; third, integrate multi-sensor fusion to achieve accurate judgment of the size and material of foreign objects, assisting in cleaning decision-making.

REFERENCES

[1] *P. Jiang, D. Ergu, F. Liu, Y. Cai, B. Ma,* A Review of Yolo algorithm developments, Procedia Computer Science, Vol. **199**, pp. 1066-1073, 2022. https://doi.org/10.1016/j.procs.2022.01.135

[2] *A. Sharma, V. Kumar, L. Longchamps,* Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and Faster R-CNN models for detection of multiple weed species, Smart Agricultural Technology, Vol. **9**, pp. 100648, 2024. https://doi.org/10.1016/j.atech.2024.100648

[3] *X. Lu, J. Ji, Z. Xing, Q. Miao,* Attention and feature fusion SSD for remote sensing object detection, IEEE Transactions on Instrumentation and Measurement, Vol. **70**, pp. 1-9, 2021. Doi: 10.1109/TIM.2021.3052575

[4] *J. Terven, D. M. Córdova-Esparza, J. A. Romero-González,* A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS, Machine learning and knowledge extraction, Vol. **5**, Iss. 4, pp. 1680-1716, 2023. https://doi.org/10.3390/make5040083

[5] *S. Rani, D. Ghai, S. Kumar,* Object detection and recognition using contour based edge detection and fast R-CNN, Multimedia Tools and Applications, Vol. **81**, Iss. 29, pp. 42183-42207, 2022. Doi: 10.1007/s11042-021-11446-2

[6] *S. H. Kang, J. S. Park,* Aligned matching: improving small object detection in SSD, Sensors, Vol. **23**, Iss. 5, pp. 2589, 2023. https://doi.org/10.3390/s23052589

[7] *C. Liu, X. Shu, L. Pan, J. Shi, B. Han,* Multiscale underwater image enhancement in RGB and HSV color spaces, IEEE Transactions on Instrumentation and Measurement, Vol. **72**, pp. 42183-42207, 2023. Doi:10.1109/TIM.2023.3298395

[8] *X. Li, M. Liu, Q. Ling,* Pixel-wise gamma correction mapping for low-light image enhancement, IEEE Transactions on Circuits and Systems for Video Technology, Vol. **34**, Iss. 2, pp. 681-694, 2023. Doi: 10.1109/TCSVT.2023.3286802

[9] *M. Muniraj, V. Dhandapani,* Underwater image enhancement by combining color constancy and dehazing based on depth estimation, Neurocomputing, Vol. **460**, pp. 211-230, 2021. https://doi.org/10.1016/j.neucom.2021.07.003

[10] *C. Chen, H. Lee, M. Chen,* Steel surface defect detection method based on improved YOLOv9, Scientific Reports, Vol. **15**, Iss. 1, pp. 25098, 2025. https://doi.org/10.1038/s41598-025-10647-1

[11] *T. Diwan, G. Anirudh, J. V. Tembhurne,* Object detection using YOLO: challenges, architectural successors, datasets and applications, Multimedia Tools and Applications, Vol. **82**, Iss. 6, pp. 9243-9275, 2023. https://doi.org/10.1007/s11042-022-13644-y

[12] *A. Vijayakumar, S. Vairavasundaram,* Yolo-based object detection models: A review and its applications, Multimedia Tools and Applications, Vol. **83**, Iss. 35, pp. 83535-83574, 2024. https://doi.org/10.1007/s11042-024-18872-y

[13] *P. K. Sekharamantry, F. Melgani, J. Malacarne,* Deep learning-based apple detection with attention module and improved loss function in YOLO, Remote Sensing, Vol. **15**, Iss. 6, pp. 1516, 2023. https://doi.org/10.3390/rs15061516

[14] *C. Xiong, T. Zayed, E. M. Abdelkader,* A novel YOLOv8-GAM-Wise-IoU model for automated detection of bridge surface cracks, Construction and Building Materials, Vol. **414**, pp.135025, 2024. https://doi.org/10.1016/j.conbuildmat.2024.135025

[15] *P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, A. Haworth,* A review of medical image data augmentation techniques for deep learning applications, Journal of medical imaging and radiation oncology, Vol. **65**, Iss. 5, pp.545-563, 2021. https://doi.org/10.1111/1754-9485.13261