Energy Proceedings
Vol 57, 2025
ISSN 2004-2965

Research on CO_2 Pipeline Leakage Detection Method Based on Lightweight Network

Zhang Lijun[1*], Li Yuchun[1,2], Nian Xingwang[1], Tan Qingxun[1], Wang Zongke[1], Zhang Jingwei[1], Zhou Xinyan[1,2], Liu Jigang[1], Pan Cong[1]

1 Daqing Oilfield Design Institute Co., Ltd, DaQing, 163712

2 State Key Laboratory of Continental Shale Oil, DaQing, 163712

(*Corresponding Author: 544421593@qq.com)

ABSTRACT

In CCUS (Carbon Capture, Utilization and Storage) technology, CO_2 pipeline transportation serves as a crucial link, and its safety directly impacts the reliability of the entire industrial chain. When a CO_2 pipeline leaks, the sudden release of high-pressure CO_2 from the pipeline into a low-pressure environment triggers the Joule-Thomson effect. This physical characteristic creates a significant temperature difference between the leakage area and the surrounding environment. The core advantage of infrared video technology lies in its high sensitivity to temperature field distribution, enabling the clear identification of low-temperature characteristic signals in the leakage area. Therefore, this paper employs infrared video technology for data collection, utilizes image denoising methods to remove impurities from infrared images, and finally adopts the lightweight network MobileNets to detect CO_2 pipeline leaks under three different operating conditions: varying temperatures and pressures, in order to enhance the accuracy of CO_2 pipeline leak detection and reduce response time. Experimental verification shows that, compared with Vgg-16 and ShuffleNet_V2 models, the MobileNetV3_Large model achieves a 100% detection accuracy rate and a faster response time relative to traditional classical models. This demonstrates that lightweight network can effectively distinguish between leakage temperature and pressure differences during CO_2 pipeline leakage, achieving efficient and accurate detection results that can better meet application needs. It plays a positive role in promoting and enhancing CCUS pipeline safety.

Keywords: CCUS, CO_2 pipeline leakage, Lightweight network, Infrared video, Image denoising

NONMENCLATURE

Abbreviations	
APEN	Applied Energy
Symbols	
n	Year

1. INTRODUCTION

CO_2 is a colorless, odorless, and tasteless gas, and its leakage is often not detected by people, but it may cause serious harm to humans and the environment. Therefore, how to achieve early detection and accurate identification of CO_2 leakage has become the key to ensuring safety. In the Carbon Capture, Utilization and Storage (CCUS) technology system, CO_2 pipeline transportation is as an important link in the CCUS process, and its leakage problem has always been a prominent issue restricting the large-scale development of CCUS. The causes of CO_2 pipeline leakage are various reasons, including pipeline aging, corrosion, and human operational errors, etc. After a CO_2 pipeline leaks, due to the Joule-Thomson effect, the temperature drops rapidly, and a large area of high-concentration CO_2 gas cloud will accumulate near the ground in the area close to the leakage point. The low-temperature CO_2 cloud that drifts and spreads with the atmosphere will cause serious asphyxiation hazards to the people who cannot evacuate in time[1]. In addition, high-concentrations of CO_2 can inhibit the photosynthesis of plants and affect their normal growth. CO_2 that leaks into the soil may alter the soil's pH level and disrupt its ecological balance, etc. To effectively address the issue of CO_2 leakage, some commonly used detection techniques have been developed at present, such as sensor

This is a paper for International CCUS Conference 2025 (ICCUSC 2025), May 28-29, 2025, Beijing, China.

detection technology and magnetic flux leakage detection method

Infrared video technology, as an important imaging technology, is based on the physical characteristic that all objects above absolute zero (>-273.15 ℃) emit infrared radiation. It can penetrate obstacles such as darkness, haze, and thick smoke for target detection and monitoring, and is widely used in fields such as national defense security, industrial testing, and medical diagnosis. However, infrared images also have obvious shortcomings, such as low contrast, low resolution, and noise interference, which directly lead to a decrease in image quality, unclear details and contour information of detected targets, and increased difficulty in identifying targets. However, image denoising algorithms can effectively remove impurities from images, restore noise-free images, and at the same time retain most of the details in the image, making the image clearer[2]. At present, many scholars have conducted in-depth research on infrared image denoising methods. For instance, references[3-4] have provided detailed introductions to infrared image denoising methods, covering various types of denoising algorithms and their application examples, which offer significant theoretical support and practical references for the development of infrared image denoising technology.

In recent years, deep learning models have been applied to tasks such as image classification, object detection, and facial recognition, while also supporting real-time applications. Among them, the Convolutional Neural Networks (CNNs) is good at processing image and video data. Its uniqueness lies in using convolutional and pooling layers to efficiently extract and learn image features, while achieving complex pattern recognition through multi-layer nonlinear transformations. During the period of 2012-2017, several classic CNN architectures emerged in the field of computer vision, such as AlexNet, VGGNet, as well as lightweight networks like MobileNet and ShuffleNet. Traditional CNN networks have problems such as large parameter count, high memory usage, and long real-time response time due to their complex structure, which greatly limits their application in industrial environments with limited resources and real-time processing. In contrast, lightweight networks have the advantages of low complexity, small parameter count, low computational load, and high accuracy, and have gradually become a key technology for mobile computer vision tasks. In reference[5], a natural gas pipeline leak aperture recognition method was proposed basing on multi convolutional dense block lightweight neural network; In reference[6], a water pipeline leak detection algorithm was proposed basing on sparse lightweight convolutional neural networks.

Therefore, this paper attempts to use infrared video technology for data acquisition, and then use image denoising methods to remove impurities from infrared images. Finally, a lightweight network is used for CO_2 pipeline leakage detection to enhance the accuracy of CO_2 pipeline leak detection and reduce response time.

2. IMAGE DENOISING

2.1 Image denoising methods

Image denoising is a key technology in image processing, which can effectively improve image quality and provide strong support for subsequent tasks such as image analysis and recognition. In actual scenarios, images are easily affected by factors such as imaging equipment, transmission environment, and lighting conditions, generating various noises, which lead to blurred image details and submerged features. Therefore, noise reduction processing becomes the core link in the image preprocessing stage. Filtering algorithms, as an important means of image denoising and enhancement, play an irreplaceable role in engineering applications and can effectively eliminate noise, extract features, and enhance images, etc. At present, the widely used filtering algorithms in various noise reduction scenarios mainly include Median Filtering (MF), Gaussian Filtering (GF), Non-Local Mean Filtering (NLMF), etc. Their specific characteristics are as follows:

Median Filtering is a non-linear smoothing technique that sets the grayscale value of each pixel to the median of the grayscale values of all pixels within a neighboring window of that point, in order to suppress noise[7]. In the field of image processing, Median Filtering is commonly used to eliminate salt-and-pepper noise while preserving the edge information of the image. The function is expressed as

$$y[n] = \mathrm{median}\left(x[n-k],\ldots,x[n],\ldots,x[n+k]\right) \quad (1)$$

Where *x* represents the original signal, *y* represents the filtered signal, *n* represents the current position, and *k* represents the window size.

Gaussian Filter is a linear smoothing filter, whose core idea is to weight and average each pixel in the image, with pixels closer to the center having higher weights. In this way, Gaussian Filtering can effectively

suppress noise in the image while preserving edge and detail information. The function is expressed as

$$G(x,y)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

Where *x* and *y* represent the coordinates of pixels in the image; σ represents the standard deviation, which controls the smoothness of the filter.

Non-Local Means Filtering is a denoising algorithm based on the global similarity of images, which eliminates noise by integrating information from all similar regions in the image. Compared with traditional local means filtering, Non-Local Means Filtering can better preserve the details of the image while effectively reducing the impact of noise. The function is expressed as

$$NL[\upsilon](i)=\sum_{j\in I}w(i,j)\upsilon(j) \quad (3)$$

Where υ represents the noisy image, $NL[\upsilon]$ represents the restored image, *i* represents any pixel in the restored image, and *w* represents the degree of similarity between the current pixel and the remaining pixels.

2.2 *Image evaluation indicators*

To further quantitatively evaluate the impact of image denoising methods on image quality, and to ensure that the denoised image retains as much of the original image's details and visual effects as possible while removing noise. Two indicators, Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM)[8], are used for evaluation.

The principle of Peak Signal to Noise Ratio detects the quality of restored image by comparing the original image with the image after adding noise[9]. The higher the value, the more similar the image is. The unit is dB and the value range is [20,40]. The function is expressed as

$$\mathrm{MSE}=\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[X(i,\ j)-X'(i,\ j)]^2 \quad (4)$$

$$\mathrm{PSNR}=10\log_{10}\left[\frac{\mathrm{MAX}_X^2}{\mathrm{MSE}}\right] \quad (5)$$

Where *X* and *X'* represent the original image and the image with added noise, respectively; *m* and *n* represent the dimensions of the image; MAX represent the maximum pixel value of an image.

Structural Similarity is a comprehensive assessment of an image from three dimensions: brightness, contrast, and structure[10]. The larger the value, the more similar the image. The value range is [0,1]. The function is expressed as

$$\mathrm{SSIM}(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \quad (6)$$

Where μ and σ represent the mean and variance of the image block, respectively; σ_{xy} represents the covariance between image block *x* and image block *y*; C_1 and C_2 represent constants and can generally be set to 0.

3. MOBILENETS LIGHTWEIGHT NETWORK

MobileNets is a lightweight convolutional neural network series designed by Google researchers specifically for mobile devices and embedded systems. The core innovation is the proposal of Depthwise Separable Convolution (DSC), which decomposes standard convolutions into two steps:

(1) Depthwise Convolution (DC) : Spatial filtering is independently performed on each input channel, and the number of output channels is the same as that of the input.

(2) Pointwise Convolution (PC) : Cross-channel information is fused through 1×1 convolution to adjust the output dimension.

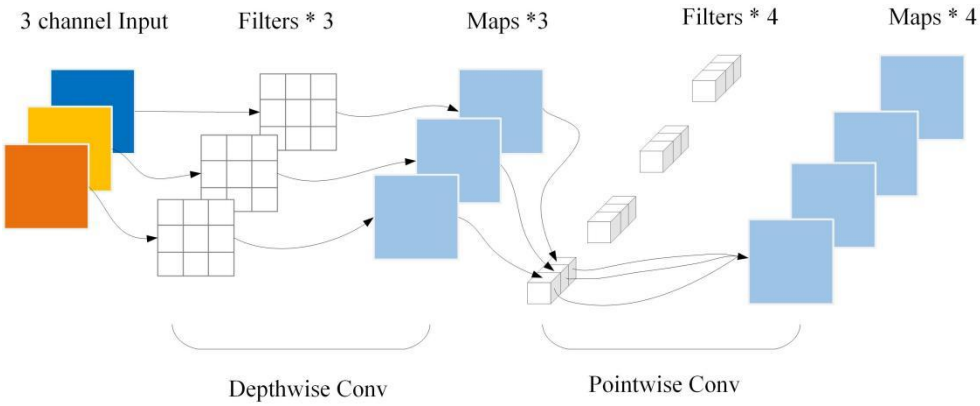


Fig. 1. The structure of Depthwise Separable Convolutional

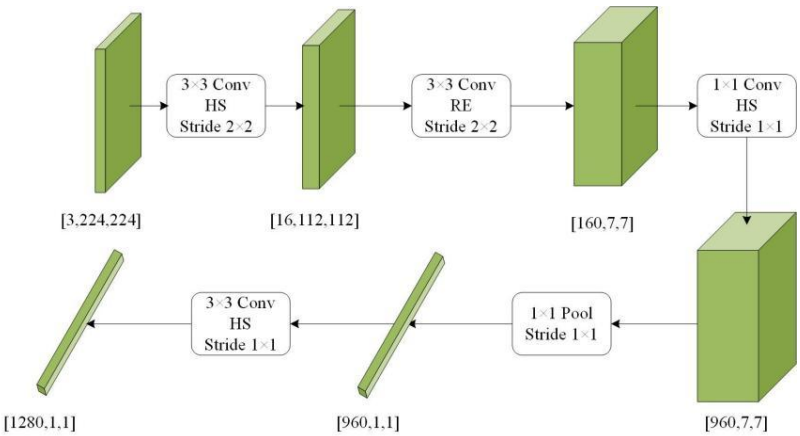


Fig. 2. Network architecture of MobileNetV3_Large

The design of the MobileNets model reduces the computational load of 3×3 convolution kernels to 1/8 to 1/9 of that of standard convolution, significantly reducing the number of parameters and computational costs. The structure of Depth Separable Convolution is shown in Figure 1.

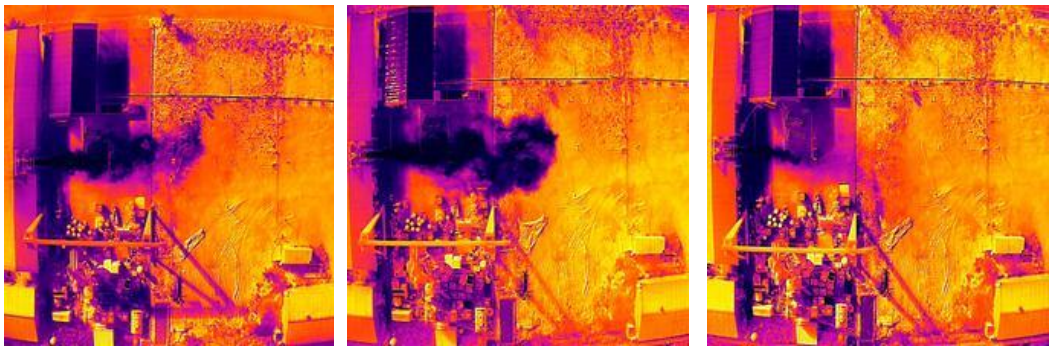
MobileNets currently comprises numerous variations that consist of MobileNetV1, MobileNetV2, and MobileNetV3. The third-generation MobileNetV3 continues the Depthwise Separable Convolution of the first-generation MobileNetV1 and the Inverted Residual structure of the second-generation MobileNetV2. At the same time, technologies such as Neural Architecture Search (NAS), attention mechanism of SE module, and the novel activation functions of H-Swish were introduced to further optimize the efficiency and performance of the model, making it a benchmark model on mobile devices. The MobileNetV3 model is divided into two versions: Large and Small. For the purpose of this study, we selected the MobileNetV3_Large model, which is suitable for high-precision tasks. The network structure of MobileNetV3_Large model is illustrated in Figure 2.

4. EXPERIMENTAL RESEARCH

Build an industrial-scale CO_2 pipeline leakage experimental setup, consisting of five major parts: a 22m-long main leakage pipeline, a 0.3m-long leakage control system, a CO_2 injection system, a phase state control system, and a data acquisition system. The size of the leakage port is 1% of the main pipeline diameter. During the test, infrared videos are captured using a drone, and the operating conditions are as shown in Table 1. Then, Python is used for video image preprocessing to convert the three-dimensional video images into 224 × 224 thumbnail images of two-dimensional images. The top view of the CO_2 pipeline leakage is shown in Figure 3.

Table 1 Comparison of different working conditions

Working Condition	Pressure /MPa	Temperature /℃	Number of samples
01	9.5	25	2800
02	10	35	2776
03	10.5	41	924



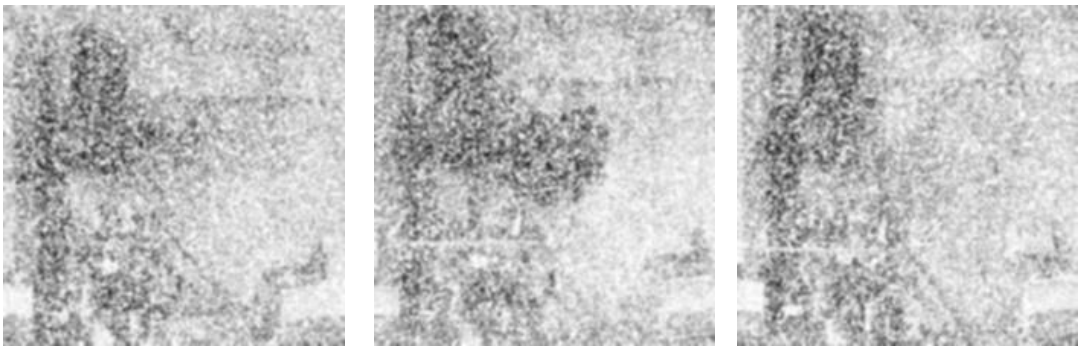
（a）Working condition 01　（b）Working condition 02　（c）Working condition 03

Fig. 3. Infrared images of CO_2 pipeline leakage under different conditions

After the original infrared image was processed using three methods: Median Filtering, Gaussian Filtering, and Non-Local Means Filtering, the evaluation results of PSNR and SSIM indicators are shown in Table 2. Through comparison, it was found that the PSNR and SSIM indicator values of the image processed by Non-Local Means Filtering were higher than those of Median Filtering and Gaussian Filtering, which is more in line with the requirements of image quality evaluation indicators. Therefore, this paper chooses the Non-Local Means Filtering method, and the processed image is shown in Figure 4.

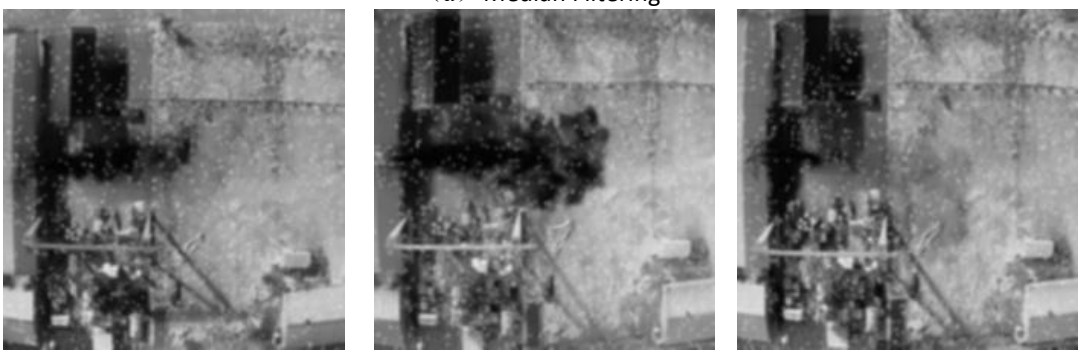
Table 2 Evaluation results of indicators

Noise reduction methods	PSNR/dB	SSIM
Median Filtering	22.51	0.6372
Gaussian Filtering	24.16	0.7636
Non-Local Means Filtering	30.97	0.8930



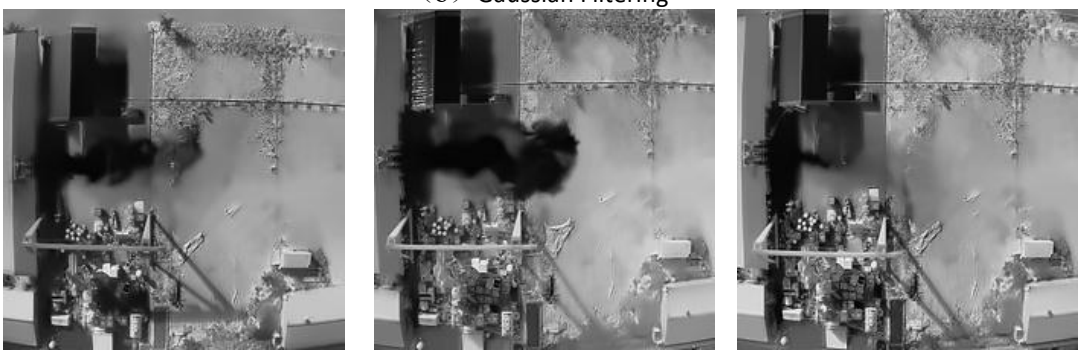
（1）Working condition 01　（2）Working condition 02　（3）Working condition 03

（a）Median Filtering



（1）Working condition 01　（2）Working condition 02　（3）Working condition 03

（b）Gaussian Filtering



（1）Working condition 01　（2）Working condition 02　（3）Working condition 03

（c）Non-Local Means Filtering

Fig. 4. Comparison of denoised images

5. RESULT ANALYSIS

In this paper, infrared video images under three different operating conditions are collected, and image samples are generated after Non-Local Means Filtering. Assuming that the number of training sets accounts for 4/5 of the total number of datasets, there are a total of 5200 sets as training samples and 1300 sets as test samples. During the training process of each model, the parameter settings are as follows: the optimizer is set SGD, the learning rate (LR) is 0.001, the number of epochs is 20, the batch size is 32, and the loss function is CrossEntropyLoss(). The experimental results are shown in Tables 3.

Table 3 The comparison of running results of different model

Model	Network parameters /MB	Accuracy /%	Iterate time/s
Vgg-16	537.17	100	925.85
ShuffleNet_V2	8.79	85.15	109.8
MobileNetV3_Large	5.4	100	129.85

In Table 3, the network parameter size of the MobileNetV3_Large model is 5.4MB, while that of the Vgg-16 model is 537.17MB. Both models achieve a recognition accuracy of 100%. In terms of single iteration time, the Vgg-16 model is seven times longer than the MobileNetV3_Large model, further demonstrating that the larger the network parameters, the longer the computation time. The ShuffleNet_V2 model has a network parameter size of 8.79MB and a single iteration time of 109.8s, which is lower than that of the MobileNetV3_Large model. However, its recognition accuracy is only 85.15%, lower than that of both the Vgg-16 and MobileNetV3_Large models. Overall, the MobileNetV3_Large model achieves high recognition accuracy in a relatively short time, fully demonstrating that after denoising processing, infrared images can be effectively recognized using the lightweight MobileNetV3_Large model. At the same time, this method reduces the requirements for hardware systems and enhances real-time responsiveness.

6. CONCLUSIONS

The intelligent recognition method for CO_2 pipeline leaks based on CCUS using infrared images is a promising technical approach. By deeply researching infrared imaging technology and image denoising methods, and combining them with the practical needs of CCUS technology, an efficient and accurate CO_2 pipeline leakage detection method can be developed. Comparative analysis of the performance of different deep learning models reveals that the MobileNetV3_Large model achieves a recognition accuracy of 100%, and its response time falls between that of the Vgg-16 and ShuffleNet_V2 models, fully demonstrating the superiority of the MobileNetV3_Large model in CO_2 pipeline leakage detection. In the future, with continuous technological advancements and expanding applications, lightweight networks are expected to play a greater role in the CCUS field, making significant contributions to achieving the goal of "carbon neutrality".

ACKNOWLEDGEMENT

The present work was funded by the Natural Science Foundation of Heilongjiang Province's key project "Research on Transient Pressure Reduction Mechanism and Resilience Crack Prevention Model of Supercritical and Dense Phase Carbon Dioxide Pipeline" (ZD2024E010).

DECLARATION OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors read and approved the final manuscript.

REFERENCE

[1] Wang Rungang. Study on leakage and diffusion of supercritical CO_2 transportation pipeline[J]. Petrochemical Industry Technology, 2016, 23(12): 28.

[2] Bai Ruxin, Luan Shangmin. Research Review of the Application of Image Denoising Algorithms in Image Processing[J]. Modern Information Technology, 2024, 8(10): 21-25+31.

[3] Zhao Zihui, Zhou Yongkang, Zeng Bangze, et al. A Review of Infrared Image Denoising[J]. Infrared Technology, 2025, 47(03): 299-306.
[4] Wang Yang, Pan Zhibin. Review of De-noise and Enhancement Technology for Infrared Image[J]. Radio Engineering, 2016, 46(10): 1-7+28.
[5] Sun Jiedi, Wang Lixuan, Wen Jiangtao, et al. Pipeline leakage aperture recognition based on lightweight neural network with the improved dense block[J]. Chinese Journal of Scientific Instrument, 2022, 43(03): 98-108.
[6] Liu Jie, Zhu Zhengwei. Pipeline leakage detection algorithm based on sparse and lightweight convolutional neural network[J]. Electronic measurement technology, 2022, 45(19): 131-135.
[7] Zhu Liuzhong, Li Jingze, Wu Lei, et al. Optical fiber defect detection method based on Gabor filter[J]. Journal of Nanjing University of Science and Technology, 2022, 46(01): 119-126.
[8] Chen Qiuyan, Zhang Xinyan, He Min, et al. Identification of Pipeline Thermal Image Leakage Based on Deep Learning[J]. Infrared Technology, 46(05): 522-531.
[9] Li Zhengde, Zhao Shuang, Chen Mingzhe, et al. Zero-shot learning-based medical image denoising algorithm[J]. Optical Technique, 2025, 51(03): 352-356.
[10] Liu Yi, Zhu Jiahui, Zheng Dichen, et al. Lightweight Low- light Image Enhancement Method Based on Multi- attention Feature Fusion[J]. Journal of Data Acquisition and Processing, 1-18[2025-07-11]. http://kns.cnki.net/kcms/detail/32.1367.TN.20250711.1339.037.html.