

Crack Image Detection and Edge Feature Detection by Introducing Lightweight Network

Quanyi Guo, Chen Chong

Electronic Information School, Wuhan University, Wuhan, Hubei, P.R.China. 430072

Abstract. In order to optimize the crack image detection technology and realize edge feature detection, an edge detection model based on Visual Geometry Group (VGG) algorithm is constructed based on multi-scale supervised model and convolutional neural network (CNN). Then, a lightweight network is introduced to optimize the model, and a significant analysis is made with the classical algorithm. Compared with the mean absolute error (MAE), the performance advantage of the optimization algorithm is further verified. The research results show that the edge detection algorithm of multi-scale supervised model proposed in this study is superior to the traditional algorithm in all datasets. The F-measure values in four different datasets are 0.86, 0.93, 0.93 and 0.88 respectively, which shows the superior performance of the algorithm. Meanwhile, the MAE values of the algorithm based on lightweight network optimization in four different datasets are 0.04, 0.03, 0.03 and 0.06 respectively. Compared with other classical algorithms, the optimization algorithm has achieved the lowest MAE values, which proves that the optimization algorithm has significant advantages and higher accuracy in edge detection tasks. This study improves the accuracy and efficiency of crack image detection, which is of great significance to promote the development and application of crack image detection technology.

Keywords: Lightweight network; Crack image detection; Edge features; Multi-scale supervision; Convolutional Neural Network.

1. Introduction

Crack detection plays an important role in the fields of material science, civil engineering and structural health monitoring. The infrastructure such as buildings, roads and bridges involved in these fields bear the economic development of society and the safety of people's lives. In the long-term operation of infrastructure, the appearance of cracks is a common sign of structural damage and potential danger. Therefore, timely and accurate detection and location of cracks is of immeasurable value in preventing structural failure, ensuring structural safety, reducing maintenance costs and prolonging service life. However, traditional crack detection methods, such as manual visual inspection, are intuitive but have many limitations. Firstly, manual visual inspection requires professionals to have rich experience and skills, and the inspection process is time-consuming and labor-intensive and inefficient. Secondly, manual inspection is easily influenced by human factors, such as fatigue and negligence, which makes it difficult to guarantee the accuracy of the test results. In addition, it is often difficult to identify complex fracture shapes and hidden fracture locations effectively by manual inspection [1-3]. In order to solve these problems, the development of efficient and accurate automatic crack image detection technology has become the focus of current research.

With the rapid development of deep learning technology, convolutional neural network (CNN) has made remarkable achievements in the field of image detection. CNN can automatically learn the feature information in the image by simulating the connection and working mode of human brain neurons, and carry out efficient classification and detection [4]. Especially, Visual Geometry Group (VGG) algorithm, as a classical CNN model, has shown great potential in image classification, target detection and other tasks with its deep network structure and good generalization ability. However, when VGG algorithm is directly applied to crack image detection, it still faces some challenges. Firstly, fracture images usually have complex and changeable shapes and scales, and feature

extraction with a single scale is difficult to meet the actual needs. In order to fully capture the multi-scale features in fracture images, it is necessary to design a model structure that can fuse the feature information of different scales. Secondly, the traditional CNN model usually has huge parameters and computation, which leads to the model consuming a lot of computing resources and time in the process of training and reasoning. In real-time and embedded system applications, there are strict requirements for the calculation and parameters of the model. Therefore, it is necessary to introduce lightweight network to optimize the model to reduce the calculation and parameters of the model and improve the reasoning speed and real-time performance of the model [5,6].

Here, CNN and multi-scale supervised model are combined to construct an edge detection model based on VGG, and the significance is analyzed with the classical algorithm. The F-measure value is compared and analyzed to evaluate the performance of the algorithm on different datasets. Secondly, the lightweight network is introduced to optimize the model, and the MAE is compared with the classical algorithm. The innovation of this study lies in the fusion of CNN and multi-scale supervision model, and the edge detection model based on VGG is constructed. This fusion method makes full use of the information at different scales, and improves the detection ability of the model for crack edge features. In addition, according to the characteristics of the field of crack image detection, a lightweight network structure is designed, which effectively reduces the computational complexity and memory occupation of the model, while maintaining a high detection accuracy. This optimization strategy not only improves the practicability and real-time performance of the model, but also provides a feasible solution for the application of crack image detection technology, and provides new ideas and methods for the further development and application of this field.

2. Literature Review

With the continuous development of deep learning technology in recent years, the research and application of lightweight network in the fields of image detection and edge feature detection have been paid more and more attention. The research of Aiadi et al. (2023) showed that the use of lightweight network can effectively reduce the computational complexity of the model and speed up the speed of image detection, especially in the environment with limited resources such as mobile terminals, which had better real-time performance [7]. Alansari et al. (2023) found that the small model size and low computing requirements of lightweight network made it very suitable for image detection applications on embedded systems, such as smart cameras and smart phones, and can realize real-time image processing and analysis [8]. Zhang et al. (2023) found that lightweight network can reduce the number of layers and channels, thus reducing the parameters and computational complexity of the model. The simplified model structure can not only reduce the computational cost, but also improve the training and reasoning speed of the model [9]. Chen et al. (2023) found that edge feature detection usually required complex calculation and analysis of images. The traditional deep neural networks often had high computational complexity, which was not suitable for edge feature detection tasks. While lightweight networks can effectively reduce the computational complexity and improve the efficiency of edge feature detection by reducing the parameters and model complexity [10].

To sum up, lightweight network plays an important role in the field of image detection and edge feature detection. It can effectively reduce the computational complexity of the model, speed up image detection, and show better real-time performance in the environment with limited resources such as mobile terminals. However, despite its potential in many fields, there are still some gaps in the research of lightweight network in crack image detection.

2.1 Research Methodology

2.2 Crack image

Crack image refers to the image of cracks formed in the surface or structure of materials. These cracks may be caused by material fatigue, stress concentration, external damage and other factors. In the

engineering field, it is very important to identify and analyze crack images for evaluating the health status of materials, predicting their life and formulating corresponding maintenance and repair strategies [11].

Crack images usually have irregular shapes and complex textures, and the size, shape and direction of cracks are varied. In addition, crack images are often influenced by factors such as illumination, angle and environment, which makes the edges of cracks blurred and even integrated with the background. Therefore, the characteristics of crack images include diverse shapes, complex textures, blurred edges and great influence by environmental factors, which bring certain challenges to the detection and analysis of crack images [12].

2.3 CNN

CNN is an artificial neural network specially used for processing data with grid structure, and it has achieved great success in the fields of image detection and speech detection. CNN usually consists of convolution layer, pooling layer, activation function, fully connected layer, etc. The convolution layer is mainly used for feature extraction, pooling layer is used for reducing feature dimension, and fully connected layer is used for mapping extracted features to output categories [13].

VGG network is a classical CNN architecture, which is famous for its concise and clear architecture and small convolution kernel size. Its basic architecture consists of a plurality of convolution layers and pooling layers alternately stacked, followed by a plurality of fully connected layers, and finally outputs the classification results. VGG network has achieved good performance in image classification, target detection and other tasks, and has become one of the classic models in the field of deep learning [14].

Let the input characteristic map be X , the convolution kernel be W and the offset be B , then the calculation equation of convolution operation is as follows:

$$C(X, W) = \sum_{n,m} X_{n,m} \cdot W_{n,m} + B \quad (1)$$

For the case of multiple input channels and multiple output channels, the convolution kernel is a three-dimensional tensor, and each channel needs to be convolution independently. ReLU activation function nonlinearly maps the output of convolution layer, and sets the negative value to zero, keeping the positive value unchanged.

In order to reduce the dimension of the feature graph, the maximum and average pooling operations are generally adopted, and the calculation equations of the maximum and average pooling operations are as follows:

$$MP(X) = \max_{n,m} X_{n,m} \quad (2)$$

$$AP(X) = \frac{1}{ij} \sum_{n,m} X_{n,m} \quad (3)$$

ij is the size of the pooled window. In the fully connected layer, if the output of the previous layer is flattened into a one-dimensional vector x , the calculation equation of the output of the fully connected layer is as follows:

$$FC(x, W, B) = xW + B \quad (4)$$

$$output = ReLU(xW + B) \quad (5)$$

2.4 Edge feature detection

Edge feature detection is an important task in the field of computer vision, which aims to identify the edges or contours of objects and scenes in images. In recent years, deep learning technology has made remarkable progress in the field of edge feature detection. Multi-scale supervised model is a deep learning model, which processes input data of different scales through multiple branches or processing paths and fuses them to improve the model's ability to capture multi-scale features of

images, thus improving the performance and generalization ability of the model in image processing tasks [15].

In this study, the multi-scale supervised model is used as the basic model of edge feature detection, and the structure of the multi-scale supervised model is shown in Figure 1.

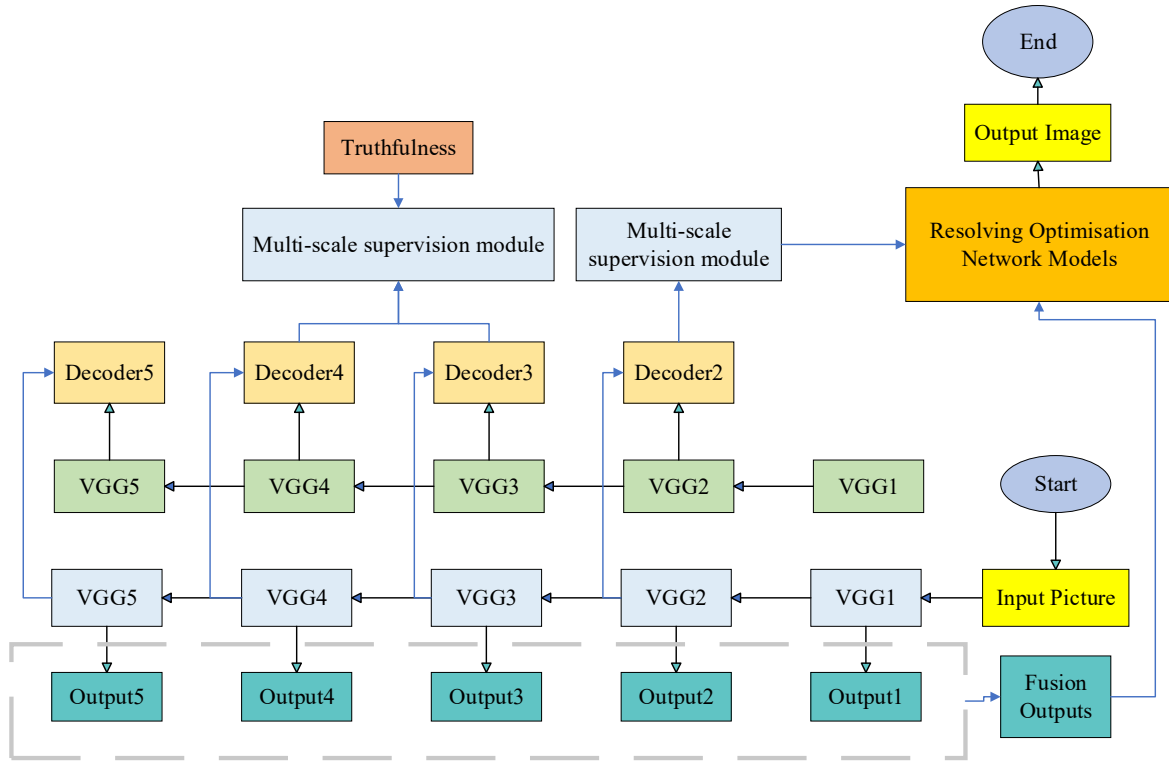


Figure 1: Structure of multi-scale supervision model

2.5 Lightweight network

Lightweight network refers to adopting a series of optimization strategies, such as reducing the number of parameters, model depth and computational complexity to maintain high performance and reduce the storage and computational overhead of the model. This design enables the lightweight network to operate efficiently in the resource-constrained environment such as mobile devices and embedded systems, and provides faster and more energy-saving intelligent services for smart phones, smart cameras and other devices [16].

2.6 Image detection model based on lightweight network optimization

Based on lightweight network optimization, this study designs a model for image detection. Firstly, the lightweight basic network structure is adopted, which combines the idea of dense connection and deeply separable modules, effectively reducing the calculation and storage overhead of the model and improving the feature representation ability. Secondly, a lightweight decoding module is designed. The features are further processed and reorganized by the depth separable module, and the number of feature channels is controlled. Meanwhile, the processed feature map is enlarged by the traditional bilinear interpolation and transposed convolution methods, and finally a one-channel saliency prediction map is output to realize the goal of lightening the model and balancing the computational cost and the model effect. The structure of image detection model based on lightweight network optimization in this study is shown in Figure 2.

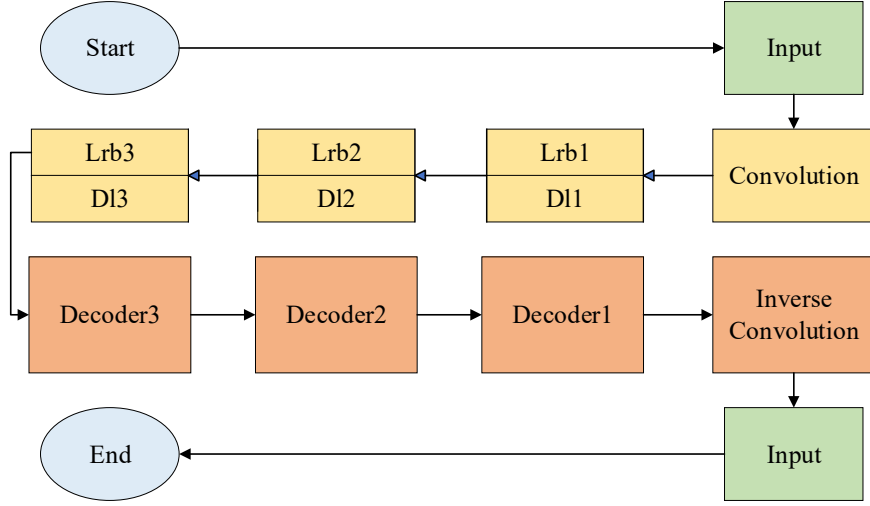


Figure 2: Image detection model structure based on lightweight network optimization
The structure of lightweight network decoder used in this study is shown in Figure 3.

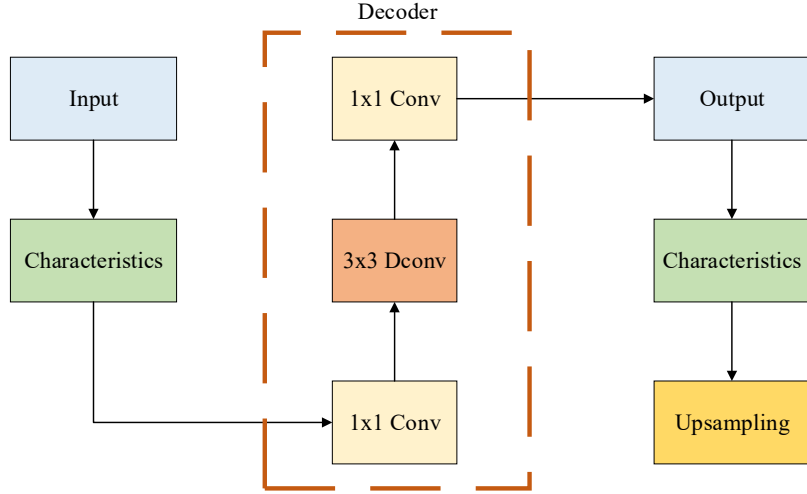


Figure 3: Lightweight network decoder structure

3. Experimental Design and Performance Evaluation

3.1 Datasets Collection

Detection of Unsanitary Targets Dataset (DUTS) is used as the main data source in this study. And it is tested on four public datasets: DUTS, Hong Kong University Image Segmentation Dataset (HKU-IS), Extended Complex Scene Saliency Dataset (ECSD) and Pascal Salient Object Dataset (Pascal-S). DUTS dataset is a commonly used image segmentation dataset, which contains images of various scenes and objects, and is suitable for evaluating the performance of image segmentation algorithms. HKU-IS, ECSSD and PASCAL-S are the benchmark datasets for image segmentation tasks, which are widely used to evaluate the generalization ability and performance of the algorithm.

In this study, Region-based Fully Participatory Networks (R-FCN), Non-Local Deep Features (NLDF), Unsupervised Common Fate (UCF), Domain Guided Relationship Learning (DGRL) and Progressive and Adaptive Graph Learning (PAGL) are compared and analyzed significantly [17-20]. In this study, the resolution of the image is set to 288×288 , and the VGG network adopts Adam optimizer, with a learning rate of 0.01, 50 iterations and a batch size of 256. The evaluation indexes are F-measure(F-m) and Mean Absolute Error (MAE).

3.2 Performance Evaluation

The comparison results of significance test between the multi-scale supervised model edge detection algorithm and the classical algorithm model used in this study are shown in Figure 4.

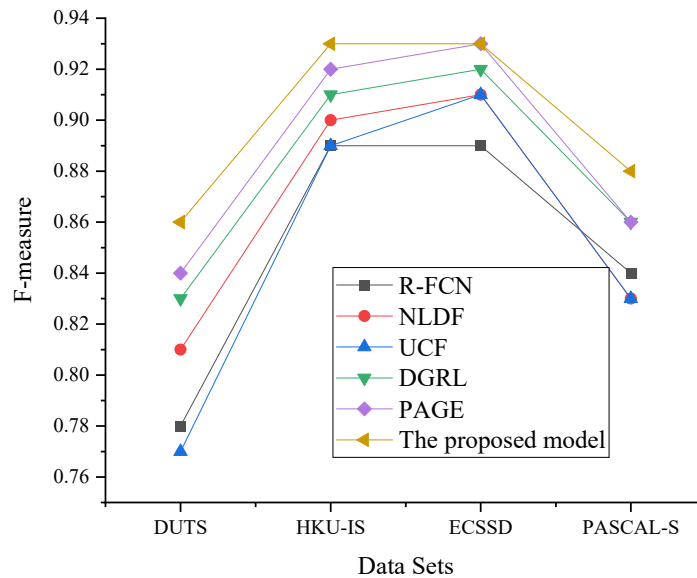


Figure 4: The comparison results of significance test between this research algorithm and the classical algorithm model

Analysis of Figure 4 shows that the F-m values of R-FCN, NLDF and UCF are 0.78, 0.81 and 0.77 respectively, while the F-m values of DGRL and PAGE are 0.83 and 0.84 respectively. The algorithm in this study performs best on this dataset, and the F-m value is 0.86. On HKU-IS dataset, the F-m values of R-FCN and UCF are both 0.89, while those of NLDF, DGRL and PAGE are 0.90, 0.91 and 0.92, respectively. The algorithm in this study is also outstanding on this dataset, and the F-m value is 0.93. On the ECSSD dataset, the F-m value of R-FCN is 0.89, the F-m values of NLDF and UCF are 0.91, and the F-m values of DGRL and PAGE are 0.92 and 0.93, respectively. The algorithm in this study is the same as PAGE on this dataset, and the F-m value is both 0.93. On PASCAL-S dataset, the F-m values of R-FCN, NLDF and UCF are 0.84, 0.83 and 0.83, respectively, and the F-m values of DGRL and PAGE are 0.86. The algorithm in this study performs best again on this dataset, and the F-m value is 0.88. To sum up, the multi-scale supervised model edge detection algorithm proposed in this study performs well on all four datasets, and the F-m value is the highest in all datasets.

The comparison results of significance test between this research algorithm and the classical algorithm model based on lightweight network optimization are shown in Figure 5.

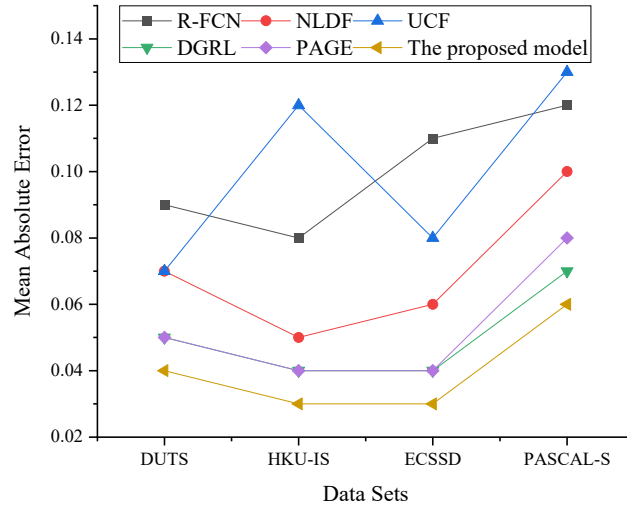


Figure 5: Comparison of the significance test between this research algorithm and the classical algorithm model for lightweight network optimization

According to the analysis of Figure 5, the MAE value of R-FCN on all datasets is relatively high, which is 0.09 on DUTS dataset, 0.08 on HKU-IS, 0.11 on ECSSD and 0.12 on Pascal-S. The MAE value of NLDF is lower than that of R-FCN, which is 0.07 in DUTS, 0.05 in HKU-IS, 0.06 in ECSSD and 0.10 in Pascal-S. The MAE values of UCF on HKU-IS and PASCAL-S datasets are relatively high, which are 0.12 and 0.13 respectively. The performance of DGRL and PAGE is similar, and the MAE values of all datasets are low. The MAE value of this algorithm is the lowest on all datasets, with DUTS of 0.04, HKU-IS of 0.03, ECSSD of 0.03 and PASCAL-S of 0.06. To sum up, the performance of this algorithm based on lightweight network optimization is significantly better than that of traditional algorithms on various datasets, with the lowest MAE value and the smallest error.

4. Conclusion

In this study, the convolution neural network and multi-scale supervised model are combined to construct an edge detection model based on VGG, and the significance of the model is analyzed with the classical algorithm to evaluate the performance of the algorithm on different datasets. Secondly, the lightweight network is introduced to optimize the model, and compared with the classical algorithm, which further proves the performance advantage of the optimization algorithm. The research results show that the edge detection algorithm of multi-scale supervised model proposed in this study is superior to the traditional algorithm in all datasets. The F-measure value is 0.86 on DUTS dataset, 0.93 on HKU-IS dataset, 0.93 on ECSSD dataset and 0.88 on PASCAL-S dataset. These results are significantly higher than other classical algorithms, which shows the superior performance of this algorithm. The research algorithm based on lightweight network optimization has the lowest MAE value on DUTS dataset, such as 0.04, 0.03 for HKU-IS, 0.03 for ECSSD and 0.06 for PASCAL-S, which shows that the algorithm has significant advantages and higher accuracy in edge detection. The deficiency of this study is that it adopts a simple model structure, and different levels of features can be optimized and integrated in the future to improve the performance of the model.

References

- [1] Khan M A M, Kee S H, Pathan A S K, et al. Image Processing Techniques for Concrete Crack Detection: A Scientometrics Literature Review[J]. Remote Sensing, 2023, 15(9): 2400.
- [2] Chen D, Huang B, Kang F. A review of detection technologies for underwater cracks on concrete dam surfaces[J]. Applied Sciences, 2023, 13(6): 3564.
- [3] Shim S. Self-training approach for crack detection using synthesized crack images based on conditional generative adversarial network[J]. Computer-Aided Civil and Infrastructure Engineering, 2024, 39(7): 1019-1041.

- [4] Taye M M. Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions[J]. *Computation*, 2023, 11(3): 52.
- [5] Tabrizchi H, Parvizpour S, Razmara J. An improved VGG model for skin cancer detection[J]. *Neural Processing Letters*, 2023, 55(4): 3715-3732.
- [6] Shah S R, Qadri S, Bibi H, et al. Comparing inception V3, VGG 16, VGG 19, CNN, and ResNet 50: a case study on early detection of a rice disease[J]. *Agronomy*, 2023, 13(6): 1633.
- [7] Aiadi O, Khaldi B, Saadeddine C. MDFNet: An unsupervised lightweight network for ear print recognition[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(10): 13773-13786.
- [8] Alansari M, Hay O A, Javed S, et al. Ghostfacenets: Lightweight face recognition model from cheap operations[J]. *IEEE Access*, 2023, 11(1): 35429-35446.
- [9] Zhang D, Hao X, Wang D, et al. An efficient lightweight convolutional neural network for industrial surface defect detection[J]. *Artificial Intelligence Review*, 2023, 56(9): 10651-10677.
- [10] Chen N, Li Y, Yang Z, et al. LODNU: lightweight object detection network in UAV vision[J]. *The Journal of Supercomputing*, 2023, 79(9): 10117-10138.
- [11] Lin S L. Research on tire crack detection using image deep learning method[J]. *Scientific reports*, 2023, 13(1): 8027.
- [12] Zhu Z, Huang D, Zhou X, et al. Image quality evaluation method for surface crack detection based on standard test chart[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2024, 39(9): 1294-1317.
- [13] Salehi A W, Khan S, Gupta G, et al. A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope[J]. *Sustainability*, 2023, 15(7): 5930.
- [14] Lu S, Xia K, Wang S H. Diagnosis of cerebral microbleed via VGG and extreme learning machine trained by Gaussian map bat algorithm[J]. *Journal of ambient intelligence and humanized computing*, 2023, 14(5): 5395-5406.
- [15] Sui B, Cao Y, Bai X, et al. BIBED-Seg: Block-in-block edge detection network for guiding semantic segmentation task of high-resolution remote sensing images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, 16(1): 1531-1549.
- [16] Lawal O M. YOLOv5-LiNet: A lightweight network for fruits instance segmentation[J]. *Plos one*, 2023, 18(3): e0282297.
- [17] Cen H. Target location detection of mobile robots based on R-FCN deep convolutional neural network[J]. *International Journal of System Assurance Engineering and Management*, 2023, 14(2): 728-737.
- [18] Zhao Y, Zou J, Liu S, et al. Terrace Extraction Method Based on Remote Sensing and a Novel Deep Learning Framework[J]. *Remote Sensing*, 2024, 16(9): 1649.
- [19] Mazzia F, Chan R. Alma Collections as a promotional tool for government documents: By Faye Mazzia and Robin Chan, University of Central Florida[J]. *Journal of Electronic Resources Librarianship*, 2022, 34(4): 341-344.
- [20] Aranguren C, Aguilera R. Application of self-learning enhancing Huff and Puff gas injection optimization in shale reservoirs through sequence-based proxy reservoir simulation and reinforcement learning[J]. *Geoenergy Science and Engineering*, 2024, 234(1): 212676.