IEICE TRANSACTIONS

on Fundamentals of Electronics, Communications and Computer Sciences

DOI:10.1587/transfun.2024EAP1051

Publicized:2024/09/27

A PUBLICATION OF THE ENGINEERING SCIENCES SOCIETY

The Institute of Electronics, Information and Communication Engineers

Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3 chome, Minato-ku, TOKYO, 105-0011 JAPAN

PAPER

An Accurate Tunnel Crack Identification Method Integrating Local Segmentation and Global Fusion Detection

Baoxian Wang[†], ***Member,*** **Ze Gao**[†], **Hongbin Xu**[††], **Shoupeng Qin**[†††], **Zhao Tan**[†††]
and **Xuchao Shi**[††††], ***Nonmembers***

SUMMARY This paper proposes a novel approach for tunnel crack identification, employing local segmentation and global fusion detection. Initially, a local segmentation network is constructed using weights from an encoding layer pre-trained on numerous non-tunnel cracks, with only a limited number of tunnel crack samples. The input image is divided, and the local segmentation network performs pixel segmentation on these sub-images, with the sub-results stitched together to ensure accurate identification of all suspicious crack pixels. Subsequently, a global fusion detector is introduced, comprising two sub-models: Sub-model 1 extracts total crack targets within the stitched results, while Sub-model 2 detects possible false alarms from regular-shaped areas. The results from both sub-models are combined to effectively reduce the false alarm rate and ensure accurate segmentation results of cracks. Experimental findings on actual tunnel images demonstrate that the "segmentation before localization" method proposed in this paper achieves superior recognition accuracy and IOU ratio compared to the Unet3+, DeeplabV3+, and "localization before segmentation" Mask-RCNN algorithms. Specifically, the proposed method yields an accuracy improvement of 3.81% over the Unet3+ network, 2.71% over the DeeplabV3++ network, and 1.93% over the Mask-RCNN network. Moreover, noise interference from bolt repair areas is effectively mitigated, enhancing the method's engineering applicability.

key words: *crack identification; local segmentation; global fusion detection; transfer learning.*

1. Introduction

Tunnel surface cracks are characterized by weak, linear features with low signal-to-noise ratio, present within the complex tunnel environment[1]. Studies show that traditional image processing methods[2]-[6] face challenges in achieving accurate crack identification results.

With the rapid development of computational methods, deep learning networks[7]-[9] have been extensively employed in crack segmentation due to their robust feature extraction capabilities. For instance, Wang al. [10] proposed the integration of a dilated convolution module into the Unet++ network to significantly enhance crack segmentation performance. Wu al.[11] introduced a crack detection method based on enhanced Retinex and VGG19. Wang al.[12] utilized SE-ResNet50 to develop the SU-ResNet++ algorithm for tunnel disease detection. Sohaib al.[13] employed multiple quantized YOLOv8 models for crack detection. Li al.[14] enhanced the capability of the YOLOv5 network to identify cracks using Crack Conv and Adapt-weight Down Sample. Zhao al.[15] modified the backbone network and incorporated a cascade structure for the R-CNN part to improve the accuracy of crack identification. The conducted experiments demonstrate that although these algorithms improve single-network crack recognition accuracy, background noises often appear in the final recognition result.

To reduce the false alarm rate and enhance the efficiency of crack identification, Yang al.[16] suggested utilizing the YOLO-SAMT network for localizing crack areas, followed by crack extraction using an enhanced k-means clustering algorithm. Xie al.[17] proposed employing the Faster R-CNN network initially to identify potential crack regions, followed by utilizing the U-Net network to extract the crack pixels. Tong al.[18] introduced a PSP-YOLO crack detection and segmentation algorithm based on YOLO V5 and PSPnet. Fan al.[19] developed an activation function called MeLU and employed a distinctive calculation method to improve the Mask-RCNN network's ability to locate and identify cracks. Yu al.[20] integrated YOLOv5 and U-Net3+ networks and significantly enhanced the crack identification speed.

Compared to a single network model, these five methods adopt a "first detect the crack regions and then segment the crack pixels" framework. This approach exhibits high computational efficiency as it only performs pixel segmentation within the crack candidate regions. However, due to cracks occupying only a few pixels and resulting in a low signal-to-noise ratio within the complex tunnel environment, this framework is prone to missed detections during the crack region detection stage. Moreover, the aforementioned methods solely concentrate on crack pixel segmentation, without effectively identifying the regions where cracks originate. Consequently, numerous cracks from certain regions, such as bolt-hole repairs, often appear in the segmentation results. Furthermore, crack identification results obtained in this method are not effectively utilized for evaluating tunnel performance. To mitigate

† The authors are with Key Laboratory of Structural Health Monitoring and Control, Shijiazhuang Tiedao University, Shijiazhuang, 050043, China.
†† The author is with Shenzhen University, Shenzhen, 518060, China.
††† The authors are with China Railway design corporation, Tianjin, 300251, China.
†††† The author is with Hebei Jinbo Elevator Intelligent Equipment Co.,Ltd., Shijiazhuang, 050043, China.
Corresponding author: Hongbin Xu(xuhongbin@semi.ac.cn)

these issues, the present study proposes a novel framework, as depicted in Figure 1. Unlike conventional crack recognition models, the proposed model adopts a "local segmentation first, followed by global fusion detection" framework.

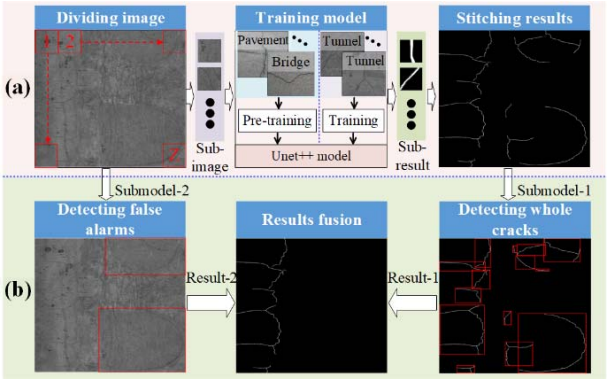

Figure. 1 Schematic diagram of the presented tunnel crack detecting framework. (a) The local segmentation module (b) The global fusion detection module

The main work of this paper involves the following: Firstly, a local segmentation module was employed to segment the original image. Subsequently, the recognition results were merged, ensuring the identification of all suspicious crack targets to the fullest extent and reducing the missed detection rate. Secondly, a two-detection framework was developed using the YOLO network. The first sub-model identifies regions corresponding to the same crack target in the segmentation results, while the second sub-model detects the bolt repair area. By merging these two recognition results, noise interference originating from bolt repairs is mitigated.

2. Proposed method

2.1 Local crack pixel segmentation model

In the presence of disturbances such as tunnel joints and water stains, manually labeling tunnel crack samples, which often have widths of only a few pixels, is challenging and time-consuming. Inspired by the concept of transfer learning, this article suggests constructing a non-tunnel crack dataset for the development of a tunnel crack segmentation model[21]. The details of this approach are outlined as follows:

Firstly, more than ten thousand non-tunnel crack images from pavement and bridge sources can be easily acquired by resizing and overlapping clipping[22]-[25]. These images are then set as dataset ***A***. Subsequently, the collected tunnel original images (each with a size of 4096×4096) are segmented without repetition, resulting in 100 tunnel crack images of size 512×512, which are set as dataset ***B***. For more effective training of the tunnel crack segmentation model with ample labeled crack samples, this paper selected non-tunnel cracks from dataset ***A*** that bear a high similarity to the images in dataset ***B***. During this selection process, the image features of crack samples are extracted using a feature encoding network pre-trained on the ImageNet dataset[26] and evaluated using the cosine similarity equation as follows:

$$f(a,b)=\frac{a\bullet b}{\|a\|\bullet\|b\|} \tag{1}$$

where '***a***' and '***b***' denote the feature vectors of a non-tunnel crack image and a tunnel crack image, respectively. By employing Eq. (1), the similarities between a non-tunnel crack image and all images from dataset ***B*** can be calculated, and their average value is denoted as '***f1***'. If '***f1***' surpasses the specified threshold '***T1***' (where '***T1***' is set to 0.7), this non-tunnel crack is selected as part of dataset ***C***, which will be utilized for pre-training the tunnel crack segmentation model.

Secondly, utilizing the abundance of non-tunnel cracks, the Unet++ network[27] is employed to train the crack recognition model, and the resulting encoding layer parameters ***G*** are preserved for the tunnel crack segmentation model. The final training equation is as follows:

$$W=M\left(B,G\right)=M\left(B,T_P\left(C\right)\right) \tag{2}$$

Where, $\boldsymbol{T_p}$ denotes the pre-training with non-tunnel cracks to obtain the parameter ***G*** of the Unet++ coding layer. ***M*** represents the operation of efficiently acquiring the final parameters ***W*** of the tunnel crack identification network through synchronous fine-tuning, which utilizes the pre-trained parameters ***G*** and the tunnel crack dataset ***B***. By utilizing parameter ***W*** to divide the original image, the resulting block images are processed by the local tunnel crack pixel segmentation model in parallel. Additionally, these identification results are merged and utilized as the preliminary recognition results. Through this approach, all potential pixels belonging to cracks are detected, thereby significantly reducing missed detections.

2.2 Global fusion detection model

1) Whole crack region extraction

Throughout the crack pixel segmentation process, there is a chance that not all pixels associated with the same crack are accurately identified, resulting in sporadic appearances of recognition outcomes. Additionally, certain background pixels may be mistakenly categorized as cracks, consequently causing false alarms within the final crack recognition results. To ensure precise extraction of the entire crack, it is essential to conduct secondary identification of the combined crack recognition results. For clarity, this secondary identification process is labeled as sub-model 1, and its details are as follows.

Initially, by incorporating the local crack recognition outcomes, a portion of global crack

recognition results can be derived. Subsequently, through analysis of the visual attributes such as curvature and the extension direction of cracks in the global recognition findings, a set of crack-region labels can be readily obtained. Then, the YOLO-V5 network[28] is employed to construct sub-model 1. After inputting the newly merged crack recognition results, potential crack regions can be detected. Finally, to mitigate potential disturbances from backgrounds, the area of the detected crack region is computed. The paper removes detection results whose area falls below a defined threshold 'st'=25mm^2, resulting in accurate recognition of the entire crack.

2) Fusion of crack recognition results

To enhance the corrosion resistance of bolts, it is common to fill bolt holes with concrete structures. However, cracks often emerge from these areas, although they do not directly impact the tunnel structure. This paper employed the YOLO-V5 network to construct sub-model 2 to address these false alarms and identify bolt-hole regions. The details of this sub-model are as follows:

Initially, by studying the shape of bolt-hole repair regions, a set of region labels for the bolt-hole repair areas can be manually acquired. Subsequently, the YOLO-V5 network is employed for sub-model 2 as well. Upon inputting the original images, all potential regions (i.e., false-alarm regions) containing bolt-holes can be identified and labeled as $\boldsymbol{O}$. Then, the coordinate range of $\boldsymbol{O}$ is set as the screening condition, and the crack identification results of sub-model 1 are further processed. Specifically, we calculated the ratio η of pixel number of the same crack inside region $\boldsymbol{O}$ to that outside region $\boldsymbol{O}$, and analyzed it according to the following formula:

$$P = \begin{cases} 0 & \eta > 0.4 \\ 1 & \eta \le 0.4 \end{cases} \tag{3}$$

Where, if the η value is not greater than 0.4, the crack target will be retained; otherwise, the crack target is removed. According to the formula (3), the crack target near the bolt-hole region can be preserved to the maximum extent, thereby avoiding the missed detection of suspected crack targets.

3) False alarm elimination

While the fusion of results effectively eliminates noise interference from bolt repairs, strip noise persists in the fusion outcomes. To address this, sub-model 2 was employed to compute the average width of the recognition target in the fusion result and subsequently filter out recognition noise, outlined as follows:

Considering the challenges posed by the multi-scale characteristics of fracture targets, accurately determining their length values is often difficult. Given that real crack width typically spans only a few pixels, this paper computes the average width, denoted as $\boldsymbol{x}$, of the identified target in the fusion result using sub-model 2. An average width threshold '$\boldsymbol{x}_{max}$' =10 is then employed to filter out these noises, effectively reducing the false alarm rate of the recognition results.

$$\begin{cases} x \times y = s \\ 2 \times (x + y) = l \end{cases} \rightarrow \quad x \approx \frac{2 \times s}{l} \tag{4}$$

where $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{l}$, and $\boldsymbol{s}$ represent the width, length, perimeter, and area (in pixel values) of the recognized target, respectively. It is worth noting that the average width value is computed by deriving the perimeter and area values of the identified target. Recognition results exceeding the threshold '$\boldsymbol{x}_{max}$' are filtered out.

3. Experimental evaluation and analysis

3.1 Experimental data and platform

During image acquisition in the actual tunnel, 600 tunnel crack images are obtained with a resolution of 4k ×4k. Subsequently, this study randomly selects 100 of these images and, through division and manual labeling, acquires 1000 local crack-labeled samples sized 512× 512, constituting the training dataset $\boldsymbol{D}$. The remaining 500 images, with a resolution of 4k×4k, are designated as the test dataset $\boldsymbol{E}$. Calculations are conducted on a single NVIDIA TESLA V100 32G GPU using the PyTorch framework. The initial learning rate for model training is set to 0.005, with 2500 iterations. Additionally, a weight decay of 0.0001 and a momentum of 0.9 are employed. The optimizer used is Adam, and the network adopts the DiceBCELoss as its loss function. To evaluate the experimental outcomes accurately, we utilize four evaluation metrics: Precision (P), Recall (R), balanced F1 score (F1), and Intersection over Union (IoU).

3.2 Threshold setting

To identify the minimum cosine similarity value required between feature vectors of non-tunnel crack images and tunnel crack images, this study manually examined 100 images from dataset $\boldsymbol{A}$ that were similar to those in dataset $\boldsymbol{B}$, forming dataset $\boldsymbol{A1}$.

The average cosine similarity for each image in dataset $\boldsymbol{A1}$ is computed using all images in dataset $\boldsymbol{B}$. It was determined that the minimum average value was 0.7, which was established as the threshold '$\boldsymbol{T1}$'. Figure 2 illustrates the detection results for various tunnel and non-tunnel data.

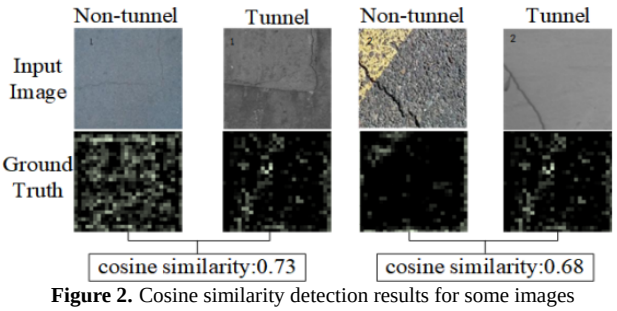

Figure 2. Cosine similarity detection results for some images

To mitigate point background noise from the recognition outcomes, this study manually tabulated the area of 100 cracked outer rectangular boxes. The ultimate threshold '***st***' = 25mm^2 was established by multiplying the minimum area of the outer rectangular box by 0.8. Figure 3 shows the detection results of the area of the external rectangular frame of cracks in real tunnels.

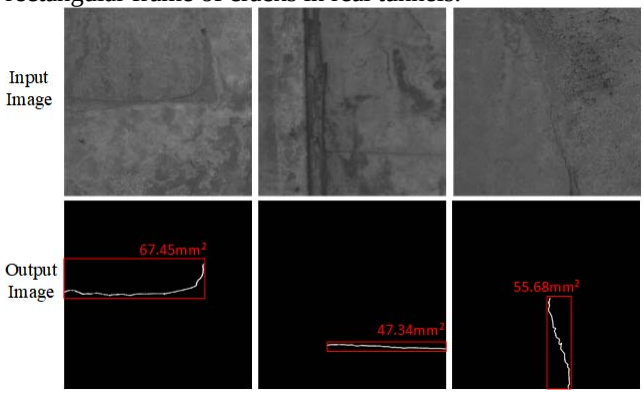

Figure 3. Detection results of the outer rectangular frame

To eliminate strip noise interference in the final recognition outcomes and enhance the model's detection precision. By referring to the relevant specifications[29], the average width of cracks in the tunnel is usually not more than 10mm. Therefore, the target average width parameter x_{max} is set as follows:

$$x_{max} = 10/r \quad (5)$$

Where, the value of parameter r is the target imaging scale factor, which is set by manual calibration before the tunnel image acquisition starts. Finally, the average width threshold 'x_{max}' is set to 10 pixels. Recognition results exceeding the threshold 'x_{max}' are filtered out. Figure 4 displays the results of average width detection for cracks in different tunnels.

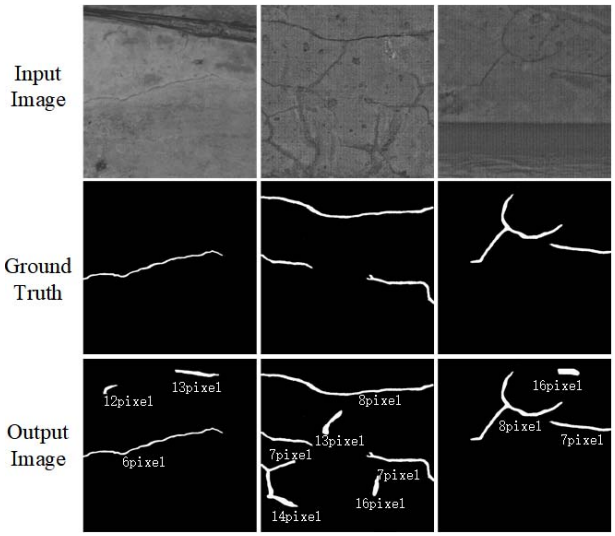

Figure 4. Detection results of average crack width in a tunnel

3.3Experiments results and discussion

1）Ablation study

Ablation experiments were conducted on the same tunnel crack image dataset using four models: the NT model, which directly trains the Unet++ network; the LT model, which utilizes transfer learning for local crack pixel segmentation; the Y-U model, which follows a "first localization and then segmentation" approach; and the U-Y model, which adopts a "first segmentation and then localization" strategy. These experiments were designed to evaluate the effectiveness of the U-Y model proposed in this paper. The experimental results are presented in Table 1.

Table 1 Ablation experiment results

Method	P(%)	R(%)	F1(%)	IoU(%)
NT	80.54	40.96	54.36	31.69
LT	81.13	41.23	54.96	40.87
Y-U	82.64	43.06	56.71	45.35
U-Y	84.06	45.63	59.26	49.13

Table 1 indicates that the model proposed in this paper outperforms other models in terms of precision (P), recall (R), balanced F1 score (F1), and Intersection over Union (IoU). Compared to the NT model, the precision of the LT model, which utilizes transfer learning, shows an improvement of 0.59%. Upon introducing the two sub-models proposed in this paper, the precision of the U-Y model demonstrates an enhancement of 2.93% compared to the LT model and a superiority of 1.42% over the Y-U model. The results of the ablation experiments validate the effectiveness of the proposed U-Y model, achieving an overall precision 3.52% higher than that of the NT model.

2）Performance comparison

Quantitative comparisons

To thoroughly assess the effectiveness of the U-Y model proposed in this paper, comparative experiments were conducted with four classical network models under identical experimental conditions. The corresponding results of these experiments are summarized in Table 2.

Table 2 Evaluation metrics of different models

Method	P(%)	R(%)	F1(%)	IoU(%)
Unet3+	80.25	40.62	53.62	33.87
DeeplabV3+	81.35	41.51	55.02	35.47
Mask-RCNN	82.13	42.03	55.68	41.59
Y-U	82.64	43.06	56.71	45.35
U-Y	**84.06**	**45.63**	**59.26**	**49.13**

The experimental results demonstrated that the performance of Y-U method is slightly better than that of Mask-RCNN. One possible reason is that YOLOv5 detection model used in Y-U method is superior to Faster-RCNN detection model used in Mask-RCNN method.

By contrast, the proposed U-Y model firstly recognizes the crack pixels within the divided images, which can minimize the likelihood of crack missed detection; and then detects the whole crack object via

YOLOv5 technique, that ensures the integrality of recognition results for the same crack. Table 2 indicates that the proposed model outperforms the other four algorithms across all four metrics. Specifically, its precision surpasses that of the Unet3+ network[30] by 3.81%, the DeeplabV3++ network [31]by 2.71%, the Mask-RCNN network[19] by 1.93% and the Y-U detection model by 1.42%. Furthermore, the F1 score and IoU value achieved by the proposed model notably exceeds those of the other four models. This observation demonstrates the U-Y model's capability to accurately identify crack targets while effectively reducing false detection rates compared to both the single network model and the "location before segmentation" model

Qualitative evaluation

To provide a clearer demonstration of the detection efficacy of each algorithm, several representative crack recognition results were selected, and the results are illustrated in Figure 5.

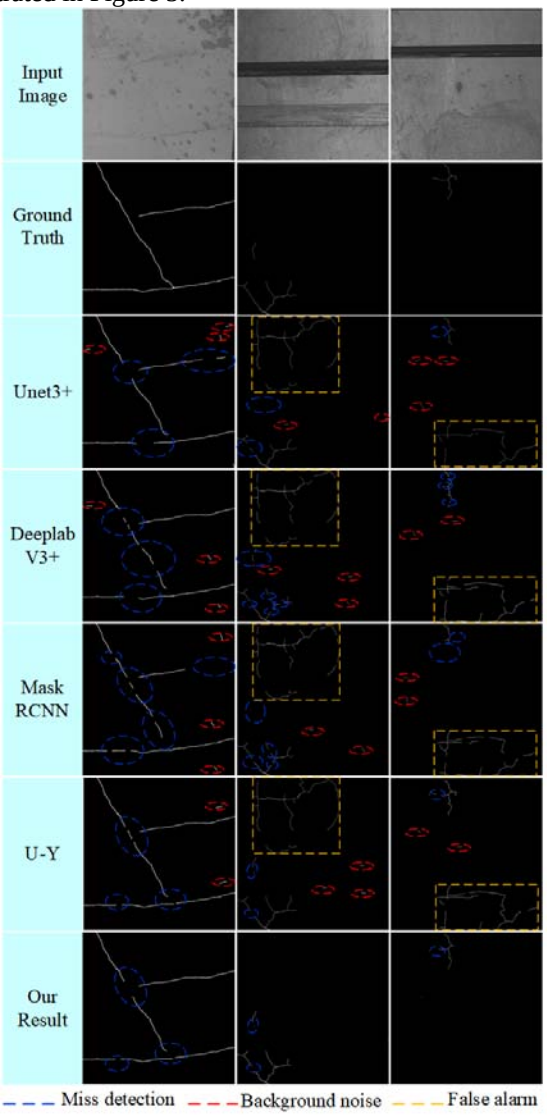

Figure 5. Some representative crack damage detecting results

As shown in the figure, the red and blue ellipses boxes indicate background noise and miss detection, respectively; The yellow rectangle boxes indicates false alarm.

In the third and fourth lines of the figure, the U-Y model, which adopts a "first segmentation and then localization" framework, shows superior recognition completeness of crack targets (see the blue oval box in Figure 5) compared to the Mask-RCNN model.

In rows 3-7 of the figure, compared with the Our Result, other models perform direct identification cracks, result in more background noise in the recognition results (see the red oval box in Figure 5). As demonstrated in row 7 of the figure, our approach efficiently suppresses background noise, by utilizing the global fusion detection module.

In the third and fourth columns of the figure, compared to other models that only identify crack targets, our approach uses Sub-model 2 to distinguish the crack generation region. Through the fusion of the crack recognition module, false alarm from bolt repairs regain is efficiently eliminated, as shown in the yellow rectangular box.

4. Discussion

Compared with the existing crack visual detection and identification methods [10-20], the proposed method has four main advantages:

(1) Improvement of training performance of crack identification model

A mass of labeled samples is essential for training a crack recognition model. In literatures [10-12], numerous labeled crack samples were required through manual labeling, which ensured the training performance of crack identification model. However, labeling crack samples is a time-consuming and labor-intensive task, and thus, it is difficult to build a crack identification model to some extent. Currently, transfer learning offers a promising solution for model training with small samples[32] and also becomes an optimal scheme for building crack identification model. Compared with the existing transfer learning methods, this work proposes a fast and effective calculation method (i.e., formula (1)), which can be used to select data from a number of non-tunnel crack samples. The selected similar crack samples can contribute to the rapid construction of tunnel crack identification model.

(2) Reducing the failure rate of crack identification

In order to enhance the efficiency of crack identification model, literatures [16-20] have suggested first locating the crack regions and then recognizing the crack pixels within the detected crack regions. The width of cracks is only a few pixels, which results in the low signal-to-noise ratio of cracks in complex tunnel environments. With this condition above, compared with the block-shaped object recognition, it is very difficult to obtain the favorable performance for recognizing the crack regions. Therefore, the existing methods [16-20] are prone to missed detection in the crack region detection step. In

contrast, the proposed model utilizes a "first segmentation and then localization" framework in sub-model 1. This framework not only minimizes the likelihood of crack missed detection, but also ensures the integrality of recognition results for the same crack.

(3) Identifying the location of cracks occurring

Currently, the most crack recognition methods based on deep learning [10-12] only regard the crack recognition as a simple semantic segmentation problem. Via adjusting the network structures or model parameters, the ability of these methods can be improved. However, these strategies only represent a basic application of deep learning in crack identification scenarios. In this paper, according to the actual scene requirements for tunnel crack disease identification, the crack identification results can be categorized into two types: The first one refers to cracks generated by the bolt-hole repair regions, while the other one refers to cracks in other areas of tunnel lining. To better identify the location of tunnel lining cracks, we proposed a detection model via combining two different networks, which can achieve more effective identification of tunnel crack diseases.

(4) Improving the accuracy of crack identification results

Currently, in the complex environmental interference, there still exist some recognizing noises for most of crack identification methods [16-20]. To address this issue, based upon the segmentation of crack pixels, we present a new crack-width estimation method (see formula (4)) via the analysis of crack stripe attributes. With the suitable threshold parameters, we can effectively filter the crack identification results, thereby significantly improving the accuracy of final crack recognition results.

To sum up, although that some conventional techniques (such as transfer learning, sub-model combination, etc.) are employed in our work, each technique has been organically combined into the developed crack identification model, thereby resulting in a new and practical detecting and recognizing framework for tunnel cracks.

5. Conclusion

This paper introduces a novel framework for crack recognition, which initiates with local crack pixel segmentation followed by global fusion detection. Leveraging a large set of labeled non-tunnel crack samples, a tunnel crack local-segmentation model is constructed using transfer learning, a distinguishing feature of the approach. Unlike existing models, the framework prioritizes crack segmentation on local image blocks, ensuring the retention of potential crack pixels. Subsequently, with the merged local crack recognition results, two sub-models are proposed to precisely localize whole crack targets while minimizing background false alarms. Experimental findings on actual tunnel crack data confirm the effectiveness of the framework in addressing challenges posed by complex tunnel environments. Comparative analyses on real tunnel crack data reveal superior performance of the framework over two single network models, with notable precision improvements of 3.81% and 2.71%, respectively, and a 1.93% accuracy increase compared to the "first location then segmentation" approach. Furthermore, the framework achieves higher F1 scores and IoU values compared to other algorithms, underscoring its efficacy in tunnel crack disease recognition.

6. Acknowledgments

This paper is supported by National Key R&D Program of China (2022YFB2603300, 2022YFB2603303), National Natural Science Foundation of China (51808358, 52178293), Laboratory Basic Research Project of China Railway (L2021G013), Shenzhen Science and Technology Program (KQTD20180412181337494), Natural Sciences Innovative Research Group project of Hebei Province (E2021210099), S&T Program of Hebei (21567625H).

References

[1] H. Liu, X. Du, Z. Xu, W. Li, and X. Shi, "Tunnel crack detection method using lightweight high-resolution features," Comput. Techno. Autom., 42, (2), PP. 144-150, 2023.

[2] C. Sun, Y. Chu, Y. Fan, and L. Dang, "Pavement crack image processing system research on VC++," Comp. Appl. Softw., 26, (8), pp. 82-85, 2009.

[3] S. Zhang, and J. Liu, "Edge detection operators and their application in cracking image,"Concr., (6), pp.25-27, 2010.

[4] Q. Li, and Q. Hu, "A pavement crack image analysis approach based on automatic image dodging," J. Highw. Trans. Res. Dev., 27, (4), pp. 1-5, 2010.

[5] F. Zhao, W. Zhou, Y. Chen, and H. Peng, "Application of improved Canny operator in crack detection," Electron. Meas. Technical., 41, (20), pp. 107-111, 2018.

[6] H. Fang, and N. He, "Detection Method of Cracks in Expressway Asphalt Pavement Based on Digital Image Processing Technology,"Appl. Sci.,13, (22),2023.

[7] G. Yang, P. Geng, and H. Ma, "DWTA-Unet: Concrete Crack Segmentation Based on Discrete Wavelet Transform and Unet," Proc. 2021 Chin. Intell. Autom. Conf., pp. 702-710, 2022.

[8] P. Geng, Z. Tan, J. Luo, T. Wang, F. Li, and J. Bei, " ACPA-Net: Atrous Channel Pyramid Attention Network for Segmentation of Leakage in Rail Tunnel Linings,"Electron., 2, (12), pp. 255, 2023.

[9] P. Geng, J. Lu, G. Yang, and H. Ma, "Crack Segmentation Based on Fusing Multi-Scale Wavelet and Spatial-Channel Attention," SDHM., 17, (1), 2023.

[10] B. Wang, S. Bai, and W. Zhao, "Pavement crack damage visual detection method based on feature reinforcement learning," J. Railw. Sci. Eng., 19, (7), pp. 1927-1935, 2022.

[11] J. Wu, and X.Zhang, "Tunnel Crack Detection Method and Crack Image Processing Algorithm Based on Improved Retinex and Deep Learning," Sens., 23, (22), pp. 9140, 2023.

[12] B. Wang, R. Wang, M. Chen, Y. Pan, and L.Wang, "Automatic Detection Method for Surface Diseases of Shield Tunnel Based on Deep Learning," J. Shang. Jiaotong Univ., pp. 1-16, 2023.

[13] M. Sohaib, S. Jamil, and J. Kim, "An Ensemble Approach for Robust Automated Crack Detection and Segmentation in Concrete Structures," Sens., 24, (1), pp. 257, 2024.

[14] Y. Li, S. Sun, W. Song, J. Zhang, and Q. Teng, "CrackYOLO: Rural Pavement Distress Detection Model with Complex

Scenarios," Electron. 13, (2), pp. 312, 2024.
[15] M. Zhao, X. Xu, X. Bao, X. Chen, and H. Yang, "An Automated Instance Segmentation Method for Crack Detection Integrated with CrackMover Data Augmentation," Sens., 24, (2), pp. 446, 2024.
[16] Z. Yang, C. Ni, L. Li, W. Luo, and Y. Qin, "Three-stage pavement crack localization and segmentation algorithm based on digital image processing and deep learning techniques," Spec. Issue Image/Signal Process. Mach. Vis. Sens. Appl., 22, (21), pp. 8459, 2022.
[17] X. Xie, H. Wang, B.Zhou, J. Cai, and F. Peng, "Research on Fast Identification and Segmentation Algorithms for Cracks of Highway Tunnels in Complex Environment," Chin. J. Undergr. Space Eng., 18, (6), pp. 1025-1033, 2022.
[18] Z. Tong, B. Lei, L. Jiang, and N. Wangu, "Fusion segmentation method for pavement crack detection," NDT., 45, (1), pp. 1-7, 2023.
[19] L. Fan, and J. Zou, "A Novel Road Crack Detection Technology Based on Deep Dictionary Learning and Encoding NetworksAppl. Sci., 13, (22), 2023.
[20] J. Yu, B. Liu, D. Yin, W. Gao, and Y. Xie, "Intelligent identification and measurement of bridge cracks based on YOLOv5 and U-Net3+," J. Hunan Univ. (Nat. Sci.)., 50, (5), pp. 65-73, 2023.
[21] W. Zhang, and T. Zhang, "Research on convolutional neural network image recognition method based on transfer learning," Mod. Inf. Technol., 7, (14), pp. 57-60, 2023.
[22] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "CrackTree: Automatic crack detection from pavement images," Pattern Recog. Lett., 33, (3), pp. 227-238, 2012.
[23] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," IEEE Trans. Intell. Transp. Syst., 17, (12), pp. 3434-34, 2016.
[24] M. Eisenbach, R. Stricker, D. Seichter, K. Amende, K.Debes, M.Sesselmann, D. Ebersbach, U.Stoeckert, and H. Gross , "How to get pavement distress detection ready for deep learning? A systematic approach," IJCNN, pp.2039-2047,2017.
[25] F. Yang, L. Zhang, S. Yu, P. Danil, X. Mei, and H. Ling, "Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection," IEEE Trans. Intell. Transp. Syst., 21, (4), pp. 1525-1535, 2019.
[26] Y. Ge, S. Jiang, F. Ye, Q. Yu, and Q. Tang, "Remote Sensing Image Retrieval Using Pre-trained Convolutional Neural Networks Based on ImageNet," Geom. Info. Sci. Wuhan Univ., 43, (1), pp. 67-73, 2018.
[27] H. Zhang, B. Peng, and W. Xu, "Road crack detection based on Unet++ and conditional generative adversarial network," J. Comput. Appl., 40, (s2), pp. 158-161, 2020.
[28] C. Bian, W. Hao, and W. Ma, "YOLOv5 building crack detection method based on Xception and SA," Comput. Technol. Dev., 33, (8), pp. 159-164, 2023.
[29] SAC/TC 529, "Specifications for operation monitoring of urban rail transit facilities - Part 3: Tunnel," GB/T 39559.3-2020, 2020.
[30] Q. Xv, Q. Liang, T. Zhao, W.Zou, and Y. Pan , "Segmentation Method of Concrete Small Cracks Based on UAV Images," Recent Adv. Comput. Sci. Commun., 17, (5), 2024.
[31] X. Wang, T. Wang, and J. Li, "Advanced crack detection and quantification strategy based on CLAHE enhanced DeepLabv3+," Eng. Appl. Artif. Intell., 126, (PB), 2023.
[32] V. Manjunatha, S. Ramalingam, T. K. Marks, and L. Davis, "Class Subset Selection for Transfer Learning using Submodularity," arXiv preprint arXiv:1804.00060, 2018.

Baoxian Wang received the Ph.D. degree from Beijing Institute of Technology, Beijing, China, in 2016. He is currently an Associate Professor with Shijiazhuang Tiedao University. His current research interests include image processing, computer vision, machine learning, and pattern recognition.

Ze Gao received the B.Eng. degree from Tangshan University, Tangshan, China, in 2020. He is currently pursuing the M.E degree in the School of Electrical and Electronic Engineering, Shijiazhuang Tiedao University, Shijiazhuang, China. His current research interests include image processing, computer vision, and pattern recognition.

Hongbin Xu received the Ph.D. degrees in civil engineering at Beijing Jiaotong University, Beijing, China, in 2021. Since 2021, he has been a postdoctoral fellow in Shenzhen University, and since 2022, he has been an associate research fellow in this school. His current research interests include in structural health monitoring, and signal processing application in railway engineering.

Shoupeng Qin is currently an engineer with the China Railway design corporation. His current research interest is the image processing application in tunnel engineering.

Zhao Tan is currently a senior engineer with the China Railway design corporation. His current research interest is the image processing application in tunnel engineering.

Xuchao Shi is Chairman of Hebei Jinbo Elevator Intelligent Equipment Co., LTD. His current research interest is the intelligent signal processing and control in elevator equipment.