Automatic Sensing and Detection for Subway Tunnel Pathologies

Xingyu Wang*, Zhengkun Zhu

School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 102616, China

*Corresponding author: Xingyu Wang, k1941172018@163.com

Abstract: Subway tunnels often suffer from surface pathologies such as cracks, corrosion, fractures, peeling, water and sand infiltration, and sudden hazards caused by foreign object intrusions. Installing a mobile visual pathology sensing system at the front end of operating trains is a critical measure to ensure subway safety. Taking leakage as the typical pathology, a tunnel pathology automatic visual detection method based on Deeplabv3+ (ASTPDS) was proposed to achieve automatic and high-precision detection and pixel-level morphology extraction of pathologies. Compared with similar methods, this approach showed significant advantages and achieved a detection accuracy of 93.12%, surpassing FCN and U-Net. Moreover, it also exceeded the recall rates for detecting leaks of FCN and U-Net by 8.33% and 8.19%, respectively.

Keywords: Visual detection; Neural network; Health monitoring; Image segmentation; Water leakage; Subway tunnel.

1. Introduction

There are several types of tunnel defects, including segment leakage, cracks, lining misalignment, track bed hollowing, concrete deterioration, and cross-section ovalization. In addition, intrusion boundary phenomena caused by construction, support structure deformation, surrounding rock deformation, etc. also occur from time to time. At the same time, sudden hazards such as track damage and falling interference objects may also occur (**Figure 1**). If these defects are not dealt with in time, they will have a serious impact on society and the economy [1]. Among them, water leakage is one of the most common defects in subway lines [2]. The running interval of subway trains is very short. The average interval of the Beijing subway is 3–5 minutes, and it may even be shortened to 2 minutes during peak hours. Therefore, timely and fast on-board inspection of tunnels is particularly important.

The fully automatic tunnel defect sensing system enables automatic sensing of subway tunnel defects. The system consists of two subsystems, namely the mobile visual defect sensing system and the processing feedback system (**Figure 2**).
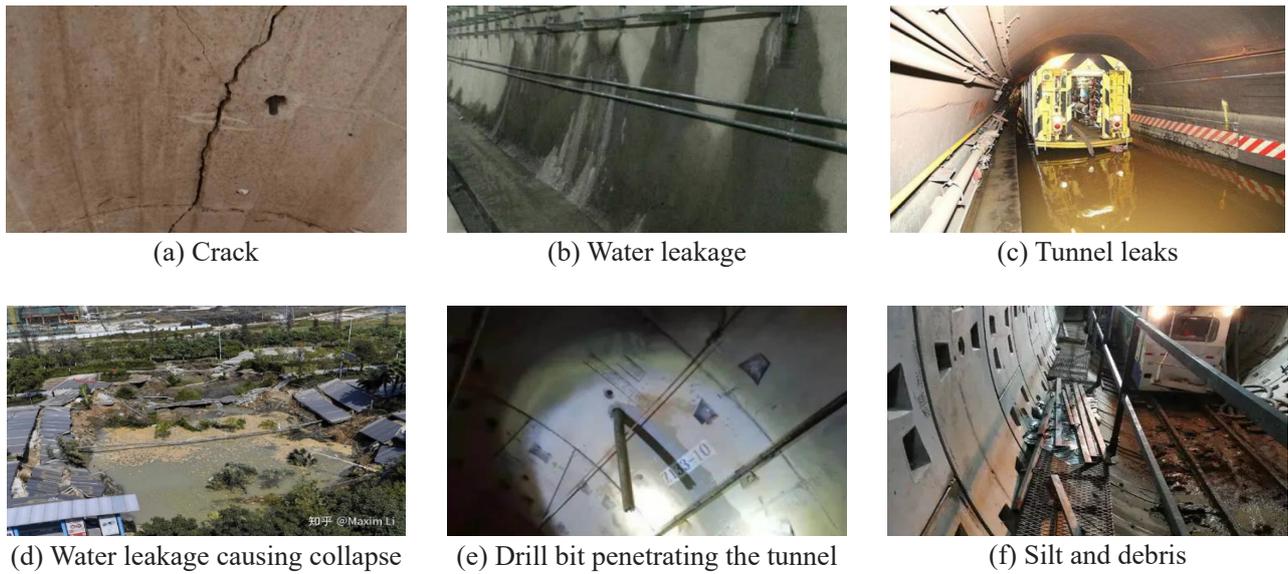
| (a) Crack | (b) Water leakage | (c) Tunnel leaks |

| (d) Water leakage causing collapse | (e) Drill bit penetrating the tunnel | (f) Silt and debris |

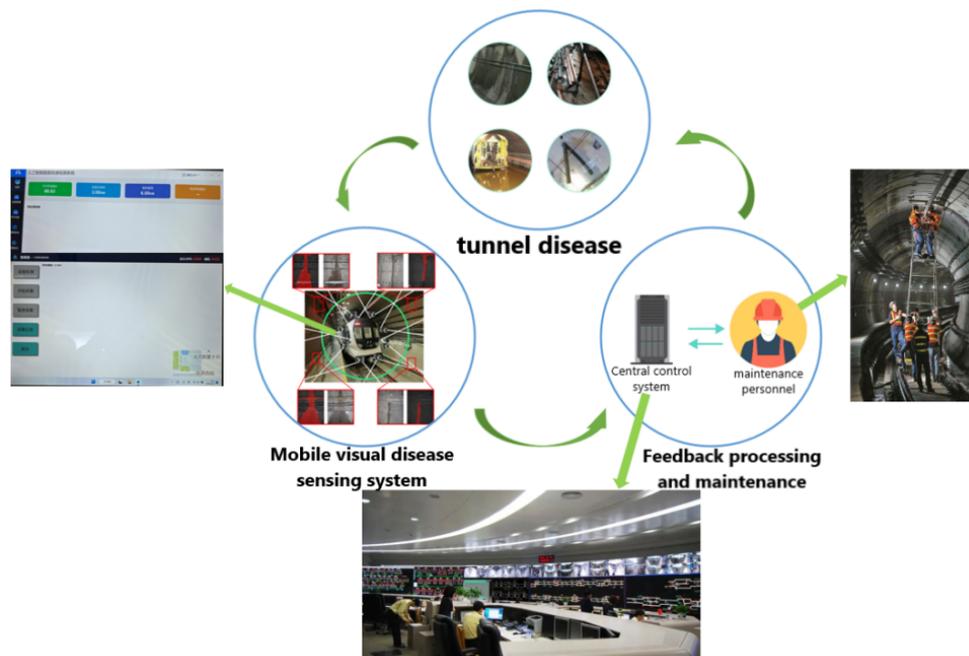**Figure 1.** Common defects of subway tunnel

Figure 2. Fully automatic tunnel disease visual perception system

In the mobile visual disease sensing system, 8 high-resolution industrial charge-coupled device (CCD) cameras are installed on the top, left and right sides, and bottom of the subway train, and the sensing system is composed of LED lighting equipment, power supplies, and lifting platforms. By taking a circular shot of the tunnel, image data of the entire tunnel is obtained [3]. The image processing technology is used to analyze and detect images to detect various diseases in subway tunnels promptly [4]. Then, multiple positioning methods are integrated to obtain disease location information in real time. The mobile defect sensing system is connected to the train operation central control system, so the location of the defect can be detected based on the train mileage and the defects detected. Second, the inertial navigation recorder is used to record the train acceleration and angular velocity, and the position of the defect will be calculated based on these two information. By

integrating the information obtained by the two aforementioned calculations, the precise location of the defect in the tunnel can be obtained and sent to the processing feedback system.

In the processing feedback system, the central control system generates and distributes maintenance tasks based on the information provided by the mobile visual defect sensing system. Maintenance personnel can then be allocated to repair the subway tunnel defects based on the type and location of the defect to ensure operational security and stability. In this process, it is important to accurately categorize the type of disease to facilitate repair work.

In recent years, the advancement of computer technology has brought rapid development to visual measurement. Machine learning methods can be used to segment the area of water leakage more accurately. Xue *et al.* [5] adopted fully convolutional (FCN) and Dai *et al.* [6] local-based FCN (R-FCN) to label, classify and detect cracks and leaking water. Huang *et al.* [7] used FCN to perform semantic segmentation of leaking water. However, small leaks may be overlooked due to their spatial invariance characteristics [8]. Hu *et al.* [9] improved the MSRCR algorithm to enhance crack images and combined the crack coordinate design algorithm and the SVM support vector machine method to classify crack types. Tian *et al.* [10] used the target detection algorithm MaskR-CNN to automatically detect water leakage in two-dimensional images, thereby producing a high-quality segmentation module. Jiang *et al.* [11] proposed a subway tunnel crack target detection scheme based on deep learning, using a deep convolutional generative adversarial network to achieve sample expansion, and using YOLOv5 and digital image processing technology to achieve intelligent identification of subway tunnel cracks. However, the network structure does not take into account computational efficiency. Orimasa *et al.* [12] used the improved SSD fully convolutional network structure to complete the intelligent identification and category labeling of targets such as cracks. Zhao *et al.* [13] proposed using the MaskR-CNN network for water leakage detection and verified that this model can achieve precise positioning and pixel segmentation of water leakage. Zhu *et al.* [14] proposed an improved YOLOv5 target detection model. According to the position and scale characteristics of the target, the original model was optimized around the detection accuracy, model parameters, and detection speed. However, due to the complexity of the environment, the detection accuracy is still not guaranteed.

There is still room for optimization in the automatic detection of subway tunnel defects in terms of coping with complex environmental challenges and detection accuracy. We proposed a fully automatic tunnel disease perception system composed of a mobile visual disease perception system and a processing feedback system called Automated Subway Tunnel Pathology Detection System (ASTPDS). This system was created based on Deeplabv3+, which is used for the semantic segmentation of tunnel diseases. The adoption of StepLR and CosLR learning rate decay algorithms makes the model more stable, converges faster, and has better generalization performance. The Adam optimizer was used to improve model training speed, performance, and generalization capabilities. The focal loss function was used to overcome positive and negative sample imbalance. This system could accurately segment leaking water areas in various complex environments and had excellent segmentation accuracy.

2. Automatic segmentation method of tunnel diseases

ASTPDS consists of two modules, in which the encoding area is used to extract high-level semantic information; and the decoding area is used to extract low-level semantic information, as shown in **Figure 3**.
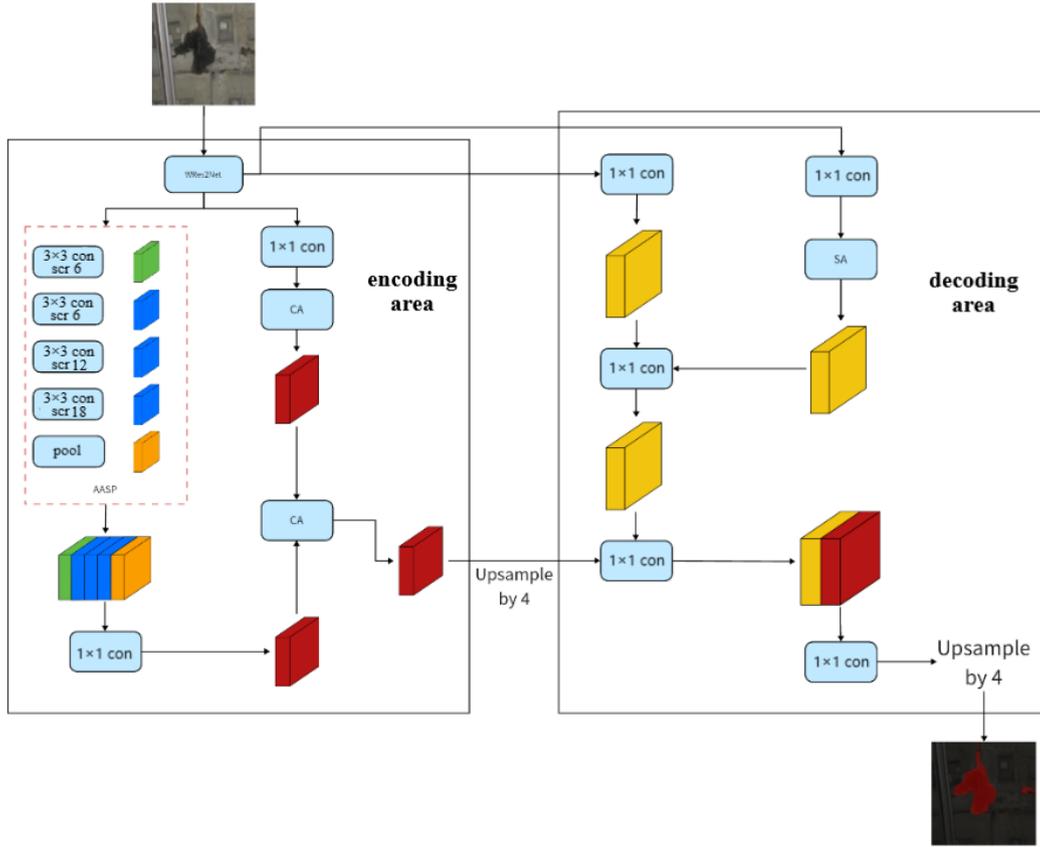
Figure 3. ASTPDS network structure diagram

The backbone network of ASTPDS consists of the WRes2Net network, which is used to extract leakage water feature information. The channel attention mechanism (CA) is used in the coding area to extract the depth of the channel to separate the convolutional layer features and achieve accurate allocation of high-response channels. Assuming that the high-level semantic information is $\psi_c^h \in I^{w \times h \times c}$, $\psi_c^h = [\psi_1, \psi_2 \ldots, \psi_c]$, among them, are the width and height of the input feature image respectively, and is the number of channels.

$$f_{CA}(g_c, \Phi_c) = s_1\{f_{c2}\{r[f_{c1}(g_c, \Phi_{c1})], \Phi_{c2}\} \tag{1}$$

In Formula (1), g_c is the feature map after ψ_c^h average pooling, Φ_c is the parameter in the channel attention module, s_1 is the sigmoid activation function, f_C represents the fully connected layer, and r is the ReLU activation function. The CA module outputs f_{CA} and weights the ASSP feature map to obtain the output feature map $\psi_c^{h\prime}$, as shown in Formula (2).

$$\psi_c^{h\prime} = \psi_c^h \cdot f_{CA} \tag{2}$$

Compared with the encoding area, the decoding area can obtain features such as the location and edge of the target information, but there is a large amount of background information, which will affect the accuracy of segmentation to a certain extent. The spatial attention mechanism (SA) is used in the decoding area to focus on the target feature area, adaptively combine high-level features with low-level features, and use high-level features to filter out background information. In this way, global information can be obtained without increasing the parameters and semantic segmentation can be improved through global convolutional networks. Besides, a two-layer convolution operation is used. The two-layer convolution kernels are $1 \times k$ and $k \times 1$ respectively, which are used to obtain key feature information.

$$A_1 = Conv_1[Conv_2(\psi_c^h, \Phi_{s1}), \Phi_{s2}] \tag{3}$$

$$A_2 = Conv_2[Conv_1(\psi_c^h, \Phi_{s1}), \Phi_{s2}] \tag{4}$$

$$f_{CA}(\psi_c^{h'}, \Phi_{s1}) = S_2(A_1 + A_2) \tag{5}$$

In Formulas (3)–(5), A_1 is the feature map completed by convolution of convolution kernels 1×5 and 5×1, A_2 is the feature map completed by the convolution of convolution kernels 5×1 and 1×5, $Conv_1$ is $5 \times 1 \times C$ convolution kernel, $Conv_2$ is the convolution kernel of $1 \times 5 \times C$, S_2 is sigmoid activation, and $\psi_c^{h'}$ is obtained by SA weighting.

Training complex deep learning models may take a long time. Adding learning rate decay methods and optimizers can increase the training efficiency of the model and achieve better training results. The learning rate decay method STEPLR is used during the training process to effectively improve the model convergence speed, prevent over-fitting, and improve model performance.

$$P = L_0 \times g^{f\frac{e}{s}} \tag{6}$$

In Formula (6), P represents the current learning rate, L_0 is the initial learning rate, g is the attenuation coefficient, e represents the current number of training rounds, and s is the size of the attenuation step. After each s training rround the learning rate will attenuate according to the attenuation coefficient g. The learning rate decay method CosLR is used to improve the convergence speed and generalization ability of the model.

$$P = 0.5 \times L_0 \times \left(1 + cos\left(\frac{e}{T}\right)\right) \tag{7}$$

In Formula (7), P represents the current learning rate, L_0 is the initial learning rate, e represents the current number of training rounds, and T is the total number of rounds of learning rate decay. At the same time, the optimizer SGD is used to update the model parameters to minimize the loss function and improve the accuracy and generalization ability of the model (Formula [8]).

$$w = w - l \times g \tag{8}$$

In Formula (8), w represents the weight parameter of the model, l represents the learning rate, and g represents the gradient of the loss function on the weight parameter. The optimizer Adam is used to adaptively update the model parameters to improve the convergence speed and generalization ability of the model.

$$m = b_1 \times m + (1 - b_1) \times g \tag{9}$$

$$v = b_2 \times v + (1 - b_2) \times g^2 \tag{10}$$

$$\tag{11}$$

$$w = \frac{w - p \times m}{(sqrt(v) + ep)}$$

In Formulas (9)–(11), w represents the weight parameter of the model, p represents the learning rate, g represents the gradient of the loss function on the weight parameter, m and v are mean estimate and uncentered variance estimate, respectively. $sqrt(v)$ represents the square root of v, while ep is a constant used to stabilize the calculation and prevent division by zero.

After testing with the same optimizer, the network performance using CosLR is better. Under the same learning decay rate, the network performance of Adam is generally better than SGD.

$$Loss = \begin{cases} -\kappa(1 - y')^\delta, y = 1 \\ -(1 - \kappa)y'^\delta \log(1 - y'), y = 0 \end{cases} \tag{12}$$

Additionally, training is divided into freezing and unfreezing phases. The focal loss function is used to

solve the imbalance problem of positive and negative samples, as shown in Formula (12).

3. Experiments and data analysis

3.1. Dataset construction

The experimental data set consisted of two parts. The first part was the lining water leakage data set of a subway tunnel in Shanghai, and the second part was the subway tunnel lining water leakage data set of a subway tunnel in Beijing collected by the mobile visual defect sensing system. The quality, shape, and dimensions of data were optimized for the training and application of deep learning models, thereby improving the accuracy and reliability of the model. The data were optimized through data source screening, data cleaning, data preprocessing, data partitioning, and data annotation.

Through data source screening and cleaning, abnormal and erroneous data were eliminated, and a total of 4,700 water leakage images were obtained. Data preprocessing included operations like statistical analysis of data, feature extraction, and normalization, in which the sizes of the images were standardized to 512×512 pixels. Next, the 4700-image data set was divided into a training data set and a test data set at a ratio of 4:1. Finally, manual annotation is used to mark the leakage area, which is regarded as the true value. The entire annotation task is completed using the LabelMe annotation tool.

3.2. Experimental program

Due to the complexity of the underground tunnel environment, water leakage exhibits characteristics such as irregular horizontal or vertical shapes, making it easy to go unnoticed, along with evidence of manual maintenance. This necessitates a network structure with excellent robustness to effectively address these challenges. Therefore, we divided the leakage data into five categories: blocky, vertical, horizontal, stain-covered, and occluded. FCN, U-Net, and ASTPDS were used for segmentation. The semantic segmentation effects were also compared to verify the advancement, accuracy, and robustness of ASTPDS.

$$P = \frac{TP}{TP+FP} \tag{13}$$

$$R = \frac{TP}{TP+FN} = \frac{(GT-FN)}{GT} \tag{14}$$

$$OAA = \frac{GT-FN+TN}{GT+TN+FP} \tag{15}$$

$$DSC = \frac{2TP}{FP+TP+GT} \tag{16}$$

In Formulas (13)–(16), *TP* refers to the number of pixels of the leakage water image that is correctly segmented, *FP* refers to the number of pixels of the leakage water image that is incorrectly segmented, *FN* refers to the number of unidentified leakage water image pixels in the segmentation result, and *TN* refers to the unidentified leakage water image pixels in the segmentation result. Number of pixels leaking water, *GT* refers to the actual number of pixels leaking water.

Four key indicators, including precision (*P*), recall rate (*R*), overall detection accuracy (*OAA*), and model segmentation performance (*DSC*), were used to evaluate the network model's effectiveness in identifying water leakage. Precision measures the accuracy of positive predictions, while recall rate assesses the model's ability to detect real instances of leakage water. Overall detection accuracy offers a comprehensive assessment,

and model segmentation performance evaluates the accuracy of delineating leakage areas. Analyzing these indicators provides valuable insights into the model's performance in detecting water leakage.

3.3. Experimental results and discussion

In order to evaluate the performance of ASTPDS, a comparative test with similar models was designed. A total of FCN, U-Net, DeeplabV3+, and other models were selected and evaluated from five technical indicators. The calculation results are shown in **Table 1**.

Table 1. Performance comparison of different algorithms

Network type	Accuracy, P (%)	Recall, R (%)	DSC (%)	OAA (%)	Time (h)
FCN	87.24	87.15	87.25	99.423	23.51
U-Net	87.56	87.29	88.17	99.641	22.62
ASTPDS	93.12	95.48	94.02	99.941	21.25

All evaluation indicators of ASTPDS were better than other models, with detection accuracy increased by 5.88% and 5.56% respectively. The recall rate increased by 8.33% and 8.19%, respectively. The DSN representing the model segmentation performance increased by 6.77% and 5.85% respectively. The results showed that ASTPDS was superior to other models in all aspects, and its accuracy was significantly higher. As shown in **Figure 4**, ASTPDS performed significantly better than other similar models in terms of ModelLoss index. In addition, as shown in **Figure 5**, after the DSC coefficient stabilizes, its segmentation performance is also significantly better than that of similar models.
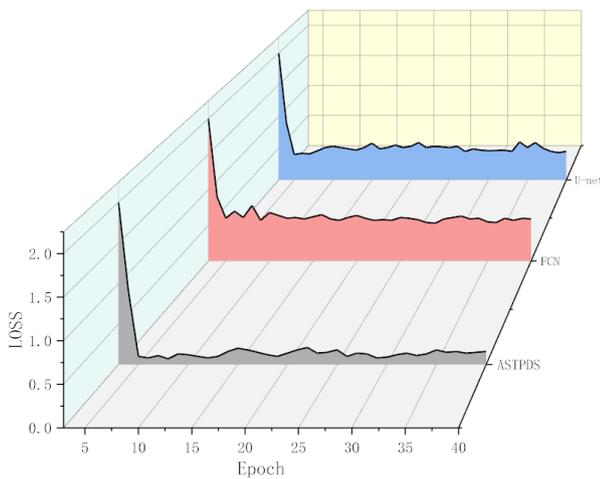

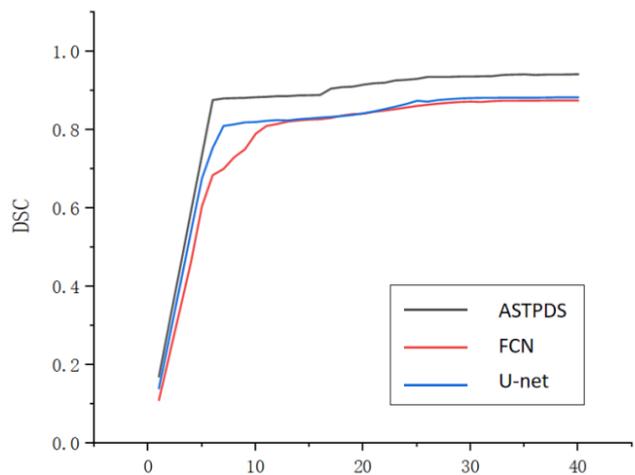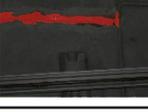**Figure 4.** Model losses for different models

Figure 5. DSC comparison of different models

The categorization of water leakage by different systems is shown in **Table 2**. It can be seen from this that compared with block leakage water, U-Net has a better segmentation effect and can identify small leakage water blocks, while ASTPDS can segment small leakage more completely. For vertical water leakage, FCN segmentation is incomplete and U-Net suffers from overfitting. For horizontal strip water leakage, FCN has the problem of over-segmentation. For leaking water covered by stain, redundant background information appeared in the U-Net segmentation area, and the outline of the segmentation area was lacking or overfitting, while the segmentation by ASTPDS was finer and more complete. For occluded leakage, when the occluded part was

unknown, all three models showed good results, but FCN still had the problem of incomplete segmentation.

Table 2. Comparison of semantic segmentation methods for different types of leakage water

4. Conclusion

Subway tunnels are often characterized by poor and uneven illumination, coupled with high levels of noise in images captured onboard trains. As a result, the accuracy of detecting leaking water, a common issue in subway tunnels, tends to be low. Addressing the challenges posed by this complex environment, this paper introduces a novel vehicle-mounted pixel-level visual sensing detection method. This method enables fully automatic and highly accurate identification and segmentation of tunnel leakage water.

Through extensive experimental testing, the proposed water leakage detection model (ASTPDS) showcased superior performance compared to other similar models. Specifically, the detection accuracy improved by 6.74%, the recall rate increased by 9.56%, and the segmentation performance saw a notable enhancement of 7.76%.

Disclosure statement

The authors declare no conflict of interest.

References

[1]   Hong J, 2023, Overview of the Application of Visual Inspection Technology for Surface Defects of Subway Tunnel. Railway Survey, 49(01): 18–22 + 32. https://www.doi.org/10.19630/j.cnki.tdkc.202201140001

[2]   Tian Y, Fan T, Tang C, 2022, Rapid Detection Technology for Surface Leakage of Subway Tunnels. Bulletin of Surveying and Mapping, 2022(09): 29–33. https://www.doi.org/10.13474/j.cnki.11-2246.2022.0259.

[3]   Xue Y, Li Y, 2018, A Method of Disease Recognition for Shield Tunnel Lining Based on Deep Learning. Journal of Hunan University. 45(3): 100–109.

[4]   Wang X, Wu Y, Cui J, et al., 2020, Shape Characteristics of Coral Sand from the South China Sea. J. Mar. Sci. Eng., 8(10): 803

[5]   Shen JH, Wang X, Liu WB, et al., 2020, Experimental Study on Mesoscopic Shear Behavior of Calcareous Sand

Material with Digital Imaging Approach. Adv. Civ. Eng., 2020: 8881264.

[6] Ren YP, Huang JS, Hong ZY, et al., 2020, Image-Based Concrete Crack Detection in Tunnels Using Deep Fully Convolutional Networks. Constr. Build. Mater., 234: 117367–117379.

[7] Huang HW, Li QT, Zhang DM, 2018, Deep Learning-Based Image Recognition for Crack and Leakage Defects of Metro Shield Tunnel. Tunn. Undergr. Sp. Tech., 77: 166–176.

[8] Garcia-Garcia A, Orts-Escolano S, Oprea S, et al., Review on Deep Learning Techniques Applied to Semantic Segmentation.

[9] Hu M, 2021, Design and Implementation of Crack Detection System for Subway Tunnel Lining based on OpenCV, dissertation, Yangzhou University. https://www.doi.org/10.27441/d.cnki.gyzdu.2021.001467

[10] Tian Y, Fan T, Tang C, 2022, Rapid Detection Technology for Surface Leakage of Subway Tunnel. Surveying and Mapping Bulletin, 546(9): 29–33. https://www.doi.org/10.13474/j.cnki.11-2246.2022.0259

[11] Jiang S, 2022, Image Detection of Crack Disease in Subway Tunnel based on Deep Learning, dissertation, East China Normal University. https://www.doi.org/10.27149/d.cnki.ghdsu.2022.004606

[12] Bai B, 2019, Research on Image Recognition Algorithm of Complex Crack Disease in Subway Tunnel, dissertation, Beijing Jiaotong University.

[13] Zhao S, Zhang D, Huang H, 2020, Deep Learning-Based Image Instance Segmentation for Moisture Marks of Shield Tunnel Lining. Tunnelling and Underground Space Technology, 95: 103–109.

[14] Zhu J, Zheng A, Lei Z, et al., 2023, Detection Method for Apparent Defects of Subway Tunnel Ancillary Facilities and Lining Based on Improved YOLOv5. Journal of Railway Science and Engineering, 20(03): 1008–1019. https://www.doi.org/10.19713/j.cnki.43-1423/u.t20220712