Yuan et al. Heritage Science (2023) 11:65
https://doi.org/10.1186/s40494-023-00895-7

Heritage Science

RESEARCH

Open Access

Automatic recognition of craquelure and paint loss on polychrome paintings of the Palace Museum using improved U-Net

Quan Yuan†, Xiang He†, Xiangna Han* and Hong Guo

Abstract

Craquelure is the most common defect on ancient polychrome paintings, which may deteriorate further to paint loss. Previous image processing methods, which can accurately recognize paint loss, have limited precision and efficiency in segmenting craquelure. This paper proposes a semantic segmentation method, Res-UNet, for the recognition of craquelure and paint loss in the Palace Museum, Beijing. The residual structure of ResNet-50 enables the avoidance of network degradation, and image features can be fully extracted. Using the unique skip connection module of U-Net, features of different levels are fused to improve segmentation accuracy and provide smoother craquelure edges. Three loss functions are combined to accelerate stable convergence. The model was tested on a newly built dataset based on 600 images. Experimental results supported by statistical tests show that Res-UNet is a capable method of craquelure recognition, with an accuracy rate of 98.19%, and F1-score of 93.42%. Hence, the proposed hybrid approach is a promising tool to support the preservation and restoration of valuable traditional Chinese polychrome architectural paintings.

Keywords The Palace Museum, Architectural polychrome painting, Craquelure, Semantic segmentation

Introduction

The Palace Museum was the imperial palace of the Ming and Qing dynasties. It was the political center of ancient China and residence of emperors for more than five centuries. The architectural polychrome paintings inside it are of great value. Beyond their artistic importance, these paintings carry cultural and political significance such as distinguishing ranks and highlighting ethics, and are important to the study of ancient technology. However, the lifetime of outdoor paintings is limited. Under the influence of periodic variations of temperature and humidity, as well as sunlight, rain, and wind, the defects of craquelure, flaking, and paint loss occur in sequence, which can destroy a painting within a decade [1]. For example, about 60 years after they had been painted, a large portion of the paintings in the Dapaojing Well Pavilion fell off, and those paintings that remained were severely cracked or warped (Fig. 1).

Craquelure and paint loss are common phenomena not only at heritage sites but also on building materials, coatings, and soil. Today, the most common methods in the survey of defects of cultural relics are still visual identification, plotting, and area calculation, whose results are too rough for use in further study. Meanwhile, crack identification methods based on digital image processing have been intensively studied in other fields, and are already used to support in-depth research such as damage degree evaluation and cracking mechanisms. For example, Ai et al. [2, 3] identified the crack propagation of coal rock under impact, and

†Quan Yuan and Xiang He have contributed equally to this work and should be regarded as co-first authors

*Correspondence:
Xiangna Han
jayna422@ustb.edu.cn
Institute for Cultural Heritage and History of Science and Technology, University of Science and Technology Beijing, Beijing 100083, China

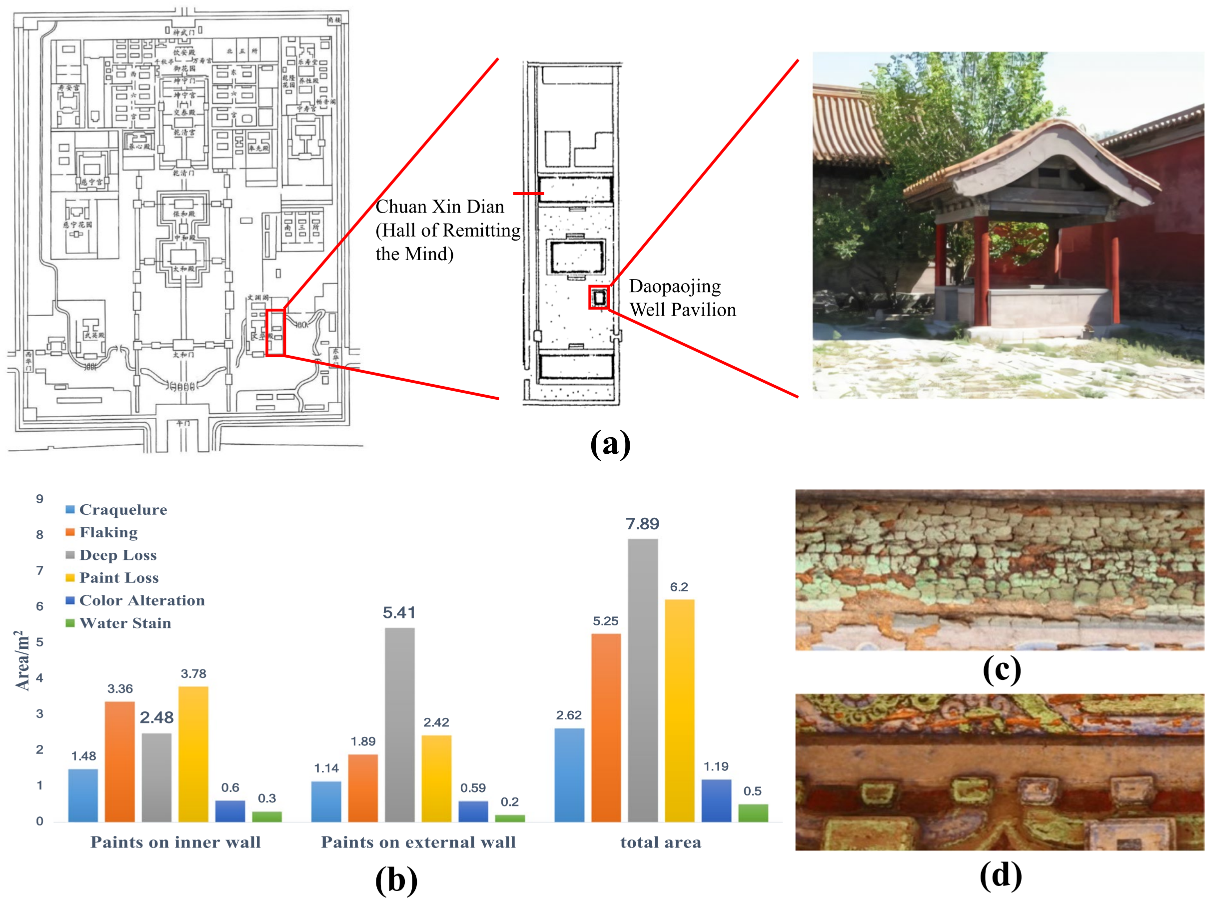

Fig. 1 **a** Geographical location map of Dapaojing Well Pavilion in the Palace Museum; **b** Statistics on the area of defects on polychrome architectural paintings of Dapaojing Well Pavilion; **c** Craquelure and **d** Paint loss

studied the relationship between crack characteristics and impact rate, stress, and other factors.

Traditional image processing methods identifying low-level features have long been used in the recognition of craquelure and paint loss. Recognition tasks are relatively easier for bulk paint loss. An improved region-growing algorithm fusing threshold segmentation realized the automatic extraction of mural paint loss [4]. Cao et al. [5] recognized mural paint loss in the Qu Tan Temple using a comprehensive method that integrated spectral features and Hu moment. Mishra [6] reviewed image processing techniques and computer vision-based methods used in classifying and quantifying damages in historical constructions. However, the end to end segmentation of craquelure is difficult. Methods based on threshold, edge detection, and graph theory can recognize crack-like networks such leaf vein, road, and vessel [7–10], but are often disturbed by noise in the background. Poor continuity and low contrast often lead to the degradation of such low-level feature-based methods [11].

With the development of deep learning, crack detection technology has made great progress. In 2016, Zhang [12] first applied deep learning to crack detection and proposed that neural networks are insensitive to complex background noise in images. With the emergence of FCN [13], a semantic segmentation neural network, Zou [14] and Shen [15] proposed DeepCrack and Crack_FCNN to achieve end-to-end crack recognition. Many kinds of networks, such as VGG, U-Net, and R-CNN, can perform the task of crack segmentation. Improving the accuracy of edge location is one of the main directions [16]. U-Net is often adopted because its encoder-decoder structure promotes more effective feature extraction, thus yielding both higher detection accuracy and efficiency [17].

Crack recognition using convolutional neural network (CNN) methods has been introduced in preserving cultural heritage, showing great potential to precisely and automatically recognize defects of relics. Kwon et al. [18] proposed a system to automatically detect and classify damage occurring in stone cultural properties using a faster region-based convolutional neural network (R-CNN) algorithm. Sizyakin et al. [19] realized the automatic recognition of cracks in paintings based on multimodal data. Mishra et al. [20] proposed a method of heritage structure defects detection based on the YOLOv5 algorithm, in view of typical damage diseases such as discoloration, exposed bricks, cracks, and spalling. Lv et al. [22] proposed an improved U-Net model to extract paint loss areas in murals. Wu et al. [23] proposed TMCrack-Net with feature fusion and a transformer for Tang Dynasty tomb chamber mural crack segmentation. However, no published research achieved pixel-level segmentation of cracks on polychrome paintings of Chinese traditional architectures using images.

This study proposes an improved Res-UNet semantic segmentation model for photography, whose loss function and training method are optimized. A dataset of craquelure and paint loss on polychrome paintings is built for model training. The model used is a new convolutional neural network that can identify craquelure on traditional Chinese architectural paintings at the pixel level. The model is used to realize the high-precision identification of defects on architectural paintings of the Palace Museum, and its performance is compared to that of other models.

Method

Data acquisition and augmentation

There is no open dataset of craquelure on ancient architectural polychrome paintings for machine vision. For this study, high-definition images were collected in the Wen Yuan Ge (Pavilion of Literary Profundity), Jian Ting (Archery Pavilion), and Dapaojing Well Pavilion in the Palace Museum, and supplementary images were collected in the Qi Sheng Wang Dian (the memorial hall of the Confucius' father) of the Temple of Confucius, Qufu, Shandong Province, and Fengguo Temple, Yixian, Liaoning Province (Fig. 2), whose paintings follow the same technique and standard.

After collection of images, preprocessing, labeling, and data augmentation were carried out to build a dataset for the training of a semantic segmentation neural network. Images were sliced, 600 images were selected, and LabelMe was used to label them. Single and combined operations such as flipping, inversion, translation, sharpening, HSV transformation, and the addition of noise were conducted to expand the dataset and avoid overfitting (Fig. 3). This last step was repeated until 21,000 polychrome paintings were obtained. The dataset was randomly divided into training and test sets at a 9:1 ratio.

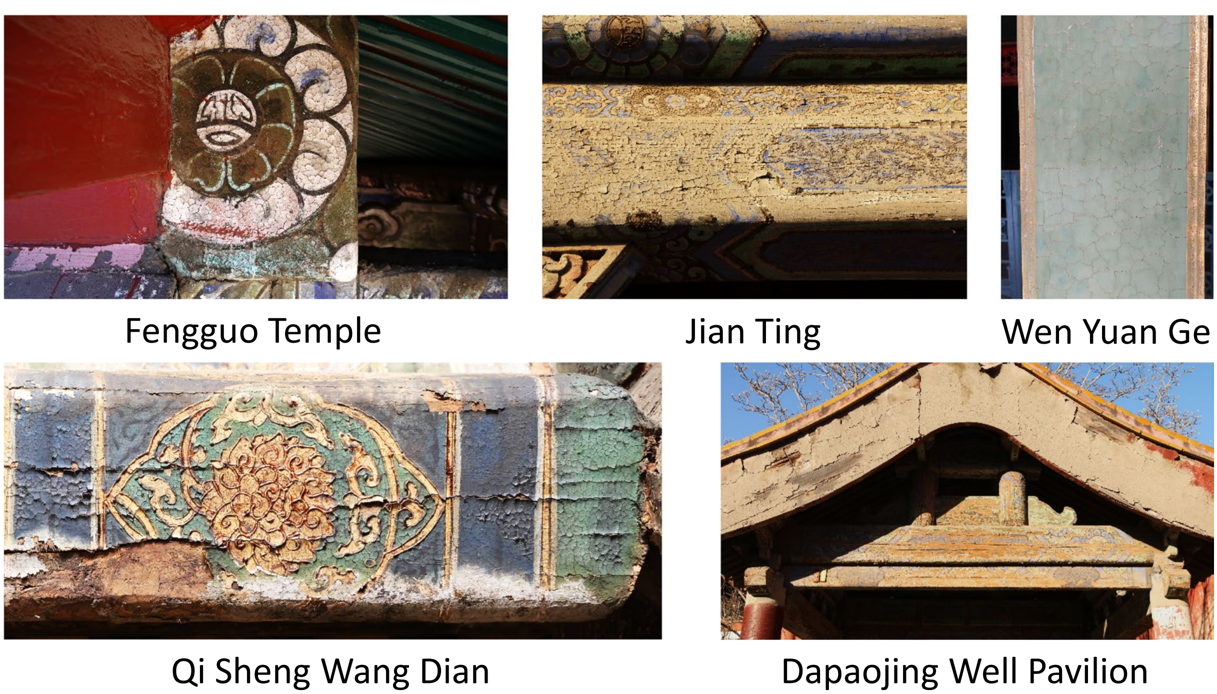

Fig. 2 Craquelure and paint loss in official architectural paintings in Ming and Qing dynasties

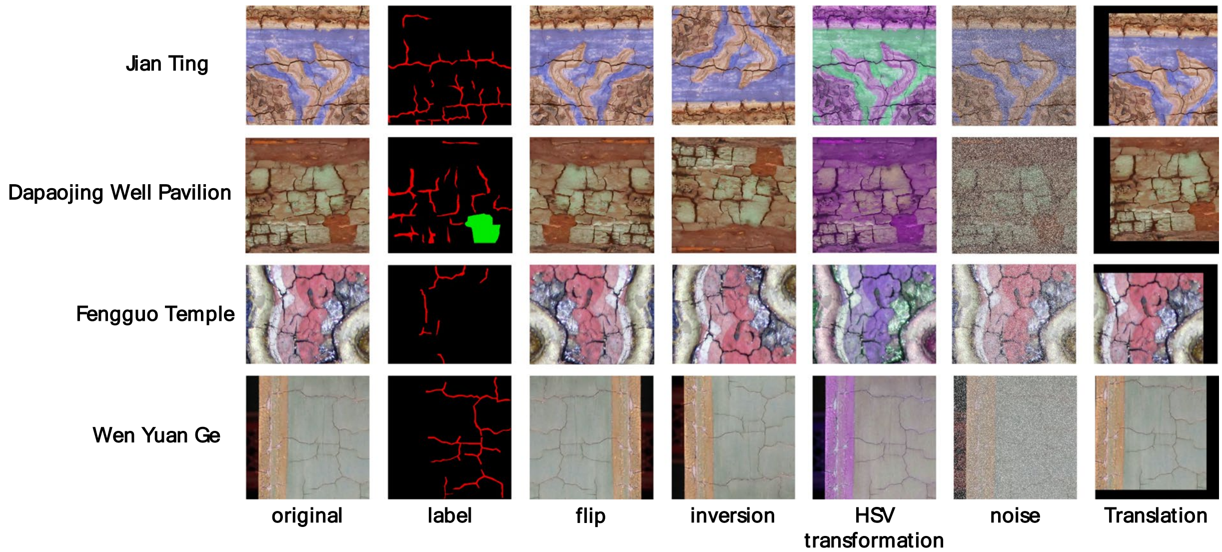

Fig. 3 Data augmentation

Model structure

U-Net is an encoder-decoder structure whose lower and deeper layers are used for pixel localization and classification, respectively. Combining features of different levels and using feature stitching and multiscale fusion to preserve segmentation details, it has shown excellent performance in image segmentation and edge detection [24]. It is worth mentioning that U-Net was initially used in the field of medical cell segmentation, where it achieved good results [25], and the intercellular space under a microscope is similar to craquelure on architectural polychrome paintings.

ResNet is widely used such as in face recognition and automatic driving, and has been proposed to solve the problems of gradient disappearance, gradient explosion, and network degradation caused by the increase of network layers. ResNet is essentially a process of continuously fitting residuals, which accelerates network convergence and improves the accuracy of the model. Commonly used ResNet models include ResNet-18,

Table 1 ResNet system structure diagram [26]

Layer name	Output size	18-layer	34-layer	50-layer	101-layer
conv1	112×112	7×7, 64, stride 2			
conv2_x	56×56	3×3 max pool, stride 2			
		$\begin{bmatrix} 3 \times 3 & 64 \\ 3 \times 3 & 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 & 64 \\ 3 \times 3 & 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3 & 128 \\ 3 \times 3 & 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 & 128 \\ 3 \times 3 & 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 4$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3 & 256 \\ 3 \times 3 & 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 & 256 \\ 3 \times 3 & 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 23$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3 & 512 \\ 3 \times 3 & 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 & 512 \\ 3 \times 3 & 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3$
	1×1	Average pool, fc, softmax			
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9

ResNet-34, ResNet-50, and ResNet-101, as shown in Table 1 [26]. According to a preliminary experiment, ResNet-50 was selected because of its acceptable segmentation speed, number of parameters, and performance in recognition tasks.

To meet the segmentation requirements of polychrome paintings and improve recognition performance for craquelures, we combine two kinds of networks in a Res-UNet network for semantic segmentation, using ResNet-50 as the encoder network for feature extraction, and U-Net as the decoder network for feature fusion, as shown in Fig. 4a. In the process of feedforward calculation, ResNet-50 subsamples five times, with 7×7 convolution and four bottleneck blocks. The bottleneck architecture is shown in Fig. 4b. Two 1×1 convolution operations are used to respectively reduce and increase the feature dimension, which greatly reduces calculation, and is beneficial to data training and feature extraction. The decoder part of Res-UNet retains the U-Net upsampling process. The unique skip connection of U-Net can achieve feature fusion at the channel level. Upsampling and 1×1 convolution reduce the number of channels and restore the resolution of the original image. Finally, the segmentation results of craquelure and painting loss are obtained.

Neural network training

Experiments were performed on a Windows platform with Nvidia GeForce GTX 1080Ti hardware, using Python 3.8.12 and a PyTorch framework. Experiments used freeze training, the epoch was set to 100, the batch size was set to 8, the learning rate was $\mathrm{Lr} = 1 \times 10^{-4}$, the minimum learning rate was $\mathrm{Lr\_min} = 1 \times 10^{-6}$, the learning rate decay method was cosine annealing, and the optimizer was Adam.

Transfer learning

Due to the problems of low quality, low quantity, difficulty in the acquisition of architectural polychrome painting photos, and similarity of cellular organs in medical images in terms of geometry and texture to craquelure and paint loss, transfer learning of the proposed model was performed on the basis of a medical cell segmentation dataset. Pretrained weights were loaded into a base model, and all layers were frozen by setting trainable = false, so as to learn the features of craquelure and paint loss, decrease training time, and reduce memory usage.

Loss function

The loss function has a great impact on the effect of model optimization. One of the most commonly used loss functions, cross-entropy (CE) loss, can stably return the gradient and effectively solve the problem of gradient disappearance during the feedback process of the network. However, this function evaluates all categories in the image equally, which may result in poor recognition when applied to a non-equilibrium classification problem such as craquelure, which accounts for a small percentage in a picture.

Therefore, we used Focal Loss [27], Dice Loss [28], and CE Loss to form mixed loss functions. The total loss can be formulated as

$$CE(p_t) = -\alpha_t \log(p_t) \tag{1}$$

$$\alpha_t = \begin{cases} \alpha & \text{if } y = 1 \\ 1 - \alpha & \text{otherwise} \end{cases} \tag{2}$$

$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise} \end{cases} \tag{3}$$

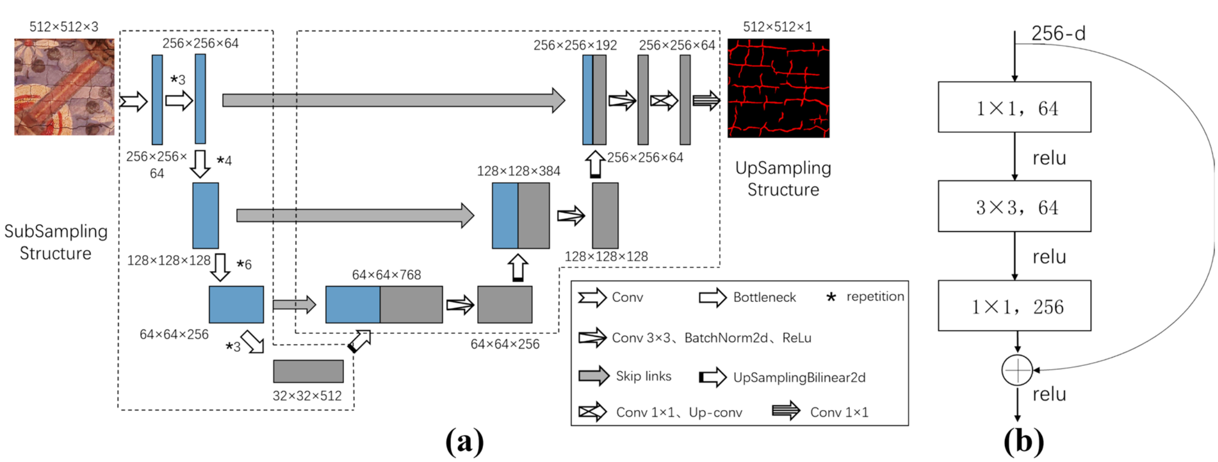

Fig. 4 **a** Res-UNet network and **b** Bottleneck structure

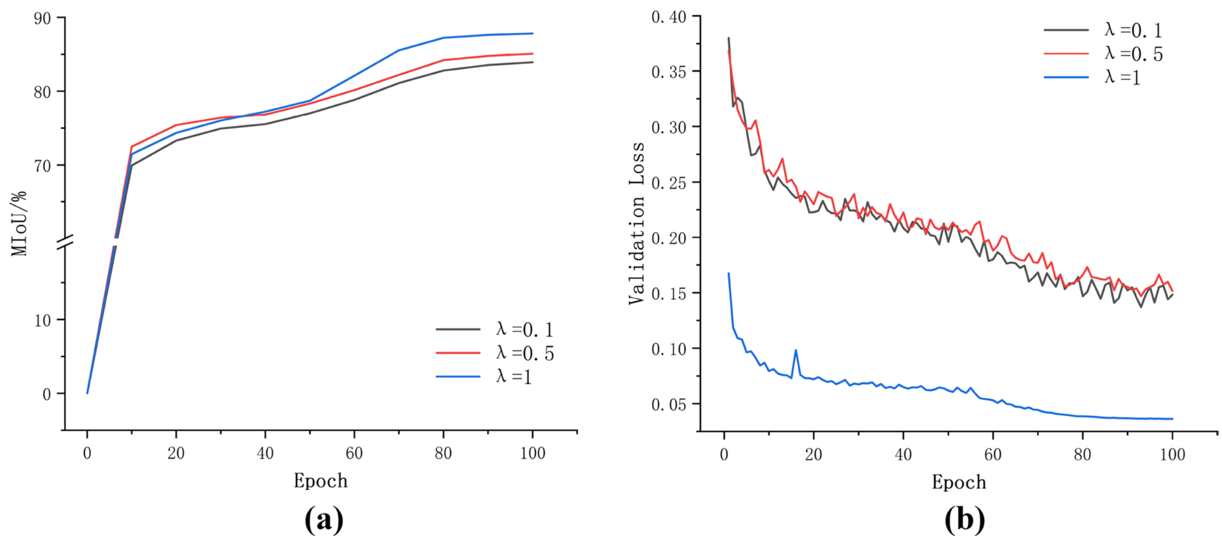

Fig. 5 The influence of λ on **a** MIoU curve and **b** validation loss curve

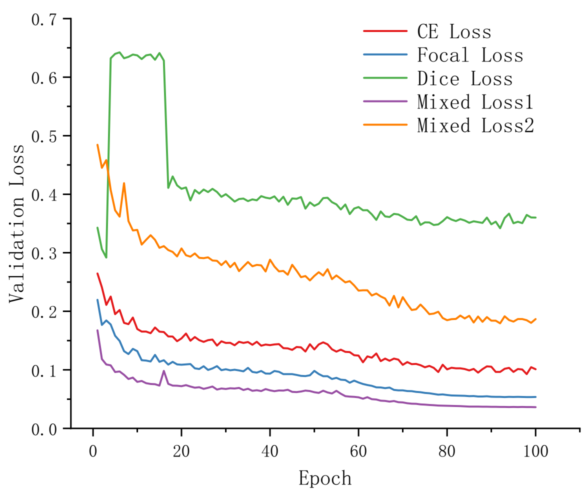

Fig. 6 Loss function curve of Res-UNet using CE Loss, Focal Loss, Dice Loss, Mixed Loss1, and Mixed Loss2 in segmentation of validation dataset during training

Table 2 IoU indexes corresponding to different loss functions

	Background	Craquelure	Paint loss
CE Loss	0.97	**0.59**	0.93
Dice Loss	0.98	**0.67**	0.93
Focal Loss	0.98	**0.69**	0.93
Mixed Loss2	0.97	**0.66**	0.93
Mixed Loss1	0.98	**0.71**	0.94

$$\text{Focal Loss}\,(p_t) = -\alpha_t(1-p_t)^{\gamma}\log(p_t) \tag{4}$$

$$\text{Dice Loss} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \tag{5}$$

$$L_{\text{Mixed1}} = \text{Dice Loss} + \lambda\text{Focal Loss} \tag{6}$$

$$L_{\text{Mixed2}} = \text{Dice Loss} + \text{Focal Loss} + \text{CE Loss} \tag{7}$$

where p_t is the prediction probability of the sample, y is the label of the sample,α_t, λ, and γ are hyperparameters, α_t and γ are used to balance positive and negative samples and adjust the weights of difficult samples respectively, λ is a coefficient adjusting the influence of Focal Loss and Dice Loss. X is the prediction map, Y is the ground truth, L_{Mixed1} is Mixed Loss1, L_{Mixed2} is Mixed Loss2.

Mixed Loss1, Mixed Loss2, Focal Loss, Dice Loss, and CE Loss were used to train the proposed model, and their effects were compared using the Intersection over Union (IoU) index and smoothness of convergence.

Comparison of recognition results

To evaluate the recognition ability of the proposed Res-UNet, two networks are trained using the same dataset. One network is U-Net, and the other is a multi-classification lightweight network, MC-DM [29], which is a full convolutional neural network model for mural segmentation, characterized by a deep separable convolution structure to fuse multiscale features and achieve a light weight. After training the three networks, RGB images of polychrome paintings of any size were input for testing and recognition. Accuracy (Acc) [30], Recall, Precision, F1-score, Mean Intersection over Union (MIoU), FPS (Frames Per Second,

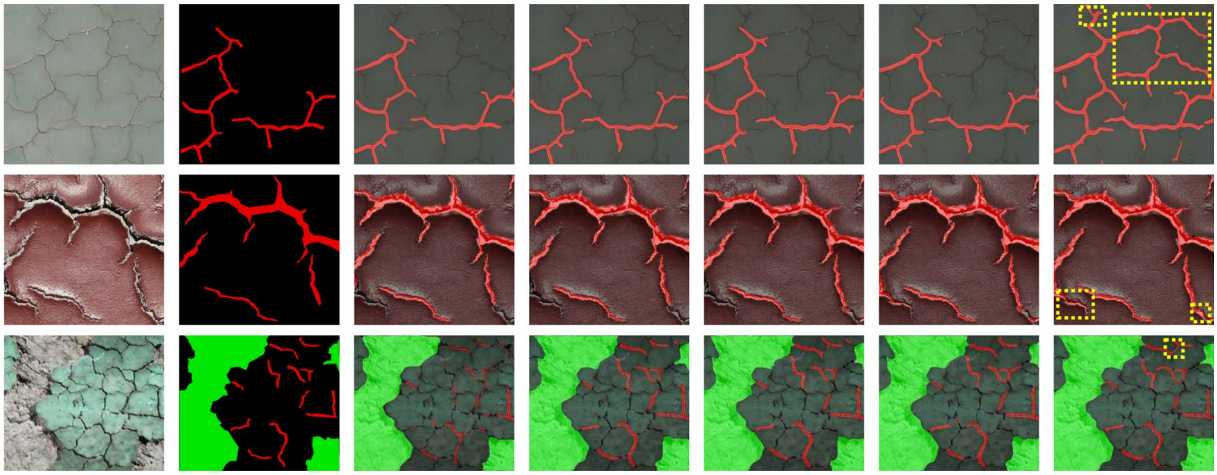

Fig. 7 Prediction results corresponding to different loss functions

indicating the number of images segmented per second), and human judgment were used to evaluate recognition ability of the three networks, and the calculation formulas are as follows:

$$\mathrm{Acc} = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

$$\mathrm{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\mathrm{Recall} = \frac{TP}{TP + FN} \tag{10}$$

$$\mathrm{F1\text{ - }score} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{11}$$

$$\mathrm{MIoU} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{TP}{TP + FN + FP} \tag{12}$$

where the TP represents the number of positive samples predicted correctly, the FP represents the negative samples predicted falsely, the TN represents the number of negative samples predicted correctly, the FN represents the positive samples missed, and k is the number of categories.

The proposed network performs the following steps: (1) adjust the image resolution to 512×512 using the resize function; (2) use the bottleneck structure in the ResNet-50 encoder network for low-level feature extraction; (3) copy and crop the low-level features obtained by each bottleneck structure, input them to the decoder network, and fuse the deep semantic features obtained by bilinear interpolation upsampling through a skip connection; (4) carry out upsampling and 1×1 convolution on the feature fusion image, restore the image resolution, reduce the channel number dimension, and obtain the semantic segmentation image.

Results and discussion

Effect of loss function

Firstly, hyperparameters in Focal Loss (Eq. 4) and Mixed Loss1 (Eq. 6) were optimized. According to preliminary experiments, Focal Loss performed better when $\alpha_t = 0.5$ and $\gamma = 2$. The influence of λ on MIoU and validation loss are shown in Fig. 5. It is clear that higher value of λ can improve network accuracy, indicating Focal Loss acted as important as Dice Loss in Mixed Loss1.

CE Loss, Dice Loss, Focal Loss, Mixed Loss1, and Mixed Loss2 were used to train the model, and the convergence and IoU were compared. The loss function curve during training is shown in Fig. 6. Mixed Loss1 outperformed three individual loss functions. It can effectively reduce the loss of the validation dataset and produce a smoother convergence curve.

The IoU indexes using the five loss functions are shown in Table 2, and recognition results of three images are shown in Fig. 7. It is clear that IoU is the highest when Mixed Loss1 is used. In the recognition of craquelure, which is the object of this study, IoU increases by 2–12%. In Fig. 7, it is seen that Mixed Loss1 also produced better results in all three images by avoiding insufficient recognition. Most concerned areas are emphasized by yellow dashed box. It is worth mentioning that the recognition result is very close to the label map when other

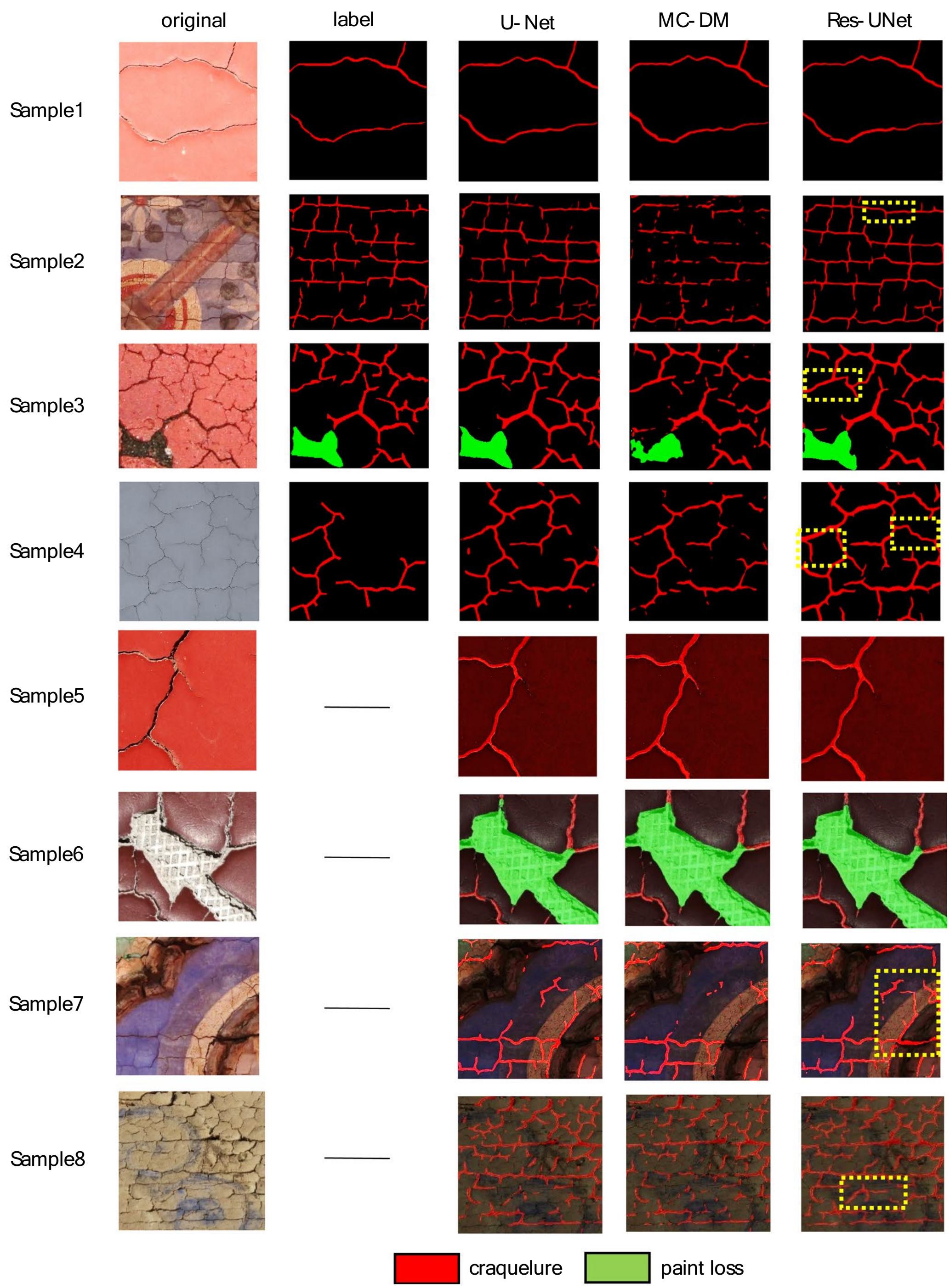

Fig. 8 Effects of three models in segmenting craquelure and paint loss on polychrome architectural paintings in the Palace Museum

loss functions are used, indicating that overfitting can be avoided by the use of Mixed Loss1.

In conclusion, the proposed loss function Mixed Loss1 retains the advantage of different functions. It can strengthen the learning of craquelure features and improve recognition accuracy. According to previous studies, Focal Loss can reduce the weight of easy classification samples, make the model pay more attention to difficult classification samples, and preserve complex edge details. Dice Loss can solve the problem of

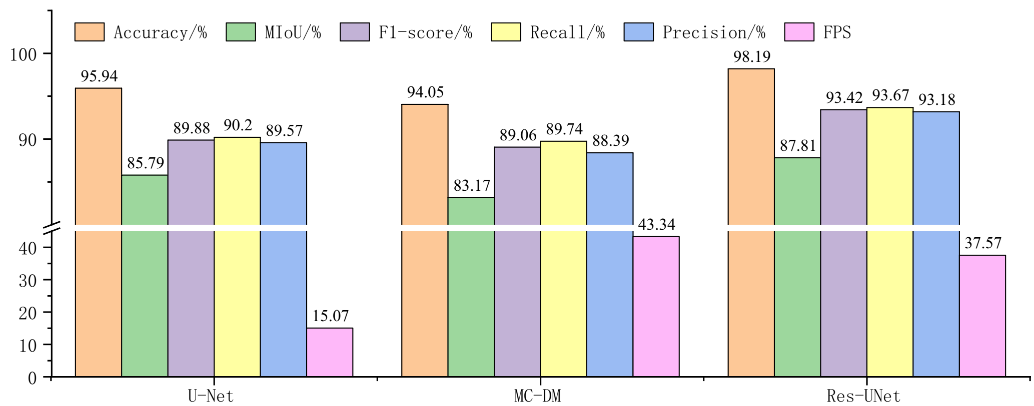

Fig. 9 Comparison of evaluation indexes of three models

imbalance in pixel categories, reduce the noise caused by Focal Loss, and improve segmentation accuracy [31]. However, Mixed Loss2 performed worse than Focal Loss and CE Loss. The performance of the network cannot be improved by simply increase loss function complexity. Weights of different loss functions should be adjusted to achieve better results.

Comparison of models

Figure 8 qualitatively compares the three models. Among eight segmented images, samples 1–4 are in the test dataset, and samples 5–8 are not. In samples 1, 5, and 6, which include simple lines and monotonous backgrounds, all models give similarly good results. However, as shown in the yellow dashed box in samples 2–4, 7, and 8, the deep separable convolution of MC-DM optimizes the feature extraction in image segmentation, but loses many details, resulting in seriously incomplete recognition and poor continuity of segmentation results. U-Net performs well, as its unique skip connection structure can be well combined with the background semantic information for multiscale segmentation, but there are still some undetected craquelures. In the Res-UNet model, the encoder structure is modified, the number of network layers is deepened, and the number of parameters is reduced, and a residual structure is used to accelerate the convergence of the model. The improved loss function effectively solves the problem of insufficient recognition of difficult-learning samples, improves segmentation accuracy, and obtains more precise and complete recognition results.

Accuracy, Recall, Precision, MIoU, and FPS were used as objective evaluation metrics for semantic segmentation, whose results are shown in Fig. 9 for three models.

The indexes of Res-UNet are better than those of the other two networks. Compared with U-Net, the values of Accuracy, Recall, Precision, F1-score, and MIoU increased by 2.25%, 3.47%, 3.61%, 3.54%, and 2.02%, respectively, and by 4.14%, 3.93%, 4.79%, 4.36%, and 4.64% compared with MC-DM. The lightweight MC-DM model had the fastest computing speed. By introducing the residual module of ResNet-50, Res-UNet reduces the number of parameters. Hence Res-UNet had much better segmentation efficiency than the original U-Net, and was slightly worse than MC-DM.

The comprehensive comparison results show that the improved Res-UNet model had a better segmentation effect than the other two models, had higher accuracy and a faster recognition speed, and effectively avoided the missed detection and misdetection problems in the other models, and the segmentation edge of the model tended to the ideal contour.

Conclusions

Craquelure and paint loss are the most common defects on traditional Chinese architectural polychrome paintings. The identification of cracks has been a problem because of the influence of original lines and patterns. In this study, a dataset was established, and U-Net was modified to accurately identify craquelure and paint loss. The improved U-Net uses ResNet-50 as the encoder network, which retains the advantages of a skip connection, avoids excessive thickening of the craquelure boundary, and deepens the network through the residual structure. In addition, the use of a mixed loss function avoided overfitting caused by category imbalance, strengthened the extraction of semantic features, and increased segmentation accuracy. Experiments showed that the proposed network could accurately and fully identify craquelure and paint loss on the polychrome paintings in the Palace Museum. Accuracy, Recall, Precision, MIoU, and FPS have achieved 98.19%, 93.67%, 93.18%, and 87.81%, 37.57, respectively. The proposed network outperformed the original U-Net and MC-DM.

As a fast, objective, and stable method, machine vision has broad application prospects in the field of cultural heritage. Based on the precise recognition of cracks, the research of defect quantification, crack monitoring, and deterioration mechanisms is underway.

Acknowledgements
The authors would like to express their great gratitude to Department of Ancient Architecture of the Palace Museum, Ms. Mengmeng Cai, Mr. Binhao Wu, and Ms. Bingbing Yang. We thank LetPub for its linguistic assistance during the preparation of this manuscript.

Author contributions
Conceptualization, X. He; Data acquisition, Q. Yuan and X.N. Han; Dataset construction, Q. Yuan; Code compiling, Q. Yuan; Funding acquisition, X. He and X.N. Han; Methodology, X. He and Q. Yuan; Resources, H. Guo; Supervision, X.N. Han; Writing—original draft, Q. Yuan; Writing—review and editing, X. He. All authors have read and agreed to the published version of the manuscript.

Funding
This work was financially supported by the National Key R&D Program of China (Grant No. 2020YFC1522404).

Availability of data and materials
The datasets generated and/or analysed during the current study are not publicly available due to intellectual property issues but are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 15 November 2022 Accepted: 28 February 2023
Published online: 31 March 2023

References

1. Qi Y. The progress of the research on conservation and restoration of colored pattern of traditional Chinese architecture. Sciences of Conservation and Archaeology. 2008;20(S1):109–13. https://doi.org/10.16334/j.cnki.cn31-1652/k.2008.s1.020. **(in Chinese)**.
2. Li CW, Ai DH, Sun XY, Xie BJ. Crack identification and evolution law in the vibration failure process of loaded coal. Journal of Geophysics and Engineering. 2017;14(4):975–86.
3. Li CW, Ai DH. Automatic crack detection method for loaded coal in vibration failure process. PLoS ONE. 2017;12(10):e0185750.
4. Cao JF, Li YF, Cui HY, Zhang Q. The application of improved region growing algorithm for the automatic calibration of shedding disease on temple murals. Journal of Xinjiang University (Natural Science Edition). 2018;35(04):429–36. https://doi.org/10.13568/j.cnki.651094.2018.04.007. **(in Chinese)**.
5. Cao PH, Lv SQ, Wang WF, Gao ZH, Hou ML. Extraction of mural paint loss regions based on spectral dimensionality reduction and Hu moment. Journal of Graphics. 2020;41(06):930–8 **(in Chinese)**.
6. Mishra M. Machine learning techniques for structural health monitoring of heritage buildings: a state-of-the-art review and case studies. Journal of Cultural Heritage. 2021;47:227–45.
7. Katyal V, Aviral. Leaf vein segmentation using Odd Gabor filters and morphological operations. Computing Research Repository. 2012;abs/1206.5157.
8. Hassan G, El-Bendary N, Hassanien AE, Fahmy A, Snasel V. Retinal blood vessel segmentation approach based on mathematical morphology. Procedia Computer Science. 2015;65:612–22.
9. Strisciuglio N, Azzopardi G, Petkov N. Detection of curved lines with B-COSFIRE filters: a case study on crack delineation. Springer. 2017;108–20.
10. Kheradmandi N, Mehranfar V. A critical review and comparative study on image segmentation-based techniques for pavement crack detection. Construction. 2022;321: 126162.
11. Ali R, Chuah JH, Talip MSA, Mokhtar N, Shoaib MA. Structural crack detection using deep convolutional neural networks. Automation in Construction. 2022;133: 103989.
12. Zhang L, Yang F, Zhang YD, Zhu YJ. Road crack detection using deep convolutional neural network. In: 2016 IEEE international conference on image processing (ICIP). IEEE. 2016;3708–12.
13. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015;3431–40.
14. Zou Q, Zhang Z, Li QQ, Qi XB, Wang Q, Wang S. DeepCrack: learning hierarchical convolutional features for crack detection. IEEE transactions on image processing. 2018;28(3):1498-512.
15. Shen JK (2019) Research on crack detection algorithm of concrete structure based on computer vision [Master]. Institute of Engineering Mechanics, China Earthquake Administration. (in Chinese)
16. Li H, Wang W, Wang M, Li L, Vimlund V. A review of deep learning methods for pixel-level crack detection. Journal of Traffic and Transportation Engineering (English Edition). 2022;9(06):945-68.
17. Zhou S, Canchila C, Song W. Deep learning-based crack segmentation for civil infrastructure: data types, architectures, and benchmarked performance. Automation in Construction. 2023;146:104678.
18. Kwon D, Yu J. Automatic damage detection of stone cultural property based on deep learning algorithm. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. 2019;42:639–43.
19. Sizyakin R, Cornelis B, Meeus L, Dubois H, Martens M, Voronin V. Crack detection in paintings using convolutional neural networks. IEEE Access. 2020;8:74535–52.
20. Mishra M, Barman T, Ramana G. Artificial intelligence-based visual inspection system for structural health monitoring of cultural heritage. Journal of Civil Structural Health Monitoring. 2022;1–18.
21. Zhang YE, Wu LG. The detection of cave mural damage based on deep learning. Yungang Research. 2022;2(01):85–90. https://doi.org/10.19970/j.cnki.ISSN2096-9708.2022.01.012. **(in Chinese)**
22. Lv SQ, Wang SH, Hou ML, Gu MY, Wang WF. Extraction of mural paint loss diseases based on improved U-Net. Geomatics World. 2022;29(01):69–74 **(in Chinese)**.
23. Wu M, Jia M, Wang J. TMCrack-Net: a U-shaped network with a feature pyramid and transformer for mural crack segmentation. Applied Sciences. 2022;12(21):10940.
24. Yang J, Faraji M, Basu A. Robust segmentation of arterial walls in intravascular ultrasound images using dual path U-Net. Ultrasonics. 2019;96:24–33.
25. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer. 2015;234–41.
26. He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016;770–8.
27. Lin TY, Goyal P, Girshick R, He KM, Dollár P. Focal loss for dense object detection. Proceedings of the IEEE international conference on computer vision. 2017;2980–8.
28. Li XY, Sun XF, Meng YX, Liang JJ, Wu F, Li JW. Dice Loss for data-imbalanced NLP tasks. Computer Research Repository. 2019;abs/1911.02855.

29. Chen ZQ, Rajamanickam L, Tian XD, Cao JF. Application of optimized convolution neural network model in mural segmentation. Applied Computational Intelligence. 2022;2022.
30. Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V, Garcia-Rodriguez J. A review on deep learning techniques applied to semantic segmentation. arXiv preprint. arXiv:170406857. 2017.
31. Zhu WT, Huang YF, Zeng L, Chen XM, Liu Y, Qian Z, et al. AnatomyNet: deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy. Med Phys. 2018;46(2):576–89.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.