

EXTRACTION OF PAINT LOSS IN ANCIENT MURALS BASED ON 3D RESIDUAL NEURAL NETWORK

S. Y. Li^{1,2}, M. L. Hou^{1,2}, P. H. Cao³, S. Q. Lyu^{1,2,*}

¹ School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, No.15 Yongyuan Road, Daxing District, Beijing, China -2108570021116@stu.bucea.edu.cn, houmiaole@bucea.edu.cn, lvshuqiang@bucea.edu.cn.

² Beijing Key Laboratory for Architectural Heritage Fine Reconstruction & Health Monitoring, No.15 Yongyuan Road, Daxing District, Beijing, China

³ Shenzhen Feima Robotics Technology Co., LTD., No.8 Heiquan Road, Haidian District, Beijing, China -15210481449@163.com.

Commission IV, WG IV/11

KEY WORDS: Mural Hyperspectral Images, Spectral Reflectance, Paint Loss, Deep Learning, Residual Network

ABSTRACT:

The ancient murals of Qutan Temple in Qinghai Province have a very serious loss of paint. Moreover, the main components of the base color paint layer in the paint loss area and the white patterns in the murals are both calcified, which are similar in color and spectral features. Thus, it is difficult to distinguish them by only using spectral features. A method of paint loss area extraction based on 3D residual network with multi-scale feature fusion is proposed. Firstly, the hyperspectral images with paint loss regions were collected by hyperspectral images. They are pre-processed to establish the training data set. Secondly, 3D residual network models are constructed using $3\times 3\times 3$, $3\times 3\times 5$ and $5\times 5\times 3$ convolution kernels to realize the extraction and fusion of spatial and spectral features at different scales of hyperspectral images. The produced mural hyperspectral dataset is used for network training to obtain the prediction model. Finally, the hyperspectral images are input into the trained model to achieve the extraction of paint loss. After comparing different methods, the experimental result shows that the proposed method can improve the extraction accuracy of mural paint loss and serve as a reference for other deteriorations extraction.

1. INTRODUCTION

Mural painting is one of the carriers of human civilization, which is a unique form of artistic expression, bearing the cultural essence of the creation era. It is one of the priorities of heritage conservation research because of its magnificent colors and rich expressions. After thousands of years, mural paintings, due to environmental changes and human factors, are under the threat of a variety of deteriorations, such as paint loss, fading, colour changing, flaking, etc. Their aesthetic value is destroyed (Cao J F, 2020). It is an essential task of survey in the conservation of murals, and computer-aided mural paint loss detection, which can quickly and accurately determine the distribution of murals' deteriorations. It is a hot spot for current research.

In recent years, many scholars have conducted a series of studies on the extraction of deteriorations in Mural paintings. Li C Y (2016) proposed an automatic mud-spot calibration algorithm, which used spatial autocorrelation to analyze the texture features of mud-spot, and achieved the calibration of mud-spot on mural paintings in Tang burial chambers by using mud-spot texture, brightness and other features through threshold segmentation under the YcbCr model. Zhang Z Y (2021) used weighted averaging to convert color images into grayscale maps before virtual restoration of mural crack lesions and extracted mural lesion masks using local optimal hierarchical clustering. Roman (2020) proposed a fast crack detection algorithm based on deep convolutional neural networks (CNN), which improved the CNN with actual crack boundaries, which allowed the network to effectively improve the classification results and achieve the extraction of cracked lesions in the "Ghent Altarpiece". This type of research is

mainly based on digital images, using the spatial features for extraction, and had not yet utilized its spectral features. Hyperspectral remote sensing has been applied to the digital conservation of cultural relics with features such as map unification and non-contact imaging (Gong M T, 2014). For example, Zhou P P (2019) used the spectral difference between stains on Chinese paintings and paper for feature band selection to remove stains and reconstruct images by color constrained Poisson editing. Li P (2018) used principal component analysis and sparse self-encoder to extract spectral features, clustered the extracted spectral features with a point sorting recognition clustering structure density clustering algorithm, and visualized the results. The analysis was achieved for the assessment of Dunhuang mural nail. Liu X Q (2019) used support vector machines to classify hyperspectral images of murals to achieve the identification of mural paint loss. Such studies mainly exploited the spectral features of hyperspectral images and lacked the utilization of spatial features of paint loss.

With the development of artificial intelligence, several scholars have improved the deep learning network so as to achieve the deteriorations recognition and extraction of cultural relics. Laurens Meeus (2019) improved the structure of U-Net network so that the network could extract image features through multiple paths, and finally the improved U-Net was used to achieve the extraction of the paint loss area of the Ghent Altarpiece. Lyu S Q (2022) improved the downsampling region of the U-Net network to enhance the learning ability of the network detail information and produced a dataset of pigment layer peeling from multiple temple murals and realized the automatic extraction of pigment layer peeling degradation using the improved network. The above research mainly focuses on digital images, and numerous scholars have also conducted

research on hyperspectral image classification. Cao P H (2020) used support vector machine (SVM) classification to classify hyperspectral images and introduced Hu moments to discriminate the shape of white patterns and paint loss areas as a way to extract paint loss detection. In summary, there are few studies on the identification and extraction of heritage deteriorations based on hyperspectral images and deep learning, which are still in the exploration stage.

The paint loss area in the mural painting and the white pattern share some similarities in terms of color and spectral features, which makes it challenging to differentiate between the two solely based on spectral analysis. Therefore, a multi-scale spatial-spectral feature fusion method is proposed for the extraction of paint loss area from murals. A 3D residual network model is constructed using hyperspectral images of some murals to produce a dataset for the extraction and fusion of spatial and spectral features at different scales of hyperspectral images. The fusion of spectral and spatial features is used to improve the accuracy and efficiency of paint loss extraction, which can provide reference for the extraction of other mural deteriorations.

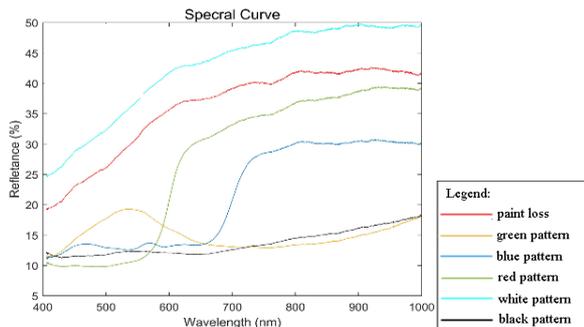
2. MATERIAL AND METHODS

2.1 Overview of the research area

Hyperspectral data were collected from the murals of Qutan Temple in Qinghai Province, which was built in the 25th year of the Ming dynasty (1392 AD). The scale of the complex is magnificent, with a total area of 1338 m² of murals and 358 m² of murals on the inner walls of the cloister (Li F S, 2020). The murals in the temple were mostly painted in the Ming and Qing dynasties and have high research value. However, because of their longevity, the surface of the murals has serious degradations. The hyperspectral images of some of the murals in the west cloister of Qutan Temple were selected for the study of shedding extraction, and their locations are shown in the red box of Figure 1(a). The average reflectance spectral curves of different color patterns are shown in Figure 1(b).



(a) Part of the mural orthophoto in the west cloister of Qutan Temple



(b) Average reflection spectrum curve of different color patterns

Figure 1. Research area information

The research region contains green pattern, blue pattern, red pattern, white pattern, black pattern and paint loss area. Paint loss is the main deterioration types of the mural, which visually appears as a white irregular region with large variability from other color patterns, which were easy to distinguish. However, the paint loss area and the white patterns on the mural are visually similar in color, with similar spectral curves. They had only small differences in reflection intensity, so it is difficult to achieve high extraction accuracy only by relying on spectral features.

2.2 Processing of Multi-scale spatial-spectral feature fusion for paint loss extraction

For the paint loss area of the mural pigment layer is similar to the reflectance spectrum of the white pattern and difficult to distinguish, a multi-scale feature fusion method is proposed to extract the paint loss area of the pigment layer taking into account the spatial and spectral features. The process of paint loss extraction is shown in Figure 2.

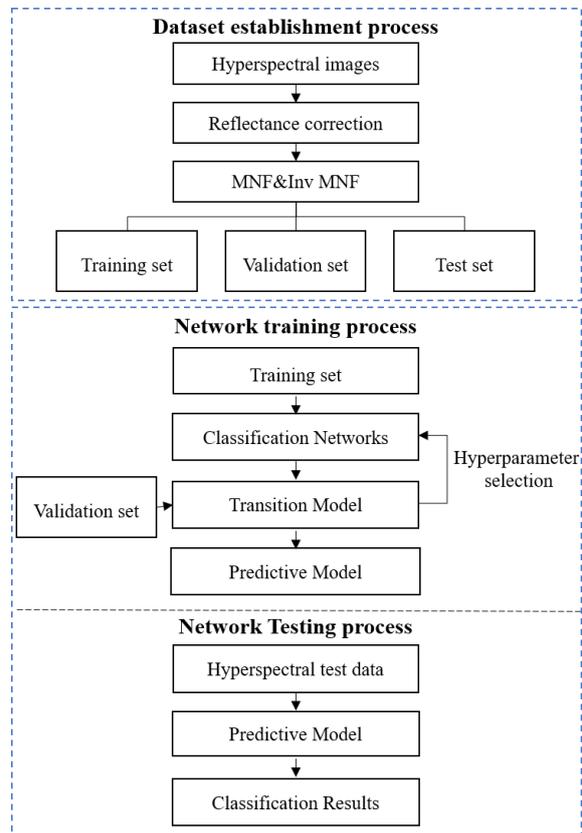


Figure 2. Paint loss extraction flow chart

Data preprocessing: The original hyperspectral image is reconstructed using equation (1) to obtain the reflectance image. The minimum noise fraction (MNF) transform is performed on it, and the first m bands with high signal-to-noise ratio are selected for MNF inversion transformation to remove hyperspectral data noise. Some of the pre-processed hyperspectral data are selected for dataset production, and the image element classes are manually calibrated and proportionally divided into a training set and a validation set for network training. The equation is as follows:

$$R = \frac{R_{Data} - R_{Dark}}{R_{White} - R_{Dark}} \times 99\% \quad (1)$$

Where R is the reflectance; R_{Data} is the hyperspectral data; R_{Dark} is the dark current data; R_{White} is the reflectance of 99% standard reflector data.

Network Training: The produced training set is fed into the deep learning network, and different hyperparameters are selected to train the network and generate the transition model. In the process of model updating, the network is optimized using the backpropagation algorithm. Using the results predicted by the transition model and the manual calibration results, the cross entropy is calculated according to equation (2) as the loss value for the network training. Through several iterations, the loss value decreases continuously, and when it tends to be stable, the model training is considered complete. The test set is input into the transition model, and the classification accuracy of the model is calculated according to equation (3). If it meets the expectation, the prediction model is saved as an over model, and if not, the model hyperparameters are adjusted and retrained.

$$Loss(y^1, y^2) = \sum_{i=1}^L y_i^1 \left(\log \sum_{j=1}^L e^{y_j^2} - y_i^2 \right) \quad (2)$$

Where y^1 are the artificially calibrated image element categories; y^2 are the model predicted image element classes; L are total number of pixels.

$$P = \frac{\sum_{i=1}^C Q_i^T}{Q^A} \quad (3)$$

Where P are the artificially calibrated image element categories; Q_i^T is the correctly classified pixels; Q^A are the total number of pixels.

Network Testing: The hyperspectral data of murals at other locations are selected and input into the trained prediction model in order to make predictions, and the final prediction results are obtained. Thus, the model is verified to have good generalization ability.

2.3 Three-dimensional convolutional neural network

Two-dimensional convolutional neural networks (2D-CNN) have been widely used in the field of digital images, such as classification of digital images (Krizhevsky A, 2012), target detection (Zoph, B 2018), etc. Compared with digital images, hyperspectral images have more bands, and a large number of parameters are introduced during the training process, which affects the computational efficiency of the network and is more prone to overfitting. While the convolution kernel of 2D-CNN is two-dimensional, which can only effectively use the spatial features of the image and lacks the utilization of the spectral features of the image. The convolution kernel of the three-dimensional convolutional neural network (Ji S W, 2013) (3D-CNN) increases one dimension compared to that of the 2D-CNN, which allows the network to utilize three different dimensional features of the input data at the same time, and its formula is shown in equation (4):

$$O^{xyz} = \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} \sum_r W_m^{pqr} i_m^{(x+p)(y+q)(z+r)} \quad (4)$$

Where O^{xyz} denotes the output value; i denotes the input value; M denotes the number of convolution kernels; P and Q are the size of the spatial dimensions of the convolution kernels; and R denotes the size of the spectral dimensions of the convolution kernels.

3D-CNN can utilize both spatial and spectral features, which reduces the occurrence of overfitting phenomenon and is more

suitable for hyperspectral image processing (Zhong Z L, 2017). While the residual network can solve the problem of gradient dispersion and gradient explosion caused by the deepening of network layers, combining the two and constructing 3D residual network for classification of hyperspectral images can significantly improve the efficiency of network training.

2.4 Three-dimensional residual networks for multi-scale feature fusion

The residual network (ResNet) is composed of residual blocks constructed by shortcuts and constant mappings. It is commonly used to solve the gradient dispersion and explosion problems due to network deepening (He K, 2016). The constant mapping is shown in equation (5).

$$F(x) = H(x) - x \quad (5)$$

Where $H(x)$ is the output value; x is the input value; $F(x)$ is the learning functions. When $F(x) = 0$, which means the constant mapping $H(x) = x$.

A Multi-scale 3D Residual Network (M3RN) with multi-scale feature fusion is designed to extract the spatial and spectral features of the mural (Jahandad, 2019). Its residual block structure is shown in Figure 3 and the network structure is shown in Figure 4.

The $3 \times 3 \times 3$, $5 \times 5 \times 3$, and $3 \times 3 \times 5$, three different spatial and spectral convolutional kernels are used in the 3D residual block for multi-scale feature fusion to extract the spatial-spectral features of hyperspectral images at different scales. Considering the small spatial size of the input data, the 7×7 convolutional kernels are not utilized. Inception structure is used to connect different convolutional kernel channels to achieve the fusion of different features, which will introduce a large number of parameters and affect the network efficiency. In order to improve the network operation speed, reduce the network parameters, and improve the network computational efficiency, a 3D residual block of multiscale feature fusion is introduced to keep the number of convolutional kernel channels constant during feature fusion, which we named add layer.

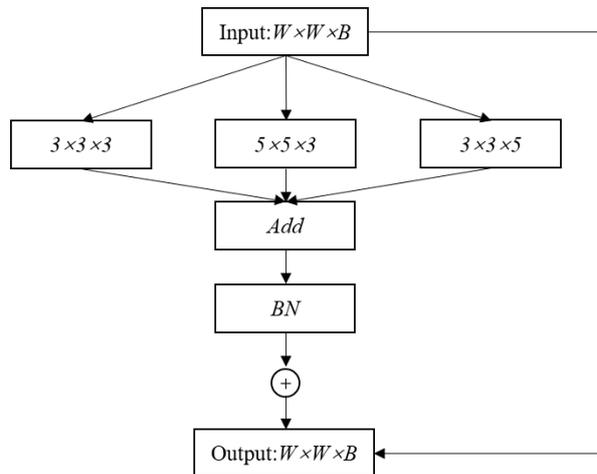


Figure 3. 3D-residual block of multi-scale feature fusion

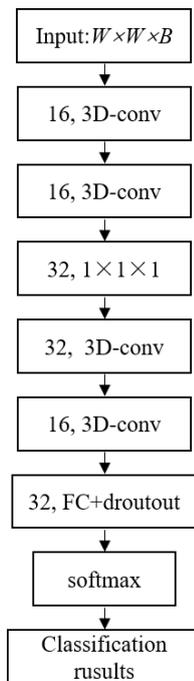


Figure 4. 3D-residual block of multi-scale feature fusion

As can be seen from Figure 4, the multiscale 3D residual network has nine layers, including input layer, output layer, fully connected layer, softmax layer, $1 \times 1 \times 1$ convolutional layer, and four layers of 3D residual blocks for multiscale feature fusion. The network imports the hyperspectral sample blocks into the network through the input layer, and then extracts the input image features through two 3D residual blocks with multi-scale feature fusion of channel number 16. And then up-dimensions them through $1 \times 1 \times 1$ convolution of channel number 32, which makes the features of different channels better fused. And this step can increase the nonlinearity of the network model and improve the generalization ability of the network. Finally, the extracted features are fed into the fully connected layer and the classification results are obtained by using the softmax function. In terms of activation function to improve the speed of network operation and the efficiency of gradient transfer, the network selects ReLu as the activation function, which is prone to overfitting phenomenon due to too many training parameters of the network. Therefore, the network adds BN in the 3D residual block of multi-scale feature fusion and dropout (Hinton G E, 2012) in the fully connected layer to avoid the overfitting phenomenon in order to improve the network accuracy and operation efficiency.

3. RESULT AND ANALYSIS

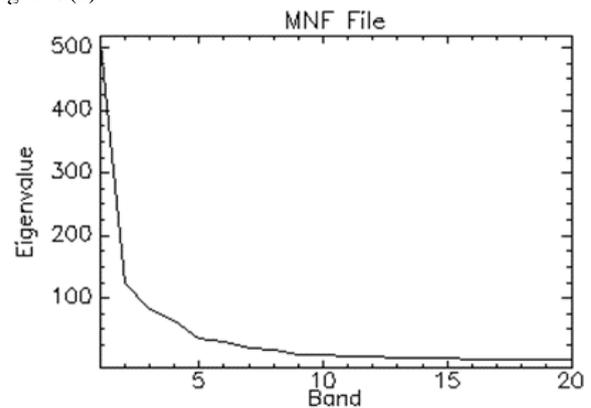
The experiment runs on a 64-bit Win10 system. The development language is Python, and the compiler is Pycharm. The interface provided by TensorFlow 2.0 is used to build the required model and adjust various parameters in the network.

3.1 Data set

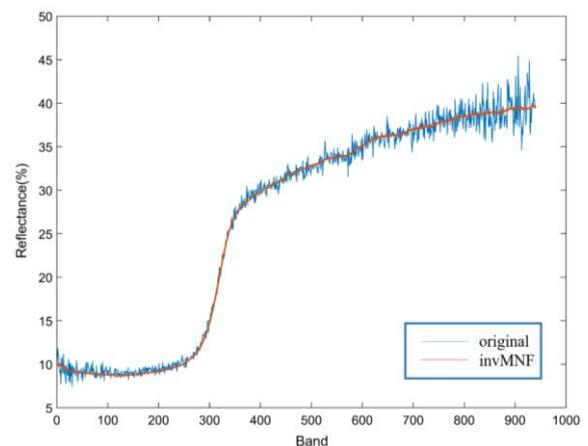
Currently, there is no hyperspectral dataset applicable to the extraction of paint loss of murals. We had made a hyperspectral dataset of murals was produced for the extraction of paint loss by referring to the hyperspectral dataset of remote sensing.

First, the acquired raw hyperspectral images were pre-processed to reduce the noise in the data. It was found that there

was a large amount of noise in the first 50 bands and the last 50 bands of the reflectance hyperspectral images, and the 51st-990th bands were manually selected for MNF transformation, and their eigenvalue distributions are shown in Figure 5(a). The eigenvalue of the 10th component in the figure is already very low, so the first 10 components are selected for MNF inversion transformation. Comparing the spectral curves of the original reflectance image and the MNF inverse transformed image, it can be found that the noise is significantly reduced, as shown in Figure 5(b).



(a) Eigenvalue curve of MNF



(b) Spectral comparison between the original image and invMNF image

Figure 5. Data processing with MNF and invMNF

The preprocessed hyperspectral image has 940 bands, which would introduce many parameters and reduce the training efficiency if it is directly input into the network. Therefore, band averaging is performed on the preprocessed hyperspectral images to reduce the number of bands while ensuring that the image information is not missing as much as possible. The size of the data after dimensionality reduction is $200 \times 200 \times 235$, where 200×200 denotes the spatial dimension and 235 denotes the number of bands.

Then, the hyperspectral images with rich color patterns and paint loss were selected for the dataset production. This part of the mural mainly included green patterns, blue patterns, red patterns, white patterns, black patterns and areas with paint loss. Finally, manual markings were performed. The pixels on the hyperspectral image that cannot be classified are labeled as unclassified. The data set is divided into a training set and a validation set in the ratio of 8:2 for the training and testing of the network.

Combined with the distribution of mural deteriorations, the mural hyperspectral images were cropped and the mural hyperspectral dataset was produced according to the above

method. As shown in Figure 6. Data set sample type and quantity is shown in table 1.

Label id	Area1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Total
Paint loss	2222	0	100	0	1747	3840	530	0	0	8439
White patterns	0	5351	3989	6582	3103	0	2336	3001	7092	31454
Red patterns	0	1232	5	2139	0	0	11	0	1137	4524
Green patterns	6184	0	1578	110	3512	5522	3997	2	2	20907
Black patterns	1594	3417	2865	1147	1320	453	431	193	1152	12572
Blue patterns	0	0	1444	0	0	0	1818	6783	431	10476
Unclassified	0	0	19	22	318	185	877	21	186	1628
Total	10000	10000	10000	10000	10000	10000	10000	10000	10000	90000

Table 1. Data set sample type and quantity

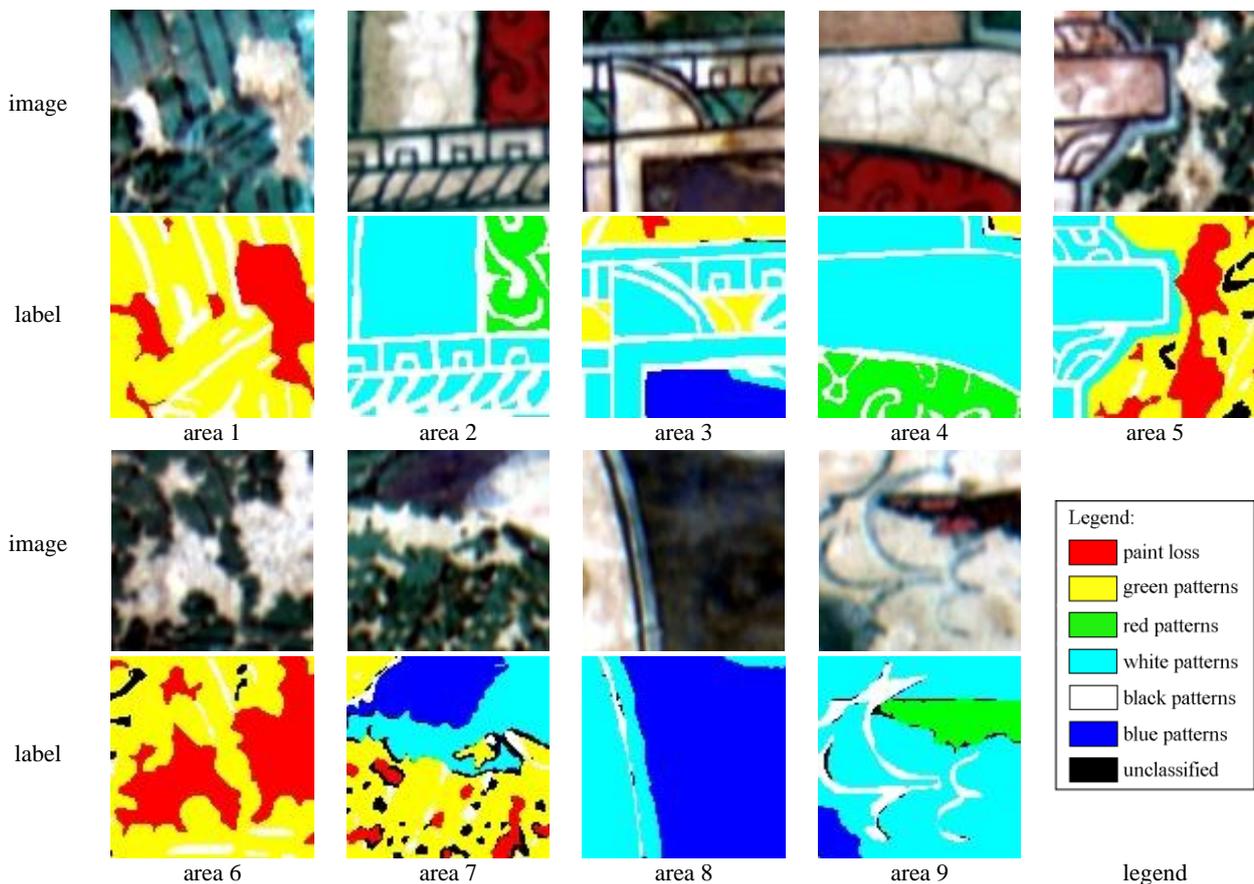


Figure 6. Data sets and corresponding labels

3.2 Network parameters selection

lr	Drop out	batch	epoch	Train set accuracy(%)	Test set accuracy(%)
1×10^{-6}	0.3	16	30	95.33	93.52
1×10^{-6}	0.2	16	30	95.55	93.06
1×10^{-5}	0.3	16	30	98.35	94.50
1×10^{-5}	0.2	16	30	97.84	93.71
1×10^{-4}	0.3	16	30	36.37	36.45
1×10^{-6}	0.3	32	30	75.77	76.50

Table 2. Network accuracy of dataset 1 based on different parameters

The hyperparameters of batch-size, dropout, and learning rate(lr) are selected using the method of grid search. The optimal parameters are determined by comparing the accuracy of the training set with the test set for different parameters. The results for different parameters are shown in Table 2.

Through different combinations of hyperparameters, the learning of 1×10^{-5} , dropout of 0.3, batch of 16 and epoch of 30 were finally selected as hyperparameters, and the accuracy of its training set was 98.35% and the accuracy of its test set was 94.50%.

The optimal parameters selected are input into the network and the network is trained using the produced hyperspectral dataset,

Label id	Paint loss	White patterns	Red patterns	Green patterns	Black patterns	Blue patterns	Total	User's accuracy (%)
Paint loss	5803	10	9	26	13	0	5861	99.01
White patterns	5	8200	21	4	70	9	8309	98.69
Red patterns	4	32	3906	5	33	6	3986	97.99
Green patterns	29	7	9	8339	131	5	8520	97.88
Black patterns	5	129	40	58	9330	8	9570	97.49
Blue patterns	0	50	2	3	15	2852	2922	97.60
Total	5846	8428	3987	8435	9592	2880	39168	
Producer's accuracy (%)	99.26	97.29	97.97	98.86	97.27	99.03		

Table 3. Extraction accuracy of dataset

and the variation of loss value and accuracy is shown in Figure 7. As can be seen from the figure, the classification accuracy keeps rising as the loss value decreases, and the final loss value is 0.048 with an accuracy of 98.35%. After the training model is finished, the validation set is fed into the model, and its classification accuracy is 96.63%. Since the accuracy difference between the training set and the validation set is not large, it proves that the network is well trained, and no overfitting occurs.

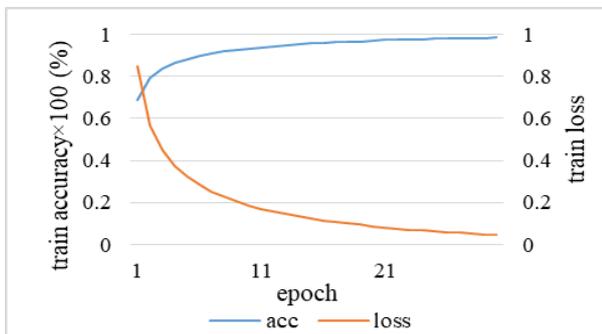


Figure 7 Accuracy and loss value of network training set

Where the blue line is the network training set accuracy and the orange line is the training loss value. As can be seen from the figure, as the network continues to iterate, the loss value of the network gradually decreases, and the training set accuracy keeps improving. When the number of iterations is greater than 20, the network loss value is less than 0.1, and the network training set accuracy is greater than 95%, and when the number of iterations continues to increase, the change of both tends to be smooth. The network was considered to have converged, and the number of iterations was set to 30 in order to prevent the network from overfitting. the prediction model was generated by training, and the test set was input to obtain the prediction results and compared with the manual labeled results, and the overall classification accuracy of the test set was calculated as 94.50% by equation (3). Since the difference between the accuracy of the training set and that of the test set is small, it indicates that the network is in good condition and no overfitting or underfitting occurred.

3.3 Experimental results

The dataset with manual annotation removed is fed into the prediction model to generate the classification results. In order to avoid the unclassified image elements interfering with the accuracy evaluation, so the confusion matrix is calculated without this class of image elements, and the results are shown in Table 3.

From Table 3, it can be calculated that the overall classification accuracy is 98.12% and the Kappa coefficient is 0.98. Since the user accuracy and producer's accuracy of paint loss and white patterns are higher than 97%, it indicates that the network has a strong ability to distinguish paint loss and white patterns with similar spectral characteristics and colour characteristics.

3.4 Comparison tests

In order to compare the extraction ability of different network models and classification methods, two ResNet networks, SSRN networks and the commonly used SVM classification methods were selected and experiments were conducted on the dataset according to the same method, and the results are shown in Figure 8.

Methods	OA(%)	AA(%)	Kappa
M3RN	98.12	98.28	0.98
SSRN	97.02	97.05	0.96
ResNet (with dropout)	84.54	86.20	0.81
ResNet (without dropout)	83.90	84.60	0.80
SVM	83.41	84.48	0.79

Table 4. The accuracy of different extraction methods

By comparison, it was found that SVM had the worst classification results, and the paint loss and white pattern were significantly misclassified. ResNet network had slightly better classification results than SVM, but there was also significant misclassification in details. SSRN and M3RN are both 3D-CNN networks, had the best classification results. For further comparison, the overall accuracy (OA), Kappa coefficient, and average accuracy (AA) of different methods (Zhang Y P, 2020) are shown in Table 4.

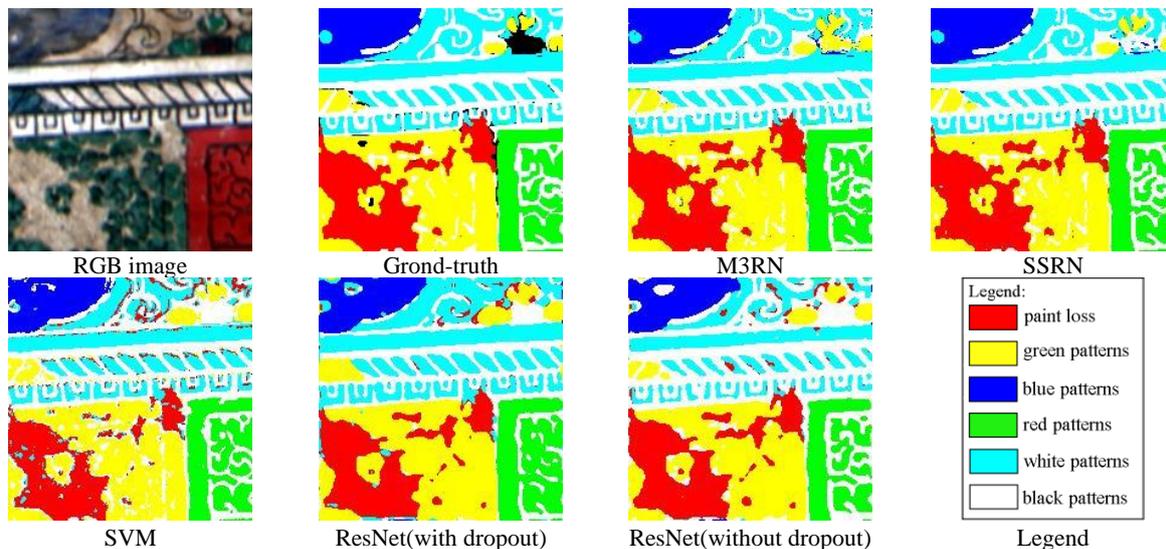


Figure 8. Extraction results using different methods

As shown in Table 4, SVM has the lowest classification accuracy, 2D-CNN network has slightly improved accuracy, and the two networks of 3D-CNN have significantly improved compared with other methods, with overall classification accuracy improved by about 13%, average classification accuracy improved by about 12%, and Kappa coefficient improved by about 0.17, which proves that 3D-CNN is more suitable for hyperspectral data sets. It is proved that M3RN can better extract the spatial and spectral features of hyperspectral data and obtain higher classification accuracy.

4. CONCLUSION

For the problem that it is difficult to distinguish the pigment layer peeling area from the white pattern in the mural painting, a 3D residual network with multi-scale feature fusion is proposed to extract the paint loss disease with higher accuracy. The hyperspectral dataset for paint loss disease extraction was firstly produced using hyperspectral images of the west cloister murals of Qutan Temple in Qinghai Province., Then the extraction of multi-scale spatial features and spectral features was realized using three different 3D convolution kernels of $3 \times 3 \times 3$, $5 \times 5 \times 3$, and $3 \times 3 \times 5$, which were applied to the residual network to determine the optimal parameters by grid search. The experimental results show that the network can improve the efficiency and accuracy of identifying and extracting the paint loss disease of murals. However, similar spectral variability exists for other degradations present in murals, and how to realize the network for the identification and extraction of different deteriorations will be a problem to be solved in future research.

ACKNOWLEDGEMENTS

This research was funded by the National Natural Science Foundation of China (No. 42171356), the Great Wall Scholars Training Program Project of Beijing Municipality Universities (CIT&TCD20180322), and the Fundamental Research Funds for Beijing University of Civil Engineering and Architecture (No.X20088).

REFERENCES

- Cao, J.F., Zhang, Z.B., Zhao, A.D., Cui, H.Y., Zhang, Q., 2020: Ancient mural restoration based on a modified generative adversarial network. *Herit. Sci.*, 8(1), 1-14.
DOI: 10.1186/s40494-020-0355-x
- Li, C.Y., Wang, H.Q., Wu, M., Pan, S.C., 2016: Automatic recognition and virtual restoration of mud spot disease of Tang Dynasty tomb murals image. *Computer Engineering and Applications*, 52(15), 233-236. (In Chinese)
DOI: 10.3778/j.issn.1002-8331.1409-0159
- Zhang, Z.Y., Shui, W.Y., Zhou, M.Q., Xu, B.K., Zhou, H., 2021: Research on disease extraction and inpainting algorithm of digital grotto murals. *Appl. Res. Comput.*, 38(8), 2495-2504. (In Chinese)
DOI: 10.19734/j.issn.1001-3695.2020.09.0395
- Sizyakin, A.R., Cornelis, B., Meeus, L., Dubois, H., Martens, M., Voronin, V., Prizurica, A., 2020: Crack Detection in Paintings Using Convolutional Neural Networks, *IEEE Access*, PP(99), 74535-74552.
DOI: 10.1109/ACCESS.2020.2988856
- Gong, M.T., Feng, P.L., 2014: A Preliminary Study on the Application of Hyperspectral Imaging Technology to the Classification and Identification of Chinese Painting Pigments—Taking Spectral Angle Mapping (SAM) as an Example. *Science of Conservation and Archaeology*. 4(8). (In Chinese)
DOI: 10.3969/j.issn.1005-1538.2014.04.011
- Zhou, P.P., Hou, M.L., Lv, S.Q., Zhao, X.S. and Wu, W.T., 2019: Virtual Restoration of Stained Chinese Paintings Using Patch-Based Color Constrained Poisson Editing with Selected Hyperspectral Feature Bands. *Remote Sens.*, 11(11), 1384.
<https://doi.org/10.3390/rs11111384>
- Li, P., Sun, M.J., Wang Z. and Chai, B.L., 2018: OPTICS-based Unsupervised Method for Flaking Degree Evaluation on the Murals in Mogao Grottoes. *Scientific Reports*, 8(1), 1-10.
DOI: 10.1038/s41598-018-34317-7

- Liu, X.Q., Hou, M.L., Dong, Y. Q., Wang, W.F., Lyu, S.Q., 2019: Extraction and evaluation of the disease of the mural paint loss. *Geomatics World*, 26(5), 22-28. (In Chinese)
DOI: CNKI: SUN: CHRK.0.2019-05-004
- Zhang, Y.P., Zhang, C.M., Bai, 2020: Hyperspectral image classification based on DenseNet-Attention model. *Journal of Graphics*, 41(6), 897-904. (In Chinese)
DOI: 10.11996/JG.j.2095-302X.2020060897
- Meeus, L., Huang, S.G., Devolder, B., Dubois, H., 2019: Deep Learning for Paint Loss Detection with a multiscale, translation invariant network. // *2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA)*.
DOI: 10.1109/ISPA.2019.8868659
- Lyu, S.Q., Wang, S.H., Hou, M.L., Gu, M.Y., Wang, W.F., 2022: Extraction of Mural Paint Loss Disease Based on Improved U-Net. *Geomatics World*, 29(01), 69-74.
- Cao, P.H., Lyu, S. Q., Wang, W. F., Gao, Z. H., Hou, M. L., 2020: Extraction of mural paint loss regions based on spectral dimensionality reduction and Hu moment. *Journal of Graphics*, 41(6), 930-938. (In Chinese)
- Li, F. S., 2020: A Preliminary Study on the Historical and Culture Value of Qutan Temple and its Cultural Relics. *The Silk Road*, (3)4, 169-172. (In Chinese)
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017: ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
<https://doi.org/10.1145/3065386>
- Zoph, B., Vasudevan, V., Shlens, J., Le, Q.V., 2018: Learning Transferable Architectures for Scalable Image Recognition. // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press*, 8697-8710.
DOI: 10.48550/arXiv.1707.07012
- Ji, S. W., Xu, W., Yang, M., Yu, K., 2013: 3D Convolutional Neural Networks for Human Action Recognition. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 35(1), 221-231.
DOI: 10.1109/TPAMI.2012.59
- Zhong, Z.L., Li, J., Luo, Z.M., Li, J., 2017: Spectral-Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2), 847-858.
DOI: 10.1109/TGRS.2017.2755542
- He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J., 2016: Deep residual learning for image recognition. // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press*, 770-778.
DOI: 10.1109/CVPR.2016.90
- Jahandad, Sam, S.M., Kamardin, K., Sjarif, N., Mohamed, N., 2019: Offline Signature Verification Using Deep Learning Convolutional Neural Network (CNN) Architectures GoogLeNet Inception-v1 and Inception-v3. *Procedia Computer Science*, 161, 475-483.
- Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R., 2012: Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors. *Comput. Sci.*, 3 (4), 212-223. <https://doi.org/10.48550/arXiv.1207.0580>