

Application of Image Style Transfer Based on Normalized Residual Network in Art Design

Jing Pu, Yuke Li*

School of Arts and Media, Sichuan Agricultural University, Ya'an, 625014, China

Abstract—With the development of computer vision technology, image style transfer technology based on deep learning has achieved vigorous development. It has been widely applied in fields such as art design, painting creation, and film and television effect production. However, existing image style transfer methods still have shortcomings, including low efficiency and weak quality of style transfer, which cannot better meet the actual needs of various art and design activities. Therefore, a residual network structure is introduced to construct an image style transfer model based on the convolutional neural networks. Meanwhile, a normalization layer is added to the residual network results to optimize the image style transfer technology. An image style transfer model based on the normalized residual network is constructed. The experimental results show that the accuracy, recall, and F1 values of the improved image style transfer model proposed in the study are 97.35%, 96.49%, and 97.52%, respectively, which can complete high-quality image style transfer. This indicates that the image style transfer model proposed in the study has good performance, which can effectively improve the efficiency and quality of image style transfer, providing effective support for various art and design activities.

Keywords—CNN; residual network; normalization; image segmentation transfer; art and design

I. INTRODUCTION

With the development of technology and social progress, various digital media have become an indispensable part of daily life. As a medium for the public to showcase their unique pursuit of individuality, the demand for artistry in images is increasing. Many scholars have begun to explore how to automatically synthesize images with artistic style through computers [1]. Style transfer emerged as the times require. Image style transfer is often regarded as a universal texture generation method for research. Therefore, image style transfer (IST) has become a research highlight. IST technology refers to the process of converting an image into an image that is similar in style or content to the target image through machine learning methods and other techniques [2, 3]. With the progress of deep learning (DL) technology, some scholars build image style migration methods based on the advantages of Convolutional neural network (CNN) feature extraction, while others propose to build image style migration methods using encoder and decoder models [4]. However, existing image style transfer methods have a slower speed during the transfer process. It cannot meet the application requirements in real scenarios. Meanwhile, the comprehensive quality of style images obtained by existing image style transfer methods is not high. In deep learning technology,

convolutional neural networks are feedforward neural networks based on convolutional operations and have deep structures. They are widely used in fields such as image perception, image analysis, image classification, and natural language processing. Therefore, based on the existing CNN structure, an improved image style transfer method is constructed [5]. The structure of CNN is optimized by introducing residual network. At the same time, in the image transfer model, besides adding residual blocks, a normalization layer is added to solve the image quality problems. An image style migration Generative model based on the deep residual network instance normalization (ResNet-IN) is constructed. It is expected to achieve higher quality image style transfer, and improve the speed and efficiency of IST.

The research mainly includes five sections. The Section I is an introduction, briefly introducing the research background. The Section II mainly introduces the research review of residual networks and image generation methods based on deep learning. The Section III constructs an image style transfer method based on a normalized residual network. The Section IV validates the performance of the constructed method. The Section V summarizes the research results and points out the future research direction.

II. RELATED WORKS

A residual network is composed of multiple residual blocks stacked together, which has the advantages of easy optimization and improved accuracy. The internal residual blocks use jump connection, which relieves the vanishing gradient caused by increasing in depth network. Therefore, it is widely used in various fields. Liu et al. conducted recognition research on Boletus images using residual networks. The Fourier transform near-infrared spectral information of the uneven stipe of Boletus edulis was collected. This information was combined into four data matrices to accurately evaluate and identify different Boletus species. According to the findings, the residual neural network (ResNet) has better performance [6]. Liu et al. proposed a data augmentation method to address the issue of limited and imbalanced sample data for transformer faults. A deep residual network with identified paths was introduced to construct a fault diagnosis model, enabling effective transmission and updating of weight parameters. According to the findings, it can effectively expand the data samples with high similarity to the original data. The ResNet has strong feature extraction ability. It enhances the accuracy of fault diagnosis [7]. The existing automatic sleep staging algorithms have too many parameters and long training time, which leads to poor sleep

staging efficiency. Therefore, Yun L et al. proposed an automatic sleep staging algorithm depending on Transfer learning for random deep residual networks (TL-SDResNet) using single channel EEG signals. Experiments have shown that this model can complete rapid training of data. The overall performance is superior to other classic algorithms, with certain practical value [8]. Yousefzadeh et al. introduced an image retrieval system based on the expanded residual CNN of Triplet loss. This system had fewer parameters and provides acceptable accuracy in top-level retrieval images. Experimental evaluation shows that the model can extract more abundant information by expanding the Receptive field, improving the accuracy of image retrieval without increasing the depth or complexity of the model [9]. Siuly et al. proposed a feature extraction method based on deep ResNet to automatically extract representative features from Electroencephalography signal data of schizophrenic patients to identify schizophrenia. This method includes three stages: average filtering for signal preprocessing, deep ResNet for extracting EEG signals, and softmax layer for schizophrenia classification. The research results indicate that it has the ability to discover important biomarkers for automatic diagnosis of schizophrenia from EEG, which will help experts develop computer-aided diagnostic systems [10].

In recent years, image generation and style transformation based on DL have been widely applied. Many breakthroughs have been made. The process of IST has a common idea, which is to analyze images with a certain style. Then corresponding mathematical or statistical models are established for that style. The image to be transmitted is modified to better fit the established model. Wang J combined k-means and semantic closed natural matting algorithms for image segmentation. Neural networks are used for style and content in images. Then through image reconstruction technology, the obtained images are synthesized to achieve the transfer of Folk costume style. The results show that the average runtime, memory usage, and styles generated by this method are 0.06s, 136.06MB, and 1, respectively, reflecting high efficiency and flexibility [11]. Huang et al. proposed an algorithm involving semantic segmentation and residual networks. VGG16 was used for feature extraction to improve the efficiency and generation speed of local IST. Experiments have shown that it is more useful than other commonly used strategies. It can be applied in many specific fields, such as beauty cameras on smartphones, computer-generated images in advertisements and movies, and computer tomography images [12]. Daneshvar et al. proposed a IST method based on neural networks. By manipulating two given camouflage modes, a mixed camouflage mode is generated. The content and style of a CNN trained on image recognition are represented by the correlation between feature maps. The subjective evaluation results indicate that this method is effective in generating successful camouflage patterns [13]. Tian et al. proposed a new layout analysis and style fusion system structure for image generation and style conversion. The Generative adversarial network and multi style ancient book background model were used to train the model. In the experiment, ancient materials such as Yi characters, ancient Chinese, Jurchen characters and ancient paintings were selected as samples for testing. According to the findings, it

improves the performance of the Generative model [14]. Chen et al. constructed a conversion method consisting of multiple convolutional filter banks to better achieve image style conversion. Each filter bank (FB) clearly represents a style. To convert the image into a specific style, the corresponding FB operates on the intermediate features generated by a single automatic encoder. The style library and automatic encoder are learned in a joint manner. The results indicate that this method can effectively achieve image style conversion [15].

In summary, relatively rich research results have been achieved on residual networks and image style transfer techniques. However, in existing research, various image style transfer techniques proposed are mainly designed for specific images, such as clothing, ancient Chinese characters, etc. The practicality cannot meet the needs of various artistic creation activities. At the same time, existing image style transfer technologies have shortcomings in terms of efficiency and quality, and cannot achieve high-quality style image transfer. Furthermore, with the development of computer vision technology, the application demand for image style transfer in various art designs is becoming increasingly widespread. Therefore, based on existing research results and the advantages of residual network structure in image feature processing, a normalized residual network based image style transfer model is constructed. It is expected to further expand the application scope of image style transfer technology through this method, improve the quality and efficiency of image style transfer, and meet a wider range of artistic design needs.

III. IST GENERATION NETWORK DESIGN BASED ON NORMALIZED RESIDUAL NETWORK

IST is crucial for in various art and design activities. Therefore, this chapter will design the corresponding IST model based on CNN. Meanwhile, the normalized ResNet is applied to optimize the image style Generative model to better realize the IST.

A. IST Based on CNN

In the computer vision, IST is a common texture production method, which extracts texture feature information from an image and inputs the extracted texture feature information into the target image to obtain a stylized image. CNN is the most effective machine learning network in image migration processing. CNN is a feedforward neural network based on convolution operation, which is extensively applied in image perception, image analysis, classification and other fields [16]. The CNN is composed of convolution layer, pooling layer and full connection layer. Fig. 1 illustrates the basic structure.

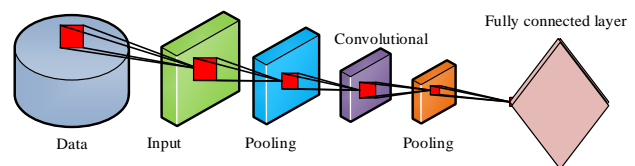


Fig. 1. Structure diagram of CNN.

In the CNN, VGG-Net network structure shows better performance in digital image processing. The research takes VGG-19 as the basic network structure for image style transfer. When the CNN completing the forward propagation task, the convolution core slides in the image matrix. By performing a convolution operation with the pixel values of the convolution kernel coverage position, the corresponding feature map is obtained, as shown in Formula (1).

$$S(i, j) = \sum_m \sum_n I(i+m, j+n)K(m, n) \quad (1)$$

In Formula (1), I stands for the input image of CNN. $S(i, j)$ stands for the coordinate position of the feature image. K represents a convolutional kernel of size $m \times n$. The convolution operation process of the input image is shown in Fig. 2.

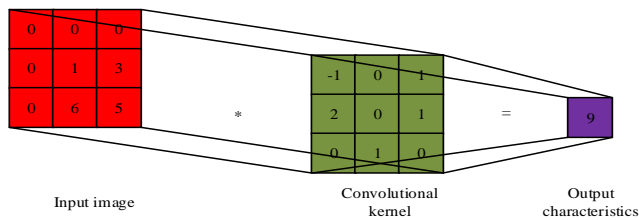


Fig. 2. Schematic diagram of convolution operation.

Firstly, the input image is defined as x . The output image is the expected image after transfer, represented as y . The convolutional layer used in the style transfer process is l . The convolutional features obtained from the input image after calculation in layer l are represented as M_{ij}^l . Among them, i stands for the i -th channel in the convolutional layer. j stands for the j -th position in the convolutional layer. The convolutional features of the image y after style transfer in the l -th layer are represented as F_{ij}^l . Therefore, the obtained image content Loss function (LF) is shown in Formula (2).

$$Loss(x, y, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - M_{ij}^l)^2 \quad (2)$$

The content Loss function can describe the difference between the style of the input image and the output image. The smaller the value, the smaller the differences between the image before and after the style transfer. Then the style features of the image are defined. The image style is represented by Gram matrix (GM), which is a Symmetric matrix composed of several vector inner products, as shown in Formula (3) [17].

$$G_{ij}^l = \sum_k F_{ik}^l F_{jk}^l \quad (3)$$

In Formula (3), G_{ij}^l represents the white noise image matrix. k represents the k -th position in the convolutional layer. The style error after image transfer is shown in Formula (4).

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2 \quad (4)$$

In Formula (4), M_l and N_l represent the height and width of the l -th layer. G^l represents the l -th layer GM of the input image x . The GM of the l -th layer of the migrated output image y is A^l . E_l stands for the GM of image style. Formula (4) has been simplified as shown in Formula (5).

$$Loss_{style}(x, y) = \sum_l \omega_l Loss_{style}(x, y, l) \quad (5)$$

In Formula (5), ω_l stands for the weight of the l -th layer. The content LF and style LF of IST are established through CNN. IST can be seen as the unity of content and style. After copying the target image for style transfer, whether an image is closer in content or more similar in style depends on the proportion of content loss and style loss in the total loss. Thus, the total LF is shown in Formula (6).

$$Loss_{total}(O_{style}, O_{content}) = \alpha Loss_{style}(O_{style}, R) + \beta O_{content}(O_{content}, R) \quad (6)$$

In Formula (6), $O_{content}$ stands for the expected content image. O_{style} stands for the expected style image. R stands for the image that needs to be generated. $Loss_{content}(O_{content}, R)$ represents the content LF. $Loss_{style}(O_{style}, R)$ is the style LF. α and β are hyperparameters used to measure the proportion of style loss and content loss in total loss. The results of IST based on CNN usually cannot preserve the local texture features of the original image. Therefore, there are issues with the representation of depth features and the accuracy of image transfer models when using this method for image style transfer.

B. Image Style Transfer Model Construction Based on Improved Normalized Residual Network

The IST network model, as a type of generative network, should have strong feature learning ability and can achieve accurate and reasonable processing of various image features. In response to the shortcomings of the above image style transfer methods, a residual network structure is introduced to construct an improved normalized residual network (ResNet-IN) image style transfer model. The residual module is added to the basic CNN structure. Deep residual neural network is a meaningful branch in the DL. ResNet can effectively solve the defect of gradient disappearance in the training of CNN by introducing residual blocks. It is widely used in fields such as image classification, object detection, and face recognition [18]. The core of ResNet is residual blocks. In the convolution layers with different depths of the CNN, ResNet introduces a special connection layer for information transmission, that is, skip connections, as shown in Fig. 3.

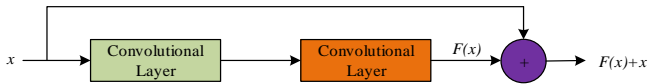


Fig. 3. Schematic diagram of residual block.

Jumping connections can cross a multi-layer network structure to establish a connection between two non-adjacent convolutional layers, thereby achieving information transmission. In Fig. 3, x represents the feature image output from the previous network layer. Then, x is input into the residual block for convolution operation. Finally, the features containing residual information and $F(x)$ are output. The expression of $F(x)$ is shown in Formula (7).

$$F(x) = W_{2\sigma}(W_1x) \quad (7)$$

In Formula (7), W_1 and W_2 stands for the weights of the weight layer, respectively. σ represents the excitation function. The final output of the residual block is represented as $H(x)$, as illustrated in Formula (8).

$$H(x) = F(x) + x \quad (8)$$

When constructing an image style transfer model, it includes a residual network module and a sampling network module. The image undergoes a two-step convolution operation to achieve down-sampling. Then it is sent to the residual network module for processing. On a convolution with a step size of 1/2, it is sampled and restored to the original size output network. The basic process is shown in Fig. 4.

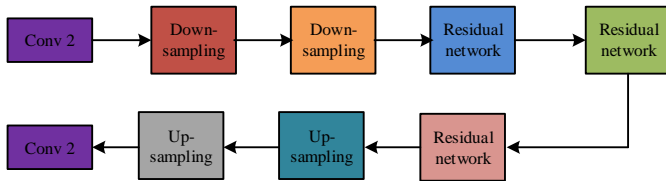


Fig. 4. The basic process of residual block sampling.

In this process, the output part of the network uses the Tanh function to limit the image pixel value between $[0, 255]$. The image style transfer generation network constructed through research is regarded as the objective function $y = f(x)$. x represents the initial image. y stands for the image after style transfer. \hat{y} stands for the output value. $\phi_z(x)$ refers to the value of the Activation function of layer z . In volume z , the output value of the Activation function is a feature mapping of $C_z \times H_z \times W_z$. C_z represents the feature maps. H_z and W_z represent the length and width of the convolutional layer. The Euclidean distance between the feature maps of the objective function to be optimized is shown in Formula (9).

$$l_{feature}^{\phi,z} = \frac{1}{C_z H_z W_z} \|\phi_z(y) - \phi_z(\hat{y})\|_2^2 \quad (9)$$

Currently, the commonly used residual generation network can achieve rapid transfer of image styles. However, the final image migration effect cannot achieve the expected results. Therefore, the normalization idea is adopted to improve the residual generation network. In the image transfer model, in addition to adding residual blocks, a normalization layer is also added. This structure is to solve the contrast problem that occurs during the implementation of style transfer in the initial image. The contrast of the migrated image generated by the IST generation network is less affected by the contrast of the content image, resulting in a higher quality IST effect. Specifically, the IST network includes basic operations such as convolution, redundancy, sampling, and normalization. $x \in \mathbb{R}^{B \times C \times W \times H}$ stands for the image. x_{bcwh} stands for the $bcwh$ -th element. b represents the image serial number. c represents the RGB channel of the image. w and h stands for spatial dimensions. The normalization result of image contrast is shown in Formula (10).

$$y_{bcwh} = \frac{x_{bcwh}}{\sum_{w=1}^W \sum_{h=1}^H x_{bcwh}} \quad (10)$$

However, the basic batch normalization layer defines normalization on all images in the dataset. This normalization method is greatly affected by noise and consumes a lot of storage, resulting in low applicability. Therefore, the Instance Normalization (IN) method is introduced to normalize each sample image one by one. The mathematical definition is shown in Formula (11) [19].

$$\begin{cases} IN(x) = \gamma \left(\frac{x - \mu(x)}{\sigma(x)} \right) + \beta \\ \mu_{bc}(x) = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W x_{bcwh} \\ \sigma_{bc}(x) = \sqrt{\frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W [(x_{bcwh} - \mu_{bc}(x))]^2} \end{cases} \quad (11)$$

The specific process of instance normalization operation is shown in Fig. 5.

Instance normalization is used to replace the task of batch normalization. It can effectively enhance the learning efficiency of the model and avoid various changes in data statistical characteristics. The effect of IST is often evaluated as Mean Intersection over Union (MIOU). The obtained style transfer images are evaluated using the MIOU indicator [20]. MIOU is the standard metric for semantic segmentation, which is typically used to calculate the intersection and union ratio of two sets. In the measurement of IST, two sets are style images and style transfer images, respectively. The calculation method is shown in Formula (12).

$$MIOU = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ij} - P_{ii}} \quad (12)$$

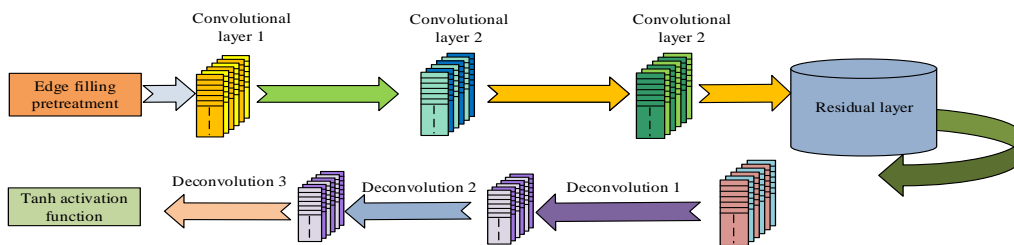


Fig. 5. Image style transfer process based on instance normalization operation.

In Formula (12), k is pixel quantity. p_{ij} represents the intersection to union ratio of each pixel. i stand for the true value. j stand for the prediction. The larger the MIoU is, the smaller the difference between the predicted and the true value. In IST, the closer the style transfer image is to the initial style image.

IV. PERFORMANCE ANALYSIS OF IST GENERATION NETWORK BASED ON NORMALIZED RESNET

A Generative model of IST based on normalized residual network is constructed on the basis of CNN. This chapter will test and verify the performance of the image style Generative model. Then the model is applied to actual IST and the transfer effect is analyzed.

A. Performance Analysis of Improved Normalized Residual Network

To better validate the performance of the image style transfer model based on an improved normalized ResNet, relevant experiments are designed to verify the availability of this method. The environmental design is illustrated in Table I.

TABLE I. TEST ENVIRONMENT SETTINGS

CPU Model	Intel@core i7-7700k CPU@4.2GHz x8
GPU model	GeFore GTX 1080 Ti/PCIe/SSE2
Python version	2.7.13
Tensorflow version	R1.0
CUDA version	8.0
Bazel version	0.7.0
Cudnn version	V5.1
Batchsize	4

Firstly, the availability of the IST based on residual network is verified. The dataset used in the model training process is Microsoft COCO, which contains over 80000 images. 5000 images are randomly selected for model training, and an additional 1500 images are selected as the test set. The image feature data is input into the residual CNN for iterative training. The LF is used at the output end of the model to predict the error. The LF and training accuracy of the proposed method under different Learning rate are illustrated in Fig. 6. From Fig. 6, when the Learning rate is 0.0001, the model is prone to over fitting. When the Learning rate is 0.01, the model precision does not get the optimal solution. When the Learning rate is 0.001, the loss value of the model is 0.2. The corresponding accuracy fluctuation under the Learning rate is the smallest and the accuracy is the best. Therefore, 0.001 is taken as the Learning rate of model training.

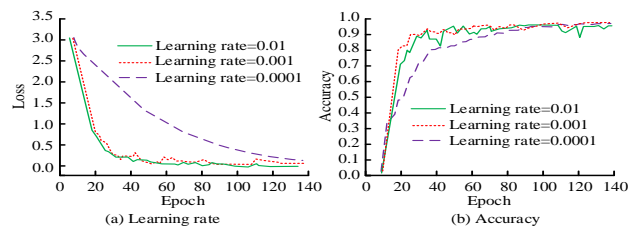


Fig. 6. Loss value and accuracy rate under different learning rate.

To better compare the performance of different methods, the IST method based on improved normalized residual network proposed in the study is compared with commonly used methods. CNN algorithm, based on batch normalization of image style transfer model (CNN-BN), Fast Neural Style Transfer (FNST), and A Neural Algorithm of Artistic Style (NAAS) algorithms are used for comparison [21]. The AUC values of several models were compared on the training and testing sets, as displayed in Fig. 7. In Fig. 7(a), on the training set, the AUC value of the ResNet-IN model reaches 0.985, which is significantly higher than the other three methods. In Fig. 7(b), on the test set, the AUC value of the ResNet-IN model reaches 0.993, which is also superior to the other three comparison methods in terms of performance. According to the AUC values, the ResNet-IN model constructed in the study has better performance.

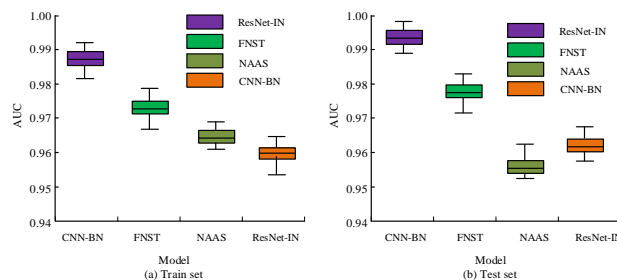


Fig. 7. AUC value of three models.

The accuracy of several common methods is verified. The changes of Loss function and accuracy of different models are obtained as shown in Fig. 8. In Fig. 8, among the common image style transfer methods, the CNN-BN method has the lowest accuracy, only 82.51%. NAAS has an accuracy of 85.69%. FNST has an accuracy of 91.26%. The ResNet-IN method used in the study has an accuracy of 95.73%. At the same time, considering the loss rates of different methods, the ResNet IN has the best loss value. This indicates that the ResNet-IN image style transfer method has the best performance, which can obtain higher accuracy transfer images.

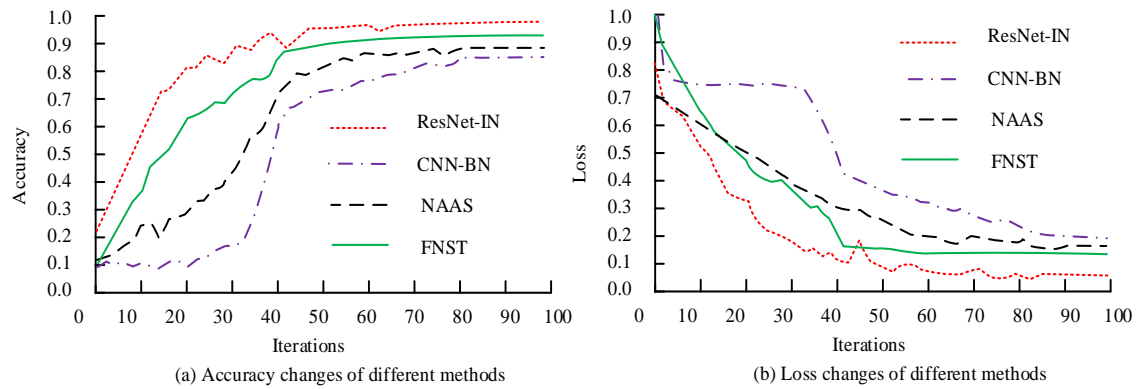


Fig. 8. Changes in accuracy and loss values of different methods.

B. Application Effect Analysis of IST Generation Network based on Normalized ResNet

To explore the transfer effect of the image style transfer model based on normalized residual network proposed in the study on different style images, eight types of images with different styles are selected for five tests. Table II illustrates the experimental results. From Table II, in this testing platform, there are small differences in the time required for each style transfer test for eight different style images. However, overall, the actual time required for image transfer with different styles is less than 3.5s, indicating a better efficiency in image style transfer.

To test the performance of the image style transfer model constructed in the research, a satisfaction survey is conducted on the transfer effects of eight different styles of images. The comprehensive satisfaction evaluation indicators include completeness, fluency, and coherence. The satisfaction score is 10 points. The comprehensive satisfaction results of the expert's evaluation are shown in Fig. 9. From Fig. 9, taking Style 1, Style 3, and Style 8 as examples, the satisfaction results are as follows. The satisfaction rates for fluency, coherence, and integrity of the transfer map in Style 1 are 7.02, 8.56, and 7.64, respectively. The satisfaction rates for fluency, coherence, and integrity of the transfer map in Style 3 are 7.86, 8.62, and 7.49, respectively. The satisfaction rates for fluency, coherence, and integrity of the transfer map in Style 8 are 7.43, 7.74, and 8.35, respectively. The experimental results show that the overall satisfaction of the image style transfer model is high, demonstrating the excellent applicability of the model.

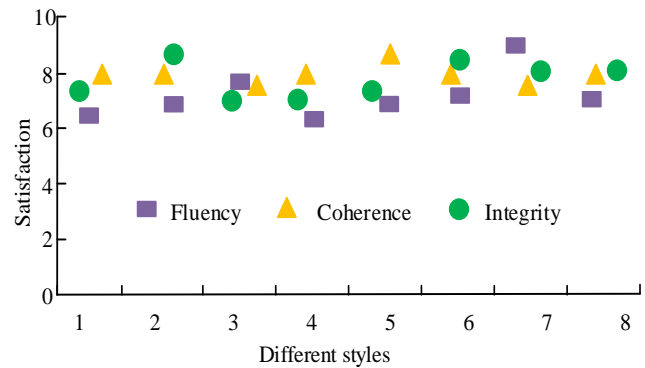


Fig. 9. The satisfaction results of different styles.

Comparing the proposed method with commonly used methods, the MIOU values of the image style transfer method obtained are illustrated in Table III. From Table III, the MIOU value of the CNN based image style transfer method is 0.728, and the image transfer effect of this method is the weakest. The MIOU value of CNN-BN is 0.82, FNST is 0.864, and NAAS is 0.858. The difference between these three methods is relatively small. The MIOU value of the ResNet-IN method is the highest, at 0.9, indicating that the image achieved using this method is closest to the original image, and the difference between the two is the smallest. The image style restoration obtained using this method is the best.

TABLE II. EFFICIENCY COMPARISON OF DIFFERENT STYLES OF IMAGES

Image Style Migration Category	Test Platform	1	2	3	4	5	Average Time of 5 Tests (S)
Style 1	Android	2.03	2.17	2.45	1.86	1.99	2.1
Style 2	Android	1.34	1.25	1.57	1.46	1.42	1.408
Style 3	Android	2.34	2.16	2.75	2.09	2.45	2.358
Style 4	Android	2.64	3.17	2.88	3.23	3.03	2.99
Style 5	Android	3.16	2.98	3.25	3.11	3.23	3.146
Style 6	Android	2.94	2.56	2.73	2.46	2.38	2.614
Style 7	Android	2.45	2.29	2.37	2.58	2.64	2.43
Style 8	Android	2.87	3.15	3.06	3.11	3.24	3.086

TABLE III. MIOU COMPARISON OF IMAGES WITH DIFFERENT STYLES

Model	Number of Tests					Average value
	1	2	3	4	5	
ResNet-IN	0.85	0.89	0.94	0.87	0.95	0.90
CNN	0.64	0.69	0.67	0.85	0.79	0.728
CNN-BN	0.78	0.84	0.85	0.81	0.82	0.82
FNST	0.87	0.86	0.80	0.91	0.88	0.864
NAAS	0.81	0.88	0.84	0.86	0.90	0.858

The performance of several methods in the application process is compared. The results are illustrated in Fig. 10. The accuracy, recall, and F1 values of the ResNet-IN model are 97.35%, 96.49%, and 97.52%, respectively. The three values of NAAS are 95.21%, 94.67%, and 95.23%, respectively. The three values of CNN-BN are 96.38%, 90.77%, and 93.50%, respectively. The three values of FNST are 95.21%, 91.42%, and 93.76%, respectively. The three values of CNN are 90.76%, 91.29%, and 94.08%, respectively. Overall, the ResNet-IN image style transfer method constructed in the study has more advantages in the application process, which can more accurately achieve IST and ensure the quality of style images.

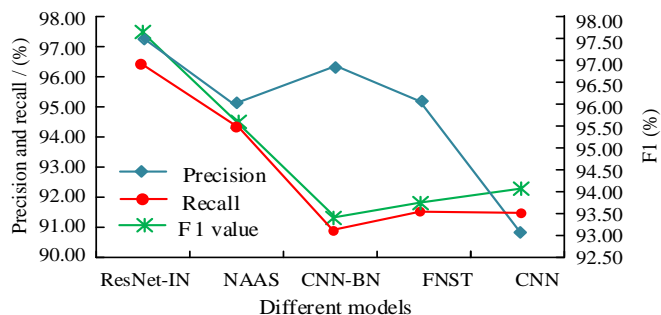


Fig. 10. Accuracy, recall rate and F-measure value of the three models.

V. CONCLUSION

Image style transfer can transfer the style features of one painting to another portrait. With the development of artificial intelligence technology, IST technology based on DL has been widely developed. However, existing image style transfer technologies still have issues of low efficiency and low image quality. Therefore, based on the existing CNN, the ResNet structure is introduced to optimize it. At the same time, the normalization idea is used to improve the contrast problem that occurs during image transfer. The IST network Generative model based on the normalized ResNet is constructed. According to the findings, the AUC values of the ResNet-IN model proposed in the study are 0.985 and 0.993 in the training and testing sets, significantly higher than those of commonly used methods. Among several common image style transfer methods, the CNN-BN method has the lowest accuracy, only 82.51%. The accuracy of NAAS is 85.69%. The accuracy of FNST is 91.26%. The accuracy of the ResNet-IN method used in the study was 95.73%. The ResNet-IN image style transfer method proposed in the study has the best performance and can obtain higher accuracy transfer images. From the perspective of image restoration obtained by different style transfer methods, the proposed image style transfer method has the highest MIOU value of 0.9.

The style transfer image achieved using this method is closest to the original image, with the smallest difference between the two. The image style restoration obtained using this method is the best. Through testing eight different style images, there are small differences in the time required for each style transfer test. Overall, the actual time required for image transfer with different styles is less than 3.5 seconds, indicating a better efficiency in IST. In summary, the IST model proposed in the study can significantly improve the efficiency and quality of image style transfer, provide support for art and design activities represented by painting, and meet the needs of image style transfer in more practical scenarios. However, there are still shortcomings in the research. There are shortcomings in the flexibility of IST. Therefore, in future research, further research is needed on the flexibility of IST. However, there are still shortcomings in the research. The existing image style transfer methods are mainly influenced by factors such as algorithm execution speed, flexibility, and image quality. Although the proposed method has optimized the execution speed and the quality of image extraction, it still has shortcomings in the continuity of image style features. In addition, this model can only complete the transfer of similar image styles after each completion, which has shortcomings in flexibility methods. Therefore, in future research, based on the continuity of image feature extraction and the flexibility of image conversion, the image style transfer ability will be further optimized.

REFERENCES

- [1] Vasavi S, Krishna K, Raman S V. Gender prediction using ensemble based wide residual network from surveillance video. *Microsystem Technologies*, 2022, 29(4): 527-537.
- [2] Arthy P S, Kavitha A. Deep Capsule Residual Networks for Better Diagnosis Rate in Medical Noisy Images. *Intelligent Automation & Soft Computing*, 2022, 36(2): 1381-1393.
- [3] Imani M. Spatial and polarimetric information fusion using residual network for polarimetric synthetic aperture radar image classification. *IET Radar, Sonar & Navigation*, 2022, 16(12): 1963-1976.
- [4] Wu Z Q, Xie B. Fine-Grained Pornographic Image Recognition with Multi-Instance Learning. *Computer Systems Science and Engineering*, 2023, 47(1): 299-316.
- [5] Liu D H, Zhao J, Xi A, Wang C, Huang X N, Lai K C, Liu C. Data Augmentation Technology Driven By Image Style Transferin Self-Driving Car Based on End-to-End Learning. *Computer Modeling in Engineering & Sciences*, 2020, 122(2): 593-617.
- [6] Liu H, Liu H G, Li J Q, Wang Y Z. Rapid and Accurate Authentication of Porcini Mushroom Species Using Fourier Transform Near-Infrared Spectra Combined with Machine Learning and Chemometrics. *ACS omega*, 2023, 8(22): 19663-19673.
- [7] Liu X Z, Xie J, Luo Y H, Yang D S. A novel power transformer fault diagnosis method based on data augmentation for KPCA and deep residual network. *Energy Reports*, 2023, 9(8): 620-627.
- [8] Yun L, Zhang X X, Zheng Y C, Wang D H, Hua L Z. Enhance the Accuracy of Landslide Detection in UAV Images Using an Improved Mask R-CNN Model: A Case Study of Sanming, China. *Sensors (Basel, Switzerland)*, 2023, 23(9): 286-294.
- [9] Yousefzadeh S, Pourreza H, Mahyar H. A Triplet-loss Dilated Residual Network for High-Resolution Representation Learning in Image Retrieval. *Journal of Signal Processing Systems*, 2023, 95(4): 529-541.
- [10] Siuly S, Guo Y H, Alcin O F, Li Y, Wen P, Wang H. Exploring deep residual network based features for automatic schizophrenia detection from EEG. *Physical and engineering sciences in medicine*, 2023, 46(2): 561-574.

- [11] Wang J. Garment image style transfer based on deep learning. *Journal of Intelligent & Fuzzy Systems*, 2023, 44(3): 3973-3986.
- [12] Huang L, Wang P, Yang C F, Tseng H W. Rapid Local Image Style Transfer Method Based on Residual Convolutional Neural Network. *Sensors and materials: An International Journal on Sensor Technology*, 2021, 33(4):1343-1352.
- [13] Daneshvar E, Tehran M A, Zhang Y J. Hybrid camouflage pattern generation using neural style transfer method. *Color Research and Application*, 2022, 47(4): 878-891.
- [14] Tian Y, Chen S, Zhao F, Lin X, Xiong H. The Layout Analysis of Handwriting Characters and the Fusion of Multi-Style Ancient Books' Background. *Journal of Computer-Aided Design & Computer Graphics*, 2020, 32(7): 1111-1120.
- [15] Chen D, Yuan L, Liao J, Yu N H, Hua G. Explicit Filterbank Learning for Neural Image Style Transfer and Image Processing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(7): 2373-2387.
- [16] Hu M, Rutqvist J, Steefel C I. Mesh generation and optimization from digital rock fractures based on neural style transfer. *Journal of Rock Mechanics and Geotechnical Engineering*, 2021, 13(4):912-919.
- [17] Chen H, Han Q, Li Q, Tong X J. Image steganalysis with multi-scale residual network. *Multimedia Tools and Applications*, 2022, 82(14):22009-22031.
- [18] Wang L, Wang L, Chen S. ESA-CycleGAN: Edge feature and self-attention based cycle-consistent generative adversarial network for style transfer. *IET Image Processing*, 2022, 16(1):176-190.
- [19] Wang X, Cheng M, Eaton J, et al. Fake node attacks on graph convolutional networks. *Journal of Computational and Cognitive Engineering*, 2022, 1(4): 165-173.
- [20] Sun Y H, Liu Y. Model-data-driven P-wave impedance inversion using ResNets and the normalized zero-lag cross-correlation objective function. *Petroleum Science*, 2022, 19(6): 2711-2719.
- [21] Iseringhausen J, Weinmann M, Huang W, Hullin M B. Computational Parquetry: Fabricated Style Transfer with Wood Pixels. *ACM Transactions on Graphics*, 2020, 39(2): 1-14.