ISSN 2616-5775 Vol. 4, Issue 8: 83-89, DOI: 10.25236/AJCIS.2021.040816

Research on Image Style Convolution Neural Network Migration Based on Deep Hybrid Generation Model

Junhu Zhou[a], Yujie Wang[b], Jiafang Gong[c], Guoqing Dong[d], Wentao Ma[e]

School of IMUT, Inner Mongolia University of Technology, Hohhot, Inner Mongolia, 010051, China
[a]751013199@qq.com, [b]961703309@qq.com, [c]825349579@qq.com, [d]1767558864@qq.com, [e]7293470@qq.com

Abstract: *The main content of image style transfer transforms the image style from one region to another. This task puts forward new needs for the traditional convolutional neural network architecture. Therefore, a deep hybrid generation model is usually used in the study of processing image style transfer. Image style transfer aims to transform the image into a new idea by image generation. This paper proposes an image-style convolution neural network migration model based on the deep mixing generation model based on background. The image quality is improved through image processing. The deep hybrid generation model mainly relies on to combine confrontation network generation and self-encoder. In this paper, unsupervised and supervised image style migrations are designed according to the different basic tasks of image style migration. On this basis, unsupervised image style migration of combative neural networks based on cyclic consistency and supervised image style migration of adversarial networks based on cross-domain self-encoder are proposed. This paper further improves the quality of created images by introducing an unsupervised and supervised image style migration standard dataset.*

Keywords: *Deep learning; Image style transfer; Convolutional neural network*

1. Introduction

In recent years, convolutional neural networks have also been applied in image recognition and image change-over. Based on this research, this is famous for researching the creative ability of convolutional neural networks. We can produce images through describe an image. With the image style migration, the graphics are improved from the details.

The main content of image style transfer is to transform the image style of one region to another. This task puts forward new needs for the traditional convolutional neural network architecture. Therefore, a deep hybrid generation model is usually used in the everyday study of processing image style transfer. The goal of image style transfer is to transform the image into a new idea.

In this paper, unsupervised and supervised image style migrations are designed according to the different basic tasks of image style migration. On this basis, unsupervised image style migration of combative neural networks based on cyclic consistency and supervised image style migration of combative networks based on cross-domain self-encoder are proposed. This paper further improves the quality of produced images by introducing an unsupervised and supervised image style migration standard dataset.

2. Review of Image Style Transfer

2.1 Review of Image Style Transfer Technology

In the informatization worldwide, data in society has reached a high-level. The generation of various data plays an incisive role in developing machine learning and deep learning technology. Deep learning the tacit information from the data can build its neural network and improve computational efficiency. In the research of making the neural network algorithm work faster, many excellent algorithms have been born. [1] Among them, the convolutional neural network uses the location information characteristics and distribution patterns between image pixels to extract features, to carry out image recognition and

ISSN 2616-5775 Vol. 4, Issue 8: 83-89, DOI: 10.25236/AJCIS.2021.040816

optimization. On this basis, the image quality can be improved, the image style can be adjusted, and the image style migration can be carried out. The schematic diagram of this work is shown in Fig.1.

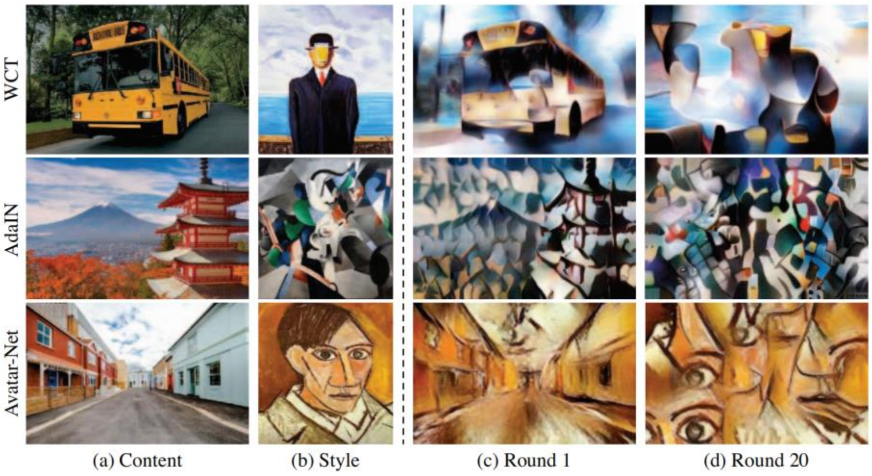

Fig.1 Image Style Transfer Diagram

2.2 Image Style Transfer Model

2.2.1 Convolutional Neural Network Model

As a classical image processing algorithm in deep learning, a convolution neural network has been widely used in image processing and recognition. In this process, a convolution neural network requires fewer parameters than other neural networks and can directly enter the original image to perform corresponding processing operations. Therefore, it has been widely and deeply applied in this field.

As a classical graphical algorithm neural network, VGG neural network is used for feature extraction and migration learning. In this, the neural network completes related operations through the hierarchical algorithm. The whole learning of the neural network is usually from forward to reverse. First, set the wanted and store it as input mode. Then sort out the training sample information and inject it into the neural network in the form of table data. This is positive. The difference between the and expected in the learning process is calculated and recorded as the error value. When the expected value does not match the real deal, the network will apply the principle of the minimum square error to replace the data and enter the reverse process. In the forward and reverse processes, the responses are weighted according to the continuous feedback process of error values, and the path weights of corresponding connection paths are correct. The actual output of the network gradually approaches the preset expectation. Output the connection weights between layers through the above learning process. For an input sample, an output is obtained by forwarding the network's reasoning, and then it is compared with the expected output sample. If there is a deviation, propagate back from the output and adjust the weight coefficient Wij. Let X be the input sample, Y be the output sample, T be the expected output sample, η be the learning rate (a positive number less than 1), f (x) be the activation function of the network, S-shaped curve is selected, and Wij is the weight coefficient of the connection between the I unit and the j unit. When forward propagation is from the input layer to the output, the output of the upper layer is the input of the next layer. [2]

The logic of this algorithm is as follows. Firstly, the image is input, and the feature extraction of the image is carried out. On this basis, the image style migration is achieved by exporting the image category. The total loss function is as follows :

$$L_{\text{total}}(\vec{p},\vec{a},\vec{x}) = \alpha L_{\text{content}}(\vec{p},\vec{x}) + \beta L_{\text{style}}(\vec{a},\vec{x}) \quad (1)$$

Where α represents the weight coefficient of image content loss function $L_{\text{content}}(\vec{p},\vec{x})$, β represents the weight coefficient of the image style loss function $L_{\text{style}}(\vec{a},\vec{x})$, α and β can adjust the proportion of content and style. The content style loss function $L_{\text{content}}(\vec{p},\vec{x})$ of the image is shown as:

ISSN 2616-5775 Vol. 4, Issue 8: 83-89, DOI: 10.25236/AJCIS.2021.040816

$$L_{\text{content}}(\vec{p},\vec{x},l)=\frac{1}{2}\sum_{i,j}\left(F_{ij}^{l}-P_{ij}^{l}\right)^{2} \tag{2}$$

Where F_{ij}^{l} represents the activation value of the content image at the position j at the i convolution layer in the first block of model VGG ,the original paper selects the output characteristics of the fourth layer and the second convolution layer, P_{ij}^{l} is the content feature representation of noise images in VGG the network. The above algorithm strategy process is shown in Fig. 2.

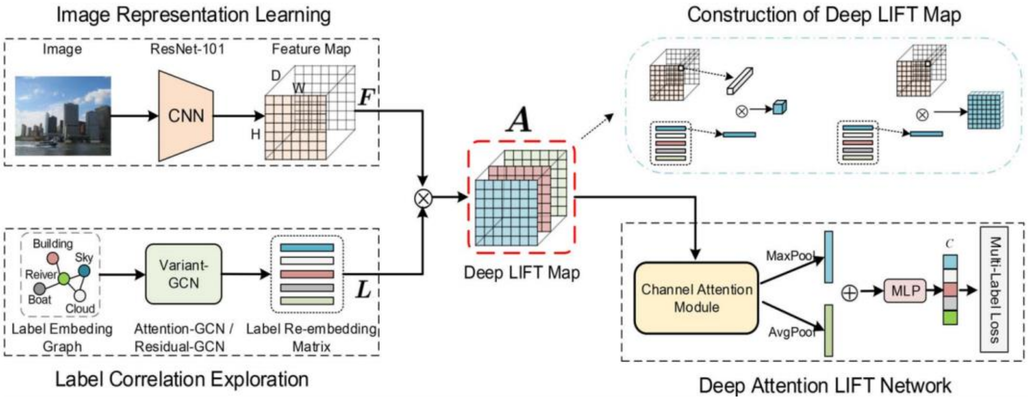

Fig. 2 Basic Architecture Diagram of Convolutional Neural Network

2.2.2 Generation of Adversarial Networks

The adversarial network generation process is as follows: by inputting the image, after being generated by the generator, the discriminator is used to discriminate the true and false points of the picture to train the neural network. In this process, the neural network uses the self-encoder, but this kind of self-encoder is unique, and its encoder and decoder are symmetrical.

The loss function of the Pix2pixGAN generative adversarial network section is :

$$L_{cGAN}(G,D)=E_{x,y}[\log \mathrm{D}(\mathrm{x},\mathrm{y})]+E_{x,z}[\log(1-\mathrm{D}(\mathrm{x},\mathrm{G}(\mathrm{x},\mathrm{z}))] \tag{3}$$

Information will be shared between the input and output of the supervised migration generator, so the L1 loss is added to ensure the similarity between the input image and the result.

$$L_{L1}(G)=E_{x,y,z}\left[\| y-G(x,z)\|_{1}\right] \tag{4}$$

The overall optimization objective is :

$$G^{*}=\arg\min_{G}\max_{D} L_{cGAN}(G,D)+\lambda L_{L1}(G) \tag{5}$$

Based on the above process, the self-encoder confrontation network can be generated, and the strategy diagram is shown in Fig. 3.

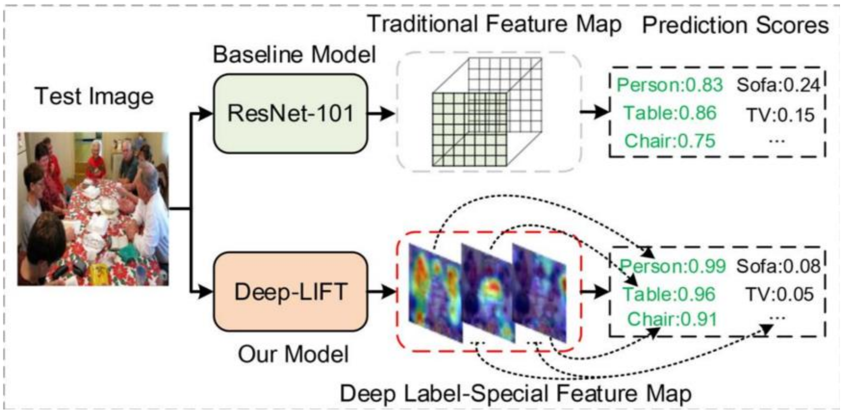

Fig 3 Schematic Diagram of Learning Strategies for Adversarial Neural Networks

ISSN 2616-5775 Vol. 4, Issue 8: 83-89, DOI: 10.25236/AJCIS.2021.040816

2.2.3 Basic Network Structure

The traditional neural network architecture usually stacks the neural network layer, which generally realizes the relevant functions, producing degradation problems. In this case, with the deepening of the number of network layers, the overall architecture of the neural network will create a systematic gradient disappearance problem. In this process, the neural network forms a gradient multiplier through the reverse propagation process, which leads to the propagation gradient becoming smaller and smaller, and finally close to zero. [3]

Therefore, the neural network used in the current graphics recognition project has introduced the architecture of the residual network. By submitting the residual block and using the jump connection strategy, the information lost in the above backpropagation process is reintroduced into the system through the jump connection, and then the image is normalized. Style migration, the process is shown in Figure 4.

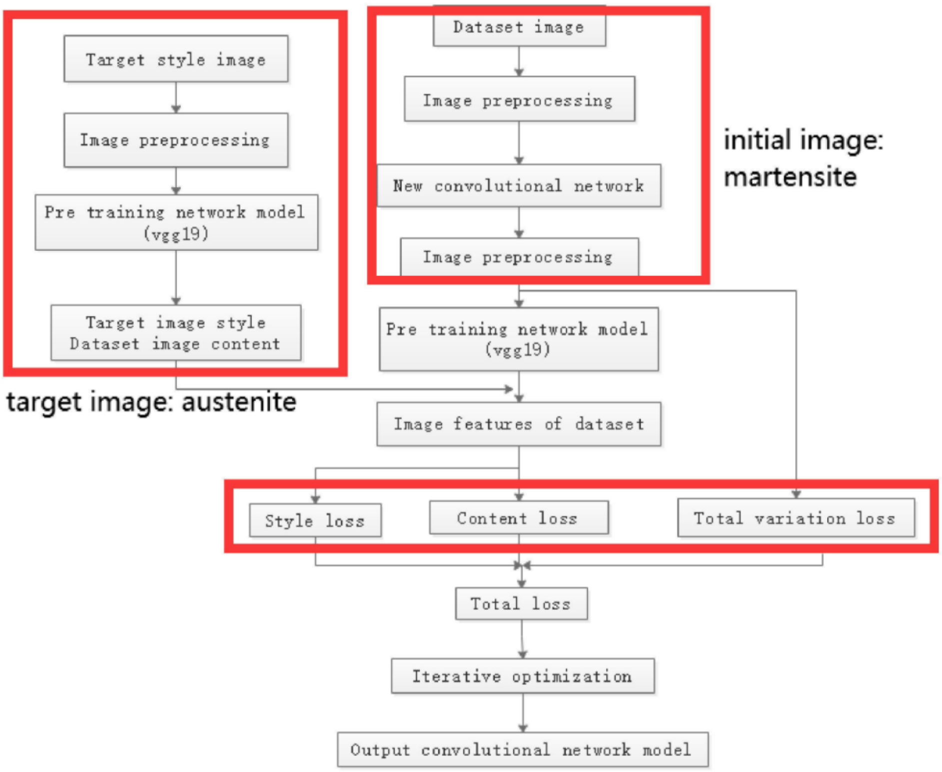

Fig. 4 Schematic Diagram of Basic Network Structure

3. Unsupervised Image Style Migration Based on Convolutional Neural Network

3.1 Generation of Confrontation Network Model Based on Cyclic Consistency

The generation of an adversarial network model based on cyclic consistency is realized by bidirectional mapping. This part usually uses two generators to work together and then judges whether the discriminator distorts the output. The principle is shown in Figure 5.

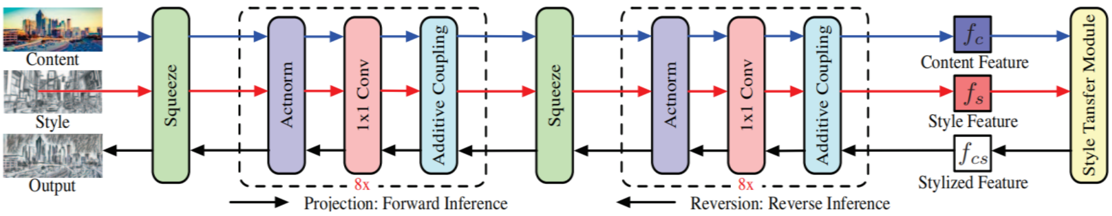

Fig.5 Schematic Generation of Adversarial Network Model Based on Cyclic Consistency

ISSN 2616-5775 Vol. 4, Issue 8: 83-89, DOI: 10.25236/AJCIS.2021.040816

3.2 The Compilation Process of Unsupervised Loss Function

The unsupervised loss function is mainly reflected in three parts. The first is to generate confrontation network loss, the second is circular consistency loss, and the third is perception loss. The loss function reflects these three parts.[4] The optimal generator and discriminator are obtained by optimizing the loss function concept to complete the training process of the neural network. When used, the output results after the conversion style can be obtained by inputting the image of the corresponding domain. The schematic diagram is shown in Figure 6.

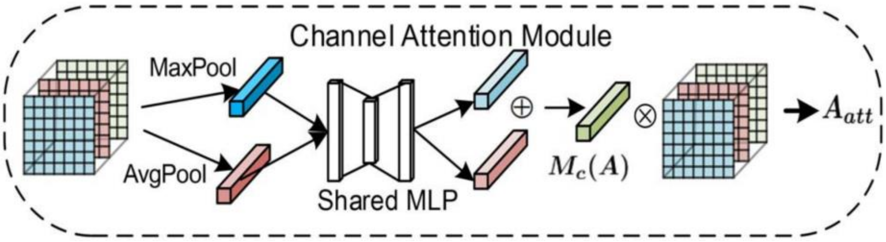

Fig.6 Unsupervised Loss Role Preparation Diagram

3.3 Unsupervised Quantization Model Generation and Dataset Invocation

Complete the call by finding the Horse2zebra dataset in the primary dataset. The calling is. First, the test training is carried out through the training set. After completing the training, the image style training of the neural network is carried out. The above process is repeated through the dual-loop network architecture. The process is shown in Figure 7. The effect diagram is shown in Fig. 8.

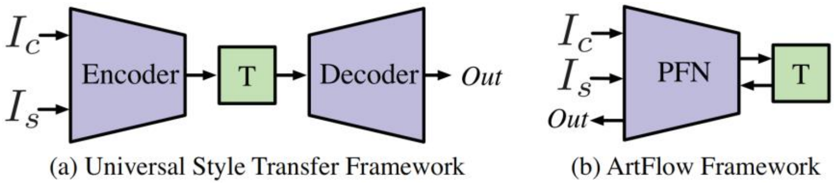

Fig. 7 Unsupervised Quantization Model Generation and Dataset Invocation Diagram

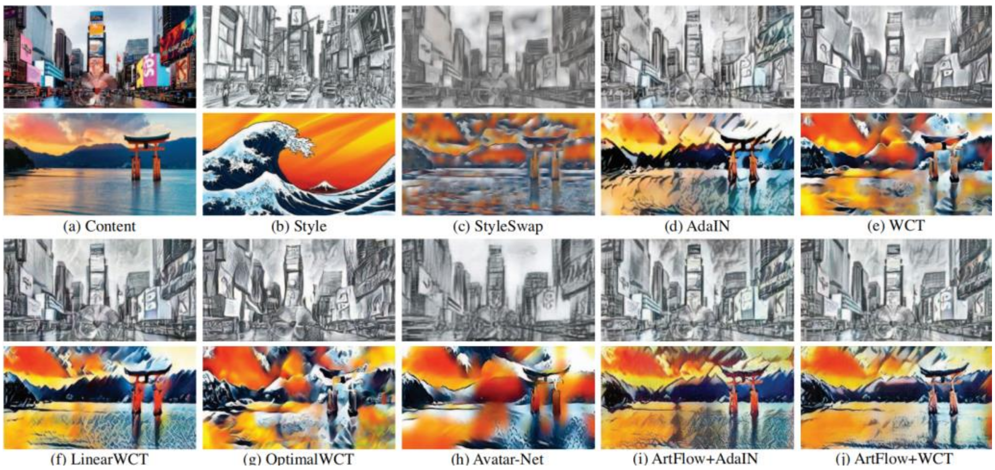

Fig. 8 Unsupervised Quantization Model Generation and Data Set Call Effect Diagram

4. Supervised Image Style Migration Based on Convolutional Neural Network

4.1 Cross Domain Autoencoder Generative Adversarial Network Model

The cross-domain self-encoder generates a confrontation network model to realize the mutual conversion of two corresponding image domains through two mappings. In the process of image coloring, the image is generally greyed out, and then the color information is encoded by the neural network and finally re-compiled for output. [5] The principle is shown in Figure 9.

ISSN 2616-5775 Vol. 4, Issue 8: 83-89, DOI: 10.25236/AJCIS.2021.040816

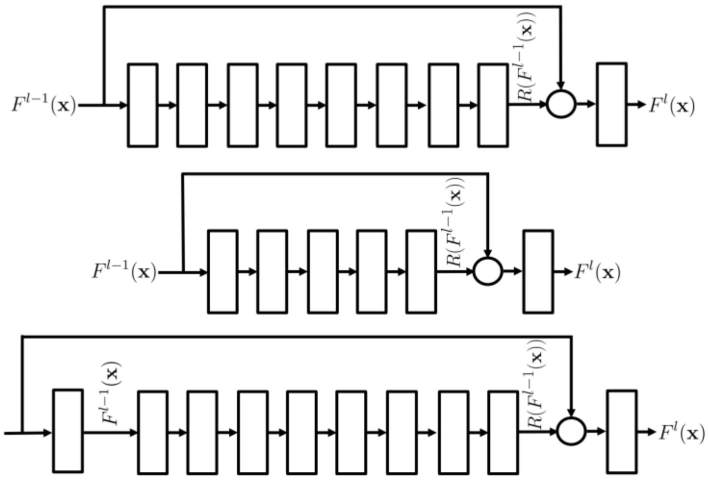

Fig.9 Schematic Diagram of Cross-domain Autoencoder Generation Countermeasure Network Model

4.2 The Compilation Process of Supervised Loss Function

The compilation process of the supervised loss function is mainly composed of three parts of loss function first is to fight against network loss, the second is to reconstruct loss, and the third is to target loss in the process of style transfer. The process of each part is introduced as follows :

Adversarial network loss: the least square method is used to generate the hostile network. Then, the loss function of the least honest way is used to replace the original loss function based on the original adversarial network to generate higher-quality images. This training process is relatively stable. [6]

Optimization process: The optimal cross-domain autoencoder output is achieved by formulating objective constraints for the loss function. In the actual application process, the corresponding high-quality images can be obtained by inputting the target image. The schematic diagram is shown in Fig. 10. [7]

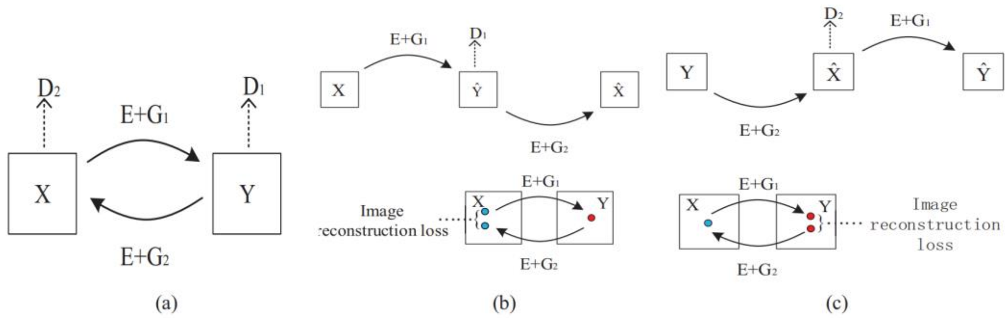

Fig. 10 Diagram of Supervised Loss Function Compiling Process

4.3 Supervised Quantization Model Generation and Dataset Invocation

Supervised quantization model generation and data set call are realized through the model network structure. A cross-domain autoencoder usually implements this process. Each encoder has eight convolution layers unique to convolutional neural networks, and the decoder has seven deconvolution layers. The neural network learning process is carried out by random gradient descent, and the optimization method is realized by momentum optimization based on Adam. A total of 100 epochs are trained. The initial learning rate of the first 50 cycles is set to 0.005, and the learning rate of the last 50 cycles is halved. Normalization in this experiment also uses sample normalization. [8]

ISSN 2616-5775 Vol. 4, Issue 8: 83-89, DOI: 10.25236/AJCIS.2021.040816

(a) Input (b) Round 1 (c) Round 20 (d) Round 40

5. Conclusion

With the high development of information technology, deep learning convolutional neural network has been highly developed in many social, digital data learning processes. The image style migration required for image quality improvement has gradually become feasible. The main content of image style transfer is to transform the image style of one region to that of another area. This task puts forward new requirements for the current traditional convolutional neural network architecture. Therefore, the depth hybrid generation model is usually used in the everyday study of processing image style migration. Image style migration aims to transform the input image into a new idea through image generation. This paper proposes an image-style convolution neural network migration model based on the deep mixing generation model based on the above background. The image quality is improved through the above image processing. The deep hybrid generation model mainly relies on the combination of confrontation network generation and self-encoder. In this paper, unsupervised and supervised image style migrations are designed according to the different basic tasks of image style migration. On this basis, unsupervised image style migration of adversarial neural networks based on cyclic consistency and supervised image style migration of adversarial networks based on cross-domain self-encoder are proposed. This paper further improves the quality of generated images by introducing an unsupervised and supervised image style migration standard dataset.

References

[1] Z. Liu et al., "Remove Appearance Shift for Ultrasound Image Segmentation via Fast and Universal Style Transfer," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 2020, pp. 1824-1828.

[2] H. Ye, W. Liu and Y. Liu, "Image Style Transfer Method Based on Improved Style Loss Function," 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 2020, pp. 410-413.

[3] P. Li, L. Zhao, D. Xu, and D. Lu, "Incorporating Multiscale Contextual Loss for Image Style Transfer," 2018 IEEE 3rd International Conference on Image, Vision, and Computing (ICIVC), Chongqing, China, 2018, pp. 241-245.

[4] S. Chelaramani, A. Jha and A. M. Namboodiri, "Cross-Modal Style Transfer," 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 2018, pp. 2157-2161.

[5] Y. Zhang, Y. Zhang and W. Cai, "A Unified Framework for Generalizable Style Transfer: Style and Content Separation," in IEEE Transactions on Image Processing, vol. 29, pp. 4085-4098, 2020.

[6] Z. Lin et al., "Image Style Transfer Algorithm Based on Semantic Segmentation," in IEEE Access, vol. 9, pp. 54518-54529, 2021.

[7] J. J. Virtusio, A. Talavera, D. S. Tan, K. Hua, and A. Azcarraga, "Interactive Style Transfer: Towards Styling User-Specified Object," 2018 IEEE Visual Communications and Image Processing (VCIP), Taichung, Taiwan, 2018, pp. 1-4.

[8] Y. He, J. Li and A. Zhu, "Text-Based Image Style Transfer and Synthesis," 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW), Sydney, NSW, Australia, 2019, pp. 43-48.