Style transfer for converting images into Chinese landscape painting based on CycleGAN

Bowen Deng

The Department of School of Software, Northwestern Polytechnical University, Xi'an, 710072, China

2020300849@mail.nwpu.edu.cn

Abstract. Chinese landscape painting is a prevalent and unique mode of artistic expression within traditional Chinese art. It boasts intricate techniques and demands a relatively high level of artistic skill. Recent advancements in artificial intelligence have ushered in the era of image style transfer techniques, making it feasible to convert landscape photographs into stunning Chinese landscape paintings. In this study, the author has developed an image translation model that enables mutual transformation between images of landscapes and Chinese landscape paintings. This technology significantly reduces the difficulty of creating Chinese landscape paintings, allowing more ordinary people in China to experience and appreciate the joy brought by traditional Chinese art. Experimental results indicate that the style transfer model based on CycleGAN has achieved significant success in this scenario. The generated artworks successfully integrate the style of Chinese landscape painting into the original images without altering the original composition and details. As a result, the original photos gain a certain level of artistic value. Additionally, this study innovatively explores the goal of reverse restoration of Chinese landscape paintings into images, highlighting both the similarities and differences with the current research, thus laying the foundation for future studies.

Keywords: Chinese Landscape Painting, Image Translation, CycleGAN.

1. Introduction

Chinese landscape painting is a traditional Chinese art form depicting natural landscapes, including mountains and rivers. It originated between 220 and 420 AD and has since become an important genre within Chinese painting [1]. Chinese landscape painting embodies the profound sentiments of the Chinese people, the cultural awareness of exploring mountains and enjoying water, the inner cultivation of associating virtue with mountains and character with water, and the perception of vastness within the span of a foot have always been the central themes in the interpretation of landscape painting. Chinese landscape painting not only aims to depict what the artist sees but also to convey the spiritual values within the artwork. Therefore, crafting Chinese landscape paintings is certainly not a straightforward endeavour. This indirectly affects the level of attention and enthusiasm that have towards learning to craft Chinese landscape painting for contemporary Chinese people.

Image translation aims to transform source domain images into target domain images through end-to-end models [2-4]. This process achieves the stylization of the source domain images under the style of the target domain, thus facilitating the transformation from source domain images to target domain

images. Image translation technology has witnessed numerous milestones in the past few years. In 2006, software called "The Painting Fool" emerged, which could observe photos, extract color information from them, and create art using real-world materials such as paint, pastels, or pencils. A breakthrough in Image Translation tasks is the rapid development of artificial intelligence technologies, particularly the advent of a series of network models represented by Generative Adversarial Networks (GANs). GANs were proposed by Goodfellow in 2014 as a type of adversarial network [5]. This network framework consists of two parts: a generative model and a discriminative model. Both models continually improve their abilities through learning. The generative model aims to generate more convincing fake data to deceive the discriminative model, while the discriminative model aims to learn how to accurately identify fake data generated by the generative model. In 2017, Efros et al introduced CycleGAN, which employs two GAN models to learn transformation functions between two domains [6]. This establishes a relationship between the two domains. The emergence of CycleGAN means that images can be transferred from one type to another. CycleGAN can effectively convert image styles to match those of artists like Monet and Van Gogh, showcasing notable success in this aspect.

In this paper, the author will explore CycleGAN-based model with the Chinese Landscape Painting Dataset to investigate the methods and effects of transferring the style of Chinese landscape painting. Furthermore, the author will evaluate and discuss the authenticity and artistic elements of the generated Chinese landscape paintings.

2. Method

2.1. Dataset preparation

In this study, the author used a set of real landscape images and a set of Chinese landscape paintings. Real landscape photos were obtained from the mountain section in *Landscape Recognition Image Dataset* on Kaggle [7]. The Chinese landscape painting dataset was sourced from *Chinese Landscape Painting Datase*t which is also available on the Kaggle website. The sample images of the collected dataset are presented in Figure 1.
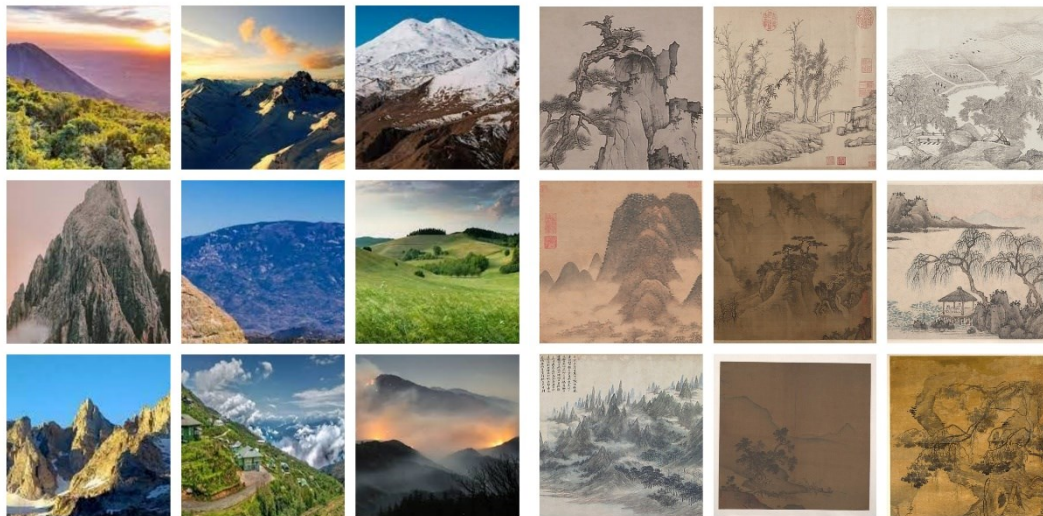

Figure 1. Photos from Landscape Recognition Image Dataset and Chinese Landscape Painting Dataset [7].

Pre-preprocessing for images is as follows: 1) The training sets of both datasets were adjusted to contain 2000 images each. 2) The dimensions of each image were uniformly processed to be 512pixels x 512pixels, and the photos were in RGB colour format. 3) Both datasets were split into training and testing sets in an 80-20 ratio.

2.2. CycleGAN

CycleGAN is an image translation model that does not require paired data in both domains. Innovatively, this model employs two generators and two discriminators to evaluate the effects of both forward and backward transformations. For image to Chinese landscape painting task in this paper, assuming there is a dataset of landscape photos, denoted as X, and a dataset of traditional Chinese landscape paintings, denoted as Y. The goal is to train a generator G, which takes a landscape photo x as input and produces a Chinese landscape painting y', denoted as G(x) = y', where x $\in$ X. Additionally, this study aims to train another generator F, which takes a Chinese landscape painting as input and produces a landscape photo, denoted as F(y) = x', where y $\in$ Y. To ensure the accuracy of both style transformations, it is also necessary to train two discriminators Dx and Dy, which respectively assess the quality of images generated by the two generators. If the generated image y' does not resemble the images y in dataset Y, discriminator Dy should assign it a low score (with the minimum score being 0). Conversely, if the image y' resembles the images y in dataset Y, discriminator Dy should assign it a high score (with the maximum score being 1). Additionally, discriminator Dy should always assign high scores to real images from dataset Y. The overall network architecture is depicted in the Figure 2:
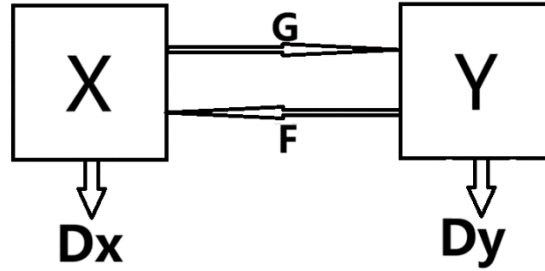
Figure 2. mappings of CycleGAN (Photo/Picture credit: Original).

The loss function of CycleGAN consists of Adversarial Loss and Cycle Consistency Loss:

$$Loss = Loss_{GAN} + Loss_{cycle} \tag{1}$$

$Loss_{GAN}$ ensures the co-evolution of the generator and discriminator, thereby ensuring the generator can produce more realistic images. Its specific components are:

$$Loss_{GAN} = \zeta_{GAN}(G, D_Y, X, Y) + \zeta_{GAN}(F, D_X, X, Y) \tag{2}$$

For further decomposition, Adversarial Loss can be further transformed into the following equation:

$$Loss_{GAN} = E_{y \sim Pdata(y)}[logD_Y(y)] + E_{x \sim Pdata(x)}[log(1 - D_Y(G(x)))] + E_{x \sim Pdata(x)}[logD_X(x)]$$
$$+ E_{y \sim Pdata(y)}[log(1 - D_X(F(y)))] \tag{3}$$

$Loss_{cycle}$ ensures that the output images of the generator have the similar content as the input images, but with a different style. Its specific components are:

$$Loss_{cycle} = E_{x \sim Pdata(x)}[\ ||F(G(x)) - x||_1] + E_{y \sim Pdata(y)}[\ ||G(F(y)) - y||_1] \tag{4}$$

## 2.3. Implementation details

In this study the author uses Pytorch framework to construct the image to painting model based on CycleGAN. For the model, Adam is chosen as optimizer for better convergence performance. The learning rate is not dynamically adjusted during the training process. Some of the training parameters are list in the following Table 1.

Table 1. Part of parameters in the style transfer model for training photos and Chinese landscape paintings.

Parameter	Value
Input image channel	3
Output image channel	3
Epochs	100
Batch size	1
Load size	286
Crop size	256

3. Results and discussion

As depicted in Figure 3, the initial and third columns in each row showcase authentic photographs, while adjacent to each real photo, one can observe its respective Chinese landscape painting following style translation. Notably, the AI-generated Chinese landscape paintings distinctly contribute background colors to the original images. These background hues seamlessly blend with the characteristic colors often found in traditional Chinese landscape paintings. However, a discernible distinction arises when contrasting these AI-generated paintings with their traditional counterparts. In traditional Chinese landscape paintings, the depth of background colors is contingent upon the historical period in which the artwork was crafted. In contrast, the background colors in the AI-generated paintings are influenced by the background's depth within the original images. Secondly, the overall outline of the mountains in generated Chinese landscape paintings remains unchanged, while in contrast, the details in the sky are noticeably reduced. These adjustments align well with the artistic techniques commonly used in Chinese landscape painting. Finally, the colour of the mountains in the generated Chinese landscape paintings tends to become darker as they rise, and trees in these generated paintings emphasize the contours of their trunks and branches. These rendering details largely meet the expressive needs of Chinese landscape painting.
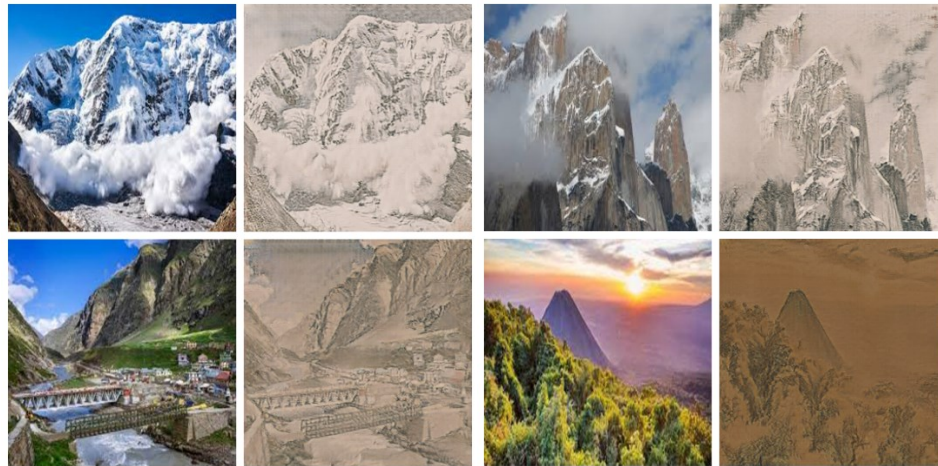
Figure 3. Comparative images of real photographs and generated Chinese landscape paintings (Photo/Picture credit: Original).

In addition, this paper also attempted style transfer from Chinese landscape paintings to real photographs. Figure 4 displays the experimental results.

Figure 4. Comparative images of real Chinese landscape paintings and generated fake photographs (Photo/Picture credit: Original).

From the Figure 4, it can be observed that the image translation model can to some extent provide a richer colour palette for landscape paintings. However, it falls far short of achieving a truly realistic effect. Considering this, the author has the following interpretation of these results:

- The deficiency in color is quite pronounced in traditional Chinese landscape paintings, particularly due to the deliberate removal of detailed sky colors during creation, which poses a certain level of difficulty in restoration.
- Chinese landscape paintings often feature Chinese poems or verses written by artists or collectors, which constitute an unavoidable negative factor in the input for the model.
- When creating Chinese landscape paintings, there is no requirement to preserve all details outside of the contours. Therefore, a considerable amount of detail is lost before the restoration process.

Considering these three key points, the author recommends that future research should take a two-fold approach. Firstly, there should be an emphasis on the identification and removal of extraneous text information from input images prior to the training process. Secondly, for restoration tasks, models such as pix2pix are preferred and warrant further exploration. These models require paired source and target domain. This would enhance the fidelity of style translation from Chinese landscape paintings to real photos. Additionally, some advanced modules such as attention mechanism [8-10] may be also considered for improving the ability to focus on key areas during model generation.

4. Conclusion

This article presents a method based on CycleGAN for the mutual translation between real photographs and traditional Chinese landscape paintings. The research results indicate that the proposed method can effectively achieve the style transfer from real landscape photographs to Chinese landscape paintings. The generated Chinese landscape paintings exhibit a strong correlation with the original images, while also adding distinctive artistic charm of Chinese landscape paintings. The author believes that this achievement may appear in practical applications in the future. Meanwhile, it is also noticeable that there is room for improvement in the authenticity when transforming Chinese landscape paintings into real images. In the future, the author will attempt to conduct more profound research on this task.

References

[1]   Duan L 2019 Semiotics for Art History: Reinterpreting the Development of Chinese Landscape Painting (in Chinese) Theoretical Studies in Literature and Art 41(04):138

[2]   Isola P Zhu J Y Zhou T and Efros A A 2017 Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).

[3]   Liu M Y Breuel T and Kautz J 2017 Unsupervised image-to-image translation networks Advances in neural information processing systems 30

[4]   Zhu J Y Zhang R Pathak D Darrell T Efros A A Wang O and Shechtman E 2017 Toward multimodal image-to-image translation Advances in neural information processing systems 30

[5]   Goodfellow I Pouget-Abadie J Mirza M et al 2014 Generative adversarial nets Advances in neural information processing systems 27

[6]   Zhu J Y Park T Isola P and Efros A A 2017 Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision pp. 2223-2232

[7]   Kaggle 2023 Landscape Recognition Image Dataset https://www.kaggle.com/datasets/utkarshsaxenadn/landscape-recognition-image-dataset-12k-images

[8]   Qiu Y et al 2022 Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training. Biomedical Signal Processing and Control 72 103323

[9]   Sinha A and Dolz J 2020 Multi-scale self-guided attention for medical image segmentation. IEEE journal of biomedical and health informatics 25(1) 121-130

[10]   Farag M M Fouad M and Abdel-Hamid A T 2022 Automatic severity classification of diabetic retinopathy based on densenet and convolutional block attention module IEEE Access 10 38299-38308