

PAPER • OPEN ACCESS

Research on Neural Style Transfer Algorithm

To cite this article: Chengsi Yao *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **569** 052107

View the [article online](#) for updates and enhancements.

You may also like

- [Research on Neural Style Transfer](#)
Xueyuan Liu
- [Improved Image Style Transfer Based on VGG-16 Convolutional Neural Network Model](#)
Chenhao Shu and Xichuan Hu
- [Research advanced in image style transfer based on deep learning](#)
Haozhuo Shi and Yingqi Wang



245th ECS Meeting • May 26-30, 2024 • San Francisco, CA

Present your work at the leading electrochemistry & solid-state science conference.

Network with academic, government, and industry influencers!

Submit abstracts by December 1, 2023

[Learn more & submit!](#)



Research on Neural Style Transfer Algorithm

Chengsi Yao^{1*}, Yuanhao Li¹, Yali Qi¹

¹ Information Engineering, Beijing Institute of Graphic Communication School, Beijing, Beijing, 102600, China

*Corresponding author's e-mail: 798218356@qq.com

Abstract. Researchers have successfully applied the convolutional neural network (CNN) to style transfer. Since then, Neural Style Transfer (NST) has received widespread attention in both scientific and industrial fields. Researchers in the field of machine vision are constantly proposing ways to optimize image style migration. This paper aims to summarize the history of style transfer before and after the rise of CNN, classify the existing classical and improved algorithms, and compare the results of some of them. Finally, after this study, we put forward some suggestions on the development trend of image style transfer.

1. Introduction

Painting is one of the most popular works of art and is highly appreciated. People who like a particular style want to combine this style with their own images, and this need makes image style transfer a research trend.

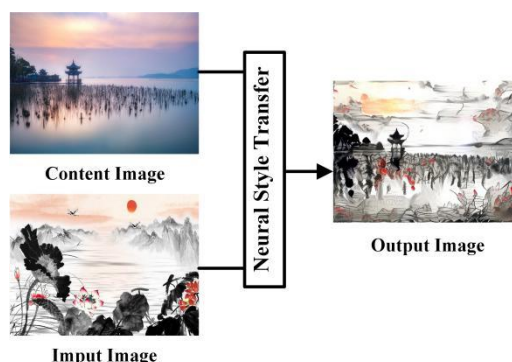


Figure 1. The content image is the west lake scenery, the style image is the ink painting, after the style transfer to get the ink painting effect of the west lake scenery.

Style transfer can be traced back to the research of image texture synthesis. All previous papers on image texture were manually modeled[1], and the most important idea used was that texture can be described by the statistical model of image local features. At the same time, the researchers also tried the transfer of oil painting style and head portrait style transfer[2-3].

With the development of computer graphics, deep learning has been greatly developed because it can be used to train object recognition models. CNN can extract features. In addition to image recognition and image classification, CNN is also used for style transfer.

In this paper, we will be mainly divided into three parts. In the first part, the style transfer algorithm without CNN. The second part, the typical NST algorithm. The third part is the summary of the improved



algorithm of style transfer. In other parts of the article, we also briefly describe the application and future development of style transfer.

2. Style transfer without neural network

Before the rise of CNN, researchers tried to transfer styles without relying on neural network algorithms.

Transfer the paintbrush of a famous artist to the photo to be rendered to get a good portrait[4]. Portrait drawing is an extension of image analogy, but its data is difficult to obtain in our practice, which has caused certain limitations to the extension research[5].

Many objects in nature have similar textures. We can extract the texture from the sample and regenerate a large amount of image data. The texture synthesis algorithm can be extended to transfer the texture obtained from different targets and re-render the image. Or combine existing textures to create new ones[6-7].

3. Neural Style Transfer

In 2012, deep learning gradually developed and rapidly gained wide attention from researchers. Neural networks can automatically extract useful features after a certain amount of training, rather than simply dividing objects into small pieces. At the time, one of the most famous object recognition networks of CNN was called VGG-19. VGG-19 network appeared in the paper of Simonyan and Zisserman in 2014[8]. The network uses 3x3 convolution kernel, maximum pooling, after the full connection layer connected to a softmax classifier.

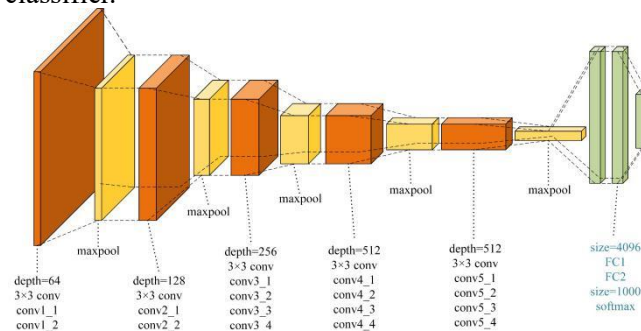


Figure 2. The first 16 layers of VGG-19 are the alternation of convolution layer and maximum pooling layer. Each convolution contains multiple convolution layers, and finally there are three fully connected layers.

VGG-19 can extract more complex features with each layer of convolution. In short, VGG-19 network is a pile of local feature recognizers. VGG-19 output can be visualized in every layer. This paper uses a picture of a cat as an example.

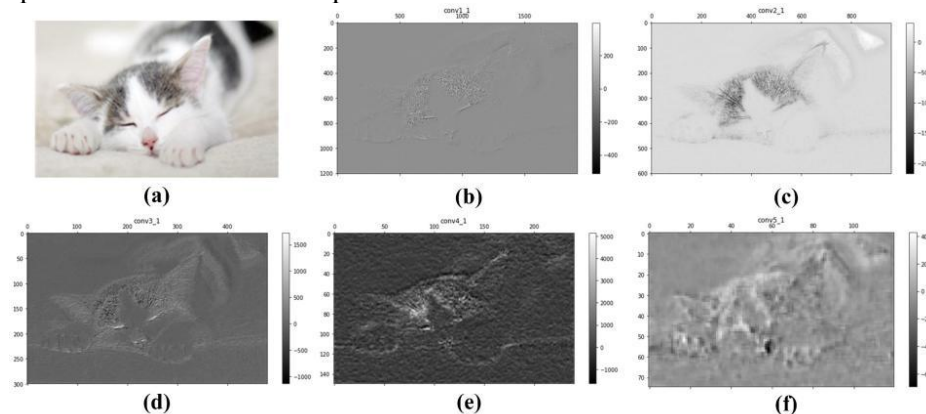


Figure 3. (a) is the original image, and (b) to (f) are the visualized output of VGG-19

3.1. Slow Neural Style Transfer

In 2015, Gatys et al. combined neural network and style transfer, which officially opened the prelude of Neural Style Transfer. Gatys et al. applied the gram matrix to different local feature maps extracted by the VGG-19 network and calculated the correlation between features to form a statistical model[9]. The computation of style features is performed on all convolution layers. A part of local features is directly taken as the content, and the content features are for a convolution layer. Finally, combine the content features and style features of the picture together to form a new picture.

Firstly, a white noise image is generated, and iteration is carried out according to the loss function. Then, gradient descent algorithm is used for back propagation and continuous optimization to obtain the minimum loss. The loss function used by Gates et al. is the sum of content loss and style loss, both of which have their own parameters.

3.1.1. Content loss

Given a picture \vec{p} , the feature map is obtained by calculating in the neural network. Each layer can get n feature maps, which is determined by the number of filters. The feature map is vectorized and the resulting vector is finally put into the matrix F . The element F_{ij} represents the activation response of the i filter in the l layer at position j . Specify the feature representation of a layer l , and generate a style-transferred image \vec{x} , so that the feature representation of the layer P_l is equal to the original feature representation of F_l . The loss function is defined as follows:

$$\mathcal{L}_{content}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - P_{ij}^l)^2 \quad (1)$$

3.1.2. Style loss

As mentioned above, the style of a picture is actually a Gram matrix obtained by calculating the correlation of features between layers. Gram matrix contains style content, such as texture, color and so on. Its definition is as follows:

$$G_{ij}^l = \sum_k F_{ik} F_{jk} \quad (2)$$

Given the style picture \vec{a} , the target generated picture is \vec{x} , and the Gram matrix corresponding to the l layer is A_l and G_l , respectively. The loss function of this layer is defined as follows:

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2 \quad (3)$$

3.1.3. Total Style Loss

$$\mathcal{L}_{style} = \sum_{l=0}^L \omega_l E_l \quad (4)$$

ω_l is the weight of each layer. The weight of style layer is the reciprocal of the number of style layers, and the weight of other layers is 0.

3.1.4. Total loss

By adding the content loss and style loss in a certain proportion, and the optimized output figure x is obtained through back propagation. Its definition is as follows:

$$\mathcal{L}_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{content}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{style}(\vec{a}, \vec{x}) \quad (5)$$

3.2. Fast Neural Style Transfer

Gatys et al.'s method takes thousands or even thousands of iterations to generate a new image each time, which takes quite a long time. In order to save time and cost, based on the work of Gatys et al., Johnson et al. proposed a style transfer algorithm that can improve the speed by three orders of magnitude, which is called Fast Neural Style Transfer[10]. Two networks are used: image transform network and loss network.

Image transform network, which requires pre-training, so that any given a content picture, can quickly generate a fixed style of migration picture. Loss network is used to calculate the perceptual loss. The perceptual loss is used to train the image transform network and update its parameters.

4. Speed and quality improvement

Although Johnson et al. greatly improved the speed of style transfer, they could only limit themselves to one style. Researchers have developed an algorithm for incorporating multiple styles into a single model[11]. Vincent Dumoulin et al. used a deep network to acquire paintings of many different styles[12]. Chen et al. proposed the concept of Stylebank Network[13]. Stylebank Network consists of stylebank layer, encoder and decoder.

Shen et al. rejected the method of stochastic gradient descent and proposed the concept of Meta Networks[14]. The Meta Networks only needs one feedforward neural transmission to acquire the style of any style image and generate the corresponding image generation network.

To improve the authenticity of the content of the generated images, some experts are beginning to introduce Generative Adversarial Networks(GAN) into style transfer[15]. Jun et al. used Cycle-Consistent Adversarial Networks to conduct image style transfer. Introduce forward and backward mappings in the source domain X and the target domain Y[16].













Group Methode	Style 1 Content 1	Style 2 Content 2	Style 3 Content 3
Gatys et al.			
Johnson et al.			
Shen et al.			
Jun et al.			

Figure 4. The comparison of several different Neural Style Transfer algorithms shows that gatys et al. take a long time, but the results are relatively stable, while the other three algorithms save time, but the final output varies with the input, sometimes good, sometimes bad.

5. Applications and prospects

Neural Style Transfer plays an important role not only in the field of art, but also in other areas. Popular photo-editing software for young people, such as meitu xiuxiu, allows you to change the style of your photos in one step. Recently, the popular face changing software is a derivative algorithm based on NST[17]. NVIDIA researchers proposed a general method of linear transfer, and theoretically obtained the expression form of the transfer matrix[18]. It plays an important role in game scene construction and

video post processing. In future studies, researchers can extend the style model to make the extraction of arbitrary styles more accurate, and put more research into video style transfer. We should gradually extend the style transfer to more engineering fields for the benefit of mankind.

Acknowledgments

This research is supported by Collaborative innovation project of green printing and publishing technology of Beijing Municipal Education Commission: PXM2016-014223-0000025 and Scientific Research Project of Beijing Education Committee (KM201710015010).

References

- [1] Portilla, J., Simoncelli, E.P. (2000) A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40: 49-70.
- [2] Semmo, A., Limberger, D., Kyprianidis, J. E., Döllner, J. (2015) Image Stylization by Oil Paint Filtering using Color Palettes. In: *Workshop on Computational Aesthetics*. Potsdam. pp. 149-158.
- [3] Shih, Y. C., Paris, S., Barnes, C., Freeman, W. T., Durand, F. (2014) Style transfer for headshot portraits. *ACM Transactions on Graphics*, 33: 1-14.
- [4] Zhao, M. T., Zhu, S. C. (2011) *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering*. ACM, New York.
- [5] Li, Y. H., Zhang, T. Y., Han, X., Qi, Y. L. (2018) Image Style Transfer in Deep Learning Networks. In: *5th International Conference on Systems and Informatics (ICSAI)*. Nanjing. pp. 660-664.
- [6] Gatys, L.A., Ecker, A.S., Bethge, M. (2015) Texture Synthesis Using Convolutional Neural Networks. <https://arxiv.org/abs/1505.07376>.
- [7] Efros, A. A., Freeman, W. T. (2001) *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM, New York.
- [8] Simonyan, K., Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. <https://arxiv.org/abs/1409.1556>.
- [9] Gatys, L. A., Ecker, A. S., Bethge, M. (2015) A Neural Algorithm of Artistic Style. <https://arxiv.org/abs/1508.06576>.
- [10] Johnson, J., Alahi, A. (2016) Perceptual Losses for Real-Time Style Transfer and Super-Resolution. <https://arxiv.org/abs/1603.08155>.
- [11] Jing, Y. C., Yang, Y. Z., Feng, Z. L., Ye, J. W., Yu, Y. Z., Song, M. L. Neural Style Transfer: A Review. <https://arxiv.org/abs/1705.04058>.
- [12] Dumoulin, V., Shlens, J., Kudlur, M. (2017) A Learned Representation For Artistic Style. <https://arxiv.org/abs/1610.07629>.
- [13] Chen, D. D., Yuan, L., Liao, J., Yu, N. H., Hua, G. (2017) StyleBank: An Explicit Representation for Neural Image Style Transfer. <https://arxiv.org/abs/1703.09210>.
- [14] Shen, F. L., Yan, S. C., Zeng, G. (2018) Neural Style Transfer via Meta Networks. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City. pp. 8061-8069.
- [15] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014) Generative Adversarial Nets. In: *28th Conference on Neural Information Processing Systems*. Montreal. 2672-2680.
- [16] Zhu, J. Y., Park, T., Isola, P., Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice. pp. 2242-2251.
- [17] Korshunova, I., Shi, W. Z., Dambre, J., Theis, L. (2017) Fast Face-swap Using Convolutional Neural Networks. <https://arxiv.org/abs/1611.09577>.
- [18] Li, X. T., Liu, S. F., Kautz, J., Yang, M. H. (2018) Learning Linear Transformations for Fast Arbitrary Style Transfer. <https://arxiv.org/abs/1808.04537>.