Neural Network Techniques for Image Style Transfer

Ziqi Zhang*

School of Computer Science, Sichuan Normal University, 610100, Chengdu, China

Abstract. With the rapid advancement of deep learning technology, neural networks have achieved remarkable results in the field of image processing. Particularly in image style transfer, neural network-based methods have become a research hotspot. Image style transfer aims to apply the stylistic features of one image to another while preserving its content information, thereby creating images with artistic value. This paper reviews the primary methods of image style transfer, including slow transfer based on image iteration and fast transfer based on model iteration and compared the experimental results of various methods. It further explores the implementation approaches and technical developments of single-style, multi-style, and arbitrary style transfer. Analyzing the strengths and weaknesses of existing methods, this paper examines efficiency improvement, image quality enhancement, and diversity augmentation, highlighting the challenges in content preservation, computational resource requirements, and transfer quality in current style transfer techniques. This paper aims to provide researchers with future research directions.

1 Introduction

A key field of research in computer vision is image style transfer, which aims to combine the style of one image with its content to produce a new image with artistic value. The rapid development of deep learning technology, especially the efficient use of convolutional neural networks (CNNs), has enabled significant progress in picture style transfer research. In complicated image contexts, traditional style transfer approaches often rely on hand-crafted features and rules that are unable to capture high-level semantic information in images. Therefore, deep learning-based style transfer methods have emerged, utilizing the powerful representation ability of neural networks. These techniques broaden the scope of style transfer's applicability while simultaneously enhancing its quality.

Numerous scholars have studied picture style transmission in recent years, putting forth a number of creative models and methods. Gatys et al., for example, presented the neural style transfer approach, which uses a pre-trained VGG model to extract content and style elements and optimizes a loss function to produce stylized images [1]. This method laid the foundation for the field of style transfer. Additionally, Huang and Belongie proposed the AdaIN method, which adjusts the statistical properties of content and style features for more

* Corresponding author: zhangban@asu.edu.pl

flexible style transfer [2]. Meanwhile, the CycleGAN algorithm enabled style transfer between unpaired images, greatly broadening its applications, particularly in art creation and image processing [3]. Moreover, attention-based style transfer methods, such as AdaAttN, have gained attention for improving the quality and expressiveness of generated images [4]. These successful methods indicate that research in image style transfer is moving toward more intelligent and flexible approaches.

Despite the success of existing style transfer methods, several challenges remain. For instance, content loss or blurring may occur during style transfer, especially when dealing with complex images. Additionally, traditional methods often have high computational resource requirements, limiting their applicability to mobile devices. Thus, maintaining image content clarity while improving transfer efficiency has become a focal point of current research.

Reviewing and evaluating the primary techniques in picture style transfer while meticulously describing the state of the field's research and its obstacles is the aim of this work. The article is organized as follows: Neural network-based image transmission techniques are introduced in Section 2, including statistical and non-statistical parameter methods; Section 3 discusses model iteration-based image transfer methods, covering single, multi, and arbitrary style transfers; Section 4 presents performance analysis of different methods through experimental results; Section 5 explores the current status and challenges of image style transfer applications; and Section 6 offers recommendations for further research as the paper's conclusion.

2 Slow Neural Network Style Transfer Based on Image Iteration

2.1 Statistical Parameter-Based Methods

In image processing, texture refers to the local repetitive features in an image (such as color, shape, patterns, etc.), and it is one of the core elements of an image's style. Traditional texture synthesis methods often rely on hand-designed features, but these methods fall short when handling complex artistic styles.

Given that CNNs can extract increasingly abstract information from images layer-by-layer, Gatys et al. proposed using CNNs to learn and describe visual textures. This hierarchical representation is effective at capturing both local and global information, making it well-suited for handling textures and styles [1].

2.1.1 Gram Matrix-Based Methods

In 2016, Gatys et al. introduced a feature-space texture model based on convolutional neural networks, which elevated traditional image style transfer methods to a new level and opened new directions for style transfer via deep learning [1]. This model represents the image's style using the Gram matrix and uses CNNs to extract features from images. The Gram matrix is defined as:

$$G_{ij}^{l} = \sum_{k} F_{ik}^{l} F_{jk}^{l} \tag{1}$$

where the feature responses of the i-th and j-th channels in the l-th layer are denoted by $\mathrm{F}_{\mathrm{ik}}^{\mathrm{l}}$ and $\mathrm{F}_{\mathrm{jk}}^{\mathrm{l}}$, respectively. Entering a content image and a style image into a convolutional neural network is the main procedure. While the style image goes through several levels of feature extraction, the content picture's high-level features are extracted after it has passed through the convolutional layers. Each feature map's texture information is captured by computing the Gram matrix at each layer. After initializing a random noise picture, the image is iteratively refined using gradient descent optimization to match the style aspects of the style image as well as the content features of the content image. The following is the detailed algorithm: In the Gatys et al. method, given a target image, a content image, and a style image, the total loss function is presented below:

$$L_{tatal}(I, I_c, I_s) = \alpha L_{content}(I, I_c) + \beta L_{style}(I, I_s) \tag{2}$$

where the weights for the style loss function $L_{style}(I, I_s)$ and the content loss function $L_{content}(I, I_c)$ are β and α. The style loss $L_{style}(I, I_s)$ is defined by the difference in Gram matrices across layers:

$$L_{style}(I, I_s) = \sum_{L=0}^{L} \omega_l E_l \tag{3}$$

where ω_l is the weight of the style loss at the l-th layer, and E_l is the style loss at the l-th layer of VGG16, defined as:

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} \left(G_{ij}^l(I_s) - G_{ij}^l(I) \right)^2 \tag{4}$$

where N_l is the number of filters in the l-th layer, M_l is the size of the feature map at the l-th layer, and $G_{ij}^l(*)$ represents the Gram matrix of image $*$ at layer l [1].

2.1.2 Maximum Mean Discrepancy

Although Gatys et al.'s method achieved impressive results in capturing style, it still has limitations in matching high-order statistical features, failing to fully represent the high-order statistics of the style image [1]. To address this, Li et al. introduced a novel approach that frames the problem as domain adaptation and uses MMD to measure and optimize the distributional difference between the generated image and the style image in the feature space [5]. Specifically, MMD replaces the traditional Gram matrix-based style loss to better capture higher-order statistical features, leading to more precise style transfer.

Despite improvements over Gatys et al.'s algorithm, these methods still fail to resolve issues such as changes in brushstrokes and missing deep positional information, which significantly affect the quality of generated images.

2.2 Non-Statistical Parameter-Based Methods

Non-statistical parameter-based techniques, as compared with statistical parameter-based methods first separate the style and content images into several regional blocks. To accomplish style transfer, the two images' most comparable regional blocks are then matched. These techniques do a good job of maintaining the image's local characteristics. Non-statistical parameter approaches include deep neural networks based on Markov Random Fields, semantic style transfer, and deep image analogy.

2.2.1 Markov Random Field

Li and Wand were the first to propose a neural network-based style transfer algorithm using MRF [6]. In terms of obtaining the style image's local texture information, this approach offers a number of advantages. An MRF loss function was used in place of the Gram matrix loss function that was employed in classic style transfer. The following is an expression for the MRF's loss function:

$$L_s = \sum_{l \in \{l_s\}} \sum_{i=1}^{m} \left\| \Psi_i \left(F^l(I) \right) - \Psi_{NN(i)} \left(F^l(I_s) \right) \right\|^2 \tag{5}$$

where $F^l(*)$ denotes the feature map taken from layer l, I is the target picture, and I_s is the style image. The created image $*$ is the input. The most comparable patch in the style picture that corresponds to the i-th local area is $\Psi_{NN(i)}\left(F^l(I_s)\right)$, while $\Psi_i\left(F^l(I)\right)$ indicates the local feature that was taken from the feature map of the generated image at position i.

This technique improves the style transfer outcomes, especially when it comes to synthesising local textures and details, by allowing the model to extract cross-layer texture information. When the style photos and text are comparable, it performs very well. However, the picture patches might not match perfectly if the content and style images differ significantly, which would produce less than ideal outcomes.

2.2.2 Semantic Style Transfer

To address the limitation of existing methods that focus only on low-level texture features and high-level structural features, while ignoring semantic information in different regions of the image, Champandard proposed a method that combines semantic segmentation with the Markov Random Field algorithm [7]. This technique guarantees that style matching takes place between regions with the same semantic meaning by applying semantic masks to both the content and style pictures. This method greatly enhances the visual quality and realism of style transfer while successfully preventing artificial texture transfer occurrences.

2.2.3 Deep Image Analogy

Liao et al. proposed a style transfer method based on deep image analogy, where they find similar regions between the content and style images using analogy analysis and transfer the style texture to those regions [8]. Nevertheless, this approach is unable to successfully maintain the image's global semantic information. To solve this issue, Gu et al. introduced feature rearrangement loss, adding constraints to the nearest neighbor algorithm to ensure that each region is matched, thereby better maintaining the global semantics of the image [9].

The above methods suffer from large computational costs and slow speeds, which makes it difficult to apply them in practical scenarios.

3 Fast Neural Network Style Transfer Based on Model Iteration

While image-iterative slow neural network style transfer generates high-quality images, most methods rely on complex computational processes, limiting their efficiency in practical applications. In order to get over these restrictions, model iteration-based rapid neural network style transfer techniques have been put forth, which optimize the generative model's structure and training methodology for more effective style transfer.

3.1 Single-Style Image Transfer

In 2016, Johnson et al. proposed a fast style transfer method that set the precedent for performing style transfer using a feedforward neural network [10]. This technique builds a loss network and an image transformation network, training the network with perceptual loss. Consequently, the stylized image incorporates the visual attributes of the target style while maintaining the primary qualities of the content image. Compared to traditional iterative optimization-based style transfer methods, fast style transfer significantly improves execution speed and is suitable for real-time applications, producing high-quality stylized images.

Similar to Johnson's method, Ulyanov et al. proposed the texture network, which focuses on real-time texture generation and image stylization [10,11]. The generative network employs a multi-scale pyramid structure. Although the method is not as adaptive to style generation as Johnson's approach, the multi-scale design allows better handling of style information at different levels. Subsequently, Ulyanov et al. further optimized the style transfer network by introducing instance normalization as a replacement for the traditional batch normalization. Instance normalization normalizes each image individually, resulting in better style transfer, especially in tasks involving single image style transfer.

Li et al. proposed the Markov Generative Adversarial Networks (MGANs), which combine the ideas of Generative Adversarial Networks (GANs) and MRF [12]. MGANs perform style transfer using a patch-level non-parametric method, excelling in the generation of coherent textures for complex images. MGANs solve the inefficiency problems of conventional style transfer techniques by using adversarial training.

With the further development of GANs, CycleGAN introduced the concept of cycle-consistency loss to address the problem of unpaired image translation [13]. This method combines unsupervised and supervised learning to perform image translation between source and target domains without requiring paired datasets.

3.2 Multi-Style Image Transfer

Conventional single-model single-style approaches necessitate independent training for every style, which is inflexible and time-consuming. In order to increase efficiency, researchers have suggested single-model multi-style approaches that increase the flexibility of style transfer through the integration of several styles into a single model. In this field, there are two basic techniques such as one binds a limited set of parameters for every style, while the other feeds the model with both style and content aspects.

In 2017, Dumoulin et al. proposed Conditional Instance Normalization (CIN), a method that enables multi-style transfer by binding a small number of parameters for each style within the same convolutional neural network [14]. By altering the normalization layer's scaling and shifting parameters, CIN replicates many styles, allowing for quick style transition and transfer without adding further convolution weights. This method can also generate images that blend multiple styles by combining different style parameters.

Another approach, proposed by Li et al., involves a style selection unit, which encodes the styles and combines the style and content features as input to the decoder to generate stylized images [12]. By controlling the combination of style-encoded features with content-encoded features, this method can generate images in various styles. This approach not only allows the model to flexibly adapt to style changes but also facilitates the extension of new styles by simply adjusting style features.

Similar to Dumoulin's approach, Chen et al. proposed the StyleBank method [15]. The above method represents many different looks using multiple convolution filter banks, each of which corresponds to a distinct style. StyleBank achieves style transfer by convolving content image features with filters specific to the desired style. One advantage of StyleBank is its support for flexible incremental training: new style convolutional banks can be trained on top of an existing model without needing to retrain the entire network.

Nevertheless, as the assortment of patterns increases, the model size usually does as well, and the photographs' clarity and details may deteriorate. To deal with this problem, Zhang et al. proposed a novel method by introducing the CoMatch layer, which matches the second-order statistical features of styles, achieving higher-quality multi-style transfer. Additionally, they introduced the Multi-Style Generation Network (MSG-Net), which enhances the capability for real-time generation [16].

In addressing artifacts and color blending issues in style transfer, Li et al. further improved the method for extracting style features and optimized the texture representation in style transfer, making the generated images appear more natural. Furthermore, Qiao Ping'an et al. proposed a more efficient solution for animation style transfer using the MC-CartoonGAN model, addressing the issues of large parameter sizes and image detail loss, significantly improving both transfer efficiency and image quality [17].

3.3 Arbitrary Style Image Transfer

While multi-style models have somewhat solved the issue of model size, generating a new style still requires additional training time. Subsequently, arbitrary style models emerged, where a single model can generate images in any style. With approaches based on MRF and adaptive normalization based on statistical distributions, respectively, the development of single-model arbitrary style transfer may be broadly classified into two directions: parametric models and non-parametric models.

3.3.1 Non-Parametric Arbitrary Style Transfer Models

Chen and Schmidt proposed a non-parametric arbitrary style transfer model based on the concept of Markov Random Fields (MRF)[18]. By matching local patches in the content image with the most similar patches in the style image using the feature space of a pre-trained VGG network, they exchange the two and utilize an image reconstruction approach to produce the stylized output. This method offers good flexibility and can generate images in various styles. However, its style transfer effect can sometimes be suboptimal because the patch swapping may not always align with the desired style.

3.3.2 Parametric Arbitrary Style Transfer Models

Adaptive Instance Normalization (AdaIN), proposed by Huang and Belongie, is a representative method for parametric arbitrary style transfer [2]. By modifying the mean and variance of the content and style picture attributes, as indicated by the following formula, this technique accomplishes style transfer:

$$\mathrm{AdaIN}(x, y) = \sigma(y)\left(\frac{x-\mu(x)}{\sigma(x)}\right) + \mu(y) \tag{6}$$

where $\mu(*)$ and $\sigma(*)$ stand for the mean and standard deviation of the feature mappings $*$, and x and y for the feature representations of the content and style images, respectively. The above technique maintains the content image's structure while achieving quick style transfer. While AdaIN can generate stylized images in real-time, its style expression is somewhat limited, often restricted to color transfer and lacking detailed and complex structural style expression.

Li et al. proposed Whitening and Coloring Transform (WCT), which adjusts the statistical distribution of the content and style features to capture more complex style characteristics [19]. Whitening the content elements to eliminate the original style information is one of the fundamental processes, followed by a coloring transformation to incorporate style information, and finally generating the stylized image through a decoder. The multi-level WCT algorithm further improves visual quality by progressively capturing detailed style features, resulting in more refined stylized outcomes. The formulas for WCT are as follows:

$$F_c' = E_c D_c^{-\frac{1}{2}} E_c^T F_c \tag{7}$$

$$F_s' = E_s D_s^{-\frac{1}{2}} E_s^T F_s \tag{8}$$

The content and style pictures' feature matrices are denoted by $\mathrm{F_c}$ and $\mathrm{F_s}$, whereas the whitened content and style feature matrices are denoted by $\mathrm{F_c}'$ and $\mathrm{F_s}'$, respectively. $\mathrm{E_c}$ and $\mathrm{E_s}$ are the eigenvector matrices of the covariance matrices for content and style features, and $\mathrm{D_c}$ and $\mathrm{D_s}$ are the diagonal matrices of these covariance matrices, with the diagonal elements representing the eigenvalues. This multi-level style transfer method excels at capturing style details but comes with high computational costs.

To better understand the characteristics and limitations of different neural network-based style transfer methods, this paper summarizes the advantages and disadvantages of the mainstream methods, including their application scenarios and performance, for comparison and selection, as shown in Table 1.

Table 1. Summary of advantages and disadvantages of image style transfer methods based on neural networks.

Method	Advantages	Disadvantages
Gram Matrix-based Method	Generates artistically styled images	High computational cost, difficult to control style and content balance
Maximum Mean Discrepancy	Suitable for unsupervised learning, achieves fine details	Long computation time, color discrepancies
Markov Random Field (MRF)	Flexibly models the local spatial relationships of images	Requires a lot of manual intervention, high computational complexity, may slow down optimization
Semantic Style Transfer	Ensures semantic consistency across regions with semantic masks, improving rationale and visual effects	Difficult to handle style consistency in complex scenes
Deep Image Analogy	Retains similar region textures through region matching	Struggles to preserve global semantic information effectively
Fast Style Transfer Method	Strong real-time performance, suitable for fast applications	The generated stylized effects are relatively simple, lacking details
Instance Normalization Method	Improves the style transfer quality of a single image, generating better results	Less effective in multi-image scenarios
MGANs (Markov-	Excels in generating complex and	High training difficulty, lacks

Generative Adversarial Networks)	coherent texture images	semantic considerations
CIN	Generates realistic images with fine details	Lacks details in high-quality textures
Style Selection Unit Method	Adapts to different style variations, generates diverse and flexible stylized images	Style fusion effect is sometimes suboptimal
StyleBank Method	Requires fewer parameters, supports flexible incremental training, suitable for style extension	As the number of styles increases, model size grows, and training costs increase
MSG-Net Method	Supports multi-style transfer with high real-time performance, suitable for creative applications	Artifacts problem not fully resolved, color mixing performance needs improvement
MC-CartoonGAN	Optimized for cartoon style transfer, better at retaining color and texture	Room for improvement in detail representation
Non-Parametric Arbitrary Style Transfer	High flexibility, suitable for generating images with various style combinations	Style block replacement may not always meet expectations, resulting in unnatural images
AdaIN	Enables fast style transfer, suitable for real-time stylization tasks	Limited style expression, coarse details
WCT	Captures detailed complex styles, generating refined images	High computational cost, more suitable for offline applications

4 Experimental Results and Analysis

The experiment's measures for assessing the quality of the produced images were the SSIM and the PSNR. These two metrics are commonly used to compare the perceptual and result-based measurements between two images.

Based on pixel differences, PSNR is a commonly used statistic for evaluating image quality that gauges how similar the reconstructed and original images are. Better image quality, with less variations from the original image, is indicated by a higher PSNR score. The PSNR calculation formula is:

$$\mathrm{PSNR} = 10 * \log_{10}\left(\frac{\mathrm{R}^2}{\mathrm{MSE}}\right) \tag{9}$$

Where R is the image's maximum pixel value and MSE is the difference between the produced and original pictures.

Another crucial statistic for evaluating image quality is SSIM, which was created to more accurately represent the perceptual features of the human visual system. To measure how similar the original and reconstructed images are, it examines the image's brightness, contrast, and structural details. Higher image quality is indicated by an SSIM value that is closer to 1. The SSIM formula is:

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (10)$$

where the average luminance of images x and y is represented by μ_x and μ_y, their variances are represented by σ_x^2 and σ_y^2, the covariance is represented by σ_{xy}, and stability-adding constants C_1 and C_2 are included.

The following tables (Table 2 and Table 3) summarize the PSNR and SSIM comparisons between typical methods:

Table 2. PSNR and SSIM Comparison Between Typical Methods [20].

Method	PSNR	SSIM	Method Category
Gatys Style [1]	11.56602	0.28774	Image Optimization-Based Style Transfer
Johnson Style [10]	11.90553	0.407876	Single-Style
Ulyanov Style [11]	12.70022	0.542995	Single-Style
Li and Wand Style [6]	12.25893	0.308762	Single-Style
Li Diverse Style [12]	12.49095	0.383528	Multi-Style
Zhang Style [16]	12.47281	0.417546	Multi-Style
Huang Style [2]	12.4078	0.309562	Arbitrary Style
Li Universal Style [19]	11.46045	0.264825	Arbitrary Style

Table 3. Timing of Different Typical Methods in Various Dimensions [20].

Method	Time (s)	Time (s)	Time (s)
	256x256	**512x512**	**1024x1024**
Gatys et al. [1]	14.32	51.19	200.3
Johnson et al. [10]	0.014	0.045	0.166
Ulyanov et al. [11]	0.022	0.047	0.145
Zhang and Dana [16]	0.019	0.059	0.23
Li et al. [12]	0.017	0.064	0.254
Chen and Schmidt [18]	0.123	1.495	-

Li et al. [19]	0.62	1.139	2.94

Tables 2 and 3 show that although Gatys et al.'s method produces lower PSNR and SSIM values, their pixel-wise optimization strategy produces comparatively higher image quality, despite having a very lengthy computation time. Ulyanov's method, on the other hand, demonstrates better PSNR and SSIM than most methods, and Johnson's method also performs reasonably well. This is because the accelerated neural network approach used in Johnson's method significantly outperforms Gatys et al.'s image optimization-based style transfer method in terms of execution speed. The arbitrary style methods, such as those by Huang and Li Universal, exhibit weaker performance in arbitrary style conversion, especially Li Universal Style, which has the lowest PSNR and SSIM. However, it offers higher versatility.

5 Current Applications

Image style transfer technology has demonstrated broad application prospects and far-reaching impact in many fields. In the entertainment industry, in the realms of art creation and image processing, style transfer enables ordinary users to easily transform photos into the style of classic artwork. In the business sector, well-known brands leverage style transfer technology to enhance the artistic appeal and uniqueness of advertisements and product displays, thereby strengthening brand attraction. In the medical field, image style transfer assists in improving the readability of medical images, aids clinical diagnosis, and increases the accuracy and usability of medical imaging. In the gaming industry, this technology provides game developers with a solution to quickly generate diverse artistic styles for game scenes, improving development efficiency. In the education sector, style transfer technology serves as a teaching tool, enhancing students' understanding of art and image processing techniques, while also promoting further research into related algorithms. Style transfer opens up new creative options for image manipulation in the realm of image processing. In conclusion, image style transfer is continuously driving innovation and development across various domains.

6 Challenges and Prospects

Despite significant progress in image style transfer technology in recent years, many challenges remain. First, in terms of transfer learning theory, how to effectively represent and extract shared features has become a central issue. This is especially true when there are significant data distribution differences between the source and target domains, as domain shifts can affect transfer performance. Additionally, many of the current advanced style transfer models are large-scale, leading to high computational complexity and storage requirements, making the realization of lightweight models (such as model compression, quantization, and knowledge distillation) an urgent challenge. Additionally, handling both the subject and backdrop of images for more sophisticated style transfer presents a problem, as style transfer overlaps with tasks like image categorization and segmentation. There is also a lack of standardized criteria for evaluating style transfer effectiveness, and the objective and subjective evaluation of content and style preservation is still an area worth further investigation. The complexity of manually adjusting model parameters, combined with the large size and computational demands of existing pre-trained models (such as VGG), limits the efficiency and flexibility of style transfer.

In the future, the development of image style transfer can be explored from multiple directions. Cross-modal style transfer is expected to expand into domains such as music and video, enriching the application scenarios. The fusion of multiple input information (such as semantic segmentation and sentiment analysis) is expected to enhance style transfer effects while improving real-time performance and interactivity, allowing users to adjust the style instantaneously and providing a more personalized experience. Furthermore, integrating attention mechanisms can better control the style transfer process, focusing on important regions in the image, thus promoting finer-grained style control. For anime-style transfer, in-depth research into its features and the development of interactive tools will meet personalized needs. Finally, combining graph neural networks with style transfer will enhance the capture of semantic and structural information in images, leading to the intelligent development of this technology and expanding its application potential in various fields.

7 Conclusion

The traditional techniques for image style transfer have been examined in this work, along with the present difficulties and potential avenues for further study. Although existing technologies have been successfully applied in many fields, there is still room for improvement. As an important branch of computer vision and deep learning, the future development of style transfer will rely on more efficient algorithms and broader practical applications. By exploring interdisciplinary collaborations and the fusion of new technologies, it is hoped that future innovations in this field will continue to drive progress and provide users with richer and more diverse visual experiences.

References

1. L. Gatys, A. Ecker, M. Bethge, A Neural Algorithm of Artistic Style. Journal of Vision. 16(12), 326-326 (2016)
2. X. Huang, S. Belongie, Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization, Proceedings of the IEEE International Conference on Computer Vision, (2017), 1510-1519
3. J.Y. Zhu, T. Park, P. Isola, et al., Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. Proceedings of the IEEE International Conference on Computer Vision. 2242-2251 (2017)
4. S. Liu, T. Lin, D. He, et al., AdaAttN: Revisit Attention Mechanism in Arbitrary Neural Style Transfer. Proceedings of the IEEE/CVF International Conference on Computer Vision. 6649-6658 (2021)
5. Y.H. Li, N.Y. Wang, J.Y. Liu, et al., Demystifying Neural Style Transfer. arXiv:1701.01036 (2017)
6. C. Li, M. Wand, Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2479-2486 (2017)
7. A.J. Champandard, Semantic Style Transfer and Turning Two-Bit Doodles into Fine Artworks. arXiv:1603.01768 (2016)
8. J. Liao, Y. Yao, L. Yuan, et al., Visual Attribute Transfer through Deep Image Analogy. arXiv:1705.01088 (2017)
9. S. Gu, C. Chen, J. Liao, et al., Arbitrary Style Transfer with Deep Feature Reshuffle. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 8222-8231 (2018)

10. J. Johnson, A. Alahi, L. Fei-Fei, Perceptual Losses for Real-Time Style Transfer and Super-Resolution. European Conference on Computer Vision. 694-711 (2016)
11. D. Ulyanov, A. Vedaldi, V. Lempitsky, Improved Texture Networks: Maximizing Quality and Diversity in Feedforward Stylization and Texture Synthesis. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 6924-6932 (2018)
12. Y. Li, C. Fang, J. Yang, et al., Diversified Texture Synthesis with Feed-Forward Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 3920-3928 (2017)
13. J.Y. Zhu, T. Park, P. Isola, et al., Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. Proceedings of the IEEE International Conference on Computer Vision. 2242-2251 (2017)
14. V. Dumoulin, J. Shlens, M. Kudlur, A Learned Representation for Artistic Style. arXiv:1610.07629 (2016)
15. D. Chen, L. Yuan, J. Liao, N. Yu, G. Hua, StyleBank: An Explicit Representation for Neural Image Style Transfer. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 1897-1906 (2017)
16. H. Zhang, K. Dana, Multi-Style Generative Network for Real-Time Transfer. arXiv:1703.06953 (2017)
17. P.A. Qiao, J.W. Li, J.L. Cao, Animation of Image Style in Multi-Channel CartoonGAN. Application Research of Computers. 38(11), 3517-3520 (2021)
18. T.Q. Chen, M. Schmidt, Fast Patch-Based Style Transfer of Arbitrary Style. Proceedings of the NIPS Workshop on Constructive Machine Learning (2016)
19. Y. Li, C. Fang, J. Yang, et al., Universal Style Transfer via Feature Transforms. Advances in Neural Information Processing Systems. 386-396 (2017)
20. S. Bagwari, K. Choudhary, S. Raikwar, P.S. Rana, S. Mighlani, A Review: The Study and Analysis of Neural Style Transfer in Image. In: Tistarelli, M., Dubey, S.R., Singh, S.K., Jiang, X. (eds) Computer Vision and Machine Intelligence. Lecture Notes in Networks and Systems, vol 586, Springer, Singapore. https://doi.org/10.1007/978-981-19-7867-8_17 (2023)