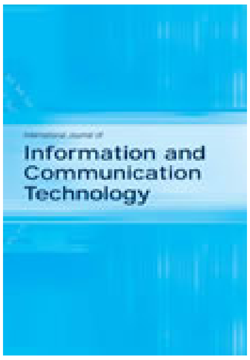


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Artistic style image migration model based on cycle-consistent generative adversarial networks

Laixi Zheng

DOI: 10.1504/IJICT.2025.10071315

Article History:

Received:	27 March 2025
Last revised:	09 April 2025
Accepted:	10 April 2025
Published online:	27 May 2025

Artistic style image migration model based on cycle-consistent generative adversarial networks

Laixi Zheng

School of Painting and Drawing,
Guangzhou Academy of Fine Arts,
Guangdong 510261, China
Email: zhenglaixi137@126.com

Abstract: Art style image migration has been increasingly important in image processing research as computer vision and deep learning technologies develop. Most well-known style migration techniques depend on paired training data, which is occasionally difficult to get in practice. Many solutions for complex art forms lose material or have style inconsistencies, which makes it challenging to fulfil high-quality content preservation. Regarding the mentioned problems, this work presents ArtCycleGAN, a cycle-consistent generative adversarial network-based art style image migration model. Pre-trained VGG19 network perceptual loss and cycle-consistent loss help to enable high quality unsupervised art style migration. ArtCycleGAN proves its validity in art style migration by experimental findings showing good performance in style similarity, content retention, and perceptual quality. This work presents a dependable and efficient approach for unsupervised art style migration as well as fresh ideas and references for picture synthesis applied with generative adversarial networks.

Keywords: art style image migration; CycleGAN; perceptual loss; unsupervised learning; image generation; content retention.

Reference to this paper should be made as follows: Zheng, L. (2025) 'Artistic style image migration model based on cycle-consistent generative adversarial networks', *Int. J. Information and Communication Technology*, Vol. 26, No. 16, pp.53–68.

Biographical notes: Laixi Zheng received his Master's degree from Guangzhou Academy of Fine Arts in 2010. Currently, he works in Guangzhou Academy of Fine Arts. His research interests include comprehensive painting research and artificial general intelligent control.

1 Introduction

Image processing and generation have advanced greatly with the fast development of computer vision and deep learning method (Jiao and Zhao, 2019). Emerging as a technique, artistic style image migration seeks to generate new images with distinctive visual effects by combining the content of one image with the artistic style of another. This approach not only gives designers and artists fresh creative tools but also demonstrates wide application possibilities in many disciplines including virtual reality, picture editing, and game creation. Achieving high-quality art style migration, however, is not a simple process; it calls for exact integration of elements of the target art style, such colours, textures, and brushstrokes, so preserving the semantic meaning of the source image. Many scholars have been motivated by this difficulty to investigate closely related approaches.

Early studies in the field of art style image migration concentrated on manual feature extraction and basic image fusion techniques, however these approaches sometimes struggle to simultaneously satisfy the needs of content preservation and style fusing (Xu and Fu, 2023). With the advent of deep learning methods-particularly generative adversarial networks (GANs)-breakthroughs in the domains of picture production and manipulation have emerged recently (Chen et al., 2024). By means of adversarial training of generators and discriminators, GANs can create realistic visuals, thereby offering fresh possibilities for artistic style migration. Nevertheless, conventional GAN techniques typically demand a lot of paired training data to understand the mapping relationships between images, which is sometimes challenging to get in useful applications. CycleGAN has surfaced as a solution for this challenge. By adding cyclic consistency loss, which allows high-quality image-to-image transformation without paired data, CycleGAN presents a fresh path for art style migration (Guo et al., 2024).

In related work, other GAN-based style migration techniques have been investigated. Although it performs well under supervised settings, Pix2Pix is a conditional generative adversarial network (cGAN) that learns the mapping relationship between input and target images through pairs of training data, hence data availability in unsupervised applications limits it (Araujo et al., 2023). On the other hand, style generative adversarial network (StyleGAN) requires high quality and quantity of training data but has great capacities in the style migration task and can create high-quality images by introducing a styled generator (Melnik et al., 2024). DeepArt also makes stylised images by means of an optimisation technique using pre-trained convolutional neural networks (CNN), however its computational efficiency is limited and processing vast-scale image data in real-time presents challenges.

Apart from GAN-based techniques, several algorithms have also shown noteworthy success in the domain of art style migration. For instance, neural style transfer (NST) maximises the feature representation of the target image so that it is near to the input image in terms of content and close to the target style image in terms of style (Jing et al., 2019). Although this approach is somewhat common in art style migration, its produced photos can suffer from blurriness of details. Attention Mechanisms have lately been included into the style migration model, which concentrates on the important areas of the produced images so improving their quality and detail performance (Liu et al., 2022). Furthermore employed for the style migration task is the Transformer architecture, whose parallel processing capacity and ability to model long distance dependencies let the model more effectively capture the stylistic traits of a picture.

Although these methods have considerable success, their major emphasis is on particular art styles or datasets and they still have restrictions when addressing challenging styles or cross-domain projects. For highly abstract art styles (e.g., Picasso style), some techniques might, for instance, fail to faithfully restore the details of the target style or lose the semantic value of the original image (Satchidanandam et al., 2023). Furthermore, current approaches provide difficulties for training efficiency and resource usage, particularly in relation to big-scale datasets. Thus, a hot and challenging question in present research is how to create a model that can fluidly migrate between several artistic forms and sustain high-quality production.

This work presents a CycleGAN-based art style image migration model ArtCycleGAN, which is able to produce high-quality art style migration in an unsupervised manner by introducing cyclic consistency loss and perceptual loss, so addressing the aforesaid issues. This paper's innovations consist in the following:

1. Optimisation of model architecture: using more effective convolutional layers and normalisation methods, ArtCycleGAN has maximised the architectural design of the generator and discriminator. These developments not only raise the model's training efficiency but also help it to learn intricate style elements, therefore enabling it to perform well in extremely abstract or complex styles.
2. Implementation of unsupervised learning: unlike conventional supervised style migration techniques, ArtCycleGAN achieves excellent style migration without the necessity of pairs of training data. This invention helps the model to be more widely employed in real-world circumstances, especially when paired data is difficult to get, therefore considerably lowering the effort and cost of data preparation.
3. Adaptability to multi-style migration: strong adaptability of the ArtCycleGAN model in several art style migration challenges is shown. The model is able to dynamically handle the migration challenges of several styles, including but not limited to Van Gogh style, Monet style, and Picasso style, by varying the weights of cyclic consistency loss and perceptual loss. This multifarious adaption offers a broad and effective answer for the topic of art style migration.
4. Perceptual loss calculation: ArtCycleGAN adds perceptual loss to help produced photos to be even more of quality. By use of a pre-trained VGG19 network, this loss extracts high-level aspects of an image, therefore optimising the produced results by means of feature differences between the produced image and the target style image.

2 Cycle-consistent generative adversarial networks

Widely applied in picture generating tasks, GAN is a model built by generator and discriminator adversarial training. Targeted especially at unsupervised image-to-image transformation tasks, CycleGAN is an expansion of the conventional GAN. CycleGAN solves the problem of unpaired data by adding the cyclic consistency loss and can achieve source-to- target domain picture transformation without paired data unlike normal GANs. Specifically, each generator guarantees that the converted image recovers as much of the original image as feasible by adding the cyclic consistency loss; the source-domain image $x \in X$ is mapped to the target-domain image $G(x) \in Y$, the target-domain image $y \in Y$ is converted to the source-domain image $F(y) \in X$.

First of all, the generator's loss function is grounded on GAN's traditional adversarial training goal. The generators in adversarial training aim to maximise the misclassification of the produced image by the discriminator (Sampath et al., 2021). The source domain generator G specifically aims to make the discriminator D_y unable to differentiate the produced image $G(x)$ from the true target domain image y. The adversarial loss of generator G therefore is:

$$L_G = \mathbb{E}_{x \sim P_x}\left[\log\left(1 - D_y(G(x))\right)\right] \tag{1}$$

where P_x represents the actual data distribution in the source domain. The generator aims to maximise the loss so that the created image is as real as feasible; the loss gauges the discriminator's capacity to discern the generated image as real. This will help to trick the discriminator.

Analogous to this, the aim of the target domain generator F is to translate the target domain picture y into the source domain image $F(y)$ so rendering the discriminator D_x incapable of discriminating between the produced image $F(y)$ and the real source domain image x. The target domain generator suffers adversarial loss of:

$$L_F = \mathbb{E}_{x\sim P_x}\left[\log\left(1 - D_x(F(y))\right)\right] \tag{2}$$

where P_y reflects the actual data distribution within the target domain. Like generator G, generator F must maximise this loss, hence the discriminator must misinterpret its produced images.

Apart from the adversarial loss, CycleGAN presents one of its main innovations-the cyclic consistency loss (Vasantrao et al., 2024). By use of the cyclic consistency loss, the picture can be restored to the original state as much as feasible following generator transformation, therefore preventing the loss of content information throughout the transformation process. The generator G transforms the source domain picture x to the target domain image $G(x)$, and subsequently the target domain generator F transforms $G(x)$) back to the source domain image $F(G(x))$. Defining the cyclic consistency loss helps to guarantee that the image's content is not lost:

$$L_{cycle}(x, F(G(x))) = \left\|x - F(G(x))\right\|_1 \tag{3}$$

This loss metric gauges the image converted by generators G and F's difference from the source domain image x. Minimising this disparity will help to produce a picture converted by the generators as near to the original image (Baraheem et al., 2023).

The generator F turns the target domain image y into the source domain image $F(y)$, then the generator G turns $F(y)$ back to the target domain image $G(F(y))$:

$$L_{cycle}(y, G(F(y))) = \left\|y - G(F(y))\right\|_1 \tag{4}$$

It gauges the variations between the target domain picture y and the generated image by generators F and G. This loss distinguishes the target domain image y from the generated image by generators F and G. Once more, the aim is to reduce this difference so as to maintain the desired image's structure and substance.

CycleGAN's generator loss is the weighted sum of its adversarial loss and cyclic consistency loss; the total losses of generators G and F are, respectively:

$$L_G^{total} = L_G + \lambda L_{cycle}\left(x, F(G(x))) + \lambda L_{cycle}(y, G(F(y))\right) \tag{5}$$

$$L_F^{total} = L_F + \lambda L_{cycle}\left(x, F(G(x))) + \lambda L_{cycle}(y, G(F(y))\right) \tag{6}$$

where λ is a hyperparameter balancing the weights of the cyclic consistency loss and the adversarial loss. Simultaneous minimising of these two components of the loss helps the generator to provide reasonable target domain images and preserve content consistency.

Additionally included in CycleGAN are two discriminators for the target and source domains (Li et al., 2020). The target domain discriminator D_y and the source domain discriminator D_x have respectively loss functions:

$$L_{D_y} = \mathbb{E}_{y \sim P_y}\left[\log D_y(y)\right] + \mathbb{E}_{x \sim P_x}\left[\log\left(1 - D_y(G(x))\right)\right] \tag{7}$$

$$L_{D_x} = \mathbb{E}_{x \sim P_x}\left[\log D_x(x)\right] + \mathbb{E}_{y \sim P_y}\left[\log\left(1 - D_x(F(y))\right)\right] \tag{8}$$

The discriminator's objective is to minimise the likelihood that a created sample is judged to be real while maximising the likelihood that a real sample is assessed to be real (Zhang et al., 2019). By means of adversarial training, the discriminator keeps raising its capacity to distinguish generated from genuine data.

CycleGAN thus maximises the generator's and the discriminator's training process on this premise. To provide better picture transformations in the training phase of the model, the generator seeks to minimise both the adversarial loss and the cyclic consistency loss. The last training objective is:

$$L_{total} = L_G^{total} + L_F^{total} + L_{D_y}^{total} + L_{D_x}^{total} \tag{9}$$

This goal maximises the generator and discriminator losses to generate more realistic and structurally consistent target images.

By adding cyclic consistency loss and bi-directional generator structure, CycleGAN finally effectively addresses the unsupervised image transformation problem. It can accomplish excellent picture transformations in several domains and does not depend on paired data. CycleGAN's creativity offers fresh concepts for unsupervised learning and advances generative adversarial network technology's ongoing growth.

3 Artistic style image migration model based on CycleGAN

3.1 Artistic style image migration

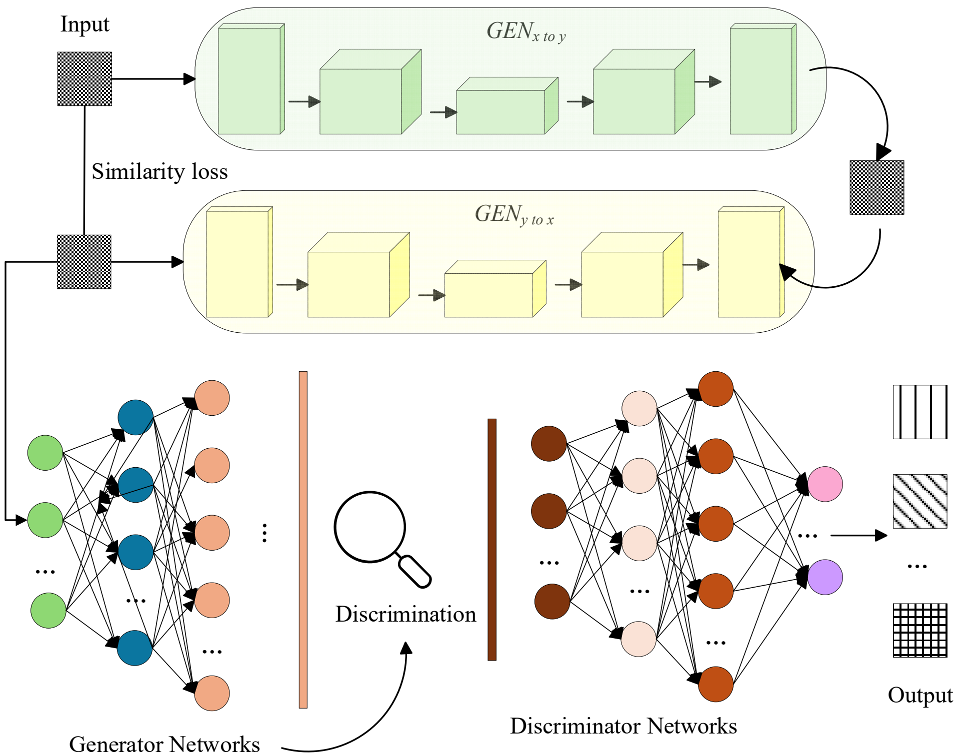
Using computer vision methods, artistic style image migration is the process of combining the content of one image with the creative style of another. The aim is to create an image that combines elements of both the creative style (e.g., colours, textures, brushstrokes, etc.) of the target image with the structural, shape, and semantic contents of the input image (Bai and Li, 2023). This method is frequently used to translate real-world images into works with an artist's style, as a Van Gogh or Monet painting from a current snapshot.

In art style picture migration, first typically via an intermediary layer in a CNN to capture the content structure of the image, the content features of the input image must be retrieved from it first. Stylistic elements then are taken from the target art style image to explain image colour distribution, texture, and detail information. Optimising the technique produces a new image that maintains the structure of the input image while including the visual characteristics of the target art style, therefore attaining the ideal balance of both content and style.

3.2 Migration modelling

Based on CycleGAN, image migration model in artistic styles several key modules make up ArtCycleGAN, each of which collaboratively helps to migrate a source image to a goal artistic style image. Combining the elements of the target art style with the content of the source image will help the model to create an image that preserves both the target style and the content information. See Figure 1 for specifically an image conversion module, a style preservation module, a cyclic consistency module, and a perceptual loss module in the model.

Figure 1 Art style image migration model ArtCycleGAN (see online version for colours)



3.2.1 Image conversion module

The execution of art style picture migration depends much on the image conversion module (Mao et al., 2019). Its principal objectives are to simultaneously translate the target domain picture y back to the source domain image $F(y)$ and convert the source domain image x to the target style image $G(x)$. This technique guarantees that the converted image stays aesthetically natural and consistent in addition to combining the content of the source image with the artistic elements of the target image.

First, using the generator G, the source image x is mapped to the target stylised image $G(x)$ by means of the formula:

$$x \to G(x) \tag{10}$$

where $G(x)$ is the target style picture acquired by generator G. This approach teaches the generator G how to translate the visual impacts of the target style from the picture elements of the source domain.

Second, generator F transforms the target style picture y back to the source domain image $F(y)$ using indicated formula:

$$y \to F(y) \tag{11}$$

This procedure guarantees that, notwithstanding style migration, the reverse conversion preserves the content elements of the original image, therefore preventing content loss or distortion resulting from over-stylisation.

3.2.2 Style preservation module

The main goal of the module on style preservation is to guarantee that the produced image reflects the visual traits of the intended artistic style. Style loss, unlike pixel-level similarity comparisons, computes the high-level feature differences of an image in a deep convolutional neural network, therefore determining its stylistic similarity. This approach can efficiently capture high-level visual characteristics including texture, colour distribution, form and artistic style of a picture, therefore enabling the produced image to not only be comparable to the source image in terms of content but also can demonstrate the uniqueness of the target creative style.

Calculating the difference between the produced picture $G(x)$ and the desired style image y in the feature maps extracted at every layer of the network helps one to acquire the style loss. The formula for figuring the style loss is:

$$L_{style} = \sum_{l} \lambda_l \left\| F_l(G(x)) - F_l(y) \right\|_2^2 \tag{12}$$

where $F_l(y)$ indicates the feature maps of the target style image y at the same level, while $F_l(G(x))$ indicates the features extracted from the generative image $G(x)$ in the l^{th} layer of the convolutional network; λ_l is the weight factor of each layer to control the influence of the features at different layers.

Furthermore, the style preservation module adds a normalised Gramme matrix to evaluate the similarity of the picture style features so ensuring that the produced image can be more precisely consistent with the target style image (Guo et al., 2023). Since the Gramme matrix can efficiently record the texture correlation between the individual channels in the image, it is extensively applied in the image style migration task.

The Gramme matrix computation yields:

$$g_l = F_l^T F_l \tag{13}$$

where g_l is the layer l Gramme matrix and F_l is the feature map matrix of layer l. Calculating the Gramme matrix allows one to capture the link between the channels in the image, therefore optimising the style loss and producing an image with stylistically closer proximity to the goal art style (Yu and Zheng, 2023).

3.2.3 Cycle consistency module

The CycleGAN model, which seeks to guarantee that the contents of the source and destination images remain constant during the image style migration process, consists fundamentally on this module. CycleGAN guarantees that, following style migration, the source image can be restored to the original image as much as possible after style conversion and inverse conversion, and the target image can be restored to its natural look after similar operations by including the cyclic consistency loss.

One may express the cyclic consistency loss of the so
urce image as:

$$L_{cycle}(x) = \mathbb{E}_{x \sim P_x}\left[\left\|F(G(x)) - x\right\|_1\right] \tag{14}$$

where $\|F(G(x)) - x\|$ signifies the L1 distance between the produced image $F(G(x))$ and the original source image x, $L_{cycle}(x)$ evaluates the recovery error of the source image x after style transformation and inverse transformation. CycleGAN guarantees that the content of the source picture is retained following transformation by reducing this loss.

The target image's cyclic consistency loss subsequently is:

$$L_{cycle}(y) = \mathbb{E}_{y \sim P_y}\left[\left\|G(F(y)) - y\right\|_1\right] \tag{15}$$

The aim is to guarantee the similarity of the produced image $G(F(y))$ to the target image y. $L_{cycle}(y)$ measures the recovery error of the target image y after source style transformation and reverse transformation; $G(F(y))$ is the reverse transformed image.

CycleGAN generates images with the target artistic style while preserving the integrity of the source image content by optimising the cyclic consistency loss, therefore enabling seamless mapping between source and target images without the necessity of paired data.

3.2.4 Perceptual loss module

Aiming to maximise the image quality by comparing the high-level feature differences between the produced image and the target image in the deep network, the perceptual loss module is absolutely important in art style image migration. Unlike the conventional pixel-level loss, perceptual loss can better maintain the high-level semantic information of the image, such as texture and shape, so producing a generated image more natural and realistic in the process of style migration.

In this model, perceptual loss is calculated by extracting the features of an image through a pre-trained VGG19 network (Nguyen et al., 2022). It is a generally utilised deep network trained in a great number of image classifications, which is capable of effectively capturing texture, shape and other high-level properties of an image. In particular, the model computes the perceptual loss by means of the middle layer feature map of the VGG19 network by differentiating the produced image from the target image.

The calculated perceptual loss is:

$$L_{perceptual}(G(x), y) = \sum_{l} \lambda_l \left\|\phi_l(G(x)) - \phi_l(y)\right\|_2^2 \tag{16}$$

where $G(x)$ is the produced image; y is the target style image; ϕ_1 is the VGG19 layer's features; λ_l is the layer's weighting coefficient. The model can guarantee that the produced image is perceptually comparable to the target image by computing the variation between the generated image and the target image in every level of characteristics.

This work computes the perceptual loss at several levels and weights the total to help to maximise the quality of the produced image. Calculating the loss on several layers of the feature map helps the model to get more perceptual information from several perspectives (Shi and Pun, 2019). One might say this as:

$$L_{total} = \sum_{l \in L} \lambda_l \left\| \phi_l(G(x)) - \phi_l(y) \right\|_2^2 \tag{17}$$

where L is the set of layers selected, often consisting of several low to high level convolutional layers. Different weight λ_l is assigned to every layer of the feature map difference computation, therefore controlling the impact of features from various layers and producing the resulting image with best overall perceptual quality.

Optimising the perceptual loss helps the produced image to remain more visually natural and preserve the high-level elements of the target image during the style migration process. The addition of this module not only solves the issue of conventional pixel-level loss avoiding quality degradation but also greatly enhances the creative impact of the image.

Through the joint activity of four modules, this model delivers effective creative style image migration overall. First, the image conversion module achieves the conversion between the input image and the target style image; second, the style preservation module ensures that the generated image can effectively retain the visual features of the target style; the cyclic consistency module further guarantees the consistency of the content between the generated image and the original image; and lastly, the perceptual loss module improves the generating of the generated image by means of high-level feature differences between the generated image and the target image in the deep network. image's perceived quality. The four modules' synergy helps the model to keep artistic style and image quality during style migration, so offering strong assistance for image editing and production.

4 Experimental results and analyses

4.1 Experimental data

The key dataset for this experiment came from the met museum dataset. The metropolitan museum of art supplies the dataset, which features artworks spanning several historical eras and cultural contexts. Renaissance, baroque, impressionist, modern, and Asian art among other artistic styles and genres abound in the approximately 40,000 pieces of art that exist overall. Rich diversity of these works offers a range of art styles for style migration, including both classic Western art and significant pieces from non-Western civilisations.

High picture resolution of the dataset is appropriate for tasks involving image production and style migration. Every piece offers excellent photos, but it also comes with comprehensive metadata including artist name, date of creation, art genre, medium, size, and information on the work. Particularly suited for training and testing of style migration models, the art forms and genres contained in the dataset span a large time range and may efficiently facilitate migration and modification between several styles.

Table 1 The met museum dataset information

Description	*Number of images*	*Art styles and categories*	*Art movements*
A collection of artworks from the Metropolitan Museum of Art, covering a wide range of historical periods and regions. The dataset is suitable for tasks like style transfer, art style analysis, and image generation	Approximately 40,000	Includes Western and non-Western art styles such as Renaissance, Baroque, Impressionism, Modern Art, and Asian Art	Renaissance, Baroque, Impressionism, Modern Art, Asian Art, etc.

4.2 Performance comparison of different model structures

This work designs Experiment 1, which compares the performance of the ArtCycleGAN model with that of several other common style migration models (including StandardGAN, Pix2Pix and StyleGAN), so enabling a thorough evaluation of the performance of the proposed Art CycleGAN model in the art style image migration task.

Used for model training and performance evaluation respectively, the dataset consists of a training set and a test set. While the conventional GAN, Pix2Pix, and StyleGAN are trained according on their traditional architectures, the ArtCycleGAN model is trained according to the architecture and training method provided in this work. Every model uses the same training set and the training time is noted.

Regarding performance assessment, this work evaluates the models using the three following criteria:

Style similarity is the first one calculating the style loss between the produced image and the desired style image helps one to determine style consistency (Ruder et al., 2018). The style loss is computed using the Gramme matrix's difference applied using the following formula:

$$L_{style} = \frac{1}{C^2} \sum_{i,j} \left(G_{ij}^{gen} - G_{ij}^{style} \right)^2 \tag{18}$$

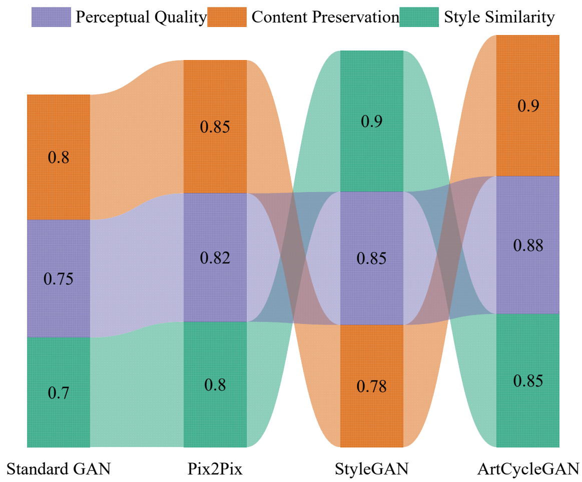
where C is a normalisation constant to scale the loss and G_{ij}^{gen} and G_{ij}^{style} respectively refer to the Gramme matrix elements of the produced image and the target style image.

- Content conservation comes next. Content preservation uses cyclic consistency loss to ensure the produced image has the content information of the source image (You et al., 2019).
- At last, perceptual quality. Perceptual loss allows one to assess the general quality of the produced image (Martin-Donas et al., 2018). The perceptual loss is computed using the pre-trained VGG19 network's feature map difference in line with this formula:

$$L_{perceptual} = \sum_{l} \lambda_l \left\| \phi_l(G(x)) - \phi_l(y) \right\|_1 \tag{19}$$

The experimental results are shown in Figure 2:

Figure 2 Comparative results of the performance of different models on image migration (see online version for colours)



ArtCycleGAN has good performance in the three criteria of style resemblance, content retention, and perceptual quality based on experimental data. ArtCycleGAN performs especially in content retention (0.90), meaning that, when moving the structural and semantic information of the original image content to the target style, the model is able to efficiently maintain them. Though StyleGAN earns the maximum 0.90, ArtCycleGAN also exhibits great similarity (0.85) and also achieves 0.88 in perceptual quality, second only to StyleGAN's 0.85. In terms of style similarity. By comparison, notably in terms of content retention and perceived quality, regular GAN and Pix2Pix perform somewhat less well in these criteria. This implies that ArtCycleGAN clearly excels in terms of overall performance and can balance style migration and content retention to preserve high quality of produced photos.

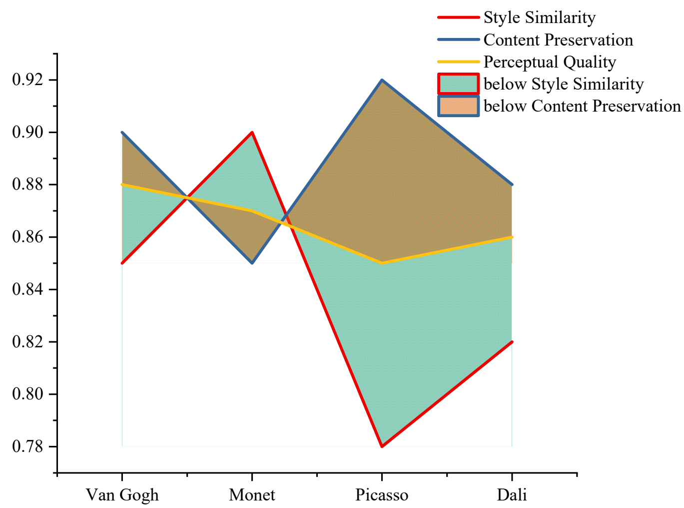
In terms of information retention and perceived quality particularly, the ArtCycleGAN model shows outstanding performance in the art style picture migration job. These findings confirm the efficiency and simplicity of the Art CycleGAN model in the unsupervised style migration task, therefore offering a dependable and quick method in the field of art style picture production.

4.3 Comparison of migration effects of different art styles

This work planned the experiment with the intention of evaluating the variations in the migration effects of several art styles so more fully evaluate the performance of the ArtCycleGAN model in the art style image migration task. Analysing the performance of the model on a range of art forms (e.g., Van Gogh style, Monet style, Picasso style and Dalí style) helps one to evaluate its performance on three main criteria: style resemblance, content retention and perceived quality. By means of these measures, one can grasp the flexibility and stability of the model in various art style migration jobs, so guiding the selection of an appropriate model for a given style in pragmatic uses. Furthermore, the outcomes of this experiment can be combined with past performance comparison studies using various model architectures to assess the whole performance of the ArtCycleGAN model more generally.

First, in order to guarantee that these photographs have clear semantic meaning and structure, a collection of high-quality source images was chosen for this work. From The Met Museum Dataset, then, a range of typical art forms was chosen: Van Gogh style, Monet style, Picasso style, Dali style. The ArtCycleGAN model is trained for every style using the same source picture set and style migration generates output images of the matching styles. Figure 3 shows the experimental findings:

Figure 3 Experimental results of migration effects of different art styles (see online version for colours)



The ArtCycleGAN model demonstrates, from the experimental results, some degree of adaptability and stability in various art style migration challenges. Regarding style similarity, the Monet style migration task performs best at 0.90, meaning the model is able to better capture the Monet style traits. With a style similarity of 0.85, the Van Gogh style migration job has great style consistency as well. Picasso's style is somewhat less comparable at 0.78, though this could be connected to the complexity and abstraction of his work. With regard to content retention, the Picasso style migration job fared the best at 0.92, suggesting that the model was able to keep the content of the source image better when moving the image to the Picasso style. Though it is somewhat less than the Picasso style, the Monet style migration work exhibits great content consistency with a content retention of 0.85. The Monet style migration challenge has a perceptual quality of 0.87, which is rather higher than the other styles and indicates that the produced images are more visually natural. With a 0.88 perceived quality, the Van Gogh style migration assignment likewise exhibits good visual quality. Though there is still space for development in some styles (e.g., Picasso style), the ArtCycleGAN model exhibits good general performance in the several art styles migration challenge. These findings offer useful references for additional model optimisation and a foundation for selecting an appropriate model for particular styles in pragmatic use.

Figure 4 and 5 respectively highlight the source images and the outcomes of ArtCycleGAN model migration to Van Gogh and Monet styles.

Figure 4 Example of van Gogh art style image migration effect (see online version for colours)

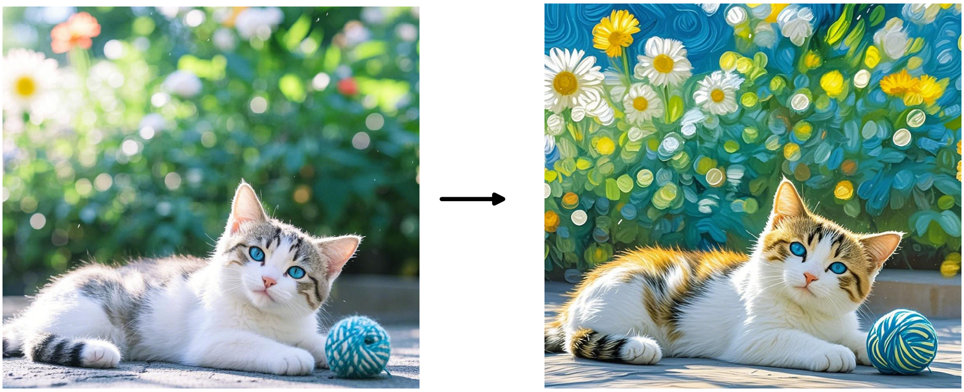


Figure 5 example of Monet art style image migration effect (see online version for colours)

5 Conclusions

This work investigates the performance and application possibilities of ArtCycleGAN, a CycleGAN-based art style picture migration model, in unsupervised image-to-image conversion tasks. The ArtCycleGAN model shows great performance in a range of art style migration tasks by including cyclic consistency loss and perceptual loss, especially it is able to effectively fuse the visual features of the target art style while keeping the content structure of the source image. Verifying its efficacy and applicability in the field of art style migration, experimental results reveal that the model obtains outstanding outcomes in three fundamental criteria: namely, style similarity, content retention, and perceptual quality.

The ArtCycleGAN model still has significant restrictions, nevertheless, even if it performs quite well in many tests. First of all, the model's performance in terms of style similarity and content retention is really inadequate when handling some sophisticated or extremely abstract art forms. Second, especially in regard to big-scale datasets, the training procedure of the model calls for substantial computational resources; so, the training time and resource consumption may constitute a bottleneck in useful applications. Furthermore, especially in some style migration jobs where the produced images may be blurry or distorted, the present model still has potential for development in the detail and texture performance of the created images. Particularly in real-time and resource-limited contexts, these problems restrict the implementation of the model across a greater spectrum of circumstances.

Future studies in the following spheres will help to solve these constraints:

1 Improved model architecture: investigate more effective model architectures, including the addition of an attention mechanism or a Transformer structure, to improve the model's capacity to capture local information and detail for the purpose of migrating sophisticated artistic styles. When considering extremely abstract forms, this will help the model perform better.

2 Optimising training strategies: to cut training time and resource consumption, look at more effective training techniques such dynamically changing the learning rate, improving hyper-parameter selection and using more complex optimising algorithms (Andonie, 2019). Furthermore, investigating how to reach quick convergence of the model with constrained computational capacity would help to improve the actual application practicality of the model.

3 Multimodal data fusion: to offer better input information, think about merging image data with additional modal information, such textual descriptions or user comments. By means of this multimodal data fusion technique, the migration of the model may be further improved, thereby enabling the creation of art-style images more in line with user expectations.

In summary, this work obtained considerable results in the art style picture migration problem with the ArtCycleGAN model, while also pointing out the limitations of the model and suggesting future research possibilities. These works not only provide new technological approaches in the subject of art style migration, but also provide important references for GANs in a wider variety of application scenarios.

Declarations

All authors declare that they have no conflicts of interest.

References

Andonie, R. (2019) 'Hyperparameter optimization in learning systems', *Journal of Membrane Computing*, Vol. 1, No. 4, pp.279–291.

Araujo, G.F., Machado, R. and Pettersson, M.I. (2023) 'Synthetic SAR data generator using Pix2pix cGAN architecture for automatic target recognition', *IEEE Access*, Vol. 11, pp.143369–143386.

Bai, S. and Li, P. (2023) 'Algorithm and simulation study of oil painting classification based on visual perception and improved embedded learning', *Journal of Intelligent and Fuzzy Systems*, Vol. 45, No. 6, pp.9979–9989.

Baraheem, S.S., Le, T-N. and Nguyen, T.V. (2023) 'Image synthesis: a review of methods, datasets, evaluation metrics, and future outlook', *Artificial Intelligence Review*, Vol. 56, No. 10, pp.10813–10865.

Chen, H., Wei, J., Huang, H., Yuan, Y. and Wang, J. (2024) 'Review of imbalanced fault diagnosis technology based on generative adversarial networks', *Journal of Computational Design and Engineering*, Vol. 11, No. 5, pp.99–124.

Guo, L., Chuah, J.H., Raymond, W.J.K., Gu, X., Yao, J. and Chang, X. (2024) 'Unsupervised feature-preserving CycleGAN for fault diagnosis of rolling bearings using unbalanced infrared thermal imaging sample', *IEEE Access*, Vol. 12, pp.28449–28461.

Guo, Z., Zhu, Z., Li, Y., Cao, S., Chen, H. and Wang, G. (2023) 'AI assisted fashion design: a review', *IEEE Access*, Vol. 11, pp.88403–88415.

Jiao, L. and Zhao, J. (2019) 'A survey on the new generation of deep learning in image processing', *IEEE Access*, Vol. 7, pp.172231–172263.

Jing, Y., Yang, Y., Feng, Z., Ye, J., Yu, Y. and Song, M. (2019) 'Neural style transfer: a review', *IEEE Transactions on Visualization and Computer Graphics*, Vol. 26, No. 11, pp.3365–3385.

Li, X., Luo, M., Ji, S., Zhang, L. and Lu, M. (2020) 'Evaluating generative adversarial networks based image-level domain transfer for multi-source remote sensing image segmentation and object detection', *International Journal of Remote Sensing*, Vol. 41, No. 19, pp.7343–7367.

Liu, P., Zhang, C., Qi, H., Wang, G. and Zheng, H. (2022) 'Multi-attention DenseNet: A scattering medium imaging optimization framework for visual data pre-processing of autonomous driving systems', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 12, pp.25396–25407.

Mao, H., She, J. and Cheung, M. (2019) 'Visual arts search on mobile devices', *ACM Transactions on Multimedia Computing, Communications, and Applications* (*TOMM*), Vol. 15, No. 2s, pp.1–23.

Martin-Donas, J.M., Gomez, A.M., Gonzalez, J.A. and Peinado, A.M. (2018) 'A deep learning loss function based on the perceptual evaluation of the speech quality', *IEEE Signal Processing Letters*, Vol. 25, No. 11, pp.1680–1684.

Melnik, A., Miasayedzenkau, M., Makaravets, D., Pirshtuk, D., Akbulut, E., Holzmann, D., Renusch, T., Reichert, G. and Ritter, H. (2024) 'Face generation and editing with stylegan: a survey', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 46, No. 5, pp.3557–3576.

Nguyen, T-H., Nguyen, T-N. and Ngo, B-V. (2022) 'A VGG-19 model with transfer learning and image segmentation for classification of tomato leaf disease', *AgriEngineering*, Vol. 4, No. 4, pp.871–887.

Ruder, M., Dosovitskiy, A. and Brox, T. (2018) 'Artistic style transfer for videos and spherical images', *International Journal of Computer Vision*, Vol. 126, No. 11, pp.1199–1219.

Sampath, V., Maurtua, I., Aguilar Martin, J.J. and Gutierrez, A. (2021) 'A survey on generative adversarial networks for imbalance problems in computer vision tasks', Journal of Big Data, Vol. 8, pp. 1-59.

Satchidanandam, A., Al Ansari, R.M.S., Sreenivasulu, A., Rao, V.S., Godla, S.R. and Kaur, C. (2023) 'Enhancing style transfer with GANs: perceptual loss and semantic segmentation', *International Journal of Advanced Computer Science and Applications*, Vol. 11, pp.321–329.

Shi, C. and Pun, C-M. (2019) 'Adaptive multi-scale deep neural networks with perceptual loss for panchromatic and multispectral images classification', *Information Sciences*, Vol. 490, pp.1–17.

Vasantrao, C.P., Gupta, N., Mishra, A.K., Bhavekar, G.S. and Gupta, M.K. (2024) 'Change detection using multispectral satellite images: a systematic review of literature', *Bulletin of Electrical Engineering and Informatics*, Vol. 13, No. 4, pp.2496–2507.

Xu, W. and Fu, Y. (2023) 'Deep learning algorithm in ancient relics image colour restoration technology', *Multimedia Tools and Applications*, Vol. 82, No. 15, pp.23119–23150.

You, C., Li, G., Zhang, Y., Zhang, X., Shan, H., Li, M., Ju, S., Zhao, Z., Zhang, Z. and Cong, W. (2019) 'CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE)', *IEEE Transactions on Medical Imaging*, Vol. 39, No. 1, pp.188–203.

Yu, L. and Zheng, Q. (2023) 'AI-Enhanced digital creativity design: content-style alignment for image stylization', *IEEE Access*, Vol. 11, pp.143964–143979.

Zhang, T-F., Tilke, P., Dupont, E., Zhu, L-C., Liang, L. and Bailey, W. (2019) 'Generating geologically realistic 3D reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks', *Petroleum Science*, Vol. 16, No. 3, pp.541–549.