International Journal for Housing Science and Its Applications

Publish September 9, 2025. Volume 47, Issue 1 Pages 214-224

https://doi.org/10.70517/ijhsa47118

Utilizing AI technology to modernize traditional oil painting techniques and promote the development of art

Chen Jiang[1,*]

[1] Academy of Fine Arts, Capital Normal University, Beijing, 100000, China

Corresponding authors: (e-mail: jkleeeiii@163.com).

Abstract As a time-honored and technically complex art form, oil painting is undergoing unprecedented transformation in its creative process and aesthetic value under the influence of AI technology. This study explores the modernization and innovation of oil painting creation from an AI technology perspective, as well as its role in driving artistic development. It primarily achieves oil painting style transfer through the optimization of the CycleGAN (Cycle-Consistent Generative Adversarial Network), proposes the introduction of spectral normalization processing, and improves the residual structure. Finally, the algorithmic model is validated through quantitative and qualitative experiments. The experiments demonstrate that the model generates oil painting images with outstanding performance in terms of clarity, diversity, and style similarity. Its IS values are 4.17% to 172.02% higher than those of the comparison methods, while its ID values are reduced by 2.99% to 63.50%. Additionally, the subjective quality evaluation and style similarity evaluation of the generated images are optimal. The model constructed in this paper improves the visual effects of image style transfer and can be used to assist artists in oil painting creation, inspiring their creative inspiration.

Index Terms CycleGAN, style transfer, AI technology, oil painting creation

I. Introduction

Oil painting originated in Europe during the 15th century, and oil painting techniques represented a revolutionary innovation in the medium of painting at the time. Artists began using oil-based pigments, which take a long time to dry and produce unique color layers and depth, thereby greatly expanding the possibilities of artistic expression [1], [2]. The core of traditional oil painting techniques lies in their precise control of color saturation and delicate texture, which is made possible by the unique properties of oil-based pigments. These pigments allow artists to create rich visual effects and a sense of depth through layered techniques [3]-[6]. From the perspective and detailed figure depiction of the Renaissance, to the dramatic portrayal of light, shadow, and movement in the Baroque and Rococo periods, to the Impressionists' spontaneous grasp of light and color, the painting techniques and styles of each period deeply reflect the cultural spirit and aesthetic trends of their time [7]-[10].

The integration of traditional oil painting techniques with modern technology is a significant issue in the development of oil painting. Currently, oil painters are exploring how to blend traditional techniques with modern technology to create works that align with contemporary trends [11], [12]. This exploration not only enriches the artistic language of oil painting but also drives its international development [13]. Traditional techniques embody the wisdom of thousands of years of painting art, while modern technology represents the cutting-edge exploration of contemporary art. The integration of the two is both a continuation and innovation of traditional culture and an active response to the development of modern art [14]-[17]. This integration is not a simple addition but involves understanding the characteristics of each and finding their common ground to achieve artistic elevation [18], [19].

With the continuous development of technology, artificial intelligence is profoundly influencing the creation of oil paintings. For example, artists use deep learning algorithms to collect a large number of existing images to form a dataset, and then perform preprocessing, classification, and recognition on the images. Kim, J et al. believe that the visual depth information of oil painting brushstrokes can effectively enhance the classification performance of machine learning models for oil paintings, and by utilizing reflection transformation imaging (RTI) image training methods, they achieved more precise classification of oil painting works [20]. Albadarneh, I. A., and Ahmad, A. developed a feature extraction method considering aspects such as oil painting colors and textures, which is suitable for the growing demand for digital oil painting authentication, offering advantages of low cost and high classification accuracy [21]. Liao, Z et al. proposed a clustering multi-kernel learning algorithm that fuses multiple features such as oil painting color, texture, and spatial layout for classification, enabling accurate identification and classification of oil painting creators and their works [22]. Guo, B., and Hao, P. utilized a convolutional neural network (CNN) to learn existing oil painting feature information, thereby obtaining the distribution of oil painting styles from both low-level and high-level features, which facilitates the exploration of the relationship between artistic styles and art history [23]. Guo, H, et al. combined visual neural network algorithms with big data analysis technology to establish an emotional expression analysis model for oil painting creation, which can automatically and accurately synthesize high-resolution images consistent with the emotional themes of oil paintings [24]. Zhang, X explored style recognition and classification techniques for oil painting images, improving the accuracy of identifying the oil painting creation process by introducing attention mechanisms and improved residual networks, providing references for oil painting appreciation and artistic creation [25].

Some artists have also utilized deep learning algorithms and other technologies to create visual effects that cannot be achieved through traditional methods, further enhancing the generation quality of oil painting works in the field of style conversion. Chang, Y. noted that traditional oil painting style conversion algorithms often overlook the distribution of image edges, leading to blurred contours in the generated images. To address this, they established an SRCNN (Super Resolution Convolutional Neural Network) model aimed at accurately generating oil paintings in the desired painting style [26]. Li, J, and others investigated image style conversion methods from graffiti to oil paintings, introducing the "Dual Sketch-to-Painting Network (DSP-Net)" to extract sparse information from graffiti images, providing additional assistance in generating paintings with artistic style [27]. Liu, X pointed out that the oil painting creation process integrated with computer image processing technology has good efficiency and creativity. By leveraging relevant mathematical models, color conversion models, etc., the artistic value of oil paintings during the creation process can be effectively enhanced [28].

Meanwhile, by leveraging the zero-sum game between the generator and discriminator in generative adversarial networks (GANs), significant assistance is provided for generating oil paintings with artistic value. Li, M et al. proposed using a cyclic generative adversarial network (CycleGAN) to perform style conversion on acquired natural images. This method does not rely on matching relationships between images and plays a crucial role in generating oil paintings with specific styles [29]. Xu, G analyzed the application of generative adversarial networks (GANs) in painting art creation. The proposed image generation model to some extent integrates the essence of painting schools, bringing profound impacts on art creation and the art industry [30]. Lv, X, and Zhang, X studied the impact of different deep learning algorithms and generator models on training time and results for specific style images. Among these, the CycleGAN not only achieved good results in Western oil painting style conversion but also effectively imitated the style of traditional Chinese landscape painting [31]. From the above research, it can be seen that artificial intelligence is already capable of practical application in the oil painting industry. The field of oil painting will undoubtedly become increasingly intertwined with AI development, and the trend toward their integration will become increasingly evident.

This paper analyzes innovative practices combining AI and oil painting, exploring how they can jointly drive the development of art. To optimize oil painting techniques, improvements were made to the CycleGAN model. For image transfer targeting the style of oil paintings, the generator was constructed by proposing improvements to the residual structure to optimize signal propagation and reduce training errors, thereby alleviating issues such as gradient vanishing and network degradation caused by increased network depth. Spectral normalization is introduced into the discriminator to effectively control the learning range of neural network parameters, thereby enhancing the stability of the network model. Finally, a combined objective and subjective evaluation criterion is proposed. The model is trained on datasets of different styles, and the evaluation metrics of different models are compared. The performance of the model is experimentally verified and analyzed from two aspects: image quality and style similarity of generated images, to investigate the image generation effects of the constructed oil painting style transfer model.

II. AI-based oil painting innovation and artistic development

With the continuous advancement of artificial intelligence technology, AI is being applied more and more widely in artistic creation. AI painting can not only simulate the styles of human artists, but also generate unique works of art through complex algorithms. AI technology has brought new creative tools and methods to traditional oil painting, providing impetus for the development of art.

II. A. A New Perspective on Cross-Border Integration

AI, as a cross-disciplinary tool, can convert complex scientific concepts and data into visual language. For example, by using AI to analyze patterns in nature, such as geometric or biological structures, artists can transform these patterns into visual elements on a canvas. Similarly, AI can process complex mathematical algorithms and visualize them, thereby providing artists with new sources of inspiration. This fusion not only demonstrates the dialogue between oil painting creation and science but also offers viewers a new way to understand and appreciate oil paintings, transforming them from mere aesthetic objects into carriers of knowledge and ideas.

II. B. Personalized Creation

AI technology can analyze an artist's past oil paintings to learn their specific style and preferences. Based on this information, AI algorithms can customize creative tools, such as adjusting the color palette, establishing specific compositions, or generating images in a particular style. Such personalized tools are more tailored to the artist's unique style and creative needs, providing them with a more personalized and style-specific oil painting experience. In this way, AI can serve not only as an auxiliary tool but also as a partner in the artist's creative process.

II. C. Diversity and interactivity

In oil painting creation, artists can explore AI technology in various ways. This includes not only learning the styles of classic works but also conducting in-depth analyses of color, texture, and form. Artists can also use AI for trial and error and adjustments. AI's rapid processing capabilities enable artists to quickly experiment with different compositions and color combinations during the creative process, thereby improving efficiency and increasing diversity. Additionally, by integrating sensors and AI algorithms, oil paintings can respond to viewers' movements or emotional states, creating dynamic visual effects that enhance immersion and engagement. This interactivity not only gives oil paintings a modern feel but also makes viewers part of the creative process, breaking down the traditional barriers between oil paintings and their audience. AI technology provides oil painters with a vast exploration space, enabling them to transcend traditional boundaries and explore new forms of artistic expression.

II. D. Experimental and exploratory

AI not only assists artists in achieving richer visual expressions, such as generating images or patterns through algorithms, but also provides new perspectives and inspiration in composition and color application. Artists can use AI to conduct various experiments and explore artistic forms that were previously difficult to achieve or even imagine. This technology encourages artists to break free from traditional thinking constraints and create unique and innovative artworks. In terms of composition, AI can help artists analyze and simulate the effects of different compositional methods, or even propose innovative compositional solutions. In terms of color application, AI can provide artists with new color combination suggestions based on a large amount of color data, or recommend color schemes by analyzing specific emotions and themes.

II. E. Openness and inclusiveness

The introduction of AI has not only simplified the oil painting creation process, making it more accessible and easier to participate in, but has also opened up new areas of artistic exploration. This means that oil painting creation is no longer limited to traditional learning paths and elite circles, but has become more open and diverse. Ordinary people can now easily create oil paintings using AI tools to express their emotions and views, which undoubtedly adds diversity and vitality to the art world.

III. Oil painting image style transfer based on CycleGAN

Image-to-image style transfer has a lot of uses in art, but making creative oil paintings by hand takes a lot of time and effort. Plus, the biggest problem with this traditional method is that it's not very forgiving and doesn't really encourage innovation. So, this paper uses AI tech to modernize traditional oil painting techniques and comes up with a style transfer method for oil paintings based on CycleGAN (CycleGAN-OPM).

III. A. CycleGAN Model

III. A. 1) Network Structure

CycleGAN is an unsupervised image conversion network. Unlike traditional generative adversarial networks (GANs), CycleGAN does not rely on paired training samples but instead uses unpaired image data for training. This feature of CycleGAN enables the conversion of image data between two styles without requiring paired images. By autonomously learning the information features from the source domain to the target domain, it establishes a mapping function between the two domains, thereby generating images that conform to the style of the target domain.

CycleGAN consists of two generators, Gx and Gy, and two discriminators, Dx and Dy. Gx converts source domain images to target domain images, while Gy converts target domain images to source domain images. The discriminators judge the authenticity of source domain and target domain images, respectively, and guide the training of the generators by judging the differences between real images and generated images.

III. A. 2) Loss Function

The loss function of CycleGAN combines the generative adversarial network (GAN) loss and the cyclic consistency loss to ensure that the generator not only generates realistic images but also maintains consistency in bidirectional conversion.

(1) Generative adversarial loss

CycleGAN uses the GAN framework, in which the generator and discriminator compete against each other. The GAN loss consists of two parts: the generator loss and the discriminator loss. The generator's goal is to generate images that can fool the discriminator into believing that the generated images are real. The discriminator's goal is to distinguish between real images and generated images. For CycleGAN's two generators Gx and Gy and discriminators Dx and Dy, their GAN loss is defined as follows:

Gx aims to generate images from domain X to domain Y $G(x)$ and make Dy believe that $G(x)$ is real, so the loss function of generator Gx is:

$$L_{GAN}(Gx, Dy, X, Y) = Ex \sim pdata(x)[\log(Dy(G(x)))] \tag{1}$$

Here, Gx is the generated image, and $Gy(G(x))$ is the output of the discriminator, indicating the authenticity of $G(x)$. Conversely, Gy generates an image Gy from the Y domain to the X domain and makes Dx believe that Gy is authentic. Therefore, the loss function of the generator Gy is:

$$L_{GAN}(Gy, Dx, X, Y) = Ey \sim pdata(y)[\log(Dx(G(y)))] \tag{2}$$

Here, $G(y)$ is the generated image, and $Gx(G(y))$ is the output of the discriminator, indicating the authenticity of $G(y)$.

(2) Cycle Consistency Loss

Cycle consistency loss is the core component of CycleGAN. In this loss, given an image y, it is first converted from domain Y to domain X via $G(y)$, then converted back to domain Y via the generator Gx, with the goal that the result $Gx(Gy(y))$ should be close to the original image y. The cycle consistency loss corresponding to the generator $G(y)$ is:

$$Lcyc(Gy, Gx) = Ey \sim pdata(y)[\| Gx(Gy(y)) - y \| 1] \tag{3}$$

In contrast, the cycle consistency loss corresponding to generator Gx is:

$$Lcyc(Gx, Gy) = Ex \sim pdata(x)[\| Gy(Gx(x)) - x \| 1] \tag{4}$$

The total loss function is:

$$\begin{aligned} L(G_x, G_y, D_x, D_y) = L_{GAN}(G_x, D_y, X, Y) \\ +LG_{AN}(G_y D_x X, Y) \\ +\lambda L_{cyc}(G_x, G_y) \end{aligned} \tag{5}$$

III. A. 3) Generators and Discriminators

The first generator in CycleGAN adopts the ResNet network architecture. The overall structure of ResNet is primarily divided into an encoder, a bottleneck layer, and a decoder. The encoder consists of a ReflectionPad, Conv2d, InstanceNorm2d, and ReLU activation function, where the ReflectionPad is a mirror padding layer that fills the three pixel matrices around the image with mirrored content. The Conv2d and InstanceNorm2d layers are used to extract image features. The residual blocks in the bottleneck layer are typically 6 or 9 in number. When the pixel size is 128×128, the network uses 6 residual blocks for task execution, and when the pixel size is 256×256, it uses 9 residual blocks. The bottleneck layer performs deeper feature extraction on the image while preserving the structural features of the input. In the decoder, ConvTranspose is used for upsampling, gradually restoring low-resolution feature images to the same resolution as the input image, while InstanceNorm2d is used to normalize the distribution of each feature map, helping to stabilize training. The final layer uses the Tanh function to limit the output to the range [-1, 1]. This ultimately achieves the generation of images with corresponding domain features.

The second generator in CycleGAN adopts the U-Net network structure, which is U-shaped and similarly divided into encoder and decoder parts, connected via skip connections. The U-Net network structure processes images in three stages: first, a shallow network (encoder) performs initial feature extraction, i.e., downsampling. Second, a deep network (decoder) performs deconvolution to capture high-level features, gradually restoring the captured image features to their original size, i.e., upsampling. Finally, the extracted features are concatenated via skip connections.

The CycleGAN discriminator uses the PatchGAN discriminator, which extracts features from the input image through multiple layers of convolution, normalization, and activation functions. The image is divided into multiple small patches, each of which is an independent sample. By performing binary classification detection on each small patch, a binary classification score is generated, and a score matrix is then generated based on this.

III. B. Image Style Transfer Methods

III. B. 1) Normalization of the model

In order to better control the output performance of the discriminator, this paper introduces spectral normalization to limit the function gradient of the neural network. Spectral normalization constrains the Lipschitz constant of the discriminator by constraining the spectral norm of the weight matrix of the discriminator network, so that no stretching occurs for any input x. Specifically, each time the network parameter matrix W is divided by the maximum singular value $\sigma(W)$:

$$W \leftarrow \frac{W}{\sigma(W)} \tag{6}$$

The advantage of this method is that it can limit the function gradient of the neural network without disrupting the proportional relationship between parameters. However, in order to calculate the maximum spectral coefficient of the discriminator neural network, singular value decomposition of the parameter matrix must be performed for each layer of the network during each training iteration. This is particularly burdensome when the weight dimension of the network is large, as the heavy computation wastes a significant amount of computational resources. Therefore, to reduce computational load and improve runtime efficiency, the power iteration method is used to approximate the singular values. Let A be a matrix; then its maximum singular value $\sigma(A)$ satisfies:

$$\sigma(A) > \frac{\| A\xi \|_2}{\| \xi \|_2} \tag{7}$$

The power iteration method can be briefly summarized as follows: first, select a non-zero initial vector $\hat{v}$; after constructing the iteration sequence, obtain the approximate solution of the unit principal eigenvector by mutual iteration according to equation (7), and then perform the following transformation:

$$\hat{u} \leftarrow \frac{W_T \hat{v}}{\left\| W_T \hat{v} \right\|_2} \tag{8}$$

$$\hat{v} \leftarrow \frac{W_T \hat{u}}{\left\| W_T \hat{u} \right\|_2} \tag{9}$$

$$\sigma(W) \approx \hat{v} W_T \hat{u} \tag{10}$$

III. B. 2) Residual structure of the model

As models become more complex, issues such as overfitting, excessive computational load, and network degradation begin to arise, leading to the development of the improved residual network ResNetV2. In the original ResNet, batch normalization was placed after the addition operation, which hindered signal propagation and caused slow error reduction during the early stages of training. Additionally, using a linear rectification function as the activation function after the addition operation resulted in constant positive outputs, which impaired the model's representational capabilities. ResNet adopts a pre-activated residual network structure, which reduces restrictions on the range of residual function values and achieves superior performance compared to previous residual networks. The pre-activated residual network structure is shown in Figure 1.

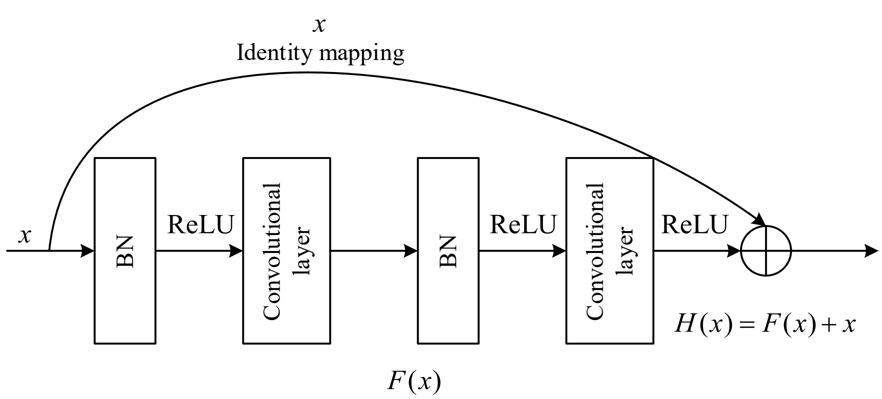

Figure 1: Improved ResNet architecture

III. B. 3) Generator and discriminator structure

In this paper, we have improved the generator under a traditional network architecture by adopting a new residual network in the encoding layer. Batch normalization normalizes individual neurons across different training datasets, while layer normalization normalizes neurons within a single layer based on the training data for that layer. For layer normalization, let the net input of the l th layer of the neural network be $z^{(l)}$, and let the total number of neurons in a layer be denoted by M. The formulas for calculating the mean and variance are as follows:

$$\mu^{(l)} = \frac{1}{M}\sum_{i=1}^{M} z_i^{(l)} \quad \sigma^{(l)} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(z_i^{(l)} - \mu^{(l)})^2} \tag{11}$$

The calculation results are expressed as:

$$\tilde{z}^{(l)} = (z^{(l)} - \mu^{(l)}) / \sqrt{\sigma^{(l)} + \varepsilon} \tag{12}$$

In cases where the batch size is small, layer normalization achieves better results than batch normalization because it extracts different features from the same samples. This effect is particularly noticeable in RNNs, but it is not as effective as BN in convolutional neural networks. Instance normalization can remove style variations by directly normalizing the statistical features of the image, and it can also accelerate model convergence, so it is commonly used in style transfer. Let t denote the index of the image, and i denote the index of the feature map. The calculation formula is as follows:

$$\mu_{ii} = \frac{1}{MW}\sum_{w=1}^{W}\sum_{m=1}^{M} x_{tiwm} \tag{13}$$

$$\sigma_{ii}^2 = \frac{1}{MW}\sum_{w=1}^{W}\sum_{m=1}^{M}(x_{tiwh} - \gamma * \mu_{ii})^2 \tag{14}$$

$$\tilde{z} = \frac{(x_{tiwh} - \mu_{ii})}{\sqrt{\sigma_{ii}^2 + \varepsilon}} \tag{15}$$

In summary, this paper incorporates an instance normalization structure into the generator network to improve its performance. The generator network structure consists of an encoding layer, a transformation layer, and a decoding layer. The encoding layer incorporates an instance normalization structure and a ReLU activation function, represented as Conv+IN+ReLU. In the transformation layer structure, nine improved ResNetV2 residual networks are used. This structure employs an instance normalization model to process feature maps, with the pre-activation function using the linear rectification function ReLU. Finally, convolution is performed using a 3×3 convolution kernel with a stride of 2 and unchanged channel count, repeating the IN-ReLU-Conv process once. The feature image obtained after the IN+ReLU+Conv operation is added to the feature image from the identity mapping. The decoder layer structure consists of two layers of De Conv-IN-ReLU and one layer of Conv-IN-ReLU.

The discriminator uses a Markov discriminator, which ultimately outputs a one-dimensional mapping to classify the image, with each output of the one-dimensional mapping corresponding to a receptive field in the original image. Additionally, the network's parameter matrix is often regularized to ensure the smoothness of the function curve. Spectral normalization functions are used to replace the batch normalization functions after the convolutional layers to achieve better regularization effects. The discriminator model consists of five layers, with convolutional layers using 4×4 convolutional kernels. The stride is set to 2 for the first three layers, with channel counts of 64, 128, and 256, respectively. The slope of the LeakyReLU function is set to 0.2.

IV. Experiments and analysis of results

IV. A. Dataset

Every artist has a specific painting style, which can be understood by studying the artist's collection of works. This paper selects the photo2vangogh and photo2monet datasets to demonstrate that the proposed model has the ability to generate high-quality visual images. All datasets contain training and testing images in two domains: one contains natural images (landscapes and scenery), and the other contains images with specific artistic styles, including Van Gogh and Monet.

IV. B. Quantitative comparison experiments and analysis

IV. B. 1) Generating image evaluation metrics

(1) Inception Score

Inception scores (IS) are one of the most commonly used and recognized metrics for evaluating the quality of generated samples. They use an image category classifier (Inception v3) to assess the quality of generated samples. Each generated image is input into the ImageNet pre-trained Inception v3 network to first obtain a probability distribution vector, which is then averaged to obtain the marginal distribution of all generated images across all categories. If the predicted IS value is high, the generated images exhibit diversity and a wide coverage range, and the distribution of the generated images across all categories should be uniform. Conversely, if the data distribution of the generated images is concentrated, it indicates that the model's output is monotonous and lacks diversity.

(2) Fréchet Inception Distance

Fréchet Inception Distance (FID) avoids the drawbacks of IS because it not only considers the similarity of individual image features but also the relationship between the features of real images and generated images. It is often applied to the evaluation of generative adversarial networks. The FID metric uses InceptionV3 as a feature extractor to extract a 2048-dimensional vector before the fully connected layer, then calculates the distance between the mean and covariance matrices of real and generated images in the feature space to assess the quality of generated images. FID is evaluated on the entire test set and the corresponding target domain image set. A lower FID score on the test data indicates that the generated data distribution is closer to the real data distribution, indicating better performance of the generative model, i.e., higher image clarity and richer diversity.

IV. B. 2) Loss Function Results

This paper analyzes the total loss function of the CycleGAN oil painting image style transfer model (CycleGAN-OPM) on the photo2monet and photo2vangogh datasets, as shown in Figure 2. The total loss of the CycleGAN-OPM model exhibits significant fluctuations in the first 50 epochs, but as the training cycle progresses, the fluctuations stabilize. This indicates that the proposed network model can be trained effectively, and the discriminator of the CycleGAN-OPM network gradually becomes unable to distinguish the authenticity of the generated images.

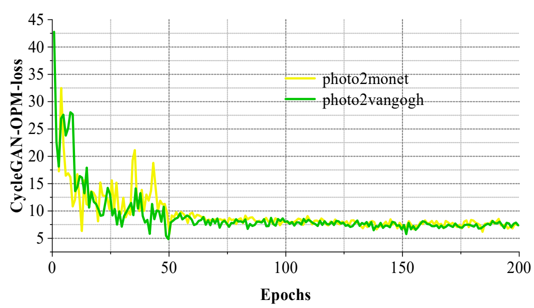

Figure 2: The total loss function of the CycleGAN-OPM oil painting migration model

IV. B. 3) Indicator assessment results

Table 1 and Table 2 present the results of IS and FID evaluation metrics for bidirectional generation using the photo2monet and photo2vangogh datasets. The comparison models include CycleGAN with and without identity mapping loss, MUNIT with and without cyclic consistency loss, and the CycleGAN-OPM model. Higher IS values indicate better image quality, while lower FID values indicate that the generated images share more visual attributes with the original images.

For the IS evaluation metric, the CycleGAN-OPM model achieves the highest IS values on the test sets of the photo2vangogh and photo2monet, vangogh2photo, and monet2photo datasets, with values of 4.64 and 5.25, and 4.81 and 3.67, respectively. For the FID evaluation metric, the FID scores obtained on the photo2vangogh and photo2monet, vangogh2photo and monet2photo datasets are lower than those of other methods, with values of 93.07 and 84.36, and 153.45 and 105.29, respectively.

Table 1: Quantitative comparison in photo2vangogh and photo2monet data sets

	photo2vangogh			photo2monet		
FID	IS		FID	IS		FID
MUNIT (loss=0.0)	4.02	0.24	255.01	4.98	0.15	195.28
MUNIT (loss=10.0)	4.36	0.74	95.94	1.93	0.42	105.03
CycleGAN (loss=0.0)	3.73	0.23	181.64	3.26	0.53	96.35
CycleGAN (loss=0.5)	4.18	0.35	156.31	5.04	0.55	93.09
CycleGAN-OPM	4.64	0.29	93.07	5.25	0.59	84.36

Table 2: Quantitative comparison in vangogh2photo and monet2photo data sets

	vangogh2photo			monet2photo		
FID	IS		FID	IS		FID
MUNIT (loss=0.0)	2.41	0.11	226.04	2.31	0.16	182.28
MUNIT (loss=10.0)	2.33	0.15	225.54	2.07	0.34	183.77
CycleGAN (loss=0.0)	3.74	0.47	182.54	2.86	0.28	141.64
CycleGAN (loss=0.5)	4.51	0.29	164.32	3.51	0.37	112.86
CycleGAN-OPM	4.81	0.52	153.45	3.67	0.45	105.29

IV. C. Qualitative comparison experiments and analysis

To further demonstrate the superiority of the CycleGAN-OPM model, this section designs and implements two user subjective evaluation experiments to assess the quality of the oil painting images generated by the model and the stylistic similarity between the generated oil painting images and the input style images. Each experiment involved 25 participants.

The generated images selected for each test were produced through three types of style conversion: landscape to oil painting, scenery to oil painting, and oil painting to oil painting.

Content refers to the objects contained in the image, such as the background and subject, as manifested by their respective shapes, contours, and structures. Style refers to the overall characteristics of the image, such as color, texture, light and shadow changes, and filter effects.

IV. C. 1) Image Quality Evaluation

(1) Evaluation Rules

Twenty sets of content map and style map combinations were randomly selected for image generation experiments. For each set of images, the CycleGAN-OPM model and the other four baseline models used the same content map and style map to generate five oil painting images of varying quality. Twenty-five participants completed a survey questionnaire to assess the overall image generation quality of the models, honestly scoring the overall quality of images generated by different models. Scoring was based on visual effects such as color harmony, line clarity, natural transitions, and image realism. Scores ranged from 1 to 5, with 5 being the highest and 1 the lowest.

(2) Experimental Results and Analysis

The scores given by the 25 participants were weighted and averaged to obtain the subjective evaluation scores for each model on this set of images. The subjective evaluation results for image generation quality are shown in Figure 3. The average score for the image quality evaluation of the CycleGAN-OPM model was 4.30, significantly higher than the scores for the image quality evaluation of the other four networks. Based on this, the CycleGAN-OPM model was used to generate oil painting images according to the three style conversion types introduced earlier, and corresponding subjective evaluation experiments were conducted on the generated image results. The average scores for the image quality evaluation of each category are shown in Figure 4. In the subjective evaluation experiments of the image quality generated by the three types of style conversion, CycleGAN-OPM obtained high scores, with evaluation scores of 4.24, 4.31, and 4.34 for each type, indicating that the image quality generated by CycleGAN-OPM is superior and closer to real images. Further comparison of the image generation experiment results of the five models across the three style conversion types shows the average scores for each category of image quality evaluation in Figure 5. In the overall subjective evaluation results for image quality, the images generated by the proposed CycleGAN-OPM method outperform those generated by other baseline models in every style conversion category.

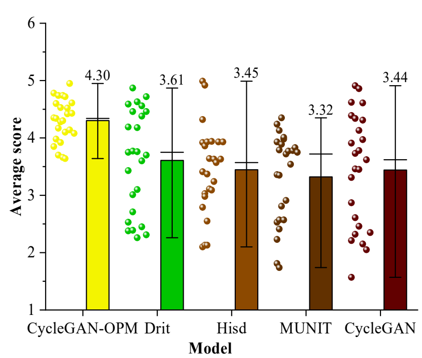

Figure 3: Generate subjective evaluation results of image quality

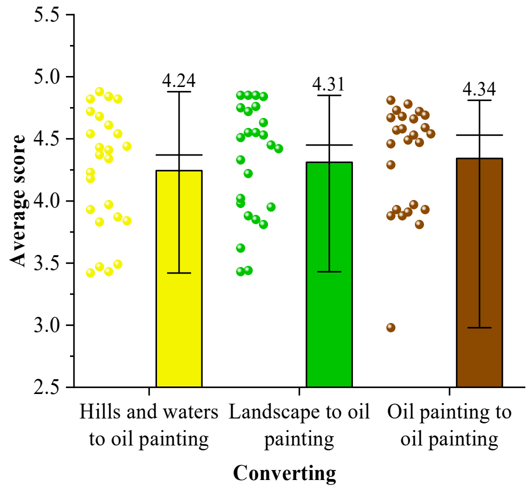

Figure 4: The score of the generation quality evaluation of each category of CycleGAN-OPM

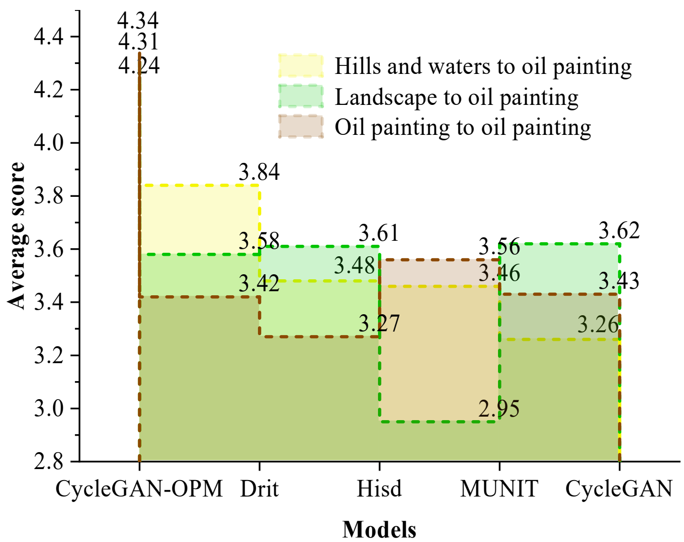

Figure 5: The comparison of average generated quality scores for each category

IV. C. 2) Image Style Similarity Evaluation

(1) Evaluation Rules

A subjective evaluation experiment was conducted to assess the similarity of content and style in the generated images. Twenty sets of content and style image combinations were randomly selected for the image generation experiment. For each set of images, the CycleGAN-OPM model and the other four baseline models generated five images of varying quality using the same content and style images. Participants were asked to score the images based on their visual perception. Higher similarity resulted in higher scores. In both comparisons, the highest score is 5 points, and the lowest is 1 point.

(2) Experimental Results and Analysis

The average scores for the similarity of content and style in the generated images are shown in Figure 6. In the subjective experiment to assess the similarity of content and style in generated oil painting images, the CycleGAN-OPM model performed well, with a subjective evaluation score of 4.06 for content similarity, the highest among the five models. Additionally, the CycleGAN-OPM model's subjective evaluation score for style similarity was 4.12, significantly exceeding the scores of other methods. The average scores for content and style similarity across different types of generated images are shown in Figure 7. The CycleGAN-OPM model can simultaneously maintain content similarity and better convey style when generating oil paintings with landscape images as content input and landscape paintings as style input.

The oil painting images generated by the CycleGAN-OPM model are more in line with human visual habits and can better convey the style information of the specified image while retaining the content of the input image, which is beneficial for artists to refer to in the process of oil painting creation.

In summary, the proposed CycleGAN-OPM model can perform style transfer while retaining more content features, and the generated distribution is closer to that of the original image. The quantitative evaluation results are consistent with the qualitative evaluation results, and the oil painting images generated by the CycleGAN-OPM model have good visual effects.

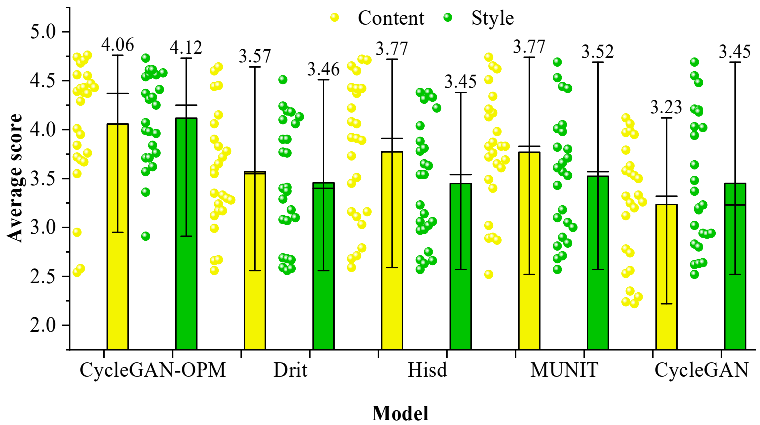

Figure 6: The average content and style similarity scores for generated images

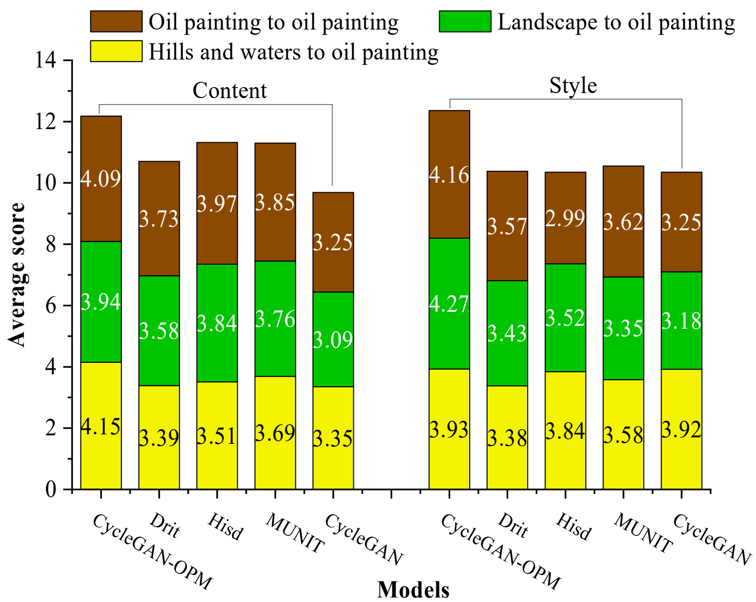

Figure 7: The content and style similarity classification average scores for generated images

V. Conclusion

This study explores the innovative application of AI technology in oil painting creation and its role in promoting artistic development. Based on the CycleGAN model, an oil painting image style transfer method is constructed, and quantitative and qualitative comparison experiments are conducted on the dataset to investigate the image generation effects of the proposed method. The main results of the experiment are as follows:

(1) In the metric evaluation results of different models, the IS values of the proposed CycleGAN-OPM model were higher than those of the comparison methods, while the FID values were lower. The IS values improved by 4.17% to 172.02%, and the FID values decreased by 2.99% to 63.50%. The experiments demonstrate that the proposed method exhibits superior style transfer performance, enabling the generation of more innovative oil painting images and aiding artists in their creative process.

(2) The average subjective evaluation score for the quality of images generated by the proposed model is 4.30, with average content and style similarity evaluation scores of 4.06 and 4.12, respectively, which are 19.11% to 29.52%, 7.69% to 25.70%, and 17.05% to 19.42% higher than those of the other five comparison models. The CycleGAN-OPM model demonstrates superior style information transmission capabilities compared to the baseline model, and the artworks it generates align more closely with human visual perception, offering intelligent auxiliary references for artists creating oil paintings.

AI's role in artistic creation is evolving from a tool to a partner, not only assisting artists technically but also providing inspiration in terms of concepts and creativity. Artists utilize AI algorithms to optimize oil painting techniques and expressive methods, refining the collaborative creation process from sketches to finished works. In the future, the integration of AI and oil painting creation will become more profound and diverse, bringing the public richer and more varied artistic forms.

References

[1] Ma, X. (2024). The Influence of Chinese Folk Art on Modern Oil Painting. Art and Performance Letters, 5, 112-119.
[2] Yan, Z., Guo, Z., & Guo, M. (2023). On the Combination of Oil Painting Techniques and Modern Illustration Art. Highlights Art Des, 4, 118-121.
[3] He, Z. (2023). Exploration on Localization of China's Oil Paintings-Research on Image Oil Painting. Art and Performance Letters, 4(8), 40-43.
[4] Zhang, J. (2021). COLOR EXPRESSIVENESS OF IMAGE OIL PAINTING BASED ON COLOR PSYCHOLOGY. Psychiatria Danubina, 33(suppl 5), 112-113.
[5] Zhang, Y. (2017, June). Function and Manifestation of Decorative Color in Oil Painting Creation. In 2nd International Conference on Contemporary Education, Social Sciences and Humanities (ICCESSH 2017) (pp. 487-489). Atlantis Press.
[6] Antunes, V., Candeias, A., Mirão, J., Carvalho, M. L., Dias, C. B., Manhita, A., ... & Manso, M. (2018). Analytical characterization of the palette and painting techniques of Jorge Afonso, the great 16th century Master of Lisbon painting workshop. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 193, 264-275.
[7] Cui, F. B. (2023). A probe into the modernity of Chinese oil painting. International Journal of Advanced Culture Technology, 11(1), 198-203.
[8] Dal Fovo, A., Tserevelakis, G. J., Klironomou, E., Zacharakis, G., & Fontana, R. (2021). First combined application of photoacoustic and optical techniques to the study of an historical oil painting. The European Physical Journal Plus, 136(7), 757.
[9] Kim, S., Park, S. M., Bak, S., Kim, G. H., Kim, C. S., Jun, J., ... & Kim, K. (2022). Investigation of craquelure patterns in oil paintings using precise 3D morphological analysis for art authentication. Plos one, 17(7), e0272078.
[10] Pakzad, Z. (2017). Color structure in the Persian painting. Rev. Eur. Stud., 9, 1.
[11] Bai, S., & Li, P. (2023). Algorithm and simulation study of oil painting classification based on visual perception and improved embedded learning. Journal of Intelligent & Fuzzy Systems, 45(6), 9979-9989.
[12] Zeng, X. (2022). Research on Cross_Contrast Neural Network Based Intelligent Painting: Taking Oil Painting Language Classification as an Example. Computational Intelligence and Neuroscience, 2022(1), 7827587.
[13] Hu, C., Ding, Y., & Li, Y. (2020, June). Image style transfer based on generative adversarial network. In 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) (Vol. 1, pp. 2098-2102). IEEE
[14] Chen, T., & Yang, J. (2021). A novel multi-feature fusion method in merging information of heterogenous-view data for oil painting image feature extraction and recognition. Frontiers in Neurorobotics, 15, 709043.
[15] Tong, Z., Wang, X., Yuan, S., Chen, X., Wang, J., & Fang, X. (2022, October). Im2oil: Stroke-based oil painting rendering with linearly controllable fineness via adaptive sampling. In Proceedings of the 30th ACM international conference on multimedia (pp. 1035-1046).
[16] Guo, W. (2022). Oil painting art style extraction method based on image data recognition. Mathematical Problems in Engineering, 2022(1), 4196174.
[17] Yao, E., & White, M. (2024). Expression of Concern: Oil painting color image enhancement recognition method based on artificial intelligence: applications of an AI model in environmental research. AQUA—Water Infrastructure, Ecosystems and Society, jws2024042.
[18] Yang, L. S. Q., Li, L., & Cheng, J. (2025). Human-AI collaborative artistic creation in oil painting and Chinese ink painting environments using multimodal visual networks. Journal of Computational Methods in Sciences and Engineering, 14727978251352134.
[19] Gong, M., & Zhang, G. (2024, May). Real-Time Interaction of Convolutional Neural Networks in Oil Painting Creation. In Proceedings of the 2024 International Conference on Machine Intelligence and Digital Applications (pp. 868-872).
[20] Kim, J., Jun, J. Y., Hong, M., Shim, H., & Ahn, J. (2019). Classification of oil painting using machine learning with visualized depth information. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 617-623.
[21] Albadarneh, I. A., & Ahmad, A. (2017). Machine learning based oil painting authentication and features extraction. International Journal of Computer Science and Network Security (IJCSNS), 17(1), 8.
[22] Liao, Z., Gao, L., Zhou, T., Fan, X., Zhang, Y., & Wu, J. (2019). An oil painters recognition method based on cluster multiple kernel learning algorithm. IEEE Access, 7, 26842-26854.
[23] Guo, B., & Hao, P. (2020, July). Analysis of artistic styles in oil painting using deep-learning features. In 2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW) (pp. 1-4). IEEE Computer Society.
[24] Guo, H., Liang, X., & Yu, Y. (2022). Application of big data technology and visual neural network in emotional expression analysis of oil painting theme creation in public environment. Journal of environmental and public health, 2022(1), 7364473.
[25] Zhang, X. (2024). Oil painting image style recognition based on ResNet-NTS network. Journal of Radiation Research and Applied Sciences, 17(3), 100992.
[26] Chang, Y. (2024). Enhancing super resolution of oil painting patterns through optimization of UNet architecture model: Y. Chang. Soft Computing, 28(2), 1295-1316.
[27] Li, J., Xue, Y., & Kraus, S. (2019, August). Scribble-to-Painting Transformation with Multi-Task Generative Adversarial Networks. In IJCAI (pp. 5916-5922).
[28] Liu, X. (2022). An improved oil painting formation using advanced image processing. Mobile Information Systems, 2022(1), 4865060.
[29] Li, M., Lin, L., Luo, G., & Huang, H. (2024, February). Monet style oil painting generation based on cyclic generation confrontation network. In International Conference on Frontiers of Applied Optics and Computer Engineering (AOCE 2024) (Vol. 13080, pp. 120-125). SPIE.
[30] Xu, G. (2024, May). The application research of generative adversarial network (GAN) technology in impressionist painting. In Proceedings of the 2024 International Conference on Digital Society and Artificial Intelligence (pp. 91-95).
[31] Lv, X., & Zhang, X. (2019, November). Generating chinese classical landscape paintings based on cycle-consistent adversarial networks. In 2019 6th International Conference on systems and Informatics (ICSAI) (pp. 1265-1269). IEEE.