Research on Engineering Structures & Materials
www.jresm.org
MIM Research Group

Research Article

Attention-enhanced CycleGAN with multi-scale VGG feature fusion for ceramic decorative pattern style transfer

Kaiyan Jin [*,a]

School of Design, Huzhou College, Huzhou, China

Article Info

Article History:

Received 18 Aug 2026

Accepted 15 Oct 2026

Keywords:

Deep learning;
Self-attention mechanism;
Ceramic decoration;
Style transfer;
Image processing;
Design efficiency

Abstract

Ceramic decorative patterns (CDP) are essential for enhancing ceramic artistic value, but traditional handmade design faces low efficiency and difficulties in maintaining structural integrity, color reproduction, and texture consistency during style transfer. To address these challenges, this study proposes an improved CycleGAN framework for high-quality CDP style transfer. The original convolution module is replaced with VGG19 to capture multi-scale semantic and fine style features. A self-attention mechanism (SAM) is integrated into the generator to model global pixel dependencies and improve pattern continuity and style consistency. In addition, an adaptive sharpening preprocessing mechanism is designed to enhance image clarity and reduce noise interference. Experimental results show that the proposed model achieves a PSNR of 25.10 dB, an SSIM of 0.87, and an FID of 12.89, outperforming the baseline model. The experiments are conducted on the Chinese Ceramic Pattern Dataset (CCPD) containing 12,000 high-quality ceramic pattern images. Practical evaluations demonstrate improved visual satisfaction, pattern integrity, and process feature restoration, while significantly reducing design time for both small and large batches. The proposed method effectively improves CDP design efficiency and artistic expression, providing technical support for ceramic art innovation.

© 2026 MIM Research Group. All rights reserved.

1. Overview

Ceramic art is an important carrier of traditional culture, and the design level of decorative patterns directly determines its artistic appeal and market value [1-2]. Traditional ceramic decorative pattern design has long relied on manual drawing or template reuse, which not only has the problems of time-consuming and labor-intensive creation efficiency, but also faces the dilemma of style solidification. This contradiction has become a key bottleneck restricting ceramic art innovation and industrial upgrading. From the perspective of ceramic craftsmanship, the technical challenges of decorative pattern design are rooted in several domain-specific processes. Underglaze blue (Mavi-Beyaz Çini), for instance, requires the pigment to be applied beneath the glaze layer, so its final visual effect depends on the interaction between cobalt pigment diffusion and glaze transparency during high-temperature firing, which corresponds computationally to a coupled problem of subsurface scattering simulation and non-rigid texture deformation. Slip decoration involves the application of liquid clay to the green body, producing relief-like textures whose computational equivalent is height-field estimation and normal-map reconstruction. Overglaze enamel decoration, including Minai and alum red, is applied on top of the fired glaze and undergoes secondary low-temperature firing, where the color rendering is highly sensitive to firing temperature and atmosphere, corresponding computationally to color gamut mapping under uncertain environmental parameters. Thermal firing deformation causes geometric distortion of the pattern, which is analogous to non-rigid registration and geometric warping in computer vision.

*Corresponding author: KaiyanJinq@outlook.com
[a]orcid.org/0009-0006-9514-7401
DOI: http://dx.doi.org/10.17515/resm2026-2215vk1808rs

These process-specific characteristics explain why general-purpose style transfer methods, which treat images as flat pixel arrays, struggle to preserve the structural integrity, color reproduction, and texture consistency of ceramic decorative patterns. In recent years, with the vigorous development of deep learning (DL) technology, style transfer (ST) and image generation algorithms have brought new solutions to traditional art creation, providing possibilities for resolving the above contradictions. Among them, CycleGAN, as a classic model for unsupervised style transfer, can achieve image transformation of different styles and types without pairing data. It ensures the reversibility of the transformation through bidirectional generators and CycleCCs, and has natural advantages in non-paired data scenarios [3]. However, existing style transfer methods still have significant limitations when dealing with ceramic decorative patterns. On the one hand, the feature extraction of traditional CycleGAN relies on simple convolutional layers, which can only capture local pixel correlations and cannot effectively model the global dependency relationship of long-distance continuous patterns in ceramic patterns. This can easily lead to pattern breakage and uneven style distribution after migration. On the other hand, existing methods lack sufficient restoration of the unique process characteristics of ceramics, often resulting in problems such as loss of details and texture distortion, making it difficult to match the artistic and process requirements of ceramic decoration [4]. Self-Attention Mechanism (SAM), as a technology that can capture long-distance dependencies in images, has shown strong potential in image processing [5]. Visual Geometry Group 19-layer Network (VGG19), as a deep CNN structure, has the advantages of simple structure and strong feature extraction ability [6]. Based on this, a ceramic decoration pattern style transfer algorithm based on CycleGAN architecture was designed. The core innovation lies in introducing self-attention mechanism (SAM) into the generator structure of CycleGAN for the first time, specifically solving the problem of insufficient global dependency modeling in traditional models. At the same time, VGG19 was used to replace the simple convolutional feature extraction module of the original CycleGAN, enhancing the ability to capture details. In addition, an adaptive sharpening enhancement preprocessing mechanism was designed to further improve the clarity of input patterns and the restoration of process features. This scheme aims to construct an unsupervised style transfer framework that adapts to the characteristics of ceramic decorative pattern data sources, achieving more refined and consistent style transfer effects, and providing technical support for ceramic art innovation.

2. Related Works

In the DL field, the application of Image Style Transfer (IST) technology has gradually matured, and many remarkable results have been achieved. These related studies revolve around three core objectives: improving migration quality, optimizing detail preservation, and adapting to specific scenarios. However, there are still significant adaptation gaps in ceramic decorative patterns, which have strong craftsmanship and artistic attributes. Fu X et al. designed an algorithm built on CycleGAN to achieve fast and effective image Art Style Transfer (AST) [7]. However, this method has not been optimized for the long-distance correlation of complex patterns, and it is prone to structural fracture when applied to continuous patterns such as ceramic entwined patterns and loop patterns. To improve the processing capabilities of complex details in ST, Zhu X et al. proposed a real-time attention ST network [8]. However, the segmentation logic of this method is more suitable for natural scene images and is difficult to adapt to the deep binding between ceramic pattern craftsmanship features and artistic styles. To realize IST, Dinesh Kumar R et al. proposed a deep CNN learning classification method for AST [9]. This study validated the advantages of VGG19 in capturing style features, but without incorporating a global dependency modeling mechanism, it cannot solve the problem of coordinating local details with global structure in ceramic patterns. To enhance the ability of neural ST in processing texture ST, Zhou Z et al. proposed a method to capture texture structure by utilizing an adaptive autocorrelation alignment loss function [10]. However, this method focuses on general texture transfer and has not been specifically optimized for unique process features such as ceramic glaze color moistening and cracking texture, resulting in a lack of ceramic art recognition for the transfer effect. To preserve more details in facial IST, Wang H Y et al. put forth an IST model built on region-guided masks [11]. But its regional guidance logic is designed for facial scenes with dense local features, which is difficult to adapt to the global consistency requirements of long-distance continuous patterns in ceramic patterns.

However, existing GAN-based methods still face problems of insufficient process adaptability, efficiency, and quality imbalance in ceramic decoration pattern transfer. This study validated the potential of GAN in small sample, domain specific image generation, but the feature differences between medical images and ceramic patterns are significant, and its model architecture cannot be directly transferred for application. To enhance the ability and efficiency of the model to extract image features in IST, Liao Y M et al. developed a DL model built on CycleGAN [12]. However, this method simplifies the feature extraction module in pursuit of efficiency, resulting in insufficient restoration of details such as brushstrokes and glaze colors in ceramic patterns. To strengthen the quality of migrated images, Garg M et al. put forward a method of generating steganographic images using a neural ST algorithm [13]. Its core advantage lies in the protection of content integrity, but the control of style consistency is weak, and the migrated ceramic patterns are prone to problems such as glaze color generalization and style disorder. To achieve artistic style conversion of complex textures and structures, Mao W et al. proposed an unsupervised framework from coarse to fine [14]. Although this method improves the transfer effect of complex textures, the two-stage process results in high computational complexity, making it difficult to meet the efficiency requirements of ceramic design batch production.

In recent years, state-of-the-art (SOTA) architectures have significantly advanced unpaired image-to-image translation and provide new references for ceramic pattern style transfer. Mochizuki et al. addressed the problem that tumor recognition in CyberKnife treatment is compromised by overlapping bone structures, and proposed a GAN-based bone suppression imaging technique that compared six models including CycleGAN, DualGAN, CUT, FastCUT, DCLGAN, and SimDCL, achieving SSIM and PSNR values of 0.96 and 36.93 dB, with SimDCL obtaining the optimal FID of 68.93 [15]. This work validates the superiority of contrastive learning-based CUT and its variants in preserving structural consistency, but it focuses on medical X-ray imaging with rigid anatomical structures and does not address the global dependency modeling required by long-distance continuous patterns in ceramic decoration. Jiang et al. tackled the problem of artifacts in CycleGAN-based unpaired medical image translation, and proposed CycleH-CUT, an unsupervised network combining cycle consistency and hybrid contrastive unpaired translation, in which a query-selected attention mechanism is adopted to select important feature queries and a boosted contrastive learning loss reweights negative patches via optimal transport, achieving an average SSIM of 0.926 on BraTS, 0.796 on OASIS3, 0.932 on IXI, and 0.890 on a private Spinal Column dataset [16]. This method demonstrates the effectiveness of query-selected attention (QS-Attn) in suppressing artifacts and enhancing feature discrimination, but its contrastive learning strategy is designed for medical image domains with relatively uniform texture distributions, and its adaptability to the multi-scale, craftsmanship-specific features of ceramic patterns remain unverified. Markham et al. addressed the scarcity of representative datasets for ID card presentation attack detection, and explored four GAN-based image translation models to generate synthetic attack samples, obtaining an EER performance increase of 0.63% for print attacks and a loss of 0.29% for screen capture attacks [17]. This study confirms the value of GAN-generated data as a complement to real samples, but its generation process does not involve style consistency constraints for artistic patterns, and the feature differences between ID card images and ceramic decorative patterns are significant. Vinay Kumar and Vimala focused on the challenges of lighting, color, and shape variations in skin lesion analysis, and proposed a customized batch normalization-assisted parallel CNN with LinkNet for skin cancer classification, achieving an accuracy of 0.975, F-measure of 0.914, and specificity of 0.986 [18]. Although this work achieves high classification performance through attention modules and hybrid feature extraction, it is a discriminative classification model rather than a generative style transfer framework, and cannot be directly applied to ceramic pattern style migration. In summary, existing SOTA methods such as CUT, FastCUT, QS-Attn, and contrastive learning-based frameworks have made notable progress in unpaired translation and artifact suppression, but they still lack targeted designs for the global continuity, process feature restoration, and artistic style consistency requirements of ceramic decorative patterns, which further confirms the necessity of the improved CycleGAN framework proposed in this study.

In summary, although current research on image style transfer and GAN in general scenarios, there are still gaps in addressing the core requirements of ceramic decorative pattern style transfer: firstly, insufficient global dependency modeling. Secondly, there is a lack of restoration of process characteristics. Thirdly, there is a poor balance between efficiency and quality. Fourthly, there is insufficient adaptability to the scene. Based on the aforementioned research gaps, the study focuses on the core pain points of ceramic decorative patterns and proposes an improved CycleGAN model that integrates self-attention mechanism (SAM) and VGG19. Through an integrated design of global dependency modeling, multi-scale, and process feature adaptation, it compensates for the insufficient adaptation of existing methods in ceramic style transfer and provides technical support for ceramic art innovation.

3. ST Model Based on Improved CycleGAN

This study uses CycleGAN as the basic framework to construct a CDP-style migration model with a two-way migration cycle. It combines VGG19 and SAM to optimize and improve the model, thereby strengthening the model's Multi-Scale Feature (MSF) extraction and global pattern dependency modeling capabilities.

3.1 CDP Style Migration Model Based on CycleGAN

ST, as a computer vision technology, can apply the style of one image to the content of another image, thereby generating a new image with a specific artistic style [19-21]. However, traditional ST methods have some limitations in CDP, such as incomplete ST, detail distortion, etc. In addition, the current ST method still has insufficient restoration of unique craftsmanship textures such as ceramic glaze texture and brushstroke marks. With its powerful generation ability and adversarial training mechanism, GAN is able to learn complex data distribution, thereby generating realistic images [22-24]. Fig.1 shows the GAN structure.

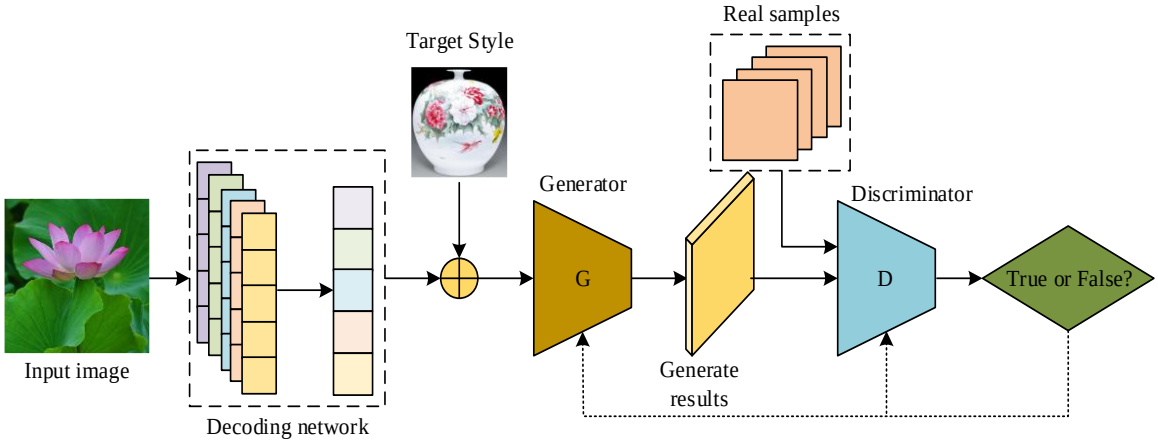


Fig. 1. GAN network structure

In Fig.1, in CDP ST, the generator receives the source style pattern as input, extracts the structural features of the pattern through the decoding network structure, injects the texture and color features of the target style, and finally outputs the style-migrated pattern. The design goal of the generator is to generate images that are as close as possible to real samples in order to fool the discriminator. To improve the basic clarity of the input CDP and reduce the impact of noise interference on the detail restoration of subsequent ST, this study introduces adaptive sharpening and enhancement processing in the generator input stage, as shown in Equation (1).

$$I_{sharp} = I_{in} + \alpha \cdot \nabla(I_{in}) \cdot exp\left(-\beta \cdot \sigma(I_{in})\right) \tag{1}$$

In Equation (1), I_{sharp} is the input image after sharpening and enhancement. I_{in} is the original input CDP image. α is the sharpening intensity coefficient. The physical meaning is to control the enhancement amplitude of pattern edge details, with a value range of [0.1, 0.5] $\nabla(I_{in})$ is the gradient map of the input image. β is the noise suppression coefficient. Used to balance the

intensity of sharpening and denoising, with a value range of [0.05, 0.2]. $\sigma(I_{in})$ is the local standard deviation of the input image. Used to quantify the noise level in a local area, the larger the standard deviation, the stronger the noise suppression. The objective function of GAN is shown in Equation (2).

$$\min_{G}\max_{D} V(D,G) = E_{x\sim P_{data}(x)}[log\, D\,(x)] + E_{z\sim P_z(z)}\left[log\left(1 - D\big(G(z)\big)\right)\right] \tag{2}$$

In Equation (2), D and G are the discriminator and generator. V is the objective function. $P_{data}(x)$ and $P_z(z)$ are the distribution of real samples and generator input noise. $D(x)$ means the possibility that the discriminator decides that x is a real sample. $G(z)$ denotes the fake sample generated by the generator. When the training reaches Nash equilibrium, the discriminator cannot distinguish real samples from generated samples, indicating that the generator has learned the style characteristics of the target domain and can thus generate high-quality transfer samples. However, although traditional GAN can achieve sample generation, it has obvious limitations in the ST task. For example, traditional GAN requires paired data from the source and target domains for training, and the style diversity of CDP makes it difficult to obtain such paired data on a large scale. In addition, traditional GAN lacks constraints on content consistency and is prone to excessive ST. As a classic model of unsupervised ST in GAN, CycleGAN does not require paired data and is highly consistent with the data source characteristics of CDP [25-26]. At the same time, CycleGAN can ensure the bidirectional consistency of ST through CycleCC. The bidirectional migration graph of CycleGAN is shown in Fig.2.

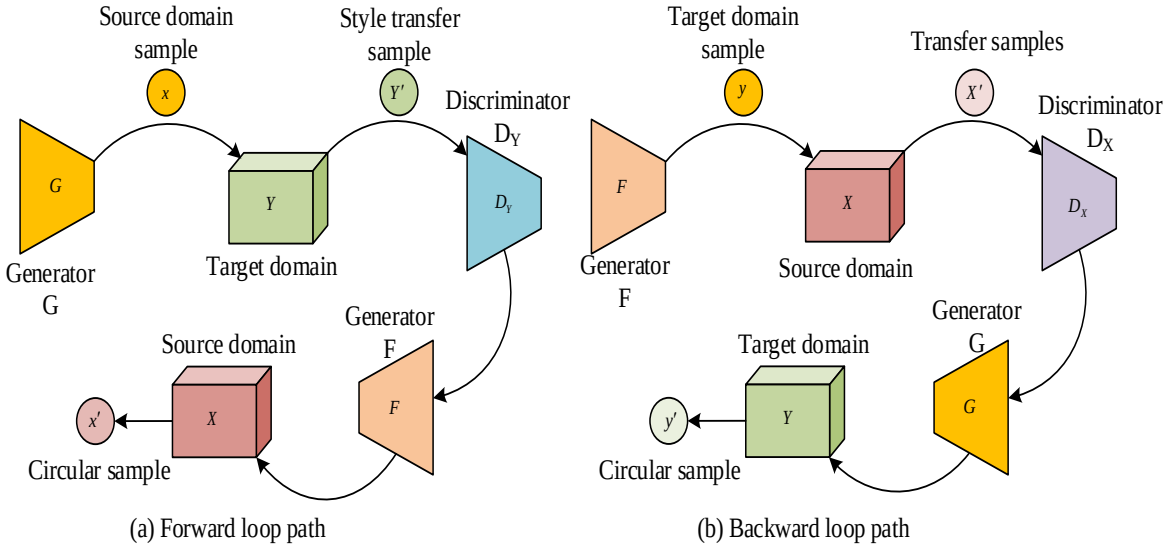


Fig. 2. Bidirectional migration cycle of CycleGAN

In Fig.2, the generator G maps the source domain sample x to the target domain Y, and the mapping generates the ST sample $Y^{'}$, and the discriminator D_Y performs authenticity determination. Then, the generator F implements reverse mapping, migrating the Y samples to the source domain X, and generating the cyclic sample x'. This forward Circular Migration Path (CMP) verifies the reversibility of content after style migration, thereby avoiding content distortion caused by style injection. In backward CMP, the target domain sample y is mapped to the X through the F to obtain the migration sample $X^{'}$, and the discriminator D_X performs authenticity determination. Then, it is migrated back to the Y through the G to generate the cyclic sample y'. This backward CMP requires a high degree of consistency between the style features of the loop samples and the original samples, which can prevent style confusion or shift during cross-domain migration. The adversarial loss of forward transfer $L_{GAN}(G, D_Y, X, Y)$ is shown in Equation (3).

$$L_{GAN}(G, D_Y, X, Y) = E_{y\sim P_Y(y)}[log\, D_Y\,(y)] + E_{x\sim P_X(x)}\left[log\left(1 - D_Y\left(G(x^{'})\right)\right)\right] \tag{3}$$

In Equation (3), $P_Y(y)$ and $P_X(x)$ are probability distributions of y and x. $D_Y(y)$ is the probability that the discriminator D_Y determines that y is a real sample. $G(x')$ is the x' generated by G. The adversarial loss of reverse transfer $L_{GAN}(F, D_X, Y, X)$ is shown in Equation (4).

$$L_{GAN}(F, D_X, Y, X) = E_{x\sim P_X(x)}[log\, D_X(x)] + E_{y\sim P_Y(y)}\left[log\left(1 - D_X\left(G(y')\right)\right)\right] \tag{4}$$

In Equation (4), $D_X(x)$ is the probability that D_X determines that x is a real sample. $G(y')$ denotes the y' generated by the G. The function of the total loss of cycle consistency $L_{cyc}(G, F)$ is shown in Equation (5).

$$L_{cyc}(G, F) = E_{x\sim P_X(x)}\left[\left\|F(G(x)) - x\right\|_1\right] + E_{y\sim P_Y(y)}\left[\left\|G(F(y)) - y\right\|_1\right] \tag{5}$$

In Equation (5), $\|\cdot\|_1$ is the Manhattan distance. The function of the total loss of identity mapping $L_{id}(G, F)$ is shown in Equation (6).

$$L_{id}(G, F) = E_{y\sim P_Y(y)}[\|G(y) - y\|_1] + E_{x\sim P_X(x)}[\|F(x) - x\|_1] \tag{6}$$

The total loss of CycleGAN model ($L(G, F, D_X, D_Y)$) training is shown in Equation (7).

$$\begin{aligned} L_{\text{total}} = {} & L_{\text{GAN}}(G, D_Y, X, Y) + L_{\text{GAN}}(F, D_X, Y, X) + \lambda_1 L_{\text{cyc}}(G, F) + \lambda_2 L_{\text{id}}(G, F) \\ & + \lambda_3 L_{\text{style}} + \lambda_4 L_{\text{PB}} + \lambda_5 L_{\text{CG}} + \lambda_6 L_{\text{LT}} + \lambda_7 L_{\text{CC}} \end{aligned} \tag{7}$$

In Equation (7), L_{total} is the unified master loss function of the proposed model. $L_{GAN}(G, D_Y, X, Y)$ and $L_{GAN}(F, D_X, Y, X)$ are the forward and reverse adversarial losses defined in Equations (3) and (4). $L_{cyc}(G, F)$ is the cycle consistency loss defined in Equation (5). $L_{id}(G, F)$ is the identity mapping loss defined in Equation (6). L_{style} is the VGG19 style loss used to constrain the distribution of multi-scale style features. L_{PB} is the pattern bone method loss, which preserves the skeleton structure of ceramic patterns. L_{CG} is the glaze color system loss, which restricts the generated image to a reasonable glaze color gamut in the CIELAB color space. L_{LT} is the texture loss, which restores local process textures such as brushstrokes and glaze cracks. L_{CC} is the continuity constraint loss based on surface parameterization, which ensures pixel consistency, gradient continuity, and feature coherence at the seam boundary. λ_1, λ_2, λ_3, λ_4, λ_5, λ_6 and λ_7 are the weight coefficients that balance the contribution of each loss term. λ_1 is the weight coefficient of the Cycle Consistency Loss (CCL), which is used to balance the importance of adversarial loss and cycle constraints. λ_2 is the weight coefficient of the identity mapping loss, which is utilized to control the strength of the identity constraint. λ_3 to λ_7 are the weight coefficients of the VGG19 style loss, pattern bone method loss, glaze color system loss, texture loss, and continuity constraint loss, respectively.

3.2 CycleGAN Migration Model Improved Based on SAM

The CDP style migration model based on CycleGAN relies on CycleCC to effectively guarantee the content structure's integrity in the two-way style migration of CDP. However, the loss function in the CycleGAN is mainly based on pixel-level differences, so the MSF separation ability of ceramic patterns is limited, resulting in a low degree of detail reduction in ST, and it is difficult to capture advanced features and style details in ceramic images. In addition, the generator in CycleGAN uses traditional conv layers as the core feature extraction unit. The local receptive field of the conv operation can only capture the correlation information of adjacent pixels, and cannot effectively model the global consistency of long-distance continuous patterns, which can easily lead to pattern breakage and uneven style distribution after migration. To solve these problems, this study introduces VGG19 and SAM to improve CycleGAN. VGG19 achieves hierarchical characterization of MSFs through progressive feature extraction of 16 conv layers, thereby being able to capture the detailed features of ceramic patterns. Fig.3 shows the VGG19's structure.

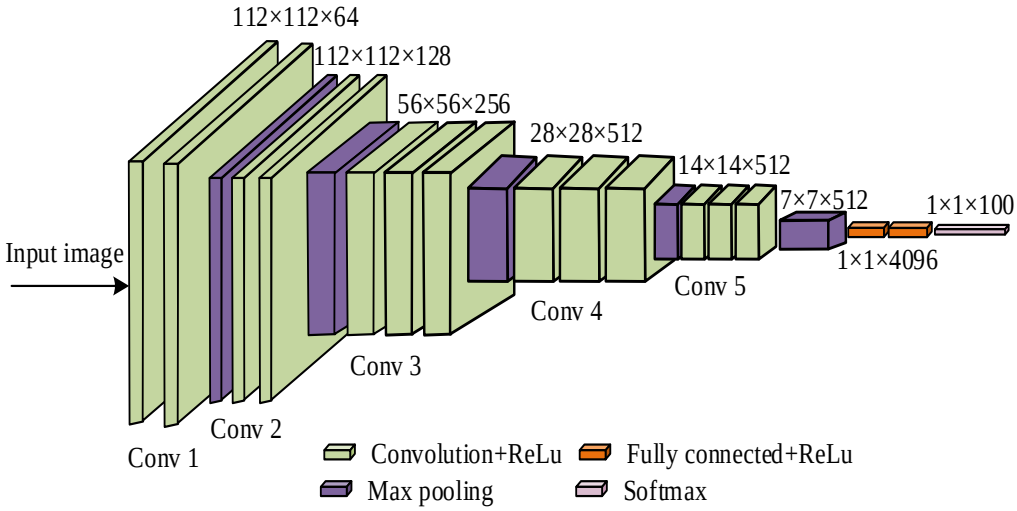


Fig. 3. Network structure of VGG19

In Fig.3, it consists of 5 conv blocks combined with a fully connected layer. Its core feature is to utilize a 3×3 small conv kernel combined with a multi-layer stacking structure to achieve MSF extraction through progressive convolution and maximum pooling operations. Therefore, this study uses the structure of conv block 3 to conv block 5 of VGG19 to replace CycleGAN's original simple convolution feature extraction module for improvement, thereby avoiding the interference of low-level features on ST. The MSF extraction formula of VGG19 is shown in Equation (8).

$$F_l = VGG_l(I) \tag{8}$$

In Equation (8), F_l and $VGG_l(\cdot)$ are the Feature Maps (Fmaps) and feature extraction operators output by the l -th conv block. I is the input image. To quantify the style characteristics of ceramic patterns, this study introduces the Gram matrix to calculate the channel correlation between Fmaps, as shown in Equation (9).

$$Gram_l = \frac{F_l^T}{C_l H_l W_l} \tag{9}$$

In Equation (9), $Gram_l$ is the Gram matrix of the Fmap of the l -th conv block. C_l is the amount of output channels of the l -th conv block. H_l and W_l are the height and width of the output Fmap of the l -th conv block. T means transpose. The style loss of VGG19 (L_{style}) is shown in Equation (10).

$$L_{style} = \sum_{l=3}^{5} \omega_l \left\| Gram_l^{gen} - Gram_l^{ref} \right\|_2^2 \tag{10}$$

ω_l is the style loss weight of the l -th conv block. $Gram_l^{gen}$ and $Gram_l^{ref}$ are the Gram matrices of the generated sample and the target style reference sample respectively after feature extraction by the l -th conv block. To further enhance the details of the mid- and high-level Fmaps extracted by VGG19 and improve the restoration of process features such as texture and strokes in ST, this study performs channel-wise detail enhancement processing on the Fmaps output by VGG19. Since the Gram matrix $Gram(F_{vgg})$ has dimensions of $C \times C$ while the feature map F_{vgg} has dimensions of $C \times H \times W$, a direct element-wise multiplication is not mathematically feasible. To resolve this dimensional mismatch, the Gram matrix is first reduced to a channel-wise energy vector by spatial average pooling, which is then broadcast along the spatial dimensions to align with F_{vgg}. The enhancement process is defined in Equation (11).

$$\begin{cases} F_{enhance}^{(c,h,w)} = F_{vgg}^{(c,h,w)} \cdot \left(1 + \gamma \cdot \dfrac{G_{norm}^{(c)}}{max_c(G_{norm}^{(c)})}\right) \\ G_{norm}^{(c)} = \dfrac{1}{H \cdot W} \sum_{h=1}^{H} \sum_{w=1}^{W} \left| F_{vgg}^{(c,h,w)} \right|^2 \end{cases} \tag{11}$$

In Equation (11), $F_{enhance}^{(c,h,w)}$ and $F_{vgg}^{(c,h,w)}$ are the enhanced feature map and the abstract style feature map output by VGG19 convolution block 5, respectively, where $c \in \{1,2,\ldots,W\}$ denote the spatial coordinates. $G_{norm}^{(c)}$ is the channel-wise normalized Gram energy, which is computed by averaging the squared activations of the c -th channel over the spatial dimensions $H \times W$, as defined in Equation (11). This operation reduces the $C \times C$ Gram matrix to a C -dimensional vector, which is then broadcast along the spatial dimensions to match the $C \times H \times W$ shape of $F_{vgg} \cdot \gamma$ is the feature enhancement coefficient. $max_c(G_{norm}^{(c)})$ is the maximum value of $G_{norm}^{(c)}$ across all channels and is used to normalize the enhancement weight into the range [0, 1]. The term $1 + \gamma \cdot \frac{G_{norm}^{(c)}}{max_c(G_{norm}^{(c)})}$ thus serves as a channel-wise scaling factor that amplifies channels with richer style energy while suppressing relatively weak channels, thereby enhancing the restoration of mid- and high-level process features such as texture and brushstrokes. γ is the feature enhancement coefficient.

SAM can calculate the correlation between the positions of the input Fmap, generate a weight matrix, and adjust the weight of each position of the Fmap to capture the global information. Therefore, this study introduces SAM modeling to improve the global pixel dependency, thereby making up for the local defects of traditional convolution and ensuring the style consistency of long-distance patterns. The SAM flowchart is listed in Fig.4.

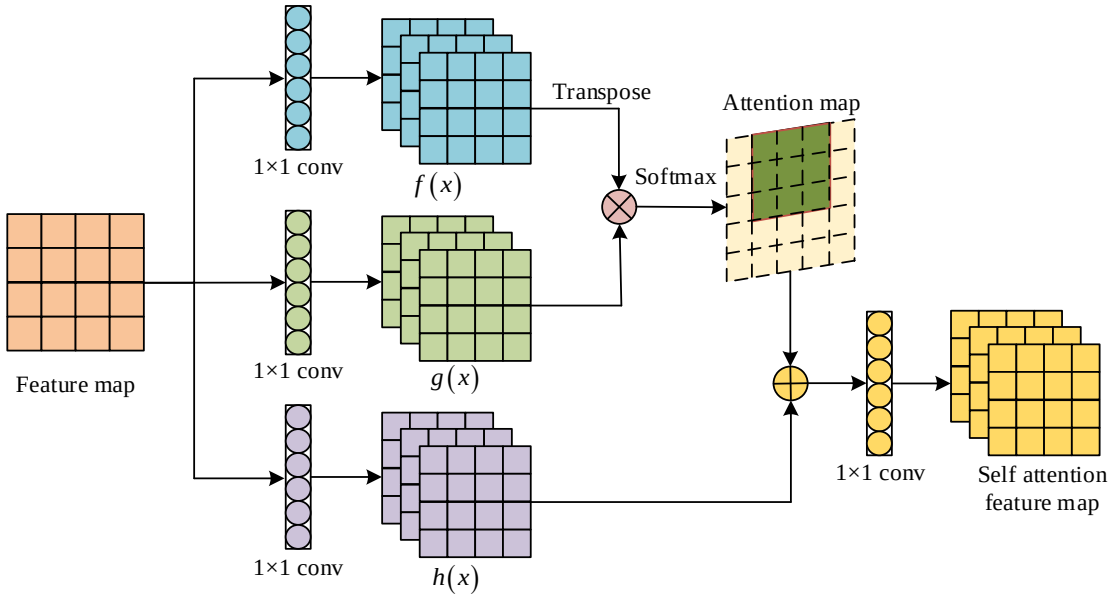


Fig. 4. SAM structure

In Fig.4, the SAM structure takes Fmaps as input, performs dimension adaptation and feature conversion on the input features through three independent 1×1 convolution branches, and generates three sets of intermediate Fmaps $f(x)$, $g(x)$, and $h(x)$. Among them, the $f(x)$ branch performs the transpose operation after 1×1 convolution. The transposed $f(x)$ and the convolved $g(x)$ are subjected to matrix multiplication to calculate the similarity between each position of the Fmap to construct a correlation strength matrix. Then, the correlation strength matrix is normalized by the Softmax function to generate a global weight matrix with a value range of [0,1]. Finally, the normalized weight matrix and the $h(x)$ branch are weighted and summed element by element to complete the aggregation of relevant features in the global scope. The mathematical formula of the intermediate Fmap is given by Equation (12).

$$\begin{cases} f(x) = W_f * X \\ g(x) = W_g * X \\ h(x) = W_h * X \end{cases} \tag{12}$$

In Equation (12), X is the input Fmap. W_f, W_g, and W_h are the 1×1 convolution kernel weight matrices of the three groups of intermediate Fmaps $f(x)$, $g(x)$, and $h(x)$. The correlation strength matrix S is calculated as Equation (13).

$$S = f(x)^T \cdot g(x) \tag{13}$$

In Equation (13), $f(x)^T$ is the transpose matrix after flattening the intermediate Fmap $f(x)$. The expression of the normalized global weight matrix (A) is given by Equation (14).

$$A = Softmax\left(\frac{S}{\sqrt{C'}}\right) \tag{14}$$

In Equation (14), $\sqrt{C'}$ is the normalization factor. The mathematical localization of the normalization factor is $\sqrt{C'} = \sqrt{d_k} \cdot \frac{1}{H \times W} \sum_{i=1}^{H \times W} \sum_{j=1}^{H \times W} S_{i,j}$, where d_k represents the channel dimension of the intermediate feature map, and $H \times W$ represents the spatial dimensions of the height and width of the feature map. The global feature aggregation formula is given by Equation (15).

$$Y = A \cdot h(x) + X \tag{15}$$

Y are the final output Fmap of SAM. Fig.5 shows the structure of the improved CycleGAN.

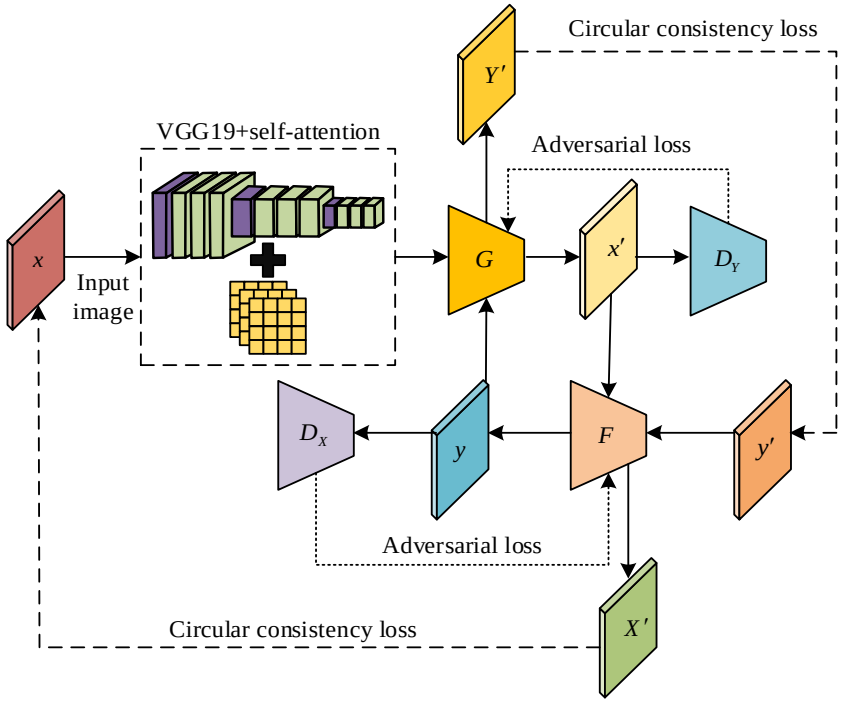


Fig. 5. Improved CycleGAN migration model structure

In Fig.5, the improved model generator adopts a two-way symmetric structure, and replaces the original simple convolution layer through VGG19 feature extraction and SAM. Among them, the VGG19 convolution block is in charge of extracting mid-to-high-level semantic features and abstract style features to avoid interference from low-level textures. SAM is used to model global pixel dependencies and solve problems such as long-distance pattern breaks and uneven style distribution. In addition, the model structure contains two discriminators, which are utilized to distinguish real and generated samples in the target domain and source domain, thereby providing

adversarial loss feedback. This structural design ensures that the content is not distorted through CycleCC, and also achieves uniform distribution of style and restoration of details through improved modules, allowing the model to achieve a migration effect of complete content, pure style, and global coordination. To solve the problem of traditional style transfer treating ceramic style as a whole and lacking targeted control, further research is being conducted to decouple and quantify the core features of ceramic style, breaking it down into three independent sub dimensions: Pattern Bone Method, Glaze Color System, and Texture. The pattern bone method loss function is shown in equation (16).

$$\begin{cases} L_{PB} = \dfrac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} \left| \tilde{S}_{gt}(i,j) - \tilde{S}_{gen}(i,j) \right| \\ \tilde{S}(i,j) = \sigma(\alpha \cdot \| \nabla G_\sigma * I(i,j) \|_2 - \tau) \\ G_\sigma(x,y) = \dfrac{1}{2\pi\sigma^2} exp\left(-\dfrac{x^2 + y^2}{2\sigma^2} \right) \end{cases} \tag{16}$$

In equation (16), L_{PB} represents the pattern skeleton loss function, H and W represent the feature map size, $\tilde{S}_{gt}$ represents the skeleton feature map of the target style map, and $\tilde{S}_{gen}$ represents the skeleton feature map of the generated map. Since conventional skeletonization operations such as morphological thinning or Medial Axis Transform (MAT) are discrete and non-differentiable, they cannot be directly embedded into the PyTorch backpropagation graph. To address this issue, this study adopts a differentiable soft-skeleton approximation based on smoothed gradient filtering, as defined in Equation (16). Specifically, the input image I is first convolved with a Gaussian kernel G_σ to suppress noise and ensure smooth gradients. The gradient magnitude $\|\nabla G_\sigma * I\|_2$ is then computed using Sobel operators, which are linear and fully differentiable. A sigmoid function $\sigma(\cdot)$ with a steepness parameter α and a threshold τ is finally applied to produce a soft skeleton map $\tilde{S}$ with continuous values in [0, 1], where pixels near the pattern ridge approach 1 and background pixels approach 0. This soft-skeleton representation preserves the structural information of ceramic pattern bones while remaining differentiable, allowing gradients to flow back to the generator during training.

Study the color gamut specificity of ceramic glaze colors and calculate the difference in color gamut distribution between the generated image and the target style image in the CIELAB color space. By calculating the color mean, variance, and color gamut coverage of both, the generated image is constrained to fall within a reasonable range of the target glaze color, avoiding color distortion in general style transfer. It should be noted that glaze color transformations during firing are non-linear, because the final chroma depends jointly on kiln temperature, atmosphere, and pigment concentration, and the resulting color shifts are not uniform across the hue circle. The CIELAB color gamut intersection $G_{gen} \cap G_{gt}$ cannot fully reproduce this non-linear firing process; instead, it provides a perceptual constraint that keeps the generated glaze color distribution within the neighborhood of the target distribution. The loss function of the glaze color system is shown in equation (17).

$$L_{CG} = \alpha' \cdot \left| \mu_{gt}^{L} - \mu_{gen}^{L} \right| + \beta' \cdot \left| \sigma_{gt}^{ab} - \sigma_{gen}^{ab} \right| + \gamma' \cdot \left(1 - \frac{G_{gen} \cap G_{gt}}{G_{gt}} \right) \tag{17}$$

In equation (17), L_{CG} represents the loss function of the glaze color system, μ^L represents the average of the L channel, σ^{ab} represents the variance of channels a and b, G represents the color gamut range, and α', β', and γ' all represent weight coefficients. The texture loss function is shown in equation (18).

$$L_{LT} = \frac{1}{N} \sum_{k=1}^{N} \sqrt{\sum_{m=1}^{M} \left(T_{gt,k}^{m} - T_{gen,k}^{m} \right)^2} \tag{18}$$

In equation (18), L_{LT} represents the texture loss function, N represents the number of local texture blocks, M represents the statistical feature dimension, and T^m represents the mth statistical feature. The continuity constraint loss based on surface parameterization is shown in equation (19).

$$L_{CC} = \frac{1}{N}\sum_{i=1}^{N}\left(\nu|I_L(i) - I_R(i)| + \vartheta|\nabla I_L(i) - \nabla I_R(i)| + \upsilon(1 - cos\,\theta_i)\right) \quad (19)$$

In equation (19), L_{CC} represents the continuity constraint loss based on surface parameterization, N represents the total number of pixels at the seam boundary, I_L and I_R represent the pixel values on the left and right sides of the seam, ∇I represents the pixel gradient, θ_i represents the angle at which VGG19 extracts features at the seam, and ν, ϑ, and υ represent the weight coefficients that control pixel consistency, gradient continuity, and feature coherence, respectively.

It should be noted that the seam boundary in Equation (19) does not refer only to the boundary of the 2D generated pattern, but also to the seam line produced by UV unwrapping when the 2D pattern is mapped onto a 3D ceramic form. To bridge the gap between 2D pattern generation and 3D ceramic surfaces, this study adopts a three-step projection pipeline: UV unwrapping, seam-aware 2D style transfer, and 3D texture mapping. First, the 3D mesh of the ceramic object (vase, bowl, plate, or jar) is parameterized into a 2D UV plane using the Least Squares Conformal Maps (LSCM) algorithm, which minimizes angular distortion while preserving local shape. The UV unwrapping inevitably introduces a seam line where the mesh is cut open, and this seam corresponds to the boundary term in Equation (19). Second, the generated 2D ceramic pattern is mapped onto the UV plane, and the continuity loss is applied along the seam line so that pixel values, gradients, and VGG19 features remain consistent across the two sides of the cut. Third, the stylized UV texture is re-projected onto the 3D mesh through inverse UV mapping, and the final rendered result is obtained under the ceramic object's surface normal and curvature. For flat, curved, and irregular ceramic forms, the UV parameterization is performed separately to ensure that the pattern density and orientation remain consistent with the physical surface.

4. ST Model Verification Based on Improved CycleGAN

4.1 Experimental Environment Setup

To validate the performance based on the improved CycleGAN, this study establishes an experimental environment on the Ubuntu 20.04 LTS. Hardware configuration includes NVIDIA GeForce GTX 1080Ti GPU (11 GB GDDR5X video memory) and Intel Core i9-13900K CPU, equipped with 64 GB of memory. The software is programmed using Python 3.8.16, and uses the PyTorch 1.13.1DL framework and CUDA Toolkit 11.7 acceleration tool. In addition, this study uses the Adam optimizer and sets the initial learning rate to 1e-4, the weight decay to 1e-5, the batch size to 16, and the number of training epochs to 500. Given that the training set contains 9,600 images (80% of the 12,000 CCPD images) and the batch size is 16, a single epoch comprises 600 iterations, and the total training process amounts to 300,000 iterations. CCL and identity mapping loss are 10 and 5. In terms of dataset, the publicly available Chinese Ceramic Pattern Dataset (CCPD) is mainly used, which is jointly released by the School of Fine Arts and the Ceramic Culture Research Center of Tsinghua University. The dataset contains 12000 high-quality ceramic pattern images, and the data sources cover classic works of Neolithic colored pottery, Tang and Song famous kilns, Ming and Qing official kilns, and folk kilns. To further enhance the style coverage and scene adaptability of the dataset, a self-built dataset was constructed to supplement the research. Collection criteria: Select ceramic artwork images from Jingdezhen Ceramic Museum's publicly available digital resources, professional ceramic design platforms, and industry journals, with priority given to modern innovative style patterns after 2010 to ensure complementarity with CCPD's traditional style. The filtering criteria include image clarity ≥ 512 × 512, no obvious noise pollution, and clear style features. Table 1 presents the quantitative distribution of style categories in both the CCPD dataset and the self-built dataset. The CCPD dataset covers five traditional style categories, namely Neolithic colored pottery, Tang and Song crackle glaze, Ming and Qing underglaze blue, Ming and Qing overglaze enamel, and folk kiln slip decoration. The self-built dataset contains three modern innovative style categories, namely modern innovative underglaze blue, modern innovative overglaze enamel, and modern comprehensive decoration. After all datasets undergo preprocessing operations like cleaning and data enhancement, they are mixed with the data in CCPD and segmented into test sets and training sets in a ratio of 2:8. The specific hardware configuration of the experiment is shown in Table 2.

Table 1. Style category distribution of CCPD and the self-built dataset

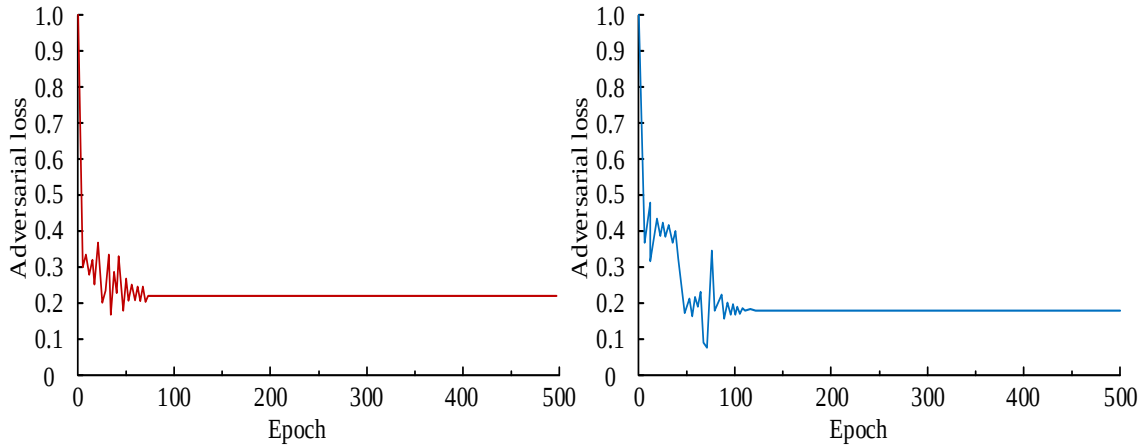
Dataset	Style Category	Number of Images	Proportion (%)
CCPD	Neolithic colored pottery	1680	14
CCPD	Tang and Song crackle glaze	2040	17
CCPD	Ming and Qing underglaze blue	3360	28
CCPD	Ming and Qing overglaze enamel	2520	21
CCPD	Folk kiln slip decoration	2400	20
CCPD Subtotal	/	12000	100
Self-built	Modern innovative underglaze blue	720	36
Self-built	Modern innovative overglaze enamel	640	32
Self-built	Modern comprehensive decoration	640	32
Self-built Subtotal	/	2000	100
Total	/	14000	–

Table 2. Experimental configuration

Environment	Configuration
Operating System	Ubuntu 20.04 LTS
GPU	NVIDIA GeForce GTX 1080Ti
CPU	Intel Core i9-13900K
Memory	64GB
GPU Memory	11 GB GDDR5X
Programming Language	Python 3.8.16
DL Framework	PyTorch 1.13.1
Acceleration tool	CUDA Toolkit 11.7

4.2 Performance Verification of the Improved CycleGAN Model

To test the training stability of the improved CycleGAN, this study analyzes the change curves of adversarial loss and CCL during model training, as displayed in Fig.6. In Fig.6(a), the adversarial loss shows a rapid downward trend in the early epochs, and the curve stabilizes after 50 epochs, and finally converges stably to 0.21. This shows that the model reaches Nash equilibrium around 50 epochs. At this time, the migration samples generated by the generator can better imitate the style features of the target domain. In Fig.6(b), the CCL quickly drops to 0.4 around 40 epochs, and then gradually converges and stabilizes at a low level of 0.18 near 80 epochs. This shows that the core structure of CDP can still remain intact after style migration, and no content distortion occurs.



(a) Adversarial loss change curve (b) Cycle consistency loss variation curve

Fig. 6. The variation curve of adversarial loss and CCL

This study uses basic CycleGAN as the benchmark to compare the image quality indicators of three types of models: CycleGAN+VGG19, CycleGAN+SAM, and improved CycleGAN's dual-module fusion, as shown in Fig.7. In Fig.7(a), under the training set, the PSNR of the improved CycleGAN is as high as 25.07 dB, and the PSNR of the basic CycleGAN, CycleGAN+VGG19, and CycleGAN+SAM are 22.31 dB, 24.15 dB, and 23.02 dB. The PSNR of the improved CycleGAN is 12.37% higher than the basic model, while the improvements of CycleGAN+VGG19 and CycleGAN+SAM are 8.24% and 3.18%. In terms of SSIM, the SSIM value of the improved CycleGAN is up to 0.86, which is 10.25% higher than the 0.78 of the basic models. The SSIM values of CycleGAN+VGG19 and CycleGAN+SAM are 0.83 and 0.81. Although they have also improved, they are 6.41% and 3.84% higher than the basic model, neither of which exceeds 10%. The PSNR (25.87 dB) and SSIM (0.88) of the improved CycleGAN also reach the highest, indicating that the fusion structure effectively improves the quality of the migrated image.

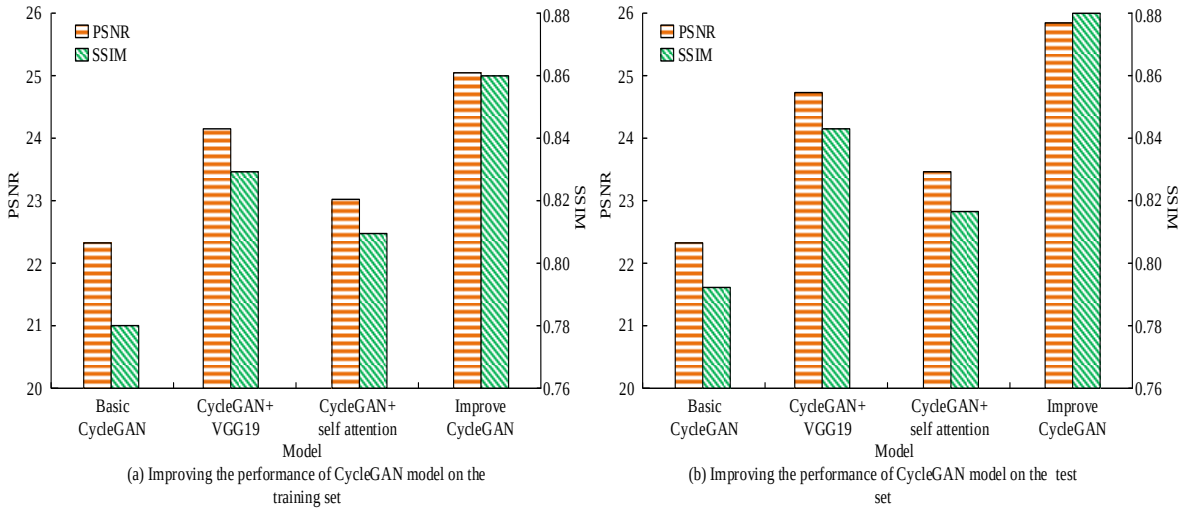


Fig. 7. Image quality before and after improvement of models

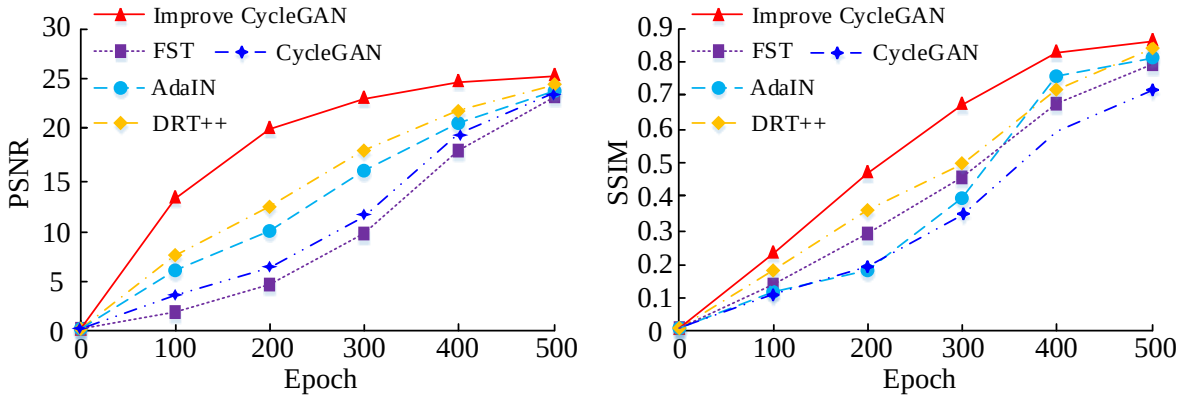
To further validate the improvement effect of each module in the improved CycleGAN, this study adds each module to the network structure to conduct ablation experiments. "√" in the experiment indicates that the module exists, and "/" indicates that it does not exist. In addition to PSNR and SSIM, this study adopts LPIPS (Learned Perceptual Image Patch Similarity) as a perceptual-level evaluation metric. LPIPS computes the L2 distance between deep features extracted by a pre-trained VGG network, which is more consistent with human perceptual judgment than pixel-level metrics. A lower LPIPS value indicates higher perceptual similarity. All models are evaluated under identical settings to ensure fair comparison. Table 3 shows the ablation experimental data. Only the base module of CycleGAN has an FID value of 28.56. After adding adaptive sharpening enhancement, the FID value of the model drops to 26.13, an increase of 8.51% compared to the baseline. After adding VGG19 alone, the FID value reaches 21.34, an increase of 25.28%. After the VGG19 overlay feature details are enhanced, the FID value is 18.72, which is 12.28% higher than the VGG19 module alone, indicating that this module can enhance the restoration of process details. After introducing the SAM module separately, the FID value is 24.78, an increase of 13.23%. After optimizing the integration of all modules, the lowest FID value of the model is only 12.89, which is 54.87% higher than the baseline module. In terms of LPIPS, the baseline CycleGAN obtains the highest value of 0.412, indicating the largest perceptual gap from the target style. Adding adaptive sharpening enhancement alone reduces LPIPS to 0.387 (a decrease of 6.07%), while VGG19 feature extraction alone reduces it to 0.324 (a decrease of 21.36%). After further introducing VGG19 feature detail enhancement, LPIPS drops to 0.286 (a decrease of 30.58% compared with the baseline). Introducing SAM alone yields an LPIPS of 0.358 (a decrease of 13.11%). With the integration of all modules, LPIPS reaches the lowest value of 0.187, a reduction of 54.61% compared with the baseline. The trend of LPIPS is consistent with that of FID, confirming that each module contributes to both pixel-level fidelity and perceptual similarity. To sum up, the addition of each module has brought gains to the performance of the model, indicating that the improved method has effectively improved the ST quality of ceramic patterns.

Table 3. Ablation experiment

Cycle GAN	Adaptive sharpening enhancement	VGG19 feature extraction	VGG19 feature detail enhancement	Self-attention	FID value	LPIPS
√	/	/	/	/	28.56	0.412
√	√	/	/	/	26.13	0.387
√	/	√	/	/	21.34	0.384
√	/	√	√	/	18.72	0.286
√	/	/	/	√	24.78	0.358
√	√	√	√	/	16.25	0.251
√	√	√	√	√	12.89	0.187

4.3 ST Performance Verification

To verify the migration performance of the improved CycleGAN, this paper compares it with the baseline CycleGAN and other advanced migration models, including Fast Style Transfer (FST), Adaptive Instance Normalization (AdaIN), Disentangled Representation for Image Translation++ (DRIT++), as well as state-of-the-art (SOTA) unpaired translation models such as Contrastive Unpaired Translation (CUT), FastCUT, Query-Selected Attention (QS-Attn), and a SwinTransformer-based generative model. Fig.8 is a comparison of the migration image quality of different migration models. In Fig.8, the baseline CycleGAN is added as a reference curve. In Fig.8(a), the PSNR value of the improved CycleGAN is up to 25.10 dB, which is 12.37%, 8.90%, 5.36%, 2.20%, 4.28%, 5.91%, 3.85%, and 6.63% higher than the PSNR values of the baseline CycleGAN (22.33 dB), FST (23.05 dB), AdaIN (23.82 dB), DRIT++ (24.56 dB), CUT (24.07 dB), FastCUT (23.70 dB), QS-Attn (24.17 dB), and SwinTransformer-based model (23.54 dB). In Fig.8(b), the SSIM of the improved CycleGAN is as high as 0.87, which is 11.54%, 10.13%, 7.41%, 3.57%, 5.75%, 7.41%, 4.82%, and 8.05% higher than the 0.78, 0.79, 0.81, 0.84, 0.82, 0.81, 0.83, and 0.80 of the baseline CycleGAN, FST, AdaIN, DRIT++, CUT, FastCUT, QS-Attn, and SwinTransformer-based model. In summary, the improved CycleGAN model has obvious advantages in image migration performance over both the baseline and SOTA models.



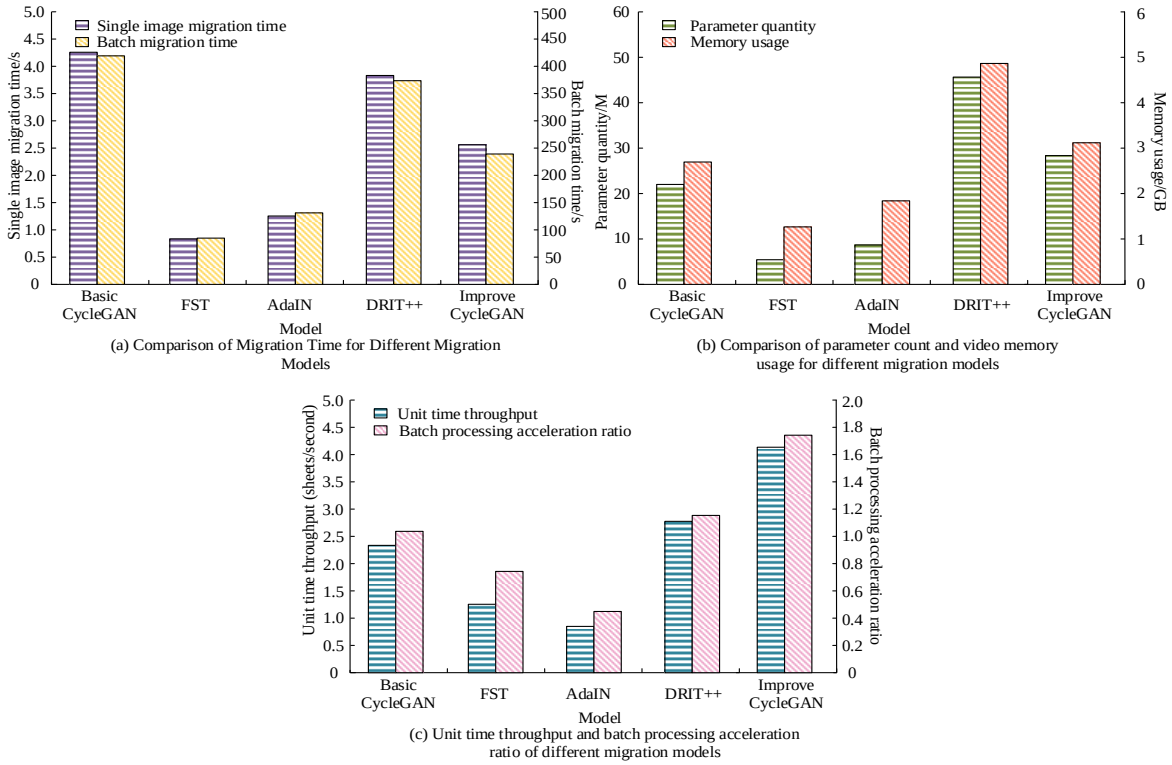
(a) Comparison of PSNR values (b) Comparison of SSIM values

Fig. 8. Transfer image quality between different models

In Table 4, the SOTA model QS-Attn exhibits the best performance (PSNR 24.17 dB, SSIM 0.83, FID 20.34), followed by CUT (24.07 dB, 0.82, 21.15), SwinTransformer-based (23.54 dB, 0.80, 23.09), and FastCUT (23.70 dB, 0.81, 22.68). All four models outperform the baseline CycleGAN (22.33 dB, 0.78, 28.56), but significantly lag behind the improved CycleGAN (25.10 dB, 0.87, 12.89). The FID of the improved CycleGAN is 36.63% lower than that of the best SOTA model QS-Attn, with a PSNR increase of 3.85% and an SSIM increase of 4.82%. This indicates that the fusion design of VGG19 multi-scale feature extraction and SAM global dependency modeling exhibits significant advantages in image quality for ceramic decorative pattern style transfer.

Table 4. Quantitative comparison of image quality among baseline, SOTA, and the improved model

Model	PSNR (dB)	SSIM	FID
Baseline CycleGAN	22.33	0.78	28.56
FST	23.05	0.79	26.87
AdaIN	23.82	0.81	24.35
DRIT++	24.56	0.84	19.42
CUT	24.07	0.82	21.15
FastCUT	23.7	0.81	22.68
QS-Attn	24.17	0.83	20.34
SwinTransformer-based	23.54	0.8	23.09
Improved CycleGAN	25.1	0.87	12.89



(a) Comparison of Migration Time for Different Migration Models

(b) Comparison of parameter count and video memory usage for different migration models

(c) Unit time throughput and batch processing acceleration ratio of different migration models

Fig. 9. Migration efficiency among various models

This study compares the migration efficiency of various migration models. In Fig.9(a), the single-graph migration time of the improved CycleGAN is 2.53 s, which is 40.75% lower than the 4.27 s of the basic CycleGAN. Although this time consumption is slightly longer than the 0.82 s and 1.25 s of the lightweight models FST and AdaIN, it is 33.60% lower than the 3.81 s of DRIT++. In terms of batch migration time, the improved CycleGAN takes 242.80 s, which is 42.24% and 35.28% lower than the 420.52 s of the basic CycleGAN and 375.24 s of DRIT++. In Fig.9(b), the parameter amount of the improved CycleGAN is 28.31M, which is an increase of 27.80% compared to the 22.15M of the basic CycleGAN, but is much lower than the 45.68M of DRIT++, which is a decrease of 38.03%. Although the parameter amount of the improved CycleGAN is slightly higher than the 5.29M and 8.59M of FST and AdaIN, the parameter amount is not excessively redundant and maintains good model complexity and balance. In terms of video memory usage, the improved CycleGAN's video memory usage is 3.12GB, which is lower than DRIT++'s 4.87GB, with a reduction of 35.93%. Although the video memory usage is slightly higher than that of FST and AdaIN, it can still run smoothly on ordinary computers, and the deployment threshold is low. In Fig.9(c), the model's unit time throughput and batch processing acceleration ratio are 0.41 pictures/s and 1.73, which are also significantly better than other models. Overall, the improved CycleGAN shows advantages in

both single-graph and batch migration efficiency. At the same time, it is also more reasonable in terms of parameter amount and memory usage, showing superior migration performance.

This study further compares the process characteristics of CDP style migration of different migration models, as shown in Table 5. The improved CycleGAN performs best in terms of glaze texture similarity, reaching 0.89, which is 43.55%, 25.35%, 18.67%, and 30.88% higher than FST, AdaIN, DRIT++, and basic CycleGAN. In terms of pattern breakage rate, the breakage rate of the improved CycleGAN is only 5.30%, which is 71.69%, 65.17%, 57.81%, and 68.52% lower than other models. In terms of classic glaze color restoration, the improved CycleGAN has the smallest color difference. The blue and white color difference and the alum red color difference are 3.24 and 4.13, which are far lower than the industry practical threshold of 5 and better than other models. The improved CycleGAN can effectively generate high-fidelity images. In terms of LPIPS, the improved CycleGAN achieves the lowest value of 0.187, which is 54.61% lower than the 0.412 of the basic CycleGAN, and 22.41%, 26.09%, 22.41%, and 29.97% lower than the 0.241, 0.253, 0.267, and 0.274 of QS-Attn, CUT, FastCUT, and the SwinTransformer-based model. Among the SOTA models, QS-Attn performs best with an LPIPS of 0.241, followed by CUT (0.253), FastCUT (0.267), and the SwinTransformer-based model (0.274). The LPIPS ranking is highly consistent with the glaze texture similarity and pattern fracture rate, further confirming that the improved CycleGAN achieves the best perceptual quality and process feature restoration.

Table 5. Transfer process characteristics of CDP styles using various models

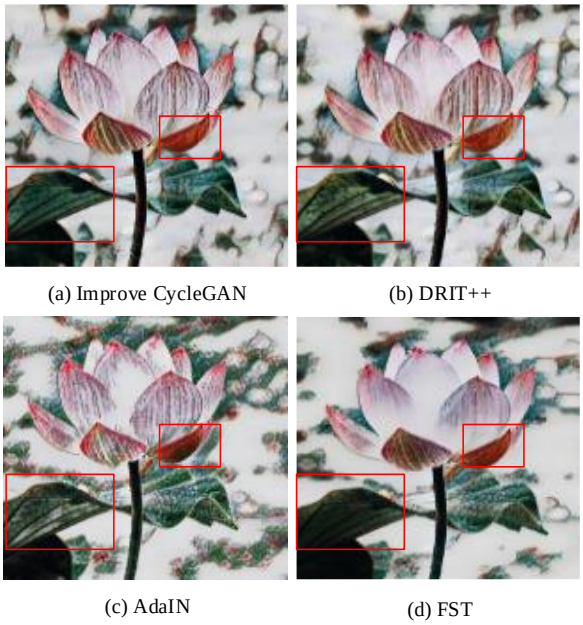
Model	LPIPS	Glazed texture similarity	Pattern fracture rate/%	Blue and white color difference	Alum red color difference
Basic CycleGAN	0.412	0.68	16.83	7.93	9.63
FST	0.386	0.62	18.72	9.80	12.32
AdaIN	0.341	0.71	15.21	8.16	10.54
DRIT++	0.278	0.75	12.56	6.74	8.90
CUT	0.253	0.77	10.24	5.87	7.35
FastCUT	0.267	0.79	11.36	6.21	7.82
QS-Attn	0.241	0.81	9.15	5.42	6.94
SwinTransformer-based	0.274	0.76	12.08	6.53	8.27
Improved CycleGAN	0.187	0.89	5.30	3.24	4.13

4.4 Practical Application Verification

To test the effect of model in practical applications, this paper compares the CDP migration results. In Fig.10(a), the improved CycleGAN's migration image has the best physical effect. The image has no blur or graininess, the pattern is continuous and complete, and the branches have no breaks. The improved CycleGAN model can generate high-quality migration images in practical applications.

This study conducts a subjective evaluation of the migration images of each migration model, with a full score of 5 points. The subjective evaluation panel consists of twelve evaluators, including four ceramic artists, three academic researchers in ceramic design, three industrial designers, and two graduate students majoring in ceramic art. A double-blind protocol is adopted, in which the evaluators are unaware of which model generated each image and the model identities are randomly coded. The inter-rater reliability is assessed by Krippendorff's Alpha, yielding a value of 0.82 for visual satisfaction, 0.79 for pattern integrity, and 0.81 for restoration degree of process characteristics, all exceeding the acceptable threshold of 0.75 and confirming the consistency of the panel scores. Table 6 shows the scores of different migration models. In the test set scoring results, the improved CycleGAN performs best in all evaluation indicators. In terms of visual satisfaction, the improved CycleGAN scores 4.86, which is 29.26%, 28.23%, and 20.59% higher than FST, AdaIN, and DRIT++. In terms of pattern integrity, the improved CycleGAN scores 4.42, which is 36.00%, 20.11%, and 9.95% higher than other models. In terms of process feature reduction, the improved CycleGAN scores 4.63, with an increase of 21.52%, 14.32%, and 9.72%. The scoring trend of the

training set is highly consistent with that of the test set. The visual satisfaction of the improved CycleGAN is as high as 4.92. Therefore, the improved CycleGAN is significantly better than other models in ceramic design transfer applications.



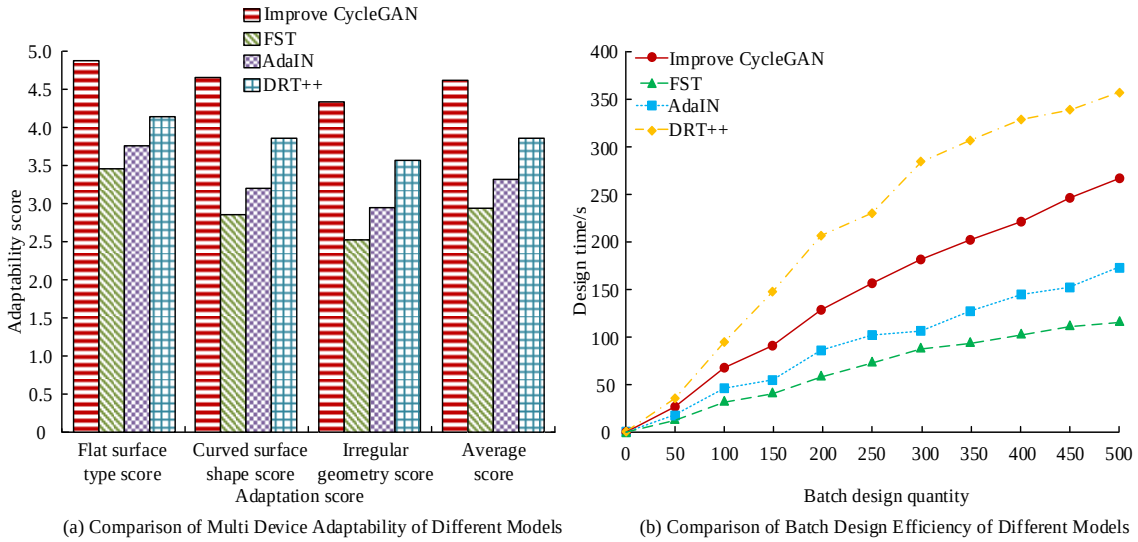
(a) Improve CycleGAN (b) DRIT++

(c) AdaIN (d) FST

Fig. 10. Migration results of CDPs using different models

Table 6. Scoring of different migration models

Dataset	Model	Subjective evaluation score		
		Visual satisfaction	Pattern integrity	Restoration degree of process characteristics
Test set	FST	3.76	3.25	3.81
	AdaIN	3.79	3.68	4.05
	DRIT++	4.03	4.02	4.22
	Improve CycleGAN	4.86	4.42	4.63
Training set	FST	3.85	3.34	3.90
	AdaIN	3.88	3.77	4.14
	DRIT++	4.11	4.10	4.30
	Improved CycleGAN	4.92	4.51	4.71



(a) Comparison of Multi Device Adaptability of Different Models (b) Comparison of Batch Design Efficiency of Different Models

Fig. 11. Multi-machine adaptability and batch design efficiency of different models

To further verify the adaptability of the improved model in actual ceramic production scenarios, this study compares the multi-device adaptability and batch design efficiency. In Fig.11(a), the improved CycleGAN shows the best adaptability in all device types. Its migration image scores for flat, curved, and irregular shapes are as high as 4.88, 4.65, and 4.32. Its average adaptation score is 4.62, which is 57.14%, 39.27%, and 20.00% higher than FST, AdaIN, and DRIT++ (2.94, 3.31, and 3.85). In Fig.11(b), the improved CycleGAN takes 28.62 s in a small batch design of 50. Although it is slower than FST's 12.58 s, it is 20.18% faster than DRIT++'s 35.84 s, which can meet the rapid response requirements of personalized customization. In a large batch design with a quantity of 500, the improved CycleGAN takes 268.20 s, 25.28% faster than DRIT++'s 358.91 s, indicating that its batch processing logic has no redundancy and can effectively support the large-scale production needs of ceramic factories. In summary, the improved CycleGAN can effectively improve production efficiency and design quality in actual production applications.

To verify the capability of the CIELAB color gamut intersection in Equation (17) to approximate non-linear firing color shifts, this study compares the generated glaze colors with real fired samples across four typical glaze categories: underglaze blue, alum red, crackle glaze, and modern comprehensive decoration. For each category, thirty real fired samples are collected from the Jingdezhen Ceramic Museum digital resources, and the corresponding generated samples are produced by the improved CycleGAN. The color difference is measured by the CIEDE2000 metric, and the hue shift is measured by the angular deviation in the CIELAB a*b* plane. Table 7 presents the quantitative results.

Table 7. Comparison between generated glaze colors and real fired samples in CIELAB space

Glaze category	Underglaze blue	Alum red	Crackle glaze	Modern comprehensive	Average
ΔE_{00}(generated vs. Real)	3.24	4.13	2.87	4.52	3.69
Hue shift (degrees)	4.15	5.82	3.46	6.73	5.04
Chroma deviation ΔC*	2.86	3.54	2.12	4.08	3.15
Gamut coverage ratio $G_{gen} \cap G_{gt}/G_{gt}$	0.91	0.87	0.93	0.84	0.89
Perceptual consistency (%)	92.4	89.6	94.1	87.3	90

As shown in Table 7, the average CIEDE2000 color difference between generated and real fired samples is 3.69, and the average gamut coverage ratio reaches 0.89, indicating that the CIELAB gamut intersection constraint in Equation (17) can approximate the non-linear firing color shift to a practically acceptable extent. Among the four categories, crackle glaze achieves the best consistency (color difference 2.87, hue shift 3.46 degrees, gamut coverage 0.93, perceptual consistency 94.1%), because its color distribution is relatively stable and less sensitive to firing atmosphere. In contrast, alum red and modern comprehensive decoration show larger deviations (color differences 4.13 and 4.52, hue shifts 5.82 and 6.73 degrees), which is consistent with the fact that copper-red and multi-process glazes are more sensitive to kiln temperature and atmosphere. Nevertheless, all four categories remain below the industry practical threshold of $\Delta E_{00} \leqq 5$, and the perceptual consistency exceeds 87%, confirming that the proposed glaze color constraint provides a reliable perceptual approximation of real firing color shifts.

5. Discussion and Interpretation

As the core value carrier of ceramic art, CDP's design innovation and efficiency improvement are of great significance to the inheritance of ceramic culture and industrial upgrading. To solve the limitations of the traditional design process that is time-consuming and labor-intensive and has fixed styles, this study proposes a CDP style migration model by improving DL and combining it with SAM.

In the experiment, the PSNR and SSIM values of the improved model were as high as 25.10 dB and 0.87, which were 12.37% and 10.25% higher than the basic CycleGAN. Compared with the method of Dinesh Kumar R et al. based only on VGG19, this paper strengthened the advanced feature extraction capabilities and also captured the long-distance dependence of patterns through the

dual-module fusion of SAM and VGG19. The research method stabilized CCL at 0.18, avoiding the common texture structure distortion phenomenon in the method of Zhou Z et al.

The current lack of unified evaluation standards in the field of ceramic pattern style transfer makes it difficult to compare the results of different studies horizontally, which restricts the overall development of the field. Therefore, this study proposes an operational ceramic pattern style transfer benchmark system (CPST-Bench), which consists of three weighted modules: objective indicators, subjective evaluation, and process adaptability testing. The objective module combines five min-max normalized indicators—PSNR, SSIM, LPIPS, FID, and glazed texture similarity—into a weighted score, with weights of 0.15, 0.20, 0.20, 0.25, and 0.20, respectively. Since lower LPIPS and FID indicate better performance, their normalized values are inverted before summation. The industry thresholds are PSNR no less than 24.0 dB, SSIM no less than 0.85, LPIPS no more than 0.25, FID no more than 20.0, and glazed texture similarity no less than 0.80. The subjective module averages the scores of at least ten ceramic design practitioners and aesthetic researchers across four dimensions—composition density, stroke charm, glaze color dryness, and process feasibility—rated on a five-point scale and weighted as 0.20, 0.30, 0.30, and 0.20, respectively, with an intraclass correlation coefficient above 0.75 to ensure consistency. The adaptability module averages three normalized sub-task scores: different machine type adaptation, different process compatibility, and batch production efficiency, where the last is quantified as one minus the ratio of actual to maximum allowable batch time, set to 500 seconds for a 500-piece batch. The final CPST-Bench score is the weighted sum of the three modules with weights of 0.40, 0.35, and 0.25. A model meets the deployment requirement only if the final score is no less than 0.80 and all five objective indicators satisfy their thresholds. This benchmark system provides a fair, reproducible, and quantifiable comparison platform for subsequent research, promotes technological iteration and resource sharing in the field, and forms lasting academic contributions.

It should be noted that the CIELAB color difference reported in Table 7 is measured between generated images and real fired samples in the sRGB color space, whereas ceramic inkjet printers operate with ceramic oxide inks (Co, Fe, Cr, Mn oxides) who's on-screen RGB values undergo non-linear color transformations after firing, so the AI output must be adapted to the printer-specific ICC profile before digital printing. This adaptation can be achieved by converting the generated sRGB image to the CIELAB space and then mapping it into the ceramic ink gamut through the ICC profile of the target printer, followed by a firing compensation curve that accounts for the color shift induced by kiln temperature and atmosphere; however, since such ICC profiles are usually device-specific and proprietary, future work should establish an open ceramic ink color database to support cross-device color consistency.

Considering the differences in hardware conditions at the design end of ceramic factories, model lightweighting is a key direction for achieving technological implementation. The research has reserved lightweight interfaces during the model architecture design phase, laying the foundation for subsequent deployment optimization.

6. Summary

CDP's ST technology has important application value in ceramic manufacturing. To meet the market demand for diversified and personalized products, this study constructed an improved CycleGAN migration model. The PSNR and SSIM values of the improved CycleGAN reached 25.10 dB and 0.87, which were better than other migration models. In addition, the parameter amount was 28.31M, and the video memory usage was 3.12GB, both of which were kept at a low level. After the integration and optimization of all modules, the lowest FID value of the model was only 12.89, which was an increase of 54.87% compared to the baseline module. In practical applications, the improved CycleGAN performed best in terms of image effects, with scores of 4.86, 4.42, and 4.63 for visual satisfaction, pattern integrity, and process feature restoration. In actual production applications, the improved CycleGAN demonstrated optimal adaptation capabilities in various device types. The average adaptation score of all device types was 4.62, which was an increase of more than 20.00% compared to the comparative model. The processing times in small batch and large batches designs were 28.62 s and 268.20 s, which could effectively support the large-scale

production needs of ceramic factories. The main contribution and discovery of the research lies in proposing an improved CycleGAN architecture that integrates SAM and VGG19. Through modules such as adaptive sharpening enhancement and feature detail enhancement, it solves the problems of loss of details and texture distortion in general style transfer models, and improves the accuracy and artistic expression of pattern transfer. In addition, the research has strengthened the cross integration with ceramic art theory, making the technological output more in line with the core aesthetic needs of ceramic art. We have established a benchmark system for the transfer of ceramic pattern styles, providing a unified standard for academic research and technical comparison in the field. In summary, the research results not only effectively improve the design efficiency and artistic quality of ceramic decorative patterns, provide technical support for ceramic art innovation, but also provide a practical path for the deep integration of traditional craftsmanship and artificial intelligence, which has important application value and academic significance.

References

[1] Chen J, Li S, Liu D, Lu W. Indoor camera pose estimation via style-transfer 3D models. Comput Aided Civ Inf. 2022;37(3):335-353. https://doi.org/10.1111/mice.12714

[2] Hu Z, Lee RKW, Aggarwal CC, Zhang A. Text style transfer: A review and experimental evaluation. ACM SIGKDD Explor Newsl. 2022;24(1):14-45. https://doi.org/10.1145/3544903.3544906

[3] Jang DK, Park S, Lee SH. Motion puzzle: Arbitrary motion style transfer by body part. ACM Trans Graph. 2022;41(3):1-16. https://doi.org/10.1145/3516429

[4] Jin D, Jin Z, Hu Z, Vechtomova O, Mihalcea R. Deep learning for text style transfer: A survey. Comput Linguist. 2022;48(1):155-205. https://doi.org/10.1162/coli_a_00426

[5] Xu K, Wang S, Jin Y, Che Q, Zhou B. Object detection-oriented style transfer network for panchromatic remote sensing image. J Appl Remote Sens. 2023;17(2):026503. https://doi.org/10.1117/1.JRS.17.026503

[6] Wang W, Yu X, Fang B, Zhao Y, Chen Y, Wei W, et al. Cross-modality LGE-CMR segmentation using image-to-image translation-based data augmentation. IEEE/ACM Trans Comput Biol Bioinform. 2022;20(4):2367-2375. https://doi.org/10.1109/TCBB.2022.3140306

[7] Fu X. Digital image art style transfer algorithm based on CycleGAN. Comput Intell Neurosci. 2022; 2022:6075398. https://doi.org/10.1155/2022/6075398

[8] Zhu X, Lin M, Yi M, Zhao H. Photorealistic attention style transfer network for architectural photography photos. Sci Rep. 2024;14(1):29584. https://doi.org/10.1038/s41598-024-81249-6

[9] Kumar RD, Julie EG, Robinson YH, Vimal S, Dhiman G, Veerasamy M. Deep convolutional nets learning classification for artistic style transfer. Sci Program. 2022; 2022:2038740. https://doi.org/10.1155/2022/2038740

[10] Zhou Z, Wu Y, Yang X, Zhou Y. Neural style transfer with adaptive auto-correlation alignment loss. IEEE Signal Process Lett. 2022; 29:1027-1031. https://doi.org/10.1109/LSP.2022.3165758

[11] Wang HY, Chien CW, Tsai MH, Lin IC. Hairstyle-and-identity-aware facial image style transfer with region-guiding masks. Multimed Tools Appl. 2024;83(17):51867-51897. https://doi.org/10.1007/s11042-023-17298-2

[12] Liao YM, Huang YF. Deep Learning-Based Application of Image Style Transfer. Math Probl Eng. 2022; 2022:1693892. https://doi.org/10.1155/2022/1693892

[13] Garg M, Ubhi JS, Aggarwal AK. Neural style transfer for image steganography and destylization with supervised image to image translation. Multimed Tools Appl. 2023;82(4):6271-6288. https://doi.org/10.1007/s11042-022-13596-3

[14] Mao W, Yang S, Shi H, Liu J, Wang Z. Intelligent typography: Artistic text style transfer for complex texture and structure. IEEE Trans Multimed. 2022; 25:6485-6498. https://doi.org/10.1109/TMM.2022.3209870

[15] Markham RP, López JME, Nieto-Hidalgo M, Tapia JE. Open-set: ID card presentation attack detection using neural style transfer. IEEE Access. 2024; 12:68573-68585. https://doi.org/10.1109/ACCESS.2024.3397190

[16] Mochizuki Z, Saito M, Suzuki T, Mochizuki K, Nemoto H, Onishi H, et al. Comparison of various GAN-based bone suppression imaging for high-accurate markerless motion tracking of lung tumors in CyberKnife treatment. Thorac Cancer. 2025;16(4): e70014. https://doi.org/10.1111/1759-7714.70014

[17] Jiang W, Qin Y, Wang X, Chen Q, Guan Q, Lu M. CycleH-CUT: An unsupervised medical image translation method based on cycle consistency and hybrid contrastive learning. Phys Med Biol. 2025;70(5):055014. https://doi.org/10.1088/1361-6560/adb2d7

[18] Vinay Kumar YB, Vimala HS. Skin cancer classification using PCNN-LinkNet model trained with monotonic and deep feature set. Biomed Mater Devices. 2026;4(3):3568-3600. https://doi.org/10.1007/s44174-025-00467-2

[19] Alshdaifat N, Owida HA, Mustafa Z, Aburomman A, Abuowaida S, Ibrahim A, et al. Automated blood cancer detection models based on EfficientNet-B3 architecture and transfer learning. Indones J Electr Eng Comput Sci. 2024;36(3):1731-1738. https://doi.org/10.11591/ijeecs.v36.i3.pp1731-1738
[20] Konstantinidis FK, Myrillas N, Tsintotas KA, Mouroutsos SG, Gasteratos A. A technology maturity assessment framework for Industry 5.0 machine vision systems based on systematic literature review in automotive manufacturing. Int J Prod Res. 2026;64(17):7269-7305. https://doi.org/10.1080/00207543.2023.2270588
[21] Wang H, Liu W, Chen J, Ji S. Transfer learning with a graph attention network and weighted loss function for screening of persistent, bioaccumulative, mobile, and toxic chemicals. Environ Sci Technol. 2025;59(1):578-590. https://doi.org/10.1021/acs.est.4c11085
[22] Sharma A, Lysenko A, Jia S, Boroevich KA, Tsunoda T. Advances in AI and machine learning for predictive medicine. J Hum Genet. 2024;69(10):487-497. https://doi.org/10.1038/s10038-024-01231-y
[23] Bintang YK, Imaduddin H. Pengembangan model deep learning untuk deteksi retinopati diabetik menggunakan metode transfer learning. JIPI (J Ilm Penelit Pembelajaran Inform). 2024;9(3):1442-1455. https://doi.org/10.29100/jipi.v9i3.5588
[24] Hosny KM, Mohammed MA. Explainable AI and vision transformers for detection and classification of brain tumor: A comprehensive survey. Artif Intell Rev. 2025;58(9):259-265. https://doi.org/10.1007/s10462-025-11221-x
[25] Wang Y, Bai J, Eshaghi MS, et al. Transfer learning in physics-informed neural networks: Full fine-tuning, lightweight fine-tuning, and low-rank adaptation. Int J Mech Syst Dyn. 2025;5(2):212-235. https://doi.org/10.1002/msd2.70030
[26] Sharma A, Lysenko A, Jia S, Boroevich KA, Tsunoda T. Advances in AI and machine learning for predictive medicine. J Hum Genet. 2024;69(10):487-497. https://doi.org/10.1038/s10038-024-01231-y