ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Generative adversarial networks are implemented in intangible cultural heritage art style transfer image enhancement technology

Wei Yang[1,*]

[1] School of Art and Design, Mingde Institute of Technology, Xi'an, 710000, Shaanxi, China

SUMMARY: *In order to improve the visual quality and style controllability of intangible cultural heritage art images, this paper proposes a generative adversarial implementation method for style transfer and image enhancement. A dataset consisting of 5240 images, including 4160 training images and 1080 testing images, is constructed from embroidery, paper cutouts, New Year paintings, and lacquer patterns. After denoising, brightness correction and label coding, the texture structure and color semantics are encoded and input into the CycleGAN framework under the constraints of PatchGAN, cycle consistency and structure-color joint constraints. The model was trained using Adam for 180 epochs with a batch size of 4 and an initial learning rate of 0.0002. Experimental results show that the PSNR value of the proposed method is 24.76 dB, the SSIM value is 0.842, and the style consistency evaluation value is 0.913. At the same time, the visual clarity and edge integrity of different heritage categories are improved. The framework provides an efficient computational solution for digital rendering, restoration, and cross-style generation of heritage images.*

KEYWORDS: *Cycle generative adversarial networks; Intangible cultural heritage image; Style transfer; Image enhancement*

1 Introduction

1.1 Intangible cultural heritage artistic style transfer and image enhancement task definition

In the digital context of intangible cultural heritage art, images are no longer just static display carriers, but structured objects after entering the computer vision system. Paper-cut, embroidery, wood-block New Year pictures, lacquer patterns and weaving patterns all have stable texture rules, color logic and symbol boundaries. These elements not only determine the separability of artistic style recognition, but also determine the computational target of image enhancement after style transfer.

Han et al. studied the style transfer method of artistic images based on deep extraction generative adversarial networks, and proposed that the style mapping path was constrained by deep structural information, so that the generated results maintained the identificability of the main contour during the texture transfer process [1]. Batziou et al. studied an artistic neural style transfer scheme combining CycleGAN and FABEMD, and proposed an adaptive information selection mechanism to coordinate the mapping strength between content preservation and style expression [2]. Zheng et al. studied the style transfer model of

---

[*] 18792971168@163.com

https://doi.org/10.65102/is2026541

cross-fusion attention and frequency domain loss cooperation, and proposed the CFA-GAN structure to constraint the quality of detail reconstruction with frequency components, so as to improve the continuity of high-frequency patterns in the generation stage [3]. Tahir et al. studied a multi-domain face image style transfer network and proposed an adversarial generation strategy for diverse style distributions, indicating that the generator needs stable feature alignment ability when dealing with cross-domain visual differences [4]. Christophe et al. studied the map style transfer method based on GAN, and proposed transfer control from the perspective of visual symbol organization and style expression consistency, which provided a reference calculation idea for image scenes with regular patterns and obvious layout structures [5]. Khowaja et al. studied the user-controllable mechanism in the hierarchical style transfer network and proposed a multi-level style adjustment strategy, so that local texture, color distribution and global visual perception could be controlled in a unified network [6].

For intangible cultural heritage images, the task of style transfer is not to simply replace colors or copy brush strokes, but to complete style generation on the basis of maintaining the semantics of patterns, modeling boundaries and technical characteristics. At the same time, the readability of low contrast regions, weak edge regions and fine-grained texture regions is improved through the enhancement process. Therefore, the tasks of intangible cultural heritage artistic style transfer and image enhancement can be defined as follows. Taking multi-category intangible cultural heritage images as input, the joint computing tasks of content structure preservation, style attribute mapping and visual quality improvement are completed under the generative adversarial framework. The core goal is to establish a learnable representation for texture, color and semantic relationships, and form a technical link suitable for digital display, repair and reconstruction, and cross-style generation. At the calculation implementation level, the input terminal needs to complete the category division, style label arrangement and resolution unification, and the output terminal requires that the generated image satisfies the calculation evaluation in terms of local texture clarity, color stability and overall style consistency, and retains the process details and hierarchical relationship feature values.

1.2 Technical ideas and research contributions of this paper

Feng et al. studied an image style transfer method based on a combined Transformer autoencoder, and proposed to improve the coupling accuracy between content representation and style representation through hierarchical feature reorganization [7]. Wei studied the artistic image style transfer method based on CycleGAN network model, and proposed to complete the stable mapping between unpaired images under the cyclic consistency constraint, which provided an executable framework for cross-category visual transfer [8]. Wang et al. studied the ESA-CycleGAN model that combines edge features and self-attention mechanism, and proposed to use edge response to enhance structure preservation ability, so that the style transfer results show clearer detail boundaries in the contour area [9]. Bao et al. studied the font style transfer model of Dongba characters based on AFGAN, and proposed to apply generative adversarial learning to the calculation of visual style of traditional culture, which provides a strong scene reference for the digital expression of intangible cultural heritage images [10]. Yang et al. studied the Transformer-GAN low-light image enhancement method, and proposed to introduce the converter representation in the generation process to improve the brightness recovery and local texture reconstruction effect [11]. Yu et al. studied the low-illumination enhancement method of unsupervised generative adversarial networks, and proposed a training strategy to optimize image quality under the condition of lack of paired samples [12].

Based on the above research paths, we take intangible cultural heritage images as the processing object, and construct a computing link consisting of sample organization, preprocessing and correction, joint feature expression, and adversarial generation and enhancement, so as to synchronously maintain pattern semantics, modeling structure, and color level in the unpaired style transfer scenario. The technical idea of this paper is not to separate style transfer and image enhancement into two separate processes, but to complete the collaborative learning of content constraint, style mapping and quality improvement in a unified generation framework.

The corresponding research contributions are in three levels: unified representation, co-training and multi-scene implementation. The intangible cultural heritage texture structure, process boundaries and color distribution are incorporated into the same representation space, and the computability of complex drawings is enhanced. The style transfer output and the image enhancement target are written into the training process of the generator, and the visual restoration ability of weak texture regions and low contrast regions is improved. Multiple categories of intangible cultural heritage images, such as paper-cutting, embroidery, New Year paintings, and lacquerwork, are incorporated into the reusable implementation process, and the model results can serve for digital display, restoration generation, and cross-style visual design. This path emphasizes the optimization of cross-domain mapping and visual enhancement with the generative adversarial network as the core, and maintains the stable transfer ability of the model on multi-class intangible cultural heritage images.

2 Related work

Focusing on image enhancement and style transfer driven by generative adversarial networks, existing research has formed a number of technical paths from visual quality restoration to cultural image reconstruction. Wang et al. studied the low-light image enhancement method based on the generative adversarial network, and proposed to improve the visibility of the dark field area through the joint learning of brightness distribution and detail response, so that the image maintained high structural stability after enhancement [13]. Jin et al. studied the event-guided dual-branch GAN low-illumination enhancement method, and proposed to introduce event information in addition to the main image branch to assist modeling, so as to improve the texture recovery ability and edge response accuracy under complex lighting conditions [14]. Wu et al. studied the FW-GAN underwater image enhancement method with multi-scale fusion, and proposed to introduce a multi-level feature aggregation mechanism in the generation stage to coordinate the mapping relationship between color deviation correction, contrast improvement and detail reconstruction [15]. Jia et al. studied an unsupervised underwater image enhancement method based on edge extraction, and proposed to combine edge information to constrain the generation process under the condition of no paired samples, so as to make the enhancement results more consistent in clarity and boundary continuity [16].

The research on cultural heritage and intangible cultural heritage image processing is pushing deep learning from general visual enhancement to object-oriented fine reconstruction. Karimi et al. studied the automatic detection method of ceramic tile defects in Portuguese cultural heritage buildings, and proposed to use deep learning to identify and classify surface damage areas, which provided a structured scheme for local anomaly analysis of traditional decorative images [17]. Stoean et al. studied the 3D restoration method of degraded cultural relics for museum digital display, and proposed an artificial intelligence-driven reconstruction process to connect historical fragments and digital restoration models, so as to make visual display have higher reduction and interactive adaptation capabilities [18]. Altaweel and Khelifi studied the reconstruction method of cultural relics based on generative artificial intelligence, and proposed to take the Roman coin as an example to realize the generative completion of the defective area, which illustrates the practicability of the generative model in historical pattern restoration and visual deduction [19]. Zhu et al. studied the transfer learning method of face restoration under the condition of small samples, and proposed to combine transfer learning to complete the detail compensation in the repair scene of the cracked face of the Terracotta Warriors and the damaged nose of the Buddha statue, so that the cultural heritage images can still maintain a good structural restoration effect under limited samples [20].

The above studies show that the generative adversarial network has formed a strong technical accumulation in visual enhancement, texture restoration, edge preservation and cultural image restoration. However, most of the methods are respectively modeling for low-light, underwater environment, lost cultural relics or local restoration scenes, and style transfer and image enhancement often exist as two adjacent but separate processing links. For intangible cultural heritage images, the patterns are repeated, the process boundaries are clear, the color levels are complex, and the categories are significantly different. Such visual objects need not only the ability of cross-style mapping, but also the synchronous maintenance of local details, weak texture areas and color stability in the enhancement stage. The research objects, main methods and adaptation scope of the existing related methods are shown in Table 1.

Table 1: Comparison of related work

Reference	Research Object	Main Method	Result Characteristics	Applicable Scope
Wang et al. [13]	Low-light image enhancement	GAN-based brightness restoration	Dark-scene details and structural clarity are improved simultaneously	Suitable for low-light enhancement, but does not directly perform style transfer
Jin et al. [14]	Image enhancement under complex illumination	Event-guided dual-branch GAN	More stable texture restoration and more sensitive edge responses	Suitable for dynamic lighting scenarios, but its cross-style mapping capability is limited
Wu et al. [15]	Underwater image enhancement	Multi-scale fusion FW-GAN	Color cast correction and detail reconstruction are performed collaboratively	Suitable for degraded image enhancement, and its multi-scale design can be used for reference
Jia et al. [16]	Unsupervised underwater enhancement	GAN with edge extraction constraints	Clarity and boundary continuity are preserved under unpaired conditions	Suitable for edge preservation, but does not directly address cultural style representation
Karimi et al. [17]	Defect detection in architectural tiles	Deep learning-based detection and classification	Strong structural specificity in local damage identification	Suitable for traditional decorative image analysis and can assist in the detection of intangible cultural heritage patterns
Stoean et al. [18]	3D restoration of degraded cultural relics	AI-driven digital reconstruction	High visual restoration fidelity and display adaptability	Suitable for cultural relic restoration, but relatively weak in 2D style transfer
Altaweel and Khelifi [19]	Cultural artifact reconstruction	Generative AI-based completion	Missing regions can be generated with the ability to infer historical patterns	Suitable for cultural relic completion, and its generative restoration pathway can be used for reference
Zhu et al. [20]	Small-sample facial restoration of cultural relics	Transfer learning-based restoration	Good structural restoration can still be maintained with limited samples	Suitable for damaged restoration, but offers limited support for multi-category style mapping

Table 1 shows that although the existing methods have achieved good computational results in brightness restoration, edge preservation, defect completion and cultural heritage reconstruction, respectively, there are still obvious differences in the modeling objectives of different studies. Low-light enhancement methods pay more attention to visibility restoration and local texture brightening, underwater enhancement methods pay more attention to color deviation correction and multi-scale detail restoration, and cultural heritage restoration methods focus on structure completion and historical visual restoration. Compared with these studies, the processing of intangible cultural heritage images requires not only the ability of cross-style mapping of the generative model, but also the simultaneous maintenance of pattern boundaries, technical details and color levels in the enhancement process. Therefore, the style transfer, edge preservation and visual enhancement are integrated into the same framework for collaborative modeling, which is more in line with the computational characteristics of intangible cultural heritage images. It also provides a clear technical basis for subsequent model design.

3 Generative adversarial implementation framework for intangible cultural heritage art style transfer

3.1 Data construction and sample organization of intangible cultural heritage art images

The data construction of this paper does not use general face or natural scene data, but establishes a dedicated image set around the task of intangible cultural heritage art style transfer and image enhancement. The sample source consists of three parts: public digital catalogue, museum authorized images and physical photography collection, covering four types of objects: paper cutting, embroidery, woodblock New Year pictures and lacquer art patterns, with a total of 5240 pieces. In order to ensure the stability of cross-category visual mapping, five metadata including process category, dominant color system, pattern density, boundary clarity and image source are kept synchronously in the acquisition stage, so that the subsequent generative adversarial training not only depends on pixel input, but also can read style-related structural information. After the original image enters the data pipeline, it is first uniformly cropped to a square view field, then scaled to 256×256 resolution, and color space alignment and channel normalization are completed. The sample normalization is expressed as Equation (1).

$$\mathrm{X_i} = \left(\frac{\mathrm{R(I_i)} - \mu}{\sigma + \varepsilon}\right) \odot \mathrm{M_i} + \lambda\,(1 - \mathrm{M_i}) \tag{1}$$

where $\mathrm{I_i}$ represents the i original image, $\mathrm{R}(\cdot)$ represents the unified resolution map after scaling and clipping,$\mu = [\mu_1, \mu_2, \mu_3]$ and $\sigma = [\sigma_1, \sigma_2, \sigma_3]$ represent the mean vector and standard deviation vector of the three color channels respectively, ε represents the smoothing term, $\mathrm{M_i}$ represents the effective pixel mask extracted from the foreground region, $\odot$ represents element-wise multiplication. Let λ denote the background compensation coefficient. The function of this formula is not simply to complete channel normalization, but to suppress the interference of invalid background regions on style distribution synchronously in the standardization process, so that the body of the pattern, the boundary lines and the color level maintain clearer statistical consistency in the input stage, so as to provide a more stable data basis for the subsequent construction of style description vectors and generative adversarial training. In order to form a unified data description of multi-source intangible cultural heritage images before entering generative adversarial training, this paper summarizes the sample organization fields and corresponding roles as shown in Table 2.

Table 2: Sample organization field design of Intangible cultural heritage art images

Field Name	Field Meaning	Value Format	Role in the Task
Sample_ID	Unique sample identifier	String-based ID	Used to index the original image, enhanced image, and style transfer result
Craft_Type	Craft category	Paper-cutting, embroidery, woodblock New Year painting, lacquer pattern	Distinguishes the style domain to which the intangible cultural heritage image belongs
Source_Type	Data source	Digital catalog, authorized museum collection image, physical-object photography	Marks the image acquisition pathway and controls source bias
Resolution_Tag	Original resolution label	High, Medium, Low	Records differences in input clarity and assists in enhancement strategy setting
Color_Mode	Dominant color mode	Warm, Cool, Mixed	Reflects color distribution characteristics and serves style mapping
Texture_Level	Texture density level	Sparse, Moderate, Dense	Describes pattern complexity and assists texture modeling
Edge_Clarity	Edge clarity	Weak, Medium, Strong	Reflects contour distinguishability and serves edge-preserving constraints
Style_Vector	Style description vector	Real-valued vector	Used as a joint conditional input for style transfer and enhancement
Domain_Label	Source/target domain label	Source, Target	Supports unpaired style transfer training
Split_Tag	Data subset label	Train, Val, Test	Controls dataset usage during the training, validation, and testing stages

In the table, the training set and the test set are hierarchically divided according to the style category, and the proportion of samples from different sources is kept consistent, so as to avoid the occurrence of a certain source in a single subset. In addition to the basic image tensor, this paper also constructs three types of style description vectors of texture, color and structure for each sample, whose fusion representation is shown in Equation (2).

$$s_i = \alpha h_i^{tex} + \beta h_i^{clr} + \gamma h_i^{str}, \quad \alpha + \beta + \gamma = 1 \tag{2}$$

Among them, h_i^{tex} represents texture coding, h_i^{clr} represents color coding, h_i^{str} represents structure coding, α, β, γ represent the fusion weights of the three types of features, and satisfy the sum of 1. The function of this formula is to organize the most recognizable style information in the intangible cultural heritage image into a unified representation, so as to facilitate the completion of style constraints in the subsequent unpaired transfer. In order to ensure that the sampling frequencies of the four types of samples remain balanced in the training process, this paper further constructs the class-balanced sampling probability, which is defined as Equation (3).

$$w_k = \frac{N}{K n_k}, \quad p_i = \frac{w_{y_i}}{\sum_{j=1}^{N} w_{y_j}} \tag{3}$$

Here, N represents the total sample size, K represents the number of classes, n_k represents the number of samples of the k class, w_k represents the class weight, and p_i represents the extraction probability of sample i. The function of this formula is to weaken the dominant influence of high-frequency categories on the update direction of the generator, so that the paper-cut contour, embroidery texture, year painting color layer and paint surface texture can obtain more stable learning opportunities in the same training cycle.

In the sample organization stage, each training batch contains source domain images, target domain images and their style description vectors, and the intra-batch structure is generated according to the dual constraints of class balance and source balance. This reduces inter-domain bias and enhances the interpretability of transfer results, while making the training update cadence overall smoother. After the above data construction and sample organization, the four types of intangible cultural heritage images form a unified computational representation in terms of source proportion, category distribution, style description and input scale. After this process, the subsequent denoising correction, style label encoding, and generative adversarial transfer enhancement can be established on a consistent data base, and the model training process is easier to maintain a balance between structure learning and style learning.

3.2 Image denoising correction and style label preprocessing process

Image denoising correction and style label preprocessing are not simple image cleaning steps, but the structure sorting process before the intangible cultural heritage art samples enter the generative adversarial network. The high-contrast contour of paper-cut, the fine line of embroidery, the large-area comprehensive color layer of New Year paintings, and the reflective area of paint patterns all have different noise distribution and color shift characteristics. If they are directly input into the unified network, it is easy to miswrite the acquisition noise as style texture in the style transfer stage. Its overall process is shown in Fig. 1. In the figure, there are four steps in turn: noise removal, brightness and color correction, pixel normalization mapping, and style label coding. Denoising is used to weaken compressed particles, background interference and local light spots, and preserve the boundary details of paper-cut contours, embroidery stitches and lacquer patterns. Brightness and color correction are used to reduce the exposure difference and color shift under different acquisition conditions. The normalized mapping is used to unify the numerical range of each channel. The style label encoding encodes the craft category, dominant color system, texture level and boundary clarity into a unified vector. After the above processing, a sample tensor containing both normalized pixel information and style semantic information is formed at the output, which can be directly sent to the generative adversarial network for training.

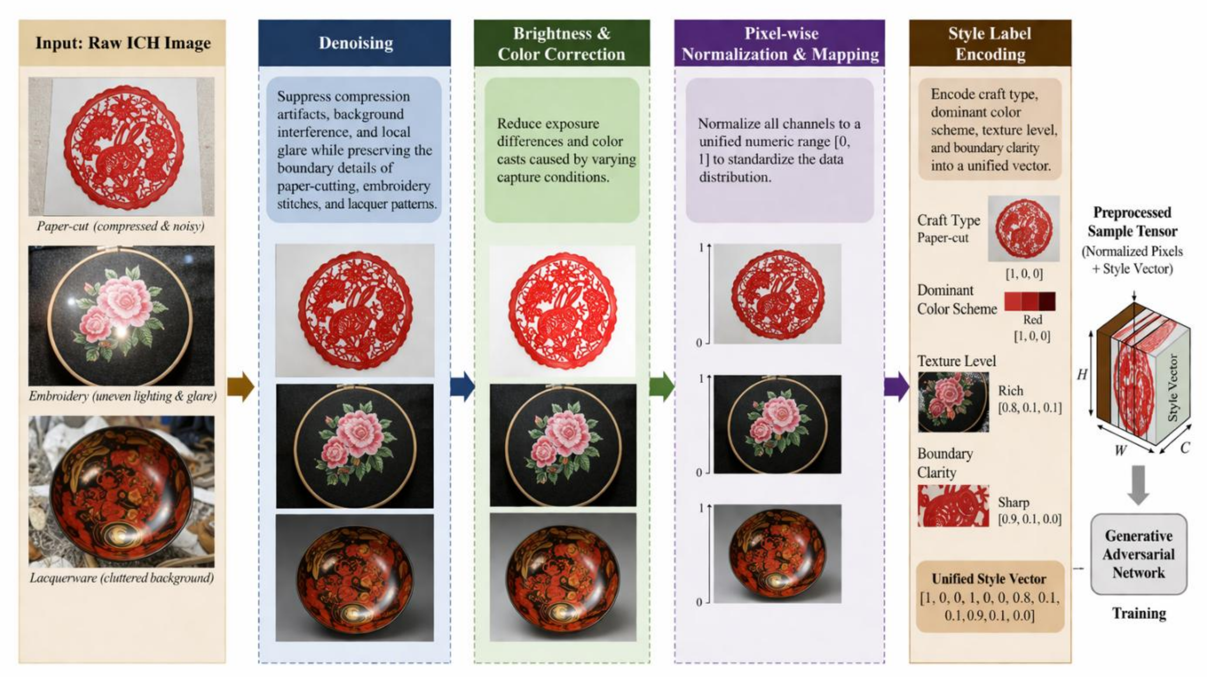

Figure 1: Process of image preprocessing and style label generation for intangible cultural heritage art

In the denoising stage, edge-preserving filtering is used to weaken compression noise and shooting particles, and the filtering result is expressed as Equation (4).

$$\mathrm{I_d}(\mathrm{p}) = \frac{1}{\mathrm{K_p}} \sum_{\mathrm{q} \in \Omega} \mathrm{G_s}\,(\|\mathrm{p}-\mathrm{q}\|)\,\mathrm{G_r}(|\mathrm{I}(\mathrm{p})-\mathrm{I}(\mathrm{q})|)\,\mathrm{I}(\mathrm{q}) \tag{4}$$

Here, $\mathrm{I}(\mathrm{p})$ represents the original value at pixel p, Ω represents the local neighborhood, $\mathrm{G_s}$ and $\mathrm{G_r}$ represent the spatial kernel and gray kernel, respectively, and $\mathrm{K_p}$ represents the normalized coefficient. The function of this formula is to smooth the local noise while preserving the detailed fluctuation of the paper-cut boundary, the embroidery stitch and the lacquer art pattern. After noise suppression is completed, the sample enters the brightness correction stage, and the correction result is expressed as Equation (5).

$$\mathrm{L_i^*} = \mathrm{L_i} + \delta(\bar{\mathrm{L}} - \mathrm{L_i}) \tag{5}$$

Here, $\mathrm{L_i}$ represents the image brightness component, $\mathrm{L_i^*}$ represents the corrected brightness result, $\bar{\mathrm{L}}$ represents the average brightness within the batch, and δ represents the brightness compensation coefficient. The function of this formula is to reduce the bright-dark shift caused by different acquisition environments, so that the subsequent enhancement process is more focused on style expression rather than exposure difference. The pixel standardization process uses Equation (6) to complete the unified mapping.

$$\mathrm{X_i^{(c)}} = \frac{\mathrm{I_d^{(c)}} - \min(\mathrm{I_d^{(c)}})}{\max(\mathrm{I_d^{(c)}}) - \min(\mathrm{I_d^{(c)}}) + \varepsilon} \tag{6}$$

Here, $\mathrm{X_i^{(c)}}$ represents the normalized result of the i image on the c channel, $\mathrm{I_d^{(c)}}$

represents the corresponding channel pixel value after denoising, $\max(I_d^{(c)})$ and $\min(I_d^{(c)})$ represent the maximum and minimum values within this channel, respectively, and ε represents the smoothing term that prevents the denominator from being zero. The function of this formula is to compress the fluctuations in the numerical range of samples from different sources into a unified interval, so that the generator and the discriminator can obtain a consistent input scale in the early stage of training, and reduce the interference of brightness differences on the direction of style transfer learning. The style label does not directly use a single category number, but is composed of process category, main color system, texture level and boundary clarity, and its coding form is shown in Equation (7).

$$y_i = \mathrm{Concat}\left(e_i^{craft}, e_i^{color}, v_i^{tex}, v_i^{edge}\right) \tag{7}$$

Here, y_i represents the style label vector of the i sample, $\mathrm{Concat}(\cdot)$ represents the vector splicing operation, e_i^{craft} represents the process category embedding, e_i^{color} represents the dominant color system embedding, v_i^{tex} represents the texture level vector, and v_i^{edge} represents the boundary sharpness vector. The function of this formula is to uniformly encode discrete category information and continuous style attributes into computable labels, so that the label information can participate in the subsequent style constraints in the form of vectors, so that the unpaired transfer process can maintain a clearer direction.

In the label organization stage, the continuous attributes and discrete attributes of the same image are written into the sample description file synchronously. The process category is used to distinguish the style domain, the dominant color system is used to constrain the color transfer direction, the texture level is used to describe the pattern density, and the boundary definition is used to describe the contour fidelity requirement. After this preprocessing, each sample not only has a unified pixel expression, but also has a style semantic index that can be directly called by the network, which enables the generator to read the visual content and label prior at the same time when learning the cross-domain mapping, and improves the consistency between the enhancement result and the target style. Therefore, the preprocessing no longer stays at the input cleaning level, but becomes the pre-computation module connecting the original intangible cultural heritage image, the style label, and the generative adversarial training. This organization also provides a stable entry for subsequent joint feature expression and transfer enhancement, and forms a unified data input interface.

3.3 Joint feature expression mechanism of texture structure and color semantics

Joint feature representation is the core link in the style transfer enhancement model of intangible cultural heritage art. Its role is not to simply superimpose texture description and color description, but to establish a stable relationship between structural information, color semantics and local style strength in a unified representation space. The paper-cut images emphasize boundary turning and hollow rhythm, the embroidery images emphasize stitch direction and texture density, the New Year wood-block pictures emphasize block surface distribution and comprehensive color layer, and the lacquer art patterns include reflection, lamination and surface fine crack structure at the same time. If only a single convolutional feature is relied on, the transfer result is prone to weakening of pattern trend, drifting of color relationship and compression of local semantic level. Therefore, we incorporate gradient direction, color statistics, regional semantics and feature weight into the expression process, so that the structure contour, dominant color configuration and regional semantics in the cultural heritage image form a unified and stable computational representation before entering the generator.

The texture orientation information is obtained through the local gradient response, whose basic representation is shown in Equation (8).

$$m(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2}, \quad \theta(x,y) = \arctan\frac{G_y(x,y)}{G_x(x,y) + \epsilon} \tag{8}$$

Here, $G_x(x,y)$ and $G_y(x,y)$ represent the transverse and longitudinal gradient components, $m(x,y)$ represents the gradient magnitude, $\theta(x,y)$ represents the gradient direction, and ϵ represents the smoothing term to prevent the denominator from being too small. The function of this formula is to convert the boundary ups and downs, texture turns and line directions in the intangible cultural heritage images into computable local geometric information, so that the paper-cut contours, embroidery stitches and line drawings of New Year paintings have more clear structural basis in the subsequent migration process. After obtaining the pixel gradient, this paper further statistics the orientation histogram by unit block, and its normalized form is shown in Equation (9).

$$h_b = \frac{v_b}{\sqrt{\|v_b\|_2^2 + \xi^2}} \tag{9}$$

Here, h_b represents the directional statistical vector of the b block, v_b represents the unnormalized gradient cumulative result, $\|v_b\|_2^2$ represents the two-norm, and ξ represents the smoothness constant. The function of this formula is to reduce the interference of brightness fluctuations on the direction features, make the intangible cultural heritage images obtained under different acquisition conditions maintain good stability in structure expression, and provide consistent boundary information for cross-style transfer. The color semantic part is modeled jointly by channel statistics and dominant color weight, and its representation is shown in Equation (10).

$$c_i = [\mu_r, \mu_g, \mu_b, \sigma_r, \sigma_g, \sigma_b, \rho_i] \tag{10}$$

Here, c_i represents the color semantic vector of the i sample, μ_r, μ_g, μ_b represent the mean of the three channels, σ_r, σ_g, σ_b represent the standard deviation of the three channels, and ρ_i represents the dominant color proportion vector. The function of this formula is to organize the comprehensive color layer, the main color configuration and the color dispersion degree into a unified color description, so that the generator does not only copy the local color when performing style transfer, but also learn more stable color relationships inside the intangible cultural heritage image. In order to avoid that the local statistics cannot fully express the regional semantics, this paper further introduces the regional embedding vector, whose definition is shown in Equation (11).

$$r_i = \sum_{k=1}^{K} \eta_k \, f(P_k), \quad \sum_{k=1}^{K} \eta_k = 1 \tag{11}$$

Here, r_i represents the region semantic vector of the i sample, P_k represents the k region block, $f(P_k)$ represents the region convolutional coding, and η_k represents the region weight. The function of this formula is to introduce the local semantics and regional distribution relationship of patterns into the joint representation, so that the model can distinguish the dense texture area, the color transition area and the border concentrated area, so as to enhance the style identifability between different intangible cultural heritage categories. In order to make the contribution ratio of various types of features in the unified space more reasonable, this paper uses attention weight to recalipay, and its calculation form is shown in Equation (12).

$$\mathrm{a_i} = \sigma(\mathrm{W_a}(\mathrm{h_b} \oplus \mathrm{c_i} \oplus \mathrm{r_i}) + \mathrm{b_a}), \quad \hat{\mathrm{a}}_\mathrm{i} = \frac{\mathrm{a_i}}{\|\mathrm{a_i}\|_1 + \varepsilon} \tag{12}$$

Here, $\mathrm{a_i}$ represents the original attention response vector of sample i, $\hat{\mathrm{a}}_\mathrm{i}$ represents the normalized attention weight, $\mathrm{W_a}$ and $\mathrm{b_a}$ represent the learnable parameters, $\sigma(\cdot)$ represents the nonlinear mapping function, $\oplus$ represents the vector concatenation operation, $\|\mathrm{a_i}\|_1$ represents the one-norm, and ε represents the smoothing term. Instead of assigning only a rough weight to the joint features, this formula first generates multi-dimensional attention responses through the learnable mapping, and then controls the relative contribution of each feature component through the normalization constraint, so that the texture, color and regional semantics can be coordinated and assigned in a unified scale. The final joint feature representation is shown in Equation (13).

$$\mathrm{z_i} = \mathrm{W_h}(\hat{\mathrm{a}}_\mathrm{i}^\mathrm{h} \odot \mathrm{h_b}) \oplus \mathrm{W_c}(\hat{\mathrm{a}}_\mathrm{i}^\mathrm{c} \odot \mathrm{c_i}) \oplus \mathrm{W_r}(\hat{\mathrm{a}}_\mathrm{i}^\mathrm{r} \odot \mathrm{r_i}) \tag{13}$$

Here, $\mathrm{z_i}$ represents the final joint feature of the i sample, $\hat{\mathrm{a}}_\mathrm{i}^\mathrm{h}$, $\hat{\mathrm{a}}_\mathrm{i}^\mathrm{c}$ and $\hat{\mathrm{a}}_\mathrm{i}^\mathrm{r}$ represent the weight sub-vectors assigned to structural features, color features and regional semantic features respectively, $\mathrm{W_h}$, $\mathrm{W_c}$ and $\mathrm{W_r}$ represent the linear mapping matrix of the corresponding components, $\odot$ represents the element-wise weighting operation, and $\oplus$ represents the feature concatenation operation. The function of this formula is to complete the weight adjustment in each branch, and then to press different types of style information into the same expression space through branch mapping, so that the input received by the generator not only contains the local pattern and comprehensive color layer information, but also contains the regional semantic relationship after weight calibration.

After the above construction, texture direction, color semantics and region semantics no longer enter the generative network in the form of separation from each other, but form a unified joint feature representation through weight constraints and mapping reorganization. In this way, the obtained input can more accurately correspond to the contour direction, technical patterns and comprehensive color layer structure in the intangible cultural heritage art image, and make the subsequent style transfer and enhanced reconstruction have a more reliable feature basis while maintaining structural integrity, color stability and local detail clarity

3.4 Design of style transfer enhancement model driven by Generative Adversarial networks

After the completion of data organization, preprocessing and joint feature representation, this paper further constructs a generative adversarial transfer enhanced model for intangible cultural heritage art images. Instead of directly following the general CycleGAN structure, the model adds style constraints and enhancement compensation within the bidirectional mapping framework, so that the transfer results maintain the stability of pattern boundaries, comprehensive color layers and local details while completing cross-domain style transfer. The model consists of a generator from source domain to target domain, a reverse generator from target domain to source domain, and two corresponding discriminators. The source domain mainly corresponds to the intangible cultural heritage images to be transferred, and the target domain corresponds to the reference style domain images. In order to avoid the style transfer process only stopping at the surface color change, we introduce texture structure guidance in the front end of the generator, and add an enhancement compensation path in the decoding stage, so that the model output can maintain the stability of pattern boundaries, comprehensive color levels and local details while completing cross-domain style transfer.

The overall structure is shown in Fig. 2. The left input of the figure includes the intangible cultural heritage image to be transferred and the corresponding joint feature information. The middle part completes content coding, style mapping and enhancement compensation in turn. Such a bidirectional structure does not simply pursue visual changes, but puts style transfer, detail recovery and structural constraints into the same generation link, so that the model always retains the core visual organization of intangible cultural heritage images during the transfer process.

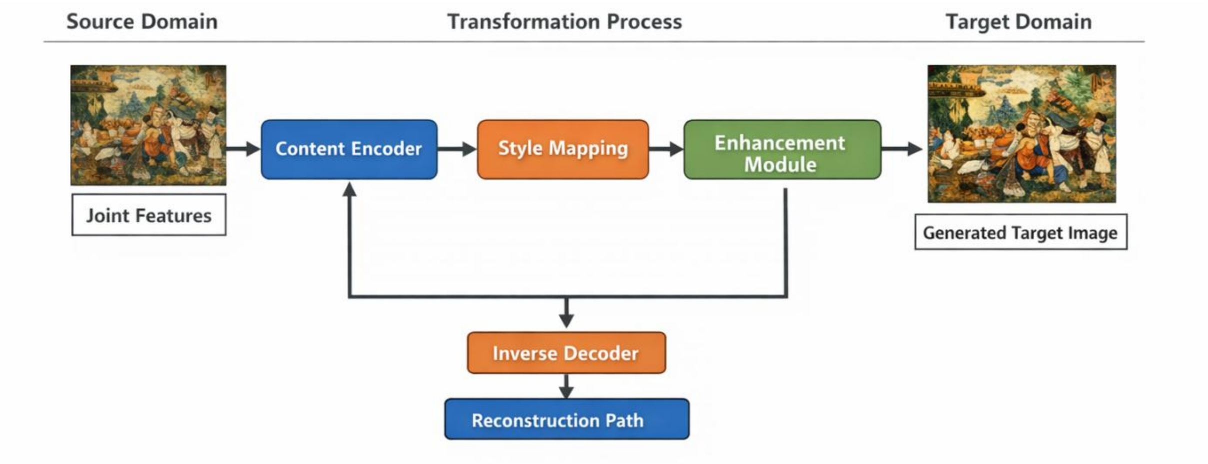

Figure 2: Enhanced model structure for style transfer of Intangible cultural heritage art

The output of the target domain after the source domain sample passes through the encoder and style mapping module can be expressed as Equation (14).

$$\hat{y}_t = G_{s\to t}(x_s, z_i) = R(T(E(x_s), z_i)) + x_s \tag{14}$$

Here, x_s represents the source domain intangible cultural heritage image, $E(\cdot)$ represents the content encoder, $T(\cdot)$ represents the style mapping module, $R(\cdot)$ represents the enhancement compensation branch, and $\hat{y}_t$ represents the target domain output after transfer enhancement. The function of this formula is to put style transfer and image enhancement into the same generation process, so that texture restoration does not rely on post-inpainting. The adversarial loss between the generator and the discriminator is defined using Equation (15).

$$L_{adv}^{t} = \mathbb{E}_{x_t\sim p_t}[\log D_t\,(x_t)] + \mathbb{E}_{x_s\sim p_s}[\log(1 - D_t(G_{s\to t}(x_s, z_i)))] \tag{15}$$

Where D_t represents the target domain discriminator, p_t and p_s represent the distribution of target domain and source domain samples respectively, and $G_{s\to t}(x_s, z_i)$ represents the generated samples. The function of this equation is to push the generated results to be close to the true target image in the style domain distribution, while maintaining the discriminative pressure of the generation process. In order to ensure that the content structure after cross-domain mapping does not drift significantly, this paper adopts the cyclic consistency constraint, whose form is shown in Equation (16).

$$\mathrm{L}_{\mathrm{cyc}} = \mathbb{E}_{\mathrm{x_s}\sim\mathrm{p_s}}[\|\mathrm{G}_{\mathrm{t}\to\mathrm{s}}(\mathrm{G}_{\mathrm{s}\to\mathrm{t}}(\mathrm{x_s},\mathrm{z_i})) - \mathrm{x_s}\|_1] + \mathbb{E}_{\mathrm{x_t}\sim\mathrm{p_t}}[\|\mathrm{G}_{\mathrm{s}\to\mathrm{t}}(\mathrm{G}_{\mathrm{t}\to\mathrm{s}}(\mathrm{x_t},\mathrm{z_i})) - \mathrm{x_t}\|_1] \tag{16}$$

Here, $\mathrm{G}_{\mathrm{t}\to\mathrm{s}}$ denotes the inverse generator and $\|\cdot\|_1$ denotes the reconstruction error. The function of this formula is to constrain the sample to return to a state close to the original input after the bidirectional mapping. Considering that the boundary details of the intangible cultural heritage image and the comprehensive color layer have strong style identification, this paper further introduces the structure-color joint constraint, whose form is shown in Equation (17).

$$\mathrm{L}_{\mathrm{sc}} = \|\mathrm{H}(\hat{\mathrm{y}}_\mathrm{t}) - \mathrm{H}(\mathrm{x_s})\|_1 + \|\mathrm{C}(\hat{\mathrm{y}}_\mathrm{t}) - \mathrm{C}(\mathrm{x_t})\|_1 \tag{17}$$

Here, $\mathrm{H}(\cdot)$ represents the structural feature extraction operator based on gradient direction, and $\mathrm{C}(\cdot)$ represents the color semantic encoding operator. The constraint compares the gradient direction feature and the color semantic vector simultaneously, so that the transfer result is consistent in edge sharpness and color level. The final training objective is shown in Equation (18).

$$\mathrm{L}_{\mathrm{total}} = \mathrm{L}_{\mathrm{adv}}^{\mathrm{t}} + \mathrm{L}_{\mathrm{adv}}^{\mathrm{s}} + \lambda_2\mathrm{L}_{\mathrm{cyc}} + \lambda_3\mathrm{L}_{\mathrm{sc}} \tag{18}$$

Here, $\mathrm{L}_{\mathrm{adv}}^{\mathrm{s}}$ represents the adversarial loss in the direction of the source domain, and λ_2 and λ_3 represent the weights of the cycle term and the joint constraint term, respectively. Through this design, model training no longer focuses on visual fidelity, but unifies structure preservation, style mapping, and enhancement restoration in intangible cultural heritage images into the same optimization goal.

In summary, the enhanced model of generative adversarial transfer constructed in this section has organized the joint feature input, bidirectional style mapping, enhanced compensation branch and multiple loss constraints into a complete training framework. The model formed in this way can not only complete the cross-style transfer of intangible cultural heritage art images, but also maintain the pattern structure, color level and local definition synchronously in the generation process.

4 Experimental results and analysis

4.1 Experimental environment and data set partition scheme

The experimental model is implemented based on Python 3.11 and PyTorch 2.1 framework, and the training and test deployment are completed in Ubuntu 22.04 environment. The hardware platform is configured with an AMD Ryzen 9 7950X processor, 64 GB memory, and an NVIDIA RTX 4090 graphics card with 24 GB of video memory capacity. The proposed configuration can support the stable training of intangible cultural heritage art images under bidirectional generation, inverse reconstruction and joint feature constraints. The total number of data sets is 5240, among which the samples of paper-cut, embroidery, wood-block New Year paintings and lacquer art patterns maintain an approximately balanced distribution. In the experiment phase, a stratified sampling strategy was used to form 4160 training samples and 1080 testing samples. In the training process, the validation monitoring is completed by setting aside in the batch, so as to complete the model adjustment and convergence judgment while maintaining the consistency of the summary data caliber. All input images are unified to 256×256 resolution, and the category proportion and source proportion are kept synchronously in the batch organization process to ensure that the style transfer enhancement results are comparable in the evaluation of different intangible cultural heritage categories. In order to ensure the reproducibility of the experiment, the random seed is set to 42, the data is loaded in parallel, the mixed precision calculation is enabled in the training process, and the generation loss, cycle consistency loss and joint constraint loss are recorded. The platform outputs PSNR, SSIM and LPIPS indicators in the test phase, which are used for performance comparison and result analysis in the subsequent chapters, and saves the experiment log for future reference.

4.2 Configuration of model training parameters and loss terms

In the model training stage, a uniform parameter configuration is used to ensure that the intangible cultural heritage art images maintain a stable convergence rhythm in the process of style transfer and enhancement. The training framework is based on Adam optimizer, the initial learning rate is set to 0.0002, the batch size is set to 4, and the number of training rounds is set to 180. The input image is unified to 256×256 resolution, and the generator encoder takes 64 channels as the initial width, and gradually expands the feature capacity during the downsampling process. The discriminator adopts a patk-level discriminative structure with ReLU activation at the generator side and Leaky ReLU activation at the discriminator side to maintain the stability of gradient propagation and local discrimination. The core training parameters are shown in Table 3.

Table 3: Model core training parameters

Parameter Item	Configuration Value
Optimizer	Adam
Initial Learning Rate	0.0002
Batch Size	4
Number of Training Epochs	180
Input Resolution	256 × 256
Initial Number of Channels	64
Discriminator Type	PatchGAN
Activation Function	ReLU / Leaky ReLU

In terms of loss configuration, adversarial loss is used to push the generated results closer to the target style domain distribution, cycle consistency loss is used to constrain the content regression after bidirectional mapping, and structure-color joint constraint is used to synchronously maintain the pattern boundary, comprehensive color hierarchy and local structure relationship. At the beginning of training, the learning rate is kept fixed to speed up the initial matching between the generator and the discriminator, and then the linear attenuation strategy is used to weaken the parameter update oscillation. The enhanced compensation branch only assumes the role of auxiliary inpainting, so its weight is controlled at a low level to avoid the local texture response being too strong and weakening the overall style expression. In the training process, a large proportion of random inactivation is not added, but instance normalization and the rhythm of the discriminator update are used to maintain the model stability. The parameter configuration formed in this way makes the model maintain a better balance between style transfer intensity, structure clarity and color stability, and also provides a unified training basis for subsequent performance evaluation and comparative analysis.

4.3 Evaluation of style transfer quality and image enhancement effect

The evaluation of style transfer quality and image enhancement effect depends not only on visual observation, but also on structure preservation and image quality indicators to judge the credibility of the generated results. PSNR is used to measure the reconstruction fidelity between the enhanced result and the reference image, SSIM is used to measure the preservation of structural hierarchy and pattern boundaries, and style consistency score is used to describe the closeness between the generated image and the target style domain in terms of comprehensive color layers and texture organization. Fig. 3 shows the variation trend of cycle consistency loss and joint structure-color constraint during the training process.

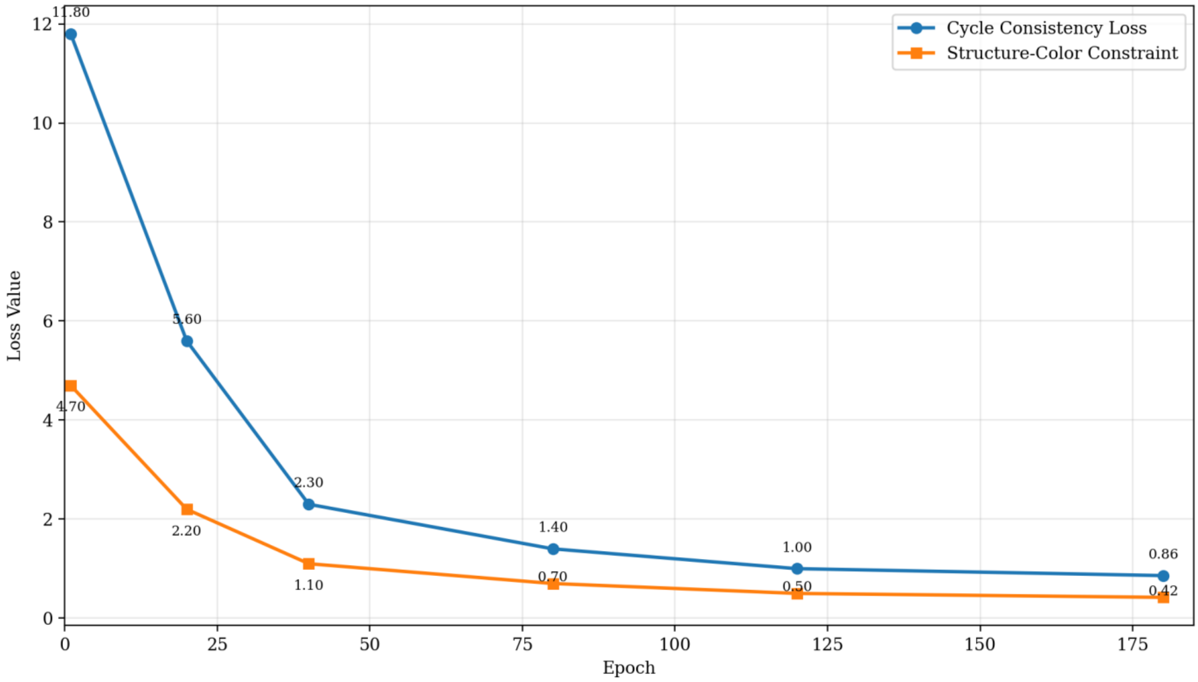

Figure 3: Trend of cycle consistency loss and structure-color joint constraints during training

Fig. 3 shows that the cycle consistency loss decreases rapidly in the first 40 rounds, from 11.8 to 2.3, and then enters the stable convergence interval, and stabilizes around 0.86 in the 180th round. The structure-color joint constraint gradually decreases from 4.7 to 0.42, which indicates that the boundary preservation and comprehensive color layer restoration tend to be stable synchronously after the bidirectional mapping is stable. The results show that the generator does not weaken the main contour and core pattern structure in the intangible cultural heritage image while completing the style transfer, but gradually establishes a coordinated expression oriented to boundary details and color levels. The comprehensive quality index changes of the four types of intangible cultural heritage samples in the testing phase are shown in Fig. 4.

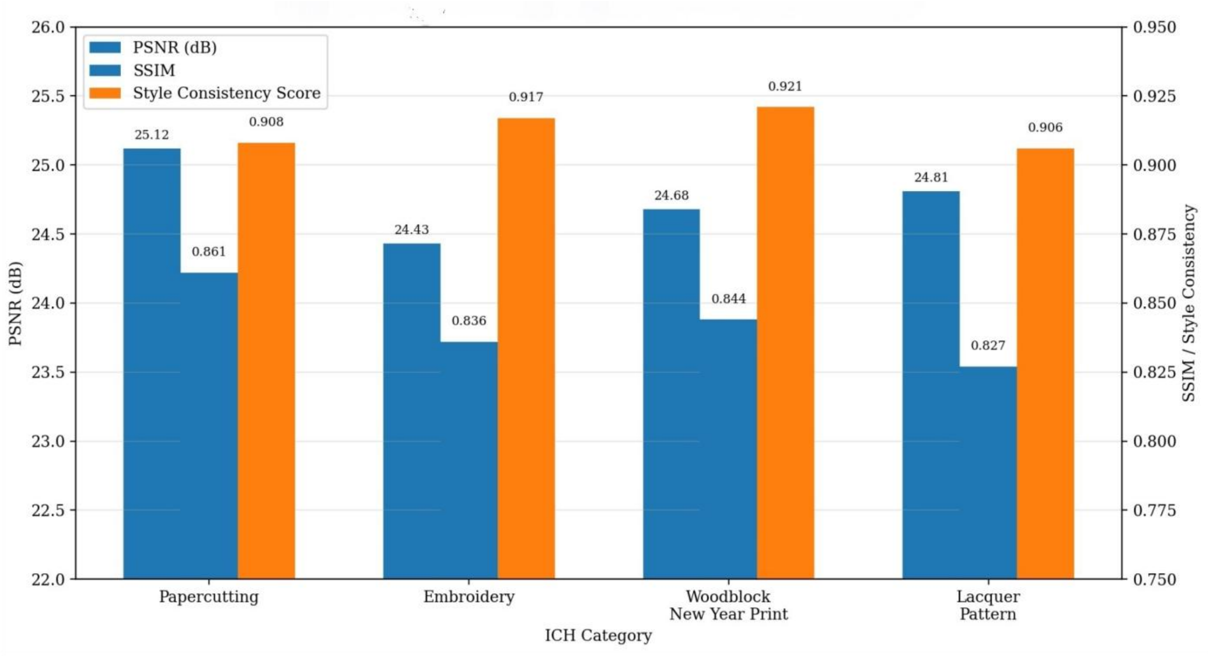

Figure 4: Performance of comprehensive quality indicators for four categories of intangible cultural heritage samples

As can be seen from Fig. 4, the average PSNR of the model on all test samples reaches 24.76 dB, the SSIM reaches 0.842, and the style consistency score reaches 0.913. When observing by type, the structure of the paper-cut sample is the most stable, and the boundary clarity is improved more obviously. Embroidery samples perform well in fine texture recovery. The New Year wood-block paintings maintained high color coordination in the comprehensive color layer transfer. The lacquer pattern samples show a strong enhancement effect in the suppression of highlight areas and the restoration of surface layers. The above results show that the proposed model can not only complete cross-style mapping, but also maintain the balance between complex patterns, color levels, and local details during the enhancement process. Therefore, the generated images have good performance in visual coherence and structural integrity. From the overall evaluation results, the style transfer quality and image enhancement effect achieve synchronous convergence in the same model, and the model output can meet the requirements of style expression and quality restoration at the same time, which provides a stable quantitative basis and comparison basis for subsequent comparison experiments and category difference analysis. And continue to support the follow-up discussion and comparative analysis.

4.4 Analysis of comparative experiments and ablation experiments

In order to verify the effectiveness of the proposed model in the intangible cultural heritage art style transfer enhancement task, this section carries out analysis from two levels of horizontal comparison and internal ablation. CycleGAN-Base and StyleUNet are selected as the reference models in the comparison experiment. The former represents the basic bidirectional transfer framework, and the latter represents the unidirectional generation structure with texture enhancement. PSNR, SSIM and style consistency score are still used to observe the comprehensive differences of different methods in three aspects of structure preservation, image enhancement and style expression. The overall performance of each model on the test set is shown in Fig. 5.

Figure 5: Comparison of comprehensive metrics of different models on the test set

It can be seen from Fig. 5 that the proposed model is superior to the reference method in three indicators, and the PSNR reaches 24.76 dB, which is higher than 23.12 dB of CycleGAN-Base and 23.68 dB of StyleUNet. SSIM reaches 0.842, which is higher than 0.816 and 0.829 of the two comparison methods. The style consistency score reached 0.913, which also remained the highest level. This indicates that the joint feature expression and enhanced compensation branch not only improve the reconstruction quality, but also enhance the preservation ability of pattern boundaries and comprehensive color layers in the style transfer task.

In order to further analyze the contribution of each component module, this paper removed the structure guidance, color semantic constraint and enhanced compensation branch respectively based on the full model, and the ablation results are shown in Table 4.

Table 4: Results of ablation experiments

Model Variant	PSNR (dB)	SSIM	Style Consistency Score
Full Model	24.76	0.842	0.913
Without Structural Guidance	24.08	0.821	0.901
Without Color Semantic Constraints	24.21	0.833	0.887
Without Enhancement Compensation Branch	23.95	0.826	0.898

Table 4 shows that after removing the structure guidance, SSIM decreases from 0.842 to 0.821, indicating that the texture boundary and contour continuity are most obviously affected. After removing the color semantic constraints, the score of style consistency decreases to 0.887, indicating that the matching degree between the comprehensive color layer and the dominant color relationship is significantly weakened. After removing the enhancement compensation branch, the PSNR decreases to 23.95 dB, indicating that the ability of local detail recovery and weak texture enhancement is inhibited. The complete model maintains the highest values on the three indicators, indicating that the modules are not simply superimposed, but form a synergy under a unified optimization objective. This result also shows that the transfer enhancement of intangible cultural heritage images cannot rely on a single generation path, but requires structure, color and detail recovery mechanisms to participate in the training process, which also provides a direct basis for the subsequent comparison and discussion of category differences. And make the follow-up discussion and comparative analysis interpretable.

4.5 Comparison of data performance under different categories of intangible cultural heritage art

Different categories of intangible cultural heritage art have obvious differences in pattern organization, color structure and boundary clarity, so the generation performance of the same model on each category also shows different characteristics. In order to further observe the category adaptation ability of the model, this paper respectively statistics the PSNR, SSIM and style consistency scores of four types of samples on the test set, such as paper-cut, embroidery, wood-block New Year paintings and lacquer art patterns, and the results are shown in Table 5.

Table 5: Comparison of test results under different intangible cultural heritage art categories

Intangible Cultural Heritage Category	PSNR (dB)	SSIM	Style Consistency Score
Paper-cutting	25.12	0.861	0.908
Embroidery	24.43	0.836	0.917
Woodblock New Year Painting	24.68	0.844	0.921
Lacquer Pattern	24.81	0.827	0.906

It can be seen from Table 5 that the paper-cut samples achieve the highest values in both indicators of PSNR and SSIM, indicating that the model has strong structure preservation ability when dealing with high contrast contours and clear boundaries. The style consistency score of the embroidery sample is the highest, indicating that the fine texture and repeated trace can complete the style transfer well under the action of joint feature expression and enhanced compensation branch. The New Year wood-block painting samples show stable performance in comprehensive color layer transfer, which indicates that the model has good coordination ability for images with large area block color and local line drawing. The PSNR of lacquer pattern samples is slightly lower, but the structure and style scores remain in the acceptable range, indicating that the high light reflection and surface layered texture put forward higher requirements for the generator. This shows that although the unified model can cover multiple types of intangible cultural heritage images, different visual structures still form recognizable performance levels in the generated results, and this hierarchical difference is directly related to boundary saliency, texture density and surface reflection characteristics. And it directly affects the stability of style mapping and enhanced recovery. It can be seen that the category comparison not only reflects the difference in model accuracy, but also further reveals the corresponding relationship between pattern structure, color hierarchy and transfer intensity. This also provides a more stable support for subsequent discussion and category analysis.

4.6 Discussion

The enhanced generative adversarial transfer model constructed in this paper is oriented to the unpaired style transfer scene of intangible cultural heritage art images. It completes training and testing on a dataset consisting of 5240 samples, and shows a relatively stable adaptation ability in four types of objects: paper-cut, embroidery, New Year wood-block paintings, and lacquer art patterns. The test results show that the average PSNR of the model reaches 24.76 dB, the SSIM reaches 0.842, and the style consistency score reaches 0.913, which indicates that the joint feature expression, enhanced compensation branch and bidirectional cyclic constraints form an effective cooperation. Compared with basic CycleGAN-Base and StyleUNet, the proposed model presents more balanced output features in three aspects: structure preservation, color coordination, and local detail recovery. The boundary definition of the paper-cut category is improved most obviously, the embroidery category maintains a high style consistency in the transfer of fine textures, the wood-block New Year pictures are stable in the comprehensive color layer mapping, and the laccase patterns show a strong ability to recover the highlighted areas and surface layered textures. The above results show that style transfer enhancement for intangible cultural heritage images cannot only rely on a single visual generation path, but need to unify the texture structure, color semantics and enhancement recovery mechanism into the same training objective. This unified modeling method makes the model still maintain relatively stable transfer intensity, structural fidelity and enhancement effect among different process categories, which indicates that a relatively coordinated computing relationship has been formed among joint feature representation, style mapping and detail restoration.

5 Conclusion

Focusing on the task of style transfer and enhancement of intangible cultural heritage art images, this paper constructs a computational implementation framework based on generative adversarial networks. At the data level, a multi-category intangible cultural heritage image set containing 5240 samples was established, and the sample organization, denoising correction and style label preprocessing were completed. At the method level, a joint feature expression mechanism and a bidirectional transfer enhancement model are designed, so that texture structure, color semantics and detail restoration can be coordinately-optimized under a unified objective. Experimental results show that the average PSNR of the model on the test set reaches 24.76 dB, the SSIM reaches 0.842, and the style consistency score reaches 0.913. It shows that the method has a good comprehensive performance in maintaining the texture boundary, comprehensive color hierarchy and local clarity. The limitations are mainly reflected in three aspects. At the sample level, although the existing data cover four types of objects, such as paper-cut, embroidery, wood-block New Year paintings, and lacquer art patterns, the proportion of extreme style samples and cross-material samples is still limited, which leads to a decrease in the stability of the transfer results when the model faces scenes with highly abstract styles and strong material contrasts. At the generation level, high-reflective lacework patterns, extremely dense embroidery areas, and locally overlapping patterns are more prone to detail compression, color offset, and insufficient boundary recovery in the mapping process, which indicates that there is still room for improvement in the control ability of the current structure-color joint constraint on complex local areas. At the deployment level, bidirectional generation and multi-branch constraints jointly increase the reasoning overhead, so that the running efficiency of the model is still limited in real-time applications and lightweight terminals. The follow-up research can be carried out in three ways: continue to expand the extreme style and cross-material intangible heritage samples to enhance the model's ability to cover complex style domains; A more fine-grained region attention mechanism and an adaptive constraint allocation strategy are introduced to improve the recovery accuracy of high reflection, dense texture and overlapping boundary regions. The lightweight generation structure and reasoning compression method are combined to promote the deployment and application of the model in digital display, repair generation and real-time interaction scenarios. On the whole, the computational link formed in this paper has connected the structure preservation, style mapping and enhancement restoration of intangible cultural heritage images into a unified technical process that can be trained. It also provides an extensible reference for generative modeling in similar cultural image tasks, and reserves sufficient space for subsequent engineering deployment and method expansion.

Funding

"Research on the Integration Path of Teaching Mode of Intangible Cultural Heritage and Illustration Design Course Driven by Artificial Intelligence" (SGH24Q394), 2024 annual project of the "14th Five-Year Plan" Education Science Plan of Shaanxi Province; Xi 'an Mingde Institute of Technology University-level Education and Teaching Reform Research Project "AIGC Driven by 'Multi-dimensional Collaboration' Innovation Practice Integration of Packaging Design Course Teaching Reform and Research (JG2024ZD09); Xi 'an Mingde Institute of Technology Mingde Cultural Research Fund Project, No. 2025MDC01[Z]

References

[1] Han X, Wu Y, Wan R. A method for style transfer from artistic images based on depth extraction generative adversarial network[J]. Applied Sciences, 2023, 13(2): 867.

[2] Batziou E, Ioannidis K, Patras I, et al. Artistic neural style transfer using CycleGAN and FABEMD by adaptive information selection[J]. Pattern Recognition Letters, 2023, 165: 55-62.

[3] Zheng X, Yang X, Zhao Q, et al. CFA-GAN: Cross fusion attention and frequency loss for image style transfer[J]. Displays, 2024, 81: 102588.

[4] Tahir R, Cheng K, Memon B A, et al. A diverse domain generative adversarial network for style transfer on face photographs[J]. 2022.

[5] Christophe S, Mermet S, Laurent M, et al. Neural map style transfer exploration with GANs[J]. International Journal of Cartography, 2022, 8(1): 18-36.

[6] Khowaja S A, Almakdi S, Memon M A, et al. Extending user control for image stylization using hierarchical style transfer networks[J]. Heliyon, 2024, 10(5).

[7] Feng J, Zhang G, Li X, et al. A compositional transformer based autoencoder for image style transfer[J]. Electronics, 2023, 12(5): 1184.

[8] Wei Y. Artistic image style transfer based on CycleGAN network model[J]. International Journal of Image and Graphics, 2024, 24(04): 2450049.

[9] Wang L, Wang L, Chen S. ESA-CycleGAN: Edge feature and self-attention based cycle-consistent generative adversarial network for style transfer[J]. IET Image Processing, 2022, 16(1): 176-190.

[10] Bao C, Li Y, Lu E. Design and implementation of dongba character font style transfer model based on AFGAN[J]. Sensors, 2024, 24(11): 3424.

[11] Yang S, Zhou D, Cao J, et al. Rethinking low-light enhancement via transformer-GAN[J]. IEEE Signal Processing Letters, 2022, 29: 1082-1086.

[12] Yu W, Zhao L, Zhong T. Unsupervised low-light image enhancement based on generative adversarial network[J]. Entropy, 2023, 25(6): 932.

[13] Wang L, Zhao L, Zhong T, et al. Low-light image enhancement using generative adversarial networks[J]. Scientific Reports, 2024, 14(1): 18489.

[14] Jin H, Wang Q, Su H, et al. Event-guided low light image enhancement via a dual branch GAN[J]. Journal of Visual Communication and Image Representation, 2023, 95: 103887.

[15] Wu J, Liu X, Lu Q, et al. FW-GAN: Underwater image enhancement using generative adversarial network with multi-scale fusion[J]. Signal Processing: Image Communication, 2022, 109: 116855.

[16] Jia Y, Wang Z, Zhao L. An unsupervised underwater image enhancement method based on generative adversarial networks with edge extraction[J]. Frontiers in Marine Science, 2024, 11: 1471014.

[17] Karimi N, Mishra M, Lourenço P B. Deep learning-based automated tile defect detection system for Portuguese cultural heritage buildings[J]. Journal of Cultural Heritage, 2024, 68: 86-98.

[18] Stoean R, Bacanin N, Stoean C, et al. Bridging the past and present: AI-driven 3D restoration of degraded artefacts for museum digital display[J]. Journal of Cultural Heritage, 2024, 69: 18-26.

[19] Altaweel M, Khelifi A. Using generative AI for reconstructing cultural artifacts: examples using Roman coins[J]. Journal of Computer Applications in Archaeology, 2024, 7(1): 301-315.

[20] Zhu J, Fang B, Chen T, et al. Face repairing based on transfer learning method with fewer training samples: application to a Terracotta Warrior with facial cracks and a Buddha with a broken nose[J]. Heritage Science, 2024, 12(1): 186.