Journal of Design Intelligence
https://www.ffspub.com/index.php/jdi
ISSN 3104-9001 (print)
E-mail: jdi.office@ffspub.com

Article

A Fast Multi-Scale Textile Pattern Generation Method Combining Layered Loss and Convolutional Attention

Mengyuan Li, Guodong Xu, Yu Chen[†]
School of Textiles and Fashion, Shanghai University of Engineering Science, Shanghai 201620, China
[†]*E-mail: ychen0918@sues.edu.cn*

Received: April 30, 2026 / Revised: May 29, 2026/ Accepted: June 11, 2026 / Published online: June 18, 2026

Abstract: Existing image generation models frequently encounter challenges such as line discontinuity, structural loss, and poor style adaptability when dealing with pattern designs that possess complex geometric structures. To address these issues, this paper proposes a novel deep neural network-based method for the digital design of traditional patterns, constructing a three-stage human–machine collaborative workflow of "structure generation—sketch translation—style transfer." By injecting adaptive noise and incorporating a multi-scale discrimination mechanism into the StyleGAN generation algorithm, and by integrating edge consistency and structure-aware losses, the continuity and completeness of line drawing generation are significantly improved. A conditional generative adversarial network (Pix2PixHD) cross-domain mapping model combined with a self-attention mechanism is employed to accurately achieve the automatic conversion of irregular hand-drawn sketches into standardized line drawings. Furthermore, a neural style transfer strategy based on multi-scale feature disentanglement is designed, jointly utilizing the Gram matrix and the Wasserstein distance, and supplemented by a convolutional attention module, to realize high-fidelity fusion between traditional structures and modern styles. Experiments verify that our method delivers superior visual quality and structural fidelity compared with state-of-the-art models. It realizes cloud motif expansion, feature extraction, sketch-to-standard line drawing translation and pattern style transfer, and offers theoretical and practical references for aesthetic and intelligent computer-aided design.

Keywords: Ethnic Clothing Patterns; Intelligent Pattern Design; Generative Adversarial Networks; Style Transfer
https://doi.org/10.64509/jdi.12.108

1 Introduction

As one of the most representative visual symbols of human civilization, ornamental patterns embody the aesthetic orientations, religious beliefs, and cultural identities of different ethnic groups. However, under the impact of modernization, many ancient patterns face the risk of disappearance due to the physical deterioration of artifacts or the loss of traditional craftsmanship. Conventional pattern design relies heavily on the personal experience and manual drawing of designers, following a linear workflow of "conception—line drawing—coloring—stylization," suffering from inefficiency and difficulties in ensuring standardization, and failing to meet the demands of large-scale customization in modern industry [1]. In recent years, generative models represented by deep neural networks, such as convolutional neural networks (CNNs) and generative adversarial networks (GANs), have introduced entirely new design paradigms to the field, offering fresh theoretical references and practical pathways for the deep integration of artificial intelligence and design practice, and have thus attracted increasing attention.

Generative adversarial networks (GANs) greatly enhance the realism of generated samples through the dynamic game between a generator and a discriminator, becoming classic generative models. Radford *et al.* [2] further proposed the deep convolutional generative adversarial network (DCGAN), which introduces convolutional operators into adversarial training to strengthen the model's feature extraction capability. Its advantage lies in the ability to learn unsupervised feature regularities; however, it exhibits pronounced training instability and mode collapse when processing high-resolution images. The StyleGAN series introduced by Karras *et al.* [3] incorporates a style-disentangled mapping network and an adaptive instance normalization (ADAIN) mechanism,

[†] Corresponding Author: Yu Chen
[*] Academic Editor: Yan Hong

achieving milestone results in fine-grained image control. Nevertheless, when dealing with line drawing patterns such as cloud motifs that demand extremely high line continuity, StyleGAN's inherent fixed noise injection lacks dynamic adjustment capability, often leading to line breakage and topological distortion of geometric structures.

In the translation from sketches to digital forms, Isola *et al*. [4] proposed the Pix2Pix framework, which uses a conditional generative adversarial network (CGAN) to achieve end-to-end image mapping, demonstrating significant advantages in preserving content consistency. However, the model's heavy reliance on paired datasets limits its generalization to the domain of ethnic patterns. To address this, Wang *et al*. [5] developed Pix2PixHD, which elevates the generation resolution to a higher level through a multi-scale generator architecture. Nonetheless, when confronted with irregular hand-drawn sketches, the locality of conventional convolutional receptive fields makes it difficult for existing models to capture the global long-range geometric dependencies of the patterns, resulting in generated line drawings that struggle to meet the structural regularity required by industrial design standards.

Neural style transfer (NST) has opened new avenues for combining ornamental patterns with textiles. Gatys *et al*. [6] successfully achieved the separation and synthesis of content and style using pre-trained convolutional neural networks and the Gram matrix. Building on this, Sun and Chen *et al*. [7] proposed a fast generation scheme combining Markov random fields (MRFs) with the Gram matrix, significantly reducing computational cost. The advantage of this method lies in its rapid iteration of design proposals; however, its core contradiction lies in an overemphasis on the statistical properties of style textures while neglecting the semantic structural integrity specific to ethnic patterns. Particularly when dealing with highly regular geometric structures carrying specific cultural connotations, it is prone to visual artifacts and logical discontinuities, failing to balance artistic appeal with regularity.

Compared with StyleGAN, we propose a data preprocessing method combining edge detection and structural optimization to construct a high-quality cloud motif line drawing dataset. Meanwhile, we optimize the network architecture of StyleGAN and embed the adaptive noise injection module as well as the multi-scale discriminative mechanism. These improvements enhance the line continuity, stability and detail fidelity of generated cloud motif images. In view of the irregularities and low efficiency existing in digital sketch conversion when using Pix2PixHD, this paper presents a sketch-to-line drawing translation method based on Conditional GAN (Pix2PixHD). It realizes automatic mapping from hand-drawn sketches to standardized cloud motif line drawings. By introducing the self-attention mechanism and perceptual loss, the model is guided to focus on main structural skeletons and fine-grained features, which greatly improves the structural standardization and line clarity of converted results. Aiming at the limitations of existing neural style transfer techniques, we put forward an improved style transfer method that integrates style feature extraction based on the VGG19 network, Gram matrix and Wasserstein distance. The Convolutional Block Attention Module (CBAM) is further introduced to optimize local details. Combined with multi-scale optimization and Total Variation (TV) loss, the proposed method achieves better visual naturalness and higher local structural fidelity for style transfer outputs.

In summary, although deep neural networks have demonstrated immense potential in image generation, existing research still exhibits shortcomings such as line breakage, structural incompleteness, and poor style adaptability when applied to ethnic costume patterns with complex topologies and specific cultural semantics. This paper proposes a novel deep neural network-based method for the digital design of traditional patterns, constructing a three-stage workflow of "structure generation—sketch translation—style transfer" and thereby realizing the conversion from hand-drawn cloud motif sketches to digital line drawings and style fusion. On this basis, we have developed an end-to-end intelligent clothing pattern design platform that implements a closed-loop design flow from creative conception to finished product export. This study not only provides efficient technical support for the digital preservation and redesign of ethnic patterns but also offers a new reference for constructing human–machine collaborative computational aesthetics design.

2 Method

This paper constructs a three-stage design framework based on deep neural networks. The framework simulates a design logic that progresses from the abstract to the concrete and can address the problems of low efficiency and poor structural fidelity inherent in traditional design workflows. It mainly comprises three aspects: structure generation, sketch translation, and style transfer. The framework diagram of the three stages is illustrated in Figure 1.

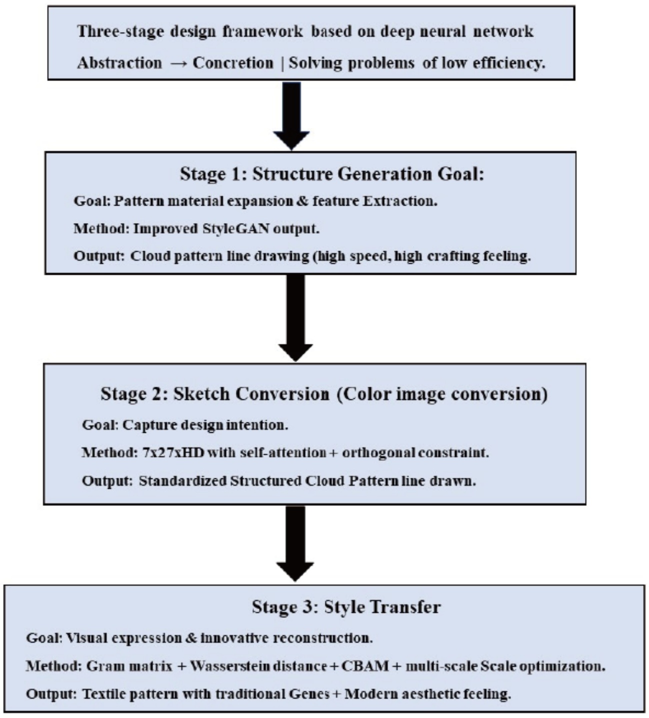


Figure 1: Three-stage structural framework diagram.

In Figure 1, Stage 1 is the structure generation stage, which utilizes an improved StyleGAN network to generate cloud motif line drawings, thereby enabling the augmentation of pattern materials and feature extraction.

Stage 2 is the sketch translation stage, where a conditional generative adversarial network (Pix2PixHD) equipped with a self-attention mechanism is introduced to achieve cross-domain mapping from hand-drawn sketches to standardized cloud motif line drawings.

Stage 3 in the figure is the style transfer stage. The style transfer employs a dual-layer reconstruction that fuses the Gram matrix and the Wasserstein distance, incorporates a convolutional block attention module (CBAM) and multi-scale optimization, and accomplishes the visual expression and innovative reconstruction of the pattern.

2.1 Cloud Motif Line Drawing Generation Algorithm based on Improved StyleGAN

To address the problem of line discontinuity in traditional generative models, this paper carries out a deep reconstruction of both the generator and the discriminator of StyleGAN.

1. Generator Adaptive Noise Injection: Unlike the conventional fixed noise injection, this paper introduces an adaptive noise injection strategy that can dynamically adjust the noise intensity according to the stage of the generation process, with its dynamic attenuation formula as shown in Equation 1:

$$\text{NoiseFactor}(\text{step}, \text{total}) = \alpha - (\alpha - \beta) \cdot \frac{\text{step}}{total_{steps}} \quad (1)$$

Here, α denotes the initial noise intensity (maximum noise factor), and β denotes the minimum noise intensity (the minimum value of the attenuated noise factor). step represents the current training resolution step (which increments gradually as the generator generates images step-by-step), and $total_{steps}$ denote the total number of steps for image generation. The noise is dynamically injected into the initial layer and each generation block of the generator. In the early training stage, a high noise intensity is adopted to improve the diversity of generated images; in the later training stage, the noise intensity is gradually reduced to enhance the stability of image details.

Multi-scale Discriminator Architecture: To enhance the model's comprehensive judgment of both local details and the overall structure, the discriminator is designed with a multi-scale auxiliary branch structure. The model performs down-sampling and feature extraction on the generated image at multiple scales, and the final output expression that synthesizes the main path and the multi-scale branches is given by Equation 2:

$$F_{\text{out}} = \frac{P_{\text{out}} + M_{\text{out}}}{2} \quad (2)$$

where P_{out} and M_{out} represent the outputs of the main discrimination path and the auxiliary branch, respectively.

2. Composite Loss Function Design: To address the issue of line breakage in cloud motifs, this paper innovatively introduces an edge consistency loss L_{edge} and a structure-aware loss L_{struct} on the basis of the Wasserstein-GP adversarial loss [8]. The edge consistency loss uses the Sobel operator to extract the gradient magnitudes of the generated image and the real image and computes their L_1 norm distance, as shown in Equation 3:

$$L_{edge} = E_{x,z}\left[\|\mathcal{E}(G(z)) - \mathcal{E}(x)\|_1\right] \quad (3)$$

The structure-aware loss utilizes the intermediate convolutional layers (relu1_2, relu2_2, relu3_3) of a pre-trained VGG19 network to extract features for constraining high-level semantics, as shown in Equation 4:

$$L_{struct} = \sum_{l \in L} E_{x,z}\left[\|\phi_l(G_\phi(z)) - \phi_l(x)\|_1\right] \quad (4)$$

In Equation 4, the total generator loss is the weighted sum of the adversarial loss and the two losses mentioned above.

The total generator loss is formulated as a weighted combination of adversarial loss and perceptual loss, which is defined as: $L'_G = L_G + \lambda L_P$, where λ denotes the weight parameter. The overall loss function is mainly composed of adversarial loss and perceptual loss. Adversarial Loss: The discriminator enhances its capability to identify real samples by maximizing the adversarial loss. In contrast, the generator minimizes this loss to make generated images approximate real ones, so as to fool the discriminator. The optimization objective of the discriminator is to maximize the probability of classifying real images as genuine and minimize the probability of misclassifying generated images as genuine. The discriminator loss is defined as follows:

$$L_D = \frac{1}{2}\left(E_{x,y}\left[(D(x,y) - 1)^2\right] + E_x\left[D(x, G(x))^2\right]\right) \quad (5)$$

In the formula, x stands for the input image and y represents the ground-truth target image. The first expectation term encourages the discriminator to produce an output close to 1 for real samples, while the second term drives its output to approach 0 for generated samples. The goal of the generator is to make the discriminator output results close to those of real images when processing generated images, to make the output of $(D(x, G(x)) - 1)$ approach 1. The generator loss is defined as follows:

$$L_G = E_x\left[(D(x, G(x)) - 1)^2\right] \quad (6)$$

Perceptual Loss: For perceptual loss, we employ the VGG network to extract feature representations of target images and generated images from multiple specific layers, and adopt the L_1 distance as the dissimilarity metric. Its formula is given below:

$$L'_G = \sum_{i \in I} E_{x,y}\left[\|\phi_i(y) - \phi_i(G(x))\|_1\right] \quad (7)$$

Here, I refers to the set of feature layers, and $\|\cdot\|_1$ represents the L_1 distance, namely the sum of absolute values of all feature elements. Through layer-wise calculation of features across different network layers, this loss term can measure the high-level semantic similarity between generated images and real images. Accordingly, the generator is able to produce results with better visual consistency with real images.

2.2 Translation of Irregular Sketches based on Pix2PixHD

To achieve efficient mapping from irregular hand-drawn sketches to standardized line drawings, this paper introduces two improvements based on the Pix2PixHD architecture.

1. Introduction of the self-attention mechanism: Traditional convolution operations are limited by local receptive fields, making it difficult to establish long-range dependencies among the complex cloud motif lines [9]. This paper embeds a self-attention module after the down-sampling block of the generator. By computing a similarity weight matrix of Query, Key, and Value, the module adaptively captures global contextual information, thereby ensuring the global structural coherence of the generated line drawing.
2. Fusion of perceptual loss: In addition to the traditional generator adversarial loss L_G, the algorithm introduces a multi-layer perceptual reconstruction loss L_P. By measuring the absolute error between the target image and the generated image in the VGG feature space, it guides the model to focus on high-frequency structural details. The resulting generator objective function is given by Equation 8:
$$L'_G = L_G + \lambda L_P \tag{8}$$
where λ is a weight parameter, i.e., the total generator loss is a weighted combination of the adversarial loss L_G and the perceptual loss L_P.

2.3 Style Transfer Algorithm for Cloud Motif Patterns Fusing Multiple Features

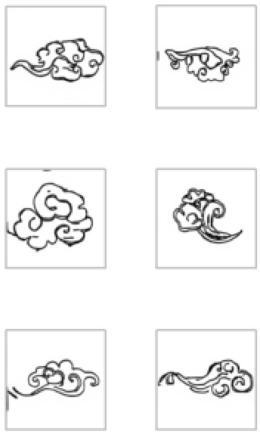
To seamlessly integrate traditional cloud motifs with modern textile pattern styles while avoiding artifacts, this paper constructs a multi-scale feature enhancement network.

1. Style loss with separation of low-level and high-level features: This paper employs the VGG19 network to extract style features. For shallow features (Layer1–3), the algorithm uses a feature-normalized Gram matrix to compute the low-level style loss, so as to accurately capture local textures and colors. The definition of the style loss function based on the Gram matrix is shown in Equation 6:
$$L_{style_g} = \frac{\sum_{i,j}\left(G_{ij}^{norm}(g) - G_{ij}^{norm}(s)\right)^2}{\sum_{i,j}\left|G_{ij}^{norm}(g) - G_{ij}^{norm}(s)\right| + \epsilon} \tag{9}$$
In Equation 9, $G_{ij}^{\text{norm}}(s)$ and $G_{ij}^{\text{norm}}(g)$ denote the feature-normalized Gram matrices of the style image and the generated image, respectively, and ϵ is a regularization term to avoid division by zero during computation.
For deep features (Layer4–5), the Gram matrix tends to lose spatial structure information. Therefore, this paper adopts the Wasserstein distance under a Gaussian distribution to measure the distribution discrepancy of deep semantics, thereby preserving complex structures, as the style loss L_{style_w} for deep features. Its calculation method is shown in Equation 10:
$$L_{style_w} = \|\mu_x - \mu_t\|^2 + \mathrm{Tr}\left(\Sigma_x + \Sigma_t - 2\left(\Sigma_t^{1/2}\Sigma_x\Sigma_t^{1/2}\right)^{1/2}\right) \tag{10}$$
In Equation 10, μ_x and μ_t denote the mean vectors of the features of the input image and the target style image, respectively, and Σ_x and Σ_t represent the covariance matrices of the features of the input image and the target style image, respectively.
2. Convolutional attention feature enhancement (CBAM) and TV loss: To suppress background interference, channel attention and spatial attention modules are cascaded inside the network, which adaptively compute feature weights, enabling the network to focus on typical texture response regions. Additionally, to suppress high-frequency noise during the generation process, the algorithm introduces a total variation (TV) loss, as shown in Equation 11:
$$L_{\text{TV}} = \sum_{i,j}\left((I_{i,j} - I_{i+1,j})^2 + (I_{i,j} - I_{i,j+1})^2\right) \tag{11}$$
The above Equation 11 incorporates a multi-scale optimization strategy with exponential moving average (EMA), where the network is progressively updated from low resolution to high resolution in a smooth manner, ensuring that the final image possesses excellent visual texture and structural integrity.

3 Experiments and Results Analysis

3.1 Dataset construction and experimental setup

Existing general-purpose datasets typically lack high-quality samples specifically for cloud motifs. This paper constructs a dedicated dataset of cloud motif line drawings and sketches. All experiments were implemented based on the PyTorch deep learning framework and conducted on a Windows 10.0 computer equipped with an NVIDIA RTX 4090 GPU (24 GB VRAM). The cuDNN acceleration library was enabled to speed up the training process, and the Adam optimizer was adopted for gradient updates during training. The dataset utilized in the experiments is the self-built cloud motif line drawing dataset. In this paper, the proposed generative model is pre-trained on the cloud motif dataset using a progressive training strategy: the image resolution is gradually increased from 32×32 to 512×512, with 500 training epochs performed at each resolution scale, and the learning rate is set to 0.0003.Partial examples from the cloud motif sketch dataset and the line drawing dataset are shown in Figure 2a and Figure 2b, respectively.




(a) Irregular sketches. (b) Line drawings.

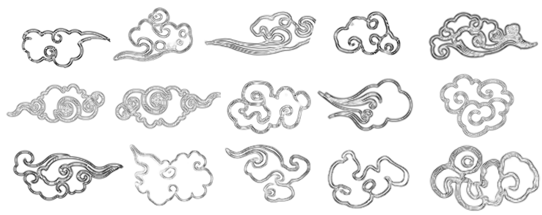
Figure 2: Display of partial cloud motif dataset: (a) irregular cloud motif sketch dataset; (b) cloud motif line drawing dataset.

Figure 2a Cloud motif sketch dataset: A deep learning model named Photo-Sketching was employed to simulate a hand-drawn style on the augmented cloud motifs. After manual screening, a dataset containing 2,300 paired sketches was constructed, primarily used for the training of Pix2PixHD. Figure 2b Cloud motif line drawing dataset: 1,200 classic cloud motif images were collected. The Canny edge detection algorithm combined with dilation and Gaussian smoothing operations was adopted to extract smooth and continuous contour lines, and the dataset was further expanded to 4,273 images through data augmentation.

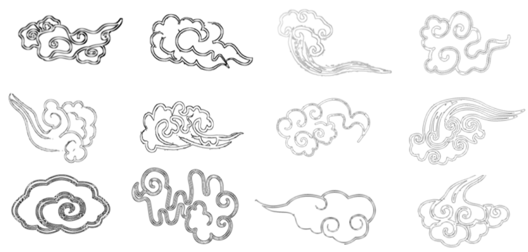
In the study, we adopted the deep learning model Photo-Sketching proposed by Li *et al.* [10] to generate simulated hand-drawn sketches. This model is built upon improved Conditional Generative Adversarial Networks (Conditional GAN) and trained on a dataset of 5,000 manually drawn sketches. It is capable of producing images whose styles are highly similar to authentic hand-drawn sketches. The original cloud motif dataset was augmented and then fed into the pre-trained Photo-Sketching model to generate corresponding sketch images. Afterwards, we performed manual screening on the generated sketches and eliminated samples that visually deviated from real hand-drawing styles, ensuring that all sketches in the final dataset achieve a high degree of hand-drawing simulation. In total, 2,300 sketch samples satisfying the visual criteria were retained as high-quality input data. The dataset was divided into a training set and a test set: 2,100 sketches were used for model training and optimization, while the remaining 200 sketches formed the test set to assess the model's generation performance and generalization ability. The chapter focuses on generating line drawings of traditional cloud motifs based on StyleGAN. The training data needs to fully cover the main stylistic characteristics of the patterns, among which shape and line information are essential for model learning. To satisfy the training requirements, we collected 1,200 representative cloud motif images from diverse sources, including online resources, art literature and academic publications. These images cover cloud motifs with varied styles from different historical periods. Despite the shape differences among cloud motifs of distinct historical backgrounds, their core stylistic features are highly consistent and have negligible influence on model generation. For this reason, we did not categorize the cloud motifs by morphological types. Instead, we selected images with typical stylistic features and extensive practical application.

3.2 Experimental Results of Cloud Motif Line Drawing Generation Using the Improved StyleGAN

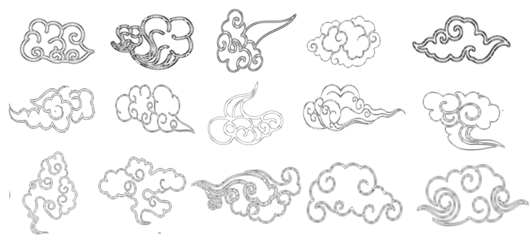
At a resolution of 512×512, the proposed algorithm was compared with WGAN-GP, Pro GAN, and the baseline StyleGAN. The cloud motif line drawings generated by different models are shown in Figure 3 below.



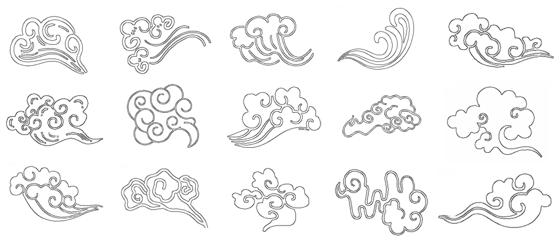
(a) WGAN-GP



(b) ProGAN



(c) StyleGAN



(d) Our

Figure 3: Samples generated by different models.

As shown in Figure 3a, WGAN-GP can reproduce the basic outline of the cloud motif pattern, but the lines are blurred at the branching points of the pattern. Figure 3b shows that Pro GAN, through progressive training, can generate the main contours of the pattern; however, in regions where the pattern is intricately interwoven, problems of line interruption and local structural loss still occur. In Figure 3c, the images generated by the baseline StyleGAN are constrained by static noise, making local regions prone to fractures. Figure 3d presents the images generated by the improved StyleGAN proposed in this paper. Through comparison, it can be concluded that the line drawings generated by the proposed algorithm exhibit natural trajectories, clear turning points, and maintain high structural integrity at branching locations.

In the comparison among different models, the proposed algorithm achieved the optimal performance metrics. The metric values of the generation effects obtained from different models in the experiment are shown in Table 1 below.

Table 1: Metric values of generation effects for different models.

Method	FID↓	SSIM↑
WGAN-GP	58.72	0.528
ProGAN	46.84	0.727
StyleGAN	35.65	0.756
Our	22.37	0.873

As can be seen from Table 1, after improving StyleGAN, the FID value of the cloud motif line drawings decreases to 22.37 (baseline StyleGAN: 35.65), and the SSIM value increases to 0.873. Ablation experiments further confirm that the edge consistency loss and the structure-aware loss effectively prevent mode collapse and enhance edge fidelity.

3.3 Experimental Results of Sketch-to-Line-Drawing Translation

This paper compares the improved Pix2PixHD model with Pix2Pix, Cycle GAN, and MUNIT in performance experiments. The sketch translation results of the experiment are shown in Figure 4.

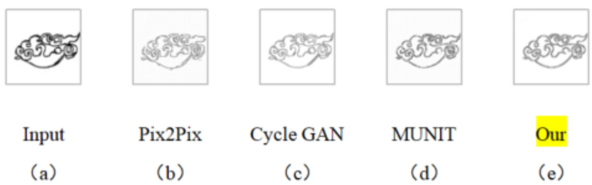


Figure 4: The sketch translation results of the experiment.

As can be seen from Figure 4, the images generated by Pix2Pix and MUNIT generally exhibit line breakage, while Cycle GAN shows obvious texture distortion. The proposed algorithm (Figure 4, Our (e)) can accurately reproduce the intent of the hand-drawn sketch when faced with complex interwoven structures, generating smooth and coherent lines; when generating simple patterns, the line structure is complete with no evident breakage, meeting visual expectations. The experimental results demonstrate that the method proposed in this paper offers significant advantages in improving line quality and structural accuracy, providing a new methodology and reference for the creation of cloud motif patterns.

This paper also conducted performance comparison experiments across different models, with the experimental results shown in Table 2.

Table 2: Performance comparison results of different models.

MODELS	FID↓	SSIM↑
Pix2Pix	40.34	0.887
Cycle GAN	47.73	0.839
MUNIT	36.4	0.865
Our	28.75	0.897

From the experimental comparison results in Table 2, it can be observed that the improved model (Our) achieves relatively outstanding performance. Specifically, the FID of the model reaches 28.75, outperforming Pix2Pix's 40.34 and MUNIT's 36.4. The SSIM of the model improves to 0.897. Ablation experiments show that removing the self-attention module causes the FID to rise to 33.15, and removing the perceptual loss leads to a significant degradation in structural consistency. To verify the effectiveness of the proposed model improvements, four groups of comparative experiments are conducted for ablation analysis. Table 1 summarizes the quantitative results of all schemes in terms of FID and SSIM.

Table 3: Quantitative Metrics Comparison in Ablation Experiments.

Scheme	FID↓	SSIM↑
Scheme a	28.75	0.897
Scheme b	33.15	0.843
Scheme c	35.84	0.812
Scheme d	37.56	0.786

(a) Scheme 1: Full model (equipped with the self-attention module and perceptual loss);
(b) Scheme 2: Model with the self-attention module removed;
(c) Scheme 3: Model with both the self-attention module and perceptual loss removed;
(d) Scheme 4: Original baseline Pix2PixHD model.

The experimental results indicate that our model (Our) performs excellently in both FID and SSIM metrics, significantly lower than Pix2Pix, Cycle GAN, and MUNIT, suggesting that in the sketch-to-line-drawing translation task, due to the introduction of the self-attention mechanism, the generated images are distribution ally closer to real images, yielding higher generation quality. That is, the proposed model demonstrates a clear advantage in processing complex patterns.

3.4 Experimental results of style transfer

In the style transfer experiments targeting textile pattern generation, traditional methods based solely on the Gram matrix suffer from severe texture confusion when handling structurally strong cloud motifs [11]. By integrating the Wasserstein distance and TV loss, the proposed method generates images that not only accurately inherit the color and texture of the target image (e.g., embroidery texture) but also globally preserve the rigorous geometric structure of the cloud motif. Noise tests show that after introducing TV loss, high-frequency noise in the image background is significantly suppressed, and smoothness is substantially improved. Figure 5 below shows the experimental results of style transfer.

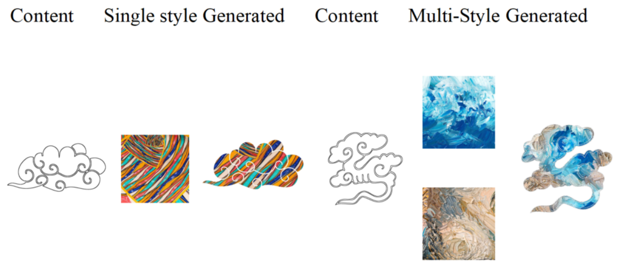


Figure 5: Composite rendering of style transfer for cloud motif patterns.

The experimental results of style transfer synthesis for cloud motif patterns shown in Figure 5 can verify the model's adaptability and stability when dealing with different style types. In single-style transfer, the model can effectively extract the key features of the input pattern and smoothly map them onto the cloud motif structure, maintaining both style consistency and the structural integrity of the pattern. In the process of multi-style transfer, the model can integrate the features of two input styles and fusing them into the cloud motif pattern. Moreover, during multi-style fusion, the proposed model produces no obvious visual conflicts; on the contrary, it demonstrates a natural transition within the cloud motif line structure, achieving a harmonious unity of traditional and modern styles and enhancing the comprehensive artistic expressiveness of the pattern.

4 Intelligent Design Platform and Application Transformation

4.1 Platform Architecture Design

In order to transform the algorithmic models proposed in this paper into implementable design tools, this paper designs and implements an intelligent clothing pattern design platform. The platform framework structure diagram is shown in Figure 6.

As can be seen from the layered architecture design diagram of the platform in Figure 6, the intelligent clothing pattern design platform comprises three layers in total, which are primarily the user layer, the application layer, and the technical layer:

1. Technical support layer: As the core of the platform, this layer integrates the optimized StyleGAN, Pix2PixHD, and the VGG19-based style transfer model proposed in this paper, primarily providing underlying computational capability and algorithm driving.
2. Application function layer: This layer mainly includes functional modules such as sketch input, automatic line drawing generation, interactive style reconstruction, and vectorized export.
3. User interaction layer: This layer is used to provide an intuitive visual interface, assisting designers in sketching, parameter adjustment, and real-time effect preview, thereby achieving a user-friendly human–computer interaction experience.

Through the above three-layer platform architecture, a high degree of system modularity and scalability can be effectively ensured.

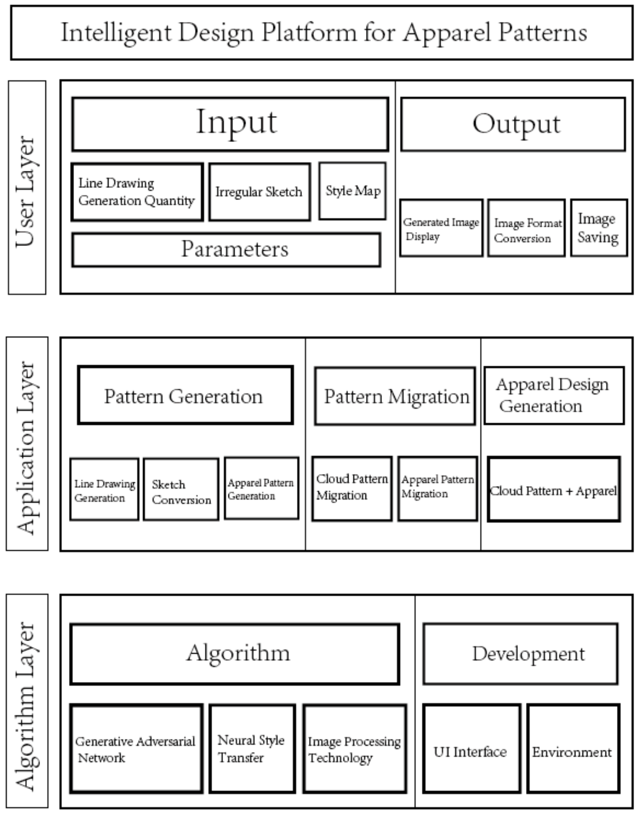


Figure 6: Architecture diagram of the intelligent clothing pattern design platform

4.2 Human–Machine Collaborative Intelligent Design

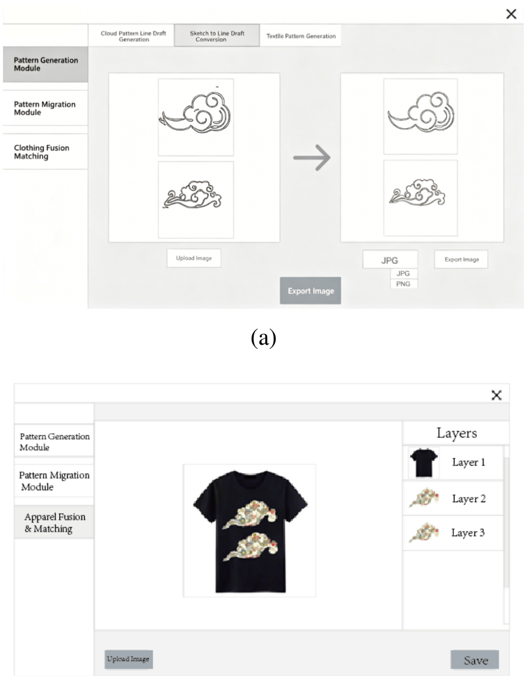
In clothing pattern design, traditional pattern design often relies on the experience of designers, which is inefficient and struggles to meet rapidly changing market demands [12]. This paper builds an intelligent clothing pattern design platform ranging from "irregular sketch input" to "stylized product output". A partial design interface of the platform is shown in Figure 7.

Figure 7 displays a partial design interface of the intelligent clothing pattern design platform, where subfigure (a) shows the line drawing conversion process, and subfigure (b) presents the style transfer effect of the generated pattern integrated with garments. The main functions of the intelligent pattern design platform are as follows:

1. **Sketch conversion and line drawing generation:** Designers only need to input simple hand-drawn sketches, which are then automatically transformed into standardized cloud motif line drawings by the improved conditional generative adversarial network, significantly lowering the threshold for digital drawing.

2. **Style transfer and fusion:** Through the hierarchical style reconstruction algorithm, users can fuse modern artistic styles with traditional line drawings, realizing innovative expression of cultural genes.

Through this intelligent design platform, highly efficient and intelligent design tools can be provided for both users and designers, enabling the conversion of hand-drawn sketches and line drawing generation, style fusion and transfer. This design platform not only offers intelligent support for the traditional clothing design workflow but also provides a certain reference for subsequent research on multi-style pattern fusion and related topics.



(a)

(b)

Figure 7: Line drawing conversion (a) and garment fusion and matching interface (b).

5 Conclusion and Future Outlook

Addressing the pain points of low digitization efficiency, difficulty in feature extraction, and susceptibility to distortion in style transfer inherent in the design of cloud motif patterns for traditional ethnic costumes, this paper proposes a systematic deep neural network solution. The main contributions are summarized as follows:

1. The proposed improved StyleGAN model, by virtue of adaptive noise injection and a joint constraint of edge and structure losses, effectively resolves the line breakage problem in the generation of complex geometric lines.
2. The proposed sketch translation algorithm incorporating a self-attention mechanism achieves highly standardized end-to-end hand-drawn sketch conversion.
3. A multi-scale neural style transfer framework is designed, which successfully expands the boundaries of modern fusion design for traditional cloud motifs by decoupling low- and high-level feature losses and introducing attention mechanisms.

Limitations and outlook: Although the proposed model perform excellently in most scenarios, it may still exhibit local feature blurring when handling extremely dense intersecting and overlapping textures. In the future, text-driven diffusion models and multimodal large language models could be introduced to construct a “text-to-image” workflow, thereby enhancing semantic precision and personalized controllability in the intelligent creation of traditional patterns.

Funding

No funding was obtained for this study.

Author Contributions

Conceptualization, Yu Chen; methodology, Mengyuan Li; software, Guodong Xu; validation, Mengyuan Li; formal analysis, Mengyuan Li; investigation, Mengyuan Li; data curation, Mengyuan Li; writing—original draft preparation, Mengyuan Li; writing—review and editing, Yu Chen; supervision, Yu Chen; funding acquisition, none. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

All the authors declare that they have no conflict of interest.

References

[1] Khanafiah, D., Situngkir, H.: Computational Batik Motif Generation: Innovation of Traditional Heritage by Fractal Computation. SSRN preprint (2009). https://doi.org/10.2139/ssrn.1346403

[2] Radford, A., Metz, L., Chintala, S.: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. In International Conference on Image and Graphics, pp. 97–108 (2017). https://doi.org/10.1007/978-3-319-71589-6_9

[3] Karras, T., Aila, T., Laine, S., Lehtinen, J.: Progressive growing of Gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196 (2017). https://doi.org/10.48550/arXiv.1710.10196

[4] Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. InProc eedings of the IEEE international conference on computer vision, pp. 2223–2232 (2017). https://doi.org/10.1109/ICCV.2017.244

[5] Park, T., Liu, M.Y., Wang, T.C., Zhu, J.-Y.: Semantic image synthesis with spatially-adaptive normalization. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 2337–2346 (2019). https://doi.org/10.1109/CVPR.2019.00244

[6] Gatys, L.A., Ecker, A.S., Bethge, M.: A neural algorithm of artistic style. Journal of Vision, **16**(12), 326 (2016). https://doi.org/10.1167/16.12.326

[7] Sun, Y., Chen, Y.: Fast textile pattern generation combining MRF-based and Gram-based methods. Industria Textila **74**(4), 439–445 (2023). https://doi.org/10.35530/IT.074.04.202254

[8] Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In Proceedings of the 34th International Conference on Machine Learning, pp. 214–223 (2017).

[9] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., *et al.*: An image is worth 16x16 words: Transformers for image recognition at scale. In The Eighth International Conference on Learning Representations, pp. 1–19 (2021)

[10] Li, M., Lin, Z., Mech, R., Yumer, E., Ramanan, D.: Photo-sketching: Inferring contour drawings from images. In 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1403–1412 (2019). https://doi.org/10.1109/WACV.2019.00154

[11] Gatys, L.A., Ecker, A.S., Bethge, M.: A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576 (2015). https://doi.org/10.48550/arXiv.1508.06576

[12] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.,: Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems, pp. 6000–6010 (2017)