

Optimizing the Digital Design Process for Lolita-Style Clothing Patterns Using the StyleGAN Image Generation Algorithm

Meng Ren, Jinjiang Wang, Gaojie Zheng

College of Fashion and Art Design, Minnan University of Science and Technology, Quanzhou 362700, China

Corresponding author: Gaojie Zheng (e-mail: baifabaizhong0305@163.com)

ABSTRACT With the increasing adoption of electromagnetic sensing, wireless imaging, and intelligent information acquisition technologies in digital fashion design and virtual manufacturing, high-quality pattern generation has become an important component of efficient visual information processing. To address the low efficiency of manual creativity and the difficulty of preserving cultural characteristics in the digital design of Lolita-style clothing patterns, this paper proposes an intelligent pattern generation framework based on an improved StyleGAN. A multi-scale style control architecture with a 12-layer mapping network and a hierarchical attention mechanism is developed to achieve fine-grained control of representative elements such as lace, ruffles, and decorative textures. Furthermore, a cultural feature constraint module combining VGG-19 feature alignment and Gram matrix style loss is introduced to preserve semantic consistency with Lolita aesthetic norms, while an improved adaptive instance normalization mechanism enhances local feature representation through grouped channel control and learnable normalization strength. A latent-space optimization strategy and seamless integration with CLO 3D further support efficient digital design and virtual garment development. Experimental results demonstrate that the proposed method achieves an average aesthetic score of 8.7, compresses the design cycle by 92.7%, and significantly reduces manual intervention while maintaining stable robustness under different noise conditions. The proposed framework provides an efficient and reliable solution for intelligent fashion pattern generation and offers methodological support for electromagnetic-enabled visual information acquisition, digital design systems, and multidisciplinary engineering applications involving advanced image processing and virtual modeling.

INDEX TERMS StyleGAN, clothing pattern generation, cultural feature constraint, digital fashion design, electromagnetic sensing

I. INTRODUCTION

As a unique fashion category that combines elements of European classical clothing with characteristics of Japanese subculture, Lolita-style clothing is designed with delicate elements such as lace, ruffles, and bows as its core, and encompasses cultural connotations from multiple sub-styles, such as Gothic darkness and sweet fantasy [1]. With the widespread adoption of digital design technology in the clothing industry, the digital creation of Lolita patterns has gradually become a key approach to improving design efficiency. However, the current model of manual drawing or simple material splicing is inefficient throughout the workflow, from creative sketches to final rendering evaluation, and struggles to meet the iterative requirements of the fast fashion industry [2]. Additionally, the design process is highly dependent on the designer's experience to maintain cultural characteristics, which can lead to issues

such as imbalanced lace density and sub-style confusion, causing the generated pattern to deviate from Lolita aesthetic norms [3]. Notably, there is a significant difference between clothing pattern generation and overall clothing style generation. In terms of the continuity of two-dimensional textures, alignment accuracy, and refined control of style elements, existing models lack effective processing for such specific needs [4,5]. Therefore, there is an urgent need to systematically optimize and reconstruct the digital design process of Lolita clothing patterns through intelligent methods, building an efficient closed-loop generation process that accurately preserves cultural characteristics to resolve the conflict between efficiency and cultural inheritance. In recent years, image generation technology based on Generative Adversarial Networks (GANs) has provided a new paradigm for clothing pattern design [6,7]. In early studies, the Deep Convolutional Generative Adversarial

Network (DCGAN) achieved texture generation through a deep convolutional structure. Fan [8] improved the structure and loss function of the standard DCGAN and combined it with real-time style transfer technology to generate clothing image models with higher clarity, more diverse details, and lower resource consumption. Liu [9], based on the improved DCGAN model, transferred the style of ethnic patterns to traditional clothing patterns in real time, improving detail expression and generation efficiency. However, due to the lack of hierarchical control over style features, it is difficult to handle the coordinated presentation of multiple elements in Lolita patterns. Progressive Growing of GANs introduces a progressive growth strategy to improve the generation quality of high-resolution images, but its single style vector input cannot achieve independent adjustment of lace and ruffles [10]. Style-Based Generator Architecture for Generative Adversarial Networks (StyleGAN) achieves fine-grained control of image style by separating the latent space and style control mechanism. For instance, Xiong [11] extended the 2D generation based on StyleGAN-Human to a usable 3D clothed human, which provided important inspiration for the preservation and editing of clothing appearance and texture under 3D/viewpoint transformation. Masukawa [12] combined conditioned StyleGAN2 with a style transfer model to generate high-fidelity clothing images, making the design details clearer. However, the original StyleGAN architecture was not optimized for the cultural characteristics of clothing patterns, and the generated results often exhibited issues such as imbalanced element proportions or blurred sub-style features. Existing GAN-based clothing pattern generation research still has two significant limitations: First, there is a lack of a cultural feature constraint mechanism. Most methods focus solely on visual similarity and ignore the implicit aesthetic rules of the Lolita style, resulting in generated patterns that are “similar in form but different in spirit.” Second, the granularity of element control is insufficient, as no dedicated control module has been designed for the hierarchical structure of Lolita patterns, making it difficult to meet designers’ customized needs for local elements. Furthermore, existing datasets mostly focus on mass clothing categories and lack sufficient data for Lolita sub-styles, which further limits the algorithm’s ability to learn cultural characteristics. In response to the above problems, this paper proposes an intelligent generation method for Lolita clothing patterns based on an improved StyleGAN. This method aims to optimize the entire digital design process, from creative input to pattern generation and evaluation feedback. It enhances the representation and hierarchical control capabilities of key visual elements by deepening the mapping network structure and introducing a hierarchical attention mechanism. At the same time, it builds a constraint module that integrates cultural characteristics to ensure the style conformity of the generated pattern. Additionally, this study explores the integration scheme between the generation algorithm and mainstream clothing design software, designing and imple-

menting a docking interface with CLO 3D. This interface simplifies the application process from pattern to clothing model, reduces designers’ manual adjustment workload, and ensures high-quality presentation of the generated pattern in a three-dimensional virtual fitting environment.

II. LOLITA FASHION PATTERN DIGITAL DESIGN SYSTEM

The traditional Lolita fashion pattern design process typically relies on designers hand-drawing or combining patterns from a library of materials. This process is time-consuming, lacks quantitative standards, and can easily lead to stylistic deviations. To overcome these bottlenecks, this paper proposes and constructs a closed-loop digital design system based on an improved StyleGAN. This system uses the designer’s creative intent as input to generate high-quality patterns that meet Lolita aesthetic standards. The results are then fed back for evaluation and iteration, forming a closed loop of “input–generation–evaluation–optimization.”

STYLEGAN ARCHITECTURE MODIFICATION FOR PATTERN GENERATION

To more effectively handle the complex structure and fine textures of Lolita clothing patterns, this paper makes targeted improvements to the original StyleGAN architecture. The core of these improvements is to enhance the style encoder’s ability to represent specific visual elements of Lolita clothing (lace texture, dense lace, and complex prints) and to optimize the generator’s precise control over these elements [13,14]. First, the mapping network is deepened. The original StyleGAN uses an 8-layer multi-layer perceptron (MLP) to map the initial noise vector $z \in \mathbb{R}^{512}$ to the intermediate latent space W . Considering the hierarchical characteristics of Lolita patterns, the mapping network is expanded to a 12-layer MLP structure, denoted as $f_{\text{map}} : z \rightarrow w$. Each layer uses the LeakyReLU activation function to enhance nonlinear expression capabilities and better capture the complex mapping relationship from abstract noise to specific style features [15]. Secondly, to achieve more refined multi-scale control, the style input dimension of the generator is adjusted from a single $W \in \mathbb{R}^{512}$ to a hierarchical style vector set $\{w_s\}_{s=1}^S$, where S represents the number of resolution levels in the generator. The style vectors $w_s \in \mathbb{R}^M$, where the dimensionality M of each level is adjusted, are responsible for controlling the visual features corresponding to that level. This design allows for the independent injection of style information at different resolution levels: the low-resolution level controls the overall layout and theme elements, while the high-resolution level focuses on generating detailed textures [16]. Furthermore, a hierarchical attention mechanism (HAM) is introduced to enhance the style vector’s ability to perceive key areas [17,18]. For the s level, its style vector w_s interacts with the feature map $F_s \in \mathbb{R}^{C \times H_s \times W_s}$ of that level to generate an attention weight A_s :

$$A_s = \text{Softmax}(\text{MLP}_{\text{att}}(w_s) \cdot \text{Flatten}(F_s)) \quad (1)$$

where MLP_{att} is a small MLP used to project the style vector into a space compatible with the flattened dimension of the feature map. The attention weights A_s are then used to weight the style vector to generate the final hierarchical style representation:

$$\hat{w}_s = w_s \odot A_s \quad (2)$$

where $\odot$ represents element-wise multiplication. This mechanism enables the generator to dynamically adjust the focus of style control according to the characteristics of the current level, thereby more accurately generating the key visual elements of Lolita clothing. To explicitly support the accurate generation of multiple sub-styles of Lolita fashion, a Multi-Sub-Style Classifier was designed at the input of the generator as an auxiliary control module. This classifier uses a lightweight convolutional neural network (CNN) architecture consisting of four convolutional layers and two fully connected layers. It takes the generator's initial latent vector z as input. Its core function is to predict the Lolita sub-style category (such as Gothic Lolita, Sweet Lolita, etc.) to which the vector z belongs based on its distribution characteristics. The predicted category probability vector is then concatenated with the original style vector via an embedding layer as additional conditional information before being fed into the subsequent mapping network and generator. The innovation of this paper's improvement to StyleGAN lies in combining general generation technology with a HAM designed specifically for Lolita clothing patterns, forming a highly customized solution to the core challenges of this specific field. Unlike general image generation or clothing style generation, the bottleneck of Lolita pattern design is not simply generating an attractive clothing image, but precisely controlling the coordinated presentation and aesthetic balance of multiple high-density, small-scale, stylized decorative elements in a two-dimensional plane. Therefore, the improvement in this paper is not a simple stacking of existing technologies, but a systematic reconstruction around this core requirement: a deeper 12-layer MLP mapping network is designed to learn a deep nonlinear mapping from abstract noise to the complex combinations of elements unique to Lolita. The layered style vectors provide independent control channels for features at different scales. Crucially, the HAM module enables a closed-loop perception-control of key visual elements by dynamically enabling the style vectors of each layer to focus on and weight the corresponding feature map regions. This mechanism enables the model to focus on details such as the density of lace and the layering of bows during the generation process, similar to a professional designer. This addresses the issues of element disproportion and chaotic layout that are common in conventional GANs when processing such high-precision, high-complexity patterns.

CULTURAL FEATURE CONSTRAINT MODULE DESIGN

Lolita clothing is more than just a collection of visual patterns; it carries profound cultural connotations and strict aesthetic standards. To ensure that the generated patterns faithfully reflect the core cultural characteristics of the Lolita style, at the semantic content level, cultural characteristics refer to the presence and rational layout of key elements that define the identity of the sub-style. Gothic Lolita requires elements such as dark colors, crosses, bats, and lace mesh, and emphasizes symmetrical and solemn layouts; while Sweet Lolita requires elements such as highly saturated pastels, strawberries, rabbits, ruffles, and puffy skirts, and emphasizes asymmetrical and lively layouts. At the visual style level, cultural characteristics are reflected in specific texture density, harmonious color matching, and the overall atmosphere of the image. To avoid generating designs that deviate from the norm or lack cultural connotations, a cultural characteristic constraint module was designed and integrated. This module guides the generation process towards a direction that is consistent with Lolita culture by introducing a specific loss function [19,20]. The core idea of the cultural feature constraint module is to align the high-level semantic content and low-level texture style of the generated pattern with the predefined Lolita cultural characteristics. To this end, this paper constructs a composite loss function L_{cultural} , which is composed of two weighted parts: content loss L_{content} and style loss L_{style} :

$$L_{\text{cultural}} = \lambda_1 L_{\text{content}} + \lambda_2 L_{\text{style}} \quad (3)$$

where λ_1 and λ_2 are the hyperparameters that balance the importance of the two loss terms. The content loss L_{content} aims to ensure that the generated pattern is consistent with the typical Lolita pattern in terms of high-level semantics. The pre-trained Visual Geometry Group 19 (VGG-19) network is used as the feature extractor because it can effectively capture the semantic information of the image [21,22]. Let G be the generator, x_{ref} be the reference Lolita pattern, and $x_{\text{gen}} = G(z)$ be the generated pattern. The feature map of the 'relu4_2' layer of VGG-19 is selected to calculate the content loss, as the features of this layer retain sufficient semantic information without being too abstract. The content loss is defined as:

$$L_{\text{content}} = \frac{1}{2} \sum_{i,j} (F_{\text{gen}}^l(i,j) - F_{\text{ref}}^l(i,j))^2 \quad (4)$$

F_{gen}^l and F_{ref}^l represent the feature maps of the generated image and the reference image at the l -th layer, respectively, and i, j are the spatial indices of the feature maps. The style loss L_{style} focuses on whether the texture and color matching of the generated pattern conform to the aesthetic norms of the Lolita style. It measures the style similarity by comparing the Gram matrices of the generated image and the reference image. The Gram matrix G^l describes the correlation between the feature map channels, and its elements are defined as:

$$G_{c,c'}^l = \sum_{i,j} F_c^l(i,j) \cdot F_{c'}^l(i,j) \quad (5)$$

where c, c' are the channel indices of the feature map. The style loss is calculated as follows:

$$L_{\text{style}} = \sum_l \frac{1}{4C_l^2 H_l^2 W_l^2} \sum_{c,c'} G_{\text{gen}}^l(c,c') - G_{\text{ref}}^l(c,c')^2 \quad (6)$$

The summation traverses multiple layers of VGG-19, where C_l , H_l , and W_l are the number of channels, height, and width of the feature map at layer l , respectively. The VGG-19 feature extraction effect is shown in Figure 1.

IMPLEMENTATION OF THE ADAPTIVE INSTANCE NORMALIZATION MODULE

Adaptive Instance Normalization (AdaIN) is a key technology for achieving style control in StyleGAN. It normalizes feature maps and applies an affine transformation based on the style vector, thereby injecting style information into the generation process. To more precisely control the complex and varied stylistic elements in Lolita fashion patterns, this paper improves the AdaIN module [23,24]. The standard AdaIN operation is defined as follows: Given a content feature map $x_c \in \mathbb{R}^{C \times H \times W}$ and a style vector $y_s \in \mathbb{R}^{2C}$, AdaIN first performs instance normalization (IN) on the content feature map:

$$\text{IN}(x_c) = \frac{x_c - \mu(x_c)}{\sigma(x_c) + \epsilon} \quad (7)$$

where $\mu(x_c)$ and $\sigma(x_c)$ are the mean and standard deviation calculated along the spatial dimensions (H, W), respectively, and ϵ is a small constant for numerical stability. Then, AdaIN uses the style vector y_s to perform an affine transformation on the normalized features:

$$T(x_c, y_s) = y_{s,\text{scale}} \cdot \text{IN}(x_c) + y_{s,\text{bias}} \quad (8)$$

where $y_{s,\text{scale}} \in \mathbb{R}^C$ and $y_{s,\text{bias}} \in \mathbb{R}^C$ are the first and second halves of the style vector y_s , respectively. To better control different types of visual elements, this paper divides the feature channels C into G groups, each corresponding to a specific element type. In this way, the style vector is also divided into G sub-vectors, each of which independently controls the affine transformation of a group of channels. This makes the style control more refined and avoids mutual interference between different element features. In the standard AdaIN, the normalization process is fixed. A learnable parameter $\alpha \in [0, 1]$ is introduced to adjust the strength of normalization:

$$\text{IN}_\alpha(x_c) = \alpha \cdot \frac{x_c - \mu(x_c)}{\sigma(x_c) + \epsilon} + (1 - \alpha) \cdot x_c \quad (9)$$

When $\alpha = 1$, it degenerates to standard instance normalization; when $\alpha = 0$, no normalization is performed, retaining

the original features. This design enables the network to learn the optimal degree of normalization to adapt to the varying sensitivity of different elements in Lolita patterns to style changes.

LATENT SPACE OPTIMIZATION AND SOFTWARE INTEGRATION

1) Latent Space Optimization

The power of StyleGAN lies in the semantic properties of its latent space (W or $W+$). By operating on the latent vector, new images can be explored and generated without retraining the network. However, original random sampling or simple interpolation often cannot guarantee the optimality of the generated results, especially when specific creative goals are required. Therefore, this paper designs a set of latent space optimization strategies to improve the quality, diversity, and controllability of the generated patterns [25]. The latent space optimization strategy mainly uses guided optimization. Guided optimization aims to find the optimal latent vector based on a specific goal. Given an objective function $J(w)$:

$$J(w) = -L_{\text{cultural}}(G(w)) - L_{\text{style-consist}}(G(w)) \quad (10)$$

Gradient ascent is used to optimize the latent vector w :

$$w^{(t+1)} = w^{(t)} + \eta \nabla_w J(w^{(t)}) \quad (11)$$

where η is the learning rate and t is the number of iterations. To prevent the optimized vector from deviating too far from the distribution of the original W space, a regularization term $R(w)$ is introduced, which is usually the square of the Euclidean distance between the vector and the mean $\bar{w}$ of the W space:

$$w^{(t+1)} = w^{(t)} + \eta \left(\nabla_w J(w^{(t)}) - \lambda_R R(w^{(t)} - \bar{w}) \right) \quad (12)$$

where λ_R is the regularization strength. In this way, starting from a randomly initialized w , the process can gradually optimize the latent vector that can generate high-quality and high-conformity patterns.

2) CLO Clothing Modeling Implementation

To achieve the seamless transition of Lolita clothing patterns from two-dimensional design to three-dimensional clothing products, this system designs and implements a docking solution with CLO 3D. The aim is to efficiently and accurately apply the high-quality two-dimensional patterns generated by the algorithm to the virtual clothing model in CLO 3D, thereby completing the rapid transition from pattern creativity to three-dimensional virtual samples and improving the digitalization and automation level of the overall design process. The system supports exporting generated Lolita patterns to standard image formats and automatically generates a corresponding Drawing Exchange



FIGURE 1. Pattern feature extraction results

Format (DXF) file based on the user-specified garment component, leveraging the built-in or user-uploaded CLO model structure information. This DXF file contains geometric information such as the garment component's precise 2D outlines, seams, and key alignment points. The mapping between the DXF file and CLO 3D's recommended parameters is achieved through a "feature-parameter bidirectional association algorithm." The system first analyzes the structural information of the Lolita pattern in the DXF file and establishes a mapping index based on the texture features of the generated pattern. It then uses a pre-trained "Lolita material parameter library" to assign optimal initial parameters to different patches through cosine similarity matching. Finally, the parameters are fine-tuned based on style features extracted by VGG-19 to ensure semantic consistency between the parameters and the pattern elements. To address the continuity constraints across paper patterns, the system designs a "multi-facet collaborative calibration module": (1) Adjacent facets are identified using seam line markers in the DXF file, and the pixel gradients and orientation angles of the textures between the facets are calculated. First, a "Pattern Repeat Unit Identification Submodule" is added to determine whether the Lolita garment pattern is of the regularly repeating type. For regularly repeating patterns, texture coordinate alignment and texture parameter adjustment are prioritized to ensure that the pattern repeat units are precisely aligned at the seams. The gradient difference threshold is adjusted to 2% to improve alignment accuracy. Only when the gradient difference still exceeds 2% after parameter adjustment is the "Repeat Unit-Based Local Poisson Fusion Algorithm" invoked. This algorithm only processes edge pixels within one repeat unit at the seam to avoid overall pattern deformation. Finally, "Seam Pattern Alignment Error" and "Repeat Unit Integrity" are added as verification indicators to ensure that the pattern maintains structural integrity at the seams to meet the needs of actual garment production; (2) For irregular-shaped facets, UV unfolding consistency verification is used to compare the matching degree of the repeated units of the unfolded texture with the adjacent facets. If the deviation exceeds 8%, parameter backtracking adjustment is triggered; (3) An abnormal parameter fuse mechanism is set. When a conflict is detected between the material parameters of a single facet and the sub-style specifications, it is automatically replaced with safe parameters of the same style in the library, and the conflict point is marked in the log, supporting designers in

fine-tuning the texture of the 3D model or avoiding texture distortion or style deviation caused by abnormal parameters. The modeling effect is shown in Figure 2.

This interface design effectively simplifies the application process of patterns to clothing models, reduces the workload of designers in manually adjusting mapping coordinates and alignment, ensures high-quality visual presentation of generated patterns in a three-dimensional virtual fitting environment, and successfully integrates the creative patterns generated by this system into the complete clothing product development process.

III. EXPERIMENTAL DESIGN

This dataset uses a hybrid of the Deep Fashion Dataset and a self-built Lolita-style dataset to balance data size and style specificity. Within the Deep Fashion Dataset, a sub-dataset related to clothing patterns was selected. A total of 12,000 images containing elements such as lace, frills, and prints were selected as the basic training data for general clothing texture features. The image resolution was uniformly adjusted to $1,024 \times 1,024$. The self-built Lolita-style dataset covers Gothic Lolita and Sweet Lolita sub-styles, totaling 660 images, with a uniform resolution of $1,024 \times 1,024$. A portion of the dataset is shown in Figure 3.

For experimental data partitioning, 12,000 general clothing texture images were selected and divided into training and test sets in a 7:3 ratio. Stratified sampling was used to ensure consistent distribution of elements such as lace, trims, and prints within each subset. Data augmentation was performed using random horizontal flipping (with a probability of 0.5), brightness and contrast fine-tuning ($\pm 10\%$), random cropping (cropping 896×896 regions from $1,024 \times 1,024$ images), and Gaussian blurring (with kernel sizes of 1–3 and a probability of 0.2) to improve the generalization ability of the model. The Adam optimizer is used in the training process, the initial learning rate is set to 1×10^{-4} , and the cosine annealing strategy is used to decay the loss of the validation set to 1×10^{-6} when the loss does not decrease for five consecutive epochs. The batch size is set to 16 based on the graphics card memory; the number of iterations is set to 200 epochs, and the convergence criterion is defined as the fluctuation of the cultural feature constraint loss on the validation set for eight consecutive epochs being less than 1×10^{-5} and the average aesthetic score of the generated patterns remaining



FIGURE 2. CLO clothing modeling

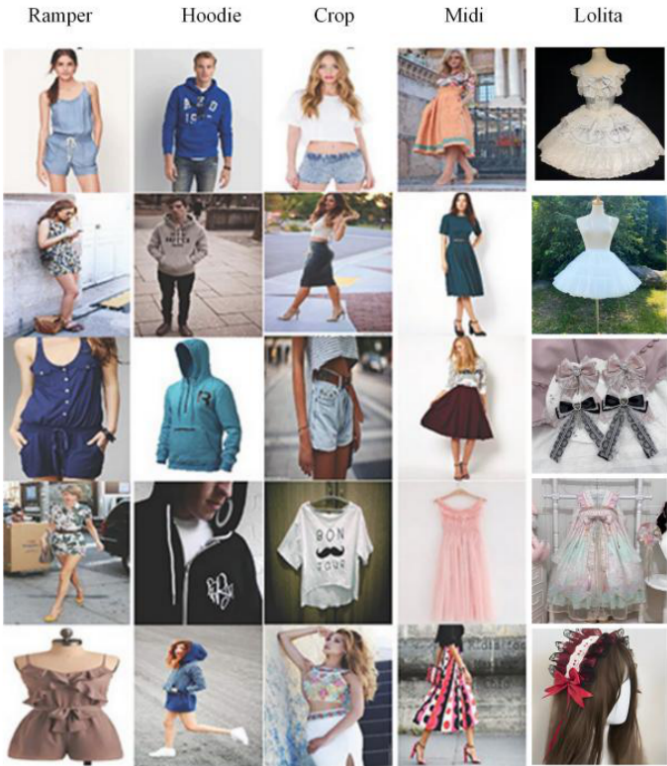


FIGURE 3. Selected experimental datasets

stable above 8.0. At the same time, graphics memory usage is controlled within 28 GB on a single graphics card to

avoid overflow. A stratified sampling method was used to partition the core Lolita dataset of 660 images: first,

the dataset was split into a 7:3 ratio based on sub-styles, resulting in a training set of 462 images and a test set of 198 images. This ensured a consistent distribution of the two sub-styles across all datasets, preventing a single sub-style sample from dominating model learning. Data augmentation ensured both element integrity and diversity, employing only random horizontal flipping, brightness/contrast fine-tuning, and small-scale Gaussian blur. This dataset served as the core reference sample for the cultural identity constraint module. During training, it was mixed with 12,000 ordinary clothing textures. The Lolita samples were used to calculate the content loss and Gram matrix style loss after VGG-19 feature extraction. During testing, the Lolita samples were used solely to verify the sub-style consistency of the generated patterns, preventing ordinary clothing textures from interfering with the assessment of Lolita cultural characteristics. All Lolita clothing pattern images used in the experiment were pre-set to a uniform resolution of 300 DPI. Based on the physical conversion relationship of “1 inch = 25.4 mm,” the “Set Scale” function in ImageJ was used to establish the pixel-to-millimeter conversion formula: $1 \text{ pixel} = 25.4 \text{ mm} / 300 \approx 0.0847 \text{ mm}$. If some images in the experiment were acquired by scanning physical objects, a reference object of known physical dimensions was placed during the scanning process. After scanning, the “Analyze → Set Scale” function in ImageJ was used to calibrate the corresponding pixel count based on the actual length of the reference object. This generated a unique pixel-to-millimeter conversion factor for that image to minimize errors caused by differences in scanning device resolution. This experiment uses the following operational definitions of process indicators, measurement procedures, and multi-operator blind design: the design cycle is defined as the cumulative time from the designer clarifying the Lolita sub-style requirements and core element list to the output of 2D pattern files and 3D CLO virtual samples that meet the acceptance criteria, accurate to 0.1 hour; the frequency of manual intervention is defined as the number of manual adjustments required by the designer during the design process, covering four scenarios: lace density correction, color calibration, layout optimization, and sub-style calibration. This experiment employed a double-blind evaluation design with two core dimensions: 1. Operator-blind: Each operator was only aware of the process type they were currently executing and was completely unaware of other operators’ execution processes or task allocation patterns. Furthermore, task numbers contained only sequential information, without any identifiers related to process type or sub-style. This prevented operators from inferring task associations based on numbers, ensuring that each operational step was independent and free of interference. 2. Evaluator-blind: Evaluators received only the final design results. All results concealed information about operator number, execution process type, and the association between task number and Lolita sub-style. Validity was determined solely based on pre-set acceptance criteria, and any background information

that could potentially trigger subjective bias was masked throughout the evaluation process to ensure objectivity. This method is compared with the original StyleGAN model and the DCGAN clothing pattern generation model. This comparison with the original StyleGAN verifies the effectiveness of the proposed targeted improvements to StyleGAN in generating and finely controlling complex elements in Lolita clothing patterns. DCGAN is widely used in the field of image generation and has a relatively simple structure. Choosing it as a comparison highlights the advantages of StyleGAN and its improved versions in multi-scale style control and texture detail generation, especially the ability of this method to depict fine elements such as lace and ruffles in the Lolita style.

IV. RESULTS ANALYSIS

PROCESS OPTIMIZATION EFFECTIVENESS

Design cycle compression is a core indicator for measuring process optimization effectiveness. The experiment selected several typical Lolita patterns and compared the design cycle compression rates of the traditional manual design process with those of this method to verify the robustness of the method in batch scenarios. The starting boundary was when the designer received the pattern design requirements and identified the sub-style and core elements, and the ending boundary was when the final pattern file was generated and met the acceptance criteria. The control group was the traditional manual process (including hand-drawing creativity, software splicing, repeated adjustments, and 3D modeling adaptation), and the experimental group was the proposed method. The results are shown in Figure 4.

The traditional workflow cycle fluctuates between 45 and 60 hours, which is significantly affected by differences in manual operations. However, the cycle of the improved method in this paper is stable in the range of 3 to 5 hours, and the mean line is almost horizontal, reflecting the algorithm’s precise control of the design rhythm. The stability of the compression performance as the sample size increases is quantified through the cumulative compression rate curve and the 95% confidence interval: when the cumulative number of patterns exceeds 30, the compression rate converges to $92.7\% \pm 1.2\%$; the confidence interval width is less than 2%, which verifies the robustness of the improved method in batch design scenarios. In the digital design process of Lolita-style clothing patterns, the frequency of manual intervention directly affects design efficiency and stability. To visually demonstrate the differences in intervention frequency between the two workflows and the optimization effects of each type of intervention, a statistical comparison of the frequency of manual intervention using different methods was conducted. The results are shown in Figure 5.

As shown in Figure 5, layout optimization and color adjustment have the highest intervention frequency in the traditional workflow, reflecting the difficulty of manually controlling the coordinated presentation of complex ele-

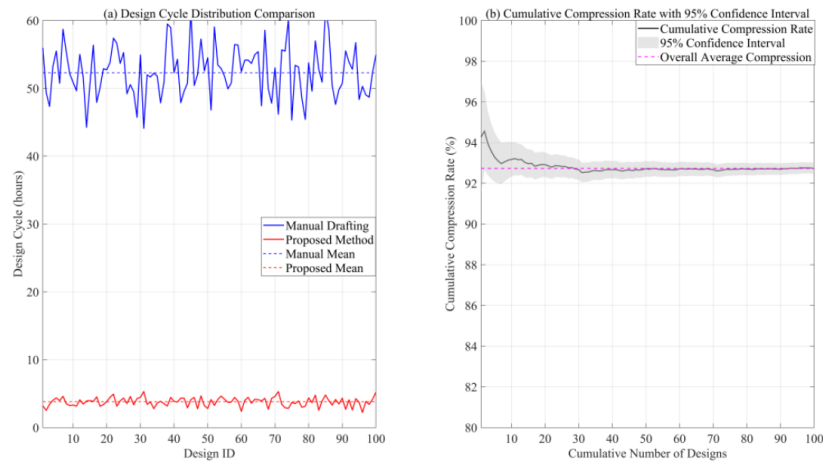


FIGURE 4. Design cycle compression rate

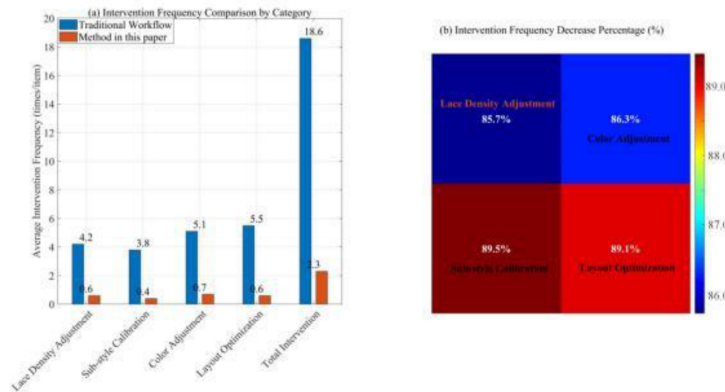


FIGURE 5. Intervention Frequency and Reduction Rate

ments. The proposed method significantly reduces the frequency of intervention in each category by dynamically adjusting the style control focus through HAM. Heatmaps quantifying the reduction in the percentage of interventions across categories show that lace density correction and sub-style calibration reduced the percentages by 85.7% and 89.5%, respectively. This confirms the precise constraint effect of the composite loss function in the cultural feature constraint module on the core elements of Lolita, reducing the need for manual calibration of aesthetic standards.

PATTERN GENERATION QUALITY

Ten professional designers with at least five years of Lolita design experience were invited to participate. The judges comprised three groups: four full-time designers from leading Lolita brands, three independent Lolita designers (with at least five years of custom design experience), and three senior editors from Lolita fashion media (with extensive experience in style appreciation and evaluation). This ensured coverage of both design practice and style research perspectives. The scoring criteria used a three-dimensional, refined scale: color harmony (0–10 points), rational element layout (0–10 points), and stylistic consistency (0–10 points).

A one-hour pre-scoring training session was held, and the scoring scale was calibrated using five standard examples. Afterward, Cronbach's α coefficient was used to test statistical consistency. The results are shown in Figure 6.

It can be seen that the median and upper quartile values of the proposed method are significantly higher, the dispersion is smaller, and the average aesthetic score of the generated pattern is 8.7 points, which is significantly higher than the 6.2 points of the original StyleGAN and the 5.1 points of DCGAN. Representative lace elements from Lolita patterns were selected to quantitatively assess element control accuracy. ImageJ image analysis software was used to select three non-overlapping sampling points in the area with the clearest lace texture. The sampling protocol is as follows: first, in the lace image to be analyzed, a complete, blur-free/noise-free clear continuous area of lace is determined through visual screening and software contrast adjustment; then, a fixed sampling rule is set based on the extension direction of the lace: if the lace extends horizontally, the "top sampling point," "middle sampling point," and "bottom sampling point" are evenly selected along the longitudinal dimension within the clear area; if the lace extends vertically, the "left sampling point,"

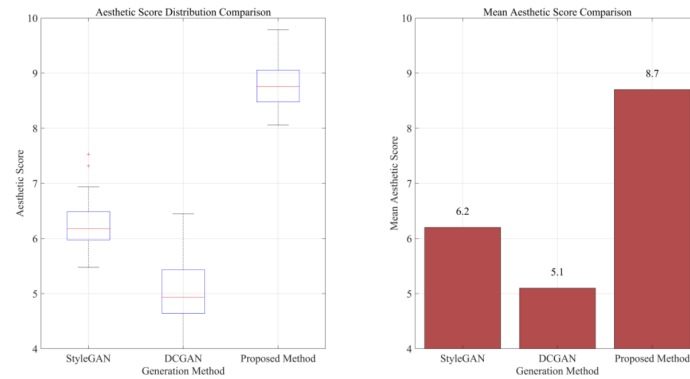


FIGURE 6. Aesthetic scoring results

“middle sampling point,” and “right sampling point” are evenly selected along the transverse dimension, and the three sampling points must be located in three different complete lace units respectively to ensure that they do not overlap; at the same time, the sampling points must avoid the edge transition zone of the lace, and only the core position of the unit with the clearest main texture of the lace is selected to reduce measurement errors. The results are shown in Figure 7.

As can be seen from Figure 7, as the lace reference width increases from 0.5 mm to 2.0 mm, the error rate of this method remains stable in the range of 2.0%–4.0%, with a fluctuation range of less than 1.2%. The error rates of the original StyleGAN and DCGAN fluctuate in the range of 9.0%–14.0% and 14.0%–18.0%, respectively. This shows that the improved AdaIN module effectively solves the problem of proportional imbalance when generating lace of different widths through grouped feature control and learnable normalized strength parameters. Mean quantization results further validate the improved method’s average lace width error rate, which was 72.9% and 79.5% lower than the original StyleGAN and DCGAN, respectively. This provides technical support for the precise generation of lace elements in Lolita patterns and, together with the composite cultural feature constraint loss function, ensures the control accuracy of pattern elements.

ROBUSTNESS ANALYSIS

This paper adds a Gaussian perturbation Δw to the initial noise vector z in the StyleGAN latent space (W space). The perturbation strength is quantified using the L2 norm $\|\Delta w\|_2$ (0.1, 0.3, 0.5, 1.0). The Peak Signal-to-Noise Ratio (PSNR) of the generated patterns compared to the original patterns is calculated for auxiliary verification. Furthermore, the quality scores of the generated patterns at different perturbation intensities and their degree of conformity to Lolita cultural characteristics are evaluated to verify the robustness of the method. The results are shown in Figure 8.

As can be seen from Figure 8, under different $\|\Delta w\|_2$, this method is significantly better than StyleGAN and DCGAN in both pattern quality and cultural fit, and the overall

performance tends to remain stable as noise increases. To further verify the role of each key module in the overall methodology, this paper designs an ablation experiment to compare and analyze the generation results when different components are removed. As shown in Figure 9, taking Gothic Lolita and Sweet Lolita as typical examples, the differences in clothing pattern generation between the complete model and different ablation settings are visually demonstrated, and the impact of each module on style modeling and structure preservation is evaluated from a visual perspective.

As shown in Figure 9, the complete method generates clothing patterns with clear structure, complete texture details, and stable style characteristics under both styles, demonstrating good overall consistency. When the cultural constraint module is removed, the generated results exhibit significant style mixing, with some Gothic and Sweet color and decorative features intermingling, indicating that this module plays a crucial role in maintaining the semantic purity of the style. Further removal of the improved AdaIN structure results in a significant degradation of pattern details and local textures, with lace and decorative layers becoming blurred, indicating that this module makes a significant contribution to the expression of fine-grained style features. Furthermore, without classifier constraints, the overall color distribution of the generated results deviates from the target style, reflecting the importance of style discrimination information for stabilizing the generated results.

ABLATION EXPERIMENTS

To quantify the contribution of each core module to the performance of the proposed method, the paper gradually removed key components and compared the differences in pattern quality, cultural consistency, and generation efficiency between the complete method and the ablation model. This clarified the functional necessity of each module. The results are shown in Table 1:

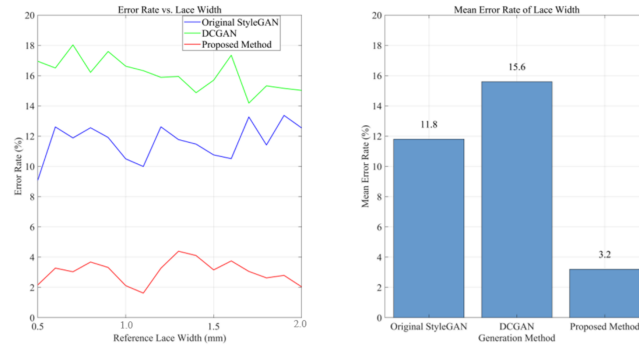


FIGURE 7. Element control accuracy

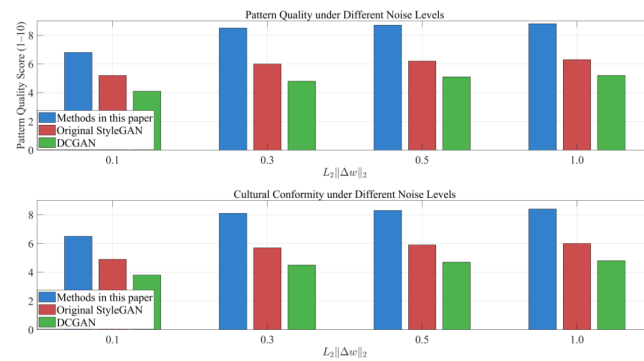


FIGURE 8. Changes under noise interference for different algorithms



FIGURE 9. Patterns generated under different noise levels

TABLE 1. Ablation Experiment Results

Model Configuration	Pattern Quality	Quality Loss Rate	Cultural Characteristics Conformity	Cultural Loss Rate	Generation Efficiency	Efficiency Loss Rate
Complete Method	100%	0%	100%	0%	100%	0%
Removed Cultural Trait Constraint Module	82.5%	17.5%	76.8%	23.2%	98.2%	1.8%
Removed Improved AdaIN Module	86.3%	13.7%	88.1%	11.9%	95.6%	4.4%
Removed Multi-Sub-Style Classifier	89.2%	10.8%	91.5%	8.5%	99.1%	0.9%

Table 1 shows that the cultural feature constraint module has the most significant impact on performance. Its removal reduces cultural feature conformity by 23.2% and pattern

quality by 17.5%. This demonstrates that this module is a core component for preserving the cultural essence of Lolita. Its use of the cultural feature constraint loss function effectively prevents style drift. Removing the improved AdaIN module results in a 13.7% loss in pattern quality and an 11.9% loss in cultural feature conformity, confirming its key role in controlling the fineness of lace textures and generating layered lace elements. The multi-sub-style classifier had the least impact, with both quality and cultural consistency losses of approximately 10% after its removal, indicating that it primarily aided in sub-style differentiation rather than generating basic patterns.

V. CONCLUSIONS

This paper constructs a Lolita fashion pattern generation framework based on an improved StyleGAN, effectively overcoming the efficiency bottlenecks of manual design and the problem of cultural identity distortion. Within the multi-scale style control system, a 12-layer MLP mapping network and a hierarchical attention mechanism achieve fine-grained hierarchical control of elements such as lace texture density and bow layering. The improved AdaIN module resolves the conflicting issue of multi-element collaborative presentation through grouped feature control and a learnable normalized intensity parameter. The cultural feature constraint module leverages VGG-19 feature alignment and Gram matrix style metrics to ensure that generated patterns maintain sub-style identity. The proposed hierarchical control and feature constraint framework exhibits excellent transferability, and its core concept provides a general technical paradigm for the intelligent generation of other culturally distinct garments. For example, for traditional garments such as cheongsam, kimono, and Hanfu, which also possess complex hierarchical structures and profound cultural connotations, the hierarchical attention mechanism of this method can be adapted to independently control the morphology and pattern of different components. The cultural feature constraint module learns and constrains the aesthetic norms of specific garments by replacing corresponding style reference images and pre-trained feature extractors. This demonstrates that this study not only solves a specific problem, but also provides a set of scalable and reusable modular methodologies, laying a technical foundation for the broader digitalization and intelligent design of traditional cultural costumes.

AUTHOR CONTRIBUTIONS

Conceptualization – Meng Ren and Gaojie Zheng; methodology – Jinjiang Wang and Gaojie Zheng; investigation – Jinjiang Wang; writing-original draft preparation – Meng Ren and Jinjiang Wang. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] Y. Zhang, "Rethinking Lolita Fashion: A Case Study in China," *Journal of Asia-Pacific Pop Culture*. 2025; 10(1):55-70. doi: 10.5325/jasia-pacipopcult.10.1.0055.
- [2] S. Gill, H. Al Houf, S. Hayes, and J. Conlon, "Evolving Pattern Practice, from Traditional Patterns to Bespoke Parametric Blocks," *International Journal of Fashion Design, Technology and Education*. 2024; 17(2):144-161. doi: 10.1080/17543266.2023.2260829.
- [3] P. Bai, "A decolonial analysis of Lolita dressing practice and fashion in Mainland China," *Global Media and China*. 2024; 9(1):52-68. doi: 10.1177/20594364231205081.
- [4] S. Y. Jang, C. Chung, and J. Ha, "A comparative study on the fashion design process utilizing shape memory textiles and conventional textiles: Implications for the industry and education," *Fashion and Textiles*. 2025; 12(1):12. doi: 10.1186/s40691-025-00421-2.
- [5] H. E. Smith and K. P. Winterich, "Trendy or Classic? Increasing Sustainability by Increasing Preference for Style Durability," *Journal of Sustainable Marketing*. 2024; 5(1):66-76. doi: 10.51300/JSM-2024-120.
- [6] J. Jung, H. Kim, and J. Park, "Deep Fashion Designer: Generative Adversarial Networks for Fashion Item Generation Based on Many-to-One Image Translation," *Electronics*. 2025; 14(2):220. doi: 10.3390/electronics14020220.
- [7] D. Zhou, H. Zhang, J. Ma, and J. Shi, "BC-GAN: A Generative Adversarial Network for Synthesizing a Batch of Collocated Clothing," *IEEE Transactions on Circuits and Systems for Video Technology*. 2024; 34(5):3245-3259. doi: 10.1109/TCSVT.2023.3318216.
- [8] L. Fan, "Design and application of an intelligent generation model for fashion clothing images based on improved generative adversarial networks," *Service Oriented Computing and Applications*. 2024; 1-14. doi: 10.1007/s11761-024-00429-4.
- [9] Y. Liu and M. Wu, "Intelligent Design of Ethnic Patterns in Clothing using Improved DCGAN for Real-Time Style Transfer," *International Journal of Advanced Computer Science & Applications*. 2023; 14(11):1034-1044. doi: 10.14569/ijcsa.2023.01411105.
- [10] A. Imtiaz, N. Pathirana, S. Saheel, K. Karunanayaka, and C. Trenado, "A Review on the Influence of Deep Learning and Generative AI in the Fashion Industry," *Journal of Future Artificial Intelligence and Technologies*. 2024; 1(3):201-216. doi: 10.62411/faith.3048-3719-29.
- [11] Z. Xiong, D. Kang, D. Jin, W. Chen, L. Bao, S. Cui, et al., "Get3DHuman: Lifting StyleGAN-Human into a 3D Generative Model using Pixel-aligned Reconstruction Priors," *Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV)*; 1-6 October 2023; Paris, France. New York, NY, USA: IEEE; 2023. p. 9253-9263. doi: 10.48550/arXiv.2302.01162.
- [12] R. Masukawa, S. Haji, T. Takagi, T. Matsui, K. Ishikawa, M. Fuchi, et al., "GAN-based Detailed Clothing Generation System," *IIAI Letters on Informatics and Interdisciplinary Research*. 2023; 3. doi: 10.52731/liir.v003.067.
- [13] D. Zhou, H. Zhang, Q. Li, J. Ma, and X. Xu, "COutfitGAN: Learning to Synthesize Compatible Outfits Supervised by Silhouette Masks and Fashion Styles," *IEEE Transactions on Multimedia*. 2023; 25:4986-5001. doi: 10.1109/TMM.2022.3185894.
- [14] H. Yan, H. Zhang, J. Shi, J. Ma, and X. Xu, "Toward Intelligent Fashion Design: A Texture and Shape Disentangled Generative Adversarial Network," *ACM Transactions on Multimedia Computing, Communications and Applications*. 2023; 19(3):107. doi: 10.1145/3567596.
- [15] W. Hu and Y. Zhang, "Research on Artistic Pattern Generation for Clothing Design Based on Style Transfer," *Journal of Combinatorial Mathematics and Combinatorial Computing*. 2025; 127:4539-4550. doi: 10.61091/jcmcc127b-248.
- [16] J. Jiang, D. Wu, H. Deng, Y. Long, W. Tang, X. Li, et al., "HAIGEN: Towards Human-AI Collaboration for Facilitating Creativity and Style Generation in Fashion Design," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2024; 8(3):107. doi: 10.1145/3678518.
- [17] L. Xiao and T. Yamasaki, "Attribute-Guided Multi-Level Attention Network for Fine-Grained Fashion Retrieval," *IEEE Access*. 2024; 12:48068-48080. doi: 10.1109/ACCESS.2024.3383785.
- [18] Y. Wan, G. Zou, C. Yan, and B. Zhang, "Dual attention composition network for fashion image retrieval with attribute manipulation," *Neural Computing and Applications*. 2023; 35(8):5889-5902. doi: 10.1007/s00521-022-07994-9.
- [19] Y. Zheng, T. Zhang, J. Bao, D. Chen, and M. Zeng, "High-fidelity instructional fashion image editing," *Graphical Models*. 2024; 135:101223. doi: 10.1016/j.gmod.2024.101223.
- [20] J. Sun, K. Cen, X. Ding, S. Haidar, and F. Zou, "A Study on Region-Aware Fashion Design Based on Probabilistic Diffusion Model," *IET Image Processing*. 2025; 19(1):e70103. doi: 10.1049/ipr2.70103.
- [21] B. Suvama and S. Balakrishna, "Enhanced content-based fashion recommendation system through deep ensemble classifier with transfer learning," *Fashion and Textiles*. 2024; 11(1):24. doi: 10.1186/s40691-024-00382-y.
- [22] M. Adel, A. Elsayed, and A. Abdelatey, "A Fashion Recommendation System Based on Transfer Learning and Deep Ensemble Model: En-

- hancing Style Prediction,” *IJCI. International Journal of Computers and Information*. 2025; 12(2):30-46. doi: 10.21608/ijci.2024.331343.1181.
- [23] T. Fontanini and C. Ferrari, “Would Your Clothes Look Good on Me? Towards Transferring Clothing Styles with Adaptive Instance Normalization,” *Sensors*. 2022; 22(13):5002. doi: 10.3390/s22135002.
- [24] Q. Zhu, H. Bai, J. Sun, C. Cheng, and X. Li, “Lpadain: Light Progressive Attention Adaptive Instance Normalization Model for Style Transfer,” *Electronics*. 2022; 11(18):2929. doi: 10.3390/electronics11182929.
- [25] S. Huang, Q. Li, J. Liao, S. Wang, L. Liu, and L. Li, “Controllable image synthesis methods, applications and challenges: A comprehensive survey,” *Artificial Intelligence Review*. 2024; 57(12):336. doi: 10.1007/s10462-024-10987-w.