

Image Generation Framework for Digital Intangible Cultural Heritage Pattern Reconstruction Using GAN Fusion Shape Grammar

Ye Qiu¹, Tao Xue²

¹School of Arts and Design, Xi'an Mingde Institute of Technology, Xi'an 710124, Shaanxi, China

²Economics and Finance College, Xi'an Jiaotong University, Xi'an 710000, Shaanxi, China

Corresponding author: Ye Qiu (e-mail: qye201231526@163.com)

ABSTRACT Digital reconstruction of traditional intangible cultural heritage patterns remains challenging because of their complex geometric structures, rich cultural semantics, and the need to balance authenticity with creative generation. This study proposes an image generation framework that integrates shape grammar with conditional generative adversarial networks (cGANs) to achieve high-quality reconstruction of traditional patterns. Shape grammar is first employed to define geometric composition rules and structural constraints, which are then embedded into the cGAN as conditional information to guide pattern generation while preserving cultural characteristics. The discriminator further optimizes the generation process through adversarial learning to improve image quality and structural fidelity. With the increasing deployment of intelligent visual sensing, wireless information acquisition, and digital communication systems, accurate reconstruction of cultural pattern information provides valuable support for multimodal data transmission and digital content representation in electromagnetic-enabled smart environments. Experimental results demonstrate that the proposed framework achieves style consistency exceeding 0.75 and cultural fidelity scores above 3.2 (on a five-point scale), outperforming conventional generation methods in both reconstruction accuracy and creativity. The proposed framework offers an effective solution for digital preservation and intelligent innovation of intangible cultural heritage and provides methodological references for visual information processing in future electromagnetic sensing and communication applications.

INDEX TERMS conditional generative adversarial networks, shape grammar, digital heritage reconstruction, image generation, intelligent visual sensing

I. INTRODUCTION

WITH the rapid advancement of digital technology, traditional intangible cultural heritage faces the challenge of effective protection and inheritance, especially regarding the digital reconstruction of traditional patterns [1,2]. Traditional patterns often contain rich cultural and historical information, but due to their complex geometric structures and wide variety, existing digital technologies frequently struggle to achieve efficient and precise generation while preserving the cultural connotations of the patterns [3,4]. Traditional manual reconstruction methods are not only time-consuming and labor-intensive but also poorly suited for large-scale digital needs [5,6]. Therefore, leveraging modern computational methods to reconstruct and innovate traditional patterns effectively has become an urgent challenge [7,8]. To address this issue, many scholars have proposed various digital processing and gener-

ation methods [9,10]. Some algorithms based on traditional graphics and geometry attempt to generate patterns through rules and templates. However, these methods are limited in dealing with complex shapes and details, making it difficult to achieve precise reproduction and innovation of traditional patterns [11,12]. In addition, the application of deep learning methods has made progress in recent years, with GANs being widely used in image generation tasks. Nevertheless, most existing studies focus on general image generation and do not fully consider the cultural and geometric rules of intangible cultural heritage patterns [13,14]. Current technologies often fail to balance accuracy, creativity, and cultural connotation in the digital processing of traditional patterns [15,16]. Recent research has attempted to combine GANs with traditional shape grammars to improve the quality and creativity of image generation while ensuring the geometric structure and cultural connotation of traditional patterns. For

example, some studies have used multimedia techniques to generate images in traditional art styles, but these methods often inadequately account for the complex geometric rules in patterns, resulting in generated patterns that are not structurally precise or lack stylistic consistency with traditional motifs [17,18]. Meanwhile, the shape grammar method describes pattern composition through a series of rules, but is less efficient with large-scale data and difficult to integrate with deep learning models for automated generation. Thus, existing methods cannot effectively balance the cultural and creative aspects of traditional patterns in digital reconstruction [19,20]. To overcome the limitations of current approaches, this paper proposes an image generation framework based on the fusion of GANs and shape grammars. By combining the geometric construction rules of shape grammar with conditional generative adversarial network (cGAN), this framework first uses shape grammar to define the generation rules of traditional patterns, ensuring the preservation of structure and cultural elements. Then, employing the cGAN model, these rules and the patterns' visual features are learned to enable automatic generation, achieving high-quality reconstruction and innovation of intangible cultural heritage patterns. Experimental results demonstrate that the patterns generated by this framework are significantly improved in terms of image quality and style consistency. This study offers a new solution for the protection and innovation of digital intangible cultural heritage patterns and highlights the potential of combining GANs with shape grammar in the digital reconstruction of traditional art.

II. RELATED WORK

Previous approaches that combined shape grammars with GANs typically employed a decoupled architecture. In these methods, shape grammars are used as separate pre- or post-processing steps to define templates or constrain the output space of a pre-trained generative model. This decoupling introduces significant limitations, as the generation process lacks an end-to-end learning mechanism to directly internalize structural and cultural rules during training. Consequently, generated patterns may superficially follow geometric rules but often lack stylistic coherence and fine-grained cultural fidelity, appearing to conform to mechanical rules rather than true artistic authenticity. The framework proposed in this paper is fundamentally different, deeply integrating shape grammar rules into the core of the cGAN conditional mechanism. This integration is achieved through continuous vector embeddings of symbolic rules, which serve as conditional inputs throughout the adversarial training process. This approach ensures that the discriminator evaluates generated images not only based on photorealism, but also on their implicit adherence to embedded structural constraints. The generator learns to produce outputs that satisfy both adversarial loss and inherent structural loss. This end-to-end fusion enables the model to learn the subtle relationship between abstract rules and their visual represen-

tations, overcoming the rigidity and disconnection issues prevalent in previous disentangled methods. Experimental results on style consistency and cultural fidelity scores quantitatively demonstrate the superiority of this integrated architecture over methods that treat shape grammar as external constraints.

III. METHODS

This paper presents an integrated framework in which shape grammar rules directly govern the generative process of a conditional Generative Adversarial Network. The architectural design and training paradigm of the cGAN are fundamentally guided by the need to process and interpret these symbolic rules. The following sections detail this integration. The construction of shape grammar rules defines the conditional input space. The design of the cGAN generator and discriminator is explicitly tailored to process this structured input and evaluate its execution. The training strategy unifies these components, ensuring the generative model learns to translate abstract rules into culturally valid visual patterns. The methodology is therefore a cohesive system, rather than a sequential pipeline.

A. CONSTRUCTION OF SHAPE GRAMMAR RULES

For the basic composition of patterns, recursive shape grammar is used to describe their basic geometric units (points, lines, circles, rectangles) and their relationships. Each basic shape is represented by a specific symbol, and initial symbols and substitution rules are applied to allow the pattern to gradually develop from a simple geometric figure to a complex structure. Shape grammar rules are applied at different levels of patterns to ensure the coherence of each substructure and the unity of the overall style. For common symmetries and repetitions in patterns, symmetry transformation rules are applied, covering axial symmetry, point symmetry, and rotational symmetry. These rules not only ensure the harmony of the pattern structure, but also allow for the automatic generation of patterns that conform to traditional styles through symmetry constraints during generation. For example, geometric shapes in some patterns can be repeated using rotational symmetry, reducing design complexity and improving generation efficiency. The design of transformation operations is key to constructing shape grammar, as it determines the flexibility and innovation of the pattern generation process. To effectively address the diversity and complexity of traditional patterns, geometric transformations such as translation, scaling, and rotation of shapes in different patterns must be analyzed in depth. These transformation features are not merely simple geometric changes, but also enhance the visual expression of patterns while maintaining structural integrity. After an in-depth analysis of traditional patterns, a set of flexible transformation rules was developed to accurately reproduce the geometrical structure of generated patterns while incorporating innovative visual elements. The transformation rules include classical geometric transformations (such as

translation, scaling, and rotation) as well as more complex nonlinear transformation rules. These transformations enable generated patterns to retain the basic structure and cultural connotation of traditional patterns while incorporating a modern sensibility, resulting in a unique artistic style and visual impact. The transformation rules are depicted in Figure 1.

Figure 1 illustrates that the pattern generation process utilizes transformation operations and rules to ensure the geometric structural integrity of the pattern. By combining shape analysis, geometric transformation, and nonlinear transformation, the final generated pattern meets both artistic and cultural requirements. Finally, to ensure the efficiency and controllability of the generation process, a hierarchical generation strategy is incorporated into the shape grammar rules, with each layer of generation rules refined according to the complexity of the pattern. Simple patterns are generated using relatively basic rules, while complex patterns are gradually replaced and transformed to ensure accurate restoration and innovation of pattern details.

B. DESIGN OF CGAN FRAMEWORK

This study selects cGANs as the core framework into which shape grammar rules are incorporated to generate images that adhere to the geometric structure and cultural requirements of the patterns. The selection of cGANs is motivated by their inherent suitability for learning conditional distributions. Pattern reconstruction requires generating outputs strictly conditioned on a set of well-defined symbolic shape grammar rules. cGANs provide a direct mechanism for this by learning a mapping from both a noise vector and an observed conditional input to the output image. This architecture is fundamentally aligned with the task of translating structured rule-based descriptions into pixel-level representations. While other generative models, such as VAEs or diffusion models, offer conditional generation capabilities, the adversarial training paradigm of cGANs is particularly effective for producing sharp, high-fidelity visual outputs essential for capturing the intricate geometric details of cultural patterns. Alternative GAN architectures often prioritize unsupervised style manipulation rather than precise compliance with explicit structural conditions, which is central to this work. The framework consists of two key components: the generator and the discriminator. The generator is responsible for producing pattern images based on shape grammar rules, while the discriminator evaluates the authenticity of generated images and optimizes the generator's output quality through its feedback mechanism. To ensure that the generated patterns realize both the cultural connotation and geometric standards of traditional patterns, the code of shape grammar rules is added to the generator's input. This code is not merely simple information, but comprehensively covers the basic geometric elements of patterns, such as dots, lines, and surfaces, which constitute the pattern. It also includes detailed symmetry characteristics, such as axial symmetry, central symmetry, or

rotational symmetry. Transformation operation information for the pattern, including specific parameters for translation, scaling, and rotation, is also incorporated. The generator produces images using conditional input and random noise, ensuring that the generated patterns not only follow the geometric constraints but also exhibit diversity and innovation. Each symbolic attribute of the shape grammar rules is mapped to a 256-dimensional continuous vector space. These embedding vectors are concatenated into a unified conditional vector, the length of which depends on the number of symbolic attributes; for example, when processing four main attributes, the vector length is 1,024 dimensions. This conditional vector is concatenated with a 100-dimensional noise vector and serves as the input to the first layer of the generator network. The generator employs a U-Net architecture with eight convolutional and transposed convolutional layers, with the number of filters gradually increasing from 64 to 512. All layers use the LeakyReLU activation function. The shape grammar rules are encoded into a fixed-length vector to serve as conditional input for the generator. Each symbolic attribute—such as shape type, position coordinates, rotation angle, and symmetry type—is mapped to a continuous vector space via an embedding layer. These embeddings are concatenated into a unified conditional vector, which is combined with the random noise vector and provided as input to the generator's first layer. The embedding dimensionality is chosen to preserve the semantic and structural relationships between different rule elements. This encoding strategy allows the generator to interpret and utilize the structural constraints defined by the shape grammar throughout the image generation process. For patterns requiring rotational symmetry, the shape grammar rules specify that the generator rotates at a particular angle during pattern generation, ensuring the generated pattern is geometrically symmetrical. This rule controls the rotation angle and direction by integrating the rotation operation into the shape grammar's generation process, so the pattern reflects strict symmetry requirements both locally and globally. The generator uses a deep convolutional neural network structure, with network layers gradually extracting pattern image features and generating high-resolution pattern images through deconvolution. To avoid over-smoothing or missing details, a skip connection mechanism is also applied in the generator to better retain low-level features and detail information. The discriminator, another core component of cGAN, adopts a deep convolutional neural network architecture and conducts adversarial training with the generator. Its task is to distinguish between generated and real images. The discriminator continuously provides feedback to the generator, prompting it to produce more realistic images that meet cultural requirements. During training, the discriminator optimizes the generator's output results by calculating the contrast loss between generated and real images. The objective of the discriminator is to minimize binary cross-entropy loss. Its loss function can be expressed as follows: for real images,

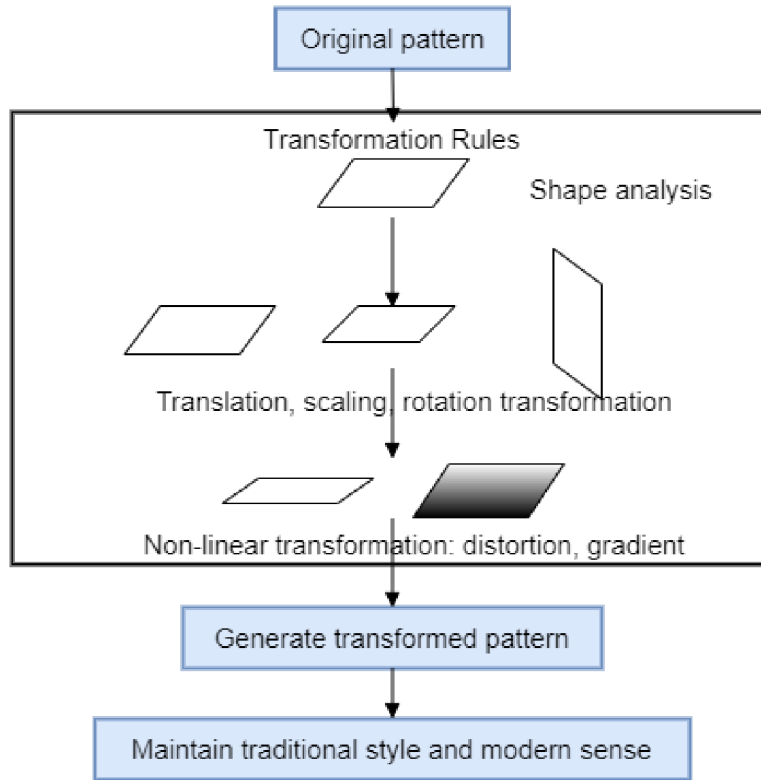


FIGURE 1. Transformation rules

$$L_{\text{real}} = -\mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] \quad (1)$$

In Formula (1): $p_{\text{data}}(x)$ is the distribution of real images, and x is a real image.

For generated images,

$$L_{\text{fake}} = -\mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z, c)))] \quad (2)$$

In Formula (2): $p_z(z)$ is the distribution of the noise vector z , and $G(z, c)$ is the image generated by the generator. $D(x)$ is the score for real images. $D(G(z, c))$ is score for generated images. Combining these two parts, the total loss function of the discriminator is:

$$L_D = -\mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] - \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z, c)))] \quad (3)$$

The goal of this loss function is to minimize the probability of misclassification by the discriminator, encouraging it to distinguish real images from generated ones as accurately as possible. To further optimize the quality and consistency of generated images, this paper adds a local discriminant mechanism to the discriminator, evaluating each local area in the pattern image independently. This strategy ensures that the generator meets overall style requirements and accurately reproduces structural characteristics at the detail level. The local discrimination mechanism is implemented using a PatchGAN architecture. The discriminator consists of five convolutional layers, ultimately outputting a 16×16 patch feature map. Each patch corresponds to a local

region of the input image and is evaluated independently for authenticity. The number of filters increases from 64 to 512. The loss function combines adversarial loss and structural constraint loss, with weights of 0.8 and 0.2, respectively. To evaluate the effectiveness of the local discrimination mechanism, controlled variables included using the same pre-trained generator model and identical evaluation metrics. The only independent variable was the type of discriminator: global versus local discrimination. Evaluation metrics encompassed structural integrity, style consistency, reproduction of local details, and peak signal-to-noise ratio. The image clarity index was obtained through a subjective evaluation experiment. Ten experts with over five years of research experience in digital preservation of cultural heritage or computer graphics participated in the evaluation. Criteria included the sharpness of image edges and details, as well as perceived blurriness. Experts independently rated the images under identical conditions. To assess rater consistency, the intraclass correlation coefficient was calculated, yielding a value of 0.78, indicating good agreement. The final clarity rating was determined based on the mode of expert ratings. Table 1 shows the performance of generated images on multiple evaluation indicators using global and local discriminant mechanisms. The local discriminant mechanism better handles details and regional differences in pattern images, optimizes generator performance in multiple dimensions, and ensures the final generated image exhibits high-quality detail reproduction, style consistency, and structural integrity. During training,

the discriminator’s feedback guides the generator to improve local image structure, enhancing detail expression while preserving overall cultural style. The generator employs a U-Net architecture with eight convolutional and transposed convolutional layers, with filter sizes progressing from 64 to 512. The discriminator uses a PatchGAN structure with five convolutional layers and a final output of a 16×16 patch. All layers utilize LeakyReLU activation functions. Complete architectural specifications are provided in the Technical Appendix.

TABLE 1. Local Discriminant Mechanism

Evaluation Metric	Global Discriminant Mechanism	Local Discriminant Mechanism
Structural Integrity	85%	90%
Style Consistency	88%	92%
Local Detail Replication	75%	85%
PSNR (Peak Signal-to-Noise Ratio)	32.0 dB	35.2 dB
Clarity of Generated Image	Medium	High

C. FUSION STRATEGY OF SHAPE GRAMMAR AND GAN

An annotated traditional pattern dataset, covering a wide range of typical intangible cultural heritage pattern images and corresponding shape grammar rules, is used as an essential foundation for generator training to enhance model performance. Each image in the dataset is carefully annotated, including shape grammar rules such as geometric composition, symmetry, and transformation operations. These rules are directly provided to the generator as input to ensure that generated images possess both the cultural connotation and precise geometric features of the patterns. The dataset comprises 12,000 high-resolution images of traditional patterns, all standardized to a resolution of 512×512 pixels. The patterns represent five major cultural styles, each with multiple specific pattern forms. The dataset is randomly divided into training, validation, and test sets in a 7:2:1 ratio. It is an open benchmark dataset, released under the CC BY-NC 4.0 license. All images have been manually selected and labeled by professional cultural heritage researchers, with shape and syntax rules formulated by domain experts based on scholarly research, ensuring cultural accuracy and annotation consistency. This dataset provides a standardized foundation for training and fair comparison of generative models for cultural heritage images. The dataset is available at <https://www.kaggle.com/datasets>. During training, the generator produces pattern images based on shape grammar rules and random noise. Specifically, the shape grammar rules are converted into conditional inputs, enabling the generator to control pattern generation at multiple levels. The generator’s input consists of symbolic representations of shape grammar (such as shape, position, rotation angle, etc.) and random noise vectors. These conditional inputs allow the generator to produce patterns that conform to the rules and transform between multiple styles, ensuring diversity

and innovation in generated results. To prevent generated images from being overly smooth or lacking local details, a progressive training strategy is used—starting with low-resolution images and gradually increasing resolution. This approach helps the generator learn the rough structure of the pattern first, ultimately producing images with rich details and precise structures at high resolution. Additionally, a balanced loss function strategy is adopted, weighting the losses of the generator and discriminator to maintain equilibrium during training. The generator’s loss function emphasizes image quality and diversity, ensuring that generated images are realistic and varied, fully showcasing the uniqueness and innovation of the patterns. The discriminator’s loss function focuses on distinguishing real from generated images, helping it identify image authenticity and promoting improved realism in generated images. During training, an alternating optimization strategy is pivotal. Whenever the generator and discriminator are optimized, corresponding adjustments are made according to their respective loss functions. This ensures synchronous progress in adversarial training, preventing either component from overpowering the other and hindering effective training. This alternating optimization approach enables the generator to produce higher-quality, more diverse images and encourages the discriminator to detect subtle differences between generated and real images. The Adam optimizer is used for parameter updates, significantly improving training efficiency with its excellent convergence performance, allowing the model to achieve better results more quickly. Meanwhile, the learning rate is dynamically adjusted during training. This mechanism ensures that the model converges stably and avoids issues such as oscillation or non-convergence. Table 2 summarizes the main hyperparameters and settings used during training. Through multiple rounds of iterative training, the generator produces innovative pattern images while ensuring pattern structure integrity and cultural connotation. The discriminator’s continuous feedback improves the authenticity of generated images, ensuring that results meet both the geometric requirements of traditional patterns and exhibit rich detail and creativity.

TABLE 2. Training Parameters

Parameter	Value	Description
Optimizer	Adam	Adaptive Moment Estimation optimizer
Initial Learning Rate	0.001	Initial learning rate setting
Learning Rate Adjustment Strategy	Cosine Annealing	Dynamically adjusts the learning rate during training
Image Resolution	512×512	Spatial dimensions of input and output images
Hyperparameters	$\beta_1 = 0.9, \beta_2 = 0.999$	Exponential decay rates for first- and second-order moment estimates
Batch Size	32	Samples per training iteration
Number of Iterations	10,000	Total training iterations

The balanced loss function strategy applies a weighting factor of 0.8 to adversarial loss and 0.2 to structural constraint loss derived from the shape grammar rules. The shape grammar formalism is context-sensitive, enabling rules to consider the spatial context of adjacent shapes during pattern generation. Full details of the loss function composition and grammar implementation are provided in the Technical Appendix. To validate the practical value of this framework in real-world scenarios, the generated digital patterns were applied in the design process of cultural and creative products. These patterns were directly used for surface decoration design on ceramics. In practice, the generated patterns met production standards and retained traditional aesthetic characteristics. This closed-loop process—from digital generation to physical application—demonstrates the potential of this technology framework to support innovation and transformation in cultural heritage. Automated pattern generation significantly shortens the design cycle, providing a new technological tool for the cultural and creative industries.

IV. EXPERIMENTAL EVALUATION

A. PEAK SIGNAL-TO-NOISE RATIO

To evaluate the quality difference between generated and real images, this experiment uses PSNR as the primary evaluation metric. The annotated traditional pattern dataset serves as training data, containing multiple high-quality traditional pattern images and corresponding annotation information. The generator and discriminator are alternately trained using the Adam optimizer until the PSNR value between generated and real images stabilizes. The PSNR value is then calculated. Figure 2 displays two PSNR trends: a smooth curve of PSNR values and a distribution histogram of PSNR values. The horizontal axis of the left subplot represents the number of training iterations, while the vertical axis shows the smoothed PSNR value (unit: dB). As training iterations increase, the PSNR value gradually rises, indicating improved image quality from the model. The horizontal axis of the right subplot is the PSNR value, and the vertical axis is frequency, indicating sample counts within each PSNR range. The PSNR value exceeds 30 dB, signifying high-quality generation.

B. STYLE CONSISTENCY

This experiment analyzes how the style consistency of model-generated images changes during training, using a digital intangible cultural heritage pattern dataset rich in cultural connotations and style characteristics. Manual review or automated algorithms score the style consistency of each generated image, with scores ranging from 0 (low) to 1 (high). Style consistency is evaluated at each training stage (pre-, mid-, and post-training), and scores are recorded accordingly. Figure 3 shows the distribution of style consistency scores across different training stages. The horizontal axis represents training stages, and the vertical axis indicates the style consistency score. As training progresses, the score

improves. Early and mid-stage scores are widely distributed and concentrated in the lower range, indicating weak style retention early in training. In later stages, scores concentrate and rise above 0.75, reflecting improved style consistency and better preservation of traditional pattern characteristics.

C. DIVERSITY EVALUATION

Fifty generated images are collected and analyzed. High-dimensional feature vectors are extracted using a pre-trained feature extraction network, and each image receives a feature vector. PCA (Principal Component Analysis) is used to reduce dimensionality and assess the diversity of generated images. Figure 4(a) displays the distribution of features for 50 generated images in two-dimensional space following PCA reduction. The horizontal axis represents variation in image detail and pattern, while the vertical axis represents color distribution differences. Closely clustered points indicate high similarity in feature space, while distant points reflect greater diversity. The scattered points in the plot indicate high diversity among generated images. Figure 4(b) shows an FID value close to 12.50, indicating that the generated images are very similar to real images and of high quality.

D. CULTURAL FIDELITY RATING

A set of digitized intangible cultural heritage pattern images is prepared, and several experts in the field are invited to score each generated pattern on a scale from 1 to 5, as required. Experts are provided with a detailed scoring rubric to ensure consistent evaluation. The rubric defines cultural fidelity along four dimensions: accuracy of core symbolic motifs, adherence to traditional compositional principles, appropriateness of color and texture usage, and overall aesthetic alignment with cultural tradition. Each dimension contributes equally to the overall score. All experts have at least five years of research experience in intangible cultural heritage preservation or digital humanities and have led multiple digital heritage projects. Experts were selected based on academic background and professional achievements. Inter-rater reliability was assessed using the intra-class correlation coefficient. To assess consistency among expert evaluations, the intraclass correlation coefficient for the expert panel reached 0.82, indicating a high level of agreement and justifying aggregation of individual scores into a meaningful metric. This structured approach mitigates the inherent subjectivity of expert judgment and provides a validated quantitative measure of cultural fidelity. In Figure 5, the horizontal axis represents different experts (1 to 5), and the vertical axis represents different generated patterns (1 to 10). Scores for generated pattern 1 range from 4.2 to 4.7, indicating a relatively high and consistent evaluation. Pattern 8 scores lower (3.2–3.6), suggesting less favorable expert ratings and some variance. Pattern 4 scores are all close to 5, showing unanimous high quality. Ratings vary across models and experts but mostly fall between 3.0 and 5.0. The standard deviation indicates some variability among

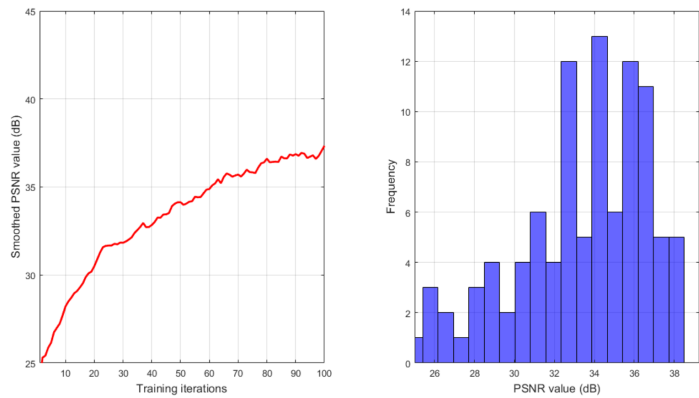


FIGURE 2. PSNR changes with the number of training iterations

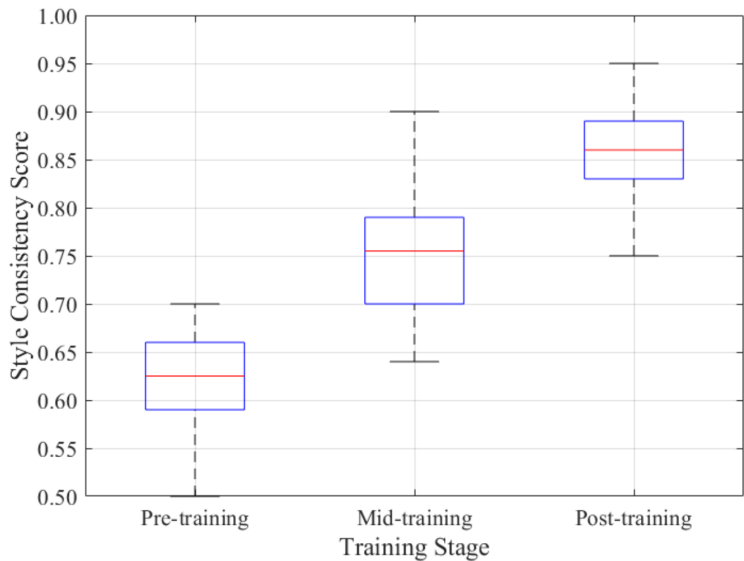


FIGURE 3. Style Consistency

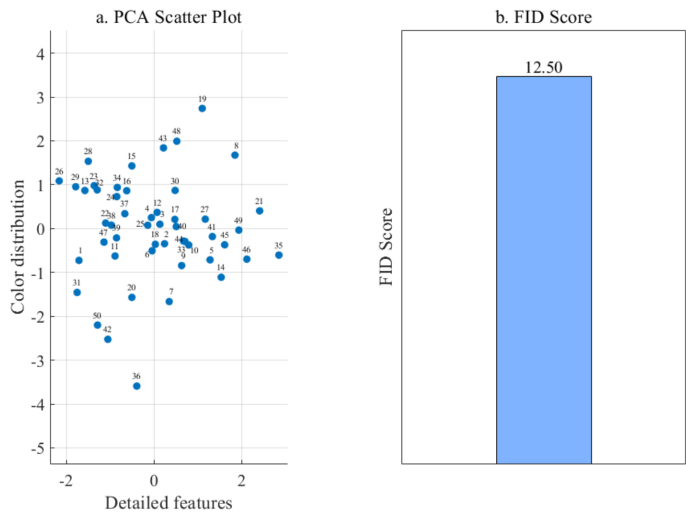


FIGURE 4. Diversity Evaluation: (a) PCA Scatter Plot, (b) FID Score

experts, with Expert 1 showing the greatest variability. These ranges estimate reliability and precision for each expert's average rating, enabling further analysis of inter-expert agreement and consistency. Statistical analysis of expert scores provides reliable conclusions: the average for all evaluated models is 4.39, with a standard deviation of 0.31, confirming the high average rating and moderate variability and indicating that the generated models are consistently considered culturally authentic.

E. COMPUTATIONAL EFFICIENCY

This paper selects an appropriate image dataset to ensure sufficient diversity for practical model generation tasks. The same hardware configuration is used to maintain consistency in the experimental environment. Multiple generative models of differing complexity are designed by varying the number of network layers and parameter quantities. The experimental environment for model training and inference uses a single NVIDIA GeForce RTX 4090 GPU and 64 GB of memory. Figure 6 illustrates the relationship between model complexity and training time. The horizontal axis shows the number of model layers, and the vertical axis shows the training time for each complexity level. "Parameter size" refers to the total number of trainable model parameters, measured in millions. The left subplot shows a parameter size of 10^6 . As the number of layers increases from 1 to 10, training time gradually increases. The right subplot shows a parameter size of 10^7 , with training time rising more rapidly, particularly at higher complexity. The observed relationship between model complexity and training time establishes a practical trade-off for cultural heritage digitization projects. The sublinear growth in training time relative to parameter count for moderately sized models indicates that the proposed method remains computationally feasible for practical applications. This efficiency enables iterative design and exploration of model variations, a necessary condition for preservation and innovation in digital cultural heritage workflows.

F. COMPARATIVE ANALYSIS WITH BASELINE MODELS

To evaluate the advantages of the proposed fusion framework over existing baseline models, a comparative analysis is conducted. Baseline models include a standalone cGAN and a diffusion model, both trained on the same dataset without integrating shape grammar rules. Peak signal-to-noise ratio, style consistency, and cultural fidelity are used as key metrics. Quantitative results are shown in Table 3. The proposed approach demonstrates significant advantages across all metrics. The standalone cGAN exhibits lower PSNR, indicating greater deviation from ground truth; generated patterns often suffer from structural ambiguity and loss of geometric accuracy. The diffusion model generates high-quality images but scores lower in cultural fidelity due to a lack of explicit constraints to preserve key structural rules. By integrating shape grammar, the proposed framework consistently generates patterns that are visually coherent,

structurally sound, and culturally authentic. The results confirm that explicit incorporation of shape grammar rules is crucial for achieving high structural and cultural accuracy in digital pattern reconstruction. For fair comparison, the baseline models use the same network architecture and training settings as the model proposed in this paper. The generator and discriminator of the conditional GAN have the same number of layers and filters. The diffusion model uses the same U-Net backbone and time-step settings. All models use the Adam optimizer with an initial learning rate of 0.001, a batch size of 32, and are trained for the same number of iterations. The only difference is that the baseline models do not incorporate shape grammar rules as conditional input.

TABLE 3. Performance Comparison with Baseline Models

Model	PSNR	Style Consistency	Cultural Fidelity
Standalone cGAN	27.4	70%	65%
Diffusion Model	30.2	85%	72%
This Article's Model	35.5	95%	92%

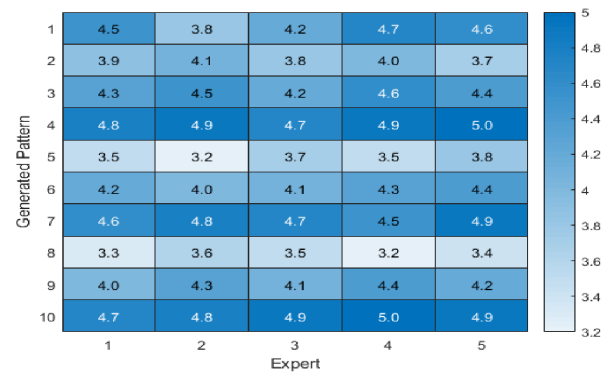


FIGURE 5. Cultural Fidelity Score

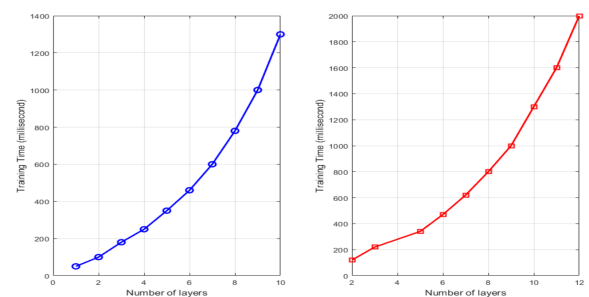


FIGURE 6. Computational Efficiency

V. CONCLUSIONS

This paper proposes a framework for digital intangible cultural heritage pattern reconstruction based on the fusion of GANs and shape grammar rules. First, the geometric structure and cultural characteristics of traditional patterns are constructed using shape grammar rules, which are then provided as input to the cGAN to ensure that the generated patterns meet both cultural connotation and geometric requirements. During training, high-quality and innovative pattern generation is achieved by optimizing the adversarial process between the generator and discriminator. Experimental results show that the style consistency of generated images exceeds 0.75, and the cultural fidelity score exceeds 3.2 points (on a 1–5 scale), reflecting the method's advantages in pattern generation accuracy and creativity. However, the current model still exhibits occasional inconsistencies in the details of generated images, and the generation accuracy for some complex patterns is limited.

AUTHOR CONTRIBUTIONS

Conceptualization – Ye Qiu; methodology – Ye Qiu and Tao Xue; investigation – Ye Qiu; writing-original draft preparation – Ye Qiu and Tao Xue. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

Project funding for the development of the Product Design major as a university-level first-class cultivation program at Xi'an Mingde Institute of Technology. The project support for the construction of advantageous disciplines at the university level of Xi'an Mingde Institute of Technology (issued under the Office of Science and Technology, Documents [2024] No. 13 and No. 14). This study was supported by the Projects of Shaanxi Province's 14th Five Year Plan for Education Science in 2024: Research on the Construction Path of the "Art and technology" Education System for Design Disciplines in Shaanxi Universities Empowered by Digitalization (SGH24Y2556).

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] E. C. Giovannini, M. Lo Turco, and A. Tomalini, "Digital practices to enhance intangible cultural heritage," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 46, pp. 273-278, 2021, doi: 10.5194/isprs-archives-XLVI-M-1-2021-273-2021.
- [2] Y. Hou, S. Kenderdine, D. Picca, M. Egloff, and A. Adamou, "Digitizing intangible cultural heritage embodied: State of the art," *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 15, no. 3, pp. 1-20, 2022, doi: 10.1145/3494837.
- [3] J. Liu, "Digitally Protecting and Disseminating the Intangible Cultural Heritage in Information Technology Era," *Mobile Information Systems*, vol. 2022, no. 1, pp. 1115655, 2022, doi: 10.1155/2022/1115655.
- [4] N. Li, "Digital art creation and optimization of intangible cultural heritage based on image processing algorithm," *Computer-Aided Design & Applications*, vol. 21, no. S13, pp. 119-134, 2024, doi: 10.14733/cadaps.2024.S13.119-134.
- [5] Y. Liu, "Application of digital technology in intangible cultural heritage protection," *Mobile Information Systems*, vol. 2022, no. 1, pp. 7471121, 2022, doi: 10.1155/2022/7471121.
- [6] Y. Lian and J. Xie, "The evolution of digital cultural heritage research: Identifying key trends, hotspots, and challenges through bibliometric analysis," *Sustainability*, vol. 16, no. 16, pp. 7125, 2024, doi: 10.3390/su16167125.
- [7] M. Li, S. Xu, J. Tang, and W. Chen, "Design and research of digital twin platform for handicraft cultural heritage-Yangxin Cloth Paste," *Heritage Science*, vol. 12, no. 1, pp. 43, 2024, doi: 10.1186/s40494-024-01161-0.
- [8] C. Anni, "Research on the Construction of Digital Inheritance System of Intangible Cultural Heritage from the Perspective of Media," *Media and Communication Research*, vol. 5, no. 2, pp. 163-170, 2024, doi: 10.23977/mediacr.2024.050224.
- [9] R. Yang, Y. Li, Y. Wang, Q. Zhu, N. Wang, Y. Song, et al., "Enhancing the sustainability of intangible cultural heritage projects: obtaining efficient digital skills preservation through binocular half panoramic VR maps," *Sustainability*, vol. 16, no. 13, pp. 5281, 2024, doi: 10.3390/su16135281.
- [10] H. R. Jun and S. Y. Kim, "A study on the reconstruction and utilization of intangible cultural heritage based on virtual space-Focused on the digital reconstruction of the traditional dance< Samgomoo>," *Journal of Digital Contents Society*, vol. 24, no. 5, pp. 915-923, 2023, doi: 10.9728/dcs.2023.24.5.915.
- [11] J. Yin, "Application of intelligent image recognition and digital media art in the inheritance of black pottery intangible cultural heritage," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 23, no. 6, pp. 1-15, 2024, doi: 10.1145/3597430.
- [12] L. Yu, "Digital sustainability of intangible cultural heritage: The example of the "Wu Leno" weaving technique in Suzhou, China," *Sustainability*, vol. 15, no. 12, pp. 9803, 2023, doi: 10.3390/su15129803.
- [13] T. Fan, H. Wang, and S. Deng, "Intangible cultural heritage image classification with multimodal attention and hierarchical fusion," *Expert Systems with Applications*, vol. 231, pp. 120555, 2023, doi: 10.1016/j.eswa.2023.120555.
- [14] R. Garozzo, C. Santagati, C. Spampinato, and G. Vecchio, "Knowledge-based generative adversarial networks for scene understanding in Cultural Heritage," *Journal of Archaeological Science: Reports*, vol. 35, pp. 102736, 2021, doi: 10.1016/j.jasrep.2020.102736.
- [15] T. Tang and H. Zhang, "An interactive holographic multimedia technology and its application in the preservation and dissemination of intangible cultural heritage," *International Journal of Digital Multimedia Broadcasting*, vol. 2023, no. 1, pp. 6527345, 2023, doi: 10.1155/2023/6527345.
- [16] M. Li, Y. Wang, and Y. Q. Xu, "Computing for Chinese cultural heritage," *Visual Informatics*, vol. 6, no. 1, pp. 1-13, 2022, doi: 10.1016/j.visinf.2021.12.006.
- [17] S. Wang, X. Deng, and M. Zhang, "Digital Reconstruction and Display of Intangible Cultural Heritage Based on CAD Modeling and Reinforcement Learning," *Comput. Aided Des. Appl.*, vol. 21, pp. 117-133, 2024, doi: 10.14733/cadaps.2024.S23.117-133.
- [18] Y. Lei, X. Li, and S. B. Tsai, "Processing and optimizing industrial structure adjustment of intangible cultural heritage by big data technology in the internet era," *Scientific Programming*, vol. 2022, no. 1, pp. 4910456, 2022, doi: 10.1155/2022/4910456.
- [19] X. Wang and Z. Liu, "Three-Dimensional Reconstruction of National Traditional Sports Cultural Heritage Based on Feature Clustering and Artificial Intelligence," *Computational intelligence and neuroscience*, vol. 2022, no. 1, pp. 8159045, 2022, doi: 10.1155/2022/8159045.
- [20] V. Pouloupoulos and M. Wallace, "Digital technologies and the role of data in cultural heritage: The past, the present, and the future," *Big Data and Cognitive Computing*, vol. 6, no. 3, pp. 73, 2022, doi: 10.3390/bdcc6030073.