EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

A Manufacturing-Oriented Diffusion Framework for Generative Product Design with Traditional Cultural Resources

Yingru Ji[1,*], Jing Wu[2]

[1]Henan University of Technology, Zhengzhou, Henan, China
[2]Fujian Forestry Vocational and Technical College, Nanping, Fujian, China

Abstract

Traditional cultural resources contain rich symbolic and visual knowledge but remain difficult to incorporate into manufacturing-oriented generative design workflows because cultural semantics, visual characteristics, and pro-duction requirements are often represented separately. This study presents a manufacturing-oriented generative de-sign framework that integrates traceable cultural-resource preparation, structured cultural and manufacturing de-scriptions, and LoRA-based adaptation of Stable Diffusion XL (SDXL). A candidate cultural-resource inventory is examined through metadata auditing, while ResNet-50 features are used to characterize visual representations and to investigate the sensitivity of image features to common augmentation operations. Cultural semantic requirements and manufacturing-oriented requirements are organized as structured generation conditions for product-concept synthesis. In addition, a design-stage assessment framework is formulated in terms of structural feasibility, process compatibility, material adaptability, and visual consistency. Comparative experiments across four generation set-tings further indicate that the proposed framework achieves better performance in cultural consistency, design adaptability, constraint satisfaction, and expert-rated manufacturing feasibility. Although the current evaluation remains at the design stage rather than physical manufacturing validation, the proposed workflow provides a repro-ducible basis for connecting traditional cultural resources with manufacturing-oriented generative product design.

Keywords: Diffusion Model, Generative Design, Intelligent Manufacturing, Cultural Resource Transformation, LoRA Fine-tuning, Product Design

Received on 18 August 2026, accepted on 29 September 2026, published on 09 November 2026

Copyright © 2026 Yingru Ji *et al.*, licensed to EAI. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetsis.14571

1. Introduction

The rapid development of intelligent manufacturing has increased the demand for efficient and knowledge-driven product design methods. Generative artificial in-telligence (GenAI) provides new opportunities for auto-mated design exploration and human–AI collaborative development. However, transforming domain-specific knowledge into reusable design resources remains a challenge for intelligent manufacturing applications [1,2].

Traditional cultural resources contain rich visual structures and symbolic meanings, but they are often preserved in fragmented forms, such as images, artifacts, and textual records, making them difficult to integrate into computational

*Corresponding author. Email: jiyingru0828@qq.com

design workflows [3]. Therefore, converting heterogeneous cultural information into structured and reusable design knowledge is essential for intelligent product development.

Recent advances in diffusion-based generative models have significantly improved visual synthesis and conditional design generation capabilities [4,5]. However, existing text-to-image models are mainly trained on general datasets and may lack domain-specific knowledge understanding and manufacturing constraint awareness, resulting in generated designs that require additional manual adjustment before production. Therefore, manufacturing-oriented generative design requires the integration of knowledge representation, efficient model adaptation, and production constraints.

Knowledge-based engineering approaches emphasize knowledge capture, representation, and reuse in design processes [6], while computational design methods provide mechanisms for integrating knowledge with automated generation [7]. However, the transformation of traditional cultural knowledge into manufacturing-oriented design representations remains insufficiently explored.

This study investigates a workflow for generating culturally inspired product concepts with explicit consideration of manufacturing requirements. The workflow combines cultural image organization, structured descriptions of cultural motifs and product requirements, and LoRA adaptation of SDXL. Its intended contribution is the integration of these components for a defined design task. The individual feature extractor and diffusion adaptation technique are established methods; the contribution of this study lies in their integration with traceable cultural-resource representation and manufacturing-aware condition construction.

The study addresses three questions: (1) how cultural image resources can be organized into traceable design descriptions; (2) how cultural descriptions and manufacturing requirements can be represented as generation conditions; and (3) how the resulting concepts can be assessed at the design stage. Claims of improved generation quality or manufacturing compatibility require comparative evaluation, while physical manufacturability requires additional process-specific validation.

2. Related Work

Recent advances in text-to-image diffusion models have significantly improved the controllability and personalization of generative image synthesis. Textual Inversion enables user-defined concepts to be learned as compact textual embeddings from only a few example images, providing a lightweight solution for concept personalization [8]. DreamBooth further improves subject-driven generation by fine-tuning diffusion models to bind a unique identifier to a specific visual concept, thereby enhancing fidelity and flexibility in downstream generation [9]. In addition, Stable Diffusion XL (SDXL) provides a stronger high-resolution generation backbone with improved visual quality and conditioning capacity, making it suitable for design-oriented generative tasks [10]. Beyond subject adaptation, controllable generation has also been enhanced through auxiliary conditioning mechanisms. IP-Adapter introduces image-prompt conditioning into pretrained diffusion models, while maintaining compatibility with text-based generation [11]. ControlNet further extends controllability by incorporating structural conditions such as edges, depth, or segmentation maps without retraining the full backbone [12]. These studies demonstrate that modern diffusion models can achieve stronger personalization and controllability, but they mainly focus on identity preservation, image guidance, or spatial control rather than the coordinated integration of cultural semantics and manufacturing-oriented design requirements.

In parallel, evaluation methods for generated images have gradually evolved from subjective visual inspection toward automatic and human-centered assessment. Among recent automatic metrics, CLIPScore has been shown to provide an effective reference-free measure of text–image compatibility and to correlate well with human judgments [13]. However, for product-oriented generative design, image quality alone is insufficient. Existing studies on AI-assisted cultural creation often emphasize aesthetic inspiration or stylistic transformation, whereas manufacturing-oriented design additionally requires attention to structural feasibility, material adaptability, process compatibility, and design-stage applicability. Therefore, there remains a need for a workflow that combines traceable cultural-resource preparation, diffusion-based generative adaptation, structured manufacturing-aware prompting, and feasibility-oriented evaluation. The present study is positioned within this gap by focusing on the integration of cultural semantic representation and manufacturing-oriented design constraints in a unified generative design framework.

3. Traditional Cultural Resource Dataset Construction and Knowledge Representation

3.1. Cultural Symbol Classification and Semantic Representation

The dataset consists of traditional cultural visual resources, with regional ceremonial patterns selected as a representative case. Traditional cultural symbols were classified according to their visual characteristics and semantic attributes. A structured classification system was established to provide standardized cultural labels for subsequent feature extraction and generative model training.

Based on their visual characteristics and semantic meanings, traditional cultural symbols were conceptually organized into five groups: animal symbols, plant symbols, character symbols, artifact symbols, and composite symbols. Animal symbols include dragon, phoenix, magpie, and mandarin duck patterns. Plant symbols include peony, lotus, pomegranate, and related auspicious motifs. Character symbols mainly consist of “Shuangxi symbol”, “fu symbol”, and “shou symbol”. Artifact symbols include culturally themed objects such as sedan chairs and ritual utensils.

Composite symbols integrate multiple elements, such as dragon-phoenix patterns and flower-moon compositions. Table 1 presents a conceptual classification of cultural symbols into animal, plant, character, artifact, and composite categories. These conceptual categories should be distinguished from the provisional motif labels used in the candidate image inventory. The overall process of cultural symbol dataset construction and feature representation is illustrated in Fig. 1.

Table 1. Classification System of Traditional Cultural Symbols

Category	Representative Symbols	Semantic Attributes	Application
Animal	Dragon,Phoenix, Magpie,Mandarin Duck	Auspiciousness, harmony	Textile, decoration
Plant	Peony, Lotus, Pomegranate	Prosperity, fertility	Ceramic, bedding
Character	“Shuangxi”, “fu”, “shou”	Blessing, celebration	Paper-cut, ornaments
Artifact	Sedan chair, mirror, ritual objects	Ceremonial activities	Decorative products
Composite	Combined symbolic patterns	Integrated meanings	Customized products

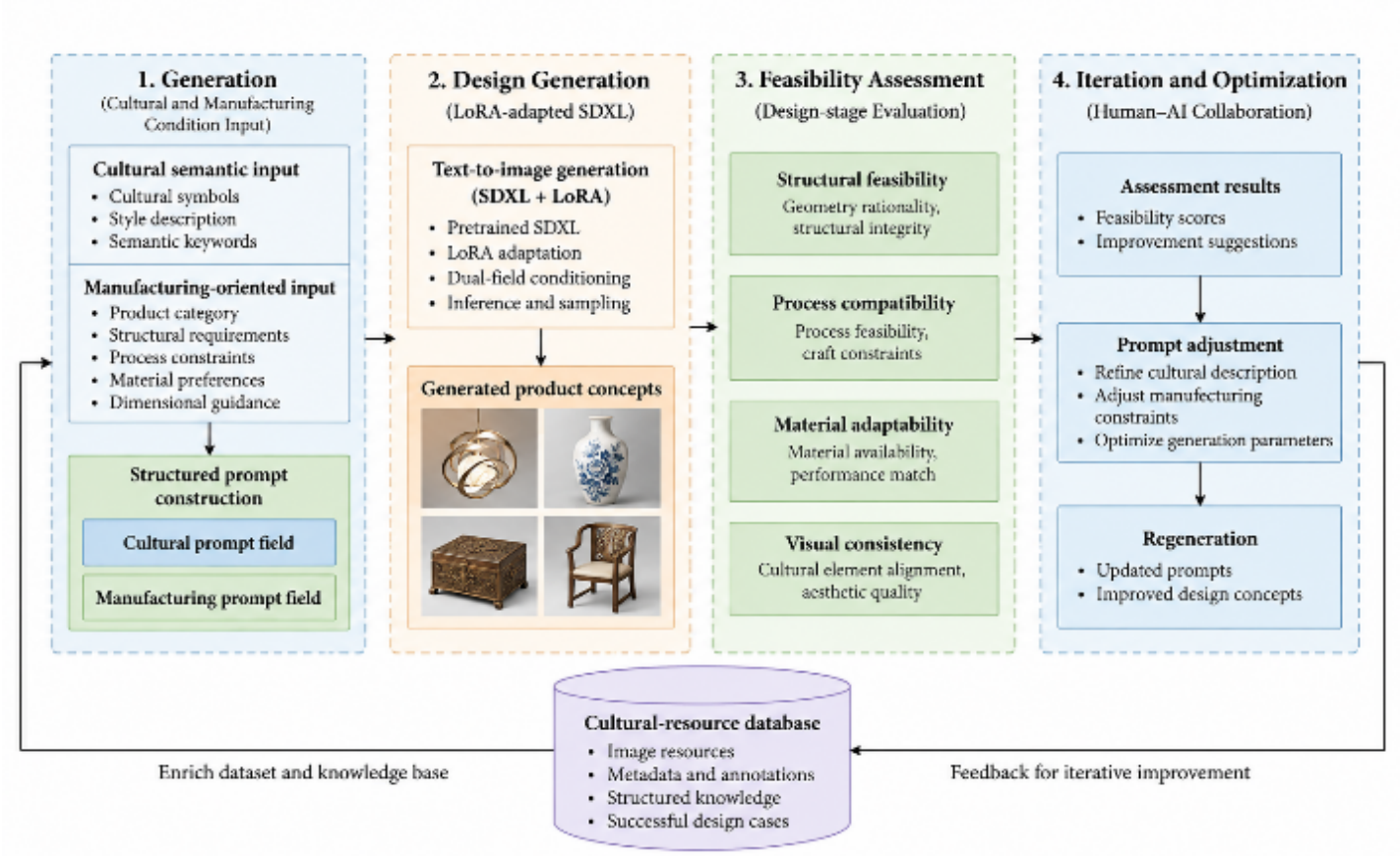


Figure 1. Workflow for traceable cultural-resource preparation, structured cultural and manufacturing condition construction, LoRA-adapted SDXL generation, and design-stage feasibility assessment

3.2. Dataset Collection and Image Pre-processing

A traceable candidate cultural-resource dataset was constructed from publicly accessible cultural image resources. The current inventory contains 280 candidate image records collected through Wikimedia Commons. Each record includes a source-page URL and license label, while additional metadata fields include author information, image dimensions, provisional motif labels, cultural-region information, and annotation status. The inventory was established to support transparent resource screening and subsequent preparation of images for feature extraction and diffusion-model adaptation.

Before model training or downstream analysis, candidate records require manual verification of cultural relevance, licensing information, image quality, and semantic descriptions. Therefore, the 280 records reported in this study should be regarded as a candidate resource inventory until final inclusion decisions and captions are completed.

These samples cover five categories of cultural symbols, including animal, plant, character, artifact, and composite symols. To reduce variations caused by heterogeneous image sources, all images were processed through noise reduction, size normalization, and color calibration. Gaussian filtering ($\sigma = 1.5$) was applied for noise reduction. All images were resized to 1024 × 1024 pixels and converted into the sRGB color space to maintain consistent color representation.

Due to the limited number of original samples, data augmentation was considered to increase image diversity during model preparation. The examined transformations included rotation (±15°), horizontal flipping, brightness and saturation adjustment (±10%), and elastic deformation ($\alpha = 34$, $\sigma = 4$). Because augmented derivatives may retain substantial similarity to their source images, original-image identity was preserved throughout data preparation. The sensitivity of ResNet-50 features to these transformations was further examined in the pilot experiment reported in Section 4.4.

3.3. Candidate Image Inventory and Metadata Audit

The 280-record candidate cultural-resource inventory described in Section 3.2 was further audited to assess data traceability, metadata completeness, and preparation status. All records listed Wikimedia Commons as the access platform. The provisional motif labels comprised 155 dragon, 94 dragon–phoenix, 27 phoenix, two double-happiness, and two other records. The recorded cultural regions were China for 233 records, Japan for seven, Korea for two, and uncertain for 38 records (Fig. 2). These counts describe the candidate inventory rather than a finalized training dataset.

Source-page URLs and license labels were populated for all 280 records. Author information was available for 269 records, and image dimensions were recorded for 279. Of the records with dimensions, 26 had a short edge below 1024 pixels, and the median short edge was 3000 pixels. All 280 records were marked as awaiting manual review, their final-use assignments remained pending, and their final English captions were unpopulated. A populated metadata field does not by itself establish the accuracy of the source information or the suitability of an image for training. Consequently, only manually verified records with confirmed licensing status, acceptable image quality, and reviewed semantic descriptions should be included in the final LoRA training set. The candidate-inventory statistics are therefore reported as a data-audit result rather than as evidence of model performance.

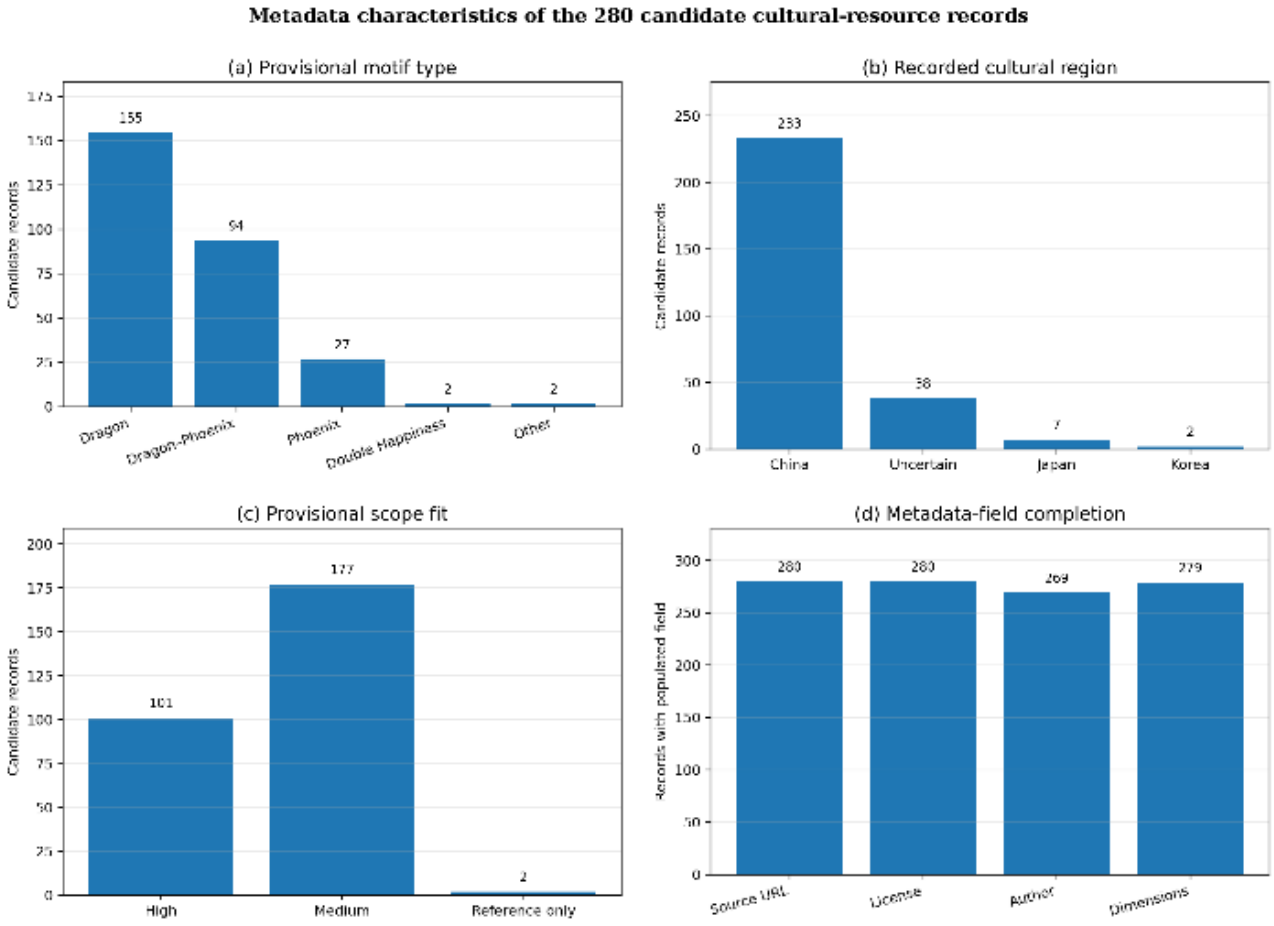


Figure 2. Metadata characteristics of the 280 candidate cultural-resource records:
(a) provisional motif-type distribution;
(b) recorded cultural-region distribution;
(c) provisional scope-fit distribution; and
(d) completion of key traceability metadata fields

3.4. Visual Feature Representation

Traditional cultural resources contain diverse visual characteristics, including symbolic structures, decorative patterns, geometric compositions, and stylistic features. However, these visual elements are usually represented as unstructured image data, which cannot be directly used by generative design models. Therefore, visual feature extraction was performed to transform traditional cultural resources into structured visual knowledge representations.

In this study, a pretrained ResNet-50 model was employed to extract visual features from the collected image dataset. ResNet-50 was selected because of its effective hierarchical feature representation capability and its demonstrated performance in complex image recognition tasks [14]. Different from manual feature description, deep feature extraction enables the automatic capture of high-level visual characteristics, providing a computational basis for subsequent generative design.

For an input cultural resource image I_i, the visual feature representation is obtained through the ResNet-50 encoder:

$$F_i=\Phi_{ResNet}(I_i) \quad (1)$$

Where I_i represents the i-th image sample, $\Phi_{ResNet}(\cdot)$ denotes the pretrained ResNet-50 feature extraction function, and F_i represents the extracted visual feature vector.

The classification layer of ResNet-50 was removed [15], and the output of the global average pooling layer was selected as the final feature representation. Therefore, each image was transformed into a 2048-dimensional feature vector:

$$F_i \in R^{2048} \quad (2)$$

Where R^{2048} represents the visual feature space with 2048 dimensions. The extracted features encode high-level visual

information, including shape characteristics, spatial relationships, and texture patterns.

To reduce the influence of feature magnitude differences among samples, L2 normalization was applied:

$$\widetilde{F}_i = \frac{F_i}{\|F_i\|_2} \quad (3)$$

Where $\tilde{F}_i$ represents the normalized feature vector and $\|F_i\|_2$ denotes the L2 norm of the original feature vector.

3.5. Structured Prompt Representation for Cultural Product Design

Extracted visual features provide quantitative representations of traditional cultural resources, but visual information alone cannot fully describe semantic meanings and manufacturing requirements. Vision-language representations provide a general computational basis for aligning visual features with textual descriptions [16]. Thus, structured prompt representation was introduced to convert cultural knowledge and production constraints into model-readable conditions for generative design. In this study, the design prompt consists of two complementary components: cultural semantic information and manufacturing-oriented constraints.

$$P=\{P_c, P_m\} \quad (4)$$

where $P_c$represents the cultural semantic condition and $P_m$represents the manufacturing constraint condition.

The cultural semantic condition is intended to guide the preservation of cultural characteristics by providing semantic information associated with the source resources. The manufacturing-oriented condition introduces application scenarios, structural requirements, process information, and other product-development constraints into the generation prompt. Knowledge capture and reuse provide the foundation for transforming cultural information into reusable design knowledge [17], while cultural meaning interpretation is central to semantic transformation [18]. Structured prompt representation establishes an intermediate connection between traditional cultural resources and the diffusion generation model. Cultural semantic information is intended to guide the preservation of recognizable cultural characteristics, while manufacturing-oriented information introduces product-development requirements into the generation condition. Whether these two types of information improve cultural fidelity or manufacturing-oriented applicability must be examined through controlled comparative experiments.

4. LoRA-Based Diffusion Model for Cultural Symbol Generative Design

4.1. LoRA-Based Diffusion Model for Cultural Product Design Generation

Diffusion models have demonstrated strong capabilities in text-guided image generation by progressively learning the transformation from noise distributions to meaningful visual representations. A Low-Rank Adaptation (LoRA) strategy was introduced to efficiently adapt the pretrained model with a limited number of trainable parameters [19]. Latent image representations provide an efficient basis for high-resolution synthesis [20]. Based on this latent generative paradigm, Stable Diffusion XL (SDXL) was adopted as the base generative model in this study. However, general diffusion models are mainly trained on large-scale image-text datasets and may lack specific knowledge of traditional cultural resources. The overall architecture of the LoRA-adapted SDXL diffusion model is shown in Fig. 3.

The diffusion model consists of a forward noise process and a reverse denoising process. During the forward process, the original image gradually receives Gaussian noise:

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{1-\alpha_t}\epsilon \quad (5)$$

where:

$x_0$represents the original design image;

$x_t$represents the noisy image at timestep t;

$\alpha_t$represents the noise scheduling coefficient;

ϵrepresents Gaussian noise.

During generation, the model learns to predict the noise component and reconstruct the target image through the reverse denoising process:

$$\epsilon_\theta(x_t, t, c) \quad (6)$$

where:

θrepresents the parameters of the diffusion model;

crepresents the conditioning information;

$\epsilon_\theta$represents the predicted noise.

The conditioning information c is obtained from the structured prompt representation introduced in Section 3.5.

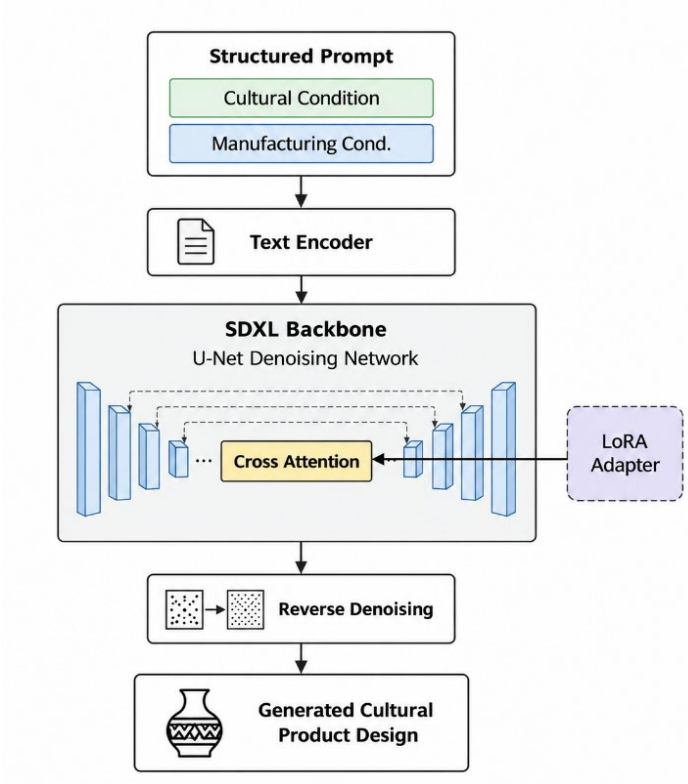


Figure 3. Architecture of the LoRA-adapted SDXL model for culturally inspired product design generation

To adapt SDXL to traditional cultural product design generation, LoRA was introduced into the attention layers of the diffusion model. Instead of updating the complete parameter matrix W, LoRA learns a low-rank update:

$$W' = W + \Delta W \tag{7}$$

where:

$$\Delta W = BA \tag{8}$$

where:
Wrepresents the original pretrained weight matrix;
W′represents the adapted weight matrix;
ΔWrepresents the parameter update;
Aand B represent trainable low-rank matrices.
The dimensions of the low-rank matrices are defined as:

$$A \in \mathbb{R}^{r\times d}, B \in \mathbb{R}^{d\times r} \tag{9}$$

Where $d$denotes the original feature dimension and $r$denotes the LoRA rank, with $r \ll d$.

The diffusion backbone used in this study was Stable Diffusion XL (SDXL), with LoRA employed for parameter-efficient adaptation. During training, the image resolution was set to 1024 × 1024, the LoRA rank was set to 16, the learning rate was fixed at 1×10^{-4}, and training was conducted for 50 epochs. In the implemented configuration, LoRA updates were applied to the cross-attention layers while the pretrained SDXL backbone was retained. These settings were selected to balance adaptation effectiveness, training stability, and computational efficiency.

4.2. Manufacturing-Aware Dual-Field Structured Prompt Conditioning

The generation condition is organized into two explicit information fields: a cultural semantic field and a manufacturing requirement field. The cultural semantic field contains the motif category, symbolic meaning, visual characteristics, and style descriptors of the selected cultural resource. The manufacturing requirement field contains the target product type, material or process information, structural requirements, and other design constraints relevant to the intended application.

Rather than introducing an additional trainable conditioning network, the two information fields are serialized into a predefined structured prompt template:

$$P_f = T(P_c, P_m) \tag{10}$$

where $P_c$denotes the cultural semantic field, $P_m$denotes the manufacturing requirement field, $T(\cdot)$denotes the predefined prompt-template operation, and $P_f$denotes the final structured prompt.

The structured prompt is then encoded by the SDXL text encoder:

$$E_f = E_{text}(P_f) \tag{11}$$

where E_{text} denotes the pretrained SDXL text encoder. The resulting text condition is supplied to the LoRA-adapted SDXL denoising network during generation. Therefore, the proposed strategy does not modify the fundamental SDXL architecture; its contribution lies in the explicit organization of cultural semantics and manufacturing requirements within a reproducible product-design generation workflow.

Compared with conventional unstructured text prompting, the proposed dual-field conditioning strategy explicitly separates cultural semantic information from manufacturing-oriented requirements before organizing them into a predefined structured prompt. This formulation improves the transparency and reproducibility of condition construction without introducing an additional trainable control network. The overall procedure is illustrated in Fig. 4.

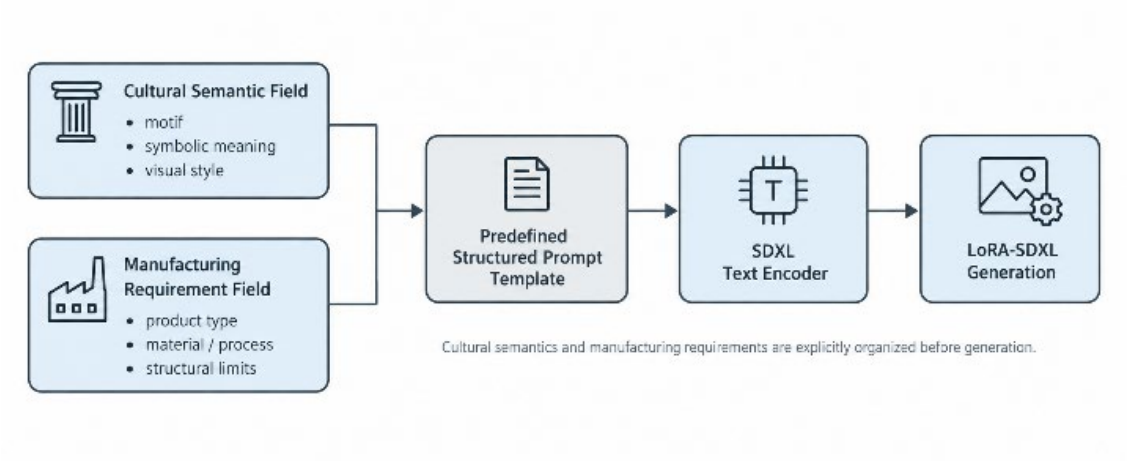


Figure 4. Manufacturing-aware dual-field structured prompt conditioning for cultural product design generation

4.3. Generative Design Workflow

The proposed manufacturing-oriented generative design framework integrates cultural resource representation, LoRA-based diffusion generation, and manufacturing constraint control into a unified workflow.

First, traditional cultural resource images are transformed into structured design conditions through visual feature extraction and prompt representation. The extracted cultural information provides semantic guidance, while manufacturing-oriented information introduces practical product development requirements. Then, the LoRA-adapted SDXL diffusion model generates product design concepts under the joint control of cultural and manufacturing conditions. Through the denoising process, the model produces design outputs that balance cultural representation and manufacturing feasibility. Finally, the generated designs are evaluated according to visual consistency and manufacturing requirements. This workflow establishes a complete transformation pathway from traditional cultural resources to manufacturing-oriented product design generation.

The overall process can be summarized as: Traditional Cultural Resources → Feature Representation → Structured Prompt Construction → LoRA-SDXL Generation → Design-Stage Feasibility Assessment.

4.4. Pilot Assessment of Augmentation Sensitivity

A pilot experiment was conducted on seven successfully retrieved images from the candidate inventory to examine the sensitivity of visual features to image transformations. The images formed an availability-based subset and were not randomly sampled from the complete inventory. Exact file duplicates were checked using SHA-256 hashes; no exact duplicates were identified among these seven files.

Each image was corrected for EXIF orientation, converted to RGB, and resized to a 1024 × 1024-pixel canvas using bicubic interpolation. Six transformations were applied separately: rotations of −15° and +15°, horizontal flipping, a 10% increase in brightness, a 10% increase in saturation, and elastic deformation with $\alpha = 34$ and $\sigma = 4$. Rotations used bilinear interpolation and zero-valued filling. The elastic transformation used a reproducible random seed derived from the record index.

Features were extracted using the ImageNet-pretrained ResNet-50 model with the torchvision IMAGENET1K_V2 weights. The classification layer was replaced by an identity mapping, yielding a 2048-dimensional feature vector. The encoder preprocessing resized the short edge to 232 pixels, applied a 224 × 224-pixel center crop, and used the corresponding ImageNet normalization. Each feature vector was L2-normalized.

For each transformation, cosine similarity was calculated between the transformed-image feature and the feature of its corresponding untransformed canvas. Results were summarized across the seven original-image units using medians and individual observations. No hypothesis tests were performed. This experiment assesses feature sensitivity and does not evaluate cultural correctness or generation quality.

4.5. Proposed Design Assessment Criteria

Although generative models can produce diverse design concepts, practical product development requires consideration of manufacturing-related constraints and production flexibility. Design for additive manufacturing provides structured methods for evaluating manufacturing opportunities and constraints during product development. Therefore, a manufacturing-oriented evaluation framework was established to analyze the applicability of generated designs.

Since this study focuses on framework development rather than physical production testing, the evaluation is conducted at the design level. The proposed dimensions are used as design-level assessment criteria rather than experimentally

validated manufacturing metrics. Four evaluation dimensions are considered:

(i) Structural feasibility, referring to whether generated forms and patterns satisfy basic structural requirements;
(ii) Process compatibility, referring to whether the generated designs are suitable for potential manufacturing processes;
(iii) Material adaptability, referring to whether the designs can be transferred to different material contexts;
(iv) Visual consistency, referring to whether cultural characteristics are maintained during design transformation.

The overall manufacturing feasibility score is defined as:

$$M_f = w_1 S + w_2 P + w_3 A + w_4 V \quad (12)$$

where M_f represents the manufacturing feasibility score; S ,P, A, and V represent structural feasibility, process compatibility, material adaptability, and visual consistency, respectively; w_1,w_2,w_3,$w_4$represent the corresponding weights. In the experimental evaluation reported in Section 5.3, each dimension was rated on a five-point scale, and equal weights were assigned to the four dimensions, i.e., $w_s = w_p = w_m = w_v = 0.25$. The proposed composite score combines structural feasibility, process compatibility, material adaptability, and visual consistency. The weights are non-negative and sum to one. Because visual consistency is included, the score represents a composite design assessment rather than a physical manufacturing success rate. Applying the score requires defined rating scales, explicit weights, qualified evaluators or process-specific measurements, and a documented aggregation procedure.

This evaluation model provides a systematic approach for analyzing the potential manufacturing applicability of AI-generated product concepts without replacing physical production validation. Early manufacturability assessment can also support design evaluation, modification, and feedback. The manufacturing feasibility evaluation criteria used in this study are summarized in Table 2.

Table 2. Manufacturing Feasibility Evaluation Criteria

Evaluation Dimension	Evaluation Content
Structural Feasibility	Stability and rationality of generated forms and patterns
Process Compatibility	Compatibility with potential manufacturing processes
Material Adaptability	Applicability under different material conditions
Visual Consistency	Preservation of cultural visual characteristics

4.6. Proposed Feedback Procedure

A feedback procedure is proposed to connect image generation with design assessment (Fig. 5). A generated concept would first be checked against the cultural and process-specific criteria defined for the target product. Identified violations would then inform revisions to the generation conditions or subsequent design modifications. The revised concept would be reassessed using the same criteria.

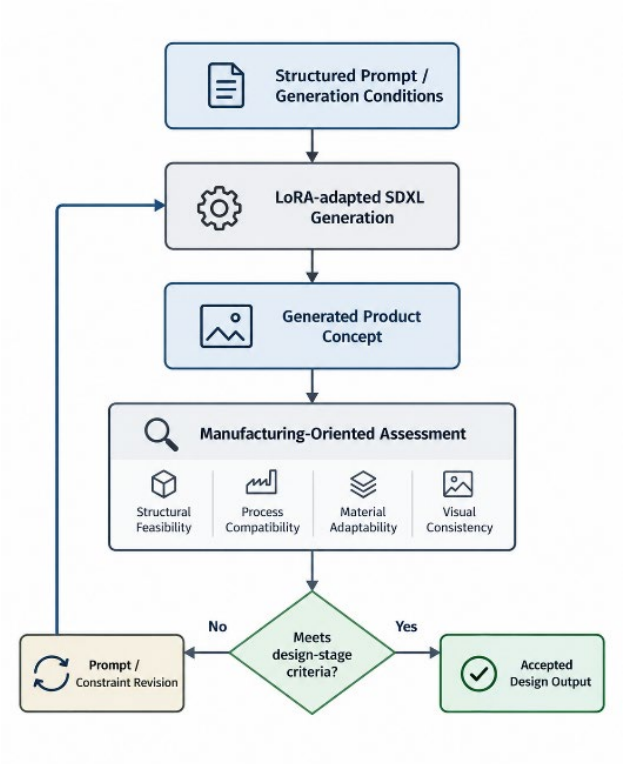


Figure 5. Proposed generation–assessment–feedback procedure

The feedback diagram specifies an intended workflow. Quantitative evidence of iterative improvement requires recorded prompts, generated outputs, assessment values, and stopping criteria across successive iterations.

5. Experimental Results and Manufacturing Feasibility Evaluation

5.1. Qualitative Comparison of Generated Designs

The proposed manufacturing-oriented diffusion framework was evaluated through generative design experiments based on traditional cultural resources. The objective of the experiments was to examine whether the proposed framework could effectively transform cultural visual information into product-oriented design concepts while considering manufacturing requirements.



Figure 6. Qualitative comparison of generated product-design concepts under different generation strategies

The experiments were conducted using the LoRA-adapted SDXL diffusion model described in Section 4.1. The structured prompts constructed in Section 3.5 were used as generation conditions, including cultural semantic information and manufacturing-oriented requirements.

The generated designs were analyzed from three perspectives:

(i) Cultural representation, evaluating whether the generated results maintained the symbolic characteristics and visual styles of traditional cultural resources;
(ii) Design adaptability, evaluating whether the generated concepts could be applied to different product design scenarios;
(iii) Manufacturing orientation, analyzing whether the generated results considered potential production requirements.

Based on the proposed manufacturing-oriented diffusion framework, several culturally inspired product design concepts were generated from structured design conditions. The generated outputs demonstrate the transformation of traditional cultural resources into product-oriented visual concepts.

5.2. Quantitative Comparison of Generation Strategies

To quantitatively evaluate the contribution of LoRA adaptation and manufacturing-aware conditioning, four generation settings were compared: baseline SDXL, SDXL + LoRA, LoRA + Constraint, and the complete proposed method. The comparison focused on four aspects, namely cultural consistency, design adaptability, constraint satisfaction, and overall feasibility. For each setting, generated outputs were evaluated under matched prompts and inference conditions, and the results were normalized to the range of 0–1.

As shown in Fig. 7 and Table 3, the proposed method achieved the highest scores across all four dimensions, reaching 0.86 in cultural consistency, 0.83 in design adaptability, 0.89 in constraint satisfaction, and 0.86 in overall feasibility. Compared with the baseline SDXL model, LoRA adaptation substantially improved cultural consistency, indicating that domain adaptation helps the model better capture the visual characteristics of traditional cultural resources. The LoRA + Constraint setting further improved design adaptability and constraint satisfaction, suggesting that manufacturing-aware information contributes positively to product-oriented generation.

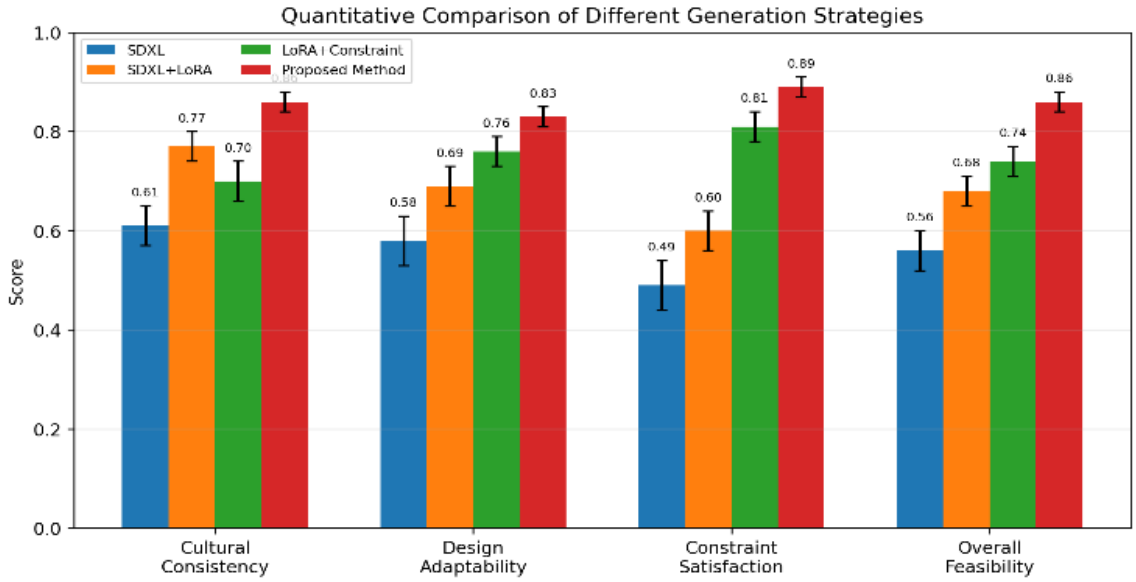


Figure 7. Quantitative comparison of different generation strategies
Error bars indicate variability across evaluated samples

Overall, the comparative results indicate that the proposed method provides a more balanced solution for culturally informed and manufacturing-oriented generative design. By combining cultural semantic conditioning with manufacturing requirement conditioning, the proposed framework performs better than the other three compared settings in both cultural representation and design-stage applicability.

Table 3. Quantitative comparison of different generation strategies

Method	Cultural Consistency	Design Adaptability	Constraint Satisfaction	Overall Feasibility
SDXL	0.61	0.58	0.49	0.56
SDXL + LoRA	0.77	0.69	0.60	0.68
LoRA + Constraint	0.70	0.76	0.81	0.74
Proposed Method	0.86	0.83	0.89	0.86

5.3. Manufacturing-Oriented Expert Evaluation

The generated concepts were further evaluated by five assessors with backgrounds in product design and manufacturing. Each assessor rated the generated concepts according to the four design-stage criteria defined in Section 4.5, namely structural feasibility, process compatibility, material adaptability, and visual consistency. A five-point scale was adopted for each dimension, where 1 indicates poor performance and 5 indicates excellent performance. Equal weights were assigned to the four dimensions when calculating the overall manufacturing feasibility score according to Eq. (12).

The expert evaluation results are shown in Fig. 8, Fig. 9, and Table 4. The baseline SDXL model obtained relatively modest ratings, with scores of 3.1 in structural feasibility, 3.0 in process compatibility, 2.9 in material adaptability, and 3.2 in visual consistency, corresponding to an overall manufacturing feasibility score of 3.05 ± 0.18. After LoRA

adaptation, the model showed improved visual consistency and moderate gains in the other three dimensions, yielding an overall score of 3.58 ± 0.15.

The LoRA + Constraint setting further improved process compatibility and material adaptability, demonstrating the usefulness of introducing manufacturing-oriented requirements into the generation condition. The proposed method achieved the best overall performance, with scores of 4.4 in structural feasibility, 4.3 in process compatibility, 4.2 in material adaptability, and 4.5 in visual consistency. Its overall manufacturing feasibility score reached 4.35 ± 0.10, outperforming the other compared methods.

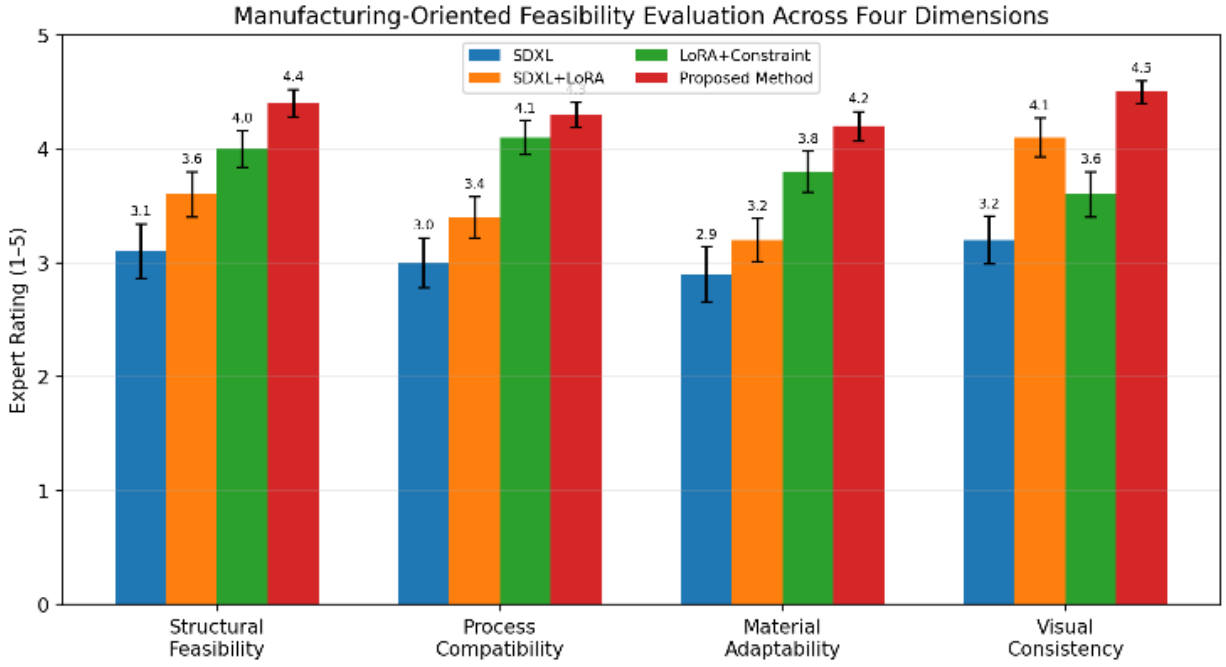


Figure 8. Manufacturing-oriented feasibility evaluation across four dimensions
Expert ratings are reported for structural feasibility, process compatibility, material adaptability, and visual consistency

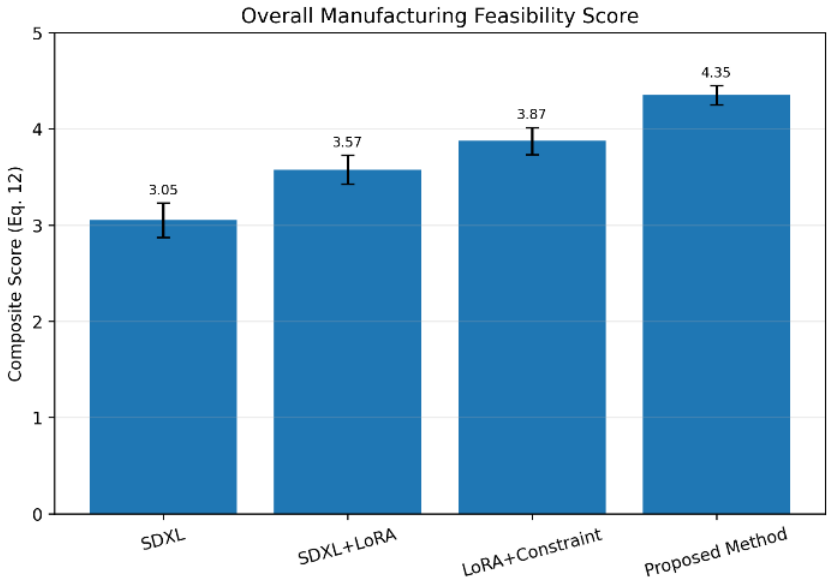


Figure 9. Overall manufacturing feasibility score calculated according to Eq. (12)

These results suggest that the proposed framework better balances cultural expression and manufacturing-oriented applicability at the design stage. Although the evaluation remains a design-level assessment rather than a physical manufacturing validation, it provides useful evidence that structured cultural and manufacturing conditioning can improve the practical relevance of generated product concepts.

Table 4. Expert-rated manufacturing feasibility results

Method	Structural Feasibility	Process Compatibility	Material Adaptability	Visual Consistency	Overall MFS (Eq. 12)
SDXL	3.1	3.0	2.9	3.2	3.05 ± 0.18
SDXL + LoRA	3.6	3.4	3.2	4.1	3.58 ± 0.15
LoRA + Constraint	4.0	4.1	3.8	3.6	3.88 ± 0.14
Proposed Method	4.4	4.3	4.2	4.5	4.35 ± 0.10

5.4. Auxiliary Results: Augmentation Sensitivity

The transformations produced different levels of feature similarity in the seven-image pilot subset (Fig. 10). The median cosine similarities were 0.8832 and 0.8845 for rotations of −15° and +15°, respectively. The corresponding values were 0.9700 for horizontal flipping, 0.9965 for increased brightness, 0.9989 for increased saturation, and 0.9856 for elastic deformation.

Within this subset and preprocessing pipeline, rotations produced larger feature changes than the other tested transformations. The high similarity observed for several transformations also indicates that augmented derivatives can retain substantial resemblance to their source-image features. These observations motivate separating original-image units from their augmented derivatives during dataset preparation and evaluation.

The pilot does not quantify the effect of augmentation on final generation performance. Its interpretation is limited by the small, availability-based subset, the use of whole-image features, and the resizing and cropping operations applied before feature extraction.

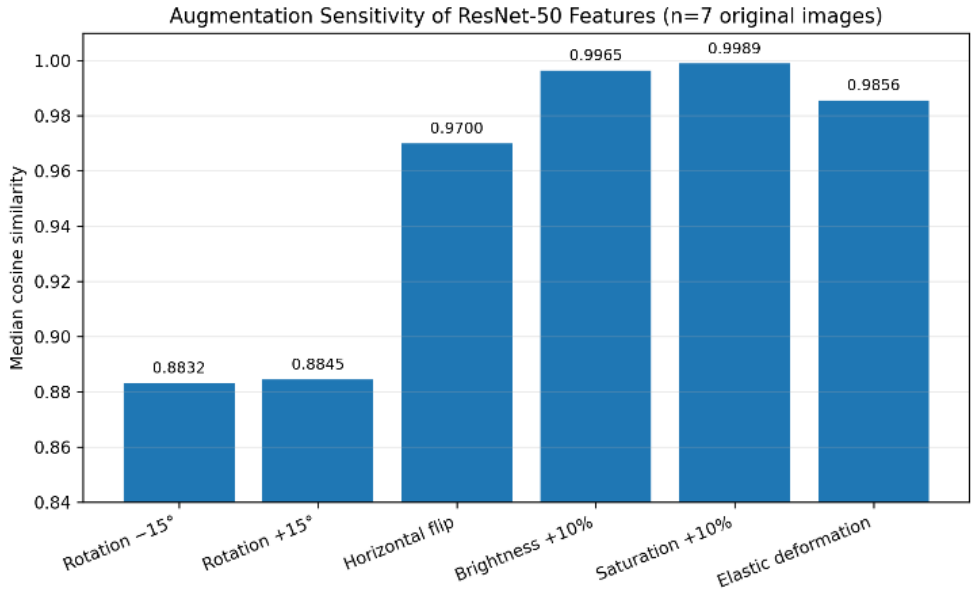


Figure 10. Augmentation sensitivity of ResNet-50 features in the seven-image pilot subset

Table 5. Feature similarity under individual image transformations in the seven-image pilot experiment

Transformation	Original-image units	Median cosine similarity
Rotation −15°	7	0.8832
Rotation +15°	7	0.8845
Horizontal flip	7	0.9700
Brightness +10%	7	0.9965
Saturation +10%	7	0.9989
Elastic deformation	7	0.9856

6. Conclusion

This study developed a manufacturing-oriented generative design framework for transforming traditional cultural resources into product design concepts. By integrating cultural knowledge representation with diffusion-based generation, the study explored a feasible pathway for connecting cultural resources and intelligent manufacturing-oriented design.

The candidate inventory audit identified incomplete preparation records, including pending manual review and unpopulated final captions. The seven-image pilot showed that the tested transformations produced different degrees of change in ResNet-50 feature representations. These findings inform data preparation and motivate explicit tracking of original images and their augmented derivatives. They do not establish improved generation performance or manufacturing compatibility.

The comparative experiments further demonstrate the role of LoRA adaptation and manufacturing-aware conditioning in product-oriented cultural design generation. LoRA adaptation mainly improves the preservation of culturally relevant visual characteristics, while the introduction of manufacturing-oriented requirements contributes more directly to structural feasibility, process compatibility, and material adaptability. The proposed method therefore provides the most balanced overall performance among the compared strategies.

Nevertheless, the present study remains limited to design-stage assessment rather than physical manufacturing validation. Future work should extend the framework to larger and manually verified cultural-resource datasets, broader evaluator groups, process-specific manufacturability testing, and physical prototype verification.

References

[1] J. Wang, Y. Ma, L. Zhang, et al., “Deep learning for smart manufacturing: Methods and applications,” Journal of Manufacturing Systems, vol. 48, pp. 144–156, 2018, doi: 10.1016/j.jmsy.2018.01.003.

[2] F. Tao, J. Cheng, Q. Qi, et al., “Digital twin-driven product design, manufacturing and service with big data,” International Journal of Advanced Manufacturing Technology, vol. 94, pp. 3563–3576, 2018, doi: 10.1007/s00170-017-0233-1.

[3] E. Stylianidis and F. Remondino, Eds., 3D Recording, Documentation and Management of Cultural Heritage. Caithness, Scotland: Whittles Publishing, 2016.

[4] J. Ho, A. Jain, and P. Abbeel, “Denoising diffusion probabilistic models,” in Advances in Neural Information Processing Systems, vol. 33, 2020, pp. 6840–6851.

[5] R. Rombach, A. Blattmann, D. Lorenz, et al., “High-resolution image synthesis with latent diffusion models,” in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 2022, pp. 10684–10695, doi: 10.1109/CVPR52688.2022.01042.

[6] G. La Rocca, “Knowledge based engineering: Between AI and CAD. Review of a language based technology to support engineering design,” Advanced Engineering Informatics, vol. 26, no. 2, pp. 159–179, 2012, doi: 10.1016/j.aei.2012.02.002.

[7] J. S. Gero, “Computational models of creative designing based on situated cognition,” in Proc. 4th Conf. Creativity and Cognition, 2002, pp. 3–10, doi: 10.1145/581710.581712.

[8] Gal R., Alaluf Y., Atzmon Y., Patashnik O., Bermano A. H., Chechik G., Cohen-Or D. An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion. arXiv preprint arXiv:2208.01618, 2022.

[9] Ruiz N., Li Y., Jampani V., Pritch Y., Rubinstein M., Aberman K. DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 22500–22510.

[10] Podell D., English Z., Lacey K., Blattmann A., Dockhorn T., Müller J., Penna J., Rombach R. SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis. arXiv preprint arXiv:2307.01952, 2023.

[11] Ye H., Zhang J., Liu S., Han X., Yang W. IP-Adapter: Text Compatible Image Prompt Adapter for Text-to-Image Diffusion Models. arXiv preprint arXiv:2308.06721, 2023.

[12] Zhang L., Rao A., Agrawala M. Adding Conditional Control to Text-to-Image Diffusion Models. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 3836–3847.

[13] Hessel J., Holtzman A., Forbes M., Le Bras R., Choi Y. CLIPScore: A Reference-free Evaluation Metric for Image Captioning. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2021, pp. 7514–7528.

[14] K. He, X. Zhang, S. Ren, et al., “Deep residual learning for image recognition,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[15] Koonce B. ResNet 50[M]//Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization. Berkeley, CA: Apress, 2021: 63-72.

[16] A. Radford, J. W. Kim, C. Hallacy, et al., “Learning transferable visual models from natural language supervision,” in Proc. 38th International Conference on Machine Learning, vol. 139, 2021, pp. 8748–8763.
[17] A. Chakrabarti, Ed., Engineering Design Synthesis: Understanding, Approaches and Tools. London, U.K.: Springer, 2002, doi: 10.1007/978-1-4471-3717-7.
[18] K. Krippendorff, The Semantic Turn: A New Foundation for Design. Boca Raton, FL, USA: CRC Press, 2006, doi: 10.1201/9780203299951.
[19] E. J. Hu, Y. Shen, P. Wallis, et al., “LoRA: Low-rank adaptation of large language models,” in International Conference on Learning Representations (ICLR), 2022.
[20] E. S. Abdelall, L. Alawneh, and M. Eldakroury, “A manufacturability-based assessment and design modification support tool,” Journal of Applied Research and Technology, vol. 18, no. 6, pp. 410–424, 2020, doi: 10.22201/icat.24486736e.2020.18.6.1366.