



Watch and Transformation of Traditional Carving Skills in the Era of AI

Xiao Li^{1,*}

¹ School of Arts & Crafts, Suzhou Art & Design Technology Institute, Suzhou 215104 Jiangsu, China

SUMMARY: *Under the background of the deep integration of artificial intelligence into the protection, dissemination and reuse of cultural heritage, it is urgent to move from empirical inheritance to data-based, intelligent and interactive expression of traditional carving skills. This paper constructs a multi-modal data system that integrates images, videos, 3D point clouds, text and oral data. Combined with deep visual coding, cross-modal knowledge representation, knowledge graph constraints, conditional generative pattern reconstruction, style transfer and intelligent interactive feedback mechanism, this paper proposes a technical path for traditional sculpture maintenance and transformation. The experimental results show that the accuracy of the proposed method in carving category recognition and knife discrimination reaches 90.6% and 88.9%, respectively. In the generation task, SSIM, PSNR, style consistency and semantic integrity reach 0.893, 30.4 dB, 0.901 and 0.918, respectively. The system also improves the scores of pattern recognition, knife understanding and style judgment by 12.8, 13.2 and 12.9 points respectively, and maintains the recognition accuracy of 80.9% under the condition of 40% noise or occlusion, which reflects good cross-scene adaptation ability and robustness. It provides more operable technical support for the digital activation, accurate inheritance and innovative transformation of traditional carving skills.*

KEYWORDS: *Artificial intelligence; Traditional carving skills; Multi-modal feature representation; Digital inheritance*

1 Introduction

With the accelerated entry of artificial intelligence technology into the field of cultural production and knowledge dissemination, the preservation mode, transmission path and innovation mechanism of traditional crafts are undergoing profound changes [1]. As a compound cultural heritage with material form, operational experience and aesthetic norms, carving skills have long relied on teacher-student inheritance, word-to-word instruction and on-site demonstration to complete the continuation of knowledge. Its core is not only reflected in the patterns of finished products and the appearance of objects, but also in the judgment of material selection, knife rhythm, hierarchical processing, pattern layout and technological context [2, 3]. With the reduction of the number of old artists, the long training cycle of systematic apprentices, and the dispersion of regional technique data, many carving experiences with local characteristics are faced with practical difficulties such as incomplete records, non-standard expressions, and limited transmission radius [4]. How to use artificial intelligence to achieve high-quality retention, structured analysis and creative transformation without weakening the value of art ontology has become an important issue in the research of

*lixiyao2026@163.com

<https://doi.org/10.65102/is2026729>

traditional sculpture and the construction of digital culture [5].

From the perspective of technology development trends, computer vision, 3D scanning, multi-modal feature learning, knowledge graph and generative artificial intelligence provide new method support for the digital expression of traditional carving skills [6]. High-precision image acquisition and point cloud reconstruction can completely preserve the texture details, spatial hierarchy and knife mark characteristics of the carving object. The deep learning model can extract morphological features, style features and process features from a large number of carving images and videos, and establish the mechanism of category recognition, pattern matching and technique analysis. Natural language processing and knowledge modeling techniques can help to transform the tacit knowledge scattered in oral experience, process documents and case descriptions into searchable and related knowledge units. Generative models further expand the application space of traditional pattern reorganization, modeling aided design and cultural creative translation, making "watch" and "transformation" no longer stay at the static archive level, but gradually enter a new stage of computable, interactive and extensible [7, 8].

However, the introduction of artificial intelligence into traditional sculpting does not simply replace experience with algorithms. Carving skills have distinct regional characteristics, material dependence and manual randomness. There are significant differences in the presentation ways of the same ornament on different wood, stone or jade materials, and the same knife method will also form unique strength trajectories and aesthetic orientations in the hands of different artists. This means that traditional carving data often have the characteristics of limited sample size, fuzzy category boundaries, label dependence on expert judgment and subtle style differences, etc. Directly applying general visual recognition or image generation models is easy to cause loss of technique semantics, misreading of cultural symbols and superficial generation results [9, 10]. Therefore, artificial intelligence research for traditional carving needs to pay attention to cultural authenticity, process interpretability and technical adaptability at the same time, and form a research path more suitable for process attributes at the levels of data construction, feature representation, model design and evaluation system [11-13].

Based on this, this paper focuses on the watch and transformation of traditional carving skills in the AI era, focusing on the organization of traditional carving multimodal data, the intelligent extraction method of carving morphology and knife characteristics, the generation mechanism for pattern reconstruction and style transfer, and the intelligent interactive application path serving for inheritance teaching and innovative design. It strives to establish a more robust connection between cultural protection and technological innovation, and provides landing research ideas for the digital inheritance, intelligent analysis and modern transformation of traditional carving skills.

2 Related work

In recent years, the application of artificial intelligence in the digital protection of cultural heritage has been deepened, and the research perspective has gradually shifted from single information collection to the collaborative promotion of recognition, modeling, generation and knowledge organization. Focusing on OCR post-processing in the digitization of Chinese historical texts, Lee et al. proposed a line segmentation and reading order sorting algorithm based on adjacent character detection, indicating that the digitization of cultural heritage has entered the stage of structured understanding from static scanning [14]. Gattiglia summarized the application boundary, governance logic and method integration of artificial intelligence in

the management of archaeology, and emphasized that AI has the dual value of tool extension and research paradigm reshaping in heritage research [15]. Corradetti et al. used large-scale multimodal models for pattern recognition and benchmarking of stone cultural relics deterioration, which promoted the transformation of heritage disease interpretation from experience driven to intelligent discrimination [16]. Malchiodi et al. compared the differences between one-class and two-class machine learning in the classification of chemical elements of ceramic samples, which provided a method reference for cultural heritage identification under the condition of small samples and weak labels [17].

In the aspect of heritage form analysis and visual modeling, Pang et al. realized heritage architectural component recognition and modeling based on local features, indicating that fine-grained component feature extraction has become an important path for automatic analysis of complex heritage objects [18]. Crabu et al. used photometric stereo method under unknown lighting conditions to reconstruct the color of rock paintings, which enhanced the visibility of carving surface details and historical information [19]. Vuoto et al. further reconstructed the classical temple through generative modeling, demonstrating the feasibility of heritage form representation and virtual transformation driven by existing data [20]. Mezina et al. proposed EnsArtNet integrated neural network to recognize painting style, indicating that the computational representation of artistic style has moved from shallow classification to more stable ensemble discrimination [21]. Li et al. took Terracotta Warriors as an example to carry out expert knowledge visualization research, which promoted the extension of cultural heritage knowledge from textual experience to graphical and interpretable expression [22].

At the level of heritage state assessment and protection decision-making, Ramirez-Arellano et al. combined Deng entropy and LSTM to classify heavily damaged architectural heritage, reflecting the advantages of temporal model in complex damage discrimination [23]. Ottoni et al. implemented the classification of cultural heritage buildings based on transfer learning and data augmentation, indicating that deep learning still has strong applicability in scenarios with limited samples [24]. Colace et al. summarized the new AI challenges faced by cultural heritage protection from the overall level, and pointed out that credibility, interpretability and application norms have become key issues [25]. At the same time, Kloiber et al. conducted a comparative study on the time cost of traditional manual cutting wood for repair, reminding researchers that they still need to respect the process labor ontology and operation logic when introducing intelligent technology [26]. Through the re-study of lacquer art objects in the Northern Wei Dynasty, Zhu et al. revealed the deep correlation between traditional process materials, processes and historical context, which provided enlightenment for the fine modeling of process knowledge [27]. In general, the existing research has laid a foundation for the digital acquisition, visual recognition, generation and reconstruction and knowledge organization of traditional carving skills. However, the research on the collaborative expression of carving tool method, operation process and implicit experience is still insufficient, which also constitutes the entry point for further discussion in this paper.

3 Watch and transformation methods of traditional carving skills in the era of AI

3.1 Multimodal data construction and knowledge representation of traditional carving techniques

The digital modeling of traditional carving skills should not stop at the image acquisition level

of finished products, but should incorporate the shape information, technological process information, cultural semantic information and inheritance experience information of the carving object into a unified data organization framework. In terms of the research object, traditional carving includes not only the pattern structure of the surface of the vessel, the relief level, the trend of the knife mark and the texture of the material, but also the dynamic operation characteristics of the lifting, turning, collecting and pressing of the knife in the process of carving, but also the knowledge elements such as the implication of the subject matter, the regional style, the process lineage and the relationship between teachers and students. A single modality can only record a part of the content, and it is difficult to fully express the process ontology of the traditional sculpture coupling of "form-technology-meaning". Therefore, it is necessary to build a multi-modal data system oriented to images, videos, 3D point clouds, text records and oral data, and form a computable, associable and searchable knowledge representation model on this basis.

As shown in Figure 1, the multimodal data and knowledge representation framework of traditional carving techniques can be divided into data acquisition layer, preprocessing layer, feature representation layer, knowledge association layer and application output layer. The data acquisition layer mainly obtains high-definition images of physical objects, process demonstration videos, 3D scanning point clouds, process literature texts and oral materials of inheritor. The preprocessing layer completed image denoising, video key frame extraction, point cloud registration, text segmentation and term standardization. The feature representation layer is responsible for extracting the contour of the pattern, the surface texture, the knife mark rhythm, the spatial fluctuation and the semantic label. The knowledge association layer connected the "pattern type, cutting characteristics, material properties, process steps and cultural meanings" through entity relationship modeling. The application output layer provides support for style recognition, technique analysis, pattern generation, assistant teaching and digital display. The key of this structure is not to simply stack data, but to establish the semantic alignment relationship between different modalities, so that visual information, action information and text knowledge can form a collaborative expression in a unified space.

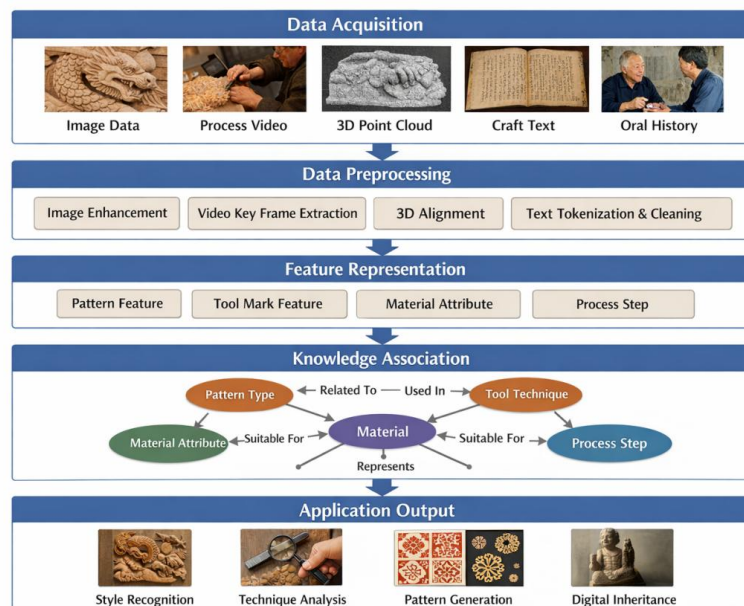


Figure 1: Framework diagram of multimodal data and knowledge representation of traditional carving techniques

In the multi-modal feature mapping stage, let the MTH modal input be $x_i^{(m)}$, and its original features can be mapped to the shared representation space after being extracted by the corresponding encoder $\phi_m(\cdot)$:

$$z_i^{(m)} = W_m \phi_m(x_i^{(m)}) + b_m \quad (1)$$

Here, W_m and b_m denote the projection parameters of the m -th mode, respectively. The function of this formula is to compress image texture features, video action features, point cloud geometric features and text semantic features into a unified dimension, which lays a foundation for subsequent cross-modal matching and fusion. Due to the obvious differences in sampling scale, information density and noise sources of traditional sculpture data, the use of unified projection can weaken the problem of dimensional inconsistency between modalities and improve the stability of knowledge representation.

In the process of cross-modal alignment, only completing dimension mapping is still not enough to ensure semantic consistency, so it is necessary to further constrain the representation proximity of the same carving sample in different modalities. The alignment loss between vision, text and action can be defined as follows:

$$\mathcal{L}_{align} = - \sum_{i=1}^N \log \frac{\exp(\text{sim}(z_i^{(v)}, z_i^{(t)})/\tau)}{\sum_{j=1}^N \exp(\text{sim}(z_i^{(v)}, z_j^{(t)})/\tau)} \quad (2)$$

Here, $z_i^{(v)}$ represents the visual modal representation, $z_i^{(t)}$ represents the textual modal representation, $\text{sim}(\cdot)$ is the similarity function, and τ is the temperature coefficient. The loss term can strengthen the association between the same pattern image, the same process description and the same knife demonstration, so that the system can not only recognize the "what" of the carving, but also understand the "how" and "what" of the carving.

In the knowledge association layer, the traditional carving technique can be expressed as a triplet set (h,r,t) , where h is the head entity, r is the relationship, and t is the tail entity, such as "curled grass grain - corresponding - rotary knife method" and "Nanmu material - suitable - bas-relief process". In order to measure the plausibility of a relationship in a knowledge graph, the following scoring function can be used:

$$f(h, r, t) = \|e_h + e_r - e_t\|_2^2 \quad (3)$$

Here, e_h , e_r and e_t represent the embedding vectors of head entity, relation and tail entity, respectively. The smaller the score value, the more stable this knowledge relationship is. In this way, technical knowledge scattered in images, texts, and oral experiences can be compressed into structured semantic networks to support subsequent retrieval, reasoning, and generation applications.

In general, the core of the multi-modal data construction of traditional carving technology revolves around the coupling relationship between carving objects, technological processes and cultural semantics, and establishes a perceptual, computable and interpretable knowledge representation system. This method not only provides high-quality input for subsequent carving form recognition and knife discrimination, but also provides a unified data basis for traditional pattern generation, style transfer and digital inheritance system construction.

3.2 Depth visual coding method for carving shape and knife feature

The effective information in the traditional carving image is not only reflected in the external

contour, but also concentrated in the turning relationship of the edge of the pattern, the depth change of the relief level, the dense organization of the knife mark texture and the fluctuation rhythm of the local area. If only a single convolution path is used to uniformly encode the whole image, it is easy to submerge the subtle knife differences of the carving surface into the large-scale morphological features, which weakens the ability of the model to recognize the process style and operation habit. Based on this, we construct a deep visual coding network for carving morphology and knife features, which integrates the overall shape, local texture, edge structure and hierarchical relief into a unified framework. Through multi-branch parallel extraction and hierarchical fusion, we form a high-quality representation that takes into account both appearance morphology and process details.

As shown in Figure 2, the carving form and knife feature extraction network is composed of an input reconstruction layer, a local texture coding branch, an edge contour enhancement branch, a relief depth perception branch, a multi-scale feature fusion layer, an attention selection layer, and a technique discrimination output layer. After the input image enters the network, size normalization, brightness balance and background suppression are completed to reduce the interference of non-process factors on the recognition results. Then, the local texture branch focuses on extracting knife details, material texture and Nick direction information. The edge contour branch highlights the pattern boundary, rotation Angle change and composition skeleton. The depth branch of relief depicts the spatial differences among the shallow, relief, and through-relief through the shadow gradient and local relief response. The three types of features enter the fusion module after the middle layer is aligned, so that the network can not only grasp the macroscopic modeling structures such as dragon grain, circled grain, and moire, but also distinguish the local surface features formed by different knife methods such as flat knife, round knife, and oblique knife.

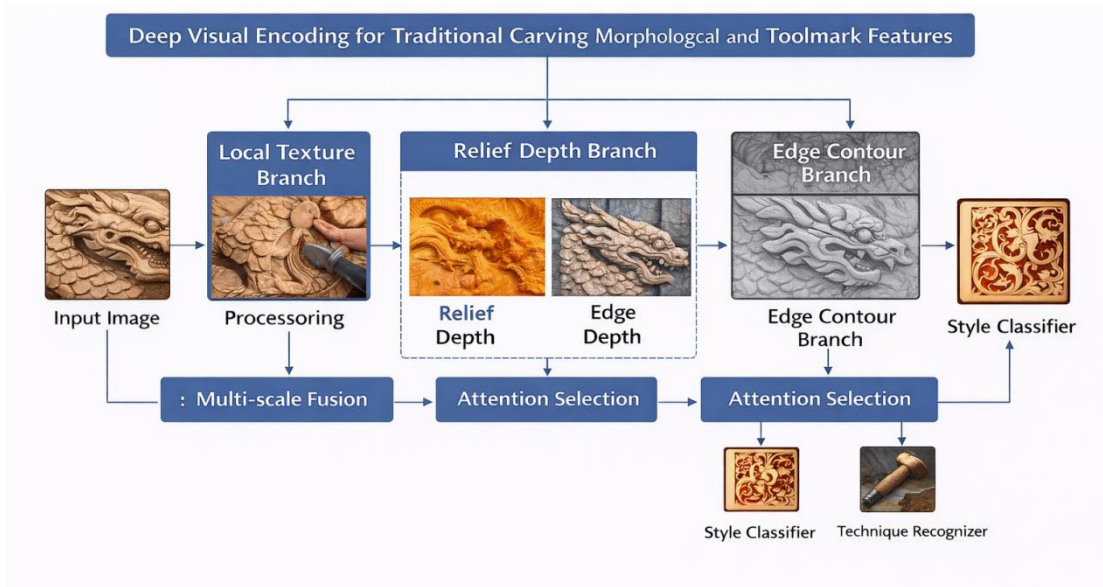


Figure 2: Network structure diagram of carving morphology and knife method feature extraction

In the local texture coding stage, let the preprocessed input image be $X \in \mathbb{R}^{H \times W \times C}$, and the convolution response of the s -th scale branch can be expressed as follows:

$$F_s = \delta(\text{BN}(W_s * X + b_s)) \quad (4)$$

Here, W_s represents the s -th scale convolution kernel parameter, $*$ represents the convolution operation, $\text{BN}(\cdot)$ is the batch normalization, and $\delta(\cdot)$ is the nonlinear activation function. The function of this formula is that different receptive fields are used to capture fine-grained knife texture and large-scale pattern structure respectively, so that the model has multi-scale perception ability in the shallow layer. For traditional carving, local knife marks have obvious directionality and rhythm, and the pattern composition is reflected by stable morphological layout. Therefore, multi-scale convolution is more suitable for expressing the superposition relationship of these two types of information than single scale.

In order to highlight the importance of contour edges and morphological turning regions, we further introduce edge enhancement response. Let the intermediate feature before fusion be U , and its edge-sensitive weight be defined as follows:

$$M_e = \sigma(W_e \cdot \nabla U + b_e) \quad (5)$$

Here, ∇U represents the gradient response of the feature map, W_e and b_e are learnable parameters, and $\sigma(\cdot)$ is the Sigmoid function. The weight map can give higher response to key areas such as grain boundaries, cutting edges and hierarchical sections, so as to suppress the interference of background wood grain, shooting noise and non-sculptural textures. For the traditional carving image, the edge is the result of the process action, such as the turning point formed by the collapsing of the knife edge, the arc boundary formed by the transition of the round knife, etc., which can be strengthened by the weight mechanism.

After multi-branch fusion, in order to realize the joint discrimination of carving categories and tool types, this paper adopts the output mode of combining global convergence and classification mapping. Let the final fused feature be F^* , then its class prediction vector can be written as follows:

$$P = \text{Softmax}(W_c \cdot \text{GAP}(F^*) + b_c) \quad (6)$$

where $\text{GAP}(\cdot)$ represents the global average pooling, W_c and b_c are the classification layer parameters, and P is the class probability distribution. Compared with the direct use of fully connected high-dimensional mapping, this method can compress redundant parameters while retaining the main structural information, and improve the generalization ability of the network on small sample process images. The final output can be used for both pattern category recognition and knife label discrimination, which provides a visual basis for subsequent style analysis, pattern generation and process teaching feedback.

Table 1 shows the carving visual feature encoding module configuration. It can be seen from the table that different modules are not simply connected in series, but spread out around the logic of "form recognition - knife characterization - hierarchical analysis - feature fusion - joint discrimination", which can better meet the feature expression requirements of the interwoven structure information and process information in traditional carving images.

Table 1: Carving visual feature encoding module configuration Table

Module Name	Input Content	Output Content	Main Function
Input Reconstruction Module	Raw carving images	Standardized input tensor	Unifies image size, performs brightness correction, and suppresses background interference
Local Texture Encoding Branch	Standardized images	Local texture features	Extracts chisel-mark directions, material textures, and fine carving detail responses
Edge Contour Enhancement Branch	Standardized images and shallow features	Contour edge features	Enhances pattern boundaries, turning structures, and compositional skeleton information
Relief Depth Perception Branch	Standardized images and grayscale gradients	Relief depth features	Characterizes spatial depth differences among shallow carving, relief carving, and openwork carving
Multi-scale Fusion Module	Multi-branch feature set	Integrated visual representation	Aligns and fuses features from different scales and branches
Attention Selection Module	Integrated visual representation	Key-region enhanced features	Adaptively highlights important pattern regions and key carving-technique responses
Joint Classification Module	Enhanced feature vectors	Style categories and technique labels	Enables joint output for carving form recognition and carving-technique classification

3.3 Sculptural Pattern Reconstruction and Style Transfer Mechanism Based on generative AI

In the process of digital transformation of traditional carving skills, pattern reconstruction and style transfer are not mechanical copies of existing patterns, but realize the completion of incomplete patterns, the adaptation of heterogeneous materials, and the translation of modern design scenes on the basis of respecting traditional composition rules, knife semantics and cultural meanings. Compared with general image generation tasks, carved patterns have the characteristics of strong symmetry, clear boundaries, stable hierarchical rhythm, many local repetitive units, and clear style constraints. Therefore, based on the above multi-modal data representation and deep visual coding results, this paper constructs a conditional generation mechanism for carving patterns, which introduces morphological features, craft style features and knowledge constraints into the generation process, so that the system has the ability of "completing", "turning into decent" and "becoming available" at the same time.

As shown in Figure 3, the carving pattern generation and style transfer process mainly includes six links: incomplete pattern input, structural semantic analysis, style condition encoding, generation and reconstruction, constraint optimization and result output. For damaged or incomplete traditional patterns, the system first uses the coding network to extract the contour skeleton, local decorative units and pattern arrangement rhythm, and then combines the pattern types, composition rules and typical knife information in the knowledge graph to form a condition vector. On this basis, the generator completes the missing region,

and injects the target carving style into the output result through the style mapping module, so that the generated drawings have clear regional aesthetic and technical characteristics while maintaining the continuity of the original structure. This process is not only suitable for the digital restoration of traditional wood carving, stone carving or brick carving patterns, but also can be used for pattern reorganization and style translation in cultural and creative design.

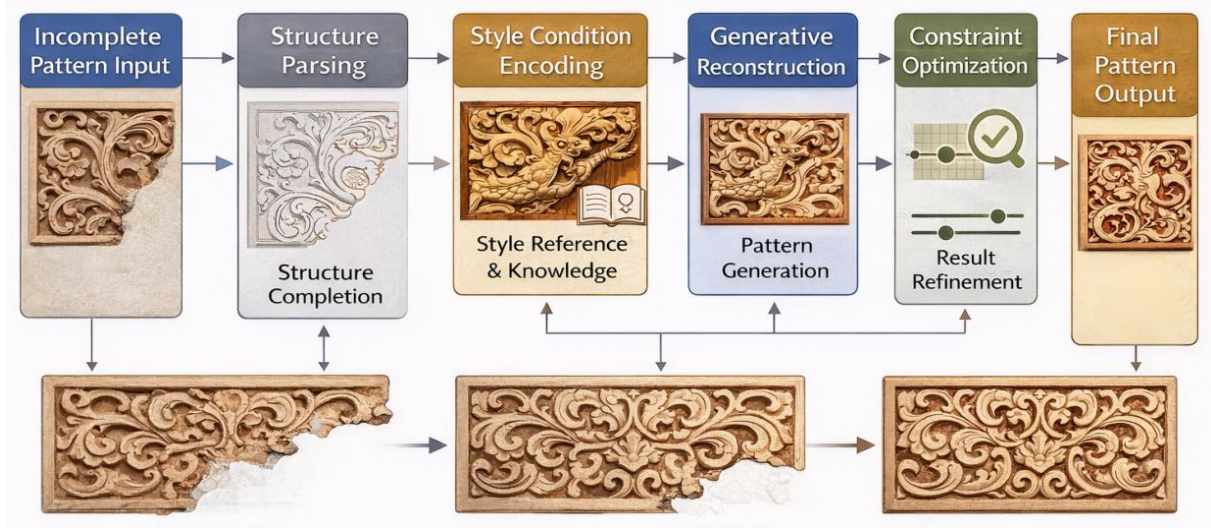


Figure 3: Flow chart of carving pattern generation and style transfer

In the conditional generation stage, suppose that the incomplete pattern input is x_m , the style condition vector is s , and the knowledge constraint embedding is k , then the generator output can be expressed as follows:

$$\hat{x} = G(x_m, s, k) \quad (7)$$

Here, $G(\cdot)$ represents the conditional generation network. This formula indicates that the generated result is controlled by the target style and process knowledge, so as to avoid that the completion result only "looks reasonable" at the pixel level, but deviates from the traditional carving context at the style level. For carving patterns, the knowledge condition is particularly important, because different types of curled grass patterns, twig patterns and kui dragon patterns have strong regularity in repetition units, turning directions and density distribution. If there is no structural prior, the model is easy to generate patterns with fine surface but distorted logic.

In order to ensure the continuity of the reconstruction of the missing region, this paper uses the mask reconstruction loss to enhance the recovery quality of the lost part. Let M be the missing region mask, then the reconstruction term is defined as follows:

$$\mathcal{L}_{rec} = \|M \odot (\hat{x} - x)\|_1 \quad (8)$$

where x is the true complete pattern and \odot represents element-wise multiplication. The loss mainly acts on the area to be repaired, so that the model pays more attention to the relationship between boundary cohesion, local repetition and pattern closure. Due to the mirrored symmetry, continuous rotation and modular splicing structures common in traditional carving patterns, this local constraint can help to reduce broken edges and false extension, and improve the structural integrity of the completion result.

In the style transfer stage, this paper uses the feature correlation matrix to constrain the consistency of the target style, and the style loss is written as follows:

$$\mathcal{L}_{sty} = \sum_{l=1}^L \|\Gamma(\hat{x}^{(l)}) - \Gamma(y^{(l)})\|_F^2 \quad (9)$$

Here, $\Gamma(\cdot)$ represents the correlation matrix of the LTH layer feature, and y is the target style reference sample. This key point restricts the overall decorative density of the pattern, the complexity of the lines and the surface texture temperament, so that the generated results present a visual order similar to the target carving style while being "recognizable in form". For example, wood carving often emphasizes rhythm and level fluctuation of knife marks, brick carving emphasizes plane organization and edge turning, and stone carving often shows a stronger sense of thickness and block surface, through which style differences in different material contexts can be stably transmitted.

In order to improve the realism and usability of the generated patterns, this paper further introduces an adversarial learning mechanism to make the generated results approximate to the real sculptural pattern distribution. The objective function can be written as follows:

$$\mathcal{L}_{adv} = \mathbb{E}_x[\log D(x)] + \mathbb{E}_{x_m, s, k}[\log(1 - D(G(x_m, s, k)))] \quad (10)$$

where $D(\cdot)$ is the discriminator. The introduction of adversarial constraints can improve the problems of edge rigidity, local smoothness and thin texture common in the generated patterns, and make the output patterns closer to the real craft samples in local details, overall level and sculptural sense. In particular, adversarial learning is helpful to enhance the visual completion of patterns in cultural and creative transformation and digital display scenarios.

Considering structure recovery, style preservation and distribution approximation, the joint optimization strategy is finally adopted:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{rec} + \lambda_2 \mathcal{L}_{sty} + \lambda_3 \mathcal{L}_{adv} + \lambda_4 \mathcal{L}_{sem} \quad (11)$$

Among them, \mathcal{L}_{sem} is used to constrain the consistency between the generated results and the semantics of the pattern category, and $\lambda_1, \lambda_2, \lambda_3$, and λ_4 are the weights of each item. Through this multi-objective collaborative optimization, the system can balance the digital restoration and style transformation of traditional carving patterns, so that the output not only retains the stability of traditional decorative language, but also has the ability to expand for contemporary design applications. On the whole, this mechanism provides an operational technical path for the intelligent transformation of traditional sculpture from "static preservation" to "generative, transferable, and redesignable".

3.4 Design of Traditional carving Intelligent interactive System for Inheritance and Innovation

After completing the multimodal data construction, visual coding and pattern generation mechanism design of traditional carving skills, it is necessary to establish an intelligent interactive system that can realize the collaborative operation of knowledge invocation, interactive feedback, assistant learning and creation support. The user interaction layer, perception analysis layer, knowledge service layer and application output layer are integrated into the same framework, so that the traditional sculpture is transformed from a static digital resource into an intelligent service object that is understandable, trainable, deducible and

creative.

From the system structure, the functional architecture shown in Figure 4 can be divided into five levels. The bottom layer is the resource collection and storage layer, which is responsible for aggregating carving images, craft videos, 3D models, literature materials, oral interviews and style samples, and accomplishing unified management. The upper layer is an intelligent parsing layer, which uses the visual coding model, pattern generation model and knowledge representation model constructed in the previous section to recognize the style of the input content, analyze the knife method, match the semantic and reconstruct the structure. The next layer is the knowledge service layer, which outputs interpretable results for users through knowledge graph retrieval, case association and semantic reasoning. The upper layer was an interactive response layer, which supported multiple forms of interaction such as image upload, voice question answering, sketch input, keyword retrieval and process playback. The top layer is the application scene layer, which provides differentiated services for teaching and training, digital display, repair assistance, pattern design and creative development. By stacking multiple functional modules, the system makes recognition, analysis, generation and feedback form a closed loop, so that users can continuously obtain targeted knowledge support and behavior guidance in the process of interaction.

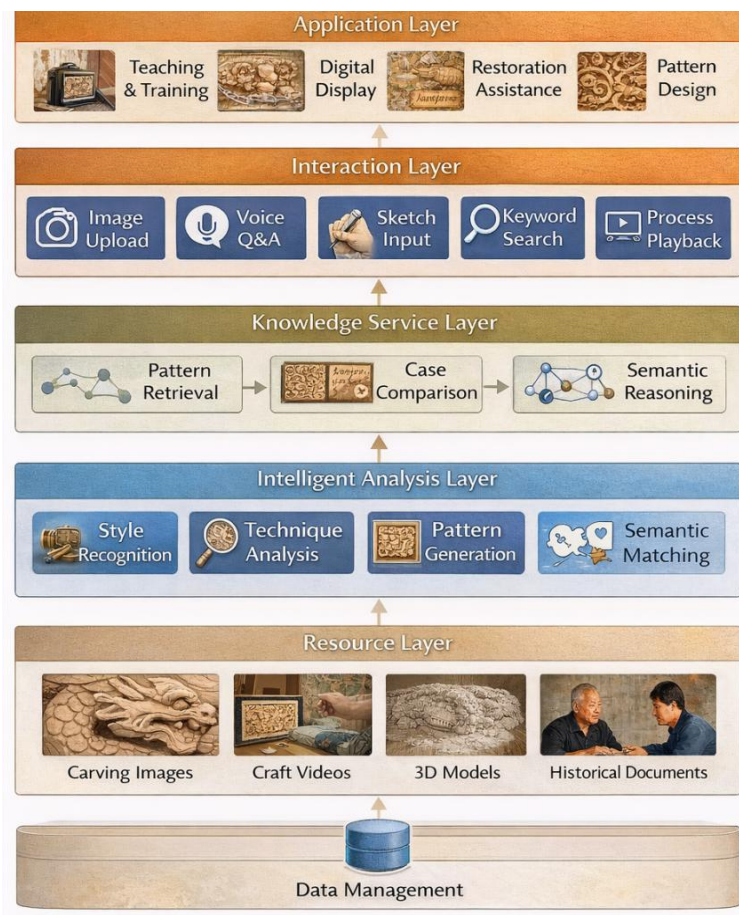


Figure 4: Functional architecture diagram of traditional carving intelligent interactive system

In the stage of user intention understanding, the system needs to quickly judge the service goal according to different input methods. Let the user input feature vector be q and the set of candidate task categories be $\{c_1, c_2, \dots, c_n\}$, then the decision probability of interaction task can be expressed as follows:

$$p(c_i|q) = \frac{\exp(\omega_i^T q + \beta_i)}{\sum_{j=1}^n \exp(\omega_j^T q + \beta_j)} \quad (12)$$

Here, ω_i and β_i are task category corresponding parameters. This formula is used to distinguish whether the user's current intention belongs to different service types such as "style recognition", "technique learning", "pattern generation", "case retrieval" or "repair assistance". Due to the large differences in traditional carving user groups, apprentices are more inclined to process learning, researchers pay more attention to knowledge retrieval, and designers pay more attention to pattern generation and style transfer. Therefore, the accurate judgment of task intention directly affects the efficiency of subsequent service path selection.

In the knowledge service layer, the system needs to match the user input with the patterns, knives, materials and cases in the knowledge base. Let the user query be represented by u and the knowledge node be represented by g_j , then the semantic retrieval score can be defined as follows:

$$s_j = \frac{u^T g_j}{\|u\|_2 \|g_j\|_2} + \eta \cdot r_j \quad (13)$$

Here, the former term represents the semantic similarity between the query and the knowledge node, r_j represents the rule relevance of the node in the current task scenario, and η is the regulation coefficient. The retrieval method can avoid the system only based on the surface keyword matching results, but integrates the semantic of patterns, process context and usage scene to sort. For example, when the user inputs "how to collect the knife at the corner of the dragon pattern wood carving", the system not only returns the dragon pattern sample, but also gives priority to presenting the technical cases related to wood properties, corner structure and round knife processing, so as to improve the pertinency of the interactive results.

In order to enhance the personalized teaching ability in the inheritance scenario, this paper further designs a recommendation mechanism driven by user ability status. Suppose the user's current ability vector is a_t and the target knowledge unit difficulty vector is d_k , then the recommended fitness function can be written as follows:

$$R_{t,k} = \exp\left(-\frac{\|a_t - d_k\|_2^2}{\sigma^2}\right) \cdot \gamma_k \quad (14)$$

Here, γ_k represents the cultural importance or teaching priority of the knowledge unit. For learners who are new to traditional carving, the system will preferentially recommend basic pattern recognition, tool cognition and shallow scoring exercises. For those with experience, complex composition analysis, style discrimination and partial restoration cases will be promoted.

Considering that traditional sculpture learning is highly dependent on practice feedback, the system also needs to continuously update its portraits based on user interaction results. Let the user state after the TTH interaction be a_t and the system feedback evaluation vector be f_t , then the updated state can be expressed as follows:

$$a_{t+1} = \mu a_t + (1 - \mu) f_t \quad (15)$$

Here, μ is the historical experience retention coefficient. The dynamic update mechanism can integrate the user's performance in the process of recognition test, technical question and

answer, sketch submission and pattern generation into the long-term portrait, so that the system can gradually master the user's strength in pattern understanding, process judgment and creative transfer.

4 Evaluation of the effectiveness of the method

4.1 Experimental data source and Platform environment Configuration

In order to verify the applicability of the proposed method in traditional carving category recognition, knife discrimination, pattern reconstruction and intelligent interaction scenes, the experiment constructs a multi-modal sample set of traditional carving from multiple channels, including physical images, process videos, 3D models, text materials and oral records. The collection objects cover three common carriers of wood carving, stone carving and brick carving, and the types of patterns include dragon grain, grass grain, moire, flower grain and geometric grain, which can better reflect the differences of traditional carving in morphological structure, surface texture and process expression. In order to ensure the consistency of the experiment, the image samples are uniformly processed by resolution normalization and color correction, the video samples are extracted by key frame extraction and time series segmentation, the 3D model is converted into a standard grid format, and the text data is entered into the knowledge base after word segmentation cleaning and terminology unification. The experimental platform configuration is shown in Table 2. The overall training environment uses a high memory graphics processing unit (GPU) and a deep learning framework to support the joint testing of visual coding, generative modeling and interactive response modules. It can be seen from Table 2 that the data scale and platform configuration used in this study can meet the needs of multi-task experiments. Among them, the number of carving image samples reaches 3600, the total duration of process video is 18.6 hours, and the 3D model and text data also maintain a good type coverage, which provides a stable basis for the effectiveness evaluation of subsequent methods.

Table 2: Experimental data set and platform environment configuration table

Item	Configuration	Details
Carving image samples	3,600 images	Including 1,500 wood carving images, 1,200 stone carving images, and 900 brick carving images, all standardized to a resolution of 512×512
Craft video samples	186 clips	Total duration of 18.6 h, covering typical processes such as starting the cut, turning the tool, closing the cut, and finishing
3D model samples	420 sets	Obtained through laser scanning and photogrammetry, and uniformly converted into standard mesh models
Textual document samples	520 documents	Including craft literature, technique descriptions, collection records, and local gazetteer materials
Oral record samples	68 records	Collected from inheritor interviews and teaching session records, with transcription and terminology annotation completed
Training / validation / test sets	7 : 1.5 : 1.5	All types of data were divided in a category-balanced manner
Operating system	Ubuntu 22.04	64-bit operating environment
Processor	Intel Xeon Silver 4314	Clocked at 2.40 GHz
Graphics processor	NVIDIA RTX 4090	24 GB VRAM
Deep learning framework	PyTorch 2.1	Running with CUDA 12.1
Interactive system development environment	Python 3.10 + Flask	Supports joint debugging of recognition, retrieval, and generation modules

4.2 Classification and technique discrimination effect analysis of traditional carving

In order to verify the effectiveness of the proposed method in the traditional sculpture visual recognition task, CNN, ResNet, ViT and multimodal baseline methods are selected as the control models, and the comprehensive evaluation is carried out from the two dimensions of sculpture category recognition and knife discrimination. The experimental results of the different methods are shown in Figure 5 and Table 3.

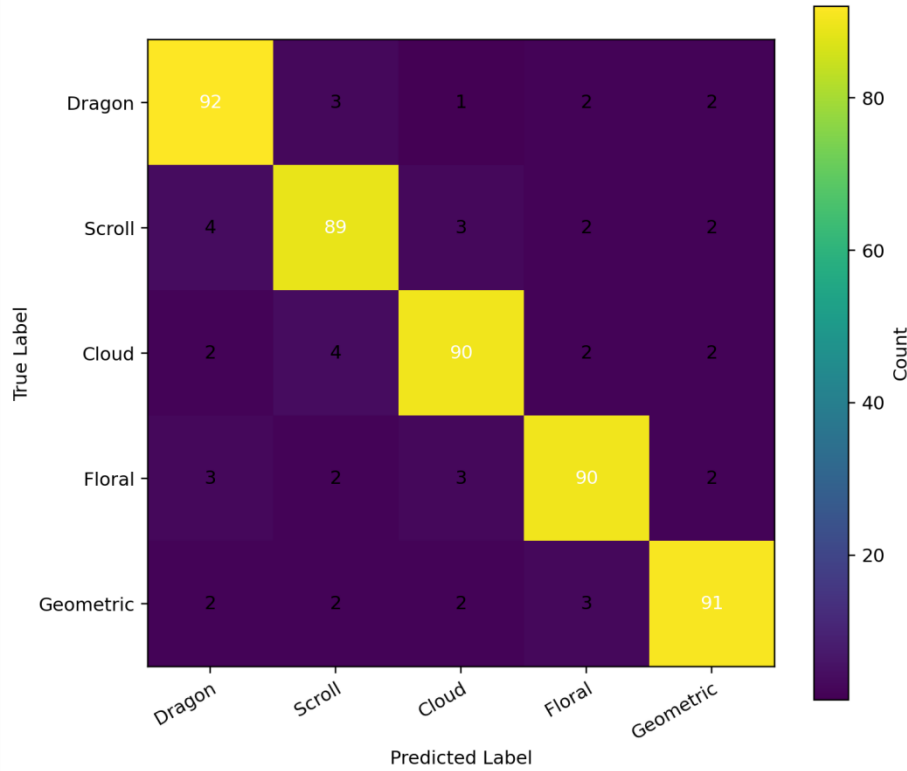


Figure 5: Confusion matrix diagram of recognition results for different carving categories

Table 3: Performance comparison table of different methods in the task of carving category recognition and knife discrimination

Method	Category Accuracy/%	Category F1/%	Technique Accuracy/%	Technique F1/%
CNN	81.2	80.4	77.6	76.8
ResNet	84.7	83.9	81.8	80.9
ViT	86.1	85.4	83.2	82.6
Multimodal Baseline	87.4	86.8	85.1	84.4
Proposed Method	90.6	89.8	88.9	87.9

It can be seen from Figure 5 that the prediction results of the proposed method on five types of samples, such as dragon grain, curly grass grain, moire grain, flower grain and geometric grain, mainly focus on the main diagonal region of the confusion matrix, indicating that the constructed deep visual coding network can better distinguish the morphological structure differences of different sculpted objects. Among them, there is still a small amount of cross misjudgment between the grass grain and the moire grain, and there is also a slight

confusion between the flower grain and the dragon grain, but the overall number of misjudgment is low, which indicates that the model has strong stability in the extraction of complex pattern edges, local repetitive units and surface knife mark information. Combined with Table 3, it can be further seen that the proposed method achieves the best results on both tasks, the sculpture category recognition accuracy reaches 90.6%, and the F1 value reaches 89.8%. The discrimination accuracy of knife method reaches 88.9%, and the F1 value is 87.9%, which is 3.2 percentage points, 3.0 percentage points, 3.8 percentage points and 3.5 percentage points higher than the baseline method of modal, respectively. On the whole, the proposed method shows obvious advantages in both category discrimination ability and technique recognition ability. In Figure 5, the number of main diagonal recognition of five types of samples is all above 89, which further indicates that the proposed method has a good recognition effect on complex visual features of traditional carving.

4.3 Analysis of the quality and style consistency of carving pattern generation

In order to further verify the effectiveness of the proposed method in the task of engraving pattern reconstruction and style transfer, we select the traditional inpainting method, GAN-based generation method and diffusion baseline method as the control model, and comprehensively evaluate the quality of structure restoration, perceptual similarity, style consistency and semantic integrity. In the experiment, the edge continuity, the closure of local decorative units, the natural degree of texture transition and the overall visual coordination under the constraint of the target style after the completion of incomplete patterns are mainly investigated. The experimental results show that although the traditional inpainting method can complete the filling of the basic area, it is easy to cause stiff boundaries and texture fracture in complex corner, revolving patterns and multi-layer decorative structures. The GAN-like methods have improved in visual naturalness, but the local details are still unstable. The diffusion baseline method performs well in the overall structure recovery, but there is still room for improvement in the preservation of traditional carving style and cultural semantic constraints. In contrast, under the combined effect of conditional generation and style constraints, the proposed method can better realize the continuous recovery of pattern structure and the stable transfer of traditional style, so that the output results are more in line with the technical expression characteristics of engraving patterns. Table 4 shows the quantitative evaluation of different methods on the quality of carving patterns generation and style consistency.

Table 4: Evaluation table of quality and style consistency of carving pattern generation

Method	SSIM	PSNR/dB	LPIPS	Style Consistency	Semantic Integrity
Traditional Inpainting	0.781	24.6	0.218	0.742	0.756
GAN-based Method	0.824	26.8	0.176	0.801	0.814
Diffusion Baseline	0.851	27.9	0.149	0.836	0.848
Proposed Method	0.893	30.4	0.112	0.901	0.918

It can be seen from Table 4 that the proposed method achieves optimal results on all evaluation indicators. Among them, the SSIM reaches 0.893, which is 0.112 higher than that of the traditional inpainting method, indicating that the generated results are closer to the real pattern in terms of structure preservation. The PSNR reaches 30.4 dB, which is 2.5 dB higher than the diffusion baseline, indicating that the reconstructed image has higher quality. LPIPS is reduced to 0.112, which shows that the proposed method has more obvious advantages in

perceptual similarity. At the same time, the style consistency and semantic integrity reach 0.901 and 0.918 respectively, which are 0.100 and 0.104 higher than those of GAN-like methods, indicating that the proposed method can not only complete the incomplete area, but also better retain the decorative logic, style characteristics and cultural semantics of traditional carving patterns. Thus, it reflects stronger generation quality and transformation adaptation ability.

4.4 Evaluation of auxiliary effects of skill inheritance and user interaction experience

In order to verify the practical application value of the traditional carving intelligent interactive system in the skill inheritance scene, this paper recruited 12 apprentice users, 12 research users and 12 design users, a total of 36 subjects to carry out the interactive test, and carried out a comprehensive evaluation from two dimensions of interactive experience and learning improvement. The three types of users completed the tasks of system familiarity, case retrieval, technical question and answer, pattern analysis and generation assistance, and the single test time was controlled at about 90 minutes. The interactive experience was evaluated on a 10-point scale from four aspects: operation fluency, functional understanding, feedback clarity and result usability. The pre-test and post-test were carried out around the three tasks of pattern recognition, knife understanding and style judgment, and the test results were uniformly converted into a percent system. In the statistical analysis, the interactive experience part was described by the median, interquartile range and discrete distribution, and the learning improvement part was compared by the mean difference before and after the test to ensure the comparability of the results before and after the use of the system for different user types. The interactive experience scoring results of different user groups are shown in Figure 6.

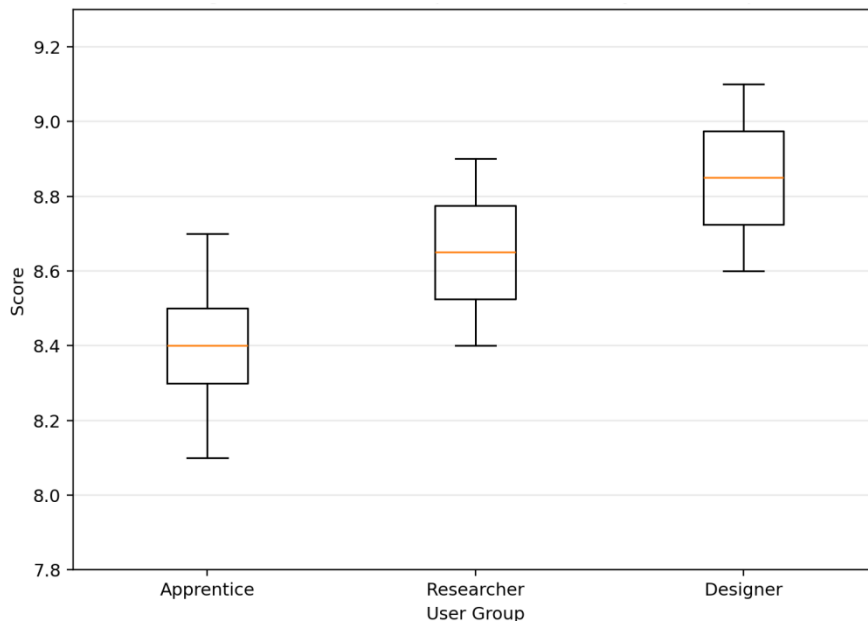


Figure 6: Box plots of interaction experience ratings for different user groups

It can be seen from Figure 6 that the interaction scores of the three types of users are in a relatively high range, and the distribution is relatively concentrated, indicating that the system has good stability in terms of operation fluency, functional understanding and feedback clarity.

Among them, design users had the highest median score of about 8.9, research users had a median score of about 8.7, and apprentice users had a median score of about 8.4. This shows that the system has stronger adaptability to research analysis and creative application scenarios in style retrieval, pattern generation and knowledge interpretation. For apprentice users, although the overall experience is also good, there is still some room for improvement in complex function understanding and multi-module switching. Overall, the lower quartiles of the three types of user ratings remained above 8.2 points, indicating that the system can stably support the inheritance learning and interactive use of different groups of people.

Figure 7 shows the improvement of learning effect before and after the use of the system.

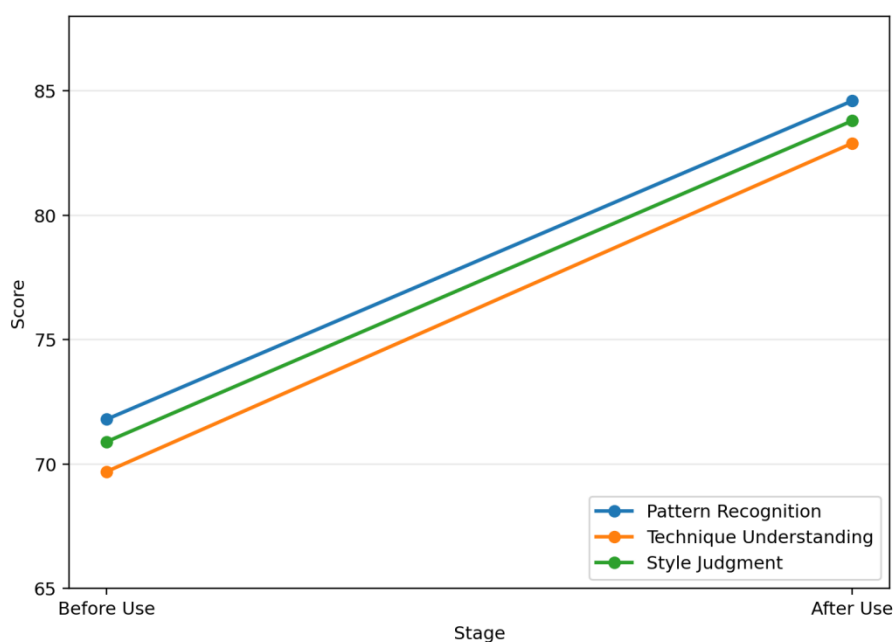


Figure 7: Line chart of learning improvement before and after system use

It can be seen from Figure 7 that after the same group of subjects completed the system-assisted learning, the users showed significant improvement in the three aspects of pattern recognition, knife understanding and style judgment. Among them, the score of pattern recognition increased from 71.8 to 84.6, an increase of 12.8 points. The knife understanding score increased from 69.7 to 82.9, an increase of 13.2 points; The score of style judgment increased from 70.9 to 83.8, an increase of 12.9 points. The improvement of the three indicators is more than 12 points, indicating that the system can not only provide static knowledge display, but also help users establish a more complete cognitive chain of traditional carving through interactive mechanisms such as case retrieval, process playback, technique explanation and feedback generation. In general, the system in this paper shows good advantages in terms of user interaction experience and learning promotion effect, which not only enhances the understandability of traditional carving knowledge transmission, but also improves the continuity and effectiveness of skill learning process, which provides a practical basis for subsequent cross-scene application verification.

4.5 Cross-scenario application capability and system robustness verification

In order to further test the adaptation ability and stable performance of the system in different

task environments, this paper starts from four application scenarios of teaching and training, digital display, repair assistance and cultural and creative design, and comprehensively evaluates five indicators: recognition accuracy, generation quality, response speed, user satisfaction and system stability. Considering that the original dimensions of different indicators are not consistent, in this section, the results are mapped to the percent system before the scene comparison. Among them, the recognition accuracy is directly based on the task recognition results. The generated quality was weighted and summed after normalization by SSIM, PSNR, style consistency and semantic integrity. The response speed is normalized by the reciprocal of the average response delay. User satisfaction was converted from 10-point questionnaire to hundred-point system. The system stability is obtained by combining task completion rate, exception recovery and interface call fluctuation during continuous operation. The comprehensive score of the scene was calculated by weighted summation, and the weights of the five indicators were set to 0.30, 0.25, 0.15, 0.15 and 0.15, respectively. The comprehensive system performance under different application scenarios is shown in Figure 8.

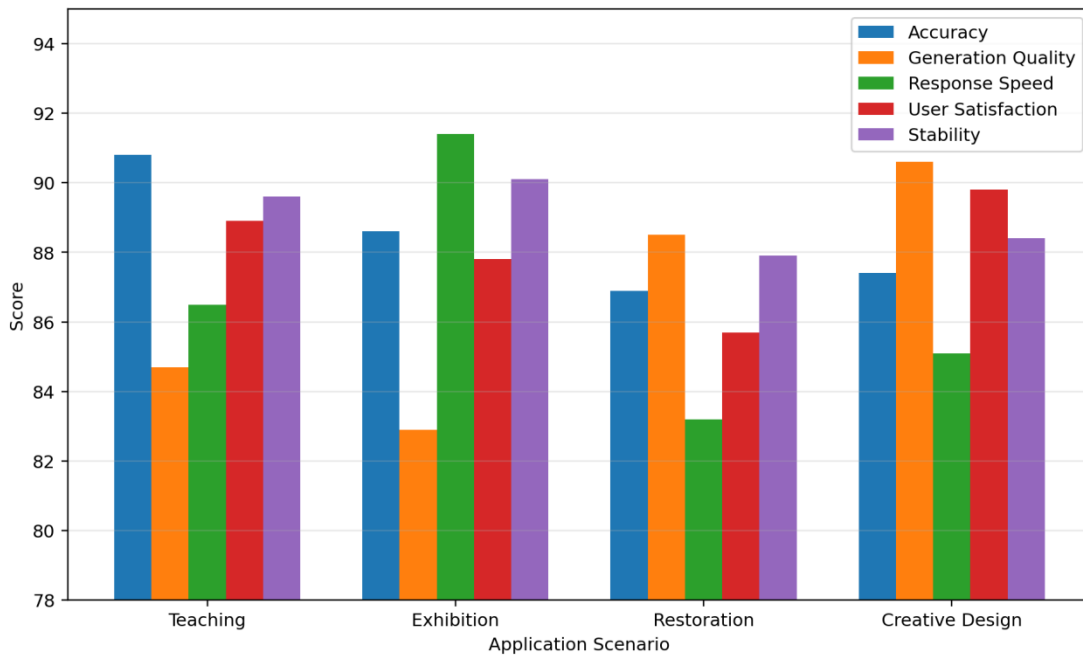


Figure 8: Bar charts for comparison of overall system performance under different application scenarios

As can be seen from Figure 8, the proposed system maintains a relatively high performance level in the four types of scenarios, and the differences in the five dimensions of recognition accuracy, generation quality, response speed, user satisfaction and system stability of each scenario are relatively clear. Among them, the teaching training scene performs better in terms of recognition accuracy and stability, reaching 90.8 and 89.6 respectively. The digital display scene is the most prominent in response speed, with a score of 91.4. The generation quality dimension of the inpainting assisted scene reaches 88.5, indicating that the system has strong adaptability in incomplete pattern completion and style maintenance. The generation quality and user satisfaction of cultural and creative design scenarios reach 90.6 and 89.8, respectively, reflecting a good innovation support ability. Overall, the average comprehensive score of the four types of scenes reaches 87.9, indicating that the system can not only serve

the inheritance of traditional sculpture knowledge, but also adapt to diverse application requirements such as display communication, repair analysis and creative transformation.

In order to analyze the contribution of each key module to the overall performance, this paper further carries out ablation experiments to remove the knowledge constraint, style coding, multi-modal fusion and interactive feedback modules respectively, and compares the task performance under each configuration. The results are shown in Figure 9.

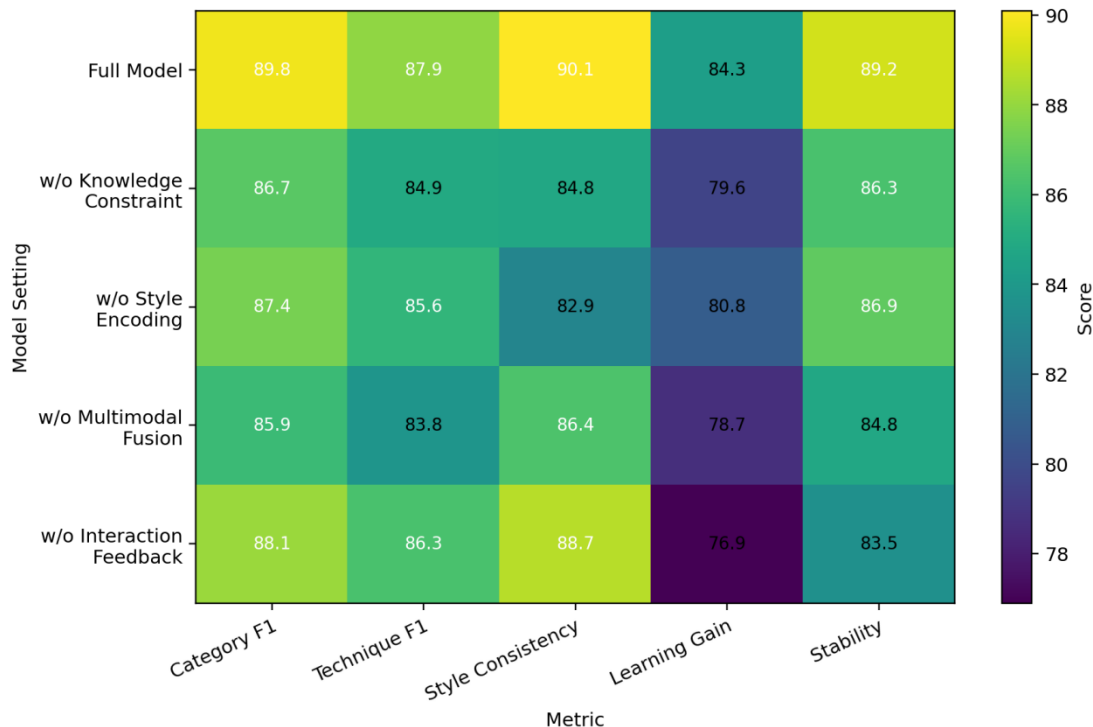


Figure 9: Matrix plot of ablation experiment performance of key modules

It can be seen from Figure 9 that the complete model remains optimal in the five indicators of category recognition F1, knife discrimination F1, style consistency, learning gain and system stability, with scores of 89.8, 87.9, 90.1, 84.3 and 89.2, respectively. After removing the knowledge constraint, the style consistency decreases from 90.1 to 84.8, which indicates that the traditional process semantic constraint plays an important role in the cultural fidelity of the generated results. After removing the style encoding, the style consistency is further reduced to 82.9, which shows that this module is particularly critical for pattern transfer. After removing the multi-modal fusion, the F1 of category recognition decreased to 85.9, indicating that the collaborative expression of vision, text and process information had a significant role in promoting the recognition task. After removing the interactive feedback, the learning gain decreased from 84.3 to 76.9, and the reduction reached 7.4 points, indicating that the feedback mechanism was the most significant support for continuous learning in the inheritance scenario. In summary, the proposed system performs stably in cross-scene tasks, and each key module plays a substantial role in the overall performance improvement.

The system recognition performance changes under different noise interference or occlusion conditions are shown in Figure 10.

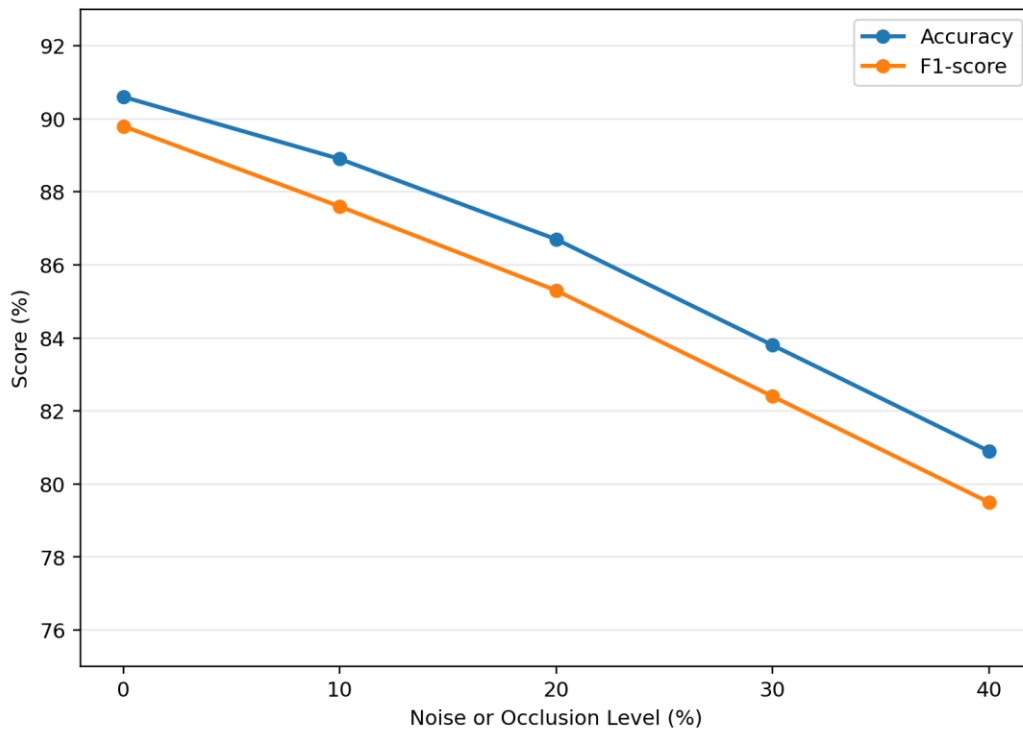


Figure 10: Graph of recognition performance variation under noise interference or occlusion conditions

As can be seen from Figure 10, with the gradual increase of noise intensity or occlusion ratio, the recognition accuracy and F1 value of the system show a gentle downward trend, but the overall fluctuation range is relatively controllable. When the interference level increased from 0% to 40%, the recognition accuracy decreased from 90.6% to 80.9%, a decrease of 9.7 percentage points. F1 value decreased from 89.8% to 79.5%, a decrease of 10.3 percentage points. Although the local pattern edge, knife detail and structure contour will be affected differently under the condition of high interference

5 Discussion

This paper focuses on the digital protection and intelligent transformation of traditional carving skills in the AI era, and forms a relatively complete technical chain from multimodal data construction, visual feature coding, pattern generation to interactive system design. Combined with the experimental results in Chapter 4, it can be seen that the constructed method shows good stability in the tasks of carving category recognition, knife discrimination, pattern reconstruction, style transfer and interactive assistance, which indicates that the image, video, 3D shape, text knowledge and process semantics are incorporated into a unified framework. It can effectively alleviate the problems of "seeing only the pattern, but not the technique" and "seeing only the finished product, but not the process" in the traditional carving research. Especially, in the pattern generation and style consistency evaluation, the proposed method takes into account both structural integrity and cultural semantic constraints, and the generated results are no longer superficial similarities at the level of general decorative patterns, but closer to the internal organization law and aesthetic order of traditional carving patterns.

From the application level, the performance of intelligent interactive systems in learning promotion and multi-scene adaptation is also of practical significance. Experiments show that users have significantly improved in pattern recognition, knife understanding and style judgment after using the system, indicating that AI technology can not only assume the functions of data storage and information retrieval, but also participate in teaching training and cognitive construction as a procedural auxiliary tool in the inheritance of traditional skills. This point has a strong pertinence for the current problems of traditional carving, such as long inheritance cycle, strong experience, and scattered teaching resources. At the same time, the system maintains high scores in teaching and training, digital display, repair assistance and cultural and creative design scenarios, which also shows that the digital transformation of traditional crafts is not a single path, but can establish a linkage mechanism between protection, dissemination, research and innovation.

Of course, there are still some limitations in this study. First, although the samples cover wood carving, stone carving and brick carving, the regional schools and material types are still not rich enough, and some minority patterns with strong local characteristics have not been fully included, which may limit the generalization ability of the model. Second, the discrimination of tool method is still mainly dependent on the reverse process characteristics of visual results, and the description of real hand movements, knife operation strength and machining sequence is not detailed enough. Third, although the generative model has improved in maintaining the traditional style, it may still appear that the local details are too regular or the cultural meaning is not fully expressed in the face of highly complex historical sculpture context. Subsequent research can further expand the scale of multi-source samples, introduce more fine-grained motion capture and process data, and further deepen the knowledge constraints, expert participation and interactive feedback mechanism, so as to improve the system's ability to express the real connotation of traditional carving skills.

6 Conclusion

Focusing on the watch and transformation of traditional carving skills, this paper forms an overall technical path of "multi-modal data construction-deep visual encoder-generative pattern reconstruction-intelligent interactive service". The research shows that the multimodal collaborative representation can effectively improve the performance of carving category recognition and knife discrimination, the knowledge constraint and style coding can significantly improve the quality of pattern generation, and the interactive feedback mechanism enhances the continuity and pertinence of inheritance learning. In the experiment, the average comprehensive score of the system in four types of application scenarios reaches 87.9, and the recognition accuracy is 80.9% under the condition of 40% noise or occlusion, which shows that the proposed method has both stability and practicability. In the future, it is necessary to further expand the regional genre samples, introduce more fine-grained motion capture and expert participation mechanism, and continue to improve the system's ability to express the complex process semantics and historical context of traditional carving.

Funding

This work was supported by 2023 Jiangsu Provincial University Philosophy and Social Science Research Project "Research on the Inheritance and Development of Intangible Cultural Heritage in Higher Vocational Colleges from a Digital Perspective", 2023SJYB1523.

References

- [1] Mishra M. Machine learning techniques for structural health monitoring of heritage buildings: A state-of-the-art review and case studies[J]. *Journal of Cultural Heritage*, 2021, 47: 227-245. DOI:10.1016/j.culher.2020.09.005.
- [2] Monna F, Rolland T, Denaire A, et al. Deep learning to detect built cultural heritage from satellite imagery. - Spatial distribution and size of vernacular houses in Sumba, Indonesia -[J]. *Journal of Cultural Heritage*, 2021, 52: 171-183. DOI:10.1016/j.culher.2021.10.004.
- [3] Koutsoudis A, Ioannakis G, Pistofidis P, et al. Multispectral aerial imagery-based 3D digitisation, segmentation and annotation of large scale urban areas of significant cultural value[J]. *Journal of Cultural Heritage*, 2021, 49: 1-9. DOI:10.1016/j.culher.2021.04.004.
- [4] Pistofidis P, Arnaoutoglou F, Ioannakis G, et al. Design and evaluation of smart-exhibit systems that enrich cultural heritage experiences for the visually impaired[J]. *Journal of Cultural Heritage*, 2023, 60: 1-11. DOI:10.1016/j.culher.2023.01.004.
- [5] Mishra M, Lourenço P B. Artificial intelligence-assisted visual inspection for cultural heritage: State-of-the-art review[J]. *Journal of Cultural Heritage*, 2024, 66: 536-550. DOI:10.1016/j.culher.2024.01.005.
- [6] Towarek A, Halicz L, Matwin S, Wagner B. Machine learning in analytical chemistry for cultural heritage: A comprehensive review[J]. *Journal of Cultural Heritage*, 2024, 70: 64-70. DOI:10.1016/j.culher.2024.08.014.
- [7] Karimi N, Mishra M, Lourenço P B. Deep learning-based automated tile defect detection system for Portuguese cultural heritage buildings[J]. *Journal of Cultural Heritage*, 2024, 68: 86-98. DOI:10.1016/j.culher.2024.05.009.
- [8] D'Orazio M, Gianangeli A, Monni F, Quagliarini E. Automatic monitoring of the bio colonisation of historical building's facades through convolutional neural networks (CNN)[J]. *Journal of Cultural Heritage*, 2024, 70: 80-89. DOI:10.1016/j.culher.2024.08.012.
- [9] Cui J, Tao N, Omer A M, et al. Attention-enhanced U-Net for automatic crack detection in ancient murals using optical pulsed thermography[J]. *Journal of Cultural Heritage*, 2024, 70: 111-119. DOI:10.1016/j.culher.2024.08.015.
- [10] Wen R, Fan F. Quantifying pigment features of Thangka Five Buddhas using hyperspectral imaging[J]. *Journal of Cultural Heritage*, 2024, 70: 120-133. DOI:10.1016/j.culher.2024.08.020.
- [11] Agirbas A. Geometric decomposition and algorithmic model of main gate muqarnas of Ankara Cenabi Ahmet Pasha Mosque[J]. *Journal of Cultural Heritage*, 2024, 70: 134-142. DOI:10.1016/j.culher.2024.09.003.
- [12] Cardarelli L. From fragments to digital wholeness: An AI generative approach to reconstructing archaeological vessels[J]. *Journal of Cultural Heritage*, 2024, 70:

250-258. DOI:10.1016/j.culher.2024.09.012.

- [13] Chen Y, Zhang L, Dong Q. Using natural language processing to evaluate local conservation text: A study of 624 documents from 303 sites of the World Heritage Cities Programme[J]. *Journal of Cultural Heritage*, 2024, 70: 259-270. DOI:10.1016/j.culher.2024.09.011.
- [14] Lee A, Yu H, Min G. An algorithm of line segmentation and reading order sorting based on adjacent character detection: A post-processing of OCR for digitization of Chinese historical texts[J]. *Journal of Cultural Heritage*, 2024, 67: 80-91. DOI:10.1016/j.culher.2024.02.001.
- [15] Gattiglia G. Managing Artificial Intelligence in Archeology. An overview[J]. *Journal of Cultural Heritage*, 2025, 71: 225-233. DOI:10.1016/j.culher.2024.11.020.
- [16] Corradetti D, Delgado Rodrigues J. Identification of stone deterioration patterns with large multimodal models. Definitions and benchmarking[J]. *Journal of Cultural Heritage*, 2025, 71: 175-183. DOI:10.1016/j.culher.2024.11.017.
- [17] Malchiodi D, Zanaboni A M, Di Gioacchino A, Bonizzoni L. One-class vs binary machine learning classification of ceramic samples described by chemical element concentrations[J]. *Journal of Cultural Heritage*, 2025, 71: 234-241. DOI:10.1016/j.culher.2024.11.015.
- [18] Pang B, Yang J, Xia T, et al. Automated heritage building component recognition and modelling based on local features[J]. *Journal of Cultural Heritage*, 2025, 71: 252-264. DOI:10.1016/j.culher.2024.12.006.
- [19] Crabu E, Pes F, Rodriguez G, Tanda G. Color reconstruction by photometric stereo with unknown lighting of rock art carvings found in two Sardinian Domus de Janas[J]. *Journal of Cultural Heritage*, 2025, 71: 320-327. DOI:10.1016/j.culher.2024.12.019.
- [20] Vuoto A, Funari M F, Karimzadeh S, Lourenço P B. Generative modelling of Monopteros and Tholos temples using existing data: The case study of Vesta temple in Tivoli[J]. *Journal of Cultural Heritage*, 2025, 71: 334-345. DOI:10.1016/j.culher.2024.12.005.
- [21] Mezina A, Burget R. EnsArtNet: Ensemble neural network architecture for identifying art styles from paintings[J]. *Journal of Cultural Heritage*, 2025, 72: 71-80. DOI:10.1016/j.culher.2025.01.005.
- [22] Li S, Jiang Y, Jing B, et al. AI-based experts' knowledge visualization of cultural heritage: A case study of Terracotta Warriors[J]. *Journal of Cultural Heritage*, 2025, 72: 81-90. DOI:10.1016/j.culher.2025.01.006.
- [23] Ramírez-Arellano A, Muñoz-Silva E M, Antonio-Cruz M, Vasquez-Gomez J I. Deng entropy and LSTM neural network to classify built cultural heritage with severe damage[J]. *Journal of Cultural Heritage*, 2025, 73: 286-294. DOI:10.1016/j.culher.2025.04.003.
- [24] Ottoni A L C, Ottoni L T C. A deep learning approach for cultural heritage building

- classification using transfer learning and data augmentation[J]. *Journal of Cultural Heritage*, 2025, 74: 214-224. DOI:10.1016/j.culher.2025.06.010.
- [25] Colace F, Gaeta R, Lorusso A, Pellegrino M, Santaniello D. New AI challenges for cultural heritage protection: A general overview[J]. *Journal of Cultural Heritage*, 2025, 75: 168-193. DOI:10.1016/j.culher.2025.07.019.
- [26] Kloiber M, Kunecký J, Hrivnák J, Eisler M, Růžička P. Time demands of traditional hand-hewing techniques for sustainable repair of wooden heritage structures: A comparative study[J]. *Journal of Cultural Heritage*, 2024, 66: 197-203. DOI:10.1016/j.culher.2023.11.009.
- [27] Zhu Z, Qin Y, Guo Z, et al. Shedding new light on lacquering crafts from the Northern Wei Dynasty (386–534 CE) by revisiting the lacquer screen from Sima Jinlong's Tomb[J]. *Journal of Cultural Heritage*, 2025, 71: 309-319. DOI:10.1016/j.culher.2024.12.002.