ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Exploration of Visual Communication Design Element Extraction and Reconstruction Techniques Based on Laplace Transform Algorithms in Digital Preservation of Cultural Heritage

Weiyuan Liu[1], *Sirui Chen*[1] *and Yanchun Jiang*[2,*]

[1] School of Art and Design,Guilin University of Electronic Technology,Guilin 541000, Guangxi Zhuang Autonomous Region
[2] Nanxiashan Hospital of Guangxi Zhuang Autonomous Region,Guilin 541000, Guangxi Zhuang Autonomous Region

SUMMARY: *The Laplace transform algorithm serves as a pivotal component in the application of technology to humanities. This paper integrates the Laplace transform algorithm with a multi-scale feature learning network to extract digital image features, leveraging the unique value of cultural heritage. By utilizing the discrete convolution kernel structure of the Laplace operator, it performs second-order differential calculations on image edge features. Combined with the encoding and decoding steps of the multiscale feature learning network, this approach enhances image segmentation accuracy. It achieves edge structure extraction and global semantic integration in digital images. Compared with similar image feature extraction methods, the proposed method achieves over 90% accuracy in feature extraction and classification performance across three datasets. Moreover, the computational time ranges from 36.53 to 75.42 seconds. This method demonstrates high precision and fast speed in image feature extraction.*

KEYWORDS: *Laplace transform; discrete convolution kernel; multiscale feature learning; cultural heritage images; feature extraction*

1 Introduction

Cultural heritage serves as tangible evidence of humanity's historical progress, playing a vital role in disseminating cultural knowledge, embodying national identity, and documenting the evolution of civilization. From ancient architectural ruins to exquisite handicrafts, from precious literary texts to traditional folk arts, each cultural heritage item bears witness to the evolution of human society, representing irreplaceable treasures [1, 2]. As the global crisis of cultural heritage loss intensifies, digital preservation has become imperative, aiding in its transmission and safeguarding [3, 4]. With the advancement of digital technology, the medium of cultural heritage has gradually detached from physical cultural sites, freeing itself from the constraints of physical space on artifacts. Through three pathways—digital scanning, digital photography, and digital modeling—it has evolved into digital media represented by digital archives and digital museums within the digital realm [5-8]. This has imposed new demands on the visual presentation of cultural heritage for external display. Digital preservation of cultural heritage is not only a long-standing initiative advocated by UNESCO but also a key indicator of a nation's technological foundation and conservation infrastructure for cultural

*18632332365@163.com
https://doi.org/10.65102/is20261020
Published: 30 April, 2026

heritage [9-11].

As an integral component of cultural heritage, visual communication design elements have gained more diverse presentation methods in the digital era, enabling more efficient information transmission, stronger communicative impact, and enhanced interactive engagement [12-14]. Visual communication design enables objects to express their image information externally and convey it to people for visual perception. For cultural heritage, especially intangible cultural heritage, visual communication design is particularly crucial as it facilitates human memory and acceptance [15, 16]. Nicholas Milzow, Associate Professor of Art at Stony Brook University, argues that visual elements subvert and challenge any attempt to define culture through pure linguistic forms. This insight urges us to advance the visualization of cultural heritage, transcending the limitations of textual representation alone [17]. Against this backdrop, extracting and reconstructing visual communication design elements from cultural heritage has become a vital technical foundation for both digital preservation and design innovation.

In advanced mathematics, the Laplace transform serves as a mathematical tool for transforming functions. Its core principle involves transforming time-domain problems into frequency-domain problems through mathematical transformation, converting linear ordinary differential equations of time functions into algebraic equations of complex-valued functions. After obtaining the desired complex-valued function, an inverse transform is applied to retrieve the target time-domain function. This technique finds extensive application in image processing, signal processing, and related fields [18-20]. For instance, Rahmawati et al. (2024) demonstrated that combining the Laplace transform with image enhancement techniques can improve fog removal and image quality in transmission map reconstruction based on dark channel prior methods under foggy conditions [21]. This illustrates that the Laplace transform serves as a crucial analytical tool for characterizing visual communication design elements in cultural heritage.

Zhang et al. (2021) combined sketch prior knowledge with hyperspectral images of cultural relic sketches to extract image bands. They then adjusted the model based on expert-drawn relic diagrams and mitigated noise interference within the U-net architecture to restore faded and deteriorated sketches [22]. Hu et al. (2021) applied image processing tools and user selection to extract cultural form design elements from traditional cultural images. They then identified key design elements within a decision matrix using fuzzy analytic hierarchy process and entropy calculations [23]. Qiu et al. (2023) designed a spatial dual-encoder model and a spectral fuzzy C-means algorithm for extracting and analyzing material properties of Sanxingdui artifacts. This approach effectively identifies surface substances and extracts structural features, enabling efficient recognition of totemic objects when combined with intelligent synthesis technology [24]. Zhang et al. (2025) developed a generative adversarial network algorithm based on U-Net architecture and recurrent consistency to extract visual elements from intangible cultural heritage in complex environments. This approach significantly improved extraction accuracy while preserving high aesthetic value and cultural identity [25]. Jiang (2025) employed adaptive convolutional neural networks to recognize traditional cultural images. Under denoising diffusion probabilistic model symbol reconstruction and color clustering methods, visual symbols (including shape, color, pattern, and five other types) were extracted and transformed from image data [26]. Hu et al. (2025) combined root mean square error constrained smoothing algorithms, collision detection, nonlinear functions, and grayscale information to extract textures and other details from digital rubbing patterns within 3D digital models of stone carving artifacts, enabling deep extraction of friction traces [27].

Belhi et al. (2020) established low-latency image completion and reconstruction methods using supervised and unsupervised deep learning algorithms, enabling the restoration of

damaged or missing cultural heritage images [28]. Nasri and Huang (2020) integrated thresholding techniques, RGB-to-HSV color space conversion, and morphological operations to remove dust and small objects from high-resolution images of ancient statues. They then converted the image color space to extract and measure missing color regions [29]. Song and Wu (2022) applied discrete wavelet transform methods, incorporating fuzzy C-means algorithms for automatic image clustering. They performed pixel-based automatic sample color migration and solved for reconstructed images after denoising with wavelet domain image restoration models, achieving virtual color restoration of cultural heritage and enhancing visual effects [30]. Süvari et al. (2023) employed laser scanning for cultural heritage surveying, constructing a virtual restoration model based on internal spatial data and augmented reality technology. This model reconstructs visual elements such as spatial layout, color, and texture of cultural heritage based on original image records [31]. Li and Chen (2024) employed virtual engine technology and fine-grained display techniques to construct an intangible cultural heritage image reconstruction platform and transform cultural visual symbols, enabling virtual presentation and cross-cultural understanding of intangible heritage visual symbols [32].

Cultural heritage possesses characteristics such as diversity and uniqueness, further highlighting the importance of digital preservation and innovative development. This paper selects the Laplacian transform algorithm as one method for extracting design elements from cultural heritage digitalization, performing feature sharpening and extraction. By approximating the Laplacian operator using discrete convolution kernels, it continuously calculates the second-order differential of edge contour structures in digital images, progressively approximating and restoring image edge features. A multi-scale feature learning network is employed to refine the extraction of design element details. The combination of multi-scale features, bilinear interpolation, and a loss function enhances the accuracy of image semantic segmentation, strengthens feature details, and reduces model training overhead.

2 Digital Preservation of Cultural Heritage Using Laplace Transform Algorithms

2.1 The Relationship Between Information Visualization and Digital Cultural Heritage (DCH)

Digital cultural heritage preservation transcends the temporal and spatial constraints of traditional physical preservation methods, while information visualization enables traditional cultural heritage conservation to break free from two-dimensional static limitations, evolving toward multidimensional, dynamic, interactive, and virtual approaches. As humanity now accesses cultural heritage information in real-time and serialized formats (e.g., multidimensional dynamic geospatial visualization of archaeological sites, network data visualization of museum collections), the accumulation of massive datasets continues to accelerate. This objectively demands greater capacity and methods to process such data, thereby enhancing our understanding of the trajectories of national civilization and social progress reflected in cultural heritage. The process of visualizing digital cultural heritage information essentially involves graphically simulating its internal constituent elements and external constraints. This approach reduces the complex abstract symbols traditionally used in cultural heritage preservation, establishing information generation and realization systems for digital heritage protection through more intuitive, multidimensional, dynamic, interactive, and virtual representations. It enables interactive operations between users and the digital cultural heritage information environment, shortening the temporal and spatial distance of cultural heritage

information transmission. It minimizes information distortion and loss caused by intermediaries, enhances cultural heritage education and museum experience effectiveness. In summary, information technology-based cultural heritage preservation serves as both an evolution and supplement to physical preservation, while physical preservation remains the foundation and source of information technology-based preservation.

Information visualization design for digital cultural heritage preservation entails creating visual representations tailored to heritage conservation. In the context of information economics, "informationization" signifies the separation of information content from its reliance on other sensory forms. Through the transformation and construction of information languages—such as textual symbols, color, music, graphics, and images—into formalized and "semantic" expressions, it achieves the objective revelation of key attributes within information content. This process facilitates information creation and distant interpretation, providing mechanisms for consumption and evaluation. Information visualization necessitates the presentation and interaction of visual elements within information deconstruction and reconstruction processes. These activities inevitably intertwine with the visual design and interaction design practices of information designers. This study endeavors to approach information visualization design for digital cultural heritage preservation as a systematic engineering project, addressing all relevant factors holistically. Information visualization design and reconstruction leverage technological and conceptual advancements from the information society to gain a more comprehensive and profound understanding of users' material and spiritual selves. By placing individuals at the center of digital cultural heritage preservation activities, it employs design philosophies and reconstruction principles to create "human-centered" designs grounded in practical reasoning. In this sense, information visualization design reconstructs a new visual culture that propels the advancement of digital cultural heritage preservation.

2.2 Analysis of Application Methods for the Laplace Transform Algorithm

2.2.1 Definition of the Laplace Transform

The core principle of the Laplace transform lies in measuring the "rate of change of the rate of change" (the slope's rate of variation) of image edge shapes through second derivatives. By leveraging the characteristic of slope "zero-crossing points," it achieves precise edge detection and feature extraction, thereby fully restoring, extracting, and reconstructing design elements within cultural heritage images. This section provides a detailed exposition of the Laplace transform's definition.

Let $f(t)$ be a function of time t such that when $t < 0.0$, $f(t) = 0.0$. For a function $f(x)$ of a real variable t defined on the interval $(0.0, +\infty)$, multiply it by e^{-pt} (where p is a complex number), then integrate t from 0.0 to $+\infty$. If this generalized integral converges, it defines a complex-valued function $F(p)$ of a complex number p:

$$F(p) = \int_0^{+\infty} f(t)e^{-pt}dt \tag{1}$$

The integral in the above equation establishes a correspondence between the function $f(t)$ and another function $F(p)$. This correspondence is known as an integral transform, and the specific transform here is called the Laplace transform, denoted by $L[f(t)]$, i.e.,

$$L[f(t)] = F(p) = \int_0^{+\infty} f(t)e^{-pt}dt \; Or \; F(p) = L[f(t)] \tag{2}$$

At this point, the function $F(p)$ is referred to as the image function of $f(t)$, or the result of the Laplace transform of $f(t)$, commonly abbreviated as the Laplace transform of $f(t)$. Conversely, the function $f(t)$ is termed the original function of $F(p)$ or the inverse Laplace transform. The inverse Laplace transform of $F(p)$ is denoted by $L^{-1}[F(p)]$, that is:

$$f(t) = L^{-1}[F(p)] \tag{3}$$

The Laplace transform is a complex-valued function of the complex variable $p = \alpha + \beta i$, where $\alpha = \mathrm{Re}(p)$ and $\beta = \mathrm{Im}(p)$.

As a significant integral transform, the Laplace transform converts a time-domain function $f(t)$ into a complex-valued function $F(p)$. Here, $f(t)$ is the original function, $F(p)$ is the image function, and p is often referred to as the complex frequency. Consequently, $F(p)$ is also termed the frequency-domain function or complex-frequency function.

2.2.2 Existence Conditions for the Laplace Transform

A function requires certain conditions to undergo a Laplace transform, namely the conditions for the existence of the Laplace transform:

1) $f(t)$ is piecewise continuous on any finite interval $t \geq 0.0$, with a finite number of discontinuities, all of which are first-order discontinuities;

2) When $t < 0.0$, $f(t) = 0.0$;

3) $f(t)$ is finite-order (or exponential), meaning there exist constants $m \geq 0.0$ and $M > 0.0$ such that

$$\left| f(t) \right| \leq Me^{mt} (t \geq 0.0) \tag{4}$$

Here, the number m is called the growth index of $f(t)$. When $f(t)$ is a bounded function and $m = 0.0$, the Laplace transform of $f(t)$ is:

$$F(p) = \int_0^{+\infty} f(t)e^{-pt}dt \tag{5}$$

In the half-plane $\mathrm{Re}(p) > m$, it must exist, and in this case, the integral on the right-hand side of the above equation converges absolutely and uniformly. Furthermore, within this half-plane, $F(p)$ is an analytic function.

2.2.3 Laplace Operator

The specific application of the Laplace transform algorithm in feature extraction and enhancement of discrete two-dimensional images for cultural heritage digitization is primarily achieved through the Laplacian operator. This involves directly applying its discrete convolution kernel form to perform edge detection and sharpening on images, thereby restoring and extracting the shape features of design elements. Since the Laplacian operator is sensitive to image noise when used alone, it is typically combined with Gaussian filtering to complete the image smoothing preprocessing. The Laplacian operator is a second-order differential operator. For an image $f(x, y)$, its Laplacian operator is defined as follows:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{6}$$

In particular, the two second-order partial derivatives in equation (6) can be expressed as follows:

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y) \tag{7}$$

$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y) \tag{8}$$

From Equations (7) and (8), it can be seen that the Laplace transform is a linear transformation.

Therefore, for discrete two-dimensional images in cultural heritage digitization, the Laplacian operator can be expressed as:

$$\nabla^2 f = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y) \tag{9}$$

Figure 1 shows the commonly used template for the Laplace operator formula. A discrete convolution kernel is employed to approximate the Laplace operator. Common kernels include the 4-neighborhood kernel and the 8-neighborhood kernel. The template corresponding to formula (9) is shown in Figure 1(a), while the other three templates are variations of Figure 1(a). The coefficients in the template represent the coefficients at each pixel position (x, y). Diagonal directions can also be incorporated into the Laplace transform definition by adding two terms to the above equation—one for each diagonal direction. Each newly added term follows a form similar to Equation (7) or Equation (8), but with coordinates aligned along the diagonal direction.

Since each diagonal direction includes a $-2f(x, y)$, the total sum subtracted from the terms in different directions is $-8f(x, y)$. Figure 1(b) shows this newly defined template. The other two templates shown in the figure are also frequently used in experiments. These templates are similarly based on the Laplace transform definition, but the sign of the intermediate coefficients is opposite to that in Figure 1(a) and Figure 1(b), respectively. Consequently, they produce identical results. However, when the Laplace-filtered image is added to or subtracted from other images, the difference in sign must be taken into account. The fundamental method of image enhancement using the Laplace transform is:

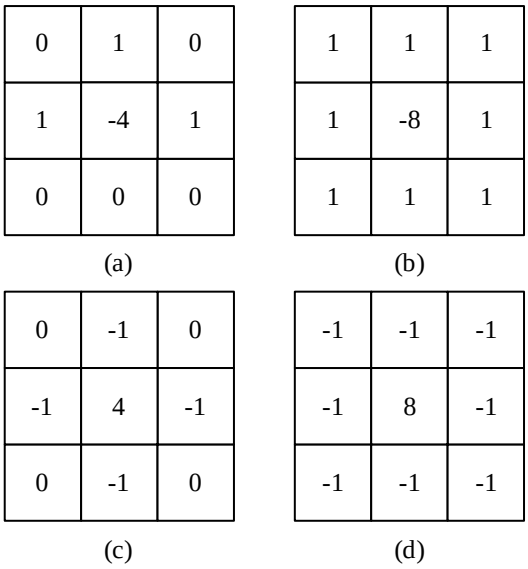

Figure 1: Common Templates of Laplace Transform

Overlaying the original image with its Laplacian transform sharpens the image while simultaneously restoring other information. Therefore, the method for enhancing images using the Laplacian transform is as follows:

$$g(x,y)=\begin{cases} f(x,y)-\nabla^2 f(x,y), & \text{If the Laplace template center is negative} \\ f(x,y)+\nabla^2 f(x,y), & \text{If the Laplace template center is positive} \end{cases} \tag{10}$$

Overall, extracting design elements from digitized cultural heritage images using discrete convolution kernels of the Laplacian operator requires: 1) Smoothing image noise with Gaussian filters to reduce convolution sensitivity; 2) Enhancing shape edges with Laplacian sharpening filters if elements exhibit low contrast; 3) Selecting scale standard deviation parameters suited to varying edge thicknesses to extract the overall contours of design elements.

2.3 Feature Extraction Based on Multi-Scale Feature Learning Network Architecture

2.3.1 Encoder Architecture Based on Multi-Scale Features

The network architecture comprises four modules: the Multi-Scale Block (MSB), Multi-Scale Encoder (MSE), Multi-Scale Decoder (MSD), and Semantic Segmentation Head (SH). To further leverage contextual information within the image feature network, a series of skip connections are stacked between the encoder and decoder for feature fusion. Detailed analysis of the modules follows:

Figure 2 illustrates the structure of the Multi-Scale Block (MSB). Achieving detailed and complete design element contours requires precise control over image segmentation accuracy. A critical challenge in semantic segmentation is extracting contextual information, necessitating the use of convolutional kernels of varying sizes to capture global context. Within this module, outputs from different kernel sizes are sequentially stacked and fused using 1×1 and 3×3 convolutional kernels with dilation rates of 1 and 2, respectively. Finally, feature fusion is

performed using a skip connection approach, where the channel dimensions transition as follows $(5,64) \rightarrow (64,128) \rightarrow (128,128)$.

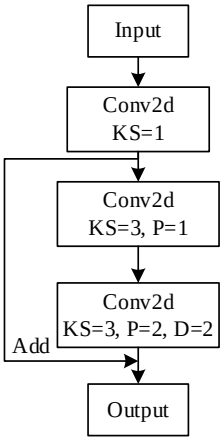

Figure 2: Multi-scale Module Structure

Figure 3 provides a detailed description of the Multi-Scale Encoder (MSE). The MSE is employed to extract image features related to cultural heritage assets at different scales. It adopts a sequential stacking + skip connection architecture, utilizing global average pooling during the downsampling stage to mitigate the additional computational overhead introduced by convolutional operations. This step has negligible impact on the final Mean Intersection over Union (MIoU). Consequently, under such conditions, the number of channels in the feature images remains largely unchanged.

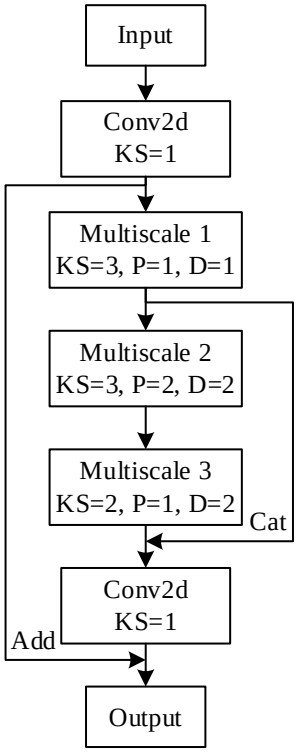

Figure 3: Detailed Description of Multi-scale Encoder

2.3.2 Decoder Architecture Based on Multi-Scale Features

Figure 4 illustrates the structural details of the Multi-Scale Decoder (MSD). Compared to upsampling operations and transposed convolutions, the MSD employs multi-scale features combined with bilinear interpolation to construct its decoder component. The primary reasons are: 1) Upsampling disregards the learnable "parameter" factor, potentially compromising segmentation accuracy in the final results. In contrast, the approach adopted in this paper offers the advantages of parameter-free learning and rapid training speed; 2) Transposed convolution incorporates learnable factors but struggles to ensure alignment with encoder information. Therefore, this paper replaces the conventional transposed convolution step with multi-scale features + bilinear interpolation.

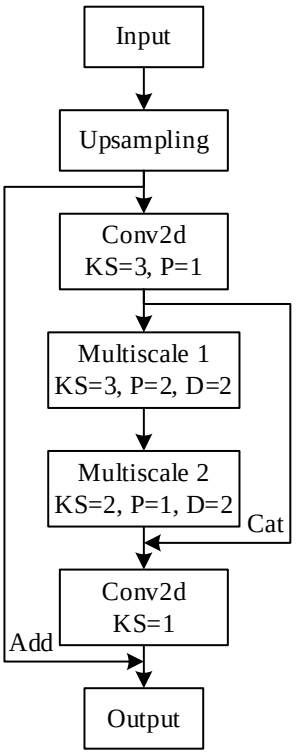

Figure 4: Details of the Multi-scale Decoder Structure

2.3.3 Loss Function Based on Multi-Scale Feature Learning Networks

Analysis of practical scenarios and datasets reveals that the criterion for selecting loss functions is to enhance the generalization capability of neural networks. Key challenges include: 1) blurred boundaries in image design elements; 2) mitigating class imbalance; and 3) optimizing the Mean Intersection over Union (MIoU) metric. To address these challenges, a combination of loss functions is employed to supervise model training, primarily including cross-entropy loss L_{wce}, *Lovász-Softmax* loss, and boundary loss L_{bd}.

In semantic segmentation tasks, two-dimensional images predominantly suffer from blurred object boundaries during segmentation. Therefore, boundary loss L_{bd} is defined as shown in Equation (11):

$$L_{bd}(\hat{y}, y) = 1 - \frac{2P^C \cdot R^C}{P^C + R^C} \tag{11}$$

In the formula, P^C and R^C denote the precision and recall of the predicted boundary feature y_{pd} relative to the true label y_{gt} for category C.

The boundary loss is given by Equation (12):

$$\begin{cases} y_{pd}^b = pool(1-y_{pd},\theta_0)-(1-y_{pd}) \\ y_{gt}^b = pool(1-y_{gt},\theta_0)-(1-y_{gt}) \end{cases} \tag{12}$$

In the formula, $pool(\cdot)$ denotes the max pooling operation on a sliding window of size θ_0.

To address the issue of imbalanced class labels, weights are determined using the reciprocal of the square root of the class occurrence frequency. The $Soft\max$ cross-entropy loss function is given by Equation (13):

$$L_{wce}(\hat{y},y) = -\sum_i \alpha_i p(y_i)\log(p(\hat{y}_i)), \alpha_i = \frac{1}{\sqrt{f_i}} \tag{13}$$

In the equation, y_i and $\hat{y}_i$ represent the true and predicted values of the category labels, respectively; f_i denotes the frequency of occurrence for category label C.

Simultaneously, the *Lovász-Softmax* loss function is employed to maximize the Mean Intersection over Union (MIoU), as shown in Equation (14):

$$L_{ls} = \frac{1}{|C|}\sum_{c\in C}\overline{\Delta J_c}(m(c)), m_i(c) = \begin{cases} 1-x_i(c), & if\ \ c = y_i(c) \\ x_i(c), & otherwise \end{cases} \tag{14}$$

In the equation, $|C|$ denotes the number of category labels; $\overline{\Delta J_c}$ represents the extended term of the Jacobian matrix index; $x_i(c)\in[0,1]$, $y_i(c)\in\{-1,1\}$ represents the true label probability and predicted label probability for the i th pixel of category label C.

Finally, the combined loss function is given by Equation (15), where w_1, w_2, w_3 represent the weights of the different loss functions:

$$L = w_1\cdot L_{ls} + w_2\cdot L_{wce} + w_2\cdot L_{bd} \tag{15}$$

3 Practical Application of Laplace Transform-Based Feature Extraction for Cultural Heritage Images

3.1 Image Segmentation Performance Experiments and Comparisons

3.1.1 Transformer Layer Parameter Experiment

To enhance the accuracy of feature extraction for design elements in digital images of cultural heritage, this study validates the application advantages of combining the Laplacian transform with a multi-scale feature + bilinear interpolation convolution method and loss function model training approach. This section conducts parameter experiments on the Transformer layer to investigate the performance of the designed model in segmentation tasks for digital images of

cultural heritage (using Thangka Buddhist paintings as an example), with particular focus on the impact of image block size and stacking depth within the Transformer layer on segmentation accuracy. Thangka Buddhist paintings, as target cultural design elements featuring complex textures and intricate details, pose challenges to a model's perceptual capabilities and contextual understanding during segmentation. By adjusting the image block size and stacking depth of the Transformer layer, we can gain insights into the model's processing capabilities for images with complex edges and rich textures, identifying optimal model configurations to enhance segmentation accuracy for thangka Buddhist paintings.

Table 1 illustrates the impact of different combinations of image block sizes and Transformer layer stacking numbers on experimental accuracy. The results reveal the influence of both image block size and Transformer layer stacking on Thangka Buddha image segmentation precision. First, examining the effect of varying image block sizes on segmentation accuracy shows that when the block size is 64×64, the model generally achieves better segmentation performance compared to other block sizes. This may stem from the 64×64 size effectively capturing both local image details and preserving contextual relationships, thereby enhancing the model's global perception capabilities. Secondly, increasing the number of stacked Transformer layers generally improves segmentation accuracy. This indicates that expanding the Transformer stack enables the model to better learn long-range dependencies and global information within digital heritage images, ultimately boosting segmentation precision. Comprehensive analysis reveals that the highest mAP (97.79%), mDice (98.83%), and mIoU (96.25%) were achieved with a 64×64 image patch size and 17 stacked Transformer layers. This configuration enables the model to better capture semantic information and contextual relationships within the Thangka Buddha image segmentation dataset, achieving higher segmentation accuracy and enhancing feature extraction for target design elements. In practical complex applications, it is also necessary to comprehensively consider the data characteristics of different cultural heritage digitized images and fine-tune the model configuration to achieve better segmentation results.

Table 1: Experimental results of Transformer layer parameters

Patch size	Stack times of transformer layers	mPA (%)	mDice (%)	MioU (%)
32×32	5	96.56	97.47	94.34
	9	96.92	97.85	94.89
	13	96.86	97.54	94.96
	17	96.78	97.57	95.08
	21	97.14	97.82	95.11
64×64	5	97.17	97.84	95.25
	9	97.26	97.88	95.32
	13	97.42	98.18	95.51
	17	**97.79**	**98.83**	**96.25**
	21	97.42	98.22	95.88
128×128	5	97.12	98.25	93.73
	9	97.28	97.67	93.82
	13	97.25	98.48	94.12
	17	97.28	97.54	94.03
	21	97.47	98.49	94.16

3.1.2 Comparative Experiments with Other Models

Based on the experimental results in Section 3.1.1, we set the image patch size to 64×64 and the number of stacked Transformer layers to 17. We conducted a performance comparison of the designed model against existing semantic segmentation and feature extraction models on

the Thangka Buddha image segmentation dataset to determine whether our model exhibits comparative advantages. Table 2 summarizes the performance comparison results across multiple models. The experimental results reveal the specific metric values achieved by each model in terms of mPA, mDice (Dice score), mIoU, and FPS. In the context of Thangka Buddha image segmentation, the edges of Buddha elements typically exhibit complex textures and intricate details, thereby imposing higher demands on segmentation model performance. Comparing the four metrics, the proposed model outperforms the other six comparison models in mPA, mDice, mIoU, and FPS, achieving 98.76%, 98.17%, 95.89%, and 45.97%, respectively. This demonstrates its capability to achieve higher-precision segmentation results in Thangka Buddha image segmentation tasks, with a certain ability to capture edge details. The computational complexity enhances image processing efficiency while reducing computational costs, offering advantages in real-time segmentation applications. The optimized model is more suitable for Thangka Buddha image segmentation and feature extraction.

Table 2: Performance comparison results of different models

Medol	mPA (%)	mDice (%)	MioU (%)	FPS
UNet	90.23	90.34	90.12	27.31
TransUNet	96.31	95.97	94.03	26.73
SegNet	91.87	91.56	86.01	30.55
SA-UNet	88.25	88.55	87.04	21.06
R2U-Net	90.12	90.21	86.27	22.24
HyperSeg	91.54	90.68	92.45	34.69
Ours	**98.76**	**98.17**	**95.89**	**45.97**

3.2 Verification of Resampling and Segmentation Accuracy for Design Elements by Category

3.2.1 Image Resampling and Statistics for Cultural Heritage

To enhance the accuracy of design element extraction, beyond model optimization, the study also implemented a series of data-level operations, specifically:

1) Counted the number of labeled samples for each ritual implement category in the Thangka Buddha image dataset, classifying categories with fewer than 250 samples as sparse classes.

2) Performed oversampling on categories with fewer samples: First, manually selecting corresponding categories for cropping and rotation operations to increase the number of original images. Second, applying sharpening and enhancement to the original images. Given the characteristics of thangka images, which require preserving their structural features, this study employed random augmentation of saturation, contrast, and brightness for data enhancement. Finally, re-annotating the processed images to maintain the number of labels per category around 250.

3) The number of samples per category after resampling was statistically analyzed, and the data distribution before and after sampling was compared. Figure 5 illustrates the data situation for different categories of ritual implements in thangka Buddha images before and after resampling. The number of labels for the three ritual implement categories—scepter, pipa, and wish-fulfilling jewel—increased from 124, 153, and 176 before resampling to 250 after resampling. The numbers for longevity vases and three-pronged skull forks increased from 195 and 257 to 275. The counts for sutra scrolls, wisdom swords, golden wheels, and ritual bowls remained unchanged at 282, 407, 421, and 465, respectively. Through resampling, the number

of Tangka Buddha segments increased, and the distribution of ritual object labels across categories became relatively balanced. This improved the segmentation accuracy of design elements, facilitating subsequent feature extraction.

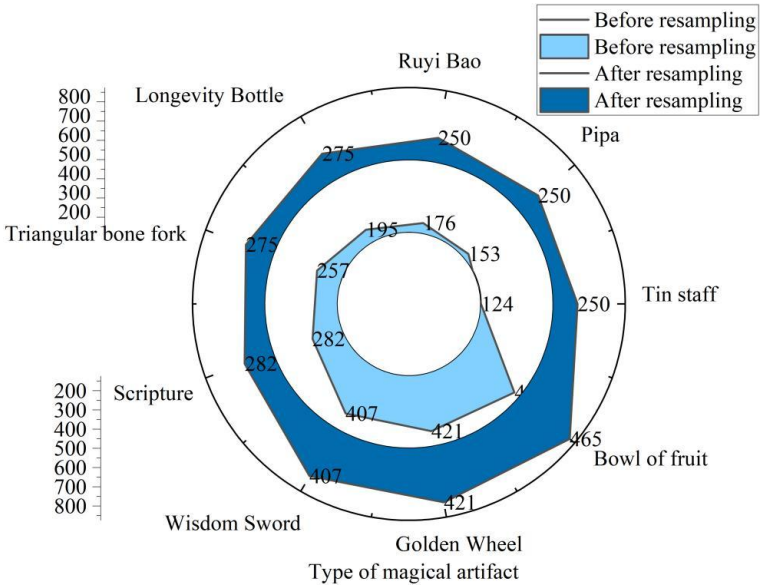

Figure 5: Distribution of different types of talismans before and after resampling

3.2.2 Verification of Category-Specific Feature Segmentation Accuracy Meeting Standards

After optimizing the Thangka Buddha image data, the segmentation accuracy for different ritual implement categories was verified to meet standards, ensuring that segmentation errors would not impact subsequent feature extraction. Table 3 evaluates the segmentation accuracy for each category of ritual implements. For the segmentation of images related to the 9 ritual implements, the Pixel Accuracy (PA) exceeded 85% across all categories, ranging from a minimum of 86.36% to a maximum of 98.23%. The Intersection over Union (IoU) exceeded 80% for all categories, with the highest value of 91.07% achieved for the Golden Wheel. The harmonic mean of precision and recall (F1-Score) also exceeded 85%, ranging from 87.84% to 96.42%. This indicates that through optimized resampling of the data, the general contours of ritual implements across different categories can be effectively separated, ensuring the integrity of element feature extraction.

Table 3: Summary of feature extraction accuracy for all categories

Types	PA (%)	IoU (%)	F1-Score (%)
Tin staff	96.34	83.18	91.86
Pipa	90.05	80.24	90.04
Ruyi Bao	93.11	80.71	89.87
Longevity Bottle	89.19	80.85	89.79
Triangular bone fork	86.36	80.83	87.84
Scripture	98.23	87.52	94.46
Wisdom Sword	94.48	88.76	95.13
Golden Wheel	97.12	91.07	96.42
Bowl of fruit	93.01	87.35	94.38

3.3 Comparison of Feature Extraction Performance Across Different Models

3.3.1 Feature Extraction Accuracy Comparison

The nine categories of thangka images depicting Buddhist deities and ritual implements were divided into three datasets based on the number of resampled labels. Specifically: - Mace, Pipa, and Wish-fulfilling Jewel form Dataset 1 - Long-life Vase, Three-pronged Skull Fork, and Sutra Scroll form Dataset 2 - Wisdom Sword, Golden Wheel, and Dharma Bowl form Dataset 3 Different image feature extraction methods were applied to each dataset, with their accuracy compared. Table 4 presents the accuracy results of various feature extraction methods across the three datasets. Among the 2,480 thangka images of Buddhist statues and ritual implements across the three datasets, the model proposed in this paper is the only one achieving accuracy exceeding 90%, demonstrating optimal performance with rates of 90.37%, 90.86%, and 90.95%, respectively. This indicates that for feature extraction in relatively complex cultural heritage images, the proposed model can more accurately and comprehensively identify the shapes of elements within them.

Table 4: Accuracy results of different methods on three datasets

Method	Epoch	Accuracy (%)		
		Dataset 1	Dataset 2	Dataset 3
BEiT	450	67.37	76.51	85.82
MAE	450	84.39	89.35	88.46
CAE	450	80.62	85.53	87.01
BYOL	450	75.72	84.87	86.94
SimSiam	450	84.61	89.78	88.26
MoCo-v3	450	83.65	80.12	87.33
Ours	450	**90.37**	**90.86**	**90.95**

3.3.2 Feature Extraction Classification Performance Comparison

The feature extraction accuracy results demonstrate that the proposed model significantly outperforms most classical methods. To further evaluate the model's generalization capability and stability, K-Nearest Neighbor classification experiments were conducted with K set to 35 and 85. This approach requires no additional classifier training, performing majority voting predictions directly based on pre-trained features. It effectively measures the model's representation robustness when independent of local feature semantic signals. Table 5 compares the KNN classification performance of different feature extraction methods across three datasets. Without relying on local feature semantic signals, the feature extraction accuracy of the proposed model remains above 90%, peaking at 91.89% and reaching a minimum of 91.26%. This significantly outperforms the BEiT model and slightly exceeds the CAE and SimSiam models. Consequently, the proposed model demonstrates markedly superior effectiveness and robustness compared to other feature extraction methods in the task of learning and extracting features from digital images of cultural heritage.

Table 5: KNN classification performance of different methods on three datasets

Method	Accuracy (%)					
	Dataset 1		Dataset 2		Dataset 3	
	K=35	K=85	K=35	K=85	K=35	K=85
BEiT	58.17	63.92	69.23	69.69	75.21	73.24
MAE	78.23	83.16	84.18	83.58	83.85	81.81
CAE	82.64	89.09	84.66	83.45	87.41	85.15
BYOL	79.21	84.33	86.98	84.13	80.83	78.76
SimSiam	85.85	87.95	88.52	89.99	88.96	88.53
MoCo-v3	85.83	84.31	75.54	77.04	86.12	87.41
Ours	**91.26**	**91.75**	**91.64**	**91.52**	**91.47**	**91.89**

3.3.3 Comparison with Manual Recognition and Extraction Efficiency

To evaluate the efficiency of model-based feature recognition and extraction versus manual recognition and extraction, this paper compares the time consumption of model-based feature recognition and extraction with that of manual recognition and extraction for ritual implements depicted in Thangka Buddhist paintings. Figure 6 presents the time consumption comparison between the two methods. The feature recognition and extraction time for both methods increases with the number of images. However, the model-based recognition and extraction time ranges from [36.53, 75.42] seconds, while manual recognition and extraction takes significantly longer, reaching 101.59–145.28 seconds. Compared to manual recognition and extraction, the model demonstrates higher efficiency in identifying and extracting image features, effectively reducing time costs.

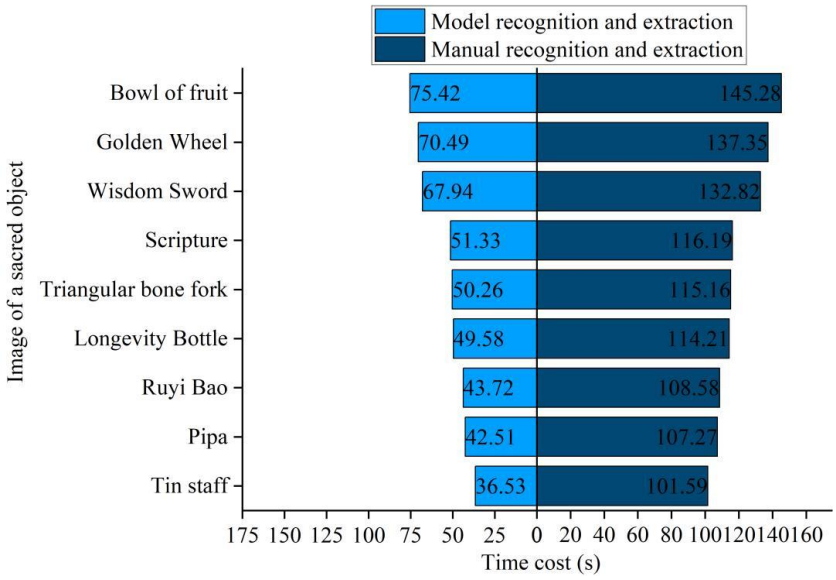

Figure 6: Comparison of time consumption for the two methods

3.4 Ablation Experiments and Effectiveness Analysis

To assess whether each improvement in the proposed Laplace Transform + Multi-Scale Feature Learning Network model contributes positively to the overall network and to more intuitively compare their respective contributions to the entire model, ablation experiments were conducted in this section. Table 6 summarizes the contributions of each improvement to the reconstruction performance. As shown in the first column of Table 6, the ablation experiments

primarily compare the following configurations: the original multiscale block (MSB), the addition of an encoder (MSE), the further addition of a decoder (MSD), and the incorporation of loss function improvements based on this foundation. Ablation tests were conducted at image block resolutions of 32×32, 64×64, and 128×128. Comparing PSNR and SSIM values visually demonstrates that each component improvement enhances network performance. The loss function modification yields the most significant gains: PSNR consistently exceeds 28 (peaking at 30.58), while SSIM surpasses 0.67 (reaching a maximum of 0.8132).

Table 6: Improving the contribution of each part to the reconfiguration effect

Method	Segmentation accuracy	PSNR	SSIM
MSB	32×32	26.41	0.6621
MSB	64×64	27.92	0.7268
MSB	128×128	25.05	0.5973
+MSE	32×32	26.83	0.6661
+MSE	64×64	28.16	0.7295
+MSE	128×128	25.53	0.6002
+MSE+MSD	32×32	28.64	0.6709
+MSE+MSD	64×64	29.76	0.7334
+MSE+MSD	128×128	26.47	0.6027
+MSE+MSD+Loss function	32×32	29.85	0.6713
+MSE+MSD+Loss function	64×64	**30.58**	**0.8132**
+MSE+MSD+Loss function	128×128	28.69	0.7336

4 Conclusion

This paper integrates the Laplace transform algorithm with a multi-scale feature learning network to perform feature extraction and reconstruction of cultural heritage images. Using a 64×64 image block size and 17 stacked Transformer layers improves segmentation accuracy. When the number of resampled labels for different image categories reaches 250 or more, the segmentation accuracy (PA), precision, recall, and harmonic mean (F1-Score) all exceed 85%, while the intersection-over-union (IoU) consistently surpasses 80%. Across three Thangka Buddha image datasets, the proposed method achieved feature extraction accuracies of 90.37%, 90.86%, and 90.95%, demonstrating robust performance. Compared to manual identification and extraction (0 efficiency, 101.59s–145.28s), this method exhibits faster feature learning and extraction speeds (36.53s–75.42s). Ablation studies confirm that PSNR and SSIM peak values reach 30.58 and 0.8132, respectively. Our optimization approach achieves effective extraction and reconstruction of design elements in cultural heritage, supporting the permanent digitization of outstanding cultural assets.

About the Author

WeiYuan Liu was born in Tianmen, Hubei. in 1980.He is currently an associate professor at the School of Art and Design, Guilin University of Electronic Technology, China, and serves as a supervisor for Master of Arts candidates. He served as the Director of the Product Design Department from 2018 to 2020. He received his education at Guilin University of Electronic Technology, and his research interests include structural design, and product innovation design. E-mail: lwy@guet.edu.cn

Sirui Chen was born in Zhangjiakou, Hebei. He is currently a postgraduate student in the School of Art and Design at Guilin University of Electronic Technology, Guilin, China. Her research interests include design engineering, seismic-resistant product development and human-centered design. E-mail:2108256251@qq.com

Yanchun Jiang,Female, Associate Chief Physician, Doctor of Medicine, Born on March 4, 1985, Works at the Department of Neurology, Nanxishan Hospital of Guangxi Zhuang Autonomous Region. E-mail:18632332365@163.com

Funding

This work was by supported by 2023 Guangxi Philosophy and Social Science Research Project (Approval Number: 23FWY022), Project Title: Research on the In-depth Development and Design Strategies of Experiential Tourism Products in Ancient Towns of Guangxi.

Sichuan Province Key Research Base Project for Philosophy and Social Sciences in Higher Education Institutions (GYSJ2023-29), Project Title: Research on the Construction of the Index System for Rural Tourism Experience Elements in the Greater Guilin Area under the Background of Rural Revitalization.

References

[1] Cerisola, S. (2019). A new perspective on the cultural heritage–development nexus: The role of creativity. Journal of Cultural Economics, 43(1), 21-56.

[2] Ferrazzi, S. (2021). The notion of “cultural heritage” in the international field: Behind origin and evolution of a concept. International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique, 34(3), 743-768.

[3] Li, L., & Tang, Y. (2023). Towards the contemporary conservation of cultural heritages: An overview of their conservation history. Heritage, 7(1), 175-192.

[4] Portalés, C., Rodrigues, J. M., Rodrigues Gonçalves, A., Alba, E., & Sebastián, J. (2018). Digital cultural heritage. Multimodal Technologies and Interaction, 2(3), 58.

[5] Fontanive, S. (2024). Scanning as a form of cultural heritage preservation. Jornal da Universidade, 11 jan. 2024, n. 176.

[6] Wijesundara, C., & Sugimoto, S. (2018). Metadata model for organizing digital archives of tangible and intangible cultural heritage, and linking cultural heritage information in digital space. Library and Information Science Research E-Journal.

[7] Liao, H. T., Zhao, M., & Sun, S. P. (2020, May). A literature review of museum and heritage on digitization, digitalization, and digital transformation. In 6th International Conference on Humanities and Social Science Research (ICHSSR 2020) (pp. 473-476). Atlantis Press.

[8] Barrile, V., Bilotta, G., & Lamari, D. (2017). 3D models of Cultural Heritage. International journal of mathematical models and methods in applied sciences, 11(2017), 1-8.

[9] Xiao, W., Mills, J., Guidi, G., Rodríguez-Gonzálvez, P., Barsanti, S. G., & González-Aguilera, D. (2018). Geoinformatics for the conservation and promotion of cultural heritage in support of the UN Sustainable Development Goals. ISPRS Journal of Photogrammetry and Remote Sensing, 142, 389-406.

[10] Georgopoulos, A. (2018). Contemporary digital technologies at the service of cultural heritage. In Heritage Preservation: A Computational Approach (pp. 1-20). Singapore: Springer Singapore.

[11] Paschalidou, E., Fafet, C., & Milios, L. (2022). A strong sustainability framework for digital preservation of cultural heritage: introducing the eco-sufficiency perspective. Heritage, 5(2), 1066-1088.

[12] Zheng, Y. (2022). Study on the application of Chinese traditional visual elements in visual communication design. Mathematical Problems in Engineering, 2022(1), 1020033.

[13] Zhao, R. (2021, May). New thinking on visual communication design under the background of digital age. In 2021 International Symposium on Artificial Intelligence and its Application on Media (ISAIAM) (pp. 82-85). IEEE.

[14] Li, G. (2023). A Visual Communication Design Study: Graphic Element Design Under Traditional Handwork. Journal of Information Processing Systems, 19(2).

[15] Huang, W., Xiang, H., Liang, W., Li, X., Lai, G., & Lin, Y. (2018, December). The design of visual communication for intangible cultural heritage in product form. In 2018 2nd International Conference on Education Innovation and Social Science (ICEISS 2018) (pp. 294-298). Atlantis Press.

[16] Xu, M. (2024). Construction and Research of Ethnic Cultural Elements in Visual Communication. Journal of Humanities, Arts and Social Science, 8(3).

[17] Luigini, A., Massari, G. A., Vattano, S., Pellegatta, C., & Luce, F. (2018, July). Visual Culture and Cultural Heritage: ViC-CH a Synthesis Between Digital Representation and Heritage Experience. In International and Interdisciplinary Conference on Digital Environments for Education, Arts and Heritage (pp. 288-302). Cham: Springer International Publishing.

[18] Verma, D. (2019). A Laplace Transformation approach to Simultaneous Linear Differential Equations. New York Science Journal, 12(7), 58-61.

[19] Rezvanifard, F., & Radmehr, F. (2024). Laplace Transform in Mathematics and Electrical Engineering: A Praxeological Analysis of Two Textbooks on the Differential Equations and Signal Processing. IEEE Transactions on Education, 67(4), 508-518.

[20] Jacobs, B. A., & Harley, C. (2018). Application of Nonlinear Time-Fractional Partial Differential Equations to Image Processing via Hybrid Laplace Transform Method. Journal of Mathematics, 2018(1), 8924547.

[21] Rahmawati, L., Rustad, S., Marjuni, A., Arief, M., Soeleman, C. S., & Shidik, G. F. (2024). Transmission Map Refinement Using Laplacian Transform on Single Image Dehazing Based on Dark Channel Prior Approach. Cybernetics and Information Technologies,

24(4).

[22] Zhang, Q., Cui, S., Liu, L., Wang, J., Wang, J., Zhang, E., ... & Liang, H. (2021, July). Deep learning for the extraction of sketches from spectral images of historical paintings. In Optics for Arts, Architecture, and Archaeology VIII (Vol. 11784, pp. 11-20). SPIE.

[23] Hu, Y., Yu, S., Qin, S., Chen, D., Chu, J., & Yang, Y. (2021). How to extract traditional cultural design elements from a set of images of cultural relics based on F-AHP and entropy. Multimedia tools and applications, 80(4), 5833-5856.

[24] Qiu, S., Zhang, P., Li, S., & Hu, B. (2023). Extraction and analysis algorithms for Sanxingdui cultural relics based on hyperspectral imaging. Computers and Electrical Engineering, 111, 108982.

[25] Zhang, Z., Wan, Z., & Zhu, H. (2025). Intelligent Extraction of Intangible Cultural Heritage Visual Elements and Application in Graphic Design Based on Generative Adversarial Network. International Journal of Information Technologies and Systems Approach (IJITSA), 18(1), 1-16.

[26] Jiang, J. (2025). Study on the Path of Visual Symbol Extraction and Design Transformation of Traditional Culture in Western Guangdong by Integrating Computer Vision Technology under the Background of Digital Intelligence. Journal of Network and Innovative Computing, 13, 25-38.

[27] Hu, C., Han, Y., Xia, G., Wang, Y., & Ma, X. (2025). Research on the extraction of inscription rubbings information of stone carved cultural relics based on 3D fine models. npj Heritage Science, 13(1), 36.

[28] Belhi, A., Al-Ali, A. K., Bouras, A., Foufou, S., Yu, X., & Zhang, H. (2020). Investigating low-delay deep learning-based cultural image reconstruction. Journal of Real-Time Image Processing, 17(6), 1911-1926.

[29] Nasri, A., & Huang, X. (2020). A missing color area extraction approach from high-resolution statue images for cultural heritage documentation. Scientific Reports, 10(1), 21939.

[30] Song, C., & Wu, X. (2022). Research on virtual color restoration of complex building system based on discrete wavelet transform. Discrete Dynamics in Nature and Society, 2022(1), 6450780.

[31] Süvari, A., Okuyucu, Ş. E., Çoban, G., & Eren Tarakci, E. (2023). Virtual reconstruction with the augmented reality technology of the cultural heritage components that have disappeared: The Ayazini Virgin Mary Church. ACM Journal on Computing and Cultural Heritage, 16(1), 1-16.

[32] Li, X., & Chen, M. (2024). A Semiotics-Based Visual Reconstruction Method for Intangible Cultural Heritage Images: A Case Study of China's Hainan< Geng Lu Bu>. Asia-pacific Journal of Convergent Research Interchange (APJCRI), 569-585.