Available online at www.HighTechJournal.org

HighTech and Innovation Journal

ISSN: 2723-9535

Vol. 7, No. 1, March, 2026

Design of Miao Embroidery Gene Decoding and Digital Activation Based on CNN-Transformer

Yuchun Huang [1*]

[1] *School of Design Art, Xiamen University of Technology, Xiamen, 361024, China.*

Received 21 October 2025; Revised 22 January 2026; Accepted 29 January 2026; Published 01 March 2026

Abstract

To overcome the challenges of decoding pattern genes and the lack of innovative designs in the digital transformation of Miao embroidery as an intangible cultural heritage, a deep learning fusion architecture based on a convolutional neural network Transformer (CNN Transformer) is proposed. This architecture aims to accurately analyze the semantics and generate stylized Miao embroidery patterns. This architecture uses a ResNet-50 backbone network to extract local pattern details and combines an improved atrous spatial pyramid pooling (ASPP) module to preserve edge structures for semantic segmentation. It simultaneously utilizes a progressive Transformer encoder and multi-scale channel attention to dynamically focus on the pattern subject, fuse features, and achieve style transfer. The experimental results showed that the segmentation accuracy of the model on four types of embroidery patterns exceeded 85% (up to 92.68%), and the structural similarity (SSIM) reached 0.957. In the style transfer task, the style fidelity index (KL divergence 0.52, Gram matrix MSE 0.019) and average accuracy (AP 0.926) of this method were significantly better than the comparison algorithms. This research method effectively solves the problems of inaccurate segmentation of complex patterns and distortion of cultural symbols in traditional style transfer methods. It provides a technical reference for analyzing and innovatively activating Miao embroidery culture.

Keywords: Intangible Cultural Heritage; Miao Embroidery; CNN; ResNet-55; Transformer; Semantic Segmentation; Style Migration.

1. Introduction

Under the background of the digital development process of cultural industry, accelerating the protection of non-heritage cultural traditions with the help of digital media technology plays an important role in enhancing the cultural soft power and promoting the development of non-heritage innovation industry. As a typical representative of traditional handicrafts of Chinese ethnic minorities, Miao embroidery carries rich ethnic cultural genes in its unique pattern symbols, color system, and craft techniques, and has high artistic value and cultural research significance [1, 2]. Miao embroidery is a Miao folk embroidery technique. Its pattern shape, texture, color, and other contents carry their unique understanding of life and art, and it is also the cultural carrier for them to express their love and faith. Among them, Miao embroidery patterns are rich and diverse in modeling features, and their forms and connotations are rooted in specific cultural contexts. Therefore, it is relatively fixed in the overall style expression [3]. Miao embroidery colors are characterized by high brightness and high saturation. Its modeling is related to the embroiderer's subjective aesthetic sense. There are also differences in the expression and modeling presented by different embroidery techniques. For example, flat embroidery is mostly used to lay out large patterns, while overlay embroidery is mostly used to create a three-dimensional sense of hierarchy. The living heritage and development of intangible cultural heritage cannot be separated from the cultural

* Corresponding author: 2024990045@xmut.edu.cn; hyc123456u@outlook.com

https://doi.org/10.28991/HIJ-2026-07-01-08

subject of “people”. At present, Miao embroidery intangible cultural heritage in the development of inheritance is faced with problems such as loss of skills, lack of innovation, low market awareness, etc., which is threatened to be lost [4, 5]. The genetic decoding of Miao embroidery in the context of transformation of non-heritage digital intelligence refers to the digital analysis and structured characterization of the craft characteristics and aesthetic laws of Miao embroidery patterns. Activation design is based on the results of gene decoding, to realize the redesign of the content of Miao embroidery in terms of style migration and element reorganization, and to realize the feasibility of its commercial packaging design. The application of digital technology to the innovative redesign of Miao embroidery is to meet the needs of permanent protection and innovative inheritance of Miao embroidery. It is also an important technical way to realize the decoding and activation design of its cultural gene.

To address the above challenges, many scholars have attempted to explore the digital path of intangible cultural heritage. Early work focused primarily on static digital archiving. This included color extraction, shape display, and pattern contour extraction of textiles via high-precision scanning and photography [6]. This method effectively preserves data permanently, but it essentially builds a "digital museum." This method lacks in-depth analysis and activation of cultural connotations and has limited potential for commercial transformation. In recent years, with the development of artificial intelligence technology, researchers have begun to attempt deeper levels of digital analysis. For example, in the digital research of Zhuang brocade intangible cultural heritage, Ni et al. constructed a feature dataset and implemented automatic generation of patterns and colors through algorithms. This verified the feasibility of feature extraction [7]. Ji et al. examined the intricate design of Zhuang clothing, incorporating convolutional neural networks and attention mechanisms to enhance the classification and recognition accuracy of local features [8]. In addition, Zhao utilized 3D modeling and animation technology to construct an interactive display system for Miao ethnic hand weaving skills, enhancing the user experience [9]. Liu et al. proposed adding a pre-trained Markov adversarial network to a recurrent generative adversarial network (GAN) to enhance texture constraints and improve the realism of digital embroidery. This approach solved the texture blur problem in existing style transfer models. The results showed that this method could learn image style features better and texture details clearer than convolutional neural networks [10].

Although the above research has made positive progress, a careful review of the published literature reveals certain gaps in the digitization of intangible cultural heritage, particularly regarding the digital protection of Miao embroidery. One of them is the one sidedness and semantic gap in feature extraction. Most existing methods focus on single-dimensional features. Traditional image processing methods have difficulty capturing the complex, high-level symbolic semantics behind Miao embroidery patterns. Methods based on a single deep learning model are limited by their inherent local receptive fields, making it difficult to capture the global spatial layout and combination rules of pattern elements. The second reason is that current research tends to view the analysis and design of intangible cultural heritage independently. Although style transfer models can generate images with embroidery styles, they lack structured adherence to the original patterns' rules, which often leads to texture blurring, detail distortion, and the problem of being "similar in shape but not in spirit". Users are unable to perform controllable and interpretable redesign, greatly limiting the effectiveness and cultural fidelity of the design. Miao embroidery patterns contain complex semiotics and aesthetic rules, and traditional manual analysis and design are not only inefficient, but also highly subjective. The existing digital tool chain has failed to provide intelligent support throughout the entire process, from cultural symbol analysis to innovative design output. This prevents it from meeting the needs of the cultural and creative industry for efficient and accurate design tools.

Therefore, in order to construct an intelligent decoding and revitalization design method for Miao embroidery cultural genes, and to improve the intelligence and digital preservation and inheritance of Miao embroidery designs, the study utilizes the residual structure in the convolutional neural network (residual network with 50 layers, ResNet-50) to extract the local texture and color features of the Miao embroidery tattoos. Moreover, the study designs a two-stage progressive Transformer encoder structure to realize the global spatial relationship grasp of the symbols of Miao embroidery patterns. It can realize multi-level style design migration and provide intelligent design tools for cultural and creative industries. Based on the basic framework of convolutional neural network Transformer (CNN-Transformer), the study designs a fusion architecture with improved methodology. The study transforms the “genetic decoding” of Miao embroidery into computable semantic segmentation and feature mapping, and the “revitalization design” into data-driven style generation. This convergence architecture idea can realize the whole chain support from “cultural symbol analysis” to “innovative design output” of Miao embroidery. The research aims to promote the transformation of non-heritage digital intelligence from “preservation” to “creation”. It can provide a new paradigm for the deep integration of traditional culture and modern design.

2. Related Works

Aiming at the limitations of fragmentation and single textual characterization of NRM data, Fan et al. proposed to construct a multimodal knowledge graph integrating text and images. In this study, representative images were screened by denoising algorithm, and the atlas data were constructed by combining global/local visual features and text features. The results indicated that it could effectively enhance the public's visual cognition of NRH and provide a new paradigm for multimodal cultural heritage digitization [11]. To better deconstruct ethnic embroidery styles, Du et al. proposed to implement their storage and retrieval with the help of knowledge graphs. The research found that the development of embroidery was influenced by both internal (gradual) and external (sudden) stimuli, which provided a new perspective of NRH evolution [12]. Aiming at the semantic understanding limitations of non-heritage knowledge retrieval, Xu et al. proposed a cloud brocade question-and-answer system fusing knowledge graph and retrieval enhancement generation techniques. The study adopted an improved pre-training model to achieve text vectorization, combined with a similarity search database to optimize semantic matching. The results indicated that the system could effectively handle fuzzy queries, generate logical answers, and provide an intelligent tool for non-heritage knowledge management [13]. In response to the difficulty of quantitatively analyzing the traditional clothing color system, AI proposed a three-stage clustering method to decode the color evolution law of Changning Miao clothing and establish a quantitative model that connects visual perception and cultural selection [14]. In terms of semantic extraction and the redesign of cultural connotations, Zhao and Xin proposed a method to address the issue of cultural connotations being easily overlooked in the redesign of embroidery patterns. This method integrated extensible semantic analysis and Pix2Pix models to achieve the entire process from cultural feature quantification to innovative pattern generation [15].

To solve the problem of 3D digitization of apparel, Yu et al. developed an autonomous scanning system based on an offline point cloud algorithm. The study combined Laplace mesh deformation with multi-software synergy to realize the transition from 2D restoration to 3D display. The results indicated that the deviation between the 3D model and the actual costume was only 0.5 cm, which verified the accuracy and practicality of the technology in digital restoration of costumes and provided a new paradigm for cultural heritage protection [16]. In the context of generative design, Xu et al. proposed a parallel convolutional neural network model (PSO-PCNN) to promote morphological innovation in traditional craftsmanship. This model effectively translates the cultural symbols of the Maonan Flower Bamboo Hat [17]. Kang et al. used a combination of the diffusion model and the fuzzy comprehensive evaluation method to design sustainable, innovative Miao ethnic wax printing patterns. They then screened the patterns according to user preferences, bridging the gap between intelligent generation and market application [18]. In addition, Zhong et al. developed a GAN restoration framework that integrated spatial channel attention mechanism to address the challenge of physical damage repair of intangible cultural heritage. This framework efficiently restored complex embroidery patterns and provided a new tool for digital restoration and preservation of intangible cultural heritage [19].

To achieve color and contour feature recognition of embroidery images, Ju et al. proposed intelligent recognition with the help of K-means clustering and Canny operator. The results indicated that the method could effectively identify the color error and accurately obtain the contour of the embroidery image [20]. Zhang et al. constructed a temporal multilayer network for co-occurrence of embroidery patterns and detected and analyzed the pattern nodes with the help of link clustering algorithm. This study found that embroidery patterns had scale-free properties and could effectively reflect community culture [21]. Facing the Miao embroidery heritage dilemma, Na et al. used computer graphics to extract pattern color genes. The study synthesized phenomenology and quantitative analysis to construct a digital design process. The results indicated that the method could successfully realize the modern translation of traditional patterns and promote the cross-border integration of non-heritage and digital technology [22]. Cong et al. proposed an intelligent design model for clothing patterns based on deep learning and interactive genetic algorithm. The results indicated that the geometric pattern deviation under this method was reduced and the creativity score was improved by more than 5%, which could effectively satisfy the design efficiency and aesthetic value [23]. To achieve accurate color separation of multicolor printed fabrics, Qian et al. designed a two-level segmentation algorithm with self-organizing mapping and efficient density-based subspace clustering algorithm. The study automatically determined the clustering center by contour coefficients, and post-processing eliminated edge mis-segmentation. The results indicated that the algorithm achieved 88.3% color separation accuracy for complex patterns, which significantly improved the adaptability for industrial applications [24]. Aiming at the problem of insufficient interpretability of deep learning models, Chiu et al. developed a hybrid model combining target detection and cluster analysis. The results indicated that the accuracy of the method on electronic embroidery detection was more than 95%, and the detection efficiency was improved by more than 20%. The method could provide key technical support for the digital transformation of manufacturing industry [25].

The mainstream ideas of current non-heritage digitization research focus on multimodal knowledge graph construction, intelligent retrieval enhancement techniques, feature extraction, and generation algorithms. Although these methods have made progress in knowledge structuring, semantic understanding, and visual restoration, there are still problems such as single representation, weak dynamic adaptability, and loss of stylistic details texture. It is difficult to effectively represent the semantic content and cultural logic of intangible cultural heritage. Therefore, the study proposes to design an improved collaborative model to realize Miao embroidery activation design and feature extraction with the help of improved CNN-Transformer. This study aims to provide a new paradigm for digital intellectualization of NRH that combines technical adaptability and cultural interpretability.

3. Genetic Decoding and Activation Design of Miao Embroidery Under Transformation of Non-Heritage Digital Intelligence

The study proposes to implement the improved CNN-Transformer fusion architecture for Miao embroidery gene decoding and activation design from two aspects: image semantic segmentation and style migration. In the part of image semantic segmentation, the study realizes image semantic feature extraction with ResNet-50 as the backbone network in CNN. Moreover, the Atrous spatial pyramid pooling (ASPP) and spatial channel message enhancement (CME) modules with incentive branching improvement are introduced to realize feature fusion. In the style migration section, the study proposes using a two-stage progressive Transformer design to implement the scale network feature tattoo style processing. This design fuses edge features with multi-scale channel attention features and enhances feature information with a bidirectional multi-stage aggregation decoder.

3.1. The Design of Multi-Scale Feature Semantic Segmentation Based on Genetic Decoding of Miao Embroidery Tattoos

Miao embroidery “living culture” non-genetic inheritance mainly rely on the main body. Changes in ideological concepts, “flux” in inheritance techniques and semantics, and changes and impacts in the geographic environment have all made the handmade Miao embroidery in Qiandongnan face difficulties in preservation. It is difficult to realize the complete preservation and digital storage of some of the handmade Miao embroidery films [26]. Miao embroidery gene decoding refers to the structured characterization of the symbolic features of Miao embroidery by means of digital technology. Considering the differences in the characteristics of Miao embroidery images, the study proposes to extract the micro-process characteristics and geometric topology of the Miao embroidery “gene” with the help of ResNet-50 in CNN. This can preserve its unique features and cultural characteristics. A multi-scale semantic segmented network architecture is also designed to realize the preservation of image tattoos. Figure 1 shows the schematic diagram of segmented network architecture design.

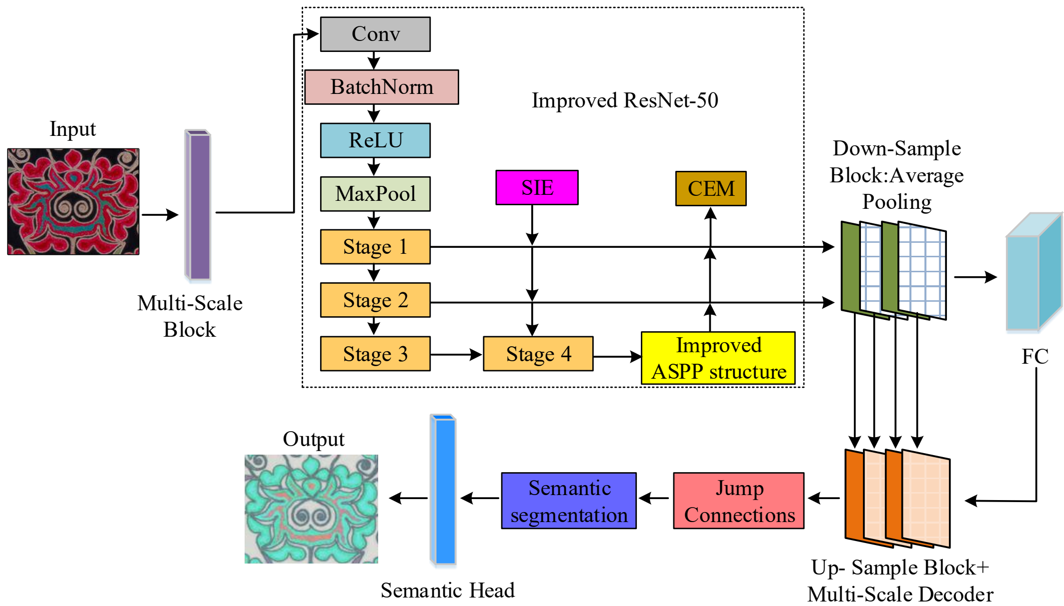


Figure 1. Schematic diagram of segmented network architecture design

In Figure 1, the backbone of this network structure includes an encoder and a decoder. Among them, the encoder backbone network is a ResNet-50 residual network. It can downsample the input image, and the resulting image features can be pooled at different scales. To maximize the retention of the edge information of the image, the decoder uses continuous up-sampling to recover the image features. The final image result is output with the help of hopping connection and semantic segmentation. The multi-scale module of the network structure designed in the study is mainly used to obtain the fused image after channel increase with the help of different convolutional kernel superposition and expansion processes. Multi-scale encoder achieves extraction of features at different scales with the help of sequential stacking and jump joins. Its global average pooling avoids additional computational overhead. The parameter-free upsampling operation of the decoder improves the training speed. Moreover, its combined placement of multi-scale features and inverse convolution is different from the common traditional convolution, which can effectively realize the alignment with the encoder information [27]. Among them, the ResNet-50 network consists of a set of 50 convolutional layers. It can solve the gradient vanishing and gradient explosion problems of the network with the help of residual module structure and jump connection, which requires less network parameters and computation [28]. Equation 1 shows the residual Y representation of the feature map of ResNet-50.

$$Y = F(X) + X \tag{1}$$

In Equation 1, Y is the original feature map and $F(X)$ is the residual function. In the case of residual block stacking, ResNet-50 has better feature extraction ability. Different from recovering the image resolution with the help of bilinear interpolation, the study inserts new elements in the middle of the neighboring elements of the feature map in an inverse convolutional manner. Moreover, adaptive learning is carried out with the help of convolutional kernel in order to realize the integration of the feature map corresponding to the encoder. Meanwhile, in order to better integrate the feature information of different scales, the study proposes to enhance the sensory field and reduce the computation with the help of ASPP. ASPP consists of multiple parallel null convolutional layers and a global average pooling branch. It captures multi-scale features and maintains the feature map resolution. All branch outputs are spliced in channel dimension and later processed by convolution to achieve dimensionality reduction [29]. The mathematical expression of ASPP is shown in Equation 2.

$$y = Concat[Conv1 * 1(x)\ Conv_{3*3}^{r=d}(x)\ GAP(x)] \tag{2}$$

In Equation 2, x is the input feature map. y is the output feature map. $Conv_{3*3}^{r=d}$ is the 3*3 convolution with a null rate of d. GAP is global average pooling. $Conv1 * 1(x)$ is the 1*1 convolution operation. Meanwhile, in order to avoid the difficulty of the fixed expansion rate setting of ASPP to adapt itself to different scales of texture structures, the study adds different ratios to the pooling layer of the ASPP spatial pyramid to avoid the loss of tiny detail features. Moreover, the excitation branch is added to enhance the semantic scale to recognize the image category information. Figure 2 shows the modified ASPP module structure.

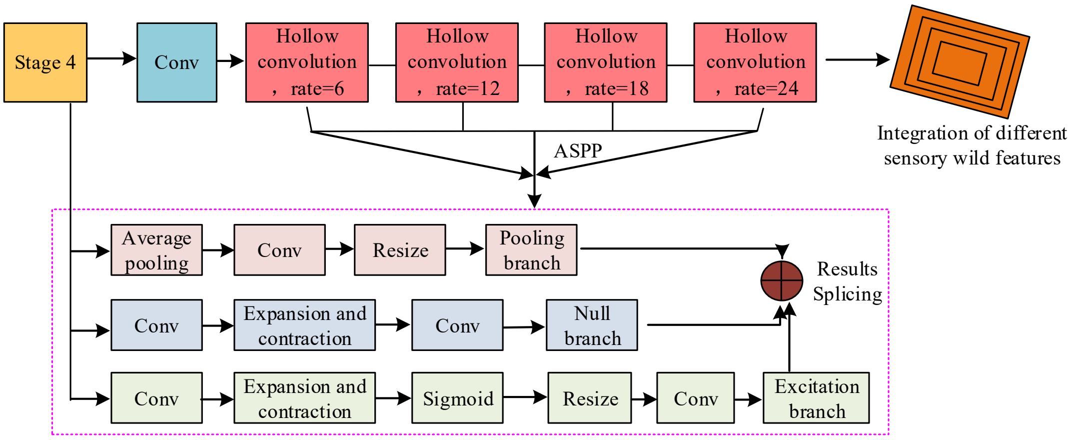


Figure 2. Modified ASPP module structure

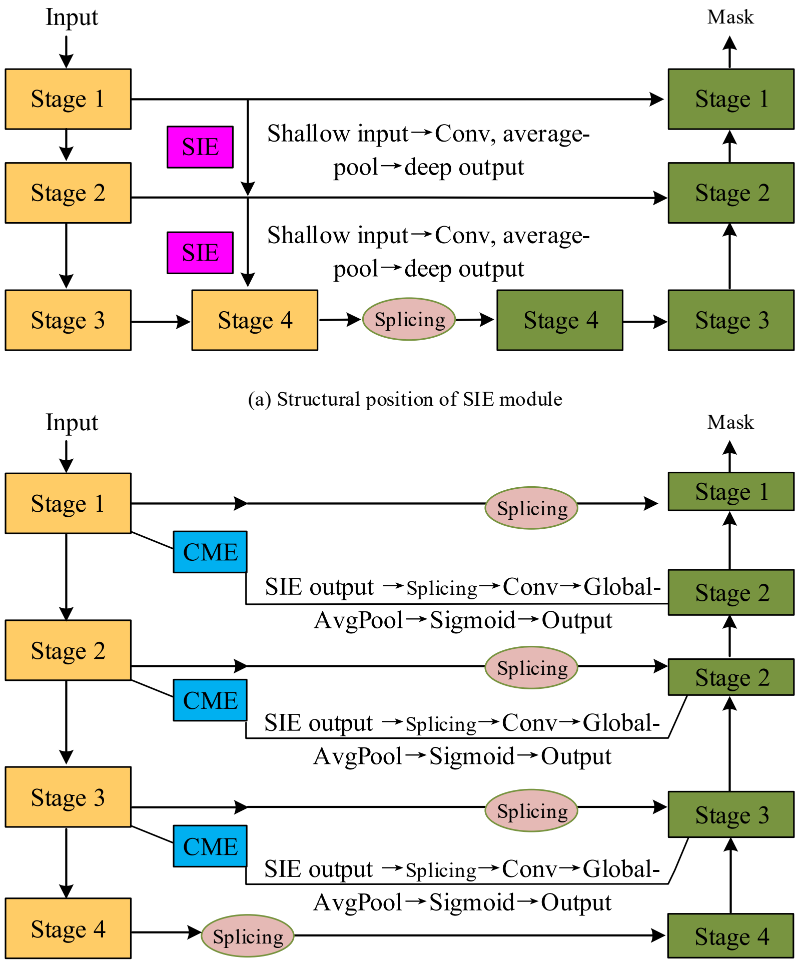
In Figure 2, the modified ASPP module structure includes feature mapping, pooling branch output, null branch output, and excitation branch output. After the encoder enters the null branch, it will first perform 1*1 convolution processing to reduce the number of channels. After that, the pooling branch will process it and achieve cross-channel information integration. The integrated feature information acquires a larger receptive field in the excitation branch. The study negatively weights the channel features with the help of activation functions to form dense contextual information. Considering the uneven distribution of spatial and semantic information in different stages of the feature map, jump-joining, and inverse convolution are not enough to extract high-quality feature information [30]. Therefore, in order to enhance the semantic information of channel information, space information enhancement (SIE) module and CME module are introduced in the coding part of the study to realize it. Equation 3 is the mathematical expression of the SIE high-resolution feature map output result S_{i+1}^{out}.

$$S_{i+1}^{out} = Cp = 2^{3\times3}(S_i^{in}) + Avg^{p=2}(S_i^{in}) \tag{3}$$

In Equation 3, i is the number of encoder layers. S_i^{in} is the high-resolution feature map input value. $Cp = 2^{3\times3}$ is the 3×3 convolution with step size 2. Avg is the average pooling layer. p is the step size. The CME module mainly squeezes the coded feature information to improve the feature recognition performance of the decoder. Its output result C_{j+1}^{out} is mathematically expressed as Equation 4.

$$C_{j+1}^{out} = CME(Decoj + 1^{out}\ Encoj^{out}) \oplus SIEoutput \tag{4}$$

In Equation 4, $Decoj + 1^{out}$ is the output value of decoding $j + 1$ stage. $Encoj^{out}$ is the output value of encoding j stage. $SIEoutpu$ is the SIE output of the encoding stage. $\oplus$ is the multiplication symbol. Figure 3 shows a schematic diagram of the location of SIE and CME modules in the network.



(b) Structural position of CME module

Figure 3. Schematic diagram of the location of SIE and CME modules in the network

In Figure 3, the SIE module is used twice in the middle of the first and third stages of the encoder to better recover the image details, respectively. CME can realize the aggregation of channel and context information by global average pooling of the squeezed feature maps, and generate the attention map weights to obtain deeper semantic information. To address the common problems of blurred boundaries and category imbalance in image semantic segmentation, the study proposes to implement model training with a combined loss function, as shown in Equation 5 [31].

$$L = w1 \cdot Lls + w2 \cdot Lwce + w2 \cdot Lbd = w1 \cdot \frac{\sum_{c\in C}\overline{\Delta Jc}(m(c))}{|C|} + w2 \cdot [-\sum_i \alpha ip(yi)\, log(p(\hat{y}i))] + w2 \cdot (1 - \frac{2P^C \cdot R^C}{P^C + R^C}) \tag{5}$$

In Equation 5, $Lwce$ is the cross-entropy loss. Lbd is the boundary loss. Lls is the maximum cross-merge ratio. Among them, C is the category. P^C and R^C are the precision and recall of boundary features with true labels. yi and $\hat{y}i$ are the true and predicted values of category labels. $|C|$ is the number of category labels. ΔJc is the expansion term of the Jacobi matrix index. $w1$ and $w2$ are the weight values of the corresponding loss functions. p is the probability, αi is the weight coefficient, and $m(c)$ is the pixel prediction error vector for category c. ResNet-50 has been widely validated and proven to provide a very robust baseline between feature extraction capability and computational cost, which is stable and reliable. The EfficientNet model's core advantage is its ability to significantly reduce computational and parameter complexity while maintaining high accuracy. This makes it ideal for deployment on mobile devices or in scenarios with limited resources. When designing research questions, the ability of the model to represent features and its stability should be prioritized over ultimate efficiency. ConvNeXt backbone networks and others have adopted design ideas from Transformers, which could result in redundant models. The research on Transformer encoders aims to better distinguish the responsibilities of CNNs and Transformers. CNNs focus on local information, while Transformers focus on global information.

3.2. Designing a Living Expression of Miao Embroidery Style Migration Based on the Transformer Architecture

In the foregoing, the study extracts the sample data features with the help of semantic segmented network. The ResNet 50 network results can capture different levels of embroidery details. To further realize the hierarchical parsing of the semantics of Miao embroidery pattern and to improve the flexibility of its feature style activation design, the study fuses its features with the help of style migration network model to realize the digital representation of the image. The past style migration network model often faces limitations such as small sensory field and deep network layers when realizing feature extraction of content images [32]. To improve the performance of the model, the study introduces the multi-head attention mechanism in Transformer architecture into ResNet 50 to realize the parallel processing of image features [33]. Figure 4 shows the Transformer architecture model.

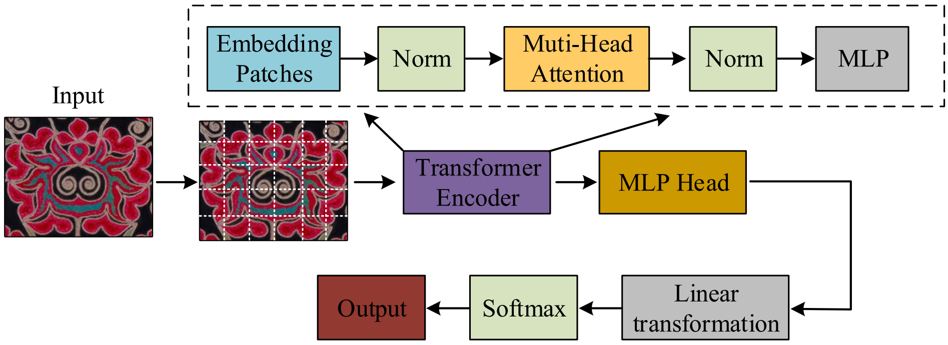


Figure 4. Transformer architecture model

In Figure 4, Transformer divides the input 2D image into different blocks. After that, the spatial location information of the blocks is preserved with the help of linear transformations and embedded into the encoder for normalization, multi-head attention, and multi-layer perceptron processing. The core module of Transformer uses multi-head attention by computing multiple independent single-head self-attention units in parallel. It also outputs the final feature vector with the help of splicing and linear transformation. The input embedding in each self-attention unit includes three vectors of query Q, key K, and value V [34]. Equation 6 shows the output process of MHSA.

$$\begin{cases} Z_i = SelfAtten(X'W_i^Q\ X'W_i^K\ X'W_i^V) \\ MultiHeadAtten(Q\ K\ V) = concat(Z_1\ Z_2\ Z_3\ ,,,\ Z_h)W^O \end{cases} \tag{6}$$

In Equation 6, h is the number of attention heads. W^O is the output mapping matrix. is the output vector. W_i^Q is the mapping matrix of Q. W_i^K is the mapping matrix of K. W_i^V is the mapping matrix of V. X' is the input embedding vector. Miao embroidery patterns have rich color gradient or texture density variation. Therefore, the study is based on Transformer and proposes to extract features from the main body and local part of the texture with the help of two-stage gradient idea. This can reduce the interference of the background on the contour features and realize the fusion of edge features with the help of multi-scale channel attention features. Figure 5 shows the schematic structure of the progressive Transformer model.

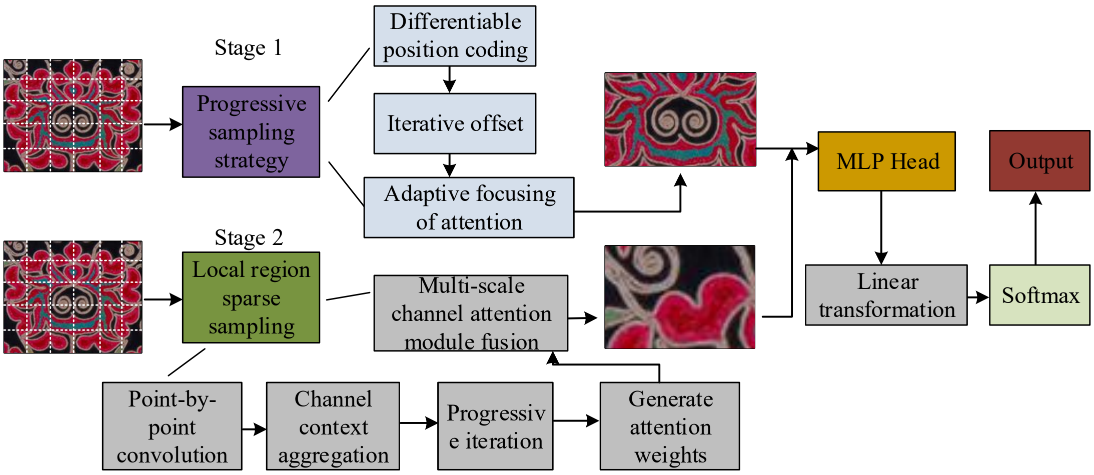


Figure 5. Schematic diagram of the structure of the progressive Transformer model

In Figure 5, the first stage adopts a progressive sampling strategy, which iteratively offsets the attention to adaptively focus on the main region of the tattoo through differentiable position coding. This avoids the background interference caused by uniform chunking in the traditional Transformer model. In the second stage, the local region is sparsely sampled, and the global and local features are fused by the multiscale channel attention module. That is, the point-by-point convolutional aggregation of the channel context is combined with the attention weights generated by progressive iteration to realize the multi-scale enhancement of edge features. The attentional feature fusion module will integrate the two-stage cues to enhance the extraction accuracy of the body and edges of the Miao embroidery pattern. The progressive

Transformer model can locate the discriminative region and predict the tattoo sampling location by iterative updating and progressive sampling. Equation 7 shows the sampling position after iterative updating.

$$Pt + 1 = Pt + Ot \quad t \in \{1\ ,,,\ N - 1\} \tag{7}$$

In Equation 7, Pt is the given sampling position. Ot is the offset matrix. t is the marker position. N is the number of repetitions. The input feature map samples the initial markers with the help of bilinear interpolation. The sampling positions are encoded and fed back into the Transformer encoder layer. The sampling positions in this model are not equally spaced. Therefore, it is necessary to project the spatial coordinates of the normalized result of the sampled positions and use the result of this projection as the position embedding [35]. The study designs bidirectional feature aggregation with the help of a bidirectional multilevel aggregation decoder, i.e., the Transformer encoder feature information is enhanced from two paths, top-down and bottom-up, as shown in Equation 8.

$$f_g = GD(Z_g^6, Z_g^{12}, Z_g^{18}, Z_g^{24}) \tag{8}$$

In Equation 8, fg is the pixel-level global feature. $(Z_g^6\ Z_g^{12}\ Z_g^{18}\ Z_g^{24})$ is the block of features 6, 12, 18, and 24. The sampled features under the bidirectional path are connected to the tensor and smooth connection is achieved with the help of convolutional stack. Equation 9 shows the implementation process of multi-scale channel attention module.

$$L(X') = BN\left(PWConv2\left(\delta\left(BN\big(PWConv1(X')\big)\right)\right)\right) \tag{9}$$

In Equation 9, the input intermediate feature values are processed by 1×1 point-by-point convolution $PWConv1$ and batch normalization BN, and then the activation function linear variation BN is performed. After that, the result is processed by 2×2 point-by-point convolution BN and batch normalization BN to output the weighted multiscale channel attention feature $L(X')$. The $L(X')$ module can capture the multiscale channel relationship by different scale convolution kernels and stabilize the training with the help of batch normalization. The fusion of attention features effectively solves the problem of semantic and scale inconsistency between features, and reduces the number of network parameters while maintaining performance. The global decision result after the first processing can be merged with the feature map to get the edge map features. After that, the study extracts the intermediate features to emphasize the edge details with the help of a bidirectional multilevel aggregation decoder. Moreover, the results are optimized with the help of loss function of Equation 10.

$$Lg = L_g^E + \gamma L_g^{side} \tag{10}$$

In Equation 10, L_g^E is the edge map loss. L_g^{side} is the lateral loss. γ is the balance coefficient. After that, the study performs the same operation on the intermediate features of the second stage with the help of a bidirectional multistage aggregation decoder to optimize the loss results. Figure 6 shows the schematic structure of the concatenated network under the fusion architecture.

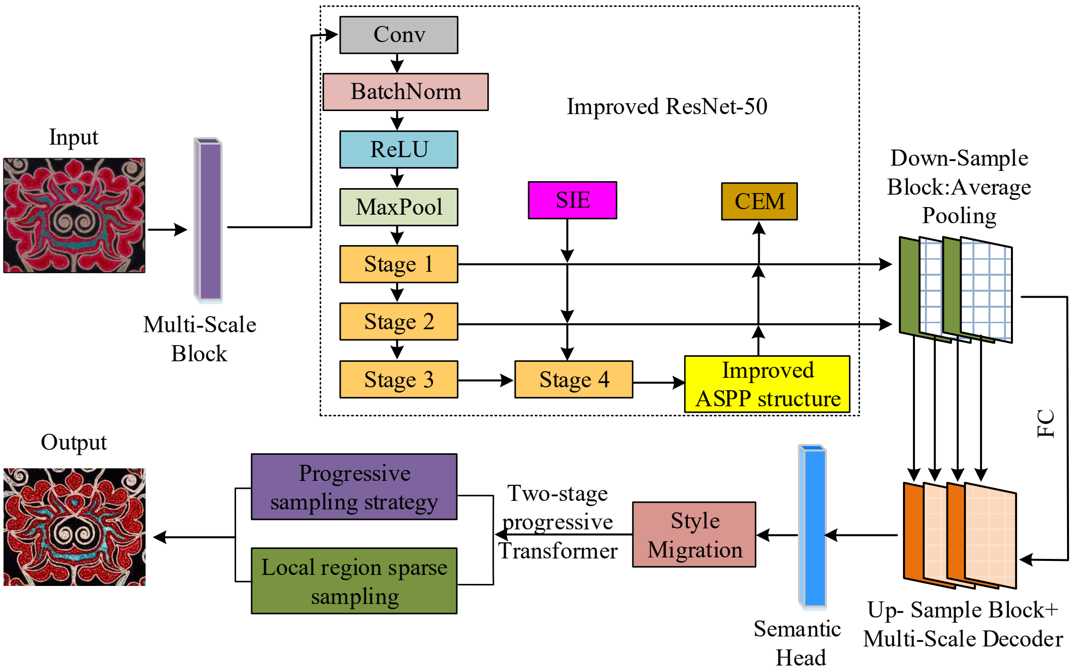


Figure 6. Schematic diagram of parallel network structure based on improved CNN-Transformer fusion architecture

In Figure 6, the 3*3 convolution operation in this network architecture can realize the fusion of contextual information. The parallel structure of convolution branch and Transformer branch can realize the synchronous connection and acquisition of image information. The channel cut structure can reduce the computational memory and computation. The key difference between embroidery and ordinary images is its unique physical texture. This texture is created by the unevenness, luster, and line texture produced by different embroidery techniques, such as flat, braided, and looped embroidery. Many generative models often lose this fine texture when imitating styles. The ResNet-50 backbone network excels at capturing local textures, stitch details, and color gradients. The Progressive Transformer uses a global attention mechanism to understand the overall structure, symbol layout, and spatial relationships of patterns, thereby ensuring their aesthetic appeal. The bidirectional, multi-level aggregation decoder and multi-scale channel attention module ensure accurate alignment and fusion of local details extracted from CNN with the global structure understood from Transformer at different scales. This method avoids the common problems of "content distortion" and "texture loss" in style transfer. It preserves the aesthetic authenticity of the original work while maintaining the style.

4. Results of Miao Embroidery Digital Intelligent Design Under the Convergence Architecture

4.1. Experimental Setup and Data Set

The study is based on a self-constructed dataset and the BSDS500 public dataset (https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html) for the Miao embroidery numerical intelligence test. Among them, the homemade dataset is collected from the Ethnic Museum of Guizhou Province (https://gzsmzmuseum.cn/appreciation-3.html) and a workshop and a tattoo design company. The dataset collects a total of 600 Miao embroidery patterns, including butterflies, fish, phoenixes, and many other patterns. Each image is manually edge labeled, and the dataset is expanded to 22,000 images by rotating, scaling, flipping, and other manipulation methods. The BSDS500 public dataset can be used to examine the image segmentation effect and contains 500 natural images. It also provides manually labeled segmentation and edge labels, which involves image segmentation and edge detection tasks. The two datasets are divided into training set and test set in the ratio of 6:4, and the effect of image numerical intelligence is examined with the help of the research model. Table 1 shows the experimental operating environment and parameter settings.

Table 1. Experimental operating environment and parameter settings

Operating environment	
Operating system-environment	Ubuntu 20.04 LTS
CPU processor	Intel Xeon E5-2678 v3 (12 cores)
GPU	NVIDIA RTX 3090 (24GB video memory)
Deep learning framework	Pytorch
Programming language	Python 3.8 + Anaconda
Parameter settings	
Positive sample threshold	0.4
Optimizer	Adam
Initial learning rate	3e-4
Image size	320×320
Batch size	16
Training batch	150
Transformer sampling iteration times	3
Attention head count	8
Convolutional scaling factor	0.5

The optimizer and the selection of the initial learning rate are based on the settings in Wojciuk et al. [36]. The selection of the Transformer sampling iteration times and attention heads parameters is based on grid search and ablation experiments on the embroidery validation set. The study evaluates the model performance with the help of metrics such as optimal dataset scale (ODS), mean intersection over union (mIoU), average precision (AP).

4.2. Miao Embroidery Image Texture Semantic Segmentation Results

The study analyzes the research-proposed converged architecture network model on the BSDS500 public dataset. It is also compared with fully convolutional network-conditional random field (FCN-CRF), object-contextual representations network (OCRNet), deep lab version 3 plus (DeepLab V3+), and residual attention-based generative adversarial network (RA-GAN). First, the loss and error situation of the above algorithm for image feature segmentation is analyzed. The results are shown in Figure 7.

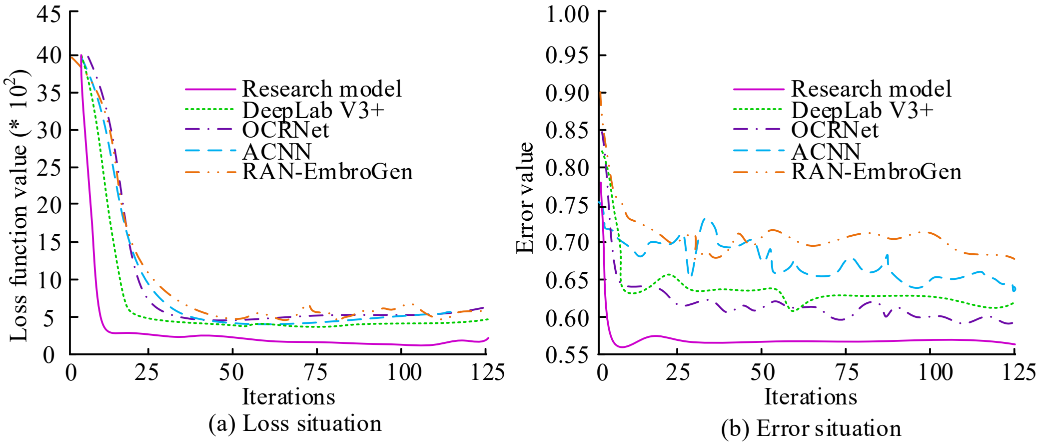


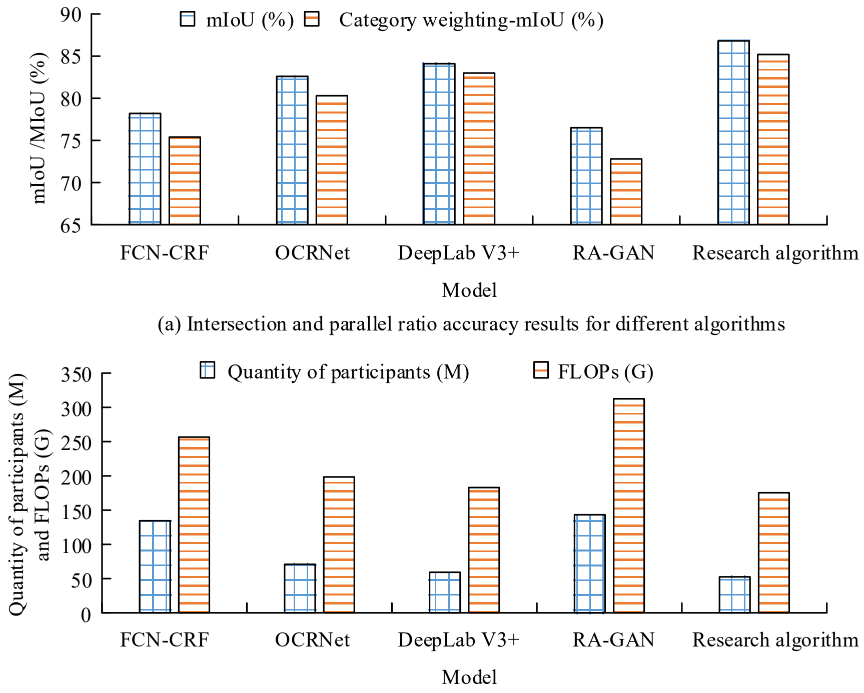
Figure 7. Image segmentation loss values and error results for different algorithms

In Figure 7(a), on the loss data, the algorithm proposed by the study exhibits smaller loss values and its loss values in the later iterations are less than 5×10^2. In contrast, the number of iterations required for the loss profiles of the other comparative algorithms to reach convergence is basically greater than 25, and the average loss value is greater than 5×10^2. In Figure 7(b), the error value of the study algorithm converges to the average value of 0.57 after the number of iterations is greater than 150. The OCRNet algorithm and DeepLab V3+ algorithm have slightly inferior error performance to the research algorithm with a value not exceeding 0.65. After that, the image segmentation accuracy of the above algorithms is compared on different Miao embroidery pattern types. The results are shown in Table 2.

Table 2. Image segmentation accuracy under different algorithm processing

Types of embroidery patterns	Model	Accuracy rate (%)	Recall (%)	F1
Geometric pattern	Research algorithm	91.99	94.12	95.15
	DeepLab V3+	89.92	91.17	90.41
	OCRNet	87.81	87.34	87.74
	FCN-CRF	82.99	86.37	85.24
	RA-GAN	79.91	84.08	81.51
Botanical pattern	Research algorithm	92.68	95.09	94.78
	DeepLab V3+	88.81	88.28	89.51
	OCRNet	85.64	85.17	86.14
	FCN-CRF	77.91	81.21	80.34
	RA-GAN	84.88	86.08	85.47
Landscape pattern	Research algorithm	90.42	88.87	90.24
	DeepLab V3+	88.44	90.77	89.81
	OCRNet	84.58	87.69	89.09
	FCN-CRF	80.59	82.82	81.95
	RA-GAN	78.82	83.27	82.66
Animal pattern	Research algorithm	87.91	88.19	88.22
	DeepLab V3+	81.98	85.11	86.51
	OCRNet	78.01	80.24	79.37
	FCN-CRF	71.89	73.17	74.36
	RA-GAN	72.89	74.18	75.36

In Table 2, the research algorithm shows more than 85% feature segmentation accuracy on all four embroidery pattern types. The segmentation accuracy is 91.99% and 92.68% for geometric pattern and botanical pattern. The next best performers are DeepLab V3+ algorithm and OCRNet algorithm. Its overall segmentation accuracy on embroidery patterns exceeds 80%, and its segmentation accuracy on geometric pattern and botanical pattern exceeds 85%. It shows that it can handle simple embroidery patterns better. The FCN-CRF algorithm and RA-GAN algorithm have slightly worse segmentation results. Their recall values on landscape pattern and animal pattern are 82.82%, 83.27%, 73.17%, and 74.18% respectively, which makes it difficult to deconstruct complex structural patterns. After that, the segmentation accuracy and computational efficiency of the above algorithms are compared. The results are shown in Figure 8.



(a) Intersection and parallel ratio accuracy results for different algorithms

(b) Quantity of participants and FLOPs results for different algorithms

Figure 8. Segmentation accuracy and computational efficiency

In Figure 8(a), the values of mIoU and mIoU under category weighting for image segmentation performed by the research algorithm are greater than those of the other comparative models, with values of 86.8% and 85.2%. In contrast, the values of the cross-combination ratios of the remaining models do not exceed 85%, and the difference in magnitude with the research model is more obvious. In Figure 8(b), the number of parameters of FCN-CRF algorithm, OCRNet algorithm, DeepLab V3+ algorithm, and RA-GAN algorithm are 134.5M, 70.8M, 59.3M, and 143.2M, respectively. The inference speed is much larger than 180G, and its computational efficiency is much larger than that of the research model. The research model's higher accuracy in processing embroidery patterns with fine structures and complex backgrounds is mainly due to the synergistic effect of the improved ASPP module and progressive Transformer. This enhanced model is able to suppress background noise and make segmentation more focused on the subject, despite the different datasets. In summary, the algorithms proposed in the study have better processing efficacy and can effectively reduce the amount of computation. Subsequently, the image segmentation quality of the above algorithm is further analyzed. The results are shown in Table 3.

Table 3. Image segmentation quality

Algorithm	FCN-CRF	RA-GAN	DeepLab V3+	OCRNet	Research Algorithm
Information entropy	6.422	6.581	6.710	7.201	7.892
Structural similarity	0.884	0.918	0.927	0.945	0.957
Mean contrast	12.525	13.212	13.844	15.355	16.639
Peak signal-to-noise ratio	28.711	30.224	31.501	33.812	34.472
Brightness relationship factor (%)	0.656	0.725	0.752	0.824	0.812
Mutual information	1.252	1.434	1.573	1.896	1.925

In Table 3, the information entropy, structural similarity, mean contrast value, peak signal-to-noise ratio (PSNR), luminance relation factor, and mutual information of the studied algorithms are 7.892, 0.957, 16.639, 34.472, 0.812, and 1.925, respectively. The overall quality of the image segmentation it exhibits is more effective. The structural similarity values of RA-GAN algorithm, DeepLab V3+ algorithm, and OCRNet algorithm are 0.918, 0.927, and 0.945, respectively. The PSNR values are also more than 30. After that, the image segmentation pixels and computation time of the above algorithms are compared. The results are shown in Figure 9.

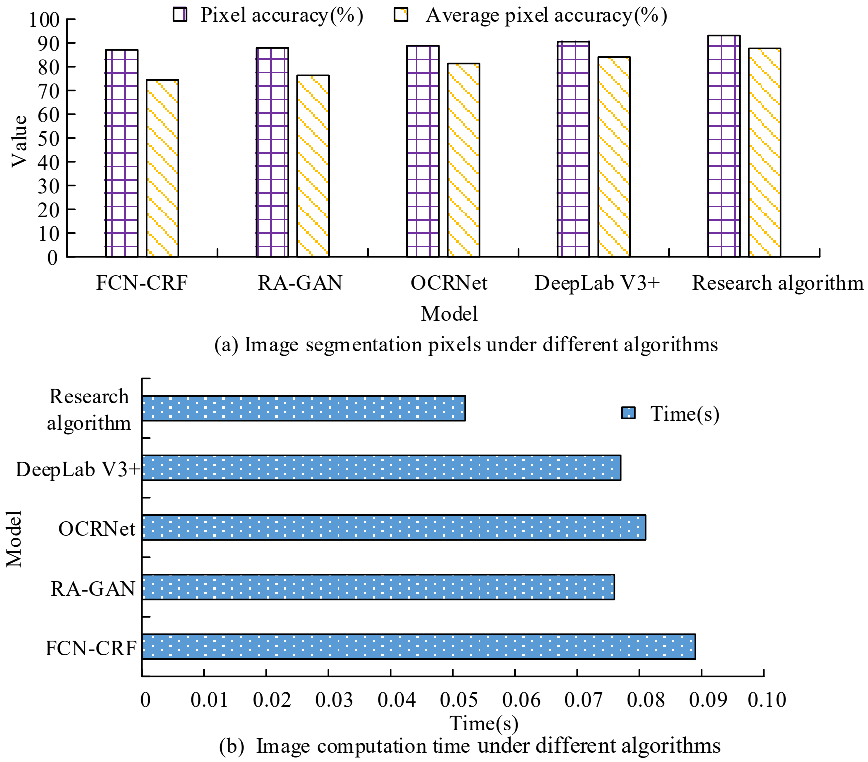


Figure 9. Image segmentation pixels and computation time

In Figure 9, the pixel accuracy and average pixel accuracy of the studied algorithms are 93.12% and 87.75%, respectively. It is much higher than 90.55% and 84.08% of the DeepLab V3+ algorithm, and the comparison difference of the more FCN-CRF algorithm and RA-GAN algorithm are more than 4% and 7%. In terms of time spent, the computation time of the research model is 0.052s, which is much less than the other comparative models.

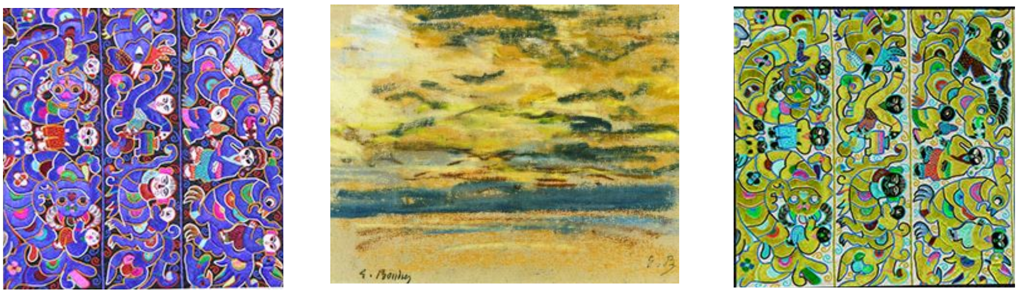
4.3. Miao Embroidery Image Style Design Result

The research proposed algorithm is compared with PatchMatch algorithm, topological semantic transfer (TST), frequency separation channel attention mechanism (FSCAM), U-Net cycle generative adversarial network (U-Net-CycleGAN). The comparison metrics include style fidelity, detail integrity, perceptual quality, and computational efficiency. "Style fidelity" is quantified using automated metrics. The KL divergence of color histograms is used to evaluate color distribution similarity, and the Gram matrix mean squared error (MSE) is used to evaluate texture and style similarity. Although subjective, "perceptual quality" is mainly measured through objective indicators such as structural similarity (SSIM) and PSNR. The results are shown in Table 4.

Table 4. Image style migration effect

Index		PatchMatch	TST	FSCAM	U-Net-CycleGAN	Research Algorithm
Style fidelity	Color-Histogram KL	1.65	0.88	0.61	0.79	0.52
	Gram-Matrix MSE	0.051	0.036	0.022	0.028	0.019
Extract content integrity	ODS	0.815	0.821	0.832	0.836	0.848
	AP	0.847	0.893	0.897	0.852	0.926
Computing efficiency	Quantity of participants (M)	95.4	112.6	89.2	143.8	52.6
	Inference time	120	185	92	105	75

In Table 4, the metric values of the research method on style fidelity are 0.52 and 0.019, which are much smaller than the other comparative algorithms. The ODS value of the research method (0.848) is greater than 0.815 for PatchMatch algorithm, 0.821 for TST algorithm, 0.832 for FSCAM algorithm, and 0.836 for U-Net-CycleGAN algorithm. The AP value is even 0.926, and the computational efficiency is faster than the other comparison models. Compared with traditional optimization based methods, the end-to-end model proposed in the study has significant advantages in efficiency. Compared with U-Net-CycleGAN, which is also based on deep learning, the research model achieved a nearly 34% improvement in style fidelity (as measured by KL divergence). This improvement is due to the fine fusion of multi-scale features by the bidirectional multi-level aggregation decoder. This decoder effectively avoids the pattern collapse and detail blurring problems that plague GAN models and better preserves the "gene" features of Miao embroidery. Figures 10 and 11 show the results of style migration and activation design of Miao embroidery for the studied algorithms.



(a) Target picture 1 (b) Target picture 2 (c) Pattern migration results

Figure 10. Migration effect of Miao embroidery style under the research algorithm

In Figure 10, Miao embroidery pattern and western art images can be better integrated, and its original style adds more romantic colors. After that, the embroidery pattern on a Miao headband is used for digitized innovative design. That is, the extracted features are processed to obtain Figure 11.

(a) Miao embroidered headband (b) Innovative product design style

Figure 11. Digitized innovations in Miao embroidery patterns

In Figure 11, the pattern on the Hmong headscarf can be applied to the design of products such as carpets after being digitally extracted. Retaining its pattern effect and nationalized design style, the application effect is better. The revised content is as follows: It further analyzes the style transfer results of the research models in study [15, 17-19], yielding Table 5.

Table 5. Model performance results based on style transfer task

Evaluation metrics		Research model	Zhao & Xin [15]	Xu et al. [17]	Kang et al. [18]	Zhong et al. [19]
Style fidelity	Color histogram KL divergence	0.48	0.71	0.65	0.58	0.82
	Gram matrix MSE	0.014	0.035	0.028	0.025	0.048
Content completeness	ODS	0.821	0.815	0.830	0.825	0.901
	AP	0.933	0.882	0.865	0.894	0.938
Generate quality (rating)	Cultural compatibility (1-10)	8.8	9.2	8.5	8.9	7.8
	Visual Aesthetics (1-10)	9.1	8.1	8.3	9.1	8.4
	Innovation (1-10)	8.8	7.8	8.6	9.0	7.0
Computational efficiency	Parameter quantity (M)	50.1	98.3	121.5	890.2	156.4
	Single graph inference time (ms)	70	102	135	210	88

As shown in Table 5, the research model performs best in restoring color distribution (KL divergence: 0.48) and texture style (Gram MSE: 0.014). This is due to the model's progressive Transformer's precise focus on the pattern subject and its ability to fuse features at multiple scales. The repair model in Zhong et al. [19] has a relatively high value in this aspect, as its main objective is to complete the structure and style consistency is not the primary task. The research model performs well with ODS and AP indicators. This indicates a high degree of alignment between the model's generated results and the contours and regions of the target content. The AP value of the model in Zhong et al. [19] is the highest (0.938), which is related to the characteristics of its repair task. Due to its explicit integration of extensible semantic analysis, the model in Zhao & Xin [15] may have an advantage in accurately expressing cultural connotations, achieving the highest simulated score for cultural fit. The model in Kang et al. [18] combines the powerful generation ability of diffusion models and user preference screening based on fuzzy evaluation, with a generation quality score exceeding 8.5. The research model is relatively balanced and comprehensive in terms of subjective ratings. It has significant advantages in terms of both parameter quantity (50.1M) and inference speed (70M).

5. Discussion

The study designed an idea for semantic segmentation and style migration of Miao embroidery based on improved CNN-Transformer fusion architecture for transformation of non-heritage digital intelligence. The results indicated that the study's proposed algorithm exhibited a small loss value. Moreover, its loss value in the late iteration was less than 5×10^2 and its error value (0.57) after stabilization was small. The error value of OCRNet algorithm and DeepLab V3+ algorithm did not exceed 0.65. The research algorithms achieved more than 90% segmentation accuracy for both geometric pattern and botanical pattern. The DeepLab V3+ algorithm and the OCRNet algorithm had an overall segmentation accuracy of more than 80% for embroidery patterns.

The research model is inspired by the Hollow Space Pyramid Pooling (ASPP) module in DeepLabV3+ to capture multi-scale contextual information. Meanwhile, the parallel fusion of the Transformer encoder and the CNN backbone network overcomes the limitations of CNN in modeling long-distance dependencies. The Progressive Transformer can dynamically focus on the subject of the pattern, which is different from how the DeepLabV3+ processes the entire feature map equally. This allows it to effectively reduce background noise interference. RA-GAN is a GAN whose core is a game between a generator and a discriminator. It is mainly used for image generation. The research model is a unified framework for semantic segmentation and style transfer, which differs fundamentally from GAN in terms of architecture and task objectives.

One of the core tasks of the research is precise semantic segmentation, which requires the model to have pixel level accurate classification ability. At present, the CNN Transformer architecture optimized for segmentation tasks has more advantages and controllability in boundary processing and structural fidelity compared to pure generative models. Although diffusion models can generate high-fidelity images, their inference process typically requires multiple iterations. This results in high computational costs and may not be efficient enough for design applications requiring rapid iterations. At the same time, the studied style transfer module achieves precise control over the degree of style transfer and detail preservation by decoupling content and style features. This decoupling is crucial for preserving the "cultural genes" of Miao embroidery's authenticity.

The FCN-CRF algorithm and RA-GAN algorithm had slightly worse segmentation results, with recall values not exceeding 85% on landscape pattern and animal pattern. The reason for these results was that the FCN-CRF algorithm's base FCN feature extraction capability was limited, and it was insufficient for complex pattern segmentation. The OCRNet algorithm had difficulty capturing small-scale details (e.g., embroidery stitches), and the DeepLab V3 algorithm fused multi-scale features. This made its null convolution computationally expensive and the edge details easy to blur. The RA-GAN model was difficult to provide better extraction of semantic features.

The research method synergized the CNN local perception and Transformer attention mechanism, which could better improve the inaccurate segmentation of tiny stitches and complex pattern edges by traditional methods. Compared to the genotype coding approach of M. Ni et al [7], the studied fusion architecture could capture both the geometric structure and cultural semantics of the tattoos, avoiding the semantic disconnection that could be caused by region substitution algorithms. Furthermore, the image segmentation mIoU and mIoU values under category weighting of the studied algorithms were greater than those of other comparative models, with values exceeding 85%. The information entropy, structural similarity, mean contrast, peak signal-to-noise ratio, luminance relation factor, and mutual information of the research algorithm were 7.892, 0.957, 16.639, 34.472, 0.812, and 1.925, respectively. The structural similarity values of RA-GAN algorithm, DeepLab V3+ algorithm and OCRNet algorithm did not exceed 0.95. They could have unclear image segmentation thresholds as well as the presence of information loss.

The pixel accuracy could reflect the percentage of the number of correctly categorized pixels. The intersection ratio related to the ratio of the area of the segmented region to the real region. Larger values of these two indicators indicated a better segmentation effect. The average pixel accuracy of the studied algorithm was 87.75%, which was much higher than that of the DeepLab V3+ algorithm, which was 84.08%. The study improved the ASPP module by multi-ratio null convolution and excitation branching. It could significantly improve the model's adaptability to the multi-scale features of Miao embroidery. This was similar to the channel attention mechanism of Ji et al [8]. However, the research method further combined with spatial pyramid pooling, which had similarity with the design idea of reinforced texture constraints proposed by Liu et al [9], while maintaining high pixel accuracy. It could better improve the texture blurring problem that exists in the image.

Color histogram KL scatter and Gram matrix mean square error could evaluate the image style fidelity effect. Among them, KL scatter could be used to assess the similarity of color distribution, the smaller value indicated the more accurate color reproduction. Gram matrix mean square error could be used to compare the similarity of texture and style between the generated image and the target image features, the smaller the value the higher the style match. The index values of the research method on style fidelity were 0.52 and 0.019, which were much smaller than other comparison algorithms. Moreover, its AP value was even 0.926, and the computational efficiency was also faster than other comparison models. The method had better results in image fusion and digital design. The model proposed in the study performed a semantic fusion of global and local tattoo features for style migration, resulting in better fidelity of the generated results. This idea echoed the color gene extraction of Na et al. [22]. However, the study further realized the decoupling control of pattern structure and style through the semantic focusing ability of Transformer. Its image migration application was significantly better than the results of interactive genetic algorithm [23].

6. Conclusion

The advent of computer technology has effectively transformed the preservation and innovation of Miao embroidery. The systematic decoding and digital reconstruction of Miao embroidery's cultural genes can be achieved through deep learning and digital image processing. A research design based on an improved CNN-Transformer fusion architecture method can improve the accuracy with which Miao embroidery patterns are decoded and the adaptability of design revitalization. However, the existing sample concentration area of the study is slightly narrow, and the feature extraction is more dependent on predefined labels. Therefore, future research should extend the cross-regional Miao embroidery data to improve generalizability. It should also explore model distillation, dynamic sparsification methods, and multimodal models to achieve semantic autocorrelation. These improvements will enhance the relevance, effectiveness, and applicability of NRM digital heritage transmission and protection.

7. Declarations

7.1. Data Availability Statement

The study is based on a self-constructed dataset and the BSDS500 public dataset for the Miao embroidery numerical intelligence test. Available online: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html (accessed on January 2026).

7.2. Funding

The author received no financial support for the research, authorship, and/or publication of this article.

7.3. Institutional Review Board Statement

Not applicable.

7.4. Informed Consent Statement

Not applicable.

7.5. Declaration of Competing Interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8. References

[1] Chen, L., Su, Z., He, X., Chen, X., & Dong, L. (2022). The application of robotics and artificial intelligence in embroidery: challenges and benefits. Assembly Automation, 42(6), 851–868. doi:10.1108/AA-07-2022-0183.

[2] Han, D., & Cong, L. (2023). Miao traditional patterns: the origins and design transformation. Visual Studies, 38(3–4), 425–432. doi:10.1080/1472586X.2021.1940261.

[3] Chen, L., Daud, W. S. A. W. M., & Arif, M. F. M. (2024). Preservation and Adaptation of Traditional Miao (Hainan) Pattern Arts: Themes, Shapes, and Colors in Contemporary Art. Asian Journal of Research in Education and Social Sciences, 6(4), 269-284.

[4] Wenji, Z., Rongrong, C., & Li, N. (2022). Design and Cultural Aspects of 20th Century Chinese Xiangjin Brocade. Fibres & Textiles in Eastern Europe, 151(3), 116–129. doi:10.2478/ftee-2022-0030.

[5] Qiao, S., & Abdramanov, Y. (2024). Exploratory sequential design for creation of the Miao's traditional patterns mobile application. Culture and History Digital Journal, 13(2), 286. doi:10.3989/chdj.2024.286.

[6] Zhu, S., & Liu, X. (2025). The Ecodesign Transformation of Smart Clothing: Towards a Systemic and Coupled Social–Ecological–Technological System Perspective. Sustainability (Switzerland), 17(5), 2102. doi:10.3390/su17052102.

[7] Ni, M., Huang, Q., Ni, N., Zhao, H., & Sun, B. (2024). Research on the Design of Zhuang Brocade Patterns Based on Automatic Pattern Generation. Applied Sciences (Switzerland), 14(13), 5375. doi:10.3390/app14135375.

[8] Ji, J., Lao, Y., & Huo, L. (2024). Convolutional neural network application for supply–demand matching in Zhuang ethnic clothing image classification. Scientific Reports, 14(1), 13348. doi:10.1038/s41598-024-64082-9.

[9] Zhao, Q. (2025). Algorithm Optimization and Innovative Design of Digital Inheritance System for Miao Family Weaving Handicraft in South Sichuan Province. Journal of Combinatorial Mathematics and Combinatorial Computing, 127b, 7171–7195. doi:10.61091/jcmcc127b-391.

[10] Liu, C., Gu, J., Yao, L., & Zhang, Y. (2025). Research on embroidery style migration model based on texture cycle GAN. International Journal of Clothing Science and Technology, 37(1), 138–153. doi:10.1108/IJCST-04-2023-0062.

[11] Fan, T., Wang, H., & Hodel, T. (2023). CICHMKG: a large-scale and comprehensive Chinese intangible cultural heritage multimodal knowledge graph. Heritage Science, 11(1), 115. doi:10.1186/s40494-023-00927-2.

[12] Du, D., Ding, J., & Liu, Y. (2025). Knowledge graph construction of Chinese embroidery evolution based on associating cultural space and critical incidents under intangible cultural heritage. Electronic Library, 43(3), 283–302. doi:10.1108/EL-02-2024-0036.

[13] Xu, L., Lu, L., Liu, M., Song, C., & Wu, L. (2024). Nanjing Yunjin intelligent question-answering system based on knowledge graphs and retrieval augmented generation technology. Heritage Science, 12(1), 118. doi:10.1186/s40494-024-01231-3.

[14] Ai, L. (2026). Decoding Color: A Three-Stage Clustering Approach to Analysis Color Evolution in Changning Miao Costumes. Color Research & Application, 51(1), e70029. doi:10.1002/col.70029.

[15] Zhao, D., & Xin, W. (2025). Research on redesign of Chinese local embroidery patterns based on extension semantics and the pix2pix model. Textile Research Journal , 95(21–22), 2691–2706. doi:10.1177/00405175251314903.

[16] Yu, Q., & Zhu, G. (2023). Digital Restoration and 3D Virtual Space Display of Hakka Cardigan Based on Optimization of Numerical Algorithm. Electronics (Switzerland), 12(20), 4190. doi:10.3390/electronics12204190.

[17] Xu, M., Tao, X., Liu, J., Pan, L., & Huang, Z. (2025). Innovation in Wickerwork Forms through the Integration of Plural Geometry and AI Algorithms, and Methods for Preserving the Cultural Symbolism of the Miao People's Flower Bamboo Hats. International Journal for Housing Science and Its Applications, 46(3), 6696–6707. doi:10.70517/ijhsa463575.

[18] Kang, X., You, W., & Xie, H. (2025). An innovative and sustainable design of intangible Miao wax printing patterns in combination of diffusion model and fuzzy TOPSIS. Humanities and Social Sciences Communications, 12(1), 1–16. doi:10.1057/s41599-025-05724-9.

[19] Zhong, C., Yu, X., Xia, H., Xie, R., & Xu, Q. (2025). Restoring intricate Miao embroidery patterns: a GAN-based U-Net with spatial-channel attention. Visual Computer, 41(10), 7521–7533. doi:10.1007/s00371-025-03821-z.

[20] Ju, F., Wang, Q., Tan, Z., & Li, Q. (2022). Intelligent Recognition of Colour and Contour from Ancient Chinese Embroidery Images. Fibres & Textiles in Eastern Europe, 151(3), 79–92. doi:10.2478/ftee-2022-0026.

[21] Zhang, Y., Zhao, H., Qi, L., Zhang, J., & Zhang, T. (2025). Research on the co-occurrence feature mining of the Qing Dynasty embroidery patterns based on temporal multilayer networks. NPJ Heritage Science, 13(1), 228. doi:10.1038/s40494-025-01766-z.

[22] Na, Z., & Sharudin, S. A. (2024). Research on Innovative Development of Miao Embroidery Intangible Cultural Heritage in Guizhou, China Based on Digital Design. GATR Journal of Business and Economics Review, 9(2), 85–94. doi:10.35609/jber.2024.9.2(1).

[23] Cong, X., & Zhang, W. (2024). Design of geometric flower pattern for clothing based on deep learning and interactive genetic algorithm. Journal of Intelligent Systems, 33(1), 57–76. doi:10.1515/jisys-2023-0269.

[24] Qian, M., Wang, Z., Huang, X., Xiang, Z., Wei, P., & Hu, X. (2022). Color segmentation of multicolor porous printed fabrics by conjugating SOM and EDSC clustering algorithms. Textile Research Journal, 92(19–20), 3488–3499. doi:10.1177/00405175221083214.

[25] Chiu, M. C., Chiang, Y. H., & Chiu, J. E. (2024). Developing an explainable hybrid deep learning model in digital transformation: an empirical study. Journal of Intelligent Manufacturing, 35(4), 1793–1810. doi:10.1007/s10845-023-02127-y.

[26] Zhao, Y., Fan, Z., Yao, H., Zhang, T., & Seng, B. (2025). Automatic Classification and Recognition of Qinghai Embroidery Images Based on the SE-ResNet152V2 Model. IET Image Processing, 19(1), 70108. doi:10.1049/ipr2.70108.

[27] Cho, H.-Y. (2021). The Language of Miao Embroidery: Exploring the Traditional "Embroidered Rear Skirt Panels" Worn by the

Miao Women of the Huawu Village. TEXTILE, 21(1), 2–31. doi:10.1080/14759756.2021.1959821.

[28] Alaeddine, H., & Jihene, M. (2021). Deep residual network in network. Computational Intelligence and Neuroscience, 2021(1), 6659083. doi:10.1155/2021/6659083.

[29] Hu, Q., Peng, Y., Xu, J., Shao, Z., Tian, Z., & Chen, J. (2025). Adaptive Stylized Image Generation for Traditional Miao Batik Using Style-Conditioned LCM-LoRA Enhanced Diffusion Models. Mathematics, 13(12), 1947. doi:10.3390/math13121947.

[30] Hou, X., Zhao, H., & Wang, C. (2024). Hierarchical segmentation for traditional cultural pattern based on iterative compression and clustering. Multimedia Systems, 30(6), 372. doi:10.1007/s00530-024-01578-4.

[31] Zhou, Y., Ren, Y., Wu, C., & Xue, M. (2024). Style Transfer of Chinese Wuhu Iron Paintings Using Hierarchical Visual Transformer. Sensors, 24(24), 8103. doi:10.3390/s24248103.

[32] Zhuo, X., Huang, D., Lin, Y., & Huang, Z. (2024). Combined query embroidery image retrieval based on enhanced CNN and blend transformer. Scientific Reports, 14(1), 27518. doi:10.1038/s41598-024-79012-y.

[33] He, W., Song, B., Zhang, N., Xiang, J., & Pan, R. (2024). Modeling and realization of image-based garment texture transfer. Visual Computer, 40(9), 6063–6079. doi:10.1007/s00371-023-03153-w.

[34] Sheng, H., Cai, S., Zheng, X., & Lau, M. (2025). Knitting Robots: A Deep Learning Approach for Reverse-Engineering Fabric Patterns. Electronics (Switzerland), 14(8), 1605. doi:10.3390/electronics14081605.

[35] Fang, B., Jiang, M., Shen, J., & Stenger, B. (2022). Deep Generative Inpainting with Comparative Sample Augmentation. Journal of Computational and Cognitive Engineering, 1(4), 174–180. doi:10.47852/bonviewJCCE2202319.

[36] Wojciuk, M., Swiderska-Chadaj, Z., Siwek, K., & Gertych, A. (2024). Improving classification accuracy of fine-tuned CNN models: Impact of hyperparameter optimization. Heliyon, 10(5), 26586. doi:10.1016/j.heliyon.2024.e26586.