J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025), 6025--6043

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Stroke Structure Reconstruction and Virtual Display System Design for Calligraphic Seal Engraving Using Artificial Intelligence

Shiyi Xu[1, ✉]

[1] The College of Educational Science and Technology, Anshan Normal University, Anshan, Liaoning, 114000, China

ABSTRACT
Aiming at the needs of reconstructing the structure of calligraphic seal cutting strokes and virtual display, this study designs a GAN technique that integrates three models, namely, “WGAN, DCGAN and CGAN”. The CycleGAN model is used to obtain the mapping relationship between learning and style migration by utilizing its cyclic consistency loss. Adaptive pre-morphing technique is introduced to process the input image to capture the outline information and morphological features of calligraphic seal carvings, and a Generative Adversarial Network-based Generative Model for Structural Reconstruction of Calligraphic Fonts (CRA-GAN) is proposed. Meanwhile, an online virtual display system is designed to provide users with a good sense of experience in the virtual display of calligraphy. The results show that the CRA-GAN model can better capture the details and global information of the fonts, and its recognition rate of the eight calligraphic fonts ranges from 90.42% to 97.38%, and the MOS rating value of the text image is > 8.5 points, and its recognition results are in line with the observation characteristics of the human eye for calligraphic images. The FID calculation result of the CRA-GAN method (204.361) of the CRA-GAN method is much lower than that of other methods, which obviously improves the diversity and visual quality of the generated calligraphic fonts. This paper evaluates the user's experience of the system from five aspects: narrative experience, emotional experience, sensory experience, cognitive experience and interactive experience, and calculates that the final score of the system is in the range of 80-100, which indicates that the user's satisfaction is very high after actually experiencing the virtual display system.

Keywords: GAN technology, adaptive pre-deformation technology, CRA-GAN, online virtual display system, calligraphy seal cutting

1. Introduction

Calligraphy, as a unique art form with Chinese characteristics, unites the wisdom and aesthetic

✉ Corresponding author.
E-mail address: 15642246875@163.com (S. Xu).
Received 20 March 2024; Revised 05 June 2024; Accepted 20 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-332

interests of the Chinese nation. Calligraphy strokes must be accurate and powerful, ink in the font between the strokes should maintain a certain thickness of the hanging wave, to avoid sloppy and sloppy, curved and powerful, is an indispensable realization of the calligrapher to show their skills [1-2]. In calligraphy, each character consists of a series of strokes. Strokes are the basic elements of a written character, and their shape and order together determine the overall appearance of a character. The order of the strokes can affect the trend of the character shape. Generally speaking, the strokes of a written character should be written in the order of horizontal, vertical, apostrophe, and downstroke, because the horizontal and vertical strokes constitute the basic skeleton of the character, and the apostrophe and downstroke strokes are supplemented on the basis of this [3-4]. In addition, the continuity, symmetry, sparseness and length of the strokes can influence the final presentation of the calligraphy. The first two are to praise the overall coordination and balance of calligraphy, while the sparse and dense distribution of strokes can increase the sense of hierarchy and change of the font, and different lengths of strokes can make the structure of the font become rich and full [5-6].

Through different combinations of strokes, different character structures are formed, which determine the shape and beauty of a character. The overall structure of a character includes the distribution and size relationship between the top, bottom, left and right directions, and a good overall structure can make the character look harmonious and balanced [7]. Among them, radicals are radicals with independent meanings in a character form, which determines a part of the character form. In calligraphy, the rational use of radical structure can enhance the recognizability and beauty of a character. The structural law of calligraphic fonts is to create an overall aesthetic sense through the organic combination and adjustment of different strokes, and to make the fonts characterized by stability, harmony and order [8].

Seal carving, as a form of expression combining calligraphy and engraving, is a unique artistic expression in Chinese traditional culture [9]. It is not only the skill of seal making, but also a kind of artistic creation rich in philosophy and aesthetic value. Calligraphy and seal carving, the former using the pen and the latter using the knife, are both the art of lines, forming a similarity to the structure of the font strokes. On the one hand, successful seal carving works can not be separated from excellent calligraphy works as a virtue to rely on; on the other hand, seal carving is another form of expression of calligraphy, which promotes the reform and development of the art of calligraphy, and the two are complementary to each other, and they are developing together [10-12].

In this paper, the basic principle and training process of GAN are introduced, and then it is improved with the help of WGAN, DCGAN and CGAN models sequentially in terms of dispersion, network structure and conditional information, and the Pix2Pix model and CycleGAN model, which are commonly used in the style transformation, are introduced. The input image is processed by introducing the adaptive pre-morphing technique to make it conform to the morphological features of calligraphic seal cutting. The Robert operator is then utilized to extract the contours of the Chinese character images. A region-aware attention generator is utilized to further improve the generative performance of the model, and the relationship between the calligraphic image and the mapping is ensured with the help of contour consistency loss function. Finally, a generative adversarial network-based calligraphic font generation model (CRA-GAN) is proposed to realize the generation of Chinese character fonts with specific calligraphic styles, and then the effect of users' application and experience of the model in this paper is explored through an online virtual display system.

2. Overview

In terms of calligraphic strokes, literature [13] extracts the stroke structure points of “character” through the rules of calligraphic writing, and connects them, and after special short stroke return, curved connection forms the whole outline, and collects the characteristics of individual stroke outlines, forming the outline method of calligraphic stroke extraction and expression. Literature [14] splits the calligraphic strokes, extracts the individual stroke structure features, and at the same time extracts the features of the whole word, and carries out quantitative evaluation in the designed in-depth stroke extraction method for intelligent evaluation of calligraphic Chinese characters, which highlights the fine granularity of the strokes in the "word". Literature [15] constructed a projector and motion sensor calligraphy stroke learning aid system, which can comprehensively display the strokes, trajectories, and handwriting of calligraphy strokes, and there are videos of professional calligraphers' writing, and guide their strokes through the plane visual data. Literature [16] considered the brush indicators such as the angle of the calligraphy brush barrel, the posture and position of the pen in writing, and proposed a dynamic output writing rendering algorithm, which automatically generates the stroke trajectory of calligraphy by predicting the stroke trajectory. All these studies are on the extraction of calligraphic stroke detail features, aiming at obtaining the font characteristic style of calligraphic works for better appreciation or learning. In addition, a number of scholars have also studied the automatic generation of calligraphic strokes based on stroke features. Literature [17] utilized the hierarchical inference pyramid method to collect the sequence of calligraphic strokes and designed a robotic calligraphy stroke inference method in the context of the complete feature set. Similarly, literature [18] developed a robotic calligraphy writing system by extracting the style and stroke features of various types of Chinese calligraphy works, learning from stroke training in a generative adversarial learning model and generating the corresponding strokes of the calligraphy style. Similarly, literature [19] used generative adversarial networks to construct the unsupervised calligraphic font generation network model SPFont, but the model not only considered the writing style and stroke structure features, but also included the content, stroke alignment, and exhibited high-quality calligraphic fonts. Literature [20] conceived a composite curve expansion stroke model CCD-BSM under the consideration of the brush and writing action of the brush, a major advantage of this model is that it does not need to be trained with a large number of samples, and experiments show that the model has a high degree of similarity in simulating the graphics of strokes. And the literature [21] constructed a stroke segmentation generative adversarial network algorithm, which can generate all strokes in the calligraphy image simultaneously and be used as an automatic evaluation of calligraphy works.

In the area of seal carving, literature [22] designs an intelligent seal carving art generation system guided by visual knowledge, database of seal images and seal images, and professional knowledge of seal carving, and optimizes the structure of seal fonts via a deformation algorithm. There are very few studies in the direction of seal carving, and we speculate that it may be caused by the difficulty in mastering the structure of seal carving strokes. Therefore, it is of cutting-edge significance to utilize artificial intelligence technology to conduct research on seal carving in calligraphy.

3. Stroke Structure Reconstruction for Improved GAN Modeling of Calligraphic Seal Engraving

3.1. Relevant Technology and Theoretical Basis

3.1.1. Basic Theory and Implementation of GANs:

1) The basic principle of GAN. GAN [23] is mainly through the two modules in the model - generator and discriminator to play with each other to improve the learning ability of the model, so as to produce better generation results. The goal of the generator is to generate results that are realistic enough that the discriminator is unable to distinguish between true and false; the goal of the discriminator is to distinguish the gap between the generated sample $G(Z)$ and the true sample X, and to judge the generated sample $G(Z)$ as false. The generator and the discriminator play games with each other, and continuously optimize in the confrontation, so that the model generates satisfactory results.

2) Training process of GAN. The training process of GAN is a continuous game between the generator and the discriminator, the generator generates realistic enough sample images for the discriminator to determine them as true, while the discriminator has to improve its ability to recognize true and false

samples, and determine the images generated by the generator as false.

The optimization function of GAN is as follows:

$$\min_{G}\max_{D} V(D,G) = E_{x\sim P_{das}(x)}[\log D(x)] + E_{z\sim P_z(z)}[\log(1-D(G(z)))] \tag{1}$$

Generator G aims to minimize this function, while discriminator D wants to maximize it. The game between the generator and the discriminator is a game in which the optimization function is extremely small.

GAN belongs to unsupervised learning, no need to label the data in advance, and the learned features are more adaptive and rich compared to supervised learning. However, the training of GAN is more difficult, and there are problems such as non-convergence and pattern collapse.

3.1.2. Improvements in GAN. In response to the shortcomings of GAN, researchers have improved and explored it in different aspects. In this section, CGAN, DCGAN and WGAN will be briefly introduced.
1) WGAN. Uses Wasserstein distance instead of JS distance.Wasserstein distance can still reflect the distance between two distributions when they do not overlap, providing meaningful gradient information and effectively avoiding the situation of gradient disappearance. In the training process, Wasserstein distance can indicate the training process, the smaller its value, the better the GAN is trained, and the higher the quality of the image produced by the generator.The proposal of WGAN solves the problem of unstable GAN training, so that the model user no longer needs to carefully balance the degree of training of the generator and the discriminator.
2) DCGAN. DCGAN applies convolutional neural networks to GAN for the first time, utilizing the strong representativeness of convolutional neural networks and the acuity of capturing image features to accelerate the training of the model and improve the stability of training.
3) CGAN. CGAN adds the conditional variable y on the basis of GAN, which introduces additional information to the model, guides the generation of data, and makes the generated samples have specific properties. Conditional variable y can be in various forms, such as category information or image data.

3.1.3. Common Models for Style Transitions. The fusion of improved models of GAN in different aspects has led to the emergence of a series of complex GAN-based models that are capable of solving problems in different domains. The tasks of calligraphic character image generation as well as texture generation are often viewed as style migration of images, and in this section, Pix2Pix and CycleGAN, the classic models in the field of image style migration, will be introduced in detail.
1) Pix2Pix. Pix2Pix [24] is a variant of CGAN, which can realize the mapping from image to image, and it is a stronger CGAN.Pix2Pix provides noise in the form of dropout, which is applied to the layer of the generator during training and testing, and at the same time, it takes the real image x as a conditional variable, which guides the generator G, so that the generated G(x) is affected by and has a correspondent relationship with x. The generator G(x) has a corresponding relationship with x, and the generator G(x) has a corresponding relationship with x. relationship.

Pix2Pix retains the adversarial loss in CGAN and adds pixel loss to it to calculate the gap between the generated image $G(x,z)$ and the real image y. The optimization function of Pix2Pix is shown in Eq. (2), where L_{cGAN} is the adversarial loss, which embodies the adversarial process between the generator G and the discriminator D; L_{L1} is the pixel loss, which restricts the generated image $G(x,z)$ to be consistent with the real image y and λ is the weight of the pixel loss. Then:

$$G^{*} = \arg\min_{G}\max_{D} \mathrm{L}_{cGAN}(G,D) + \lambda \mathrm{L}_{L1}(G) \tag{2}$$

$$\mathrm{L}_{cGAN}(G,D) = E_{x,y}[\log D(x,y)] + E_{x,z}[\log(1-D(x,G(x,z)))] \tag{3}$$

$$\mathrm{L}_{L1}(G) = E_{x,y,z}[\| y - G(x,z) \|] \tag{4}$$

The generator of Pix2Pix has chosen the U-Net network architecture, which is able to share both the high-level semantic information and the underlying semantic information, retaining the pixel-level detail information at different resolutions and maintaining the high resolution of the image.

2) CycleGAN. To address the drawback that Pix2Pix must be trained using paired data, CycleGAN [25] is the most classical unsupervised image stylization model, and is often referred to as the unpaired version of Pix2Pix because the training data do not need to be paired. CycleGAN makes it possible for the model to perform style-specific transformations without using paired data by using a closed-loop dyadic network structure, as well as cyclic consistency loss.

In order to be able to compute the loss efficiently and constrain the generator to produce satisfactory images, CycleGAN adds a generator F. Generator G converts image x into image $\hat{Y}$ and then feeds it into generator F to produce image $\hat{x}$, and computes the loss between image x and image $\hat{x}$ as a way to constrain the generator. The loss between picture x and picture $\hat{x}$ is called cyclic consistency loss.

The binding of $\tilde{Y}$ generated by the intermediate process is small. In order to achieve satisfactory generation of $\tilde{Y}$ as well, CycleGAN adds another loop, the generator process of $y \to \hat{X} \to \hat{y}$. The two loops ensure that both generator G and generator F are boosted.CycleGAN has two discriminators, D_Y and D_X, and the model inputs $\hat{Y}$ and y into discriminator D_Y to determine true or false, and $\hat{X}$ and x into discriminator D_X to determine true or false.

3.2. CRA-GAN based model for generating calligraphic characters

In this paper, we propose a Chinese Calligraphy Character Generation Method (CRA-GAN) based on contour and region-aware attention, which uses CycleGAN (Cycle Generative Adversarial Network) as the backbone network, and adds contour information, region-aware attention, and adaptive pre-morphing to improve the quality of the calligraphic character generation.

3.2.1. Adaptive Pre-deformation. The conversion process between standard and calligraphic fonts can be seen as a complex deformation process. The specific process of Adaptive Pre-deformation (AP) module is as follows:

1) Given a Chinese character x_i in the source font dataset X of size N_x, firstly, the maximum outer matrix of the image of size 128*128, i.e., the Chinese character region, is calculated by the edge detection algorithm, and then the height (h_{x_i}) and width (w_{xi}) of this Chinese character region are recorded, and the average literal face occupancy ratio (avg_crp_X) and the average height-to-width ratio (avg_hwr_x) of the source Chinese character image set X based on the source Chinese character image set 6 can be computed by the following equations:

$$avg\_crp_X = \frac{1}{N_X}\sum_{i}^{N_X}\frac{h_{x_i} * w_{x_i}}{128*128} \tag{5}$$

$$avg_h wr_x = \frac{1}{N_X}\sum_{i}^{N_X}\frac{h_{x_i}}{w_{x_i}} \tag{6}$$

Similarly, we can derive the corresponding average literal face occupancy (avg_crp_γ) and average aspect ratio (avg_hwr_γ) for the target font dataset Y of size N_Y, as shown below:

$$avg\_crp_Y = \frac{1}{N_Y}\sum_{j}^{N_Y}\frac{h_{y_j} * w_{y_j}}{128^*128} \tag{7}$$

$$avg\_wr_Y = \frac{1}{N_Y}\sum_{j}^{N_Y}\frac{h_{y_j}}{w_{y_j}}, \tag{8}$$

where h_{y_i} and w_{y_j} denote the height and width of the Chinese character region of y_j in the target font dataset Y, respectively.
2) Using the above four statistical values, adaptive pre-morphing is realized according to the following three cases: a) if $avg\_crp_x > avg\_crp_Y$, firstly, the resize() function is used to shrink the aspect ratio of each x_i to make it consistent with avg_hwr_Y, and then the white background is filled to make the image size 128*128; b) if $avg\_crp_X < avg\_crp_Y$, each image x_i is enlarged according to avg_hwr_Y, and then the image is cropped to make the image size 128*128; c) if $avg\_crp_x = avg\_crp_Y$, i.e., keep the image itself without adaptive pre-distortion.
3) According to the above four statistical formulas, a predeformation operation can be performed on each of the input source Chinese character images so that the source image is roughly aligned with the Chinese character region of the corresponding target calligraphy image.

Through the adaptive predeformation operation, the Chinese characters in the source font set are roughly aligned with the literal regions of the Chinese characters in the target font set. It can be seen that the literal regions of the source characters after the pre-morphing are similar to the literal regions of the corresponding target characters.

3.2.2. Contour extraction. Chinese characters are composed of strokes, and a Chinese character outline is the set of points formed by the edges of a Chinese character's strokes. We note that the background noise of the Chinese calligraphy character images used for training is very small, and the outline of the calligraphy can be easily extracted with some simple edge detection operators at a very low computational cost. In this paper, we utilize the simple but efficient Robert's operator [26] to extract the contours of Chinese character images. Specifically, the Roberts operator uses the difference between two neighboring pixels in the diagonal direction to approximate the gradient magnitude to detect the edges, and it employs the 2×2 operator. The difference calculation formula is as follows:

$$\Delta f_x(x,y) = f(x,y) - f(x+1,y+1) \tag{9}$$

$$\Delta f_y(x,y) = f(x+1,y) - f(x,y+1) \tag{10}$$

Its gradient magnitude is calculated as:

$$F(x,y) = [(\Delta f_x(x,y))^2 + (\Delta f_y(x,y))^2]^{\frac{1}{2}} \tag{11}$$

Therefore, the threshold T is chosen appropriately and if $F(x,y) > T$, pixel (x,y) is considered as an edge point.

3.2.3. Area-aware attention generator. Since the content region of a calligraphic character provides important information about the content of the calligraphic character, we incorporate region-aware attention into the generator to capture the content region to further improve the generation performance.

The region-aware attention generator follows the encoder-decoden architecture is composed of a parameter sharing encoder, converter, content mask decoder, attention mask decoder and fusion module. The parameter sharing encoder encodes a four-channel input image to produce a feature representation. A converter consisting of nine residual blocks stacked together converts the feature representation of the source font domain to the feature representation of the target font domain. The content mask decoder tends to capture the foreground content of the Chinese characters, while the attention mask decoder tends to capture the foreground content region as well as the background region. The content mask decoder tends to capture the foreground content of Chinese characters, while the attention mask decoder tends to capture the foreground content area as well as the background area. The fusion module in the region-aware attention generator is shown in Figure 1 (wherein "任" is the content of the Chinese character in the input image), and is designed to fuse the foreground content mask $\left(\left\{C_y^f\right\}_{f=1}^{n-1}\right)$, the foreground content area attention mask $\left(\left\{A_y^f\right\}_{f=1}^{n-1}\right)$ and the background area attention mask (A_y^b) and the input character x in the source font domain to obtain the calligraphic character area and the background region of the calligraphic character image, and then fuse to generate the target calligraphic character $G_\gamma(x)$, wherein n is the number of masks used in the model. Default $n = 10$, the mathematical formula is expressed as follows:

$$G_Y(x) = \sum_{f=1}^{n-1}(C_y^f \otimes A_y^f) + x \otimes A_y^b \tag{12}$$

where $\otimes$ denotes the pixel-level cumulative multiplication.

Similarly, inputting an image y, the output image $G_x(y)$ of generator G_x can be represented as:

$$G_X(y) = \sum_{f=1}^{n-1}(C_x^f \otimes A_x^f) + y \otimes A_x^b \tag{13}$$

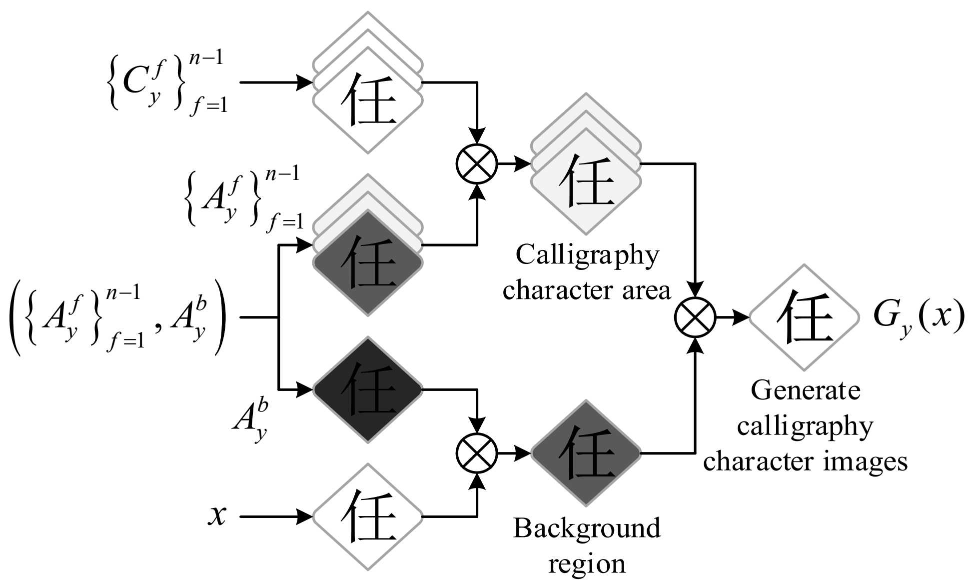

Fig. 1. The region senses the fusion module in the attention generator

3.2.4. Loss function design. The model training process combines adversarial loss, cyclic consistency loss and contour consistency loss. For mapping $G_Y: X \to Y$, G_Y generates calligraphic images $G_\gamma(\tilde{x})$ that look as similar as possible to the images in domain Y, while D_γ distinguishes between the real y and generated $G_\gamma(\tilde{x})$ as much as possible, and the other mapping $G_X: Y \to X$ is the same, so the adversarial loss can be expressed as:

$$\begin{aligned} &L_{adv}(G_X, G_Y, D_X, D_Y) = \\ &E_{y\sim P(y)}[\log D_Y(y)] + E_{\tilde{x}\sim P(\tilde{x})}[\log(1 - D_Y(G_Y(\tilde{x})))] \\ &+E_{\tilde{x}\sim P(\tilde{x})}[\log D_X(\tilde{x})] + E_{\hat{x}\sim P(\hat{x})}[\log(1 - D_X(\hat{x}))] \end{aligned} \tag{14}$$

where G_γ, G_χ aims to minimize the adversarial loss and D_γ, D_χ maximizes it as much as possible.

Considering that the content of the reconstructed image and the real original image should be consistent, i.e., $G_x(G_Y(\tilde{x})) \approx \tilde{x}$, $G_Y(G_X(y)) \approx y$, the L_1 loss is used to constrain the content:

$$\begin{aligned}L_{cyc}(G_X,G_Y) &= E_{\tilde{x}\sim P(\tilde{x})}[\| G_X(G_Y(\tilde{x})) - \tilde{x} \|_1] \\ &+ E_{y\sim P(y)}[\| G_Y(G_X(y)) - y \|_1]\end{aligned} \tag{15}$$

where $G_X(G_Y(\tilde{x})$ and $G_Y(G_X(y))$ are reconstructed images and $\|\cdot\|_1$ denotes the L_1-parameter number.

After extracting the calligraphic contours, a large amount of redundant information is removed from the Chinese character images, on the other hand, the indispensable structural information of the strokes is also highlighted. Therefore, we propose contour consistency loss $L_{contour}(G_x,G_y)$ to ensure the contour consistency between $G_x(G_Y(\tilde{x}))$ and $\tilde{x}$, $G_Y(G_X(y))$ and y. $L_{contour}(G_X,G_Y)$ can be expressed as:

$$\begin{aligned}L_{contour}(G_X,G_Y) &= E_{\tilde{x}\sim P(\tilde{x})}[\| CE(G_X(G_Y(\tilde{x}))) - CE(\tilde{x}) \|_1] \\ &+ E_{y\sim P(y)}[\| CE(G_Y(G_X(y))) - CE(y) \|_1]\end{aligned} \tag{16}$$

where CE represents the contour extraction operation, which is used to extract the contours of the reconstructed image and the original Chinese character image. Finally, our total loss function is defined as:

$$\begin{aligned}L_{CRA-GAN}(G_X,G_Y,D_X,D_Y) &= L_{ad_v}(G_X,G_Y,D_X,D_Y) \\ &+ \lambda_1 L_{cyc}(G_X,G_Y) + \lambda_2 L_{contour}(G_X,G_Y)\end{aligned} \tag{17}$$

where λ_1,λ_2 is the weight of each loss.

3.3. Training results and performance of the CRA-GAN model

3.3.1. Experimental setup：

1) Data sets. In this paper, a total of eight calligraphy datasets containing very different handwriting thicknesses and writing styles were collected, of which the size of the dataset used for training and testing is shown in Table 1.

Table 1. Data set size for training and testing

Font name	Training set size	Test set size	Font name	Training set size	Test set size
Wxz	5421	1361	Yzq	5499	1371
Lgq	5487	1379	Sgt	4938	1205
Smh	1029	907	Mf	5474	1347
Htj	4938	220	Oyx	2893	719

2) Extraction of skeleton image and contour image. We enhance the calligraphic fonts using expansion and erosion operations, and then obtain the corresponding skeleton image and contour image by skeleton algorithm and contour algorithm.

3) Non-exact pairing for Chinese character recognition

In order to better capture the global structural information of calligraphic Chinese characters, this paper proposes to use the Chinese character recognition technique to generate a partially imprecise pairing dataset. Although there is room for improvement in the recognition accuracy of the adopted Chinese character recognition model, good recognition results can be achieved for most of the calligraphic fonts. Even if they are incorrectly recognized, the incorrect fonts are very similar in terms of the shape of the characters, thus providing more information about the characters to the model network.

4) Baseline Model. In order to verify the effectiveness of the CRA-GAN model proposed in this paper, it is compared with the following baseline model:

CycleGAN: GAN model based on unpaired dataset.

UGATIT: GAN model based on adaptive layer instance normalization and unpaired data.

We use Robert as the optimizer with parameters 0.5 and 0.999 for the generator and discriminator optimization subproblems.We set the four hyperparameters for the weights of cyclic consistency loss, cyclic skeleton consistency loss, cyclic contour consistency loss, and soft label loss to 1, 0.0001, 0.001, and 0.001, respectively.

5) Evaluation index. We use three evaluation indexes, namely, structural similarity (SSIM), peak signal-to-noise ratio (PSNR), and mean square error (MSE), to evaluate the experimental effect of the model.

3.3.2. Comparison of CRA-GAN model with existing models. The performance comparisons of different models in eight calligraphy font generation tasks are shown in Table 2.CRA-GAN achieves excellent performance in all metrics. Specifically, the mean square error (MSE) of this paper's model, CRA-GAN, is the smallest among the eight calligraphic fonts, ranging from 0.002-0.097, all of which are less than 0.1. The MSEs of the CycleGAN model and the UGATIT model range from 0.092-0.154 and 0.103 respectively -0.181. In terms of Peak Signal to Noise Ratio (PSNR), the larger value represents a better recognition of the model. The results of the survey show that the peak signal-to-noise ratio (PSNR) of the CRA-GAN model is increased by 2.457-7.498 and 1.652-5.223 over the CycleGAN and UGATIT models, respectively; and the recognition effect on structural similarity (SSIM) is increased in the order of 0.007 -0.346 and 0.083-0.257. The comparison results of the recognition performance of the three metrics in eight kinds of calligraphy show that the CRA-GAN model proposed in this paper has the smallest error and the best effect in recognizing the structure of calligraphic seal-carving murals. This is directly related to the fact that the CRA-GAN model can well alleviate the problem of pattern collapse and achieve better generation results.

Table 2. Different models are compared in eight types of calligraphy fonts

Index	Mission	Wxz	Lgq	Smh	Htj	Yzq	Sgt	Mf	Oyx
MSE	CycleGAN	0.111	0.106	0.113	0.095	0.092	0.110	0.103	0.154
	UGATIT	0.181	0.107	0.152	0.103	0.112	0.115	0.126	0.169
	CRA-GAN	0.086	0.044	0.091	0.036	0.002	0.081	0.005	0.097
PSNR	CycleGAN	6.484	6.964	6.503	9.228	7.592	7.737	9.767	9.942
	UGATIT	10.617	10.579	8.262	9.244	6.729	6.594	10.572	9.287
	CRA-GAN	13.963	11.092	10.089	11.717	11.952	10.915	12.224	12.459
SSIM	CycleGAN	0.361	0.191	0.556	0.195	0.38	0.361	0.554	0.57
	UGATIT	0.174	0.285	0.679	0.332	0.342	0.289	0.45	0.383
	CRA-GAN	0.431	0.519	0.762	0.541	0.599	0.488	0.599	0.607

3.3.3. Ablation experiments. In this section, we conduct a series of ablation experiments to validate the effectiveness of the skeleton, outline, and soft labels obtained from the Chinese character recognition techniques introduced in this paper. For this purpose, we consider four calligraphic font generation tasks “wxz, Lgq, Smh and Htj”. Taking CycleGAN as the baseline model, we consider the performance of adding skeleton information, outline information and soft label information into the model one by one, where the CRA-GAN model proposed in this paper corresponds to the model after adding all the skeleton, outline and soft label information into CycleGAN. The results of the ablation experiments are shown in Table 3. It can be seen that the structural similarity (SSIM) and peak signal-to-noise ratio (PSNR) values of the model are gradually improved after the skeleton, silhouette and soft label information are added to the model sequentially, among which the improvement is more significant when the skeleton and silhouette information are added; and the MSE is gradually reduced. These experimental results also illustrate the effectiveness of skeleton and outline information in generating calligraphic Chinese fonts.

Table 3. Ablation experiment results

Index	Mission	Wxz	Lgq	Smh	Htj
MSE	CycleGAN	0.237	0.151	0.126	0.191
	CycleGAN+skeleton	0.203	0.128	0.113	0.157
	CycleGAN+skeleton+ outline	0.154	0.124	0.107	0.125
	CycleGAN+ skeleton+outline+ Soft label	0.106	0.101	0.082	0.102
PSNR	CycleGAN	6.001	6.23	6.633	8.091
	CycleGAN+skeleton	6.807	6.772	6.923	8.204
	CycleGAN+skeleton+ outline	9.807	8.242	8.923	11.204
	CycleGAN+ skeleton+outline+ Soft label	10.499	9.995	9.773	12.466
SSIM	CycleGAN	0.145	0.503	0.437	0.377
	CycleGAN+skeleton	0.194	0.507	0.432	0.389
	CycleGAN+skeleton+ outline	0.249	0.512	0.526	0.413
	CycleGAN+ skeleton+outline+ Soft label	0.356	0.523	0.547	0.449

3.3.4. Evaluation of Chinese character image quality. Currently, there is a lack of common evaluation methods for both the calligraphic images and the results of the generated adversarial model, in order to quantitatively evaluate the experimental effects, the DR, MOS, and the initial distance (FID) are used as the metrics for evaluating the results of the model proposed in the paper.The DR value indicates the probability that the generated calligraphic fonts can be correctly recognized, which is used to measure the accuracy of the generated fonts, and the MOS value indicates the subjective evaluation score of the generated results of the three models, which is used to measure the visual quality of the generated images. The MOS value represents the subjective evaluation score of the human eye of the generated results, which is used to measure the visual quality of the generated images. Ten images of each of the eight styles of calligraphic fonts, namely, 1 WeiBei, 2 KaiShu, 3 XingShu, 4 XingKai, 5 CursiveShu, 6 ShuShu, 7 ClericalShu, and 8 LiuShu, were randomly disrupted and subjectively evaluated by the human eye in a normal light environment, and each observer was required to recognize the fonts in the images and score the quality of the images. (Each observer is required to recognize the image fonts and score the image quality (1~10 points), the higher the score, the better the visual quality of the human eye. The average of the results from 30 observers was chosen as the final result.

The comparison of the results of the recognition rate of the generated fonts is shown in Figure 2. As can be seen from the figure, the recognition rate of various calligraphy fonts generated by the algorithm in this paper is higher than that of the other two models, except for the cursive (90.42%) and Shu font (90.71%), which are complicated in their own structure and more personalized, resulting in lower recognition rates, the correct recognition rate of other fonts reaches more than 95%, and for the calligraphy fonts such as Regular, Clerical, and Willow fonts which have clear font frames and neat structure, the correct recognition rate is More than 96%.

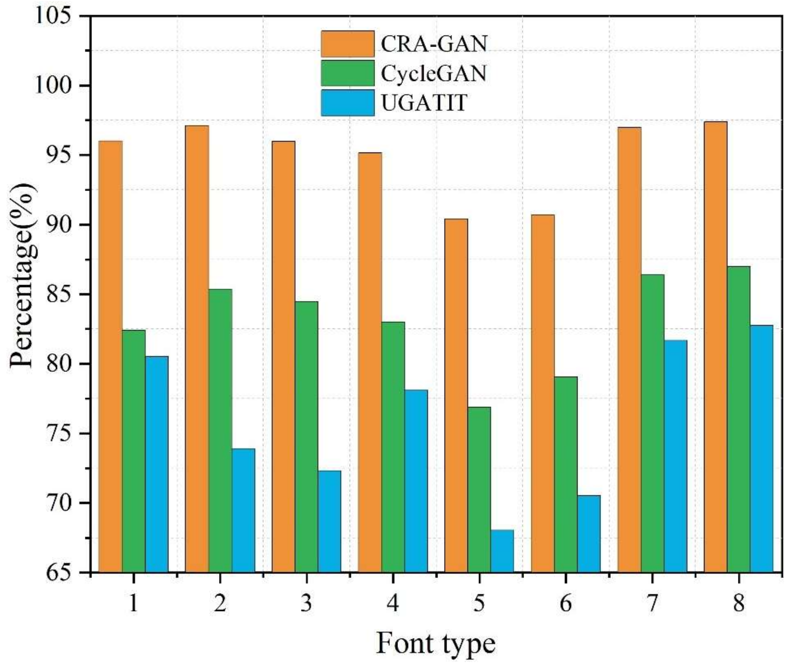

Fig. 2. The result comparison of the resulting font recognition rate

The comparison of the MOS value scoring results of the generated fonts is shown in Fig. 3. In terms of visual quality, it can be seen that the MOS value scoring results of the CRA-GAN results are centered on more than 8.5 points, which is much higher than the other comparative models, indicating that the generated images are in line with the observation characteristics of the human eye for calligraphic images.

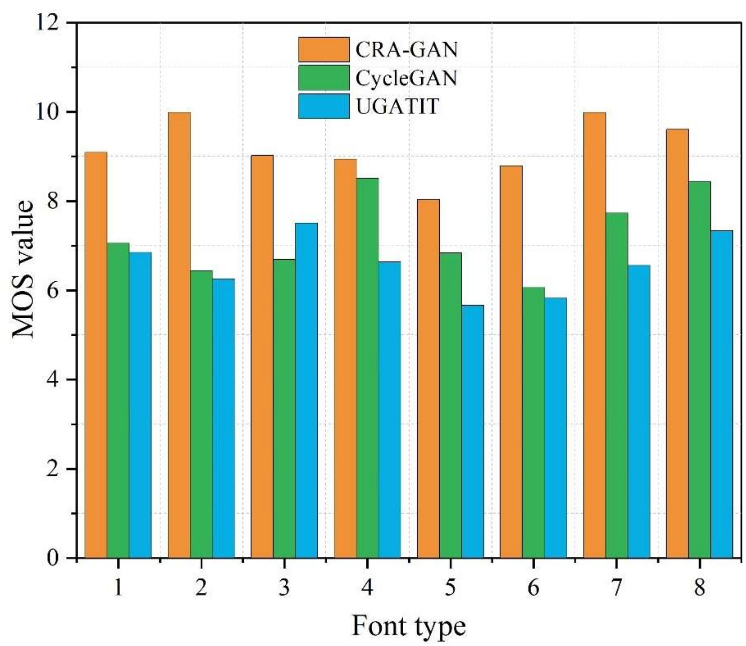

Fig. 3. The resulting color MOS value was compared

FID is used to measure 2 metrics of GAN networks: quality and diversity of generated images by extracting features from the middle layer and then modeling the distribution of these features with a normal distribution; lower FID means higher quality and diversity of images. Compared to IS, FID has better robustness to noise. Its calculation formula is as follows:

$$F = \| \mu_r - \mu_g \|^2 + T_r(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}) \quad (18)$$

Where: F is the result of FID calculation; T_r is the sum of elements on the diagonal of the matrix.

The FID calculation results are shown in Table 4.The FID calculation results of the CRA-GAN algorithm (204.361) are much lower than those of the CycleGAN method (857.4332) and the UGATIT method (1034.27), which indicates that the CRA-GAN algorithm has significantly improved in generating both the diversity of calligraphic fonts and visual quality. Overall, the CRA-GAN algorithm is not only capable of generating calligraphy images in many different styles, but also has a high recognition rate and good image quality, which is especially suitable for the field of generating fonts with strong structure such as Clerical Script and Liu Style. It is also found that the image quality and recognition rate do not exactly correspond to each other, with the highest recognition rate for the official script fonts but the best visual evaluation results for the Liu style images, and the lowest recognition rate for the cursive script fonts but the worst evaluation quality for the Shu style images, indicating that the human eye's viewing of the calligraphy images does not entirely depend on whether it can correctly recognize the fonts in the calligraphic works, and also further demonstrating the importance of the emotional styles in calligraphy in the generation of the calligraphic images and the It also further demonstrates the importance of emotional style in calligraphy in the generation and study of calligraphy images.

Table 4. FID calculation results

Index	Model		
	CycleGAN	UGATIT	CRA-GAN
FID value	857.4332	1034.27	204.361

4. Design of the virtual calligraphy display system

4.1. Overall architecture and technology

1)SSH framework. The framework of the artwork online display system is developed using the Struts+Spring+Hibernate framework based on the MVC pattern. the MVC pattern application is based on the model, view, controller as the core, the controller completes the information acceptance, calls different models according to the different needs to solve the user request, and then the results of the system processing is fed back to the view, and the user output is provided through the interface. Output for the user.

2) Database design. The design of this system chooses to use Oracle database, which has powerful data management functions, can permanently save a large amount of data to ensure the reliability and sharing of data, and its distributed processing functions can support the data warehouse operation, in which the information guidelines, view update guidelines, guaranteed access guidelines, and logical and physical independence guidelines can provide strong support for the design of the system.

3) B/S model. The system is carried out in B/S mode, which is the most popular system design framework, and can better adapt to the needs of Internet technology development. Users can access the resources through web browsers, which is more convenient to use, and users can access the online display system of art works on computers, cell phones and other terminal devices, which reduces the development cost of the system, completes the development in one go, and enables the access to the same database from different geographical locations, which helps in information protection and rights management. It helps information protection and permission management.

4) Virtual Reality Modeling Language. VRML is a virtual reality modeling language released by the alliance, which is an open, industry-standard, extensible scene description language, and the technology provides solutions for the three-dimensional animation effects in web pages and user interaction problems. The technology is to provide three-dimensional graphics library on the user side,

and the web page runs according to the need for coloring and rendering, effectively reducing the amount of data transmitted by the network.

5) Online display system design case. The online display system based on interactive virtual technology is based on product management, providing online display and exhibition services, as well as analyzing data on exhibits, collecting users' browsing information, providing an efficient information display platform for product exhibition and trading, etc., and at the same time providing information support for decision-making. The functions of the system include audio-visual module, interactive module and social module.

The system adopts the development mode of three-tier architecture, based on J2EE platform, according to the MVC design and development ideas, to realize the modularization of the system and component development, users can access the system through the B/S mode. At the beginning of the system design, it must pay attention to the establishment of the user's psychological model, and collect the user's demand, user's behavioral habits and typical psychological activities through the way of designing and researching, so as to provide the basis for the system design. Figure 4 shows the overall structure of the online display system.

The online display system needs to deal with graphic engine issues such as graphic rendering, lighting, and interaction in three dimensions. The user is using a multi-user system, which can realize automatic analysis and calculation according to the browsing experience and habits of the customers, realize business intelligence pairing according to the interests of the users, and establish an intelligent navigation system. The system is built based on the VRML multi-user environment to provide users with a real experience environment, which makes the selection and evaluation of goods more convenient.

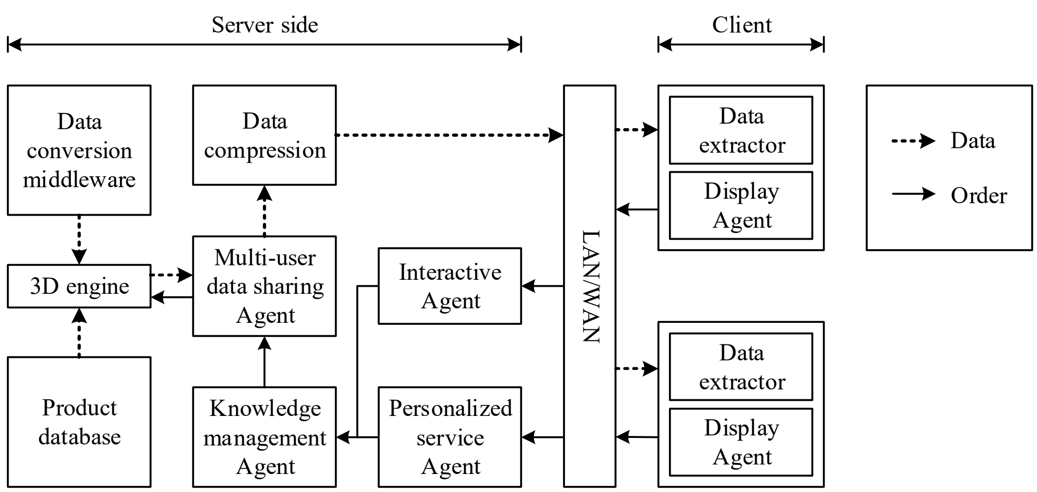

Fig. 4. shows the overall structure of the system online

4.2. *User experience results and analysis*

In order to clarify the effect of the artificial intelligence CRA-GAN model proposed in this paper on the reconstruction and virtual display of calligraphy seal cutting, this paper carries out a questionnaire survey on the audience after actually experiencing the virtual display system. A total of 100 questionnaires were collected, of which 100 were valid. The user's choices in the questionnaire were assigned scores: 100 for agree; 80 for relatively agree; 60 for can't say; 40 for not quite agree; and 0 for disagree. In the basic information of the statistical questionnaire, 48% are male and 52% are female, the age distribution of the respondents is between 18-52 years old, and their positions are students, teachers, and public officials. The degree of their hobby of calligraphy as the standard division into calligraphy professional audience, calligraphy learning audience and general tour audience three, its distribution accounted for 32%, 33% and 35% of the total number of people. The results of the experience factor optimization statistics are shown in Table 5.

Table 5. Experience factors select the results of the results

Experience classification	Primary weight	Experience factor	Secondary weight	Total weight
Narrative experience	0.2	Narrative activity	0.154	0.0308
		Narrative accuracy	0.069	0.0138
		Narrative attraction	0.114	0.0228
		Narrative richness	0.019	0.0038
		Narrative coherence	0.039	0.0078
		Narrative substitution	0.227	0.0454
		Storytelling	0.105	0.021
		Narrative linearity	0.13	0.026
		Narrative memory	0.143	0.0286
Emotional experience	0.2	Novelty	0.203	0.0406
		Sense of discovery	0.157	0.0314
		Satisfaction	0.139	0.0278
		Pleasure sense	0.127	0.0254
		Sympathetic sense	0.077	0.0154
		Comfort	0.163	0.0326
		Sense of manipulation	0.134	0.0268
Sensory experience	0.2	Environmental authenticity	0.203	0.0406
		Interface clarity	0.107	0.0214
		Dubbing comfort	0.109	0.0218
		Immersion	0.131	0.0262
		Device prompt note	0.035	0.007
		Visual immersion	0.175	0.035
		Information accuracy	0.085	0.017
		Tactile perception	0.155	0.031
Cognitive experience	0.2	Knowledge learnability	0.303	0.0606
		Whole and part	0.419	0.0838
		System applicability	0.278	0.0556
Interactive experience	0.2	The accessibility of interaction	0.411	0.0822
		Clarity of prompt	0.212	0.0424
		The rationality of feedback	0.145	0.029
		Interesting interaction	0.232	0.0464
Total	1	31	5	1

By counting the results of the questionnaire, the statistical results of each score of the questionnaire are shown in Table 6. The total score is the sum of the scores of the corresponding experience factors multiplied by the corresponding weights, and generally 100-80 points are considered to be very satisfactory to the user, 80-60 points are considered to be relatively satisfactory to the user, and 60 points or less are considered to be unsatisfactory to the user. The calculated total score is 89.31728, which is between 100-80 points, indicating that users are highly satisfied after actually experiencing the virtual display system.

Table 6. Questionnaire score statistics

Questionnaire problem	Experience factor	Total weight value W	Score C
Very vivid	Narrative activity	0.0308	88.23
Very accurate	Narrative accuracy	0.0138	92.09
That's attractive to me	Narrative attraction	0.0228	87.96
The story is rich	Narrative richness	0.0038	91.43
The story is consistent	Narrative coherence	0.0078	96.25
It's very native	Narrative substitution	0.0454	89.57
Very interesting	Storytelling	0.021	83.88
The story is characterized by novel and nonlinear	Narrative linearity	0.026	89.57
The historical knowledge of the narrative is easy to remember	Narrative memory	0.0286	92.82
The show is interesting, I think it's new	Novelty	0.0406	90.96
Let me be interested in exploring the contents of the scene	Sense of discovery	0.0314	89.69
Show content to satisfy my expectations	Satisfaction	0.0278	90.59
The experience is interesting and pleasant	Pleasure sense	0.0254	88.71
The experience content can resonate with me	Sympathetic sense	0.0154	93.9
The content of the system makes me feel comfortable and easy to be affected	Comfort	0.0326	87.24
The interaction has a sense of manipulation	Sense of manipulation	0.0268	92.77
The picture is very real and is very similar to the actual scene	Environmental authenticity	0.0406	89.36
The interface between the 18 and virtual scenarios is clear	Interface clarity	0.0214	86.48
The score made me feel comfortable	Dubbing comfort	0.0218	92.81
I feel immersed in my helmet with a helmet	Immersion	0.0262	93.06
The note makes my operation more accurate	Device prompt note	0.007	86.22
The picture makes me feel like I'm in my place	Visual immersion	0.035	86.18
The scene of the scene is accurate and clear	Information accuracy	0.017	88.67
Handle operation convenient, manipulative and feedback	Tactile perception	0.031	89.22
The history of history is rich and easy to learn	Knowledge learnability	0.0606	87.45
The knowledge of the system intermediary is connected to the interconnection	Whole and part	0.0838	90.35
I think the system can be accepted by many people and I will recommend more people	System applicability	0.0556	89.92
Interaction is intuitive and easy to learn	The accessibility of interaction	0.0822	91.3
The tips for interaction are clear and easy to understand	Clarity of prompt	0.0424	84.72
I can feel the feedback of the interaction and have no wrong operation	The rationality of feedback	0.029	85.37
Interaction makes me feel interesting	Interesting interaction	0.0464	87.94

For the five experience categories evaluated by the system, the scores were counted, and the results of the scores of each experience category in the questionnaire are shown in Table 7. It was found that the scores for the five experiences of "narrative experience, emotional experience, sensory experience, cognitive experience and interactive experience" were 90.2, 90.55, 89, 89.24 and 87.33 respectively, which were all in the range of 80-100 points, indicating that the users had a good experience. From the statistics of the scores of each experience category, the audience's scores for the narrative experience part are slightly higher than the other four parts of the experience, indicating that the narrative experience part of the system has been recognized by the audience.

Table 7. Questionnaire experience classification score statistics

Experience classification	Score
Narrative experience	90.2
Emotional experience	90.55
Sensory experience	89
Cognitive experience	89.24
Interactive experience	87.33

5. Conclusion

In this paper, using the GAN family of models, we propose a calligraphic font generation model based on generative adversarial network (CRA-GAN) to address the shortcomings of the existing methods for generating fonts for seal engraving, and construct an online display system for the virtual experience of calligraphy. The results show that:

1) The calligraphy fonts generated by the CRA-GAN model proposed in this paper can better capture the details and global information of the fonts, and have good generation effect, and its accuracy of the recognition rate of the eight calligraphy fonts is above 90%. The MOS scores of the text images processed by this paper are concentrated in more than 8.5 points, which is much higher than other comparative models, and the recognition results are in line with the observation characteristics of the human eye for calligraphy images.The CRA-GAN method has significantly improved the diversity of the calligraphic fonts and the visual quality of the generated calligraphic fonts, and the result of its FID is 204.361, which is much lower than that of the other methods.

2) This paper evaluates the user's system experience from five aspects: narrative experience, emotional experience, sensory experience, cognitive experience and interactive experience. We refine the experience factors of the five experience items and design a questionnaire. The weight of each influence factor on the final result is determined by the merit comparison method, and the final score of the system is calculated to be 89.31728, which shows that the users' satisfaction is very high after actually experiencing the virtual display system.

References

[1] Liu, H., Lyu, R., Yang, G., & Ze, Y. (2023, December). Identifying Chinese Calligraphy Styles by Simple Stroke Yi. In 2023 16th International Symposium on Computational Intelligence and Design (ISCID) (pp. 180-183). IEEE.

[2] Shi, X. (2019). Chinese calligraphy as “force-form”. Journal of Aesthetic Education, 53(3), 54-70.

[3] Haider, H. (2021). Calligraphic architecture: Stroke to form, space and surface. How does Arabic calligraphy influence the design process of Zaha Hadid in her creation of architectural forms?. Fields: Journal of Huddersfield Student Research, 7(1), 1-21.

[4] Xing, L. (2018). Design and Application of a Fuzzy Evaluation Algorithm for Stroke Force in Calligraphy Teaching. International Journal of Emerging Technologies in Learning, 13(7).

[5] Pan, Y., Lyu, R., Nie, Q., & Meng, L. (2020, November). Study on the emotional image of calligraphy strokes based on sentiment analysis. In 2020 5th international conference on communication, image and signal processing (CCISP) (pp. 264-269). IEEE.

[6] Chen, M., & Zhang, X. (2024, January). Progressive Chinese calligraphy stroke extraction. In 2024 International Conference on Informatics Education and Computer Technology Applications (IECA) (pp. 47-53). IEEE.

[7] Amir, A., & Kamel, W. (2019). Representations of the Shape Penetration in Calligraphy Compositions. International Design Journal, 9(1), 351-358.

[8] Shi, Z., Xu, B., Zheng, X., & Zhao, M. (2017). A Chinese character structure preserved denoising method for Chinese tablet calligraphy document images based on KSVD dictionary learning. Multimedia Tools and Applications, 76, 14921-14936.

[9] Hui, G. (2019). Miyoseki Shimura's enthusiasm for Chinese calligraphy and seal cutting. Voice of Friendship, 1.

[10] Choi, J. (2018). Study of Evolution Pattern of Seal. Indian Journal of Public Health Research & Development, 9(8).

[11] Li, Z. (2022). An Analysis of Deng Sanmu's Seal Cutting Emulation-Centering on Ancient Seal, Qin Seal and Han Seal. International Journal of Advanced Culture Technology, 10(4), 219-225.

[12] Xu, S. (2024). Integration of traditional elements and modern education-The inheritance and reform of seal cutting art education. Arts Educa, 38.

[13] Wang, Y., & Zhang, X. (2023, May). Calligraphy image processing with stroke extraction and representation. In International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2022) (Vol. 12604, pp. 226-232). SPIE.

[14] Li, M., & Ren, G. (2023). Intelligent Evaluation Method of Calligraphy Characters Based on Deep Stroke Extraction. Advances in Computer, Signals and Systems, 7(10), 99-106.

[15] Matsumaru, T., & Narita, M. (2017). Calligraphy-stroke learning support system using projector and motion sensor. Journal of Advanced Computational Intelligence and Intelligent Informatics, 21(4), 697-708.

[16] Liu, Y., & Zhu, Q. (2017, July). Dynamic simulation of stroke trajectories in Chinese calligraphy based on writing momentum. In Ninth International Conference on Digital Image Processing (ICDIP 2017) (Vol. 10420, pp. 925-933). SPIE.

[17] Ma, Z., & Su, J. (2017). Stroke reasoning for robotic Chinese calligraphy based on complete feature sets. International Journal of Social Robotics, 9(4), 525-535.

[18] Lin, G., Guo, Z., Chao, F., Yang, L., Chang, X., Lin, C. M., ... & Shang, C. (2021). Automatic stroke generation for style-oriented robotic Chinese calligraphy. Future Generation Computer Systems, 119, 20-30.

[19] Chen, F., Wang, C., Yao, X., & Sun, F. (2024). SPFont: Stroke potential features embedded GAN for Chinese calligraphy font generation. Displays, 85, 102876.

[20] Guo, D., Ye, L., Yan, G., & Min, H. (2023). CCD-BSM: composite-curve-dilation brush stroke model for robotic chinese calligraphy. Applied Intelligence, 53(11), 14269-14283.

[21] Bi, F., Han, J., Tian, Y., & Wang, Y. (2022). SSGAN: generative adversarial networks for the stroke segmentation of calligraphic characters. The Visual Computer, 1-10.

[22] Zhang, K., Zhang, R., Yin, Y., Li, Y., Wu, W., Sun, L., ... & Pan, Y. (2022). Visual knowledge guided intelligent generation of Chinese seal carving. Frontiers of Information Technology & Electronic Engineering, 23(10), 1479-1493.

[23] Lei Liu,Jose J. Salazar,Honggeun Jo,Maša Prodanović & Michael J. Pyrcz. (2025). Minimum acceptance criteria for subsurface scenario-based uncertainty models from single image generative adversarial

networks (SinGAN). Computational Geosciences(1),6-6.

[24] Afaf Saad,Noha Ghatwary,Safa M. Gasser & Mohamed S. ElMahallawy. (2025). Automatic image generation and stage prediction of breast cancer immunobiological through a proposed IHC-GAN model. BMC Medical Imaging(1),6-6.

[25] XUANZHAODONG,YalinWang,JayShah,ValentinaGhisays,JiLuo,YinghuaChen... & YiSu. (2025). End-to-End 3D CycleGAN Model for Amyloid PET Harmonization. Alzheimer's & Dementia(S8),e095403-e095403.

[26] Wei Yangzhe. (2021). Detection of lane line based on Robert operator. Journal of Measurements in Engineering(3),156-166.