International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-13

Article

The Facilitating Mechanism of Tang Calligraphy on the Formation of Shodo Style in Japan during the Heian Period in the Framework of Cross-cultural Exchange

Yang Wang [1,*], Ringah anak Kanyan [1] and Ge Kong [2]

[1] Faculty of Applied and Creative Arts, Universiti Sarawak, Kuching, Sarawak, 94300, Malaysia
[2] School of Arts, Qilu Institute of Technology, Jinan, Shandong, 250200, China
* Correspondence author: 22010360@siswa.unimas.my

Abstract: Tang calligraphy, as the southern part of East Asian calligraphy, had a profound influence on the formation and evolution of Shodo styles in Japan during the Heian period. In the context of cross-cultural communication, the study is to organize and complete the calligraphy style dataset including Tang Dynasty calligraphy images and Japanese Shudo images. The stack self-encoder in deep learning algorithm is used as the feature extraction algorithm, and the Softmax classifier is used for style recognition, and the recognition test is conducted on the expanded dataset for the characteristics of the five calligraphic Chinese character styles. On this basis, the dual influence of Tang Dynasty calligraphy on the aesthetics and style of Japanese Shudo fonts is further explored. The experimental results show that the recognition accuracy of this paper's method on five calligraphic Chinese character styles reaches 98.95%, and the recognition accuracy on cursive and running script styles is significantly higher than that of other recognition methods, which proves the validity of this paper's method on the recognition of calligraphic styles. At the level of influence, Tang Dynasty glyphs (which are one of the factors influencing the foundation of Japanese Shudo aesthetics), replacement lines and unified structures have a significant impact on Japanese Shudo font styles.

Keywords: stack self-coding; calligraphic style recognition; Tang Dynasty stroke structure; Japanese Shudo

1. Introduction

In the process of overseas dissemination of Chinese calligraphy, it has had the most profound influence on Japan within the cultural circle of Chinese characters. In the two-way interactive communication history between China and Japan, one-way dissemination of Chinese calligraphy to Japan is the mainstream [1]. The long history of Chinese-Japanese calligraphy exchanges has led to a better development of calligraphy in Japan. The famous Japanese Chinese newspaper, Chinese Herald, constantly publishes exhibitions of Chinese calligraphy held in Japan, as well as academic symposiums of Chinese and Japanese calligraphy researchers, and the rich academic research results in turn provide a solid theoretical foundation for the integration of calligraphy into the practice of Chinese language teaching in Japan [2-3]. With the promotion of international Chinese language education, Chinese calligraphy has also become an important part of Chinese language promotion due to its unique cultural connotations, and a large number of studies on the overseas dissemination of Chinese calligraphy have appeared in recent years [4]. However, there is a relative lack of research on the dissemination of Chinese calligraphy in Chinese language teaching and Chinese language teaching in Japan.

In the process of overseas dissemination of ancient Chinese calligraphy, it had the most profound influence on Japan within the cultural circle of Chinese characters. Japanese calligraphy originated from the introduction of Chinese characters, and before the late Tang Dynasty, Japanese calligraphy and

Chinese calligraphy maintained a trend of development in step with each other, and after the late Tang Dynasty, Japan combined with the characteristics of its own country and developed localized fonts, such as kana, on the basis of Chinese characters [5-7]. The Heian Period is the 391-year period between the transfer of the Japanese capital to Heijo-kyo (794 A.D.) and the establishment of the Kamakura regime (1185 A.D.) [8]. Because of the cultural similarities between the Nara Period and the Heian Period, which were also influenced by China, there were Japanese calligraphers who were deeply influenced by Chinese calligraphy, such as Mitsumoto, Kukai, Emperor Saga, and Tangerine Yisei [9-11]. Among them, the two monks who visited the Tang Dynasty as emissaries to the Tang Dynasty, Mr. Mitsumoto and Mr. Kukai, introduced a lot of Chinese calligraphic works from China [12]. During this period, extra-territorial regions studied Tang culture through official or private means of exchange, with the most frequent cultural exchanges occurring in Japan and Korea [13-15]. Therefore, it is important to further investigate the promotion mechanism of Tang calligraphy on the formation of calligraphy style in Japan during the Heian period, with a view to providing useful reference for promoting the construction of a more open, inclusive and standardized modern international Chinese language education system and enhancing the international influence of Chinese culture.

The study firstly collects different Tang Dynasty calligraphy and Japanese Shodo images on the web, and the collected images are preprocessed. Then based on the stack self-encoder technique in deep learning, combined with Softmax classifier for calligraphy style recognition, through five kinds of Tang Dynasty styles of calligraphy Chinese characters as the original dataset, by adjusting the values of the number of paths N and structural elements M of morphology neural network, to validate the effect of this paper's method on the accuracy of the style recognition of the Chinese characters of calligraphy and to Comparison. Finally, comparing the relationship between the two types of calligraphy (Tang Dynasty Calligraphy and Japanese Shodo) in terms of font aesthetics and font style characteristics, which is a reference value for understanding the influence mechanism of Tang Dynasty Calligraphy in the East Asian cultural circle.

2. Calligraphy image pre-processing

2.1. Training process for images

First of all, we need to get the calligraphy/shudo pictures of Tang Dynasty and Heian Period of Japan from the website, due to the picture type is too complex, in order to facilitate the training of the network model for the picture, the picture is preprocessed, and the image preprocessing methods usually used include inverse phase change, size normalization, binary, picture de-noising and other operations. For the content of this study, the aim is to obtain images with style black on white background and better quality. The process of acquiring the image after processing and inputting it to the network model to start training is shown in Fig. 1.

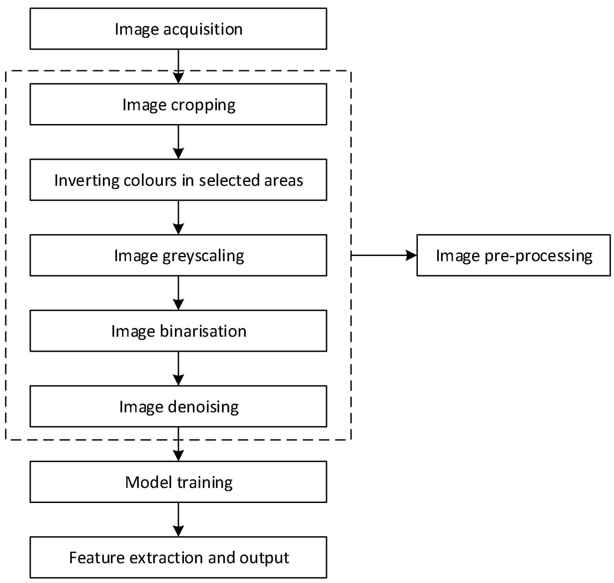

Figure 1. Image Training Flowchart

2.2. Image Preprocessing

2.2.1. Image inverse color change

Some of the web postings are intercepted to form a single word image with white characters on a black background, and in order to process the image in a uniform way, this part of the image is subjected to black and white inverse color change. Image inverse color change is the image of black and white pixel inversion, for the generalized inverse color transformation can also be applied to color images, that is, all the pixels of the value of the complementary. The principle of operation is to subtract the RGB value from 255, and the RGB value range is 0-255.

2.2.2. Image Grayscaling

There are four main methods for grayscaling color images: the maximum value method, the average value method, the component method and the weighted average method. Among them, the maximum value method is by selecting the maximum value of the three component pixels R, G and B in the color image, which will be used as the pixel value for grayscaling process. The average method is to find the average value of the R, G, and B component pixels in a color image and use it as the pixel value for grayscaling. Component method is to take the pixel values of R, G, and B components in the color image as the grayscale value of the grayscale image, and select one of them for image grayscaling according to the specific needs. The weighted average method is to weight the three components of R, G, and B in the color image with different weights and use them as the gray value of the grayscale image.

2.2.3. Image binarization

Image binarization is a very important image processing technique in digital image processing, before obtaining the binarized image, the original picture will be grayscaled first, and the grayscale map has many levels of color depths between black and white, the binarization is the grayscale map all the grayscale values are divided into 0 and 255. the image after the binarization process can still react to the whole image and local features and make the The entire image appears in black and white.

Assuming that the coordinate values are used to represent the grayscale of the image, let $f(x,y)$ denote the points on the original image that already exists, and $g(x,y)$ denote the points on the binarized image, with the upper-left corner of the image as the origin of the coordinates, the direction of the x direction is to the right of the horizontal coordinates with the origin as the starting point, and the direction of the y coordinates is to the downward of the vertical coordinates with the origin as the starting point; the width of the image is w, and the height is h; T represents the threshold value we set. After the image is binarized, the pixel points on the original picture are processed as 0 or 255, the binarization process is shown in equation (1):

$$g(x,y)=\begin{cases}0, & f(x,y)<T\\ 255, & f(x,y)\geq T\end{cases} \tag{1}$$

In particular, $0\leq x\leq w$ and $0\leq y\leq h$.

2.2.4. Image denoising

Noise in an image is the presence of unnecessary or redundant disturbing information in the image data that can affect the quality of the image. High quality images are obtained by performing noise reduction on low quality images. The noise reduction process removes useless information from the image while maintaining the integrity of the original information as much as possible, reducing the factors that affect the clarity of the image.

For noise in an image, two commonly used methods of image denoising are: median filtering and mean filtering. The use of filters can make the target image in the noise suppression effect can also try to retain the details of the image features. The quality of the image filtering process can directly affect the reliability and effectiveness of the subsequent analysis and processing of the image.

2.3. Data set construction

In this paper, a complete dataset of six styles of calligraphy and Shudo based on 3,000 commonly used kanji characters has been compiled, covering the two major systems of Chinese calligraphy of the

Tang Dynasty (including Regular Script, Running Script, Cursive Script, and Clerical Script) and Japanese Shudo of the Heian Period (including Wabi-sabi Shudo and Kana Calligraphy, etc.). The dataset used in this paper is based on the collection of calligraphy and Shudo works from different historical periods, such as inscriptions, handwriting, and sutras, using the Internet and literature. Among them, the same font contains different contents written by different calligraphers, which ensures the richness of the morphology of calligraphic word images and makes the dataset more universal.

After finishing, the final collected dataset contains six styles of calligraphic fonts, each with about 7000 characters. Among them, some styles are collected in limited numbers due to fewer historical circulating works, for example, 3427 single images are collected for a certain category. The total number of all the calligraphy and Shudao images total 38239. The network model training needs to be based on a large number of data feature learning to complete the style recognition task, and a large number of high-quality images are the basis for completing the recognition and generation of calligraphy and Shudao styles.

3. Calligraphic style recognition based on stack self-coding

3.1. Self-Encoder Neural Networks

AE is a shallow ANN proposed by Rumelhart, which is an unsupervised learning network. A single-layer AE is a neural network structure consisting of an input layer with d neurons, a hidden layer with h neurons, a decoding layer with d neurons and an activation function $f(\cdot)$. During training, it maps $x \in R^d$ to the hidden layer and obtains $h \in R^{d'}$ (d' is the number of neurons in the hidden layer). The network structure corresponding to this step is called "encoder". Then, h is mapped by the "decoder" to the output layer with the same number of neurons as the input layer to obtain $\hat{x} \in R^d$, which is the "reconstruction process". These two steps can be expressed as follows:

$$h = f\left(W_h x + b_h\right) \tag{2}$$

$$\hat{x} = f\left(W_{\hat{x}} x + b_{\hat{x}}\right) \tag{3}$$

where W_h and $W_{\hat{x}}$ denote the weights of the input-implicit and implicit-output layers, respectively, and b_h and $b_{\hat{x}}$ denote the deviations of the input-implicit and implicit-output layers, respectively.The Sigmoid function is the most commonly used activation function, viz:

$$f(\cdot) = 1/(1+e^{-x}) \tag{4}$$

It should be noted that this paper has the following provisions to simplify the calculations and reduce the parameters of the model:

$$W_h = W'_{\hat{x}} = W \tag{5}$$

Therefore, only three sets of parameters need to be learned: W, b_h and $b_{\hat{x}}$. The ultimate goal is to minimize the error between the input values and the reconstructed values as much as possible, i.e.,:

$$\underset{W,b_h,b_{\hat{x}}}{\arg\min}\left[J(x,\hat{x})\right] \tag{6}$$

When x is input, W, b_h and $b_{\hat{x}}$ will determine the output value of $\hat{x}$, and $J(x,\hat{x})$ is the error between the input value and the reconstructed value. Therefore, the weight update rule can be defined as (where η is the learning rate):

$$W = W - \eta \frac{\partial J(x,\hat{x})}{\partial W} \tag{7}$$

$$b_h = b_h - \eta \frac{\partial J(x,\hat{x})}{\partial b_h} \tag{8}$$

$$b_{\hat{x}} = b_{\hat{x}} - \eta \frac{\partial J(x,\hat{x})}{\partial b_{\hat{x}}} \tag{9}$$

When training is complete, the decoding layer is removed along with the parameters it contains, and only the feature representations learned at the implicit layer are retained and used as inputs for later higher-level classification or prediction tasks. Note that throughout the decoding process, the model relies only on the implicit layer information h and encodes it as feature input to the decoding layer.

The words into x imply that the feature representation of the implicit layer retains enough of the original input information and that this nonlinear transformation of the training weights and biases can be considered feasible and efficient. Therefore, stacking the AEs in this training manner will result in effective raw data information and feature representations, which is the reason why the AE technique is chosen to extract deep features from hyperspectral data.

SAE is an unsupervised deep learning network consisting of multiple AEs stacked layer by layer, which has more powerful data feature representation capability than shallow ANN.The programming of SAE neural network is to execute each AE sequentially in order from front to back, and the output of the previous AE is used as the input of the next AE, and similarly the decoding process of the SAE neural network is to reverse the order of executing each AE.

The training method of SAE is unsupervised layer-by-layer greedy training, each time an implicit layer is trained to complete the AE optimization, and then begin to train the next implicit layer, until the last implicit layer is trained. In summary, SAE neural network is multiple AEs to accomplish the task of layer-by-layer feature extraction, with the aim of making the final obtained feature representation more representative and discriminative.

3.2. Stacked Self-Coding Feature Extraction

The stack self-coding algorithm increases the number of hidden layers on the basis of sparse self-coding to extract deeper features. The feature extraction is carried out, using double hidden layer, the weight matrix of the hidden layer is obtained, the weight matrix output from the second hidden layer is the second layer of deep features we need, and then continue the training of Softmax classifier, using fine-tuning function to adjust the error. The deep feature extraction and classification steps are as follows:

(1) Take the pre-processed calligraphy works as the original data as the input data, train the deep network structure and parameters according to the training algorithm of sparse self-encoder, and calculate the output of the first hidden layer by combining these parameters with the input data.

(2) Take the output of the first hidden layer as the input of the second hidden layer, and process the data of the second hidden layer using the same method as in step (1).

(3) After obtaining the second layer depth features, the features are put into the classifier for classification and recognition. This topic uses both SVM and Softmax classifiers for classification.

(4) If Softmax classifier is used, the training set data labels are used to train the parameters of the classifier and the cost function values and the bias values for each parameter are calculated based on the trained parameters.

(5) The network parameters of the two hidden layers and the Softmax classifier are used as inputs, and the cost function is minimized by iterative computation using the L-bfs algorithm, and the parameters when the number of iterations is completed or convergence has been reached are used as the optimal classification parameters.

(6) Based on the optimal parameters, the test set data is subjected to style identification using the Softmax classifier.

3.3. Selection of the Consideration Function

The goal of the self-encoding technique is to minimize the error between the reconstructed value and the input value, the reconstruction error can be calculated by the calculation of the cost function Cost, and this error can be minimized by the iterative optimization of the L-bfs function.

Different cost functions of the self-encoder can lead to different performance of the deep network, and the global minimum mean square error (MSE) cost function is commonly used among the commonly used cost function methods. The MSE function is shown in Equation (10) when the number of samples is n, x_i denotes the i th input and y_i denotes the i th output:

$$J_{MSE}(\theta)=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{2}\left\|y_i-x_i\right\|\right) \tag{10}$$

If there is only cost function MSE, the self-encoder is prone to overfitting phenomenon during the working process, so the attenuation coefficient term is added, this attenuation coefficient term we have mentioned in Chapter 2, where we set the coefficient term as $J_{weight}(\theta)$, and the sparse self-encoder adds the sparse term $J_{sparse}(\theta)$ on top of that, so A complete cost function formula should profile as:

$$J_{cost}(\theta) = J_{MSE}(\theta) + J_{sparse}(\theta) + J_{weight}(\theta) \tag{11}$$

Although good recognition results have been achieved using the traditional MSE cost function, this cost function is more sensitive to non-Gaussian noise and less local information extraction, in order to suppress this drawback, the local maximum correlation entropy (MCC) cost function is used.

In MCC cost function, correlation entropy is the more important method. Literature uses correlation entropy to measure the local similarity, this method has been widely used in the field of pattern recognition, and in the existing literature reflects the effect of superiority over the MSE cost function. The defining equation of correlation entropy is shown in Eq. It is a generalized algorithm to calculate the similarity between two random variables:

$$V_\sigma(X,Y) = E\left[\kappa_\sigma(X-Y)\right] \tag{12}$$

$E[\cdot]$ is the expectation function and κ_σ is a Gaussian kernel function that satisfies Mercer's theory. The use of kernel function in correlation entropy makes it possible to map the input data to a higher dimensional space, in the traditional kernel function the samples are correlated with each other, whereas the kernel function used here makes each sample uncorrelated.

$$\kappa_\sigma(X-Y) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(X-Y)^2}{2\sigma^2}\right) \tag{13}$$

MCC is a measure of local correlation entropy, while MSE is calculated globally, so using the MCC cost function can show better adaptability and accuracy. Moreover, the joint probability density of the data in the experiment is often unknown, so the correlation entropy is defined as:

$$\hat{V}_\sigma(X,Y) = \frac{1}{n}\sum_{i=1}^{n}\kappa_\sigma(x_i - y_i) \tag{14}$$

This leads to the final definition of the MCC cost function, where n is the number of samples and m is the dimension of each sample, which is substituted for the MSE cost function in $J_{cost}(\theta)$. And it can be seen that when the correlation entropy is maximized, the cost function takes the minimum value.

$$J_{MCC} = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\kappa_\sigma\left(x_i^j - y_i^j\right) \tag{15}$$

3.4. Softmax Classifier

Softmax regression arose from logistic regression and is more suitable for multi-categorization tasks. In the Softmax regression setting, it is assumed that there are k classes, k is not less than 2, and the vector y is the label. When the training set is $\left\{\left(x^{(1)}, y^{(1)}\right), \cdots, \left(x^{(n)}, y^{(n)}\right)\right\}$, it can be obtained as $y^{(d)} \in \{1, 2, \cdots, k\}$. Assuming that the sample point obeys a certain probability partition, let the parameter be Φ: $p\left(y = i \middle| \Phi\right) = \phi_i, (i = 1, 2, \cdots, k)$, which denotes the probability that the sample point d will be recognized as the i th class.

Express this probability division in the form of exponential respectively as:

$$P\left(y = i \middle| x; \theta\right) = \frac{\exp(\eta_i)}{\sum_{j=1}^{k}\exp(\eta_j)} \tag{16}$$

This expression is analogous to a Sigmoid function, where η is the assumed natural parameter, such that $\eta = \theta^T x$, when the probability of the sample belonging to a certain class and the parameter θ are linearly related, there is:

$$P\left(y=i\left|x;\theta\right.\right)=\frac{\exp\left(\theta_i\cdot x^{(i)}\right)}{\sum_{j=1}^{k}\exp\left(\theta_j\cdot x^{(i)}\right)} \tag{17}$$

When k is 2, it is in the form of logistic regression. When there are n sample points in the training set, the logistic expression of the likelihood function is obtained:

$$l(\theta)=\sum_{i=1}^{n}\log\prod_{j=1}^{k}P\left(y=i\left|x;\theta\right.\right)+\frac{\lambda}{2}\sum_{i=1}^{k}\sum_{j=1}^{m}\theta_{ij}^2 \tag{18}$$

In the formula, λ is the attenuation coefficient, and the Bayesian regularization idea is used to make $l(\theta)$ a convex function, and the gradient descent method is applied to find the extreme value of this function, and finally the iterative formula of the parameters is obtained:

$$\begin{gathered}\theta_j=\theta_j-\alpha\nabla_{\theta_j}l(\theta)\\ \nabla_{\theta_j}l(\theta)=-\frac{1}{n}\sum_{i=1}^{n}\left[x^{(i)}\left(1\left(y^{(i)}=j\right)-P\left(y=i\left|x;\theta\right.\right)\right)\right]+\lambda\theta_j\end{gathered} \tag{19}$$

$\nabla_{\theta_j}l(\theta)$ is the result of the partial derivative of the likelihood function on θ_j . α is the learning rate parameter, each iteration of the calculation of the degree of parameter change is controlled by it, when α the smaller the better the effect, but the disadvantage is to reduce the operational efficiency, how to choose the appropriate value is a problem that needs to be studied.

4. Experimental results and analysis of calligraphy style recognition

4.1. Network training

In this experiment, the cross-entropy loss function is chosen for the loss function of the network, and the optimization algorithm adopts Adam's algorithm, with momentum factors of β_1 =0.9, β_2 =0.99, the learning rate is set to 0.001, the batch size is 32, and the number of iteration rounds in the training set is set to 50. The network is trained. The training set is divided into five training subsets according to the five Tang Dynasty calligraphic Chinese character styles: regular script, seal script, cursive script, running script, and official script, respectively.

Using the cross-entropy loss function, the loss value L_t of the current recognition network H_t is calculated by the predicted probability of each training sample belonging to the style category and the sample category label; the Adam optimization algorithm is used, with the learning rate set to 0.001, and the structural elements, the weights of the convolutional layer and the fully-connected layer ω_t , and the bias θ_t in the network H_t are updated by the loss value L_t , to obtain the network after the current iteration.

The trend of the recognition accuracy on the calligraphic Chinese character style task with the number of iteration rounds during the training process on both the training and validation sets is shown in Fig. 2. It can be seen that the curve is basically flat when epoch is 37, and the recognition accuracy of the network model on the two types of datasets for calligraphic Chinese character styles no longer changes, indicating that the network has converged at this time. The final recognition accuracy reaches 99%, and the recognition effect is optimal.

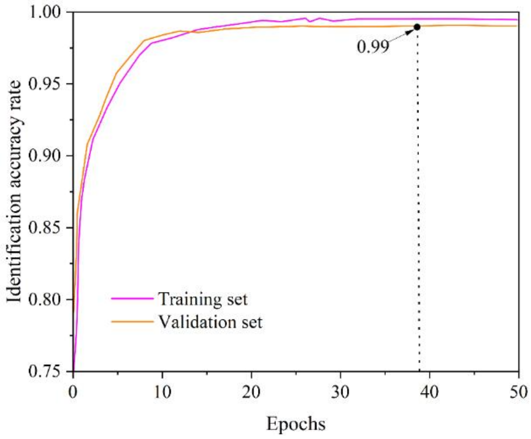

Figure 2. The trend of recognition accuracy as a function of the number of iterations

The trend of the loss function with the number of iteration rounds during the training process on the training and validation sets is shown in Fig. 3. From the visualized curves, it can be seen that the loss function of the training and validation sets gradually tends to 0 with the increase of the number of iteration rounds, at which time all the parameters in the network have been optimized through training.

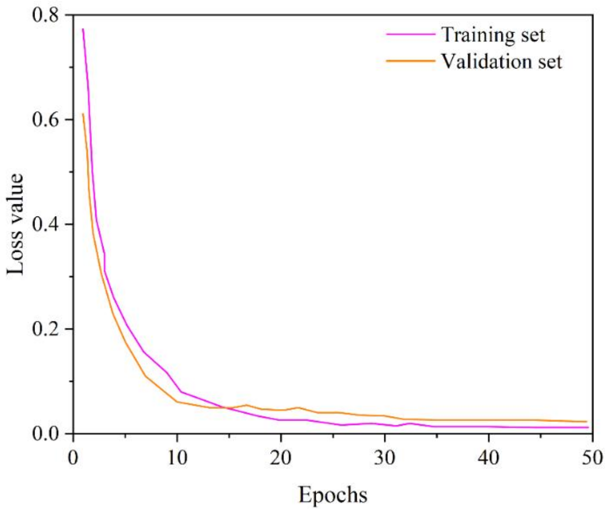

Figure 3. The trend of the loss function as the number of iterations increases

4.2. Experimental results and analysis

1. Experimental environment

All experiments in this paper are based on the following hardware and software platforms:

(1) Hardware: the processor CPU model is i7-9700K, the main frequency frequency is 3.60GHz, and the GPU is NVIDIA GeForce RTX 2080.

(2) Software: the operating system is Window 10, configured with TensorFlow 2.1.0 as the programming framework, using PyCharm programming software, the programming environment is python 3.7.

2. Calligraphy Chinese Character Style Recognition

The model combining stack self-encoder and Softmax is used to test on the test set, to demonstrate the effect of different values of N and M on the accuracy rate of calligraphy Chinese character style recognition by adjusting the number of paths N and the number of structural elements M in the morphological neural network in the inflated pooling subnetwork and to analyze the results of the recognition of each calligraphic style. The recognition accuracy is shown in Table 1.

According to the results of the table, it can be seen that the number of model paths N increases from 1 to 5, the change in the recognition accuracy rate of calligraphic Chinese character styles is more obvious, in which the improvement of the accuracy rate is higher from even to odd paths than from odd to even paths. For example, when M is 4, 8, 16 and 32, respectively, N increases from 2 to 3, the recognition accuracy is improved by 0.09%, 0.13%, 0.14%, 0.14%, 0.14%, and N increases from 4 to

5, the recognition accuracy is improved by 0.10%, 0.11%, 0.26%, 0.04%, respectively, while N increases from 1 to 2, or from 3 to 4, the recognition accuracy improves less significantly, by about 0.06% to 0.09%.

Table 1. Accuracy rate of character recognition for Chinese calligraphy styles

M	N							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
4	98.19	98.28	98.37	98.39	98.49	98.51	98.58	98.59
8	98.29	98.36	98.49	98.57	98.68	98.75	98.77	98.34
16	98.39	98.44	98.58	98.69	98.95	98.81	98.82	98.79
32	98.43	98.51	98.65	98.73	98.77	98.79	98.79	98.79

Fig. 4 shows the trend of recognition accuracy with N and M. It can be seen that the overall recognition accuracy is best at 98.95% when N=5 and M=16.

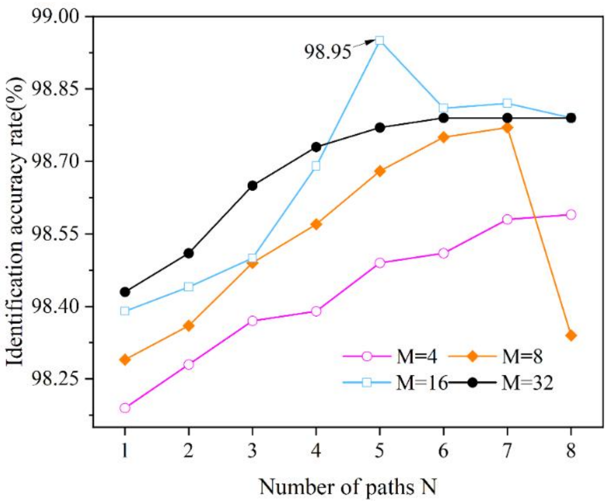

Figure 4. The trend of recognition accuracy as a function of N and M

In the following, a confusion matrix is used to evaluate the recognition of Tang Dynasty calligraphic Chinese character styles, which is calculated to obtain the recognition results for each calligraphic style. Through this matrix, various classification metrics can be calculated, the common ones are accuracy, precision, recall and F1 score, the higher the value of these metrics, the higher the accuracy and reliability of recognizing as a certain category. The style recognition method proposed in this paper is tested on a test set and the confusion matrix obtained is shown in Table 2.

Table 2. Results of recognizing the Chinese calligraphic and character styles

Ground truth	**Predicted value**					**Total**
	Regular script	**Seal Script**	**Cursive Script**	**Running script**	**Official script**	
Regular script	21183	0	8	78	10	21279
Seal Script	5	15073	11	5	36	15130
Cursive Script	21	7	15179	165	6	15378
Running script	155	2	121	14472	14	14764
Official script	10	29	6	7	14468	14520
Total	21374	15111	15325	14727	14534	81071

The precision rate, recall rate and F1 score calculated from the above table are shown in Table 3 and Fig. 5. From Table 3 and Fig. 5, it can be seen that the style recognition method in this paper obtains very high recognition rates on all five styles of calligraphic Chinese characters, where the average precision rate, recall rate and F1 score are 99.22%, 99.16% and 99.20%, respectively, which are up to the demand of practical applications. Due to the clear and simple structural features of seal script and clerical script, the precision rates of these two styles are 99.81% and 99.64% respectively, which are slightly higher than the other styles, and other metrics show the same results.

Table 3. Identification results of calligraphic Chinese character styles (%)

metric	Regular script	Seal Script	Cursive Script	Running script	Official script	Average Recognition Rate
Precision	99.18	99.81	99.14	98.33	99.64	99.22
Recall	99.62	99.68	98.76	98.03	99.73	99.16
F1 score	99.39	99.75	98.96	98.19	99.69	99.20

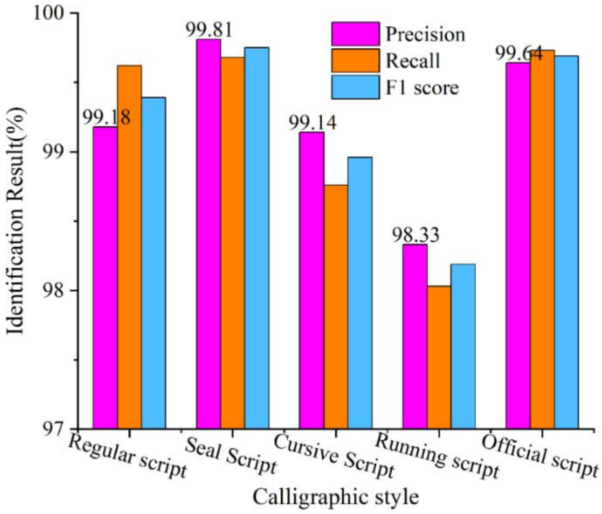

Figure 5. Results of recognizing the Chinese calligraphic and character styles

In order to better demonstrate the effectiveness of this paper's method on the task of calligraphic Chinese character style recognition, this paper's method is compared with other methods, and the results are shown in Table 4.

Inception V4 network, a recognition accuracy of 98.98% is obtained on running script; using DenseNet network for style recognition, the method proposed in this paper improves the recognition rate by 2% on cursive style and 2.30% on running script style; using the fine-tuned AlexNet network on the existing CADAL standard character library dataset for style recognition, the average recognition rate of five calligraphic styles recognition is the lowest. The method proposed in this paper effectively improves the recognition accuracy of calligraphic styles, especially in the more difficult to recognize cursive and running script styles, reaching 99.05% and 98.49%, respectively, which is significantly higher than other algorithms.

Table 4. Precision rate of Chinese character style recognition in calligraphy

metric	Regular script	Seal Script	Cursive Script	Running script	Official script
Inception V4	99.01	99.02	95.99	98.98	98.99
DenseNet	98.53	99.17	97.05	96.19	98.96
AlexNet	95.7	99.1	93.63	93.41	98.31
Methodology of This Article	99.14	99.74	99.05	98.49	99.61

5. The Influence of Tang Calligraphy on the Shodo Style of the Heian Period in Japan

5.1. The Influence of Tang Dynasty Calligraphy and Structure on the Aesthetics of Shudo Scripts

Table 5 shows the frequency and cumulative number of choices of aesthetics and style of each Tang Dynasty calligraphy to the Japanese Shudo style in the Heian period of Japan. It can be seen that the frequency of selection of aesthetics and style of Tang Dynasty calligraphic characters is higher than that of Nihonshudo, indicating that Tang Dynasty calligraphy and structure are positively proportional to the aesthetics and style of Nihonshudo fonts, i.e., the more times a font is selected for evaluating its aesthetics, the more often the corresponding stylistic vocabulary is selected as well. From Nihonshudo to Tang Dynasty calligraphic characters, the frequency of selection of aesthetics and style increases sequentially, indicating that with the addition of Tang Dynasty penmanship and structure, the "aesthetics" of Nihonshudo characters has been enhanced, and the sense of style has been enriched.

Table 5. Statistics of selected times of aesthetic and style words

	Font Aesthetic Selection(Frequency/ Total Frequency)	Aesthetic Change/%	Style Selection Frequency/Total Frequency	Style Change/%
Regular script	1022/2.011	-7.536	1863/3.732	-28.173
Japanese Calligraphy	957/1.743		1285/2.656	
Seal Script	1042/2.028	-6.462	2045/3.949	-32.402
Japanese Calligraphy	963/1.834		1328/2.681	
Cursive Script	1049/2.064	-6.028	2029/3.815	-28.042
Japanese Calligraphy	977/1.825		1339/2.689	
Running script	1039/2.023	-3.745	1865/3.432	-26.574
Japanese Calligraphy	993/1.858		1442/2.673	
Official script	1054/2.031	-4.482	1841/3.574	-24.346
Japanese Calligraphy	993/1.842		1465/3.012	

Figure 6 shows the mean values of the P_r for the aesthetic choices of the two groups of samples. It can be seen that the aesthetic choice curves of Tang Dynasty calligraphy and Japanese calligraphy overlap to a large extent, especially the extreme values of the aesthetic choice P_r. In different font models, the positions of the curves are relatively consistent. That is, after the operations of "line replacement" and "structure unification", the results of the aesthetic choices of the characters do not fluctuate much, indicating that the replacement of lines and the unification of structure have limited influence on the aesthetic perception of Chinese characters. This also implies that the shape of Tang Dynasty characters (the basic form of characters) is one of the fundamental factors influencing the aesthetic appreciation of Japanese calligraphy.

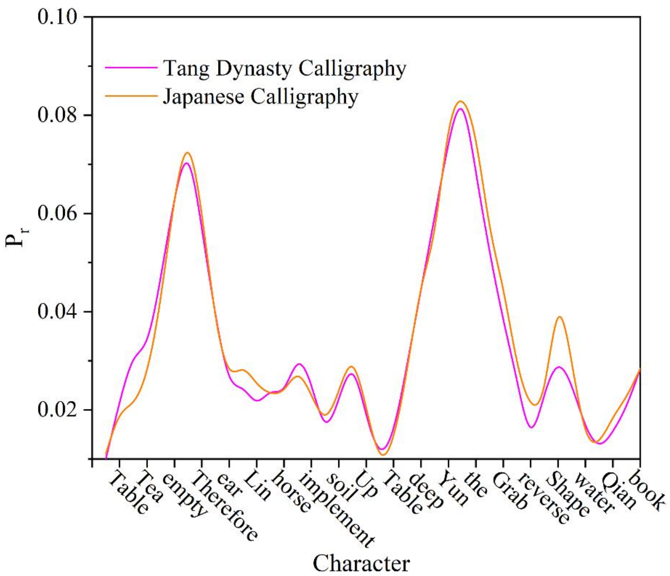

Figure 6. Results of aesthetic choices for different font styles

5.2. The Influence of Tang Dynasty Calligraphy and Structure on Shudo Script Style

Figure 7 shows the P_r mean values of the stylistic choices of the two groups of font samples. It can be seen that there is a big difference between the stylistic vocabulary choices of Tang Dynasty calligraphy and Japanese Shudo, indicating that "replacement of lines" and "unified structure" have a significant effect on the style of Japanese Shudo fonts.

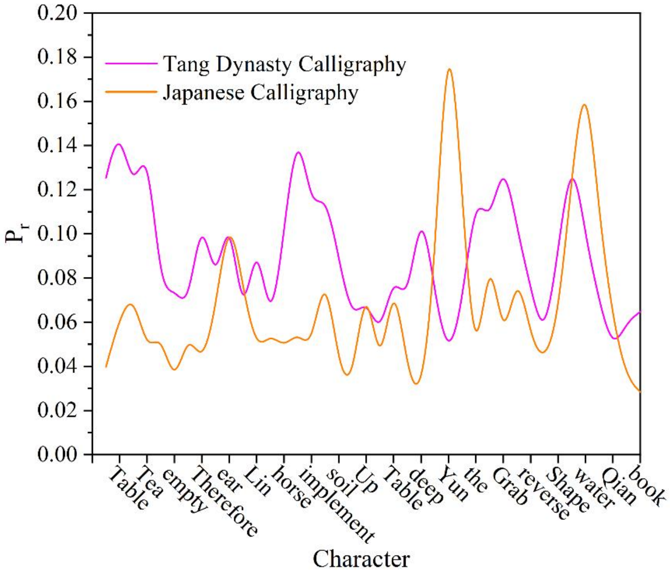

Figure 7. The style selection result of each font model

In order to further compare the effects of "Replacement Lines" and "Unified Structure" on the styles of fonts, Pearson's correlation was used to quantify the styles of different calligraphic styles. The correlation coefficients between the aesthetics and styles of the two groups of font samples are shown in Tables 6 and 7, respectively. The correlation between the aesthetics of each font is 0.6~0.8, which indicates that the aesthetics of the two groups of font samples are relatively close to each other.

Table 6. Font aesthetic relevance

Font	Correlation	
	Tang Dynasty Calligraphy	Japanese Calligraphy
Yan Style	0.641	0.668
Liu Style	0.733	0.746
Ouyang Style	0.682	0.646
Zhao Script	0.631	0.623
Light Gold Script	0.692	0.687
Overall	0.665	0.658

Table 7. Font style correlation

Font	Correlation	
	Tang Dynasty Calligraphy	Japanese Calligraphy
Yan Style	-0.051	0.255
Yan Style	0.079	0.219
Liu Style	-0.085	0.169
Ouyang Style	-0.309	0.143
Zhao Script	0.231	0.009
Light Gold Script	0.076	0.100

6. Conclusion

Calligraphy style recognition has important research and application value in the fields of calligraphy learning and appreciation, calligraphy font retrieval, and digital library. Under the perspective of cross-cultural communication, the study selects Tang Dynasty calligraphy images and Japanese Shudo images as datasets and processes them, uses a model combining stack self-encoder and Softmax to recognize the style of Tang Dynasty calligraphy, and verifies the effectiveness of the style recognition method through experiments. On this basis, the promotion mechanism of Tang Dynasty calligraphy on Japanese Heian Period Shudo style is analyzed. It is experimentally verified that when N=5 and M=16, the average recognition accuracy is the highest, reaching 98.95%, and the recognition accuracy is improved compared with other recognition methods in cursive and running script styles, reaching 99.05% and 98.49%, respectively. At the level of stylistic influence, with the addition of Tang penmanship and structure, the "sense of beauty" of Japanese Shudo scripts has been enhanced, and the sense of style has been enriched, and the “replacement of lines” and “unification of structure” have significant effects on the style of Japanese Shudo scripts. “Replacement of lines” and “unification of structure” had a significant impact on the style of Japanese Shudo scripts.

References

1. Liu, X. (2026). Cross-Cultural Aesthetics and Educational Practice: Exploring the International Dissemination Path of Chinese Calligraphy and Seal Carving in Global Art Education. International Journal of Educational Development, 2(3), 24-45.
2. Liang, L. (2024). Research on the Transmission Path and Cultural Adaptation of Calligraphy Art in Southeast Asia. History, 6(10), 28-32.
3. Teng, J. (2018). The History of Sino-Japanese Cultural Exchange. Routledge.
4. Shyu, S. C. (2021). Seventeenth-century Chinese émigrés and Sino-Japanese cultural exchanges. In The Tokugawa World (pp. 470-484). Routledge.
5. SHIMIZU, Y. (2018). Copying in Japanese Art: Calligraphy, Painting, and. Japanese Art: Critical and Primary Sources, 4, 323.
6. Zeng, W. (2021, January). The influence and communication of Chinese calligraphy in South Korea. In The 6th International Conference on Arts, Design and Contemporary Education (ICADCE 2020) (pp. 720-723). Atlantis Press.
7. Miao, H. (2022). The New Result of Studies in the History of Medieval Calligraphy: A Review of The Manual of Calligraphy by Sun Guoting of the Tang: A Comprehensive Study on the Manuscript and Its Author. Contemporary Social Sciences, 7(3).
8. Carpenter, J. T. (2007). Chinese calligraphic models in Heian Japan: copying practices and stylistic transmission. In The Culture of Copying in Japan (pp. 168-208). Routledge.
9. Wang, Z., Liao, M., & Maekawa, Z. (2016). A study on quantitative evaluation of calligraphy characters. Computer Technology and Application, 2016(7), 103-122.
10. Yu, S. (2025). Changes in Calligraphy through Cultural Interaction: The Cross-Cultural Transmission of" Tang-like" Calligraphy in the Edo Period. Mediterranean Archaeology and Archaeometry, 25(3).
11. Morley, B. A. (2016). Poetry and diplomacy in early Heian Japan: The embassy of Wang Hyoryŏm from Parhae to the Kōnin court. Journal of American Oriental Society, 136(2), 343-369.
12. Ding, L. (2023). The Transmission of Medieval Chinese Paintings to Japan: Paintings on the “Book Road” and Their Reception. In Beyond the Silk and Book Roads (pp. 97-125). Brill.
13. Minguzzi, D. (2024). Mid-Tang Exchange Poetry, the Kingdom of Parhae, and the Reception of Bai Juyi in Early Heian Japan, Part 1: Parhae Envoys and the Yuan-Bai Style of Exchange Poetry. Sungkyun Journal of East Asian Studies, 24(2), 257-284.
14. Kotyk, J. (2020). The medieval Chinese vision of Japan: Buddhist perspectives in the Tang and Song periods. Studies in Chinese Religions, 6(4), 360-385.
15. Keyworth, G. A. (2023). Vowing the Buddhist Canon along the Silk Road (s): A Study of Colophons to Manuscripts from Dunhuang and Japan. In Beyond the Silk and Book Roads (pp. 38-96). Brill.