www.cad-journal.net
Computer-Aided Design AND APPLICATIONS

Interactive Art Style Font Modeling and Implementation Based on Artificial Intelligence

Feng Zhang[1] and Haihong Xie[2]

[1]Department of Visual Communication Design, Zhengzhou Academy of Fine Arts, Zhengzhou, Henan 450000, China, zhangfeng821013@163.com
[2]Department of Visual Communication Design, Zhengzhou Academy of Fine Arts, Zhengzhou, Henan 450000, China, xiehaihong@zzuli.edu.cn

Corresponding author: Haihong Xie, xiehaihong@zzuli.edu.cn

Abstract. Stylized font is to apply some special artistic styles to traditional fonts and present the original traditional rigid text fonts in a certain artistic form. Artificial intelligence can simulate human thinking and decision-making processes, but its decisions are based on a large amount of data and algorithmic calculations, rather than individual subjective judgments and inspirations. Although artificial intelligence can also produce some unique results, its creativity and uncertainty are limited and predictable, and its innovation and flexibility are relatively low compared to the inspiration of human art and design. In this article, an AI-based high-resolution image reconstruction algorithm for artistic fonts is proposed to improve the modeling stage of artistic font computer-aided design (CAD), which not only meets the requirements of font variability, but also strengthens the human-computer interaction (HCI) experience of user participation. The learning method of artistic font style based on this model has more advantages in recognition accuracy, and it is feasible to be applied in interactive artistic font CAD modeling. Dropout method added to the model can effectively prevent over-fitting, strengthen the generalization ability of the network, and effectively complete the task of font style learning. Under the promotion of HCI concept, AI can also be used to automate Type design. This method uses machine learning and computer vision technology to automatically recognize and process various elements in Type design, such as characters, glyphs, font styles, etc. In general, the combination of artificial intelligence and artistic style Type design can provide designers and artists with more creativity and flexibility, and also bring new possibilities and challenges to Type design.

Keywords: Artificial Intelligence; Artistic Fonts; Human-Computer Interaction; Computer Aided Design
DOI: https://doi.org/10.14733/cadaps.2024.S7.193-206

1 INTRODUCTION

As one of the practice bases, the core of interactive art is to create more novel experiences. Stylized font is to apply some special artistic styles to traditional fonts and present the original traditional rigid text fonts in a certain artistic form. With the continuous development of technology, the fusion of hyperspectral and multispectral images has become an important topic. However, existing methods often overlook the variability between images, which can lead to inaccurate fusion results. We propose a new method that combines hyperspectral and multispectral images through coupling tensor decomposition, while considering the variability between images. And better preserve the information of the original image. Borsoi et al. [1] proposed a new method to fuse hyperspectral and multispectral images through coupling tensor decomposition, while considering the variability between images. Preprocess hyperspectral and multispectral images to obtain better fusion results. Perform coupling tensor decomposition on the preprocessed images to obtain their respective decomposition results. Determine the similarity region between two decomposition results by comparing their similarity. Fusion hyperspectral and multispectral images based on similar regions. It conducted experiments using different datasets, including hyperspectral images and multispectral images. By comparing different methods, we found that our fusion and better preserve the information of the original image. These artistic fonts not only give the audience a more pleasant visual experience, but also easily leave a deeper impression on the audience. The essence of interactive art design is the communication between people, works and equipment, and if you want to play the application value of AI and integrate it into interactive art design, you must follow the principle of humanization. Gao et al. [2] It is a complex problem to synthesize artistic glyph images through one-stage few shot learning, which requires the combination of computer graphics and machine learning and other knowledge in many fields. In computer graphics, artistic font image synthesis usually involves font design, rendering and synthesis. These processes require a lot of manual interaction and adjustment to produce high-quality results. However, through machine learning technology, we can automate this process and produce more efficient and high-quality results. In machine learning, few shot learning is a model that can learn complex relationships from a small number of samples. This model can be applied to the synthesis of artistic glyph images, learning the relationships between glyphs through a small amount of sample data, and generating new glyph images. Different styles of fonts can not only express the writing style of the font founder, but also express different artistic conception and aesthetic feeling according to the different use environment. More and more users are increasingly demanding personalized font expression and fast font generation technology and methods. The research on faster and more convenient font generation strategy has been paid more and more attention by the font industry. Traditional methods of making stylized fonts are all done manually by professional designers, which is undoubtedly a very time-consuming and laborious process, so the main starting point of this article is to consider whether CAD methods based on AI can be used to help people design and generate stylized fonts. Research on data acquisition algorithms has become increasingly important. Hu [3] explores how to use image processing and artificial intelligence technology for data collection to improve collection efficiency and accuracy. Image processing technology can be used for preprocessing and enhancing images for better data collection. In the preprocessing stage, image enhancement technology can remove noise and interference in the image, improve the clarity and contrast of the image. In addition, image segmentation technology can divide an image into multiple regions to better extract the required information. By using computer vision technology and deep learning models, objects, people, scenes, etc. in images can be automatically recognized and relevant information can be extracted. For example, in intelligent security systems, abnormal behavior and threats can be automatically identified through image recognition technology, and alerts or other measures can be triggered. Through these techniques, models can be trained to automatically recognize targets in images and extract the required information.

DenseNet and Inception models are commonly used Convolutional neural network models in the field of deep learning. The Inception model, on the other hand, was proposed by the Google team and has the characteristics of multi-scale processing and parallel processing, which can improve the performance of the model. Jalali and Lee [4] can use the integration of DenseNet and Inception models to improve the performance of traditional Asian character recognition in high cursive script with adaptive constraints. Ensemble learning can improve the generalization performance and robustness of models by fusing the results of multiple models. In Ensemble learning, adaptive constraint is an important technology, which can automatically adjust the weight of models according to the prediction results of models, so as to better integrate the results of multiple models. Font stylization not only involves the extraction of detailed features such as texture and color, but also involves the reconstruction of the overall style of the image. In animation design, 3D reality technology and CAD technology complement each other. Jing and Song [5] use 3D reality technology to create realistic 3D spaces and lighting effects, making animated characters more realistic and vivid. CAD technology provides precise shapes and structures, making animated characters more in line with physical laws and design requirements. By combining these two technologies, more exquisite and vivid animation works can be created. The animation tools in CAD software can create animation sequences, add actions and expressions to animated characters, and make them more vivid. CAD software can create precise 2D graphics and 3D models, providing precise shapes and structures for animated characters. The character recognition algorithm based on these feature extraction methods has a high recognition rate for character images with high quality image background or simple deformation, but it has certain limitations for the recognition of character images with serious deformation and complex image background, and the recognition rate is relatively low. Artificial neural network (ANN) is a computational model simulating the human brain nervous system, which is composed of multiple neurons connected with each other, and has strong learning and adaptive capabilities. In automatic crack detection of concrete structures, deep learning neural networks can be used for image analysis and processing. Image processing is one of the key technologies for detecting cracks in concrete structures. Kim et al. [6] can accurately detect the position and shape of cracks by performing preprocessing, segmentation, and feature extraction on images of concrete structures. By using neural networks and image processing techniques, cracks in concrete structures can be automatically detected, greatly improving detection efficiency and accuracy. By analyzing the detected cracks, the type, degree, and cause of the cracks can be determined in order to take corresponding treatment measures.

As a written carrier of language, words can be seen everywhere in our daily life, and with the pursuit of aesthetic feeling and artistic nature, words have also appeared in various styles. In different applications, using different styles of words can set off different effects and bring a better impression. With the growth of image processing technology, people try to apply image style transfer technology to font style design. This method can quickly transfer the style to any character by designing a small number of style characters, which can greatly reduce the workload of text style designers. The existing font stylization methods based on feature extraction and texture synthesis mainly include manual stylization method, statistical stylization method and deep learning method. Traditional feature extraction methods use a single numerical value or a single feature vector to represent the target, and the extracted feature information can not meet the requirements of deformed font image recognition. In this article, a high-resolution image reconstruction algorithm of artistic fonts combined with AI is proposed to provide technical support for interactive artistic font CAD modeling. Compared with traditional methods, this article has made the following innovations:

⊖ By studying the similarities and differences between AI and font design, this article explores the combination of AI and interactive artistic style font design, and analyzes the new design mode brought by intelligent design.

⊜ Expand the multi-dimensional expression of artistic fonts in graphic design through AI and CAD, from the behavior judgment in the program to the feedback of the linkage between behavior and vision, and then deeply explore the variable vision of algorithmic fonts, so as to provide an

integrated rational design method for graphic design of fonts, and strengthen the interactive experience of users' participation while meeting the requirements of font variability.

(3) In the network model, a jump connection module integrating attention mechanism is added, which is used to project the features in the encoder to the decoder, so as to avoid generating structural errors by reducing the information loss of the decoder.

This article introduces the interactive artistic font CAD modeling method based on AI algorithm. Then the font style learning and high-resolution reconstruction performance of the algorithm are verified on public data sets. Finally, the contribution of this study to the study of art design HCI is summarized and the future research direction is pointed out.

2 RELATED WORK

Kolařík et al. [7] achieved automatic and accurate segmentation results using high-resolution 3D image data. The network adopts a U-Net structure, consisting of two parts: encoder and decoder. The encoder extracts the features of the image through multiple down-sampling operations, and the decoder restores the feature map to the same size as the input image through up-sampling operations. Dense connections have been added to the network, connecting the output of each decoder with the output of the corresponding encoder layer, enabling the network to better capture the detailed information of the image. A skip connection has been added to the network, which directly connects the feature map of the encoder layer to the corresponding position of the decoder layer, enabling the network to better integrate feature information of different scales. Through the above structural optimization, the network has achieved excellent performance in brain and spine segmentation tasks, and can automatically and accurately segment the contours of the brain and spine, improving the accuracy and efficiency of segmentation. Normalization of images is one of the important steps to improve recognition accuracy in the OCR process. Konovalenko et al. [8] adjust the image to a uniform size, which can be achieved through interpolation algorithms (such as Bilinear interpolation). The image is converted from color to Grayscale, which can be realized by various gray conversion algorithms (such as weighted average, maximum and minimum). Adjust the pixel value range of Grayscale to a uniform range, usually 0 to 1 or -1 to 1. This can be achieved through linear transformation. The Grayscale is converted to a Binary image, so that the text area becomes black and the non text area becomes white. This can be achieved through various binary algorithms such as fixed thresholds, adaptive thresholds, etc. Liu and Yang [9] explored the creative centered Computer-aided design teaching mode of contemporary art. Computer-aided design is widely used in contemporary art. It can not only improve the efficiency of artistic creation, but also add more expressiveness and imagination to artistic works. However, how to combine Computer-aided design with artistic creativity in teaching is a challenging problem. This model regards Computer-aided design as a tool to realize the idea of artistic creativity. In the teaching, the basic skills and methods of Computer-aided design are demonstrated through cases and examples, so that students can understand the basic operation and characteristics of Computer-aided design. Then, through practical projects, students are allowed to apply their learned skills to achieve their creative ideas, cultivating their innovative thinking and practical abilities.

Pulse coupled neural network (SNN) is a simulated neural network model based on biological neural systems, which has wide applications in the field of image processing. SNN can be used for image recognition tasks, such as handwritten digit recognition, facial recognition, etc. By converting the image into a pixel matrix and inputting it into SNN, connections between neurons can be trained to recognize patterns in the image. Liu et al. [10] divided the image into multiple different regions. By using the pulse synchronization phenomenon of SNN, the pixels in the image can be divided into different clusters, thereby achieving image segmentation. SNN can be used for image restoration tasks, such as denoising, deblurring, etc. By training SNN to restore the original pixel values of an image, the quality of the image can be improved. SNN can also be used for image compression tasks, by representing images as activation sequences of neurons, image data

can be compressed and efficient data transmission and storage can be achieved. Building footprint refers to the projected area of a building on the surface of the Earth, and is an important parameter in urban planning, architectural design, and environmental science research. Liu et al. [11] evaluated the trained model using test data, and evaluated the performance of the model. Use the trained model to extract building footprints from new high-resolution satellite images and obtain data such as the projected area of the building. When using ResNet for building footprint extraction, the image can be used as input and the projected area of the building can be used as output. By training neural networks, it is possible to learn to extract features related to the projected area of buildings from images, thereby accurately extracting building footprints. This method learns the detailed information of high-resolution images from low resolution satellite images by constructing a multi-scale residual deep neural network model, thereby achieving the goal of super-resolution. Lu et al. [12] utilized multi-scale residual networks to extract feature information at different scales. And fuse this feature information into a high-resolution image through residual connections. Specifically, the multi-scale residual deep neural network model includes multiple convolutional layers and residual blocks, each of which contains multiple convolutional layers and skip connections, used to extract feature information at different scales. When training the model, we use a large number as training data to optimize model parameters by minimizing the difference between high-resolution images and model outputs. In this way, achieving super-resolution of satellite images. Peng et al. [13] collected two sets of high-resolution satellite images, one as a training set and the other as a testing set. At the same time, in order to enable the model to better learn the change information of images, the training set images can be preprocessed with the test set images, such as registration, alignment, and other operations. It adopts an improved UNet++model to transform the model for end-to-end change detection.

Specifically, feature maps can be output in the middle layer of the UNet++model and transmitted to the classification module for detecting changes in the image. At the same time, techniques such as residual connections or skip connections can be used to better preserve the detailed information of the image. The training set is used to train the model, and Loss function such as Cross entropy are used to minimize the Loss function of the classification module, so as to optimize the model parameters. Evaluate the performance of the trained model using a test set, calculate the accuracy, recall rate, and other indicators of the model to evaluate its performance. Apply the trained model to the actual scene of end-to-end change detection in high-resolution satellite images. Specifically, the model can be applied to satellite images of the same region at different times, identifying the parts that have undergone changes, and outputting the location and type of changes. It should be noted that when training the model, it is necessary to select appropriate super parameters such as Learning rate and optimizer, as well as appropriate structural parameters such as network structure and layers, so as to obtain better model performance. In addition, in practical applications, it is necessary to deploy and optimize the model to better meet practical needs. Yang et al. [14] The artistic text benchmark is a standard used to evaluate the performance of text recognition algorithms. It presents the text image to the recognition algorithm and evaluates the accuracy of the algorithm's output. This benchmark typically uses a large number of different text images, including various fonts, font sizes, lighting conditions, and backgrounds. Through this test, the performance of different recognition algorithms can be compared and which algorithm performs best in specific application scenarios can be determined. Through pattern analysis, machine learning algorithms can identify rules and patterns in data, and use this information to make decisions or predictions. Machine intelligence trading refers to the use of machine learning algorithms for financial transactions. This technology uses Analysis of algorithms to analyze Market data and make investment decisions. Machine intelligence trading can improve trading returns by analyzing a large amount of data and learning market patterns. IEEE pattern analysis and machine intelligence trading are two different fields, but both use machine learning algorithms to process data and make decisions. Pattern analysis focuses on discovering patterns from data and classifying them, while machine intelligence trading focuses on using algorithms for financial transactions. Computer graphics and image aided design are widely used in art design teaching. They can not only provide efficient and accurate drawing

and design tools, but also help students and designers better express their creativity and ideas. Computer graphics and image aided design can provide a variety of tools and skills for art design teaching. Zhang and Rui [15] can use CAD software for accurate drawing and modeling, image editing software for photo editing and composition, and 3D software for creating and rendering 3D models and scenes. These tools help students and designers create more quickly and accurately, while also providing more visual effects and ways of expression. Computer graphics and image aided design provide more creative expressions for art design. These technologies not only stimulate the creativity of students and designers, but also enable them to better express their ideas and design concepts.

3 INTERACTIVE ARTISTIC FONT CAD MODELING

3.1 Artistic Font Style Learning

The pixels in the image are not isolated, and each pixel is under a certain shape background. Therefore, when extracting features, the shape context also extracts the background of the pixel as the feature of the pixel and digitizes it. Although neural network can approximate any nonlinear function, sometimes the function representing the internal relationship of sample data is too complex, so most shallow structure algorithms used for regression and classification at present, this type of network is not very adaptable to complex classification problems and its model result stability effect is not very good. In the field of signal processing, it is usually processed by orthogonal linear transformation. However, there are many limitations and shortcomings in signal transformation decomposition using fixed orthogonal basis functions. It has a good application effect in the fields of pattern recognition, image reconstruction and image compression, which improves the compression efficiency, reduces the cost of signal processing and improves the efficiency. In AI-based HR image reconstruction, sparse representation of images has become a very important part of HR reconstruction. As a classic and effective statistical mathematical model, sparse coding technology has been successfully applied to image processing and computer vision, such as image compression, image denoising and image HR reconstruction, and achieved superior results. Sparse coding technology is to represent a natural image or signal as a linear combination of a few atoms selected from an overcomplete dictionary, as shown in Figure 1. Its image signal y can be approximately represented by the linear combination of three atoms in the overcomplete dictionary D. This stage of solving the optimization problem is called "sparse coding".

Convolutional neural network (CNN), as a powerful unsupervised learning model, applies normalized structure to both generator and discriminator for the first time, which enables it to extract features more comprehensively. However, it does not provide an image output template or semantic tags for segmentation, which makes the model only complete image reconstruction and generate chaotic pictures according to the extracted features. In practice, the following three constraints are usually used to solve the sparse representation problem:

$$\min_{\alpha} \|D\alpha - X\|_2^2 \quad s.t \ \|\alpha\|_0 \le L \tag{1}$$

$$\min_{\alpha} \|\alpha\|_0^2 \quad s.t \ \|D\alpha - X\|_2^2 \le \xi^2 \tag{2}$$

$$\min_{\alpha} \left\{ \|D\alpha - X\|_2^2 + \gamma \|\alpha\|_0^2 \right\} \tag{3}$$

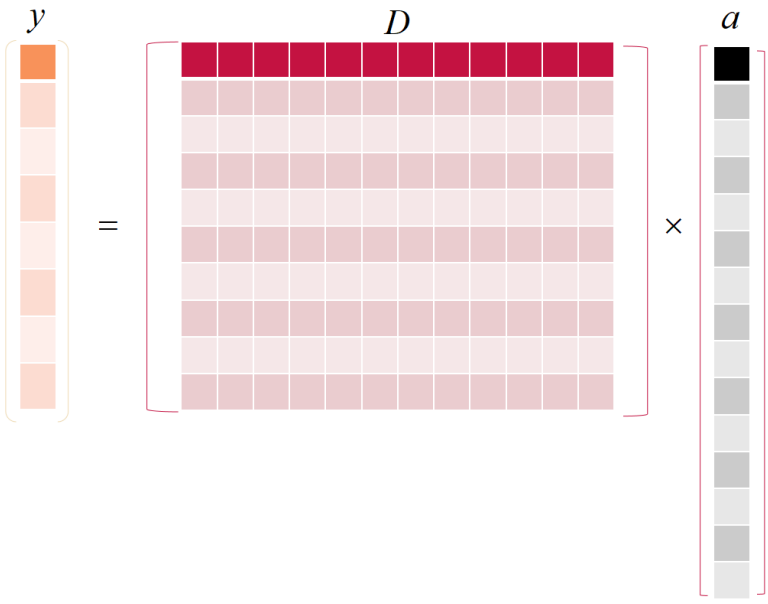

Figure 1: Signal sparse representation model.

Among them, the constraint condition of the first case is sparsity L; In the second case, the constraint condition is to solve the residual of the signal; In the third case, sparsity L and signal residual are considered comprehensively.

In the standardization operation, the example standardization rather than batch standardization is adopted because of the particularity of font style migration. Because of the different strokes of each font, when generating style fonts, the generation effect of fonts mainly depends on the example of a certain font, not the whole sample. Batch standardization based on the whole sample is obviously not suitable for font style migration. Stylistic pattern synthesis refers to the combination of many pattern materials through certain constraint rules to synthesize a higher form of pattern or other results. Due to the lack of clear constraints, the special font images generated by a single stylized model have the defect of blurred images. Therefore, how to provide an efficient font stylization method to realize the diversification and clarity of special effects of generated fonts is an urgent problem in the field of font stylization. The main principle of text matching image interaction art design is to automatically generate design works through design suggestions. This design method utilizes artificial intelligence technology to automatically generate corresponding image works based on user input text information, achieving text to image conversion. This design approach can help designers and users better understand and express their ideas, improve design efficiency, and also bring different sensory and psychological experiences. For example, in a smart home system, users can design the layout and decoration of a room by inputting text information. The system will automatically generate corresponding 3D images, allowing users to have a more intuitive understanding of the design effect. This not only improves design efficiency, but also allows users to more conveniently participate in the design, achieving the integration of human-computer interaction and interpersonal interaction. And AI can judge excellent design works according to different needs by sorting out massive data and comparing the works of excellent designers. Figure 2 shows the algorithm flow of artistic font style learning.

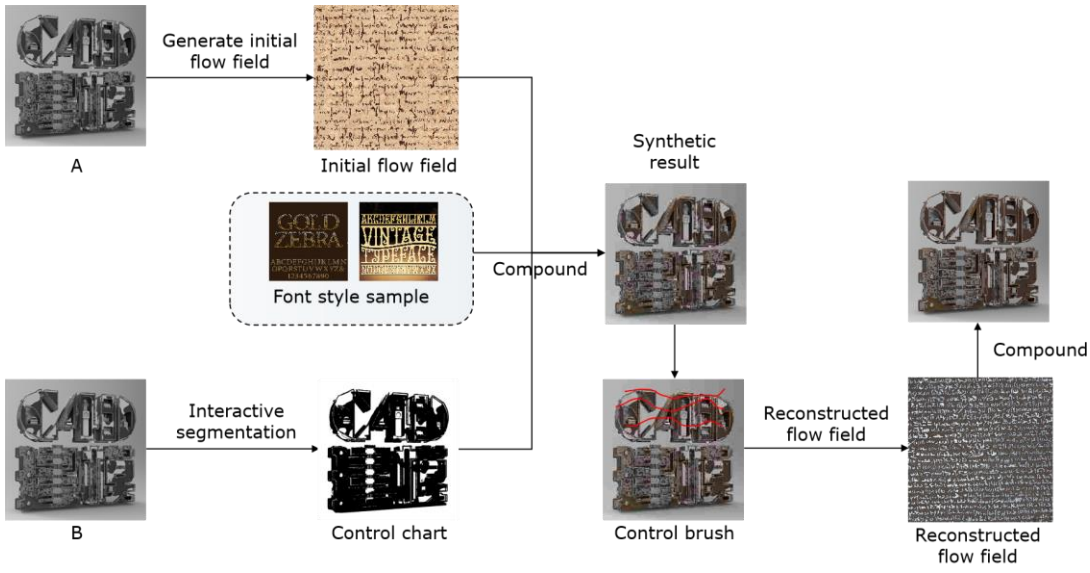

Figure 2: Algorithm flow of artistic font style learning.

In the stage of implementation, the generator is mainly composed of three parts, namely, feature encoder, feature migrator and feature decoder. In the feature encoder, the main function is to extract the style features of style fonts and the content features of fonts to be migrated, and it is composed of convolution layer, instance standardization layer and Relu activation function layer.

$$x_k = f_k\left(x_{k-1}, u_{k-1}, v_{k-1}\right) \tag{4}$$

Observation equation:

$$y_k = h_k\left(x_k, w_k\right) \tag{5}$$

The existing pattern synthesis technology rarely takes into account the semantic information of patterns, but the work in this article is to divide characters into different sets of strokes. In the stage of stroke replacement, not only the similarity in shape between strokes and patterns is emphasized, but also the semantic correlation between patterns and characters is concerned. It is difficult to depict rich textures with a single texture generator. Therefore, it is needed to establish a hierarchical calligraphy texture structure according to the average gray value of texture, because the average gray value of single pixel sequence is very effective to quantify the hierarchy of texture. Then, for each level of calligraphy texture, a texture generator is constructed to describe the calligraphy texture characteristics of this level. AI and interactive art design are inseparable, and they are mutually influenced and promoted. Sorting out the relationship between the two can provide a theoretical basis for the writing of the next paper and also provide some guiding opinions for the future growth of HCI.

3.2 High Resolution Reconstruction of Artistic Font Image

In the stage of image formation, the quality of the obtained image is reduced due to hardware equipment and external objective environment. The influence of hardware equipment includes aberration and distortion caused by the process size of optical imaging system and sensor. The influence of external environment includes the imaging system and the shooting scene, poor focusing and system noise. These factors all affect the image quality, making the image blurred, distorted and the resolution lower. Because the input image itself has a certain fractal structure, the input image itself can be directly used as an example picture in the reconstruction process.

When the image itself can't provide more details reconstruction information, users can also be allowed to provide additional sample images for details reconstruction.

The design principle of a good pattern font needs to meet the following two requirements. One is that the pattern text cannot affect the readability of the original text, and the other is that the pattern used to synthesize the text should have a certain semantic correlation with the text. The three-layer CNN network avoids the problem that the parameters are difficult to transfer because of the deep network, and can converge quickly. The subjects of interaction are often audience and works, audience and creators, and audience and audience. This form of interaction can be the interaction between the audience and the work, such as voice interaction, or the exchange of ideas between the creator and the audience. Between adjacent corners, the system can delete redundant input data by uniform sampling. The down-sampled input data not only saves storage space, but also improves the efficiency of calculating calligraphy shape by this method.

In some scales (resolution), the features in the image are not easy to be seen, but they are easy to be detected in another scale (resolution). Multi-scale technology of image is a special image analysis method, which is also called image multi-resolution technology. Using this technology, images can be expressed in multi-scale and processed separately based on different scales. Using multi-scale technology to express and analyze images is conducive to extracting image features more effectively. Figure 3 shows the stage of high-resolution detail reconstruction of an artistic style font.

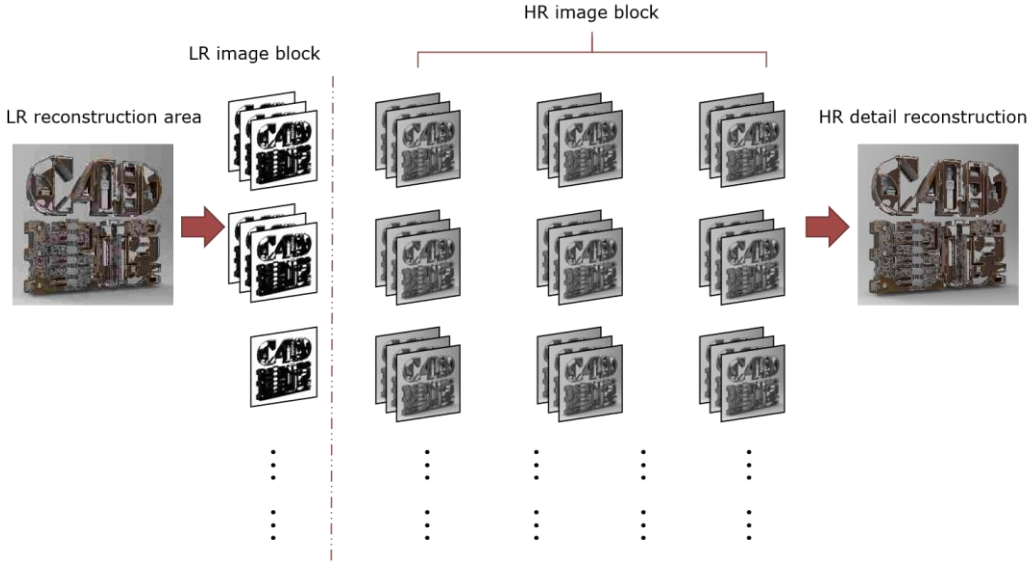

Figure 3: Reconstruction of font details.

The spatial relationship characteristics of an image mainly describe the spatial relative position or relative direction relationship between multiple targets in the image, which is often used to describe the content of the image. In practical applications, low-pass filtering processing and image sub-sampling operation are often carried out at the same time to construct multi-scale expression of images. According to the image acquisition model established in this article, the reconstruction stage of artistic font HR can be described as:

$$f(x) \approx f_k(x) = \sum_y [g_k(y) - n_k(y)] h_k^{BP}\left(m_k'(x) - y\right) \tag{6}$$

g_k represents the LR image of the k th frame, and $g_k(y)$ represents the gray value of the LR image at the position $y=[s,t]^T$; f represents the ideal HR image, and $f(x)$ represents the gray value of the ideal HR image at the position $x=[u,v]^T$; f_k represents the HR image obtained from g_k solution; n_k is the superimposed noise defined on g_k ; h_k^{BP} can be considered as the approximation of the inverse stage of the point spread function; $m_k^{'}$ represents the pixel mapping relationship from HR image f to LR image g_k .

After sparse coding, the dictionary of LR image blocks is expressed as follows:

$$D_h = \arg\min_{\{D^h,A\}} \left\| X^h - D^h A \right\|_2^2 + \lambda \|A\|_1 \tag{7}$$

$$D_l = \arg\min_{\{D^l,A\}} \left\| Y^l - D^l A \right\|_2^2 + \lambda \|A\|_1 \tag{8}$$

Where $X^h = \{x_1, x_2, \cdots, x_n\}$ is HR image block set, $Y^l = \{y_1, y_2, \cdots, y_n\}$ is LR image block set, and A is sparse matrix.

The edge reconstructed image I_e and the detail reconstructed image I_d can be fused by using the perceptual image W_d , so as to obtain the final super-resolution image I_h . The fusion process can be represented by the following formula:

$$I_h(x) = \begin{cases} W_d(x) \cdot I_d(x) + [1 - W_d(x)] \cdot I_e(x), & x \in \bigcup_n R_n \\ I_e(x), & x \notin \bigcup_n R_n \end{cases} \tag{9}$$

Context awareness diagram is established in the coordinate system of LR image, so when applying this model, it is needed to first convert the context awareness weight diagram into the coordinate system of HR image. In the dictionary training stage, the initial value of clustering is determined according to the distance distribution of samples, which makes dictionary training more stable and accurate. In the reconstruction stage, the non-local similarity constraint is added to the image, so that the algorithm can converge to the optimal value faster and output the reconstructed image of HR.

4 RESULT ANALYSIS AND DISCUSSION

In order to highlight the performance difference between the improved algorithm and the traditional algorithm, this section takes a representative picture as an example to process the image, and compares the processing renderings of each link. The selected pictures not only have complex backgrounds, but also have different text sizes. Under such complex conditions, the performance differences of the algorithms can be tested. The function of the feature decoder is to reconstruct the style font, which consists of deconvolution layer, instance standardization, activation function and the last convolution layer. The deconvolution layer deconvolves the feature map after feature migration to restore the image details, and then outputs the target font of 64*64*1 through the last convolution operation. The test results of artistic font recognition using traditional CNN model are shown in Figure 4. The test results of artistic font style learning using this model are shown in Figure 5.

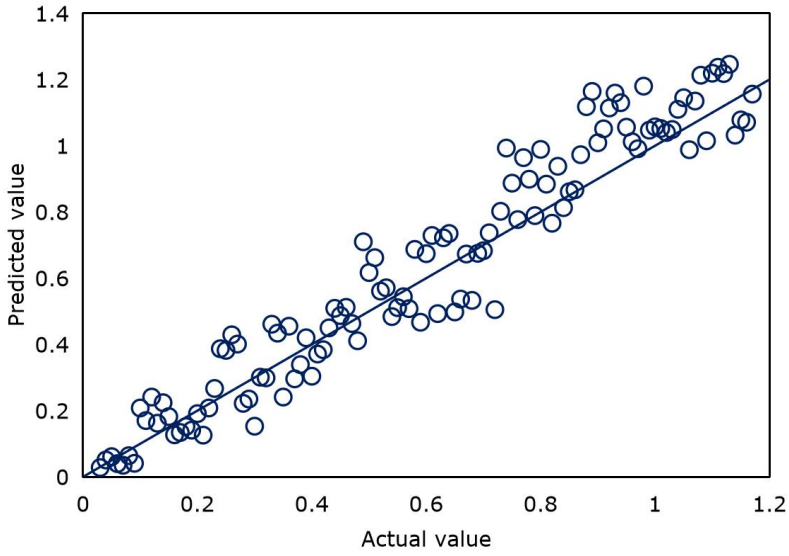

Figure 4: Art font style learning of CNN model.

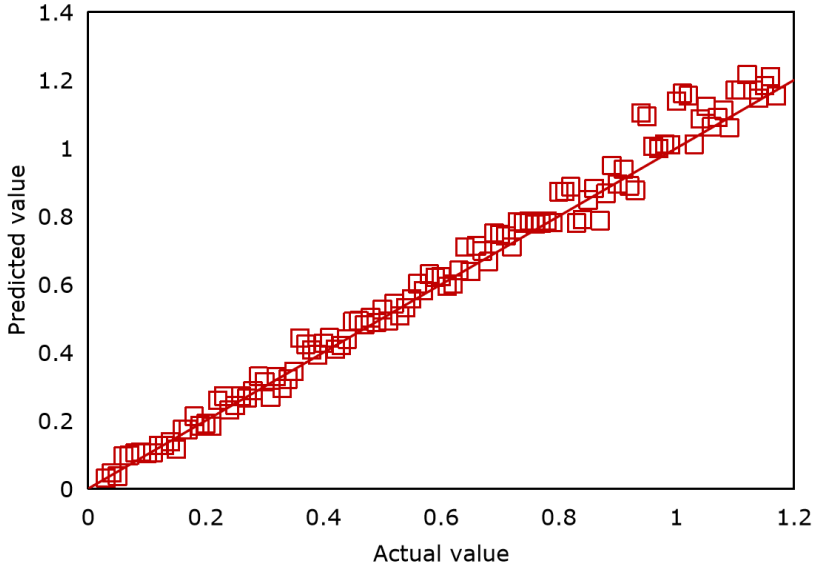

Figure 5: Art font style learning of the model in this article.

It can be analyzed that the artistic font style learning method based on this model has more advantages in recognition accuracy and is feasible to be applied in interactive artistic font CAD modeling. Considering that the reference font and the generated font should have similar structures to represent the same content, this article directly connects the lower layer with rich details in the content encoder to the corresponding decoder layer. Interlayer jump link is added between the content encoder and the result decoder, so that the bottom information of the font image can be preserved as much as possible, and the formation and structure information of the font image can be avoided from being lost in the down-sampling process.

Based on interactive graphic characters, visual forms can be randomly constructed according to participants' behavior data. This randomness can be continuously accumulated and changed with the extension of time, and dynamic expression breaks the fixed cycle law preset by designers, that is, it gives the computer autonomy within a limited range, and creates a complete dynamic reading process through interactive triggering to visual generation. When describing a large image, we can calculate the maximum value or average value of the features of each region on the image for aggregation statistics, thus effectively reducing the dimension of the image. Before fitting measures are added, the change curve of recognition accuracy is shown in Figure 6.

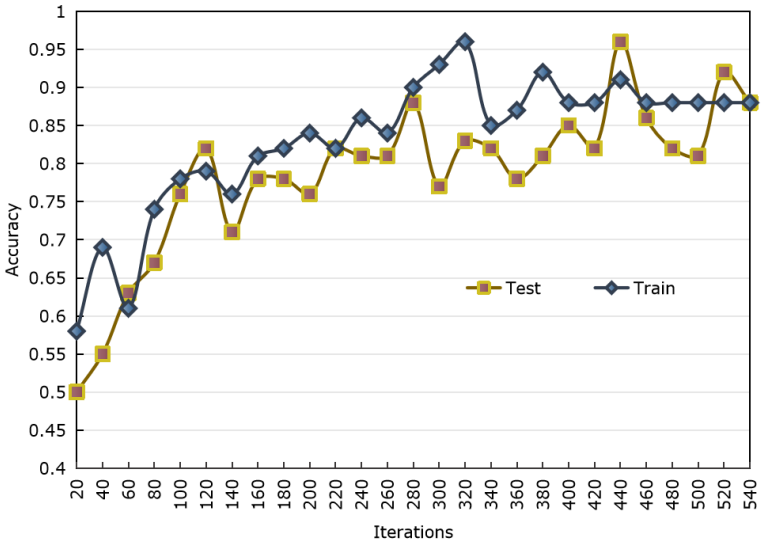

Figure 6: Change in accuracy before adding overfitting measures.

Sample-by-sample mean subtraction is mainly used in those stable data sets, that is, the statistical properties of those data are the same among each dimension. For example, in natural pictures, this can reduce the influence of brightness on the data, because we seldom use the information of brightness. It can be observed that the expressive force of the high-resolution reconstruction model of font image in this article is worse than that in the same training set, and there is an over-fitting phenomenon.

In the training stage of general neural network, a difficult problem is how to choose an appropriate weight to assign the whole network for initialization training. If the initialization weight is not ideal, it will usually lead to the slow convergence of the whole neural network and the poor effect of the resulting model. Dropout method is adopted to prevent over-fitting, and the accuracy transformation after processing is shown in Figure 7.

According to the comparison between the curve of accuracy change in the figure and that before the over-fitting measures are added, it can be seen that through repeated iterative training, the Dropout method added in this article can effectively prevent over-fitting, strengthen the network generalization ability and effectively complete the task of font style learning. The network effectively extracts the font skeleton information by constructing a font skeleton synthesis network composed of an encoder composed of expansive structure and an efficient channel attention decoder, and proposes a style correction network based on multi-scale feature fusion of attention. By combining multi-scale features with attention mechanism, the weight information of different scales is obtained, and the network generation effect is effectively corrected. Finally, the self-built data set proves that this method can generate high-quality calligraphy images only by training a very small number of data sets, and has good generating effect.

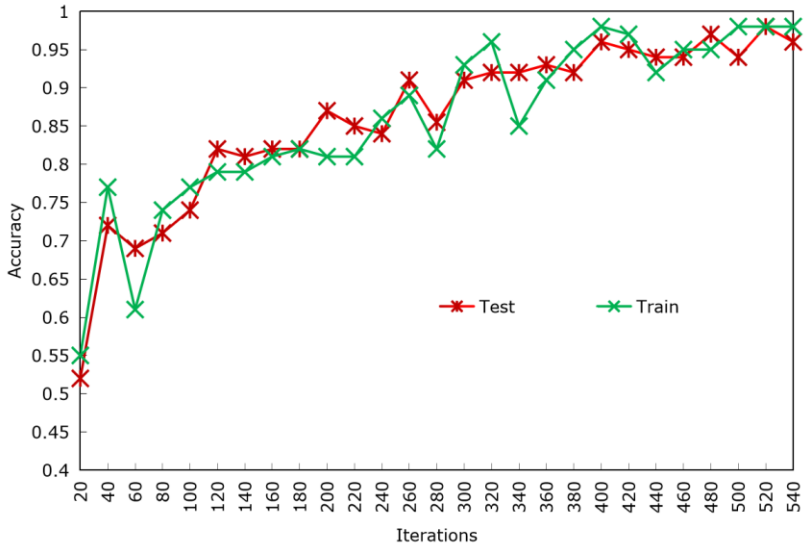

Figure 7: Change in accuracy after adding overfitting measures.

5 CONCLUSIONS

The growth of CAD and AI poses a subversive challenge to the existing rules and procedures of character design. Learning to master the operating principle of computer system can help artistic imagination to really become active between various technologies and media. Traditional feature extraction methods use a single numerical value or a single feature vector to represent the target, and the extracted feature information can not meet the requirements of deformed font image recognition. Artificial intelligence plays an important role in interactive art design, bringing more creativity and interactivity to designers and audiences, as well as a more unique and shocking visual and experiential experience for art works. Artificial intelligence has also reconstructed the visual thinking of designers, allowing them to more freely unleash their creativity and imagination. In this article, a high-resolution image reconstruction algorithm of artistic fonts combined with AI is proposed to provide technical support for interactive artistic font CAD modeling. The results show that the learning method of artistic font style based on this model has more advantages in recognition accuracy, and it is feasible to be applied in interactive artistic font CAD modeling. The Dropout method added to the model can effectively prevent over-fitting, strengthen the generalization ability of the network, and effectively complete the task of font style learning. In the future, by increasing the amount of data and optimizing the model code, we can explore the internal relationship between the types of sample sets and the depth model, and at the same time establish a more perfect deep learning model.

Feng Zhang, https://orcid.org/0009-0007-7313-3882
Haihong Xie, https://orcid.org/0009-0006-0966-1018

REFERENCES

[1] Borsoi, R.-A.; Prévost, C.; Usevich, K.; Brie, D.; Bermudez, J.-C.; Richard, C.: Coupled tensor decomposition for hyperspectral and multispectral image fusion with inter-image variability, IEEE Journal of Selected Topics in Signal Processing, 15(3), 2021, 702-717. https://doi.org/10.1109/JSTSP.2021.3054338

[2] Gao, Y.; Guo, Y.; Lian, Z.; Tang, Y.; Xiao, J.: Artistic glyph image synthesis via one-stage few-shot learning, ACM Transactions on Graphics (TOG), 38(6), 2019, 1-12. https://doi.org/10.1145/3355089.3356574
[3] Hu, S.: Research on data acquisition algorithms based on image processing and artificial intelligence, International Journal of Pattern Recognition and Artificial Intelligence, 34(06), 2020, 2054016. https://doi.org/10.1142/S0218001420540166
[4] Jalali, A.; Lee, M.: High cursive traditional Asian character recognition using integrated adaptive constraints in ensemble of DenseNet and Inception models, Pattern Recognition Letters, 131(11), 2020, 172-177. https://doi.org/10.1016/j.patrec.2020.01.013
[5] Jing, Y.; Song, Y.: Application of 3D reality technology combined with cad in animation modeling design, Computer-Aided Design and Applications, 18(S3), 2020, 164-175. https://doi.org/10.14733/cadaps.2021.S3.164-175
[6] Kim, J.-J.; Kim, A.-R.; Lee, S.-W.: Artificial neural network-based automated crack detection and analysis for the inspection of concrete structures, Applied Sciences, 10(22), 2020, 8105. https://doi.org/10.3390/app10228105
[7] Kolařík, M.; Burget, R.; Uher, V.; Říha, K.; Dutta, M.-K.: Optimized high resolution 3d dense-u-net network for brain and spine segmentation, Applied Sciences, 9(3), 2019, 404. https://doi.org/10.3390/app9030404
[8] Konovalenko, I.-A.; Kokhan, V.-V.; Nikolaev, D.-P.: Optimal affine image normalization approach for optical character recognition, Computer Optics, 45(1), 2021, 90-100. https://doi.org/10.18287/2412-6179-CO-759
[9] Liu, F.; Yang, K.: Exploration on the teaching mode of contemporary art computer aided design centered on creativity, Computer-Aided Design and Applications, 19(S1), 2021, 105-116. https://doi.org/10.14733/cadaps.2022.S1.105-116
[10] Liu, H.; Liu, M.; Li, D.; Zheng, W.; Yin, L.; Wang, R.: Recent advances in pulse-coupled neural networks with applications in image processing, Electronics, 11(20), 2022, 3264. https://doi.org/10.3390/electronics11203264
[11] Liu, P.; Liu, X.; Liu, M.; Shi, Q.; Yang, J.; Xu, X.; Zhang, Y.: Building footprint extraction from high-resolution images via spatial residual inception convolutional neural network, Remote Sensing, 11(7), 2019, 830. https://doi.org/10.3390/rs11070830
[12] Lu, T.; Wang, J.; Zhang, Y.; Wang, Z.; Jiang, J.: Satellite image super-resolution via multi-scale residual deep neural network, Remote Sensing, 11(13), 2019, 1588. https://doi.org/10.3390/rs11131588
[13] Peng, D.; Zhang, Y.; Guan, H.: End-to-end change detection for high resolution satellite images using improved UNet++, Remote Sensing, 11(11), 2019, 1382. https://doi.org/10.3390/rs11111382
[14] Yang, S.; Wang, W.; Liu, J.: Te141k: artistic text benchmark for text effect transfer, IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(10), 2020, 3709-3723. https://doi.org/10.1109/TPAMI.2020.2983697
[15] Zhang, B.; Rui, Z.: Application analysis of computer graphics and image aided design in art design teaching, Computer-Aided Design and Applications, 18(S4), 2021, 13-24. https://doi.org/10.14733/cadaps.2021.S4.13-24