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

Intelligent Recognition Algorithm for Calligraphy Fonts Based on Texture Mapping

Mingke Liu[1] and Yue Zhang[2]

[1]Public Arts Education Centre, Xinyang Vocational and Technical College, Xinyang, Henan 464000, China, Liumingke@xyvtc.edu.cn
[2]Public Arts Education Centre, Xinyang Vocational and Technical College, Xinyang, Henan 464000, China, Zhangyue20230216@163.com

Corresponding author: Mingke Liu, Liumingke@xyvtc.edu.cn

Abstract. From the visual sense, the visual feeling of each font is different, so the different writing methods of each font can be expressed by a texture feature. Therefore, the method of texture mapping can be applied to the study of calligraphy font recognition in this article. Because Convolutional Neural Network (CNN) can extract the deep features of characters in the process of font recognition, reduce the amount of calculation, and effectively solve the unique characteristics of calligraphy fonts, this article can effectively recognize characters by CNN. In this article, the intelligent recognition algorithm of calligraphy font based on texture mapping and CAD is studied, and the rapid recognition model of calligraphy font features is constructed by CNN, and the extraction method of this feature parameter is improved. The experimental results based on MNIST data set show that the recognition accuracy of the improved calligraphy font recognition method based on deep learning (DL) proposed in this article reaches 98.6% in the test set, which effectively improves the recognition accuracy of the original method. Applying CNN model based on texture mapping and CAD to font recognition can not only solve the problem of fast recognition of calligraphy fonts, but also broaden the application field of neural network.

Keywords: Texture Mapping; Artificial Intelligence; CAD; Font Recognition
DOI: https://doi.org/10.14733/cadaps.2024.S3.197-210

1 INTRODUCTION

Character is the main tool for human communication. With the rapid growth of AI, it has become a very important research field to use computers to process and identify character information. Font recognition belongs to the category of character recognition, and it is also a difficult problem in the field of character recognition, especially calligraphy font recognition, which is a challenging theoretical research topic in character recognition. Character recognition based on gesture motion has significant advantages over traditional language character recognition. Alam et al. [1]

analyzed the motion Gesture recognition recognition of language writing characters. It analyzed the normalization level of the character dataset of the convolutional network through model testing and validation by collecting databases. When using deep neural networks and depth sensors for trajectory-based air writing recognition, it is necessary to pay attention to the accuracy of data preprocessing and feature extraction, as well as selecting appropriate deep learning models and parameter settings. At the same time, it is also necessary to standardize and standardize the input air writing trajectory data to ensure the stability and accuracy of the model. In addition, other technologies can be combined to improve the accuracy and stability of air writing recognition, such as using sensor fusion technology to fuse multiple sensor data together, or using data augmentation technology to expand and enhance the data. These technologies can improve the robustness and generalization ability of the model, thus being better applied in practical application scenarios. From the visual sense, the visual experience of each font is different, therefore, different writing methods of each font and different text images can be represented by a texture feature. Cursive character recognition is a CAD computer aided vision system problem. Currently, many algorithms for visual computer simulations still have certain problems in designing character content. With the construction of the Mixture model of English handwritten fonts, Albattah and Albahli [2] analyzed the accuracy of the training set of the hybrid deep learning model of machine learning. It is necessary to use a CNN model for feature extraction. Different CNN architectures can extract different features. For example, independent CNN can extract global features, while hybrid CNN can extract local and global features. After feature extraction, different features can be fused. In intelligent Handwriting recognition using different independent and hybrid CNN architectures, appropriate CNN architectures and parameter settings need to be selected according to specific application scenarios. Therefore, the method of texture mapping can be applied to the study of calligraphy font recognition in this article. It is an arduous task to generate the corresponding visual image from the object expressed by graphics, which requires very complicated operations. For texture images, especially 2D matrix bitmaps, it directly corresponds to the raster device and can be displayed directly on the raster display with a small amount of operations or even without calculation, which can save a lot of time in operation. Texture mapping is the process of establishing the corresponding relationship between the 3D object surface and the 2D image space pixel coordinates. Alemayoh et al. [3] conducted a survey on the specific characters of numerical characters in the range of data collected from images. The prototype of the learning method for the trained model deep network was implemented using sensor motion data. Using IMU and force sensors to capture handwritten motion data for deep learning-based character recognition. Its challenging task is that the trajectory of handwritten motion and force sensor data contain multiple features. Use the trained model to recognize and classify new handwritten motion data. For each input handwritten motion data, the model can output the corresponding character category or sequence. For handwritten text recognition, data preprocessing is very important. First, the data needs to be cleaned and pretreated to remove noise and Outlier, and the data needs to be standardized to make the data meet the requirements of the model. When designing a compact deep learning model, it is necessary to consider the accuracy and efficiency of the model. Convolutional neural network (CNN) or Recurrent neural network (RNN) is usually used as the model infrastructure. When designing the model, it is necessary to select the appropriate network structure, layers, Activation function and other parameters according to the actual situation. Annanurov and Noor [4] conducted a computer simulation assisted convolutional model performance link analysis for the temporary digitization of archives. It constructs an offline recognition model under hardware performance. In model training, a large amount of training data is used, and Backpropagation is used to optimize the weight and bias of the model. During the training process, overfitting was avoided by using regularization and increasing the dataset to avoid overfitting.

Traditional research methods mainly study characters through the mode of "preprocessing+feature extraction+classifier", but there are still many difficulties and shortcomings in the research of character recognition because of its various types, confusing characters and complex structure changes. MNIST and EMNIST are two commonly used datasets for handwritten character recognition, both of which provide extensive data support for the recognition and research of handwritten characters. Baldominos et al. [5] conducted a visual analysis application of the innovative contribution dataset in the writing of handwritten characters. It conducts data augmentation analysis of neural networks through deep learning. The purpose of using the MNIST dataset for handwritten digit recognition is to predict the number represented by each handwritten digit image through machine learning algorithms. Accuracy is usually used as a performance indicator to evaluate handwritten digit recognition algorithms, that is, to predict the proportion of correct digital images in the total number. Deep learning (DL) of artificial intelligence (AI) is a learning method that simulates the dynamic process of most human brain nerves by computer and draws corresponding conclusions from the data. The reasonable application of this method in calligraphy character recognition and texture mapping can effectively enhance the feature recognition effect of AI on character image information, and build a more efficient image processing capability for AI. Different types of calligraphy have different positions of font information, and the font information belongs to the micro-structure part of the character strokes, so its anti-interference ability is poor. Therefore, it is difficult to recognize calligraphy fonts. ANN is connected by a large quantity of neurons, which can imitate the information processing function of human brain to process high-complexity information, and at the same time, it can abstract the model of similar information and classify the newly received information. Neural network is a pattern recognition method, which can obtain the texture mapping relationship in a given input-output data pair through training and learning, and can still make very good predictions for the data that have not been learned.

Because of the complex structure, huge character set, high character similarity and diverse font styles, the recognition research of calligraphy fonts is still facing great challenges. Image recognition plays a fundamental and important role in the task of computer vision. It is a technology that uses computer vision-related methods and technologies to process images, so that computers can identify information such as the categories and positions of various targets and objects in images. As a new DL model, CNN has the characteristics of multi-core, parameter sharing, local features and global features, etc. In the process of font recognition, it can extract the deep features of characters, reduce the computational complexity, and effectively solve the unique characteristics of calligraphy fonts. Therefore, CNN can be used to effectively identify calligraphy fonts, and the following improvements have been made:

⊖ In this article, an improved method based on optimization strategy is proposed. LeNet-5 network can be improved by selecting appropriate optimizer, setting Loss function, and using regularization and Dropout methods to improve the problem of network overfitting.

⊜ In text and image processing, noise is eliminated from the image, color channels are normalized, and the text in the image is located and adjusted in size. Only the preprocessed image can serve as a training sample set for the neural network.

The article first summarizes the relevant research on text recognition, and then proposes an intelligent recognition model for calligraphy fonts based on texture mapping and CAD; Subsequently, the feasibility of the model was verified on the MNIST dataset; Finally, the achievements and contributions of this article are summarized.

2 RELATED WORKS

Written letter analysis based on Gesture recognition has been greatly improved with the help of computer system model construction. Chang et al. [6] enhanced the model of gesture writing algorithm based on initial module and integrated structure of aviation model character analysis. It analyzed and constructed the character fluidity of spatial recognition gestures, and evaluated the template of data signals. By constructing recognition and writing methods for aerial models, an effective signal template has been constructed. Cloud based CAD parameterization can be used for design space exploration and design optimization in numerical simulation. By parameterizing CAD models, the design space can be represented as a combination of a set of parameters, which can then be optimized through numerical simulation. Guerrero et al. [7] conducted numerical simulations of automatic framework script recognition for spatial design optimization. By designing a parameterized computer-aided program platform, it analyzed the optimization of resource calculation for quantitative spatial exploration. Put the parameterized design into the cloud for calculation, and use the powerful computing power of cloud computing for large-scale calculation and processing. This can greatly improve computational efficiency and speed, while also saving local computer resources. At present, the CAD Assistive technology of digital information is developing rapidly, and the information transformation based on video character recognition has become an important theoretical tool for scene information. Guo and Li [8] conducted a digital network scene assisted frame extraction recognition with the smallest unit of the computer. Computer animation production software can help artists create complex animations. Artists can use computer technology to design virtual scenes, characters, objects, and control their behavior and motion through tools such as keyframes, paths, and constraints. These technologies and methods can help artists more efficiently carry out artistic design and production, and can save a lot of time and cost. At the same time, these technologies can also help artists create more complex and exquisite works, thereby improving the quality and efficiency of artistic works. Optical character recognition (OCR) is a technology that converts printed text into digital text. It can be used to recognize and process text in images. The Nave Bayes classifier is a classifier based on Probability theory, which can be used to classify and recognize data.

Hubert and Sudaryono [9] carried out Optical character recognition of preferential images in the image publicity of enterprise media. By using the naive Bayesian algorithm of images for feature classification, it found the accuracy of the model for automatically promoting enterprise promotional images. Use OCR algorithm to recognize and extract text from promotional images. Commonly used OCR algorithms include feature-based OCR algorithms, deep learning-based OCR algorithms, etc. These algorithms can recognize text, numbers, letters, etc. in promotional images and convert them into numerical text. Jalali and Lee [10] analyzed the robustness of alternative functions for image samples and analyzed their feature selection. It constructs a visual system-based network parameter gradient descent adaptive adjustment model. By using functional learning simulations using computer neural networks, it analyzed the spatial structure balance ability of character classification in traditional literature. It should be noted that in the process of integrating adaptive constraints into deep learning models for character recognition, appropriate parameter tuning and optimization are required to ensure the accuracy and robustness of the model. In addition, it is necessary to select and process the input data to ensure the quality and recognizability of the data. CAD aided image zonal Optical character recognition requires normalization of affine projection at any angle. The accuracy loss of affine image projection recognition requires standardized significant image loss analysis. Scaling and feature normalization are two commonly used methods. Scaling can be achieved by scaling the Image scaling to the standard template size, while feature normalization can be achieved by standardizing each feature. Konovalenko et al. [11] considered the optimal affine normalization problem based on this criterion. For different OCR application scenarios, different affine image normalization methods can be selected. For example, in license plate recognition, commonly used methods are scaling and character segmentation. In document recognition, commonly used methods are scaling and feature normalization. The affine image normalization method in Optical character recognition needs to be selected and optimized according to specific application scenarios to ensure the quality and efficiency of image preprocessing. Liu [12] analyzed the design analysis of CAD based product 3D pattern software for product parameterization. It constructs a technical architecture for visualizing interfaces in large-scale production design processes for enterprises. After establishing a three-dimensional digital model, a CAD system can be used for structural optimization, including analysis of product structure rationality and stability, material selection and optimization, and optimization of structural shape. By optimizing the structure, the quality and performance of the product can be improved, while reducing development costs. Use manufacturing processes and 3D digital models generated by CAD systems for product manufacturing. During the manufacturing process, CAD systems can be used for manufacturing process management and quality monitoring to ensure that the quality and performance of products meet design requirements.

Liu and Yang [13] designed and created an open database for creative teaching based on modularity. Through the open language learning environment of this database, it is possible to access the modular design principles of developing web pages. Through cross-border integration, students can understand knowledge and skills in different fields and cultivate more comprehensive abilities. It can guide students to combine Computer-aided design technology with other disciplines or fields, such as culture, science and technology, society, etc., so that students can understand the needs and characteristics of different fields and improve their comprehensive application ability. Ott et al. [14] analyzed and recognized a sensor character recognition system based on online characters. By collecting information from handwritten datasets, the accuracy of classifier effects for different uppercase and lowercase characters was analyzed and constructed. When using machine learning to enhance online Handwriting recognition from IMU, it is necessary to pay attention to the accuracy of data preprocessing and feature extraction, and select appropriate machine learning models and parameter settings. At the same time, it is also necessary to standardize and standardize the input handwritten motion data to ensure the stability and accuracy of the model. In addition, the real-time performance and delay time of the model should also be considered to facilitate effective online Handwriting recognition in practical application scenarios. Due to the wide variety of Chinese strokes, different characters may have high similarity, such as "day" and "mouth", which increases the difficulty of recognition algorithms. The writing method of handwritten Chinese varies from person to person, and each person may have their own writing habits and styles, which increases the difficulty of recognition algorithms. In practical application scenarios, handwritten Chinese may be subject to various noise and interference, such as stroke breakage, stroke concatenation, etc., which requires recognition algorithms to have strong anti-interference capabilities. Offline handwritten Chinese recognition is a challenging task that requires research and development of efficient recognition algorithms and technologies to achieve accurate and fast handwritten Chinese recognition.

Shen et al. [15] constructed an offline Chinese handwriting pattern through pattern recognition analysis of offline handwriting. Offline handwritten Chinese recognition requires high real-time performance, as it requires accurate recognition and classification of input handwritten characters in a short period of time. Wang and Chen [16] proposed a visual writing expression tool that reflects the author's emotions. By analyzing and measuring the visualization of the author's modality changes during the writing process, a semi structured emotional analysis contact method was constructed. After calculating the user's emotions, it is necessary to present them visually. Various visualization tools and technologies can be used, such as graphical interfaces, animations, virtual reality, etc. Through visual presentation, users can have a more intuitive understanding of their emotional state and establish interaction with computers. Yahya et al. [17] conducted a large amount of dataset algorithm accuracy analysis, which constructed training feature error image impact analysis for redundant targets. This method eliminates the limitation of target independent variables between data and the negative impact of noise. It uses data augmentation technology to expand and enhance the data, or uses model compression technology to reduce the computational complexity and latency of the model. These technologies can improve the efficiency and accuracy of digital recognition, thus being better applied in practical application scenarios. Edge depth learning based on ToF sensor Real time Cursive character recognition is a method that uses ToF sensor and depth learning technology to realize real-time Cursive character recognition. This method can recognize handwritten characters quickly and accurately, and process them in real time on Edge device. Zhang et al. [18] analyzed the real-time inference recognition task of sensor arrays. This method constructs the performance of learning algorithms by analyzing and designing different written letters. Other technologies can also be combined to improve recognition efficiency and accuracy, such as using data augmentation technology to expand and enhance data, or using model compression technology to reduce model computation and latency.

At present, there are many methods to recognize characters, but the recognition rate is still not high enough for handwritten characters and characters with large character sets. Aiming at the low recognition rate of handwritten conjoined characters in the existing network, this article proposes an improved method based on optimization strategy. For LeNet-5 network, the problem of over-fitting is improved by selecting appropriate optimizer, setting loss function, and using regularization and Dropout methods, and the accuracy of calligraphy font recognition is further improved.

3 INTELLIGENT RECOGNITION OF CALLIGRAPHY FONT BASED ON TEXTURE MAPPING AND CAD

3.1 The Application of CNN in Character Recognition

Texture mapping is a widely used technology in computer graphics. It can map a texture image to the surface of an object to increase the surface detail and complexity. In intelligent font recognition, texture mapping can be used to enhance the feature representation of fonts, thereby improving recognition accuracy. The specific application method can be to first image process the font and convert the text into image data. Then, the texture image is fused with the text image, and through a certain mapping algorithm, the text has more visual features. These features can be recognized by machine learning algorithms for more accurate intelligent font recognition. CAD (Computer-aided design) technology can also play an important role in intelligent font recognition. Through CAD technology, fonts can be transformed into 3D models, allowing for analysis and recognition of fonts from more perspectives and dimensions. The specific application method can be to first use CAD technology to convert text into a 3D model, and then use computer vision and deep learning algorithms to analyze and process the model. By analyzing the shape, contour, proportion and other features in the model, and combining them with machine learning algorithms for training, intelligent recognition of fonts can be achieved. Overall, both texture mapping and CAD have their own advantages and application scenarios in font intelligent recognition. By combining these two and utilizing current advanced deep learning algorithms, intelligent font recognition can be more precise, providing people with more efficient and accurate text recognition services. However, the specific application methods still need to be designed and optimized according to the actual situation to achieve the best results.

Character recognition is a hot research work in the field of pattern recognition, and the earliest research is mainly on the recognition of English and Arabic numerals. Character recognition can be regarded as a complex pattern recognition system, and English letters can be regarded as a slightly simple pattern recognition because of the limited quantity of letters. With the improvement of computer performance and the enrichment of large-scale data sets, CNN has made a major breakthrough in the field of computer vision. Its speed and accuracy far exceed those of self-encoder, deep belief network and other artificial neural networks, and gradually replace the traditional artificial neural networks to become the basic network structure of mainstream DL.

In order to solve the problem of small samples, a new data amplification method is proposed in this article. That is, the original sample is subjected to ripple deformation, rotation, translation and other ways to amplify the sample in real time. In order to further shorten the training time, Adam and SGD are comprehensively used in the training process of the network, which realizes the combination of rough training and fine tuning of the network. In this article, the gray processing can not only improve the efficiency of the algorithm at pixel level, but also make the gray difference between the lines of the text and the middle blank lines obvious through the unique pixel value, which will improve the good foundation for the subsequent text segmentation. The CNN recognition model of calligraphy font based on texture mapping and CAD is shown in the figure 1.

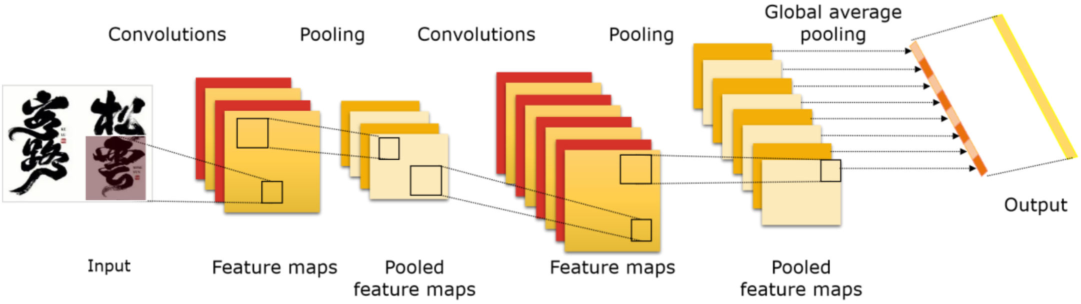

Figure 1: CNN recognition model of calligraphy font.

Local connection means that the convolution kernel only covers a local area of the input data volume during each convolution operation, and each neuron of the feature map is locally connected with the input data. This characteristic considers the spatial structure of the data and ensures that the convolution kernel can respond strongly to the local input pattern in space after learning. In the process of character recognition, it usually begins with the processing of pictures, because pictures, as an intermediate medium from paper characters to computer digital information, have good properties, which not only completely reflect all the information of paper characters, but also can be completely recognized by computers, and there are many ways to choose flexibly in the process of acquisition and transmission. The CNN function is defined as:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l\right) \tag{1}$$

Represent the output feature map in a layer.

$$F_j^{(n)} = \sum_i w_{ij}^{(n)} * F_i^{(n-1)} + b_j^{(n)} \tag{2}$$

Where: $*$ is the 2D convolution; $w_{ij}^{(n)}$ and $b_j^{(n)}$ are the convolution filter and bias, respectively; $F_j^{(n)}$ is the j-th output feature map in the n layer.

3.2 Intelligent Recognition of Calligraphy Font Based on Texture Mapping

For the application scene of calligraphy font recognition, excessive image processing to achieve text data amplification will lead to problems such as missing key structure and inverted character structure, which will make the font structure change too different, resulting in the amplification data generated by this kind of processing method no longer suitable for font recognition. Therefore, this article only uses image scaling, image rotation, convex deformation and other image processing methods that will not excessively affect the font characteristics to expand the character database. The segmentation method adopted in this article is the global threshold segmentation method, and the threshold is set to 50, that is, pixels with gray value less than 50 will be considered as noise, while pixels with gray value greater than 50 will be retained. On this basis, if the background in the image is all black, that is, the gray value is 0, and the target part is all white, that is, the gray value is 275. In this way, the binarization of an image can be completed. There are many ways to realize the image graying effect in the calligraphy font image, and the commonly used graying algorithms are as follows:

Maximum method:

$$g(i,j) = \max\{R(i,j), G(i,j), B(i,j)\} \tag{3}$$

Average method:

$$g(i,j) = \left(R(i,j) + G(i,j) + B(i,j)\right)/3 \tag{4}$$

Weighted average method:

$$g(i,j) = R(i,j) \times 0.229 + G(i,j) \times 0.587 + B(i,j) \times 0.114 \tag{5}$$

Among them, $g(i,j)$ represents the grayscale pixel value of the (i,j) coordinate position of the grayscale image; $R(i,j)$ represents the pixel value of the (i,j) position of the color image red channel; $G(i,j)$ represents the pixel value of the (i,j) position of the color image green channel; $B(i,j)$ represents the color image blue channel (i,j) position of the pixel value Pixel values. In this article, the weighted average method is used to gray the digital image. Assume that the input picture size is I * I; The convolution kernel size is K* K; Step size is S; ; The quantity of pixels filled is P; The calculation formula of the size O of the characteristic map output by that conv layer is:

$$O = \left(I - K + 2P\right)/S + 1 \tag{6}$$

In this article, ReLU nonlinear function can speed up training. Moreover, it uses overlapping pooling layer, Dropout and data enhancement technology to solve the over-fitting problem in deep CNN, so that the neural network that appears later can be deeper and deeper and achieve better performance. The down-sampling layer locally averages the feature mapping plane of the conv layer, but it also reduces the resolution of the feature mapping plane, which is why it is called down-sampling layer. This operation can reduce the sensitivity of the network to graphic translation and deformation. After image scaling, bilinear interpolation assignment method is used to ensure that there will be no sawtooth and unclear structure after image scaling, and the structural characteristics of the original character image will be preserved. Figure 2 shows the convolution operation process.

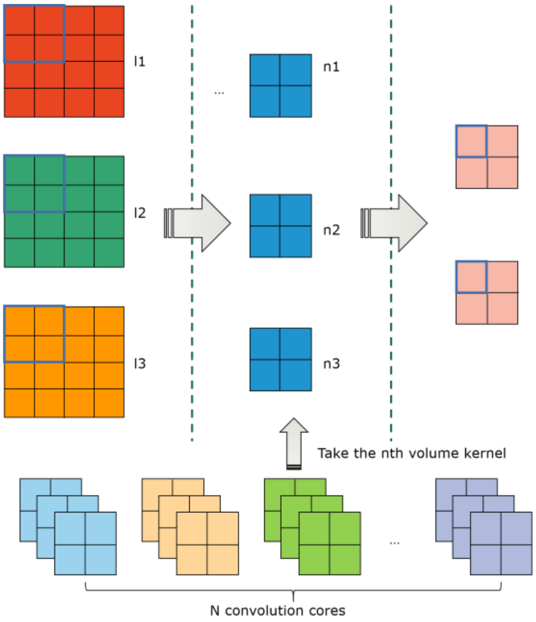

Figure 2: Process diagram of convolution operation.

In order to analyze and process the image, we must first get the feature information of the image. These features must clearly reflect the shape features of the graphics contained in the picture and be quite different from other shapes. Image denoising is to reduce the noise information in digital images through certain technical means. Noise information will lead to inaccurate text positioning in text images and affect the accuracy of subsequent text recognition. For the same window function, its passband width is inversely proportional to the window length. Therefore, if you want a high-resolution font image, the window length should be as long as possible. The convolution calculation is expressed as:

$$\mathrm{s}(\mathrm{t}) = \mathrm{x}(\mathrm{t}) * w(\mathrm{t}) = \sum_{\tau=-\infty}^{\tau=+\infty} x(\tau) w(t-\tau) \tag{7}$$

Where $\mathrm{x}(\mathrm{t})$ is the input feature; $w(\mathrm{t})$ is feature mapping. When processing 2D matrix data, it can be expressed as:

$$\mathrm{s}(\mathrm{i, j}) = \sum_{M=0}^{M} \sum_{n=0}^{N} \left(w_{m,m} x_{i+m} + w_b \right) \tag{8}$$

Because the input layer of the training network is complex, considering that the model should not be too complex, after many model trainings, the hidden layer finally adopts two-layer structure, each layer has 20 neurons.

By preprocessing the image to remove the noise in the image, the image to be recognized becomes a binary image with only black and white pixels, which provides a good foundation for subsequent text segmentation. Normalization is an important link in pretreatment. Because the original images are very different in size and shape, it is necessary to normalize them to make them have the same size and shape, so as to facilitate the process of feature extraction and recognition.

4 RESULT ANALYSIS AND DISCUSSION

Data sets play a vital role in the application of neural networks. The reason why neural network can achieve such great success in visual problems is inseparable from the increasingly large-scale, more accurate and more scientific annotation data sets. When implementing a neural network project, it is more important to deal with the data set needed by the project besides building the network structure of the model. In this chapter, the algorithm is mainly tested in MNIST handwritten character image database. The MNIST handwritten character image database is established and maintained by Corinna Cortes of Google Lab, Christopher J.C.Burges of Microsoft Research Institute and Yann Le Cun of new york University. It is one of the most widely used image collections in the field of image processing. The database contains 50,000 training images, 10,000 verification images and 10,000 test images. Each image is a gray image with 28*28 pixels, which contains a set of 10 handwritten characters from 0 to 9.

The calligraphy image is preprocessed, and the feature vector of each character is extracted by using the corresponding feature extraction method, and the feature vector is stored in the character feature database. The purpose of preprocessing is to remove noise, absorb character deformation, strengthen useful information and compress redundant information, and deal with the position, size and shape of characters as much as possible, which makes preparations for feature extraction. This article uses a certain data augmentation method to preprocess the training image. Based on the characteristics of the sample images used in the experiment, and without changing the text structure, this article adopts the data amplification method of ripple distortion combined with translation, rotation and scale scaling. The parameters of each layer of CNN are shown in Table 1.

Number	*Layer type*	*Output size*
1	Input layer	32×32×1
2	Conv layer 1	32×32×64

3	Pool layer 1	16×16×64
4	Conv layer 2	16×16×128
5	Pool layer 2	8×8×128
6	Conv layer 3	8×8×256
7	Pool layer 3	8×8×256
8	Conv layer 5	8×8×256
9	Pool layer 4	4×4×256
10	Conv layer 6	4×4×512
11	Batch standardization layer 1	4×4×512
12	Conv layer 7	4×4×512
14	Batch standardization layer 2	4×4×512
15	Pool layer 5	2×2×512
16	Conv layer 8	2×2×512
17	Dropout layer	2×2×512
18	Sofumax regression layer	1×6676

Table 1: Description of parameters of CNN layers.

The padding setting of all conv layers is zero-padding mode, and the step size is set to 1. In order to improve the speed of calligraphy font recognition, the sum of all feature vectors of each character is obtained to establish an index, and then a multi-template character morphological feature database is designed according to the result of feature vector processing. For CNN, the quantity of feature maps in the conv layer can be adjusted. The more the number, the more local features can be extracted, but at the same time, the network scale is getting bigger and bigger, and the training is more time-consuming. In this chapter, many experiments were conducted, and the data of two experiments were selected to draw a data map, as shown in Figure 3.

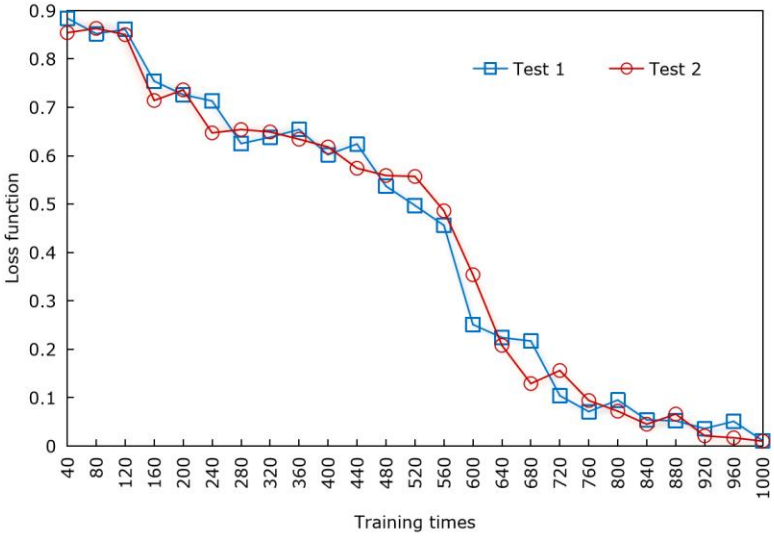

Figure 3: Loss function of calligraphy font recognition model.

The model in this article has gone through 700 iterations, and the loss function of calligraphy font recognition decreases rapidly with the iteration, and gradually tends to 0. The weight sharing mechanism is adopted because the local statistical features of the image have nothing to do with the position and are repetitive in the whole image, and the information learned in local areas can also be applied to other parts of the image, that is, the same convolution kernel deconvolves the whole image, which is equivalent to filtering every part of the image.

In this article, the convolution feature is used because the image has a "static" property, which means that two different image regions may have the same feature. Therefore, when describing a large image, we can calculate the maximum or average value of the features of each region on the image to make aggregate statistics, thus effectively reducing the dimension of the image. Before fitting measures are added, the change curve of recognition accuracy is shown in Figure 4.

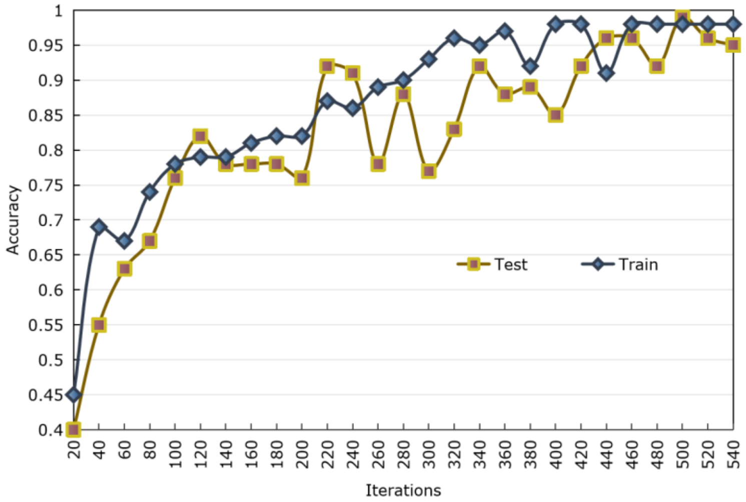

Figure 4: Changes of accuracy before fitting measures are added.

It can be observed that the expressive force of CNN in the text test set is worse than that in the same training set, and there is an over-fitting phenomenon. In order to prevent over-fitting, this article adopts Dropout and L2 regularization methods to prevent over-fitting, and the experimental results are shown in Figure 5.

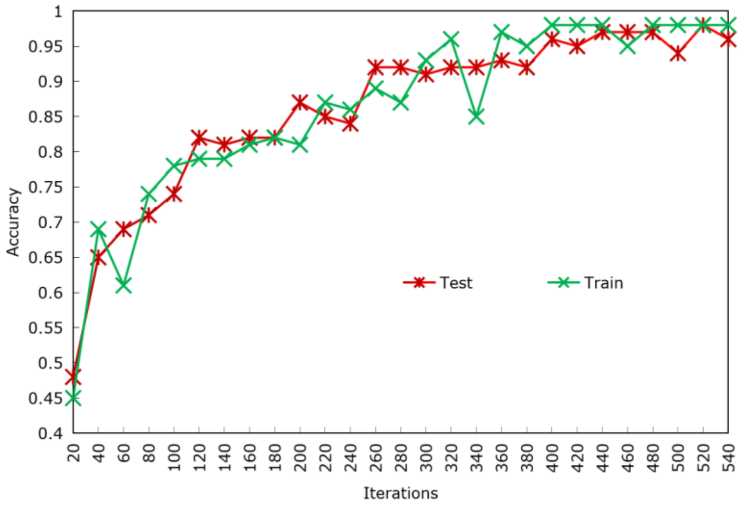

Figure 5: Changes in accuracy after adding over-fitting measures.

According to the comparison between the curve of accuracy change in the figure and that before adding the over-fitting measure, it can be known that through repeated iterative training, the Dropout and L2 regularization method added in this article can effectively prevent over-fitting, strengthen the generalization ability of the network and effectively complete the task of font recognition. Characters with different writing styles have different changes according to their stroke distribution. In fact, dividing handwritten characters into elastic grids is a nonlinear normalization of text images, which can effectively reflect the common characteristics of handwritten characters.

In order to verify the accuracy of the whole model, this article selects 8000 unusual character combinations to form a test set to test the accuracy of the whole model. Input the picture into the trained model, judge its font by the output probability of the model, and then compare it with the marked font. The accuracy of different algorithms on the training set is shown in Figure 6. The accuracy of different algorithms on the test set is shown in Figure 7.

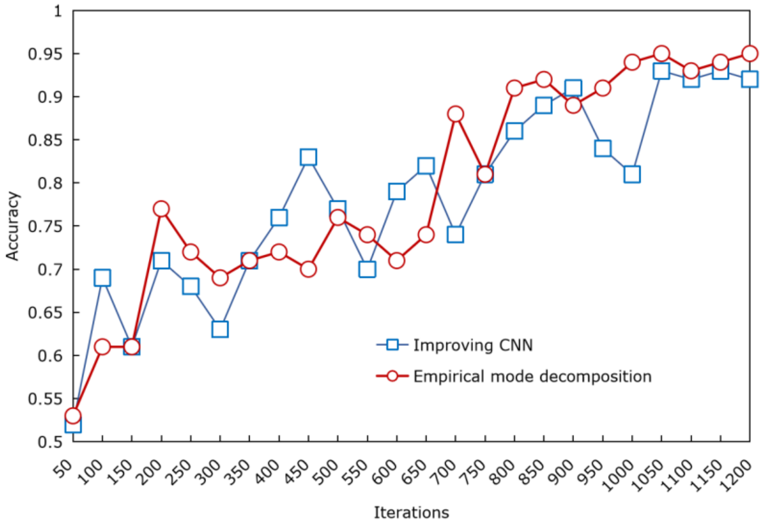

Figure 6: Accuracy of different algorithms on training sets.

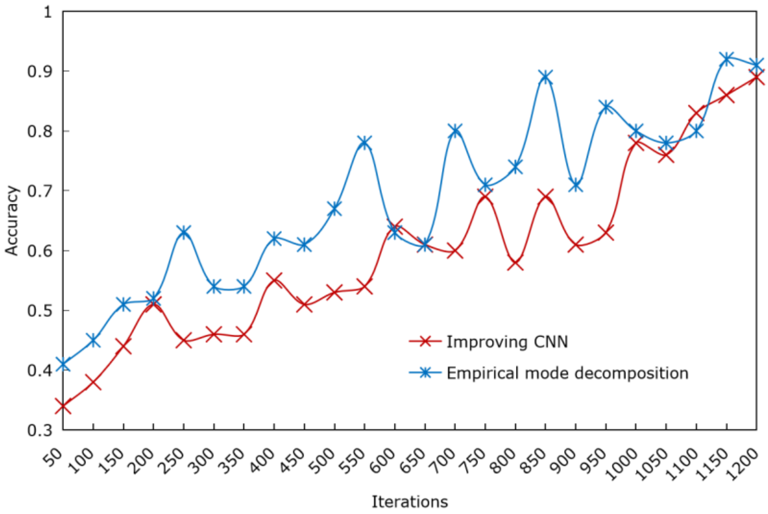

Figure 7: Accuracy of different algorithms on test sets.

Because the network structure of CNN is complex, it is not easy to modify it, and training is time-consuming. The advantage is that convolution network has better effect on character recognition and is more suitable for image recognition. In the process of comparing the character recognition rate of convolutional network with that of other methods, the performance of convolutional network is more excellent. The recognition accuracy of the intelligent calligraphy font recognition algorithm designed in this article is better than that of the traditional empirical mode decomposition algorithm, which effectively improves the recognition accuracy of calligraphy fonts.

5 CONCLUSIONS

Calligraphy font recognition is a technique to identify the calligraphy font to which the characters in the image belong by analyzing the character images. The existing optical character recognition technology has very high recognition efficiency and accuracy. However, the disadvantage of optical character recognition technology is that it can only extract the text content information on the image, but can't extract the font information of the text. From the visual sense, the visual feeling of each font is different. Therefore, the different writing methods of each font and different text images can be represented by a texture feature. In this article, the intelligent recognition algorithm of calligraphy font based on texture mapping and CAD is studied, and the rapid recognition model of calligraphy font features is constructed by CNN, and the extraction method of this feature parameter is improved. In the process of model training, this article comprehensively uses strategies such as data amplification, combination of various optimization methods and batch standardization, and verifies the effectiveness of the network through a large quantity of comparative experiments. Applying the CNN model constructed in this article to the field of calligraphy font recognition can not only solve the problem of rapid recognition of calligraphy font, but also broaden the application field of ANN. In the process of calligraphy font recognition, the multi-level classifier improves the recognition rate, but it also affects its recognition speed. It is need to conduct in-depth research on how to balance the recognition speed and accuracy.

Mingke Liu, https://orcid.org/0009-0008-1346-6391
Yue Zhang, https://orcid.org/0009-0003-9524-2749

REFERENCES

[1] Alam, M.-S.; Kwon, K.-C.; Alam, M. A.; Abbass, M.-Y.; Imtiaz, S.-M.; Kim, N.: Trajectory-based air-writing recognition using deep neural network and depth sensor, Sensors, 20(2), 2020, 376. https://doi.org/10.3390/s20020376

[2] Albattah, W.; Albahli, S.: Intelligent arabic handwriting recognition using different standalone and hybrid CNN architectures, Applied Sciences, 12(19), 2022, 10155. https://doi.org/10.3390/app121910155

[3] Alemayoh, T.-T.; Shintani, M.; Lee, J.-H.; Okamoto, S.: Deep-learning-based character recognition from handwriting motion data captured using imu and force sensors, Sensors, 22(20), 2022, 7840. https://doi.org/10.3390/s22207840

[4] Annanurov, B.; Noor, N.: A compact deep learning model for Khmer handwritten text recognition, IAES International Journal of Artificial Intelligence, 10(3), 2021, 8938-8938. https://doi.org/10.11591/ijai.v10.i3.pp584-591

[5] Baldominos, A.; Saez, Y.; Isasi, P.: A survey of handwritten character recognition with mnist and emnist, Applied Sciences, 9(15), 2019, 3169. https://doi.org/10.3390/app9153169

[6] Chang, W.-D.; Matsuoka, A.; Kim, K.-T.; Shin, J.: Recognition of uni-stroke characters with hand movements in 3d space using convolutional neural networks, Sensors, 22(16), 2022, 6113. https://doi.org/10.3390/s22166113

[7] Guerrero, J.; Mantelli, L.; Naqvi, S.-B.: Cloud-based CAD parametrization for design space exploration and design optimization in numerical simulations, Fluids, 5(1), 2020, 36. https://doi.org/10.3390/fluids5010036

[8] Guo, S.; Li, X.: Computer aided art design and production based on video stream, Computer-Aided Design and Applications, 18(S3), 2020, 70-81. https://doi.org/10.14733/cadaps.2021.S3.70-81
[9] Hubert, P.-P.; Sudaryono, R.: Classifying promotion images using optical character recognition and nave bayes classifier, Procedia Computer Science, 179(1), 2021, 498-506. https://doi.org/10.1016/j.procs.2021.01.033
[10] 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
[11] 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
[12] Liu, F.: Fast industrial product design method and its application based on 3d CAD system, Computer-Aided Design and Applications, 18(S3), 2020, 118-128. https://doi.org/10.14733/cadaps.2021.S3.118-128
[13] 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
[14] Ott, F.; Wehbi, M.; Hamann, T.; Barth, J.; Eskofier, B.; Mutschler, C.: The onhw dataset: Online handwriting recognition from imu-enhanced ballpoint pens with machine learning, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 4(3), 2020, 1-20. https://doi.org/10.1145/3411842
[15] Shen, L.; Chen, B.; Wei, J.; Xu, H.; Tang, S.-K.; Mirri, S.: The challenges of recognizing offline handwritten Chinese: A Technical Review, Applied Sciences, 13(6), 2023, 3500. https://doi.org/10.3390/app13063500
[16] Wang, C.-M.; Chen, Y.-C.: Design of an interactive mind calligraphy system by affective computing and visualization techniques for real-time reflections of the Writer's emotions, Sensors, 20(20), 2020, 5741. https://doi.org/10.3390/s20205741
[17] Yahya, A.-A.; Tan, J.; Hu, M.: A novel handwritten digit classification system based on convolutional neural network approach, Sensors, 21(18), 2021, 6273. https://doi.org/10.3390/s21186273
[18] Zhang, J.; Peng, G.; Yang, H.; Tan, C.; Tan, Y.; Bai, H.: Real-time finger-writing character recognition via tof sensors on edge deep learning, Electronics, 12(3), 2023, 685. https://doi.org/10.3390/electronics12030685