Proceedings of the 5th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/100/2025.18118

Research on Minority Character Recognition

- Taking Small Seal Font as an Example

Chuyan Zeng[1,a,*]

[1]School of Economics and Management, Tiangong University, Tianjin, China
a. 2210650206@tiangong.edu.cn
**corresponding author*

Abstract: Text recognition is one of the important fields of computer vision and is widely used in automated office, assisted reading and other fields. With the continuous development of deep learning, the visual language text recognition method generated by the combination of optical character recognition and natural language processing involves the technology of recognizing text information from images or videos, which greatly improves the machine's understanding ability and interaction efficiency. However, for some niche fonts, such as Urdu, Xixia, and Xiaozhuan, these characters have complex structures, diverse strokes, and sloppy writing, making them a challenging field in text recognition. Taking Small Seal Scripts, also called Xiaozhuan, as an example, this paper summarizes the advantages and disadvantages of the current Xiaozhuan text recognition algorithm, and proposes algorithm improvement suggestions for the segmentation of Xiaozhuan characters on seals. Xiaozhuan fonts are quite different from modern Chinese characters, mainly reflected in the stroke structure, writing style, and the correlation between characters. It has smoother lines and many glyphs are unique. This complexity makes it difficult to directly apply traditional Optical Character Recognition (OCR) technology. Therefore, many studies have combined image preprocessing, character segmentation, and deep learning models for automatic recognition. Future research should focus more on the generalization and lightweight of the model so that it can run on devices with limited computing resources.

Keywords: Character recognition, Small Seal Scripts, deep learning.

1. Introduction

Now is the information age, people are faced with not only the number of information is large and messy, in terms of other information carriers, the advantages of the text are not only easy to save, but also easy to transmit information, it is so that the information in time and space have been rapidly spread. People need to recognize a very large number of words in their lives, with the continuous development of science and technology, text recognition applications have been expanded accordingly.

Automatic text recognition is a task that is both labor-intensive and difficult to accomplish. The recognition of English characters has entered the experimental stage since the late 1950s. Relying on the uniqueness of the 26 letters of the English alphabet and the small number of word combinations, the recognition of handwriting can achieve a good effect. However, the recognition of handwriting

DOI: 10.54254/2755-2721/100/2025.18118

and other texts is still not well solved. After the 1980s, a large number of research institutions with text recognition as the core emerged around the world and made outstanding contributions.

As for Chinese character recognition, due to the complexity and particularity of Chinese, and the late start of character recognition in China, the products currently available on the market from research institutions can maintain a relatively ideal recognition rate, but there are also some shortcomings that need to be improved.

At present, the main methods of text recognition are optical character recognition (OCR) and deep learning-based algorithms. Optical character recognition uses electronic devices to extract the shape, size and other features of text images to recognize characters; while deep learning-based algorithms, such as convolutional neural networks (CNN) and recurrent neural networks (RNN), learn text features and perform recognition by training large amounts of data. In addition, there are some algorithms that combine deep learning and other technologies, such as convolutional recurrent neural networks (CRNN), which can process text sequences of indefinite length and improve recognition accuracy.

For some niche scripts, special algorithms need to be designed for recognition. For example, the data volume of Xixia script is extremely limited, mainly from existing documents and cultural relics. The images obtained are noisy and often incomplete. By improving the DenseNet network and introducing spatial channel reconstruction convolution to replace the traditional 3×3 convolution of the original model, the feature representation ability of the network can be improved [1]. There are many diacritical marks in Arabic and the writing is sloppy, with uneven spacing within words and between words, which brings great difficulties to Arabic recognition. The Arabic video text recognition system based on MDLSTM network and CTC output layer uses a non-segmentation method, which is better than the current recognition method in the line-level recognition rate of Arabic characters [2]. Urdu characters have complex shapes and lack publicly available datasets. Noor ul Sehr Zia et al. proposed an end-to-end handwriting recognition system based on a new CNN-RNN architecture with n-gram language modeling, and constructed a new unconstrained dataset called NUST-UHWR, which achieved the lowest character error rate of 5.28% in Urdu handwriting recognition [3].

This article takes the Xiaozhuan font as an example to explore the recognition method of niche fonts and puts forward feasible improvement suggestions for the existing algorithms. Xiaozhuan fonts appear in various books, cultural relics, etc., with the characteristics of multiple shapes of one word and stroke adhesion. At present, some models have good efficiency in Xiaozhuan font recognition, but the problem of character adhesion on the seal has always been a difficulty in image segmentation in the model. Since handwritten fonts also have the problem of character adhesion, you can refer to how the handwritten character segmentation algorithm solves the problem of character adhesion.

2. Text recognition methods

2.1. Traditional text recognition methods

2.1.1. Statistical decision-making

This method is based on probability theory and mathematical statistics, and has many typical representatives, such as the Bayes classification algorithm. This method has strong anti-interference ability, but has high requirements for feature selection.

2.1.2. Structural feature method

This method analyzes the glyph structure of the text in depth, splits the characters into many primitives by feature extraction, and sorts them according to certain corresponding criteria to form

DOI: 10.54254/2755-2721/100/2025.18118

the characteristics of the characters. Then, the characters are mapped to the structural space composed of primitives by mapping, and finally the recognition results are obtained. This method relies on a lot of storage and computing support to describe and compare the features, but the recognition accuracy is generally high.

2.1.3. Template recognition method

This method requires the establishment of a complete standard matching character library, which is used as a matching template. During the recognition process, the character to be recognized is matched with the data in the library, and the classification with the smallest final difference is selected as the category to which the character to be tested belongs. This method has strong limitations. It has good support for handwriting, but is not suitable for handwriting recognition of different styles.

2.2. Steps of text recognition based on machine learning

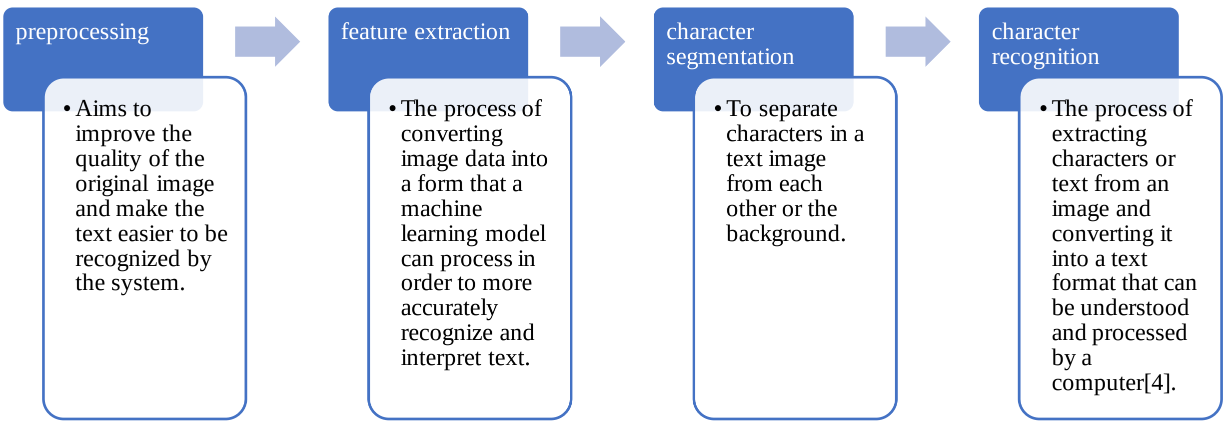

Figure 1: Steps of text recognition based on machine learning.

Figure 1 shows the steps of text recognition based on machine learning [4].

3. A review of text recognition algorithms based on machine learning

3.1. Recognition method of small seal scripts in tiles based on neural network

Due to the particularity of the small seal script that the same meaning has different glyphs, this study proposed a sample grouping algorithm based on mean nearest neighbor, NMSG. Before learning the algorithm, the tile end digital samples were grouped to solve the problem of multiple glyphs for one character to a certain extent. The recognition rate in the self-owned tile end image library was as high as 86.9%. However, due to limited experimental data, this study only selected 10 characters and a total of 100 samples for the experiment. As the amount of data increases, the recognition results cannot be guaranteed [5]. Figure 2 shows different forms of the same character.

DOI: 10.54254/2755-2721/100/2025.18118

Figure 2: Different forms of the character “长” [5]

3.2. A style-independent method for learning small seal character root structure sequences

In order to solve the problem of numerous styles and unbalanced samples of small seal characters, a style-independent small seal character root structure sequence learning method is proposed to achieve small sample recognition of small seal characters. By utilizing the consistency of the structure of small seal characters in different styles, style transfer learning of multiple styles and standard style small seal characters is achieved, and the common spatial structure and root radical information are learned from pictures of small seal characters in different styles. Figure 3 shows the spatial structure of Chinese characters.

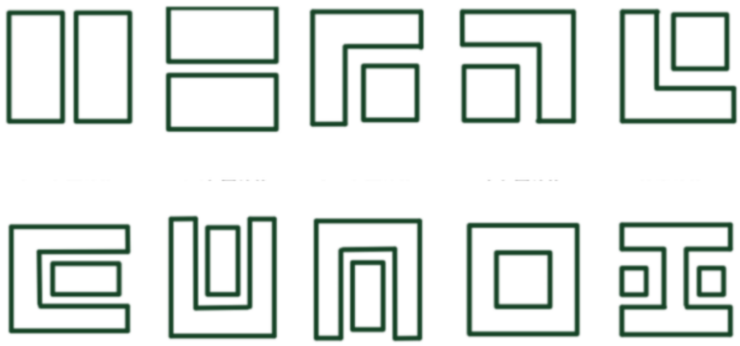

Figure 3: Spatial structure of Chinese characters[6].

Although this model can achieve a higher recognition accuracy rate for small-sample seal characters than existing network models, its accuracy rate for seal characters of unknown categories still needs to be improved. In addition, this model is currently only applicable to printed seal characters images[6]. Figure 4 shows the framework diagram of the learning network for style independent small seal character root structure.

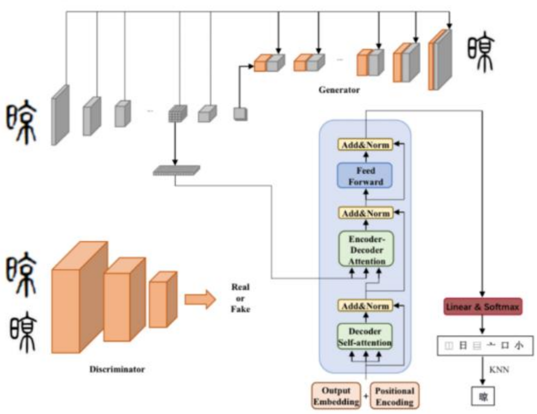

Figure 4: Framework diagram of the style-independent small seal character root structure learning network [6].

DOI: 10.54254/2755-2721/100/2025.18118

3.3. A Zero-Sample small seal character recognition model based on pictographic radical sequence learning

Although there are commercial small seal characters, there are still a large number of small seal characters that are not included in "Shuowen Jiezi", which means that small seal recognition faces the challenge of unseen writing styles or zero samples. TSeal is a zero-shot small seal recognition model based on pictographic radical sequence learning. TSeal uses a pictographic feature encoder to extract pictographic features from small seal images, and then uses a radical decoder based on the Transformer model to learn the radical structure sequence from the pictographic features of small seals. Finally, the K nearest neighbor algorithm (KNN) is used to find the corresponding Chinese character from the radical sequence dictionary, thereby achieving zero-shot small seal recognition of unseen characters. Figure 5 lists the structure and radical sequence of a character.

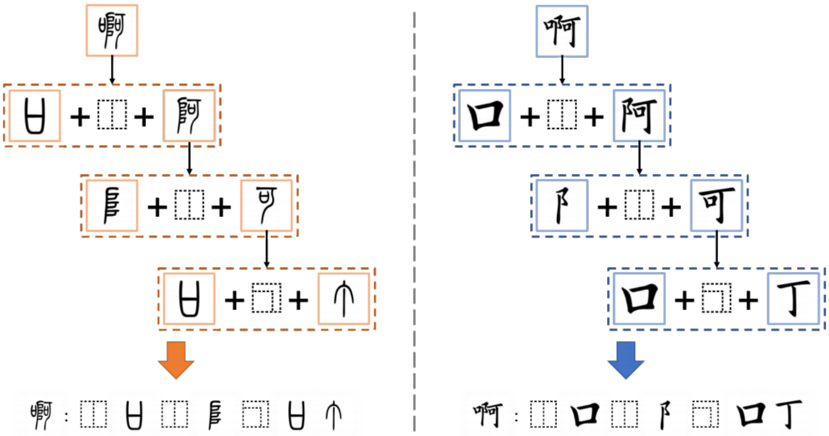

Figure 5: The structure and radical sequence of the character “啊” [7].

To address the challenge of learning semi-/unstructured pictographic radicals, a feature-level collaborative Xiaozhuan recognition model UGTSeal was designed using TSeal as the baseline network. UGTSeal can achieve an accuracy of 93.82% for Xiaozhuan recognition with unseen writing styles, 71.94% for zero-shot recognition with known writing styles but unseen Xiaozhuan characters, and 66.09% for zero-shot recognition with unseen writing styles and Xiaozhuan characters.

To further improve the accuracy of zero-shot recognition of seal script, the Vision Transformer (ViT) architecture was introduced into UGTSeal to replace the original pictographic feature encoder, and the VGTSeal model was designed. A spatial attention loss function was proposed to make the model pay more attention to the learning of the spatial structure of seal script. For zero-shot recognition of unseen writing styles and seal script, the accuracy was improved to 72.61%, which is significantly higher than the existing zero-shot recognition methods. However, the model is still only applicable to printed seal script images [7].

3.4. Adversarial Sequence-to-Sequence Domain Adaptation(ASSDA)

ASSDA is a new "adversarial sequence-to-sequence domain adaptation method" for robust text image recognition. It focuses on the alignment of local character regions, improving the problem that traditional methods can only handle overall image alignment.

This method uses adversarial training and adopts a two-level alignment mechanism: global level alignment and local character level alignment, so that the model can better handle the conversion between different domains. The innovation of this method is that it introduces an attention mechanism to automatically locate the character region in the recognition of sequence images, and further combines the domain adaptation technology to align the source domain and the target domain at the local character level. And a spatial normalization network is designed to deal with the problem of geometric deformation in different scenarios. This algorithm can be added to the previous two small seal character recognition algorithms to solve the problem that the model can only be applied to printed text [8].

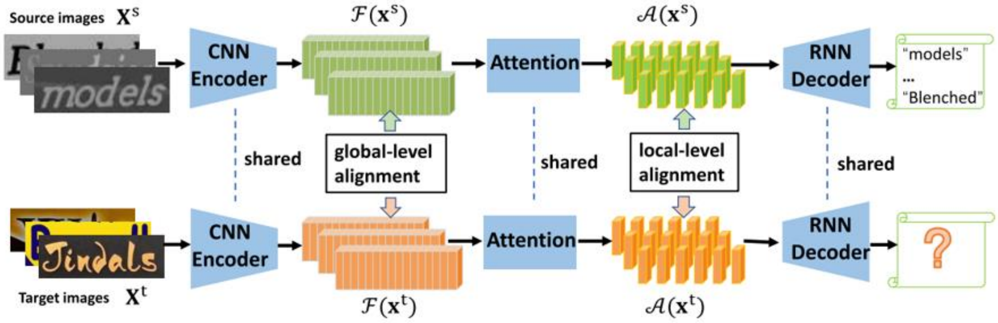

Figure 6: The structure of ASSDA [8]

The structure of ASSDA consists of the following components (Figure 6): a CNN encoder to map the input images into a sequence of high-level feature vectors, an attention unit between the encoder and decoder to adaptively focus on the location of the character, and an RNN decoder to convert encoded features into output strings recurrently. The paper tackles the domain shift on two levels, the global-level alignment and the local-level alignment, where two domain classifiers are built on two levels and trained in an adversarial training manner.

3.5. A Chinese Seal Recognition System(CSRS)

Small seal characters often appear on seals. CSRS is a high-precision Chinese seal recognition system, which consists of the following three parts: 1. Siamese-MTL, which can effectively solve the similarity measurement problem and improve the generalization of the model; 2. ABG, which can generate a large number of seal images with different backgrounds for effective training; 3. A new Siamese network training method based on center constraints. The study also established two large-scale seal databases, including 15,000 Chinese seal images and 1,700 background images, which have great practical application potential in seal recognition. However, the running speed is slow and it takes up a lot of memory [9]. Figure 7 shows an example of an impression image.

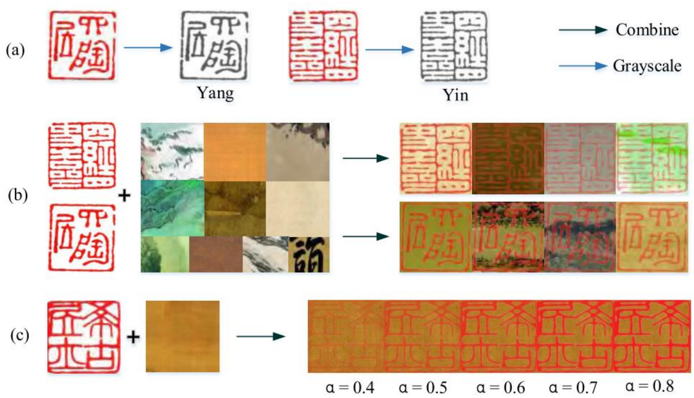

Figure 7: Examples of seal images. (a) The carving techniques of Yin and Yang. (b)Seals with different backgrounds. (c) Seals with different combine ratio[9].

DOI: 10.54254/2755-2721/100/2025.18118

3.6. SEL-RefineMask

The SEL-RefineMask algorithm combines Feature Pyramid Network Selector (SEL-FPN) with RefineMask to segment and identify seals. SEL-FPN can intelligently select feature layers of different scales, reduce the number of anchor boxes, and improve the accuracy and efficiency of segmentation. In the segmentation module, the improved RefineMask is introduced for high-quality character segmentation. For character recognition, a visual transformer (ViT) is used to classify the segmented seal character images and convert them into simplified Chinese characters. The author also constructed two public datasets: SACB (for seal segmentation, containing 12,000 seal images) and SSF (for recognition of small seal fonts in seals, containing 1,000 character categories). On the SACB dataset, the segmentation results (AP value) of SEL-RefineMask are 3.19% higher than other baseline methods. In the character recognition experiment on the SSF data set, the ViT model achieved a classification accuracy of 91%, significantly better than traditional SVM, AlexNet and other models. This algorithm performs well in segmenting adhesive characters and removing seal borders, but the recognition accuracy still needs to be improved [10]. Figure 8 shows the specific model structure.

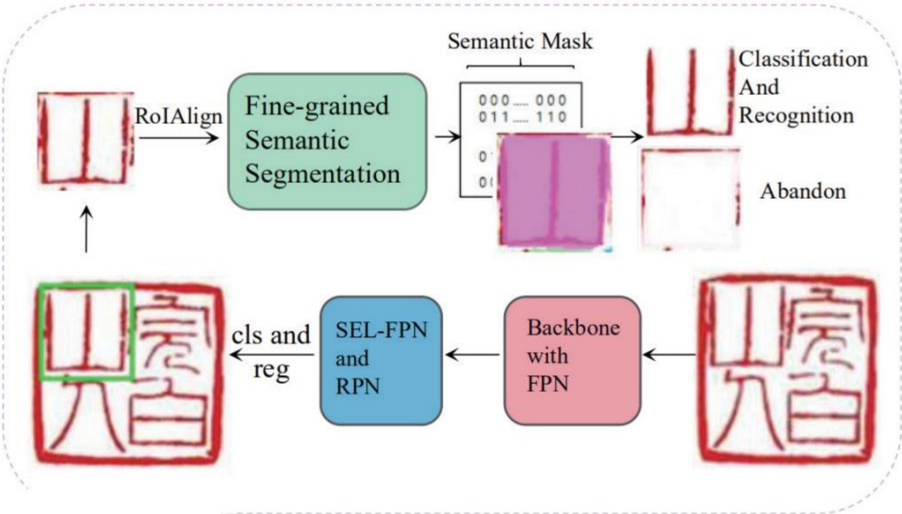

Figure 8: Overview of segmentation framework[10].

3.7. DB-EAC and LSTR

The main challenges of traditional seal text recognition are: text curves, irregular arrangement, occlusion, blur and other interferences. The DB-ECA model is a seal text area detection model based on DBnet. By adding the efficient channel attention module (ECA), it solves the hierarchical conflict problem in the feature pyramid and improves the detection effect of multi-scale targets. LSTR model: A lightweight seal text recognition model that combines convolutional neural network (CNN) and self-attention mechanism to extract character features and uses CTC loss function to align character sequences, thereby improving the recognition ability of blurred and deformed seal texts. Compared with the five models in the past three years, LSTR has the highest accuracy of 91.29%, and has the advantages of fewer parameters and faster running speed. However, the decoding method of the CTC loss function used in it has a certain degree of randomness, which will affect the alignment effect of features and labels [11]. Figure 9 shows the LSTR structure diagram. Table 1 summarizes the characteristics and advantages of different machine learning based text recognition algorithms.

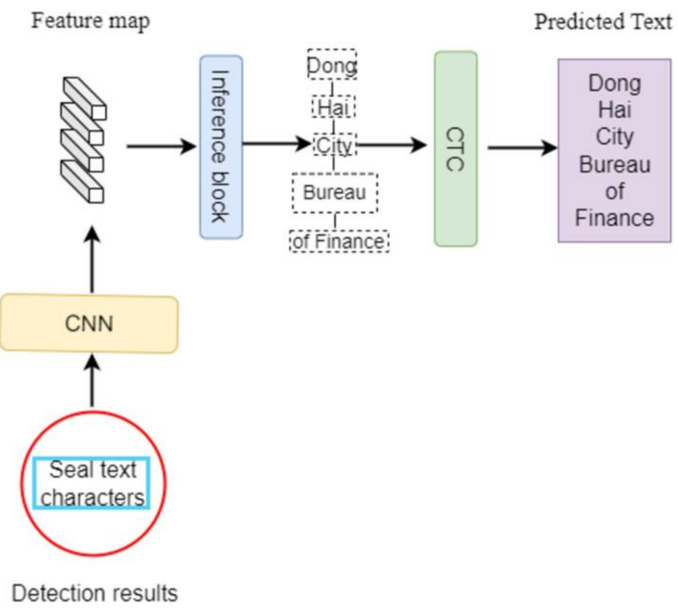

Figure 9: LSTR structure diagram[11].

Table 1: Text recognition algorithm based on machine learning.

Algorithm	Features	Disadvantages
MNSG	Before the algorithm learns, the digital samples are grouped	When the amount of data increases, the recognition results cannot be guaranteed
Root structure sequence learning	Realize small sample recognition of small seal characters	For unknown types of small seal characters, the accuracy still needs to be improved. The model is currently only applicable to printed small seal characters.
Pictographic radical sequence learning	Achieve zero-sample recognition of unseen small seal characters	The model is currently only applicable to printed small seal pictures
ASSDA	Algorithms for robust text image recognition	When the curve angle is too large, the algorithm will interfere with the capture of character-level information due to complex problems.
CSRS	Generate a large number of sealed images with different backgrounds for effective training	It runs slower and takes up more memory.
SEL-RefineMask	Can split the concatenated characters	The accuracy of character segmentation still needs to be improved
DB-EAC and LSTR	Small number of parameters and fast running speed	The method of decoding the CTC loss function has a certain degree of randomness

4. Algorithm Improvements

As mentioned above, there is a problem of sticking of small seal characters on seals. There is an optimized automatic segmentation algorithm for handwritten characters that has a better segmentation effect for sticky characters. This method is based on the principle of the dripping algorithm and

DOI: 10.54254/2755-2721/100/2025.18118

combines CFS (color filling segmentation) to perform preliminary segmentation; then the width of the continuous black pixel points of the segmented character is used to determine whether it is a sticky character. If it is a sticky character, the black pixel position is scanned between 0.2 times the width and 0.8 times the width of the segmented character image; the starting dripping point of the dripping algorithm is determined by combining the middle position of the segmented image to solve the problem of inaccurate positioning of the starting dripping point in special cases. If this algorithm can be used in the CSRS and LSTR algorithms, it can improve the recognition efficiency of small seal characters on seals [12]. Figure 10 shows the correct and incorrect starting and ending points.

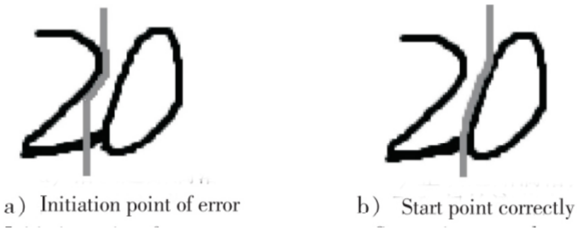

Figure 10: Correct and wrong start drop point[12].

However, since the seal characters on the seal pursue overall artistry, there will be some stroke adhesion in the horizontal direction, which is a challenge for the dripping algorithm suitable for vertical segmentation. In the future, researchers should consider how to add the dripping algorithm in both horizontal and vertical directions to improve the efficiency of seal character recognition. Figure 11 shows a stamp image with characters stuck together.

Figure 11: A stamp image with characters stuck together.

5. Conclusion

This paper summarizes the text recognition algorithms of small seal characters. By comparison, a variety of advanced small seal character recognition technologies and models are demonstrated, such as the recognition method of small seal characters in Wazhong based on neural networks and the zero-sample small seal character recognition method. From these methods, it can be seen that small seal characters have complex and variable font structures and scarcity of literature data sets. Deep learning has become the mainstream method for small seal character recognition. Since small seal characters are only a branch of seal script, future research should focus on the generalization of the model and pay attention to how to improve the recognition research of characters similar to small seal characters, such as large seal characters, and enable it to run efficiently on devices with limited computing resources, expanding the application scope of text recognition technology in low-resource environments such as mobile devices.

References

[1] Ma, J., Yan, Q., Ma, Z., et al. (2024). A review of Xixia character recognition methods. Journal of Huazhong University of Science and Technology (Natural Science Edition), 1-17. https://doi.org/10.13245/j.hust.241103.

[2] Zayene, O., Touj, S. M., Hennebert, J., Ingold, R., & Ben Amara, N. E. (2018). Multi-dimensional long short-term memory networks for artificial arabic text recognition in news video. IET Computer Vision, 12(5), 710-719. https://doi.org/10.1049/iet-cvi.2017.0468.

[3] Zia, N., Naeem, M. F., Raza, S. M. K., Khan, M. M., Ul-Hasan, A., & Shafait, F. (2022). A convolutional recursive deep architecture for unconstrained urdu handwriting recognition. Neural Computing & Applications, 34(2), 1635-1648. https://doi.org/10.1007/s00521-021-06498-2.

[4] Chen, X., Lu, L., Wang, T., et al. (2024). A review of text recognition methods based on visual language. Journal of Beijing Institute of Graphic Communication, 32(06): 35-43. DOI: 10.19461/j.cnki.1004-8626.2024.06.009.

[5] Zhou, Z. (2014). Recognition method of small seal characters in tiles based on neural network. Northwest University.

[6] Li, J. (2022). Small sample recognition of seal script based on Chinese character root structure sequence learning. Hangzhou Dianzi University. DOI:10.27075/d.cnki.ghzdc.2022.000593.

[7] Liu, J. (2023). Research on zero-sample small seal recognition based on pictographic feature consistency. Hangzhou Dianzi University. DOI:10.27075/d.cnki.ghzdc.2023.001716.

[8] Zhang, Y., Nie, S., Liang, S., & Liu, W. (2021). Robust text image recognition via adversarial sequence-to-sequence domain adaptation. IEEE Transactions on Image Processing, 30, 3922-3933. https://doi.org/10.1109/TIP.2021.3066903.

[9] Wang, Z., Lian, J., Song, C., Zheng, W., Yue, S., & Ji, S. (2019). CSRS: A chinese seal recognition system with multi-task learning and automatic background generation. IEEE Access, 7, 96628-96638. https://doi.org/10.1109/ACCESS.2019.2927396.

[10] Dun, Z., Chen, J., Qu, M., & Jiang, B. (2022). SEL-RefineMask: A seal segmentation and recognition neural network with SEL-FPN. JIPS(Journal of Information Processing Systems), 18(3), 411-427.

[11] Huang, B., Bai, A., Wu, Y., Yang, C., & Sun, H. (2024). DB-EAC and LSTR: DBnet based seal text detection and lightweight seal text recognition. PloS One, 19(5), e0301862-e0301862. https://doi.org/10.1371/journal.pone.0301862.

[12] Huang Y., Zheng J., Cao C. (2021). An optimized automatic segmentation algorithm for handwritten characters. Journal of Jimei University (Natural Science Edition), 26(02): 152-159. DOI: 10.19715/j.jmuzr.2021.02.10.