Creativity and Innovation
2025, VOL. 9, NO. 2, 61-65
DOI: 10.47297/wspciWSP2516-252712.20250902

Advances in Offline Handwritten Chinese Character Recognition : a Survey

Junhao Liu

School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China

ABSTRACT

Offline Chinese character recognition has played an indispensable role over the past decade. While it has been continuously integrated into daily life, the technology behind it has also been continuously iterated and updated. Aiming to analyse the important results of the current development of this technology as well as its development prospects, this essay will focus on the excellent results of the past development of Offline Handwritten Chinese Character Recognition (OHCCR) as well as the current status, and present an overview and insights of these technologies. Finally, the problems encountered in the research of this technology and the prospects are presented.

KEYWORDS

Offline Handwritten Chinese Character Recognition; Neural Network; Deep Learning; Feature Extraction

1 Introduction

The continuous development of AI and the lightness of its volume have brought AI-related technologies closer to the lives of the general public, including the OHCCR, in recent years. There is an increasing demand for the digitisation of data such as traditional paper documents, historical archives and handwritten notes, and how to accurately and efficiently convert handwritten Chinese characters into digital text has become one of the key technologies to promote digital transformation and intelligent information processing.

Offline Chinese handwriting recognition is used in many scenarios, including automatic recognition of documents in the office, automatic correction of answer content in education, and automatic analysis of patient medical records. This technology makes life easier and saves time when recognising text. Further research into this technology can improve the robustness and accuracy of Chinese character recognition results, and even extend the recognition range so that the model can recognise more rare characters and characters within dialects. In some application scenarios, digital transformation can increase document processing efficiency by 30 to 50 percent and significantly reduce manual input errors. Meanwhile, according to the IEEE Xplore database, the number of research papers related to Chinese character recognition has increased by about 300% over the past decade, reflecting the continued attention and investment in the field from both academia and industry.

Some of the related techniques have been summarised in previous studies by Shen et.al [1] and Ye-eum [2], and since then HCCR has undergone a number of developments and technological updates. The traditional approach to this technique is good at capturing local features, but does not perform well in low resolution environments and with different handwriting styles.When combined with the more popular convolutional neural networks (CNNs), the technique has shown higher accuracy, which is one of the reasons why this year's results are better than the traditional method. Some researchers have fused techniques such as classifiers from traditional methods with CNNs and have also achieved better results. The purpose of this paper is to review the main technical routes of offline Chinese handwriting recognition, and to analyse the research progress and application effects of various methods in terms of feature extraction, neural network models and hybrid strategies. At the same time, we will also discuss the shortcomings and limitations of the existing methods and give an outlook on the future research direction. By comparing the advantages and disadvantages of various techniques, we hope to provide a reference for subsequent research and promote the further development of the field.

2 An Overview of Offline Handwritten Chinese Character Recognition Techniques

This part of the paper intends to make an overall review on OHCCR, and describes and introduces key techniques, including traditional feature extraction methods, the most widely developed convolutional neural network combination methods, methods that combine several techniques, and more innovative methods in other fields that incorporate the latest state-of-the-art technology.

2.1 Feature Extraction-Based Approaches

Traditional methods for offline handwritten Chinese character recognition usually include three main steps:

preprocessing, feature extraction and classification. Firstly, in the pre-processing stage, the captured image needs to be binarised to convert the grey scale image to black and white to distinguish the foreground from the background, and at the same time, techniques such as median filtering and mean filtering are applied to remove the noise, and to unify the size and orientation of the characters through normalisation, centring and rotational normalisation, etc. In the case of continuous strokes or composite Chinese characters, segmentation will also be carried out to extract the individual strokes or the basic constituents.

The features in the feature extraction part are basically human-determined or designed, and they usually include structural features, such as the start and end points of the strokes of Chinese characters, radicals, etc., as well as statistical features, such as area density, edge distribution, and histograms of horizontal and vertical projections.

Early feature extraction efforts focused on reflecting local features of Chinese characters, especially strokes and radicals, such as the Gabor method [3], which was mostly used in the early days to simulate the response of simple cells in the human visual cortex to information of different orientations and frequencies, and thus to extract local texture features in an image. The features extracted by using Gabor filters for multiple directions at fixed sampling points can better reflect the texture features and local information of Chinese characters. Differently, Chang et.al [4] tried to use Euclidean Distance Transform (EDT) with Gradient Oriented Tracing to construct the skeleton of the characters and combined with curvature segmentation and other methods to extract the strokes.

The classifier needs to further classify the extracted features so that the final output can match the most original Chinese character, but due to the large individual differences in handwritten Chinese characters, the same character written by different people may have significant morphological differences.

One of the most widely used early classifiers is SVM [5], which optimises the feature representation at different scales to ensure maximum information retention, and nested subset classification based on Mahalanobis distance balances its classification and generalisation capabilities. Overall, this method significantly improves the offline recognition performance of handwritten Chinese characters through better feature representations and robust classification strategies, especially in cursive and ligatures, but on the other hand, this method lacks the ability to deal with large-scale information and shows obvious inefficiencies when faced with large-scale information. Another widely used classifier is MQDF [6], which achieves high-performance character recognition by splitting the high-dimensional feature vector into multiple low-dimensional parts and computing the modified Mahalanobis distance on each part independently, which not only reduces the computational cost but also improves the discriminative accuracy in the case of insufficient training samples.

With the introduction of a large number of deep learning methods, the vulnerability of traditional methods becomes more and more obvious: geometric features are unstable, feature extraction is more wired, and computational complexity is higher; these vulnerabilities have led to a sharp decline in the use of traditional methods in recent years, but there are still methods proposed. In 2016, for the complex structure problem of handwritten Chinese characters, Zhu et. al [7] proposed a "substructure string recognition" method, i. e., using the inherent composition laws of characters, decomposing Chinese characters into multiple substructures, automatically discovering the substructure patterns by density clustering, thus transforming the single character recognition problem into a substructure string recognition problem, and finally combining the substructure string recognition with the traditional method. This method makes full use of the structural prior of Chinese characters and solves the confusion problem between similar characters to some extent, but performs poorly on Chinese characters with complex structures. On the other hand, Figure 1. displays the process of Discriminative Quadratic Feature Learning (DQFL) [8]. This method extends the quadratic features by both statistical and spatial correlations based on traditional gradient direction features, which effectively improves the feature dimensions and discriminative ability, and then obtains the results by discriminative dimensionality reduction, which provides feature representations that are highly discriminative with low computational complexity, and provided a highly efficient and accurate recognition method at that time.

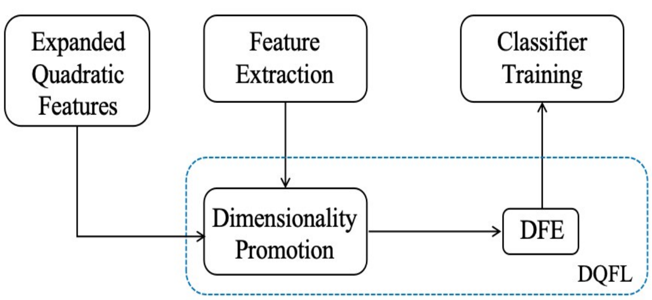

Figure1 The process of Dicriminative Quadratic Feature Learning

2.2 Neural Network-Based Approaches

The most popular methods in recent years are mostly deep learning models based on traditional neural networks, which usually have high recognition accuracy and strong expressiveness, but on the other hand consume more computational resources and have higher model complexity. Compared to the deliberately designed feature extraction method of traditional methods, deep learning models can learn multi-layered hierarchical feature representations directly from the original image, and some deep learning models themselves have multi-layered structures that can be jointly trained to learn feature representations in all modules, resulting in stronger overall expressiveness and robustness.

Previous work involving Convolutional Neural Networks (CNNs) has mostly taken advantage of their ability to extract features directly from the original image, classify the resulting features directly, and train the CNN model directly to achieve better performance. Cheng et. al [9] used two kinds of supervised signals to train deep convolutional neural networks, on the one hand, using traditional classification supervision, i. e. softmax and cross-entropy loss, and on the other hand, introducing a ternary-based loss function, and at the same time, the classification loss ensures the overall category differentiation and locally performs similarity ranking, makes the trained model have an advantage in the face of confusable fonts. In the same period, Yang et. al [10] dynamically updated the sampling probability of each sample according to the soft maximum output value through a preset quota updating function, which enabled the training process to pay more attention to the samples of Chinese characters that are difficult to recognise or confusing, and reduced the learning of the mastered samples, thus improving the learning efficiency and recognition performance, and alleviating the interference of data imbalance and noise on the training.

Subsequently, the focus for offline Chinese handwriting recognition gradually shifted to further optimising the network structure. The traditional CNN model uses softmax for classification, which requires a large number of labelled samples, but cannot satisfy the recognition of a large number of out-of-the-box Chinese characters. To solve this problem, Li et.al [11] proposed a method called Deep Matching Network, shown in Figure 2.It replaces the parameters of the softmax regression layer with the features extracted from the stencil image, and adopts a local softmax regression strategy to reduce the computation and memory; in this way, the model can learn the matching relationship between the stencil characters and the handwritten characters, so as to recognise the Chinese characters that are not present in the samples, thus showing better generalisation.

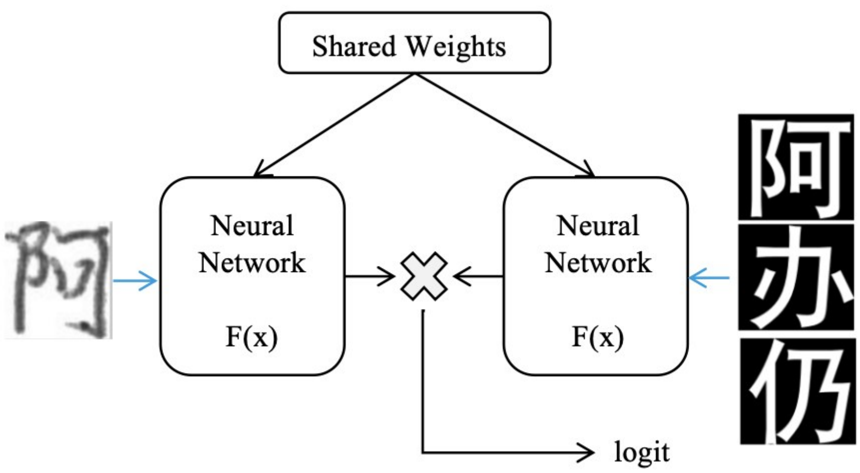

Figure 2 The Structure of Deep Matching Network

On the other hand, Aleskerova et. al [12] designed a hierarchical CNN that combines training samples with similar appearance features using two classifiers, divides all the samples into a number of clusters, and then classifies each cluster a second time using a second classifier, thus locally improving the recognition accuracy; this approach reduces the burden of processing large data, but is sensitive to the design of the parameters, and the case of a large number of clusters may result in a large burden on the computational resources of the second classifier.

There are also many excellent methods proposed in recent years.It's worth noting that in 2022, Gan et.al [13] proposed a new method, Pyramid Graph Transformer (PyGT), which improves the computational efficiency of the model by treating characters as skeleton graphs, combining graph convolution with the transformer, and ensures a comprehensive modelling of the geometric features of the characters through the graph attention mechanism, and performs well in the face of the change of writing order and the large scale of the The performance is excellent in the face of the change of writing order and the large scale of the character sets.In order to solve the problem of the difficulty of Chinese character feature extraction in complex scenarios, Chen et. al [14] proposed a Chinese character recognition scheme based on multidimensional representations, which recognises Chinese characters with only a small number of key features, and applies a character similarity algorithm to eliminate the feature noise, which improves the recognition accuracy in complex scenarios and has a strong recognition performance.

In general, nowadays, OHCCR in CNN-based development is more inclined to improve the model accuracy and robustness, transform the network structure to lighten the model, and at the same time improve the generalisation of the model to non-given samples, so as to better cope with the problem of similar Chinese characters and different writing styles.

3 Other Innovative Methods

For offline handwritten Chinese character recognition, traditional methods rely on manual feature extraction, which has average performance but considerable computational resources; while CNNs possess higher performance. Combining traditional techniques and CNN's not only improves the performance of handwritten Chinese character recognition system, but also reduces the demand for additional computational resources, which can effectively reduce the impact due to writing differences.In 2017, Zhang et.al [15] used directMap to decompose the stroke direction of Chinese characters into multiple feature maps, which were inputted into the CNN; in order to adapt to different writing styles, the introduction of an adaptation layer to reduce the difference between training and testing data and enhance the generalisation ability of the model. This method improves the recognition accuracy, reduces the dependence on data enhancement, and improves the computational efficiency, obtaining 97.37% recognition accuracy on the ICDAR 2013 dataset.

Some of the methods intend to use a different approach than CNN in automatically extracting features within the convolutional layer. Sun et.al [16] used the Mahalanobis distance of vector division as the feature separation method, and computed the covariance matrix of pre-separated low-dimensional vectors to approximate the Mahalanobis distance for feature matching, which is more effective than CNN's feature processing method for high level features, and also maintains a more efficient feature matching method under the limited computational resources and training samples. training samples also maintains higher recognition accuracy.

4 Experiments

This section describes the OHCCR evaluation approach with some of the commonly used data sets.

4.1 Evaluation Metrics

The vast majority of offline handwritten Chinese character recognition tasks measure the reliability of the method using the accuracy of the model on a particular dataset, i.e., Accuracy, which is calculated using the following formula, where NC stands for the number of correct samples and NT stands for the total number of samples.

$$\text{Accuracy} = \frac{NC}{NT}$$

In addition, some of the methods will also use Normalised Edit Distance (NED) as an evaluation metric to measure the overall ability of the model in long text recognition.A lower NED means that the model recognises the text with less error, and is able to better deal with noise and distortion in long text and complex backgrounds.

4.2 Related Database

The following are some of the datasets with their corresponding relevant data.

Dataset	Year	Total Sample	Content
Hit-MW[17]	1996	500,370	3,755 Chinese characters and multiple symbols
HCL2000[18]	2009	3,755,000	3,755 Chinese characters from 1,000 different writers
CASIA-HWDB 1.0-1.2[19]	2010	3,900,000	7,185 Chinese Characters and 171 symbols
ICDAR 2013[20]	2013	224,419	3,755 Chinese Characters from 300 different writers
IAHCC[21]	2017	449,268	3,811 Chinese Characters ,52 English letters and 10 numbers
IHCCD[22]	2024	112,650	3,755 Chinese Characters

5 Conclusion

This paper provides a comprehensive overview of the evolution of OHCCR, examining both traditional feature extraction methods and modern deep learning techniques. It highlights how traditional methods, despite their effectiveness in certain conditions, face challenges in handling variations in handwriting styles and low-resolution images. In contrast, the introduction of CNNs and hybrid models has considerably improved recognition accuracy and robustness. However, several challenges remain. One of the primary hurdles is the inherent variability in handwriting styles, Another major challenge is the limited ability to handle complex character structures. To address these issues requires ongoing research into more advanced models and feature extraction techniques.

References

[1] Shen L, Chen B, Wei J, et al. The challenges of recognizing offline handwritten Chinese: a technical review[J]. Applied Sciences, 2023, 13(6): 3500.

[2] Ye-eun K. A Review of the Current Status of AI Research in Handwritten Chinese Character Recognition[C]//SHS Web of Conferences. EDP Sciences, 2024, 185: 01004.

[3] Hamamoto Y, Uchimura S, Masamizu K, et al. Recognition of handprinted Chinese characters using Gabor features[C]//Proceedings of 3rd international conference on document analysis and recognition. IEEE, 1995, 2: 819-823.

[4] Chang H H, Yan H. Analysis of stroke structures of handwritten Chinese characters[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 1999, 29(1): 47-61.

[5] Zhang J, Ding X, Liu C. Multi-scale feature extraction and nested-subset classifier design for high accuracy handwritten character recognition [C]//Proceedings 15th International Conference on Pattern Recognition. ICPR-2000. IEEE, 2000, 2: 581-584.

[6] Sun F, Omachi S, Kato N, et al. Fast and precise discriminant function considering correlations of elements of feature vectors and its application to character recognition[J]. Systems and Computers in Japan, 1999, 30(14): 33-42.

[7] Zhu Y, An X, Zhang K. A handwritten Chinese character recognition method combining sub-structure recognition[C]//2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR). IEEE, 2016: 518-523.

[8] Zhou M K, Zhang X Y, Yin F, et al. Discriminative quadratic feature learning for handwritten Chinese character recognition[J]. Pattern Recognition, 2016, 49: 7-18.

[9] Cheng C, Zhang X Y, Shao X H, et al. Handwritten Chinese character recognition by joint classification and similarity ranking[C]//2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR). IEEE, 2016: 507-511.

[10] Yang W, Jin L, Tao D, et al. DropSample: A new training method to enhance deep convolutional neural networks for large-scale unconstrained handwritten Chinese character recognition[J]. Pattern Recognition, 2016, 58: 190-203.

[11] Li Z, Wu Q, Xiao Y, et al. Deep matching network for handwritten Chinese character recognition[J]. Pattern Recognition, 2020, 107: 107471.

[12] Aleskerova N, Zhuravlev A. Handwritten Chinese characters recognition using two-stage hierarchical convolutional neural network[C]//2020 17th international conference on frontiers in handwriting recognition (ICFHR). IEEE, 2020: 343-348.

[13] Gan J, Chen Y, Hu B, et al. Characters as graphs: Interpretable handwritten Chinese character recognition via Pyramid Graph Transformer[J]. Pattern Recognition, 2023, 137: 109317.

[14] Cheng Chen, Ming Jiang, Min Zhang. A Chinese Character Recognition Solution Based on Multidimensional Representation [J]. Software Engineering, 2024, 27 (08): 24-29. DOI: 10.19644/j.cnki.issn2096-1472.2024.008.006.

[15] Zhang X Y, Bengio Y, Liu C L. Online and offline handwritten Chinese character recognition: A comprehensive study and new benchmark[J]. Pattern Recognition, 2017, 61: 348-360.

[16] Sun F, Omachi S, Kato N, et al. Fast and precise discriminant function considering correlations of elements of feature vectors and its application to character recognition[J]. Systems and Computers in Japan, 1999, 30(14): 33-42.

[17] Su T, Zhang T, Guan D. Corpus-based HIT-MW database for offline recognition of general-purpose Chinese handwritten text[J]. International Journal of Document Analysis and Recognition (IJDAR), 2007, 10: 27-38.

[18] Zhang H, Guo J, Chen G, et al. HCL2000-A large-scale handwritten Chinese character database for handwritten character recognition[C]// 2009 10th international conference on document analysis and recognition. IEEE, 2009: 286-290.

[19] Liu C L, Yin F, Wang D H, et al. CASIA online and offline Chinese handwriting databases[C]//2011 international conference on document analysis and recognition. IEEE, 2011: 37-41.

[20] Yin F, Wang Q F, Zhang X Y, et al. ICDAR 2013 Chinese handwriting recognition competition[C]//2013 12th international conference on document analysis and recognition. IEEE, 2013: 1464-1470.

[21] Ren H, Wang W, Lu K, et al. An end-to-end recognizer for in-air handwritten Chinese characters based on a new recurrent neural networks [C]//2017 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2017: 841-846.

[22] Ji J M,Shao Y X and Ji T Z. 2024. IHCCD:dataset for identification of irregular handwritten Chinese characters. Journal of Image and Graphics, 29(11):3345-3356