MDPI

Review

The Challenges of Recognizing Offline Handwritten Chinese: A Technical Review

Lu Shen [1,*], Bidong Chen [1], Jianjing Wei [1], Hui Xu [1], Su-Kit Tang [1,*] and Silvia Mirri [2]

1   Faculty of Applied Sciences, Macao Polytechnic University, Macao SAR 999078, China
2   Department of Computer Science and Engineering, University of Bologna, 40126 Bologna, Italy
*   Correspondence: lu.shen@mpu.edu.mo (L.S.); sktang@mpu.edu.mo (S.-K.T.)

Abstract: Offline handwritten Chinese recognition is an important research area of pattern recognition, including offline handwritten Chinese character recognition (offline HCCR) and offline handwritten Chinese text recognition (offline HCTR), which are closely related to daily life. With new deep learning techniques and the combination with other domain knowledge, offline handwritten Chinese recognition has gained breakthroughs in methods and performance in recent years. However, there have yet to be articles that provide a technical review of this field since 2016. In light of this, this paper reviews the research progress and challenges of offline handwritten Chinese recognition based on traditional techniques, deep learning methods, methods combining deep learning with traditional techniques, and knowledge from other areas from 2016 to 2022. Firstly, it introduces the research background and status of handwritten Chinese recognition, standard datasets, and evaluation metrics. Secondly, a comprehensive summary and analysis of offline HCCR and offline HCTR approaches during the last seven years is provided, along with an explanation of their concepts, specifics, and performances. Finally, the main research problems in this field over the past few years are presented. The challenges still exist in offline handwritten Chinese recognition are discussed, aiming to inspire future research work.

Keywords: handwritten Chinese recognition; deep learning; convolutional neural network; recurrent neural network; long short-term memory; connectionist temporal classification

1. Introduction

Character recognition converts text information on paper into computer text through electronic equipment. It has many applications in office automation, building electronic databases, machine translation, intelligent paper marking, etc. With the development of information technology, people have become accustomed to relying on computers to complete text information processing tasks. However, due to the massive amount of text information generated daily, the standard manual input of text information has been unable to meet the needs of people to obtain information quickly. Therefore, how to obtain text information speedily and recognize it accurately through computer technology has become an important research direction in pattern recognition.

The Chinese recognition can be divided into printed Chinese recognition and handwritten Chinese recognition according to the recognition object [1,2]. According to the different data collection methods, it can be subdivided into online handwritten Chinese recognition and offline handwritten Chinese recognition. The classification of Chinese recognition and the scope of our study is shown in Figure 1. Printed Chinese recognition is the recognition of printed paper material. The recognized Chinese is neat and clear, which can meet the requirements of Optical Character Recognition (OCR) [3]. However, the OCR technology has poor robustness for correcting and recognizing uneven illumination, blur, font deformation, and tilted text, which cannot meet the requirements of handwritten text recognition with diverse styles. Online handwritten Chinese recognition mainly identifies

the Chinese written on the electronic screen, such as the handwritten input of mobile phones, tablets, and other electronic devices. It records the timing information of Chinese writing and the relative position information of strokes [4]. Online handwritten Chinese recognition technology developed at the beginning of the twenty-first century with the rise of the electronic information industry. The technology and products related to printed Chinese recognition and online handwritten Chinese recognition have already matured. In addition, the object processed by offline handwritten Chinese recognition is the 2D picture of offline handwritten Chinese captured by a scanner or camera, which loses the timing information. Generally speaking, offline handwritten Chinese recognition is more complex than online handwritten Chinese recognition. At present, its recognition rate still cannot reach satisfactory results. Therefore, the recognition of offline handwritten Chinese remains the mainstream research direction.
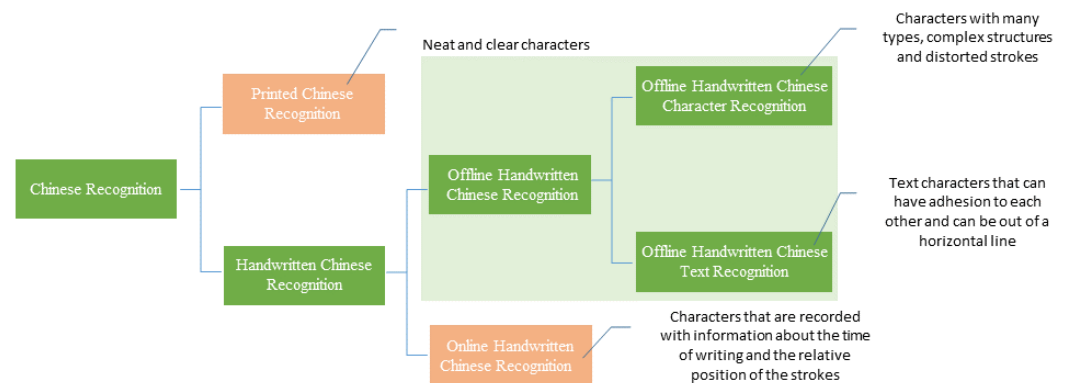

Figure 1. Classification of Chinese recognition. The green rectangles from left to right depict the research directions of this paper from broad to specific. The orange rectangles represent relevant research directions but are not described in detail in this paper. The light blue area is the main work of this paper.

In offline handwritten Chinese recognition, depending on the number of Chinese characters contained in the input image, i.e., whether the task is to recognize single Chinese characters or lines of Chinese characters, Offline Handwritten Chinese Recognition can be further divided into Offline Handwritten Chinese Character Recognition (Offline HCCR) and Offline Handwritten Chinese Text Recognition (Offline HCTR). In the HCCR task, the grayscale image of a single handwritten Chinese character is studied and assigned to a corresponding category based on the information in the picture. A line or a paragraph of characters is included for text recognition. The entire section is recognized based on extracting individual characters in conjunction with contextual information. As shown in Figure 2, handwritten Chinese characters have more deformations in strokes than traditional printed Chinese characters and are less scalable for traditional algorithms such as template matching. Therefore, recognizing handwritten Chinese characters (HCC) and handwritten Chinese text (HCT) is currently more challenging.

Specifically, the difficulties of offline handwritten Chinese recognition are mainly manifested in: (1) Compared with other languages, there are more types of Chinese characters. According to the criteria specified in the National Standard Simplified Chinese Character Set of the People's Republic of China (GB2312-80), there are 6763 categories of commonly used first-level and second-level Chinese characters [5]. However, it still falls short of fulfilling the demands for Chinese information processing, and the subsequent GBK and GB18030 character sets have been supplemented to address further requirements. Since each individual character in Chinese is a different class, the number of Chinese character categories in the recognition task is enormous compared to languages such as English, where only a few dozen classes of letters need to be identified. This makes the complexity of the algorithm required to recognize Chinese higher; (2) Chinese characters have intricate structures. The structure can be divided into upper and lower, left and right,

semi-enveloped, and mosaic structures. The average number of strokes in traditional and simplified Chinese characters is 16 and 11, respectively, and the recognition rate is reduced due to excessive strokes. (3) There are many similar Chinese characters. These characters almost have the same spatial structure and have comparable stroke patterns, such as "己" and "已", "乌" and "鸟", "戊" and "戌", etc. The recognition of similar characters in the feature extraction process will have high requirements for the algorithm. (4) Different people have distinct writing styles. The handwriting of various persons varies widely, and even the script of the same person may differ in different states. At the same time, the strokes of some Chinese characters are easily deformed in writing, which will significantly increase the difficulty of recognition. Based on these problems, although offline handwritten Chinese recognition has been studied for decades, the recognition effect still does not reach people's expectations. It is still a very challenging pattern recognition problem. This paper will review the offline handwritten Chinese recognition technology to summarize and analyze the latest academic and technical progress.
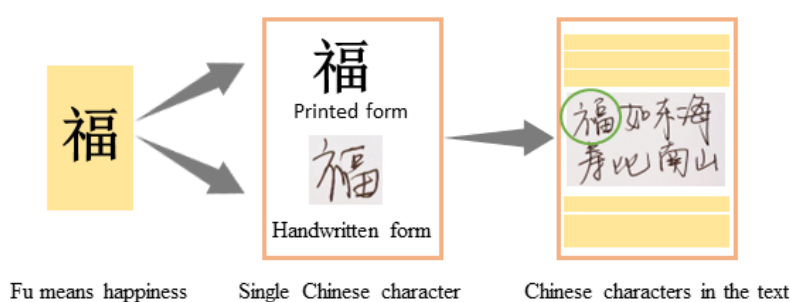

Figure 2. Comparison between printed Chinese and handwritten Chinese.

By searching for handwritten character recognition-related reviews, we found that these papers outline recognition methods for different languages separately and do not systematically describe the recognition of Chinese, e.g., literature [6–8]. In 2016, Ref. [9] provided a detailed analysis and summary of handwritten Chinese recognition methods. However, unfortunately, it was written in Chinese seven years ago and could not summarize the latest methodological techniques. After this paper, finding another summary article on offline handwritten Chinese recognition is difficult. Therefore, this paper reviews the recognition techniques for offline HCC and offline HCT, summarizes and analyzes the latest academic progress in this field since 2016, composes the main research issues, and suggests future works for the challenges that currently face handwritten Chinese recognition. The main contributions are summarized as follows:

- We review the technical advancements since 2016 based on the different approaches and research problems of HCCR and HCTR, meeting the demand for an up-to-date technological survey in this field as an essential research topic of pattern recognition;
- In order to serve as a reference for experimental techniques and recognition effects of future research, we count the number of publications on various research methods used in HCCR and HCTR over the last seven years and demonstrate the best recognition accuracy on the same publicly available dataset for each year;
- For researchers to better grasp the study trends of various issues in handwritten Chinese recognition in recent years, we compile the main research concerns in this field since 2016 and visualize them according to their study hotness;
- We present the existing challenges in recognizing handwritten Chinese to provide inspiration and suggestions for future research work.

The rest of this paper is organized as follows: Section 2 introduces common public standard datasets and evaluation metrics for handwritten Chinese recognition. Section 3 summarizes the traditional machine learning methods, deep learning methods, traditional techniques combined with deep learning, and some new methods with other knowledge for offline HCCR, respectively. Section 4 introduces the recognition methods for offline

HCTR. In Section 5, we discuss and analyze the current research work and present some prospects for future work. Finally, we conclude the whole paper in Section 6.

2. Databases and Model Evaluation Metrics

This chapter focuses on the HCC and HCT datasets separately in chronological order. We summarize the earlier datasets and detail the datasets that are commonly used for experiments and the recently released large datasets. We also introduce model evaluation metrics in handwritten Chinese recognition to facilitate the reader's understanding of the comparison of different approaches that will be presented in the following chapters.

2.1. Standard Databases

Data play a critical role in pattern recognition, and a high-quality dataset can improve the effectiveness of model training and the accuracy of prediction results. Table 1 summarizes the commonly used publicly available datasets for offline handwritten Chinese recognition and gives detailed descriptions. To facilitate the study, the corresponding download links are provided in Table 2.

Table 1. Introduction of some commonly used handwritten Chinese datasets.

Dataset	Year	Total Sample	Content	Writers	Description
HIT-MW [10]	1996	500,370	3755 Chinese characters and 94 common symbols, punctuation, English letters, etc.	130	The samples written by 50 people are segmented, and the samples written by 80 people are not segmented.
SCUT-IRAC [11]	2006	186,444	8664 lines of text (including letters, punctuation and Chinese characters), each sample has 10.16 text lines, each text line has 21.51 characters, each sample includes 218.57 characters	780	The samples are written without a standard written horizontal line reference, making them suitable for conducting experiments with Chinese text line segmentation; text data distribution equilibrium; data with miswriting and erasing.
CASIA-HWDB 1.0-1.2 [12]	2010	3,900,000	7356 classes (7185 Chinese characters and 171 symbols)	1020	At the character level, all of the data have been segmented and annotated, and each dataset has been partitioned into standard training and test subsets.
CASIA-HWDB 2.0-2.2 [12]	2010	1,350,000	5090 pages	1020	At the character level, all of the data has been segmented and annotated, and each dataset has been partitioned into standard training and test subsets.
SCUT-COUCH [13]	2011	more than 3,600,000	5401 Big5 traditional Chinese characters, 1384 traditional Chinese characters corresponding to GB2312-80 level 1 font, 44,208 phrases, 2010 Chinese pinyin, 184 symbols (including letters, numbers, and common symbols), 8809 lines of online text	more than 190	Database includes Chinese pinyin, phrases, and symbols in a variety of styles and sources.
ICDAR 2013 [14]	2013	224,419	3432 text image samples	60	Each image has the word's bounding box label and text content.

Table 2. Download links for some popular handwritten Chinese datasets.

Dataset	Link
HIT-MW [10]	http://www.hcii-lab.net/data/hcclib/hcclib_chn.htm, (accessed on 2 March 2023)
SCUT-IRAC [11]	https://code.google.com/archive/p/hit-mw-database/ (accessed on 2 March 2023),
CASIA-HWDB 1.0-1.2 [12]	http://www.nlpr.ia.ac.cn/databases/handwriting/Download.html, (accessed on 2 March 2023)
CASIA-HWDB 2.0-2.2 [12]	http://www.nlpr.ia.ac.cn/databases/handwriting/Download.html, (accessed on 2 March 2023)
SCUT-COUCH [13]	http://www.hcii-lab.net/data/SCUTCOUCH/CN/download.html, (accessed on 2 March 2023)
ICDAR 2013 [14]	https://rrc.cvc.uab.es/?ch=2&com=downloads, (accessed on 2 March 2023)

The first offline HCT dataset is HIT-MW [10], established by the Harbin Institute of Technology in 2006. It contains a total of 853 handwritten text images written by 780 people and 186,444 Chinese character samples. The corpus data were obtained from reports published by the People's Daily. The Chinese characters in this dataset were written in various styles, and the images contain scribbles and skewed lines of text. Therefore, this database is mainly used for document segmentation, HCCR, and HCTR tasks. All images in the database are stored as binary images for researchers to use.

The CASIA-HWDB dataset [12], subsequently collected by the Institute of Automation of the Chinese Academy of Sciences, is the primary publicly available and widely used dataset in research on offline handwritten Chinese recognition. The CASIA-HWDB dataset contains six sub-datasets, divided into the character sets HWDB1.0–1.2 and text sets HWDB2.0–2.2. Character sets 1.0–1.1 contain commonly used Chinese characters, and 1.2 is a non-useful Chinese character dataset. The three datasets contain 7356 characters (7185 Chinese characters and 171 other symbols) in nearly 3.9 million images. The text dataset is 52,230 images of text lines containing 2703 classes of characters, which are segmented by the line from 5090 texts containing 1,350,000 characters and divided into a training set of 41,781 images and a test set of 10,449 images. The contents of the text are mostly news reports on the Internet and a part of ancient poems, and the text dataset also provides detailed inter-character segmentation annotation information. Since the test set in CASIA-HWDB is relatively close to the training set in font style and text content, most experiments use the ICDAR 2013 [14] competition set as the test set to judge the performance of the algorithm. The dataset is written by 60 people and contains 3432 text image samples. In addition, the SCUT-COUCH dataset [13] can also be used for HCCR and HCTR tasks. It is a handwritten dataset covering single words, phrases, text lines, numbers, letters, and symbols published by the South China University of Technology.

Creating a dataset requires a significant investment of human and material resources, and collecting and annotating a large number of samples are challenging tasks. Therefore, there are few research efforts to create high-quality datasets with sample diversity and sufficient sample size. Two large handwritten datasets have been released in recent years, the HCC dataset HITHCD-2018 [15], collected by the Harbin Institute of Technology in 2019, and the HCT dataset SCUT-HCCDoc [16], collected by South China University of Technology in 2020. The HITHCD-2018 dataset includes 21,000 Chinese character categories and 20 million Chinese character images, significantly expanding the scope of the previous HCC dataset and covering the entire list of GBK character set specifications. The SCUT-HCCDoc dataset contains 12,253 images of documents taken by the camera, with 116,629 lines of text and 1,155,801 characters. The data in this dataset are unconstrained text data due to the different conditions of camera capture, different scenes of text sources, differences in the length and orientation of text lines, etc. This dataset is also more diverse.

2.2. Performance Evaluation Methodology

Among the performance evaluation methods of offline HCCR, the model's performance is mainly measured based on the label accuracy of the model on the test set. In the testing process, the accuracy of the samples is calculated as shown in Equation (1), where N_C represents the number of correctly classified samples, and N_T represents the sum of all the numbers to be classified:

$$\text{Accuracy} = N_C / N_T \tag{1}$$

In the HCTR experiment, the recognition results of text lines can be evaluated by three metrics: Accurate Rate (*AR*), Correct Rate (*CR*), and Character Error Rate (CER). The three metrics are defined as:

$$AR = (N_t - N_d - N_s - N_i)/N_t \tag{2}$$

$$CR = (N_t - N_d - N_s)/N_t \tag{3}$$

$$\text{CER} = (N_s + N_i + N_d)/N_t \tag{4}$$

where N_t is the total number of character samples in the evaluation set. N_s, N_i, and N_d denote the total number of substitution, insertion, and deletion errors, respectively.

3. Handwritten Chinese Character Recognition

In this chapter, we present the progress of recognizing offline handwritten Chinese characters using traditional techniques, deep learning methods, methods combining traditional techniques with deep learning, and novel methods combining other domains in the last seven years, respectively, and analyze and compare the performance of these methods.

3.1. Recognition Methods Based on Traditional Technology

The first to break the barrier of the HCCR task was a recognition model for handwritten Japanese and Chinese characters designed by Fuji et al. in 1981 [17]. Chinese researchers started to study HCCR in depth after 1980, and in late 1989, Tsinghua University successfully developed a system that could recognize 3755 frequently used Chinese characters on an enormous scale [18]. A few years later, in 1997, the THOCR-97 integrated Chinese character recognition system, also designed by Tsinghua University, achieved an accuracy rate of 95.8% while limiting the neatness of the writing [19].

The traditional HCCR method is mainly divided into three steps: pre-processing the image, extracting the features with distinguishing ability, and recognizing according to specific classification rules. The recognition process and methods commonly used in each step are shown in Figure 3.
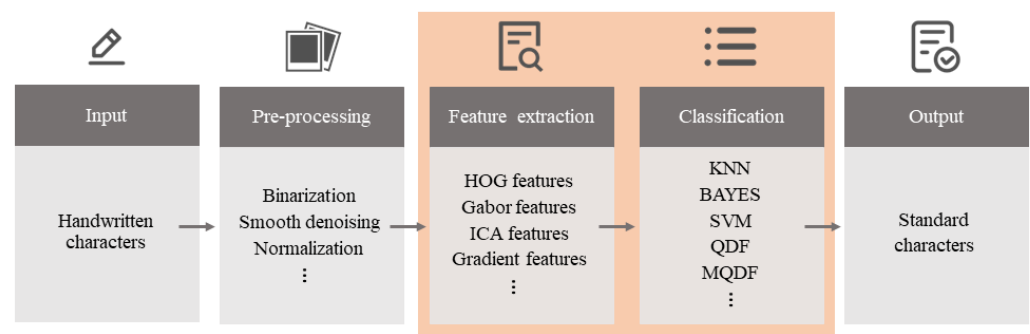

Figure 3. Traditional offline HCCR process and typical methods for each step. The orange rectangle shows some features extracted with traditional techniques and a few commonly used machine learning classifiers.

Image pre-processing directly affects the effect of subsequent feature extraction and recognition classification. The information on handwritten Chinese characters in real life is not ideally standardized, neat, and transparent. Therefore, we need to discard information irrelevant to recognition when we perform effective information extraction. Typical pre-processing operations include binarization, smooth denoising, and sample normalization. The image binarization process transforms the whole character image into only black and white. Since the background of the handwritten Chinese character image is relatively simple, this method can clearly extract the target character from the background and noise information. Normalization operation is to normalize images of different scales to the same scale without changing the aspect ratio to eliminate the effect of scale differences.

Feature extraction is a critical step in traditional Chinese character recognition. One of the mainstream methods is based on structural features, which extract features mainly by analyzing Chinese characters' composition structure, strokes, and radicals. The other one is based on statistical features. Due to structural feature extraction's difficulty and noise sensitivity, statistical-based feature extraction methods show better performance. For example, HOG features [20], Gabor features [21], ICA features [22], and Gradient features [23] have obtained high recognition rates in the HCCR task.

In terms of classifier selection, the commonly used models are K-Nearest Neighbor (KNN) [24], Bayes Classifier [25], Support Vector Machine (SVM) [26], Quadratic Discriminant Function (QDF) [27], and Modified Quadratic Discriminant Function (MQDF) [28]. Among them, SVM is a fast and efficient method for classifying small sample sets. SVM has strong robustness and generalization ability, but it is more sensitive to missing data, and its efficiency decays faster when there are too many data samples. The classification efficiency of the naive Bayesian classifier is more stable, more tolerant to missing data, and more robust, but the requirement of data independence is very high, and the sample data often do not meet such requirements in practice. Hence, the classification prediction effect is not very satisfactory. In contrast, the MQDF classifier based on Gradient features has achieved better results in recognition tasks and has been widely used in practice.

With the rise of deep learning, little research has been carried out in the past seven years to do offline HCCR tasks based on the traditional "pre-processing + feature extraction + classification" framework. In 2016, Zhu et al. [29] proposed a method to recognize handwritten Chinese characters by combining the substructures of Chinese characters. They used a density-based clustering method to obtain substructures and converted the problem of recognizing a single character consisting of multiple substructures into the recognition of substructure strings, with the final extracted features fed into the MQDF classifier. It is worth noting that Zhou et al. [30] proposed a feature learning method used to do HCCR experiments in the same year. The method is called discriminative quadratic feature learning (DQFL), which mainly uses the statistical and spatial correlation of features to increase the dimensionality of features, and then discriminative feature extraction (DFE) is used to reduce the dimensionality. The combination of dimensionality enhancement and dimensionality reduction makes the feature representation more discriminative and nonlinear, thus significantly improving the recognition accuracy. Both methods above use MQDF to complete the classification task at the end, but it has a high storage requirement for MQDF. In 2018, Wei et al. [31] used sparse coding to compact the parameters of MQDF, constructed two compact MQDF classifiers with maximum likelihood-based and K-SVD methods, and recognized unconstrained handwritten Chinese characters efficiently. These methods are summarized and compared in Table 3. In 2021, Ma et al. [32] attempted to use the proposed efficient 2D-GMM-HMM based on the Kaldi toolkit for HCCR experiments. This system is not only an improvement of the 1D-GMM-HMM system in recognition accuracy but also differs from the 1D-GMM-HMM, which does a one-dimensional alignment. 2D-GMM-HMM can segment Chinese characters into basic components in horizontal and vertical directions according to hidden states.

Table 3. Summary and comparison of traditional methods for recognizing handwritten Chinese characters since 2016. "n/a" indicates no mention.

Method	MDQFL [30]	Combined Sub-Structure [29]	Compact MQDF [31]
Year	2016	2016	2018
Experimental platform	server with 2 Intel 2.6 GHz CPUs and 4 NVIDIA Tesla C2075 GPUs	n/a	a PC with an Intel(R) Core(TM) i7-2600 3.4 GHz CPU and 4 GB memory
Feature	gradient direction histogram (GDH)	weighted direction code histogram	gradient element
Feature processing	quadratic correlation (ascending dimension), discriminative feature extraction (descending dimension)	LDA (descending dimension)	normalization-cooperated gradient feature, Gaussian blurring, sparse coding
Classifier	DLQDF	MQDF	MQDF-KSVD
Dataset	ICDAR-2013	CASIA	CASIA-HWDB1.1
Evaluation metric	accuracy	character recognition error rate	accuracy
Value	94.92	2.57	90.08
Advantages	Improve classification accuracy; Test is faster.	Sub-structure discovery of Chinese characters; Determining the presence of similar characters to improve recognition rate.	Compact classifiers

The traditional HCCR framework has not received much attention in recent years, but this does not mean that the traditional techniques are obsolete. Some research works still show that traditional feature extraction methods and classifiers can further improve the performance of widely used neural networks in this field. These approaches we will present in Section 3.3.

3.2. Deep Learning Based Recognition Methods

Deep learning is a machine learning method based on traditional neural networks, designed to simulate the human brain's learning process with a multi-level structure and strong abstract learning capability. In contrast to feature extraction using traditional techniques, the deep learning-based approach can usually obtain the specific feature information directly from the original input image. This also enables deep learning-based approaches to provide end-to-end offline handwritten Chinese recognition solutions. The process is shown in Figure 4. From the perspective of network structure, the typical network models currently used for recognizing HCC are deep convolutional neural network (DCNN), ResNet, GoogleNet, deep residual network (DRN), etc. The current research on HCCR focuses on three issues: improving recognition accuracy, creating fast and compact models, and experimenting with other machine learning methods, such as transfer learning. We will also present the applications of deep learning methods in offline HCCR since 2016 from these three aspects and analyze and summarize their performance.
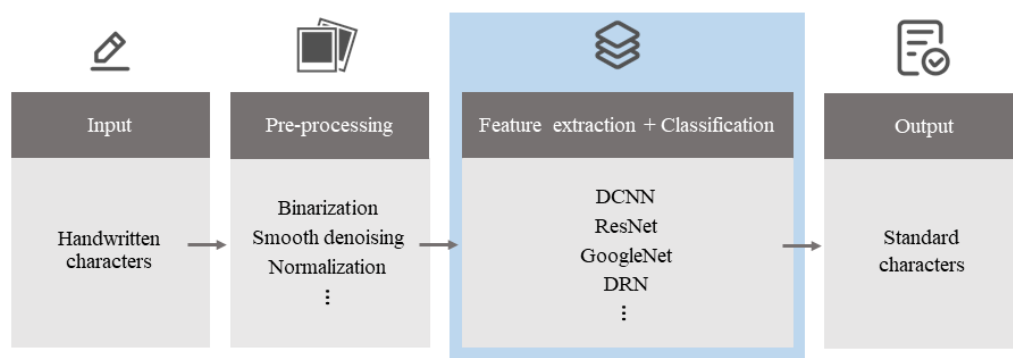
Figure 4. The process of recognizing offline HCCs using deep learning-based approaches. The blue rectangle shows some typical deep learning models.

3.2.1. Improve the Accuracy of Recognition

In recent years, researchers have focused on improving the performance of HCCR methods by optimizing the network model, combining structural information of Chinese characters, and expanding the dataset.

In 2016, Feng et al. [33] studied the correlation between the area size of Chinese character strokes and the receptive fields, hyperparameters, and the total number of feature maps in the convolution network and proposed a method for selecting the receptive field size of handwritten Chinese character recognition. The theoretical and experimental results have significant reference values for the rational and effective selection of the receptive field size of the DCNN model in HCCR application. Wu et al. [34] tried multi-stage feature extraction to recognize HCC. The experiments concluded that the recognition effect of multi-stage feature extraction is better than that of single-stage feature extraction and also summarized the suitable position of multi-stage feature extraction in the network. In 2017, Yang et al. [35] proposed an iterative refinement module implemented by an attention-based recurrent neural network, which focuses on improving the current prediction by updating attention with previous predictions to improve the classification accuracy of similar characters further. In the same year, Maidana et al. [36] explored 18 models, including commonly used convolutional neural network (CNN) structures and their fused models with SVM, to identify 200 classes of HCC in ICDAR 2013. ZFNet, a single network adapted as AlexNet, achieved the highest recognition accuracy of 98.2%. The subsequent fusion to SVM did not significantly improve the performance of most of the individual networks in the experiment but provides a good reference for fusion methods that aggregate the performance of two or more networks. Refs. [37,38] are innovative in the data preprocessing stage. Zhong et al. [37] use the spatial transformer network (STN) [39] to learn directly from the data to achieve the normalization of character shapes. Zhuang et al. [38] preprocess characters by median filtering images to achieve data smoothing and noise reduction.

Refs. [40,41] are improvements to the GoogLeNet network to identify HCC. In 2019, Bi et al. [40] changed the input and output layers of the original structure of GoogLeNet, then added layers for batch normalization, fragmented convolutional layers, and finally performed a profound transformation of the network. The improved model achieved 98.2% recognition accuracy in the CASIA-HWDB1.1 dataset. Later, Min et al. [41] proposed a shallow GoogLeNet network that maintains the depth of the initial structure while reducing the number of training parameters while targeting the misidentification problem of HCC by recognizing similar character sets again. In 2020, Aleskerova et al. [42] designed a two-stage hierarchical convolutional neural network to identify HCC in response to the problems of speed and accuracy of classification of large categories such as HCCR running on, for example, a CPU or the like. The first stage of the method can identify and distinguish different subsets of data, and the second stage network is trained for the classification problem within the

corresponding set. The recognition accuracy is not particularly outstanding due to its simple architecture.

In addition to the approaches mentioned above to improve the network structure, several studies have experimented with different loss functions. In 2016, Cheng et al. [43] used classification and similarity ranking to maximize inter-class differences and minimize intra-class differences. They used softmax and conditional log-likelihood loss as the model's loss function, which is effective in the HCCR recognition task. Similarly, to capture more inter-class and intra-class information, Zhang et al. [44] proposed a central loss-based metric learning character recognition algorithm in 2017, which combines metric learning and the ResNet network to achieve 97.03% accuracy on the ICDAR 2013 dataset. In addition, using a modified ResNet network to identify HCC is Chen et al. [45], who proposed a discriminant weighting method for cross-entropy loss calculation to handle recognition errors in the training phase in 2019. The method achieves the highest known accuracy of 98.79% on the ICDAR 2013 dataset when using a sparse training technique and the specific condition of utilizing the testing mini-batch mean and variance for batch normalization. Zeng et al. [46] combined triplet loss and softmax with cross-entropy loss as the loss function of the CNN model for local discriminant training of HCCR. Furthermore, they used the Conditional Random Field (CRF) to achieve global optimization. In addition, Xiao et al. [47] proposed two new loss functions to accomplish the HCCR task in 2019. They created the character template to address the inherent similarity between Chinese characters. Depending on the complexity of the classification, instance loss can reduce category differences while severely penalizing outlier instances of handwritten Chinese characters.

Other scholars have worked on improving the accuracy of HCCR in terms of the specific structure of Chinese characters, the handwriting styles of different people, and the expansion of the dataset. In 2017, Luo et al. [48] added local features of Chinese characters to the traditional CNN model. This multi-supervised training method learns both global and regional features of HCC and achieves a then state-of-the-art accuracy of 97.42% without data augmentation and model ensemble. It also shows the potential of combining structural information of Chinese characters for recognition. Xu et al. [49] also considered the local features of Chinese characters and proposed a convolutional neural network combining multiple attention mechanisms, MCANet. It maps the last convolutional feature into multiple attention graphs, using contrast loss to restrict different attention mechanisms to focus on different subfeature spaces, eventually improving the accuracy to 97.66%. Zhang et al. [50] proposed an adaptive feature learning (AFL) algorithm to solve the difference of writing styles of different authors. The model improves the accuracy of the model by combining the prior knowledge of printed data and the author independent semantic feature information. Liu et al. [51] proposed a writing style confrontation network (WSAN) to weaken the impact of recognition accuracy caused by different writing styles. The network mainly consists of two classifiers: character classifier and writer classifier. By applying joint optimization to the top of the two classifiers, the loss value of the writer classifier is maximized, and the loss value of the character classifier is minimized, thus reducing the impact of writing style in handwritten Chinese character recognition. Song et al. [52] improved the recognition performance of the CNN model by extending the dataset through random elastic deformation, shear transformation, and rotation in a small range. Luo et al. [53] segmented the character images into sub-images based on the separability of the Chinese character structure. They increased the number of training data by recombining the sub-images. By using this training dataset that includes reconstructed and rotated characters, extracting features using DRN, and applying the center loss function, the final recognition accuracy of this method reaches 97.53% in the ICDAR 2013.

3.2.2. Create Fast and Compact Models

Networks tend to be larger and deeper to achieve higher recognition accuracy. However, with limited computational resources, researchers must consider the computational cost and memory consumption of the model.

The issues of model recognition speed and storage capacity in HCCR tasks were addressed by Xiao et al. [54] in 2017 by proposing a Global Supervised Low-Rank Expansion (GSLRE) method and an Adaptive Drop-Weight (ADW) strategy to build faster and more compact CNNs. Compared to the baseline model, their proposed method loses only a fraction of a percent of accuracy, but computational cost and storage parameters can be reduced by a factor of ten or more. In 2018, Li et al. [55] improved Global Average Pooling to address the problem of simply summing spatial information with the same weights at the expense of accuracy degradation. They added an attention mechanism to the Global Average Pooling method, called Global Weighted Average Pooling (GWAP), so that the fully connected layer parameters are reduced without the cost of accuracy degradation. In 2020, Melnyk et al. [56] made improvements to the GWAP proposed in the literature [55]. They proposed a high-performance identification network with interpretability. The network can learn the in-depth feature information and can visualize it. In addition, they introduced Global Weighted Output Average Pooling (GWOAP) to improve the performance of the model. In the same year, Jiang et al. [57] proposed CWCCNN-V1, a neural network consisting of the encoder and the classifier, which has a smaller depth and fewer parameters with guaranteed accuracy. Xu et al. [58] proposed a lightweight network named LightweightNet for HCCR. The network structure uses the low-dimensional features of the fully connected layer to save a large amount of memory. It constructs multiple compact modules to extract the features of the convolution layer and uses the extraction rules based on accuracy sensitivity to save the calculation time. The experimental results also proved the model's advantages in speed and performance. Inspired by the lightweight network shuffleNetV2, Xu et al. [59] recently proposed an efficient network model combining the Multiple scale convolution shuffle module (MSCS) for acquiring in-depth features of images and the attention features spatial aggregation (ASA) module for compressing important features. The method was tested at ICDAR 2013, requiring only 22.9 MB of storage, and achieving an accuracy of 97.63%.

3.2.3. Few-Shot Learning

Current approaches are mainly tested on large public datasets. Handwriting recognition tasks in real applications are much more complex than in public datasets. We need to develop unique datasets based on specific research tasks. However, manual labeling is not only cumbersome but also expensive. Therefore, researchers try to perform relevant handwriting recognition tasks using a small number of labeled or unlabeled samples.

A new neural network architecture was constructed by Zhu et al. [60], employing semi-supervised learning, which involves using a lot of labeled source data and little target domain data for model training. The algorithm mainly uses the domain adaptation in the transfer learning to transform the classification classifier. Finally, the effectiveness of the model was proved by testing on different data sets. Wang et al. [61] proposed a dense connection structure radical analysis network for analyzing Chinese character radicals and two-dimensional structures, namely DenseRAN. In this network, an encoder converts the image into high-level visual features. Then, a decoder composed of recurrent neural networks is used to obtain the relevant radicals and two-dimensional structures to improve the model's accuracy in recognizing Chinese characters. Moreover, the method can also handle some unseen Chinese characters in the training dataset. In 2019, Wang et al. [62] similarly exploited the decomposition knowledge of Chinese characters to implement the few-shot/zero-shot HCCR task. They proposed a novel radical aggregation network (RAN) consisting of a radical mapping encoder (RME), a radical aggregation module (RAM), and a character analysis decoder (CAD) to cope with the problems of insufficient representation of character radical features and inflexible decoding algorithms that affect recognition performance in [61]. RME can guarantee recognition performance for seen characters and recognize unseen characters effectively. In addition, with a small data sample, Li et al. [63] proposed a matching network to connect template characters and handwritten characters in 2020. The features taken from the template character image in place of the

softmax regression layer's parameters, the network can generalize well to new Chinese characters that are not part of the training set. At the same time, they used a special structure to accomplish the prediction of similarity between HCC and template images [64].

A novel zero-shot HCCR method combining the hierarchical knowledge of Chinese characters was proposed by Cao et al. [65]. First, the tree layout of primitives is obtained by manipulating the relationship between characters and their primitives, such as radicals and structures. Then, this paper proposes a new zero-shot hierarchical decomposition embedding method for encoding tree layouts into semantic vectors. Finally, a framework based on Convolutional Neural Network (CNN) is constructed to obtain relevant semantic vectors for learning radicals and character structures, thereby obtaining recognition results. Since different Chinese characters share some common radicals and structures, the zero-shot based method is able to identify new categories without extracting any labeled samples from them. The method achieved competitive performance on conventional experimental setups and significantly outperforms state-of-the-art methods on zero-shot experimental setups. In fact, character and radical-based methods are less satisfactory for recognizing some rare characters. Chen et al. [66] proposed a stroke-based character recognition method based on zero-shot learning. As shown in Figure 5, this method decomposes characters into stroke sequences through network learning and recognizes some rare characters based on matching strategies. The experimental results show good performance in character zero-shot and radial zero-shot scenarios.
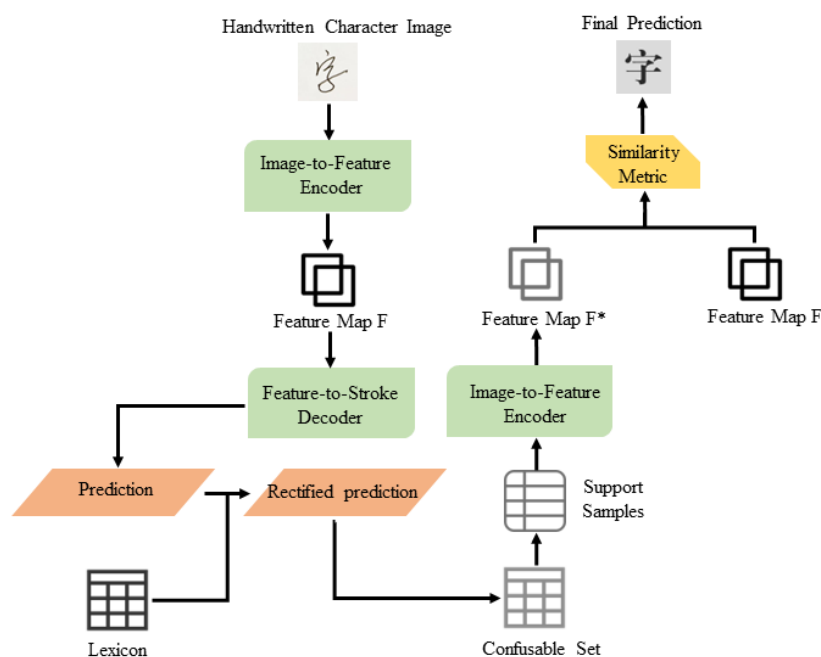

Figure 5. Schematic diagram of the HCCR algorithm based on zero-shot learning [66].

3.3. Combination of Deep Learning and Traditional Techniques

In addition to the mainstream use of deep learning methods to recognize HCC, some researchers have tried to combine traditional techniques to improve the performance of handwriting recognition systems further. The recognition process is shown in Figure 6. Specifically, for example, gradient features extracted by traditional techniques are fed into a deep network together with the original image [67]. In addition, features learned by deep networks are combined with traditional classifiers, such as MQDF [68]. However, since deep learning-based methods have enabled HCCR to achieve high recognition accuracy, only a few methods combining traditional techniques with deep learning networks have been developed in recent years.
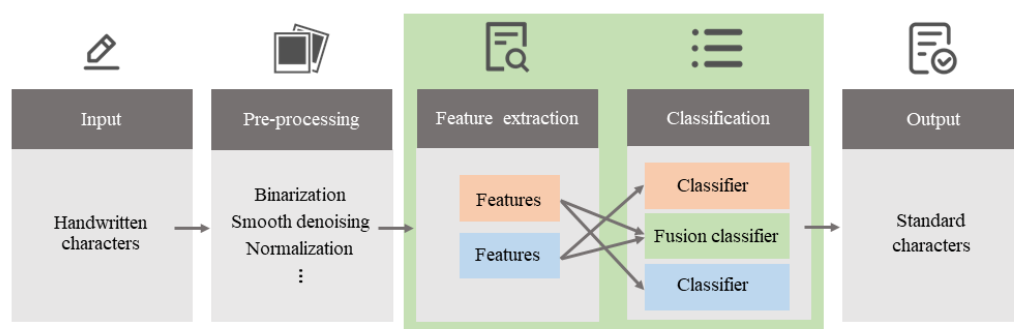
Figure 6. The process of recognizing offline HCCs using traditional techniques and deep learning. The larger green rectangle indicates the combination of different features and classifiers in the method. The orange rectangle indicates the traditional technique. The blue rectangle indicates the deep learning method. The smaller green rectangle indicates the combination of machine learning classifier and neural network classification.

Liu et al. [69] fused MQDF with a deep belief network (DBN) as a new classifier cascaded model. When recognition confidence after using MQDF is below the threshold, DBN is used for re-identification to obtain the final recognition result. Zhang et al. [70] improved the model performance using specific domain knowledge. Specifically, they achieved 97.37% recognition accuracy on the ICDAR 2013 dataset in 2017 by integrating the traditional normalized cooperative directional decomposition feature map (directMap) and deep convolutional neural network (convNet) without data augmentation or model ensemble. In addition, the authors proposed an adaptive layer to reduce the mismatch between the specific source and test domains. New writing styles can be adapted to particular writers. Table 4 summarizes the recognition performance of different deep learning methods and some methods combining traditional techniques on the ICDAR 2013 dataset. Since some methods do not state their training time and testing speed in the literature, the comparison is more concerned with the recognition performance and model size of different methods.

### 3.4. Novel Methods of Applying Other Knowledge

Faced with the problems of large training data requirements, many parameters, and high consumption of computational resources brought by deep learning methods-based recognition of HCC, some researchers have applied rather innovative methods of other concepts in the face of small samples to bring new ideas to the HCCR task.

A handwritten Chinese character recognition system based on image alignment technology was proposed by Li et al. [71] in 2016. The algorithm builds a nearest neighbor classifier based on template matching and improves the modeling ability of different text types by utilizing the average image transformation as a basic module, and then adopts a fuzzy entropy-based metric function. However, the fuzzy entropy metric necessitates a separate calculation for the similarity relation per pixel stack across all image pixels. The system may not outperform techniques based on feature extraction methods or traditional information entropy, which is a problem that needs to be addressed. Simultaneously, more emphasis should be placed on improving character recognition accuracy for complex structures.

Table 4. Recognition performance of different deep learning methods and some methods combining traditional techniques on the ICDAR 2013 dataset. "n/a" indicates no mention.

Method	Year	Experimental Platform	Framework	Accuracy (%)	Storage (MB)
STN-Residual-34 [37]	2016	Intel(R) Xeon(R) E5-2680v3 2.50GHz CPU, one NVIDIA Titan X GPU	Torch7	97.37	92.3
ResNet + center loss [44]	2017	GTX TITAN X GPU	Caffe	97.03	n/a
M-RBC + IR [35]	2017	NVIDIA K40 GPU	n/a	97.37	n/a
DRN + Radical region (b) [48]	2017	GeForce GTX 1080 GPU	Caffe	97.42	125.8
DirectMap + ConvNet+ Adaptation [70]	2017	NVIDIA Titan X 12G GPU	Theano	97.37	23.5
HCCR-CNN12Layer + GSLRE 4X + ADW [54]	2017	single-thread CPU	Caffe	97.4	3
DenseRAN [61]	2018	one GPU	n/a	96.66	n/a
Cascaded model (quantization) [55]	2018	GTX TITAN X GPU	Tensorflow	97.11	3.3
AFL [50]	2018	n/a	n/a	98.29	18.2
A novel RAN [62]	2019	n/a	n/a	96.97	n/a
WSAN [51]	2019	n/a	n/a	97.27	28.7
HCCR14Layer + template-instance loss [47]	2019	n/a	n/a	97.45	n/a
MCANet [49]	2019	n/a	Pytorch	97.66	n/a
Inception-ResNet + DWL + Pruning [45]	2019	NVIDIA Titan Xp GPUs	Caffe	98.79	n/a
HCCR-CNN12Layer + pm + sep [58]	2019	NVIDIA Titan X 12G GPU	Caffe	97.42	5.95
Deep matching network [64]	2020	GTX1060 (3G) card	Tensorflow	95.58	n/a
HDE-Net [65]	2020	n/a	Pytorch	97.14	n/a
Melnyk-Net [56]	2020	NVIDIA GeForce GTX 1080 Ti with 11 GB memory	Tensorflow	97.61	24.9
Stroke-based method [66]	2021	NVIDIA RTX 2080Ti GPU with 11 GB memory	Pytorch	96.74	n/a
MSCS + ASA [59]	2022	NVIDIA GeForce GTX 2080 Ti with 11 GB memory	Tensorflow	97.63	22.9

Figure 7 shows the process of HCC learning with concept learning, first proposed by Xu et al. [72] in 2019. Different from traditional deep learning methods, concept learning only needs one sample to achieve related recognition tasks. The method first builds a meta-stroke library with prior knowledge. Then, combined with the stroke extraction method and Bayesian learning, a conceptual model of Chinese characters is constructed. Finally, a character generation model is constructed through Monte Carlo-Markov chain sampling, to obtain the recognition and classification results of the target characters. However, one limitation of the method is that the number of concept models will increase with the complexity of the characters to achieve better performance, which will take a long time to build and fit.
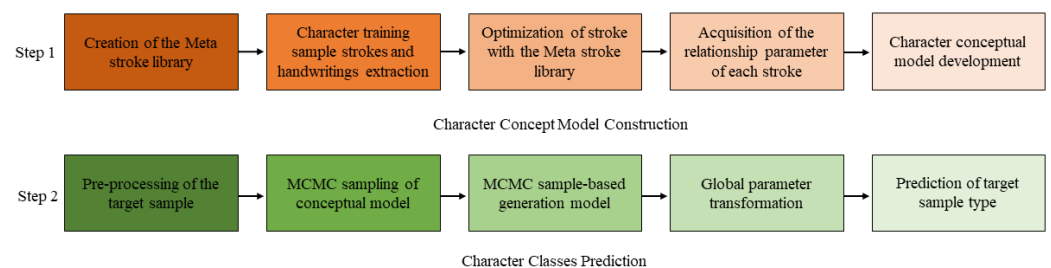
Figure 7. Framework of an HCCR method based on concept learning [72].

A discrete dynamic recognition system based on iteration and auxiliary functions employing linear combinations was constructed by Yu et al. [73] in 2020. The algorithm improves the system structure by adding bevel and translation to construct the font surface, to solve the convergence problem of the flatness of the Chinese character image surface. To accurately recognize handwritten characters, it is critical to learn an appropriate distance metric to measure the difference between data inputs. Existing distance metric learning methods either produce unacceptable error rates or provide results with little interpretability.

In 2021, Dong et al. [74] proposed an interpretable distance metric learning method. First, an algorithm named MetChar was proposed to optimize the weight distribution of fixed components. Then, the author proposed an algorithm named HybridSelection to select components and input them into MetChar to learn the distance metric of handwritten Chinese characters. Their work is inferior to the neural network-based approach in recognition accuracy but shows easy interpretation capabilities and good learning efficiency.

Figure 8 summarizes the number of papers generated by the HCCR study using the different approaches described in Sections 3.1–3.4 for each year from 2016 to 2022 and shows the annual variation of the highest recognition accuracy achieved on the ICDAR 2013 dataset without model ensemble and data augmentation. We can see that the deep learning method represented by the blue bar area has become the mainstream method for HCCR at present due to its ability to achieve high end-to-end recognition performance in large category classification tasks and the continued research by scholars on neural network model improvement. The HCCR recognition accuracy also reached more than 98% in 2018 and 2019, making the last two years of research on HCCR less hot and focusing on more research questions. For example, Ref. [66] focuses on the zero-shot problem that has not been fully solved in this field, and Refs. [56,59] sacrifice a small amount of recognition accuracy but achieve fewer storage requirements and faster inference time.
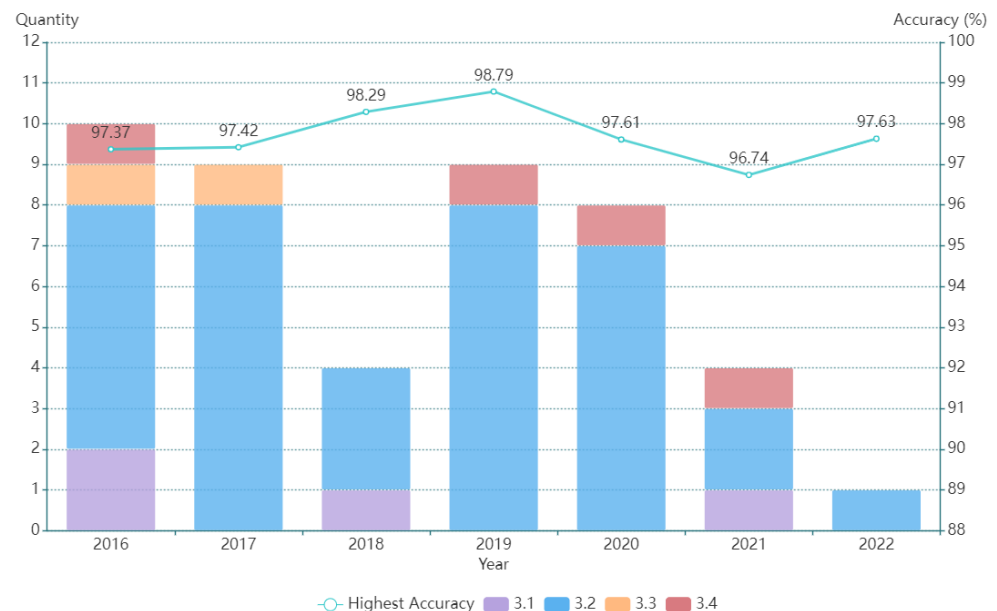

Figure 8. The number of papers per year from 2016 to 2022 on HCCR with different methods and the highest recognition accuracy achieved on the ICDAR 2013 dataset without model ensemble and data augmentation. The references corresponding to the highest recognition accuracy in each year are [37,45,48,50,56,59,66].

4. Handwritten Chinese Text Recognition

Compared with single-character recognition described in Section 3, HCTR is more complex and relatively less accurate due to the unconstrained nature of text lines and the adhesion between characters. It can be further divided into line-level HCTR and page-level

HCTR depending on whether the recognition object is a cropped image of a text line or an entire page. This chapter will analyze and summarize the methods for recognizing line-level HCT and page-level HCT in the past seven years.

4.1. Line-Level Text Recognition

Line-level text recognition methods are mainly divided into two categories: one is to perform pre-segmentation, use a single-character classifier for recognition, and combine it with the context to generate text lines. The other is to recognize the text directly without character segmentation. Segmentation-based approaches mostly use over-segmentation and depth detection networks. Since HCCs tend to form overlapping and consecutive strokes between characters in the unrestrained writing process, it will significantly impact the cutting results, thus limiting the final recognition accuracy. The over-segmentation-based methods also face the problem of data sparsity. Therefore, there has been more research on segmentation-free based methods in the past few years. Clearly, the segmentation-free approach loses the detection information of individual characters, so some scholars combine segmentation and segmentation-free approaches and use deep detection networks to achieve end-to-end recognition of text while preserving the location information of characters. We will present both types of approaches in the following.

4.1.1. Segmentation-Based Recognition

The segmentation-based text recognition approach is based on individual Chinese character recognition. Both character shape modeling and linguistic context modeling play a significant role. Most of the previous segmentation-based approaches are based on over-segmentation strategies. This approach usually involves building several modules, including character over-segmentation, character classification, language modeling, and geometric modeling. Then, they are integrated to find the optimal path.

In 2016, Wang et al. [75] used deep knowledge containing character or non-character labels, class labels, and the heterogeneous CNN that can be trained by deep knowledge to recognize HCT. The text will be over-segmented, so the HCT deep knowledge contains two parts: segmentation labels and classification labels for each segmented character. They proposed two types of heterogeneous CNN models, the cascading CNN containing 2-class CNN and all-class CNN, and the negative-awareness CNN. The proposed method has the advantage of incorporating the language model deeper into the path search when combined with language models compared to Long Short-Term Memory (LSTM) and Hidden Markov Model (HMM) frameworks that are good at handling time sequence tasks. Subsequently, during the period of back-off N-gram language models (BLMs) as the standard language model for the HCTR task, Wu et al. [76] first tried to recognize HCT using the neural network language model (NNLM) based on over-segmented text. They use CNN-based neural networks as a shape model for text recognition, which is used to accomplish over-segmentation of characters, classification, and as a geometric context model, and combined with hybrid NNLMs to improve the performance of the recognition system. Their approach achieved the best recognition result then and is often used as a comparison for the recognition results of other research methods on HCTR afterward. The research approaches since then have been focused on segmentation-free text recognition. In 2020, Wang et al. [77] proposed an HCTR method based on weakly supervised learning, as shown in Figure 9. Firstly, the text image is segmented by a CNN-based segmentation algorithm to generate a series of original text segments. Then, it is combined into one or more consecutive text segments to form a corresponding candidate character pattern and form a character candidate region. Finally, the wave beam search method is used to search the optimal path in the candidate region. The language and character classification models are combined to obtain the final recognition result. Their final experimental results show the competitive recognition performance of the method compared with other state-of-the-art methods. There is also room to further improve the performance by combining more powerful language models.
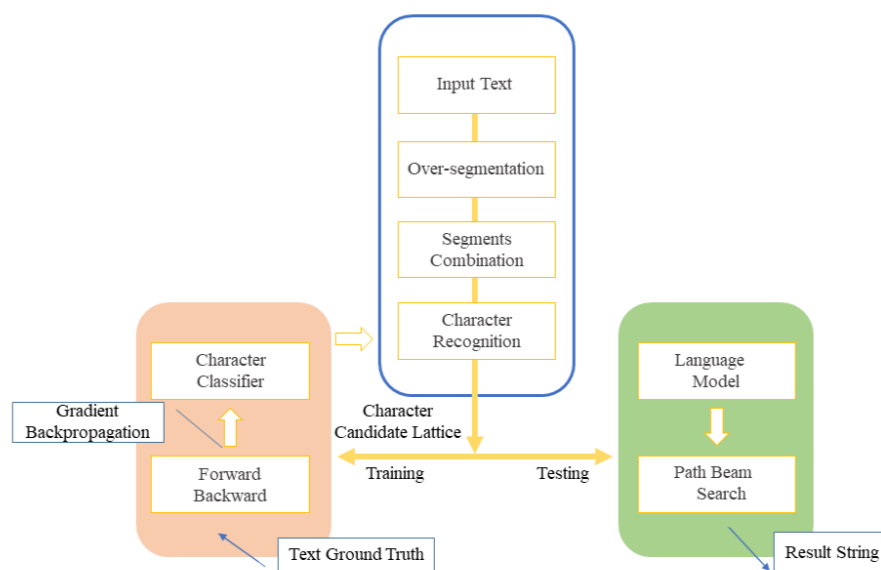
Figure 9. HCTR algorithm based on weakly-supervised learning [77].

Unlike over-segmentation-based methods, Peng et al. [78,79] used the full convolutional network-based approach for character detection as character segmentation, providing a new idea for segmentation-based HCTR. They proposed a fully convolutional network (FCN) that enables fast and accurate implementation of HCTR [78]. The structure extends three modules for detecting, localization, and classifying characters in text after a backbone network generates feature maps. Their proposed network model not only produces recognition results compared to the mainstream approach Recurrent Neural Network (RNN)/LSTM but also contains segmentation results that can be used for other subsequent applications, such as removing certain information. Compared with over-segmentation-based methods such as [75], this method can somewhat solve the problem of difficulty in segmenting characters due to overlapping characters, etc. Still, the recognition efficiency is not very satisfactory compared to other methods. In 2022, they further proposed a segmentation-annotation-free approach to segment and recognize text [79]. Compared with [78], they proposed a weakly supervised learning method that trains the network with only transcript annotations, which greatly reduces the cost of manual annotation while ensuring the output of segmented characters. In addition, they devised a contextual regularization method that incorporates integrated contextual information in the training of the FCN, thus significantly improving the recognition performance. The latest research is Hu et al. [80], who proposed a new retrieval-based approach that dynamically retrieves relevant content of the recognized text from the Internet to train an adaptive language model (LM) that can be integrated into the whole recognition process. Their strategy goes through a two-pass recognition process. After the initial recognition result is obtained by completing the first pass using the baseline system, the second pass recognition builds on this by retrieving relevant content on the Internet and building a domain-related language model integrated into the path search. Eventually, the method does not use much synthetic data compared to [79]. It achieves recognition results that can compete with [76], which has a large amount of character-level annotation and uses NNLM.

4.1.2. Segmentation-Free Recognition

The segmentation-free recognition method focuses on continuous feature extraction of the input text images. The possible sequences of characters represented in the sequential features are then matched with those of single characters. The best-matched text sequence is derived as the output of the whole recognition according to the evaluation criteria of the objective function. The HMM based on the Gaussian Mixture Model (GMM) proposed by

Su et al. [81] became the representative algorithm in handwritten text recognition. As the length of recognized characters increases, the HMM-based approach generates too many parameters that limit the recognition performance. Subsequently, offline handwritten text recognition based on or combined with neural networks has evolved more significantly. In recent years, some approaches using the Multi-dimensional LSTM (MDLSTM) Model [82], Convolutional Recurrent Neural Network (CRNN) [83], Connectionist Temporal Classification (CTC) algorithm [84], and attention mechanism [85], etc., have improved the model recognition accuracy to different degrees.

To apply MDLSTM-RNN to HCTR with an accuracy of 83.5% was first proposed by Messina et al. [86]. Shen et al. [87] pointed out that the current data augmentation method can significantly improve the accuracy of the recognition task of printed documents. However, it is not suitable for the task of handwriting recognition. Therefore, the team utilizes the baseline model MDLSTM-RNN [88] for generating synthetic line images. By training and testing on the CASIA dataset, the character error rate of the model trained with synthetic images and real images is reduced by 10.4%. Bluche et al. [89] proposed using character decomposition techniques to speed up HCTR for the slow system of [86]. Their system consists of an optical model based on MDLSTM-RNN and a language model. They use Cangjie [90], Wubi [91], and an arbitrary encoding to decompose characters for experiments without significant modification to the baseline MDLSTM-RNN. Wu et al. [92] improved the MDLSTM-RNN network. In addition, the network based on SMDLSTM-RNN applies fewer MDLSTM layers than the network based on MDLSTM-RNN [86], consumes much fewer resources and computation, and improves the accuracy by 3.14% without adding a corpus.

Du et al. conducted a series of studies combining neural networks with HMM. In 2016, they proposed a deep neural network (DNN-HMM) based on the Bayesian decision for handwritten text recognition [93]. The model extracts the gradient features based on the classifier of DNN, and then the HMM models the text lines sequentially. Finally, the feature language model is integrated with the DNN-HMM feature model, and the final recognition result can be obtained by Bayesian decision. Later, they extended their work in [94] to investigate the key issues of feature extraction, character modeling, and language modeling when using GMM-HMM, DNN-HMM, and DCNN-HMM to implement HCTR. They proposed a hybrid neural network Hidden Markov Model (NN-HMM) and proved its effectiveness for offline HCTR [95]. In 2020, they proposed a Chinese text recognition method (CNN-PHMM) that combines a writer-aware CNN network and parsimonious HMM [96]. First, the state binding algorithm is used by PHMM to reduce the total amount of HMM states. Second, WCNN integrates each convolutional layer with an adaptive layer composed of writer-dependent vectors and optimizes it by combining it with other network parameters to obtain the final recognition result. Though the method segments characters, their approach reduces the CER for recognition of the ICDAR 2013 dataset to the best result at that time based on the compact design of the output layers and the writer-aware convolutional layers. Due to the high dependence of this approach on a large amount of writer-specific data and it needing multiple-pass decoding, which is time consuming, they proposed another generalized writer adaptation scheme in 2022 that can achieve fast sentence-level unsupervised adaptation [97]. Trained by identification loss (IDL), they proposed a style extractor network (SEN) consisting of a convolutional layer, a recurrent neural network with gated recurrent units (GRU), where the writing style is represented by a one-dimensional vector integrated by the output of the GRU. The style information extracted by the SEN is fed into the writer-independent recognizer to achieve adaptation. Thus, the performance of the recognizer using SEN is improved. Notably, in 2021, Wang et al. [98] also proposed an intelligent residual attention gate module combining residual and attention frameworks based on fully convolutional neural networks. The module enhances the ability to extract meaningful features from text images by weighting to increase the importance of representative features and reduce the influence of background or noise, significantly improving the performance of the current HCTR.

Different neural networks have different advantages and can be integrated by model merging. In recent years, RNNs and their modified networks have been used more often in combination with convolutional networks to recognize handwritten text. In 2019, Xiu et al. [99] proposed a neural network structure combining a Multi-level Multimodal Fusion module and an Attention-based LSTM module to fully use visual feature information and linguistic semantic information. Zhang et al. [100] proposed a Bidirectional Recurrent Neural Network (BiRNN) for text recognition. The network first obtains the semantic feature map by the convolution layer and then converts the semantic information into time series information through the conversion module and processes it through Bidirectional-LSTM. The results prove that the algorithm can analyze and predict characters by linking forward and backward contexts and effectively solve the problem of sample imbalance. In 2020, Xie et al. [101] proposed a data enhancement method for HCTR and constructed a recognition network model combining CNN and LSTM (CNN-ResLSTM). The experiment proved that data preprocessing and data augmentation methods can effectively improve the model performance.

In addition to the character encoding method discussed in [89], Hoang et al. [102] developed a new encoding technique called LOgographic DEComposition Encoding (LODEC). It can perform one-to-one mapping of Chinese. Instead of splitting characters into scattered stroke elements, the proposed encoding limits the decomposition to predefined radical components to encode many Chinese characters with a small number of basic elements. They also proposed a deep learning architecture called LODENet that extracts radical-based features from the input data, which is then decoded by a transformation network for radical-based features. In 2021, Ngo et al. [103] first proposed to use the RNN-Transducer model for offline HCTR. The network mainly comprises a visual feature encoder, language context encoder, and joint decoder. Specifically, CNN is used to obtain relevant visual features, and then the visual features are encoded according to the BLSTM module. On the other hand, relevant language context information is obtained through the Embedding layer and the LSTM module. Finally, the visual and language features are coupled and decoded into the final recognition result through the fully connected and softmax layers. In addition to the above approaches, the visual characteristics of glyphs and the semantics of Chinese characters are also exploited by [104]. In 2022, Zhan et al. proposed the Glyph-Semanteme fusion embedding (GSE) module according to the correlation between glyphs and semantics. Specifically, the decoder prediction obtains the recognition result by extracting the glyph and semantic embedding of Chinese characters and automatically calculating the symbol speech fusion embedding of characters according to the parametric gating fusion strategy. Furthermore, two kinds of generalized systems engineering, character-level generalized systems engineering (CGSE) and text-level generalized systems engineering (TGSE), are applied in the decoder stage to generate predictions. On the standard benchmark ICDAR-2013 HCTR competition dataset, the method achieves a character-level recognition accuracy of 96.65%, which demonstrates the effectiveness of the proposed glyph-semantic fusion embedding.

Figure 10 counts the number of papers on line-level HCTR research by segmentation and segmentation-free based approaches from 2016 to 2022, as well as the highest AR achieved on the ICDAR 2013 dataset with and without LM per year. In general, segmentation-free approaches have become more mainstream in recent years due to their avoidance of expensive character segmentation annotations and character segmentation errors, combined with writer-aware, language models, etc., to improve recognition performance. Current research on segmentation-based approaches aims to achieve high-performance end-to-end recognition while providing character location information. In addition to using a single strategy, Zhu et al. [105] proposed an attention combination model that can combine the advantages of segmentation and segmentation-free recognition methods. Both provide character position information for more accurate recognition of characters, avoid segmentation errors, and enable end-to-end recognition. They first obtained recognition results for multiple recognition methods. Then, a combined model of

one-dimensional gated convolutional neural networks containing multiple encoders and a decoder could infer the final recognition results. Their experimental results show the effectiveness of combining explicit and implicit segmentation methods. Table 5 summarizes the recognition accuracy of different methods on the ICDAR 2013 dataset for line-level text.
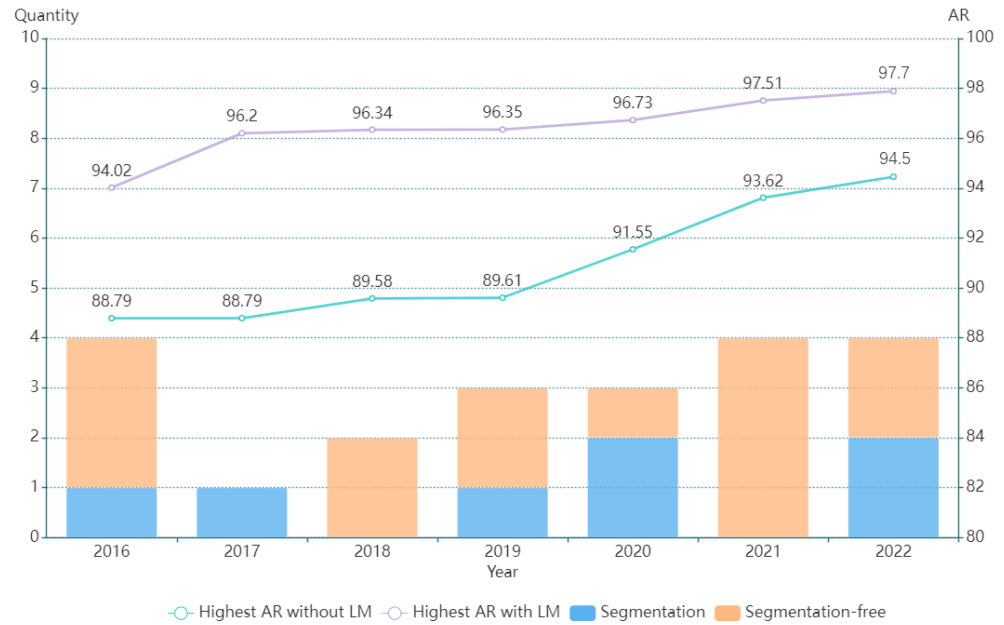

Figure 10. The number of papers per year from 2016 to 2022 on line-level HCTR with different methods and the highest AR (100%-CER) achieved on the ICDAR 2013 dataset with and without LM. The references corresponding to the highest AR (without LM) per year are [75,78,79,95,101,106]. The references corresponding to the highest AR (with LM) per year are [75,76,79,95,96,99,106].

Table 5. AR (100%-CER) of line-level text recognition by different methods on the ICDAR 2013 dataset. "n/a" indicates not mentioned, "×" indicates used, "√" indicates not used.

Method	Year	Experimental Platform	Framework	Method Type	Without LM	LM	Data Synthesis
Negative-awareness CNN (trigram) [75]	2016	n/a	n/a	over-segmentation	88.79	94.02	×
DNN-HMM [93]	2016	n/a	n/a	segmentation-free	83.89	93.5	×
DCNN-HMM [94]	2017	n/a	Caffe	segmentation-free	n/a	95.93	×
HRMELM [76]	2017	Intel Core i7-4790 3.60 GHz CPU, NVIDIA Titan X GPUs	Caffe	over-segmentation	n/a	96.2	×
NN-HMM [95]	2018	n/a	Caffe	segmentation-free	89.58	96.34	×
FCN (with SRM) [78]	2019	GTX 1080Ti GPU	n/a	segmentation	89.61	95.51	√

Table 5. *Cont.*

Method	Year	Experimental Platform	Framework	Method Type	Without LM	LM	Data Synthesis
Attention-Based LSTM [99]	2019	n/a	Tensorflow	segmentation-free	88.74	96.35	×
Attention combination of sequence models [105]	2020	two TITAN Xp GPUs	Pytorch	combination of segmentation and segmentation-free	90.86	94	√
CNN-ResLSTM [101]	2020	i7-8700K CPU, single Titan X GPU	Caffe	segmentation-free	91.55	96.72	√
Weakly supervised learning [77]	2020	NVIDIA GeForce GTX 1080Ti GPUs	Pytorch	over-segmentation	n/a	95.11	×
WCNN-PHMM [96]	2020	NVIDIA GeForce GTX 1080Ti GPUs	Kaldi, Pytorch	segmentation	91.36	96.73	×
Residual-attention [98]	2021	Intel Core i9-9900K 3.60 GHz CPU, NVIDIA RTX 2080Ti GPUs	Tensorflow	segmentation-free	91.30	96.51	×
CNN-CTC-CBS [106]	2021	two NVIDIA Titan V GPUs	Pytorch	segmentation-free	93.62	97.51	√
Attention-based model + GSE [104]	2022	Intel XEON E5-1650 with 3.5 GHz, one NVIDIA GTX 2080Ti GPU	Tensorflow	segmentation-free	89.87	96.65	×
Common + Retrieval LM + Improvements [80]	2022	n/a	n/a	over-segmentation	n/a	95.48	×
Segment-annotation-free [79]	2022	NVIDIA GTX 1080Ti GPU with 11 GB of memory	Pytorch	segmentation	94.50	97.70	√

4.2. Page-Level Text Recognition

In addition to recognizing line-level text, driven by the critical value of text recognition for fields such as industry and education, scholars have recently conducted several studies on recognizing page-level handwritten text. Refs. [107–110] recognize traditional ancient texts read from top to bottom, and Ref. [111] recognizes texts under standard test paper formatting specifications, which are restricted to a specific layout. Faced with unconstrained text, Kundu et al. [112] extracted text lines from unconstrained handwritten documents using Generative Adversarial Networks (GAN). Yan et al. [113] identified the failure of existing methods to correctly identify Chinese text when adding keywords and word replacement as a two-dimensional problem. Therefore, the team proposed a Structural Attention Network (SAN) model to learn irregular structures to recognize text correctly with insertions and swaps. A critical attempt and advancement of unconstrained page-level HCTR was in 2022; Peng et al. [114] proposed a new end-to-end weakly supervised page-level HCTR method, PageNet. The method consists of three modules, a module for detecting and recognizing characters, a reading order module for determining intercharacter relationships and determining head and tail characters, and a graph-based decoding module for outputting the detection and recognition results, trained in a weakly

supervised learning framework. This method achieves solving for the the HCTR page-level reading order problem under weak supervision and can handle unconstrained text. It also provides new ideas and inspirations to other researchers to solve the page-level text recognition problem.

5. Discussion

In this section, we summarize and compare the different approaches of HCCR and HCTR, compile the main concerns about improving the effectiveness of recognition methods in this field over the last seven years, discuss the existing challenges and future research work, and aim to provide directions and methodological references for the subsequent research.

5.1. Methods Comparison

Tables 6 and 7 show the advantages and disadvantages of the different methods for recognizing offline HCC and offline HCT, respectively. For character recognition, compared to extracting features with traditional techniques and then using machine learning classifiers, deep learning methods can achieve end-to-end recognition in this enormous class classification task. Along with the continuous improvement of neural network models, deep learning-based methods have achieved good recognition performance in HCCR even beyond that of humans and are now the dominant methods in the field. Some new techniques that combine other concepts and focus more on small samples have inferior recognition accuracy compared to deep learning methods but can be easily interpreted and have good learning efficiency. The focus of text recognition methods is gradually shifting from segmentation-based to segmentation-free based that can provide good end-to-end recognition performance. Combining the advantages of both, i.e., achieving end-to-end recognition while providing character location information, has also emerged in the last two years to show effectiveness. Combining models with writer awareness and language models is also a common practice in this field to improve recognition.

Table 6. Summary of the advantages and disadvantages of methods for recognizing offline HCC.

Method Type	Advantages	Disadvantages
Traditional technology	Low hardware requirement; Relatively easy to implement; Good effect for specific scenes; Greater interpretability.	Difficult to capture advanced semantic features; Poor accuracy; Poor generalizability and robustness.
Deep learning	End-to-end training available; High accuracy; Stronger generalizability and robustness.	Highly influenced by dataset; High computational requirement; Weak interpretability.
Traditional techniques combined with deep learning	The features extracted by traditional techniques can be added with prior knowledge; Fusion classifiers with neural networks and machine learning classification algorithms can be used to improve recognition accuracy.	The steps are relatively independent and lack a global optimization scheme for control; Less efficient.
Other concept	Research can be based on small samples; Easy to interpret and has good learning efficiency.	The recognition accuracy is inferior.

Table 7. Summary of the advantages and disadvantages of methods for recognizing offline HCT.

Method Type	Advantages	Disadvantages
Segmentation	Provides the detection information of individual characters.	The segmentation effect is affected by character overlap and adhesion, which limits recognition accuracy; Over-segmentation faces the problem of sparse data.
Segmentation-free	Avoid costly character segmentation annotations and segmentation errors to achieve good end-to-end recognition.	Loss of character location information.

5.2. Methods Focus

Figure 11 summarizes the research issues researchers have focused on for offline handwritten Chinese recognition from 2016 to 2022. For offline HCCR, most researchers work on improving recognition accuracy, such as modifying the model structure, using different loss functions, weakening different writing styles, and extending the dataset for training through data augmentation or character structure reorganization. Some researchers have also tried to recognize handwritten characters using a small number of labeled or unlabeled samples or to recognize unseen characters in the training set using knowledge of the decomposition of the characters. There are also issues such as combining powerful or adaptive language models, using encoding approaches that combine glyphs and semantics, extending the training dataset by data synthesis, and using writer adaptation methods to improve recognition accuracy. Other research issues are recognition based on weakly supervised learning and recognition of irregular text. In contrast to previous approaches based on segmentation or segmentation-free, recent HCTR research approaches are dedicated to end-to-end text recognition while providing character location information.
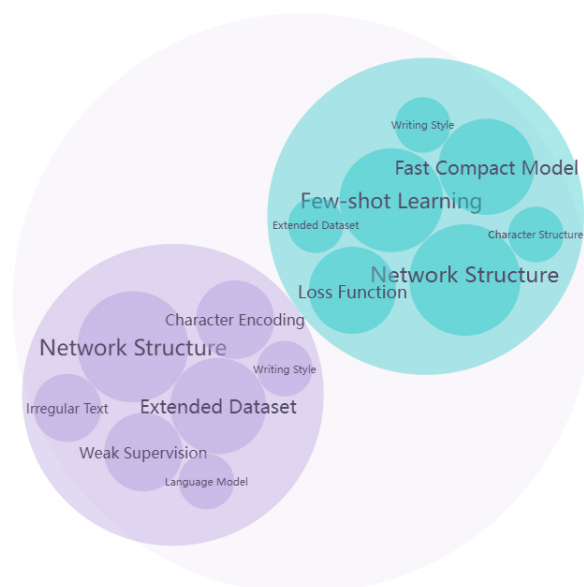

Figure 11. Research questions focused on handwritten Chinese recognition from 2016 to 2022. The blue circle represents research on HCCR, and the purple circle represents research on HCTR. The size of each sub-circle and label represents the number of studies on the issue.

5.3. Challenges and Future Work

As seen from Figures 8 and 10, the current accuracy of offline HCCR and offline line-level HCTR with and without LM reach over 98%, 97%, and 94%, respectively, which is close to or even exceeds human recognition performance in terms of accuracy. However, considering the practical applications of handwritten Chinese recognition, some issues in this field still deserve attention and research.

1. Offline handwritten Chinese recognition in different scenarios. The recognition objects of current research are mostly characters and text in simple scenarios. However, in practical life, we must process handwritten Chinese in more complex scenes influenced by lighting, shooting, and different applications.

2. Create unconstrained handwritten datasets with large categories and multiple styles. Adequate and diverse data are essential for designing and training handwritten Chinese recognition models. Therefore, more unconstrained handwriting databases containing large classes and various styles are currently needed. For most current experiments, researchers have only trained and tested models on two public

datasets, CASIA-HWDB and ICDAR 2013. However, they still do not contain a large enough variety of Chinese characters and sample size.

3. Ultra-large category offline handwritten character recognition. The current studies focus on recognizing the 3755 classes of Chinese characters commonly used in the GB 2312-80 regulation, but the total number of Chinese characters exceeds 50,000. In the future, the objects of ultra-large category recognition can include more categories of simplified and traditional Chinese characters and English, Japanese, and other significant languages commonly used worldwide to achieve universal recognition of offline handwritten characters.

4. Unconstrained offline page-level handwritten Chinese text recognition. Most current studies on offline HCTR focus on line-level text that only focuses on the recognition performance of cropped text lines. The few studies on page-level HCTR are also limited to pages with specific formats, such as ancient books or standard test papers. However, modifications and random irregular writing can significantly affect the recognition results in practical application scenarios.

5. Create fast and compact practical models. Research on handwritten Chinese recognition based on deep learning often accompanies many parameters that consume much time to train the model. This makes deploying the models in embedded systems such as Raspberry Pi and cell phones challenging.

To facilitate further work, Table 8 summarizes the existing challenges, applications, and future research on handwritten Chinese recognition, expecting to provide some references for researchers.

Table 8. A summary of existing challenges, applications, and future research on offline handwritten Chinese recognition.

Existing Challenges	Applications	Future Research
Less research on recognition tasks in complex scenes	Recognize handwritten medical manuscripts with adherent or simplified characters; Recognize irregular or mutilated characters written by people with functional impairment; Recognize similar characters in jobs requiring high-precision recognition; Recognize handwritten characters in natural scenes under different lighting and photo angles.	Specific datasets can be trained for different application scenarios and targeted to address recognition challenges.
Lack of datasets containing sufficient variety and number of samples	Designing and training recognition models using unconstrained handwriting datasets with large categories and multiple styles to adapt to recognizing handwritten Chinese in different scenarios.	Create datasets with more types of Chinese characters and larger sample sizes; Create specific datasets for different recognition application scenarios; Use some data augmentation and data synthesis techniques to create handwritten Chinese and expand datasets.
Recognition requirements for offline handwritten characters with ultra-large categories	Recognize more categories of traditional characters used in antiquarian research or Hong Kong, Macau, and Taiwan; Recognize multilingual text.	Balancing recognition effectiveness, processing speed, and model size for ultra-large category classification
Difficulty in recognizing unconstrained offline page-level handwritten text	Recognize page-level text that contains modifications, overlapping of upper- and lower-line characters, irregular text lines with unequal line spacing, and different writing styles.	Try to combine traditional techniques to detect irregular text lines. Consider more page-level text scenarios in practical applications to improve the performance of unconstrained HCTR.
Demand for fast and compact practical models	Deploy models in embedded systems such as Raspberry Pi and cell phones for easy daily use.	Optimize the network structure, improve the underlying code of the model, and explore relevant software and hardware combinations; Combine different compression and acceleration algorithms to create models with a low number of parameters and a reasonable recognition rate.

6. Conclusions

Offline handwritten Chinese recognition is one of the hot research topics in computer vision and pattern recognition, which is often applied to the entry of handwritten documents in banking, medical, postal, and education. In this paper, we introduce the different techniques of offline HCCR and offline HCTR from 2016 to 2022 and compare the advantages and disadvantages of various recognition methods and their test performance on public datasets. We compose the main concerns for improving the recognition methods in this field, presenting the existing challenges and future work. This paper aims to review the research progress of offline handwritten Chinese recognition in the past seven years and provide ideas and inspirations for future research by summarizing and analyzing various methods and problems.

Author Contributions: Conceptualization, L.S.; methodology, L.S.; software, L.S.; validation, L.S.; formal analysis, L.S. and B.C.; investigation, L.S., B.C., and J.W.; resources, L.S., S.-K.T., and S.M.; data curation, L.S.; writing—original draft preparation, L.S.; writing—review and editing, L.S., S.-K.T., and S.M.; visualization, L.S. and H.X.; supervision, S.-K.T. and S.M.; project administration, S.-K.T. and S.M.; funding acquisition, S.-K.T. and S.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, C.L.; Yin, F.; Wang, D.H.; Wang, Q.F. Online and offline handwritten Chinese character recognition: Benchmarking on new databases. In *Pattern Recognition*; Elsevier: Amsterdam, The Netherlands, 2013; Volume 46, pp. 155–162.
2. Chen, L.; Wang, S.; Fan, W.; Sun, J.; Naoi, S. Beyond human recognition: A CNN-based framework for handwritten character recognition. In Proceedings of the 2015 IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 3–6 November 2015; IEEE: New York, NY, USA, 2015; pp. 695–699.
3. Mori, S.; Suen, C.Y.; Yamamoto, K. Historical review of OCR research and development. *Proc. IEEE* **1992**, *80*, 1029–1058. [CrossRef]
4. Liu, C.L.; Yin, F.; Wang, D.H.; Wang, Q.F. Chinese handwriting recognition contest 2010. In Proceedings of the 2010 Chinese Conference on Pattern Recognition (CCPR), Chongqing, China, 21–23 October 2010; IEEE: New York, NY, USA, 2010; pp. 1–5.
5. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
6. Memon, J.; Sami, M.; Khan, R.A.; Uddin, M. Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). *IEEE Access* **2020**, *8*, 142642–142668. [CrossRef]
7. Ruiz-Parrado, V.; Heradio, R.; Aranda-Escolastico, E.; Sánchez, A.; Vélez, J.F. A bibliometric analysis of offline handwritten document analysis literature (1990–2020). *Pattern Recognit.* **2021**, *125*, 108513. [CrossRef]
8. Sinwar, D.; Dhaka, V.S.; Pradhan, N.; Pandey, S. Offline script recognition from handwritten and printed multilingual documents: A survey. *Int. J. Doc. Anal. Recognit. (IJDAR)* **2021**, *24*, 97–121. [CrossRef]
9. Jin, L.W.; Zhong, Z.Y.; Yang, Z.; Yang, W.X.; Xie, Z.C.; Sun, J. Applications of deep learning for handwritten Chinese character recognition: A review. *Acta Autom. Sin.* **2016**, *42*, 1125–1141.
10. Su, T.; Zhang, T.; Guan, D. Corpus-based HIT-MW database for offline recognition of general-purpose Chinese handwritten text. *Int. J. Doc. Anal. Recognit. (IJDAR)* **2007**, *10*, 27–38. [CrossRef]
11. Wang, J.; Li, W.; Wang, J. Fault tolerant recognition method of handwritten chinese characters based on double weights elliptical neuron. In Proceedings of the International Conference on Intelligent Computing, Kunming, China, 6–19 August 2006; Springer: Berlin/Heidelberg, Germany, 2006; pp. 370–376.
12. Liu, C.L.; Yin, F.; Wang, D.H.; Wang, Q.F. CASIA online and offline Chinese handwriting databases. In Proceedings of the 2011 International Conference on Document Analysis and Recognition, Beijing, China, 18–21 September 2011; IEEE: New York, NY, USA, 2011; pp. 37–41.

13. Jin, L.; Gao, Y.; Liu, G.; Li, Y.; Ding, K. SCUT-COUCH2009—A comprehensive online unconstrained Chinese handwriting database and benchmark evaluation. *Int. J. Doc. Anal. Recognit. (IJDAR)* **2011**, *14*, 53–64. [CrossRef]
14. Yin, F.; Wang, Q.F.; Zhang, X.Y.; Liu, C.L. ICDAR 2013 Chinese handwriting recognition competition. In Proceedings of the 2013 12th International Conference on Document Analysis and Recognition, Washington, DC, USA, 25–28 August 2013; pp. 1464–1470.
15. Su, T.; Pan, W.; Yu, L. HITHCD-2018: Handwritten Chinese Character Database of 21K-Category. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 1378–1383.
16. Zhang, H.; Liang, L.; Jin, L. SCUT-HCCDoc: A new benchmark dataset of handwritten Chinese text in unconstrained camera-captured documents. *Pattern Recognit.* **2020**, *108*, 107559. [CrossRef]
17. Kobayashi, K.; Yoda, F.; Yamamoto, K.; Nambu, H. Recognition of handprinted kanji characters by the stroke matching method. *Pattern Recognit. Lett.* **1983**, *1*, 481–488. [CrossRef]
18. Govindan, V.K.; Shivaprasad, A.P. Character recognition—A review. *Pattern Recognit.* **1990**, *23*, 671–683. [CrossRef]
19. Jin, L.; Wei, G. Handwritten Chinese character recognition with directional decomposition cellular features. *J. Circuits Syst. Comput.* **1998**, *8*, 517–524. [CrossRef]
20. Meygret, A.; Levine, M.D.; Roth, G. Robust primitive extraction in a range image. In Proceedings of the 11th IAPR International Conference on Pattern Recognition. Vol. III. Conference C: Image, Speech and Signal Analysis, The Hague, The Netherlands, 30 August–1 September 1992; IEEE Computer Society: New York, NY, USA, 1992; Volume 1, pp. 193–196.
21. Hamamoto, Y.; Uchimura, S.; Masamizu, K.; Tomita, S. Recognition of handprinted Chinese characters using Gabor features. In Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 4–16 August 1995; Volume 2, pp. 819–823.
22. Chang, C.H. Word class discovery for postprocessing Chinese handwriting recognition. In Proceedings of the COLING 1994 Volume 2: The 15th International Conference on Computational Linguistics, Kyoto, Japan, 5–9 August 1994.
23. Chang, H.H.; Yan, H. Analysis of stroke structures of handwritten Chinese characters. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **1999**, *29*, 47–61. [CrossRef] [PubMed]
24. Chen, M.Y.; Kundu, A.; Zhou, J. Off-line handwritten word recognition using a hidden Markov model type stochastic network. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16*, 481–496. [CrossRef]
25. Shi, D.; Shu, W.; Liu, H. Feature selection for handwritten Chinese character recognition based on genetic algorithms. In Proceedings of the SMC'98 Conference—1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 98CH36218), San Diego, CA, USA, 14 October 1998; Volume 5, pp. 4201–4206.
26. Zhang, J.; Ding, X.; Liu, C. Multi-scale feature extraction and nested-subset classifier design for high accuracy handwritten character recognition. In Proceedings of the 15th International Conference on Pattern Recognition, Brisbane, Australia, 3–7 September 2000; ICPR-2000; IEEE: New York, NY, USA, 2000; Volume 2, pp. 581–584.
27. Kawatani, T. Handwritten Kanji recognition with determinant normalized quadratic discriminant function. In Proceedings of the Proceedings 15th International Conference on Pattern Recognition, Brisbane, Australia, 3–7 September 2000; ICPR-2000; IEEE: New York, NY, USA, 2000; Volume 2, pp. 343–346.
28. Sun, F.; Omachi, S.; Kato, N.; Aso, H. Fast and precise discriminant function considering correlations of elements of feature vectors and its application to character recognition. *Syst. Comput. Jpn.* **1999**, *30*, 33–42. [CrossRef]
29. Zhu, Y.; An, X.; Zhang, K. A handwritten Chinese character recognition method combining sub-structure recognition. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; pp. 518–523.
30. Zhou, M.K.; Zhang, X.Y.; Yin, F.; Liu, C.L. Discriminative quadratic feature learning for handwritten Chinese character recognition. *Pattern Recognit.* **2016**, *49*, 7–18. [CrossRef]
31. Wei, X.; Lu, S.; Lu, Y. Compact MQDF classifiers using sparse coding for handwritten Chinese character recognition. *Pattern Recognit.* **2018**, *76*, 679–690. [CrossRef]
32. Ma, J.; Wang, Z.; Du, J. An Open-Source Library of 2D-GMM-HMM Based on Kaldi Toolkit and Its Application to Handwritten Chinese Character Recognition. In Proceedings of the International Conference on Image and Graphics, Haikou, China, 6–8 August 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 235–244.
33. Feng, S.; Guo, P. Research of the properties of receptive field in handwritten Chinese character recognition based on DCNN model. In Proceedings of the Eighth International Conference on Digital Image Processing (ICDIP 2016), Chengu, China, 20–22 May 2016; Volume 10033, pp. 198–203.
34. Wu, X.; Shu, C.; Zhou, N. Multi-stage Feature Extraction in Offline Handwritten Chinese Character Recognition. In Proceedings of the Pattern Recognition, Rome, Italy, 24–26 February 2016; Tan, T., Li, X., Chen, X., Zhou, J., Yang, J., Cheng, H., Eds.; Springer: Singapore, 2016; pp. 474–485.
35. Yang, X.; He, D.; Zhou, Z.; Kifer, D.; Giles, C.L. Improving offline handwritten Chinese character recognition by iterative refinement. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; Volume 1, pp. 5–10.
36. Maidana, R.G.; dos Santos, J.M.; Granada, R.L.; de Morais Amory, A.; Barros, R.C. Deep neural networks for handwritten Chinese character recognition. In Proceedings of the 2017 Brazilian Conference on Intelligent Systems (BRACIS), Uberlandia, Brazil, 2–5 October 2017; pp. 192–197.

37. Zhong, Z.; Zhang, X.Y.; Yin, F.; Liu, C.L. Handwritten Chinese character recognition with spatial transformer and deep residual networks. In Proceedings of the 2016 23rd international conference on pattern recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 3440–3445.

38. Zhuang, Y.; Liu, Q.; Qiu, C.; Wang, C.; Ya, F.; Sabbir, A.; Yan, J. A Handwritten Chinese Character Recognition based on Convolutional Neural Network and Median Filtering. In Proceedings of the Journal of Physics: Conference Series, Diwaniyah, Iraq, 21–22 April 2021; Volume 1820, p. 012162.

39. Jaderberg, M.; Simonyan, K.; Zisserman, A. Spatial transformer networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 2017–2025.

40. Bi, N.; Chen, J.; Tan, J. The handwritten Chinese character recognition uses convolutional neural networks with the googlenet. *Int. J. Pattern Recognit. Artif. Intell.* **2019**, *33*, 1940016. [CrossRef]

41. Min, F.; Zhu, S.; Wang, Y. Offline Handwritten Chinese Character Recognition Based on Improved Googlenet. In Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition, Xiamen, China, 26–28 June 2020; pp. 42–46.

42. Aleskerova, N.; Zhuravlev, A. Handwritten Chinese characters recognition using two-stage hierarchical convolutional neural network. In Proceedings of the 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), Dortmund, Germany, 8–10 September 2020; pp. 343–348.

43. Cheng, C.; Zhang, X.Y.; Shao, X.H.; Zhou, X.D. Handwritten Chinese character recognition by joint classification and similarity ranking. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; pp. 507–511.

44. Zhang, R.; Wang, Q.; Lu, Y. Combination of ResNet and center loss based metric learning for handwritten Chinese character recognition. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; Volume 5, pp. 25–29.

45. Chen, L.; Peng, L.; Yao, G.; Liu, C.; Zhang, X. A modified inception-ResNet network with discriminant weighting loss for handwritten chinese character recognition. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 1220–1225.

46. Zeng, X.; Xiang, D.; Peng, L.; Liu, C.; Ding, X. Local discriminant training and global optimization for convolutional neural network based handwritten Chinese character recognition. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; Volume 1, pp. 382–387.

47. Xiao, Y.; Meng, D.; Lu, C.; Tang, C.K. Template-instance loss for offline handwritten Chinese character recognition. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 315–322.

48. Weike, L.; Sei-Ichiro, K. Radical region based CNN for offline handwritten Chinese character recognition. In Proceedings of the 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR), Nanjing, China, 26–29 November 2017; pp. 542–547.

49. Xu, Q.; Bai, X.; Liu, W. Multiple comparative attention network for offline handwritten chinese character recognition. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 595–600.

50. Zhang, Y.; Liang, S.; Nie, S.; Liu, W.; Peng, S. Robust offline handwritten character recognition through exploring writer-independent features under the guidance of printed data. *Pattern Recognit. Lett.* **2018**, *106*, 20–26. [CrossRef]

51. Liu, H.; Lyu, S.; Zhan, H.; Lu, Y. Writing Style Adversarial Network for Handwritten Chinese Character Recognition. In Proceedings of the International Conference on Neural Information Processing, Vancouver, BC, Canada, 8–14 December 2019; pp. 66–74.

52. Song, X.; Gao, X.; Ding, Y.; Wang, Z. A handwritten Chinese characters recognition method based on sample set expansion and CNN. In Proceedings of the 2016 3rd International Conference on Systems and Informatics (ICSAI), Shanghai, China, 19–21 November 2016; pp. 843–849.

53. Luo, W.; Zhai, G. Offline handwritten Chinese character recognition based on new training methodology. In Proceedings of the International Forum on Digital TV and Wireless Multimedia Communications, Shanghai, China, 8–9 November 2017; pp. 235–244.

54. Xiao, X.; Jin, L.; Yang, Y.; Yang, W.; Sun, J.; Chang, T. Building fast and compact convolutional neural networks for offline handwritten Chinese character recognition. *Pattern Recognit.* **2017**, *72*, 72–81. [CrossRef]

55. Li, Z.; Teng, N.; Jin, M.; Lu, H. Building efficient CNN architecture for offline handwritten Chinese character recognition. *Int. J. Doc. Anal. Recognit. (IJDAR)* **2018**, *21*, 233–240. [CrossRef]

56. Melnyk, P.; You, Z.; Li, K. A high-performance CNN method for offline handwritten Chinese character recognition and visualization. *Soft Comput.* **2020**, *24*, 7977–7987. [CrossRef]

57. Jiang, Y.; Song, Y. High-Accuracy Offline Handwritten Chinese Characters Recognition Using Convolutional Neural Network. *J. Comput.* **2020**, *31*, 12–23.

58. Xu, T.B.; Yang, P.; Zhang, X.Y.; Liu, C.L. LightweightNet: Toward fast and lightweight convolutional neural networks via architecture distillation. *Pattern Recognit.* **2019**, *88*, 272–284. [CrossRef]

59. Xu, X.; Yang, C.; Wang, L.; Zhong, J.; Bao, W.; Guo, J. A sophisticated offline network developed for recognizing handwritten Chinese character efficiently. *Comput. Electr. Eng.* **2022**, *100*, 107857. [CrossRef]

60. Zhu, Y.; Zhuang, F.; Yang, J.; Yang, X.; He, Q. Adaptively transfer category-classifier for handwritten Chinese character recognition. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Macau, China, 14–17 April 2019; pp. 110–122.

61. Wang, W.; Zhang, J.; Du, J.; Wang, Z.R.; Zhu, Y. Denseran for offline handwritten chinese character recognition. In Proceedings of the 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, USA, 5–8 August 2018; pp. 104–109.

62. Wang, T.; Xie, Z.; Li, Z.; Jin, L.; Chen, X. Radical aggregation network for few-shot offline handwritten Chinese character recognition. *Pattern Recognit. Lett.* **2019**, *125*, 821–827. [CrossRef]

63. Li, Z.; Wu, Q.; Xiao, Y.; Jin, M.; Lu, H. Deep matching network for handwritten Chinese character recognition. *Pattern Recognit.* **2020**, *107*, 107471. [CrossRef]

64. Li, Z.; Xiao, Y.; Wu, Q.; Jin, M.; Lu, H. Deep template matching for offline handwritten Chinese character recognition. *J. Eng.* **2020**, *2020*, 120–124. [CrossRef]

65. Cao, Z.; Lu, J.; Cui, S.; Zhang, C. Zero-shot handwritten Chinese character recognition with hierarchical decomposition embedding. *Pattern Recognit.* **2020**, *107*, 107488. [CrossRef]

66. Chen, J.; Li, B.; Xue, X. Zero-shot Chinese character recognition with stroke-level decomposition. *arXiv* **2021**, arXiv:2106.11613.

67. Zhong, Z.; Jin, L.; Xie, Z. High performance offline handwritten chinese character recognition using googlenet and directional feature maps. In Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR), Tunis, Tunisia, 23–26 August 2015; pp. 846–850.

68. Wang, Y.; Li, X.; Liu, C.; Ding, X.; Chen, Y. An MQDF-CNN hybrid model for offline handwritten Chinese character recognition. In Proceedings of the 2014 14th International Conference on Frontiers in Handwriting Recognition, Hersonissos, Greece, 1–4 September 2014; pp. 246–249.

69. Liu, L.; Sun, W.; Ding, B. Offline handwritten Chinese character recognition based on DBN fusion model. In Proceedings of the 2016 IEEE International Conference on Information and Automation (ICIA), Ningbo, China, 1–3 August 2016; pp. 1807–1811.

70. Zhang, X.Y.; Bengio, Y.; Liu, C.L. Online and offline handwritten chinese character recognition: A comprehensive study and new benchmark. *Pattern Recognit.* **2017**, *61*, 348–360. [CrossRef]

71. Li, F.; Shen, Q.; Li, Y.; Parthaláin, N.M. Handwritten Chinese character recognition using fuzzy image alignment. *Soft Comput.* **2016**, *20*, 2939–2949. [CrossRef]

72. Xu, L.; Wang, Y.; Li, X.; Pan, M. Recognition of handwritten Chinese characters based on concept learning. *IEEE Access* **2019**, *7*, 102039–102053. [CrossRef]

73. Yu, W.; Li, Y.; Peng, H.; Zhang, L. Image iterative method for handwritten Chinese character recognition. In Proceedings of the Journal of Physics: Conference Series, Warsaw, Poland, 2–4 July 2020; Volume 1684, p. 012101.

74. Dong, B.; Varde, A.S.; Stevanovic, D.; Wang, J.; Zhao, L. Interpretable distance metric learning for handwritten chinese character recognition. *arXiv* **2021**, arXiv:2103.09714.

75. Wang, S.; Chen, L.; Xu, L.; Fan, W.; Sun, J.; Naoi, S. Deep knowledge training and heterogeneous CNN for handwritten Chinese text recognition. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; pp. 84–89.

76. Wu, Y.C.; Yin, F.; Liu, C.L. Improving handwritten Chinese text recognition using neural network language models and convolutional neural network shape models. *Pattern Recognit.* **2017**, *65*, 251–264. [CrossRef]

77. Wang, Z.X.; Wang, Q.F.; Yin, F.; Liu, C.L. Weakly supervised learning for over-segmentation based handwritten Chinese text recognition. In Proceedings of the 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), Dortmund, Germany, 8–10 September 2020; pp. 157–162.

78. Peng, D.; Jin, L.; Wu, Y.; Wang, Z.; Cai, M. A fast and accurate fully convolutional network for end-to-end handwritten Chinese text segmentation and recognition. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 25–30.

79. Peng, D.; Jin, L.; Ma, W.; Xie, C.; Zhang, H.; Zhu, S.; Li, J. Recognition of Handwritten Chinese Text by Segmentation: A Segment-annotation-free Approach. *IEEE Trans. Multimed.* **2022**. [CrossRef]

80. Hu, S.; Wang, Q.; Huang, K.; Wen, M.; Coenen, F. Retrieval-based language model adaptation for handwritten Chinese text recognition. *Int. J. Doc. Anal. Recognit. (IJDAR)* **2022**, 1–11. [CrossRef]

81. Su, T.H.; Zhang, T.W.; Guan, D.J.; Huang, H.J. Off-line recognition of realistic Chinese handwriting using segmentation-free strategy. *Pattern Recognit.* **2009**, *42*, 167–182. [CrossRef]

82. Thireou, T.; Reczko, M. Bidirectional long short-term memory networks for predicting the subcellular localization of eukaryotic proteins. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2007**, *4*, 441–446. [CrossRef]

83. Shi, B.; Bai, X.; Yao, C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 2298–2304. [CrossRef] [PubMed]

84. Graves, A. Connectionist temporal classification. In *Supervised Sequence Labelling with Recurrent Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 61–93.

85. Chorowski, J.K.; Bahdanau, D.; Serdyuk, D.; Cho, K.; Bengio, Y. Attention-based models for speech recognition. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 577–585.

86. Messina, R.; Louradour, J. Segmentation-free handwritten Chinese text recognition with LSTM-RNN. In Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR), Tunis, Tunisia, 23–26 August 2015; pp. 171–175.

87. Shen, X.; Messina, R. A method of synthesizing handwritten chinese images for data augmentation. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; pp. 114–119.

88. Graves, A.; Schmidhuber, J. Offline handwriting recognition with multidimensional recurrent neural networks. *Adv. Neural Inf. Process. Syst.* **2008**, *21*, 545–552.

89. Bluche, T.; Messina, R. Faster segmentation-free handwritten Chinese text recognition with character decompositions. In Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 23–26 October 2016; pp. 530–535.

90. Liu, C.L.; Lin, J.H. Using structural information for identifying similar Chinese characters. In Proceedings of the ACL-08: HLT, Short Papers, Columbus, OH, USA, 15–20 June 2008; pp. 93–96.

91. Sacher, H. Interactions in Chinese: Designing interfaces for Asian languages. *Interactions* **1998**, *5*, 28–38. [CrossRef]

92. Wu, Y.C.; Yin, F.; Chen, Z.; Liu, C.L. Handwritten Chinese Text Recognition Using Separable Multi-Dimensional Recurrent Neural Network. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; Volume 1, pp. 79–84.

93. Du, J.; Wang, Z.R.; Zhai, J.F.; Hu, J.S. Deep neural network based hidden Markov model for offline handwritten Chinese text recognition. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 3428–3433.

94. Wang, Z.R.; Du, J.; Hu, J.S.; Hu, Y.L. Deep convolutional neural network based hidden markov model for offline handwritten Chinese text recognition. In Proceedings of the 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR), Nanjing, China, 26–29 November 2017; pp. 816–821.

95. Wang, Z.R.; Du, J.; Wang, W.C.; Zhai, J.F.; Hu, J.S. A comprehensive study of hybrid neural network hidden Markov model for offline handwritten Chinese text recognition. *Int. J. Doc. Anal. Recognit. (IJDAR)* **2018**, *21*, 241–251. [CrossRef]

96. Wang, Z.R.; Du, J.; Wang, J.M. Writer-aware CNN for parsimonious HMM-based offline handwritten Chinese text recognition. *Pattern Recognit.* **2020**, *100*, 107102. [CrossRef]

97. Wang, Z.R.; Du, J. Fast writer adaptation with style extractor network for handwritten text recognition. *Neural Netw.* **2022**, *147*, 42–52. [CrossRef]

98. Wang, Y.; Yang, Y.; Ding, W.; Li, S. A Residual-Attention Offline Handwritten Chinese Text Recognition Based on Fully Convolutional Neural Networks. *IEEE Access* **2021**, *9*, 132301–132310.

99. Xiu, Y.; Wang, Q.; Zhan, H.; Lan, M.; Lu, Y. A Handwritten Chinese Text Recognizer Applying Multi-level Multimodal Fusion Network. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 1464–1469.

100. Zhang, X.; Yan, K. An Algorithm of Bidirectional RNN for Offline Handwritten Chinese Text Recognition. In Proceedings of the Intelligent Computing Methodologies, Nanchang, China, 3–6 August 2019; Lecture Notes in Computer Science; Huang, D.S., Huang, Z.K., Hussain, A., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 423–431. ._39. [CrossRef]

101. Xie, C.; Lai, S.; Liao, Q.; Jin, L. High Performance Offline Handwritten Chinese Text Recognition with a New Data Preprocessing and Augmentation Pipeline. In Proceedings of the Document Analysis Systems, Wuhan, China, 26–29 July 2020; Lecture Notes in Computer Science; Bai, X., Karatzas, D., Lopresti, D., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 45–59. ._4. [CrossRef]

102. Hoang, H.T.; Peng, C.J.; Tran, H.V.; Le, H.; Nguyen, H.H. LODENet: A Holistic Approach to Offline Handwritten Chinese and Japanese Text Line Recognition. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 4813–4820.

103. Ngo, T.T.; Nguyen, H.T.; Ly, N.T.; Nakagawa, M. Recurrent neural network transducer for Japanese and Chinese offline handwritten text recognition. In Proceedings of the International Conference on Document Analysis and Recognition, Lausanne, Switzerland, 5–10 September 2021; pp. 364–376.

104. Zhan, H.; Lyu, S.; Lu, Y. Improving offline handwritten Chinese text recognition with glyph-semanteme fusion embedding. *Int. J. Mach. Learn. Cybern.* **2022**, *13*, 485–496. [CrossRef]

105. Zhu, Z.Y.; Yin, F.; Wang, D.H. Attention Combination of Sequence Models for Handwritten Chinese Text Recognition. In Proceedings of the 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), Dortmund, Germany, 8–10 September 2020; pp. 288–294.

106. Liu, B.; Sun, W.; Kang, W.; Xu, X. Searching from the Prediction of Visual and Language Model for Handwritten Chinese Text Recognition. In Proceedings of the International Conference on Document Analysis and Recognition, Lausanne, Switzerland, 5–10 September 2021; pp. 274–288.

107. Yang, H.; Jin, L.; Sun, J. Recognition of chinese text in historical documents with page-level annotations. In Proceedings of the 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, USA, 5–8 August 2018; pp. 199–204.

108. Yang, H.; Jin, L.; Huang, W.; Yang, Z.; Lai, S.; Sun, J. Dense and tight detection of chinese characters in historical documents: Datasets and a recognition guided detector. *IEEE Access* **2018**, *6*, 30174–30183. [CrossRef]

109. Xie, Z.; Huang, Y.; Jin, L.; Liu, Y.; Zhu, Y.; Gao, L.; Zhang, X. Weakly supervised precise segmentation for historical document images. *Neurocomputing* **2019**, *350*, 271–281. [CrossRef]

110. Ma, W.; Zhang, H.; Jin, L.; Wu, S.; Wang, J.; Wang, Y. Joint layout analysis, character detection and recognition for historical document digitization. In Proceedings of the 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), Dortmund, Germany, 8–10 September 2020; pp. 31–36.

111. Su, T.; You, H.; Liu, S.; Wang, Z. FPRNet: End-to-End Full-Page Recognition Model for Handwritten Chinese Essay. In Proceedings of the International Conference on Frontiers in Handwriting Recognition, Hyderabad, India, 4–7 December 2022; pp. 231–244.

112. Kundu, S.; Paul, S.; Bera, S.K.; Abraham, A.; Sarkar, R. Text-line extraction from handwritten document images using GAN. *Expert Syst. Appl.* **2020**, *140*, 112916. [CrossRef]

113. Yan, S.; Wu, J.W.; Yin, F.; Liu, C.L. Recognizing Handwritten Chinese Texts with Insertion and Swapping Using a Structural Attention Network. In Proceedings of the International Conference on Document Analysis and Recognition, Lausanne, Switzerland, 5–10 September 2021; pp. 557–571.

114. Peng, D.; Jin, L.; Liu, Y.; Luo, C.; Lai, S. PageNet: Towards End-to-End Weakly Supervised Page-Level Handwritten Chinese Text Recognition. *Int. J. Comput. Vis.* **2022**, *130*, 2623–2645. [CrossRef]