

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR



TRABAJO FIN DE MÁSTER

**CHINESE CALLIGRAPHY:
CHARACTER STYLE RECOGNITION BASED
ON FULL-PAGE DOCUMENT**

**Máster Universitario en
Investigación e Innovación en TIC**

Autor: Wen, Yuanbo

Tutor: Sigüenza Pizarro, Juan Alberto

**Departamento de Ingeniería Informática,
Escuela Politécnica Superior,
Universidad Autónoma de Madrid**

FECHA: Noviembre, 2018

Acknowledgements

Before the beginning of the thesis, I want to thank my tutor Professor Juan Alberto Sigüenza. My tutor gave me a lot of help on this road since knowing him till the completion of the thesis. I used to have various problems, including language and the pressure of the course. When I wanted to give it up, my tutor used his wisdom and patience to give me the courage to complete the thesis. He patiently accompanied me to complete the thesis and shared various ideas to guide me do the research.

In addition, I would like to thank Professor Manuel Sánchez-Montañés Isla for helping me to improve my thesis and code patiently. So many times, we discussed my thesis after his classes and guide me how to learn deep learning with Python. Also, Professor Luis Lago give us his courses of machine learning which is the key to complete my thesis.

Finally, I want to thank my family and my friend Wenji Duo, who have supported me throughout the process and I will always be grateful for your help.

Abstract

Calligraphy plays a very important role in the history of China. From ancient times to modern times, the beauty of calligraphy has been passed down to the present. Different calligraphy styles and structures have made calligraphy a beauty and embodiment in the field of writing. However, the recognition of calligraphy style and fonts has always been a blank in the computer field. The structural complexity of different calligraphy also brings a lot of challenges to the recognition technology of computers. In my research, I mainly discussed some of the main recognition techniques and some popular machine learning algorithms in this field for more than 20 years, trying to find a new method of Chinese calligraphy styles recognition and exploring its feasibility.

In our research, we searched for research papers 20 years ago. Most of the results are about the content recognition of modern Chinese characters. At first, we analyze the development of Chinese characters and the basic Chinese character theory. In the analysis of the current recognition of Chinese characters (including handwriting online and offline) in the computer field, it is more important to analyze various algorithms and results, and to analyze how to use the experimental data, besides how they construct the data set used for their test.

The research on the method of image processing based on Chinese calligraphy works is very limited, and the data collection for calligraphy test is very limited also. The test of dataset that used between different recognition technologies is also very different. However, it has far-reaching significance for inheriting and carrying forward the traditional Chinese culture. It is very necessary to develop and promote the recognition of Chinese characters by means of computer technique. In the current application field, the font recognition of Chinese calligraphy can effectively help the library administrators to identify the problem of the classification of the copybook, thus avoiding the recognition of the calligraphy font which is difficult to perform manually only through subjective experience.

In the past 10 years of technology, some techniques for the recognition of single Chinese calligraphy fonts have been given. Most of them are the pre-processing of calligraphy characters, the extraction of stroke primitives, the extraction of style

features, and the final classification of machine learning. The probability of the classification of the calligraphy works. Such technical requirements are very large for complex Chinese characters, the result of splitting and recognition is very large, and it is difficult to accurately divide many complex font results. As a result, the recognition rate is low, or the accuracy of recognition of a specific word is high, but the overall font recognition accuracy is low.

We understand that Chinese calligraphy is a certain research value. In the field of recognition, many research papers on the analysis of Chinese calligraphy are based on the study of calligraphy and stroke. However, we have proposed a new method for dealing with font recognition. The recognition technology is based on the whole page of the document. It is studied in three steps: the first step is to use Fourier transform and some Chinese calligraphy images and analyze the results. The second is that CNN is based on different data sets to get some results. Finally, we made some improvements to the CNN structure. The experimental results of the thesis show that the full-page documents recognition method proposed can achieve high accuracy with the support of CNN technology, and can effectively identify the different styles of Chinese calligraphy in 5 styles. Compared with the traditional analysis methods, our experimental results show that the method based on the full-page document is feasible, avoiding the cumbersome font segmentation problem. This is more efficient and more accurate.

Contents

1. Introduction	1
2. Related works and State of art	3
3. Objectives	14
4. Data	16
4.1. Data preparation	16
4.1.1 Dataset_A	16
4.1.2 Dataset_B	20
4.2. Dataset and Pre-processing	21
5. Technique	22
5.1. Method with Fourier transform	22
5.1.1 Analysis with Fourier transform	22
5.1.2 Analysis with Euclidean distance	25
5.2. Method with Convolutional Neural Network	32
5.2.1 Feature Learning, Layers, and Classification	36
5.2.2 Classification Layers	38
6. Results	39
6.1. The first experimental protocol	39
6.1.1 Experiment and results	39
6.1.2 Evaluation	41
6.2. The second experimental protocol	44
6.2.1 Experiment and results	44
6.2.2 Evaluation	45
6.3. The third experimental protocol	46
6.3.1 Experiment and results	47
6.3.2 Evaluation	50
7. Discussion	52
8. Conclusion	56
9. Future work	58
REFERENCES	59

1.Introduction

In recent years, with the rapid development of computers, pattern recognition technology has made continuous progress, greatly improving the information interaction capability between human and machine. Computer text recognition is an important branch of pattern recognition, including digital character recognition, Western character recognition, and oriental character recognition. The research on character recognition in numbers and English is the most comprehensive, the recognition method is mature, and accuracy of the recognition is very high. In the course of research, it was found that the recognition technique of Oriental characters is relatively poor, and the relative recognition accuracy is low and the limitations are very large. The identification of Chinese characters is also relatively difficult.

After the development of Chinese characters in the past few years, they are divided into two major categories. As we know the study of handwritten Chinese character recognition (HCCR) for more than 50 years [1, 2]. The first category is the research that includes different calligrapher. They analysis the results and do some experiments with the recognition of differences handwriting-style, they use the data of the handwriting but the same Chinese characters. The second major category is that based on the type of Experimental data, handwriting recognition can be divided into online and offline [3,4]. Online handwritten Chinese characters are based on recording and analyzing the trajectory of the pen tip movement to dynamically identify and judge the problems and possible probabilities that need to be input, and finally obtain the output results. Such methods are mostly used in handwriting recognition systems of self-calculations and in current mobile communication devices [5,6,7]. The standard of measurement to this day, more is to judge the accuracy of the input and the speed of recognition. However, the other category is offline HCCR, which is a static text analysis, most of which need to be classified into image recognition classification. Its applications are also more extensive, including signature recognition in emails and some electronic signature recognition of bank checks. China has organized three competitions, the National Institute of Model Identification (NLPR) of the Institute of Automation of the Chinese Academy of Sciences (CASIA) in CCPR-2010 [8], ICDAR-2011 [9], and ICDAR-2013 [10]. The data of these several games are also presented in the following table for your reference. The database it uses is also relatively uniform and standardized.

In the course of the research, we also found that if we only follow the previous research results, there is not much content that can be referenced, and there are few Chinese characters based

on the entire document to identify the Chinese character style. In the research process of this article, we explored some basic methods and compared whether it is possible to identify Chinese characters based on the full-page document, and it is aimed at the recognition of Chinese calligraphy. This is a big challenge.

These include some basic issues that need to be resolved in advance, such as the establishment of the database, because the test data used is not the same, so the database used in many articles which is not available in our research. In addition, we only target the recognition of the main Chinese calligraphy characters in five styles, and there are not many reference texts or characters that can be preserved in the ancient books of Chinese calligraphy. More are more works by modern people based on past historical books. They make the copy of the style of the character which is similar to that of ancient calligraphy writers. And you also can find the preserved Chinese characters, there is a lot of clarity is not enough, and so on, these series of questions will be the problems that we need to solve and explore later.

2.Related works and State of art

Calligraphy is a traditional art unique to China. Chinese characters are created by the working people which began to use picture books. After years of professional bookmakers' achievements, they have been continuously improved and developed into Chinese characters. The processing and refining of various artistic techniques has produced calligraphy. Throughout the ages, Chinese characters have been written on specific rice papers with a brush, and they are required to be in one go [11,12]. They must never be modified or written back, and they must be accurate, beautiful and conform to the requirements of the basic structure of Chinese characters. In other words, calligraphy refers to writing in the style, structure and chapters of the book according to the characteristics of the text and its meaning, making it an aesthetically pleasing work of art. Especial we heard of seal carving, ornate paperweights, and inkstones.

As we known, "Chinese calligraphy", is mainly divided into "soft pen calligraphy" and "hard pen calligraphy", which is a unique traditional art and culture in China. Throughout the ages, Chinese characters were created by the Chinese ancient calligrapher. After that, the ancestors wrote the Chinese characters they thought were beautiful on the sand. This was the earliest calligraphy, and the ancestors invented the writing with a brush to create calligraphy. Throughout the ages, it has been written with brush writing. As for other writing forms, such as hard pens and finger-books, the writing rules are not completely different from the writing rules of writing brushes, but they are basically the same. Based on Suo Ben's source (after the nail bones), here is a focus on the rules of writing Chinese calligraphy. [13,14].

We know that Chinese characters and languages have a history of more than 4,000 years of development. The main Chinese characters that have been passed down to the present have evolved from the beginning to the five basic types [15]. They are also the main Chinese learning and copying five types of characters. traditional character we have scripts, scripts, scripts, cursive fonts, and five types of calligraphy. There are several main categories of Chinese traditional calligraphy:

Xing (semi-cursive), It is the development of the origin of the Li, a type of font between the regular and the cursive, in order to make up for the slow writing speed of the script and the difficulty in recognizing the cursive. The style of Xing is not smooth as Li. In other words, it is a variant of the Kai [16,17].

Cao (cursive), Formed in the Han Dynasty, it was developed for the convenience of writing on the basis of Li. Among the five styles, cursive is the most abstract artistic trait [18,19].

Zhuan (seal), Preserving the obvious features of ancient hieroglyphics [20].

Li (clerical), Evolved from the script of Zhuan. It is much more convenient to write than the script, saving a lot of valuable time for future generations, and it is also of great value in academics. Li originated from the Qin Dynasty and reached its peak in the Eastern Han Dynasty [21,22].

Kai (regular), The script is also called Zhen Shu or Zheng Shu. It gradually evolved from the Li, and it became more simplified. The glyphs were changed from flat to flat, and the strokes of Han Li were saved in the strokes [23,24].

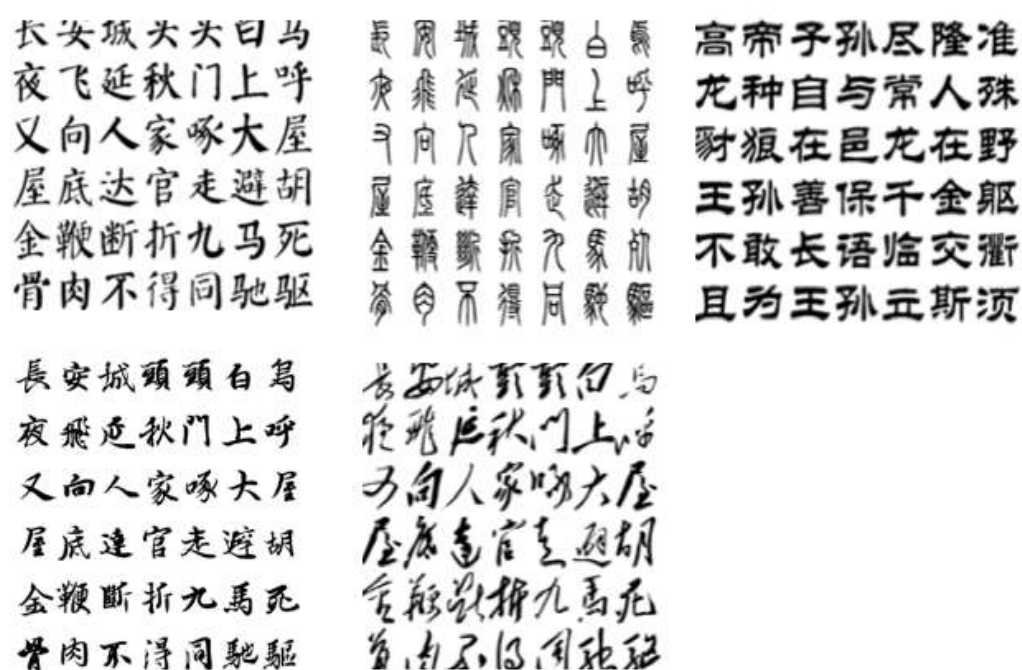


Figure 1. First row (from left to right: Kai, Li, Zhuan)
Second row (from left to right: Xing, Cao)

The four most important items in Chinese calligraphy are brush-pen, ink, paper and inkstone. It is the main carrier of the 5000-year inheritance of Chinese culture. It has always been praised by Chinese literati. Call them treasures, not because of rarity and price, but because of the book, the poems are made.

-Brush pen

The brush pen refers to the Hu-pen. (Huzhou, Zhejiang Province) Everyone knows that the pen is used for writing records. In the era when there was no technology, the ancients used the method of engraving and knotting. With the development of the times, the brush has gradually appeared, and the brush can be divided into hard, soft, and equal. It is also very skillful when choosing a brush. The tip of the pen should be sharp and neat.

-Ink

Ink refers to Huimo, (Huizhou, now Anhui Province) is a unique art in China. It has a history of more than 3,000 years. There are many kinds of inks. The more famous ones are loose smoke ink, oil smoke ink and mixed ink. Because of the differences in process and specific materials between different inks, the effect of painting and calligraphy is very different. It is precisely this situation that has also created the uniqueness of the changes in the ink paintings of Chinese paintings.

-Paper

Paper refers to rice paper, (now in Anhui Province, the ancient county of Ningguo Prefecture in Gifu Prefecture, the name of the paper is in the name of Fucheng Xuancheng). Currently, the papers for domestic painting are all rice paper, because it has strong water absorption, often Presenting the best paintings in front of people's eyes, so Xuan paper is the best choice for many calligraphy and painters.

-Inkstone

The platform is also necessary to use. This is the object used to load the ink. And since ancient times, the shape of this object has a lot of rich design, and also as one of the collectors' collections.

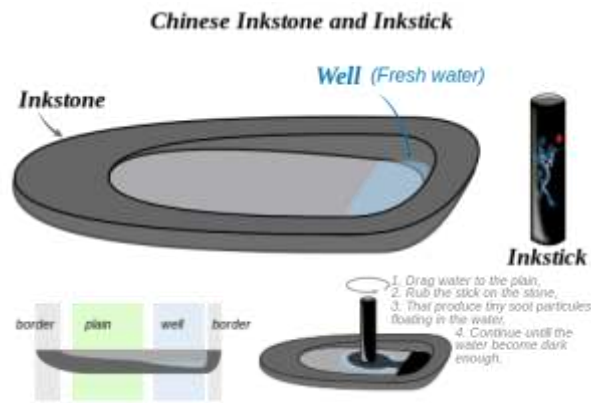


Figure 2. Chinese inkstone and inkstick

Traditionally, the way to obtain calligraphy is usually divided into two methods [25]. The first way, through the writing of ancient inscriptions, engraves the characters, copies them to paper through technical ways, and then repairs the missing fonts by technical processing. Other way is for people learn by the example of modern calligraphers. This is also one of the sources of the main fonts. Because the past inscriptions damaged the subspecies, it also added a lot of difficulty to the repair work, and also made the characters retained to modern times not very complete. But by analyzing the correct strokes, stroke order, and font results, people have also imitated many other font art.



Figure 3. Copybook (for calligraphy)

The second is through the ancient calligraphy, the five major styles are also the main five major Character, through continuous imitation and connection, generate more than 5,000 major Chinese characters in Chinese characters, thus Dada learned and tried to figure out [26,27]. Most of these copybooks are modernly generated, consistent with the basic style of preservation that has been handed down in ancient times. It is also used by most people as a sample of research.



Figure 4. Copybook (handwriting Practice)

In recent research, more and more paper research has created robotic arms by imitating calligraphy fonts and improved the calligraphy integrity of the font. This research can generate more font libraries on the basis of the future. However, to provide the degree of imitation of fonts through machine learning, it is largely dependent on the classification and recognition technology of calligraphy fonts, and by continuously learning the existing font system, the generation of different fonts that imitate new Chinese characters is generated. It is also a very meaningful but challenging task [28,29,30].

In the beginning of studying Chinese calligraphy Chinese characters, we first reviewed the Chinese character recognition research in the past 20 years [31,32]. It is found that the technology has mainly focused on the single recognition of Chinese characters, and it is mainly divided into two parts.

The first part is the segmentation of Chinese characters. The segmentation of Chinese characters is a necessary basis for recognition, and can greatly improve the accuracy of the recognition rate. Through research, we found that the technology of splitting mainly focuses on the following three methods.

In the past 20 years, the main technology for the segmentation, which are as below:

Publication	Method	Dataset	Accuracy(%)
2005 Xianghui wei etc.[33]	Genetic algorithm	428 images, each of two Chinese characters	88.9
2005 Zhizhen Liang etc.[34]	Metasynthetic	921 postal strings(7913 characters)	87.6
2001 Shuyan Zhao etc.[35]	Two-stage	1000 strings containing 8011 characters	81.6
1999 Yi-Hong Tseng etc.[36]	Recognition-based	125 text-line images containing 1132 characters	95.58
1998 Lin Yu Tseng etc.[37]	Heuristic merging,DP	900 handwriting and 100 printing	95

Table 1 Some result of segmentation

1. The segmentation of unconstrained handwritten

Unconstrained handwritten text line segmentation method to accurately segment the text lines. There are three common situations in Chinese text lines for natural writing: naturally separated characters, overlapping characters, and glue characters are processed separately. For naturally separated characters, they are segmented by histogram projection [38]; for overlapping characters, they are segmented by setting the isolation points in the overlapping area; for the sticky characters, the adhesion points are first detected, and then Split the character at the glue point and finally split it as an overlapping character [39]. The method of the invention can ensure high segmentation accuracy while realizing fast segmentation of text lines.

2. Coarse segmentation

The core of the coarse segmentation technique is to generate appropriate segmentation paths and reduce inappropriate paths based on vertical projections and background cues [40]. In the past, people used the Hilditch refinement algorithm to weaken the background area. Colleagues, also to determine the road between charging characters. The key to segmentation is the need to accurately partition multiple blocks from the generated histogram. The size of the block is the width of the character, and the accuracy of this estimate also affects the correctness.

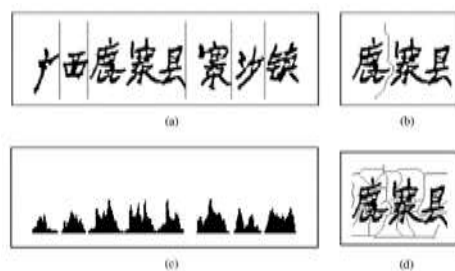


Figure 5. "Coarse segmentation. (a) The paths identified by vertical projection, (b) a path generated by tracing the background skeleton, (c) vertical projection, (d) background skeleton." [40]

3. Fine segmentation

The fine segmentation method is used after the coarse segmentation [41,42]. This is the candidate mural and segmentation point for the segmentation point based on the rough segmentation. The fine style path and mode are presented below.

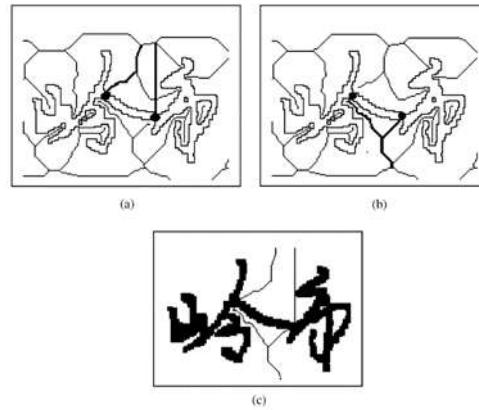


Figure 6. “Examples of fine segmentation paths. (a) Upper-Parts (dark segments), (b) Lower-Parts (dark segments), (c) fine segmentation paths.” [42]

In the past 20 years, other the main technology for the recognition, which are as below:

Publication	Method	Dataset	Accuracy(%)
2005 Jia Zeng etc.[40]	MRF	9 pairs of similar Chinese in ETL9B	92
2003 Feng Lin etc.[41]	Statistical Character Structure Modeling	KAIST Hanja1 and Hanja2	98
2001 Cheng-Lin Liu etc.[42]	Model-based stroke extraction and matching	783 characters	97
2000 Qing Wang etc.[43]	HMMRF	470 Chinese from 3775	88

Table 2. Some results of stroke extraction in literature [43]

1. Stroke-Based Approaches

In the past literature, there are many ways to extract strokes, such as the description method of the stroke font using dot matrix Chinese characters, and the algorithm based on this description method. Or, according to the composition of different strokes of a single Chinese character, each stroke feature is extracted, and matching recognition is performed according to this feature. The specific process is as follows:

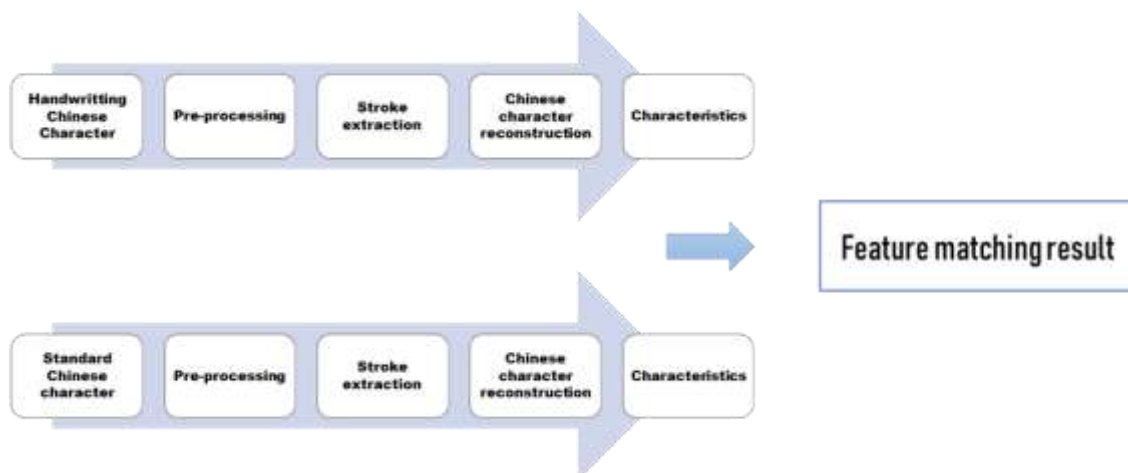


Figure 7. Handwritten Chinese character recognition process

2. The radical-based approach

This method is a new method for online handwritten Chinese character recognition. Different from other methods, this method is a new attempt for the statistical classification of free radicals and the excessive segmentation of characters into candidate free radicals. It can be said that this method can make good use of the Chinese characters of the left and right structures, and also extends to other structures. The results obtained also confirm the superiority of this method.



Figure 8. Examples of similar characters (tomb, dusk, curtain and subscription) decomposed into radicals

3. Holistic approaches

The holistic approaches consider a whole piece of data as wavelet texture analysis for extracting font features for classification; using filters to extract global text features of texts as classification features; using wavelet packets for multi-level decomposition to extract font texture features; Extract features using manual experience. It can be seen that these methods have an important drawback. If the entire layout of a text contains different font information, the overall analysis method cannot be used in this scenario.

Year	Author	Features	Classifier
2006	Fu Chang [44]	Vector feature	Multi-Trees, Sub-Vec, SVM
2005	Hailong Liu etc [51]	Gradient feature	Multi-scheme enhanced MQDF
2005	Jian-xiong Dong [52]	Directional Feature	SVM
2002	Rui Zhang etc [46]	Directional Features	MCE enhanced MQDF
2002	Yong Ge etc [49-51]	Gabor features and their spatial derivatives	CDHMM
2002	Bing Feng etc [53]	Five features	HMM
2001	Yan Xiong etc [47]	Five global features	HMM
2000	Mingrui Wu etc [52]	Directional line element feature	Contextual stochastic model
2000	C.H Wang etc [53]	Three kinds of statistical feature	Single CSN network

Table 3. holistic-based recognition of offline handwritten Chinese characters

After reviewing the Chinese handwriting recognition technology in the past 20 years, referring to the results of some Chinese handwriting recognition contests, the accuracy of different technologies for offline and online Chinese character recognition is summarized in the following table. The name is 12th International Conference on Document Analysis and Recognition (ICDAR 2013) [54,55]. There is a good advantage in the competition, the standards of the database used by all technologies are the same. Based on the new databases CASIA-OLHWDB and CASIA-HWDB [56,57]. The following table summarizes the technical and competition results for offline and online.

Results on ICDAR-2013 Offline HCCR Competition Database			
No.	Method	Accuracy	Training Data
1	Human Performance	96.13%	N/A
2	Traditional Method: DFE + DLQDF	92.72%	1.0+1.1
3	CCPR-2010 Winner: HKU	89.99%	1.0+1.1
4	ICDAR-2011 Winner: IDSIAAnn-2	92.18%	1.1
5	ICDAR-2013 Winner: Fujitsu	94.77%	1.1
6	Multi-Column DNN (MCDNN)	95.79%	1.1
7	ATR-CNN Voting	96.06%	1.1
8	HCCR-Gabor-GoogLeNet	96.35%	1.0+1.1
9	HCCR-Ensemble-GoogLeNet-4	96.64%	1.0+1.1
10	HCCR-Ensemble-GoogLeNet-10	96.74%	1.0+1.1
11	CNN-Single	96.58%	1.0+1.1+1.2
12	CNN-Voting-5	96.79%	1.0+1.1+1.2
13	DirectMap + ConvNet ours	96.95%	1.0+1.1
14	DirectMap + ConvNet + Adaptation	97.37%	1.0+1.1
15	DirectMap + ConvNet + Ensemble-2 ours	97.07%	1.0+1.1
16	DirectMap + ConvNet + Ensemble-3	97.12%	1.0+1.1

Table 4. Different methods for ICDAR-2013 offline HCCR competition [58]

Results on ICDAR-2013 Online HCCR Competition Database			
No.	Method	Accuracy	Training Data
1	Human Performance	95.19%	N/A
2	Traditional Method: DFE + DLQDF	95.31%	1.0+1.1
3	CCPR-2010 Winner: SCUT-HCII-2	92.39%	1.0+1.1
4	ICDAR-2011 Winner: VO-3	95.77%	1.0+1.1+Others
5	ICDAR-2013 Winner: Uwarwick	97.39%	1.0+1.1+1.2
6	DropSample-DCNN	97.23%	1.0+1.1
7	DropSample-DCNN-Ensemble	97.51%	1.0+1.1
8	DirectMap + ConvNet	97.55%	1.0+1.1
9	DirectMap + ConvNet + Adaptation	97.91%	1.0+1.1
10	DirectMap + ConvNet + Ensemble-2	97.60%	1.0+1.1
11	DirectMap + ConvNet + Ensemble-3	97.64%	1.0+1.1

Table 5. Different methods for ICDAR-2013 online HCCR competition [58]

Through the development of Chinese character recognition technology, whether it is handwritten or printed technology, the recognition accuracy in this field has been very mature [59,60]. With the development of modern electronic equipment, we can see various input methods applied to computers, banks, tablet and various voice input recognition [61,62,63]. The recognition technology for handwritten signatures is also becoming more and more accurate. Based on this research, we also noticed a branch, which is the recognition technology of Chinese calligraphy characters [64,65,66]. Therefore, according to this request, we have consulted the questions of the recent 20 years, but unfortunately, there is not much information, and there are fewer documents that can be referenced [67,68,69,70,71,72]. The following articles are compiled in the paper.

Year	Author.[Ref]	Name	Method	Accuracy(%)
2016	Chen, Z , Yu-Sheng [73]	Machine Learning for Calligraphy Styles Recognition	Single Character	78.5
2016	Li Boqi [9]	Convolution Neural Network for Traditional Chinese	Single Character	88.6
2017	Zhang Jiulong; Guo Luming; Yang Su; Sun Xudong; Li Xiaoshan [74]	Detecting Chinese calligraphy style consistency by deep learning and one-class SVM	Single Character with Calligraphy Font	85.5
2017	Gao Pengcheng, Gu Gang, Wu Jiangqin, Wei Baogang [75]	Chinese calligraphic style representation for recognition	Single Character	94.22

Table 6. references for the recognition of Calligraphy styles

Chen, Z, et al. [73] a variety of classification problems based on machine learning, the identification of individual Chinese characters, including Softmax Regression, Support Vector Machine (SVM), k-Nearest Neighbors (kNN) and Random Forests. Based on these techniques to compare the advantages and disadvantages of each technology and Finally, the technology with relatively high recognition rate is summarized. The contribution of this document can refer to more identification technologies, improve their recognition technology according to their needs, and most importantly, the recognition of Chinese characters. All databases are data samples created.

Li Boqi, et al. [9] In this paper, the CNN technology and convolutional neural network in the mainstream image recognition field are adopted. Also chose the mainstream Chinese traditional calligraphy style. They will train convolutional neural networks on five types of data sets. The database used contains 15,000 instances. And built 8 different models for testing. The recognition rate of each font analyzed by evaluating the performance of different models.

Zhang junlong, et al. [74] The paper compares three techniques, including deep network function extraction features, One-Class SVM classification and kNN technology to extract features and classifications, where the test data type is the window font and the data sample

has 5000. The data type also has handwritten calligraphy, and the test results show that the three techniques have different accuracy rates for different data samples.

Gao pengcheng, et al. [75] This paper presents an automatic character style representation method. The three features extracted are global features, GIST and local features, and scale-invariant feature transforms. The most advanced classifier is used to correct the quadratic discriminant function for identification. The data set is an unconstrained the reality calligraphic character dataset (CCD) and SCL (standard calligraphic character library). The left side is the depth feature extracted by the Deep Convolutional Neural Network (CNN). Using MQDF to compare with support vector machines and neural networks, the three classifiers are evaluated, and the best result is deep learning.

3. Objectives

The challenge of computer for calligraphy style lies in the diversity and complexity of this visual art. These factors are visually and artistically present by many unreliable factors. But because of the complexity of this technology, the recognition technology makes more sense.

Based on the analysis of past papers, we see that most of the research is based on the recognition of a single Chinese character. So, we make a hypothesis “Is possible to perform an automatic recognition of Chinese calligraphic styles using full-page documents”.

Our objectives are to find a more efficiency and fast algorithm application with image or text recognition to avoid cumbersome font splitting. And verify the results of the two algorithms and prove that based on the full-page documents is a feasible and effective method.

Based on this aim, we analyzed the existing technical situation and derived objectives into the following three parts:

- Building a database

As we considered, for example, the complexity of the model is also important to classify the style of calligraphy, but more importantly, the acquisition of big data, or build a database. The more data can improve the training accuracy of the model a lot. In the past experiments, we were unable to find data suitable for our research, so the important thing is to establish our own database.

- Frequency analysis

Extending research in the field of image recognition to methods of recognizing text images. In this field, frequency analysis is usually used, which is called the Fourier transform. This way separating the images after high and low frequencies, you can add more methods in extracting features. The technique of Fourier transform is a method we tried in this article, and we hope that we can summarize certain rules.

- Machine learning techniques

Machine learning plays an important role in the classification problem. By extracting features, we can use the various classification methods in machine learning to classify the Chinese Calligraphy. We choose CNN to do the classification in my research.

4. Data

During initial literature review, to be honest, finding the proper database make us confused for a quite long time. We could find so many databases because that all of them are based on individual single Chinese Character. For example, a Calligraphic Character Dictionary called CADAL, as is introduced by Zhang and Nagy [56] Although CADAL contains thousands of calligraphy images, we couldn't choose to use that database. Because the dataset is not created by full-page. All that datasets are prepared to recognize by single Chinese Character. Besides, it is not the Chinese calligraphy that we wanted.

4.1. Data preparation

We prepare the two sets of datasets. We evaluated the performance of each dataset with each model and analyzed the recognition ratio of each style.

4.1.1 Dataset_A

Our aim is to focus on five styles of Chinese calligraphy. As we know, the only way we have to do that create the database by ourselves. After analysis our requirement:

1. To create five styles of Chinese calligraphy
2. To make all the different Chinese calligraphic characters in one page
3. Each style make sure that has more than 1000 pictures (full-page)

To decide using CNN for recognition and classification, at least we need to have 5000 pictures for the dataset. But face on the fact that nobody created them before. So that it become the first challenge to prepare them.

After several times meeting with my tutor. We find out one way to generate those full-page Chinese calligraphy with python automatic.

At the first step, we create full-page Chinese calligraphy from this website www.akuziti.com/mb/. In this website, we can choose the main five styles of Chinese Calligraphic character that we want. Also, we need to give the size of this picture, in that case, we decide to create size 1000 x 1000 pixels with 272 Chinese Characters. Make sure all the characters are different in the full-page. The Figure 9. is the example that we create by the website.

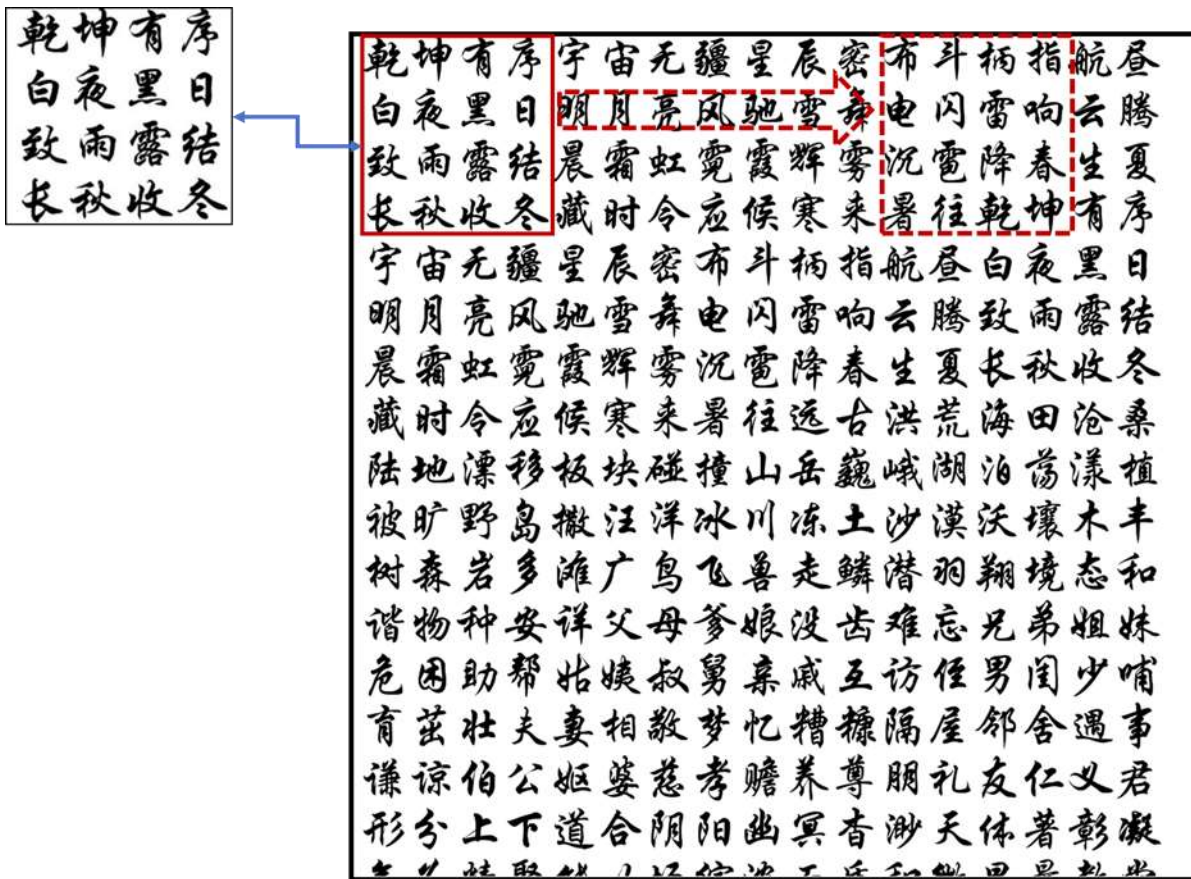


Figure 10. cut full-page document into some pieces

The procedure is below as the Figure 11.

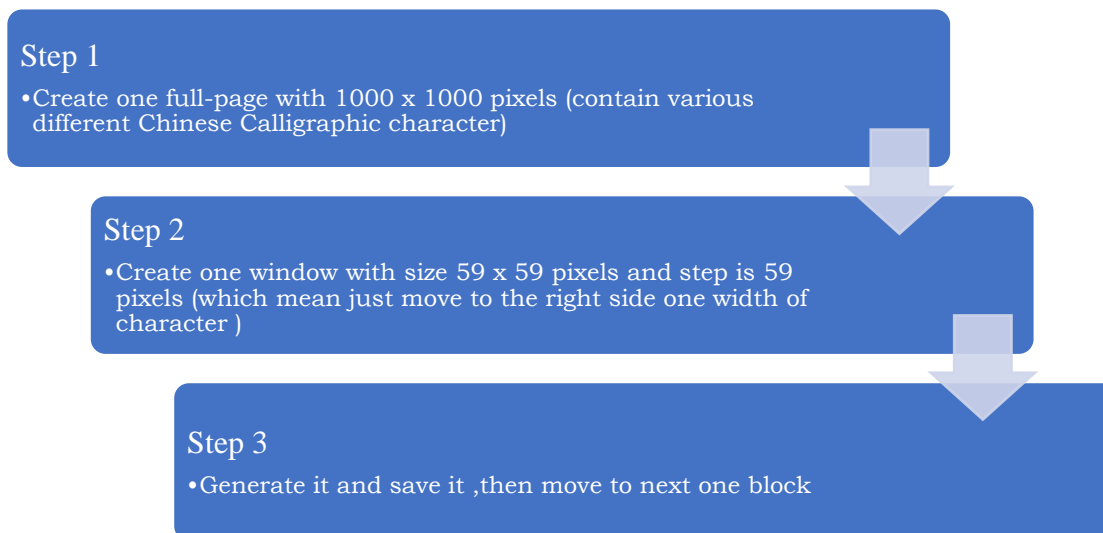


Figure 11. The procedure of generate dataset

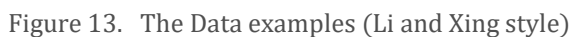
序有乾坤
日黑有夜
结露有雨
冬收有秋
长致有白

et has the following features:

1. Those pictures are the same size.

3. Those pictures without angle rotation.

In Figure 13. we try to make each character is different. In this example, it included 16 single Chinese characters. All of them are different meaning and different script. In order to let CNN can learn more features from this dataset. At least, each one style of Chinese Calligraphy we prepared 1050 pictures. Actually, we only using 1000 pictures for each one style.



4.1.2 Dataset_B

In order to test the CNN performance with complicated pictures. We also prepared another dataset. All this data download from the internet. It has some scanning documents or photos by people take and write. It looks like disorderly and unsystematic. Figure 14. are some examples with Li and Xing style?

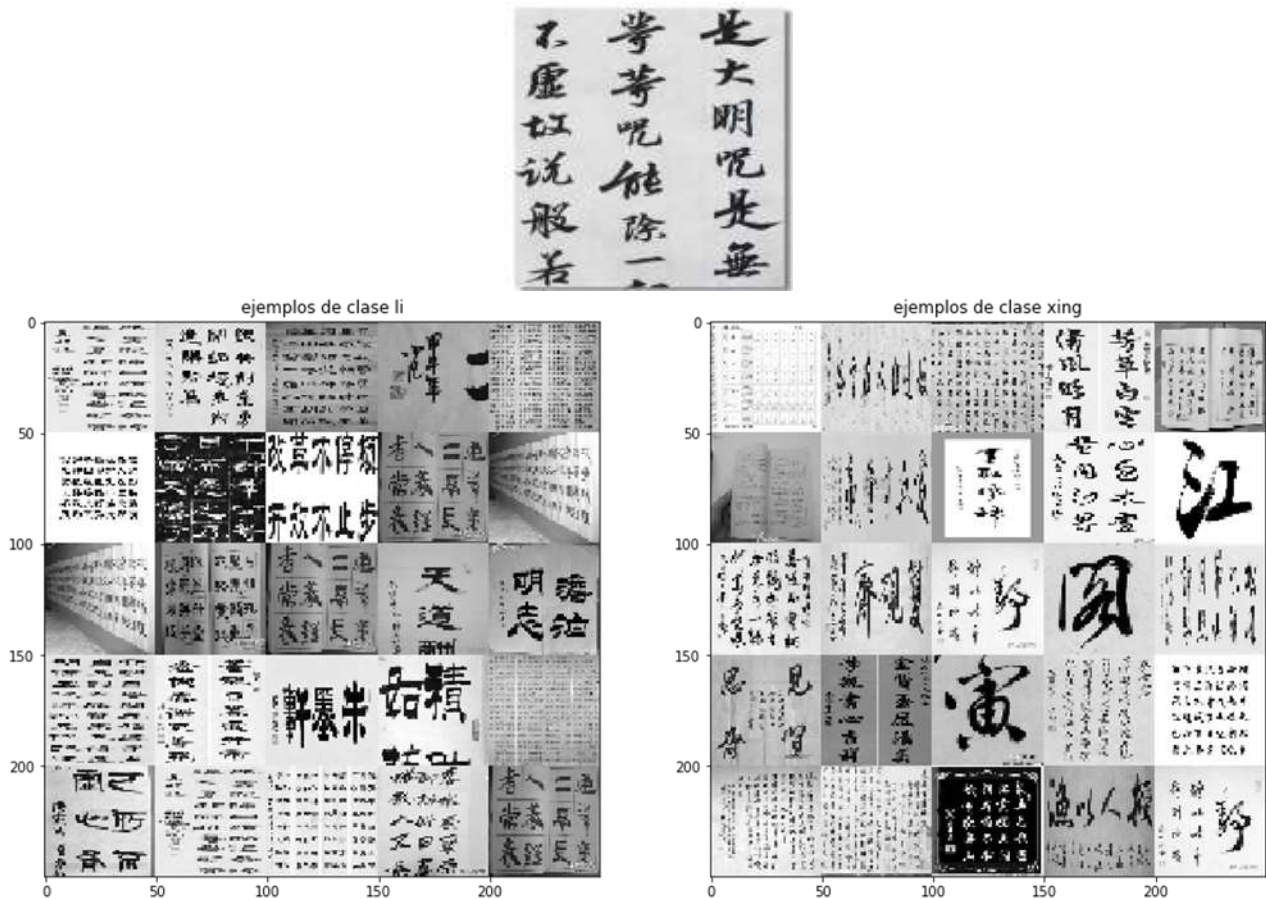


Figure 14. The Data examples (Li and Xing style)

Those datasets contain 500 pictures, with 100 pictures for each style. We can observe that dataset has the following features:

1. Those pictures are disorganized and varied.
2. Those pictures contain inconsistent characters size in full-page
3. Those pictures which background color is also inconsistent.
4. Some pictures with angle rotation.

5. Some pictures with seals or ink stains

Xia et al. [57] suggested that when doing preprocessing on Chinese calligraphy images, man-made seals and background (which sometimes contains ink stains) should be removed.

However, for using CNN, we just want to test if the capability of CNN on handling these conditions. Therefore, we didn't remove any seals or ink stains.

4.2. Dataset and Pre-processing

After we finished all that dataset. We used the dataset which is separated into 3 parts. In machine learning, the data used to train the model is generally divided into the following three parts, so that it can effectively solve the problem of lacking sufficient data to verify the model and evaluate the prediction effect of the model. The commonly used methods are as follows:

The first one is the easiest and easy to understand it. We can divide the dataset into two parts, one part for training and another part for verification, which is the training set and test set we often called.

We know that when more data used for model training, the result usually goes better. So that the resulting model effect will also be affected. Based on this background, someone proposed the cross-validation method, which is cross-validation.

In our work we also prepare to divide the data according to the standard test dataset. Here is a common data set partitioning principle:

- The training set is used to estimate the model and determine the characteristics learned by the model;
- The Validation set is used to determine the network structure or control the complexity of the model parameters, and make the adjustments in time;
- The test set is selected model, and make a performance optimization which can be achieved.

Thus, the typical division should be that the training set accounts for 50% of the total sample, while the others each account for 25%. Three sections were randomly selected from the samples.

5. Technique

In our literature search, we found that there are recent works which implement CNN technique on Chinese character recognition. However, we didn't find examples of using CNN based on Global Page. All of them are dealt with each single Chinese Character.

In our research we decide to use two different methods and recognize 5 styles of Chinese calligraphy based on full-page document. And compare the results and see which method is more effective.

5.1. Method with Fourier transform

5.1.1 Analysis with Fourier transform

We tried to find a new approach for Chinese calligraphy recognition but based on the full-page instead of individual Chinese character. In this paper, we used the method of Fourier transform with python. In this part, we contain two parts, one is the Fourier transform of wavelet filters, another part is Euclidean distance which to distinguish five styles of Chinese character.

The method is based on the Fourier transform of wavelet filters and the method of Euclidean distance analysis applied to full-page document previously localized.

The Fourier transform is very important for image processing, mainly to decompose the image into sine and cosine components. Thus, after the image transformation, more changes and processing are done. For example, high-frequency information extraction, contour processing or object recognition technology of images. At present, our technology of applying Fourier transform is very mature, and we have derived a series of other related filters based on this technology.

Our article focuses on digital images, which belong to the Discrete Fourier Transform (DFT) category. Therefore, we mainly use DFT to be a sampled Fourier transform, in which only part of the frequency is included in the transformed information. The purpose of this is to describe the image information in the spatial domain, so that a set of arrays can be obtained. The number of frequencies corresponds exactly to the number of pixels in the spatial domain image.

Image processing is a very broad and promising direction in the field of artificial intelligence. In the process of using OpenCV, we know that the digitization of images is an important part of processing it. It is expected that Fourier transform can help you understand the coding in

images. Information. The processing of images and even video involves more or less Fourier transforms. The core idea of this transformation is that all waveforms can be superimposed by a series of simple but different frequency sinusoids.

After the Fourier transform image, the high frequency and low frequency information can be effectively distinguished, and the richer technical means can be used to extract the features and information we need. For example, the white part after the Fourier transform (that is, the low-frequency part with larger amplitude) indicates the characteristic of slow change in the image, or the characteristic that the gray level changes slowly (low frequency part). The black portion after the Fourier transform (that is, the high-frequency portion with a low amplitude) indicates the characteristic of fast change in the image, or the characteristic that the gray scale changes rapidly (high-frequency portion).

For method of the Fourier transform. In this method we only choose a few test data with good quality figures. To find out some interesting parts of that five styles of Chinese calligraphy.

The First we create some samples of different styles of Chinese Character on full-page document. Figure 15 below as five single style of Chinese calligraphy, which are Cao, Kai, Li, Xing, Zhuan. It's similar like the dataset_A but only the different is the just 20 examples of image. With those images, we could easy to find some regulars and do some basic analysis and comparisons.

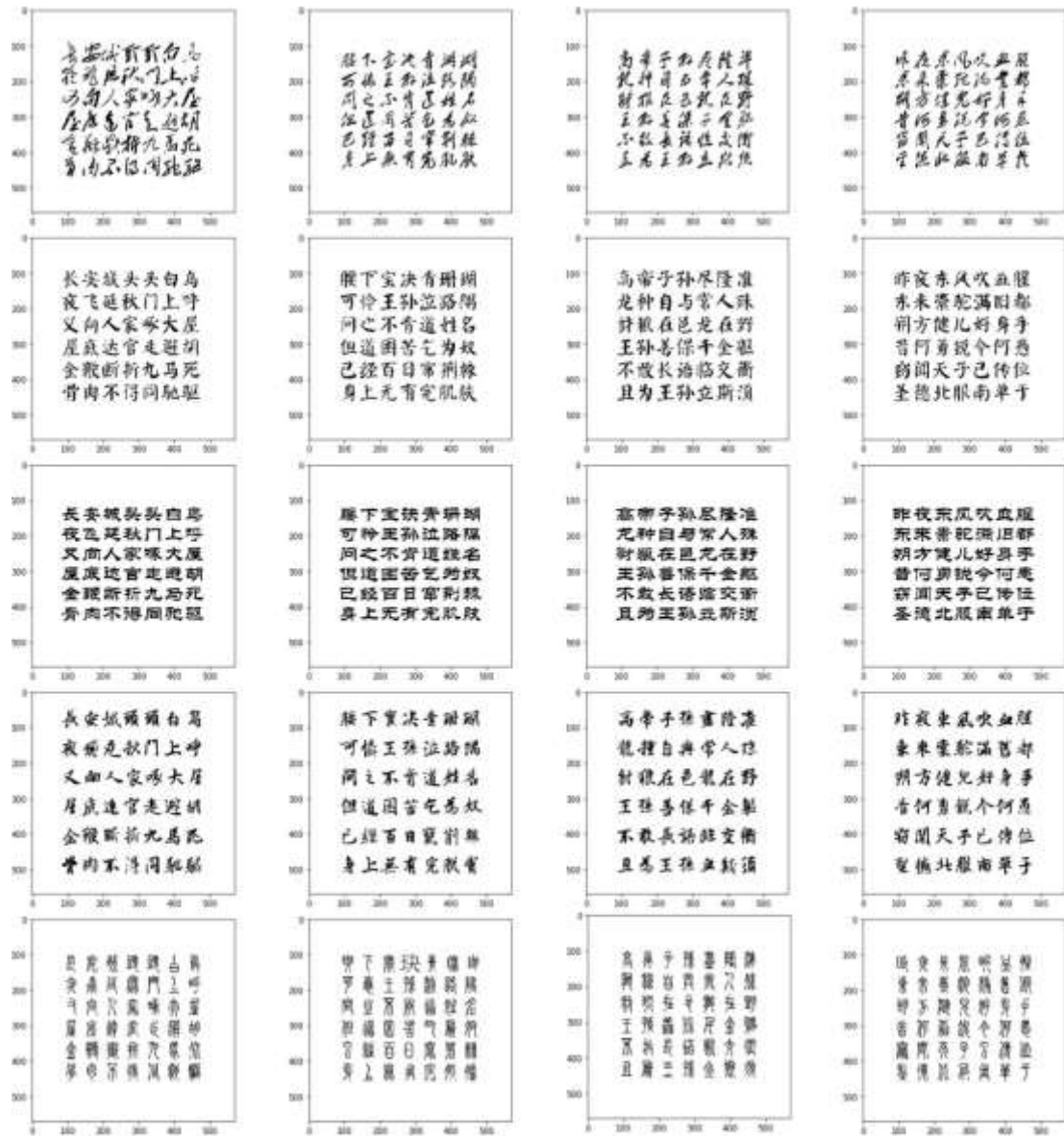


Figure 15. The Data example for FT

As shown in Figure 15, from first line to the last line, each line we create 4 photos with the same style of Chinese Character but different single words. The first line is *Cao style*. The second line is *Kai style*, the third line is *Li style*, the fourth line is *Xing style* and the last line is *Zhuan style*.

Color picture information is too large, and sometimes we only use the information in the grayscale image. In order to improve the speed of operation, we will use a grayscale image naturally, and sometimes even the grayscale image is too large. It also can use Binary images to extract some information.

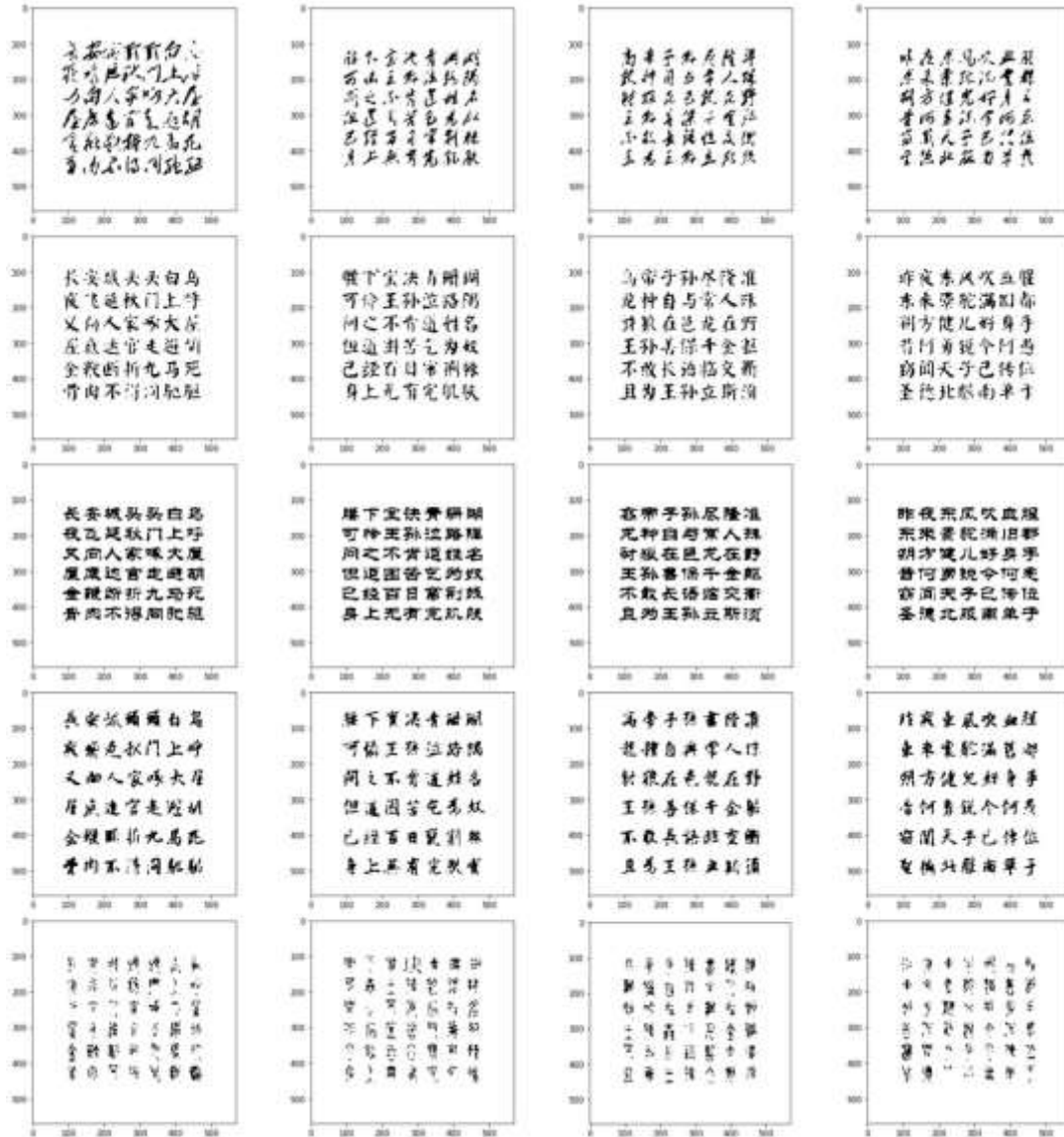


Figure 16. After convert gray and Global Thresholding ($v = 127$)

From figure 16, we can see images which are changed a lot, which are not clear as before. Which mean the computer can process the data much higher efficiency. But we also know we maybe lost some important information for one style. (i.e. Zhuan style)

5.1.2 Analysis with Euclidean distance

For a square image of size $N \times N$, the two-dimensional DFT is given by:

$$F(k, l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) e^{-i2\pi(\frac{ki}{N} + \frac{lj}{N})}$$

The $f(i, j)$ in the above formula is called an image in the spatial domain, and the basis function corresponding to each point $F(k, l)$ in the Fourier space is used as an exponential term. The formula means that the value of each point $F(k, l)$ is obtained by multiplying the spatial image by the corresponding basis function and summing the results.

The Euclidean distance is the "normal" linear distance between two points in the Euclidean space. It can be used for image matching in image processing. Through this distance, the space can be measured, so that the similarity between the pictures can be judged. We know that the dimensions of image processing are broader, and the generalized term for the Euclidean norm is the L2 norm or L2 distance.

Usually we know that in one dimension, we can simply say that the distance between any two points on the solid line is the absolute value of the numerical difference of its coordinates. If p and q are two points on a solid line, then the distance between them is:

$$\sqrt{(q - p)^2} = |q - p|$$

The two dimensions, In the Euclidean plane, if $p = (p_1, p_2)$ and $q = (q_1, q_2)$ then the distance is given by:

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}$$

Three dimensions, In the three-dimensional Euclidean space, the distance is:

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2}$$

The core of the Euclidean distance algorithm is: Let the image matrix have n elements (n pixels), and use n element values (P_1, P_2, \dots, P_n) to form the feature set of the image (all in the pixel matrix) Pixels), the feature group forms an n -dimensional space (Euclidean distance is for multidimensional space), and the feature code (each pixel) in the feature group constitutes the value of each dimension, which is P_1 (the first pixel) Corresponding to one dimension, P_2 (the second pixel) corresponds to two dimensions, \dots , P_n (the n th pixel) corresponds to n dimensions. In the n -dimensional space, the two image matrices form a point, and then the distance between the two points is calculated by the mathematical Euclidean distance formula. The smallest distance is the best matching image.

Then for an n -dimensional space, the distance is:

$$d(\mathbf{p}, \mathbf{q}) = d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}.$$

In this section, we will use OpenCV to find the Fourier Transform of images. from the document OpenCV 3.0 manual said “*Fourier Transform is used to analyze the frequency characteristics of various filters. For images, 2D Discrete Fourier Transform (DFT) is used to find the frequency domain. A fast algorithm called Fast Fourier Transform (FFT) is used for calculation of DFT*” [61]. According to the theoretical part of the manual, combined with Python's own OpenCV processing library, you can perform fast Fourier transform or more image editing operations on the image, which can be said to save a lot of operations for subsequent image processing. The effect is also quite ideal.

For In this technique we use a sinusoidal signal, $x(t) = A\sin(2\pi ft)$ to sample the image information. f is the frequency of the signal, and each picture uses its frequency domain to get the peak at f . The same frequency domain is obtained in the discrete signal replacement, the interval is $[-\pi, \pi]$ or $[0, 2\pi]$ (or $[0, N]$ for the N point DFT, and is periodically transformed). The image is a superposition of the signals sampled in both directions. The frequency at which the image can be generated is then expressed in the X and Y directions. Then, based on the transformation result of this image, the Euclidean distance is used to judge the similarity between the images. [61]

Figure 17. show the original photo after FT, the Result look like below:

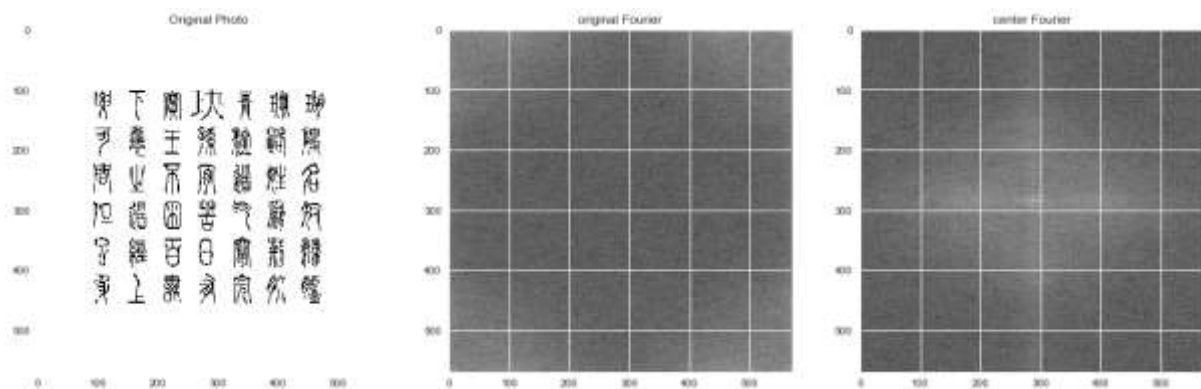


Figure 17. Photo after FT (original Fourier and Center Fourier)

After analyzing the transformed picture, we observe that in the white area at the center, this part is the area showing the low frequency after the image transformation. The high-frequency

area and the low-frequency area must be included in each picture information, this information is very good for processing pictures, and finding certain rules between each image or sample can be performed in the frequency domain. After that, some operations we could do, such as adding filters, or use high-pass filtering or some of its band filters to denoise or extract contours or reconstruct images. These can be used as image features to provide a basis for analysis for subsequent processing.

Before we start to use the Euclidean distance, we have to make them normalization. It is a widely used technique in the fields of computer vision and pattern recognition. The so-called image normalization is to convert the original image to be processed into a corresponding unique standard form through a series of transformations (the standard form image has invariant characteristics for affine transformation such as translation, rotation, scaling, etc.). In recent years, the image normalization technique based on moments has received widespread attention. The basic working principle is as follows: Firstly, the parameters of the transformation function are determined by using the moments in the image which are invariant to the affine transformation, and then determined by this parameter. The transformation function transforms the original image into a standard form of image (the image is independent of the affine transformation). In general, the moment-based image normalization process consists of four steps: coordinate centering, x-shearing normalization, scaling normalization, and rotation normalization.

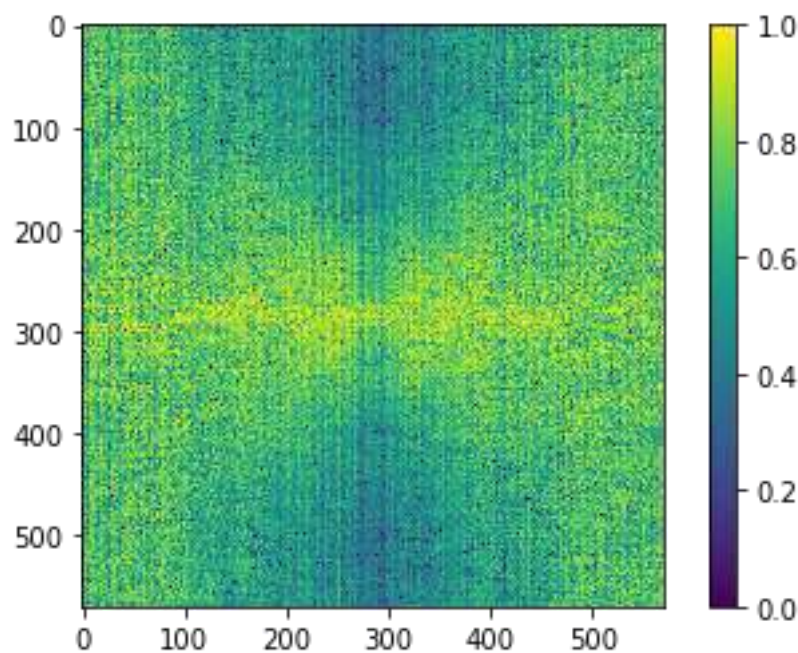


Figure 18. Normalization (image processing) after FFT

Figure 18. Normalization (image processing) in image processing. We do data normalization when seeking for relations. Normalization is a process that changes the range of pixel intensity values. The image matrix in python between 0.0 to 1.0. Based on distance comparison, the best result is to normalize the compared objects to a fixed range, preventing the data from being too large or too small. The results of such comparisons are relatively stable and easier to understand and find the rules.

After normalization, we can do FFT in order to find more similar parts in those five kinds of Chinese calligraphy.

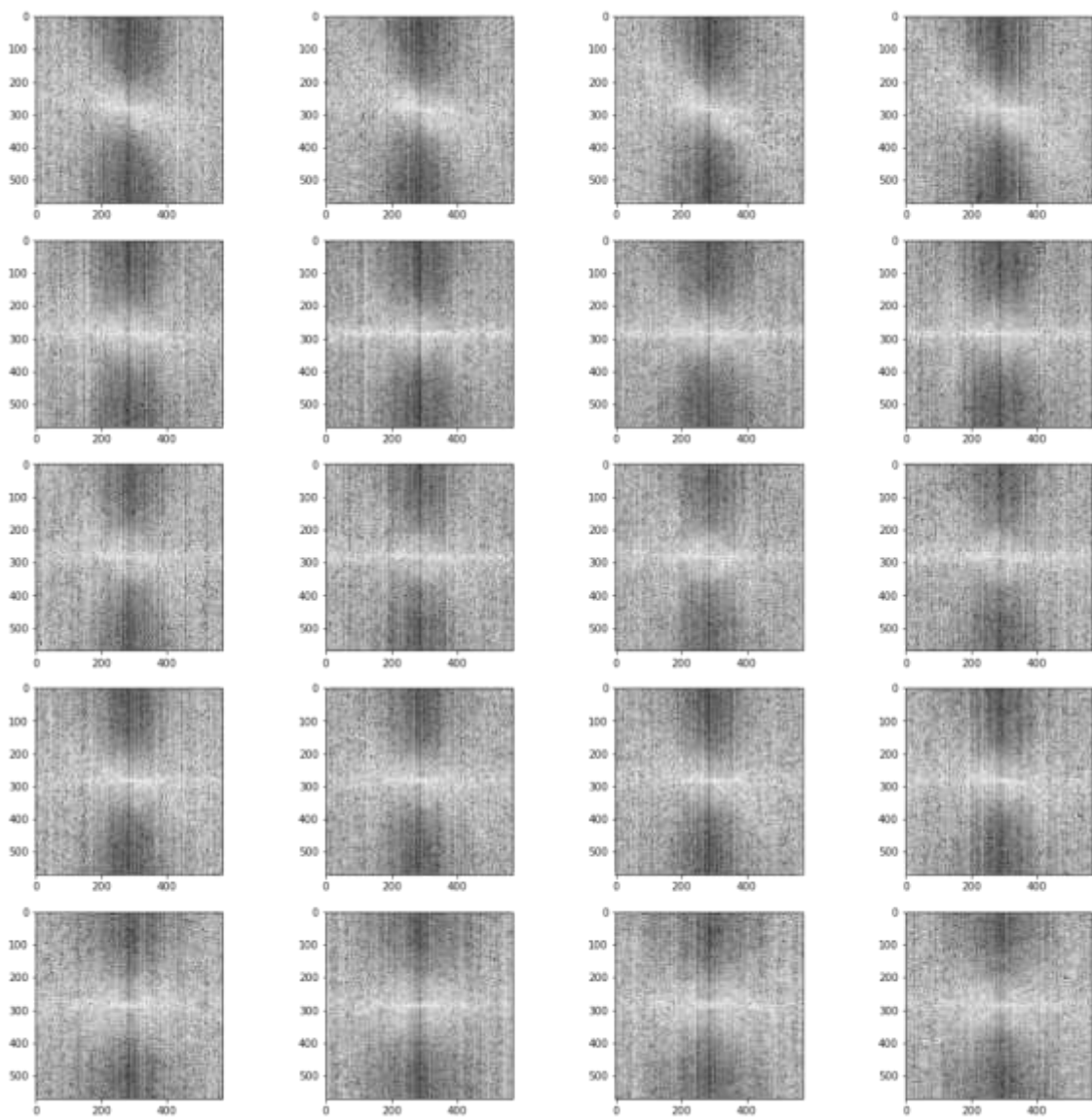


Figure 19. 5 sets of Data after FFT (Center Fourier black and white)

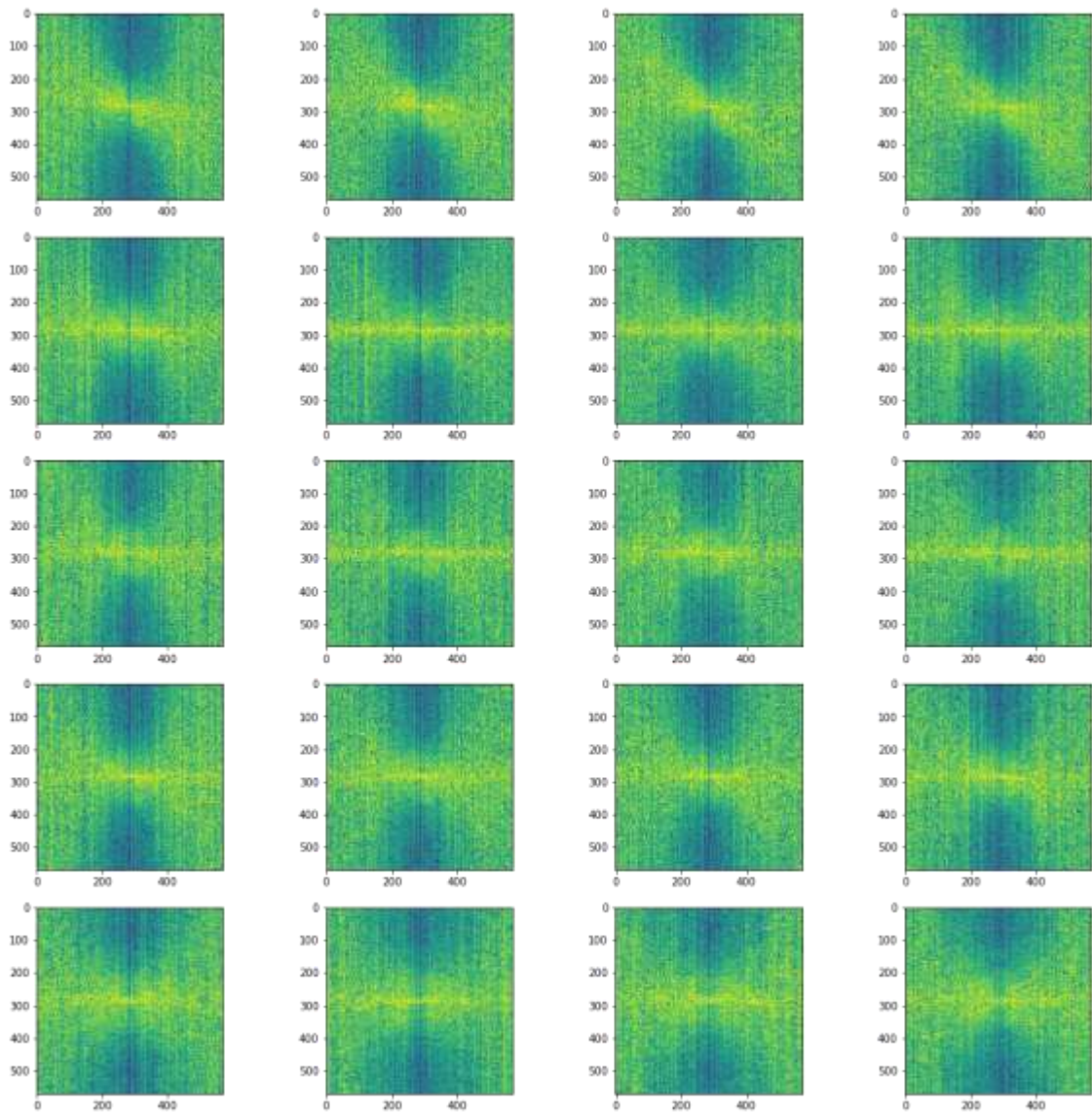


Figure 20. 5 sets of Data after FFT (Center Fourier color)

As we can see from figure 21 and figure 22, no matter black or color images. It seems that have some regulars for each sets of data. In distance-based classification, according to the problem of distance classification, the purpose of standardizing each feature vector is to change it. When calculating the distance, the range of each feature value is not too broad. As we mentioned before, using the Euclidean distance we contained two situations:

1. The Euclidean distance between the same style of Chinese calligraphy
2. The Euclidean distances between each sets style of Chinese calligraphy

In this article, we extract one of example and the result, to compare with others images. We find the results as below, (i.e. using image 0(**Cao style**) and compared with other data)

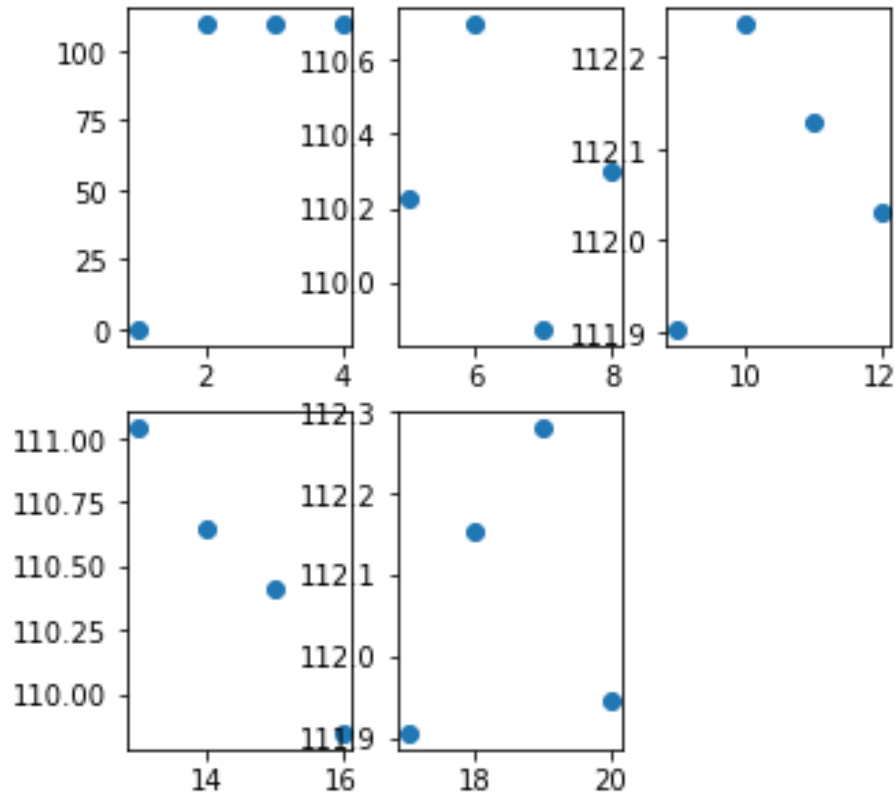


Figure 21. using image 0(Cao style) and compared with other data

0	1	2	3	4	5	6	7	8	9
0	109.6058	110.1992	109.8859	110.2243	110.7002	109.8728	110.2975	111.9032	112.2377
10	11	12	13	14	15	16	17	18	19
112.1281	112.0298	111.0424	110.6462	110.4108	109.8462	111.9051	112.1541	112.28	111.9462

Table 7. (0-3) showed result of the Euclidean distance between the same style of Chinese calligraphy

(4-19) showed result of the Euclidean distance between each sets style of Chinese calligraphy

In order to compare distribution of all data, we plot in the same Figure 22. Now we can see it hard to find some rules or the regulars which can perfect to separate them. Further, see the result and analysis, the X-axis represents 0 - 20 all data sequences, and the Y-axis represents distance of image 0 to other images. We couldn't use Euclidean distance to completely separate even in the same sets of data. For example, we can't say when Euclidean distance between 109.5 to 110.5 is class 1 (Cao style). Because of image 6 and 7 or image 15, it also between 109.5 to 110.5, they are totally belonged to other styles of Chinese Calligraphy.

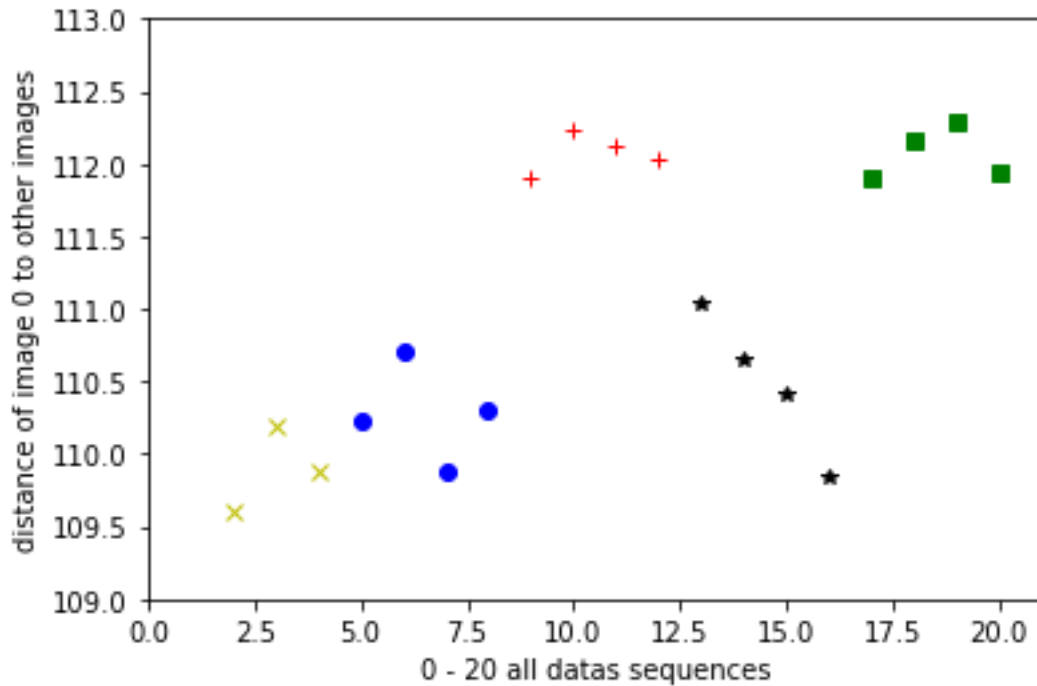


Figure 22. Euclidean distance distribution of image 0(Cao style) to others

And for another thing, if we only consider two classifications for this result. we can do some analysis with two classes (Cao style and Li style). As the result, we know the points from 0 to 3 are belonged to Cao style, from 8 to 11 are belonged to Li style. In this plot, they are totally separated by the Euclidean distance. Because of the points from 8 to 11 which distance is more than 111.5.

For that reason, we try other method to do the classification. But in the future work, we will continue to use other classification combination with Fourier analysis. If it's possible to find other dismission to use SVM to do the classification.

5.2. Method with Convolutional Neural Network

In this article, we use Convolutional Neural Network to recognize different traditional Chinese calligraphy styles.

We know that the structure of a neural network is like this:

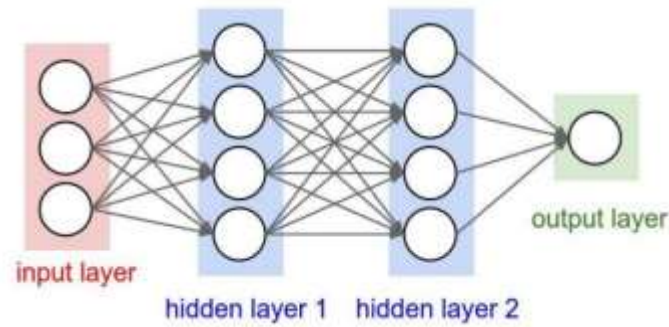


Figure 23. the structure of a neural network

In the field of image recognition technology, convolutional neural networks (CNN) have become a mainstream technology, and a wide range of applications in object detection and facial recognition, more things to be developed. [62]

In image processing, pictures are stored as a matrix and in a computer, CNN in deep learning can process and classify pictures (eg, Xing, Cao, Li, Zhuan, Kai). Stored in a computer we can consider a matrix with these dimensions (h = height, w = width, d = size) depending on the resolution, the value of the image will vary. In general, a matrix of RGB, such as the example of the following figure, is an image of a $6 \times 6 \times 3$ array (3 fingers of RGB values), and if the image processed by grayscale is a $6 \times 6 \times 1$ array.

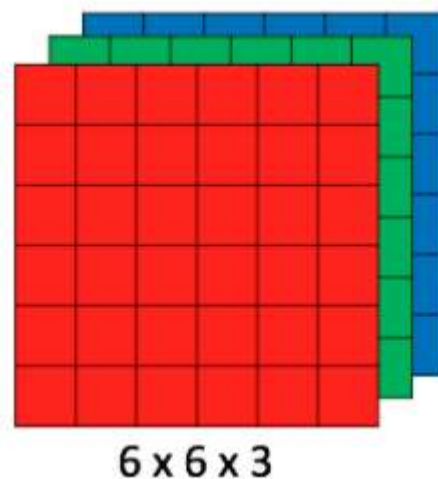


Figure 24. Array of RGB Matrix

Technically, deep learning CNN models to train and test, normally each input image will pass it through a series of convolution layers with filters (Kernels), Pooling, fully connected layers (FC) and apply Softmax function to classify an object. The below figure is a complete flow of CNN to process an input image and classifies the objects based on values.

The basic structure of the CNN model is not complicated. Generally, each input image is first subjected to basic image processing by convolution layers Kernels, followed by the necessary Pooling layer, and the final processed image needs to pass through the convolution layer of the fully connected layer (FC). The function applied to the Softmax layer is the key to classifying images.

According to this structure, we have drawn the following basic image processing classification.

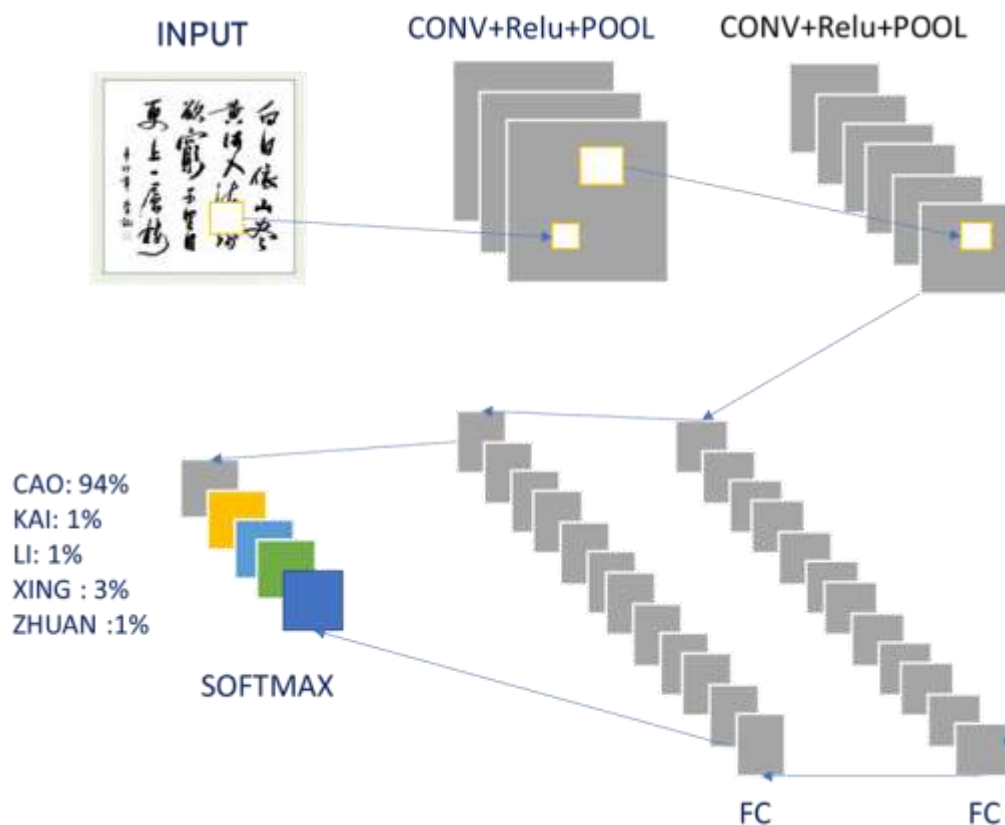


Figure 25. Neural network with many convolutional layers

As is mentioned in the previous section, in the beginning we trained our models on 320 samples and used 80 samples for validation. We used a batch size of 128. The configuration of this model which are contained two layer of convolutional layer, one max pooling and fully connected layer. As we see, it's quite simple model for CNN.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 48, 48, 32)	320
conv2d_2 (Conv2D)	(None, 46, 46, 64)	18496
max_pooling2d_1	MaxPooling2 (None, 23, 23, 64)	0
flatten_1 (Flatten)	(None, 33856)	0
dense_1 (Dense)	(None, 128)	4333696
dense_2 (Dense)	(None, 5)	645
Total params: 4,353,157		
Trainable params: 4,353,157		
Non-trainable params: 0		

Table 8. CNN configuration of a simple model (Train on 320 samples, validate on 80 samples)

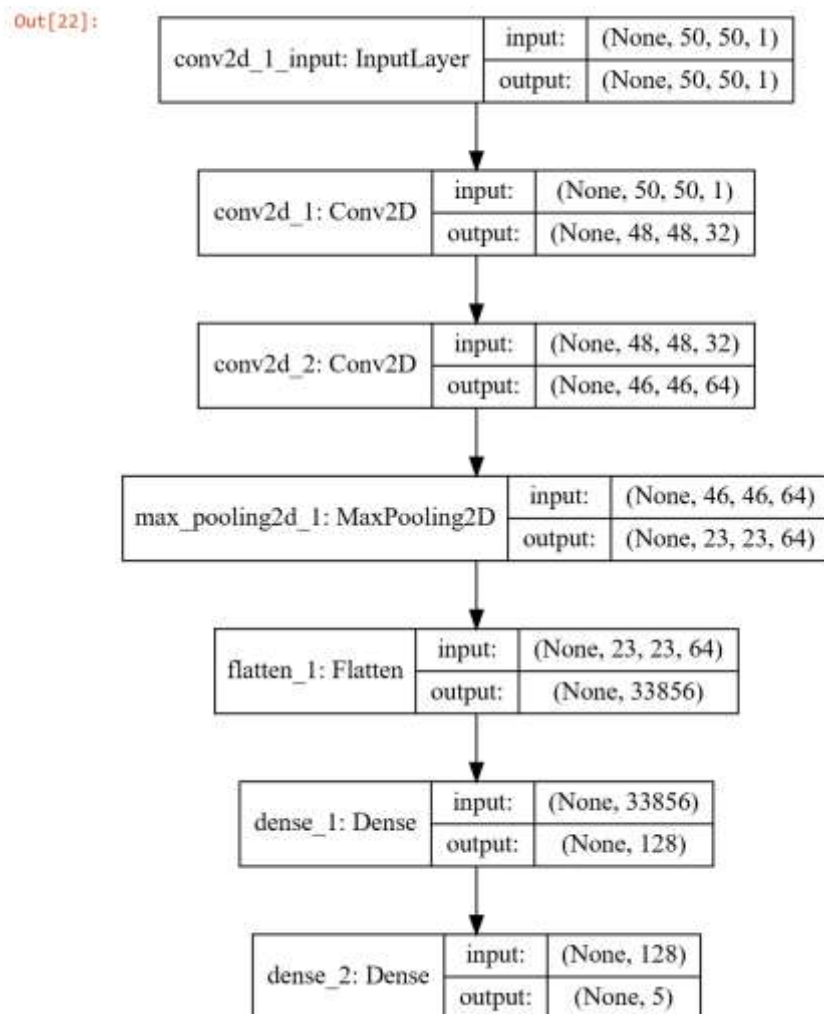


Figure 26. diagram of CNN

5.2.1 Feature Learning, Layers, and Classification

Convolution Layer

The convolutional layer is a layer that processes an image as a primary layer, and calculates the upper layer by convolution and window-by-window sliding. There is a convolution kernel involved, and the convolution operation is an inner product operation of a matrix of two convolution kernel sizes. The size and number of convolution kernels of each layer can be defined by themselves, but in general, according to the experimental experience, a small number of convolution kernels will be set in the convolution layer closer to the input layer. The convolution layer sets the number of convolution kernels. In this article, we have 2 convolutional layers. The Convolution Kernel in CNN is no different from the traditional Convolution Kernel. Still taking an image as an example, the Convolution Kernel is convolved in turn with image blocks at different locations of Input.

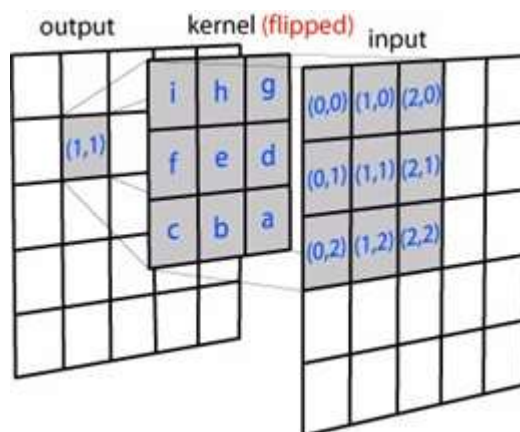


Figure 27. Convolution and kernel

Strides

Stride is the number of pixels shifts over the input matrix. Decide how many steps you can slide to the edge. For example, the stride is 1 it means we move the filters to 1 pixel at a time. And if the stride is 2, it means move the filters to 2 pixels at a time and so on.

Rectified linear unit (ReLU)

The activation function is one of the key elements of a neural network. Without them, our neural network can only be a combination of linear functions, and it can only have limited scalability, and it does not have the power to go beyond logical regression. Introducing nonlinear elements will give you more flexibility in the learning process and the ability to create complex representations. The activation function also has a major impact on learning speed, which is one of the main selection criteria. In Figure 28 below, some common activation

functions are shown. Currently, the most popular hidden layer may be ReLU. But if we want to deal with the two-class problem, especially if we want the value returned from the model to be in the range of 0 to 1, we sometimes still use the sigmoid activation function.

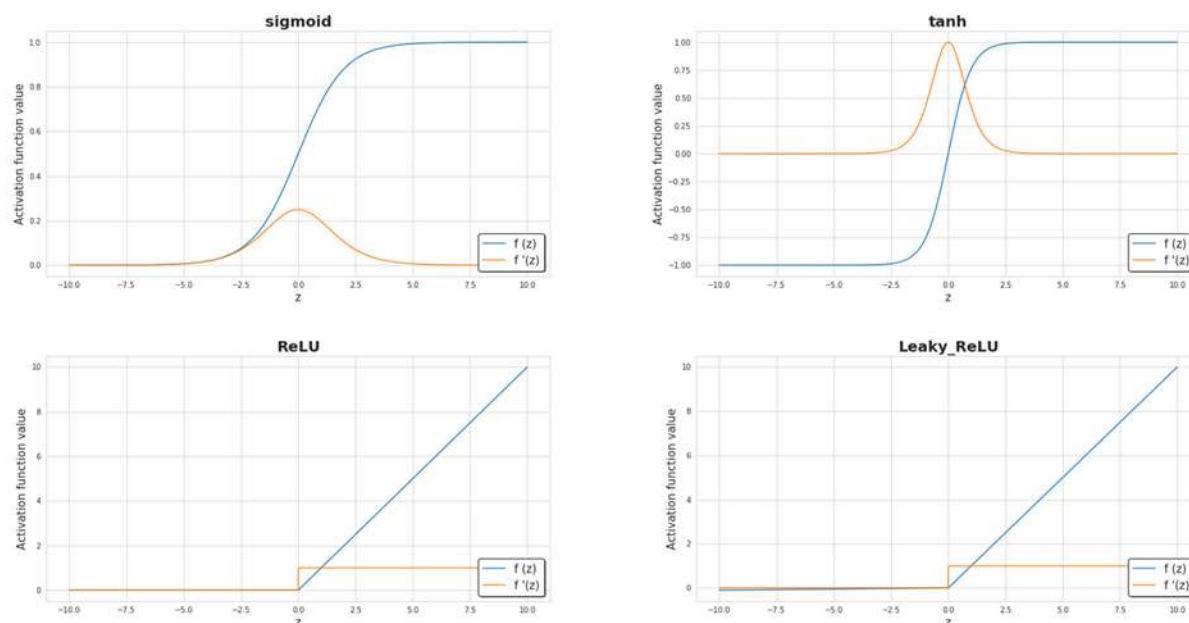


Figure 28. The most popular activation function and its derived diagrams

Pooling

The essence of Pooling is actually sampling. Pooling For the feature map you enter, choose some way to compress it. As shown in the figure below, it means that the value in the neighborhood of the Feature Map is selected and the maximum value is output to the next layer. This is called Max Pooling. So a Feature Map is compressed. Generally speaking, there are two kinds of pooling methods: Max Pooling: take the largest value in the sliding window, Average Pooling: take the average value of all the values in the sliding window. Pooling layers section would reduce the number of parameters when the images are too large. [62]

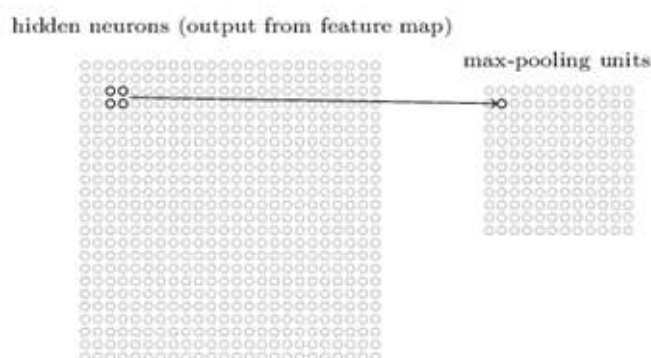


Figure 29. The most popular activation function and its derived diagrams

However, not all cases of Max Pooling have a good effect. Sometimes some surrounding information will have a certain effect on the recognition of a particular feature. It is not worthwhile to abandon this part of the "not important" information at this time. Therefore, the specific situation has to be specifically analyzed.

5.2.2 Classification Layers

After learning features in many layers, the architecture of a CNN shifts to classification.

Fully Connected Layer

The layer we call as FC layer, there are three parameters of full connection layer affects the model. First is total number of layers (length) of the fully connected layer, the second is number of neurons in a single fully connected layer (width), and the last one is activation function. Like the matrix flattened into vector and feed it into a fully connected layer.

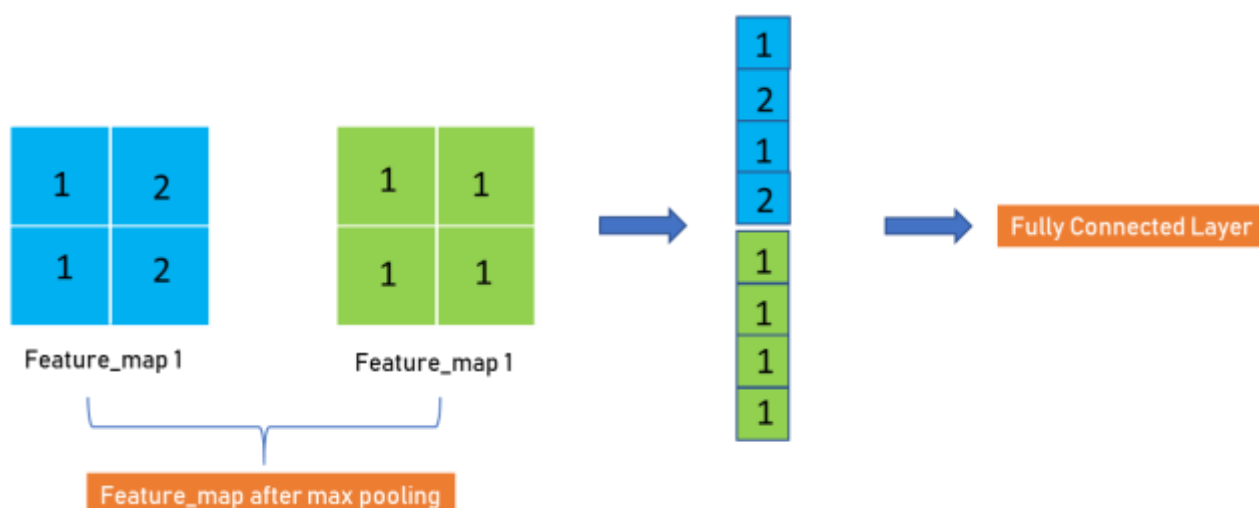


Figure 30. The process of flatten

In simply said, we can think of a fully connected layer as a "classifier" and in the entire convolutional neural network. This explains the role of each level. In general, the operation of the convolutional layer, the merged layer, and the active functional layer that mapping the raw data to the hidden layer feature space, so the fully connected layer is functional that mapping the "distribution feature representation " to the sample tag space.

6. Results

As is mentioned in the previous section, we trained our models on different samples and used different samples of validation. We had created 3 models for the next experiments.

In those experiments we will test our 3 models with different dataset. We make 3 experiments to compare those results and to see which one is the best. For those 3 models, 20 epochs are enough to observe the validation accuracy stabilize during the last 1-2 epochs.

6.1. The first experimental protocol

The model we used is Model_1 with dataset_B (500 examples). The examples are following as Figure 30.



Figure 31. Examples of data (dataset_B)

From this figure 31. We can see those data are all kinds different types of pictures, black and white, large fonts, a whole page, scanned and even with angled. All of them are download from the internet.

6.1.1 Experiment and results

Accuracy – From theory describe “For a given test data set, the ratio of the number of samples correctly classified by the classifier to the total number of samples.” For our model, we have got 0.29 which means our model 1 is not good for deal with those data. The point is that Train accuracy is low (47.5%)

Accuracy = $\frac{TP+TN}{TP+FP+FN+TN}$

Result of Model 1 (Train on 320 samples, validate on 80 samples, test on 100 samples)					
Train loss	Train accuracy	Val loss	Val accuracy	Test loss	Test accuracy
1.29236813784	0.475	1.69200890064	0.325	1.68353269577	0.29

Table 9. CNN configuration of the basic model (Train on 320 samples, validate on 80 samples)

The confusion matrix is another method to judge the degree of classification in addition to the ROC curve and AUC. In machine learning, especially in statistical classification, the confusion matrix is also called the error matrix. Each list of the matrix reaches the classifier's class prediction for the sample. Each row of the two matrix expresses the real category to which the version belongs. It is called the confusion matrix because it is easy to see whether the machine learning will sample. The category is confused.

Here are a few concepts that need to be explained first:

TP (True Positive): true is 0, prediction is also 0

FN (False Negative): true 0, predicted as 1

FP (False Positive): true is 1, prediction is 0

TN (True Negative): true is 1, prediction is also 1

Every model we had do the test, the evaluation with confusion matrix can show the predicted with each Chinese calligraphic style. The table 9 show us the result of model 1 with 100 examples test.

		Predicted				
		cao	kai	li	xing	zhuan
Actual	cao	12	0	4	0	0
	kai	8	3	8	0	0
	li	5	0	11	0	0
	xing	12	1	13	0	0
	zhuan	9	1	10	0	3

Table 10. confusion matrix with model 1(100 examples with all different pictures)

According to the table 10. We observe that it contains 100 examples. Also, we get the result of Cao style, which we got 12 images are predicted correct of 16 totals. Other 4 images are

predicted incorrect. CNN thinks they are more similar with Li style. The same way we can observe others styles of Chinese calligraphy. In this case, Xing and Zhuan style are worse performance than others.

6.1.2 Evaluation

In deep learning, there are several indicators used to evaluate his quality against the model. The data evaluated by this model can explain whether the prediction of the data is accurate. The established model must pass the following evaluation test.

The following Table 11. shows the results of the model 1

	precision	recall	f1-score	support
0	0.26	0.75	0.39	16
1	0.60	0.16	0.25	19
2	0.24	0.69	0.35	16
3	0.00	0.00	0.00	26
4	1.00	0.13	0.23	23
avg / total	0.42	0.29	0.22	100

Table 11. Precision/ Recall/ Fi-score /support (Train on 320 samples, validate on 80 samples)

Precision - In fact, it's very simple. The accuracy rate is for our predictions. From theory describe “*It shows how many of the samples with positive predictions are true positive samples.*” [62]. There are two possibilities for the prediction to be positive. One is to predict the positive class as a positive class (TP), and the other is to predict the negative class as a positive class (FP). We have got 1.00 precision which is pretty good for the style 4 (Zhuan) but it doesn't mean this model 1 is fit to the Zhuan style because from the formula we will understand why we got 1.00. also, we need to see the Recall.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

Recall (Sensitivity) - The recall rate is for our original sample, From theory describe “*it indicates how many positive examples in the sample are predicted correctly.*” [62]. There are also two possibilities. One is to predict the original positive class as a positive class (TP), and the other is to predict the original positive class as a negative class (FN). We have got recall of 0.75 for style 0 (Cao) which is good for this model as it's above 0.5. but the average recall just is 0.29.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

F1 score – From theory describe “*F1 Score is the weighted average of Precision and Recall.*” [62]. When we evaluate, if we want to be idealized, we want the higher the precision and the higher the recall, but this will not happen because the two indicators are contradictory in some cases. In our case, F1 average score is 0.22.

$$\text{F1 Score} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision})$$

	actual positive	actual negative
predicted positive	TP	FP
predicted negative	FN	TN

(a) Confusion Matrix

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{True Positive Rate} = \frac{TP}{TP+FN}$$

$$\text{False Positive Rate} = \frac{FP}{FP+TN}$$

(b) Definitions of metrics

Figure 32. (a) Confusion Matrix (b) Definitions of metrics

AUC is a model evaluation index and we know it can only be used for evaluation of two-category models. However, we could treat as each style can be two-category models. For example, for the Cao style we could say if it is Cao style otherwise no. in this case that we could draw the AUC.

AUC (Area Under Curve) is the area under the ROC curve. If the model is accurate as described above, the AUC is 1. However, in real life, especially in the industrial world, there will not be such a perfect model. Generally, the AUC is between 0.5 and 1. The higher the AUC, the better the discrimination of the model. The AUC of Cao _CNN in the figure 33. above is 0.70.

In the real world, since the data is one by one, the threshold is discretized, and the curve presented is jagged. Of course, the more data, the finer the threshold also the “curve” much smoother the “curve”.

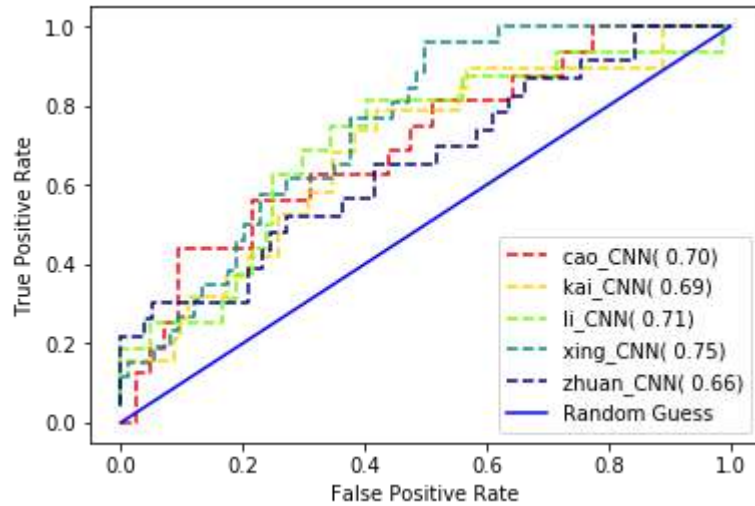


Figure 33. ROC (receiver operating characteristic curve) AUC (Area Under Curve)

The relativeness in the definition suggests that Zhuan samples are more difficult to be recognized in this model. But due to results, we will check some examples of the mislabeled samples are shown in Figure 34.

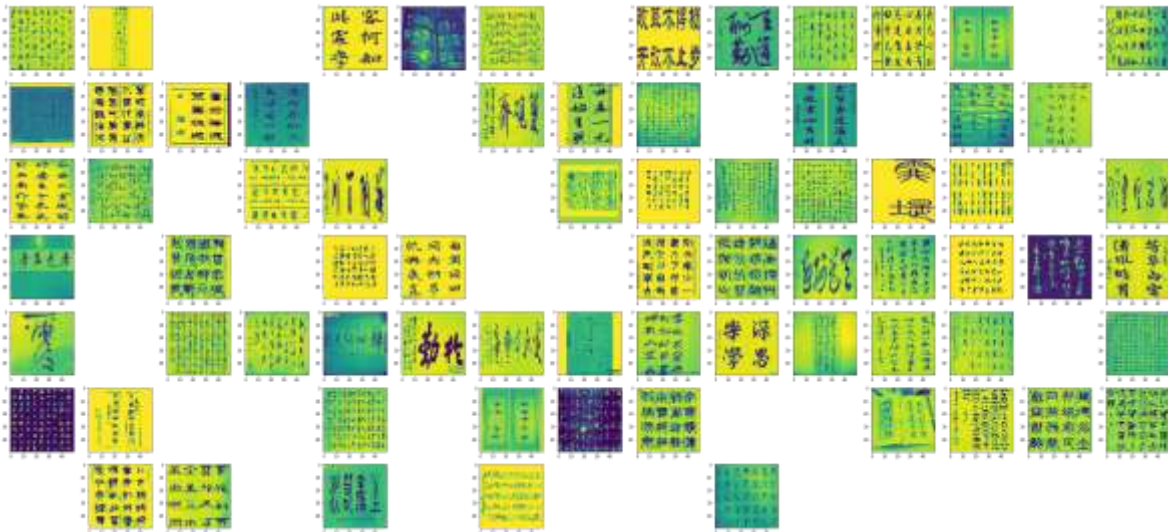


Figure 34. Predict incorrect of the images

To analysis the result of Figure 34. We could observe that there are no common features. But the most important thing is all of them are very difficult to recognize even by human also. With this samples make us the new idea to improve the data examples first and do more experiment. For checking if the result will improve more or no.

6.2. The second experimental protocol

6.2.1 Experiment and results

The model we used is Model_1 with same pictures but different styles (total in 500 examples).

In this experiment we just want to explore if Model 1 would be overfitting in this situation.

The examples are following as Figure 35.

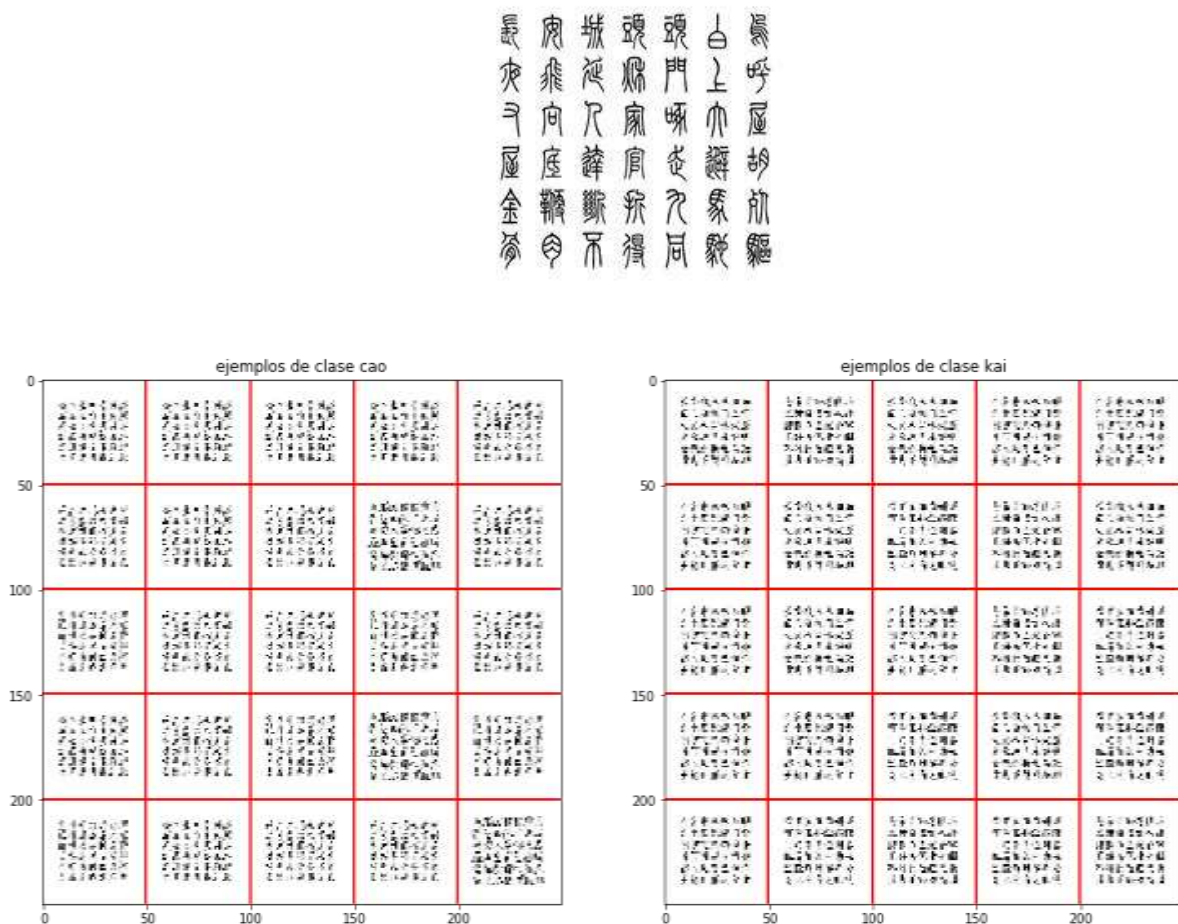


Figure 35. Examples of Chinese calligraphy styles

Add more data and modify the quality of pictures which are much clear. As we see from the Figure 35. The examples of images are Cao and Kai style.

The result as we expected, we use the train on 320 samples, validate on 80 samples, test on 100 samples, the result as Table 12. show the result is the overfitting, nothing is predicted wrong.

Result of Model 1(Train on 320 samples, validate on 80 samples, test on 100 samples)					
Train loss	Train accuracy	Val loss	Val accuracy	Test loss	Test accuracy
0.028243242	1	0.03289739	1	0.0270118	1

Table 12. Result of Model 1

Also, we could know the evaluation with confusion matrix, it can show the predicted with each Chinese calligraphic style. The table 13 show us the result of model 1 with 100 examples test. Obviously, we know it is overfitting.

		Predicted				
		cao	kai	li	xing	zhuan
Actual	cao	17	0	0	0	0
	kai	0	19	0	0	0
	li	0	0	21	0	0
	xing	0	0	0	25	0
	zhuan	0	0	0	0	18

Table 13. confusion matrix with model 1(100 examples with same pictures)

6.2.2 Evaluation

	precision	recall	f1-score	support
0	1.00	1.00	1.00	17
1	1.00	1.00	1.00	19
2	1.00	1.00	1.00	21
3	1.00	1.00	1.00	25
4	1.00	1.00	1.00	18
avg / total	1.00	1.00	1.00	100

Table 14. Recognition performance Precision/ Recall/ Fi-score /support

(Train on 320 samples, validate on 80 samples)

The definition of overfitting that is comparing the performance on training (e.g., accuracy) vs. the performance on testing or validation and this is the only way.

Evaluation with ROC Figure 36. We can't say it's overfitting because this only evaluation for the model. But for the result of table 14. The ROC will think this is the good model for recognize those Chinese calligraphy styles.

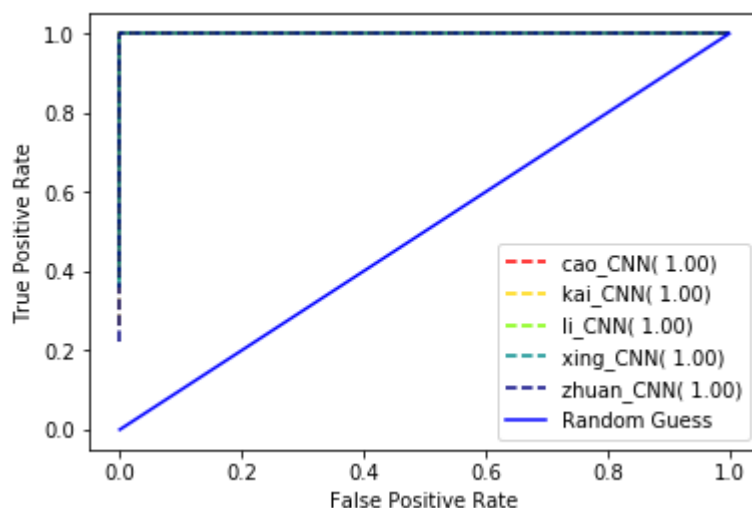


Figure 36. ROC (receiver operating characteristic curve) , AUC (Area Under Curve)

6.3. The third experimental protocol

After the second experiment, as we know, there are several steps for reducing overfitting:

1. Add more data
2. Use data augmentation
3. Use architectures that generalize well
4. Add regularization (mostly dropout, L1/L2 regularization are also possible)
5. Reduce architecture complexity.

The first step is of course to collect more data. However, in most cases we will not be able to. Then the next step is data augmentation: something that is always recommended to use.

Data augmentation includes things like randomly rotating the image, zooming in, adding a color filter etc. Data augmentation only happens to the training set and not on the validation or test set. It can be useful to check if you are using too much data augmentation.

6.3.1 Experiment and results

In this experiment, we add regularization of CNN in order to make sure the accuracy will be increase.

According to other experiments before, we use dataset_B first and try to improve the CNN network performance and reducing overfitting.

The step 1, using model 1 without dropout to observe the result and performance. That experiment we already did it in the first one. The table show the basic model 1 that we used before.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 48, 48, 32)	320
conv2d_2 (Conv2D)	(None, 46, 46, 64)	18496
max_pooling2d_1	MaxPooling2 (None, 23, 23, 64)	0
flatten_1 (Flatten)	(None, 33856)	0
dense_1 (Dense)	(None, 128)	4333696
dense_2 (Dense)	(None, 5)	645
Total params: 4,353,157		
Trainable params: 4,353,157		
Non-trainable params: 0		

Table 15. CNN Model 1

With the dataset_B we could observe that the train accuracy is just only 58.3%. So, we know that the low train accuracy can impact the validation and test accuracy.

Result of Model 1(Train on 96 samples, validate on 24 samples, test on 30 samples)					
Train loss	Train accuracy	Val loss	Val accuracy	Test loss	Test accuracy
1.298266331	0.58333333	1.532536626	0.29166657	1.5614792	0.30000001

Table 16. The result of train accuracy 58%

The step 2, add dropout layer and create new model CNN, in this time we named it Model 2.

Layer (type)	Output Shape	Param #
conv2d_9 (Conv2D)	(None, 98, 98, 32)	320
conv2d_10 (Conv2D)	(None, 96, 96, 64)	18496
max_pooling2d_5	MaxPooling2 (None, 48, 48, 64)	0
dropout_6 (Dropout)	(None, 48, 48, 64)	0
flatten_5 (Flatten)	(None, 147456)	0
dense_9 (Dense)	(None, 128)	18874496
dense_10 (Dense)	(None, 5)	645
Total params: 18,893,957		
Trainable params: 18,893,957		
Non-trainable params: 0		

Table 17. CNN Model 2

And the train accuracy is changed to 55%

The step 3, I added another dropout after (Dense) dense_11, which is named Model 3.

Layer (type)	Output Shape	Param #
conv2d_11 (Conv2D)	(None, 98, 98, 32)	320
conv2d_12 (Conv2D)	(None, 96, 96, 64)	18496
max_pooling2d_6	MaxPooling2 (None, 48, 48, 64)	0
dropout_7 (Dropout)	(None, 48, 48, 64)	0
flatten_6 (Flatten)	(None, 147456)	0
dense_11 (Dense)	(None, 128)	18874496
dropout_8 (Dropout)	(None, 128)	0
dense_12 (Dense)	(None, 5)	645
Total params: 18,893,957		
Trainable params: 18,893,957		
Non-trainable params: 0		

Table 18. CNN Model 3

This time we get the good result, so it proved that dropout can make the accuracy is up. The train accuracy up to 73%. The only thing is the bad example also can impact the train accuracy. In this case we need to consider another dataset.

Result of Model 3(Train on 96 samples, validate on 24 samples, test on 30 samples)					
Train loss	Train accuracy	Val loss	Val accuracy	Test loss	Test accuracy
1.170002381	0.73958333	1.383834958	0.41666666	1.4763973	0.40000001

Table 19. CNN Model 3

The results in Table 19. reveal that the recognition accuracy heavily depends on the similarity between training data and testing data. So, in order to development the model 3, we use the dataset_A and train the model 3 again.

Finally, the last step we use model 3 (two dropout layers) and using data with 5000 clear pictures and different Characters in full-page documents.

Layer (type)	Output Shape	Param #
conv2d_11 (Conv2D)	(None, 98, 98, 32)	320
conv2d_12(Conv2D)	(None, 96, 96, 64)	18496
max_pooling2d_6	MaxPooling2 (None, 48, 48, 64)	0
dropout_7 (Dropout)	(None, 48, 48, 64)	0
flatten_6 (Flatten)	(None, 147456)	0
dense_11 (Dense)	(None, 128)	18874496
dropout_8 (Dropout)	(None, 128)	0
dense_12 (Dense)	(None, 5)	645
Total params: 18,893,957		
Trainable params: 18,893,957		
Non-trainable params: 0		

Table 20. CNN Model 3 with dataset_A

And the train accuracy is changed to 99.8%, the test accuracy is 96.7%

Result of Model 3(Train on 3200 samples, validate on 800 samples, test on 1000 samples)					
Train loss	Train accuracy	Val loss	Val accuracy	Test loss	Test accuracy
0.099894739	0.9984375	0.175736851	0.96625	0.167294774	0.967

Table 21. The result of CNN Model 3 with dataset_A

The test results of confusion matrix:

		Predicted				
		cao	kai	li	xing	zhuan
Actual	cao	197	0	0	3	0
	kai	0	193	2	5	0
	li	0	1	199	0	0
	xing	15	5	2	178	0
	zhuan	0	0	0	0	200

Table 22. The test results of CNN Model 3

The best test result is Zhuan style, also the worst test result is Xing style which some of them are predicted as Cao style. Actually, even by the human it's hard to recognize those two styles, because they are too similar. Other reason it's very complicated to write Xing style and Cao style. So that we can understand why we get this result.

6.3.2 Evaluation

From the theory “*Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.*” [63]. We have got 1.00 precision which is pretty good for the style 3(Xing) and 4 (Zhuan) but the bad for the style 1(Kai).

	precision	recall	f1-score	support
0	0.91	0.98	0.94	200
1	0.90	0.98	0.94	200
2	0.96	1.00	0.98	200
3	1.00	0.78	0.87	200
4	1.00	1.00	1.00	200
avg / total	0.95	0.95	0.95	1000

Table 23. CNN Model 3 with 5000 examples

Defined as the area under the ROC curve, it is obvious that the value of this area will not be greater than 1. Since the ROC curve is generally above the line $y=x$, the value of AUC is between 0.5 and 1. The higher the AUC, the better the discrimination of the model. The AUC of all style in the Figure 37. Almost near 1.00.

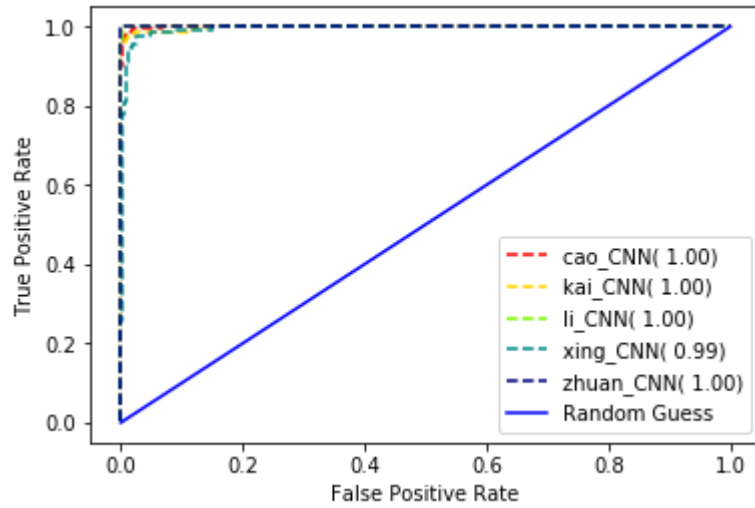


Figure 37. ROC (receiver operating characteristic curve) , AUC (Area Under Curve)

The table 24. Show the best test result is Zhuan style, and the worst test result is Xing style. Others are almost the same performance.

Calligraphy Style	accuracy
Cao	0.985
Kai	0.965
Li	0.995
Xing	0.89
Zhuan	1
avg / total	0.967

Table 24. Accuracy of five Calligraphy style

7. Discussion

At the beginning of the article, we learned that the following references are about solving Chinese calligraphy characters recognition. Some of these methods are worth learning. According to the data they test and the techniques used, we can analyze them and makes a comparison with our technology.

Year	Author.[Ref]	Name	Method	Accuracy(%)
2016	Chen, Z , Yu-Sheng [73]	Machine Learning for Calligraphy Styles Recognition	Single Character	78.5
2016	Boqi Li [9]	Convolution Neural Network for Traditional Chinese	Single Character	88.6
2017	Zhang Jiulong; Guo Luming; Yang Su; Sun Xudong; Li Xiaoshan [74]	Detecting Chinese calligraphy style consistency by deep learning and one-class SVM	Single Character with Calligraphy Font	85.5
2017	Gao Pengcheng, Gu Gang, Wu Jiangqin, Wei Baogang [75]	Chinese calligraphic style representation for recognition	Single Character	94.22
2018	Our	Our	Based on full-page document	96.7

Table 25. Different methods for Chinese calligraphy recognition

Chen, Z, et al. [73] In this paper , they use a variety of classification problems based on machine learning, the identification of individual Chinese characters, including Softmax Regression, Support Vector Machine (SVM), k-Nearest Neighbors (kNN) and Random Forests. Based on these techniques to compare the advantages and disadvantages of each technology and Finally, the technology with relatively high recognition rate is summarized. The contribution of this document can refer to more identification technologies, improve their recognition technology according to their needs, and most importantly, the recognition of Chinese characters. All databases are data samples created.

Li Boqi, et al. [9] In this paper, the CNN technology and convolutional neural network in the mainstream image recognition field are adopted. Also chose the mainstream Chinese traditional calligraphy style. They will train convolutional neural networks on five types of data sets. The database used contains 15,000 instances. And built 8 different models for testing. The recognition rate of each font analyzed by evaluating the performance of different models.

Zhang junlong, et al. [74] In this paper, when practicing Chinese calligraphy for beginners, some of the characters produced are incorrectly styled, and the detected samples are distinguished and some abnormal values are selected. The technique used is the current popular neural network model, which uses the model's stacked auto-encoder extraction and finally trains a class of support vector machines to distinguish between correct styles and foreign style characters. The resulting algorithm is good, and this method has a certain inspiration for recognizing the style of Chinese calligraphy characters.

Gao pengcheng, et al. [75] The research has an innovative approach to automatic character style representation, which can be considered as the best example of the current Chinese calligraphy single font recognition. They use three features to represent calligraphic characters, namely global features, GIST and local features, and scale-invariant feature transformations. The data set used is the unconstrained Real World Calligraphy Character Dataset (CCD) and SCL (Standard Calligraphy Character Library). The two results of support vector machine and neural network are compared for identification. In our experiments, all three features were evaluated using all three classifiers, and the depth feature was found to be the best feature of calligraphy style recognition.

In the process of image feature information extraction, CNN has its own way and method. Each convolution layer contains a part of the image information after processing the image. In the Chinese calligraphy structure, there are actually many features that can be used and distinguished, but for the CNN structure of the black box, we can ignore this part. This is also why CNN has an advantage in processing image recognition for loads.

However, in this article we still try to extract some of the examples of functionality through Figure 38. It showed one example of features extracted by CNN (32 kernels). Although we can't know why CNN extracted these features.

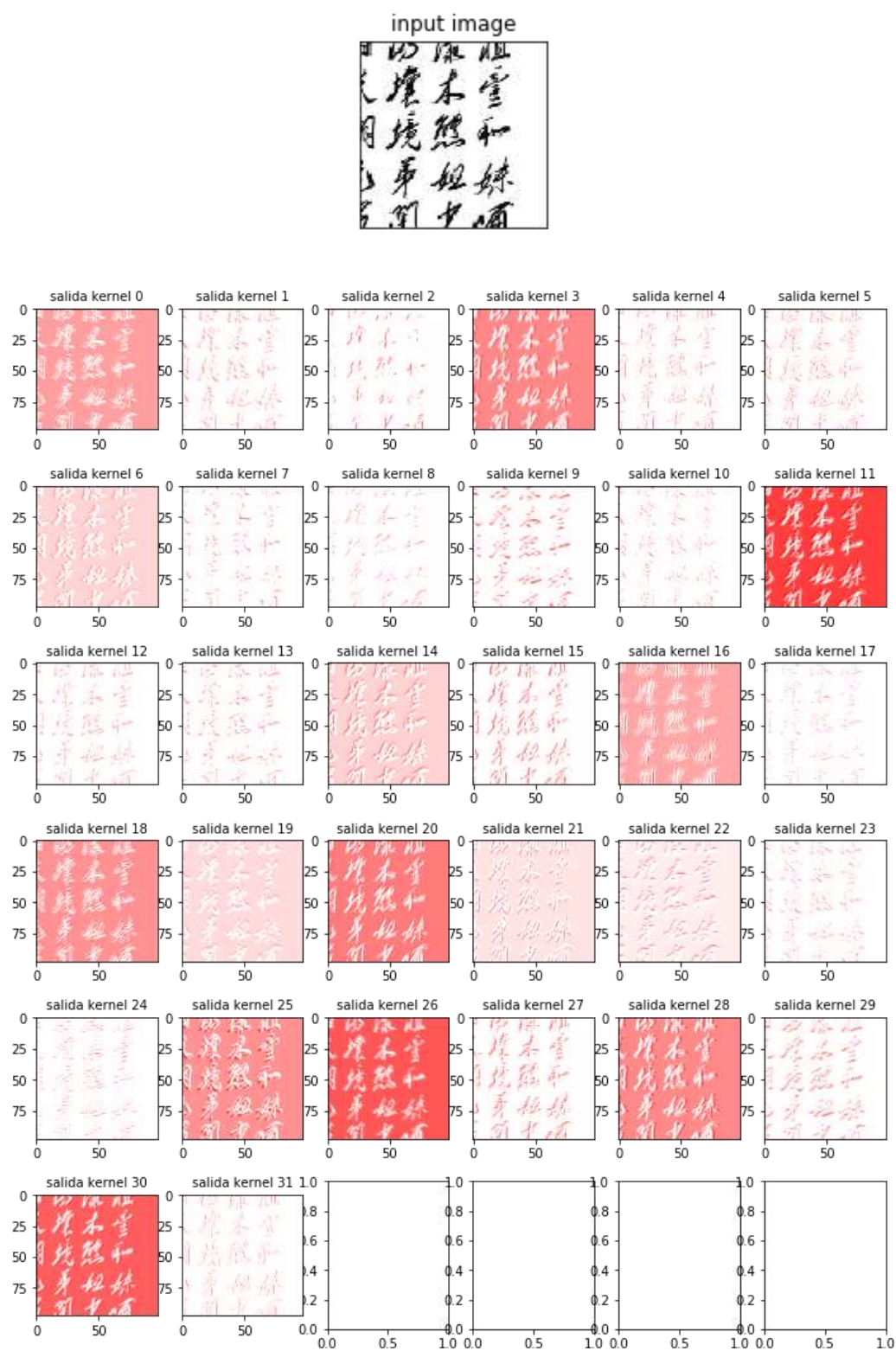


Figure 38. One example of features extracted by CNN (32 kernels)

Comparing with others techniques. Our research has some advantages than others:

1. Avoid cumbersome Chinese calligraphy segmentation

As mentioned earlier in the article, segmentation is a relatively complicated task with a relatively large amount of computation. Need to consider the special case of different handwritten character.

2. Improved recognition efficiency

Based on the full-page document, we improved recognition efficiency. For a large number of document recognition. The step of reducing unnecessary segmentation has also improved a lot in computer calculations.

3. Strong scalability

Relying on CNN's technical theory, we can continuously improve and add new technologies to improve recognition accuracy. For example, because you don't have to consider the features extracted by CNN, you can continuously modify and add new feature learning in the quality of the image. Increase the amount of data learning, including blurred, noisy, slanted pictures, to enrich the accuracy of identifying fonts.

8. Conclusion

From the experiments above, we may conclude that:

In this article, Chinese calligraphic style representation for recognition method was proposed. Based on full-page document with the CNN, we get the accuracy of Model_3 with dropout which is much higher than the other two (Model_1 and Model_2).

Our experiments show that Chinese calligraphic style recognition task, achieving 96.7% average recognition accuracy. In this Model_3, according to the f1-score we know the best performance is zhuan style (1.00), also we know the worst of one is Xing style. Other index also shows the same result.

For the reason that we could explain, in the structure of zhuan style and the recognition of zhuan style are clear and simple, compared to the Xing style is indeed easier to identify than others, and not just the machine, even the human itself is the same result. So that the experimental results are basically the same as our predictions.

From Experiment 1 that we could get the conclusion: Using the Model_1, It works with CNN that can do the recognition in full-page document. But it observes that a few data can affect with the training accuracy a lot. For the reason that is still need more data to training this model and make the Model_1 get more higher training accuracy.

From Experiment 2 that we could get the conclusion: In order to increase more data, we can't use the same pictures to be the dataset. Because it will be caused overfitting. In this case it needs to be avoid. Add some layers are more necessary in this experiment.

From Experiment 3 that we could get the conclusion: The quality and quantity of data is more important than the network structure in my experimental result. For the first experiment, with all various different pictures that it hard to let CNN to learn each style feature. But when we changed Dataset_B to Dataset_A, we got the best result is from our data that we created by ourselves. But Simple CNN structure with Dataset_B is works. Besides, with Experiment 3 we also find out that modify the layers of CNN network, can improve a lot with performance of the CNN. Even the dataset is so complicated.

All the experiments show that using machine learning to simulate the recognition of human visual observation images is an efficient method based on the full-page document. Using CNN can well avoid feature analysis and quickly give accurate analysis. No need to understand the

meaning of a single Chinese character, and give classification or recognition to the five style of Chinese calligraphy.

9.Future work

At the end, for the Method with Fourier analysis. We could go further. We think it should have some ways to do classification much better and easier. We have so many things need to be solved.

As we know, so many techniques can do classification. for example, we know all techniques from Softmax Regression, Support Vector Machine (SVM), k-Nearest Neighbors (k-NN) and Random Forests, which should be a good way to do classification after the FFT image. although if we can't classify the 5 classes completely in the same time. But we still can compare two styles of them.

On one hand, for example, discuss the good feature for exact from the calligraphic style information. if we can find the key to identify the font from a better calligraphy structure, it can more accurately identify more different characters. More often than not, it is easy to overlook that there are many changes in Chinese characters, which makes the researchers themselves should study the Chinese character culture itself rather than focusing on classification.

One the other hand, In the future, it has been suggested that a combination of non-stroke local features can be used to refine the style descriptor. This method has not been verified yet, but it can be tried as a direction. In that case we need to create the standard database which can support researchers compared their results with others results of reports in the same situation. This should be another subject to do and more useful.

REFERENCES

- [1] Liu K, Huang Y S, Suen C Y, “Robust Stroke Segmentation Method for Handwritten Chinese Character Recognition,” Proceedings of the Fourth International Conference on Document Analysis and Recognition, pp. 211-215, 1997
- [2] P. Ahmed and C. Y. Sum, “Computer recognition of totally un- constrained handwritten Zip codes,” Int. J. Pattern Recognition Artificial intell., vol. 1, pp.1-15, 1987
- [3] Yu K, Wu J, Zhuang Y, Skeleton-Based Recognition of Chinese Calligraphic Character Image, Advances in Multimedia Information Processing-PCM 2008, pp. 228-237, 2008.
- [4] Leung H, Wong S T S, Ip H H S, In the Name of Art, IEEE Signal Processing Magazine, 25(4): 49-54, 2008.
- [5] Li X, Ding X, Writer Identification of Chinese Handwriting Using Grid Microstructure Feature, Advances in Biometrics, pp. 1230-1239, 2009.
- [6] Zhuang Y, Lu W, Wu J, Latent Style Model: Discovering Writing Styles for Calligraphy Works, Journal Of Visua lCommunication And Image Representation, 20(2):84-96, 2009.
- [7] Han C C, Chou C H, Wu C S, An Interactive Grading and Learning System for Chinese Calligraphy, Machine Vision And Applications, 19(1):43-55, 2008.
- [8] “Convolutional Neural Networks (LeNet) Deep Learning 0.1 documentation”. Deep Learning 0.1. LISA Lab. Retrieved 31 August 2013.
- [9] Li Boqi, CS231N final project. Convolution Neural Network for Traditional Chinese Calligraphy Recognition. 2016
- [10] Xia, Z. Yang, K. Wang, “Chinese calligraphy word spotting using elastic HOG feature and derivative dynamic time warping”, Journal of Harbin Institute of Technology (New Series) 21(2):21-27 · March 2014
- [11] Lu W, Wu J, Wei B, Zhuang Y, “Efficient Shape Matching for Chinese Calligraphic Character Retrieval,” Zhejiang University Press, pp. 873-884, 2011
- [12] R. Gonzales, R. Woods., Digital Image Processing, Addison-Wesley Publishing Company, 1992, pp 81 - 125.

- [13] B. Horn Robot Vision, MIT Press, 1986, Chaps 6, 7.
- [14] A. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, 1989, pp 15 - 20.
- [15] Shi D M, Damper R I, Gunn S R. Offline handwritten Chinese character recognition by radical decomposition. ACM Transactions on Asian Language Information Processing. 2003, 1: 27-48
- [16] Kim I J, Kim J H. Statistical character structure modeling and its application to handwritten Chinese character recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2003, 25: 1422-1436
- [17] Lu Y, Shridhar M. Character segmentation in handwritten words-An overview. Pattern Recognition. 1996, 29: 77-96
- [18] Wei X H, Ma S P, Jin Y J. Segmentation of connected Chinese characters based on genetic algorithm. In: ICDAR'05: Proceedings of the Ninth International Conference on Document Analysis and Recognition, Seoul, Korea, Vol 1. IEEE Computer Society, 2005, 645-649
- [19] Liang Z Z, Shi P F. A metasynthetic approach for segmenting handwritten Chinese character strings. Pattern Recognition Letters. 2005, 26: 1498-1511
- [20] Zhao S Y, Chi Z R, Shi P F, et al. Two-stage segmentation of unconstrained handwritten Chinese characters. Pattern Recognition, 2003, 36: 145-156
- [21] Zhao S Y, Chi Z R, Shi P F, et al. Handwritten Chinese character segmentation using a two-stage approach. In: Proceedings of the sixth International Conference on Document Analysis and Recognition (ICDAR'01), Seattle, WA, Vol 1. IEEE Computer Society, 2001, 179-183
- [22] Tseng Y H, Lee H J. Recognition-based handwritten Chinese character segmentation using a probabilistic Viterbi algorithm. Pattern Recognition Letters. 1999, 20: 791-806
- [23] C.-L.Liu,F.Yin,D.-H.Wang,Q.-F.Wang, Chinese handwriting recognition contest 2010, in: Proceedings of Chinese Conference on Pattern Recognition (CCPR), 2010.
- [24] Gao J, Ding X q, Wu Y S. A segmentation algorithm for handwritten Chinese character strings. In: Proceedings of the Fifth International Conference on Document Analysis and Recognition (ICDAR'99), India, Vol 1. IEEE Computer Society. 1999, 633-636

- [25] F.Yin,Q.-F.Wang,X.-Y.Zhang,C.-L.Liu, ICDAR2013 Chinese handwriting recognition competition, in: Proceedings of International Conference on Document Analysis and Recognition(ICDAR), 2013,pp.1095–1101.
- [26] M. Hamanaka, K. Yamada, and J. Tsukumo, “On-Line Japanese Character Recognition Experiments by an Off-Line Method Based on Normalization-Cooperated Feature Extraction,” Proc. Third Int’l Conf. Document Analysis and Recognition, pp. 204-207, 1993.
- [27] M. Hamanaka and K. Yamada, “On-Line Character Recognition Adaptively Controlled by Handwriting Quality,” Proc. Seventh Int’l Workshop Frontiers in Handwriting Recognition, pp. 23-32, 2000.
- [28] T.K. Ho, J.J. Hull, and S.N. Srihari, “Decision Combination in Multiple Classifier Systems,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 1, pp. 66-75, Jan. 1994.
- [29] Zhao S Y, Chi Z R, Shi P F, et al. Two-stage segmentation of unconstrained handwritten Chinese characters. Pattern Recognition, 2003, 36: 145-156
- [30] Xiong Y, Huo Q, Chan C K. A discrete contextual stochastic model for the off-line recognition of handwritten Chinese characters. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23: 774-782
- [31] Charles C. Tappert ,“The State of the Art in On-Line Handwriting Recognition”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 8,1990
- [32] Dong J X, Krzyzak A, Suen C Y. An improved handwritten Chinese character recognition system using support vector machine. Pattern Recognition Letters. 2005, 26: 1849-1856
- [33] Wei X H, Ma S P, Jin Y J. Segmentation of connected Chinese characters based on genetic algorithm. In: ICDAR’05: Proceedings of the Ninth International Conference on Document Analysis and Recognition, Seoul, Korea, Vol 1. IEEE Computer Society, 2005, 645-649
- [34] Liang Z Z, Shi P F. A metasynthetic approach for segmenting handwritten Chinese character strings. Pattern Recognition Letters. 2005,26: 1498-1511
- [35] Zhao S Y, Chi Z R, Shi P F, et al. Two-stage segmentation of unconstrained handwritten Chinese characters. Pattern Recognition, 2003,36: 145-156

- [36] Zhao S Y, Chi Z R, Shi P F, et al. Handwritten Chinese character segmentation using a two-stage approach. In: Proceedings of the sixth International Conference on Document Analysis and Recognition (ICDAR'01), Seattle, WA, Vol 1. IEEE Computer Society, 2001, 179-183
- [37] Tseng Y H, Lee H J. Recognition-based handwritten Chinese character segmentation using a probabilistic Viterbi algorithm. *Pattern Recognition Letters*. 1999, 20: 791-806
- [38] C. Hong, G. Loudon, Y. Wu, and R. Zitserman, "Segmentation and Recognition of Continuous Handwriting Chinese Text," *Proc. Int'l Conf. Computer Processing of Oriental Languages*, pp. 630-633, 1997.
- [39] Liu H L, Ding X Q. Handwritten character recognition using gradient feature and quadratic classifier with multiple discrimination schemes. In: Proceedings of the Ninth International Conference on Document Analysis and Recognition, Korea, Vol 1. IEEE Computer Society, 2005, 19-23 42
- [40] Zeng J, Liu Z Q. Markov random fields for handwritten Chinese character recognition. In: ICDAR'05: Proceedings of the Ninth International Conference on Document Analysis and Recognition, Korea, Vol 1. IEEE Computer Society, 2005, 101-105
- [41] Kim I J, Kim J H. Statistical character structure modeling and its application to handwritten Chinese character recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2003, 25:1422-1436
- [42] Liu C L, Kim I J, Kim J H. Model-based stroke extraction and matching for handwritten Chinese character recognition. *Pattern Recognition*, 2001, 34: 2339-2352
- [43] Wang Q, Chi Z r, Feng D D, et al. Hidden Markov random field based approach for off-line handwritten Chinese character recognition. In: Proceedings of 15th International Conference on Pattern Recognition, Vol 2. 2000, 347-350
- [44] Fu C. Techniques for solving the large-scale classification problem in Chinese handwriting recognition. In: Summit on Arabic and Chinese Handwriting, College Park, USA. 2006, 87-92
- [45] R.T. Hummel and S.W. Zucker, "On the Foundations of Relaxation Labeling Processes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 5, no. 3, pp. 267-287, 1983.

- [46] Zhang R, Ding X Q. Minimum classification error training for handwritten character recognition. In: Proceedings of 16th International Conference on Pattern Recognition, Vol 1. 2002, 580-583
- [47] P. Ahmed and C. Y. Sum, "Computer recognition of totally un- constrained handwritten Zip codes," *Int. J. Pattern Recognition Artificial intell.*, vol. 1, pp.1-15, 1987
- [48] Long-Long Ma, Cheng-Lin Liu, A new radical-based approach to online handwritten Chinese character recognition, 2008 19th International Conference on Pattern Recognition, 8-11 Dec. 2008
- [49] Ge Y, Huo Q. A comparative study of several modeling approaches for large vocabulary offline recognition of handwritten Chinese characters. In: Proceedings of 16th International Conference on Pattern Recognition, Vol 3. 2002, 85-88
- [50] Ge Y, Huo Q. A study on the use of CDHMM for large vocabulary off-line recognition of handwritten Chinese characters. In: Proceedings of Eighth International Workshop on Frontiers in Handwriting Recognition, Canada, Vol 1. 2002, 334-338
- [51] Ge Y, Huo Q, Feng Z D. Offline recognition of handwritten Chinese characters using Gabor features, CDHMM modeling and MCE training. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol 1. 2002, I-1053-I1056
- [52] Dong J X, Krzyzak A, Suen C Y. An improved handwritten Chinese character recognition system using support vector machine. *Pattern Recognition Letters*. 2005, 26: 1849-1856
- [53] Feng B, Ding X Q, Wu Y S. Chinese handwriting recognition using hidden Markov models. In: Proceedings of 16th International Conference on Pattern Recognition, Vol 3. 2002, 212-215
- [54] Liu C L, Marukawa K M. Pseudo two-dimensional shape normalization methods for handwritten Chinese character recognition. *Pattern Recognition*. 2005, 38: 2242-2255
- [55] Wei X H, Ma S P, Jin Y J. Segmentation of connected Chinese characters based on genetic algorithm. In: ICDAR'05: Proceedings of the Ninth International Conference on Document Analysis and Recognition, Seoul, Korea, Vol 1. IEEE Computer Society, 2005, 645-649
- [56] Liang Z Z, Shi P F. A metasynthetic approach for segmenting handwritten Chinese character strings. *Pattern Recognition Letters*. 2005, 26: 1498-1511

- [57] Zhao S Y, Chi Z R, Shi P F, et al. Two-stage segmentation of unconstrained handwritten Chinese characters. *Pattern Recognition*, 2003, 36: 145-156
- [58] Zhao S Y, Chi Z R, Shi P F, et al. Handwritten Chinese character segmentation using a two-stage approach. In: *Proceedings of the sixth International Conference on Document Analysis and Recognition (ICDAR'01)*, Seattle, WA, Vol 1. IEEE Computer Society, 2001, 179-183
- [59] Xu-Yao Zhang, Yoshua Bengio, Cheng-Lin Liu, et al. *Online and Offline Handwritten Chinese Character Recognition: A Comprehensive Study and New Benchmark*. 2016.
- [60] Fu C. Techniques for solving the large-scale classification problem in Chinese handwriting recognition. In: *Summit on Arabic and Chinese Handwriting*, College Park, USA. 2006, 87-92
- [61] Liu H L, Ding X Q. Handwritten character recognition using gradient feature and quadratic classifier with multiple discrimination schemes. In: *Proceedings of the Ninth International Conference on Document Analysis and Recognition*, Korea, Vol 1. IEEE Computer Society, 2005, 19-23 42
- [62] Zhang R, Ding X Q. Minimum classification error training for handwritten character recognition. In: *Proceedings of 16th International Conference on Pattern Recognition*, Vol 1. 2002, 580-583
- [63] Xiong Y, Huo Q, Chan C K. A discrete contextual stochastic model for the off-line recognition of handwritten Chinese characters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23: 774-782
- [64] Feng B, Ding X Q, Wu Y S. Chinese handwriting recognition using hidden Markov models. In: *Proceedings of 16th International Conference on Pattern Recognition*, Vol 3. 2002, 212-215
- [65] Ge Y, Huo Q. A comparative study of several modeling approaches for large vocabulary offline recognition of handwritten Chinese characters. In: *Proceedings of 16th International Conference on Pattern Recognition*, Vol 3. 2002, 85-88
- [66] Ge Y, Huo Q. A study on the use of CDHMM for large vocabulary off-line recognition of handwritten Chinese characters. In: *Proceedings of Eighth International Workshop on Frontiers in Handwriting Recognition*, Canada, Vol 1. 2002, 334-338

- [67] Ge Y, Huo Q, Feng Z D. Offline recognition of handwritten Chinese characters using Gabor features, CDHMM modeling and MCE training. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol 1. 2002, I-1053-I- 1056
- [68] Wu M R, Zhang B, Zhang L. A neural network based classifier for handwritten Chinese character recognition. In: Proceedings of 15th International Conference on Pattern Recognition, Vol 2.2000, 561-564
- [69] Wang C H, Xiao B H, Dai R W. A new integration scheme with multilayer perceptron networks for handwritten Chinese character recognition. In: Proceedings of 15th International Conference on Pattern Recognition, Vol 2. 2000, 961-964
- [70] Charles C. Tappert ,“The State of the Art in On-Line Handwriting Recognition”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 8,1990
- [71] Liu C L, Kim I J, Kim J H, “Model-Based Stroke Extraction and Matching for Handwritten Chinese Character Recognition,” Pattern Recognition, 34(12):2339-2352, 2001.
- [72] Yu K, Wu J, Zhuang Y, “Skeleton-Based Recognition of Chinese Calligraphic Character Image,” Advances in Multimedia Information Processing-PCM 2008, pp. 228237, 2008.
- [73] Chen, Yu-Sheng, Guangjun Su, and Haihong Li. "Machine Learning for Calligraphy Styles Recognition." ,2016
- [74] Zhang Jiulong; Guo Luming; Yang Su; Sun Xudong; Li Xiaoshan, Detecting Chinese calligraphy style consistency by deep learning and one-class SVM, 2017
- [75] Gao Pengcheng, Gu Gang, Wu Jiangqin, Wei Baogang,Chinese calligraphic style representation for recognition,18 January 2017

Additional Bibliography

- [76] B. Horn Robot Vision, MIT Press, 1986, Chaps 6, 7.
- [77] A. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, 1989, pp 15 - 20.
- [78] (<https://homepages.inf.ed.ac.uk/rbf/HIPR2/fourier.htm>)
- [79] Steven W. Smith, Ph.D, The Scientist and Engineer's Guide to Digital Signal Processing

- [80] Xu S, Lau F C M, Cheung W K, Pan Y, “Automatic Generation of Artistic Chinese Calligraphy,” IEEE Intelligent Systems, 20(3): 32-39, 2005.
- [81] Xu S, Jiang H, Lau F C M, Pan Y, “An Intelligent System for Chinese Calligraphy,” AAAI’07 Proceedings of the 22nd national conference on Artificial intelligence, pp. 1578-1583, 2007.
- [82] Wong H T F, Ip H H S, “Virtual Brush: A ModelBased Synthesis of Chinese Calligraphy,” Computers & Graphics, 24(1): 99-113, 2001.
- [83] Mi X F, Tang M, Dong J X, “Droplet: A Virtual Brush Model to Simulate Chinese Calligraphy and Painting,” Journal Of Computer Science And Technology, 19(3):393404, 2004.
- [84] Zhuang Y, Lu W, Wu J, “Latent Style Model: Discovering Writing Styles for Calligraphy Works,” Journal Of Visual Communication And Image Representation, 20(2):84-96, 2009.
- [85] Han C C, Chou C H, Wu C S, “An Interactive Grading and Learning System for Chinese Calligraphy,” Machine Vision And Applications, 19(1):43-55, 2008.
- [86] Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T.: Decaf: a deep convolutional activation feature for generic visual recognition. In: Proceedings of The 31st International Conference on Machine Learning, pp. 647–655 (2014)
- [87] Oquab, M., Bottou, L., Laptev, I., et al.: Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1717–1724 (2014)