



Filtering Approaches and Mish Activation Function Applied on Handwritten Chinese Character Recognition

Zhong Yingna^a, Kauthar Mohd Daud^a, Kohbalan Moorthy^b
and Ain Najiha Mohamad Nor^a

^a Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia, Bangi, Selangor, 43600

^b Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan, Pahang, 26600

✉ p110117@siswa.ukm.edu.my, kauthar.md@ukm.edu.my, kohbalan@ump.edu.my, ainnajiha46@gmail.com

KEYWORDS

activation functions;
convolutional
neural network;
CNN; filtering
approaches;
machine learning.

ABSTRACT

Handwritten Chinese Characters (HCC) have recently received much attention as a global means of exchanging information and knowledge. The start of the information age has increased the number of paper documents that must be electronically saved and shared. The recognition accuracy of online handwritten Chinese characters has reached its limit as online characters are more straightforward than offline characters. Furthermore, online character recognition enables stronger involvement and flexibility than offline characters. Deep learning techniques, such as convolutional neural networks (CNN), have superseded conventional Handwritten Chinese Character Recognition (HCCR) solutions, as proven in image identification. Nonetheless, because of the large number of comparable characters and styles, there is still an opportunity to improve the present recognition accuracy by adopting different activation functions, including Mish, Sigmoid, Tanh, and ReLU. The main goal of this study is to apply a filter and activation function that has a better impact on the recognition system to improve the performance of the recognition CNN model. In this study, we implemented different filter techniques and activation functions in CNN to offline Chinese characters to understand the effects of the model's recognition outcome. Two CNN layers are proposed given that they achieve comparative performances using fewer-layer CNN. The results demonstrate that the Weiner filter has better recognition performance than the median and average filters. Furthermore, the Mish activation function performs better than the Sigmoid, Tanh, and ReLU functions.



1. Introduction

Handwritten Chinese Characters (Hànz), go by a variety of names, including «Sinograms» (from the Greek term for China), «Hanja» (from Korean), and «Kanji» (from Japanese). They can be grouped as one of the most sophisticated writing systems used nowadays. Since these characters are considered very difficult to write, a standard ballpoint pen or most gel pens cannot achieve the tapering and broadening strokes that give beauty to Chinese characters. However, the appearance of these characters can be enhanced by means of soft lead and a fountain pen with a high-quality nib. Although everything in life is heading towards electronic and technological developments, handwritten Chinese characters remain to have a variety of functions, especially in financial bills, reports, and many others. With the advancement of technology, most characters have been computerized. Nevertheless, in writing, Chinese characters are quite challenging due to varying writing styles among individuals and the numerous identical characters they feature.

Given the advances in technology, a new approach has been developed using deep learning. In the field of science and technology, deep learning is considered one of the key research techniques. Deep learning derives from building artificial neural networks based on the inherent logic of sample data. Then, it is possible to forecast similar data types, making deep learning suitable for tasks such as speech recognition, image classification, and text recognition. For Handwritten Chinese characters, image classification and text recognition are two essential aspects. A convolutional neural network (CNN) is a sort of artificial neural network that is mainly used for image recognition and processing because of its capacity to spot different patterns in visual data. Although a CNN is a powerful tool, its training process necessitates millions of tagged data points.

There has been numerous research on offline Handwritten Chinese Character Recognition (HCCR). However, it is still challenging to achieve high accuracy for actual recognition due to the intricacy of Chinese characters and the wide variety of handwriting styles. Therefore, the handwriting component must be simplified by preprocessing the data to improve the HCCR's recognition accuracy. The primary target of this paper is to improve recognition performance by applying a filter and activation function that has a better effect on the recognition of Chinese offline characters.

This paper is organized as follows: the related methods for Handwritten Chinese Character Recognition are overviewed in Section 2; the definition and concept of convolution neural network is presented in Section 3. In Section 4, we describe the proposed method and the employed dataset. We then report and discuss the results in Section 5. Finally, the paper is concluded in Section 6.

2. Handwritten Chinese Character Recognition Methods

There are two distinct recognition methods for handwritten Chinese characters: online and offline. Online and offline recognition generally use distinct standardization and feature extraction techniques; however, they typically use the same classification techniques. Online handwriting recognition is known to be based on pen trajectory as it takes the input, such as the stroke of a pen on a digital surface during the run time. On the other hand, in offline recognition, the character image is analyzed and divided into various classes (Liu et al., 2004). Online handwriting data is dynamic as it stores the points of writing in the function of time, making it easier and faster to distinguish the different characters than offline characters. The difficulty of both character recognition systems differs because of the diverse input data. Furthermore, online recognition is more straightforward than offline recognition

because writers can edit their work after submission and vary their writing styles in response to recognition results, thus fostering robust engagement and adaptability. Due to many characters and varied writing styles, offline HCCR research is still complex. Since online character recognition has produced excellent results with high performance, such as the research of Zheng et al. (1997), which yielded an identification result of 98.8 %; currently, the research focus has shifted to less restricted HCCR (Liu et al., 2004).

In contrast to online HCCR, offline recognition is concentrated on examining and categorizing character images (grayscale or binary images) into various classes. Offline HCCR techniques can be divided into two categories: classical approaches and deep learning approaches. Before deep learning techniques were effectively applied, traditional recognition methods were typical. CNN is recognized as the most well-liked deep learning method for recognition problems. Feature extraction, classification, and data preprocessing are the three primary parts of the traditional HCCR system.

- Several feature extraction methods have been utilized in picture processing, such as the Gabor and Histogram of Oriented Gradient (HOG) feature gradational quality. Studies have revealed that in HCCR, the use of the Gabor feature can produce competitive results (Daugman, 1988; Ge et al., 2002). Several investigations also attained greater accuracy using Gabor Feature (Alom et al., 2017). Moreover, gradient features can be extracted from grayscale or binary images. According to research, using the gradient feature extraction approach is an efficient way to boost HCCR's efficiency (Shi et al., 2002). The HOG feature is recommended in computer vision as an essential quality for effective performance (Dalal and Triggs, 2005). Another study proposed the application of the HOG method to enhance the performance of CNN (Zhong et al., 2015).
- Numerous statistical classifiers have been used extensively in recent years, and the classification phase for HCCR comprises the modified quadratic discriminant function (MQDF), discriminative learning quadratic discriminant function (DLQDF), and nearest prototype classifier (NPC) (Dai et al., 2007). Several experiments have received significant acclaim for applying standard classification techniques to improve HCCR performance (Liu and Ding, 2005; Liu, 2006; Dai et al., 2007; Zhong et al., 2015). Although there has not been much progress in recent years, the conventional techniques that depend on these classifiers have no further progress. Therefore, the limitations of conventional Chinese character recognition techniques leave a significant performance gap with humans.
- Data preprocessing is a crucial strategy for improving the effectiveness of the recognition system. The benefit of CNN is that it does not rely on input (Kavitha and Srimathi, 2019); however, the noise will affect recognition accuracy (Zhuang et al., 2021). A CNN and median filtering model for HCCR was created by (Zhuang et al., 2021) to smooth and diminish picture noise. Their study showed that the generated model's output improved by 1.8 % compared to the CNN model without median filtering. Filtering is a typical method for removing noise in data preprocessing. In image preprocessing, the images are modified and enhanced to extract the essential information by removing undesirable features with augmentation or modification. According to Kumar et al. (2021), the noise removal approach can improve the accuracy and quality of the images. Image noise can be reduced using basic filtering methods including the Wiener, median, and average (mean) filters. The researchers discovered that the Wiener filter performs better than the median and average filters for all types of Gaussians, Poisson, and Speckle noise (Shetti and Patil, 2017; Dass and Saini, 2022). In contrast, the median filter outperforms the other two filters in terms of Salt & Pepper noise (Patidar et al., 2010; Srinivas and Panda, 2013; Arya and Samwel, 2017).

2.1. Median Filter

The median is a type of non-linear filter and is the most used order-statistics filter. Due to their superior noise removal and much lower distortion than linear filters of equivalent size, median filters are preferred (Ng and Ma, 2006). This type of noise removal is a typical data preparation method that improves the results of further processing, including boundary recognition on images. Since it can preserve boundaries while performing noise removal operations under specific conditions, the median filter is commonly used to process digital images. The ability of the median filter to reduce the effects of extremely high input noise values is a crucial advantage, even though there are relatively few outliers noise values (Church et al., 2008; Srivinas and Panda, 2013). The median filtering technique replaces a value with the median of its nearby values closer to the real value, thus eliminating specific noisy regions. The following equation determines the median value:

$$f(x, y) = \underset{(s, t) \in S_{xy}}{\text{median}} \{g(s, t)\}$$

Where S_{xy} refers to the set of coordinates inside the square filter's or mask's rectangular window, the function $f(x, y)$ can be determined by calculating the median using the area's pixel value.

2.2. Average Filter

The average filter is another type of linear filter that has been used to remove noise in data. The average filter is also referred to as the mean filter. It smooths the image by lessening the intensity variance between a pixel's neighboring pixels. This filter is typically used for noise suppression and performs poorly at retaining edges (Srivinas and Panda, 2013). This filtering method operates like the median filter. The average filtering approach replaces the unrepresentative pixels in the surrounding area by replacing a point's value with the mean of its surrounding values. The mean value is calculated using the formula below:

$$f(x, y) = \frac{1}{m * n} \sum_{(s, t) \in S_{xy}} g(s, t)$$

Where (m, n) refers to the filter size, and the function $f(x, y)$ is obtained by calculating the average value using pixel values in the area.

2.3. Wiener Filter

The objective of a wiener filter is to eliminate the noise that distorts signals. The wiener filter employs filtering from a certain angle and employs a statistical analytical method. Typically, filters are created for the variable frequency response. Suppose the spectral characteristics of the initial signal and the noise are known. Thus, the linear time-invariant (LTI) filter whose output is as close as possible to the original signal is desired (Kazubek, 2003). The following formula describes the Wiener filter:

$$G(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 P_s(u, v) / P_n(u, v)}$$

Where P_n/P_s is the reverse of the signal-to-noise ratio, $P_n(u, v)$ is the power spectral density of noise, and $P_s(u, v)$ is regarded as the ungraded image. Meanwhile, $H(u, v)$ refers to the functions of degradation, while $H^*(u, v)$ refers to the complex conjugate of a degradation function.

3. Convolutional Neural Networks

The deep learning hypothesis offers a fresh approach to making strides and reducing the gap between machines and humans. It has the potential to anticipate the behaviour of complex data due to the inclusion of numerous hidden layers inherited from the regular neural network architecture (Aljojo, 2022). Deep learning approaches have primarily replaced traditional approaches to HCCR because of their success in a variety of pattern recognition tasks, such as character recognition (Ciresan et al., 2011; Goodfellow et al., 2013), image recognition (Ciresan et al., 2011; Szegedy et al., 2016), and human face recognition (Taigman et al., 2014). The most popular deep learning method, convolutional neural networks (CNN), has been widely used, particularly for image recognition applications, due to the properties of CNN, which are the weight-sharing method and the local-connection strategy (Chen et al., 2015). Another outstanding study, conducted by Yadav et al. (2018), employs CNN for Hindi characters to automate manual tasks and therefore save time.

The successful implementation of CNN on HCCR has been reported for the first time using the multi-column deep neural network (MCDNN) model (Arya and Semwal, 2017). MCDNN was created by combining several deep neural network columns. Following that, Ciregan and Schmidhuber (2012) introduced a spatially sparse CNN that took first place in the ICDAR2013 competition as its performance successfully surpassed humans. A CNN-based system was introduced by Chen et al. (2015), and it was also more effective than humans at recognizing Chinese characters (CASIA).

Researchers have documented several beneficial and significant accomplishments CNN application in HCCR. A significant aspect of CNN models is their efficiency in recognition tasks. The research on this study topic has been continuous, and as a result, the HCCR's accuracy has been consistently improved in recent years in several different ways.

The modified model of the adaptive drop-weight (ADW) and global supervised low-rank expansion (GSLRE) approach has outperformed the best single network of CNN. It also significantly improved running time and reduced storage than the best single network of CNN (Xiao et al., 2017). A research team developed a 9-layer CNN baseline model that achieved 97.30 % recognition accuracy (Zhang et al., 2017). Moreover, the researchers created a 12-layer convolutional neural network model with a recognition accuracy of 97.59 %. Therefore, a larger and more complex neural network performs better than a smaller and less complex neural network.

Moreover, a study proposed a cascaded model that combined the fire module with the global weighted average pooling (GWAP) technique in a fully connected layer (Chen et al., 2015; Iandola et al., 2016). Their model performs at 97.14 %, and it takes less time to classify a character image than Xiao and Jin et al. (2017).

Furthermore, Global Weighted Output Average Pooling (GWOAP) was used to optimize from the GWAP and allowed production class activation maps (CAMs) for finding the most identifiable area of a picture. Melnyk et al., (2020) built an enhanced system based on that of Li et al. (2018). The outcome demonstrated an increase in accuracy of 0.02 % and almost negligible storage as the model HCCR-CNN12Layer (Xiao et al., 2017).

A few CNN models with the three loss functions (cross-entropy, a combination of cross-entropy and Euclidean distance, and a combination of cross-entropy and average variance similarity) were

proposed by Zou et al. (2019). The authors' research concludes that the loss functions of the combination between cross-entropy and average variance similarity produced the highest accuracy for HCCR. The results indicate that applying similarity ranking can simultaneously reduce intra-class variation and increase inter-class variation, thus leading to a considerably superior recognition result.

Bi et al. (2019) made some advanced improvements based on GoogLeNet (Szegedy et al., 2016). The proposed improved models successfully outperformed all the CNN models available then, in terms of recognition accuracy. Although the recognition accuracy increased and the training time decreased, it affected the number of parameters and storage.

Gan et al. (2020) created a CNN model, in-air HCCR, in which it uses Channel Pruning (CP), Drop-Weights (DW), and Incremental Network Quantization (INQ) techniques. The updated model results showed that the proposed model has visible advantages over other advanced techniques for offline HCCR (Min et al., 2020), even though both models were tested on the offline ICDAR-2013 dataset with the same recognition accuracy.

Based on the research of Szegedy et al. (2016), Min et al. (2020) constructed an improved shallow GoogLeNet. The authors added an error elimination algorithm to the model to remove incorrect data. The proposed model not only improves accuracy but also minimizes training parameters.

A new approach for HCCR has been developed by Li et al. (2020) called template matching, which makes it possible to determine how much character images resemble template images. According to their research findings, the proposed approach can recognize the «open set» or, more specifically, distinguish the characters that are not used during the training phase. The authors' research revealed that the proposed method could identify the «open set,» or characters that are seldom used for training.

A two-stage hierarchical deep CNN was designed by Aleskerova and Zhuravlev (2020). Even though the final accuracy was not as great as earlier results from the HCCR, the proposed hierarchical model outperformed the standard CNN in recognition accuracy. Additionally, the study found that a hierarchical technique may handle the classification problem with so many categories better than a typical one-stage CNN.

For handwritten Chinese text recognition (HCTR), Liu et al. (2020) suggested a CNN-only model. While the fully connected layers of the model architecture are different, all the convolution layers are based on the recognized VGG-16 network (Xiao et al., 2017). The CNN+CTC method was used in the HCTR domain for the first time, whereby CTC was used as the loss function. The proposed method outperformed all other methods in the character error rate benchmark on the ICDAR 2013 without a language model. The outcome indicated that CNN is effective for modelling a model with many classes, and the CTC loss function is suitable for categorization issues.

A residual-attention CNN-based model was proposed by Wang et al. (2021), which extracts the representative features to reduce the weight of unimportant features. A customized CNN model with only three layers—two convolutional, two pooling, and one fully connected—was proposed by Ameri, Alameer et al. (2021) for handwritten Chinese number classification. This model outperforms deep CNN models with larger architectures such as SqueezeNet (Iandola et al., 2016), MobileNet (Sandler et al., 2018), and GoogleNet (Amodei et al., 2016).

One study also created a «privately tailored» CNN-based recognition algorithm (Li and Li, 2021). The authors constructed a user-specific dataset by training the model with CASIA-HWDB1.1 and manually correcting the incorrect pre-identified character labels. Their system's recognition accuracy for English letters, Chinese characters, Arabic numerals, and punctuation marks may reach over 98 %.

A better CNN model was created by Liu and Ding (2005) to identify identical handwritten Chinese characters. The proposed model's accuracy was 97 %, showing that it outperformed existing

classification techniques such as SVM, regular neural network, and logistic regression in classifying visually comparable characters.

A model that combines attention features spatial aggregation (ASA) and multiple scale convolution shuffle (MSCS) was proposed by Xu et al. (2022), of which the ASA module is mainly used for model compression, and the MSCS module is primarily used for feature extraction. In terms of inference time and storage, the proposed model produced a competitive performance that was the best in single networks. A summary of the more detailed findings from the study on current performance improvement is provided in Table 1.

Table 1. Summary of Literature on Convolutional Neural Networks (CNN) in HCR

References	Method	Accuracy	Dataset	Script
(Xiao et al., 2017)	CNN-9Layer; CNN-12Layer	97.30 % 97.59 %	ICDAR-2013 Offline HCCR competition	Chinese
(Chen et al., 2015)	Cascaded CNN Model + WAP	97.14 %	ICDAR-2013 Offline HCCR competition	Chinese
(Melnik et al., 2020)	CNN-15Layer + GWOAP +CAMs	97.61 %	ICDAR-2013 Offline HCCR competition	Chinese
(Li et al., 2018)	CNN + Similarity Ranking Function	95.58 %	CASIA- HWDB1.1	Chinese
(Zou et al., 2019)	GoogLeNet + Several Adjustments	98.20 %	CASIA- HWDB1.1	Chinese
(Gan et al., 2020)	CNN + CP2 +DW + INQ (8-bit)	96.97 %	ICDAR-2013 Offline HCCR competition	Chinese
(Min et al., 2020)	Improved Shallow GoogLeNet + Error Elimination Algorithm	97.48 %	HWDB1.1- 1CASIA- HWDB1.1	Chinese
(Li et al., 2020)	Siamese Network + CNN Based + Template Matching	92.31 %	ICDAR-2013 Offline HCCR competition	Chinese
(Aleskerova and Zhuralev, 2020)	Two-stage Hierarchical CNN	92.10 %	ICDAR-2013 Offline HCCR competition	Chinese
(Liu et al., 2020)	Residual attention + fully CNN	97.32 %	CASIA-HWDB test set	Chinese
(Wang et al., 2021)	Customized CNN-5Layer	99.10 %	Handwritten Chinese Number	Chinese
(Amodei et al., 2016)	Data Collection + Label Image + CNN	98.137 %	CASIA-HWDB1.1 (for training); Collect pictures to build the test dataset	Chinese; English; Arabic Digit; Punctuation Marks
(Liu and Ding, 2005)	Enhanced CNN + Adam	97 %	Self-collect dataset	Chinese
(Xu et al., 2022)	MSCS + ASA	97.63 %	ICDAR-2013 Offline HCCR competition	Chinese

3.1. Activation Functions

The activation layer follows the convolutional layer and introduces non-linearity to the system, allowing the network to fit the data accurately. In deep learning networks, the activation layer is essential since it increases the choice function needed by the multi-layer network and the network's overall non-linearity without affecting the receptive field of the convolutional layer (Hao et al., 2020). Different neural networks perform differently when different activation functions are used (Hao et al., 2020; Sun, 2021). The most used activation functions include Tanh, Sigmoid, and ReLU. Rectified linear units (ReLU) are recommended as they train neural networks more quickly than other functions (Sakib et al., 2019).

The ReLU function is a popular activation function used in deep learning neural networks as it is simple to compute and has the desirable property of introducing non-linearity into the output of the neuron. Therefore, it allows for modelling more complex relationships between inputs and outputs. However, it has certain limitations, such as when a neuron becomes inactive during training, it hinders the network's learning process. Mish is a new activation function first presented by Misra (2019), which features a combination of Sigmoid and ReLU functions. Due to its advantages over other activation functions, such as improved training speed, accuracy, and robustness, Mish functions have outperformed other highest-performing networks, especially in computer vision tasks such as object identification and image categorization. The Mish activation function has the following characteristics: smoothness, continuity, self-regularization, non-monotonic, and differentiation.

The Mish function was added to LENET- 5 CNN for the Minist dataset (Zhang et al., 2017). The authors compared the performance of Mish, ReLU, and Sigmoid activation functions in LENET-5 CNN. The study shows that Mish had higher accuracy and convergence speeds than the ReLU and Sigmoid functions. The accuracy of Mish was 13.93 % greater than Sigmoid's and 3.54 % higher than ReLU's. Meanwhile, ReLU had a higher accuracy and convergence speed than the Sigmoid activation function. The authors concluded that the Mish activation function is efficient and accurate in determining the Minst dataset.

The study by Guo et al. (2020) lays the groundwork for future research that will examine the impact of parameters and functions, such as the activation functions, loss functions, and CNN model learning rate, on the performance of CNN. According to their research on the HWDB dataset, the Mish activation function outperforms the Sigmoid, Tanh, ReLU, and Mish functions regarding recognition rate and stability. Table 2 shows the performance of different activation functions.

Table 2. Performance Comparisons of Different Activation Functions

References	Model	Dataset	Accuracy (%)			
			Sigmoid	Tanh	ReLU	Mish
(Zhang et al., 2017)	LENET-5 CNN	MNIST	84.16	-	94.55	98.09
(Guo et al., 2020)	CNN-based	HWDB	10	85	87	90
(Sriporn et al., 2020)	NasNetMobile + Nadam	Malaria Image	-	-	88.64	91
	Inception-V3 + Nadam	Malaria Image	-	-	93.5	95.21
	Xception + Nadam	Malaria Image	-	-	98.07	99.28
	AlexNet + Nadam	Malaria Image	-	-	81.64	82.71
	VGG-16 + Nadam	Malaria Image	-	-	84.35	85
	ResNet-50 + Nadam	Malaria Image	-	-	92.5	93.07

(continued)

Table 2. Performance Comparisons of Different Activation Functions (continued)

References	Model	Dataset	Accuracy (%)			
			Sigmoid	Tanh	ReLU	Mish
(Maniatopoulos and Mitianoudis, 2021)	CNN3ConvLayer	MNIST Fashion	-	98.35	98.75	98.81
	CNN5ConvLayer	MNIST	-	89.79	89.56	90.37
	CNN5ConvLayer	CIFAR-10	-	67.85	68.29	69.38
	ResNet-v1-56	CIFAR-100	36.47	63.88	69.64	69.88
(Zhu et al., 2021)	ResNet-50	CIFAR-10	88.8	87.5	89.6	92.4
	DenseNet-201	CIFAR-10	79.82	81.84	86.89	88.17
	SE-Inception ResNet-v2	CIFAR-10	76.38	74.23	76.12	80.71
	ResNet	MNIST	98.87	99.24	98.76	99.47
	DenseNet	MNIST	98.42	99.07	98.91	99.04
	SENet	MNIST	98.81	98.64	98.01	98.94
(Liu and Di, 2021)	MobileNet	CIFAR-10	-	-	84.12	85.27
	Resnet 32	CIFAR-10	-	-	91.78	92.29
	ShuffleNet	CIFAR-10	-	-	87.05	87.31
	MobileNet	CIFAR-100	-	-	49.21	51.93
	Resnet 32	CIFAR-100	-	-	68.45	69.44
	ShuffleNet	CIFAR-100	-	-	57.98	59.19

- indicates not tested in the research experiment.

Researchers typically used deep learning approaches to extract meaningful features from those datasets. Although deep learning-based frameworks successfully improved classification accuracy, the computational complexity was affected. Given the ubiquity of methods developed in HCCR and the availability of significant character datasets, CNN has been utilized extensively for optical character recognition. It has been observed that researchers are increasingly improving CNN for the recognition of offline Chinese characters. Furthermore, the recognition rate is significantly reduced when dealing with real and unconstrained handwritten samples due to the complexity of Chinese characters and the diversity of styles. Therefore, due to these factors, the problem of improving the recognition accuracy of HCCR is still a sophisticated challenge.

4. Materials and Methods

Considering that this work involves Chinese character information and computation methods, in this section, we introduce the proposed two CNN layers for improving the accuracy of the recognition model. Several filtering techniques have been applied to the data, and the effects are compared for each filtering technique. Moreover, according to the extensive literature study, Mish activation produced higher accuracy and convergence speed. Therefore, the Mish activation function is incorporated into two layers of CNN. Apart from the Mish function and the default ReLU function used throughout the experiments, the Sigmoid and Tanh activation functions are also included here for comparison.

Figure 1 illustrates the experimental design of the proposed model. The design is divided into two parts: processing data with different filters and data training on the CNN model. In the first part of the experiment design, three different filters (median, average, and Wiener filters) were applied to the preprocessed data, which is discussed further on in this section. The second part of this experiment is to generate recognition results by inputting the processed data into CNN.

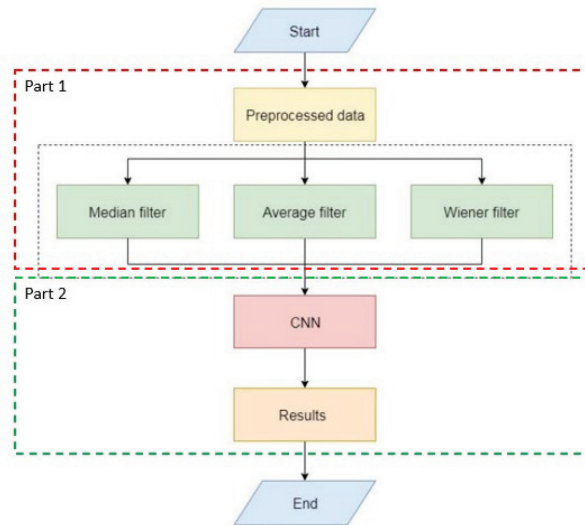


Figure 1. The experiment design for filtering (Proposed Model)

As shown in Figure 2, the training data and testing data used in this experiment were processed by the Wiener filter. Then, four CNN models with different activation functions were trained and tested.

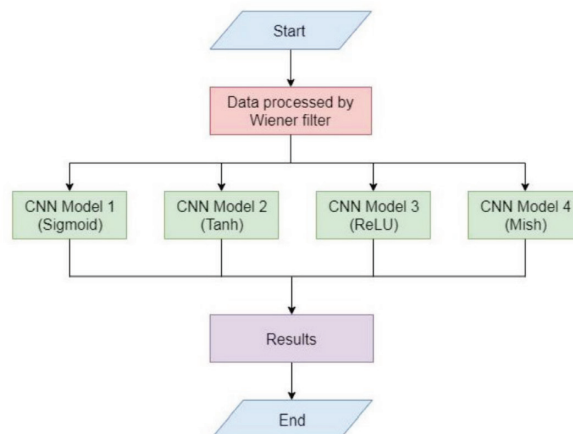


Figure 2. The experiment design for activation functions

4.1. Dataset

This research used the offline CASIA-HWDB 1.1 to train and test the proposed model experiment (Liu et al., 2011). The dataset can be downloaded for free on this website <http://www.nlpr.ia.ac.cn/databases/handwriting/Download.html>. The dataset consists of binary images of characters handwritten using Aoto pens. A summary of the character dataset statistics is given in Table 3. There are 171 alphanumeric and symbol characters, 3,755 Chinese characters, and 1 172,907-character samples in the HWDB1.1 dataset. The 300 writers are divided into 240 and 60 for training and testing data, respectively.

Table 3. The statistics of character datasets

Dataset	Writers		Chinese Characters / Class		Symbol Samples / Class	Total Samples
	Train	Test	Train	Test		
HWDB1.1	240	60	897,758 / 3,755	223,991 / 3755	51,158 / 171	1,172,907

Given that the dataset contains an enormous number of image samples, the training and testing process may take a very long period, necessitating a powerful computer. Therefore, only 20 classes were selected in the train and test dataset, as shown in Figure 3 (Zhang et al., 2017).

CASIA-HWDB1.1 DATASET					No. of Character	
Character Samples	00001(丁)	00002(七)	00003(万)	00004(丈)	00005(三)	100(train)
	00006(上)	00007(卜)	00008(不)	00009(与)	00010(丑)	
	00011(专)	00012(且)	00013(世)	00014(丘)	00015(丙)	60(test)
	00016(业)	00017(丛)	00018(东)	00019(丝)	00020(丢)	
Total	160*20=3200					

Figure 3. Character samples used in the first study

In addition, 10-character datasets from the CASIA-HWDB1.1 dataset were selected for the second experiment, which had the same amount of character samples as the research of Guo et al. (2020). The specifics of the datasets used in the second investigation are provided in Figure 4. This experiment consisted of six sets. Each set consisted of 180 training samples and 20 testing samples for each character, which accounted for 2 000 character samples.

Sets	Character Samples					No. of each Character
Set-0	00031(丹)	00032(为)	00033(主)	00034(丽)	00035(举)	180(train)
	00036(乃)	00037(久)	00038(么)	00039(义)	00040(之)	
Set-1	00018(东)	00029(临)	00034(丽)	00065(事)	00075(亡)	
	00079(亦)	00083(享)	00090(什)	00095(今)	00100(仓)	
Set-2	00024(个)	00034(丽)	00051(乙)	00081(亨)	00082(亩)	
	00091(仁)	00105(付)	00121(伊)	00140(伶)	00173(尧)	
Set-3	00029(临)	00038(么)	00041(乌)	00046(兵)	00047(兵)	
	00057(书)	00059(乱)	00080(产)	00086(亮)	00088(人)	
Set-4	00011(专)	00033(主)	00049(乖)	00057(书)	00106(仙)	
	00143(似)	00149(佐)	00163(使)	00168(伺)	00170(依)	
Set-5	00023(丧)	00029(临)	00036(乃)	00037(久)	00047(兵)	20(test)
	00106(仙)	00153(余)	00175(侧)	00182(促)	00200(催)	

Figure 4: Character samples used in the second study.

4.2. Data Preprocessing

Before convolutional neural network training, data pre-processing is essential. Data pre-processing is carried out in almost every data-related operation to clean and alter the data, convert text into binary or numerical values, and organise missing items. It plays a key function in the fields of machine learning and deep learning (Sakarkar et al., 2021). In our experiment, the preprocessing stage included the following operations: grayscale conversion, binarization, and scaling. The overview of data preprocessing used in this study is shown in Figure 5.

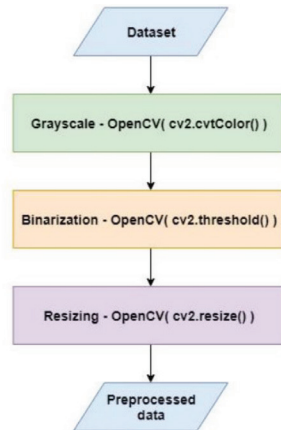


Figure 5. The overview of the pre-processing

5. Results and Discussion

5.1. Filtering Approaches

As mentioned in the previous section, the total 3 200 samples of the data set were split into training and testing sets, 2 000 and 1 200 samples, respectively. Table 4 provides the results of different filters according to different epochs.

Table 4. The result of different filters according to different epochs

Table of Accuracy of Different Filters According to Different Epochs							
Accuracy	Epoch						
	10	100	300	500	1000	1500	2000
Median	21.08 %	84.17 %	87.67 %	89.08 %	89.42 %	90.17 %	91.08 %
Average	5.08 %	83.08 %	88.25 %	89.75 %	89.75 %	89.83 %	90.08 %
Wiener	78.00 %	92.83 %	93.5 %	94.00 %	94.17 %	94.17 %	93.58 %

As shown in Table 4, the accuracy of the average filter was the lowest compared to the median and Wiener filters in ten epochs. The median filter outperformed the average filter, while the Wiener filter could recognize more than 3/4 of 1 200 samples. As the number of epochs increased to 100, there was a sharp increase in accuracy for the median and average filters, which reached 84.17 % and 83.08 %, respectively. Furthermore, the accuracy of the Wiener filter reached more than 90 % by 100 epochs.

The accuracy of all three filters grew slowly after 1000 epochs. From 1000 epochs to 1 500 epochs, the accuracy of the median filter increased by 0.75 %, the average filter increased by 0.08 %, and the accuracy of the Wiener filter stayed consistent, which shows that the Wiener filter is relatively converged. Nevertheless, by 2000 epochs, the accuracy of the median filter was 91.08 %, the average filter achieved 90.08 %, and the Wiener filter achieved 93.58 %, which shows some fluctuations during training epochs as shown in Figure 6.

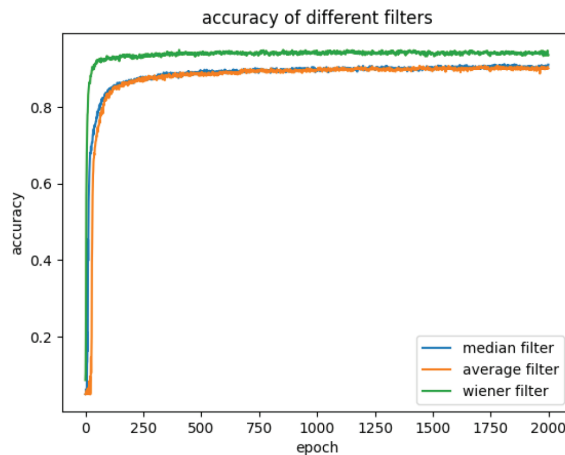


Figure 6. Curve of test accuracy of each filter with the increase of training epochs

Table 5 presents each filter's fluctuation range, total training time, and maximum and minimum results during the training in 2 000 epochs. The minimum accuracy refers to the minimum value found during the rest epochs since the epoch of the maximum accuracy. The maximum accuracy of the median filter was 91.17 %, and of the average filter was 90.75 %, while the maximum accuracy of the Wiener filter can reach 94.92 %. The fluctuation range between the maximum and minimum of all filters was within 2 %. The fluctuation range between the maximum and the accuracy of the last epoch (i.e., 2 000 epoch) also shows that all filters were relatively converged since there was no noticeable improvement in accuracy. As for the Wiener filter, it reached the highest accuracy rate of 94.92 % at the 855 epoch, which illustrates that it can converge relatively within 1 000 epochs.

Table 5. The fluctuation range of different filters

Table of Fluctuation Range of Different Filters						
	Accuracy/ Epoch			Fluctuation Range		Time (second)
	Maximum	Minimum	Final	Max-Min	Max-Final	
Median	91.17 %/1853	89.75 %/1940	91.08 %/2000	1.42 %	0.08 %	5302
Average	90.75 %/1445	88.75 %/1964	90.08 %/2000	2.00 %	0.67 %	5037
Wiener	94.94 %/855	93.25 %/1469	93.58 %/2000	1.67 %	1.33 %	1463

Therefore, based on the comparison and discussion presented above, the accuracy of the median filter was higher than the average filter. The median filter's ultimate accuracy rate was stabilized at around 91.08 %, which was 1.00 % higher than the average filter. Considering the advantage of the median filter over the average filter, which is more robust to the image noise, the median filter performs better on edge preservation (Church et al., 2008). Therefore, the recognition model may perform better on the data processed with the median filter, as the median filter may reach a slightly higher accuracy than the average filter.

Compared with the other two, the Wiener filter achieved the best performance for the recognition model. The final accuracy rate was 93.58 %, with a fluctuation range of about 1.33 % between the maximum accuracy rates. Meanwhile, the Wiener filter can converge to a higher accuracy than the other two at the fastest speed. The Wiener filter takes the least training time when compared to the other two filters. Nevertheless, the median and average filters have no significant difference in training time due to the different characteristics of filters, where Wiener and average are linear filters, whereas the median filter is non-linear.

Therefore, this research concludes that the Wiener filter is optimal. The dataset processed with the Wiener filter is used for further experiments. Furthermore, as mentioned above, the Wiener filter can converge within 1 000 epochs. Thus, the model is trained with 1 000 epochs. Table 6 illustrates the comparative study of the results of our proposed model and the research results from Zhuang et al. (2021). The results show that median and average filters achieved less accuracy than the former research. However, the final accuracy of the Wiener filter is 1.3 % higher than that of Zhuang et al. (2021).

Table 6. Comparisons with other research using filtering approach

<i>Model</i>	<i>Accuracy</i>
Median Filter (Zhuang, Liu et al., 2021)	92.28 %
Median Filter (This Research)	91.08 %
Average Filter (This Research)	90.08 %
Wiener Filter (This Research)	93.58 %

5.2. Convolutional Neural Network

Given that the Wiener filter can converge within 1 000 epochs, therefore in this experiment, 1 000 epochs were used to train and evaluate the proposed CNN model. In this experiment, 1 800 training samples and 200 test samples were used. The accuracy rate of different activation functions is given in Table 10. Figure 20 illustrates the curve of test accuracy of each activation function.

As shown in Table 7, by ten epochs, the Sigmoid function achieved the lowest accuracy, 74.50 %, while the other three were able to reach more than 80.00 %. After 100 epochs, the accuracy rate of the Sigmoid function exceeded the Tanh function. However, it was still lower than the ReLU and Mish activation functions.

Table 7. Accuracy of various activation functions according to different epochs

Epoch							
	10	100	200	400	600	800	1000
Sigmoid	74.50 %	93.00 %	93.50 %	94.50 %	94.50 %	95.00 %	95.50 %
Tanh	84.50 %	92.00 %	95.00 %	96.00 %	96.00 %	96.50 %	95.50 %
ReLU	89.00 %	96.00 %	96.00 %	95.00 %	95.00 %	95.00 %	96.00 %
Mish	86.00 %	95.50 %	96.00 %	96.50 %	96.50 %	96.50 %	97.00 %

Although the accuracy of the Mish function at 200 epochs was the same as ReLU, the accuracy of Mish was higher than ReLU in 100 epochs onwards. Mish function was able to achieve 97 % accuracy at 400 epochs. The final accuracy rate at 1000 epochs of the Sigmoid, Tanh, ReLU, and Mish functions was 95.50 %, 95.50 %, 96.00 %, and 97.00 %, respectively.

From Figure 7 it can be seen that the Mish function's test accuracy curve was almost always on top of the other three functions. The final accuracy of the Sigmoid function was able to stay at around 95.50 %, with a fluctuation range between 94.50 % and 95.50 %. The final accuracy of the Tanh function was able to stay at around 95.50 %, with a fluctuation range between 95.00 % and 96.50 %. The final accuracy of the ReLU function was able to stay at around 96.00 %, with a fluctuation range of between 95.50 % and 96.50 %. Meanwhile, the final accuracy of the Mish function was able to stay at around 97.00 %, with a fluctuation range between 96.50 % and 97.50 %.

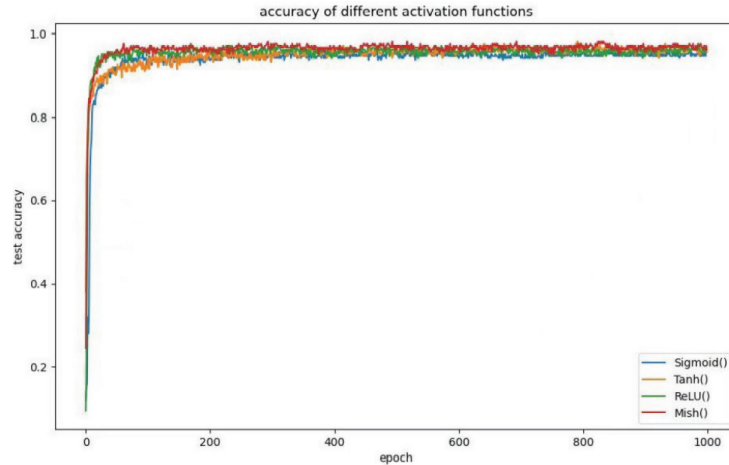


Figure 7. Curve of test accuracy of each activation function

Consequently, the Sigmoid function had the lowest recognition result, which stabilized at around 95.50 %. The accuracy of the Tanh function was able to stay at around 95.50 %; however, it performed better than the Sigmoid function, as the accuracy was 0.50 % higher than the Sigmoid function according to the fluctuation range. The accuracy of the ReLU function was able to stay at around 96.00 %, which performed better than the Sigmoid and Tanh functions. The Mish activation function performed best in this experiment, as it was approximately 1.50 %, 1.50 %, and 1.00 % higher than the Sigmoid, Tanh, and ReLU functions.

Besides, we also conduct another experiment using several datasets to compare with the research of Guo et al. (2020). In this experiment, the datasets from Set-1 to Set-5 were utilized. All the character samples in these datasets were selected at random. We trained and tested these data with 1 000 epochs and calculated the average of 5 different datasets. The average of each activation function for all datasets was shown in Table 8.

Table 8. Comparison with Guo et al. (2020) using several activation functions

CNN Model	Activation Functions	Accuracy
2ConvLayer-model (Guo et al., 2020)	Sigmoid	10 %
	Tanh	85 %
	ReLU	87 %
	Mish	90 %
Proposed model	Sigmoid	95.8 %
	Tanh	94.1 %
	ReLU	96.2 %
	Mish	96.6 %

Table 8 compares the findings from the research by Guo et al. (2020) with those from our models. From the table, all the activation functions perform better in our study, and all four functions achieve higher accuracy than those of Guo et al. (2020). Moreover, both studies achieved the highest accuracy with the Mish activation function. Mish activation function continuously differentiates; thus, it can avoid singularities and preserves a few negative information to avoid the dying ReLU phenomenon (Misra, 2019; Zhang et al., 2017).

Besides, in their research, the authors suggested that the Sigmoid function was unsuitable for classifying the image under the application of CNN. However, the model with the Sigmoid function was able to achieve competitive recognition accuracy in this study. This finding also accords with the statement of the research that different activation functions have different performances in different neural networks (Hao et al., 2020; Sun, 2021). Apart from that, the authors' research's final accuracy rate (after parameters fine turning) was stabilized at 93 %. Nevertheless, the results of all four CNN models in our experiment were still higher than those of Guo et al. (2020), which illustrates that our CNN model has a competitive performance in recognizing handwritten Chinese characters.

6. Conclusion

The median, average, and Wiener filters are methods examined in this study. The results show that the Wiener filter performs the best as its accuracy is 2.5 % greater than that of the other two filters. The model described in this study still has several constraints. Although this study compares three widely used filtering methods, additional standard, and enhanced filtering techniques can be applied. Therefore, additional filtering methods may be added for comparison in future research. It has been proposed to use the Mish activation function in the network to improve the recognition precision of our CNN model. The results show that our modified model's accuracy increased by 1.0 % and is superior to the baseline model created using the default ReLU, Tanh, and Sigmoid functions. As a result, the outcomes of the experiments carried out for this study, and the comparisons with other studies, show that the performance of the developed convolutional neural network for handwritten Chinese character recognition is competitive.

7. Funding

This research was funded by Geran Galakan Penyelidik Muda (GGPM) from Universiti Kebangsaan Malaysia (grant number: GGPM-2021-039).

References

- Aleskerova N. and A. Zhuravlev, 2020. Handwritten Chinese Characters Recognition Using Two-Stage Hierarchical Convolutional Neural Network. 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR). IEEE. <https://doi.org/10.1109/ICFHR2020.2020.00069>
- Aljojo, N. (2022). Predicting Financial Risk Associated with Bitcoin Investment by Deep Learning. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 11(1), 5-18. <https://doi.org/10.14201/adcaij.27269>

- Alom, M. Z., Sidike, P., Taha. T. M, and Asari, V. K. (2017). Handwritten bangla digit recognition using deep learning. *arXiv preprint arXiv:1705.02680*.
- Ameri, R., Alameer, A., Ferdowsi, S., Abolghasemi, V., & Nazarpour, K. (2021). Classification of handwritten Chinese numbers with convolutional neural networks. *2021 5th International Conference on Pattern Recognition and Image Analysis (IPRIA)*. IEEE. <https://doi.org/10.1109/IPRIA53572.2021.9483557>
- Amodei, D., Ananthanarayanan, S., Anubhai, R., Bai, J., Battenberg, E., Case, C., Casper, J., Catanzaro, B., Cheng, Q., & Chen, G. (2016). Deep Speech 2: End-to-end speech recognition in English and Mandarin. *International Conference on Machine Learning*, PMLR.
- Arya, M. C. and Semwal A., 2017. Comparison on Average, Median and Wiener Filter using Lung Images. *IMAGE 1*(37.42), 3.45.
- Bi, N., Chen, J., & Tan, J. (2019). The handwritten Chinese character recognition uses convolutional neural networks with the GoogLeNet. *International Journal of Pattern Recognition and Artificial Intelligence*, 33(11), 1940016. <https://doi.org/10.1142/S0218001419400160>
- Chen, L., Wang, S., Fan, W., Sun, J., & Naoi, S. (2015). Beyond human recognition: A CNN-based framework for handwritten character recognition. *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*. IEEE. <https://doi.org/10.1109/ACPR.2015.7486592>
- Church, J., Chen, C. Y., & Rice, S. V. (2008). A spatial median filter for noise removal in digital images. In *IEEE SoutheastCon 2008* (pp. 618-623). IEEE. <https://doi.org/10.1109/SECON.2008.4494367>
- Church, J. C., Chen, Y., & Rice, S. V. (2008). A spatial median filter for noise removal in digital images. *IEEE SoutheastCon 2008*. IEEE. <https://doi.org/10.1109/SECON.2008.4494367>
- Ciregan, D., Meier, U., & Schmidhuber, J. (2012). Multi-column deep neural networks for image classification. *2012 IEEE conference on computer vision and pattern recognition*. IEEE. <https://doi.org/10.1109/CVPR.2012.6248110>
- Ciresan, D. C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2011a). Flexible, high performance convolutional neural networks for image classification. *Twenty-Second International Joint Conference on Artificial Intelligence*.
- Ciresan, D. C., Meier, U., Gambardella, L. M., & Schmidhuber, J. (2011b). Convolutional neural network committees for handwritten character classification. *2011 International Conference on Document Analysis and Recognition*. IEEE. <https://doi.org/10.1109/ICDAR.2011.229>
- Dai R., Liu C. and Xiao B., 2007. Chinese character recognition: history, status and pro-spects. *Frontiers of Computer Science in China*, 1(2), 126-136. <https://doi.org/10.1007/s11704-007-0012-5>
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. IEEE.
- Dass, R., & Saini, J. (2022). Assessment of de-noising filters for brain MRI T1-weighted contrast-enhanced images. In *Emergent Converging Technologies and Biomedical Systems* (pp. 607-613). Springer. https://doi.org/10.1007/978-981-16-8774-7_50
- Daugman, J. G. (1988). Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on acoustics, speech, and signal processing* 36(7), 1169-1179. <https://doi.org/10.1109/29.1644>
- Gan, J., Wang, W., & Lu, K. (2020). Compressing the CNN architecture for in-air handwritten Chinese character recognition. *Pattern Recognition Letters*, 129, 190-197. <https://doi.org/10.1016/j.patrec.2019.11.028>

- Ge, Y., Huo, Q., & Feng, Z. D. (2002). Offline recognition of handwritten Chinese characters using Gabor features, CDHMM modeling, and MCE training. *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE. <https://doi.org/10.1109/ICASSP.2002.5743976>
- Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., & Shet, V. (2013). Multi-digit number recognition from street view imagery using deep convolutional neural networks. *arXiv preprint arXiv:1312.6082*.
- Guo, H., Ai, L., & Chen, S. (2020). Application of convolutional neural network in handwritten Chinese character recognition. In *2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)* (Vol. 1, pp. 1278-1281). IEEE. <https://doi.org/10.1109/ICIBA50161.2020.9277290>
- Hao, W., Yizhou, W., Yaqin, L., & Zhili, S. (2020). The role of activation function in CNN. *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*. IEEE. <https://doi.org/10.1109/ITCA52113.2020.00096>
- Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J. & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. *arXiv preprint arXiv:1602.07360*.
- Kavitha, B. R., & Srimathi, C. (2019). Benchmarking on offline handwritten Tamil character recognition using convolutional neural networks. *Journal of King Saud University - Computer and Information Sciences*.
- Kazubek, M. (2003). Wavelet domain image denoising by thresholding and Wiener filter-ing. *IEEE Signal Processing Letters* 10(11), 324-326. <https://doi.org/10.1109/LSP.2003.818225>
- Kumar, A. A., Lal, N., & Kumar, R. N. (2021). A comparative study of various filtering techniques. *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*. IEEE. <https://doi.org/10.1109/ICOEI51242.2021.9453068>
- Li, Y., & Li, Y. (2021). Design and implementation of handwritten Chinese character recognition method based on CNN and TensorFlow. *2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*. IEEE. <https://doi.org/10.1109/ICAICA52286.2021.9498146>
- Li, Z., Teng, N., Jin, M., & Lu, H. (2018). Building efficient CNN architecture for offline handwritten Chinese character recognition. *International Journal on Document Analysis and Recognition (IJDAR)*, 21(4), 233-240. <https://doi.org/10.1007/s10032-018-0311-4>
- Li, Z., Xiao, Y., Wu, Q., Jin, M., & Lu, H. (2020). Deep template matching for offline handwritten Chinese character recognition. *The Journal of Engineering*, 2020(4), 120-124. <https://doi.org/10.1049/joe.2019.0895>
- Liu, B., Xu, X., & Zhang, Y. (2020). Offline handwritten Chinese text recognition with convolutional neural networks. *arXiv preprint arXiv:2006.15619*.
- Liu, C. L., Yin, F., Wang, D. H., & Wang, Q. F. (2011). CASIA online and offline Chinese handwriting databases. In *2011 International Conference on Document Analysis and Recognition* (pp. 37-41). IEEE. <https://doi.org/10.1109/ICDAR.2011.17>
- Liu, C. L. (2006). High accuracy handwritten Chinese character recognition using quadratic classifiers with discriminative feature extraction. *18th International Conference on Pattern Recognition (ICPR'06)*. IEEE.
- Liu, C. L., Jaeger, S., & Nakagawa, M. (2004). Online recognition of Chinese characters: The state-of-the-art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2), 198-213. <https://doi.org/10.1109/TPAMI.2004.1262182>

- Liu, H., & Ding, X. (2005). Handwritten character recognition using gradient feature and quadratic classifier with multiple discrimination schemes. *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*. IEEE.
- Liu, X., & Di, X. (2021). TanhExp: A smooth activation function with high convergence speed for lightweight neural networks. *IET Computer Vision* 15(2), 136-150. <https://doi.org/10.1049/cvi2.12020>
- Maniatopoulos, A., & Mitianoudis, N. (2021). Learnable Leaky ReLU (LeLeLU): An alternative accuracy-optimized activation function. *Information*, 12(12), 513. <https://doi.org/10.3390/info12120513>
- Melnyk, P., You, Z., & Li, K. (2020). A high-performance CNN method for offline handwritten Chinese character recognition and visualization. *Soft Computing*, 24(11), 7977-7987. <https://doi.org/10.1007/s00500-019-04083-3>
- Min, F., Zhu, S., & Wang, Y. (2020). Offline handwritten Chinese character recognition based on improved GoogLeNet. In *Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition*. <https://doi.org/10.1145/3430199.3430202>
- Misra, D. (2019). Mish: A self regularized non-monotonic activation function. *arXiv pre-print arXiv:1908.08681*.
- Ng, P. E., & Ma, K. K. (2006). A switching median filter with boundary discriminative noise detection for extremely corrupted images. *IEEE Transactions on Image Processing*, 15(6), 1506-1516. <https://doi.org/10.1109/TIP.2005.871129>
- Patidar, P., Gupta, M., Srivastava, S., & Nagawat, A. K. (2010). Image denoising by various filters for different noise. *International Journal of Computer Applications*, 9(4), 45-50. <https://doi.org/10.5120/1370-1846>
- Sakarkar, G., Kolekar, M. K. H., Paithankar, K., Patil, G., Dutta, P., Chaturvedi, R., & Kumar, S. (2021). Advance approach for detection of DNS tunneling attack from network packets using deep learning algorithms. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 10(3), 241-266.
- Sakib, S., Ahmed, N., Kabir, A. J., & Ahmed, H. (2019). An overview of convolutional neural network: Its architecture and applications. <https://doi.org/10.20944/preprints201811.0546.v4>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, C. L. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2018.00474>
- Shetti, P. P., & Patil, A. (2017). Performance comparison of mean, median, and Wiener filter in MRI image denoising. *International Journal for Research Trends and Innovation*, 2, 371-375.
- Shi, M., Fujisawa, Y., Wakabayashi, T., & Kimura, F. (2002). Handwritten numeral recognition using gradient and curvature of gray scale image. *Pattern Recognition*, 35(10), 2051-2059. [https://doi.org/10.1016/S0031-3203\(01\)00203-5](https://doi.org/10.1016/S0031-3203(01)00203-5)
- Srinivas, R. & Panda, S. (2013). Performance analysis of various filters for image noise removal in different noise environment. *International journal of advanced computer research*, 3(4), 47.
- Sriporn, K., Tsai, C. F., Tsai, C. E., & Wang, P. (2020). Analyzing malaria disease using effective deep learning approach. *Diagnostics*, 10(10), 744. <https://doi.org/10.3390/diagnostics10100744>
- Sun, Y. (2021). The role of activation function in image classification. In *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*. IEEE. <https://doi.org/10.1109/CISCE52179.2021.9445868>



- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2016.308>
- Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2014.220>
- Wang, Y., Yang, Y., Ding, W., & Li, S. (2021). A residual-attention offline handwritten Chinese text recognition based on fully convolutional neural networks. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2021.3115606>
- Xiao, X., Jin, L., Yang, Y., Yang, W., Sun, J., & Chang, T. (2017). Building fast and compact convolutional neural networks for offline handwritten Chinese character recognition. *Pattern Recognition*, 72, 72-81. <https://doi.org/10.1016/j.patcog.2017.06.032>
- Xu, X., Yang, C., Wang, L., Zhong, J., Bao, W., & Guo, J. (2022). A sophisticated offline network developed for recognizing handwritten Chinese character efficiently. *Computers and Electrical Engineering*, 100, 107857. <https://doi.org/10.1016/j.compeleceng.2022.107857>
- Yadav, M., Purwar, R. K., & Jain, A. (2018). Design of CNN architecture for Hindi characters. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 7(3), 47-61. <https://doi.org/10.14201/ADCAIJ2018734762>
- Zhang, X. Y., Bengio, Y., & Liu, C. L. (2017). Online and offline handwritten Chinese character recognition: A comprehensive study and new benchmark. *Pattern Recognition*, 61, 348-360. <https://doi.org/10.1016/j.patcog.2016.08.005>
- Zheng, J., Ding, X., & Wu, Y. (1997). Recognizing online handwritten Chinese characters via farg matching. In *Proceedings of the Fourth International Conference on Document Analysis and Recognition*. IEEE.
- Zhong, Z., Jin, L., & Xie, Z. (2015). High performance offline handwritten Chinese character recognition using GoogLeNet and directional feature maps. In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*. IEEE. <https://doi.org/10.1109/ICDAR.2015.7333881>
- Zhu, H., Zeng, H., Liu, J., & Zhang, X. (2021). Logish: A new non-linear non-monotonic activation function for convolutional neural networks. *Neurocomputing*, 458, 490-499. <https://doi.org/10.1016/j.neucom.2021.06.067>
- Zhuang, Y., Liu, Q., Qiu, C., Wang, C., Ya, F., Sabbir, A., & Yan, J. (2021). A handwritten Chinese character recognition based on convolutional neural network and median filtering. In *Journal of Physics: Conference Series* (Vol. 1820, No. 1, pp. 012162). IOP Publishing. <https://doi.org/10.1088/1742-6596/1820/1/012162>
- Zou, J., Zhang, J., & Wang, L. (2019). Handwritten Chinese character recognition by convolutional neural network and similarity ranking. *arXiv preprint arXiv:1908.11550*.