Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2022.030303
ISSN 2790-2692 Vol. 3, Issue 3: 17-27

Print Character Recognition Method Based on Convolutional Neural Network

Yuan Wang*

School of Art and Design, Xi'an Eurasia University, Xi'an, China

Yuan_WY2021@163.com

**corresponding author*

Keywords: Convolutional Neural Networks, Printed Characters, Handwritten Characters, Character Recognition

Abstract: With the rapid development of the economy, digital transformation has become a future trend. Character recognition (CR) technology, as a key technology to realize digitalization, has always been a hot research topic that people pay attention to. CR can be widely used in the fields of transportation, finance, industry, economy and artificial intelligence. In-depth study of CR technology is of great practical significance for promoting the development of science and technology and economy. In this context, this paper studies the printed CR method based on CNN. Combining wavelet theory and CNN, this paper studies the extraction and recognition methods and implementation ways of printed characters in detail. The algorithm applied in this paper is recognized on the DSP (DM642) chip, the recognition accuracy is over 99.93%, and the recognition time is controlled within 5ms, which meets the actual requirements. Comprehensive analysis verifies the feasibility and superiority of the recognition algorithm applied in this paper.

1. Introduction

In recent years, with the continuous development of information technology, the status of CR technology in society is increasing day by day. At the same time, pattern recognition has received more and more attention from researchers. At present, printed CR technology has been widely used in our production and life, including express single number recognition, banknote number recognition, car license plate number recognition, bank bill recognition and other fields. For the research on printed CR, there are many mature algorithms, but there is no general algorithm that can solve any situation that may occur in the process of CR. Therefore, it is very important to develop an efficient CR system for numbers and letters [1-2].

In related research, Mohd et al. introduced a Quran Optical CR (OCR) system based on Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) [3]. Good word recognition rate (WRR) and CR rate (CRR) performance are achieved in the experiments. Alkhateeb introduced an efficient method based on deep learning methods to design a system for recognizing isolated handwritten Arabic characters [4]. The CNN model was developed and trained with Arabic HCs in offline mode using three Arabic handwritten CR datasets. To validate the proposed system, various experiments are performed using the AHCR, AHCD, Hijja datasets.

This paper presents a CR algorithm based on improved CNN. Since the traditional CNN will misjudge similar characters during recognition, this paper proposes an improved CNN CR algorithm after many studies and tests. First, build a first-level CNN network model, change the size of the input layer and convolution kernel based on the structure of the traditional CNN network model, improve the CR speed, and realize "rough classification" of characters; secondly, build a second-level CNN network model, in The traditional CNN network model is improved in series, and the characters that are easily confused can be re-identified to improve the accuracy of CR and achieve character "segmentation".

2. Design Research

2.1. Research Difficulties of CR

CR, especially the recognition of handwritten characters (HC), has always been a difficult point and a hot spot in the field of CR. Although handwritten CR has reached a recognition rate of more than 95% in some datasets, in practical applications, the recognition accuracy rate is often There will be a significant drop, and there is still some distance from the requirements of industrial applications [5-6]. The main factors that affect CR are:

(1) The background and noise are complex. Although the background of HCs is relatively simple, there are some special noises. For example, when writing and starting the pen, there is often a lot of jitter, and the strokes are scribbled or deformed. However, printed CR is usually in a complex external environment. For example, house number recognition usually has a complex background, which is often affected by external factors such as light and weather, and the house number is faded, stained, blocked, etc.

(2) Topology is a variable in script CR. HCs are often scribbled, such as stroke thickness, font size, deformation, hyphenation, hyphenation, missing strokes, etc., which directly affect CR.

(3) There are many types of characters and many similar characters. The standard set of Chinese characters GB2312-80 stipulates that there are 3755 types of common first-level Chinese characters, and 3008 types of second-level Chinese characters, reaching a total of 6763 types, and some Chinese characters differ by only one or two strokes. The huge number of categories and The extremely similar structure of some characters makes it difficult to recognize Chinese characters.

At present, the most widely used CR technology at home and abroad is still the traditional method based on the combination of structural features and statistical features and based on artificial neural network classification. However, due to the inherent properties of the above-mentioned characters, the traditional CR method has entered a bottleneck period. Even if the traditional recognition algorithm is improved, the CR cannot be greatly improved. People begin to explore new technologies for CR.

2.2. CNN Theory

In the traditional neural network CR process, people usually use the BP back-propagation algorithm to train the feedforward neural network structure [7-8]. Although some success can be achieved with ordinary feedforward neural networks, there are still many problems. According to the traditional pattern recognition method, the original data must be preprocessed and manually selected and extracted before classification, and it is often the most difficult to determine the best classification features of the original data. If you ignore this step and directly input each pixel of the original image as a feature point, when the input image is too large, the data to be processed is unprecedentedly huge, which will lead to dimensional disaster, which is obviously not feasible in reality [9-10]. Moreover, various characters in natural scenes, especially HCs, have relatively large changes. Traditional shallow neural networks have no inherent invariance to rotation, displacement, deformation, etc., which makes recognition more difficult.

The CNN improves the multi-level feedforward hierarchical neural network, and obtains a certain scale, displacement and deformation tolerance by means of original image input, weight sharing and pooling downsampling, which effectively avoids the above-mentioned problems. However, it has natural advantages in image classification [11-12].

2.3. CR Classification

At present, CR can be divided into print recognition and handwriting recognition [13-14]. According to the type of handwriting recognition, handwriting methods are divided into two forms: online and offline. Offline form, mainly through handwritten images collected by scanners, cameras and cameras (offline handwriting recognition); online form, handwritten dynamic information left on the interactive interface through electronic pens or touch screens, stored in a computer in a certain data format In (online handwriting recognition), character classification is shown in Figure 1 [15-16]. Handwriting input is widely used in the fields of finance and education, which greatly improves people's convenient lifestyle.

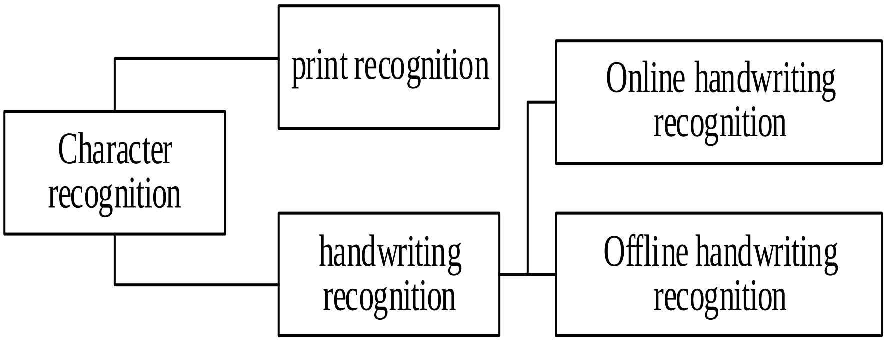

Figure 1. Classification of CR

2.4. Basic Structure of CNN

CNNs are designed to process two-dimensional data, which are invariant to scaling, rotation, displacement, etc. The artificial neural network is combined with the convolutional structure to form a CNN, in which the feature extraction process of the input data is completed in the convolutional layer, which does not require too much preprocessing, and the data dimensionality reduction is in the pooling layer. Such extraction and dimensionality reduction reduces the computational load of the network model [17-18].

(1) Convolutional layer

The convolutional layer plays a key role in the CNN and is the core of the CNN. The convolution layer applies the convolution operation. The convolution operation symbol is represented by "*". The general form of continuous function convolution is:

$$s(t)=(x*w)(t)=\int x(a)w(t-a)da \tag{1}$$

The discrete function convolution formula is expressed as:

$$s(t)=(x*w)(t)=\sum_{a=-\infty}^{+\infty} x(a)w(t-a) \tag{2}$$

Among them, x is the input, w is the convolution kernel, and s is the feature map. In the convolution operation, assuming that the image is X(i,j) and the convolution function is K(i,j), the expression of the convolution function is:

$$S(i,j)=(X*K)(i,j)=\sum_{m}\sum_{n} X(m,n)K(i-m,j-n) \tag{3}$$

(2) Character segmentation

The main purpose of character location of printed characters is to find the area where the number character is located from the image, extract the character from the entire number image, and obtain information such as the position, shape and size of the character. In the previous section, the position of the character has been roughly positioned through morphological processing. In order to determine the precise position of each character, this article uses the projection method to determine the specific position of a single character. The number image can be regarded as a "pixel matrix" of size (W*H), g(x,y) is the pixel value of a point in the pixel matrix, the horizontal projection V(i) and the vertical projection H(i) The calculation method can be expressed as:

$$V(i)=\sum_{x=0}^{W} g(x,i) \tag{4}$$

$$H(i)=\sum_{y=0}^{H} g(i,y) \tag{5}$$

From the above equations (1) and (2), the projection map of the number image in the horizontal and vertical directions can be obtained. Let the size of the location area where the character is located be (h*w).

3. Experimental Study

3.1. Image Preprocessing

The printed CR system designed in this paper takes the number image as the research object. The system performs image preprocessing, character extraction, CR and other processes on the collected number image and outputs the recognized text result, as shown in Figure 2.

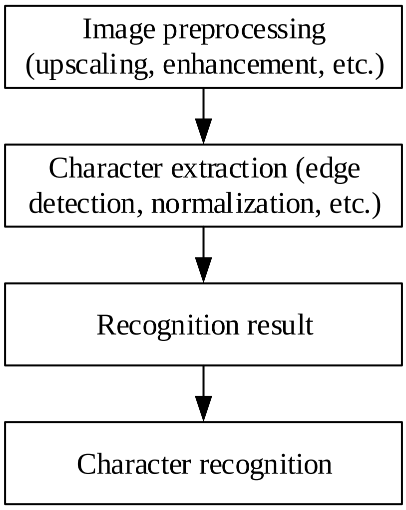

Figure 2. Steps of image processing

In practical applications, no matter whether the number image is acquired by a camera or by other scanners, some external environmental factors such as noise and contamination will inevitably be damaged during the acquisition process, which will seriously affect the accuracy of CR. Therefore, before the characters are extracted and recognized, the image needs to be preprocessed accordingly, so as to do the preparatory work for the later segmentation and recognition work. The image preprocessing operation contains rich contents. Aiming at the research object of the number image, this subject completes the preprocessing of the number image through a series of operations such as image enlargement, image skew correction, and image enhancement.

3.2. CNN Training

CNN network training mainly includes two processes: CNN forward propagation and weight update. Forward propagation refers to a sample of a certain size as input, which is transmitted through two convolutional layers and two lower pooling layers, and finally through a fully connected layer. Feature connection, the process of classification through the activation layer. Backpropagation is the inverse transmission of the error value between the model output and the label, calculating the error corresponding to each layer of the model, and updating the weights and biases in the network.

(1) Forward propagation algorithm

The propagation formula from the input layer to the first convolutional layer can be expressed as:

$$a^2 = \sigma(z^2) = \sigma(a^1 * W^2 + b^2) \tag{6}$$

where, σ: activation function, *: convolution operator, b: bias. The data output by the convolutional layer is passed to the downsampling layer. If the dimension of the feature map input to the downsampling layer is N ×N, and the size of the window dimension of the pooling layer is F ×F , the dimension of the matrix output by the sampling layer is N / F ×N /F . The formula of the fully connected layer in the forward propagation process can be expressed as:

$$a^l = \sigma(z^l) = \sigma(W^1 a^{l-1} + b^l) \tag{7}$$

(2) Back propagation algorithm

The backpropagation algorithm is calculated based on the gradient descent criterion, and its purpose is to minimize the value of the loss function. The essence of backpropagation is the process of gradient calculation. According to the gradient of each layer and the adjustment of bias and weight, the final output result is obtained.

(3) Network training process

Before the CNN network is trained, the network model must be constructed. After establishing the network model, first set the initial conditions such as the learning rate, the size of each batch of samples during training, etc., and then send the samples into the network structure for training, in the order described above, Step by step, repeat the training, and stop training when the set conditions are reached.

3.3. Network Structure Parameter Setting of CNN

The network structure parameter settings of the improved CNN are shown in Table 1. The network parameter setting of the CNN plays a decisive role in the entire CNN. During the entire model training process, this paper has conducted comparisons through many experiments, and compared the weight parameter W and bias parameter b of each layer of the model. Random initialization, in which the number of samples in batch training set in this paper is 720, which is a multiple of the number of neurons in the fully connected layer. After experimental comparison, when the number of training samples is 720, the obtained mean square error is the smallest. The test results are the best. The specific parameter settings of the CNN network are shown in Table 2.

Table 1. List of improved model parameter values

Improved CNN structure	Convolution kernel size	Step size	Feature map size	Number of feature maps
Input layer			14×14	1
Layer C1	3×3	1	12×12	6
S2 layer	2×2	2	6×6	6
Layer C3	3×3	1	4×4	6
S4 layer	2×2	2	2×2	6
Layer C5		1	1×1	24
Output layer			1×1	23

Table 2. List of improved CNN network parameters

Parameter name	Size/method
Opts.batchsize	53
Opts.maxepoch	1000
Opts.alpha	1
Cnn.lr	0.001
Optimzer	Sgd

4. Experiment Analysis

4.1. Comparative Experiment of CNN_small and CNN Model

The comparative experimental analysis of CNN_small and CNN models will be introduced here. Through the experiments of CNN_small and CNN models, a comparative analysis of the recognition rate, recall rate and F1 value of the two models is carried out, as shown in Table 3:

Table 3. Comparative analysis of CNN_small and CNN models

Model	Evaluation indicators					
	Test set (%)			Validation set (%)		
	Recognition rate	Recall	F1 value	Recognition rate	Recall	F1 value
CNN_small model	95.1	94.6	94.6	95.5	95	95
CNN model	97	96.6	96.6	97	96.6	96.6

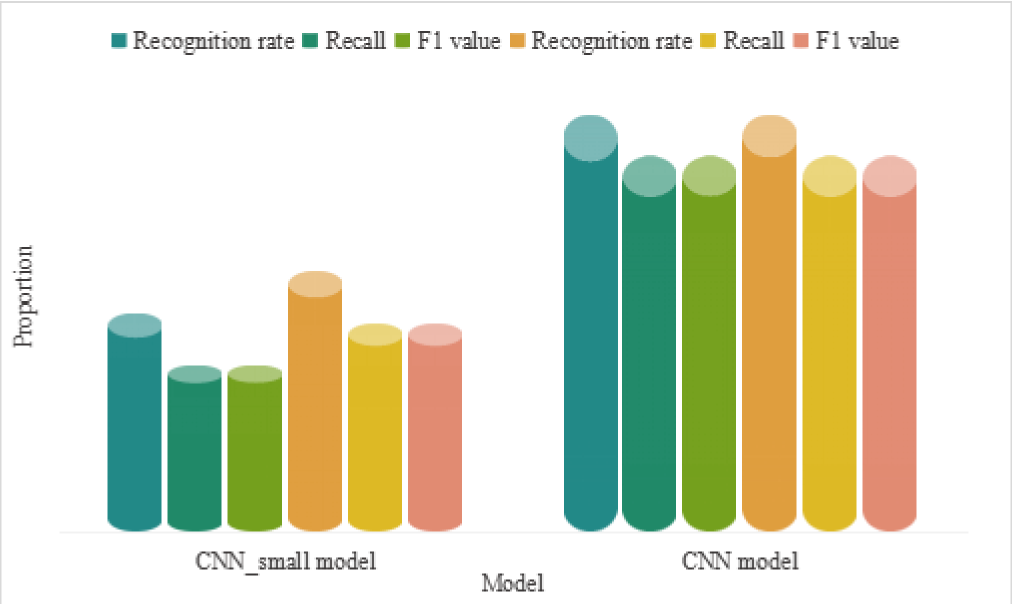

Figure 3. Comparison of CNN_small and CNN models

Figure 3 shows that the recognition rate, recall rate and F1 value of the CNN model on the test set and validation set were increased by 2%, 2%, 2%, 2%, 1% and 1% respectively through the debugging of various parameters.

4.2. Comparative Analysis Based on Different Fonts

Based on the comparative analysis of the model effects of different fonts, it is found that with the increase of the number of fonts, the recognition rate, recall rate and F1 value of the model are relatively reduced, thus showing an inverse relationship between the number of fonts and the recognition rate, recall rate and F1 value of the model. As shown in Table 4:

Table 4. Comparative analysis of different fonts

Font type	Test set (%)			Validation set (%)		
	Recognition rate	Recall	F1 value	Recognition rate	Recall	F1 value
10	100	99.8	99.8	99.6	99.1	99.3
30	99.9	99.1	99.1	99	98.8	98.7
50	98.5	98.2	98.2	98.1	98	98
80	98	97.2	97.2	98	97.8	97.9
90	97	96.6	96.6	97	96.6	96.6

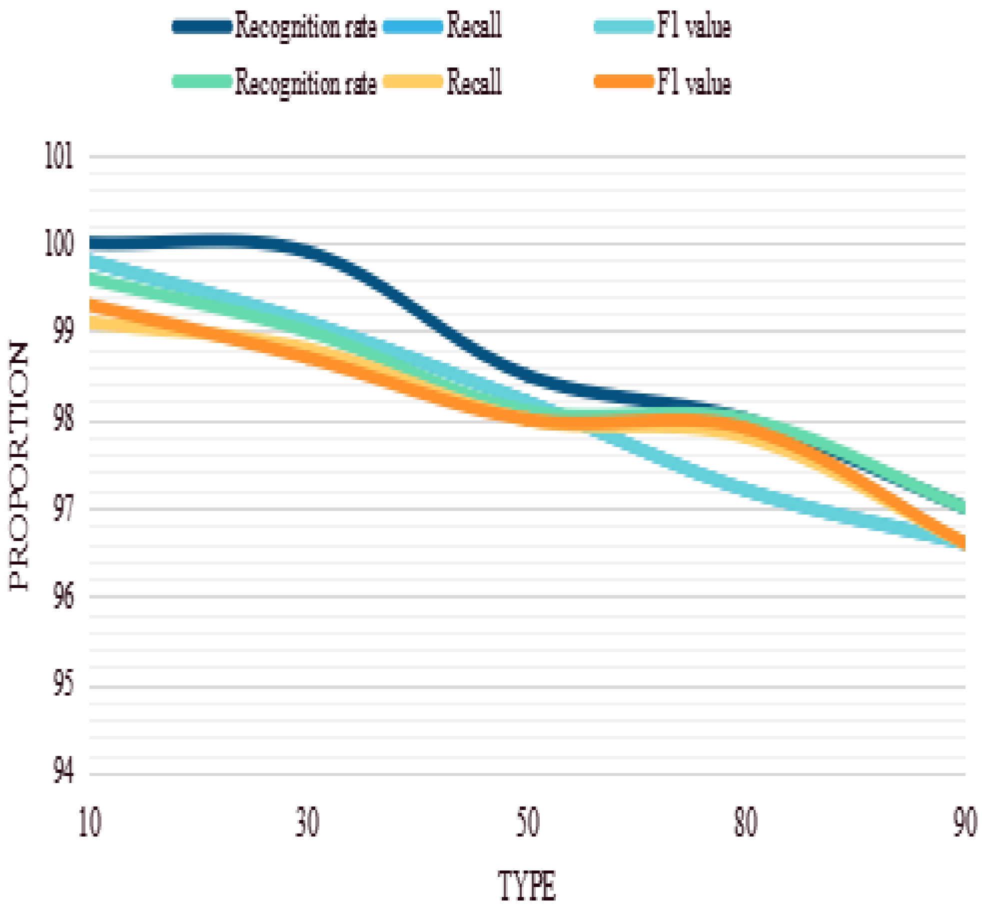

Figure 4. Comparison of different fonts

As shown in Figure 4, with the number of fonts, the recognition rate, recall rate and F1 value of the model change significantly. When the number of fonts increases, the recognition rate, recall rate and F1 value of the model decrease, and the two show an inverse relationship.

4.3. Analysis and Summary of System Test Results

In order to verify the feasibility of the algorithm applied in this paper, 1171 number images are randomly selected from the collected data set for simulation in the built C++Builder environment. Each number image has undergone complete preprocessing, character extraction, and CR. After system testing, the entire system can achieve 100% recognition accuracy for 1171 number images. In order to verify the practicability of the algorithm, the algorithm in this paper is used to identify 10,000 number images collected on the DSP (DM642) chip, and the overall accuracy rate can reach more than 99.93%. The LeNet-5 network and the first-level CNN improved by the LeNet-5 network in this paper recognize the 10,000 number image datasets taken in this paper on the DM642 chip, and the accuracy is shown in Figure 5.

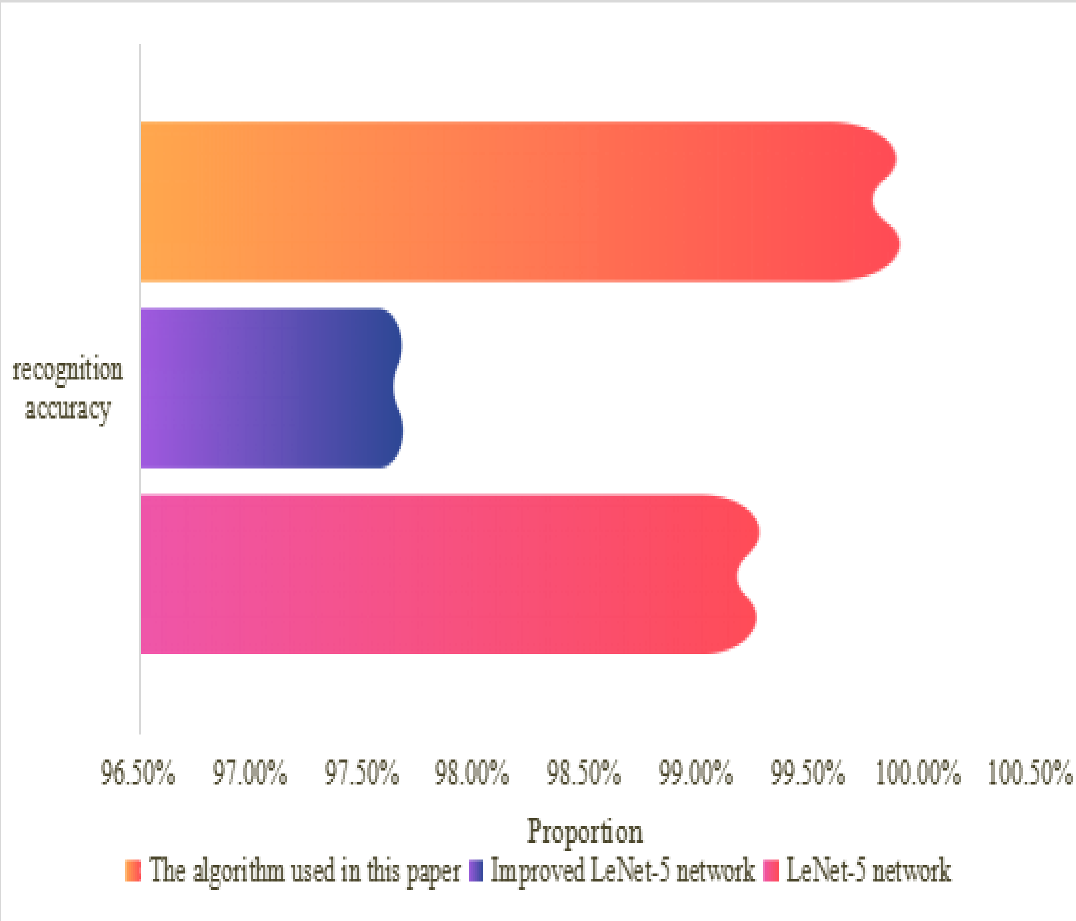

Figure 5. Comparison and analysis of recognition rates

After repeated tests, the above three methods are compared. The traditional LeNet-5 network is applied to the DM642 chip for recognition, and the recognition accuracy is 99.32%. The recognition time of a single number image is within 50ms, which cannot meet the requirements of the actual project. , so this method cannot be used; the improved LeNet-5 network is used for identification on the DM642 chip, the identification accuracy is 97.69%, the identification time of a single number image is within 5ms, and the identification time meets the project requirements, but The recognition accuracy rate does not meet the actual requirements, so this method cannot be used; and the algorithm applied in this paper is recognized on the DSP (DM642) chip, the recognition accuracy rate is over 99.93%, and the recognition time is controlled within 5ms, which meets the actual requirements. Comprehensive analysis verifies the feasibility and superiority of the recognition algorithm applied in this paper.

5. Conclusion

With the continuous development of society, although some artificial intelligence technologies such as face recognition have been rapidly improved, printed CR still plays a vital role in today's society. As a practical application problem in image classification, printed CR has a wide range of application backgrounds and application values in many fields such as finance, transportation, and education. Based on this, this paper is devoted to researching a recognition system, optimizing the existing printed CR algorithm, increasing the accuracy of CR and improving the efficiency of CR. Printed CR has been involved in all aspects of production and life, saving the consumption of working time and promoting the development of society.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Chandio A A, Asikuzzaman M, Pickering M R. Cursive CR in Natural Scene Images using a Multilevel CNN Fusion. IEEE Access, 2020, PP(99):1-1.

[2] Nayak V. Automatic number plate recognition. International Journal of Advanced Trends in Computer Science and Engineering, 2020, 9(3):3783-3787. https://doi.org/10.30534/ijatcse/2020/195932020

[3] Mohd M, Qamar F, Al-Sheikh I, et al. Quranic Optical Text Recognition Using Deep Learning Models. IEEE Access, 2021, PP(99):1-1. https://doi.org/10.1109/ACCESS.2021.3064019

[4] Alkhateeb J H, Turani A A, Abuhamdah A, et al. An Effective Deep Learning Approach for Improving Off-Line Arabic Handwritten CR. International Journal of Computer Systems & Software Engineering, 2021, 6(2):104-112.

[5] Phan T H, Tran D C, Hassan M F. Vietnamese CR based on CNN model with reduced character classes. Bulletin of Electrical Engineering and Informatics, 2021, 10(2):962-969. https://doi.org/10.11591/eei.v10i2.2810

[6] Jokic P, Emery S, Benini L. Improving Memory Utilization in CNN Accelerators. IEEE embedded systems letters, 2020, PP(99):1-1.

[7] Shimoda M, Sada Y, Kuramochi R, et al. SENTEI: Filter-Wise Pruning with Distillation towards Efficient Sparse CNN Accelerators. IEICE Transactions on Information and Systems, 2020, E103.D(12):2463-2470. https://doi.org/10.1587/transinf.2020PAP0013

[8] Zanghieri M, Benatti S, Burrello A, et al. Robust Real-Time Embedded EMG Recognition Framework Using Temporal Convolutional Networks on a Multicore IoT Processor. IEEE Transactions on Biomedical Circuits and Systems, 2020, 14(2):244-256. https://doi.org/10.1109/TBCAS.2019.2959160

[9] Konovalenko I A, Kokhan V V, Nikolaev D P. Optimal affine image normalization approach for optical CR. Computer Optics, 2021, 45(1):90-100. https://doi.org/10.18287/2412-6179-CO-759

[10] Mahesha C. Telugu Optical CR Using Cloud Computing and Python. International Journal of Computer Science Engineering and Information Technology Research, 2021, 11(1):95-106. https://doi.org/10.24247/ijcseitrjun20219

[11] Phan T H, Tran D C, Hassan M F. Vietnamese CR based on CNN model with reduced character classes. Bulletin of Electrical Engineering and Informatics, 2021, 10(2):962-969. https://doi.org/10.11591/eei.v10i2.2810

[12] Hubert, Phoenix P, Sudaryono R, et al. Classifying Promotion Images Using Optical CR and Nave Bayes Classifier. Procedia Computer Science, 2021, 179(1):498-506.

https://doi.org/10.1016/j.procs.2021.01.033

[13] Sadouk L. Handwritten Phoenician CR and its Use to Improve Recognition of Handwritten Alphabets with Lack of Annotated Data. International Journal of Advanced Trends in Computer Science and Engineering, 2020, 9(1):171-181. https://doi.org/10.30534/ijatcse/2020/26912020

[14] Khobragade R N, Koli N A, Lanjewar V T. Comparison of Feature Extraction methods for the study of Cursive Handwritten Marathi Compound CR. Journal of engineering sciences, 2020, 11(6):1070-1079.

[15] Reddy N, Ur E, Mohammed S W. A Novel Geometrical Scale and Rotation Independent Feature Extraction Technique for Multi-lingual CR. International Journal of Advanced Computer Science and Applications, 2020, 11(11):231-239. https://doi.org/10.14569/IJACSA.2020.0111130

[16] Husen M N. Automated Car Plate Detection and CR with Parking Fee Management Using Point Feature Matching. Journal of Advanced Research in Dynamical and Control Systems, 2020, 12(SP8):214-220. https://doi.org/10.5373/JARDCS/V12SP8/20202518

[17] Beltran-Hernandez J, Ruiz-Pinales J, Lopez-Rodriguez P , et al. Multi-Stroke handwriting CR based on sEMG using convolutional-recurrent neural networks.. Mathematical biosciences and engineering: MBE, 2020, 17(5):5432-5448. https://doi.org/10.3934/mbe.2020293

[18] Venugopal A, Madanan M, Venugopal A. Devanagari Cr Using Feature Set Selection Algorithm. International Journal of Psychosocial Rehabilitation, 2020, 24(8):15710-15715.