Research on License Plate Character Recognition Technology Based on Image Processing and Deep Learning

Chun Chen[1], Xiaolei Zhong[2,*]
[1]Honghe Vocational and Technical College, Honghe 661100, China
[2]Yunnan University Dianchi College, Kunming 650228, China
*Corresponding author: Xiaolei Zhong(e-mail: zhongxiaolei@hust.edu.cn)

2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA) | 978-1-6654-1606-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/EEBDA53927.2022.9744806

***Abstract*—Character recognition methods are applied in many fields, greatly improving work efficiency in daily life[1], such as license plate retrieval, invoice printing recognition, lottery betting codes, tax reports, etc. Digital recognition has been widely used in the field of computer vision and image recognition, and deep learning algorithms are currently popular image recognition algorithms. Deep learning has been widely studied and applied in target recognition and speech content recognition. With the rapid increase in production requirements and computer data processing speed, the application of character recognition in actual production and life is becoming more and more common[2]. It is also extremely important for automatic retrieval and real-time, fast and accurate character input. However, traditional pattern recognition and feature extraction algorithms cannot well meet the requirements of real-time and correctness in production. At the same time, due to the vigorous development of deep learning, character recognition technology based on deep learning has advantages that traditional recognition algorithms cannot match. This paper proposes a barcode recognition algorithm based on a deep neural network combined with a global optimization method. It uses a convolutional recurrent network to extract the characteristics of each character in the barcode and classify it. Compared with the traditional method, it has stronger adaptability and generalization. Chemical energy.**

Keywords*—Character recognition, deep learning, character recognition, recognition algorithms

I. Introduction of Related Principles

Nowadays, character recognition is basically applied to the research of printed characters and handwritten characters. For print character recognition, through the image acquisition of the character[3-4], the initial image dot matrix information is binarized, the interference noise is filtered, and the normalization is preprocessed, and then the inherent characteristics of the character itself are extracted by the fractal moment method, and it can perform unique recognition calculations to realize the automatic recognition of characters. For handwritten characters, the initial image dot matrix information is binarized, interference noise is filtered, and the normalization is preprocessed, and then the method of directly extracting the edge contour features of the characters is used to complete the character feature extraction.

This topic studies the computer recognition method of character images. In reality, data is either printed or handwritten characters. We perform pre-processing based on their respective characteristics, extract appropriate features, and perform recognition algorithms to complete the character recognition of written text to the computer. The whole system includes four parts: image acquisition, image preprocessing, feature extraction and character recognition.

II. The Process of Character Recognition

Character recognition requires many key technologies. There are image data reading, image gray value, binarization, image adjustment, discrete noise point removal, character cutting, character refinement, feature extraction in preprocessing. We only study the technologies of character recognition based on a single image, so we only introduce the image gray value, binarization, discrete noise point removal, character cutting, normalization. The process is shown in Figure 1.

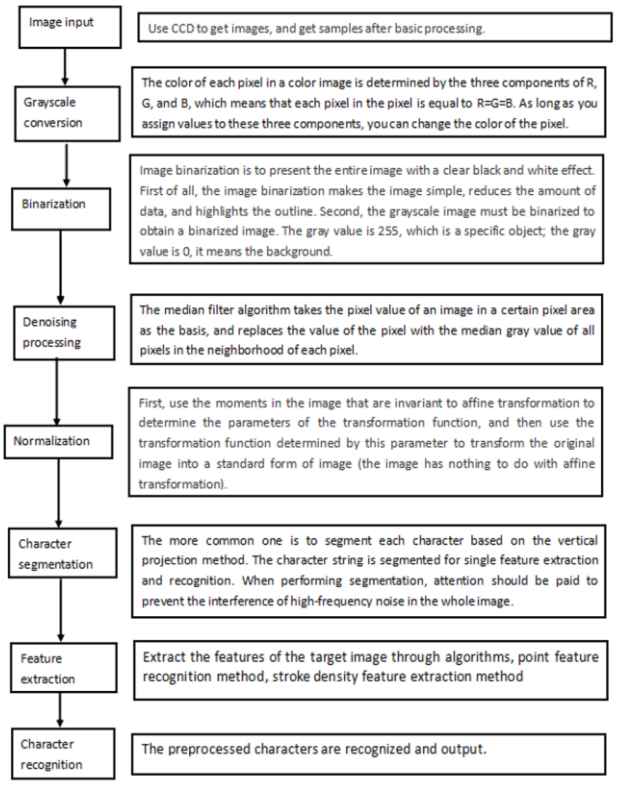

Figure 1. Flow chart of character recognition

III. Image Processing Technology

Digital image processing technology is a multi-purpose technology. It is based on computer technology and takes the visual effect of the image as the core. It performs a

series of processing on the image to get the effect as we want. The research on digital image processing technology mostly focuses on the fields of image digitization, image enhancement, image restoration, image segmentation. At first, digital image processing technology was mainly used in the newspaper industry. With the rapid development of computers, image processing technology is also flourishing. The old antiques that were produced in the 1920s also changed with the times. So far, image processing technology has become indispensable. In the financial industry, big data processing has always been an arduous and demanding task. With the help of image processing technology, data can be accurately and quickly stored in a computer for storage or calculation.

A. Image acquisition

In image acquisition, the clarity of the image must be ensured to obtain a higher-quality image. We use the CCD camera connection to sample, it not only to obtain the images with high efficiency, but also with strong promotion. The sampling of the image is not a matter of choice, but depends on the focal length and pixel resolution of the camera.

1) Pixel resolution

Pixel resolution refers to the minimum number of pixels for a measured object to display its completeness and its useful information.

Generally speaking, the number of pixels in an image and the quality of the image are directly related to the image sampling interval. The smaller the interval, the better the image quality. The more pixels, the larger the relative image space array.

2) Lens focal length

The choice of focal length (F) is determined by three parameters: the size of the CCD sensor, the distance from the lens to the object being captured, and the imaging size required by the focal length of the lens, as shown in formula (1).

$$F = V \times {}^{D}\!/_{V} \tag{1}$$

For example, a 1/4CCD sensor v (vertical) = 2.70mm, h (horizontal) = 3.6mm, object distance) = 300mm, assuming the average height of the ticket V = 150mm, as shown in equation (2)

$$F = V \times \left({}^{D}\!/_{V}\right) = 2.70 \times \left({}^{300}\!/_{150}\right) = 5.4ram \tag{2}$$

That is, you can choose a lens with a focal length less than 5.4.

B. Image preprocessing

1) Grayscale transformation

Grayscale transformation is also called image point operation (only for a certain pixel of the image) to change the grayscale value point by point. It is the simplest technology among all image processing technologies. Its transformation form is as follows:

s=T(r) s=T(r) where T is the gray scale transformation function; r is the gray scale before transformation; s is the pixel after transformation.The relationship between the gray scale changes is shown in Figure 2.

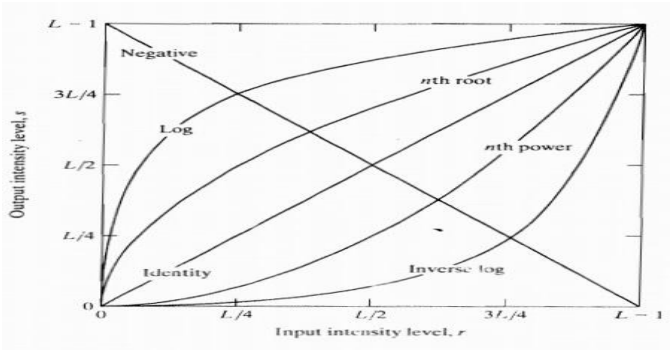

Figure 2. Relationship diagram of gray change

The image gray scale transformation has the following functions:

(a) It can improve the quality of the image and optimize the grayscale contrast of the image.

(b) It can highlight image features under gray contrast and show more image details.

(c) It can effectively change the histogram distribution of the image to make the pixel distribution more uniform.

The figure above shows the graphs of several common gray-scale transformation functions. According to the graph of this function, the effects of these kinds of gray-scale transformations can be known. For example, logarithmic transformation and power-law transformation can achieve image gray scale expansion or compression. There are three main functions used for image gray scale transformation:

Linear function, logarithmic function (logarithmic and antilog transformation), power law function. This time we study the algorithm of linear function to change the gray value. For linear transformation, we can use this formula

$$S = a \cdot r + b \tag{3}$$

Among them, a is the slope of the straight line, and b is the intercept on the y-axis. Let r is the gray level before transformation, and s is the gray level after transformation.

In the case of under-exposure or over-exposure, the image grayscale may display abnormally, blurry or single grayscale. It will effectively improve the visual effect of the image by using a linear single-valued function which can linearly expand each pixel in the image

2) Dichotomies

Image binarization is a basic technology in image processing. The image must be preprocessed before the image binarization operation[5-6], including image grayscale and image restoration. It needs to refer the histogram when thresholding, and the thresholding method is divided into global thresholding method and local thresholding method. When thresholding method is processed, it cannot be done arbitrarily,nor can it be converted into a binary image by randomly obtaining a threshold. However, an appropriate threshold must be selected, otherwise a lot of information will be lost after processing and an incomplete image will be obtained.

C. Segmentation method based on deep learning neural network

In recent years, neural network recognition technology based on deep learning has received extensive attention and has been used for image segmentation. The basic idea of the neural network segmentation method is to train the multilayer perceptron to obtain a linear decision function, and then use the decision function to classify the pixels to achieve the purpose of segmentation. This method requires a lot of training data. Neural network has a large number of connections, it is easy to introduce spatial information, and can solve the problem of noise and unevenness in the image.

1) Focal loss

Focal loss is an improvement of the standard cross-entropy evaluation function. Focal loss is achieved through deformation, so the weight loss allocated to well-classified samples is lower. Ultimately, this ensures that there is no classification imbalance. In this loss function, as the confidence of the correct category increases, the cross-entropy loss gradually decays to zero as the scale factor increases. In the training process, the scale factor will automatically reduce the contribution of simple samples and pay more attention to complex samples. The formula is as follows:

$$FL(p_t) = -(1-p_t)^{\gamma}\log(p_t) \tag{4}$$

2) Dice loss

Dice coefficient is a set similarity measurement function, usually used to calculate the similarity of two samples (the value range is [0, 1]), the formula is as follows:

$$DSC = \frac{2|X \cap Y|}{|X|+|Y|} \tag{5}$$

3) Boundary loss

A variant of boundary loss is used for highly imbalanced segmentation tasks. The form of this loss is a spatial contour, not an area distance measure. This method solves the problem of area loss in highly unbalanced tasks. The formula is as follows:

$$Dist(\partial G, \partial S) = \int_{\partial G} \left\| y\partial s(p) - p \right\|^2 dp \tag{6}$$

4) Weighted cross entropy

In a variant of cross entropy, in the case of a type of imbalance, all positive samples are weighted with a specific coefficient. The formula is as follows:

$$WCE(p, \hat{p}) = -(\beta p \log(\hat{p}) + (1-p)\log(1-\hat{p})) \tag{7}$$

5) IoU-balance loss

IoU-balanced classification loss aims to increase the gradient of high IoU samples while reducing the gradient of low IoU samples. In this way, the positioning accuracy of the machine learning model is improved.

$$IoU = TP/(TP+FP+FN) \tag{8}$$

6) Lovász-Softmax loss

This loss is based on the sub-module loss of the convex Lovasz extension, which directly optimizes the average loU of the neural network.

$$\text{loss}(f) = \frac{1}{|c|}\sum_{c\in C}\overline{\Delta J_c}(m(c)) \tag{9}$$

7) Other losses

(a)Hausdorff distance (HD) loss: use convolutional neural network to estimate Hausdorff distance.

(b)Distance penalty CE loss: It is used to guide the network in areas that are difficult to separate.

(c)TopK loss: used to ensure that the network pays attention to complex samples during the training process.

(d)Sensitivity-specificity (SS) loss: Calculate the weighted sum of the mean square deviation of specificity and sensitivity.

D. Normalization

The constant distance of the image is used to eliminate the influence of other functions on the image transformation, so as to obtain a high-quality image, which is close to or higher than the standard image.

We use the constant distance of the image to find the parameters, put the function values of the image in one stage, center the coordinates, use the secondary parameters to obtain the transformation function, and use the function to calculate the standard form of the transformation image.

Normalization is a way to simplify calculations. It uses the invariant moments of the image to find a set of parameters so that it can eliminate the influence of other transformation functions on the image transformation. The purpose is to:

(1) Ensure that the input function value can be used normally, so as to avoid unused amount, which may cause image damage.

(2) Ensure that there are small values in the output data, so as not to cause image quality changes.

(3) Normalization can prevent neuron output saturation caused by excessive net input absolute value.

E. Character segmentation

The character segmentation process is shown in Figure

3.

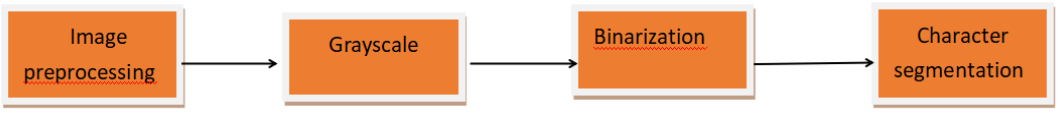

Figure 3. Character segmentation flowchart

It is an operation to divide an image displayed in a character string into individual characters. It separates characters, letters, numbers, etc. from the image so that the characters can be correctly recognized.

(1) Improved connected domain segmentation method

Layout segmentation is a common way of character segmentation. It uses a connected domain to perform a specific algorithm for connected regions of an image. After the image layout is corrected, the image layout is blurred and then the layout is analyzed. This method is effective, low-cost, and it is universally applicable to various occasions. Cutting the characters is a good way to get a clear and recognizable image.

(2) Character segmentation algorithm based on vertical projection

In order to effectively and correctly identify the characters, the vertical projection character segmentation algorithm can be used to cut the binarized image before the horizontal direction, it eliminates the interference of information such as the border, and then perform the vertical segmentation.

IV. Experimental Results and Analysis

The most common character recognition is the license plate recognition. We randomly selected a license plate and used MATLAB to preprocess and recognize the image. After simple pre-processing such as image cropping and binarization, the characters are cut to eliminate the interference of edge information such as borders for easy recognition. The character image after cutting becomes a single character, which lays the foundation for recognition. It can be found that in the rendering of character cutting, the image pixels are different, so we need to go through normalization to solve this problem.

After the normalized image, the problem of different sizes and different pixels of a single character image is solved, so that the accuracy and reliability of the subsequent character recognition are guaranteed.

Finally, we conclude that the recognition rate of the characters after these processing is greatly improved. The following figure shows the character similarity comparison chart, and it can be seen that the character recognition rate is quite high.

In the course of the experiment, we finally performed the effect simulation after preprocessing the handwritten characters, and the preliminary results are shown in Figure 4 below:

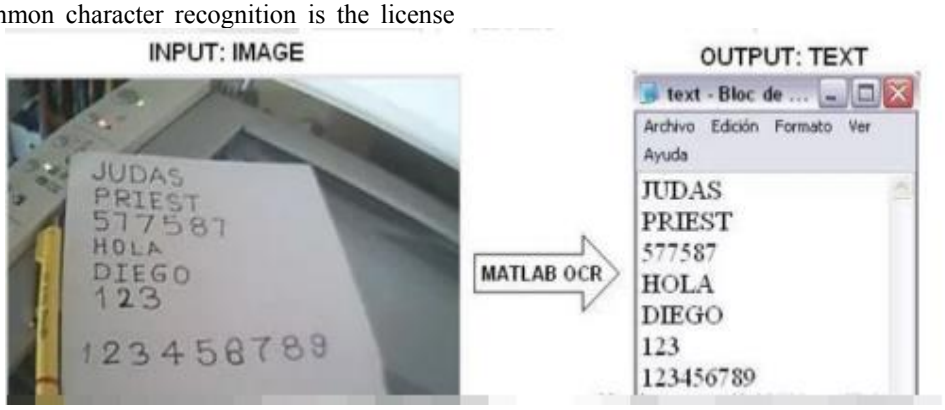

Figure 4. Character recognition effect diagram

It can be found that the simulation results are effective. The character recognition rate is guaranteed, and the accuracy rate is remarkable.

In the test of printed character recognition, we mainly use images, license plates and text as the research objects, and black as the background of the image. It is concluded that the recognition rate of printed characters is stable above 95.15%, which can accurately obtain the characters from the written version to the computer; while the recognition rate of handwritten characters is lower than that of the printed characters, there may be a problem of recognition errors, but it can also be successful 91.15% of the characters were recognized. In general, the recognition rate of the character recognition system reaches 96.25%, which is an ideal data and achieves the expected effect. The effect diagram is shown in Figure 5 below.

Recognition rate of all characters			
Character	Recognition rate %	Character	Recognition rate
1	98.1	9	96.8
2	97.3	6	95.4
3	95.1	7	97.9
4	90.3	8	99.1
Average recognition rate	96.25		
Recognition rate of printed characters			
A	95.3	E	91.3
B	94.6	F	95.2
C	95.5	9	97.3
D	96.6	6	96.5
WEN	94.3	YU	90.2
YUN	97.4	CHUAN	98.1
Average recognition rate	95.15		
Handwritten character recognition rate			
A	92.3	E	89.9
B	91.3	F	92.4
C	93.1	9	95.1
D	95.3	6	91.5
WEN	89.1	YU	87.6
YUN	88.2	CHUAN	88
Average recognition rate	91.15		

Figure 5. Experimental data graph of character recognition rate

It can be concluded from the above experimental data that the recognition rate of printed text is generally higher than that of handwritten text. The printed body can also complete the task with high recognition rate by the template matching method after pre-processing. However, the handwritten characters can not achieve the expected effect for the template matching method, and the point feature and structure feature extraction by the neural network has obvious effects on the basis of the genetic algorithm. However, in text recognition, there are also some things that cannot be recognized or are recognized incorrectly. The main reason is that the details of the recognition process are not perfect, and more experimental research is needed.

V. Conclusion

This research takes printed characters and handwritten characters as sample objects, and uses deep learning-based template matching methods and genetic algorithms for recognition. Most of this achievement can be achieved by machines, instead of manually storing data, which greatly improves the accuracy and utilization rate, and makes the machineization of data flourish. This article mainly starts research on character recognition technology based on deep learning. In response to the current demand for printed character recognition in production, relying on the anti-theft door number character recognition project, the focus is on character recognition technology under complex backgrounds, which can be divided into image preprocessing, character target positioning and character recognition. Content recognition based on deep learning network framework.

Acknowledgements

This work was financially supported by the first batch of the Ministry of Education Industry-University Cooperation and Education Project in 2018, China (Project Number: 201801037032).

References

[1] Wang Hongzhu, Cui Yongjie. Research on Tomato Color Detection Based on Computer Vision. Agricultural Technology and Equipment, 2017(1): 49-51

[2] Zhao Yiqun, Liu Fu, Kang Bing. Vehicle identification method based on license plate detection. Journal of Jilin University, 2019, 37(2): 169-173

[3] Yang Dingding, Chen Shiqiang, Liu Jingyi. License plate location algorithm based on license plate background and character color characteristics. Computer Applications and Software, 2018, 35(12): 217-221

[4] Cui Bin, Liu Chen, Li Ming. Research on the Application of Digital Image Processing in License Plate Recognition. Technological Innovation and Application, 2020(32):166-167

[5] Deng Yunsheng, Zheng Chenxia, Yin An. Comparative study and implementation of license plate positioning and character segmentation methods. Journal of Lanzhou University of Arts and Sciences (Natural Science Edition), 2019, 33(6): 78-83

[6] Yao Wenfeng, Zhen Tong, Lv Zongwang, et al. Research on segmentation and recognition technology of license plate characters. Modern Electronic Technology, 2020, 43(19): 65-6