A Multi-table Image Recognition System Based on Deep Learning and Edge Detection

Dong Xiao, Haiyang Sun, Zhonglin Bao, Peng Yang and Wenrui Zhang

May 1, 2019

A Multi-table Image Recognition System Based on Deep Learning and Edge Detection

Dong Xiao[1,2] , Haiyang Sun[1], Zhonglin Bao[1], Yang Peng[1], Wenrui Zhang[1]

Northeastern University , Shenyang 110819
E-mail: xiaodong@ise.neu.edu.cn

Abstract: In order to recognize the information in multi-table images, we use deep learning and edge detection and pr opose a multi-table image recognition system based on deep learning and edge detection. After the image is preprocess ed to the standard image, the table area in the image sample is located by Hough transform, morphological function an d edge detection technology, and the cells in the table area are extracted from the image sample. The projection analysi s method separates and extracts a single character in the cell, and finally uses a recognition model based on the deep le arning convolution neural network to recognize each single character image. After experiments, we find that the syste m has the advantages of fast and accuracy, which brings great convenience and benefit to the companies engaged in inf ormation field.

Key Words: Deep Learning, Edge Detection, Table Image, Hough Transform, Morphological Function, Projection Analysis

1 INTRODUCTION

The multi-table image recognition system involves image preprocessing, segmentation of table images, segmentation of strings, and recognition of characters.

1.1 Image Preprocessing

Image preprocessing includes grayscale processing, binarization, denoising processing, and tilt correction of the image. Traditional linear mixed algorithms of gray processing includes component method, maximum method, average method and weighted average method. Another type of gray processing method is a nonlinear conversion algorithm, which mainly includes logarithmic transformation, power transformation, exponential transformation, and piecewise function transformation. For example, Rasche Karl and Geist Robert proposed the multi-dimensional scaling matrix (MDS) algorithm in 2005[1]. Later Song Liming proposed a color-to-gray conversion algorithm using Gaussian mixture model. [2]

The binarization methods of images can be roughly divided into full threshold algorithms, local threshold algorithms and other algorithms. The full threshold algorithm is commonly found by the histogram method and the largest inter-class variance method proposed by Otsu. According to the method of edge detection in local information The local threshold algorithm is divided into local image function based method, image filtering, multi-scale method, reaction-diffusion equation based method, boundary curve fitting method, active contour method, etc. In addition to the two threshold algorithms, there are Markov random field based methods, regional growth methods, waterline algorithms, and minimum description length methods. In recent years, a large number of improved algorithms have emerged, such as a binarization method based on adaptive crack width proposed by Yaming Xu and Peng Tian [3]. Di Lu, Xin Huang proposed a binarization algorithm based on regional contrast enhancement [4]. In this study, the adaptive threshold algorithm is used to implement the binarization processing of project images.

Common image denoising algorithms include spatial domain filtering, transform domain filtering, partial differential equations, variational methods, and morphological noise filters. The Common spatial image denoising algorithms include neighborhood averaging, median filtering, low-pass filtering, etc. Image denoising methods include Fourier transform, Walsh-Hadmard transform, cosine transform, K-L transform, and wavelet transform. The partial differential equations have the equations of Perona and Malik. Rudin Osher and Fatemi proposed the variational method which is widely used in the total variation TV model. Wavelet shrinkage methods include threshold shrinkage and proportional shrinkage. In recent years, Hongyan Li proposed a new wavelet threshold denoising algorithm based on the classical wavelet threshold denoising algorithm [5] and Yongcai Liu and Yidong Bao proposed the bilateral filtering images combined with total variation in noise method research in 2017[6].In our system, we use the spatial domain-based median filtering method to achieve denoising processing.

Image tilt correction includes tilt detection and correction. The methods of tilt detection mainly include projection map based methods, Hough transform methods, the nearest neighbour clustering methods, vectorization methods, cross correlation methods based on Fourier transforms. The correction includes horizontal correction and vertical correction. The horizontal correction method is divided into three types: straight line detection method (mainly Hought transform method and rotary projection method), characteristic straight line method and direction image

This work is supported by National Nature Science Foundation under Grant National Natural Science Foundation of China (Grant Nos. 61773105,41371437, 61203214). Fundamental Research Funds for the Central Universities (Grant Nos. N160404008).

method: straight line detection method. The methods of vertical correction mainly include straight line fitting method and rotary projection method. In recent years, Lei Feng and Yingnan Geng proposed a fast correction of oblique document images based on hough transform in 2017 [7], and Zhiqing Guo proposed a biochip image tilt correction algorithm based on roundness discrimination[8]. We uses the tilt correction method based on Hough transform in our system.

1.2 Table Image Segmentation

There are six image segmentation methods widely used in the field of image segmentation: threshold method, region growing method, edge detection method, artificial neural network method, variable model method and fuzzy set theory based methods. Being applied to segment conventional images, where table images exist with multiple lines intersecting, these method form two new algorithms: the method of Maximum Attributive Zone (MAZ) [9] and the global threshold segmentation algorithm [10]. Based on the background estimation and compensation of the form document image, Rencan Nie proposed a global threshold segmentation algorithm for visual table images in 2016 [11]. We use the Hough transform combined with the etch expansion and edge detection techniques in the morphological function to realize the segmentation of multi-table images in our system.

1.3 String Segmentation

Traditional character segmentation algorithms can be divided into algorithms that directly segment characters and adaptive segmentation clustering algorithms based on classifiers. The most basic method which is based on the vertical segmentation, can be divided into three categories: template matching method, vertical projection method and vertical segmentation method. On this basis, four improved methods are generated: contour feature & vertical projection method, CC & vertical projection method, template matching & vertical projection adaptive method and a method based on the combination of the inherent features of the license plate characters and projection. In addition to the methods of segmentation and improvement, there are several algorithms: horizontal projection based methods, cluster analysis character segmentation methods, Renyi entropy & mathematical morphology edge detection, a method based on associated and complementary color features, a method based on differential projection and excellent segmented characters and a method for accurately segmenting characters using feedback. For handwritten English character segmentation, Rui Ma et al. proposed a handwritten English string segmentation based on post-recognition processing [12]. To solve the adhesion of English strings, Jiayu Luo proposed a fast segmentation algorithm for confined English strings [13]. To segment Uyghur and Chinese characters, Abdusalam Dawut proposed an image segmentation method based on mixed Gaussian model fitting threshold and region growing [14].

1.4 Character Recognition

The traditional handwritten Chinese character recognition system mainly includes three parts: data preprocessing, feature extraction and classification recognition. Feature extraction can be divided into structural features and statistical features. The best features are currently statistical features. The most commonly used models of classifiers include improved quadratic decision function, support vector machine, hidden Markov model, and differential learning, decision function and learning vector quantization, etc. For text line recognition, there are mainly two recognition methods based on the segmentation strategy and the non-segmentation strategy. The method of projection, connected domain analysis and other methods are used to segment the text lines.

Deep learning, as new development of neural network models, was formed in the 1940s. From the original MP model to the new deep neural network model deep belief network(DBN) and its training methods. And in recent years, the new algorithms such as logistic regression, neural network, convolutional neural network and KNN (K-nearest neighbor) algorithm have been formed. With the continuous development of deep learning, depth models such as deep convolutional neural network (CNN), deep belief network (DBN), cascaded automatic coding machine (SAE), and deep recurrent neural network (RNN) are gradually applied to the field of handwritten text recognition. And achieved a lot of breakthrough development. The most widely studied is the deep convolutional neural network. By using the backpropagation algorithm (BP) for the convolution structure, LeCun proposed a convolutional neural network model with local connections between layers and a multi-layer structure. Krizhevsky, who was innovating the reserch on CNN and using a series of new training techniques such as ReLU nonlinear activation function and Dropout have promoted the development of CNN. Aiming at the characteristics of handwritten Chinese character recognition, many improved CNN training methods are also proposed, such as CNN-based end-to-end identification method and a multi-column CNN model (MCDNN) proposed by IDSIA Lab. In recent years, there have been many other deep learning models. One of the models worthy of attention is Deep Reinforcement Learning (DRN). For example, the GoogleDeep Mind team proposed the deep Q network in 2015 [15]. Qiao Junfei proposed an adaptive depth Q learning strategy for handwritten digit recognition in 2018[16]. This paper used Tensorflow deep learning based convolutional neural network for character recognition.

2 MODEL ESTABLISHMENT

2.1 Image Preprocessing

The image preprocessing includes four steps: gray level processing, binarization, noise removal processing and tilt correction to obtain the standard image samples. Grayscale processing is to establish the correspondence between brightness Y and R, G and B color in YUV color space according to the change relationship between RGB and YUV color space：

$$Y = 0.3R + 0.59G + 0.11B \quad (1)$$

We can use this brightness value which represents the grayscale value of the image to obtain the ideal grayscale image.

Image binarization is to convert the gray value of pixels on the image to 0 or 255 by setting a threshold,Adaptive threshold algorithm realizes binarization by comparing the value I of input pixel with a value C and determining the output value according to the comparison result. Each pixel of adaptive binarization has a different comparison value C, which is calculated by subtracting the delta difference from a block range centered on that pixel. There are two common methods of computing C: Average minus delta(Boxfilter is used to filter all pixels with the same weight around them), Gaussian weighted sum minus delta(GaussionBlur is used, the weight of pixels around (x, y) is obtained by gaussian equation according to the distance from the center point).

The median filtering method based on spatial domain is used to realize de-noising. Firstly, the probability density function of pepper and salt noise should be defined:

$$P(z) = \begin{cases} P_a & (z = a) \\ P_b & (z = b) \\ 1 - P_a P_b & (others) \end{cases} \quad (2)$$

For the input image of size M*N, it is assumed that the probability of each pixel turning into pepper noise or salt powder noise is Pa and Pb, respectively. Suppose both a and b are saturated values, which are equal to the maximum and minimum allowed in the digitized image, to an 8-bit image, this means a=255 (white), a=0 (black). In the actual processing, a matrix corresponding to the input image of size is generated, and each element in the matrix is a random number between 0 and 1, so that an input image corresponding to the elements in the two intervals can be specified. The elements become (white) and (black). In the median filter, half of a set of values is less than or equal to K, and half of the set is greater than or equal to K. This is the median K of an image. To perform median filtering on a pixel, first, you must sort the pixel points in the mask and the pixel values in the field to determine the median value, and assign the median value to the pixel. When all the pixel values in a field are the same, any one of them can be used as the median value.

For the tilt correction based on Hough transform, a straight line can be represented by an equation: y= k*x +b in a Cartesian coordinate system. The main idea of Hough transform is to use the correspondence between the parameter space and the variable space in this equation. In the variable space, (x, y) is known as the quantity; in the parameter space, (k, b) is used as the variable coordinate, or vice versa. The point where the line y=k_1×x+b_1 is mapped in the parameter space is (k_1,b_1) , and the point where the parameter space line y_1=k×x_1+b is mapped in the variable space is (x_1,y_1). The over variable (x, y) has an infinite number of lines corresponding to the parameter space point. Therefore, it can be inferred that in the variable space, the number of points which are distributed on the same line is N, and each point uniformly extracts M lines according to the inclination angle in the parameter space, then the M*N lines have an overlap number in the parameter space. The point at which M is reached, the angle of inclination of the line in the variable space corresponding to the point is the angle of inclination of the character image. Select polar coordinates:

$\rho = \cos(\theta) \times x + \sin(\theta) \times y$, ρ is the distance from the line to the origin, and θ is the angle between the line of the origin and the X axis.

$$\rho = \cos(\theta) \times x + \sin(\theta) \times y = A\sin(\alpha + \theta) \quad (3)$$

The character in the table has obvious boxes or obvious vertical and horizontal lines. When the Hough transform is performed, these lines will be extracted and finally get the wrong tilt angle. We control the extracted line by threshold and solve the wrong angle. The specific method is to display the tilt angle of the extracted straight line, select the threshold range according to the tilt angle of the specific image, remove the tilt angle outside the threshold range, and use the least squares algorithm to obtain the appropriate tilt angle from the tilt angle within the threshold. Let the angle of inclination left be θ_1， θ_2 ....θ_n, the most suitable angle is θ; $(\theta - \theta_1)^2 + (\theta - \theta_2)^2 + \cdots + (\theta - \theta_n)^2$ is the smallest, and the most suitable angle is $\theta = \frac{\theta_1 + \theta_2 + \cdots + \theta_n}{n}$ (4) calculated by least squares method.

2.2 Table Positioning and Cell Extraction

Line detection: Extract horizontal and vertical lines through Hough changes and morphological operations. Use tilt correction and Hough transform to get a horizontal vertical table. First, the edge of the image is detected, and then each linear region on the image is described by a two-dimensional vector, and the linear region counter on the image is mapped to the storage unit in the parameter space. ρ is the distance from the straight line area to the origin, so for an image with a diagonal length n, the value of ρ ranges from (0, n), and the value of θ ranges from (0, 360). For all pixels (x, y), we find ρ at each angle of θ, thereby accumulating the number of occurrences of ρ. The ρ above a certain threshold is a straight line.

Erosion and Dilation: The two most basic morphological operations of morphology are Dilation and Erosion. Dilation adds pixels to the boundaries of objects in the image, and Erosion is just the opposite. The amount of pixels added or deleted depends on the size and shape of the structural elements used to process the image, respectively. Dilation: The value of the output pixel is the maximum value of all pixels in the size and shape of the structural element, that is, the highlighted part of the image is expanded, and the effect picture has a larger highlighted area than the original image. In a binary image, if any one of the pixels of the input image within the range of the kernel is set to a value of 1, the corresponding pixel of the output image will be set to 1. The latter applies to any type of image (such as grayscale, BGR, etc.).The expansion mathematical expression is as follows：

$$dst(x,y) = \max\ scr(x+x',y+y')((x',y') \neq 0(5)$$

Erosion: The application of the etching operation is also the same. The value of the output pixel is the minimum value of all the pixels in the size and shape of the structural element, that is, the highlighted part in the original image is corroded, and the highlight area of the effect picture is smaller than the original picture. The Erosion mathematical expression is as follows:

$$dst(x,y) = \min\ scr(x+x',y+y')((x',y') \neq 0) \quad (6)$$

Table positioning and extraction: Use the bitwise_and function to cross horizontal and vertical lines to locate the intersection. This function is an 'AND' operation on binary data.

Then use the findcontour function to find the contour, and judges whether it is a table according to the shape and size of the contour. Specifically, the area of the obtained region is ignored if it is smaller than a certain value, indicating that the miscellaneous line is not a table. The topological analysis of the digital binary image is needed. The algorithm determines the surrounding relationship of the boundary of the binary image, that is, determines the outer boundary, the hole boundary and their hierarchical relationship. These boundaries have a one-to-one correspondence with the original image. The boundary is used to represent the original image. The input binary image is an image of 0 and 1, and the pixel value of the image is represented by f i, j). Each line scan is terminated by two cases: f(i, j-1)=0, f(i, j)=1 (f(i, j) is the starting point of the outer boundary); f(i , j)>=1, f(i,j+1)=0 (f(i,j) is the starting point of the hole boundary). Then mark the pixels on the border starting from the starting point. Here a unique identifier is assigned to the newly discovered boundary, called NBD. Initially NBD=1, each time a new boundary is found plus one. In this process, when f (p, q) =1, f (p, q+1) =0, f (p, q) is set to -NBD.

After confirming the shape and size of the table and then approximating the area into a shape, use the approxPolyDP function to approximate the specified point set. The function is implemented using the Douglas-Peucker algorithm to give a curve consisting of line segments to find similar curves with fewer points. The starting curve is an ordered set of points or lines, distance dimension $\varepsilon > 0$. Initially give all points between the first and last points, automatically marking the first and last points to be saved. Then find the farthest point from the line segment consisting of the first point and the last point as the end point; this point is obviously the farthest on the curve of the approximate line segment between the end points. If the point is closer to ε than the line segment, then any points that are not currently marked will be saved, while curves that are not simplified or worse than ε can be discarded. If the farthest point distance of the offline segment is greater than ε, the point must be retained. The algorithm recursively calls itself at the first point and the farthest point, and then recursively calls itself at the farthest point and the last point. When the recursion is complete, a new output curve can be generated that includes all the points marked as reserved.

Finally, use the boundingRect function to convert this area into a rectangle containing the shape of the input. Use this

function to include the specified set of points, so that a most suitable forward rectangle is formed to frame the currently specified set of points, and the minimum rectangle of the vertical boundary of the outline is calculated, and the rectangle is parallel to the upper and lower boundaries of the image.

Cell segmentation and extraction: Perform the operation shown in the previous step for each table image to obtain a horizontal and vertical line intersection graph, traverse the points of each row, and find the relationship between the x and y values of each cell:

YList [yt] + 3:YList[yt +1], XList [xt] + 3:XList [xt +1](7)

The above series is used to crop each cell from the binary image by using the srcimg function, and finally obtain a picture of the cell we want for the next character segmentation and recognition operation.

2.3 Character Segmentation and Recognition

The single character segmentation in the cell is extracted by projection analysis and then normalized. The projection analysis method divides the characters, the cutting process is to binarize the cell image, and analyzes the distribution histogram of the pixels of the binarized image to find the boundary point of the adjacent characters for segmentation, and projection, what is reflected is the number of pixels in the digital area in the vertical direction. You only need to judge each column of the projection to find the dividing point.

First, transform the image to change it from black on white to white on black; second, define an array to store the number of white pixels in each column of pixels; Then, traversing the binarized image, recording the white (where is digital region) pixels in each column in the array; Finally, draw the projection map according to the gray value in the array, and then find the division point between adjacent characters according to the content of the array, and then divide the characters.

We need to standardize the segmented character pictures before character recognition. The character picture standardization is to convert the segmented character picture into a standard 28×28 pixel standard picture, and the character part is displayed in the part of 20×20 pixels, which can greatly improve the recognition accuracy.

2.4 Identification System Establishment

The establishment of the identification system is divided into three steps of identifying, training and creating a training set.

In terms of identification, based on the open source project of tensorflow, the neural network is built; using the training data and the test data, the neural network will learn independently, adjust the network parameters, and make the overall network shift to the direction of error reduction; Using the Convolutional Neural Network Algorithm

(CNN), it is possible to effectively use the positional relationship between pixels to identify while effectively reducing the amount of computation.

In terms of training, the key to convolutional neural networks lies in the convolution kernel, and the purpose of training is to find a suitable convolution kernel; during training, the network reads in picture information and tag information (true values). In the early stages of training, the accuracy of the training is equal to the reciprocal of the category, ie, the results are completely random. The real tag information can correct the network error. By gradually fine-tuning the network parameters, the error rate will be reduced in the direction of decreasing (gradient descent method). After a lot of corrections, the network will be able to accurately map the pictures to the correct category; After the network is established, the image to be identified is filled into the pre-prepared placeholders in the network for calculation, and the obtained result will show the most similar type of the image and the trained type. Passing a picture which is to be identified, after standardization processing and handing it to the network calculation, we can get a prediction result, which is the identification value of the picture.

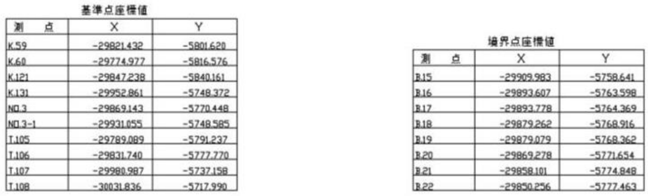
When create the training set, the sample classification of the characters (including but not limited to numbers) which need to be identified is organized and archived, and the same standardized processing is performed, so that it becomes a standard picture of 28*28 pixels, and the gray value of the picture is converted into hexadecimal. The hexadecimal form is written to the training/test set. At the same time, according to the selected picture, the same number of training/test label set files are produced, and the classification of the pictures is also represented by 2-digit hexadecimal numbers.

3 EXPERIMENT AND ANALYSIS

The system uses Python as the programming environment, combined with the OpenCV function library and the NumPy function library and the OS function library to experiment. We used the model to process 283 multi-table images in batches. Each picture contains 5-20 of the number of tables, the structure type is similar, but the size shape is not uniform. The specific processes and functions are described below.

3.1 Picture Preprocessing

We use the OpenCV function library and the NumPy library function in Python to preprocess the picture. Taking Fig.1 as an example, we know that the graph is an image containing two tables , one part of the original images. Firstly, as known that the color image can be grayscale by using the component method, the maximum value method, the mean method and the weighted average method of grayscale processing, or we can directly use the cvtcolor function with the OpenCV library for grayscale processing. After that, by calling the adaptiveThreshold function in the OpenCV library to achieve an automatic threshold of two value, we got a clearer line and text, as shown in Fig.2. Denoising processing using the median filtering method, by calling the medianBlur function in the OpenCV library, the sample picture noise is small. There was not big difference of effect after denoising, so we can save this step operation. Tilt correction requires the use of Hoff Transform. First of all, we should detect the line in the picture through the Hough transform, then, calculate the tilt angle of the line and achieve the rotation correction of the picture. As for picture standard, tilt correction after the effect is very small, we can save this step of operation.



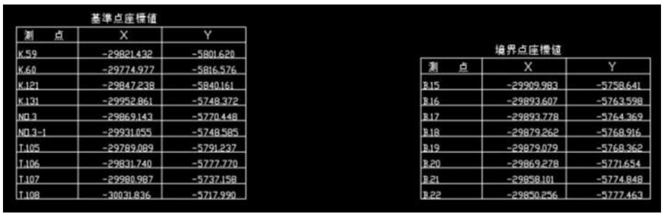
基準点座標値

測 点	X	Y
K.59	-29821.432	-5801.620
K.60	-29774.977	-5816.576
K.121	-29847.238	-5840.161
K.131	-29952.861	-5748.372
NO.3	-29869.143	-5770.448
NO.3-1	-29931.055	-5748.585
T.105	-29789.089	-5791.237
T.106	-29831.740	-5777.770
T.107	-29980.987	-5737.158
T.108	-30031.836	-5717.990

境界点座標値

測 点	X	Y
B.15	-29909.983	-5758.641
B.16	-29893.607	-5763.598
B.17	-29893.778	-5764.369
B.18	-29879.262	-5768.916
B.19	-29879.079	-5768.362
B.20	-29869.278	-5771.654
B.21	-29858.101	-5774.848
B.22	-29850.256	-5777.463

Fig.1 Samples diagram



基準点座標値

測 点	X	Y
K.59	-29821.432	-5801.620
K.60	-29774.977	-5816.576
K.121	-29847.238	-5840.161
K.131	-29952.861	-5748.372
NO.3	-29869.143	-5770.448
NO.3-1	-29931.055	-5748.585
T.105	-29789.089	-5791.237
T.106	-29831.740	-5777.770
T.107	-29980.987	-5737.158
T.108	-30031.836	-5717.990

境界点座標値

測 点	X	Y
B.15	-29909.983	-5758.641
B.16	-29893.607	-5763.598
B.17	-29893.778	-5764.369
B.18	-29879.262	-5768.916
B.19	-29879.079	-5768.362
B.20	-29869.278	-5771.654
B.21	-29858.101	-5774.848
B.22	-29850.256	-5777.463

Fig.2 Binarization pictures diagram

3.2 Positioning of tables and cell extraction

After getting the standard two-valued image, we need to extract the horizontal and vertical lines based on morphological operation, and detect the horizontal and vertical lines through the getStructuringElemen function. In order, the corrosion erode and expansion dilate are processed in the morphological function, and the clear horizontal chart (Fig. 3) and the vertical chart (Fig. 4) are obtained, and the horizontal and vertical lines of Fig.3 and Fig.4 are added together:

$$mask = horizontal + vertical \quad (8)$$

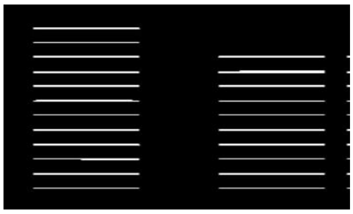


Fig.3 Horizontal diagram

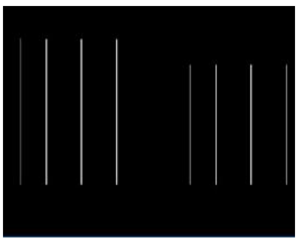


Fig.4 Longitudinal diagram

The combination diagram of the mask is obtained by the above formula (Fig. 5). After the intersection of the Horizontal and vertical lines is obtained by the bitwise_and function, then the edge detection technique is processed on the mask graph. After that, we should find the contour through findContours and determine whether the contour shape and size are tables, and use theapproxPolyDP function to approximate the area to become a shape. The true value indicates that the resulting area is a closed area, and the boundingRect function is used to convert the area to a rectangle which contains the input shape, and finally gets a separate table (Fig.6). Similarly, by deposing a single table to get the intersection diagram, we condense it into a single pixel (Fig.7) and start traversing the points of each row to find the relationship between the X, Y values of each cell:

YList [yt] + 3:YList[yt +1], XList [xt] + 3:XList [xt +1](9)

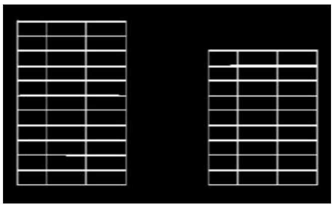
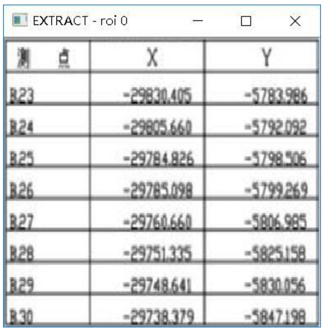


Fig.5 Transverse longitudinal line intersection diagram



EXTRACT - roi 0

测 点	X	Y
B23	-29830.405	-5783.986
B24	-29805.660	-5792.092
B25	-29784.826	-5798.506
B26	-29785.098	-5799.269
B27	-29760.660	-5806.985
B28	-29751.335	-5825.158
B29	-29748.641	-5830.056
B30	-29738.379	-5847.198

Fig.6 Single table split diagram

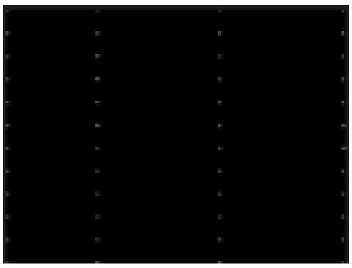


Fig.7 Table intersection diagram

We used the srcimg function to crop each cell from a two-valued image to get the cell image shown in Figure 8.

Here is just an example of a part of the original image. For the original image, we can process a multi-table picture with 20 tables of varying sizes by changing the scale value in the Hough line detection and adjusting it to 61.

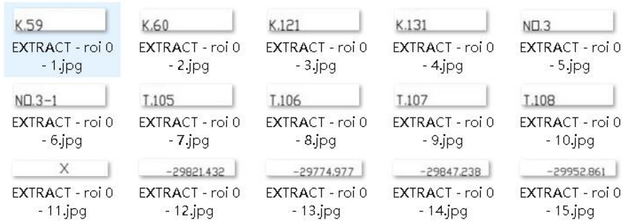


Fig.8 Cell split diagram

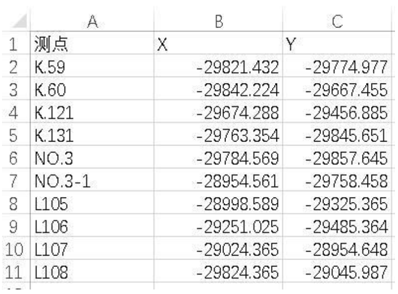
3.3 Character segmentation and recognition

The main method is to extract the single character segmentation in the cell through the projection analysis method and standardize the processing.

First of all, we need to transform the image so that it changes from white to black and white, and define an array named perPixelValue to store the number of white pixels in each column of pixels. Then, we need to traverse the binary image and record the white which is the numeric area pixels in the array in each column. Finally, the projection graph is drawn according to the grayscale value in the array. And through the contents of this array, we found the segmentation point between adjacent characters to split the characters. Next, using the characteristics of image matrix form under NumPy function, we standardized it to get character image becomes standard images of 28x28 pixel. Figure 9 is a split character image and a standardized character image.

Fig.9 Character Standardized diagram

We used a trained neural network to identify these character pictures, imported them into Excel and put them into the corresponding ranks according to the properties of the cells to which they belong, so as to achieve complete recognition processing. The corresponding Excel information is shown in Fig.10.



	A	B	C
1	测点	X	Y
2	K.59	-29821.432	-29774.977
3	K.60	-29842.224	-29667.455
4	K.121	-29674.288	-29456.885
5	K.131	-29763.354	-29845.651
6	NO.3	-29784.569	-29857.645
7	NO.3-1	-28954.561	-29758.458
8	L105	-28998.589	-29325.365
9	L106	-29251.025	-29485.364
10	L107	-29024.365	-28954.648
11	L108	-29824.365	-29045.987

Fig.10 Identified information is imported into Excel

Through statistical recognition results, we found that the accuracy of printing numbers and alphabet recognition reached 99.8%. At the same time, we experimented with other types of images. The statistics showed that the accuracy of handwritten digital recognition reached 98.6%, handwritten English reached 97.4%, printed Chinese characters reached 94.2%. According to the accuracy of the recognition, it is not difficult to conclude that the multi-table recognition system based on deep learning and edge detection can not only efficiently complete the segmentation of multi-table images, but also accurately identify the digital letters of the printed handwriting and the Chinese characters of the printed body, which has strong practicability. However，we found that the drawback is that the segmentation technology is not perfect, and there's still a need for progress on studying the situation of linking words. Therefore, the accuracy of handwritten Chinese characters is low, so we should further optimize the recognition model.

4 CONCLUSION

A multi-table image recognition system based on deep learning and edge detection mentioned in this paper has the perfect efficiency of table recognition processing. Likewise, the accuracy of character recognition is high, and the recognition processing speed is fast. This system improves the recognition efficiency and accuracy when batching the table images, which brings great convenience and benefit to the companies engaged in information recognizing and entering.

REFERENCES

[1] Rasche,Karl, Geist,Robert, Westall James, Detail preserving reproduction of color images for monochromats and dichromats, IEEE computer graphics and applications, Vol.25, No.3, 22-30, 2005.

[2] Song Min-gli, Wang Hui-qiong, Chen Chun, Gaussian Mixture Model Based Approach on Color Transfer, Journal of Computer-Aided Design & Computer Graphics, Vol.20, No.11, 1471-1476 , 2008.

[3] XU Yar-ning, TIAN Peng, Adaptive Binary Research Based on the Width of Cracks, Journal of Geodesy and Geodynamics, Vol.36, No.03, 239-243 , 2016.

[4] LU Di, HUANG Xin, LIU Chang-yuan, LIN Xue, ZHANG Hua-yu, YAN Jun, Binarization Method Based on Local Contrast Enhancement, Journal of Electronics & Information Technology, Vol.36, No.03, 239-243 , 2016.

[5] Li Hong-yan, Zhou Yun-long, Tian Feng, Li Song, Sun Tian-bao, Wavelet-based vibration signal de-noising algorithm with a new adaptive threshold function, Chinese Journal of Scientific Instrument, Vol.36, No.10, 2200-2206 , 2015.

[6] LIU Yong-cai, BAO Yi-dong, WANG Xiao-ping, CHEN Wen-liang, On logic-based intelligent systems, Modern Electronic Technique, Vol.40, No.17, 36-39 , 2017.

[7] FENG Lei, GENG Ying-nan, Fast Correction of Skew Document Image Based on Hough Transform, Journal of Inner Mongolia Normal University, Vol.46, No.02, 219-222 , 2017.

[8] Nikola, Kasabov, GUO Zhi-qing, JIA Zhen-hong, YANG Jie,, The Biochip Image Tilt Correction Algorithm Based on Roundness Criterion, Vol.38, No.06, 58-61 , 2017.

[9] WANG Quan, WANG Lai-jing, WANG Bo, MAZ segmentation approach to the table-form image, Journal of Xidian University, Vol.35, No.02, 293-296 , 2008.

[10] Huang Li-chen, Zhang Jian, Zhou You, Tan Fang-fang., An Improved Image and Text Binarization Algorithm Based on Non-Uniform Illumination, Hunan University of Science and Technology, Vol.27, No.06, 40-45 , 2013.

[11] NIE Ren-can, HE Min, ZHOU Dong-ming, YU Jiang, DING Xing-li, Global threshold segmentation algorithm for visual form images, Laser & Infrared, Vol.47, No.02, 234-238 , 2017.

[12] MA Rui,GU Yun-hua,YAN Yun-yang, An Approach Based on Recognition and Postprocessing for Handwritten Letters, Journal of Wuhan University of Technology, Issue 16, 34-38 , 2010.

[13] Abdusalam Dawut, YANG Yi, Askar Hamdulla, Research on the extraction of Uyghur-Chinese mixed text lines from images with complex backgrounds, Laser Journal, Vol.35, No.07, 5-10 , 2014.

[14] LUO Jia, Study on a Fast Segmentation Algorithm of Adhesion English String, Computer Technology and Development, Vol.24, No.08, 59-62 , 2014.

[15] Mnih V, Kavukcuoglu K, Silver D, Rusu A A, Veness J,Bellemare M G, Human-level control through deep reinforcement learning, Nature, Vol.518, No.7540, 529-533 , 2015.

[16] LU Gang, Recognition of multi-fontstyle characters based on convolutional neural network, Journal of Zhejiang Normal University, Vol.34, No.04, 425-428 , 2011.