The 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
18-21 September, 2019, Metz, France

Machine Learning Based Medicine Distribution System

Huiling Xia [1], Chunzhi Wang [1], Lingyu Yan [1] ,Xinhua Dong,Yichao Wang
[1]School of Computer Science, Hubei University of Technology
Huiling Xia, China
[1]2246735918@qq.com, [1]chunzhiwang@vip.com, [1]yanlingyu@hbut.edu.cn, xydongxh@126.com, echarwang@163.com

Abstract—**In this thesis, an intelligent medicine recognition method for medicine vending machines has been proposed. Considering that the name of the medicine is generally the largest character in the medicine box, the medicine name recognition can be converted into the largest character recognition in the picture in this thesis. First, we used Support Vector Machine (SVM) and Connected Component to determine the text region, and found the largest connected region to locate the medicine name. Secondly, we use the method of "fragment link" for text segmentation, which divides the text into two elements: fragment and link, and combines the whole word according to the set rules. Finally, we can directly use Optical Character Recognition (OCR)software for character recognition. Experiments show that this method has a high accuracy of medicine identification.**

Keywords—medicine recognition, Medicine Identification Center (MIC), fragment, link

I. INTRODUCTION

•With the rapid development of artificial intelligence, machine learning and intelligent recognition has been applied to human life. Today, vending machines have long been a new form of retail commerce. However, there are many vending machines for drinks and snacks in the market. Considering that many consumers' demand for medicines is generally sudden, and there are often many hospitals, queues, or urgent use of medicines at night, this often makes consumers feel very distressed, and many pharmacies are costly. In terms of problems, it is generally not available 24 hours a day, which often leads to the difficulty of buying medicines. Self-administration makes it necessary to check medicines. Self-administration makes it necessary to check medicines.

Since the microcomputer for setting the photograph is located at the exit of the medicine vending machine, the light is dark, which makes scanning the barcode less practical. Therefore, we take the method of extracting the name of the medicine on the medicine box to check the medicine. As early as the 1970s, image text extraction [1,3,9,14] and text recognition [2,4,7,10] began to attract people's attention. Based on this, this thesis presents an intelligent medicine recognition for medicine vending machines based on text recognition in images. The model consists of two parts: Medicine Identification Center (DIC) and Terminal, which is shown in Figure 1. It not only brings convenience to patients, but also greatly improves the working efficiency and management level of the hospital.

The rest part is shown below:

II: Approach. The working principle of Medicine Identification Center is introduced in detail, and the combined algorithm is introduced.

III: Experiment. We test the data set by using the algorithm, and finally analyze the experimental results.

IV: Conclusion. We summarize the article and the areas to be improved or a prospect of the future.

II. APPROACH

A. Detection method based on Connected Component

•The Connected Component generally refers to the image region composed of foreground pixels with the same pixel value and adjacent positions. Connected Component Analysis refers to finding and marking each connected region in the image. The detection method based on the Connected Component extracts the connected region in the image as the candidate of the text region through the Maximum Stable Extreme Region method, and then distinguishes the connected region of the text through the SVM classifier [5]. The core of this method is to find the region with the same gray value inside the text. By finding the connected region [8] with the same gray value, the region belonging to the text can be found. Then, the largest connected region is the region where we need the medicine name. Its core is to locate the text area [6,11,12,13,15].

B. The fragment link method

•This is a fast arbitrary text detection method. The idea is to break the text down into two basic elements, fragments and links. A fragment is a partial bounding box that covers a small piece of word or text box, represented by an angled rectangle[1,3].The height of the rectangle is close to the height of the whole word, the link exists between two adjacent fragments and represents the connection relation of fragment time. Linked fragments belong to the same word, and unlinked words belong to different words. During detection, the method detects fragments and links intensively in the whole graph, and forms the connected fragments into a whole word bounding box to obtain the detection results.

C. *The mounting position of microcomputer*

•The outside image of a medicine box is obtained by a micro-camera installed at the exit of the medicine vending machine, whenever a medicine has been taken out by the patient. The specific installation diagram is shown in Fig 1：

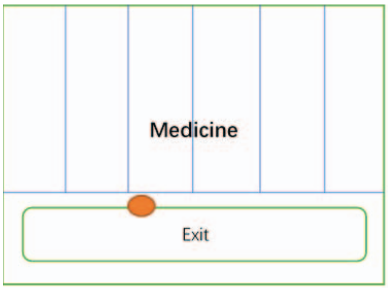

Fig.1. installation diagram of microcomputer

Note: represent the microcomputer

D. *Medicine Identification Center(MIC)*

When a medicine is about to be taken out, a tiny camera takes a picture ， and the image is sent to a Medicine Identification Center (DIC). The workflow of MIC is shown in Figure 2.

E. *Medicine Recognition method Based On OCR and SVM*

- Step 1: We grayed and binarized the image, let the picture become simple, easy to deal with.

For an image to be grayed, we conduct the research on RGB(A), and R, G and B in RGB(A) are respectively represented Red, green, blue, and possibly Alpha transparency. There are two methods of transformation, and choose the one that performs better. That is:

Grayed R = R * 0.3 before treatment + G * 0.59 before treatment + B * 0.11 before treatment

Grayed G = R * 0.3 before treatment + G * 0.59 before treatment + B * 0.11 before treatment

Grayed B = R * 0.3 before treatment + G * 0.59 before treatment + B * 0.11 before treatment

Thus, we can get each pixel gray value of a picture. And then we binarize it, calculate the average value avg of gray value of all pixel points in the pixel point matrix, then compare the each pixel with avg. If it's less than or equal to avg, it's 0 (black), larger than avg are 255(white).

- Step 2: the training set data are trained with SVM algorithm for classification modeling, and the optimal penalty factor C and g of kernel function are determined. SVM can be divided into hard interval SVM and soft interval SVM. Here, the soft interval C-SVM is used to improve the recognition accuracy. The description about the algorithm of nonlinear classification (Fig. 3):

As is shown in Figure 3, the interval $M=\frac{2}{\sqrt{W.W}}$. Under the condition of linear segmentation, the optimal function is $\min\frac{1}{2}|w|^2$. In the case of linear indivisibility, the soft interval C-SVM is equivalent to adding a penalty function after the original optimization function (C is the penalty factor), with the constraint condition:

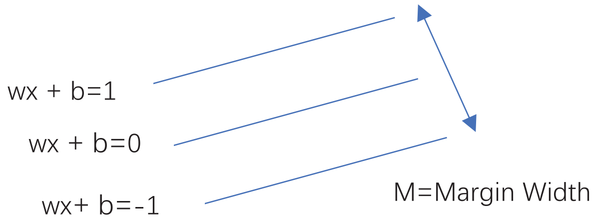

Fig. 3. The definition of maximum interval

$$\min\frac{1}{2}|w|^2+C\sum_{i=1}^{R}\varepsilon_i$$

$$s.t., y_i\left(w^T x_i+b\right)\geq 1-\varepsilon_i, \varepsilon_i\geq 0$$

$$y_i=\{-1,1\}$$

It is transformed into dual problem, then be solved.

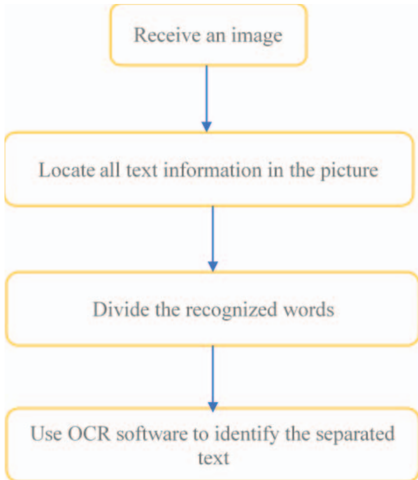

Fig.2. the workflow flowchart of MIC

- Step 3: for the connected region obtained in step 1, find the maximum connected region, that is, the connected region where the medicine name is located. In fact, the connected region where the medicine name is located is the second largest connected region, because the largest connected region must be the background of the picture.

- Step 4: We decompose the contents of the connected region obtained in step 3 into segments and links to distinguish words and characters. Because both fragment and link are local, fragment only accounts for a small part of the whole word, and it only needs to detect the local features of the image. Because links exist between two linked fragments, they are also local. Fragments and links can be densely detected on pictures of any size to form words of any length, which can be used to detect long words or long words of non-horizontal words.

- Step 5: The information obtained in step 4 is identified with OCR software to obtain the name of the medicine.

III. Experimental Reesult and Analysis

A. *Experimental database introduction*

The experiment used a self-made database. All the image data was taken from the medicine vending machine. It consists of ten medicines. There were 10 pictures of each medicine, 100 in total. After processing, the resolution of each image is 450*300 =135000, which is an RGB tricolor image. Each medicine in the data set was divided into two groups, the first seven as training sets and the last three as test sets. There are 70 samples in the training set and 30 samples in the test set.

Hardware: Lenovo, PC, windows 10, intel (i7). Software: Python3.6、Matlab.

B. *Experimental Process of MIC*

Image preprocessing

Numbering each picture by category. The number is the number of each category. The first 7 pictures of each medicine were placed in the training set to form the sample training set. The last three pictures were placed in the test set to form the test sample set. The medicine image was preprocessed into 256*256 format and Matlab image cutting function was used. As shown in Figure 4:

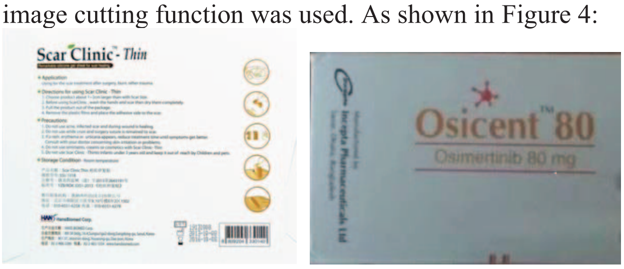

Fig.4. The medicine images

Grayscale and binarization of image

The picture in FIG. 4 is first grayed and then binarized, and the results are shown in Fig. 4.

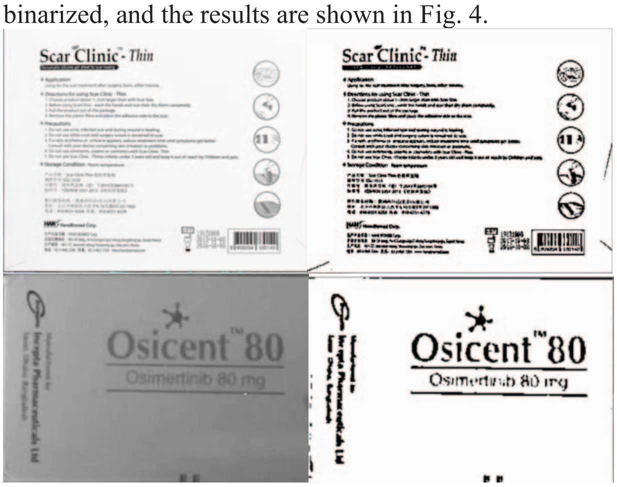

Fig.5. The medicine image after gray level binarization

Find the largest text connected region

MSER algorithm was used to find the region with the same gray value inside the text, and the connected region in the image was extracted as the candidate region of the text. Then, the connected region of the text and non-text was distinguished by SVM, so as to find the largest connected region of the text. Thus, the maximum connected text region in Fig. 5 can be obtained, as shown in Fig. 6 :

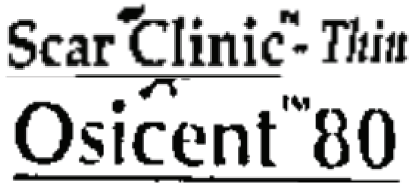

Fig.6. The largest connected text region obtained

Text fragment links

Given an image of size I, the fragment link will predict the fragment and link at each location on the image. Next, the detected fragment is combined by the detected link to get the detection result of the whole word. The detection result of the whole word will be represented by the rectangular box, where is the center point of the rectangular box, and is the length and width of the rectangular box.

Fragment links are shown in Fig 7, where a yellow border represents a fragment and a green line between adjacent borders represents a link between two fragments, and the whole word can be determined after each fragment is linked.

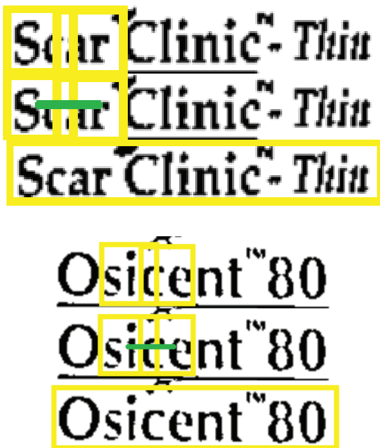

Fig.7. Fragment link diagram

Identify the text linked from the fragment

After obtaining the detection information linked from the connected region segment, the OCR software package is used to identify the medicine name. The results are shown in Figure 8:

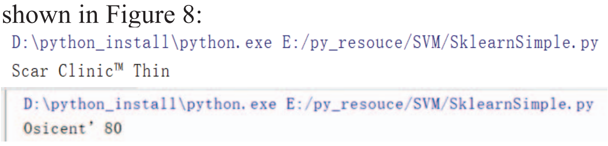

Fig.8. Fragment link diagram

Then we compare the proposed method with the method based only on the connected region, and the method based on the connected region identifies the result as shown in Figure 9:

```
D:\python_install\python.exe ″E:/py_resouce/SVM/Connected Component.py″
ScarClinic- Thin
D:\python_install\python.exe ″E:/py_resouce/SVM/Connected Component.py″
Osicent'″80
```

Fig.9. Recognition result based on Connected Component

As can be seen from Fig.8 the results of identification, the accuracy of this method is higher, size comparison modest words all can normal testing, only like Angle smaller font, such as the mark in the gray binarization process effect is not obvious, may lead to error appear in the test identification, but generally small words such as superscript information will not affect medicine information, so our experiment results from the overall still is very meaningful. It can effectively solve the problem that patients may make mistakes when taking medicine by themselves, and can also relieve the pressure of the hospital window.

IV. CONCLUSION

This thesis proposes an intelligent medicine recognition method for medicine vending machines. The text area was determined by SVM and connected region to determine the medicine name. The text was segmented by fragment link method and then identified by OCR. The experimental results show that the method can better identify the medicine name. However, for small fonts, due to the small number of pixel points after binarization, it is not possible to get a more accurate result. In this aspect, we still need to improve. Due to time constraints, the text in the medicine boxes studied in this thesis does not consider the problem of color text binarization to 255(white). The problem will be to study the adaptive threshold of the main color image later, and try to solve this problem as soon as possible. The results will also be compared with SVM and DNN to better demonstrate the efficiency of our proposed method.

ACKNOWLEDGMENT

Thanks for the funding supported by the National Natural Science Foundation of China, grant number 61772180, Major Technological Innovation Project of Hubei Province 2019(Research and Application of Intelligent Optimizing Key Technologies of Transmission Network Based on Deep Evolutionary Neural Network, Fiber Communication Technology Co., Ltd.), Green Industry Science and Technology Leadership Program of Hubei University of technology (No.CPYF2018005，YXQN2017002).

REFERENCES

[1] Gomez, Lluis, and Dimosthenis Karatzas. "Multi-script text extraction from natural scenes." 2013 12th International Conference on Document Analysis and Recognition. IEEE, 2013.

[2] Zhang, Chengquan, et al. "Look More Than Once: An Accurate Detector for Text of Arbitrary Shapes." arXiv preprint arXiv:1904.06535 (2019).

[3] Phan, Trung Quy, Palaiahnakote Shivakumara, and Chew Lim Tan. "Detecting text in the real world." Proceedings of the 20th ACM international conference on Multimedia. ACM, 2012.

[4] Shi, Baoguang, Xiang Bai, and Serge Belongie. "Detecting oriented text in natural images by linking segments." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.

[5] Codella, Noel, et al. "Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images." International workshop on machine learning in medical imaging. Springer, Cham, 2015.

[6] Gupta, Ankush, Andrea Vedaldi, and Andrew Zisserman. "Synthetic data for text localisation in natural images." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

[7] Mohammad, Saif M. "Sentiment analysis: Detecting valence, emotions, and other affectual states from text." Emotion measurement. Woodhead Publishing, 2016. 201-237.

[8] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[9] Yin, Xuwang, et al. "Effective text localization in natural scene images with MSER, geometry-based grouping and AdaBoost." Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). IEEE, 2012.

[10] Jaderberg, Max, et al. "Reading text in the wild with convolutional neural networks." International Journal of Computer Vision 116.1 (2016): 1-20.

[11] Neumann, Lukas, and Jiri Matas. "A method for text localization and recognition in real-world images." Asian Conference on Computer Vision. Springer, Berlin, Heidelberg, 2010.

[12] Neumann, Lukas, and Jiri Matas. "Text localization in real-world images using efficiently pruned exhaustive search." 2011 International Conference on Document Analysis and Recognition. IEEE, 2011.

[13] Neumann, Lukáš, and Jiří Matas. "Real-time scene text localization and recognition." 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012.

[14] Pach, Jakub Leszek, and Piotr Bilski. "A robust binarization and text line detection in historical handwritten documents analysis." International Journal of Computing 15.3 (2016): 154-161.

[15] Soni, Rituraj, Bijendra Kumar, and Satish Chand. "Text detection and localization in natural scene images based on text awareness score." Applied Intelligence 49.4 (2019): 1376-1405.