ISSN: 2688-8645

Bill Information Recognition System based on Deep Learning

Shiji Zheng[1,*], Jiahao Xie[1]

[1]School of Information Engineering, Jiangmen Polytechnic, Jiangmen, 529000, China

Abstract

This research is based on image processing technology and deep learning technology, through intelligent image preprocessing, target region location and cutting, CNN deep learning classification and recognition technology, carries on the fusion depth application in the project, realizes the intelligent bill information recognition and storage with high precision and high efficiency, and can greatly improve the efficiency of enterprise informatization in the practical application.

Keywords

Bill recognition; Image processing; Convolution neural network; Deep learning.

1. Introduction

The basis of bill recognition technology is OCR technology. Due to the wide circulation of bills, in order to deal with these documents quickly, many European and American countries have studied OCR technology for a long time, but it was not until the 1990s that English document recognition technology began to be put into practical application. The breakthrough of Chinese character recognition technology is in the 1860s. With the proposal of template matching method and the continuous optimization and improvement in the later stage, the accuracy of Chinese character recognition has been continuously improved.

Deep learning is an algorithm that can obtain the relationship model between data from massive data. the more classical one is convolution neural network, which can extract the characteristics of data through learning and establish a model to accurately judge and predict unknown events. With the continuous development of deep learning and image processing technology, the recognition of all kinds of complex format bills has become possible. However, from a global point of view, the bill identification system still has the following shortcomings: the cost of purchasing equipment at the initial stage is high, and a lot of manpower and material resources are needed for later maintenance. Therefore, it is necessary to improve the accuracy and recognition efficiency of the bill information recognition system on the basis of the original technology.

2. Convolutional Neural Network

Convolutional neural network is the most widely used deep learning model. It can learn the features comprehensively and effectively through training. Compared with the traditional neural network, convolutional neural network takes advantage of the characteristics of convolution algorithm, and greatly reduces the complexity of the network structure by means of partially connected, local receptive field, shared weights and pooling. At present, convolutional neural network has been widely used in natural language processing, climate prediction, artificial intelligence and other fields. Because convolutional neural network can learn the characteristics of different levels from a large number of data, it will have a very broad application prospect in the field of network intrusion detection. The classical convolution neural network model is mainly composed of five parts: input layer, convolution layer, pooling layer, full connection layer and classification layer, as shown in Figure 1.

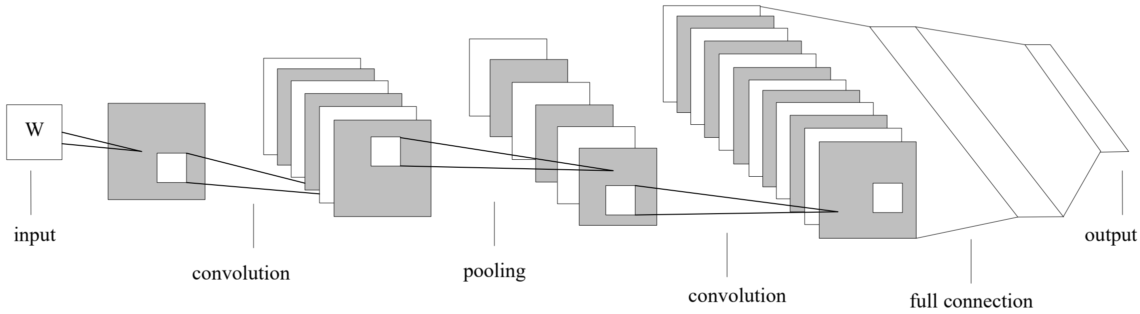

Figure 1. Classical convolution neural network model

2.1. Convolution Layer

The convolution layer is the most important part of convolution neural network. After the feature image is input into the convolution layer, it will convolute with the convolution kernel. The convolution formula is as follows:

$$\mathrm{x}_j^l = \mathrm{f}\left(\sum_{\mathrm{i}\in \mathrm{P}_j} \mathrm{x}_i^{l-1} * \mathrm{k}_{ij}^l + \mathrm{b}_j^l\right) \tag{1}$$

$f\left(\sum_{i\in P_j} x_i^{l-1} * k_{ij}^l + b_j^l\right)$ is a tanh function,P_j is a local receptive field,$\mathrm{x}_i^{l-1}$is the value of the l-1 feature on the i window, the position on the 1st floor is (i, j), k_{ij}^l is the weight of the convolution kernel,$\mathrm{b}_j^l$is the offset of the feature.

During convolution operation in convolution layer, the convolution kernel slides on the whole input feature according to the set step size, and simultaneously multiplies and sums the parts corresponding to the local receptive field to perform convolution operation until the convolution kernel slides out of the input feature.

2.2. Pooling Layer

The role of pooling layer is to reduce the dimension of features and filter out redundant features, so as to reduce the computation and improve the generalization ability of the network. According to the different sampling methods, the pooling layer can be divided into maximum pooling and mean pooling. The pooling process can be expressed as follows:

$$x_j^i = f\left(pool\left(x_i^{l-1}\right) + b_j^l\right) \tag{2}$$

x_i^{l-1} is the value of the i window in the layer l-1 input feature,b_j^l is the offset of the j window in layer l, and pool represents the sampling function.

2.3. Full Connection Layer

The convolution neural network generally configures a full connection layer after the convolution layer and pooling layer. The neurons in the full connection layer are connected to each other in the previous layer. The calculation formula is as follows:

$$x^l = f(u^l) \tag{3}$$

$$u^l = W^l x^{l-1} + b^l \tag{4}$$

$f(u^l)$ is the activation function,W^l is the weight of layers l-1 to 1,b^l is the offset of layer 1,x^{l-1} is the output feature of layer l-1.

ISSN: 2688-8645

2.4. Classification Layer

The effect of convolutional neural network model largely depends on the selection of features and classification layer classifier. In general, if you have good features, even if you choose a simple classifier, such as Support Vector Machine (SVM), you can get good results. However, when SVM is used in large-scale detection applications, it is usually constrained by time and space, and its efficiency is not high. In this paper, Softmax is used as classifier[3]. Softmax is suitable for multi classification. Its expression function is as follows:

$$h_0(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | (x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta \\ \cdots \\ p(y^{(i)} = k | x^{(i)}; \theta \end{bmatrix} = \frac{1}{\sum_{i=1}^{k} e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \cdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \tag{5}$$

The output of the function is a k-dimensional vector (the sum of the vector is 1), which is used to represent the probability value of the k estimates. That is to say, for a given input x vector of k dimension, the probability p (y = i| x) of each category j is calculated to achieve the purpose of estimating the probability value of each classification output result of input x.

3. System Structure and Testing

The goal of this study is to solve the problem of information identification of all kinds of Chinese bills, and to realize the intelligence and automation of information acquisition, so as to improve the information level and work efficiency of enterprises to a great extent. The system structure is shown in figure 2.

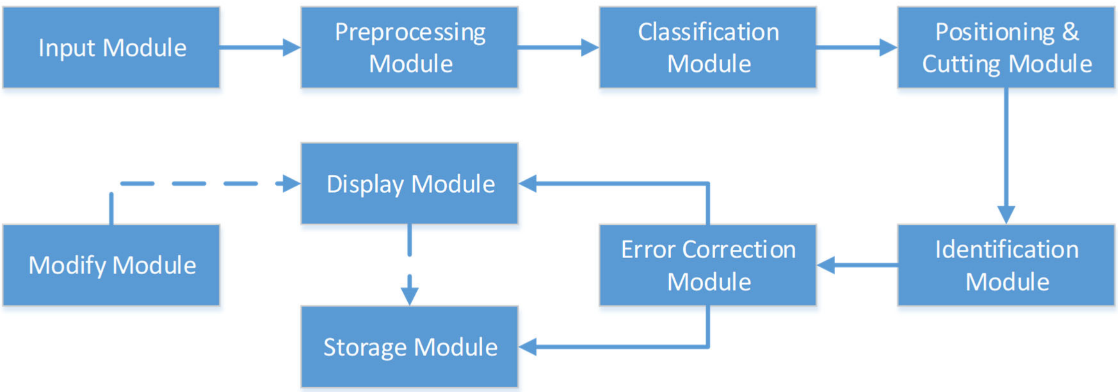

Figure 2. System structure

(a) Efficient intelligent image preprocessing module. According to the characteristics of bills, a new bill seal removal algorithm based on RGB channel and geometry is designed.

(b) The positioning and cutting module of accurately locating the region where the text is located and accurately cutting individual Chinese characters, English letters, numbers, special symbols and other characters from the text area.

(c) Real-time image classification and recognition module with high recognition rate and fast running speed.

(d) The training sample generation module with high quality and complete categories.

(f) In order to ensure the high stability of the recognition system and improve the recognition accuracy, the high-performance error correction module is realized by combining Huawei Cloud, Tencent Cloud and other technologies.

First of all, use the mobile phone or camera, scanner to collect the image of the bill, and then carry on the intelligent preprocessing of the bill image, including normalization, judging whether the bill needs to be flipped, binarization, morphological processing and so on. Then the text location algorithm is used to complete the information text location, and the improved projection cutting method is used to complete the single character cutting. For character recognition, because the information recognized in this study are numerals and Chinese printed characters, neural network is used as the recognition algorithm of this system. After testing, the overall efficiency of the system bill recognition is very high, the sum of the highest average response time of each part is 0.2s/, the overall recognition accuracy of the test is 99.9%, and the recognition accuracy of normal uncontaminated and folded bills is 100%.

The recognition interface and results of the system test are shown in figure 3.

Figure 3. System test results

4. Conclusion

In order to solve the problem of low recognition rate of bill characters, this study improves the accuracy and efficiency of bill classification and character information recognition by innovating image processing technology, deep learning and neural network technology. the application and popularization of the research results will certainly produce good economic and social benefits.

References

[1] Khalid Aboura, Rami Al-Hmouz. An Overview of Image Analysis Algorithms for License Plate Recognition[J]. Organizacija , 2017, 50(3):285-295.

[2] Bai X, Yang M, Shi B, et al. Deep learning for scene text detection and recognition[J]. SCIENTIA SINICA Informationis, 2018, 48(5): 531-544.

[3] Yoo H J. Deep Convolution Neural Networks in Computer Vision[C]. Ieie Transactions on Smart Processing & Computing, 2015, 4(1):35-43.

[4] Huang G , Liu Z, Maaten LVD, et al. Densely Connected Convolutional Networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.

[5] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. in: Proceedings of Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.