Research on Image Recognition Method of Supermarket Commodities Based on Convolutional Neural Network

Shugang Liu[1, a], Xing Chen[1, b]

[1]School of Computer Science, NORTH CHINA ELECTRIC POWER University, Baoding 071003, China.

[a]274786113@qq.com, [b]1943203577@qq.com

Abstract

At present, image recognition has become an important research direction in the field of image research. Image recognition technology has also received great attention in the field of machine vision, which is of great significance. In recent years, deep learning has made great progress in terms of images, text and speech[1] .Image recognition technology mainly includes image feature extraction and classification recognition[2].In the identification process, feature extraction is the key to determining the performance of recognition. Deep learning establishes a hierarchical model structure similar to the human brain, extracting the characteristics of the data layer by layer and constructing the model. Applying deep learning to image recognition can further improve the accuracy of image recognition[3,4]. In this paper, the image recognition algorithm based on convolutional neural network is deeply studied. Considering that different activation functions and downsampling methods have a great influence on the recognition performance, the appropriate activation function and downsampling method are selected through experiments, and these activation functions and downsampling methods are discussed and analyzed[5]. Finally, the image recognition results based on convolutional neural network are compared with the results of traditional image recognition methods, and the effectiveness of the method is proved.

Keywords

Deep learning, image recognition, activation function, convolutional neural network.

1. Introduction

With the continuous development of science and technology, unmanned supermarkets have begun to enter the people's field of vision, and there have been "unmanned supermarkets" in Beijing, Shanghai, Hangzhou, Chengdu and other places. Commodity identification is an important step in the implementation of unmanned supermarkets.

The entire implementation process of the entire unmanned supermarket intelligent project is: installing camera equipment on the shelf, collecting the shelf image during the shopping process, and inputting the captured image into the image recognition system[6]. The system uses image processing technology to identify the type, number, and location information of each item in the image, so that the situation of the item purchased by the customer can be further derived[7]. This research is a preliminary exploration of intelligent management of supermarkets and has certain research prospects.

2. Research Background and Content

As an aspect of intelligent management of supermarkets, the research on supermarket intelligent shelf image recognition has high research value. Image is the main source of human acquisition and exchange of information. In recent years, image processing technology has been applied to various

fields of human life and work. With the continuous expansion of human activities, the application field of image processing will continue to expand.

It is believed that in the near future, smart shelf image recognition technology will be applied to the intelligent management of supermarkets closely related to people's lives. With the application of this technology, the labor cost and capital cost of supermarket management will be greatly reduced. The overall benefits of the supermarket, so the research of smart shelf image recognition technology is very meaningful.

This thesis will carry out research on the image recognition of unmanned supermarket shelves, introduce the knowledge of unmanned supermarket management, and the direction in which the intelligent management of unmanned supermarkets can be further studied[8,9]. The research significance and application prospects of the image recognition of unmanned supermarket shelves are introduced in detail.

These research contents are based on deep learning, and the customer's shopping information is obtained by image recognition and analysis of the collected images, for example, what goods the customer bought, the location and the quantity.

3. Design of Commodity Image Recognition Scheme

(1) Take a photo when opening the smart shelf door.

(2) Take a photo after closing the smart shelf door.

(3) After closing the door, perform image analysis and send data to the server.

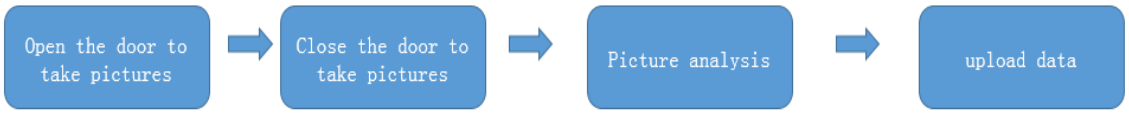

Fig 1. Commodity image recognition scheme flow chart

4. Image Recognition Based on Convolutional Neural Network

4.1 Image Recognition Algorithm Selection

In order to overcome the shortcomings of the lack of learning ability in the shallow machine model, the traditional neural network enhances the recognition performance by increasing the network depth and the number of hidden layer nodes, but the depth increase makes the network more sensitive to the initial weight and easy to fall into[10]. Locally optimal, leading to serious over-fitting and other issues.

CNN is the first deep learning architecture to successfully train multi-layer neural networks. It is a multi-layer perceptron specially designed to recognize two-dimensional images[11].

The advantages of CNN are mainly reflected in the following two points:

(1) Directly taking a two-dimensional image as input avoids complicated pre-processing and data reconstruction processes;

(2) By means of spatial relationship and network sharing mode, the number of training parameters needed in the network is greatly reduced while reducing the complexity of the model, and the generalization performance of the network is greatly enhanced[12].

4.2 Image Recognition

The structure of CNN generally needs to be set according to the size of the input image. Since the depth of CNN has a great influence on the final classification result, the experiment first uses 32*32 size images for experiments, and then expands based on this framework[13].

The size of the convolution kernel selected by the experiment is 5*5. Since the size of the convolution kernel determines the size of the local connection region, selecting the appropriate convolution kernel size has a great influence on the recognition performance of CNN[14], if the convolution kernel Too

small a size can result in an inability to extract valid local features from the input image, and if the size is too large, the complexity of the features extracted from the input image exceeds the representation capability of the convolution kernel.

The convolution operation is as shown in the figure. The window moves in steps of 1. The convolution operation of the input layer is performed by using six convolution kernels of size 5*5. The convolution result is added with 6 offset vectors. An activation function is calculated to obtain 6 different feature maps in the convolutional layer.

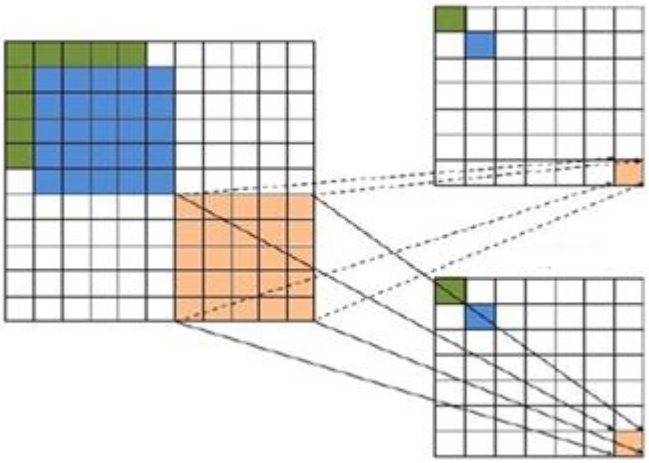

Fig 2. Convolution budget diagram

4.3 Common Activation Function

In the CNN structure, except for the input layer neurons, there is no activation function, and all other layers of neurons can have some activation function as a transfer function. CNN uses the activation function to select the features extracted from the network in a nonlinear mapping manner to avoid the problem of insufficient linear operation expression[14]. In practice, the common activation functions can be generally divided into saturated nonlinear functions ,such as Sigmoid function and Saturated nonlinear function ,Re LU, Leaky Re LU function.

4.4 Classifier Selection

The experimental classifier uses the Softmax regression classifier, which is based on the Logistic Regression classifier. Logistic regression is a very efficient classifier. It generally solves the problem of two classifications. If you encounter multiple classification problems, you can only use multiple independent binary classifiers to make one-to-one discrimination[15].

Softmax regression is a generalization of logistic regression on multi-classification problems. It can complete classification tasks for multiple categories at one time, and Softmax regression can converge to global minimum, avoiding over-fitting caused by local convergence. Softmax Regression is a supervised single-layer neural network classifier that is currently used in conjunction with deep learning methods for classification and identification problems.

5. Conclusion

The experiment was carried out under the environment of Intel Core i5, 2.30GHz CPU, 2G memory Windows operating system and Matlab operating platform. The number of iterations of this experiment is 500, and the learning rate is set to 0.01.Based on the CNN structure described above, the classifier selects the Softmax regression method to test the recognition error rate of CNN under different activation functions.

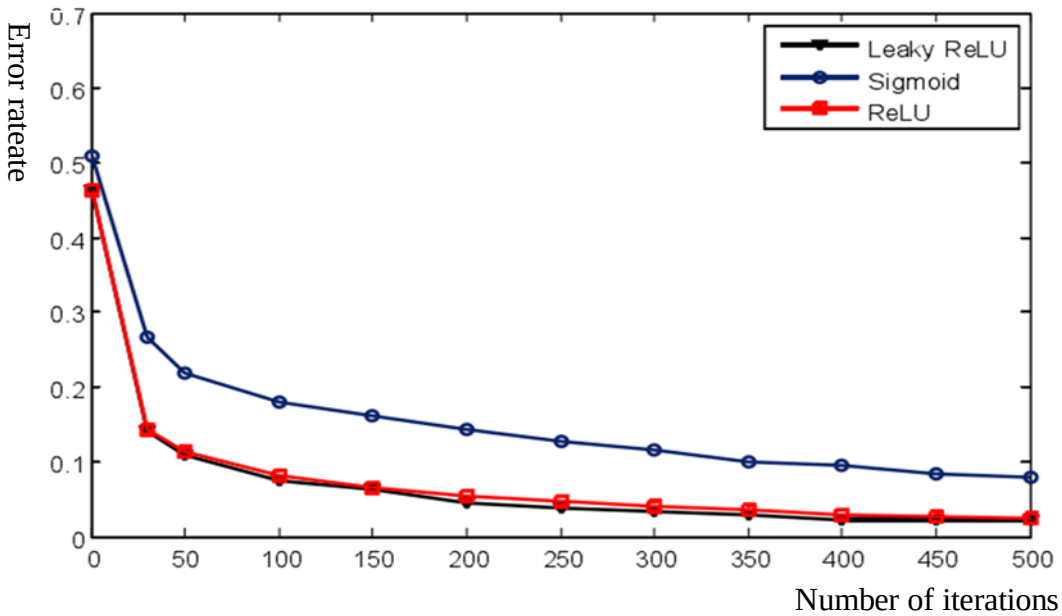

Fig 3. Comparison of CNN results under different activation functions

After 500 iterations, the recognition error rates of CNN with Sigmoid, Re LU and Leaky Re LU as activation functions were 7.80%, 2.46% and 2.04%, respectively. It can be seen that the recognition error rate obtained by using the nonlinear correction function as the activation function in CNN is significantly lower than the recognition error rate using Sigmoid as the activation function.

This experiment uses CNN to collect and identify commodity images, and analyzes the influence of different experimental conditions on the recognition results in detail. The recognition performance of CNN under different activation functions is analyzed through experiments. Through the comparison of experimental results, in order to further improve the recognition performance of the algorithm, the activation function with more accurate recognition results is selected, which proves the effectiveness of using convolutional neural network for commodity image recognition, and shows the advantage of CNN in commodity image recognition.

References

[1] Weigel VB. Deep learning for a digital age: Technology's untapped potential to enrich highereducation. John Wiley & Sons, Inc, 001.

[2] Ngiam J, Coates A, Lahiri A, Prochnow B, Le QV, Ng AY. On optimization methods for deep earning.In: Proceedings of the 28th International Conference on Machine Learning (ICML-11), 2011.

[3] Lee H, Pham PT, Largman Y, Ng AY. Unsupervised feature learning for audio classification usingconvolutional deep belief networks. In: NIPS, 2009.

[4] World Health Organization.Global status report on noncommunicable diseases 2010[R]. Geneva:WHO, 2011.

[5] Bengio Y, Delalleau O. On the expressive power of deep architectures[C]. Algorithmic Learning Theory, Berlin Heidelberg, 2011.

[6] Roux N, Bengio Y. Representational power of restricted Boltzmann machines and deep beliefnerworks [J]. Neural Computation, 2008.

[7] Larochelle H, Bengio Y, Louradour J, et al. Exploring strategies for training deep neural networks [J].Journal of Machine Learning Research, 2009.

[8] Breakthrough Technologies 2013[N]. MIT Technology Review, 2013.

[9] BENGIO Y.Learning deep architectures for AI [J]. Foundations and Trends in Machine Learning, 2009.

[10] Braverman M. Poly-Logarithmic independence fools bounded-depth boolean circuits [J].Communications of the ACM, 2011,
[11] Wu Jian-xin, Rehg J M. Where am I: Place instance and category recognition using spatial PACT[C].Computer Vision and Pattern Recognition, Anchorage, Alaska, 2008.
[12] Le Roux,Nicolas, Yoshua Bengio.Representationalpower of restricted boltzmann machines anddeep belief networks [J]. Neural Computation, 2008.
[13] Li Deng, Dong Yu, John Platt. Scalable stacking andlearning for building deep architectures [C].ICASSP. Kyoto, 2012.
[14] Bengio Y. Learning deep architectures for AI [J]. Foundations and trends in Machine Learning, 2009.
[15] Bengio Y.Learning deep architectures for AI [J]. Foundations and Trends in Machine Learning, 2010.