Study on neural networks in optical character recognition

WU Jin-Song

Beijing Institute of Economics and Management, Beijing,102602,China

Wjsong101@sina.com

Keywords: Optical character recognition; neural network

Abstract. Optical character recognition is already widely applied to industrial applications, and its recognition accuracy is also increasing with a variety of algorithms. This paper introduces three kinds of neural network structures named BP neural network, DBN network, CNN network. Through experimental comparison, analyze the performance of three networks on the optical character recognition.

Introduction

Optical character recognition (OCR) has been successfully applied to the industrial field, intelligent transportation, Information security, and so on. There are many methods about the optical character recognition presently, and they are the method based on template matching, structure analysis and neural network method, and so on [1-8].

Starting in the 80 s, there are many researches focuses on artificial neural network, and it is widely used in pattern recognition. The successful use of propagation algorithm (BP algorithm) proves that the neural network has powerful learning ability. But BP algorithm has its own disadvantages, such as easy to fall into local optimum, the limitation on the input data format, etc. Although in theory, the more complex neural network can fit the training model more accurate, but the problem that the complex neural network bring the increase of the computational complexity has been affecting the further development of the neural network

In 2006, Hinton [2] proposed a fast learning depth belief network (DBN) algorithm, and then the research of deep learning becomes more and more popular. At the same time, optical character recognition based on DBN also obtained a better recognition result than BP network. Patrice [3] brought the convolutional neural network (CNN) into optical character recognition, and got good results. With the development of the deep learning structure and the related algorithms, there are more and more algorithms about image recognition based on deep learning.

In this paper, three basic neural networks named BP, DBN have been studied, and the performance of these different structures is compared. The advantage, the disadvantage and the future direction are analyzed.

Back Propagation network

Back Propagation（BP） is a kind of multilayer feed forward network, and it is currently one of the most widely used neural network model [4]。

BP has the basic three layer perceptron structure, mainly can be divided into input layer, hidden layer and output layer. The forward calculation process of BP network is divided into information dissemination and error back propagation two process BP. Input layer neurons is responsible for receiving input information from the outside world, and transmitted to the middle layer neurons; hidden layer, namely information collection processing, can be seen as feature extraction, the structure can be designed for single hidden layer or multi-hidden layer structure. Output layer is the classification and decision layer, and it is to receive the characteristic information of the hidden layer transfer after processing, and the output is the classification results. When the actual output is different to the desired output, the algorithm should entered into the stage of error back propagation. The weight can be modified according to the gradient descent methods, and the process is from the

output layer to the hidden layer, and the last, to the input layer. Through constant iterative forward propagation and error propagation direction iteration process, the weight is modified, until it reaches an acceptable error range or the iterations times researched the predefined number. The network model is shown in figure 1.

The design of BP network is mainly concentrated in the network layer number, number of hidden layer neurons, the selection of initial weights and the selection of learning rate and expected error. And it has proved that it can approach any rational function when it has the deviation and at least one S hidden layer and a linear output layer of the network. Under normal conditions, the increase of the number of network layers and neurons can reduce the error of the network, but it is also increase the complexity of the network, and increase the calculation cost and the result may be trapped in local optimal value. So, the number of the neurons of hidden layer will be increased to improve the training accuracy.
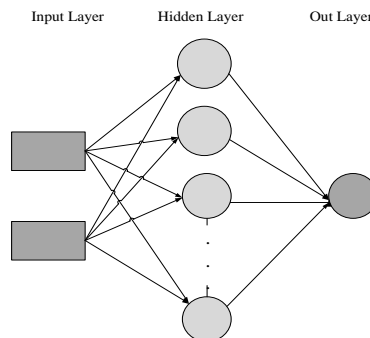

Fig.1 the basic structure of BP network

Although BP network has been widely used, but there are some defects and shortcomings. First of all, due to reverse the process of error propagation, mostly algorithms use the gradient descent method to make the weights convergence. However, the gradient descent method is easy to fall into local optimum, and we cannot get a global optimal value. Secondly, there is no theoretical basis of the network layer number of hidden layer and the number of units, they are often gotten based on experience or through repeated experiments. So, there are much redundancy, and the redundancy increase the burden of network learning.

DBN network

DBN[5] is proposed by Hinton using multilayer boltzmann unit into the depth of the neural network method. There is no link between nodes of each layer, the first layer is the input data layer, and it also is called visible layer. The second layer is the hidden layer (h). If all of the nodes are random binary variable node (0 or 1 value only), at the same time, assuming that the probability distribution p (v, h) meet the Boltzmann distribution, we call this model is Restricted Boltzmann Machine (RBM).

Classic DBN network structure is made up of several layers of RBM and a deep neural network constructed by BP neural network. RBM is a single random neural network, as shown in figure 2, it is essentially a probability graph model. Between input layer and hidden layer are all connected, but the inside layer without mutual connection between neurons. Each neuron has just two states, that is to say, the 1 represent activation, and 0 represent the inactivated. And the probability meet sigmoid function. There are many advantages of RBM, and the basic is that the layers are independent of each other. And it is can sampled randomly. The contrastive divergence algorithm make the computing of the weight more complexes, so there are other method to solve this problem, such as CD-n algorithm.
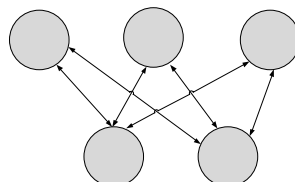

Fig. 2 the basic structure of RBM

Convolution Neural Network

Convolution neural network is a multilayer neural network, each layer is composed of multiple 2D plane, and each plane is composed of multiple independent neurons. Convolution is the similar to the multilayer perceptron neural network, can be divided into input layer, hidden layer and output layer. The input information of the convolution neural network is often 2D data, such as an image information. The hidden layer of each layer is divided into two parts, C and S, C layer can be regarded as a feature extraction. And the feature mapping can be gotten through a filter and a deviation. S layer can be gotten by down sampling, and then the weighted sum are calculated, at last, a sigmod function are used to get the result. After a preset layers, each layer S for the column vector of each neuron consolidation, as the input to the output layer into a normal neural network. The network learning process is adopted to feedback algorithm for training.
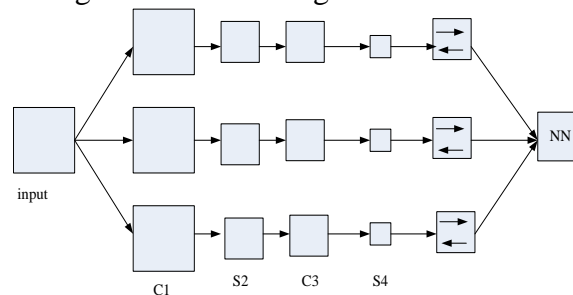

Fig .3 basic structure of convolution neural network

Experiment and Analysis

In order to verify the effectiveness and differences of these kinds of neural network structure in character recognition. Matlab is used to test the MNIST dataset, and the number of training dataset is 30000 and the testing sample training data is 5000.

Based on BP algorithm, using one hidden layer and two hidden layers, each layer with different number of hidden nodes training 20 times as a result, as shown in table 1. It can be seen from the experimental results that. The good performance is not related to the number of the layer. And the increase of the number of hidden layer can reduce network error, improve accuracy, but also make the network complex, thereby increasing the network training time and appeared the tendency of over fitting. It will also increase the risk of the nodes in a local optimum.

Table.1 result of different layer in based on BP

Hidden layer	neurons	Error rate
1	20	0.0728
1	100	0.0504
1	200	0.0440
1	250	0.0470
2	200	0.1226

Based on the algorithm of DBN, 2 hidden layer with 100 nodes is used to test data, and get the error rate of 0.0318, the error rate is 0.025 when the three layers with 100 nodes are used. We can get the conclusion from the two experimental results that DBN network overall result is better than BP network. The reason is that, DBN network has a preliminary training process, and the initial value of each layer node weight is decided through the RBM preliminary. However, the weight in the BP network is random, and it reduced the probability of in local optimum by using the gradient descent algorithm in the pro-training procedure.
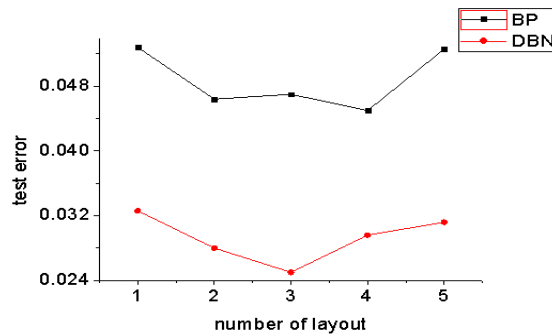
Fig. 4 the comparison results between BP and DBN

Experiments were also produced in BP and DBN network at the same time with the layer number from 1 to 5 hidden layer, each layer has 100 hidden nodes. The result is shown in figure 4. With the same test data, the error rate of DBN is markedly lower than the BP network. At the same time, the figure shows that is not in the same network layers, the more, the better the results of the test.

Algorithm based on CNN, choose 2 C layer, 2 S layer, and error rate of the test data of is 0.0134. CNN not only contains the characteristics of the global features, but also includes the local characteristics. Otherwise, the mapping of the process using the multiple features of training, has a better robustness. The network can use the 2D data of image directly, don't need to be converted into a column. And it can be more convenient application in image processing

Summary

There are many applications about optical character recognition based on neural network. However, for the off-line handwritten recognition, many deficiencies still exist in the process of practical application. With the development of deep learning theory, there are more and more new methods of neural network in image recognition, speech recognition, natural language processing, and this can also provide a new way to solve optical character problems. But, so far, there is not corresponding theory support of the number of layers and the number of neurons in the neural network layers, and this is future direction in academic.

References

[1] Rehman A, Saba T. Artificial Intelligence Review, 42(2014), p.253-273.

[2] Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh. Neural computation 18.7 (2006): 1527-1554.

[3] Simard, Patrice Y., Dave Steinkraus, and John C. Platt. null. IEEE, 2003.p.958

[4] Le Cun, B. Boser, et al. "Handwritten digit recognition with a back-propagation network." Advances in neural information processing systems. 1990.

[5] Hinton G E, Osindero S, Teh Y W.. Neural computation, 18(2006), p.1527-1554.

[6] Poultney C, Chopra S, Cun Y L. in:Advances in neural information processing systems. 2006: 1137-1144.

[7] AREL, ITAMAR, DEREK C. ROSE, THOMAS P. KARNOWSKI. Computational Intelligence Magazine, IEEE, 5(2010), p. 13-18.

[8] Larochelle H, Bengio Y. in: Proceedings of the 25th international conference on Machine learning. ACM, 2008: 536-543.