Character Recognition Algorithm Based on Neural Network

Qingqun Lan
Department of Urban Rail Transit and Information Engineering, Anhui Communications Vocational and Technical College, Hefei 230000, Anhui, China

Abstract
Character recognition is a topic with great research value and it includes all typical issues in the field of pattern recognition: data acquisition, processing and selection, selection of input sample expression as well as the guided training on the recognizer with sample set. Artificial neural network (ANN) is a mathematical model which simulates the behavior characteristics of neural network and performs distributed and parallel information processing. Along with the progress in neural network and logic techniques, people have used many new methods and means on the study of character recognition. This paper proposes a character recognition method based on improved BP neural network. This method extracts features of training samples according to certain methods, builds the classification analysis system, divides each character to be recognized into 8*8 blocks for digitalization and represents each block with a vector. The input vector is defined as an input vector matrix and the vector represents a certain character or number. Then, it conducts feature extraction on certain unknown sample to be detected with the same method and the system will judge its classification according to the sample feature extracted. The training of neural network shall train the arrays which represent characters or numbers respectively in the input end under supervision and correspond to the specific location in the output end. The test result on this character recognition method shows that this algorithm has relatively high recognition accuracy, a faster processing speed and strong robustness in the applications.

Key words: Character Recognition, BP Neural Network, Feature Extraction.

1. INTRODUCTION

Artificial neural network recognition is a non-linear adaptive information processing system consisted of numerous processing units. It is proposed on the basis of the research achievements of modern neuroscience and it tries to perform information processing by simulating the method the brain neural network processes and memorizes information. In the artificial neural network, the neuron processing units can represent different objects, e.g. numbers, letters, features and some meaningful abstract patterns(A. Namane and A. Guessoum et al., 2014). The connection weight between the neurons reflects the connection strength of the units whilst the information presentation and processing is shown in the connection relation of the network processing units. BP (Back Propagation) algorithm is an error back-propagation algorithm. The input pattern of its network is a pattern forward propagation process from the input layer to the output layer while the error signal between the expected output and the actual output is an error back propagation process from the output layer to the intermediate layer and then the input layer. During the network memory training process with alternately repeated pattern forward-propagation and error back-propagation, the network converges; in other words, it is the learning convergence process where the global error of the network converges to the minimum. With such strengths as parallel computation, distributed storage and strong classification capacity which surpass some conventional techniques, neural network recognition can not only accomplish the processing of a large number of complicated data in a quick and real-time manner, but it can also achieve satisfactory recognition result in the end(Najmeh and Hamid, 2015; Chacko and Dhanya, 2015).

Character recognition (CR) acquires the text and image information through such optical input methods as scanning and camera-shooting and analyzes the text form feature with various pattern recognition algorithms. Tausheck, a German scientist has come up with the concept of optical character recognition (OCR) in 1929 officially, however, it is not realized until electronic computer came into being. GNagy and A.L.spizt and others have put forward a character recognition strategy in their research on the printed documents with bad-quality recognition image. The process that the computer recognizes characters is similar to that of human beings(Xiangang and Yuning et al., 2015). Firstly, process the character image and extract the main features which can represent the characters. Then, form a corresponding relation between the feature and the character coding and store in the computer, which is called training. The recognition process is compare the input character image after processing with all the characters in the computer and find out the most similar character, which will be the final recognition result(Sheila and Jair et al., 2013). The statistical features of many characters have formed the template information which is about the character stereotype information and constitute the recognition characters after extraction, learning and classification. The template information is stored in the
recognition system. In the recognition of unknown image, firstly extract the same statistical features. Then, compare it with the character stereotype knowledge stored in the recognition system. Finally, determine the final character classification according to the comparison result and achieve the purpose of recognition. In order to recognize different characters, train all possible character samples on the recognizer, however, it is not an easy thing to complete such training work and obtain the expected result. So far, character recognition technique has been developed and used in the practical issues like voice, image, robot and artificial intelligence. The key to solve these issues is to perform complicated and huge real-time data processing. Neural network pattern recognition technology has made great progress in the real-time applications of pattern recognition and character recognition is a greatly significant field of such applications (Fikriye and Figen, 2012; Yifang, 2016).

This paper, firstly, introduces the development overview of applying neural network in the pattern recognition and summarized the advantages of neural network pattern recognition from the comparison with the traditional pattern recognition. Then, it introduces the binarization processing of character recognition and the multi-layer perceptron of neural network, brings in the most extensively-used BP neural network, proposes heuristic improvements for BP algorithm and applies in the character recognition with a purpose to reduce the oscillation in the learning process, reinforce the convergence, enhance the stability and improve the learning training speed and recognition accuracy. The test experiment proves that the method of this paper has better classification recognition effect.

2. BINARIZATION PROCESSING OF CHARACTER RECOGNITION

Image graying and binarization is aimed to change every pixel in the image into 0 or 255 to facilitate the computation. Meanwhile, it can also reduce some noises and then it cuts the image into the characters one by one, extracts the features of every character and generates the feature vector or feature matrix.

Binarization image refers to the image with only two gray levels. Binarization processing is to transform the digital image into the pixel matrix represented with 0 and 1. The key to binarization transform of character image is to identify the proper threshold to segment the character and the background and its result shall have excellent shape preservation without losing any useful shape information or generating extra vacancy. The key of binarization methods is the selection of the threshold \( T \), the value-selecting method of which is determined by the local features of the pixels or the global features. The threshold includes global threshold and local threshold. Threshold is the scale to separate the target and the background. To select the proper threshold is to preserve as much image information as possible and to reduce as much disturbance of background and noises as possible, which is the principle to select the threshold. Image binarization sets a certain threshold and divides the pixels of the gray image into two parts: the pixel cluster bigger than the threshold and the pixel cluster smaller than the threshold. For example, if the input gray image function is \( f(x,y) \) and the output binarization image function is \( g(x,y) \), then

\[
g(x, y) = \begin{cases} 
0 & f(x, y) < \text{Threshold} \\
255 & f(x, y) \geq \text{Threshold} 
\end{cases}
\] (1)

It can be seen from the above formula that the binarization processing of the image is to set the gray level of the points in the image as 0 or 1, that is to say to make the entire image into the black-and-white image. In other words, the binarized image which reflects the global and local features can also be obtained from the gray image with 128 brightness levels with proper threshold selection. On the basis of the binarization image, conduct further image processing and obtain some geometrical features or other features of the image (Olarik and Mahir et al., 2015; Szilárd and Yves et al., 2015). The following Fig.1 is the index color image binarization and Fig.2 is RGB image binarization.

![Figure 1. Index color image binarization](a) Index color original image  (b) Binarized image
3. DESIGN OF CHARACTER RECOGNITION CLASSIFIER

3.1. Multi-layer Perceptron Model of Neural Network

Neural network has four models: layered network, layered network with intra-layer connection, layered network with feedback connection and interconnection network.

(1) Forward network

The output of forward network is only determined by the current input and the weight matrix and it is not relevant to the previous output state of the network. The neurons in the network are arranged in layers and every neuron is only connected with the neurons of the previous layer. With neurons arranged in layers, the input layer, the intermediate layer (also called “hidden layer”, which can be composed of several layers) and the output layer are formed. There is no feedback link and delay is not taken into consideration in the mapping relation from the input to the output. Get the connection weight from learning and then finish the appointed task, which can only be used in the associative mapping and its classification. After the orderly propagation through various layers, the input pattern is finally output in the output layer (Cuihong and Ana et al., 2015). Both the perceptron network and BP network belong to forward network, the structure of which is shown as Fig.3.

(2) Forward network with feedback from output to input

The forward network with feedback has taken the delayed elements between the input and the output into consideration and the network system has several steady states. When the network starts to move from a certain initial state, the network system can always converge to a certain steady equilibrium state. Taking delay into account, feedback network is a dynamic system and its model is a dynamic equation (differential equation). In the optimization computation, confirm the weight not from learning, but using analysis algorithm on the target function. Firstly, set the initial state of the network. Then the system moves. If stable, it reaches a steady state in the end and the corresponding output is the solution to the optimization problem. In the meanwhile, it can also be used in the associative memory and to constrain the solution to the optimization problem (Junchuan and Xiao et al., 2012). Its structure is shown as Fig.4.

(3) Forward network with intra-layer connection

This network divides the intra-layered neurons into different groups and every group will move as a whole. In this way, the interaction between the neurons in the same layer is realized through the inter-constraint of the neurons in the layer. Through the mutual combination of the intra-layered neurons, the lateral inhibition or excitatory mechanism among the neurons in the same layer can be realized. From this perspective, it can restrict the neurons which move at the same time in every layer or divides the neurons into several groups and make every group computes as a whole (Somesh and Manu, 2012). Such interconnection in the same layer is aimed to
restrict the number of neurons with excitement or suppression at the same time in the same layer so as to complete the specific functions. Its structure is indicated as Fig.5.

Figure 4. Structure of forward network with feedback

Figure 5. Structure of forward network with mutual connection

(4) Mutually-connected network

In the mutually-connected network, every neuron may be mutually connected and all the neurons can be used as the input and the output. When the signal goes through a certain neuron in the forward network with no feedback, the process is over. Different from the network with no feedback, the signal in such network, propagates back and forth among the neurons repeatedly; in other words, the network is located in a constantly changing dynamic process. Starting from a certain initial state, the signal will reach certain equilibrium state after several changes. According to the network structure and the neuron characteristics, the operation of the network may enter into periodic oscillation or other chaotic equilibrium state (Kalyan and Rohit et al., 2015). The structure of the mutually-connected network is shown in Fig.6 and it can be seen obviously that there may be connection between the neurons.

Figure 6. Structure of mutually-connected network

3.2. BP Algorithm

BP neural network is short for error back-propagation neural network. It is made of an input layer, one or more hidden layers and an output layer and every layer contains several neurons. Basic BP algorithm includes two aspects: signal forward propagation and error back propagation. In other words, the calculation of the actual output goes from the input to the output while the modification of weight and threshold proceeds from the output to the input. The neurons in every layer are only connected with the neighborhood neurons and there is no connection between the neurons in the same layer. There is no feedback connection between the layers of each layer and they form the feedback neural network system with a hierarchical structure. The calculation of the layered feedforward neural network can only solve the linearly separable problems while the network to
solve non-linear problems must be the multi-layered neural network with hidden layer(s). The information of the above network propagates from the input unit to the hidden unit and finally the output unit. Such forward network with hidden layer(s) has an important feature that the units in the hidden layer(s) can randomly form their own input presentation and the weights between the units of the input layer and the hidden layer determines whether each unit in the hidden layer is activated. BP algorithm is slow in learning and the predictive ability training capability (learning capability) of its network is contradictory (Parshuram and Ravinda, 2015; Amit and Rahul et al., 2013). The structure of BP network is shown as Fig. 7. Here, \( n \) is the number of nodes in the input layer, \( q \) the number of nodes in the hidden layer and \( m \) the number of nodes in the output layer. Every connection corresponds to a weight. The modification of these weight learning or training network can modify the network function mapped from the input to the output.

In Fig. 7, the input vector is \( X = (x_1, x_2, \ldots, x_n)^T \), the output vector of the hidden layer is \( Y = (y_1, y_2, \ldots, y_j, \ldots, y_m)^T \), the output vector of the output layer is \( O = (o_1, o_2, \ldots, o_k)^T \), the expected output vector is \( D = (d_1, d_2, \ldots, d_j, \ldots, d_l)^T \) and the weight matrix from the input layer to the hidden layer is represented with \( V \). \( V = (v_1, v_2, \ldots, v_j, \ldots, v_m) \). Here, column vector \( v_j \) represents the corresponding weight vector to the \( j \)th neuron in the hidden layer. The weight matrix between the hidden layer and the output layer is \( W \) and \( W = (W_1, W_2, \ldots, W_k, \ldots, W_l) \). Here, the column vector \( W_j \) refers to the corresponding weight vector to the \( k \)th neuron in the output layer.

### 3.3. Feature Extraction and BP Network Recognition Procedures

1. **Initialization.** Select three-layered structure network with reasonable structure, set all the adjustable parameters such as weight and threshold, give random small value to all weights and set the initial value for the threshold.

2. **Provide training data set, namely provide the input vector \( x \) and the expected output \( y \) and perform forward calculation. Provide output for the \( j \)th node in the hidden layer according to Formula (2).

   \[
   o_j(t) = f(g_j(t)) = \frac{1}{1 + \exp(-g_j(t))}
   \] (2)

   Provide output for the \( k \)th node in the output layer according to Formula (3) and the error signal of this node is

   \[
   e_k(t) = y_k^d(t) - y_k(t)
   \] (3)

3. **Reverse calculation of the local gradient information of the node.** For the nodes of the output layer,

   \[
   \delta_k(t) = e_k(t)
   \] (4)

   For the nodes in the hidden layer,
\[ \delta_j(t) = o_j(t)[1-o_j(t)] \sum_{k=1}^{n_{out}} \delta_k(t)w_{jk}(t) \]  

(5)

4. Adjust the weight and modify the weight from the output node back to the hidden layer according to the following formula according to the error back propagation direction.

\[ W_j(t+1) = W_j(t) + \mu \delta_j y_j \]  

(6)

Here, \( \mu \) is the gain bigger than 0, \( \delta_j \) is the error of node \( j \) and \( \delta_j \) is calculated with the following formula according to its different forms.

\[ \delta_j = y_j (1-y_j) (y_j - y_{j^+}) \quad \text{as } j \text{ is the output node} \]

\[ \delta_j = y_j (1-y_j) \sum_k \delta_k W_{jk} \quad \text{as } j \text{ is the hidden node} \]  

(7)

5. Calculate the actual output \( y \).

\[ y_j = f \left( \sum_i W_{ij}x_i \right) \]  

(8)

Here, \( f \) function is Sigmoid function.

6. Return and repeat Step (2). Input new samples until the fitness function value reaches the preset requirement or the error meets the requirements.

4. TEST RESULTS OF FEATURE EXTRACTION AND RECOGNITION

Starting from the processing time, the classification accuracy and the prepared training samples, it is a practical processing method to determine the number of neuron nodes through multiple experiments with the above-mentioned method. As the neurons in the hidden layer can be randomly adjusted, the BP neural network of the single-hidden layer can approximate any random non-linear mapping. There is only 1 neuron in both the input layer and the output layer. And there is also only 1 hidden layer, the neurons in which can be selected within the range of 9-16 according to the above-said design experiment formula and practical situation. The BP neural network with variable neurons in the hidden layer below can determine the number of hidden layers through error and training steps and test the impact the number of neurons in the hidden layer has on the network performance. Character recognition extracts the features of training samples according to certain rules. Build the classification analysis system, conduct feature extraction on a certain unknown sample to be tested with the same method and the system will judge which classification it belongs to according to the extracted sample feature. This paper has combined the 64-dimensional outer feature vector with the width of 8 and the height of 8 in the character normalization and the 64-dimensional penetration feature vector to get the 128-dimensional character feature vector. The learning process is to introduce the feature vectors into the neural network classifier for training in order to determine the weight and then the object to be recognized is introduced into the well-training network to get the recognition result.

Training data is the pixel points of every image and every training sample is a 8×8 matrix, in this way, the array needs to store the data points in \( n \) columns and 64 rows and the corresponding target value is the numerical value. As for the sample numbers, predict with BP neural network training and recognize the figures to be tested. These figures may not be neat, so it seems quite difficult to directly read and recognize the samples. The method adopted is to perform neural network training on such font figures, recognize the read-in test character image with the well-trained neural network, add or not add noises to the tested character image and conduct classification and recognition with BP neural network.

Conduct the experiment with the method of this paper in three different moments and adjusted neural network parameters and the BP network recognition result and analysis are shown in Fig.8-Fig.14 and Tab.1-Tab.3.
Figure 8. The characters to be recognized in the image

ABCDEFabcdef123450

Sakkal Majalla

Figure 9. The results of image character recognition

Get it easy

Happy new year

Add user defined phrase

Figure 10. The characters to be recognized in the PDF file

Get it easy

Happy new year

Add user defined phrase

Figure 11. The results of PDF file character recognition

Figure 12. Performance of BP network
Figure 13. Training state of BP network

Figure 14. Regression of BP network

Table 1. Number of neurons, training error & learning steps in the hidden layer at certain moment I

<table>
<thead>
<tr>
<th>Number of Neurons</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Training Error</td>
<td>0.1572</td>
<td>3.2923</td>
<td>4.5671</td>
<td>0.0868</td>
<td>0.1364</td>
<td>0.0835</td>
<td>0.0878</td>
<td>0.0862</td>
</tr>
</tbody>
</table>

Table 2. Number of neurons, training error & learning steps in the hidden layer at certain moment II

<table>
<thead>
<tr>
<th>Number of Neurons</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Training Error</td>
<td>0.0858</td>
<td>0.1166</td>
<td>0.0903</td>
<td>0.1027</td>
<td>0.0885</td>
<td>0.0868</td>
<td>0.0872</td>
<td>4.9133</td>
</tr>
<tr>
<td>Learning Step</td>
<td>1238</td>
<td>1164</td>
<td>2000</td>
<td>2000</td>
<td>373</td>
<td>225</td>
<td>92</td>
<td>721</td>
</tr>
</tbody>
</table>

Table 3. Number of neurons, training error & learning steps in the hidden layer at certain moment III

<table>
<thead>
<tr>
<th>Number of Neurons</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Training Error</td>
<td>0.0892</td>
<td>0.0870</td>
<td>0.0893</td>
<td>7.833</td>
<td>0.0889</td>
<td>0.0915</td>
<td>0.0897</td>
<td>0.0848</td>
</tr>
<tr>
<td>Learning Step</td>
<td>639</td>
<td>262</td>
<td>71</td>
<td>1326</td>
<td>2000</td>
<td>425</td>
<td>55</td>
<td>67</td>
</tr>
</tbody>
</table>
It is not good if there are too many or too few nodes in the hidden layer. In the network learning, it is the best if there are about 10 training samples in the connection of every neuron or it is proper to have two to three times of the neurons in the output layer. The former can set the number of neurons in the hidden layer according to the samples which can be prepared while the latter according to the number of classifications. Therefore, 15 neurons in the hidden layer is the perfect choice, which will have small error, short learning steps and steady performance. Besides, it can also be seen that not more neurons in the hidden layer will lead to better network performance.

5. CONCLUSIONS

Artificial neural network has provided new methods for character recognition. It is the self-organizing self-learning ability, generalization ability and highly-parallel non-linear and operation ability of neural network that makes pattern recognition the most successful application field of the current neural network. The neural network character recognition method of this paper has normalized character into the corresponding character feature vector, introduced the feature vector into BP neural network classifier for training to identify the weight (namely the learning process) and then put the object to be recognized into the well-trained network to get the recognition result. This method can improve the character recognition efficiency.

REFERENCES


