Recognition of mnist handwritten digits database based on svm and bp neural network

Zhang Zheming

1. Introduction

Arabic numerals are one of the symbols of human civilization and the main medium for people to communicate and communicate. As the information carrier of world economic development, digital shows the way of thinking and cognition of people of all ethnic groups in the world. In our daily life, we have to do a lot of document processing every day, such as the processing of documents and forms, the processing of tax bills, bank checks, remittance slips, credit card bills, and the sorting of post office letters, etc. How to use computer characters Recognition and document processing technology to free people from these heavy manual labor has become an urgent problem to be solved.

Computational text recognition, namely Optical Character Recognition (OCR), is a bridge for information communication between computers and people. Computer character recognition has always been a very active research topic in the field of image recognition. It is considered to be a typical image recognition application research topic, which includes the basic theory of pattern recognition and the application of many typical techniques. With the advent of the information age, how to realize the automatic identification and processing of digital identification will be directly related to the development of my country's information industry. Digit recognition is a small category of pattern recognition problem, which has great research value both theoretically and technically. Among them, digital recognition, especially handwritten digital recognition, is regarded as the "most difficult field" in the field of recognition because it requires high recognition accuracy and fast recognition speed, and has become a hot research topic at home and abroad.

Digital recognition is a branch of pattern recognition, involving pattern recognition, image processing, digital signal processing, natural language understanding, artificial intelligence, fuzzy mathematics, computer and other disciplines. In the handwritten digit recognition system, there are printed and handwritten characters, so digital recognition can be divided into two categories: printed digit recognition and handwritten digit recognition. With the development of computer technology and network technology, a large amount of digital information needs to be collected into the computer in a certain way before being processed, such as the zip code on the letter, the digital information on various bank notes, the admission ticket number on the test paper, The entry of handwritten numbers such as handwritten express orders in the logistics industry

can be recognized by handwritten number recognition technology, which improves efficiency and saves manpower. Handwritten digit recognition integrates the knowledge of image processing, pattern recognition, machine learning and other fields. It is a complex interdisciplinary problem. The recognition system usually consists of three parts: image preprocessing, feature extraction and classification and recognition.

2. Related work

In 1929, German scientist Taushek used optical template matching recognition, creating a precedent for using machines to recognize text symbols. In order to input a lot of materials into computers for information processing, European and American countries have started the study of Western Optical Character Recognition (OCR) since the 1950s. The meaning of OCR has evolved to use optical technology to scan text and characters. computer code. From the 1960s to the 1970s, countries around the world started the research on OCR one after another. In the early stage of the research, most of the researches focused on the method of character recognition, and the recognized characters were only numbers from 0 to 9. mM company first developed OCR products, in 1965 in the New York World's Fair exhibited the company's OCR product - mMI287. This product at that time could only recognize printed numbers, English letters and some symbols, and it had to be in the specified font. In the late 1960s, Hitachi and Fujitsu also developed their own OCR products. The world's first automatic letter sorting system to realize handwritten zip code recognition was developed by Japan's Toshiba Corporation, and NEC Corporation also launched the same system two years later. In 1974, the automatic sorting rate of letters in Japan reached about 92%. China's research work on OCR technology started relatively late. It only began to study the recognition technology of numbers, English letters and symbols in the 1970s, and began to study Chinese character recognition in the late 1970s. In 1986, the National 863 Program in the field of information organized Tsinghua University, Beijing Institute of Information Engineering, and Shenyang Institute of Automation to jointly develop Chinese OCR software. In 1989, Tsinghua University took the lead in launching the first set of Chinese OCR software in China. So far, China's OCR technology has officially moved from the laboratory to the market. In recent years, the domestic research on digital recognition is still quite in-depth, and great achievements have been made, so that the recognition rate of the system continues to rise. At present, the recognition rate of printed numbers has reached more than 99%, and the recognition rate of online handwritten numbers has reached 99%. In principle, handwritten digit recognition methods can be roughly divided into three categories: classification based on statistical features, classification based on structural features, and neural network-based algorithms. Although there have been many years of research, with the intersection of computer science,

psychology and artificial intelligence. With the integrated development of disciplines and the drive of market demand, handwriting recognition technology will achieve greater prospects.

3. Steps and methods for recognizing handwritten digits

3.1 Introduction to Mnist

The Mnist-image handwritten digital image recognition database is a handwritten database jointly established by Yann LeCun, Courant Institute of New York University, Corinna Cortes, Google Labs, Christopher JC Burges of New York, Microsoft Research, and Raymond. It contains a training set of 60,000 examples and a test set of 10,000 examples. It is a subset of the larger set provided by NIST. Figures are size-normalized and centered around a fixed-size image. It is a good database for those who wish to experiment with learning techniques and pattern recognition methods on real data while spending minimal preprocessing and formatting.

3.2 Step for recognizing handwritten digits

Usually, handwriting that can be exploited by a single feature extraction method

Handwritten digit recognition integrates the knowledge of image processing, pattern recognition, machine learning and other fields, and is a complex interdisciplinary problem.

(1) Image preprocessing

The first step in handwritten digit recognition is image preprocessing. Usually, the handwritten digital images to be recognized often bring some noise during the scanning process, and the digital images obtained with different scanning resolutions have different qualities. In addition, it is also necessary to correctly segment the handwritten numbers in the entire document image, and the size and font of the segmented numbers are often different, so normalization is required. Image preprocessing plays a very important role in the whole handwritten digit recognition. If the preprocessing is done well, the part that reflects the essential characteristics of the text can be retained, which makes the later recognition relatively easy, the recognition accuracy rate is high, and the recognition speed is fast.

(2) Feature extraction

The purpose of feature extraction is to extract the essential attributes used to distinguish other digital categories from digital images after image preprocessing, and quantify them to form feature vectors. Common handwritten digital features are: statistical features and structural features.

Statistical features are features extracted from digital image lattices after mathematical transformation on the basis of binary or gray value lattice images. Statistical features are often used in conjunction with neural network classifiers or statistical matching methods.

Structural features mainly describe the geometric structure of handwritten numerals, and focus on embodying the essential characteristics of numeral structures. This method has a strong ability to distinguish similar words, and can obtain classification results with high recognition accuracy, but it is easily disturbed by factors such as noise.

(3) Classification and identification

Classification and recognition is to use the trained classifier to classify and recognize the handwritten digits after feature extraction. The recognition principle of the classifier is to define a set of surfaces or hypersurfaces on the feature space through its topological structure and built-in parameters, and use this set of surfaces or hypersurfaces to divide the feature space into different regions, so as to achieve the purpose of classification and recognition.

Artificial neural network classifier: It is a data-driven nonlinear model constructed according to the principle of human brain organization and activity. It is composed of several basic elements such as neuron structure model, network connection model, network learning algorithm, etc. It is a system with some intelligent functions. The artificial neural network classifier has the characteristics of self-adaptation, generalization and nonlinear mapping, but the training time is often too long, and the convergence speed and generalization ability also need to be further studied. Common artificial neural network classifiers include RBF network and BP network.

The support vector machine classifier is a classifier that improves the classification performance by maximizing the classification interval or edge according to the principle of structural risk minimization. The SVM classifier selects and trains a set of feature subsets called support vectors, so that the linear division of the support vector set is equivalent to the division of the entire data set, which reduces the computational complexity and ensures the accuracy of classification and recognition. Since the SVM classifier has a strong ability of small-scale subdivision, the SVM classifier is used to automatically recognize handwritten digits, and an improvement is made. The separability strength of each category in the feature space is used to determine the optimal kernel parameters. The training process is accelerated, the recognition speed of handwritten digits is improved, and a good recognition accuracy is ensured.

3.3BP neural network

3.3.1BP neural network principle and model

The essence of the neural network is a mapping function from a given input x to an output function y. The coefficients of the function are the parameters θ that we want to get through network training, and the coefficients of the function are determined by the neural network learning algorithm. When x is input, a corresponding y is calculated and output by the neural network. As for whether the result y of the training output satisfies our expected results, this is what we

need to do by improving the performance of the model and optimizing the learning algorithm.

3.3.2Principles of Neural Networks.

BP is a feedback learning network. The learning process of the algorithm consists of two parts, the first is the forward propagation of information, and the second is the back propagation through the error. The data information is input into the neural network through the input layer of the neural network, and each unit of the input layer transmits the data to each neuron of the hidden layer for internal processing of the data. The hidden layer is also called the middle layer, and its function is mainly for information Processing and exchange, its structure can be a single-layer structure or a multi-layer structure. The more layers, the higher the complexity of the neural network. The hidden layer transmits the internally processed data information to each unit of the output layer. The processed data information is output in the output layer, thus completing a training of forward propagation [5]. Backpropagation is a process in which God compares the output value of the network with the expected value, calculates the error, and uses the error to solve the problem layer by layer.

The model of the BP neural network is that many logical units are organized according to different levels. First, the calculation is performed layer by layer according to the forward propagation. After the output of each layer is weighted and summed, it is used as the input variable of the next layer. A three-layer neural network. The data that needs to be processed and calculated is input in the first layer of the neural network, that is, the input layer, and the prediction result after the training data is output by the output layer. The result output by the output layer is then compared with the expected value.

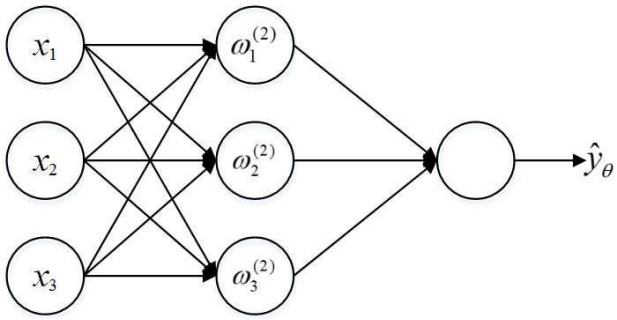


Figure 1 Schematic diagram of neural network topology

The BP algorithm has the advantages of simplicity, easy execution, and small computational complexity in use, and is a relatively common and relatively mature algorithm for neural network training. This paper uses BP neural network to realize the recognition of handwritten digits, introduces the specific implementation steps, and realizes it through matlab programming. After the experiment, by adding Dropout, the accuracy of the neural network training results is improved to a certain extent. The complexity of the neural network is not high, so the optimization effect is not very obvious. Due to the BP neural network, there are

shortcomings such as low learning efficiency, local optimal points may appear during training, and improper parameter selection is prone to overfitting and slow convergence. The author will continue to strengthen the related research of machine learning and deep learning in the future study, learn more advantageous neural network algorithms, study convolutional neural networks, and use some algorithms commonly used in machine learning, such as Bayesian algorithm, fuzzy reasoning Algorithm, deep belief network is integrated into it, which further improves the accuracy of neural network.

3.4 SVM

Statistical learning theory is based on a solid theoretical foundation, which provides a unified framework for solving the problem of limited sample learning. It can incorporate many existing methods into it, and is expected to solve many problems that were difficult to solve before, such as neural network structure selection problem, local minimum point problem, etc. At the same time, on the basis of this theory, a new general learning method - support Vector machine (referred to as SVM), has initially shown a lot of performance better than existing methods. Some scholars believe that it is becoming a new research hotspot after the neural network research, and will promote the significant development of machine learning theory and technology.

The SVM method is based on the dimension theory of statistical learning theory and the principle of structural risk minimization. According to the limited sample information, the complexity of the model, that is, the learning accuracy and learning ability of a specific training sample, that is, the ability to correctly identify any sample. In order to obtain the best promotion ability, it is necessary to seek the best compromise between them.

The basic idea of SVM can be summarized as firstly transform the input space into a high-dimensional space through nonlinear transformation, and then find the optimal linear classification surface in this new space, and this nonlinear transformation is achieved by defining an appropriate inner product function realized.

The classification function obtained by SVM is similar in form to a neural network, and its output is a linear combination of several intermediate layer nodes, and each intermediate layer node corresponds to the inner product of the input sample and a support vector machine, so it is also called support vector network. Since the final decision function actually only includes the inner product and the summation of the support vector, the computational complexity of recognition depends on the number of support vectors.

Another key point is that since the dimension of the transformation space may be very high, the VC dimension of the linear discriminant function in this space may also be very large, which will lead to the unsatisfactory effect of the classifier. However, as long as a smaller VC dimension can be constructed in a high-

dimensional space, a better generalization ability can be obtained.

Further, regarding the generalization ability of the optimal classification surface and the generalized optimal classification surface, the following conclusions are drawn:

If a set of training samples can be separated by an optimal classification surface or a generalized optimal classification surface, the expected upper bound on the classification error rate for the test samples is the ratio of the average support vector in the training samples to the total number of training samples

The generalizability of SVM is also independent of the dimension of the transformation space. As long as an inner product definition can be properly selected and an optimal or generalized optimal classifier with relatively few support vectors is constructed, it can be well generalized. sex.

Here, statistical learning theory uses a completely different idea from the traditional method, that is, instead of trying to reduce the dimension of the original input space, that is, feature selection and feature transformation, as in the traditional method, it tries to increase the dimension of the input space in order to achieve higher dimensionality. In the dimensional space, the problem becomes linearly separable or nearly linearly separable because the increase in dimension only changes the inner product operation, and does not increase the complexity of the algorithm as the number of dimensions increases, and the generalization ability in high-dimensional space does not increase. It is not affected by the dimension, so this method is feasible.

4 Conclusion

The SVM is applied to the recognition of handwritten digits, and good results are obtained. Experiments show that SVM can also have a higher recognition rate in the case of fewer samples. Compared with BP neural network, SVM has the advantages of simple algorithm, no prior knowledge, easy control, easy training, and good stability. The hidden layer design of BP neural network depends on the experience of the experimenter and is unstable. Although SVM can better solve the recognition problem in the case of small samples, it is still beneficial to improve the recognition rate if more training samples can be collected. All the papers have carried out feature extraction on the experimental samples, but the effect of feature extraction on improving the recognition rate is very limited through the experiments in this paper. Therefore, this paper does not carry out feature extraction. In addition, the research in this paper is mainly aimed at the recognition rate of SVM. For the speed and speed of SVM training and recognition, how to reduce the recognition time while ensuring the high recognition rate of SVM is the focus of future research.

References

[1] Vapnik V N. Statistical Learning Theory [M]. Xu Jianhua, Zhang Xuegong, Translator. Beijing: Electronic Industry Press, 2004.
[2] Astorino A, Gorgone E, Gaudioso M, et al. Data preprocessing in semi - supervised SVM classification [J]. Optimization, 2011, 60(1-2):143-151.
[3] Li Wenchi. Application of SVM in Handwritten Digit Recognition [J]. Journal of Quanzhou Normal University (Natural Science), 2010, 28(4): 18-21.
[4] Tian Yiran. Research and implementation of handwritten digit recognition technology [D]. Changchun: Jilin University, 2015.
[5] Zhou Zheng. A review of the development status of BP neural network [J]. Shanxi Electronic Technology, 2008 (02): 90-92.
[6] Xing Yuan. The application of deep learning in handwritten digit recognition [D]. Suzhou: Soochow University, 2017.
[7] Xiao Jian. OCR Character Recognition Based on Learning [J]. Computer Age, 2018 (07): 48-51

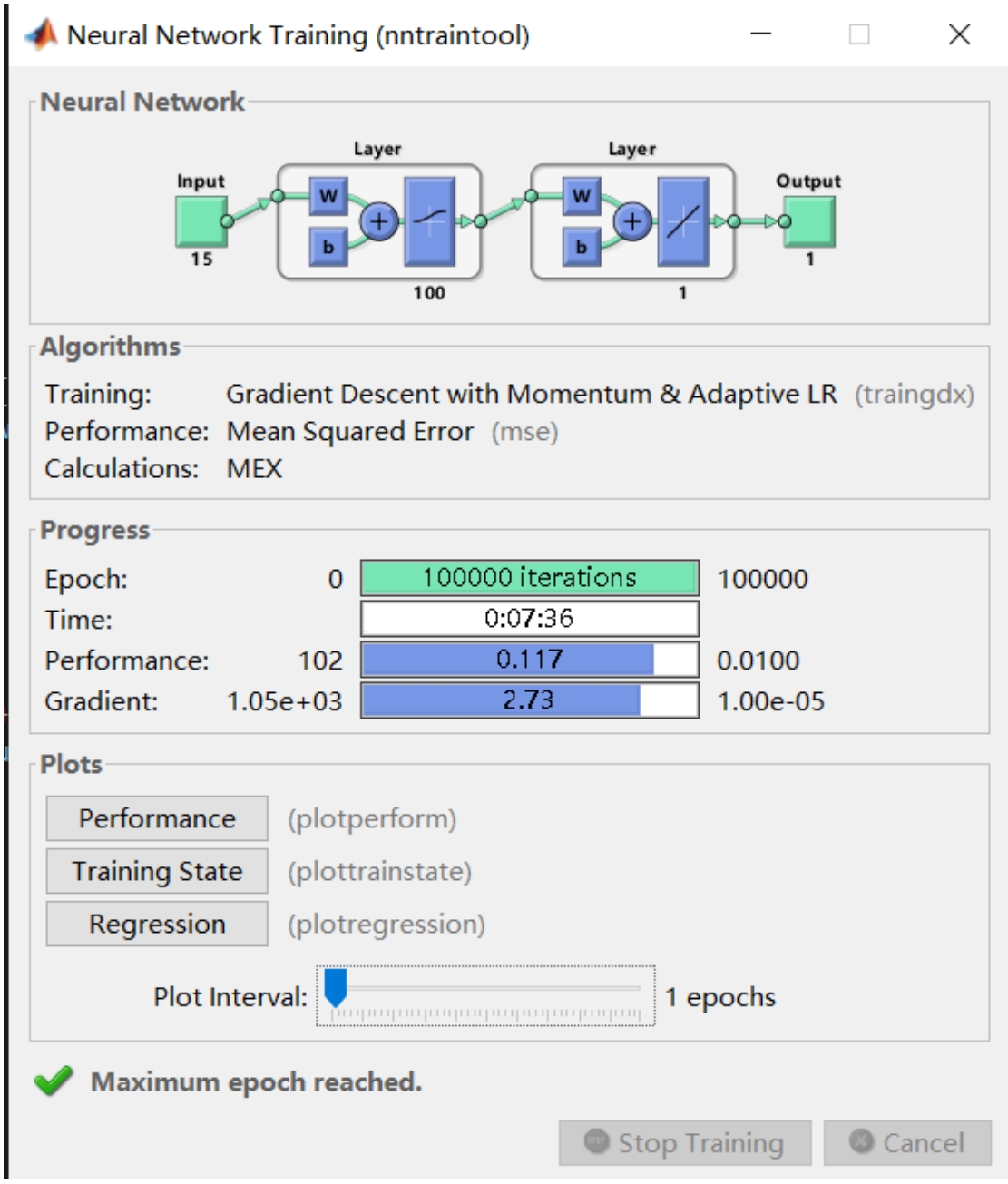
Appendices

BP neural network

```
clear all;
'LOADING.....'
kk_max = 10;
iii_min=1;
jjj_min=1;
iii_max = 10;
jjj_max = 10;
p=[];
count=1;
for kk = 0:kk_max-1
    m=strcat('appr_',int2str(kk),'.bmp');
    bw=imread(m,'bmp');
    for iii = iii_min-1:iii_max-1
        for jjj = jjj_min-1:jjj_max-1
            bww = bw(iii*64+1:iii*64+64,jjj*64+1:jjj*64+64);
            [feature,pic,flag]=getFeature(bww);
            %if flag==1
            %if sum(feature)~=0
                p(:,count)=feature;
                t(count)=kk;%+10*flag;
                count=count+1;
        end
    end
end

'LOAD OK.'
t;
save E52PT p t;

clear all;
load E52PT p t;

for i=1:15
pr(i,1)=min(p(i,:));
pr(i,2)=max(p(i,:));
end;
net=newff(pr,[100 1],{'logsig' 'purelin'}, 'traingdx', 'learngdm');
net.trainParam.epochs=100000;
net.trainParam.goal=0.01;
net.trainParam.show=1000;
net.trainParam.lr=0.05;
```

```
net=train(net,p,t);
'TRAIN OK.'

save E52net net;

for times=0:-1
    clear all;
    load E52net net;
    zzz=input('FileNo:', 's');
    xxx=str2num(input('LineNo:', 's'));
    yyy=str2num(input('colmNo:', 's'));
    test=strcat('appr_',zzz,'.bmp');
    x=imread(test,'bmp');
    bw=x((xxx-1)*64+1:(xxx-1)*64+64,(yyy-1)*64+1:(yyy-1)*64+64);

    [pp,bwp]=getFeature(bw);

    [a,Pf,Af]=sim(net,pp);
    imshow(bwp);
    a=round(a)
end

clear all;
kk_max = 10;
iii_min=1;
jjj_min=1;
iii_max = 5;
jjj_max = 5;
for kk = 0:kk_max-1

    load E52net net;
    m=strcat('appr_',int2str(kk),'.bmp');
    bw=imread(m,'bmp');

    for iii = iii_min-1:iii_max-1
        for jjj = jjj_min-1:jjj_max-1
            bww = bw(iii*64+1:iii*64+64,jjj*64+1:jjj*64+64);
            [pp,bwp]=getFeature(bww);
            [a,Pf,Af]=sim(net,pp);
            %imshow(bwp);
            a=round(a);
            b(iii*jjj_max+jjj+1,kk+1)=mod(a,10);
        end
    end
```

end


Neural Network Training (nntraintool)
Neural Network
Input
15
Layer
W
b
100
Layer
W
b
1
Output
1
Algorithms
Training: Gradient Descent with Momentum & Adaptive LR (traingdx)
Performance: Mean Squared Error (mse)
Calculations: MEX
Progress
Epoch: 0 100000 iterations 100000
Time: 0:07:36
Performance: 102 0.117 0.0100
Gradient: 1.05e+03 2.73 1.00e-05
Plots
Performance (plotperform)
Training State (plottrainstate)
Regression (plotregression)
Plot Interval: 1 epochs
Maximum epoch reached.
Stop Training
Cancel

SVM

```
clear;clc;
tr_feats = loadMNISTImages('train-images');
tr_label = loadMNISTLabels('train-labels');
te_feats = loadMNISTImages('t10k-images');
te_label = loadMNISTLabels('t10k-labels');

svmstr = sprintf('-t 1 -d 5 -r 1 -g 3 -c 2 -b 1');
model = fitcsvm(tr_label,tr_feats,svmstr);


classid = unique(te_label);
[predictlabel,accuracy,dec_values] = svmpredict(te_label,te_feats,model,'-b 1');
t_label = sort(te_label);
te_count = hist(t_label,unique(t_label));
pre_count = te_count;
for i = 1:size(te_feats,1)
    if te_label(i) ~= predictlabel(i)
        pre_count(te_label(i)+1) = pre_count(te_label(i)+1)-1;
    end
 end
 for i = 1:length(classid)
     fprintf('Classifying for Digit :   %i \n', classid(i));
     fprintf('accuracy on %i : %.2f%% (%i/%i)
\n',classid(i),pre_count(i)*100/te_count(i),pre_count(i),te_count(i));
 end
```

```
te_count = hist(t_label,unique(t_label));%每个类各有多少样本
pre_count = te_count;%每个类正确判断了多少个样本
for i = 1:size(te_feats,1)
    if te_label(i) ~= predictlabel(i)
        pre_count(te_label(i)+1) = pre_count(te_label(i)+1)-1;
    end
end
for i = 1:length(classid)
    fprintf('Classifying for Digit :  %i \n', classid(i));
    fprintf('accuracy on %i : %.2f%% (%i/%i) \n',classid(i),pre_count(i)*100/te_count(i
end
```

命令行窗口

```
accuracy on 7 : 96.21% (989/1028)
Classifying for Digit :  8
accuracy on 8 : 96.41% (939/974)
Classifying for Digit :  9
accuracy on 9 : 95.24% (961/1009)
>>
```