

# CSF images fast recognition model based on improved convolutional Neural Network

WENMING HUANG<sup>1, a</sup>, JINQIANG LENG<sup>2, b</sup> and zhenrong deng<sup>3, c</sup>

<sup>1,2,3</sup>Guangxi Key Laboratory of Image

Guilin University of Electronic Technology

Guilin, 541000 P.R. China

E-mail: <sup>a</sup>wmhuang@guet.edu.cn, <sup>b</sup>lengjqiang@126.com, <sup>c</sup>zhrdeng@guet.edu.cn

**Keywords:** convolution neural network; cerebrospinal fluid; rectie function; identification; classification of linear support vector machine.

**Abstract.** The sparseness of feature is an important characteristic determining feature, which directly affects the accuracy of image recognition<sup>[1]</sup>. By studying the traditional convolution neural network, we find that the learning of image features of cerebrospinal fluid cell easily overfits, but using rectie activation function instead of sigmoid activation functions, the features extracted are more sparse and have faster convergence rate in the process of training. Then features extracted are classified through a linear support vector machine. The experiments show that the improved model can enhance significantly the image recognition efficiency of cerebrospinal fluid, where two, three, four categories are respectively increased by 9.78%, 6.53%, 11.69%, and the average recognition time of a single image is also reduced 0.32s.

## 1 INTRODUCTION

Cerebrospinal Fluid [2], CSF, is a kind of liquid with colorless, lucid and glutinous, but the cell's topological structure is chaotic. If cut the image which the object is independent and single. It is very difficult to guarantee the precision and the completeness of segmentation image, so for some quantitative characteristics of the definition, the computer cannot distinguish accurately, thus the result of training the classifier and predicting the result will be inaccurate.

Convolution neural network algorithm in the field of graphic image recognition in recent years has obtained very result, this method is based on animal visual system development. In the field of domestic scholars on the basis of modern immunology and cytology, the Cerebrospinal Fluid cell function and morphological features, the combination of this paper proposes a new classification algorithm, non-pathological and abnormal Cerebrospinal Fluid cell sorting. Risk by analyzing the emergence of different types of cell and cell classification as a result, you can very well composite type of diagnostic cytology. this concept in the central nervous system disease clinical research and disease diagnosis is of great significant.

Bengio [3] explains the influence of data sparsity on algorithm when learning algorithm deeply. Most of images are converted to the vectors, which make any changes that are likely to lead to significant increase in the amount of calculation about data high-density. Conversely, if the data is sparse, the characteristic of high-robustness will constrain the trace change of input. The experimental results show the improved convolutional neural network make the characteristic of extractive cerebrospinal fluid, which reduce the computational complexity.

In the first section we introduce the related research work, the second section mainly introduce the network structure, the training process, the classifier's recognition process. And in section third present the experimental results and discussion.

## 2 RELATED WORK

Hubel and Wiesel [4] took the lead to study visual cortical cells of cats and found the concept of receptive field. In 1984 the Japanese scholar Fukushima presented neurocognitive machine on the

basis of receptive field. It is a variant of the neural network can be regarded as the first implementation of a neural network. Van Ooyen and Niehuis introduced a new parameter to the improve recognizability of neurocognitive machine. However, as a inhibiting signal, the parameter prevents neuro from exciting repetitive characteristics of incentives. Most neural network train information in weights memory. According to Hebel learning rule, the more times the characterized training are, the more likely to be checking in later identification process. There is significant breakthrough in machine learning of Neural Networks, which is the research focus in recent years.

Convolution neural network[5] has become an important part of the current image recognition and speech analysis research in the field. It is a hot topic in artificial neural networks. One of its advantages is the weight share, which reduces the number of parameters thus reducing the complexity of network structure, especially when the input image data is more complicated. Convolution neural network is a multi-layered perceptron, this network has strong robustness for image zoom, pan and tilt deformation.

Support Vector Machine (SVM) [6]was presented by Cortes and Vapnik et al in 1995.SVM started off as a model that did binary classification for linearly separable problems. Interval maximization is its basic learning strategy.SVM performs well in multi-classification problems. Multi-classification strategy mainly consist of one-to-the-rest, one-to-one, and directed-acyclic-graph (DAG SVM).This article takes the one-to-the-rest strategy.

### 3 MODEL DETAILS

#### 3.1 NETWORKS STRUCTURE

As shown in Figure 1,the whole network structure includes input layer, convolution layer, sub-sampling layer, the output layer. The feature through the training the whole network, is input to input to Linear support vector machine to classify ,finally get the classification accuracy, and recognition time for every picture. Cerebrospinal fluid cell image is the input layer, feature mapping layer is also called the convolution layer.

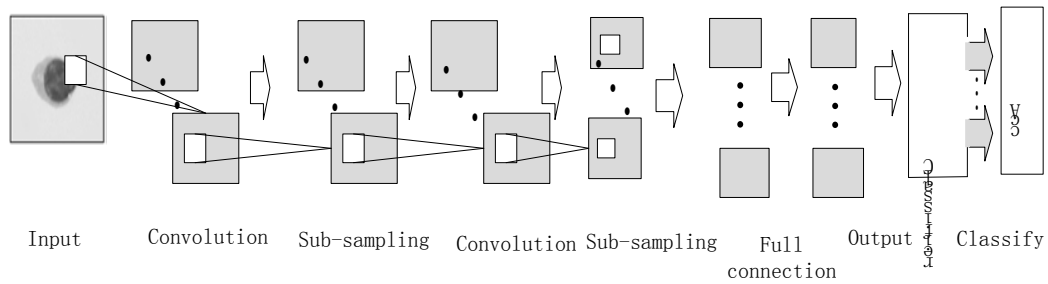


Fig1: The network structure of Convolution layer:

The input pictures operate with a convolution kernel, then input to activation function and obtained the feature map as this layer. This layer can make the different angles to learn image features, which has strong robustness to displacement .

$$X_j^l = f(\sum_{i \in M_j} X_i^{l-1} * F_{ij}^l + B^l) \quad (1)$$

where  $l$  means layer,  $F$  means convolution kernel,  $M_j$  means offset.

The output of this layer can be used as the input of the next layer. The size of Convolution kernel is  $h * h$  . If the size of input image is  $m * n$  , then the convoluted picture size is  $(m - h + 1) * (n - h + 1)$  .

Sub-sampling layer :abandoning part of secondary feature information to reduce the data dimension of the feature map and enhance feature's robustness to image distortion and offset.

$$X_j^l = f(\frac{1}{n} \sum_{i \in M_i} X_i^{l-1} * F_{ij}^l + B^l) \quad (2)$$

The size of sub-sampling is 2 pixels, which means getting the maximum pixel in the adjacent 4 pixels, so after sub-sampling layer, the map of 28\*28 dimensional reduce to 14\*14 dimension.

### 3.2 MODIFIED ACTIVATION FUNCTION

Olshausen and Field [7], for the first time, introduced sparsity in sparse coding of computational neuroscience disciplines. At the same time, it is an important key point of the depth of the convolution neural network[8] and deep confidence network[9], and sparse penalty term has been used in neurology and machine learning.

The standard sigmoid function expression is:

$$y = \frac{1}{1 + e^{-Net}} \quad (3)$$

Its output value does not have sparse row, and some penalty factors are needed to control its sparsity, like  $l_1$ ,  $l_2$  regular. But rectie function is linear correction, and the formula is:  $y = \max(0, x)$ . Cortical neurons in nerve disciplines rarely is still in maximum, and Douglas and Mahowald M et al point out the neuronal excitation value is similar to rectifier function. Its function is that let it equal to zero if the calculated value is less than 0, otherwise keep the original value unchanged. The large number of experiments show[1], the trained network is fully equipped with appropriate sparsity.

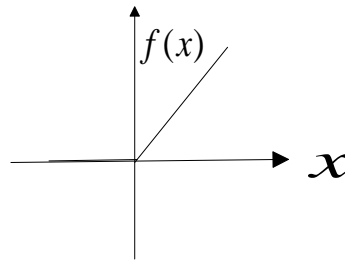


Fig.2 rectier function

Rectier activation function allows a cyber data to obtain sparse representation more easily. In the weights by unified initialization, the continuous output value of 50% hidden units is true zero. Because of this kind of linear relationship, the active paths of the gradient flow in neurons become easier to study. The calculation is also convenient: The calculation of activation of exponential function is not needed, thus the sparsity can be better utilized.

### 3.3 UPDATE THE WEIGH

The step to train convolutionl neural network is similar to BP algorithm.

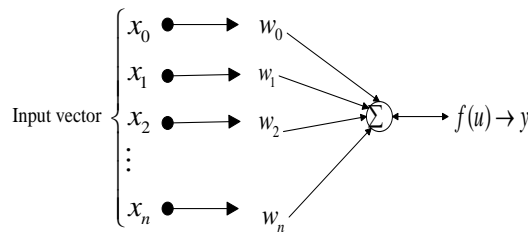


Fig. 3 Structure diagram of a neuron

The weights sharing between each layer of network is an important characteristic of neural networks. Between convolution neural network the learning uses learning method of error correction type, whose the basic idea is to use the difference between the desired output value with the actual output value of neurons as the reference standard criteria of network weights adjustment. The error for the entire network is:

$$E = \sum E_i \quad (4)$$

$$E_i = \frac{1}{2} \sum_{j=1}^m (y_{ij} - o_{ij})^2 \quad (5)$$

where  $E_i$  represents the error of sample  $i$ , and in order to minimize the error, constantly make iteration adjustment on the entire network weights to achieve the optimal solution.

Suppose the input sample set is  $X = (x_0, \dots, x_N)$ , Suppose the input sample set is  $X = (x_0, \dots, x_N)$ , each node of the hidden layer is  $H = (h_0, \dots, h_L)$ , and the output vector of the entire network is  $Y = (y_0, \dots, y_M)$ . The output expression of each unit in the middle is:

$$h_j = f\left(\sum_{i=0}^{N-1} V_{ij} x_i + \partial_j\right) \quad (6)$$

where  $V_{ij}$ : the weight of output unit  $i$  to hidden unit  $j$ ,  $\partial_j$ : the threshold value expression of hidden layer.

$$y_k = f\left(\sum_{j=0}^{L-1} V_{jk} x_j + \theta_k\right) \quad (7)$$

where  $W_{jk}$ : the weight of hidden unit  $j$  to output unit  $k$ ,  $\theta_k$ : the threshold value of output unit.

(1) Initialization: Set weight  $V_{ij}$ ,  $W_{jk}$  and threshold value  $\theta_k$  and  $\partial_j$  to the random numbers between (0,1), and initialize the learning rate  $\alpha$ .

(2) Updating vector: On the basis of the sample set  $x$ , calculate the output vector  $H$  and the actual output vector  $Y$  of the intermediate layer.

(3) Calculating the error value: For the difference between the output vector element  $y_k$  and the target vector element  $d_k$ , and altogether we need to calculate  $M$  such output error:

$$\xi_k = (d_k - y_k) y_k (1 - y_k) \quad (8)$$

where  $D = (d_0, \dots, d_M)$  represents the target output vector of the training; Similarly, the middle hidden layer also needs to calculate  $L$  unit error term, and the error term for each unit is

$$\xi_j = h_j (1 - h_j) \sum_{k=0}^{M-1} \xi_k W_{jk} \quad (9)$$

(4) Adjusting the weights and thresholds:

Adjust the weights:

$$W_{jk}(n+1) = W_{jk}(n) + \Delta W_{jk}(n) \quad (10)$$

$$V_{ij}(n+1) = V_{ij}(n) + \Delta V_{ij}(n) \quad (11)$$

where

$$\Delta W_{jk}(n) = \frac{\alpha}{1+L} (\Delta W_{jk}(n-1) + 1) \xi_k h_j \quad (12)$$

$$\Delta \partial_j(n) = \frac{\alpha}{1+N} (\Delta V_{ij}(n-1) + 1) \xi_k h_j \quad (13)$$

(5) The function of sum of errors of the entire sample is:

$$C = \sum C_s \quad (14)$$

The error function of the single sample:

$$C_s = \frac{1}{2} \sum_{j=1}^m (y_{ij} - o_{ij})^2 \quad (15)$$

When  $C \leq \varepsilon$ , then exit the iteration, otherwise continue the iteration training from formula (1). But then the weights and thresholds are just obtained through learning without initializing again. When exiting the iteration training, the weights and thresholds of each network layer have reached stability, and classify through classifiers and output the classification results.

### 3.4 TRAINING THE CLASSIFIER

In this paper, the processing object is a multi classification of a smaller number (less than four classifications), so only use simple one against rest strategies to train linear support vector machine (Linear Support Vector Machine, LSVM). The training set is assumed given a feature space:  $T = \{(z_1, y_1), (z_2, y_2), \dots, (z_N, y_N)\}$

where  $z_i \in R^n, y_i \in Y = \{1, 2, \dots, L\}, i = 1, 2, \dots, N$ ,  $z_i$  is the  $i$ th eigenvector, and  $y_i$  is the class label of  $z_i$ . LSVM's goal is to learn the linear function  $\{w_c^T z | c \in Y\}$ , then the class prediction for the training set  $z$  can be represented as follows:

$$y = \max_{c \in Y} w_c^T z \quad (16)$$

Use OVR (OVR is multiple classification comparison, and every time solve a binary-class problem, dividing it into: a class for another class.) to train  $L$  binary-class LSVMs, in which each needs to solve a non-limiting convex optimization problem:

$$\min \left\{ J(w_c) = \|w_c\|^2 + C \sum_{i=1}^n \ell(w_c; y_i^c; z_i) \right\} \quad (17)$$

Where, when  $y_i = c, y_i^c = 1$  (representing positive samples), otherwise  $y_i^c = -1$ , where

$$\ell(w_c; y_i^c; z_i) = [\max(0, w_c^T z_i - y_i^c)]^2 \sqrt{a^2 + b^2} \quad (18)$$

is the hinge loss function. Use LBFGS<sup>[10]</sup> algorithm can optimize equation (10).

The recognition process for images is supervised, and tagging samples is needed. The process is as follows:

Test sample set  $P = \{p_1, p_2, \dots, p_N\}$  is given, where  $p_i$  represents the  $i$ th test sample; Category labels  $L = \{l_1, l_2, \dots, l_N\}$  corresponding to the test sample set  $P$  are given, where  $l_i$  represents the class label of the  $i$ th sheet of test sample; By identifying the model make class prediction for the given test sample set  $P$ , and the predicted class label we get is  $M = \{m_1, m_2, \dots, m_N\}$ . Function:

$$f(m_i) = \begin{cases} 1 & m_i = l_i \\ 0 & \text{otherwise} \end{cases} \text{ represents that the identification of a test sample is correct or not, then the}$$

recognition accuracy of the entire test sample set can be defined as:

$$\text{Accuracy} = \sum_{i=1}^N f(m_i) / N \times 100\% \quad (19)$$

## 4 RESULTS AND DISCUSSION

We train the model under 64 bit Win8.1 operating system, Intel (R) i5-4590 3.3.GHz, 8.00G memory, Matlab 2013a.

The experiment object: images of lymphocyte, images of erythrocyte, images of mononuclear cell, images of neutrophils, the samples are shown in following figure.

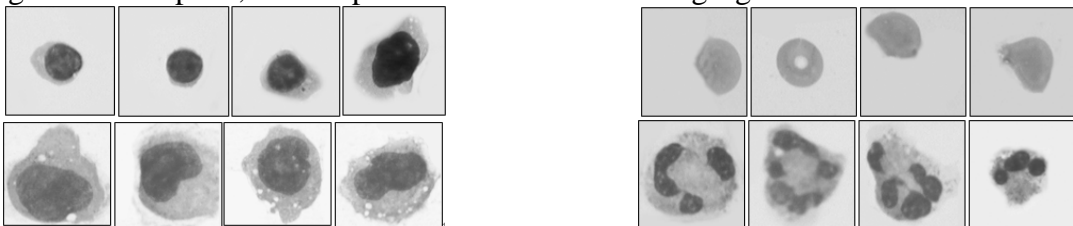


Fig.4 Images of lymphocytes, erythrocyte, monocytes, neutrophils (from left to right, and up to down)

Table.1 Numbers of train samples and test samples

sample of Cerebrospinal fluid cell	Numbers of train samples	Numbers of test samples
images of lymphocyte	360	60
images of erythrocyte	220	60
images of monocyte	230	60
images of neutrophil	117	60

Transform each image with filling zero to 28\*28pixel. The learning rate was 0.01 and followed by 100 iterations to 1000 iterations. Taking the highest accuracy rate as the best classification effect.

Experimental comparison :the different results of three models with different kinds of cells in cerebrospinal fluid.

(1)the ScSPM model: the sparse coding for training on cerebrospinal fluid cell image, take the feature to the Pyramid model, finally input the result to the linear support vector machine to obtain the final accuracy rate.

(2) the CNN[3]model: convolutional neural networks before improvement.

Table.2 The prediction results of two kinds of samples

model	Time per image(sec)			acc(%)		
	Two	Three	Four	Two	Three	Four
ScSPM	1.8	1.8	2.8	90.32	76.63	73.43
CNN	1.45	1.45	1.45	88.23	78.02	70.92
The modified CNN	<b>0.8</b>	<b>1.2</b>	<b>1.4</b>	<b>98</b>	<b>84.5</b>	<b>82.61</b>

The results of table 2 show that, with the increasing of cerebrospinal fluid cell image category, the recognition accuracy is reducing and the single recognition time is also increasing. The improved model in time and accuracy significantly has better results than the obvious model with no improvement and the sparse coding model. Especially for normal and abnormal categories, the improved model accuracy is close to 100%, which will have important application value in medicine.

Throughout the experiment, the model parameters can also cause a wave of the identification accuracy, such as when the learning rate rises from 0.01 to 1, for binary classification cell image problem, the highest accuracy increases to 99% from 98%. Different number of iterations can also cause a change of the identification accuracy, and it is shown in the following figure.

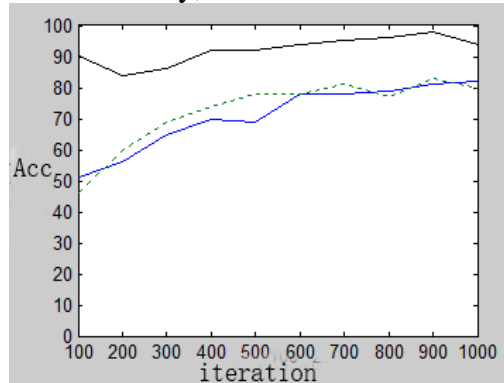


Fig.5 Different accuracy rate with different nums of iters.

The black line represents the accuracy rate trend of two class, the dashed line represents the accuracy rate trend of three class, the blue line represents the accuracy rate trend of four class.

## 5. CONCLUSION

Increasing the density of the image data, sparsity is an important feature of the property. A good feature has been a key point of the image classification. This article take measures to improve the convolution neural network which be used to classify cerebrospinal fluid cell,. The model with rectifier function can be used to get features more sparse than what with sigmoid function. Finally through LSVM classification, the result show the efficiency has improved with the modified model in the cerebrospinal fluid cell image recognition.

Throughout the network training process is to adjust the parameters of the artificial process, if the model can be obtained optimal parameters by learning, then the efficiency will be higher. The follow-up work will also be considered how to adaptively adjust the parameters in order to improve test efficiency.

## REFERENCES

- [1]. Glorot X, Bordes A, Bengio Y. Dee-p sparse rectifier networks[C]//Proce-edings of the 14th International Conference on Artificial Intelligence and Statistics. JMLR W&CP Vol ume. 2011, 15: 315-323.
- [2]. Pollay M. Cerebrospinal Fluid in Diseases of the Nervous System[J]. Neurosurgery, 199 3, 32(2): 325.
- [3]. Bengio Y. Learning deep architectures for AI[J]. Foundations and trends® in Machine L earning, 2009, 2(1): 1-127.
- [4]. Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architectu re in the cat's visual cortex[J]. The Journal of physiology, 1962, 160(1): 106-154.
- [5]. Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[M]//Comp uter Vision–ECCV 2014. Springer International Publishin-g, 2014: 818-833.
- [6]. Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other ker nel-based learning methods[M]. Cambridge university press, 2000.
- [7]. Olshausen B A, Field D J. Sparse coding of sensory inputs[J]. Current opinion in neu robiology, 2004, 14(4): 481-487.
- [8]. Lee H, Grosse R, Ranganath R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations[C]//Proceedings of the 26th Annual International Conference on Machine Learning. ACM, 2009: 609-616.
- [9]. Hinton G, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neur al computation, 2006, 18(7): 1527-1554.
- [10]. Liu D C, Nocedal J. On the limited memory BFGS method for large scale optimiz ation[J]. Mathematical programming, 1989, 45(1-3): 503-528.