CHEMICAL ENGINEERING TRANSACTIONS
VOL. 51, 2016
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet
Guest Editors: Tichun Wang, Hongyang Zhang, Lei Tian

ISBN 978-88-95608-43-3; **ISSN** 2283-9216
DOI: 10.3303/CET1651078

Study on Chaotic Neural Network and its Application in Blood Cell Classification of Medical Image

Limin Zhao*, Peng Yue

College of Management, Xinxiang Medical University, Xinxiang 453002, Henan, China.
zhaoxxmu@163.com

Automatic analysis and classification of blood cell image has become the main stream of blood cell automatic detection technology. However, the automatic classification of blood cell image is not very good, since the large change of blood cell concentration and the shape and the variety of blood cell diseases are very great. For improving the recognizing accuracy rate, we have developed the automatic procedure of the cell extraction which uses the strong classifier power of chaotic neural network (CNN). For validating the efficiency of our method, comparative work has been performed for SVM and CNN on the HEp-2 set of blood cell image which consist of 721 images which belongs to twelve types and widely used. Experiment result shows that the procedure proposed in this paper outperforms the existing algorithms in both recognition precision rates.

1. Introduction

Blood cell testing, also blood routine examination, refers to count and detect the blood white cells, red cells, platelets, haemoglobin and related data in the medical image. Along with the development of medical microscopic image technology, Blood cell imaging technology has laid the foundation for realizing the automation of blood cell testing which major work are classified into blood cell diagnosis and blood cell automation classification. Blood cell diagnosis check whether the existence of cancer cells by analyzing the image of the cell while blood cell automation classification are automatically count and classify the various types of cell according to their shape, colour and other characteristics.

At present, with the development of computer image processing and pattern recognition technology, automatic analysis and classification of blood cell image has become the main stream of blood cell automatic detection technology. The main advantage of this technology is that it uses image processing and analysis technology to classifying the blood cell using pattern recognition by extracting cell shape, color, texture and other characteristics for cell classification. The automatic analysis system of blood cell image has been a very active field in the field of medical information engineering.

However, because of the large change of blood cell concentration, the shape and complexity, and the variety of blood cell diseases and other reasons, the automatic classification of blood cell image is not very good, so the most difficult problem is the recognition between two neighboring cells in their development line. Rezatofighi, et al. (2010) have presented the solution for a system of blood cell recognition applying multilayer perceptron (MLP), but the accuracy is far from good (39% of misclassifications) in the recognition of 16 types of cells. In addition, Umpon and Gader (2002) have reported the recognition of twelve types of cells by using multilayer perceptron at a misclassification of 42%. Long, etal. (2005) have presented the first results of the recognition of 17 blood cells, and the averaged recognition error compared with that of a human expert was 18.7%. The most important task in blood cell concentration is selecting the proper diagnostic features, describing the image by the numerical values, and enabling the automatic system to perform the recognition (Rezatofighi, 2010; Umpon, 2002; Long, 2005). There are many ways on how the features can be generated for the image of the cells extracted from the bone marrow aspirate. Usually, we can generate many (over a hundred) of the features using texture, geometrical, and statistical analyses of the cell image. Not all of them are equally good and applicable in the recognition process. Thus, the important problem is finding out the optimal set of the most important features leading to the highest efficiency of the recognition. Many different approaches have been proposed for the

Please cite this article as: Zhao L.M., Yue P., 2016, Study on chaotic neural network and its application in blood cell classification of medical image, Chemical Engineering Transactions, 51, 463-468 DOI:10.3303/CET1651078

selection of the most important features, which include correlation analysis, analysis of the mean and variance of the data belonging to different classes, principal component analysis etc. In this paper the selection of the best features are obtained by R package of R Functions for Generalized Simulated Annealing (GenSA).
After obtaining the features of blood cell image, the Chaotic Neural Network is used to classifying the blood cell. The chaotic neural network (CNN) is a powerful solution to the classification problems. In this paper, it has been used for the recognition and classification of cells, but it can also be applied to feature selection (Wang, 2004a). The main advantage of the CNN used as a classifier is its very good generalization ability, extremely powerful searching procedure, and very leading to the global optimal solution (Wang, 2004b)
The main task of this work is the development of an algorithm with the highest possible recognition rate of the blood cells based on the image. The first step in the whole procedure is the extraction of the individual cells from the image of the smear image. We have developed the automatic system of the cell extraction which has been implemented in R. The image processing system analyses the blood cell image and performs all morphological operations and, as a result, produces the separated cell images. Secondly, the choice of the numerical features describing the image is obtained using GenSA. The features are the numerical values that should be stable for all representatives of the same class but substantially differ for cells belonging to different classes and we have chosen the size of the cells, shape, granulation, distribution, and intensity of colours. Thirdly, the Chaotic Neural Network is used to classifying and counting the various types of blood cells. Finally, experiment result shows that the procedure proposed in this paper outperforms the existing algorithms in both recognition precision rates.

2. Chaotic neural network

There have been extensive research interests in theory and applications of Hopfield-based type neural networks. Since the original Hopfield neural network (HNN) (Wang, 2004b) can be easily trapped in local minima, stochastic simulated annealing (SSA) has been combined with the HNN. Besides, chaotic neural networks (Lee, 2004) have also attracted much attention because chaotic neural networks have a richer spectrum of dynamic behaviors, such as stable fixed points, periodic oscillations, and chaos, in comparison with static neural network models. Beksac etal. (2007) demonstrated the search ability of chaotic neural networks. Theera-Umpon and Gader (2002) proposed chaotic simulated annealing (CSA) by starting with a sufficiently large negative self-coupling in the neurons and then gradually decreasing the self-coupling to stabilize the network. They called this model the transiently chaotic neural network (TCNN). Because the TCNN restricts the random search to a subspace of the chaotic attracting set, which is much smaller than the entire state space, it can search more efficiently.
By adding decaying stochastic noise into the TCNN, Wang et al.(2004a) proposed a new approach to simulated annealing, i.e., stochastic chaotic simulated annealing (SCSA), using a noisy chaotic neural network (NCNN). Compared with CSA, SCSA performs stochastic searching both before and after chaos disappears and is more likely to find optimal or suboptimal solutions. This novel method has been applied successfully to solving several challenging optimization problems, including the traveling salesman problem (TSP) and the channel assignment problem (CAP). The TCNN model is described as follows in：

$$x_{jk} = \frac{1}{1+e^{-y_{jk}(t)/\varepsilon}} \tag{1}$$

$$y_{jk}(t+1) = k y_{jk}(t) - z(t)\left(x_{jk}(t) - I_0\right) + n(t) + \alpha\left(\sum_{i=1, i\neq j}^{N} \sum_{l=1, l\neq k}^{M} \omega_{jkil} x_{jk}(t) + I_{jk}\right) \tag{2}$$

$$z(t+1) = (1-\beta_1) z(t) \tag{3}$$

$$n(t+1) = (1-\beta_2) n(t) \tag{4}$$

In the above equations, the notation are that x_{jk} is output of neuron of jk, y_{jk} is input of neuron of jk, ω_{jkil} is the connection weight from neuron jk to neuron il, $z(t)$ is self-feedback connection weight, $n(t)$ is random noise injected into the neurons, k, α, β_1 and β_2 are the parameter of damping factor of nerve membrane, positive scaling parameter for inputs, damping factor for neuronal self-coupling and damping factor for stochastic noise, I_0 is positive parameter, ε is steepness parameter of the output function and E is the energy function.

3. Simulated Annealing

Simulated annealing is a well-established stochastic technique originally developed to model the natural process of crystallization and later adopted to solve optimization problems. As with a greedy search, it accepts all changes that lead to improvements in the fitness of a solution. However, it differs in its ability to allow the probabilistic acceptance of changes which lead to worse solutions; i.e., reversals in fitness. The probability of

accepting a reversal is inversely proportional to the size of the reversal with the acceptance of smaller reversals, being more probable. This probability also decreases as the search continues, or as the system cools, allowing eventual convergence on a solution. It is defined by Boltzman's equation:

$$P(\Delta E)\infty e^{\frac{-\Delta E}{T}} \quad (5)$$

In which ΔE is the difference in energy (fitness) between the old and new states and T is the temperature of the system. In the virtual environment, the temperature of the system is lowered after certain predefined number of accepted changes, successes, or total changes, attempts, depending on which is reached first. The rate at which temperature decreases depends on the cooling schedule. In the natural process, the system cools logarithmically; however, this is so times consuming that many simplified cooling schedules have been introduced for practical problem solving. The following simple cooling model is widely used:

$$\mathrm{T(K+1)} = \frac{T(K)}{1+\sigma} \quad (6)$$

In which $\mathrm{T(K+1)}$ is the current temperature, $\mathrm{T(K)}$ is the previous temperature, and σ dictates the cooling rate, simulated annealing has been applied to such problems as the well-known traveling salesman problem and function optimization.

4. Results of experiments

4.1 Experimental setup

The database used in this work is the HEp-2 set of blood cell image provided by ICPR 2012 for classifying the type of blood cell (Theera-Umpon and Gader, 2002). This data consist of 721 images of HEp-2 which belongs to twelve types. For convenience, all image are processed and transformed into the 64×64 sub image which adopt the segment methods of 4 pixel× 4 pixel and get 16× 16=256 block for each image. More than 2000 blood cells taking part in the experiments have been acquired this way. Twelve families of cells have been used in the experiments. The available blood cell image data is divided into two subsets: One subset that contains two third of the data set has been used in learning and the remaining part (one third) has been used for testing only and 5 times cross validation is used to perform 10 pass for getting the averaging experiment results. The learning procedure of CNN uses the package of General regression neural network (GRNN) which is run on a3.1 GHz personal computer with 8GB random access memory.

4.2 Results of the SA experiments

Since the feature number of blood cell image is very large, we have applied the SA for feature selection. To get an objective assessment of the results, we have performed 25 runs of the SA procedure at different initial temperature with the different initial population and different divisions of the data for the learning and testing subset. The best results of the SA have been compared with the results obtained with the application of all features and also with the best results corresponding to the 20 features selected using linear CNN ranking.

Table 1: Results of the blood cell recognition error rate in the best run of SA

Cell type	SA	CNN(20	All feature
1	16.21%	18.36%	23.67%
2	9.63%	15.43%	18.64%
3	21.58%	24.21%	30.02%
4	15.31%	18.62%	19.97%
5	18.63%	24.37%	28.94%
6	16.37%	18.36%	21.33%
7	12.35%	15.63%	19.23%
8	8.36%	12.68%	17.25%
9	21.58%	25.74%	29.37%
10	11.23%	12.58%	14.37%
11	13.58%	16.37%	18.92%
12	29.49%	31.25%	37.29%
average	16.19%	19.47%	23.25%

Table 1 depicts the results of the relative recognition error of the cells using the proposed SA method and the selection using linear CNN ranking, with the application of the 20 best features and no ranking of features. All errors have been computed as the difference between the results of our system and the human expert score used as the reference. The errors have been calculated as the ratio of the number of misclassified samples to the total number of cells belonging to the appropriate class. The average error of recognition was calculated as the ratio of the number of misclassified samples to the total number of samples in the last row of the Table 1.As shown in Table 1, although the best results of the SA is the best in three above cases, the case of CNN with 20 features is almost the same as that of SA. So the followed recognition process adopts the 20 features obtained by SA as the major feature for CNN classification. Table 2 presents the mean and standard deviation of recognition accuracy rate for 12 types of blood cell and the last row give the mean result of 12 types of blood cell. As table 2 shown, the recognition accuracy rate of CNN is stable for all types of blood cell.

Table 2: The results of the blood cell recognition accuracy rate using CNN

Cell type	Mean of recognition accuracy rate	Standard deviation of recognition accuracy rate
1	82.23%	2.31%
2	83.62%	3.12%
3	79.21%	3.93%
4	83.54%	2.92%
5	84.57%	1.82%
6	86.67%	2.84%
7	81.25%	3.21%
8	85.36%	2.17%
9	86.97%	1.12%
10	82.14%	2.39%
11	81.51%	5.36%
12	76.81%	2.39%
average	82.82%	2.80%

4.3 Comparing study of experiment

This paper is a study of automatic cell classification in the classification process, the classification of nuclear extract follow-up is essential, it is the relationship between the training samples and the accuracy of the classification results, as everyone knows the training model is good or bad depends entirely on the training model of training sample training. According to the following sampling strategy, two indicators of over fitting rate and under fitting rate are used to determine the classification results in which results obtained by over fitting rate on the classification are very large and the smaller it is and the better is it, so its value is not too large. Table 2 presents the statistical measures of the accuracy of the CNN approach to the recognition of the considered blood cells. The statistics (the mean value and the standard deviation) have been estimated based on 25 independent runs of the CNN. We see that the results of the individual runs of the CNN are close to each other. (Standard deviations take small values.) However, the accuracy of recognition varies from class to class. For some classes, the relative misclassification rate is rather high (Yin and Liu, 2015). This is due to the high similarity among the neighbours for these classes.

Additionally, the data for the experiments have been collected from 48 patients suffering from different kinds of leukaemia at various stages of development of the illness are tested in this paper for validating the efficiency of CNN. Since the data used in classification have been gathered in the hospital within a few years, and their image quality was also very different which result in a great variation among the cells belonging to the same type and make the recognition problem extremely difficult. However, the important advantage of the application of our algorithm is the simultaneous determination of the image quantity and the classification of cell type since our method adopt the rational process step for selection the representative the feature set. It is interesting to compare the number of features selected by the CNN and other algorithm such as SVM (Osowski, 2007; Siroic, 2007) for the recognition between different classes of cells. Table 3 and Table 4 present the numbers and the precision rate of SVM and CNN. According to the above table 4 and Table 4, the overall precision of CNN is superior to those of SVM. Such as, for the first type, the number of blood cell is 21, the precision of SVM is

80.95 while that of CNN is 95.24 which are far higher than SVM. Another example is the 7th type of blood cell which precision rate is the almost lowest for CNN as well as SVM, the precision rate of CNN is 86.67% while that of SVM is 80.00% which is less 6.67% than CNN. Furthermore, for the other 11 type of blood cell the above case is also true which can be explained by the last row of Table 3 and Table 4. The mean precision rate of SVM is 84.98% while that of CNN is 94.47%.

Table 3: The experiment result of SVM

Cell type (number of cell)	Recognized 1	Recognized 2	Recognized 3	Recognized 4	Recognized 5	Recognized 6	precision rate
1(21)	17	0	1	0	2	1	80.95%
2(23)	20	0	1	0	1	1	86.96%
3(14)	12	0	2	0	0	0	85.71%
4(27)	24	1	0	1	1	0	88.89%
5(18)	15	0	1	0	1	1	83.33%
6(17)	13	0	0	2	1	1	76.47%
7(15)	12	1	1	0	1	0	80.00%
8(28)	24	1	0	1	1	1	85.71%
9(26)	22	0	1	1	1	1	84.61%
10(25)	23	0	1	0	1	0	92.00%
11(19)	17	1	0	0	1	0	89.47%
12(20)	16	0	1	1	1	1	80.00%
total(253)	215	4	9	6	12	7	84.98%

Table 4: The experiment result of CNN

Cell type(number of cell)	Recognized 1	Recognized 2	Recognized 3	Recognized 4	Recognized 5	Recognized 6	precision rate
1(21)	20	0	1	0	0	0	95.24%
2(23)	22	0	0	0	0	1	95.65%
3(14)	14	0	2	0	0	0	100.00%
4(27)	26	1	0	1	1	0	96.30%
5(18)	17	0	0	0	1	0	94.44%
6(17)	15	0	0	0	1	1	88.24%
7(15)	13	0	1	0	1	0	86.67%
8(28)	26	1	0	1	0	0	92.86%
9(26)	25	0	0	1	0	0	96.15%
10(25)	23	0	1	0	1	0	92.00%
11(19)	19	0	0	0	0	0	100.00%
12(20)	19	0	1	10	0	0	95.00%
Total(253)	239	2	6	13	5	2	94.47%

4. Conclusion

The automatic classification of blood cell image is very important and is vital for blood routine examination. However, the large change of blood cell concentration and the shape and the variety of blood cell diseases are very great, so the current algorithms for classification of blood cell image cannot achieve the effect of practical use. For improving the recognizing rate we proposed the new methods of recognizing the type of blood cell which used the combination of SA as feature selection and CNN as classifier. Since SA can obtain the better features which exactly represent the blood cell image and CNN has strong fault tolerance and stability, so the new algorithm proposed in this paper has the better classification ability, comparative study on the data set of HEp-2 set of blood cell image which consist of 721 images which belongs to twelve types are performed on CNN and SVM and experiment result shows that our proposed method is better than that of SVM, so our method is more practical and more accuracy in classifying the blood cell.

Reference

Beksac M., Beksac M. S., Tippi V. B., Duru H. A., Karakas M. U., Nurcakar A., 1997, An Artificial Intelligent Diagnostic System On Differential Recognition Of Hematopoietic Cells From Microscopic Images, Cytom.,30(3), 145–150

Lee R. S. T., 2004, A Transient-Chaotic Auto Associative Network (TCAN) Based On Lee Oscillators, IEEE Trans. Neural Netw., 15(5), 1228–1243, DOI: 10.1109/TNN.2004.832729..

Long X., Clevelnd W. L., Yao Y. L., 2005, A New Preprocessing Approach For Cell Recognition, IEEE Transaction Information Technology in Biomedicine, 9(3), 407-412, DOI: 10.1109/TITB.2005.847502.

Osowski S., Markiewicz T., 2007, Support Vector Machine For Recognition Of White Blood Cells Of Leukaemia, Kernel Methods in Bioengineering, Signal and Image Processing, 93–123, DOI: 10.4018/978-1-59904-042-4.ch004

Rezatofighi S.H., Khaksari K., Soltanian-Zadeh H., 2010, Automatic Recognition Of Five Types Of White Blood Cells In Peripheral Blood, in Proceeding of International Conference Image Analysis and Recognition, Povoa de Varzim, Portugal.

Siroic, R., Osowski, S., Markiewicz, T., & Siwek, K., 2007, Support Vector Machine And Genetic Algorithm For Efricient Blood Cell Recognition. IEEE Instrumentation & Measurement Technology Conference 1-6, DOI: 10.1109/IMTC.2007.379321.

Theera-Umpon N., Gader P., 2002, System-Level Training Of Neural Networks For Counting White Blood Cells, IEEE Trans. Syst., Man, Cybern. 32(1), 48–53, DOI: 10.1109/TSMCC.2002.1009139.

Umpon N.T., Gader P.D., 2002, System-Level Training Of Neural Networks For Counting White Blood Cells, IEEE Transaction System, Man & Cybernetics, 32(1), 48-53, DOI: 10.1109/TSMCC.2002. 1009139.

Wang L., Li S., Tian F., Fu X., 2004a, A Noisy Chaotic Neural Network For Solving Combinatorial Optimization Problems: Stochastic Chaotic Simulated Annealing, IEEE Transaction System, Man & Cybernetics 34(5), 2119–2125, DOI: 10.1109/TSMCB.2004.829778.

Wang R. L., Tang Z., Cao Q. P., 2004b, A Hopfield Network Learning Method For Bipartite Subgraph Problem, IEEE Trans. Neural Netw., 15(6), 1458–1465, DOI: 10.1109/TNN.2004.836234.

Yin L., Liu Y.G., 2015, Biclustering of the gene expression data by coevolution cuckoo search. International Journal Bioautomation, 19(2): 161-176