Research on Parameter Optimization of Neural Network*

Guifang Wu(1), Yumin Ren(2), Yan Li(1), Hoonsung Kwak(3), Seyoung Jang(3)
(1) Electronic Information Engineering College, Henan University of Science & Technology, Luoyang 471003, China
(2) Dept. of Children's Clinic, Luoyang Dongfang Hospital, Luoyang 471003, China
(3) Dept. of Computer Engineering, Engineering College, Chonbuk National University of Korea, Jeonju 561756, Korea
E-mail: Easyfancy@126.com

Abstract

Based on researching parameter optimization method of neural network deeply, a new parameter optimization method is presented and applied to surface defect online inspection system of cold rolled strips. The method takes advantages of small-samples fully, and can get a group of neural network parameters which can mostly express the neural network under a certain specific condition. The method is advantageous for its simplicity, easy to maintain and fast, it can be applied to many fields too, such as iron-steel industry, medicine. Experiments showed that a best recognition effect by using the parameters for neural network which are achieved by the new parameter optimization method can be got among all the parameters optimized randomly for surface defect of cold rolled strips.

1. Introduction

It is always a serious problem for some researchers to select a group optimal parameters of neural network, because there are no fixed criteria to choose a set of optimized parameters for neural network, but at the same time, whether the choice of the parameters is optimum or not is a very important factor to influence the precision of neural work. When using the neural network, many researchers chose the parameter for it by experience or finding a better group of parameters from certain of possible parameters as reference [1] and [2]. If parameters are selected completely by experience, the unique individuality of the problem will not be involved, and when the individuality is extrusive, a big error of neural network will occur between theoretic results and practical results. If parameters are selected by comparing results of experiments group by group under certain scheduled parameters, the work will be too tedious and fussy to make it impossible. At the same time, the training neural network is always very complicated, therefore it will take long time for every experiment if traversing every possible combination of all the parameters manually, and the efficiency will be too low to make the experiment impossible.

There are different automatic detection systems for different objects in the field of surface detection as reference [3]-[5]. The surface defect online inspection system of cold rolled strips is a set of real-time inspection system combined with the technologies of image process, pattern recognition and automatic control, it can work under a hostile environment to do continuous real-time inspection of cold rolled strips all day long. The

* This work is financially supported by Talent Science Research Foundation of Henan University of Science & Technology of China(No.09001121), all the data is acquired from Hainan Limited Corp. of Wuhan Iron & Steel Group of China.

recognition of surface defects of cold rolled strips is a major part of the system and it is very important, the effect of recognition influences the performance of the whole system to some deep extent. Whether the system can recognize the types of defects which have been detected by the system, is directly concerned with the system's intelligence and practicability.

When applying neural network to make recognition of the defects, the choice of optimized parameters is a most important tache as reference [6], it can decide the recognition result to some deep extent. Therefore, in order to settle the parameter optimization problem, a new parameter optimization method of neural network is presented. The method originated from practical problem which need be settled, it considered not only the time cost of large samples but also the authenticity whether the parameters can reflect the problem's essence.

2. Theory for parameter optimization

The theory parameter optimization method of neural network is aiming at finding inner parameters for describing practical problem based on small-samples as reference [7] and [8], so that the parameters can meet the solution of the problem, and it makes possible for avoiding blindly finding possible parameters of neural network. Figure 1 shows the flowchart of the new parameter optimization method based on small-samples.

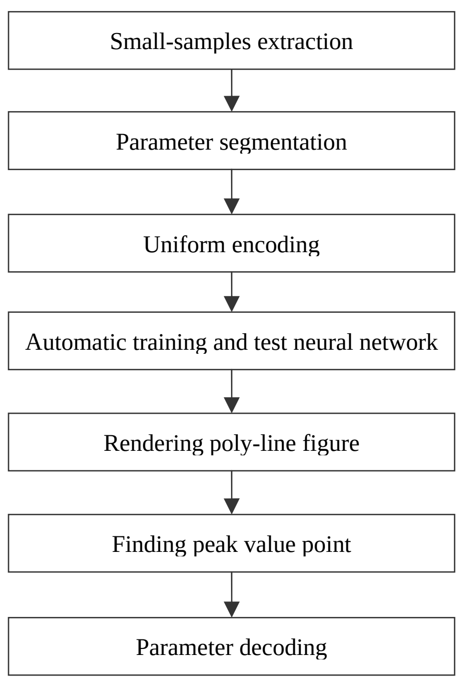

Figure 1. Flowchart of parameter optimization method

2.1 Acquirement of small-samples

To make parameters reflect the essence of practical problem, a certain proportion of samples is extracted from all the samples to form a new sample aggregation, named small-samples, and let the proportion be η. We supposed that the original sample aggregation is P_0 as equation (1).

$$P_0 = \{X_1, X_2, \cdots, X_n\} \tag{1}$$

The amount of every item inside the original sample aggregation is $p_0^1, p_0^2, \cdots, p_0^n$ respectively. At the same time, we supposed that the new sample aggregation is P_s as equation (2)

$$P_s = \{X_1^{'}, X_2^{'}, \cdots, X_n^{'}\} \tag{2}$$

Its amount is N, therefore the extraction rule is defined as following:

Extracting every class of samples by its proportion from sample library, and defining the number of every sample as p_s^i, therefore it can describe as equation (3).

$$p_s^i = \frac{p_0^i}{\sum_{j=1}^{n} p_0^j} N \qquad i = 1, 2 \cdots n \tag{3}$$

Consequently the distribution law of extracted new small-samples P_s is the same as original samples P_0, in another word that the extracted small-samples can stand for the original samples to some extent, and the parameters of neural network selected under such small-samples can express the map relationship of the neural network under the whole original samples. Such small-samples selected by the method is so called proportional samples. A group of training small-samples p_{s1} and testing small-samples p_{s2} can be extracted in the practical application.

2.2 Parameter segmentation and encoding method

In practical engineering project, firstly it is necessary to know all the parameters of neural network, and choose the parameters which are related with the essence of problem among all the parameters. Now defining the number of selected parameters as p, and all parameters are independent and have no relation with samples theoretically, in another word that the selected parameter variables can reflect the essence characters of the practical problem. After determined parameters which need be optimized, it is time to segment all the parameters under a scope of experience values or theoretical values.

Supposed that all these p parameters are divided into such parts respectively as $C_1, C_2, \cdots, C_p$, then combining all the parameters under unrepeated rule, and there are total M combinations, it can be described as equation (4):

$$M = \prod_{i=1}^{P} C_{C_i}^1 \tag{4}$$

What come on next? We must encoding the parameters for every combination, and assign a number to it as 1, 2, ⋯, M. There are mainly 2 steps in this stage as following:

Firstly, determining the order of every parameter and fixing its position from the front to the back. Figure 2 shows the exact order of every parameter in the same group.

P_1	P_2	…	P_M

Figure 2. Exact order of every parameter

Secondly, a uniform serial number is given to every combination. When a specific number is given to a front parameter, the back parameter will traverse all possible values, and the number will increase 1 step by step till all M possibilities is finished, and the encoding finished.

2.3 Training and visual method in getting result

After encoding all the parameters, a automatic program must be developed, and every group of parameters is used as parameters of neural network respectively to do training and testing. The procedure can be described as following:

1. Take out one specific group of parameters as parameters of neural network.

2. Training the neural network with the training small-samples of P_{s1}.

3. When the training finished, saved the trained neural network model for testing.

4. Testing the neural network with testing small-samples p_{s2}.

5. Making a statistic of its recognition rate.

6. Repeating step 1 to 5 again till all M group of parameters are traversed for one time.

7. Using the M recognition rate values under different M group of parameters of neural network to plot a poly-line figure on a coordinate system.

8. Finding the peak recognition rate value of the figure, and writing down its related group number of parameters which is selected as the final parameters of neural network under the original samples.

3. Instance for processing parameter optimization

In order to explain the theory of parameter optimization method, an instance is given as followings. Table 1 is sample distribution table which are all acquired by surface defect online inspection system of cold rolled strips from locale of certain production line in China. From the table, it can be seen that there are total 12 classes of defects of cold rolled strips, and there are total 6360 original training samples and 3180 original testing samples.

Table 1. Samples distribution table of system

Item	scratches	coil breaks	point sticks	feathers	white spots	roll imprints	edge foldings	rusts	emulsion marks	orange peels	edge cracks	other	Total
Number of training samples	1223	1398	1203	1153	463	295	193	121	89	80	70	71	6360
Number of testing samples	612	699	602	577	232	147	96	61	45	40	35	35	3180

Table 2. Distribution of extracted small sample

Sample type	scratches	coil breaks	point sticks	feathers	white spots	roll imprints	edge foldings	rusts	emulsion marks	orange peels	edge cracks	other	Total
Number of small-samples p_{s1}	98	112	96	92	37	24	15	10	7	6	6	6	509
Number of small-samples p_{s2}	49	56	48	46	19	12	8	5	4	3	3	3	254

According the extraction rule of small-samples of parameter optimization method described in preceding part, a training small-samples aggregation P_{s1} and a testing small-samples aggregation P_{s2} will be get from original training samples and testing samples under a given extraction proportion of 8% and 4% respectively. The distribution of extracted small-samples is shown as Table 2.

For LVQ3 neural network as reference [9], the number of neurons in compete layer must be determined firstly, and according the project, it is assigned 106. Secondly, make sure that there are three parameters which are related with the essence of described system, they are initial learn rate $\eta(0)$, relative learn rate ε and window width *m*. According to the experience of LVQ3 neural network, the experience scope of all the three parameters can be described as equation (5).

$$\begin{cases} 0.01 \le \eta(0) \le 0.1 \\ 0.05 \le \varepsilon \le 0.3 \\ 0.1 \le m \le 0.5 \end{cases} \tag{5}$$

Therefore, all these three parameters can be segment within their own experience scopes. $\eta(0)$ is segmented to 5 values such as 0.01, 0.03, 0.05, 0.08 and 0.1, ε is segmented to 4 values such as 0.05, 0.1, 0.2 and 0.3, and *m* is segmented to 5 values such as 0.1, 0.2, 0.3, 0.4 and 0.5. Consequently, there are total *M* combinations shown as equation (6).

$$M = C_5^1 \times C_4^1 \times C_5^1 = 100 \tag{6}$$

After segmentation, a group of number is given to every combination such as 1, 2, ⋯, 100, and every encoding number stands for a specific group of parameter values. The method of encoding here is to give a specific value to the parameters of $\eta(0)$, ε and *m* respectively, and when the front parameter is given a value, the back parameter traverses all the possible value of itself once from the back to front, the number will increase for 1 step when a different value is given to any parameter. For example, No.5th combination of the parameters stands for the value of (0.01, 0.05, 0.5), it means that parameter $\eta(0)$ is 0.01, parameter ε is 0.05, parameter *m* is 0.5, for the same reason, No. 16th combination of the parameters stands for the value of (0.01, 0.3, 0.1), and every other combination of parameters can be worked out by the same way.

For BP neural network, there are two main parameters must be determined, such as initial learn rate $\eta(0)$ and inertial item coefficient α as reference [1] and [2]. According to experience, the parameter $\eta(0)$ is always between 0.9 and 1.0, while parameter α is between 0.7 and 0.9. Therefore the parameter $\eta(0)$ can be segmented to 10 items such as 0.9, 0.91, 0.92, L , 0.99, and α can be segmented to 11 items such as 0.7, 0.72, 0.74, L , 0.9. Consequently there are total 110 combinations by these tow parameter, and gave a number of 1, 2, L and 110 as their corresponding parameter combination number, and its encoding method is the same as LVQ3. For example, No. 5th combination of the parameters stands for (0.9, 0.78) and No. 16th parameter combination stands for (0.91, 0.78). Therefore all other parameter combination can be worked out by the same way.

Next, it is time to train the LVQ3 and BP neural network with different parameter combination using training small-samples P_{s1}, and work out its recognition rate using test small-samples P_{s2}. Then the poly-line will be plotted under coordinates of recognition rate and parameter combination number as shown in Figure 3 and Figure 4.

It can be seen from Figure 3 that the coordinate (48, 96.2%) stands for the peak point in the figure, it means that No. 48 parameter combination can describe the practical problem better than other parameter combination of LVQ3 neural network, and the test result can reach the highest recognition rate of 96.2%. By decoding the parameter combination, every parameter value can be expressed by equation (7).

$$\begin{cases} \eta(0) = 0.05 \\ \varepsilon = 0.1 \\ m = 0.3 \end{cases} \tag{7}$$

For the same reason, the experiments of BP neural work can get a figure as Figure 5, and the coordinate of (2, 92.32%) shows a peak point of recognition rate, it means that No. 2 parameter combination can get a highest recognition rate of 92.3% among all the parameter combinations, and every parameter value can be get as equation (8) by decoding.

$$\begin{cases} \eta(0) = 0.9 \\ \alpha = 0.72 \end{cases} \tag{8}$$

Therefore, it can be known from above that equation (7) stand for the optimized parameters of LVQ3 neural network and equation (8) stand for the optimized parameters of BP neural network when applying to surface defect recognition of cold rolled strips.

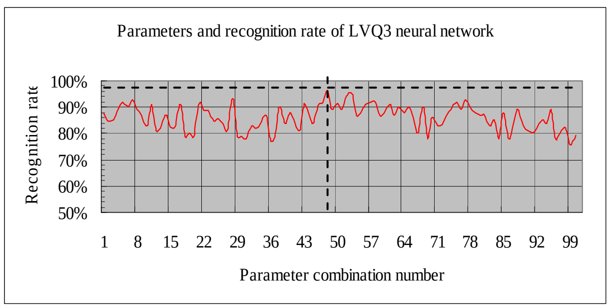

Figure 3. Curve of optimized parameter optimization of LVQ3 neural network

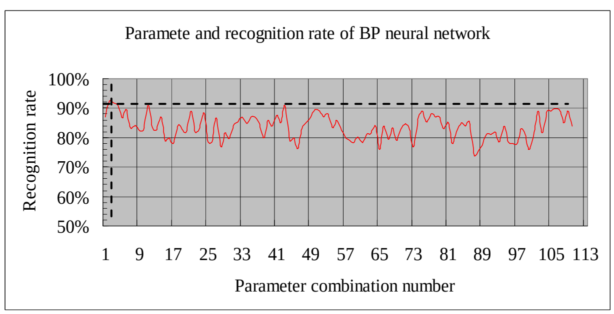

Figure 4. Curve of optimized parameter optimization of BP neural network

4. Experiment results

According to the parameters optimized by the method presented above, all parameters of LVQ3 neural network can be achieved just like Table 3, and many experiments will be done to construct surface defect recognition system of cold rolled strips under all the parameters of LVQ3 neural network in Table 3. Meanwhile, in order to verify the recognition effect under optimized parameters, several contrast similar system must be constructed with different such three parameters as initial learn rate

$\eta(0)$, relative learn rate ε and window width m. Then inputting all training and testing samples in Table 1 to all the different recognition systems with three main different parameters, and work out their recognition result such as number of correct recognition, recognition rate, etc. Therefore a train result table under different neural network parameters can be drawn as Table 4, and a test result table under different neural network parameters can be drawn as Table 5.

In Table 4 and Table 5, No.1 group of parameters of LVQ3 neural network are get by the new parameter optimization method, and other group of parameters from No.2 to No.8 are selected combined randomly. To display the result intuitionisticly, a histogram of recognition rate under different neural network parameters is drawn as Figure 5.

It can be seen from the two tables of Table 4 and Table 5 that the recognition rate under No.1 group of parameters reaches to 94.28% and 89.69% respectively, and they are both highest in each table, but under other 7 group of parameters, the recognition rate of LVQ3 neural network, the highest training recognition rate is 90.19% under No.6 group of parameters, and the highest testing recognition rate is 87.77% still under No.6 group of parameters, they are both lower than the recognition rate under No.1 parameters. At the same time, the lowest recognition rates appeared in No.2 group of parameters, they are 81.64% in training and 78.40% in testing, and they are respectively 12.64% and 11.29% lower than No.1 group. And is can be seen from Table 4 and Table 5 that the average recognition rate of other 7 groups of parameters except for No.1 group of parameters is 86.16% and 82.75% respectively, and it is 8.12% and 6.94 lower than highest item.

Table 3. Parameters of LVQ3 neural network

Item	Neuron number of input layer	Neuron number of competion layer	Neuron number of output layer	Initial learn rate $\eta(0)$	Relative learn rate ε	Window width m	Iteration times
Value of parameter	22	1276	12	0.05	0.1	0.3	51040

Table 4. Training results under different neural network parameters

No.	Neural network type	Training condition			Training result		
		$\eta(0)$	ε	m	Number of correct	Recognition rate	Difference with No.1
1	LVQ3	0.05	0.1	0.3	5996	94.28%	0.00%
2	LVQ3	0.01	0.05	0.1	5192	81.64%	12.64%
3	LVQ3	0.1	0.3	0.5	5404	84.97%	9.31%
4	LVQ3	0.08	0.2	0.4	5381	84.61%	9.67%
5	LVQ3	0.03	0.3	0.2	5488	86.29%	7.99%
6	LVQ3	0.06	0.08	0.3	5684	89.37%	4.91%
7	LVQ3	0.01	0.3	0.1	5475	86.08%	8.20%
8	LVQ3	0.05	0.1	0.5	5736	90.19%	4.09%

Table 5. Testing results under different neural network parameters

No.	Neural network type	Testing condition			Testing result		
		$\eta(0)$	ε	m	Number of correct	Recognition rate	Difference with No.1
1	LVQ3	0.05	0.1	0.3	2852	89.69%	0.00%
2	LVQ3	0.01	0.05	0.1	2493	78.40%	11.29%
3	LVQ3	0.1	0.3	0.5	2598	81.70%	7.99%
4	LVQ3	0.08	0.2	0.4	2530	79.56%	10.13%
5	LVQ3	0.03	0.3	0.2	2645	83.18%	6.51%
6	LVQ3	0.06	0.08	0.3	2791	87.77%	1.92%
7	LVQ3	0.01	0.3	0.1	2609	82.04%	7.65%
8	LVQ3	0.05	0.1	0.5	2755	86.64%	3.05%

Therefore, the experiments here fully proved that using the optimized parameters of neural network by the new parameter optimization method as the parameters of neural network, the classification information can be expressed more accurately, a higher recognition rate can be got when applying to recognize surface defect of cold rolled strips, and it can get a higher recognition rate when applying to medicine recognition field, such as heart disease, lung disease and brain disease, especially for children's disease, because it required higher accuracy.

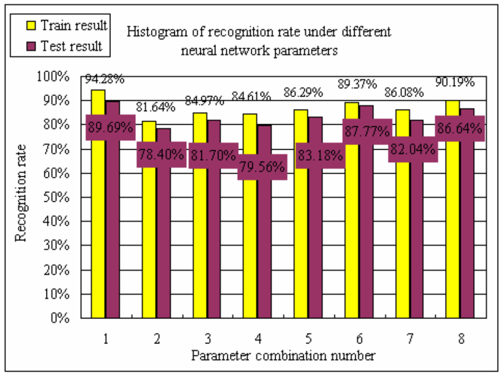

Figure 5. Histogram of recognition rate under different parameters

5. Conclusion

When applying neural network to deal with data, the parameters of neural network influence recognition effect to a deep content, and it determines the recognition effect directly in some special field which requires higher accuracy such as children's heart disease, lung disease, brain disease, etc., therefore, the parameter optimization was researched deeply in the paper. While neural network involves training and testing of multi-samples, it is difficult to select optimized parameters only by experiences, and whether the experience values of the parameters can reflect the essence of practical problem was not settled effectively. After researched selection law of samples and essence of problem requiring settled, the new parameter optimization method of neural network presented here can give optimized parameters of neural network based on considering experience scope of every parameter of neural network. Experiments show that using the parameters optimized in the new method as parameters of neural network, it can reach optimum effect of recognition, and at the same time it can improve recognition rate of samples.

6. References

[1] K. Cai, J. T. Xia, L. T. Li, et al., Analysis of the electrical properties of PZT by a BP artificial neural network, Computational Materials Science, 2005, 34(2), pp. 166-172.

[2] K. Xu, J. W. Xu, X. J. Ban, Research on Pattern Recognition Method Of Automatic Surface Quality Inspection System for Cold Rolled Strips, Iron & Steel, 2002, 37(6), pp. 28-31.

[3] I. Moriyama, Y. Tanabe, I. Ishimaru, et al., Surface inspection system with automatic discrimination of microscratches and particles on chemical mechanical polished wafers, Proceedings of the 1999 IEEE International Symposium on Semiconductor Manufacturing, 1999, pp. 139-142.

[4] Hayati S., Balaram J., Seraji H., et al., Remote surface inspection system, Proceedings of 1993 IEEE International Conference on Robotics and Automation, 1993, 3, pp. 875-882.

[5] S. M. Kim, Y. C. Lee; S. C. Lee, Vision Based Automatic Inspection System for Nuts Welded on the Support Hinge, SICE-ICASE International Joint Conference, 2006, 10, pp. 1508-1512

[6] Pedrycz, W., Keun-Chang Kwak, Linguistic models as a framework of user-centric system modeling, IEEE Transactions on Systems, Man and Cybernetics, 2006, 36(4), pp. 727 – 745.

[7] B. Tang, S. W. Luo, H. Huang, High performance face recognition system by creating virtual sample, Proceedings of the 2003 International Conference on Neural Networks and Signal Processing, 2003, 2, pp. 972 – 975.

[8] M. Moszynski, M. Balcerzyk, M. Kapusta et al., CdWO/sub 4/ crystal in gamma-ray spectrometry, IEEE Transactions on Nuclear Science, 2005, 52(6), pp. 3124 – 3128.

[9] Y. Satyanarayana, Sanjeev Naithani, R. Anu, Seafloor sediment classification from single beam echo sounder data using LVQ network, Marine Geophysical Researches, 2007, 28(2), pp. 95-99.

Authors

Guifang Wu (1978-), born Jiangxi province of China, graduated and received his PhD in 2006 in Mech-tronical engineering from University of Science and Technology Beijing, and then worked in Henan University of Science and Technology. From 2007 to 2008 he worked and studied in Chonbuk National University of Korea as a visit scholar. He published a book and over 12 science papers till now, and there are 1 paper was indexed by SCI, 5 papers were indexed by EI, and 6 papers were indexed by ISTP.

Yumin Ren (1982-), born Henan province of China, graduated in 2005 in Department of Nursing Studies of the Henan University of Science and Technology, and then worked in Luoyang Dongfang Hospital as a nurse. After worked, she researches children disease.