ATLANTIS PRESS
Advances in Computer Science Research (ACSR), volume 76
7th International Conference on Education, Management, Information and Mechanical Engineering (EMIM 2017)

The BP Neural Network Design Applied on the Classification of the Apples

Xianjiang Meng[1, a *] and Xianli Meng[2,b]

[1] School of Communication and Engineering, Shenyang Ligong University, Shenyang, 110159, China

[2] Heilongjiang Branch China Mobile Group Telecommunication Design Institute Co. Ltd, Harbin, 150080, China

[a]76619815@qq.com, [b]kissny@sina.com

Keywords: Neural network; Spectral analysis; Biological classification; BP

Abstract. A kind of artificial neural network (BP-ANN) was used to classify the different kinks of apples with the reflected spectrum. The spectrum (from 460nm to 720nm) of some specks on the surface of the apples was measured with the spectrometer. A kind of BP-ANN with single hidden layer was devised to identify two sorts of the apples (the scar and the rotten) by the biological surface spectrum. To get the optimized parameters of the neural network, it was studied that the performance of the ANN with the different neurons numbers of the hidden layer, the different ranges of the output, the error of network with the different MSE(Minimum Squares Error) goal. Finally a multi-stage ANN was devised to identify the four sorts of apples, the fleckless, the pumped, the scar, and the rotten. 20,20,20 and 20 were selected as training samples respectively, and 10,10,10 and 10 were selected as testing samples respectively. This BP-ANN can achieve 87.5% accuracy if 15% noise was added.

Introduction

The spectral analysis has been focused for many years, especially the spectral analysis technology has been introduced into the biological area[1,2,3]. The main spectral analysis methods are the principal component analysis, the parameter method, and the sample irrelevant method and so on [4]. With the development of neural artificial networks technology, its applications were focused on in recent years, and the research on the spectral analysis with the artificial neural network had ever been discussed [5,6]. The artificial neural network was proved to be an effective method in non-linear recognition and biological applications [7, 8]. In recent years, many researchers has focused on the spectral analysis with neural network in biological engineering, or even medical area. This paper gave a method to identify the visible spectrum of micro areas on the biological surface with the back propagation artificial neural network.

In this paper，First we discusses the recognition of the scar and rotten apples with the BP-ANN, analyses the parameters of the ANN to get a better recognition, and then devise a three stage ANN to classify four types of apples, the fleckless, the pushed, the scar and the rotten. Finally analyses the performance of the system if it is added certain percent of the noise. It is proved that the system could classify the sorts of the apples with high accuracy, even with high percent of noise.

Principle

The principle of classification of different apples is the spectrums of the different are changing with the different sorts of the apples. If the apples are not the normal, the reflected spectrum of the speck apples over the normal one must change. The type of the speck is different, the spectrum diffracted is different. The surface characteristics is changing according to the type of the speck, so the reflected intensity accumulated is different. We could identify the type of the apples through the detection of the relative reflected intensity.

The apple's surface can be considered as a diffuse reflection. Its reflected spectral rate is:

$$\beta(\lambda) = \alpha(\lambda).\frac{\rho_{sp}(\lambda)}{\rho_{st}(\lambda)} \quad (1)$$

$\rho_{st}(\lambda)$ is the reflected spectrum of the standard White Plate, $\rho_{st}(\lambda)$ is the reflected spectrum of the speck apples, $\alpha(\lambda)$ is an adjust value, it can be gained through the measurement of the standard White Plate with more precise instrument.

We can get the reflected spectral rate of the speck apple micro area $\beta_{sb}(\lambda)$, the reflected spectral rate of the normal apple micro area is $\beta_{so}(\lambda)$, so the relative reflected spectral ratio $\beta_{sr}(\lambda)$ is:

$$\beta_{sr}(\lambda) = \frac{\beta_{sb}(\lambda)}{\beta_{so}(\lambda)} \quad (2)$$

The spectrum of the relative reflected spectral ratio will be used to classify the different sorts of the apples.

Experiments

The measuring system is shown in Fig. 1, the light of the light source irradiates on the surface of the apples, the reflected light go into the grating monochromator. The grating in the monochromator is driven by the step motor, the refractive light is received by the windows of the photomultiplier, and then the signal is converted into the current. The current is amplified and converted into digital signal by A/D, and then put into computer for data processing, the data was sampled every 5nm.A bromide-tungsten light (color temperature is 2856K) was used as the standard light source, The Supply voltage is 7.9V, and the current is 7A.

The neural networks are simulated on the computer (Intel Core i5 3.7GHz, RAM 4G, Windows 10). The BP-ANN programs are accomplished by the neural networks tool box of MATLAB 7.0 from Math Works.

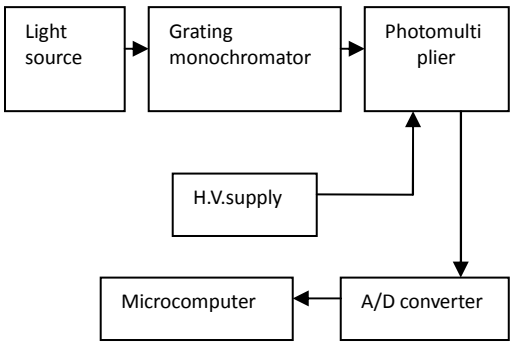

Figure 1. The testing system diagram

In this paper, the data was obtained in the wavelength from 460nm to 720nm, every sample at an interval of 5nm, so 53 samples all together. The neural networks have a single hidden layer, the output has 1 node. We use such a BP-ANN to classify the different types of the apples.

Data Processing and Analysis

The Micro Area Spectrum of Four Sorts Apples. From 460nm to 720nm, four kinds of apples' micro area spectrum, the speckless, the rotten , the scar and the bumped are measured. From the spectrum it can been seen that the speckless apples spectrum are similar, the relative intensity ratio are all close to 1, To the rotten apples spectrum, the relative intensity ratio are below 1, the relative intensity ratio becomes bigger from 460nm to 720nm, and there is a crest at the length of 650nm, To the scar apples spectrum, the relative intensity ratio is almost below 1, and the curve almost

linear from 460nm to 720nm, to the bumped apples spectrum, the relative intensity ratio is above 1, and with the increasing of wavelength, the ratio becomes bigger. We can know there is a little difference among four kinds of apples, we can not classify them with people's eyes, but we can distinguish different kind of apples with the neural network according their spectrum.

The Single Stage Neural Network. From the analysis of the apples spectrum, the spectrum of the scar and the rotten are the closest, so we chose these two kinds spectrum to establish ANN to classify them. To determine the parameters of the neural network, we analyze establishing process of the network to classify the scar and the rotten.

The Influence to the Neural Networks Performance by the Number of Hidden Units. In the BP-ANN, the traingdx.m is used as the training function, the transfer function in the hidden layer is sigmoid function. The linear function is used as the output function. The maximum training times is 20000. The MSE (Minimum Squares Error) goal is 0.01. There are 40 input values which were taken as the training samples, and 20 ones were taken as the testing samples.

In table 1, the value is the classifying capability of the networks with the different neurons number of the hidden layer. SSE (Sum Squared Error) and MAE (Mean Absolute Error) are the main parameters.

Table 1 The performances of ANN with the different neuron number of the hidden layer

In

Neuron number of hidden layer	4	5	6	7	8
Training sample' SSE	0.1105	0.1428	0.1629	0.1489	0.0957
Training sample' MAE	0.1530	0.1159	0.0977	0.0918	0.1006
Testing sample' SSE	0.6954	0.2764	0.3564	0.3254	0.3276
Testing sample' MAE	0.3045	0.2057	0.1604	0.2033	0.1065

Table 1, we can see that when the SSE and MAE of the training samples are almost the same with the different nodes of the hidden layer, the error of the testing samples was different. If the node numbers of the hidden layer was 5, the SSE of the testing samples reach minimum, and the network has good recognition and generalization capability. If the training times is 2000, the MSE goal is 0.0, we can get the values in Table 2.

Table 2 The influence to the precision of the ANN by the different numbers of the hidden layer

neuron number of hidden layer	4	5	6	7	8
Training sample' SSE	0.111	0.109	0.110	0.073	0.063

In Table 2, we can see the SSE of training samples is getting smaller with the increasing of the neuron number of the hidden layer, that is to say the networks recognition precision to the training samples is improving. Considering the generalization of the networks, we take 5 or 6 nodes as the hidden layer. If the hidden layer has 5 nodes, and the MSE of the training samples is set 0.01, 0.001, 0.0001 respectively, we can get the SSE, MAE of the training samples and testing samples as Table

Table 3 The error of the network with different MSE

MSE	0.01	0.001	0.0001
Testing sample' SSE	0.1428	0.0265	0.0195
Training sample' MAE	0.1159	0.0983	0.1045
Testing sample' SSE	0.2764	0.5967	0.7231
Training sample' MAE	0.2057	0.2988	0.1566
Times	925	3946	5963

In Table 3, we find that the SSE and MAE of the training samples is decreasing with the decreasing of the MSE, but the SSE and MAE of the testing samples, and the cycle times is also increasing. It can be concluded that we can not get both in generalization capability and recognition precision, but we can optimize between them.

Table 4 The performances of ANN with different output region

Output region	[0.1 0.9]	[0.3 0.7]	[0.4 0.6]	[-1 1]
Training sample's	0.1595	0.1577	0.1462	0.2562
Training sample' MAE	0.0716	0.0857	0.0766	0.0932
Testing sample' SSE	0.3504	0.1636	0.1285	2.4402
Testing sample' MAE	0.1628	0.1132	0.0925	0.4182
Times	743	103	20	2000
accuracy	100%	96.15%	42.30%	100%

The Influence to the Neural networks Performance by the Output Region. Theoretically the ANN can complete any nonlinear classification if a sigmoid is used between the input layer and hidden layer, and a linear function is used between the hidden layer and output layer. Now we create a neural network with 5 nodes in hidden layer, MSE is 0.01, the maximum training times is 2000. The output target is [0.1 0.9], [0.3 0.7], [0.4 0.6] and [-1 1]. The result of new ANN to the testing samples is shown in Table 4.

In Table 4, the SSE and MAE of all samples decrease When the output region width decreases, but with the decreasing of the output region width, the recognition capability of the ANN becomes lower. As the weights are initialized at [-1 1], the wider is the output region, the bigger is the variation of the weights, the more times needed to train the networks, but recognition capability is improving. So the networks performance is greatly influenced by the output region.

The Establishment of Three-stage BP Neural Network and Its Performance. In order to classify four types of the apples with good performance, a three-stage ANN is established, as in Fig. 2. In each stage, the neural networks have a single hidden layer, the hidden layer has 5 nodes, the input layer has 53 nodes, and the input is the relative reflected spectral ratio of the certain kind of speck apple spectrum over the normal one, the output layer has 1 node, the output target is [0.1 0.9]. The four sorts of apple samples, the fleckless, the pushed, the scar, and the rotten, 20,20,20 and 20 are selected as training samples, and the samples 10,10,10 and 10 respectively are selected as testing samples.

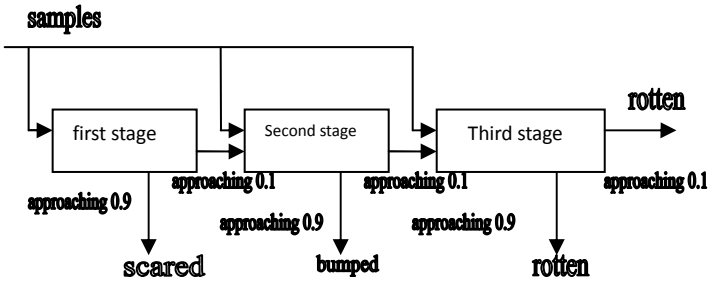

Figure 2. The three-stage ANN

The networks has 5 neurons in the hidden layer, [0.9 0.1] is the output region, the MSE goal is 0.01, 80 samples were used to train the networks, and 40 samples were used as testing samples. When there is no noise, we get the accuracy 100% of all sorts of the apples. When we add 1%, 5%, 10%, 15% of the noise to the input, we get the accuracy of the networks to the testing samples were 100%, 97.5%,92.5%,87.5%, as shown in Table 5. So we can say this BP-ANN can achieve 87.5% accuracy if 15% noise was added.

Table 5 The accuracy under certain percent of noise

Noise Percentage	1%	5%	10%	15%
Error judgment Number	0	1	3	5
Correct judgment Number	40	39	37	40
Percentage of accuracy	100%	97.5%	92.5%	87.5%

Conclusion

As discussed above, the BP-ANN have good nonlinear recognition capability, it is can be used as a method to identify the spectrum of biological surface. So this analysis method did some contribution to the research of the neural network and the spectrum-analyzing technology in the biological engineering.

Acknowledgements

This paper is funded by the National Nature and Science Fund of China. Thanks to my workmate PHD. Weijie Shi's cooperation.

References

[1] Aparecido Carlos Gonçalves, Linilson Rodrigues Padovese: Identification of Lubricant Contamination by Biodiesel Using Vibration Analysis and Neural Network, Industrial Lubrication and Tribology,Vol.64 No.2,2015.pp. 68-75.
[2] Wang, J.Paliwal: Generalisation Performance of Artificial Neural Networks for Near Infrared Spectral Analysis, Biosystems Engineering, Vol.94,No.1, 2012, pp. 1537-1541.
[3] Pang Tao-tao,Yao Jian-bin and Du Li-ming: Artificial Neural Networks for the Identification of Infrared Spectra of Ilex Kudingcha, Spectroscopy and Spectral Analysis ,China, vol.27,No.7,pp.1336-1339, 2007,pp. 1537-1541 .
[4] Appadu, Appanah Rao: Applications and Spectral Analysis of some Optimized High Order Low Dispersion and Low Dissipation Schemes, APPLIED MATHEMATICS & INFORMATION SCIENCES,Vol.8:No.3, 2015,pp. 993-1001.
[5] Irmak A., Jones J W.,Batchelor,W D., Irmak, S,Boote and K J Paz: Artificial Neural Network Model as a Data Analysis Tool in Precision Farming, Transactions of the ASABE - American Society of Agricultural Biological Engineers, Vol.49,No.6, 2013, pp. 2027-2032.
[6] Kovalenko Igor V, Rippke Glen R, Hurburgh and Charles R: Dimensionality Reduction of Near Infrared Spectral Data Using Global and Local Implementations of Principal Component Analysis for Neural Network Calibrations , Journal of Near Infrared Spectroscopy,Vol.15 ,No.1, 2014, pp. 328-335.
[7] Hirane Tatsuya Toryu, Tetsuya Nakano Hidehiro and Miyauchi Arata: Evolutionary and Other Biological Inspirations - Analysis for Characteristics of GA-Based Learning Method of Binary Neural Networks, Lecture Notes in Computer Science, Vol.36, 2011,pp. 325-330.
[8] Guo Shangjiang, Huang Lihong: Biological physics - Stability Analysis of a Delayed Hopfield Neural Network, Physical Review - Section E - Statistical Nonlinear and Soft Matter Physics ,Vol. 7,No.6, 2013,pp. 1540-1545.