Chapter 89
Throat Polyp Detection Based on the Neural Network Classification Algorithm

Shan Qin, Baoju Zhang, Wei Wang, and Sijie Cheng

Abstract This paper realizes the judgment that whether patients have throat polyp by normalization processing, principal component analyzing and Neural Network Classifying the extracted audio data. This implementation replaces the traditional approach to diagnosis of throat polyps. Conventional laryngoscopy need to cutout, clamp or puncture from the patient to remove the lesions to do pathological examinations, which is so hurt to the patient. The test for throat polyp prediction with the neural network classification algorithm are carried out. The results shows that the correct rate of prediction is stable under different number of samples and different random measurement matrices.

Keywords Principal component analysis • Neural network • Normalization

89.1 Introduction

Throat polyp detection is a field that demands more investigation. It is common to have throat polyps and to be completely unaware of them, particularly if they are fairly small. Traditionally, the methods of diagnosis are indirect laryngoscope, video-laryngoscope, and stroboscope light [1]. These polyps then break off and disappear inside the body or clear up by themselves. However, throat polyps can increase in size to the extent that they affect a person's ability to speak. Furthermore, most of these methods need special instrument, and mainly depend on the experience of the pathologists. Also, the patients will feel uncomfortable pain usually. It would be desirable if throat polyps could be detected based on the patient vowel voices only [2].

Traditional pattern recognition techniques such as Bayesian classifier, known as the optimal classifier, could be used if the voice samples follow certain distribution, and this belongs to model-based statistical processing. In [3], the statistical characteristic root-mean square-delay spread and standard deviation were employed to

S. Qin (✉) • B. Zhang • W. Wang • S. Cheng
College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China
e-mail: videolab@163.com

describe the speech frequency domain characteristic and used as two antecedents. The Fuzzy logic system was used to make polyp patients' diagnosis. The results demonstrated that the proposed method could detect the throat polyps with low prob-48 ability of miss detection and 0 % false alarm rate. In [4], some methods of speech analysis for the diagnosis of the laryngeal function have been discussed. In humans' voices, the voice amplitude is highly bursty, and we believe that no statistical model can really demonstrate the uncertain nature of the voice [5].

Because of the complexity and unpredictability of voice, the data and information in many aspects, such as analysis, making discissions are non linear connection and complex. As an important branch of artificial intelligence, Neural network implements a mapping function from the input to the output, which is proved by mathematical theory that three layer neural network can approximate any nonlinear continuous function with arbitrary precision [6]. This makes it particularly suitable for solving complex problems, namely that has strong nonlinear mapping ability.

This paper builds a BP neural network, which realizes the judgment that whether patients have throat polyp. Some popular industry technologies, including normalization processing, principal component analyzing and neural network classifying are appropriately combined.

89.2 Theory

89.2.1 Principal Component Analysis

There are random variables $X_1, X_2, \ldots, X_p$, whose standard deviations of the sample is recorded as $S_1, S_2, \ldots, S_p$. First, standardization transformation is $C_j = a_{j1 \times 1} + a_{j2 \times 2} + \ldots + a_{jp \times p}$, $j = 1, 2, \ldots, p$. We have the following definitions:

If $C_1 = a_{11 \times 1} + a_{12 \times 2} + \ldots + a_{1p \times p}$, and $Var(C_1)$ is the biggest, C_1 is called the first principal component;

If $C_2 = a_{21 \times 1} + a_{22 \times 2} + \ldots + a_{2p \times p}$, and $(a_{21}, a_{22}, \ldots, a_{2p})$ perpendicular to the $(a_{11}, a_{12}, \ldots, a_{1p})$, and $Var(C_2)$ is the biggest, C_2 is called the second principal component;

Similarly, there is a third, fourth, fifth ... the main ingredient, and at most p points.

Principal component analysis is a statistical approach allows to reduce the dimensionality, which is implemented by an orthogonal transformation translating the related component of the original random vector into uncorrelated components of new random vector. On the respect of algebra, it means that the covariance matrix of original random vector is transformed into a diagonal matrix. In geometry, it means that the original coordinate system is transformed into a new orthogonal coordinate system, which point to the P orthogonal directions of the

sample points that spread most open. Then, the multidimensional variable system will reduce the dimension [7]. The math algorithm is as follows.

P dimensional random vectors of standard collection of the original data is $x = (x_1, x_2, \ldots, x_p)^T$, n samples is $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^T, i = 1, 2, \ldots, n$

When $n > p$, construct the sample matrix, and make the standard transformation to the sample array element as follows:

$$Z = \frac{x_{ij} - \bar{x}_j}{s_j}, i = 1, 2, \ldots, n; j = 1, 2, \ldots, p \tag{89.1}$$

$$\bar{x}_j = \frac{\sum_{i=1}^{n} x_{ij}}{n}, s^2 j = \frac{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2}{n - 1} \tag{89.2}$$

and the standardization matrix is Z.

Second, demand the correlation coefficient matrix of Z.

$$R = \left[r_{ij} \right]_p xp = \frac{Z^T Z}{n - 1} \tag{89.3}$$

$$r_{ij} = \frac{\sum z_{kj} \cdot z_{kj}}{n - 1}, i, j = 1, 2, \ldots, p \tag{89.4}$$

Third, demand the characteristic equation of sample correlation matrix R, which is $|R - \lambda I_p| = 0$, have p characteristic roots to determine the main ingredient. Then, according to $\frac{\sum_{j=1}^{m} \lambda_j}{\sum_{j=1}^{p} \lambda_j} \geq 0.85$ to determine the value of m, and ensure the utilization rate of information is more than 85 %, Solution of equations, $Rb = \lambda_j b$, for each λ_j, $j = 1, 2, \ldots, m$, get the unit eigenvector b_j^o.

Last, conversion the normalized indicator variables to main component $U_{ij} = z_i^T b_j^o, j = 1, 2, \ldots, m$, U_1 is the first main ingredient.

89.2.2 Neural Network Algorithm

Artificial neural network has the characteristics of self-adaption, self-organization and self-learning. It is already well known that an ANN consists of a number of artificial neurons and connections among them. An artificial neuron is generally regarded as a nonlinear device with multiple inputs and a single output. An ANN model is shown in Fig. 89.1.

Where $x_n(t)$ is the output of the n‑th neuron at time t which is also the n‑th input to the i‑th neuron at the same time, w_{in} is the weight representing the connection strength between the n‑th and i‑th neuron, net(t) is the net total input to the i‑th neuron at time t, $a_i(t)$ is the activation of the i‑th neuron at time t which is a function
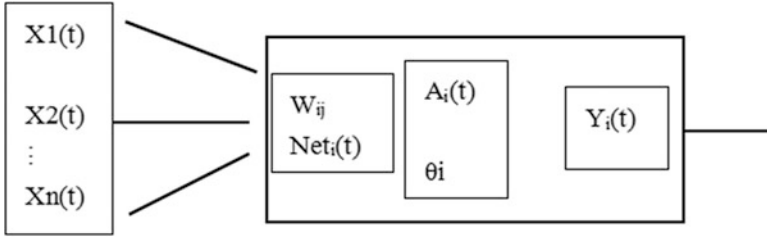
Fig. 89.1 An ANN mode

of $net_i(t)$, θ_i is the threshold of the i‑th neuron, and $y_i(t+1)$ is the output which is a nonlinear function of $a_i(t)$ and θ_i, as shown in Eqs. 89.5, 89.6, and 89.7 respectively [8].

$$net_i(t) = \sum_{n=1}^{N} W_{in}x_n(t) \qquad (89.5)$$

$$a_i(t) = g(net_i(t)) \qquad (89.6)$$

$$y_i(t+1) = f(a_i(t), \theta_i) \qquad (89.7)$$

BP neural network algorithm is the most widely used, which is based on the error back propagation to adjust the network weights and thresholds constantly, and establishes a network with the minimum of sum of squared errors. Actually, the data flow information is positive, but the error propagation is reversed [9].

89.3 Experiment and Simulation

MATLAB is a useful industrial, educational and research tool, which can enough to help users find what can be done and what not to do, which can help develop and broaden the field of neural networks work. In this lab, each person has two sound samples, /a:/and /i:/. In MATLAB, the function of premnmx is used to normalize the original data between −1 and 1, and the function of princomp is used to extract the characteristic values. We take five characteristic values from /a:/and /i:/ sample of each person in a column as one input of net. There are 18 training samples, and 15 testing samples [10].
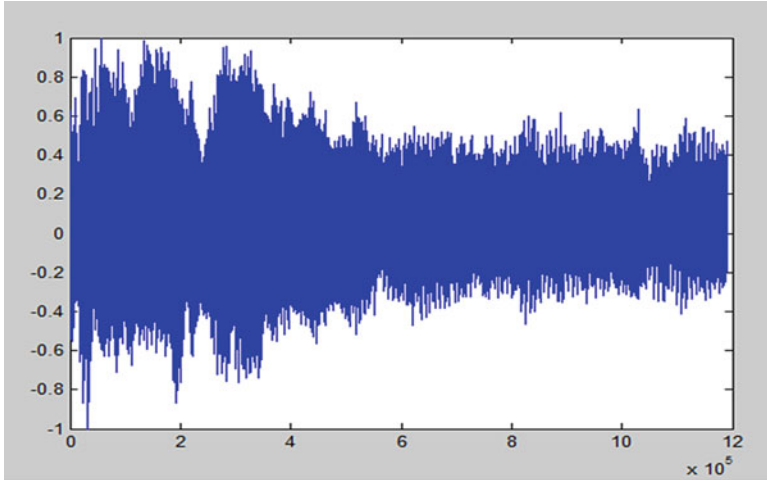
Fig. 89.2 The data distribution of data normalization

89.3.1   The Result of Experiment

89.3.1.1   The Result of the Data Normalization

The main idea of data normalization is to define the data within a fixed range, by setting up a normalization factor, which is subtracted by the original data, and at same time, to eliminate systematic errors in the experimental data. In this experiment, the collected data samples of voice normalized between -1 and 1. The normalized data will be very convenient, and ensure a faster convergence during running programs in subsequent processing. Figure 89.2 is a sample in which the data distribution of data normalization.

89.3.1.2   The Result of Principal Component Analysis

In the samples, the original data is a matrices of n rows and 1 column, after matrix deformation, they become a matrix of m rows and 100 columns. Principal component analysis was performed on the raw data, and we get the triangular matrix of 100 rows and 100 columns and achieve the effect of reducing the dimensions. In the MATLAB, we can obtain the characteristics of each vector value directly.

Dimension reduction will cause a loss of information. However, the loss of information is rarely. Because the main part of the information is extracted. Figure 89.3 shows the main ingredients of one sample.
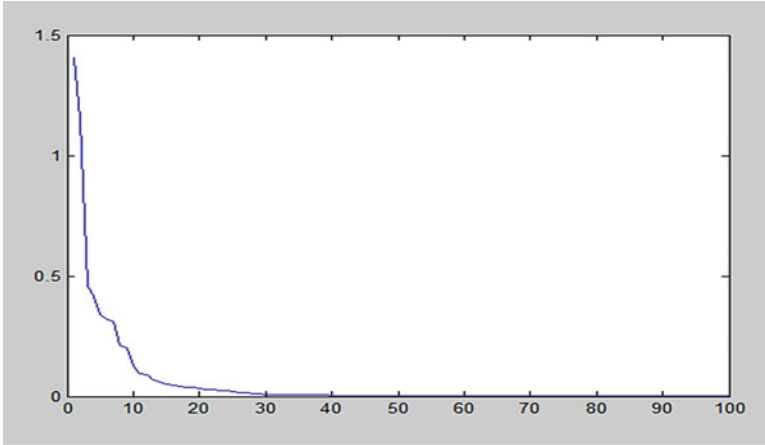
Fig. 89.3 The main ingredients of one sample

89.3.1.3 The Result of Neural Network Classifying

The experiment choose Back Propagation network algorithm, which is the most widely used. And learning rules is quasi-Newton method.

Compared with the standard steepest descent method, quasi-Newton method has a high convergence speed in the vicinity of the area closed to the optimal solution, and we can improve the learning speed of neural networks. In addition, iterative direction of Quasi-Newton method has conjugation, which has a limited secondary termination. In fact, this property is one of evaluation standards to judge a algorithm whether is good or bad. If a convergence algorithm does not have this property, then it would be difficult to have the super linear convergence speed.

The experiment uses eight-layer hidden layer. Increasing the number of layers can reduce the learning error and improve the training accuracy. But, network becomes complex, and training time of the network weights will be longer.

The expected outputs of training sample is [1,1,1,1,1,1,1,1,1,1,0,0,0,0,0,0,0,0,0,0], and the actual output is [0.9975, 0.9973, 0.9973, 0.9975, 0.9974, 0.9961, 0.9902, 0.9975, 0.9975, 0.9971, 0.0611, 0.0339,0.0200, 0.0560, 0.0339, 0.0339, 0.0129, −0.0103], in which illness output is 1, not illness output is 0. The result of training is shown in Figs. 89.4 and 89.5.

89.3.2 The Result of Simulation

There are 15 testing samples. The expected output is [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0], and the real output is [0.9975 0.9975 0.9967 0.9797 0.0339 0.9975 0.0070 0.0273 0.9267 0.9975 0.9922 0.0069 0.9975 0.0345 0.3247]. The result is shown in Figs. 89.6 and 89.7.
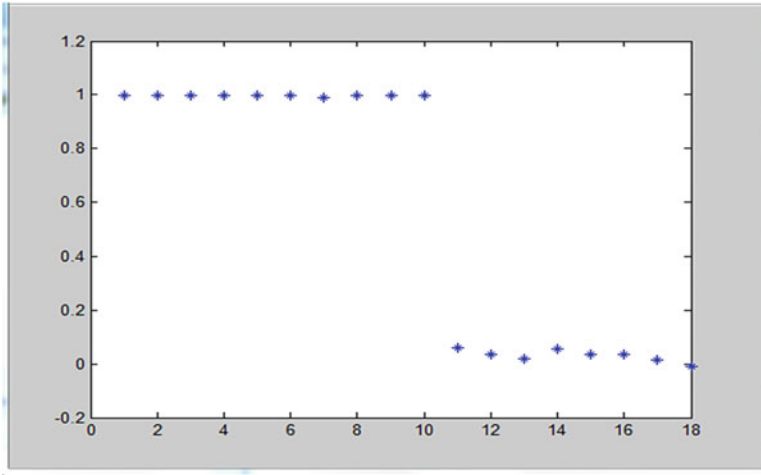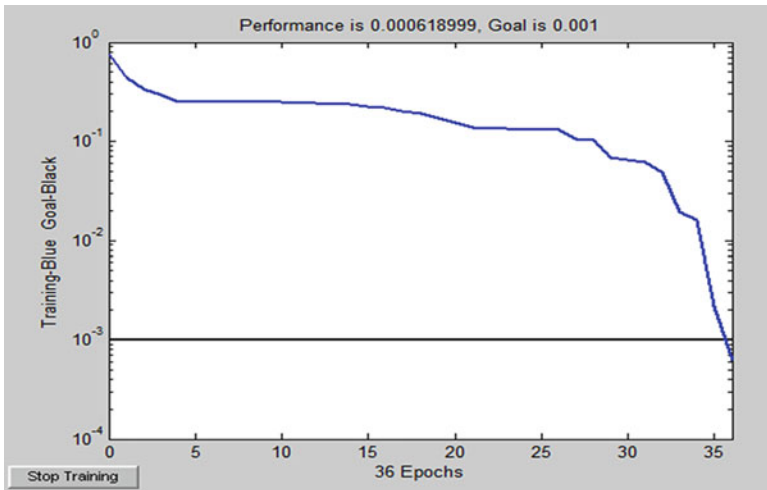
Fig. 89.4 The output of training

Fig. 89.5 The target error of training

89.3.3   Error Analysis

The results of simulation displays that the correct rate of prediction is believable under different number of samples and different random measurement matrices. By calculating, It can be drawn that the of the sick can reach 70 %, and the judgment accuracy of the health rate is up to 60 %. So, the judgment accuracy of wether is sick or health is about 67 %.
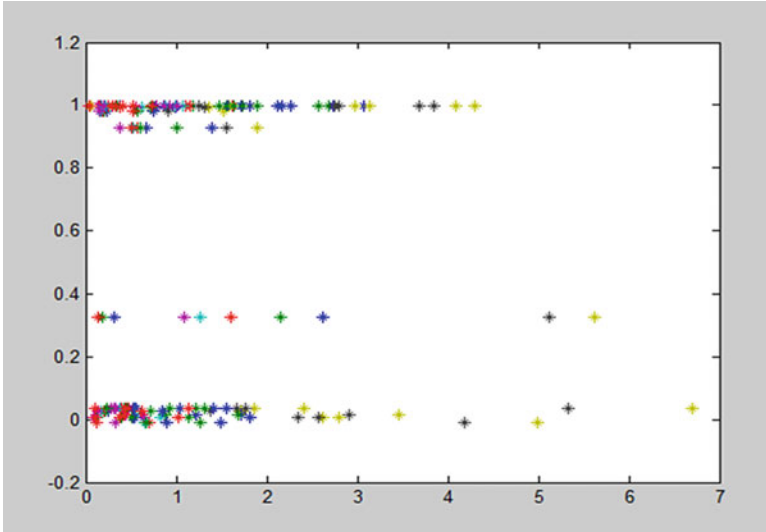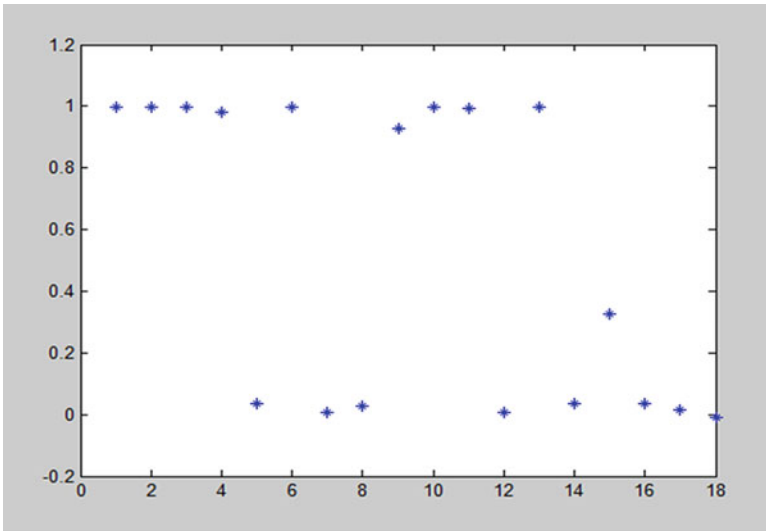
Fig. 89.6 The result of simulation of 10-dimensional

Fig. 89.7 The result of simulation of 1-dimensional

Conclusion

In this experiment, we analyze the data, extract its characteristic values, and then use the neural network learning to train the samples. We get the classification data. The simulation results of test samples are also ideal, indicating that the experiment can be replicated in reality to alleviate the pain of patients in the process of diagnosis.

The results shows that the correct rate of prediction is stable under different number of samples and different random measurement matrices. But However, more voice data should be sampled in order to reach a better diagnosis result, and the test accuracy still need to be improved by improved algorithm. We will continue to study in the future.

References

1. Zhong Z, Chen Z, Liang Q, Xiao S (2012) Throat polyps detection based on patient voices. Lect Notes Electr Eng 202:531–539
2. Wei Wang, Zhangliang Chen, Jiasong Mu, Tingting Han (2014) Throat polyp detection based on compressed big data of voice with support vector machine algorithm. EURASIP J Adv Signal Process. doi:10.1186/1687-6180-2014-1
3. Zhong Z, Jiang T, Zhang WS, Yao H (2010) Analysis speech of polypus patients based on channel parameters and fuzzy logic systems. In: Proceedings of the 2010 seventh international conference on fuzzy systems and knowledge discovery, Yantai, pp 529–532
4. Choi JM, Sung MW, Park KS (2002) New method in acoustic analysis for the diagnosis of the laryngeal functions. Proc Second Jt EMBS/BMES Conf 10:135–136
5. Budhaditya S, Pham D-S, Lazarescu M, Venkatesh S (2009) Effective anomaly detection in sensor networks data streams. In: Ninth IEEE international conference on data mining, pp 722–727
6. Wei Wang et al (2014) Intelligent throat polyp detection with separable compressive sensing. EURASIP J Adv Signal Process 2014:6
7. Pham D-S, Venkatesh S, Lazarescu M, Budhaditya S (2012) Anomaly detection in large scale data stream networks. Data Min Knowl Discov. doi:10.1007/s10618-012-0297-3
8. Dardari D et al (2008) Threshold-based time of arrival estimators in UWB dense multipath channels. IEEE Trans Commun 56(8):1366–1378. doi:10.1109/TCOMM.2008.050551
9. Zhong yixin, Chi huisheng (1992) A survey of artificial neural networks. Acta Electron Sin 10:20–25
10. He qingbi, Hu jiuyong (2004) An overview of data mining. J Southwest Univ Natl 29:328–330