Cancer Diagnosis Using Artificial Neural Networks

M. Baradaran Nia, Sh. Shogian and M. H. Zarifi

Department of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

Summary
Artificial Neural Networks are used to analyze electrical impedance spectroscopy (EIS) data taken from patients bladders. Using this system, malignant areas from non-malignant areas in the urinary bladder of the patient can be separated very rapidly. Four different structures of artificial neural networks (ANNs) are used. The results show that both LVQ4b and SFAM networks show minimum error rate equal to 9.52% in test phase. Also using the MLP network make it possible to have no-prediction state which makes it easy to deal with critical practical applications.

Key words:
Artificial neural networks, electrical impedance spectroscopy, bladder cancer diagnosis

1. Introduction

Artificial neural networks are simplified models of the biological nervous system and therefore have drawn their motivation from the kind of computing performed by a human brain. An ANN is a massively parallel distributed processing system made up of highly interconnected neural computing elements that vase the ability to learn and thereby acquire knowledge and make it available for use [1]. Since detecting cancerous area from the normal area using data obtained by electrical impedance spectroscopy has a close relation with soft computing, it gives us the idea of using artificial neural network as an appropriate tool.
Electrical impedance spectroscopy (EIS) is a non-invasive screening technique to separate malignant areas from non-malignant areas in different living tissues such as the urinary bladder. This is a result of the electrical impedance spectrum of the tissue being a function of tissue structure at the cellular level. Electrical impedance spectroscopy technique involves driving electrical currents through electrodes into the body, measurements of the resulting potentials by the other electrodes and then calculation of the transfer impedance. There are several researchers working in this field such as a study which concerns the relation between tissue structures and imposed electrical flow in cervical neoplasia carried out by Brown et al. to compare the impedance of normal and abnormal cervical tissues [2]. Another study investigated virtual biopsies in Barrett's Oesophagus using electrical impedance measurements [3]. The aim of their study was to show the possibility of differentiating two types of epithelia (squamous and columnar) in terms of their electrical impedances. Furthermore, they have considered the inflammation effects on Barrett's oesophagus using low frequency system [4]. Stray capacitance is present between all connecting cables of the probe especially between pairs of drive and receive cables and its effect must be reduced to achieve accurate measurements. Such evidence provides the rationale for studying more about the electrical properties of biological tissue to understand the detection of bladder cancer. For this purpose it will be better to discuss the electrical properties and impedances of living tissue. Furthermore, EIS involves driving electrical currents through electrodes into the body, measurement of the resulting potentials by other electrodes and then calculation of the transfer impedance. Different tissues may have distinguishing characteristics in the shape of the impedance spectrum.
In this paper, the data acquired from EIT systems is used to train and test four different supervised ANN structures for diagnosis bladder cancer. Two error-based and two prototype-based networks are studied. Each ANNs' error rate, the rate of wrong decisions in the test phase, is minimized by adjusting their parameters by try and error. Then they are compared with their error rate in the test phase by considering the error-based network capability of making a third choice as no-prediction.
The results obtained from comparison of these four different ANN structures, also their compatibility of implementation on a Field Programmable Gate Array (FPGA) which has great potential to be used in EIT systems due to its characteristics such as producing fast prototypes of complex hardware designs, fast tests, modification accomplishments, and being updated by means of single software modifications with an effective production cost, motivate us to focus our future work on implementation the ANNs on an FPGA for bladder cancer detection.

2. Electrical Impedance Spectroscopy

2.1 Tetra-polar technique

The most common form of measuring tissue impedance is the tetra-polar or 4-electrode technique. This technique can

Manuscript received July 5, 2008.
Manuscript revised July 20, 2008.

measure transfer impedance (the ratio of measured voltage to applied current) of the urinary bladder. In the tetra-polar technique, a known current is driven between two electrodes and the resulting voltage is measured between the other two electrodes. This is used because it is designed to minimize the effect of electrode impedance. The principle is to use a constant current source with infinite output impedance and a voltage-measuring amplifier with infinite input impedance. Because no current flows across the receiving electrodes in the 4-electrode measurement technique, the electrode impedance does not contribute to the measurements and thus the measurement is a true reflection of the tissue transfer impedance.
Both the delivered current and recorded voltage are supplied using a 4-electrode probe applied to the urothelium of the target bladder tissue. The technique requires that the probe be placed on the surface of the bladder tissue.

2.2 EIT System Architecture

The architecture of an EIS system is illustrated in Fig. 1. It consists of two isolated sections, the probe circuit at the patient side and the main interface circuit to process and import the collected data to a host computer. The building blocks of the system are a current source circuit, a digital-controlled oscillator, a four-electrode probe, an instrumentation amplifier, and a phase-sensitive demodulator. The system is controlled via an I/O data acquisition card by a routine implemented in LabView running on the host PC.

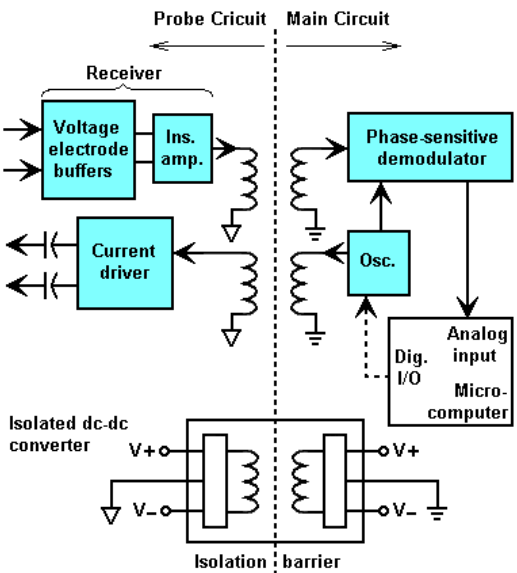

Fig. 1 EIS system architecture

3. Artificial Neural Network

3.1 Review

An Artificial Neural Network (ANN) is a computational model which its performance is derived from human brain. It is an information-processing system that has certain performance characteristics in common with biological neural networks. An ANN consists of some nodes which are connected via weights. Each node receives data from behind nodes, adds it and passes data through a nonlinear function, and then propagates data to proceeding nodes. There are two phases in the ANN performance: training phase and test phase. In training phase, input patterns are presented to the ANN and weights are adjusted and fixed to learn these patterns. Indeed, the ANN learns input patterns in learning phase. In the test phase, the patterns which are not used in training phase are presented to the ANN and the ANN's outputs are used to evaluate its performance [5]. If the evaluation of ANN's performance is satisfactory, it can be used in its own specific application.

3.2 Artificial Neural Network Structures

Neural Networks have been used in a wide range of applications, including: pattern classification, pattern completion, function approximation, optimization, prediction and automatic control. In a general view, ANNs can be categorized into two groups as supervised and unsupervised learning. If the outputs of the input patterns used in the training phase of the ANN are available via a specific experiment, the ANN is said to be supervised and otherwise the ANN is said to be unsupervised. The patients to whom the data is available through electrical impedance spectroscopy technique were examined to determine that their cancerous cells are malignant or nonmalignant. So supervised learning networks are used here.
Also, supervised ANNs can be divided into two groups: error-based and prototype-based. In the error-based network the main effort is to minimize the cost function which is defined in the basis of error between the desired output and the network output. The prototype-based network tries to minimize the distance between the inputs patterns and the prototypes which are assigned to each cluster. The MLP and RBF networks as the error-based networks and the SFAM and LVQ networks as the prototype-based networks have been used is this research and the results are compared to choose the best compatible network for this case study.
Among all the various supervised neural network structures, the Multilayer Perceptrin (MLP) network has the highest practical value. It is a feed-forward layered network with one input layer, one output layer, and some

hidden layers [5]. The method for training the MLP is based on the minimization of a suitable cost function, and is called the backpropagation algorithm. The first version of this algorithm, which was based on the gradient descent method, was independently proposed by Werbos [6] and Parker [7]. Many modifications have been put forward by others. In this paper two variants of the backpropagation algorithm are used: backpropagation with plummeting learning rate factor (BPLRF) and backpropagation with declining learning-rate (BPDLR). For more information about these algorithms refer to [8-9].
The construction of a Radial Basis Function (RBF) network, in the most basic form, involves three layers with entirely different roles. The inputs layer is made up of source nodes that connect the network to its environment. The second layer applies a nonlinear transformation from the input space to the hidden space; in most applications the hidden space is of high dimensionality. The output layer is linear, supplying the response of the network to the activation pattern applied to the input layer [10-11].
Linear Vector Quantization (LVQ) originally was introduced By Linde et al. [12] and Gray [13] as a tool for image data compression and later was adapted by Kohonen [14] for pattern recognition. The main idea is to divide the input space into number of distinct regions, called decision regions. Like other networks, LVQ network has some variants. Here LVQ1, LVQ2, LVQ3 and different types of LVQ4 have been used. For more information about types of LVQ network refer to [15].
Simplified Fuzzy Artmap (SFAM), the abridged model of fuzzy adaptive resonance theory, is a prototype-based network which can handle both binary and analogue data in a supervised manner. In addition to the high practical potential of the SFAM network, its intricacy prevents the others from using it.

3.3 Data Acquisition and Arrangement

Data set is obtained from measuring the urinary bladder of 63 suspected cases. These measurements are executed in 7 frequencies. 42 suspected cases have nonmalignant areas and the others have malignant areas. So we can divide data set into two groups, first one represents a 42×7 matrix corresponding to the impedances of nonmalignant areas of 42 cases in 7 different frequencies and the other represents a 21×7 matrix corresponding to the impedances of malignant areas of 21 cases in the same 7 frequencies.
To apply the data set to each ANN, firstly we arrange it as a matrix which we call it data matrix. The data matrix has 63 rows, each row showing the number of suspected cases and 7 columns, each column showing corresponding frequency in which we have measured suspected cases' impedances. The component of the data matrix demonstrates measured impedances. Arrangement of the data matrix's rows is as below: Two first rows of this matrix represent the impedances corresponding to the first and second suspected cases in the first group (nonmalignant group), third row of the data matrix represents the impedances corresponding to the first suspected case in the second group (malignant group), fourth and fifth rows of the data matrix represents the impedances corresponding to the third and fourth suspected cases in the first group, sixth row of the data matrix represents the impedances corresponding to the second suspected case in the second group and so forth. So we have a data matrix with dimension 63×7. We use 42 rows of this matrix as train set for training ANN and 21 rows of this matrix as test set to test implemented ANN.

4. Result and Discussion

Four different ANN structures are used to be trained by the data matrix. Two of these structures, MLP and RBF, are error-based and the others, SFAM and LVQ, are prototype-based networks. Two variants of backpropagation algorithm, BPLRF and BPDLR, are used to train the MLP network. Also six different types of LVQ network are used. The number of epochs in training phase for all networks is 100 except the SFAM network which converges in third epoch. For error-based networks the number of neurons in output layer is chosen to be equal to the number of clusters i.e. 2. The desired output for malignant cases is $[0$ and for nonmalignant cases is $[1$. The absolute difference between the elements of the network output can be used to categorize the input pattern into three states: malignant, nonmalignant and no-prediction. If the absolute difference between the output elements is less than a predefined value (like), the pattern input is categorized as no-prediction. It is clear that no-prediction state can help us to make critical decisions in practical cases. In prototype-based networks no-predication state is not available. By choosing $\gamma = ($, MLP's evaluation in test phase shows 3 wrong decisions in which one of them is no-prediction state. The RBF networks exhibits 8 wrong decisions in which 3 of them are no-prediction states. Fig. 2 illustrates the Total Sum Square Error (TSSE) curves of the MLP and RBF networks in training phase. In addition to the less value of the TSSE for the MLP in the training phase, it also converges faster than the RBF.

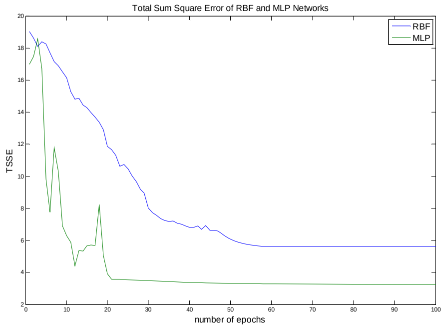

Fig. 2 Total sum square error for RBF and MLP networks

Table I shows the results obtained from different types of LVQ networks. The parameters of training phase are available for further simulations. All used LVQ networks are pattern mode except LVQ4b in which weight adjustment is batch mode. The latter LVQ network has minimum error rate in test phase.

Table 1: The Results of LVQ Network

	α_0[a]	α_{end}[b]	ε[c]	TM[d]	ER[e]
LVQ1	0.01	0.001	-	pattern	33.33%
LVQ2	0.01	0.001	-	pattern	33.33%
LVQ3	0.8	0.1	0.25	pattern	14.28%
LVQ4a1	0.1	0.001	-	pattern	33.33%
LVQ4b	0.8	0.01	-	batch	9.52%
LVQ4c1	0.9	0.1	-	pattern	14.28%

[a] α_0: α in the first epoch
[b] α_{end}: α in the last epoch
[c] ε: stabilizing constant factor
[d] TM: Training Mode; the mode of adjusting weights in training phase; pettern: weight adjustment in each iteration, batch: weight adjustment in each epoch
[e] ER: Error Rate; the rate of false decisions to correct decisions in test phase

Table II demonstrates SFAM network parameters through which 19 prototypes in training phase are created.

Table 1: The Results of SFAM Network

	ρ[a]	ε[b]	β[c]	α[d]	ER[e]	NOP[f]
SFAM	0.5	0.001	0.3	0.001	9.52%	19

[a] ρ: vigilance factor
[b] ε: vigilance factor increment parameter
[c] β: learning factor
[d] α: tie-breaker parameter
[e] ER: Error Rate
[f] NOP: Number of Prototypes

Table III summarizes all of the ANN's performances. The minimum error rate in test phase is attained as 9.52% with both the SFAM and LVQ4b networks. These two networks have the minimum error rate and are winners in the sense of minimum error rate. But by taking account that the error-based networks have the capability of making a third choice as no-prediction state, we suggest to use the MLP network with BPLRF and BPDLR algorithms as variant of the main backpropagation algorithm.

Table 1: The Comparison of different ANN Structures

	LVQ4b	SFAM	MLP	RBF
error rate	9.52%	9.52%	14.28%	38.1%

5. Conclusion

We used four different ANN structure to predict the malignancy of the bladder cancer. We suggest using the MLP network for both its acceptable error rate in the test phase and its capability of making a third choice as no-prediction state. So the critical decision in practical cased can be made easily. The future prospect of this research is to focus on system-on-a-chip (SoC) applications using ANN implementation with the purpose of obtaining cancerous cell detection on a single integrated circuit

References

[1] L. Fausett, *Fundamentals of Neural Networks*. Prentice-hall, 1994.
[2] B . Brown , J . Tidy , K . Boston , A . Blackett , R . Smallwood , F . Sharp, "Relation between tissue structure and imposed electrical current flow in cervical neoplasia," *The Lancet*, vol. 355, 2000, pp. 892-895.
[3] C. A. González-correa, B. H. Brown, R. H. Smallwood, N. Kalia, C. J. Stoddard, T. J.stephenson, S. J. Haggie, D. N. Slatter, K. D. Bardhan, "Virtual Biopsies in Barrett's Oesophagus using an Impedance probe," *Annals New York Academy of Sciences*, vol.873, 1999, pp. 313-321.
[4] C. A. Gonzalez-Correa, B. H. Brown, R. H. Smallwood, T. J. Stephenson, C. J. Stoddard, K. D. Bardhan, "Low frequency electrical bioimpedance for the detection of inflammation and dysplasia in Barrett's oesophagus," *Physiological Measurement*, vol. 24, 2003, pp. 291-296.
[5] G.Alizadeh, J. Frounchi, M. Baradaran Nia, S. Asgarifar, and M. H. Zarifi, "An Artificial Neural Network for Prediction of Cetane Number in Diesel Fuel Implemented on a FPGA," in *Proc. IEEE International Conference on Computer and Communication Engineering*, Kuala Lumpur, Malaysia, 2008, to be published.
[6] P.J. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in Behavioral Science," Ph.D thesis, Harvard Univ., Cambridge, MA., 1974.
[7] D.B. Parker, "Learning Logic," Invention Report s81-64, File 1, Office of Technology Licensing, Stanford University, 1982.
[8] M.T. Vakil-Baghmisheh, "Farsi Character Recognition Using Artificial Neural Networks," Ph.D thesis, Faculty of Electrical Engineering, Ljubljana Univ., 2002.
[9] M.T. Vakil-Baghmisheh, N. Pavešić, "Backpropagation with Declining Learning Rate," in *Proc. 10th Electrothechnical and Comp. Scie. Conf.,* vol. B, Portorož, Slovenia, 2001, pp. 297-300.
[10] S. Haykin, *Neural Networks, Comprehensive Foundation.* New Dehli: Prentice Hall of India, 2006.
[11] M.T. Vakil-Baghmisheh, N. Pavešić, "Training RBF network with selective backpropagation," *Neurocomputing*, vol. 62, 2004, pp. 39-64.
[12] Y. Linde, A. Buzo, R.M. Hray, "An algorithm for vector quantizer design," *IEEE Trans. Communication*, vol. 28, 1980, pp. 84-95.
[13] R.M. Gray, "Vector quntization," *IEEE ASSP Magazine*, 1984, pp. 4-29.
[14] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, 1990, pp. 1464-1480.
[15] M.T. Vakil-Baghmisheh,N. Pavešić, "Premature clustering phenomenon and new training algorithms for LVQ," *Pattern Recognition*, vol. 36, 2003, pp. 1901-1912.