

An Artificial Neural Network Hardware for Bladder Cancer

Javad Frounchi

*Microelectronic and Microsensor Research Laboratory
Faculty of Electrical and Computer Engineering
University of Tabriz, Tabriz, Iran
E-mail: jfrounchi@tabrizu.ac.ir
Tel: +98-411-3393722; Fax: +98-411- 3300819*

Ghader Karimian

*Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran
Tel: +98-411-3393722; Fax: +98-411- 3300819*

Ahmed Keshtkar

*Medical Physics Department, Medical School
Tabriz University of Medical Sciences, Tabriz, Iran
Tel: +98-411-3393722; Fax: +98-411- 3300819*

Abstract

An Artificial Neural Network was implemented on a FPGA to analyze electrical impedance spectroscopy (EIS) data taken from a patient bladder. Using this system, malignant areas from non-malignant areas in the urinary bladder of a patient can be separated very rapidly. Two variants of backpropagation algorithm were used to train a multilayer perceptron (MLP) neural network. By adjusting the ANN's parameters, the maximum error and error percentage of test phase in 300 epochs were reduced to 0.33297 and 0.39035%, respectively. The ANN architecture has been implemented on a Virtex-4 LX25 FPGA from Xilinx. The number of occupied slices on the FPGA is 10136 and the design covers 94% of the chip.

Keywords: Artificial Neural Network, Electrical Impedance Spectroscopy, FPGA, Backpropagation.

1. Introduction

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 has 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.

Among all the various NN structures, the MLP has the highest practical value. The method for training the MLP is based on the minimization of a suitable cost function, and is called the backpropagation algorithm. So far, many variants of backpropagation algorithms have been introduced.

In this paper, two variants of the backpropagation algorithm are used as main algorithms: backpropagation with plummeting learning rate factor (BPLRF) and backpropagation with declining learning-rate (BPDLR). To show that these algorithms have better results, some other algorithms are also implemented and results are compared. The BPLRF and BPDLR algorithms have been implemented on a FPGA to detect cancerous area of a target bladder.

FPGAs are being increasingly used for a variety of computationally intensive applications, mainly in the realm of Digital Signal Processing (DSP) and communications [2]. Due to rapid increases in the technology, current generation of FPGAs contain a very high number of Configurable Logic Blocks (CLBs) [3], and are becoming more feasible for implementing a wide range of applications. The high nonrecurring engineering (NRE) costs and long development time for ASICs are making FPGAs more attractive for application specific DSP solutions [4].

In this paper, we have implemented our system on a veirtex-4 LX25 FPGA from Xilinx. The principles of an EIS system and its application to characterize different living tissues are described in section two. Section three deals with multilayer perceptron neural networks and their training techniques. The simulation results of various ANN training algorithms are presented in section four. The implementation of the ANN architecture on the FPGA is described in section five.

2. Electrical Impedance Spectroscopy

Electrical impedance spectroscopy (EIS) is a non-invasive screening technique to separate malignant areas from non-malignant areas in different living tissue 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. Different tissues may have distinguishing characteristics in the shape of the impedance spectrum.

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 [5]. Another study investigated virtual biopsies in Barrett's Oesophagus using electrical impedance measurements [6]. 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 [7]. Finally, the feasibility of adapting this minimally invasive technique to screen for bladder cancer has been explored by some researchers in biomedical field [8-9]. In these studies, the electrical impedance spectrum of the bladder tissue obtained by measuring impedance at different frequencies were used for tissue characterization.

Biological tissues have a complex electrical impedance, which is a function of frequency, and their components have resistive and charge storage (capacitance) properties. The real part of the impedance measurement is used in this work because higher signal to noise ratio (SNR) can be obtained than one obtained if the complex or imaginary part is used. Lu et al in 1996 showed that the simulated imaginary part of the electrical impedance is affected to a greater extent than the real part [10]. 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.

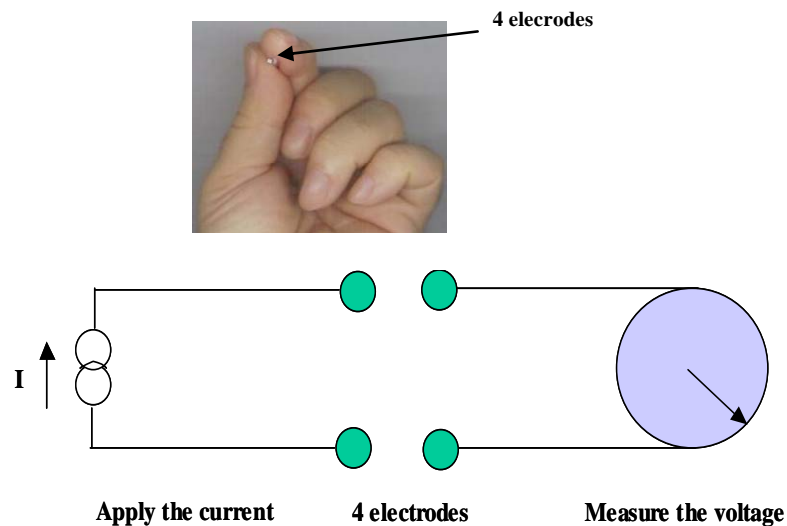
2.1. Tetra-Polar Technique

The most common form of measuring tissue impedance is the tetra-polar or 4-electrode technique. This technique can 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 as shown in figure1. In this measurement technique, the effect of electrode impedance is minimized. 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. In contrast, in the bi-polar or 2-electrode technique the voltage across each electrode is measured in addition to the voltage developed across the tissue. Thus this results in the incorrect analysis of the tissue electrical impedance and it is usually impossible to separate the tissue impedance from the electrode impedance, particularly with small electrodes, which have higher impedance. Therefore, the 4-electrode technique allows more accurate analysis of the measured electrical impedance than the 2-electrode technique.

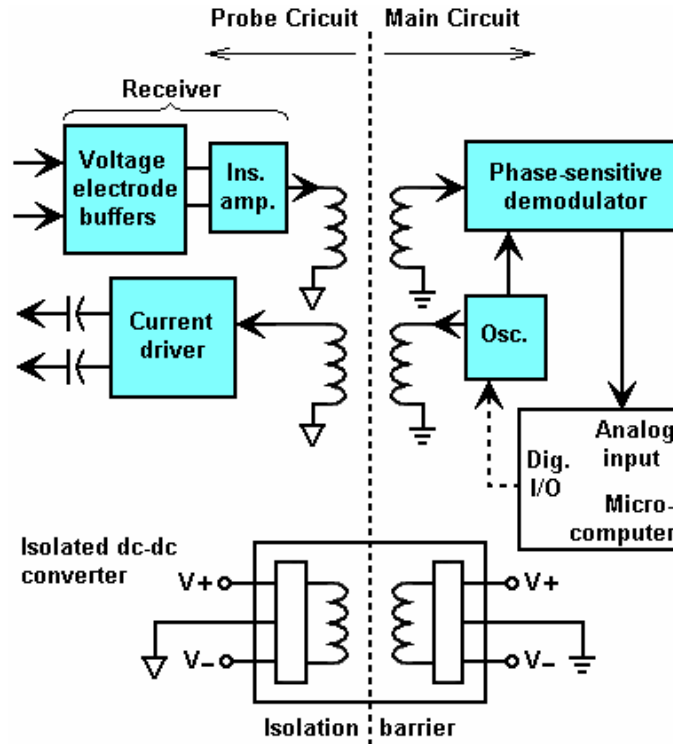
Both the delivered current and recorded voltage are supplied using a 4-electrode probe applied to the urothelium of the target bladder tissue. The technique requires that the probe be placed on the surface of the bladder tissue.

Figure 1: 4-Electrodes and their connections to both the current power supply and the voltage measurement system



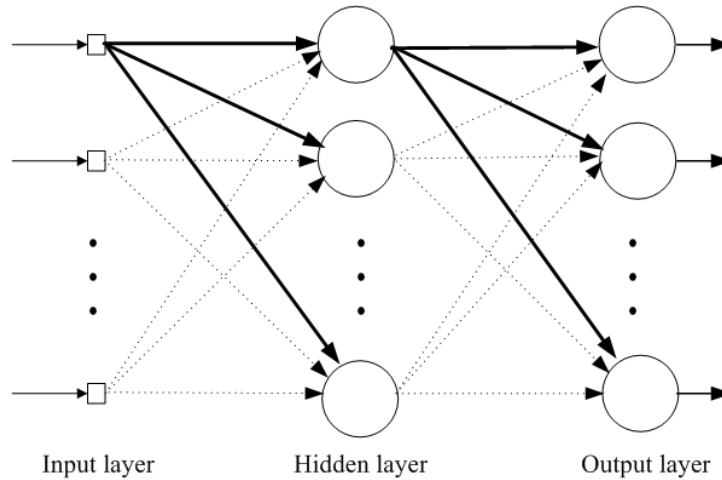
2.2. EIS System Architecture

The architecture of our EIS system is illustrated in figure 2. 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 composed of 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.

Figure 2: EIS system architecture

3. Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational models which their performance is derived from human brain. An ANN is an information-processing system that has certain performance characteristics in common with biological neural networks. ANNs have been developed as generalizations of mathematical models of human cognition or neural biology [1]. An ANN consists of some interconnected processing elements, called nodes, which are connected to each other via a combination of adaptable interconnections, called weights. Figure 3 shows the structure of an MLP ANN. Each node in an artificial neural network represents one biological neuron in biological neural network. Data is propagated through the network, layer by layer. Besides input and output layers, there can be other middle layers, called hidden layers. Each node receives data from preceding nodes, adds it and passes data through a nonlinear function, and then propagates data to following nodes. There are two phases in ANN performance: training phase and test phase. Training phase starts with assigning initial values of parameters to the ANN and presenting input patterns. The process proceeds with adaptation of weights until the ANN learns these patterns. In a test set, the patterns which are not used in training phase are presented to ANN and ANN's outputs are used to evaluate ANN's performance.

Figure 3: Multilayer perceptron with one hidden layer

Neural Networks have been used in a wide range of applications, including: pattern classification, pattern completion, function approximation, optimization, prediction and automatic control. Among all the various NN structures, the MLP has the highest practical value. It is a feed-forward layered network with one input layer, one output layer, and some hidden layers [11]. An MLP network is shown in figure 3. 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 [12] and Parker [13]. Many modifications have been put forward by others.

In this paper two variants of the backpropagation algorithm, BPLRF and BPDLR, are used as main algorithms. Moreover, the other variations of backpropagation algorithm are used for training MLP network and compared with BPLRF and BPDLR algorithms to show that the former algorithms have better results. These variations are Levenberg-Marquardt backpropagation, gradient descent with momentum backpropagation, gradient descent with adaptive learning rate backpropagation and gradient descent with momentum and adaptive learning rate backpropagation.

The backpropagation with plummeting learning rate factor (BPLRF) has a minor difference with basic backpropagation algorithm. In this algorithm we end the training with diminished training parameters, as explained in the following: About ten epochs before ending the training, the learning rate and momentum factors are decreased suddenly to very small values e.g. 10 times smaller. In this way, a sudden decrease in the values of cost function and recognition errors occurs [14].

Simulation results obtained by basic backpropagation show that in the beginning of training, the convergence could be faster with a relatively larger learning-rate factor, but in the middle of the process, a larger learning-rate factor stimulates instability and prevents convergence. The backpropagation with declining learning-rate factor (BPDLR) is based on exploiting this point. It starts training with a relatively large but constant step size, but before destabilizing the network it decreases step size monotonically [15].

The data set is obtained from measuring urinary bladder of 63 suspected cases. These measurements are executed in 7 frequencies. 42 suspected cases have malignant areas and the others have non-malignant areas. So we can divide the data set into two groups, first one represents a 42×7 matrix corresponding to the impedances of non-malignant 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.

In our work, an MLP network with three layers is used. Number of neurons in the first layer corresponding to the number of inputs is 7. The output layer has one neuron. Its output value shows the degree of malignancy as a number between 0 and 1. If output of network is exactly 1, the suspected case has certainly malignant area and if output number is exactly 0, the case has certainly non-

malignant area. Since the output value is a number between 0 and 1, the more this number is close to 1, the more the case has malignant area.

To apply the data set to the ANN, firstly we arrange it as a matrix called data matrix. The data matrix has 63 rows, each row shows the number of suspected cases and 7 columns, each column shows corresponding frequency in which we have measured suspected cases' impedances. Component of the data matrix demonstrates measured impedances. Arrangement of data matrix rows is as below: Two first rows of this matrix represents impedances corresponding to the first and second suspected cases in the first group (malignant group), the third row of data matrix represents impedances corresponding to the first suspected case in the second group (non-malignant group), the fourth and fifth rows of data matrix represents impedances corresponding to the third and fourth suspected cases in the first group, the sixth row of data matrix represents 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 40 rows of this matrix as a train set for training the ANN and 23 row of this matrix as a test set to test the trained ANN.

4. Evaluation

We simulated various training algorithms for this data set to get the best result. Firstly, Levenberg-Marquardt backpropagation, gradient descent with momentum backpropagation, gradient descent with adaptive learning rate backpropagation and gradient descent with momentum and adaptive learning rate backpropagation are used. MATLAB 7.2 artificial neural network toolbox is used to implement these algorithms. Simulation is run for 300 and 5000 epochs in training phase. Neuron activation function is assumed to be sigmoid and the number of neurons in hidden layer is 5. Table 1 shows the simulation results. Three important parameters of the ANN in test phase including maximum of error and error percentage and running time for each algorithm are given.

Table 1: Test Phase Results for Various Backpropagation Algorithms

	NE ^a =300			NE =5000		
	ME ^b	EP ^c	RT ^d	ME ^b	EP ^c	Running time
GDM ^e	0.5360	2.5038	1.7079	0.4038	1.1504	14.7142
GDA ^f	0.5668	2.4157	1.7242	0.2557	0.4690	15.5854
GDMA ^g	0.3944	1.0550	1.7230	0.2375	0.4454	14.7578
LM ^h	1.7017	6.0328	3.1998	0.6478	1.5009	37.3907

^a NE: number of epochs in train phase

^b ME: max-error; maximum error in test phase

^c EP: error-percent; error percent in test phase

^d RT: running time of algorithm in train phase

^e GDM: gradient descent with momentum backpropagation

^f GDA: gradient descent with adaptive learning rate backpropagation

^g GDMA: gradient descent with momentum and adaptive learning rate backpropagation

^h LM: Levenberg-Marquardt backpropagation

Table 2 and Table 3 report the simulation results for simple backpropagation algorithm with and without using BPLRF and BPDLR algorithms for 5000 and 300 epochs, respectively, in training phase for different η s. η in Table 2 and 3 for BPLRF and BPDLR algorithms represent the initial value of learning rate. Algorithm running time for 5000 epochs in both cases is approximately 10.2 seconds. In 300 epochs running, this time is approximately 0.8 second which is a reasonable running time. The learning rate bigger than 0.6 second and smaller than 0.1 second does not show good results. From Table 2, it is clear that using BPLRF and BPDLR algorithms decrease maximum error and error percent for all η s in test phase. Table 3 shows that using BPLRF and BPDLR algorithms decrease maximum error and error percent for $\eta > 0.3$ in test phase.

Table 2: Test Phase Results of MLP with and without BPLRF and BPDLR algorithms in 5000 epochs

	BPLRF & BPDLR		without BPLRF & BPDLR	
	ME ^a	EP ^b	ME ^a	EP ^b
$\eta = 0.6$	0.00598	0.00010	0.00878	0.00028
$\eta = 0.45$	0.00869	0.00020	0.01224	0.00049
$\eta = 0.35$	0.01239	0.00040	0.01688	0.00091
$\eta = 0.2$	0.02783	0.00200	0.03420	0.00361
$\eta = 0.1$	0.07221	0.01462	0.06927	0.01551

^a ME: max-error; maximum error in test phase

^b EP: error-percent; error percent in test phase

Table 3: Test Phase Results of MLP with and without BPLRF and BPDLR algorithms in 300 epochs

	BPLRF & BPDLR		without BPLRF & BPDLR	
	ME ^a	EP ^b	ME ^a	EP ^b
$\eta = 0.6$	0.33297	0.39035	0.65688	0.90674
$\eta = 0.45$	0.34339	0.43183	0.61225	0.79826
$\eta = 0.35$	0.35080	0.47012	0.53954	0.64975
$\eta = 0.2$	0.35097	0.54733	0.41773	0.52868
$\eta = 0.1$	0.36448	0.74503	0.38907	0.68120

^a ME: max-error; maximum error in test phase

^b EP: error-percent; error percent in test phase

Figure 4 shows the effect of backpropagation with plummeting learning rate factor (BPLRF) algorithm. A sudden decrease in learning rate in 297th epoch causes a sudden decrease in total sum square error (TSSE) in the last epochs of training phase.

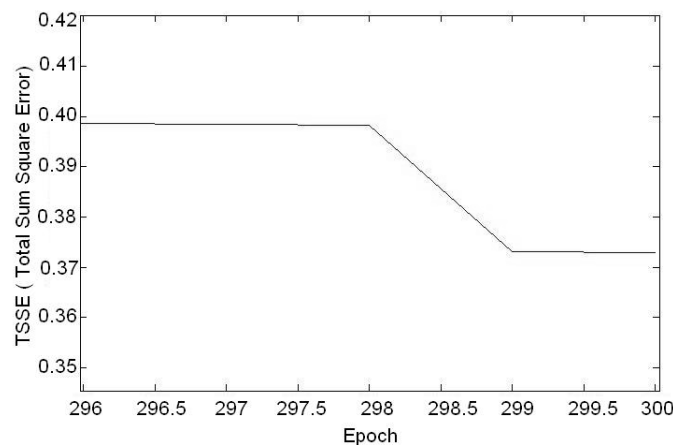
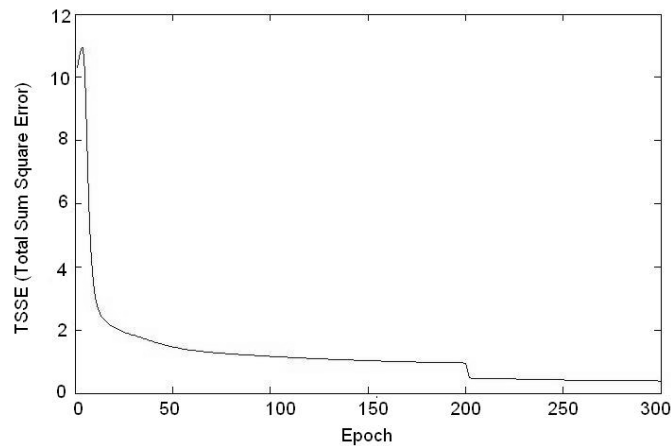
Figure 4: Sudden TSSE decrement in 297th epoch shows the effect of using the BPLRF algorithm

Figure 5 illustrates the effect of backpropagation with declining learning-rate factor (BPDLR). Starting to decrease the value of TSSE in the 200th epoch shows the effect of starting to decrease learning rate monotonically.

Figure 5: Sudden TSSE decrement in 200th epoch shows the starting of the BPDLR algorithm and monotonically declining of TSSE after 200th epoch shows its effect.



5. Implementation

The ANN architecture has been implemented on a Virtex-4 LX25 FPGA from Xilinx. The FPGA features 24,192 logic cells, 48 18×18-bit signed multipliers and 72 Block Select RAMs. The Xilinx ISE 8.1i has been used to implement and synthesis the system. Table 4 shows the utilization of hardware resources on the chip. The multipliers are implemented using the DSP48 blocks of the FPGA. The number of occupied slice is 10136 therefore the design occupies 94% of the FPGA.

Table 4: Utilization of hardware resources

FPGA Utilization Summary			
Logic Utilization	Used	Available	Utilization
Number of Slice flip flops	2021	21,504	9%
Number of 4 input LUTs	17830	21,504	82%
Number of occupied Slices	10136	10,752	94%
Number of bonded IOBs	67	240	27%

In Figure 6, the FPGA evaluation board is shown which is connected to a PC via an electronic interface so that the patient data can be downloaded to the FPGA from the PC. The ANN evaluation of the patient data for a specific bladder area is computed by the FPGA and the result is shown as a number between zero and one on the board LCD.

Figure 6: The FPGA evaluation board under test

6. Conclusion

To implement an artificial neural network, a reconfigurable computing platform such as FPGA is an appropriate alternative which keep both benefits of hardware speed and software flexibilities. In this work, we implemented an ANN on a FPGA to analyze electrical impedance spectroscopy (EIS) data taken from a patient bladder.

A Virtex-4 LX25 FPGA from Xilinx was employed in our system. The design occupies the 94% of the available slices of the FPGA. Using this system, malignant areas from non-malignant areas in the urinary bladder of the patient can be separated very quickly.

References

- [1] L. Fausett, *Fundamentals of Neural Networks*, Prentice-hall, 1994.
- [2] A. R. Omondi, J. C. Rajapakse, *FPGA Implementations of Neural Networks*, Springer, 2006.
- [3] B. L. Hutchings and B. E. Nelson, "Gigaop DSP on FPGA," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001.
- [4] A. Alsolaim, J. Becker, M. Glesner, and J. Starzyk, "Architecture and Application of a Dynamically Reconfigurable Hardware Array for Future Mobile Communication Systems," *International Symposium on Field Programmable Custom, Computing Machines (FCCM)*, 2000.
- [5] B. H. Brown, J. A. Tidy, et al. "Relation between tissue structure and imposed electrical current flow in cervical neoplasia." *Lancet* 355(9207): 892-5. 2000
- [6] C. A. Gonzales-Correa, B. H. Brown, et al. "Virtual Biopsies in Barrett's Oesophagus using an Impedance probe." *Annals New York Academy of Sciences* 873: 313-321.1999
- [7] C. A. Gonzales-Correa, B. H. Brown, et al. "Low frequency electrical bioimpedance for the detection of inflammation and dysplasia in Barrett's oesophagus." *Physiol.Meas.* 24: 291-296. 2003
- [8] A. Keshtkar, A Keshtkar, R. H. Smallwood, 'Electrical impedance spectroscopy and the diagnosis of bladder pathology', *Physiol. Meas.*, 27, 585-596. 2006
- [9] A. Keshtkar, 'Virtual bladder biopsy using bio-impedance spectroscopy at 62.500 Hz–1.024 MHz'. *Measurement*, 40(6)585-590. 2007
- [10] L. Lu, L. Hamzaoui, et al., "Parametric modeling for electrical impedance spectroscopy system." *Med Biol Eng Comput* 34(2): 122-6. 1996
- [11] S. Haykin, *Neural Networks*, Second Edition Printice-Hall of India, 2006.
- [12] P. J. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in Behavioral Science", Ph.D Thesis, Harvard University, Cambridge, MA., 1974.
- [13] D. B. Parker, "Learning Logic," Invention Report s81-64, File 1, Office of Technology Licensing, Stanford University, March 1982.
- [14] M. T. Vakil_baghmisheh, "Farsi Character Recognition Using Artificial Neural Networks", Ph.D Thesis, Faculty of Electrical Engineering, University of Ljubljana, 2002.
- [15] M. T. Vakil-Baghmisheh, N.Pavešić, "Backpropagation with Declining Learning Rate," *Proc. Tenth Electrothechnical and Comp. Scie. Con.*, vol. B, Portorož, Slovenia, 2003, pp. 297-300, 2001.