Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 3159-0040
Vol. 3 Issue 2, February - 2016

A New Expert Diagnosis System For Cardiac Arrhythmia Based On Gda-Anfis

[1]Münire Migal Avcı
Elazig Education and Research Hospital,
Department of Tuberculosis and Pulmonology
Elazig, Turkey
e-mail: leavci@turk.com

[2]Derya Avcı
[1]Firat University, Engineering Faculty, Department of Electrical and Electronic Engineering, 23119 Elazig, Turkey
e-mail: derya2344@hotmail.com

Abstract—Cardiac arrhythmia is the condition in which the heart's normal rhythm is disrupted. In this paper, an expert diagnosis system based on General Discriminant Analysis (GDA) and Adaptive Network Based on Fuzzy Inference System (ANFIS) for the cardiac arrhythmia is introduced. This expert diagnosis system deals with combination of the feature extraction and classification. This expert cardiac arrhythmia diagnosis system has two stages, which feature extraction – reduction and classification stages. In feature extraction – reduction stage, the cardiac arrhythmia features were obtained from UCI Repository of Machine Learning Databases. Then, number of these features was reduced from 279 to 10 by using General Discriminant Analysis (GDA). In classification stage, these reduced features are given to inputs ANFIS classifier. The correct diagnosis performance of the GDA-ANFIS automatic diagnosis system for cardiac arrhythmia diseases is estimated by using classification accuracy. The classification accuracy of this GDA-ANFIS expert diagnosis system for diagnosis of cardiac arrhythmia diseases was obtained about 84.10 %.

Keywords—Expert Systems; General Discriminant Analysis (GDA), Adaptive Network Based on Fuzzy Inference System (ANFIS); cardiac arrhythmia database; classification accuracy; sensitivity and specifity analysis.

1. Introduction

Cardiac arrhythmia is the status in which the heart's normal rhythm is disrupted. In human body, the heart pumps blood containing oxygen, nutrients, immune cells, and regulatory molecules to the body organs [1]. The rhythm of the heart is set by a small region of cardiac muscle cells in the right atrium named the sinoatrial (SA) node that acts as a spontaneous pacemaker, but is under the control of nerves and circulating hormones that affect the heart rate via a host of control circuits that maintain adequate blood pressure and oxygenation.

The normal heart rhythm is named sinus rhythm. Each beat spontaneously generated from the SA node in sinus rhythm produces a propagating wave of bioelectricity that spreads throughout the four chambers of the heart in a coordinated fashion. All impulse propagates throughout the atria before being channeled through the atrioventricular (AV) node to the ventricles. This electrical wave (impulse) triggers intracellular calcium processes that produce the contractions of the cardiac muscle that pump the blood to the organs of the human body. The slow (about 120-200 ms) conduction time through the AV node permits adequate time for atrial contraction and ventricular filling. The electrical impulse propagates through specialized conducting bundles called the His-Purkinje system and from there to the ventricles upon emerging from the AV node. The His-Purkinje system permits rapid conduction to all areas of the ventricles. Therefore is responsible for ensuring effective ventricular contraction. Although there is substantial individual variation, normally, the heart beats at a rate of approximately 75 beats per minute and pumps about 5 liters of blood per minute.

The heart rhythm is generally monitored by an electrocardiogram (ECG), which measures the voltage differences between points on the surface of the human body. The normal cardiac rhythm on an ECG, where the P wave is associated with the excitation of atria, the QRS complex is associated with the excitation of ventricles, and the T wave is associated with the relaxation of the ventricles. The duration of the excitation phase of the ventricles is associated with the QT interval and the duration of the time for an excitation to travel from the atria through the AV node to the ventricles is associated with the PR interval.

Abnormal cardiac rhythms are named cardiac arrhythmias. The abnormal propagation of a wave of cardiac excitation, or some combination of the two and Cardiac arrhythmias are associated with abnormal initiation of a wave of cardiac excitation. Cardiac arrhythmias can manifest themselves in many different ways. It is still not always possible to determine the mechanism of an arrhythmia.

Arrhythmias can be classified in several ways and one useful classification that will be utilized here is reentrant versus non-reentrant arrhythmias. The cardiac tissue is repetitively excited by a propagating wave circulating around an obstacle (anatomical reentry) or circulating freely in the tissue as a spiral or scroll wave (functional reentry) in reentry. So, there is a strong spatial component to reentrant arrhythmias. Either a sufficiently large spatial extent is needed to support the initiation and continuation of the arrhythmia, or an appropriate geometry must be present to allow a reentrant circuit. Non-reentrant

arrhythmias have a strong geometric componenent in which either propagation is inhibited at particular anatomical sites or one or more pacemakers form at abnormal namely ectopic locations [1].

The heart rate also can classify to arrhythmias. Tachyarrhythmias are rhythms that the heart rate is faster than normal, usually taken as greater than 100 beats per minute. They are further classified based on where they arise as either ventricular or supraventricular (anywhere above the ventricles, including the SA node, the atria, and the AV node).

Tachyarrhythmias may cause to reduced blood flow to the organs of the body leading to reduced ability to exercise, faintness, or, in cases in which the blood flow is too low, death. Tachyarrhythmias can emerge from an accelerated sinus rhythm, an accelerated rhythm from an abnormal ectopic site, or from the interactions of multiple ectopic sites. Moreover, more usually tachyarrhythmias are believed to emerge from reentrant arrhythmias, in which the period of the oscillation is set by the time an excitation takes to travel in a circuitous path, rather than the period of oscillation of a pacemaker.

Bradycardias are rhythms that heart rate is decreased below 60 beats per minute. Bradycardias may emerge because the sinus rhythm is abnormally slow, the sinus rhythm is absent and a secondary slower pacemaker (e.g. in the AV node) takes over, or the normal sinus rhythm is inhibited in its passage through the heart. Not all bradycardic rhythms are associated with impaired cardiac function, since well-conditioned athletes may have normal heart rates below 60 beats per minute.

A physician commonly determines decisions by evaluating the current test results of a patient or the physician compares patient with other patience with the same condition by referring to the previous decisions. Therefore, diagnose of the cardiac arrhythmia for a physician is very difficult matter [2]. For this aim, an expert diagnosis system based on General Discriminant Analysis (GDA) and Adaptive Network Based on Fuzzy Inference System (ANFIS) for the cardiac arrhythmia is used in this study for help to physician on diagnose of the cardiac arrhythmia diseases.

The paper is organized as follows. In Section 2, General Discriminant Analysis (GDA) is introduced. In Section 3, proposed GDA-ANFIS expert system for diagnosis of the cardiac arrhythmia diseases is explained. In Section 4, the obtained results using GDA-ANFIS is given. Finally, in Section 5, the discussion and conclusion are presented.

2. Generalized Discriminant Analysis

The Generalized Discriminant Analysis (GDA) is similar to Linear Discriminant Analysis's (LDA) [3], [4]. The aim of GDA is to maximize the quotient between the inter-classes inertia and the intra-classes inertia in a mapped feature space. If there is a K-class problem and letting M_s be the number of samples in class k, a set of training patterns from the *C* classes can be given as below:

$$M=\sum_{k=1}^{K} M_k \qquad x_{ki}, \quad k=1,2,...,K \quad i=1,2,...,M_k \quad (1)$$

Φ: $R^L \rightarrow T$, which is a nonlinear mapping the set of training samples in the mapped feature space, can be represented as $\{\Phi(x_{ki}), s = 1, 2, ., K; i = 1, 2,., M_k\}$. The K_b and K_w of the training set can be calculated as below:

$$K_w=(1/K)\sum_{k=1}^{K}\frac{1}{M_k}\sum_{i=1}^{M_k}\phi(x_{ki})\phi(x_{ki})^T \quad (2)$$

$$K_b=(1/K)\sum_{k=1}^{K}(\mu_k-\mu)(\mu_k-\mu)^T \quad (3)$$

GDA calculates the eigenvalues $\lambda \geqslant 0$ and eigenvectors $m \in T \setminus \{0\}$ satisfying

$\lambda K_w l = K_b l$ (4)

here all solutions *m* lie in the span of $\Phi(x_{11}), ., \Phi(x_{si}), .$ and there exist coefficients p_{ki} such that

$$l=\sum_{k=1}^{K}\sum_{i=1}^{M_k} p_{ki}\phi(x_{ki}) \quad (5)$$

the dot product of a sample *i* from class *t* and the other sample *j* from class *g* in the feature space is performed by using kernel techniques, which are shown as $(A_{zd})_{tg}$, such as radial basis kernel, polynomial kernel etc. as below:

$$(A_{zd})_{tg}=\Phi(x_{tg})\cdot\Phi(x_{gy})=k(x_{tz},x_{gd})=e^{-|x_{tz}-x_{gd}|^2/p} \quad (6)$$

If B is a $N \times N$ matrix defined on the class elements by $(B_{tg})_{(t=1,....k),(g=1,.....k)}$, here B_{tg} is a matrix composed of dot products between vectors from class *p* and *q* in feature space:

$$B_{tg}=(A_{zd})_{(z=1,....,M_t),(d=1,....,M_g)} \quad (7)$$

a $N \times N$ block diagonal matrix can be defined as below:

$H=(H_k)_{k=1,.,}k$ (8)

there V_k is $M_k \times M_k$ a matrix with terms all equal to $1/M_k$.

The solution of Eq(4) can be accomplished by substituting Eqs.(2), (3) and (5) into (4) and taking inner-product with vector $\Phi(x_{zd})$ on both sides:

$$\lambda DD_t = DHD_t \quad (9)$$

there t represents a column vector with entries t_{ki}, $k = 1, . , K$, $i = 1, ., M_k$. The eigenvectors of the matrix $(DD)^{-1}$ DHD are calculated. If these eigenvectors of $(DD)^{-1}$ DHD are found, the solution of a in Eq. (9) is completed. Moreover, the matrix B might not be reversible. The eigenvector is found **a** by first diagonalising matrix D [4]. Firstly W important

eigenvectors are calculated, a projection matrix can be structured as below:

$$F = [a_1\ a_2\ a_3 ... a_J] \qquad (10)$$

The projection of x in the W-dimensional GDA space is given by:

$\mathbf{y}=O_x F$ (11)

here

$O_x=[O(x,x_{11}).u(x,x_{ki}).u(x,x_{KMk})]$(12)

3. Developed Expert Diagnosis System based on GDA - ANFIS for The Cardiac Arrhythmia

Developed expert diagnosis system based on General Discriminant Analysis (GDA) and Adaptive Network Based on Fuzzy Inference System (ANFIS) for the cardiac arrhythmia diseases used in this study consists of two stages: The General Discriminant Analysis (GDA) phase and classificiation by using ANFIS classifier phase.

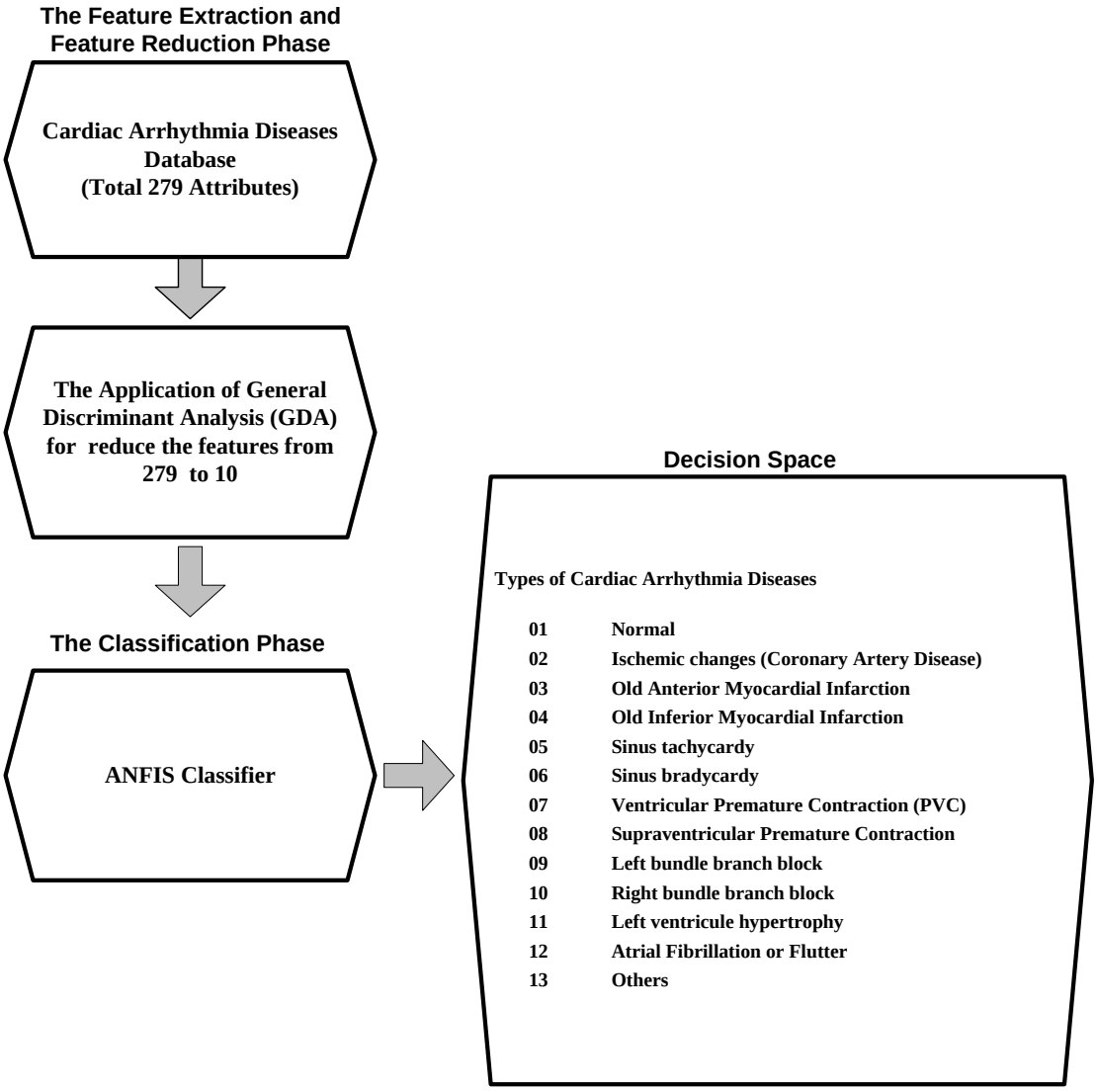

Fig.1. The block diagram of expert diagnosis system based on GDA - ANFIS for the cardiac arrhythmia diseases used in this study.

1. The Feature Extraction and Feature Reduction Phase: In this phase, the cardiac arrhythmia diseases database obtained from the UCI Repository of Machine Learning Databases was used [5]. Then, General Discriminant Analysis (GDA) mentioned in Section 2 was conducted to this database for feature reduction for reduce the features (attributes) from 279 to 10. Features of this database were given in below:

- Number of Instances in Database: 452
- Number of Attributes in Database: 279

The most important attributes of the cardiac arrhythmia diseases database used in this study can be given as below in Table 1:

Table 1. Some attributes of the cardiac arrhythmia diseases database.

Age: Age in years , linear
Sex: Sex (0 = male; 1 = female)
Height: Height in centimeters
Weight: Weight in kilograms
QRS duration: Average of QRS duration in msec.
P-R interval: Average duration between onset of P and Q waves in msec.
Q-T interval: Average duration between onset of Q and offset of T waves in msec.
T interval: Average duration of T wave in msec.
P interval: Average duration of P wave in msec.
Vector angles in degrees on front plane of.
Heart rate: Number of heart beats per minute
Average widths of channels in msec.
Number of intrinsic deflections.
Existence of ragged R wave,
Existence of diphasic derivation of R wave,
Existence of ragged P wave.
Existence of diphasic derivation of P wave.
Existence of ragged T wave.
Existence of diphasic derivation of T wave.
QRSA , Sum of areas of all segments divided by 10, (Area= width * height / 2).
QRSTA = QRSA + 0.5 * width of T wave * 0.1 * height of T wave. (If T is diphasic then the bigger segment is considered).
Amplitude * 0.1 milivolt of channel

In this cardiac arrhythmia diseases database, there are 13 classes: Normal has number of instances is 245, Ischemic changes (Coronary Artery Disease) has number of instances is 44, Old Anterior Myocardial Infarction has number of instances is 15, Old Inferior Myocardial Infarction has number of instances is 15, Sinus tachycardy has number of instances is 13, Sinus bradycardy has number of instances is 25, Ventricular Premature Contraction (PVC) has number of instances is 3, Supraventricular Premature Contraction has number of instances is 2, Left bundle branch block has number of instances is 9, Right bundle branch block has number of instances is 50, Left ventricule hypertrophy has number of instances is 4, Atrial Fibrillation or Flutter has number of instances is 5 and others has number of instances is 22.

2. The Classification Phase: Thus, the size of obtained feature vector at final of feature extraction and feature reduction phase is 452 x 10. The half of this 452 x 10 size feature vector is used for classification phase. Namely, 226 x 10 feature vector is given to the inputs of ANFIS classifier.

The rest of this 452 x 3 size feature vector is used in phase of testing of correct diagnosis performance of GDA-ANFIS for diagnosis of the cardiac arrhythmia diseases method used in this study.

In structure of the ANFIS classifier are used both artificial neural network and fuzzy logic [6-8]. ANFIS classifier is formed if-then rules, couples of input-output and learning algorithms of neural network. These are used for training of ANFIS classifier [7-11].

In this experimental study, the ANFIS classifier has 10 inputs ($x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}$) and 1 output (y). A typical rule set with base fuzzy if-then rules for a first order Sugeno fuzzy model can be expressed as below:

If x_1 A_1 and x_2 B_1 and x_3 C_1 and x_4 D_1 and x_5 E_1 and x_6 F_1 and x_7 G_1 and x_8 H_1 and x_9 K_1 and x_{10} L_1 then

$$f_1 = px_1 + ppx_2 + qx_3 + qqx_4 + sx_5 + ssx_6 + rx_7 + rrx_8 + tx_9 + ttx_{10} \quad (13)$$

Here, p, pp, q, qq, s, ss, r, rr, t, tt and u are linear output parameters. The structure of this ANFIS classifier is formed by using five layer and 1024 if-then rules. The readers can acquired more knowledge about ANFIS from [11].

The parameters of the ANFIS classifier used in this study can be given as below in Table 2:

Table 2. The parameters of the ANFIS classifier used in this study.

The number of layers : 5

Input: 10

Rules number: 1024

Output: 1

Type of Input Membership Functions: Bell-shaped

Training Parameters and Learning Rule: Hybrid Learning Algorithm (Back-propagation for nonlinear

parameters (a_i, c_i) and Least squre errors for linear parameters (p_i,

pp_i, q_i, qq_i, s_i, ss_i, r_i, rr_i, t_i, tt_i, u_i))

Sum-squared Error: 0.00000001

Reaching Epochs Number to Sum-squared Error: 2018

4. The Obtained Results Using GDA-ANFIS Expert Diagnosis Method

The correct diagnosis performance of the GDA-ANFIS expert system for diagnosis of the cardiac arrhythmia diseases is estimated by using classification accuracy performance evaluation method.

In these experimental studies, the half of this 452 x 10 size feature vector and the rest of this 452 x 3 size feature vector are used for training and testing of GDA-ANFIS expert diagnosis system for the cardiac arrhythmia diseases respectively.

***The Results of Diagnosis Accuracy Analysis for GDA-ANFIS*:** The correct diagnosis rates of the cardiac arrhythmia diseases were calculated for obtaining of results of classification accuracy analysis for GDA-ANFIS expert diagnosis system for the cardiac arrhythmia diseases in this study by using Eq.(14) [12-20].

$$diagnosis\_accuracy(C) = \frac{\sum_{k=1}^{|C|} assess(c_k)}{|C|}, \qquad c_k \in C \tag{14}$$

$$assess(c) = \begin{cases} 1, & if \quad classify(c) = c.d \\ 0, & otherwise \end{cases}$$

there, C is the set of database to be classified, which is the test set. $c_k \in C$, $c_k.d$ is the class of item c. The classify(c_k) returns the classification of c_k by ANFIS classifier.

The obtained classification accuracy by using GDA-ANFIS expert diagnosis system for cardiac arrhythmia diseases are given in Table 3.

Table 3. The obtained classification accuracy by using GDA-ANFIS expert diagnosis system for cardiac arrhythmia diseases.

Index of type of the cardiac arrhythmia diseases	Number of correct diagnosis	Number of Incorrect diagnosis	Percentage of correct diagnosis rate (%)
01	117	6	94.87
02	20	2	90.90
03	5	3	62.50
04	5	2	71.42
05	4	3	57.14
06	9	3	75.00
07	1	1	50
08	0	1	0
09	2	2	100
10	20	5	80
11	2	0	100
12	2	0	100
13	9	2	81.81
Total	**195**	**31**	**84.10**

Obtained the confusion matrix of the proposed GDA-ANFIS expert diagnosis system for cardiac arrhythmia diseases is given in Table 4.

Tablo 4. Obtained the confusion matrix of the proposed GDA-ANFIS expert diagnosis system for cardiac arrhythmia diseases.

	01	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**
01	117	-	1	1	-	-	1	-	-	3	-	-	-
02	-	20	-	-	1	-	-	-	1	-	-	-	-
03	-	2	5	1	-	-	-	-	-	-	-	-	-
04	-	-	-	5	-	-	-	1	-	-	-	1	-
05	-	-	-	-	4	-	2	-	-	-	1	-	-
06	-	-	1	-	-	9	-	-	1	-	-	-	1
07	-	-	-	-	-	-	1	-	1	-	-	-	-
08	-	-	-	-	-	-	-	-	-	-	-	-	1
09	-	-	-	1	-	-	-	-	2	-	1	-	-
10	-	1	-	-	-	-	-	-	3	20	-	1	-
11	-	-	-	-	-	-	-	-	-	-	2	-	-
12	-	-	-	-	-	-	-	-	-	-	-	2	-
13	-	-	-	-	-	1	-	-	1	-	-	-	9

It was performed experimental studies on the cardiac arrhythmia diseases database mentioned in Section 3 to evaluate the robustness of GDA-ANFIS expert diagnosis system for the cardiac arrhythmia diseases.

5. Discussion and Conclusion

In this study, GDA-ANFIS expert diagnosis system for the cardiac arrhythmia diseases was used. Then, performance evaluation technique was applied to obtained classification results. This technique is classification accuracy [12-20]. As shown from these results, the GDA-ANFIS expert diagnosis system for the cardiac arrhythmia diseases obtains very promising results in classifying the possible the cardiac arrhythmia patients. This statement of the GDA-ANFIS expert diagnosis system for the cardiac arrhythmia diseases is clearly seen from the above results in Tables 3-4.

Therefore, the used GDA-ANFIS expert diagnosis system for the cardiac arrhythmia diseases can be very helpful to the physicians for their final decision on their patients. The physicians can perform very accurate decisions by using such an efficient tool.

The results show that GDA-ANFIS based a learning method can assist in the diagnosis of the cardiac arrhythmia diseases. In future studies of the cardiac arrhythmia diseases diagnostic, different

feature extraction and classifier methods will be used for increasing of correct accurate.

References

1. www.scholarpedia.org/article/Cardiac_arrhythmia, access date: 17 August 2010.

2. Polat, K., and Gunes, S. (2006). Hepatitis disease diagnosis using a new hybrid system based on feature selection (FS) and artificial immune recognition system with fuzzy resource allocation, *Digital Signal Processing*, *Volume 16, Issue 6*, *November 2006*, *Pages 889-901.*

3. http://en.wikipedia.org/wiki/pattern_recognition , access date: 25 November 2007.

4. G Baudat, F.E. Anouar, (2000) Generalized discriminant analysis using a kernel approach, Neural Computation 12 (10), 2385– 2404.

5. ftp://ftp.ics.uci.edu/pub/machine-learning arrhythmia database (last accessed: 17 August 2010).

6. Blake, C. L., & Merz, C. J. (1996). UJI Reporsitory of Machine Learning Databases.

Available from http://www.ics.uci.edu./~mlearn/MLReporsitory.html.

7. F. Tang and H. Tao, (2006). Fast Linear Discriminant Analysis Using Binary Bases

The 18th International Conference on Pattern Recognition (ICPR'06).

8. P.N., Belhumeur, J. Hespanda, and D. Kiregeman, (1997). Eigenfaces

vs.Fisherfaces: Recognition Using Class Specific Linear Projection, *IEEE PAMI*.

9. Avci, E. and Turkoglu, I. (2003). Modelling of Tunnel Diode by Adaptive-Network-

Based Fuzzy Inference System, International Journal of Computational Intelligence,

Volume:1, Number:1 , pp. 231-233.

8. Kosko, B., (1991), Neural Networks and Fuzzy Systems, A Dynamical Systems Approach, Englewood Ciffs., NJ: Prentice Hall.

9. Jang, J. S. R., Sun, C. T., (1995), Neuro-Fuzzy Modeling and Control, proceedings of the IEEE, vol. 3, No. 3.

10. Jang, J. S. R., May. (1993), "ANFIS: Adaptive network based fuzzy inference systems", IEEE Trans. Syst., Man. And Cybern., vol. 23, pp. 665-685.

11. Avci, E., Turkoglu I. and Poyraz M. (2005). Intelligent Target Recognition Based on

Wavelet Adaptive Network Based Fuzzy Inference System, Lecture Notes in

Computer Science, Springer-Verlag , vol. 3522 / 2005 , pp. 594-601.

12. Comak, E., Arslan, A., & Turkoglu, I. (2006). A decision support system based on support vector machines for diagnosis of the heart valve diseases, *Computers in Biology and Medicine, Volume 37, Issue 1, January 2007, Pages 21-27.*

13. Avci, E. ve Akpolat, Z. H. (2006). Speech Recognition Using A Wavelet Packet

Adaptive Network Based Fuzzy Inference System," *Expert Systems with Applications,*

31(3), 495-503.

14. Polat, K., and Gunes, S. (2007). An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease, Digital Signal Processing, Volume 17, Issue 4, Pages 702-710.

15. Polat K. and Gunes, S. (2008), Principles component analysis, fuzzy weighting pre-

Processing and artificial immune recognition system based diagnostic system for

diagnosis of lung cancer, Expert Systems with Applications, Volume 34, Issue 1,

Pages 214-221.

16. Watkins, A., (2001). AIRS: A resource limited artificial immune classifier, Master

thesis, Mississippi State University.

17. Avci, E., (2007), A New Optimum Feature Extraction and Classification Method for

Speaker Recognition: GWPNN, *Expert Systems with Applications,* 32(2), 485-498.

18. Avci, E., (2007). An Expert System Based on Wavelet Neural Network-Adaptive Norm Entropy For Scale Invariant Texture Classification," *Expert Systems with*

Applications, 32(3), 919-926.

19. Avci, E., (2008). Comparison of Wavelet Families for Texture Classification By

Using Wavelet Packet Entropy Adaptive Network Based Fuzzy Inference System",

Applied Soft Computing, 8 (1), 225-231,

20. Avci, E., (2007). An Automatic System for Turkish Word Recognition using Discrete

Wavelet Neural Network Based on Adaptive Entropy, *Arabian Journal for Science*

and Engineering, Cilt:32, No:2B, 239-249.