European Online Journal of Natural and Social Sciences 2013; www.european-science.com
Vol.2, No.3 Special Issue on Accounting and Management.
ISSN 1805-3602

# Presenting a New Strategy to Extract Data Clustering Heartbeat Samples by Using Discrete Wavelet Transform

**Marjan Nobakht[1*], Jalal Javadi Moghaddam[2], Ahmad Bagheri[2]**
[1] Department of computer Engineering, University of Guilan, University campus 2, Guilan, Iran
[2] Department of mechanical Engineering, The University of Guilan, Guilan, Iran
*E-mail: marjannobakht@gmail.com

**Abstract**

This paper presents the improvement of detection system for normal and arrhythmia electrocardiogram classification. This classification is done to aid the ANFIS (Adaptive Neuro Fuzzy Inference System). The data used in this paper obtained from MIT-BIH normal sinus ECG database signal and MIT-BIH arrhythmia database signal. The main goal of our approach is to create an interpretable classifier that provides an acceptable accuracy. In this model, the feature extraction using DWT (Discrete Wavelet Transform) is obtained. The last stage of this extraction is introduced as the input of ANFIS model. In this paper, the ANFIS model has been trained with Quantum Behaved Particle Swarm Optimization (QPSO). In this study, for training of proposed model, four sample data have been used which result in acceleration of training data. On the test set, we achieved an outstanding sensitivity and accuracy 100%. Experimental results show that the proposed approach is very fast and accurate in improving classification. Using the proposed methodology and telemedicine technology can manage patient of heart disease.

**Keywords:** ECG signal, discrete wavelet transform, Adaptive Neuro-Fuzzy Interface System, Quantum behaved Particle Swarm

## Introduction

According to the recent survey from the World Health Organization (WHO), cardiovascular disease causes 17.3 million deaths each year globally, ranking No.1 in the leading causes of mortality (Liang et.al, 2014) Electrocardiography is commonly used by the physicians in cardiology since it consist of effective, functioning, simple, non-invasive, and low cost tool to the diagnosis of cardiovascular disease(Ozbay & Gulay, 2010). An ECG recording is a measure of the activity of the heart from electrodes placed at specific locations on the torso (Bwgacem et.al, 2003). According to ECG’s shape and feature, disease could be diagnosed. Early detection and treatment of the heart diseases can rescue the patient’s life or prevent permanent damages on tissues of the heart. A correct diagnosis of a heart disease might require manual inspection of many hours of ECG heartbeats by expert physicians. This is tedious and time consuming with high possibility of missing critical information (Kutlu &. Kntalp, 2012).

Figure 1 shows ECG signal. The majority of the professionals clinically believed that the useful information in the ECG is originated in the intervals and also the amplitudes that defined by its features (characteristic wave peaks and time durations). The improvement of precise and rapid methods for automatic ECG feature extraction is of chief importance, particularly for the examination of long recording. Therefore, automatic detection and classification systems can enable cardiologists to detect the cardiac diseases, timely. Therefore, it is leading to an extended quality of life for the patients which is very important (Abawajy et.al, 2013).

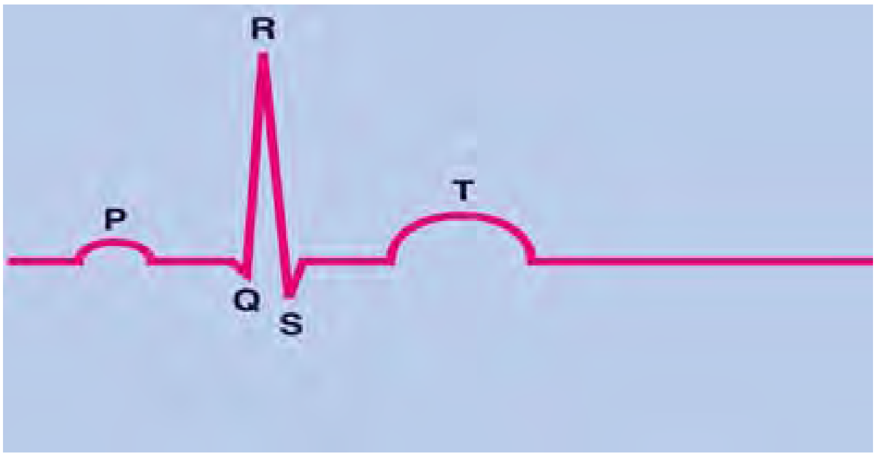


**Figure 1. Diagram of the Human Heart and An example of Normal ECG Trace**

Automatic classification of ECG data sets is critical for clinical diagnosis of cardiovascular diseases such as abnormal cardiac rhythm and arrhythmia.

Abnormality of the ECG shape is usually called arrhythmia. Arrhythmia is a common term for any cardiac rhythm that differs from normal sinus rhythm. Automatic arrhythmia detection and classification of arrhythmia are important in clinical cardiology, especially when performed in real time. Several research and various methods of automatic arrhythmia detection have been developed, such as use of neural network method (Silipo & Marchesi, 1998; Ozbay & Karlik, ,2001; Dokur & Olmez, ,2001). Neuro fuzzy method (Ozbay et.al, 2006; Gacek. & Pedrycz, 2006; Engin , 2004; Acharya et.al, 2003).and feature extraction method (Ceylan, & Ozbay ,2007; Yu & Chou ,2008; Froese et.al 2006). The studies in (Pannizzo & Furman, 1998; Afonso et.al 1999; Christov et.al 2006). Proposed some signal detection methods for discriminating cardiac arrhythmia in the time or frequency domains. The study in (Li et.al, 1995). Proposed a wavelet decomposition algorithm for analysing ECG beat. Using the multi scale feature of wavelet transforms, the QRS complex can be distinguished from high P or T waves, noise, and baseline drift. Rai et al (Rai & Trivedi, 2012).proposed the Wavelet Transform for feature extraction and Back Propagation Neural Network for classifying and analysing heartbeats (Vimala et.al, 2013). Proposed a DWT for feature extraction and neuro-fuzzy methods uses for classification. In this study, feature extraction the ECG signal done with DWT and output of DWT have been utilized as input of ANFIS. The proposed method is implemented by MATLAB software.

**Theory**

In the proposed approach, the ECG classification contains some steps as shown in Figure 2.For this paper, normal sinus ECG database and arrhythmia database is utilized. It can be seen that the whole theory is divided into two basic parts: feature extraction and classification. The first stage of proposed model is feature selection for ECG that it is done using feature extraction technique. A novel method is presented for extracting the features vector for each sample of selected database using an algorithm that exploits the statistics data derived from Discrete Wavelet Transform. In the next step, the procedure of the classification process using an Adaptive Neuro Fuzzy Inference System modelling is presented. Output of ANFIS is decimal data and it should be convert to the integer data with an order of a round function in the sake of classification.

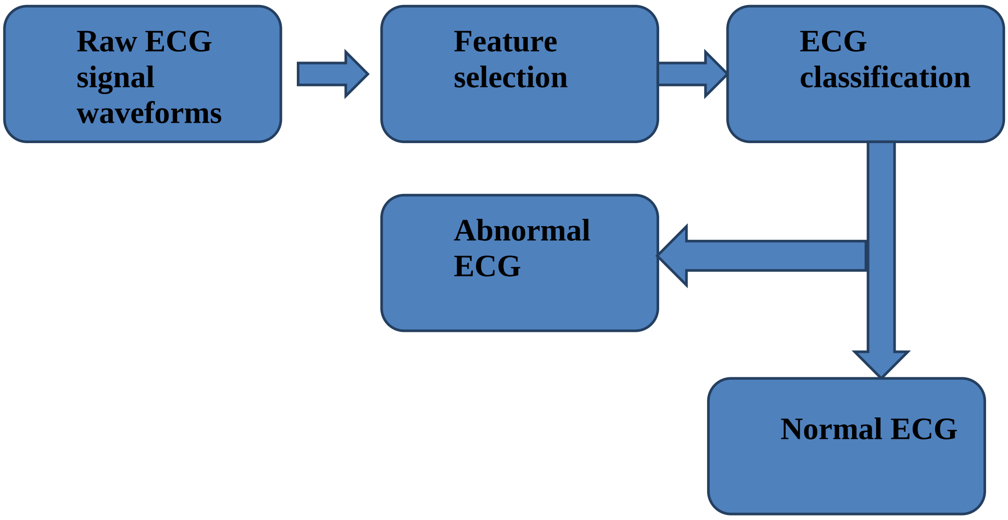


**Figure 2. General overview of ECG signal filtering and classification**

***DWT***

Wavelet analysis consists of a versatile collection of tools for the analysis and manipulation of signal. Wavelet transform equipped with fast numerical algorithms enabling real-time implementation of a variety of signal processing task, such as data compression, extracting of parameters for recognition and diagnostic, transformation and manipulation of data (Coifman et.al, 2010).The Wavelet Transform (WT) is designed to address the problem of non-stationary ECG signals. It derived from a single generating function called the mother wavelet by translation and dilation operations. The main advantage of the WT is that it has a varying window size, being broad at low frequencies and narrow at high frequencies, therefore uses this transform causes to an optimal time-frequency resolution in all frequency ranges. The WT of a signal is the decomposition of the signal over a set of functions obtained after dilation and translation of an analysing wavelet. The ECG signals which consisting of many data points, can be compressed into a few features by performing spectral analysis of the signals with the WT. These features characterize the behaviour of the ECG signals. Using a smaller number of features to represent the ECG signals is particularly important for recognition and diagnostic purposes. The ECG signals were decomposed into time-frequency representations using DWT. The DWT technique has been widely used in signal processing tasks in recent years. The major advantage of the DWT is that it provides a good time resolution. Good resolution at high frequency and good frequency resolution at low frequency. Because of its great time and frequency localization ability, the DWT can reveal the local characteristics of the input signal (Srivastava1 & Prasad, 2013).

***Feature Extraction with DWT***

The DWT wavelet types have been chosen in the features extraction and the ECG signals were decomposed into time-frequency representations using single-level one-dimensional wavelet decomposition. In this study, The Daubechies wavelet filters (db4) has been selected and the number of decomposition levels is chosen to be 5. In the result of decomposition, the ECG signals were decomposed into the details coefficients D1-D5 and one final approximation coefficient, A5.

The results shows that the Daubechies wavelet of order 4 (db4) is more capable to detect the changes of ECG signal. In this work, from the original of ECG signal, four standards of measurement statistical parameters are utilized. These parameters can train the model, perfectly. Moreover, the fast training of the proposed model can be considered. These parameters can be obtained by the following:

- Maximum.
- Minimum.
- Mean.
- Standard deviation.

Finally, 4 wavelets for each of the ECG signals based on the feature have been obtained.

**ANFIS Classifier**

ANFIS is an adaptive network which permits the usage of neural network topology together with fuzzy logic. It is not only includes the characteristics of the two methods, but also eliminates some disadvantages of their lonely-used case. Actually, ANFIS is such as the fuzzy inference system with this difference that in the ANFIS mechanism a feed-forward back propagation method is employed and tries to minimize the resulted error. Consequent parameters are calculated forward while premise parameters are calculated backward. Several fuzzy inference system have been described by different researchers (Wang et.al, 1993; Tsukamoto, 1979, Mamdani et.al 1974). The basic structure of a fuzzy inference system maps input characteristics to input membership functions (mf), input mf to rules, rules to a set of output characteristics, output characteristics to output mf, and the output mf to a single-valued output or a decision associated with the output. The steps for implementing ANFIS are shown in fig.3 (Vijilal et.al 2005). ANFIS can then purify the fuzzy if–then rules and membership functions to describe the input–output behaviour of a complex system. Jang showed that even if human expertise is not available, it is possible to intuitively set up practical membership functions and employs the neural training process to generate a set of fuzzy if–then rules that approximate a desired data set.

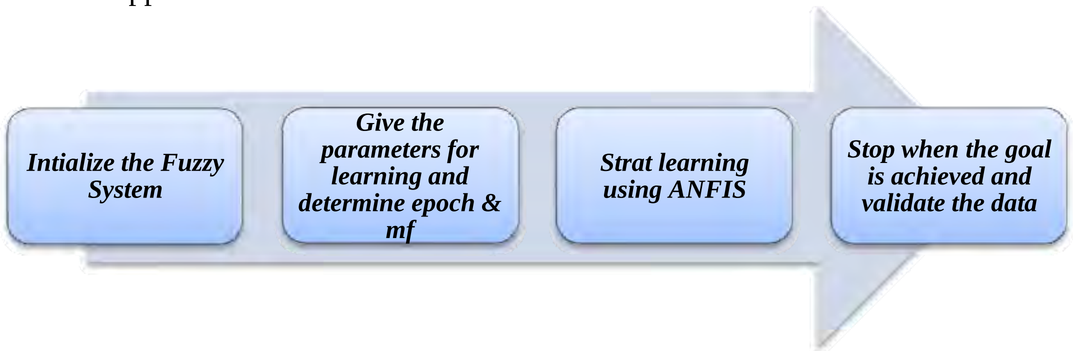


**Figure 3. Program structure of ANFIS**

In the proposed ANFIS method, the consequent parameters are identified by the least squares estimation in the forward pass of the hybrid learning algorithm. In the backward pass, the Quantum behaved particle swarm Optimization algorithm is utilized. In this paper, four inputs and for each input, five membership functions are considered. Fig. 4 shows the structure of this system. Training data is used to prepare the network and to decide the input and output ranges according to the training function. The test data is applied in the next step. It is noted that in this paper, the ANFIS output is considered in the form of number (Sahu et.al, 2013).

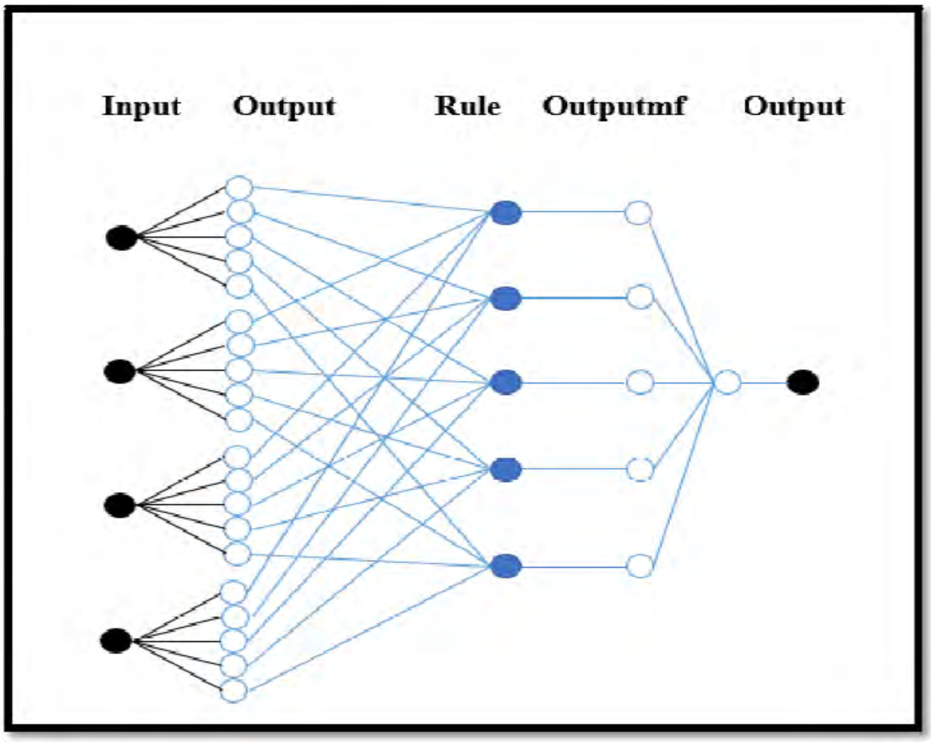


**Figure 4.Model structure**

**PSO**

Particle Swarm Optimization (PSO) (Eberhart & Kennedy, 1995a Eberhart & Kennedy, 1995b). is an evolutionary optimization algorithm proposed by Kennedy and Eberhart in the mid-1990s while attempting to simulate the choreographed, graceful motion of swarms of birds as part of a socio-cognitive study investigating the notion of ‘collective intelligence’ in biological populations.

Every swarm particle of PSO explores a possible solution. It adjusts its flight according to its own and its companion’s flying experience. The personal best position is the best solution that is found by the particle in the course of flight. The best position of the whole flock is the global best solution. The former is called personal best, and the latter global best. Every swarm continuously updates itself through the above mentioned best solutions. Thus a new generation of community comes into being, which has moved closer towards a better solution, ultimately converging onto the optimal solution. In practical operation, the fitness function, which is determined by the optimization problem, assesses the extent to which the particle is good or bad. If the scale of swarm is N, then the position of the ith, (" i = 1, 2, 3. . . N) Particle is expressed as Xi. The ‘‘best” position discovered by the i-th particle is expressed as $pBest_i$. The index of the position of the particle of the swarm, with best solution is expressed as gBest. Therefore, a swarm particle – i will update its own speed and position according to the following Eqs. (1) and (2)

$$V_{(i+1)} = w \times V_i + \left\{ c_p \times r_1 \times (pBest_i - X_i) \right\} + \left\{ c_g \times r_2 \times (gBest - X_i) \right\} \quad (1)$$

$$X_{(i+1)} = X_i + V_{(i+1)} \quad (2)$$

Where Cp is the Cognitive learning rate and Cg is the social learning rate. The factors r1 and r2 are randomly generated within the range (Wei Liang, 2014) and w is the inertia factor. The equation is essentially made up of three parts. The first part is the former speed of the swarm, which shows the present state of the swarm; the second part is the cognition modal, which expresses the thought of the individual swarm particle itself; the third part is the social modal. These three parts together determine the solution space searching ability. The first part has the ability to balance the whole and search a local part. The second part causes the swarm to have a strong ability to search the whole and avoid local minimum. The third part reflects the information sharing among the swarms. Under the influence of all the three parts, the swarm can reach an effective and best position (Omkar et.al 2009).

**Quantum behaved Particle Swarm Optimization**

The main disadvantage of PSO is that global convergence cannot be guaranteed (Bergh & F. V. D, 2001). To deal with this problem, concept of a global convergence guaranteed method called as Quantum behaved PSO (QPSO), was developed and reported at conferences (Sun et.al, 2005; Sun et.al, 2004; Cai et.al.,2008).Work was also done to develop a complete mathematical concept and the parameter control method of the QPSO (Cai et.al, 2008).

In the quantum model of a PSO, the state of a particle is depicted by wave function $\Psi_{(x,\ t)}$, instead of position and velocity.

The dynamic behaviour of the particle is widely divergent from that of the particle in traditional PSO systems in that the exact values of X and V cannot be determined simultaneously. We can only learn the probability of the particle's appearing in position x from probability density function $\left|\Psi_{(x,t)}\right|^2$, the form of which depends on the potential field the particle lies in. The particles move according to the following iterative equations:

$$X_{(t+1)} = P_i - \beta \times (mBest - X_t) \times \ln(1/u) \text{ if } k \geq 0.5 \qquad (3)$$

$$X_{(t+1)} = P_i + \beta \times (mBest - X_t) \times \ln(1/u) \text{ if } k<0.5 \qquad (4)$$

where,

$$P_i = \varphi \times pBest + (1-\varphi) \times gBest_i \qquad (5)$$

$$mBest = \frac{1}{N}\sum_{i=1}^{N} pBest_i \qquad (6)$$

mBest is the Mean Best position defined as the mean of all the best positions of the population, k, u and φ are random number distributed uniformly on [0, 1] respectively. Considering that the number of iterations and population size are common requirements in every evolutionary algorithm, ß, called Contraction–Expansion coefficient, it is the only parameter in QPSO algorithm. It can be tuned to control the convergence speed of the algorithms.

QPSO is very easy to be understood and implemented and it has already been tried and tested in various standard optimization problems with excellent results (Sun et.al, 2004; Sun & Feng, 2004; Sun & T. H, 2008). Moreover, QPSO algorithm is proven to be more effective than traditional algorithms in most cases (Sun et.al, 2004; Sun & Feng, 2005; Sun & Feng 2004).

**Simulation result**

In this work, we concentrate on the classification on the normal and arrhythmia database. The sampling frequency of the ECG signals in this database is Fs=360Hz. Arrhythmia database and normal database have been labelled by 1 and 2. 17 data are selected for normal class and 20 data are chosen for arrhythmia class. It was selected 4 data for training and the rest of data is considered for the test.

It should be noted that in the proposed ANFIS, a small set of data is considered for training. It can be understood that these set of data are able to classify the test data without any errors. It is classified the test data into two classes 1 and 2.

Figure 5-8 plotted for better showing result. The scatter plot is shown in figure 5. It is used when a variable exists that is below the control of the experimenter. One of the aspects of a scatter plot, however, is its ability to show nonlinear relationships between variables. Output of ANFIS is decimal data so we have to use round function for converting decimal data to integer data, because we have two classes one and two. Finally, all data is classified into two classes 1 and 2.

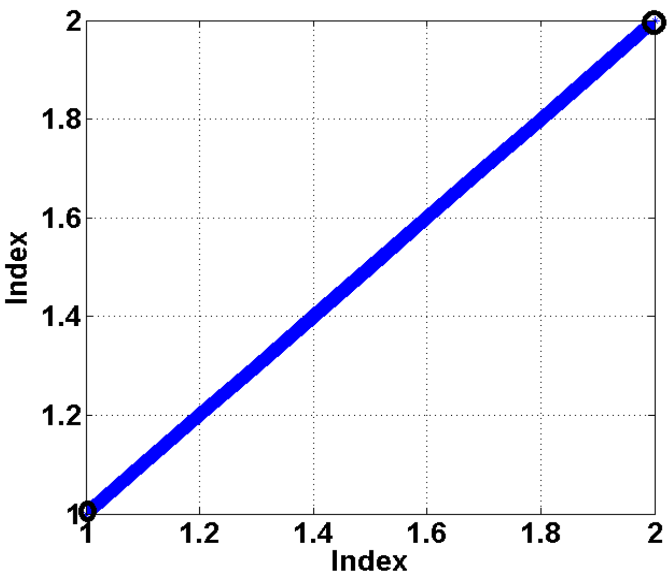


**Figure 5. Scatter diagram**

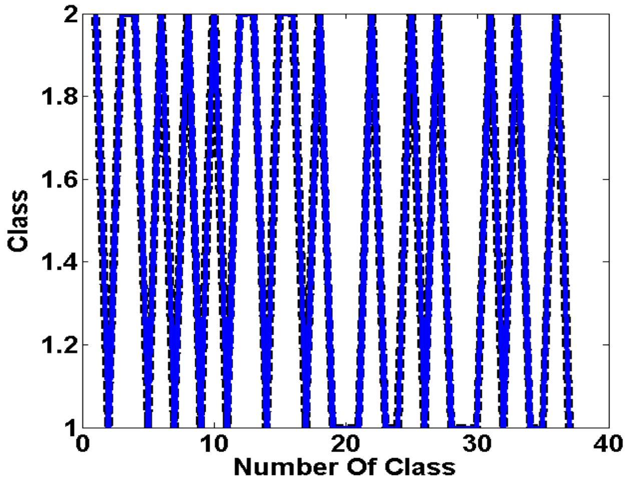


**Figure 6. Result of classification**

Figure 6 shows this matter that the black lines are represented the raw data and the blue lines are shown the output of the ANFIS. These two lines are covered each other exactly as shown in that figure and therefore the superiority of the proposed method can be understood.

Figure 7 shows the gBest. When a horizontal line is fixed, it is indicated that the better results are obtained. Population size and number of generation are set as 1000 in the proposed modelling strategy. These values were chosen because they have been classified all data with 100% of accuracy.

In Figure 8 ‘O’ is actual value of data and ‘*’ is output of ANFIS. All of data are distributed in two classes one and two. According to this figure actual value and output of ANFIS are completely overlapped for all of data.

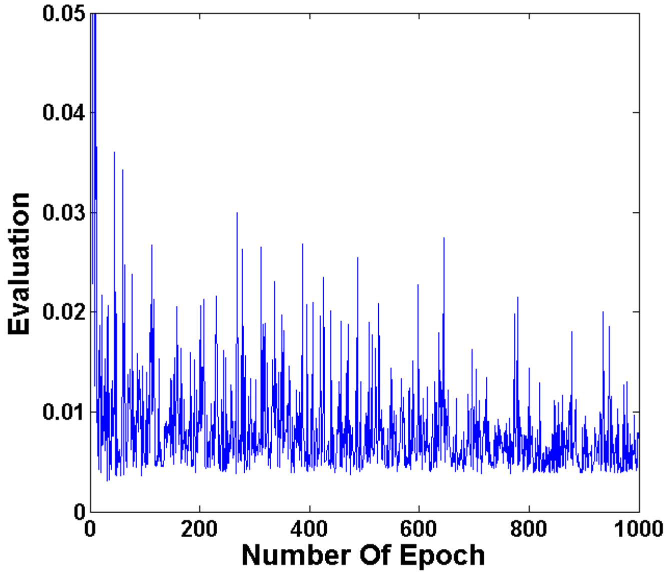


**Figure 7.gBest**

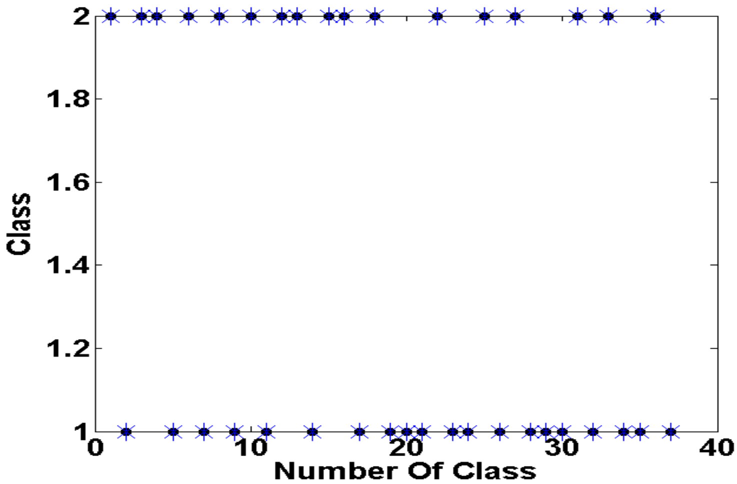


**Figure 8. Result of classification**

**Conclusion**

The abnormality detection of the ECG signal based on ANFIS and DWT is 100% efficient. The source of the ECGs data was obtained by MIT-BIH Arrhythmia database that 17 data are selected for normal class and 20 data are chosen for arrhythmia class. In this study, a novel strategy for data clustering using ANFIS has been proposed in which the statistical parameters were extracted from DWT to cover the important diagnostic information of cardiac event. The feature extraction criterion enhanced the reliability and decreased the structural complexity of the ECG classifier. This ANFIS model is trained with Quantum Behaved Particle Swarm Optimization. This strategy have been used a few data for training, furthermore, this model is simple and quickly. This model can be used to classify different types of heart disease by electrocardiogram.

**References**

Abawajy,, J.H., A.V.Kelarev, &M.Chowdhury(2013). Multistage Approach for Clustering and Classification of ECG Data. Computer Methods and Programs in Biomedicine.

Acharya, U.R., P.S. Bhat, S.S. Iyengar, A. Rao,& S. Dua(2003).Classification of heart rate data using artificial neural network and fuzzy equivalence relation. Pattern Recognition 36, 61–68.

Afonso, V.X., W.J. Tomkins, T.Q. Nguyen, & S. Luo(1999), ECG beat detection using filter banks, IEEE Transaction on Biomedical Engineering 46, 192–202.

Bergh, F. V. D. (2001). An analysis of particle swarm optimizers. Ph.D. Thesis, Department of Computer Science, University of Pretoria, Pretoria, South Africa.

Cai, Y. et al. (2008). Optimizing the codon usage of synthetic gene with QPSO algorithm. Journal of Theoretical Biology.

Ceylan, R., & Y. Ozbay (2007). Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. Expert Systems with Applications 33, 286–295.

Christov, I., G. Gómez-Herrero, V. Krasteva, I. Jekova, A.Gotchev., & K. Egiazarian (2006). Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification. Medical Engineering & Physics 28 876–887.

Dokur, Z., & T. Olmez (2001). ECG beat classification by a novel hybrid neural network. Computer Methods and Programs in Biomedicine 66, 167–181.

Eberhart, R. C., & Kennedy, J. A. (1995a). New optimizer using particles swarm theory. In Sixth international symposium on micro machine and human science, Nagoya, Japan (pp. 39–43).

Eberhart, R. C., & Kennedy, J. A. (1995b). Particle swarm optimization. In Proceedings of the IEEE international conference on neural networks, Perth, Australia (pp.1942–1948).

Engin, M. (2004).ECG beat classification using neuro-fuzzy network. Pattern Recognition Letters 25, 1715–1722.

Froese, T., S. Hadjiloucas, R.K.H. Galvao, V.M. Becerra, & C.J. Coelho (2006). Comparison of extra systolic ECG signal classifiers using discrete wavelet transforms. Pattern Recognition Letters 27, 393–407.

Gacek, A., & W. Pedrycz (2006).A granular description of ECG signals. IEEE Trans. on Biomed. Eng. 53 (10), 1972–1982.

Hari Mohan, R., & Trivedi, A. (2012).ECG Signal Classification using Wavelet Transform and Back Propagation Neural Network. 5$^{th}$ international conference on computers and device for communication.

Kezi, C., & Selva V., P. Kanagasabapathy, & S. Johnson (2005).Adaptive Neuro Fuzzy Inference System for Extraction of FECG, IEEE Indicon 2005 Conference, Chennai, India, I I - 1 3.

Kumar Sahu, N., & Ayub, Sh., & J. P. Saini (2013). Detection of Normal ECG and Arrhythmia Using Adaptive Neuro-Fuzzy Interface System. International Journal of Advanced Research in Computer Science and Software Engineering, 925-932.

Kutlu, Y. & D. Kntalp (2012). Feature extraction for ECG heartbeats using higher order statistics of WPD coefficient.Computer Methods and Programs in Biomedicine, 105,257-267

Li, C.W.,& C.X. Zheng, C.F. Tai (1995). Detection of ECG characteristic points using wavelet transforms. IEEE Transaction on Biomedical Engineering 42, 21–28.

Liang, W., Zhang, Y., Jindong T., & Li, Y. (2014). A Novel Approach to ECG Classification Based upon Two-Layered HMMs in Body Sensor Networks. Sensors, 5994-6011.

Mamdani, E. H (1974). Applications of fuzzy algorithms for control for a simple dynamic plant. Proc. IEEE. 121(12):1585–1588.

Omkar, S.N., Rahul Khandelwal, T.V.S. Ananth, G. Narayana Naik, & S. Gopalakrishnan (2009). Quantum behaved Particle Swarm Optimization (QPSO) for multi-objective design optimization of composite structures. Expert Systems with Applications 36, 11312–11322.

Ozbay, Y., & B. Karlik (2001). A recognition of ECG arrhythmias using artificial neural networks. Proceedings of the 23rd Annual Conference, IEEE/EMBS,I˙stanbul, Turkey, Oct. 25–28.

Ozbay, Y., & Tezel, G.(2010). A new method for classification of ECG arrhythmias using neural network with adaptive activation function. Digital Signal Processing, 1040-1049.

Ozbay, Y., R. Ceylan, & B. Karlik (2006). A fuzzy clustering neural network architecture for classification of ECG arrhythmias. Computers in Biology and Medicine 36, 376–388.

Pannizzo, F., & S. Furman (1998). Frequency spectra of ventricular tachycardia and sinus rhythm in human intracardiac electrograms: application to tachycardia for cardiac pacemakers. IEEE Transaction on Biomedical Engineering 35, 421–425.

Silipo, R., & C. Marchesi (1998). Artificial neural networks for automatic ECG analysis, Signal Processing 46, 1417–1425.

Srivastava1, V.K., Prasad, D. (2013). DWT - Based Feature Extraction from ECG Signal. American Journal of Engineering Research, 44-53.

Sun, J., Feng, B., & Xu, W. B. (2004). Particle swarm optimization with particles having quantum behavior. IEEE Proceedings of Congress on Evolutionary Computation, 325–331.

Sun, J., Feng, B., & Xu, W. B. (2005). Adaptive parameter control for quantum behaved particle swarm optimization on individual level. In Proceedings of the 2005 IEEE international conference on systems, man and cybernetics, Piscataway, NJ (pp. 3049–3054).

Sun, J., Xu, W. B., & Feng, B. (2004). A global search strategy of quantum-behaved particle swarm optimization. Cybernetics and Intelligent Systems Proceedings of the 2004 IEEE Conference, 111–116.

Sun, T. H. (2008). Applying particle swarm optimization algorithm to roundness measurement. Expert Systems with Applications, 36(2P2) 3428–3438.

Tsukamoto, Y (1979). An approach to fuzzy reasoning method. Gupta, M. M., Ragade, R. K., and Yager, R. R., (Eds.), Advances in Fuzzy Set Theory and Applications, 137:149. North-Holland, Amsterdam.

Vimala, K., & V.Kalaivani (2013). Classification of Cardiac Vascular Diseasef ECG Signals for Enhancing Modern Health Care SCENARIO. Health Informatics- An International Journal (HIIJ) 63-72.

Wang, L.-X.,& Mendel, J.-M. (1993). Generating fuzzy rules by learning from examples. IEEE Trans. Syst. Man Cybern. 22 (6):1414–1427.

Yu, S.N., & K.T. Chou (2008). Integration of independent component analysis and neural networks for ECG beat classification. Expert Systems with Applications 34, 2841–2846.