Classification of Cardiac Arrhythmias with TSK Fuzzy System using Genetic Algorithm

Naser Safdarian, Keivan Maghooli and Nader Jafarnia Dabanloo

Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
n.safdarian@srbiau.ac.ir , k_maghooli@srbiau.ac.ir , jafarnia@srbiau.ac.ir

Abstract

Detection of cardiac arrhythmias, particularly ventricular fibrillation (VF), and ventricular tachycardia (VT) have been highly regarded and has done several works in this field. In this study, a method based on the Takagi-Sugeno-Kang (TSK) fuzzy system for ECG arrhythmia detection and classification of normal sinus rhythm (NSR), ventricular fibrillation (VF) and ventricular tachycardia (VT) has been used. ECG arrhythmia signals have been obtained from MIT-BIH database. At the first, preprocessing is performed on the signals to get a signal without any noise. Then two features of ECG signals include an average period T (i.e. the time interval between two R peaks) and amplitude of QRS complex, are used as inputs to fuzzy classifier. The triangular membership functions for converts crisp input values (features of ECG signals) to the fuzzy values are used to provide the fuzzy system. Using genetic algorithms, optimization rules with membership functions by minimizing the error function, and convert them to proper rules and membership functions for classifying arrhythmias do with high accuracy. Finally, we achieved the classification accuracy for normal signals (NSR) 91.66%, for VT signals 92.86% and for VT signals equal to 100%. We obtained the overall accuracy of the classifier 93.33%. Also, sensitivity for NSR signals is equal 92.30%, for VT signals is 93.33% and for VF signals is equal to 100%. Specificity for NSR, VT and VF signals is equal to 94.44%, 93.57% and 100% respectively. The simply of propose method can be considered as its major advantage.

Keywords: *Cardiac Arrhythmias, ECG Signal Processing, Classification of ECG Signal, TSK Fuzzy System, Genetic Algorithm*

1. Introduction

There are changes in heart activity due to the existence of various cardiac arrhythmias, and when these arrhythmias occur for a long time, they are very dangerous for humans. Many of these arrhythmias have obvious signs in the ECG signal (especially in QRS complex) and the sequence of beats with the timing and shape are unusual. The dangerous arrhythmias e.g. ventricular fibrillation and ventricular tachycardia as ventricular arrhythmias are very dangerous that cause abnormal changes in the Ventricle [6]-[11]. In ventricular tachycardia arrhythmia, the heart rate usually reaches 120 to 200 beats per minute and its wave shape may be similar to the previous complex or can be different from beat to beat. Stable VT is an emergency condition that can predict cardiac arrest and immediate treatment is required. VT is usually with an acute myocardial infarction or ischemic heart, electrolyte severe disorders and conditions that can cause QT interval prolongation. Hence, this arrhythmia is a serious threat for health of the heart.

Ventricular fibrillation is the stage before death and almost always is seen in the dying hearts. VF is the most common type of ventricular arrhythmia that young people suffer sudden death. In fact, when the ventricular fibrillation occurs, cardiac output cannot be complete, and cardiac resuscitation and electrical defibrillation should be done quickly. Therefore, it can result seriously harm in the body. Thus need accurate identification of cardiac arrhythmias from the ECG signal through automated methods was always considered by the researchers and many studies have been done in this area. Today's intelligent diagnostic systems are most important industries in medical studies. In this area, fuzzy logic for classifying operations, especially in cardiology is a suitable method [7]. The purpose of this study, design a system based on the fuzzy logic and genetic algorithms for classification of normal cardiac sinus rhythm (NSR), VT and VF as two ventricular arrhythmias.

2. Review of Previous Research

In 1999, Minami et al applied the Fourier transform based on frequency domain techniques for classification of ventricle arrhythmias such as ventricular tachycardia and ventricular fibrillation [13]. Also in 1990, NV Thakor et al presented a method using sequential hypothesis testing algorithm to identify VT and VF [14]. In 2009, Phan Anh Phong et al presented the method based on TSK fuzzy system for classification of cardiac arrhythmias [4]. In that study, two features of the ECG signal, mean period (T) and pulse width of QRS as the input of the fuzzy classifier was used, and fuzzy rules had been determined by an expert. In 2010, Yun-Chi Yeh et al presented diagnosis of cardiac arrhythmias using fuzzy C-Means method [1, 2]. They used Normal heart rhythm (NORM), arrhythmias and abnormal heart beats, including left bundle branch block (LBBB), right bundle branch block (RBBB), ventricular premature contractions (VPC) and atrial premature contractions (APC) for their Studied. In that study, they only used the features available in the QRS complex. In 2010, N. Ö zlem Ö ZCAN et al presented a method for classifying cardiac arrhythmias that SVM has been combined with fuzzy (FSVM) [3]. In a research, N. Ö zlem Ö ZCAN et al used ANFIS toolbox of MATLAB to extract the rule bases. They used QRS complex duration and heart rate as two features of the ECG signal.

Generally in most research, the signal of normal heart with cardiac arrhythmias has been studied and some of these researches presented acceptable accuracy in the classification results.

3. Introduce arrhythmias using in this classification

An Electrocardiogram (ECG) signal represents the electrical activity of heart muscle that it is recorded on the body surface. Figure 1 shows the various components of an ECG signal. P wave represents atrial depolarization, while the QRS complex represent ventricular depolarization and the T wave represent ventricular repolarization [8].

Ventricular fibrillation and ventricular tachycardia are two types of dangerous cardiac arrhythmias that the occurrence of these arrhythmias threat human life. Critical cardiac events often occur outside the hospital. So to solve this problem automated external defibrillation (AED) was provided. Since completion of VT and VF require rapid response and deployment of high-energy shocks on the heart, accuracy of the algorithm for the detection of these arrhythmias is very important. Therefore, automatic recognition should also be consistent with the experts [8].

In this section, algorithms based on time domain [14, 15] or frequency domain [16], or the combination of time and frequency techniques to classify cardiac arrhythmias will be reviewed. Usually, methods in time domain are superior from the frequency domain methods, because they are simple calculations, while frequency method for classification of VF and VT are much more reliable. In addition Time domain features are usually considered qualitative features [5]. From a medical perspective, quality Features such as RR interval and QRS amplitude are generally for understanding the characteristics against Features of the frequency domain. Unfortunately, the most ECG signals are inherently contain noise and hence the extracted features in the time domain are unknown.

This is due to high sensitivity of ECG signals to electrode movement and muscle activity. In addition, electrical incoming interference from the network can affect the recorded ECG. Although this artifacts can be reduced by filtering techniques, but their complete removal is almost impossible. In detection QRS of the ECG signal, we will have the wrong detection when the peak amplitude of R wave was combined with amplitude noise. In addition, missing a peak when this peak is relatively lower than other peaks is more probable [4].

Both features that presented in this study are common for various forms of VT and VF signals. It is difficult to extract features from noisy ECG signals in time domain for classification of VT, VF, NSR signals. Therefore we use such signals for evaluation and increase validation of the proposed classifier.

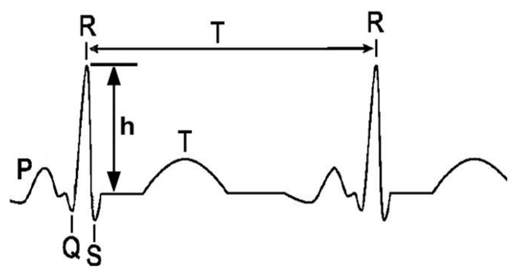

Figure 1. ECG Signal Components: P wave, QRS complex and T wave

4. Preprocessing and Feature Extraction of ECG Signals

In this study, signals malignant ventricular arrhythmias database (VFDB) obtained from the MIT-BIH, has been used [9]. For pre-processing of the signals we used noise removal algorithms that remove the DC component and their deviation from the baseline, and finally, noise power of the signals were removed by a 60 Hz notch filter. Results of this process were shown in Figure 2.

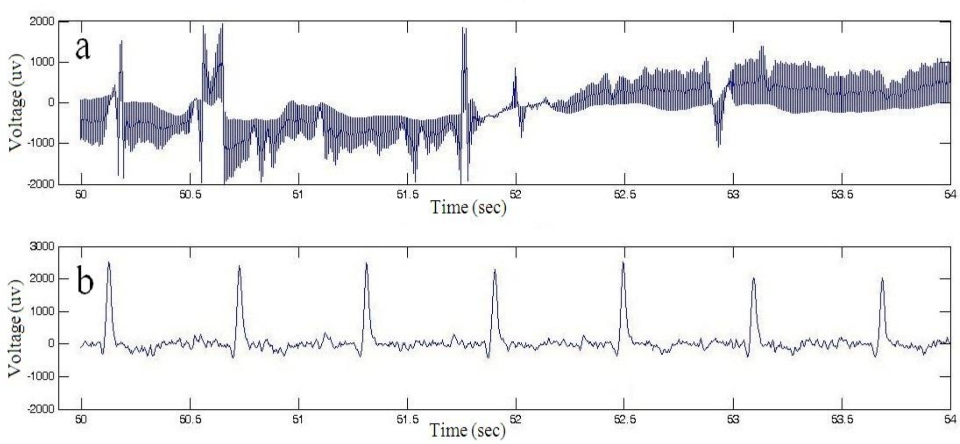

Figure 2. a) First Signal Obtained from MIT-BIH; b) Clean Signal

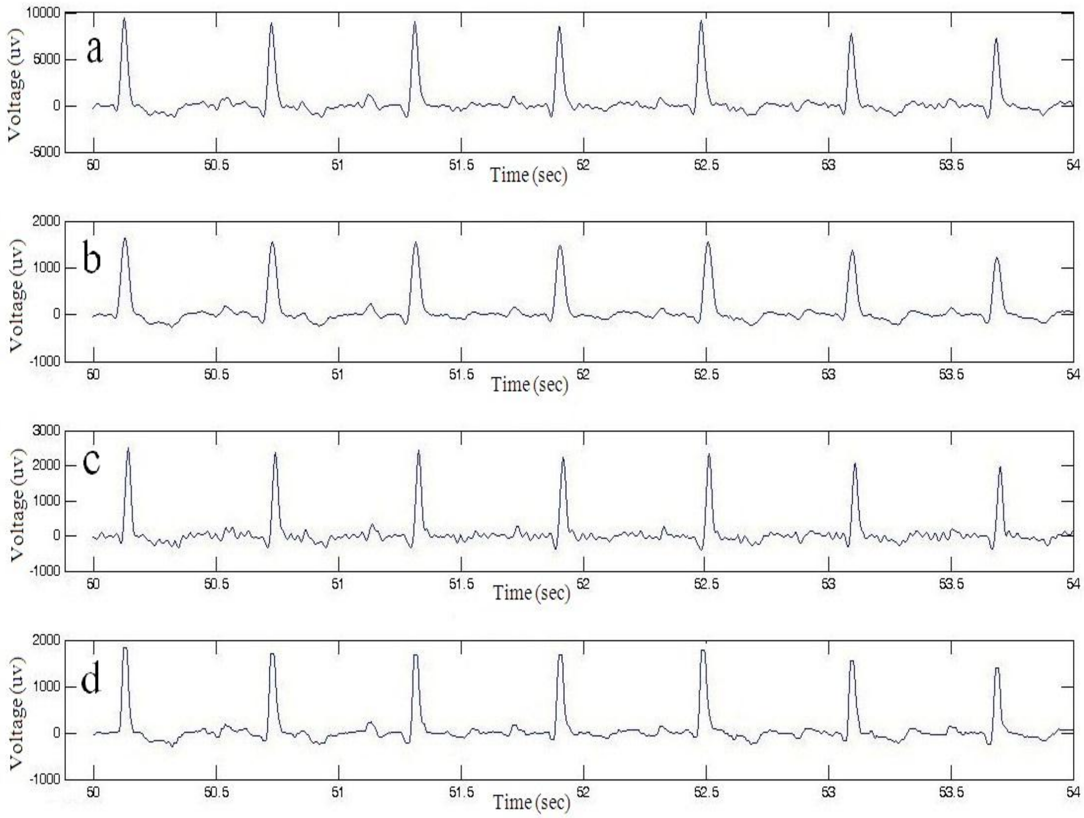

Figure 3. Passing Clean Signal Through Four Types Smoothing Filter

a) Kaiser Filter , b) Moving Average Filter , c) Butterworth Filter, d) Median Filter

Clean signals were passed through different smoothing filters that this operation is shown in Figure 3. Then we extracted two features that are common in the ECG signal for classifying cardiac arrhythmias.

1) The period average T between the R wave (R-R Interval Duration);
2) Amplitude of the QRS complex;

Figure 1 shows these features. To extract features, the R wave and the time of its occurrence at any rate must determine. Figure 4 shows the result of this operation performed on a sample of the normal signal.

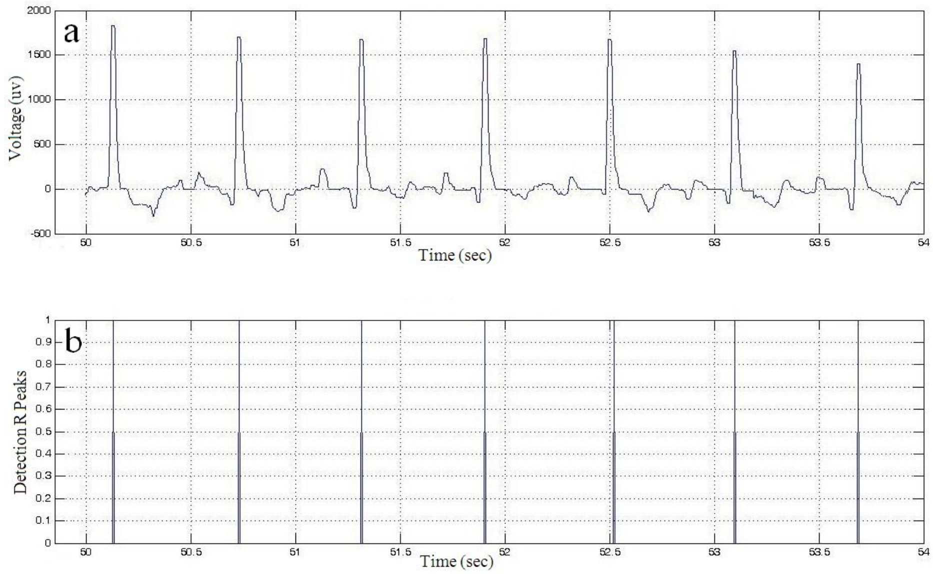

Figure 4. a) R Waves Extracted from the Sample Normal Signal, and b) Determine the Time Position of R Waves Extracted

5. Fuzzy Classifier System Design

The fuzzifier converts crisp inputs to the fuzzy sets that called Small, Medium and Large. In this study, each antecedent fuzzy set was defined by triangular fuzzy membership function (Figure 5 and Figure 6). The membership function of Medium is require three parameters (point of the left L_i , central point m_i and point of the right r_i), and membership functions of small and large are require two parameters. In this study, we used the TSK fuzzy classifier system.

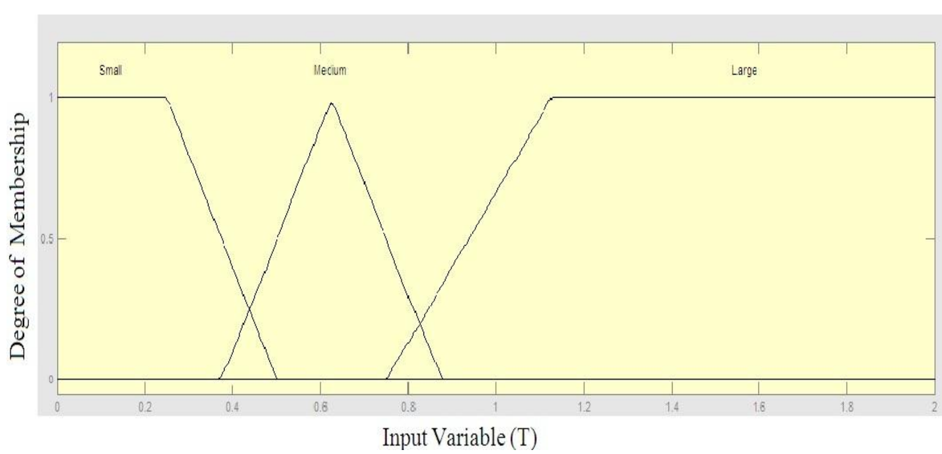

Figure 5. Triangular Membership Function for T (first feature of ECG signal) as the First Input of Fuzzy System

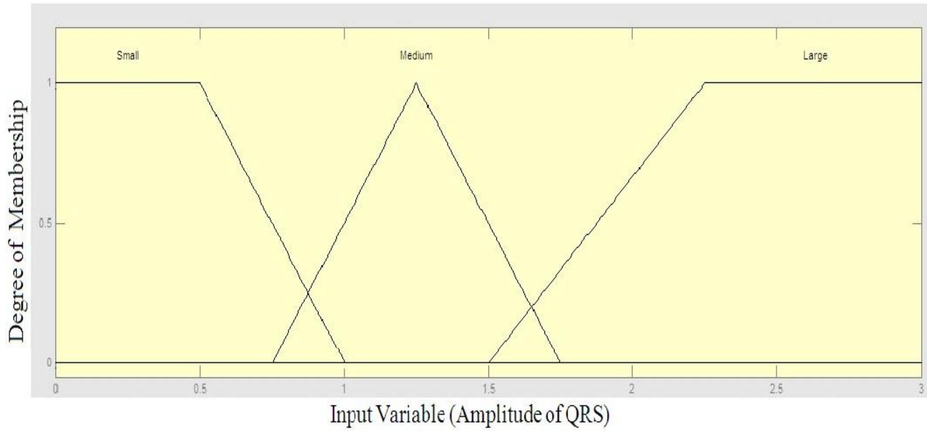

Figure 6. Triangular Membership Function for Amplitude of QRS Complex (second feature of ECG signal) as the Second Input of Fuzzy System

The structure of this classifier is a set of nine fuzzy IF-THEN rules base that they are as follows:

$$x1 \text{ is } Fj1(r) \text{ and } x2 \text{ is } Fj2(r) \text{ , THEN } y \text{ is } Class(r) \tag{1}$$

Where class(r) = {1,2,3} as " NSR,VF,VT " , r=1,2,..,9 and j_1=j_2={1,2,3} as " Small , Medium , Large " . x_1 and x_2 are two features that extracted from the ECG signal as input fuzzy systems, and y is the output of the TSK fuzzy system that is define as:

$$y = ax1 + bx2 + c \tag{2}$$

Where the 'a', 'b' and 'c' represented output coefficients of the fuzzy system.

6. Genetic Algorithm to Optimize the Classifier Parameters

This section is discussed on how to formulate the fuzzy classifier by using genetic algorithms. At this stage, the parameters of the triangular membership functions, fuzzy system output parameters and the basic rules are randomly selected. Then these cases were converted by using genetic algorithm as the most efficient mode for optimum performance of classifier.

A genetic algorithm was generated with a random population of chromosomes and by specific genetic operators goes to the better chromosome. Population is undergoes evolution in a form of natural selection. During the repetitions consecutive that is called generation, chromosomes in the population assessed and as suitable solutions are proposed, and based on this assessment is formed a new population of chromosomes using selection mechanism and particular genetic operators (such as mutation and crossover) [8]. Genetic algorithms to create the next generation selected some people have called the parents of the current population and uses them to create the next generation. This algorithm chose parents that have better fitness.

Generally genetic algorithm uses the following steps to achieve the premier generation:

a) algorithm selects initial population as randomly;

b) The next population from the current population creates by algorithm. To obtain the next generation, algorithm runs the following:

- Vote out every member of the population with calculate fitness value;
- Scaled scores to better usage from the scores and population;
- Selected the next generation of parents, reproduction with changing organ of a parent, genetic mutation or combination of parents for reproduction;

Select parents based on their fitness;

- Replacing the old with the new generation;

c) The algorithm stops when a predetermined stop criterion reached.

Also, when designing a fuzzy system using genetic algorithms, the operation of the coding system in the chromosome is very important. In our design, two inputs exists (features of ECG) and each fuzzy input variable (x1, x2) in three membership functions were partitioned. Consequently, these cases will result 14 parameters in chromosome (for two input, there are 3 membership functions that all parameters will determine with 7 value). Moreover, there are 9 rules in the rule base, that part of THEN in this rules will result 9 parameters in chromosome. Also, the output of TSK fuzzy system has 9 parameters (for three classes, there are 3 output of y and each output have 3 parameters that called 'a',' b' and 'c'). As a result, that outputs can also result 9 parameters in chromosome. Hence, altogether 32 parameters (or 32 genes) were needed to determine chromosome. Chromosome structure was shown in Figure 7.

$L^1_{x_1}$	$m^1_{x_1}$	$r^1_{x_1}$	...	$L^3_{x_2}$	$m^3_{x_2}$	$r^3_{x_2}$	a_1	b_1	c_1	...	a_3	b_3	c_3	r_1	...	r_9

Figure 7. Chromosome Structure Designed

It should be noted that the input and output parameters of fuzzy system were encoded as 8-bit and parameters for part of THEN in all rules were encoded as 2-bit. Thus, the parameters are generally in 202-bit encryption. Another important issue is the choice of fitness function. A good fitness function can be reflects the aim of the system. How describing the fitness function is an important issue. Generally, the fitness function can be expressed as the number of classes that have been classified correctly [8]. Fitness function was proposed as follows:

$$fitness = \frac{AC_{NSR} + AC_{VF} + AC_{VT}}{3} \quad (3).$$

Where, AC is represents the percent of classes are correctly classified.

After evaluating of the each chromosome and its correlation with the fitness, present population is done process of reproduction for the generation of next population. Changes in the genetic operators (crossover and mutation) play an important role in genetic algorithms. In this study, both genetic operators -crossover and mutation- were used for obtain research purposes.

7. Classification of Cardiac Arrhythmias

As noted, the initial parameters of fuzzy membership functions, also the output parameters of TSK fuzzy system and second part of the fuzzy rules (THEN) are set as random. Then by rules expressed, the classification of ECG signals using the database as input of fuzzy system is done, and finally the overall accuracy of classification was presented as a percentage. In continue, our aim is maximum this accuracy.

The initial values of the chromosome (which was stated in the previous section), was presented to the genetic algorithm. By applying crossover as a genetic operator, new chromosomes were generated and Classification operations again with a new range of values and parameters, by all ECG signals were used as input was expressed. Finally, the overall accuracy of the classifier was calculated. The above steps continue for several times, until overall accuracy reach its maximum. It should be noted that after some time using genetic crossover operator, accuracy of classifier in the early stages is increased then decreased that after this reduction, rather than the crossover, genetic mutation operator was used. After using several different mutations in the chromosome and extracted the parameters from the chromosome, most optimal the chromosome was considered. Termination of genetic algorithm was achieved the maximum accuracy in the classification of ECG signals. In fact, after each time the genetic mutation (or crossover) operator, parameters of chromosome were extracted and then classify the signals with the new values was done. When the accuracy of the classifier was reached optimum value and it did not exceed from this, the genetic algorithm was stopped.

Now, we reach to the optimal conditions for the input range (the range of the triangular membership functions), and output values of the TSK fuzzy system and parameters in the second part of fuzzy rules (THEN). Therefore, fuzzy rules with the correct parameters for classification operation were available and classification operation with high accuracy and lowest error rate was done.

8. Result

In this study, accuracy of classification in each class is defined as number of signals that they have correct class label by classifier per total number of signals in same class. Also, overall classification accuracy is defined as total number of signals that correctly detected by proposed classifier, per the total number of signals have been considered. Also, sensitivity and specificity were defined according to table 1 and equations 4 and 5.

Table 1. Table For Sensitivity And Specificity In Proposed Model

		Algorithm Label		
		NSR	**VT**	**VF**
Reference Label	**NSR**	*NSR-NSR*	*NSR-VT*	*NSR-VF*
	VT	*VT-NSR*	*VT-VT*	*VT-VF*
	VF	*VF-NSR*	*VF-VT*	*VF-VF*

NSR Signals :

$TP = NSR - NSR$

$FN = NSR - VT, NSR - VF$

$FP = VT - NSR, VF - NSR$

$TN = VT - VT, VT - VF, VF - VT, VF - VF$

VT Signals :

$TP = VT - VT$

$FN = VT - NSR, VT - VF$

$FP = NSR - VT, VF - VT$

$TN = NSR - NSR, NSR - VF, VF - NSR, VF - VF$

VF Signals :

$TP = VF - VF$

$FN = VF - NSR, VF - VT$

$FP = NSR - VF, VT - VF$

$TN = NSR - NSR, NSR - VT, VT - NSR, VT - VT$

(4).

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

(5).

In the proposed method, after all these steps the classification accuracy for normal sinus rhythm (NSR) is equal 91.66%, for VT signals is 92.86% and for VF signals is equal to 100%, and overall accuracy of the classifier is equal to 93.33%. Also, sensitivity for NSR signals is equal 92.30%, for VT signals is 93.33% and for VF signals is equal to 100%. Specificity for NSR, VT and VF signals is equal to 94.44%, 93.57% and 100% respectively. This result shows good performance for classification of normal sinus rhythm, ventricular tachycardia and ventricular fibrillation.

It should be noted that to examine amount of the features discussed in proposed method, in general we used from five cycles of each ECG signals and finally after calculating the results for this five cycle, by averaging from this results we reached to the final amount for tow feature of each ECG signal, and these values were used for final conclusions.

9. Conclusion and Discussion

In this study, a method based on TSK fuzzy classifier using genetic algorithm to optimize triangular membership functions, outputs and rules of fuzzy system to determine ECG signals include of normal sinus rhythm (NSR), ventricular fibrillation (VF) and ventricular tachycardia (VT) was proposed. In this section, accuracy of result from proposed method has been compared with several previous researches in this field. Also, all this research used the VFDB database from MIT-BIH. Table 2 shows this comparison.

Table 2. Comparison Accuracy Results In Proposed Method with Several Previous Researches

Algorithm	NSR	VF	VT
	AC (%)	AC (%)	AC (%)
Rate and Irregularity Analysis [10]	100	87.85	87.75
VF-Filter Leakage [10]	97.55	89.22	89.71
Fuzzy Rule-Based Method [12]	94.30	78	82
Type-2 Fuzzy System [6]	100	84	90.91
Interval Type-2 TSK Fuzzy System [4]	100	93.3	92
Proposed Algorithm	**91.66**	**100**	**92.86**

According to the table 2, the results of the accurate classified in VF signals compared to other studies done, show that proposed method classified this signals with maximum accuracy that show the superiority of the accuracy in proposed method for classification of these signals. Also, according to the results obtained from the classification of VT signals, in this case the accuracy of the proposed method is superior than other studies, but need to improve this was felt. In addition, according to the table 2 can be observed that the accuracy of the method proposed for classification of NSR signals despite the high ratio, is not superior to the accuracy results of other studies. It should be noted that is not obtain the maximum accuracy for the classification of signals NSR and VT, due to overlap with features extracted from each of the these signals. To increase and improve the classification accuracy of these signals, the characteristics of these signals that do not overlap with together for input of the classifier to be extracted.

Generally, accuracy of the results obtained in this study (except results for classification of NSR signals) is better than the accuracy of other studies. But in general need to improve the classification results in signals NSR and VT and also obtain maximum accuracy, sensitivity and specificity for those signals.

It should be noted that combination of several methods does not essentially improve the quality of a technical work. But in fact the goal of use the genetic algorithm is optimize parameters of fuzzy system, where that the initial parameters of fuzzy system were obtained as approximately by studies the medical references. In fact, after several times used genetic algorithms and reached to the new parameters of fuzzy systems at each stage, the operation classified was performed, and finally the genetic algorithm converted parameters of fuzzy system to the optimal position. Also, in this work we only in first step of designing fuzzy systems used medical references to determine parameters and in continue with introduced intelligent approach to achieve the good classifier with high accuracy, that this can be an advantage for this study. Also the simply of proposed method can be considered as its major advantage.

In next section, we present suggestions for future studies to improve the classifier performance.

10. Suggestions and Future Studies

There are several suggestions for future studies. In order to improve the performance of classifier, more feature extraction of ECG signals and use these features for increase inputs to increase the validation and accuracy of classifier, and use the other types of fuzzifier membership functions and applying more rules in fuzzy system design to enhance accuracy of proposed classifier are suggest.

Acknowledgements

With special thanks from Bioelectric Group of Biomedical Engineering Department, Science and Research Branch, Islamic Azad University of Tehran.

References

[1] Y. –C. Yeh and H. –J. Lin, “Cardiac Arrhythmia Diagnosis Method Using Fuzzy C-Means Algorithm on ECG Signals”, International Symposium on Computer, Communication , Control and Automation, **(2010)**, pp. 272-275.
[2] Y. –C. Yeh, et. Al., “A novel fuzzy c-means method for classifying heartbeat cases from ECG signals”, Measurement, vol. 43, no. 10, **(2010)**, pp.1542–1555.
[3] N. Özlem ÖZCAN, Fikret GÜRGEN, “Fuzzy Support Vector Machines for ECG Arrhythmia Detection”, International Conference on Pattern Recognition, **(2010)**, pp. 2973-2976.
[4] P. A. Phong, K. Q. Thien, “Classification of Cardiac Arrhythmias Using Interval Type-2 TSK Fuzzy System”, International Conference on Knowledge and Systems Engineering, **(2009)**, pp. 1-6.
[5] S. Z. Mahmoodabadi, A. Ahmadian, M. D. Abolhassani, J. Alireazie and P. Babyn, “ECG Arrhythmia Detection Using Fuzzy Classifiers”, Fuzzy Information Processing Society. NAFIPS '07, Annual Meeting of the North American, **(2007)**, pp. 48-53.
[6] W. W. Tan, C. L. Foo and T. W. Chua, “Type-2 Fuzzy System For ECG Arrhythmic Classification”, In Proceedings of the IEEE International Conference on Fuzzy Systems, London UK, **(2007)**, pp. 859-864.
[7] A. Messaoud, M. B. Messaoud, A. Kachouri and F. Sellami, “Fuzzy logic based system for classification of atrial fibrillation cardiac arrhythmias”, 13th IEEE International Conference on Electronics, Circuits and Systems, ICECS '06, **(2006)**, pp. 347-350.
[8] T. W. Chua, W. W. Tan, “GA Optimisation of Non-Singleton Fuzzy Logic System for ECG Classification”, IEEE Congress on Evolutionary Computation, **(2007)**, pp. 1677- 1684.
[9] www.PhysioNet.org, MIT/BIH Database Distribution, Massachusetts Inst. Technol., Cambridge, MA.
[10] X. S. Zhang, Y. S. Zhu, N. V. Thakor and Z. Z. Wang, “Detecting ventricular tachycardia and fibrillation by complexity measure”, IEEE Transactions on Biomedical Engineering, vol. 46, no. 5, **(1999)**, pp. 548-555.
[11] U. Ayesta, L. Serrano and I. Romero, “Complexity measure revisited: a new algorithm for classifying cardiac arrhythmias”, Proceedings of the 23rd Annual EMBS International Conference, **(2001)**, pp. 1589–1591.
[12] E. Chowdhury and L. C. Ludeman, “Discrimination of cardiac arrythmias using a fuzzy rule-based method”, Computers in Cardiology, **(1994)**, pp. 549–552.
[13] K. Minami, H. Nakajima and T. Toyoshima, “Real-time discrimination of ventricular tachyarrhythmia with Fourier transform neural network”, IEEE Transactions on Biomedical Engineering, pp. 46, no. 2, **(1999)**, pp. 179-185.
[14] N. V. Thakor, Y. S. Zhu and K. Y. Pan, “Ventricular Tachycardia and Fibrillation Detection by A Sequential Hypothesis Testing Algorithm,” IEEE Transactions on Biomedical Engineering, vol. 37, no. 9, **(1990)**, pp. 837-843.

[15] J. Ruiz, E. Aramendi, S. Ruiz de Gauna, A. Lazkano, L. A. Leturiondo and J. J. Guti'errez, “Distinction of Ventricular Fibrillation and Ventricular Tachycardia Using Cross Correlation”, Computers in Cardiology, vol. 30, **(2003)**, pp. 729-732.
[16] K. M. Ropella, A. V. Sahakian, J. M. Baerman and S. Swiryn, “The Coherence Spectrum: A Quantitative Discriminator of Fibrillatory and nonfibrillatory Cardiac Rhythms”, Circulation, vol. 80, **(1989)**, pp. 112-119.