GA-NN APPROACH FOR ECG FEATURE SELECTION IN RULE BASED ARRHYTHMIA CLASSIFICATION

*Gözde Aslantaş, Fikret Gürgen, Albert Ali Salah**

Abstract: Computer-aided ECG analysis is very important for early diagnosis of heart diseases. Automated ECG analysis integrated with experts' opinions may provide more accurate and reliable results for detection of arrhythmia. In this paper, a novel genetic algorithm-neural network (GA-NN) approach and a rule extraction method are proposed for contribution to automated ECG analysis. The goal of this paper is to classify arrhythmia in a rule-based manner to aid the cardiologists based on ECG features extracted from electrodes attached to specific positions on the skin and recorded by an electrocardiogram. Genetic algorithm (GA) - Neural Network (NN) approach is used for feature selection and arrhythmia classification. Accuracy and interpretability of results are more important than speed in this application. Based on selected features, rule extraction algorithms are performed to interpret classification results in human readable format.

Key words: *ECG, Genetic algorithm (GA), Neural network (NN)*

Received: December 1, 2012
Revised and accepted: ??

1. Introduction

Alterations that disrupt the regular functioning of the heart cycle may cause a cardiac arrhythmia. These anomalies might be a potential reason for a heart disease. Thus, early detection of arrhythmia can save lives. Electrocardiogram (ECG) is widely used for diagnosis of such abnormalities. ECG output must be integrated with medical assessment to provide more meaningful results. ECG analysis as a part of clinical assessment requires full attention and quickness as well as the knowledge and experience. Due to these kinds of requirements, automated ECG analysis integrated with experts' opinions provides more robust and reliable results to detect abnormal patterns.

Motivated by this reason, in this paper, we aim to detect arrhythmic patterns in a rule-based manner, and thus we intend to aid the cardiologists. We propose genetic algorithm-neural network approach (GA-NN) for ECG feature selection in

*Gözde Aslantaş, Fikret Gürgen, Albert Ali Salah
Department of Computer Engineering, Boğaziçi University, Istanbul, 34342, Turkey, gozde.aslantas@boun.edu.tr, gurgen@boun.edu.tr, salah@boun.edu.tr

rule-based arrhythmia classification. GA-NN approach performs feature selection and classification simultaneously. By this way, GA-NN approach finds relevant features which contribute to classification mostly. In this approach, neural network is the main classifier while genetic algorithm performs feature selection. Neural network weights are encoded in chromosomes and optimal neural network model is selected according to its performance on classification. Objective function is based on performance of the candidate networks which is related to classification error. We mainly utilize global search ability of genetic algorithm in weights space. At the end of this iterative search, a final model is obtained with the features having non-zero weights. Thus, we have a pruned classified data which can be processed so as to be interpreted in human readable format. We are mainly interested in producing robust and reliable arrhythmia detection results to be applicable in diagnostic decision support systems. So, rule extraction methods are applied to classification output. Applying solely rule extraction methods to original raw data will increase the complexity in produced rules as fully described in Halford *et al.* [37]. GA-NN approach is used as a preclassifier and feature selector to produce rules concisely.

Finding determinant features among the whole input is an important step for ruleset construction. Rules are useful as long as they are clearer, simpler and more precisely. For this reason, extraction of relevant features is an inevitable preprocessing for rule producing.

The organization of the paper is as follows: section 2 presents the related work. In section 3, we derive the proposed GA-NN approach in detail. Rule Extraction is explained in section 4. Several experiments performed on UCI dataset [1] and a new proprietary ECG database introduced in this paper, with the results are presented in section 5. In section 6, the proposed approach is discussed and section 7 concludes our study with a summary of the empirical results and a future work.

2. Related Work

ECG output is just a laboratory test result and must be integrated with a clinical assessment. Reasoning methods of clinicians may vary while performing analysis during these assessments. Moreover, an analysis of ECG tracing requires a deep attention and quickness. So, an automated ECG analysis attached with an ECG output might be a supportive element for the clinicians while concluding their evaluations. In literature, there are many completed and ongoing works to improve methods for automated ECG analysis that detect abnormal patterns. UCI Arrhythmia Repository [1] and MIT-BIH [2] are two most commonly used databases for ECG Analysis. Some researches that use these databases are summarized in Tab. I.

ECG tracing consists of waves which are referred as letters: P, Q, R, S, T, U. This scaled paper has horizontal and vertical lines. Along the horizontal lines, time is measured and voltage is measured along the vertical lines. Electrical activity during the cycle is plotted with waves. ECG features are derived from the waves, complexes, intervals, amplitudes and regularity of the waves.

P wave indicates atrial depolarization and duration of the P wave should not exceed 0.12 s. for normal rhythm.

Ref.	**Year**	**Protocol**	**Algorithm**	**Results**	**Dataset**
Yeap *et al.* [29]	1990	5% tra. 5% test. and 100% test.	ANN BP	98.36% sens. 67.80% sens.	AHA
Hu *et al.* [5]	1993	3-fold CV	51-25-2 MLP	≈90% acc.	MIT-BIH
Silipo and Marchesi [6]	1998	15% tra. 85% test.	ANN with BP	≈90% sens.	MIT-BIH and VALE
Chazal and Reilly [7]	2000	10-fold CV	LDA	69.3-74.7% acc.	Frank Lead ECG Data[8]
Gao *et al.* [9]	2005	t-test	Bayesian ANN	≈76% sens.	UCI
Niwas *et al.* [10]	2005	58% tra.42% test.	ANN	≈99% acc.	MIT-BIH
Zhang and Zhang [11]	2005	2/3 tra.1/3 test.	PCA-SVM	≈99% acc.	MIT-BIH
Song *et al.* [12]	2005	CV	LDA-SVM	99.35% avg. acc.	MIT-BIH
Uyar [13]	2006	10-fold CV	PCA-SVM	83.7% acc.	UCI
Kara and Okandan [25]	2007	24 of 72 normal - 28 of 52 AF signals test.	ANN BP	100% acc	MIT-BIH
Asl *et al.* [14]	2008	2/3tra. 1/3 test.	GDA-SVM	≈99% acc.	MIT-BIH
Oliveira *et al.* [15]	2010	75% tra. 25% test.	Bayesian Networks	≈99% sens.	MIT-BIH and QT database
Ozcan [16]	2010	10-fold CV	FSVM	85.71% acc.	UCI and Real ECG
Jadhav *et al.* [28]	2011	90% tra. 10% test.	ANN BP	86.67% sens.	UCI
Homaeinezhad *et al.* [17]	2012	4035 beats tra. 3150 beats test.	SVM-KNN-four MLP-BP (Neuro-SVM-KNN)	98.06% acc.	MIT-BIH

Tab. I *Previous work using various data.*

P-R interval includes the time required for atrial depolarization and the onset of ventricular depolarization. This interval is measured from the onset of the P wave to the beginning of the QRS complex. The normal value is in the range of 0.12-0.20 s. Longer P-R intervals are seen in the cases of AV block while shorter ones in different arrhythmias.

QRS complex indicates the total ventricular depolarization time which is measured from the onset of the Q wave to the offset of the S wave. The upper limit for normal is 0.10 s. Heart rate can be computed from two successive QRS complexes or R-R interval. Normally, heart rate ranges from 60 to 90 beats per minute.

S-T segment is the duration between the end of ventricular depolarization which is indicated by QRS complex and the beginning of the T wave.

T wave indicates the ventricular repolarization.

Q-T interval represents the duration of electrical systole and it is measured from the onset of the Q wave to the end of the T wave.

U wave is a deflection which follows the T wave preceding the next P wave and usually shows the same polarity as the T wave.

Arrhythmia is diagnosed based on these features. Interpretation of ECG features as a part of the analysis of ECG tracings is essential for arrhythmia recognition. Before detection of abnormal patterns in ECG plot, understanding the normal patterns is very important in terms of strengthening the ability of reasoning. A plot of normal ECG is given in Figure 1.

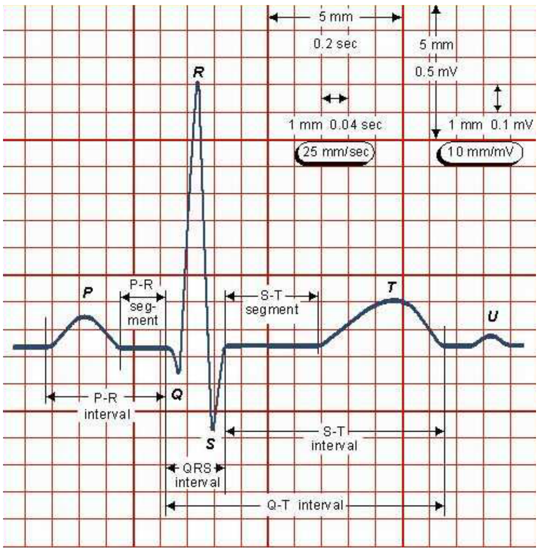

Fig. 1 *Waves on a normal ECG plot [3].*

The irregularity of the waves may cause an abnormality in heart beats. As an example of such abnormality is given in Figure 2 where R-R intervals are inconsistently irregular and the waves between each QRS complex are indistinct.

Fig. 2 *Waves on an abnormal ECG plot [4].*

There are some values belonging to featues of the waves in normal ECG samples. Any difference in such measurements may indicate an arrhythmic pattern. Additionally, any disruption in the order of waves and missing wave might cause abnormalities. In Figure 3 there are no P waves and QRS complex is wide which indicates a ventricular tachycardia.

Fig. 3 *Waves on an abnormal ECG plot [4].*

A preprocessing step is required to detect characteristic waves in researches using ECG signal data such as MIT-BIH database. Pan and Tompkins [18] developed an algorithm to detect QRS complexes. The optimized implementation of QRS detection in C language is presented in [19]. ECG filtering method is used to extract heartbeat and RR intervals in the study of Niwas *et al.* [10]. Asl *et al.* [14] applied QRS detection algorithm [18, 19]. Wavelet transform analysis is also widely used to detect P wave, QRS complex and T wave in ECG signal. Song *et al.* [12] utilized the wavelet transform based method in detection of the QRS complex which is proposed by Park *et al.* [21]. Kadambe *et al.* [20] present a dyadic wavelet transform based QRS detector which computes local peaks across two successive dyadic scales and determines the presence of the QRS complex. Homaeinezhad *et al.* [17] applied an ECG detection and delineation method implemented by Ghaffari *et al.* [22]. Lin *et al.* [23] used a Morlet wavelet decomposition to extract features. Benitez *et al.* [24] present Hilbert based transform with an average detection rate of 99%. Kara and Okandan [25] extracted features using wavelet coefficients and Welch method [26]. In the study by Oliviera *et al.* [15], features such as distance between two consecutive QRS complexes, QRS complex shape are provided by hidden Markov model based framework which is developed by Andreão *et al.* [27].

Various approaches that address arrhythmia detection problem can be seen in previous works. Neural network based methods are very common in this field as they are able to solve nonlinear classification problems. Niwas *et al.* [10] studied arrhythmia classification using a multilayer feedforward neural network trained with backpropagation algorithm. In this work, QRS duration, RR-Interval features which were extracted with the heartbeat detection by Pan and Tompkins [18] and spectral entropy that describe heart rate variability comprised the feature set. Including normal beat, 10 different beat types were classified with 99.02% overall accuracy using chosen feature sets. Multilayer perceptron (MLP) is an extensive model of neural networks consisting of one input layer, one output layer, and one or more hidden layers in between. Jadhav *et al.* studied a comparison study of multilayer perceptron (MLP) neural network model, generalized feedforward neural network and modular neural network model using UCI arrhythmia dataset and among three models, the study resulted in with outperforming of MLP ANN. More than one model of classifiers can be used together as one model is used alone. As an example of this, Hu *et al.* [5] used MLPs in a cascade relationship for beat classification of one normal and 12 abnormal classes.. First model of MLP is used for classifying as a normal or an abnormal beat and the second model is used to categorize abnormal beats into 1 to 12 abnormal classes. This composite MLP classifer is better than using one MLP classifer alone for multiclass problem. However, in such a composite usage of the neural networks, the overall performance of the results is principally relies on how well the first neural network is. Because the second neural network model is fed by the output of the first one. Silipo and Marchesi [6] concentrated on the unknown or ambiguous events in ECG analysis and presented a combination of ANN approach with the capability of uncertainty management. This management is provided by adding uncertainty criteria for the beats classified which does not exist in the training set or for the ectopic beats based on pre-defined thresholds. The ANN approach enhanced with the uncertainty management capability is very good at dealing with the ambiguous nature of ECG signal. But convergence to an adequate error value might be a problem as the dimension of the data is getting higher when arrhythmic classes in the training set are increased which may affect the reliability in a negative way. Another study involving the capability of uncertainty management is held by Gao *et al.* [9] and is based on a Bayesian ANN. In [9], a two-class arrhythmia classification problem is dealt with using a Bayesian ANN approach supported with a dual threshold method. The dual threshold method is applied by defining two thresholds as a boundary of a probability range in which uncertain cases with high-risk classification output are lying. This adaption can be used for the suppression of false alarms but this study is potentially more useful with given noiseless inputs.

Support Vector Machine (SVM) based approaches are also very common. The effect of the classifiers has been optimized when it is used in conjunction with dimension reduction techniques. Uyar *et al.* [13] studied arrhythmia classification using SVM with principal component analysis (PCA) and independent component analysis (ICA). According to experimens, SVM result was improved with PCA more than it is used with ICA. Song *et al.* [12] used linear discriminant analysis (LDA) for dimension reduction before applying SVM on data. In discriminating six types of arrhythmia beats, SVM associated with LDA has overperformed SVM with PCA. Besides the ECG signals, there are researches using heart rate variability (HRV) data. Asl *et al.* [14] classified six different types of arrhythmia cases based on HRV data. SVM is used as a classifier and the generalized discriminant analysis (GDA) is used for feature reduction. SVM-GDA technique showed a good performance in classification. However, some arrhythmia types such as left bundle branch block (LBBB) and righ bundle branch block (RBBB) could not be detected with the features extracted from the HRV signal. Along with using SVM in conjunction with a feature reduction technique, SVM is combined with various approaches in order to increase the performance. As an example of this usage, Özcan *et al.* [16] combined fuzzy approach with SVM to deal with the effects of outliers in classification.

There are also hybrid approaches such as a combination of SVM, k-NN and four different MLP models proposed by Homaeinezhad *et al.* [17]. A combined form of genetic algorithms and neural networks has increasingly been used in the literature. The common reason of this increase is the idea that two of them may provide an efficiency for solving more problems than either of them alone. The way to use this combination is varying according to the problem. In the study of Amma [30] and Jiang *et al.* [31], the optimization technique of genetic algorithm is used. Genetic algorithms might also be used for learning purpose. Ölmez *et al.* [32] utilized genetic algortihm to train neural network. Zhou and Li [33] studied premature ventricular contraction (PVC) detection in ECG signals using genetic algorithm trained perceptrons.

General classification schemes used in the literature for arrhythmia detection problem are implemented for comparison in this study. k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Naive Bayes and Bayesian Networks are those which are used as comparison. The idea behind selecting k-NN and SVM classifiers is analyzing the results of locally and globally search method for arrhythmia classification. Naive Bayes and Bayesian Networks are selected to study the results of probabilistic approach. Dimensionality reduction techniques are used such as recursive feature extractor (RFE-SVM), correlation based feature selection (CFS), principal component analysis (PCA) and factor analysis (FA). These classifiers are applied both original dataset and the reduced data sets to observe the effect of the dimension reduction techniques on the improvement of the classifiers. The flow diagram of the comparison is given in Figure 4. The performance of these techniques on original dataset and reduced data sets are compared with the proposed approach.

3. GA-NN Approach

In GA-NN approach, genetic algorithm has an optimizing role whilst neural network plays the role of classifier. Therefore, the type of the combination used in this study might be considered as *supportive combination* according to Schaffer *et al.* [34] because the way to use genetic algorithm is to assist neural network in feature selection. The main idea behind the GA-NN approach is utilizing global search ability of the GA in ECG feature space while performing arrhythmia classification with neural network. By doing so, it is intended to find which ECG features are more effective in determining arrhythmia patterns.

In our implementation, weights of neural networks are encoded in a chromosome. ECG features are represented as inputs to neural network and output nodes indicate class labels which are normal and abnormal. The encoding schema is given in Figure 5.

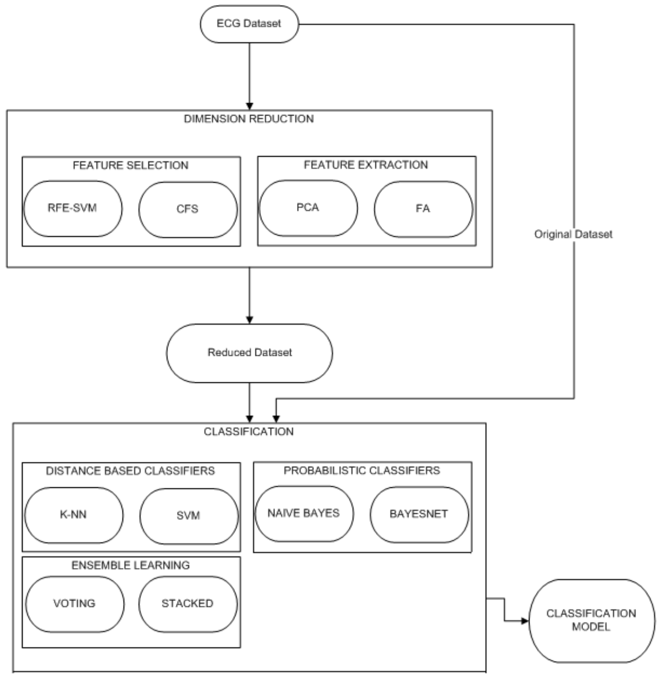

Fig. 4 *General classification schemes used in the literature for arrhythmia detection problem.*

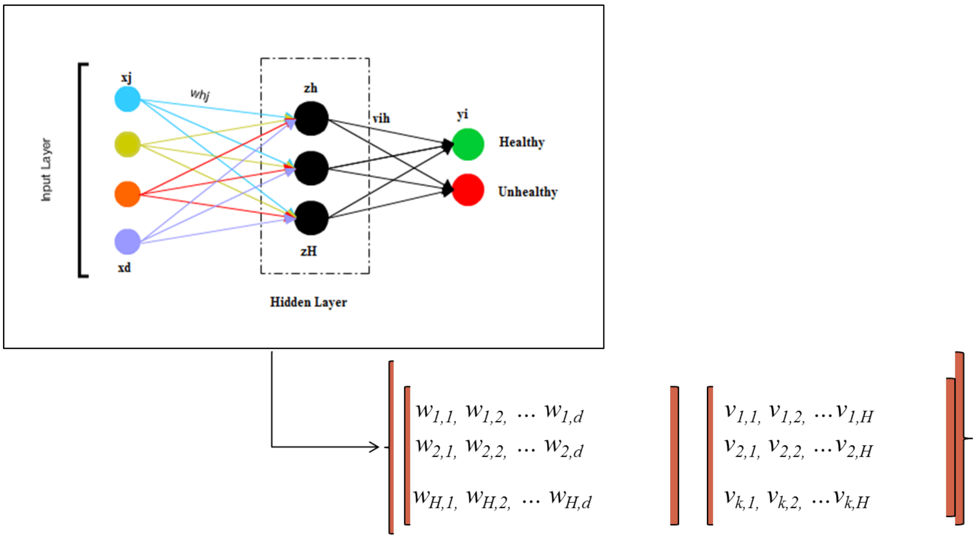

Fig. 5 *Encoding schema of a chromosome.*

Neural networks with the same topology in a population are considered as a candidate solution. Each candidate solution is assigned a fitness value which is an assessment that indicates how good a solution is. Fitness values are calculated based on the performance of the solutions. During the iterations, weights are optimized and fitness values of the solutions are computed. The possibility of the solutions to be in next generation is inversely proportional to their fitness values. The new generation is intended to be better than their predecessors. Parent solutions are selected using roulette wheel selection. Crossover is applied to produce offsprings. These offsprings are mutated so as to provide variation. Additionally, a second mutation operator is applied which prunes the network from irrelevant features by zeroing their weights and fully described by Sexton *et al.* [35]. The best offsprings produced as a result of these operations comprise of the 90% of the next generation. The remaining part is filled by the best solutions from old generation in order to carry best characteristics from ancestors. The solution with the best fitness value of all its generation is compared with the one from old generation. The iteration is stopped when the best solution of the current generation is not better than the best solution of the previous.

As a candidate solution, a neural network is evaluated according to its performance on validation data. The performance is generally related to the error which is the difference between the actual output and the predicted output. The fitness value [35] is obtained from the fitness function is as below:

$$Min\{f = \sum(O_i - \hat{O}_i)^2 + C\sqrt{\frac{\sum_{i=1}(O_i - \hat{O}_i)^2}{N}}\} \qquad (1)$$

where O is the target class, $\hat{O}$ is the estimated class of the instance i and N is the number of the instances. Here, C represents the number of nonzero weights in the network. Each network is assigned a probability based on its fitness value [35]. The probability is computed as in Equation 2

$$P(X = x) = \frac{f_{cur} - f_{bad}}{\sum_{i=1} f_i - f_{bad}} \tag{2}$$

where f_i indicates the fitness value of the solution i, f_{bad} is the worst fitness value which is greater than the rest of the fitness values in the population and f_{cur} is the fitness value of the current solution.

Backpropagation algorithm is used for setting initial weights to the neural network. Training parameters are predefined. All networks in a population are tested using the same parameters.

The aim is to detect presence or absence of arrhythmia accurately. In order to achieve this, selection of network parameters improving the accuracy of the classifier is important. So, using trial-and-error method, some promising parameter sets are specified and for each parameter set, neural networks are trained and tested so as to find optimum parameter set. Parameter selection is summarized in Figure 6. Both feature selection and classification are performed simultaneously using the neural network with selected parameter set.

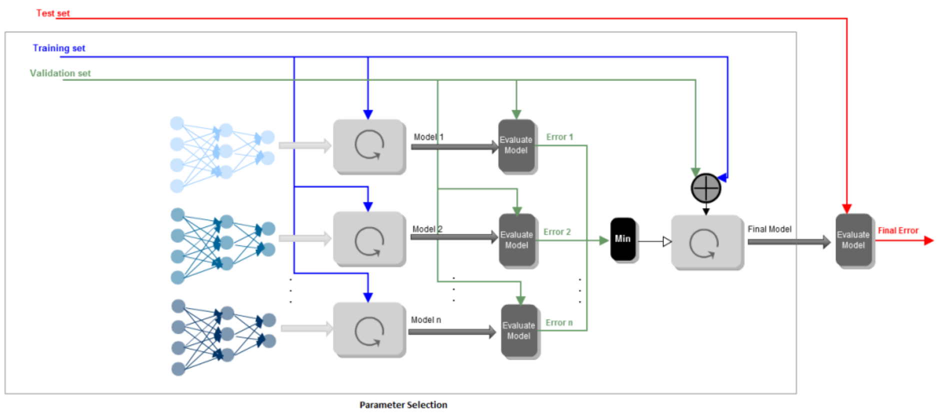

Fig. 6 *A diagram of parameter selection [36].*

4. Rule Extraction

The last stage is the rule extraction. Producing robust and reliable arrhythmia classification results to be applicable in diagnostic decision support systems and to aid cardiologists with medical assessments is the main interest. Rule extraction from classification output is important in terms of producing the results in human readable format. Applying solely rule extraction methods will increase the complexity in rules as fully described by Halford *et al.* [37]. So, genetic algorithm-neural network approach is used as a preclassifier and feature selector to produce rules concisely. Main flow of the system is summarized in Figure 7. C4.5, RIPPER, PART and HotSpot methods are used to perform rule extraction.

- *C4.5* is a decision tree algorithm using information entropy for deciding the best split. Rules are generated by adding the conditions on each path from each leaf node to the root node.
- In *RIPPER*, rule set is produced by repeatedly adding rules that contain features with the highest information gain. Post-prunning is performed for optimization.
- *Part* builds a partial C4.5 decision tree in each iteration and the leaf node having the maximum coverage is selected as the rule of that iteration. Next iterations are performed on the instances which are not covered by the previous rule.
- *HotSpot* learns association rules corresponding to the target class which is abnormal class in our study.

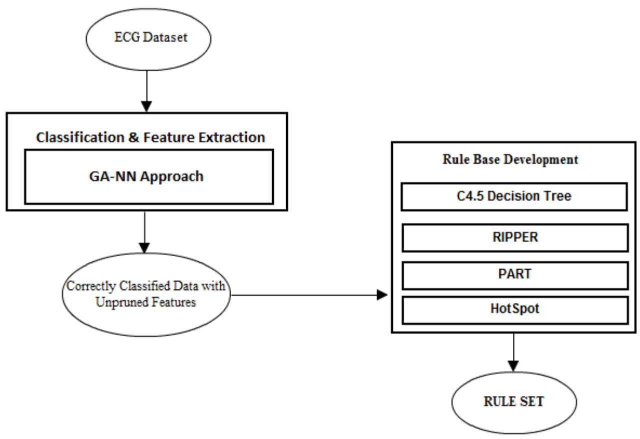

Fig. 7 *Main flow of the system.*

There are many researches focusing on rule generation for arrhythmia cases along with determinant features which are empirically known in literature. Based on this knowledge, features that comprise the rules are convenient. Additionally, output of each rule extraction method is tested on validation set and performance metrics such as accuracy, sensitivity, f-score, precision and MCC (Matthews correlation coefficient) are calculated in order to get insight how well these rules draw the characteristic differences between arrhythmia and normal cases. In this study, the significance of rule extraction for ECG analysis is to comply with experts' opinions so that it is intented to assist cardiologists in determininig the cases.

5. Experimental Details

Genetic Algorithm assisted Neural Network approach is implemented in Matlab [39]. In experiments, UCI Arrhythmia dataset is used for comparison and real ECG data is used for verification. A population of neural networks with the same parameter set is initially created. The structure of the neural networks is comprised of one input layer, one hidden layer and one output layer. Each of them is trained with scaled conjugate gradient backpropagation algorithm, namely 'trainscg'. Genetic algorithm steps are applied to the population iteratively. At each iteration, neural network weights are applied crossover, mutation and mutation2 operators. Parent selection is performed based on roulette wheel selection. Crossover rate is chosen as 90% and mutation rate is set as 10%. Fitness value of the offsprings is obtained from objective function and based on fitness value evaluation, probability of each solution to be in the next generation is computed. 90% of the new population is populated by the offsprings with the highest probability. The rest of the population is completed by the best solutions from the old generation in order to carry the best characteristics of the ancestors. Model selection is performed based on its performance. The resulting model is tested on validation set and verified on real ECG dataset.

5.1 Experimental Setup

The dataset used in this study is obtained from UCI Repository [1]. UCI Arrhythmia dataset originally contains 452 instances with 279 attributes. There are 16 arrhythmia cases associated with each instance. 15 of them indicate anomalies and one of them is normal rhythm. 206 attributes are numeric and the rest is nominal.The first four features are about personal information such as age, sex, heigth and weight. The rest presents different measurements from the ECG paper.

About 0.33% of the feature values in the dataset are missing. Class distribution of this dataset is very unfair and instances of classes 11, 12 and 13 do not exist in the dataset while instances of class 01 which indicates normal rhythm is frequent. So, alll anomalies are grouped into one class as abnormal rhythm and the resulting dataset contains inputs of two groups: Normal and Abnormal. Missing values among the features are removed. This leads a reduction to 278 features and 420 instances.

The proposed approach aims to identify determinant features taking into account of weights of the features. The irrelevant ones are eliminated by zeroing the relevant weights. Data is normalized in range of -1 and +1 in order to provide efficiency.

For comparison purpose, selected classifier techniques are applied to reduced data sets besides the original dataset. 5, 7, 10, 13, 20 and 30-feature datasets are obtained by applying RFE-SVM, CFS and PCA techniques to original dataset. 3, 5, 7 and 8-feature data sets are provided by applying FA technique.

The second dataset is a set of real ECG data. This data is obtained from Kardiosis ECG Tool of the manufacturing firm TEPA [38]. This dataset is smaller than UCI Arrhythmia dataset. There are 20 records, 13 of them are normal and the rest is abnormal. Only relevant features which are obtained as a result of the GA-NN approach are computed from these records because this dataset is intended to verify the proposed method.

5.2 GA-NN Classification

Genetic algorithm is used to assist neural network for performing feature selection and classification simultaneously. According to the experiments, 86.75% of accuracy is obtained with the resulting model. There is also approximately 95% decrease in the number of features. The original input feature number is 278 while the resulting input feature is 12. k-NN, SVM, Naive Bayes and BayesNet methods are applied to UCI Arrhythmia dataset for comparison purpose. In these experiments 4 techniques are used for dimension reduction: Recursive Feature Elimination with Support Vector Machine (RFE-SVM), Correlation Based Feature Selection (CFS), Principal Component Analysis (PCA) and Factor Analysis (FA). 10 fold cross validation is used in classification experiments. The classifiers with their best performance in terms of sensitivity are summarized in Tab. II.

k-NN$_5$ + RFE-SVM$_{13}$	**SVM+ CFS$_{20}$**	**Naive Bayes$_5$ + Original**	**BayesNet+ FA$_8$**	**Proposed+Original**
0.7179	0.3571	0.7158	0.7128	0.9646

Tab. II *Classifier+Dataset sensitivity comparison.*

According to these results, GA-NN approach provides the highest hit rate for the target class. Naive Bayes and Bayesian methods draw similar performance. As a proof of the efficiency of feature selection, GA-NN outperforms Naive Bayes for original dataset.

The highest f-score results of all classifiers are given in Tab. III.

k-NN$_5$ + RFE-SVM$_{13}$	**SVM+ CFS$_{20}$**	**Naive Bayes$_5$ + FA$_8$**	**BayesNet+ CFS$_{20}$**	**Proposed+Original**
0.7044	0.3933	0.7778	0.7904	0.8916

Tab. III *Classifier+Dataset f-score comparison.*

GA-NN approach comes the first scoring the highest value. BayesNet comes the second with 0.7904 f-score value.

The highest MCC and accuracy results of the classifiers are given in Table 5.2 and in Tab. IV respectively. In terms of taking both posivite and negative class into account, GA-NN provides the highest result. Probabilistic approaches perform better than k-NN and SVM. Due to the value that is obtained is close to zero, SVM can be said to predict a little better than random prediction.

k-NN$_5$ + CFS$_{20}$	SVM+ RFE-SVM$_{13}$	Naive Bayes$_5$ + FA$_8$	BayesNet+ CFS$_{20}$	Proposed+Original
0.5247	0.0432	0.6209	0.6701	0.7375

Tab. IV *Classifier+Dataset MCC comparison.*

k-NN$_5$ + CFS$_{20}$	SVM+ RFE-SVM$_{13}$	Naive Bayes$_5$ + FA$_8$	BayesNet+ CFS$_{20}$	Proposed+Original
0.7857	0.5333	0.8095	0.8333	0.8675

Tab. V *Classifier+Dataset accuracy comparison.*

When we consider the ratio of correctly classified instances, GA-NN comes the first achieving 86.75% accuracy. As overall, it can be said that GA-NN approach outperforms others for all performance metrics. Probabilistic approaches draw a consistent picture.

GA-NN approach is tested on real ECG dataset obtained from TEPA [38]. An accuracy of 85% rate is obtained as a result.

5.3 Comparison with State of the Art

The classifiers used for arrhythmia classification problem in the literature shows diversity. For comparison purpose, some of the classifiers which are frequently used in this problem are implemented. The proposed approach presents both feature reduction and classification. So, to be fair in comparison and feel the effect of dimension reduction to the classification, we obtained reduced data sets in arbitrary dimensions using four different techniques. RFE-SVM, CFS, PCA and FA are the methods used for dimension reduction purpose. After applying RFE-SVM, CFS and PCA, we obtained 5, 7, 10, 13, 20 and 30-featured data sets and using FA we provided 3, 5, 7 and 8-featured data sets. k-NN, SVM, Naive Bayes and BayesNet are experimented on the reduced resulting data sets and the original dataset. For k-NN classifier, 6 different k values are chosen which are 5, 7, 9, 13 and 15. K-NN is executed for each k value. Furthermore, stacked and voted combinations of k-NN and SVM are tried on data sets which they perform the best in order to experiment the effect of combination to the performance. Every result of each classifier on each data set is evaluated based on the performance metrics which are accuracy, sensitivity, f-score, precision and MCC.

According to experiments, proposed approach outperforms others in terms of the performance metrics. The results show that dimension reduction has an optimized effect on classification. An increasing trend in sensitivity is seen among classifiers when reduced data sets are used. In addition to this, it is observed that performance of k-NN and SVM has increased when these techniques are combined as voting and stacked. Furthermore, it can be said that probabilistic techniques draw consistent performance in classification results.

5.4 Rule Extraction

Based on selected features, rule set is extracted using C4.5, RIPPER, PART and HotSpot algorithms. The performance measures including accuracy, sensitivity, f-score, precision and MCC are computed for the rule extraction methods. The comparison of the rule extraction methods based on the performance measures is given in Tab. VI.

Method	Accuracy	Sensitivity	F-score	Precision	MCC
C4.5	87.8%	81%	85.2%	89.8%	75%
RIPPER	88.5%	85.6 %	86.6%	87.6%	76.5%
PART	**94.2**	87.9%	93%	98.7%	88.6%

Tab. VI *Performance results of rule sets.*

Among these measures, accuracy indicates the ratio of correctly classified data to whole data while others focus on target class. The target class in this study refers to abnormal class. According to the measures, PART povides highest measures for all metrics. C4.5 and RIPPER have similarities but RIPPER outperforms C4.5 especially in terms of sensitivity which is related to its ability to identify target class. The resulting rules produced by PART are given in Tab. VII.

1	If V1_Avg_QRSA > 1 then CLASS = ABNORMAL
2	If AVL_Amplitude_T_wave ≤ -0.8 and V1_N_intrinsic_deflections ≤ 8 then CLASS = ABNORMAL
3	If heartrate ≤ 57 and Tinterval ≤ 165 then CLASS = ABNORMAL
4	If Tinterval > 221 and DI_Avg_QRSTA ≤ 25.5 then CLASS = ABNORMAL
5	If heartrate > 94 and Tinterval > 148 then CLASS = ABNORMAL
6	If V3_Avg_width_S_wave ≤ 28 and V3_Amplitude_S_wave > -6.8 and DI_Avg_QRSTA > 17.7 then CLASS = ABNORMAL
7	If V1_N_intrinsic_deflections > 24 and V3_Avg_width_S_wave > 40 then CLASS = ABNORMAL
8	If QRSduration > 107 and V1_Avg_QRSA ≤ -25 then CLASS = ABNORMAL
9	If V2_Avg_width_S_wave > 44 and V1_N_intrinsic_deflections > 4 and heartrate > 64 and V3_Avg_width_S_wave ≤ 56 then CLASS = ABNORMAL

Tab. VII *PART rule set.*

The resulting rule set consists of 9 rules. QRSA of V1 Lead, amplitude of S and T waves, heartrate, duration of T wave and QRS complex, width of S wave, QRSTA of D1 Lead and number of intrinsic deflections are selected features. The rules are not in a successive order because Part produces rules from the repeated generated partial decision trees.

6. Discussion

In this paper, we proposed GA-NN approach for ECG feature selection and rule-based arrhythmia classification. The type of GA-NN combination might be considered as *supportive combination* according to Schaffer *et al.* [34] because the way to use genetic algorithm is to assist neural network in feature selection.

Arrhythmia detection problem has been a noteworthy research topic during many years. Most of the works studied in the literature are related to improving automated ECG analysis. Thereby ECG interpretation, and thus arrhythmia diagnosis are aimed to be performed in a more robust and reliable way.

Neural network is a popular technique in arrhythmia classification. Applying solely a classifier to the original data may reduce efficiency in terms of performance metrics. According to experimental results, it is seen that using classifier with a feature reduction technique is more efficient. In this proposed approach, we use GA-NN combination in order to overcome this problem. Feature reduction and classification are performed simultaneously. By doing so, features which mostly contribute to the classification are determined. Neural network is the main classifier whilst genetic algorithm has an optimizing role assisting for feature reduction. We utilize global search ability of the genetic algorithm in weight space so as to avoid local minima. Features are evaluated according to their weights which are changed during the genetic algorithm steps. At the end of this iterative process, the aim is to find the most efficient and pruned neural network classfier.

In this study, it can be said that robustness and reliability are provided and the resuts might be helpful for cardiologists while evaluating arrhythmia cases. Genetic algorithm is a supportive technique in terms of improving the ability of the neural network classifier. This proposed method in the sense of acceleration may need to be improved, because the iterative approach can take time to find the best artificial neural network classifier.

7. Conclusions

The heart has a unique place among other muscles in terms of being capable of automatic rhythmic contraction. The impulses that stimulate muscular contraction arise in the conduction system of the heart and thereby, the blood circulation takes place in a specific order. Any disruption in the order of this activity may cause a cardiac arrhythmia which is worth considering in terms of being a potential reason for a heart disease. Therefore, computer-aided ECG analysis is very important for early diagnosis of heart diseases. ECG tracing is just a laboratory result which needs to be supported with a clinical assessment to be meaningful and useful. At this point, ECG interpretation by an expert is indispensable. This process is very important and requires full attention. Automated ECG analysis is supportive during this interpretation process.

In this study, GA-NN approach is proposed for feature selection and rule based arrhythmia classification. We also performed a comparison amongst the proposed approach and the other most common techniques. The proposed approach outperformed others according to experimental results. Moreover, a real ECG data is used for verifying the proposed approach and we obtain high accuracy rate. As a result, based on selected features, rule extraction algorithms are performed to interpret classification results in human readable format. Accuracy and interpretability are taken into account while doing this study. For future work, we may accelerate the performance of the approach as an optimization.

The significance of data obtained from ECG cannot be undervalued. However, this is also important that interpretation of ECG in conjunction with the clinical assessments will be more effective in arrhythmia diagnosis than deciding only on the basis of ECG data.

Our contibutions in this paper are to perform rule-based arrhythmia classification and producing a result from rules in order to aid cardiologists during ECG interpretation. This study enables the performing ECG feature selection and the producing concise rules using the determinant features.

References

[1] UCI Machine Learning Repository, http://archive.ics.uci.edu/ml, 2010

[2] MIT-BIH Arrhythmia Database, http://www.physionet.org/physiobank/database/html/mitdbdir/mitdbdir.htm, 1997

[3] Normal ECG Image, http://www.swharden.com/blog/images/ecg2.jpg, 2008

[4] Becker E. B.: Fundamentals of Electrocardiography Interpretation, Anesthesia Progress, 53, 2, 2006, pp. 53-64

[5] Hu Y. H., Tompkins W. J., Urrusti J. L., Afonso V. X.: Applications of Artificial Neural Networks for ECG Signal Detection and Classification, Journal of Electrocardiology, 26, 1993, pp. 66-73

[6] Silipo R., Marchesi C.: Artificial Neural Networks for Automatic ECG Analysis, IEEE Transactions on Signal Processing, 46, 5, 1998, pp. 1417-1425

[7] de Chazal P., Reilly R.B.: A Comparison of the ECG Classification Performance of Different Feature Sets, Computers in Cardiology 2000, 2000, pp. 327 -330

[8] Frank E.: An Accurate, Clinically Practical System for Spatial Vectorcardiography, Circulation, 13, 5, 1956, pp. 737-749

[9] Gao D., Madden M., Chambers D., Lyons G.: Bayesian ANN Classifier for ECG Arrhythmia Diagnostic System: a Comparison Study, Proceedings of International Joint Conference on Neural Networks, 2005

[10] Niwas S. I., Kumari R. S. S., Sadasivam V.: Artificial Neural Network Based Automatic Cardiac Abnormalities Classification, Proceedings of the Sixth International Conference on Computational Intelligence and Multimedia Applications, 2005, pp. 41-46

[11] Zhang H., Zhang L.: ECG Analysis Based on PCA and Support Vector Machines, International Conference on Neural Networks and Brain, 2005. ICNN&B '05., 2, 2005, pp. 743-747

[12] Song M. H., Lee J., Cho S. P., Lee K. J., Yoo S. K.: Support Vector Machine Based Arrhythmia Classification Using Reduced Features, International Journal of Control, Automation and Systems, 3, 4, 2005, pp. 571-579

[13] Uyar A.: Intelligent Arrhythmia Classification Based on Support Vector Machines, Boğaziçi University, 2006

[14] Asl B. M., Setarehdan S. K., Mohebbi M.: Support Vector Machine-Based Arrhythmia Classification Using Reduced Features of Heart Rate Variability Signal, Artificial Intelligence in Medicine, 44, 1, 2008, pp. 51-64

[15] Oliveira L. S. C., Andreão R. V., Sarcinelli-Filho M.: The Use of Bayesian Networks for Heart Beat Classification, Brain Inspired Cognitive Systems 2008, 657, 2010, pp. 217-231

[16] Özcan N. Ö., Gürgen F.: Fuzzy Support Vector Machines for ECG Arrhythmia Detection, Proceedings of the 2010 20th International Conference on Pattern Recognition, 4, 2010, pp. 2973-2976

[17] Homaeinezhad M. R., Atyabi S. A., Tavakkoli E., Toosi H. N., Ghaffari A., Ebrahimpour R.: ECG Arrhythmia Recognition via a Neuro-SVM-KNN Hybrid Classifier with Virtual QRS Image-Based Geometrical Features, Expert Systems with Applications, 39, 2, 2012, pp. 2047-2058

[18] Pan J., Tompkins W. J.: A Real-Time QRS Detection Algorithm, IEEE Transactions on Biomedical Engineering, 3, 1985, pp. 230-236

[19] Hamilton P. S., Tompkins W. J.: Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database, IEEE Transactions on Biomedical Engineering, 12, 1986, pp. 1157-1165

[20] Kadambe S., Murray R., Boudreaux-Bartels G. F.: Wavelet Transform-Based QRS Complex Detector, IEEE Transactions on Biomedical Engineering, 46, 7, 1999, pp. 838-848

[21] Park K., Lee K., Yoon H.: Application of a Wavelet Adaptive Filter to Minimise Distortion of the ST-Segment, Medical and Biological Engineering and Computing, 36, 5, 1998, pp. 581-586

[22] Ghaffari A., Homaeinezhad M., Khazraee M., Daevaeiha M.: Segmentation of Holter ECG Waves via Analysis of a Discrete Wavelet-Derived Multiple SkewnessKurtosis Based Metric, Annals of Biomedical Engineering, 38, 4, 2010, pp. 1497-1510

[23] Lin C. H., Du Y. C., Chen T.: Adaptive Wavelet Network for Multiple Cardiac Arrhythmias Recognition, Expert Systems with Applications, 34, 4, 2008, pp. 2601-2611

[24] Benitez D., Gaydecki P. A., Zaidi A., Fitzpatrick A. P.: The Use of the Hilbert Transform in ECG Signal Analysis, Computers in Biology and Medicine, 31, 5, 2001, pp. 399-406

[25] Kara S., Okandan M.: Atrial Fibrillation Classification with Artificial Neural Networks, Pattern Recognition, 40, 11, 2007, pp. 2967-2973

[26] Welch, P.: The Use of Fast Fourier Transform for the Estimation of Power Spectra: a Method Based on Time Averaging over Short, Modified Periodograms, IEEE Transactions on Audio and Electroacoustics, 15, 2, 1967, pp. 70-73

[27] Andreão R.V., Dorizzi B., Boudy J.: ECG Signal Analysis through Hidden Markov Models, IEEE Transactions on Biomedical Engineering, 53, 8, 2006, pp. 1541-1549

[28] Jadhav S.M., Nalbalwar S. L., Ghatol A. A.: Artificial Neural Network Based Cardiac Arrhythmia Disease Diagnosis, Process Automation, Control and Computing (PACC), 2011 International Conference on, 2011, pp. 1-6

[29] Yeap T. H., Johnson F., Rachniowski M.: ECG Beat Classification by a Neural Network, Engineering in Medicine and Biology Society, 1990., Proceedings of the Twelfth Annual International Conference of the IEEE, 1990, pp. 1457-1458

[30] Amma N. G. B.: Cardiovascular Disease Prediction System Using Genetic Algorithm and Neural Network, 2012 International Conference on Computing, Communication and Applications (ICCCA), 2012, pp. 1-5

[31] Jiang W., Kong S. G., Peterson G. D.: ECG Signal Classification Using Block-Based Neural Networks, 2005 IEEE International Joint Conference on Neural Networks (IJCNN '05), 2005, pp. 326-331

[32] Ölmez T.: Classification of ECG Waveforms by Using RCE Neural Network and Genetic Algorithms, Electronics Letters, 33, 18, 1997, pp. 1561-1562

[33] Zhou J., Li L.: Using Genetic Algorithm Trained Perceptrons with Adaptive Structure for the Detection of Premature Ventricular Contraction, Computers in Cardiology 2004, 2004, pp. 353-356

[34] Schaffer J.D., Whitley D., Eshelman L. J.: Combinations of Genetic Algorithms and Neural Networks: a Survey of the State of the Art, 1992 International Workshop on Combinations of Genetic Algorithms and Neural Networks (COGANN-92), 1992, pp. 1-37

[35] Sexton R. S., Dorsey R. E., Sikander N. A.: Simultaneous Optimization of Neural Network Function and Architecture Algorithm, Decision Support Systems, 36, 3, 2004, pp. 283-296

[36] Gutierrez-Osuna R.: Intelligent Sensor Systems, Lecture Topic: Validation, http://courses.cs.tamu.edu/rgutier/ceg499_s02/l13.pdf, 2002

[37] Halford G. S., Wilson W. H., Phillips S.: Processing Capacity Defined by Relational Complexity: Implications for Comparative, Developmental, and Cognitive Psychology, Behavioral and Brain Sciences, 21, 6, 1998, pp. 803-831

[38] TEPA Inc http://www.tepa.com.tr/, 2007

[39] Mathworks, http://www.mathworks.com, 2011