Diagnosis of the Heart Rhythm Disorders by Using Hybrid Classifiers

Sule Yucelbas, Gulay Tezel, Cuneyt Yucelbas, Seral Ozsen

Abstract—In this study, it was tried to identify some heart rhythm disorders by electrocardiography (ECG) data that is taken from MIT-BIH arrhythmia database by subtracting the required features, presenting to artificial neural networks (ANN), artificial immune systems (AIS), artificial neural network based on artificial immune system (AIS-ANN) and particle swarm optimization based artificial neural network (PSO-NN) classifier systems. The main purpose of this study is to evaluate the performance of hybrid AIS-ANN and PSO-ANN classifiers with regard to the ANN and AIS. For this purpose, the normal sinus rhythm (NSR), atrial premature contraction (APC), sinus arrhythmia (SA), ventricular trigeminy (VTI), ventricular tachycardia (VTK) and atrial fibrillation (AF) data for each of the RR intervals were found. Then these data in the form of pairs (NSR-APC, NSR-SA, NSR-VTI, NSR-VTK and NSR-AF) is created by combining discrete wavelet transform which is applied to each of these two groups of data and two different data sets with 9 and 27 features were obtained from each of them after data reduction. Afterwards, the data randomly was firstly mixed within themselves, and then 4-fold cross validation method was applied to create the training and testing data. The training and testing accuracy rates and training time are compared with each other.

As a result, performances of the hybrid classification systems, AIS-ANN and PSO-ANN were seen to be close to the performance of the ANN system. Also, the results of the hybrid systems were much better than AIS, too. However, ANN had much shorter period of training time than other systems. In terms of training times, ANN was followed by PSO-ANN, AIS-ANN and AIS systems respectively. Also, the features that extracted from the data affected the classification results significantly.

Keywords—AIS, ANN, ECG, hybrid classifiers, PSO.

I. INTRODUCTION

NEURAL networks (NNs) [1]-[3] are strong mathematical models inspired by the human brain. Particle Swarm Optimization (PSO) which is one of the heuristic methods is successfully applied to the training of ANN due to various reasons such as the small number of parameters to be set, easy realization and capable of treatment with real numbers [4]. AIS are algorithms that developed based on the theoretically and observed immune functions which applied to complex problem domains.

According to [5], a PSO-ANN system has been proposed for the classification of EEG signals. In that study, 5 and 10 cells for hidden layer have been used and been observed a better classification performance with neural network having a number of 5 cells. In [6], a radial basis function neural networks learned by PSO has been proposed for detecting of abnormalities in the ECG beats which were taken from MIT-BIH arrhythmia database. Also, K-means, Kohonen and K-nearest neighbor algorithms have been used for comparison and the proposed method was found to be much faster than others.

When looking at the studies with AIS-ANN, in [7], a new AIS based on SOM (self organization map) neural networks has been developed. Analyzes were used to estimate the silicon content and was added that the results are fascinating. In addition to these, in [8], AIS-ANN hybrid system was formed to extract rules from classification problems. 96.4% and 96.8 accuracy rates have been achieved for heart diseases and hepatitis, respectively. Also, results were compared with other approaches until then.

II. MATERIALS AND METHODS

In this study, the performance analyzes of used systems was performed on total of 6 class data taken from MIT-BIH ECG arrhythmia database. They are normal sinus rhythm (NSR), atrial premature contraction (APC), sinus arrhythmia (SA), ventricular tachycardia (VTK), ventricular trigeminy (VTI) and atrial fibrillation (AF). Normal sinus rhythm was used to create double groups with other data set. In total, 5 different data sets were formed with this method. The resulting 5 data sets are NSR-APC, NSR-SA, NSR-VTI, NSR-VTK and NSR-AF. Training and testing data separation process was performed using 4-fold cross-validation method. The following Table I shows the features of data used in the study.

The classification processes for all systems were repeated for five times. The accuracy of classification results obtained in each run were collected and divided into 5 to achieve an average classification accuracy. Practiced process steps are shown in Fig. 1.


### TABLE I

<table>
<thead>
<tr>
<th>Arrhythmia name</th>
<th>Registration No</th>
<th>Acronyms</th>
<th>Data is received from the time interval (min)</th>
<th>The number of samples taken in the time interval</th>
<th>The number of R-R interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Sinus Rhythm</td>
<td>103</td>
<td>NSR</td>
<td>1.09-17.21</td>
<td>349921</td>
<td>100</td>
</tr>
<tr>
<td>Atrial Premature Contraction</td>
<td>101</td>
<td>APC</td>
<td>9.54-24.31</td>
<td>315721</td>
<td>60</td>
</tr>
<tr>
<td>Sinus Arrhythmia</td>
<td>113</td>
<td>SA</td>
<td>12.27-22.10</td>
<td>209881</td>
<td>60</td>
</tr>
<tr>
<td>Ventricular Tachycardia</td>
<td>200</td>
<td>VTK</td>
<td>1.41-5.38</td>
<td>85321</td>
<td>60</td>
</tr>
<tr>
<td>Ventricular Trigeminy</td>
<td>119</td>
<td>VTI</td>
<td>2.38-4.51</td>
<td>47881</td>
<td>60</td>
</tr>
<tr>
<td>Atrial Fibrillation</td>
<td>202</td>
<td>AF</td>
<td>29.35-30.06</td>
<td>10801</td>
<td>39</td>
</tr>
</tbody>
</table>

---

1) **Discrete Wavelet Transform (DWT)**

Wavelet Transform (WT), attempts to provide the optimal resolution in terms of time and frequency because it uses small window for high frequencies and large size window for low frequencies [10], [11]. The following equation of continuous wavelet transform is given [10].

\[
W(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t) \psi \left( \frac{t-b}{a} \right) dt
\]  

(1)

where the parameters: \(x(t)\) is the signal itself, \(W(a, b)\) is wavelet transform, \(\psi\) is mother wavelet function, \(a\) is scale and \(b\) is scrolling at time. Mother wavelet \(\psi\), is obtained by \(a\) and \(b\) scales [12], [13].

\[
\psi_{a,b}(t) = |a|^{-1/2} \psi \left( \frac{t-b}{a} \right)
\]  

(2)

Resolution is done suitable both of the frequency and time by mother wavelet \(\psi\) [14]. When a scale has a great value, the mother wavelet expands and the scale is selected a small value while obtaining low frequency details, mother wavelet narrows and high frequency details are obtained. There are forms of continuous or discrete wavelet transform. However, for this study, Continuous wavelet transform will produce coefficient more than necessary to identify the signal and will cause a lot of difficulties in the calculation because all of these parameters are necessary for restructuring of the signal. Therefore, Discrete Wavelet Transformation (DWT) which limits of transformation and change of scale provides a great convenience for the method to use. Scale and shift parameters have been bringing into disjoint with \(x\) and \(y\) for Continuous Wavelet transform and Discrete Wavelet transform (DWT) basic expression can be expressed as follows:

\[
W(j, k) = \sum_{j} \sum_{k} x(n) 2^{-j/2} \psi(2^{-j/2} n - k)
\]  

(3)

where \(x\) (n) is signal itself and \(\psi\) is mother wavelet. With multi-frequency filter banks which are balanced and logarithmic structure, WT analysis can be performed [13]. Detail coefficients are obtained with high pass filter and approximate coefficients are obtained with low pass filter for sign with DWT. These operations are repeated until a target resolution is reached at the obtained detail or approximate frequency bands. As can be seen in Fig. 1, discrete \(x[n]\) signal has been passed through high-pass filter (HPF) for detail coefficients (D\(i[n]\)) and has been passed through low-pass

---

### A. QRS Detection

In order to operate on the received signals, the signal must be separated by significant parts primarily for diagnostic block. As is known, electrocardiographic signals are signals of a certain period. However, this period can vary from person to person, such as time, even for the same record for the same person may vary. For this reason, QRS detection algorithms of the existing literature are implemented and algorithm to detect the R peaks with minimum error is chosen. An RR interval represents a period for the system. RR interval of the received signal, part of the system to be used for diagnosis, is equal to the number of pattern. Divided by the RR intervals from the ECG signals, data can be obtained for block the signal patterns of diagnostic. This data is a test data for diagnostic block. Firstly, using these techniques on the data, samples of unnecessary influence in the classification of the signal is discarded and the signal is compressed [9].

### B. Feature Extraction

DWT was used to reduce the number of features. In this study, 3 and 5 degree of Daubechies-2 (db-2) was selected to obtained approximation coefficients (cA) at the output of low-pass filter. In this way, 200 features were reduced to 9 (cA.9) and 27 (cA.27) by using fifth and third order approximation coefficients, respectively.

---

![Fig. 1 Process Steps](image-url)
filter (LPF) for approximate coefficients \((A_i[n])\). \(g[n]\) and \(h[n]\), mathematical expression of this separation process in time, respectively, high-pass and low-pass filters is expressed as follows:

\[
h_{i+1}(k) = \sum_n h_i[n]h[2k - n] \quad (4)
\]

\[
g_{i+1}(k) = \sum_n h_i[n]g[2k - n] \quad (5)
\]

At each level of decomposition, half band filters provides to occur signals with half of the frequency band.

---

**C. Artificial Neural Networks (ANN)**

The structure of the nervous system has been fundamental to the development of Artificial Neural Networks (ANN). An artificial neural network occurs from cells are connected with each other in layers and weights connecting them to each other. One of the basic elements of an artificial neural network coefficients are between connections which adjust proportionally the exchanging of data with each other of nerve cells. During training of ANN, these coefficients are adjusted the appropriate values by training algorithms and ANN is made original for related problem. The data transmission between the input and output of the cell, which is developed for the problem, is performed by cell activation function. This function determines the basic structure of the cell and the data in the cell takes shape according to this function [15].

In the implementation phase, the number of neurons of the hidden layer (HLNN) and the learning rate (LR) were incorporated in as variable parameters of ANN system. In the literature, because of achieving the best classification accuracy when the momentum constant \((M_c)\) was received 0.8 or 0.9, it was kept constant at 0.8 in this study [9], [16]. Finally, the number of iterations was taken as 10,000 and all procedures were performed according to these parameters.

ANN structure used in this study was composed of 3 layers which were an input layer, a hidden layer and an output layer. 20, 30, 40 and 50 values were tested for HLNN. Also, 0.4, 0.8, 1.2 and 1.6 values were tested for LR. The activation function used for all layers was Logarithmic sigmoid. In addition, the target stop criterion was taken 0.001.

**D. Particle Swarm Optimization (PSO)**

An information sharing mechanism is used by Particle swarm optimization algorithm. A group of \(n\) particles fly in a particular speed in the D-dimensional search space. Each particle in the search process takes notice of its own search history and the best point within the group of other particles and location varies based on this. Both of particle position and velocity vary according to the following equations in the typical PSO method:

\[
v_{j}(k+1) = \omega v_{j}(k) + c_1 r_1(\max - x_{j}(k)) + c_2 r_2(\max - x_{j}(k)) \quad (6)
\]

\[
x_{j}(k+1) = x_{j}(k) + v_{j}(k+1) \quad (7)
\]

\(X_j\) is the position vector of the \(j\)-th particle, \(V_j\) is the velocity vector. Consider \(P\) the best position of \(j\)-th particle during its search process, and \(P_g\) as the whole particle swarm’s best position during the current search. \(\omega\) is the inertia coefficient and the integration of the inertia coefficient allows for settings of coefficients that provide convergence and control the intended stability between exploitative and explorative approach. \(c_1\) and \(c_2\) are called learning factor, which makes particles have the function of self-summary and learn to the best of the swarm, and get close to the best position of its own as well as within the swarm. \(Rand()\) is the random number distributed in \((0, 1)\). Each particle’s velocity is limited to between the maximum and minimum speed interval, \(V_{\text{max}}\) and \(V_{\text{min}}\) [17].

---

**E. Neural Network Learned by Particle Swarm Optimization (PSO-NN)**

In this study, back propagation algorithm of neural network was learned by PSO algorithm for PSO-NN hybrid model. The dimension size was taken 3 and the inertial coefficient firstly was taken 0.9 but it was multiplied by 0.975 for each loop and...
the result was appointed as the new value. Learning factors (c1 and c2) were taken as 2. Finally Vmax and Vmin were taken as between -1 and 1. While PSO part was created, particle number (PN) and number of hidden layer neurons (HLNN) were incorporated in as variable parameters of PSO-NN system. 10, 30, 50 and 100 values were tested for PN. Also, 3, 4, 5 and 6 values were tested for HLNN and all procedures were performed according to these parameters [18].

Steps in the process of training for PSO-NN are provided in Fig. 3. Getting started with the program, weights that contain numeric expression of the connections between layers are randomly assigned as the number of particles initially for learning of Neural Network. Training process is started. Fitness values are calculated and weights are reloaded. After this stage, typical PSO operations begin. When the target fitness value for the gbest is reached the minimum error value, training process is finished and testing process is started [18].

F. Artificial Immune Systems (AIS)

In this study, AIS system based on clonal selection algorithm was used. Clonal selection algorithm operates according to five basic principles:

1. Firstly, a population (P) is generated and this population occurs from candidate solutions, (the memory cells (M) and remaining population (Pr)) (P=Pr+M).
2. At the second step, the best n elements in the population are selected depending on the sensitivity criteria and Pn population is generated [19].
3. The best individuals selected for depending on the sensitivity to the antigen are cloned (proliferation). Linearly, a high number of clones are comprised from individuals with high sensitivity. Similarly, a low number of clones are comprised from individuals with low sensitivity. Thus, a clone set C is formed.
4. A hyper-mutation operation is performed to population C consisting of clones. Direct proportion with sensitivity is also in the process of hyper-mutation. The set obtained after hyper-mutation is called population C* [19].
5. After cloning and hyper-mutation operations, memory set M is again created by making a selection process to add individuals to the population. After the selection process, location of some cells in the population P is expected to leave for some cells in set of C*. The number individuals (d) of population P are replaced by the newly produced individuals to provide diversity in the population. Displacement of low-sensitivity individuals are more likely [19].

In the implementation phase, the stopping criterion (SC) and the number of antibodies (M) were incorporated in as variable parameters of AIS system. The value of SC was changed between 0.92 and 0.98 in the interval of 0.02. Also, the value of M was selected as 30, 50, 70 and 100 and all procedures were performed according to these parameters [18].

G. Neural Network Learned by Artificial Immune Systems (AIS-NN)

In this study, back propagation algorithm of neural network was learned by AIS algorithm for AIS-NN hybrid model. In the implementation phase, the target minimum error (TME) and the number of antibodies (M) were incorporated in as variable parameters of AIS-NN system. 0.15, 0.09, 0.05 and 0.02 values were chosen for TME. Also, 10, 30, 50 and 100 values were used for M [18].

Steps in the process of training for AIS-NN are provided in Fig. 4. Getting started with the program, weights that contain numeric expression of the connections between the layers are randomly assigned as the number of antibodies (M) initially for learning of Neural Network. Training process is started. Average errors are calculated and weights are reloaded. After this stage, typical AIS operations begin. When the TME value for the antibodies is reached to minimum error, training process is finished and testing process is started [18].

H. Training and Testing Errors Calculation Techniques

Training algorithm, created for AIS in the system, gives memory antibodies, generated as an output after the training phase, and their class information. At step of training and testing phase, these outputs are used to identify both training and testing classes of antigens in classification.
Each antigen class was determined as $s_{\text{test}}$ with both the training and test classification processes and classification accuracy of the algorithm was calculated as in (8) and (9) [20].

\[
\text{classification accuracy} = \sum_{i=1}^{N_t} \frac{\text{correct } s_{i}(i)}{N_t}
\]

(8)

\[
\text{correct } s_{i}(i) = \begin{cases} 1 & \text{if } s_{\text{test}}(i) = s(i) \\ 0 & \text{otherwise} \end{cases}
\]

(9)

In here:
- $s_{\text{test}}(i)$: Training or test antigen class determined by the algorithm,
- $N_t$: Total data (antigen) number in training or test set,
- $s(i)$: The real class of training or test antigen.

The performances of other systems except AIS were calculated using 4 different statistical criteria which were rounding expressed as ROUND, mean squared error (MSE), mean absolute error (MAE) and root mean square error (RMSE).

\[
\text{MSE} = \text{error} = \frac{1}{p} \sum_{p=1}^{P} (y_p - y_{dp})^2
\]

(10)

\[
\text{MAE} = \text{error} = \frac{1}{p} \sum_{p=1}^{P} |y_p - y_{dp}|
\]

(11)

\[
\text{RMSE} = \text{error} = \sqrt{\frac{1}{p} \sum_{p=1}^{P} (y_p - y_{dp})^2}
\]

(12)

III. EXPERIMENTAL RESULTS

In this study, 4 different classifiers (ANN, AIS, PSO-NN and AIS-NN) were used to classify ECG signals taken from MIT-BIH ECG database. For each system, some of the parameters specific to that system were applied as variables.

In the process, 4 different evaluation methods were used to find classification error (MAE, ROUND, MSE and RMSE) and forms of the errors with subtracting from 100 were shown as accuracy rates (MA-A, R-A, MS-A and RMS-A, respectively). Also, average results obtained each fold were shown in all tables because 4-fold cross-validation was used to increase robustness. Table II expresses to classification results of ANN system for different learning rates (LR) and number of neurons in the hidden layer (HLNN) for all data sets (both cA_9 and cA_27). Analyzing Table II, it is seen that the highest classification accuracy rates were found to be 100% for R-A accuracy. Secondly best results were obtained for MS-A. Furthermore, the best accuracy values of all evaluation criteria for both training and testing data sets are seen for NSR-APC data set. Also, spent time and the number iteration for this data set are generally less than other data sets. In addition, it is observed that the best results for all data sets were obtained for 30 HLNN and 0.8 LR in general for ANN.

Table III illustrates classification results of all data sets (both cA_9 and cA_27) for selected values of the stopping criteria (SC) and number of antibodies (M) in AIS system.

Table III shows that the best results for both training and testing data sets were obtained for NSR-VTK in general. But the least number of memory antibodies (MANN) is reached with NSR-APC data set. It is important that this parameter is lower to identify the achievement of algorithm for AIS systems. In addition, the best results were seen generally for 0.98 SC and 50 M for AIS.

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>OBTAINED CUMULATIVE TRAINING AND TESTING RESULTS FOR ALL DATA SETS IN ANN SYSTEM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NSR-APC</strong></td>
<td><strong>NSR-SA</strong></td>
</tr>
<tr>
<td>MA-A</td>
<td></td>
</tr>
<tr>
<td>cA_9</td>
<td>98.71</td>
</tr>
<tr>
<td>cA_27</td>
<td>98.46</td>
</tr>
<tr>
<td>R-A</td>
<td></td>
</tr>
<tr>
<td>cA_9</td>
<td>100</td>
</tr>
<tr>
<td>cA_27</td>
<td>100</td>
</tr>
<tr>
<td>MS-A</td>
<td></td>
</tr>
<tr>
<td>cA_9</td>
<td>99.91</td>
</tr>
<tr>
<td>cA_27</td>
<td>99.9</td>
</tr>
<tr>
<td>RMS-A</td>
<td></td>
</tr>
<tr>
<td>cA_9</td>
<td>98.15</td>
</tr>
<tr>
<td>cA_27</td>
<td>98.19</td>
</tr>
<tr>
<td><strong>TIME</strong></td>
<td><strong>TRAINING</strong></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>MA-A</td>
<td></td>
</tr>
<tr>
<td>cA_9</td>
<td>1.8</td>
</tr>
<tr>
<td>cA_27</td>
<td>1.44</td>
</tr>
<tr>
<td>R-A</td>
<td></td>
</tr>
<tr>
<td>cA_9</td>
<td>94.25</td>
</tr>
<tr>
<td>cA_27</td>
<td>103.25</td>
</tr>
<tr>
<td><strong>HLLNN</strong></td>
<td><strong>TRAINING</strong></td>
</tr>
<tr>
<td>BOTH</td>
<td>30</td>
</tr>
<tr>
<td><strong>LR</strong></td>
<td><strong>TRAINING</strong></td>
</tr>
<tr>
<td>BOTH</td>
<td>0.8</td>
</tr>
</tbody>
</table>
TABLE III
OBTAINED CUMULATIVE TRAINING AND TESTING RESULTS FOR ALL DATA SETS IN AIS SYSTEM

<table>
<thead>
<tr>
<th></th>
<th>NSR-APC</th>
<th>NSR-SA</th>
<th>NSR-VTI</th>
<th>NSR-VTK</th>
<th>NSR-AF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCURACY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RATE</td>
<td>cA_9</td>
<td>96.43</td>
<td>97.89</td>
<td>96.31</td>
<td>96.04</td>
</tr>
<tr>
<td></td>
<td>cA_27</td>
<td>97.34</td>
<td>97.47</td>
<td>97.07</td>
<td>97.53</td>
</tr>
<tr>
<td>TIME</td>
<td>cA_9</td>
<td>613.2</td>
<td>855.26</td>
<td>2059.47</td>
<td>748.48</td>
</tr>
<tr>
<td></td>
<td>cA_27</td>
<td>521.89</td>
<td>184.49</td>
<td>591.90</td>
<td>332.75</td>
</tr>
<tr>
<td>MAN</td>
<td>cA_9</td>
<td>6.75</td>
<td>16.75</td>
<td>13.25</td>
<td>16.75</td>
</tr>
<tr>
<td></td>
<td>cA_27</td>
<td>10.5</td>
<td>10.5</td>
<td>11.0</td>
<td>12.75</td>
</tr>
<tr>
<td>ITERATION</td>
<td>cA_9</td>
<td>15962.25</td>
<td>619092.5</td>
<td>82292.5</td>
<td>514.75</td>
</tr>
<tr>
<td></td>
<td>cA_27</td>
<td>390.75</td>
<td>246210.75</td>
<td>751.75</td>
<td></td>
</tr>
<tr>
<td>SC BOTH</td>
<td></td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M BOTH</td>
<td></td>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table IV expresses the results of different target minimum error (TME) and number of antibodies (M) for all data sets (both cA_9 and cA_27) in AIS-NN system.

The highest classification accuracy rates were found for R-A and MS-A respectively in Table IV. Furthermore, NSR-APC with cA_9 and NSR-AF with cA_27 are the most successful data sets for AIS-NN. However, NSR-SA is seen that better than the others in terms of duration and iteration. In addition, it is observed that the best results can be obtained for 0.02 TME and 10 M in general for AIS-NN.

Table V consists of classification results different number of neurons in the hidden layer (HLNN) and number of particles (PN) for all data sets (both cA_9 and cA_27) in PSO-NN system.

In Table V, the highest classification accuracy rates were found for R-A and MS-A as in other tables, respectively. The highest achievement was reached with NSR-SA and NSR-VTI for training data sets. Besides that, NSR-APC was the best data set for testing. Also, NSR-AF is better than the others in terms of duration and iteration in Table V. In addition, it is observed that the best results can be obtained for 3 HLNN and 50 PN in general for PSO-NN.

As a result, AIS-ANN and PSO-ANN hybrid classification systems were as successful as ANN system. But the best achievement was obtained with ANN for all data sets. Also, the results of the hybrid systems were much better than AIS, too. Furthermore, the data set which the best results were mostly obtained for all systems is NSR-APC in terms of test process.

However, ANN had much shorter period of training time than other systems. In terms of training times, ANN was followed by PSO-ANN, AIS-ANN and AIS systems respectively. Also, the features that extracted from the data affected the classification results significantly.
ACKNOWLEDGMENT

This study is supported with 13101011 project number by the Scientific Research Projects of Selcuk University.

REFERENCES