Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2022) 30: 1788 – 1803

doi:10.55730/1300-0632.3905

Development of a hybrid system based on ABC algorithm for selection of appropriate parameters for disease diagnosis from ECG signals

Ersin ERSOY, Erkan BOSTANCI, Mehmet Serdar GÜZEL*
Department of Computer Engineering, Faculty of Engineering, Ankara University, Ankara, Turkey

Received: 25.10.2021 • **Accepted/Published Online:** 02.06.2022 • **Final Version:** 22.07.2022

Abstract: The number of people who die due to cardiovascular diseases is quite high. In our study, ECG (electrocardiogram) signals were divided into segments and waves based on temporal boundaries. Signal similarity methods such as convolution, correlation, covariance, signal peak to noise ratio (PNRS), structural similarity index (SSIM), one of the basic statistical parameters, arithmetic mean and entropy were applied to each of these sections. In addition, a square error-based new approach was applied and the difference of the signs from the mean sign was taken and used as a feature vector. The obtained feature vectors are used in the artificial bee colony algorithm; whose fitness function is a multilayer perceptron neural network. It has been shown that the hybrid system we have developed achieves high classification accuracy by selecting the most appropriate parameters for the detection of arrhythmia from ECG signals.

Key words: ECG, artificial bee colony, artificial neural network, machine learning, optimization, feature selection, arrhythmia

1. Introduction

Arrhythmia (rhythm disorder) is a type of disease that shows abnormal heartbeats. Because arrhythmias can increase or decrease blood pressure, they can lead to paralysis, stroke and even death. Nonperiodic segments and waves in the ECG signal, or the onset and end times being longer or shorter than certain values are signs of arrhythmia. In ECG measurements, such arrhythmias are observed as waveform deformations or irregularities [1].

Artificial neural network (ANN) models are frequently used in different ways for arrhythmia diagnosis and classification processes from ECG signals [2]. ANN is formed by connecting more than one cell called neuron. ANN models are an artificial intelligence approach that can detect the relationships between the inputs and outputs of the samples given to them, learn these relationships, and generalize over the results obtained. The multilayer perceptron neural network (MPNN) model is a widely used model due to its ease of use and high success rates in classification of signals. Artificial bee colony (ABC), a population-based method, is a method proposed by Karaboğa for the optimization of numerical functions [3]. This heuristic algorithm, inspired by the intelligent foraging behavior model of bees, determines the most appropriate value sought according to the fitness function of the problem. Many hybrid ANN-based approaches have been used for the detection of arrhythmias from ECG signals: An improved ABC has been used for ECG analysis and heartbeat type classification. In this algorithm, unlike the classical ABC, improvements have been made in the search approaches of the worker

*Correspondence: mguzel@ankara.edu.tr

bees and the search process has been started from a certain location instead of the random processes in the algorithm and it has been observed that it gives better results [4, 5]. In a study conducted in 2021, especially for the classification of stressful ECG signs, mental health was tried to be evaluated by using ECG signs. In the study, the multicore SVM structure was further developed using genetic algorithm (GA), ABC and improved particle swarm optimization (PSO). As a result, it has been revealed that the proposed study gives better one [6]. In a study conducted in 2021, especially for the classification of stressful ECG signs, mental health was tried to be evaluated by using ECG signs. In the study, the multicore SVM structure was further developed using GA, ABC, and PSO. As a result, it was revealed that the proposed study gave better results than the existing studies [7].

In this study, ECG signals were divided into segments and waves using temporal boundaries, and the feature vector of each segment from the signal similarity methods convolution, correlation, covariance, signal peak to noise ratio (PNRS), structural similarity index (SSIM), arithmetic mean which is one of the basic statistical parameters variance and entropy were applied. In addition, the square error, which is a measure of prediction quality, was applied directly on the signals, unlike its conventional applications in the literature. We have also employed the squared error in a novel way than the commonly used form as the difference between the estimated and expected values. In our study, it was obtained by taking the difference of the mean sign obtained from the signs of healthy individuals and the current sign processed. In other words, the value that should be included in the definition of square error was used as the mean sign value, and the estimated value was used as healthy ECG and ECG with arrhythmia. The ABC algorithm has selected the most suitable feature vectors among the processed feature vectors to maximize the overall accuracy in arrhythmia diagnosis. In the developed hybrid system, ABC algorithm used MPNN as a fitness function and sent the overall accuracy rate obtained after processing the selected feature vectors in the ANN to the ABC algorithm as a result.

2. Material and method

2.1. Dataset

Atrial fibrillation database includes 25 long-term ECG recordings of human subjects. The individual recordings are each 10 h in duration. The original analog recordings were made at Boston's Beth Israel Hospital using ambulatory ECG recorders with a typical recording bandwidth of approximately 0.1 Hz to 40 Hz. Normal sinus rhythm database includes 18 long-term ECG recordings of subjects referred to the Arrhythmia Laboratory at Boston's Beth Israel Hospital. Subjects included in this database were found to have had no significant arrhythmias; they include 5 men, aged 26 to 45 and 13 women, aged 20 to 50.

2.2. ECG sign and features

The "Physionet ECG databases" database was used as an ECG signal. "MIT-BIH Normal Sinus Rhytm (NS) Database"[1] was used for healthy ECG sign. "MIT-BIH Atrial Fibrillation (AF) Database"[2] was used for arrhythmia sign. Figure 1 shows the P, Q, R, S, T peaks of the waves observed in the normal ECG signature and the PR, QRS, ST, QT intervals.

Intervals in ECG signals have some temporal characteristics which can be described as follows:

P wave: Normally, the width of the P wave is less than 0.11 s, regardless of the lead type [8, 9]. PR interval:

[1]Pyshio.net NS database website https://archive.physionet.org/physiobank/database/nsrdb/

[2]Pyshio.net AF database website https://archive.physionet.org/physiobank/database/afdb/

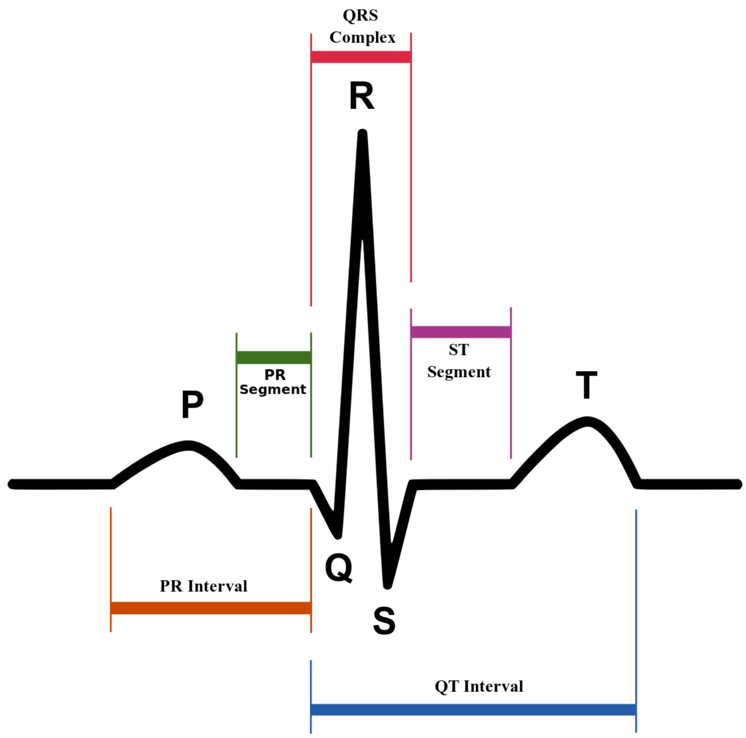


Figure 1. Intervals for ECG sign.

The PR interval is defined by the time between the onset of the P wave and the onset of the QRS complex. In adults, the normal value for the PR interval is 0.12–0.20 s [10].

QRS complex: Q wave is the first negative wave after P wave, R wave is the first positive wave, and S wave is the negative wave after R. Differently, it is less than 0.04 s and does not exceed 25 percent of the total QRS time. QRS complex has a maximum duration of 0.11 s [11].

ST segment: The ST segment is the interval between the J junction where the QRS complex ends and the beginning of the T wave. Its duration varies inversely with heart rate and is between 0 and 0.15 s [11].

T wave: Reflects repolarization of the ventricles. The duration of the normal T wave in adults is 0.10–0.25 s [12].

QT interval: The QT interval is a maximum of 0.44 s. The time from the beginning of the Q wave to the end of the T wave in the QRS complex is the QT interval [13].

PT interval: It is the interval from the beginning of the p point to the end of the t point. It is a cycle of blood coming to and pumping from the heart.

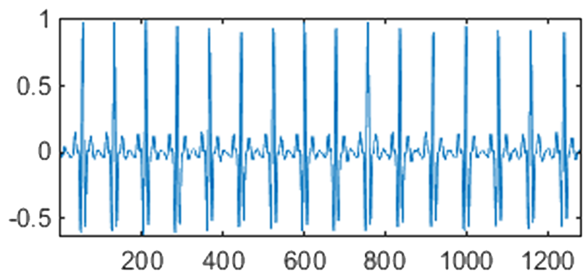
2.3. Feature extraction based on the calculation of temporal intervals from ECG signals

The feature vectors of the segments and waves in the ECG signal were calculated using the temporal distance to the R point. Pan-Tompkins algorithm was used to detect the R point in the ECG mark and the ECG mark P, PR, QRS, QT, ST, PT, and T were obtained.

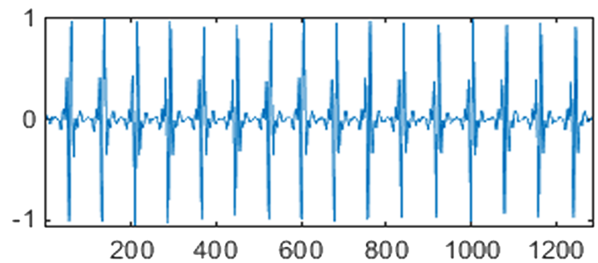
The Pan-Tompkins algorithm consists of several stages [14]. The first step is to apply a bandpass filter to filter out the noise in the ECG signals. Figure 2a shows the 10-second portion of the signal attenuated by the band-pass filter. The bandpass filter used in the Pan-Tompkins algorithm is used. A band-pass filter is obtained by using a low and high-pass filter. For the high-pass filter, the sampling frequency is 200 Hz, the cutoff frequency is 11 Hz, the offset amount is 5 samples, that is, 25 ms. The cut-off frequency of the high-pass filter is 200 Hz, the sampling frequency is 5 Hz, the shift amount is 16 samples, that is, 80 ms [15].

As depicted in Figure 2b, in the derivation phase, the filtered ECG signal was applied to the derivative receiver in order to make the QRS more pronounced and suppress the low-frequency components, and an ECG signal free of low-frequency components was obtained. Finally, smoothing is performed with the squaring and sliding window integration in Figure 2c [2].

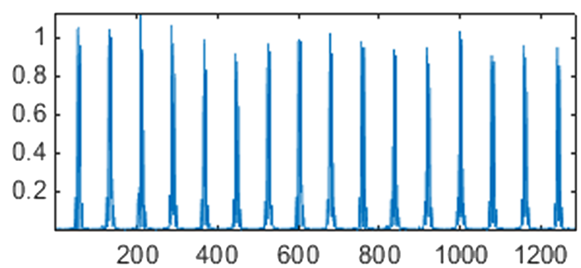
In Figure 3, QRS points detected on the ECG mark filtered by the Pan-Tompkins algorithm are shown.



(a) Band-pass filter applied.



(b) Derivative filter applied.



(c) Squared ECG signal.

Figure 2. Pan-Tompkins algorithms filters

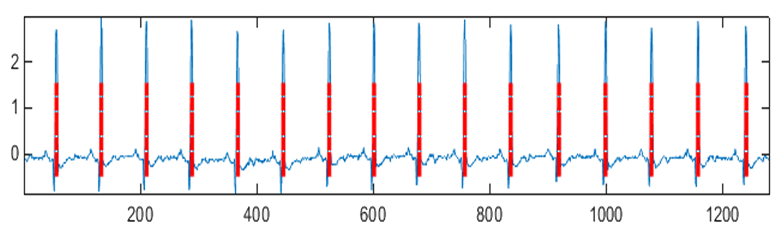


Figure 3. QRS points shown on the normal ECG signal.

After determining the R point of the ECG signal, the specified duration of the segments and waves and their distance to the r point are calculated and separated. After this separation, feature vectors are obtained by applying the following operations:

Convolution: In its simplest form, it is the process of passing a signal through another signal. f(t) and g(t) represent functions of the signal, in other words, it is to find out how a signal applied to a system will be output (Equation 1).

$$\begin{aligned} y(t) &= f(t) * g(t) \\ y(t) &= \int_{-\infty}^{\infty} f(T).g(t-T)\,dT \end{aligned} \tag{1}$$

Correlation: The correlation coefficient (r) is the measure of the relationship between two variables and varies between −1 and +1. In signal processing, crosscorrelation is characterized as a function of displacement of one with respect to the other, as f and g are continuous functions (Equation 2) [17].

$$(f \star g)(T) = \int_{-\infty}^{\infty} \overline{f(t)} g(t+T)\,dt \tag{2}$$

Covariance: The event that two or more variables change with two or more groups is called covariance or covariate. X, Y represent the independent variables and E represents the expected value (Equation 3) [18].

$$\begin{aligned} K_{XY}(t_1,t_2) &= conv(X_{t1},Y_{t2}) \\ conv(X_{t1},Y_{t2}) &= E[(X(t_1)-\mu_X(t_1))(Y(t_2)-\mu_Y(t_2))] \end{aligned} \tag{3}$$

Structural similarity index (SSIM): SSIM is generally used to measure the similarity between two images. It is based on the comparison of 3 features called brightness, contrast and structure of two images. Here, x, y compared images μ_x, μ_y, σ_x, σ_y and σ_{xy} indicate the pixel density mean, standard deviation and common variance of the estimated image (Equation 4).

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{4}$$

Signal peak to noise ratio (PNRS): PSNR is an engineering term that indicates the ratio of the maximum possible power of a signal to the power of the noise on the signal. Here, MAX_I is the maximum possible pixel value of the image, and MSE is the mean square error (Equation 5) [19].

$$PSNR = 20 \times \log_1 0(MAX_I) - 10 \times \log_1 0(MSE) \tag{5}$$

Arithmetic mean: It is the result obtained by dividing the sum of the items in the sample by the number of items (Equality 6).

$$\widehat{x} = \frac{1}{n}(\sum_{i=1}^{n}) = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} \tag{6}$$

Variance: It is the sum of the differences of all data in the distribution with the arithmetic mean divided by the number of data (Equation 7).

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n} \tag{7}$$

Entropy: It is a unit of measure that expresses the disorder of the elements in the data set with real numbers (Equation 8) [20].

$$H(X) = -\sum_{i=1}^{n} P(x_i)\log P(x_i) \tag{8}$$

Fourier transform: It is the name given to the mathematical operation to express a waveform as a combination of sine and cosine functions. Both transformations map an object between two orthogonal spaces. An object in x space is defined in ω space by Fourier transform (FT). Differential equations in x space are expressed as linear equations in ω space. After finding the solution of this equation in ω space, its equivalent in x space is obtained by inverse transformation. With the inverse transformation, the equivalent of the solution in the x space is obtained (Equation 9) [21].

$$F(x) = \left\{ \frac{1}{2\pi} \int_0^\infty 2F(\omega) \mathrm{e}^{i\omega t}\, d\omega \right\} \tag{9}$$

Mean square error-based approach: The approach we propose is similar to the quadratic calculation method. However, the calculation of the difference is obtained from the difference of the processed sign with the mean sign. The mean sign is the mean sign obtained in the healthy ECG sign n values on a vector, Y represents the observed value of the predicted variable, $\widehat{Y}$ represents the predicted value (Equation 10).

$$\begin{gathered} MSE = \frac{1}{n} \sum_{j=0}^{n-1} [Y(i) - \widehat{Y}(i)]^2 \\ RMSE(x) = \sqrt{MSE(x)} = \sqrt{E(\widehat{x} - x)^2} \end{gathered} \tag{10}$$

While obtaining the feature vectors, they are used as combinations of each other in order to expand the search space for selection space of the optimization algorithm. The procedures applied only for the P wave are listed in Table 1:

By repeating these operations for each segment and wave, a total of 105 + 1 (class) were obtained.

Table 1. The procedures applied only for the P wave.

p__fft	P wave is processed with fast Fourier transform (FFT). Its absolute value and squared are taken to save complexity.
p__fft__sum	The total value of the p__fft value to use in the classifier.
p__fft__mean	The mean value of the p__fft value for use in the classifier.
p__var	It is the variance value of the processed P wave.
p__fft__var	To use in the classifier p__fft value is obtained by applying the variance with the formula in Equation 2.
p__xcorr__max	Crosscorrelation of the processed P wave with the Port shown in Equation 3 is applied and its maximum value is taken.
p__conv__max	It is the maximum value of the processed P wave convolution with Port.
p__conv__sum	It is the total value of the processed P wave convolution operation with Port.
p__xcov__max	It is the maximum value of the processed P wave crosscovariance with the Port.
p__xcov__sum	It is the total value of the processed P wave crosscovariance operation with the Port.
p__rmse	Calculated as the square root of the difference of the processed P wave with the Port.
p__rmse__fft	Calculated by the square of the difference of the FFT states of the processed P wave and the Port wave.
p__peaksnr	PSNR of the processed P wave and Port wave
p__ssimval	It is the structural similarity index of the processed P wave and the Port wave.
p__mean	The arithmetic mean of the processed P wave.
p__entropy	It is the “Shannon” entropy value of the processed P wave.

3. Hybrid system developed with ABC And ANN

3.1. Artificial bee colony

ABC algorithm is an algorithm developed by Karaboğa inspired by the intelligent search of food by bees and is based on population logic [3]. ABC algorithm working system is as follows [22–25]:

- The number of food sources and the number of worker and onlooker bees are the same. An attendant bee is sent to each food source.
- The class of the bee of the depleted source changes and becomes a scout bee and looks for a new food source.
- The locations of the nutrient sources represent possible solutions to the optimization problem. The amount of nutrients in the sources also corresponds to the quality of the solution. The solution quality is also the fitness value in the algorithm.
- In ABC algorithm, the aim is to find the position of the source with the highest amount of nutrients in the search space and to determine the values that give the maximum or minimum solution to the problem.

3.2. Multilayer perceptron neural network

An MPNN model is structurally composed of three consecutive layers: the input layer, the hidden layer, and the output layer. The number of neurons in the input layer is the number of features to be trained. The number of neurons in the output layer also represents the desired properties. The number of hidden layers can be changed according to the problem and increase the capability of the network. The hidden layer is the classifier. It transmits the values from the input layer to the output layer as a result. Increasing the number of neurons in the hidden layer increases the capability of the network, but may adversely affect the uptime [26–28].

3.3. Developed hybrid system

In the hybrid system proposed in our study, the ABC performs the feature vector selection process. The ABC sends the feature vectors it chooses to the fitness function, that is, to the MPNN. MPNN works with selected feature vectors and the final correct classification success of the model was evaluated with the help of statistical criteria. The most basic criteria for this assessment are specificity, sensitivity, and overall classification accuracy. The overall classification accuracy is sent to the ABC algorithm as a result of the fitness function. The ABC algorithm works to maximize this accuracy value. The hybrid system we developed is schematized in Figure 4.
Training parameters:
ABC algorithm training parameters:

- Number of colonies NC = 20
- Number of food sources NC/2 = 10
- The number of iterations that will be abandoned as a result of not improving the trials is 100.
- Stop criterion = 2500 iterations
- Number of parameters to optimize = 7 (P, PR, PT, QRS, QT, ST, T)

MPNN parameters:

- Number of hidden layers= 10
- Train percentage = %70
- Test Percentage = %15
- Validation = %15
- Number of input feature vector = 7
- Number of outputs = 1

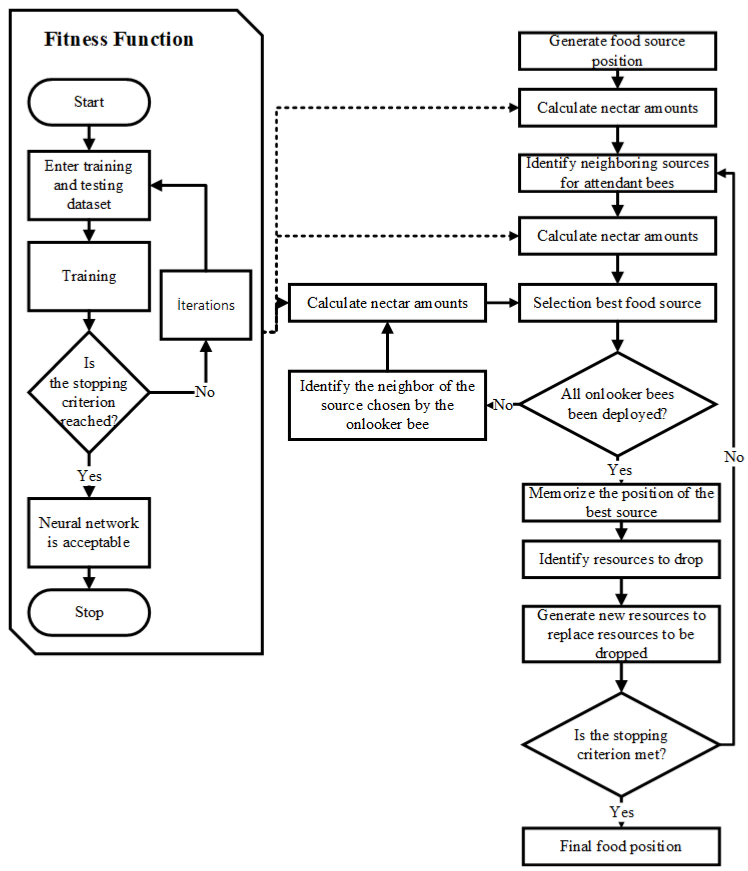


Figure 4. Neural network based optimization system flowchart.

4. Experimental findings and discussion

The data set we obtained consists of 128,682 rows and 105 + 1 (class). As a result of the proposed method, the optimization algorithm made the following choices:

- The crosscorrelation maximum value of the P wave,
- The crosscovariance maximum value of the PR interval,
- The crosscorrelation maximum value of the PR interval,
- Fourier transform sum of QRS wave,
- The square error value of the fourier transform signals of the QRS wave,
- Variance value of QRS wave,
- The crosscorrelation maximum value of the T wave.

As seen in Table 2, the confusion matrix contains information about the actual and predicted classifications made by a classification system. It can also be defined as the ratio of true positives to false positives.

Table 2. Confusion matrix tables.

	Positive prediction	**Negative prediction**
Positive class	True positive (**TP**)	False negative (**FN**)
Negative class	False positive (**FP**)	True negative (**TN**)

The criteria generally used in the evaluation are accuracy, precision and recall. There are shown in Equations 11–13.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{11}$$

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

$$Recall = \frac{TP}{TP + FN} \tag{13}$$

The ABC algorithm is the best nutrient-oriented algorithm based on random processes. For this reason, the hybrid system was run 20 times and the most selected methods were used. When the methods selected by the optimization are tested in the ANN, the results of the system results are shown in Table 3. The ROC curve obtained as a result of testing the developed model is shown in Figure 5. Accuracy was found to be 99.98%, precision 99.97% and recall 99.98%

Table 3. Confusion matrix using selected areas.

Class	True positive	False negative
True positive	64325	16
False negative	8	64333

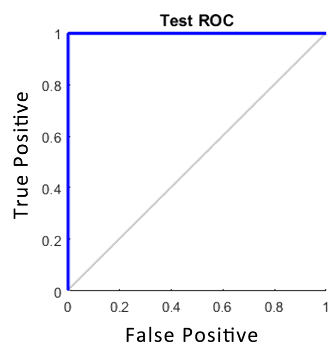


Figure 5. ROC curve of the developed model.

To verify, the methods chosen by the optimization were applied separately to all intervals and segments of the ECG signal, and datasets were obtained to compare the optimization performance. For example, if the optimization algorithm chose the crosscorrelation method, only crosscorrelation was applied to all ranges and segments and the data obtained was tested for classifier success in the neural network. This study has been applied for all the methods selected by the system we propose. The overall accuracy values obtained are as follows:

Only crosscorrelation (xcorr max) process maximum was applied to P, PR, PT, QRS, QT, ST, T intervals and segments. The obtained data set was used in the neural network and the classifier success was observed. The test results are as in Table 4 and the ROC curve is given in Figure 6. Accuracy was found to be 96.74%.

Neural network test results obtained by P, PR, PT, QRS, QT, ST, T using FFT are as in Table 5 and ROC curve is given in Figure 7. Accuracy was found to be 85.11%.

The neural network test results of the square error (fft_rmse) of the values obtained by the P, PR, PT, QRS, QT, ST, T using FFT are as in Table 6 and the ROC curve is given in Figure 8. Accuracy value is 92.6%.

Table 4. Confusion matrix of the trained network with the data obtained by applying crosscorrelation.

Class	**True positive**	**False negative**
True positive	64089	252
False negative	1034	63307

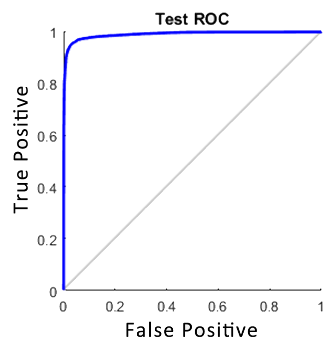


Figure 6. ROC curve obtained using crosscorrelation maximum values of all segments.

Table 5. Confusion matrix of the trained network with the data obtained by applying FFT.

Class	**True positive**	**False negative**
True positive	64039	302
False negative	18849	45492

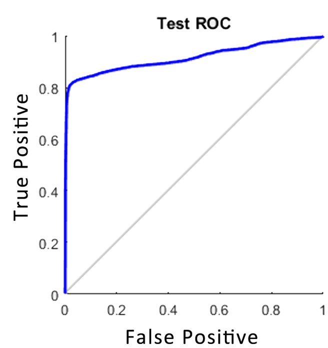


Figure 7. ROC curve obtained using the values of FFT of the signals.

Neural network test results of P, PR, PT, QRS, QT, ST, T using variance (var) values are as in Table 7 and ROC curve is given in Figure 9. Accuracy was found to be 84.52%.

The neural network test results of the data obtained by applying the P, PR, PT, QRS, QT, ST, T crosscovariance process are as in Table 8 and the ROC curve is given in Figure 10. Accuracy was found to be 99%.

Table 6. Confusion matrix of the applying FFT and square error to all intervals and segments.

Class	**True positive**	**False negative**
True positive	61792	2549
False negative	6935	57406

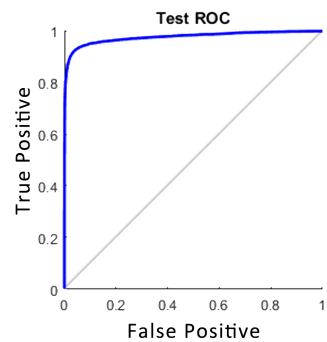


Figure 8. ROC curve obtained using the values of the square error and FFT of the signals obtained by the P, PR, PT, QRS, QT, ST,T.

Table 7. Confusion matrix of the trained network with the data obtained by applying variance to all intervals and segments.

Class	**True positive**	**False negative**
True positive	56794	7547
False negative	12372	51969

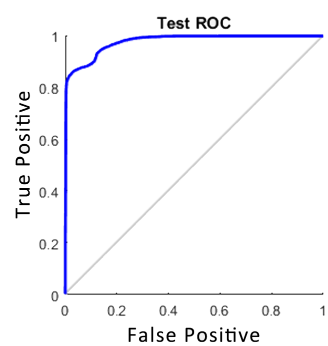


Figure 9. ROC curve obtained using P, PR, PT, QRS, QT, ST, T variance values.

Table 8. Confusion matrix of the trained network with the data obtained by applying crosscovariance to all intervals and segments.

Class	**True positive**	**False negative**
True positive	64194	147
False negative	1041	63300

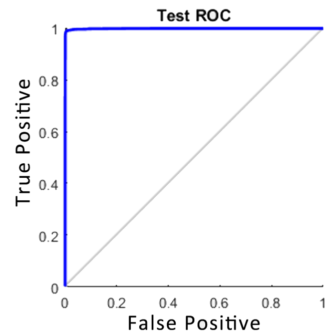


Figure 10. ROC curve obtained using P, PR, PT, QRS, QT, ST, T crosscovariance values.

One-way analysis of variance (ANOVA) was used to compare the obtained results. The selected methods were run 30 times separately and their fitness values were recorded. ANOVA test fitness values and fields representing methods are shown in Table 9.

ANOVA test is a statistical method used when the value to be compared is more than two. ANOVA test results are given in Table 10.

- Sum of squares (SS)
- Degrees of freedom (DoF)
- MS is the squared error value for each variation
- F is the mean squared error rate
- The probability that the test statistic is greater than the calculated value is defined as the P-value.

Table 9. Classes which are used for ANOVA testing.

1	ABC algorithm choices
2	Applying FFT to all ranges and segments
3	Applying fft_rmse to all ranges and segments
4	Applying var to all ranges and segments
5	Applying xcorr_max to all ranges and segments
6	Applying xcov_max to all ranges and segments

Table 10. ANOVA test results.

SS	**dF**	**MS**	**F**	**P-Value > F**
6920.41	5	1384.08	131.27	3.945 $\times e^{-57}$
1834.64	174	10.54		
8755.06	179			

As it can be seen in Table 10, it can be said that there is a significant difference between the groups since 0.05>> $3.945 \times e^{-57}$ A posthoc test should be done to see which groups differ. Posthoc test results are given in Table 10. ANOVA tests are an analysis method used to determine significant difference in more than two groups. The ANOVA test gives information about whether the groups are different from each other, but not which groups are different from each other. To see which groups are different from each other, posthoc tests were used. In the posthoc analysis, the groups were compared in pairs. We have two hypotheses to use in

ANOVA analysis.

H_0: There is no significant difference between the selected methods.

H_1: There is a significant difference between the selected methods.

As seen in Table 11, the number 1 group representing the methods selection by the ABC algorithm for each interval and segment has a significant difference from the groups 2, 3, 4, and 5. There is no significant difference between crosscovariance values representing group 6 and applied to all intervals and segments and ABC selections as $0.05 < 0.984$.

As a result, while the H_1 hypothesis was valid for the 6th group, the H_0 hypothesis was valid for the other groups. The process of applying crosscovariance to all waves and segments with the methods selection by the ABC algorithm is shown in Table 12 in terms of sensitivity and precision. Due to the closeness of the accuracy values, precision and recall values were calculated. To summarize, the overall accuracy and success according to methods for diagnosing atrial fibrillation is given in Figure 11. Table 13 shows the selection methods and related ranges and segments.

Table 11. Posthoc test results.

Group	Other group	P
1	2	0.0000
	3	0.0000
	4	0.0000
	5	0.0031
	6	0.9842
2	3	0.0000
	4	1.0000
	5	0.0000
	6	0.0000
3	4	0.0000
	5	0.0000
	6	0.0000
4	5	0.0000
	6	0.0000
5	6	0.0312

Table 12. ABC selections and crosscovariance.

	ABC algorithm selection	Crosscovariance
Accuracy(%)	99.98	99
Precision(%)	99.97	98.4
Recall(%)	99.98	99.77

5. Discussion

It has been observed in the literature that it provides higher accuracy, precision and sensitivity than classifications made using ABC algorithm on ECG signals. The multicore SVM structure developed in 2021 to classify

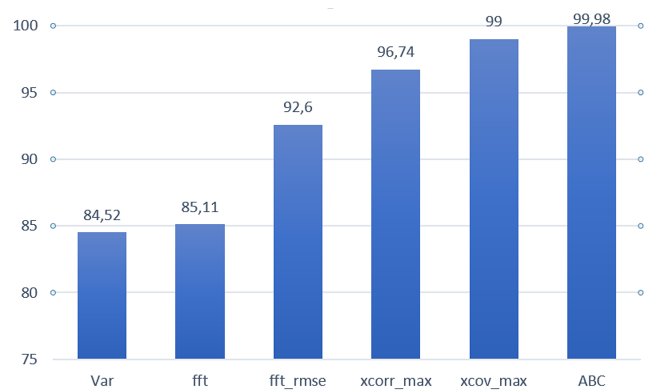


Figure 11. Overall accuracy success of atrial fibrillation diagnosis by methods.

Table 13. ABC selections and crosscovariance.

Wave or segment	**Process**
P	Crosscorrelation
PR	Crosscovariance
QRS	Fourier Transform (FT) Square error method with FT Variance Value
T	Crosscorrelation
Accuracy: 99.98%, Precision: 99.97%, Recall: 99.98%	

stressful ECG signals was further developed using GA, ABC and improved PSO, achieving 98.93% accuracy, 96.83% precision, 96.83% sensitivity has been done [7]. In order to classify the ECG signals, certain processes were applied to the signal and the classifier success was tried to be increased. For this, linear discriminant analysis (LDA) was used and the dimensionality of the data was reduced. Finally, well-known optimization techniques such as GA, genetic programming (GP) and ABC have been used to classify the ECG signal. Experimental results show that the accuracy of GA, GP and ABC classifiers is 94.41% (CI), 91.2% (GP), and 90.8% (ABC) [32]. Our study has achieved higher performance than these previous studies. The success of the model we developed is high, but the working time is long. A disadvantage of our approach can be seen as the extra computational time since the iterations of the ANN took a considerable amount of time.

6. Conclusion

The ECG sign is one of the most important sources of information about the parts of the heart. Each interval and segment of this sign represents a different part of the heart. Separating and processing these intervals and segments separately will give more accurate results. Our study has proven that instead of applying the same process to each range and segment obtained, applying the most appropriate process to the range and segment and classifying it gives better results. The system we propose overlaps with the trait vectors that characterize diseases. As a result, when defining atrial fibrillation, rapid, irregular, fibrillation waves of different shapes and sizes occur instead of P waves. Higher-than-expected heart rate and wide QRS complexes and irregular RR intervals are seen during atrial fibrillation [29–31]. In addition to the high success rate of the hybrid system we developed, it similarly chose P, PR, QRS, and T waves for atrial fibrillation, and focused on P and QRS for diagnosis, supporting that it selected the correct set of features. Similar studies to be conducted on other types of arrhythmias in the future and choosing the most appropriate procedure according to the range and segment will increase the diagnostic accuracy in cases with limited measurement time and insufficient data.

References

[1] Ngouala CRN. ”Comparative Study Of Arrhythmia Classification From Electrocardiographic Signals”, Modeling Simulation and Data Analysis 2021.

[2] Murawwat S, Hafiz MA, Sana I, Malik MI, Raahemifar K. ”Denoising and classification of Arrhythmia using MEMD and ANN”, Alexandria Engineering Journal 2021; 61 2807–2823. doi: 10.1016/j.aej.2021.08.014

[3] Akay B, Karaboğa D. ”A modified artificial bee colony algorithm for real-parameter optimization”, Information Sciences 2012; 192: 120–142. doi:10.1016/j.ins.2010.07.015

[4] Dilmac S, Korurek M. ”A New ECG Arrhythmia Clustering Method Based on Modified Artificial Bee Colony Algorithm”, Comparison with GA and PSO Classifiers 2013; IEEE INISTA, doi: 10.1109/INISTA.2013.6577616

[5] Küpeli I, Sarucan A, Küpeli S. ”Solution of the Distributed Permutated Flow Shop Scheduling Problems and Artificial Bee Colony Algorithm”, El-Cezerî Journal of Science and Engineering 2020; 7 (2): 549-562, doi :10.31202/ecjse.670424

[6] Babu GC, Shankar MG, Priyanka GS, Velusamy S, Vidyavathi K.”Performance Analysis Of Linear Discriminant Analysis (Lda) With Optimization Techniques For Arrhythmia Classification From Ecg Signals” International Journal of Aquatic Science 2021; ISSN: 2008-8019 12(3)

[7] Malhotra V, Sandhu MK. ”Improved ECG based Stress Prediction using Optimization and Machine Learning Techniques”, EAI Transactions on Scalable Information Systems 2021; ISSN 2032-9407, doi: 10.4108/eai.6-4-2021.169175

[8] Dilaveris PE, Gialafos JE. ”P-wave dispersion: a novel predictor of paroxysmal atrial fibrillation”, Ann Noninvasive Electrocardiol 2001; 6 (2): 159–165, doi: 10.1111/j.1542-474X.2001.tb00101.x

[9] Albert MR, Luis AR, Irene RD, Maria G, Joan S.”Analysis of the Association Between Electrocardiographic P-wave Characteristics and Atrial Fibrillation in the REGICOR Study”, Rev Esp Cardiol 2017; 70 (10): 841–847, doi: 10.1016/j.rec.2017.02.019

[10] Yiheng Y, Cheng C, Penghong D, Suman T, Lianjun G. ”The ECG Characteristics of Patients With Isolated Hypomagnesemia”, Frontiers in Physiology 2021; doi: 10.3389/fphys.2020.617374

[11] Tang X, Hu Q, Tang W. ”A Real-Time QRS Detection System With PR/RT Interval and ST Segment Measurements for Wearable ECG Sensors Using Parallel Delta Modulators” IEEE Transactions On Biomedical Circuits And Systems 2018; 12 (4): 751-761, doi: 10.1109/TBCAS.2018.2823275

[12] Wijaya C. ”Abnormalities State Detection from P-Wave QRS Complex and T-Wave in Noisy ECG”, Journal of Physics Conference Series 2019; 1230 (1).doi:10.1088/1742-6596/1230/1/012015

[13] Kumar A, Riaz SU, Rajesh, Kazmi SHM, Kumar R et al. ”Prolongation of QT Interval in ECG: A Hidden Complication of Cirrhotic Liver Disease” Annals Abbasi Shaheed Hospital and Karachi Medical and Dental College 2019; ISSN 1563-3241 24(1)

[14] Pan J, Tompkins WJ. ”A Real-Time QRS Detection Algorithm”, IEEE Transactions on Biomedical Engineering 1985; 32 (3): 230-236, doi: 10.1109/TBME.1985.325532

[15] Indik JH, Pearson EC, Fried K, Woosley RL. ”Bazett and Fridericia QT correction formulas interfere with measurement of drug-induced changes in QT interval”, Heart Rhythm 2006; 3 (9) 1003-1007 doi: 10.1016/j.hrthm.2006.05.023

[16] Deng B, Tao R, Wang Y. ”Convolution theorems for the linear canonical transform and their applications”, Science in China Series F: Information Sciences 2006; 49 (5) 592–603

[17] Mikki S, Sarkar D, Antar Y. ”Near-Field Cross-Correlation Analysis for MIMO Wireless Communications”. IEEE Antennas And Wireless Propagation Letters 2019; 18 (7) 1357 - 1361, doi: 10.1109/LAWP.2019.2916627

[18] Demirsoy S, Akçapinar H. ”Kuzularda Büyümeyi Etkileyen Çevresel Faktörlerin Kovaryans Analizi İle İncelenmesi”, Lalahan Hayvancilik Arastirma Enstitusu Dergisi 1997; 37 (1) 37-55 (in Turkish).

[19] Bağlan S. ”JPEG XR Başarimi”, PhD, Istanbul Technical University 2014; Istanbul, Turkey

[20] Bulut F. ”Different Mathematical Models for Entropy in Information Theory”, Bilge International Journal of Science and Technology Research 2017; 1 (2): 167-174 ISSN: 2587-0742

[21] Minami K, Nakajima H, Toyoshima T. ”Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network”, IEEE Transac-tions on Biomedical Engineering 1999; 46, 179-185, doi: 10.1109/10.740880

[22] Akay B. ”Performance analysis of artificial bee colony algorithm on numerical optimization problems”, Phd. Erciyes University 2009; Kayseri, Turkey

[23] Xue Y, Jiang J, Zhao B, Ma T. ”A self-adaptive artificial bee colony algorithm based on global best for global optimization”, Soft Computing 2018; 22: 2935–2952, doi: 10.1007/s00500-017-2547-1

[24] Rao H, Shi X, Rodrigue AK, Feng J, Xia Y. ”Feature selection based on artificial bee colony and gradient boosting decision tree”, Applied Soft Computing Journal 2019; 74: 634-642 doi: 10.1016/j.asoc.2018.10.036

[25] Alpaslan F, Egrioglu E, Aladag HC, Ilter D, Dalar Z. ”Comparison of Multiplicative Neuron Models Using ABC and BP Training Algorithms: An Application to Istanbul Gold Change”, Anadolu University Journal Of Science and Engineering 2013; 14 (3): 315-328.

[26] Bishop CM. ”Neural Networks for Pattern Recognition”, Oxford University Press Inc. 1995; New York NY USA.

[27] Savalia S, Emamian V. ”Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Networks” Bioengineering 2018; 5 (2) doi: 10.3390/bioengineering5020035

[28] Konakoglu B. ”Cok Katmanli Algilayici Yapay Sinir Agi ile Jeodezik Elipsoidal Koordinatlarin (ϕ, λ, h) 3 Boyutlu Global Kartezyen Koordinatlara (X, Y, Z) Donusumu”, Gumushane University Journal of Science Institute 2020; 10 (3): 702-710 (in Turkish), doi: 10.17714/gumusfenbil.712100

[29] Camm AJ, Kirchhof P, Gregory YHL, Schotten U, Savelieva I. ”Atrial fibrillation treatment guide”, European Heart Journal 2010; 31 (19): 2369–2429, doi: 10.1093/eurheartj/ehq278

[30] Kaypakli O, Koca H, Şahin DY, Okar S, Karataş F. ”Association of P wave duration index with atrial fibrillation recurrence after cryoballoon catheter ablation”, Journal of Electrocardiology 2018; 51 (2): 182-187. doi: 10.1016/j.jelectrocard.2017.09.016

[31] Palano F, Adduci C, Cosentino P, Silvetti G, Boldini F. ”Assessing Atrial Fibrillation Substrates by P Wave Analysis A Comprehensive Review”, High Blood Pressure & Cardiovascular Prevention 2020; 27:341–347, doi: 10.1007/s40292-020-00390-1

[32] Malhotra V, Sandhu MK. ”Improved ECG based Stress Prediction using Optimization and Machine Learning Techniques”, EAI Transactions on Scalable Information Systems 2021; ISSN 2032-9407, doi: 10.4108/eai.6-4-2021.169175