[Turkish Journal of Electrical Engineering and Computer Sciences](https://journals.tubitak.gov.tr/elektrik)

Volume 26 | Number 6 Article 4

1-1-2018

# Comparative analysis of MABC with KNN, SOM, and ACO algorithms for ECG heartbeat classification

SELİM DİLMAÇ

ZÜMRAY ÖLMEZ

TAMER ÖLMEZ

Turkish Journal of Electrical Engineering & Computer Sciences

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

Research Article

Turk J Elec Eng & Comp Sci
(2018) 26: 2819 – 2830
© TÜBİTAK
doi:10.3906/elk-1712-328

# Comparative analysis of MABC with KNN, SOM, and ACO algorithms for ECG heartbeat classification

**Selim DİLMAÇ***, **Zümray DOKUR**, **Tamer ÖLMEZ**
Department of Electronics and Communication Engineering, Faculty of Electrical and Electronics Engineering, İstanbul Technical University, İstanbul, Turkey

**Received:** 24.12.2017 • **Accepted/Published Online:** 09.07.2018 • **Final Version:** 29.11.2018

**Abstract:** In this paper, we proposed a classification method based on a nature-inspired algorithm, i.e., modified artificial bee colony (MABC). This method was applied to electrocardiogram (ECG) heartbeat classification. ECG data was obtained from MITBIH database. Eight different types of heartbeats (N, j, V, F, f, A, a, and R) were analyzed. For a better classification result, both time domain and frequency domain features were used. Feature selection was done by divergence analysis. MABC classification accuracy and heartbeat sensitivity values were compared with the results of other methods. Among other classifiers, k-nearest neighbor (KNN), Kohonen's self-organizing map (SOM), and ant colony optimization (ACO) were the best performing ones, and therefore their results are presented. The MABC classifier achieved 97.18% accuracy on the analyzed dataset, as well as high sensitivity values for heartbeat types.

**Key words:** ECG heartbeat, data classification, ABC algorithm, nature-inspired

## 1. Introduction

Electrocardiogram (ECG) is a record of heart's electrical activity. Most of the cardiac disorders, called arrhythmia, may appear anytime during a day. In 24-h long term records, over 100,000 beats per day need to be analyzed. Manual analysis of electrocardiographic signal by cardiologists through visual inspection requires too much time. As a result, computer-aided, automated electrocardiographic signal analysis for arrhythmia classification has been an active research topic for the last couple of decades.

For this purpose, various methods are used in the literature including linear and nonlinear classifiers [1–9]. In addition to the classical methods such as the nearest mean classifier, k-means, k-nearest neighbor, naive Bayesian, and neural network classifier, there are many newly proposed nature-inspired algorithms, as well as some other hybrid methods for ECG arrhythmia classification. Some of them are ant colony optimization (ACO) [10,11], genetic algorithm (GA) [12,13], particle swarm optimization (PSO) [14,15], and artificial bee colony (ABC) [16,17]. Nature-inspired algorithms may imitate a process in nature or may use an approach simulating some intelligent behaviors of social animal groups in nature such as bird flocks and ant or bee colonies. These kinds of algorithms are called swarm intelligence (SI) algorithms.

This study focuses on the ABC algorithm which is one of the recent swarm intelligence methods. The ABC algorithm was first proposed by Karaboga, inspired by the foraging behavior of honey bee colonies [18]. Since then, Karaboga and other researchers have used this algorithm in many application areas including data

*Correspondence: sdilmac@arroweurope.com

clustering [16,24–26] and developed its variants and derivatives [19–23]. Also, there are detailed survey studies about the ABC algorithm [27,28].

ABC is a simple and flexible algorithm; it uses few control parameters. It can converge to the global best solution because of its random search mechanism. Unlike some classical methods, it does not get stuck at local good solutions. The k-nearest neighbor (KNN) and neural network (NN) classifiers are widely used classical methods to solve the data classification problem. KNN is a supervised method which stores all training samples in memory and makes the classification of test samples by using a similarity measure. It is a simple and relatively high-accuracy algorithm [29,30]. It does not have a training phase; however, during the test phase, it must make distance calculations between the test sample and all training samples, which makes it not so suitable for real-time systems. Self-organizing map (SOM) is a type of artificial neural network (ANN). It was introduced by Kohonen in 1982 and is used in many applications [31,32]. SOM network is trained using unsupervised learning to produce a low-dimensional representation of the input space of the training samples, which is called a map. By using the class information of output nodes in the map, classification of unknown patterns can be performed. Ant colony optimization (ACO) is a clustering algorithm based on the ant system which was developed by Dorigo in 1991. It is a multiagent system which has interactions between its agents, converging to an optimal solution. It is applied to various real-life problems [33,34].

In this study, the ABC algorithm was modified by using a new control parameter, i.e., scout conversion threshold ratio (SCTR) [16], and also the fitness function was updated. A classifier based on the modified ABC algorithm, called MABC, was applied to the ECG dataset. MABC algorithm-based classifier result was compared with two other classification methods, i.e., KNN and SOM. During this comparison, the effect of iteration number and network dimension in the SOM algorithm was evaluated. Experimental results showed that the MABC algorithm classifier has the highest classification accuracy on this dataset among the compared methods.

The rest of this paper is organized as follows. In Section 2, we present the ABC algorithm and its modified version, as well as other classification methods which are used in this study. ECG dataset details and feature extraction method are introduced in Section 3. In Section 4, classification results of the three methods on ECG dataset and discussion are presented. Finally, we conclude the paper in Section 5.

## 2. Classification methods

### 2.1. The ABC algorithm and its modified version

The ABC algorithm is a nature-inspired, population-based algorithm. Honey bee swarms have a specific behavior for foraging. They can find a food source by random search, evaluate its value, memorize its location, and share that information with other colony members, so those other bees can go directly to the food source without random search. By using this method, they can exploit the richest food sources in the shortest time and making the least effort possible.

Positions of food sources around the hive represent possible solutions of the problem. The nectar amount in a food source represents the fitness of these solutions. There are three types of bees in the ABC algorithm: employed bees, onlooker bees, and scout bees. Employed and onlooker bees execute the exploitation process, while scout bees execute the exploration process [14].

In the ABC algorithm, each employed bee works on a food source. It represents the location of a possible solution. That location is defined by a p-dimensional vector in the solution space. During the best solution search process, Eq. (1) is used.

$$x_{gm}^{q}(t+1) = x_{gm}^{q}(t) + (\varphi_{gm} \times (x_{gm}^{q}(t) - x_{gm}^{r}(t))) \tag{1}$$

r: randomly selected another employed bee, r$\neq$q; m: randomly selected dimension;

$\varphi$: random number between $[-1, 1]$; t: iteration number in the search process; g: cluster number.

In the original ABC algorithm, during the search process if there is no fitness improvement at a location of an employed bee during the last "limit" iteration steps, that location is abandoned, and that employed bee becomes a scout bee. Scout bee restarts to search at a random location. This mechanism avoids being stuck at local good solutions and enables the algorithm to find the global best solution. However, it has also a negative effect on performance. When an employed bee finds a good location, even it could be the best possible location; there may be no improvement on fitness value after a high number of iteration, bigger than the "limit". The original ABC algorithm abandons that location and the search starts at a random location. This causes the location of that good solution to be forgotten. To avoid this problem, we introduced a new control parameter called the scout conversion threshold ratio (SCTR). In the modified ABC algorithm, a location is not abandoned if that location has a good fitness value. SCTR can be chosen between [0.5, 0.99]. In this study, we used 0.7 for it. By using this modification, higher classification accuracy can be achieved. In a previous study, we tested the original ABC algorithm- and MABC algorithm-based classifiers on a different dataset. In that dataset, there were three beat types: N, V, and A. Arrhythmic beats V and A have subtypes, we used each of them as a different class, so we worked on a 7-class dataset. A total of 8848 normal and arrhythmic heartbeats were analyzed. Only time domain features were used as they were sufficient to classify those beat types. The ABC algorithm-based classifier had 98.33% accuracy, while the MABC classifier achieved 99.30% accuracy on that dataset [16].

In the original ABC algorithm-based classifier, minimizing the fitness function is aimed to find the best cluster centers. Fitness is defined as the sum of distances of the training patterns to the nearest cluster center. This may result in low classification accuracy, when there are unbalanced samples in different classes of datasets. Heartbeat types have such a characteristic. Number of normal beats is much higher than that of arrhythmic beats in ECG. In this study, we used classification accuracy as fitness function. That modification improved the classification result.

Flowchart of the MABC algorithm is provided in Figure 1.

### 2.2. Other classification methods: SOM, KNN, and ACO

SOM is a type of artificial neural network, also called Kohonen's network. Training of this network is an unsupervised learning process, producing low-dimensional (usually 2-dimensional) representation of input space of the training samples, called a map. Structure of SOM can be found in [31]. SOM is different from other neural networks as it uses a neighborhood function to update weights of each node during training process. Once the training phase is completed, a class label is assigned to each node, according to maximum number of class training samples which are closest to that node. Test samples are classified with class label of the nearest output node's label.

KNN classifier is an instance-based learning method, which stores all training sample vectors. It is a very simple and effective method especially for high-dimensional problems. It classifies the new unknown test samples based on similar training samples. Similarity measure is usually the Euclidean distance. All the sample

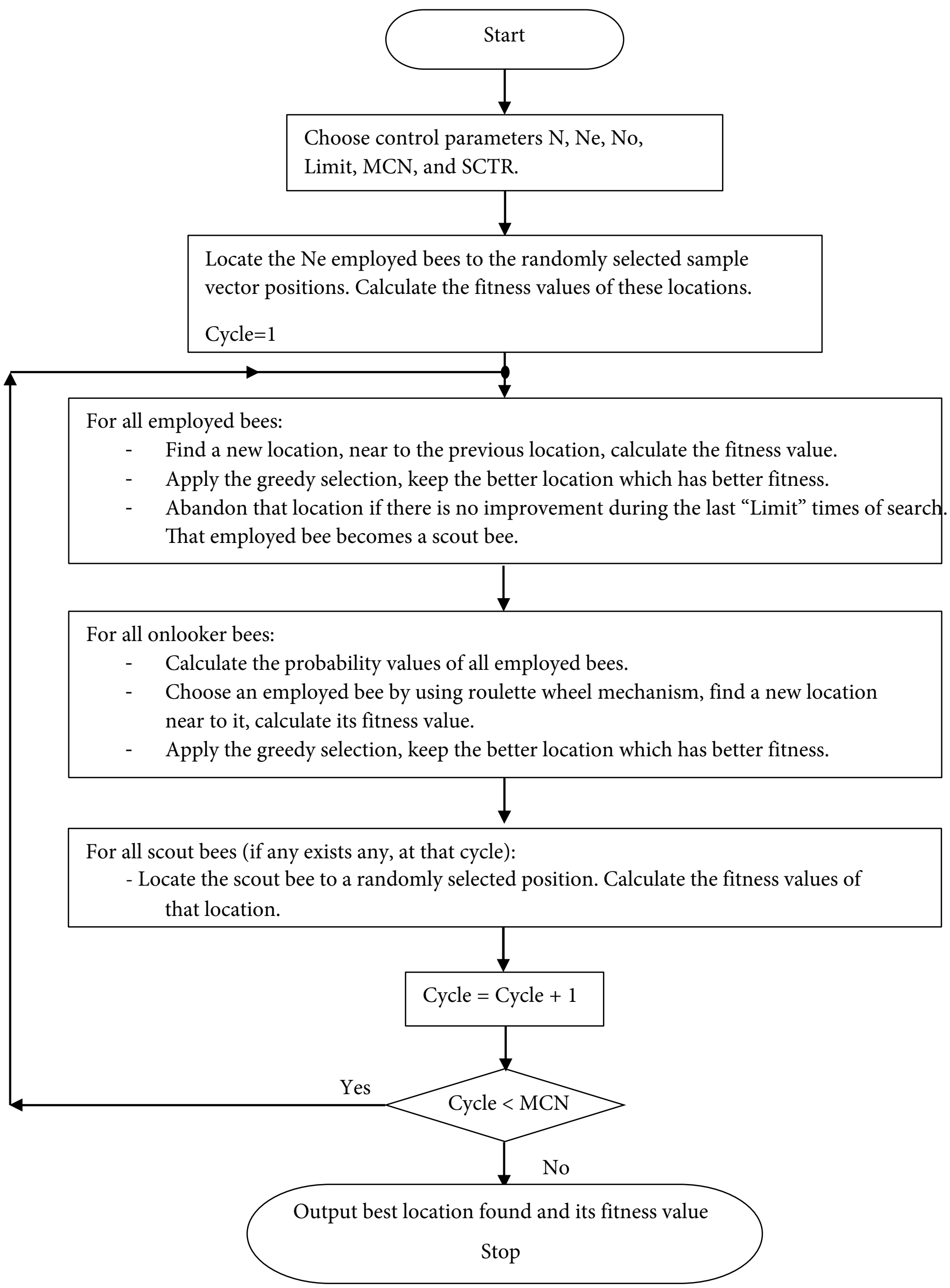



**Figure 1**. Flowchart of the MABC algorithm.

vectors in the training set are used for classification of test samples. Therefore, it has a disadvantage in terms of required memory space and test time.

The ACO algorithm is a nature-inspired problem-solving method. For the clustering problem, ants visit other nodes randomly and they lay some pheromone on that path. After several iterations, total pheromone on a path between close nodes raises, while the pheromone on longer paths evaporates. Finally, certain number of clusters are formed.

## 3. ECG dataset and feature extraction

Heartbeats can be classified as normal and arrhythmic beats. Arrhythmic beats are indications of heart diseases; they are defined by their features. In this study, the analyzed ECG data is obtained from the MIT-BIH database, which was developed by Massachusetts Institute of Technology. Detailed description of the database can be found in [35] and it can be downloaded from physionet website (http://physionet.org/physiobank/database/mitdb/).

### 3.1. ECG dataset

The following heartbeat samples were obtained from the MIT-BIH database and classified in this study: normal beat (N), junctional escape beat (j), premature ventricular contraction (PVC - V), fusion of PVC and normal beat (F), fusion of paced and normal beat (f), atrial premature beat (APB - A), aberrated APB (a), and right bundle branch block (RBBB - R). Different types of heartbeat samples are shown in Figure 2. A total of 16 records of MIT-BIH were used (listed as follows: 100, 105, 116, 119, 201, 202, 205, 208, 209, 210, 213, 217, 222, 223, 231, and 233). These records were chosen to include ventricular, supraventricular, and junctional arrhythmias and conduction abnormalities. Some of the heartbeat type samples are limited, so especially those records were used to increase the number of the analyzed arrhythmic beats as much as possible. Randomly chosen samples were included in the training and test sets. About 50% of the arrhythmic heartbeat samples were used for training. As there were many normal beat samples, about 10% of them were used. Number of beats for each heartbeat type is shown in Table 1.

**Table 1**. Heartbeat type sample numbers.

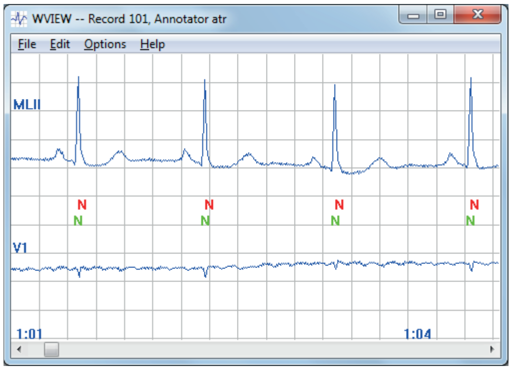
| | Beat types | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | N | j | V | F | f | A | a | R |
| Training | 2534 | 700 | 109 | 378 | 436 | 107 | 338 | 67 | 399 |
| Test | 8735 | 6896 | 110 | 379 | 436 | 108 | 339 | 68 | 399 |
| Total | 11269 | 7596 | 219 | 757 | 872 | 215 | 677 | 135 | 798 |

### 3.2. Feature extraction

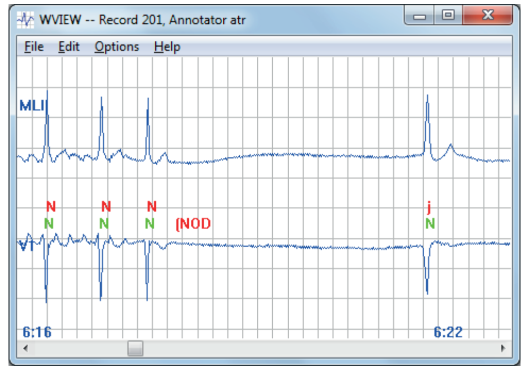
ECG signals include noises such as motion artefact and power line noise. The Pan–Tompkins algorithm-based Physionet “physio toolkit, ecgpuwave” software was used to filter the noises and also to detect the beginning and end, as well as the peak points of p, qrs, and t waves in ECG signal. A feature extraction software program was developed, and 38 different time domain features were calculated [16].

In addition to time domain features, some types of heartbeats such as fusion beats need frequency domain features to have better classification accuracy. The discrete Fourier transform (DFT) of a window, which included 256 sampled data points around the R-wave peak, was used for this purpose. The DFT generated 128 coefficients, about 40 of which had significance.

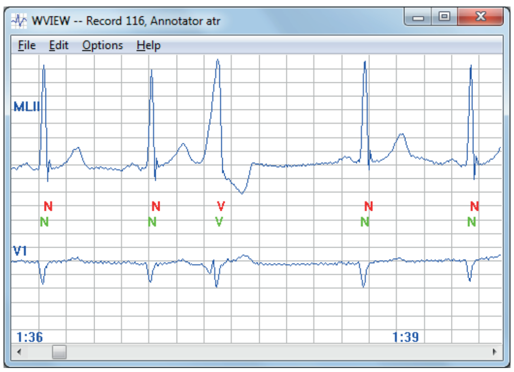
In order to find the most divergent features, we used divergence analysis in this study. By utilizing the within-class (W) and between-class (B) covariance matrixes, divergence analysis provides the divergence values of features [36]. For a given feature, if it makes the between-class distribution higher (samples in different clusters are dissimilar to each other) and the within-class distribution is lower (samples in the same cluster are similar to each other), then that feature is a high-divergent one.



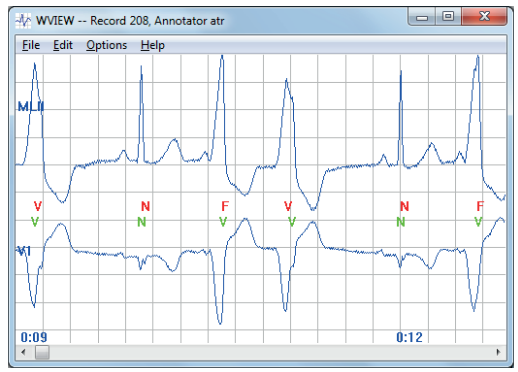

a. “N” type heart beat samples in record 101



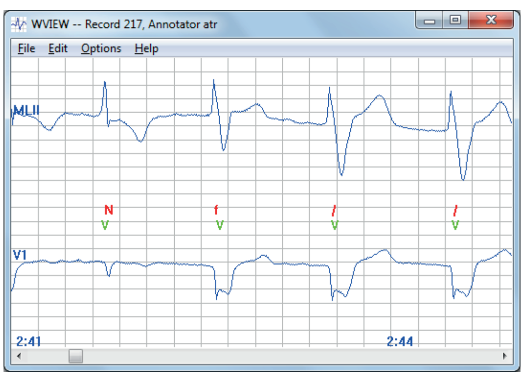

b. “j” type heart beat sample in record 201



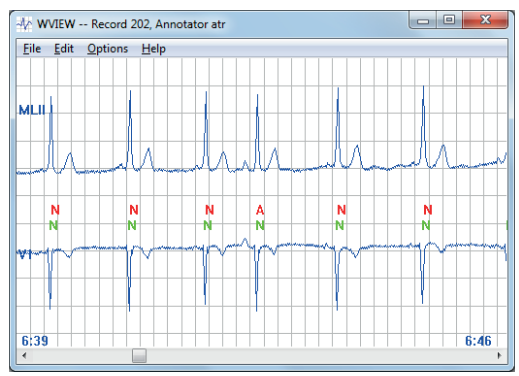

c. “V” type heart beat sample in record 116



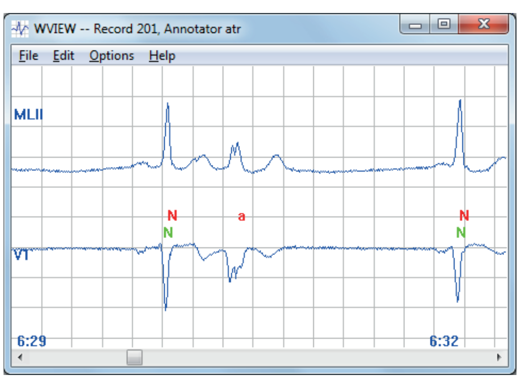

d.“F” type heart beat samples in record 208



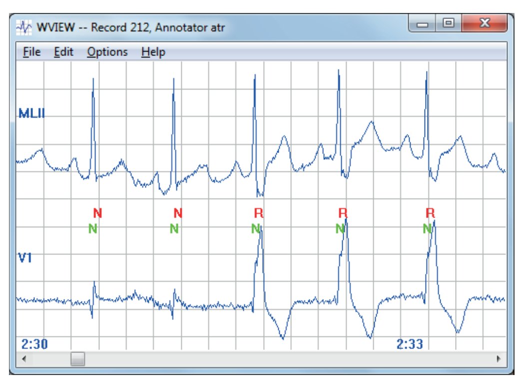

e. “f” type heart beat sample in record 217




f. “A” type heart beat sample in record 202




g.“a” type heart beat sample in record 201




h.“R” type heart beat samples in record 212

**Figure 2**. Different types of heartbeat samples.

By using divergence analysis, we reduced the 38 time domain and 40 frequency domain features down to a total of 15 features. Time domain features can be listed as follows: (1) P-wave existence, (2) R-peak height, (3) QRS area, (4) absolute QRS area, (5) minimum value between R-to-R peaks, (6) time interval between R(t)-to-R(t-1), and (7) time interval between R(t+1)-to-R(t). Frequency domain features are DFT coefficients 3, 23, 2, 10, 4, 5, 1, and 16.

## 4. Results and discussion

In this study, the developed classification software based on the MABC algorithm was executed on a computer which has 2.53 GHz Intel Core i5 processor and 3 GB RAM. The software was developed and run on Mathworks MATLAB. Once the cluster centers were calculated in training phase, 8735 test samples from eight different clusters were classified in 3.52 seconds.

There are three main control parameters in the standard ABC algorithm: bee number in hive (N), MCN, and limit. The MABC algorithm has random search characteristics. Therefore, for a combination of a control parameter set, the program is executed 5 times on training set. The best result cluster center set is used to make classification on the test set. Confusion matrix result of the MABC classifier on ECG test dataset is shown in Table 2.

**Table 2**. Confusion matrix result of the MABC classifier on the ECG test set.

| | | Predicted heartbeat types | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | j | V | F | f | A | a | R |
| True heartbeat types | N | 6807 | 0 | 0 | 17 | 19 | 49 | 4 | 0 |
| | j | 0 | 105 | 0 | 2 | 0 | 3 | 0 | 0 |
| | V | 0 | 0 | 354 | 25 | 0 | 0 | 0 | 0 |
| | F | 69 | 1 | 3 | 346 | 6 | 3 | 6 | 2 |
| | F | 6 | 3 | 0 | 3 | 94 | 0 | 2 | 0 |
| | A | 15 | 0 | 0 | 0 | 0 | 324 | 0 | 0 |
| | a | 1 | 1 | 0 | 0 | 0 | 3 | 60 | 3 |
| | R | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 399 |

Classification accuracy is defined as the number of total true positive test samples over total test samples. According to the confusion matrix in Table 8, MABC classification accuracy on this dataset was found as 97.18%.

In addition to system classification accuracy, there are three other standard metrics to quantify the performance of the system: sensitivity (Se), specificity (Sp), and positive predictivity (Ppr) [16]. We used them to evaluate the results of our method. Performance statistics of the MABC classifier is shown in Table 3.

**Table 3**. Performance statistics of the MABC classifier.

| (%) | N | j | V | F | f | A | a | R |
|---|---|---|---|---|---|---|---|---|
| Sensitivity | 98.71 | 95.45 | 93.40 | 79.36 | 87.04 | 95.58 | 88.24 | 100.0 |
| Specificity | 95.05 | 99.94 | 99.96 | 99.43 | 99.71 | 99.31 | 99.86 | 99.94 |
| Pos. predictivity | 98.68 | 95.45 | 99.16 | 88.04 | 78.99 | 84.82 | 83.33 | 98.76 |

Kohonen's SOM algorithm has two control parameters: iteration number and number of nodes. In order to find out the effect of the network size (node number) and iteration number on classification accuracy, 36 tests were run by using D = 10, 14, 20, and 30; iteration no = 10000, 20000, 30000. For each parameter set, 3 tests were run. Kohonen's SOM classifier test results by using different control parameters are presented in Table 4. Average classification accuracy increases by higher iteration numbers. However, calculation time also considerably increases by larger network size and higher iteration number. Also, classification increase rate gets lower after 20 × 20 network. We used the best of these test results to compare them with those of other methods. The best accuracy achieved with the SOM classifier was 94.88%. Confusion matrix result of the SOM classifier on the ECG test dataset is shown in Table 5.

**Table 4**. Kohonen's SOM classifier test results by using different control parameters.

| Network size | Iteration | Test1 (%) | Test2 (%) | Test3 (%) | Average (%) | Aver(Aver) (%) |
|---|---|---|---|---|---|---|
| 10 × 10 | 10 K | 85.62 | 75.00 | 81.38 | 80.67 | 79.72 |
| 10 × 10 | 20 K | 76.88 | 94.29 | 73.18 | 81.45 | |
| 10 × 10 | 30 K | 63.76 | 86.86 | 80.54 | 77.05 | |
| 14 × 14 | 10 K | 80.53 | 80.91 | 93.87 | 85.10 | 86.75 |
| 14 × 14 | 20 K | 91.75 | 79.42 | 88.01 | 86.39 | |
| 14 × 14 | 30 K | 87.08 | 86.47 | 92.74 | 88.76 | |
| 20 × 20 | 10 K | 94.02 | 89.58 | 94.42 | 92.67 | 92.17 |
| 20 × 20 | 20 K | 90.67 | 94.13 | 90.44 | 91.75 | |
| 20 × 20 | 30 K | 92.18 | 90.90 | 93.20 | 92.09 | |
| 30 × 30 | 10 K | 94.47 | 94.13 | 94.41 | 94.34 | 94.01 |
| 30 × 30 | 20 K | 93.98 | 94.33 | 93.04 | 93.78 | |
| 30 × 30 | 30 K | 92.32 | 94.49 | 94.88 | 93.90 | |

**Table 5**. Confusion matrix result of Kohonen's SOM classifier on the ECG test set.

| | Predicted heartbeat types | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | j | V | F | f | A | a | R |
| True heartbeat types | N | 6612 | 2 | 1 | 22 | 41 | 215 | 2 | 1 |
| | j | 7 | 103 | 0 | 0 | 0 | 0 | 0 | 0 |
| | V | 15 | 0 | 360 | 4 | 0 | 0 | 0 | 0 |
| | F | 38 | 0 | 0 | 394 | 3 | 1 | 0 | 0 |
| | f | 15 | 2 | 0 | 3 | 86 | 1 | 0 | 1 |
| | A | 38 | 1 | 0 | 1 | 0 | 299 | 0 | 0 |
| | a | 11 | 0 | 0 | 3 | 0 | 1 | 52 | 1 |
| | R | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 382 |

Confusion matrix results of the KNN and ACO classifiers on the ECG test dataset are shown in Tables 6 and 7, respectively. Classification accuracy of the KNN classifier on the ECG dataset was 96.73%, while it was 96.84% for the ACO classifier. Classification accuracy and classification time of classifiers are provided as bar chart in Figure 3.

Classification accuracy and sensitivity values of beat types using different classifiers on the test set are given in Table 8. Classification time of different classification algorithms are provided in Table 9.

**Table 6**. Confusion matrix result of the KNN classifier on the ECG test set.

| | Predicted heartbeat types | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | j | V | F | f | A | a | R |
| True heartbeat types | N | 6637 | 0 | 0 | 54 | 12 | 189 | 4 | 0 |
| | j | 0 | 110 | 0 | 0 | 0 | 0 | 0 | 0 |
| | V | 0 | 0 | 376 | 2 | 1 | 0 | 0 | 0 |
| | F | 3 | 0 | 1 | 427 | 2 | 0 | 1 | 2 |
| | f | 0 | 2 | 0 | 0 | 105 | 0 | 1 | 0 |
| | A | 6 | 1 | 0 | 0 | 0 | 332 | 0 | 0 |
| | a | 1 | 0 | 0 | 2 | 1 | 1 | 63 | 0 |
| | R | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 399 |

**Table 7**. Confusion matrix result of the ACO classifier on the ECG test set.

| | Predicted heartbeat types | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | j | V | F | f | A | a | R |
| True heartbeat types | N | 6674 | 0 | 0 | 55 | 6 | 161 | 0 | 0 |
| | j | 0 | 110 | 0 | 0 | 0 | 0 | 0 | 0 |
| | V | 0 | 0 | 373 | 5 | 1 | 0 | 0 | 0 |
| | F | 2 | 1 | 1 | 423 | 4 | 1 | 2 | 2 |
| | f | 1 | 1 | 0 | 1 | 103 | 1 | 1 | 0 |
| | A | 23 | 0 | 0 | 0 | 0 | 316 | 0 | 0 |
| | a | 2 | 0 | 0 | 2 | 2 | 1 | 61 | 0 |
| | R | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 399 |

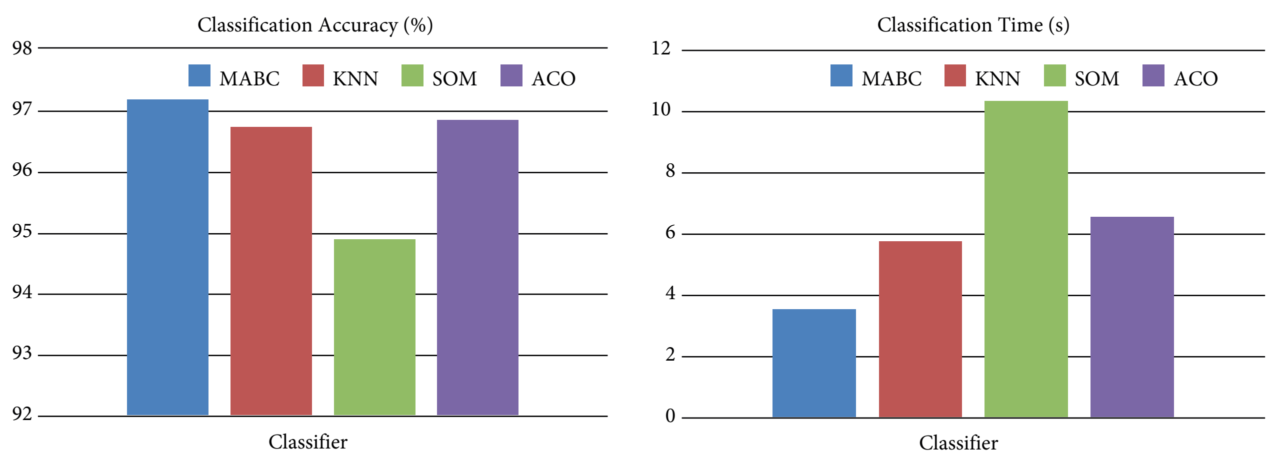



**Figure 3**. Classification accuracy and classification time of classifiers.

## 5. Conclusion

In this paper, a nature-inspired algorithm, the modified ABC (MABC), classifier was developed and applied on an ECG dataset for heartbeat classification. It has few control parameters and an easy-to-use and flexible algorithm. It detects and classifies heartbeat types automatically, which helps medical doctors, cardiologists to analyze long-term ECG records.

**Table 8**. Classification accuracy and sensitivity values of beat types using different classifiers on the test set.

| Classifier | Accuracy (%) | Sensitivity (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | j | V | F | f | A | a | R |
| MABC | 97.18 | 98.71 | 95.45 | 93.40 | 79.36 | 87.04 | 95.58 | 88.24 | 100.0 |
| KNN | 96.73 | 96.24 | 100.0 | 99.21 | 97.94 | 97.22 | 97.94 | 92.65 | 100.0 |
| SOM | 94.88 | 95.88 | 93.64 | 94.99 | 90.37 | 79.63 | 88.20 | 76.47 | 95.74 |
| ACO | 96.84 | 96.78 | 100.0 | 98.42 | 97.02 | 95.37 | 93.22 | 89.71 | 100.0 |

**Table 9**. Classification time of different classifiers on the test set.

| Classifier | Classification time (s) |
|---|---|
| MABC | 3.52 |
| KNN | 5.72 |
| SOM | 10.32 |
| ACO | 6.53 |

The result of the MABC on the ECG dataset was compared with those of other classifiers. Among other classifiers, the best performing methods were KNN, Kohonen's SOM, and ACO, whose results are presented in this article.

Feature set has an important effect on classification accuracy while analyzing ECG signals. By using the most divergent time and frequency domain features in the MABC algorithm, high classification accuracy (97.18%) was achieved by the MABC classifier, better than that of the other analyzed methods, i.e., KNN, SOM, and ACO classifiers. In total, 8 different types of heartbeats are classified with high-sensitivity values.

KNN has high-sensitivity values for arrhythmic beats, but it has relatively low sensitivity for normal beats. Another disadvantage of KNN is calculation time on test phase. It has to make distance calculation between test sample and training samples. Also, it needs much higher memory capacity to store the training samples, while the MABC needs to store only the cluster centers.

The SOM classifier has good sensitivity for some beat types, but it has low sensitivity for f- and a-type beats. Also, it has relatively low sensitivity for normal beats. The ACO has high sensitivity for most of the arrhythmic beats, and relatively low sensitivity for normal beats and A-type beats. Although it has lower sensitivity for normal beats, it has the second best accuracy result among the analyzed methods.

The MABC classifier can do classification on test samples quite fast, which enables this method to be used in real-time systems with high accuracy. In addition to the ECG dataset, the MABC classifier is applied to other datasets such as IRIS, PIMA, and BUPA. The results show that the MABC can be used to classify different types of datasets successfully.

In this study, eight different types of heartbeats were classified. These are the most common beat types in the MIT-BIH database. In future studies, classification accuracy can be analyzed for more numbers of beat types. In order to increase the accuracy and beat type sensitivity values, new features can be used. Wavelet transforms can be used to extract new features. Some heartbeat types such as V-type (PVC) beats can have multiple morphologies, in other words one class can consist of more than one cluster. For those types of beats, multiple clusters can be used for the same class sample data to improve the sensitivity.

Considering the high classification speed, in future studies the MABC classifier can be used to analyze and classify other biological signals in real time such as EMG and EEG. For this purpose, various feature extraction methods must be applied on datasets to find out the optimum feature set to obtain maximum divergence.

## References

[1] Chen G, Jaradat SA, Banerjee N, Tanaka TS, Ko MSH, Zhang MQ. Heartbeat classification using projected and dynamic features of ECG signal. Biomed Signal Process Control 2017; 31: 165-173.

[2] Osowski S, Linh TH. ECG beat recognition using fuzzy hybrid neural network. IEEE T Bio-med Eng 2001; 48: 1265-1271.

[3] Dokur Z, Ölmez T. ECG beat classification by a novel hybrid neural network. Comput Meth Prog Bio 2001; 66: 167-181.

[4] Ölmez T. Classification of ECG waveforms by using RCE neural network and genetic algorithms. Electron Lett 1997; 33: 1561-1562.

[5] Dong X, Wang C, Wenjie S. ECG beat classification via deterministic learning. Neurocomputing 2017; 240: 1-12.

[6] Mateo J, Torres AM, Aparicio A, Santos JL. An efficient method for ECG beat classification and correction of ectopic beats. Comput Electr Eng 2016; 53: 219-229.

[7] Wen C, Lin TC, Chang KC, Huang CH. Classification of ECG complexes using self-organizing CMAC. Measurement 2009; 42: 399-407.

[8] Abawajy JH, Kelarev AV, Chowdhury M. Multistage approach for clustering and classification of ECG data. Comput Meth Prog Bio 2013; 112: 720-730.

[9] Zidelmal Z, Amirou A, Ouldabdeslam D, Merckle J. ECG beat classification using a cost sensitive classifier. Comput Meth Prog Bio 2013; 111: 570-577.

[10] Korürek M, Nizam A. Clustering MIT-BIH arrhythmias with ant colony optimization using time domain and PCA compressed wavelet coefficient. Digit Signal Process 2010; 20: 1050-1060.

[11] Doğan B, Korürek M. A new ECG beat classification method based on kernelized fuzzy C-Means and hybrid ant colony optimization for continuous domains. Appl Soft Comput 2012; 12: 3442-3451.

[12] Vafaie MH, Ataei M, Koofigar HR. Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals. Biomed Signal Process Control 2014; 14: 291-296.

[13] Khazaee A, Ebrahimzadeh A. Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features. Biomed Signal Process Control 2010; 5: 252-263.

[14] Korürek M, Doğan B. ECG beat classification using particle swarm optimization and radial basis function neural network. Expert Syst Appl 2010; 37: 7563-7569.

[15] Shadmand S, Mashoufi B. A new personalized ECG signal classification algorithm using block-based neural network and particle swarm optimization. Biomed Signal Process Control 2016; 25: 12-23.

[16] Dilmaç S, Korürek M. ECG heartbeat classification method based on modified ABC algorithm. Appl Soft Comput 2015; 36: 641-655.

[17] Dilmaç S, Korürek M. A new arrhythmia classification method by using modified ABC (MABC) algorithm. In: ICMCA 2013 Fourth International Conference on Mathematical and Computational Applications; 11-13 June 2013; Manisa, Turkey. pp. 390-398.

[18] Karaboğa D. An idea based on honey bee swarm for numerical optimization. Technical report TR06, Erciyes University, Engineering Faculty, Computer Engineering Dept, 2005.

[19] Karaboğa D, Ökdem S, Öztürk C. Cluster based wireless sensor network routing using Artificial Bee Colony algorithm. Wirel Netw 2012; 18: 847-860.

[20] Taşgetiren F, Pan Q, Suganthan PN, Chen AHL. A discrete artificial bee colony algorithm for the total flow time minimization in permutation flow shops. Inform Sciences 2011; 181: 3459-3475.

[21] Tsai PW, Pan JS, Liao BY, Chu SC. Enhanced artificial bee colony optimization, Int J Innov Comput Inf Control 2009; 5: 5081-5092.

[22] Maeda M, Tsuda S. Reduction of artificial bee colony algorithm for global optimization. Neurocomputing 2015; 148: 70-74.

[23] Banitalebi A, Aziz MIA, Bahar A, Aziz ZA. Enhanced compact artificial bee colony. Inform Sciences 2015; 298: 491-511.

[24] Karaboğa D, Öztürk C. A novel clustering approach: artificial bee colony (ABC) algorithm. Appl Soft Comput 2011; 11: 652-657.

[25] Zhang C, Ning J, Ouyang D. An artificial bee colony approach for clustering. Expert Syst Appl 2010; 37: 4761-4767.

[26] Yan X, Zhu Y, Zou W, Wang L. A new approach for data clustering using hybrid artificial bee colony algorithm. Neurocomputing 2012; 97: 241-250.

[27] Bolaji AL, Khader AT, Albetar MA, Awadallah MA. artificial bee colony algorithm, its variants and applications: a survey. J Theor Appl Inf Technol 2013; 47: 434-459.

[28] Karaboğa D, Görkemli B, Öztürk C, Karaboğa N. A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artif Intell Rev 2014; 42: 21-57.

[29] Zhang Z. Introduction to machine learning: k-nearest neighbors. Ann Transl Med 2016; 4: 218-225.

[30] Saini I, Singh D, Khosla A. QRS detection using k-nearest neighbor algorithm (KNN) and evaluation on standard ECG databases. J Adv Res 2012; 4: 331-344.

[31] De A, Guo C. An adaptive vector quantization approach for image segmentation based on SOM network. Neurocomputing 2015; 149: 48-58.

[32] Rousi E, Anagnostopoulou C, Tolika K, Maherras P. Representing teleconnection patterns over Europe: a comparison of SOM and PCA methods. Atmos Res 2015; 152: 123-137.

[33] Korürek M, Nizam A. A new arrhythmia clustering technique based on ant colony optimization. J Biomed Inform 2008; 41: 874-881.

[34] Tomera M. Ant colony optimization algorithm applied to ship steering control. Computer Science 2014; 35: 83-92.

[35] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol 2001; 20: 45-50.

[36] Korürek M, Yüksel A, Dokur Z, Ölmez T. Dimension reduction by a novel unified scheme using divergence analysis and genetic search. Digit Signal Process 2010; 20: 1535-1546.