

Features Optimization for ECG Signals Classification

Alan S. Said Ahmad^{1,*}

Department of Physics, Faculty of Science
University of Zakho
Kurdistan Region, Iraq

Majd Salah Matti², Adel Sabry Essa³

Department of Computer Science, Faculty of Science
University of Zakho
Kurdistan Region, Iraq

Omar A.M. ALhabib⁴

College of Science
International University of Erbil
Kurdistan Region, Iraq

Sabri Shaikhow⁵

Department of Cardiology
Azadi Teaching Hospital Duhok
Kurdistan Region, Iraq

Abstract—A new method is used in this work to classify ECG beats. The new method is about using an optimization algorithm for selecting the features of each beat then classify them. For each beat, twenty-four higher order statistical features and three timing interval features are obtained. Five types of beat classes are used for classification in this work, atrial premature contractions (APC), normal (NOR), premature ventricular contractions (PVC), left bundle branch (LBBB) and right bundle branch (RBBB). Cuttlefish algorithm is used for feature selection which is a new bio-inspired optimization algorithm. Four classifiers are used within CFA, Scaled Conjugate Gradient Artificial Neural Network (SCG-ANN), K-Nearest Neighborhood (KNN), Interactive Dichotomizer 3 (ID3) and Support Vector Machine (SVM). The final results show an accuracy of 97.96% for ANN, 95.71% for KNN, 94.69% for ID3 and 93.06% for SVM, these results were tested on fourteen signal records from MIT-HIH database, where 1400 beats were extracted from these records.

Keywords—Features optimization; cuttlefish; ECG; ANN-SCG; ID3; KNN; SVM

I. INTRODUCTION

Automatic diagnosis of electrocardiogram (ECG) it is very important in the field of heart disease diagnosis, that is why feature extraction and classification it is an important step to achieve a good diagnosis [1, 2].

Many techniques have been proposed to classify ECG beat using data preprocessing, feature extraction, and classification algorithms. Some of these techniques are, Ali Kraiem and Faiza Charfi have used C4.5 technique to classify ECG beats using morphological features from signals denoised by band pass filter [3]. Ataollah Ebrahimzadeh and Ali Khazaei they have used wavelet transform and time interval features with radial base function for classification of five types of beats [4]. Ataollah Ebrahim and Ali Khazaei they have proposed a method for using morphological and time features with support vector machine for classification of 5 beat types [5]. Yakup Kutluc and Damla Kuntalp used nearest neighborhood (KNN) for classification of 5 beat types and higher order statistic

features [6]. Ali Khazaei have used genetic algorithm with radial base function for classification of 3 beat types and morphological and timing interval features [7]. Ebrahimzadeh, Shakiba and Khazaei used higher order statistics with time interval features with radial base function and bees algorithm for classification of 5 beats [1]. Raju, Rao and Jagadesh used discrete wavelet transform as features and PCA with ANN to classify 5 beat types [8]. Inbalatha and Kalaivani have used wavelet transform with principle component analysis for features and K-nearest neighborhood for classification of 2 beat types [9]. Alan and Majd have proposed a method of using higher order statistics with time intervals as features and Artificial Neural Networks to classify 5 arrhythmia beat types [10].

Another strategy has been proposed in this work. This technique is comprising of four stages. To start with, ECG flag preprocessing utilizing denoising dependent on Discrete Wavelet Transform (DWT). The database of the flag records that are utilized is MIT-BIH database [11] in which two atrial untimely constrictions (APC) records are chosen, three ordinary (NOR) records, three untimely ventricular compressions (PVC) records, three remaining group branch (LBBB) records and three right package branch (RBBB) records are utilized. Second, highlights extraction from each flag's beat and standardized for advancement and grouping, these twenty-four higher request factual and three planning interim highlights will be utilized. Third, include choice by utilizing Cuttlefish improvement calculation. Fourth, grouping utilizing Artificial Neural Network Scaled Conjugate Gradient (ANN-SCG) classifier calculation. Figure 1 delineates this work.

The remainder of this paper is as follows: section 2 contains an overview of the preprocessing technique used, Section 3 talks about the feature extraction process, and section 4 explains the usage of optimization with classification, while section 5 describes the datasets used, and sections 6 illustrated the results and discussion. Finally, section 7 describes the final conclusions.

*Corresponding Author.

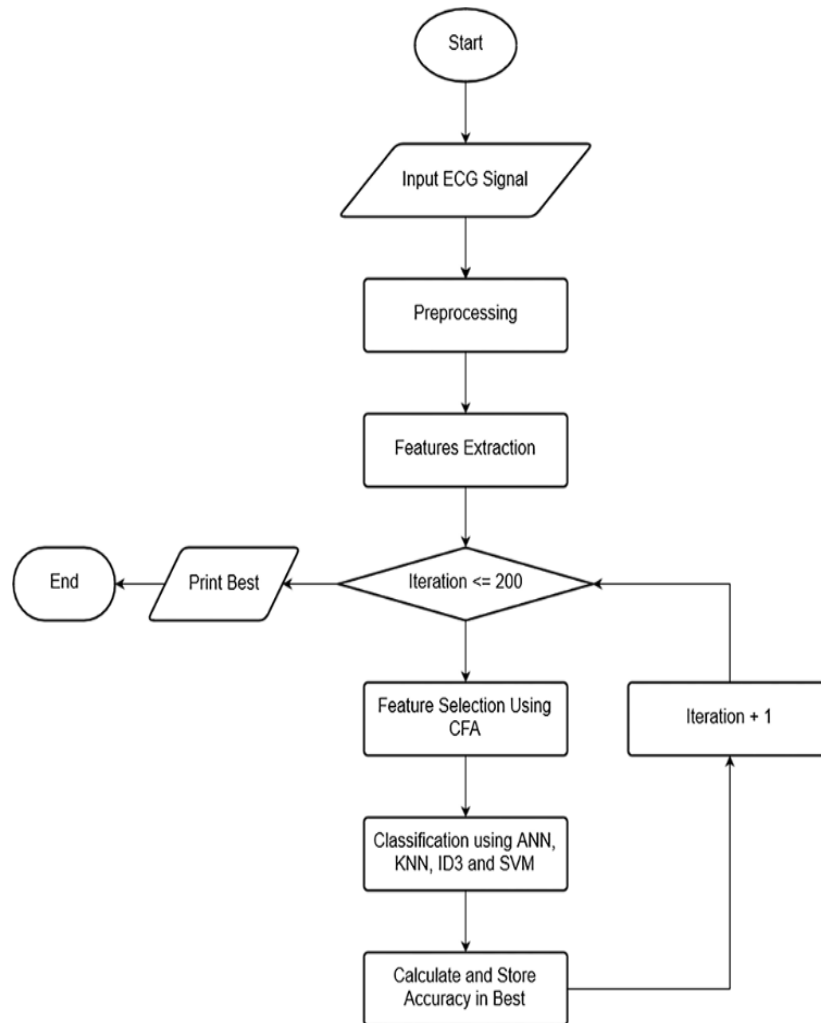


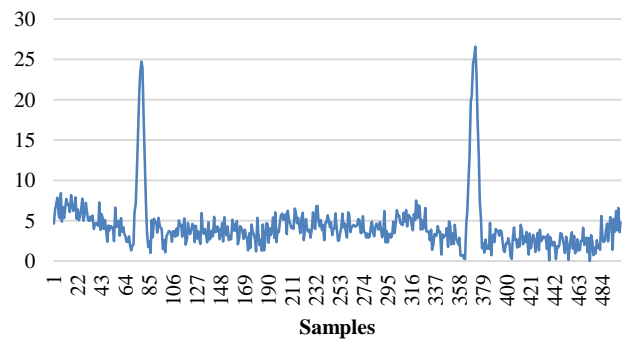
Fig. 1. The proposed system's diagram.



Fig. 2. Block Diagram of DWT.

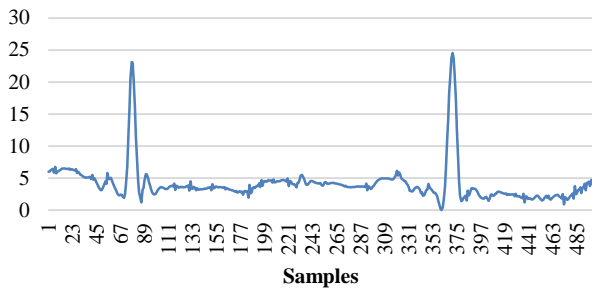
II. PREPARING SIGNALS FOR PREPROCESSING

Noise elimination from signals is very important and challenging step in signal processing [12]. Many techniques are available for noise elimination as filtering, thresholding and others. The used denoising technique in this work is wavelet shrinkage DWT method for its effective denoising results and minimum computation complexity [13]. Where denoising refers to removing noise from signal [14, 15] as in figure 3. DWT is consisting of three main steps, first signal analysis to its approximation and detailed coefficients. Second, thresholding the details coefficients. Third, synthesis of the analyzed signal to its original signal [16] as illustrated in the block diagram in figure 2.



Noisy Signal

*Corresponding Author.



Denoised Signal

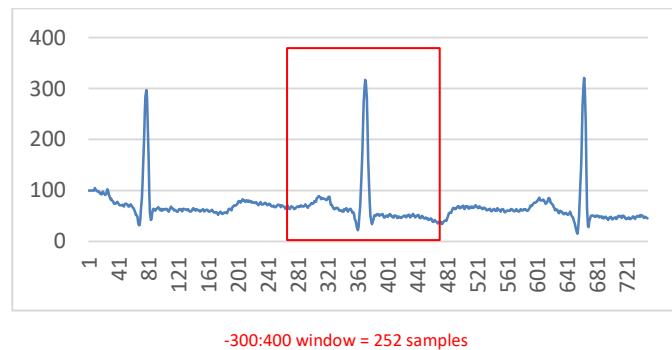
Fig. 3. The denoising process on any signal.

III. EXTRACTION OF FEATURES

A. Higher Order Statistics

Higher order statistics it's a very popular and good tool used to extract some features from nonlinear signals as ECG signal. The first and second order statistics are not sufficient for representing nonlinear signals. So the third and fourth order statistics will be used to represent each beat selected from ECG record. While variance and mean are contained in the first and second order statistics, higher order statistics contain higher order moments and cumulants [1, 6].

To extract these features, R peak must be detected at first. In any ECG signal a windows of size -300 ms to 400 ms is used which represents a 252 samples, this 252 samples will be normalized to mean of zero and standard deviation of unity. Then these 252 samples will be grouped into eight small groups starting from 30-45, 45-83, 84-112, 112-122, 122-145, 150-205, 207-225 and 230-252. Then for each small group second, third and fourth order of the cumulant is calculated. As a result we will have 24 statistical features as detailed in figure 3 [1, 6].



30-45	45-83	84-112	112-122	122-145	150-205	207-225	230-252
2nd 3rd 4th order statistics	2nd 3rd 4th order statistics	2nd 3rd 4th order statistics	2nd 3rd 4th order statistics	2nd 3rd 4th order statistics	2nd 3rd 4th order statistics	2nd 3rd 4th order statistics	2nd 3rd 4th order statistics

Fig. 4. Explanation of the statistical features.

B. Timing Features

Three timing features are extracted using equations (1), (2) and(3) for every R peak, these features are: previous time interval RR_{prev} and next time interval RR_{next} and time interval IR_i . Where IR_i is the current time

interval ratio, T_i is the current R peak, T_{i-1} is the previous R peak and T_{i+1} is the next R peak [4, 5]. Figure 4 explains the three used timing intervals. The final number of feature will be 27.

$$RR_{prev} = T_i - T_{i-1} \quad (1)$$

$$RR_{next} = T_{i+1} - T_i \quad (2)$$

$$IR_i = \frac{RR_{prev}}{RR_{next}} \quad (3)$$

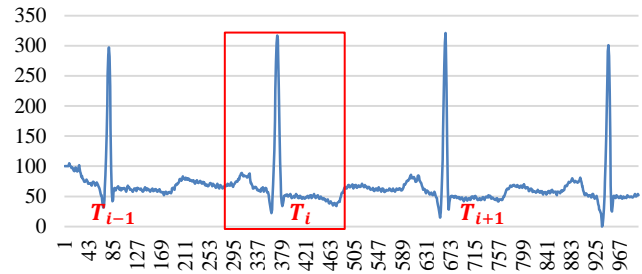


Fig. 5. Explanation of the timing features.

IV. CUTTLEFISH OPTIMIZATION ALGORITHM (CFA)

It is an optimization algorithm that is inspired by the behavior of color changing mechanism of the cuttlefish, to find the optimal solution for any problem [17, 18].

The skin of cuttlefish is comprising of three distinct layers of cells including chromatophores, leucophores and iridophores. Each layer gives distinctive shading when the light occurrence on the skin. CFA utilizes two fundamental procedures which are reflection and perceivability. The reflection procedure is the reenactment of the light reflection system, while perceivability is the recreation of perceivability of coordinating examples. These two procedures are utilized to find the worldwide ideal arrangement in the calculation. The graph in figure 6, outlining the skin structures (chromatophores, iridophores and leucophores) with two model states (upper, lower) and three unmistakable beam follows (1, 2, 3), demonstrates the modern means by which cuttlefish can change reflective shading [19].

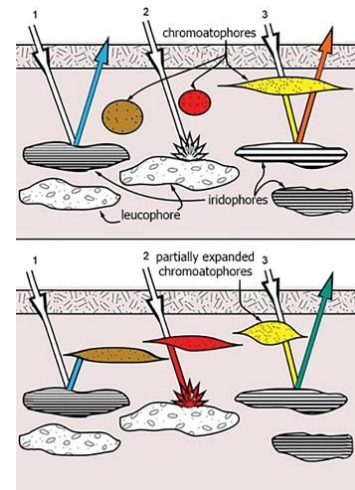


Fig. 6. The main three skin structures (chromatophores, iridophores and leucophores) of CFA.

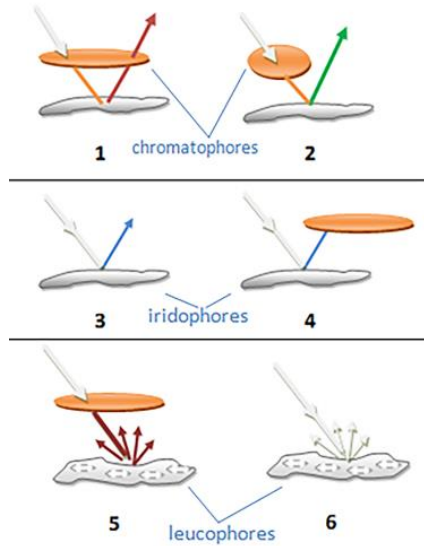


Fig. 7. Reorder of the 6 cases in figure 6.

A. Initialization

The algorithm starts by creating and initializing the population P of size N with random subsets, where each subset is consisting of two parts *selectedFeatures* and *unselectedFeatures*. Followed by the calculation of fitness value for each subset in the population. Finally, keep the best solution or (best subset) in *bestSubset* and *avBestSubset*, then remove one feature from *selectedFeatures* in *bestSubset*. After initialization the CFA will perform its main function which consists of six cases, these cases are illustrated in figure 7, and in the next steps:

B. Group 1, Case 1 and 2

In the first group which include cases 1 and 2, the algorithm will start by descending sort of the population P based on fitness value for each subset in the population. Part of the population will be selected for applying CFA equations starting at $i = 1, 2, \dots, K$. Where K is a random number selected between 0 and $N/2$.

New subset for each i will be generated by combining Reflection (R) and Visibility (V) as in equation (4), where R is the degree of reflection and V is the visibility degree of the final view of the matched pattern. The main reflection and visibility equations of the algorithm for the current cases are as follows:

$$newSubset_i = Reflection_i \oplus Visibility_i \quad (4)$$

$$Reflection_i = random(subset[R]) \subset P_i(selectedFeatures) \quad (5)$$

$$Visibility_i = random(subset[V]) \subset P_i(unselectedFeatures) \quad (6)$$

where $Reflection_i$ is a subset with size of R and its elements are selected randomly from *selectedFeatures* and $Visibility_i$ is another subset with size of V whose elements are created randomly from *unselectedFeatures*. Where R and V values can be calculated as follows:

$$R = random(0, size\ of\ selectedFeatures)$$

$$V = size\ of\ selectedFeatures - R$$

C. Group 2, Case 3 and 4

In the second group which include cases 3 and 4, to calculate the Reflection we have only to remove one element from *selectedFeatures* of *bestSubset* as in equation (6) where the visibility can be calculated by selecting only one feature from *unselectedFeatures* of *bestSubset* as in equation (7) and finally by combining the results of these two equations a new subset will be produced.

These two equations will be repeated T times, where T is a small fixed number selected from the size of *selectedFeatures* of *bestSubset*. If *newSubset* is better than *bestSubset* then replace *newSubset* with *bestSubset*.

$$Reflection = bestSubset(selectedFeatures) - bestSubset(selectedFeatures)[R] \quad (6)$$

$$Visibility = bestSubset(unselectedFeatures)[V] \quad (7)$$

where R is the index of the removed feature from *selectedFeatures* of *bestSubset*, and V is the index of selected feature from *unselectedFeatures* of *bestSubset*. The calculation of R and V are as follows:

$$R = random(0, size(bestSubset(selectedFeatures)))$$

$$V = random(0, size(bestSubset(unselectedFeatures)))$$

D. Group 3, Case 5

In the third group which is about case 5, the new subset will be generated by the combination of the equations (9) and (10) into equation (8). These equations will be repeated to size of *selectedFeatures* of *avBestSubset* times. The values of reflection are equals to *selectedFeatures* in *avBestSubset* only and the value of visibility will be a single value selected from *selectedFeatures* in *avBestSubset* according to index i .

In this way, we can produce R new subsets, the value of R is equal to the size of *selectedFeatures* each subset representing the matched pattern by removing one feature from *selectedFeatures* at each time. If *newSubset* is better than *bestSubset* then replace *newSubset* with *bestSubset*.

$$newSubset_i = Reflection - Visibility_i \quad (8)$$

$$Reflection = AVbestSubset - selectedFeatures \quad (9)$$

$$Visibility_i = AVbestSubset - selectedFeatures[i] \quad (10)$$

where, i is the index of the feature in *selectedFeatures* that will be removed. $i = 1, 2, \dots, R$. Where R is the size of *selectedFeatures*.

E. Group 4, Case 6

In the fourth gathering which is case 6, any approaching shading from the earth will be reflected as it very well may be spoken to by any irregular arrangement. In the underlying calculation, this case is utilized to produce irregular arrangements. Likewise, we utilize this case as an irregular generator procedure to create arbitrary arrangements. The quantity of ages is equivalent to $N-K$.

The new age will be begun at the area K in the wake of arranging the populace P in diving request. On the off chance

*Corresponding Author.

that the newSubset is superior to the current avBestSubset, the current avBestSubset is supplanted with newSubset. The procedure of irregular arrangements is the equivalent as that utilized with the introduction procedure which is depicted previously.

F. Stopping Condition

The algorithm will stop when the iterations number reaches its max limit which is 50 iterations.

G. Fitness Function

CFA needs to evaluate every subset of the population and assign a fitness value to it. In this work four different learner algorithms or classifiers are used where each is used separately with the CFA.

- ANN

Artificial Neural Network it is a very popular technique in the field of classification and machine learning, because its performance is very good [20]. The general structure of ANN is composed three main layers one input layer, one or more hidden layers and one output layer. In this work the used structure of the network is one input layer with N input neurons because the CFA will control the number of input layer, for the hidden layer we have two layers each one is of 40 neurons and one output layer with 5 neurons due to the five classes we have, as in figure 8. Scaled Conjugate Gradient classifier algorithm is used which is a supervised classification algorithm. The advantages of this classifier are its fast speed and it does not require a lot of memory [21, 22].

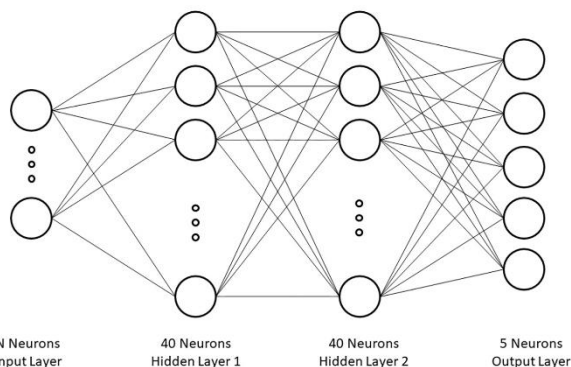


Fig. 8. The structure of the used neural network.

- KNN

The K-Nearest Neighbor is statistical classification algorithms based on closest training examples in the feature space [23]. The KNN algorithm has few parameters to set for classification [23]. These parameters are K and Distance metric. Where of K is a value which is selected in a way that it gives maximum classification accuracy [23]. Whereas the Distance metric parameter used in this work is Minkowski.

The Minkowski provides a concise, parametric distance function that generalizes many of the distance functions available as Euclidean, Manhattan, Chebyshev, ... etc. The advantage is that mathematical results can be shown for a

*Corresponding Author, whole class of distance functions, and the user can adapt the distance function to suit

the needs of the application by modifying the Minkowski parameter [24].

- ID3

The Interactive Dichotomizer 3 (ID3) it is a decision tree learner algorithm. Decision Trees (DTs) are accurate and small, in which results in a reliable and fast classification results. Because of its speed and reliability, decision trees are popular classification tool [25]. DTs are used to build classification rules in the form of top down Decision Tree [26]. Where the leaf contains class names and non-leaves are decision nodes [26].

- SVM

Support Vector Machines in general are binary classifiers, but you may want to classify your data into more than two classes [27]. To solve this problem multiclass SVM is needed which is done by using the One Against One(OAO) approach.

The OAO strategy, otherwise called (pairwise coupling), (all sets) or (round robin), comprises in building one SVM for each combine of classes. Hence, for an issue with n classes, n(n-1)/2 SVMs are prepared to recognize the examples of one class from the examples of another class. Grouping of an obscure example is finished by the greatest casting a ballot, where each SVM votes in favor of one class [27].

V. DATASETS AND EVALUATION METERING

A. Datasets Used

MIT-BIH database [11] is used to use some ECG data records. In table 1 you have 5 classes in the first column, each class represents a specific arrhythmia disease. Eight records have were selected from the database and distributed on each class [1] as in table 1. From each record signal used 65 beats were selected for training and 35 for testing, more details are in table 1.

TABLE I. SUMMERY OF THE DATA USED AND ITS DIVISIONS

Class name	Selected records	Used beats in training	Used beats in testing
NOR	100m, 105m, 215m	65, 65, 65	35, 35, 35
PVC	207m, 209m, 232m	65, 65, 65	35, 35, 35
APC	106m, 223m	65, 65	35, 35
LBBB	109m, 111m, 214m	65, 65, 65	35, 35, 35
RBBB	118m, 124m, 212m	65, 65, 65	35, 35, 35
		910 beat	490 beat

B. Evaluation Criteria

For this work four equations have been used to evaluate the performance:

- The sensitivity:

$$Se = \frac{T_P}{T_P + F_N} \times 100\% \quad (11)$$

- The specificity:

$$Sp = \frac{T_N}{T_N + F_P} \times 100\% \quad (12)$$

- The positive predictivity:

$$Pp = \frac{T_P}{T_P + F_P} \times 100\% \quad (13)$$

- The accuracy: It is very important for testing the performance [2].

$$Acc = \frac{T_P}{Total\ number\ of\ beats} \times 100\% \quad (14)$$

Where T_P (true positive) represents the correct classified beats number for any class, F_N (false negative) represents the incorrect classified beats number in the other used classes, T_N (true negative) represents the correct classified beats number for all other classes and F_P (false positive) represents the incorrect classified beats number for any class.

VI. DISCUSSION AND RESULTS

The final results of this work will be illustrated in this section. As a total number of beats 1400 beat is selected for training and testing the proposed system. Fourteen ECG records were used and distributed on five classes as in table 1, 100 beat is extracted from 106 and 223 records for APC, a total of 300 beats were selected from NOR records (100, 105 and 215) 100 from each, for PVC records (207, 209 and 232) 100 beat is selected from each record, for LBBB records (109, 111 and 214) 100 beat is selected from each record, and for RBBB records (118, 124 and 212) 100 beat is selected from each record.

In this work CFA is used optimization for features and ANN classifier is used to classify features into five classes. Figure 9 gives a good vision of the performance of the proposed method in this work for 50 iterations of CFA and Figure 10 gives the comparison with the works of others.

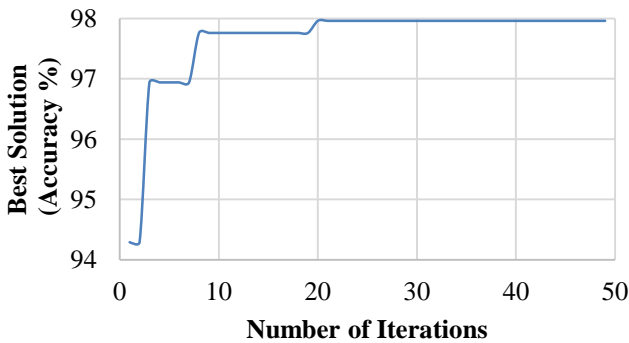


Fig. 9. The progress of best solution in every iteration.

As illustrated in figure 9 the final accuracy was 97.96% after training, which is a very accurate and high result when this result is compared with others in the same field. Also the third column in table 3 explains the accuracy of all class types used, 97.14% for APC, 98.10% for LBBB, 97.14% for NOR, 98.10% for PVC and 99.05% for RBBB.

*Corresponding Author.

Table 2 gives the complete statistics of

the total number of beats used for these variables : true positive (T_P), false negative (F_N), true negative (T_N) and false positive (F_P).

TABLE II. FULL DETAILS ABOUT THE NUMBER OF BEATS USED AND THE CORRECTLY AND INCORRECTLY CLASSIFIED BEATS FOR EACH CLASS

	Total Beats	TP	TN	FP	FN
APC	70	68	412	2	8
LBBB	105	103	377	2	8
NOR	105	102	378	3	7
PVC	105	103	377	2	8
RBBB	105	104	376	1	9

The sensitivity (Se), specificity (Sp) and positive predictivity (Pp) for each class type is detailed in table 3.

TABLE III. THE CLASSES USED AND THE SENSITIVITY, SPECIFICITY AND ACCURACY FOR EACH CLASS

Class Type	Se %	Sp %	Pp (Acc) %
APC	89.47	99.52	97.14
LBBB	92.79	99.47	98.10
NOR	93.58	99.21	97.14
PVC	92.79	99.47	98.10
RBBB	92.04	99.73	99.05
Overall	92.13	99.48	97.96

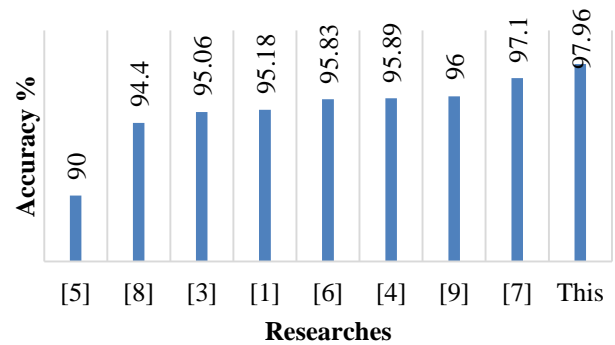


Fig. 10. Comparison with the work of other researchers.

As stated before, four classifiers have been used in this work ANN, KNN, ID3 and SVM each gave different accuracy after training phase as listed in table 4.

TABLE IV. THE COMPARISON OF ACCURACY OBTAINED FROM THE FOUR USED CLASSIFIERS

ANN	KNN	ID3	SVM
97.96 %	95.71 %	94.69 %	93.06 %

VII. FINAL CONCLUSIONS

This work is about using a hybrid technique for classifying ECG signals into five classes, using 27 features (24 statistical with 3-time interval features). Fourteen signals were selected from MIT-BIH database [11], the records were selected according to five ECG arrhythmia classes. Four classifiers are

used to show a comparison of each one, ANN-SCG, KNN, ID3 and SVM are used as classifiers.

1) From the experiments of this work it is shown that the usage of CFA optimization algorithm with ANN-SCG learner algorithm gives best classification results than ID3, KNN and SVM.

2) The comparison is done with different works, each work used different technique but ANN-SCG with CFA is the best choice.

REFERENCES

- [1] Ebrahimzadeh, A., B. Shakiba, and A. Khazaei, Detection of electrocardiogram signals using an efficient method. *Applied Soft Computing*, 2014. 22: p. 108-117.
- [2] Li, H., et al., Genetic algorithm for the optimization of features and neural networks in ECG signals classification. *Scientific Reports*, 2017. 7: p. 41011.
- [3] Kraiem, A. and F. Charfi, Arrhythmia Classification from ECG Signals Using Data Mining Approaches. 2009: p. 7.
- [4] Zadeh, A.E. and A. Khazaei, An Efficient Technique for Classification of Electrocardiogram Signals. *Advances in Electrical and Computer Engineering*, 2009. 9(3): p. 5.
- [5] Zadeh, A.E. and A. Khazaei, High Efficient System for Automatic Classification of the Electrocardiogram Beats. *Annals of Biomedical Engineering*, 2011. 39(3): p. 16.
- [6] Kutlu, Y. and D. Kuntalp, Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients. *Computer Methods and Programs in Biomedicine*, 2012. 105(3): p. 257-267.
- [7] Khazaei, A., Automated Cardiac Beat Classification Using RBF Neural Networks. *IJ.Modern Education and Computer Science*, 2013. 5: p. 7.
- [8] Raju, N.N.S.V.R., V.M. Rao, and B.N. Jagadesh, Identification and Classification of Cardiac Arrhythmia using Neural Network. *Helix*, 2017. 7(5): p. 6.
- [9] Mandala, S., et al. Performance analysis of wavelet-based denoising techniques for ECG signal. in *2017 5th International Conference on Information and Communication Technology (ICoICT)*. 2017.
- [10] Ahmad, A.S.S., et al., ECG Signal Classification Using Scaled Conjugate Gradient Learner Algorithm. *International Journal of Medical Research & Health Sciences*, 2018. 7(5): p. 7.
- [11] Goldberger, A.L., et al., PhysioBank, PhysioToolkit, and PhysioNet. Components of a New Research Resource for Complex Physiologic Signals, 2000. 101(23): p. 215-220.
- [12] Alyasseri, Z.A.A., et al., Hybridizing B-hill climbing with wavelet transform for denoising ECG signals. *Inf. Sci.*, 2018. 429(C): p. 229-246.
- [13] Srivastava, M., C.L. Anderson, and J.H. Freed, A New Wavelet Denoising Method for Selecting Decomposition Levels and Noise Thresholds. *IEEE Access*, 2016. 4: p. 3862-3877.
- [14] Ahmad, A.S.S., et al., Denoising of Arrhythmia ECG Signals. *International Journal of Medical Research & Health Sciences*, 2018. 7(3): p. 11.
- [15] Sadooghi, M.S. and S.E. Khadem, A New Performance Evaluation Scheme for Jet Engine Vibration Signal Denoising. *Mechanical Systems and Signal Processing*, 2016. 76(77): p. 201-212.
- [16] Donoho, D.L. and I.M. Johnstone, Adapting to Unknown Smoothness via Wavelet Shrinkage. *Journal of the American Statistical Association*, 1995. 90(432): p. 1200-1224.
- [17] Eesa, A.S., A.M.A. Brifciani, and Z. Orman, A New Tool for Global Optimization Problems - Cuttlefish Algorithm. *International Journal of Computer and Information Engineering*, 2014. 8(9): p. 5.
- [18] Eesa, A.S., Z. Orman, and A.M.A. Brifciani, A Novel Feature Selection Approach Based on The Cuttlefish Optimization. *Expert Systems with Applications*, 2014. 42: p. 10.
- [19] Kreit, E., et al., Biological versus electronic adaptive coloration: how can one inform the other? *Journal of The Royal Society Interface*, 2012.
- [20] AlHamouz, S. and A. Abu-Shareha, Hybrid Classification Approach Using Self-Organizing Map and Back Propagation Artificial Neural Networks for Intrusion Detection. in *2017 10th International Conference on Developments in eSystems Engineering (DeSE)*. 2017.
- [21] Khadse, C.B., M.A. Chaudhari, and V.B. Borghate, Comparison between three back-propagation algorithms for power quality monitoring. in *2015 Annual IEEE India Conference (INDICON)*. 2015.
- [22] Mohamad, N., et al. Comparison between Levenberg-Marquardt and Scaled Conjugate Gradient training algorithms for Breast Cancer Diagnosis using MLP. in *2010 6th International Colloquium on Signal Processing & its Applications*. 2010.
- [23] Saini, I., D. Singh, and A. Khosla, Delineation of ECG Wave Components Using K-Nearest Neighbor (KNN) Algorithm: ECG Wave Delineation Using KNN. in *2013 10th International Conference on Information Technology: New Generations*. 2013.
- [24] Chouikhi, H., M.F. Saad, and A.M. Alimi, Improved fuzzy possibilistic C-means (IFPCM) algorithms using Minkowski distance. in *2017 International Conference on Control, Automation and Diagnosis (ICCAD)*. 2017.
- [25] Salzberg, S.L., C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Machine Learning*, 1994. 16(3): p. 235-240.
- [26] Eesa, A.S., Z. Orman, and A.M.A. Brifciani, A new feature selection model based on ID3 and bees algorithm for intrusion detection system. *Turkish Journal of Electrical Engineering and Computer Science*, 2015. 23(2): p. 8.
- [27] Faziludeen, S. and P.V. Sabiq, ECG beat classification using wavelets and SVM. in *2013 IEEE Conference on Information & Communication Technologies*. 2013.

*Corresponding Author.