Available online at **www.ijmrhs.com**

ISSN No: 2319-5886

International Journal of Medical Research & Health Sciences, 2018, 7(5): 11-17

ECG Signal Classification Using Scaled Conjugate Gradient Learner Algorithm

Alan S. Said Ahmad[1]*, Majd S. Matti[1], Omar A. M Al-Habib[1,2] and Sabri Shaikhow[3]

[1] *Faculty of Science, University of Zakho, Kurdistan Region, Iraq*

[2] *College of Science, International University of Erbil, Kurdistan Region, Iraq*

[3] *Department of Cardiology, Azadi Teaching Hospital, Duhok, Kurdistan Region, Iraq*

**Corresponding e-mail: alan.ahmad@uoz.edu.krd*

ABSTRACT

This study is based on classifying the ECG signal into five types of classes by using statistical and timing intervals features. First, the data signals were denoised and prepared for classification. Second, 24 higher order statistical features with 3 timing interval features were extracted from each selected beat. In this work, we have 5 types of classes, atrial premature contractions (APC), normal (NOR), premature ventricular contractions (PVC), left bundle branch block (LBBB) and right bundle branch block (RBBB) were used for classification. Third, each beat was classified according to one of these classes by using the learner algorithm scaled conjugate gradient (SCG) artificial neural network (ANN). SCG is a fast algorithm and suitable in cases of less memory and ANN is a machine learning algorithm that is based on the biological neural system. The experimental results of this work show an accuracy of 96% on 1400 beats taken from 14 records from MIT/BIH arrhythmia database.

Keywords: ECG, Higher order statistics, Feature extraction, Classification, ANN, Scaled conjugate gradient

INTRODUCTION

Electrocardiogram (ECG) is the graphical representation of human heart's electrical activity [1]. ECG signal analysis and classiication is a very important issue for heart diseases diagnosis [2]. In the available literature, there are several methods and techniques used for the classification of ECG signal based on statistical, morphological, and timing intervals features as Ebrahimzadeh, et al., have used wavelet approximation with three timing intervals as features and radial base function (RBF) neural network for classification [3]. Kraiem, et al., have used some morphological features and decision trees as C4.5 and CHAID for classification [4]. Zadeh, et al., used morphological and timing features with support vector machine (SVM) for classification [5]. Kutlu used nearest neighborhood (KNN) for classification and higher order statistics as features [6]. Khazaee, et al., have proposed a method of using genetic algorithm (GA) with RBF for classification and morphological and timing intervals as features [7]. Ebrahimzadeh, et al., in 2014 used higher order statistics with three timing intervals for features and a hybrid algorithm of RBF and Bees algorithm (BA) for classification [8]. Kulkarni, et al., used some timing and statistical features with KNN as classification algorithm [9]. Inbalatha, et al., used discrete wavelet transform (DWT) and principle component analysis (PCA) as features and KNN for classification algorithm [10].

In this work, we have proposed a method for ECG beat classification. This method consists of three steps. First is data pre-processing of the ECG signals by denoising them using discrete wavelet transformation. The used data is from MIT-BIH database, two records for atrial premature contractions (APC), three records for normal (NOR), three records for premature ventricular contractions (PVC), three records for left bundle branch block (LBBB) and three records for right bundle branch block (RBBB) are used [11]. Second, some features are extracted and normalized to be prepared for classification. These features are twenty-four higher order statistical features and three timing interval features have been used. Third, the classification was using artificial neural network using scaled conjugate gradient learner algorithm. The block diagram of this work is illustrated in Figure 1.

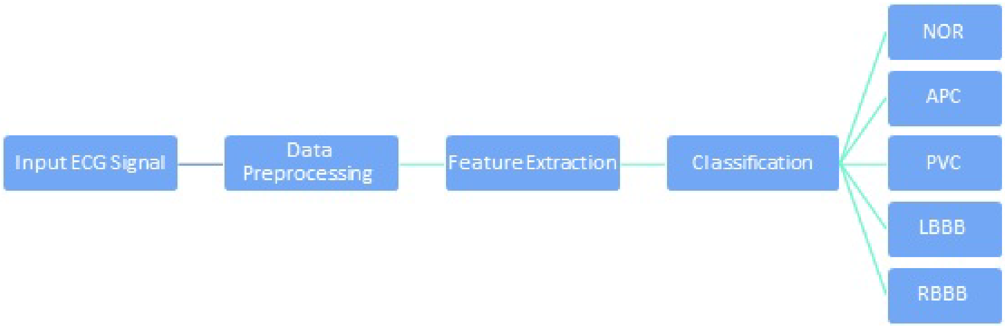

Figure 1 Block diagram of the proposed system

MATERIALS AND METHODS

Data Pre-processing

Noise elimination is a very important step in ECG signals pre-processing. Such noises are electrical activity of muscles (EMG) and instability of electrode-skin contact affect the process of clean data extraction [12]. To overcome the problem of noise, DWT technique is used for this purpose [13]. DWT mainly consists of two steps; first is decomposition of the input signal into approximation coefficients and detailed coefficients. Second, reconstruction of the decomposed signal back to its origin as in Figure 2 [13].

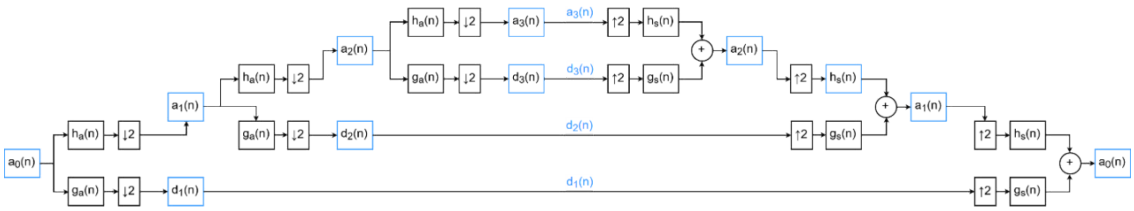

Figure 2 Block diagram of decomposition and reconstruction

In Figure 2, $h_a(n)$ and $g_a(n)$ are the decomposition low pass filter and high pass filter respectively, $h_s(n)$ and $g_s(n)$ are the reconstruction low pass filter and high pass filter respectively [8]. Only 3 levels of decomposition are illustrated in Figure 2, but for denoising 7 levels of decomposition are used with the Daubechies db5 wavelet filter.

Feature Extraction

Feature extraction is an important process before detection or classification, because these features will identity for that thing to be classified or detected. In this work, two types of features are used, viz; higher order statistics features and timing feature.

Higher order statistics features

Higher order statistics have achieved great importance in the ield of bio-signal processing like ECG signal which is a non-linear signal. First two order statistics are not sufficient for representing non-linear signals. Hence third and fourth order statistics are used in this analysis. While the first and second order statistics contain mean and variance, higher order statistics contain higher order moments and non-linear combinations of higher order moments which are known as cumulants [6,8].

The extraction of these features is usually started by detecting the R peak in the ECG signal by using a window of -300 m to the left of the R peak and 400 m to the right of the R peak which forms 252 samples representing a whole heartbeat. The 252 samples will undergo data normalization of mean of zero and standard deviation of unity to reduce the DC offset and eliminates the amplitude variance. After that these 252 samples will be divided into eight sections as 30-45, 45-83, 84-112, 112-122, 122-145, 150-205, 207-225 and 230-252. Then for each section second, third and fourth order of the cumulant is calculated. Thus, the total number of the statistical features is equal to 8 × 3=24 as in Figure 3 [6,8].

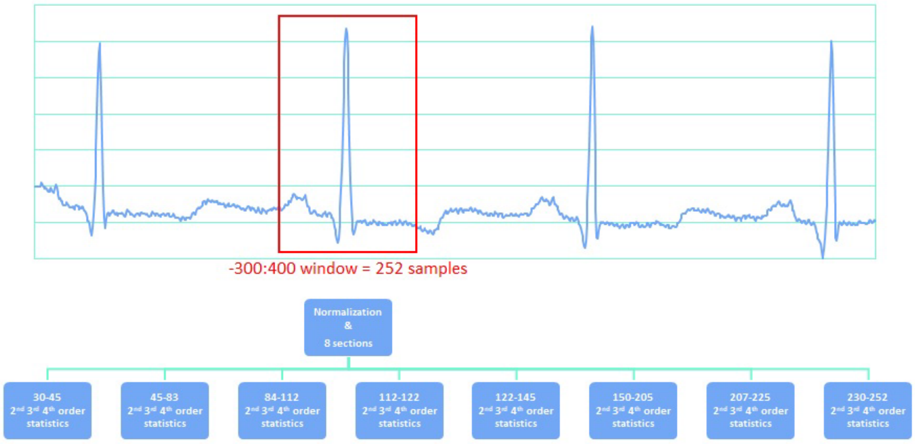

Figure 3 Block diagram explaining statistical features

Timing feature

In addition to the 24 statistical features, 3 timing features are calculated for each heartbeat as the next time interval $T_{(i+1)}$ to T_i as in equation 1 and previous time interval $T_{(i-1)}$ to T_i in equation 2, and time interval ratio (IR) as in equation 3, 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 [3,5]. The three timing intervals are explained in Figure 4. The final length of the feature vectors is 27.

$$RR_{prev} = T_i - T_{i-1} \tag{1}$$

$$RR_{next} = T_{i+1} - T_i \tag{2}$$

$$IR_i = \frac{RR_{prev}}{RR_{next}} \tag{3}$$

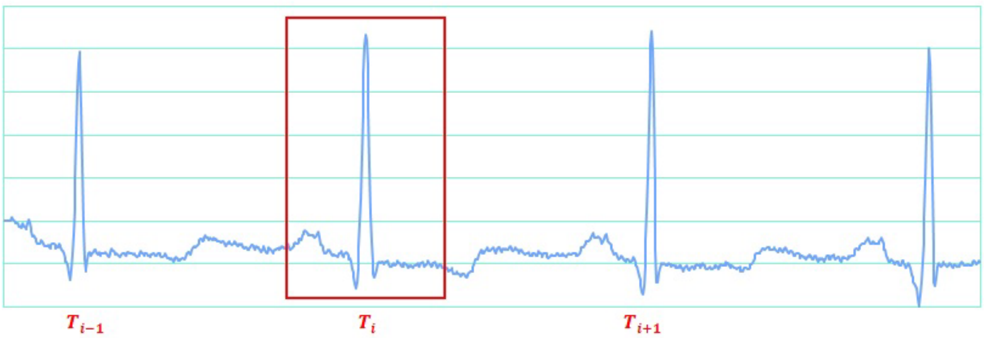

Figure 4 Block diagram explaining timing features

Classification

Artificial neural network (ANN) is a machine learning algorithm which is widely applied in the application fields of classification, because it has proven performance [14]. In this work, the structure of the used network is composed of 27 input neurons representing the 27 features, two hidden layers each of 40 neurons and 5 output neurons because we have used 5 classes, as illustrated in Figure 5. The used learner is scaled conjugate gradient algorithm which is a supervised learning algorithm, which can solve problems effectively because its fast and uses less memory [15,16].

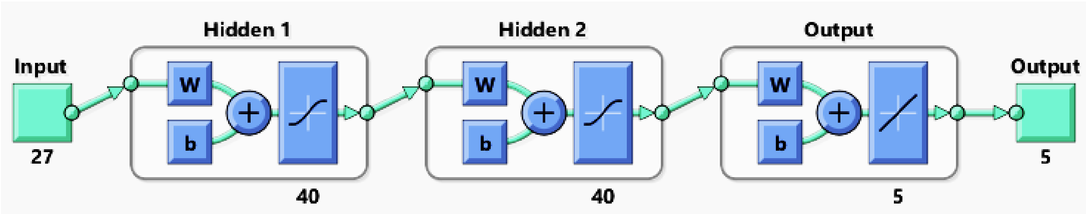

Figure 5 The structure of the used neural network

Datasets and Evaluation Metering

Datasets used

Data records used in this research is from MIT-BIH database [11], as illustrated in Table 1 there are 5 classes each class represents a specific abnormal change in the heart rhythm. Eight records have been used as a total and for each class some records were chosen to represent that class as in Table 1 [8]. For the used records, every signal 50 random beats were used for training and another random 50 beats were selected for testing, as in Table 1.

Table 1 The number of beats used for each signal

Class	Records	Number of beats used for	
		Training	Testing
NOR	100 m, 105 m, 215 m	50, 50, 50	50, 50, 50
PVC	207 m, 209 m, 232 m	50, 50, 50	50, 50, 50
APC	106 m, 223 m	50, 50	50, 50
LBBB	109 m, 111 m, 214 m	50, 50, 50	50, 50, 50
RBBB	118 m, 124 m, 212 m	50, 50, 50	50, 50, 50
		700 beat	700 beat

Evaluation metering

Four evaluation metering methods were used in this work as the classification accuracy (Acc) which is very important to show the performance of the work [17], sensitivity (Se), specificity (Sp) and positive predictivity (Pp) were used for analysis and evaluation of the proposed system.

$$Se = \frac{T_p}{T_p - F_N} \times 100\% \tag{4}$$

Where T_P (true positive) is the number of correctly classified beats of any class, and F_N (false negative) is the number of incorrectly classified beats in the all other classes.

$$Sp = \frac{T_N}{T_N - F_P} \times 100\% \tag{5}$$

Where T_N (true negative) is the number of correctly classified beats of all other classes, and F_P (false positive) is the number of incorrectly classified beats of a specific class.

$$Pp = \frac{T_p}{T_p - F_p} \times 100\% \tag{6}$$

$$Acc = \frac{N_T - N_E}{N_T} \times 100\% \tag{7}$$

Where N_E represents the total number of incorrectly classified beats, and N_T is the total number of beats.

RESULTS AND DISCUSSION

In the experimental results in this work, 1400 beat as a total were selected to test the proposed system. As in Table 1 total of 14 ECG records were used and distributed on 5 classed, 106 and 223 for APC where each record 100 beat is selected from (100, 105 and 215) records used for NOR with a total of 300 beats were selected 100 from each (207, 209 and 232) records were used for PVC with 100 beat from each record, for LBBB (109, 111 and 214) records were used with a total of 300 selected beats, and (118, 124 and 212) records for RBBB were used with 100 beat is selected from each record.

We have tested the learner algorithm many times, and in every time different beats were selected for training and testing of a different number of hidden layers with different number of neurons were tried too. Finally, the number of hidden layers was fixed at two with 40 neurons each as in Figure 5.

In the training phase of the ANN, the measuring criteria used is MSE. The learner algorithm SCG runs 1000 epochs then finally gave us 0.00241 as the minimum MSE. The progress of MSE with respect to the number of epochs is shown in Figure 6.

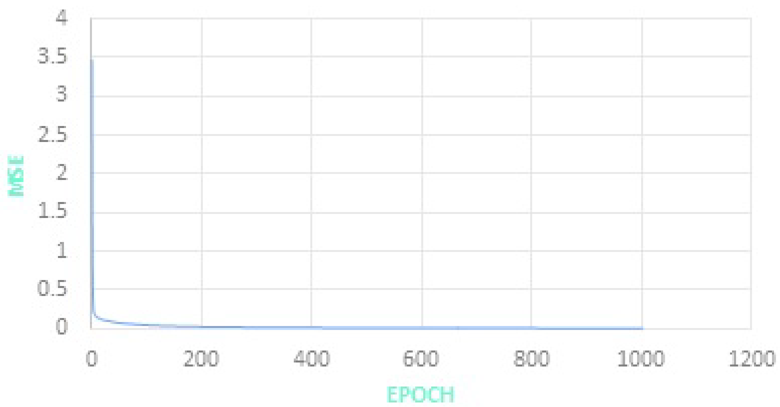

Figure 6 The MSE progress in every epoch

Table 2 describes and gives the number of beats for each of the variables stated in the evaluation metrics as T_P (true positive), F_N (false negative), T_N (true negative) and F_P (false positive). The sensitivity (Se), specificity (Sp) and positive predictivity (Pp) are given for each class type in details, as listed in Table 3.

Table 2 The total, true positive, true negative, false positive and false negative beats for each class type

	Total	T_P	T_N	F_P	F_N
APC	100	100	572	0	28
LBBB	150	148	524	2	26
NOR	150	143	529	7	21
PVC	150	134	538	16	12
RBBB	150	147	525	3	25

As shown in Figure 7 the accuracy of this work after training phase was 96% which is a very good result compared with the other research stated in the literature in the field of ECG signal classification. In addition, Table 4 gives a detailed description about the accuracy of every class type, 100% detection rate for APC, 98.7% for LBBB, 95.3% for NOR, 89.3% for PVC and 98% for RBBB.

Table 3 The sensitivity, specificity and positive predictivity for each class

Class	Se%	Sp%	Pp%
APC	78.13%	100.00%	100.00%
LBBB	85.06%	99.62%	98.67%
NOR	87.20%	98.69%	95.33%
PVC	91.78%	97.11%	89.33%
RBBB	85.47%	99.43%	98.00%
Average	85.52%	98.97%	96.27%

Table 4 The classification accuracy of each signal

Class	Acc%
APC	100.00%
LBBB	98.70%
NOR	95.30%
PVC	89.30%
RBBB	98.00%
Overall Accuracy = 96.0%	

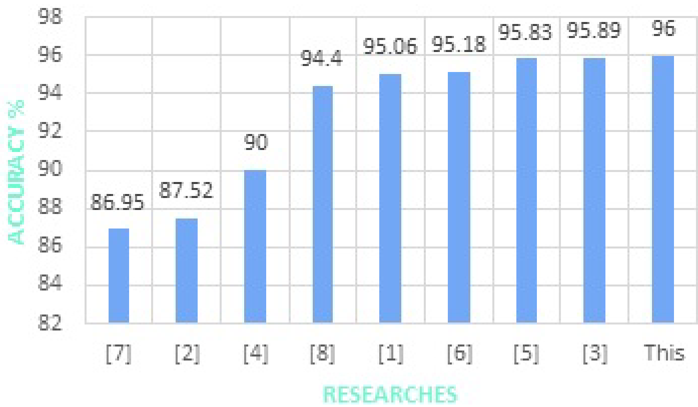

Figure 7 Comparison with the work of others

CONCLUSION

This work is about classifying ECG signals based on 5 classes and using 27 features. There were 24 statistical features and 3 timing intervals were extracted. Five classes each represent a special abnormal changes ECG signal record. Records were selected from MIT-BIH database [11]. Artificial neural network using scaled conjugate gradient learner algorithm is used as a classiier.

- From the experiments of this work, it was clear that the learner algorithm is fast and gives good results.
- In the comparison with the other researchers in the field of classification some researchers have used SVM, others used KNN and others used ANN each with different learner algorithm but ANN with scaled conjugate gradient learner algorithm is the best choice.

DECLARATIONS

Conflict of Interest

The authors have disclosed no conflict of interest, financial or otherwise.

REFERENCES

[1] Alyasseri, Zaid Abdi Alkareem, et al. “Hybridizing β-hill climbing with wavelet transform for denoising ECG signals.” *Information Sciences,* Vol. 429, 2018, pp. 229-46.

[2] Zhang, Jin, et al. “ECG signals denoising method based on improved wavelet threshold algorithm.” *Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*. IEEE, 2016.

[3] Ebrahimzadeh, Ataollah, and Ali Khazaee. “An efficient technique for classification of electrocardiogram signals.” *Advances in Electrical and Computer Engineering,* Vol. 9, No. 3, 2009, pp. 89-93.

[4] Kraiem, Ali, and Faiza Charfi. “Arrhythmia Classification from ECG signals using Data Mining Approaches.” 2009, p. 7.

[5] Zadeh, Ataollah Ebrahim, and Ali Khazaee. “High efficient system for automatic classification of the electrocardiogram beats.” *Annals of Biomedical Engineering,* Vol. 39, No. 3, 2011, pp. 996-11.

[6] Kutlu, Yakup, and Damla Kuntalp. “Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients.” *Computer Methods and Programs in Biomedicine,* Vol. 105, No. 3, 2012, pp. 257-67.

[7] Khazaee, Ali. "Automated cardiac beat classification using RBF neural networks." *International Journal of Modern Education and Computer Science,* Vol. 5, No. 3, 2013, p. 42.

[8] Ebrahimzadeh, Ataollah, B. Shakiba, and Ali Khazaee. "Detection of electrocardiogram signals using an efficient method." *Applied Soft Computing,* Vol. 22, 2014, pp. 108-17.

[9] Kulkarni, A., S. Lale, P. Ingole, et al. "*Analysis of ECG Signals." SSRG International Journal of Electronics and Communication Engineering*, Vol. 3, No. 4, 2016, p. 3.

[10] Inbalatha, M, Kalaivani, S. "Performance appraisal of dwt and pca based cardiac ecg arrhythmias diagnosis with K-NN classifier". *International Journal of Electrical, Electronics and Data Communication*, Vol. 5, No. 2, 2017, p. 7.

[11] Golberger, A., et al. "PhysioBank, PhysioToolkit, and PhysioNet: Component of a New Research Resource for Complex Physiologic Signals." *Circulation,* Vol. 101, No. 23, 2000, pp. 215-20.

[12] Šarlija, Marko, Fran Jurišić, and Siniša Popović. "A convolutional neural network based approach to QRS detection." *Image and Signal Processing and Analysis (ISPA), 2017 10th International Symposium on*. IEEE, 2017.

[13] Donoho, David L., and Iain M. Johnstone. "Adapting to unknown smoothness via wavelet shrinkage." *Journal of the American Statistical Association,* Vol. 90, No. 432, 1995, pp. 1200-24.

[14] AlHamouz, Sadeq, and Ahmad Abu-Shareha. "Hybrid Classification Approach Using Self-Organizing Map and Back Propagation Artificial Neural Networks for Intrusion Detection." *Developments in eSystems Engineering (DeSE), 2017 10th International Conference on*. IEEE, 2017.

[15] Mohamad, Nadiah, et al. "Comparison between Levenberg-Marquardt and scaled conjugate gradient training algorithms for breast cancer diagnosis using MLP." *Signal Processing and its Applications (CSPA), 2010 6th International Colloquium on*. IEEE, 2010.

[16] Khadse, Chetan B., Madhuri A. Chaudhari, and Vijay B. Borghate. "Comparison between three back-propagation algorithms for power quality monitoring." *India Conference (INDICON), 2015 Annual IEEE*. IEEE, 2015.

[17] Li, Hongqiang, et al. "Genetic algorithm for the optimization of features and neural networks in ECG signals classification." *Scientific Reports,* Vol. 7, 2017, p. 41011.