A Genetic Algorithm-Neural Network
Wrapper Approach for Bundle Branch Block Detection

Ragheed Allami[1], Andrew Stranieri[1], Venki Balasubramanian[1], Herbert F Jelinek[2]

[1]Centre for Informatics and Applied Optimisation, Federation University, Victoria, Australia
[2] School of Community Health, Charles Sturt University, NSW, Australia

Abstract

A Bundle Branch Block (BBB) is a delay or obstruction along electrical impulse pathways in the heart. The automated detection and classification of a BBB is important for prompt, accurate diagnosis and treatment of heart conditions, especially in accurate identification, of left BBB. This work proposes a new wrapper based hybrid approach for the detection of BBB that uses a Genetic Algorithm (GA) in combination with Artificial Neural Networks (ANN) to improve classification accuracy.

Nineteen temporal features and three morphological features were extracted and normalized for each heartbeat from standard ECG recordings obtained from the MIT-BIH Arrhythmia database.

The GA-ANN Hybrid resulted in improved sensitivity, specificity and accuracy (98%, 98% and 98% respectively) compared to the Principal Components Analysis (PCA)-ANN method (55%, 98% and 77% respectively) in the presence of noise. The GA-ANN Hybrid provides a better, more accurate identification for presence of BBB from ECG recordings leading to more timely diagnosis and treatment outcomes.

1.   Introduction

An Electrocardiogram (ECG) records the electrical impulses of the heart and indicates rhythm anomalies for diagnostic purposes [1,2]. A typical ECG tracing of the cardiac cycle consists of a P wave, QRS complex, and T wave [3]. Good performance of an ECG analyzing system depends heavily upon the accurate and reliable detection of the QRS complex, as well as the T and P waves [4]. A Bundle Branch Block (BBB) is a delay or obstruction along electrical impulse pathways of the heart manifesting in a prolonged QRS interval usually greater than 120ms. The automated detection and classification of a BBB is important for prompt, accurate diagnosis and treatment to reduce morbidity and mortality.

Applying traditional methods to extract ECG features leads to a large number of features with many features being trivial for classification. Therefore, selection of significant features plays a vital role in the classification, particularly when the ECG data is noisy. Current reduction/optimization methods to solve the feature reduction problem in ECG classification include Genetic Algorithm (GA) with Support Vector Machine (SVM) [5], Principal Component Analysis (PCA) with SVM, Linear Discriminant Analysis (LDA) with SVM [6], Cartesian Genetic Programming (CGP) with Neural Network (NN) [7], Firefly and Particle Swarm Optimization (FFPSO) technique with Levenberg Marquardt Neural Network (LMNN) [8]. Existing methods were tested on noise-free ECG data, which produced accurate classification results [9]. However, these methods may not provide the same high accuracy in the presence of a noisy ECG.

In the past decade, research in the field of artificial neural networks (ANNs) and genetic algorithms (GAs) has gained popularity [10] because of the computational abstractions of biological information processing systems. In general, in GA-ANN hybrid approaches, ANNs are used as learning systems and GAs as optimization systems, but as many researchers have discovered, they may be combined in a number of different ways resulting in highly successful adaptive systems [11].

In our proposed method, a genetic algorithm is used to extract the key features and the neural network learns to combine data on these selected features to optimize detection of BBB on noisy ECG data.

2.   Materials

The [12] is commonly used in ECG signal analysis because it includes both normal and pathological ECG signals. The MIT-BIH Arrhythmia database contains 48 ECGs of 30 minute recording lengths. Furthermore, it has two channel ambulatory ECG recordings in digital format with 360 Hz sampling rate. Our proposed algorithms were evaluated using the MIT-BIH Arrhythmia database. We applied the proposed algorithm on those records known to

have noisy ECG recordings including 101, 103, 105, 106, 109, 111, 118, 124, 207, 212 and 234. The classes of beats considered are Normal beats (N), Right Bundle Branch Block (RBBB) and Left Bundle Branch Block (LBBB). Table 1 illustrates the number of beats in each MIT-BIH record used in this study.

Table 1. MIT-BIH data records explanation.

Record No.	N	LBBB	RBBB
101	1860	-	-
103	2082	-	-
105	2526	-	-
106	1507	-	-
234	2700	-	-
109	-	2492	-
111	-	2123	-
207	-	1457	86
118	-	-	2166
124	-	-	1531
212	-	-	1825

3.   Methodology

The description of the features extracted from the ECG signals and the suggested design for the classification are illustrated in figure 1. Each phase in figure 1 is described below.
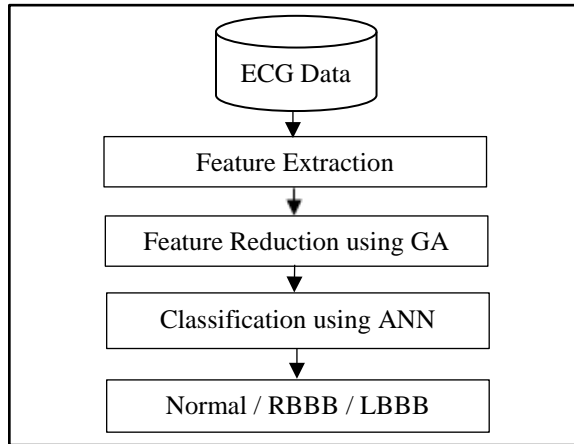

Figure 1. General block diagram of the suggested method.

3.1.   Feature extraction

For each signal 19 temporal features including PR interval, PQ interval, RR interval and PT interval and three morphological features were identified. Additional features that were included are QR, ST and RS intervals as well as difference value in these features such as V(Q)-V(S). The morphological features included the minimum

and the maximum points of each beat of the ECG signal after normalizing the patient ECGs by Equation 1:

$$f(t)=f(t)-min(t)/max(t)-min(t) \qquad (1)$$

The minimum and maximum points between the first and the second R peak were computed first and then normalized to values between 0 and 1. All features were identified for each beat and submitted to separate vectors as either N, LBBB or RBBB. The next phase involved feature reduction.

3.2.   Feature reduction using GA

GA is an optimization technique that replicates natural survival of the fittest where feature subsets with the best performance are included in the generation of new feature subsets. The next generation of subsets preserve favourable characteristics while unfavourable characteristics are omitted, leading to species progressive evolution [13]. GA iterates and evolves a population by forming a new population at every step through selection, recombination, mutation and finally applying a fitness function. The selection process directly selects subsets of features to form a current population. Each subset (chromosome) is evaluated using a fitness function. Within each subset, a crossover operation or recombination creates a new feature subset. Mutation applied to each feature subset to produce modified subsets. The fitness function then measures the quality of the solution expressed as the percentage of individuals correctly classified.

3.3.   Classification using ANN-GA

The Cascade Forward Neural Network (CFNN) used in the current study is trained with back propagation of errors [14] similar to feed-forward networks, but includes a weight connection from the input to each layer and from each layer to successive layers. The main advantage of CFNN is that each layer is related to all previous layers of the neurons.

The ANNs are capable of learning the desired mapping between the inputs and outputs signals of the system without knowing the exact mathematical model of the system and are therefore excellent estimators of nonlinear systems.

The network was trained by a continual readjustment process to the weight and the threshold value, in order to reduce the network error to a pre-set minimum of 1% or to stop at a pre-set number of cycles. Then for the forecasting, held out samples or test data was presented to the trained network to obtain the test results according to a standard approach [15]. GA was used to optimize the input features of the ANN to obtain smaller and more

effective inputs. Data in the GA were represented by n-bit binary vectors. Thus, the search space corresponds to an n-sized Boolean space. For each generation, the evaluation of the data, i.e. input feature subset required for the training of the ANN and use of the result in an objective function [16]. The classification system is described in figure 2 below that emphasizes the GA optimization.
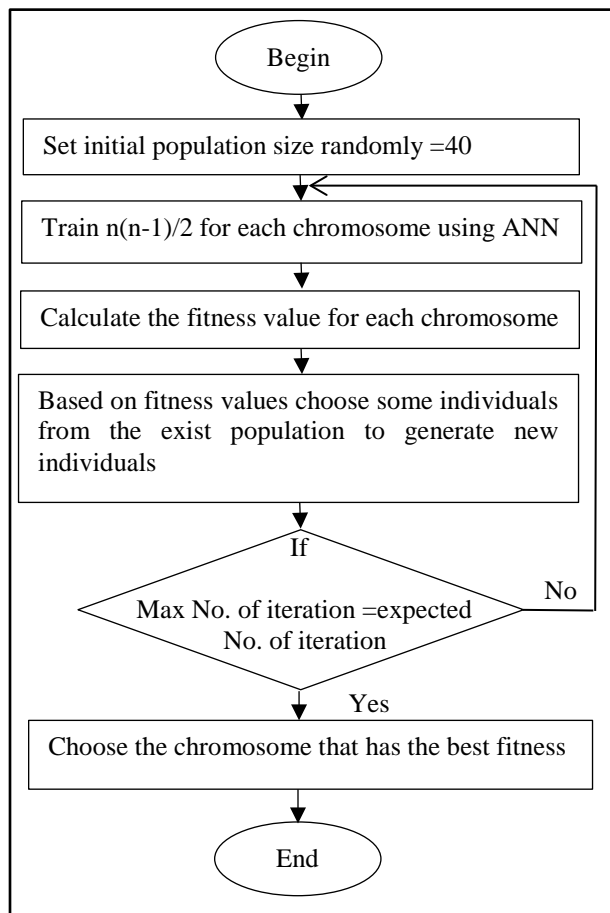

Figure 2. GA training process.

4.   Results

For the evaluation of the GA-ANN classification model half of the MIT-BIH dataset records were used for training, while the other half were used for evaluation.

Table 2 illustrates that 13 of the original 22 features were retained for the classification task following feature reduction. The results for classifying N, LBBB (LL) and RBBB (LR) were compared against PCA-ANN results and shown as sensitivity, specificity and accuracy in Table 6.

Table 2. Main features identified for the optimal classification.

Feature No.	Description	Feature No.	Description
1	V(R1)	8	V(R2)-V(R1)
2	V(S)	9	X(R1)-X(S)
3	X(T)	10	X(S)-X(T)
4	V(P)	11	X(P)-X(T)
5	X(Q)	12	X(P)-X(Q)
6	X(R2)	13	X(Q)-X(R2)
7	V(R2)		

The superiority of the GA-ANN over the PCA-ANN in BBB detection when combined with normal beats is shown in Table 3. The PCA-ANN distinguished LBBB from RBBB at 100% accuracy, but its performance decreased substantially when presented with RBBB and normal beats and to a lesser degree when comparing LBBB and normal.

Table 3. Arrhythmia classification results for sensitivity using GA-ANN compared with PCA-ANN.

Method	N	LR,LL	LR,N	LL,N
GA-ANN	100	100	93.4	100
PCA-ANN	77.33	100	66.67	90.33

GA-ANN and PCA-ANN perform equally in identifying BBB beats, but PCA-ANN loses its specificity for identifying normal beats, whereas GA-ANN is weaker for differentiating of LBBB to normal beats as depicted below in Table 4.

Table 4. Arrhythmia classification results for specificity using GA-ANN compared with PCA-ANN.

Method	N	LR,LL	LR,N	LL,N
GA-ANN	100	100	100	94.86
PCA-ANN	92.1	100	100	100

Table 5 illustrates that the accuracy of BBB detection is 100% for both classification methods. However, GA-ANN performs better than PCA-ANN when the input set was combined with normal beats.

Table 5. Arrhythmia classification results for Accuracy using GA-ANN compared with PCA-ANN.

Method	N	LR,LL	LR,N	LL,N
GA-ANN	100	100	95.67	99.17
PCA-ANN	91.67	100	83.33	96.67

Sensitivity, specificity and accuracy are reported below in Table 6.

Table 6. Parameter comparison of different methods.

Method	Sensitivity %	Specificity %	Accuracy %
GA+ANN	98.35	98.71	98.71
PCA+ANN	55.33	98.2	77.66

5. Conclusion

The GA improved the diagnostic accuracy by ignoring redundant and noisy features to determine the most significant features. Our results indicate that the ANN has the ability to self-learn, organize the dataset, and detect possible interactions by the predictor variables. For all datasets and classifiers that were evaluated, the GA-ANN approach provided improvements in the classification accuracy and sensitivity even with noisy data, except for specificity, which was 2% less accurate when compared to PCA-ANN. One of the reasons behind the successful performance of the GA-ANN model is its capability to integrate various optimal solutions given by the GA method to enhance the generality of the final solution.

Acknowledgements

References

[1] Wasilewski J, Polonski L. An introduction to ECG interpretation. In: Gacek A, Pedrycz W, editors. ECG signal processing, classification and interpretation: a comprehensive framework of computational intelligence. London: Springer, 2012:1-20.

[2] Li P, Chan KL, Fu S, Krishnan SM. An abnormal ECG beat detection approach for long-term monitoring of heart patients based on hybrid kernel machine ensemble. In: Oza NC, Polikar R, Kittler J, Roli F, editors. Multiple classifier systems. USA: Springer Berlin Heidelberg, 2005:346-355.

[3] Magdi Amien BM, Japir TE. A reliable arrhythmias-recognition scheme via wavelet and multiclass support vector machine. International Journal of Soft Computing and Engineering 2014;4:119-123.

[4] Yochum M, Renaud C, Jacquir S. Automatic detection of P, QRS and T patterns in 12 leads ECG signal based on CWT. Biomedical Signal Processing and Control 2016;25:46-52.

[5] Nasiri JA, Naghibzadeh M, Yazdi HS, Naghibzadeh B. ECG arrhythmia classification with support vector machines and genetic algorithm. Computer Modeling and Simulation, Third European Symposium 2009; 187-192.

[6] Martis RJ, Acharya UR, Min LC. ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. Biomedical Signal Processing and Control 2013;8:437-448.

[7] Ahmad AM, Khan GM, Mahmud SA. Classification of arrhythmia types using cartesian genetic programming evolved artificial neural networks. International Conference on Engineering Applications of Neural Networks. Springer Berlin Heidelberg, 2013;282-291.

[8] Kora P, Krishna KS. Hybrid firefly and particle swarm optimization algorithm for the detection of bundle branch block. International Journal of the Cardiovascular Academy 2016;2:44-52.

[9] Pasolli E, Melgani F. Genetic algorithm-based method for mitigating label noise issue in ECG signal classification. Biomedical Signal Processing and Control 2015;19:130-136.

[10] Mitra M, Samanta RK. Cardiac arrhythmia classification using neural networks with selected features. Procedia Technology 2013;10:76-84.

[11] Ding YR, Cai YJ, Sun PD, Chen B. The use of combined neural networks and genetic algorithms for prediction of river water quality. Journal of Applied Research and Technology 2014;12:493-502.

[12] Mark R, Moody G. Mit-bih arrhythmia database. URL http://ecg. mit. edu/dbinfo. 1997.

[13] Nazir M, Majid-Mirza A, Ali-Khan S. PSO-GA based optimized feature selection using facial and clothing information for gender classification. Journal of Applied Research and Technology 2014;12:145-152.

[14] Wang W, Murphey Y, Watta P. A computational framework for implementation of neural networks on multi-core machine. Procedia Computer Science 2015;53:82-91.

[15] Ding S, Xu X, Zhu H, Wang J, Jin F. Studies on optimization algorithms for some artificial neural networks based on genetic algorithm (GA). Journal of Computers 2011;6:939-946.

[16] Goldberg DE. Genetic algorithms in search optimization and machine learning. USA: Addison-Wesley, 1989.

Address for correspondence:

Ragheed Allami
University Drive,
Mt Helen VIC 3350, Melbourne, Australia
ragheed.dawood@yahoo.com