Optimized Feature Selection with Mutual Information for ECG based Bio-Metric Recognition system using Genetic Algorithm

Kiran Kumar Patro[1], Prof. P Rajesh Kumar[2]

[1]*UGC-SRF, Department of ECE, AUCE (A), Andhra University, Visakhapatnam, A.P, India*

[2]*Professor, Department of ECE, AUCE (A), Andhra University, Visakhapatnam, A.P, India*

Abstract: The ECG signal of an individual has several features. Some features vary with respect to time and body conditions, whereas some features are invariant and unique for each individual. Development of an Optimized feature selection algorithm is necessary to detect appropriate features and classify them as invariant and unique on the basis of biometric recognition of ECG signals. In this system, a total of 72 characteristic features are extracted, based on time interval, amplitude, slope and angles between fiducials from each heartbeat. The calculated features were fed to an algorithm, which reduces features by classifying vital features and avoiding random and correlated features. In this paper, a new intelligent Metaheuristic based feature optimization method; Genetic Algorithm (GA) is proposed to reduce the feature space based on a fitness function with mutual information. The obtained optimized features are trained with Machine learning classification algorithms, namely ANN, SVM and KNN for identification. The Biometric Identification system is tested with public available open database MIT-BIH ECG ID from Phisionet. Finally, the proposed algorithm enhances identification accuracy when compared to direct classification algorithms.

Keywords: Electrocardiogram (ECG); Biometric; Metaheuristic; Genetic Algorithm (GA); ANN; SVM; K-NN; ECG ID database.

1. Introduction

Biometrics literally translates to measurement of biological data. It is the measurement or the recording of human characteristics. Biometric technologies are fields of information security that are developing at a fast pace, gradually entering all spheres of human activity. Presently, three efficient biometric technologies, which are fingerprint, iris or retina, and face recognition, are being used for identification. Current security systems make use of passwords or PINs, hand geometry, voice, writing and typing dynamics, etc.,which are also useful based on the purpose and range of application. However, these systems can still be falsified, stolen and forged. Face recognition can be fooled by a picture, fingerprints can be recreated and voice can be imitated or pre-recorded [1], [2]. On the other hand, an ECG signal is a life indicator, and can be used as a tool for aliveness detection. ECG signal is controlled by anatomic features of the human heart, which are inside of human body, therefore it cannot be falsified or copied. The shape of ECG beat with its characteristic points shown in Fig.1. Many studies have introduced ECG signal as the most accurate Biometric personal identification [3], [4]. ECG based recognition was first introduced by Biel et al. [5] , in which as many as 12 features have been proposed with a certain hardware. 95% recognition rate has been achieved by testing 3 lead ECG data from 20 subjects at rest.

Adrina et al. [6] proposed ECG based biometric system with 50 subjects of lead-1 ECG and used different classification schemes based on three measures such as correlation coefficient, residual difference and wavelet distance and reported an accuracy of 80%.

Irvine et al. [7] proposed a method for human identification with eigen space analysis to

ECG traces and used Minimum distance classifier (MD) and achieved 91% identification accuracy.

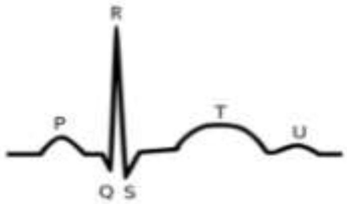


Fig. 1. ECG beat

This work mainly focused on feature extraction and optimized feature selection of ECG signal. In this, a new effective way of extracting fiducial based feature set are introduced. Moreover, optimized features are selected through Genetic Algorithm in which a finest mutual information based fitness function is proposed. Finally, the obtained optimized feature set is applied to Machine learning classification algorithms for evaluation of the system performance.

The research article is framed into 5 sections: Section.1 will give a short introduction to the background of different biometric technologies and also the advantages of ECG based biometrics. Section.2 describes steps for ECG based personal identification system and also extraction of features. In section.3 feature optimization technique is introduced. Further, section.4 presents identification results. Finally, section.5 includes the final conclusion.

2. Methodology

The biometric personal identification system uses a standard scheme, which comprises ECG data collection, Data Pre-processing, Feature extraction, Feature reduction or Optimization, and finally, Cardiac cycle classification shown in Fig. 2.

Among these stages, feature extraction and feature optimization methods can have great influence on biometric system performance. Many studies used non-fiducial based features, which are extracted from frequency content of the ECG signal (Wavelet coefficients) [8]. In recent studies, 256 biorthogonal wavelet coefficients are extracted from ECG beat by Wan and Yao [9]. Five levels of db3 DWT coefficients are used by Belgacem et al. [10].

In this study Fiducial based features are extracted from ECG beat. A time domain based analysis is used for detection of characteristic points P, Q, R, S and T [11]. After detection, it is identified that QT, ST and QRS intervals are varying accordance with heart rate, therefore finest PQRST fragments were detected for feature extraction. The best PQRST fragments were selected by their least difference from the mean of the QT, ST and QRS intervals. Only these 6 best fragments are used for data extraction [12] [13].

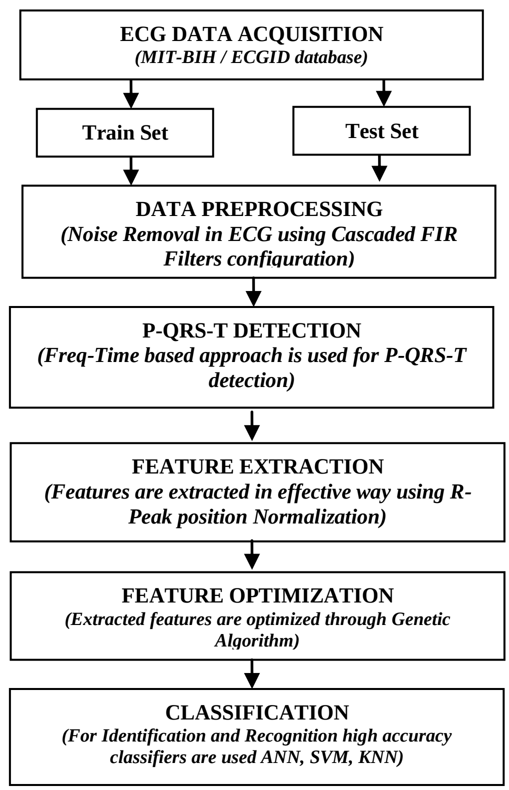


Fig. 2. Methodology for ECG Biometric System

Each pulse was broken ranging from P'-peak to T'-peak shown in Fig. 3. This broken beat pulse still not appropriate for extracting ECG feature set, so that the position has to be normalized by taking a standard origin point for R'-peak as shown in Fig. 4.

Based on time and amplitude relation between ECG fiducials, a total of 72 features are extracted, which are related to Amplitude (y), Time duration (x), Difference between amplitudes, Distance, Slope, Angles, Area, ratio and some are Miscellaneous features tabulated in Table 1.

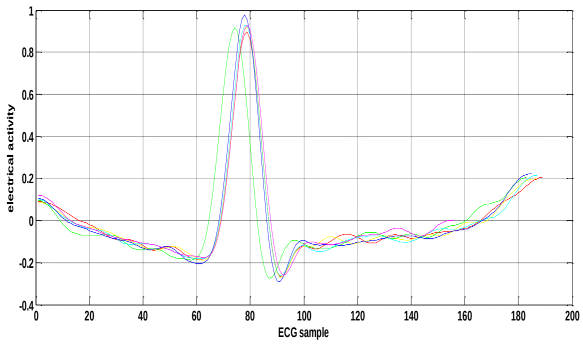


Fig. 3. ECG pulse of best six PQRST fragments

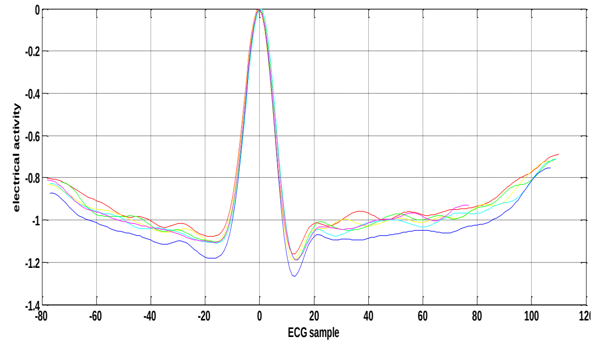


Fig. 4. Best 6-PQRST fragments with the Position Normalized

Table 1. The set of features extracted with respect to the position of 'R' as origin

Features	Features Description			
I.Time Features	1.Px 5.PQ 9.ST	2.Qx 6.PT 10.PS	3.Sx 7.QS 11.PT/QS	4.Tx 8.QT 12.QT/QS
II.Amplitude Features	13.Py 17.PQ 21.QS 25.ST/QS 29.PQ/PS 33.RS/QT	14.Qy 18.QR 22.PS 26.RS/QR 30.PQ/QR 34.ST/PQ	15.Sy 19.RS 23.PT 27.PQ/QS 31.PQ/RS 35.ST/QT	16.Ty 20.ST 24.QT 28.PQ/QT 32.RS/QS
III. Distance Features	36.PQ 40.QS	37.QR 41.PR	38.RS 42. ST/QS	39.ST 43.RS/QR
IV.Slope Features	44.PQ 48.QS 52.PR	45.QR 49.PT	46.RS 50.PS	47.ST 51.QT
V. Angle Features	53.PQR 57.RSQ	54.QRS 58.RTS	55.RST	56.RQS
VI. Miscellaneous Features	59.QRS area 63.S angl/QTtime 67.(Q/T) angl 71.QRSxcentroid	60.QRSarea/RS^2 64.S angl/PQ dis 68.QRSarea/QRamp 72.QRSycentroid	61.(R/S)angl 65.(R/Q) angl 69.QRS perimeter	62.R angl/QStime 66.(R/T)angl 70.QRS in radius

3. Feature Optimization

The data set (N=72) extracted from ECG signals may lead to a curse of dimensionality problem [14]. Feature optimization techniques reduce the computational cost and increases classification accuracy when classifying high dimensional data (N=72). Smaller feature subsets are selected from a superset of original features by feature selection to avoid irrelevant features in the data set.

All the features extracted from the ECG signals may not be marked as fiducial features. Thus the features which are fiducial are to be sort out from the non-fiducial ones. Also the correlated and over-fitted features should be sorted out to avoid the redundancy in the data and increase computational speed [15], [16].

Optimization is the selection of best element based on certain criteria, from a set of available alternatives. Optimization involves determining best available values of some objective function given a defined domain, including a variety of different types of objective functions and different types of domains. These algorithms are problem independent and it finds an optimized solution in a best way [17].

3.1 Genetic Algorithm

In the domain of artificial intelligence, a Genetic Algorithm (GA) is a search strategy that imitates the process of natural selection. Genetic algorithms are a part of the larger class of evolutionary algorithms (EA), which generate solutions in order to optimize problems using inspiration from natural evolution techniques, such as inheritance, mutation, selection, and crossover [18].

Genetic Algorithm optimizes the data to generate the best set of features by calculating the entropy in the overall data. The set of best fitting features are generated by determining the features which are least deviated from their mean value. In each generation, more robust set of features are generated until the desired fitness level is achieved or the number of maximum generation is reached.

The Working mechanism of Genetic Algorithm is described as follows:

1. The algorithm begins by creating a random initial population (Data set N=72).
2. The algorithm uses the individuals in the current generation to create the next population.
3. Scores each member of the current population by computing fitness values and converts them into a more usable range of values.
4. Based on fitness values, it selects members which are referred to as Parents.
5. Lower fitness of individuals in current population is chosen as elite. Elite individuals are passed into the next generation.
6. It produces the children from parents by making random changes to a single parent mutation or by combining the vector entries of a pair of parents' crossover.
7. It replaces the current population with the children to form the next generation.
8. The algorithm stops when stopping criteria (No. of Generations or Fitness limit) are met.

3.2 Fitness Function with Mutual Information

The Fitness function is used to evaluate the quality of every population, which must be designed before searching for the optimal values of the feature selection. Mutual Information is used to know the relationship between two variables shown in Fig. 5. It measures the dependence of one feature on another feature. The idea of mutual information is linked with entropy.

Fig. 5. Mutual Information

In Fig. 5 joint area is noted as Mutual Information M (X; Y) Entropy associated with Feature (x) is

$$H(X) = \sum_i P(x_i) \log_2 \frac{1}{P(x_i)} \quad (1)$$

Entropy associated with Feature (y) is

$$H(Y) = \sum_i P(y_i) \log_2 \frac{1}{P(y_i)} \quad (2)$$

Joint entropy associated with two features is noted as

$$H(X,Y) = \sum_i P(x_i, y_i) \log_2 \frac{1}{P(x_i, y_i)} \quad (3)$$

Mutual Information M (X; Y) with joint entropy

$$M(X;Y) = H(X) + H(Y) - H(X,Y) \quad (4)$$

$$M(X;Y) = \sum_{x_i \in X} \sum_{y_i \in Y} P(x_i, y_i) \log \frac{P(x_i, y_i)}{P(x_i)P(y_i)} \quad (5)$$

The Fitness function used is the Shannon Entropy function [19].

$$S = mean\_MIxy - C * mean\_MIxx \quad (6)$$

Where,
mean_MIxy is the Mutual information between the features and output. This function eliminates relevant data between the features and output**.**

$$mean\_MIxy = \frac{mean\_MIxy}{feat\_numb} \quad (7)$$

mean_MIxx is the Mutual information between the features. This function eliminates redundant data between the features. C is an arbitrary constant.

$$mean\_MIxx = \frac{mean\_MIxx - feat\_numb}{feat\_numb^2 - feat\_numb} \quad (8)$$

The selection function chooses parents for the next generation by their scaled values of the fitness function based on mutual information between features. To sort out redundant and correlated features, fitness limits are set to be *S_max* and *S_min*. The features (Parents) with fitness greater than *S_max* and lower than *S_min* are allowed to pass to the next generation. The mixed information between different features and their targets are also generated. After performing necessary operations related to crossover and mutation of Genetic algorithm, more robust set of features are generated over each generation until the desired fitness value is achieved.

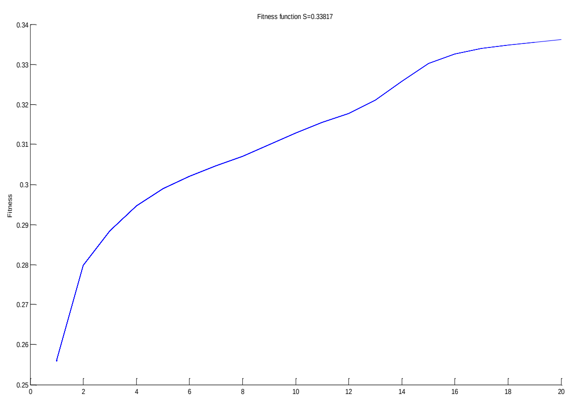


Fig. 6. Fitness Curve of GA

The Fig .6 represents the fitness curve for which target fitness level is achieved at successive generations. Finally, at 20th generation 35 best features are generated as shown in Table 2 (Feature number follows from Table 1).

Table 2. Best set of features reproduced at 20^{th} generation using GA

Type	Optimized Features Description
1.Time Features	1.Px 2.Qx 3.Sx 4.Tx 9.ST 10.PS 11.PT/QS
2.Amplitude Features	13.Py 14.Qy 17.PQ 21.QS 22.PS 24.QT 25.ST/QS 26.RS/QR 27.PQ/QS 28.PQ/QT 30.PQ/QR 32.RS/QS 33.RS/QT 34.ST/PQ
3.Distance Features	36.PQ 38.RS
4.Slope Features	47.ST 49.PT 50.PS 51.QT
5.Angle Features	54.QRS 61.(R/S)angle 65.(R/Q) angle 66.(R/T) angle
6.Miscellanious Features	67.(Q/T) angle 68.QRSarea/QRamp 69.QRS perimeter

4. Identification Results

The Performance of the proposed Genetic algorithm for ECG biometric application tests with public available ECG ID database of Phisionet website [23]. Five ECG records taken under different conditions for each individual are considered. Out of five, two records are used for training and the remaining for testing. For biometric identification Lead-I ECG is chosen, because it is easily measured and not sensitive. The best sets of optimized features from Genetic Algorithm are used to train Machine learning classification algorithms, namely Artificial Neural Networks (ANN), Support Vector Machines (SVM) and K- Nearest Neighbor (KNN) for Identification. The performance of Biometric Personal identification can be estimated by examining the performance parameters. All the performance parameters are calculated using confusion matrix for true identification of person.

Table 3. Confusion Matrix for True Identification of 20 Persons using GA+ANN methodology

	P3	P10	P25	P30	P32	P34	P36	P52	P53	P59	P72	P1	P2	P24	P28	P11	P26	P9	P63	P71
P3	**3**	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
P10	0	**2**	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0
P25	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P30	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P32	0	0	0	0	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P34	0	0	0	0	0	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P36	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0
P52	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0
P53	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0
P59	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0
P72	0	1	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0
P1	0	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0
P2	0	0	0	0	0	0	0	0	0	2	0	0	**2**	0	0	0	0	1	0	0
P24	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	0	0	0
P28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0
P11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	0
P26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0
P9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0
P63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**	0
P71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**
Match	100%	66.7%	100%	100%	66.7%	66.7%	100%	100%	100%	33.3%	33.3%	100%	100%	66.7%	100%	100%	100%	66.7%	100%	100%

Table 4. Confusion Matrix for True Identification of 20 Persons using GA+SVM methodology

	P3	P10	P25	P30	P32	P34	P36	P52	P53	P59	P72	P1	P2	P24	P28	P11	P26	P9	P63	P71
P3	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P10	0	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P25	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P30	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P32	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P34	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P36	0	0	0	0	0	0	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0
P52	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0
P53	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0
P59	0	0	0	0	0	0	1	0	0	**1**	0	0	0	0	0	0	0	0	0	0
P72	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0
P1	0	1	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0
P2	0	0	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0
P24	1	0	0	0	0	2	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0
P28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0
P11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	0
P26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0
P9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**	0	0
P63	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	0	0	**3**	0
P71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**
Match	66.7%	66.7%	100%	100%	33.3%	33.3%	66.7%	100%	100%	33.3%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

According to confusion matrix, the True Positive Rate of overall true identification of person involved in a testing session with ANN Classifier using GA was 82.80% from Table 3.

From confusion matrix, the True Positive Rate of overall true identification of person involved in testing session with SVM Classifier using GA was 84.482% from Table 4.

Table 5. Confusion Matrix for True Identification of 20 Persons using GA+KNN

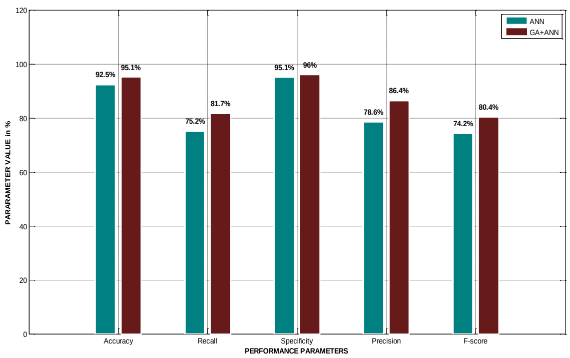
	P3	P10	P25	P30	P32	P34	P36	P52	P53	P59	P72	P1	P2	P24	P28	P11	P26	P9	P63	P71
P3	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P10	0	**3**	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
P25	0	0	**3**	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
P30	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P32	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P34	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P36	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0
P52	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0
P53	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	2	0	0	0	0	0
P59	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0
P72	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	1
P1	0	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0
P2	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	0	0	0	0
P24	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	0	0	0
P28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0
P11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	0
P26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0
P9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**	0	0
P63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**	1
P71	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	**1**
Match	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	33.3%	100%	66.7%	66.7%	33.3%	100%	100%	100%	100%	33.3%

According to confusion matrix, the True Positive Rate of overall true identification of person involved in testing with K-NN Classifier using GA was 86.2069% from Table 5.

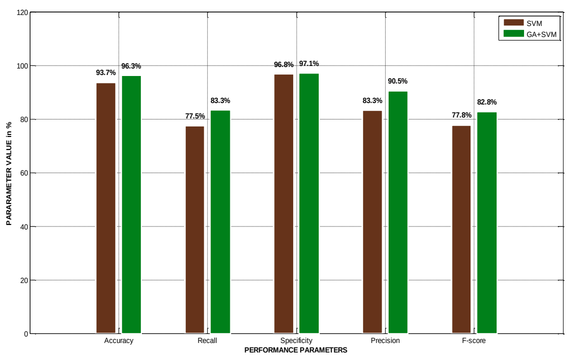
Table 6. Performance Metrics (%)

S.No	Techniques	True_Classification Rate(TPR)	True_Identification Rate(TPR)	Overall Classification Accuracy	Recall	Specificity	Precision	F Score
1	ANN	71.1823	81.010	92.4532	75.210	95.1303	78.6427	74.2122
2	SVM	73.1527	82.7586	93.7094	77.510	96.7934	83.263	77.8418
3	KNN	72.9064	81.0345	92.7241	77.9762	96.2739	79.2631	76.8515
4	GA+ANN	74.8768	82.80	95.1034	81.6667	96.00	86.3929	80.369
5	GA+SVM	77.5862	84.4820	96.2759	83.333	97.0909	90.5357	82.7857
6	GA+KNN	75.6158	86.2069	96.7931	88.333	97.3669	90.9167	87.0357

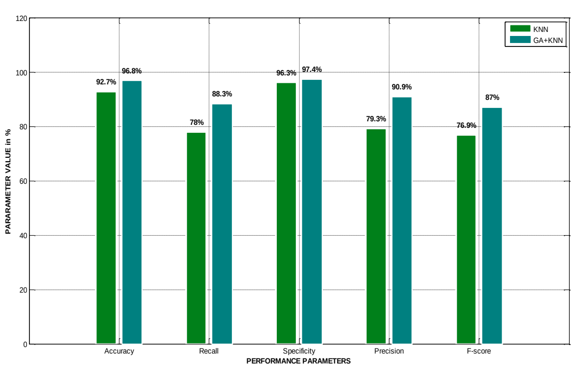
From Table 6, Optimized feature selection methods shows enhanced results when compared to the features directly imposed on Classification algorithms. Meanwhile, it also improves the quality of classification by increasing parameters like accuracy, precision, recall, specificity and F-score.



(a) GA with ANN Classifier



(b) GA with SVM Classifier



(c) GA with K-NN Classifier

Fig. 7. Comparisons of performance metrics with Machine learning classification algorithms

Fig.7 shows comparison of different performance metrics for Biometric recognition system using ANN, SVM, K-NN classifiers, with and without optimization. t is observed that GA based feature optimization methodology shows improved results when compared to direct classification of ECG data. ANN with GA achieved an improvement of 2.65% in Classification accuracy compared to ANN. Similarly, SVM with GA shows an improvement of 2.56% in Classification accuracy compared to SVM. Finally, KNN with GA shows an improvement of 4.069% in Classification accuracy compared to KNN.

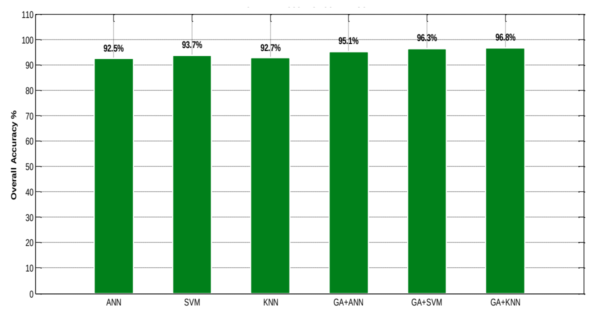


Fig. 8. Overall Accuracy Comparisons

Fig. 8 shows comparison of overall accuracy for biometric system for different techniques. It is observed that proposed Metaheuristic based optimization Technique (GA) shows enhanced results when compared to direct classification algorithms, in which GA shows an improvement of 4.339%. Finally, the proposed GA+KNN show an accuracy of 96.7931% for the biometric personal identification system when compared to previous literature shown in Table 7.

Table 7. Comparison with existing Literature

Author	Techniques	Accuracy
Adrian, et. Al 2008 [6]	Percent Residual difference (PRD) Correlation Coefficient (CC)	70% (PRD) 80% (CC)
Irvine, et. Al 2008 [7]	Minimum Distance Classifier (MD)	91%
Tatiana, et. Al 2005 [20]	LDA and Majority Vote Classifier	96%
X.Tang, et. Al 2014 [21]	Quantum Neural Network (QNN)	91.70%
F. Silva, et. Al 2017 [22]	SVM with Genetic Algorithm	93.1034%
Proposed Study*	ANN, SVM, KNN with Genetic Algorithm	96.793% (GA+KNN)

5. Conclusion

In this research, Metaheuristic based feature optimization method has been proposed for selecting the best suitable features for classification to avoid random, correlated and over-fitted features. In this paper, Genetic Algorithm (GA) has been introduced to reduce the feature space based on a fitness function with mutual information between features. The ECG signals considered for analysis were taken from 20 subjects each five individuals each, for duration of six months from the MIT-BIH ECG ID database. Performance of Biometric personal identification system has been tested with different classification algorithms such as ANN, SVM and KNN, with and without feature optimization. Experimental results show that the proposed feature optimization methods considerably improve the quality of classification by improving certain parameters like accuracy, precision, recall, specificity and F-score. Genetic Algorithm shows an improvement of 4.339% in accuracy, when compared to direct classification algorithms. Finally, the proposed GA+KNN technique for classification shows an accuracy of 96.7931%.

References

[1] R. Scott Blaul, John Chirillo, “Implementing Biometric Security”, Inianapolis: Wiley publications 2003.

[2] JH Connel, RM Bolle, Pankanti S, “Guide to Biometrics”, Newyork: Springer Verlag 2003.

[3] Israel SA, Irvine JM, Cheng A, Wiederhold BK, “ECG to identify individuals”, Pattern Recognition, vol. 38 , pp.133-142, 2005.

[4] Israel SA, Irvine JM, Cheng A, Wiederhold BK, “The Heartbeat the living Biometric. Newyork”, USA: Wiley-IEEE Press;2009.

[5] Biel. L, Pettersson. O, Philipson. L, Wide. P , “ECG analysis: a new approach in human identification.”, IEEE Transactions on Instrumentation and Measurement, vol. 50, no. 3, pp. 808-812, Jun. 2001 .

[6] Adrina, D.C Chan, M Hamdy, “Wavelet distance measure for person identification using electrocardiograms”, IEEE transactions on Instrumentation and measurement, vol. 57, no. 2, pp. 248-253, 2008.

[7] Irvine JM, Israel SA, Scruggs WT,” Robust human identification from Cardiovascular function”, Pattern Recognition, vol. 41, pp. 3427-3435, 2008.

[8] Hatzinakos D, Agrafioti F, Gao J, “Heart Biometrics: Theory, methods and applications” Biometrics B, vol. 3, pp. 199-216, 2011.

[9] Wan Y, Yao J, “ A Neural network to identify human subjects with Electrocardiogram signals”, World congress on Engineering and Computer Science; 2008; USA.

[10] Belgcem N, Ali A, Fouenier, “ ECG based human authentication using wavelets and random forests”, International Journal of Cryptography Information security, vol. 2, pp. 1-11, 2012.

[11] Kiran Kumar patro, P. Rajesh Kumar, “ A Novel Frequency-Time Based Approach for the Detection of Characteristic Waves in Electrocardiogram Signal”, Springer India, vol. 372, pp.57-67, 2015.

[12] LugovayaT.S.,“Biometric human identification based on electrocardiogram”, [Master's thesis] Faculty of Computing Technologies and Informatics, Electrotechnical University "LETI", Saint-Petersburg, Russian Federation; June 2005.

[13] Kiran Kumar patro, P. Rajesh Kumar, “ Effective Feature extraction of ECG for Biometric Application”, Journal of Procedia of Computer Science, Elsevier vol. 115, pp.296-306, 2017.

[14] Devi A, Misal A, “A survey on Classifiers used in heart valve disease Detection”, International journal of Advance Research in Electrical, Electronics and Instrumentation Engineering, vol. 2, no. 1, 2013.

[15] Khan M, Quadri, “Effects of using filter based feature selection on the performance of Machine learners using different datasets” International Journal of Information Security, vol. 5, no. 2, pp. 597-603, 2013.

[16] Adejoke B, Omotoyo M., “Review of Feature selection Methods in Medical Image Processing”, IOSR Journal of Engineering, vol. 4, no. 1, pp. 01-05, 2014.

[17] SK Singhi, Liu H, “Feature Subset selection bias for Classification Learning”, 23rd International Conference on Machine Learning, pp. 01-08, USA 2006.

[18] Gen M, Yun Y, “Soft Computing Approach for reliability optimization: State of art survey”, In Proceedings of the Reliability engineering and system safety, pp. 1008-1026, 2006.

[19] Xu Y, Jones G, Wang B, “ A Study on Mutual Information based Feature selection for text Categorization”, International Journal of Computational Information Systems, vol. 3, no. 3 , pp. 1007-1012.

[20] Tatiana, T.S Lugovaya, “Biometric human identification based on Electrocardiogram”, Proc. XII th Russian Conference on Mathematical Methods of Pattern Recognition, pp. 387-390, Moscow, 2005.

[21] X. Tan, L Shu, “Classification of Electrocardiogram signal with RS and Quantum Neural Networks”, International Journal of Multimedia and Ubiquitous Engineering , vol. 9, no. 2, pp. 363-372, 2014.

[22] F. G. Silva Teodoro, S. M. Peres and C. A. M. Lima, "Feature selection for biometric recognition based on electrocardiogram signals," *IEEE* International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, 2017, pp. 2911-2920.

[23] http://www.physionet.org/cgi-bin/atm/ATM – “MIT-BIH ECG ID Database.