Biomedical Engineering: Applications, Basis and Communications, Vol. 29, No. 3 (2017) 1750023 (10 pages)
DOI: 10.4015/S1016237217500235

World Scientific
www.worldscientific.com

A SYSTEM FOR ACCURATELY PREDICTING THE RISK OF MYOCARDIAL INFARCTION USING PCG, ECG AND CLINICAL FEATURES

Maryam Zarrabi*, Hossien Parsaei*,**, Reza Boostani†, Azam Zare‡, Zhila Dorfeshan§, Khalil Zarrabi¶ and Javad Kojuri‖

*Department of Medical Physics and Engineering, School of Medicine
Shiraz University of Medical Sciences, Shiraz, Iran

†Department of CSE & IT, Faculty of ECE
Shiraz University, Shiraz, Iran

‡Tarbiat Modares University, Tehran, Iran

§Alzahra Hospital, Shiraz, Iran

¶Department of Cardiac Surgery, Nemazee Hospital
Shiraz University of Medical Sciences, Shiraz, Iran

‖Education Development and Research Centre
Shiraz University of Medical Sciences, Shiraz, Iran

Accepted 25 June 2017
Published 17 July 2017

ABSTRACT

Myocardial infarction (MI) also known as heart attack is one of the prevalence cardiovascular diseases. MI that is due to the blockade in the coronary artery is caused by the lack of blood supply (ischemia) to heart tissue. Determining the risk of MI and hospitalizing the victim immediately can prolong patient's life and enhance the quality of living through appropriate treatment. To make this decision more accurate, in this study, a decision support system is proposed to classify patients with hard chest pain (sign of MI) into high and low risk groups. Such a system can also assist in managing the limitation of bed in the care units such as cardiac care unit by deciding on admitting a subject with a hard chest pain whom refers to a hospital or not. Despite several efforts in this issue, the so far published results demonstrated that distinguishing these patients using just electrocardiogram (ECG) features is not promising. **In addition, these methods did not focus on classifying the patients with high and low risks of MI**. In this regard, auxiliary features from phonocardiogram (PCG) signals and clinical data were elicited to create a discriminative feature set and ultimately improve the performance of the decision making system. In this research, ECG (from 12 leads), PCG signal and clinical data were acquired from 83 patients two times (morning and evening) in the first day. Since the number of elicited features from the raw data of each patient is high, the irrelevant and non-discriminative features were eliminated by sequential forward selection. The selected features were applied to k-nearest neighbor classifier resulted in 98.0% sensitivity, 100% specificity and 99.0% accuracy over the patients. The results illustrate that neither clinical data nor ECG features nor PCG features are lonely enough for estimating the risk of MI. Employing features from different modalities can improve the performance such that the developed multimodal-based system overperformed single modal-based systems. The obtained results are promising and suggest that using this system might be useful as a means for altering the risk of MI in patients.

Keywords: Feature extraction; Dimension reduction; Classification; Myocardial infarction.

**Corresponding author: Hossein Parsaei, Department of Medical Physics and Engineering, Shiraz University of Medical Sciences, Shiraz, Iran. Tel/Fax:(+98) 7132349332; E-mail: hparsaei@sums.ac.ir

INTRODUCTION

Cardiac infarction is one of the leading causes of death and disability in developed countries.[1] As 30% of mortality is approximately related to cardiac diseases, heart failure is among the costliest cardiovascular disorders and hence it represents a tremendous load to overall healthcare costs.[2] One of the most prevalence cardiovascular diseases is myocardial infarction (MI) that is caused by the lack of blood supply (ischemia) to heart tissue. Heart ischemia is due to the blockade in the coronary artery. When ischemia happens, the blood perfusion to the heart is interrupted and ultimately the myocardium is damaged leading to the occurrence of MI.[2]

Although a high number of patients with a hard chest pain are referred to hospitals everyday, a few of them could be admitted in the cardiac care unit (CCU), mainly due to the limitation in CCU beds. To decide whether a patient should be admitted or not, specialists assess electrocardiogram (ECG) and clinical features while this decision is not promising enough. Early diagnosis of MI can prolong patient's life and enhance the quality of living through appropriate treatment. In addition, identification and treatment of high risk subjects before a brutal hard attack is vital as it has a great impact on reducing the side effects of MI.

The simplest and widely used technique for early diagnosing patients with the risk of MI is visual inspection of ECG signals and assessing other clinical data such as sport test, patient's history, laboratory data. In fact, ECG reveals the electrical activity of heart and is mostly recorded by placing three electrodes for early diagnosis; nevertheless for an accurate screening, 12 electrodes should be placed for assessing the heart muscles' activity. ECG signal is a quasi-periodic signal carrying a low degree of stochastic behavior for normal subjects while ECG signal behaves a bit chaotic when arrhythmia is happened. As far as cardiovascular disorders change the amplitude, morphology and frequency bands of ECG components,[3] several attempts have been made to detect different ECG arrhythmia.[3–13] ECG is used as a diagnostic means of cardiovascular diseases[4] for these four main reasons: (I) it is easily recorded in a non-invasive manner, (II) ECG recording is fairly cheap, (III) ECG setups are mostly portable and above all, (IV) ECG signals can manifest 97.0% of all cardiac diseases. Therefore, for every patient with a sign of heart disease, an ECG test is completed at the first step. Nevertheless, it should be stated that standard ECG results suffer from the lack of sensitivity to act as the only screening tool for diagnosing MI.[5]

Towards the detection of MI using ECG analysis, Baxt *et al.*[6] used an artificial neural network (ANN) with four layers to recognize the presence of MI. ECG features as well as clinical data were arranged in a feature vector for each subject. The network was trained on 706 patients with anterior chest pain. The overall sensitivity and specificity on the trained network was 97.0% with a striking result on recognizing the presence of MI but not on predicting. Yang *et al.*[7] designed an ANN using QRS and S–T features as input parameters to diagnose inferior MI. They used 592 ECGs (208 from patients with inferior myocardial infarction, 300 from normal subjects and 84 left ventricular hypertrophy cases). The sensitivity of this system was 88.0%. These researchers showed that increasing the number of hidden layers increased the computation time of the system while did not assist with improving the accuracy of the system. They concluded that using new parameters of ECG might enhance the performance of the system. Heden *et al.*[8] developed a system using multilayered perceptron ANN to diagnose MI. They recorded 12-lead ECG of 1313 patients with chest pain. Amplitudes and durations of the ECG components such as Q–R and S–T were applied to the system. The sensitivity for the network was 81.4%. Focusing on diagnosing only anterior MI is one of the reasons of high number of false negative errors of the system presented in this study. Another reason is using the limited number of variables of ECG and uncommon features as input values. Menown *et al.*[3] aimed to diagnose acute MI using S–T elevation with different definitions. Moreover, they added QRS complex features such as Q waves duration, S–T depression, T wave inversion, bundle branch block, axes deviation and left ventricular hypertrophy to S–T elevation and improved overall classification. The accuracy of classification estimated using 1190 subjects was 84.1%. Although ST segment elevation is one of the important parameters to diagnose MI, using additional features of ECG signals may assist with improving the accuracy of this system. Haraldsson *et al.*[9] used ANN along with 12-lead ECG to detect acute MI. They analyzed all 12-lead ECGs using Hermitian functions and the coefficients were elicited and applied to a Bayesian neural network. They used 1119 ECGs from patients who admitted to the coronary care unit and discharged with AMI diagnosis to train and test the system. The accuracy was 94.0%, but again a limited number of ECG features were used.

Heart sounds reflect the blood turbulence created when the heart valves snap shut. Cardiac auscultation is one of the widely used tools for listening/recording these sounds that provide important auditory data regarding

the condition of the heart. There is a possibility to record these sounds in the form of an electrical signal, termed as PCG, using an array of dynamic microphones. The PCG signal is mainly used for the diagnosis of the heart valves and artery stenosis. The heart cycle is composed of four sounds that the first two components $(S1, S2)$ can be heard and two other components $(S3, S4)$ have lower amplitudes which make them hard to be heard.[10] PCG sensors are much sensitive than the stethoscope diaphragm and provides a signal with much higher quality.[11] Some attempts have been made to investigate the PCG features before, through and after occurring MI. In this regard, Price and Brown[12] tried to measure the intensity of the heart sounds after MI and found a significant reduction in the intensity of the heart sounds compared to the normal state. Moreover, Adolph *et al.*[13] analyzed frequency content of the heart sounds and evaluated the clinical utility of this technique for MI diagnosis using 158 subjects with different cardiac diseases. An accuracy rate of only 54.0% has been reported.

Although PCG sounds are synchronously occurred with the ECG components, they convey complementary information for the ECG features. A few efforts for simultaneous processing of ECG and PCG signals have been deployed to identify ischemic heart from the normal one,[5] but the results were not as remarkable as expected.

Fatemian[14] studied the applicability of ECG and PCG as two cardiac biometrics. Identification and authentication are proposed and analyzed in both scenarios. A total of 98 subjects were chosen by two publicly available (MIT-BIH and PTB) and University of Toronto (U of T) rest and exercise databases.

Christer Ahlstrom[15] tried to develop PCG signal processing techniques (e.g. homomorphic filtering, Shannon energy, variance fractal dimension, etc.) using ECG signals. A number of data sets have been used in this thesis while the maximum number of subjects in all six group of datasets was 77.

Noorzadeh[16] analyzed the morphology of fetal ECG using multimodality. PCG signal was used as another signal modality. Nine pregnant healthy women participated in the study. The result shows a good detection of R-peaks in ECG.

The performance of the decision support systems presented so far is not high enough to be used in clinical environments,[6–8] hence the existing remarkable gold-standard diagnosis manner is still the specialist's opinion, which suffers from the subjectivity. In practice, besides the observation of ECG and PCG, physicians pay much attention to clinical features of patients like sport test, history of patient and laboratory data. To build up a coverall set of informative features, all of these modalities need to be deployed to achieve a higher accuracy in correct diagnosis of patients with chest pain.[17] Nevertheless, the manner of fusing the elicited features from different modalities is the critical step in the diagnostic process.

Even though some of the mentioned studies demonstrated ECG/PCG *features change just before and after occurring* MI, *they did not focus on classifying the patients with high and low risks of* MI. *In other words, these studies did not focus on predicting the occurrence of* MI *in advance. None of the former studies attempts follow up their patients' data for* 12 h *and all of them just use* MIT/BIH *or* PTB datasets[14,18] *which do not include the follow up labels of these patients whether experiencing an* MI *after the chest pain or not.*

In this study, the first step was to collect a dataset from those who referred to the Alzahra heart hospital (Shiraz, Iran) with the sign of chest pain. Since they monitored 12 h later, their information in terms of ECG, PCG and clinical were collected two times in the morning and evening, meanwhile they were in the hospital campus. After recording and preprocessing the collected raw data, different features were extracted from PCG and ECG signals while the clinical data of each patient converted to 42 feature values. Regarding the scattered nature of features, k-nearest neighbor (KNN) classifier was employed to locally assign the labels to the test vectors.

The rest of this paper is organized as follows. Section 2 describes the collected dataset and presents the methodology. Section 3 explains the evaluation method and demonstrates the achieved results. In Sec. 4, results of the proposed approach and state-of-the-art methods on this dataset were compared and their pros and cons were discussed. Finally, the paper is concluded in Sec. 5.

MATERIALS AND METHODS

The presented method is to estimate the class label of a patient (i.e. under the risk of MI/ not under the risk). The concept of risk in medicine varies according to different health settings and different times. In this work, risk means the probability of developing MI. As with any pattern recognition method, the developing process consists of four main steps (Fig. 1): data collection, feature extraction from different modalities (ECG, PCG and clinical data), dimension reduction, and classification. Following is a description of each step.

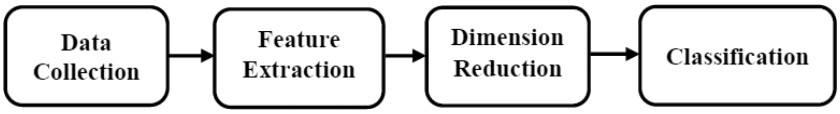

Fig. 1 The steps of developing MI risk detection algorithm.

Data Collection

With the cooperation of Alzahra heart hospital (Shiraz, Iran), 83 patients participated in this research whose age ranged from 20 to 80 years old (Mean age: 62 ± 10). They contained 29 males (35%) and 54 females (65%). These patients are referred to the hospital with the sign of chest pain which is the most common symptom of patients before MI. However, clinical examination conducted by cardiologists reported that 44 patients were high risk patients and the remaining 39 were identified as low risk patients. Three different types of data were recorded from the patients: (a) patient's history, physical examination results and the laboratory data, (b) 12 leads ECG signals and (c) PCG signal.

The first set of information is collected according to the form shown in Fig. 2 that contains quantitative and qualitative parameter. All of the data were quantified and considered as a part of the feature vector (42 values) for each subject.

The ECG signals were recorded by CARDIAX ECG (version 4.21.6),[19] containing 12 leads which records

Name:	**Code:**		**Tel:**	
Gender:	**Age:**		**Date:**	
Main Presenting Symptoms:				
Ischemia (Typical) chest pain ☐	Atypical chest pain ☐	Dyspnea ☐	Cold sweating ☐	Nausea/Vomiting ☐
Weakness ☐	Palpitation ☐	Dizziness/Syncope ☐	Aborted SCD or cardiac arrest ☐	other
Risk Factors:				
Diabetes mellitus ☐	Hypertension ☐	Dyslipidemia ☐	IHD ☐	Kidney disease ☐
Peripheral artery disease ☐	Family Hx of premature CAD ☐	Cigarette ☐	Waterpipe smoking ☐	Opium ☐
Family hx:				
Past Cardiac hx:				
Physical examination:				
Blood pressure:	Heart rate:	Pulmonary edema ☐	Peripheral edema ☐	
Lab Data:				
Wbc:	Hb:	Plt:	Bun:	Creat:
Na:	K:	FBS:	TG:	Chol:
LDL:	HDL:	CPK-MB:	Troponin:	
Echocardiography:				
LVEF:		RWMA:		
Angiography:				

Fig. 2 A form of collecting clinical history, physical examination and laboratory data information.

ECG signals with sampling rate of 2 kHz. ECG signals were recorded for 60 s. The recorded signals were first filtered via a 5th order band-pass Butterworth where the lower and upper cutoff frequencies are 0.5 Hz and 150 Hz, respectively.

The third group of data is PCG signal. To record PCG, a digital stethoscope (HANBYUL model) was used which records the signals with the sample rate of 8 kHz. As with ECG signals, PCG signals were also recorded for 60 s. After computerized recording of discrete PCG signals, they were filtered by a 5th order band-pass Butterworth filter with the lower and upper cutoff frequencies of 20 Hz and 150 Hz, respectively.

Feature Extraction

Each subject is presented by a quantitative vector containing 143 features. The first 42 features were extracted from patient's history, physical examination results and the laboratory data. Clinical data acquired using the form shown in Fig. 2 are all converted to quantitative values. 96 features were extracted from the 12-lead ECG signals and five features from PCG signals.

For the second part of the features (i.e. features of ECG signals), from each ECG lead the eight parameters duration of P, P–R, Q–R–S and S–T along with the amplitudes of P, QRS, ST and T (as introduced in Refs. 16 and 17) are computed and are used as features. Consequently, 96 feature values are computed from 12 ECG leads recorded for each patient. These features are employed because they provided a good discrimination ability in identifying MI.[20] It should be mentioned that the CARDIAX software is able to automatically determine these amplitudes and durations; therefore, no more software is written to elicit ECG features.

For PCG features, totally five features were extracted from each PCG trial. The first three features are mean, median and sum of PCG spectrum. For estimating the power spectrum (PS) of a given PCG signal, we used Welch method[21] by which the expected value of the PS estimate is

$$\hat{p}(f_n) = \frac{L}{UK} \sum_{k=1}^{K} |A_k(n)|^2, \tag{1}$$

$$U = \frac{1}{L} \sum_{j=0}^{L-1} W^2(j), \tag{2}$$

where L is the length of data segments, K is the number of segments, A_k is the finite Fourier transforms and $W(j)$ is the data window. In this work, we used Hamming window.

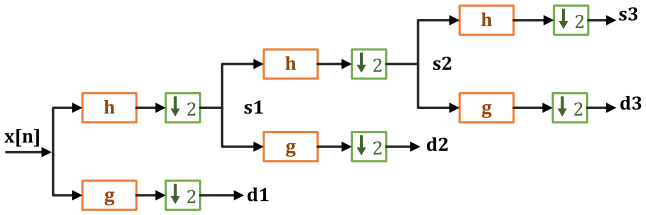

Fig. 3 A three-level discrete wavelet decomposition of a signal $x[n]$. Sm represents approximation coefficients and dm symbolizes detailed coefficients obtained at level m.

The two remaining PCG features were calculated using wavelet decomposition of PCG signal.

Wavelet transform (WT) of a signal is an efficient method for representing both time and frequency information of a given signal. In fact, the WT of a signal provides a two-dimensional time–frequency representation of the signal. For computing the discrete wavelet transform (DWT) of a signal (known as wavelet coefficients), the given signal $x[n]$ is successively passed through a series of low-pass and high-pass filters. Passing the samples through a low-pass filter results in approximation coefficients $s[n]$, and filtering the signal with a high-pass filter provides detailed coefficients $d[n]$. Mathematically, DWT can be expressed as

$$s[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k], \tag{3}$$

$$d[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k], \tag{4}$$

where $h[n]$ represent the impulse response of low-pass filter.

This process is continued to further extract wavelet coefficients. Figure 3 illustrates how an input signal $x[n]$ is passed through three successive approximation and detail filters. The outputs contain the contents of the signal at different frequency sub-bands. The approximations are the low-frequency components and the details are the high-frequency components.[22] The features used in this work are the mean of approximation coefficients and the mean of detail coefficients which were determined. Daubechies (db6) was employed as mother wavelet.

Dimension Reduction

Typically, in a pattern recognition problem, the number of features should be large enough to well describe a pattern; in contrast, this number should not be too large to prevent the curse of dimensionality. The phrase

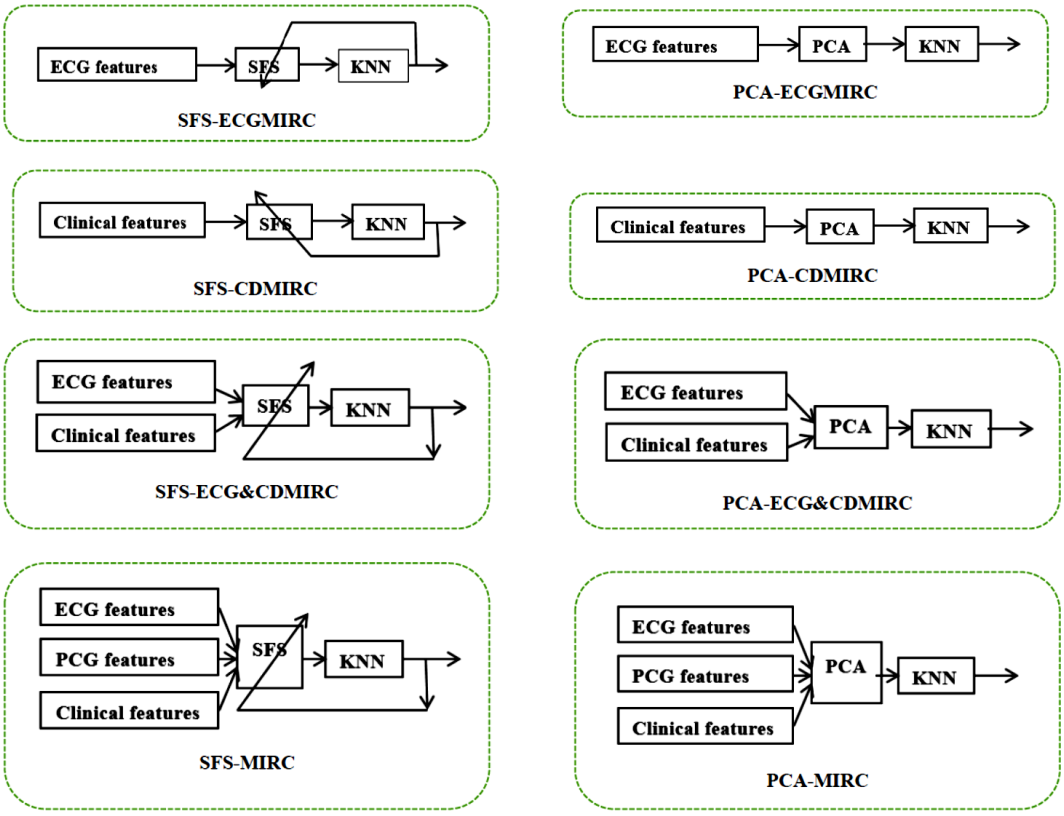

Fig. 4 Classification algorithms schemes used in this work.

Note: The systems that did not use any dimension reduction techniques such as PCGMIRC, ECGMIRC and CDMIRC are not depicted in this figure.

"curse of dimensionality" describes the fact that the time and number of samples required to compute an approximate solution for a pattern recognition problem grows exponentially with the dimension of the feature space. Ideally, a feature vector should be comprised of a low number of uncorrelated features that their values significantly changed from one class to another one. In fact, independent features with high inconsistency (from one class to the other one) should be selected or extracted to accurately classify input patterns.

In this study, the dimensionality of the feature space was reduced in two successive steps. In the first step, a filter approach feature selection method was used to select mutually informative features. By using Pearson correlation coefficient, we found that the feature P wave amplitude along with the four features that are related to the duration of an ECG complex (P wave duration, PR duration, QRS duration, and ST duration) were very correlated for 12-leads. Therefore, we considered the values measured for these five features using only one lead, Lead II. We used this method for reducing the number of ECG features because our experience showed that the value for these four features are correlated over the 12 leads. By using this simple filter-based feature selection technique, the number of features of ECG signal used was reduced from 96 to 41.

In the second step of feature reduction, two dimension reduction methods were examined that one of them is the feature extraction and the other is the feature selection. Principal component analysis (PCA) was used as the feature extraction while sequential forward selection (SFS) was used as the feature selection scheme.

In PCA, the existing features are transformed to a new coordinate system which are orthogonal to each other and the axes corresponding to the larger variances are selected. The most effective features in the given data are the first d principal components that account for a specific portion, say $\beta\%$, of the variance in the data. The value of d is the dimension of the new feature space. The best value for the parameter β was found empirically through cross-validation by which the accuracy of the system was estimated using several values of β (ranging from 90 to 99) and then the value on which the performance of the system is maximized was set as the value for β.

In SFS, the main goal is to select the most discriminative subset of features iteratively. In other words, the

objective was to choose a subset of d features, such that the resulting subset of features does not degrade the performance of the system significantly. The SFS algorithms start with an empty feature set and then repeatedly add the most significant feature to the subset until there is no improvement in the objective function. SFS is used as a wrapper approach here where the classification accuracy feedbacks the algorithm to select the new feature at each iteration.

Classification

Classification, by definition, is the problem of identifying the class label of a new pattern (sample), by using a set of data that their class labels are known (i.e. training data). A classifier is a function that shatters input space into some subspaces where samples of each subspace take a certain label. In this study, the KNN classifier[19] was deployed to classify the patients with high risk of MI from those (low risk patients) who do not need to be admitted at CCU. The KNN classifier is a non-parametric supervised classifier that assign a class label to the input pattern based on the class labels represented by the K-closest neighbors of the vector.[23] We used the KNN classifier because this classifier does not have any training phase and more importantly it is simple non-parametric suboptimal classifier. Although the functionality of this classifier seems easy, it has derived from the Bayesian decision rule where the probability function of each class is estimated in a non-parametric manner. It can be mathematically proved that the error probability of KNN falls with the following bounds:

$$P_{\text{error}}(\text{Bayes}) \leq P_{\text{error}}(1\text{NN}) \leq 2P_{\text{error}}(\text{Bayes}),$$

where 1NN is the KNN with $K = 1$.

There are two parameters that need to be defined when using KNN: the distance metric for finding the neighbors and the number of neighbors (K). In this study, Euclidean distance was chosen and the parameter K was estimated through cross-validation by which the accuracy of the classifier was estimated for different values of K and then the value of K which maximized the classification accuracy was used as the best value for this parameter. In this work, this value was found to be one. In other words, a 1NN classifier (known as nearest neighbor classifier) is used.

Classification strategies

The classification schemes sketched in Fig. 4 were employed in this study. As shown in one method, all features including ECG, PCG and clinical are given to a single KNN; in another strategy, each type of feature is fed to a separate single classifier and for each type of feature, SFS is executed to select the discriminative values of that set to the corresponding KNN.

Since the wrapper approach was used, the accuracy of the KNN classifier is fed back to the SFS algorithm[24,25] for selecting discriminative features to discriminate low risk from high risk cardiac patients.

In addition to the two classification techniques, another system was designed in this work such that the system uses a reduced number of uncorrelated features of ECG signal selected via the PCA and a KNN classifier to classify high risk patients from the low risk group.[26,27] The ECG-based MI risk classifier developed using PCA is called PCA-ECGMIRC and that developed using SFS algorithm is called SFS-ECGMIRC.

Two other classifiers use the 42 features extracted from the clinical data (i.e. patient's history, physical examination results and the laboratory data) to determine the risk of MI. As with ECG features, one classifier employs PCA for dimension reduction and the other one uses SFS for this task. The clinical data-based MI risk classifier that uses PCA is called PCA-CDMIRC and that use SFS method is called SFS-CDMIRC. No dimension reduction method was used for PCG data, since just five features are used. The MI risk classifier that uses only PCG features is called PCGMIRC.

The last two classifiers in which one uses both ECG and clinical features and the other uses all three groups of data (ECG signals, PCG waveforms and clinical findings) for classification task. Again, one classifier employs only PCA for dimension reduction and the other one uses SFS to select the most effective features. The MI risk classifier that uses ECG and clinical features along with PCA is called PCA-CD&ECGMIRC. The system that employs these two set of information, but uses the SFS algorithm is called SFS-CD&ECGMIRC. The MI risk classifier that uses all three groups of information (ECG, PCG and clinical data) and PCA as dimension reduction is called PCA-MIRC and that uses the same information (set of features), but the SFS method for dimension reduction is called SFS-MIRC. Therefore, 12 classifiers were developed and studied.

RESULTS

Considering the label provided by the cardiologists (experts) for each subject (feature vector) as the gold-standard (ground truth), the performance of the nine developed MI risk detection systems was evaluated in

Table 1. Performance of the Developed System in Determining the Risk of MI.

Method	Sensitivity	Specificity	Accuracy
PCGMIRC	56.8	76.9	66.3
CDMIRC	21.4	81.3	53.3
ECGMIRC	84.1	89.7	86.7
CD&ECGMIRC	93.2	94.9	94.0
MIRC	95.0	95.7	95.3
PCA–CDMIRC	35.7	81.3	60.0
PCA–ECGMIRC	81.8	84.6	83.1
PCA–CD&ECGMIRC	93.2	100.0	96.4
PCA–MIRC	95.4	100.0	97.6
SFS–CDMIRC	27.5	88.7	57.0
SFS–ECGMIRC	87.5	94.4	90.1
SFS–CD&ECGMIRC	97.7	100.0	98.8
SFS–MIRC	98.0	100.0	99.0

terms of correctly classifying patient with high risk of MI from those with low risk. Three performance indices were used for this purpose: sensitivity, specificity, and accuracy. These three indices are given by

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100,$$

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100,$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100,$$

where the parameters TP, TN, FP, and FN are defined as follows:

TP: Number of high-risk patients correctly identified as high risk.
TN: Number of low-risk patients correctly identified as low risk.
FP: Number of low-risk subjects incorrectly identified as high risk.
FN: Number of high-risk patients incorrectly identified as low risk.

The classification performances of the developed MI prediction systems are summarized in Table 1. The numbers presented in this table were obtained by running a Leave-one-out cross-validation on the collected feature vectors. To illustrate the effectiveness of the dimension reduction step, the results for three classifiers that use only raw features are reported in Table 1. In this table, ECGMIRC and PCGMIRC stand for the classifier that uses 41 ECG and 5 PCG features discussed in the previous sections, respectively; CDMIRC is the classifier that was developed using 42 clinical data features, CD&ECGMIRC is the model that uses both ECG and clinical data features, and finally MIRC is the classifier that uses all features extracted from ECG, PCG and clinical data.

DISCUSSION

Cardiac infarction is one of the leading causes of death and disability in developed countries. One of the most prevalence cardiovascular diseases is MI. Rapid identification of the subjects with the risk of MI leads to the early implementation of appropriate therapies that may save the patient's life. Therefore, preventing, early diagnosis of MI is critical in today's societies.

In recent years, there has been an explosion of interest in the application of pattern recognition and machine learning in health system, biomedicine, and biomedical engineering. This rapid rise has led to the development of several algorithms for biomedical signal processing, medical image processing, clinical decision support system and medical data analysis.[28–32] In this paper, we investigated the application of these techniques in categorizing patients with hard chest pain into high risk of MI or low risk of MI. Specifically, our objective was to develop an accurate machine learning-based system for identification of patient with high risk of MI. In general, evaluation results as presented in Table 1 show that the developed system could be able to identify such patients and hence using such system could be effective.

In terms of feature selection technique, the obtained results show the importance of the SFS in improving the results. As shown, the systems developed using the SFS algorithm outperformed the other methods. The PCA-MIRC that used PCA for dimension reduction provided the same specificity as SFS-MIRC, but the former provided the lower sensitivity and ultimately the lower accuracy. The main reason that the SFS-based systems outperformed the PCA-based systems is that the PCA in an unsupervised dimension reduction system and the objective in PCA is keeping the variance of the data not classifier accuracy while the objective of the SFS feature selection strategy is maximizing classifier performance.

In terms of using information, neither clinical data nor ECG features nor PCG features are lonely enough for estimating the risk of MI. Using all sets of these features resulted in a system with encouraging performance. As shown in PCA-based systems, the system that uses both clinical data and ECG information resulted in around 12.0% improvement in sensitivity in comparison to the system that uses only ECG features and at least 54.0% in comparison to the system that uses only clinical data. When PCG features are added to

ECG and clinical data, the sensitivity was improved by around 2%. One can say that recording and extracting PCG features increase the computation time and cost of the system, therefore it might not worthwhile to use PCG at all. But, 2.0% improvement in a large number of potentially MI patients referring to the hospital is a significant number.

In terms of the discriminant (prediction) ability of the features used in this study, based on the results obtained using the SFS algorithm, for CD features, the most important feature is diastolic blood pressure from physiological examination, HDL and Na from Lab data, diabetes mellitus. For ECG, the important features are QRS amplitude of leads $V3, V2$, and $V6$; S amplitude of lead $V6, T$ amplitude of lead $V3$ and lead II; and finally S–T duration measured on lead II.

The specificity of the systems can be improved by increasing the value of K in KNN classifier, but increasing K causes a reduction in sensitivity. As the risk of classifying a patient with the risk of MI as normal is higher than labeling a normal one as high risk patient, in this study, the value of K that provided the highest sensitivity and accuracy was used in developing each classifier discussed in the text and summarized in Table 1.

We have compared the performance of the developed and studied method with a conventional method that was used for diagnosing MI, the extracted features were used in a PCA algorithm and then fed to an ANN with eight layers (sigmoid function). The process was repeated 100 times where an overall accuracy of 89.9% and standard deviation of 3.7 were obtained. It shows that the method developed and presented here has a better performance (97.6%) in comparison with the conventional method (89.9%).

CONCLUSIONS

Early accurate diagnosis of MI or even predicting the risk of MI can prolong patient's life and enhance the quality of living through appropriate treatment. In this paper, we proposed a model predicting the risk of MI with the help of ECG findings along with PCG signal, clinical history, physical examination and laboratory data. Such a system can also assist with managing the limitation of beds at CCU. To make this decision accurate, several features extracted from PCG, ECG and clinical features are fused in order to build a rich set of features for the decision maker. Due to the high variance of features, a KNN classifier is employed to locally decide on the test samples. Numerical and comparative results illustrate that neither clinical data nor ECG features nor PCG features are lonely enough for estimating the risk of MI. Employing features from different modalities can significantly enhance the performance such that the developed multimodal-based system overperformed previously single modal research approaches. The results are promising and show that the developed system can assist with determining the patients with the risk of MI among who reported chest pain.

ACKNOWLEDGMENT

The paper has been extracted from parts of the M.Sc. thesis of Maryam Zarrabi supported by the Research Council of Shiraz University of Medical Sciences (93-01-01-9029).

REFERENCES

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, de Ferranti S, Després J-P, Fullerton HJ, Howard VJ, Huffman MD, Isasi CR, Jiménez MC, Judd SE, Kissela BM, Lichtman JH, Lisabeth LD, Liu S, Mackey RH, Magid DJ, McGuire DK, Mohler ER III, Moy CS, Muntner P, Mussolino ME, Nasir K, Neumar RW, Nichol G, Palaniappan L, Pandey DK, Reeves MJ, Rodriguez CJ, Rosamond W, Sorlie PD, Stein J, Towfighi A, Turan TN, Virani SS, Woo D, Yeh RW, Turner MB, Executive summary: Heart disease and stroke statistics — 2016 update: A report from the American Heart Association, *Circulation* **133**:447, 2016.
2. Kasper DL, Fauci AS, Hauser S, Longo D, Jameson JL, Loscalzo J, Harrison's Principles of Internal Medicine, 19th edn, McGraw-Hill Education/Medical, Columbus, Ohiao, 2015.
3. Menown I, Mackenzie G, Adgey A, Optimizing the initial 12-lead electrocardiographic diagnosis of acute myocardial infarction, *Eur Heart J* **21**(4):275, 2000.
4. Moses D, A survey of data mining algorithms used in cardiovascular disease diagnosis from multi-lead ECG data, *Kuwait J Sci* **42**(2):206, 2015.
5. Phanphaisarn W, Roeksabutr A, Wardkein P, Koseeyaporn J, Yupapin P, Heart detection and diagnosis based on Ecg and EPcg relationships, *Med Devices* **4**:133, 2011.
6. Baxt WG, Analysis of the clinical variables driving decision in an artificial neural network trained to identify the presence of myocardial infarction, *Ann Emerg Med* **21**(12):1439, 1992.
7. Yang TF, Devine B, Macfarlane PW, Use of artificial neural networks within deterministic logic for the computer ECG diagnosis of inferior myocardial infarction, *J Electrocardiol* **27**:188, 1994.
8. Hedén B1, Ohlsson M, Rittner R, Pahlm O, Haisty WK Jr, Peterson C, Edenbrandt L, Agreement between artificial neural networks and experienced electrocardiographer on electrocardiographic diagnosis of healed myocardial infarction, *J Am Coll Cardiol* **28**(4):1012, 1996.

9. Haraldsson H, Edenbrandt L, Ohlsson M, Detecting acute myocardial infarction in the 12-lead ECG using Hermite expansions and neural networks, *Artif Intell Med* **32**(2):127, 2004.
10. Choi S, Jiang Z, Comparison of envelope extraction algorithms for cardiac sound signal segmentation, *Expert Syst Appl* **34**(2):1056, 2008.
11. Sa-Ngasoongsong A1, Kunthong J, Sarangan V, Cai X, Bukkapatnam ST, A low-cost, portable, high-throughput wireless sensor system for phonocardiography applications. *Sensors* **12**(8):10851, 2012.
12. Price W, Brown A, Alterations in intensity of heart sounds after myocardial infarction, *Br Heart J* **30**(6):835, 1968.
13. Adolph RJ, Stephens JF, Tanaka K, The clinical value of frequency analysis of the first heart sound in myocardial infarction, *Circulation* **41**(6):1003, 1970.
14. Fatemian SZ, A wavelet-based approach to electrocardiogram (ecg) and phonocardiogram (pcg) subject recognition, Master's thesis, University of Toronto, 2009.
15. Ahlström C, Nonlinear phonocardiographic signal processing, Ph.D. thesis, Linköping University, Sweden, 2008.
16. Noorzadeh S, Extraction of fetal ECG and its characteristics using multi-modality, Ph.D. thesis, Université Grenoble, Alpes, 2015.
17. May AL, Kuklina EV, Yoon PW, Prevalence of cardiovascular disease risk factors among US adolescents, 1999–2008, *Pediatrics* **129**(6):1035, 2012.
18. Arif M, Malagore IA, Afsar FA, Detection and localization of myocardial infarction using K-nearest neighbor classifier, *J Med Syst* **36**(1):279, 2012.
19. www.dgdena.com.
20. Mazomenos E, Chen T, Acharyya A, Bhattacharya A, Rosengarten J, A time-domain morphology and gradient based algorithm for ECG feature extraction, *Proc IEEE Int Conf Ind Technol(ICIT)*, p. 117, 2012.
21. Welch PD, The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Trans Audio Electroacoust* **15**(2):70, 1967.
22. Mallat SG, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans Pattern Anal Machine Intell* **11**(7):674, 1989.
23. Fix E, Hodges Jr, Discriminatory analysis-nonparametric discrimination: Consistency properties, *Int Stat Rev* **57**(3):238, 1989.
24. Liu H, Motoda H, *Feature Selection for Knowledge Discovery and Data Mining*, Kluwer Academic Publishers, Norwell, MA, 1998.
25. Llamedo M, Martínez JP, Heartbeat classification using feature selection driven by database generalization criteria, *IEEE Trans Biomed Eng* **58**(3):616, 2011.
26. Zong R, Zhi Y, Yao B, Gao J, Stec AA, Classification and identification of soot source with principal component analysis and back-propagation neural network, *Aust J Forensic Sci* **46**(2):224, 2014.
27. Lópeza M, Ramíreza J, Górriza JM, Álvareza I, Salas-Gonzaleza D, Segoviaa F, Chavesa R, Padillaa P, Gómez-Ríob M, Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease, *Neurocomputing* **74**(8):1260, 2011.
28. Chiu CC, Linben TH, Liau BY, Using correlation coefficient in ECG waveform for arrhythmia detection, *Biomed Eng Appl Basis Commun* **17**:147, 2005.
29. Parsaei H, Stashuk D, EMG signal decomposition using motor unit potential train validity, *IEEE Trans Neural Syst Rehabil* **21**(2):265, 2013.
30. Parsaei H, Nezhad FJ, Stashuk DW, Hamilton-Wright A, Validation of motor unit potential trains using motor unit firing pattern information, *Proc 31st IEEE Annu Int Conf Eng Med Biol Soc(EMBS)*, p. 974, 2009.
31. Parsaei H, Stashuk DW, Adaptive motor unit potential train validation using MUP shape information, *Med Eng Phys* **33**:581, 2011.
32. Amiri S, Movahedi MM, Kazemi K, Parsaei H, 3D cerebral MR image segmentation using multiple-classifier system, *Med Biol Eng Comput* **55**(3):353, 2017.