Biomedical Signal Processing and Control 57 (2020) 101747

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Automated myocardial infarction identification based on interbeat variability analysis of the photoplethysmographic data

Abhishek Chakraborty, Deboleena Sadhukhan, Saurabh Pal, Madhuchhanda Mitra*

Department of Applied Physics, University of Calcutta, Kolkata, West Bengal, India

ARTICLE INFO

Article history:
Received 16 May 2019
Received in revised form 8 October 2019
Accepted 26 October 2019

Keywords:
Automated diagnosis
Myocardial infarction (MI)
Photoplethysmography (PPG)
PPG derivatives
Feature extraction
Variability analysis
Expert health systems

ABSTRACT

Background and objective: Myocardial infarction (MI) remains a major cause of mortality around the world for decades. Timely detection followed by instant medical intervention is strongly recommended to minimize MI related death threats. Generally, electrocardiogram (ECG) based automated techniques are preferred to ensure early diagnosis of MI. Recently, Photoplethysmogram (PPG) signal is evolving as a promising diagnostic tool for several cardiac monitoring applications because of its low cost, reliable and easy acquisition technology. However, the use of PPG for MI detection has not been much explored till date. Hence, in the present work, a method for MI detection is proposed based on the use of the PPG signal only.

Method: The pathophysiological alteration due to MI induces a beat-to-beat variation in the PPG beat morphology. Different time-plane parameters from the PPG signal and its derivatives are extracted to represent the individual beat morphologies and their variations are then analyzed as features over the cardiac cycles. By using different feature optimization techniques, finally five features are selected that presents discriminating variability to identify MI.

Results: The proposed method is evaluated with PPG records collected from 52 hospitalized MI patients and 52 normal subjects. The optimized five feature set, representing the inter beat morphological variations exhibits significant performance along with SVM (linear) classification technique with an average sensitivity, positive predictivity and detection accuracy of 92.70 %, 98.10 %, and 95.40 %, respectively.

Conclusions: Compared to other related works, use of PPG signal for MI detection is studied in this research for the first time. The promising result obtained establishes the utility of PPG signal for MI detection with a potential of implementation in the personal healthcare systems.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Myocardial infarction (MI), commonly known as heart attack, is one of the most fatal cardiovascular diseases (CVDs), causing permanent damage to the cardiac muscles following occlusion of the coronary arteries. Till date, MI remains a critical health-care issue and considered to be a leading cause of mortality around the world. According to the estimation of World Health Organization (WHO), cardiovascular diseases (CVDs) are responsible for 31 % of the global deaths in 2016 and out of which 85 % deaths are caused by heart attack and stroke [1].

Myocardial infarction (MI) occurs due to full or partial blockage of the coronary arteries. The blockage in the arteries then interrupts the blood circulation and due to the lack of necessary oxygen and nutrients, the heart muscle (myocardium) begin to lose its perfusion. Ultimately, the affected myocardial cells stop functioning and experiences irreversible episodes of injury, and necrosis to cause heart failure [2,3].

According to the physicians, timely detection of the blockage followed by immediate hospitalization of the patient is critically important to minimize MI related mortality rate [4]. Although, in most of the remote rural areas, the primary medical services required to deal with cardiovascular complications are still found to be widely compromised due to inadequate medical infrastructure and limited availability of trained medical personnel [5]. As a result, development of different automated MI detection techniques are evolving as a necessary solution to avoid sudden death threats and to save a huge amount of post-treatment expenditure, especially for the underprivileged populations around the world. These tech-

* Corresponding author.
E-mail addresses:
acaphy_rs@caluniv.ac.in (A. Chakraborty), dsaphy_rs@caluniv.ac.in (D. Sadhukhan), spaphy@caluniv.ac.in (S. Pal), mmaphy@caluniv.ac.in (M. Mitra).

https://doi.org/10.1016/j.bspc.2019.101747
1746-8094/© 2019 Elsevier Ltd. All rights reserved.

niques are capable of handling a large number of bio-signal records with minimum infrastructure and can be used as a promising tool for early diagnosis of cardiac diseases. Most of the automated techniques use ECG signal for its analysis because it is the most popular noninvasive method that has been clinically approved.

Electrocardiogram (ECG) is a reliable and noninvasive way to represent the electrical response of the heart muscles. Standard ECG system uses twelve leads to identify the signature of MI and to locate the infraction zones from multiple angles [6]. After MI, anatomical alterations of the infracted heart muscle introduces abnormal changes into the ECG waveform, which appears in the form of elevated or depressed ST segment, tall T-waves, flattened T-waves, inverted T-waves or biphasic T-wave and variation of Q-wave etc [7]. Numerous researches have been proposed till date that uses different computerized methods to analyze these changes in the ECG signal for MI detection [8–15]. Most of the proposed methods can be broadly categorized into two types: 1) morphological analysis of the ECG signal [8,9] and 2) analysis of the ECG signal using different transformation based techniques [10–13]. The algorithms proposed in [8,9], uses morphological analysis of the ECG signal to extract different time-domain features (e.g., Q-wave, QT interval, ST elevation or ST depression, and T-wave detection) through accurate detection of fiducial points. However, variant nature of ECG signal morphology under wide pathophysiological alteration and the presence of different types of noises often impose serious limitations for accurate detection of fiducial points and hence affect the signal analysis. Use of different complex signal processing techniques such as, wavelet transform [10–12] and Fourier transform [13] makes the signal analysis more computationally intensive for real-time implementation in an embedded platform. In some cases, hidden Markov model [14] and convolution Neural Network (CNN) based advanced methods [15] are also proposed for MI detection with reasonable accuracy.

However, the use of ECG for cardiac monitoring applications is imposed with certain limitations. 1) Attachment of multiple electrodes on different body parts, use of gel and associated electrical arrangements, might hamper the patient comfort and mobility; 2) ECG system with reduced number of leads often fail to establish the presence of MI; 3) ECG often shows normal or semi normal trace for 20 % patients, diagnosed with MI [16]; 4) different non-cardiac conditions also introduces abnormalities in the ECG trace [17]; 5) ECG based automated analysis often uses complicated signal processing techniques [10,11,14,15], high feature dimension [8,9,11] and intensive classification techniques [8,9,14,15], which limits the application of ECG on low cost, battery operated, personal health monitoring devices.

Now a day, Photoplethysmogram (PPG) is evolving as a potential diagnostic tool because of its effortless, portable, cost-effective acquisition technology and expert independent operational simplicity, which allows for both patient mobility and comfort. Photoplethysmography is a non-invasive electro-optical technique, used to detect relative blood volume changes in the peripheral body sites (i.e. finger tip, earlobe, and toe). Recent PPG acquisition module uses light emitting diodes (LED) to illuminate the tissues and a photodiode to measure the amount of reflected or transmitted light. Inherently, PPG signal consist of a pulsatile ac component (signifies average blood volume change) superimposed on a quasi-dc component (representing respiration and sympathetic nervous activity etc.). The onset of each pulse indicates the commencement of blood ejection from the heart to the aorta. The end of ejection is manifested in the form of dicrotic notch, indicating closure of the aortic valve [18] as shown in Fig. 1(a). Although, PPG signal morphology is found to be sensitive to age, gender and pathophysiological changes [18]. In such cases effective signal enhancement and quantification are carried out using first (FDPPG) and second derivative (SDPPG) of the PPG signals with minimal computational complexity. Usually, FDPPG wave consist of a maximum peak, a minimum peak and a local extreme point, whereas in general four systolic waves ('a', 'b', 'c' and 'd' waves) and one diastolic wave ('e' wave) related to the dicrotic notch are found to be associated with the SDPPG waveform as illustrated in Fig. 1(b) and Fig. 1(c) respectively [19].

Photoplethysmogram (PPG) carries potential information related to the cardiovascular events and the blood circulatory system. As a result, apart from its conventional use for the measurement of blood oxygen saturation, PPG signal has now been successfully employed in the analysis of several cardiovascular parameters, such as, heart rate [20], blood pressure (BP) [21], cardiac outputs (CO) [22] and also for the detection of cardiac abnormalities such as, Atrial fibrillation [23] and Premature Ventricular Contractions [24] etc. Moreover, owing to its portability, unobtrusive nature and inexpensive acquisition technology, PPG signal has been successfully integrated with several wearable devices [25] to ensure easy and early diagnosis of diseases at the comfort of home.

The full or partial blockage of the coronary arteries due to MI is responsible for inadequate blood flow to the heart muscles; which eventually lead to their damage [2,3]. This in turn affects the pumping ability of the heart and the ejected blood volume rate also. The resultant change in the blood volume is then reflected in the PPG signal morphology. Regardless of this fact, only some preliminary research work is reported so far, that has shown the utility of PPG signal for MI detection [26–29]. The method in [26] uses only five PPG signal data to show that the estimated spectral parameters and HRV parameters is different for normal and cardiovascular patients. In [27], only twelve PPG records are used and a statistical analysis is carried out to indicate that the mean value and distribution of the relative crest time is significantly different for the cardiovascular patients compared to the healthy volunteers. Different amplitude-interval features of the PPG signal second derivative (SDPPG) are used in [28,29], to differentiate MI patients from normal. All the above mentioned works present only primary research to establish the fact that, PPG signal features can be applied as a basic non-invasive tool to indicate the presence of MI. Although, there are certain inherent drawbacks of the above mentioned researches as mentioned below:

1) The amplitude and duration features that has been used in the above mentioned literatures are also affected by other intrinsic physiological parameters such as, ageing, gender, vascular conditions, arterial stiffness [18] etc. This in turn might affect the discriminating ability and robustness of these features to indicate the presence of MI only. 2) Moreover, most of the above mentioned algorithms are validated over smaller dataset only and also 3) none of the work tried to develop a complete implementable MI detection model utilizing suitable feature optimization and classification techniques.

Basically, MI induces imbalance in the cardiac cycle which causes interbeat variation in the morphology of the cardiac signal. Till date, several researchers have shown [30–35] interbeat variability of the ECG signal features as an alternative way out to indicate the presence of MI. Those researches use different indices such as, variability of the QRS complex [30,31] variability of the Q–T interval [32,33] and the T wave amplitude [33] for efficient quantification and characterization of MI. In addition, beat-to-beat variability analysis of the ventricular repolarization using T-wave spectral variance (TSV) [34] and T-wave spatial variance (SVT) [35] are also reported to study MI related changes. Outcome of the above mentioned literatures are indicative of the fact that, MI induces beat-to-beat variability in the ECG signal morphology. ECG represents the electrical activity of the heart as a function of its pumping action. Any disruption in the cardiac cycle due to MI

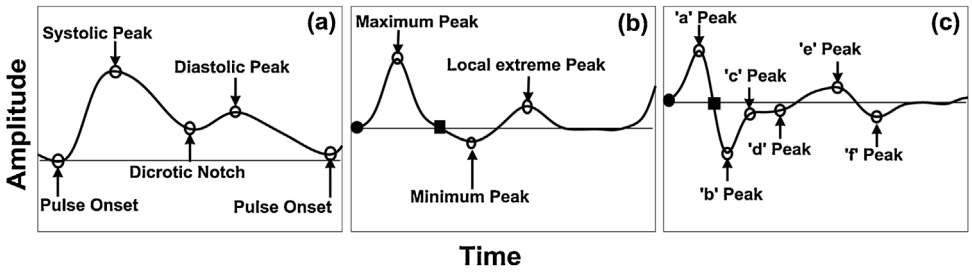

Fig. 1. Illustrative description of the a) PPG, (b) FDPPG and (c) SDPPG signals with all characteristic points. The first and the second zero crossing points in the FDPPG and SDPPG signals are indicated by filled circle and square symbols respectively.

related changes will have an adverse effect on the pumping action of the heart and also on the blood ejection rate. The changes in the effective blood volume on every cycle due to MI are then found to be embedded on beat-to-beat morphological features of the PPG signal.

In this research, the complicacy realated to ECG based techniques is avoided and a complete PPG based method is proposed to re-establish the utility of the PPG signal for MI detection. The use of PPG signal ensures easy implementation on home monitoring applications to facilitate preliminary level MI diagnosis at the comfort of home. In the proposed method, beat-to-beat PPG feature variability is taken as features for the first time to present discriminating separation between normal and MI class towards the development of a complete, classifiable MI detection model. To the best of our knowledge, beat-to-beat variability information of the PPG signal features to study MI related changes has not been reported yet. Compared to the absolute PPG features, which are affected by other physiological parameters, interbeat morphological variation of the features extracted from the PPG signal and its derivatives are found to be solely indicative of the imbalance created due to MI. Consequently, the main aim of the present study is to develop a complete, high performing and automated MI detection algorithm, based on the variability information of the PPG parameters only, which can be efficiently applied beyond the hospital boundary as an alternative for preliminary and accurate diagnosis of MI.

The main contributions of the proposed method are summarized as follows.

1) Instead of ECG, the proposed method uses only PPG signal to indicate the presence of MI. The use of PPG also ensures easy implementation on the device level for home monitoring application.
2) In order to indicate the presence of MI, beat-to-beat variance values of each of the extracted parameter from the PPG signal and its derivatives are used as features in this method. The used variance values of the parameters are much more robust compared to the absolute or mean value of the time-plane parameters. Moreover, it is worth mentioning that to the best of our knowledge, variability of the PPG signal parameters is reported in this research for the first time for MI detection.
3) After extensive evaluation using several machine learning methods, the proposed method presents promising results using SVM (linear) classifier and with five distinctive features. The promising results obtained in this study have laid the foundations of effective use of the PPG signal variability for easy MI screening.
4) Till date, no standard, annotated PPG signal database is available for public use. Hence, the proposed algorithm is evaluated on the PPG signal data, collected from fifty two healthy subjects and from fifty two MI patients admitted in the hospital.

The rest of this paper is organized as follows. Methodology of the proposed algorithm along with a description of the used database is described in section 2. Performance analysis of the proposed algorithm and its comparison with the other relevant works are presented in section 3. A complete discussion about the robustness and efficiency of the algorithm is presented in section 4. Finally, the conclusion is drawn in section 5 along with possible future work prediction.

2. Methodology

Detail block diagram of the proposed MI detection technique is illustrated in Fig. 2. The proposed algorithm can be divided into four major sections: i) fiducial point detection from the PPG signal and its derivatives followed by time-plane parameter extraction; ii) computation of variance value for each of the extracted parameters over a fixed number of beats to form a feature set; iii) optimization of the extracted feature set by combing the results obtained from the feature selection methods; and finally, iv) classification of the normal and MI patients using suitable classifier with optimized performance.

2.1. PPG Data acquisition

Since the use of PPG for cardiac monitoring is comparatively a new field of research, no standard database is available for MI affected PPG data. So for the development of the proposed technique, real PPG data was collected from actual subjects using BIOPAC MP 45 at 250 Hz sampling frequency. During data acquisition, all the participants were requested to stay calm in supine position and PPG signal data was collected from the left hand index finger for the duration of 5 min. Prior to data collection, all the healthy subjects were advised to avoid any kind of exercise, caffeinated or alcoholic beverage and smoking and were asked to remove any kind of metallic rings or electronic gadgets. Fifty two healthy male volunteers (with age ranging from 25 to 65) with no medical history of any cardiac diseases were participated in the data collection process at the institutional laboratory. PPG signal data was also collected from fifty two male MI patients (with age ranging from 35 to 70), admitted in the department of cardiology, Medical College and Hospital, Kolkata. Before data collection at the hospital, medical records of each patient was examined to confirm about the type of MI they are diagnosed with (e.g. Anterior, Antero-septal, Inferior, Infero-lateral and Infero-posterior). The entire research work was approved by the ethics committee of the hospital (vide approval no. MC/Kol/IEC/Non-spon/314/05–2016 dated 25 June 2016) and was executed under trained medical supervision. Also, as per institutional policy and international protocol, written informed consent was taken from each participant after explaining the entire procedure in detail.

Demographical details of all the participants including age, height, weight and blood pressure information are listed in Table 1.

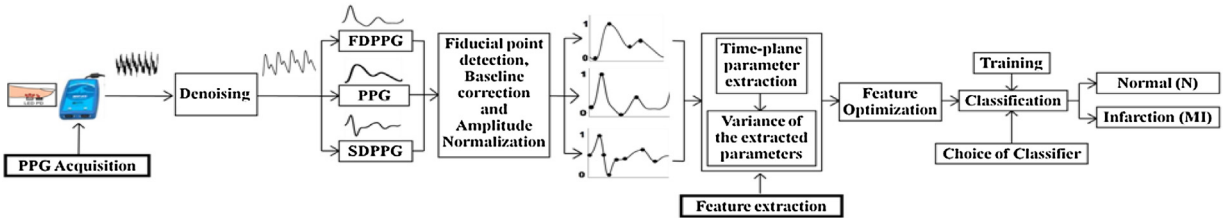

Fig. 2. Detailed block schematic of the proposed methodology. After pre-processing, different time-plane parameters are extracted from the PPG signal and its derivatives via accurate detection of fiducial points. The variance values of the extracted parameters over a fixed number of beats are taken as features. The obtained feature set is then minimized and finally an optimum classification technique is chosen with least number of features to identify the presence of MI.

Table 1
Demographic detail of all the subjects.

Description	Normal	MI
Number of subjects	52	52
Age in year (mean ± standard deviation)	30.691 ± 8.219	51.961 ± 9.822
Height in meter (mean ± standard deviation)	1.527 ± 0.226	1.581 ± 0.129
Weight in kg (mean ± standard deviation)	69.357 ± 11.768	71.843 ± 6.795
Systolic blood pressure in mmhg (mean ± standard deviation)	121.359 ± 19.486	115.216 ± 15.23
Diastolic blood pressure in mmhg (mean ± standard deviation)	79.724 ± 12.364	74.333 ± 8.926

As most of the MI patient participated in the present study are aged person, the fact is reflected in the age information of Table 1. Apart from the age information, other parameters are found to be compatible between two groups.

2.2. Preprocessing

PPG signal is influenced by numerous factors such as, light at the photo-detector, amount of pressure variation on the sensor tip, poor blood perfusion of the peripheral tissues, presence of motion artifacts and even the position of the PPG sensor with respect to the heart [36,37]. These factors may cause signal deterioration at various steps of signal acquisition and hence affect the signal analysis. In the present work, PPG data acquisition was carried out with utmost care to minimize the above mentioned effects, especially the motion artifact. The use of derivative based technique makes the algorithm more sensitive to the presence of high frequency noises. Therefore, the digitally acquired PPG signals are initially denoised using a 6th order Butterworth low-pass filter with a cut-off frequency of 15 Hz, to remove any high frequency noise components and to retain essential clinical bandwidth of the PPG signal.

2.3. Fiducial point detection

Accurate detection of fiducial points followed by time-plane parameter extraction has been considered as the most crucial step for automated analysis of any bio-signal. A robust, accurate yet simple algorithm is proposed for accurate detection of fiducial points from the PPG signal and its derivatives. Initially, the algorithm identifies the fiducial points of the first (FDPPG) and second derivative (SDPPG) of the PPG signal. Then, fiducial points of the PPG signal are identified using the fiducial point locations of FDPPG and SDPPG respectively. Instead of using any computationally intensive transform technique, the proposed algorithm uses signal derivatives and simple mathematical techniques such as amplitude threshold, slope reversal and an empirical formula based approaches. Robustness of the proposed fiducial point detection algorithm is verified over a wide pathophysiological variety with significant results. Detail methodologies about the fiducial point detection and parameter extraction are presented in [38].

2.4. Baseline modulation correction and normalization

Usually most of the amplitude features are calculated with respect to a common reference value and any abrupt change in the signal baseline might lead to wrong feature information. However, considering different factors such as, irregular breathing, unusual body movement and unsteady finger tip contact on the sensor, baseline of the PPG signal often varies even within a single beat duration as can be seen from an example beat of Fig. 3(a). To eliminate this effect, the first pulse onset point is considered as the baseline point of each cardiac beat as indicated in Fig. 3(a). Then, all the samples in between the onset points are modified using an empirical formula as discussed in detail in [39]. After baseline correction, the corrected PPG signal and its derivatives are also subjected to amplitude normalization in the range 0–1 using Eq. 1.

$$V_{norm} = \left(\frac{V_{signal} - V_{signal(\min)}}{V_{signal(\max)} - V_{signal(\min)}} \right) \quad (1)$$

Here, V_{signal} = baseline corrected PPG and FDPPG/SDPPG signal respectively, $V_{signal\,(min)}$ = minimum value of the baseline corrected PPG and FDPPG/SDPPG signal respectively, $V_{signal\,(max)}$ = maximum value of the baseline corrected PPG and FDPPG/SDPPG signal respectively and V_{norm} = Normalized PPG and FDPPG/SDPPG signal. A typical PPG signal after baseline modulation correction and amplitude normalization is shown in Fig. 3(b).

2.5. Feature extraction

In the recent days, different characteristic parameters of the PPG signal are used to estimate different physiological parameters and abnormalities. A detail characterization of the SDPPG signal is given in [40]; along with the definitions of different SDPPG feature. In addition, arterial stiffness estimation is carried out in [41], based on PPG systolic, diastolic and dicrotic notch related parameters. A new set of PPG, FDPPG and SDPPG parameters are proposed in [42] to improve the accuracy of the cuffless and continuous BP estimation. Biometric authentication and hemoglobin level prediction is done in [43,44] respectively, using PPG, FDPPG and SDPPG parameters. Even

Intradialytic Hypotension (IDH) episodes are also characterized using different time domain PPG and FDPPG parameters in [45].

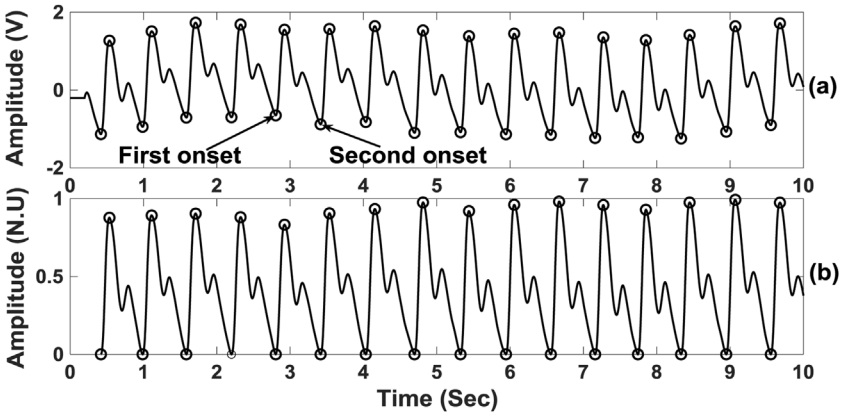

Fig. 3. (a) A typical PPG record before baseline modulation correction and amplitude normalization and (b) the same PPG record after baseline modulation correction and amplitude normalization.

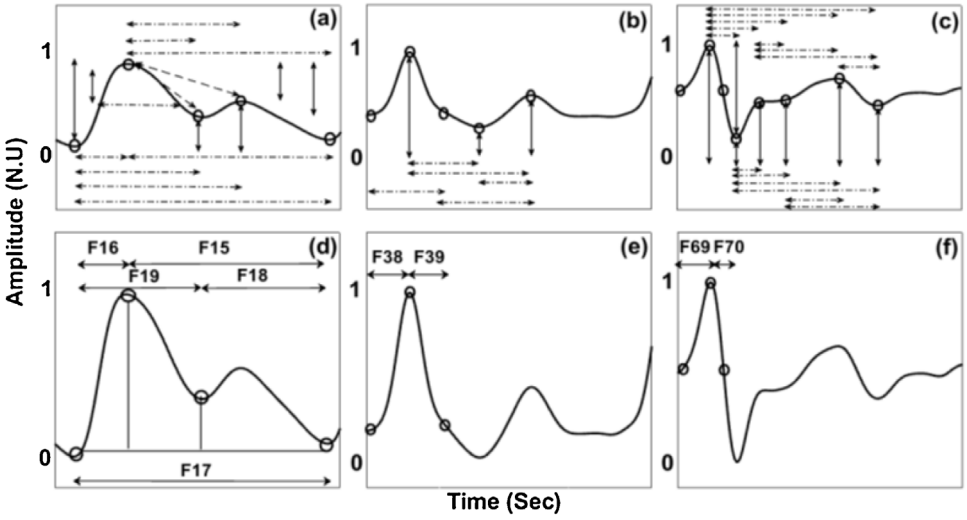

Fig. 4. Detail features of the (a) PPG, (b) FDPPG and (c) SDPPG signals with all detected characteristic points. Solid lines in the pictures indicate the amplitude features, dashed lines indicate the slope features, and dot-dash lines indicate the duration features. Area feature zones of (d) PPG, (e) FDPPG and (f) SDPPG are marked with solid lines in the pictures.

Hence we extracted all the standard PPG, FDPPG and SDPPG signal parameters, used in different applications [41–45] to ideally quantify the individual beat morphologies. Details of the extracted features are presented in Fig. 4 and also listed in Table 2 respectively.

Myocardial infarction causes rapid and irreparable damage to a portion of the heart muscles, which in turn affects the pumping action of the heart with an imbalance in the cardiac cycle [46] as well. The resultant imbalance in the pumping action of the heart also induces irregularities in the blood ejection rate from the left ventricle. Consequently, it is hypothesized that this imbalance will induce a beat-to-beat variability in the PPG signal cycle also. Generally, absolute or mean values of the extracted morphological parameters are found to be affected by different intrinsic factors such as, heart rate variations, age, vascular conditions, even body temperature [18] and are less reliable for accurate characterization of MI. Whereas, inter-beat variation among the PPG signal beats arises exclusively due to the imbalance introduced into the cardiac cycle as a result of infarction. Hence, in this work, instead of absolute or mean values of the morphological parameters, variation of the all the individual beat parameters, extracted from the PPG signal and its derivatives are studied to indicate the presence of MI.

Beat to beat variations of three selected parameters of the PPG, FDPPG and SDPPG signals are presented as an example in Fig. 5 for both normal and MI records to reveals the fact that, compared to the normal record, variation of the parameter is significantly higher for the MI record.

Generally, the variance is defined as a measurement of the deviation from the mean value in a data set and is expressed by the following equation:

$$VOF = \left(\frac{1}{N-1} \sum_{i=1}^{N} \left| f_i - \bar{f} \right|^2 \right) \tag{2}$$

where

$$\bar{f} = \frac{1}{N} \left(\sum_{i=1}^{N} f \right)$$

Here, f_i denotes the amplitude/time parameter value of the i^{th} PPG/FDPPG/SDPPG data cycle considering a total of N number of cycles and $\bar{f}$ represents the mean of f for N number of cycles. In the present research, variance values of 72 different morphological beat parameters, calculated over 30 PPG data cycles are used as features.

Table 2
Summary of the features used in the proposed method.

PPG signal features			
Features used	Descriptions	Features used	Descriptions
F1: ttc	1 st onset to systolic peak duration	F16: atcppg	1st onset to systolic peak area
F2: ttn	1 st onset to notch duration	F17: attppg	1st onset to 2nd onset area
F3: ttdp	1 st onset to diastolic peak duration	F18: antppg	Notch to 2nd onset area
F4: tstrtftr	1 st onset to 2nd onset duration	F19: atnppg	1st onset to notch area
F5: tcrtdn	Systolic peak to notch duration	F20: didppg	Systolic peak to 2nd onset amplitude difference
F6: tcrtdp	Systolic to diastolic peak duration	F21: aidppg	Systolic peak to 1st onset amplitude difference
F7: tctst	Systolic peak to 2nd onset duration	F22: mrslppg	1st onset to systolic peak rising slope
F8: thw	Pulse width	F23: mrsrppg	Systolic peak to 2nd onset falling slope
F9: vsp	Systolic peak amplitude	F24: slp1	Systolic peak to notch slope
F10: vdn	Dicrotic notch amplitude	F25: slp2	Systolic peak to diastolic peak slope
F11: vdp	diastolic peak amplitude	F26: spoc	Systolic peak output curve
F12: vhwh	Systolic peak half amplitude	F27: dpdc	Diastolic peak downward curve
F13: AI	Augmentation index	F28: IAR	Inflection point area ratio
F14: AAI	Alternative augmentation index	F29: rtftr	Ratio of ttc and tctst
F15: actppg	Systolic peak to 2nd onset area	–	–
PPG first derivative (FDPPG) features			
F30: tfdp1tp2	Maximum to minimum peak duration	F36: vfdp2	Minimum peak amplitude
F31: tfdp1tp3	Maximum peak to local extreme peak duration	F37: vfdp3	Local extreme peak amplitude
F32: tfdp2tp3	Minimum peak to local extreme peak duration	F38: actfdppg	1st zero crossing point to maximum peak area
F33: tfdfzctszc	1st to 2nd zero crossing point duration	F39: atcfdppg	Maximum peak to 2nd zero crossing point area
F34: tfdszctp3	2nd zero crossing to local extreme peak duration	F40: mrslfdppg	1st zero crossing point to maximum peak slope
F35: vfdp1	Maximum peak amplitude	F41: mrsrfdppg	Maximum peak to minimum peak slope
PPG second derivative (SDPPG) features			
F42: tsdp1tp2	"a" to "b" peak duration	F58: vsdp2	"b" peak amplitude
F43: tsdp1tp3	"a" to "c" peak duration	F59: vsdp3	"c" peak amplitude
F44: tsdp1tp4	"a" to "d" peak duration	F60: vsdp4	"d" peak amplitude
F45: tsdp1tp5	"a" to "e" peak duration	F61: vsdp5	"e" peak amplitude
F46: tsdp1tp6	"a" to "f" peak duration	F62: vsdp6	"f" peak amplitude
F47: tsdp2tp3	"b" to "c" peak duration	F63: bbya	Ratio of vsdp2 and vsdp1
F48: tsdp2tp4	"b" to "d" peak duration	F64: cbya	Ratio of vsdp3 and vsdp1
F49: tsdp2tp5	"b" to "e" peak duration	F65: dbya	Ratio of vsdp4 and vsdp1
F50: tsdp2tp6	"b" to "f" peak duration	F66: ebya	Ratio of vsdp5 and vsdp1
F51: tsdp3tp4	"c" to "d" peak duration	F67: AGI1	((vsdp2 - vsdp23 - vsdp4 - vsdp5) / vsdp1)
F52: tsdp3tp5	"c" to "e" peak duration	F68: AGI2	((vsdp2 – vsdp5) / vsdp1)
F53: tsdp3tp6	"c" to "f" peak duration	F69: actsdppg	1st zero crossing point to "a" peak area
F54: tsdp4tp5	"d" to "e" peak duration	F70: atcsdppg	"a" peak to 2nd zero crossing point area
F55: tsdp4tp6	"d" to "f" peak duration	F71: mrslsdppg	1st zero crossing point to "a" peak slope
F56: tsdp5tp6	"e" to "f" peak duration	F72: mrsrsdppg	"a" peak to "b" peak slope
F57: vsdp1	"a" peak amplitude	–	–

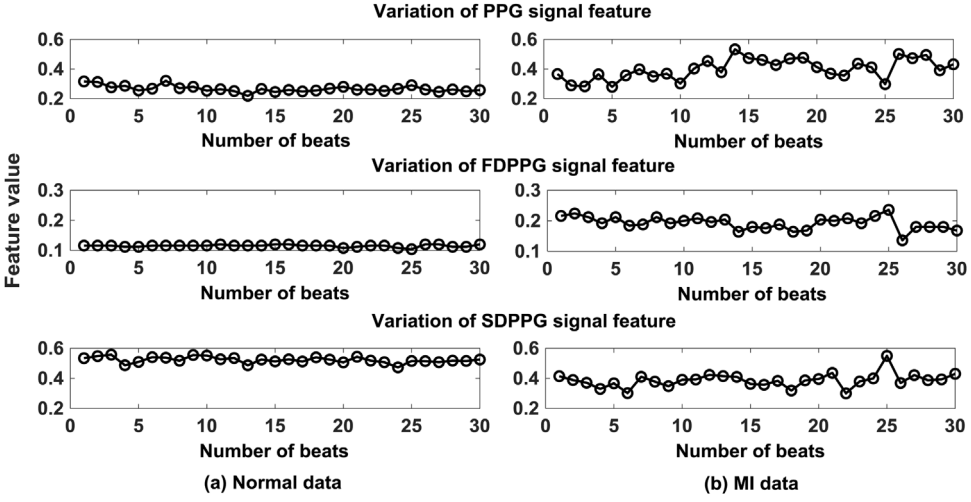

Fig. 5. Beat-by- beat variation of PPG, FDPPG and SDPPG signal features for a) Normal and b) MI records with respect to the number of beats are shown in the first, second and third window respectively. It is evident, that the variation of the parameters is significantly higher for the MI record compared to the normal.

2.6. Feature selection

In order to select the optimal feature set which shows high discrimination ability, four filter-based methods [gain ratio (GR), one-R (1R), relief F (RLF) and symmetrical uncertainty (SU)] and two wrapper-based [Naive bayes (NB) and logistic regression (LR)] feature selection methods are used and compared against each other to determine optimized feature subset. The steps involved in the feature selection are summarized as follows:

1) At first, the calculated variance data for each of the 72 parameters obtained from 52 MI and 52 normal subjects are applied at the input of each feature selection module.

2) Based on the input, each of the feature selection algorithms presents a feature subset in terms of ranking.
3) To facilitate mutual comparison among the selection algorithms, all the selected features by each of the algorithm are accumulated in a table.
4) If any feature is selected by any algorithm, a selection score is "1″ is assigned to that feature. Otherwise, a zero score is given to indicate that the feature is not selected by the corresponding algorithm.
5) Now, the selection scores of each feature are added to obtain a total score. The total score of each feature signifies the fact that how many times the feature is selected by the selection modules.

In order to present the selection result in a concise form, only those features with selection score greater than three are presented in Table 3. This causes an effective reduction of the feature dimension from 72 to 42, a shown in Table 3.

Table 3 reveals that, except for the two features (F44, F59), rest of the features presents the same total selection score. Hence, the results of Table 3 is combined with the box-plot for further optimization of the feature set to identify the features showing maximum variation in the values for the healthy and the MI data. Finally, based on the total selection score and the result obtained from the box-plot, only eight features, showing high discrimination ability are considered for the classification stage. The selected eight features are indicated by bold in Table 3 and the box plots of the selected eight features are shown in Fig. 6 respectively.

The mean and standard deviation (SD) of each of the selected eight features are computed over 30 beats for each of the normal and MI records and are presented in Table 4. It is evident from Table 4 that the variance value of each of the selected feature is higher for the MI record as compared to the normal records, which is quite in agreement with the proposed hypothesis. Significance of the selected eight feature set is also verified by two-sample *t*-test and the resultant significantly low p-value (<0.0001) of the selected features is also listed in Table 4.

2.7. Classification

The final stage of the proposed method involves the use of classification technique to design a suitable decision boundary for the separation of two different classes of normal and MI respectively. The mean, standard deviation and the p-values of each of the eight selected features, as presented in Table 4 clearly exhibits distinctive categorization between the normal and the MI class. Moreover, the median values of the box-plot of each features as depicted in Fig. 6 also demonstrates significant contrast between the two classes. To verify the robustness of the selected features, initially, all the selected eight features are applied at the input of six traditional state-of-art classification techniques [Decision Tree (DT), Quadratic discriminate (QD), Logistic regression (LR), Linear SVM (LS), Nonlinear SVM (NLS), k-Nearest Neighbor (kNN)]. In the following sections efficiency of these methods is described in detail.

3. Experimental results

Initially, 72 parameters are extracted from the PPG signal and its derivatives and the calculated variance value of each of these parameters are used as features in this research. Six different feature selection algorithms are used to reduce the feature dimension and to maximize the classification performance. Each feature selection algorithm provides a unique, optimized set of features in terms of ranking. Finally, the results obtained from each of the feature selection algorithm are combined and the importance of each feature is presented in terms of a total selection score instead of individual ranking as shown in Table 3. The total selection score is defined as the maximum frequency of a particular feature selected by the selection algorithms. The feature selection result is then combined with statistical analysis to generate an optimized feature set of eight features.

Overall performance evaluation of the proposed model is carried out by means of five different benchmark parameters. In addition to the conventional statistical evaluators such as, accuracy (Acc), sensitivity (Se), and specificity (Sp), the overall diagnostic ability of the proposed model is also assessed by mean of receiver operating characteristic (ROC) curve and area under the curve (AUC) respectively. Details of the performance evaluators are formulated as follows.

$$Sensitivity(Se) = \frac{TP}{TP + FN} \times 100\% \tag{3}$$

$$Predictivity(Sp) = \frac{TN}{TN + FP} \times 100\% \tag{4}$$

$$Accuracy(Acc) = \frac{TP + TN}{TP + FN + TN + FP} \times 100\% \tag{5}$$

Here TP = number of MI beats correctly detected as MI, TN = number of normal beats correctly identified as normal, FN is the number of MI beats incorrectly detected as normal, FP is the number of normal beats incorrectly detected as MI.

3.1. Classification result

The classification result obtained from each classifier is summarized in Table 5 and also shown in the ROC curve of Fig. 7. Finally, based on the results as presented in Table 5, and Fig. 7, support vector machine (linear) is chosen as an optimum classification technique in the present work. Support vector machine is a state-of-art supervised learning technique that searches for the optimized hyper plane to maximize the distance between the support vectors of two distinct classes. SVM has been successfully used in different PPG based researches to characterize hypovolemia, euvolemia [47] and type-2 diabetes [48] respectively.

Moreover, in order to save computational time and complexity, further optimization of the selected eight feature dimension is carried out by varying the number of features at the input of the chosen SVM classifier. The performance is listed in Table 6. Evidently, the classification performance is significant with seven selected features as shown in Table 6. However, for the sake of further optimization, only five features (as indicated by bold italics in Table 3) are chosen as the final features for efficient discrimination of normal and MI classes. The efficiency of the proposed classifier over the standard automated MI detection techniques is presented in the performance comparison section.

In this research, the proposed method uses variance value of all the parameters, calculated over thirty beats only to carry out faster analysis. Signal segment with shorter length often fail to reflect distinct feature variance and eventually affect the classification accuracy. However, the classification results of each of the classifier as presented in Table 5 and Table 6 indicate acceptable performance of the algorithm with 30 beats.

3.2. Performance evaluation

Statistical significance of the proposed model is verified using a generalized k-fold cross validation technique. Generally, the whole dataset is randomly divided into k number of equal sized groups in k-fold cross validation method. Then one group among the k-equal sized groups is chosen for testing purpose and the rest of the data groups are assigned to train the classifier. The whole process is repeated for k times considering varying random partitions of

Table 3
Feature selection results.

Feature selection algorithms	Features													
	F2	F3	F5	F6	F10	F11	***F13***	F14	F17	F21	F25	F27	**F31**	F32
Gain ratio	1	1	1	1	0	1	***1***	1	1	1	1	1	**1**	1
One R	1	1	1	1	1	1	***1***	1	1	1	1	1	**1**	1
Relief	1	1	1	1	1	1	***1***	1	1	1	1	1	**1**	1
Symmetrical uncertainty	1	1	1	1	1	1	***1***	1	1	1	1	1	**1**	1
Logistic regression	0	0	0	0	0	0	***0***	0	0	0	0	0	**0**	0
Naive bayes	0	0	0	0	1	0	***0***	0	0	0	0	0	**0**	0
Total Score	4	4	4	4	4	4	***4***	4	4	4	4	4	**4**	4

Feature selection algorithms	Features													
	F34	F37	F38	F41	**F43**	***F44***	F45	f46	**F47**	***F48***	F49	F50	F51	F52
Gain ratio	***1***	1	1	1	**1**	***1***	1	1	**1**	***1***	1	1	1	1
One R	***1***	1	1	1	**1**	***1***	1	1	**1**	***1***	1	1	1	1
Relief	***1***	1	1	1	**1**	***1***	1	1	**1**	***1***	1	1	1	1
Symmetrical uncertainty	***1***	1	1	1	**1**	***1***	1	1	**1**	***1***	1	1	1	1
Logistic regression	***0***	0	0	0	0	***1***	0	0	0	*0*	0	0	0	0
Naive bayes	***0***	0	0	0	0	***1***	0	0	0	*0*	0	0	0	0
Total Score	***4***	4	4	4	***4***	***6***	4	4	**4**	***4***	4	4	4	4

Feature selection algorithms	Features													
	F53	F54	F55	F56	***F59***	F61	F62	F64	F65	F66	F67	F68	F69	F72
Gain ratio	1	1	1	0	***1***	1	1	1	1	1	1	1	1	1
One R	1	1	1	1	***1***	1	1	1	1	1	1	1	1	1
Relief	1	1	1	1	***1***	1	1	1	1	1	1	1	1	1
Symmetrical uncertainty	1	1	1	1	***1***	1	1	1	1	1	1	1	1	1
Logistic regression	0	0	0	1	***1***	0	0	0	0	0	0	0	0	0
Naive bayes	0	0	0	0	***1***	0	0	0	0	0	0	0	0	0
Total Score	4	4	4	4	***6***	4	4	4	4	4	4	4	4	4

"1" indicates selection of the feature by the algorithm and "0" indicates no selection.
Selected eight features are shown in bold and the top five features are denoted by bold italics.

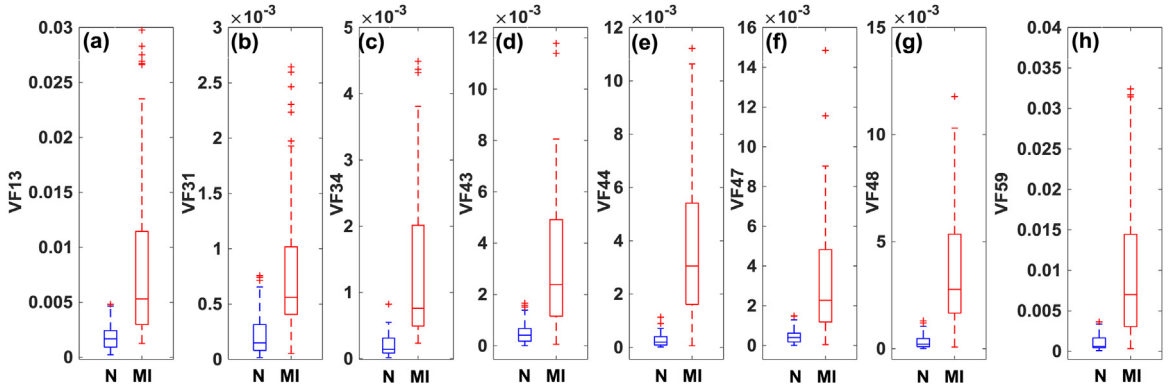

Fig. 6. Box plots of the selected eight features showing the statistical distribution for the Normal (N) and Myocardial infarction (MI) data. The central mark in each box denotes the median, the edges represent the 25th and 75th percentiles and the whiskers denotes the outliers. Variability of the features a) F13, b) F31, c) F34, d) F43, e) F44, f) F47, g) F48 and h) F59 are shown in the box plots respectively.

Table 4
Mean, standard deviation and p-value of the selected features.

Selected features	Normal		Infarction (MI)		p-value
	Mean	SD	Mean	SD	
F13	0.00178	0.00116	0.00883	0.00816	1.4189×10^{-8}
F31	0.00021	0.00019	0.00086	0.00071	6.1850×10^{-9}
F34	0.00020	0.00016	0.00137	0.00127	1.7328×10^{-9}
F43	0.00051	0.00045	0.00316	0.00274	5.1108×10^{-10}
F44	0.00030	0.00027	0.00370	0.00269	8.1890×10^{-15}
F47	0.00050	0.00041	0.00326	0.00302	2.6347×10^{-9}
F48	0.00033	0.00032	0.00359	0.00265	3.4596×10^{-14}
F59	0.00106	0.00091	0.01011	0.00942	4.4803×10^{-10}

Table 5
Classification result with eight selected features.

Classification algorithms	Acc (%)	Se (%)	Sp (%)	AUC (%)
Decision Tree (DT)	93.30	92.31	94.23	93.15
Quadratic discriminate (QD)	90.40	88.46	92.30	91.0
Logistic regression (LR)	90.40	90.38	90.38	90.2
Linear SVM (LS)	**95.20**	**92.31**	**98.1**	**95.16**
Nonlinear SVM (NLS)	94.23	92.31	96.15	94.15
k-Nearest Neighbor (kNN)	93.30	88.46	98.07	94.0

Table 6
Performance of the chosen SVM (linear) classifier with different number of selected features.

No of features	Acc (%)	Se (%)	Sp (%)
Eight	95.20	92.31	98.1
Seven	96.5	92.31	100
Six	94.2	92.31	96.15
Five	**95.2**	**92.31**	**98.1**

Table 7
Performance evaluations of the SVM (linear) classifier with five -fold cross validation.

No of folds	Acc (%)	Se (%)	Sp (%)
One	96.15	92.31	100
Two	96.15	92.31	100
Three	95.20	94.23	96.15
Four	95.20	92.31	98.10
Five	94.23	92.31	96.15
Average	95.40	92.70	98.10

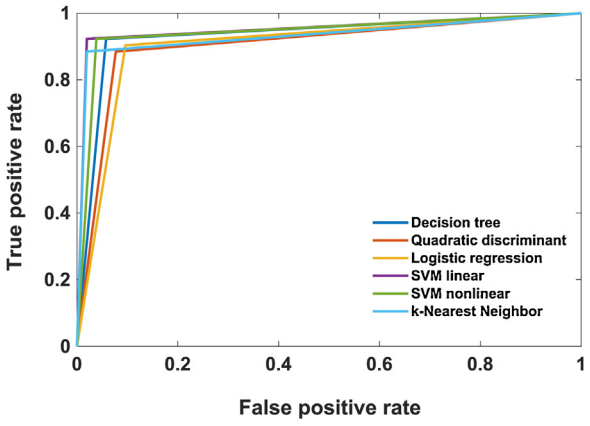

Fig. 7. The Receiver operating characteristic (ROC) curves obtained for different classification techniques used in Table 5 with eight selected features.

the dataset. In the present research, as the analysis is carried out with comparatively less number of records (52 normal and 52 MI patients), hence, only five fold cross validation is used for generalization of the proposed model and the result is shown in Table 7. Five-fold cross validation performance of the chosen SVM classifier with five selected features as presented in Table 7 can be summarized with average sensitivity, positive predictivity and detection accuracy of 92.70 %, 98.10 %, and 95.40 %, respectively. Furthermore, it is worth mentioning that all the classification results as listed in Table 5 and Table 6 are also based on fivefold cross validation method.

3.3. Validation of the classification technique

Initially, the proposed SVM (linear) classifier is chosen based on the performance with eight selected features. Then the selected classifier performance is optimized by varying the number of features at the input. Finally, optimized performance of the SVM classifier with five selected features is chosen for the final model. However, for further generalization, the chosen SVM classification technique with selected number of features is compared against other standard classification techniques, that are commonly used for MI detection. For each of the employed classification technique, fivefold cross validation technique is used for training and testing purpose. After evaluation over five fold, the average performance is listed in Table 8.

Table 8
Performance comparison for different classification techniques with five selected features.

Classification algorithms	Acc (%)	Se (%)	Sp (%)	AUC (%)
Decision Tree (DT)	93.3	92.31	94.23	93.09
Quadratic discriminate (QD)	92.3	88.46	96.15	93
Logistic regression (LR)	92.3	92.31	92.31	91.16
Nonlinear SVM (NLS)	95.2	92.31	98.1	95.16
k-Nearest Neighbor (kNN)	96.2	92.31	100	96.5
Proposed	95.40	92.70	98.10	96

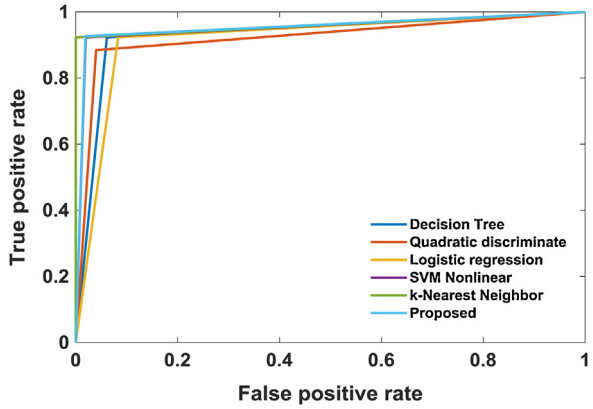

Fig. 8. The Receiver operating characteristic (ROC) curves obtained for different classification techniques used in Table 8 with five selected features.

It is evident from Table 8 that apart from the adopted SVM (linear) technique, other classifiers such as, k-nearest neighbor (kNN) with k = 10 and SVM (nonlinear) also presents superior or compatible performance. The noteworthy performance of the chosen classifier is also indicated in the receiver operating characteristic curves (ROC curve) of Fig. 8 with higher value of area under curves (AUC).

3.4. Performance evaluation in terms of training data size

The sole purpose of the present research is to establish the fact that PPG contains significant cardiovascular information, which can be used as a preliminary alternative to indicate the presence of MI. The proposed algorithm uses different standard classification techniques to quantify the discrimination ability of the extracted feature set. It can also be seen from Table 8 that more than one classifier presents compatible performance compared to the chosen SVM (linear) classifier. However, it is worth mentioning that the essence of the proposed research does not aim for optimization of the machine learning techniques. Furthermore, optimization against variable signal length and large data size is also beyond the scope of this research due to the unavailability of additional database and the use of fixed sample size. Therefore, based on the available data, accuracy and AUC performances of the proposed SVM (linear) classifier including other two high performing classification techniques [SVM (non-linear) and k-Nearest Neighbor] are presented in Fig. 9, against different percentage of training data size. In this case, no cross-validation techniques of the training and test

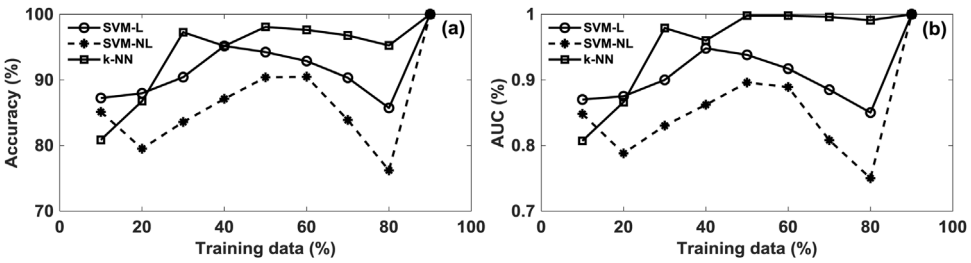

Fig. 9. Performance of three high performing classification techniques [SVM (linear), SVM (non-linear) and k-Nearest Neighbor] against the percentage of training dataset is presented in terms of (a) accuracy and (b) AUC plot respectively.

dataset are used to present the performance. It is evident from Fig. 9 that with increasing percentage of training data size, the k-Nearest Neighbor method presents more stable performance compared to other SVM based methods. However, considering the difficulties associated with the choice of 'k' values including the problems of computational burden and memory requirement, in this research the overall performance of the SVM classifier with linear kernel function is chosen to present the model.

3.5. Performance comparison

To the best of our knowledge, only a limited number of literatures are available that uses PPG based analysis for MI detection. Moreover, the proposed MI detection method, based on variability analysis of the PPG, FDPPG and SDPPG parameters has not been reported in any literature till date. Since, the primary motivation of the proposed research work is to investigate the potential of PPG signal for the detection of MI, the efficiency of the proposed technique are also justified by comparing the result with other ECG based automated MI diagnosis techniques. However, differences in used data sets, number of ECG leads, number of features used, validation process, differences in features types and the used classification tools of the mentioned literatures also present serious obstacle for exact comparison. A descriptive and reasonable comparison of the proposed algorithms with the available state-of-the art works on automated MI detection is presented in Table 9. The comparison report as listed in Table 9 reveals the fact that, compared to most of the reported ECG and PPG based methods, the proposed method presents significantly high detection accuracy using PPG signal only.

4. Discussion

In recent days, with the advancement in technologies, the necessity and demand of different automated health monitoring devices are increasing day by day to facilitate expert independent self diagnosis at the comfort of home. As a result, for the past two decades, most of the research attention is focused on towards the development of different computer assisted, automated techniques that uses the properties of different bio-signals (e.g. ECG, PCG and PPG etc) for non-invasive analysis of cardiac diseases. Till date, ECG has been used as a gold standard tool for the diagnosis of MI because of its high accuracy, reliability and versatility in the detection and characterization of MI from multiple angles. Although, starting from signal acquisition to analysis, multi-lead, automated ECG systems often come up with severe limitations, which includes complicated operator dependent arrangements, the use of high feature dimension (8, 9, 11), complicated feature extraction techniques (10, 11, 14, 15) and classification methodologies (8, 9, 14, 15). Therefore, ECG based methods are not at all suitable to be implemented on hand-held, battery operated portable devices with limited computational and memory resources. On the contrary, recent researches exhibit the fact that, the potential cardiovascular information embedded in the PPG signal can be used for the analysis of cardiac diseases [23,24,28,29] as well. However, it is important to note that the use of PPG signal for diagnosing MI is still on experimental level and has not been clinically accepted yet. Moreover, compared to ECG, PPG analysis can only suggest the presence of infarction but cannot provide a detailed analysis about the disease severity and the zone of infarction. Hence, PPG can only be used to provide a preliminary level screening of MI, especially on the home monitoring platforms, due to its low cost and easy acquisition techniques.

Until now, to the best of our knowledge, only few PPG signal based research [26–29] is carried out that uses conventional absolute parameters for the detection of MI. These features are highly affected by different subject specific intrinsic parameters and the methods are evaluated on limited number of test subjects only. In the proposed research, absolute value of the PPG features, extracted from the normal and MI patients are also investigated. However, the discrimination ability of these features is found to be less prominent due to strong functional dependence on different pathophysiological parameters (such as, age, gender, vascular condition etc). Moreover, absolute feature are also used to indicate the presence of other cardiac diseases as well [23,24]. Hence, the use of absolute features to indicate the sole presence of MI might be inappropriate. On the other hand, it has already been reported in different ECG based literatures [30–35] that compared to the absolute features, beat-to-beat feature variability can be a potential indicator of MI. Basically, ECG is indicative of the electrical activity of the heart and its morphology is a highly related to the pumping action of the heart as well. Any damages in the cardiac muscles due to MI induced changes causes imbalance in the pumping action of the heart. The resultant imbalances in the pumping action will induce irregular changes in the ECG beats as well as in the ejected blood volume rate also. The ejected blood volume change is measured at different body extremities in terms of the PPG signal. Therefore, imbalances in the cardiac cycle, due to MI not only induce beat-to-beat changes in the ECG morphology but also impose changes in the beta-to-beat morphological variation of the PPG signal. In this research, the interbeat morphological variation of the features extracted from PPG signal and its derivatives presents prominent discrimination ability with satisfactory detection accuracy, compared to other PPG based state-of-the-art methods. This in turn justifies the proposed hypothesis with a positive assertion that the proposed beat-to-beat PPG feature variability can be used as a potential indicator of MI and the superior outcome is reflected in the performance comparison Table 9. Moreover, Table 9 ravels the fact that, detection accuracy of the proposed method presents compatible performance with other ECG based methods also.

In order to avoid biases of the machine learning approaches towards specific features and to construct an impartial feature subset, six different feature selection algorithms are adopted in the present research. Then based on the selection score, only those

Table 9
Performance comparison of the proposed method with relevant studies.

Signal	Ref.(Year)	Features used		Classifier	Performance		
		Number	Feature type		Acc	Se	Sp
ECG (12 lead)	[8] (2006)	12/lead 144 (Total)	Time plane features	Rough set	–	95.8	100
ECG (12 lead)	[9] (2012)	6 /lead 72 (total)	ST segment polynomial features	Multiple instance learning	–	92.5	89.1
ECG (3 lead)	[10] (2014)	2 /lead 6 (total)	Cross Wavelet Trans-form,cross spectrum coherence	Heuristically determined mathematical formula	97.6	97.3	98.8
ECG (12 lead)	[11] (2015)	5 /lead 60 (total)	Energies and eigen values of multi-scale Discrete Wavelet Transform	KNN	96	93	99
ECG (3 lead)	[13] (2018)	2/lead 6 (total)	Phase of DFT	Nonlinear SVM LR and Threshold based	94.1	94.6	92
ECG (4 lead)	[14] (2012)	4	Hidden Markov model and log likelihood	Gaussian mixture model and SVM	82.50	85.71	79.82
ECG (4 lead)	[15] (2018)	ECG beats processed using Fuzzy Information Granulation		2D convolution neural net and Lead Asymmetric Pooling	96.00	95.40	97.37
PPG (from ear lobe)	[26] (2001)	–	Spectral analysis and HRV	Classification is not performed, only visually displays the difference in spectral characteristic for normal and MI (only 10 records are taken)			
PPG (from finger-tip)	[27] (2012)	1	Time domain signal processing	Classification is not performed, only shows that the 'relative crest time' is below 0.2 and has a tight distribution for healthy cases (24 MI and 60 normal records are taken data)			
PPG (from finger-tip)	[28] (2016)		SDPPG time-plane features	Logistic regression and statistical analysis	–	90.6 81.2 (32 MI and 32 H data)	87.5 84.4
PPG (from finger-tip)	[29] (2017)		SDPPG time-plane features	Binary logistic regression and statistical analysis	90.6	93.9 (32 MI and 32 H data)	87.5
PPG (from finger-tip)	Proposed	5	Variance values of the extracted Time-plane parameters from PPG, FDPPG and SDPPG	SVM (Linear)	95.40	92.70 (52 MI and 52 H)	98.10

features are considered, that are selected by most of the selection algorithms. However, the feature selection algorithms presents only 40 % reduced version of the initial feature dimension and most of the optimized feature presents the same total selection score. Hence the feature selection results are combined with statistical analysis to determine the optimum feature subset of eight features.

The proposed model uses a single segment of PPG signal, collected from each record to initiate faster analysis at every step of fiducial point detection, feature extraction and finally came up with reasonable accuracy. The method can also be implemented with multiple signal records, taken from each of the subject as well to ensure higher feature variability and to facilitate better classification accuracy. However, in that case, the resultant memory requirements and the computational burden will be much higher.

In this research, MI induced changes are mostly reflected by one PPG signal feature (F13), one FDPPG feature (F34) and three SDPPG features (F44, F48 and F59) respectively. Which in turn establishes the discriminating ability of the SDPPG signal features for MI detection, as mentioned in previous literatures as well [28,29]. Due to the inadequacy of data, only five fold cross validation technique is adopted to evaluate the performance of the classifier. Although, higher cross validation methods might increases the computational burden.

Considering high sensitivity of the PPG signal to motion artifact, care was taken during PPG signal acquisition. Since the present research is based on the morphological analysis of the PPG signal and its derivatives, it is acknowledged that the presence of motion artifacts and other cardiac arrhythmias [e. g. Premature Ventricular Contraction (PVC) or fibrillation] might induce abrupt morphological variations in the PPG signal waveform. The consequent, uneven changes will also be reflected in the FDPPG and SDPPG waveform and in the process, degrade the performance of

the algorithm. Hence, efficiency of the proposed method is kept limited for the patient with no physical activities, as mentioned in the data acquisition section also. The denoising technique adopted in the proposed algorithm is not adequate to deal with such situations. In future, during practical implementation of the proposed method, some advanced filtering techniques is essential for effective removal of the motion artifact corruption and abnormal beats from the PPG signal segments prior to analysis.

In this research, beat-to-beat PPG feature variability, created due to MI is studied and found to be more indicative of MI than any other PPG features. The discrimination ability of the features is shown by different standard classification techniques [Decision Tree (DT), Quadratic discriminate (QD), Logistic regression (LR), Linear SVM (LS), Nonlinear SVM (NLS), k-Nearest Neighbor (kNN)]. It can also be seen from Table 8 that apart from the proposed SVM (linear) classifier, two other classifier also presents notable performance. That means, other than the proposed SVM (linear) method, any one of the two classifiers [Nonlinear SVM (NLS) and k-Nearest Neighbor (kNN)] can also be used to indicate the presence of MI. Finally, considering computational overhead, SVM (linear) is reported as the classification technique for the present model. Moreover, it is important to note that, no standard database is available for MI affected PPG signal and we were granted permission to collect PPG signal data from MI affected patients for six month duration only. The algorithm is then evaluated only on the hospital collected 52 PPG signal records, which is quite robust enough to validate the utility of the proposed technique but inadequate for validation or optimization of the machine learning methods. Hence, the proposed research doesn't aim towards the validation or optimization of the machine learning methods. However, optimized performances of the proposed SVM (linear) classifier including two other high performing classification techniques [SVM (non-linear) and k-Nearest Neighbor] are presented in Fig. 9 by varying the training data percentage. The promising result obtained from this research validates the fact that PPG signal can be successfully employed for MI detection, but performance optimization with other advanced classification techniques and tuning of the classification methods is still an open problem for the researcher and will be considered in our future works as well.

The basic purpose of the present research work is to develop a complete model using PPG signal features only for MI detection, which can be implemented in automated healthcare systems. However the developed model has not been implemented in hardware level and the present scope of the work is concerned only with implementation and test on the software level in MATLAB platform using a personal computer. Instead of simple mathematical techniques, the fiducial point detection and detail feature extraction methodology, used in the proposed method imposes certain computational burden. Therefore, for each segment of the PPG signal, the optimal signal length is kept limited to 30 beats only, in order to carry out faster analysis with acceptable accuracy at the cost of less computational burden. However, in case of real life practical implementation of the method, consistent evaluation of the PPG signal record, after a certain interval, taken from the same subject will be required to evaluate and specify the onset of MI.

5. Conclusion

In this research, a PPG based novel approach is proposed for preliminary level, automated screening of MI. To the best of our knowledge, a complete MI detection model based on the PPG signal has not been reported yet. Compared to multi-lead ECG system, a single lead PPG signal is used in this research to ensure easy, unobtrusive, less expensive and operator independent signal acquisition at the comfort of home. The proposed parameter extraction algorithm relies on simple derivative-based method and uses less computationally intensive mathematical techniques such as amplitude threshold, slope reversal and an empirical formula based approach. This, in turn, ensures easy implementation of the algorithm in PPG based health monitoring systems. Instead of using absolute feature, beat-to-beat variance values of different morphological parameters, extracted from the PPG signal and its derivatives are reported in the proposed work for the first time to quantify MI related changes. The chosen beat-to-beat PPG feature variability, created due to MI is found to be more indicative of MI than any other PPG features. The algorithm uses only five discriminating features out of seventy two as an optimum feature set and presents significant accuracy of 95.40 %. The competency of the model is validated and justified using normal and MI records under medical supervision.

It is expected that the promising results obtained from the PPG based method opens a new horizon for early detection of MI using less expensive, simple automated techniques. However, it is worth mentioning that, the proposed PPG based MI detection method can only provide preliminary diagnosis for early screening of MI, which can be used as a primary reference before taking any major diagnostic decision. Hence, it can be applied as an affordable alternative for the massive rural population around the world to reduce the risk of mortality due to MI. In addition, the proposed research justifies the fact that, PPG signal can be efficiently used as an important tool for cardiac disease analysis. Although, further detail analysis of the PPG signal feature from multiple aspects is required to extract other cardiac abnormality information caused by MI. In future, revision of the proposed work with increased number of subjects along with advanced signal processing and classification tools are required to improvise the detection ability.

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgements

The authors would like to express their deepest gratitude to Dr. S. Guha, MD, DM (cardiology) (Professor and HOD), other medical practitioners and the entire support staffs of the department of cardiology, Medical College and Hospital, Kolkata for their valuable suggestions and assistance in the development of PPG database and validation of the proposed algorithm.

References

[1] World Health Organization, Cardiovascular Diseases (CVDs), 2018 (Accessed on 14 April 2019) http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)/.
[2] K. Thygesen, J.S. Alpert, A.S. Jaffe, M.L. Simoons, B.R. Chaitman, H.D. White, Third universal definition of myocardial infarction, Circulation 126 (6) (2012) 2020–2035, http://dx.doi.org/10.1161/CIR.0b013e31826e1058.
[3] D. Sopic, A. Aminifar, A. Aminifar, D. Atienza, Real-time event-driven classification technique for early detection and prevention of myocardial infarction on wearable systems, IEEE Trans. Biomed. Circuits Syst. 12 (5) (2018) 982–992, http://dx.doi.org/10.1109/TBCAS.2018.2848477.
[4] C.P. Cannon, et al., Relationship of symptom-onset-to-balloon time and door-to-balloon time with mortality in patients undergoing angioplasty for acute myocardial infarction, J. Am. Med. Assoc. 283 (2000) 2941–2947, http://dx.doi.org/10.1001/jama.283.22.2941.
[5] B.S. Chandra, C.S. Sastry, S. Jana, Reliable resource-constrained telecardiology via compressive detection of anomalous ECG signals, Comput. Biol. Med. 66 (2015) 144–153, http://dx.doi.org/10.1016/j.compbiomed.2015.09.005.

[6] A.K. Dohare, V. Kumar, R. Kumar, Detection of myocardial infarction in 12 lead ECG using support vector machine, Appl. Soft Comput. J. 64 (2018) 138–147, http://dx.doi.org/10.1016/j.asoc.2017.12.001.

[7] K.A. Ammer, J.A. Kors, B.P. Yawan, R.J. Rodeheffer, Defining unrecognized myocardial infarction: a call for standardized electrocardiographic diagnostic criteria, Am. Heart J. 148 (2) (2004) 277–284, http://dx.doi.org/10.1016/j.ahj.2004.03.019.

[8] S. Mitra, M. Mitra, B.B. Chaudhuri, A rough-set-based inference engine for ECG classification, IEEE Trans. Instrum. Meas. 55 (6) (2006) 2198–2206, http://dx.doi.org/10.1109/TIM.2006.884279.

[9] L. Sun, Y. Lu, K. Yang, S. Li, ECG analysis using multiple instance learning for myocardial infarction detection, IEEE Trans. Biomed. Eng. 59 (12) (2012) 3348–3356, http://dx.doi.org/10.1109/TBME.2012.2213597.

[10] S. Banerjee, M. Mitra, Application of cross wavelet transform for ECG pattern analysis and classification, IEEE Trans. Instrum. Meas. 63 (2) (2014) 326–333, http://dx.doi.org/10.1109/TIM.2013.2279001.

[11] L.N. Sharma, R.K. Tripathy, S. Dandapat, Multiscale energy and eigenspace approach to detection and localization of myocardial infarction, IEEE Trans. Biomed. Eng. 62 (7) (2015) 1827–1837, http://dx.doi.org/10.1109/TBME.2015.2405134.

[12] S. Padhy, S. Dandapat, Third-order tensor based analysis of multilead ECG for classification of myocardial infarction, Biomed. Signal Process. Control 31 (2017) 71–78, http://dx.doi.org/10.1016/j.bspc.2016.07.007.

[13] D. Sadhukhan, S. Pal, M. Mitra, Automated identification of myocardial infarction using harmonic phase distribution pattern of ECG data, IEEE Trans. Instrum. Meas. 67 (10) (2018) 2303–2313, http://dx.doi.org/10.1109/TIM.2018.2816458.

[14] P.C. Chang, J.J. Lin, J.C. Hsieh, J. Weng, Myocardial infarction classification with multi-lead ECG using hidden Markov models and Gaussian mixture models, Appl. Soft Comput. 12 (10) (2012) 3165–3175, http://dx.doi.org/10.1016/j.asoc.2012.06.004.

[15] W. Liu, et al., Real-time multilead convolutional neural network for myocardial infarction detection, IEEE J. Biomed. Health Inform. 22 (5) (2018) 1434–1444, http://dx.doi.org/10.1109/JBHI.2017.2771768.

[16] K. Channer, F. Morris, Clinical review: ABC of clinical electrocardiography Myocardial ischaemia, Br. Med. J. 324 (2002) 1023–1026, http://dx.doi.org/10.1136/bmj.324.7344.1023.

[17] C. Van Mieghem, M. Sabbe, D. Knockaert, The clinical value of the ECG in noncardiac conditions, Chest 125 (4) (2004) 1561–1576, http://dx.doi.org/10.1378/chest.125.4.1561.

[18] J. Allen, Photoplethysmography and its application in clinical physiological measurement, Physiol. Meas. 28 (3) (2007) 1–39, http://dx.doi.org/10.1088/0967-3334/28/3/R01.

[19] M. Elgendi, On the analysis of fingertip photoplethysmogram signals, Curr. Cardiol. Rev. 8 (1) (2012) 14–25, http://dx.doi.org/10.2174/157340312801215782.

[20] M.T. Islam, I. Zabir, S.T. Ahamed, M.T. Yasar, C. Shahnaz, S.A. Fattah, A time-frequency domain approach of heart rate estimation from photoplethysmographic (PPG) signal, Biomed. Signal Process. Control 36 (2017) 146–154, http://dx.doi.org/10.1016/j.bspc.2017.03.020.

[21] X. He, R.A. Goubran, X.P. Liu, Secondary peak detection of PPG signal for continuous cuffless arterial blood pressure measurement, IEEE Trans. Instrum. Meas. 63 (6) (2014) 1431–1439, http://dx.doi.org/10.1109/TIM.2014.2299524.

[22] C. Butter, et al., Cardiac resynchronization therapy optimization by finger plethysmography, Heart Rhythm 1 (5) (2004) 568–575, http://dx.doi.org/10.1016/j.hrthm.2004.07.002.

[23] A. Sološenko, A. Petrėnas, V. Marozas, L. Sörnmo, Modeling of the photoplethysmogram during atrial fibrillation, Comput. Biol. Med. 81 (2017) 130–138, http://dx.doi.org/10.1016/j.compbiomed.2016.12.016.

[24] A. Sološenko, A. Petrėnas, V. Marozas, Photoplethysmography-based method for automatic detection of premature ventricular contractions, IEEE Trans. Biomed. Circuits Syst. 9 (5) (2015) 662–669, http://dx.doi.org/10.1109/TBCAS.2015.2477437.

[25] F. Miao, X. Wang, L. Yin, Y. Li, A wearable sensor for arterial stiffness monitoring based on machine learning algorithms, IEEE Sens. J. 19 (4) (2019) 1426–1434, http://dx.doi.org/10.1109/JSEN.2018.2880434.

[26] V.S. Murthy, S. Ramamoorthy, N. Srinivasan, S. Rajagopal, M.M. Rao, Analysis of photoplethysmographic signals of cardiovascular patients, Proc. of the 23rd IEEE EMBS 3 (2001) 2204–2207, http://dx.doi.org/10.1109/IEMBS.2001.1017209.

[27] G. Angius, et al., Myocardial infarction and antiphospholipid syndrome, a first study on finger PPG waveforms effects, Comput. Cardiol. (2010) 39 (2012) 517–520.

[28] N. Mahri, K.B. Gan, M.A. Mohd Ali, M.H. Jaafar, R. Meswari, Analysis of myocardial infarction signals using optical technique, J. Med. Eng. Technol. 40 (4) (2016) 155–161, http://dx.doi.org/10.3109/03091902.2016.1153740.

[29] N. Mahri, K.B. Gan, R. Meswari, M.H. Jaafar, M.A.Mohd. Ali, Utilization of second derivative photoplethysmographic features for myocardial infarction classification, J. Med. Eng. Technol. 41 (4) (2017) 298–308, http://dx.doi.org/10.1080/03091902.2017.1299229.

[30] S.A. Ben-Haim, et al., Beat-to-beat electrocardiographic morphology variation in healed myocardial infarction, Am. J. Cardiol. 68 (8) (1991) 725–728, http://dx.doi.org/10.1016/0002-9149(91)90643-Y.

[31] L. Sörnmo, B. Wohlfart, J. Berg, O. Pahlm, Beat-to-beat QRS variability in the 12-lead ECG and the detection of coronary artery disease, J. Electrocardiol. 31 (1998) 336–344, http://dx.doi.org/10.1016/S0022-0736(98)90019-X.

[32] T. Murabayashi, et al., Beat-to-beat QT interval variability associated with acute myocardial ischemia, J. Electrocardiol. 35 (1) (2002) 19–25, http://dx.doi.org/10.1054/jelc.2002.30250.

[33] M.A. Hasan, D. Abbott, M. Baumert, Beat-to-beat QT interval variability and T-wave amplitude in patients with myocardial infarction, Physiol. Meas. 34 (9) (2013) 1075–1083, http://dx.doi.org/10.1088/0967-3334/34/9/1075.

[34] P.D. Arini, E.R. Valverde, Beat-to-beat electrocardiographic analysis of ventricular repolarization variability in patients after myocardial infarction, J. Electrocardiol. 49 (2) (2016) 206–213, http://dx.doi.org/10.1016/j.jelectrocard.2015.12.003.

[35] M.P. Bonomini, P.D. Arini, Modulation of spatial variance of ventricular repolarization after myocardial infarction, Biomed. Signal Proc. Control. 34 (2017) 214–219, http://dx.doi.org/10.1016/j.bspc.2017.01.014.

[36] J.A. Sukor, S.J. Redmond, N.H. Lovell, Signal quality measures for pulse oximetry through waveform morphology analysis, Physiol. Meas. 32 (3) (2011) 369–384, http://dx.doi.org/10.1088/0967-3334/32/3/008.

[37] A. Reisner, P.A. Shaltis, D. McCombie, H.H. Asada, Utility of the photoplethysmogram in circulatory monitoring, Anesthesiology 108 (5) (2008) 950–958, http://dx.doi.org/10.1097/ALN.0b013e31816c89e1.

[38] A. Chakraborty, D. Sadhukhan, M. Mitra, An automated algorithm to extract time plane features from the PPG signal and its derivatives for personal health monitoring application, IETE J. Res. (2019), http://dx.doi.org/10.1080/03772063.2019.1604178.

[39] A. Chakraborty, D. Sadhukhan, M. Mitra, Accurate detection of Dicrotic notch from PPG signal for telemonitoring applications, Int. J. Biomed. Eng. Technol. (2018), http://dx.doi.org/10.1504/IJBET.2018.10022653.

[40] M. Elgendi, Detection of c, d, and e waves in the acceleration photoplethysmogram, Comput. Methods Programs Biomed. 117 (2014) 125–136, http://dx.doi.org/10.1016/j.cmpb.2014.08.001.

[41] D.G. Jang, U. Farooq, S.H. Park, C.W. Goh, M. Hahn, A knowledge-based approach to arterial stiffness estimation using the digital volume pulse, IEEE Trans. Biomed. Circuits Syst. 6 (4) (2012) 366–374, http://dx.doi.org/10.1109/TBCAS.2011.2177835.

[42] Wan-Hua Lin, Hui Wang, Oluwarotimi Williams Samuel, Gengxing Liu, Zhen Huang, Guanglin Li, New Photoplethysmogram Indicators for Improving Cuffless and Continuous Blood Pressure Estimation Accuracy, Physiol. Meas. 39 (2) (2018), http://dx.doi.org/10.1088/1361-6579/aaa454.

[43] A.R. Kavsaoğlu, K. Polat, M. Recep Bozkurt, A novel feature ranking algorithm for biometric recognition with PPG signals, Comput. Biol. Med. 49 (2014) 1–14, http://dx.doi.org/10.1016/j.compbiomed.2014.03.005.

[44] A.R. Kavsaoɔlu, K. Polat, M. Hariharan, Non-invasive prediction of hemoglobin level using machine learning techniques with the PPG signal's characteristics features, Appl. Soft Comput. J. 37 (2015) 983–991, http://dx.doi.org/10.1016/j.asoc.2015.04.008.

[45] V.R. Nafisi, M. Shahabi, Intradialytic hypotension related episodes identification based on the most effective features of photoplethysmography signal, Comput. Methods Programs Biomed. 157 (2018) 1–9, http://dx.doi.org/10.1016/j.cmpb.2018.01.012.

[46] J. Allen, C.P. Oates, T.A. Lees, A. Murray, Photoplethysmography detection of lower limb peripheral arterial occlusive disease: a comparison of pulse timing, amplitude and shape characteristics, Physiol. Meas. 26 (5) (2005) 811–821, http://dx.doi.org/10.1088/0967-3334/26/5/018.

[47] N. Reljin, et al., Using support vector machines on photoplethysmographic signals to discriminate between hypovolemia and euvolemia, PLoS One 13 (2018), doi:10.1371/journal.pone.0195087. https://doi.org/10.1371/journal.pone.0195087.

[48] N. Nirala, R. Periyasamy, B.K. Singh, A. Kumar, Detection of type-2 diabetes using characteristics of toe photoplethysmogram by applying support vector machine, Biocybern. Biomed. Eng. 39 (2019) 38–51, http://dx.doi.org/10.1016/j.bbe.2018.09.007.