



# Effortless detection of premature ventricular contraction using computerized analysis of photoplethysmography signal

SANDIPAN DHAR<sup>1,\*</sup>, ABHISHEK CHAKRABORTY<sup>1</sup>, DEBOLEENA SADHUKHAN<sup>2</sup>, SAURABH PAL<sup>1</sup> and MADHUCHHANDA MITRA<sup>1</sup>

<sup>1</sup>Department of Applied Physics, University of Calcutta, Kolkata, West Bengal, India

<sup>2</sup>Department of Applied Mechanics, Indian Institute of Technology Madras, Chennai, India  
e-mail: sandipan086@gmail.com; acaphy\_rs@caluniv.ac.in; deboleena.rainbow@gmail.com; spal76@gmail.com; mmaphy@caluniv.ac.in

MS received 28 September 2021; revised 25 November 2021; accepted 26 November 2021

**Abstract.** Abnormality in the heart rhythm owing to premature ventricular contraction (PVC) often causes fatal cardiac consequences. Normally, early detection of PVC uses long-term Electrocardiogram (ECG) monitoring (Holter) techniques. In recent days, Photoplethysmography (PPG) based approaches are also being adopted for PVC detection. Primarily, the lower cost and effortless acquisition of PPG makes it suitable for long-term continuous monitoring applications. However, the PPG-based PVC detection method has not been standardized yet and it is an open research problem to date. In this research, a less complicated and automated PVC detection method is proposed that uses PPG signal-based analysis only. Instead of any computationally intense time-plane features, the overall morphology of each of the PPG beats is quantified using two simple statistical parameters. Variations of these two parameters are then employed as features to identify the abrupt morphological changes caused by PVC. The proposed features are also used to identify and eliminate noisy data segments and minimize the rate of false detections. Finally, a simple threshold-based criterion is used to identify the presence of PVC beats among the normal beats. After evaluation over the PPG signal records obtained from the MIMIC dataset, the proposed method exhibits sensitivity, specificity and accuracy of 99.23%, 99.68% and 99.05%, respectively. Compared to other ECG or PPG-based methods, the methodological simplicity and the overall noteworthy outcome associated with the proposed PPG-based PVC detection technique show immense potential for implementation in personalized health monitoring applications.

**Keywords.** Computerized disease detection; expert health system; photoplethysmography; premature ventricular contraction (PVC); statistical feature extraction.

## 1. Introduction

In recent decades, the dire influences of cardiovascular diseases (CVDs) have been considered as a leading cause of increased mortality in our society and even today it is being treated as a serious health concern [1]. Certain cardiac abnormalities are found to be associated with the rhythmic inconsistency of the cardiac cycles, commonly known as arrhythmia [2]. Out of many, premature ventricular contraction (PVC) is regarded as a specific kind of cardiac arrhythmia, found to be present in normal people as well, without having a medical history of any cardiac diseases [3]. The problem of PVC originates due to the presence of ectopic centers in the ventricles, which in turn causes premature heartbeats as a result of premature contraction of the ventricle [4]. Recent research shows that, apart from its

apparent benign nature, the effects of PVC often lead to serious life threatening cardiac disorders such as heart failure [5], atrial fibrillation [6] and ventricular fibrillation [7], etc.

Hence, early detection of PVC plays an important role to reduce cardiac disease-related mortality rates.

The conventional approaches for arrhythmia detection are typically based on long-term monitoring (Holter) and analysis of the ECG signal [8]. However, for the sake of accuracy and to avoid manual overload, recent ECG based methods are employing different computerized approaches. Most often these automated methods use different types of feature, such as ECG morphological features [9], PCA based features [10], frequency-based features [11], Convolutional neural networks (CNNs) [12] and also different classification methodologies such as linear discriminants [9, 10] artificial neural networks [11], decision trees, support vector machines and random forest [12], etc. for

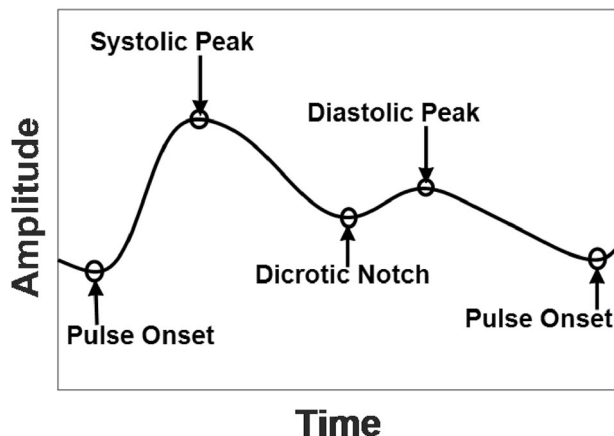
\*For correspondence

automated identification of arrhythmic beats. Apart from the complicity involved in these feature extraction and classification techniques, the use of the ECG signal itself also imposes certain limitations as mentioned below.

(1) ECG acquisition is relatively costly and complicated; (2) ECG requires multiple electrodes to be placed at different parts of the body via conductive gel, which imposes a direct effect on patient comfort and mobility; (3) proper acquisition of ECG signal requires skilled operators and (4) ECG signal quality is found to be compromised by improper placement of ECG electrodes, the presence of multiple noises and movement-related artefacts [13, 14]. These limitations sometimes impose a major restriction on the use of computationally intense ECG based techniques for long term monitoring applications.

Consequently, as a competent alternative, different non-invasive and unique properties of Photoplethysmogram (PPG) signal are now being extensively investigated, because of its suitable, inexpensive acquisition technology and operative independent operational facilities. Typically, a non-invasive and electro-optic technique is exploited in the Photoplethysmography method in order to extract the information related to the variation in the blood volume at different human body extremities (i.e. earlobe, fingertip and toe). Usually, the low-frequency morphology (approximately 1 Hz) of the PPG wave includes a pulsating AC-part (corresponds to the average changes in blood volume with each cardiac cycle) overlapped on a quasi-DC part (indicating the influence of respiration and vasomotor activity, etc.). The pulsatile part of the PPG single is characterized by some clinically significant fiducial points, as illustrated in a single PPG beat waveform of figure 1 [15, 16].

Contemporary researches have revealed the fact that, analysis of the PPG signal [17] features essentially leads to the extraction of a variety of cardiovascular factors for instance heart rate (HR) [18], cardiac outputs (CO) [19] and



**Figure 1.** A representative Photoplethysmography (PPG) signal beat along with its detected characteristic points.

blood pressure (BP) [20], etc. Moreover, PPG signal features also present suitable promise to identify different cardiac disorders such as myocardial infarction (MI) [21, 22], premature ventricular contraction (PVC) [23–26].

To date, only a handful of literature is available that exploits the clinical aspects of the PPG signal to indicate the presence of premature ventricular contraction (PVC) [23–26]. The method proposed in [23] uses the correlation characteristics of a 2-lead ECG and PPG to analyze the arrhythmic heartbeats. Then, peak-to-peak interval (PPI) of the pulse and the pulse height features are used for accurate quantification of the arrhythmic pulse over a very small dataset, taken from three subjects only. The inherent potential of the PPG signal in terms of amplitude, rhythm, and pulse analysis for PVC detection is explained in [24] without any quantification. In [25], PVC beats are initially marked by the ECG signal and the turbulence onset, slope and shape-related parameters are computed from the PPG beats. The extracted PPG features are then used in a linear classifier to carry out PVC classification over a PPG dataset obtained from 27 patients with 4131 PVC beats. Six temporal (peak-to-peak intervals) and power-derived features (the power ratios) are used in [26] to classify PVC beats using artificial neural network. The methodology uses the R-R interval of the synchronously recorded ECG signal for manual annotation of the PVC beats in the corresponding PPG signal. The entire algorithm is evaluated over the twenty-six PPG dataset with 3280 PVC beats and finally shows the classification of two distinct types of PVC (PVC with large or small pulse amplitude) with reasonable accuracy. However, most of the above-mentioned research either uses (1) the detection of multiple fiducial points [23, 25], (2) the extraction of multiple features [25, 26], or (3) uses a complicated classification model [26].

In this research, the detection of a number of fiducial points and the extraction of computational intense time plane features are avoided and the overall PPG beat morphology is represented in terms of two easy-to-extract statistical parameters via accurate detection of PPG onset points only. The calculated beat-to-beat difference of these two parameters present discriminating separation for the PVC beats as compared to the normal beats and hence are used as features to identify the occurrence of PVC by means of a simple threshold-based classification technique. Compared to the above-mentioned researches [23–26], both the adopted feature extraction and classification techniques present much reduced computational complexity with a promise to be implemented in personal health monitoring devices. Furthermore, the same extracted features, combined with two less complicated conditions are also used to identify and discard the PPG beats degraded with motion artefacts. The adopted artefact removal methodology is very simple and requires no advanced signal analysis techniques. The efficiency of the developed algorithm method is validated over the standard dataset obtained from the MIMIC database.

The remainder of the article is structured as follows. The entire methodological details regarding the developed algorithm and also particulars about the selected database are explained in Section 2. The experimental results obtained and evaluation of the proposed algorithm with other current state-of-the-art works is described in Section 3. Detailed discussion related to the robustness and competence of the algorithm is carried out in section 4. Finally, the synopsis of the work and the prospect of the proposed method are concluded in Section 5.

## 2. Methodology

The entire methodology followed in the developed PVC detection algorithm is presented in the block diagram of figure 3. The whole endeavor has been classified into four key sections: (i) denoising of the acquired PPG signal and detection of PPG onset points, (ii) extraction of PPG signal features; (iii) elimination of noisy beats using the extracted features; and at the end, (iv) PVC beat classification based on the obtained features.

### 2.1 Database used

The detection of PVC using PPG signal features is quite a new area of research and therefore, no standard PPG based PVC dataset is available to date. In the present research, the presence of PVC is identified using twenty five PPG signal records of 1 to 2-hour duration, collected from the Physionet MIMIC Database [27] with a sampling frequency of 125 Hz. However, to minimize the errors in the detected fiducial points and the calculated features, the collected PPG datasets are again re-sampled to 250 Hz. While collecting all the PPG signal records, the corresponding ECG signal records are also taken into account to facilitate the identification of the PVC beats as shown in figure 2. In this case, primarily the PVC induced abnormalities are identified in the ECG signal

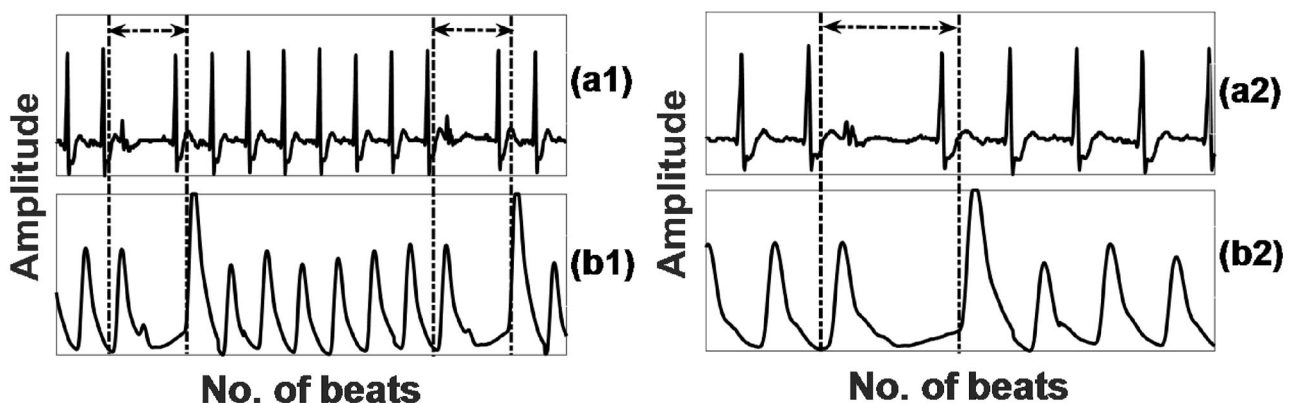
beats by investigating the nature of the ‘QRS-complex’, ‘R-R interval’ and the polarity of the R-peaks respectively. Each of the ECG based PVC beats are then utilized for manual annotation of the PVC beats in the corresponding PPG signal. In the present research, after examining the entire database, near about 2,423 PVC beats are taken into account for analysis. After detection of the PPG based PVC beats, the rest of the PPG beats with normal sinus rhythm are considered as the non-PVC (N).

It is to be mentioned that, during the investigation, PPG signal records obtained from other subjects of the database has not been taken into account due to the presence of massive signal corruptions, pathological variation or due to the absence of exploitable signal information. Detail description about the used database is given in table 1 below.

### 2.2 Preprocessing and normalization

During the course of PPG data acquisition, the signal quality is vastly compromised by various factors. Mostly, the amount of light captured by the photo-detector, stability of the index finger pressure on the acquisition sensor, inadequate blood perfusion at the peripheral tissues and the episodes of motion artefact noises often impose detrimental changes in the signal morphology and make it incompetent for further processing [28], etc.

Although, the techniques used for the detection of onset points of the PPG signal follows signal derivative-based methodologies. Consequently, the algorithm becomes sensitive to high-frequency noise and artefacts. Hence, to minimize the effect of high-frequency noise, the PPG signals used in the present research are de-noised using a finite impulse response (FIR) band-pass filter having a cut-off frequency ranging from 0.5 to 15 Hz. The use of a band pass filter not only facilitates effective removal of the high-frequency noise elements but also helps to maintain the clinical bandwidth of the PPG signal.



**Figure 2.** Two typical ECG signal based PVC segments (a1 and a2) and the effect of PVC on the corresponding PPG signals (b1 and b2) are presented with highlighted marked zones.

**Table 1.** Details about the used PPG dataset collected from MIMIC dataset.

Record no.	Collected PVC beats	Record no.	Collected PVC beats
039 m	0	404 m	0
041 m	0	408 m	9
055 m	1	439 m	12
211 m	0	442 m	1295
212 m	164	444 m	7
218 m	0	449 m	6
221 m	10	466 m	3
224 m	0	471 m	1
225 m	0	474 m	2
230 m	0	482 m	48
237 m	37	484 m	67
252 m	0	485 m	761
253 m	0	–	–
Total	212	Total	2211

After denoising, in order to facilitate further evaluation of all the PPG records under the same scale, amplitude normalization of the denoised PPG signal within the range of 0 to 1 is carried out using equation (1).

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

In the above equation,  $V_{signal}$  = denoised PPG data,  $V_{signal(min)}$  [or  $V_{signal(max)}$ ] = the minimum (or the maximum) reference value of the PPG data,  $V_{norm}$  = the normalized PPG data.

### 2.3 Beat segmentation

In order to quantify sudden irregular variation in the PPG signal due to PVC, careful inspection of the PPG waveform morphology via beat-by-beat segmentation is required. Physiologically, the pulse onset corresponds to the beginning of the blood ejection into the aorta as a result of ventricular contraction. Whereas, the systolic peak signifies the maximum changes in the blood volume and also indicates the end of blood ejection in the PPG waveform [16]. In this research, the onset-to-onset segmentation of the PPG beats is carried out using a robust, accurate and less computationally intensive technique, as described in [29]. At first, the mentioned algorithm facilitates the identification of all the characteristic points from the PPG first derivative (FDPPG).

Accurate positions of these detected FDPPG fiducial points are then employed to locate the precise location of the PPG onset points. As a whole, the detection of PPG onset uses no computationally rigorous transformation methodologies. Rather, the methodology follows signal derivatives; amplitude thresholding, slope reversal and empirical formula based less complicated mathematical

techniques in order to confirm its easy realization in the hardware level. A visual illustration of beat segmentation is presented in figure 3.

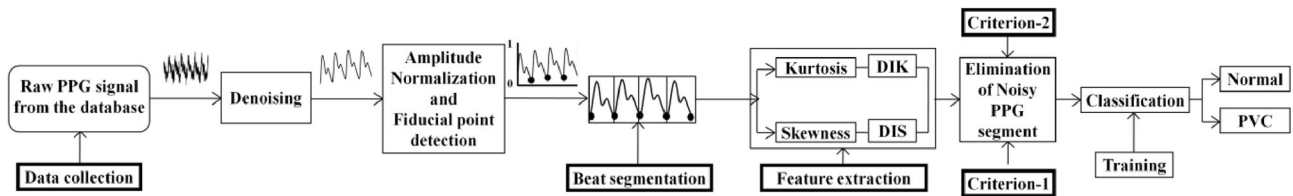
### 2.4 Feature extraction

Premature ventricular contraction (PVC) causes irregularity in the heartbeat, which can be realized in the form of palpitation in the chest. Usually, in the case of PVC, the ventricular contraction begins even when it is partially filled with blood by the atria and thus resulting in a skipped beat. Such discrepancies due to premature contraction of the ventricle present a distinct signature in the electrical response (ECG wave) of the heart as reported by many studies [9–12]. Moreover, this sudden imbalance in the cardiac cycle due to PVC should also affect the mechanical pumping activity of the heart and must impose improper variation on the ejected blood volume rate also. Now, with each cardiac cycle, these resulting changes in the blood volume can be quantified at multiple marginal sites of the body in the form of PPG signal. Therefore, rapid and irregular variations in the morphology of the PPG beat can be thought of as an efficient alternative for the non-invasive detection of PVC.

The primary objective of the proposed research is to present a simple and highly implementable PVC detection technique based on nominal analysis of the PPG signal. In general, PPG signal analysis can be categorized into two parts. (1) Analysis of the PPG signal via extraction of time domain features. (2) Frequency domain analysis of the PPG signal via some computationally intense transform based technique.

However, feature extraction in the time domain necessitates precise detection of a number of fiducial points from the PPG waveform. The required fiducial point detection algorithm in turn enhances the computational load of the overall analysis. Whereas, apart from its inherent complexity, the adopted transform based technique on the other hand creates confusion about the selection of the coefficients while representing the signal. Consequently, careful inspection is carried out in the present research while selecting the feature type and dimension.

Generally, every PPG beat contains a prominent primary peak and its time-plane morphology is found to be analogous to the probability distribution curve. Hence, variation of two statistical features, kurtosis and skewness has been chosen to describe the morphological changes in the PPG signal due to PVC. Inherently, kurtosis is used to quantify the peakedness of any distribution. Whereas, skewness is used to measure the symmetry (or asymmetry) of the data values around the sample mean [30]. The occurrence of PVC episodes is associated with abrupt changes in the PPG signal. Hence, immediately after the PVC episodes, these two parameters are found to capture the resulting effects on the PPG beat morphology very efficiently. Most



**Figure 3.** Block diagram representation of the entire methodology for PVC detection. Subsequent to pre-processing and normalization, the onset points of the PPG signal are detected in order to facilitate the calculation of two statistical parameters from each of the PPG beats. The calculated beat-to-beat difference values (DIK and DIS) of the extracted parameters are then considered as features. The extracted features are also used to eliminate noisy beats with two simple criteria. Finally, the features are used to classify the PVC beats via a less complicated threshold-based classification technique.

importantly, these parameters can be easily calculated based on the algebraic values of the PPG samples, involves the detection of PPG onset only and requires minimal infrastructure to implement. In the case of a PPG beat  $y$  ( $n$ ) of length  $N$ , the kurtosis and skewness values are calculated by the following equation

$$Kurtosis(K) = \frac{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^4}{\left\{ \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 \right\}^2} \quad (2)$$

$$Skewness(S) = \frac{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^3}{\left\{ \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2} \right\}^3} \quad (3)$$

Here,  $y_i = i^{\text{th}}$  sample value and  $\bar{y} =$  mean value of the corresponding PPG beat.

Beat-to-beat extracted values of kurtosis and skewness, as plotted in figures 4(b) and figure 4(c) indicate stable variation of these values for normal beats and abrupt variation for PVC beats. Although, depending on the pathophysiological conditions, sometimes beat-to-beat kurtosis and skewness values of few normal PPG pulses present compatible variation with the PVC beats. Hence, beat-to-beat differences of these two parameters, denoted by DIK (difference in kurtosis) and DIS (difference in skewness) are computed and taken as features in this research to identify the presence of PVC beats in the PPG signal. In Figures 4(b) And 4(c), differences of kurtosis (DIK) and skewness (DIS) values between the PVC beat ( $k_2$  or  $s_2$ ) and the previous ( $k_1$  or  $s_1$ ) or next normal beat ( $k_3$  or  $s_3$ ) are shown by an example. Evidently, in the presence of a PVC beat, the corresponding DIK and DIS values present a discriminating higher range compared to the normal beats. Finally, the extracted DIK and DIS feature values are evaluated using a simple threshold-based classification technique for the identification of PVC beats. The generalized threshold condition is determined after trying the algorithm over the entire datasets.

The calculated DIK and DIS feature values, obtained from the normal PPG and PVC induced PPG beats are also

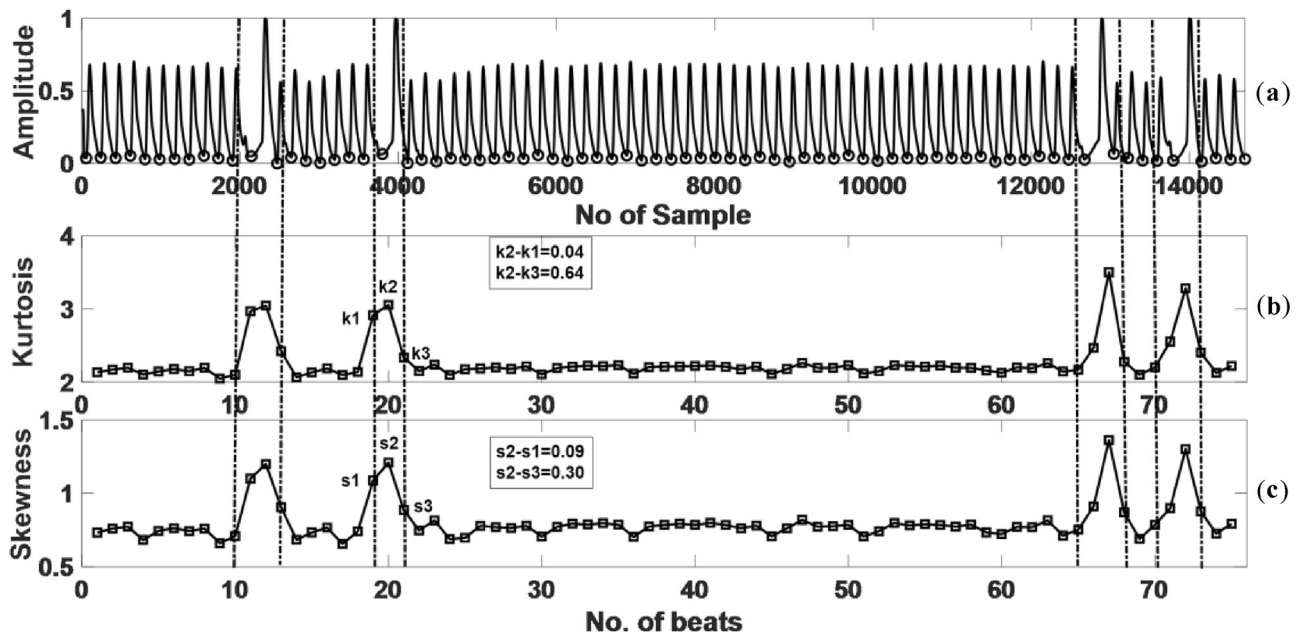
represented in the box plot of figure 5. Evidently, the median value of figure 5 presents discriminating difference for the PVC beats compared to the normal beat. A two-sample t-test is also carried out in order to justify the supremacy of the chosen statistical features and the outcome is indicated in figure 5 with minimized p-values ( $<0.0001$ ). The obtained low p-values for both of the features also indicate the potential of these extracted features to classify the existence of PVC beats from the normal.

### 2.5 Elimination of the artefact corrupted PPG beats

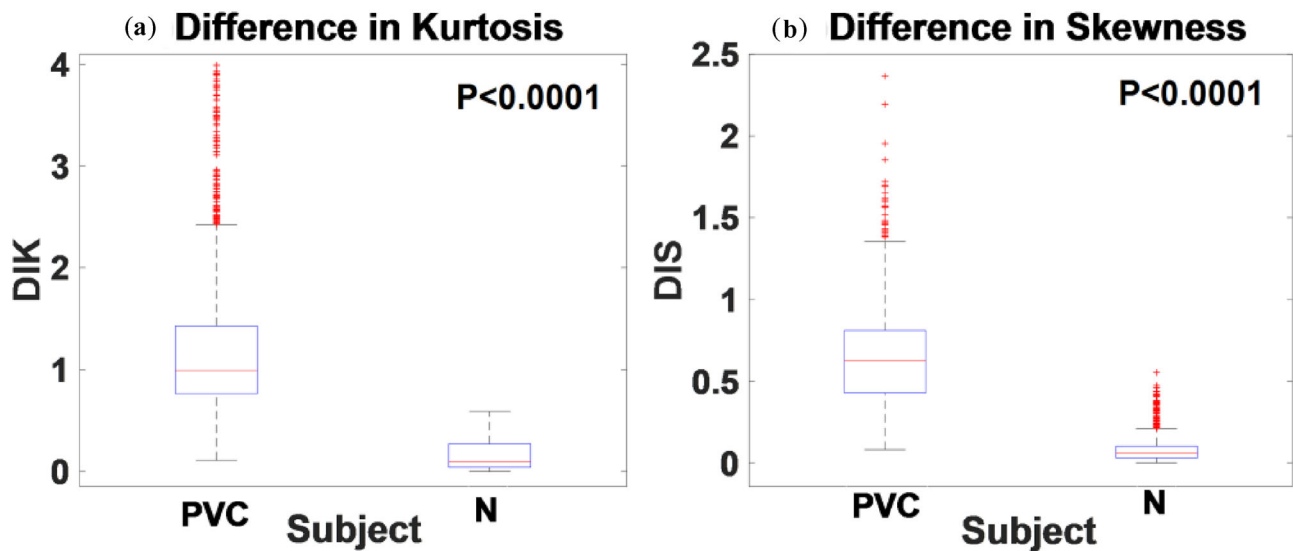
As stated earlier, PPG acquisition methodologies are highly affected by motion artefacts (MA). Unusual movement of the subjects often introduces sudden and intermittent artefacts in the acquired PPG signal with random morphological variations as indicated within the marked zone of figure 6(a). Eventually, the altered wave morphology associated in the MA induced PPG signal waveform causes random variation in the extracted feature values, which might often lead to false detection of the PVC beats. As a whole, the effect of MA corruption distorts the embedded clinical information by diminishing the signal quality. In the proposed method, instead of using any advance motion artefact removal technique, the data segment corrupted with MA is removed by means of simple and convenient criteria as discussed below.

**Criterion 1:** After evaluation over the whole database, it can be seen that the calculated PPG onset-to-onset time intervals are either lower than 300 ms or higher than 1500 ms respectively. This corresponds to the upper and lower threshold values of an individual's heart rate as 200 beats/minute or 40 beats/minute and also facilitates rejection of artefact corrupted beats.

**Criterion 2:** The difference in kurtosis (DIK) and skewness (DIS) values between two consecutive beats must lie above some threshold values [above 1 for DIK and above 0.8 for DIS]. Both of the threshold values are chosen empirically and are indicated by dashed line in figures 6(b) and also in figure 6(c), respectively.



**Figure 4.** a A typical amplitude normalized PPG based PVC record is presented, where the occurrence of PVC beats are marked within the dashed lines. Figure b and c indicates the calculated values of Kurtosis and c Skewness parameter against the PPG beats. From the above figure, it is evident that these parameters present discriminating higher values for the PVC beats compared to the normal beats.

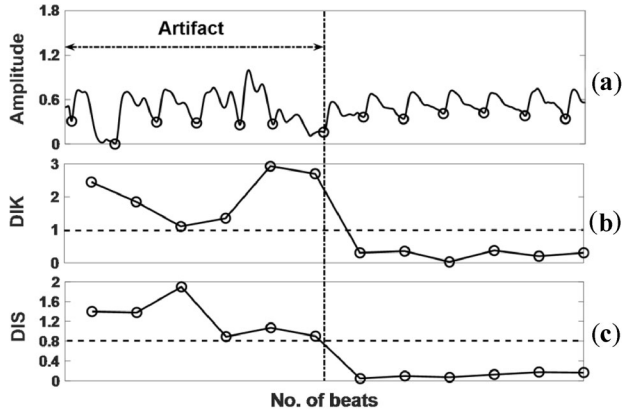


**Figure 5.** Statistical distribution of the extracted features, difference in a Kurtosis and b Skewness values are shown in the box-plot for normal (N) and PVC beats respectively. The median is indicated in the box by the middle line, the edges represent the 75th and 25th percentiles and the outliers are indicated by the whiskers. Moreover, the calculated p-values of these features are also indicated in the figures.

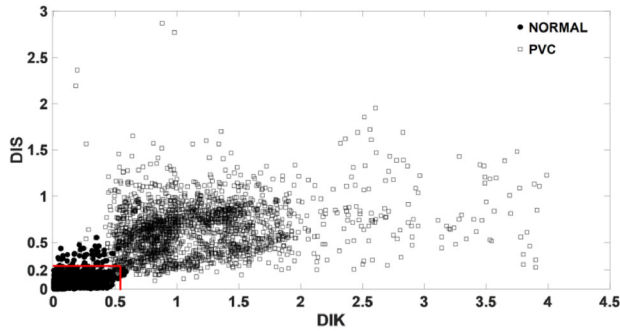
### 2.6 Classification

The median values of each of the proposed feature, obtained from the box-plot distribution characteristics of

figure 5 and also the obtained p-values of each of the extracted features present considerable discrimination between the normal and the PVC class. In the proposed algorithm, instead of any complicated classification



**Figure 6.** **a** Detected PPG beats in the presence of motion artefact (MA) corruption. **b** Corresponding variation in kurtosis (DIK) and **c**) skewness values obtained from the PPG beats. Allowable threshold ranges corresponding to each of the parameters are marked in the figure by dotted lines. Evidently, these two parameters present wide variation for the MA corrupted section (marked as the artefact zone), compared to other MA free zone.



**Figure 7.** The values of DIK and DIS for all the records are represented in the form of a two dimensional scatter plot. Every PVC beats are pointed out by the square symbol whereas black filled circle symbol is used to represent the normal beats. The chosen threshold-based boundary values for the classification of PVC beats are indicated by the red solid lines.

techniques, the adopted features are simply classified by means of a simple threshold-based decision rule method. The decision rule adopted to classify with feature values (DIK and DIS) into one of the two classes is formulated as follows

$$\text{Beat}_n = \begin{cases} \text{PVC, } K_n - K_{n-1} > (DIK_{th}) \text{ or } K_n - K_{n+1} > (DIK_{th}) \text{ and} \\ S_n - S_{n-1} > (DIS_{th}) \text{ or } S_n - S_{n+1} > (DIS_{th}) \\ \text{Normal, else} \end{cases} \quad (4)$$

The intra-class variation between the adopted features (DIK and DIS) is also represented in figure 7 in the form of a two-dimensional scatter plot. Clearly, figure 7 shows that in comparison with the normal class, the proposed attributes present much ample variation for the PVC.

### 3. Experimental results

In general, the efficiency of the developed algorithm is exhibited by means of four standard parameters, form instance, detection accuracy (Acc), sensitivity (Se), specificity (Sp) and also the receiver operating characteristic (ROC) curve respectively. Details of the performance evaluators are formulated as follows.

$$\text{Sensitivity (Se)} = \frac{TP}{TP + FN} \times 100\% \quad (5)$$

$$\text{Specificity (Sp)} = \frac{TN}{TN + FP} \times 100\% \quad (6)$$

$$\text{Accuracy (Acc)} = \frac{TP + TN}{TP + FN + TN + FP} \times 100\% \quad (7)$$

Where, TP is the sum of appropriately identified PVC beats, TN is the sum of appropriately identified healthy beats, FN signifies the sum of PVC beats mistakenly identified as healthy, FP represents the sum of healthy beats erroneously identified as PVC. Sensitivity (Se) signifies the capability of an analysis to categorize positive cases and Specificity (Sp) signifies the probability of being true positive when the test result is positive.

#### 3.1 Performance evaluation

The occurrence of PVC episodes is highly intermittent in nature. Hence, in the present research, around four million PPG beats from the adopted PPG dataset are investigated to find out the presence of PVC. After evaluation over the entire dataset, 2,423 PPG based PVC beats are taken into account. Moreover, implication of the developed model is also evaluated statistically by means of a 5-fold cross-validation method, considering less availability of the data. Moreover, to reduce the class imbalance, only 2,423 normal bets against 2,423 PVC beats are finally chosen for validation. It is worth mentioning that after detecting the PVC beats from a record under investigation, an equal number of normal beats are also taken from the same record. If any record shows the absence of the PVC beats, no normal beat is taken from that record.

Now, as per the procedure followed in cross-validation, the entire database is initially divided into five identical sections. Then, corresponds to every iteration, one section is kept for testing and the remaining four parts are employed for the validation purpose. Depending on the

arbitrary divisions of the dataset, the whole process is iterated five times. During cross-validation, the selected feature threshold values are kept unaltered in each fold, to make the training and validation process independent of each other. Following evaluation over the adopted dataset, the overall outcome of the developed algorithm is presented in Table 2 along with the pre-determined threshold values for DIK and DIS values. Evidently, Table 2 shows high average detection accuracy (Acc) of 99.05 %, sensitivity (Se) of 99.23% and specificity (Sp) of 99.68%, respectively.

### 3.2 Validation of the classification technique

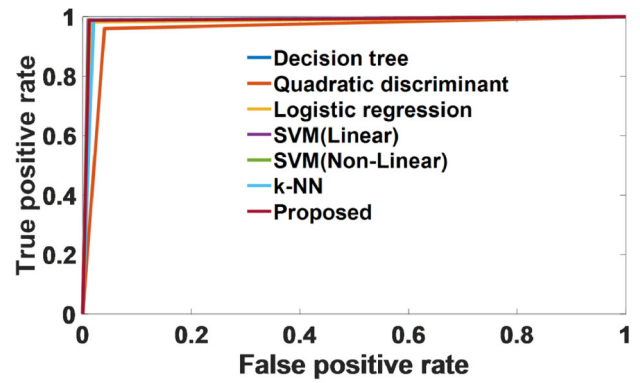
The proposed research uses beat-to-beat differences of two statistical parameters and a computationally simple threshold-based decision rule method for classification and finally come up with high accuracy. However, for further generalization, a standard comparison of the chosen threshold-based classification technique against other state-of-the-art classification methods is carried out to verify the discrimination efficiency of the proposed features. It is to be motioned that, for every adopted classification technique, training and testing process is carried out using five-fold cross-validation method. The average performance after final evaluation over fivefold is depicted in the final ROC plot of figure 8 and given in Table 3 as well. The significantly high accuracy, as indicated by each of the classifiers including the proposed one in Table 3 justifies the discrimination efficiency of the adopted features.

### 3.3 Performance comparison

It is to be mentioned that, to date only a few types of research have followed the analysis of the PPG signal to address the problem of PVC detection. Hence, in the present research, the performance of the method is compared only with a few PPG-based methods. In addition, the results obtained from the developed model are also compared with other automated ECG centric PVC detection methods to establish its proficiency. Although variation in the number and type of the used data sets, adopted feature dimension, differences in the validation process and most importantly the employed classification techniques of the suggested literature creates an imbalance for exact comparison. The detailed assessment of the proposed algorithm in terms of its performance with other contemporary literature is listed

**Table 2.** Performance of the selected threshold based classification technique.

Classification method	Chosen threshold values	True class	Acc (%)	Se (%)	Sp (%)	AUC (%)
Threshold based	DIK = 0.53 DIS = 0.23	PVC N	99.05	99.23	99.68	100



**Figure 8.** The Receiver operating characteristic (ROC) plots for the proposed classification methods as mentioned in Table 3.

**Table 3.** Performance comparison of the selected classifier with other standard classification techniques.

Classification method	Acc (%)	Se (%)	Sp (%)	AUC (%)
Decision tree	98.80	98.56	99.05	99.0
Quadratic discriminant	98.00	98.54	98.24	100
Logistic regression	99.01	98.92	99.19	100
SVM (Linear)	99.00	98.96	99.02	100
SVM (Non-linear)	98.98	98.92	99.05	100
k-NN (k=10)	98.76	98.83	98.69	99
Proposed	99.05	99.23	99.68	100

in Table 4. Clearly, the descriptions incorporated in Table 4 are indicative of the fact that compared to other related ECG or PPG based literature; the proposed algorithm shows high efficiency in terms of its accuracy and robustness using two easy-to-calculate PPG features only.

## 4. Discussion

Recent researches reveal that the underlying cardiovascular attributes of the PPG signal can be a facilitator of PVC detection [23–26]. Although, it is worth mentioning that PPG-based techniques to identify PVC beats is comparatively a new area of research. In this study, a simple, automated yet robust PVC detection method is proposed based on the variation analysis of two simple statistical parameters, extracted from the PPG signal. The major steps



**Table 4.** Performance comparison with other reported PVC detection techniques.

Signal used	Ref. (year)	Feature used	Classifier	Performance (%)		
				Acc	Se	Sp
ECG	[9] (2004)	Morphology indices	Linear discriminant	80.5	43.0	99.3
ECG	[10] (2013)	ECG beat and Z-score	Probabilistic neural network	99.71	97.98	99.10
ECG	[11] (2006)	Wavelet transform	Artificial neural networks (ANN)	–	76.5	85.8
ECG	[12] (2019)	Convolutional neural networks (CNNs)	Decision trees, Support vector machines and Random forest	99.77	99.30	–
PPG	[23] (2009)	Time plane	Threshold based	98.90	91.50	99.0
PPG	[25] (2013)	Time plane	Linear classifier	99.3	90.5	99.9
PPG	[26] (2015)	Temporal and power-derived features	Artificial Neural Networks (ANN) N	99.7	99.8	93.3
			PVC-1	99.8	93.2	99.9
			PVC-2	99.8	92.4	99.9
Proposed	–	Statistical	Threshold based	99.05	99.23	99.68

involved in the proposed approach are kept simple for targeted implementation in portable, handy, battery-operated health monitoring devices with limited hardware capabilities. The proposed algorithm comes up with certain distinct advantages as summarized below.

- (1) The suggested PVC detection methodology uses PPG signal acquired from a single channel to ensure its easy implementation for home monitoring applications.
- (2) No feature selection method is used and only two simple statistical parameters, calculated from accurate detection of PPG onset points are used. Then, unique beat-to-beat difference using these two parameters is considered as features and reported in this research for the first time.
- (3) Instead of any complicated transformation methods, accurate detection of PPG onset algorithm follows much less computationally-intensive techniques.
- (4) The proposed set of features presents such significant contrast between the two classes, a simple threshold based classification technique is found to be sufficient for the identification of the PVC beats with high accuracy.
- (5) Instead of using any advanced signal processing approaches, the motion artefact corrupted beats are eliminated in this method based on a simple criterion on the same extracted features. The proposed criteria impose no additional computational burden.

In the case of beat-by-beat normalization of the PPG beats, the amplitude of each of the normal or the PVC beats will be in the same range of 0 to 1. In such a case, the calculated feature values obtained from each of the PVC and the normal beats might not produce a discriminating outcome. Hence, in the present algorithm, instead of beat-by-beat normalization of the PPG beats, the entire PPG signal record under investigation is normalized within a range of 0 to 1. This will not only facilitate to provide a discriminating feature values within a range of 0 to 1 but

also helps to estimate a generalized threshold value for each of the adopted features.

In this research, rather than selecting the kurtosis values and skewness values from the PPG beat as features, beat-by-beat differences between these two parameters are taken as features to identify the presence of PVC.

Apart from simplicity in the adopted approach, this will impose a certain restriction to identify the onset of a particular PVC episode. Hence, it is important to mention that if any PVC beat with large or small amplitudes exist in a PPG record, the present algorithm is only capable of indicating the presence of that PVC beats but not the location of that particular beat.

In real life automated analysis, this is a fact that the occurrence of motion artefact might induce rapid morphological variations in the PPG waveform. Although, the derivative-based technique followed for the detection of PPG onset is usually responsive to motion artefact. The resultant uneven amplitude variation under motion artefact corruption eventually degrades the overall efficiency of the algorithm. Moreover, the identified noisy data segments, using the extracted features are discarded from the data strip in the proposed method. This might impose a problem for datasets with a small duration. In the future, instead of removing the noisy data segments, the use of some advanced motion artefact reduction algorithm should be used to facilitate the recovery of the PPG beats from the noisy segments.

The algorithm efficiency is tested over the PPG dataset obtained from the Physionet MIMIC databases only. Although, the presence of PVC beats in the MIMIC dataset is limited. In the future, further evaluation of the proposed algorithm is required over a huge number of other PPG based PVC records. In addition, it has been observed from the adopted PPG dataset that, within the chosen interval, few PPG record contains no PVC beats at all compared to other records. Investigation of those records over some additional time interval is required to study the occurrence

of PVC beats in the PPG signal. However, this will increase the additional computational burden and will also add extra memory requirements. In the proposed research, careful inspection of the ECG signal beats is carried out from the beginning, to identify the presence of the PVC beats only. Although, other arrhythmia types, for example premature atrial contraction, atrial fibrillation, atrial flutter, etc could also affect PPG signal. The effect of these arrhythmia types on the proposed features and consequent classification outcome has not been considered in this research and will be studied in the future.

The discrimination efficiency of the extracted features is verified by a number of other standard classification techniques. A high value of accuracy reported by each of the classification techniques indicates the robustness of the proposed features and also demonstrates the usefulness of the PPG signal features for the detection of PVC beats. Although considering computational overheads and low-cost health monitoring applications, only the threshold-based classification method is chosen in this research. However, it is acknowledged that further optimization in the performance of the proposed method against other new and superior classifiers are an open area of research to date.

All the experimentation of the algorithm is performed and justified by executing it on the software level and no real-time hardware platform is developed in the present study. The evaluation result is prepared using MATLAB on a personal computer. As a whole, the overall process of fiducial point identification, extraction of features and at last, classification using a simple thresholding criterion takes around 10 seconds on average to identify the presence of the PVC beat, while evaluated over a PPG signal record of one-hour duration. The execution time is calculated by taking the entire database under consideration and only the average value is reported. Clearly, the high computational efficiency of the algorithm establishes its utility to be used in state-of-the-art health monitoring devices to serve a wide variety of populations and to prevent cardiac disease-related mortality.

## 5. Conclusion

A simple and automated algorithm is developed in this research, for primary level detection of the PVC using PPG signal features only. In comparison with other ECG-based approaches, the acquired PPG signal using a single channel provides a much easier, non-invasive and inexpensive way out. The proposed onset detection algorithm also follows signal derivative, amplitude threshold and slope reversal based less complicated techniques. Evidently, the use of these simple methodologies establishes immense promise for hardware implementation of the algorithm. The proposed algorithm employs a beat-to-beat variation of only two statistical parameters from the PPG signal, kurtosis and skewness as features and a simple threshold-based decision

rule method for the detection of PVC with an accuracy of 99.05%. The proposed model is examined using numerous PPG records, acquired from the ICU patients of the standard MIMIC database.

The proposed simple methodologies, high-speed execution and the promising outcome of the algorithm ensure its compatibility in cutting-edge health monitoring devices which can be applied to minimize the mortality rate due to PVC. The entire methodology is also indicative of the fact that PPG signal features can be applied for the efficient detection of PVC. In the future, further modification of this algorithm over a larger subject population along with advanced signal processing techniques is also suggested in order to improve the detection efficiency.

## Acknowledgement

The authors express their deepest gratitude to TEQIP-Phase III, University College of Technology, University of Calcutta, India for providing financial assistance.

## References

- [1] Mendis S, Puska P and Norrving B 2011 Global atlas on cardiovascular disease prevention and control. World Health Organization, pp 2–14
- [2] Al-Yarimi F A M, Ali Munassar N M and Al-Wesabi N F 2020 Electrocardiogram stream level correlated patterns as features to classify heartbeats for arrhythmia prediction. *Int. J. Computat. Math. Electr. Electron. Eng.* 54(5)
- [3] Zarei R, He J, Huang G and Zhang Y 2016 Effective and efficient detection of premature ventricular contractions based on variation of principal directions. *Digital Signal Process. Rev. J.* 50: 93–102
- [4] Munoz Del Carpio, Syed F F, Noheria A, Cha Y M, Friedman P A, Hammill S C, Munger T M, Venkatachalam K L, Shen W K and Packer D L 2011 Characteristics of premature ventricular complexes as correlates of reduced left ventricular systolic function: Study of the burden, duration, coupling interval, morphology and site of origin of PVCs. *J. Cardiovas. Electrophys.* 22: 791–798
- [5] Ephrem G, Levine M, Friedmann P and Schweitzer P 2013 The prognostic significance of frequency and morphology of premature ventricular complexes during ambulatory Holter monitoring. *Ann. Noninvasive Electrocardiol.* 18(2): 118–125
- [6] Watanabe H, Tanabe N, Makiyama Y, Chopra S S, Okura Y, Suzuki H, Matsui K, Watanabe T, Kurashina Y and Aizawa Y 2006 ST-segment abnormalities and premature complexes are predictors of new-onset atrial fibrillation. *Niigata Prevent. Med. Study, Am. Heart J.* 152(4): 731–735
- [7] Santoro F, Biase L D, Hranitzky P, Sanchez J E, Santangeli P, Perini A P, Burkhardt J D and Natale A 2014 Ventricular fibrillation triggered by PVCs from papillary muscles: Clinical features and ablation. *J. Cardiovascular Electrophys.* 25(11): 1158–1164

- [8] Bouaziz F, Oulhadj H, Boutana D and Siarry P 2019 Automatic ECG arrhythmias classification scheme based on the conjoint use of the multi-layer perceptron neural network and a new improved metaheuristic approach. *IET Signal Process.* 13(8): 726–735
- [9] Hickey B, Heneghan C and De Chazal P 2004 Non-episode-dependent assessment of paroxysmal Atrial Fibrillation through measurement of RR interval dynamics and atrial premature contractions. *Ann. Biomed. Eng.* 32(5): 677–687
- [10] Wang J S, Chiang W C, Hsu Y L and Yang Y T C 2013 ECG arrhythmia classification using a probabilistic neural network with a feature reduction method. *Neurocomputing* 116: 38–45
- [11] Inan O T, Giovangrandi L and Kovacs G T A 2006 Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE Trans. Biomed. Engineering* 53(12): 2507–2515
- [12] Khatibi T and Rabinezhadsadatmahaleh N 2019 Proposing feature engineering method based on deep learning and K-NNs for ECG beat classification and arrhythmia detection. *Australas Phys. Eng. Sci. Med.*
- [13] Barhatte A, Dale M and Ghongade R 2019 Cardiac events detection using curvelet transform. *Sadhana* 44: 47
- [14] Zheng Y, Poon C C Y, Yan B P and Lau J Y W 2016 Pulse Arrival Time Based Cuff-Less and 24-H Wearable Blood Pressure Monitoring and its Diagnostic Value in Hypertension. *J. Med. Syst.* 40 (9):
- [15] Habbu S, Dale M and Ghongade R 2019 Estimation of blood glucose by non-invasive method using photoplethysmography. *Sadhana* 44: 135
- [16] Allen J 2007 Photoplethysmography and its application in clinical physiological measurement. *Physiol. Measure.* 28(3): 1–39
- [17] Elgendi M 2012 On the analysis of fingertip photoplethysmogram signals. *Current Cardiol. Rev* 8(1): 14–25
- [18] Islam M T, Zabir I, Ahamed S T, Yasar M T, Shahnaz C and Fattah S A 2017 Frequency domain approach of heart rate estimation from Photoplethysmographic signal. *Biomed. Signal Process. Control* 36: 146–155
- [19] Butter C *et al.* 2004 Cardiac resynchronization therapy optimization by finger plethysmography. *Heart Rhythm* 1(5): 568–575
- [20] He X, Goubran R A and Liu X P 2014 Secondary peak detection of PPG signal for continuous cuffless arterial blood pressure measurement. *IEEE Trans. Instrum. Measure.* 63(6): 1431–1439
- [21] Sadhukhan D, Dhar S, Pal S and Mitra M 2019 Automated Screening of Myocardial Infarction based on Statistical Analysis of Photoplethysmographic data. *IEEE Trans. Instrum. Measure.* 1–1
- [22] Chakraborty A, Sadhukhan D, Pal S and Mitra M 2020 Automated myocardial infarction identification based on interbeat variability analysis of the photoplethysmographic data. *Biomed. Signal Process. Control* 57
- [23] Suzuki T, Kameyama K I and Tamura T 2009 Development of the irregular pulse detection method in daily life using wearable photoplethysmographic sensor. In: *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009* 6080–6083
- [24] Shelley KH 2007 Photoplethysmography: beyond the calculation of arterial oxygen saturation and heart rate. *Anesthesia Analgesia* 105(SUPPL. 6)
- [25] Gil E, Laguna P, Martínez J P, Barquero-Pérez Ó, García-Alberola A and Sörnmo L 2013 Heart rate turbulence analysis based on photoplethysmography. *IEEE Trans. Biomed. Eng.* 60(11): 3149–3155
- [26] Sološenko A, Petrenas A and Marozas V 2015 Photoplethysmography-Based Method for Automatic Detection of Premature Ventricular Contractions. *IEEE Trans. Biomed. Circuits Syst.* 9(5): 662–669
- [27] Goldberger A L, Amaral L A, Glass L, Hausdorff J M, Ivanov P C, Mark R G, Mietus J E, MoodyPeng G B C K and Stanley H E 2000 PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101(23): 215–220
- [28] Sukor J A, Redmond S J and Lovell N H 2011 Signal quality measures for pulse oximetry through waveform morphology analysis. *Physiol. Measure.* 32(3): 369–384
- [29] Chakraborty A, Sadhukhan D, Pal S and Mitra M 2018 Accurate detection of Dicrotic notch from PPG signal for telemonitoring applications. *Int. J. Biomed. Eng. Technol.* 37(2): 121–137
- [30] Joanes D N and Gill C A 1998 Comparing Measures of Sample Skewness and Kurtosis. *J. R. Statist. Soc. Ser. D (The Statistician)* 47: 183–189