RESEARCH

A novel signal-adaptive multi-feature extraction algorithm for arrhythmia detection

L. B. Vinutha[1*] , P. S. Ramkumar[2] and Rajashekar Kunabeva[1]

Abstract

Background: The significant features like an amplitude and intervals of electrocardiograph or P-QRS-T wave represent the functionality of the heart. Accurate extraction of these features helps in capturing characteristics of the signal helpful for the detection of cardiac abnormalities. In this paper, a novel signal folding-based algorithm is proposed to obtain detailed information about the complex morphology of signal. It explores the denoising and feature extraction of the specific ECG signals.

Results: The experimental study conducted using MIT-BIH Arrhythmia database ECG records with known conditions of left bundle branch block, right bundle branch block, Wolff-Parkinson-White syndrome beats has been considered. Heart rate values for selected ECG records from MIT-BIH dataset and synthetic signals from ECG simulator yielded the same values and thus validate our approach.

Conclusion: The proposed algorithm determines the heart rate, percentage leakage around the peak and is capable of folding a signal very efficiently based on detected R peaks and period-dependent gate(window).

Keywords: ECG, Arrhythmia, Left bundle branch block (LBBB), Right bundle branch block (RBBB), Wolff-Parkinson-White syndrome (WPW), Signal folding

Background

The flow of electrical impulses throughout the heart forms the heartbeat expressed in beats per minute (bpm). If these impulses do not flow in a right path, it leads to various cardiac abnormalities called as arrhythmia. Normal heart beat range is 60–100 bpm. If the heart beats too fast (HR > 100 bpm), it is known as a tachycardia (Jones 2021), and if the heart beats too slow (HR < 60 bpm), it is known as bradycardia (Jones 2021). Since heart rate is one of the deciding factors for detecting arrhythmias, it is considered as one of the vital parameters.

The electrocardiogram (ECG) is one of the most important and non-invasive tool to monitor the heart functionalities. P wave of ECG signal reflects the atrial depolarization, QRS complex represents the ventricular depolarization and T wave represents the ventricular

repolarization. The variations in the morphology of P-QRS-T waves may specify underlying abnormalities. Hence, the detection of morphology variations of these waves is vital for clinical diagnosis. In our work, a novel signal folding (SF)-based approach is used to obtain subtle morphology variations of three arrhythmias namely left bundle branch block (LBBB), right bundle branch block (RBBB) and Wolff-Parkinson-White syndrome (WPW). The heart beats are averaged, and features are extracted from the averaged signal instead of extracting features from the single beat. Baseline wander with the frequency range of 0–0.5 Hz and power line interference of 50/60 Hz are the significant noise which corrupt and change the characteristics of ECG signal. Hence, it is necessary to remove these noises sufficiently before further processing. In this paper, ECG is de-noised in frequency domain to improve the SNR by square root of N which is not possible in time domain. After noise removal, signal is smoothened by FIR filter with Kaiser window because of its greater stability by using FDA tool.

*Correspondence: lbvinutha@gmit.ac.in
[1] Department of E& CE, GM Institute of Technology, Davangere, India
Full list of author information is available at the end of the article

Vinutha *et al. Bull Natl Res Cent*   (2021) 45:164

Page 2 of 12

LBBB is caused due to the pause in the passage way. In LBBB, the impulse first travels down the right bundle branch. Then, the impulse activates the interventricular septum from right to left. Finally, the impulses activate the left ventricle (Coviello 2020). RBBB is due to the blockage in the path where impulses move throughout the heart to make a heartbeat. The block in the route, where electrical impulses move through right bundle branch, is known as *RBBB* (Coviello 2020) (Figs. 1, 2, 3, 4).

In Wolff-Parkinson-White syndrome (WPW), there is an adjunct transmission passage way between atria and ventricles. Electrical impulses are passed so quickly through ventricles that AV node cannot control it (Coviello 2020).

Several methods have been proposed to extract features from the ECG signal. In Tribhuvanam et al. (2019), authors have extracted features from the single ECG beat only and can be extended for frequency domain features also. In Ramakrishnan and Yogeswaran (2017), heart rate and time domain features have been extracted from the ECG signal. Features are extracted in time domain. In Deriche et al. (2019), 13 features have been extracted from the signal based on Pan–Tompkins technique and classified with accuracy of more than 92%. In Peshave and Shastri (2014), authors have used DWT and thresholding technique with fixed window size to extract features where the window is not adaptive. In Sujan (2015), authors have extracted features well but it is necessary to extract morphological variations to increase sensitivity. In Riasi and Mohebbi (2015), authors have extracted QRS complex and R peaks based on DWT.

In the proposed method, Raw ECG data from standard database have been taken and denoised to remove baseline wander and power line interference in frequency domain by applying FFT algorithm. After preprocessing, heart rates are calculated in frequency domain instead of complex time domain technique and compared with standard MIT-BIH arrhythmia records. R peaks are extracted by using Pan–Tompkins method. Based on R peak location and heart rate, adaptive window has set to segment the beats and segmented beats are averaged to extract features instead of extracting features from the single beat.

MIT-BIH arrhythmia database from PhysioNet website has been taken to validate the performance of the proposed method which consists of 48 records each of 30 min, 1 min and 10 s duration, having two leads (modified limb lead II and V1, V2, V4 or V5), and each record has been sampled at a rate of 360 samples per second (https://physionet.org/content/mitdb/1.0.0/). In this database, every beat is annotated in terms of timing
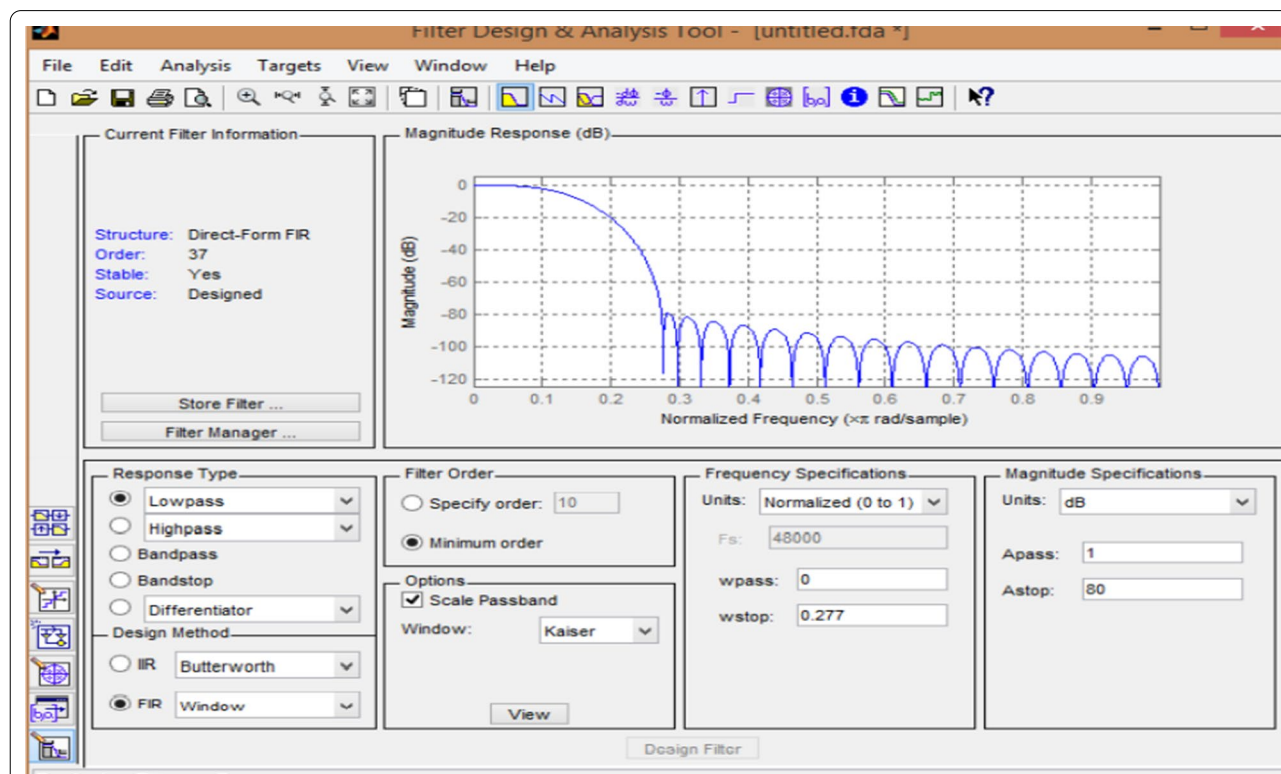
Fig. 1 FDA Tool Diagram: FDA tool is used to design a low-pass FIR filter having order 37 with the normalized cutoff frequency 0.277
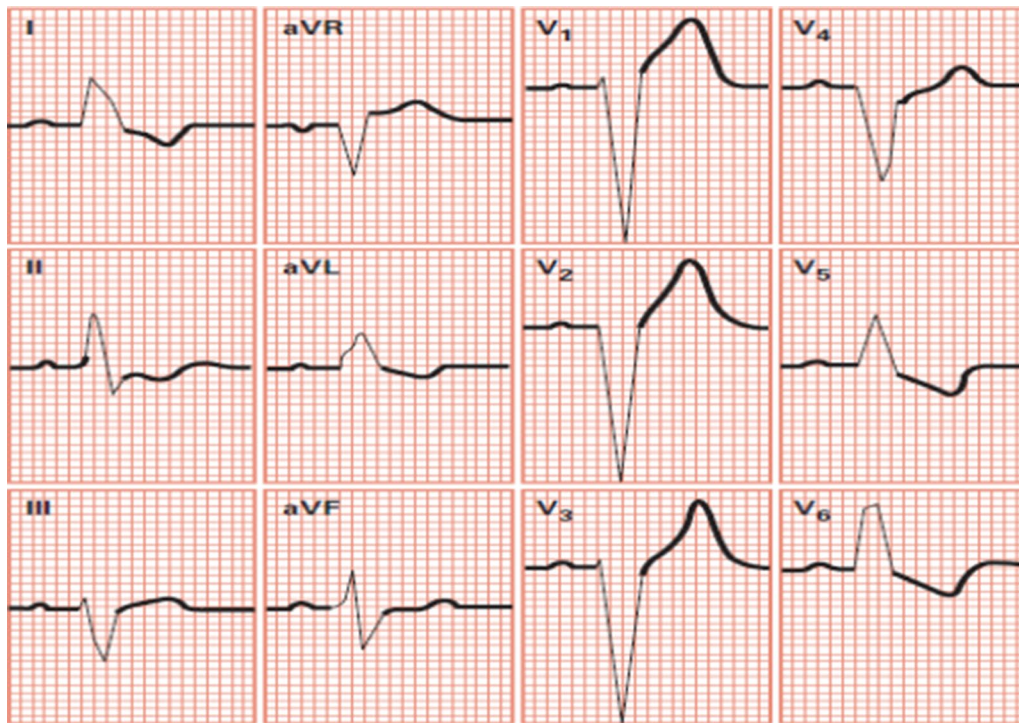
Fig. 2 Morphology of LBBB arrhythmia: Figure shows various morphologies of LBBB arrhythmia with different leads
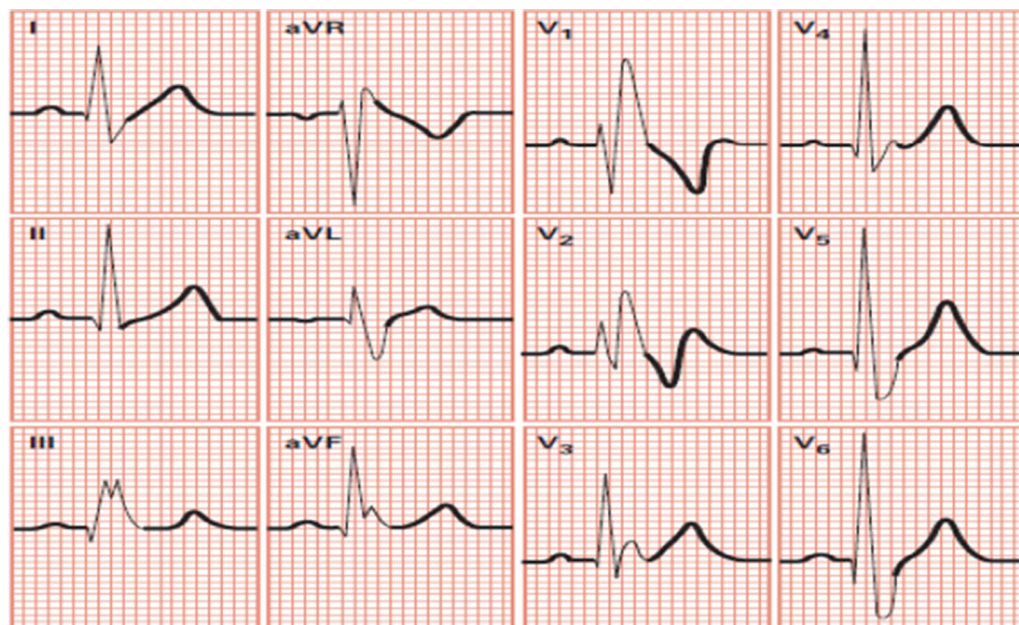
Fig. 3 Morphology of RBBB arrhythmia: Figure shows various morphologies of RBBB arrhythmia with different leads

Vinutha *et al. Bull Natl Res Cent*   (2021) 45:164

Page 4 of 12

Fig. 4 Morphology of WPW arrhythmia: Figure shows various morphologies of WPW arrhythmia with different leads

information and rhythm. Twenty-eight ECG records and also one record from ECG simulator have been considered to calculate heart rates, and specifically LBBB, RBBB and WPW beats have been taken to segment and average the signal.

The rest of the paper is organized as follows: Segment II explains proposed method, Segment III explains the results. Segment IV is concerned with discussion and conclusion.

Methods

The proposed methodology includes several stages: Preprocessing stage to improve the signal quality, feature extraction stage includes extracting heart rate, percentage leakage around the peak in frequency domain and P width and interval detection from the segmented and averaged beats. The main concern of the paper is to extract features such as heart rate and a new parameter percentage leakage around the peak in simple frequency domain and to extract the features from the segmented and overlapped beats so that the noise will be averaged out and subtle features can be detected instead of considering single beat. The proposed algorithm was implemented through MATLAB software. The complete flowchart of the proposed method is shown in Fig. 5.

A. *ECG signal*

To test the proposed method, 28 records of MIT-BIH arrhythmia database from PhysioNet website (https://physionet.org/content/mitdb/1.0.0/) have been used to find heart rate and percentage leakage around the peak. ECG record number 230 containing more WPW rhythm, 111 containing more LBBB beats and 212 for RBBB beats have been considered to fold the signal.

B. *Frequency domain representation*

Raw ECG signal taken from the MIT-BIH arrhythmia record is converted from time to frequency domain by applying FFT algorithm

III. *De-noising*

In the proposed method, noise removal is achieved in frequency domain by converting time domain signal into frequency domain. Noise such as baseline wander and power line interference have been removed efficiently. Firstly, baseline wander is removed by assigning zeros to frequencies from 0 to 0.5 Hz in frequency domain. Power line interference(PLI) is removed by applying simple notch filter with stop band of 45–75 Hz by using threshold (mean + 3*standard deviation) so that signal frequencies within that stop band would not be affected. Then, the signal is smoothened by using FIR filter with Kaiser window from 0.5 to 5 Hz since person will not alive or will be mortal below lower limit (30 bpm) and beyond upper limit (300 bpm) considering only heart rate. FIR filter with the cutoff frequency of 5 Hz is designed by using standard FDA tool.

The FIR filter is designed as follows:

Filter—Low-pass filter

Window—Kaiser window

Normalized frequency = 0.277 Hz

Total number of samples taken from MIT-BIT arrhythmia record (N): 650,000

Because of symmetry property of FFT algorithm, it is sufficient to take only 325,000 samples ($N/2 = 650,000/2$). Lower limit (30 bpm) and upper limit (300 bpm) for heart rate behind which man will be mortal.

$$\text{Frequency for normal heart rate}(f_{hr}) = 72/60 = 1.2\,\text{Hz}$$

$$\text{Lower frequency}(f_L) = \text{lower limit}/60 = 30/60 = 0.5\,\text{Hz}$$

$$\text{Higher frequency}(f_H) = \text{upper limit}/60 = 300/60 = 5\,\text{Hz}$$

$$\text{Frequency Resolution} = 5.5385 * 10^4\,\text{bits/sample}$$
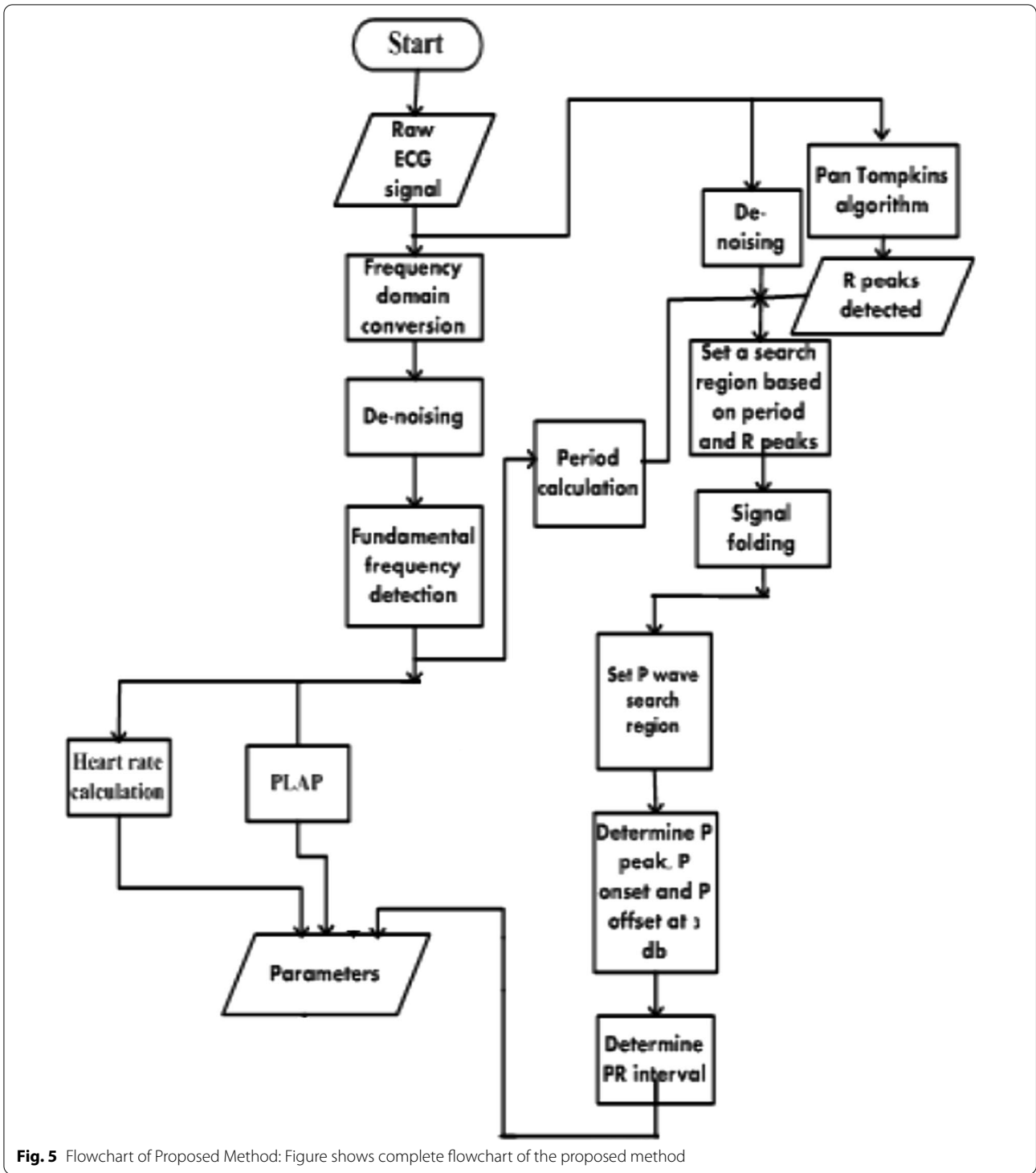
Fig. 5 Flowchart of Proposed Method: Figure shows complete flowchart of the proposed method

Vinutha *et al. Bull Natl Res Cent*   (2021) 45:164

Page 6 of 12

Number of samples after which fundamental frequency lie
N_{samp} = fundamental frequency/frequency resolution

$$= 1.2/\left(5.5385 * 10^{-4}\right) = 2166 \, \text{samples}. \tag{1}$$

Number of samples after 0.5 Hz frequency

$$N \, \text{lower frequency} = \left(f_L * N_{samp}\right)/f_{hr} \tag{2}$$
$$= (0.5 * 2166)/1.2 = 902.5 \, \text{samples}$$

Including 10 harmonics leads to $f_{L10} = 10* N_{lower \, frequency} = 9025$ samples)

Number of samples after 5 Hz frequency

$$N \, \text{upper frequency} = \left(f_H * N_{samp}\right)/f_{hr} \tag{3}$$
$$= (5 * 2166)/1.2 = 9025 \, \text{samples}$$

Including 10 harmonics leads to $f_{H10} = 10* N_{upper \, frequency} = 90250$ samples)

De-noised signal is then fed to the designed low-pass filter (LPF) with the normalized frequency

$$F_{norm} = f_{H10}/(N/2) \tag{4}$$
$$= (90250/325000) = 0.277$$

to obtain smoothened signal.

IV. *Feature extraction*

Feature extraction is one of the major stages to classify normal and abnormal ECG beats. In the proposed method, frequency domain features such as HR and PLAP are extracted and three specific records are used to segment the beats based on adaptive window, and then these segmented beats are averaged to get p width and intervals.

(1) *Heart rate (HR)*

Heart rate is calculated by taking fundamental frequency or maximum frequency of the ECG signal which represents the averaged R-R interval in time domain and can be calculated as (Vinutha et al. 2016)

$$HR = F1 \max *60 \tag{5}$$

where $F1\max$ is the maximum frequency (Hz). Where, ledge and uedge are the points at 3db on either side of the fundamental frequency peak shown in Fig. 6

A simple and accurate frequency domain approach has been used to extract HR. Further, R peak detection has been performed using Pan–Tompkins algorithm in time domain. R peaks and heart rates are used for novel SF algorithm. In this approach, signal is folded based on the period-dependent window which is adaptive based on the signal to get averaged information and subtle variations of the signal. This approach averages out all the noise and also capable of showing detailed morphology of the signal.

(2) *Period calculation*

In order to set a window, period is calculated based on the heart rate which in turn depends on maximum frequency. Period is calculated as follows

$$\text{Period} = \text{sampling frequency/maximum frequency} \tag{6}$$

Window is adaptive and adjustable to all types of arrhythmias.

Fig. 6 Fundamental frequency and 3db points on FFT signal: Fundamental frequency in this figure represents the average heart rate of the signal where most of the information is preserved in fundamental frequency and standard 3db points have taken on either side of fundamental frequency peak to find out the wobbling of the heart rate

Vinutha *et al. Bull Natl Res Cent*   (2021) 45:164

Page 7 of 12

E. *Signal folding (SF) algorithm*

The proposed method uses novel signal folding method. In this method, instead of extracting features from the single beat, the averaged features could be extracted such as p width and PR interval, by dividing and folding(averaging) the signal based on the R peak location and adaptive period to get the subtle variations in the signal so that accurate results will be obtained. The window for particular signal is set as follows:

Adaptive period value is halved as calculated in formula 6. By taking R peak location as a reference, the window is set prior and after the R peak location by considering halved

overlapped on one another to get the subtle morphological variations in signal instead of considering single beat. The noise removed ECG signal is divided and averaged to get accurate results. In this paper WPW, LBBB, RBBB beats are folded in order to obtain subtle information, small variations, and clear signature of these arrhythmias.

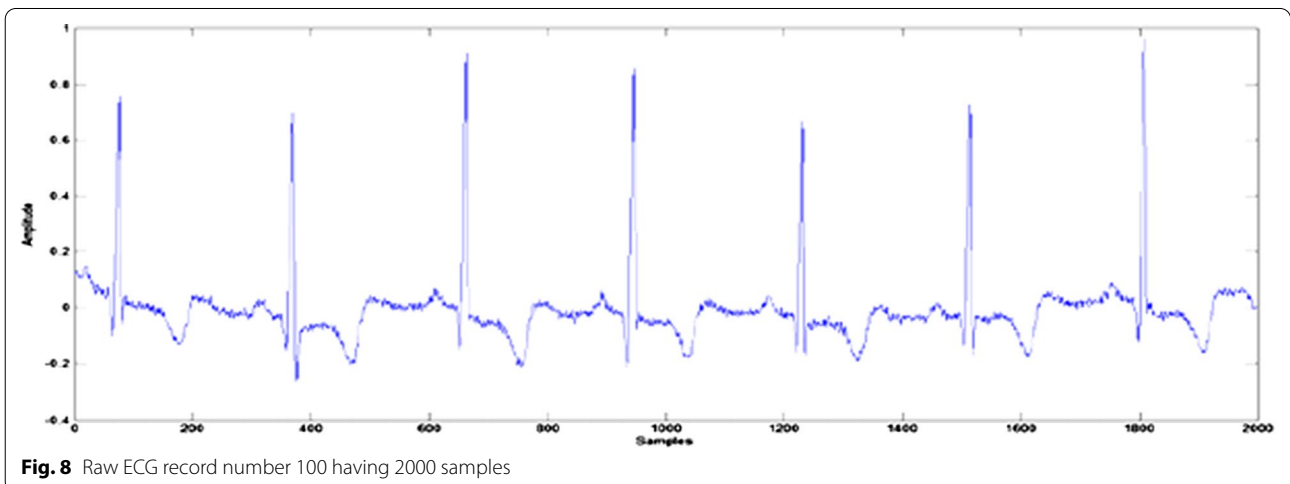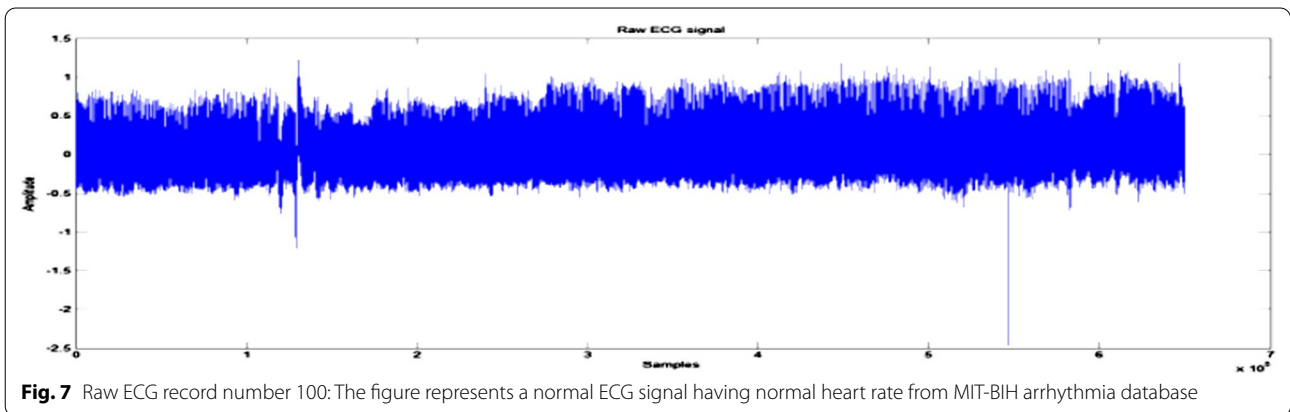### F. *Percentage leakage at each point around peak (PLAP)*

The novel feature provides information about duration of the heart beat resides at each point around the peak. It gives the amount of time heart beat spends at each point around peak.
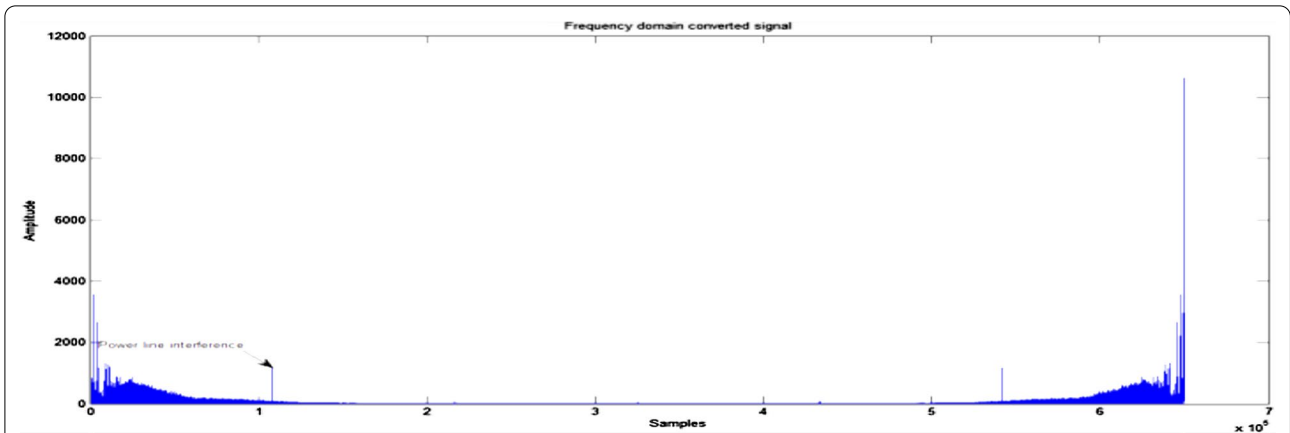
$$\text{PLAP} = (\text{Peak points around mean heart rate}(P_{\mathrm{m}}) \\ /\text{Sum of all peak points around the mean heart rate}(S_{\mathrm{m}})) * 100 \tag{9}$$

period value on either side of the location. After setting adaptive window, 30 min baseline removed ECG signal is divided based on the period-dependent gate(adaptive) and

Results

See Figs. 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 and Tables 1, 2.

Fig. 7 Raw ECG record number 100: The figure represents a normal ECG signal having normal heart rate from MIT-BIH arrhythmia database

Fig. 8 Raw ECG record number 100 having 2000 samples

Vinutha *et al. Bull Natl Res Cent*   *(2021) 45:164*

Page 8 of 12

Fig. 9 FFT signal of record number 100 showing PLI: The raw ECG signal of record number 100 is converted to frequency domain by using fast Fourier transform which is showing power line interference noise at 60HZ
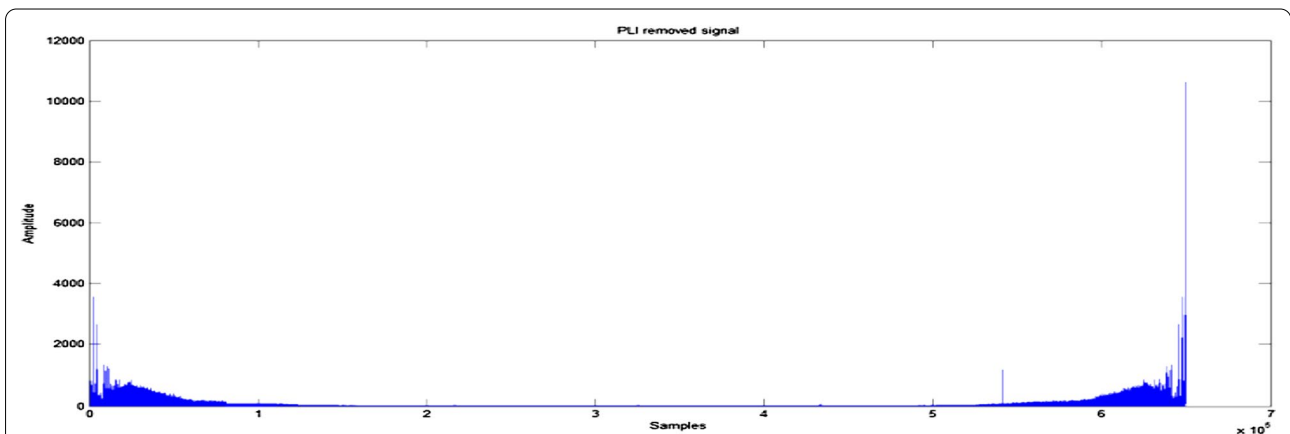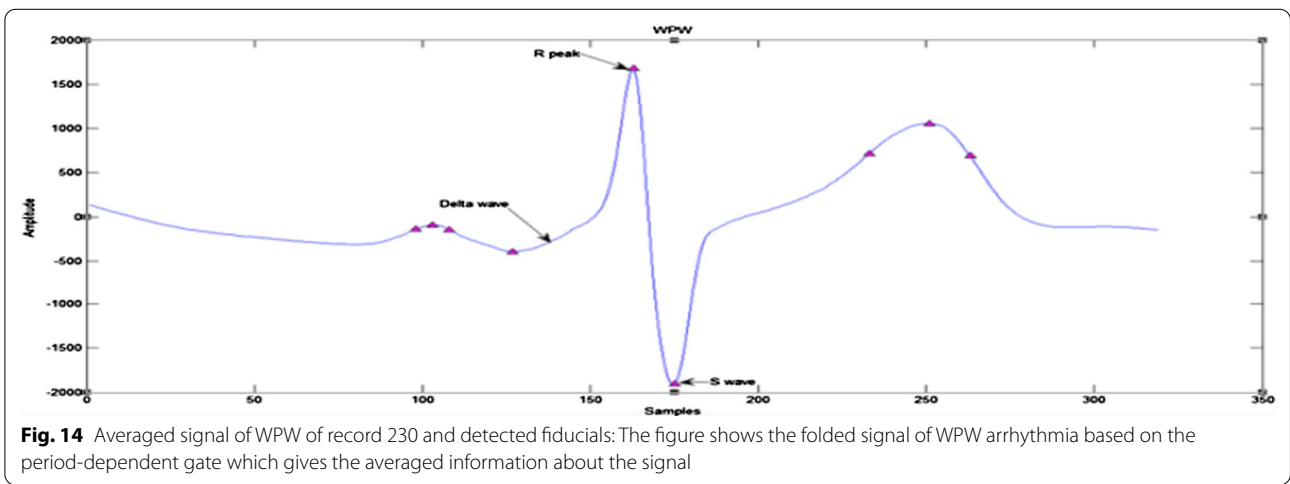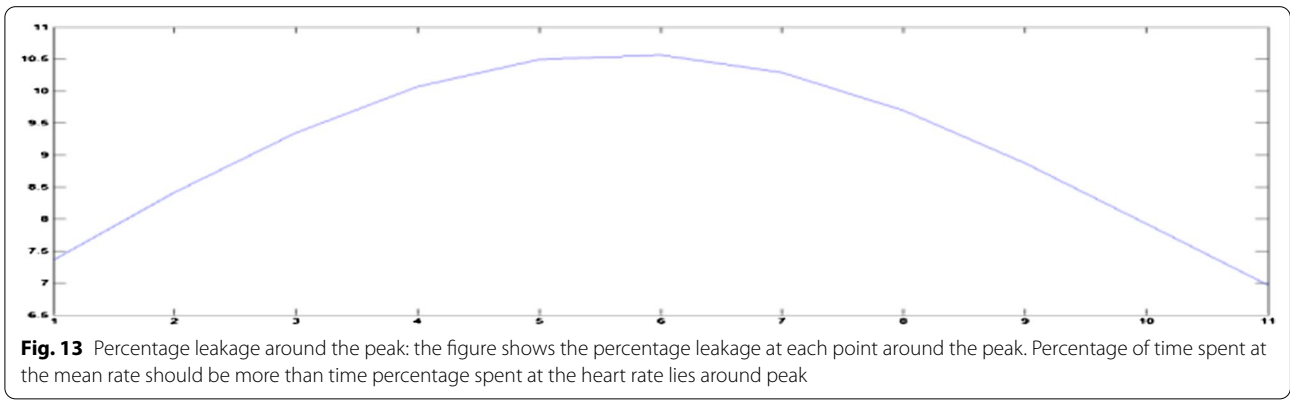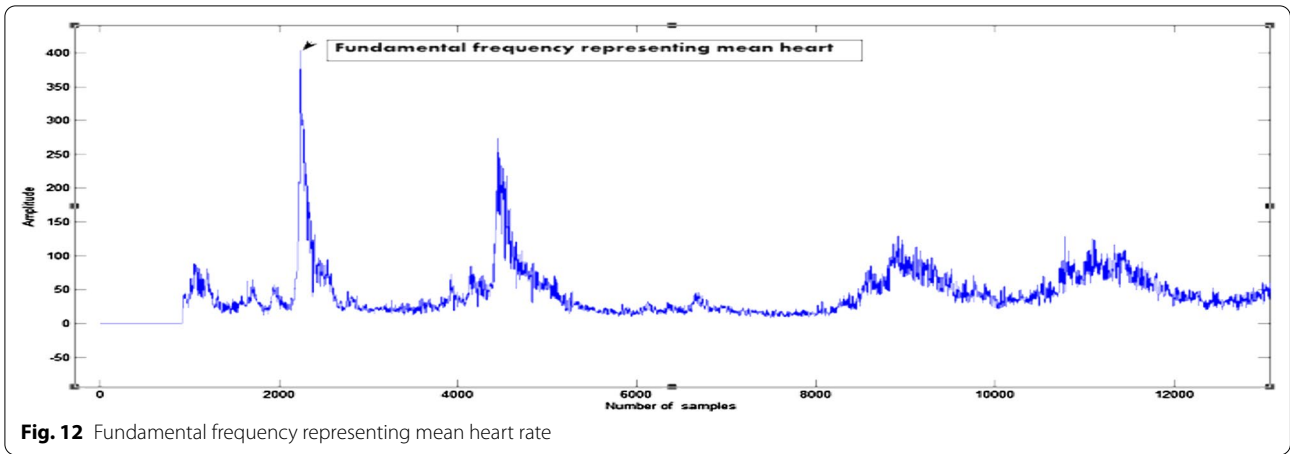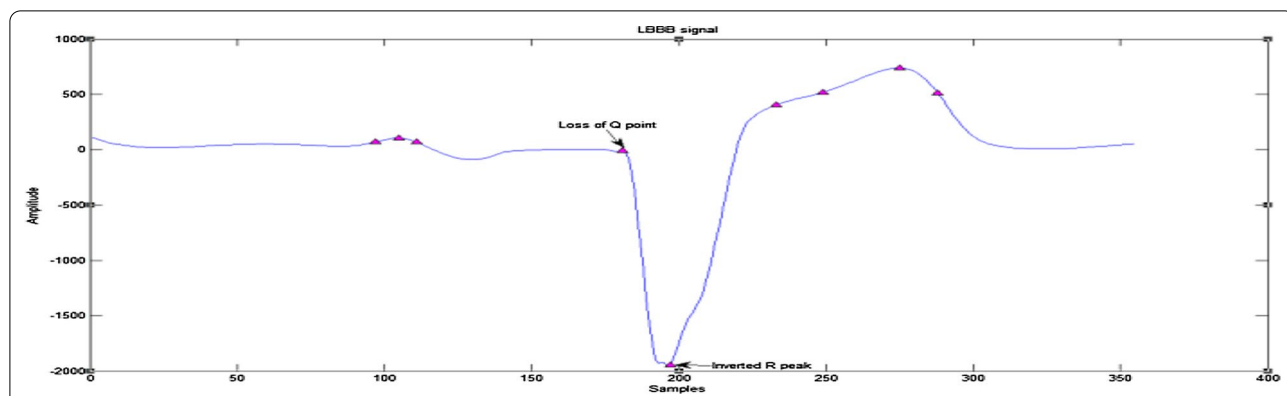

Fig. 10 PLI removed signal of record number 100: The figure represents power line interference removed signal in frequency domain at 60 Hz
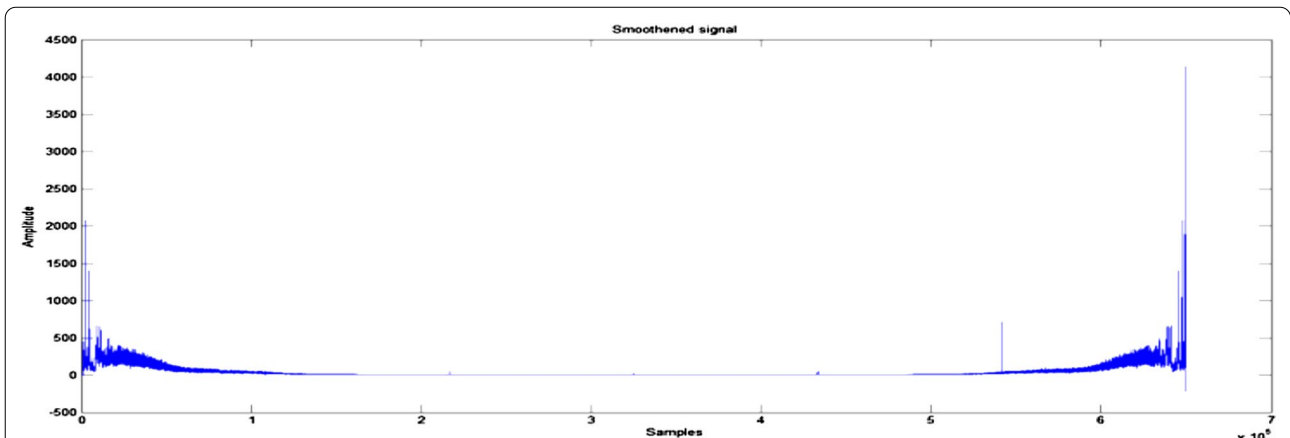

Fig. 11 Smoothened signal by FIR filter: Figure represents low-pass filtered signal with the Kaiser window having normalized cutoff frequency 0.277

Fig. 12 Fundamental frequency representing mean heart rate

Fig. 13 Percentage leakage around the peak: the figure shows the percentage leakage at each point around the peak. Percentage of time spent at the mean rate should be more than time percentage spent at the heart rate lies around peak

Fig. 14 Averaged signal of WPW of record 230 and detected fiducials: The figure shows the folded signal of WPW arrhythmia based on the period-dependent gate which gives the averaged information about the signal

Fig. 15 Averaged signal of LBBB of record 111 and detected fiducials: The figure shows the folded signal of LBBB arrhythmia based on the period-dependent gate which gives the averaged information about the signal
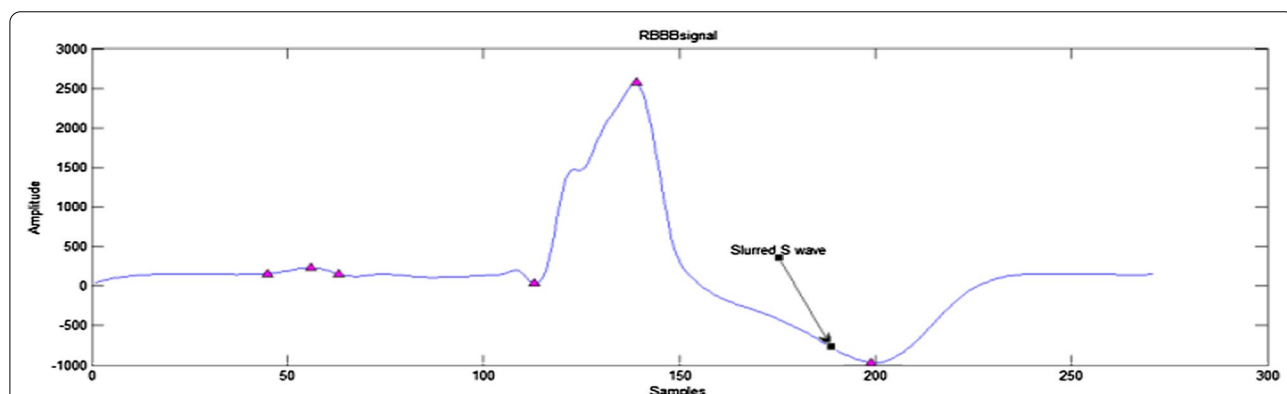
Fig. 16 Averaged signal of RBBB of record 212 and detected fiducials: The figure shows the folded signal of RBBB arrhythmia based on the period-dependent gate which gives the averaged information about the signal

Discussion

To test the proposed algorithm, 30 min ECG signals from MIT-BIH Arrhythmia database were used. Heart rate is calculated for 28 records in frequency domain and compared with annotated values of heart rates of standard MIT-BIH database. Specific records containing more LBBB beats, RBBB beats and WPW beats are folded based on detected R peaks and period-dependent gate. The raw estimates of HR and PLAP results are summarized. LBBB, RBBB, and WPW beats have been folded showing the signatures of these arrhythmias efficiently. Table 1 illustrates the calculated heart rates from the proposed algorithm compared with heart rates of MIT-BIH database and heart rate variability calculations. Table 2 shows the heart rates of only normal beats and compared with heart rates of MIT-BIH database. It is evident that proposed algorithm is capable of calculating heart rates accurately for 26 records out of tested 28 records. In MIT-BIH arrhythmia database directory, every beat is annotated and folded to get the averaged information. Significant features such as heart rate, percentage leakage around the peak and R amplitudes have been taken. Raw ECG signal is shown in Fig. 7 representing number of samples on X axis and amplitude on Y axis. Raw ECG signal having 2000 samples is shown in Fig. 8. Frequency domain representation of raw ECG signal is shown in Fig. 9 with PLI peak at 60 Hz. Figure 10 shows PLI removed signal by notch filter. Figure 11 shows smoothened signal by applying FIR filter with Kaiser window. Figure 12 shows the

Vinutha *et al. Bull Natl Res Cent*   (2021) 45:164

Page 11 of 12

Table 1 Heart rates from the proposed method compared with standard MIT-BIH records (Vinutha et al. 2016)

Record number	Heart rate(bpm)	Heart rate from MIT-BIH database(bpm)
100(Normal)	74.5324	70–89
102	73.2032	72–78
104	74.5988	69–82
105	83.8365	78–102
106	132.0848	49–121
107	71.2095	68–82
108	56.2565	44–78
109(LBBB)	84.2352	77–101
111(LBBB)	83.8365	64–82
112	83.7368	74–91
113	31.2351	48–87
114	55.4922	51–82
116	80.5468	74–86
117	51.8038	48–66
118(RBBB)	72.3060	54–91
121	60.2772	55–83
122	87.0597	67–97
123	50.6740	41–65
124(RBBB)	57.3863	47–64
205	82.3744	80–216
208	127.9312	91–134
212(RBBB)	91.8114	63–108
215	111.6823	81–124
220	69.1161	58–74
228	83.8365	50–88
230	77.5562	63–99
231(RBBB)	62.8026	49–69
234	90.1832	84–99
ECG6120 (From simulator generating 120 bpm)	120.37	120 (From simulator)

fundamental frequency representing mean heart rate. Figure 13 shows the percentage leakage at each point around the peak. The percentage of time spent at the mean rate should be more than time percentage spent at the heart rate that lies around the peak. Figures 14, 15 and 16 are the folded signals of WPW, LBBB, RBBB beats.

Conclusions

The proposed algorithm determines the heart rate, percentage leakage around the peak and is capable of folding a signal very efficiently based on detected R peaks and period-dependent gate(window). As case study, three ECG arrhythmia records have been folded efficiently to get averaged widths and intervals. The features of WPW are delta wave at Q point and wide QRS complex. The proposed algorithm folded the signal very efficiently and showed the signature of WPW correctly. The features of LBBB and RBBB beats are loss of Q point, wide QRS complex and slurred S wave, M shaped R peak, respectively, where signal folding showed signature of those beats very effectively. Features and intervals calculated from the folded signal give the averaged information than extracting features from each beat directly.

In this paper, only features have been extracted from the ECG signals such as HR and PLAP in frequency domain, and specific signals are segmented based on adaptive window and averaged to extract features. The proposed method is used for only ambulatory ECG. Further, the work can be extended to use these features to detect remaining arrhythmias.

Table 2 Heart rates from the proposed method compared with standard MIT-BIH records of normal beats

Record number	Number of samples	Type of beats	Heart rate (bpm)	Heart rate (MIT-BIH database)
100	1 to 650,000	Normal beats	74.53	70–89 bpm
105	1 to 650,000	Normal beats	83.8365	78–102 bpm
205	1 to 650,000	Normal with ventricular tachycardia	132.0848	75–216 bpm
212	1 to 650,000	Normal beats	91.8114	63–108 bpm
230	1 to 650,000	Normal beats	77.2968	63–99 bpm
222	1 to 191,381	Normal beats	76.2812	49–84 bpm
223	1 to 42,685	Normal beats	90.5014	75–94 bpm

Abbreviations
ECG: Electrocardiogram; LBBB: Left bundle branch block; RBBB: Right bundle branch block; WPW: Wolff-Parkinson-White syndrome; HR: Heart rate; PLAP: Percentage leakage around the peak.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the MIT-BIH arrhythmia database containing ECG records.

Declarations

Ethics approval and consent to participate
Dataset is taken from https://physionet.org/content/mitdb/1.0.0/.

Consent for publication
Not applicable.

Competing interests
Not applicable.

Author details
[1]Department of E& CE, GM Institute of Technology, Davangere, India. [2]Applied Cognition Systems Pvt. Limited, Bengaluru, India.

References
Coviello JS (2020) ECG interpretation made incredibly easy, 5th edn. Lippincott Williams & Wilkins, Philadelphia

Deriche M, Aljabri S, Al-Akhras M, Siddiqui M, Deriche N (2019) An optimal set of features for multi-class heart beat abnormality classification. In: 16th International multi-conference on systems, signals & devices (SSD'19), IEEE

https://physionet.org/content/mitdb/1.0.0/

Jones SA (2021) "ECG Notes" interpretation and management guide. FA Davis, Philadelphia

Peshave JD, Shastri R (2014) Feature extraction of ECG signal. In: International conference on communication and signal processing, April 3–5, 2014, India

Ramakrishnan S, Yogeswaran R (2017) Design and analysis of feature extraction algorithm for ECG signals using adaptive threshold method. In: IEEE

Riasi A, Mohebbi M (2015) Prediction of ventricular tachycardia using morphological features of ECG signal. In: International symposium on artificial intelligence and signal processing

Sujan KSS (2015) Performance analysis for the feature extraction algorithm of an ECG signal. In: IEEE sponsored 2nd international conference on innovations in information, embedded and communication systems (ICIIECS)

Tribhuvanam S, Nagaraj HC, Naidu VPS (2019) ECG abnormality classification with single beat analysis. In: International conference on vision towards emerging trends in communication and networking

Vinutha LB, Ramkumar PS, Sushmitha BS (2016) Heart rate and heart rate variability estimation using spectral analysis. Int J Innov Res Sci Eng Technol 5(10):1–7

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.