CrossMark

A novel LMS algorithm for ECG signal preprocessing and KNN classifier based abnormality detection

C. Venkatesan[1] · P. Karthigaikumar[2] · R. Varatharajan[3]

Abstract ECG signal abnormality detection is useful for identifying heart related problems. Two popular abnormality detection techniques are ischaemic beat classification and arrhythmic beat classification. In this work, ECG signal preprocessing and KNN based arrhythmic beat classification are performed to categorize into normal and abnormal subjects. LMS based adaptive filters are used in ECG signal preprocessing, but they consume more time for processing due to long critical path. To overcome this problem, a novel adaptive filter with delayed error normalized LMS algorithm is utilized to attain high speed and low latency design. Low power design is achieved in this design by applying pipelining concept in the error feedback path. R-peak detection is carried out in the preprocessed signal using wavelets for HRV feature extraction. Arrhythmic beat classification is carried out by KNN classifier on HRV feature extracted signal. Classification performance reveals that the proposed DWT with KNN classifier provides the accuracy of 97.5% which is better than other machine leaning techniques.

Keywords ECG noise removal · Adaptive filter · Least mean square algorithm · Discrete wavelet transform · Machine learning · KNN classifier

1 Introduction

Cardiovascular diseases are mainly responsible for heart attacks worldwide and their diagnosis from electrocardiogram (ECG) has greater importance in medical applications. Myocardial

✉ R. Varatharajan
varathu21@yahoo.com

1 Faculty of Information and Communication Engineering, Anna University, Chennai, Tamilnadu, India

2 Department of Electronics and Communication Engineering, Karpagam College of Engineering, Coimbatore, Tamilnadu, India

3 Department of Electronics and Communication Engineering, Sri Ramanujar Engineering College, Chennai, India

🖉 Springer

ischaemia and cardiac arrhythmias are the two popular heart diseases. In myocardial ischaemia, blood supply reduction alters the morphology of ECG signal whereas cardiac arrhythmias indicate some critical information that leads to sudden cardiac death [8, 20]. Hence arrhythmic beat classification provides important information about cardiac condition of human beings. Electrocardiogram (ECG) morphological analysis has been used in the past for the assessment of arrhythmias at expense of more computational elements. To overcome the complexity involved in morphology based signal processing, heart rate variability (HRV) analysis become very much popular in ECG processing applications [7]. HRV is used to measure the variations in R-R interval or inter-beat interval (IBI), which is defined as duration observed between consecutive R waves of an electrocardiogram in milliseconds. Heart rate in beats-per-minute (bpm) is computed by $HR = (60000)/(RR$ interval) considering R-R interval in milliseconds.

The preprocessing of ECG signal is usually carried out using adaptive filters, wavelets and empirical mode decomposition (EMD) based techniques. Adaptive finite impulse response (FIR) filters widely preferred in many preprocessing applications, but they could not be utilized due to its high computational cost. However, advances in very large scale integration (VLSI) technology make the adaptive filtering computations in single chip with less number of elements. Among various types of adaptive filtering methods, least mean square (LMS) based adaptive filtering is commonly used for wide range of signal processing applications [24].

In the standard normalized LMS (NLMS) algorithm, a step size parameter can be used to control the convergence performance of the filter [3]. A proper step size improves the convergence performance while reducing the MSE. Variable step size based LMS algorithms have been applied in the weight update of adaptive filters to improve the performance with low computational power [16]. To improve the convergence with low steady state error, error normalized LMS (ENLMS) algorithm has been introduced [13]. The instantaneous data vector is normalized in NLMS algorithm, however the error vector is normalized in ENLMS algorithm. The critical path is generally reduced by pipelined structures with delay elements so that the desired sample period can be attained [11, 19]. In this work, a delayed error normalized LMS (DENLMS) adaptive filter is used for ECG noise removal with less computational complexity [5, 21].

Wavelet transform based signal processing is preferred over filtering methods due to its capability of extracting time-frequency domain information retrieval from the non-stationary signals. Since ECG is non-stationary in nature, wide variety of wavelets has been applied in ECG signal preprocessing and feature extraction [17]. Wavelet decomposition and reconstruction are performed to extract the important features of the ECG signal. The number of coefficients and level of decomposition decide the performance of any wavelet. Coiflet wavelet is one of the popular wavelet that is useful in finding R-peaks in ECG signal [15]. R-peaks are used in finding the HRV of the ECG signals. Though Coiflets are formed from Daubechies wavelet, they are more symmetric and has linear phase and its characteristics are much better than Daubechies wavelet and Spline wavelet [9]. Discrete wavelet transform (DWT) based HRV feature extraction is utilized in this work as it reduces the required number of computational elements [14]. Though many wavelets are available for R-peak detection and HRV feature extraction, Coiflet is chosen in this work to extract QRS complex better and results in determining maximum possible R-peaks [2, 6].

Machine learning techniques are useful in many complex applications where the conventional algorithms are infeasible [18]. Since medical expert systems are part of the portable and smart healthcare monitoring devices used in day to day life, large amount of clinical data are generated and they have to be processed and classified appropriately. The K-nearest neighbor algorithm (KNN) is used for classification of feature extracted ECG data which is a simple

machine learning algorithm compared to other machine learning techniques such as Support Vector Machine (SVM), decision tree classifier and deep learning [4].

Time domain and frequency domain parameters are used in many machine learning based ECG classification applications to identify sympathetic and parasympathetic activities [22]. In general, the low frequency range is normally affected by both sympathetic and parasympathetic activities, where high frequencies are largely found in parasympathetic activity [1]. The extracted time domain and frequency domain features are applied into the machine learning classifier. In this work, KNN classifier is applied for performing classification using DWT based feature extracted time domain and frequency domain features.

This paper focuses on DENLMS adaptive filter based ECG signal preprocessing and DWT based HRV feature extraction. Arrhythmic beat classification is performed by KNN classifier on HRV features and HRV parameters extracted by the DWT. The performance comparison is made with similar machine learning based classifiers. The paper is organized into five sections including the introductory section. ECG signal preprocessing using DENLMS algorithm is focused in Section 2. DWT based feature extraction is explained and KNN Classifier based arrhythmia classification on HRV is described in Section 3. Section 4 discusses the obtained results and Section 5 concludes the proposed work.

2 ECG signal preprocessing using DENLMS algorithm

The LMS algorithm is the most preferred choice in adaptive filters due to its computational simplicity. In each iteration of the standard LMS algorithm, the FIR filter coefficients are modified based on the error output and step size. In this work, delayed error normalized LMS (DENLMS) algorithm is used with the weight update equation

$$w(n+1) = w(n) + \mu_{en}e(n)x(n) \tag{1}$$

Where $x(n)$, $e(n)$, $w(n)$ and $w(n+1)$ represent input, error, old weight and updated weight respectively. The error normalized step size μ_{en} is written based on squared norm of error vector instead of input vector as

$$\mu_{en} = \left[\frac{\mu}{p + e^t(n)e(n)} \right] \tag{2}$$

where parameter p is set to avoid the high denominator value and small step size.

From Eq. (2), if the error is high, the value of step size decreases which may affect the convergence. However, the increase in step size parameter leads to faster convergence, since e(n) decreases in magnitude. The algorithm initially converges slowly due to small μ value. However, when the algorithm reaches the convergence and error $e(n)$ starts to become less in magnitude, the step size happen to be larger, leads to a faster convergence of filter.

3 DWT based feature extraction and abnormality detection using KNN classifier

In DWT based feature extraction, the R-peaks are detected to determine the HRV signal features. Arrhythmic beat classification is performed to detect abnormalities in ECG signal

using KNN classifier. R-peak detection techniques are mainly based on heart rate function which is widely used to calculate the RR interval. Dividing one minute by the instantaneous heart rate gives the RR interval of the given ECG signal. Consecutive RR intervals of the ECG signal are calculated from starting interval of the heart rate function. Among the Meyer, Biorthogonal, reverse Biorthogonal and Coiflet wavelets, Coiflet is chosen to extract R-peaks. In practical RR interval measurement system, correlation technique with timing resolution ± 1 ms is used. The accurate RR interval measurement can be obtained by high performance digital signal processor or customized processor. Considerable amount of baseline trend in ECG signal is removed using the adaptive filtering based preprocessing.

Time domain and frequency domain features can be derived from the extracted HRV features. Figure 1 depicts the components involved in the arrhythmic beat classification based on preprocessing and HRV feature extraction. The proposed ANFIS evaluation consists of i) Collection of raw electrocardiogram (ECG) signal from MIT-BIH database ii) Preprocessing of ECG signal using adaptive filter iii) R-peak detection iv) Frequency domain feature extraction from HRV signal v) KNN classification into normal and abnormal.

There are 14 well known time domain and frequency domain HRV features. Time domain HRV parameters used are RR mean (ms), RR Std (ms), HR mean (bpm), HR Std (bpm), RMSSD (ms), NN50, pNN50, RR Triangular Index, TINN (ms).In this work, six frequency domain features have been utilized. They are VLF Power (ms^2), LF Power (ms^2), HF Power (ms^2), LF norm, HF norm and LF/HF Ratio [12]. The three frequency bands used are VLF, LF, HF and a frequency ratio LF/HF. In addition to these parameters, LF norm and HF norm are calculated in terms of normalized units.

The K-nearest neighbor algorithm (KNN) is a simple machine learning algorithm compared to similar machine learning approaches [23]. KNN classifier is based on grouping of closest
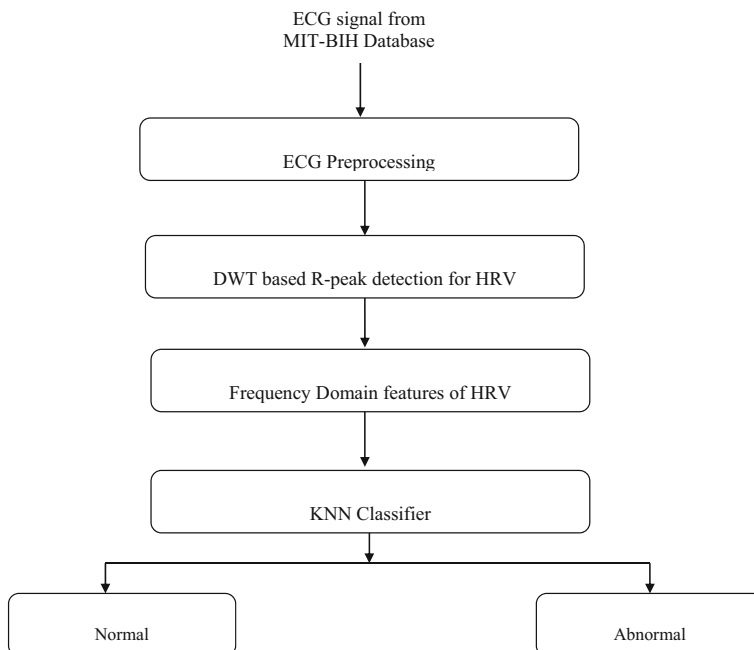
Fig. 1 Steps involved in abnormality detection

Table 1 Comparison of various classification techniques

Technique	Physiological signals and features	Classification accuracy (%)
ANN Approach [18]	ECG, Blood volume	95.5%
Neuro fuzzy system [12]	ECG	94%
KNN Classifier [4]	HRV	96%
DWT with KNN classifier (proposed)	HRV	97.5%

training points of data in the considered feature space. The grouping is done by majority of voting to the nearest neighbor points. In KNN, K value indicates Euclidian and it takes value from 2 to 7 in this work. MIT-BIH database has been chosen for arrhythmic beat classification and abnormality detection. The detected the R waves and RR intervals are used for HRV frequency domain analysis. The frequency domain parameters listed in Table 1 are calculated for the HRV extracted preprocessed data for arrhythmic beat classification.

4 Results and discussion

The simulation experiments are conducted based on MIT-BIH arrhythmia database ECG signal and real-time recorded ECG signals. In real-time ECG recording, ECG data has been collected from both the healthy persons and the people with some heart related problems. The different age groups between 18 and 40 have been chosen with different backgrounds for this study. The database ECG signal is originally obtained by placing the electrodes on the chest in the first channel V1 that is the standard practice in ECG recording. The recorded signal has been digitized at 540 Hz sampling rate with 11-bit resolution over a 10 mV range. All the simulation experiments were conducted using MATLAB® v. 7.10. In the simulation scenario, we choose μ for all the filters as 0.01 and the filter length as 21 for reducing the simulation time. The number of iterations for the experiment is 1000. Five sample data (ECG data record 105, ECG data record 106, ECG data record 107, ECG data record 108 and ECG data record 109) of the database have been used. Moreover, speech signal is considered in simulation to test the performance of DENLMS algorithm for different type of signals. Figure 2 shows the MSE comparison of different algorithms for noise SNR of 2.5 dB.

Filter convergence vary in accordance to step size, it converges faster for high step size. MSE reduces in accordance with the increase in SNR of noise. MSE of 19.5 dB has been achieved for the iteration time110 in DENLMS algorithm which is superior to NLMS, transform domain (TDLMS) and delayed NLMS (DNLMS) algorithms.

The filtering performance of this algorithm is observed using five MIT-BIH data base ECG signals. Figure 3 shows the extraction of clear ECG from corrupted ECG data (record 104).

The denoised ECG is processed using the R-peak detection algorithm to determine the possible beats per minute. The beat rate of 136, 129 and 74 are detected using Meyer wavelet, Biorthogonal wavelet and Coiflet wavelet. Since the human heart rate values vary between 60 and 100 bpm, Coiflet wavelet provides the correct beat rate of 74. The obtained result is shown in Fig. 4.

The frequency domain parameters such as VLF (ms^2), LF (ms^2), HF (ms^2), frequency ratio LF/HF, LF norm and HF norm are noted for all the feature extracted ECG data. National instruments (NI) biomedical kit has been utilized for the computation of frequency domain values. The obtained values are compared among normal and abnormal persons in Fig. 5. These values are averaged to analyze the risk using arrhythmic beat classification.
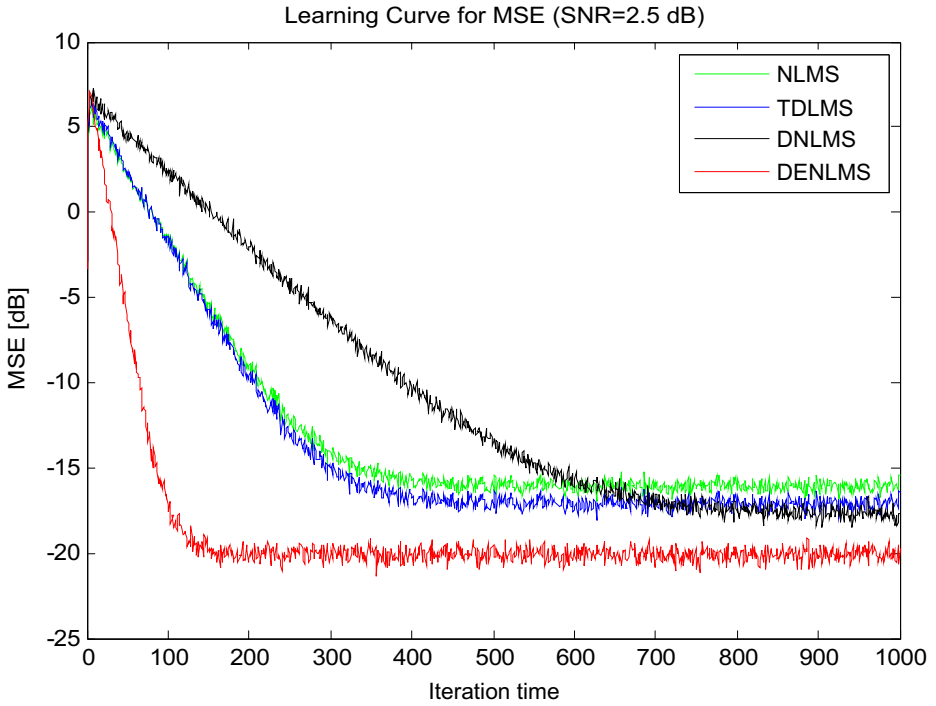
Fig. 2 MSE comparison of LMS algorithms for noise with SNR = 2.5 dB

In KNN classification, frequency domain features have been computed from the total of 160 HRV data. Different K values from 2 to 7 are applied over the different frequency band
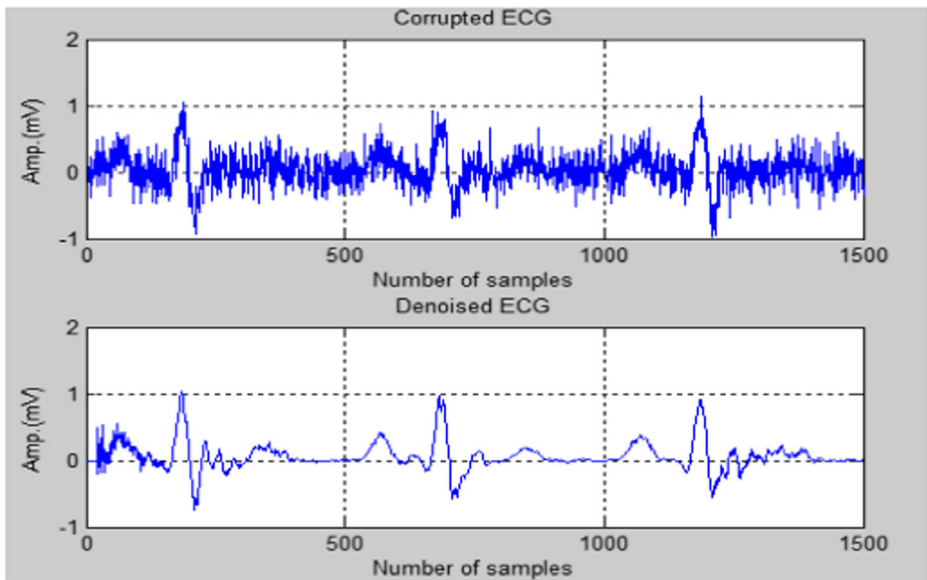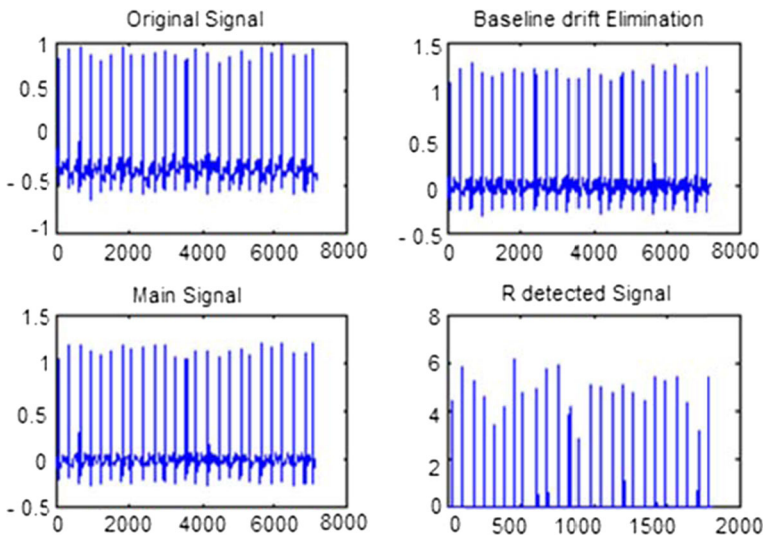

Fig. 3 Original and denoised ECG signals using DENLMS algorithm

Fig. 4 R-peak detection using Coiflet

values for better classification results. Threshold values are fixed for the frequency domain features VLF (ms^2), LF (ms^2), HF (ms^2), frequency ratio LF/HF, LF norm and HF norm while categorizing into normal and abnormal subjects. One hundred twenty-eight training data (80%) with signal length of 30 s in 1000 epochs have been used to train the KNN classifier. After training, 32 testing data (20%) were used to validate the accuracy of the classifier.

In the proposed KNN classification based abnormality detection method, only R peak detection has been carried out after performing ECG signal preprocessing. Several existing techniques are time-consuming and require complex computations. In addition, morphological ECG features are not feasible while dealing with noisy data. Various techniques are compared in Table 1 to categorize the arrhythmic risk abnormal and normal subjects. Some of the exiting classification techniques chosen which are based on ANN, neuro fuzzy and conventional KNN classifiers [10]. In these techniques, ECG, blood volume, HRV parameters were used. The maximum classification accuracy of 96% has been achieved using these techniques. But the experimental result of KNN based classifier gives a maximum accuracy of 97.5% on classifying normal and arrhythmic risk abnormal subjects.
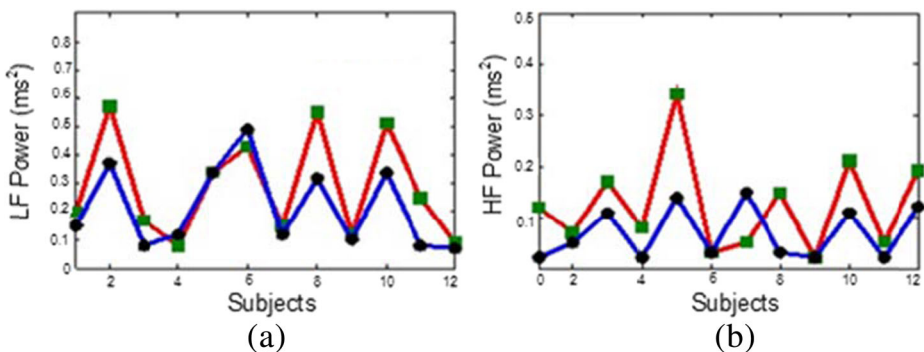
Fig. 5 Comparison of normal and abnormal subjects based on (**a**) LF power (**b**) HF Power

5 Conclusion

ECG signal preprocessing is carried out by DENLMS algorithm based adaptive filter. It is observed that MSE value reduces in accordance with the increase in SNR of noise. MSE of 19.5 dB has been achieved for the iteration time110 in DENLMS algorithm which is the performance improvement of 24.35% while comparing with the TDLMS algorithm. In R-peak detection, Coiflet wavelet is used to for better HRV feature extraction with maximum R-peaks. Time domain and frequency domain features are applied to KNN classifier for arrhythmic beat classification which is simpler than other machine learning approaches. The exiting classification techniques are based on ANN, neuro fuzzy and conventional KNN classifiers. In these techniques, ECG, blood volume, HRV parameters were used. The maximum classification accuracy of 96% has been achieved using these techniques. But the experimental result of KNN based classifier with DWT gives a maximum accuracy of 97.5% on classifying normal and arrhythmic risk abnormal subjects.

References

1. Chawla M (2011) PCA and ICA processing methods for removal of artifacts and noise in electrocardiograms: a survey and comparison. Appl Soft Comput 11(2):2216–2226. https://doi.org/10.1016/j.asoc.2010.08.001
2. Elgendi M, Jonkman M, Boer FD (2009) R wave detection using Coiflets wavelets. 2009 I.E. 35th annual northeast bioengineering conference 1-2. https://doi.org/10.1109/nebc.2009.4967756
3. Gupta A, Joshi S (2008) Variable step-size LMS algorithm for fractal signals. IEEE Trans Signal Process 56 (4):1411–1420. https://doi.org/10.1109/tsp.2007.909374
4. Jayalalith S, Susan D, Kumari S, Archana B (2014) K-nearest neighbour method of analysing the ECG signal (To Find out the Different Disorders Related to Heart). J Appl Sci 14(14):1628–1632. https://doi.org/10.3923/jas.2014.1628.1632
5. Jubairahmed L, Satheeskumaran S, Venkatesan C (2017) Contourlet transform based adaptive nonlinear diffusion filtering for speckle noise removal in ultrasound images. Cluster Comput 1–10. https://doi.org/10.1007/s10586-017-1370-x
6. Kohler B, Hennig C, Orglmeister R (2002) The principles of software QRS detection. IEEE Eng Med Biol Mag 21(1):42–57. https://doi.org/10.1109/51.993193
7. Kumar M, Weippert M, Vilbrandt R, Kreuzfeld S, Stoll R (2007) Fuzzy evaluation of heart rate signals for mental stress assessment. IEEE Trans Fuzzy Syst 15(5):791–808. https://doi.org/10.1109/tfuzz.2006.889825
8. Langley P, Bowers E, Wild J, Drinnan M, Allen J, Sims A, Murray A (2003) An algorithm to distinguish ischaemic and non-ischaemic ST changes in the Holter ECG. Comput Cardiol 2003:235–238. https://doi.org/10.1109/cic.2003.1291135
9. Nouira I, Abdallah A, Kouaja I, Bedoui M (2012) Comparative study of QRS complex detection in ECG. Proc World Acad Sci Eng Technol 71:1447–1451
10. Osowski S, Hoai L, Markiewicz T (2004) Support vector machine-based expert system for reliable heartbeat recognition. IEEE Trans Biomed Eng 51(4):582–589. https://doi.org/10.1109/tbme.2004.824138
11. Parhi K, Messerschmitt D (1989) Pipeline interleaving and parallelism in recursive digital filters. I. Pipelining using scattered look-ahead and decomposition. IEEE Trans Acoust Speech Signal Process 37 (7):1099–1117. https://doi.org/10.1109/29.32286
12. Petković D, Ćojbašić Ž, Lukić S (2013) Adaptive neuro fuzzy selection of heart rate variability parameters affected by autonomic nervous system. Expert Syst Appl 40(11):4490–4495. https://doi.org/10.1016/j.eswa.2013.01.055
13. Rahman MZ, Shaik RA, Reddy DV (2012) Efficient and simplified adaptive noise cancellers for ECG sensor based remote health monitoring. IEEE Sensors J 12(3):566–573. https://doi.org/10.1109/jsen.2011.2111453

14. Rooijakkers MJ, Rabotti C, Oei SG, Mischi M (2012) Low-complexity R-peak detection for ambulatory fetal monitoring. Physiol Meas 33(7):1135–1150. https://doi.org/10.1088/0967-3334/33/7/1135
15. Satheeskumaran S, Sabrigiriraj M (2014) A new LMS based noise removal and DWT based R-peak detection in ECG signal for biotelemetry applications. Natl Acad Sci Lett 37(4):341–349
16. Satheeskumaran S, Sabrigiriraj M (2016) VLSI implementation of a new LMS based algorithm for noise removal in ECG signal. Int J Electron 103:975–984
17. Stolojescu C, Railean I, Moga S, Isar A (2010) Comparison of wavelet families with application to WiMAX traffic forecasting. 2010 12th international conference on optimization of electrical and electronic equipment 932-937. https://doi.org/10.1109/optim.2010.5510403
18. Szilagyi S, Szilagyi L (2003) Wavelet transform and neural-network-based adaptive filtering for QRS detection. Proceedings of the 22nd annual international conference of the IEEE engineering in medicine and biology society (Cat. No.00CH37143) 2:1267-1270. https://doi.org/10.1109/iembs.2000.897966
19. Ting L, Woods R, Cowan C (2005) Virtex FPGA implementation of a pipelined adaptive LMS predictor for electronic support measures receivers. IEEE Trans Very Large Scale Integr (VLSI) Syst 13(1):86–95. https://doi.org/10.1109/tvlsi.2004.840403
20. Tsipouras M, Fotiadis D, Sideris D (2005) An arrhythmia classification system based on the RR-interval signal. Artif Intell Med 33(3):237–250. https://doi.org/10.1016/j.artmed.2004.03.007
21. Venkatesan C, Karthigaikumar P, Varatharajan R (2018) FPGA implementation of modified error normalized LMS adaptive filter for ECG noise removal. Cluster Comput 1-9. https://doi.org/10.1007/s10586-017-1602-0
22. Wiggins M, Saad A, Litt B, Vachtsevanos G (2008) Evolving a Bayesian classifier for ECG-based age classification in medical applications. Appl Soft Comput 8(1):599–608. https://doi.org/10.1016/j.asoc.2007.03.009
23. Zhai J, Barreto A (2006) Stress detection in computer users based on digital signal processing of noninvasive physiological variables. 2006 international conference of the IEEE engineering in medicine and biology society 1355-1358. https://doi.org/10.1109/iembs.2006.4397662
24. Zhang J, Zhao H (2010) A novel adaptive bilinear filter based on pipelined architecture. Digit Signal Process 20(1):23–38. https://doi.org/10.1016/j.dsp.2009.06.006

Mr. C. Venkatesan received Bachelor of Engineering in Electrical and Electronics Engineering from Anna University, Chennai and Master of Engineering in Applied Electronics from the Government College of Technology, Coimbatore. He is pursuing Ph.D in Information and Communication Engineering under Anna University, Chennai. He has a teaching experience over 10 years in various academic positions. His research interests include VLSI Signal Processing, Bio-medical Signal Processing, Low power VLSI, and Artificial Intelligence based healthcare applications. He has published several research papers in refereed national and international journals. He is also member of ISTE (MISTE), member of International Association of Engineers (MIAENG) and member of International Association of Computer sciences and Information Technology (MIACSIT).

Dr. P. Karthigaikumar received his Bachelor of Engineering degree in Electrical and Electronics Engineering from the Bharathiar University, India in 1999 and his Master of Engineering degree with Distinction in Applied Electronics from Bharathiar University, India in 2002. He completed Ph.D. degree in Information and Communication Engineering under Anna University, India in 2011, focusing on FPGA and ASIC implementation of Media Security processor. He is a member of IEEE (MIEEE), senior member of Association of Computer Electronics and Electrical Engineers (ACEEE), member of International Association of Engineers (MIAENG) and member of International Association of Computer sciences and Information Technology (MIACSIT). He joined Karunya University; Coimbatore, India in 2000 and worked for 13 years. He is now Professor and Head in Electronics and Communication Engineering in Karpagam College of Engineering, Coimbatore, India. He has published more than 60 papers in journals and conferences. He received IETE K S Krishnan award for the best system oriented research paper in the year 2010. He applied for 2 Indian patents and is published in Indian Patent Journal. He is a reviewer for different reputed journals like Elsevier, Wiley, Inderscience etc., and he has been the Guest editor for few special issues in Hindawi, Elsevier, Inderscience, Springer. His research interest includes FPGA implementation of Media security algorithm and Signal Processing algorithm.

Dr. R. Varatharajan received his B.E., M.E. and Ph.D. degrees all in Electronics and Communication Engineering from Anna University and Bharath University, India. His main area of research activity is Medical Image processing, Wireless Networks and VLSI Physical Design. He has served as a reviewer for Springer, Inderscience and Elsevier journals. He has published many research articles in refereed journals. He is a member of IEEE, IACSIT, IAENG, SCIEI and ISTE wireless research group. He has been serving as Organizing Chair and Program Chair of several International conferences and in the Program Committees of several International conferences. Currently he is working as a Associate professor in the Department of Electronics and Communication Engineering at Sri Ramanujar Engineering College, Chennai, India.