Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 55 (2019) 82–94

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

An efficient detection of congestive heart failure using frequency localized filter banks for the diagnosis with ECG signals

Ankit A. Bhurane [a], Manish Sharma [b,*], Ru San-Tan [c], U. Rajendra Acharya [d,e,f]

[a] Department of Electronics and Communication, Indian Institute of Information Technology, Nagpur (IIITN), India
[b] Department of Electrical Engineering, Institute of Infrastructure, Technology, Research and Management (IITRAM), Ahmedabad, India
[c] National Heart Centre Singapore, Singapore
[d] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore 599489, Singapore
[e] Department of Biomedical Engineering, School of Science and Technology, SUSS, Singapore
[f] School of Medicine, Faculty of Health and Medical Sciences, Taylor's University, 47500 Subang Jaya, Malaysia

Received 13 October 2018; received in revised form 11 December 2018; accepted 26 December 2018
Available online 8 January 2019

Abstract

Congestive heart failure (CHF) refers to the condition in which the heart is unable to maintain the required blood flow under normal heart pressure. CHF is one of the major causes of death worldwide and is commonly caused by coronary artery disease, diabetes and high blood pressure. It typically affects the elderly population. The diagnosis of CHF is mostly based on clinical assessment of symptoms, signs, imaging findings, and invasive intracardiac pressure measurement. The electrocardiogram (ECG) is neither sensitive nor specific for diagnosis of CHF, and the analysis of the ECG signal for the possible presence of CHF is manually intensive and requires adequate skills and expertise for discerning subtle abnormalities in the electrical activity of the heart that may be associated with CHF. We hypothesized that this task of recognizing the multiparametric patterns of ECG signal aberrations that might occur in CHF could be expedited and optimized using machine learning. In this paper, we present an automated approach for the diagnosis of CHF using ECG signals. The proposed approach was tested on four different sets of normal and CHF ECG signals obtained from established public databases. The experiments were performed using short (2 second (s)) ECG segments. Five different features (fuzzy entropy, Renyi entropy, Higuchi's fractal dimension, Kraskov entropy and energy) were extracted from the wavelet decomposition of ECG segments using frequency localized filter banks. For training and classification, we employed quadratic support vector machine (QSVM). A 10-fold cross-validation technique was used for evaluation. Accuracy $\geqslant$ 99.66%, sensitivity $\geqslant$ 99.82%, and specificity $\geqslant$ 99.28%, were obtained across all four data sets used. The system can be deployed in hospitals to facilitate the diagnosis of CHF. The proposed system can reduce the time requirement and error rate associated with manual reading of large ECG signals.

Keywords: Electrocardiogram signals; Congestive heart failure; Wavelets

1. Introduction

According to the National Vital Statistics Reports, heart diseases are the leading cause of mortality (Kenneth et al., 2011). Congestive heart failure (CHF) is a syndrome where

* Corresponding author.
E-mail addresses: manishsharma@iitram.ac.in (M. Sharma), tan.ru.san@singhealth.com.sg (R. San-Tan), aru@np.edu.sg (U.R. Acharya).

https://doi.org/10.1016/j.cogsys.2018.12.017
1389-0417/

the heart is unable to pump out sufficient blood flow output at without a rise in chamber pressure (Figueroa & Peters, 2006). The poor cardiac output results in poor blood flow and oxygen carriage to other body parts, which manifests as tiredness, difficulty in performing day-to-day activities. The increased intracardiac pressure causes accumulation of fluids in the lungs and peripheries, which lead to breathlessness and edema (swelling), respectively. The CHF adversely impacts the quality of life. As the illness deteriorates, more health care including hospitalizations are required, which can significantly impact on the finance of the patient family (Levenson, McCarthy, Lynn, Davis, & Phillips, 2000).

The diagnosis of CHF is clinical generally based on the assessment of symptoms, signs, imaging findings, and invasive intracardiac pressure measurement. The simplest non-invasive approach are electrocardiogram (ECG) and photoplethysmography (PPG) (Moraes et al., 2018). The heart activity is analyzed by the cardiologist who visually examines the electrical patterns of the ECG, looks for the subtle abnormalities in the signal may be associated with CHF. However, the ECG is neither sensitive nor specific for diagnosis of CHF. The well-known disadvantages of such visual examination are excess time consumption, prone to manual errors, and fatigue. We hypothesized that the task of recognizing the multiparametric patterns of ECG signal aberrations that might occur in CHF could be accelerated and optimized using machine learning techniques. An automated tool to examine CHF would not only assist the cardiologists in the diagnosis by minimizing the manual input, but would also help in cross validation of the data under consideration. The main contribution of this paper is the automated classification of normal and CHF ECG signals using nonlinear features (fuzzy entropy, Renyi entropy, Higuchi's fractal dimension, Kraskov entropy and energy) extracted from wavelet coefficients obtained from an optimal class of frequency localized filter banks.

2. Related works

A review of various techniques employed in the characterization of heart failure till 2017 is discussed in the work by Tripoliti, Papadopoulos, Karanasiou, Naka, and Fotiadis (2017). The computer aided diagnosis approaches for CHF detection can be broadly classified on the basis of the domain of feature extraction. Most of the contribution in the literature is based on the time-domain, frequency domain, and wavelet-domain techniques supported by machine learning algorithms. Approaches based on entropy of binary occupancies of sequential spectrum were studied by Kamath (2015) and Kamath (2012). Time-domain statistical features like mean RR, mean IHR, NN50 and pNN50 combined with back-propagation neural network have proven to be promising for CHF detection (Khaled, Owis, & Mohamed, 2006). One of the frequency domain approaches evaluate the usage of low, high and total power spectrum was presented by Pecchia, Melillo, Sansone, and Bracale (2011). A method based on both time and frequency was studied by Orhan (2013). The authors have analysed the efficiency of equal frequency in amplitude and equal width in time discretization followed by linear regression. Studies based on higher order spectra and usefulness of genetic algorithm in feature selection are also available in the literature (Yu & Lee, 2012). The usage of optimum-path forest (OPF) classifier has shown better performance than the multilayer artificial neural network and SVM classifiers for ECG analysis (Albuquerque et al., 2018; Luz, Nunes, Albuquerque, Papa, & Menotti, 2013). One of the recent works on wavelet-based approach includes the usage of dual-tree complex-wavelet transform for CHF classification (Sudarshan et al., 2017). Other recent advances include the usage of non-linear analysis (Marques et al., 2018), cloud-based systems (Hussein et al., 2018), augmented reality and virtual reality based diagnosis of heart diseases (Hemanth, Köse, Deperlioglu, & Albuquerque, 2018). More recently, an approach using a deep convolutional neural network (DCNN) to classify normal and CHF signals was presented in the work by Acharya et al. (2018). In this work, the authors employed an 11-layer DCNN to detect the CHF using raw ECG signals without any preprocessing. Although for a large dataset, the DCNN would work well, the initial setup of DCNN is more difficult as compared to the SVMs.

The ECG signals for CHF have lower amplitude as compared to the normal ECG signals. Also, due to abnormal heart rate and rhythm, the structure of ECG signal is different. Wavelets and filter banks have been widely used for the analysis of physiological signals including ECG (Addison, 2005; Sharma, Tan, & Acharya, 2018; Sharma, Achuth, Deb, Puthankattil, & Acharya, 2018; Sharma, Bhurane, & Acharya, 2018; Sharma, Goyal, Achuth, & Acharya, 2018; Sharma, Sharma, Pachori, & Acharya, 2018; Sharma, Pachori, & Acharya, 2017; Sharma & Pachori, 2017). The shape of ECG signals can be well represented using shifted and scaled versions of wavelets. Recently, a multiresolution approach using wavelet packet transform for classification of ECG signals was presented by Fujita et al. (2017). In this approach authors have used the segmented ECG signals followed by wavelet packet decomposition, feature extraction, and K-nearest neighbor classifier.

Wavelet have found there usage in several application including secured sharing of medical records (Faeq Hussein et al., 2018). Unlike classical wavelets, time-frequency localized filters (Bhati, Sharma, Pachori, & Gadre, 2017; Sharma, Dhere, Pachori, & Gadre, 2017; Sharma, Achuth, Pachori, & Gadre, 2017; Sharma, Deb, & Acharya, 2018) provide better representation of the inherent dynamics of normal and CHF ECG signals. The localized wavelet filter banks offer several advantages, including compact frequency spread, smaller magnitude ripples and sharper roll-off (Sharma, Dhere, Pachori, & Acharya, 2017; Sharma, Gadre, & Porwal, 2015; Sharma, Kolte, Patwardhan, & Gadre, 2010; Sharma, Sharma,

Pachori, & Gadre, 2019). The salient features of time-frequency optimized wavelet filter banks motivated us to use them for the automated classification of CHF signals. The novelty of this paper is the usage of nonlinear features (fuzzy entropy, Renyi entropy, Higuchi's fractal dimension, Kraskov entropy and energy) extracted from wavelet coefficients of frequency-localized filter bank for diagnosis of CHF. It is worth noting that we have not used standard and traditional Daubechies (dB) wavelet FBs as they cannot be considered as optimal FBs (Sharma, Vanmali, & Gadre, 2013; Shah, Sharma, Deb, & Pachori, 2019). The extracted features are fed to support vector machine (SVM) with quadratic kernel for training and classification (Vapnik, 1995). The details of the proposed approach is presented in the subsequent sections.

2.1. Notations and abbreviations

The notations $h_0(n)$, $h_1(n)$, $g_0(n)$, $g_1(n)$ denote the impulse response sequences of analysis and synthesis low-pass and high-pass filters of underlying wavelet filter bank, similarly $p(n)$ denotes impulse response of the product filter. The notations $X(\omega)$ and $X(z)$ represent frequency response and z transform of the sequence $x(n)$. $\mathbf{x}(\mathbf{n})$ and $\mathbf{w}(\mathbf{n})$ represent input time sequence and wavelet sub-bands respectively. Symbol $\langle 2\pi \rangle$ denotes the time period taken over interval of 2π seconds. The notation $\delta(n)$ represents unit impulse sequence. Symbols $(\downarrow 2)$ and $(\uparrow 2)$ represent down-sampling and up-sampling by the factor of two, respectively. Table 1 shows the list of abbreviations used in the work along with their definitions.

3. Materials and methods

3.1. Database

In our experiments, we have used the ECG samples from subjects with moderate or severe symptomatic CHF. The ECG signals from three different public databases namely Physionet (2016) and Iyengar, Peng, Morin, Goldberger, and Lipsitz (1996), Normal Sinus Rhythm Database (NSRDB) obtained from arrhythmia laboratory at Boston's Beth Israel Hospital, and Boston's Beth Israel Hospital (BIDMC) CHF Database (Goldberger et al., 2000) were used. Fantasia and NSRDB databases consist of normal ECG recordings from 40 and 18 subjects respectively, whereas the BIDMC consists of CHF signals from 18 subjects. The BIDMC-CHF data was recorded using ambulatory ECG recorders with a recording frequencies range from 0.1 Hz to 40 Hz. Each ECG signal was recorded at a sampling rate of 250 samples per second with a voltage resolution of 12-bits in a range of 10 millivolts magnitude (BIDMC, 2016; Baim et al., 1986). During the recording, the limb leads were not used. Instead, the sensing electrodes were placed at appropriate positions on the trunk (Goldberger et al., 2000; Malmivuo & Plonsey, 1995). These ECG recordings obtained from public databases are segmented into 2 s duration. In order to evaluate the performance on even and uneven datasets, the normal EEG signals from two database and a common dataset of EEG signals for CHF are arranged into four sets (two balanced and two unbalanced) as given in Table 2. A balanced dataset consists of equal number of samples (30,000 each) for normal and CHF EEG signals, whereas an unbalanced dataset comprises uneven samples for normal and CHF EEG signals. Samples from each database are shown in Fig. 1.

Table 1
List of abbreviations used in the work.

Abbreviation	Definition
CHF	Congestive heart failure (CHF)
ECG	Electrocardiogram
SVM	Support vector machine
QSVM	Quadratic support vector machine
PPG	Photoplethysmography
OPF	Optimum-path forest
DCNN	Deep convolutional neural network
NSRDB	Normal Sinus Rhythm Database
BIDMC	Boston's Beth Israel Hospital
DWT	Discrete wavelet transform
FLWFB	Frequency-localized wavelet filter banks
FE	Fuzzy entropy
RE	Renyi entropy
HFD	Higuchi's Fractal dimension
KE	Kraskov entropy
E	Energy
ROC	Receiver operating characteristics

3.2. Proposed approach

The proposed approach used in this work is shown in Fig. 2. The ECG signals are initially decomposed using discrete wavelet transform (DWT). The resultant wavelet coefficients are fed to the feature extraction engine. The features acts as input data to the classifier. The details of each step are as follows.

3.2.1. Frequency-localized wavelet filter banks (FLWFB)

Wavelets have been widely used to analyze the biomedical signals. Traditional Daubechies wavelet FBs have the limitation of not having optimized time and frequency spread (Sharma, Bhati, Pillai, Pachori, & Gadre, 2016). The time-frequency localized filter banks have been proven to perform better in various signal processing applications including denoising, compression, and classification. We employed these frequency-localized filter banks for wavelet decomposition. This section gives an overview of FLWFB.

Table 2
Set of database used in this work.

Set	Fantasia	NSRDB	BIDMC	Total
A	0	70,308	30,000	100,308
B	80,000	0	30,000	110,000
C	0	30,000	30,000	60,000
D	30,000	0	30,000	60,000

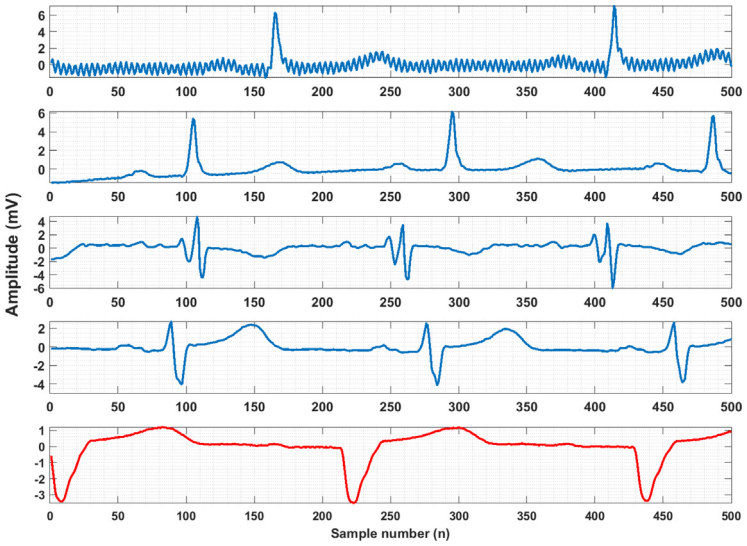

Fig. 1. Sample sequences from the database used in this work. From the top: Fantasia, Fantasia (balanced), NSRDB, NSRDB (balanced), and CHF.

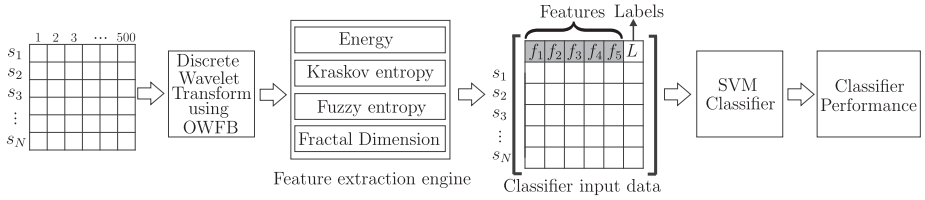

Fig. 2. Block diagram of the proposed approach.

Consider the classical two-band filter-bank structure shown in Fig. 3. Let $h_0[n]$ be the impulse response of analysis low-pass filter with $0 \leqslant n \leqslant N-1$. Now, the aim is to minimize the mean squared bandwidth of $h_0[n]$ by retaining the orthonormality and regularity constraint. Thus, the problem can be formulated as,

$$\min_{h_0[n]} B_\omega = \frac{1}{2\pi E} \int_{<2\pi>} \omega^2 |H_0(\omega)|^2 d\omega \tag{1}$$

such that,

$$\sum_{n=0}^{N} h_0[n] h_0[n-2k] = \delta[n] \tag{2}$$

$$\sum_{n=0}^{N} (-1)^k k^l h_0[k] = 0; \tag{3}$$

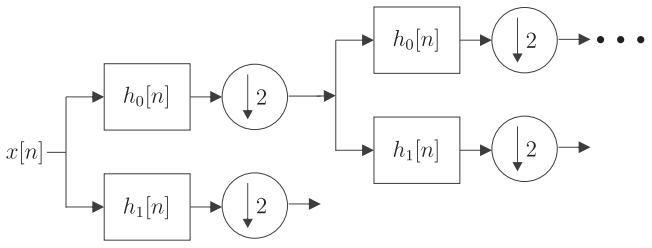

Fig. 3. Structure of two-band filter bank.

where $k = 0, 1, \ldots, \frac{N}{2} - 1$ and $l = 0, 1, \ldots, M-1$. The problem can be expressed in terms of the impulse response of the product filter $P(z) = H_0(z) H_0(z^{-}1)$ as,

$$min_{p[n]} B_\omega = \frac{1}{2\pi E} \int_0^\pi \omega^2 P(\omega) d\omega \tag{4}$$

such that,

$$p[2k] = \delta[n] \tag{5}$$

$$p[0] + 2\sum_{n=0}^{N} (-1)^k p[k] = 0 \tag{6}$$

$$\sum_{n=0}^{N} (-1)^k k^{2l} p[k] = 0; \tag{7}$$

where $k = 0, 1, \ldots, \frac{N}{2} - 1$ and $l = 0, 1, \ldots, M-1$. and $P(\omega) \geqslant 0$.

Equation for B_ω gives the coefficients of analysis low-pass filter $h_0[n]$ from which we can obtain other filters $h_1[n]$, $g_0[n]$, and $g_1[n]$.

3.2.2. Stages of proposed approach

Our approach comprises the following three stages.

1. Stage-1: Wavelet decomposition FLWFB: The raw ECG signals are subjected to a multiresolution decomposition using the frequency localized discrete wavelet transform (DWT). In this process the ECG signals are

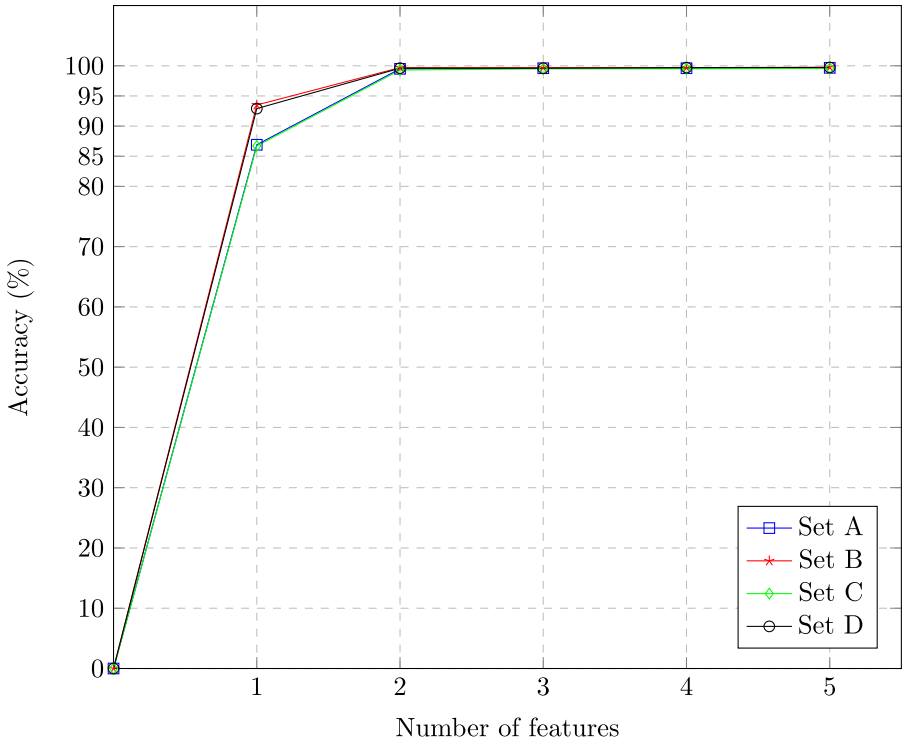

Fig. 4. Number of features versus accuracy for all datasets.

filtered using multiple stages of two-channel orthogonal filter banks to obtain seven the sub-bands of each ECG segment.In this work we choose 6-level wavelet decomposition using $N = 13$ (length-14) filter with coefficients given in Table 3.

2. Stage-2: Feature extraction engine: In this stage, the feature extraction engine extracts five different features as listed below.

(a) **Fuzzy entropy (FE):** The sample entropy for classification of CHF signals was given by Zhao et al. (2015). For an L length data $\mathbf{x}$ we need to form a vector new sequence as,

$$X_i^m = \{x[i], x[i+1], \ldots, x[i+m-1]\} \tag{8}$$

where X_i^m represents m consecutive values of $\mathbf{x}$. Let $A_i^m[r] = \frac{1}{(L-m)} X_j^{m+1}$ and $B_i^m[r] = \frac{1}{(L-m)} X_j^m$ for all $1 \leqslant j \leqslant L-m$. Then, the sample fuzzy entropy is defined as,

$$F = -\ln\left(\frac{\sum_{i=1}^{L-m} A_i^m[r]}{\sum_{i=1}^{L-m} B_i^m[r]}\right) \tag{9}$$

where the parameter m is the embedded dimension that is empirically set as $m = 3$.

(b) **Renyi entropy (RE):** The RE is defined as (Renyi, 1961),

$$R = \frac{A}{1-A} \log\left(\|x[n]\|_A\right) \tag{10}$$

where $A > 0$ and $A \neq 1$. The RE features are extracted from each subband of wavelet decomposition. The features obtained from individual subbands are concatenated to form a single vector.

(c) **Higuchi's Fractal dimension (HFD):** The fractal dimension is a measure of signal complexity. HFD was given by Higuchi (1988) and has been widely used for signal analysis due to its efficiency and speed of calculation (Kesic & Spasic, 2016). The fractal dimension is calculated as follows. For an L-length sequence $\mathbf{x}$, a new sequence is formed as,

$$X_k^m = \{x[m]; x[m+k]; x[m+2k], \ldots, x[m+(n-m)/k]\}. \tag{11}$$

Table 3
Length-14 filter coefficients used in this work.

n	$h_0[n]$
0	0.2358
1	0.6485
2	0.6319
3	0.1275
4	−0.2497
5	−0.1059
6	0.1346
7	0.0656
8	−0.0824
9	−0.0363
10	0.0603
11	−0.0008
12	−0.0234
13	0.0085

Table 4
Rank values for each feature and subbands.

Feature #	Subband #	Rank # for Dataset			
		A	B	C	D
1	1	21	21	21	21
	2	5	5	5	5
	3	3	26	3	26
	4	33	6	33	25
	5	6	25	34	34
	6	34	34	31	6
	7	31	13	28	2
2	1	28	35	13	13
	2	13	2	6	28
	3	12	7	12	30
	4	2	30	2	33
	5	30	28	30	35
	6	4	18	4	18
	7	10	33	10	14
3	1	32	14	32	7
	2	8	8	26	8
	3	26	3	8	3
	4	9	31	9	31
	5	25	9	25	9
	6	29	4	29	11
	7	20	11	20	4
4	1	1	27	1	27
	2	11	10	11	10
	3	22	20	22	20
	4	7	17	7	29
	5	23	29	35	17
	6	35	24	23	24
	7	27	1	18	1
5	1	18	12	27	12
	2	19	23	19	23
	3	14	22	14	22
	4	17	32	17	32
	5	24	19	24	19
	6	15	15	15	15
	7	16	16	16	16

Table 5
Training and prediction time for four datasets.

Dataset	Training time (s)	Prediction time (obs/s)
A	669.26	65,000
B	668.36	79,000
C	298.63	58,000
D	240.47	83,000

where X_k^m represents shifted elements of **x**. The length of X_k^m is then defined as,

$$\mathbf{l}=\frac{1}{k}\left[\left(\sum_{i=1}^{\lfloor (n-m)/k\rfloor}|x[m+ik]-x[m+(i-1)k]|\right)\frac{L-1}{\lfloor (n-m)/k\rfloor k}\right]. \tag{12}$$

The average length vector **l** is used to find the HFD by finding the polynomial that fits the curve of ln(**l**). We treat the wavelet coefficients as input sequence and calculate the HFD to obtain the HFD feature vector.

(d) **Kraskov entropy (KE):** It is a measure of interdependencies in the physiological time series data and is beneficial in reducing the systematic errors. KE is defined as (Kraskov, Stögbauer, & Grassberger, 2004),

$$K=-\psi(k)+\psi(L)+\log V+\frac{d}{L}\sum\log\epsilon \tag{13}$$

where V gives the volume of number of elements, ϵ is the function of distance between k neighbor elements.

(e) **Energy (E):** The energy features are extracted from the L-length wavelet coefficients **w** obtained at Stage-1. The energy feature is defined as,

$$E=\sum(\mathbf{w}(\mathbf{n}))^2 \tag{14}$$

The expression gives the amount of energy of wavelet time-series $w(n)$.

Table 7
Classification performance obtained for four datasets using SVM classifier.

Measure	Value for class			
	A	B	C	D
Accuracy	0.9966	0.9977	0.9958	0.9972
Sensitivity	0.9969	0.9986	0.9955	0.9980
Specificity	0.9959	0.9952	0.9962	0.9964
Precision	0.9983	0.9982	0.9962	0.9964
False positive rate	0.0041	0.0048	0.0038	0.0036
False negative rate	0.0031	0.0014	0.0045	0.0020
F1 score	0.9976	0.9984	0.9958	0.9972
Matthews correlation coefficient	0.992	0.9941	0.9917	0.9944

Table 6
Confusion matrices obtained for A, B, C and D datasets.

		Set							
		A		B		C		D	
		Predicted label		Predicted label		Predicted label		Predicted label	
		0	1	0	1	0	1	0	1
True Label	0	70186	122	79855	145	29885	115	29893	107
	1	215	29785	113	29887	135	29865	61	29939

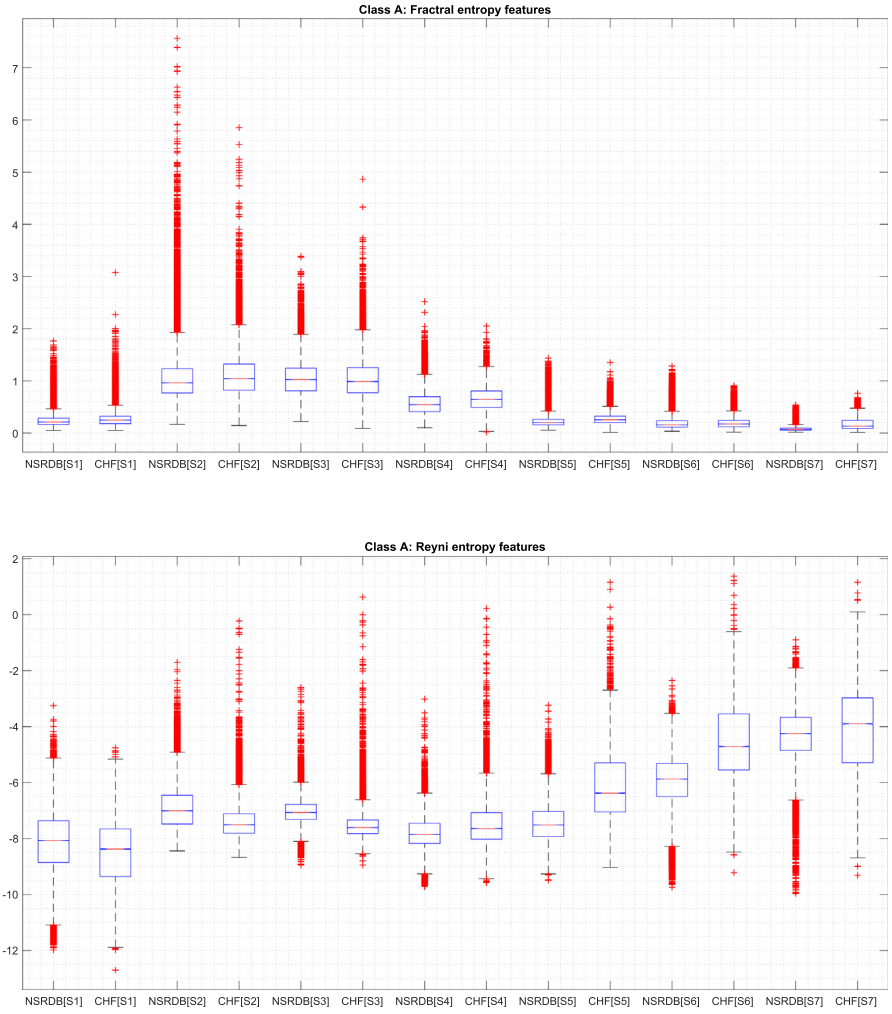

Fig. 5. Box plots for FE (top) and RE (bottom) features of Class-A database.

Finally all the five features are concatenated to form a feature vector. We do this for all the signals in the database. For feature ranking, we used t-test algorithm (Devore & Devore, 2012; Student, 1908). The rank values of all five features extracted from all seven subbands are listed in Table 4. The combined feature vector is fed to the SVM classifier.

3. Stage-3: Classification: The feature vectors obtained from Stage-2 are fed to support vector machine (SVM) classifier. SVM was initially proposed by Vapnik (1995) as a statistical learning tool. SVM aims to identify a hyperplane that can classify a data with maximum separation. Initially the data is transformed to a higher dimensional space using a d-degree polynomial kernel function given by,

$$K(x,y) = \left(\sum_{i=1}^{n} x_i y_i + c\right)^d \tag{15}$$

where (x,y) are feature vectors and $c \geqslant 0$. We used quadratic polynomial kernel SVM i.e. when $d = 2$. Post transformation, a hyperplane that optimally classifies the data into two classes is found using,

$$H = \sum_{i=1}^{n} w_i K(x, s_i) + b$$

where K is the kernel as defined in Eq. (15), w_i are the weights, s_i are the support vectors, and b is a bias. Depending on the sign of H, the input data is classified into two classes.

4. Results

We implemented the proposed approach on personal computer equipped with i7 6th generation processor and 16 Gb RAM. The average training time required to train the system is given in Table 5. We evaluated the data using 10-fold cross validation (Kohavi, 1995). The values of confusion matrices and performance parameters namely, overall accuracy, sensitivity and specificity are given in Tables 6 and 7. The error rates are as low as 0.45%, 0.33%, 0.44%

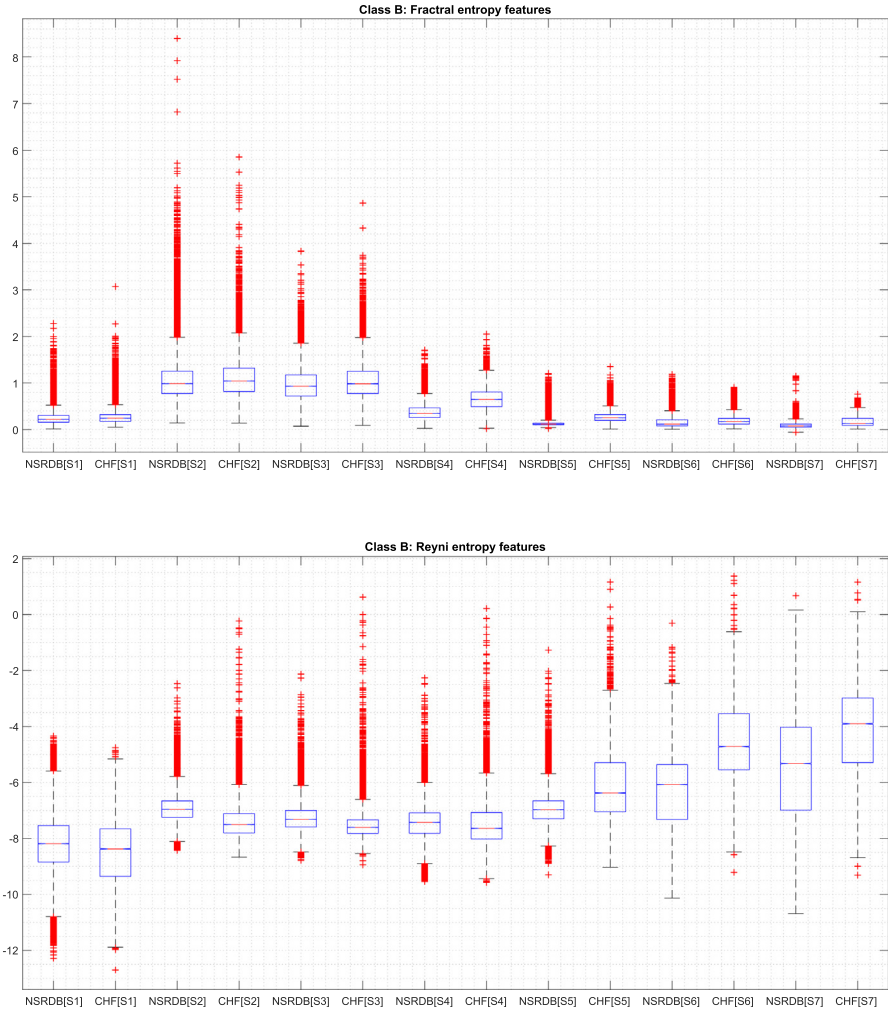

Fig. 6. Box plots for FE (top) and RE (bottom) features of Class-B database.

and 0.15% for sets A, B, C, and D respectively. The receiver operating characteristics (ROC) curves attain close to perfect classification point of (0,1) for all four datasets. We achieved highest accuracy, lowest training time and fastest prediction for Set D compared to A, B, and C.

5. Discussion

The frequency-localized filters are proven to perform better compared to the conventional set of wavelet filters (Sharma et al., 2016). Further, the features such as FE and RE act as a powerful tools to distinguish between normal and CHF ECG signals. We employed a set of these nonlinear features in wavelet domain for the automated diagnosis of CHF ECG signals. The statistics of the top two extracted features can be observed from the box plots shown in Figs. 5–8. For FE features, the box plots show that the variation in the feature values is comparatively higher in the middle subbands due to the variation in the ECG signals in different bands. This can be observed from the lower and upper whisker of the box plots. Further, the notches of normal and CHF features are close to each other. On the other hand, for RE features, the gap between the notches of normal and CHF features is comparatively higher indicating the statistical difference in the two data features. Also, this gap is more in the higher level subbands indicating that the RE feature difference is well captured in the higher level subbands. Thus RE features contribute significantly as the second best feature. It can be noted that the trend is consistent across all the four sets of data.

The performance comparison of the results obtained using the proposed approach with the state-of-the-art techniques are presented in Table 9. The values clearly show that our approach surpasses the performances of the existing techniques. Moreover, we are able to achieve the speed training speed of 669.26 s, 668.36 s, 298.63 s, and 240.47 s for sets A, B, C, and D respectively which is around 1/4th the time taken by the approach presented in Acharya et al. (2018).

The proposed approach has following salient features.

- The abnormal CHF signals can be differentiated from the normal ECG signals using the proposed set of wavelet-based non linear features.

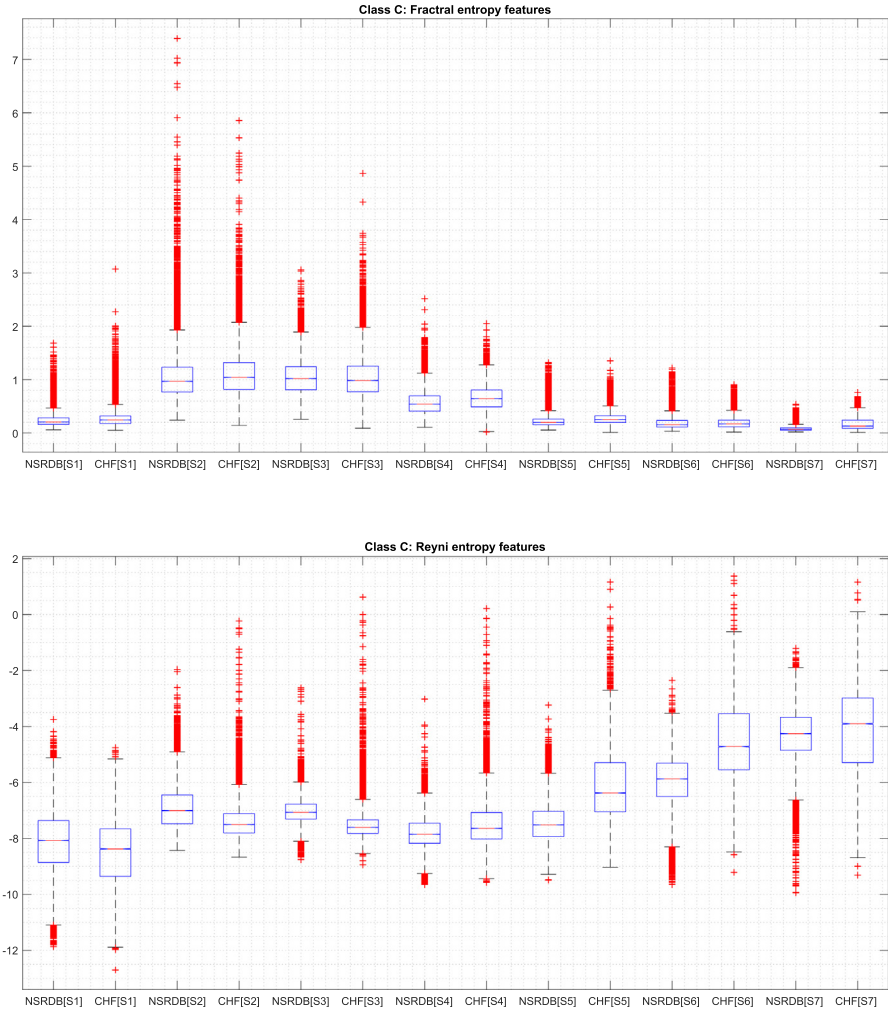

Fig. 7. Box plots for FE (top) and RE (bottom) features of Class-C database.

- Only two features (FE and RE) are found to be contributing significantly to achieve accuracies $> 99\%$. Incremental increase in the accuracy is contributed by the other three features (HFD, KE, and E) across all four datasets. This can be observed from the plots of accuracy versus number of features shown in Fig. 4. We extracted five different features from all the wavelet subbands. The sequence of respective progressive accuracies obtained for all the four datasets is tabulated in Table 8.
- The filters fit into the classical quadrature-mirror structure of two-channel orthogonal wavelet filter banks that are optimal unlike the traditional standard dB filter banks.
- A single set of filters once generated is applicable for all the four datasets.
- The process does not require any type of pre-processing prior, thus the ECG signals can be directly subjected to feature extraction.
- The proposed approach does not consider any data dependent parameters.

These features are well suited for the diagnosis of CHF. Although the approach is efficient, the proposed method has the following limitations.

- The optimal modeling of the SVM is a time consuming process. The training and prediction time is high for single features. However, it reduces with the inclusion of more features.
- One needs to find optimal lengths and ZMs of the filter banks for the given application as the selection of optimal filter banks is application dependent.
- The difficulty in selecting an appropriate kernal function in the SVM.
- The memory requirement increases with the increase in the training samples.

In future, we intend to use various deep learning methods to classify without extracting the features using some advanced techniques and data-sets (Faust, Hagiwara, Hong, Lih, & Acharya, 2018; Oh, Ng, Tan, & Acharya, 2018; Yildirim, Plawiak, Tan, & Acharya, 2018).

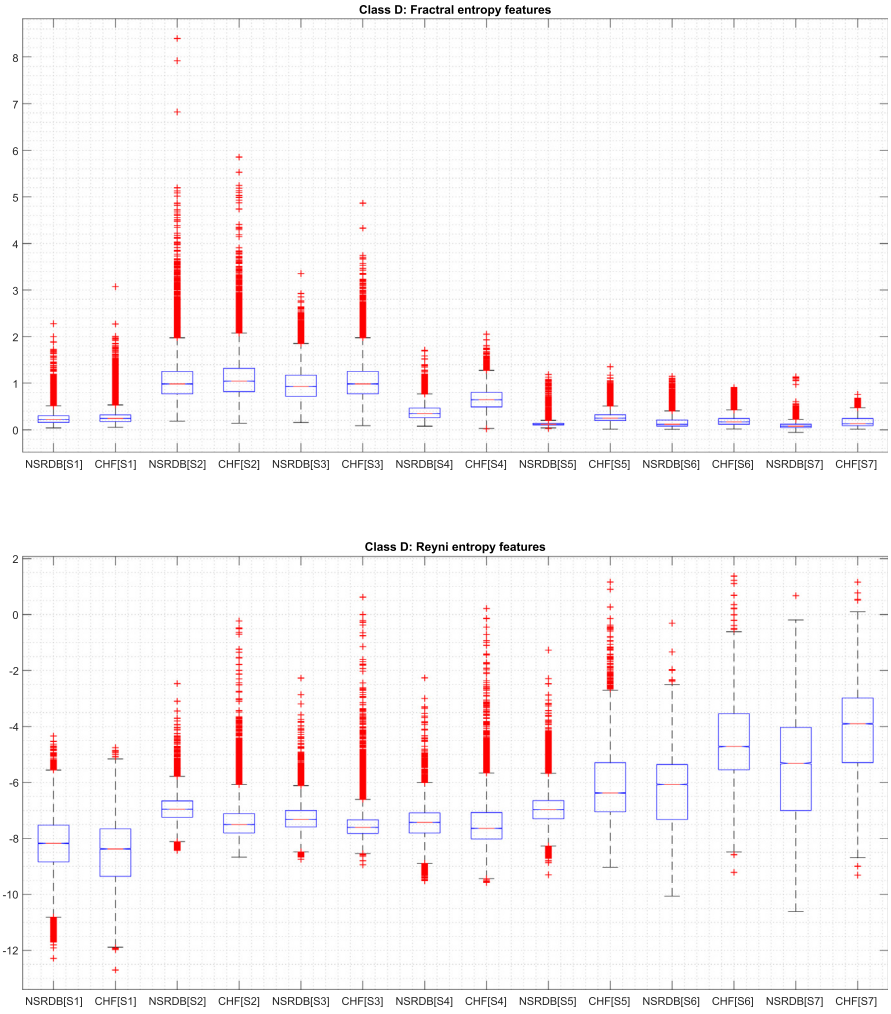

Fig. 8. Box plots for FE (top) and RE (bottom) features of Class-D database.

Table 8
Progressive accuracies for all the datasets.

# Features	Class			
	A	B	C	D
	NSRDB and BIDMC	Fantasia and BIDMC	NSRDB and BIDMC (Balanced)	Fantasia and BIDMC (Balanced)
1 (FE)	86.90%	93.50%	86.70%	92.90%
2 (FE + RE)	99.50%	99.70%	99.30%	99.60%
3 (FE + RE + HFD)	99.60%	99.70%	99.50%	99.60%
4 (FE + RE + HFD + KE)	99.60%	99.70%	99.50%	99.70%
5 (FE + RE + HFD + KE + E)	99.66%	99.77%	99.58%	99.72%

6. Conclusions

We proposed a wavelet-based approach for the automated ECG diagnosis of CHF. Raw ECG signals are subjected to wavelet decomposition using frequency-localized orthogonal filter banks. The resultant wavelet coefficients are given as an input to the feature extraction engine that extracts different nonlinear features. The extracted features are classified using SVM with quadratic kernel function. We tested our approach on four different datasets and obtained an accuracy of 99.85%. The proposed method compares favorably with other state-of-the-art techniques, and can be used for automated and efficient diagnosis of CHF. Recently, Sharma, Tan, et al. (2018), Sharma, Agarwal, and Acharya (2018), Sharma and Acharya (2018) have developed novel filter banks for the analysis

Table 9
Summary of related work on automated detection of CHF using ECG signals.

Author	Validation	Duration	# ECG Data	Approach	Performance
Sudarshan et al. (2017)	10-fold cross-validation	2 s	All four sets A, B, C, and D	Dual tree complex wavelet transform with KNN classifier	Set A Acc: 98.42% Set B Acc: 99.87% Set C Acc: 97.94% Set D Acc: 99.86%
Acharya et al. (2018)	10-fold cross-validation	2 s	All four sets A, B, C, and D	11-layer deep convolutional neural network	Set A Acc: 95.98% Set B Acc: 98.97% Set C Acc: 94.40% Set D Acc: 98.33%
Masetic and Subasi (2016)	10-fold cross-validation	2.77s	Set A and B	Dual tree complex wavelet transform with KNN classifier	Set A Acc: 100% Set B Acc: 100%
This work	10-fold cross-validation	2 s	All four sets A, B, C, and D	FLWFB and Quadratic SVM	**Set A Acc: 99.77 % Set B Acc: 99.77% Set C Acc: 99.66 % Set D Acc: 99.85% Sen: 99.87 % Spec: 99.80 %**

of physiological signals, these newly designed filter bank may be used for the detection of congestive heart failure in future. Our development model can also be used for the early diagnosis of congestive heart failure, coronary artery disease, and myocardial infarction.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cogsys.2018.12.017.

References

Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H., Adam, M., & Tan, R. S. (2018). Deep convolutional neural network for the automated diagnosis of congestive heart failure using ecg signals. *Applied Intelligence*. https://doi.org/10.1007/s10489-018-1179-1.

Addison, P. S. (2005). Wavelet transforms and the ecg: A review. *Physiological Measurement, 26*(5), R155<http://stacks.iop.org/0967-3334/26/i=5/a=R01>.

Albuquerque, V., Nunes, T., Pereira, D., Luz, E., Menotti, D., Papa, J., & Tavares, J. (2018). Robust automated cardiac arrhythmia detection in ecg beat signals. *Neural Computing and Applications, 29*, 679–693. https://doi.org/10.1007/s00521-016-2472-8.

Baim, D. S., Colucci, W. S., Monrad, E. S., Smith, H. S., Wright, R. F., Lanoue, A., ... Braunwald, E. (1986). Survival of patients with severe congestive heart failure treated with oral milrinone. *Journal of the American College of Cardiology, 7*(3), 661–670. https://doi.org/10.1016/S0735-1097(86)80478-8<http://www.sciencedirect.com/science/article/pii/S0735109786804788>.

Bhati, D., Sharma, M., Pachori, R. B., & Gadre, V. M. (2017). Time-frequency localized three-band biorthogonal wavelet filter bank using semidefinite relaxation and nonlinear least squares with epileptic seizure EEG signal classification. *Digital Signal Processing, 62*, 259–273.

BIDMC (2016). The BIDMC congestive heart failure database [Online; accessed 15-Sept-2018]. <https://www.physionet.org/physiobank/database/chfdb/>.

Devore, J., & Devore, R.P. (2012). Statistics: The exploration and analysis of data.

Faeq Hussein, A., N, A., Ramirez-Gonzalez, G., Abdulhay, E., Manuel, J., Tavares, R. S., & Albuquerque, V. (2018). A medical records managing and securing blockchain based system supported by a genetic algorithm and discrete wavelet transform. *Cognitive Systems Research, 52*, 1–11. https://doi.org/10.1016/j.cogsys.2018.05.004.

Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine, 161*, 1–13. https://doi.org/10.1016/j.cmpb.2018.04.005<http://www.sciencedirect.com/science/article/pii/S0169260718301226>.

Figueroa, M. S., & Peters, J. I. (2006). Congestive heart failure: Diagnosis, pathophysiology, therapy, and implications for respiratory care. *Respiratory Care, 51*(4), 403–412. Available from: arXiv: http://rc.rcjournal.com/content/51/4/403.full.pdf<http://rc.rcjournal.com/content/51/4/403>.

Fujita, H., Sudarshan, V. K., Adam, M., Oh, S. L., Tan, J. H., Hagiwara, Y., ... Acharya, U. R. (2017). Characterization of cardiovascular diseases using wavelet packet decomposition and nonlinear measures

of electrocardiogram signal. In S. Benferhat, K. Tabia, & M. Ali (Eds.), *Advances in artificial intelligence: From theory to practice* (pp. 259–266). Cham: Springer International Publishing.

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation, 101*(23), e215–e220. https://doi.org/10.1161/01.CIR.101.23.e215, Circulation Electronic Pages: <http://circ.ahajournals.org/content/101/23/e215.full> PMID:1085218.

Hemanth, J. D., Köse, U., Deperlioglu, O., & Albuquerque, V. (2018). An augmented reality-supported mobile application for diagnosis of heart diseases. *The Journal of Supercomputing*. https://doi.org/10.1007/s11227-018-2483-6.

Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena, 31*(2), 277–283. https://doi.org/10.1016/0167-2789(88)90081-4<http://www.sciencedirect.com/science/article/pii/0167278988900814>.

Hussein, A. F., Kumar, A., Burbano-Fernandez, M., Ramirez-Gonzalez, G., Abdulhay, E., & de Albuquerque, V. H. C. (2018). An automated remote cloud-based heart rate variability monitoring system. *IEEE Access*, 1. https://doi.org/10.1109/ACCESS.2018.2831209.

Iyengar, N., Peng, C. K., Morin, R., Goldberger, A. L., & Lipsitz, L. A. (1996). Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 271*(4), R1078–R1084. https://doi.org/10.1152/ajpregu.1996.271.4.R1078, pMID: 8898003. Available from: arXiv: https://doi.org/10.1152/ajpregu.1996.271.4.R1078.

Kamath, C. (2012). A new approach to detect congestive heart failure using sequential spectrum of electrocardiogram signals. *Medical Engineering and Physics, 34*(10), 1503–1509. https://doi.org/10.1016/j.medengphy.2012.03.001<http://www.sciencedirect.com/science/article/pii/S1350453312000495>.

Kamath, C. (2015). Entropy measures of irregularity and complexity for surface electrocardiogram time series in patients with congestive heart failure. *Journal of Advances in Computer Research, 6*(4), 1–11<http://jacr.iausari.ac.ir/article_14757.html>.

Kenneth, M., Kochanek, D., Jiaquan Xu, M., Sherry, B., Murphy, L., Arialdi, M., ... Kung, H.-C. (2011). Deaths: Final data for 2009. *National Vital Statistics Reports, 60*(3), 1–117<https://www.cdc.gov/nchs/data/nvsr/nvsr60/nvsr60_03.pdf>.

Kesic, S., & Spasic, S. Z. (2016). Application of higuchi's fractal dimension from basic to clinical neurophysiology: A review. *Computer Methods and Programs in Biomedicine, 133*, 55–70. https://doi.org/10.1016/j.cmpb.2016.05.014<http://www.sciencedirect.com/science/article/pii/S0169260715302923>.

Khaled, A., Owis, M., & Mohamed, A. S. A. (2006). Employing time-domain methods and poincaré plot of heart rate variability signals to detect congestive heart failure 6.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th international joint conference on artificial intelligence - Volume 2, IJCAI'95* (pp. 1137–1143). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.<http://dl.acm.org/citation.cfm?id=1643031.1643047>.

Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E, 69*(6), 066138.

Levenson, J. W., McCarthy, E. P., Lynn, J., Davis, R. B., & Phillips, R. (2000). The last six months of life for patients with congestive heart failure 48, S101–S109. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1532-5415.2000.tb03119.x>.

Luz, E., Nunes, T., Albuquerque, V., Papa, J., & Menotti, D. (2013). Ecg arrhythmia classification based on optimum-path forest. *Expert Systems with Applications, 40*, 3561–3573. https://doi.org/10.1016/j.eswa.2012.12.063.

Malmivuo, J., & Plonsey, R. (1995). Bioelectromagnetism: Principles and applications of bioelectric and biomagnetic fields.

Marques, J., Cortez, P., Madeiro, J., Albuquerque, V., Fong, S. C., & Schlindwein, F. (2018). Nonlinear characterization and complexity analysis of cardiotocographic examinations using entropy measures. *The Journal of Supercomputing*. https://doi.org/10.1007/s11227-018-2570-8.

Masetic, Z., & Subasi, A. (2016). Congestive heart failure detection using random forest classifier. *Computer Methods and Programs in Biomedicine, 130*(C), 54–64. https://doi.org/10.1016/j.cmpb.2016.03.020.

Moraes, J. L., Rocha, M. X., Vasconcelos, G. G., Vasconcelos, E., Hugo, V., Albuquerque, V., & Alexandria, A. (2018). Advances in photoplethysmography signal analysis for biomedical applications. *Sensors, 18*. https://doi.org/10.3390/s18061894, Id.

Oh, S. L., Ng, E. Y., Tan, R. S., & Acharya, U. R. (2018). Automated diagnosis of arrhythmia using combination of cnn and lstm techniques with variable length heart beats. *Computers in Biology and Medicine, 102*, 278–287. https://doi.org/10.1016/j.compbiomed.2018.06.002<http://www.sciencedirect.com/science/article/pii/S0010482518301446>.

Orhan, U. (2013). Real-time chf detection from ecg signals using a novel discretization method. *Computers in Biology and Medicine, 43*(10), 1556–1562. https://doi.org/10.1016/j.compbiomed.2013.07.015<http://www.sciencedirect.com/science/article/pii/S0010482513001893>.

Pecchia, L., Melillo, P., Sansone, M., & Bracale, M. (2011). Discrimination power of short-term heart rate variability measures for chf assessment. *IEEE Transactions on Information Technology in Biomedicine, 15*(1), 40–46. https://doi.org/10.1109/TITB.2010.2091647.

Physionet (2016). Fantasia Database [Online; accessed 15-Sept-2018]. <https://physionet.org/physiobank/database/fantasia>.

Renyi, A. (1961). On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability. Contributions to the theory of statistics* (Vol. 1, pp. 547–561). Berkeley, California: University of California Press<https://projecteuclid.org/euclid.bsmsp/1200512181>.

Shah, S., Sharma, M., Deb, D., & Pachori, R. B. (2019). An automated alcoholism detection using orthogonal wavelet filter bank. *2019 International conference on machine intelligence and signal analysis advances in intelligent systems and computing* (Vol. 748, pp. 473–483). Singapore: Springer. https://doi.org/10.1007/978-981-13-0923-6_41.

Sharma, M., & Acharya, U. R. (2018). Analysis of knee-joint vibroarthographic signals using bandwidth-duration localized three-channel filter bank. *Computers and Electrical Engineering, 72*, 191–202. https://doi.org/10.1016/j.compeleceng.2018.08.019<http://www.sciencedirect.com/science/article/pii/S0045790618311017>.

Sharma, M., Achuth, P., Deb, D., Puthankattil, S. D., & Acharya, U. R. (2018). An automated diagnosis of depression using three-channel bandwidth-duration localized wavelet filter bank with eeg signals. *Cognitive Systems Research, 52*, 508–520<http://www.sciencedirect.com/science/article/pii/S1389041718302298>.

Sharma, M., Achuth, P. V., Pachori, R. B., & Gadre, V. M. (2017). A parametrization technique to design joint time–frequency optimized discrete-time biorthogonal wavelet bases. *Signal Processing, 135*, 107–120.

Sharma, M., Agarwal, S., & Acharya, U. R. (2018). Application of an optimal class of antisymmetric wavelet filter banks for obstructive sleep apnea diagnosis using ecg signals. *Computers in Biology and Medicine, 100*, 100–113. https://doi.org/10.1016/j.compbiomed.2018.06.011<http://www.sciencedirect.com/science/article/pii/S0010482518301598>.

Sharma, M., Bhati, D., Pillai, S., Pachori, R. B., & Gadre, V. M. (2016). Design of time—frequency localized filter banks: Transforming non-convex problem into convex via semidefinite relaxation technique. *Circuits, Systems, and Signal Processing, 35*(10), 3716–3733. https://doi.org/10.1007/s00034-015-0228-9.

Sharma, M., Bhurane, A. A., & Acharya, U. R. (2018). MMSFL-OWFB: A novel class of orthogonal wavelet filters for epileptic seizure detection. *Knowledge-Based Systems, 160*, 265–277. https://doi.org/

10.1016/j.knosys.2018.07.019<http://www.sciencedirect.com/science/article/pii/S0950705118303721>.

Sharma, M., Deb, D., & Acharya, U. R. (2018). A novel three-band orthogonal wavelet filter bank method for an automated identification of alcoholic eeg signals. *Applied Intelligence, 48*(5), 1368–1378. https://doi.org/10.1007/s10489-017-1042-9.

Sharma, M., Dhere, A., Pachori, R. B., & Acharya, U. R. (2017). An automatic detection of focal EEG signals using new class of time–frequency localized orthogonal wavelet filter banks. *Knowledge-Based Systems, 118*, 217–227.

Sharma, M., Dhere, A., Pachori, R. B., & Gadre, V. M. (2017). Optimal duration-bandwidth localized antisymmetric biorthogonal wavelet filters. *Signal Processing, 134*, 87–99.

Sharma, M., Gadre, V. M., & Porwal, S. (2015). An eigenfilter-based approach to the design of time-frequency localization optimized two-channel linear phase biorthogonal filter banks. *Circuits, Systems, and Signal Processing, 34*(3), 931–959.

Sharma, M., Goyal, D., Achuth, P., & Acharya, U. R. (2018). An accurate sleep stages classification system using a new class of optimally time-frequency localized three-band wavelet filter bank. *Computers in Biology and Medicine, 98*, 58–75. https://doi.org/10.1016/j.compbiomed.2018.04.025<http://www.sciencedirect.com/science/article/pii/S0010482518301069>.

Sharma, M., Kolte, R., Patwardhan, P., & Gadre, V. (2010). Time-frequency localization optimized biorthogonal wavelets. In *2010 International Conference on Signal Processing and Communications (SPCOM), 2010* (pp. 1–5).

Sharma, M., & Pachori, R. B. (2017). A novel approach to detect epileptic seizures using a combination of tunable-q wavelet transform and fractal dimension. *Journal of Mechanics in Medicine and Biology, 17*(07), 1740003. https://doi.org/10.1142/S0219519417400036.

Sharma, M., Pachori, R. B., & Acharya, U. R. (2017). A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension. *Pattern Recognition Letters, 94*, 172–179. https://doi.org/10.1016/j.patrec.2017.03.023<http://www.sciencedirect.com/science/article/pii/S0167865517300995>.

Sharma, M., Sharma, P., Pachori, R. B., & Acharya, U. R. (2018). Dual-tree complex wavelet transform-based features for automated alcoholism identification. *International Journal of Fuzzy Systems, 20*(5), 1297–1308. https://doi.org/10.1007/s40815-018-0455-x<https://link.springer.com/article/10.1007/s40815-018-0455-x>.

Sharma, M., Sharma, P., Pachori, R. B., & Gadre, V. M. (2019). Double density dual-tree complex wavelet transform based features for automated screening of knee-joint vibroarthrographic signals. In *Machine intelligence and signal analysis. Advances in intelligent systems and computing* (vol. 748, pp. 279–290). Singapore: Springer.

Sharma, M., Tan, R. S., & Acharya, U. R. (2018). A novel automated diagnostic system for classification of myocardial infarction ecg signals using an optimal biorthogonal filter bank. *Computers in Biology and Medicine*. https://doi.org/10.1016/j.compbiomed.2018.07.005<http://www.sciencedirect.com/science/article/pii/S0010482518301884>.

Sharma, M., Vanmali, A. V., & Gadre, V. M. (2013). *Construction of wavelets: Principles and practices in wavelets and fractals in earth system sciences*. CRC Press, Taylor and Francis Group.

Student (1908). The probable error of a mean. *Biometrika, 6*(1), 1–25<http://www.jstor.org/stable/2331554>.

Sudarshan, V. K., Acharya, U., Oh, S. L., Adam, M., Tan, J. H., Chua, C. K., ... Tan, R. S. (2017). Automated diagnosis of congestive heart failure using dual tree complex wavelet transform and statistical features extracted from 2s of ecg signals. *Computers in Biology and Medicine, 83*, 48–58. https://doi.org/10.1016/j.compbiomed.2017.01.019<http://www.sciencedirect.com/science/article/pii/S0010482517300264>.

Tripoliti, E. E., Papadopoulos, T. G., Karanasiou, G. S., Naka, K. K., & Fotiadis, D. I. (2017). Heart failure: Diagnosis, severity estimation and prediction of adverse events through machine learning techniques. *Computational and Structural Biotechnology Journal, 15*, 26–47. https://doi.org/10.1016/j.csbj.2016.11.001<http://www.sciencedirect.com/science/article/pii/S2001037016300460>.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. Berlin, Heidelberg: Springer-Verlag.

Yildirim, Özal, Plawiak, P., Tan, R.-S., & Acharya, U. R. (2018). Arrhythmia detection using deep convolutional neural network with long duration ecg signals. *Computers in Biology and Medicine, 102*, 411–420. https://doi.org/10.1016/j.compbiomed.2018.09.009<http://www.sciencedirect.com/science/article/pii/S0010482518302713>.

Yu, S.-N., & Lee, M.-Y. (2012). Bispectral analysis and genetic algorithm for congestive heart failure recognition based on heart rate variability. *Computers in Biology and Medicine, 42*(8), 816–825. https://doi.org/10.1016/j.compbiomed.2012.06.005<http://www.sciencedirect.com/science/article/pii/S0010482512000972>.

Zhao, L., Wei, S., Zhang, C., Zhang, Y., Jiang, X., Liu, F., & Liu, C. (2015). Determination of sample entropy and fuzzy measure entropy parameters for distinguishing congestive heart failure from normal sinus rhythm subjects. *Entropy, 17*(9), 6270–6288. https://doi.org/10.3390/e17096270<http://www.mdpi.com/1099-4300/17/9/6270>.