J. Med. Biol. Eng. (2017) 37:276–287
DOI 10.1007/s40846-017-0222-8

ORIGINAL ARTICLE

A Spectral Exponent-Based Feature of RR Interval Data for Congestive Heart Failure Discrimination Using a Wavelet-Based Approach

Suparerk Janjarasjitt[1]

Received: 13 April 2016 / Accepted: 6 September 2016 / Published online: 24 March 2017

Abstract Congestive heart failure (CHF) is a common chronic condition which affects several millions of people around the world. Heart rate data which are crucial information needed for diagnosis and classification of CHF are similar to several physiological signals that exhibit an extraordinary range of patterns and behaviors. In this study, the spectral exponents of RR interval data of two groups of subjects, i.e., subjects with CHF and subjects with normal sinus rhythm, obtained using a wavelet-based approach are examined where the second order Daubechies wavelets are used. The spectral exponent of RR interval data is determined from a slope of logarithm of variance of wavelet coefficients ($\log_2 \mathrm{var}(d_{m,n})$) versus levels of wavelet-based decomposition (m) graph. In particular, the spectral exponent is estimated from the levels of wavelet-based decomposition ranges between $m = 1$ and $m = 3$ corresponding to finer-scale components of RR interval data. The minimum of spectral exponents of epochs of RR interval data is proposed as a quantitative feature for discriminating a subject with CHF. The computational results show that the subjects with CHF can be perfectly discriminated from the subjects with normal sinus rhythm using the spectral exponent-based feature. Furthermore, the perfect CHF discrimination can be achieved using the RR interval data with epoch size as short as 128 beats (approximately 2 min).

Keywords Congestive heart failure · RR interval data · Normal sinus rhythm · Discrimination · Spectral exponent · Wavelet analysis

1 Introduction

Heart failure (HF), also called congestive heart failure, is a chronic and progressive condition in which the heart cannot pump enough blood through to meet the body's needs for blood and oxygen [1–3]. A complex clinical syndrome of heart failure results from any structural or functional impairment of ventricular filling or ejection of blood [4]. Heart failure is a common but serious condition. In the United States, there are about 5.1 million people who have heart failure [5]. In 2009, heart failure is a contributing cause of one in nine death in the United States [5]. Symptoms of heart failure may include shortness of breath; rapid or irregular heartbeat; fatigue and weakness; swelling in your legs, ankles and feet; persistent cough or wheezing [6]. Common causes of heart failure are coronary artery disease, high blood pressure, and diabetes [3].

To diagnose heart failure, medical history and symptoms are reviewed and the physical examination is performed. One or more diagnostic tests such as blood tests, chest X-ray, electrocardiogram (ECG), echocardiogram, radionuclide ventriculography may be ordered for further diagnosis. There is no single diagnostic test for heart failure [4]. The ACCF/AHA Stages of Heart Failure [7] and the New York Heart Association (NYHA) Functional Classification [7, 8] are commonly used heart failure classification systems that provide complementary information about the presence and severity of heart failure [4]. The ACCF/AHA Stages of Heart Failure categorize patients into four stages, i.e., A–D [7], emphasizing the

✉ Suparerk Janjarasjitt
suparerk.j@ubu.ac.th

[1] Department of Electrical and Electronic Engineering, Ubon Ratchathani University, 85 Sathonlamak Road, Warin Chamrap, Ubon Ratchathani 34190, Thailand

development and progression of disease [4]. The NYHA Functional Classification categorize patients into four classes, i.e., I–IV [8], focusing on exercise capacity and the symptomatic status of the disease [4]. Four NHYA classes indicate the severity of symptoms.

ECG which records the electrical activity of the heart is one of the most common diagnostic tests. ECG helps in diagnosis of heart rhythm and also heart damage. Heart rate variability (HRV) analysis is one of conventional and fundamental methods used for assessing heart's health and diagnosing heart diseases. HRV may be evaluated using a number of methods [9] including time domain methods, statistical methods, geometric methods, frequency domain methods, and nonlinear methods. HRV has also been applied for diagnosing and classifying congestive heart failure. Measures obtained from various HRV analysis techniques were investigated. Several measures (for example, [10–16]) showed their potentials for application on congestive heart failure classification. Even though intriguing results obtained in these studies, multiple features were used and complicated classifiers were applied.

In the past few decades, concepts and computational tools derived from the study of complex systems including chaos theory, nonlinear dynamics and fractals have gained increasing interests and been widely applied to various applications in biology and medicine. One of the main reasons is that physiological signals and systems can exhibit an extraordinary range of patterns and behaviors [17] that may defy concepts and properties of linear systems such as superposition theorem. In particular, nonlinear phenomena are certainly involved in the underlying mechanism of heartbeat and its variability [9]. Underlying dynamics of heartbeat is determined by complex interactions of hemodynamic, electrophysiological, and humoral variables as well as by the autonomic and central nervous regulations [9]. Complex systems analysis and concepts have therefore played a remarkable role in cardiology. A number of measures that have been applied to examine characteristics of HRV include spectral exponent, scaling exponent, correlation dimension, and Lyapunov exponent [9].

Fractals is one of several complex systems concepts that have been used to characterize heart's underlying dynamics and behaviors. Typically, the mathematical concept of fractals is associated with irregular objects that manifest a geometric property called scale-invariance or self-similarity [17, 18]. Fractal forms are composed of subunits resembling the structure of the macroscopic object [17] which in nature can emerge from statistical scaling behavior in the underlying physical phenomena [19]. The His-Purkinje conduction system is a complex anatomic structure exhibiting fractal like geometry [17, 20]. There have been evidences that biological and physiological systems exhibit scale-invariant or scale-free behaviors. Fractal properties of biological and physiological systems can be strikingly different in their nature, origin, and appearance [21]. Such scale-invariant or scale-free behavior is a tendency of a complex system to develop long-range correlations in time and space [22–24].

A spectral exponent γ is a measure that characterizes the power law behavior of $1/f$ processes. The $1/f$ processes is an important class of statistical scale-invariant or self-similar random processes [19, 25, 26]. The spectral exponent specifies the distribution of power in $1/f$ processes from low to high frequencies. The wavelet transform is a natural tool for characterizing self-similar or scale-invariant signals and plays a significant role in the study of self-similar signals and systems [19, 25, 26], in particular $1/f$ processes [19, 25, 26]. In addition to the power spectral density traditionally used, the spectral exponent can be determined using a wavelet-based approach [19, 25, 26] whereas the wavelet-based approach can be used to estimate the power spectral density [27]. From the wavelet-based representations for $1/f$ processes [19, 25, 26], it was proved that the spectral exponent of a nearly-$1/f$ signal is characterized by a linear relationship between logarithm of variance of wavelet coefficients and scale (or level of wavelet decomposition). The wavelet-based approach was shown to allow an unbiased estimate of the spectral exponent [27].

The spectral exponent is directly related to the self-similarity parameter, i.e., Hurst exponent H [19, 25, 26]. The spectral exponent is also related to the scaling exponent α [28] which characterizes the long-range correlation of a signal. It was computationally shown that the relationship between the spectral exponent and the scaling exponent is $\alpha = (\gamma + 1)/2$ [29]. Furthermore, it was shown that the spectral exponent obtained using the wavelet-based approach provides the better performance on long-range correlation characterization than the scaling exponent obtained using the detrended fluctuation analysis (DFA) method [29]. The main advantage of the DFA method is that it can detect long-range correlations in non-stationary signal [30]. Investigations of cardio-pulmonary system including congestive heart failure and HRV are one of earliest applications of the DFA method. In [30, 31], the DFA method was applied to examine the long-range correlation behaviors of heart rate time series of subjects with normal sinus rhythm and those of subjects with congestive heart failure. From both studies, it was found that the scaling exponents of heart rate time series of both subjects were different. Furthermore, the crossover phenomena was observed [30].

The wavelet-based approach for spectral exponent estimation has been applied to a number of electrophysiological signals associated with various physiological and pathological states such as epileptic intracranial EEG (or

electrocorticogram) signals [32] and sleep EEG signals [33]. In addition to the spectral exponents of signals that are typically determined from a wide range of scales, a variety of quantitative features have been extracted from slopes of the logarithm of variance of wavelet coefficients versus scale graph over specific ranges of scales for classification and detection purposes [34]. In this study, a single quantitative feature extracted from the spectral exponent obtained using a wavelet-based approach is applied for congestive heart failure classification. A minimum of spectral exponents is the feature based on spectral exponent examined. Since a single quantitative feature is used, a simple logical comparison, i.e., a thresholding technique, is applied for discriminating a subject with congestive heart failure.

2 Materials and Methods

2.1 Data and Subjects

Two sets of RR interval data are examined in this study. The first set belongs to the RR interval data of subjects with normal sinus rhythm (NSR group) that were obtained from the MIT-BIH Normal Sinus Rhythm Database (NSRDB) available online at http://physionet.org/physiobank/database/nsrdb/ [35]. Another set belongs to the RR interval data of subjects with congestive heart failure (CHF group) that were obtained from the BIDMC Congestive Heart Failure Database (CHFDB) available online at http://physionet.org/physiobank/database/chfdb/ [35].

The MIT-BIH Normal Sinus Rhythm Database includes 18 long-term ECG recordings of subjects referred to the Arrhythmia Laboratory at Boston's Beth Israel Hospital (now the Beth Israel Deaconess Medical Center). The subjects included 5 men (aged from 26 to 45 years) and 13 women (aged from 20 to 50). All subjects included in this database were found to have had no significant arrhythmias.

The BIDMC Congestive Heart Failure Database includes 15 long-term ECG recordings of subjects with severe congestive heart failure (NYHA class 3–4). The subjects included 11 men (aged from 22 to 71 years) and 4 women (aged from 54 to 63 years). This group of subjects was part of a larger study group receiving conventional medical therapy prior to receiving the oral inotropic agent, milrinone [35].

The original ECG recordings of subjects in the MIT-BIH Normal Sinus Rhythm Database were digitized at 128 samples per second while the original ECG recordings of subjects in the BIDMC Congestive Heart Failure Database were digitized at 250 samples per second. Each ECG recording is approximately 20 h in duration. The beat annotations in both databases were obtained by automated analysis.

2.2 A Wavelet-Based Approach and a Spectral Exponent-Based Feature

A quantitative feature of RR interval data used in this study is obtained from the derivation of the wavelet-based representations for $1/f$ processes [19, 25, 26]. Typically, models of $1/f$ processes are represented using a frequency domain characterization. The dynamics of $1/f$ processes exhibit the power-law behavior [36] and can be characterized in the form of [25, 26]

$$S_x(\omega) \sim \frac{\sigma_x^2}{|\omega|^\gamma} \tag{1}$$

over several decades of frequency ω, where $S_x(\omega)$ is the Fourier transform of the signal $x(t)$ and γ is the spectral exponent. In [19, 25, 26], it was proved that the spectral exponent of $1/f$ processes can be determined from the corresponding wavelet coefficients.

The steps for computing the spectral exponent of RR interval data are as follows:

1. Decompose RR interval data into M levels using the wavelet-basis expansions to obtain the wavelet coefficients $d_{m,n}$ where m denotes a level, i.e., $m = 1, 2, \ldots, M$;
2. Compute the variance of wavelet coefficients $d_{m,n}$ corresponding to each level m;
3. Take the logarithm to base 2 of the corresponding variances of wavelet coefficients;
4. Determine the spectral exponent γ by estimating the slope of a $\log_2\text{var}(d_{m,n})$-$m$ graph over a specified range of levels m.

The minimum of spectral exponents of RR interval data is proposed as the quantitative feature for discriminating a subject with congestive heart failure.

2.3 Data Analysis and Classification

The RR interval data of subjects from both NSR and CHF groups are partitioned into four different sizes of epochs: 64 (2^6), 128 (2^7), 4096 (2^{12}), and 8192 (2^{13}) beats or samples. These epoch sizes are equivalent to approximately 1, 2, 56, and 114 min long of RR interval data, respectively. In the computational analysis, the RR interval data are partitioned without overlap at the epoch sizes of 64 and 128 while at the epoch sizes of 4096 and 8192, the RR interval data are partitioned with 50% overlap. Remark that the total number of epochs of RR interval data of each subject is varied corresponding to its length of RR interval data originally obtained from the NSRDB and CHFDB

databases. In general, the RR interval data span approximately 12 h (or a half day).

The wavelet-based approach is applied to epochs of RR interval data using the second order Daubechies wavelets. The Daubechies wavelet family is one of the most commonly used wavelet families and has several great characteristics including orthogonality and finite compact support. The number of vanishing moments for wavelet function of the second order Daubechies is 2. The epochs of RR interval data are decomposed into the highest levels corresponding to their epoch sizes. The corresponding levels of wavelet-based decomposition of RR interval data M are 4, 5, 10, and 11 for the epoch sizes of 64, 128, 4096, and 8192, respectively. However, the level of wavelet-based decomposition used in the spectral exponent estimation ranges between $m = 1$ and $m = 3$. Here, the corresponding spectral exponent is denoted by $\gamma_{(1,3)}$. This range of levels of wavelet-based decomposition is chosen because it provides the best discrimination results. A linear least-squares regression technique is used for estimating the spectral exponent γ, i.e., the slope of a $\log_2\text{var}(d_{m,n})$-$m$ graph.

To evaluate the performance of the proposed feature, a simple binary classification is performed. By aiming to obtain a perfect classification of CHF subjects, the highest value of proposed features of the CHF subjects is used as a threshold. The classification performance is evaluated using the following three measures: accuracy (Ac), sensitivity (Se), and specificity (Sp) that are, respectively, given by

$$Ac = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}), \quad (2)$$

$$Se = \text{TP}/(\text{TP} + \text{FN}), \quad (3)$$

$$Sp = \text{TN}/(\text{TN} + \text{FP}) \quad (4)$$

where TP, TN, FP, and FN denote a number of true positives, a number of true negatives, a number of false positives, and a number of false negatives. A receiver operating characteristic (ROC) curve, that is a two-dimensional graph in which the true positive rate (sensitivity) is plotted on the X-axis and the false positive rate (1-specificity) is plotted on the Y-axis [37], is also used to evaluate the classification performance. Since the number of subjects is small, the leave-one-out cross validation is further applied to validate the classification performance.

In addition, the performance of the proposed feature is compared to another two nonlinear features, the approximate entropy $ApEn$ and the sample entropy $SampEn$, and their sum referred to as SUM_IE [10]. The feature SUM_IE is the best individual quantitative feature among the other two features SUM_TD, a combination of time-domain features, and SUM_FD, a combination of frequency-domain features, examined in Ref. [10]. All computational experiments are performed using MATLAB® (MathWorks®, Natick, Massachusetts, USA) running on a Mac OS X notebook computer with 2.6 GHz Intel Core i5 CPU and 8 GB (1600 MHz, DDR3) memory.

3 Results

3.1 Characteristics of $\log_2\text{var}(d_{m,n})$-$m$ Graphs of RR Interval Data

Plots of average $\log_2\text{var}(d_{m,n})$ of 128-beat epochs of RR interval data corresponding to each subject from the NSR group are shown in Fig. 1a while those corresponding to

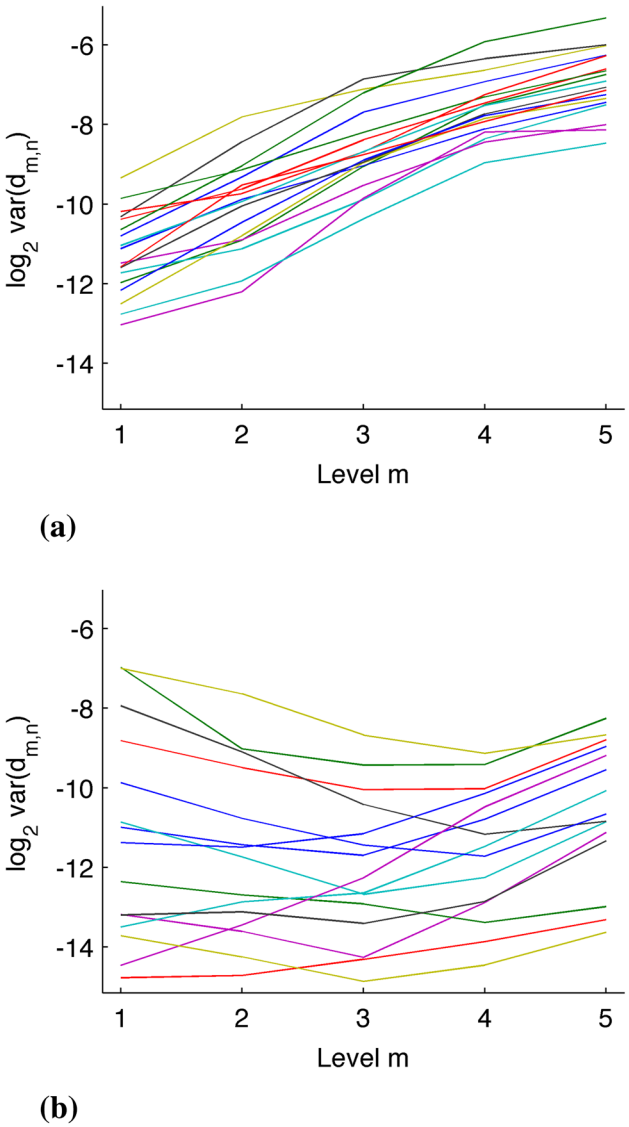


Fig. 1 Plots of average $\log_2\text{var}(d_{m,n})$ of 128-beat epochs of RR interval data. **a** Subjects with normal sinus rhythm. **b** Subjects with congestive heart failure

each subject from the CHF group are shown in Fig. 1b. A distinguishing characteristic of 128-beat epochs of RR interval data differentiating between the NSR and CHF group is observed. The value of $\log_2\text{var}(d_{m,n})$ of 128-beat epochs of RR interval data of the NSR subjects tends to increase as the level m increases. On the contrary, the plot of $\log_2\text{var}(d_{m,n})$ of 128-beat epochs of RR interval data of the CHF subjects generally tends to lie flat.

Figure 2 shows plots of average $\log_2\text{var}(d_{m,n})$ of 4096-beat epochs of RR interval data corresponding to each subject from the NSR group and those corresponding to each subject from the CHF group. With the wider range of levels of wavelet-based decomposition, a further intriguing characteristic of 4096-beat epochs of RR interval data is revealed. The plot of $\log_2\text{var}(d_{m,n})$ of 4096-beat epochs of RR interval data of the NSR subjects tends to be straight or its slope tends to slightly change. The slope of $\log_2\text{var}(d_{m,n})$-$m$ graphs of 4096-beat epochs of RR interval data of the CHF subjects however substantially varies, in particular compared between those corresponding to smaller and larger levels of wavelet-based decomposition.

Furthermore, the distribution of $\log_2\text{var}(d_{m,n})$ of 128-beat epochs of RR interval data of all NSR and CHF subjects are shown in Fig. 3. Also, the box plots shown in Fig. 4 show the distribution of $\log_2\text{var}(d_{m,n})$ of 4096-beat epochs of RR interval data of all NSR and CHF subjects, respectively. The trends of $\log_2\text{var}(d_{m,n})$-$m$ graphs of RR interval data of subjects from both NSR and CHF groups are evidenced. Tables 1 and 2 summarize the means and the standard deviations of $\log_2\text{var}(d_{m,n})$ of 128-beat epochs

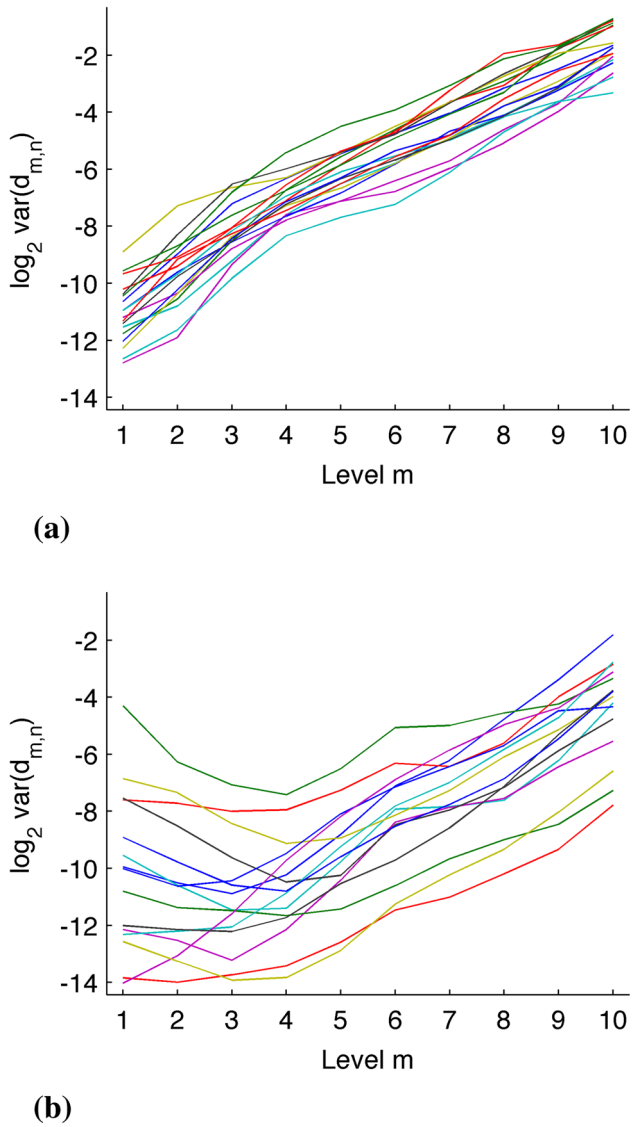


Fig. 2 Plots of average $\log_2\text{var}(d_{m,n})$ of 4096-beat epochs of RR interval data. **a** Subjects with normal sinus rhythm. **b** Subjects with congestive heart failure

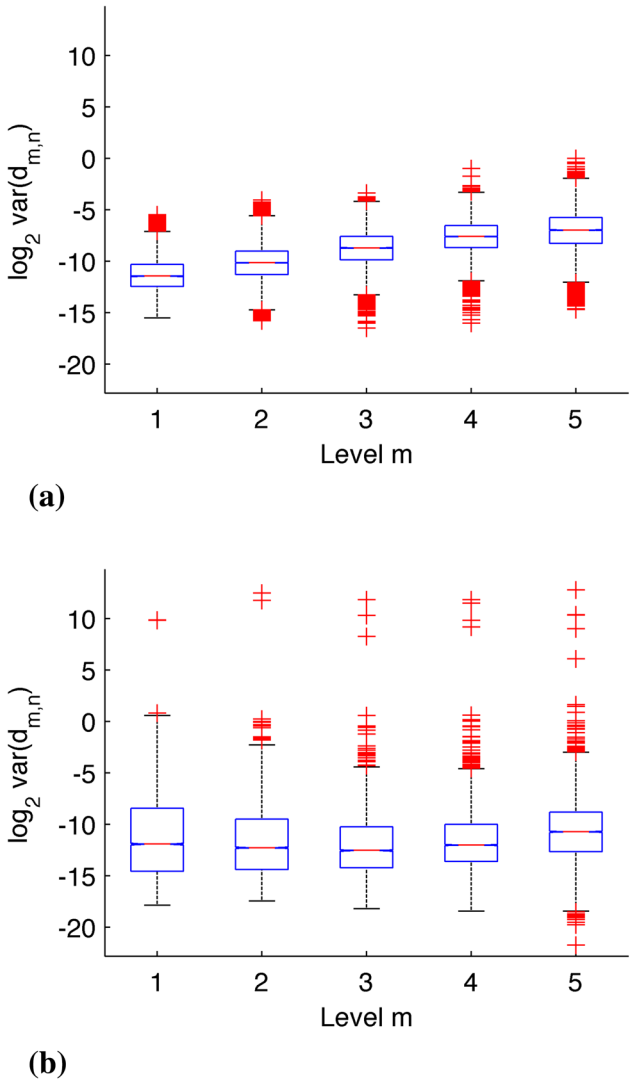
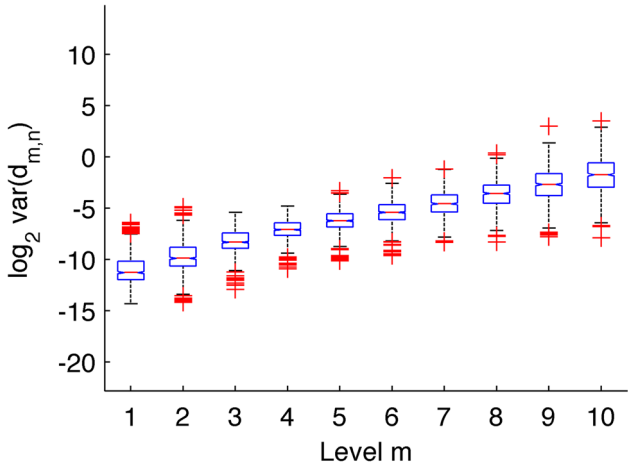
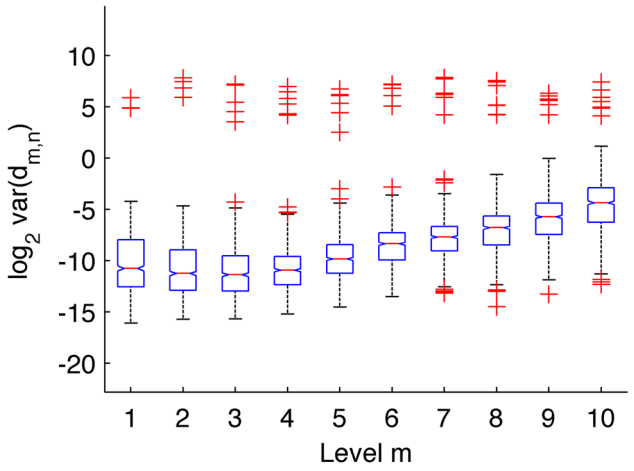


Fig. 3 Distribution of $\log_2\text{var}(d_{m,n})$ of 128-beat epochs of RR interval data. **a** Subjects with normal sinus rhythm. **b** Subjects with congestive heart failure



(a)



(b)

Fig. 4 Distribution of $\log_2\text{var}(d_{m,n})$ of 4096-beat epochs of RR interval data. **a** Subjects with normal sinus rhythm. **b** Subjects with congestive heart failure

Table 1 The statistical values (mean ± SD) of $\log_2\text{var}(d_{m,n})$ of 128-beat epochs of RR interval data

Level *m*	Subject group	
	NSR	CHF
1	−11.3007 ± 1.6831	−11.4150 ± 3.4308
2	−10.1249 ± 1.8257	−11.8220 ± 2.9184
3	−8.7636 ± 1.7090	−12.1238 ± 2.6711
4	−7.6503 ± 1.6448	−11.7084 ± 2.5401
5	−7.0356 ± 1.9289	−10.6717 ± 2.7109

of RR interval data of NSR and CHF subjects and those of 4096-beat epochs of RR interval data of subjects from NSR and CHF subjects, respectively.

Table 2 The statistical values (mean ± SD) of $\log_2\text{var}(d_{m,n})$ of 4096-beat epochs of RR interval data

Level *m*	Subject group	
	NSR	CHF
1	−11.0810 ± 1.4629	−10.3232 ± 3.1946
2	−9.7728 ± 1.5125	−10.8161 ± 3.0295
3	−8.2449 ± 1.1623	−11.1230 ± 2.7658
4	−7.0858 ± 0.9785	−10.8236 ± 2.6168
5	−6.2300 ± 1.1012	−9.7712 ± 2.5648
6	−5.4623 ± 1.1903	−8.4622 ± 2.5385
7	−4.5854 ± 1.2567	−7.8088 ± 2.4032
8	−3.6345 ± 1.4119	−6.9810 ± 2.4576
9	−2.7463 ± 1.5912	−5.8914 ± 2.5480
10	−1.8033 ± 1.8129	−4.5737 ± 2.6548

3.2 Discrimination of Congestive Heart Failure Patients

The spectral exponents $\gamma_{(1,3)}$ of epochs of RR interval data of the NSR and CHF subjects at the epoch sizes of 64, 128, 4096, and 8192 beats are compared in the box plots shown in Figs. 5a, b, c and d, respectively. The spectral exponents $\gamma_{(1,3)}$ of epochs of RR interval data of subject numbers 1 through 18 on the left-hand side of the vertical dotted line are belonged to the NSR subjects while the spectral exponents $\gamma_{(1,3)}$ of epochs of RR interval data of subject numbers 1–15 on the right-hand side of the vertical dotted line are belonged to the CHF subjects. As evidenced from the plots of $\log_2\text{var}(d_{m,n})$, the spectral exponents $\gamma_{(1,3)}$ of epochs of RR interval data of the CHF subjects tend to be less than those of epochs of RR interval data of the NSR subjects.

The box plots shown in Figs. 6a, b, c, d, respectively, compare between the spectral exponent-based features, i.e., the minimums of spectral exponents $\gamma_{(1,3)}$, of epochs of RR interval data of corresponding NSR subjects and the spectral exponent-based features of epochs of RR interval data of corresponding CHF subjects at the epoch sizes of 64, 128, 4096, and 8192 beats. At the epoch sizes of both 128 and 4096 beats, the spectral exponent-based features of epochs of RR interval data of the NSR and CHF subjects are completely separated. On the other hand, such characteristic does not hold at the other two epoch sizes. The means, minimums, and maximums of the spectral exponent-based features of epochs of RR interval data of the NSR and CHF subjects at all epoch sizes are shown in Table 3. At the thresholds of −2.1199, −1.7954, −0.2795, and 0.5306, the corresponding classification results and also the performance of congestive heart failure classification are summarized in Table 4. Furthermore, the results

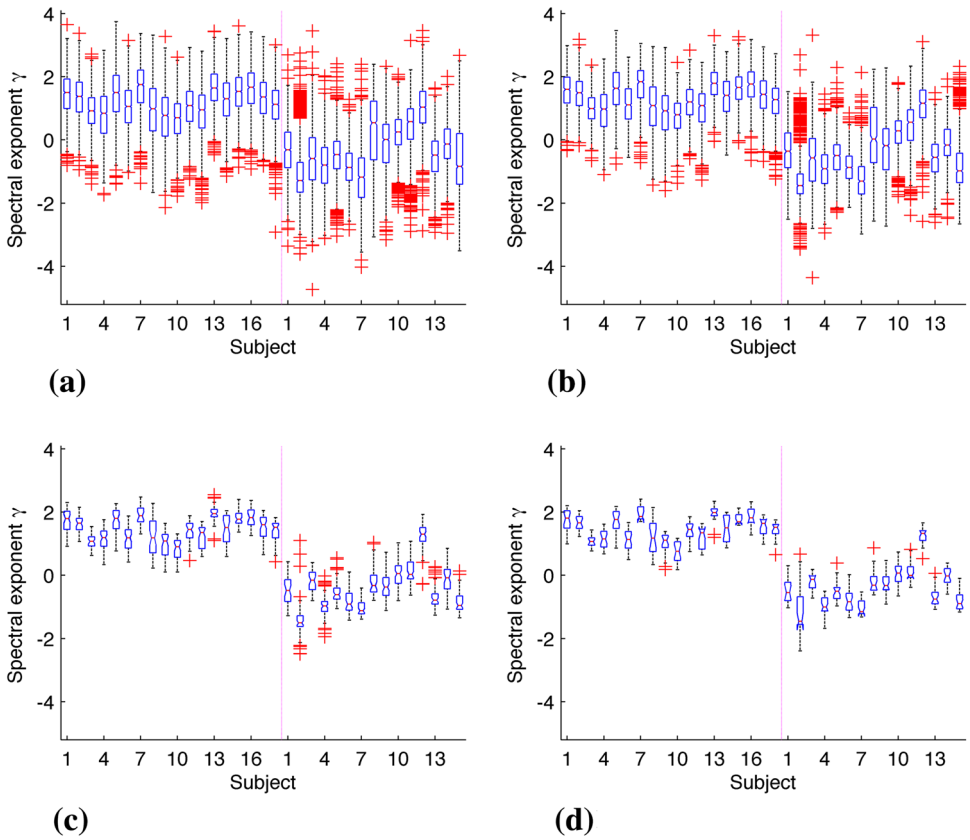


Fig. 5 Comparison of the spectral exponents of RR interval data of all subjects corresponding to various epoch sizes. **a** Epoch size of 64 beats. **b** Epoch size of 128 beats. **c** Epoch size of 4096 beats. **d** Epoch size of 8192 beats

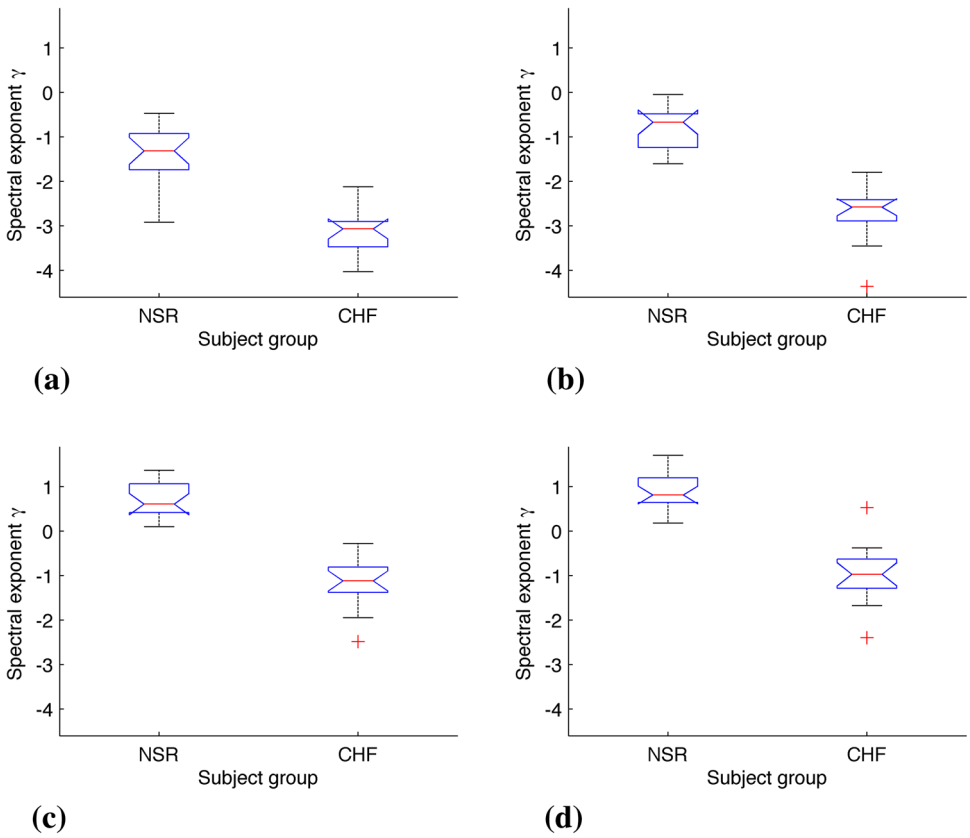


Fig. 6 Comparison of the corresponding minimum spectral exponents of RR interval data of each subject group corresponding to various epoch sizes. **a** Epoch size of 64 beats. **b** Epoch size of 128 beats. **c** Epoch size of 4096 beats. **d** Epoch size of 8192 beats

Table 3 Statistical values of the spectral exponent-based features of epochs of RR interval data

Epoch size	Mean		Min		Max	
	NSR	CHF	NSR	CHF	NSR	CHF
64	−1.3674	−3.2139	−2.9171	−4.7305	−0.4693	−2.1199
128	−0.7635	−2.6977	−1.6023	−4.3561	−0.0463	−1.7954
4096	0.6852	−1.1785	0.0998	−2.4830	1.3683	−0.2795
8192	0.8543	−0.9542	0.1808	−2.3941	1.7025	0.5306

of leave-one-out cross validation are summarized in Table 5. The one observation of CHF subjects left out in the cross validation leads to a decrease in accuracy and sensitivity while the specificity remains the same.

Figure 7a–d shows the ROC curves of congestive heart failure classification using the spectral exponent-based features of epochs of RR interval data with the size of 64, 128, 4096, and 8192 beats, respectively. The ROC curves of congestive heart failure classification using the features of epochs of RR interval data with the size of 64, 128, 4096, and 8192 beats based on the *ApEn* (plotted in blue), the *SampEn* (plotted in black) and the *SUM_IE* are, respectively, shown in Fig. 8a–d. The corresponding areas under ROC curves shown in Figs. 7a–d and 8a–d are compared in Table 6. Also, the average computational times of spectral exponents, *ApEn* and *SampEn* of epochs of RR interval data with epoch sizes of 64, 128, 4096, and 8192 beats are summarized in Table 6. As expected, the areas under ROC curves of congestive heart failure classification using the spectral exponent-based features of epochs of RR interval data with the size of 128 and 4096 beats are 1.0. The best area under ROC curve of congestive heart failure classification among three nonlinear features, i.e., *ApEn*, *SampEn*, and *SUM_IE*, is 0.9370 achieved using the feature based on sample entropy *SampEn* of epochs of RR interval data with the size of 128 beats.

4 Discussion

From the computational results, at the larger size of epochs, i.e., 4096 beats, it is observed that the epochs of RR interval data of subjects with congestive heart failure have a shorter-range correlation compared to those of RR interval data of subjects with normal sinus rhythm in which their corresponding plots of $\log_2 \mathrm{var}(d_{m,n})$ tend to be straight. The so-called *crossover phenomena*, the phenomena that a slope of corresponding plot changes, is also observed in the plots of $\log_2 \mathrm{var}(d_{m,n})$ of epochs of RR interval data. This is consistent with the previous findings that can be evidenced in the power spectrum [31] and also the plots of the DFA [30]. The crossover points are about at the levels of wavelet-based decomposition $m = 3$ and $m = 4$.

Table 4 Performance of the congestive heart failure classification using the spectral exponent-based features

Epoch size	TP	TN	FP	FN	*Ac*	*Se*	*Sp*
64	15	16	2	0	0.9394	1.0	0.8889
128	15	18	0	0	1.0	1.0	1.0
4096	15	18	0	0	1.0	1.0	1.0
8192	15	14	4	0	0.8788	1.0	0.7778

Table 5 Results of leave-one-out cross validation of the congestive heart failure classification using spectral exponent-based features

Epoch Size	TP	TN	FP	FN	*Ac*	*Se*	*Sp*
64	14	16	2	1	0.9091	0.9333	0.8889
128	14	18	0	1	0.9697	0.9333	1.0
4096	14	18	0	1	0.9697	0.9333	1.0
8192	14	14	4	1	0.8485	0.9333	0.7778

The spectral exponents of epochs of RR interval data of subjects with congestive heart failure determined from the range of lower levels of wavelet-based decomposition, i.e., between $m = 1$ and $m = 3$, are significantly lower than those of epochs of RR interval data of subjects with normal sinus rhythm with $p \ll 0.0001$. In fact, there is only the 12th subject with congestive heart failure whose spectral exponents $\gamma_{(1,3)}$ are not separable from those of epochs of RR interval data of subjects with normal sinus rhythm. The level of wavelet-based decomposition ranging from $m = 1$ and $m = 3$ corresponds to finer-scale (high-frequency) components of underlying dynamics of heartbeat. The spectral exponents of epochs of RR interval data of subjects with congestive heart failure and those of subjects with normal sinus rhythm determined from the range of higher levels of wavelet-based decomposition, i.e., between $m = 4$ and $m = 10$, corresponding to large-scale (low-frequency) components of underlying dynamics of heartbeat are approximately identical.

Furthermore, analogous to the power spectral density, the plots of $\log_2 \mathrm{var}(d_{m,n})$ of epochs of RR interval data of subjects with normal sinus rhythm indicates that the corresponding power of epochs of RR interval data constantly

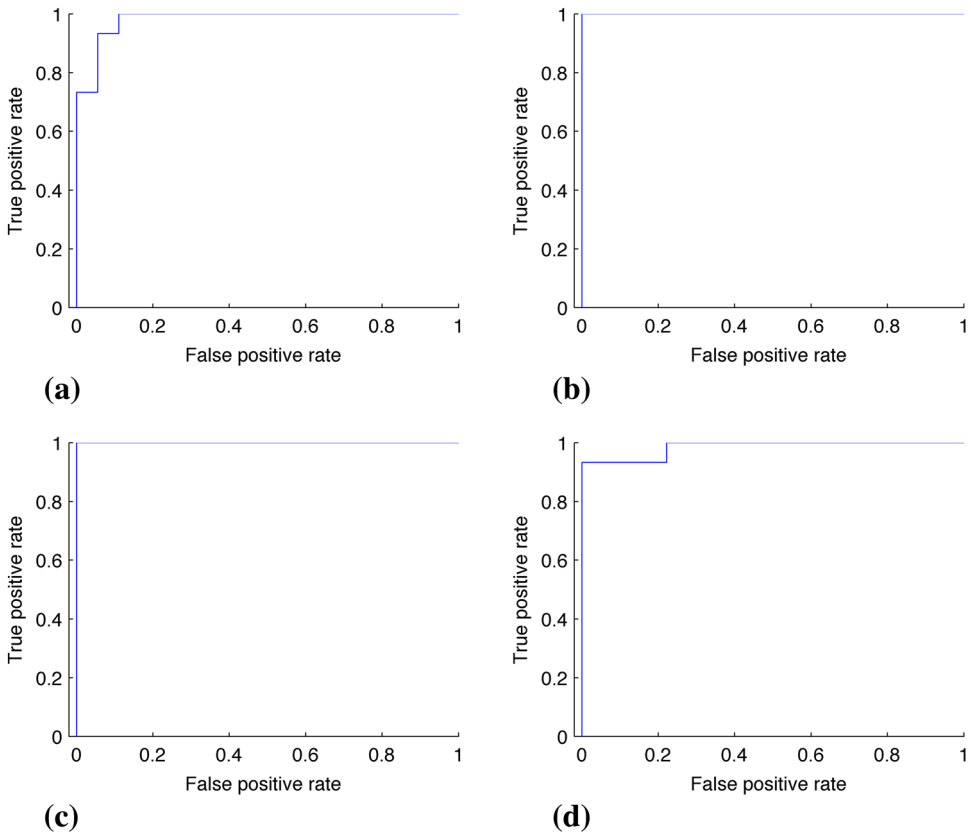


Fig. 7 ROC curves of congestive heart failure classification using the spectral exponents corresponding to various epoch sizes. **a** Epoch size of 64 beats. **b** Epoch size of 128 beats. **c** Epoch size of 4096 beats. **d** Epoch size of 8192 beats

increases as the scale becomes larger or the frequency decreases. Similarly, from the plots of $\log_2 \mathrm{var}(d_{m,n})$ of epochs of RR interval data at the larger scales, i.e., between the levels $m = 4$ and $m = 10$, it is shown that the power of epochs of RR interval data of subjects with congestive heart failure increases as the scale becomes larger. At the lower scales, i.e., between $m = 1$ and $m = 4$, the power of epochs of RR interval data of subjects with congestive heart failure however generally remains the same or slightly decreases as the scale becomes larger. In particular, the spectral exponents $\gamma_{(1,3)}$ of epochs of RR interval data of subjects with congestive heart failure are mostly negative. This suggests that the power of epochs of RR interval data of subjects with congestive heart failure decreases as the frequency decreases. It is also observed that the spectral exponent of RR interval data of both subjects with congestive heart failure and subjects with normal sinus rhythm increases as the epoch size increases. This implies that as the time window becomes longer there is a greater increase in larger-scale (lower-frequency) components of underlying dynamics of heartbeat compared to the finer-scale (higher-frequency) components of underlying dynamics of heartbeat.

The computational results show that the feature of spectral exponents of RR interval data can be used to discriminate subjects with congestive heart failure excellently. The epoch size used in RR interval data segmentation has however an influence on the discrimination result. The segmentation of RR interval data into epochs with size ranging between 128 (2^7) and 4096 (2^{12}) beats results in the perfect discrimination of subjects with congestive heart failure. This is because the underlying dynamics of heartbeat associated with congestive heart failure does not manifest in too short epochs of RR interval data, i.e., epoch size of 64 beats (approximately 1 min). On the other hand, too long epochs of RR interval data taint the distinct underlying dynamics of heartbeat associated with congestive heart failure.

In Ref. [10], the perfect CHF classification was obtained with the accuracy, the sensitivity, and the specificity of 100.0, 100.0, and 100.0%, respectively, using three short-term HRV features including a combination of time-domain features *SUM_TD*, a combination of frequency-domain features *SUM_FD*, and a combination of nonlinear features *SUM_IE*. The three features were however based on nine measures while an SVM classifier with polynomial kernel function was used [10]. The approach proposed in Ref. [10] was compared and shown to be superior to a number of methods including [14–16]. The genetic algorithm and the *k*-nearest neighbor classifier was applied in Ref. [14] and the best performance with the sensitivity and the specificity of 100.0 and 94.74%, respectively. A

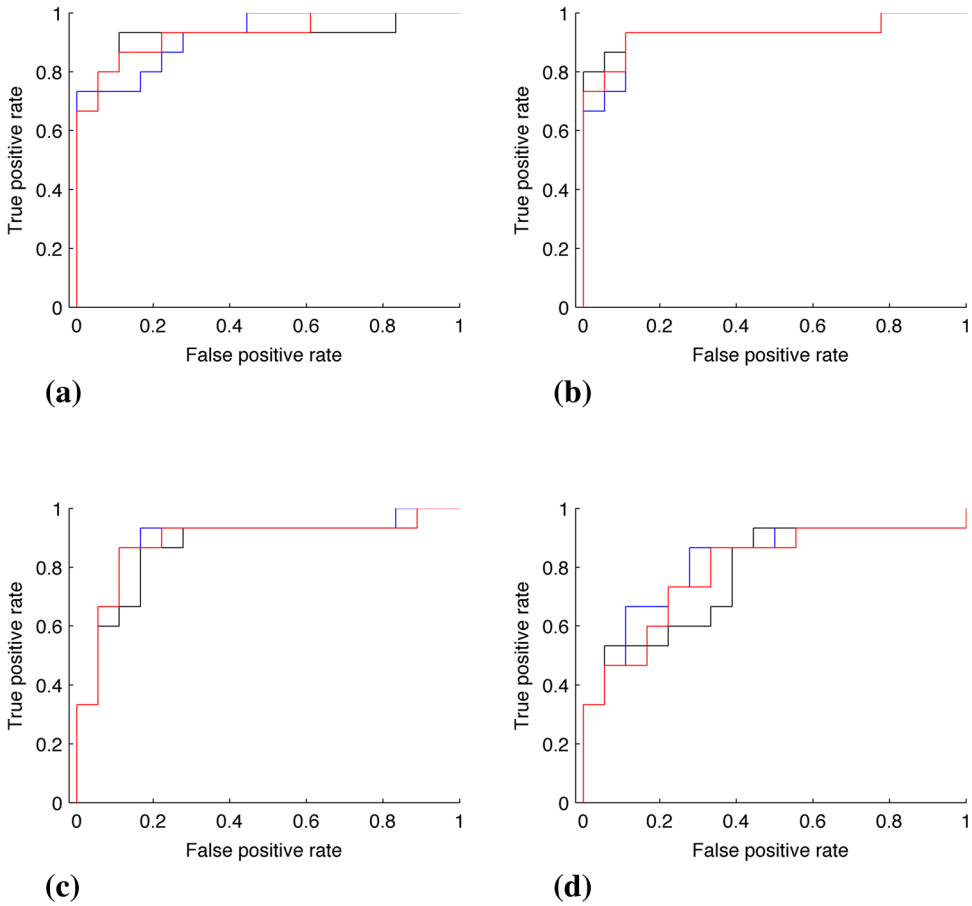


Fig. 8 ROC curves of congestive heart failure classification using the *ApEn* (plotted in *blue*), *SampEn* (plotted in *black*) and *SUM_IE* (plotted in *red*) corresponding to various epoch sizes. **a** Epoch size of 64 beats, **b** Epoch size of 128 beats, **c** Epoch size of 4096 beats, and **d** Epoch size of 8192 beats

Table 6 Comparison of performance of the congestive heart failure classification

Epoch Size	Area under curve				Computational time (ms)		
	γ	*ApEn*	*SampEn*	*SUM_IE*	γ	*ApEn*	*SampEn*
64	0.9815	0.9259	0.9222	0.9296	1.8	11.9	8.9
128	1.0	0.9222	0.9370	0.9296	1.9	46.1	32.6
4096	1.0	0.8815	0.8667	0.8852	2.2	55,916.4	41,993.2
8192	0.9852	0.8185	0.7778	0.7926	2.4	254,830.9	202,822.9

combination of conventional HRV and wavelet entropy measures is used and classified using the genetic algorithm and the *k*-nearest neighbor classifier [14]. A classifier based on classification and regression tree method was developed and applied for congestive heart failure classification in Refs. [15, 16]. The sensitivity of 100.0% and the specificity of 89.74% were achieved.

However, when the nonlinear feature *SUM_IE* and its two components, i.e., *ApEn* and *SampEn*, are separately applied for the congestive heart failure classification, the corresponding classification performances are reduced. Furthermore, the computational times of three nonlinear features, i.e., *ApEn*, *SampEn*, and *SUM_IE*, are substantially higher than the computational times of the spectral exponents-based features. The computational times of spectral exponents of epochs of RR interval data with epoch size of 128 and 4096 beats providing the perfect classification performance are only 1.9 and 2.2 ms. The computational results thus suggest that the wavelet-based approach can be applied to a time series of RR intervals for an instant processing and also the spectral exponent-based features are an excellent quantitative feature for CHF classification.

5 Conclusions

In this study, the spectral exponent of RR interval data obtained using the wavelet-based approach is demonstrated to be, in general, an excellent measure used for

discriminating patients with congestive heart failure. Further, the minimum spectral exponent of RR interval data is shown to be a quantitative feature of spectral exponents of RR interval data that results in the perfect discrimination of congestive heart failure patients. The subjects with congestive heart failure can be perfectly discriminated from the subjects with normal sinus rhythm using a simple binary classification scheme with the single feature, i.e., the minimum spectral exponent of epochs of RR interval data. The epoch size of RR interval data can be as short as 128 beats or approximately 2 min. This therefore suggests that the diagnosis of congestive heart failure can be successfully achieved instantly using such feature.

The promising computational results obtained in this study can be further examined and applied to advance this field and also applications in medicine including cardiology. Even though the excellent computational results are obtained in this study, it will be applied to larger sets of RR interval data associated with larger groups of subjects in the future work. In particular, groups of subjects need to include all classes of congestive heart failure. Studies in the future work can examine various aspects of congestive heart failure. It will be examined to verify if features of spectral exponents can be used to specify a class of congestive heart failure or severity of congestive heart failure symptoms. Effects of physiological states on features of spectral exponents will also be examined. RR interval data recorded from a specific physiological state may lead to a better congestive heart failure discrimination. In addition, this may lead to a reduction of length of RR interval data needed for the congestive heart failure discrimination.

Acknowledgements This work was supported by the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission.

Compliance with Ethical Standards

Conflict of interest None.

References

1. The American Heart Association. (2012). About heart failure. http://www.heart.org/HEARTORG/Conditions/HeartFailure/AboutHeartFailure/About-Heart-Failure.
2. National Heart, Lung, and Blood Institute. (2015). What is heart failure? http://www.nhlbi.nih.gov/health/health-topics/topics/hf.
3. MedlinePlus. (2015). Heart failure. https://www.nlm.nih.gov/medlineplus/heartfailure.html.
4. Yancy, C. W., Jessup, M., Bozkurt, B., Butler, J., Casey, D. E., Drazner, M. H., et al. (2013). 2013 ACCF/AHA guideline for the management of heart failure: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation, 128*(16), e240–e327.
5. Go, A. S., Mozaffarian, D., Roger, V. L., Benjamin, E. J., Berry, J. D., Borden, W. B., et al. (2013). Heart disease and stroke statistics—2013 update: A report from the American Heart Association. *Circulation, 127*(1), e6–e245.
6. Mayo Clinic. (2016). Disease and conditions heart failure. http://www.mayoclinic.org/diseases-conditions/heart-failure/basics/definition/CON-20029801?p=1.
7. Hunt, S. A., Abraham, W. T., Chin, M. H., Feldman, A. M., Francis, G. S., Ganiats, T. G., et al. (2009). 2009 focused update incorporated into the ACC/AHA 2005 guidelines for the diagnosis and management of heart failure in adults: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines: Developed in collaboration with the International Society for Heart and Lung Transplantation. *Circulation, 119*(14), e391–e479.
8. The Criteria Committee of the New York Heart Association. (1994). *Nomenclature and criteria for diagnosis of diseases of the heart and great vessels* (9th ed.). Boston: Little & Brown.
9. Electrophysiology, Task Force of the European Society of Cardiology the North American Society of Pacing. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation, 93*(5), 1043–1065.
10. Liu, G., Wang, L., Wang, Q., Zhou, G. M., Wang, Y., & Jiang, Q. (2014). A new approach to detect congestive heart failure using short-term heart rate variability measures. *PLoS ONE, 9*(4), e93399.
11. Moore, R. K. G., Groves, D., Kearney, M. T., Eckberg, D. L., Callahan, T. S., Shell, W. E., et al. (2004). HRV spectral power and mortality in chronic heart failure (CHF): 5 year results of the UK heart study. *Heart, 90,* A6.
12. Guzzetti, S., Magatelli, E., Borroni, E., & Mezzetti, S. (2001). Heart rate variability in chronic heart failure. *Autonomic Neuroscience, 90,* 102–105.
13. Nolan, J., Batin, P. D., Andrews, R., Lindsay, S. J., Brooksby, P., Mullen, M., et al. (1998). Prospective study of heart rate variability and mortality in chronic heart failure: Results of the United Kingdom heart failure evaluation and assessment of risk trial (UK-Heart). *Circulation, 98,* 1510–1516.
14. Isler, Y., & Kuntalp, M. (2007). Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure. *Computers in Biology and Medicine, 37,* 1502–1510.
15. Pecchi, P., Melillo, L., Sansone, M., & Bracale, M. (2011). Discrimination power of short-term heart rate variability measures for CHF assessment. *IEEE Transactions on Information Technology, 15,* 40–46.
16. Melillo, P., Fusco, R., Sansone, M., Bracale, M., & Pecchia, L. (2011). Discrimination power of long-term heart rate variability measures for chronic heart failure detection. *Medical & Biological Engineering & Computing, 49,* 67–74.
17. Goldberger, A. L. (2006). Complex systems. *Proceedings of the American Thoracic Society, 3,* 467–472.
18. Mandelbrot, B. B. (1982). *The fractal geometry of nature*. San Fracisco: WH Freeman.
19. Wornell, G. W. (1995). *Signal processing with fractals: A wavelet-based approach*. Upper Saddle River, NJ: Prentice Hall.
20. Goldberger, A. L., Bhargava, V., West, B. J., & Mandell, A. J. (1985). On a mechanism of cardiac electrical stability: The fractal hypothesis. *Biophysical Journal, 48,* 525–528.
21. Havlin, S., Buldyrev, S. V., Goldberger, A. L., Mantegna, R. N., Ossadnik, S. M., Peng, C.-K., et al. (1995). Fractals in biology and medicine. *Chaos, Solitons & Fractals, 6,* 171–201.
22. Bak, P. (1997). *How nature works*. Oxford: Oxford University Press.
23. Barabasi, A. L., & Stanley, H. E. (1995). *Fractal concepts in surface growth*. Cambridge: Cambridge University Press.
24. Bassingthwaighte, J. B., Liebovitch, L. S., & West, B. J. (1994). *Fractal physiology*. New York: Oxford University Press.

25. Wornell, G. W. (1993). Wavelet-based representations for the 1/f family of fractal processes. *Proceedings of IEEE, 81,* 1428–1450.
26. Wornell, G. W. (1991). *Synthesis, analysis, and processing of fractal signals*. Ph.D. Thesis, Massachusetts Institute of Technology, Massachusetts.
27. Abry, P., Goncalves, P., & Flandrin, P. (1993). Wavelet-based spectral analysis of 1/*f* processes. In IEEE international conference on acoustics, speech, and signal processing (pp. III-237–III-240).
28. Goldberger, A. L., Amaral, L. A., Hausdorff, J. M., Ivanov, P. Ch., Peng, C.-K., & Stanley, H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences of the United States of America, 99,* 2466–2472.
29. Janjarasjitt, S. (2014). Computational validation of fractal characterization by using the wavelet-based fractal analysis. *Journal of the Korean Physical Society, 64,* 780–785.
30. Peng, C.-K., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos, 5,* 82–87.
31. Peng, C.-K., Mietus, J., Hausdorff, J. M., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1993). Long-range anticorrelations and non-Gaussian behavior of the heartbeat. *Physical Review Letters, 70,* 1343–1346.
32. Janjarasjitt, S., & Loparo, K. A. (2015). Examination of scale-invariant characteristics of multi-channel ECoG data for epileptic seizure localization. *Journal of Medical and Biological Engineering, 35,* 278–284.
33. Janjarasjitt, S. (2011). Wavelet-based fractal analysis of sleep EEG. In *ICICS2011—8th international conference on information, communications and signal processing*.
34. Janjarasjitt, S. (2015). Spectral exponent characteristics of intracranial EEGs for epileptic seizure classification. *IRBM, 36,* 33–39.
35. Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. Ch., Mark, R. G., et al. (2000). Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation, 101*(23), e215–e220.
36. Watters, P. A. (1998). Fractal structure in the electroencephalogram. *Complexity International, 5,* 1–8.
37. Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27,* 861–874.