Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 145 (2018) 97–104

Procedia Computer Science

www.elsevier.com/locate/procedia

Postproceedings of the 9th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2018 (Ninth Annual Meeting of the BICA Society)

Efficient Methods for Calculating Sample Entropy in Time Series Data Analysis

Ronakben Bhavsar[a,*], Na Helian[a], Yi Sun[a], Neil Davey[a], Tony Steffert[b], David Mayor[a]

[a] *University of Hertfordshire, Hatfield, AL10 9AB, UK*
[b] *The Open University, Milton Keynes, MK7 6AA, UK*

Abstract

Recently, different algorithms have been suggested to improve Sample Entropy (SE) performance. Although new methods for calculating SE have been proposed, so far improving the efficiency (computational time) of SE calculation methods has not been considered. This research shows such an analysis of calculating a correlation between Electroencephalogram(EEG) and Heart Rate Variability(HRV) based on their SE values. Our results indicate that the parsimonious outcome of SE calculation can be achieved by exploiting a new method of SE implementation. In addition, it is found that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in a time domain.

Keywords: Sample Entropy; EEG; HRV; Time Series Data Analysis; Pearson Correlation.

1. Introduction and Related Work

Nonlinear dynamic analyses have been widely used to study the complex behaviours and different structures of biological systems [2]. Nonlinear dynamic analysis proves to be a powerful approach for the assessment of different physiological time series as it can determine the hidden patterns related to underlying mechanisms [3] [13]. The chaotic behaviour of cardiac system and brain waves indicate nonlinearity [1]. With the given nature of nonlinearity, Electroencephalogram (EEG) and Heart Rate Variability (HRV) turn out to be appropriate for nonlinear time series analysis [1]. The different types of nonlinear complex measures of variability are Lyapunov exponent, Correlation Dimension D2, Approximate Entropy (AE), Sample Entropy (SE), Multiscale entropy (MSE), Poincare plots, Detrended Fluctuation Analysis (DFA) and many more.

* Corresponding author. Tel.: +44-757-436-9688.
E-mail address: r.bhavsar2@herts.ac.uk

1877-0509
10.1016/j.procs.2018.11.016

SE has been used widely to investigate various biological conditions in human body like arrhythmia studied through ECG (Electrocardiogram) [3], Alzheimer′s patients' EEG background activity [1], analyzing human postural sway data [12] and studying HRV in the case of obstructive sleep apnoea syndrome [2]. SE is also used to detect the termination of a particular medical condition like seizures [15] and to test the effect of a therapy like ketogenic diet used for controlling intractable seizures [14]. These studies have concluded that SE is robust quantifier of complexity, which offers an accurate nonlinear metric for quantification [3]. It gives a good dynamical signature and is a helpful tool that provides insights into various biological time series [1],[12]. Therefore, SE is considered as an effective method for investigating different types of time series data.

In recent years, different algorithms attempting to improve SE have been proposed. Quadratic Sample Entropy (QSE) was introduced to reduce the influence of arbitrary constants of sequence comparision and tolerance on SE, as well as to reduce the skewing of results when either the top or the bottom of the conditional probabilities was very small or very large [9]. Another attempt to improve SE was with the introduction of Fuzzy Entropy (FuzzyEn) [5], using the concept of fuzzy sets in order to determine a fuzzy measurement of similarity of two vectors based on their shapes. Multi Scale Entropy (MSE) established by [14], was a useful extension of SE to multiple time scales, in recognition of the likelihood that dynamical complexity of biological signals may operate across a range of temporal scales.

In this work, the type of nonlinear complex measures of variability exploited is Sample Entropy (SE). The aim of the research is not to propose another new method derived from SE, but efficient method improving the computational time for the SE calculation. The computational time for SE calculation using the new and original SE methods will be compared on calculating the correlation between SE values of EEG and HRV in time domain.

2. Dataset Information

Our datasets consist of EEG and ECG recordings from 15 participants. This data was obtained over 5 minute time slots in a relaxed state with eyes opened. 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) for EEG recording were used, following the standard 10%-20% system [7], as shown in Fig.1. The values of 10% and 20% shown in Fig.1. refer to the distances between adjacent electrodes: either 10% or 20% percentage of the total front-to-back or right-to-left distance over the skull - front-to-back distance is based on the measurement from the Nasion (point between forehead and nose) to Inion (lowest point of the skull from the back of the head indicated by a prominent bump), and right-to-left distance is based on the measurement between the left and right pre-auricular points.

The sampling rate used for EEG was 250Hz, and the reference was linked to ear electrodes. For ECG data, one electrode was positioned on the volar surface of each forearm (with an additional electrode as ground on the dominant side) to record the electrical activity of heart over time, and the sampling rate was 256Hz.

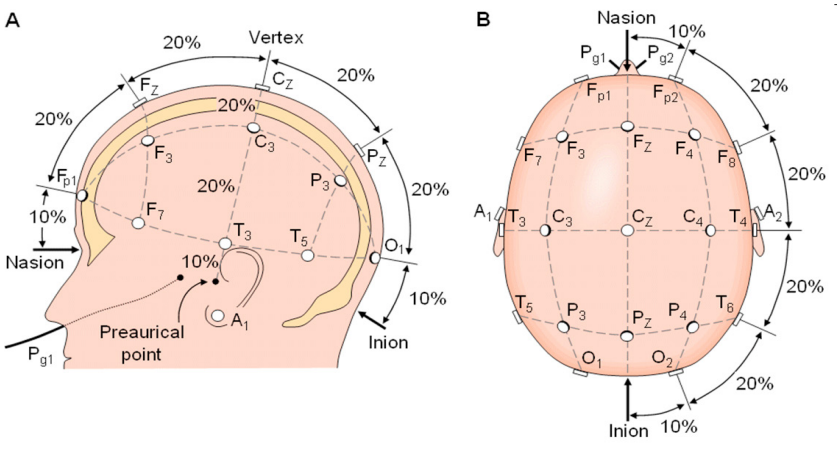

Figure 1. The international 10%-20% system seen from A (left side of the head) and B (above the head). The letter F, T, C, P, O, A, Fp and Pg stands for frontal, temporal, central, parietal, occipital, earlobes, frontal polar, and nasopharyngeal, respectively. The figure is obtained from [7].

3. Sample Entropy and Proposed Implementation

SE is considered as an effective method for investigating different types of time series data. A lower SE value indicates a high frequency of similarity in time series [13].

For a time series of length n, $SampleEntropy(m,r,n)$ can be defined as the negative logarithm of conditional probability that two sequences are similar for m point [10] within a tolerance value r, excluding any self-matches [13]. The equation can be represented according to [13], as:

$$SE(m,r,n) = -\ln\left(\frac{A}{B}\right), \tag{1}$$

where, m is the length of sequences to be compared , r is the tolerance value for accepting matches, n is the length of original data, and A and B are defined as follow:

$$A = \left\{\frac{[(n-m-1)(n-m)]}{2}\right\}A^m(r), \text{ and } B = \left\{\frac{[(n-m-1)(n-m)]}{2}\right\}B^m(r)$$

where, $A^m(r)$ is the probability that the two sequences match for $m+1$ points, and $B^m(r)$ is the probability that the two sequences match for m points. Each SE value indicates relative consistency with respect to any value of (m,r). That is, if a record has a lower SE value than another record for a part of fixed m and r values, it will be lower for any part of fixed m and r values [10]. SE is independent of the data length and shows an elimination of self-matching. In order to approximate the conditional probabilities of matches, SE uses a point-wise approach [13].

3.1. An Example of SE Calculation

In this section, we explain how the SE is calculated in practise by giving an example with a simple time series. Let the input time series be $x(n) = \{0.1, 0.1, 0.2, 0.5, 0.22\}$, with $m = 2$, $r = 0.2$, $n = 5$.

The value of m specifies the length of sequences to be considered for SE. Thus, the default value of m is 2(i.e. the maximum length of sequence considered is 2). The value of r represents the tolerance value below that a match is deemed. The input point sequence for $A^m(r)$ is n points, while $B^m(r)$ considers $n-1$ points of the input sequences. That means, for $A^m(r)$ and $B^m(r)$ the input point sequence is $\{0.1, 0.1, 0.2, 0.5, 0.22\}$ and $\{0.1, 0.1, 0.2, 0.5\}$, respectively.

As shown in Table 1, to calculate the probability for A and B, the number of matches obtained for respective sequences of $A^m(r)$ and $B^m(r)$ is counted as "1". For a particular length of sequence m, the point matches are calculated by calculating the absolute difference between the points in the sequences. The difference should be below the tolerance value r (in this case 0.2). The calculation of similar segments can be summarized as: $|x(i) - x(j)| < r$. Considering the following sequence $(x_k(i), x_k(j)) = [(0.1, 0.1), (0.2, 0.5)]$, where i and j are the point sequence , and k is the index for these point sequences. To test the match, $(|0.1 - 0.2|, |0.1 - 0.5|) = (0.1, 0.4)$ is calculated. It can be observed that $x_1(i)$ and $x_1(j)$ (i.e. 0.1 and 0.2) satisfy the condition but $x_2(i)$ and $x_2(j)$ (i.e. 0.1 and 0.5) do not satisfy the condition because the absolute difference is greater than the tolerance value r. Since the point sequence is not a complete match under the tolerance value r, this sequence is not considered as a match.

Table 1 represents the point sequence match at a given length of sequence for $m(0$ to $2)$ for the tolerance value $r(0.2)$. In third and fourth columns, "1" represents a match and "0" represents no match at tolerance value r.

According to equation (1), SE value can be calculated as follows:

SE (0, 0.2, 5) = p (0) = − ln (A[0]/((n*n-1)/2)) = − ln(6/10) = 0.5108

SE (1, 0.2, 5) = p (1) = − ln (A[1]/B[0])) = − ln(1/3) = 1.0986

SE (2, 0.2, 5) = p (2) = − ln (A[2]/B[1])) = − ln(0/0) = inf

From the SE values obtained from the above examples, it can be seen that a low SE value is obtained at $m = 0$, which increases with the increase of m. This indicates that for a longer point sequence, the similarity decreases for this time series.

Table 1. Point sequences at $m = 0; 1;$ and 2 along with the count of match obtained for $A^m(r)$ and $B^m(r)$. Here, X represent that the point sequence was not considered for $B(m)$. Columns $A(m)$ and $B(m)$ indicates count for the total number of matches obtained .

Length of sequence for m	**Point Sequences**	**Point Matches at** $r = 0.2$			
		$A^m(r)$	$B^m(r)$	$A(m)$	$B(m)$
m=0	[0.1, 0.1]	1	1	6	3
	[0.1, 0.2]	1	1		
	[0.1, 0.5]	0	0		
	[0.1, 0.22]	1	X		
	[0.1, 0.2]	1	1		
	[0.1, 0.5]	0	0		
	[0.1, 0.22]	1	X		
	[0.2, 0.5]	0	0		
	[0.2, 0.22]	1	X		
	[0.5, 0.22]	0	X		
m=1	[(0.1, 0.1),(0.1, 0.2)]	(1, 1)	(1, 1)	1	1
	[(0.1, 0.1),(0.2, 0.5)]	(1, 0)	(1, 0)		
	[(0.1, 0.1),(0.5, 0.22)]	(1, 0)	X		
	[(0.1, 0.2),(0.2, 0.5)]	(1, 0)	(1, 0)		
	[(0.1, 0.2),(0.5, 0.22)]	(1, 0)	X		
	[(0.2, 0.5),(0.5, 0.22)]	(0, 0)	(0, 0)		
m=2	[(0.1, 0.1, 0.2),(0.1, 0.2, 0.5)]	(1, 1, 0)	(1, 1, 0)	0	0
	[(0.1, 0.1, 0.2),(0.2, 0.5, 0.22)]	(1, 0, 1)	X		
	[(0.1, 0.2, 0.5),(0.2, 0.5, 0.22)]	(1, 0, 0)	(1, 0, 0)		

3.2. Efficient and Parsimonious way for Sample Entropy Calculation

SE measures the probabilities of matches for a time series data using point-wise approach. This can be time consuming when long sequence of points need to be compare, and can be done more efficiently. Computation time for SE can be reduced without losing much information from the signals by using the three methods proposed in this section. Thus, calculation time for SE could be shortened and the computational expense would be more cheaper. Fig.2. illustrates on how these three methods work.

SE-Method 1 is about shortening the time series signal without loss of too much information for the point-wise approach. For example, instead of considering the original data length(n) of the 5-minute signal (=75000 data points), it could be shortened by dividing 1.1 on the original data length (75000/1.1=68181 equivalent to 28 seconds data points). Binary chop [8] is performed in order to find out at which point the most accurate result for SE could be obtained.

SE-Method 2 is about SE calculation on a moving window, calculating SE on individual windows to find out which window gives the SE values that are most similar to the original SE value. For example, using a 2 seconds moving window, SE is calculated for a window size of 500 points (2*(250Hz) = 500 data points).

SE-Method 3 is to calculate the mean for a given window first before performing SE. This data window could be as long as a minute or as short as a second. For example, if the mean of each 1 sec data (250 points) is gathered, then it will give us a reduced length of $n = 300$ data points on which to perform SE calculation, and not $n = 75000$. This way the SE computational time should be reduced dramatically.

4. Experimental Results

The EEG signals were pre-processed to remove artefacts caused by electrical activity in muscles, including of the eye, jaw and other muscle movements using Independent Component Analysis (ICA) as mentioned in [6]. It is relatively straightforward to remove these artefacts using ICA.

To extract HRV from ECG signals, the method designed by Lin et al. [11] was adopted. The results of the automatic analysis were reviewed and then any errors in ECG R-wave detection and QRS complex (combination of three graph-

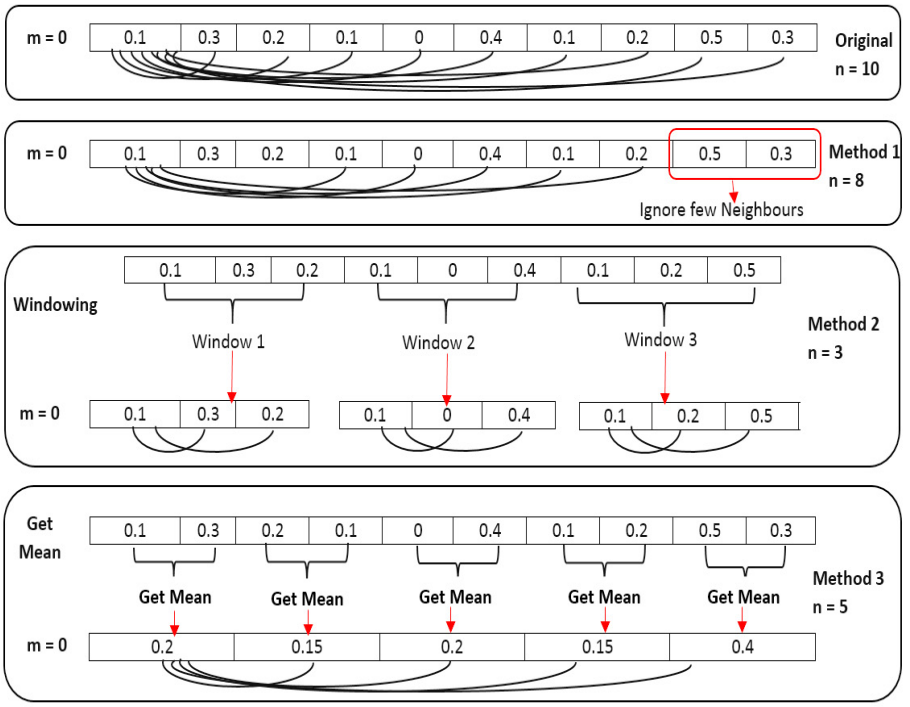

Figure 2. An example of how the SE can be calculated efficiently.

ical deflections: Q, R and S waves) labelling were removed manually. R-R interval data obtained from the edited time sequence of R-wave and QRS labelling was then transferred to a personal computer. In order to remove artefacts from extracted HRV signal, each R-R interval was compared against a local average interval. If a R-R interval differs from the local average more than a specified threshold (in seconds) value, then that R-R interval is defined as an artefact and is replaced with an interpolated value using a cubic spline interpolation.

4.1. Experiments using three proposed SE Calculation Methods

For each of the five minutes of EEG data; the following three experiments have been undertaken, results are shown in Table 2. For the purpose of comparison, SE values of the original SE performance is also shown. All code is run on a personal computer: Windows 7 Enterprise, Intel(R) Core (TM) i7-3770T, 64-bit Operating System, 16 GB RAM.

Experiment 1 is the implementation of SE-Method 1, by restricting number of neighbours for comparisons on SE calculation. It is found that ignoring the last 25 to 30 seconds of data still achieves as accurate results as if they are included, but with improving computational time by 10 seconds.

Experiment 2 is about experimenting SE-Method 2, 10 different window size are considered (i.e. 2 Sec, 10 Sec, 20 Sec, 30 Sec, 40 Sec, 50 Sec, 60 Sec, 70 Sec, 80 Sec and 90 Sec windows) on which to perform the SE calculation, to find out which window gives the SE values that are most similar to the original SE value, as shown in the Table 2. It is found that the smaller the window size, the shorter the calculation time. In addition, the most similar results to the original SE calculation results is the smallest window size.

Experiment 3 is demonstrating SE-Method 3, calculating the mean of each window of a second of data (i.e 1 Sec= mean of 250 points). The experiment is done with 8 different window sizes (i.e. 0.06 Sec, 0.12 Sec, 0.25 Sec, 0.55 Sec, 1 Sec, 2 Sec, 3 Sec and 4 Sec) on which to calculate the mean, as shown in Table 2. The SE is then performed on the mean values of the signal. It is found that bigger the window size, the shorter the calculation time. Moreover, the best match to the original SE calculation results is at the mean of each 1 seconds window.

Experiment 1-3 demonstrated a strong positive correlation between the results obtained using the original and each of the new three SE approaches with the correlation values of 0.99, 0.68, and 0.96 for SE-Method 1, SE Method 2 and

Table 2. Computation time for the SE calculation using the Original approach and our three Experimental methods.

Experiments	Details	Computation Time
SE-Original	Original Performance	75 Sec
SE-Method 1	Shortening the neighbour comparison	62 Sec
SE-Method 2	2 Seconds Moving Window	0.002 Sec
	10 Seconds Moving Window	0.08 Sec
	20 Seconds Moving Window	0.38 Sec
	30 Seconds Moving Window	0.72 Sec
	40 Seconds Moving Window	1.30 Sec
	50 Seconds Moving Window	1.94 Sec
	60 Seconds Moving Window	3 Sec
	70 Seconds Moving Window	4 Sec
	80 Seconds Moving Window	6 Sec
	90 Seconds Moving Window	9 Sec
SE-Method 3	Mean of Each 0.06 Seconds Window	0.38 Sec
	Mean of Each 12 Seconds Window	0.11 Sec
	Mean of Each 25 Seconds Window	0.02 Sec
	Mean of Each 50 Seconds Window	0.007 Sec
	Mean of Each 1 Seconds Window	0.003 Sec
	Mean of Each 2 Seconds Window	0.008 Sec
	Mean of Each 3 Seconds Window	0.01 Sec
	Mean of Each 4 Seconds Window	0.02 Sec

SE-Method 3, respectively, along with the probability of 0. Whilst SE-Method 1 and SE-Method 2 do not improve the trend and the computational time for SE calculation, SE-Method 3 clearly works best because it provides the most predictive value, and trend for SE performance to those provided by original SE performance with improving computational time.

4.2. Experiment 4

The aim of experiment 4 is to compare the performance of new methods and original methods for SE calculation. The previous three experiments shows that SE-Method 3 is the best one in terms of improving SE calculation time without losing much information. Hence, Only SE-Method 3 is considered for this experiment because correlation coefficient works on similar length of the signals, and SE-Method 3 gives us the same length of samples for EEG as of HRV. In order to demonstrate correlation between EEG and HRV, for each of the five minutes of EEG and HRV data; the following steps have been undertaken for both original and the new approach (SE-Method 3) of SE calculation.

1. For each electrode's data, divide data into 10 equal bins to perform SE calculation. This process has been repeated for each participant.
2. Compute correlation coefficients on the 10 SE values of the EEG and 10 SE values of the HRV obtained in step 1.

Pearson′s correlation is used to perform the correlation coefficients. It measures how closely two different observables are related to each other. Pearson′s correlation co-efficient R ranges between 1 (when the matching entities are exactly the same) and -1 (when the matching entities are inverses of each other). A value of zero indicates no relationship existing between the entities.

Once the electrode's SE correlation performance are gathered, the best performance electrodes have been ranked-where, the ranking has been given based on electrode correlation values, the bigger the value, the higher the rank. The three best performance electrodes' results have been looked closely. Some common electrode rankings are found for all the participants investigated, highlighted in yellow colour, as shown in the Fig.3.

It is found that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in time domain. Moreover, the new approach (SE-Method 3) of SE calculation is giving more focused result than the original SE calculation.

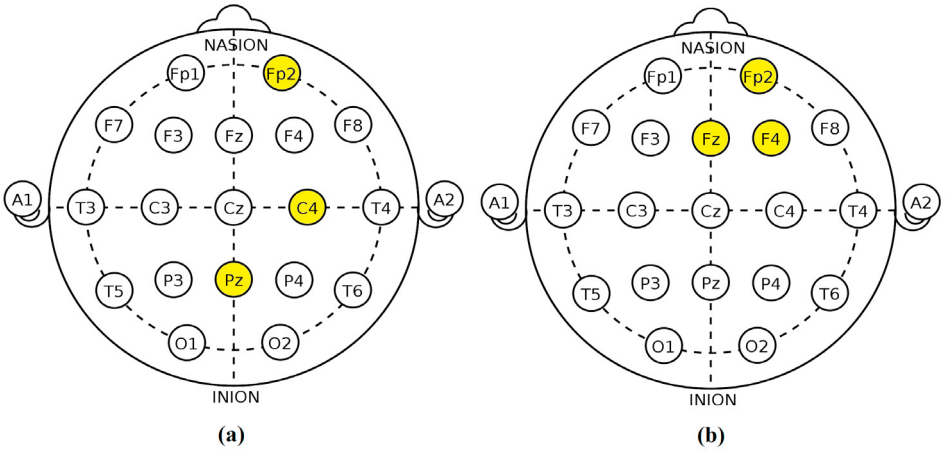

Figure 3. Electrode Ranking based on correlation performance between SE values of EEG and HRV, showing best three performing electrodes across participants, highlighted in yellow colour: (a) Ranking based on the original approach for SE calculation, and (b) Ranking based on the new approach (SE-Method 3) for SE calculation.

5. Discussion and Conclusion

The main conclusion of this work is that parsimonious results for SE can be achieved using the proposed new methods of pre-processing the data prior to SE calculation. SE-Method 3 clearly works best for improving the performance because it gives good predictive values without changing the trends visible in SE calculated using the original standard approach. SE-Method 1 provides SE values very close to those obtained using the original SE approach, but it does not improve computational time much. Similarly, SE-Method 2 is not robust because neither the trends nor computational time are improved significantly.

The second conclusion from this work is that there is a strong positive correlation (R=0.96, Probability = 0) between results obtained using the original and the new (SE-Method 3) SE approaches. Also, we found low positive correlations between SE values of EEG and HRV in the time domain. The results shown in our previous work suggested that the electrical activity in the frontal lobe of the brain is correlated with the HRV. It shows that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in time domain, which is in consistant with our previous finding on frequency domain in paper [4].

In summary, SE-Method 1 and SE-Method 2 do not improve the trend or the computational time for SE calculation. SE-Method 3 does not give values similar to those provided by the original SE approach, but it does provide the most predictive value for SE performance. Although the result is not exactly similar as the original SE performance, the trend is. Therefore, we believe the most efficient way for SE calculation is SE-Method 3.

References

[1] D Abásolo, R Hornero, P Espino, D Alvarez, and J Poza. Entropy analysis of the eeg background activity in alzheimer's disease patients. *Physiological measurement*, 27(3):241, 2006.

[2] Haitham M Al-Angari and Alan V Sahakian. Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. *Biomedical Engineering, IEEE Transactions on*, 54(10):1900–1904, 2007.

[3] Raul Alcaraz and Jose Joaquin Rieta. A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. *Biomedical Signal Processing and Control*, 5(1):1–14, 2010.

[4] Ronakben Bhavsar, Neil Davey, Yi Sun, and Na Helian. An investigation of how wavelet transform can affect the correlation performance of biomedical signals. 2018.

[5] Weiting Chen, Zhizhong Wang, Hongbo Xie, and Wangxin Yu. Characterization of surface emg signal based on fuzzy entropy. *IEEE Transactions on neural systems and rehabilitation engineering*, 15(2):266–272, 2007.
[6] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.
[7] George H Klem, Hans Otto Lüders, HH Jasper, C Elger, et al. The ten-twenty electrode system of the international federation. *Electroencephalogr Clin Neurophysiol*, 52(3):3–6, 1999.
[8] Donald E Knuth. Sorting and searching, 2nd edn. the art of computer programming, vol. 3, 1998.
[9] DE Lake. Improved entropy rate estimation in physiological data. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 1463–1466. IEEE, 2011.
[10] Douglas E Lake, Joshua S Richman, M Pamela Griffin, and J Randall Moorman. Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 283(3):R789–R797, 2002.
[11] Che-Wei Lin, Jeen-Shang Wang, and Pau-Choo Chung. Mining physiological conditions from heart rate variability analysis. *IEEE Computational Intelligence Magazine*, 5(1):50–58, 2010.
[12] Sofiane Ramdani, Benoît Seigle, Julien Lagarde, Frédéric Bouchara, and Pierre Louis Bernard. On the use of sample entropy to analyze human postural sway data. *Medical engineering & physics*, 31(8):1023–1031, 2009.
[13] Joshua S Richman and J Randall Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6):H2039–H2049, 2000.
[14] Tetsuya Takahashi, Raymond Y Cho, Tomoyuki Mizuno, Mitsuru Kikuchi, Tetsuhito Murata, Koichi Takahashi, and Yuji Wada. Antipsychotics reverse abnormal eeg complexity in drug-naive schizophrenia: a multiscale entropy analysis. *Neuroimage*, 51(1):173–182, 2010.
[15] Cheol Seung Yoo, Dong Chung Jung, Yong Min Ahn, Yong Sik Kim, Su-Gyeong Kim, Hyeri Yoon, Young Jin Lim, and Sang Hoon Yi. Automatic detection of seizure termination during electroconvulsive therapy using sample entropy of the electroencephalogram. *Psychiatry research*, 195(1):76–82, 2012.