Journal of Medical Systems (2018) 42: 109
https://doi.org/10.1007/s10916-018-0961-2

IMAGE & SIGNAL PROCESSING

A Novel Feature Level Fusion for Heart Rate Variability Classification Using Correntropy and Cauchy-Schwarz Divergence

Ateke Goshvarpour[1] · Atefeh Goshvarpour[1,2]

Received: 30 January 2018 / Accepted: 16 April 2018 / Published online: 30 April 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Heart rate variability (HRV) analysis has become a widely used tool for monitoring pathological and psychological states in medical applications. In a typical classification problem, information fusion is a process whereby the effective combination of the data can achieve a more accurate system. The purpose of this article was to provide an accurate algorithm for classifying HRV signals in various psychological states. Therefore, a novel feature level fusion approach was proposed. First, using the theory of information, two similarity indicators of the signal were extracted, including correntropy and Cauchy-Schwarz divergence. Applying probabilistic neural network (PNN) and k-nearest neighbor (kNN), the performance of each index in the classification of meditators and non-meditators HRV signals was appraised. Then, three fusion rules, including division, product, and weighted sum rules were used to combine the information of both similarity measures. For the first time, we propose an algorithm to define the weights of each feature based on the statistical *p*-values. The performance of HRV classification using combined features was compared with the non-combined features. Totally, the accuracy of 100% was obtained for discriminating all states. The results showed the strong ability and proficiency of division and weighted sum rules in the improvement of the classifier accuracies.

Keywords Feature level fusion · Correntropy · Cauchy-Schwarz divergence · Heart rate variability · Meditation

Introduction

The variations in R-R intervals are called heart rate variability (HRV), which is usually served as an index of the autonomic nervous system activities. Many applications have been introduced to analyze the heart signal in medicine. HRV analysis not only has received much attention in detection, prediction, classification, or even treatment of cardiovascular diseases but also for the understanding of the psychological disorders and states [1–3]. To address this goal, several computer-aided algorithms recruited for offline/online monitoring of HRV. Generally, in the design of a typical HRV monitoring procedure, several components may be incorporated, including data acquisition step, pre-processing the data, feature extraction algorithms, feature reduction, and classification (pattern classification or recognition) modules. However, it is noted that except for the data acquisition step, a scheme may comprise of one or more component, but not necessarily all.

Depending on the utilization, different criteria are considered in designing algorithms. For instance, in offline computations, usually a high precision algorithm is contemplated, and in others, specifically in online applications, benchmarks also focus on the computational speed. In some literature, more attention has been placed on the assumption of the non-stationarity of the bio-signals [4, 5]. This category of research highlights the use of nonlinear techniques because it is believed that using traditional linear approaches may fail in demonstrating the important info. To date, several nonlinear indices have been introduced, some inspired by chaos theory or predictability of the signal, and the others based on the statistical methodologies. Correntropy (corrEn) was recently introduced [6] as a new nonlinear index to assess the similarity of a time-series. This approach integrates time structure and

This article is part of the Topical Collection on *Image & Signal Processing*

✉ Atefeh Goshvarpour
af_goshvarpour@sut.ac.ir

Ateke Goshvarpour
ak_goshvarpour@sut.ac.ir

[1] Department of Biomedical Engineering, Faculty of Electrical Engineering, Sahand University of Technology, Tabriz, Iran

[2] Khayyam Boulevard, Khayyam 20, Number 54, Mashhad, Khorasan Razavi, Iran

statistical distribution of the signal. Cauchy-Schwarz divergence (CSD) is another similarity index [7] which quantifies the difference between the information content of the data. The commonality of these two algorithms (corrEn and CSD) is that both utilize information theoretic descriptors. Usually, a kernel-based approach is applied to both measures. It is used to handle the harmful effects of the outliers and improve the accuracy. Eliminating the outliers and applying the useful data, enables a faster convergence.

Recently, meditation attracted the scientific attention of researchers extensively. Therefore, physiological effects of meditation techniques on the circulatory system have been scrutinized using HRV signals. Earlier, HRV dynamics during meditation were explored using nonlinear features, including Poincare plots [8, 9], fractal scaling [10], multi-fractal analysis [11], dynamical complexity [12], recurrence plot analysis [13], Hilbert transform [14], higher order spectra [15], Lyapunov exponents [16, 17], and some entropy measures [16, 18]. Though, there is not any published article, exploring the differences in HRV fluctuations similarity in the normal state and meditation practices.

In this paper, two similarity indices were inspected to examine the resemblance of the HRV signal in two different meditation techniques. One of them utilizes a probability density function (PDF) and the other applies a nonlinear approach. The former is known as CSD and the latter is recognized as "corrEn". Since the measures are strongly dependent on the kernel bandwidth, the role of choosing kernel sizes was also evaluated in this study. In addition, the proficiency of each similarity index in the HRV classification is appraised. We also examine the ability of feature level fusion in classification performances. To this end, three fusion rules, including division, product, and weighted sum rules are proposed. For the first time, we propose a weighted sum rule, which recruits the *p*-values of the statistical test to define the weights.

Material and methods

HRV recordings during spontaneous normal breathing (SNB), and pre/during the meditation practices, including Chinese Chi meditation (CCM) and Kundalini Yoga meditation (KYM) available at the Physionet database were studied [14]. Two similarity indices were extracted to characterize HRV signals (Feature Extraction). A novel feature level fusion approach was introduced and the features combined using three different fusion rules. The normalized combined/non-combined features were input the classifier to separate psychological states. The effect of parameterization of the features and the classifier was considered. The total procedure is diagrammed in Fig. 1.

Data selection

We studied the HRV signals of healthy meditators and non-meditators available in Physionet database [14, 19]. The meditation techniques comprise KYM (as trained by Yogi Bhajan) and CCM.

Wearing a Holter, the electrocardiogram (ECG) signal recordings of four well qualified Kundalini yoga performing group were obtained for about one and a half hours. A 15 min ECG signal was recorded primarily as pre KYM (pKYM). Then, successive breathing and chanting exercises were done during the Kundalini yoga practice (dKYM), while participants sit in a cross-legged posture.

HRV recordings of CCM were accessible in two states pre and during the CCM performances (pCCM and dCCM). Most of the participants were skilled in meditation and having started the CCM for about one to three months before the trial. Wearing a Holter monitor, each of participants done one hour of CCM. Sitting and listening to the taped instructions of the master, the volunteers breathe spontaneously while imagining the opening and closing of a perfect lotus in their stomach.

HRV signals of non-mediating control group accessible in Physionet database [14, 19] was employed. The participants of this group comprise of 11 healthy subjects which breathe spontaneously during sleep hours. The general health level of the groups was comparable [14]. The sampling rate was fixed at 128 Hz.

Participant information is summarized in Table 1.

Feature extraction

Correntropy

A nonlinear mapping of the signal to the kernel space provides a similarity parameter that is called corrEn [20]. Applying corrEn, the data is first transformed into a high-dimensional reproducing kernel Hilbert space and then the "conventional" correlation is computed. This feature takes advantage of the statistical distribution and the time structure of the data, simultaneously. Precisely, corrEn is capable of preserving nonlinear specifications and higher order data moments [21]. Additionally, it is robust in dealing with impulsive noise and its computational complexity is low [22]. These benefits make it an appropriate signal processing technique for characterizing nonlinear dynamics [23]. Previously, this method has been successfully used as a nonlinear method for analyzing biosignals [24, 25]. It was shown that corrEn is linked with the regularity of HR signals during sleep disturbances [25].

Considering two time-series $X \in R^{M\times N}$; $X=[x_1, x_2, \Lambda, x_N]$ and $Y \in R^{M\times N}$; $Y=[y_1, y_2, \Lambda, y_N]$. M denotes the dimension and N represents the sample number. CorrEnt is defined as Eq. (1):

$$V(X,Y) = E[< \Phi(X), \Phi(Y) >] = E[\kappa_\sigma(X,Y)] \quad (1)$$

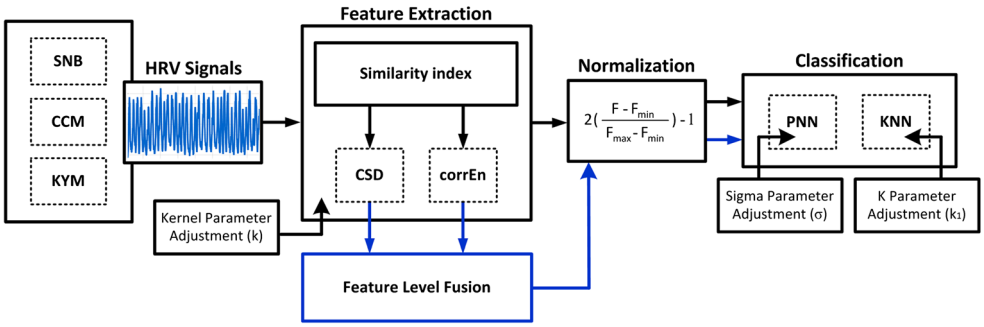


Fig. 1 The proposed methodology

where a radial kernel function is shown by κ_σ, in which σ is a kernel size (bandwidth). A nonlinear mapping is introduced by Φ, which maps data to a high dimensional kernel Hilbert space. E represents the mean operator and $<\cdot,\cdot>$ is the inner product. As a result, the corrEn can be rewritten as follows:

$$V(X,Y) = \frac{1}{N}\sum_{i=1}^{N}\kappa_\sigma(x_i,y_i) \tag{2}$$

The most common kernel function in corrEn is the Gaussian kernel, which is formulated as (3):

$$\kappa_\sigma(x_i,y_i) = \exp\left(\frac{\|x_i - y_i\|^2}{2\sigma^2}\right) \tag{3}$$

where $\|\cdot\|$ denotes Euclidean norm.

In this study, the kernel size (k) was set in the range of 1 to 10 with the step size of 1; k = 1, 2, 3, …, 10.

The Cauchy-Schwarz divergence

Consider the probability density functions $f(X)$ and $g(Y)$ for two time-series X and Y. Basically, CSD is derived from the Cauchy-Schwarz inequality [26] for inner products and calculates the distance between the density functions. It also recruited as a similarity index. The Cauchy–Schwarz divergence is effortlessly computable. It is affine transformations invariant in terms of input data transformation [27]. Earlier, for examining the time series similarity, an algorithm was presented that used Cauchy–Schwarz divergence [28]. The algorithm was tested on various databases, including the heart signals of arrhythmia patients. The usefulness of the proposed algorithm in the databases with varying length was shown by its high speed and high accuracy.

Equation (4) shows the CSD formulization:

$$CSD(f,g) = -\ln\frac{< f,g >}{\|f\|\|g\|} \tag{4}$$

Since the probability densities are in the range of zero and one, the logarithm argument is non-negative and its maximum value is 1. If f=g, then CSD = 0, else it is a positive value. A few problems occur with such an approach: (1) the signals are discrete, so the actual densities (f and g) is not clear. (2) The computational cost of statistical density estimators is very high because they require rather large sample sizes. To overcome these problems, usually, some kind of density estimation is performing [29].

In this study, the estimation of both corrEn and CSD was realized using incomplete Cholesky decomposition [30]. In addition, different kernel sizes (k) were evaluated (k was in the range of 1 to 10 with the step size of 1; k = 1, 2, 3, …, 10).

Feature level fusion

Feature level fusion approaches were adopted in this study. Consider a feature vector of **A** for corrEn measures and **B** for CSD. Three different fusion rules were examined as follows.

1- Product rule.

$$\mathrm{Pr}\_f = \mathbf{A} \times \mathbf{B} \tag{5}$$

2- Division rule.

$$Dr\_f = \mathbf{A}/\mathbf{B} \tag{6}$$

Table 1 Subject characteristics in groups of meditators and non-meditators

	Kundalini yoga meditation pKYM & dKYM	*Chinese Chi meditation* pCCM & dCCM	*Spontaneous nocturnal breathing (SNB)*
Number of subjects	4	8	11
Age range	20–52	26–35	20–35
Mean age	33	29	29
Gender	2 women, 2 men	5 women, 3 men	8 women, 3 men

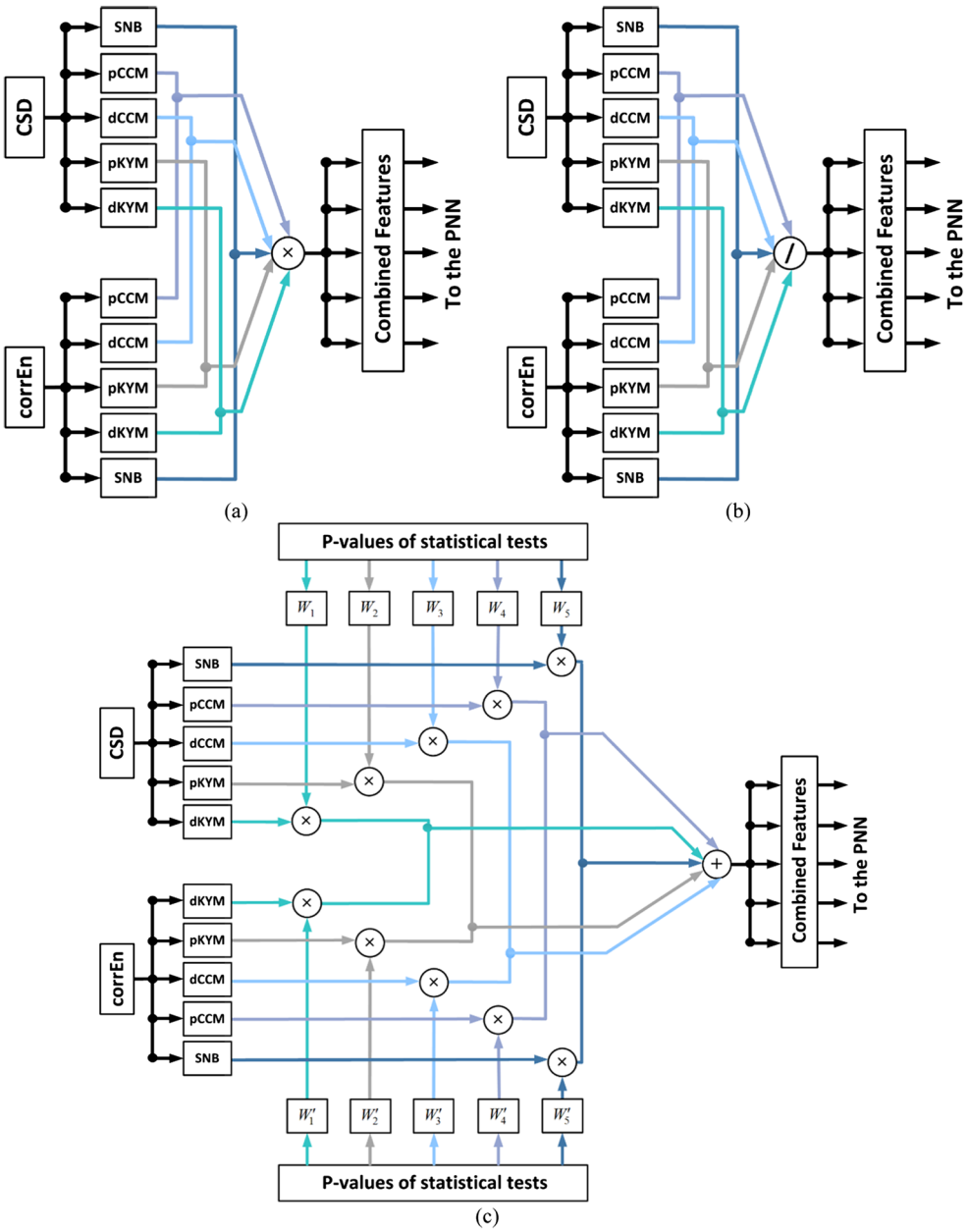


Fig. 2 Feature level fusion rules (**a**) product; (**b**) division; and (**c**) weighted sum

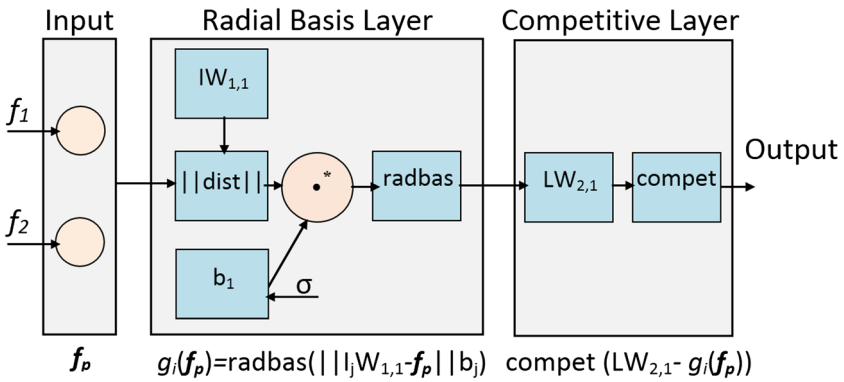


Fig. 3 The architecture of PNN

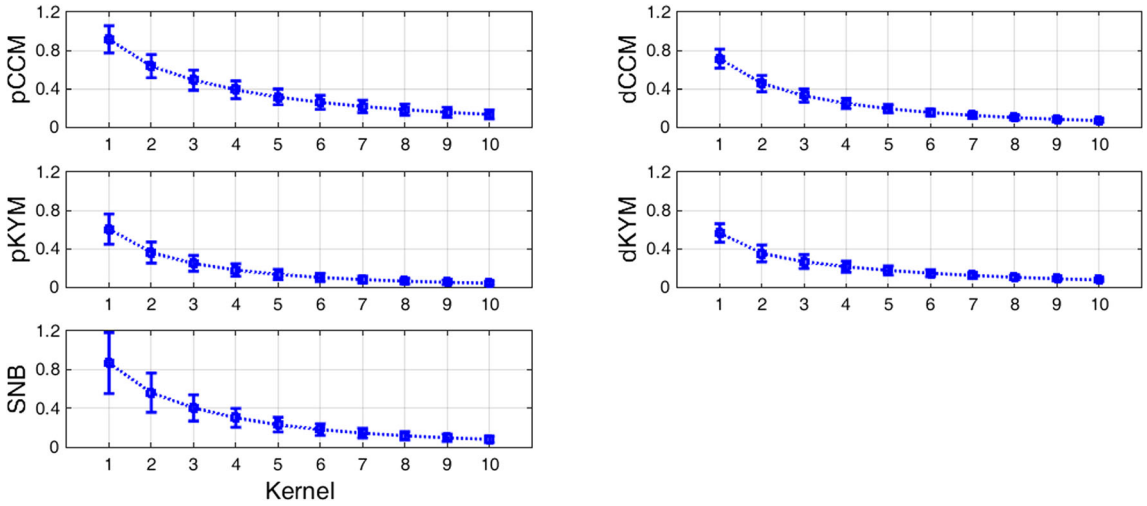


Fig. 4 Variations of corrEn in different states. X-axis shows the kernel bandwidth and y-axis depicts the corrEn values

3- Weighted sum rule.

$$Wsr\_f = \sum_{i=1}^{5} W_i \mathbf{A} + W_i' \mathbf{B} \quad (7)$$

where W denotes the weights and $i = 1,2,..5$ correspond to the five predefined classes (pCCM, dCCM, pKYM, dKYM, and SNB).

In weighted sum rule, the following assumption is adopted: the contribution of some features is more than the other measures. To address this problem, the indices were weighted regarding their significance in each class. This significance was defined based on the statistical p-values. If the corresponded p is lower than 1×10^{-4}, then W=2. For $1 \times 10^{-4} < p < 1 \times 10^{-3}$, W=1.5; for $1 \times 10^{-3} < p < 1 \times 10^{-2}$, W=1; for $0.01 < p < 0.05$, W=0.5; and for $p \geq 0.05$, W=0 were considered.

Figure 2 shows the three feature level fusion rules, schematically.

Classification

Probabilistic neural network

The HRV classification was performed using probabilistic neural network (PNN) to separate SNB from each pCCM, dCCM, pKYM, and dKYM state. Before inputting the features to the PNN, they were normalized (Norm_F) as follows:

$$Norm\_F = 2\left(\frac{F - F_{\min}}{F_{\max} - F_{\min}}\right) - 1 \quad (8)$$

The efficiency of PNN in biomedical pattern recognition and signal classification has been established in the previous

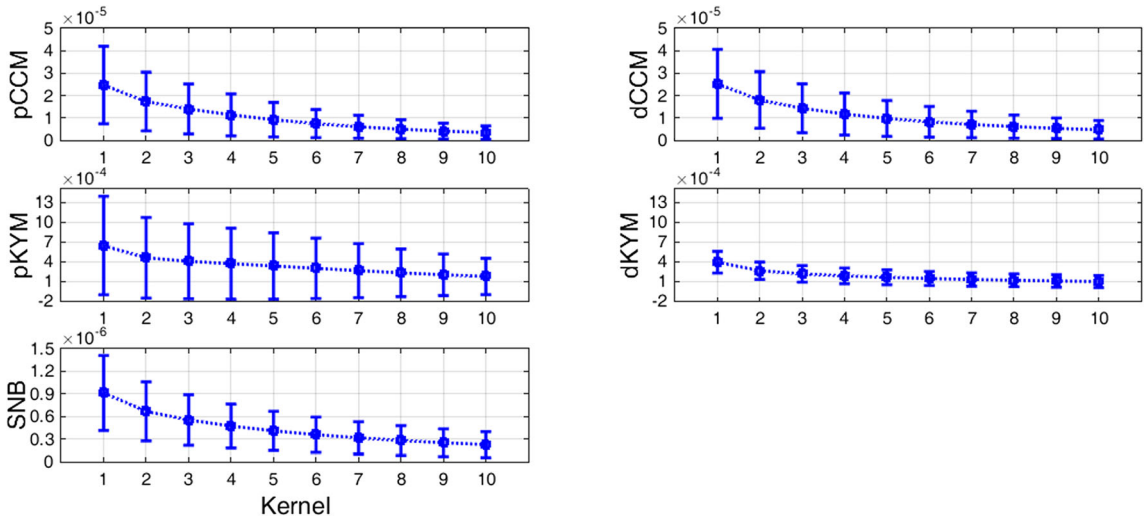


Fig. 5 Variations of CSD in different states. X-axis shows the kernel bandwidth and y-axis depicts the CSD values

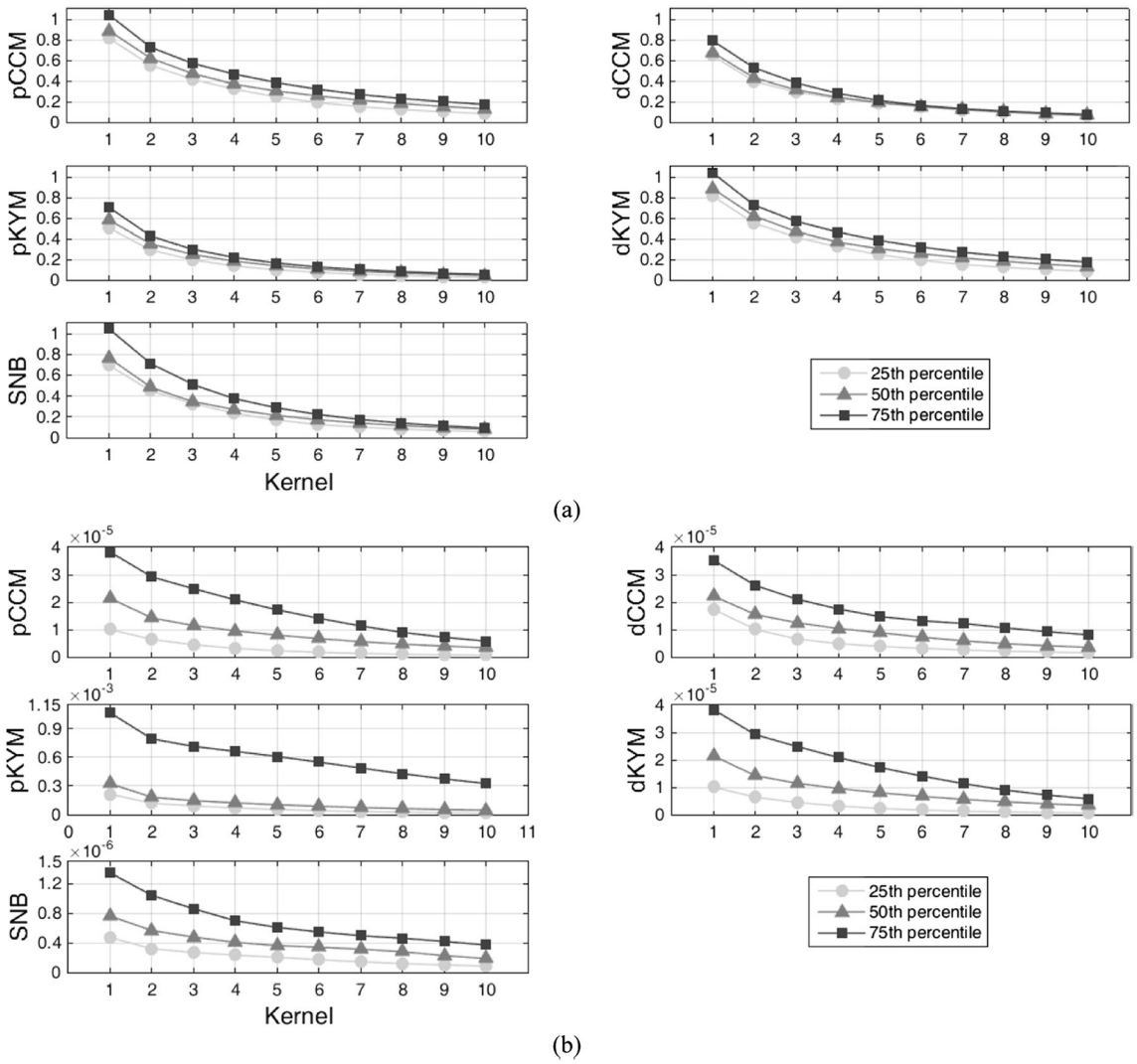


Fig. 6 The percentiles of the features in different states. **a** corrEn and (**b**) CSD. X-axis shows the kernel bandwidth and y-axis depicts the feature values

literature [31]. The PNN classifier is allocated to the group of feed forward based networks which applies the radial basis function using the sigma adjustment values (σ). To better understand how this classifier works, Fig. 3 shows the PNN architecture, schematically.

In this study, to discriminate between SNB and each pCCM, dCCM, pKYM, and dKYM group, a 5-fold cross-validation fashion was adopted. Randomly dividing data into 5 equal parts the cross-validation is performed, while 4 parts are allocated to the training and one for the test. Additionally, the performance of PNN was appraised for 20 different sigma values (in the range of 0.05 to 1 with the step size of 0.05; σ = 0.05, 0.1, 0.15, …, 1). To evaluate the classification performances, accuracy was calculated as Eq. (9):

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \tag{9}$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively.

K-nearest neighbor

K-nearest neighbor (kNN) is a supervised and nonparametric procedure. In this techniques, by computing Euclidean distance, the new sample query result is categorized based on the closeness of the k-nearest samples accessible in the feature space. Applying a short time for training, it results in a good performance [32, 33]. In this study, different k values (k_1) were tested to evaluate the classification results (k_1 was set in the range of 2 to 11 with the step size of 1; $k_1 = 2, 3, \ldots, 11$).

Results

Variation of mean corrEn values with varying kernel sizes were shown in Fig. 4 for HRV signals during different states.

As seen in Fig. 4, the lowest variations in the corrEn were observed for the large kernel sizes in all states. Considering the lower kernel sizes, the differences of corrEn values in

Table 2 Statistical differences between the groups (* denotes $p < 0.05$, which shows significant difference)

Kernel size	pCCM vs. SNB		dCCM vs. SNB		pKYM vs. SNB		dKYM vs. SNB	
	CSD	corrEn	CSD	corrEn	CSD	corrEn	CSD	corrEn
k = 1	$2.65 \times 10^{-5*}$	0.35	$7.9 \times 10^{-4*}$	0.24	$1.5 \times 10^{-3*}$	0.18	$1.5 \times 10^{-3*}$	0.026*
k = 2	$2.65 \times 10^{-5*}$	0.31	$3.6 \times 10^{-3*}$	0.27	$1.5 \times 10^{-3*}$	0.14	$1.5 \times 10^{-3*}$	0.040*
k = 3	$2.65 \times 10^{-5*}$	0.15	$3.6 \times 10^{-3*}$	0.15	$1.5 \times 10^{-3*}$	0.10	$1.5 \times 10^{-3*}$	0.14
k = 4	$5.29 \times 10^{-5*}$	0.09	$3.6 \times 10^{-3*}$	0.20	$1.5 \times 10^{-3*}$	0.08	$1.5 \times 10^{-3*}$	0.18
k = 5	$1.85 \times 10^{-4*}$	0.07	$3.6 \times 10^{-3*}$	0.35	$1.5 \times 10^{-3*}$	0.040*	$1.5 \times 10^{-3*}$	0.23
k = 6	$1.85 \times 10^{-4*}$	0.040*	$9.1 \times 10^{-3*}$	0.49	$1.5 \times 10^{-3*}$	0.040*	$1.5 \times 10^{-3*}$	0.41
k = 7	$1.85 \times 10^{-4*}$	0.033*	0.026*	0.66	$1.5 \times 10^{-3*}$	0.026*	$1.5 \times 10^{-3*}$	0.57
k = 8	$1.85 \times 10^{-4*}$	0.033*	0.033*	0.66	$1.5 \times 10^{-3*}$	0.040*	$1.5 \times 10^{-3*}$	0.75
k = 9	$3.18 \times 10^{-4*}$	0.026*	0.033*	0.66	$1.5 \times 10^{-3*}$	0.026*	$1.5 \times 10^{-3*}$	0.95
k = 10	$3.18 \times 10^{-4*}$	0.020*	0.050	0.66	$1.5 \times 10^{-3*}$	0.026*	$1.5 \times 10^{-3*}$	0.85

different psychological states were more pronounced. For k = 1, the mean corrEn values were 0.9146, 0.7128, 0.6039, 0.5635, and 0.8679 for pCCM, dCCM, pKYM, dKYM, and SNB, respectively. The corresponding corrEn values were decreased to 0.1317, 0.067, 0.04, 0.0741, and 0.0801 for k = 10. Therefore, the lowest corrEn values were assigned to the HRV signals of KYM (both pKYM and dKYM) states.

Figure 5 shows variations of mean CSD values with varying kernel bandwidth in different states.

Adopting k = 1, for pCCM, dCCM, pKYM, dKYM, and SNB, the average CSD values were 2.47×10^{-5}, 2.524×10^{-5}, 6.418×10^{-4}, 3.902×10^{-4}, and 9.131×10^{-7}, respectively; which decreased to 3.44×10^{-6}, 4.72×10^{-6}, 1.722×10^{-4}, 9.52×10^{-5}, 2.259×10^{-7}, for k = 10. The largest CSD values were obtained for KYM states (pKYM and dKYM) and the minimum value was found for SNB.

Totally, the variations of corrEn and CSD across lower kernel values are larger than those across higher kernel sizes. Additionally, diverse similarity measures were attained for different psychological states.

The distribution of the features were evaluated and the median and statistical parameters of dispersion are shown in Fig. 6.

Since the distribution of the features is not-normal (Fig. 6), the Wilcoxon ranksum test was used to compare two states for statistical differences. Table 2 lists the *p*-values.

Figure 7 summarizes the performance of the classifier using the corrEn similarity measure.

In discriminating pCCM from SNB, the accuracy of 100% was achieved using k = 10 and σ = 0.05. Figure 7a shows that the average accuracy was maximum for σ = 0.05 and minimum for σ = 1. Discrimination between pCCM and SNB showed the highest average recognition rate of 85% for σ = 0.05, which dropped to 56% for σ = 1. The maximum average classification results of 72.81% was obtained for k = 10.

For k = 9 and k = 10 with σ = 0.05, the accuracy was 100% for dCCM vs. SNB classification. In this case, the highest average rate of 89.38% and the lowest mean rate of 55% were reached for σ = 0.05 and σ = 0.6, respectively. Additionally, the highest average recognition rate was 62.81% for k = 6.

Figure 7b shows the KYM vs. SNB classification results. The pKYM vs. SNB classification was performed with the highest rate of 100% for σ = 0.05 and k = 1,2,3, and 4. In this case, the maximum correct rate (90.83%) and minimum correct rate (66.67%) were achieved for σ = 0.05 and σ = 0.55 to 1, respectively. In addition, the maximum average accuracy (72.92%) was obtained for k = 6.

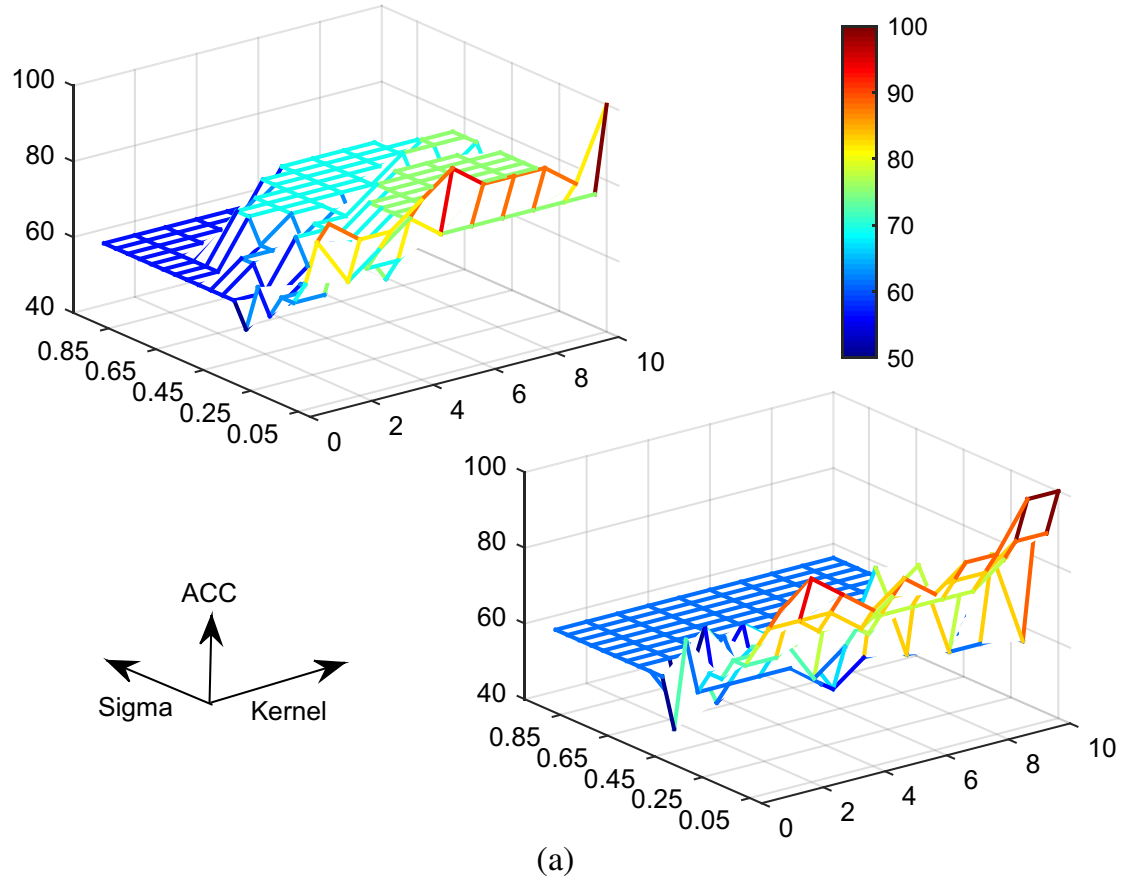
Figure 7c shows the classification results of two classes (pCCM vs. dCCM and pKYM vs. dKYM). For pCCM and dCCM classification, the highest rate of 100% was obtained for σ = 0.05 and k = 7–9. The maximum average rates were 93.85% for σ = 0.05 and 76.92% for k = 1 and k = 2.

Compared to the other states, higher recognition rates were achieved for pKYM and dKYM classification. For σ = 0.05 and k = 1,2,5,6,7,8,9,10; for σ = 0.1 and k = 1,2,6,7,8,9,10; for σ = 0.15 and k = 1,2,7,8,9,10; for σ = 0.2 and k = 2,8,9,10; and for σ = 0.25 and k = 10, the accuracy was 100%. In addition, the average maximum correct rates were 95.71% for σ = 0.05 and 82.86% for k = 10.

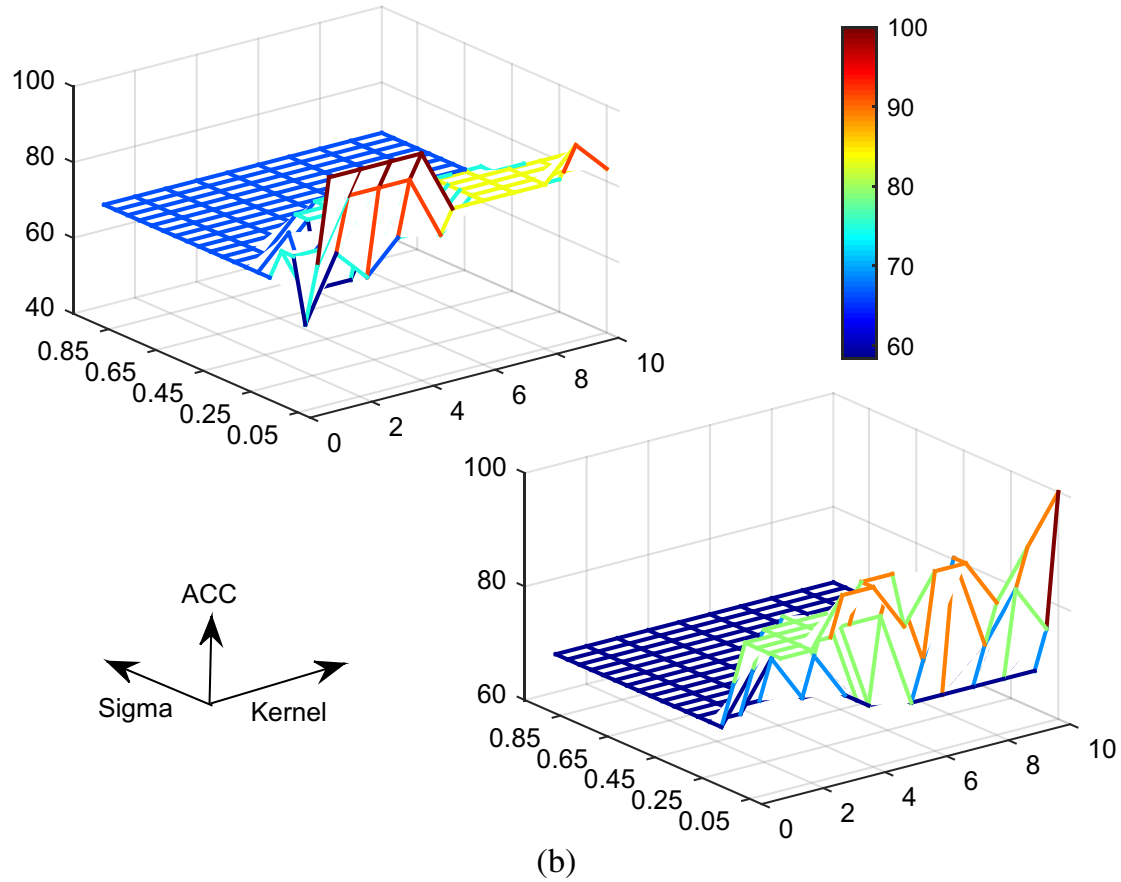
The performances of the CSD similarity index for classifying different states are shown in Fig. 8.

Figure 8a shows the classification results of CCM vs SNB; the top frame illustrates the pCCM vs. SNB and the bottom displays the dCCM vs. SNB. For the former, the highest rate of 100% was achieved for σ = 0.05 and k = 1. In this case, the maximum average accuracies were 90.63% for σ = 0.05 and 80.94% for k = 6, 7, 8. For the latter, the recognition rate of 100% was obtained for σ = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3 and k = 1, for σ = 0.05, 0.1, 0.15 and k = 2, and for σ = 0.05 and k = 3. In this case, the highest average rates of 95.63 and 85.63% were obtained for σ = 0.05 and k = 1, respectively.

The KYM vs. SNB classification results are shown in Fig. 8b; the top frame illustrates the pKYM vs. SNB and the


100
80
60
40
0.85
0.65
0.45
0.25
0.05
0
2
4
6
8
10
100
90
80
70
60
50
ACC
Sigma
Kernel

(a)


100
80
60
40
0.85
0.65
0.45
0.25
0.05
0
2
4
6
8
10
100
90
80
70
60
ACC
Sigma
Kernel

(b)

◀ **Fig. 7** PNN classification accuracies using corrEn for different k and varying σ parameter. **a** *top:* pCCM vs. SNB, *bottom:* dCCM vs. SNB; (**b**) *top:* pKYM vs. SNB, *bottom:* dKYM vs. SNB; (**c**) *top:* pCCM vs. dCCM, *bottom:* pKYM vs. dKYM

bottom presents the dKYM vs. SNB. For pKYM vs. SNB, the maximum average rates were accomplished for σ = 0.05 and k = 1 with the rates of 87.5 and 75%, respectively. However, the accuracy of 100% was attained using σ = 0.05 and k = 1. For dKYM vs. SNB, the average classification accuracies have increased sharply. The mean recognition rate of 100% was obtained for σ = 0.05 and σ = 0.1. Considering kernel sizes, the highest average rate of 84.58% was accomplished for k = 1. In addition, for all sigma values and k = 1 and k = 2, the accuracy was 100%.

For pCCM vs. dCCM classification (Fig. 8c, top frame), the highest mean accuracies were 93.08 and 73.08% for σ = 0.05 and k = 8, respectively. However, the recognition rates were 100% for σ = 0.05 and k = 2, 3, 4. For pKYM vs. dKYM classification (Fig. 8c, bottom frame), the average classification accuracies were maximum for σ = 0.05 (85.71%) and k = 1 and k = 2 (69.29%). However, the recognition rate was 85.71% for σ = 0.05 and all kernel sizes.

The feature level fusion was performed using three fusion rules. To simplify, we only reported the maximum accuracy and the parameters for which the maximum accuracy was obtained in Table 3. For comparison, the classification results of any non-fusion feature are also summarized in Table 3.

The results showed that two fusion approaches, including weighted sum and division rules, outperformed the other schemes. For these two fusion methodologies, a greater number of selected parameters have resulted in the accuracy of 100%. However, regarding the maximum accuracy of the classification, the product rule resulted in a lower accuracy rate than the other fusion and non-fusion strategies. Additionally, the results showed that discrimination of two classes of dKYM and dCCM performed better than other classes. Exactly, in the most cases, the highest accuracy was obtained. According to the classification results, the best values for σ and k are 0.05 and 1, respectively.

The results of the data classification using kNN are given in Table 4.

The results showed that applying division rule as a fusion approach, the accuracy of 100% was achieved for discrimination of all states. The second best results were obtained with the other two fusion rules. In these cases, the maximum

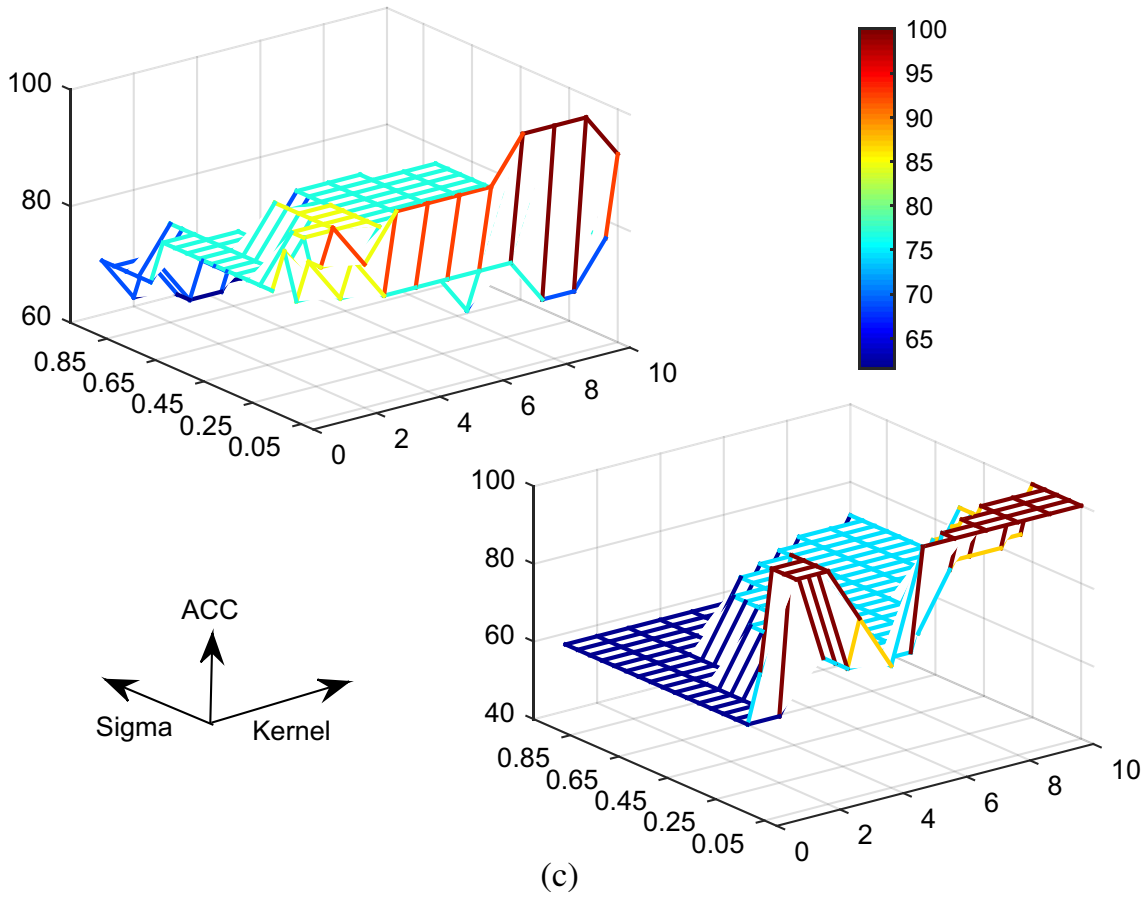
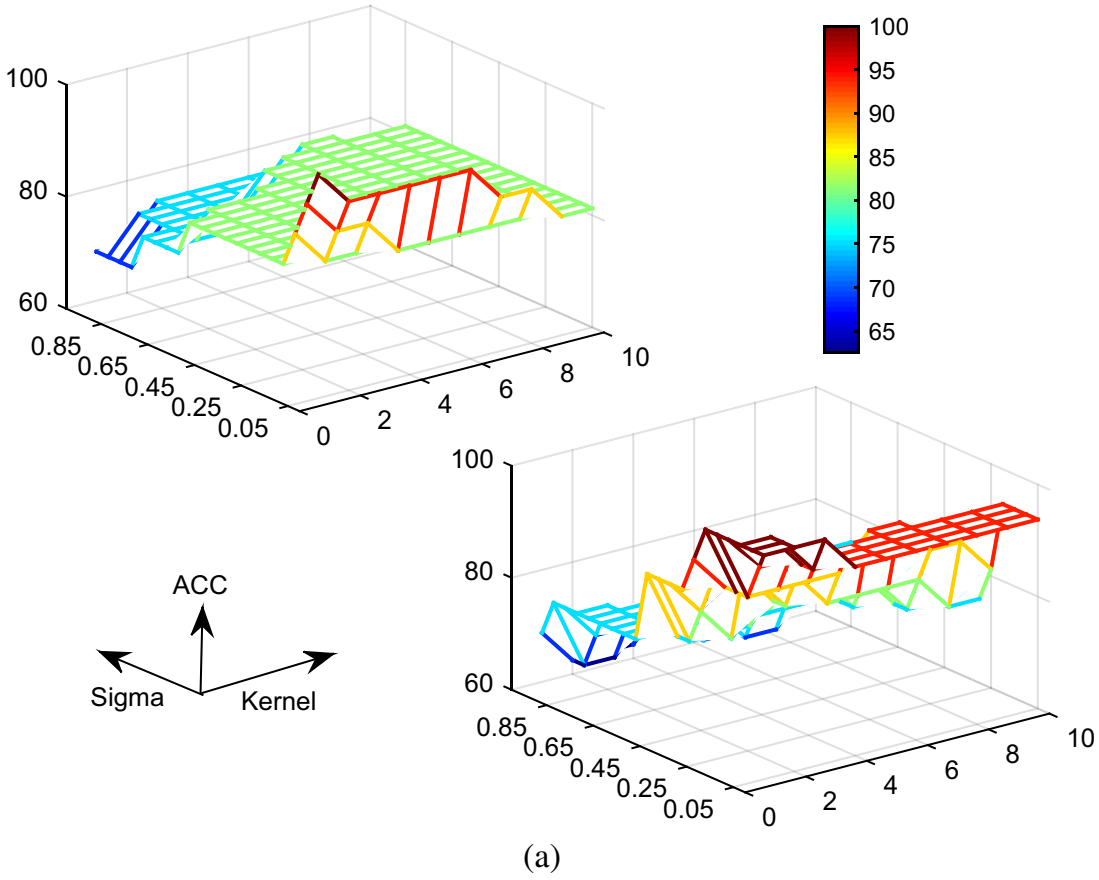
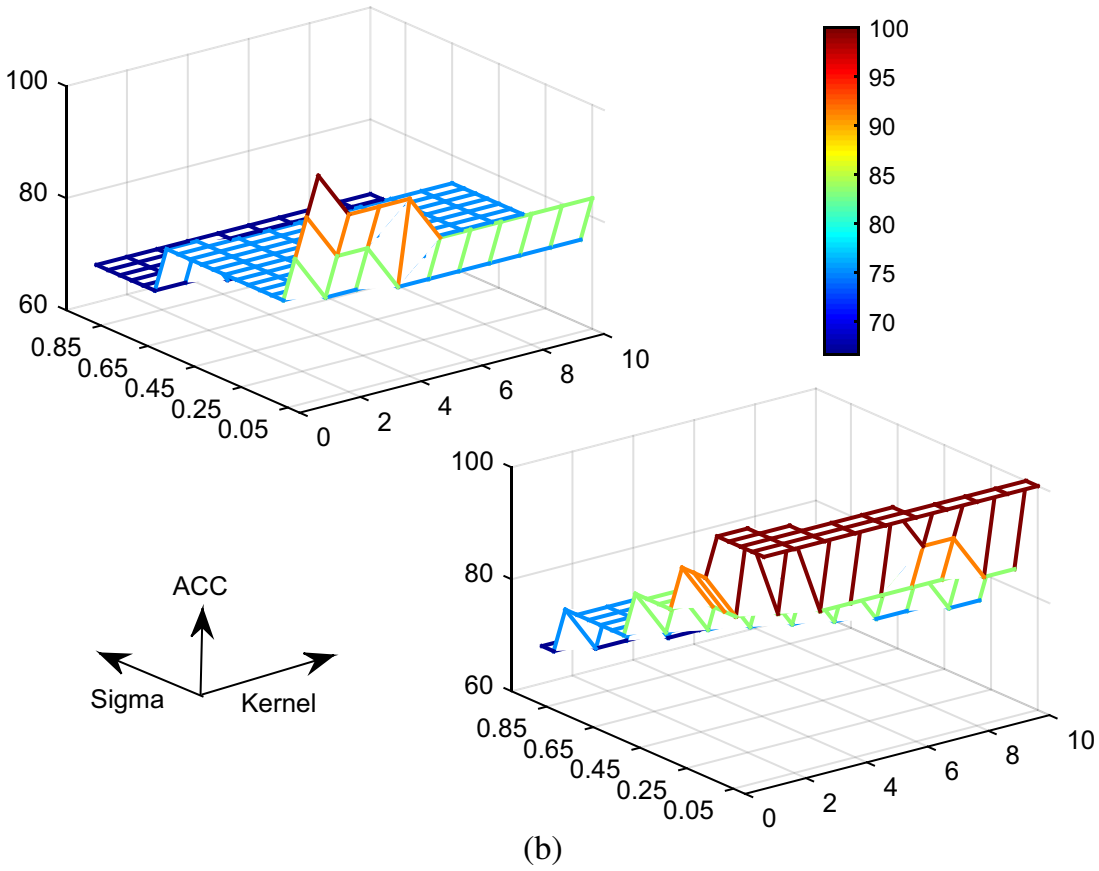


Fig. 7 (continued)


100
80
60
0.85
0.65
0.45
0.25
0.05
0
2
4
6
8
10
100
95
90
85
80
75
70
65
ACC
Sigma
Kernel

(a)


100
80
60
0.85
0.65
0.45
0.25
0.05
0
2
4
6
8
10
100
95
90
85
80
75
70
ACC
Sigma
Kernel

(b)

◀**Fig. 8** PNN classification accuracy using CSD for different k and varying σ parameter. **a** *top:* pCCM vs. SNB, *bottom:* dCCM vs. SNB; (**b**) *top:* pKYM vs. SNB, *bottom:* dKYM vs. SNB; (**c**) *top:* pCCM vs. dCCM, *bottom:* pKYM vs. dKYM

accuracy of 100% was found for discriminating all classes, except for pCCM. Regarding the maximum accuracy, a lower accuracy rate was gotten for corrEn. Moreover, the kNN results revealed that dKYM and pKYM classes were classified with higher recognition rates, compared with other classes. According to Table 4, the best value for k was 5.

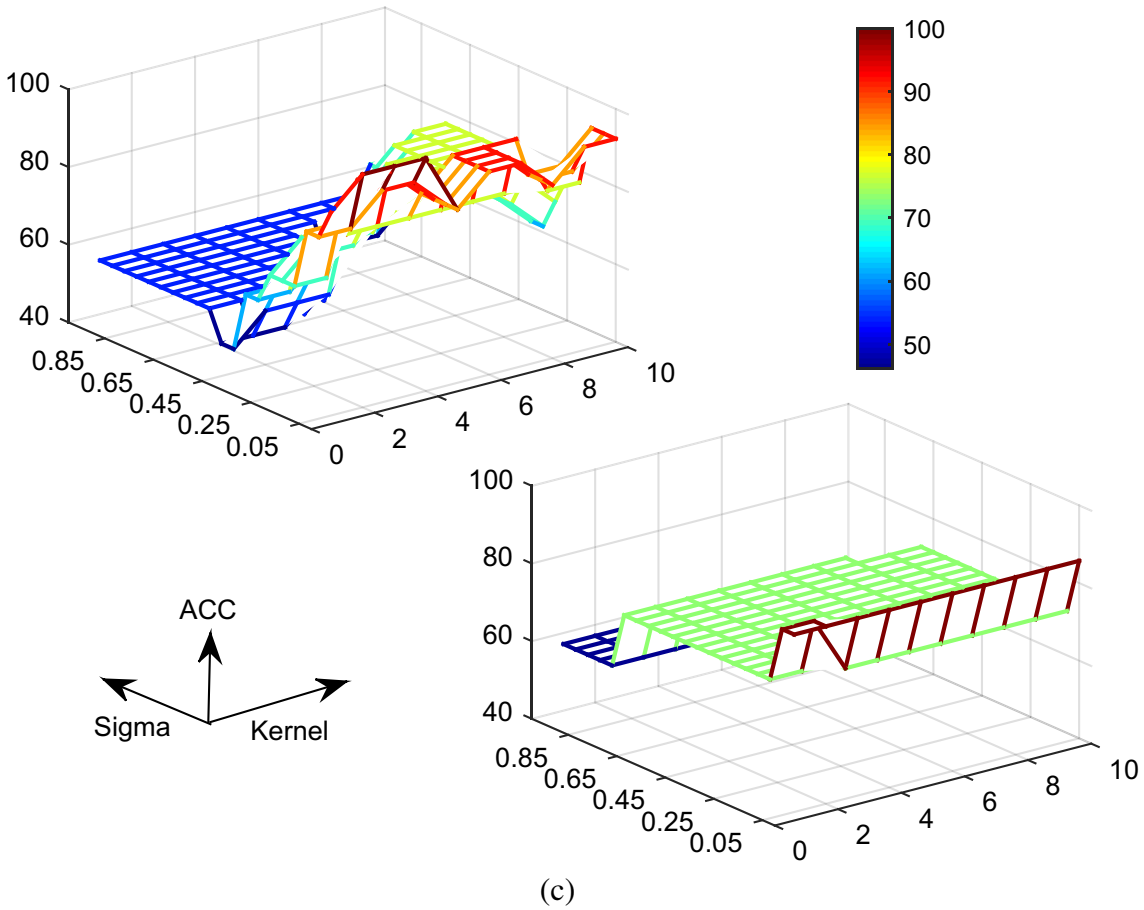
Comparing the performance of two classifiers (kNN and PNN), it can be concluded that:

- Better classification performances were achieved using PNN.
- In terms of classification accuracy, for both classifiers (PNN and kNN), division rule outperformed the other strategies, in which the maximum accuracy of 100% was achieved for classifying all states.
- For both classifiers, the best classification results were achieved for dKYM.

Discussions

In this paper, a comparison between two information theoretic similarity indices (CSD and corrEn) was done to classify HRV signals. In addition, three feature level fusion approaches were examined. To this effect, the product rule, division rule, and weighted sum rule were evaluated. An innovative approach was introduced to define the weights of each feature in the weighted sum fusion rule. The discriminative ability of the features was determined using the *p*-values of the statistical test. The system performances were evaluated by changing the kernel size of the features and adjustable parameter of the classifiers (sigma for PNN and k_1 for kNN).

It has previously been shown that HRV characteristics are affected by meditation. Li et al. [12] studied the HRV complexity in CCM and KYM groups. They concluded that the dynamical complexity of HRV declines during meditation. Song et al. [11] showed more HRV regularities and lower degree of multifractalities during CCM. Analyzing recurrence plots confirmed that the nonlinear HRV interactions were dropped during meditation [13].



(c)

Fig. 8 (continued)

Table 3 The highest accuracy of the combined and non-combined features in classification of SNB and each psychological state using PNN

		SNB vs. ↓			
		pCCM	dCCM	pKYM	dKYM
Fusion					
Division rule	Max ACC (%)	100	100	100	100
	k & σ	k = 1 & σ = 0.05	k = 1, 2 & σ = 0.05	k = 1 & σ = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4	k = 1 & σ = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35
	k & σ		k = 2 & σ = 0.2, 0.15	k = 2 & σ = 0.05, 0.1, 0.15	k = 2 & σ = 0.05, 0.1
	k & σ			k = 3 & σ = 0.05	k = 3 & σ = 0.05
Product rule	Max ACC (%)	93.75	93.75	91.67	100
	k & σ	k = 1, 2, 3, 4, 5, 6, 7 & σ = 0.05	k = 1 & σ = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4	k = 1, 2 & σ = 0.05	k = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 & σ = 0.05
	k & σ	k = 1, 2 & σ = 0.1	k = 2 & σ = 0.05, 0.1		k = 1, 2, 3, 4, 5, 6 & σ = 0.1
	k & σ		k = 3 & σ = 0.05		k = 1, 2 & σ = 0.15
	k & σ				k = 1 & σ = 0.2
Weighted sum rule	Max ACC (%)	100	100	100	100
	k & σ	k = 1, 2, 3, 4, 6 & σ = 0.05	k = 1 & σ = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3	k = 1 & σ = 0.05	k = 3, 4, 5, 6, 7, 8, 9, 10 & σ = 0.05, 0.1
	k & σ		k = 2 & σ = 0.05, 0.1		k = 3, 4, 5, 6 & σ = 0.15
	k & σ		k = 3 & σ = 0.05		k = 3 & σ = 0.2
CSD					
	Max ACC (%)	100	100	100	100
	k & σ	k = 1 & σ = 0.05	k = 1 & σ = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3	k = 1 & σ = 0.05	k = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 & σ = 0.05, 0.1
	k & σ		k = 2 & σ = 0.05, 0.1, 0.15		k = 1, 2, 3, 4, 5, 6 & σ = 0.15
	k & σ		k = 3 & σ = 0.05		k = 1, 2, 3 & σ = 0.2
	k & σ				k = 1, 2 & σ = 0.25
corrEn					
	Max ACC (%)	100	100	100	100
	k & σ	k = 10 & σ = 0.05	k = 9, 10 & σ = 0.05	k = 1, 2, 3, 4 & σ = 0.05	k = 10 & σ = 0.05

Max ACC, Maximum Accuracy; *k*, kernel size of the features; *σ*, the PNN adjustable parameter

Our results are consistent with the previous findings. It has shown that the similarity measures of HRV signals were affected by the psychological state. The results showed that the lowest corrEn values were achieved for the HRV signals of pKYM and dKYM states. The lower corrEn might be attributed to the lower stress levels and lower sympathetic reactions in the dKYM. In contrast, the highest corrEn has belonged to the pCCM and SNB. The highest CSD was attained for pKYM and the lowest was obtained for the SNB. As the kernel size increases, the mean similarity measures decrease, as well as their variations.

The critical role of parameter selection in classification performances has been also shown. Both kernel size of the features and sigma/k_1 parameters affected the classification accuracies. In addition, our results showed the benefit of feature level fusion strategies in HRV classification. Previously, the role of feature level fusion in increasing the performance of the classification of psychological data has been shown [34, 35].

By combining the similarity indices derived from corrEn and CSD, valuable information of HRV signals is achieved which led to superior classification results. The CSD was unable to show the nonlinear similarity of the signals, which is accessible with corrEn. Consequently, the information of these two features is complimentary. Although the high classification performances (100%) were achieved by each feature, fusion strategies (especially division rule) obtained a higher number of maximum accuracies (Tables 3 and 4). Totally, the best results were obtained using PNN.

The classification of HRV signals during mediation has not been documented sufficiently. Previously, a system was proposed to classify HRV signals of Samadhi and non-Samadhi groups based on time and frequency features [36]. The

Table 4 The highest accuracy of the combined and non-combined features in classification of SNB and each psychological state using kNN

		SNB vs. ↓			
		pCCM	dCCM	pKYM	dKYM
Fusion					
Division rule	Max ACC (%)	100	100	100	100
	k & k_1	k = 1,2,3,4,5,6,7,8 & k_1 = 2	k = 1,2,3,4,5 & k_1 = 2	k = 2,3,4,5,6,7,8,9,10 & k_1 = 2	k = 2,3,4,5,6,7,8,9,10 & k_1 = 2
	k & k_1	k = 4,6,7 & k_1 = 3	k = 3,5 & k_1 = 4	k = 4,5,6,7,8,9,10 & k_1 = 3	k = 4,5,6,7,8,10 & k_1 = 3
	k & k_1	k = 5,6 & k_1 = 4	k = 4 & k_1 = 5	k = 3,5,6,7,8,9,10 & k_1 = 4	k = 4,5,6,7,8,9,10 & k_1 = 4
	k & k_1	k = 5 & k_1 = 5		k = 5,6,7,9,10 & k_1 = 5	k = 4,6,10 & k_1 = 5
Product rule	Max ACC (%)	93.33	100	100	100
	k & k_1	k = 1,2,3,4,5,6,7,9 & k_1 = 2	k = 1 & k_1 = 2,3,4,5,6,9	k = 1,2,3,4,5,6 & k_1 = 2	k = 1,2,3,4,5,7,9,10 & k_1 = 2
	k & k_1	k = 1,2,3,4,5,6,7 & k_1 = 3		k = 1,2,3,4,5 & k_1 = 3	k = 1,2,3,4,5,6,7,8,9,10 & k_1 = 3
	k & k_1	k = 1,2,3,4,5,6,10 & k_1 = 4		k = 1,2,4,5,6 & k_1 = 4	k = 1,2,3,4,5,6,7,8,9 & k_1 = 4
	k & k_1	k = 1,2,3,4,5,6,7,8,9,10 & k_1 = 5,6,7,8,9,10,11		k = 1,2,3,4 & k_1 = 5	k = 1,2,3,4,5 & k_1 = 5
	k & k_1			k = 1,2,3 & k_1 = 6,8	k = 1,2,3,4,5,6,7,8 & k_1 = 6
	k & k_1			k = 1 & k_1 = 7	k = 1,2,5 & k_1 = 7
	k & k_1				k = 4 & k_1 = 8
Weighted sum rule	Max ACC (%)	93.33	100	100	100
	k & k_1	k = 1,2,3,4,5 & k_1 = 2,3,4,5,6,7,8,9,10,11	k = 1 & k_1 = 2,3,4,5,6,8	k = 1,2 & k_1 = 2	k = 3,6,7,8,9 & k_1 = 2
	k & k_1			k = 1,2,3 & k_1 = 3,4	k = 3,4,5,6,7,8,9,10 & k_1 = 3
	k & k_1			k = 3 & k_1 = 6	k = 3,4,5,6,7,8,9 & k_1 = 4
	k & k_1				k = 7,9 & k_1 = 5
	k & k_1				k = 8,9 & k_1 = 6
CSD					
	Max ACC (%)	78.95	78.95	100	100
	k & k_1	k = 1,7 & k_1 = 2	k = 5 & k_1 = 8	k = 1,2,3 & k_1 = 2	k = 1,2,4,5,6,7,8,9,10 & k_1 = 2
	k & k_1	k = 1 & k_1 = 3		k = 1,2,3 & k_1 = 3	k = 1,2,3,4,5,6,7,8,9,10 & k_1 = 3
	k & k_1			k = 1,3 & k_1 = 4	k = 1,2,6 & k_1 = 4
	k & k_1				k = 1,2,8,9 & k_1 = 5
	k & k_1				k = 1,2 & k_1 = 6,7
corrEn					
	Max ACC (%)	68.42	68.42	80	80
	k & k_1	k = 10 & k_1 = 8,11	k = 7 & k_1 = 6	k = 6 & k_1 = 5	k = 9 & k_1 = 2
	k & k_1			k = 6,7,9 & k_1 = 9	k = 2 & k_1 = 4
	k & k_1				k = 1 & k_1 = 6

Max ACC, Maximum Accuracy; *k*, kernel size of the features; k_1, the kNN adjustable parameter.

accuracy of 94.8% was attained using Fisher discriminant analysis. In another study, Lyapunov exponent and entropy were calculated to characterize HRV dynamics during CCM and pCCM [16]. Three machine learning systems were evaluated, including Quadratic, fisher, and kNN. Quadratic classifier achieved the maximum accuracy of 92.31%.

To sum up, the results of this study established that new combined similarity indices of HRV signals could be served as an appropriate quantity to correctly discriminate different psychological states.

One of the limitations of this study is the number HRV signals for evaluating the proposed algorithm. It is suggested that the framework be assessed on a big data in the future. In addition, the number of signals in different categories is not the same in the existing database. This imbalance classes can affect classifier recognition rates. In the future, the class imbalance problem should be taken into account.

Conclusion

Classification of HRV signals in different psychological states has recently become a challenging topic. In this study, we attempted to characterize the HRV signals in different states, including pCCM, dCCM, pKYM, dKYM, and SNB. To address this issue, two features were extracted which quantified the similarity of the time-series, known as CSD and corrEn. In addition, a novel feature level fusion approach was presented based on the *p*-values of the similarity indices. The experimental results on the two forms of meditation (KYM and CCM) confirmed the proficiency of the proposed feature level fusion strategy for classifying HRV in the psychological states. Specifically, the results suggested that combination of CSD and corrEn information of the HRV signals based on weighted sum and division rules can improve the accuracy of the classifiers. In addition, the role of parameter selection on classification performances has been shown. The maximum accuracy of 100% was achieved.

Funding "This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors".

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Karimi Moridani, M., Setarehdan, S. K., Motie Nasrabadi, A., and Hajinasrollah, E., Analysis of heart rate variability as a predictor of mortality in cardiovascular patients of intensive care unit. *Biocybern. Biomed. Eng.* 35:217–226, 2015.
2. MM Hassan, S., Huda, J. Y., Jelinek, H. F., and Almogren, A., Multistage fusion approaches based on a generative model and multivariate exponentially weighted moving average for diagnosis of cardiovascular autonomic nerve dysfunction. *Inform. Fusion* 41: 105–118, 2018.
3. Bevilacqua, F., Engstrom, H., and Backlund, P., Changes in heart rate and facial actions during a gaming session with provoked boredom and stress. *Entertain. Comput.* 24:10–20, 2018.
4. Huikuri, H. V., Makikallio, T., Airaksinen, K., Mitrani, R., Castellanos, A., and Myerburg, R. J., Measurement of heart rate variability: A clinical tool or a research toy? *J. Am. Coll. Cardiol.* 34:1878–1883, 1999.
5. Acharya, U. R., Kannathal, N., OngWai, S., Luk Yi, P., and TjiLeng, C., Heart rate analysis in normal subjects of various age groups. *Biomed. Eng. Online* 3:24–28, 2004.
6. Santamaria, I., Pokharel, P. P., and Principe, J. C., Generalized correlation function: Definition, properties and application to blind equalization. *IEEE Trans. Signal Process.* 54:2187–2197, 2006.
7. Hasanbelliu, E., Giraldo, L. S., and Principe, J. C., Information theoretic shape matching. *IEEE Trans. Pattern Anal. Mach. Intell.* 36:2436–2451, 2014.
8. Goshvarpour, A., Goshvarpour, A., and Rahati, S., Analysis of lagged Poincare plots in heart rate signals during meditation. *Digit. Signal Process.* 21:208–214, 2011.
9. Goshvarpour, A., and Goshvarpour, A., Poincare indices for analyzing meditative heart rate signals. *Biomed. J.* 38:229–234, 2015.
10. Alvarez-Ramirez, J., and E Rodriguez, J. C., Echeverria fractal scaling behavior of heart rate variability in response to meditation techniques. *Chaos Solitons Fractals.* 99:57–62, 2017.
11. Song, R., Bian, C., and Ma, Q. D. Y., Multifractal analysis of heartbeat dynamics during meditation training. *Physica A.* 392:1858–1862, 2013.
12. J Li, J. H., Zhang, Y., and Zhang, X., Dynamical complexity changes during two forms of meditation. *Physica. A.* 390:2381–2387, 2011.
13. Goshvarpour, A., and Goshvarpour, A., Recurrence plots of heart rate signals during meditation. *I. J. Image Graph. Signal Process.* 2: 44–50, 2012.
14. Peng, C.-K., Mietus, J. E., Liu, Y., Khalsa, G., Douglas, P. S., Benson, H., and Goldberger, A. L., Exaggerated heart rate oscillations during two meditation techniques. *Int. J. Cardiol.* 70:101–107, 1999.
15. Goshvarpour, A., and Goshvarpour, A., Comparison of higher order spectra in heart rate signals during two techniques of meditation: Chi and Kundalini meditation. *Cogn. Neurodyn.* 7:39–46, 2013.
16. Goshvarpour, A., and Goshvarpour, A., Classification of heart rate signals during meditation using Lyapunov exponents and entropy. *I. J. Intell. Syst. Appl.* 2:35–41, 2012.
17. Goshvarpour, A., and Goshvarpour, A., Chaotic behavior of heart rate signals during Chi and Kundalini meditation. *I. J. Image Graph. Signal Process.* 2:23–29, 2012.
18. Gao, J., J Fan, B. W. W., Zhang, Z., Chang, C., Hung, Y. S., Fung, P. C., and Sik, H. H., Entrainment of chaotic activities in brain and heart during MBSR mindfulness training. *Neurosci. Lett.* 616:218–223, 2016.
19. Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E., PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 101(23):e215–e220, 2000.
20. Liu, W., Pokharel, P. P., and Principe, J. C., Correntropy: A localized similarity measure. In: The IEEE International Joint Conference on Neural Network Proceedings, pp 4919–4924, 2006.
21. Garde, A., Sornmo, L., and Jané, R., Correntropy-based analysis of respiratory patterns in patients with chronic heart failure. In: Proceedings of the 31st annual international conference of the

IEEE engineering in medicine and biology society: Engineering the future of Biomedicine, pp 4687–4690, 2009.

22. Pokharel, P. P., Liu, W., and Principe, J. C., A low complexity robust detector in impulsive noise. *Signal Process.* 89:1902–1909, 2009.
23. Liu, W., and PP Pokharel, C. P. J., Correntropy: Properties and applications in non-Gaussian signal processing. *IEEE Trans. Signal Process.* 55(11):5286–5298, 2007.
24. Melia, U., Guaita, M., Vallverdú, M., Montserrat, J. M., Vilaseca, I., Salamero, M., Gaig, C., Caminal, P., and Santamaria, J., Correntropy measures to detect daytime sleepiness from EEG signals. *Physiol. Meas.* 35(10):2067–2083, 2014.
25. Guaita, M., Melia, U., Vallverdú, M., Caminal, P., Vilaseca, I., Montserrat, J. M., Gaig, C., Salamero, M., and Santamaria, J., Regularity of cardiac rhythm as a marker of sleepiness in sleep disordered breathing. *PLoS One* 10(4):e0122645, 2015.
26. Kullback, S., *Information theory and statistics*. New York: Wiley, 1959.
27. Czarnecki, W. M., and Tabor, J., Multithreshold entropy linear classifier: Theory and applications. *Expert. Syst. Appl.* 42:5591–5606, 2015.
28. Nguyen, C., Lovering, C., and Neamtu, R., Ranked time series matching by interleaving similarity distances. In: IEEE International Conference on Big Data (Big Data), 11–14 Dec. 2017, Boston, MA, USA, 2017.
29. Czarnecki, W. M., and Tabor, J., Extreme entropy machines: Robust information theoretic classification. *Pattern. Anal. Appl* 2(2):383–400, 2017.
30. Seth, S., and Príncipe, J. C., On speeding up computation in information theoretic learning. In: International Joint Conference on Neural Networks (IJCNN 2009), 14–19 June 2009, Atlanta, GA, USA, pp 2883–2887, 2009.
31. Ghongade, R., Deshmukh, M., and Joshi, D., Arrhythmia classification using morphological features and probabilistic neural networks. Computational Intelligence on Power, Energy and Controls with their impact on Humanity (CIPECH), 80–84, 2014.
32. Han, J., Pei, J., and Kamber, M., *Data mining: Concepts and techniques. 3rd edition*, Elsevier, 2011.
33. Larose, D. T., *Discovering knowledge in data: An introduction to data mining*. John Wiley & Sons, 2014.
34. Goshvarpour, A., Abbasi, A., and Goshvarpour, A., Fusion of heart rate variability and pulse rate variability for emotion recognition using lagged Poincare plots. *Australas. Phys. Eng. Sci. Med.* 40: 617–629, 2017.
35. Goshvarpour, A., Abbasi, A., Goshvarpour, A., Daneshvar, S., A novel signal-based fusion approach for accurate music emotion recognition. Biomedical Engineering Applications, Basis and Communications, 28: 1650040 [10 pages], 06, 2016.
36. Phongsuphap, S., and Pongsupap, Y., Analysis of heart rate variability during meditation by a pattern recognition method. *Comput. Cardiol.* 38:197–200, 2011.