Belief Rényi Divergence of Divergence and Its Application in Time Series Classification

Lang Zhang, Fuyuan Xiao, Senior Member, IEEE

Abstract—Time series data contains the amount of information to reflect the development process and state of a subject. Especially, the complexity is a valuable factor to illustrate the feature of the time series. However, it is still an open issue to measure the complexity of sophisticated time series due to its uncertainty. In this study, based on the belief Rényi divergence, a novel time series complexity measurement algorithm, called belief Rényi divergence of divergence (BRéDOD), is proposed. Specifically, the BRéDOD algorithm takes the boundaries of time series value into account. What is more, according to the Dempster-Shafer (D-S) evidence theory, the time series is converted to the basic probability assignments (BPAs) and it measures the divergence of a divergence sequence. Then, the secondary divergence of the time series is figured out to represent the complexity of the time series. In addition, the BRéDOD algorithm is applied to sets of cardiac inter-beat interval time series, which shows the superiority of the proposed method over classical machine learning methods and recent well-known works.

Index Terms—D-S evidence theory, uncertainty, belief Rényi divergence, belief divergence of divergence, complexity, time series analysis, classification

I. Introduction

COMPLEXITY measurement is a valuable approach to explore the characteristic of information in time series data [1–4]. Specifically, it reflects the development or states of a subject [5, 6]. However, measuring the complexity of time series has become a critical challenge [7, 8]. With the advent of Artificial Intelligence, deep learning methods have been heavily used for feature extraction. It has also been proved that they extract complexity features of time-series data, based on a large number of neurons [9]. While deep learning approaches can yield more sophisticated complexity measures, a significant limitation is their reduced interpretability [10, 11]. The practical value of data features is compromised if they cannot be interpreted meaningfully.

Numerous studies suggest a direct correlation between the amount of information and its complexity. As the information increases, time series often becomes more uncertain, exhibiting higher levels of entropy [12–15]. By this observation, the complexity of time series and dynamics of the system could be calculated by utilizing the uncertainty of information [16–18]. This means that it is necessary to model the uncertainty of time series in complex environments to explain its characteristics [19]. Nevertheless, without a fixed criterion, it is still an open issue to figure out the uncertainty of time series. Recently, several well-known works have been proposed to deal with uncertainty problem in various fields, such as decision making [20–23], fault diagnose [24], clustering [25, 26], reliability analysis [27, 28], pattern classification [29, 30], and so on [31–34].

Here, D-S evidence theory provides an effective way to address the uncertainty problems in an interpretable way [35–37]. It is extended in different aspects of random generalized D-S structure [38], permutation set [39, 40] and quantum evidence theory [41, 42], and widely used in output control [43], conflict management [44], evidence reasoning [45, 46], and so on [47–49]. It is noticed that D-S evidence theory offers a robust and framework for quantifying the uncertainty inherent in a time series to enable the assessment of its complexity [50, 51]. Specifically, it provides an efficacious approach for representing segments of time series as combinations of multiple subsets and singleton sets through the construction of BPAs. This methodology facilitates the extraction of segmental features of time series. Furthermore, it presents an interpretable method for measuring the complexity of time series, since BPAs illustrate the feature of uncertainty with assigning probabilities to different categories [52]. So, the interval characterization of time series data can be expressed based on D-S evidence theory reasonably.

Next, it is necessary to calculate the discrepancy of BPAs as it contributes to complexity measurement [53]. Recently, Xiao et al. [54] proposed generalized evidential divergences to measure the discrepancy between BPAs. Huang et al. [55] introduce belief f-divergence to improve the performance of complexity evaluation based on BPAs measurement. Moreover, belief Rényi divergence can quantifies the differences between BPAs for time series analysis [56, 57]. Hence, divergence shows well-performance qualities for discrepancy measurement [58–62]. Studying several belief divergence methods, belief Rényi divergence has outstanding properties to measure the discrepancy of BPAs. Specifically, the BPA degenerates to a probability distribution with all hypotheses containing one event, and the belief Rényi divergence correspondingly degenerates to a traditional Rényi divergence. Moreover, based on the parameter setting of belief Rényi divergence, it is related to the Kullback-Leibler divergence, Hellinger distance or χ^2 divergence, which leads to a more flexible belief divergence. However, there is still a limitation. For fussy

(Corresponding author: Fuyuan Xiao.)
L. Zhang and F. Xiao are with the School of Big Data and Software Engineering, Chongqing University, Chongqing 401331, China. (e-mail: xiaofuyuan@cqu.edu.cn; doctorxiaofy@hotmail.com)

time series, it is hard to extract the complexity features by using divergence measurement only once.

To address the limitation above, a novel belief divergence of divergence algorithm, called BR*é*DOD is proposed. At first, a divergence sequence of time series can be figured out according to belief R*é*nyi divergence. Next, the complexity of the time series can be calculated by measuring the discrepancy of BPAs that are produced by the divergence sequence. Based on belief divergence of divergence, the deeper feature of time series can be obtained. Furthermore, the BR*é*DOD algorithm is applied to cardiac inter-beat interval time series, which demonstrates the effectiveness in handling a real-world time series issue in complexity measurement, and a pattern classification problem is addressed.

Main contributions of this paper are listed as follows:

1) Time series data is converted to mass function by means of D-S evidence theory, which provides an interpretable approach to carry out time series feature extraction.
2) The complexity of the time series is effectively figured out according to the proposed BR*é*DOD algorithm by measuring the belief R*é*nyi divergence of divergence.
3) An application in cardiac inter-beat interval time series classification is carried out with BR*é*DOD algorithm, which shows higher classification accuracy than classical machine learning methods and several well-known works. Hence, the superiority of the proposed method has been discussed.

The organization of this study is shown as follows. Section II introduces the fundamental concepts of this study. In Section III, a novel belief R*é*nyi divergence of divergence algorithm for complexity measure of time series is proposed and a specific example is put forward. In Section IV, the processing of complexity measurement is demonstrated based on cardiac inter-beat interval time series, and an application of pattern classification is carried out to illustrate the out-performance of the BR*é*DOD algorithm. Section V makes a conclusion of this study.

II. Preliminaries

This section briefly presents the basic concepts, including D-S evidence theory and belief R*é*nyi divergence.

A. D-S Evidence Theory

D-S evidence theory is generalized by typical probability theory, which deals with uncertainty problems more effectively and flexibly.

Definition 1 (Framework of discernment).

Let Θ be the discernment with mutually exclusive events which can be defined as:

$$\Theta = \{e_1, e_2, \ldots, e_n\}. \tag{1}$$

Then, its power set 2^Θ can be defined as follows:

$$2^\Theta = \{\varnothing, \ldots, \{e_n\}, \ldots, \{e_1, e_2\}, \ldots, \Theta\}, \tag{2}$$

where $\varnothing$ indicates the empty set.

Definition 2 (Mass function).

With discernment Θ, a mass function m can be defined as:

$$m : 2^\Theta \to [0, 1]. \tag{3}$$

It abides the rule of

$$\sum_{E \in 2^\Theta} m(E) = 1 \quad \text{and} \quad m(\varnothing) = 0. \tag{4}$$

Here, E is a focal element with $m(E) > 0$. Specifically, the information value of mass function contributes to uncertainty measurement [63, 64].

Definition 3 (Dempster's rule of combination).

Let m_1 and m_2 be two BPAs. The rule of Dempster's combination is describe as:

$$m(A) = \begin{cases} \dfrac{1}{1-k} \displaystyle\sum_{P \cap Q = A} m_1(P) m_2(Q), & A \neq \varnothing, \\ 0, & A = \varnothing, \end{cases} \tag{5}$$

and

$$k = \sum_{P \cap Q = \varnothing} m_1(P) m_2(Q), \tag{6}$$

where P and Q are focal elements and k is regarded as a conflict coefficient.

B. Belief Entropy and Divergence

Definition 4 (Deng entropy).

Deng entropy is a new belief entropy in processing the measurement of conflict between evidence [37]. The Deng entropy can be calculated as:

$$E_D = -\sum_i m(E_i) \log \frac{m(E_i)}{2^{|E_i|} - 1}, \tag{7}$$

where $m(E_i)$ is a BPA, and $|\cdot|$ represents the cardinality of $m(E_i)$.

Definition 5 (Belief R*é*nyi divergence).

Let m_1 and m_2 be two BPAs. The belief R*é*nyi divergence between two mass function can be defined as [57]:

$$\mathcal{BRD}_\alpha(m_1 || m_2) = \frac{1}{\alpha - 1} \ln \sum_{i=1}^{2^\Theta} \frac{m_1(E_i)^\alpha m_2(E_i)^{1-\alpha}}{2^{|E_i|} - 1}, \tag{8}$$

where α is the parameter with $\alpha \in (0, 1) \cup (1, +\infty)$ and $|E_i|$ represents the cardinality of E_i.

III. Complexity Measurement Based on Belief R*é*nyi Divergence of Divergence

The time series complexity analysis algorithm, BR*é*DOD algorithm is divided into two components. The first component figures out the divergence sequence of two types of time series, and the second component calculates the divergence of divergence sequence. The flowchart of BR*é*DOD algorithm for time series processing is described as Fig. 1.

A. Measure the Divergence of Time Series

In the first component, the time series is described as $\mathcal{S} = \{x_1, \ldots, x_i, \ldots, x_N\}$ with length N. Two types

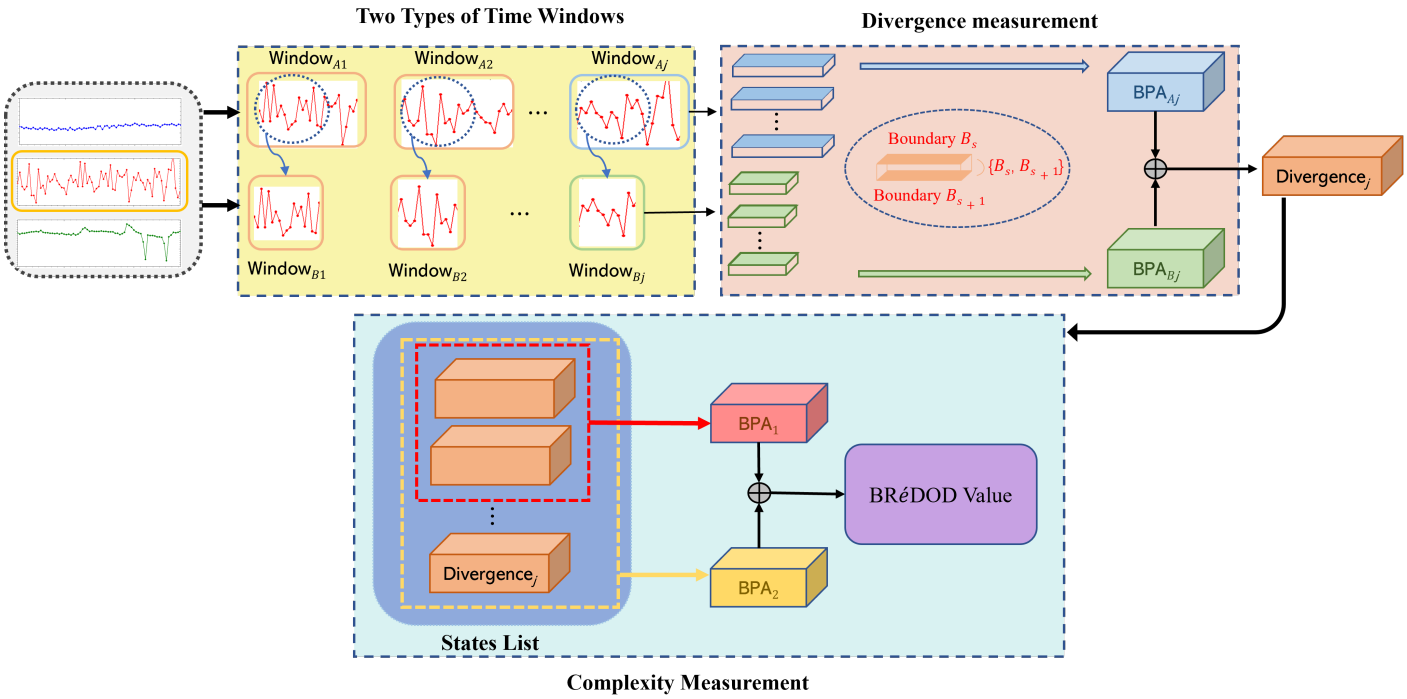


Fig. 1. Flowchart of the BRéDOD algorithm for time series complexity.

of consecutive and non-overlapping time windows are obtained by dividing the time series as type A $\left\{w_{Aj}^{(\tau)}\right\}$ and type B $\left\{w_{Bj}^{(\tau)}\right\}$. Here, $w_{Aj}^{(\tau)} = \left\{x_{(j-1)\tau+1}, \ldots, x_{(j-1)\tau+\tau}\right\}$ is of length τ, where j is the window index which ranges from 1 to N/τ. It can be regarded that type B is the truncation of type A in each window as $w_{Bj}^{(\tau)} = \left\{x_{(j-1)\tau+1}, \ldots, x_{(j-1)\tau+\upsilon}\right\}$, where $\upsilon < \tau$. The value selecting of υ and τ are discussed in Section IV.

The time interval of each window ranges from x_{min} to x_{max} where x_{min} and x_{max} indicate the lower and upper boundaries of time series $\mathcal{S}$, respectively. The time interval is equally split into k slices. Each slice contains two boundaries (e.g., B_s and B_{s+1} of the sth slice), which represents the specific state. If data points are in the same slice, then it can be considered that they are in the same states.

Definition 6 (Belief divergence sequence).

As depicted in Fig. 1, each time window is conceptualized as a discernment framework, where each slice comprises two focal elements: one signifies the specific state, while the other denotes the state of uncertainty.

Let the total number of data points over w_{ij} between B_s and B_{s+1} be p_{ij}. Specifically, let data points γ of length q_{ij} fall on the border B_s coincidentally. The BPAs based on each time window can be defined as:

$$m_{ij}: \begin{cases} m_{ij}^{(\tau)}(\{B_s\}) = \dfrac{q_{ij}}{|w_{ij}|}, & \text{if } \gamma \text{ falls on the boundary } s, \\ m_{ij}^{(\tau)}(\{B_s, B_{s+1}\}) = \dfrac{p_{ij}}{|w_{ij}|}, & \text{otherwise,} \end{cases} \tag{9}$$

where $i \in \{A, B\}$. Considering that all points are in the interval, the divergence $Div_j^{(\tau)}$ in each corresponding window is figured out based on Eq. (8):

$$\begin{aligned} Div_j^{(\tau)} &= \mathcal{BRD}_\alpha(m_{Aj}||m_{Bj}) \\ &= \frac{1}{\alpha-1}\ln\sum_{i=1}^{k}\left(\frac{m_{Aj}\,(E_i)^\alpha\, m_{Bj}\,(E_i)^{1-\alpha}}{2^{|E_i|}-1}\right), \end{aligned} \tag{10}$$

where E_i represents $\{B_i\}$ or $\{B_i, B_{i+1}\}$. Then a belief Rényi divergence sequence $\left\{Div_j^{(\tau)}\right\}$ of original time series is constructed.

The belief Rényi divergence is a generalized divergence. Choosing different α, the equation will degenerate to special divergence expression, including Hellinger distance, Kullback-Leibler divergence and χ^2 divergence [57]:

- When $\alpha \to \infty$, $Div_j^{(\tau)}$ degenerates the difference value between the logarithms of one focal element of BPA.
- When $\alpha = \frac{1}{2}$, $Div_j^{(\tau)}$ has the property of symmetry. Specifically, $Div_j^{(\tau)}$ degenerates an equation related to Hellinger distance with $\forall|E_i| = 1$.
- When $\alpha \to 1$, $Div_j^{(\tau)}$ degenerates an equation of Kullback-Leibler divergence with $\forall|E_i| = 1$.
- When $\alpha = 2$, $Div_j^{(\tau)}$ degenerates an equation related to the χ^2 divergence with $\forall|E_i| = 1$.
- When $\alpha \to 0$, $Div_j^{(\tau)} = 0$ with $\forall|E_i| = 1$.

Utilizing different α, divergence values will be obtained during calculation, which leads to a more flexible complexity measurement algorithm. The following proof shows that these properties still hold true in Eq. (10).

- Property 1. When all the data points fall on the

boundaries of time window, the $Div_j^{(\tau)}$ degenerates to:

$$Div_j^{(\tau)} = \frac{1}{\alpha-1}\ln\sum_{i=1}^{k}\left(\frac{q_{Aj}}{|w_{Aj}|}\right)^{\alpha}\cdot\left(\frac{q_{Bj}}{|w_{Bj}|}\right)^{1-\alpha}. \tag{11}$$

Proof. Given two BPAs m_{Aj} and m_{Bj}. According to Eq. (9), each focal element in BPAs is a singleton.

$$\begin{aligned} Div_j^{(\tau)} &= \mathcal{BRD}_\alpha(m_{Aj}||m_{Bj}) \\ &= \frac{1}{\alpha-1}\ln\sum_{s=1}^{k}\frac{m_{Aj}\left(\{B_s\}\right)^{\alpha}m_{Bj}\left(\{B_s\}\right)^{1-\alpha}}{2^{|\{B_s\}|}-1} \\ &= \frac{1}{\alpha-1}\ln\sum_{s=1}^{k}\left(\frac{q_{Aj}}{|w_{Aj}|}\right)^{\alpha}\cdot\left(\frac{q_{Bj}}{|w_{Bj}|}\right)^{1-\alpha}. \end{aligned}$$

■

• Property 2. When $\alpha\to\infty$, the $Div_j^{(\tau)}$ is regarded as:

$$Div_j^{(\tau)} = \ln\frac{m_{Aj}(E_\mu)}{m_{Bj}(E_\mu)}, \tag{12}$$

where $\mu = \arg\max\limits_{1\le i\le k}\left\{\frac{m_{Bj}(E_i)}{2^{|E_i|}-1}\left(\frac{m_{Aj}(E_i)}{m_{Bj}(E_i)}\right)^{\alpha}\right\}$.

Proof. $m = \max_i\left\{\frac{m_{Bj}(E_i)}{2^{|E_i|}-1}\cdot\left(\frac{m_{Aj}(E_i)}{m_{Bj}(E_i)}\right)^{\alpha}\right\}$ and $\mu = \arg\max_{1\leqslant i\leqslant k}\left\{\frac{m_{Bj}(E_i)}{2^{|E_i|}-1}\cdot\left(\frac{m_{Aj}(E_i)}{m_{Bj}(E_i)}\right)^{\alpha}\right\}$.

$$\frac{1}{\alpha-1}\ln m \le Div_j^{(\tau)} \le \frac{1}{\alpha-1}(\ln m + \ln n)$$

That is:

$$\begin{aligned} Div_j^{(\tau)} &\geqslant \lim_{\alpha\to\infty}\frac{1}{\alpha-1}\ln\max_i\left\{\frac{m_{Bj}(E_i)}{2^{|E_i|}-1}\cdot\left(\frac{m_{Aj}(E_i)}{m_{Bj}(E_i)}\right)^{\alpha}\right\} \\ &= \lim_{\alpha\to\infty}\frac{1}{\alpha-1}\ln\frac{m_{Bj}(E_\mu)}{2^{|E_\mu|}-1} + \lim_{\alpha\to\infty}\frac{1}{\alpha-1}\ln\left(\frac{m_{Aj}\left(E_\mu\right)}{m_{Bj}\left(E_\mu\right)}\right)^{\alpha} \\ &= 0 + \lim_{\alpha\to\infty}\frac{\alpha}{\alpha-1}\ln\frac{m_{Aj}\left(E_\mu\right)}{m_2\left(E_\mu\right)} \\ &= \ln\frac{m_{Aj}\left(E_\mu\right)}{m_{Bj}\left(E_\mu\right)}. \end{aligned}$$

Similarly,

$$\begin{aligned} &Div_j^{(\tau)} \\ &\leqslant \lim_{\alpha\to\infty}\frac{1}{\alpha-1}\left(\ln n + \ln\max_i\left\{\frac{m_{Bj}(E_i)}{2^{|E_i|}-1}\cdot\left(\frac{m_{Aj}(E_i)}{m_{Bj}(E_i)}\right)^{\alpha}\right\}\right) \\ &= \lim_{\alpha\to\infty}\frac{\ln n}{\alpha-1} + \lim_{\alpha\to\infty}\frac{1}{\alpha-1}\ln\frac{m_{Bj}(E_\mu)}{2^{|A\mu|}-1} + \lim_{\alpha\to\infty}\frac{\alpha}{\alpha-1}\ln\frac{m_{Aj}(E_\mu)}{m_{Bj}(E_\mu)} \\ &= 0 + 0 + \lim_{\alpha\to\infty}\frac{\alpha}{\alpha-1}\ln\frac{m_{Aj}\left(E_\mu\right)}{m_{Bj}\left(E_\mu\right)} \\ &= \ln\frac{m_{Aj}\left(E_\mu\right)}{m_{Bj}\left(E_\mu\right)}. \end{aligned}$$

Therefore, $\ln\frac{m_{Aj}(E_\mu)}{m_{Bj}(E_\mu)} \leqslant Div_j^{(\tau)} \leqslant \ln\frac{m_{Aj}(E_\mu)}{m_{Bj}(E_\mu)}$. So, $Div_j^{(\tau)} = \ln\frac{m_{Aj}(A_\mu)}{m_{Bj}(A_\mu)}$. ■

• Property 3. When $\alpha=\frac{1}{2}$, the $Div_j^{(\tau)}$ follows the rule:

$$Div_j^{(\tau)} = \mathcal{BRD}_\alpha(m_{Aj}||m_{Bj}) = \mathcal{BRD}_\alpha(m_{Bj}||m_{Aj}), \tag{13}$$

which means that it has the property of symmetry.

Proof. Given two BPAs m_{Aj} and m_{Bj}, according to Eq. (10):

$$\mathcal{BRD}_\alpha(m_{Aj}||m_{Bj}) = -2\ln\sum_{i=1}^{k}\left(\frac{\sqrt{m_{Aj}\left(E_i\right)m_{Bj}\left(E_i\right)}}{2^{|E_i|}-1}\right).$$

$$\mathcal{BRD}_\alpha(m_{Bj}||m_{Aj}) = -2\ln\sum_{i=1}^{k}\left(\frac{\sqrt{m_{Bj}\left(E_i\right)m_{Aj}\left(E_i\right)}}{2^{|E_i|}-1}\right).$$

Therefore,

$$\mathcal{BRD}_\alpha(m_{Aj}||m_{Bj}) = \mathcal{BRD}_\alpha(m_{Bj}||m_{Aj}).$$

■

Specifically, when all the data points fall on the boundaries, the belief Rényi divergence is related to the Hellinger distance. Here, $Div_j^{(\tau)}$ is regarded as:

$$Div_j^{(\tau)} = -2\ln\left(\frac{2-Hel^2(m_{Aj},m_{Bj})}{2}\right). \tag{14}$$

Proof. Given two BPAs m_{Aj} and m_{Bj}. According to Eq. (9), each focal element in BPAs is a singleton.

$$\begin{aligned} &Div_j^{(\tau)} \\ &= -2\ln\sum_{i=1}^{k}\frac{\sqrt{m_{Aj}\left(E_i\right)m_{Bj}\left(E_i\right)}}{2^{|E_i|}-1} \\ &= -2\ln\sum_{i=1}^{k}\frac{m_{Aj}\left(E_i\right)+m_{Bj}\left(E_i\right)-\left(m_{Aj}\left(E_i\right)^{\frac{1}{2}}-m_{Bj}\left(E_i\right)^{\frac{1}{2}}\right)^2}{2} \\ &= -2\ln\left(\frac{2-Hel^2\left(m_{Aj},m_{Bj}\right)}{2}\right), \end{aligned}$$

where

$$Hel^2\left(m_{Aj},m_{Bj}\right) = \sum_{i=1}^{k}\left(m_{Aj}\left(E_i\right)^{\frac{1}{2}}-m_{Bj}\left(E_i\right)^{\frac{1}{2}}\right)^2.$$

■

• Property 4. When $\alpha\to 1$ and all the data points fall on the boundaries, the belief Rényi divergence degenerates to a Kullback-Leibler divergence. Here, $Div_j^{(\tau)}$ is regarded as:

$$Div_j^{(\tau)} = \mathcal{D}_{\mathcal{KL}}(m_{Aj}||m_{Bj}). \tag{15}$$

Proof. Given two BPAs m_{Aj} and m_{Bj}. According to Eq. (9), each focal element in BPAs is a singleton.

$$\begin{aligned} Div_j^{(\tau)} &= \lim_{\alpha\to 1}\frac{\frac{\partial}{\partial\alpha}\left[\ln\sum_{i=1}^{k}m_{Aj}\left(E_i\right)^{\alpha}m_{Bj}\left(E_i\right)^{1-\alpha}\right]}{\frac{\partial}{\partial\alpha}(\alpha-1)} \\ &= \sum_{i=1}^{k}m_{Aj}(E_i)(\ln m_{Aj}(E_i) - \ln m_{Bj}(E_i)) \\ &= \sum_{i=1}^{n}m_{Aj}(E_i)\ln\frac{m_{Aj}(E_i)}{m_{Bj}(E_i)} \\ &= \mathcal{D}_{\mathcal{KL}}(m_{Aj}||m_{Bj}). \end{aligned}$$

■

• Property 5. When $\alpha=2$ and all the data points fall on the boundaries, the belief Rényi divergence is the function of χ^2 divergence. The $Div_j^{(\tau)}$ can be expressed as:

$$Div_j^{(\tau)} = \ln\left(1+\chi^2(m_{Aj},m_{Bj})\right). \tag{16}$$

Proof. Given two BPAs m_{Aj} and m_{Bj}. According to

Eq. (9), each focal element in BPAs is a singleton.

$$\begin{aligned} Div_j^{(\tau)} &= \ln \sum_{i=1}^{n} \frac{m_{Aj}(E_i)^2}{m_{Bj}(E_i)} \cdot \frac{1}{2^{|E_i|}-1} \\ &= \ln \sum_{i=1}^{n} \left(2m_{Aj}(E_i) - m_{Bj}(E_i) + \frac{\left(m_{Aj}(E_i) - m_{Bj}(E_i)\right)^2}{m_{Bj}(E_i)} \right) \\ &= \ln \left(1 + \chi^2(m_{Aj}, m_{Bj})\right). \end{aligned}$$

■

• Property 6. When $\alpha \to 0$, and all data points fall on the boundaries, the $Div_j^{(\tau)}$ is regarded as:

$$Div_j^{(\tau)} = 0. \tag{17}$$

Proof. Given two BPAs m_{Aj} and m_{Bj}. According to Eq. (9), each focal element in BPAs is a singleton.

$$\begin{aligned} Div_j^{(\tau)} &= -\ln \sum_{i=1}^{n} \frac{m_{Bj}(E_i)}{2^{|E_i|}-1} \\ &= -\ln \sum_{i=1}^{n} m_{Bj}(A_i) \\ &= 0. \end{aligned}$$

■

B. Measure the Complexity of Time Series

In the second component, belief Rényi divergence is also used to measure the complexity of time series by calculating the divergence of sequence $\left\{Div_j^{(\tau)}\right\}$. Here, the divergence sequence is processed as two parts. The first part is the first half of the sequence, and the second part is the entire sequence.

Definition 7 (Belief divergence of divergence).

As illustrated in Fig. 1, the first half segment of the state list, and the whole state list serves as distinct discernment frameworks. Specifically, each individual state is construed as a focal element within these frameworks.

Let r be the quantity in the lth state of the sequence and there are $\mathcal{L}$ states. So, the BPAs based on the belief Rényi divergence is described as follows:

$$\begin{cases} m_{\mathcal{S}_1}(\{B_l\}) = \dfrac{2r}{N/\tau}, \\ m_{\mathcal{S}_2}(\{B_l\}) = \dfrac{r}{N/\tau}, \end{cases} \tag{18}$$

where $m_{\mathcal{S}_1}$ and $m_{\mathcal{S}_2}$ represent the BPAs of the first part and second part of sequence, separately. In this case, the complexity of the whole time series $\Phi^{(\tau)}$ can be obtained by calculating the divergence between $m_{\mathcal{S}_1}$ and $m_{\mathcal{S}_2}$ with belief Rényi divergence as follows:

$$\begin{aligned} \Phi^{(\tau)} &= \mathcal{BRD}_\alpha(m_{\mathcal{S}_1} || m_{\mathcal{S}_2}) \\ &= \frac{1}{\alpha - 1} \ln \sum_{l=1}^{\mathcal{L}} \frac{m_{\mathcal{S}_1}(\{B_l\})^\alpha \, m_{\mathcal{S}_2}(\{B_l\})^{1-\alpha}}{2^{|\{B_s\}|}-1} \\ &= \frac{1}{\alpha - 1} \ln \sum_{l=1}^{\mathcal{L}} \left(\frac{2r}{N/\tau}\right)^\alpha \cdot \left(\frac{r}{N/\tau}\right)^{1-\alpha}. \end{aligned} \tag{19}$$

Note that each element in sequence is at boundary which means that each focal element in BPAs is a singleton.

Algorithm 1: Complexity analysis algorithm for time series based on Rényi divergence

Input: Time series $\mathcal{S} = \{x_1, \ldots, x_N\}$;
Output: Complexity result $\Phi^{(\tau)}$;
Split the time series $\mathcal{S}$ into two types of time windows $\left\{w_{Aj}^{(\tau)}\right\}$ and $\left\{w_{Bj}^{(\tau)}\right\}$;
Determine the lower and upper sides of time interval $\{x_{min}, x_{max}\}$;
Divided each time window into k slices;
Count the number of data points on or between boundaries;
for i=1;i$\leq N/\tau$ do
  Figure out the BPAs m_{Ai} and m_{Bi} of each time window by using the Eq. (9);
end
for i=1;i$\leq N/\tau$ do
  Calculate the divergence $Div_j^{(\tau)}$ in each corresponding window by using Eq. (10);
end
Generate the BPAs $m_{\mathcal{S}_1}$ and $m_{\mathcal{S}_2}$ of the divergence sequence with Eq. (18);
Calculate the complexity of time series $\Phi^{(\tau)}$ by using Eq. (19);
return $\Phi^{(\tau)}$.

Here, $\Phi^{(\tau)}$ is to be computed under different parameter α of belief Rényi divergence and time scales τ.

• Property 7. When $\alpha \to \infty$, based on Eq. (12), the $\Phi^{(\tau)}$ is regarded as:

$$\Phi^{(\tau)} = \ln \frac{m_{\mathcal{S}_1}(E_\mu)}{m_{\mathcal{S}_2}(E_\mu)}, \tag{20}$$

where $\mu = \arg \max_{1 \leq i \leq n} \left\{ \frac{m_{\mathcal{S}_1}(E_i)}{2^{|E_i|}-1} \left(\frac{m_{\mathcal{S}_1}(E_i)}{m_{\mathcal{S}_2}(E_i)}\right)^\alpha \right\}$.

• Property 8. When $\alpha = \frac{1}{2}$, based on Eqs. (13)-(14), the $\Phi^{(\tau)}$ follows the rule:

$$\Phi^{(\tau)} = \mathcal{BRD}_\alpha(m_{\mathcal{S}_1} || m_{\mathcal{S}_2}) = \mathcal{BRD}_\alpha(m_{\mathcal{S}_2} || m_{\mathcal{S}_1}). \tag{21}$$

$$\Phi^{(\tau)} = -2 \ln \left(\frac{2 - Hel^2(m_{\mathcal{S}_1}, m_{\mathcal{S}_2})}{2} \right). \tag{22}$$

• Property 9. When $\alpha \to 1$, based on Eq. (15), $\Phi^{(\tau)}$ is regarded as:

$$\begin{aligned} \Phi^{(\tau)} &= \mathcal{D}_{\mathcal{KL}}(m_{\mathcal{S}_1} || m_{\mathcal{S}_2}) \\ &= \sum_{i=1}^{n} m_{\mathcal{S}_1}(E_i) \ln \frac{m_{\mathcal{S}_1}(E_i)}{m_{\mathcal{S}_2}(E_i)}. \end{aligned} \tag{23}$$

• Property 10. When $\alpha = 2$, based on Eq. (16), the $\Phi^{(\tau)}$ can be expressed as:

$$\begin{aligned} \Phi^{(\tau)} &= \ln \left(1 + \chi^2(m_{\mathcal{S}_1}, m_{\mathcal{S}_2})\right) \\ &= \ln \left(1 + \sum_{i=1}^{n} \frac{\left(m_{\mathcal{S}_1}(E_i) - m_{\mathcal{S}_2}(E_i)\right)^2}{m_{\mathcal{S}_2}(E_i)} \right). \end{aligned} \tag{24}$$

• Property 11. When $\alpha \to 0$, based on Eq. (17), the $\Phi^{(\tau)}$ is regarded as:

$$\Phi^{(\tau)} = 0. \tag{25}$$

The pseudocode of time series complexity analysis algorithm based on belief Rényi divergence is shown in Algorithm 1.

IV. Application in Cardiac Inter-beat Interval Time Series Classification

In this section, an implementation of BR*é*DOD algorithm for a practical classification problem is carried out based on cardiac inter-beat interval time series data.

A. Data Description and Processing

In this study, the cardiac inter-beat interval time series are taken into consideration, which comes from PhysioNet (https://physionet.org). Three different databases of long-term ECG (Electrocardiography) with 20-24 hours records are selected as follows:

1) BIDMC Congestive Heart Failure Database (CHF;)
2) MIT-BIH Normal Sinus Rhythm Database (Healthy);
3) Long Term AF Database (AF).

The databases contain 15, 18, and 84 subjects, respectively. As for CHF and Healthy databases, each subject is truncated into five-time series by choosing the first 500 data points from every 10,000 data points. As for AF databases, time series of onset stages are extracted according to the annotation in PhysioNet. Then, 75 records are adopted whose lengths exceed 500 points. Hence, 240 cardiac inter-beat interval time series are chosen from 117 subjects, where 75, 90, and 75 are from CHF, Healthy, and AF, respectively.

Next, several parameters are set by processing the original data. First, all the data points $\{x_i\}$ are ranked and divided into 1,000 segments. In order to avoid the noise influence and detection error, the $1st$ and $999th$ of 1,000-quantiles of the ranked segments are taken into consideration as $x_{min} = 0.3$ and $x_{max} = 1.6$. This interval is divided equally into 55 parts with $k = 55$. Second, to briefly demonstrate the performance of BR*é*DOD model, 140 data points in each time series are chosen with $N = 140$, and the time window scale are considered as $\tau = 10$ and $\upsilon = 5$. So, there are 14-time windows $(\frac{N}{\tau})$ in total for each time series. At the same time, the parameter α in belief Rényi divergence varies to show the performance adequately. In addition, the influence of parameters will be discussed in the following sections.

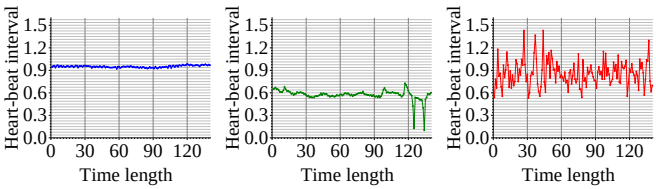


Fig. 2. The original time series with 140 data points of specific instances. (a) Original time series of CHF subject. (b) Original time series of Healthy subject. (c) Original time series of AF subject.

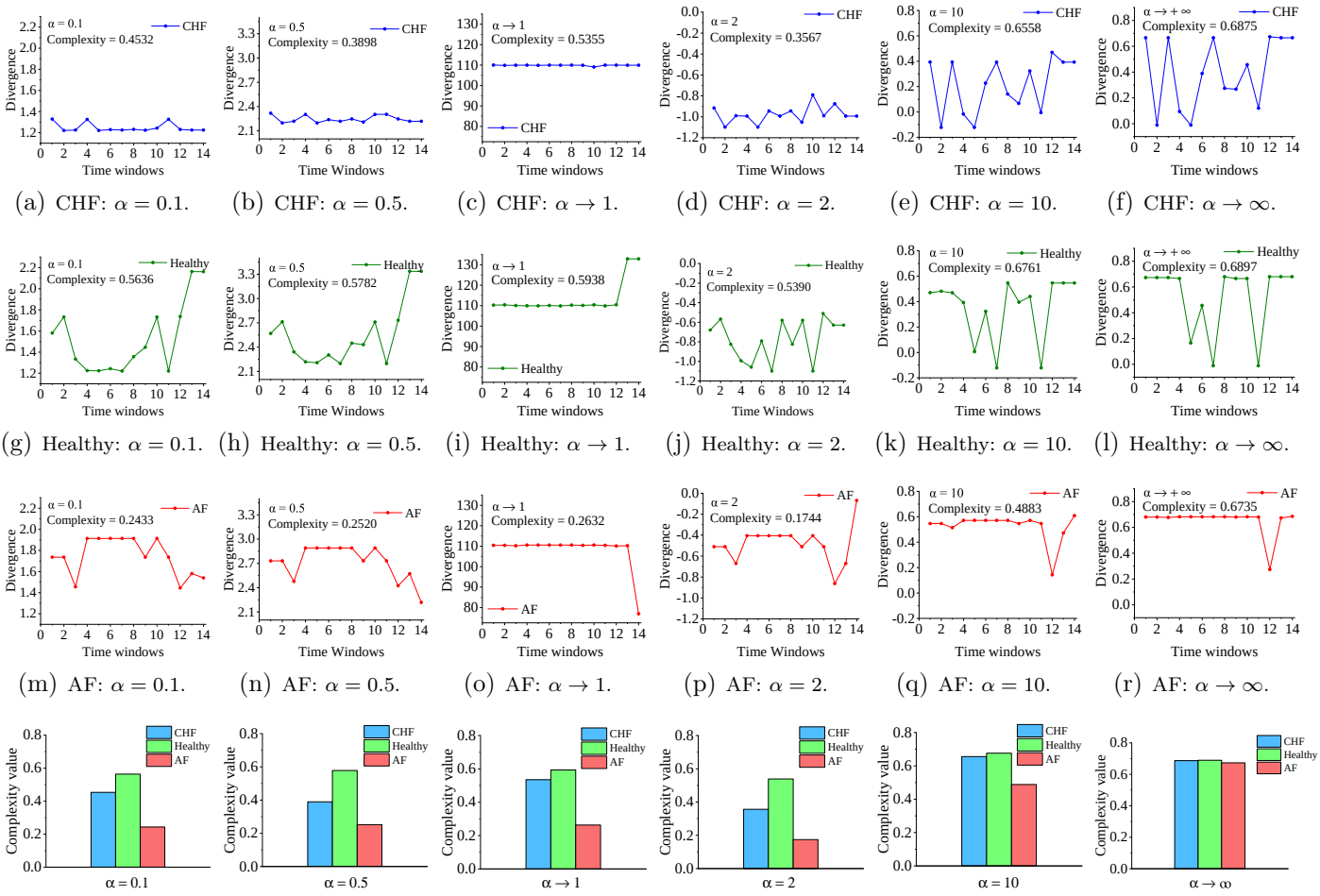
B. Implementation of the BR*é*DOD Algorithm

Three representative time series are carried out to show the process of BR*é*DOD algorithm, and final complexity values are figured out.

In the beginning, the original time series of CHF, Healthy, and AF subjects are shown in Fig. 2 based on the annotations. Note that each subject has different heart-beat rate profiles. Compared with Healthy subjects, the

TABLE I
The divergence sequence of CHF, Healthy and AF in each time window.

	$\alpha = 0.1$			$\alpha = 0.5$			$\alpha \to 1$			$\alpha = 2$			$\alpha = 10$			$\alpha \to +\infty$		
	CHF	Healthy	AF	CHF	Healthy	AF	CHF	Healthy	AF	CHF	Healthy	AF	CHF	Healthy	AF	CHF	Healthy	AF
Win_1	1.329	1.580	1.736	2.318	2.569	2.731	109.999	110.248	110.415	-0.916	-0.680	-0.510	0.393	0.471	0.546	0.666	0.673	0.680
Win_2	1.221	1.733	1.736	2.197	2.713	2.731	109.861	110.382	110.415	-1.099	-0.568	-0.511	-0.122	0.481	0.546	-0.011	0.673	0.680
Win_3	1.225	1.333	1.455	2.218	2.340	2.478	109.907	110.045	110.195	-0.990	-0.825	-0.671	0.3930	0.469	0.514	0.666	0.673	0.677
Win_4	1.326	1.224	1.914	2.303	2.217	2.890	109.967	109.905	110.554	-0.993	-0.993	-0.406	-0.017	0.393	0.571	0.094	0.666	0.682
Win_5	1.221	1.223	1.914	2.197	2.207	2.890	109.861	109.881	110.554	-1.099	-1.059	-0.406	-0.122	0.005	0.571	-0.011	0.166	0.682
Win_6	1.229	1.243	1.914	2.238	2.303	2.890	109.942	110.052	110.554	-0.945	-0.791	-0.406	0.227	0.323	0.571	0.389	0.457	0.682
Win_7	1.224	1.221	1.914	2.217	2.197	2.890	109.905	109.861	110.554	-0.993	-1.099	-0.406	0.393	-0.122	0.571	0.666	-0.011	0.682
Win_8	1.231	1.357	1.914	2.247	2.449	2.890	109.952	110.230	110.554	-0.945	-0.580	-0.406	0.141	0.546	0.571	0.274	0.680	0.682
Win_9	1.223	1.446	1.736	2.208	2.431	2.731	109.884	110.106	110.415	-1.052	-0.825	-0.511	0.067	0.395	0.546	0.267	0.666	0.680
Win_{10}	1.243	1.732	1.914	2.303	2.710	2.890	110.052	110.376	110.554	-0.791	-0.580	-0.406	0.323	0.440	0.571	0.457	0.666	0.682
Win_{11}	1.326	1.221	1.736	2.303	2.197	2.731	109.968	109.861	110.415	-0.990	-1.099	-0.511	-0.005	-0.122	0.546	0.120	-0.011	0.680
Win_{12}	1.230	1.736	1.445	2.247	2.731	2.424	109.965	110.415	110.091	-0.876	-0.511	-0.862	0.469	0.546	0.143	0.673	0.680	0.274
Win_{13}	1.224	2.163	1.581	2.217	3.337	2.572	109.905	132.869	110.253	-0.993	-0.629	-0.671	0.393	0.546	0.472	0.666	0.680	0.673
Win_{14}	1.224	2.162	1.540	2.217	3.337	2.217	109.905	132.869	76.907	-0.993	-0.629	-0.069	0.393	0.546	0.609	0.666	0.680	0.685



(a) CHF: $\alpha = 0.1$. (b) CHF: $\alpha = 0.5$. (c) CHF: $\alpha \to 1$. (d) CHF: $\alpha = 2$. (e) CHF: $\alpha = 10$. (f) CHF: $\alpha \to \infty$.

(g) Healthy: $\alpha = 0.1$. (h) Healthy: $\alpha = 0.5$. (i) Healthy: $\alpha \to 1$. (j) Healthy: $\alpha = 2$. (k) Healthy: $\alpha = 10$. (l) Healthy: $\alpha \to \infty$.

(m) AF: $\alpha = 0.1$. (n) AF: $\alpha = 0.5$. (o) AF: $\alpha \to 1$. (p) AF: $\alpha = 2$. (q) AF: $\alpha = 10$. (r) AF: $\alpha \to \infty$.

(s) Complexity: $\alpha = 0.1$. (t) Complexity: $\alpha = 0.5$. (u) Complexity: $\alpha \to 1$. (v) Complexity: $\alpha = 2$. (w) Complexity: $\alpha = 10$. (x) Complexity: $\alpha \to \infty$.

Fig. 3. Demonstration of BR*é*DOD algorithm for time series on three representative subjects with varied α.

heart-beat interval of CHF subjects tends to be more stable, while that of AF subjects varies considerably. In this case, there will be disparate features in different subjects by figuring out their complexity values.

According to Eqs. (9)-(17), Table I and Fig. (3a)-(3r) show divergence sequences of three representative subjects with varied α. After the first component of the BR*é*DOD algorithm, note that each time series has its own feature and the information of each time window is extracted. The divergence value shows the inner discrepancy during a period of time. Specifically, the divergence sequence of CHF subjects is always flat while that of Healthy and AF subjects rises and down. Moreover, the divergence of Healthy sequence fluctuates at a lower divergence value and that of AF subjects fluctuates at a higher divergence value. When parameter α tends to 1 and infinite, it can be found that there is a little discrepancy in the divergence sequence of time series.

Then, based on Eqs. (18)-(25), the complexity value of each time series is obtained after the second component of the BR*é*DOD algorithm. It can be found that the proposed method measures the complexity of inter-beat interval time series reasonably because healthy people have more sophisticated abilities to regulate control of the heart while pathological groups are poor in regulation and control. In this case, the complexity value of Healthy subjects is supposed to be larger than that of CHF and AF subjects. Specifically, when α fixes at 0.5, the secondary BPAs of divergence sequence of CHF, Healthy, and AF subjects, and complexity results are shown in Table II. According to Figs. (3v)-(3x), the complexity value of the time series tends to be the same with the increasing of α, which means that it is hard to distinguish among different time series when α is too large.

C. Classification in Cardiac Inter-beat Interval Time Series Based on BR*é*DOD

In this part, the classification accuracy in time series will be figured out to demonstrate the effective and efficient performance of the proposed method. When $\upsilon = 5$, $\tau = 10$, and $k = 55$, the evaluation criteria, sensitivity of pathological subjects, and specificity of healthy subjects, can be defined as follows:

TABLE II
The BPAs and complexity value of three instances.

Subject	BPA	Complexity
CHF	$m_{\mathcal{S}_1}(B_0)=\frac{1}{7}, m_{\mathcal{S}_1}(B_1)=\frac{2}{7}, m_{\mathcal{S}_1}(B_2)=\frac{1}{2}, m_{\mathcal{S}_1}(B_3)=\frac{1}{7}, m_{\mathcal{S}_1}(B_5)=\frac{1}{7}, m_{\mathcal{S}_1}(B_6)=\frac{1}{7}$ $m_{\mathcal{S}_2}(B_0)=\frac{1}{14}, m_{\mathcal{S}_2}(B_1)=\frac{2}{14}, m_{\mathcal{S}_2}(B_2)=\frac{1}{14}, m_{\mathcal{S}_2}(B_3)=\frac{2}{14}, m_{\mathcal{S}_2}(B_5)=\frac{1}{14}$ $m_{\mathcal{S}_2}(B_6)=\frac{3}{14}, m_{\mathcal{S}_2}(B_7)=\frac{1}{7}, m_{\mathcal{S}_2}(B_8)=\frac{1}{14}, m_{\mathcal{S}_2}(B_{10})=\frac{1}{14}$	0.3898
Healthy	$m_{\mathcal{S}_1}(B_0)=\frac{1}{7}, m_{\mathcal{S}_1}(B_1)=\frac{1}{7}, m_{\mathcal{S}_1}(B_2)=\frac{1}{2}, m_{\mathcal{S}_1}(B_3)=\frac{1}{7}, m_{\mathcal{S}_1}(B_4)=\frac{1}{7}, m_{\mathcal{S}_1}(B_5)=\frac{1}{7}, m_{\mathcal{S}_1}(B_6)=\frac{1}{7}$ $m_{\mathcal{S}_2}(B_0)=\frac{1}{14}, m_{\mathcal{S}_2}(B_1)=\frac{1}{14}, m_{\mathcal{S}_2}(B_2)=\frac{1}{14}, m_{\mathcal{S}_2}(B_3)=\frac{1}{14}, m_{\mathcal{S}_2}(B_4)=\frac{1}{14}, m_{\mathcal{S}_2}(B_5)=\frac{1}{14}$ $m_{\mathcal{S}_2}(B_6)=\frac{1}{7}, m_{\mathcal{S}_2}(B_7)=\frac{1}{14}, m_{\mathcal{S}_2}(B_8)=\frac{1}{14}, m_{\mathcal{S}_2}(B_9)=\frac{1}{14}, m_{\mathcal{S}_2}(B_{11})=\frac{1}{14}, m_{\mathcal{S}_2}(B_{12})=\frac{1}{7}$	0.5782
AF	$m_{\mathcal{S}_1}(B_0)=\frac{2}{7}, m_{\mathcal{S}_1}(B_2)=\frac{1}{7}, m_{\mathcal{S}_1}(B_3)=\frac{4}{7}$ $m_{\mathcal{S}_2}(B_0)=\frac{2}{7}, m_{\mathcal{S}_2}(B_2)=\frac{1}{14}, m_{\mathcal{S}_2}(B_3)=\frac{6}{14}, m_{\mathcal{S}_2}(B_{11})=\frac{1}{14}, m_{\mathcal{S}_2}(B_{12})=\frac{1}{14}, m_{\mathcal{S}_2}(B_{13})=\frac{1}{14}$	0.2520

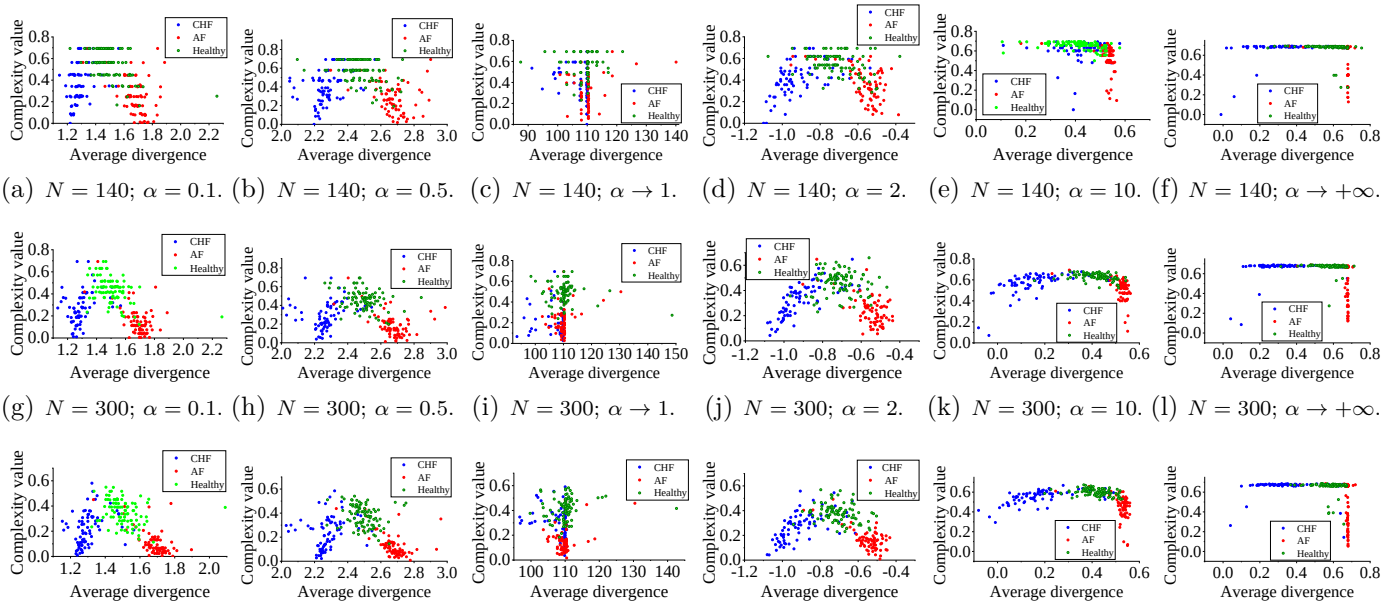
$$\begin{aligned}
&\text{Specificity: } V_{sp}=\frac{T_H}{T_H+F_H},\\
&\text{Sensitivity: } V_{se}=\frac{T_P}{T_P+F_P},\\
&\text{Accuracy: } V_{ac}=\frac{T_H+T_P}{T_H+F_H+T_P+F_P},
\end{aligned} \tag{26}$$

where T_H and F_H represent the amount of healthy subjects that are classified correctly and falsely, respectively. Besides, T_P and F_P represent the amount of pathology subjects that are classified correctly and falsely. In particular, the BR*é*DOD algorithm places its emphasis on feature extraction, rendering it highly adaptable to various classification models. This inherent flexibility enhances the algorithm's compatibility with different approaches.

In this section, a classification method known as Spectral clustering was employed to partition time series data-set into three distinct categories. This classification is based on the analysis of two-dimensional eigenvalues obtained through BR*é*DOD, specifically considering the average divergence sequence and complexity value.

In this case, 240 subjects of cardiac inter-beat interval time series are taken into consideration to show the performance of BR*é*DOD algorithm. Here, data points $N=140, N=300$, and $N=500$ with varied α are used to demonstrate the effect of time series length.

According to Eq. (26), the specificity, sensitivity, and accuracy in different lengths of time series are figured out. Based on Table III, it can be found that different lengths of time series will affect the result. There will



(a) $N=140;\ \alpha=0.1.$ (b) $N=140;\ \alpha=0.5.$ (c) $N=140;\ \alpha\to 1.$ (d) $N=140;\ \alpha=2.$ (e) $N=140;\ \alpha=10.$ (f) $N=140;\ \alpha\to+\infty.$

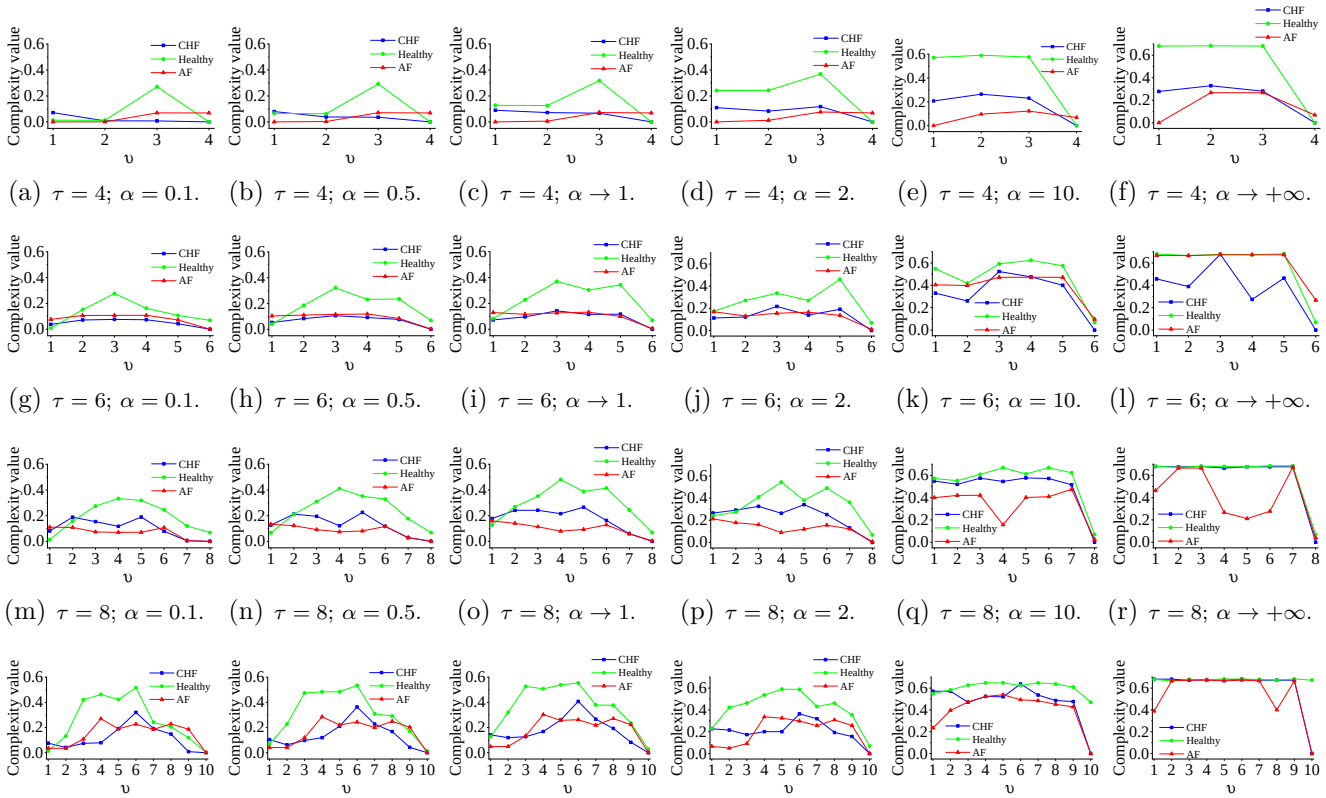
(g) $N=300;\ \alpha=0.1.$ (h) $N=300;\ \alpha=0.5.$ (i) $N=300;\ \alpha\to 1.$ (j) $N=300;\ \alpha=2.$ (k) $N=300;\ \alpha=10.$ (l) $N=300;\ \alpha\to+\infty.$

(m) $N=500;\ \alpha=0.1.$ (n) $N=500;\ \alpha=0.5.$ (o) $N=500;\ \alpha\to 1.$ (p) $N=500;\ \alpha=2.$ (q) $N=500;\ \alpha=10.$ (r) $N=500;\ \alpha\to+\infty.$

Fig. 4. The scatter plot with the relationship between belief Rényi divergence and complexity value in different time series length.

TABLE III
The sensitivity, specificity and accuracy of the application based on BR*é*DOD.

	$\alpha = 0.1$			$\alpha = 0.5$			$\alpha = 1$			$\alpha = 2$			$\alpha = 10$			$\alpha \to +\infty$		
Data points N	140	300	500	140	300	500	140	300	500	140	300	500	140	300	500	140	300	500
V_{se} in CHF subjects	.6800	.7467	.8133	.6933	.7467	.6533	.0000	.0000	.0000	.6667	.6667	.6400	.8400	.0800	.8267	.7600	.8000	.8133
V_{se} in AF subjects	.8533	.8933	.9200	.8267	.9067	.9067	.0000	.0000	.0533	.8267	.9067	.9200	.0800	.7600	.4667	.2533	.3600	.4000
V_{sp} in Healthy subjects	.8222	.9000	.8111	.8444	.9000	.9111	.9667	.9889	.9556	.7889	.8111	.8889	.8667	.9333	.9556	.8889	.9556	.9111
Accuracy	.8044	.8600	.8455	.8056	.8956	.8789	.4933	.4944	.4911	.7811	.7989	.8411	.6800	.8567	.8045	.7144	.7778	.7722



(a) $\tau = 4$; $\alpha = 0.1$. (b) $\tau = 4$; $\alpha = 0.5$. (c) $\tau = 4$; $\alpha \to 1$. (d) $\tau = 4$; $\alpha = 2$. (e) $\tau = 4$; $\alpha = 10$. (f) $\tau = 4$; $\alpha \to +\infty$.

(g) $\tau = 6$; $\alpha = 0.1$. (h) $\tau = 6$; $\alpha = 0.5$. (i) $\tau = 6$; $\alpha \to 1$. (j) $\tau = 6$; $\alpha = 2$. (k) $\tau = 6$; $\alpha = 10$. (l) $\tau = 6$; $\alpha \to +\infty$.

(m) $\tau = 8$; $\alpha = 0.1$. (n) $\tau = 8$; $\alpha = 0.5$. (o) $\tau = 8$; $\alpha \to 1$. (p) $\tau = 8$; $\alpha = 2$. (q) $\tau = 8$; $\alpha = 10$. (r) $\tau = 8$; $\alpha \to +\infty$.

(s) $\tau = 10$; $\alpha = 0.1$. (t) $\tau = 10$; $\alpha = 0.5$. (u) $\tau = 10$; $\alpha \to 1$. (v) $\tau = 10$; $\alpha = 2$. (w) $\tau = 10$; $\alpha = 10$. (x) $\tau = 10$; $\alpha \to +\infty$.

Fig. 5. Behavior of the BR*é*DOD model with the influence of time window scale υ and τ.

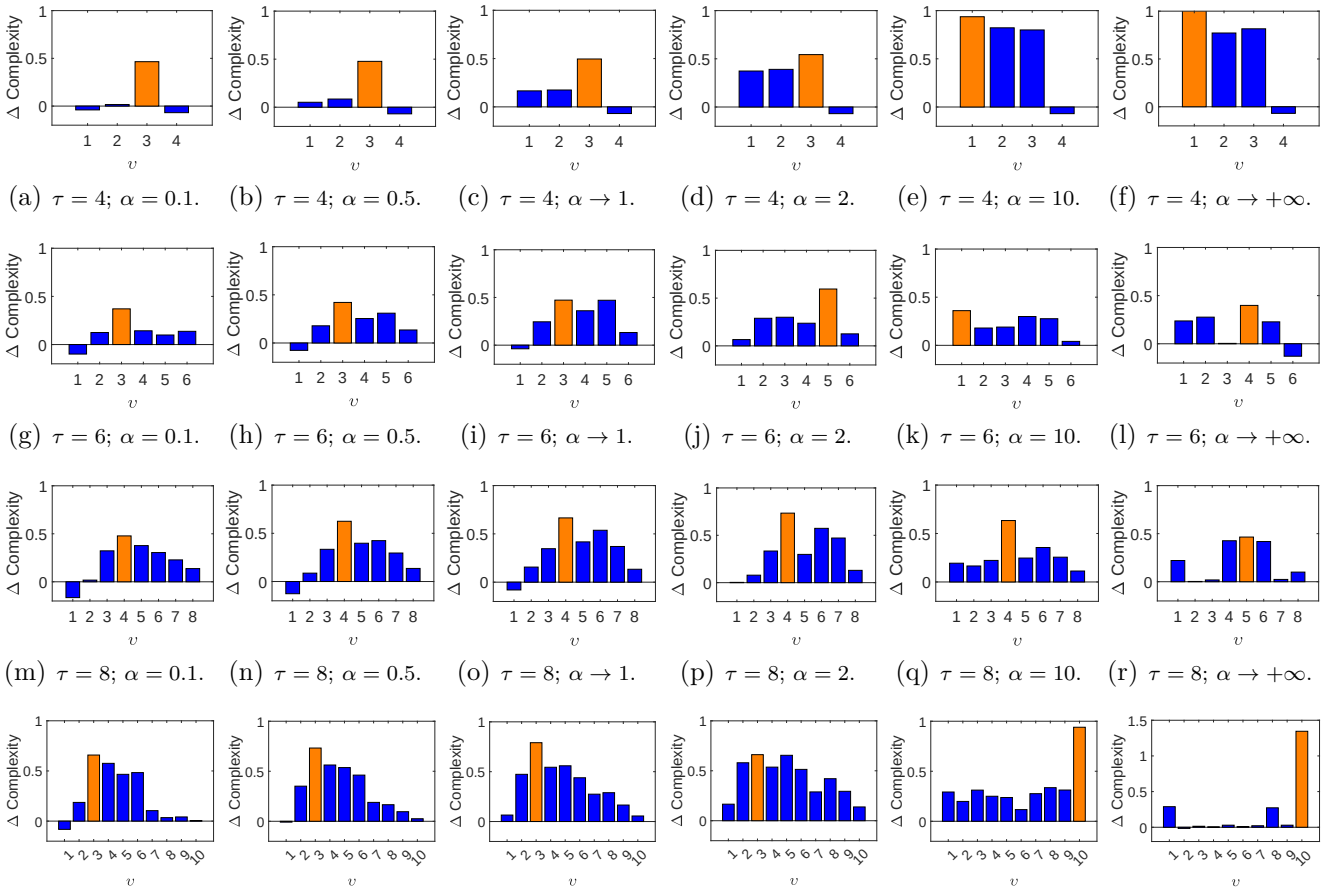
be a lack of performance in shorter time series, because insufficient quantity may cause missing information, which is intuitive that it's easier to extract features with larger data sets. As a result, the proposed BR*é*DOD method is able to distinguish the healthy subjects and pathology subjects at the same time with a longer length of time series. Considering the influence of α, there will be the highest classification accuracy with $\alpha = 0.5$, while the classification accuracy achieves the minimum with α tends to 1. Fig. 4 shows the relationship between the average value of divergence sequence and complexity value in varied lengths of time series. Each picture tends to be an inverted U shape, which means that the effectiveness of classification is further demonstrated. However, it is hard to distinguish the discrepancy among time series when α tends to be 1 and infinite.

D. The Influence of Parameter Setting

In this section, several parameters in BR*é*DOD model are taken into consideration to illustrate their influence. Specifically, the length of the time series is fixed as 300.

1) Parameters υ and τ : Before considering the influence of υ and τ, k is fixed as 55.

As is shown in Fig. 5, the complexity value of three representative subjects is getting more and more different with τ increasing from 4 to 10. When τ is at 4, there is a slight discrepancy in their complexity value. From Fig. 5, it can be found that a healthy subject reaches the highest complexity value most of the time. In addition, there will be an increase in complexity value before it declines with the change of υ. Specifically, the complexity value of all the subjects tends to zero when $\upsilon = \tau$, which means that the divergence of two BPAs tends to be 0 when they are



(a) $\tau = 4$; $\alpha = 0.1$. (b) $\tau = 4$; $\alpha = 0.5$. (c) $\tau = 4$; $\alpha \to 1$. (d) $\tau = 4$; $\alpha = 2$. (e) $\tau = 4$; $\alpha = 10$. (f) $\tau = 4$; $\alpha \to +\infty$.

(g) $\tau = 6$; $\alpha = 0.1$. (h) $\tau = 6$; $\alpha = 0.5$. (i) $\tau = 6$; $\alpha \to 1$. (j) $\tau = 6$; $\alpha = 2$. (k) $\tau = 6$; $\alpha = 10$. (l) $\tau = 6$; $\alpha \to +\infty$.

(m) $\tau = 8$; $\alpha = 0.1$. (n) $\tau = 8$; $\alpha = 0.5$. (o) $\tau = 8$; $\alpha \to 1$. (p) $\tau = 8$; $\alpha = 2$. (q) $\tau = 8$; $\alpha = 10$. (r) $\tau = 8$; $\alpha \to +\infty$.

(s) $\tau = 10$; $\alpha = 0.1$. (t) $\tau = 10$; $\alpha = 0.5$. (u) $\tau = 10$; $\alpha \to 1$. (v) $\tau = 10$; $\alpha = 2$. (w) $\tau = 10$; $\alpha = 10$. (x) $\tau = 10$; $\alpha \to +\infty$.

Fig. 6. Δ complexity value with the influence of time window scale υ and τ.

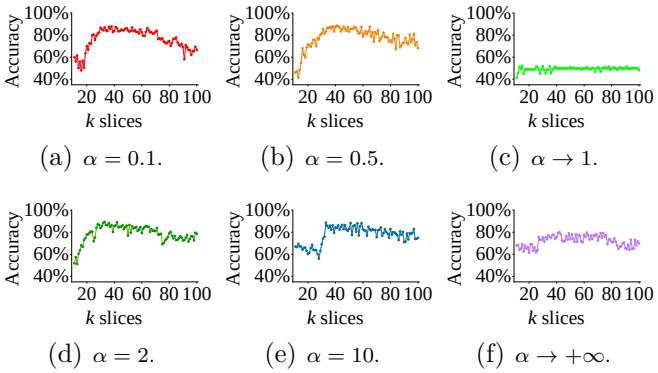
the same. It's worth noting that the computed complexity value remains low when the value of υ is small, because the elements in the corresponding BPA for υ are quite limited, which will lead to a divergence result tending to 0. When considering the changing of α, it can be found that as the value of α increases, the complexity values of the three representatives become more and more difficult to distinguish.

The setting of υ and τ is important, since them represent the truncation approaches of time window. Here, Δ complexity (the numerical value difference of complexity between healthy and pathology subjects) is utilized to analyze how to choose parameters υ and τ. The higher Δ complexity is, the easier to distinguish healthy and pathology subjects. As shown in Fig. 6, the orange bar of each sub-figure represents the optimal value of υ at a fixed τ, as it has the highest Δ complexity. It is worth noting that, for each column, Δ complexity gradually increases as τ increases with a fixed α, which means that τ is recommended to choose the larger value. However, with the increase of τ, the number of elements in the mass function built by time windows will also increase by a geometric multiple, and then a trade-off will be needed. As for υ, it usually be selected empirically, since the optimal υ will change with τ and α changing. In addition, different time series have different characteristics. The selection of parameters needs to be adaptive based on different time series data sets. Considering this data set, the best time window length setting should be $\tau = 10$ and $\upsilon = 3$, since $\alpha = 0.5$ based on Table III.

2) Parameter k: Here, υ and τ are fixed as 5 and 10 before considering the influence of k.

As for the 240 sets of cardiac inter-beat interval time series, Fig. 7 displays how classification accuracy varies with the changing of k. Note that accuracy is low when there are fewer slices, while more slices may cause the instability of the BR*é*DOD model, and accuracy will decline at the same time. It shows that $30 < k < 60$ can lead to better results in pattern classification problems. Specifically, the highest accuracy can be achieved as 89.56% with $k = 41$ and $\alpha = 0.5$. When considering α, it can be found that the classification accuracy goes up and then goes down with the increase of k. Particularly, the performance of BR*é*DOD model is the worst when $\alpha \to 1$ as classification accuracy hovers around 50%.

Hence, from the information above, the length of time series N, the scale of time window υ and τ, the segment slices k and coefficient in belief Rényi divergence α are able to affect the performance of the BR*é*DOD model in different degrees.



(a) $\alpha = 0.1$. (b) $\alpha = 0.5$. (c) $\alpha \to 1$.

(d) $\alpha = 2$. (e) $\alpha = 10$. (f) $\alpha \to +\infty$.

Fig. 7. The pattern classification accuracy of the BR*é*DOD with the influence of k.

TABLE IV
The patter classification accuracy based on different methods.

	$N = 140$	$N = 300$	$N = 500$
K-means	0.3822	0.3711	0.4078
Spectral clustering	0.4044	0.4100	0.4056
MSE	0.6738	0.7024	0.7248
EOE	0.8230	0.8789	0.8850
BEOE	0.8730	0.8800	0.8912
BR*é*DOD	0.8056	0.8956	0.8789

E. Comparison

Several classical and recent methods are applied to these cardiac inter-beat interval time series classification.

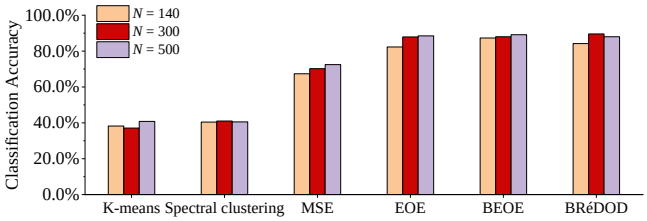


Fig. 8. The pattern classification accuracy in cardiac inter-beat interval time series.

The pattern classification accuracy based on different methods is shown in Table IV and Fig. 8. From Table IV, the accuracy reaches the peak at the value 0.8956 based on BR*é*DOD when $N = 300$, which overtakes other methods referred to.

It can be found that BR*é*DOD is far out-performance than K-means and Spectral clustering. No matter the length of the time series, these two classical machine learning methods are less accurate. According to MSE [65], EOE [66] and BEOE [51], BR*é*DOD achieves highest accuracy as 0.8956 with $N = 300$. In addition, BR*é*DOD has outstanding performance in both short-time series and long-time series. Though EOE and BEOE are able to reach high accuracy, they need more data to increase the accuracy.

Hence, the above information illustrates the superiority of the BR*é*DOD in complexity measure for pattern classification. Not only BR*é*DOD model overtakes the classical machine learning method, including K-means and Spectral clustering, but also surpasses several well-known methods, including MSE, EOE, and BEOE. In this case, the belief divergence of divergence reflects the complexity of the time series well.

V. Conclusion

As complexity analysis plays an important part in many research fields, measuring the complexity of time series is significant. A new algorithm, belief BR*é*DOD divergence of divergence was proposed to figure the complexity of time series. The main innovation of BR*é*DOD was that it measured the inner divergence of a time series twice, which can effectively reflect the feature of time series data. In addition, the boundaries of the time series value were taken into consideration to generate BPAs. Besides, the superiority of BR*é*DOD algorithm for complexity measure and pattern classification was demonstrated based on cardiac inter-beat interval time series with the highest classification accuracy. Nonetheless, the BR*é*DOD model exhibits a few limitations. Its extensive parameter set necessitates a careful selection process to identify the optimal configuration. Besides, there is a reduced likelihood of encountering boundary points when dealing with time series data distributed over a wide range. So, a more adaptive model is needed to select boundary points to construct appropriate mass functions. In summary, the BR*é*DOD algorithm has introduced an innovative approach for assessing the complexity of time series data. In future research, addressing the parameter-related issues and applying this method to a broader range of real-time series data should be considered.

VI. Acknowledgments

The authors greatly appreciate the reviewers' suggestions and editor's encouragement. This research is supported by the National Natural Science Foundation of China (No. 62003280), Chongqing Talents: Exceptional Young Talents Project (No. cstc2022ycjh-bgzxm0070), Natural Science Foundation of Chongqing, China (No. CSTB2022NSCQ-MSX0531), Chongqing Overseas Scholars Innovation Program (No. cx2022024), and Fundamental Research Funds for the Central Universities (No. 2023CDJXY-035).

References

[1] D. Yao, G. Cong, C. Zhang, X. Meng, R. Duan, J. Bi, A linear time approach to computing time series similarity based on deep metric learning, IEEE Transactions on Knowledge and Data Engineering 34 (10) (2022) 4554–4571.

[2] H. K. Lee, J. Lee, S. B. Kim, Boundary-focused generative adversarial networks for imbalanced and

multimodal time series, IEEE Transactions on Knowledge and Data Engineering 34 (9) (2022) 4102–4118.

[3] W. Wu, L. He, W. Lin, Y. Su, Y. Cui, C. Maple, S. Jarvis, Developing an unsupervised real-time anomaly detection scheme for time series with multi-seasonality, IEEE Transactions on Knowledge and Data Engineering 34 (9) (2022) 4147–4160.

[4] F. Liu, X. Zhou, J. Cao, Z. Wang, T. Wang, H. Wang, Y. Zhang, Anomaly detection in quasi-periodic time series based on automatic data segmentation and attentional lstm-cnn, IEEE Transactions on Knowledge and Data Engineering 34 (6) (2022) 2626–2640.

[5] J. W. Lai, K. H. Cheong, Superposition of COVID-19 waves, anticipating a sustained wave, and lessons for the future, BioEssays 42 (12) (2020) 2000178.

[6] Z. Cao, C.-H. Chuang, J.-K. King, C.-T. Lin, Multi-channel EEG recordings during a sustained-attention driving task, Scientific data 6 (1) (2019) 19.

[7] G. Li, B. Choi, J. Xu, S. S. Bhowmick, K.-P. Chun, G. L.-H. Wong, Efficient shapelet discovery for time series classification, IEEE Transactions on Knowledge and Data Engineering 34 (3) (2022) 1149–1163.

[8] J. Ircio, A. Lojo, U. Mori, S. Malinowski, J. A. Lozano, Minimum recall-based loss function for imbalanced time series classification, IEEE Transactions on Knowledge and Data Engineering (2023) 1–11doi:10.1109/TKDE.2023.3268994.

[9] Y. Liu, S. Garg, J. Nie, Y. Zhang, Z. Xiong, J. Kang, M. S. Hossain, Deep anomaly detection for time-series data in industrial iot: A communication-efficient on-device federated learning approach, IEEE Internet of Things Journal 8 (8) (2021) 6348–6358.

[10] W. Miao, Z. Xu, J. Geng, W. Jiang, Ecae: Edge-aware class activation enhancement for semisupervised remote sensing image semantic segmentation, IEEE Transactions on Geoscience and Remote Sensing 61 (2023) 1–14.

[11] Z.-g. Liu, Y.-m. Fu, Q. Pan, Z.-w. Zhang, Orientational distribution learning with hierarchical spatial attention for open set recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 45 (7) (2022) 8757 – 8772.

[12] J. Deng, Y. Deng, Maximum entropy of random permutation set, Soft Computing 26 (21) (2022) 11265–11275.

[13] L. Chen, Y. Deng, Entropy of random permutation set, Communications in Statistics-Theory and Methods (2023) 1–19.

[14] M. Lei, K. H. Cheong, Node influence ranking in complex networks: A local structure entropy approach, Chaos, Solitons & Fractals 160 (2022) 112136.

[15] Y. Cui, X. Deng, Plausibility entropy: A new total uncertainty measure in evidence theory based on plausibility function, IEEE Transactions on Systems, Man, and Cybernetics: Systems (2023) DOI: 10.1109/TSMC.2022.3233156.

[16] Z. Wang, C. Mu, S. Hu, C. Chu, X. Li, Modelling the dynamics of regret minimization in large agent populations: a master equation approach, in: Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI-22), 2022, pp. 534–540.

[17] C. Chu, Y. Li, J. Liu, S. Hu, X. Li, Z. Wang, A formal model for multiagent q-learning dynamics on regular graphs, in: Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI-22), 2022, pp. 194–200.

[18] Z. Wang, D. Hou, C. Gao, J. Huang, Q. Xuan, A rapid source localization method in the early stage of large-scale network propagation, in: Proceedings of the ACM Web Conference (WWW-22), 2022, p. 1372.

[19] H. Zhang, S. Zhong, Y. Deng, K. H. Cheong, LFIC: Identifying influential nodes in complex networks by local fuzzy information centrality, IEEE Transactions on Fuzzy Systems 30 (8) (2022) 3284–3296.

[20] J. Zhan, H. Jiang, Y. Yao, Three-way multiattribute decision-making based on outranking relations, IEEE Transactions on Fuzzy Systems 29 (10) (2020) 2844–2858.

[21] R. R. Yager, Multi-criteria decision making with interval criteria satisfactions using the golden rule representative value, IEEE Transactions on Fuzzy Systems 26 (2) (2017) 1023–1031.

[22] Y. Song, Q. Fu, Y.-F. Wang, X. Wang, Divergence-based cross entropy and uncertainty measures of Atanassov's intuitionistic fuzzy sets with their application in decision making, Applied Soft Computing 84 (2019) 105703.

[23] H. Liao, Z. Zhang, Z. Xu, A. Banaitis, A heterogeneous regret-theory-based method with Choquet integral to multiattribute reverse auction, IEEE Transactions on Engineering Management 69 (5) (2022) 2248–2259.

[24] D. Meng, H. Wang, S. Yang, Z. Lv, Z. Hu, Z. Wang, Fault analysis of wind power rolling bearing based on EMD feature extraction, CMES-Computer Modeling in Engineering & Sciences 130 (1) (2022) 543–558.

[25] Z. Lin, Z. Kang, L. Zhang, L. Tian, Multi-view attributed graph clustering, IEEE Transactions on Knowledge and Data Engineering 35 (2) (2023) 1872–1880.

[26] Z.-W. Zhang, Z.-G. Liu, A. Martin, K. Zhou, Bsc: Belief shift clustering, IEEE Transactions on Systems, Man, and Cybernetics: Systems 53 (3) (2022) 1748–1760.

[27] X. Gao, X. Su, H. Qian, X. Pan, Dependence assessment in human reliability analysis under uncertain and dynamic situations, Nuclear Engineering and Technology 54 (3) (2022) 948–958.

[28] J. Mi, Y.-F. Li, W. Peng, H.-Z. Huang, Reliability analysis of complex multi-state system with common cause failure based on evidential networks, Reliability Engineering & System Safety 174 (2018) 71–81.

[29] J. Jin, Y. Li, C. L. P. Chen, Pattern classification with corrupted labeling via robust broad learning

system, IEEE Transactions on Knowledge and Data Engineering 34 (10) (2022) 4959–4971.
[30] Z.-G. Liu, G.-H. Qiu, S.-Y. Wang, T.-C. Li, Q. Pan, A new belief-based bidirectional transfer classification method, IEEE Transactions on Cybernetics 52 (8) (2021) 8101–8113.
[31] J. Chu, H. Wang, J. Liu, Z. Gong, T. Li, Unsupervised feature learning architecture with multi-clustering integration RBM, IEEE Transactions on Knowledge and Data Engineering 34 (6) (2022) 3002–3015.
[32] C. Qiang, Z. Li, Y. Deng, Multifractal analysis of mass function, Soft Computing (2023) 1–14.
[33] X. Xu, D. Zhang, Y. Bai, L. Chang, J. Li, Evidence reasoning rule-based classifier with uncertainty quantification, Information Sciences 516 (2020) 192–204.
[34] H. Fujita, A. Gaeta, V. Loia, F. Orciuoli, Hypotheses analysis and assessment in counter-terrorism activities: a method based on OWA and fuzzy probabilistic rough sets, IEEE Transactions on Fuzzy Systems 28 (2020) 831–845.
[35] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, in: Classic works of the Dempster-Shafer theory of belief functions, Springer, 2008, pp. 57–72.
[36] G. Shafer, A mathematical theory of evidence, in: A mathematical theory of evidence, Princeton university press, 1976.
[37] Y. Deng, Uncertainty measure in evidence theory, SCIENCE CHINA Information Sciences 63 (11) (2020) 210201.
[38] R. R. Yager, Generalized Dempster–Shafer structures, IEEE Transactions on Fuzzy Systems 27 (3) (2019) 428–435.
[39] Y. Deng, Random permutation set, International Journal of Computers Communications & Control 17 (1) (2022) 4542.
[40] L. Chen, Y. Deng, K. H. Cheong, The distance of random permutation set, Information Sciences 628 (2023) 226–239.
[41] F. Xiao, Quantum X-entropy in generalized quantum evidence theory, Information Sciences 643 (2023) 119177.
[42] F. Xiao, Generalized quantum evidence theory, Applied Intelligence 53 (11) (2023) 14329–14344.
[43] L. Chang, L. Zhang, C. Fu, Y.-W. Chen, Transparent digital twin for output control using belief rule base, IEEE Transactions on cybernetics 52 (10) (2021) 10364–10378.
[44] C. Fu, B. Hou, M. Xue, L. Chang, W. Liu, Extended belief rule-based system with accurate rule weights and efficient rule activation for diagnosis of thyroid nodules, IEEE Transactions on Systems, Man, and Cybernetics: Systems 53 (1) (2022) 251–263.
[45] M. Zhou, Y.-Q. Zheng, Y.-W. Chen, B.-Y. Cheng, E. Herrera-Viedma, J. Wu, A large-scale group consensus reaching approach considering self-confidence with two-tuple linguistic trust/distrust relationship and its application in life cycle sustainability assessment, Information Fusion 94 (2023) 181–199.
[46] P. Liu, Y. Li, X. Zhang, W. Pedrycz, A multiattribute group decision-making method with probabilistic linguistic information based on an adaptive consensus reaching model and evidential reasoning, IEEE Transactions on Cybernetics 53 (3) (2022) 1905–1919.
[47] L. Fei, Y. Wang, An optimization model for rescuer assignments under an uncertain environment by using Dempster-Shafer theory, Knowledge-Based Systems (2022) 109680.
[48] S.-W. Tang, Z.-J. Zhou, C.-H. Hu, J.-B. Yang, Y. Cao, Perturbation analysis of evidential reasoning rule, IEEE Transactions on Systems, Man, and Cybernetics: Systems 51 (8) (2021) 4895–4910.
[49] L. Ni, Y.-w. Chen, O. de Brujin, Towards understanding socially influenced vaccination decision making: An integrated model of multiple criteria belief modelling and social network analysis, European Journal of Operational Research 293 (1) (2021) 276–289.
[50] H. Fujita, Y.-C. Ko, A heuristic representation learning based on evidential memberships: Case study of UCI-SPECTF, International Journal of Approximate Reasoning 120 (2020) DOI: 10.1016/j.ijar.2020.02.002.
[51] H. Cui, L. Zhou, Y. Li, B. Kang, Belief entropy-of-entropy and its application in the cardiac interbeat interval time series analysis, Chaos, Solitons & Fractals 155 (2022) 111736.
[52] C. Qiang, Y. Deng, K. H. Cheong, Information fractal dimension of mass function, Fractals 30 (06) (2022) 2250110.
[53] D. Han, J. Dezert, Y. Yang, Belief interval-based distance measures in the theory of belief functions, IEEE Transactions on Systems, Man, and Cybernetics: Systems 48 (6) (2018) 833–850.
[54] F. Xiao, GEJS: A generalized evidential divergence measure for multisource information fusion, IEEE Transactions on Systems, Man, and Cybernetics: Systems 53 (4) (2022) 2246–2258.
[55] J. Huang, X. Song, F. Xiao, Z. Cao, C.-T. Lin, Belief f-divergence for EEG complexity evaluation, Information Sciences 643 (2023) 119189.
[56] Y. Huang, F. Xiao, Z. Cao, C.-T. Lin, Higher order fractal belief Rényi divergence with its applications in pattern classification, IEEE Transactions on Pattern Analysis and Machine Intelligence 45 (12) (2023) 14709–14726.
[57] C. Zhu, F. Xiao, Z. Cao, A generalized Rényi divergence for multi-source information fusion with its application in EEG data analysis, Information Sciences 605 (2022) 225–243.
[58] C. Shui, R. Pu, G. Xu, J. Wen, F. Zhou, C. Gagné, C. X. Ling, B. Wang, Towards more general loss and setting in unsupervised domain adaptation, IEEE Transactions on Knowledge and Data Engineering (2023) 1–12doi:10.1109/TKDE.2023.3266785.
[59] Y. Song, Y. Gu, R. Zhang, G. Yu, Brepartition: Optimized high-dimensional knn search with bregman

distances, IEEE Transactions on Knowledge and Data Engineering 34 (3) (2020) 1053–1065.
[60] Y. Huang, F. Xiao, Higher order belief divergence with its application in pattern classification, Information Sciences 635 (2023) 1–24.
[61] L. Chen, Y. Deng, K. H. Cheong, Permutation Jensen–Shannon divergence for random permutation set, Engineering Applications of Artificial Intelligence 119 (2023) 105701.
[62] J. Zeng, F. Xiao, A fractal belief kl divergence for decision fusion, Engineering Applications of Artificial Intelligence 121 (2023) 106027.
[63] Y. Deng, Information volume of mass function, International Journal of Computers Communications & Control 15 (6) (2020) 3983.
[64] L. Chen, Y. Deng, K. H. Cheong, Probability transformation of mass function: A weighted network method based on the ordered visibility graph, Engineering Applications of Artificial Intelligence 105 (2021) 104438.
[65] M. Costa, A. L. Goldberger, C.-K. Peng, Multiscale entropy analysis of complex physiologic time series, Physical review letters 89 (6) (2002) 068102.
[66] C. F. Hsu, S.-Y. Wei, H.-P. Huang, L. Hsu, S. Chi, C.-K. Peng, Entropy of entropy: Measurement of dynamical complexity for biological systems, Entropy 19 (10) (2017).