ICJK
Chinese Journal of Information Fusion
http://dx.doi.org/10.62762/CJIF.2024.999646

REVIEW ARTICLE

# Complex Evidence Theory for Multisource Data Fusion

**Fuyuan Xiao**[1,*], **Junhao Wen**[1], **Witold Pedrycz**[2,3,4] **and Masayoshi Aritsugi**[5]

[1] School of Big Data and Software Engineering, Chongqing University, Chongqing 401331, China
[2] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2R3, Canada
[3] Systems Research Institute, Polish Academy of Sciences, 00-901 Warsaw, Poland
[4] National Information Processing Institute, 00-608 Warsaw, Poland
[5] Faculty of Advanced Science and Technology, Kumamoto University, Kumamoto 860-8555, Japan

## Abstract

**Data fusion is a prevalent technique for assembling imperfect raw data coming from multiple sources to capture reliable and accurate information. Dempster–Shafer evidence theory is one of useful methodologies in the fusion of uncertain multisource information. The existing literature lacks a thorough and comprehensive review of the recent advances of Dempster–Shafer evidence theory for data fusion. Therefore, the state of the art has to be surveyed to gain insight into how Dempster–Shafer evidence theory is beneficial for data fusion and how it evolved over time. In this paper, we first provide a comprehensive review of data fusion methods based on Dempster–Shafer evidence theory and its extensions, collectively referred to as classical evidence theory, from three aspects of uncertainty modeling, fusion, and decision making. Next, we study and explore complex evidence theory for data fusion in both closed world and open world contexts that benefits from the frame of complex plane modelling. We then present classical and complex evidence theory framework-based multisource data fusion algorithms, which are applied to pattern classification to compare and demonstrate their applicability. The research results indicate that the complex evidence theory framework can enhance the capabilities of uncertainty modeling and reasoning by generating constructive interference through the fusion of appropriate complex basic belief assignment functions modeled by complex numbers. Through analysis and comparison, we finally propose several challenges and identify open future research directions in evidence theory-based data fusion.**

**Keywords**: multisource data fusion, Dempster–Shafer evidence theory, complex evidence theory, quantum evidence theory, uncertainty modeling, conflict management, belief function, decision making, pattern classification.

## 1 Introduction

With the development of the information age, large amounts of data are being generated, gathered and disposed. Interfering based on solely on single source

**Academic Editor:**
Fengbao Yang

**Submitted:** 30 July 2024
**Accepted:** 28 September 2024
**Published:** 30 September 2024

**Vol.** 1, **No.** 2, 2024.
10.62762/CJIF.2024.999646

***Corresponding author:**
✉ Fuyuan Xiao
xiaofuyuan@cqu.edu.cn

**Citation**
Xiao, F., Wen, J., Pedrycz, W., & Aritsugi, M. (2024). Complex Evidence Theory for Multisource Data Fusion. *Chinese Journal of Information Fusion*, 1(2), 134–159.

is no longer sufficient. Data fusion, also known as multisource data fusion, is a necessary technique to assemble various kinds of data from multiple sources to capture reliable and accurate information [1–4]. Nevertheless, uncertain, imprecise, imbalance, and incomplete or even false data are inevitable on account of the impacts of the environment and the complexity of the goals [5–7]. Such kinds of problems increase difficult to multisource data fusion. To improve the performance of the fusion system, various data fusion methods have been presented [8, 9], and applied in a wide variety of areas [10–14], including artificial intelligence, target tracking and recognition, smart engineering management, IoT systems, financial systems, medical diagnosis, and so on [15–17].

Existing works provide an overview of the effort related to data fusion from different perspectives. Some papers [18, 19] investigate data fusion in smart Internet of Things (IoT). Some papers [20] investigate multisensor data fusion techniques. A comprehensive survey on data fusion in remote sensing is presented in [21]. Some papers [22] investigate data fusion in machine learning. A survey in [23] presents mobile agent itinerary planning for information fusion in wireless sensor networks. A survey on information fusion for edge intelligence is presented in [24]. A recently published survey conducts a comprehensive statistical analysis of the current theoretical and application achievements of multisource information fusion [25]. Dempster–Shafer evidence theory (DSET) [26, 27] is one of useful methodologies in the fusion of uncertain multi-source information [28]. To the best of our knowledge, although researchers have conducted reviews and surveys of data fusion from different perspectives, no attempt has been made to provide a comprehensive overview on the DSET for data fusion by carefully analyzing the abovementioned data fusion literature.

Therefore, in this paper, a comprehensive and systematic survey on data fusion in DSET is conducted. We first review the basic concepts and knowledge of classical DSET, then study the axioms of Dempster's rule of combination and the characteristics of DSET that are desirable for data fusion. Whereafter, we review classical DSET and its extensions, collectively referred to as classical evidence theory, for data fusion from three aspects of uncertainty modeling, fusion, and decision making. Next, we explore complex evidence theory for data fusion in both closed world and open world contexts that benefits from the frame of complex plane modelling. After that, we present classical and complex evidence theory framework-based multisource data fusion algorithms. These algorithms are applied to pattern classification to demonstrate their applicability through comparison with other related well-known methods. Finally, we discuss challenges and open directions for future research. The objectives of this work are to: 1) offer a general and synthetic review of classical and complex evidence theories; 2) provide a profound analysis and discussion of existing work in classical and complex evidence theory framework-based data fusion domains; and 3) identify remaining challenges and open directions for future research in this field.

The paper is organized as follows. Section 2 reviews classical DSET and its typical generalized theories. In Section 3, we analyze classical evidence theory for data fusion. In Section 4, we study and explore complex evidence theory for data fusion in both closed and open world contexts. In Section 5, classical and complex evidence theory framework-based multisource data fusion algorithms are presented. In Section 6, several challenges and open future research directions are discussed. Finally, Section 7 concludes this work.

## 2 Review of Classical DSET and Its Generalized Theories

In this section, we first review the basic concepts, knowledge and limitations of classical Dempster–Shafer evidence theory (DSET). Next, we review two typical generalizations of DSET, namely, DSmT: Dezert-Smarandache Theory, and GET: generalized evidence theory. In addition, we analyze the characteristics of the classical DSET and its generalized theories, then assess their differences.

### 2.1 DSET: Dempster–Shafer Evidence Theory [26, 27]

#### 2.1.1 *Basic Concepts and Knowledge of Classical DSET*

The classical DSET, also called the theory of belief functions, was first presented by Dempster [26] and later developed by Shafer [27]. As a generalization of Bayesian probability theory, DSET is more flexible and effective to express and process uncertainty [29–31], which is applied in many fields, such as evidential reasoning [32–34], belief rule-base expert system [35–39], fault diagnosis [40, 41], software risk evaluation [42], and other aspects [43, 44]. The main concepts of DSET are introduced below [26, 27].

**Definition 1** (*Frame of discernment*). Let $\phi_i$ be an arbitrary nonempty event. A frame of discernment

(FOD), denoted as $\Phi$, is defined as:

$$\Phi = \{\phi_1, \ldots, \phi_i, \ldots, \phi_g, \ldots, \phi_n\}, \tag{1}$$

where $\forall i, g = \{1, \ldots, n\}$, $\phi_i$ and $\phi_g$ are two arbitrary nonempty events and $\phi_i \cap \phi_g = \emptyset$.

**Definition 2** (*Power set*). Let $2^\Phi$ be the power set of $\Phi$, denoted as:

$$2^\Phi = \{\emptyset, \{\phi_1\}, \{\phi_2\}, \ldots, \{\phi_n\}, \{\phi_1, \phi_2\}, \ldots, \{\phi_1, \phi_2, \ldots, \phi_i\}, \ldots, \Phi\}, \tag{2}$$

where $\emptyset$ is an empty set.

**Definition 3** (*Hypothesis or proposition*). $\forall \psi_j \in 2^\Phi$, $\psi_j$ is defined as a hypothesis or proposition.

**Definition 4** (*Mass function in DSET*). In FOD $\Phi$, a mass function $m$ is defined as a mapping:

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

satisfying the following:

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{\psi_j \subseteq \Phi} m(\psi_j) = 1, \tag{4}$$

where $m$ is also called a basic belief assignment (BBA).

**Definition 5** (*Focal element in DSET*). Let $m$ be a BBA defined in Definition 4. $\forall \psi_j \subseteq \Phi$, if $m(\psi_j) > 0$, $\psi_j$ is defined as a focal element.

**Definition 6** (*Belief function*). A belief function $Bel$, mapping from $2^\Phi$ to $[0, 1]$, is defined by

$$Bel(\psi_j) = \sum_{\psi_k \subseteq \psi_j | \psi_j \in 2^\Phi} m(\psi_k). \tag{5}$$

**Definition 7** (*Plausibility function*). A plausibility function $Pl$, mapping from $2^\Phi$ to $[0, 1]$, is defined by

$$Pl(\psi_j) = \sum_{\psi_k \cap \psi_j \neq \emptyset | \psi_j, \psi_k \in 2^\Phi} m(\psi_k) = 1 - Bel(\bar{\psi_j}), \tag{6}$$

where $\bar{\psi_j} = \Phi - \psi_j$.

Clearly, $\forall \psi_j \in 2^\Phi$, $Pl(\psi_j) \geq Bel(\psi_j)$, where $Bel(\psi_j)$ and $Pl(\psi_j)$ are the lower and upper limit functions to support $\psi_j$, respectively.

**Definition 8** (*Dempster's rule of combination*). Let $m_1$ and $m_2$ be two independent BBAs in FOD $\Phi$ with propositions $\psi_k, \psi_h \subseteq \Phi$, respectively. Dempster's rule of combination (DRC), represented in the form $m_1 \oplus m_2$, is defined by

$$m_1 \oplus m_2(\psi_j) = \begin{cases} \frac{1}{1-K} \sum\limits_{\psi_k \cap \psi_h = \psi_j} m_1(\psi_k) m_2(\psi_h), & \psi_j \neq \emptyset, \\ 0, & \psi_j = \emptyset, \end{cases} \tag{7}$$

with

$$K = \sum_{\psi_k \cap \psi_h = \emptyset} m_1(\psi_k) m_2(\psi_h), \tag{8}$$

where $K$ is the conflict coefficient between $m_1$ and $m_2$ [45].

Note that Eq. (7) is feasible under the condition $K < 1$.

### 2.1.2 Axioms and Characteristics

Notably, DRC is conducive to data fusion since it has a set of attractive axioms that are illustrated below [46]:

**Axiom A1: Compositionality**.

$m_1 \oplus m_2(\psi_j)$ is a function of only $\psi_j$, $m_1$, and $m_2$.

**Axiom A2: Commutativity**.

$$m_1 \oplus m_2 = m_2 \oplus m_1.$$

**Axiom A3: Associativity**.

$$(m_1 \oplus m_2) \oplus m_3 = m_1 \oplus (m_2 \oplus m_3).$$

**Axiom A4: Conditioning**.

If $m_2(\psi_h) = 1$, then

$$m_1 \oplus m_2(\psi_j) = \begin{cases} \sum\limits_{\psi_k \subseteq \bar{\psi_h}} m_1(\psi_j \cup \psi_k), & \text{for all } \psi_j \subseteq \psi_h, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

**Axiom A5: Internal Symmetry**.

Let $\{\theta_1, \theta_2, \ldots, \theta_j, \ldots, \theta_{2^\Phi}\}$ be an arbitrary permutation of hypothesis $\{\psi_1, \psi_2, \ldots, \psi_j, \ldots, \psi_{2^\Phi}\}$. Consider BBAs $m_t$ and $\dot{m}_t$ $(t = 1, 2)$:

$$\begin{aligned} m_t =& [m_t(\psi_1), m_t(\psi_2), m_t(\psi_1, \psi_2), m_t(\psi_3), \ldots, \\ & m_t(\psi_1, \psi_2, \ldots, \psi_{2^\Phi})]; \\ \dot{m}_t =& [\dot{m}_t(\theta_1), \dot{m}_t(\theta_2), \dot{m}_t(\theta_1, \theta_2), \dot{m}_t(\theta_3), \ldots, \\ & \dot{m}_t(\theta_1, \theta_2, \ldots, \theta_{2^\Phi})]. \end{aligned}$$

Then, $m_1 \oplus m_2 = \dot{m}_1 \oplus \dot{m}_2$.

**Axiom A6: Autofunctionality**.

For $\forall \theta_j \in \Phi$ and $\theta_j \neq \Phi$, $m_1 \oplus m_2(\theta_j)$ does not rely on $m_1(\theta_h)$ for all $\theta_h \subseteq \bar{\theta}_j$.

By analyzing the above definitions and axioms, the characteristics of the classical DSET can be summarized as follows:

**C**1: BBA $m$ in DSET has the ability to model partial or complete ignorance.

**C**2: Compared to the Bayesian decision model, the belief function does not need experts to offer prior probabilities.

**C**3: DRC satisfies the associative law and commutative law and provides flexible and facilitating reasoning to handle uncertainty in the fusion of multisource data.

**C**4: The belief interval $[Bel(\psi_j), Pl(\psi_j)]$ in DSET provides upper and lower probabilities by means of a belief function and plausibility function.

Consequently, as a form of nonprior generalization of Bayesian inference, DSET offers a general framework to support decision making through reasoning under uncertainty.

### *2.1.3 Restraints in DSET*

DSET provides a mass function to express uncertainty quantitatively, and DRC for reasoning to ensure fusion. Specifically, the aim of Dempster's rule of combination is to aggregate and combine information modeled in mass functions or BBAs into a distinct function. Although DSET has many advantages that are desirable for data fusion, it suffers the following restraints that limit its application:

**R**1: **Restraint on the frame of discernment.** In terms of the FOD of DSET, the elements are assumed to be exhaustive and exclusive. However, for a variety of fusion problems, the internal essence of hypotheses may be vague and imprecise, so the elements in the FOD may overlap. On the other hand, for dynamic fusion problems, the number of elements in the FOD changes over time, accompanied by the amendment of available knowledge. Hence, relaxing these assumptions by taking into account nonexclusivity among elements, as well as the evolution of element quantity in the FOD, is a key issue in data fusion to describe the problems in actual applications more realistically.

**R**2: **Constraint on independent evidence fusion.** In DSET, when using DRC, the evidence to be fused assumed to be independent; however, dependency among evidence is ubiquitous in practical applications. How to overcome this limitation to make the combination rule able to handle uncertainty is another key issue in data fusion.

**R**3: **Counterintuitive result when fusing conflicting evidence.** Because the BBAs are generated based on uncertain input variables modeled in a variety of forms from different sources, conflicts among multiple sources may exist due to the impacts of subjective and objective uncertainties. Nevertheless, counterintuitive results occur when fusing highly conflicting evidence through DRC. How to manage conflicting evidence to improve the fusion quality with a high decision level is another key issue in data fusion.

Thus far, several studies have contributed solutions to the abovementioned issues; however, no consensus has been reached about what the best approach is for the restraint relaxation of the FOD, dependent evidence fusion, and conflict management. Next, we survey two typical generalization frameworks of DSET that were presented to address these issues, namely, DSmT: Dezert-Smarandache Theory [47], and GET: generalized evidence theory [48]. The main concepts of these frameworks will be introduced below. In addition, a corresponding analysis of their characteristics, compared with those of classical DSET, will be presented.

## 2.2 DSmT: Dezert-Smarandache theory [47]

**Definition 9** (*Frame*). Let $\Phi = \{\phi_1, \ldots, \phi_j, \ldots, \phi_n\}$ be a finite set of $n$ exhaustive elements called a frame. If $\Phi$ is congenitally not closed, namely, it is an open world or frame, $\phi_{n+1}$ can always be included in $\Phi$ as a new closed world or frame: $\{\phi_1, \ldots, \phi_j, \ldots, \phi_n, \phi_{n+1}\}$.

In contrast to classical DSET, there are constraints on $\phi_j$ and $\Phi$ in DSmT other than exhaustivity. Therefore, the frame of DSmT releases the restraints where the events must be exclusive and constant in the FOD of DSET. As a result, DSmT offers a more realistic and flexible structure of the frame model.

**Definition 10** (*Hyper-power set*). The hyper-power set of $\Phi$ is denoted as

$$D^{\Phi} \triangleq (\Phi, \cup, \cap), \tag{10}$$

where the hyper-power set $D^{\Phi}$ is defined as the set of all subsets from $\Phi$ with union and intersection operators, i.e., $\cup$ and $\cap$, such that:

- $\emptyset, \phi_1, \ldots, \phi_n \in D^{\Phi}$;
- If $\psi_k, \psi_h \in D^{\Phi}$, then $\psi_k \cap \psi_h \in D^{\Phi}$ and $\psi_k \cup \psi_h \in D^{\Phi}$;
- No other elements belong to $D^{\Phi}$ except those obtained through rules (1) and (2).

Remarkably, given a finite frame $\Phi$, $D^{\Phi}$ has the following characteristics:

- When $|D^{\Phi}| \geq |2^{\Phi}|$, $D^{\Phi}$ is called the hyper-power set of $\Phi$.
- When all the elements of $\Phi$ are known (or are assumed) to be truly exclusive, $D^{\Phi}$ becomes the classical power set $2^{\Phi}$.

**Definition 11** (*Super-power set*). The super-power set of $\Phi$ is denoted as

$$S^{\Phi} \triangleq [\Phi, \cup, \cap, c(\cdot)], \tag{11}$$

where the super-power set $S^{\Phi}$ is defined as the set of all subsets from $\Phi$ with union, intersection and complementation operators, i.e., $\cup$, $\cap$, and $c(\cdot)$, such that:

- $\emptyset, \phi_1, \ldots, \phi_n \in S^{\Phi}$;
- If $\psi_k, \psi_h \in S^{\Phi}$, then $\psi_k \cap \psi_h \in S^{\Phi}$ and $\psi_k \cup \psi_h \in S^{\Phi}$;
- If $\psi_k \in S^{\Phi}$, then $c(\cdot) \in S^{\Phi}$;
- No other elements belong to $S^{\Phi}$ except those obtained through rules (1), (2) and (3).

**Definition 12** (*Mass function in DSmT*). In frame $\Phi$, a mass function $m$ in DSmT is defined as a mapping:

$$m : G^{\Phi} \to [0, 1], \tag{12}$$

satisfying the following properties:

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{\psi_j \in G^{\Phi}} m(\psi_j) = 1, \tag{13}$$

where $G^{\Phi}$ is a fusion space, which may be $2^{\Phi}$, $D^{\Phi}$ or $S^{\Phi}$, in accordance with the model selection for $\Phi$.

Note that $m$ is also called a generalized basic belief assignment (GBBA).

Comparison of Definition 4 with Definition 12 indicates that compared to the classical BBA in DSET, $m$ in DSmT has the following interpretations and properties:

- In DSmT, $m$ can be modeled in the power set $2^{\Phi}$, hyper-power set $D^{\Phi}$, and super-power set $S^{\Phi}$, while in DSET, $m$ can be modeled only in the power set $2^{\Phi}$.
- When $G^{\Phi} = 2^{\Phi}$, where all the elements in $2^{\Phi}$ are known and exclusive, GBBA $m$ in DSmT reduces to the classical BBA in DSET.

**Definition 13** (*Generalized belief function in DSmT*). A generalized belief function $GBel$ in DSmT, mapping from $G^{\Phi}$ to $[0, 1]$, is defined by

$$GBel(\psi_j) = \sum_{\psi_k \subseteq \psi_j | \psi_j \in G^{\Phi}} m(\psi_k). \tag{14}$$

**Definition 14** (*Generalized plausibility function in DSmT*). A generalized plausibility function $GPl$ in DSmT, mapping from $G^{\Phi}$ to $[0, 1]$, is defined by

$$GPl(\psi_j) = \sum_{\psi_k \cap \psi_j \neq \emptyset | \psi_j, \psi_k \in G^{\Phi}} m(\psi_k). \tag{15}$$

Clearly, $\forall \psi_j \in G^{\Phi}$, $GPl(\psi_j) \geq GBel(\psi_j)$, where $GBel(\psi_j)$ and $GPl(\psi_j)$ are the lower and upper limit functions of $\psi_j$, respectively.

In DSmT, several proportional conflict redistribution (PCR) rules are presented for data fusion. The idea behind the PCR fusion rules is to proportionally shift total or partial conflict masses to nonempty sets involved in the conflicts in regard to the masses allocated by sources [49]. In particular, as discussed in [49], PCR5, which takes the conjunctive normal form of partial conflict into consideration, is regarded as the most mathematical and effective PCR fusion rule. Thus, we introduce the basic concept of PCR5 below.

**Definition 15** (*PCR5 rule*) [49]. Let $m_1$ and $m_2$ be two GBBAs in frame $\Phi$. The PCR5 rule in DSmT is defined by:

$$m_1 \oplus m_2(\psi_j) = \begin{cases} m_{\cap}(\psi_j) + \\ \sum\limits_{\substack{\psi_j \cap \psi_k = \emptyset \\ \psi_k \neq \emptyset}} \left[ \frac{m_1(\psi_j)^2 m_2(\psi_k)}{m_1(\psi_j) + m_2(\psi_k)} + \frac{m_2(\psi_j)^2 m_1(\psi_k)}{m_2(\psi_j) + m_1(\psi_k)} \right], & \psi_j \neq \emptyset, \\ 0, \quad \psi_j = \emptyset, \end{cases} \tag{16}$$

with

$$m_{\cap}(\psi_j) = \sum_{\psi_h \cap \psi_k = \psi_j} m_1(\psi_h) m_2(\psi_k), \tag{17}$$

where $\psi_j, \psi_h, \psi_k \in G^{\Phi}$.

Notably, in contrast to the classical DRC, the PCR5 rule is quasi-associative and maintains the neutral influence of vacuous belief assignment. This is because the conjunctive normal form of each partial conflict does not cover $\Phi$, as $\Phi$ is a neutral element for conflict. Therefore, no mass is assigned to $\Phi$ after redistributing the conflict mass.

In summary, the generalized DSmT inherits the merits of classical DSET and has its own attractive characteristics, as follows [47]:

**C**1: The frame $\Phi$ in DSmT relaxes the assumptions on the FOD in DSET, except for the exhaustivity of $\Phi$. Specifically, the frame $\Phi$ in DSmT also contains the elements with conjunctions and/or disjunctions and negations/complements of pure hypotheses.

**C**2: GBBA $m$ in DSmT is capable of expressing partial or complete ignorance with not only the power set, but also the hyper- and super-power sets.

**C**3: The generalized belief function in DSmT also does not need experts to provide prior probabilities, in contrast to the Bayesian decision model.

**C**4: DSmT affords better fusion rule of PCR5, which can effectively cope with conflicting evidence compared to the classical DRC in DSET.

**C**5: The generalized belief interval $[GBel(\psi_j), GPl(\psi_j)]$ in DSmT also provides upper and lower probabilities. In addition, DSmT presents a method to work with imprecise quantitative or qualitative information without the limitation of interval-valued belief structures.

### 2.3 GET: Generalized Evidence Theory [48]

**Definition 16** (*Mass function in GET*). In the FOD $\Phi$, a mass function $m$ in GET is defined as a mapping:

$$m : 2^{\Phi} \to [0, 1], \tag{18}$$

satisfying:

$$\sum_{\psi_j \in 2^{\Phi}} m(\psi_j) = 1, \tag{19}$$

in which $m$ is also called a GBBA.

**Definition 17** (*Focal element in GET*). Let $m$ be a GBBA defined in Definition 16. $\forall \psi_j \in 2^{\Phi}$, if $m(\psi_j) > 0$, $\psi_j$ is defined as a focal element.

**Definition 18** (*Generalized belief function in GET*). A generalized belief function $GBel$ in GET, mapping from $2^{\Phi}$ to $[0, 1]$, is defined by

$$\begin{aligned} GBel(\psi_j) &= \sum_{\psi_k \subseteq \psi_j | \psi_j \in 2^{\Phi}} m(\psi_k), \\ GBel(\emptyset) &= m(\emptyset). \end{aligned} \tag{20}$$

**Definition 19** (*Generalized plausibility function in GET*). A generalized plausibility function $GPl$ in GET, mapping from $2^{\Phi}$ to $[0, 1]$, is defined by

$$\begin{aligned} GPl(\psi_j) &= \sum_{\psi_k \cap \psi_j \neq \emptyset | \psi_j, \psi_k \in 2^{\Phi}} m(\psi_k), \\ GPl(\emptyset) &= m(\emptyset). \end{aligned} \tag{21}$$

**Definition 20** (*Generalized combination rule in GET*). Let $m_1$ and $m_2$ be two independent GBBAs with propositions $\psi_k, \psi_h \in 2^{\Phi}$, respectively, defined in Definition 16. The generalized combination rule (GCR), represented in the form $m_1 \oplus m_2$, is defined as follows:

$$m_1 \oplus m_2(\psi_j) = \begin{cases} \frac{[1-m_1 \oplus m_2(\emptyset)] \sum\limits_{\psi_k \cap \psi_h = \psi_j} m_1(\psi_k) m_2(\psi_h)}{1-K}, \\ \psi_j \neq \emptyset, \\ m_1(\emptyset) m_2(\emptyset), \quad \psi_j = \emptyset, \end{cases} \tag{22}$$

with

$$K = \sum_{\psi_k \cap \psi_h = \emptyset} m_1(\psi_k) m_2(\psi_h), \tag{23}$$

where $m(\emptyset) = 1$ if and only if $K = 1$.

The generalized GET inherits the merits of classical DSET and has the following attractive characteristics [48]:

**C**1: The structure of GBBA in GET has the ability not only to model partial or complete ignorance but also to express the uncertainty caused by the incompleteness of FOD, so it can handle the uncertainty problem in an open world.

**C**2: The generalized belief function in GET also does not need experts to provide prior probabilities, in contrast to the Bayesian decision model.

**C**3: The generalized combination rule in GET not only satisfies the associative law and commutative law but also has the capability to reason with multisource data in the face of uncertainty, even under the condition that $m(\emptyset) > 0$.

**C**4: The generalized belief interval $[GBel(\psi_j), GPl(\psi_j)]$ in GET provides upper and lower probabilities and can be applied to the open world not just the closed world.

**C**5: When $m(\emptyset) = 0$, GET reduces to the classical DSET.

Moreover, DSET has been extended in terms of other aspects [50–52].

## 3 Classical Evidence Theory for Data Fusion

Our methodology expresses a deep understanding of the surveyed papers with regards to evidence theory for data fusion. The process involves three steps, including uncertainty modeling, fusion, and decision making. Figure 1 shows a process of data fusion in the context of evidence theory.

### 3.1 Uncertainty Modeling

Uncertainty is ubiquitous in the real world and is found in almost all areas of scientific research [53–55]. Uncertainty in data fusion can generally be classified into the following two categories:

**U**1: **Aleatory uncertainty.** This uncertainty originates from natural variability of the physical world and reflects its inherent randomness. Aleatory uncertainty exists naturally without connection to human knowledge. This kind of uncertainty cannot be removed or decreased by gathering more information.

**U**2: **Epistemic uncertainty.** This uncertainty occurs because humans lack knowledge of the physical

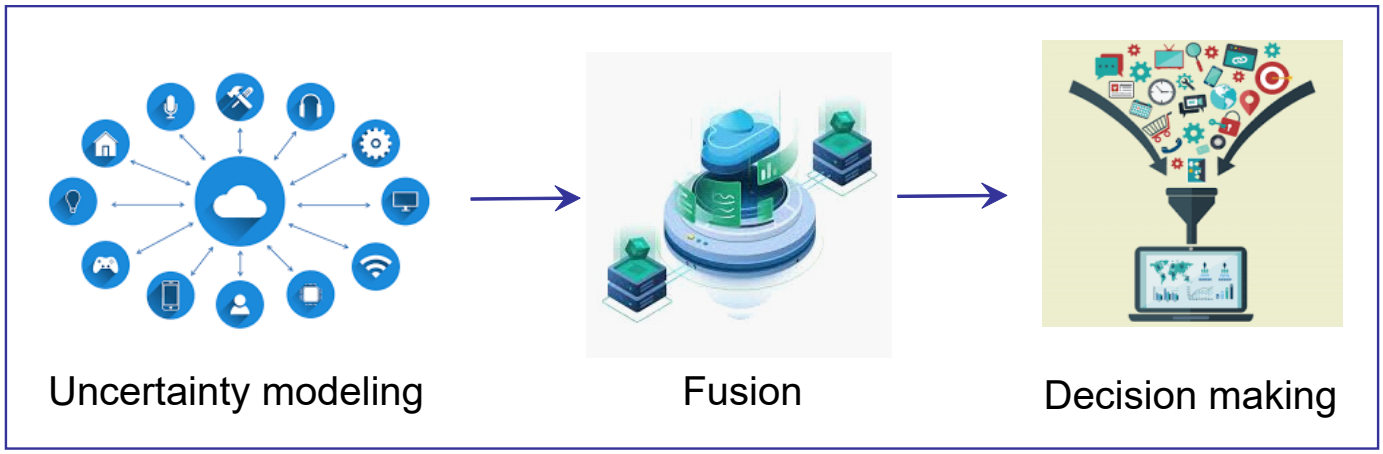


**Figure 1.** A process of data fusion in the context of evidence theory.

world and the ability to measure and model the physical world. In contrast to aleatory uncertainty, this kind of uncertainty can be reduced and even eliminated with the aid of more information and appropriate methods.

For example, phrases "I am 80% sure that ..." and "I think there is a 85% change that ..." express epistemic and aleatory uncertainty, respectively. In various areas of science and engineering, these uncertainty makes tasks more complicated and influences decision making in many adverse ways. Modeling and handling of these kinds of uncertainty plays an important role. The mass function/BBA, belief function and plausibility function in DSET offer more flexible and realistic expression and formalization of available knowledge with regards to possible values of uncertain input variables. Especially, for mass function, instead of filling in the missing value by a certain estimation, it can provide a straightforward way to quantify such a kind of ignorance state. In this way, any external operation in terms of the missing value is not required.

For evidence theory-based data fusion, BBA, as a basic unit of evidence theory framework processing, is the key issue that need to be addressed first. Therefore, how to generate appropriate BBAs from multisource information has been intensively studied in recent years. For example, the BBAs can be generated in accordance with multiple attributes of dataset. According to whether prior sample knowledge is used in the process of BBA generation, it can be mainly divided into the following three types:

(1) **Unsupervised.** For instance, paper [56] investigates unsupervised segmentation of hidden Markov fields corrupted by correlated non-Gaussian noise; paper [57] presents a belief shift clustering method for dealing with object data; paper [58] investigates neural network-based evidential clustering for BBA generation.

(2) **Semi-supervised.** For example, paper [59] studies a semi-supervised evidential label propagation algorithm for graph data clustering; paper [60] researches disagreement based semi-supervised learning approaches with belief functions; paper [61] investigates a fast semi-supervised evidential clustering.

(3) **Supervised.** For instance, paper [62] investigates evidential calibration of binary SVM classifiers; paper [63] studies evidential classifiers, including logistic regression and its nonlinear generalizations of multilayer feedforward neural networks; paper [64] investigates an evidential classifier based on Dempster-Shafer theory and deep learning.

### 3.2 Fusion

After obtaining BBAs from multisource information, other key issue is about how to fuse these BBAs to better support decision-making.

In addition to the typical DRC and promotion generalized combination rules discussed in Section 2, various research has been conducted from other perspectives to improve the fusion performance. To summarise, there are three dominating classifications: evidential combination rule-based data fusion; evidence pretreatment-based data fusion; and hybrid evidential conflict models for data fusion. We will

explain evidence theory-based data fusion methods from these three aspects in the following sections.

### 3.2.1 *Evidential combination rule-based data fusion*

In this section, we survey several existing evidential combination rules that have been widely applied in data fusion. Additionally, we compare existing evidential combination rules and summarized their properties.

(1) Evidential combination rules

Let $m_1$ and $m_2$ be two independent BBAs with hypotheses $\psi_k$ and $\psi_h$ in FOD $\Phi$, respectively.

**Definition 21** (*Smets's combination rules*) [46]. Smets's conjunctive combination rule, represented in the form $m_1 \textcircled{\cap} m_2$, is defined by

$$m_1 \textcircled{\cap} m_2(\psi_j) = \begin{cases} \sum\limits_{\psi_k \cap \psi_h = \psi_j} m_1(\psi_k) m_2(\psi_h), & \psi_j \subseteq \Phi, \\ m(\emptyset), & \psi_j = \emptyset, \end{cases} \tag{24}$$

and Smets's disjunctive combination rule, represented in the form $m_1 \textcircled{\cup} m_2$, is defined by

$$m_1 \textcircled{\cup} m_2(\psi_j) = \begin{cases} \sum\limits_{\psi_k \cup \psi_h = \psi_j} m_1(\psi_k) m_2(\psi_h), & \psi_j \subseteq \Phi, \\ m(\emptyset), & \psi_j = \emptyset. \end{cases} \tag{25}$$

In Eq. (24) and Eq. (25), $\emptyset$ has the following interpretations:

- $\emptyset$ is an empty set belonging to $2^\Phi$ in a closed world with $m(\emptyset) = 0$.
- $\emptyset$ is one or several hypotheses in an open world that does not belong to $2^\Phi$.

**Definition 22** (*Yager's combination rule*) [65]. Yager's combination rule, represented in the form $m_1 \perp m_2$, is defined by

$$m_1 \perp m_2(\psi_j) = \begin{cases} \sum\limits_{\psi_k \cap \psi_h = \psi_j} m_1(\psi_k) m_2(\psi_h), & \psi_j \subset \Phi, \\ 0, & \psi_j = \emptyset, \\ \left[ \sum\limits_{\psi_k \cap \psi_h = \Phi} m_1(\psi_k) m_2(\psi_h) \right] + K, & \psi_j = \Phi, \end{cases} \tag{26}$$

with

$$K = \sum_{\psi_k \cap \psi_h = \emptyset} m_1(\psi_k) m_2(\psi_h), \tag{27}$$

where $K$ is the conflict coefficient between $m_1$ and $m_2$.

In contrast to that under DRC, the conflict $K$ is delivered to the whole set $\Phi$ in Yager's combination rule.

**Definition 23** (*Dubois and Prade's combination rule*) [66]. Dubois and Prade's hybrid combination rule is defined by

$$m_1 \oplus m_2(\psi_j) = \begin{cases} \sum\limits_{\psi_k \cap \psi_h = \psi_j} m_1(\psi_k) m_2(\psi_h) + \\ \sum\limits_{\substack{\psi_k \cup \psi_h = \psi_j \\ \psi_k \cap \psi_h = \emptyset}} m_1(\psi_k) m_2(\psi_h), & \psi_j \neq \emptyset, \\ 0, & \psi_j = \emptyset. \end{cases} \tag{28}$$

The mass satisfying $\psi_k \cup \psi_h = \psi_j$ and $\psi_k \cap \psi_h = \emptyset$ is delivered to the subsets of $\Phi$ in Dubois and Prade's combination rule.

**Definition 24** (*Unified combination rule*) [67, 68]. Let $w(\psi_j)$ be a coefficient, where $w(\psi_j) \geq 0$ and $\sum\limits_{\psi_j \subseteq \Phi | \psi_j \neq \emptyset} w(\psi_j) = 1$. The unified combination rule is defined by

$$m_1 \oplus m_2(\psi_j) = \begin{cases} \sum\limits_{\psi_k \cap \psi_h = \psi_j} m_1(\psi_k) m_2(\psi_h) + w(\psi_j) K, & \psi_j \neq \emptyset, \\ 0, & \psi_j = \emptyset, \end{cases} \tag{29}$$

with

$$K = \sum_{\psi_k \cap \psi_h = \emptyset} m_1(\psi_k) m_2(\psi_h). \tag{30}$$

Note that the conflict $K$ in the unified combination rule is distributed to the subsets of $\Phi$, which is different from DRC.

**Definition 25** (*Weighted product combination rule*) [69]. Let $w(\psi_h, \psi_k)$ be a measure of intersection or set agreement. The weighted product combination rule is defined by

$$m_1 \oplus m_2(\psi_j) = \begin{cases} \kappa \sum\limits_{\psi_k \cap \psi_h = \psi_j} w(\psi_h, \psi_k) m_1(\psi_k) m_2(\psi_h), & \psi_j \neq \emptyset, \\ 0, & \psi_j = \emptyset, \end{cases} \tag{31}$$

where $\kappa$ is a normalization factor.

The weighted product combination rule is associative if satisfying $w(\psi_j \cap \psi_h, \psi_k) = w(\psi_j, \psi_h \cap \psi_k)$.

**Definition 26** (*Mahler's weighted combination rule*) [70]. Mahler's weighted combination rule is defined by

$$m_1 \oplus m_2(\psi_j) = \begin{cases} \kappa \sum\limits_{\psi_k \cap \psi_h = \psi_j} \frac{Bel(\psi_j)}{Bel(\psi_h) Bel(\psi_k)} m_1(\psi_k) m_2(\psi_h), & \psi_j \neq \emptyset, \\ 0, & \psi_j = \emptyset, \end{cases} \tag{32}$$

where $\kappa$ is a normalization factor and $Bel$ is the belief function of Eq. (5).

Unlike DRC, the weighted product combination rule and Mahler's weighted combination rule introduce $\kappa$, which is associated with the functions of $w$ and $Bel$, respectively, as a normalization factor, rather than the conflict $K$.

**Table 1.** Comparison of different combination rules in evidence theory.

| Combination rules | Axioms | | | | | |
|---|---|---|---|---|---|---|
| | **A1** | **A2** | **A3** | **A4** | **A5** | **A6** |
| Dempster [26] | yes | yes | yes | yes | yes | yes |
| Smets [46] | yes | yes | yes | yes | yes | yes |
| Yager [65] | yes | yes | no | yes | yes | yes |
| Dubois and Prade [66] | yes | yes | no | yes | yes | no |
| Inagaki and Lefevre et al. [67, 68] | yes | yes | yes | yes | yes | no |
| Zhang [69] | yes | yes | Under conditions | yes | yes | yes |
| Mahler [70] | yes | yes | yes | yes | yes | yes |
| Dezert and Smarandache [49] | yes | yes | yes | yes | yes | no |
| Deng [48] | yes | yes | yes | Under conditions | yes | yes |
| Jiang and Zhan [71] | yes | yes | yes | Under conditions | yes | yes |
| Xiao [72, 73] | yes | yes | yes | yes | yes | yes |
| GCECR | yes | yes | yes | Under conditions | yes | yes |

**Definition 27** (*Jiang and Zhan's combination rule*) [71]. Jiang and Zhan's combination rule, represented in the form $m_1 \oplus m_2$, is defined by

$$m_1 \oplus m_2(\psi_j) = \begin{cases} \frac{1}{1-K} \sum\limits_{\substack{\psi_k \cap \psi_h = \psi_j \\ \psi_k, \psi_h \subseteq \Phi}} m_1(\psi_k) m_2(\psi_h), & \psi_j \neq \emptyset, \\ \frac{1}{1-K} m_1(\emptyset) m_2(\emptyset), & \psi_j = \emptyset, \end{cases} \tag{33}$$

with

$$K = \sum_{\substack{\psi_k \cap \psi_h = \emptyset \\ \psi_k \cup \psi_h \neq \emptyset}} m_1(\psi_k) m_2(\psi_h), \tag{34}$$

where $m_1 \oplus m_2(\emptyset) = 1$ if $K = 1$ or $\sum\limits_{\psi_j \neq \emptyset} m(\psi_j) = 0$.

Jiang and Zhan's combination rule overcomes the shortcomings of the generalized combination rule in GET and has the following characteristics:

- When $m_1 \oplus m_2(\emptyset) = 0$ in Eq. (33), Jiang and Zhan's combination rule reduces to the classical DRC.
- $m_1(\emptyset)$ and $m_2(\emptyset)$ are combined through the orthogonal sum operation.
- In Eq. (33), $\frac{1}{1-K}$ is a process of normalization that is a generalization of $\frac{1}{1-K}$ in Eq. (8) of the classical DRC.
- When $m_1 \oplus m_2(\emptyset) = 0$, $K$ in Eq. (34) reduces to $K$ in Eq. (8).
- If the sum of the GBBAs of all nonempty sets is zero or $K = 1$, the whole belief is reallocated to $\emptyset$.

(2) Comparison and analysis

In accordance with the axioms A1-A6 of DRC in Section 2, we compare the combination rules described in Section 3.2. The results are summarized in Table 1. Table 1 indicates that Smets [46], and Mahler [70] satisfy axioms A1-A6: compositionality, commutativity, associativity, conditioning, internal symmetry, and autofunctionality, as does Dempster's rule [26]. By contrast, the combination rules of Yager [65] and Dubois and Prade [66] do not satisfy axiom A3: associativity; Zhang [69]'s combination rule satisfies axiom A3 under the condition that $w(\psi_j \cap \psi_h, \psi_k) = w(\psi_j, \psi_h \cap \psi_k)$; Deng [48] and Jiang and Zhan [71]'s combination rules satisfy axiom A4 when returning to a closed world, because of the combination of the empty set expressing uncertainty in an open world. Furthermore, the combination rules of Dubois and Prade [66], Inagaki and Lefevre et al. [67, 68], and Dezert and Smarandache [49] do not satisfy axiom A6: autofunctionality. The characteristics of different combination rules can be used to select an appropriate rule to handle multisource data fusion problems according to the specific application [74].

### 3.2.2 *Evidence pretreatment-based data fusion*

In this section, we review evidence pretreatment-based data fusion methods from several aspects, including evidential distance, Pignistic probability distance, correlation coefficient, belief divergence, belief entropy, and belief information quality.

(1) Evidential distance.

The classical evidential distance proposed by Jousselme et al. [75] is a useful tool to measure differences between different source data modeled by BBAs.

**Definition 28** (*Jousselme et al.'s distance*) [75]. Let $m_1$ and $m_2$ be two BBAs on $\Phi$, where $\psi_h$ and $\psi_k$ ($\psi_h, \psi_k \subseteq$

$\Phi$) are the hypotheses corresponding to $m_1$ and $m_2$, respectively. Jousselme et al.'s distance between $m_1$ and $m_2$, denoted as $d_{JGB}(m_1, m_2)$, is defined by

$$d_{JGB}(m_1, m_2) = \sqrt{\frac{(\vec{\mathbf{m}}_1 - \vec{\mathbf{m}}_2)^T \underline{\underline{D}} (\vec{\mathbf{m}}_1 - \vec{\mathbf{m}}_2)}{2}}, \tag{35}$$

where $\vec{\mathbf{m}}_1$ and $\vec{\mathbf{m}}_2$ are the vectors of BBAs $m_1$ and $m_2$, respectively; $(\vec{\mathbf{m}}_1 - \vec{\mathbf{m}}_2)^T$ is the transposition of $(\vec{\mathbf{m}}_1 - \vec{\mathbf{m}}_2)$; and $\underline{\underline{D}}$ represents a $2^{n-1} \times 2^{n-1}$ matrix with elements

$$\underline{\underline{D}}(\psi_h, \psi_k) = \frac{|\psi_h \cap \psi_k|}{|\psi_h \cup \psi_k|}. \tag{36}$$

Jousselme et al.'s distance has several desirable properties: nonnegativeness, symmetry, nondegeneracy, and triangle inequality. Since Jousselme et al.'s distance is a true metric [76] in favor of managing conflict in data fusion, several researchers have improved upon it [77].

(2) Pignistic probability distance.

The Pignistic probability transformation (PPT) function [78] can be used to measure conflict in data fusion, as will be detailed in Section 3.3. On the basis of the PPT function, Liu [79] presents a Pignistic probability distance to measure conflict between BBAs.

**Definition 29** Liu's Pignistic probability distance [79] is defined by

$$difBetP(m_1, m_2) = \max_{\psi_j \subseteq \Phi} \{|BetP_{m_1}(\psi_j) - BetP_{m_2}(\psi_j)|\}, \tag{37}$$

where $|\cdot|$ is the absolute value function.

Liu's Pignistic probability distance is also called the distance between betting commitments of BBAs.

(3) Correlation coefficient.

**Definition 30** Jiang's correlation coefficient [80] is defined by

$$C_J(m_1, m_2) = \frac{c(m_1, m_2)}{\sqrt{c(m_1, m_1) c(m_2, m_2)}}, \tag{38}$$

where $c(m_1, m_2) = \sum_{\psi_h \subseteq \Phi} \sum_{\psi_k \subseteq \Phi} m_1(\psi_h) m_2(\psi_k) \frac{|\psi_h \cap \psi_k|}{|\psi_h \cup \psi_k|}$.

(4) Belief divergence.

**Definition 31** Belief divergence, also called belief Jensen–Shannon (BJS), is defined by

$$D_X(m_1, m_2) = \frac{1}{2} \left\{ \sum_{\psi_j \subseteq \Phi} m_1(\psi_j) \log \left[ \frac{2m_1(\psi_j)}{m_1(\psi_j) + m_2(\psi_j)} \right] + \sum_{\psi_j \subseteq \Phi} m_2(\psi_j) \log \left[ \frac{2m_2(\psi_j)}{m_1(\psi_j) + m_2(\psi_j)} \right] \right\}. \tag{39}$$

Inspired by BJS divergence, various kinds of belief divergences for data fusion are exploited.

(5) Belief entropy.

**Definition 32** Deng entropy [81] is defined by

$$E_D(m) = -\sum_{\psi_j \subseteq \Phi} m(\psi_j) \log \frac{m(\psi_j)}{2^{|\psi_j|} - 1}, \tag{40}$$

in which $|\psi_j|$ is the cardinality of $\psi_j$.

When $m$ becomes a probability distribution, Deng entropy degenerates into the classical Shannon entropy. Furthermore, several basis properties and applications of Deng entropy are discussed in [82–84]. Details of other belief entropies can be found in [85–88].

(6) Belief information quality.

As a complementary of belief entropy, the belief information quality is proposed to measure the certainty/quality of information [89].

**Definition 33** Li et al.'s information quality of BBA $m$ [89] is defined by

$$IQ_{LD}(m) = \sum_{\psi_j \subseteq \Phi} \left\{ \frac{m(\psi_j)}{2^{|\psi_j|} - 1} \right\}^2, \tag{41}$$

in which $|\psi_j|$ is the cardinality of $\psi_j$.

### 3.2.3 *Other evidential conflict models for data fusion*

The classical discounting method proposed by Shafer [27] has been extended to manage conflicts in data fusion by taking into account the reliability of sources. Let us recall the basic definition.

**Definition 34** (*Shafer's discounting method*) [27]. Let $m_t$ be an arbitrary independent BBA in FOD $\Phi$. Shafer's discounting method is defined by

$$m_t^{\alpha}(\psi_j) = \begin{cases} \alpha_t m_t(\psi_j), & \psi_j \subset \Phi, \\ 1 - \alpha_t + \alpha_t m_t(\psi_j), & \psi_j = \Phi, \end{cases} \tag{42}$$

where $\alpha_t$ is a discounting coefficient used to generate a new BBA $m_t^{\alpha}$.

In Eq. (42), the values of $\alpha_t$ have the following interpretations:

- $\alpha_t = 0$ indicates that data source $t$ is completely unreliable.
- $\alpha_t = 1$ indicates that data source $t$ is completely reliable.

Clearly, in the process of fusion, discounting is beneficial for handling the conflict from multisource data in accordance with their reliability.

**Definition 35** Liu's two-dimensional conflict model [79] is defined by

$$cf(m_1, m_2) =< K, difBetP >, \tag{43}$$

where $K$ is the classical conflict coefficient of Eq. (8) and $difBetP$ is the Pignistic probability distance of Eq. (37).

Let $\xi$ be the threshold of conflict tolerance. If and only if $K > \xi$ and $difBetP > \xi$, $m_1$ and $m_2$ are conflicting.

**Definition 36** Daniel's plausibility conflict [90] is defined by

$$Pl_C(m_1, m_2) = \sum_{\phi_i \in \Phi_{PIC}(m_1, m_2)} \frac{1}{2} \left| Pl_P[m_1(\phi_i)] - Pl_P[m_2(\phi_i)] \right|, \tag{44}$$

where

$$\Phi_{PIC}(m_1, m_2) = \left\{ \phi_i \in \Phi \,\middle|\, \left[ Pl_P[m_1(\phi_i)] - \frac{1}{n} \right] \left[ Pl_P[m_2(\phi_i)] - \frac{1}{n} \right] < 0 \right\}, \tag{45}$$

and

$$Pl_P[m(\phi_i)] = \frac{Pl(\{\phi_i\})}{\sum\limits_{\phi_i \in \Phi} Pl(\{\phi_i\})}. \tag{46}$$

**Definition 37** Lefèvre and Elouedi's combination with adapted conflict (CWAC) rule [91], which adapts the weight between Dempster's rule and the conjunctive rule by means of Jousselme et al.'s distance, is defined by:

$$\begin{aligned} m_1 \oplus m_2(\psi_j) =& d_J(m_1, m_2)[m_1 \textcircled{\cap} m_2(\psi_j)] + \\ & [1 - d_J(m_1, m_2)][m_1 \oplus m_2(\psi_j)], \end{aligned} \tag{47}$$

where $m_1 \textcircled{\cap} m_2(\psi_j)$ is defined as Eq. (24) and $m_1 \oplus m_2(\psi_j)$ is defined as Eq. (7).

When $m_1 \textcircled{\cap} m_2(\emptyset) = 1$, $m_1 \oplus m_2(\emptyset) = 1$.

From the above discussion of hybrid evidential conflict models for data fusion, it can be learned that appropriately constructing a hybrid model by considering different aspects is a feasible and effective way to handle conflict in the fusion process.

### 3.3 Decision making

The outcome of evidence theory-based data fusion is related to the belief functions. Since belief functions have multiple interpretations, it is necessary to consider not only how to make a decision by defining the probabilistic transformation function for belief functions, but also that the selection of a suitable transformation function should be explained and justified. In this section, we survey several typical solutions for decision making on the basis of belief functions. Their advantages and limitations are also discussed.

(1) Pignistic probability transformation.

The classical Pignistic probability transformation (PPT) function presented by Smets and Kennes [78] can transform a BBA into a probability distribution.

**Definition 38** Smets and Kennes's PPT [78] is defined by

$$BetP_m(\phi_i) = \sum_{\psi_j \subseteq \Phi, \phi_i \in \psi_j} \frac{|\phi_i \cap \psi_j|}{|\psi_j|} \frac{m(\psi_j)}{1 - m(\emptyset)}, \tag{48}$$

where $|\psi_j|$ is the cardinality of subset $\psi_j$.

In Eq. (48), the values of $m(\emptyset)$ have the following interpretations:

- $m(\emptyset) = 0$ indicates a closed world.
- $m(\emptyset) > 0$ indicates an open world.

This process represents a kind of average assignment and is sometimes too conservative to produce appropriate distributions [92]. As a result, many researchers have attempted to improve the model from various perspectives.

(2) Sudano and Martin's probability transformation.

**Definition 39** Sudano and Martin's probability transformation by means of a mapping that is proportional to all plausibilities, denoted as PraPl [93], is defined by

$$PraPl(\phi_i) = Bel(\phi_i) + \xi Pl(\phi_i), \tag{49}$$

with

$$\xi = \frac{1 - \sum\limits_{\psi_j \subseteq \Phi} Bel(\phi_i)}{\sum\limits_{\psi_j \subseteq \Phi} Pl(\phi_i)}. \tag{50}$$

The above definition indicates that Sudano and Martin's probability transformation is based on a belief function and plausibility function. However, when certain singletons are not included in the subsets of focal elements, the PraPl probability transformation function cannot make a reasonable assignment.

(3) Cobb and Shenoy's probability transformation.

**Definition 40** Cobb and Shenoy's probability transformation function [78] is defined by

$$PnPl(\phi_i) = \frac{Pl(\phi_i)}{\sum\limits_{g} Pl(\phi_g)}. \tag{51}$$

Cobb and Shenoy's probability transformation method is a kind of plausibility normalization.

(4) Cuzzolin's probability transformation.

**Definition 41** Cuzzolin's probability transformation function [94], denoted as CuzzP, is defined by

$$CuzzP(\phi_i) = m(\phi_i) + \frac{\Delta(\phi_i)}{\sum\limits_{\phi_i \in \Phi} \Delta(\phi_i)} TNSM, \tag{52}$$

in which

$$\Delta(\phi_i) = Pl(\phi_i) - m(\phi_i), \tag{53}$$

and

$$TNSM = 1 - \sum_{\phi_i \in \Phi} m(\phi_i) = \sum_{\psi_j \in 2^\Phi || \psi_j | > 1} m(\psi_j). \tag{54}$$

The CuzzP probability transformation function considers the proportional redistribution of the total nonspecific mass (TNSM). However, CuzzP has some limitations. When $\phi_i \in \Phi$ and $\phi_i \cap \psi_j = \emptyset$, the uncertain information included in TNSM will also be allocated to $\phi_i$, where such kind of assignment is not intuitive. In addition, when $m$ is reduced to a probability distribution, Eq. (52) of the CuzzP probability transformation function is infeasible since $\Delta(\phi_i)$ of Eq. (53) is equal to 0, which makes no sense in mathematical form.

(5) DSmP probability transformation.

**Definition 42** In the DSmT theoretical framework, a Pignistic probability transformation function [49] presented by Dezert and Smarandache and denoted as DSmP is defined by

$$DSmP(\psi_j) = \sum_{\psi_h \in G^\Phi} \frac{\sum\limits_{\psi_k \subseteq \psi_j \cap \psi_h, |\psi_k|=1} m(\psi_k) + \varepsilon |\psi_j \cap \psi_h|}{\sum\limits_{\psi_k \subseteq \psi_h, |\psi_k|=1} m(\psi_k) + \varepsilon |\psi_h|} m(\psi_h), \tag{55}$$

where $\varepsilon$ is an adjustable parameter that is equal to or greater than 0 and $G^\Phi$ represents the hyper-power set consisting of the integrity limitations in DSmT.

In the DSmP probability transformation function, $\varepsilon$ is applied to integrate the classical PPT with the proportional belief transformation methods.

Additional decision-making approaches with belief functions can be found in [95].

## 4 Complex evidence theory for data fusion

Traditional evidence theory based on real numbers for data fusion has been found to not be applicable in some complex applications to represent data fluctuations at a given phase of time during their execution. Notably, CET, as a generalization of classical DSET, was presented by Xiao [72, 73] to be a solution. CET extends the classical DSET into the complex plane and is capable of modeling and handling uncertainty by means of complex numbers. The main concepts of CET are introduced below [72, 73].

### 4.1 CET for data fusion in a closed world

**Definition 43** (*Complex mass function*) A complex mass function (CMF) $\mathbb{M}$ in $\Phi$ is defined as a mapping:

$$\mathbb{M}: \quad 2^\Phi \to \mathbb{C}, \tag{56}$$

satisfying

$$\begin{aligned} &\mathbb{M}(\emptyset) = 0, \\ &\mathbb{M}(\psi_j) = \mathbf{m}(\psi_j) e^{i\theta(\psi_j)}, \quad \psi_j \subseteq \Phi, \\ &\sum_{\psi_j \subseteq \Phi} \mathbb{M}(\psi_j) = 1, \end{aligned} \tag{57}$$

where $i = \sqrt{-1}$, $\mathbf{m}(\psi_j) \in [0, 1]$ represents the magnitude/amplitude of $\mathbb{M}(\psi_j)$, and $\theta(\psi_j)$ denotes a phase term.

In Eq. (57), $\mathbb{M}(\psi_j)$ can be represented as

$$\mathbb{M}(\psi_j) = x + yi, \quad \psi_j \subseteq \Phi, \tag{58}$$

with

$$|\mathbb{M}(\psi_j)| = \mathbf{m}(\psi_j) = \sqrt{x^2 + y^2}, \quad \psi_j \subseteq \Phi, \tag{59}$$

where $\sqrt{x^2 + y^2} \in [0, 1]$.

$\mathbb{M}$ is also called a complex BBA (CBBA). As $|\mathbb{M}(\emptyset)| = 0$ indicates a closed world, a CBBA is capable of modelling and quantifying uncertainty with regards to data sources in the framework of complex plane for a closed world.

**Definition 44** (*Focal element in CET*). Let $\mathbb{M}$ be a CBBA defined in Definition 43. $\forall \psi_j \subseteq \Phi$, if $|\mathbb{M}(\psi_j)|$ or $\mathbf{m}(\psi_j) > 0$, $\psi_j$ is called a focal element in CET.

Comparison of Definitions 4-5 with Definitions 43-44 indicates that $\mathbb{M}$ in CET has the following interpretations and properties:

- The CBBA $\mathbb{M}$ in CET can be expressed by not only complex numbers but also positive real numbers, while $m$ can only be expressed by positive real numbers in DSET.
- In contrast to DSET, the value of $|\mathbb{M}(\psi_j)|$ or $\mathbf{m}(\psi_j)$ represents the degree to which the evidence supports $\psi_j$.
- When the focal elements of $\mathbb{M}$ reduce to positive real numbers, the CBBA $\mathbb{M}$ in CET degrades into the classical BBA in DSET.

**Definition 45** (*Commitment degree in CET*). The commitment degree $\mathbb{C}om(\psi_j)$ in CET committed to proposition $\psi_j$ is defined by

$$\mathbb{C}om(\psi_j) = \frac{|\mathbb{M}(\psi_j)|}{\sum\limits_{\psi_h \subseteq \Phi} |\mathbb{M}(\psi_h)|} = \frac{\mathbf{m}(\psi_j)}{\sum\limits_{\psi_h \subseteq \Phi} \mathbf{m}(\psi_h)}, \quad \psi_j \subseteq \Phi, \tag{60}$$

where $\mathbf{m}(\psi_j)$ and $\mathbf{m}(\psi_h)$ are the magnitudes of $\mathbb{M}(\psi_j)$ and $\mathbb{M}(\psi_h)$, respectively.

**Definition 46** (*Generalized belief function in CET*). A generalized belief function $GBel$ in CET, mapping from $2^\Phi$ to $[0,1]$, is defined by

$$GBel(\psi_j) = \sum_{\psi_h \subseteq \psi_j} \mathbb{C}om(\psi_h), \quad \psi_j \subseteq \Phi, \tag{61}$$

where

$$\mathbb{C}om(\psi_h) = \frac{|\mathbb{M}(\psi_h)|}{\sum\limits_{\psi_k \subseteq \Phi} |\mathbb{M}(\psi_k)|} = \frac{\mathbf{m}(\psi_h)}{\sum\limits_{\psi_k \subseteq \Phi} \mathbf{m}(\psi_k)}. \tag{62}$$

**Definition 47** (*Generalized plausibility function in CET*). A generalized plausibility function $GPl$ in CET, mapping from $2^\Phi$ to $[0,1]$, is defined by

$$\begin{aligned} GPl(\psi_j) &= 1 - GBel(\bar{\psi_j}) = 1 - \sum_{\psi_h \subseteq \bar{\psi_j}} \mathbb{C}om(\psi_h) \\ &= \sum_{\psi_h \cap \psi_j \neq \emptyset} \mathbb{C}om(\psi_h), \end{aligned} \tag{63}$$

where $\bar{\psi_j}$ is the complement of $\psi_j$, such that $\bar{\psi_j} = \Phi - \psi_j$.

Comparison of Definitions 6-7 with Definitions 46-47 indicates that the functions of $GBel$ and $GPl$ in CET have the following properties:

- Similar to those in DSET, $GBel(\psi_j)$ and $GPl(\psi_j)$ in CET are the lower and upper limit functions of $\psi_j$, respectively.
- When focal elements of $\mathbb{M}$ reduce to positive real numbers, $GBel(\psi_j)$ and $GPl(\psi_j)$ in CET degrade into the classical $Bel(\psi_j)$ and $Pl(\psi_j)$ in DSET, respectively.

**Definition 48** (*Complex evidence combination rule*) Let $\{\mathbb{M}_1, \ldots, \mathbb{M}_q, \ldots, \mathbb{M}_t\}$ be a set of independent CBBAs in FOD $\Phi$, where proposition $\psi_j \in 2^\Phi$. The complex evidence combination rule (CECR), denoted as $\mathbb{M}_1 \oplus \cdots \oplus \mathbb{M}_q \oplus \cdots \oplus \mathbb{M}_t$ is defined by:

$$\mathbb{M}_1 \oplus \cdots \oplus \mathbb{M}_q \oplus \cdots \oplus \mathbb{M}_t(\psi_j) = \begin{cases} \frac{1}{1-\mathbb{K}} \sum\limits_{\substack{\cap \psi_h = \psi_j \\ \psi_h \subseteq \Phi}} \prod\limits_{1 \leq q \leq t} \mathbb{M}_q(\psi_h), & \psi_j \neq \emptyset, \\ 0, & \psi_j = \emptyset, \end{cases} \tag{64}$$

with

$$\mathbb{K} = \sum_{\cap \psi_h = \emptyset} \prod_{1 \leq q \leq t} \mathbb{M}_q(\psi_h), \tag{65}$$

in which $\mathbb{K}$ is the complex evidence conflict coefficient (CECC) among these CBBAs.

Since $|\mathbb{M}(\emptyset)| = 0$ indicating a closed world, the CECR can merge arbitrary multiple CBBAs to provide uncertainty reasoning for data fusion in a closed world.

### 4.2 Generalized CET for data fusion in an open world

In this section, CET is generalised for data fusion in an open world, called as GCET. The basic concepts of GCET are presented as follows.

**Definition 49** (*Generalized complex mass function*) A generalized complex mass function (GCMF) $\mathbb{M}$ in FOD $\Phi$ is defined as a mapping:

$$\mathbb{M}: \quad 2^\Phi \to \mathbb{C}, \tag{66}$$

satisfying

$$\begin{aligned} &\mathbb{M}(\psi_j) = \mathbf{m}(\psi_j) e^{i\theta(\psi_j)}, \quad \psi_j \in 2^\Phi, \\ &\sum_{\psi_j \in 2^\Phi} \mathbb{M}(\psi_j) = 1, \end{aligned} \tag{67}$$

where $i = \sqrt{-1}$, $\mathbf{m}(\psi_j) \in [0,1]$ represents the magnitude of $\mathbb{M}(\psi_j)$, and $\theta(\psi_j)$ denotes a phase term.

In Eq. (67), $\mathbb{M}(\psi_j)$ can be represented as

$$\mathbb{M}(\psi_j) = x + yi, \quad \psi_j \in 2^\Phi, \tag{68}$$

with

$$|\mathbb{M}(\psi_j)| = \mathbf{m}(\psi_j) = \sqrt{x^2 + y^2}, \quad \psi_j \in 2^\Phi, \tag{69}$$

where $\sqrt{x^2 + y^2} \in [0,1]$.

$\mathbb{M}$ is also called a generalized CBBA (GCBBA). Since $|\mathbb{M}(\emptyset)| = 0$ indicating a closed world and $|\mathbb{M}(\emptyset)| > 0$ indicating an open world, a GCBBA is qualified to representing and quantifying uncertainty with respect to data sources in the framework of complex plane for both closed world and open world.

**Definition 50** (*Focal element in GCET*). Let $\mathbb{M}$ be a GCBBA defined in Definition 49. $\forall \psi_j \in 2^\Phi$, if $|\mathbb{M}(\psi_j)|$ or $\mathbf{m}(\psi_j) > 0$, $\psi_j$ is called a focal element in GCET.

Comparison of Definitions 43-44 with Definitions 49-50 indicates that, in contrast to the CBBA in CET, $\mathbb{M}$ in GCET has the following interpretations and properties:

- It is unnecessary for $|\mathbb{M}(\emptyset)| = 0$ in GCET, such that $|\mathbb{M}(\emptyset)| \geq 0$, while $|\mathbb{M}(\emptyset)|$ must be equal to 0 in CET.
- $\emptyset$ can be a focal element as $|\mathbb{M}(\emptyset)| > 0$ in GCET, but $\emptyset$ cannot be a focal element in CET.
- When $|\mathbb{M}(\emptyset)| > 0$, it is utilized to model an open world in GCET, indicating that $\emptyset$ is a focal element or the union of focal elements not within the FOD, rather than the empty set in CBBA in CET.
- When $|\mathbb{M}(\emptyset)| = 0$, the GCBBA $\mathbb{M}$ in GCET degrades into the CBBA in CET.

**Definition 51** (*Commitment degree in GCET*). The commitment degree $\mathbb{C}om(\psi_j)$ in GCET committed to proposition $\psi_j$ is defined by

$$\mathbb{C}om(\psi_j) = \frac{|\mathbb{M}(\psi_j)|}{\sum\limits_{\psi_h \in 2^\Phi} |\mathbb{M}(\psi_h)|} = \frac{\mathbf{m}(\psi_j)}{\sum\limits_{\psi_h \in 2^\Phi} \mathbf{m}(\psi_h)}, \quad \psi_j \in 2^\Phi. \tag{70}$$

**Definition 52** (*Generalized belief function in GCET*). A generalized belief function $GBel$ in GCET, mapping from $2^\Phi$ to $[0,1]$, is defined by

$$GBel(\psi_j) = \begin{cases} \sum\limits_{\psi_h \subseteq \psi_j} \mathbb{C}om(\psi_h), & \psi_j \neq \emptyset, \\ \mathbb{C}om(\emptyset), & \psi_j = \emptyset, \end{cases} \tag{71}$$

in which

$$\mathbb{C}om(\emptyset) = \frac{|\mathbb{M}(\emptyset)|}{\sum\limits_{\psi_h \in 2^\Phi} |\mathbb{M}(\psi_h)|} = \frac{\mathbf{m}(\emptyset)}{\sum\limits_{\psi_h \in 2^\Phi} \mathbf{m}(\psi_h)}. \tag{72}$$

**Definition 53** (*Generalized plausibility function in GCET*). A generalized plausibility function $GPl$ in GCET, mapping from $2^\Phi$ to $[0,1]$, is defined by

$$GPl(\psi_j) = \begin{cases} \sum\limits_{\substack{\psi_h \cap \psi_j \neq \emptyset \\ \psi_h \cup \psi_j \neq \emptyset}} \mathbb{C}om(\psi_h), & \psi_j \neq \emptyset, \\ \mathbb{C}om(\emptyset), & \psi_j = \emptyset. \end{cases} \tag{73}$$

Comparison of Definitions 46-47 with Definitions 52-53 indicates that the functions of $GBel$ and $GPl$ in GCET have the following properties:

- Similar to CET, $GBel(\psi_j)$ and $GPl(\psi_j)$ in GCET are the lower and upper limit functions of $\psi_j$, respectively.
- It is unnecessary for $GBel(\emptyset) = 0$ and $GPl(\emptyset) = 0$ in GCET, such that $GBel(\emptyset) \geq 0$ and $GPl(\emptyset) \geq 0$, while $GBel(\emptyset)$ and $GPl(\emptyset)$ must be equal to 0 in CET.
- When $|\mathbb{M}(\emptyset)| = 0$, $GBel(\psi_j)$ and $GPl(\psi_j)$ in GCET degrade into the $GBel(\psi_j)$ and $GPl(\psi_j)$ in CET, respectively.

**Definition 54** (*Generalized complex evidence combination rule*) Let $\{\mathbb{M}_1, \ldots, \mathbb{M}_q, \ldots, \mathbb{M}_t\}$ be a set of independent GCBBAs in FOD $\Phi$, where proposition $\psi_j \in 2^\Phi$. The generalized complex evidence combination rule (GCECR), denoted as $\mathbb{M}_1 \oplus \cdots \oplus \mathbb{M}_q \oplus \cdots \oplus \mathbb{M}_t$ is defined by:

$$\mathbb{M}_1 \oplus \cdots \oplus \mathbb{M}_q \oplus \cdots \oplus \mathbb{M}_t(\psi_j) = \begin{cases} \frac{1}{1-\mathbb{K}} \sum\limits_{\substack{\cap \psi_h = \psi_j \\ \psi_h \in 2^\Phi}} \prod\limits_{1 \leq q \leq t} \mathbb{M}_q(\psi_h), & \psi_j \neq \emptyset, \\ \frac{1}{1-\mathbb{K}} \prod\limits_{1 \leq q \leq t} \mathbb{M}_q(\emptyset), & \psi_j = \emptyset, \end{cases} \tag{74}$$

with

$$\mathbb{K} = \sum_{\substack{\cap \psi_h = \emptyset \\ \cup \psi_h \neq \emptyset}} \prod_{1 \leq q \leq t} \mathbb{M}_q(\psi_h). \tag{75}$$

where $\mathbb{M}_1 \oplus \cdots \oplus \mathbb{M}_q \oplus \cdots \oplus \mathbb{M}_t(\emptyset) = 1$ if $\mathbb{K} = 1$ or $\sum\limits_{\psi_j \neq \emptyset} \mathbb{M}(\psi_j) = 0$.

$\mathbb{K}$ denotes the generalized complex evidence conflict coefficient (GCECC) among GCBBAs $\{\mathbb{M}_1, \ldots, \mathbb{M}_q, \ldots, \mathbb{M}_t\}$.

GCECR has the following characteristics:

- When $\mathbb{M}_1 \oplus \cdots \oplus \mathbb{M}_q \oplus \cdots \oplus \mathbb{M}_t(\emptyset) = 0$, GCECR reduces to the CECR. Since $|\mathbb{M}(\emptyset)| = 0$ indicating a closed world and $|\mathbb{M}(\emptyset)| > 0$ indicating an open world, the GCECR can merge arbitrary multiple GCBBAs to facilitate uncertainty reasoning for data fusion, both in the closed world and open world contexts.
- $\mathbb{M}(\emptyset)$ of GCBBAs are fused by the operation of orthogonal sum.
- In Definition 54, the factor $\frac{1}{1-\mathbb{K}}$ is a process of normalization that is a generalization of $\frac{1}{1-\mathbb{K}}$ in Eq. (65) of the CECR.
- When $\mathbb{M}_1 \oplus \cdots \oplus \mathbb{M}_q \oplus \cdots \oplus \mathbb{M}_t(\emptyset) = 0$, $\mathbb{K}$ in Eq. (75) reduces to $\mathbb{K}$ in Eq. (65).
- If the sum of the GCBBAs of all nonempty sets is zero or GCECC is equal to 1, the whole belief is reallocated to $\emptyset$.

### 4.3 Analysis of the characteristics of CET and GCET

The CET and GCET inherit the merits of classical DSET and GET, respectively. They have the following attractive characteristics:

**C**1: The structure of $\mathbb{M}$ in CET and GCET can model partial or complete ignorance using complex numbers rather than real numbers, enhancing its effectiveness in addressing uncertainty modelling challenges, particularly for signal and image data. Furthermore, in GCET, $\mathbb{M}$ can represent the uncertainty arising from the incompleteness of the FOD, enabling it to model uncertainty in an open world context.

**C**2: The generalized belief function in CET and GCET also do not need experts to provide prior probabilities, in contrast to the Bayesian decision model.

**C**3: The CECR satisfies axioms A1-A6: compositionality, commutativity, associativity, conditioning, internal symmetry, and autofunctionality, as does DRC shown in Table 1. Whereas, similar to Deng [48] and Jiang

and Zhan [71]'s combination rules, GCECR satisfy axiom A4 when returning to a closed world, due to the processing of the empty set representing uncertainty in an open world.

**C**4: The CECR in CET and GCECR in GCET adhere to the associative law and commutative law, providing flexible and straightforward approaches to uncertainty reasoning for the process of data fusion in the complex plane. In contrast to CECR, GCECR facilitates uncertainty reasoning not only in a closed world but also in an open world context.

**C**5: The generalized interval $[GBel(\psi_j), GPl(\psi_j)]$ in CET and GCET, consisting of the generalized belief and plausibility functions, also provides upper and lower probabilities.

**C**6: When GCBBAs reduce to CBBAs, GCECR degenerates into CECR. Overall, CET offers an effective approach to uncertainty modelling and reasoning in a closed world context, whereas GCET is capable of uncertainty modelling and reasoning in both open world and closed world contexts.

In some cases, GCET has greater capability than CET to model and handle the uncertainty problem in data fusion. Take the recent novel Coronavirus (COVID-19) as an example; this virus is distinctly beyond the FOD due to lack of human knowledge. In this situation, because of the exhaustiveness assumption of the FOD, CET is not applicable, whereas GCET can model a focal element outside of the FOD just for the case of COVID-19 to handle such uncertainty in the open world.

Besides, to compare CET and GCET with the typical theoretical frameworks DSET, DSmT, and GET, the characteristics are summarized in Table 2. All of these evidence theories 1) can model partial or complete ignorance quantitatively; 2) do not require prior probabilities; 3) have reasoning ability; 4) satisfy the associative law and commutative law; 5) can be regarded as upper and lower probabilities; and 6) can handle uncertainty in the fusion process in a closed world context. In addition, DSmT, GET and GCET can handle uncertainty in the fusion process in an open world. On the other hand, DSmT can handle uncertainty under the FOD modeled by a hyper-power set or super-power set rather than the power set, while CET and GCET can handle uncertainty in the fusion process on the complex plane.

## 5 Algorithm and application

Pattern classification has attracted much attention in recent years [96, 97]. In this section, we focus on the closed world, and present classical and complex evidence theory framework-based multisource data fusion algorithms. Then, we apply these fusion algorithms to pattern classification to demonstrate their practicabilities.

### 5.1 Evidence theory framework-based weighted multisource data fusion algorithms

In this section, classical evidence theory framework-based weighted multisource data fusion (ETF-WMSDF) algorithms are devised based on evidential distance, Pignistic probability distance, correlation coefficient, belief divergence, belief entropy, and belief information quality for decision making, respectively.

**Problem statement:** Let $\{\phi_1, ..., \phi_i, ..., \phi_n\}$ be a set of objects to be recognized in FOD $\Phi$. Let $\mathcal{M} = \{m_1, ..., m_q, ..., m_t\}$ be a set of BBAs modeled from multisource. $\delta$ represents a threshold that is set in advance. The data fusion algorithms try to merge these given BBAs to make a decision.

Step 1: The weight for $m_q$ is calculated as:

$$W(m_q) = \begin{cases} \frac{t-1}{\sum_{p=1}^{t} d_{JGB}(m_q, m_p)}, & \text{Method A;} \\ \frac{t-1}{\sum_{p=1}^{t} difBetP(m_q, m_p)}, & \text{Method B;} \\ \frac{t-1}{\sum_{p=1}^{t} D_X(m_q, m_p)}, & \text{Method C;} \\ \sum_{p=1}^{t} C_J(m_q, m_p) - 1, & \text{Method D;} \\ e^{-E_D(m_q)}, & \text{Method E;} \\ IQ_{LD}(m_q), & \text{Method F.} \end{cases} \tag{76}$$

Note that "Methods A and B" denote the weighted methods based on belief distance functions defined in Definitions 28 and 29, respectively; "Method C" denotes the weighted method based on belief divergence defined in Definition 31; "Method D" denotes the weighted method based on belief correlation coefficient defined in Definition 30; "Method E" denotes the weighted method based on Deng entropy defined in Definition 32; "Method F" denotes the weighted method based on information quality defined in Definition 33.

Correspondingly, these ETF-WMSDF algorithms

**Table 2.** Summary of the characteristics of typical evidence theories.

| Characteristics | Evidence theories | | | | |
|---|---|---|---|---|---|
| | Traditional | Generalization | | | |
| | DSET [26, 27] | DSmT [47] | GET [48] | CET [72, 73] | GCET |
| Model partial or complete ignorance quantitatively | yes | yes | yes | yes | yes |
| Regardless of prior probabilities | yes | yes | yes | yes | yes |
| Reasoning | yes | yes | yes | yes | yes |
| Associative law | yes | yes | yes | yes | yes |
| Commutative law | yes | yes | yes | yes | yes |
| Upper and lower probabilities | yes | yes | yes | yes | yes |
| Hyper-/Super-power sets | no | yes | no | no | no |
| Closed world | yes | yes | yes | yes | yes |
| Open world | no | yes | yes | no | yes |
| Complex plane | no | no | no | yes | yes |

are denoted as ETF-WMSDF$_A$, ETF-WMSDF$_B$, ETF-WMSDF$_C$, ETF-WMSDF$_D$, ETF-WMSDF$_E$, and ETF-WMSDF$_F$.

Step 2: The weight of $m_q$ is normalised as:

$$\widetilde{W}(m_q) = \frac{W(m_q)}{\sum_{q=1}^{t} W(m_q)}. \tag{77}$$

Step 3: According to the normalised weight, a weighted average evidence $\widetilde{m}$ is generated as:

$$\widetilde{m} = \sum_{q=1}^{t} \widetilde{W}(m_q) \times m_q. \tag{78}$$

Step 4: $\widetilde{m}$ is fused $t-1$ times with DCR to obtain a final BBA:

$$\dot{m} = ((\widetilde{m} \oplus \widetilde{m})_1 \oplus \cdots \oplus \widetilde{m})_{(t-1)}. \tag{79}$$

Step 5: According to PPT function [78], we get:

$$\mathrm{Bet}(\{\phi_i\}) = \sum_{\phi_i \subseteq \psi_j} \frac{\dot{m}(\psi_j)}{|\psi_j|}. \tag{80}$$

Step 6: The largest $\mathrm{Bet}(\{\phi_\varphi\})$ is chosen by:

$$\varphi = \arg\max_{1 \le i \le n} \{\mathrm{Bet}(\{\phi_i\})\}. \tag{81}$$

Step 7: The target is determined as:

$$\begin{cases} \text{if } \mathrm{Bet}(\{\phi_\varphi\}) \ge \delta, & \phi_\varphi \text{ is the target,} \\ \text{if } \mathrm{Bet}(\{\phi_\varphi\}) < \delta, & \text{Cannot be determined.} \end{cases} \tag{82}$$

The ETF-WMSDF Algorithm 1 is as follows.

**Algorithm 1:** ETF-WMSDF.

**Input:** $\Theta = \{\phi_1, ..., \phi_i, ..., \phi_n\}$; $\mathcal{M} = \{m_1, ..., m_q, ..., m_t\}$; Threshold $\delta$;
**Output:** A decision;
**for** $q = 1; q \le t$ **do**
  Calculate the weight $W(m_q)$ for $m_q$ by Eq. (76);
**end**
**for** $q = 1; q \le t$ **do**
  Obtain the normalised weight $\widetilde{W}(m_q)$ by Eq. (77);
**end**
**for** $q = 1; q \le t$ **do**
  Calculate a weighted average evidence $\widetilde{m}$ by Eq. (78);
**end**
Obtain the fused BBA $\dot{m}$ by Eq. (79);
**for** $i = 1; i \le n$ **do**
  Calculate $\mathrm{Bet}(\{\phi_i\})$ by Eq. (80);
**end**
Select $\varphi = \arg\max_{1 \le i \le n} \{\mathrm{Bet}(\{\phi_i\})\}$ by Eq. (81);
**if** $\mathrm{Bet}(\{\phi_\varphi\}) \ge \delta$ **then**
  The target $\leftarrow \phi_\varphi$;
**else**
  Cannot be determined.
**end**

### 5.2 Complex evidence theory framework-based multisource data fusion algorithms

In this section, a complex evidence theory framework-based multisource data fusion (CETF-MSDF) algorithm is introduced [98].

**Problem statement:** Let $\{\phi_1, ..., \phi_i, ..., \phi_n\}$ be a set of objects to be recognized in FOD $\Phi$. Let $\mathcal{M} = \{\mathbb{M}_1, ..., \mathbb{M}_q, ..., \mathbb{M}_t\}$ be a set of CBBAs modeled from multisource. $\delta$ represents a threshold that is set in advance. The data fusion algorithms try to merge these given CBBAs to make a decision.

Step 1: The CBBAs of $\mathbb{M}$ are fused with CECR to obtain

a final CBBA:

$$\dot{\mathbb{M}} = ((\mathbb{M}_1 \oplus \mathbb{M})_2 \oplus \cdots \oplus \mathbb{M}_t). \tag{83}$$

Step 2: According to CPPT function [98], we get:

$$\mathrm{CBet}(\{\phi_i\}) = \sum_{\phi_i \subseteq \psi_j} \frac{\dot{\mathbb{M}}(\psi_j)}{|\psi_j|}. \tag{84}$$

Step 3: The largest $\mathrm{CBet}(\{\phi_\varphi\})$ is chosen by:

$$\varphi = \arg\max_{1 \le i \le n} \{\mathrm{CBet}(\{\phi_i\})\}. \tag{85}$$

Step 4: The target is determined as:

$$\begin{cases} \text{if } \mathrm{CBet}(\{\phi_\varphi\}) \ge \delta, & \phi_\varphi \text{ is the target,} \\ \text{if } \mathrm{CBet}(\{\phi_\varphi\}) < \delta, & \text{Cannot be determined.} \end{cases} \tag{86}$$

The CETF-MSDF Algorithm 2 is as follows.

**Algorithm 2:** CETF-MSDF.

**Input:** $\Theta = \{\phi_1, ..., \phi_i, ..., \phi_n\}$; $\mathcal{M} = \{\mathbb{M}_1, ..., \mathbb{M}_q, ..., \mathbb{M}_t\}$; Threshold $\delta$;
**Output:** A decision;
Obtain the fused BBA $\dot{\mathbb{M}}$ by Eq. (83);
**for** $i = 1; i \le n$ **do**
  Calculate $\mathrm{CBet}(\{\phi_i\})$ by Eq. (84);
**end**
Select $\varphi = \arg\max_{1 \le i \le n} \{\mathrm{CBet}(\{\phi_i\})\}$ by Eq. (85);
**if** $\mathrm{CBet}(\{\phi_\varphi\}) \ge \delta$ **then**
  The target $\leftarrow \phi_\varphi$;
**else**
  Cannot be determined.
**end**

### 5.3 Application to pattern classification

In this section, the ETF-WMSDF and CETF-MSDF algorithms are applied to pattern classification to demonstrate their practicabilities. Then, the ETF-WMSDF and CETF-MSDF algorithms are compared with related well-known works to reveal their performances.

#### *5.3.1 Descriptions of the datasets*

In this section, the performances of the ETF-WMSDF and CETF-MSDF algorithms are validated over five real-world datasets from the UCI machine learning repository (http://archive.ics.uci.edu/ml/).

- Iris dataset:
  - three classes of Iris flowers;
  - total 150 instances: each class has 50 instances;
  - without missing values;
  - each instance has the 4 attributes.
- Wine dataset:
  - three classes of Wine;
  - total 178 instances: one class is with 59, the other class is with 71 instances, and another is with 48 instances;
  - without missing values;
  - each instance has the 13 attributes.
- Heart dataset:
  - two classes of heart disease;
  - total 270 instances: one class is with 150, and the other class is with 120 instances;
  - without missing values;
  - each instance has 13 attributes.
- Parkinson's dataset:
  - two classes of Parkinson's disease;
  - total 195 instances: one class is with 48, and the other class is with 147 instances;
  - without missing values;
  - each instance has the 22 attributes.
- Australian dataset:
  - two classes of Australian credit approval;
  - total 690 instances: one class is with 383, and the other class is with 307 instances;
  - a few missing values;
  - each instance has 14 attributes.

#### *5.3.2 Implementation of ETF-WMSDF and CETF-MSDF algorithms*

In this experiment, each attribute from a dataset is considered as independent source to provide information. The missing values can be modelled as "complete ignorance" by $m(\Phi)$ of BBA and $\mathbb{M}(\Phi)$ of CBBA in the framework of DSET and CET, respectively. To implement the ETF-WMSDF and CETF-MSDF algorithms, several BBAs and CBBAs are first obtained according to training instances of each dataset using the extended methods of [98],

**Table 3.** Comparison of classification accuracies and standard deviations generated by ETF-WMSDF and CETF-MSDF algorithms.

| Dataset | ETF-based weighted multisource data fusion algorithms | | | | | | CETF-based multisource data fusion algorithm |
|---|---|---|---|---|---|---|---|
| | ETF-WMSDF$_A$ | ETF-WMSDF$_B$ | ETF-WMSDF$_C$ | ETF-WMSDF$_D$ | ETF-WMSDF$_E$ | ETF-WMSDF$_F$ | CETF-MSDF |
| Iris | 95.33±3.40% | 96.00±4.90% | 96.00±3.40% | 96.67±4.00% | 96.00±2.49% | 96.67±2.49% | 96.67±2.12% |
| Wine | 93.75±4.38% | 94.23±5.07% | 93.87±3.26% | 95.41±2.96% | 96.47±4.63% | 96.47±3.00% | 97.65±2.39% |
| Heart | 83.70±2.16% | 83.33±2.03% | 83.33±2.16% | 84.44±1.39% | 84.44±6.35% | 84.81±6.79% | 87.41±2.30% |
| Parkinson's | 83.65±6.38% | 83.18±6.84% | 82.66±8.54% | 79.62±7.14% | 80.02±5.65% | 80.02±10.04% | 83.00±3.51% |
| Australian | 82.77±2.85% | 83.21±2.87% | 83.08±3.11% | 87.38±2.56% | 86.35±2.78% | 85.50±12.88% | 88.84±8.77% |
| Average | 87.84±3.83% | 87.99±4.34% | 87.79±4.09% | 88.70±3.61% | 88.66±4.38% | 88.69±7.04% | 90.71±3.82% |

**Table 4.** Comparison of classification accuracies and standard deviations generated by different methods.

| Dataset | Classifiers | | | | | | | | ETF-based fusion methods | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NaB | NMC | kNN | REPTree | SVM | SVM-RBF | MlP | RBFN | kNN-DST | NDC | EvC | ETF-WMSDF$_D$ | CETF-MSDF |
| Iris | 94.67% | 90.67% | 95.33% | 92.00% | 94.67% | 94.67% | 93.33% | 92.67% | 95.33% | 94.00% | 94.67% | 96.67% | **96.67**% |
| Wine | 95.51% | 70.44% | 70.19% | 84.92% | 96.62% | 96.63% | 94.93% | 95.49% | 93.84% | 96.63% | 97.17% | 95.41% | **97.65**% |
| Heart | 82.59% | 60.37% | 57.78% | 70.74% | 83.70% | 82.96% | 75.19% | 81.85% | 76.30% | 82.59% | 83.70% | 84.44% | **87.41**% |
| Parkinson's | 68.75% | 70.77% | **83.02**% | 80.94% | 70.13% | 81.03% | 74.39% | 82.05% | 78.01% | 70.26% | 81.64% | 79.62% | 83.00% |
| Australian | 79.56% | 64.21% | 67.40% | 80.59% | 80.29% | 79.86% | 82.32% | 82.61% | 78.41% | 80.01% | 80.60% | 87.38% | **88.84**% |
| Average | 80.47% | 67.63% | 73.85% | 79.18% | 81.75% | 83.06% | 80.03% | 84.10% | 81.72% | 81.25% | 84.10% | 88.70% | **90.71**% |
| Std | 8.46% | 13.00% | 13.50% | 7.71% | 7.89% | 6.15% | 7.26% | 4.30% | 6.94% | 7.62% | 5.44% | 6.49% | 5.61% |

respectively. Specifically, for CBBAs generation, a transformation function of $e^{i\theta}$ is employed to convert the real values of the datasets into complex values. Here, the $\theta$ is a phase parameter varying within $[0, 2\pi]$.

Then, with reference to each testing instance, the ETF-WMSDF and CETF-MSDF algorithms are applied to fuse these generated BBAs and CBBAs, respectively, and classify the testing instance to a certain pattern. In addition, a five-fold cross validation is carried out: 80% of each dataset are randomly selected as training instances, while the rest of 20% of each dataset serves as the testing instances. We repeat this process five times, and average the accuracies of all classes, in which the results are summarized in Table 3.

It is noticed that the average classification accuracy generated by ETF-WMSDF$_A$, ETF-WMSDF$_B$, ETF-WMSDF$_C$, ETF-WMSDF$_D$, ETF-WMSDF$_E$, ETF-WMSDF$_F$, and CETF-MSDF algorithms over the five UCI datasets are 87.84±3.83%, 87.99±4.34%, 87.79±4.09%, 88.70±3.61%, 88.66±4.38%, 88.69±7.04%, and 90.71±3.82%, respectively.

### 5.3.3 *Comparison*

The ETF-WMSDF and CETF-MSDF algorithms with the best performance are compared with several well-known related works to verify their performances: 1) state-of-the-art classifiers: Naïve Bayes (NaB) [99], nearest mean classifier (NMC) [100], k-nearest neighbor (kNN) [101], Decision Tree (REPTree) [102], support vector machine (SVM) [103], SVM with radial basis function (SVM-RBF) [103], multilayer perceptron (MlP) [104], and RBF network (RBFN) [105], and 2) evidence theory framework-based fusion methods: k-nearest neighbor DS theory (kNN-DST) [106], normal distribution-based classifier (NDC) [107], and evidential calibration (EvC) [62].

In this comparison experiment, the same five-fold cross validation is carried out. The results of classification accuracies in terms of different datasets obtained by different methods are shown in Table 4, in which the optimal performance is highlighted in bold. The ETF-WMSDF$_D$ algorithm has classification accuracies: 96.67%, 95.41%, 84.44%, 79.62%, and 87.38% in terms of the Iris, Wine, Heart, Parkinson's and Australian datasets, respectively. The CETF-MSDF algorithm has classification accuracies: 96.67%, 97.65%, 87.41%, 83.00%, and 88.84% in terms of the Iris, Heart, Hepatitis, Parkinson's and Australian datasets, respectively. The CETF-MSDF algorithm obviously outperforms ETF-WMSDF$_D$ algorithm as well as other well-known methods for all but the Parkinson's dataset. For five UCI datasets, the NaB, NMC, kNN, REPTree, SVM, SVM-RBF, MlP, RBFN, kNN-DST, NDC, EvC, and ETF-WMSDF$_D$ algorithms have the following average classification accuracies: 84.22%±10.00%,

**Table 5.** Comparison of different accuracies from the maximal accuracy in terms of different datasets.

| Dataset | Classifiers | | | | | | | | ETF-based fusion methods | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NaB | NMC | kNN | REPTree | SVM | SVM-RBF | MlP | RBFN | kNN-DST | NDC | EvC | ETF-WMSDF$_D$ | CETF-MSDF |
| Iris | 2.00% | 6.00% | 1.34% | 4.67% | 2.00% | 2.00% | 3.34% | 4.00% | 1.34% | 2.67% | 2.00% | 0.00% | **0.00**% |
| Wine | 2.14% | 27.21% | 27.46% | 12.73% | 1.03% | 1.02% | 2.72% | 2.16% | 3.81% | 1.02% | 0.48% | 2.24% | **0.00**% |
| Heart | 4.82% | 27.04% | 29.63% | 16.67% | 3.71% | 4.45% | 12.22% | 5.56% | 11.11% | 4.82% | 3.71% | 2.97% | **0.00**% |
| Parkinson's | 14.27% | 12.25% | **0.00**% | 2.08% | 12.89% | 1.99% | 8.63% | 0.97% | 5.01% | 12.76% | 1.38% | 3.40% | 0.02% |
| Australian | 9.28% | 24.63% | 21.44% | 8.25% | 8.55% | 8.98% | 6.52% | 6.23% | 10.43% | 8.83% | 8.24% | 1.46% | **0.00**% |
| Accumulate | 32.50% | 97.12% | 79.86% | 44.39% | 28.17% | 18.43% | 33.42% | 18.91% | 31.69% | 30.09% | 15.80% | 10.07% | **0.02**% |

71.29%±10.45%, 74.74%±13.07%, 81.84%±6.90%, 85.08%±9.73%, 87.03%±7.13%, 84.03%±8.71%, 86.93%±5.91%, 84.38%±8.38%, 84.70%±9.63%, 87.56%±6.95%, and 88.70%±6.49%. However, the CETF-MSDF algorithm has 90.71%±5.61% average classification accuracy, which is higher than those of the other methods. These results demonstrate that the CETF-MSDF algorithm has the highest average classification accuracy over these five UCI datasets.

In addition, the differences between the average classification accuracy of each method and that of the optimal performance are calculated in Table 5. The differences across five datasets are accumulated for further evaluation of their relative performance. As shown in Figure 2, the overall accumulated difference across the five datasets for the CETF-MSDF algorithm is only 0.02%. However, the NaB, NMC, kNN, REPTree, SVM, SVM-RBF, MlP, RBFN, kNN-DST, NDC, EvC and ETF-WMSDF$_D$ algorithms have total accumulated differences of 32.50%, 97.12%, 79.86%, 44.39%, 28.17%, 18.43%, 33.42%, 18.91%, 31.69%, 30.09%, 15.80% and 10.07%, respectively. These results reveal that the superiority of the CETF-MSDF algorithm.

The proposed CETF-MSDF algorithm outperforms other methods because it utilizes CBBA to effectively model and enhance data features through the phase parameter on the occasion of generating CBBAs. To be specific, by fusing appropriate CBBAs expressed by complex numbers, constructive interference will be produced to strengthen modeling data features. Nevertheless, the computational complexity of the CETF-MSDF algorithm is higher comparing with these related methods, which limits its applicability in real-time scenarios. However, in scenarios where accuracy is critical or in certain complex number-based situations, the proposed CETF-MSDF algorithm is the preferred choice. To further improve the processing efficiency of the proposed algorithm for real-time applications, the phase can be modeled in an interpretable manner at an early stage, just like the magnitude.

## 6 Challenges and open future research directions

In this section, several challenges and open future research directions are summarized and discussed.

### 6.1 (C)BBA generation with large, heterogeneous and multi-modal data

As science and technology continue to develop, many applications have become heterogeneous sensor-based, so large, heterogeneous and multi-modal data are inevitable [108]. These data have the characteristics of large volume, high variety, but low value. How to generate appropriate BBAs, and even CBBAs, with these large, heterogeneous and multi-modal data to facilitate decisions is a challenging problem in data fusion. It remains an open issue to fuse such heterogeneous and multidimension data.

### 6.2 Combination of dependent evidence

As discussed in a previous section, DRC in DSET and the CECR in CET require independence among multiple pieces of evidence. However, dependency among some types of data is unavoidable. Several researchers have studied the fusion of dependent evidence from the perspective of combination rule modification and belief structure improvement, but some limitations of these methods restrict their application in data fusion. Therefore, it is necessary to consider how to determine the degree of dependence and how to develop the classical and complex evidence combination rules to fuse dependent evidence.

### 6.3 Conflict management

In evidence theory, when fusing conflicting data, the classical DRC and CECR may generate counterintuitive results, which impacts the effectiveness of these

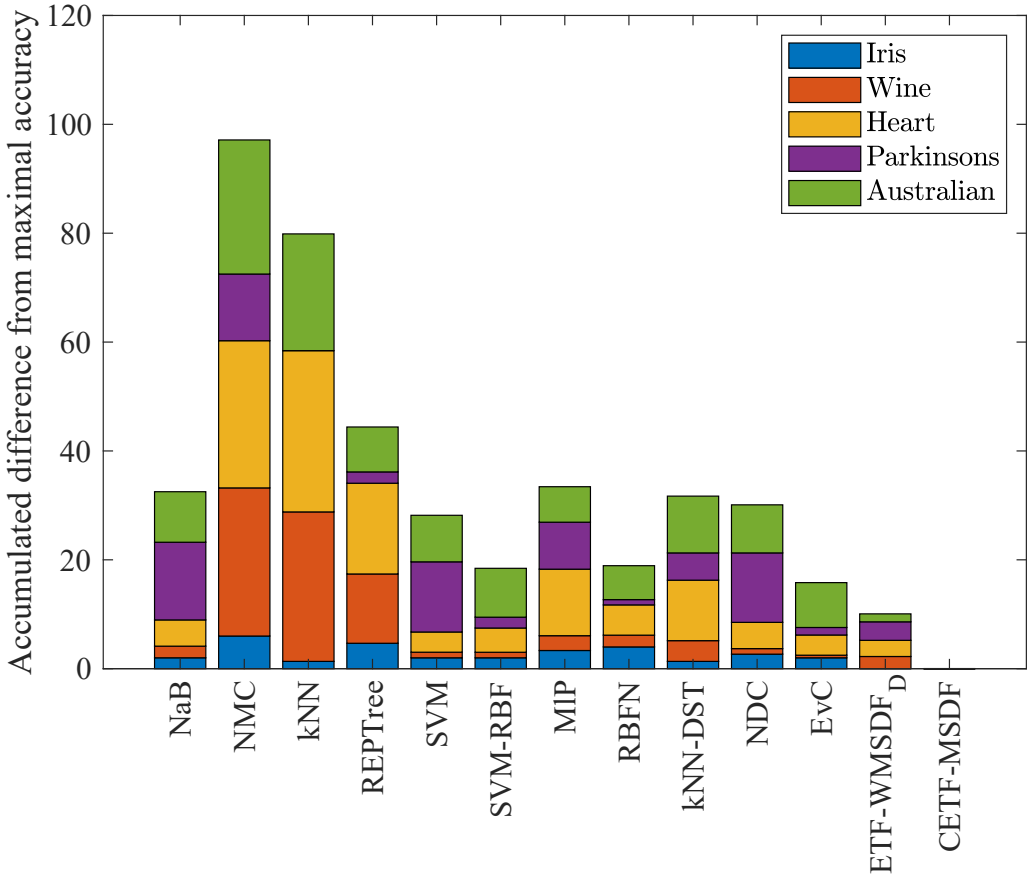


**Figure 2.** Comparison of accumulated differences from maximal accuracy for different methods.

combination rules in real-world applications of data fusion. In recent decades, the problem of conflict management in the classical DRC has been extensively studied and discussed. Various methods have been proposed, as discussed in Section 3. Careful summary and analysis indicates that no distinct conclusions have been reached. Different applications may require different or even hybrid solutions to manage the conflicting information according to the specific situation, especially for large, heterogeneous and multi-modal data. Furthermore, the current presentation of the CECR in CET requires new strategies to measure and manage conflicts from multisource data modeled in a complex plane.

### 6.4 Open world

The other limitation of evidence theory is that its FOD must be fully complete. However, in real-world applications, the targets to be detected may be unknown, for example, the detection of unknown diseases, recognition of aircraft types, and classification of unknown elements. Recently, Xiao [109] proposes a generalized quantum evidence theory (GQET) based on a quantum mechanical framework. GQET can facilitate the uncertainty reasoning of data fusion not only in the closed world (when the squared amplitude of the generalized quantum basic belief assignment for the empty set is zero $|\mathbb{Q}_{\mathbb{M}}(|\emptyset\rangle)|^2 = 0$, indicating the FOD is complete), but also in the open world (when the squared amplitude of the generalized quantum basic belief assignment for the empty set is nonzero $|\mathbb{Q}_{\mathbb{M}}(|\emptyset\rangle)|^2 > 0$, indicating the FOD is incomplete). Particularly, in a closed world, GQET degenerates, called as quantum evidence theory (QET) [109]. In summary, it provides a prospective method to uncertainty representation and reasoning in both of closed and open worlds. Hence, when the FOD is incomplete due to limited knowledge, how to select an appropriate theory, design the combination rule and manage conflicts from multisource data deserve further research and solutions.

### 6.5 Complexity and real time

In evidence theory, as the number of FODs increases, the power set of the FOD will increase exponentially; how to fuse evidence under an FOD with an abundance of elements remains an open issue. On the other

hand, big data has the characteristic of high velocity. The integration of evidence theory for real-time applications requires effective solutions. Our recent work has proposed a novel quantum Dempster's rule of combination, which constructs quantum circuits using quantum logical gates, significantly reducing the computational complexity of Dempster's rule of combination without information loss. This approach offers a promising way to handle such complexity and real-time challenges, making it worthy of further investigation.

## 7 Conclusion

In this paper, we conducted a comprehensive review of the literature on Dempster–Shafer evidence theory (DSET) for data fusion. We first introduced the basis concepts and knowledge of classical DSET and studied the axioms of Dempster's rule of combination and the characteristics and restraints of DSET. We further provided a review of the classical DSET and its extensions, collectively referred to as classical evidence theory, for data fusion from three aspects, namely, uncertainty modeling, fusion, and decision making. Particularly, in the fusion section, three main kinds of solutions for evidence theory-based data fusion were summarized, including evidential combination rule-based data fusion, evidence pretreatment-based data fusion, and other hybrid evidential conflict models for data fusion, and the typical methods and techniques were described. Next, we studied complex evidence theory for data fusion that benefits from the frame of complex plane modelling in both closed world and open world contexts. Furthermore, we presented classical and complex evidence theory framework-based multisource data fusion algorithms, which were applied to pattern classification. Through comparison with other well-known methods, complex evidence theory framework-based multisource data fusion algorithm showed its superiority to handle pattern classification problem in the complex plane. It also revealed the applicability and limitation of complex evidence theory framework-based multisource data fusion algorithm. On the basis of this survey, analysis and discussion, we present a number of challenges and open issues to help guide future research directions on evidence theory-based data fusion.

## Data Availability Statement

Data will be made available on request.

## Funding

This work was supported in part by the National Natural Science Foundation of China under Grant 62473067; in part by the Chongqing Talents: Exceptional Young Talents Project under Grant cstc2022ycjh-bgzxm0070; in part by the Chongqing Overseas Scholars Innovation Program under Grant cx2022024.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Ethical Approval and Consent to Participate

Not applicable.

## References

[1] Zhang, L., Xie, Y., Xidao, L., & Zhang, X. (2018, May). Multi-source heterogeneous data fusion. In *2018 International conference on artificial intelligence and big data (ICAIBD)* (pp. 47-51). IEEE. [CrossRef]

[2] Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., ... & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Information Fusion, 96*, 156-191. [CrossRef]

[3] Lai, J. W., Chang, J., Ang, L. K., & Cheong, K. H. (2020). Multi-level information fusion to alleviate network congestion. *Information Fusion, 63*, 248-255.[CrossRef]

[4] Yager, R. R. (2004). A framework for multi-source data fusion. *Information Sciences, 163*(1-3), 175-200. [CrossRef]

[5] Yang, J. B., Xu, D. L., Xu, X., & Fu, C. (2023). Likelihood analysis of imperfect data. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 53*(8), 5046-5057. [CrossRef]

[6] Cao, B., Li, C., Song, Y., Qin, Y., & Chen, C. (2022). Network intrusion detection model based on CNN and GRU. *Applied Sciences, 12*(9), 4184. [CrossRef]

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

[8] Judah, A., & Hu, B. (2022). An advanced data fusion method to improve wetland classification using multi-source remotely sensed data. *Sensors, 22*(22), 8942. [CrossRef]

[9] Charte, D., Charte, F., García, S., del Jesus, M. J., & Herrera, F. (2018). A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy,

models, software and guidelines. *Information Fusion, 44*, 78-96. [CrossRef]

[10] Li, T., Song, Y., & Fan, H. (2023). From target tracking to targeting track: A data-driven yet analytical approach to joint target detection and tracking. *Signal Processing, 205*, 108883. [CrossRef]

[11] Deng, X., Xue, S., & Jiang, W. (2023). A novel quantum model of mass function for uncertain information fusion. *Information Fusion, 89*, 619-631. [CrossRef]

[12] Hussain, L. A., Singh, S., Mizouni, R., Otrok, H., & Damiani, E. (2023). A predictive target tracking framework for IoT using CNN–LSTM. *Internet of Things, 22*, 100744. [CrossRef]

[13] Kang, B., & Zhao, C. (2024). Deceptive evidence detection in information fusion of belief functions based on reinforcement learning. *Information Fusion, 103*, 102102. [CrossRef]

[14] Liu, Z., Chen, F., Xu, J., Pei, W., & Lu, G. (2022). Image-text retrieval with cross-modal semantic importance consistency. *IEEE Transactions on Circuits and Systems for Video Technology, 33*(5), 2465-2476. [CrossRef]

[15] Wang, X., Zhu, D., Li, G., Zhang, X. P., & He, Y. (2022). Proposal-Copula-Based Fusion of Spaceborne and Airborne SAR Images for Ship Target Detection∗∗. *Information Fusion, 77*, 247-260. [CrossRef]

[16] Chenghai, L. I., Ke, W. A. N. G., Yafei, S. O. N. G., Peng, W. A. N. G., & Lemin, L. I. (2024). Air target intent recognition method combining graphing time series and diffusion models. *Chinese Journal of Aeronautics*. [CrossRef]

[17] Zhang, Y., Wang, X., Jiang, Z., Li, G., & He, Y. (2022). An efficient center-based method with multilevel auxiliary supervision for multiscale SAR ship detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15*, 7065-7075. [CrossRef]

[18] Lau, B. P. L., Marakkalage, S. H., Zhou, Y., Hassan, N. U., Yuen, C., Zhang, M., & Tan, U. X. (2019). A survey of data fusion in smart city applications. *Information Fusion, 52*, 357-374. [CrossRef]

[19] Ding, W., Jing, X., Yan, Z., & Yang, L. T. (2019). A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion. *Information Fusion, 51*, 129-144. [CrossRef]

[20] Deng, X., Jiang, Y., Yang, L. T., Lin, M., Yi, L., & Wang, M. (2019). Data fusion based coverage optimization in heterogeneous sensor networks: A survey. *Information Fusion, 52*, 90-105. [CrossRef]

[21] Ghamisi, P., Rasti, B., Yokoya, N., Wang, Q., Hofle, B., Bruzzone, L., ... & Benediktsson, J. A. (2019). Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of the art. *IEEE Geoscience and Remote Sensing Magazine, 7*(1), 6-39. [CrossRef]

[22] Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). A survey on machine learning for data fusion. *Information Fusion, 57*, 115-129. [CrossRef]

[23] El Fissaoui, M., Beni-hssane, A., Ouhmad, S., & El Makkaoui, K. (2021). A survey on mobile agent itinerary planning for information fusion in wireless sensor networks. *Archives of computational methods in engineering, 28*(3), 1323-1334. [CrossRef]

[24] Zhang, Y., Jiang, C., Yue, B., Wan, J., & Guizani, M. (2022). Information fusion for edge intelligence: A survey. *Information Fusion, 81*, 171-186. [CrossRef]

[25] Xinde, L. I., DUNKIN, F., & DEZERT, J. (2023). Multi-source information fusion: Progress and future. *Chinese Journal of Aeronautics*. [CrossRef]

[26] Dempster, A. P. (2008). Upper and lower probabilities induced by a multivalued mapping. In *Classic works of the Dempster-Shafer theory of belief functions* (pp. 57-72). Berlin, Heidelberg: Springer Berlin Heidelberg. [CrossRef]

[27] Shafer, G. et al. (1976). *A mathematical theory of evidence*, volume 1. Princeton University Press Princeton.

[28] Zhang, Z., Ye, S., Zhang, Y., Ding, W., & Wang, H. (2022). Belief combination of classifiers for incomplete data. *IEEE/CAA Journal of Automatica Sinica, 9*(4), 652-667. [CrossRef]

[29] Fujita, H., & Ko, Y. C. (2020). A heuristic representation learning based on evidential memberships: Case study of UCI-SPECTF. *International Journal of Approximate Reasoning, 120*, 125-137. [CrossRef]

[30] Yager, R. R. and Liu, L. (2008). *Classic works of the Dempster–Shafer theory of belief functions*, volume 219. Springer.

[31] Xiong, L., Su, X., & Qian, H. (2021). Conflicting evidence combination from the perspective of networks. *Information Sciences, 580*, 408-418. [CrossRef]

[32] Liu, P., Li, Y., Zhang, X., & Pedrycz, W. (2022). 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), 1905-1919. [CrossRef]

[33] Xu, X., Zheng, J., Yang, J. B., Xu, D. L., & Chen, Y. W. (2017). Data classification using evidence reasoning rule. *Knowledge-Based Systems, 116*, 144-151. [CrossRef]

[34] Tang, S. W., Zhou, Z. J., Hu, C. H., Yang, J. B., & Cao, Y. (2019). Perturbation analysis of evidential reasoning rule. *IEEE transactions on systems, man, and cybernetics: systems, 51*(8), 4895-4910. [CrossRef]

[35] Zhang, B., Zhang, Y., Hu, G., Zhou, Z., Wu, L., & Lv, S. (2020). A method of automatically generating initial parameters for large-scale belief rule base.

*Knowledge-Based Systems, 199*, 105904. [CrossRef]

[36] Fu, C., Hou, B., Xue, M., Chang, L., & Liu, W. (2022). 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), 251-263. [CrossRef]

[37] Zhou, Z. J., Hu, G. Y., Hu, C. H., Wen, C. L., & Chang, L. L. (2019). A survey of belief rule-base expert system. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 51*(8), 4944-4958. [CrossRef]

[38] Chang, L., Zhang, L., Fu, C., & Chen, Y. W. (2021). Transparent digital twin for output control using belief rule base. *IEEE Transactions on cybernetics, 52*(10), 10364-10378. [CrossRef]

[39] Cao, Y., Zhou, Z., Hu, C., He, W., & Tang, S. (2020). On the interpretability of belief rule-based expert systems. *IEEE Transactions on Fuzzy Systems, 29*(11), 3489-3503. [CrossRef]

[40] Xu, X., Guo, H., Zhang, Z., Shi, P., Huang, W., Li, X., & Brunauer, G. (2024). Fault diagnosis method via one vs rest evidence classifier considering imprecise feature samples. *Applied Soft Computing, 161*, 111761. [CrossRef]

[41] Xu, X., Guo, H., Zhang, Z., Yu, S., Chang, L., Steyskal, F., & Brunauer, G. (2024). A cloud model-based interval-valued evidence fusion method and its application in fault diagnosis. *Information Sciences, 658*, 119995. [CrossRef]

[42] Chen, X., & Deng, Y. (2024). Evidential software risk assessment model on ordered frame of discernment. *Expert Systems with Applications, 250*, 123786. [CrossRef]

[43] Zhou, M., Zheng, Y. Q., Chen, Y. W., Cheng, B. Y., Herrera-Viedma, E., & Wu, J. (2023). 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*, 181-199. [CrossRef]

[44] Fei, L., Liu, X., & Zhang, C. (2024). An evidential linguistic ELECTRE method for selection of emergency shelter sites. *Artificial Intelligence Review, 57*(4), 81. [CrossRef]

[45] Zadeh, L. A. (1986). A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination. *AI magazine, 7*(2), 85-85. [CrossRef]

[46] Smets, P. (2002). The combination of evidence in the transferable belief model. *IEEE Transactions on pattern analysis and machine intelligence, 12*(5), 447-458. [CrossRef]

[47] Dezert, J., & Smarandache, F. (2006). *DSmT: A new paradigm shift for information fusion*. Infinite Study.

[48] Deng, Y. (2015). Generalized evidence theory. *Applied Intelligence, 43*(3), 530-543. [CrossRef]

[49] Smarandache, F., Dezert, J., & Tchamova, A. (Eds.). (2023). Advances and Applications of DSmT for Information Fusion (Collected Works. Volume 5).

[50] Deng, Y. (2022). Random permutation set. *International Journal of Computers Communications & Control, 17*(1). [CrossRef]

[51] Deng, J., Deng, Y., & Yang, J. B. (2024). Random permutation set reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. [CrossRef]

[52] Deng, X., & Jiang, W. (2023). A framework for the fusion of non-exclusive and incomplete information on the basis of D number theory. *Applied Intelligence, 53*(10), 11861-11884. [CrossRef]

[53] Kouatli, I. (2020). The use of fuzzy logic as augmentation to quantitative analysis to unleash knowledge of participants' uncertainty when filling a survey: case of cloud computing. *IEEE Transactions on Knowledge and Data Engineering, 34*(3), 1489-1500. [CrossRef]

[54] Akcora, C. G., Gel, Y. R., Kantarcioglu, M., Lyubchich, V., & Thuraisingham, B. (2019). Graphboot: Quantifying uncertainty in node feature learning on large networks. *IEEE Transactions on Knowledge and Data Engineering, 33*(1), 116-127. [CrossRef]

[55] Fei, L., & Wang, Y. (2022). An optimization model for rescuer assignments under an uncertain environment by using Dempster–Shafer theory. *Knowledge-Based Systems, 255*, 109680. [CrossRef]

[56] An, L., Li, M., Boudaren, M. E. Y., & Pieczynski, W. (2018). Unsupervised segmentation of hidden Markov fields corrupted by correlated non-Gaussian noise. *International journal of approximate reasoning, 102*, 41-59. [CrossRef]

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

[58] Denoeux, T. (2021). NN-EVCLUS: Neural network-based evidential clustering. *Information Sciences, 572*, 297-330. [CrossRef]

[59] Zhou, K., Martin, A., Pan, Q., & Liu, Z. (2018). SELP: Semi-supervised evidential label propagation algorithm for graph data clustering. *International Journal of Approximate Reasoning, 92*, 139-154. [CrossRef]

[60] He, H., Han, D., & Dezert, J. (2020). Disagreement based semi-supervised learning approaches with belief functions. *Knowledge-Based Systems, 193*, 105426. [CrossRef]

[61] Antoine, V., Guerrero, J. A., & Xie, J. (2021). Fast semi-supervised evidential clustering. *International Journal of Approximate Reasoning, 133*, 116-132. [CrossRef]

[62] Xu, P., Davoine, F., Zha, H., & Denoeux, T. (2016). Evidential calibration of binary SVM classifiers.

*International Journal of Approximate Reasoning, 72*, 55-70. [CrossRef]

[63] Denœux, T. (2019). Logistic regression, neural networks and Dempster–Shafer theory: A new perspective. *Knowledge-Based Systems, 176*, 54-67. [CrossRef]

[64] Tong, Z., Xu, P., & Denoeux, T. (2021). An evidential classifier based on Dempster-Shafer theory and deep learning. *Neurocomputing, 450*, 275-293. [CrossRef]

[65] Yager, R. R. (1987). On the Dempster-Shafer framework and new combination rules. *Information sciences, 41*(2), 93-137. [CrossRef]

[66] Dubois, D., & Prade, H. (1988). Representation and combination of uncertainty with belief functions and possibility measures. *Computational intelligence, 4*(3), 244-264. [CrossRef]

[67] Inagaki, T. (2002). Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory. *IEEE Transactions on Reliability, 40*(2), 182-188. [CrossRef]

[68] Lefevre, E., Colot, O., & Vannoorenberghe, P. (2002). Belief function combination and conflict management. *Information fusion, 3*(2), 149-162. [CrossRef]

[69] Zhang, L. (1994). Representation, independence, and combination of evidence in the Dempster-Shafer theory. In *Advances in the Dempster-Shafer theory of evidence* (pp. 51-69). [CrossRef]

[70] Mahler, R. P. (1996). Combining ambiguous evidence with respect to ambiguous a priori knowledge. I. Boolean logic. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 26*(1), 27-41. [CrossRef]

[71] Jiang, W., & Zhan, J. (2017). A modified combination rule in generalized evidence theory. *Applied Intelligence, 46*(3), 630-640. [CrossRef]

[72] Xiao, F. (2020). Generalization of Dempster–Shafer theory: A complex mass function. *Applied Intelligence, 50*(10), 3266-3275. [CrossRef]

[73] Xiao, F. (2020). Generalized belief function in complex evidence theory. *Journal of Intelligent & Fuzzy Systems, 38*(4), 3665-3673. [CrossRef]

[74] Chen, X., & Deng, Y. (2023). A novel combination rule for conflict management in data fusion. *Soft Computing, 27*(22), 16483-16492. [CrossRef]

[75] Jousselme, A. L., Grenier, D., & Bossé, É. (2001). A new distance between two bodies of evidence. *Information fusion, 2*(2), 91-101. [CrossRef]

[76] Jousselme, A. L., & Maupin, P. (2012). Distances in evidence theory: Comprehensive survey and generalizations. *International Journal of Approximate Reasoning, 53*(2), 118-145. [CrossRef]

[77] Han, D., Dezert, J., & Yang, Y. (2016). Belief interval-based distance measures in the theory of belief functions. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 48*(6), 833-850. [CrossRef]

[78] Smets, P., & Kennes, R. (1994). The transferable belief model. *Artificial intelligence, 66*(2), 191-234. [CrossRef]

[79] Liu, W. (2006). Analyzing the degree of conflict among belief functions. *Artificial intelligence, 170*(11), 909-924. [CrossRef]

[80] Jiang, W. (2018). A correlation coefficient for belief functions. *International Journal of Approximate Reasoning, 103*, 94-106. [CrossRef]

[81] Deng, Y. (2020). Uncertainty measure in evidence theory. *Science China Information Sciences, 63*(11), 210201. [CrossRef]

[82] Abellán, J. (2017). Analyzing properties of Deng entropy in the theory of evidence. *Chaos, Solitons & Fractals, 95*, 195-199. [CrossRef]

[83] Deng, Y. (2020). Information volume of mass function. *International Journal of Computers Communications & Control, 15*(6).

[84] Liao, H., Ren, Z., & Fang, R. (2020). A Deng-entropy-based evidential reasoning approach for multi-expert multi-criterion decision-making with uncertainty. *International Journal of Computational Intelligence Systems, 13*(1), 1281-1294. [CrossRef]

[85] Zhao, T., Li, Z., & Deng, Y. (2024). Linearity in Deng entropy. *Chaos, Solitons & Fractals, 178*, 114388. [CrossRef]

[86] Cui, Y., & Deng, X. (2023). Plausibility entropy: A new total uncertainty measure in evidence theory based on plausibility function. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 53*(6), 3833-3844. [CrossRef]

[87] Qiang, C., Deng, Y., & Cheong, K. H. (2022). Information fractal dimension of mass function. *Fractals, 30*(06), 2250110. [CrossRef]

[88] Zhu, L., Zhou, Q., Deng, Y., & Cheong, K. H. (2024). Fractal-based basic probability assignment: A transient mass function. *Information Sciences, 652*, 119767. [CrossRef]

[89] Li, D., Deng, Y., & Cheong, K. H. (2021). Multisource basic probability assignment fusion based on information quality. *International Journal of Intelligent Systems, 36*(4), 1851-1875. [CrossRef]

[90] Daniel, M. (2010, June). Conflicts within and between belief functions. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (pp. 696-705). Berlin, Heidelberg: Springer Berlin Heidelberg. [CrossRef]

[91] Lefevre, E., & Elouedi, Z. (2013). How to preserve the conflict as an alarm in the combination of belief functions?. *Decision Support Systems, 56*, 326-333. [CrossRef]

[92] Abellán, J., & Bossé, É. (2016). Drawbacks of uncertainty measures based on the pignistic transformation. *IEEE Transactions on Systems, Man,*

*and Cybernetics: Systems, 48*(3), 382-388. [CrossRef]

[93] Martin, L., & Sudano, J. J. (2006, July). Yet another paradigm illustrating evidence fusion (YAPIEF). In *2006 9th international conference on information fusion* (pp. 1-7). IEEE. [CrossRef]

[94] Cuzzolin, F. (2007). Two new Bayesian approximations of belief functions based on convex geometry. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 37*(4), 993-1008. [CrossRef]

[95] Han, D., Dezert, J., & Duan, Z. (2015). Evaluation of probability transformations of belief functions for decision making. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 46*(1), 93-108. [CrossRef]

[96] Liu, Z. G., Fu, Y. M., Pan, Q., & Zhang, Z. W. (2022). Orientational distribution learning with hierarchical spatial attention for open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 45*(7), 8757-8772. [CrossRef]

[97] Liu, Z. G., Qiu, G. H., Wang, S. Y., Li, T. C., & Pan, Q. (2021). A new belief-based bidirectional transfer classification method. *IEEE Transactions on Cybernetics, 52*(8), 8101-8113. [CrossRef]

[98] Xiao, F., & Pedrycz, W. (2022). Negation of the quantum mass function for multisource quantum information fusion with its application to pattern classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 45*(2), 2054-2070. [CrossRef]

[99] Hu, B. G. (2013). What are the differences between Bayesian classifiers and mutual-information classifiers?. *IEEE transactions on neural networks and learning systems, 25*(2), 249-264. [CrossRef]

[100] Veenman, C. J., & Reinders, M. J. (2005). The nearest subclass classifier: A compromise between the nearest mean and nearest neighbor classifier. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27*(9), 1417-1429. [CrossRef]

[101] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory, 13*(1), 21-27. [CrossRef]

[102] Freund, Y., & Mason, L. (1999, June). The alternating decision tree learning algorithm. In *icml* (Vol. 99, pp. 124-133).

[103] Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST), 2*(3), 1-27. [CrossRef]

[104] Castro, C. L., & Braga, A. P. (2013). Novel cost-sensitive approach to improve the multilayer perceptron performance on imbalanced data. *IEEE transactions on neural networks and learning systems, 24*(6), 888-899. [CrossRef]

[105] CHEN S, G. M., & Grant, P. M. (1991). Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks. *IEEE Transactions on Neural Networks, 2*(2), 302-309. [CrossRef]

[106] Denoeux, T. (1995). A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE transactions on systems, man, and cybernetics, 25*(5), 804-813. [CrossRef]

[107] Xu, P., Deng, Y., Su, X., & Mahadevan, S. (2013). A new method to determine basic probability assignment from training data. *Knowledge-Based Systems, 46*, 69-80. [CrossRef]

[108] Wang, Y. M., Pan, X. H., He, S. F., Dutta, B., García-Zamora, D., & Martínez, L. (2022). A new decision-making framework for site selection of electric vehicle charging station with heterogeneous information and multigranular linguistic terms. *IEEE Transactions on Fuzzy Systems, 31*(2), 485-499. [CrossRef]

[109] Xiao, F. (2023). Generalized quantum evidence theory. *Applied Intelligence, 53*(11), 14329-14344. [CrossRef]

**Fuyuan Xiao** received the D.E. degree in Computer Science and Communication Engineering from the Graduate School of Science and Technology, Kumamoto University, Japan. She is a professor with the School of Big Data and Software Engineering, Chongqing University. She has published many papers in prestigious journals and conferences, including IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Cybernetics, IEEE Transactions on Fuzzy Systems, and IEEE Transactions on Systems, Man, and Cybernetics: Systems. She also serves as a reviewer for several prestigious journals, such as IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Fuzzy Systems, IEEE Transactions on Cybernetics, and IEEE Internet of Things Journal, etc. Her research interests include information fusion, uncertain information modeling and decision, intelligent information processing, and quantum artificial intelligence. (E-mail: xiaofuyuan@cqu.edu.cn)

**Junhao Wen** received a Ph.D. degree from Chongqing University in 2008, where he is a professor with the School of Big Data & Software Engineering. His research interests include service computing, edge computing, and software dependable engineering. He has published over 80 refereed journal and conference papers in the above areas. He has over 30 research and industrial grants and developed many commercial systems and software tools. (E-mail: jhwen@cqu.edu.cn)

**Witold Pedrycz** is a Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. He is also with the Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland. His main research directions involve computational intelligence, granular computing, and machine learning. Prof. Pedrycz is a recipient of several awards, including Norbert Wiener Award from the IEEE Systems, Man, and Cybernetics Society, IEEE Canada Computer Engineering Medal, a Cajastur Prize for Soft Computing from the European Centre for Soft Computing, a Killam Prize, a Fuzzy Pioneer Award from the IEEE Computational Intelligence Society, and 2019 Meritorious Service Award from the IEEE Systems Man and Cybernetics Society. He serves as an Editor-in-Chief for Information Sciences and WIREs Data Mining and Knowledge Discovery (Wiley) and a Co-Editor-in-Chief of International Journal of Granular Computing (Springer) and Journal of Data, Information, and Management (Springer). He is a Foreign Member of the Polish Academy of Sciences and a Fellow of the Royal Society of Canada. (E-mail: wpedrycz@ualberta.ca)

**Masayoshi Aritsugi** received the B.E. and D.E. degrees in Computer Science and Communication Engineering from Kyushu University, Japan, in 1991 and 1996, respectively. From 1996 to 2007, he was with the Department of Computer Science, Gunma University, Japan. Since 2007, he has been a Professor with Kumamoto University, Japan. His research interests include database systems and parallel/distributed data processing., Prof. Aritsugi is a Senior Member of IPSJ and IEICE and a member of ACM, IEEE and DBSJ. He was a recipient of the COMPSAC 2015 Best Paper Award, the Best Paper Award in Image Processing and Understanding in 13th IEEE International Conference on Signal Processing (ICSP2016), and the Best Paper Award in the 2019 IEEE International Cyber Science and Technology Congress (CyberSciTech). (E-mail: aritsugi@cs.kumamoto-u.ac.jp)