Evidential Reasoning with Divisive Hierarchical Clustering for Multi-source Information Fusion

Kezhu Zuo, Xinde Li, *Senior Member, IEEE*, Le Yu, Kaixuan Wu, Siyuan Li, Yilin Dong, and Zhijun Li, *Fellow, IEEE*

***Abstract*—Dempster–Shafer (DS) evidence theory provides a powerful framework for modeling uncertainty, reasoning, and combining information from multiple sources. However, it may yield counter-intuitive results when handling conflicting evidence, thereby affecting decision reliability and limiting practical applications. To address this issue, this work proposes a novel Evidential Reasoning rule with Divisive Hierarchical Clustering (ER-DHC), consisting of two main modules: evidence clustering and cluster fusion. At first, a new divisive hierarchical algorithm is introduced for evidence clustering, comprising coarse-grained and fine-grained division. In the coarse-grained stage, evidence with different decision preferences is grouped into separate clusters, thus preventing high intra-cluster conflicts and laying a solid foundation for evidence clustering. The fine-grained division adaptively refines cluster structures using an inflection point detection method, thereby enhancing clustering quality. On this basis, a new cluster fusion strategy is developed, involving intra-cluster fusion via classical Dempster's rule and inter-cluster fusion using a fuzzy preference relation-based weighted approach. This fusion strategy can degenerate into classical DS fusion and weighted fusion, while also introducing a new clustering fusion perspective, offering better flexibility. Finally, the proposed ER-DHC method is applied to the multi-source information fusion system, with experimental results demonstrating improved performance of target classification.**

***Index Terms*—Dempster–Shafer evidence theory, evidence clustering, uncertainty reasoning, fuzzy preference relation, multi-source information fusion.**

I. Introduction

In recent years, multi-source information fusion (MSIF) has been widely applied in various pattern classification tasks, including human activity recognition [1], emotion recognition [2], fault diagnosis [3], electroencephalography (EEG) analysis [4], etc. By combining more comprehensive information from multiple sources such as sensors, MSIF typically produces more reliable classification results compared to single source of information. MSIF can occur at different levels, including data, feature, and decision levels [5]. In this work, we focus on the decision-level fusion, it allows for the fusion of heterogeneous information and demonstrating excellent anti-interference capabilities [6].

As a theoretical method for decision-level fusion, Dempster-Shafer (DS) evidence theory [7] has been widely applied in MSIF systems due to its ability to model and reason under uncertainty [8]. In DS evidence theory, each information source is represented by a basic belief assignment (BBA), multiple BBAs are combined using Dempster's rule (also known as DS rule) to produce the fusion result. Unlike probability distributions, BBA can model uncertain and fuzzy information [9], making DS evidence theory more flexible and better suited for practical applications. However, when BBAs strongly support different targets, high conflict may arise, and fusion with DS rule can yield counter-intuitive results, undermining decision reliability [10].

To address the fusion of conflicting evidence, various methods have been proposed, which can be broadly classified into two strategies. The first strategy involves modifying the classical DS rule, leading to several alternative fusion rules, such as Yager's rule [11], Dubois-Prade rule [12], and proportional conflict redistribution-5 (PCR5) rule [13]. Yager's rule discards the normalization factor to handle highly conflicting information, while the Dubois-Prade rule introduces a disjunctive fusion rule that allocates conflict belief to total ignorance. The PCR5 rule distributes the conflict belief proportionally to the relevant focal elements. However, this strategy (i.e., adopting a different fusion rule) may result in the loss of the associativity property and often requires higher computational complexity than the original DS rule. The second strategy focuses on modifying the BBAs before combining them. Generally, this involves evaluating the BBAs using one or more criteria, then generating weights or discount factors based on the evaluation results, employing methods such as the normalization strategy [14] or the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) method [1]. These weights or discount factors are then used to adjust the BBAs before fusion [15]. Within this strategy, the choice of evaluation criteria is critical and can typically be grouped into two categories. The first category relies on belief divergence metrics. A lower average divergence indicates a higher credibility of the BBA, and vice versa. Representative works include the belief Jensen–Shannon divergence [14], $\mathcal{RB}$ divergence [16], and higher order fractal belief divergence [17]. The second category measures the distance between BBAs to reflect the conflict degree. A greater average distance indicates a higher

This work was supported in part by the National Natural Science Foundation of China under Grant 62233003 and 62073072, and by the Key Projects of Key Program of Jiangsu Province under Grant BE2020006 and Grant BE2020006-1, and by Shenzhen Science and Technology Program under Grant JCYJ20210324132202005 and JCYJ20220818101206014.

Corresponding author: Xinde Li.

Kezhu Zuo and Siyuan Li are with School of Cyber Science and Engineering, Southeast University, Nanjing 210096, China (Email: kezhuzuo@seu.edu.cn).

Xinde Li, Le Yu, and Kaixuan Wu are with School of Automation, Southeast University, Nanjing 210096, China (Email:xindeli@seu.edu.cn).

Yilin Dong is with College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China (Email: yldong@shmtu.edu.cn).

Zhijun Li is with School of Mechanical Engineering, Tongji University, Shanghai 201804, China (Email: zjli@ieee.org).

level of conflict, and vice versa. Representative works include Jousselme's distance [18], belief interval distance [8], and belief Hellinger distance [19]. Moreover, several studies have incorporated multiple evaluation criteria to modify BBAs to improve performance [1], [14], [20], achieving promising results. Despite these advancements, there are still limitations for evidence fusion, mainly manifesting in two aspects.

- The effectiveness of BBA modification strategy largely depends on the performance of conflict metrics. If the employed metric performs poorly, it may lead to limited improvement in fusion outcome.
- Existing studies typically categorize evidence combination into two types: the classical DS rule for low-conflict scenarios and weighted fusion for high-conflict cases. However, both approaches follow fixed combination procedures and lack the flexibility to adapt to varying degrees of evidence conflict.

To address the mentioned limitations, we propose a novel and more adaptable evidence reasoning rule (named ER-DHC), which integrates evidence clustering with fusion to enhance both performance and flexibility. First, a two-stage divisive hierarchical algorithm is developed for evidence clustering. In the coarse-grained division stage, BBAs are grouped based on their decision preferences, such that BBAs supporting different target classes are assigned to separate clusters, thus suppressing high intra-cluster conflicts. In the fine-grained division stage, the k-means algorithm is applied to further divide the clusters based on the conflict degree among BBAs, thereby optimizing cluster structure. Second, building upon the clustering results, a novel a two-stage fusion mechanism is introduced for decision-making. In the intra-cluster fusion stage, where conflicts are relatively low, BBAs are directly combined using the classical DS rule. In the inter-cluster fusion stage, where conflicts are potentially high, a weighted fusion strategy based on fuzzy preference relations is designed to manage the effects of this conflict. Notably, ER-DHC provides a flexible evidence fusion method. When all BBAs are grouped into a single cluster, it naturally degenerates into the classical DS rule. Conversely, when each BBA forms an individual cluster, it degenerates into a weighted fusion approach in [14], [21]. In addition, ER-DHC introduces a new clustering fusion perspective. This flexibility allows ER-DHC to dynamically adapt to varying levels of conflict, making it a robust solution for practical multi-source fusion tasks.

The combination of clustering and fusion has been investigated in several prior studies [22]–[24]. For instance, Li et al. [25] proposed a clustering ensemble method based on evidence theory, where mass functions are constructed by estimating the neighborhood class distribution of each sample, and the DS rule is employed to fuse base clustering results to enhance ensemble quality. Similarly, Liu et al. [26] addressed clustering with incomplete data by constructing multiple BBAs per sample and applying a discounting-based fusion strategy, thereby improving robustness in incomplete scenarios. While these approaches leverage both clustering and fusion techniques, they primarily aim to improve clustering performance. In contrast, the proposed ER-DHC is fundamentally a decision-level fusion algorithm, where clustering is employed not as the end goal, but as a mechanism to structure and guide the fusion process, where classical DS rule is used for intra-cluster fusion and a weighted strategy is applied for inter-cluster fusion, aiming to enhance the reliability of the fusion outcome rather than clustering quality. The primary contributions are summarized as follows:

- A novel divisive hierarchical algorithm is proposed for evidence clustering, including coarse-grained and fine-grained divisions, which considers not only the conflict level between BBAs, but also their decision preferences. Furthermore, an inflection point detection method is introduced to enable adaptive clustering by leveraging conflict mutations between BBAs.
- Based on the resulting evidence clusters, a flexible evidence fusion approach is proposed, incorporating intra-cluster and inter-cluster fusion. This two-stage fusion strategy adjusts the weight of conflict evidence according to the size of the respective cluster, thereby reducing its impact on the fusion.
- The proposed ER-DHC method is applied to the MSIF system. For each information source, a classifier is trained to generate probability distributions, which are transformed into BBAs based on the performance of classifier. Subsequently, the ER-DHC is executed to combine the BBAs for decision-making, and its effectiveness is verified on several public datasets.

In the rest of this article, Section II introduces the basics of DS evidence theory and fuzzy preference relations. Section III offers a detailed description of the ER-DHC. Section IV illustrates ER-DHC with several numerical examples. Section V explores the application of ER-DHC in the MSIF system. The final section concludes this work.

II. Preliminaries

A. Basics of DS Evidence Theory

In DS evidence theory, supposing that Φ is a finite set of complete propositions, denoted by

$$\Phi = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}. \tag{1}$$

The set Φ is called a frame of discernment (FoD) in DS evidence theory, and the elements are mutually exclusive within FoD. The power set 2^Φ of Φ is defined as follows:

$$2^\Phi = \{\emptyset, \sigma_1, ..., \sigma_n, \{\sigma_1, \sigma_2\}, ..., \{\sigma_1, \sigma_2, ..., \sigma_h\}, ..., \Phi\}. \tag{2}$$

A mass function, also known as a BBA, is characterized by specifying belief values for all possible subsets within a FoD. DS rule was proposed by Shafer [7] to combine BBAs over the same FoD. For two BBAs m_1 and m_2, it is represented as $m = m_1 \oplus m_2$, defined by $m(\emptyset) = 0$, and $\forall A \in 2^\Phi, A \neq \emptyset$,

$$m(A) = \frac{1}{1-K} \sum_{A_p \cap A_q = A} m_1(A_p)\, m_2(A_q) \tag{3}$$

with

$$K = \sum_{A_p \cap A_q = \emptyset} m_1(A_p)\, m_2(A_q) \tag{4}$$

where K denotes the conflict factor.

B. Basics of Fuzzy Preference Relation

Fuzzy preference relation (FPR) represents the preference between different targets [27], [28]. Assume that the set of targets is $\{m_1, m_2, ..., m_S\}$, a FPR matrix is defined as

$$FPR = \begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_S \end{array} \begin{array}{c} \begin{array}{cccc} m_1 & m_2 & \cdots & m_S \end{array} \\ \begin{bmatrix} 0.5 & p_{12} & \cdots & p_{1S} \\ p_{21} & 0.5 & \cdots & p_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ p_{S1} & p_{S2} & \cdots & 0.5 \end{bmatrix} \end{array}. \tag{5}$$

In FPR matrix, the element $p_{ij} \in [0,1](i = 1, ..., S; j = 1, ..., S)$ indicates the preference value of BBA m_i relative to BBA m_j, which is satisfied as follows:

$$p_{ii} = 0.5, \quad p_{ij} + p_{ji} = 1. \tag{6}$$

$p_{ij} > 0.5$ means that m_i is superior to m_j, and $p_{ij} = 0.5$ denotes that m_i and m_j are equally advantageous.

III. ER-DHC METHOD

This section presents the proposed ER-DHC algorithm in detail, including evidence clustering and fusion.

A. Divisive Hierarchical Clustering of Evidence

Given an MSIF system with S BBAs, the objective is to group BBAs with low conflicts into one cluster. However, traditional methods such as k-means [29], DBSCAN [30], and mean shift [31] require predefined parameters like the number of clusters or neighborhood radius, which may not ensure satisfactory clustering outcomes. Inspired by the hierarchical clustering [32], a new divisive clustering strategy is proposed. At first, S BBAs are initialized as a cluster, represented as

$$\mathcal{C}_1 = \{m_1, m_2, ..., m_S\}. \tag{7}$$

Then, a two-stage division is carried out, including coarse-grained and fine-grained divisions.

1) Coarse-grained Division: The conflict between BBAs can be reflected by their decision preferences (i.e., supported targets), which can be determined using the maximum pignistic probability [33], defined as

$$\mathcal{P}(m_i) = \arg\max_{\sigma_h \in \Phi}(BetP_i(\sigma_h)) \tag{8}$$

with

$$BetP_i(\sigma_h) = \sum_{\sigma_h \in A | A \in 2^\Phi} \frac{m_i(A)}{|A|}. \tag{9}$$

For BBAs with inconsistent preferences (i.e., supporting different targets), the conflict may be significant, and they should be assigned to different clusters. Conversely, BBAs with the same decision preference are grouped into one cluster. $\forall h \in \{1, ..., n\}$, the h-th cluster generated in the coarse-grained division stage can be defined as

$$\mathcal{C}_h = \{m_i | \mathcal{P}(m_i) = \sigma_h, i \in \{1, ..., S\}\}. \tag{10}$$

2) Fine-grained Division: To ensure that the intra-cluster conflict remains within an acceptable range, a threshold T_d is introduced, and the conflict between any two BBAs in $\mathcal{C}_h$ is expected to be less than T_d. Then, belief cosine similarity [3] is chosen as a conflict metric. Compared with other measures such as Jousselme's distance [18], Belief interval distance [8], Belief Jensen–Shannon (BJS) divergence [14], and $\mathcal{RB}$ divergence [16], belief cosine similarity possesses several desirable properties, including nonnegativity, symmetry, boundedness, extreme consistency, and sensitivity to refinement [3], which make it a strong candidate for conflict measurement. The belief cosine similarity d_C is defined as

$$d_C(m_1, m_2) = 1 - \left(\frac{\langle \vec{m}_1, \vec{m}_2 \rangle}{||\vec{m}_1|| \cdot ||\vec{m}_2||} \right)^2 \tag{11}$$

with

$$\langle \vec{m}_1, \vec{m}_2 \rangle = \sum_{p=1}^{2^n-1} \sum_{q=1}^{2^n-1} m_1(A_p) m_2(A_q) \frac{|A_p \cap A_q|}{|A_p \cup A_q|} \tag{12}$$

$$||\vec{m}|| = \left(\sum_{p=1}^{2^n-1} \sum_{q=1}^{2^n-1} m(A_p) m(A_q) \frac{|A_p \cap A_q|}{|A_p \cup A_q|} \right)^{\frac{1}{2}}. \tag{13}$$

Then, a conflict measurement set CM is constructed for $\mathcal{C}_h$ as

$$CM_h = \{d_C(m_i, m_j) | m_i, m_j \in \mathcal{C}_h, i < j\}. \tag{14}$$

By checking the maximum value of CM_h, $\mathcal{C}_h$ can be determined whether to split or not, defined as

$$\max(CM_h) < T_d. \tag{15}$$

If condition (15) is not met, $\mathcal{C}_h$ is further split. Binary k-means is a widely adopted strategy in hierarchical clustering, as it enables recursive exploration of clustering structures in an interpretable and computationally efficient manner [34]. Moreover, MSIF systems typically involve a small number of BBAs (e.g. no more than 10), and this number is further reduced within each cluster after coarse-grained division. In such cases, using a larger k value may lead over-fragmentation cluster structure. Furthermore, setting a larger k increases the risk of violating the conflict constraint. For example, if a cluster $\mathcal{C}_h$ contains only two BBAs whose conflict exceeds T_d, choosing $k \geq 3$ would prevent further splitting under the current scheme, thereby degrading the overall clustering quality. Therefore, setting $k = 2$ is generally recommended. The resulting division of $\mathcal{C}_h$ can be expressed as

$$\mathcal{C}_h^1, \ \mathcal{C}_h^2 = k\text{−means}(\mathcal{C}_h). \tag{16}$$

Due to the potential uncertainty, Euclidean distance is often inadequate for effectively measuring the conflict between BBAs. Instead, belief cosine similarity is used in the k-means algorithm as a more suitable metric. For clarity, the process of cluster splitting is shown in Algorithm 1.

The same splitting procedure is applied to $\mathcal{C}_h^1$ and $\mathcal{C}_h^2$ iteratively until all clusters satisfy (15), i.e., the termination criterion is that for any cluster, the maximum intra-cluster conflict is less than T_d.

Algorithm 1: Cluster Splitting Based on k-means and Belief Cosine Similarity

Input: Cluster $\mathcal{C}_h$
Output: Two sub-clusters: $\mathcal{C}_h^1$, $\mathcal{C}_h^2$

Randomly select two BBAs from $\mathcal{C}_h$ as initial cluster centers: $\bar{m}_h^1(\cdot)$ and $\bar{m}_h^2(\cdot)$;
repeat
  Initialize empty clusters $\mathcal{C}_h^1$ and $\mathcal{C}_h^2$;
  for $m_i \in \mathcal{C}_h$ **do**
    Compute $d_C(m_i, \bar{m}_h^1)$ and $d_C(m_i, \bar{m}_h^2)$;
    if $d_C(m_i, \bar{m}_h^1) < d_C(m_i, \bar{m}_h^2)$ **then**
      Assign m_i to $\mathcal{C}_h^1$;
    else
      Assign m_i to $\mathcal{C}_h^2$;
    end
  end
  `// Update the cluster centers`
  for $p = 1; p \leq N$ **do**
    Update $\bar{m}_h^1(A_p) = \frac{1}{|\mathcal{C}_h^1|} \sum_{m_i \in \mathcal{C}_h^1} m_i(A_p)$;
    Update $\bar{m}_h^2(A_p) = \frac{1}{|\mathcal{C}_h^2|} \sum_{m_i \in \mathcal{C}_h^2} m_i(A_p)$;
  end
until Cluster centers $\bar{m}_h^1$ and $\bar{m}_h^2$ no longer change;
return $\mathcal{C}_h^1$, $\mathcal{C}_h^2$

3) Inflection Point Detection for Adaptive Threshold Determination: In evidence clustering, selecting the threshold T_d is crucial. Generally, a small conflict between two BBAs corresponds to a lower d_C value, while significant conflict results in a higher d_C value. Thus, evidence conflict can be reflected in variations of d_C. The inflection point detection method identifies where a sequence of values undergoes the most significant change, making it suitable for determining the threshold. Specifically, the T_d is set at the conflict where the rate of change is maximized, as follows:

- *Step 1:* Considering that d_C satisfies symmetry, i.e., $d_C(m_i, m_j) = d_C(m_j, m_i)$, $\forall i, j \in \{1, ..., S\}$, compute all pairwise conflict measurements for S BBAs and sort them in ascending order, defined as

$$\mathcal{D} = sort(\{d_{ij} | d_{ij} = d_C(m_i, m_j), i < j\}). \tag{17}$$

- *Step 2:* The first derivative is computed to measure conflict changes by

$$G_t^1 = \mathcal{D}_{t+1} - \mathcal{D}_t, \quad \forall t \in \{1, ..., |\mathcal{D}| - 1\}. \tag{18}$$

The second derivative is computed to identify inflection point, which indicates a sharp change in the rate of conflict increase, defined as

$$G_t^2 = G_{t+1}^1 - G_t^1, \quad \forall t \in \{1, ..., |G_t^1| - 1\}. \tag{19}$$

The inflection point is determined as the index where the second derivative reaches its maximum:

$$t* = \arg\max_t (G_t^2) + 2. \tag{20}$$

- *Step 3:* The threshold is set as $t*$-th value in set $\mathcal{D}$ as

$$T_d = \mathcal{D}_{t*}. \tag{21}$$

In the two-stage evidence clustering strategy, coarse-grained division considers the decision preferences of the evidence. BBAs with different decision preferences typically indicates significant conflict and is therefore assigned to separate clusters, effectively preventing high intra-cluster conflict at an early stage. Based on this initial separation, the fine-grained division further refines the cluster structure. A conflict set $\mathcal{D}$ is constructed for all BBAs, where the inflection point corresponds to a conflict mutation and is used as an adaptive threshold. This means that BBAs with low conflict are grouped together to avoid excessive fragmentation, while BBAs with conflicts exceeding the mutation point are separated into distinct clusters to optimize intra-cluster conflicts, thereby improving the quality of clustering.

B. A Novel Evidence Fusion Algorithm

Suppose S BBAs are grouped into r clusters, i.e., $\mathcal{C}_1, ..., \mathcal{C}_r$. Then, the combination proceeds in two stages: intra-cluster and inter-cluster fusion.

1) Intra-cluster Fusion: Since threshold T_d ensures that intra-cluster conflict is acceptable, BBAs in $\mathcal{C}_h (h = 1, ..., r)$ are directly fused using the DS rule. Let $\mathcal{C}_h = \{m_{h1}, m_{h2}, ..., m_{hv}\}$, the intra-cluster fusion is defined as

$$\hat{m}_h = m_{h1} \oplus m_{h2} \oplus \cdots \oplus m_{hv}. \tag{22}$$

The fusion result $\hat{m}_h$ serves as the representative BBA for $\mathcal{C}_h$. If $\mathcal{C}_h$ contains only one BBA, i.e., $\mathcal{C}_h = m_{h1}$, the representative BBA is simply $\hat{m}_h = m_{h1}$.

2) Inter-cluster Fusion: Essentially, inter-cluster fusion involves combining the representative BBAs of each cluster. Given the potentially high conflict among clusters, it is essential to assign an appropriate weight to $\hat{m}_h$ and adopt a weighted fusion strategy to ensure reliable decision-making. The specific process is as follows:

- *Step 1:* The average conflict of each BBA is calculated using belief cosine similarity as

$$\widetilde{d_C}(m_i) = \frac{1}{S-1} \sum_{j=1, j \neq i}^{S} d_C(m_i, m_j). \tag{23}$$

- *Step 2:* The support degree $Sup(\cdot)$ of each BBA is acquired by

$$Sup(m_i) = \frac{1}{\widetilde{d_C}(m_i)}. \tag{24}$$

- *Step 3:* On the basis of support degree of any m_i and m_j, a FPR matrix is constructed following the method described in [35], where each element p_{ij} in FPR is calculated as

$$p_{ij} = \frac{1}{S} \sum_{v=1}^{S} (f_{iv} + f_{vj}) - 0.5 \tag{25}$$

with

$$f_{ij} = \frac{Sup(m_i)}{Sup(m_i) + Sup(m_j)}. \tag{26}$$

Note that the FPR matrix captures the support degree preferences among S BBAs and has a size of $S \times S$. Then, based on the FPR matrix, the credibility degree $R(\cdot)$ of each BBA is defined as

$$R(m_i) = \frac{2}{S^2} \sum_{j=1}^{S} p_{ij}. \tag{27}$$

- *Step 4:* The fusion weight w_h of representative BBA $\hat{m}_h$ for cluster $\mathcal{C}_h$ is defined as

$$w_h = \sum_{m_i \in \mathcal{C}_h} R(m_i). \tag{28}$$

- *Step 5:* The weighted average BBA $\tilde{m}$ of r representative BBAs is calculated as

$$\tilde{m} = \sum_{h=1}^{r} (w_h \cdot \hat{m}_h). \tag{29}$$

The final fusion result is generated using the DS rule as

$$m_{fusion} = (((\tilde{m} \oplus \tilde{m})_1 \oplus \tilde{m})_2 \oplus \cdots \oplus \tilde{m})_{r-1}. \tag{30}$$

For clarity, the pseudo-code of ER-DHC is presented in Algorithm 2.

Considering the potential high-conflict scenarios during inter-cluster fusion, a robust weighting mechanism is essential. Since the weight of each cluster (i.e., its representative BBA) is derived from the credibility degrees of BBAs within the cluster, as defined in (28), accurately estimating these credibility values is critical. Inspired by [35], [36], a FPR strategy is employed to compute credibility based on the support degrees of BBAs. This strategy leverages pairwise comparisons to quantify the extent to which one BBA is preferred over another. Such mechanism captures more detailed relative relationships, reflecting both the individual support level of a BBA and its comparative standing among others. As a result, it produces more balanced and robust credibility estimates. In high-conflict scenarios, conflicting BBAs may exhibit abnormally high or low support level. When using absolute support directly, such extremes can result in disproportionately large or small credibility values, undermining the reliability of weight assignment. In contrast, the FPR-based approach mitigates the influence of extreme cases by considering pairwise support constraints for mutual comparison, thereby enhancing the reliability of weight outcome.

3) Properties of ER-DHC: In previous studies, two strategies have been used to combine BBAs, as follows:

- *Classical DS fusion* [7]: S BBAs are directly combined using DS rule, defined as

$$m_{DS} = m_1 \oplus m_2 \oplus \cdots \oplus m_S. \tag{31}$$

- *Weighted fusion* [14], [21]: The weights $w_1, w_2, ..., w_S$ of BBAs are calculated using a metric such as belief divergence or distance. The weighted average BBA $\tilde{m}$ is then determined and fused $S-1$ times, defined as:

$$\tilde{m} = w_1 m_1 + w_2 m_2 + \cdots + w_S m_S. \tag{32}$$

$$m_{wf} = (((\tilde{m} \oplus \tilde{m})_1 \oplus \tilde{m})_2 \oplus \cdots \oplus \tilde{m})_{S-1}. \tag{33}$$

Algorithm 2: ER-DHC

Input: S BBAs m_1, m_2,..., m_S
Output: Fusion result m_{fusion}

Initialize cluster as $\mathcal{C}_1 = \{m_1, m_2, \ldots, m_S\}$;
// Step 1: Coarse-grained division
for $i = 1; i \leq S$ **do**
  Calculate decision preference $\mathcal{P}(m_i)$ using (8);
end
Divide BBAs into sub-cluster $\mathcal{C}_h$ using (10);
// Step 2: Determine threshold T_d
Compute sorted pairwise conflict measurements for S BBAs using (17);
Obtain threshold T_d based on the maximum value of second derivative using (18) - (21);
// Step 3: Fine-grained division
while $\mathcal{C}_h$ with $|\mathcal{C}_h| \geq 2$ **do**
  Compute CM_h for cluster $\mathcal{C}_h$ using (14);
  if $\max(CM_h) < T_d$ **then**
    Stop splitting $\mathcal{C}_h$;
    continue;
  end
  Apply k-means method to split $\mathcal{C}_h$;
  Replace $\mathcal{C}_h$ with two new sub-clusters ;
end
Generate evidence clustering result: $\mathcal{C}_1, ..., \mathcal{C}_r$;
// Step 4: Intra-cluster fusion
for $h = 1; h \leq r$ **do**
  Obtain intra-cluster fusion result $\hat{m}_h$ using (22);
end
// Step 5: Inter-cluster fusion
for $i = 1; i \leq S$ **do**
  Calculate average conflict of m_i by (23);
  Obtain credibility degree of m_i using (24) - (27);
end
for $h = 1; h \leq r$ **do**
  Calculate the fusion weight w_h of $\hat{m}_h$ using (28);
end
Obtain the weighted average BBA $\tilde{m}$ using (29);
Generate inter-cluster fusion result m_{fusion} by (30);

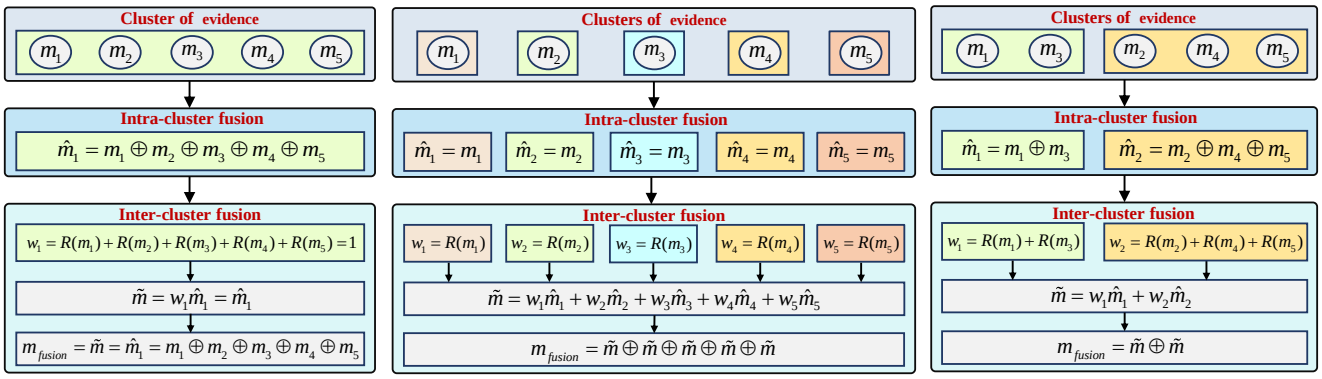
ER-DHC offers a new fusion perspective with two properties.

- ***Property 1:*** When $r = 1$, all S BBAs are grouped into a single cluster, denoted as $\mathcal{C}_1 = \{m_1, m_2, ..., m_S\}$, indicating no significant conflict among them. In this case, ER-DHC degenerates to the classical DS fusion.
  Proof: In intra-cluster fusion stage, from (22), we have

$$\hat{m}_1 = m_1 \oplus m_2 \oplus \cdots \oplus m_S. \tag{34}$$

In inter-cluster fusion stage, according to (28), the fusion

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2025.3601509



(a) Case 1: $r = 1$ (b) Case 2: $r = S$ (c) Case 3: $1 < r < S$

Fig. 1. The fusion process of ER-DHC with different clustering results.

weight w_1 of $\hat{m}_1$ is calculated as

$$\begin{aligned} w_1 = \sum_{m_i \in \mathcal{C}_1} R(m_i) =& R(m_1) + R(m_2) + \cdots + R(m_S) \\ =& \frac{Sup(m_1)}{\sum_{r=1}^{S} Sup(m_r)} + \frac{Sup(m_2)}{\sum_{r=1}^{S} Sup(m_r)} + \cdots \\ &+ \frac{Sup(m_S)}{\sum_{r=1}^{S} Sup(m_r)} = \frac{\sum_{r=1}^{S} Sup(m_r)}{\sum_{r=1}^{S} Sup(m_r)} = 1. \end{aligned} \tag{35}$$

According to (29), the weighted BBA $\tilde{m}$ is derived as

$$\tilde{m} = \sum_{h=1}^{r} (w_h \cdot \hat{m}_h) = w_1 \cdot \tilde{m}_1 = \hat{m}_1. \tag{36}$$

Based on (30), we have

$$m_{fusion} = \tilde{m} = \hat{m}_1 = m_1 \oplus m_2 \oplus \cdots \oplus m_S. \tag{37}$$

The fusion is identical to that of classical DS fusion m_{DS}.

- ***Property 2:*** When $r = S$, that is, S BBAs are grouped into S clusters. Each BBA is treated as an individual cluster, and $\mathcal{C}_1 = \{m_1\}, \mathcal{C}_2 = \{m_2\}, ..., \mathcal{C}_S = \{m_S\}$, ER-DHC degenerates to the weighted fusion in [14], [21].
Proof: In the intra-cluster fusion stage, we have

$$\hat{m}_1 = m_1, \hat{m}_2 = m_2, ..., \hat{m}_S = m_S. \tag{38}$$

In the inter-cluster fusion stage, the fusion weight w_h of $\hat{m}_h$ is obtained as

$$w_1 = R(m_1), w_2 = R(m_2), ..., w_S = R(m_S). \tag{39}$$

According to (29), the weighted BBA $\tilde{m}$ is derived as

$$\begin{aligned} \tilde{m} &= \sum_{h=1}^{r} (w_h \cdot \hat{m}_h) \\ &= w_1\hat{m}_1 + w_2\hat{m}_2 + \cdots + w_S\hat{m}_S \\ &= w_1 m_1 + w_2 m_2 + \cdots + w_S m_S. \end{aligned} \tag{40}$$

Using (30), we derive the fusion result as

$$m_{fusion} = (((\tilde{m} \oplus \tilde{m})_1 \oplus \tilde{m})_2 \oplus \cdots \oplus \tilde{m})_{S-1}. \tag{41}$$

The fusion is identical to that of weighted fusion m_{wf}.

Moreover, when $1 < r < S$, we provide a new clustering fusion technique, as presented in (22) - (30). Assuming five BBAs m_1, m_2, m_3, m_4, and m_5 are grouped into r cluster(s), which can be summarized into three cases:

- *Case 1:* $r = 1$. Five BBAs are clustered into one cluster, i.e. $\mathcal{C}_1 = \{m_1, m_2, m_3, m_4, m_5\}$. The specific fusion process is shown in Fig. 1(a).
- *Case 2:* $r = S$ ($S = 5$). Five BBAs are clustered into five clusters, i.e. $\mathcal{C}_1 = \{m_1\}$, $\mathcal{C}_2 = \{m_2\}$, $\mathcal{C}_3 = \{m_3\}$, $\mathcal{C}_4 = \{m_4\}$, and $\mathcal{C}_5 = \{m_5\}$. The specific fusion process is shown in Fig. 1(b).
- *Case 3:* $1 < r < S$. For example, five BBAs are clustered into two clusters, where $\mathcal{C}_1 = \{m_1, m_3\}$ and $\mathcal{C}_2 = \{m_2, m_4, m_5\}$. The specific fusion process is shown in Fig. 1(c).

In summary, ER-DHC not only degenerates into classical DS fusion and weighted fusion, but also introduces a new clustering combination perspective, offering greater flexibility. When evidence conflicts arise, BBAs with different decision preferences are assigned to separate clusters, ensuring the number of clusters is greater than 1, i.e., $r > 1$. In this case, ER-DHC can no longer be degenerated into the classical DS fusion for any T_d, preventing counter-intuitive results in handling conflicting evidence.

4) Computational Complexity of ER-DHC: Let S, N, and r denote the number of BBAs, the number of focal elements, and the number of resulting clusters, respectively. The ER-DHC comprises five main steps, with corresponding computational complexity summarized as follows:

- *Step 1: Coarse-grained division.* The dominant operation of this step is the computation of pignistic probability. For all S BBAs, the overall complexity is $\mathcal{O}(S \times N)$.
- *Step 2: Threshold determination.* This step involves calculating all pairwise conflict measures among the S BBAs. Each pairwise conflict requires vector operations of length N, resulting in a complexity of $\mathcal{O}(S^2 \times N^2)$.
- *Step 3: Fine-grained division.* Clusters are recursively split until the maximum intra-cluster conflict falls below a predefined threshold. At each recursion level, the conflict computation incurs a complexity of $\mathcal{O}(S^2 \times N^2)$, and the average recursion depth is $\mathcal{O}(\log S)$. Thus, the total complexity is $\mathcal{O}(S^2 \times N^2 \times \log S)$.
- *Step 4: Intra-cluster fusion.* BBAs within each cluster are fused using the classical DS rule. Aggregating across all clusters, the total complexity is $\mathcal{O}(S \times N^2)$.
- *Step 5: Inter-cluster fusion.* The r cluster-representative BBAs are fused using a weighted strategy. The dominant

cost arises from calculating average conflict among r BBAs, resulting in complexity of $\mathcal{O}(r^2 \times N^2)$.

Overall, fine-grained division (Step 3) dominates the computational cost. Therefore, the total time complexity of ER-DHC is $\mathcal{O}(S^2 \times N^2 \times \log S)$.

IV. Numerical Examples and Analysis

A. Implementation Details of ER-DHC

To better show the implementation details of ER-DHC, a numerical example is presented for illustration.

Example 1: Suppose the FoD is $\Phi = \{\sigma_1, \sigma_2, \sigma_3\}$, six BBAs m_1, m_2, m_3, m_4, m_5, and m_6 are given in Table I.

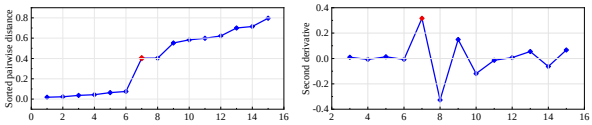
TABLE I
BBAs in Example 1

BBA	σ_1	σ_2	σ_3	$\sigma_1 \cup \sigma_2$
m_1	0.50	0.08	0.10	0.32
m_2	0.60	0.09	0.01	0.30
m_3	0.01	0.40	0.38	0.21
m_4	0.43	0.06	0.01	0.50
m_5	0.40	0.19	0.01	0.40
m_6	0.30	0.70	0.0	0.0

Step 1: Coarse-grained division. The decision preference of each BBA is calculated using (8) as $\mathcal{P}(m_1) = \mathcal{P}(m_2) = \mathcal{P}(m_4) = \mathcal{P}(m_5) = \sigma_1$, and $\mathcal{P}(m_3) = \mathcal{P}(m_6) = \sigma_2$. Hence, two clusters are generated as $\mathcal{C}_1 = \{m_1, m_2, m_4, m_5\}$ and $\mathcal{C}_2 = \{m_3, m_6\}$ in coarse-grained division.

Step 2: Threshold Determination. A threshold T_d is determined using the inflection point detection, as follows:

- *Step 2-1:* The sorted pairwise conflict measurement set for six BBAs is obtained using (17) as $\mathcal{D} = \{0.019, 0.023, 0.037, 0.044, 0.063, 0.075, 0.402, 0.403, 0.553, 0.583, 0.6, 0.622, 0.7, 0.715, 0.796\}$ and presented in Fig. 2(a).
- *Step 2-2:* The second derivative is computed for set $\mathcal{D}$, as shown in Fig. 2(b).



(a) Sorted conflict measure set $\mathcal{D}$ (b) Second derivative of $\mathcal{D}$

Fig. 2. Sorted pairwise conflict measurement and its second derivative.

- *Step 2-3:* From Fig. 2(b), the second derivative at 7-th point is the highest (i.e. inflection point), therefore the threshold is determined as $T_d = \mathcal{D}_7 = 0.402$.

Step 3: Fine-grained division. According to T_d, we determine whether $\mathcal{C}_1$ and $\mathcal{C}_2$ further split.

- *Step 3-1:* Conflict measure set constructed using (14) for $\mathcal{C}_1$ is $CM_1 = \{0.019, 0.037, 0.063, 0.044, 0.075, 0.023\}$. Since $\max(CM_1) = 0.075 < T_d$, $\mathcal{C}_1$ does not require further splitting.
- *Step 3-2:* Similarly, conflict measurement set for $\mathcal{C}_2$ is constructed as $CM_2 = \{0.402\}$. Since $\max(CM_2) = 0.402 = T_d$, condition (15) is not met, therefore $\mathcal{C}_2$ needs to be split. Using the k-means algorithm, $\mathcal{C}_2$ is split into two new sub-clusters: $\mathcal{C}_2^1 = \{m_3\}$ and $\mathcal{C}_2^2 = \{m_6\}$.
- *Step 3-3:* Since $\mathcal{C}_1$, $\mathcal{C}_2^1$, and $\mathcal{C}_2^2$ all meet condition (15), the splitting process ends. The updated clusters are $\mathcal{C}_1 = \{m_1, m_2, m_4, m_5\}$, $\mathcal{C}_2 = \{m_3\}$, and $\mathcal{C}_3 = \{m_6\}$.

Step 4: Intra-cluster fusion. The representative BBAs are generated using (22), where $\hat{m}_1 = m_1 \oplus m_2 \oplus m_4 \oplus m_5$, $\hat{m}_2 = m_3$, and $\hat{m}_3 = m_6$, as follows :

$$\hat{m}_h = \begin{array}{c} \\ \hat{m}_1 \\ \hat{m}_2 \\ \hat{m}_3 \end{array} \begin{array}{c} \begin{array}{cccc} \sigma_1 & \sigma_2 & \sigma_3 & \sigma_1 \cup \sigma_2 \end{array} \\ \begin{bmatrix} 0.911 & 0.056 & 0.0 & 0.033 \\ 0.010 & 0.40 & 0.380 & 0.210 \\ 0.30 & 0.70 & 0.0 & 0.0 \end{bmatrix} \end{array}.$$

Step 5: Inter-cluster fusion.

- *Step 5-1:* The average conflict of m_i is calculated as

$$\widetilde{d_C}(m_1) = 0.280, \widetilde{d_C}(m_2) = 0.307, \widetilde{d_C}(m_3) = 0.647,$$
$$\widetilde{d_C}(m_4) = 0.274, \widetilde{d_C}(m_5) = 0.237, \widetilde{d_C}(m_6) = 0.508.$$

- *Step 5-2:* The support degree of m_i is acquired as

$$Sup(m_1) = 3.565, Sup(m_2) = 3.262,$$
$$Sup(m_3) = 1.545, Sup(m_4) = 3.647,$$
$$Sup(m_5) = 4.217, Sup(m_6) = 1.967.$$

- *Step 5-3:* According to the support degree, the FPR matrix is obtained using (5), (25) and (26) as $FPR =$

$$\begin{bmatrix} 0.5 & 0.521 & 0.698 & 0.495 & 0.460 & 0.643 \\ 0.479 & 0.5 & 0.677 & 0.473 & 0.439 & 0.621 \\ 0.302 & 0.323 & 0.5 & 0.297 & 0.262 & 0.445 \\ 0.505 & 0.527 & 0.703 & 0.5 & 0.466 & 0.648 \\ 0.540 & 0.561 & 0.738 & 0.534 & 0.5 & 0.683 \\ 0.357 & 0.379 & 0.555 & 0.352 & 0.317 & 0.5 \end{bmatrix}.$$

Based on the FPR matrix, the credibility degree of each BBA is obtained using (27) as

$$R(m_1) = 0.184, R(m_2) = 0.177, R(m_3) = 0.118,$$
$$R(m_4) = 0.186, R(m_5) = 0.198, R(m_6) = 0.137.$$

- *Step 5-4:* Since $\mathcal{C}_1 = \{m_1, m_2, m_4, m_5\}$, $\mathcal{C}_2 = \{m_3\}$, and $\mathcal{C}_3 = \{m_6\}$, the weight w_h of $\hat{m}_h$ is derived by (28) as

$$w_1 = R(m_1) + R(m_2) + R(m_4) + R(m_5) = 0.745,$$
$$w_2 = R(m_3) = 0.118, w_3 = R(m_6) = 0.137.$$

- *Step 5-5:* The final fusion result is generated using (29) and (30) as

$$m_{fusion}(\sigma_1) = 0.973, m_{fusion}(\sigma_2) = 0.027,$$
$$m_{fusion}(\sigma_3) = 0.0, m_{fusion}(\sigma_1 \cup \sigma_2) = 0.0.$$

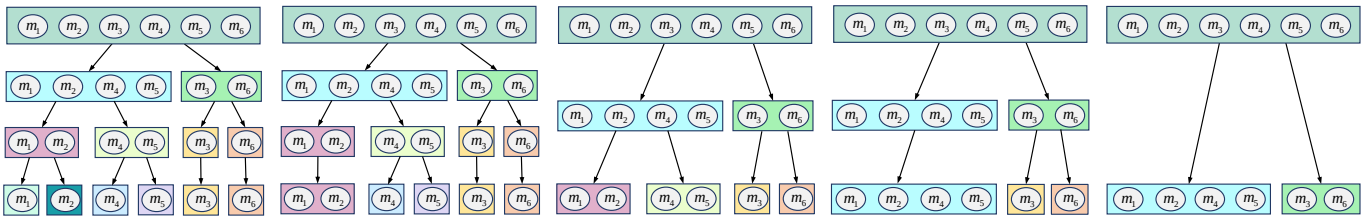
Moreover, ER-DHC is compared with several existing methods. As shown in Table II, Xiao's approach achieves the best performance among the seven existing methods, with the maximum belief value of 0.859 for σ_1, whereas ER-DHC further improves this, reaching 0.973, proving its superiority.

Besides belief cosine similarity, ER-DHC can also employ other conflict metrics. As shown in Table III, the use of belief cosine similarity yields the best performance.

B. Impact of Threshold on Evidence Clustering and Fusion

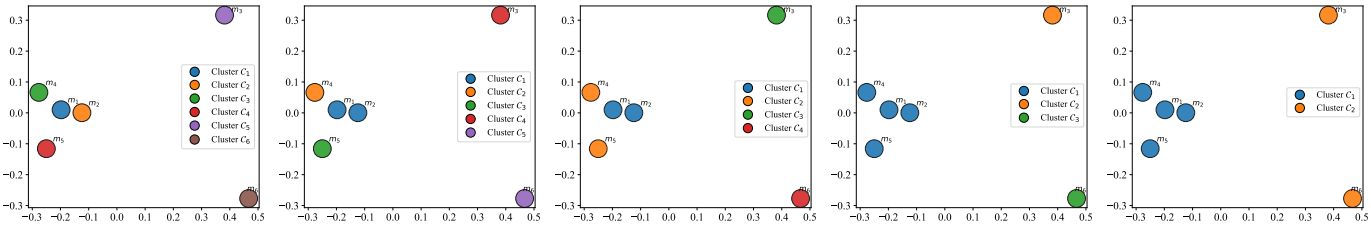
Example 2: For the six BBAs listed in Table I, different thresholds are employed to test the proposed two-stage divisive hierarchical clustering method.

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2025.3601509



(a) $0 \leq T_d \leq 0.019$ (b) $0.019 < T_d \leq 0.023$ (c) $0.023 < T_d \leq 0.075$ (d) $0.075 < T_d \leq 0.402$ (e) $0.402 < T_d \leq 1$

Fig. 3. Evidence clustering process with different thresholds.



(a) $0 \leq T_d \leq 0.019$ (b) $0.019 < T_d \leq 0.023$ (c) $0.023 < T_d \leq 0.075$ (d) $0.075 < T_d \leq 0.402$ (e) $0.402 < T_d \leq 1$

Fig. 4. Visualization of clustering results with different thresholds (Using principal component analysis (PCA) to reduce dimensionality for each BBA).

TABLE II
FUSION RESULTS OF DIFFERENT METHODS IN EXAMPLE 1

Methods	σ_1	σ_2	σ_3	$\sigma_1 \cup \sigma_2$
DS rule [7]	0.622	0.378	0.0	0.0
PCR5 rule [13]	0.566	0.434	0.0	0.0
Murphy's rule [37]	0.767	0.227	0.0	0.006
Xiao [16]	0.859	0.136	0.0	0.005
Tang et al. [38]	0.731	0.264	0.0	0.005
Gao et al. [39]	0.796	0.199	0.0	0.005
Dong et al. [40]	0.819	0.177	0.0	0.004
ER-DHC	**0.973**	0.027	0.0	0.0

TABLE III
FUSION RESULTS UNDER DIFFERENT CONFLICT METRICS IN EXAMPLE 1

Metrics	σ_1	σ_2	σ_3	$\sigma_1 \cup \sigma_2$
Jousselme's distance [18]	0.967	0.033	0.0	0.0
Belief interval distance [8]	0.968	0.032	0.0	0.0
BJS divergence [14]	0.964	0.033	0.0	0.0
$\mathcal{RB}$ divergence [16]	0.843	0.157	0.0	0.0
Belief cosine similarity [3]	**0.973**	0.027	0.0	0.0

1) Cluster Structure: The corresponding clustering process and results are shown in Figs. 3 and 4, which can be summarized into the following three cases:

- ***Case 1: Smaller threshold.*** Evidence clustering becomes relatively fragmented, even BBAs with high similarity cannot be grouped together, as shown in Figs. 3(b)-(c) and Figs. 4(b)-(c). In extreme cases, the recursive splitting may continue until each BBA forms an individual cluster, thus losing clustering significance, as shown in Fig. 3(a) and Fig. 4(a).
- ***Case 2: Larger threshold.*** A larger T_d may increase intra-cluster conflict, but does not result in higher intra-cluster conflict. As shown in Fig. 3(e) and Fig. 4(e), when $T_d = 1$, m_3 and m_6 are clustered together, with a maximum intra-cluster conflict of 0.402. This is attributed to coarse-grained division ensures that evidence with different decision preferences is grouped into separate clusters. As a result, high intra-cluster conflicts are avoided inherently, rather than relying solely on the T_d.
- ***Case 3: Adaptive selection threshold.*** Using the inflection point detection method, the threshold is set as $T_d = 0.402$. As shown in Fig. 3(d) and Fig. 4(d), m_1, m_2, m_4, and m_5 are clustered together with good compactness, while m_3 and m_6 remain in separate clusters. The maximum intra-cluster conflict is 0.075, resulting in an optimized intra-cluster conflict compared to larger thresholds. Therefore, the clustering result in this case can be considered as the optimal one. This indicates that the proposed adaptive threshold selection method contributes to balance clustering compactness and intra-cluster conflict management.

2) Sensitivity Analysis: As T_d increases from 0 to 1, the number of clusters shows a decreasing trend (from 6 to 2), exhibiting a stable monotonic trend. A smaller threshold imposes stricter conflict constraints, making the clustering more sensitive to threshold variations. However, such thresholds may lead to overly fragmented clusters, diminishing the clustering effect and is therefore not recommended. In contrast, a larger threshold imposes looser constraints on intra-cluster conflict and exhibits better robustness. From Figs. 3(d) and (e), when $0.075 < T_d \leq 1$, the cluster structure remains relatively stable and changes only once, specifically involving m_3 and m_6, while m_1, m_2, m_4, and m_5 consistently belong to the same cluster, indicating the proposed evidence clustering strategy maintains good robustness against threshold fluctuations.

C. The Role of Coarse-grained Division

Example 3: Suppose the FoD is $\Phi = \{\sigma_1, \sigma_2, \sigma_3\}$, six BBAs m_1, m_2, m_3, m_4, m_5, and m_6 are given in Table IV, where m_5 has obvious conflicts with m_3, m_4, and m_6.

In evidence clustering, coarse-grained division may split evidence that is overall similar but differs slightly in its maximum value (e.g., m_1 and m_2 in Example 3). As shown in Table V, when $T_d = 0.55$ (obtained via inflection point detection), m_1 and m_2 are grouped into the same cluster under single-stage fine-grained division (i.e., without coarse-grained division), whereas they assigned to different clusters

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2025.3601509

TABLE IV
BBAs IN EXAMPLE 3

BBA	σ_1	σ_2	σ_3	Φ
m_1	0.32	0.28	0.10	0.30
m_2	0.29	0.31	0.10	0.30
m_3	0.45	0.15	0.16	0.24
m_4	0.60	0.09	0.02	0.29
m_5	0.02	0.90	0.08	0.0
m_6	0.65	0.06	0.01	0.28

when coarse-grained division is applied. From a similarity perspective, this may slightly reduce clustering compactness. However, from a decision-making perspective, such separation is reasonable, as it ensures that all BBAs within a cluster share consistent decision preferences, which facilitates intra-cluster fusion. Since clustering in ER-DHC serves as a preparatory step for fusion, differences in clusters structure are acceptable as long as the impact on fusion results remains limited. Notably, evidence like m_1 and m_2 often reflects relatively high uncertainty (e.g., similar support for σ_1 and σ_2), and thus contributes relatively little to the final fusion. Therefore, assigning them to different clusters has limited impact on the outcome. As shown in Table V, even under different clustering strategies, the fusion results are nearly identical when $T_d = 0.55$, yielding the same final decision (i.e., σ_1).

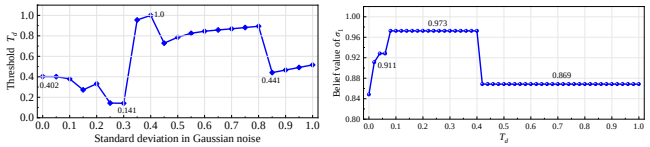
In ER-DHC, coarse-grained division brings two main advantages: *1)* Conflict control. By separating BBAs with different decision preferences in advance, it effectively suppresses high intra-cluster conflicts. This reduces the dependence on the precise value of threshold T_d and enhances the robustness of clustering. For instance, in noisy scenarios, unreasonable thresholds (e.g., $T_d = 0.98$) may arise from inflection point detection. As shown in Table V, single-stage fine-grained division under this threshold groups all six BBAs into one cluster, leading to high intra-cluster conflict of 0.95 and an unreasonable fusion result and decision (i.e., σ_2). In contrast, when coarse-grained division is applied, conflicting BBA m_5 is isolated from m_3, m_4, and m_6, reducing the maximum intra-cluster conflict to 0.55, thus achieving more appropriate cluster formation and yielding reasonable a reasonable fusion result and decision (i.e., σ_1). *2)* Computational efficiency. Applying fine-grained division directly to the initial cluster $\mathcal{C}_1 = \{m_1, m_2, \ldots, m_S\}$ requires repeated conflict computation and k-means operations, incurring a high complexity of $\mathcal{O}(S^2 \times N^2)$. Coarse-grained division reduces this initial cost to $\mathcal{O}(S \times N)$, significantly improving efficiency. As shown in Table V, coarse-grained division results in considerable time savings for both $T_d = 0.55$ and $T_d = 0.98$.

In summary, although coarse-grained division may occasionally reduce similarity-based clustering quality, it plays a crucial role in controlling intra-cluster conflict and improving computational efficiency, thereby enhancing the overall robustness and practicality of the ER-DHC.

D. Robustness of Inflection Point Detection

In ER-DHC, the inflection point detection method is employed to determine the threshold T_d. This approach is generally effective and robust to minor noise, as it relies on the overall trend of the sorted conflict curve rather than on individual fluctuations. However, as illustrated in Fig. 5(a), it remains inherently sensitive to high-intensity noise or outliers, which may shift the position of the inflection point and result in imperfect threshold selection.

Regardless of noise intensity, its influence on threshold selection can manifest in two directions: yielding a threshold lower or higher than the ideal value. Both cases may degrade clustering quality, thereby affecting the final fusion performance. Nevertheless, the coarse-grained division effectively suppresses high intra-cluster conflicts, mitigating the negative impact of imperfect threshold and helps maintain clustering quality to some extent. As shown in Fig. 5(b), the fusion performance (using the six BBAs from Table I) declines under extreme thresholds (e.g., $T_d = 0.0$ or $T_d = 1.0$) but remains superior to that of the classical DS rule and several existing methods across all tested thresholds, as summarized in Table II. This robustness stems from two characteristics of ER-DHC: *1)* When T_d is large, the mechanism of coarse-grained division avoids direct fusion of high conflict evidence. *2)* When T_d is small, cluster structure becomes relatively fragmented. In extreme cases, each BBA may form its own cluster. According to the properties of ER-DHC, it will degrade into the weighted fusion as described in [14], [21], which still outperforms classical DS rule and several existing methods.



(a) Impact on threshold T_d (b) Variation of fusion results

Fig. 5. Effect of noise and outliers on threshold detection and fusion performance under different thresholds. Based on the set $\mathcal{D}$ in Example 1, two outliers (0 and 1) are first added, followed by the introduction of Gaussian noise with varying standard deviations. In (a), the inflection point detection method is applied to the noisy conflict set to determine T_d, illustrating how increasing noise intensity affects its stability. In (b), the fusion results under different thresholds are presented, reflecting the robustness of ER-DHC.

In summary, although inflection point detection is inherently sensitive to strong noise and outliers, ER-DHC maintains superior performance even under non-ideal threshold settings.

V. Application in Multi-Source Information Fusion

In this section, the ER-DHC is applied to a multi-source information fusion (MSIF) system to validate its effectiveness. It is assumed that each information source contributes a single BBA, and the total number of sources, denoted by S, typically falls within a moderate range (e.g., 4 to 10). This setting is motivated by two practical considerations. First, when S is too small (e.g., fewer than 4), evidence clustering becomes less meaningful. Second, when S becomes excessively large, the computational cost of ER-DHC may become prohibitive, potentially limiting its practical application. Hence, ER-DHC is generally suitable for medium-scale fusion scenarios.

A. ER-DHC for MSIF

In the MSIF systems, consider a scenario with n target classes, and the target data is collected from

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2025.3601509

TABLE V
CLUSTERING AND FUSION RESULTS IN EXAMPLE 3

Coarse-grained division	T_d	Clustering result	Intra-cluster conflict (max)	Clustering time (ms)	Fusion result				Decision
					σ_1	σ_2	σ_3	Φ	
×	0.55	$\mathcal{C}_1=\{m_1,m_2,m_3,m_4,m_6\}$, $\mathcal{C}_2=\{m_5\}$	0.313	14.67±2.474	0.963	0.036	0.001	0.0	σ_1
✓	0.55	$\mathcal{C}_1=\{m_1,m_3,m_4,m_6\}$, $\mathcal{C}_2=\{m_2\}$, $\mathcal{C}_3=\{m_5\}$	0.256	10.58±1.749	0.962	0.035	0.002	0.001	σ_1
×	0.98	$\mathcal{C}_1=\{m_1,m_2,m_3,m_4,m_5,m_6\}$	0.95	3.518±0.523	0.202	0.776	0.022	0.0	σ_2
✓	0.98	$\mathcal{C}_1=\{m_1,m_3,m_4,m_6\}$, $\mathcal{C}_2=\{m_2,m_5\}$	0.55	2.471±0.495	0.811	0.186	0.003	0.0	σ_1

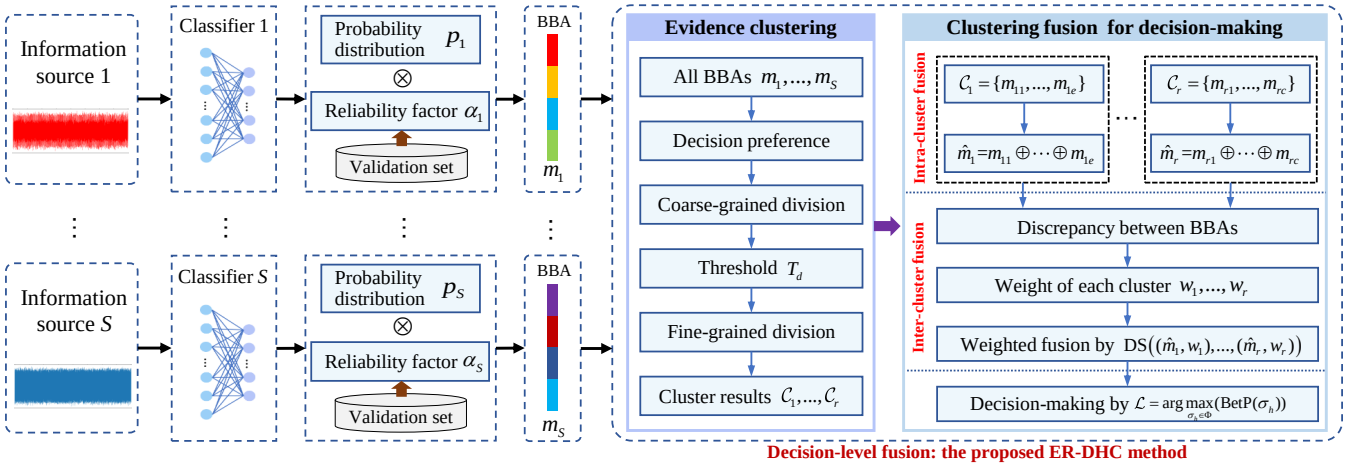


Fig. 6. Overview of the ER-DHC method for the MSIF system. For target data from each information source, a classifier is employed to generate a probability distribution, which is then transformed into a BBA. Subsequently, S BBAs are combined using the ER-DHC method for decision-making.

S information sources (e.g., S sensors). The FoD is $\Phi=\{\sigma_1,\ldots,\sigma_h,\ldots,\sigma_n\}$, where σ_h denotes the h-th class. A separate classifier is trained for each information source. For a given target data to be identified, the S trained classifiers generate S probability distributions, denoted by $p_1,...,p_S$.

1) Transforming Probability Distributions into BBAs: According to the maximum value of p_i, the target class (denoted by E_i) supported by i-th classifier is acquired and defined as

$$E_i=\arg\max_{\sigma_h\in\Phi}(p_i(\sigma_h)). \tag{42}$$

Since different classifiers may exhibit distinct performances, leading to p_i with varying degrees of reliability. To evaluate the reliability of p_i, a validation set is partitioned from the sample data. Then, the precision of i-th classifier with respect to each class of target is obtained using the validation set, and a reliability factor α_i for p_i is defined as

$$\alpha_i=\frac{TP_{E_i}}{TP_{E_i}+FP_{E_i}} \tag{43}$$

where TP_{E_i} and FP_{E_i} denote the number of samples correctly and incorrectly classified as class E_i by i-th classifier on the validation set, respectively. The reliability factor α_i finely considers the precision of each classifier on specific class E_i, not just the overall accuracy. Then, p_i is transformed into a BBA m_i using α_i, defined as

$$\begin{cases} m_i(\sigma_h)=\alpha_i\cdot p_i(\sigma_h) \\ m_i(X)=1-\alpha_i \end{cases} \tag{44}$$

with

$$X=E_1\cup E_2\cup\cdots\cup E_S. \tag{45}$$

The focal element X represents the possible decisions of S classifiers. We transform the probability value of class σ_h to X based on the reliability factor to alleviate the impact of differences in classifier performance.

2) BBAs Fusion for Decision-making: The ER-DHC is employed to combine S BBAs, and the class $\mathcal{L}$ of target is determined by the maximum pignistic probability as

$$\mathcal{L}=\arg\max_{\sigma_h\in\Phi}(BetP_{fusion}(\sigma_h)). \tag{46}$$

For convenience, an overview of ER-DHC for the MSIF system is presented in Fig. 6.

B. Experimental Setup

1) Datasets: The following 8 datasets from UCI repository are used for conducting experiments, as shown in Table VI.

TABLE VI
DATASETS USED IN EXPERIMENTS

Datasets	Classes	Attributes	S	Instances
Covertype (CovE)	6	54	9	581012
Magic Telescope (MaT)	2	10	5	19020
Dry Bean (DryB)	7	16	8	13611
Cardiotocography (CarG)	10	21	7	2126
Room Estimation (RoomE)	4	18	5	10129
Eye State (EyeS)	2	14	7	14980
Statlog	6	36	6	6435
Default of Credit (DocD)	2	22	11	30000

2) Constructing Information Sources: We adopt the methods outlined in literature [6], [17], [41] to construct multiple information sources. For sensor datasets, such as RoomE dataset containing data from 5 sensors, each sensor is treated

TABLE VII
ACCURACY OF DIFFERENT FUSION ALGORITHMS (IN %) (THE BEST PERFORMANCE IS REPRESENTED IN BOLD)

Classifier	Datasets	DS	MV	AF	PCR5	BJSD-DE	MGDD	CEMF	MGBS	**ER-DHC**
MLP	CovE	48.77±0.91	48.57±0.25	50.14±0.95	54.93±1.65	48.34±0.28	48.97±0.73	54.19±1.30	48.13±0.66	**56.68±0.28**
	MaT	71.66±1.10	73.82±1.18	74.34±1.13	74.24±1.05	**75.87±1.20**	74.45±1.23	74.76±1.36	65.19±0.32	73.61±1.29
	DryB	86.75±0.97	89.31±0.88	89.70±0.81	**89.92±0.8**8	89.42±0.76	89.46±0.96	89.68±0.87	88.70±0.27	89.24±0.81
	CarG	66.35±1.63	63.73±1.74	69.85±1.65	67.17±1.16	67.54±1.23	69.28±1.17	69.66±1.44	65.97±0.63	**71.84±1.12**
	RoomE	96.73±0.22	95.25±0.46	96.32±0.32	97.69±0.33	97.21±0.29	97.35±0.22	92.24±0.24	97.93±0.40	**98.66±0.15**
	EyeS	64.30±3.04	61.87±3.13	65.02±3.42	66.54±3.64	63.59±3.46	63.49±3.33	64.92±3.76	55.39±0.24	**67.83±3.32**
	Statlog	85.94±0.57	86.52±0.27	86.75±0.38	85.66±0.34	85.61±0.35	87.69±0.45	86.72±0.37	87.87±0.07	**88.62±0.59**
	DocD	78.07±1.36	77.91±1.20	78.07±1.07	78.18±1.15	78.17±1.08	78.09±1.27	78.07±1.06	77.89±1.41	**78.94±1.22**
SVM	CovE	48.04±0.61	47.94±0.12	52.52±0.71	51.43±1.24	47.94±0.12	48.05±0.11	49.60±1.06	38.19±0.56	**53.94±0.12**
	MaT	72.82±1.22	71.98±1.07	71.92±1.10	72.24±0.61	71.13±1.05	72.71±1.25	72.24±1.45	72.67±0.13	**73.08±1.00**
	DryB	85.57±0.61	86.35±1.17	88.76±0.50	88.32±0.58	87.28±0.98	88.69±0.81	88.65±0.80	88.41±0.23	**89.02±1.08**
	CarG	55.27±1.55	52.54±0.74	61.99±0.91	60.91±0.98	55.93±1.11	58.89±1.46	60.24±1.35	56.26±0.44	**62.97±1.77**
	RoomE	94.25±0.34	93.78±0.40	95.27±0.27	92.64±0.33	95.59±0.38	95.28±0.30	94.96±0.31	85.03±0.24	**97.04±0.29**
	EyeS	61.16±1.60	61.49±1.50	62.78±1.72	61.07±3.11	62.44±1.43	62.73±1.24	63.11±0.79	55.48±0.57	**64.97±1.97**
	Statlog	82.01±0.52	82.32±0.61	83.78±0.42	86.00±0.41	**87.30±0.48**	87.06±0.45	84.92±0.53	83.78±0.14	86.93±0.53
	DocD	74.47±0.13	77.60±0.16	75.47±0.24	77.60±0.16	76.38±0.27	75.13±0.35	74.69±0.18	71.04±0.27	**77.69±0.36**
DT	CovE	46.58±0.35	46.92±0.31	47.84±0.42	**48.45±0.55**	45.76±0.38	46.66±0.40	46.53±0.45	46.41±0.37	47.99±0.34
	MaT	68.08±2.19	70.29±1.73	70.98±2.10	71.19±1.39	70.56±1.67	71.56±1.67	70.71±1.91	64.98±0.13	**72.34±1.78**
	DryB	87.73±1.07	87.76±0.88	86.84±1.06	86.07±0.94	87.28±0.96	87.84±1.00	86.73±1.15	85.84±0.68	**88.47±0.81**
	CarG	62.79±2.03	60.91±1.58	68.01±1.88	63.31±0.37	64.77±1.52	66.74±1.48	66.98±1.12	64.35±0.40	**70.51±1.35**
	RoomE	97.91±0.22	97.29±0.09	97.95±0.06	96.70±0.08	97.80±0.15	**98.76±0.25**	97.84±0.15	94.17±0.18	98.06±0.16
	EyeS	55.20±0.17	56.43±1.54	55.20±0.17	55.48±0.57	55.58±0.53	55.58±0.53	55.20±0.17	55.10±0.06	**57.05±1.19**
	Statlog	86.34±0.45	85.14±0.53	84.59±0.55	84.62±0.77	88.32±0.65	88.60±0.72	87.56±0.70	84.01±0.25	**89.08±0.51**
	DocD	71.95±0.31	71.59±0.45	72.47±0.38	74.55±0.53	75.71±0.38	75.84±0.49	77.40±0.16	**78.29±0.67**	77.88±0.48

as an independent information source. For non-sensor datasets, the attributes of target are divided into S subsets in order, each subset is then treated as an individual information source. For example, the CovE dataset contains 54 attributes, which are divided into 9 subsets, each consisting of 6 attributes and representing a information source. The number of information sources (i.e. S) is also presented in Table VI.

3) Base Classifier for Generating BBAs: For each information source, a base classifier is trained for generating BBAs. According to the literature [6], [17], classical machine learning algorithms such as multi-layer perceptron (MLP), support vector machine (SVM), and decision tree (DT) are used as base classifiers. For fair comparison, all classifiers are trained with default parameters from the sklearn library. For each dataset, 10% of the samples are selected as the validation set, and 5-fold cross-validation is applied to the remaining samples.

C. Comparison with Existing Fusion Algorithms

To demonstrate the superiority of the ER-DHC, we compared it with 8 existing fusion algorithms, including DS rule [7], Majority Voting (MV) [42], Average Fusion (AF) [37], PCR5 rule [13], Belief Jensen–Shannon Divergence and Deng Entropy (BJSD-DE) for fusion [14], Multi-Granularity Distance for Decision (MGDD) [21], Cross Entropy of Mass Function (CEMF) [39], and Evidence Combination with Multi-granularity Belief Structure (MGBS) [41].

In terms of classification accuracy, as shown in Table VII, the proposed ER-DHC method outperforms other fusion algorithms in most cases. It consistently achieves higher accuracy than the DS rule, MV, and AF methods across all eight datasets. This performance gain is attributed to the incorporation of a conflict measure between BBAs and the subsequent clustering-based fusion strategy. In contrast, the DS rule, MV, and AF methods directly combine BBAs without evaluating their reliability or applying any modification prior to fusion. The MGDD and CEMF methods enhance the classical DS rule by introducing belief distance and belief cross-entropy to assess BBAs, respectively, resulting in improved accuracy in many cases. The BJSD-DE method further advances this idea by integrating both uncertainty and discrepancy measures to determine BBA weights, and it generally outperforms MGDD and CEMF. Compared to MGDD, BJSD-DE, CEMF, and MGBS, the ER-DHC method not only considers conflict between BBAs, but also exploits the clustering characteristics of BBAs to adaptively adjust their fusion weights. This two-stage approach effectively suppresses the influence of conflicting BBAs and contributes to superior classification performance.

In terms of fusion time, as shown in Table VIII, the MV and AF methods are the most time-efficient across all eight datasets due to their inherent simplicity, with a computational complexity of $\mathcal{O}(S \times N)$. The BJSD-DE, MGDD, and CEMF methods require additional time to assess the reliability of BBAs, resulting in a higher complexity of $\mathcal{O}(S^2 \times N^2)$ compared to DS and PCR5 rules, which remain at $\mathcal{O}(S \times N^2)$. MGBS offers a relatively better trade-off between accuracy and efficiency by approximating the original BBAs and reducing the number of focal elements to $\mathcal{N}$ ($\mathcal{N} < N$), yielding a complexity of $\mathcal{O}(S \times \mathcal{N}^2)$. In contrast, ER-DHC introduces a clustering-based fusion strategy that significantly improves classification performance, albeit with a higher computational cost of $\mathcal{O}(S^2 \times N^2 \times \log S)$. This cost primarily arises from recursive clustering and pairwise conflict assessments, which ensure robust handling of conflicting evidence.

In terms of scalability, although ER-DHC exhibits superior classification performance, its scalability remains a critical concern, particularly in large-scale or resource-constrained environments. To improve its practicality in such scenarios, several optimization strategies can be adopted. For example, in systems with homogeneous sources (e.g., identical sensors), preliminary data-level fusion can be applied to reduce the

TABLE VIII
TIME COST OF DIFFERENT FUSION ALGORITHMS (IN MS) (THE BEST PERFORMANCE IS REPRESENTED IN BOLD)

Classifier	Datasets	DS	MV	AF	PCR5	BJSD-DE	MGDD	CEMF	MGBS	**ER-DHC**
MLP	CovE	1.366±0.230	0.048±0.015	**0.021±0.009**	1.427±0.228	3.103±0.341	4.294±0.378	4.480±0.391	0.446±0.147	14.87±1.360
	MaT	0.174±0.069	0.036±0.012	**0.025±0.007**	0.104±0.058	0.397±0.097	0.500±0.110	0.572±0.114	0.220±0.078	3.618±0.438
	DryB	1.656±0.261	0.034±0.024	**0.022±0.012**	1.931±0.619	2.963±0.336	4.455±0.565	4.159±0.386	0.366±0.123	13.92±1.532
	CarG	4.373±0.315	0.051±0.013	**0.040±0.018**	4.275±0.311	6.112±0.388	8.760±0.454	7.591±0.437	0.557±0.126	18.08±1.605
	RoomE	0.385±0.215	0.027±0.007	**0.015±0.007**	0.366±0.208	0.856±0.314	1.224±0.368	1.070±0.347	0.202±0.157	4.842±0.695
	EyeS	0.166±0.129	0.029±0.010	**0.016±0.003**	0.087±0.097	0.499±0.221	0.740±0.281	0.859±0.295	0.186±0.138	3.148±0.688
	Statlog	0.999±0.340	0.033±0.012	**0.019±0.008**	1.038±0.346	1.785±0.447	2.518±0.517	2.690±0.532	0.344±0.204	8.781±1.116
	DocD	0.331±0.202	0.062±0.019	**0.021±0.006**	0.191±0.154	1.308±0.392	1.816±0.452	2.123±0.485	0.369±0.213	5.287±0.830
SVM	CovE	1.348±0.190	0.041±0.007	**0.016±0.013**	1.350±0.185	2.722±0.256	4.014±0.290	4.077±0.298	0.389±0.101	14.10±1.256
	MaT	0.143±0.062	0.034±0.003	**0.012±0.007**	0.087±0.050	0.355±0.103	0.500±0.118	0.701±0.144	0.163±0.070	3.783±0.458
	DryB	1.561±0.281	0.045±0.013	**0.044±0.068**	1.588±0.277	2.905±0.373	4.314±0.441	4.267±0.446	0.377±0.152	12.98±1.377
	CarG	4.028±0.293	0.052±0.023	**0.041±0.021**	3.881±0.277	5.421±0.324	7.346±0.304	6.481±0.317	0.512±0.123	18.34±1.617
	RoomE	0.413±0.103	0.035±0.011	**0.011±0.007**	0.384±0.101	0.901±0.143	1.150±0.153	1.307±0.162	0.176±0.068	4.598±0.343
	EyeS	0.167±0.145	0.010±0.004	**0.002±0.004**	0.060±0.070	0.539±0.256	0.953±0.344	0.792±0.302	0.201±0.160	3.091±0.692
	Statlog	1.068±0.357	0.049±0.009	**0.017±0.004**	1.101±0.363	1.810±0.445	2.545±0.583	2.516±0.515	0.337±0.200	9.454±1.170
	DocD	0.221±0.097	0.109±0.011	**0.005±0.007**	0.113±0.056	1.062±0.240	1.474±0.266	1.735±0.304	0.250±0.107	5.143±0.677
DT	CovE	1.359±0.151	0.054±0.014	**0.014±0.002**	1.468±0.164	2.999±0.237	4.390±0.275	4.323±0.271	0.447±0.106	15.82±1.344
	MaT	0.150±0.063	0.034±0.008	**0.014±0.004**	0.097±0.049	0.392±0.093	0.490±0.101	0.630±0.120	0.176±0.069	3.826±0.432
	DryB	1.636±0.418	0.036±0.016	**0.015±0.003**	1.799±0.439	3.191±0.562	4.225±0.622	4.250±0.622	0.388±0.214	12.56±1.379
	CarG	4.875±0.329	0.057±0.012	**0.043±0.017**	4.527±0.310	6.709±0.401	9.065±0.444	8.452±0.449	0.638±0.132	18.75±1.769
	RoomE	0.433±0.109	0.045±0.027	**0.008±0.003**	0.397±0.103	0.998±0.150	1.221±0.155	1.454±0.167	0.191±0.073	5.279±0.404
	EyeS	0.153±0.137	0.035±0.008	**0.023±0.007**	0.103±0.118	0.588±0.269	0.598±0.265	0.932±0.335	0.200±0.160	3.405±0.724
	Statlog	1.001±0.342	0.035±0.005	**0.023±0.012**	1.020±0.344	1.869±0.457	2.598±0.528	2.659±0.532	0.360±0.211	8.728±1.173
	DocD	0.307±0.078	0.059±0.015	**0.011±0.004**	0.188±0.062	1.123±0.137	1.668±0.176	1.905±0.182	0.332±0.080	5.672±0.413

TABLE IX
ABLATION EXPERIMENT FOR THE ER-DHC

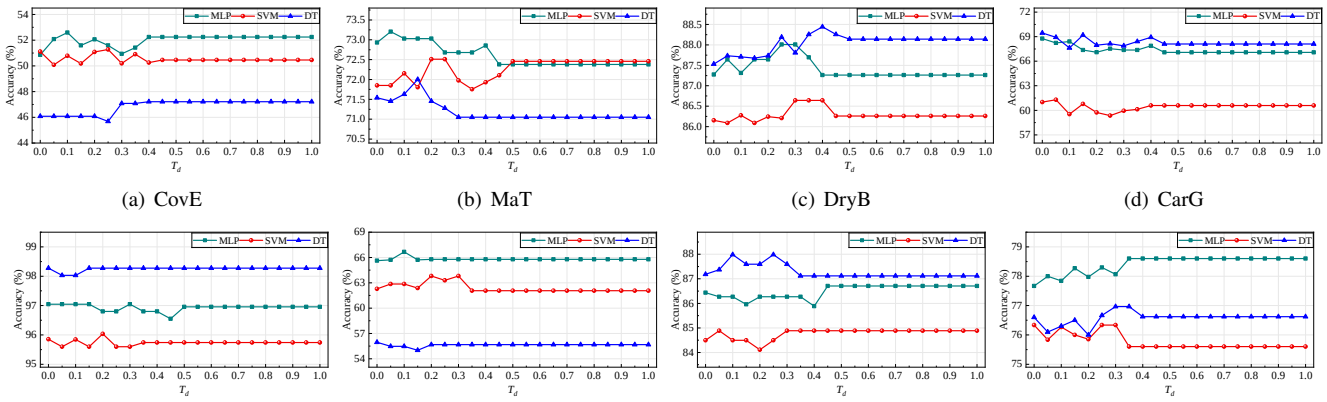
Dataset	Evidence clustering		Clustering fusion		MLP as base classifier		SVM as base classifier		DT as base classifier	
	Coarse-grained division	Fine-grained division	Intra-cluster fusion	Inter-cluster fusion	Accuracy (%)	Time of fusion (ms)	Accuracy (%)	Time of fusion (ms)	Accuracy (%)	Time of fusion (ms)
CovE	✓	×	✓	✓	52.24±0.62	6.625±0.537	50.46±0.42	6.164±0.451	47.21±0.06	6.268±0.485
	×	✓	✓	✓	52.41±0.22	19.57±1.482	51.38±0.30	19.25±1.270	47.09±0.12	18.48±1.461
	✓	✓	✓	×	48.77±0.91	10.35±0.952	48.04±0.61	10.18±1.103	46.58±0.35	10.95±0.899
	✓	✓	✓	✓	56.68±0.28	14.87±1.360	53.94±0.12	14.10±1.256	47.99±0.34	15.82±1.344
MaT	✓	×	✓	✓	72.38±0.27	0.718±0.296	72.46±0.34	0.638±0.273	71.05±0.16	0.683±0.284
	×	✓	✓	✓	72.11±0.25	3.851±0.696	72.32±0.41	3.622±0.670	70.29±0.25	3.427±0.642
	✓	✓	✓	×	71.66±1.10	2.364±0.575	72.82±1.22	2.685±0.680	68.08±2.19	2.249±0.538
	✓	✓	✓	✓	73.61±1.29	3.618±0.438	73.08±1.00	3.783±0.458	72.34±1.78	3.826±0.432
DryB	✓	×	✓	✓	87.26±1.38	7.282±0.900	86.26±1.57	7.577±0.866	88.14±1.97	7.769±0.963
	×	✓	✓	✓	88.50±1.50	18.21±1.576	88.12±1.28	18.53±1.597	88.33±1.84	18.28±1.625
	✓	✓	✓	×	86.75±0.97	10.53±1.112	85.57±0.61	10.68±1.395	87.73±1.07	10.03±1.187
	✓	✓	✓	✓	89.24±0.81	13.92±1.532	89.02±1.08	12.98±1.377	88.47±0.81	12.56±1.379
CarG	✓	×	✓	✓	67.08±1.04	9.472±0.988	60.58±1.69	8.447±0.812	68.10±0.78	9.046±0.974
	×	✓	✓	✓	66.38±1.41	21.08±2.129	60.48±1.50	22.55±2.174	67.01±0.82	22.90±2.334
	✓	✓	✓	×	66.35±1.63	15.52±0.255	55.27±1.55	14.87±0.669	62.79±2.03	15.42±1.834
	✓	✓	✓	✓	71.84±1.12	18.08±1.605	62.97±1.77	18.34±1.617	70.51±1.35	18.75±1.769
RoomE	✓	×	✓	✓	96.95±0.43	2.646±0.514	95.74±0.67	2.631±0.527	98.27±0.21	2.637±0.459
	×	✓	✓	✓	97.76±0.37	6.690±0.861	95.45±0.52	6.718±0.869	98.42±0.46	6.235±0.762
	✓	✓	✓	×	96.73±0.22	3.284±0.547	94.25±0.34	3.256±0.318	97.91±0.22	3.239±0.243
	✓	✓	✓	✓	98.66±0.15	4.842±0.695	97.04±0.29	4.598±0.343	98.06±0.16	5.279±0.404
EyeS	✓	×	✓	✓	65.78±3.54	0.928±0.328	62.06±1.91	1.101±0.350	55.67±0.52	0.909±0.323
	×	✓	✓	✓	66.21±2.38	4.572±0.779	61.97±1.59	5.042±0.782	56.24±1.27	4.986±0.794
	✓	✓	✓	×	64.30±3.04	2.251±0.218	61.16±1.60	2.873±0.186	55.20±0.17	2.641±0.199
	✓	✓	✓	✓	67.83±3.32	3.148±0.688	64.97±1.97	3.091±0.692	57.05±1.19	3.405±0.724
Statlog	✓	×	✓	✓	86.71±1.00	3.493±0.694	84.88±0.65	3.526±0.709	87.12±0.74	3.663±0.710
	×	✓	✓	✓	86.45±1.18	11.92±1.060	85.19±0.75	12.37±1.080	87.17±0.46	11.65±1.075
	✓	✓	✓	×	85.94±0.57	6.592±1.208	82.01±0.52	6.423±1.736	86.34±0.45	6.284±1.591
	✓	✓	✓	✓	88.62±0.59	8.781±1.116	86.93±0.53	9.454±1.170	89.08±0.51	8.728±1.173
DocD	✓	×	✓	✓	78.60±0.16	2.896±0.469	75.60±0.28	2.550±0.465	76.62±0.14	2.486±0.371
	×	✓	✓	✓	78.13±0.19	9.222±0.752	76.53±0.25	9.125±0.799	74.95±0.19	8.651±0.689
	✓	✓	✓	×	78.07±0.06	3.877±0.624	74.47±0.13	3.366±0.603	71.95±0.31	3.298±0.383
	✓	✓	✓	✓	78.94±1.22	5.287±0.830	77.60±0.16	5.143±0.677	77.88±0.48	5.672±0.413

number of BBAs before executing ER-DHC on heterogeneous information. Moreover, methods such as simplifying BBAs by reducing the number of focal elements, or introducing clustering constraints such as limiting recursion depth or enforcing a minimum cluster size, can further lower computational complexity, thus improving the applicability of ER-DHC.

D. Ablation Study

To highlight the contribution of each component in the ER-DHC method, we conduct an ablation study. As shown in Table IX, using only coarse-grained or fine-grained division in the evidence clustering stage leads to varying classification accuracy, whereas the two-stage division strategy yields better

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2025.3601509



(a) CovE (b) MaT (c) DryB (d) CarG

(e) RoomE (f) EyeS (g) Statlog (h) DocD

Fig. 7. Recognition performance of ER-DHC with different thresholds T_d.

performance. Compared to using only coarse-grained division, the two-stage approach enables more refined clustering, ensuring that the conflict between any two BBAs within a cluster remains within a reasonable range. On the other hand, compared to using only fine-grained division, the two-stage strategy incorporates decision preferences, ensuring that BBAs with significant conflicts are placed in different clusters. This avoids unreasonable clustering results that could arise from an inappropriate threshold T_d, thus supporting more rational evidence clustering. Moreover, integrating the coarse-grained division module significantly reduces the fusion time compared to using only fine-grained division, as it lowers the computational complexity of the clustering process.

In the fusion stage, omitting the proposed inter-cluster fusion strategy causes the combination process to degenerate into classical DS fusion, resulting in lower classification accuracy on eight datasets compared to the full ER-DHC method. In contrast, incorporating the proposed inter-cluster fusion strategy consistently improves classification accuracy over DS fusion across most datasets, regardless of whether MLP, SVM, or DT classifiers are employed. By assigning lower weight factors to conflicting BBAs, the inter-cluster fusion strategy mitigates their influence on the final decision, thereby enhancing overall fusion accuracy.

TABLE X
RECOGNITION ACCURACY OF ER-DHC WITH DIFFERENT THRESHOLDS

Dataset	Classifier	Min	Max	Avg	Std	Range
CovE	MLP	50.86	52.59	52.01	0.45	1.73
	SVM	47.08	48.11	47.24	0.27	1.03
	DT	45.68	47.21	46.86	**0.56**	1.53
MaT	MLP	72.38	73.20	72.60	0.28	0.82
	SVM	71.75	72.51	72.26	0.28	0.76
	DT	71.05	72.00	71.20	0.26	0.95
DryB	MLP	87.27	88.01	87.48	0.30	0.74
	SVM	86.09	86.64	86.29	0.15	0.55
	DT	87.53	88.44	88.07	0.20	0.91
CarG	MLP	67.08	68.76	67.38	0.49	1.68
	SVM	59.39	61.28	60.44	0.45	**1.89**
	DT	67.62	69.45	68.27	0.44	1.83
RoomE	MLP	96.55	97.04	96.93	0.12	0.49
	SVM	95.60	96.03	95.74	0.10	0.44
	DT	98.03	98.28	98.25	0.07	0.25
EyeS	MLP	65.62	66.67	65.81	0.20	1.05
	SVM	62.08	63.81	62.39	0.55	1.73
	DT	55.00	55.95	55.63	0.17	0.95
Statlog	MLP	85.88	86.71	86.50	0.27	0.83
	SVM	84.11	84.88	84.77	0.21	0.78
	DT	87.12	87.98	87.29	0.28	0.86
DocD	MLP	77.68	78.60	78.40	0.32	0.92
	SVM	75.60	76.33	75.78	0.28	0.73
	DT	76.01	76.97	76.58	0.22	0.96

E. Sensitivity Analysis of ER-DHC

In ER-DHC, the threshold T_d plays a central role in the evidence clustering process. To evaluate the influence of different threshold values on fusion performance, we assess the classification results across various T_d values. Since $d_C \in [0, 1]$, the range of T_d is accordingly set from 0 to 1.

A smaller T_d enforces lower conflict within clusters, but an excessively small threshold may lead to overly fragmented clustering, potentially resulting in clusters containing only a single BBA. In such cases, ER-DHC degenerates into a weighted fusion method as presented in [14], [21]. Conversely, a larger T_d may increase intra-cluster conflicts. To alleviate the impact of an inappropriate threshold, a coarse-grained division is first applied based on BBAs' decision preferences. This step ensures that BBAs with evident conflict are placed into different clusters, reducing sensitivity to the choice of T_d.

Fig. 7 illustrates the classification performance of ER-DHC under various thresholds. Notably, when T_d is relatively large (e.g., $T_d > 0.5$), the classification accuracy remains stable across all eight datasets. This stability arises because, beyond a certain threshold, clustering results no longer change due to the regulating effect of the coarse-grained division, indicating the robustness of the method. Although minor fluctuations in classification accuracy are observed for smaller threshold values (e.g., $T_d < 0.3$), the overall performance remains robust. Table X reports the maximum, minimum, and average classification accuracy, as well as the standard deviation and range across various thresholds. The maximum standard deviation is 0.56% (on the CovE dataset using DT classifier), and the largest range is 1.89% (on the CarG dataset using SVM classifier). Furthermore, on the MaT, DryB, RoomE, Statlog, and DocD datasets, the accuracy range remains below 1.0% regardless of

This article has been accepted for publication in IEEE Transactions on Fuzzy Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2025.3601509

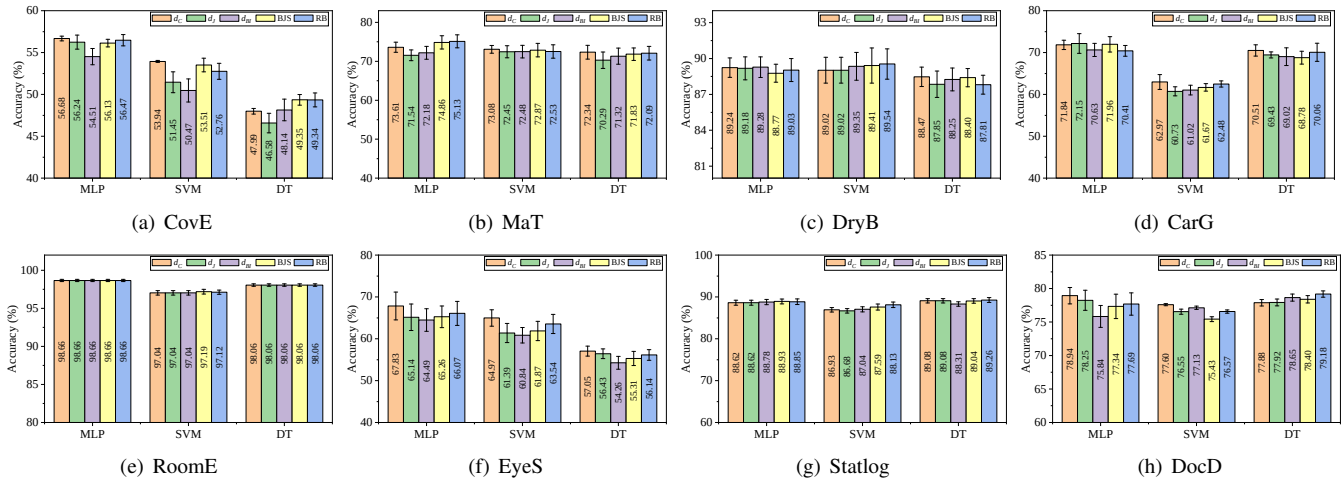


Fig. 8. Recognition performance of ER-DHC with different conflict metrics.

whether MLP, SVM, or DT is used as the base classifier.

F. Different Conflict Metrics for ER-DHC

Besides belief cosine similarity, ER-DHC can employ other conflict metrics, such as Jousselme's distance (d_J) [18], belief interval distance (d_{BI}) [8], BJS divergence [14], and $\mathcal{RB}$ divergence [16]. Since ER-DHC focuses on conflict management, belief entropy is not considered.

As shown in Fig. 8, the use of belief cosine similarity generally yields slightly better performance compared to the other four conflict metrics, which justifies its selection as the metric for ER-DHC. When the base classifiers demonstrate strong classification performance on the target dataset, the differences among the five conflict measures become negligible. This is because, for each target sample, each base classifier outputs a probability distribution p_i, which is subsequently transformed into a BBA m_i. In m_i, only one disjunctive focal element (i.e., X) is included in addition to singleton focal elements. According to (44), the value of $m_i(X)$ depends on the performance of the i-th base classifier on the validation set. For example, on the RoomE dataset, where all base classifiers achieve high classification precision (exceeding 90%), the belief values assigned to X are relatively small (typically less than 0.1). As a result, the five conflict metrics yield very similar clustering results and similar fusion weight w_h, leading to comparable classification accuracy. In contrast, on the CovE dataset, where the base classifiers perform less effectively, the belief values assigned to X tend to be larger (generally above 0.2). In such cases, the five conflict measures show varying performance, which leads to more noticeable differences in the final classification accuracy.

G. Limitations and Future Work

1) Limitations of Inflection Point-Based Threshold Selection: In ER-DHC, the conflict threshold for evidence clustering is determined via inflection point detection. While generally effective, this method may underperform when the conflict distribution exhibits weak curvature or is perturbed by high-intensity noise, potentially resulting in suboptimal clustering and fusion performance. To enhance robustness across diverse scenarios, future work will explore more adaptive threshold selection approaches. Promising directions include techniques such as the Silhouette coefficient, Elbow method [43], and Gaussian mixture model [44].

2) Impact of BBA Construction on MSIF Performance: In MSIF systems, the reliability of initial BBAs affects the overall fusion and decision-making performance. The current implementation derives BBAs from class probability distributions produced by individual classifiers. However, in real-world applications involving incomplete or noisy data, this approach may fail to capture the underlying uncertainty effectively. Future work will focus on developing more general and robust BBA construction techniques. One promising direction is to integrate Dirichlet distributions [45] into the classifier training process to explicitly model uncertainty, thereby enhancing the quality and credibility of the generated BBAs.

VI. Conclusion

In this article, we propose a novel evidential reasoning rule consisting of two modules: evidence clustering and cluster fusion. In the evidence clustering module, we introduce a two-stage divisive hierarchical clustering method. The coarse-grained division prevents high intra-cluster conflicts by considering the decision preferences of BBAs, ensuring the rationality of the cluster structure, while the fine-grained division further refines the clusters to enhance clustering quality based on conflict mutations. Subsequently, we propose a new evidence fusion algorithm, It not only degenerates into classical DS fusion and weighted fusion, but also provides a new clustering fusion perspective, exhibiting greater flexibility. Finally, we apply the ER-DHC to the MSIF system, and experimental results demonstrates that our approach contributes to enhancing classification accuracy.

References

[1] Y. L. Dong, X. D. Li, J. Dezert, R. Zhou, C. Zhu, L. Cao, M. O. Khyam, and S. S. Ge, "Multisource weighted domain adaptation with evidential reasoning for activity recognition," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 4, pp. 5530–5542, Apr 2023.

[2] W. Sheng and X. D. Li, "Multi-task learning for gait-based identity recognition and emotion recognition using attention enhanced temporal graph convolutional network," *Pattern Recognition*, vol. 114, Jun 2021.

[3] F. Xiao, Z. Cao, and A. Jolfaei, "A novel conflict measurement in decision-making and its application in fault diagnosis," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 1, pp. 186–197, Jan 2021.

[4] Z. Liu, F. Xiao, C.-T. Lin, and Z. Cao, "A robust evidential multisource data fusion approach based on cooperative game theory and its application in eeg," *IEEE Transactions on Systems Man Cybernetics-Systems*, vol. 54, no. 2, pp. 729–740, Feb 2024.

[5] X. D. Li, F. Dunkin, and J. Dezert, "Multi-source information fusion: Progress and future," *Chinese Journal of Aeronautics*, vol. 37, no. 7, pp. 24–58, Jul 2024.

[6] Z.-G. Liu, Y. Liu, J. Dezert, and F. Cuzzolin, "Evidence combination based on credal belief redistribution for pattern classification," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 4, pp. 618–631, Apr 2020.

[7] G. Shafer, "A mathematical theory of evidence," *Princeton, NJ, USA: Princeton Univ. Press*, 1976.

[8] D. Han, J. Dezert, and Y. Yang, "Belief interval-based distance measures in the theory of belief functions," *IEEE Transactions on Systems Man Cybernetics-Systems*, vol. 48, no. 6, pp. 833–850, 2018.

[9] Y. L. Dong, X. D. Li, and J. Dezert, "Evidential reasoning with hesitant fuzzy belief structures for human activity recognition," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 12, pp. 3607–3619, 2021.

[10] Z. Zhang, Y. Zhang, H. Tian, A. Martin, Z. Liu, and W. Ding, "A survey of evidential clustering: Definitions, methods, and applications," *information fusion*, vol. 115, Mar 2025.

[11] R. Yager, "On the dempster-shafer framework and new combination rules," *Information Sciences*, vol. 41, no. 2, pp. 93–137, MAR 1987.

[12] D. Dubois and H. Prade, "On the unicity of dempster rule of combination," *International Journal of Intelligent Systems*, vol. 1, no. 2, pp. 133–142, 1986.

[13] F. Smarandache and J. Dezert, "Advances and applications of dsmt for information fusion," *American Research Press*, 2004-2015.

[14] F. Xiao, "Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy," *Information Fusion*, vol. 46, pp. 23–32, Mar 2019.

[15] Z. Zhang, H. Wang, W. Jiang, and J. Geng, "A target intention recognition method based on information classification processing and information fusion," *Engineering Applications of Artificial Intelligence*, vol. 127, no. B, Jan 2024.

[16] F. Y. Xiao, "A new divergence measure for belief functions in d-s evidence theory for multisensor data fusion," *Information Sciences*, vol. 514, pp. 462–483, 2020.

[17] Y. Huang, F. Xiao, Z. Cao, and C.-T. Lin, "Higher order fractal belief rnyi divergence with its applications in pattern classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 12, pp. 14 709–14 726, Dec 2023.

[18] A.-L. Jousselme, D. Grenier, and E. Bossé, "A new distance between two bodies of evidence," *Information Fusion*, vol. 2, no. 2, pp. 91–101, 2001.

[19] C. Zhu and F. Xiao, "A belief hellinger distance for d-s evidence theory and its application in pattern recognition," *Engineering Applications of Artificial Intelligence*, vol. 106, Nov 2021.

[20] K. Zuo, X. Li, J. Dezert, and Y. Dong, "Weighted fusion of multiple classifiers for human activity recognition," in *26th International Conference on Information Fusion*, Charleston, United states, 2023.

[21] Y. Zhao, Z. Zhang, and F. Xiao, "A multi-granularity distance with its application for decision making," *Information Sciences*, vol. 661, Mar 2024.

[22] F. Li, J. Wang, Y. Qian, G. Liu, and K. Wang, "Fuzzy ensemble clustering based on self-coassociation and prototype propagation," *IEEE Transactions on Fuzzy Systems*, vol. 31, no. 10, pp. 3610–3623, 2023.

[23] L. Chen, P. D'Urso, and Y. Deng, "Occm-rps: Ordered credal c-means clustering based on random permutation set," *Information Sciences*, p. 122329, 2025.

[24] F. Li, J. Wang, L. Zhang, Y. Qian, S. Jin, T. Yan, and L. Du, "k-hyperedge medoids for clustering ensemble," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 17, 2025, pp. 18 271–18 278.

[25] F. Li, Y. Qian, J. Wang, and J. Liang, "Multigranulation information fusion: A dempster-shafer evidence theory-based clustering ensemble method," *Information Sciences*, vol. 378, pp. 389–409, 2017.

[26] Z. Liu and S. Letchmunan, "Enhanced fuzzy clustering for incomplete instance with evidence combination," *ACM Transactions on Knowledge Discovery from Data*, vol. 18, no. 3, pp. 1–20, 2024.

[27] T. Tanino, "Fuzzy preference orderings in group decision-making," *Fuzzy Sets and Systems*, vol. 12, no. 2, pp. 117–131, 1984.

[28] X. Deng, X. Lu, F. T. S. Chan, R. Sadiq, S. Mahadevan, and Y. Deng, "D numbers extended consistent fuzzy preference relations," *Knowledge-Based Systems*, vol. 73, pp. 61–68, Jan 2015.

[29] T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu, "An efficient k-means clustering algorithm:: Analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, Jul 2002.

[30] J. Shen, X. Hao, Z. Liang, and Y. Liu, "Real-time superpixel segmentation by dbscan clustering algorithm," *IEEE Transactions on Image Processing*, vol. 25, no. 12, pp. 5933–5942, Dec 2016.

[31] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, May 2002.

[32] F. Corpet, "Multiple sequence alignment with hierarchical-clustering," *Nucleic Acids Research*, vol. 16, no. 22, pp. 10 881–10 890, Nov 25 1988.

[33] P. Smets, "Decision making in the tbm: the necessity of the pignistic transformation," *International Journal of Approximate Reasoning*, vol. 38, no. 2, pp. 133–147, 2005.

[34] R. Kashef and M. S. Kamel, "Enhanced bisecting k-means clustering using intermediate cooperation," *Pattern Recognition*, vol. 42, no. 11, pp. 2557–2569, 2009.

[35] J. Qian, X. Guo, and Y. Deng, "A novel method for combining conflicting evidences based on information entropy," *Applied Intelligence*, vol. 46, no. 4, pp. 876–888, 2017.

[36] C. Zhu, B. Qin, F. Xiao, Z. Cao, and H. M. Pandey, "A fuzzy preference-based dempster-shafer evidence theory for decision fusion," *Information Sciences*, vol. 570, pp. 306–322, 2021.

[37] C. Murphy, "Combining belief functions when evidence conflicts," *Decision Support Systems*, vol. 29, no. 1, pp. 1–9, Jul 2000.

[38] Y. Tang, Y. Chen, and D. Zhou, "Measuring uncertainty in the negation evidence for multi-source information fusion," *Entropy*, vol. 24, no. 11, Nov 2022.

[39] X. Gao, L. Pan, and Y. Deng, "Cross entropy of mass function and its application in similarity measure," *Applied Intelligence*, vol. 52, no. 8, pp. 8337–8350, Jun 2022.

[40] Y. Dong, N. Jiang, R. Zhou, C. Zhu, L. Cao, T. Liu, and Y. Xu, "A novel multi-criteria conflict evidence combination method and its application to pattern recognition," *Information Fusion*, vol. 108, Aug 2024.

[41] K. Zuo, X. Li, L. Yu, T. Shen, Y. Dong, and J. Dezert, "Evidence combination with multi-granularity belief structure for pattern classification," *Information Sciences*, vol. 690, Feb 2025.

[42] T. B. Chandra, K. Verma, and B. K. Singh, "Coronavirus disease (covid-19) detection in chest x-ray images using majority voting based classifier ensemble," *Expert Systems with Applications*, vol. 165, Mar 2021.

[43] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Information Sciences*, vol. 622, pp. 178–210, Apr 2023.

[44] M.-S. Yang, C.-Y. Lai, and C.-Y. Lin, "A robust em clustering algorithm for gaussian mixture models," *Pattern Recognition*, vol. 45, no. 11, pp. 3950–3961, 2012.

[45] Z. Han, C. Zhang, H. Fu, and J. Zhou, "Trusted multi-view classification with dynamic evidential fusion." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 2, pp. 2551–2566, 2023.