Spatio-temporal credible evidence fusion for UAV type recognition in distributed urban low-altitude sensing networks

Spatio-temporal credible evidence fusion for UAV type recognition in distributed urban low-altitude sensing networks

Yihan CUI[a], Yan LIANG[a], Chaoxiong MA[a,*], Jie SHI[b], Paolo BRANDIMARTE[c] and Yangyang SUN[a]

[a]School of Automation, Northwestern Polytechnical University, 710072 Xi'an, PR China
[b]School of Power and Energy, Northwestern Polytechnical University, 710072 Xi'an, PR China
[c]Department of Mathematical Sciences, Politecnico di Torino, 10129 Turin TO, Italy

ARTICLE INFO

Keywords:
UAV recognition
Decision-level fusion
Credible evidence fusion
Integrated recognition
Spatio-temporal fusion
Evidence consistency

ABSTRACT

In continuous recognition of urban low-altitude unmanned aerial vehicles (UAVs), the widespread deployment of low-cost sensors and autonomous detection technologies has led to multi-view, high-dimensional data exhibiting strong temporal correlations. However, large-scale evidence that exhibits spatio-temporal inconsistencies caused by environmental clutter, building occlusion, and measurement errors leads to high computational complexity, underutilization of informative data, and degraded system stability. Decision-level fusion methods relying on pairwise comparison metrics impose a high computational burden and underutilize the contributions of high-quality evidence. The lack of an effective measure for temporal evidence consistency fluctuations restricts system performance in achieving efficient, accurate, and stable recognition. To address these challenges, this paper proposes a spatio-temporal credible evidence fusion method, STCEF, for distributed UAV type recognition in urban low-altitude within a collaborative spatio-temporal fusion framework. Spatially, a conditional credibility-based fusion method under a fusion-center-based strategy is designed to reduce computational complexity while improving the utilization of high-quality evidence. Temporally, sliding-window Rényi entropy quantifies uncertainty fluctuations to dynamically adjust evidence fusion weights, thereby suppressing transient interference caused by variations in class and belief assignments. Simulation results demonstrate that the proposed method enhances computational efficiency while achieving more accurate and stable type recognition, improving accuracy by 2.1% and mean belief by 10% compared with conventional methods.

1. Introduction

The proliferation of low-cost unmanned aerial vehicles (UAVs) poses significant challenges to urban low-altitude security, necessitating robust integrated recognition systems to mitigate these risks [1]. In urban environments, building occlusion, electromagnetic clutter, and multipath reflections often yield a low signal-to-noise ratio (SNR) and weakly discriminative single-frame observations of low-slow-small UAVs under long-range and resolution-limited conditions [2]. Leveraging short-term temporal stability, continuous observation provides integrated information to shift low-altitude monitoring from instantaneous detection to collaborative multi-sensor persistent recognition [3]. Specifically, heterogeneous radar and electro-optical sensors leverage complementary characteristics to enhance system recognition robustness [4].

Distributed low-cost sensor networks in urban environments effectively mitigate single-node blind spots. Such collaborative architectures also provide a feasible basis for incorporating auxiliary low-cost information sources, such as GNSS-derived positioning constraints and acoustic cues, in future integrated UAV recognition systems [5]. Constrained by communication bandwidth and local computing power, each node typically performs local feature extraction, with decision-level fusion subsequently performed at the central node [6]. However, continuous multi-node sensing generates large-scale data streams that require efficient real-time processing. Moreover, occlusion, viewpoint variations, and transient interference frequently cause non-critical nodes to produce low-belief evidence

⋆This work is supported by the National Natural Science Foundation of China under Grant No. 62573357.
*Corresponding author.
Email addresses: cuiyihan@mail.nwpu.edu.cn (Yihan CUI); liangyan@nwpu.edu.cn (Yan LIANG); chaoxiongma@mail.nwpu.edu.cn (Chaoxiong MA); shijie_2019100616@mail.nwpu.edu.cn (Jie SHI); paolo.brandimarte@polito.it (Paolo BRANDIMARTE); sunyangyang@mail.nwpu.edu.cn (Yangyang SUN)
ORCID(s): 0000-0002-3532-3913 (Yihan CUI); 0009-0007-4228-5824 (Chaoxiong MA)

or misassociated observations [7]. The influx of low-quality data reduces the utilization of informative evidence and undermines decision stability [8]. Therefore, balancing computational efficiency with information utility within collaborative frameworks remains a central challenge for persistent urban surveillance [9].

Currently, decision-level integrated UAV type recognition is primarily implemented using Bayesian statistics [10], data-driven deep learning (DL) [11], and evidential reasoning based on uncertainty theory [12]. Bayesian methods rely heavily on accurate priors and well-specified likelihood models [13]. With insufficient prior knowledge and low sample credibility, the Bayesian update process struggles to isolate conflict sources, and conflicting information is often redistributed across all classes to yield an ambiguous posterior. Moreover, frequent recursive posterior updates under large-scale multi-sensor inputs increase computational costs, thereby hindercing real-time performance. DL methods leverage extensive parameter learning for feature extraction and classification, though their performance remains constrained by training data quality. With low-credibility inputs, weakly discriminative UAV features restrict model precision and induce performance degradation [14]. Lacking explicit mechanisms to quantify conflict and uncertainty, these methods are overly sensitive to outliers and produce overconfident misclassifications, undermining system stability [15].

Compared with data- and feature-level modeling that requires precise labels and prior knowledge, decision-level recognition based on evidential reasoning quantifies the support and conflict among multi-sensor evidence, making the fusion process highly interpretable. This approach provides reliable support for urban low-altitude safety assessment [16]. Evidence theory employs basic belief assignments (BBAs) to represent data uncertainty. By measuring evidence similarity, it adjusts fusion weights to balance contributions from different credibility levels, thereby providing interpretable representations of observational discrepancies among multiple sensors. It has been widely applied in target recognition [17], classification [18], pattern recognition [19], fault diagnosis [20], and multi-criteria decision-making [21]. In continuous recognition tasks, the conflict resolution mechanism in evidential reasoning suppresses the impact of outliers, ensuring decision stability. Moreover, the framework also reduces computational costs by optimizing combination rules, which is essential for large-scale evidence fusion. Research in evidence theory aiming to enhance computational efficiency, information utilization, and fusion stability can be broadly categorized into three components: rule and combination-logic optimization, evidence uncertainty measurement, and evidence reliability assessment.

First, rule optimization and combination-logic simplification are the primary approaches to improve fusion efficiency and achieve real-time reasoning for large-scale evidence. Simplifying combination rules and averaging evidence credibility significantly reduce computational complexity [22–24]. Hierarchical strategies address the computational burden by optimizing fusion procedures, including clustering-based dimensionality reduction at the structural level [25, 26]. While these studies achieve high fusion efficiency, simplifying structures and reducing rule dimensions inevitably introduce a trade-off between efficiency and information utilization. Fine-grained evidence uncertainty measures are essential for multi-sensor evidence utilization and serve as the foundation for conflict resolution. Constructing refined belief measures and analyzing correlations can enhance fusion efficiency and the utilization of uncertain information [27–31]. However, precise modeling often increases computational complexity, making it challenging to meet real-time recognition requirements. Evidence reliability assessment is crucial for removing outliers and identifying critical fusion information, especially when addressing sensor anomalies such as interference, deception, and faults in complex urban detection environments. Quantifying evidence contributions supports the construction of reliable temporal models and enhances overall system robustness [32–34]. Although typical discounting methods suppress conflicts, averaging often diminishes the utilization of informative evidence.

While existing studies improve fusion efficiency, information utilization, and stability through rule simplification, uncertainty quantification, and reliability assessment, most static or offline mechanisms are still insufficient for handling the dynamic complexity of temporal evidence in distributed sensing. In spatial fusion, traditional evidence discounting and divergence-based weighting methods usually estimate evidence reliability from prior credibility, conflict degree, or pairwise distance. Although these methods can suppress conflicting evidence, their computational burden increases rapidly with the evidence scale, and majority-supported weighting may dilute high-quality evidence from critical sensors. In temporal fusion, sequential D-S fusion provides a recursive mechanism for continuously arriving evidence, Murphy Average Cost [22] mitigates conflict through evidence averaging, and Time Decay [33] introduces historical attenuation to improve temporal continuity. Recent methods such as DEF-NNs [35] and information entropy [36] further consider dynamic evidence uncertainty or entropy-weighted sequential fusion. However, these methods are generally not developed within a hierarchical spatio-temporal fusion framework. When applied to distributed UAV recognition, strong inconsistency may arise simultaneously from heterogeneous spatial sensing conditions and temporal belief fluctuations. Such coupled inconsistency is difficult to fully handle using pairwise discounting, averaging, fixed decay, or single-stage uncertainty modeling. Therefore, continuous UAV recognition still requires a task-oriented spatio-temporal fusion mechanism that can reduce large-scale computation, preserve informative evidence, and maintain temporally stable yet responsive decisions. Considering the multi-view, multi-temporal, and high-dimensional characteristics of evidence in continuous recognition, extracting high-quality information from redundant data while ensuring real-time, robust fusion presents several critical challenges.

1) Balancing multi-objective trade-offs within a unified framework. In distributed continuous recognition, fusion speed, information utilization, and system stability are inherently coupled. Most existing frameworks rely on static metrics or single-stage temporal processing, thereby overlooking hierarchical spatio-temporal coordination under spatial heterogeneity and temporal dynamics, and thus failing to achieve global optimization.
2) Limitations in efficiency and utilization of pairwise-comparison models. Similarity- or conflict-based fusion often incurs quadratic complexity with respect to evidence scale. Furthermore, majority-supported weighting logic tends to marginalize high-quality evidence from critical sensors or misclassify such data as outliers. This leads to effective information dilution by abundant low-credibility evidence, hindering improvements in recognition accuracy.
3) Fusion risks induced by temporal inconsistency. Unlike offline tasks, continuous recognition relies on real-time sensor streams where measurement errors cause fluctuations in belief assignments. Additionally, tracking errors driven by environmental clutter and occlusions introduce misassociated evidence. Conventional discounting and static measures are insufficiently sensitive to the magnitude and trend of temporal fluctuations, which necessitates adaptive weighting strategies to suppress decision risk while maintaining responsiveness to real class changes.

Consequently, to address the challenges of large-scale and spatio-temporally inconsistent evidence in persistent urban low-altitude monitoring, a spatio-temporal credible evidence fusion method, STCEF, is proposed for rapid, reliable, and stable integrated UAV type recognition. The novelty of STCEF is twofold. Spatially, the proposed conditional credibility mechanism uses the fusion center to update class-event probabilities and guide credibility estimation, rather than relying on empirical discounting or pairwise evidence distances. This reduces large-scale pairwise computation and helps preserve critical high-quality evidence. Temporally, sliding-window Rényi entropy measures the relative uncertainty fluctuation between current and recent historical evidence, enabling adaptive weighting between temporal stability and responsiveness.

The remainder of this paper is organized as follows: Section 2 analyzes the challenges in continuous UAV type recognition in urban low-altitude environments and discusses the necessity of a spatio-temporal collaborative mechanism. Section 3 presents the details of the STCEF. Section 4 evaluates the proposed method's effectiveness through simulations. Finally, Section 5 concludes the paper.

2. Problem formulation

In urban low-altitude UAV recognition, a single sensing node often cannot maintain reliable and continuous observation because of building occlusion, electromagnetic clutter, and multipath reflections. To alleviate sensing blind spots and extend surveillance coverage, distributed collaborative networks comprising heterogeneous radar and electro-optical (EO) sensors are typically deployed to exploit multi-view complementarity, as illustrated in Fig.1. Radar sensors primarily characterize target scattering and motion attributes, whereas EO sensors capture appearance-related information, such as texture and shape. The complementary sensing mechanisms and feature spaces of these sensors provide a physical basis for improving information utilization and recognition robustness at the decision level.

Constrained by communication bandwidth and local computing resources, distributed collaborative monitoring networks often adopt a two-tier edge-center processing workflow to minimize transmission overhead. Individual detection nodes perform local signal processing to generate evidence, which is then transmitted to a fusion center for decision-level fusion. However, in continuous monitoring, the accumulation of multi-feature evidence such as radar cross section (RCS), velocity, altitude, and range alongside high-frequency temporal sampling significantly increases evidence scale and imposes a severe computational burden. Furthermore, increased evidence volume does not guarantee superior fusion quality. Due to location and viewing-angle limitations, non-critical nodes frequently output low-belief, low-quality evidence that tends to dilute high-quality information contributions. Additionally, target pose variations and environmental occlusions cause evidence fluctuations over time, resulting in belief instability and conflict accumulation, which compromises the temporal consistency of fusion results. Therefore, while distributed

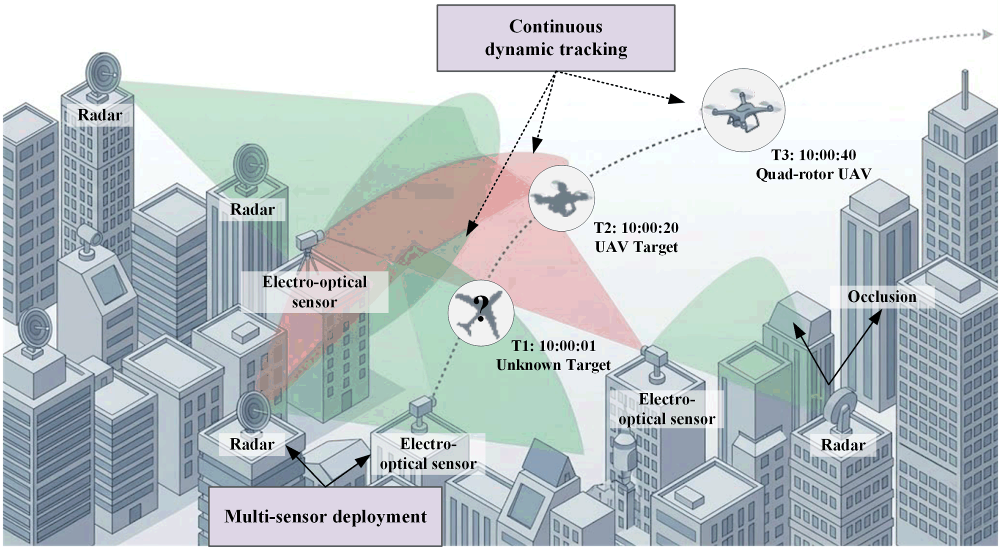


Figure 1: Continuous UAV recognition in urban environment

collaborative networks leverage heterogeneous complementarity to improve coverage, systems face the challenge of balancing computational efficiency, information utilization, and system stability.

Given massive temporal evidence, traditional fusion frameworks exhibit an inherent trade-off between fusion speed, information utilization, and system stability. This stems from metric competition across spatial and temporal dimensions, where simplifying structures improves efficiency at the expense of detail, precise modeling enhances information utilization while increasing computational cost, and conflict resolution improves stability by averaging out informative advantages. Most existing evidence fusion methods rely on static metrics and fixed rules, which often limit them to optimizing a single objective such as conflict resolution or complexity reduction. For persistent UAV recognition, these constraints induce performance limitations as the evidence scale expands and temporal inconsistency grows.

To optimize efficiency, information utilization, and recognition stability, the global fusion process is decomposed into spatial and temporal subtasks. Efficiency denotes the computational scale and resource consumption per frame, focusing on minimizing evidence comparisons and time complexity. Information utilization quantifies the capacity to retain useful information during fusion, emphasizing the prioritizing of high-quality evidence to maximize information gain. Stability denotes the temporal resistance to disturbances in continuous outputs, aiming to avoid decision risks triggered by transient anomalies.

In distributed collaborative monitoring networks, evidence exhibits distinct spatio-temporal distributions. At a single frame, spatial evidence expands in scale and dilutes information due to multi-attribute inputs, while temporal evidence across frames introduces decision risks driven by fluctuations and anomalies in continuous observation. Efficiency and information utilization are best addressed during spatial fusion, which is characterized by viewpoint variations, diverse sensing mechanisms, and unknown source priors. Meanwhile, stability is mainly managed through temporal fusion, where occlusions, pose variations, and association bias are the dominant factors. This hierarchical framework mitigates conflicts among competing fusion requirements.

Accordingly, a credible fusion framework is required to jointly balance computational efficiency, information utilization, and temporal stability under large-scale and inconsistent evidence. The proposed spatio-temporal collaborative framework decomposes the continuous recognition task into spatial and temporal dimensions and addresses their respective fusion requirements through coordinated optimization. Specifically, credible evidence fusion provides the methodological foundation, spatial evidence consistency fusion improves computational efficiency and the utilization of high-quality evidence, and temporal evidence consistency fusion enhances the stability of continuous recognition results.

2.1. Credible evidence fusion

In evidence theory, the frame of discernment (FoD) is denoted as $\Omega = \{\hat{A}_1, \hat{A}_2, \cdots, \hat{A}_n\}$, where $\hat{A}_j$ denotes the jth singleton class event. At each time, exactly one class event corresponds to the unknown ground-truth class. The power set is denoted as $2^{\Omega} = \{\emptyset, A_1, A_2, \cdots, A_J\}$, where $J = 2^n - 1$ is the number of non-empty propositions. When $|A_k| > 1$, A_k constitutes a composite class; when $|A_k| = 1$, A_k constitutes a singleton class, where $k = 1, 2, \cdots, J$.

m is the basic belief assignment (BBA) of the FoD Ω. At time t, the BBA provided by the ith evidence source is denoted as $m_i^{(t)} = \left[m_i^{(t)}(\emptyset), m_i^{(t)}(A_1), \cdots, m_i^{(t)}(A_J)\right]^T$, $m_i^{(t)} : 2^{\Omega} \to [0, 1]$, which satisfies

$$m_i^{(t)}(\emptyset) = 0, \qquad \sum_{A \subseteq \Omega} m_i^{(t)}(A) = 1 \tag{1}$$

When $m_i(A_k) > 0$, A_k is called a focal element of m_i. The subset A_k that maximizes $m_i(A_k)$ is called the principal focal element of m_i. The pignistic probability transformation redistributes the mass assigned to each composite proposition equally among its singleton elements. Let $BetP_{m_i}(\hat{A}_j)$ denote the pignistic probability of class event $\hat{A}_j$ under BBA m_i. It is defined as

$$BetP_{m_i}(\hat{A}_j) = \sum_{\substack{B \subseteq \Omega \\ \hat{A}_j \in B}} \frac{m_i(B)}{|B|} \tag{2}$$

Therefore, when the fusion result is m, the probability of event $\hat{A}_j$ is represented as

$$p(\hat{A}_j) = BetP_m(\hat{A}_j) \tag{3}$$

In UAV type recognition, targets belong to one real type, meaning that composite propositions indicate insufficient discriminability, incomplete observation, or low belief rather than mixed classes. This allows the fusion result to serve as an estimated class-consensus reference without known ground truth and support center-guided credibility optimization in spatial fusion.

In the fusion stage, the Dempster-Shafer combination rule (DCR) is typically employed to integrate multiple pieces of evidence. When there are N pieces of evidence to be fused,

$$DCR(m_1, m_2, \cdots, m_N) = m_1 \oplus m_2 \oplus \cdots \oplus m_N \tag{4}$$

Here, $\oplus$ is the DCR operator. For any $A_k \in 2^{\Omega}$, the combination of two evidence items m_i and m_ℓ is defined as

$$(m_i \oplus m_\ell)(A_k) = \begin{cases} 0, & \text{if} \quad A_k = \emptyset, \\ \dfrac{\sum\limits_{B \cap C = A_k} m_i(B) m_\ell(C)}{1 - \sum\limits_{B \cap C = \emptyset} m_i(B) m_\ell(C)}, & \text{if} \quad A_k \in 2^{\Omega} \setminus \{\emptyset\}. \end{cases} \tag{5}$$

$\sum_{B \cap C = A_k} m_i(B) m_\ell(C)$ is the conjunctive sum, and $\sum_{B \cap C = \emptyset} m_i(B) m_\ell(C)$ is the conflict measure between m_i and m_ℓ.

To address the counter-intuitive results that may arise from traditional DCR fusion, the credible evidence fusion (CEF) method introduces the following enhanced procedures within the aforementioned operational framework: construction of an Evidence Difference Measure Matrix (EDMM), evidence credibility calculation, evidence preprocessing, and credible fusion.

$$\begin{cases} Cred = f\left(EDMM(m_1, m_2, \cdots, m_N)\right) \\ \bar{m} = DCR\left(Cred; m_1, m_2, \cdots, m_N\right) \end{cases} \tag{6}$$

Here, $Cred = [Cred_1, Cred_2, \cdots, Cred_N]^T$, and $DCR(Cred; m_1, m_2, \cdots, m_N)$ denotes credibility-based evidence preprocessing followed by Dempster's rule of combination. $m_1, m_2, \cdots, m_N$ denote the evidence to be fused, $Cred_i$ denotes the credibility weight of evidence source i, and $\bar{m}$ signifies the credible fusion result. The detailed calculation is given as follows. For each class event $\hat{A}_j \in \Omega$, the event-conditional evidence $m_{\hat{A}_j}$ is constructed to represent the hypothesis that $\hat{A}_j$ is the current target class:

$$m_{\hat{A}_j}(B) = \begin{cases} 1, & B = \hat{A}_j \\ 0, & B \neq \hat{A}_j \end{cases} \quad B \in 2^{\Omega} \tag{7}$$

This maps the class event $\hat{A}_j$ into the same evidence space as m_i. Thus, the difference between evidence m_i and class event $\hat{A}_j$ can be measured by comparing m_i with $m_{\hat{A}_j}$. For N evidence items and n class events, the event-evidence evaluation matrix is constructed as

$$D_{EEM} = \begin{bmatrix} \tilde{d}_{11} & \tilde{d}_{12} & \cdots & \tilde{d}_{1N} \\ \tilde{d}_{21} & \tilde{d}_{22} & \cdots & \tilde{d}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{d}_{n1} & \tilde{d}_{n2} & \cdots & \tilde{d}_{nN} \end{bmatrix} \in \mathbb{R}^{n \times N} \tag{8}$$

where

$$\tilde{d}_{ji} = d(m_i, m_{\hat{A}_j}) \tag{9}$$

Here, $\tilde{d}_{ji}$ denotes the difference between evidence m_i and the event-conditional evidence $m_{\hat{A}_j}$. A smaller $\tilde{d}_{ji}$ indicates that m_i provides stronger support for class event $\hat{A}_j$. The support degree of evidence m_i for class event $\hat{A}_j$ is then defined as

$$Support_{ji} = e^{-\tau \tilde{d}_{ji}} \tag{10}$$

where $\tau > 0$ is the distance coefficient. The exponential function is monotonically decreasing with respect to $\tilde{d}_{ji}$. Therefore, when m_i is closer to $m_{\hat{A}_j}$, $Support_{ji}$ becomes larger. When $\tilde{d}_{ji} = 0$, $Support_{ji}$ reaches its maximum value of 1. In this study, τ is preset according to the scale of the adopted evidence-difference measure and remains fixed throughout the experiments.

For a given class event $\hat{A}_j$, the conditional credibility of evidence m_i is obtained by normalizing the support degrees of all evidence sources:

$$p(c_i \mid \hat{A}_j) = \frac{Support_{ji}}{\sum_{l=1}^{N} Support_{jl}} \tag{11}$$

Here, c_i denotes the credibility event associated with evidence m_i, and $p(c_i \mid \hat{A}_j)$ represents the relative credibility of m_i conditioned on class event $\hat{A}_j$. This normalization satisfies $\sum_{i=1}^{N} p(c_i \mid \hat{A}_j) = 1$. The coefficient τ scales the evidence-event deviation $\tilde{d}_{ji}$ and controls how rapidly the conditional credibility decreases. A larger τ strengthens the distinction between consistent and inconsistent evidence, whereas a smaller τ preserves more multi-view observations and reduces the risk of over-penalizing useful but locally inconsistent evidence.

Compared with conventional DCR, existing CEF methods use evidence-difference measures to evaluate evidence credibility, alleviating counterintuitive fusion results under high conflict and improving uncertainty handling. However, they remain limited in spatio-temporal fusion. Spatially, conventional pairwise similarity calculations incur high computational cost, and support-based weighting may weaken minority high-quality evidence. Temporally, multi-frame comparison is costly, whereas adjacent-frame fusion cannot capture evidence fluctuations. This makes random noise difficult to distinguish from real evidence changes, thereby reducing temporal consistency and decision stability.

2.2. Spatial evidence consistency fusion

The core of spatial fusion involves reliably quantifying evidence inconsistency under unknown priors and classes to enable rapid fusion and high-quality information extraction.While pairwise comparisons based on local consistency can accurately capture evidence disparities when overall quality is high, they become highly inefficient when processing massive multi-view and multi-frame evidence. Specifically, similarity and conflict-based pairwise metrics induce quadratic computational complexity that strictly conflicts with real-time constraints. Furthermore, indiscriminate weighted averaging allows abundant low-quality inputs to dilute high-belief evidence from critical views, causing recognition results to drift toward the average and reducing discriminative utility, as illustrated in Example 2.1.

Table 1
BBAs for nodes

Nodes	Sensors	$m(A)$	$m(B)$	$m(A,B)$
High credibility nodes	$\{S_1, S_2\}$	0.8	0.1	0.1
Low credibility nodes	$\{S_3, \cdots, S_N\}$	0.1	0.4	0.5

Table 2
Number of pairwise comparisons

Number of evidence	10	20	50	100	1000
Number of computations	45	190	1225	4950	499500

Table 3
Fusion results with different numbers of evidence

Number of evidence	$\{A\}$	$\{B\}$	$\{A,B\}$
2	0.8	0.1	0.1
3	0.638	0.169	0.193
5	0.339	0.297	0.364
7	0.273	0.326	0.401
8	0.231	0.344	0.425
10	0.173	0.369	0.458

Example 2.1.
Consider an urban low-altitude sensing network with N distributed sensor nodes $S = \{S_1, S_2, \cdots, S_N\}$ within the recognition frame $\Omega = \{A, B\}$. Let A represent the ground-truth type of the target. At each frame, each node performs local processing and transmits aggregated evidence $m_i, i = 1, 2, \cdots, N$, to the fusion center. Owing to building occlusions and field-of-view constraints, only a subset of nodes operate at favorable viewpoints and provide high-belief evidence supporting the true type A. The basic belief assignments (BBAs) of all nodes are listed in Table 1.

The RB-based pairwise comparison method [28] evaluates the mutual support degree by calculating the divergence or distance between any two pieces of evidence m_i and m_ℓ within the network. At each frame, the relationship between the number of comparisons C and the number of input evidence items N satisfies $C = N(N-1)/2$, as summarized in Table 2. The fusion results over the N node inputs are detailed in Table 3.

As shown in Table 2 and Table 3, the per-step computation count in the pairwise paradigm grows nonlinearly with the evidence scale. As the number of low-credibility inputs increases, the belief degree of the ground-truth class A declines and converges toward an indistinguishable composite class.

Example 2.1 indicates that pairwise comparisons fail to simultaneously achieve computational efficiency and effective information utilization. Consequently, within the hierarchical spatio-temporal fusion framework, spatial fusion should optimize the disparity comparison mechanism to establish a rapid fusion strategy preserving high-quality information.

To overcome these efficiency and utilization limitations, the proposed STCEF method formulates spatial fusion as a fusion-center-based consistency optimization. By combining class-event-conditioned evidence evaluation with fusion-center feedback, this approach derives evidence credibility without full pairwise comparisons. This design supports fast large-scale processing and improves high-quality information utilization, thereby providing reliable inputs for subsequent temporal fusion.

2.3. Temporal evidence consistency fusion

The core of temporal fusion lies in the temporal invariance of target types, which enhances stability by suppressing fluctuations. Unlike static combination rules where commutativity and associativity make fusion results independent of evidence order, continuous recognition strictly requires exploiting temporal sequences. In distributed urban sensing, measurement noise, clutter, target motion, occlusion, and track handover usually appear as local belief fluctuations rather than persistent type changes. As an extension of spatial fusion, leveraging historical context through temporal order establishes a robust mechanism to mitigate these belief fluctuations arising from building occlusions, sensor

Table 4
Shannon and Rényi entropy results

Metric	Shannon entropy	Rényi entropy
$H(t-1)$	0.394	0.205
$H(t)$	1.089	1.079
Change rate	176.1%	427.4%

faults, or target pose changes. This temporal consistency constraint suppresses anomalous evidence while maintaining stable decisions for continuous recognition tasks.

Entropy provides a quantitative measure of uncertainty in the spatial fusion result and can therefore be used to characterize temporal stability. Compared with Shannon entropy [37], Rényi entropy [38] adjusts its sensitivity to the belief distribution through the order parameter α, allowing uncertainty changes in dominant and ambiguous propositions to be distinguished more flexibly. For a BBA $m = [m(A_1), m(A_2), \cdots, m(A_J)]$ defined over the $J = 2^n - 1$ non-empty focal elements, Shannon entropy H_S and Rényi entropy H_R are defined as follows.

$$H_S(m) = -\sum_{k=1}^{J} m(A_k) \log m(A_k) \tag{12}$$

$$H_R(m) = \frac{1}{1-\alpha} \log \left(\sum_{k=1}^{J} [m(A_k)]^{\alpha} \right), \qquad \alpha > 0,\ \alpha \neq 1 \tag{13}$$

Example 2.2.

Consider a typical interference recognition scenario with the FoD $\Omega = \{A, B\}$. The BBA is represented over the propositions $\{A\}$, $\{B\}$, and Ω, and the assigned masses sum to one. During the detection process, occlusion changes the evidence from the high-belief state $m_{t-1} = [0.9, 0.05, 0.05]$ to the ambiguous state $m_t = [0.3, 0.3, 0.4]$, where the three entries correspond to $m(A)$, $m(B)$, and $m(\Omega)$, respectively. Shannon entropy and Rényi entropy with $\alpha = 2$ are calculated using Eqs.12 and 13, and the results are listed in Table 4.

Table 4 shows that Rényi entropy is more sensitive to changes in class belief, thereby enabling earlier detection of recognition risk. However, in distributed urban low-altitude recognition, absolute entropy values exhibit dynamic drift influenced by target types, observation sets, and environmental noise intensity. Relying on single-frame entropy for temporal fusion adjustments may lead the system to misclassify normal fluctuations as anomalies. In particular, direct weight assignment based on the high-sensitivity Rényi entropy can amplify minor disturbances into excessive adjustments, leading to unnecessary output fluctuations.

To address this, an adaptive weighting mechanism based on the relative entropy variation between the current evidence and the sliding-window historical reference is developed to preserve Rényi entropy sensitivity while maintaining global system stability. Initially, a sliding-window establishes a local temporal reference by fusing recent historical evidence. Subsequently, the entropy deviation between the current spatial fusion result and the historical reference is measured to quantify fluctuation intensity using the relative change rate. When the degree of deviation increases, the rising uncertainty in the current input prompts the system to reduce the fusion weight of current evidence and enhance the historical contributions to suppress transient class changes. As the deviation subsides, the system gradually restores current evidence weights to achieve an adaptive balancing between historical and current information. By using a relative deviation measure between the sliding-window reference and the current entropy, the method establishes a stable local baseline. This mechanism suppresses random outliers in spatial recognition, strengthens temporal robustness, and reduces decision risks induced by spatio-temporal inconsistency.

The proposed STCEF achieves credible evidence fusion by balancing efficiency, information utilization, and stability. Spatially, consistency optimization based on conditional credibility reduces computational complexity and enhances reliable information utilization. Temporally, Rényi entropy quantifies fluctuations to mitigate transient interference and ensure decision stability. STCEF decomposes the complex recognition task into two modules that provide a reliable fusion process for continuous and credible UAV type recognition. The research framework is shown in Fig.2. Specifically, with multi-sensor inputs, spatial fusion first addresses efficiency through evidence credibility calculation, DCR-based fusion-center computation, and event probability calculation. Subsequently, sliding-window Rényi entropy weighting is applied to yield stable and integrated recognition results.

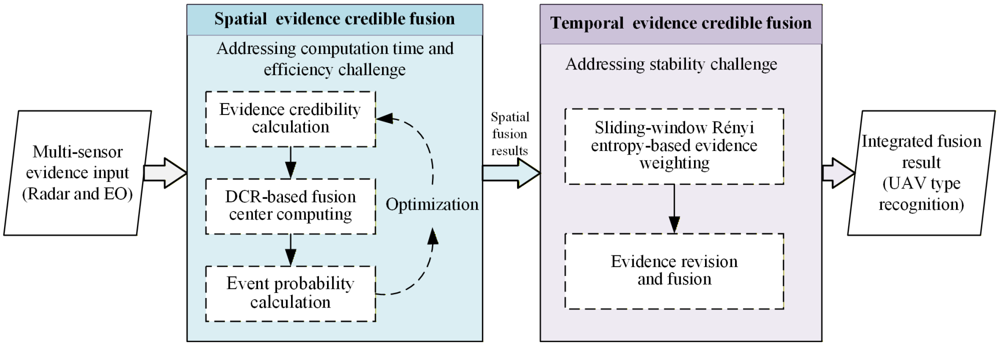


Figure 2: The overall research framework

3. Spatio-temporal credible evidence fusion

STCEF establishes a spatio-temporal hierarchical structure to address the high complexity and instability of fusing large-scale, inconsistent evidence. Spatially, conditional credibility replaces full pairwise comparisons by evaluating evidence with respect to class events and updating class-event probabilities through the fusion center, thereby improving information utilization and computational efficiency. Temporally, sliding-window Rényi entropy monitors fusion fluctuations to dynamically adjust the current-evidence weight, effectively suppressing interference and ensuring temporal stability.

The STCEF algorithm flowchart is shown in Fig.3, where the evidence fusion process is divided into a spatial fusion stage and a temporal fusion stage. In the spatial fusion stage at time t, multi-source evidence $m_1^t, m_2^t, \cdots, m_N^t$ from different sensors at the same frame is processed using credible evidence fusion, which optimizes the fusion weights and yields spatially fused evidence $\bar{m}^t$ with improved credibility and decision consistency. In the temporal stage, the sequence of spatially fused evidence $\bar{m}^1, \bar{m}^2, \cdots, \bar{m}^t$ from frame 1 to t is processed using Rényi entropy variation to derive an adaptive weight for the current evidence $\bar{m}^t$. The weighted current evidence is then combined with the historical reference h^t to produce the integrated result m^t at time t.

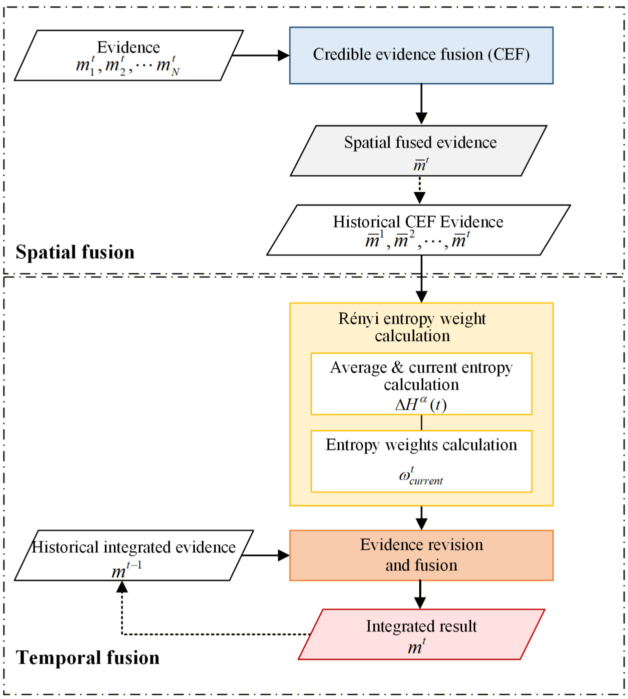


Figure 3: Processing flow of the proposed method STCEF

Overall, this spatio-temporal collaborative mechanism provides a reliable and complementary framework for continuous recognition tasks. This section also analyzes the algorithmic complexity of spatial fusion, temporal fusion, and the complete spatio-temporal fusion process to clarify the scalability of the proposed method under large-scale sensor-node conditions.

3.1. Spatial evidence consistency fusion based on conditional credibility

The core of spatial fusion is to quantify evidence credibility under unknown priors and target classes, thereby enabling efficient fusion and high-quality information extraction. Within the CEF framework, evidence credibility, the fusion center, and event probability form a closed-loop update. The current event probability is used to estimate evidence credibility, and credibility-guided DCR fusion generates the fusion center. The fusion center is then mapped into an updated event probability through the Pignistic probability function. This process is described by Eq.14 - 16.

$$Cred_i = \sum_{j=1}^{n} p\left(c_i \middle| \hat{A}_j\right) p\left(\hat{A}_j\right) \tag{14}$$

$$m_{\oplus} = DCR\left(Cred; m_1, m_2, \cdots, m_N\right) \tag{15}$$

$$p\left(\hat{A}_j\right) = BetP_{m_{\oplus}}\left(\hat{A}_j\right) \tag{16}$$

In Eq.14, $Cred_i$ is calculated by combining the conditional credibility $p(c_i \mid \hat{A}_j)$ with the current class-event probability $p(\hat{A}_j)$. A larger $Cred_i$ indicates that evidence m_i is more consistent with the current event distribution and should contribute more to fusion. Eq.15 uses the credibility set $Cred$ to perform DCR-weighted fusion and obtain the fusion center $m_{\oplus}$. The fusion center $m_{\oplus}$ represents the current global consensus among the input evidence. This center-oriented evaluation avoids repeated pairwise comparisons, reduces computational burden, and suppresses evidence with low consistency. Eq.16 maps $m_{\oplus}$ into the event probability $p\left(\hat{A}_j\right)$, which is then used in the next credibility update. In the absence of prior probabilities, the event probability is initialized as a uniform distribution according to Eq.17.

$$p^{(0)}(\hat{A}_j) = \frac{1}{n} \tag{17}$$

Eqs.14 - 17 define the closed-loop credibility update. Once the event probability becomes stable, the corresponding fusion center is retained as the spatial fusion result for subsequent temporal processing.

Example 3.1.

Low oil pressure, intake system leakage, and a stuck solenoid valve are three common faults in automotive systems, denoted by $\{A\}$, $\{B\}$, and $\{C\}$, respectively. To enable drivers to promptly and accurately identify the fault cause, five sensors installed at different locations in the vehicle continuously report information on the system operating status. These reports are typically converted into evidence representations to facilitate fusion and support decision-making. Suppose that, at a certain point in time, a low oil pressure fault occurs in the automotive system; Sensors 1-4 operate normally, whereas Sensor 5 provides an abnormal report due to random interference or a sensor fault. The evidence reported by the five sensors is presented in Table 5. To mitigate the effect of unreliable sensor reports, particularly from Sensor 5, these five pieces of evidence need to be fused credibly to accurately infer the fault in the automotive system.

Based on the evidence inputs in Table 5, m_1, m_2, m_3, and m_4 mainly support the proposition $\{A\}$, whereas m_5 assigns a high belief degree to $\{C\}$ and conflicts with the remaining evidence. As shown in Fig.4, the credibility of m_5 decreases rapidly, while the credibilities of m_1, m_2, m_3, and m_4 gradually become stable. This result indicates that the closed-loop credibility update suppresses conflicting evidence and produces a stable fusion center after a small number of updates.

3.2. Temporal evidence consistency fusion based on Rényi entropy quantification

The core of temporal fusion lies in the temporal invariance of target types, which mitigates decision risk by suppressing fluctuations arising from building occlusions, sensor faults, or target pose changes. Although Section 3.1 improves recognition performance through consistency optimization, discrete spatial outputs may still exhibit anomalies or misclassifications in dynamic environments, which undermines the stability of continuous recognition.

Table 5
Basic probability assignments for the spatial fusion example

Evidence	$\{A\}$	$\{B\}$	$\{C\}$	$\{A, B, C\}$
m_1	0.70	0.10	0.00	0.20
m_2	0.70	0.00	0.00	0.30
m_3	0.65	0.15	0.00	0.20
m_4	0.75	0.00	0.05	0.20
m_5	0.00	0.20	0.80	0.00

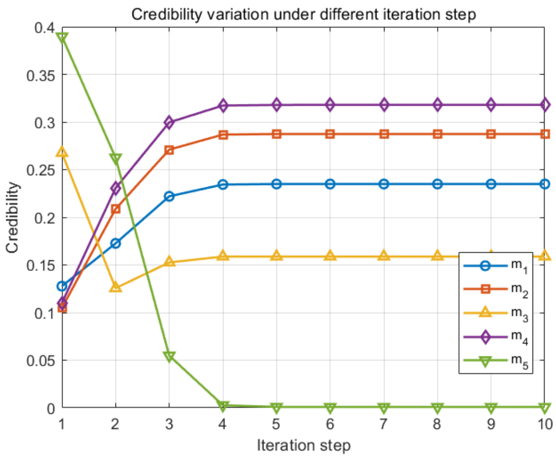


Figure 4: Credibility variation under different iteration steps

Therefore, a sliding-window Rényi entropy strategy quantifies the uncertainty of the belief distribution following spatial fusion at each frame. A historical entropy baseline is then constructed from the fused evidence within the sliding window, and the relative change rate of the current entropy is used to quantify fluctuation intensity and adapt the fusion weight of current evidence. By replacing absolute entropy values with relative deviation metrics, the proposed STCEF preserves Rényi entropy sensitivity while avoiding excessive weight adjustments, thereby enabling stable and credible recognition under dynamic interference. The proposed temporal fusion method consists of two parts, Rényi entropy-based weight calculation and evidence revision.

3.2.1. Sliding-window Rényi entropy-based evidence weighting

Based on the recognition framework Ω, at time t, the multi-sensor evidence set is denoted as $\{m_1^t, m_2^t, \cdots, m_N^t\}$. Spatial credible fusion is first applied to obtain the fused evidence $\bar{m}^t$:

$$\bar{m}^t = DCR\left(Cred^t; m_1^t, m_2^t, \cdots m_N^t\right) \tag{18}$$

Up to time t, the spatial fusion results $\bar{m}^t$ form the set $\bar{m}(t) = \{\bar{m}^1, \bar{m}^2, \cdots, \bar{m}^t\}$.

- Rényi entropy calculation

  At time t, let $\bar{m}^t$ denote the spatially fused BBA defined over the $J = 2^n - 1$ non-empty focal elements of 2^Ω.

$$H_\alpha(\bar{m}^t) = \begin{cases} -\sum_{k=1}^{J} \bar{m}^t(A_k)\log(\bar{m}^t(A_k)), & \alpha = 1 \\ \frac{1}{1-\alpha}\log\left[\sum_{k=1}^{J}\left(\bar{m}^t(A_k)\right)^\alpha\right] & \alpha \neq 1 \end{cases} \tag{19}$$

  Here, $\alpha > 0$ represents the Rényi entropy order parameter, which controls the sensitivity to the probability distribution. When $\alpha = 1$, the entropy is defined by its limiting form, namely Shannon entropy.

- Sliding-window design

For the current time t, the preceding K spatial fusion results are selected to form a sliding time window, denoted as $\mathcal{H}^t = \{\bar{m}^{t-K}, \bar{m}^{t-K+1}, \cdots, \bar{m}^{t-1}\}$. The evidence within this window is fused to obtain the historical fusion evidence h^t.

$$h^t = DCR(\bar{m}^{t-K}, \bar{m}^{t-K+1}, \cdots, \bar{m}^{t-1}) \tag{20}$$

When $2 \le t \le K$, fewer than K historical spatial fusion results are available, and the previous integrated result m^{t-1} is used as the historical reference. When $t > K$, the historical reference is constructed from the preceding K spatial fusion results according to Eq.20. A larger K incorporates more historical information and generally yields smoother outputs, whereas a smaller K provides faster response to changes in the current evidence.

- Dynamic discount weights calculation

  The Rényi entropies of the historical reference and current spatial fusion result are respectively calculated as $H_{history} = H_\alpha(h^t)$ and $H_{current} = H_\alpha(\bar{m}^t)$. Their relative variation is denoted by $\Delta H_\alpha(t)$:

$$\Delta H_\alpha(t) = \frac{H_{current} - H_{history}}{\sqrt{\left(H_{history}\right)^2 + \varepsilon^2}} \tag{21}$$

  Here, $\varepsilon > 0$ is a preset small constant introduced only to avoid numerical instability when the historical entropy approaches zero.

$$\omega_t = \frac{1}{1 + \exp\left[\gamma\left(\Delta H_\alpha(t) - \theta\right)\right]} \tag{22}$$

  Here, γ controls the slope of the Sigmoid mapping, and θ specifies the entropy-variation threshold for weight adjustment. A larger $\Delta H_\alpha(t)$ indicates that the current spatial fusion result is more uncertain than the historical reference, resulting in a smaller current-evidence weight ω_t. When $\Delta H_\alpha(t)$ decreases, the contribution of the current evidence is gradually restored.

3.2.2. *Evidence revision and fusion*

Based on the entropy variation, the current-evidence weight ω_t is obtained. The spatial fusion result $\bar{m}^t$ is discounted using ω_t to generate the corrected evidence m^{t*}, which is subsequently combined with the historical reference h^t using Dempster's rule of combination.

$$m^{t*}(A) = \begin{cases} \omega_t \bar{m}^t(A), & A \in 2^\Omega \setminus \{\Omega\}, \\ 1 - \omega_t + \omega_t \bar{m}^t(\Omega), & A = \Omega. \end{cases} \tag{23}$$

For all proper subsets of Ω, the assigned masses are multiplied by ω_t, while the remaining mass is transferred to the total ignorance proposition Ω. This discounting operation preserves the normalization of m^{t*}. The corrected evidence is then combined with the historical reference h^t:

$$m^t = DCR(h^t, m^{t*}) \tag{24}$$

At time t, the evidence set for historical fusion results is $m(t)$, $m(t) = \{m^1, m^2, \cdots, m^t\}$. The fusion result m^t at each time point is sequentially output as the system's integrated recognition result.

The complete STCEF procedure is summarized in Algorithm 1. The spatial stage applies the closed-loop credibility update defined by Eqs.14–16, while the temporal stage directly follows Eqs.19–24.

Algorithm 1: Spatio-temporal credible evidence fusion

Input: Frame of discernment $\Omega = \{\hat{A}_1, \hat{A}_2, \cdots, \hat{A}_n\}$; multi-sensor evidence sequence $\{m_i^t\}_{i=1,t=1}^{N,T}$; sliding-window length K; parameters α, γ, and θ;

Output: Spatial fusion sequence $\{\bar{m}^1, \bar{m}^2, \cdots, \bar{m}^T\}$; integrated fusion sequence $\{m^1, m^2, \cdots, m^T\}$; temporal weights $\{\omega_1, \omega_2, \cdots, \omega_T\}$.

Construct the event-conditional evidence items;
for $t = 1$ *to* T **do**
  Construct the event evaluation matrix D^t_{EEM};
  Calculate the conditional credibility $p(c_i \mid \hat{A}_j)$ according to Eqs.8–11;
  Initialize the event probability according to Eq.17;
  repeat
    Calculate the evidence credibility according to Eq.14;
    Obtain the spatial fusion center according to Eq.15;
    Update the event probability according to Eq.16;
  until *the event-probability distribution becomes stable*;
  Retain the resulting fusion center as $\bar{m}^t$;
end
$m^1 \leftarrow \bar{m}^1$;
$\omega_1 \leftarrow 1$;
for $t = 2$ *to* T **do**
  if $t \leq K$ **then**
    $h^t \leftarrow m^{t-1}$;
  else
    $h^t \leftarrow DCR\left(\bar{m}^{t-K}, \bar{m}^{t-K+1}, \cdots, \bar{m}^{t-1}\right)$;
  end
  Calculate $H_\alpha(\bar{m}^t)$ and $H_\alpha(h^t)$ according to Eq.19;
  Calculate $\Delta H_\alpha(t)$ according to Eq.21;
  Calculate ω_t according to Eq.22;
  Discount $\bar{m}^t$ according to Eq.23 to obtain m^{t*};
  $m^t \leftarrow DCR(h^t, m^{t*})$;
end
return $\{\bar{m}^1, \bar{m}^2, \cdots, \bar{m}^T\}$, $\{m^1, m^2, \cdots, m^T\}$, *and* $\{\omega_1, \omega_2, \cdots, \omega_T\}$.

3.3. Computational complexity analysis

In distributed low-altitude UAV collaborative recognition tasks, the computational complexity is mainly determined by the number of spatial evidence items in a single frame, the number of focal elements in the FoD, and the temporal sliding-window length. Let N denote the number of input evidence items at each frame, n the number of singleton class events in the FoD Ω, $J = 2^n - 1$ the number of non-empty focal elements, K the Rényi entropy sliding-window length, and T the temporal sequence length. Let $G(J)$ denote the computational complexity of calculating the evidence difference between two J-dimensional BBAs. Here, $i = 1, 2, \cdots, N$ is the evidence index, $j = 1, 2, \cdots, n$ is the class-event index, and $k = 1, 2, \cdots, J$ is the focal-element index. For the asymptotic comparison used in this paper, the finite update overhead of the closed-loop credibility calculation is absorbed into a bounded factor.

3.3.1. Spatial fusion complexity

For DS fusion, sequential combination of N evidence items requires approximately $N - 1$ applications of Dempster's rule. Since each pairwise combination over J focal elements has a complexity of $\mathcal{O}(J^2)$, the complexity of a single DS spatial fusion operation is

$$C_{DS} = \mathcal{O}\left(NJ^2\right) \tag{25}$$

For the Murphy Average Cost (MAC) method, the N pieces of evidence are first averaged or preprocessed, with a complexity of $\mathcal{O}(NJ)$, followed by Dempster's rule of combination. Therefore, the complexity of a single MAC spatial fusion operation is

$$C_{MAC} = \mathcal{O}\left(NJ + NJ^2\right) \tag{26}$$

CEF methods based on pairwise comparison metrics, such as RB and PBLBJS, usually require constructing an evidence-difference matrix to evaluate the similarity, conflict degree, or credibility-support relationships among evidence items. For N input evidence items, the number of pairwise comparisons is $N(N-1)/2$. Therefore, the construction complexity of the evidence-difference matrix is $\mathcal{O}\left(N^2G(J)\right)$. Considering the complexity $\mathcal{O}(NJ)$ of evidence weighting or discounting and the complexity $\mathcal{O}(NJ^2)$ of Dempster's rule of combination, the complexity of a single spatial fusion operation for CEF methods based on pairwise comparison metrics is

$$C_{pair} = \mathcal{O}\left(N^2G(J) + NJ + NJ^2\right) \tag{27}$$

Here, $N^2G(J)$ arises from pairwise evidence-difference calculations. As the number of input evidence items increases, this term substantially increases the computational overhead.

The proposed ICEF spatial fusion module adopts fusion-center-guided conditional credibility estimation. Its computational process is jointly formulated in terms of evidence credibility $Cred_i$, the fusion center $m_{\oplus}$, and class-event probability $p(\hat{A}_j)$. In a single ICEF-based spatial fusion operation, the event-evidence evaluation matrix has a size of $n \times N$, and its construction complexity is $\mathcal{O}\left(nNG(J)\right)$. The complexity of calculating conditional credibility is $\mathcal{O}(nN)$, the complexity of evidence weighting or discounting is $\mathcal{O}(NJ)$, the complexity of calculating the fusion center is $\mathcal{O}(NJ^2)$, and the complexity of calculating class-event probabilities is $\mathcal{O}(nJ)$. After absorbing the finite closed-loop update overhead into a bounded factor, the complexity of a single ICEF-based spatial fusion operation is

$$C_{ICEF} = \mathcal{O}\left(nNG(J) + nN + NJ + NJ^2 + nJ\right) \tag{28}$$

Compared with CEF methods based on pairwise comparison metrics, ICEF does not need to construct an $N \times N$ evidence-difference matrix. Its spatial credibility assessment is completed through an $n \times N$ event-evidence evaluation matrix, thereby avoiding the quadratic pairwise-comparison term with respect to the evidence scale.

3.3.2. Temporal fusion complexity

After spatial fusion, the temporal module processes the sequence of spatially fused results, denoted as $\{\bar{m}^1, \bar{m}^2, \cdots, \bar{m}^T\}$. Therefore, the single-step complexity of temporal fusion is independent of the number of spatial evidence items N at each frame and is mainly determined by the number of focal elements J and the sliding-window length K.

For temporal DS fusion (TDS), at each time step, the historical fusion result is combined once with the current spatial fusion result using Dempster's rule of combination. The complexity of a single fusion operation is

$$C_{TDS} = \mathcal{O}\left(J^2\right) \tag{29}$$

For Time Decay (TD), evidence discounting and Dempster's rule of combination are required. Therefore, the complexity of a single fusion operation is

$$C_{TD} = \mathcal{O}\left(J + J^2\right) \tag{30}$$

For temporal fusion based on Rényi entropy, each time step requires constructing the sliding-window historical reference, calculating the entropy of the current evidence and historical reference, calculating the relative entropy variation and Sigmoid weight, discounting the current evidence, and applying Dempster's rule of combination. When the sliding-window length is K, the single-step complexity is

$$C_{RE} = \mathcal{O}\left(KJ^2 + J + J^2\right) \tag{31}$$

Here, KJ^2 represents the complexity of sliding-window historical fusion, J represents the linear complexity of Rényi entropy calculation and evidence discounting, and J^2 corresponds to the complexity of Dempster's rule of combination.

3.3.3. *Spatio-temporal fusion complexity*

The complete single-frame spatio-temporal fusion process consists of one spatial fusion operation and one temporal fusion operation. Therefore, for a continuous recognition sequence of length T, the spatio-temporal fusion complexity is obtained by combining the single-frame spatial complexity and the single-step temporal complexity.

Accordingly, the complexity of classical DS-based spatial fusion combined with TDS-based temporal fusion is

$$C_{DS+TDS} = \mathcal{O}\left[T\left(NJ^2 + J^2\right)\right] \tag{32}$$

The complexity of MAC-based spatial fusion combined with TD-based temporal fusion is

$$C_{MAC+TD} = \mathcal{O}\left[T\left(NJ + NJ^2 + J + J^2\right)\right] \tag{33}$$

The complexity of CEF methods based on pairwise comparison metrics combined with temporal fusion based on Rényi entropy is

$$C_{pair+RE} = \mathcal{O}\left[T\left(N^2G(J) + NJ + NJ^2 + KJ^2 + J + J^2\right)\right] \tag{34}$$

The complexity of the proposed STCEF method is

$$C_{STCEF} = \mathcal{O}\left[T\left(nNG(J) + nN + NJ + NJ^2 + nJ + KJ^2 + J + J^2\right)\right] \tag{35}$$

The above complexity analysis shows that the computational differences among different methods mainly arise from the spatial credibility-assessment strategy and the temporal correction mechanism. Spatially, DS and MAC have relatively low basic computational overhead, but they lack mechanisms for measuring anomalous evidence and differences in evidence quality. CEF methods based on pairwise comparison metrics can characterize differences among evidence items, but they require constructing an $N \times N$ evidence-difference matrix, whose complexity includes $\mathcal{O}\left(N^2G(J)\right)$ and thus increases substantially under large-scale node inputs.

In contrast, at the spatial level, STCEF adopts the ICEF mechanism, replacing full pairwise comparisons with an $n \times N$ event-evidence evaluation matrix and fusion-center-guided credibility updating, thereby reducing the computational overhead of credibility assessment under large-scale spatial evidence. Temporally, Rényi entropy sliding-window fusion introduces additional steps, including historical-reference construction, entropy calculation, and weight mapping. However, this process only operates on a single time series after spatial fusion, and its complexity is independent of the number of spatial evidence items N, mainly depending on the fixed parameters K and J.

Therefore, STCEF is more suitable for complex scenarios with large-scale spatial evidence, uneven evidence quality, and a requirement for continuous and stable recognition. Subsequent experiments measure the running time of each module to verify the computational applicability of the proposed method.

4. Simulation

Three simulations are designed to validate the effectiveness of the proposed method. Section 4.1 evaluates the computational efficiency and accuracy of spatial evidence consistency fusion. Section 4.2 verifies the accuracy of temporal evidence consistency fusion. Section 4.3 assesses the overall performance of STCEF in a collaborative radar-EO scenario for UAV type recognition.

4.1. Spatial evidence consistency fusion

This section evaluates the proposed spatial evidence consistency fusion method in terms of computational efficiency and recognition accuracy. The ICEF mechanism implements the fusion-center-guided closed-loop credibility update described in Section 3.1. To evaluate computational efficiency, evidence sets with different input scales are used to compare pairwise credibility methods with their ICEF-enhanced counterparts. The target recognition frame is $\Omega = \{A, B, C\}$, and each BBA is defined over the $2^{|\Omega|} - 1$ non-empty focal elements. The compared methods include the classical DS rule, Dismp [27], RB [28], and PBLBJS [29]. Their computation times under different evidence scales are reported in Table 6.

As shown in Table 6, the computation time of pairwise comparison algorithms increases rapidly as the number of inputs grows. By introducing the ICEF mechanism, computational efficiency is substantially improved at larger input

Table 6
Computation time of spatial methods across different input scales (s)

Method	Number of Evidence									
	3	5	10	30	50	100	200	500	1000	5000
DS	0.0018	0.0021	0.0028	0.0038	0.0052	0.0087	0.0127	0.0302	0.0569	0.3339
Dismp	0.0006	0.0009	0.0019	0.0121	0.0283	0.1002	0.3731	2.3436	9.5450	250.2063
RB	0.0004	0.0006	0.0020	0.0088	0.0198	0.0584	0.2243	1.3456	5.6890	135.9080
PBLBJS	0.0007	0.0005	0.0013	0.0068	0.0319	0.0842	0.5506	1.6432	3.3031	6.1982
ICEF-Dismp	0.0012	0.0019	0.0034	0.0102	0.0187	0.0351	0.0779	0.1631	0.3261	1.6819
ICEF-RB	0.0011	0.0018	0.0053	0.0106	0.0224	0.0306	0.0696	0.1550	0.2933	1.6518
ICEF-PBLBJS	0.0012	0.0015	0.0028	0.0086	0.0164	0.0296	0.0584	0.1545	0.3046	1.4551

Table 7
Mass function distributions of five sensors

Type/Number	A	B	C	A,B	A,C	B,C	A,B,C
Mass1	0.40	0.28	0.30	0.00	0.02	0.00	0.00
Mass2	0.01	0.90	0.08	0.00	0.01	0.00	0.00
Mass3	0.63	0.06	0.01	0.00	0.30	0.00	0.00
Mass4	0.60	0.09	0.01	0.00	0.30	0.00	0.00
Mass5	0.60	0.09	0.01	0.00	0.30	0.00	0.00

Table 8
Fusion results based on different combination rules

Type/Method	A	B	C	A, B	A, C	B, C	A, B, C
DS	0.8657	0.0168	0.1167	0	0.0008	0	0
Dismp	0.9595	0.0014	0.0193	0	0.0198	0	0
RB	0.9888	0.0015	0.0073	0	0.0024	0	0
PBLBJS	0.9892	0.0002	0.0085	0	0.0021	0	0
ICEF-Dismp	**0.9926**	0.0001	0.0041	0.0001	0.0030	0.0001	0
ICEF-RB	**0.9899**	0.0013	0.0062	0	0.0026	0	0
ICEF-PBLBJS	**0.9934**	0.0001	0.0034	0	0.0031	0	0

scales through event-evidence evaluation and fusion-center-guided credibility updating. Further simulations evaluate the recognition accuracy of different algorithms.

Based on the experimental data in [39], several spatial fusion methods are compared to evaluate the recognition accuracy. The target recognition framework $\Theta = \{A, B, C\}$ involves classifying target types based on mass functions provided by spatially distributed sensors $\{S_1, S_2, S_3, S_4, S_5\}$, as detailed in Table 7. As shown in Table 8, the incorporation of ICEF yields a marked increase in belief across all pairwise comparison methods, thereby confirming its effectiveness in enhancing spatial evidence consistency.

4.2. Temporal consistency fusion

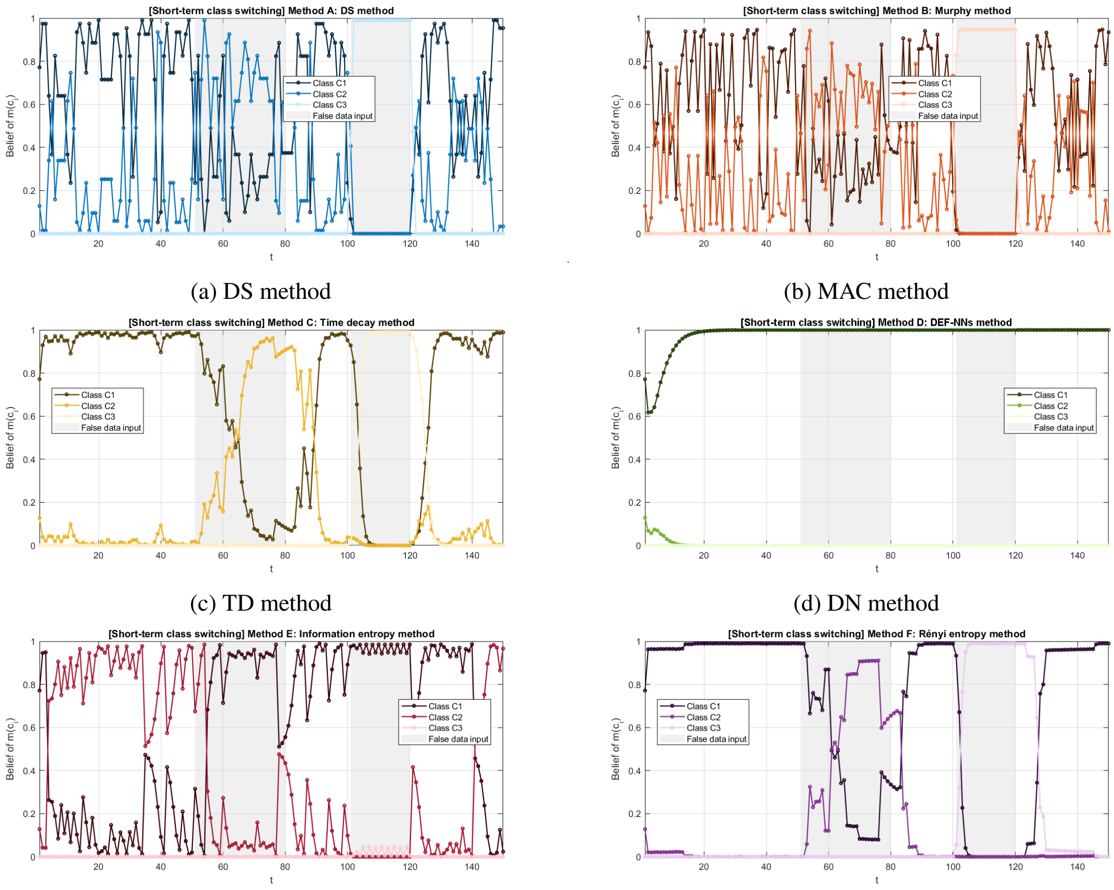
This section evaluates the temporal evidence consistency fusion method based on Rényi entropy, denoted as RE, using 150 frames of radar sequential recognition evidence. The compared methods include Temporal DS (TDS), Murphy Average Cost (MAC) [22], Time Decay (TD) [33], DEF-NNs (DN) [35], and information entropy (IN) [36]. The target recognition framework includes three typical UAV types: medium-small UAV (MSUAV), light-small UAV (LSUAV), and micro UAV (MUAV). The true class is set to MSUAV during frames 1–50, LSUAV during frames 51–80, MSUAV during frames 81–100, MUAV during frames 101–120, and MSUAV during frames 121–150. The BBAs are generated following [27] using statistical characteristics derived from radar RCS and kinematic attributes.

As shown in Fig.5, TDS and MAC exhibit obvious belief fluctuations during abnormal evidence intervals, indicating that they are sensitive to conflicting evidence. TD reduces part of the fluctuation, but still shows a certain response delay near class transition boundaries. DN maintains a high $m(c_1)$ for a long period, showing strong stability, but also indicating insufficient response to subsequent class changes. IN presents evident oscillations, suggesting relatively weak temporal stability. In contrast, RE can suppress abnormal fluctuations while preserving the response to class changes, resulting in a more balanced overall performance.

Fig.6 and Fig.7 further show the local variation of $m(c_1)$. TDS and MAC fluctuate strongly in the local intervals, indicating poor continuity of fusion results. TD shows smaller fluctuations, but its belief decrease and recovery are both delayed to some extent. DN almost always maintains a high $m(c_1)$, which preserves the initial-class output but fails to reflect subsequent class switching. IN fluctuates frequently in the local intervals, indicating insufficient fusion stability. RE can moderately adjust the support for $m(c_1)$ during abnormal intervals and recover rapidly after the class returns, showing better stability and responsiveness.

The statistical results in Table 9 are consistent with the curve observations. DN achieves the highest mean belief of 0.977, but its recognition accuracy is only 0.67, indicating that its high-belief output is not sufficiently adaptive to class changes. IN obtains the lowest accuracy, suggesting insufficient stability and reliability. TD maintains a relatively high mean belief, but its accuracy is 0.79, showing a certain response delay. RE achieves the highest accuracy of 0.85 and maintains a relatively high mean belief of 0.753, demonstrating better overall performance in terms of recognition accuracy, belief stability, and response capability.

Finally, to evaluate the computational efficiency of the temporal methods, their computation times across different sequence lengths are compared, as shown in Table 10. The results indicate that the RE method requires more computation time than TDS, MAC, TD, DN, and IN due to the additional calculation of Rényi entropy-based sliding-window weights. However, even when the input scale increases to 5000, its computation time remains at just 0.0674 s, which is well within an acceptable range. Combined with the spatial efficiency analysis, this demonstrates that STCEF maintains good computational scalability while significantly improving continuous recognition stability.



(a) DS method

(b) MAC method

(c) TD method

(d) DN method

(e) IN method

(f) Proposed RE method

Figure 5: Belief curves of six temporal fusion methods

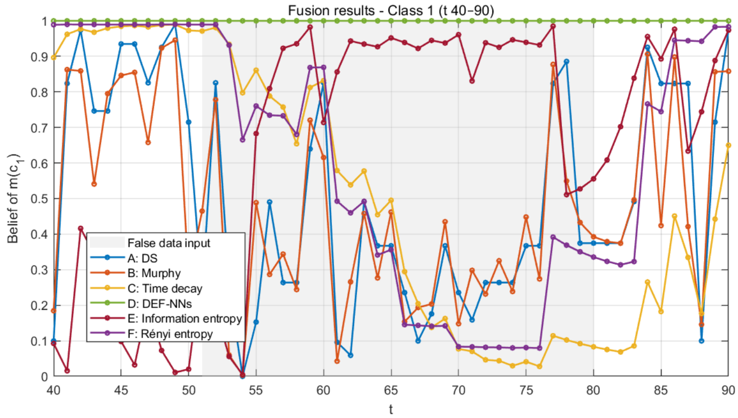


Figure 6: Detailed recognition results for Type 1 (MSUAV) during frames 40-90

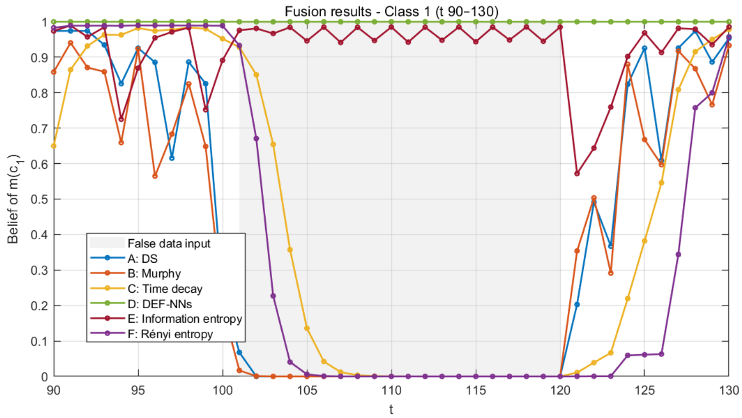


Figure 7: Detailed recognition results for Type 1 (MSUAV) during frames 90-130

Table 9
Recognition accuracy and mean belief of temporal fusion methods

Method	Accuracy	Mean belief
TDS	0.84	0.602
MAC	0.77	0.542
TD	0.79	0.729
DN	0.67	**0.977**
IN	0.31	0.551
RE	**0.85**	0.753

4.3. Spatio-temporal evidence consistency fusion

4.3.1. Simulation scenario

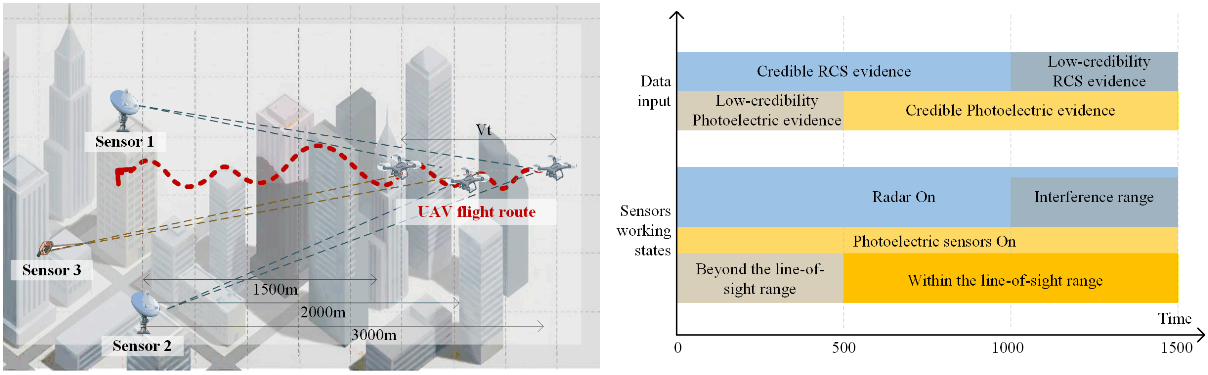
To evaluate the overall efficacy of the proposed STCEF method for integrated UAV type recognition, a simulation scenario for low-altitude UAV type recognition in an urban area is established. The experiment focuses on three UAV types consisting of MSUAV, LSUAV, and MUAV, with the "Wolf" model serving as the ground truth. Input evidence comprising target classes and associated belief levels is derived from multiple features of two radars and an EO sensor, including RCS, optical pixels, and kinematics (e.g., velocity, altitude, and maneuvers). Mean recognition belief and accuracy are adopted as the primary metrics to compare fusion methods and support the subsequent temporal fusion experiments and consistency analysis.

As illustrated in Fig.8a, an unauthorized low-altitude UAV takes off approximately 3 km away and follows the red trajectory through urban buildings, thereby constituting a potential air-defense threat. Two radars with a 10 Hz update frequency and one EO sensor operating at 60 FPS are deployed for continuous monitoring of the target. The simulation generates 1500 frames of recognition evidence. After temporal alignment, the input at each frame consists

Table 10
Computation time of temporal methods across different sequence lengths

Method	Sequence Length									
	3	5	10	30	50	100	200	500	1000	5000
TDS	0.00009	0.00001	0.00003	0.0001	0.0001	0.0002	0.0003	0.0006	0.0011	0.0051
MAC	0.00005	0.00004	0.00008	0.0002	0.0004	0.0006	0.0010	0.0023	0.0049	0.0220
TD	0.00003	0.00003	0.00006	0.0002	0.0003	0.0005	0.0009	0.0021	0.0041	0.0194
DN	0.00005	0.00003	0.00004	0.0001	0.0002	0.0003	0.0006	0.0013	0.0024	0.0118
IN	0.00007	0.00005	0.0001	0.0003	0.0005	0.0009	0.0015	0.0036	0.0071	0.0333
RE	0.0001	0.00001	0.0002	0.0006	0.0010	0.0017	0.0033	0.0074	0.0146	0.0674

of three aggregated evidence inputs, which are generated locally by three sensors using multi-attribute features such as RCS, optical pixels, and kinematics. The fusion center performs spatial fusion on these inputs and sequentially outputs integrated results after temporal fusion.

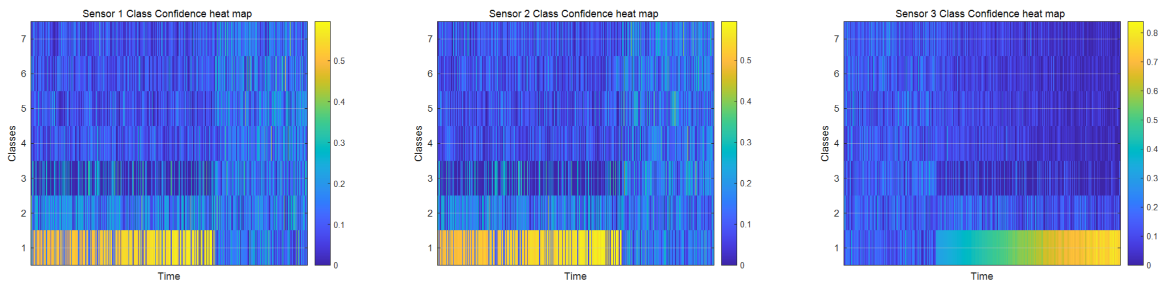


(a) Integrated simulation scenario for UAV recognition

(b) Sensors working status and data input

Figure 8: Simulation scenario

The working statuses of sensors at distinct deployment locations exhibit staged variations throughout the monitoring process, primarily driven by building occlusions and environmental clutter, as presented in Fig.8b. Throughout the recognition process, both radars and the EO sensor remain active and continuously provide evidence to the fusion system. Specific intervals are designed to simulate degraded conditions. During frames 1 to 500, the EO sensor yields low-belief evidence as the target is situated beyond the effective line-of-sight range. Conversely, during frames 1001 to 1500, radars exhibit reduced belief attributed to simulated interference. These variations are intended to emulate restricted or disturbed sensor environments and provide a basis for evaluating the robustness of the spatio-temporal fusion framework.



(a) Radar 1

(b) Radar 2

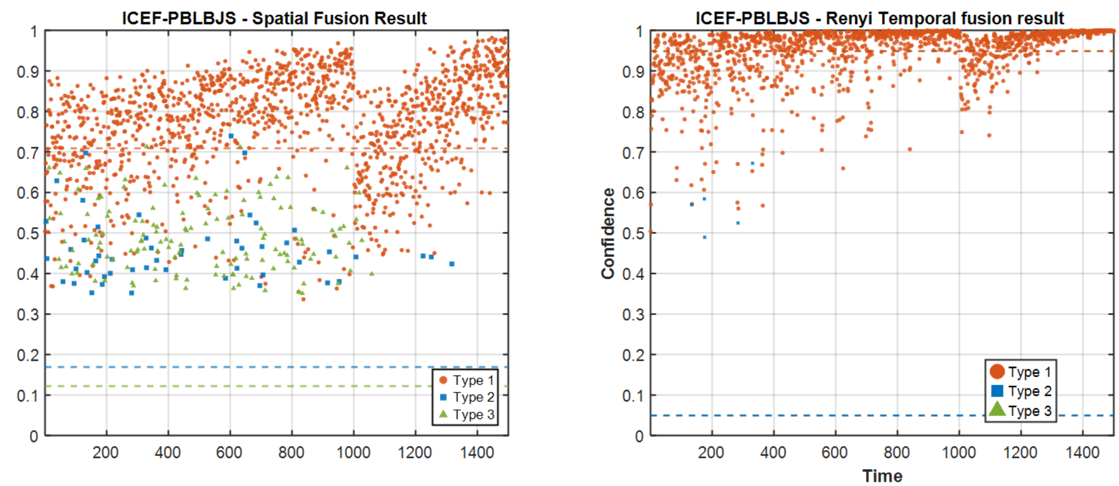
(c) EO sensor

Figure 9: Temporal class belief distribution of three sensors

Given that detection and classification capabilities are distance-dependent, the belief heatmaps for the three sensors are illustrated in Fig.9. From frames 1 to 1000, the recognition beliefs of Radars 1 and 2 regarding the true target exhibit a progressive increase. Similarly, the electro-optical sensor shows a gradual rise in belief from frames 500 to 1500. This trend is consistent with physical sensor characteristics where data quality improves as the target range decreases. Based on these 1500 frames of evidence, comparative experiments are conducted using various spatio-temporal fusion methods.

4.3.2. Spatio-temporal evidence integration

To assess fusion performance at distinct processing stages, the 1500-frame sequential recognition scenario described previously is utilized for comparative experiments. Spatial consistency fusion is evaluated using the DS, MAC, RB, PBLBJS, ICEF-RB and ICEF-PBLBJS. Temporal consistency fusion incorporates the temporal DS combination (TDS), Time Decay (TD), and the proposed RE method. The processing follows a hierarchical structure where multi-sensor evidence at each frame first undergoes spatial fusion to generate a unified spatial evidence output. The resulting spatial evidence sequence then serves as input for the temporal fusion algorithms to derive the integrated target recognition result. The belief distributions of the results obtained using the proposed STCEF method are illustrated in Fig.10a to 10b. The recognition accuracy and mean belief for the six spatial fusion methods are listed in Table 11, while the corresponding metrics for the comparative temporal fusion strategies are provided in Table 12.



(a) Spatial fusion result of STCEF

(b) Spatio-temporal fusion result of STCEF

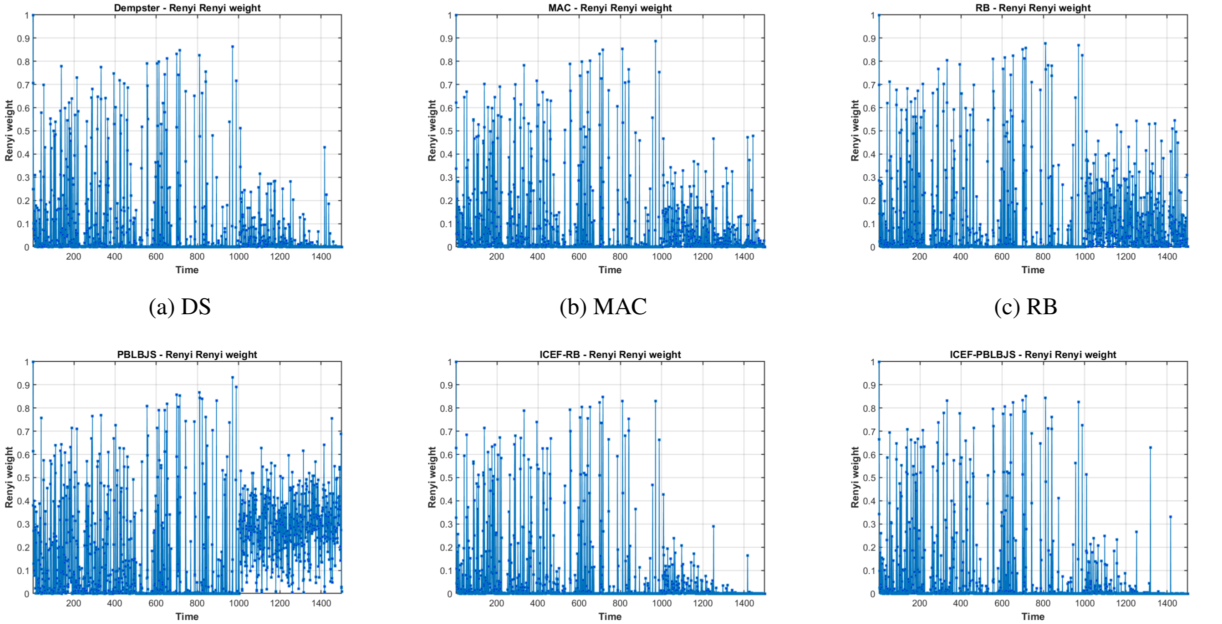
Figure 10: Spatio-temporal evidence fusion results over 1500 frames

As detailed in Table 11, ICEF-PBLBJS achieves the highest mean belief across the 1500-frame sequence, while ICEF-RB obtains the highest recognition accuracy. Both ICEF-enhanced methods outperform their corresponding original methods, indicating that the proposed spatial credibility strategy can mitigate uncertainty during evidence fusion and provide more stable and reliable outputs. Building on these spatial results, subsequent analyses examine the adaptability and temporal consistency of the various approaches in dynamic scenarios.

Table 12 presents the mean recognition belief and accuracy obtained by processing the spatial fusion outputs in Table 11 using different temporal fusion algorithms. The ICEF-enhanced spatial outputs generally achieve improved temporal fusion performance, demonstrating that the proposed spatial fusion mechanism provides stable and high-quality evidence for subsequent temporal processing.

Furthermore, the weight variation results in Fig.11 clearly reveal the dynamic regulation mechanism of the sliding-window Rényi entropy method. When target-type switching or abnormal interference causes significant inconsistency between the current evidence and the historical state, the mapped current-evidence weight changes markedly, thereby reducing the adverse influence of abrupt abnormal evidence on temporal fusion.

Based on this adaptive weighting mechanism, the ICEF-PBLBJS+RE configuration achieves the highest mean belief of 0.9490, while ICEF-RB+RE obtains the highest recognition accuracy of 0.9973. For the ICEF-PBLBJS+RE configuration, compared with DS combined with TDS, the mean belief increases by approximately 10% and the



(a) DS (b) MAC (c) RB

(d) PBLBJS (e) ICEF-RB (f) ICEF-PBLBJS

Figure 11: RE method fusion weight variation

Table 11
Performance of spatial fusion methods

Method	Mean Belief	True Type Accuracy
DS	0.6725	0.8687
MAC	0.6443	0.8620
RB	0.6318	0.8327
PBLBJS	0.5415	0.6740
ICEF-RB	0.6890	**0.8820**
ICEF-PBLBJS	**0.7088**	0.8747

Table 12
Mean belief and accuracy of three temporal fusion methods

Method	TDS		TD		RE	
	Mean belief	Accuracy	Mean belief	Accuracy	Mean belief	Accuracy
DS	0.8627	0.9760	0.8343	0.9847	0.9303	0.9960
MAC	0.8445	0.9753	0.8003	0.9740	0.9131	0.9953
RB	0.8325	0.9587	0.7888	0.9653	0.9026	0.9940
PBLBJS	0.7358	0.8480	0.7028	0.8460	0.7930	0.9087
ICEF-RB	0.8784	0.9807	0.8301	0.9780	0.9404	0.9973
ICEF-PBLBJS	0.8896	0.9793	0.8593	0.9867	**0.9490**	0.9967

recognition accuracy increases by approximately 2.1%. These results demonstrate that integrating high-quality spatial fusion with adaptive temporal weighting improves overall spatio-temporal performance.

4.4. Parameter analysis and ablation experiments

To further verify the effects of key parameters and functional modules in the proposed STCEF method on fusion performance, this section conducts parameter sensitivity analysis and ablation experiments. First, a single-factor analysis is performed for the Rényi entropy order α, the Sigmoid slope γ, and the entropy-change threshold θ in the temporal fusion stage to illustrate the effects of parameter variations on temporal fusion weights, recognition accuracy, and mean belief. Subsequently, the effect of the sliding-window length K on historical information utilization and response lag is analyzed. Finally, module ablation experiments are conducted to verify the contributions of spatial credible fusion, temporal Rényi entropy weighting, and the complete STCEF framework.

4.4.1. Rényi entropy parameter analysis

In the temporal fusion stage, Rényi entropy is used to measure the uncertainty variation between the current spatially fused result and the historical fusion result. Its order parameter α determines the sensitivity of the entropy measure to the belief distribution structure; the Sigmoid slope parameter γ controls the response speed of the weight function to the entropy change rate; and the threshold parameter θ is used to set the critical position at which changes in current evidence uncertainty trigger weight adjustment. Together, these three parameters affect the dynamic weight allocation between current evidence and historical evidence. This section adopts the single-factor analysis method to examine the effects of α, γ, and θ on the temporal fusion results, respectively, and selects $\alpha = 3$, $\gamma = 6$, and $\theta = 0.1$ as the baseline parameters for subsequent experiments.

The experimental inputs in this section include two parts. The first part consists of four typical three-class belief inputs: high-belief evidence $m_h = [0.90, 0.05, 0.05]$, moderate-belief evidence $m_m = [0.7, 0.15, 0.15]$, ambiguous-belief evidence $m_a = [0.45, 0.30, 0.25]$, and average-belief evidence $m_{average} = [0.33, 0.33, 0.34]$, which are used to analyze the entropy response characteristics of α under different belief structures.

The second part uses the temporal evidence sequence for three-class UAV recognition over 150 frames, consistent with Section 4.2, as the input. The experimental outputs include Rényi entropy variation curves, temporal belief curves, recognition accuracy, and the mean belief of the true class under different parameter settings.

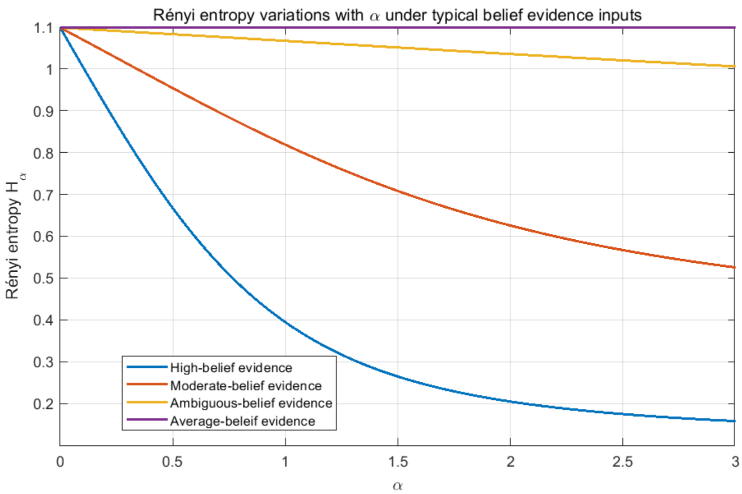
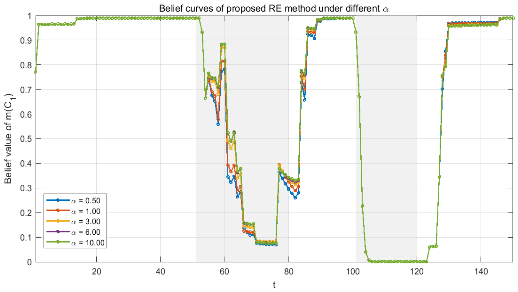


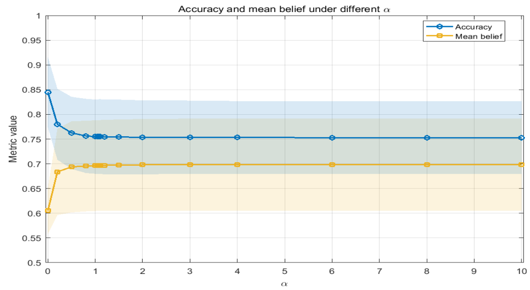
Figure 12: Rényi entropy variations with α under typical belief evidence inputs

As shown in Fig.12, as α increases, the Rényi entropy of the high-belief evidence decreases most markedly, that of the moderate-belief evidence decreases gradually, whereas the ambiguous-belief and average-belief evidence maintain relatively high entropy values. This result indicates that α can regulate the sensitivity of the entropy measure to the dominant-class belief, and a larger α helps enlarge the uncertainty difference between high-belief and ambiguous states.

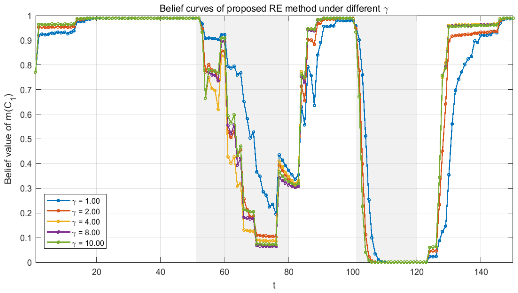
Furthermore, Fig.13 presents the statistical results and temporal belief curves under different settings of α, γ, and θ. As shown in Fig.13a and 13b, as α increases, the mean belief gradually improves and becomes nearly stable when $\alpha \geq 3$, whereas the recognition accuracy changes slightly. Therefore, $\alpha = 3$ is selected in this paper to balance entropy discrimination capability and temporal fusion stability. As shown in Fig.13c and 13d, when γ is small, the weight function changes slowly and insufficiently downweights anomalous inputs; as γ increases, the weight function drops rapidly near the threshold, thereby suppressing the influence of false inputs more promptly. The experimental results show that as γ increases from a small value to a moderate range, the recognition accuracy improves significantly and tends to stabilize when $\gamma = 6$. Therefore, $\gamma = 6$ is set in this paper to achieve good anomaly suppression capability and stable outputs. Fig.13e and 13f show the effect of the threshold θ on the recognition results. A smaller θ is more sensitive to entropy changes, while a larger θ tends to retain the current evidence. From the experimental results, within



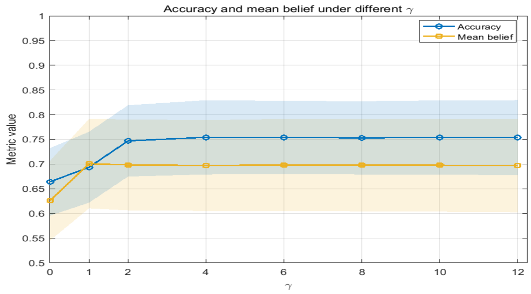
(a) Belief curves under different α



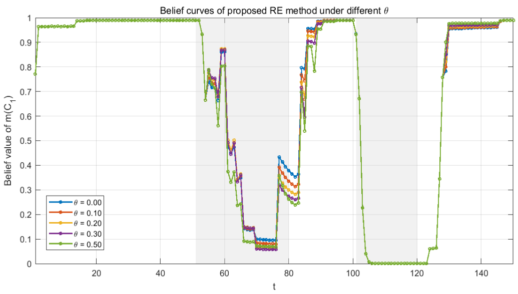
(b) Performance under different α



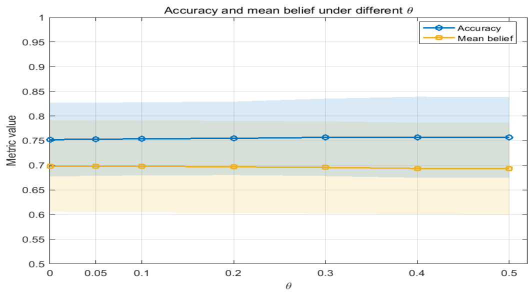
(c) Belief curves under different γ



(d) Performance under different γ



(e) Belief curves under different θ



(f) Performance under different θ

Figure 13: Parameter sensitivity analysis of the RE method

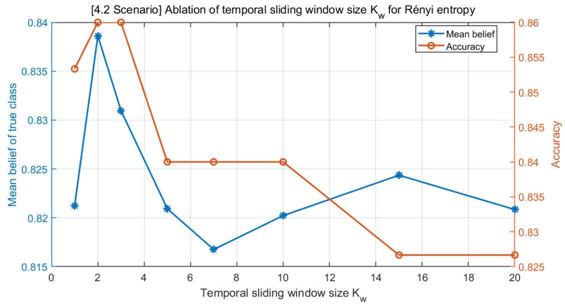
the range of $\theta \in [0, 0.5]$, the recognition accuracy and mean belief change slightly, indicating that the RE method exhibits good stability with respect to θ. Considering both false input suppression and current evidence utilization in normal intervals, this paper selects $\theta = 0.1$.

In summary, α, γ, and θ affect entropy sensitivity, weight variation speed, and the downweighting trigger position, respectively. The parameter sensitivity experiments show that the RE method can maintain stable performance over a wide parameter range, among which $\alpha = 3$, $\gamma = 6$, and $\theta = 0.1$ achieve a relatively balanced effect among false input suppression, mean belief improvement, and temporal response capability.

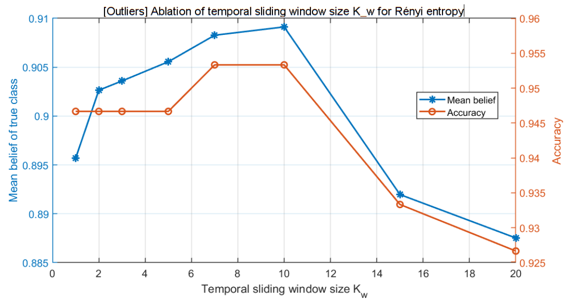
4.4.2. Sliding-window parameter K analysis

The sliding-window length K is used to construct the historical reference state for Rényi entropy, and its size directly affects the stability and responsiveness of temporal fusion. A smaller K enables a faster response to changes in current evidence, but provides weaker suppression of short-term anomalous inputs; a larger K can enhance historical smoothing, but may cause response lag. This paper tests the effect of $K = \{1, 2, 3, 5, 7, 10, 15, 20\}$ on fusion performance, using recognition accuracy and the mean belief of the true class as evaluation metrics.

Two input scenarios are designed in the experiment. The first scenario uses the temporal input sequence for three-class UAV recognition over 150 frames in Section 4.2, which includes class changes and false input intervals. The second scenario is a short-term outlier input scenario, where the true class remains c_1, while evidence for incorrect classes is introduced during frames 51–55, 76–80, and 101–105 to simulate short-term outlier interference.



(a) Section 4.2 evidence input



(b) Outliers evidence input

Figure 14: Ablation of sliding-window length K

As shown in Fig.14a, in the input scenario of Section 4.2, when $K = 3$, the method maintains high recognition accuracy while achieving a relatively high mean belief; as K further increases, the fusion result exhibits a certain response lag, leading to performance degradation. Fig.14b indicates that, in the short-term outlier scenario, better anomaly suppression is achieved when K is set to 7–10, whereas an excessively large window also leads to performance degradation.

In summary, K reflects the trade-off between responsiveness and stability. Considering that the main experiments in this paper place greater emphasis on the response to class changes in continuous recognition, while also requiring suppression of false input effects and improved computational speed, $K = 3$ is selected as the baseline setting. This value achieves a relatively balanced effect among accuracy, mean belief, and response lag.

4.4.3. Module ablation experiment

To further verify the effectiveness of the spatial ICEF mechanism and temporal RE mechanism in the STCEF framework, this paper conducts module ablation experiments to separately analyze the contributions of the spatial fusion module and temporal fusion module to recognition performance.

First, RB and PBLBJS are used as spatial baseline methods to examine the contribution of the ICEF mechanism under three temporal fusion strategies, namely TDS, TD, and RE. The results are shown in Table 13. It can be observed that, after introducing ICEF, both the mean belief and recognition accuracy of RB and PBLBJS are improved. Among them, PBLBJS shows the most significant improvement after ICEF enhancement, with the mean belief increasing by 19.67%–22.27% and the recognition accuracy increasing by 9.68%–16.63%. This indicates that the ICEF mechanism can effectively alleviate spatial evidence conflicts, improve the quality of spatial fusion results in a single frame, and provide more reliable inputs for subsequent temporal fusion.

Second, with the spatial fusion method fixed, TDS is used as the temporal fusion baseline to compare the performance changes of TD and RE. The results are shown in Table 14. RE improves both the mean belief and recognition accuracy under all spatial fusion methods, indicating that the dynamic weighting mechanism based on Rényi entropy can more effectively suppress temporal disturbances and maintain stable recognition.

In summary, the spatial ICEF mechanism mainly contributes to improving the credible quality of fused evidence in a single frame, whereas the temporal RE mechanism mainly contributes to enhancing stability and disturbance resistance during continuous recognition. By integrating these two mechanisms, STCEF achieves higher mean belief and recognition accuracy, verifying the complementary effects of spatial credibility optimization and temporal consistency correction.

4.4.4. Extreme interference scenario experiment

To further verify the robustness of the proposed STCEF method under complex conditions for low-altitude surveillance, this section constructs an extreme interference test scenario involving persistent occlusion and severe electromagnetic clutter. The experiment sets an evidence sequence over 90 frames, with the true target class set to c_1, and evaluates the fault tolerance of the method by introducing anomalous evidence in different stages. During frames 1–20 and 61–90, Sensors 1 and 2 output false evidence with high belief supporting c_2 to simulate the case

Table 13
Ablation experiment for spatial fusion modules

Spatial Baseline	Temporal Method	Δ Mean Belief	Δ Accuracy
RB → ICEF-RB	TDS	+0.0459 (+5.51%)	+0.0220 (+2.29%)
	TD	+0.0413 (+5.24%)	+0.0127 (+1.32%)
	RE	+0.0378 (+4.19%)	+0.0033 (+0.33%)
PBLBJS → ICEF-PBLBJS	TDS	+0.1538 (+20.90%)	+0.1313 (+15.48%)
	TD	+0.1565 (+22.27%)	+0.1407 (+16.63%)
	RE	+0.1560 (+19.67%)	+0.0880 (+9.68%)

Table 14
Ablation experiment for temporal fusion modules

Spatial Method	Temporal Method	Δ Mean Belief	Δ Accuracy
DS	w/ TD	-0.0284 (-3.29%)	+0.0087 (+0.89%)
	w/ RE	+0.0676 (+7.84%)	+0.0200 (+2.05%)
MAC	w/ TD	-0.0442 (-5.23%)	-0.0013 (-0.13%)
	w/ RE	+0.0686 (+8.12%)	+0.0200 (+2.05%)
RB	w/ TD	-0.0437 (-5.25%)	+0.0066 (+0.69%)
	w/ RE	+0.0701 (+8.42%)	+0.0353 (+3.68%)
PBLBJS	w/ TD	-0.0330 (-4.48%)	-0.0020 (-0.24%)
	w/ RE	+0.0572 (+7.77%)	+0.0607 (+7.16%)
ICEF-RB	w/ TD	-0.0483 (-5.50%)	-0.0027 (-0.28%)
	w/ RE	+0.0620 (+7.06%)	+0.0166 (+1.69%)
ICEF-PBLBJS	w/ TD	-0.0303 (-3.41%)	+0.0074 (+0.76%)
	w/ RE	+0.0594 (+6.68%)	+0.0174 (+1.78%)

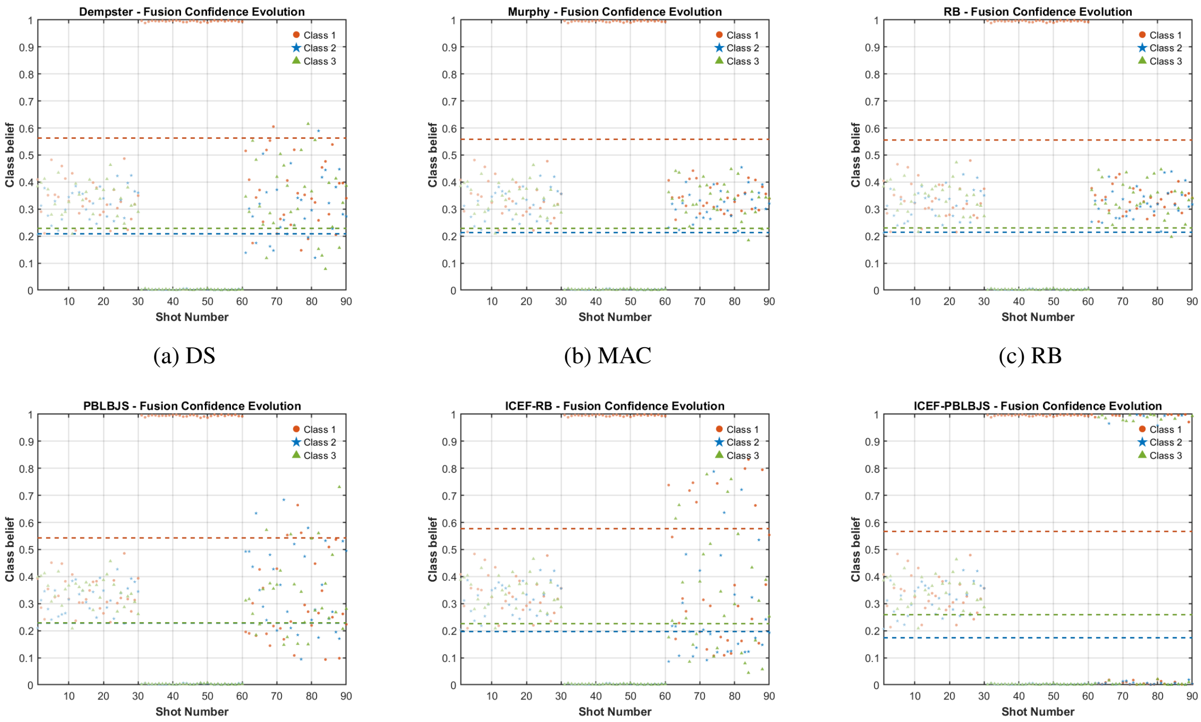
Table 15
Spatial fusion results under the extreme scenario

Method	Mean belief	Accuracy
DS	0.5628	**0.5778**
MAC	0.5583	**0.5778**
RB	0.5555	**0.5778**
PBLBJS	0.5428	0.5333
ICEF-RB	**0.5771**	**0.5778**
ICEF-PBLBJS	0.5670	0.5667

where multiple sensors are disturbed by false targets; during frames 21–30, all three sensors output random ambiguous evidence to simulate persistent occlusion; normal evidence is provided during the remaining periods.

The spatial fusion results are shown in Fig.15 and Table 15. Under extreme interference, all compared spatial methods remain susceptible to dominant high-belief false evidence, with recognition accuracy ranging from approximately 0.53 to 0.58. After ICEF is introduced, the mean beliefs of RB and PBLBJS increase by approximately 3.89% and 4.46%, respectively. The recognition accuracy of ICEF-PBLBJS increases by approximately 6.26% compared with PBLBJS, whereas RB and ICEF-RB obtain the same accuracy in this scenario. These results show that credible spatial fusion can mitigate part of the anomalous influence, but when most sensors simultaneously support an incorrect class, the correction capability of single-frame spatial fusion remains limited.

The temporal fusion results are shown in Fig.16 and Table 16. The performance of conventional temporal methods, including TDS and TD, is still constrained by the quality of the preceding spatial fusion results. Under the effects of early incorrect evidence and strongly conflicting inputs, DEF-NNs is susceptible to the accumulation of historical errors and shows obvious degradation in some tests. In contrast, the weighting method based on Rényi entropy maintains



(a) DS (b) MAC (c) RB

(d) PBLBJS (e) ICEF-RB (f) ICEF-PBLBJS

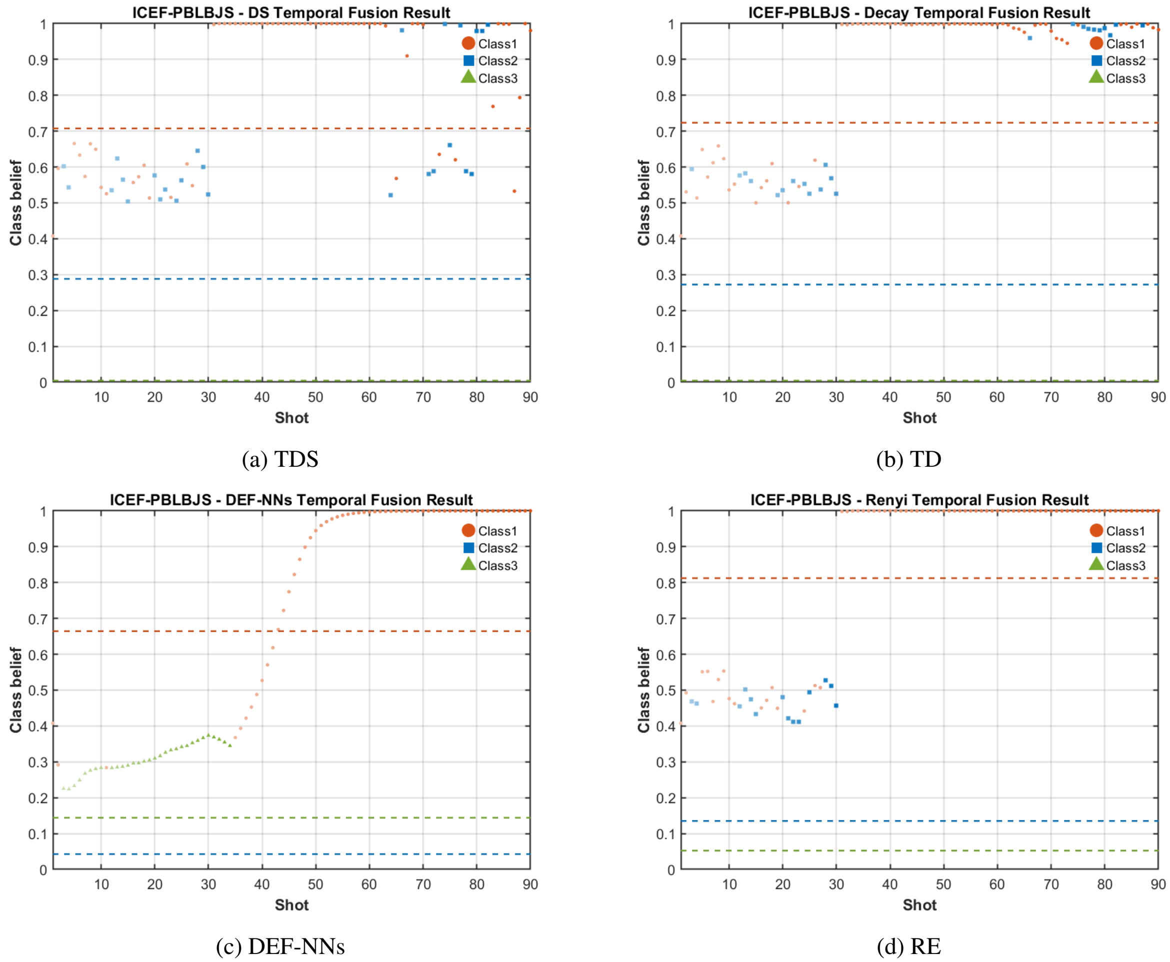
Figure 15: Spatial fusion results of different methods under the extreme interference scenario

Table 16
Temporal fusion results under the extreme scenario

Method	TDS		TD		DEF-NNs		RE	
	Mean Belief	Accuracy	Mean Belief	Accuracy	Mean Belief	Accuracy	Mean Belief	Accuracy
DS	0.6908	0.7444	0.6895	0.7556	0.6559	0.6556	0.8121	0.8444
MAC	0.6833	0.7333	0.6816	0.7333	0.6671	0.6667	0.8122	0.8444
RB	0.6796	0.6889	0.6778	0.7000	0.1130	0.0222	0.8118	0.8444
PBLBJS	0.6583	0.6778	0.6555	0.6778	0.0816	0.0000	0.8106	0.8444
ICEF-RB	0.7098	0.7667	0.7049	0.9556	0.6416	0.6333	0.8123	0.8444
ICEF-PBLBJS	0.7077	0.7111	0.7236	0.7444	0.6644	0.6556	**0.8123**	**0.8444**

stable outputs under different spatial inputs, with the mean belief remaining around 0.81 and the accuracy reaching 0.8444 in all cases. Based on the average results of the six groups of spatial inputs, compared with TDS and TD, the Rényi method improves the mean belief by approximately 17.96% and 17.87%, respectively, and improves the accuracy by approximately 17.22% and 10.94%, respectively. When ICEF-PBLBJS is used as the spatial input, the weighting method based on Rényi entropy improves the mean belief by approximately 14.78%, 12.26%, and 22.26% compared with TDS, TD, and DEF-NNs, respectively, and improves the accuracy by approximately 18.75%, 13.43%, and 28.80%, respectively. As shown by the Rényi weight variation curve in Fig.17, when the input enters the random ambiguous stage or the stage with false inputs from most sensors, the weight curve shows obvious adjustments, demonstrating that the method can effectively measure the uncertainty difference between current evidence and historical fusion results.

In summary, under strong anomalous input conditions, the STCEF method significantly enhances the fault tolerance and overall robustness of the fusion system through the synergistic effect of spatial credibility correction and temporal Rényi entropy weighting.



(a) TDS (b) TD

(c) DEF-NNs (d) RE

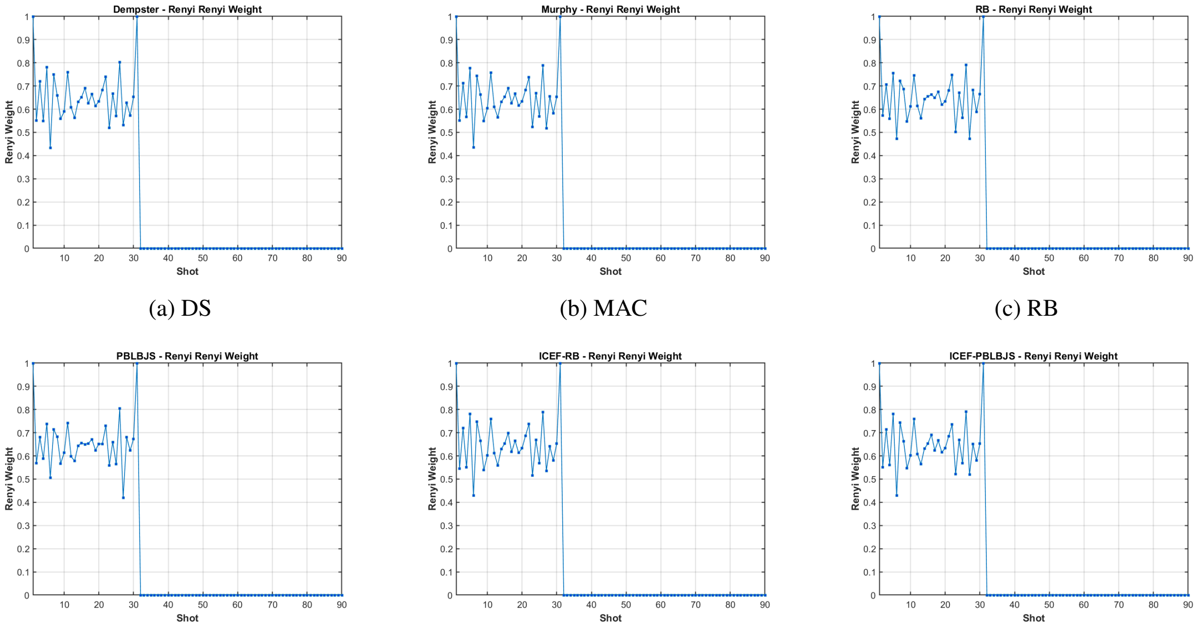
Figure 16: Results of ICEF-PBLBJS with different temporal fusion methods

5. Conclusion

This paper develops a spatio-temporal credible evidence fusion framework for urban low-altitude UAV type recognition under large-scale and inconsistent decision-level evidence. Simulation results demonstrate that the fusion-center-based conditional credibility fusion mechanism spatially achieves high computational efficiency while maintaining reliable recognition accuracy under multiple evidence inputs. Meanwhile, the sliding-window Rényi entropy temporally preserves recognition stability and accuracy under dynamic interference conditions. In conclusion, the STCEF method effectively balances fusion efficiency, information utilization, and recognition stability, thereby providing support for low-altitude security monitoring and decision-making.

However, STCEF still has certain limitations in its applicability. The spatial fusion stage relies on a fusion center to represent overall consistency. If the majority of sensors are simultaneously affected by strong interference, occlusion, spoofing, or association errors and output high-belief incorrect evidence, the fusion center may deviate from the true class, thereby weakening the effectiveness of credibility weight correction. The temporal fusion stage utilizes historical information to suppress short-term fluctuations, but this may cause a response lag in scenarios such as rapid changes in target type, track handovers, or reappearance after occlusion. Furthermore, parameters including the sliding-window length K, Rényi entropy order α, Sigmoid slope γ, and threshold θ affect the balance between temporal stability and responsiveness. Therefore, adaptive parameter optimization will need to be further explored in subsequent research.

In summary, STCEF can serve as a decision-level reliable recognition module for urban low-altitude security surveillance. It provides support for early warning of illegal UAVs, target type confirmation, and anomalous evidence



(a) DS (b) MAC (c) RB

(d) PBLBJS (e) ICEF-RB (f) ICEF-PBLBJS

Figure 17: Weight variation curves of the Rényi method under the interference scenario

indication near airport perimeters, urban core areas, major event zones, and critical infrastructure. In UAV traffic management scenarios, this method can also incorporate GNSS, acoustic, and other auxiliary information sources. By combining these with track monitoring and identification modules, it provides an interpretable fusion basis for non-cooperative target recognition, the detection of sudden changes in recognition status, and low-altitude situational awareness. Future work will focus on sensor availability modeling, the processing of highly conflicting evidence, and validation using real measurement data covering multiple target types, ultimately improving the method's engineering applicability in complex low-altitude environments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant No. 62573357.

The authors would like to thank the anonymous reviewers for their detailed comments which helped to improve the quality of the paper.

References

[1] Yinhan Wang, Jiang Wang, Shipeng Fan, and Yuchen Wang. Quick intention identification of an enemy aerial target through information classification processing. *Aerospace Science and Technology*, 132:108005, 2023.

[2] Shuaijun Wang, Fan Jiang, Bin Zhang, Rui Ma, and Qi Hao. Development of uav-based target tracking and recognition systems. *IEEE Transactions on Intelligent Transportation Systems*, 21(8):3409–3422, 2020.

[3] X Liang, Y Qi, G Chen, and G Meng. Research on a distributed strategy for uav detection and tracking of specific targets. *Artificial Intelligence Review*, 58(12):388, 2025.

[4] Federica Vitiello, Flavia Causa, Roberto Opromolla, and Giancarmine Fasano. Radar/visual fusion with fuse-before-track strategy for low altitude non-cooperative sense and avoid. *Aerospace Science and Technology*, 146:108946, 2024.

[5] Jiale Wang, Chuang Shi, Fu Zheng, Cheng Yang, Xu Liu, Shuo Liu, Ming Xia, Guifei Jing, Tuan Li, Wu Chen, et al. Multi-frequency smartphone positioning performance evaluation: insights into a-gnss ppp-b2b services and beyond. *Satellite Navigation*, 5(1):25, 2024.

[6] Hongzhi Wu, Miao wang, Jingshi Wang, and Guoqing wang. Distributed information fusion based trajectory tracking for usv and uav clusters via multi-agent deep learning approach. *Aerospace Systems*, 7(2):193–207, 2024.
[7] Jianfeng Ren and Xudong Jiang. A three-step classification framework to handle complex data distribution for radar uav detection. *Pattern Recognition*, 111:107709, 2021.
[8] Mohammadreza Doostmohammadian, Amin Taghieh, and Houman Zarrabi. Distributed estimation approach for tracking a mobile target via formation of uavs. *IEEE Transactions on Automation Science and Engineering*, 19(4):3765–3776, 2022.
[9] Lingjiao Fu, Yifang Shi, Dongliang Peng, and Ihsan Ullah. Unified framework for multi-sensor distributed fusion with memory configuration. *Aerospace Science and Technology*, 150:109184, 2024.
[10] Shimin Gong, Meng Wang, Bo Gu, Wenjie Zhang, Dinh Thai Hoang, and Dusit Niyato. Bayesian optimization enhanced deep reinforcement learning for trajectory planning and network formation in multi-uav networks. *IEEE Transactions on Vehicular Technology*, 72(8):10933–10948, 2023.
[11] Farhad Samadzadegan, Farnaz Ashtari Mahini, and Mehrnaz Gholamshahi. A vision-based explainable deep learning approach for multi-class drone detection and recognition. *Aerospace Science and Technology*, 171:111648, 2026.
[12] Zirui Liao, Shaoping Wang, Jian Shi, Zhiyong Sun, Yuwei Zhang, and Muhammad Baber Sial. Cooperative situational awareness of multi-uav system based on improved d-s evidence theory. *Aerospace Science and Technology*, 142:108605, 2023.
[13] Naixin Chen, Lei Chen, Zhen Gao, and Chunli Zhu. Causal inference-enhanced uav detection and identification for low-altitude air city transport. *IEEE Transactions on Intelligent Transportation Systems*, 26(12):22690–22703, 2025.
[14] Qinbo Sun and Zhaohui Dang. Deep neural network for non-cooperative space target intention recognition. *Aerospace Science and Technology*, 142:108681, 2023.
[15] Guidong He, Peng Lei, and Jun Wang. Outlier-robust feature selection and classification for low-altitude uavs and ground vehicles based on radar micro-doppler signatures. *IEEE Transactions on Aerospace and Electronic Systems*, 61(6):18383–18395, 2025.
[16] Qianlei Jia, Jiayue Hu, and Weiguo Zhang. A fault detection method for fads system based on interval-valued neutrosophic sets, belief rule base, and d-s evidence reasoning. *Aerospace Science and Technology*, 114:106758, 2021.
[17] Xiaojiao Geng, Yan Liang, and Lianmeng Jiao. Multi-frame decision fusion based on evidential association rule mining for target identification. *Applied Soft Computing*, 94:106460, 2020.
[18] Zhun-Ga Liu, Quan Pan, Jean Dezert, and Arnaud Martin. Combination of classifiers with optimal weight based on evidential reasoning. *IEEE Transactions on Fuzzy Systems*, 26(3):1217–1230, 2017.
[19] Zhen Zhang, Hong peng Tian, and Jin shuai Jin. Multiple adaptive over-sampling for imbalanced data evidential classification. *Engineering Applications of Artificial Intelligence*, 133:108532, 2024.
[20] Xin Guan, Haotian Yu, and Xiao Yi. Optimized methods for basic probability assignments in evidence theory: Applications to fault diagnosis. *Engineering Applications of Artificial Intelligence*, 160:111737, 2025.
[21] Yong Deng, Felix T.S. Chan, Ying Wu, and Dong Wang. A new linguistic mcdm method based on multiple-criterion data fusion. *Expert Systems with Applications*, 38(6):6985–6993, 2011.
[22] Catherine K. Murphy. Combining belief functions when evidence conflicts. *Decision Support Systems*, 29(1):1–9, 2000.
[23] Xinyang Deng, Yang Yang, and Jihao Yang. A novel discrete evidence fusion approach by considering the consistency of belief structures. *Engineering Applications of Artificial Intelligence*, 96:103994, 2020.
[24] Pengfei Zhang, Qinli Zhang, Jingxin Liu, Dexian Wang, Hua Ye, Xiabing Zhang, and Tianrui Li. Information fusion and feature selection for multi-source data utilizing dempster-shafer evidence theory and k-nearest neighbors. *Information Sciences*, 718:122408, 2025.
[25] Sangsoo Jang, Chanho Kim, Hansol Nam, Dohoon Kim, Deokgeon Kim, Kyungtaek Lee, and Kyu Hong Kim. Demand-driven standardization for multirotor uavs. *Aerospace Science and Technology*, 168:111223, 2026.
[26] Yihan Cui, Yan Liang, Jie Shi, Chaoxiong Ma, Huixia Zhang, and Shupan Li. An interpretable hierarchical framework for rapid uav type recognition based on multi-sensor data and knowledge. *IEEE Sensors Journal*, 25(12):22239–22249, 2025.
[27] Zhun ga Liu, Jean Dezert, Quan Pan, and Grégoire Mercier. Combination of sources of evidence with different discounting factors based on a new dissimilarity measure. *Decision Support Systems*, 52(1):133–141, 2011.
[28] Fuyuan Xiao. A new divergence measure for belief functions in d–s evidence theory for multisensor data fusion. *Information Sciences*, 514:462–483, 2020.
[29] Hongfei Wang, Xinyang Deng, Wen Jiang, and Jie Geng. A new belief divergence measure for dempster–shafer theory based on belief and plausibility function and its application in multi-source data fusion. *Engineering Applications of Artificial Intelligence*, 97:104030, 2021.
[30] Yang Zhang, Ao Xiong, Yu Xiao, and Ziyang Chen. A new combination method based on pearson coefficient and information entropy for multi-sensor data fusion. *Information and Software Technology*, 161:107248, 2023.
[31] Shuai-Wen Tang, Zhi-Jie Zhou, Guan-Yu Hu, You Cao, Peng-Yun Ning, and Jie Wang. Evidential reasoning rule with likelihood analysis and perturbation analysis. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(2):1209–1221, 2022.
[32] Xuan Zhao, Mengyi Zhang, Zilong Xiao, and Bingyi Kang. Evaluating the reliability and relative weight of the evidence using approximate evidential mutual information. *Engineering Applications of Artificial Intelligence*, 133:108409, 2024.
[33] Yafei Song, Xiaodan Wang, Lei Lei, and Yaqiong Xing. Credibility decay model in temporal evidence combination. *Information Processing Letters*, 115(2):248–252, 2015.
[34] Cheng li Fan, Yafei Song, Lei Lei, Xiaodan Wang, and Shi Bai. Evidence reasoning for temporal uncertain information based on relative reliability evaluation. *Expert Systems with Applications*, 113:264–276, 2018.
[35] Weiping Ding, Tao Hou, Jiashuang Huang, Hengrong Ju, and Shu Jiang. Dynamic evidence fusion neural networks with uncertainty theory and its application in brain network analysis. *Information Sciences*, 691:121622, 2025.
[36] Zixuan Zhou and Fuyuan Xiao. Conflict management in sequential evidence combination. *Information Sciences*, 734:122958, 2026.
[37] Jianhua Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information theory*, 37(1):145–151, 2002.

[38] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum renyi entropies, a new generalization and some properties. *Journal of Mathematical Physics*, 54(12), 2013.

[39] Fuyuan Xiao. Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy. *Information Fusion*, 46:23–32, 2019.