EMBRACING ALEATORIC AND EPISTEMIC UNCERTAINTY FOR TRUSTED OPEN-WORLD MULTI-VIEW CLASSIFICATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-view classification fuses heterogeneous information to achieve more comprehensive data descriptions. To deliver trustworthy decisions, trusted multi-view classification (TMVC) has emerged as a crucial research paradigm. However, existing TMVC methods assume closed-set environments and inevitably fail when novel categories emerge in open-world settings. This stems from two key limitations: they neglect data-inherent aleatoric uncertainty, leaving cross-view alignment vulnerable to noise; and their fusion strategies lack epistemic coherence, producing overconfident and unreliable predictions that are catastrophically amplified for unknown samples. To address these, we propose a novel framework jointly Embracing aleatoric and epistemic Uncertainty for trusted Open-world multi-view classification (EU^{2}O). Specifically, we quantify view-specific aleatoric uncertainty and align cross-view representations to obtain a robust joint representation, which serves as a reliable reference for subsequent opinion aggregation. Besides, we propose an Uncertainty-Driven Proportional Conflict Redistribution (UD-PCR) operator that aggregates multi-view subjective opinions by adaptively redistributing conflict mass in proportion to calibrates belief masses and uncertainty. We also theoretically prove its effectiveness, particularly in open-world scenarios. A novel Jousselme consistency loss further enforce cross-view epistemic alignment at a linear complexity. Experiments on multiple benchmarks demonstrate superior closed-set accuracy and open-world rejection.

1 INTRODUCTION

Multi-view data Liu et al. (2026), collected from diverse sensors, modalities, or feature representations, is ubiquitous in real-world applications. Multi-view learning Li et al. (2026) aims to exploit complementary and consistent information across heterogeneous views and has achieved remarkable progress in classification Liu et al. (2025), clustering Wei et al. (2026); Dong et al. (2025a), and retrieval Tian et al. (2025). However, existing methods Wang et al. (2025c); Wen et al. (2026) mainly emphasize predictive accuracy while overlooking the uncertainty inherent in their decisions, limiting their deployment in safety-critical scenarios. To address this issue, Han et al. Han et al. (2022) introduced trusted multi-view classification (TMVC), which models prediction uncertainty through Dirichlet distributions and fuses view-specific opinions via Dempster–Shafer theory Dempster (1968). Subsequent works further improved opinion aggregation through evidence accumulation Liu et al. (2022) and consistent-complementary evidence decomposition Liu et al. (2024).

Despite their initial success, existing TMVC methods rely on the closed-set assumption that all categories are known in advance Wang et al. (2025a); Shen et al. (2025), which is often violated in real-world applications such as autonomous driving, medical diagnosis, and industrial inspection Cheng et al. (2024); Wang et al. (2024); Niu et al. (2024); Jiang et al. (2025). In open-world scenarios, these methods tend to assign unknown samples to known classes with overconfident predictions and unreliable uncertainty estimates Huang et al. (2025). This limitation arises from two aspects: 1) view-specific features contain inherent aleatoric uncertainty Zhang et al. (2024), which introduces unreliable evidence and compromises cross-view fusion; and 2) Dempster–Shafer-based aggregation lacks epistemic coherence across views and can only preserve or reduce uncertainty as evidence accumulates, potentially suppressing the uncertainty signals required for unknown detec-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

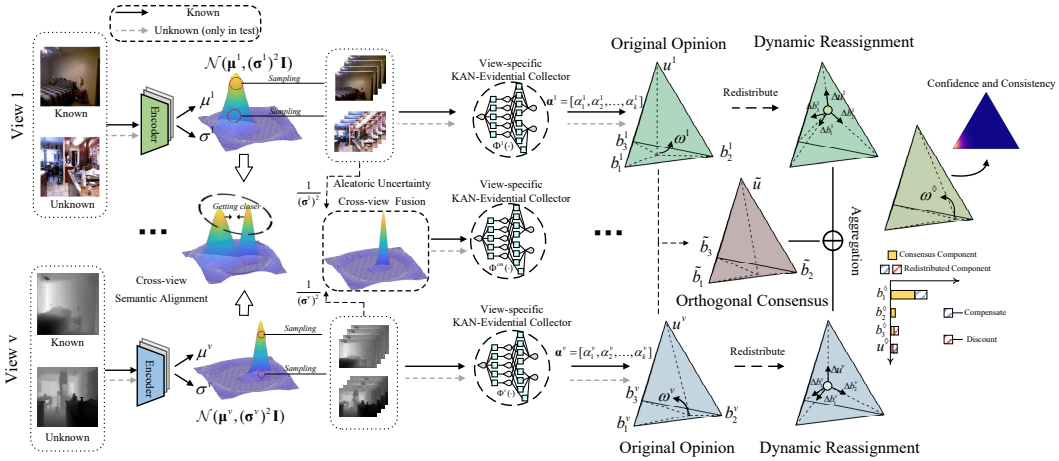


Figure 1: Overview of the proposed EU^{2}O framework. Each view is first modeled as a Gaussian distribution to capture aleatoric uncertainty, and the EU^{2}O aligns and fuses these distributions into a robust joint representation. KAN-Evidential networks then collect evidence from both view-specific and joint representations to construct subjective opinions. Finally, UD-PCR aggregates all opinions by decomposing orthogonal consensus and adaptively redistributing conflict mass in proportion to epistemic uncertainty, producing the final belief masses and uncertainty.

tion (see Appendix A.1 for theoretical analysis). Therefore, a unified framework that jointly models aleatoric uncertainty and preserves epistemic uncertainty is essential for trusted open-world multi-view classification.

In this paper, we propose a novel framework jointly embracing aleatoric and epistemic uncertainty for trusted open-world multi-view classification as shown in Fig. 1. Specifically, to quantify view-specific aleatoric uncertainty, each view's feature is modeled as a multivariate Gaussian distribution. We leverage an information-theoretic strategy to facilitate cross-view alignment and learn a robust compact joint representation based on aleatoric uncertainty as a reference for opinion aggregation. After that, KAN-Evidential networks collect view-specific evidence from these representations, and subjective opinions are constructed under subjective logic Jøsang (2016) to assess the epistemic uncertainty of each view. An Uncertainty-Driven Proportional Conflict Redistribution (UD-PCR) strategy aggregates multi-view opinions by dynamically calibrating belief masses and uncertainty according to view-level reliability and epistemic coherence degree. We also theoretically prove its effectiveness, particularly in open-world scenarios. Besides, a novel Jousselme consistency loss further enforces cross-view epistemic alignment at linear complexity. Extensive experiments on widely adopted multi-view and multi-modal benchmarks demonstrate superior closed-set accuracy and open-world rejection over state-of-the-art trusted multi-view methods. The main contributions of EU^{2}O can be listed as follows:

- In this paper, we investigate the rarely explored yet critical open-world problem in multi-view classification, and propose a unified framework jointly embracing aleatoric and epistemic uncertainty (EU^{2}O) for trusted decision-making under open-world settings.
- We propose an Aleatoric Uncertainty-Driven Cross-view Fusion (AU-DCF) module that facilitates distribution alignment across views and learns a robust joint representation as the reference for opinion aggregation. A novel UD-PCR strategy is further proposed to calibrate belief masses and uncertainty mass, and is theoretically established to be effective particularly in open-world scenarios.
- A novel Jousselme consistency loss is proposed to enforce cross-view epistemic consistency at linear complexity. Extensive experiments on multi-view and multi-modal benchmarks validate the effectiveness of the proposed framework.

2 Related Work

2.1 Trusted Multi-View Classification

Traditional multi-view classification methods Lin et al. (2025); Wang et al. (2025b) mainly focus on improving decision accuracy while ignoring prediction uncertainty. Han et al. Han et al. (2022) introduced trusted multi-view classification (TMVC) by modeling class probabilities with Dirichlet distributions and aggregating view-specific opinions via Dempster–Shafer theory. Subsequent works explored evidence accumulation Liu et al. (2022), conflict-aware opinion fusion Xu et al. (2024), consistency-complementarity decoupling Liu et al. (2024), and hierarchical opinion aggregation Shi et al. (2026) to improve uncertainty-aware multi-view fusion. However, existing TMVC methods mainly rely on closed-set assumptions and face two limitations in open-world scenarios. First, view-specific representations are typically modeled as deterministic embeddings, overlooking view-dependent aleatoric uncertainty. Second, Dempster–Shafer-based fusion mechanisms tend to monotonically reduce uncertainty when aggregating multiple opinions, which is beneficial for known samples but may suppress the epistemic uncertainty required for unknown detection.

2.2 Open-World Multi-View Learning

Open-set recognition aims to identify unknown categories while maintaining known-class classification performance, with representative solutions based on calibration Bendale & Boult (2016), generative modeling Neal et al. (2018), and distance-based approaches Oza & Patel (2019). Recent multi-view extensions include OpenViewer Du et al. (2025), which introduces openness-aware unknown generation, and TOMC Dong et al. (2025b), which formulates open multi-view classification and detects unknowns through dissonance modeling. Nevertheless, existing open-world multi-view methods mainly handle uncertainty at the decision level and lack explicit modeling of representation-level aleatoric uncertainty and epistemic coherence. Consequently, unreliable view representations and inconsistent evidence aggregation remain challenging, motivating a unified uncertainty-aware framework for trusted open-world multi-view classification.

3 Methodology

3.1 Problem Formulation

In open-world multi-view classification, the training set comprises V views from K known categories. Each instance of the v-th view is denoted as $\mathbf{x}^v \in \mathbb{R}^{d_v}$, where d_v is the feature dimensionality of the v-th view. Labels for known categories are encoded as one-hot vectors $\mathbf{y} \in \{0,1\}^K$. During testing, the classifier encounters $\tilde{N}$ samples from both known and previously unseen categories. The objectives are twofold:

1) To accurately classify samples from known categories.

2) To assign high uncertainty to samples from unknown categories for reliable rejection.

3.2 Aleatoric Uncertainty-Driven Cross-view Fusion

Existing TMVC methods use deterministic point embeddings for view-specific features, overlooking inherent data-acquisition noise and view-dependent aleatoric uncertainty. Consequently, cross-view alignment often operates on degraded features without view-quality awareness. To address this, we model each view's representation as a multivariate Gaussian distribution, where the mean encodes semantic content and the variance captures view-specific aleatoric uncertainty. Formally, given an instance $\mathbf{x}^v$ from the v-th view,

$$\begin{aligned} p(\mathbf{z}^v \mid \mathbf{x}^v) &\sim \mathcal{N}\Big(\boldsymbol{\mu}^v, {\boldsymbol{\sigma}^v}^2 \mathbf{I}\Big), \\ \boldsymbol{\mu}^v = f_v^{\mu}(\mathbf{x}^v), &\quad \boldsymbol{\sigma}^v = f_v^{\sigma}(\mathbf{x}^v), \\ \mathbf{z}^v = \boldsymbol{\mu}^v + \boldsymbol{\epsilon} \odot \boldsymbol{\sigma}^v, &\quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \end{aligned} \tag{1}$$

where f_v^{μ} and f_v^{σ} are view-specific fully connected layers, $\mathbf{I}$ is the identity matrix, and $\odot$ denotes element-wise multiplication. The variance $\boldsymbol{\sigma}^v$ reflects the aleatoric uncertainty of the v-th view, while $\mathbf{z}^v$ is the sampled latent representation used in subsequent alignment and fusion.

The view-specific Gaussian distributions reside in heterogeneous feature spaces and may encode semantic gap. Directly fusing these distributions without alignment introduces specific-view noise into subsequent opinion aggregation. To obtain semantically coherent representations, we adopt an information bottleneck principle where each view's representation should preserve information consistent across views while discarding view-specific redundancies. This is achieved by:

$$\mathcal{L}_{\text{vib}} = \sum_{v_1 \neq v_2} \mathrm{KL}(p(\mathbf{z}^{v_1} \mid \mathbf{x}^{v_1}) \parallel p(\mathbf{z}^{v_2} \mid \mathbf{x}^{v_2})) + \sum_{v=1}^{V} \mathrm{KL}(p(\mathbf{z}^{v} \mid \mathbf{x}^{v}) \parallel \mathcal{N}(\mathbf{0}, \mathbf{I}))\,. \tag{2}$$

The first term maximizes cross-view mutual information by pulling distributions of the same instance closer, preserving semantically consistent information. The second term acts as a compression bottleneck, constraining each representation toward a standard Gaussian prior to suppress view-specific noise and prevent variance collapse.

After cross-view alignment, the view-specific distributions share consistent semantics but still exhibit varying view quality. To obtain a robust joint representation, we fuse the sampled representations via an aleatoric uncertainty-driven weighting strategy, as follows:

$$\hat{\mathbf{z}} = \sum_{v=1}^{V} w^v \mathbf{z}^v, \quad w^v = \frac{\exp\left(-\frac{1}{d}\|\boldsymbol{\sigma}^v\|_1\right)}{\sum_{u=1}^{V} \exp\left(-\frac{1}{d}\|\boldsymbol{\sigma}^u\|_1\right)}. \tag{3}$$

Views with higher aleatoric uncertainty receive lower weights, suppressing their contribution to $\hat{\mathbf{z}}$. The joint representation also serves as a reliable reference for subsequent opinion aggregation.

3.3 View-Specific Opinion Representation

3.3.1 KAN-Evidential Evidence Collection

Conventional trusted multi-view classification employs MLP-based evidence collectors to map representations to class evidence. However, MLPs rely on fixed activation functions, which may insufficiently capture the complex mapping from heterogeneous view representations to evidence scores. To enhance expressiveness, we adopt Kolmogorov–Arnold Networks (KAN) as the evidence collector. Formally, the view-specific representation $\mathbf{z}^v$ and the cross-view joint representation $\hat{\mathbf{z}}$ are each mapped to evidence vectors via dedicated KAN-evidential collectors:

$$\mathbf{e}^v = \Phi^v(\mathbf{z}^v), \qquad \mathbf{e}^{\text{cross}} = \Phi^{\text{cross}}(\hat{\mathbf{z}}), \tag{4}$$

where Φ^v and Φ^{cross} denote view-specific and cross-view evidence collectors, respectively. Each output $\mathbf{e} \in \mathbb{R}^K_{\geq 0}$ is enforced to be non-negative by a Softplus activation at the output layer, and each entry e_k quantifies the amount of support for the k-th category.

3.3.2 Subjective Logic and Opinion Construction

Having obtained view-specific evidence, a principled mechanism is required to jointly model class probabilities and epistemic uncertainty. Subjective logic provides a theoretical framework that associates evidence with a Dirichlet distribution, enabling second-order uncertainty quantification beyond point estimates. The probability density function of the Dirichlet distribution over the K-dimensional simplex is given by

$$\mathrm{Dir}(\mathbf{p} \mid \boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{k=1}^{K} \alpha_k\right)}{\prod_{k=1}^{K} \Gamma(\alpha_k)} \prod_{k=1}^{K} p_k^{\alpha_k - 1}, \tag{5}$$

where $\mathbf{p} = (p_1, \ldots, p_K)$ is a probability vector on the K-dimensional unit simplex $\mathcal{S}_K$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$ are the concentration parameters, and $\Gamma(\cdot)$ is the gamma function. For the v-th view, the evidence $\mathbf{e}^v$ parameterizes the Dirichlet distribution as follows:

$$\boldsymbol{\alpha}^v = \mathbf{e}^v + \mathbf{1}, \tag{6}$$

where adding $\mathbf{1}$ ensures the distribution is non-sparse. The Dirichlet strength is $S^v = \sum_{k=1}^{K} \alpha_k^v$, reflecting the total evidence observed from the v-th view.

According to subjective logic, the Dirichlet distribution is mapped to a multinomial opinion $\omega^v = (\mathbf{b}^v, u^v, \mathbf{a}^v)$, where $\mathbf{b}^v = [b_1^v, \ldots, b_K^v]$ is the belief mass vector assigned to each class, u^v is the uncertainty mass, and $\mathbf{a}^v$ is the base rate encoding prior class preferences. The mapping is given by

$$b_k^v = \frac{e_k^v}{S^v}, \qquad u^v = \frac{K}{S^v}, \qquad a_k^v = \frac{1}{K}, \tag{7}$$

and subjective logic dictates that both b and u must be non-negative and their sum should equal one:

$$\sum_{k=1}^{K} b_k^v + u^v = 1, \quad b_k^v \geq 0, \ \forall k \in \{1, \ldots, K\}, \quad u^v \geq 0. \tag{8}$$

Eq. 7 captures the inverse relationship between evidence and uncertainty: as more evidence is collected from the v-th view, S^v increases and u^v decreases; conversely, limited or conflicting evidence yields high u^v, signaling that the opinion is unreliable. The base rate is set to a uniform prior, indicating no class preference before evidence is observed. The projected probability of class k incorporates both belief and prior-weighted uncertainty as $p_k^v = b_k^v + a_k^v u^v$. To this end, each view yields a subjective opinion that explicitly quantifies its epistemic uncertainty, forming the basis for the reliable fusion stage.

3.4 Uncertainty-Driven Proportional Conflict Redistribution

The Dempster–Shafer combination rule and its variants adopt in existing TMVC unconditionally reduce uncertainty as more opinions are aggregated, conflating the uncertainty mass from conflicting views with that from consistent ones. Besides, they treat all views equally without leveraging epistemic coherence to guide the fusion process. To this end, we propose the Uncertainty-Driven Proportional Conflict Redistribution (UD-PCR), which dynamically redistributes the generated conflict mass into class-specific beliefs and overall uncertainty in proportion to view epistemic uncertainty.

Definition 1 (UD-PCR Fusion)**.** *Let $\omega^A = (\mathbf{b}^A, u^A)$ and $\omega^B = (\mathbf{b}^B, u^B)$ be subjective opinions of two views for a given instance, where the base rate $\mathbf{a}$ is fixed to a uniform prior and omitted for brevity. The UD-PCR fused opinion $\omega^{A\Diamond B} = (\mathbf{b}^{A\Diamond B}, u^{A\Diamond B})$ is defined as:*

$$b_k^{A\Diamond B} = \tilde{b}_k + \Delta b_k^A + \Delta b_k^B, \qquad u^{A\Diamond B} = \tilde{u} + \Delta u. \tag{9}$$

where the orthogonal consensus terms are defined as follows:

$$\tilde{b}_k = b_k^A b_k^B + b_k^A u^B + b_k^B u^A, \qquad \tilde{u} = u^A u^B. \tag{10}$$

The orthogonal consensus $\{\tilde{b}_k, \tilde{u}\}$ corresponds to the portion of beliefs that can be directly combined via Dempster's combination rule, representing the information on which the two views do not conflict. The residual conflict mass is quantified as:

$$C = \sum_{i \neq j} b_i^A b_j^B. \tag{11}$$

This conflict mass C captures the portion of cross-view belief that supports divergent classes, and is dynamically redistributed in proportion to each view's epistemic uncertainty:

$$\Delta b_k^A = b_k^A \Big(\sum_{j \neq k} b_j^B\Big) \frac{(1-u^A)u^B}{Z}, \qquad \Delta b_k^B = b_k^B \Big(\sum_{i \neq k} b_i^A\Big) \frac{(1-u^B)u^A}{Z}, \qquad \Delta u = C \frac{(1-u^A)(1-u^B)}{Z}. \tag{12}$$

where the normalization factor is $Z = 1 - u^A u^B$.

The redistribution mechanism in Eq. 12 reflects the following principle. For view A, the term $\sum_{j \neq k} b_j^B$ measures the exact degree to which view B contradicts class k. The product $b_k^A \big(\sum_{j \neq k} b_j^B\big)$ thus represents the specific conflict mass generated between view A's support for class k and view B's support for divergent classes. This conflicting portion is then scaled by view A's confidence $(1-u^A)$ and view B's uncertainty u^B. When view B is highly uncertain (large u^B), the mechanism penalizes its unreliability and redistributes the conflict mass predominantly back to view A's beliefs. Conversely, when both views are confident and consistent, the conflict $C \to 0$ and the orthogonal consensus dominates the aggregated opinion.

Definition 2 (Extension to V Views)**.** *Given V view-specific opinions $\{\omega^v\}_{v=1}^V$ and the cross-view opinion ω^{cross} derived from $\hat{\mathbf{z}}$, the final fused opinion is obtained by sequential UD-PCR aggregation. Views are strictly ordered by decreasing epistemic uncertainty, i.e., ascending Dirichlet strength S^v:*

$$\omega^{\diamondsuit} = \omega^{\sigma(1)} \oplus \omega^{\sigma(2)} \oplus \cdots \oplus \omega^{\sigma(V)} \oplus \omega^{cross}, \tag{13}$$

where σ is the permutation that sorts $\{S^v\}$ in ascending order, and $\oplus$ denotes the UD-PCR operator defined in Definition 1. Fusing uncertain views first prevents premature commitment to overconfident predictions, while the cross-view opinion ω^{cross} serves as a reliable reference for the final decision.

After fusion, the final evidence can be recovered via the inverse mapping $e_k^{\diamondsuit} = b_k^{\diamondsuit} \cdot K/u^{\diamondsuit}$, from which the final class probabilities and overall uncertainty are obtained. We now state the theoretical properties of the UD-PCR operator.

Proposition 1. *For any two valid subjective opinions $\omega^A = (\mathbf{b}^A, u^A)$ and $\omega^B = (\mathbf{b}^B, u^B)$, the aggregated opinion $\omega^{\diamondsuit} = \omega^A \oplus \omega^B$ remains a valid subjective opinion. That is, $b_k^{\diamondsuit} \geq 0$ for all $k \in \{1, \ldots, K\}$, $u^{\diamondsuit} \geq 0$, and $\sum_{k=1}^K b_k^{\diamondsuit} + u^{\diamondsuit} = 1$.*

Proposition 2. *If $u^A \to 0$ and $u^B \to 1$, then $\mathbf{b}^{\diamondsuit} \to \mathbf{b}^A$ and $u^{\diamondsuit} \to 0$. Physically, when one view is fully confident while the other is completely vacuous, the fused opinion inherently reduces to the confident view.*

Proposition 3. *If $u^A \to 0$, $u^B \to 0$, and the belief distributions are orthogonal such that $\langle \mathbf{b}^A, \mathbf{b}^B \rangle = 0$, then $u^{\diamondsuit} \to 1$. When two highly confident views completely contradict each other, the fusion mechanism falls back to maximum epistemic uncertainty.*

Proposition 4. *If $u^A \to 1$ and $u^B \to 1$, then $\mathbf{b}^{\diamondsuit} \to \mathbf{0}$ and $u^{\diamondsuit} \to 1$. When all views provide vacuous evidence, the fused opinion remains maximally uncertain, thereby preventing spurious confidence on unknown samples.*

The proofs of Propositions 1–4 are provided in the Appendix A.3. Together, these properties demonstrate that UD-PCR faithfully captures epistemic states across views: it increases confidence only for reliable and consistent opinions while preserving or enlarging uncertainty under conflicting or insufficient evidence, which is essential for open-world recognition of unknown samples.

3.5 Training Objective

3.5.1 Evidential Classification Loss

The evidential deep learning paradigm trains the evidence collectors by minimizing the negative log-likelihood of the Dirichlet distribution. For the v-th view, given its Dirichlet parameter $\boldsymbol{\alpha}^v = \mathbf{e}^v + \mathbf{1}$ and the one-hot label $\mathbf{y}$, the adjusted cross-entropy loss is

$$\mathcal{L}_{\text{ace}}^v = \sum_{k=1}^{K} y_k \left(\psi(S^v) - \psi(\alpha_k^v) \right), \tag{14}$$

where $\psi(\cdot)$ is the digamma function and $S^v = \sum_{k=1}^K \alpha_k^v$ is the Dirichlet strength.

To suppress evidence for incorrect classes, a Kullback–Leibler regularization term is introduced by shrinking the misleading evidence toward zero:

$$\mathcal{L}_{\text{KL}}^v = \text{KL}(\text{Dir}(\mathbf{p} \mid \tilde{\boldsymbol{\alpha}}^v) \parallel \text{Dir}(\mathbf{p} \mid \mathbf{1})), \tag{15}$$

where $\tilde{\boldsymbol{\alpha}}^v = \mathbf{y} + (1 - \mathbf{y}) \odot \boldsymbol{\alpha}^v$ retains the ground-truth evidence while zeroing out non-target classes. The uniform Dirichlet $\text{Dir}(\mathbf{p} \mid \mathbf{1})$ serves as the reference distribution. The evidential classification loss for the v-th view is formulated as follows:

$$\mathcal{L}_{\text{acc}}^v = \mathcal{L}_{\text{ace}}^v + \lambda_t \, \mathcal{L}_{\text{KL}}^v, \tag{16}$$

where $\lambda_t = \min(1, t/T)$ is an annealing coefficient that gradually increases with training epoch t up to a step threshold T, preventing premature convergence to a flat distribution in early training. The same loss is symmetrically applied to the fused evidence $\mathbf{e}^{\diamondsuit}$ to obtain $\mathcal{L}_{\text{acc}}^{\diamondsuit}$. The total evidential loss is averaged over all V views and the fused opinion as follows:

$$\mathcal{L}_{\text{acc}} = \frac{1}{V+1} \left(\sum_{v=1}^{V} \mathcal{L}_{\text{acc}}^v + \mathcal{L}_{\text{acc}}^{\diamondsuit} \right). \tag{17}$$

3.5.2 JOUSSELME CONSISTENCY LOSS

While the $\mathcal{L}_{\text{acc}}$ aligns each individual opinion with the ground truth, it does not enforce cross-view epistemic agreement. To this end, we adopt the Jousselme distance, a principled metric defined over the power-set of the frame of discernment. Given the mass functions of two opinions $\mathbf{m}^i = [b_1^i, \ldots, b_K^i, u^i]^\top$ and $\mathbf{m}^j$ over the focal elements $\{\{1\}, \ldots, \{K\}, \mathcal{F}\}$, where $\mathcal{F}$ is the full frame, the squared Jousselme distance is as follows:

$$d_J^2(\omega^i, \omega^j) = (\mathbf{m}^i - \mathbf{m}^j)^\top \mathbf{D} \, (\mathbf{m}^i - \mathbf{m}^j), \tag{18}$$

where $\mathbf{D} \in \mathbb{R}^{(K+1)\times(K+1)}$ is the Jaccard similarity matrix with entries $D_{pq} = |A_p \cap A_q| / |A_p \cup A_q|$ for focal elements A_p, A_q. For singletons $\{k\}$ and the full frame $\mathcal{F}$, evaluating the quadratic form yields a closed-form expression as follows:

$$d_J^2(\omega^i, \omega^j) = \sum_{k=1}^{K} \big(b_k^i - b_k^j\big)^2 + \frac{K-2}{K} \big(u^i - u^j\big)^2, \tag{19}$$

where the precise mathematical derivation from Eq. 18 to Eq. 19 is detailed in the Appendix (see Appendix A.2). The Jousselme consistency loss is then applied to V views as follows:

$$\mathcal{L}_{\text{con}} = \frac{2}{V(V-1)} \sum_{i=1}^{V} \sum_{j=i+1}^{V} d_J^2(\omega^i, \omega^j). \tag{20}$$

Compared with Eq. 18, Eq. 19 achieves linear $\mathcal{O}(K)$ complexity for each view pair.

3.5.3 OVERALL OBJECTIVE

The complete training objective combines the evidential classification loss, the Jousselme consistency loss, and the cross-view alignment loss $\mathcal{L}_{\text{vib}}$ (Eq. 2) as below:

$$\mathcal{L} = \mathcal{L}_{\text{acc}} + \gamma \, \mathcal{L}_{\text{con}} + \beta \, \mathcal{L}_{\text{vib}}, \tag{21}$$

where γ and β are trade-off parameters.

4 EXPERIMENTS

4.1 DATASETS, COMPETITORS AND EVALUATION METRICS

We evaluate EU^{2}O on six publicly available multi-view/multi-modal benchmarks, including Flower17, VGGFace2-50, Youtube, Scene, MVSA-Single Niu et al. (2016), and NYU Depth V2 Silberman et al. (2012).Table 4 summarizes their statistics in supplementary material. To construct the open-world setting, we follow Scheirer et al. (2012) and partition each dataset by an openness ratio $openness = 1 - \sqrt{2K/(K+N)}$, where K and N denote the numbers of known and total categories, respectively.

We compare EU^{2}O with seven state-of-the-art trusted multi-view classification methods, including TMC Han et al. (2022), TMDLO Liu et al. (2022), RCML Xu et al. (2024), CCML Liu et al. (2024), TMCEK Liang et al. (2025), SAEML Xu et al. (2025), and GTMC-HOA Shi et al. (2026). For fair comparison, all methods are evaluated using their official code with recommended hyperparameter settings, and each experiment is repeated over five random runs. Detailed descriptions of the datasets and comparison methods are provided in Appendix A.4.

For closed-set evaluation, we report classification accuracy (ACC), macro-averaged F1 score (Macro-F1), precision, and recall. For open-set evaluation, following Dong et al. (2025b), we adopt OSCR Dhamija et al. (2018), CCR@FPR=0.5%, CCR@FPR=1%, and Open AUC Wang et al. (2022). Detailed definitions of all evaluation metrics are provided in Appendix A.4.

4.2 PERFORMANCE COMPARISON

4.2.1 CLOSED-SET CLASSIFICATION

Table 1 reports the closed-set results. EU^{2}O surpasses all baseline methods on five of the six datasets with consistent gains across all four metrics. We attribute this to two factors. (1) The AU-DCF module quantifies view-specific aleatoric uncertainty and suppresses low-quality views during fusion,

Table 1: Closed-set ($openness = 0$) classification performance across six benchmarks. The best results are **bolded**, and the second-best are underlined.

Dataset	Flower17				VGGFace2-50				Youtube			
Metrics	Accuracy	Macro F1	Precision	Recall	Accuracy	Macro F1	Precision	Recall	Accuracy	Macro F1	Precision	Recall
TMC (21, ICLR)	63.60$_{\pm 1.68}$	60.72$_{\pm 1.99}$	64.63$_{\pm 3.30}$	63.60$_{\pm 1.68}$	61.30$_{\pm 0.33}$	60.96$_{\pm 0.31}$	62.48$_{\pm 0.37}$	61.27$_{\pm 0.33}$	61.75$_{\pm 1.97}$	60.52$_{\pm 2.21}$	64.78$_{\pm 2.55}$	61.75$_{\pm 1.97}$
TMDLO (22, AAAI)	64.56$_{\pm 1.03}$	60.59$_{\pm 1.40}$	62.56$_{\pm 2.41}$	64.56$_{\pm 1.03}$	49.38$_{\pm 0.50}$	48.23$_{\pm 0.54}$	55.40$_{\pm 0.42}$	49.21$_{\pm 0.49}$	79.35$_{\pm 1.89}$	79.21$_{\pm 1.92}$	80.52$_{\pm 1.62}$	79.35$_{\pm 1.89}$
RCML (24, AAAI)	58.24$_{\pm 2.24}$	54.56$_{\pm 2.66}$	62.89$_{\pm 2.90}$	58.24$_{\pm 2.24}$	58.94$_{\pm 0.53}$	58.57$_{\pm 0.50}$	60.64$_{\pm 0.52}$	58.89$_{\pm 0.53}$	71.45$_{\pm 0.83}$	71.15$_{\pm 0.96}$	74.47$_{\pm 1.10}$	71.45$_{\pm 0.83}$
TMCEK (25, ICML)	60.66$_{\pm 2.41}$	57.52$_{\pm 3.02}$	64.94$_{\pm 2.65}$	60.66$_{\pm 2.41}$	61.14$_{\pm 0.54}$	60.86$_{\pm 0.50}$	62.51$_{\pm 0.55}$	61.11$_{\pm 0.54}$	72.75$_{\pm 1.23}$	72.41$_{\pm 1.38}$	75.37$_{\pm 1.40}$	72.75$_{\pm 1.23}$
CCML (25, MM)	68.24$_{\pm 0.76}$	66.76$_{\pm 0.98}$	68.83$_{\pm 1.60}$	68.24$_{\pm 0.76}$	59.57$_{\pm 0.29}$	59.44$_{\pm 0.29}$	61.44$_{\pm 0.26}$	59.55$_{\pm 0.29}$	69.30$_{\pm 1.22}$	68.72$_{\pm 1.22}$	70.30$_{\pm 1.22}$	69.30$_{\pm 1.22}$
SAEML (25, MM)	70.51$_{\pm 0.97}$	68.62$_{\pm 1.31}$	71.35$_{\pm 2.03}$	70.51$_{\pm 0.97}$	65.96$_{\pm 0.40}$	65.89$_{\pm 0.34}$	66.80$_{\pm 0.29}$	65.94$_{\pm 0.41}$	**86.55**$_{\pm 1.09}$	86.46$_{\pm 1.15}$	**86.96**$_{\pm 0.95}$	**86.55**$_{\pm 1.09}$
GTMC-HOA (26, TMM)	43.24$_{\pm 4.14}$	36.58$_{\pm 5.17}$	41.41$_{\pm 8.00}$	43.24$_{\pm 4.14}$	61.02$_{\pm 0.36}$	60.97$_{\pm 0.37}$	61.84$_{\pm 0.39}$	60.97$_{\pm 0.37}$	78.30$_{\pm 2.74}$	77.34$_{\pm 3.57}$	79.34$_{\pm 2.59}$	78.30$_{\pm 2.74}$
Ours	**81.54**$_{\pm 1.53}$	**81.13**$_{\pm 1.71}$	**82.29**$_{\pm 1.55}$	**81.54**$_{\pm 1.53}$	**70.53**$_{\pm 0.39}$	**70.53**$_{\pm 0.39}$	**71.21**$_{\pm 0.43}$	**70.53**$_{\pm 0.41}$	**86.55**$_{\pm 1.96}$	**86.48**$_{\pm 2.05}$	86.97$_{\pm 1.83}$	**86.55**$_{\pm 1.96}$
Dataset	Scene				MVSA-Single				NYU Depth V2			
Metrics	Accuracy	Macro F1	Precision	Recall	Accuracy	Macro F1	Precision	Recall	Accuracy	Macro F1	Precision	Recall
TMC (21, ICLR)	72.11$_{\pm 0.77}$	69.96$_{\pm 0.61}$	71.81$_{\pm 0.83}$	71.15$_{\pm 0.69}$	67.36$_{\pm 1.53}$	58.67$_{\pm 2.21}$	69.67$_{\pm 2.46}$	56.24$_{\pm 1.86}$	69.52$_{\pm 1.17}$	67.24$_{\pm 1.51}$	68.45$_{\pm 2.35}$	68.50$_{\pm 1.00}$
TMDLO (22, AAAI)	64.48$_{\pm 0.40}$	59.15$_{\pm 0.33}$	64.66$_{\pm 0.97}$	62.31$_{\pm 0.26}$	63.62$_{\pm 0.52}$	49.29$_{\pm 0.77}$	**73.96**$_{\pm 1.46}$	48.28$_{\pm 0.53}$	66.94$_{\pm 1.06}$	62.84$_{\pm 1.34}$	73.80$_{\pm 2.33}$	60.30$_{\pm 1.31}$
RCML (24, AAAI)	70.18$_{\pm 0.61}$	67.33$_{\pm 0.85}$	70.53$_{\pm 1.33}$	68.92$_{\pm 0.65}$	68.40$_{\pm 1.74}$	59.95$_{\pm 2.13}$	71.47$_{\pm 2.83}$	57.24$_{\pm 1.77}$	69.93$_{\pm 1.17}$	67.63$_{\pm 1.62}$	69.39$_{\pm 2.80}$	68.50$_{\pm 0.88}$
TMCEK (25, ICML)	70.80$_{\pm 0.51}$	67.69$_{\pm 0.49}$	71.62$_{\pm 1.63}$	69.42$_{\pm 0.55}$	**71.33**$_{\pm 1.66}$	65.29$_{\pm 1.88}$	69.26$_{\pm 2.50}$	63.52$_{\pm 1.86}$	68.10$_{\pm 1.88}$	65.79$_{\pm 2.37}$	69.13$_{\pm 2.43}$	64.77$_{\pm 2.15}$
CCML (25, MM)	70.75$_{\pm 0.91}$	67.91$_{\pm 1.12}$	71.02$_{\pm 0.80}$	69.61$_{\pm 0.92}$	68.94$_{\pm 2.12}$	61.71$_{\pm 2.56}$	69.73$_{\pm 2.43}$	59.14$_{\pm 2.35}$	70.48$_{\pm 1.82}$	68.66$_{\pm 1.70}$	70.03$_{\pm 2.99}$	69.66$_{\pm 2.34}$
SAEML (25, MM)	77.67$_{\pm 1.16}$	76.70$_{\pm 1.21}$	**77.67**$_{\pm 0.90}$	77.12$_{\pm 1.08}$	67.71$_{\pm 1.63}$	62.50$_{\pm 1.80}$	64.16$_{\pm 1.83}$	61.53$_{\pm 1.86}$	70.75$_{\pm 1.16}$	68.94$_{\pm 3.00}$	71.17$_{\pm 4.54}$	68.17$_{\pm 1.98}$
GTMC-HOA (26, TMM)	77.52$_{\pm 1.09}$	**76.86**$_{\pm 0.90}$	77.05$_{\pm 0.91}$	77.01$_{\pm 0.85}$	63.01$_{\pm 1.32}$	59.18$_{\pm 1.22}$	60.27$_{\pm 1.18}$	59.00$_{\pm 1.13}$	69.05$_{\pm 1.61}$	61.75$_{\pm 4.58}$	65.05$_{\pm 6.61}$	61.56$_{\pm 3.99}$
Ours	**77.69**$_{\pm 0.81}$	76.79$_{\pm 0.65}$	77.17$_{\pm 0.68}$	**77.22**$_{\pm 0.68}$	70.56$_{\pm 2.11}$	**66.07**$_{\pm 2.00}$	67.36$_{\pm 2.57}$	**65.20**$_{\pm 1.62}$	**73.33**$_{\pm 1.97}$	**72.41**$_{\pm 1.94}$	**76.46**$_{\pm 3.98}$	**70.48**$_{\pm 1.06}$

yielding a robust joint representation that serves as a reliable reference for opinion aggregation. (2) The Jousselme consistency loss enforces cross-view epistemic alignment at linear complexity.

4.2.2 Open-set Recognition

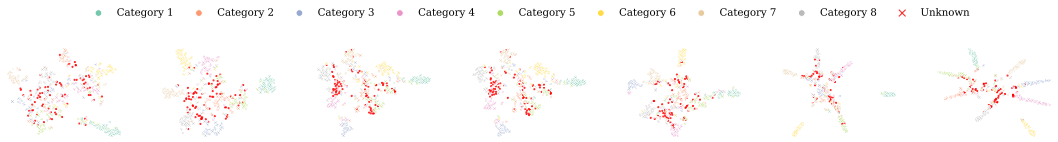
Table 2 presents the open-set results under $openness = 0.3$. EU^{2}O consistently outperforms all baselines on the majority of datasets, with the most pronounced advantages at low FPR thresholds. Two factors account for this. (1) The UD-PCR strategy dynamically redistributes conflict mass in proportion to view-level epistemic uncertainty; for unknown samples that elicit conflicting or vacuous opinions across views, UD-PCR preserves high uncertainty rather than suppressing it, addressing the epistemic collapse problem of conventional Dempster–Shafer fusion. (2) KAN-evidential collectors provide stronger expressiveness than MLP-based collectors in mapping heterogeneous representations to evidence, yielding more faithful uncertainty estimates.

Table 2: Open-set classification performance across benchmarks with $openness = 0.3$. The best results are **bolded**, and the second-best are underlined.

Dataset	Flower17				VGGFace2-50				Youtube			
Metrics	CCR@0.5%	CCR@1%	OSCR	AUC	CCR@0.5%	CCR@1%	OSCR	AUC	CCR@0.5%	CCR@1%	OSCR	AUC
TMC (21, ICLR)	3.25$_{\pm 3.02}$	5.50$_{\pm 3.76}$	57.23$_{\pm 7.70}$	64.57$_{\pm 9.28}$	13.50$_{\pm 4.72}$	18.44$_{\pm 5.14}$	59.40$_{\pm 3.33}$	72.52$_{\pm 2.16}$	8.33$_{\pm 4.86}$	15.17$_{\pm 8.14}$	67.11$_{\pm 10.07}$	73.70$_{\pm 7.34}$
TMDLO (22, AAAI)	7.25$_{\pm 5.94}$	8.25$_{\pm 6.15}$	62.81$_{\pm 6.81}$	70.64$_{\pm 6.18}$	10.26$_{\pm 3.94}$	14.60$_{\pm 4.48}$	55.57$_{\pm 4.41}$	71.55$_{\pm 2.57}$	11.83$_{\pm 8.32}$	16.50$_{\pm 9.54}$	70.97$_{\pm 7.06}$	76.29$_{\pm 5.56}$
RCML (24, AAAI)	4.00$_{\pm 3.00}$	6.25$_{\pm 3.95}$	53.06$_{\pm 9.26}$	62.72$_{\pm 7.24}$	13.08$_{\pm 3.82}$	18.13$_{\pm 3.64}$	57.63$_{\pm 3.55}$	71.68$_{\pm 2.19}$	6.33$_{\pm 7.39}$	9.67$_{\pm 9.87}$	63.86$_{\pm 11.00}$	69.54$_{\pm 8.47}$
TMCEK (25, ICML)	4.25$_{\pm 3.41}$	6.75$_{\pm 4.08}$	54.99$_{\pm 9.25}$	63.51$_{\pm 7.86}$	14.65$_{\pm 3.84}$	20.12$_{\pm 3.85}$	59.59$_{\pm 3.80}$	72.72$_{\pm 2.45}$	6.17$_{\pm 7.12}$	8.67$_{\pm 9.71}$	64.45$_{\pm 10.13}$	70.18$_{\pm 7.96}$
CCML (25, MM)	6.00$_{\pm 6.09}$	8.00$_{\pm 6.10}$	63.66$_{\pm 5.76}$	69.32$_{\pm 4.95}$	13.47$_{\pm 3.40}$	19.45$_{\pm 3.56}$	58.66$_{\pm 3.62}$	72.00$_{\pm 2.31}$	7.50$_{\pm 7.07}$	11.50$_{\pm 9.92}$	63.80$_{\pm 11.07}$	70.79$_{\pm 8.63}$
SAEML (25, MM)	13.25$_{\pm 8.16}$	18.75$_{\pm 10.16}$	70.83$_{\pm 6.88}$	76.45$_{\pm 4.80}$	**21.92**$_{\pm 3.88}$	27.82$_{\pm 3.41}$	67.70$_{\pm 2.29}$	77.02$_{\pm 1.74}$	13.17$_{\pm 8.32}$	18.83$_{\pm 8.73}$	**78.93**$_{\pm 4.13}$	81.04$_{\pm 3.52}$
GTMC-HOA (26, TMM)	3.25$_{\pm 4.15}$	5.00$_{\pm 6.89}$	53.86$_{\pm 10.08}$	61.97$_{\pm 9.60}$	10.68$_{\pm 5.73}$	14.30$_{\pm 6.82}$	64.49$_{\pm 3.44}$	75.54$_{\pm 2.39}$	**14.83**$_{\pm 7.61}$	19.17$_{\pm 10.01}$	75.52$_{\pm 7.62}$	79.06$_{\pm 7.10}$
Ours	**18.00**$_{\pm 9.80}$	**23.50**$_{\pm 9.76}$	**76.79**$_{\pm 8.18}$	**79.73**$_{\pm 5.76}$	21.79$_{\pm 5.01}$	**30.08**$_{\pm 4.18}$	**71.80**$_{\pm 2.19}$	**79.57**$_{\pm 1.85}$	14.33$_{\pm 9.85}$	**20.00**$_{\pm 11.84}$	**78.93**$_{\pm 4.61}$	**81.43**$_{\pm 3.98}$
Dataset	Scene				MVSA-Single				NYU-Depth-V2			
Metrics	CCR@0.5%	CCR@1%	OSCR	AUC	CCR@0.5%	CCR@1%	OSCR	AUC	CCR@0.5%	CCR@1%	OSCR	AUC
TMC (21, ICLR)	4.67$_{\pm 3.75}$	8.08$_{\pm 4.87}$	61.65$_{\pm 13.70}$	65.49$_{\pm 12.08}$	2.09$_{\pm 1.52}$	3.83$_{\pm 2.64}$	46.55$_{\pm 5.69}$	52.54$_{\pm 4.73}$	1.54$_{\pm 3.08}$	1.73$_{\pm 3.46}$	48.74$_{\pm 6.76}$	53.14$_{\pm 7.18}$
TMDLO (22,AAAI)	2.55$_{\pm 3.50}$	5.60$_{\pm 4.69}$	61.63$_{\pm 12.83}$	66.09$_{\pm 10.31}$	1.83$_{\pm 1.24}$	4.27$_{\pm 2.66}$	45.96$_{\pm 4.76}$	52.02$_{\pm 3.76}$	0.74$_{\pm 0.91}$	1.10$_{\pm 0.90}$	**52.58**$_{\pm 5.90}$	**53.56**$_{\pm 4.87}$
RCML (24, AAAI)	4.37$_{\pm 3.36}$	6.70$_{\pm 5.10}$	59.85$_{\pm 12.94}$	64.14$_{\pm 10.83}$	2.66$_{\pm 1.67}$	4.48$_{\pm 2.82}$	48.16$_{\pm 5.19}$	54.19$_{\pm 4.23}$	1.54$_{\pm 3.08}$	1.92$_{\pm 3.85}$	48.08$_{\pm 6.83}$	52.74$_{\pm 7.42}$
TMCEK (25, ICML)	4.47$_{\pm 3.62}$	7.80$_{\pm 5.52}$	60.91$_{\pm 12.41}$	65.44$_{\pm 10.30}$	2.10$_{\pm 1.23}$	4.10$_{\pm 1.75}$	52.60$_{\pm 4.04}$	58.52$_{\pm 3.07}$	1.89$_{\pm 2.50}$	1.89$_{\pm 2.50}$	47.62$_{\pm 9.10}$	52.38$_{\pm 8.56}$
CCML (25, MM)	4.97$_{\pm 2.73}$	8.12$_{\pm 4.30}$	63.00$_{\pm 12.12}$	68.28$_{\pm 10.31}$	2.56$_{\pm 1.50}$	4.54$_{\pm 2.91}$	47.97$_{\pm 5.83}$	53.98$_{\pm 4.97}$	1.35$_{\pm 2.69}$	1.95$_{\pm 2.49}$	49.17$_{\pm 6.67}$	53.50$_{\pm 6.98}$
SAEML (25, MM)	7.00$_{\pm 4.36}$	**11.14**$_{\pm 6.22}$	66.81$_{\pm 12.95}$	69.51$_{\pm 11.59}$	0.30$_{\pm 0.25}$	1.65$_{\pm 1.07}$	51.86$_{\pm 3.41}$	58.39$_{\pm 1.99}$	2.54$_{\pm 4.52}$	3.62$_{\pm 5.91}$	49.01$_{\pm 9.75}$	53.54$_{\pm 9.76}$
GTMC-HOA (26, TMM)	5.31$_{\pm 3.66}$	7.58$_{\pm 5.13}$	64.79$_{\pm 11.96}$	67.36$_{\pm 10.40}$	0.78$_{\pm 0.49}$	1.59$_{\pm 0.96}$	47.36$_{\pm 4.01}$	54.48$_{\pm 2.49}$	2.52$_{\pm 2.13}$	3.31$_{\pm 2.21}$	49.70$_{\pm 6.79}$	53.39$_{\pm 5.09}$
Ours	**7.46**$_{\pm 3.78}$	10.51$_{\pm 5.34}$	**68.13**$_{\pm 12.78}$	**71.19**$_{\pm 10.74}$	**5.18**$_{\pm 2.26}$	**7.62**$_{\pm 2.73}$	**56.22**$_{\pm 3.24}$	**63.16**$_{\pm 1.63}$	**4.00**$_{\pm 5.29}$	**4.96**$_{\pm 7.19}$	48.96$_{\pm 10.01}$	53.36$_{\pm 9.80}$

4.3 Ablation Study

To verify the contribution of each component, we conduct ablation experiments on MVSA-Single under $openness = 0.1$. Table 3 reports the results with different combinations of the three core modules. We obtain that all three components contribute positively to open-set rejection. Removing the AU-DCF module causes the largest drop, with CCR@0.5% falling from 4.83% to 3.02%. This can be attributed to the fact that without aleatoric uncertainty modeling, view-specific noise degrad the joint representation and cross-view alignment. Removing UD-PCR reduces CCR@1% from 7.24% to 5.57%, confirming that conflict-aware redistribution for belief and uncertainty is essential for preserving reliable uncertainty estimation for unseen sample at low FPR. Removing $\mathcal{L}_{\text{con}}$ also degrades performance. Besides, additional fusion strategy comparison are in supplementary (see Appendix A.5: Supplementary Experiments).



(a) TMC (b) TMDLO (c) RCML (d) TMCEK (e) CCML (f) GTMC-HOA (g) Ours

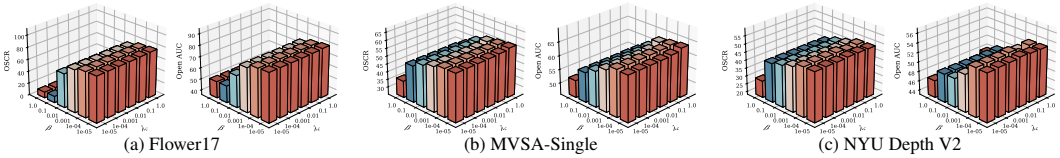
Figure 2: t-SNE visualizations on the YouTube dataset with openness $= 0.05$. Colored points represent samples from eight known categories, while red crosses represent unknown samples.

Table 3: Ablation study on the MVSA-Single dataset with $openness = 0.1$.

Components			CCR at FPR of	
AU-DCF	UD-PCR	$\mathcal{L}_{Jou}$	0.5%	1%
✓	-	✓	$3.47_{\pm 2.51}$	$5.57_{\pm 2.82}$
✓	✓	-	$3.85_{\pm 1.98}$	$6.28_{\pm 2.75}$
-	✓	✓	$3.02_{\pm 2.05}$	$5.82_{\pm 3.01}$
✓	✓	✓	$\mathbf{4.83}_{\pm 2.05}$	$\mathbf{7.24}_{\pm 2.48}$

4.4 T-SNE VISUALIZATION OF EVIDENCE

Fig. 2 presents the t-SNE visualization of evidence vectors for seven methods on the Youtube dataset under $openness = 0.05$. For the baselines, unknown samples intermingle with known-class clusters without clear separation. Our method produces compact intra-class clusters with clear inter-class separation and exhibits a distinct structural pattern for unknown samples. By reassigning high epistemic uncertainty to unknown instances, our method compresses their evidence vectors toward the origin . Consequently, unknown samples concentrate densely in the central region, remaining clearly separated from the surrounding known categories. This separation validates its effectiveness



(a) Flower17 (b) MVSA-Single (c) NYU Depth V2

Figure 3: Parameter sensitivity analysis on three datasets with $openness = 0.3$

.

4.5 PARAMETER SENSITIVITY ANALYSIS

Fig. 3 reports OSCR and Open AUC under varying trade-off weights γ and β (Eq. 21) on Flower17, MVSA-Single, and NYU Depth V2 ($openness = 0.3$). Both parameters are sampled from $\{10^{-5}, 10^{-4}, 10^{-3}, 0.01, 0.1, 1\}$. The alignment weight β exhibits a pronounced sensitivity pattern where OSCR and Open AUC peak at $\beta = 10^{-4}$ and degrade on both sides. When β is too small ($\leq 10^{-5}$), the AU-DCF module provides insufficient distribution alignment. When β is too large (≥ 0.1), the compression bottleneck dominates, forcing representations toward the standard Gaussian prior and discarding discriminative information needed for classification. The Jousselme consistency weight γ exerts a milder influence.

5 CONCLUSION

In this paper, we investigate the rarely explored open-world problem in trusted multi-view classification, and propose EU^{2}O, a unified framework jointly embracing aleatoric and epistemic uncertainty for trustworthy decision-making. Specifically, the proposed AU-DCF module align view distribution among views by information bottleneck strategy and produces a robust joint representation based on aleatoric uncertainty. Besides, we propose a novel UD-PCR to achieve more reliable multi-view opinion aggregation. A Jousselme consistency loss enforces cross-view epistemic alignment at $\mathcal{O}(K)$ complexity. Extensive experiments on six benchmarks demonstrate its effectiveness.

AI USE STATEMENT

We used generative AI tools to assist with language polishing and to improve the clarity and readability of the manuscript. All AI-assisted content was carefully reviewed and verified by the authors. The authors take full responsibility for the final content of this work.

ETHICS STATEMENT

This work does not involve human subjects, private data collection, or potentially harmful applications. All experiments are conducted on publicly available datasets.

REPRODUCIBILITY STATEMENT

To facilitate reproducibility, we provide detailed algorithm descriptions, theoretical proofs, implementation details, and hyperparameter settings in the supplementary material. The datasets used in this work are publicly available, and the source code will be released upon acceptance.

REFERENCES

Abhijit Bendale and Terrance E. Boult. Towards open set deep networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

Junda Cheng, Wei Yin, Kaixuan Wang, Xiaozhi Chen, Shijie Wang, and Xin Yang. Adaptive fusion of single-view and multi-view depth for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10138–10147, 2024.

Arthur P Dempster. A generalization of bayesian inference. *Journal of the Royal Statistical Society: Series B (Methodological)*, 30(2):205–232, 1968.

Akshay Raj Dhamija, Manuel Günther, and Terrance Boult. Reducing network agnostophobia. *Advances in Neural Information Processing Systems (NeurIPS)*, 31, 2018.

Zhibin Dong, Dayu Hu, Jiaqi Jin, Siwei Wang, Xinwang Liu, and En Zhu. Selective cross-view topology for deep incomplete multi-view clustering. *IEEE Transactions on Image Processing*, 34:4792–4805, 2025a.

Zhicheng Dong, Xiaodong Yue, Yufei Chen, and Yuxian Zhou. Trusted open-world multi-view classification with dynamic opinion aggregation. In *Proceedings of the ACM International Conference on Multimedia*, pp. 1181–1189, 2025b.

Shide Du, Zihan Fang, Yanchao Tan, Changwei Wang, Shiping Wang, and Wenzhong Guo. Openviewer: Openness-aware multi-view learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 16389–16397, 2025.

Zongbo Han, Changqing Zhang, Huazhu Fu, and Joey Tianyi Zhou. Trusted multi-view classification with dynamic evidential fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):2551–2566, 2022.

Haojian Huang, Chuanyu Qin, Zhe Liu, Kaijing Ma, Jin Chen, Han Fang, Chao Ban, Hao Sun, and Zhongjiang He. Trusted unified feature-neighborhood dynamics for multi-view classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 17413–17421, 2025.

Xi Jiang, Jian Li, Hanqiu Deng, Yong Liu, Bin-Bin Gao, Yifeng Zhou, Jialin Li, Chengjie Wang, and Feng Zheng. Mmad: A comprehensive benchmark for multimodal large language models in industrial anomaly detection. In *Proceedings of the International Conference on Learning Representations (ICLR)*, volume 2025, pp. 87273–87295, 2025.

Audun Jøsang. *Subjective logic*, volume 3. Springer, 2016.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014. URL https://api.semanticscholar.org/CorpusID:6628106.

Haobin Li, Yijie Lin, Peng Hu, Mouxing Yang, and Xi Peng. Community-aware multi-view representation learning with incomplete information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 48(3):3517–3529, 2026.

Xinyan Liang, Shijie Wang, Yuhua Qian, Qian Guo, Liang Du, Bingbing Jiang, Tingjin Luo, and Feijiang Li. Trusted multi-view classification with expert knowledge constraints. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2025.

Yuena Lin, Yiyuan Wang, Gengyu Lyu, Yongjian Deng, Haichun Cai, Huibin Lin, Haobo Wang, and Zhen Yang. Enhance multi-view classification through multi-scale alignment and expanded boundary. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2025.

Chengliang Liu, Jie Wen, Yong Xu, Bob Zhang, Liqiang Nie, and Min Zhang. Reliable representation learning for incomplete multi-view missing multi-label classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(6):4940–4956, 2025.

Wei Liu, Xiaodong Yue, Yufei Chen, and Thierry Denoeux. Trusted multi-view deep learning with opinion aggregation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 7585–7593, 2022.

Yadong Liu, Chengliang Liu, Jie Wen, Li Shen, Bob Zhang, and Yong Xu. Learning compact semantic information and reliable pseudo-labels for incomplete multi-view multi-label classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–15, 2026.

Ying Liu, Lihong Liu, Cai Xu, Xiangyu Song, Ziyu Guan, and Wei Zhao. Dynamic evidence decoupling for trusted multi-view learning. In *Proceedings of the ACM International Conference on Multimedia*, pp. 7269–7277, 2024.

Lawrence Neal, Matthew Olson, Xiaoli Fern, Weng-Keen Wong, and Fuxin Li. Open set learning with counterfactual images. In *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.

Shuai Niu, Jing Ma, Liang Bai, Zhihua Wang, Li Guo, and Xian Yang. Ehr-knowgen: Knowledge-enhanced multimodal learning for disease diagnosis generation. *Information Fusion*, 102:102069, 2024.

Teng Niu, Shiai Zhu, Lei Pang, and Abdulmotaleb El-Saddik. Sentiment analysis on multi-view social data. In *International Conference on MultiMedia Modeling*, pp. 15–27, 2016.

Poojan Oza and Vishal M. Patel. C2ae: Class conditioned auto-encoder for open-set recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

Walter J Scheirer, Anderson de Rezende Rocha, Archana Sapkota, and Terrance E Boult. Toward open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(7): 1757–1772, 2012.

Xiao Shen, Zhihao Chen, Shirui Pan, Shuang Zhou, Laurence T Yang, and Xi Zhou. Open-set cross-network node classification via unknown-excluded adversarial graph domain alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 20398–20408, 2025.

Long Shi, Chuanqing Tang, Huangyi Deng, Cai Xu, Lei Xing, and Badong Chen. Generalized trusted multi-view classification framework with hierarchical opinion aggregation. *IEEE Transactions on Multimedia*, 2026.

Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 746–760. Springer, 2012.

Likai Tian, Jian Zhao, Zechao Hu, Zhengwei Yang, Hao Li, Lei Jin, Zheng Wang, and Xuelong Li. Ccin: Compositional conflict identification and neutralization for composed image retrieval. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3974–3983, 2025.

Junjie Wang, Bin Chen, Bin Kang, Yulin Li, Weizhi Xian, Yichi Chen, and Yong Xu. Ov-dquo: Open-vocabulary detr with denoising text query training and open-world unknown objects supervision. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 7762–7770, 2025a.

Xiaoli Wang, Anqi Huang, Yongli Wang, Guanzhou Ke, Xiaobin Hong, and Jun Liu. Global-semantic alignment distillation for partial multi-view classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 21287–21295, 2025b.

Yiming Wang, Qun Li, Dongxia Chang, Jie Wen, Fu Xiao, and Yao Zhao. A category-driven contrastive recovery network for double incomplete multi-view multi-label classification. *IEEE Transactions on Multimedia*, 27:4008–4017, 2025c.

Yuqi Wang, Jiawei He, Lue Fan, Hongxin Li, Yuntao Chen, and Zhaoxiang Zhang. Driving into the future: Multiview visual forecasting and planning with world model for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14749–14759, 2024.

Zitai Wang, Qianqian Xu, Zhiyong Yang, Yuan He, Xiaochun Cao, and Qingming Huang. Openauc: Towards auc-oriented open-set recognition. *Advances in Neural Information Processing Systems (NeurIPS)*, 35:25033–25045, 2022.

Xiaohui Wei, Yuting Chen, Feiping Nie, Haibo Liu, Qiya Song, and Lin Xiao. View-adaptive multi-granularity anchor learning for multi-view clustering. *IEEE Transactions on Image Processing*, 35:3113–3126, 2026.

Jie Wen, Lian Zhao, Xiaohuan Lu, Chengliang Liu, Li Shen, Chao Huang, and Yong Xu. Disentangling consistent and specific information for double incomplete multi-view multi-label classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–14, 2026.

Cai Xu, Jiajun Si, Ziyu Guan, Wei Zhao, Yue Wu, and Xiyue Gao. Reliable conflictive multi-view learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 16129–16137, 2024.

Cai Xu, Ziqi Wen, Jie Zhao, Wanqing Zhao, Jinlong Yu, Haishun Chen, Ziyu Guan, and Wei Zhao. Beyond equal views: Strength-adaptive evidential multi-view learning. In *Proceedings of the ACM International Conference on Multimedia*, pp. 1278–1287, 2025.

Qingyang Zhang, Haitao Wu, Changqing Zhang, Qinghua Hu, Huazhu Fu, Joey Tianyi Zhou, and Xi Peng. Provable dynamic fusion for low-quality multimodal data. In *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 41753–41769. PMLR, 2023.

Wang Zhang, Ziwen Martin Ma, Subhro Das, Tsui-Wei Lily Weng, Alexandre Megretski, Luca Daniel, and Lam M Nguyen. One step closer to unbiased aleatoric uncertainty estimation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 16857–16864, 2024.

A APPENDIX

APPENDIX

In the supplemental material:

- Appendix A.1: Limitation of Dempster–Shafer Fusion in TMC.
- Appendix A.2: Closed-Form Derivation of the Jousselme Consistency Loss.
- Appendix A.3: Theoretical Properties of UD-PCR.
  - Proof of Proposition 1.
  - Proof of Proposition 2.
  - Proof of Proposition 3.
  - Proof of Proposition 4.
- Appendix A.4: Experimental Settings.
  - Dataset Statistics.
  - Comparison Methods.
  - Implementation Details.
  - Evaluation Metrics.
- Appendix A.5: Supplementary Experiments.
  - Robustness Analysis under Multi-view Corruption.
  - Comparison of Different TMVC Fusion Strategies.
  - Visualization of Dirichlet Evidence Distributions.
  - Analysis of Uncertainty Density Distribution.

A.1 LIMITATION OF DEMPSTER–SHAFER FUSION IN TMC

Consider two subjective opinions $\omega^A = (\mathbf{b}^A, u^A)$ and $\omega^B = (\mathbf{b}^B, u^B)$. The Dempster–Shafer fusion rule used in TMC computes the fused uncertainty as

$$u^{A\oplus B} = \frac{u^A u^B}{1 - C}, \tag{22}$$

where $C = \sum_{i \neq j} b_i^A b_j^B$ denotes the conflict mass between the two opinions.

Since

$$C \leq \left(\sum_i b_i^A\right)\left(\sum_j b_j^B\right) = (1 - u^A)(1 - u^B), \tag{23}$$

we have

$$1 - C \geq u^A + u^B - u^A u^B \geq u^A, \tag{24}$$

and similarly,

$$1 - C \geq u^B. \tag{25}$$

Therefore,

$$u^{A\oplus B} = \frac{u^A u^B}{1 - C} \leq u^A, \qquad u^{A\oplus B} \leq u^B, \tag{26}$$

which yields

$$\boxed{u^{A\oplus B} \leq \min(u^A, u^B)}. \tag{27}$$

This result shows that DS fusion in TMC can only preserve or reduce uncertainty during evidence aggregation. Such a property is desirable under the closed-set assumption, where accumulating complementary evidence from multiple views is expected to increase prediction confidence. However, in open-world scenarios, test samples may belong to previously unseen categories. These unknown samples can induce conflicting or unreliable evidence across views, for which high epistemic uncertainty is required for reliable rejection. Since DS fusion monotonically suppresses uncertainty as evidence is aggregated, the uncertainty associated with unknown samples may be underestimated, leading to overconfident predictions toward known classes and unreliable unknown detection.

Algorithm 1 Training procedure for EU^{2}O

Input: Known training set $\mathcal{D}_{\text{train}} = \{\{\mathbf{x}_i^v\}_{v=1}^V, \mathbf{y}_i\}_{i=1}^N$; hyperparameters γ, β, annealing step T

Output: Optimized network parameters $\Theta = \{\{f_v^\mu, f_v^\sigma\}_{v=1}^V, \{\Phi^v\}_{v=1}^V, \Phi^{\text{cross}}\}$
while not converged **do**
  Sample a mini-batch of size B
  for each view $v = 1, \ldots, V$ **do**
    $\boldsymbol{\mu}_i^v, \boldsymbol{\sigma}_i^v \leftarrow f_v^\mu(\mathbf{x}_i^v), f_v^\sigma(\mathbf{x}_i^v)$
    $\mathbf{z}_i^v \leftarrow \boldsymbol{\mu}_i^v + \boldsymbol{\epsilon} \odot \boldsymbol{\sigma}_i^v, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ ▷ Eq. 1
  end for
  Compute $\mathcal{L}_{\text{vib}}$ via Eq. 2
  $\hat{\mathbf{z}}_i \leftarrow \sum_{v=1}^V w_i^v \mathbf{z}_i^v$ with w_i^v via Eq. 3
  for each view $v = 1, \ldots, V$ **do**
    $\mathbf{e}_i^v \leftarrow \Phi^v(\mathbf{z}_i^v), \quad \boldsymbol{\alpha}_i^v \leftarrow \mathbf{e}_i^v + \mathbf{1}$
    $S_i^v \leftarrow \sum_{k=1}^K \alpha_{i,k}^v$
    $\omega_i^v \leftarrow (\mathbf{b}_i^v, u_i^v)$ via Eq. 7
  end for
  $\mathbf{e}_i^{\text{cross}} \leftarrow \Phi^{\text{cross}}(\hat{\mathbf{z}}_i), \quad$ construct ω_i^{cross}
  Sort $\{\omega_i^v\}_{v=1}^V$ by S_i^v ascending to obtain permutation σ
  $\omega_i^\Diamond \leftarrow \omega_i^{\sigma(1)}$
  for $v = 2, \ldots, V$ **do**
    $\omega_i^\Diamond \leftarrow \omega_i^\Diamond \oplus \omega_i^{\sigma(v)}$ ▷ Definition 1
  end for
  $\omega_i^\Diamond \leftarrow \omega_i^\Diamond \oplus \omega_i^{\text{cross}}$
  Compute $\mathcal{L}_{\text{acc}}$ via Eq. 17
  Compute $\mathcal{L}_{\text{con}}$ via Eq. 20
  $\mathcal{L} \leftarrow \mathcal{L}_{\text{acc}} + \gamma\, \mathcal{L}_{\text{con}} + \beta\, \mathcal{L}_{\text{vib}}$
  Update Θ by gradient descent on $\mathcal{L}$
end while

A.2 Closed-Form Derivation of the Jousselme Consistency Loss

For two subjective opinions $\omega^i = (\mathbf{b}^i, u^i)$ and $\omega^j = (\mathbf{b}^j, u^j)$ over K classes, their mass functions over the focal elements $\{\{1\}, \ldots, \{K\}, \mathcal{F}\}$ are $\mathbf{m}^i = [b_1^i, \ldots, b_K^i, u^i]^\top$ and $\mathbf{m}^j = [b_1^j, \ldots, b_K^j, u^j]^\top$, where $\mathcal{F}$ denotes the full frame of discernment with $|\mathcal{F}| = K$. The squared Jousselme distance is the quadratic form:

$$d_J^2(\omega^i, \omega^j) = (\mathbf{m}^i - \mathbf{m}^j)^\top \mathbf{D}\, (\mathbf{m}^i - \mathbf{m}^j), \tag{28}$$

where $\mathbf{D} \in \mathbb{R}^{(K+1)\times(K+1)}$ is the Jaccard similarity matrix with entries $D_{pq} = |A_p \cap A_q| \,/\, |A_p \cup A_q|$ for focal elements A_p, A_q. The entries are evaluated case by case:

- For two distinct singletons $\{i\}, \{j\}$ with $i \neq j$: $D_{ij} = |\{i\} \cap \{j\}| \,/\, |\{i\} \cup \{j\}| = 0/2 = 0$.
- For a singleton with itself: $D_{ii} = |\{i\} \cap \{i\}| \,/\, |\{i\} \cup \{i\}| = 1/1 = 1$, for $1 \le i \le K$.
- For a singleton $\{i\}$ and the full frame $\mathcal{F}$: $D_{i,K+1} = D_{K+1,i} = |\{i\} \cap \mathcal{F}| \,/\, |\{i\} \cup \mathcal{F}| = 1/K$.
- For the full frame with itself: $D_{K+1,K+1} = |\mathcal{F} \cap \mathcal{F}| \,/\, |\mathcal{F} \cup \mathcal{F}| = K/K = 1$.

Thus the matrix takes the form:

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 1/K \\ 0 & 1 & \cdots & 0 & 1/K \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1/K \\ 1/K & 1/K & \cdots & 1/K & 1 \end{bmatrix}. \tag{29}$$

Computing the matrix-vector product entry-wise, for $1 \le k \le K$:

$$[\mathbf{D}(\mathbf{m}^i - \mathbf{m}^j)]_k = (b_k^i - b_k^j) + \frac{1}{K}(u^i - u^j), \tag{30}$$

and for the $(K+1)$-th entry:

$$[\mathbf{D}(\mathbf{m}^i - \mathbf{m}^j)]_{K+1} = \frac{1}{K}\sum_{k=1}^{K}(b_k^i - b_k^j) + (u^i - u^j). \tag{31}$$

Expanding the quadratic form yields:

$$\begin{aligned} d_J^2(\omega^i, \omega^j) &= \sum_{k=1}^{K}(b_k^i - b_k^j)\Big[(b_k^i - b_k^j) + \frac{1}{K}(u^i - u^j)\Big] \\ &\quad + (u^i - u^j)\Big[\frac{1}{K}\sum_{k=1}^{K}(b_k^i - b_k^j) + (u^i - u^j)\Big] \\ &= \sum_{k=1}^{K}(b_k^i - b_k^j)^2 + \frac{2}{K}(u^i - u^j)\sum_{k=1}^{K}(b_k^i - b_k^j) + (u^i - u^j)^2. \end{aligned} \tag{32}$$

Subjective logic imposes the mass conservation constraint $\sum_{k=1}^{K} b_k^v + u^v = 1$ for any valid opinion ω^v. Taking the difference between the two views gives:

$$\sum_{k=1}^{K}(b_k^i - b_k^j) = -(u^i - u^j). \tag{33}$$

Substituting Eq. 33 into Eq. 32:

$$\begin{aligned} d_J^2(\omega^i, \omega^j) &= \sum_{k=1}^{K}(b_k^i - b_k^j)^2 + \frac{2}{K}(u^i - u^j)\big[-(u^i - u^j)\big] + (u^i - u^j)^2 \\ &= \sum_{k=1}^{K}(b_k^i - b_k^j)^2 - \frac{2}{K}(u^i - u^j)^2 + (u^i - u^j)^2 \\ &= \sum_{k=1}^{K}(b_k^i - b_k^j)^2 + \frac{K-2}{K}(u^i - u^j)^2. \end{aligned} \tag{34}$$

Eq. 34 recovers the closed-form expression stated in Eq. 19 of the main text, with an $\mathcal{O}(K)$ evaluation cost per view pair.

A.3 Theoretical Properties of UD-PCR

The UD-PCR fusion operator is defined in Definition 1 of the main text. For two subjective opinions $\omega^A = (\mathbf{b}^A, u^A)$ and $\omega^B = (\mathbf{b}^B, u^B)$, the orthogonal consensus $\{\tilde{b}_k, \tilde{u}\}$ follows Eq. 10, the conflict mass C follows Eq. 11, and the redistribution terms $\{\Delta b_k^A, \Delta b_k^B, \Delta u\}$ follow Eq. 12 with normalization $Z = 1 - u^A u^B$. The fused opinion $\omega^{A\Diamond B}$ is given by Eq. 9. The proofs of Propositions 1–4 are provided below.

Proposition 1 (Closure). For any two valid subjective opinions ω^A and ω^B, the fused opinion $\omega^{A\Diamond B}$ satisfies $\sum_{k=1}^{K} b_k^{A\Diamond B} + u^{A\Diamond B} = 1$.

Proof. Summing Eq. 9 over all classes:

$$\sum_{k=1}^{K} b_k^{A\Diamond B} + u^{A\Diamond B} = \Big(\sum_{k=1}^{K}\tilde{b}_k + \tilde{u}\Big) + \Big(\sum_{k=1}^{K}\Delta b_k^A + \sum_{k=1}^{K}\Delta b_k^B + \Delta u\Big). \tag{35}$$

From Eq. 12, summing Δb_k^A over k:

$$\sum_{k=1}^{K} \Delta b_k^A = \sum_{k=1}^{K} b_k^A \Big(\sum_{j \neq k} b_j^B \Big) \frac{(1-u^A)u^B}{Z} = C \cdot \frac{(1-u^A)u^B}{Z}, \tag{36}$$

and symmetrically $\sum_k \Delta b_k^B = C \cdot (1-u^B)u^A/Z$. The total redistribution is:

$$\begin{aligned} &\sum_{k=1}^{K} \Delta b_k^A + \sum_{k=1}^{K} \Delta b_k^B + \Delta u \\ &\quad = C \frac{(1-u^A)u^B + (1-u^B)u^A + (1-u^A)(1-u^B)}{Z}. \end{aligned} \tag{37}$$

The numerator simplifies to $1 - u^A u^B = Z$, hence the total redistribution equals C. Substituting into Eq. 35:

$$\sum_{k=1}^{K} b_k^{A \Diamond B} + u^{A \Diamond B} = \sum_{k=1}^{K} \tilde{b}_k + \tilde{u} + C. \tag{38}$$

Expanding Eq. 10 and Eq. 11:

$$\begin{aligned} \sum_{k=1}^{K} \tilde{b}_k + \tilde{u} + C &= \Big(\sum_{k=1}^{K} b_k^A + u^A \Big) \Big(\sum_{k=1}^{K} b_k^B + u^B \Big) \\ &= 1 \cdot 1 = 1. \end{aligned} \tag{39}$$

□

Proposition 2 (Unilateral Absolute Obedience). If $u^A \to 0$ and $u^B \to 1$, then $\mathbf{b}^{A \Diamond B} \to \mathbf{b}^A$ and $u^{A \Diamond B} \to 0$.

Proof. From $\sum_k b_k^B = 1 - u^B$ and $b_k^B \geq 0$, we have $0 \leq b_k^B \leq 1 - u^B$. As $u^B \to 1$, the squeeze theorem gives $b_k^B \to 0$ for all k. Consequently, $C = \sum_{i \neq j} b_i^A b_j^B \to 0$, which implies $\Delta b_k^A, \Delta b_k^B, \Delta u \to 0$. Evaluating the orthogonal consensus under the limits:

$$\begin{aligned} \tilde{u} &= u^A u^B \to 0 \cdot 1 = 0, \\ \tilde{b}_k &= b_k^A b_k^B + b_k^A u^B + b_k^B u^A \to 0 + b_k^A \cdot 1 + 0 = b_k^A. \end{aligned} \tag{40}$$

Thus $u^{A \Diamond B} = \tilde{u} + \Delta u \to 0$ and $b_k^{A \Diamond B} = \tilde{b}_k + \Delta b_k^A + \Delta b_k^B \to b_k^A$. □

Proposition 3 (Epistemic Collapse under Orthogonal Conflict). If $u^A \to 0$, $u^B \to 0$, and $\langle \mathbf{b}^A, \mathbf{b}^B \rangle = 0$, then $u^{A \Diamond B} \to 1$.

Proof. Under the stated limits, the conflict mass from Eq. 11 evaluates to:

$$\begin{aligned} C = \sum_{i \neq j} b_i^A b_j^B &= \Big(\sum\nolimits_i b_i^A \Big) \Big(\sum\nolimits_j b_j^B \Big) - \sum\nolimits_k b_k^A b_k^B \\ &= (1-u^A)(1-u^B) - 0 \to 1. \end{aligned} \tag{41}$$

The redistribution term Δu from Eq. 12 becomes:

$$\Delta u = C \cdot \frac{(1-u^A)(1-u^B)}{1 - u^A u^B} \to 1 \cdot \frac{1 \cdot 1}{1 - 0} = 1. \tag{42}$$

Since $\tilde{u} = u^A u^B \to 0$, we obtain $u^{A \Diamond B} = \tilde{u} + \Delta u \to 1$. □

Proposition 4 (Asymptotic Collapse under Bilateral Vacuity). If $u^A \to 1$ and $u^B \to 1$, then $\mathbf{b}^{A \Diamond B} \to \mathbf{0}$ and $u^{A \Diamond B} \to 1$.

Proof. Let $\epsilon_A = 1 - u^A$, $\epsilon_B = 1 - u^B$, with $\epsilon_A, \epsilon_B \to 0^+$. From $\sum_k b_k^v = \epsilon_v$ and $b_k^v \geq 0$, we have $b_k^A \leq \epsilon_A$, $b_k^B \leq \epsilon_B$ for all k.

The conflict mass is bounded by:

$$0 \leq C = \sum_{i \neq j} b_i^A b_j^B \leq \Big(\sum\nolimits_i b_i^A\Big)\Big(\sum\nolimits_j b_j^B\Big) = \epsilon_A \epsilon_B. \tag{43}$$

The normalization factor satisfies $Z = 1 - (1 - \epsilon_A)(1 - \epsilon_B) = \epsilon_A + \epsilon_B - \epsilon_A \epsilon_B$, which is $\Theta(\epsilon)$ where $\epsilon = \max(\epsilon_A, \epsilon_B)$. From Eq. 12, each redistribution term is bounded:

$$\begin{aligned} \Delta u &\leq \frac{(\epsilon_A \epsilon_B)^2}{\epsilon_A + \epsilon_B - \epsilon_A \epsilon_B} = \mathcal{O}(\epsilon^3) \to 0, \\ \Delta b_k^A &\leq \frac{\epsilon_A \epsilon_B \cdot \epsilon_A}{\epsilon_A + \epsilon_B - \epsilon_A \epsilon_B} = \mathcal{O}(\epsilon^2) \to 0, \end{aligned} \tag{44}$$

and symmetrically $\Delta b_k^B \to 0$.

For the orthogonal consensus, as $u^A, u^B \to 1$:

$$\begin{aligned} \tilde{u} &= u^A u^B \to 1, \\ \tilde{b}_k &= b_k^A b_k^B + b_k^A u^B + b_k^B u^A \leq \epsilon_A \epsilon_B + \epsilon_A + \epsilon_B \to 0. \end{aligned} \tag{45}$$

Summing all contributions, $u^{A \Diamond B} = \tilde{u} + \Delta u \to 1$ and $b_k^{A \Diamond B} = \tilde{b}_k + \Delta b_k^A + \Delta b_k^B \to 0$. □

A.4 Experimental Settings

A.4.1 Datasets

We evaluate on six multi-view and multi-modal benchmarks. Table 4 summarizes their statistics.

Flower17 consists of different flower categories that are common in the UK. The images are available in a variety of views, scales, and photometric variations. Seven visual descriptors are extracted for each sample. **VGGFace2-50** is a large-scale face recognition dataset with 50 identities, where each face is described by four deep features. **Youtube** is a video game dataset including both visual (Cuboids Histogram, Histogram of Motion Estimate, and HOG features) and audio features (MFCC, Volume Streams, and Spectrogram Streams). **Scene** provides 15 scene categories with three handcrafted visual features. **MVSA-Single**Niu et al. (2016) is a sentiment analysis benchmark pairing product images with textual reviews across three categories (*Positive*, *Negative*, *Neutral*). Since the raw annotations contain inconsistent image–text label pairs, we retain only samples whose visual and textual labels agree, yielding 2,592 valid instances from the original 4,869. **NYU Depth V2** Silberman et al. (2012) is utilized for scene recognition and comprises depth and RGB images. Following Zhang et al. (2023), we adopt the commonly used 9 out of the 27 scene categories and treat the remaining categories as *Others*.

Table 4: Summary of the six benchmark datasets.

Dataset	Samples	Views	Classes	Dimensionality
Flower17	1,360	7	17	1,360, 1,360, 1,360, 1,360, 1,360, 1,360, 1,360
VGGFace2-50	34,027	4	50	944, 576, 512, 640
Youtube	2,000	6	10	2000, 1024, 64, 512, 64, 647
Scene	4,485	3	15	20, 59, 40
MVSA-Single	2,592	2	3	2048, 768
NYU Depth V2	1,449	2	10	2048, 2048

For open-world evaluation, we follow Scheirer et al. (2012) and partition each dataset by an openness ratio $openness = 1 - \sqrt{2K/(K+N)}$, where K and N denote the numbers of known and total categories, respectively.

Table 5: Ablation study of evidence fusion strategies on the Scene dataset with openness $= 0.3$.

Fusion strategy	CCR@0.5%	CCR@1%	OSCR	AUC
MEAN	$5.01_{\pm 4.28}$	$\underline{8.81}_{\pm 6.04}$	$61.12_{\pm 14.02}$	$65.11_{\pm 12.49}$
RCML	$\underline{5.13}_{\pm 3.70}$	$8.52_{\pm 5.83}$	$61.01_{\pm 13.15}$	$\underline{65.49}_{\pm 11.14}$
TMC	$4.84_{\pm 3.80}$	$8.17_{\pm 4.96}$	$\underline{61.57}_{\pm 13.70}$	$65.43_{\pm 12.11}$
TMCEK	$5.03_{\pm 3.66}$	$\mathbf{9.08}_{\pm 6.55}$	$60.99_{\pm 13.40}$	$65.35_{\pm 11.47}$
Ours	$\mathbf{5.38}_{\pm 3.84}$	$8.45_{\pm 5.04}$	$\mathbf{61.96}_{\pm 13.83}$	$\mathbf{65.81}_{\pm 12.20}$

A.4.2 Comparison Methods

We compare EU^{2}O against seven state-of-the-art TMVC methods. **TMC** Han et al. (2022) pioneers TMVC by modeling class probabilities with Dirichlet distributions and fusing multi-view opinions via Dempster–Shafer theory. **TMDLO** Liu et al. (2022) formulates multi-view fusion as opinion aggregation with evidence accumulation and theoretically guarantees decreased uncertainty after fusion. **RCML** Xu et al. (2024) averages all evidence to mitigate the impact of inter-view disagreement during opinion aggregation. **CCML** Liu et al. (2024) dynamically decouples consistent and complementary evidence from multi-view opinions to address potential sematic vagueness. **TMCEK** Liang et al. (2025) introduces expert knowledge constraints into evidence collection and employs a distribution-aware subjective opinion mechanism for reliable confidence estimation. **SAEML** Xu et al. (2025) incorporates strength-adaptive evidential modeling that weights each view according to its estimated reliability rather than treating all views equally. **GTMC-HOA** Shi et al. (2026) adopts a hierarchical opinion aggregation framework that progressively fuses view-specific opinions across multiple granularity levels.

For fair comparison, all methods are evaluated using their official code with recommended hyperparameter settings, and each experiment is repeated over five random runs.

A.4.3 Implementation Details

Evidence collectors adopt a two-layer KAN architecture with hidden dimensions $[256, 128]$, grid size 5, and spline order 3, followed by Softplus activation to ensure non-negative outputs. For the MVSA-Single, image features are extracted by a pretrained ResNet-152 and text features by a pretrained BERT-base, then projected to a unified 256-dimensional latent space. For NYU Depth V2, both RGB and depth images are encoded by a pretrained ResNet. The remaining datasets use publicly available pre-extracted features following their original benchmarks. The model is trained with the Adam optimizer Kingma & Ba (2014), a batch size of 200, learning rate selected from $\{10^{-4}, 3\times 10^{-4}, 5\times 10^{-4}, 10^{-3}, 3\times 10^{-3}\}$, and weight decay 10^{-5}. All experiments are implemented in PyTorch on a single NVIDIA GeForce RTX 4080 SUPER GPU with 32 GB memory. For fair comparison, each method is evaluated using official code with recommended hyperparameters, and all experiments are repeated over five random runs.

A.4.4 Evaluation Metrics

For closed-set evaluation, we report classification accuracy (ACC), macro-averaged F1 score, precision, and recall over the known categories.

For open-set evaluation, we adopt four widely-used metrics following Dong et al. (2025b) to assess the model's rejection capability. Given a test set containing N_k known and N_u unknown samples, we define the confidence score of the i-th sample as $\mathcal{S}_i = 1 - u_i$, where u_i is the uncertainty mass derived from the fused opinion. For a specific confidence threshold $\tau \in [0, 1]$, the Correct Classification Rate (CCR) and False Positive Rate (FPR) are formally defined as:

$$\begin{aligned} \mathrm{CCR}(\tau) &= \frac{1}{N_k}\sum_{i\in\mathcal{I}_k}\mathbb{I}(\hat{y}_i = y_i \wedge \mathcal{S}_i \geq \tau), \\ \mathrm{FPR}(\tau) &= \frac{1}{N_u}\sum_{i\in\mathcal{I}_u}\mathbb{I}(\mathcal{S}_i \geq \tau), \end{aligned} \tag{46}$$

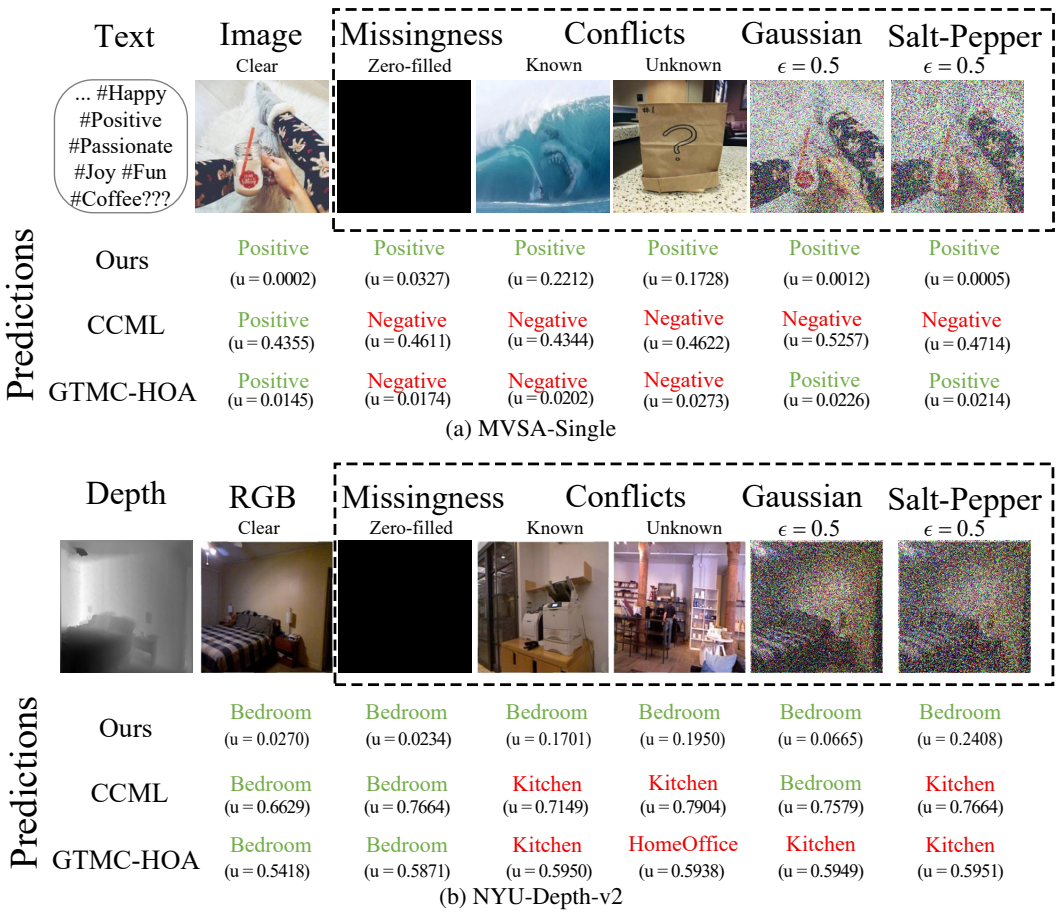


Figure 4: Visualization results on different datasets. The text or depth view remains uncorrupted, whereas the RGB image view is subjected to five extreme perturbations: zero-filled missingness, semantic conflicts caused by replacement with a random sample from another known class or the unknown class, and heavy Gaussian and salt-and-pepper noise. For MVSA-Single, *Positive* and *Negative* are known classes, while *Neutral* is treated as unknown. For NYU-Depth-v2, *Others* denotes the unknown category. Green and red text indicate correct and incorrect predictions, respectively, and u denotes the estimated uncertainty. Our method maintains accurate predictions with low uncertainty under the examined corruptions.

where $\mathcal{I}_k$ and $\mathcal{I}_u$ denote the index sets of known and unknown samples, respectively, and $\mathbb{I}(\cdot)$ is the indicator function. Based on Eq. 46, the open-set performance is comprehensively evaluated by:

- **OSCR** Dhamija et al. (2018): The Open-Set Classification Rate, computed as the area under the CCR vs. FPR curve. It comprehensively evaluates the trade-off between accurate classification of knowns and effective rejection of unknowns.

- **CCR@FPR:** The CCR value strictly evaluated at specific low FPR operating points (e.g., FPR at 0.5% and 1%).

- **Open AUC** Wang et al. (2022): The Area Under the Receiver Operating Characteristic (AUROC) curve for unknown detection, where known and unknown samples serve as positive and negative classes, respectively, based on the confidence score $\mathcal{S}_i$.

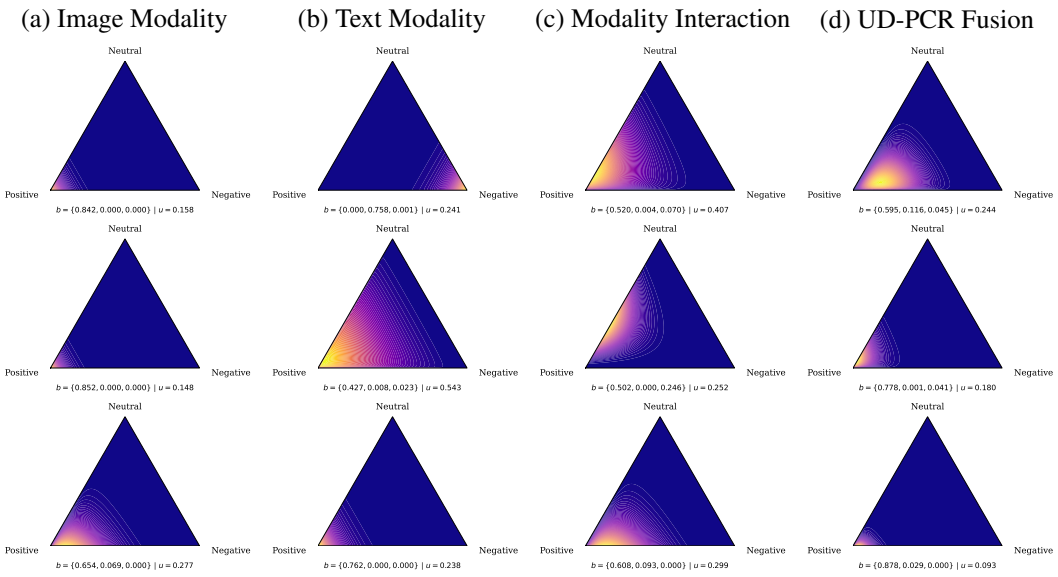


Figure 5: Visualization of Dirichlet evidence distributions across different views and fusion stages. Each row represents a distinct sample from the MVSA-Single dataset. Columns (a) and (b) display the subjective belief masses and uncertainty from the independent visual (ω^1) and textual branches ω^2, respectively. Column (c) exhibits the distribution derived from early cross-modal interaction (ω^{cross}). Column (d) is the final robust prediction formulated by our UD-PCR fusion ($\omega^{\Diamond}$).

A.5 Supplementary Experiments

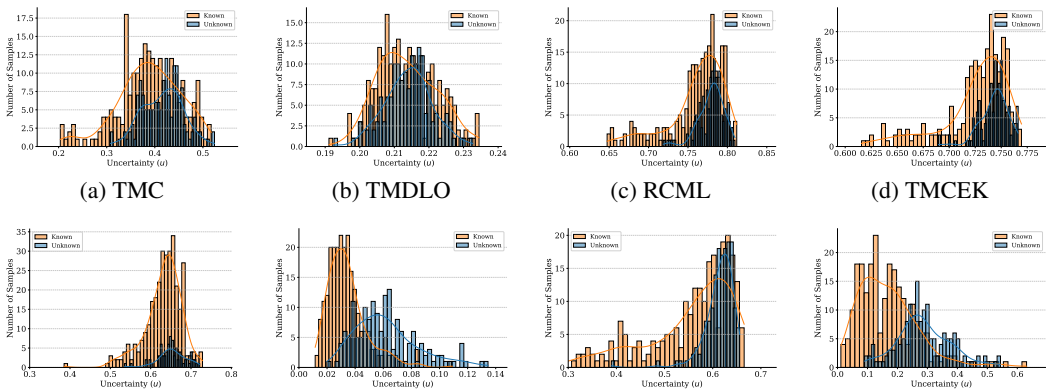
A.5.1 Robustness Analysis under Multi-view Corruption

Fig. 4 reports prediction and uncertainty under five types of RGB corruption on MVSA-Single and NYU Depth V2, with the other modality kept intact. The corruptions include zero-filled missingness, semantic conflicts by a sample from a different known category or from the unknown category, Gaussian noise, and Salt-Pepper noise.

On clean inputs, our method assigns low uncertainty ($u = 0.027$ on NYU Depth V2), while CCML and GTMC-HOA already produce substantially higher values ($u = 0.663$ and $u = 0.542$). Under Gaussian and Salt-Pepper noise, which introduces additional aleatoric uncertainty into the RGB view, the fused uncertainty increases moderately, consistent with the AU-DCF module down-weighting the corrupted view in the joint representation. When the RGB view is zeroed out, uncertainty increases further as the fused opinion relies on a single valid view. Under semantic conflict, where the RGB view is replaced by a sample from a different known category, uncertainty increases substantially. This behavior aligns with the UD-PCR epistemic collapse property (Proposition 3): when two views support divergent classes with low individual uncertainty, the fused opinion converges to high epistemic uncertainty rather than forcing an arbitrary consensus. When the RGB view is replaced by a sample from the unknown category, uncertainty is comparably elevated, confirming that unknown samples elicit vacuous opinions which UD-PCR preserves rather than suppresses, consistent with the bilateral vacuity property (Proposition 4). Across all corruption types, the uncertainty response tracks the nature and severity of the perturbation, demonstrating that EU^{2}O produces uncertainty estimates that reflect prediction reliability in open-world settings.

A.5.2 Comparison of Different TMVC Fusion Strategies

To highlight the effect of the UD-PCR fusion strategy, we replace the AU-DCF, $\mathcal{L}_{\text{con}}$ and KAN modules with a standard MLP-based evidence collector and compare different aggregation strategies on Scene under $openness = 0.3$. Table 5 shows that UD-PCR achieves the highest OSCR (61.96%) and Open AUC (65.81%). Among the baselines, MEAN fusion scores the highest CCR@1% (8.81%) but the lowest Open AUC, indicating that equal-weight averaging masks conflicting signals and undermines unknown rejection. TMC and RCML, both variants of Dempster–Shafer com-



(a) TMC (b) TMDLO (c) RCML (d) TMCEK

(e) CCML (f) SAEML (g) GTMC-HOA (h) Ours

Figure 6: Uncertainty distributions of known and unknown test samples on the YouTube dataset with openness $= 0.1$. Better open-world uncertainty estimation is characterized by less overlap between the known and unknown distributions and a clear shift of unknown samples toward higher uncertainty.

bination, obtain comparable OSCR but trail UD-PCR, consistent with their tendency to suppress uncertainty when views diverge. These results validate that dynamically routing conflict mass in proportion to view-level uncertainty, rather than enforcing unconditional uncertainty reduction, better preserves epistemic fidelity in open-world settings.

A.5.3 Visualization of Dirichlet Evidence Distributions

Fig. 5 visualizes Dirichlet belief distributions on the 3-class probability simplex for three MVSA-Single samples. Columns (a)–(d) correspond to the Image view, Text view, cross-view interaction, and UD-PCR fused opinion. The simplex corners represent Positive, Negative, and Neutral, and the heatmap encodes Dirichlet density.

The three samples illustrate three distinct fusion regimes of UD-PCR. When the two views are conflictive (row 1, $u = 0.244$), UD-PCR redistributes the conflicting belief mass into uncertainty, and the fused opinion preserves elevated uncertainty rather than forcing a concordant prediction. When one view is highly confident while the other is ambiguous (row 2, $u = 0.180$), the fused opinion tilts toward the confident view and achieves sharper concentration on the correct class. When both views agree on the correct class with low uncertainty (row 3, $u = 0.093$), UD-PCR further concentrates belief mass toward the correct corner while suppressing residual density in the simplex interior. Across all three regimes, the fused opinions place dominant belief on the correct class, confirming that UD-PCR adaptively resolves inter-view conflict and exploits cross-view agreement to produce well-calibrated uncertainty.

A.5.4 Analysis of Uncertainty Density Distribution

Fig. 6 plots the uncertainty density of known and unknown test samples on Youtube ($openness = 0.1$). For most competing methods, the two distributions exhibit strongly overlapping modes or long overlapping tails. Consequently, it is difficult to determine a stable rejection threshold that simultaneously retains correctly classified known samples and rejects unknown samples. Our method produces a clear rightward shift of the unknown distribution while keeping the known distribution concentrated in the low-uncertainty region. The visibly reduced overlap creates a wider admissible threshold interval for open-set rejection, which is consistent with the superior OSCR and Open AUC results. This separation is consistent with the UD-PCR strategy dynamically calibrating belief and uncertainty mass rather than collapsing it into a spurious concordant prediction. The clear distinction between known and unknown uncertainty prevents the model from assigning overconfident predictions to unfamiliar entities, addressing the open-world vulnerability.