

2025.05.13

Best Vision Paper

Presenter

고재훈

Reliable Conflictive Multi-View Learning

Cai Xu, Jiajun Si, Ziyu Guan, Wei Zhao, Yue Wu, Xiyue Gao

AAAI 2024, Outstanding Paper Award



Contents

I Problem

II Background

III Method

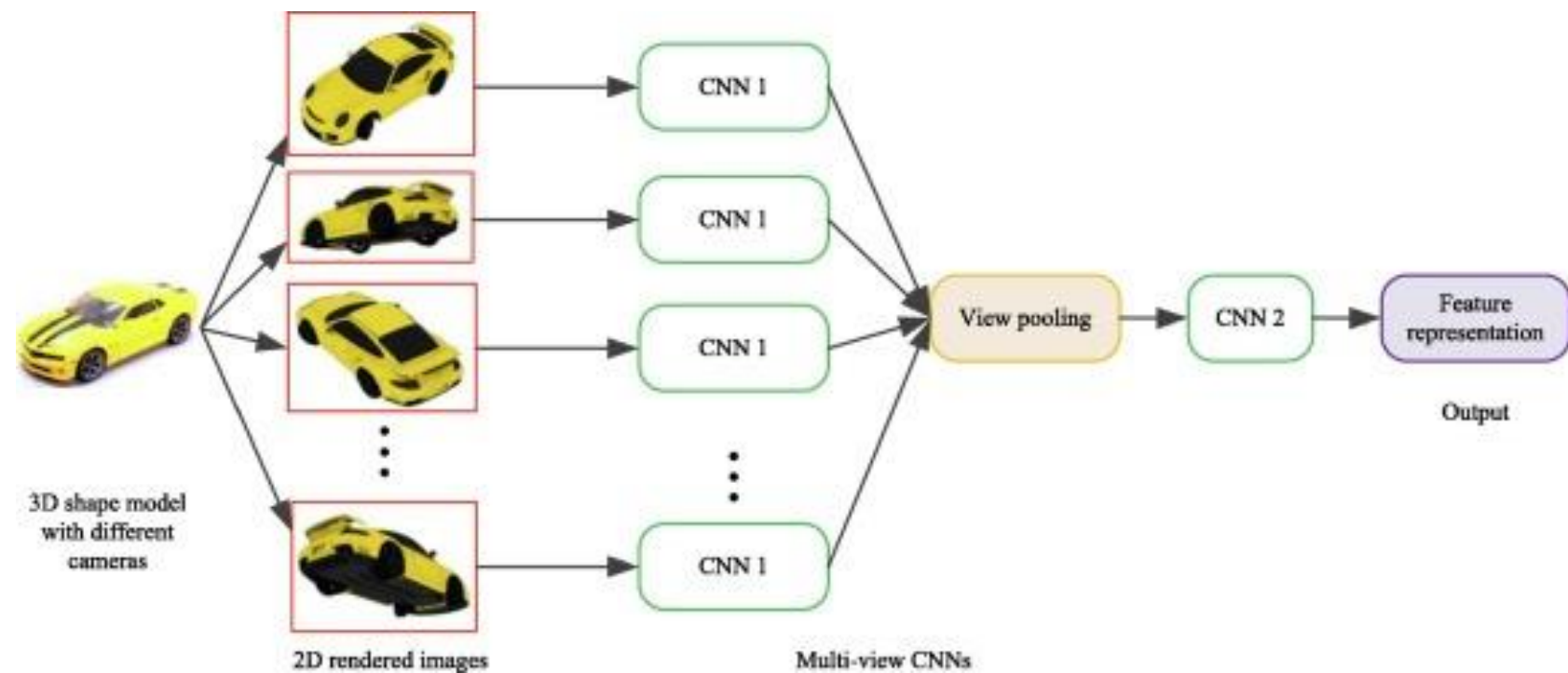
IV Experiments

V Conclusion

Problem

1. 기존의 Multi-View Learning은 모든 View가 엄격하게 정렬되어있다고 가정하고, 충돌되는 데이터를 제거하거나 교체하는 방식을 사용해왔음
2. 하지만 현실에서는 이러한 가정이 항상 보장되지는 않고, 충돌되는 View가 존재하는 상황에서 판단을 내려야함
3. 판단 결과와 신뢰도를 함께 활용하는 Reliable Conflictive Multi-view Learning(RCML) 문제와, 이를 해결하기 위한 Evidential Conflictive Multi-view Learning (ECML) 방법을 제안함

Background Multi-View Learning



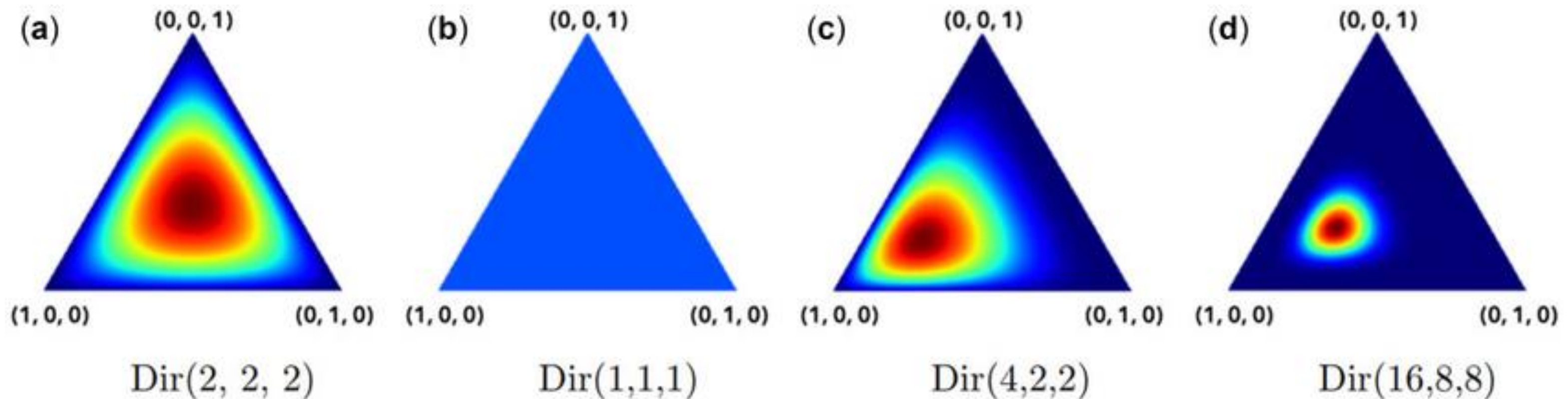
- 단일 대상에 대한 다양한 관점(Multiple Views)으로부터 얻은 정보를 통합하여 학습 결과를 도출하는 방식
- 다양한 관점에는 각도, 특징, 모달리티 등이 포함될 수 있음.

Background Dirichlet Distribution

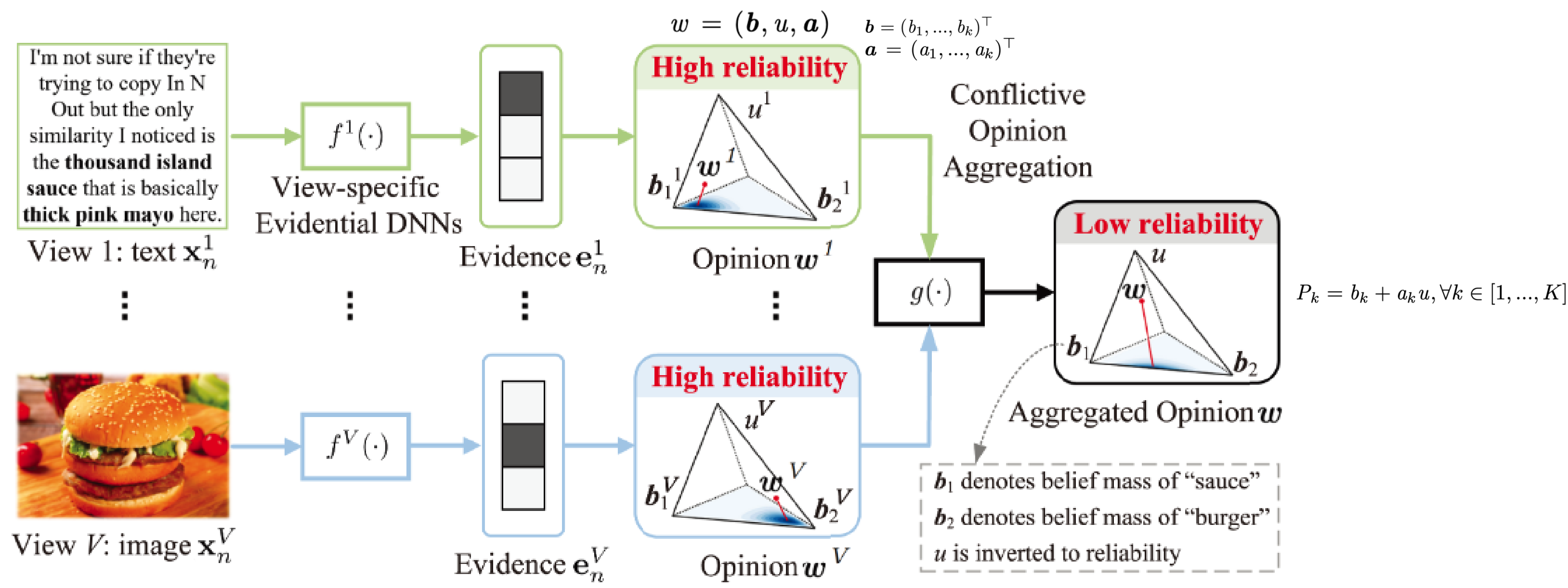
1. 합이 1인 K차원 양의 실수 벡터의 확률 분포 $\mathcal{S}_K = \left\{ \mathbf{p} \mid \sum_{k=1}^K p_k = 1 \text{ and } 0 \leq p_1, \dots, p_k \leq 1 \right\}$

2. ECML에서는 각 클래스별 확률 조합의 확률 분포를 모델링하는 데 사용

$$D(\mathbf{p}|\boldsymbol{\alpha}) = \begin{cases} \frac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^K p_k^{\alpha_k-1}, & \text{for } \mathbf{p} \in \mathcal{S}_K \\ 0, & \text{otherwise,} \end{cases} \quad \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)^\top$$

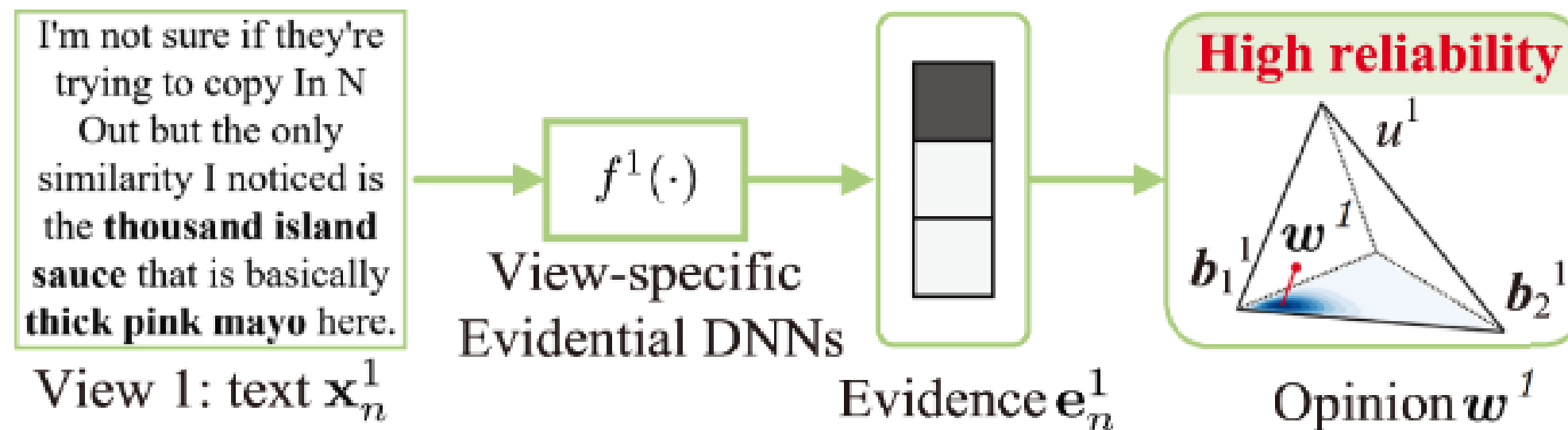


Method Overview



ECML에서는 각 View에서 얻은 증거를 종합하여 판단 결과의 신뢰도를 계산

Method View-specific Evidential Deep Learning



1. 각 DNN의 softmax를 ReLU로 대체 $\{\mathbf{e}_n^v\}$

2. Dirichlet strength $S = \sum_{k=1}^K (e_k + 1) = \sum_{k=1}^K \alpha_k$ 계산

$$3. b_k = \frac{e_k}{S} = \frac{\alpha_k - 1}{S}, u = \frac{K}{S} \Rightarrow \sum_{k=1}^K b_k + u = 1, \forall k \in [1, \dots, K]$$

Method Evidential Multi-view Fusion via Conflictive Opinion Aggregation

Definition 1. Conflictive Opinion Aggregation

$$\mathbf{w}^{A \diamond B} = \mathbf{w}^A \diamond \mathbf{w}^B = (\mathbf{b}^{A \diamond B}, u^{A \diamond B}, \mathbf{a}^{A \diamond B})$$

$$b_k^{A \diamond B} = \frac{b_k^A u^B + b_k^B u^A}{u^A + u^B}$$

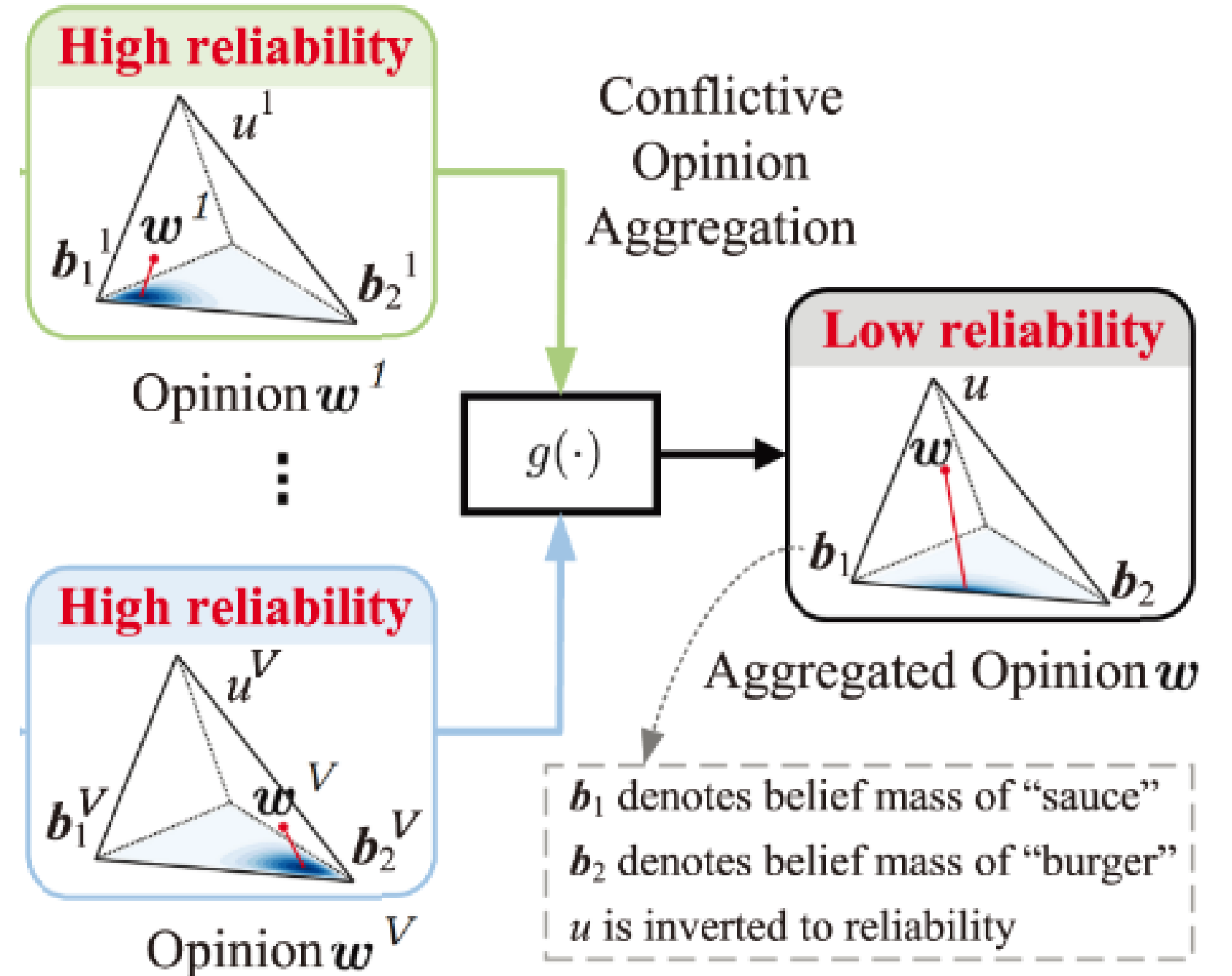
$$u^{A \diamond B} = \frac{2u^A u^B}{u^A + u^B}, \mathbf{a}_k^{A \diamond B} = \frac{\mathbf{a}_k^A + \mathbf{a}_k^B}{2}$$

Definition 2. Conflictive Degree

$$c(\mathbf{w}^A, \mathbf{w}^B) = c_p(\mathbf{w}^A, \mathbf{w}^B) \cdot c_c(\mathbf{w}^A, \mathbf{w}^B)$$

$$c_p(\mathbf{w}^A, \mathbf{w}^B) = \frac{\sum_{k=1}^K |p_k^A - p_k^B|}{2}$$

$$c_c(\mathbf{w}^A, \mathbf{w}^B) = (1 - u^A)(1 - u^B)$$

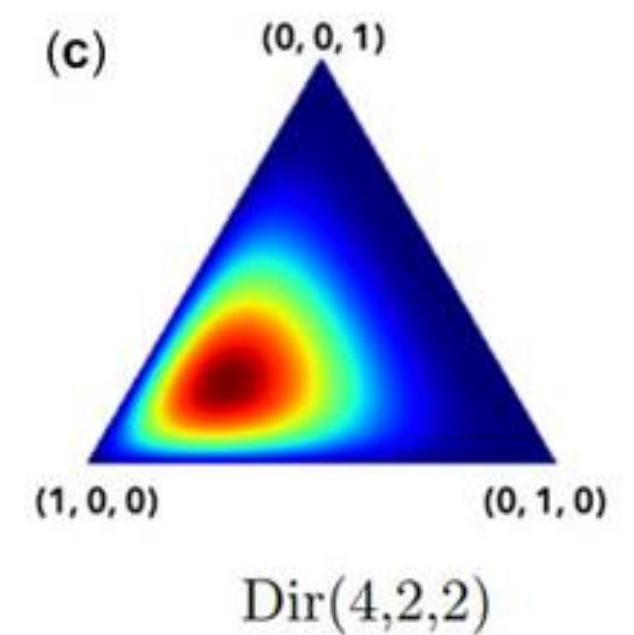
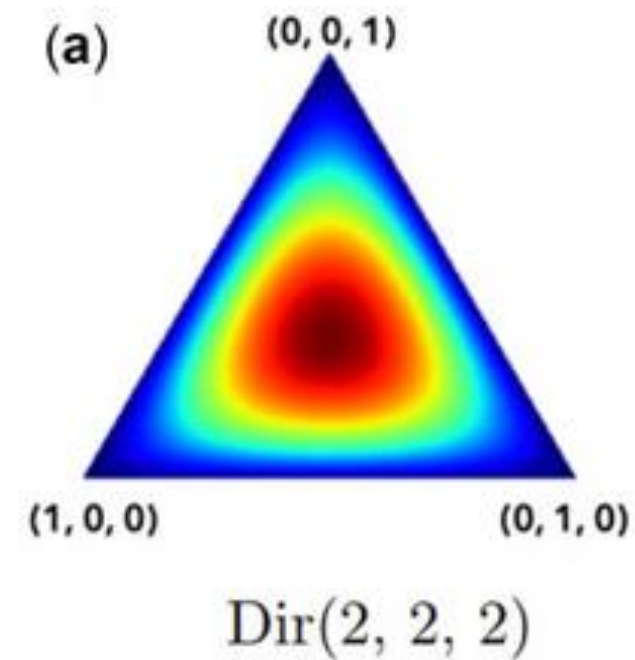


Method Loss Function

Adapted Cross-Entropy : 정답에 대한 증거가 많아지도록 유도

$$L_{ace}(\boldsymbol{\alpha}_n) = \int \left[\sum_{j=1}^K -y_{nj} \log p_{nj} \right] \frac{\prod_{j=1}^K p_{nj}^{\alpha_{nj}-1}}{B(\boldsymbol{\alpha}_n)} d\mathbf{p}_n$$

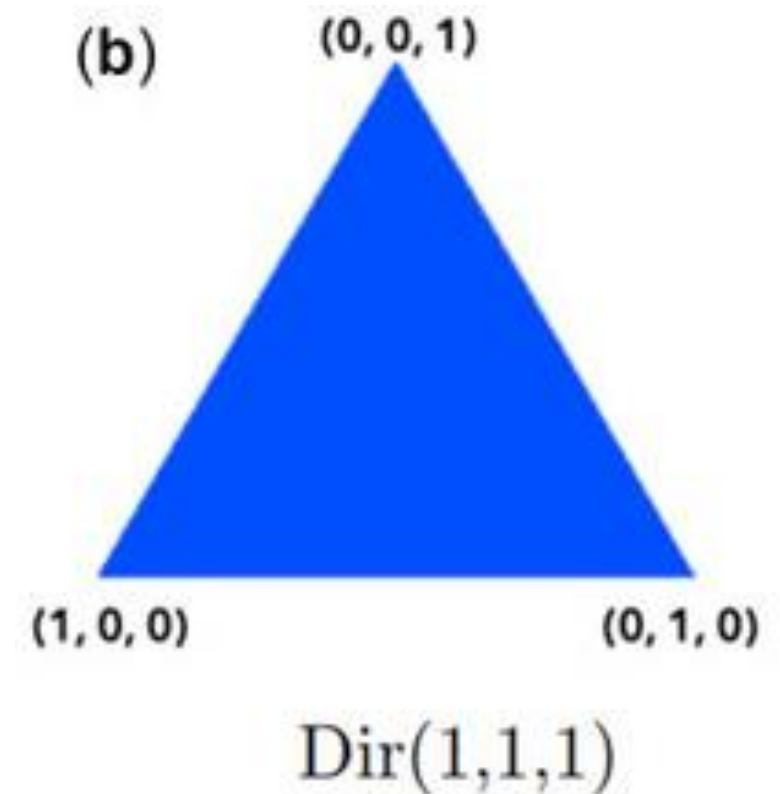
$$D(\mathbf{p}|\boldsymbol{\alpha}) = \begin{cases} \frac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^K p_k^{\alpha_k-1}, & \text{for } \mathbf{p} \in \mathcal{S}_K \\ 0, & \text{otherwise,} \end{cases}$$



Method Loss Function

KL Divergence : 잘못된 예측에 대해 높은 증거 값을 갖는 것을 방지

$$L_{KL}(\boldsymbol{\alpha}_n) = KL[D(\boldsymbol{p}_n|\tilde{\boldsymbol{\alpha}}_n) \parallel D(\boldsymbol{p}_n|\mathbf{1})]$$



Consistency Loss : 서로 다른 View가 일관되게 예측하도록 유도

$$L_{con} = \frac{1}{V-1} \sum_{p=1}^V \left(\sum_{q \neq p}^V c(\boldsymbol{w}_n^p, \boldsymbol{w}_n^q) \right) \quad c(\boldsymbol{w}^A, \boldsymbol{w}^B) = c_p(\boldsymbol{w}^A, \boldsymbol{w}^B) \cdot c_c(\boldsymbol{w}^A, \boldsymbol{w}^B)$$
$$c_p(\boldsymbol{w}^A, \boldsymbol{w}^B) = \frac{\sum_{k=1}^K |p_k^A - p_k^B|}{2}$$
$$c_c(\boldsymbol{w}^A, \boldsymbol{w}^B) = (1 - u^A)(1 - u^B)$$

Experiments

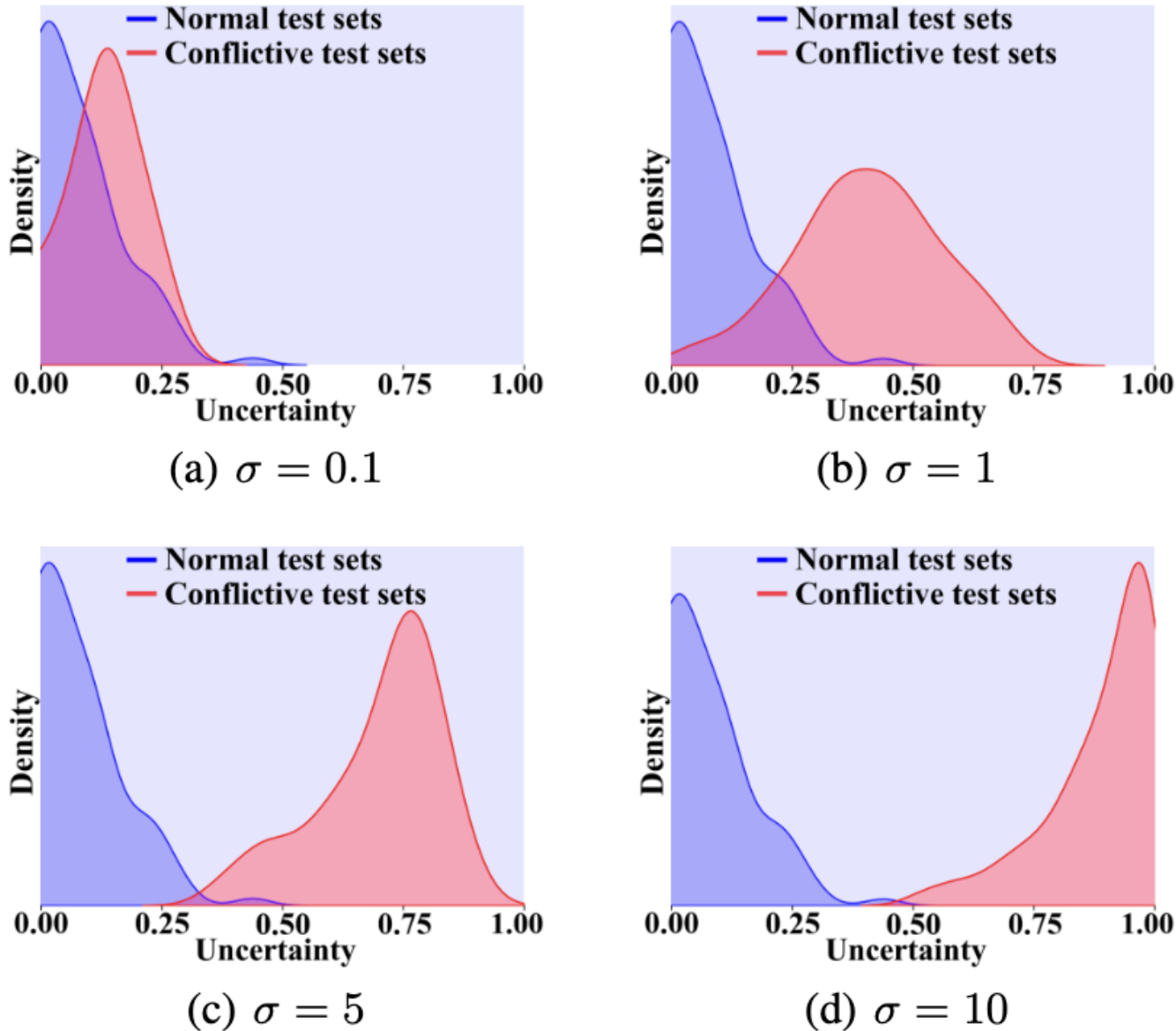
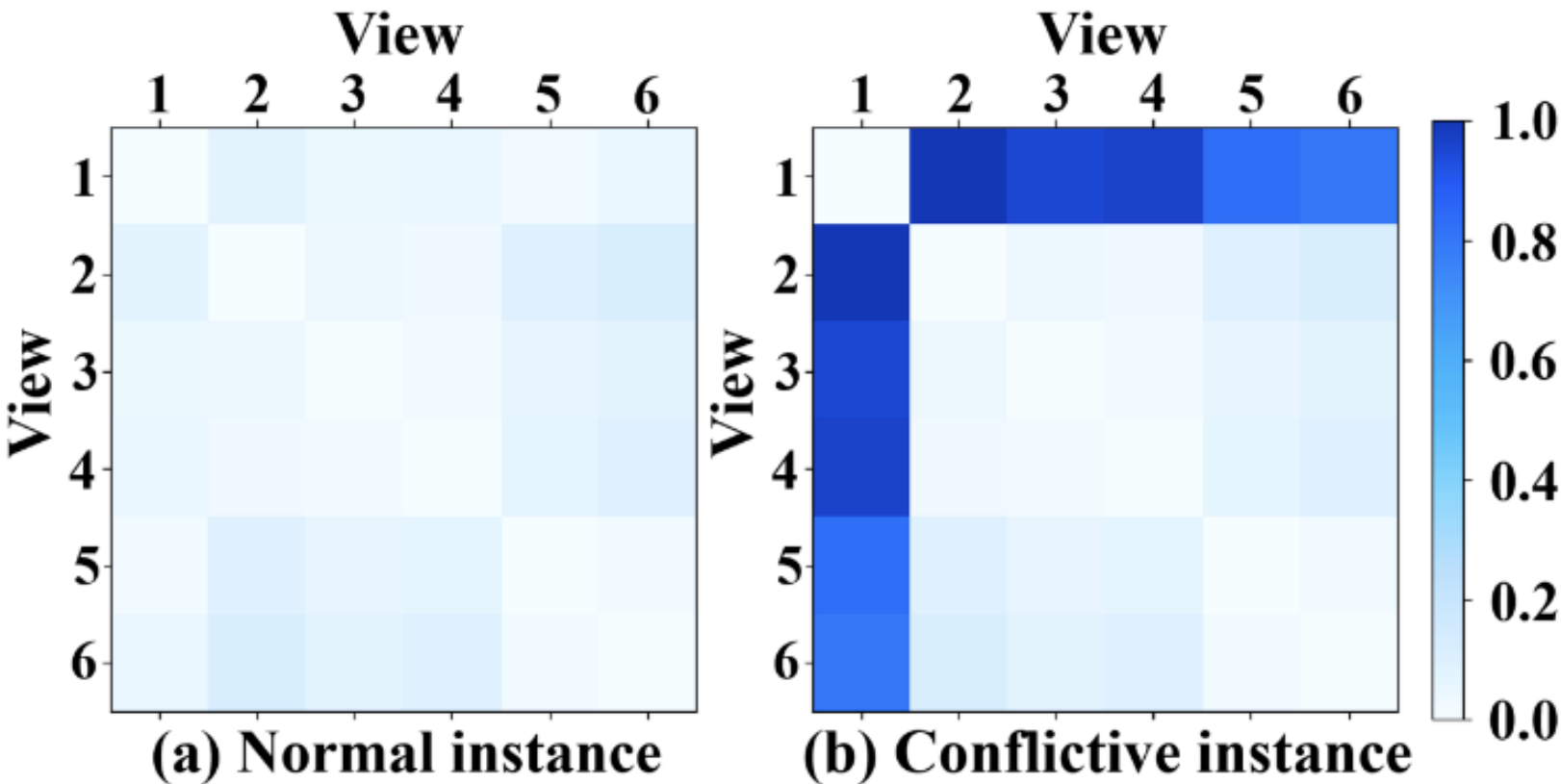
Result for Normal Dataset

Data	DCCAE	CPM-Nets	DUA-Nets	TMC	TMDL-OA	Ours	$\Delta\%$
HandWritten	95.45 \pm 0.35	94.55 \pm 1.36	98.10 \pm 0.32	98.51 \pm 0.13	99.25 \pm 0.45	99.40 \pm 0.00	0.15
CUB	85.39 \pm 1.36	89.32 \pm 0.38	80.13 \pm 1.67	90.57 \pm 2.96	95.43 \pm 0.20	98.50 \pm 2.75	3.21
HMDB	49.12 \pm 1.07	63.32 \pm 0.43	62.73 \pm 0.23	65.17 \pm 2.42	88.20 \pm 0.58	90.84 \pm 1.86	2.99
Scene15	55.03 \pm 0.34	67.29 \pm 1.01	68.23 \pm 0.11	67.71 \pm 0.30	75.57 \pm 0.02	76.19 \pm 0.12	0.82
Caltech101	89.56 \pm 0.41	90.35 \pm 2.12	93.43 \pm 0.34	92.80 \pm 0.50	94.63 \pm 0.04	95.36 \pm 0.38	0.77
PIE	81.96 \pm 1.04	88.53 \pm 1.23	90.56 \pm 0.47	91.85 \pm 0.23	92.33 \pm 0.36	94.71 \pm 0.02	2.57

- HandWritten:** 0부터 9까지의 손으로 쓴 아라비아 숫자 이미지 데이터셋
- CUB:** 다양한 종류의 새 이미지 데이터셋으로, 보통 이미지 자체 특징과 새의 속성(attribute) 정보를 다른 관점으로 사용.
- HMDB:** 사람의 다양한 행동을 담은 비디오 클립 데이터셋
- Scene15:** 15가지 다른 실내 및 실외 장면 카테고리(예: 부엌, 숲, 사무실 등)의 이미지 데이터셋.
- Caltech101:** 101가지 사물 카테고리(비행기, 얼굴, 오토바이 등)의 이미지 데이터셋.
- PIE (CMU PIE Face Database):** 다양한 조명, 표정, 포즈 변화를 포함하는 사람 얼굴 이미지 데이터셋

Experiments Result for Conflictive Dataset

Data	DCCAE	CPM-Nets	DUA-Nets	TMC	TMDL-OA	Ours	$\Delta\%$
HandWritten	82.85 \pm 0.38	83.34 \pm 1.07	87.16 \pm 0.34	92.76 \pm 0.15	93.05 \pm 0.05	94.40 \pm 0.05	1.45
CUB	63.57 \pm 1.28	68.82 \pm 0.17	60.53 \pm 1.17	73.37 \pm 2.16	74.43 \pm 0.26	76.50 \pm 1.15	2.78
HMDB	29.62 \pm 1.79	42.62 \pm 1.43	43.53 \pm 0.28	47.17 \pm 0.15	67.62 \pm 0.28	70.84 \pm 1.19	4.76
Scene15	25.97 \pm 2.86	29.63 \pm 1.12	26.18 \pm 1.31	42.27 \pm 1.61	48.42 \pm 1.02	56.97 \pm 0.52	17.66
Caltech101	60.90 \pm 2.32	66.54 \pm 2.89	75.19 \pm 2.34	90.16 \pm 2.50	90.63 \pm 2.05	92.36 \pm 1.48	1.91
PIE	26.89 \pm 1.10	53.19 \pm 1.17	56.45 \pm 1.75	61.65 \pm 1.03	68.16 \pm 0.34	84.00 \pm 0.14	23.24



Conclusion Why Best Paper?

- 충돌적 다중 뷰 데이터에 대한 결정 결과와 관련 신뢰도를 명시적으로 제공해야 하는 RCML 문제를 새롭게 정의함.
- 증거적 학습과 새로운 충돌적 의견 집계 전략을 사용하는 ECML 방법을 제안함.
- 다중 뷰들의 충돌 시 불확실성 증가 가능성을 효과적으로 모델링함
- 6개 데이터셋에 대한 실험을 통해 정확도, 신뢰도 추정, 강건성 측면에서 ECML의 효과를 실증적으로 검증함.