ICML | 2025
The Forty-second International Conference on Machine Learning

TIANFU JINCHENG LABORATORY
天府锦城实验室
(前沿医学中心)

Deep Fuzzy Multi-view Learning for Reliable Classification

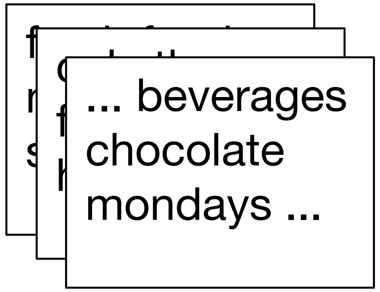
Siyuan Duan, Yuan Sun, Dezhong Peng, Guiduo Duan, Xi Peng, Peng Hu #

College of Computer Science, Sichuan University, Chengdu, China
Tianfu Jincheng Laboratory, Chengdu, China

Github repo

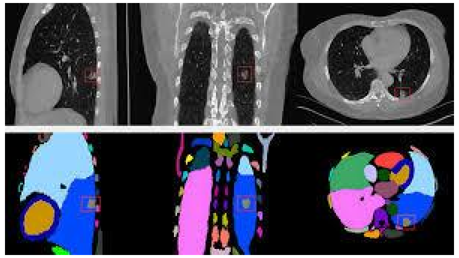
Background – Multimodal Data

Images

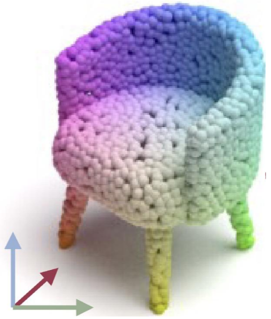


Texts

Videos



Medical Images

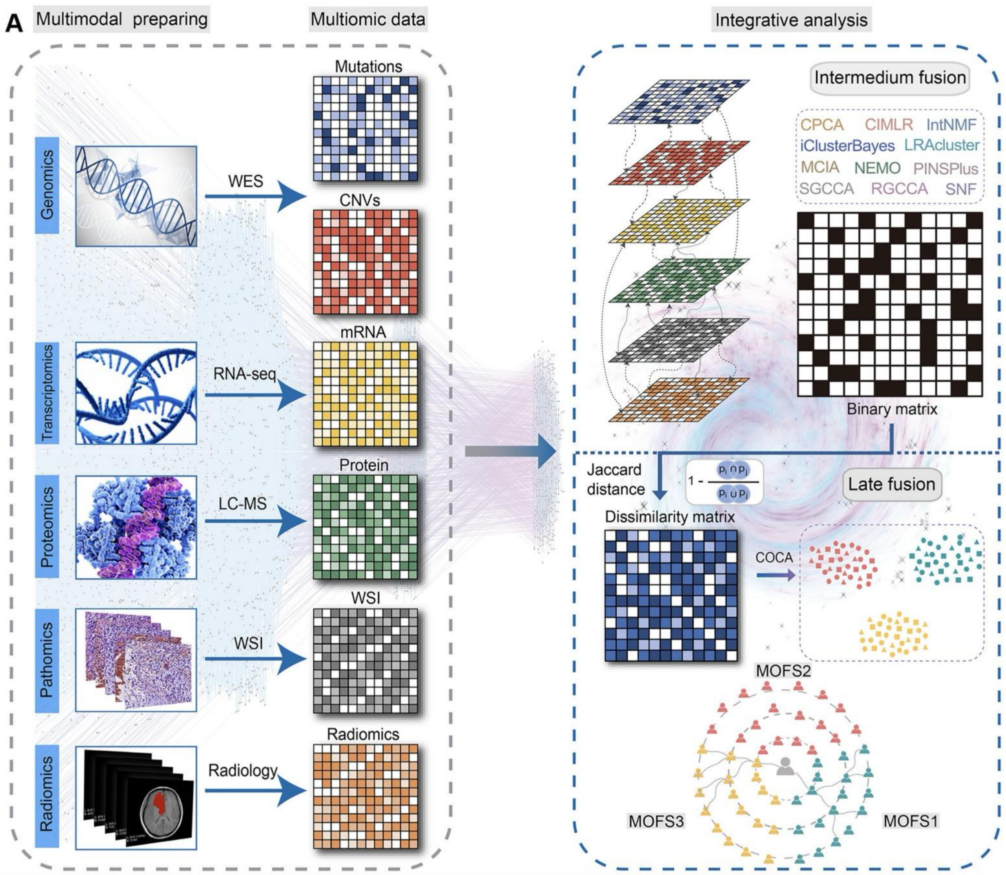


3D cloud point

Audio

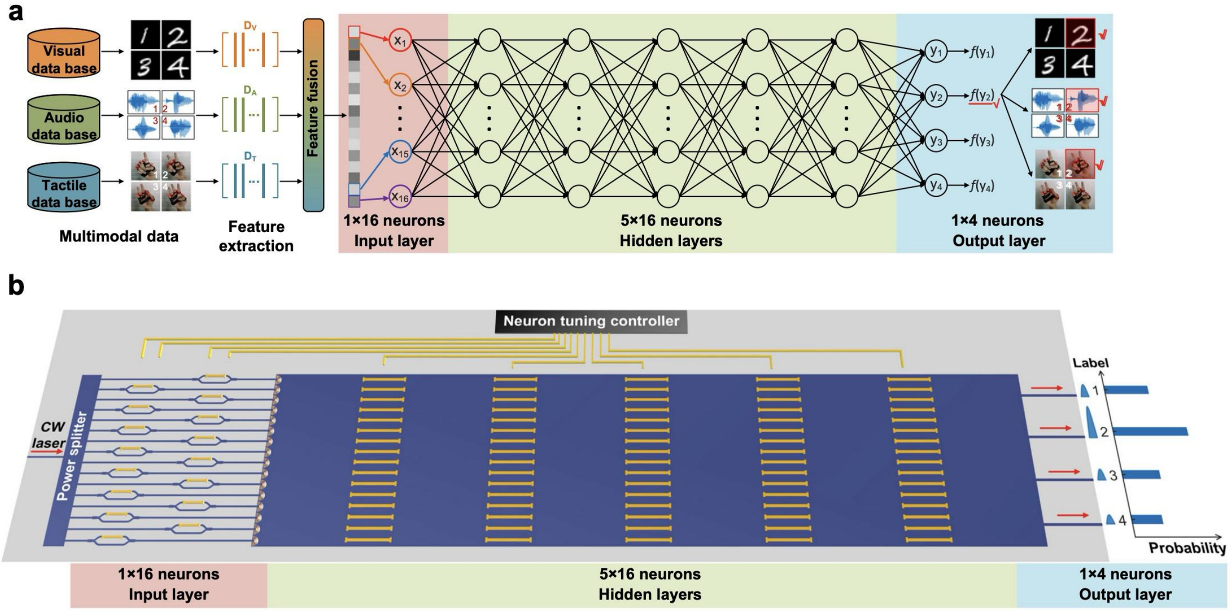
Background - Multimodal Application Scenarios

Smart Healthcare



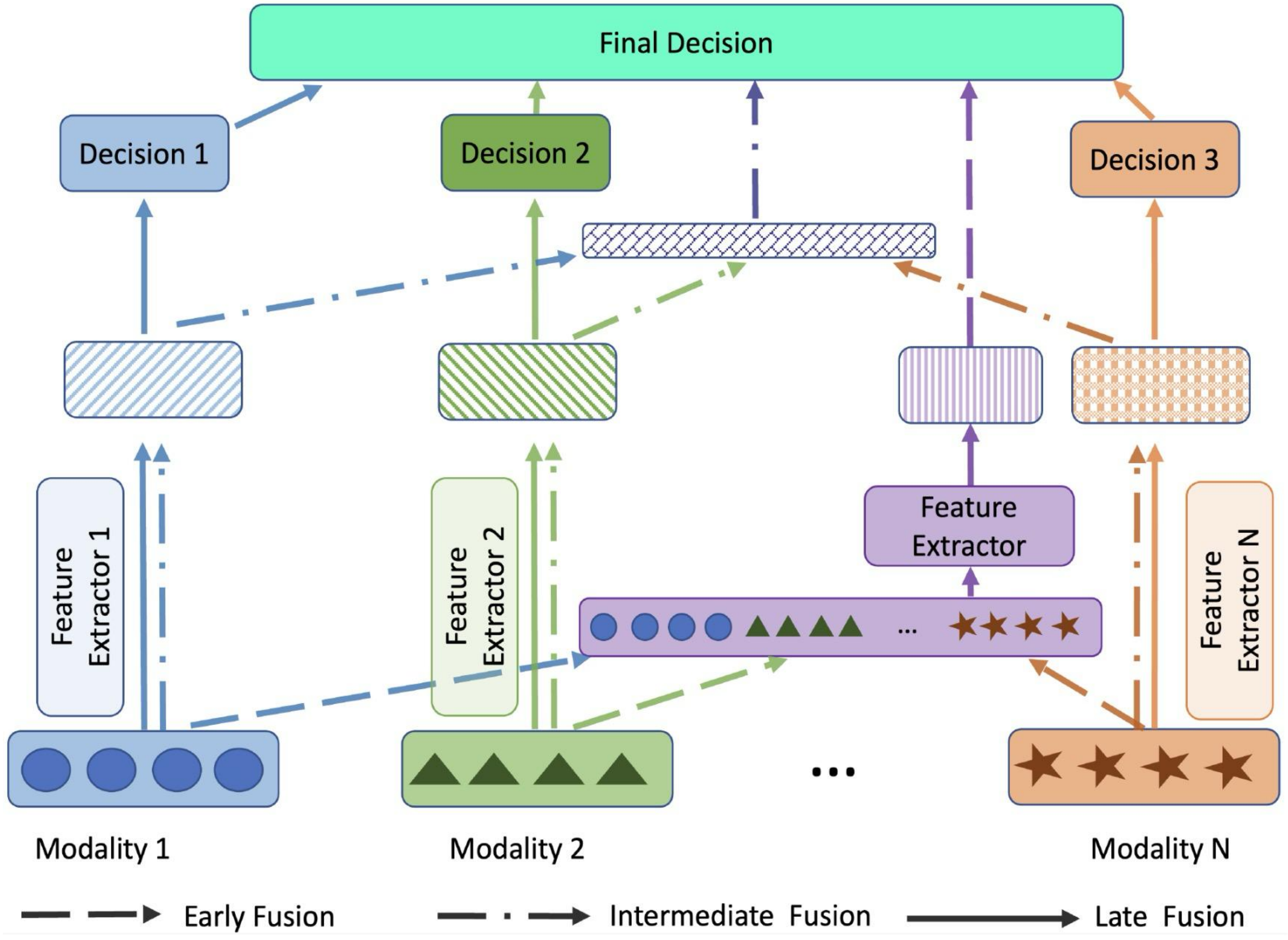
Liu, Zaoqu, et al. "Multimodal fusion of radio-pathology and proteogenomics identify integrated glioma subtypes with prognostic and therapeutic opportunities." *Nature Communications* 16.1 (2025): 3510.

Chip Design and Production



Cheng, Junwei, et al. "Multimodal deep learning using on-chip diffractive optics with in situ training capability." *Nature Communications* 15.1 (2024): 6189.

Background - Multi-view/modal Classification



Zhao, Fei, Chengcui Zhang, and Baocheng Geng. "Deep multimodal data fusion." *ACM computing surveys* 56.9 (2024): 1-36.

Motivations - Conflicting Multi-view Data

Normal multi-view instance

Text view

...This Bathroom is really clean ...

RGB view

Depth view

Conflicting multi-view instance

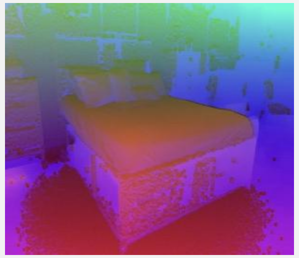
Text view

...This Bathroom is really clean ...

RGB view

Conflicting !

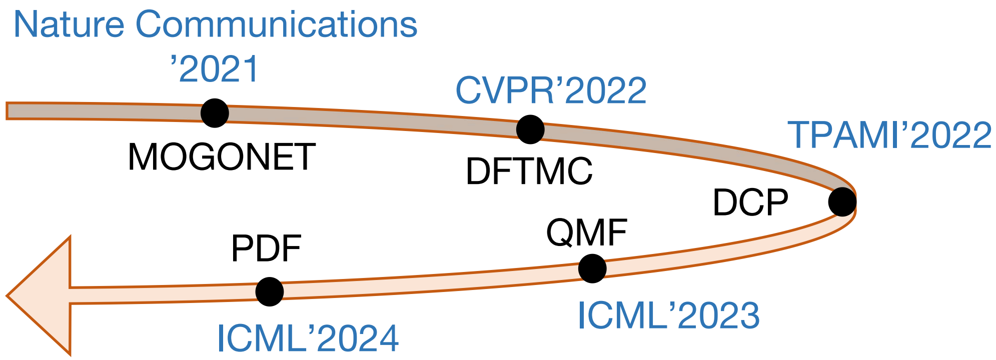
Depth view

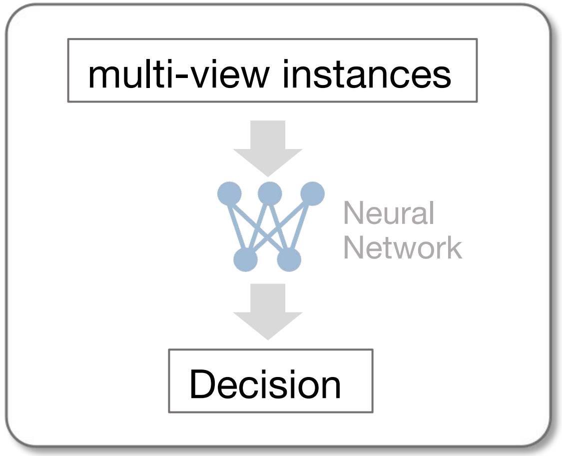


Key question: How to correctly and reliably classify conflicting multi-view instances ?

Related Works - Multi-view/modal Learning

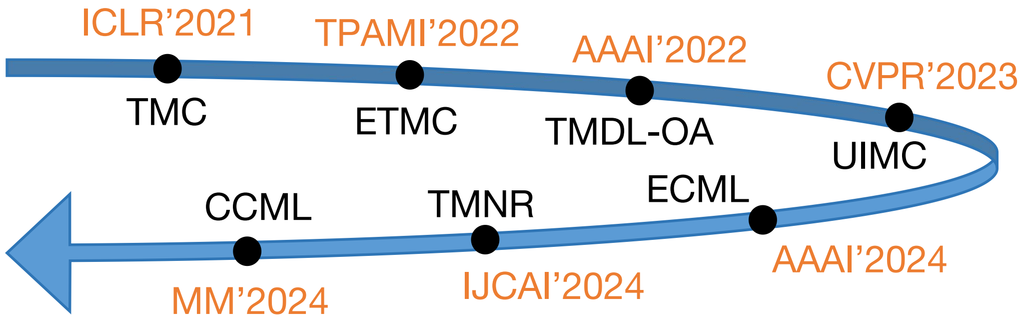
1.Untrusted Methods

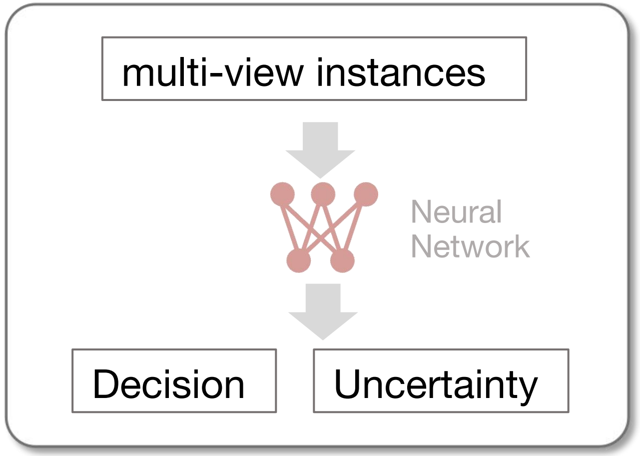




Shotcoming:
Can't estimate uncertainty.

2.EDL*-based Trusted Methods





Shotcoming

1.Ignore the global conflicts between views, which often leads to incorrect classification.

2.Ignore the inherent conflicts between views and under-estimates the uncertainty of conflicting multi-view instances.

*EDL: Evidential Deep Learning

Related Works - Uncertainty Estimation

- Existing Methods：
  - Bayesian（high computational overhead）、
  - MC-dropout （high computational overhead）、
  - Deep Ensembles （high computational overhead）、
  - Evidential Deep Learning （EDL） （low computational overhead，but not accurate！ ）

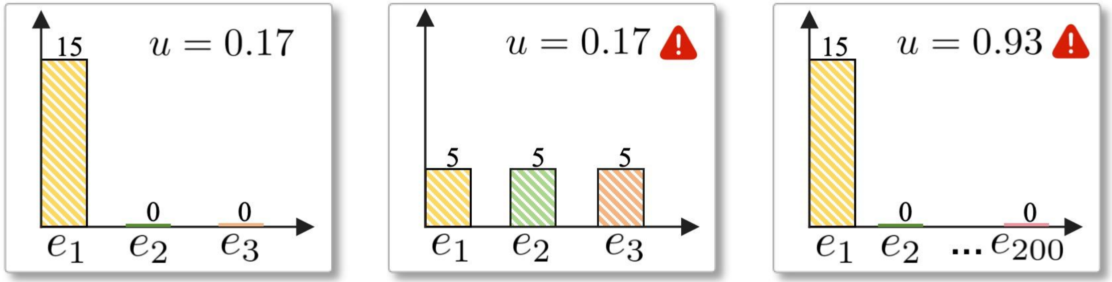
Uncertainty

Class count

$$u = \frac{K}{\sum_{i}^{K} e_i + K}$$

Total evidence

EDL's Uncertainty calculation formula



Confident prediction ✓ | Conflicting evidence — Underestimated | Confident prediction — Overestimated

Duan, Siyuan, et al. "Fuzzy Multimodal Learning for Trusted Cross-modal Retrieval." *Proceedings of the Computer Vision and Pattern Recognition Conference*. 2025.

Preliminary – Category Credibility Modeling

- **Fuzzy Membership**: a numerical value used to measure the degree to which an element belongs to a fuzzy set, and its value range is between 0 and 1. The membership of the v-th view of the i-th sample to different categories:

$$m_{i1}^v, m_{i2}^v, \ldots, m_{iK}^v \quad \text{(only provide possibility measure)}$$

- **Necessity**: the certainty that the sample does not belong to other categories:

$$e_{ik}^v = 1 - \max\{m_{il}^v \mid l \neq k\},\ k = 1, \ldots, K, \quad \text{(provide necessity measure)}$$

- **Category Credibility:**

$$c_{ik}^v = \frac{1}{2}(m_{ik}^v + 1 - \max\{m_{il}^v \mid l \neq k\}),\ k = 1, 2, \ldots, K,$$

Method - Category Credibility Learning

➢ Membership Modeling:

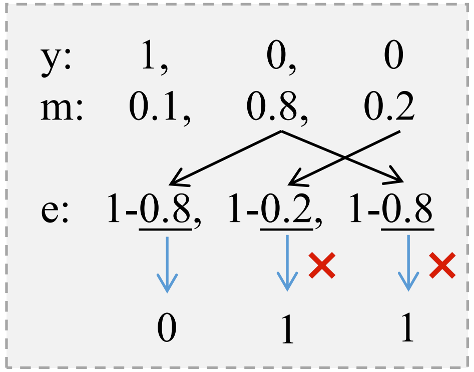
$$\mathbf{m}_i^v = \mathrm{ReLU}\left(\frac{\mathbf{a}_i^v}{||\mathbf{a}_i^v||_p}\right)$$

$\mathbf{a}_i^v \rightarrow$ logits of neural network

➢ Category Credibility Learning:

$$c_{ik}^v = \frac{1}{2}(m_{ik}^v + 1 - \max\{m_{il}^v \mid l \neq k\}),$$

$||\mathbf{c}_i^v - \mathbf{y}_i^v||_2$ or $-\mathbf{y}_i^v \cdot \log(\mathbf{c}_i^v) - (1 - \mathbf{y}_i^v) \cdot \log(1 - \mathbf{c}_i^v)$



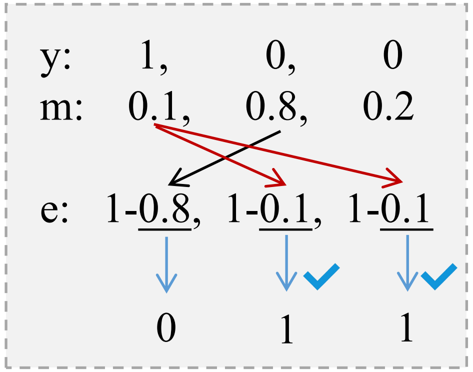
if $y_{ik}^v = 0$ and $y_{il}^v = 0$

$m_{ik}^v \rightarrow 0$

$1 - \max\{m_{il}^v \mid l \neq k\} \rightarrow 0$

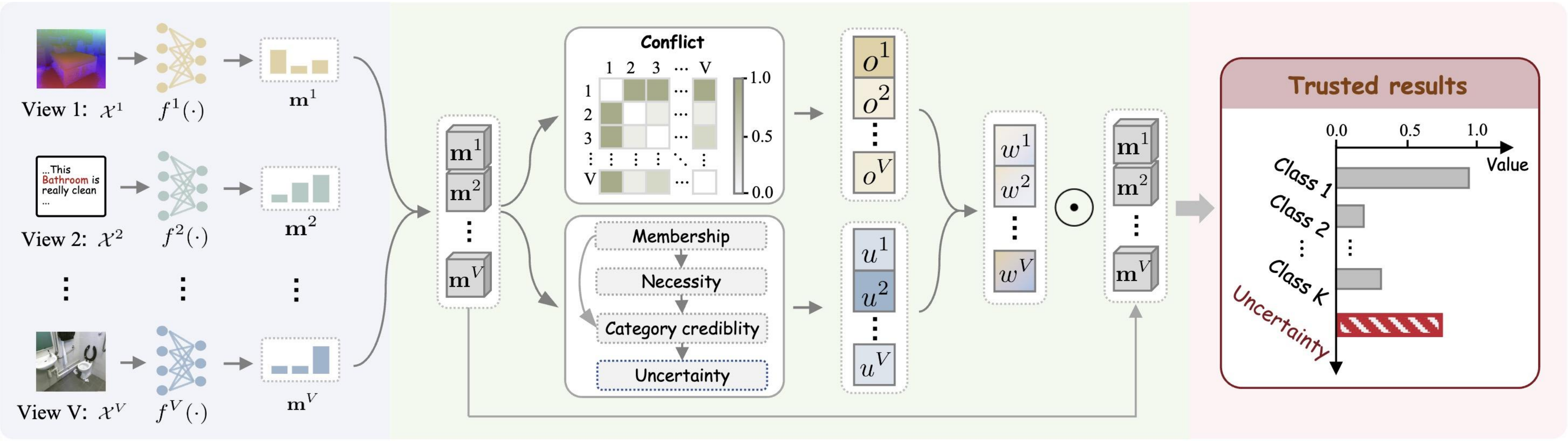
$$r_{ik}^v = \begin{cases} \dfrac{m_{ik}^v + 1 - \max\{m_{il}^v \mid l \neq k\}}{2}, & \text{if } y_{ik}^v = 1, \\ \dfrac{m_{ik}^v + 1 - m_{il}^v}{2}, & \text{if } y_{ik}^v = 0,\ l = \arg\max_k\ y_{ik}^v, \end{cases}$$

$r_{ik}^v \downarrow$ Category credibility during training



$$\mathcal{L}_{ccl} = \frac{1}{N_b}\sum_{i=1}^{N_b} -\mathbf{y}_i^v \cdot \log(\mathbf{r}_i^v) - (1 - \mathbf{y}_i^v) \cdot \log(1 - \mathbf{r}_i^v),$$

Method - Framework



Conflict:

$$o_i^v = \frac{1}{V-1}\sum_{j\neq v}^{V}\left(1 - \frac{\mathbf{m}_i^v \cdot \mathbf{m}_i^j}{||\mathbf{m}_i^v|| \cdot ||\mathbf{m}_i^j||}\right)$$

Uncertainty:

$$u_i^v = \frac{\sum_{k=1}^{K} H(c_{ik}^v)}{K \cdot \ln 2}$$
$$= \frac{\sum_{k=1}^{K} -c_{ik}^v \cdot ln(c_{ik}^v) - (1-c_{ik}^v)\cdot ln(1-c_{ik}^v)}{K \cdot \ln 2}$$

Dual-reliable Multi-view Fusion:

$$w_i^v = \frac{g\left((1-u_i^v)(1-o_i^v)\right)}{\sum_{v=1}^{V} g\left((1-u_i^v)(1-o_i^v)\right)},$$

$$\mathbf{m}_i^a = \sum_{v=1}^{V} w_i^v \cdot \mathbf{m}_i^v,$$

Loss function:

$$\mathcal{L}_{total} = \mathcal{L}_{ccl}(\mathbf{r}^a, \mathbf{y}) + \sum_{v=1}^{V} \mathcal{L}_{ccl}(\mathbf{r}^v, \mathbf{y})$$

Experiments - Setup

Datasets

Table 1. A summary of datasets used for evaluation.

DATASET	SIZE	CATEGORIES	DIMENSIONALITY
HW	2000	10	240; 76; 216; 47; 64; 6
MSRC	210	7	1302; 48; 512; 100; 256; 210
NUSOBJ	30000	31	65; 226; 145; 74; 129
FASHION	10000	10	784; 784; 784
SCENE	4485	15	20; 59; 40
LANDUSE	2100	21	20; 59; 40
LEAVES	1600	100	64; 64; 64
PIE	680	68	484; 256; 279

How to add conflict ?

① We add Gaussian noise with different standard deviations to some testing instances.

② We alter the information in a random view for a subset of instances, making the view's label inconsistent with the true label.

- Training set:test set=8:2；
- Take the best performance at each training time;
- 10 runs with different random seeds.

Experiments – Comparison with SOTA

Tab1. Performance on normal datasets. Methods marked with ★ are trusted.

METHODS.	HW	MSRC	NUSOBJ	FASHION	SCENE	LANDUSE	LEAVES	PIE
DFTMC	98.75±0.39	96.90±2.14	-	-	63.10±3.60	34.95±1.69	69.92±2.54	91.40±3.50
DCP-CV	98.75±0.59	92.86±2.61	32.19±9.48	97.96±0.16	76.70± 2.15	71.71±2.09	95.62±1.38	86.32±4.87
DCP-CG	99.00±0.47	95.24±3.69	43.65±1.10	98.11± 0.23	77.79±1.73	75.74±0.98	98.19±0.46	90.59±1.99
QMF	98.72±0.48	97.86±1.28	45.41±0.43	98.93±0.32	68.58±1.49	47.86±2.55	95.69±1.25	92.06±1.64
PDF	98.40±0.37	97.14±1.78	46.78±0.33	98.95±0.19	70.25±1.21	45.17±2.66	98.03±0.71	92.57±1.66
DUA-NETS★	98.10±0.32	84.67±3.03	27.75±0.00	91.08±0.17	65.01±1.55	45.24±1.85	90.31±1.25	90.56±0.47
TMC★	98.51±0.15	91.70±2.70	38.77±0.81	95.40±0.40	67.71±0.30	31.69±3.93	86.81±2.20	91.85±0.23
ETMC★	98.75±0.00	92.86±3.01	44.23±0.76	96.21±0.36	71.61±0.28	43.52±3.19	91.44±2.39	93.75±1.08
TMDL-OA★	98.55±0.45	95.00±1.67	27.88±0.67	86.52±0.04	75.57±0.02	25.02±2.10	75.28±3.57	92.33±0.36
UIMC★	98.25±0.00	98.81±1.19	43.42±0.12	98.13±0.13	77.70±0.00	57.95±0.61	95.31±0.71	91.69±2.16
ECML★	98.72±0.39	94.05±1.60	42.62±0.42	97.93±0.35	76.19±0.12	60.10±2.01	92.53±1.94	94.71±0.02
TMNR★	97.20±0.63	94.05±3.24	34.52±0.85	94.10±0.50	68.10±1.15	27.38±1.88	90.13±1.53	89.53±1.89
CCML★	97.60±0.62	96.90±2.39	41.43±0.71	95.16±0.41	73.87±1.83	60.86±1.93	97.72±0.92	93.97±1.67
FUML★	**99.20±0.36**	**99.76±0.75**	**48.23±0.42**	**98.96±0.25**	**79.41±1.34**	**76.71±0.46**	**99.78± 0.27**	**96.18±1.24**
IMPROVE	**△ 0.20**	**△ 0.95**	**△ 1.45**	**△ 0.01**	**△ 1.62**	**△ 0.97**	**△ 1.59**	**△ 1.47**

Tab2. Performance on conflicting datasets. Methods marked with ★ are trusted.

METHODS	HW	MSRC	NUSOBJ	FASHION	SCENE	LANDUSE	LEAVES	PIE
DFTMC	53.65±20.07	60.24±23.45	-	-	36.01± 2.78	7.88±0.94	1.10±0.12	3.97±0.82
DCP-CV	98.20±0.56	84.76±7.00	28.10±7.80	92.72±2.41	66.22±2.12	59.98±1.93	76.94±1.36	67.06±2.15
DCP-CG	98.70±0.64	90.00±1.78	38.61±1.29	90.38±2.17	66.44±0.32	61.83±2.48	79.06±1.22	69.56±3.77
QMF	97.52±0.86	95.95±1.52	42.72±0.67	92.69±0.78	59.53±1.63	40.17±2.67	77.47±1.46	82.50±2.81
PDF	94.35±1.21	94.52±3.02	43.57±0.36	90.73±0.53	58.75±1.03	39.40±1.94	76.34±1.26	74.93±2.76
DUA-NETS★	87.16±0.34	78.57±4.45	25.64±0.25	83.03±0.18	26.18±1.31	37.22±0.56	65.62±2.19	56.45±1.75
TMC★	92.76±0.15	86.20±4.90	36.00±0.78	84.76±0.78	42.27±1.61	19.67±1.88	70.25±2.55	61.65±1.03
ETMC★	93.85±1.26	87.14±4.54	40.45±0.81	86.48±1.05	56.90±1.70	36.05±2.50	74.19±1.74	73.82±4.77
TMDL-OA★	92.45±0.05	84.52±2.20	27.02±0.75	74.55±0.07	48.42±1.02	21.71±1.83	62.28±3.70	68.16±0.34
UIMC★	97.72±0.18	96.43±1.19	41.72±0.31	89.71±0.25	67.88±0.48	50.43±0.46	79.84±0.92	70.66±2.04
ECML★	94.52±0.79	90.00±2.78	39.89±0.59	84.02±0.51	56.97±0.52	50.31±1.81	74.88±1.89	84.00±0.14
TMNR★	92.78±1.01	90.71±4.19	30.88±0.58	85.76±0.81	60.00±1.43	23.95±1.92	74.09±1.99	80.59±3.26
CCML★	93.22±1.09	94.29±2.18	37.38±0.65	83.84±1.01	62.08±1.34	52.48±2.74	78.87±2.31	83.24±2.79
FUML★	**98.78±0.36**	**98.81±1.60**	**47.08±0.32**	**96.68±0.32**	**72.71±1.75**	**69.14±2.43**	**94.44±1.18**	**88.01±2.53**
IMPROVE	**△ 0.08**	**△ 2.38**	**△ 3.51**	**△ 3.96**	**△ 4.83**	**△ 7.31**	**△ 14.60**	**△ 4.01**

Experiments – Uncertainty Effectiveness Analysis

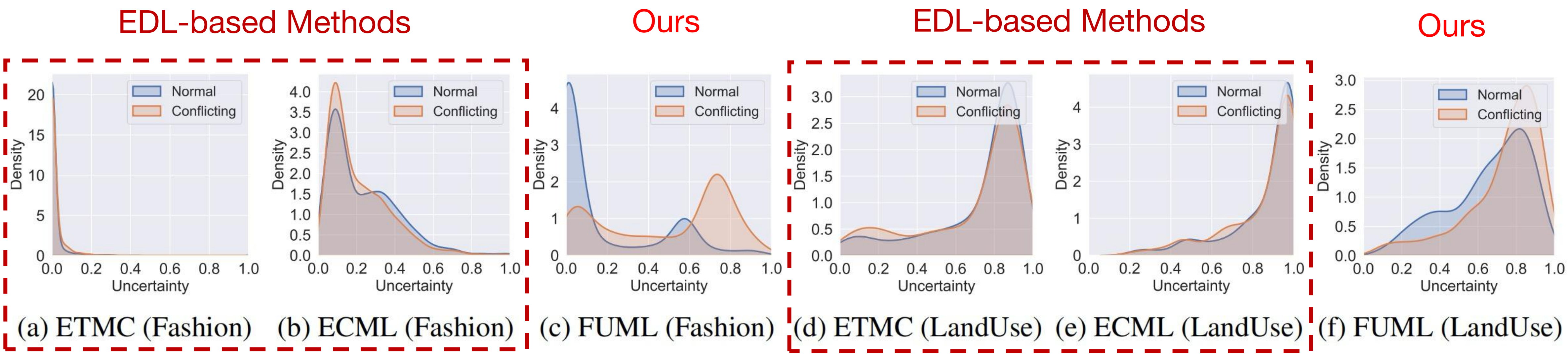
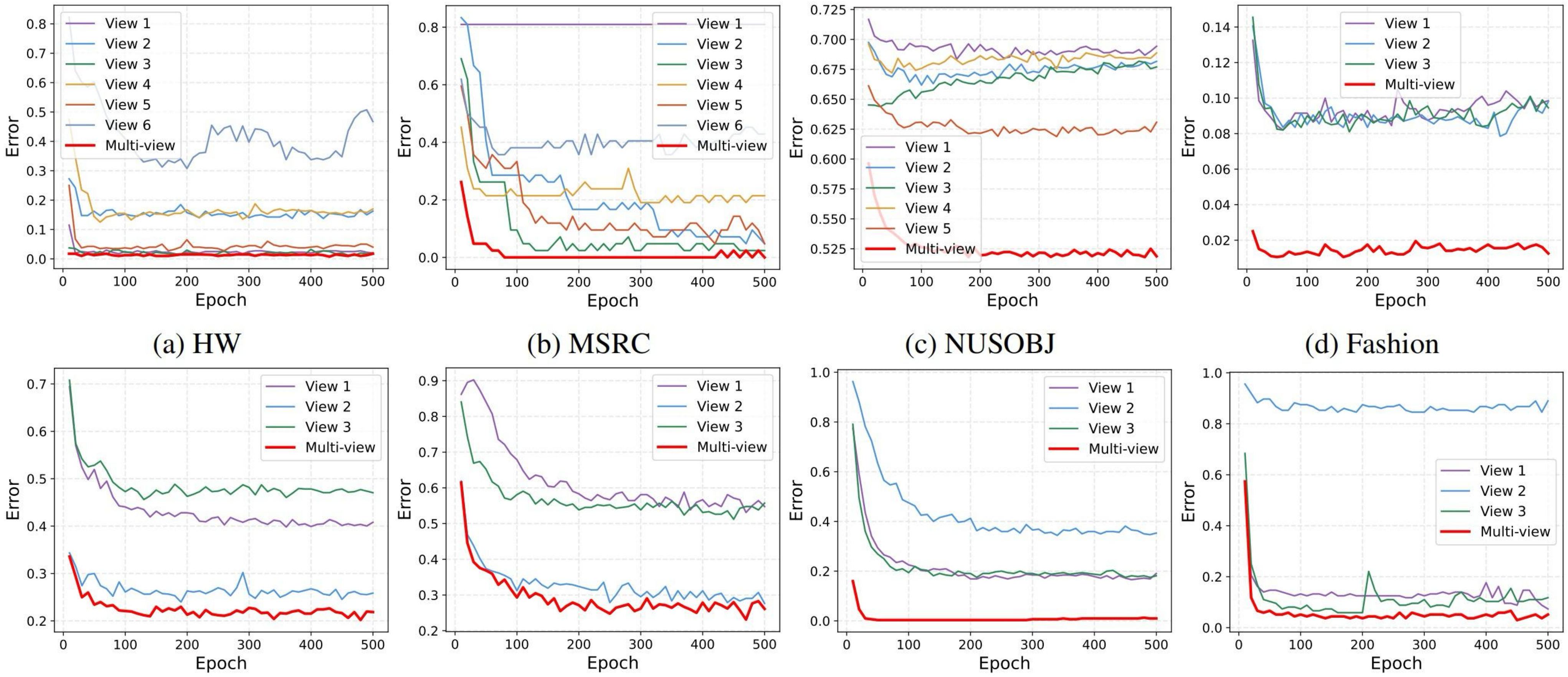


Figure 3. Density of uncertainty on the normal and conflicting testing sets of Fashion-MV and LandUse datasets.

Experiments - Multi-view Fusion Effectiveness Evaluation



(a) HW (b) MSRC (c) NUSOBJ (d) Fashion

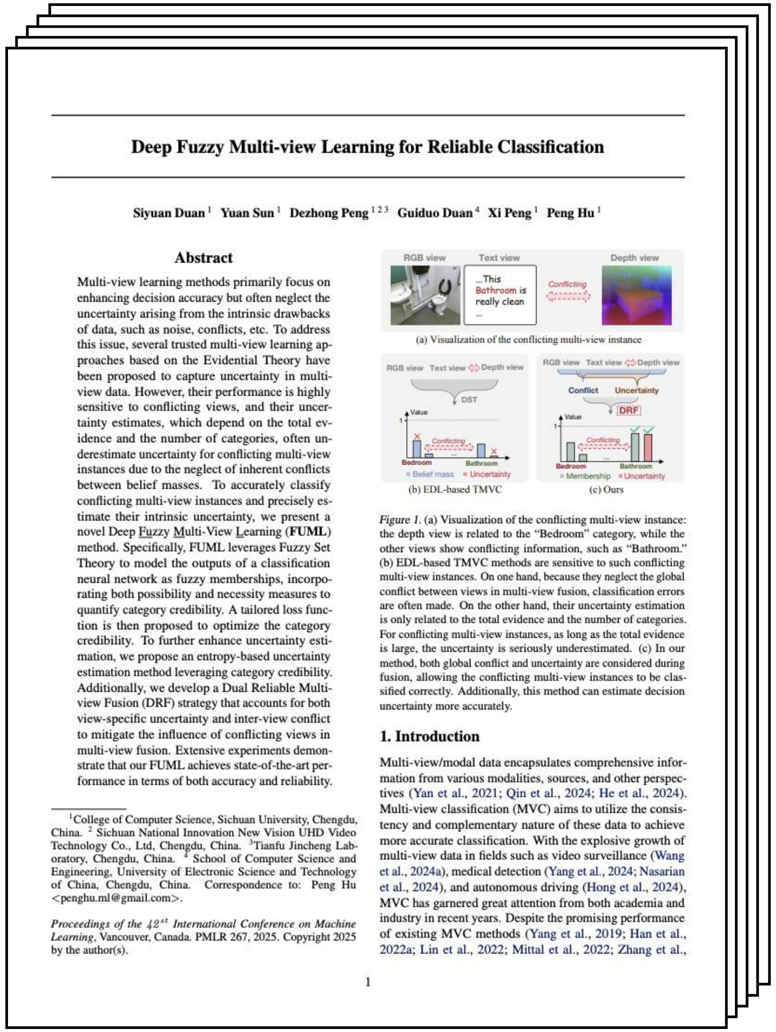
(e) Scene (f) LandUse (g) Leaves (h) PIE

Experiments - Ablation Study

Table 6. Classification accuracy (ACC), Precision (Prec.), and F-score of FUML with different combination rules on the normal and conflicting testing sets. All metrics are expressed as percentages (%). The best results are highlighted in **boldface**.

METHOD \ DATASET			FASHION (NORMAL)			LANDUSE (NORMAL)			FASHION (CONFLICTING)			LANDUSE (CONFLICTING)		
$\mathcal{L}^a_{ccl}$	$\mathcal{L}^v_{ccl}$	RULE	ACC↑	PREC.↑	F-SCORE↑	ACC↑	PREC.↑	F-SCORE↑	ACC↑	PREC.↑	F-SCORE↑	ACC↑	PREC.↑	F-SCORE↑
√	×	DRF	98.66	98.66	98.48	76.29	76.89	76.09	96.33	96.31	96.30	68.00	68.72	67.53
×	√	DRF	98.73	98.73	98.73	75.71	76.14	75.24	96.32	96.32	96.30	67.76	68.41	67.33
√	×	CONCAT	95.71	95.71	95.70	72.26	72.44	71.55	88.45	88.28	88.75	59.69	60.54	58.54
√	√	AVG	98.50	98.51	98.51	76.19	77.14	76.07	96.15	96.13	96.13	67.71	68.22	67.42
√	√	DRF	**98.96**	**98.97**	**98.96**	**76.71**	**77.57**	**76.48**	**96.68**	**96.68**	**96.68**	**69.14**	**70.21**	**69.19**

Conclusions


Deep Fuzzy Multi-view Learning for Reliable Classification

Siyuan Duan [1] Yuan Sun [1] Dezhong Peng [1 2 3] Guiduo Duan [4] Xi Peng [1] Peng Hu [1]

Abstract

Multi-view learning methods primarily focus on enhancing decision accuracy but often neglect the uncertainty arising from the intrinsic drawbacks of data, such as noise, conflicts, etc. To address this issue, several trusted multi-view learning approaches based on the Evidential Theory have been proposed to capture uncertainty in multi-view data. However, their performance is highly sensitive to conflicting views, and their uncertainty estimates, which depend on the total evidence and the number of categories, often underestimate uncertainty for conflicting multi-view instances due to the neglect of inherent conflicts between belief masses. To accurately classify conflicting multi-view instances and precisely estimate their intrinsic uncertainty, we present a novel Deep Fuzzy Multi-View Learning (**FUML**) method. Specifically, FUML leverages Fuzzy Set Theory to model the outputs of a classification neural network as fuzzy memberships, incorporating both possibility and necessity measures to quantify category credibility. A tailored loss function is then proposed to optimize the category credibility. To further enhance uncertainty estimation, we propose an entropy-based uncertainty estimation method leveraging category credibility. Additionally, we develop a Dual Reliable Multi-view Fusion (DRF) strategy that accounts for both view-specific uncertainty and inter-view conflict to mitigate the influence of conflicting views in multi-view fusion. Extensive experiments demonstrate that our FUML achieves state-of-the-art performance in terms of both accuracy and reliability.

(a) Visualization of the conflicting multi-view instance

(b) EDL-based TMVC

(c) Ours

Figure 1. (a) Visualization of the conflicting multi-view instance: the depth view is related to the "Bedroom" category, while the other views show conflicting information, such as "Bathroom." (b) EDL-based TMVC methods are sensitive to such conflicting multi-view instances. On one hand, because they neglect the global conflict between views in multi-view fusion, classification errors are often made. On the other hand, their uncertainty estimation is only related to the total evidence and the number of categories. For conflicting multi-view instances, as long as the total evidence is large, the uncertainty is seriously underestimated. (c) In our method, both global conflict and uncertainty are considered during fusion, allowing the conflicting multi-view instances to be classified correctly. Additionally, this method can estimate decision uncertainty more accurately.

1. Introduction

Multi-view/modal data encapsulates comprehensive information from various modalities, sources, and other perspectives (Yan et al., 2021; Qin et al., 2024; He et al., 2024). Multi-view classification (MVC) aims to utilize the consistency and complementary nature of these data to achieve more accurate classification. With the explosive growth of multi-view data in fields such as video surveillance (Wang et al., 2024a), medical detection (Yang et al., 2024; Nasarian et al., 2024), and autonomous driving (Hong et al., 2024), MVC has garnered great attention from both academia and industry in recent years. Despite the promising performance of existing MVC methods (Yang et al., 2019; Han et al., 2022a; Lin et al., 2022; Mittal et al., 2022; Zhang et al.,

[1]College of Computer Science, Sichuan University, Chengdu, China. [2] Sichuan National Innovation New Vision UHD Video Technology Co., Ltd, Chengdu, China. [3]Tianfu Jincheng Laboratory, Chengdu, China. [4] School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China. Correspondence to: Peng Hu <penghu.ml@gmail.com>.

Proceedings of the 42st International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025. Copyright 2025 by the author(s).

1

- We reveal and address the conflict sensitivity issue in existing EDL-based TMVC methods, proposing FUML, a novel framework based on Fuzzy Set Theory for enhancing classification and uncertainty estimation.
- We develop a Dual-reliable Multi-view Fusion (DRF) strategy that effectively reduces the impact of conflicting views, embracing more robust multi-view classification.
- We propose an entropy-based uncertainty qualification technique, enabling more accurate uncertainty estimation for conflicting multi-view instances.
- We conduct extensive experiments comparing our FUML against 13 state-of-the-art MVC baselines on eight widely-used benchmarks, demonstrating superior accuracy, reliability, and robustness.

四川大學
SICHUAN UNIVERSITY

End
Thanks!