Multimodal Learning with Uncertainty Quantification based on Discounted Belief Fusion

Grigor Bezirganyan[1] grigor.bezirganyan@univ-amu.fr

Sana Sellami[1] sana.sellami@univ-amu.fr

Laure Berti-Équille[2] laure.berti@ird.fr

Sébastien Fournier[1] sebastien.fournier@univ-amu.fr

[1]Aix-Marseille Univ, CNRS, LIS Marseille, France

[2]IRD, ESPACE-DEV Montpellier, France

Scan for Paper

Context

- **Multimodal AI** combines diverse data (images, text, audio) for critical decisions.
- **Conflict** between modalities can arise when modalities **confidently disagree**.
- **Unaddressed conflict** leads to unreliable predictions—risky in safety-critical tasks.

Limitations

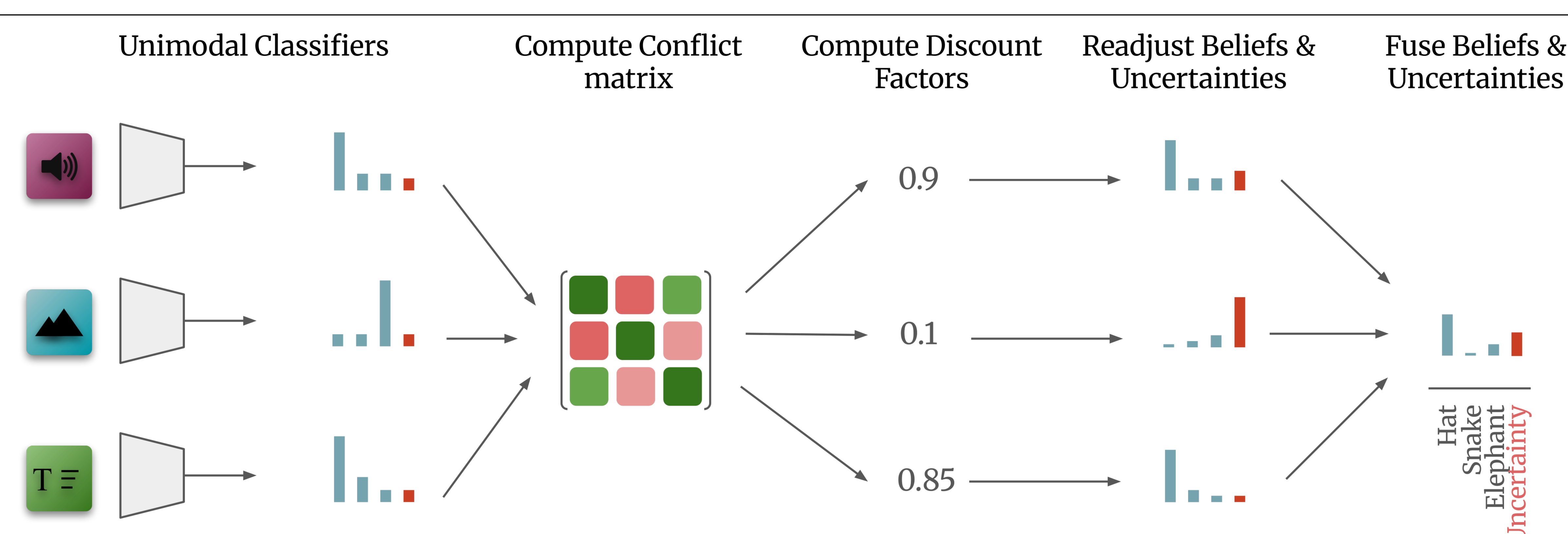
Under conflicting modalities, current SOTA methods:

- Provide counter-intuitive decisions
- Underestimate the uncertainty
- The outcomes depend on modality fusion order (non order-invariant).
- Fail to distinguish between conflicting and non-conflicting samples.

Contributions

- Propose **Discounted Belief Fusion (DBF)**, a novel uncertainty-aware multimodal fusion approach.
- Introduce a conflict-based discounting mechanism to reallocate uncertainty from conflicting modalities.
- Ensure fusion is **order-invariant** and scalable to multiple modalities.
- Demonstrate superior uncertainty quantification in conflicting scenarios compared to SOTA methods.
- Provide theoretical analysis and extensive experimental validation.

Discounted Belief Fusion



1. The **degree of conflict** DC between two subjective opinions $\boldsymbol{\omega}^1$ and $\boldsymbol{\omega}^2$ is defined as:

$$\mathrm{PD}\left(\boldsymbol{\omega}^1, \boldsymbol{\omega}^2\right) = \frac{\sum_{k=1}^{K} |p_k^1 - p_k^2|}{2},$$
$$\mathrm{CC}\left(\boldsymbol{\omega}^1, \boldsymbol{\omega}^2\right) = \left(1 - u^1\right)\left(1 - u^2\right),$$
$$\mathrm{DC}\left(\boldsymbol{\omega}^1, \boldsymbol{\omega}^2\right) = \mathrm{PD}\left(\boldsymbol{\omega}^1, \boldsymbol{\omega}^2\right) \cdot \mathrm{CC}\left(\boldsymbol{\omega}^1, \boldsymbol{\omega}^2\right),$$

2. **Conflict** and **Agreement** Matrix computation using degree of conflict:

$$C_{ij} = \mathrm{DC}(\boldsymbol{\omega}^i, \boldsymbol{\omega}^j), \quad \boldsymbol{A} = (1 - \boldsymbol{C}^{\lambda})^{1/\lambda}.$$

3. Compute the **Discounting Factors**

$$\eta^v = \prod_{i=1}^{V} A_{vi}, \quad \forall v \in [1, \dots V].$$

4. Readjust **belief and uncertainty masses**.

$$\boldsymbol{b}' = \eta \boldsymbol{b}, \quad u' = 1 - \eta + \eta u.$$

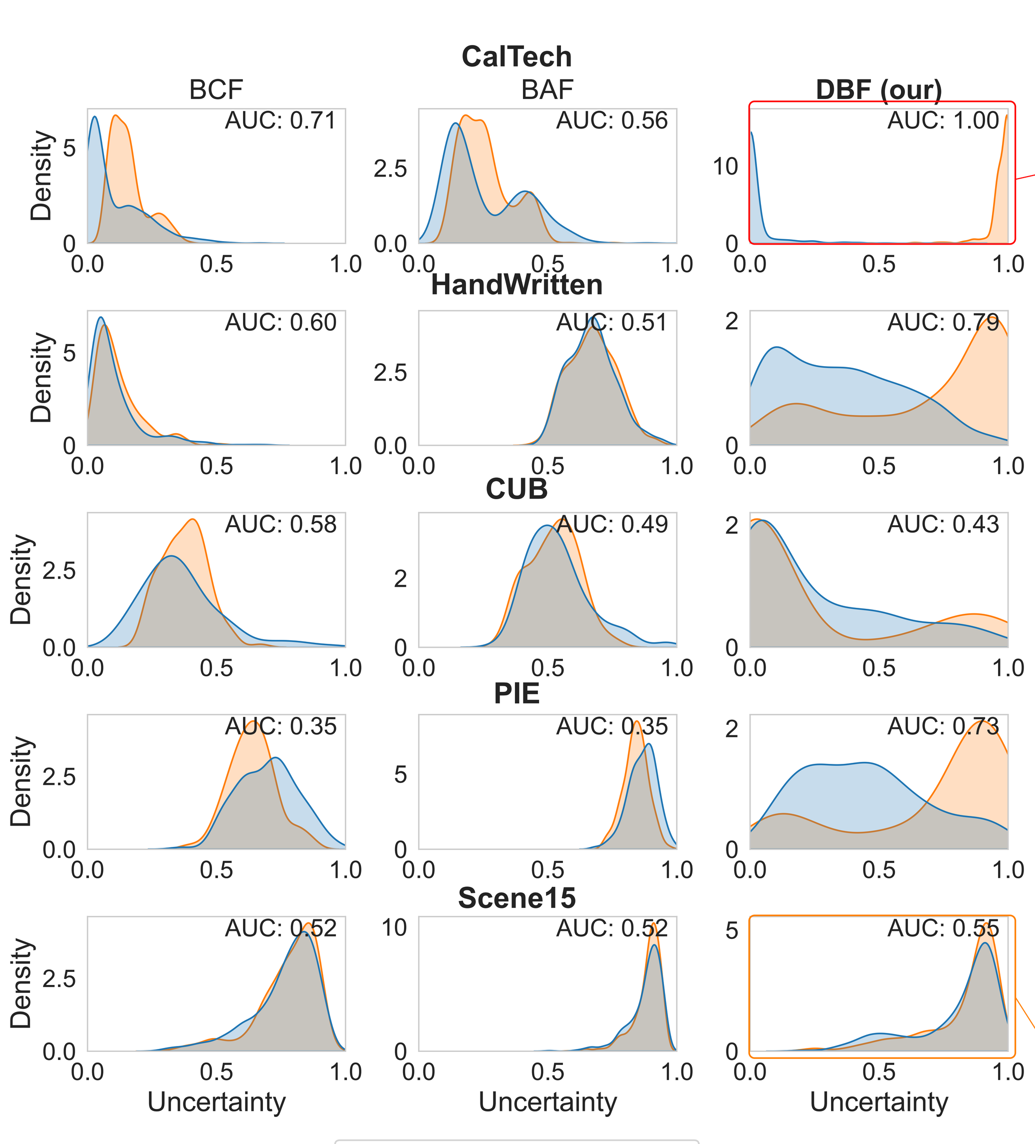
5. **Generalized Belief Averaging**

$$\boldsymbol{b}^{\diamond V} = \frac{\sum_{v=1}^{V} \left(\boldsymbol{b}^v \prod_{i \neq v} u^i\right)}{\sum_{v=1}^{V} \left(\prod_{i \neq v} u^i\right)},$$
$$u^{\diamond V} = \frac{V \prod_{v=1}^{V} u^v}{\sum_{v=1}^{V} \left(\prod_{i \neq v} u^i\right)},$$
$$\boldsymbol{a}^{\diamond V} = \frac{\sum_{v=1}^{V} \boldsymbol{a}^v}{V}.$$

Results

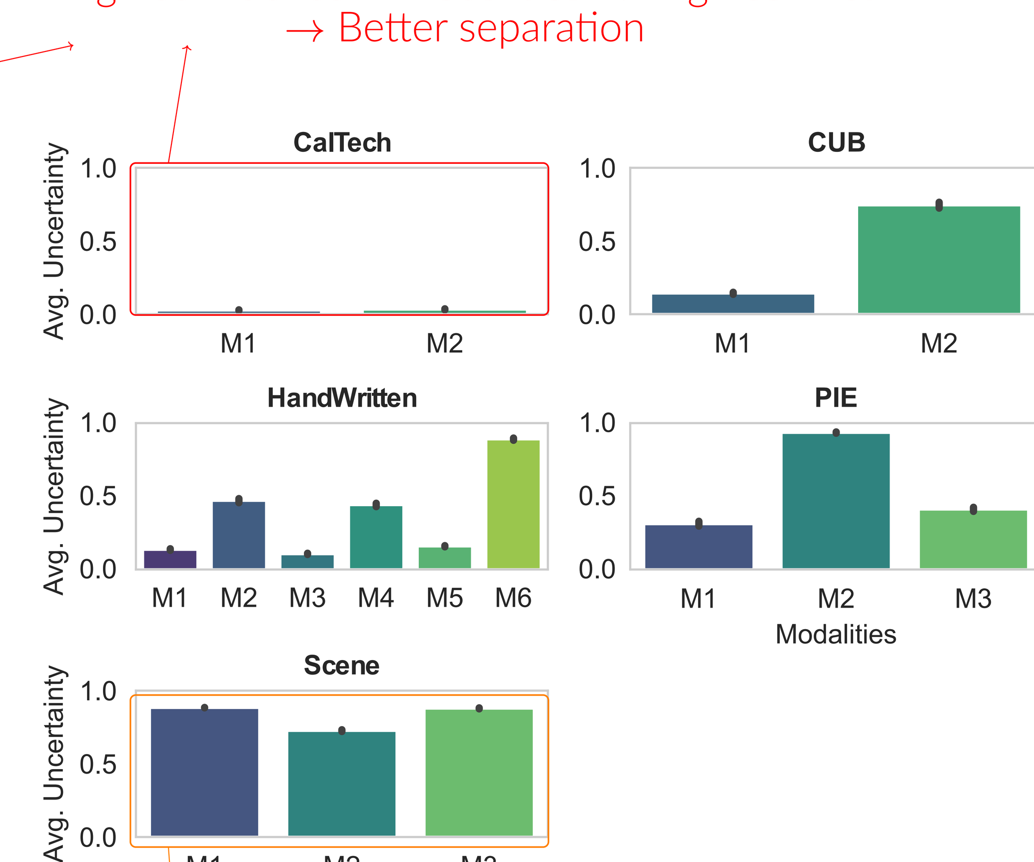
Fusion Method	b_1	b_2	b_3	u
Modality 1	0.99	0.00	0.01	0
Modality 2	0.00	0.99	0.01	0
BCF	0.0000	0.0000	1.0000	0
CBF	0.4950	0.4950	0.0100	0
BAF	0.4950	0.4950	0.0100	0
DBF $\lambda = 1$	0.0050	0.0050	0.0001	0.9900
DBF $\lambda = 3$	0.1533	0.1533	0.0031	0.6903

Unlike other fusion methods, **DBF increases the uncertainty mass under conflict**



DBF manages to **Distinguish between conflicting and non-conflicting samples** based on uncertainty values

High confidence in modalities → High conflict → Better separation



To have high degree of conflict, **modalities need to be confident** in their predictions

Low confidence in modalities → Low conflict → Bad separation