A Conflict-Guided Evidential Multimodal Fusion for Semantic Segmentation

Lucas Deregnaucourt[1] Hind Laghmara[1] Alexis Lechervy[2] Samia Ainouz[1]

[1]INSA Rouen Normandie, Univ Rouen Normandie, Université Le Havre Normandie, Normandie Univ, LITIS UR 4108, F-76000 Rouen, France

[2]Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

1. Context
2. A short introduction to the Dempster-Shafer Theory
3. Method
4. Experiments

Autonomous driving

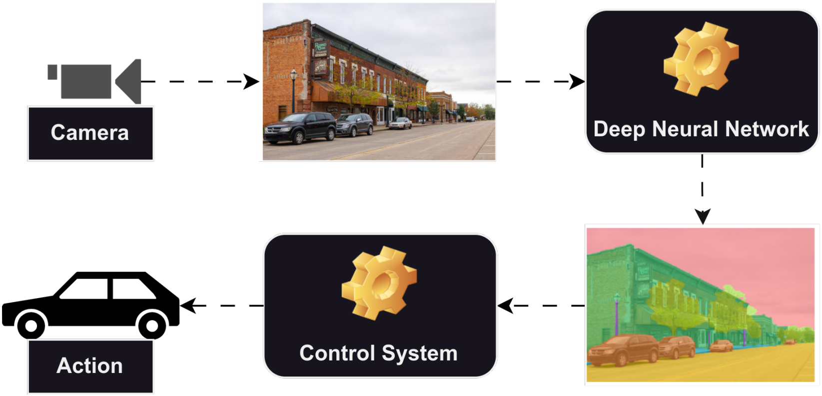


Figure 1: Autonomous driving flowchart

Problematic

→ **Task:** multimodal semantic segmentation for autonomous driving.

Problematic

→ **Task:** multimodal semantic segmentation for autonomous driving.

→ Current state of the art models **overtrust** the most informative sensor, leading to a **high drop of performances** in case of **sensor failure**.

Problematic

→ **Task:** multimodal semantic segmentation for autonomous driving.

→ Current state of the art models **overtrust** the most informative sensor, leading to a **high drop of performances** in case of **sensor failure**.

→ This **lack of robustness** is one of the key challenges of autonomous driving systems.

Our approach

Motivation

If an expert is highly in conflict with the others, it is reasonable to consider that either this expert struggles to make a decision or there is a sensor failure. Therefore, it is suitable to weaken the implication of this expert in the fusion process according to its conflict with the others.

→ We propose an adaptive multimodal late fusion pipeline to handle sensor failures for semantic segmentation.

Our approach

Motivation

If an expert is highly in conflict with the others, it is reasonable to consider that either this expert struggles to make a decision or there is a sensor failure. Therefore, it is suitable to weaken the implication of this expert in the fusion process according to its conflict with the others.

→ We propose an adaptive multimodal late fusion pipeline to handle sensor failures for semantic segmentation.

→ The proposed parameter-free fusion is grounded in the Dempster-Shafer Theory.

Our approach

Motivation

If an expert is highly in conflict with the others, it is reasonable to consider that either this expert struggles to make a decision or there is a sensor failure. Therefore, it is suitable to weaken the implication of this expert in the fusion process according to its conflict with the others.

→ We propose an adaptive multimodal late fusion pipeline to handle sensor failures for semantic segmentation.

→ The proposed parameter-free fusion is grounded in the Dempster-Shafer Theory.

→ The experts will be discounted according to their distance-based conflicts before the fusion.

Visual representation of the DST

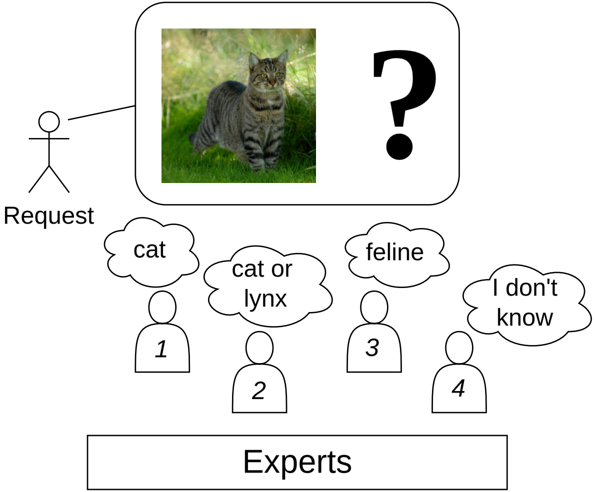


Figure 2: Multi-expert assumptions on image classification

Basic Belief Assignment

Basic Belief Assignment $\begin{cases} m(.) : 2^{\Omega} \to [0,1] \\ \sum_{A \subseteq \Omega} m(A) = 1 \end{cases}$

$\Omega = \{\omega_1, \omega_2, \ldots, \omega_K\}$: **frame of discernment**

Set of exhaustive and exclusive hypotheses

Basic Belief Assignment

Basic Belief Assignment $\begin{cases} m(.) : 2^{\Omega} \rightarrow [0,1] \\ \sum_{A \subseteq \Omega} m(A) = 1 \end{cases}$

$\Omega = \{\omega_1, \omega_2, \ldots, \omega_K\}$: **frame of discernment**

Set of exhaustive and exclusive hypotheses

$m_1(\{cat\}) = 0.7$	$m_2(\{cat\}) = 0.2$	$m_3(\{cat\}) = 0.1$	$m_4(\Omega) = 1$
$m_1(\{lynx\}) = 0.2$	$m_2(\{lynx\}) = 0.15$	$m_3(\{lynx\}) = 0.03$	
$m_1(\Omega) = 0.1$	$m_2(\{cat, lynx\}) = 0.5$	$\vdots$	
	$m_2(\Omega) = 0.15$	$m_3(\{puma\}) = 0.04$	
		$m_3(\{cat, lynx\}) = 0.05$	
		$\vdots$	
		$m_3(\{cat, lynx, \cdots, puma\}) = 0.4$	
		$m_3(\Omega) = 0.1$	

Figure 3: Mass assignments to sets of classes

Information fusion

Dempster's rule

Let m_1 and m_2 be two mass functions on Ω. The Dempster's rule to fuse them is defined as follows[1][2]*:*

$$m_{12}(A) = (m_1 \oplus m_2)(A) = \frac{1}{1-\kappa} \sum_{B \cap C = A} m_1(B) m_2(C) \quad (1)$$

$\forall A \subseteq \Omega \setminus \{\emptyset\}$

with κ the degree of conflict between the two sources of evidence:

$$\kappa = \sum_{B \cap C = \emptyset} m_1(B) m_2(C).$$

[1] A.P. Dempster. "Upper and lower probabilities induced by a multivalued mapping". In: The Annals of Mathematical Statistics 38.2 (1967), pp. 325–339

[2] Glenn Shafer. A mathematical theory of evidence. Vol. 42. Princeton university press, 1976

Probability transformation

To make a precise decision, we need to transform the mass function m into a probability vector:

$$^{1}BetP(\omega_k) = \sum_{\omega_k \in A \subseteq \Omega} \frac{m(A)}{|A|} \quad (2)$$

[1] P. Smets. "The combination of evidence in the transferable belief model". In: Transactions on pattern analysis and machine intelligence 12.2 (1990), pp. 447–458

[2] Jean Dezert and Florentin Smarandache. "A new probabilistic transformation of belief mass assignment". In: CoRR abs/0807.3669 (2008). arXiv: 0807.3669. URL: http://arxiv.org/abs/0807.3669

Probability transformation

To make a precise decision, we need to transform the mass function m into a probability vector:

$$^{1}BetP(\omega_k) = \sum_{\omega_k \in A \subseteq \Omega} \frac{m(A)}{|A|} \tag{2}$$

$$^{2}DSmP_{\varepsilon}(\omega_k) = \sum_{A \subseteq \Omega} m(A) \frac{|\omega_k \cap A| \left(m(\omega_k) + \varepsilon\right)}{\sum_{a \in A} m(a) + \varepsilon |A|} \tag{3}$$

[1] P. Smets. "The combination of evidence in the transferable belief model". In: Transactions on pattern analysis and machine intelligence 12.2 (1990), pp. 447–458

[2] Jean Dezert and Florentin Smarandache. "A new probabilistic transformation of belief mass assignment". In: CoRR abs/0807.3669 (2008). arXiv: 0807.3669. URL: http://arxiv.org/abs/0807.3669

Practical issues

Curse of dimensionality:

$$\Omega = \{\omega_1, \ldots, \omega_K\} \implies K \text{ classes.}$$

$$2^\Omega = \{\omega_1, \ldots, \omega_K, \{\omega_1, \omega_2\}, \ldots, \Omega\} \implies 2^K - 1 \text{ sets of classes.}$$

Practical issues

Curse of dimensionality:

$$\Omega = \{\omega_1, \dots, \omega_K\} \implies K \text{ classes.}$$

$$2^\Omega = \{\omega_1, \dots, \omega_K, \{\omega_1, \omega_2\}, \dots, \Omega\} \implies 2^K - 1 \text{ sets of classes.}$$

Common solution: Only the global uncertainty $m(\Omega)$ and the singletons $\{\omega_i\}_{i=1,\dots,K}$ are considered.

1. Context
2. A short introduction to the Dempster-Shafer Theory
3. Method
4. Experiments

Evidential Conflict-Guided Late Fusion

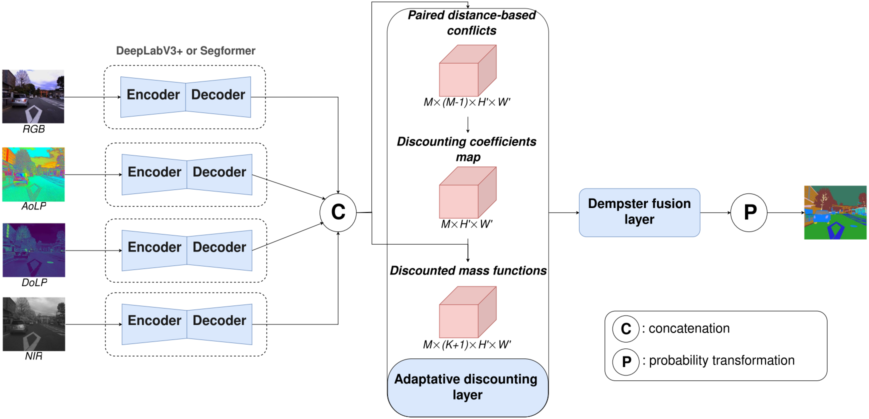


Figure 4: **ECoLaF architecture**. Each modality is associated to an independant encoder-decoder model such as DeepLabV3+ or Segformer. The mass functions of each modality are weakened by the adaptative discounting layer and fused by the Dempster's rule. The final decision is made after converting the mass functions into probabilities.

Adaptative discounting layer

Measure of conflict

Given M mass functions $m_{1,\dots,M}$, we compute the paired distances[1]:

$$d_{i,j} = dist(m_i, m_j) = \sqrt{\frac{1}{2}(m_i - m_j)^T D(m_i - m_j)} \quad (4)$$

where $D \in \mathbb{R}^{|\Omega|+1} \times \mathbb{R}^{|\Omega|+1}$, $D(A, B) = \frac{|A \cap B|}{|A \cup B|}$

[1] Anne-Laure Jousselme, Dominic Grenier, and Éloi Bossé. "A new distance between two bodies of evidence". In: Information fusion 2.2 (2001), pp. 91–101

Adaptative discounting layer

Measure of conflict

Given M mass functions $m_{1,\dots,M}$, we compute the paired distances[1]:

$$d_{i,j} = dist(m_i, m_j) = \sqrt{\frac{1}{2}(m_i - m_j)^T D(m_i - m_j)} \quad (4)$$

where $D \in \mathbb{R}^{|\Omega|+1} \times \mathbb{R}^{|\Omega|+1}$, $D(A,B) = \frac{|A \cap B|}{|A \cup B|}$

D=

	A	B	C	...	Ω
A	1	0	0	...	$\frac{1}{\lvert\Omega\rvert}$
B	0	1	0	...	$\frac{1}{\lvert\Omega\rvert}$
C	0	0	1	...	$\frac{1}{\lvert\Omega\rvert}$
⋮	...	...	...	⋱	$\frac{1}{\lvert\Omega\rvert}$
Ω	$\frac{1}{\lvert\Omega\rvert}$	$\frac{1}{\lvert\Omega\rvert}$	$\frac{1}{\lvert\Omega\rvert}$	...	1

[1] Anne-Laure Jousselme, Dominic Grenier, and Éloi Bossé. "A new distance between two bodies of evidence". In: Information fusion 2.2 (2001), pp. 91–101

Adaptative discounting layer

Measure of conflict

Given M mass functions $m_{1,\dots,M}$, we compute the paired distances[1]:

$$d_{i,j} = dist(m_i, m_j) = \sqrt{\frac{1}{2}(m_i - m_j)^T D(m_i - m_j)} \quad (4)$$

The conflict[2] between m_i and m_j can be obtained by:

$$Conf_{i,j} = \left(1 - \frac{2|\Omega|+1}{(|\Omega|+1)^2}\right) \times d_{i,j} \quad (5)$$

[1] Anne-Laure Jousselme, Dominic Grenier, and Éloi Bossé. "A new distance between two bodies of evidence". In: Information fusion 2.2 (2001), pp. 91–101

[2] Arnaud Martin. "About conflict in the theory of belief functions". In: Belief Functions: Theory and Applications: Proceedings of the 2nd International Conference on Belief Functions. Springer. 2012, pp. 161–168

Adaptative discounting layer

Measure of conflict

The conflict[2] between m_i and m_j can be obtained by:

$$Conf_{i,j} = \left(1 - \frac{2|\Omega|+1}{(|\Omega|+1)^2}\right) \times d_{i,j} \tag{4}$$

The paired conflicts are averaged to obtain the conflict asociated to each mass function m_i:

$$Conf_i = \frac{1}{M-1} \sum_{j=1, i \neq j}^{M} Conf_{i,j} \tag{5}$$

[2] Arnaud Martin. "About conflict in the theory of belief functions". In: Belief Functions: Theory and Applications: Proceedings of the 2nd International Conference on Belief Functions. Springer. 2012, pp. 161–168

Adaptative discounting layer

Mass functions discounting

Given a mass function m_i and a discounting coefficient $\alpha_i \in [0,1]$, the discounting procedure is defined as follows:

$$\begin{cases} m_i^{\alpha_i}(\omega_k) = \alpha_i m_i(\omega_k) \\ m_i^{\alpha_i}(\Omega) = 1 - \alpha_i(1 - m_i(\Omega)) \end{cases} \tag{6}$$

Adaptative discounting layer

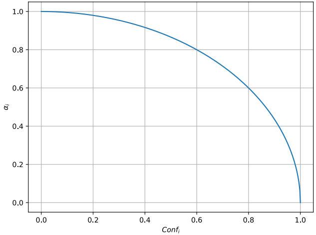
Mass functions discounting

Given a mass function m_i and a discounting coefficient $\alpha_i \in [0, 1]$, the discounting procedure is defined as follows:

$$\begin{cases} m_i^{\alpha_i}(\omega_k) = \alpha_i m_i(\omega_k) \\ m_i^{\alpha_i}(\Omega) = 1 - \alpha_i(1 - m_i(\Omega)) \end{cases} \quad (6)$$

We can compute the discounting coefficient α_i associated to m_i from $Conf_i$[1]:

$$\alpha_i = (1 - Conf_i^2)^{\frac{1}{2}} \quad (7)$$

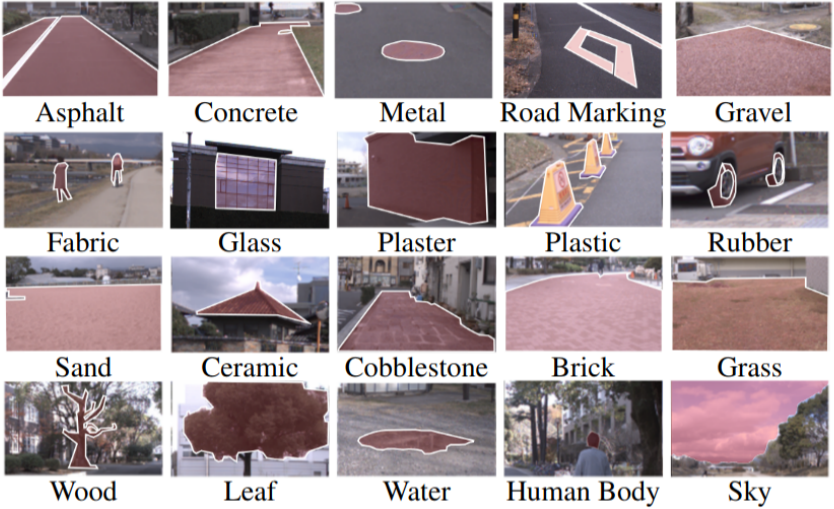


[1] Arnaud Martin, Anne-Laure Jousselme, and Christophe Osswald. "Conflict measure for the discounting operation on belief functions". In: 2008 11th International Conference on Information Fusion. 2008, pp. 1–8.

1. Context
2. A short introduction to the Dempster-Shafer Theory
3. Method
4. Experiments

Dataset

MCubeS



Quantitative results

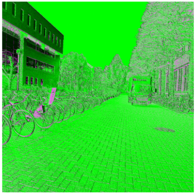
RGB	AoLP	DoLP	NIR	convolution-based models: MCubeSNet	convolution-based models: ECoLaF-DeepLabV3+	transformers-based models: CMNeXt	transformers-based models: ECoLaF-Segformer
✓				30.79	43.49	42.32	**46.48**
	✓			3.63	**21.45**	2.1	10.45
		✓		1.66	**35.44**	3.42	19.84
			✓	1.00	**32.81**	2.15	16.79
mean for 3 failures				9.27	**33.30**	12.50	23.39
✓	✓			38.10	43.36	**48.81**	46.48
✓		✓		35.98	44.95	**49.00**	48.11
✓			✓	33.16	44.39	**48.36**	45.01
	✓	✓		4.60	**36.35**	1.43	27.61
	✓		✓	1.67	**36.81**	1.74	23.14
		✓	✓	1.12	**41.53**	3.15	27.19
mean for 2 failures				19.11	**41.23**	25.42	36.26
✓	✓	✓		41.54	45.26	**49.06**	48.75
✓	✓		✓	40.61	44.25	**49.78**	47.77
✓		✓	✓	39.53	45.57	**50.02**	49.85
	✓	✓	✓	2.74	**41.72**	5.05	33.31
mean for 1 failure				26.11	44.20	38.48	**44.92**
✓	✓	✓	✓	43.26	45.74	**51.54**	49.85
average mean				24.44	**41.12**	31.99	38.61

Table 1: Performances comparison of using different combinations of modalities in mIoU(%) on MCubeS dataset. Bold values represent the best performances to the nearest rounding for each combination of modalities.

Qualitative results

All modalities are available


(a) RGB


(b) AoLP


(c) DoLP


(d) NIR


(e) MCubeSNet


(f) CMNeXt

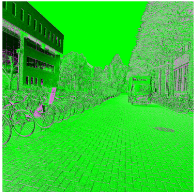

(g) ECoLaF-DeepLabV3+


(h) Ground truth

Visual results

Partial RGB occlusion

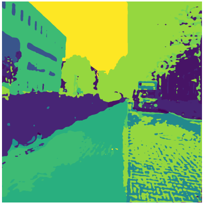
(a) RGB



(b) AoLP

(c) DoLP

(d) NIR



(e) MCubeSNet

(f) CMNeXt

(g) ECoLaF-DeepLabV3+

(h) Ground truth

Conclusion

→ Robust models are required in real life applications to face hazards.

→ Parameter-free techniques are fully adaptive, improving models robustness. On the contrary, parameters-guided fusions tend to overtrust the most informative sensor, leading to a fragile robustness.

→ Our results show that the proposed ECoLaF pipeline offers a good tradeoff between performances and robustness in case of sensor failure.

Code is available!

pip install ecolaf

https://github.com/deregnaucourtlucas/ECoLaF