Evidential Machine Learning

Theory and Applications to Medical Image Processing

Thierry Denœux

Université de technologie de Compiègne, CNRS, Heudiasyc, France
Institut Universitaire de France, Paris, France

Sustainable Photonics Conference
July 17, 2026 Hangzhou, China

From theory to medical applications

1. Why?

Uncertainty is central in medical image analysis: ambiguity, data scarcity, domain shift, imperfect labels.

2. How?

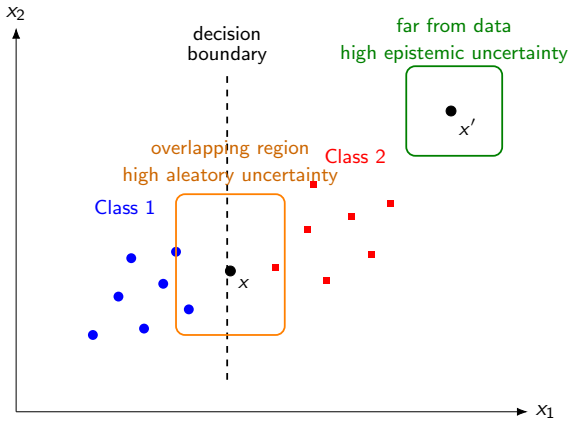
Design uncertainty-aware neural network models based on evidence theory

3. What for?

Tumour segmentation with uncertainty quantification and multimodal information fusion.

Evidence theory $\longrightarrow$ evidential machine learning $\longrightarrow$ deep evidential segmentation $\longrightarrow$ multimodal fusion

Classification uncertainty



Two qualitatively different situations

- Aleatory uncertainty: class overlap or intrinsically ambiguous voxels.
- Epistemic uncertainty: the test case is far from the training distribution.

- Probabilistic classifiers do not distinguish between these two situations.
- Evidence theory, an extension of probability theory, makes it possible to better represent both kinds of uncertainty.

Evidence theory: belief and plausibility

Consider a classification problem. Let Y be the class variable with unknown value in $\Theta = \{\theta_1, \ldots, \theta_c\}$. (E.g.: classification of voxels as gadolinium-enhancing tumor, necrotic and non-enhancing tumour core, peritumoral edema, background tissue).

Mass function

A mass function is a representation of evidence as a mapping

$$m : \mathcal{P}(\Theta) \to [0, 1], \qquad m(\varnothing) = 0, \qquad \sum_{A \subseteq \Theta} m(A) = 1.$$

The quantity $m(A)$ is the probability that the evidence tells us exactly that $Y \in A$, and nothing more.

Belief funtion

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(B)$$

Degree of support for A.

Plausibility function

$$\mathrm{Pl}(A) = \sum_{B \cap A \neq \varnothing} m(B) = 1 - \mathrm{Bel}(A^c)$$

Degree of compatibility of A with the evidence.

Evidence theory: combination of evidence

Dempster's rule

Independent pieces of evidence are pooled by intersecting focal sets:

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

where

$$\kappa = \sum_{B \cap C = \varnothing} m_1(B) m_2(C) \quad \text{is the degree of conflict.}$$

Interpretation

- κ measures the disagreement between the two pieces of evidence.
- Combination is possible iff $\kappa < 1$.

Evidence theory: discounting

A mechanism for weakening a mass function provided by a source that is not fully reliable.

Simple discounting

- A source of information provides evidence represented by a mass function m.
- The source may be reliable or not. Our degree of belief that the source is reliable is $\beta \in [0, 1]$.
- Combining this meta-evidence with m yields a discounted mass function

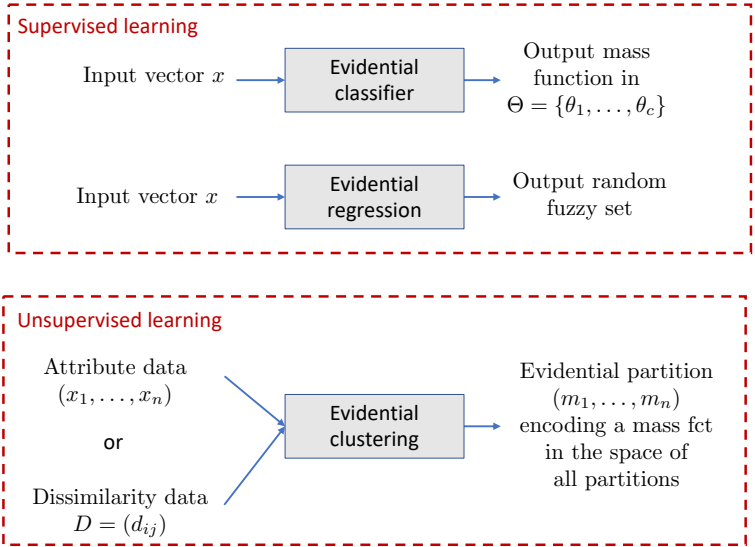
$$^{\beta}m(A) = \begin{cases} \beta m(A) & \text{if } A \subset \Theta \\ \beta m(\Theta) + 1 - \beta & \text{if } A = \Theta. \end{cases}$$

Contextual discounting

- Our degree of belief that the source is reliable given that the truth is θ_k is β_k.
- Combining this meta-evidence' with m yields a discounted plausibility function

$$^{\boldsymbol{\beta}}\mathrm{Pl}(A) = 1 - (1 - \mathrm{Pl}(A)) \prod_{\theta_k \in A} \beta_k, \quad \text{with} \quad \boldsymbol{\beta} = (\beta_1, \ldots, \beta_c).$$

Evidential machine learning



From probabilities to evidential predictions

Probabilistic classifier

For a voxel or image x, output

$$p(\theta_1), \ldots, p(\theta_c).$$

All mass is assigned to single classes.

Evidential classifier

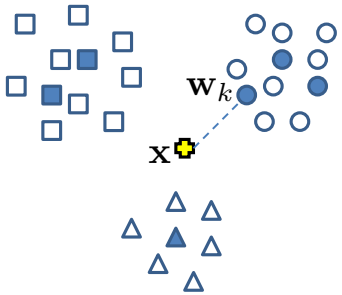
For the same input, output a mass function such as

$$m(\{\theta_k\}), \quad k = 1, \ldots, c, \qquad m(\Theta).$$

The mass $m(\Theta)$ represents residual ignorance.

Decision can be based on plausibility, or lower/upper expected utility.

Evidential neural networks: prototypes as evidence



Principle

- Learn prototypes $\mathbf{w}_1, \ldots, \mathbf{w}_K$ in feature space.
- Each prototype $\mathbf{w}_k$ has membership degree $u_{kq} \in [0,1]$ to class θ_q, with $\sum_q u_{kq} = 1$.
- It supplies a piece of evidence.
- Close prototype: strong evidence. Far prototype: weak evidence.

Prototype k

$$s_k(x) = \exp\{-\gamma_k \|x - w_k\|^2\}, \quad \gamma_k \geq 0,$$

$$m_k(\{\theta_q\}) = \alpha_k u_{kq} s_k(x),$$

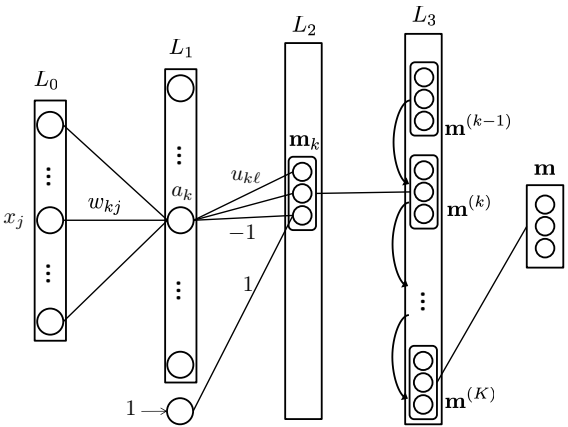
$$m_k(\Theta) = 1 - \alpha_k s_k(x).$$

Combination

$$m = m_1 \oplus \cdots \oplus m_K.$$

Each prototype contributes evidence; the output summarises all prototype evidence.

ENN architecture and learning



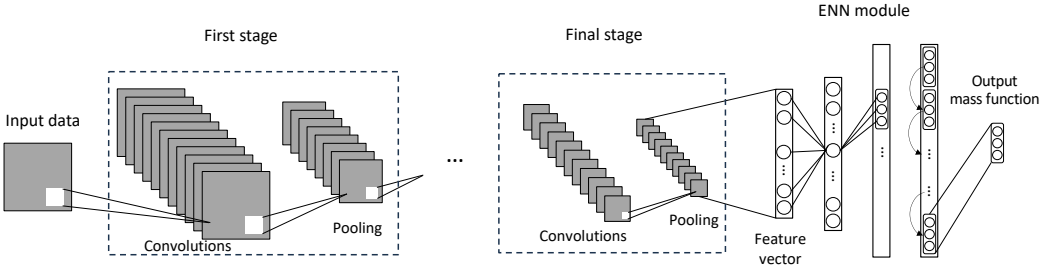
Learning

- Parameters: prototype locations, receptive-field widths, reliabilities and class memberships.
- A standard criterion is a regularised MSE between output probabilities and class indicator vectors:

$$\mathrm{MSE}_\lambda(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{q=1}^{c}(p_{iq} - y_{iq})^2 + \lambda\sum_{k=1}^{K}\alpha_k.$$

- Cross-entropy can replace MSE.

Embedding ENN modules in deep networks



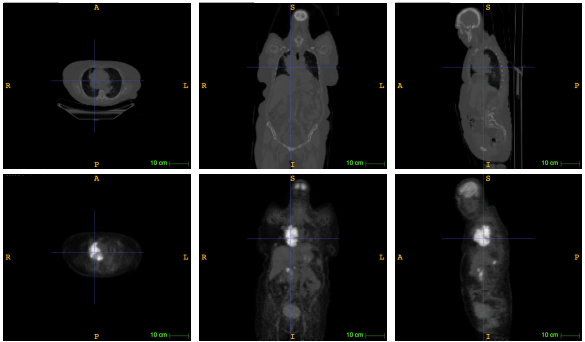
Deep evidential classifier

- A CNN maps raw input x to a feature vector z.
- The ENN module operates in the learned feature space.
- The output is a mass function, not only a probability distribution.

Semantic segmentation

The ENN module is applied pixelwise or voxelwise: each voxel receives a belief function over anatomical classes.

Application: lymphoma segmentation in 3D PET/CT



CT and PET slices for one patient in axial, sagittal and coronal views.

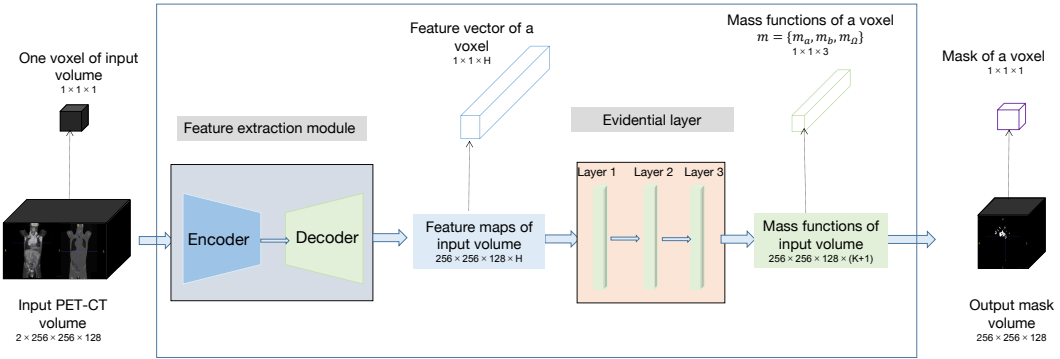
Clinical and technical context

- Diffuse large B-cell lymphoma segmentation.
- Input: registered PET and CT volumes.
- PET highlights metabolic activity; CT provides anatomical information.
- Main difficulty: blurred boundaries and bright non-tumour organs.

Reference

Huang, Ruan, Decazes, Denœux.
Lymphoma segmentation from 3D PET-CT images using a deep evidential network. IJAR, 2022.

Lymphoma segmentation: architecture



Idea

A U-Net backbone extracts features from the PET/CT volume; the evidential layer maps each voxel feature vector to a mass function over $\{\text{background}, \text{tumour}\}$.

Lymphoma segmentation: what is gained?

Segmentation accuracy

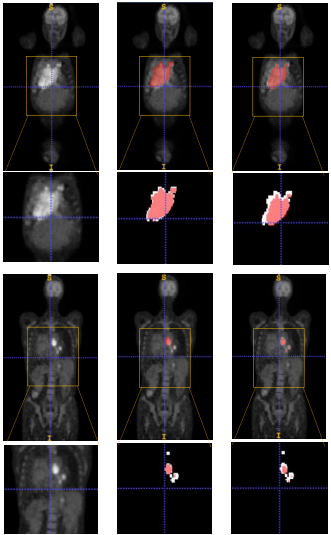
Model	Dice	Sens.	Prec.
UNet	0.753	0.782	0.896
nnUNet	0.817	0.838	0.879
VNet	0.820	0.831	0.901
SegResNet	0.825	0.832	0.876
ENN-UNet	**0.846***	0.830	0.879

Uncertainty quantification

- The evidential layer outputs $m(\Theta)$ at each voxel.
- This gives a direct map of local ignorance.
- Performance is improved without losing interpretability of the output.

Table values: means over five runs reported in Huang et al., with a dataset of 173 patients (IJAR, 2022). * p-value $= 1.7 \cdot 10^{-4}$ (Kruskal-Wallis test).

Lymphoma segmentation: qualitative result



Observation

- White: manual segmentation; red: automatic segmentation.
- ENN-UNet better overlaps the ground truth on the two illustrated cases.
- The gain is especially visible for the small isolated lymphoma.

Message

The evidential layer does not only give an additional uncertainty output; it can also improve segmentation quality.

From data-level to decision-level fusion

Usual strategy

Concatenate modalities, e.g.

$$x = (x^{\text{PET}}, x^{\text{CT}})$$

and let a deep network learn a fused representation.

Limitation

This assumes that all modalities are equally useful and reliable everywhere.

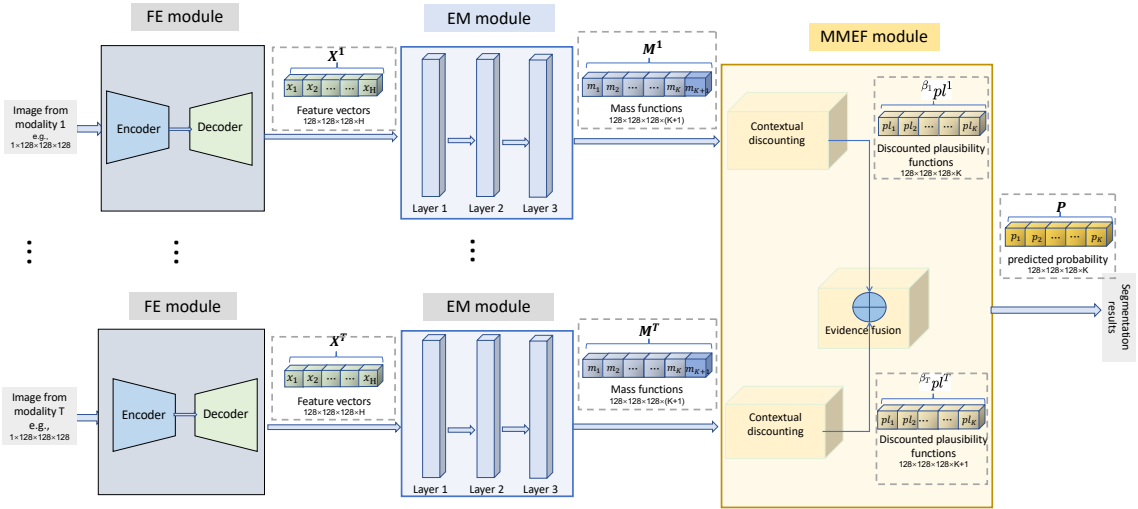
Evidential view

Each modality is a source of evidence. Its reliability may depend on the class and can be learned from data.

Reference

Huang, Ruan, Decazes, Denœux.
Deep evidential fusion with uncertainty quantification and reliability learning for multimodal medical image segmentation. Information Fusion, 2025.

Deep evidential fusion



Contextual reliability learning

For modality t and class θ_k, let β_k^t be the learned reliability coefficient.

Contextual discounting of plausibilities

$$^{\beta^t}\mathrm{Pl}_n^t(\{\theta_k\}) = 1 - \beta_k^t + \beta_k^t\,\mathrm{Pl}_n^t(\{\theta_k\}), \qquad k = 1, \ldots, c.$$

Decision-level fusion

$$^{\boldsymbol{\beta}}\mathrm{Pl}_n(\{\theta_k\}) \propto \prod_{t=1}^{T} {}^{\beta^t}\mathrm{Pl}_n^t(\{\theta_k\}), \qquad p_n(\theta_k) = \frac{^{\boldsymbol{\beta}}\mathrm{Pl}_n(\{\theta_k\})}{\sum_{\ell=1}^{c} {}^{\boldsymbol{\beta}}\mathrm{Pl}_n(\{\theta_\ell\})}.$$

The coefficients β_k^t provide insight into the contribution of each modality to the segmentation.

Application to PET-CT lymphoma segmentation

Segmentation quality and reliability

Model	ECE ↓	Brier ↓	NLL ↓	Dice ↑
UNet	0.056	0.065	0.310	0.770
UNet-MC	0.053	0.062	0.400	0.801
UNet-Ensemble	0.063	0.064	0.343	0.802
ENN-UNet	0.050	0.062	0.191	0.805
RBF-UNet	0.051	0.061	0.193	0.802
MMEF-UNet	0.045	0.056	0.180	0.811

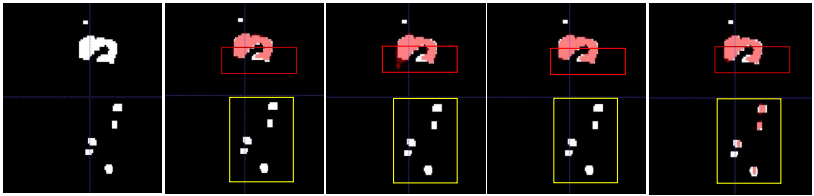
Learned reliability coefficients

Modality	Background	Lymphoma
PET	0.999	0.996
CT	0.863	0.975

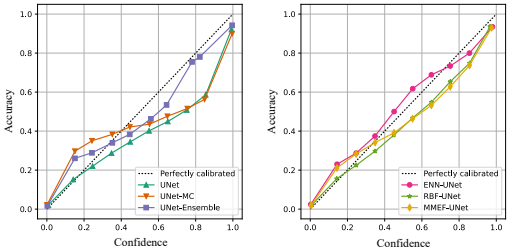
Interpretation

PET has the highest learned reliability, consistent with its role in highlighting metabolic activity.

Calibration and segmentation example



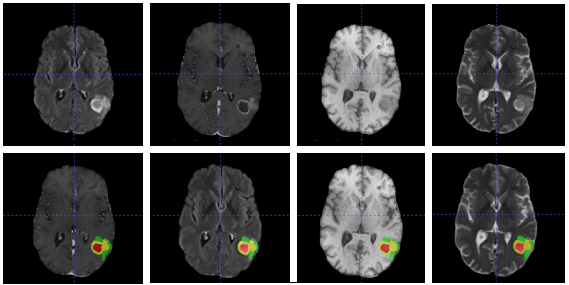
The white and red regions represent, respectively, the ground truth and the segmentation result. Red and yellow boxes highlight the main differences in segmenting large and small isolated tumours.



Message

The decision-level evidential fusion model improves both segmentation accuracy and calibration in this PET-CT example.

Application to multi-MRI brain tumour segmentation



Data

- BraTS2021 challenge
- Training, validation, and test sets with, respectively, 1251, 219, and 570 cases.
- Four modalities: FLAIR, T1Gd, T1, and T2 with $240 \times 240 \times 155$ voxels.

Learning task

- Annotations of scans: gadolinium (Gd)-enhancing tumor (ET), necrotic and non-enhancing tumour core (NCR/NET), and peritumoral edema (ED).
- The task is to segment the images into three overlapping regions: ET, tumor core (TC, the union of ET and NRC/NET), and whole tumour (WT, the union of ET, NRC/NET, and ED).
- Performances evaluated with respect to these three overlapping regions.

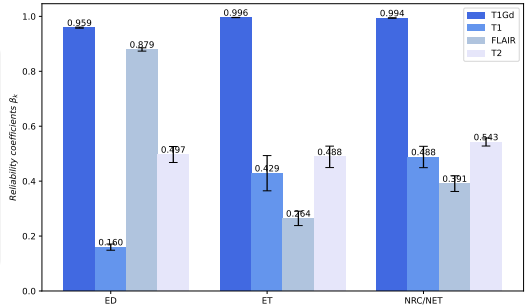
Results

Dice scores, nnFormer feature extractor

Model	ET	TC	WT	Mean
nnFormer	0.839	0.878	0.915	0.877
nnFormer-MC	0.837	0.877	0.914	0.876
ENN-nnFormer	0.836	0.882	0.914	0.878
MMEF-nnFormer	**0.854**	**0.911**	0.914	**0.893**

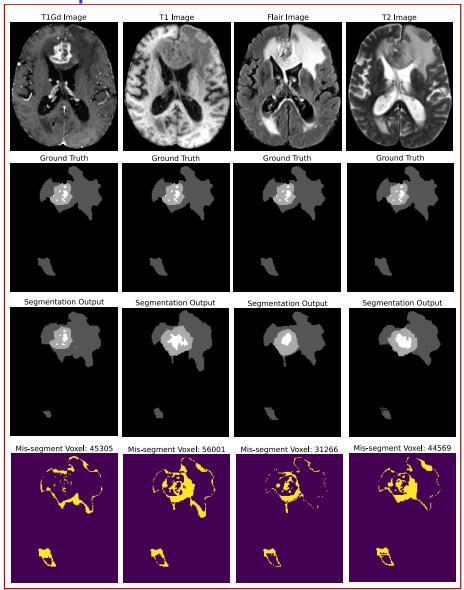
Detailed regions

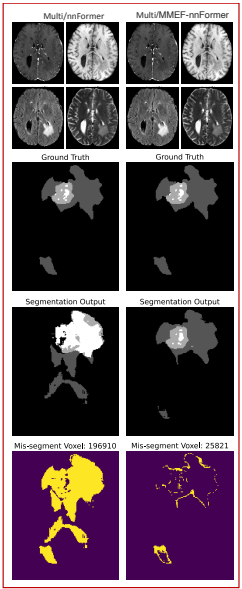
For NRC/NET, Dice increases from 0.740 to 0.805 with MMEF-nnFormer.



Learned reliability coefficients for the four MRI modalities. Consistent with domain knowledge. E.g., T1Gd useful for delineating tumour boundaries, FLAIR images sensitive to ED).

Example





- First and second rows: input modalities and the tumour ground truth.
- Third and last rows: segmentation output and the mis-segmented voxels (highlighted in yellow).
- Left red block: single-modality input using nnFormer,
- Right red block: multimodal input using nnFormer (left column) and MMEF-nnFormer (right column).

Take-home messages

1. Uncertainty quantification

Evidence theory makes it possible to represent support, plausibility, ignorance and conflict within the same mathematical framework.

2. Learning

ENN modules turn distances to prototypes into pieces of evidence and can be embedded in modern deep networks.

3. Medical image processing

- For tumour segmentation, evidential machine learning provides uncertainty maps and can improve accuracy.
- In multimodal settings, reliability learning gives insight into the role of each modality.

Main references

https://www.hds.utc.fr/~tdenoeux

T. Denœux
A neural network classifier based on Dempster-Shafer theory
IEEE Transactions on Systems, Man and Cybernetics A, 2000

Z. Tong, Ph. Xu and T. Denœux
An evidential classifier based on Dempster-Shafer theory and deep learning
Neurocomputing, 2021

L. Huang, S. Ruan, P. Decazes, T. Denœux
Lymphoma segmentation from 3D PET-CT images using a deep evidential network
International Journal of Approximate Reasoning, 2022.

L. Huang, S. Ruan, P. Decazes, T. Denœux
Deep evidential fusion with uncertainty quantification and reliability learning for multimodal medical image segmentation
Information Fusion, 2025.