ISCC Paper Presentation

Deep Evidential Fusion with Uncertainty Quantification for Multimodal Medical Image Segmentation

Group G
Hudie Sun - Jinxing Lai

University of Technology of Compiègne

November 3, 2024

Table of Contents

Table of Contents

Introduction

Multimodal Medical Imaging:

- **PET/CT/MRI**: Provides complementary information.
  - **PET**: Highlights metabolic activity.
  - **CT**: Shows anatomical structures.
  - **Combined**: Allows for more accurate diagnoses.

Problem:

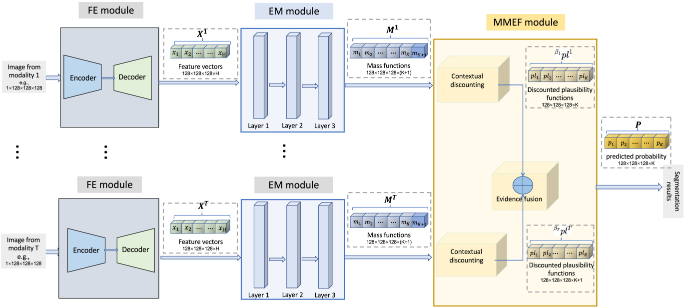
- Traditional fusion methods assume equal reliability across modalities.
- Real-world data variation:
  - Quality, resolution, and reliability differ between modalities.
- Result: Potential segmentation errors.

Aim:

- Propose a **deep evidential fusion framework**.
- Utilise **Dempster-Shafer Theory** (DST):
  - Model uncertainty and reliability.
  - Aim for more accurate and explainable segmentation.

Table of Contents

Proposed Framework (Architecture)



Key Components of the Framework:

- The proposed framework consists of three main modules:
  - **Feature Extraction module (FE)**
  - **Evidence Mapping module (EM)**
  - **Multi-modality Evidence Fusion module (MMEF)**
- Each modality has its own FE and EM module.

Table of Contents

Table of Contents

Feature Extraction (FE)

- **Deep learning models**: UNet, nnUnet, nnFormer, etc.
- **Independent feature extraction per modality.**
- **Example:**
  - **Input**: $128 \times 128 \times 128$ **single-channel** image.
  - **Output**: **H**-channel image (same spatial size).
- **H values used in the paper:**
  - PET-CT lymphoma dataset: $H = 2$
  - Multi-MRI BraTS2021 dataset: $H = 4$

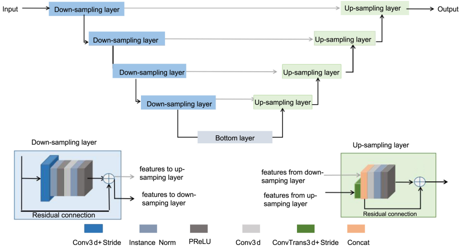


Figure: U-NET Structure

Table of Contents

Evidential Neural Network (ENN) as the EM Module

- Transforming extracted features using **Evidential Neural Network (ENN)**.
- Output: mass functions representing evidence about the class of each voxel.
- Tensor Size: $128 \times 128 \times 128 \times (K+1)$
  - One mass for each class θ_k
  - One mass for Θ

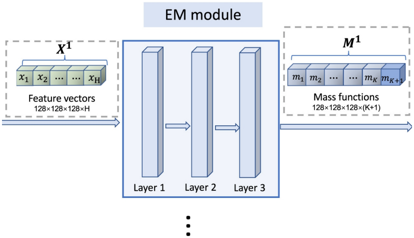


Figure: EM-ENN Module Overview

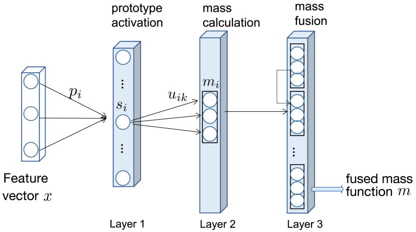


Figure: Evidential Neural Network (ENN) Structure

Prototypes p_1^t in Feature Space

- Modalities: $t \in \{1, \ldots, T\}$
- Prototypes: $p_1^t, \ldots, p_I^t$
- Prototypes are obtained using **k-means** in the feature space extracted by the Feature Extraction (FE) module.
- **Number of Prototypes (I)**:
  - PET-CT Lymphoma Dataset: $I = 10$
  - Multi-MRI BraTS2021 Dataset: $I = 20$

Similarity Measure s_i^t

- **(First Layer) Activation of Unit** i:

$$s_i^t = \alpha_i^t \exp\left(-\gamma_i^t \|x - p_i^t\|^2\right)$$

- Learnable parameters:
  - $\gamma_i^t > 0, \quad \alpha_i^t \in [0, 1]$
  - Initialisation used in this paper:
    - γ_i^t: 0.5
    - α_i^t: 0.01
- s_i^t reflects the similarity between input feature vector x and prototype p_i^t.

m_i^t: mass function with discount rate $(1 - s_i^t)$

- **Second Layer**: Computes mass function m_i^t derived from the prototype p_i^t.
- **Focal Sets**:
  - Singletons θ_k, $k = 1, \ldots, K$
  - Universal set Θ
- **Mass Calculation**:

$$m_i^t(\{\theta_k\}) = u_{ik}^t s_i^t, \quad k = 1, \ldots, K$$

$$m_i^t(\Theta) = 1 - s_i^t$$

- **Learnable parameter**:
  - Membership Degree u_{ik}^t
  - Initialised with uniform random numbers and normalization (in this paper)

$m^t = \oplus_{i=1}^{I} m_i^t$: mass function fusion

- **Third Layer**: Combine I mass functions using **Dempster's Rule**.
- Fusing mass functions to summarise evidence provided by I prototypes
- **Dempster's Rule**:

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

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

Table of Contents

Multi-modality Evidence Fusion (MMEF)

- To **fuse** the evidence gathered from each modality.
- MMEF performs fusion at the **contour function level**, not the mass function level.
- Helps facilitate **plausibility-probability transformation**.
- **T discounting vectors** $\beta = (\beta^1, \ldots, \beta^T)$, $\beta^t = (\beta^t_1, \ldots, \beta^t_K)$, represent reliability of modality t for class θ_k.
- Initialisation used in this paper: **KT reliability coefficients** β^t_k set to 0.5.

Multi-modality Evidence Fusion (MMEF) (continued)

1. **Evidence Fusion on Contour Function Level**
   - **1.1 Contour Function for Voxel n and Modality t:**

$$pl_n^t(\theta_k) = m_k^t(\{\theta_k\}) + m_n^t(\Theta)$$

   - **1.2 Discounted Contour Function for Voxel n and Modality t:**

$${}^{\beta^t}pl_n^t(\theta_k) = 1 - \beta_k^t + \beta_k^t\, pl_n^t(\theta_k)$$

   - **1.3 Combined Contour Function at Voxel n:**

$${}^{\beta}pl_n(\theta_k) \propto \prod_{t=1}^{T} {}^{\beta^t}pl_n^t(\theta_k), \quad k = 1, \ldots, K$$

2. **Transform Plausibility to Predicted Probability**

$${}^{\beta}p_n(\theta_k) = \frac{{}^{\beta}pl_n(\theta_k)}{\sum_{l=1}^{K} {}^{\beta}pl_n(\theta_l)} = \frac{\prod_{t=1}^{T}\left(1 - \beta_k^t + \beta_k^t pl_n^t(\theta_k)\right)}{\sum_{l=1}^{K}\prod_{t=1}^{T}\left(1 - \beta_l^t + \beta_l^t pl_n^t(\theta_l)\right)}, \quad k = 1, \ldots, K$$

Table of Contents

Loss Function

Optimisation Loss Function:

$$loss = loss_s + loss_f$$

loss$_s$: Evaluates segmentation performance of **each modality** and **aggregates** results.

$$loss_s = \sum_{t=1}^{T} \left[1 - \frac{2\sum_{n=1}^{N}\sum_{k=1}^{K} m_n^t(\{\theta_k\}) \times G_{kn}}{\sum_{n=1}^{N}\sum_{k=1}^{K} (m_n^t(\{\theta_k\}) + G_{kn})} \right]$$

loss$_f$: Quantifies segmentation performance **after combining all** T **modalities**.

$$loss_f = 1 - \frac{2\sum_{n=1}^{N}\sum_{k=1}^{K} {}^{\beta}p_n(\theta_k) \times G_{kn}}{\sum_{n=1}^{N}\sum_{k=1}^{K} {}^{\beta}p_n(\theta_k) + G_{kn}}$$

Note:

- N: Number of voxels.
- $G_{kn} = 1$ if voxel n belongs to class θ_k, otherwise $G_{kn} = 0$.

Table of Contents

Training Process

Learnable Parameters:

- **FM Module**: Weights of the deep learning models
- **EM Module**: $\alpha_i^t, \gamma_i^t, u_{ik}^t$
- **MMEF Module**: β

Training Steps:

1. Train the **FE** module independently.
2. Fix **FE** weights, optimise **EM** and **MMEF** modules.
3. Fine-tune the combined model (**FE** + **EM** + **MMEF**) for a few epochs.

Table of Contents

Table of Contents

Dataset Description: PET-CT Lymphoma

PET-CT Lymphoma Dataset:

- **Modalities**: PET and CT
- **Data**: 173 patients with large B-cell lymphoma
- **Split**: 138 training, 17 validation, 18 testing
- **Labels**: Background, lymphoma region

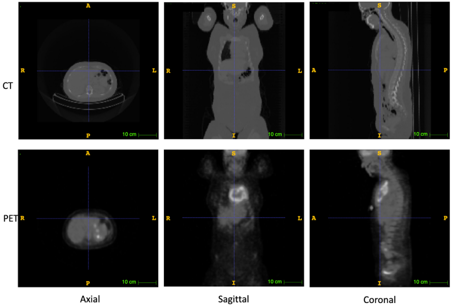


Figure: Example of PET-CT Lymphoma Image

Dataset Description: Multi-MRI Brain Tumor

Multi-MRI Brain Tumor Dataset (BraTS2021):

- **Modalities**: FLAIR, T1Gd, T1, T2
- **Split**: 834 training, 208 validation, 209 testing
- **Labels**: Peritumoral edema (ED, green), Enhancing Tumor (ET, yellow), necrotic tumor core or non-enhancing tumor (NCR/NET, red).

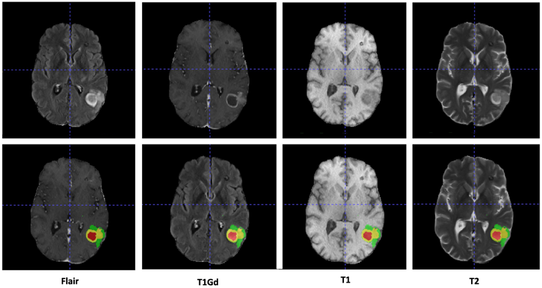


Figure: Example of MRI Brain Tumor Image

Table of Contents

Evaluation Metrics

Segmentation Accuracy: Dice Score

- Measures overlap between predicted and ground truth regions.
- Formula: $\text{Dice} = \frac{2TP}{FP+2TP+FN}$

Uncertainty Quantification: Calibration Metrics

- **Expected Calibration Error (ECE)**:
  - Measures the gap between predicted probability and actual accuracy.
  - Output probabilities are binned; for each bin E_b, calculate:
  - Accuracy: $\text{acc}(E_b) = \frac{1}{|E_b|} \sum_{n \in E_b} \mathbb{1}(S_n = G_n)$
  - Confidence: $\text{conf}(E_b) = \frac{1}{|E_b|} \sum_{n \in E_b} P_n$
  - ECE: $\text{ECE} = \sum_{b=1}^{B} \frac{|E_b|}{N} |\text{acc}(E_b) - \text{conf}(E_b)|$
- **Brier Score (BS)**:
  - Measures the accuracy of probability predictions.
  - Formula: $\text{BS} = \frac{1}{N} \sum_{n=1}^{N} (P_n - G_n)^2$
- **Negative Log-Likelihood (NLL)**:
  - Penalizes incorrect predictions, more for higher-confidence errors.
  - Formula: $\text{NLL} = -\frac{1}{N} \sum_{n=1}^{N} [G_n \log P_n + (1 - G_n) \log(1 - P_n)]$

Table of Contents

Models for PET-CT Lymphoma Segmentation

- **1. UNet with Softmax (Baseline)**:
  - Standard UNet with a softmax layer for voxel classification.
- **2. UNet with Monte-Carlo (MC) Dropout and Deep Ensemble**:
  - Uses **MC dropout** and deep ensembling to enhance uncertainty quantification.
- **3. ENN-UNet**:
  - UNet with Evidential Neural Network (ENN) as decision module.
  - Replaces softmax layer with EM (Evidential Mapping) module.
- **4. RBF-UNet**:
  - UNet with Radial Basis Function (RBF) module for decision-making.
  - Produces output belief functions similar to ENN-UNet.
- **5. MMEF-UNet**(ours):
  - It consists of encoder–decoder feature extraction (FE) modules for deep feature representation, evidence mapping (EM) modules to convert features into mass functions, and a multimodal evidence fusion (MMEF) module to integrate evidence across modalities.

Results for PET-CT Lymphoma Segmentation

Table 1
Means and standard errors of segmentation quality and reliability measures for MMEF-UNet and the referenced uncertainty quantification methods on the lymphoma dataset. The best results are in bold and the second best are underlined.

Model	ECE↓	Brier score ↓	NLL↓	Dice score ↑
UNet	$0.056 \pm 3.6 \times 10^{-3}$	$0.065 \pm 3.9 \times 10^{-3}$	$0.310 \pm 8.8 \times 10^{-2}$	$0.770 \pm 3.2 \times 10^{-2}$
UNet-MC	$0.053 \pm 4.6 \times 10^{-3}$	$0.062 \pm 4.9 \times 10^{-3}$	$0.400 \pm 8.7 \times 10^{-2}$	$0.801 \pm 1.1 \times 10^{-2}$
UNet-Ensemble	$0.063 \pm 7.6 \times 10^{-2}$	$0.064 \pm 4.0 \times 10^{-3}$	$0.343 \pm 7.2 \times 10^{-2}$	$0.802 \pm 6.7 \times 10^{-3}$
ENN-UNet	$\underline{0.050} \pm 3.5 \times 10^{-3}$	$0.062 \pm 3.9 \times 10^{-3}$	$\underline{0.191} \pm 1.4 \times 10^{-2}$	$\underline{0.805} \pm 7.1 \times 10^{-3}$
RBF-UNet	$0.051 \pm 3.3 \times 10^{-3}$	$\underline{0.061} \pm 0.9 \times 10^{-3}$	$0.193 \pm 1.3 \times 10^{-2}$	$0.802 \pm 6.9 \times 10^{-3}$
MMEF-UNet (ours)	$\mathbf{0.045} \pm 1.3 \times 10^{-3}$	$\mathbf{0.056} \pm 2.7 \times 10^{-3}$	$\mathbf{0.180} \pm 1.3 \times 10^{-2}$	$\mathbf{0.811} \pm 3.4 \times 10^{-2}$

Figure: Benchmark

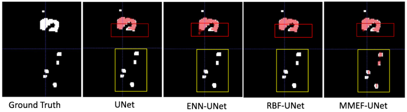


Figure: Examples of visualized segmentation results

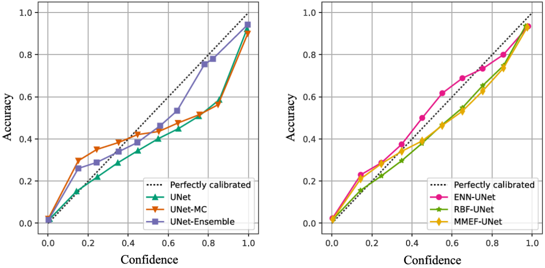


Figure: Calibration plots for probabilistic(left) and evidential (right) deep segmentation models

Analysis of reliability coefficients.

Table 2
Estimated reliability coefficient β_k^t (means and standard errors) after training for the background and lymphoma classes and the two modalities. Higher values correspond to greater contribution to the segmentation.

β_k^t	Background	Lymphomas
PET	$0.999 \pm 8.9 \times 10^{-3}$	$0.996 \pm 4.5 \times 10^{-3}$
CT	$0.863 \pm 1.8 \times 10^{-2}$	$0.975 \pm 8.9 \times 10^{-3}$

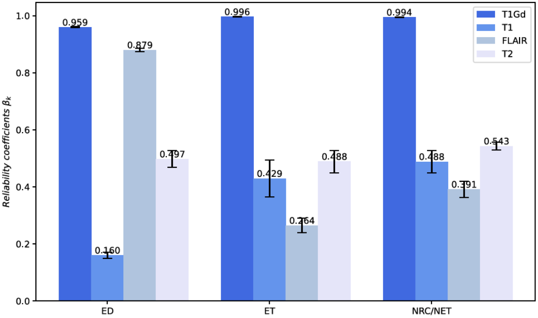


Figure: Estimated reliability coefficients β_k after training of MMEF-nnFormer for classes ED, ET, and NRC/NET in the four modalities

Table of Contents

Summary and Future Directions

Summary

- Proposed a decision-level fusion architecture for multi-modality medical image segmentation.
- Extracts UNet features from each modality, maps them to Dempster–Shafer mass functions, applies contextual discounting for reliability, and fuses using Dempster's rule.
- Evaluated on PET-CT (lymphoma) and multi-MRI (brain tumor) datasets, showing improved segmentation accuracy and uncertainty quantification over pixel-level fusion methods and models with softmax layers.
- Reliability coefficients align with domain knowledge, offering insights into the fusion process.

Future Directions

- **Broader Applications**: Extend DST-based fusion to other biomedical data (signals, biomarkers, genomics) and fields like remote sensing (e.g., fusion of Lidar, SAR, hyperspectral data) and human–machine interaction.
- **Enhanced Fusion Methods**: Consider combining entire mass functions rather than only contour functions, enabling richer outputs for more advanced decision strategies (e.g., partial classification).