Meta Evidential Transformer for Few-Shot Open-Set Recognition

Hitesh Sapkota, Krishna Prasad Neupane, and Qi Yu

Rochester Institute of Technology

July, 2024

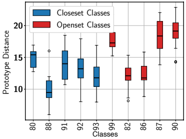
Overview

- Few Shot Open Set Recognition (FSOSR) aims to detect instances from unknown classes by utilizing a small set of labeled instances from closed-set classes.
- FSOSR can benefit many critical domains such as fraud detection and surveillance systems, where labeled data is typically limited and novel samples may occur frequently.
- Existing techniques (e.g., PEELER and SnaTCHer) are not designed to handle the challenging scenarios, where the open-set samples share similarities with closed-set samples.
- To tackle these challenging cases, we propose **Meta Evidential Transformer (MET)** that leverages
  - an evidential open-set loss to learn more compact closed-set representations during training process,
  - an evidence-to-variance ratio (EVR) to identify challenging open-set samples during inference process

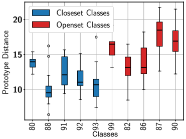
FSOSR: Key Challenges

Ferrets (88) Golden Retriever (82) Malamute (83)

(a) Open-set sample (golden retriever) shares similar features with closed-set samples including ferrets and malamute.



(b) SnaTCHer



(c) MET

- Challenging learning scenarios:
  1. General: Novel samples share some similarity with closed-set ones.
  2. Extreme: A closed-set class is very distinct from other closed-set classes.
- An open-set golden retriever in the above figure may be misclassified as closed-set classes (e.g., Ferrets and Malamute) due to their mutual similarities.
- Existing techniques are not designed to handle such challenging situations leading to sub-optimal FSOSR performance

Solution: Learn Compact Closed-Set Representation

- **Key idea:** Expose the model to "strong opponent classes" chosen from the closed-set classes to learn more compact closed-set representations
- **Meta Training:**
  1. Learn to assign a high uncertainty to the opponent classes serving as a training-time open-set sample by leveraging novel open-set evidential loss.
  2. Use a novel open-set score instead of entropy as a higher entropy does not tell whether a sample is close to multiple-set classes or far from all of them (open-set case)
- **Meta Testing:**
  1. Due to the normalization effect of attention layers in the transformer, the open-set sample in the extremely challenging situation will likely be assigned to a special closed-set class
  2. Propose **Evidence-to-variance ratio (EVR)** to identify those samples during inference

Training: Evidence Guided Training

- **Dataset Construction:**
  - Construct a meta-training set consisting of only training classes with no overlap with samples from meta-test classes
  - Choose a set of opponent classes from existing known closed-set classes to serve as open-set classes aiming to learn more compact representations.
- **Training:**

$$\boldsymbol{\theta}^* = \arg\min_{\theta} \Bigg\{ \sum_{(\mathbf{x}_j, y_j) \in T_i^{tr} | y_j \in C^s} \mathcal{L}_{close}\left(y_j, P_{\boldsymbol{\theta}}(.|\mathbf{x}_j, S_i^{tr})\right) + \lambda \sum_{(\mathbf{x}_j, y_j) \in T_i^{tr} | y_j \in C^u} \mathcal{L}_{open}\left(P_{\boldsymbol{\theta}}(.|\mathbf{x}_j, S_i^{tr})\right) \Bigg\} \quad (1)$$

  - **Open-set samples:** Shrink total evidence towards zero for the opponent open-set classes by minimizing the following open-set loss

$$\mathcal{L}_{open}(\cdot) = \sum_{j | y_j \in C^u} KL[\mathrm{Dir}(\mathbf{p}_j | \boldsymbol{\alpha}_j) || \mathrm{Dir}(\mathbf{p}_j | (1, ..., 1)^\top)] \quad (2)$$

Inference: Evidential Cross-Attention

- **For the challenging FSOSR task** involving special class, the evidence distribution within the closed-set classes exhibit high variance and relatively low maximum evidence. As such, **Evidence-to-variance ratio (EVR)** remains low in these difficult tasks

$$\mathrm{EVR}_i = \frac{1}{|\mathcal{Q}_i^{te}|} \sum_{j \in \mathcal{Q}_i^{te}} \frac{\max_{n \in N}[e_{jn}]}{\mathrm{var}_{n \in N}[e_{jn}]} \tag{3}$$

- For easy tasks, the ratio remains high resulting in the EVR for challenging tasks being lower than that of easy tasks. By leveraging EVR, the attention can be adjusted as:

$$A_i[c_1, c_2] = \left\{ \begin{array}{ll} A_i[c_1, c_2] \times \frac{\epsilon}{\mathrm{EVR}_i} & \text{if } cond == true \\ A_i[c_1, c_2] & \text{else} \end{array} \right\}$$
$$cond = \{(c_1 == c || c_2 == c) \; \& \; c_1 \neq c_2\} \tag{4}$$

- In challenging cases, by drastically altering the attention weights, the original prototype will significantly differ from the altered prototype representation leading to improved OSR.

FSOSR Performance

Approaches	MiniImageNet 5-way		TieredImagenet 5-way	
	1-shot	5-shot	1-shot	5-shot
ProtoNet	51.63 ± 0.47	60.26 ± 0.56	58.48 ± 0.50	63.46 ± 0.24
RelationNet	53.14 ± 0.67	62.22 ± 0.78	60.85 ± 0.68	64.42 ± 0.57
OpenMAX	71.67 ± 0.87	76.75 ± 0.80	62.27 ± 0.55	70.92 ± 0.52
FEAT (Probability)	45.00 ± 0.70	53.82 ± 0.78	57.14 ± 0.57	63.94 ± 0.52
Feat (Distance)	67.71 ± 0.92	75.32 ± 0.84	61.52 ± 0.58	70.77 ± 0.52
PEELER	60.36 ± 0.72	68.45 ± 0.78	58.24 ± 0.65	66.14 ± 0.74
SnaTCHer	67.37 ± 0.91	77.99 ± 0.76	71.00 ± 0.66	79.49 ± 0.47
TANE	73.23 ± 0.25	81.15 ± 0.18	74.89 ± 0.64	80.45 ± 0.49
MET	$\mathbf{76.93 \pm 0.59}$	$\mathbf{84.90 \pm 0.41}$	$\mathbf{78.77 \pm 0.46}$	$\mathbf{84.37 \pm 0.35}$

- We use AUROC as an evaluation metric for FSOSR which is higher the better.

Ablation Study

Transformer	Evidential Loss	EVR	AUROC	
			1-shot	5-shot
✓			67.37	77.99
✓	✓		74.35	81.47
✓	✓	✓	**76.93**	**84.90**

- Ablation study on MiniImageNet dataset shows that:
  1. Evidential loss brings performance improvement of more than 6.5% in a 1-Shot setting whereas more than 3% in a 5-Shot setting.
  2. EVR further boosts the performance by more than 3% in 5-Shot and more than 1.5% in 1-Shot setting.

Conclusions

- Propose a new evidential open-set loss to learn more compact closed-set representations by leveraging similar closed-set classes as opponent open-set classes.
- Propose a novel evidence-to-variance ratio (EVR) to identify challenging open-set samples.
- Propose a uniquely designed evidence-based cross-attention mechanism.
- Show state-of-the-art FSOSR performance in multiple real-world datasets.

Poster

More detailed information will be in the Poster with **ID: 1308 (Hall C 4-9 #2314)**.