南京航空航天大學
Nanjing University of Aeronautics and Astronautics

模式分析与机器智能
工业和信息化部重点实验室
MIIT Key Laboratory of
Pattern Analysis & Machine Intelligence

TALISMAN： Targeted Active Learning for Object Detection with Rare Classes and Slices using Submodular Mutual Information

Suraj Kothawade , Saikat Ghosh , Sumit Shekhar , Yu Xiang , and Rishabh Iyer
University of Texas at Dallas

ECCV 2022

1 Introduction

While deep models perform well in terms of overall accuracy, they often struggle in performance on rare yet critical data slices. For e.g., detecting objects in rare data slices like "motorcycles at night" or "bicycles at night" for self-driving applications.

Current AL based acquisition functions are not well-equipped to mine rare slices of data from large real-world datasets, since they are based on uncertainty scores or global descriptors of the image.

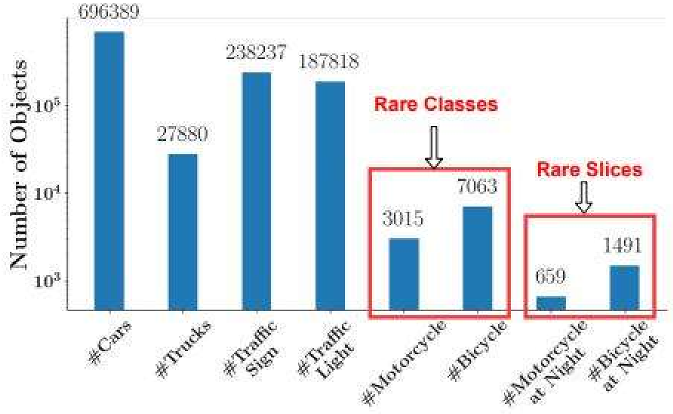

Fig. 1: Problem Statement: Rare classes and Rare slices in BDD100K [31]. Motorcycle and bicycle classes have the least number of objects, thereby making them *rare classes*, on which the model performs the worst in terms of average precision (AP). Further, motorcycle/bicycle objects at night are *rarer*, thereby making them *rare slices* on which the model performs the worst.

1 Introduction- Contributions

(a) provides a mechanism to encode the similarity between an unlabeled image and a small query set of targeted examples (e.g., images with "motorcycles at night" RoIs)

(b) mines these examples in a scalable manner from a large unlabeled set using the recently proposed submodular mutual information functions

key difference : Talisman does the selection by targeting rare slices using only a few exemplars

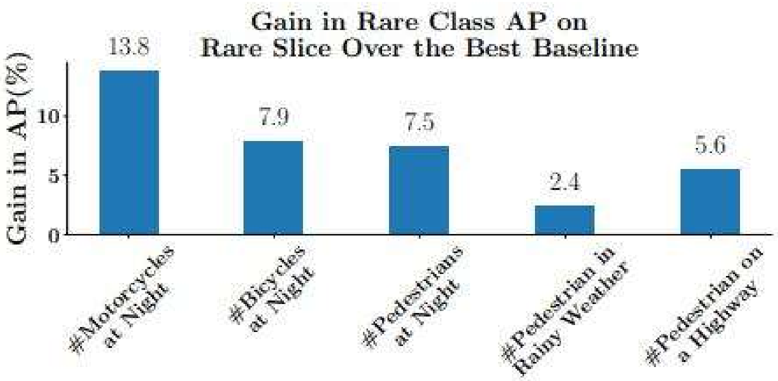

Fig. 2: Efficiency of TALISMAN over the best-performing baseline on a variety of rare slices in BDD100K

1 Introduction

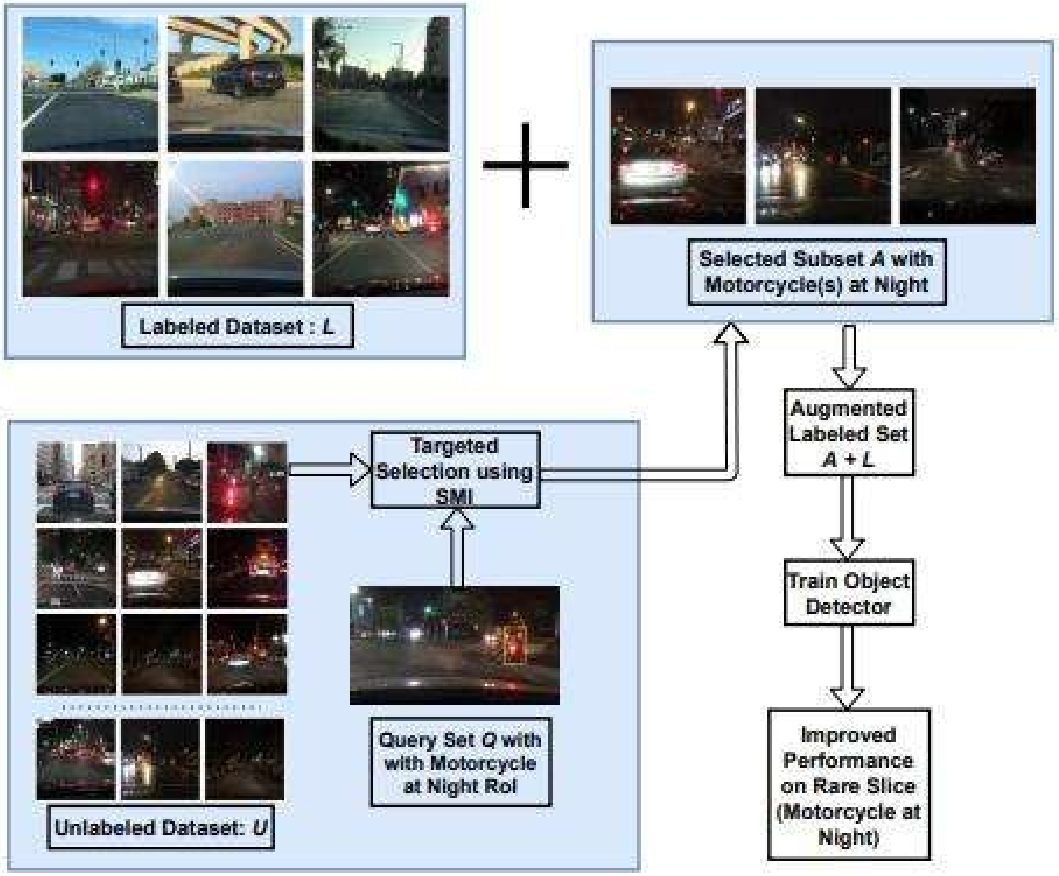

Fig. 3: Targeted Selection using TALISMAN for one round of targeted active learning. Motorcycles at night is a rare slice in the labeled data. We mine images from the unlabeled set that semantically similar to the RoIs in the query set by using the submodular mutual information (SMI) functions. These images are then labeled and added to the labeled data to improve performance on the rare slice.

2 Methodology-Submodular Functions

unlabeled set : $\mathcal{U} = \{1, 2, 3, \cdots, n\}$

set function : $f : 2^{\mathcal{U}} \to \mathbb{R}$

submodular function : $f(\mathcal{A} \cup x) - f(\mathcal{A}) \geq f(\mathcal{B} \cup x) - f(\mathcal{B}), \forall \mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{U}$ and $x \notin \mathcal{B}$.

optimization problem : $\max_{\mathcal{A}:|\mathcal{A}| \leq \mathcal{B}} f(\mathcal{A})$

a constant factor approximation : $1 - \frac{1}{e}$

2 Methodology-Submodular Mutual Information (SMI)

$$I_f(\mathcal{A}; \mathcal{Q}) = f(\mathcal{A}) + f(\mathcal{Q}) - f(\mathcal{A} \cup \mathcal{Q}).$$

$$\max_{\mathcal{A} \subseteq \mathcal{U}, |\mathcal{A}| \leq B} I_f(\mathcal{A}; \mathcal{Q})$$

Specific SMI Functions Used In TALISMAN

Facility Location(FL) : models representation (i.e., it picks the most representative points or “centroids”)

FLMI : the FL based SMI function , models representation as well as query relevance

Graph Cut(GC) : models diversity and representation, and has modeling properties similar to Fl

GCMI : the SMI variant of Gc, maximizes the pairwise similarity between the query set and the unlabeled set

Table 1: Instantiations of different submodular functions.

(a) Instantiations of Submodular functions.

SF	$f(\mathcal{A})$
FL	$\sum_{i \in \mathcal{U}} \max_{j \in \mathcal{A}} S_{ij}$
GC	$\sum_{i \in \mathcal{A}, j \in \mathcal{U}} S_{ij} - \sum_{i,j \in \mathcal{A}} S_{ij}$

(b) Instantiations of SMI functions.

SMI	$I_f(\mathcal{A}; \mathcal{Q})$
FLMI	$\sum_{i \in \mathcal{Q}} \max_{j \in \mathcal{A}} S_{ij} + \sum_{i \in \mathcal{A}} \max_{j \in \mathcal{Q}} S_{ij}$
GCMI	$2 \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{Q}} S_{ij}$

2 Methodology-TALISMAN Framework

Algorithm 1 TALISMAN: Targeted AL Framework for Object Detection (Illustration in Fig. 4)

Require: Initial labeled set of data points: $\mathcal{L}$, large unlabeled dataset: $\mathcal{U}$, small query set $\mathcal{Q}$, object detection model $\mathcal{M}$, batch size: B, number of selection rounds: N.

for selection round $i = 1 : N$ **do**

  Train model $\mathcal{M}$ on the current labeled set $\mathcal{L}$ and obtain parameters θ_i

  Compute $S \in \mathbb{R}^{|\mathcal{Q}| \times |\mathcal{U}|}$ such that: $S_{qu} \leftarrow \textsc{TargetedSim}(\mathcal{M}_{\theta_i}, \mathcal{I}_q, \mathcal{I}_u)$, $\forall q \in \mathcal{Q}, \forall u \in \mathcal{U}$ {Algorithm 2}

  Instantiate a submodular function f based on S.

  $\mathcal{A}_i \leftarrow \text{argmax}_{\mathcal{A} \subseteq \mathcal{U}, |\mathcal{A}| \leq B} I_f(\mathcal{A}; \mathcal{Q})$ {Greedy maximization of SMI function to select a subset $\mathcal{A}$}

  Get labels $L(\mathcal{A}_i)$ for batch $\mathcal{A}_i$ and $\mathcal{L} \leftarrow \mathcal{L} \cup L(\mathcal{A}_i)$, $\mathcal{U} \leftarrow \mathcal{U} - \mathcal{A}_i$

end for

Return trained model $\mathcal{M}$ and parameters θ.

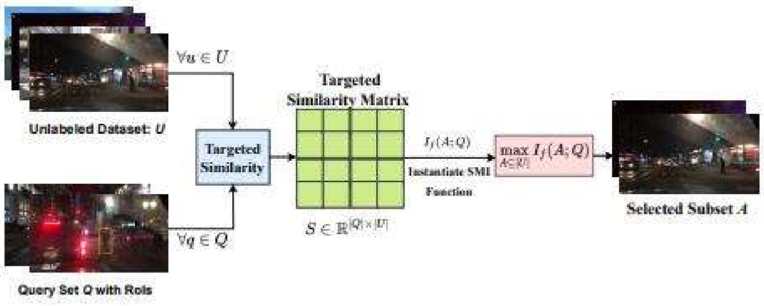

Fig. 4: Architecture of TALISMAN during one round of targeted active learning. We illustrate the targeted similarity computation in Fig. 5.

2 Methodology-Targeted Similarity Computation

Algorithm 2 TARGETEDSIM: Targeted Similarity Matching (Illustration in Fig. 5)

Require: Local feature extraction model F_θ, $\mathcal{I}_q \in \mathcal{Q}$ with T RoIs and $\mathcal{I}_u \in \mathcal{U}$ with P region proposals.

1: $\mathcal{E}_q \leftarrow F_\theta(\mathcal{I}_q)$ $\{\mathcal{E}_q \in \mathbb{R}^{T\times D}\}$

2: $\mathcal{E}_u \leftarrow F_\theta(\mathcal{I}_u)$ $\{\mathcal{E}_u \in \mathbb{R}^{P\times D}\}$

3: $\mathcal{X}_{qu} \leftarrow$ COSINE_SIMILARITY$(\mathcal{E}_q, \mathcal{E}_u)$ $\{\mathcal{X}_{qu} \in \mathbb{R}^{T\times P}$. Compute Cosine similarity along the feature dimension$\}$

4: $S_{qu} \leftarrow \max(\mathcal{X}_{qu})$ {Element-wise Max, S_{qu} represents the score between the best matching proposal $j \in P$ to some query RoI $i \in T$}

5: **Return** Similarity score S_{qu}

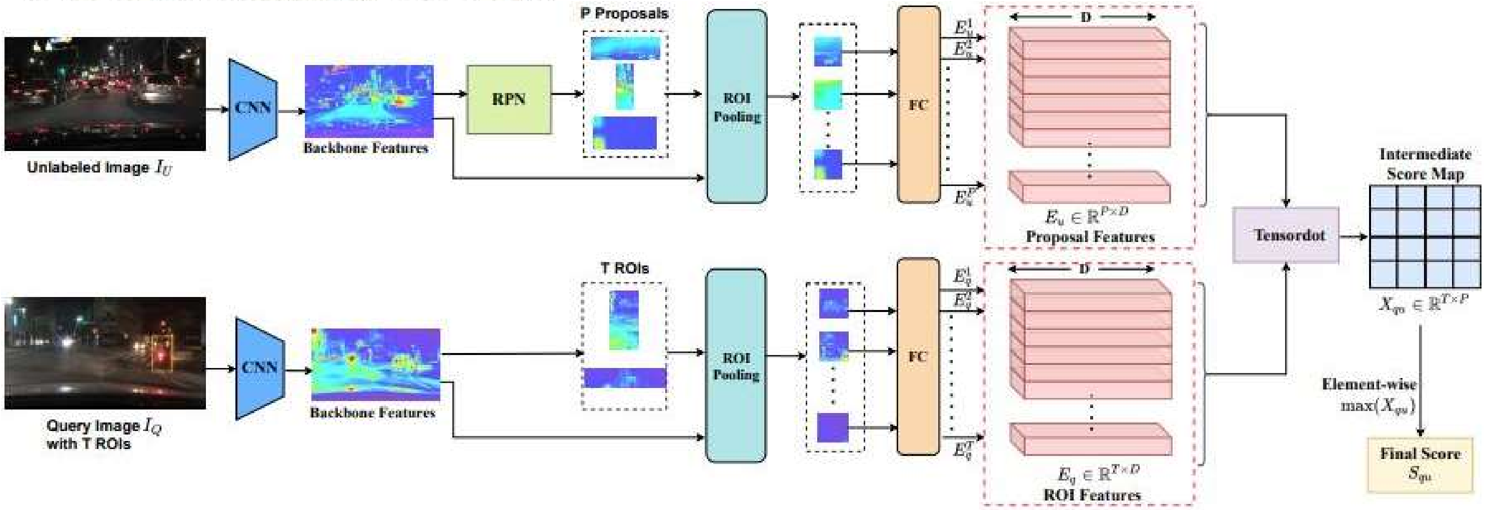

Fig. 5: TARGETEDSIM: Targeted Similarity computation in TALISMAN.

3 Experiment- Setup

Metrics: ①the mean average precision (mAP) ②average precision (AP) of the rare slice ③the number of data points selected that belong to the rare slice

Datasets: ① PASCAL VOC07+12 ② BDD100K

Baselines in all scenarios: Entropy, Targeted Entropy (T-Entropy), Least Confidence(Least-Conf), Margin, Fass, Coreset, Badge and Random sampling

For all experiments on both datasets, train a Faster RCNN model based on a ResNet50 backbone

3 Experiment- Rare Classes

Dataset : VOC07+12 dataset

initial labeled set L : create a class imbalance at an object level to simulate the rare classes

imbalance ratio : $\rho \leq (\mathcal{B}_j^{\mathcal{L}} / \mathcal{C}_i^{\mathcal{L}})$

$\mathcal{B}_j^{\mathcal{L}}$ the number of objects from a frequent class j

$\mathcal{C}_i^{\mathcal{L}}$ the number of objects from a rare (infrequent) class i

unlabeled set U : remaining data

query set Q : containing 5 randomly chosen data points representing the rare classes (RoIs)

choose two classes to be rare : 'boat' and 'bottle'

3 Experiment- Rare Classes

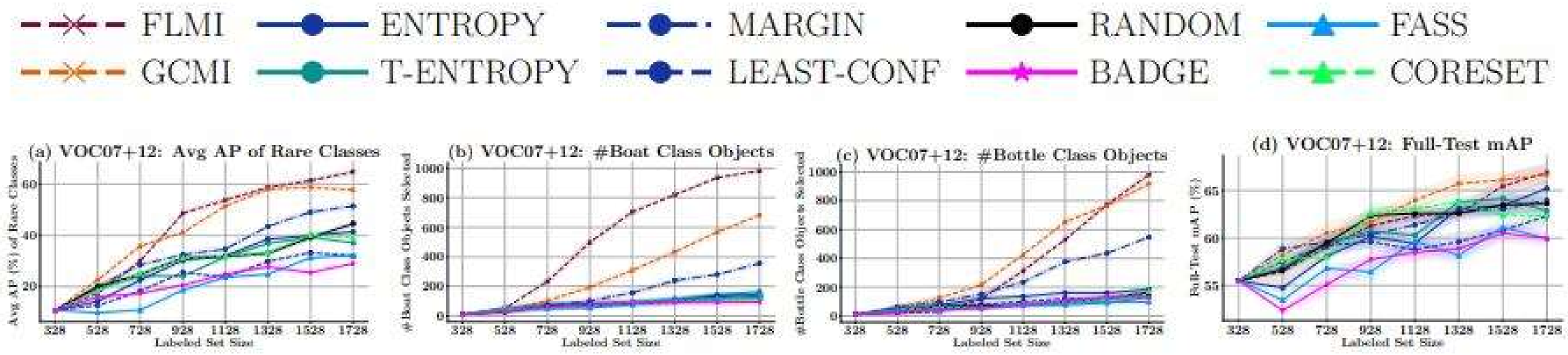

Fig. 6: Active Learning with rare classes on VOC07+12. Plot (a) shows the average AP of the rare classes, plots (b-c) show the number of boat and bottle objects selected respectively, plot (d) shows the mAP on the VOC07+12 test set. We observe that the SMI functions (FLMI, GCMI) outperform other baselines by $\approx 8\% - 10\%$ average AP of the rare classes.

Gcmi and Flqmi are able to select more data points that contain regions with objects belonging to the rare classes
and give a fair treatment to multiple rare classes at the same time by selecting significant number of objects belonging to both the rare classes

3 Experiment- Rare Slices

dataset: BDD100K

$|\mathcal{O}_c^A|$ **: the number of objects in L that belong to class c and attribute A**

initial labeled set L : have a rare slice made of a class and an attribute. For instance, motorcyles (class) atnight (attribute), class-balanced, attribute-imbalanced

imbalance ratio of class: $\rho \leq (|\mathcal{O}_i^{\bar{A}}|/|\mathcal{O}_i^A|)$

unlabeled set U : remaining data

query set Q : containing 5 randomly chosen data points representing the rare classes (RoIs).

choose two classes to be rare : 'boat' and 'bottle'

3 Experiment- Rare Slices

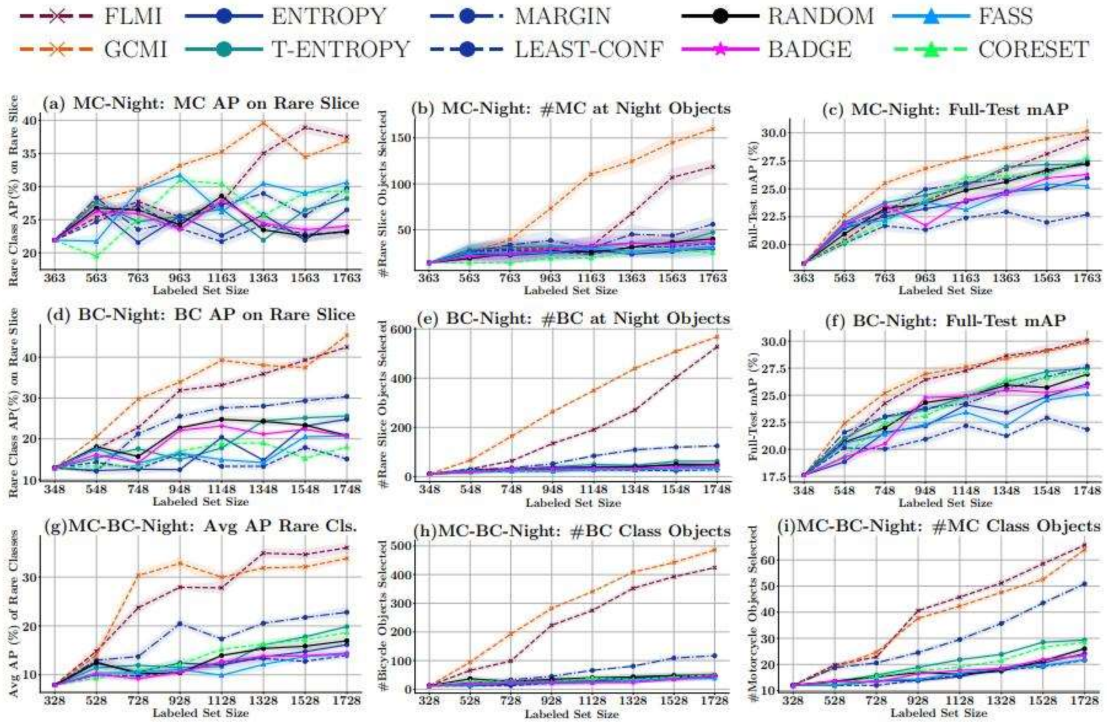

Fig. 7: Active Learning with Motorcycle (**MC**) at Night (top row) and Bicycle (**BC**) at Night (bottom row) rare slices on BDD100K. Left side plots (a,d,g) show the AP of the rare class on the rare slice of data, center plots (b,e) show the number of objects selected that belong to the rare slice, and right side plots (c,f) show the mAP on the full test set of BDD100K. We observe that the SMI functions (FLMI, GCMI) outperform other baselines by $\approx 5\% - 14\%$ AP of the rare class on the rare slice. In (h,i), we show that TALISMAN selects more objects from multiple rare slices in comparison to the existing methods.

Fig. 9: A false negative motorcycle at night (left) fixed to a true positive detection (right) using TALISMAN.

3 Experiment- Rare Slices

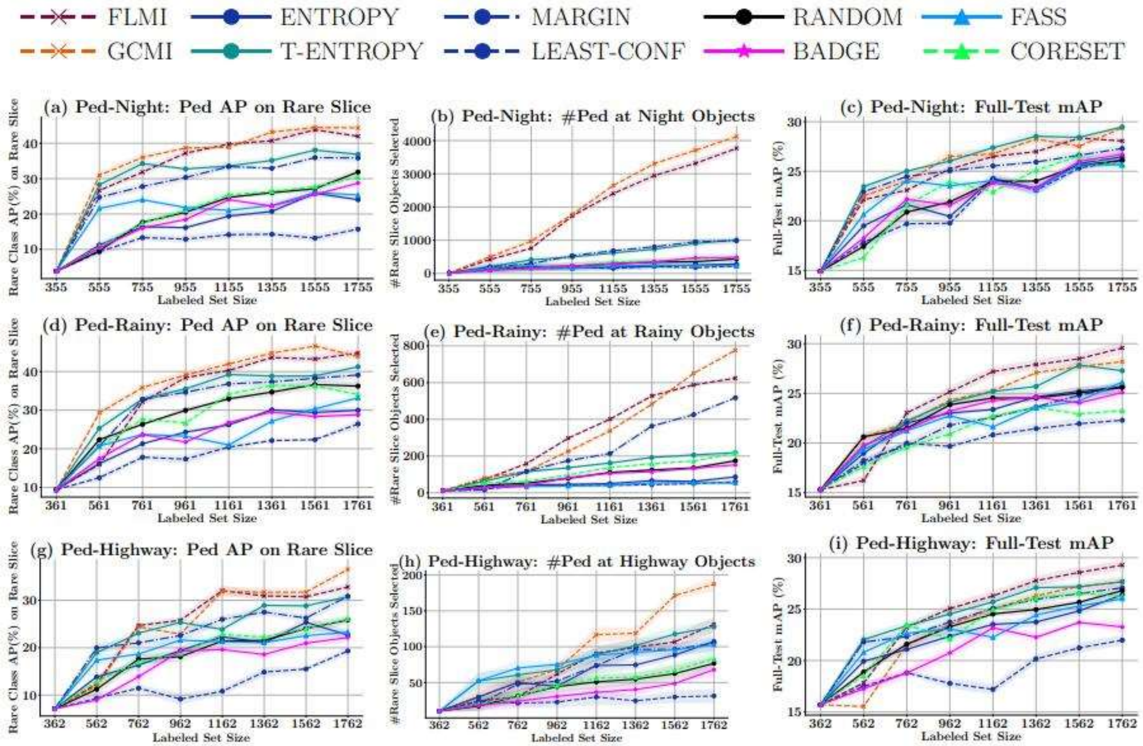

Fig. 8: AL with Pedestrian (**Ped**) at Nighttime (top row), Pedestrian in Rainy Weather (middle row), and Pedestrian on a Highway (bottom row) rare slices on BDD100K. Left side plots (a,d,g) show the AP of the rare class on the rare slice of data, center plots (b,e,h) show the number of objects selected that belong to the rare slice, and right side plots (c,f,i) show the mAP on the full test set of BDD100K. We observe that the SMI functions (FLMI, GCMI) outperform other baselines by $\approx 5\% - 10\%$ AP of the pedestrian class on the rare slice.

THANKS