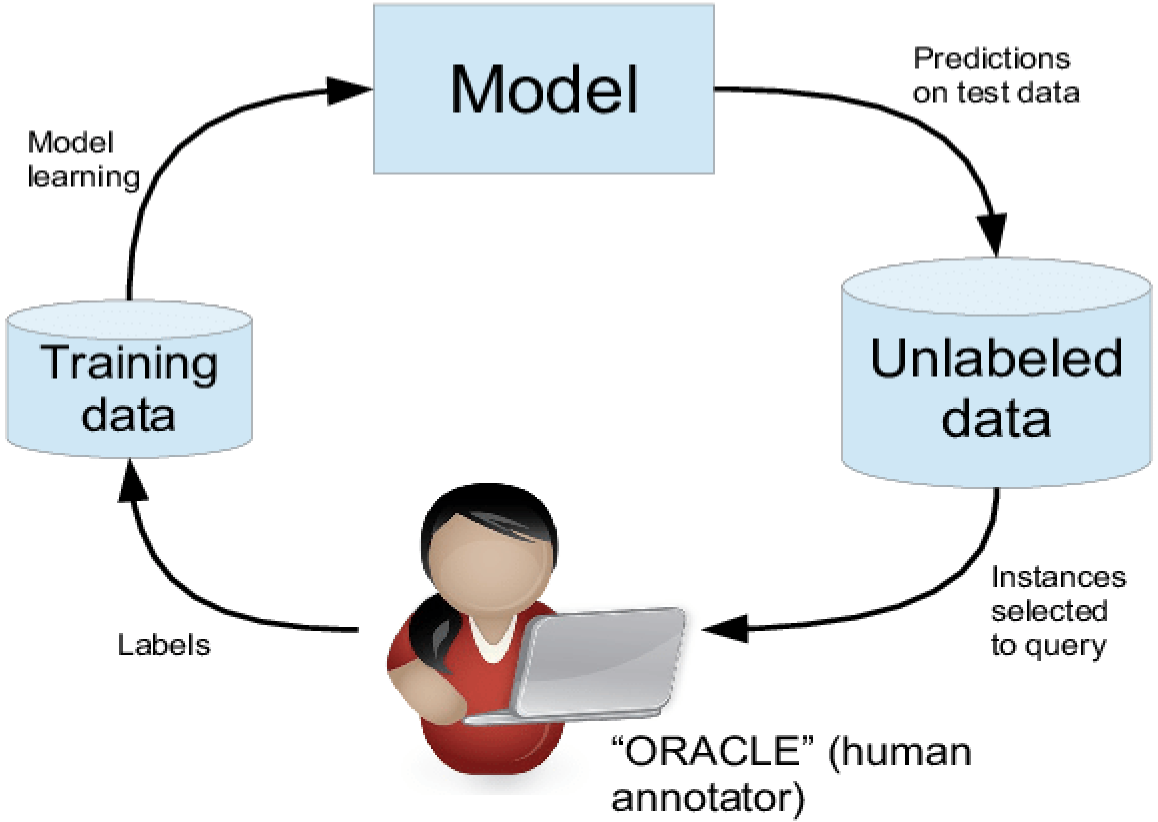
TALISMAN
Targeted Active Learning for Object Detection with Rare Classes and Slices using Submodular Mutual Information

Suraj Kothawade*, Saikat Ghosh, Sumit Shekhar, Yu Xiang, Rishabh Iyer

*suraj.kothawade@utdallas.edu

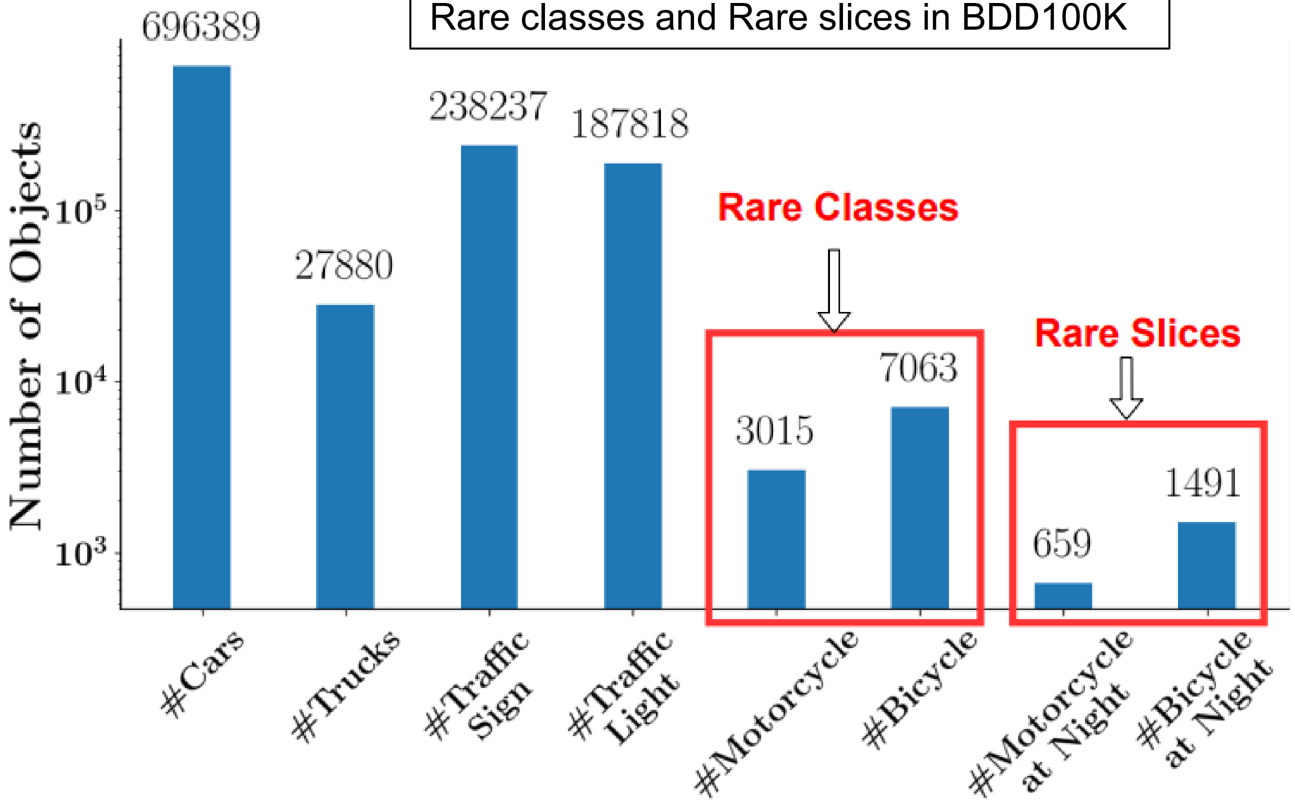
Active Learning



Castillo et al., Industrial statistics and manifold data

Rare Classes and Rare Slices

Rare classes and Rare slices in BDD100K

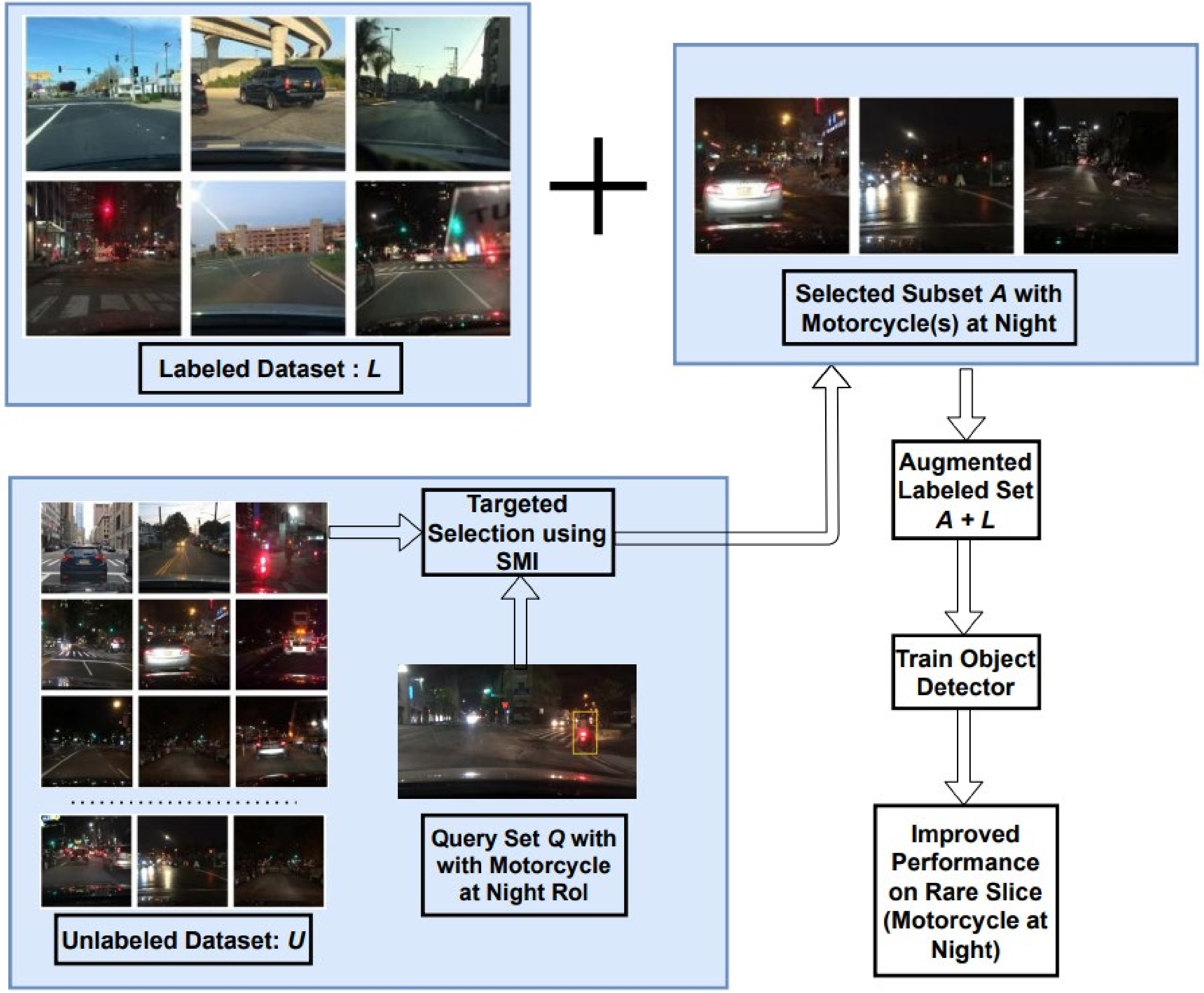


Pedestrian in the dark snapshot from a self-driving car**

- Motorcycle and bicycle classes have the least number of objects, thereby making them rare classes, on which the model suffers.
- Further, motorcycle/bicycle objects at night are rarer, thereby making them rare slices on which the model performs the worst.

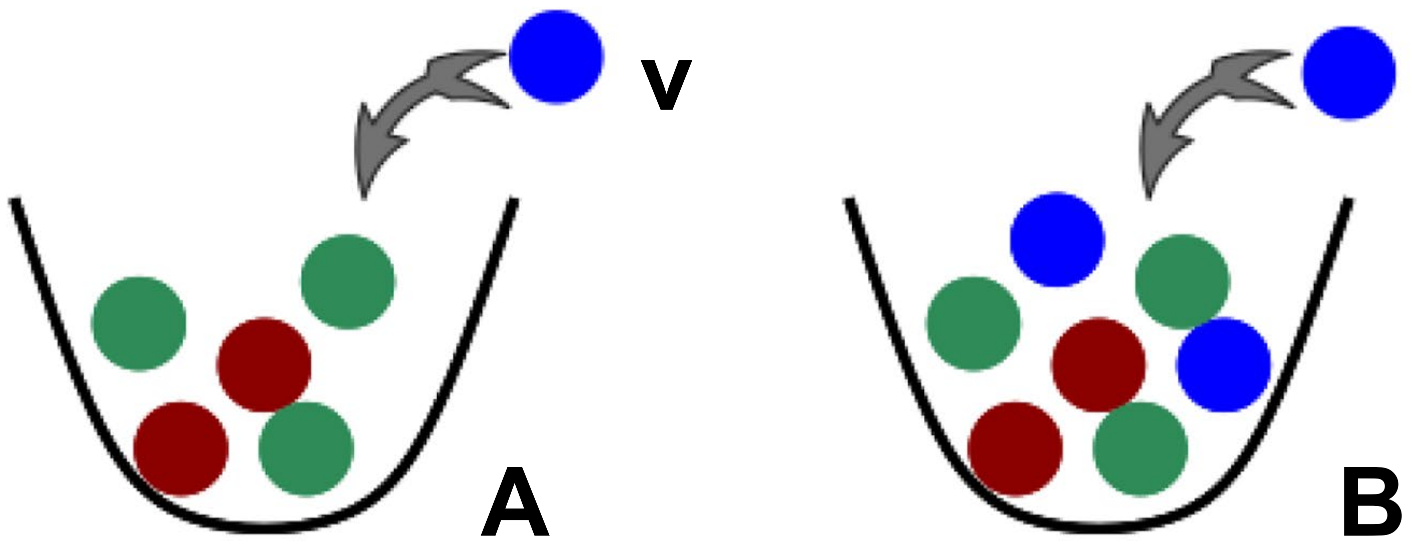
**Uber self-driving car crash in Tempe, Arizona.

Targeted Active Learning



Submodular Functions

$$f(A \cup v) - f(A) \geq f(B \cup v) - f(B), \text{ if } A \subseteq B$$



$f = \#$ of distinct colors of balls in the urn.

Information Theoretic Concepts

- **Entropy:** Given a set of random variables $X_1 \cdots, X_n$, the Entropy of a **subset** of random variables: $H(X_A) = -\sum_{X_A} P(X_A) \log P(X_A)$. Note that entropy is **submodular.**

Information Theoretic Concepts

- **Entropy:** Given a set of random variables $X_1 \cdots, X_n$, the Entropy of a **subset** of random variables: $H(X_A) = -\sum_{X_A} P(X_A) \log P(X_A)$. Note that entropy is **submodular.**
- **Mutual Information:** Given a set of random variables, $X_1, \cdots, X_n$ and sets $A, B \subseteq V$, the Mutual Information $\mathrm{I}(X_A; X_B) = H(X_A) + H(X_B) - H(X_{A \cup B})$

Can we replace H with any submodular function?

Can we replace H with any submodular function?

YES!

This gives us the Submodular Mutual Information!

Submodular Information Measures (SIM)

- Given a set of data points $V = \{1, \cdots, n\}$**,** and sets $A, Q \subseteq U$**,** the **Submodular Mutual Information (SMI)** $I_F(A; Q) = F(A) + F(Q) - F(A \cup Q)$**,** where the information of a **set** of points is $F(A)$ and F is a submodular function.

Submodular Mutual Information (SMI)

(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	$2 \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{Q}} S_{ij}$

Iyer at al., ALT 2021

TALISMAN: Targeted Active Learning Framework for Object Detection

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

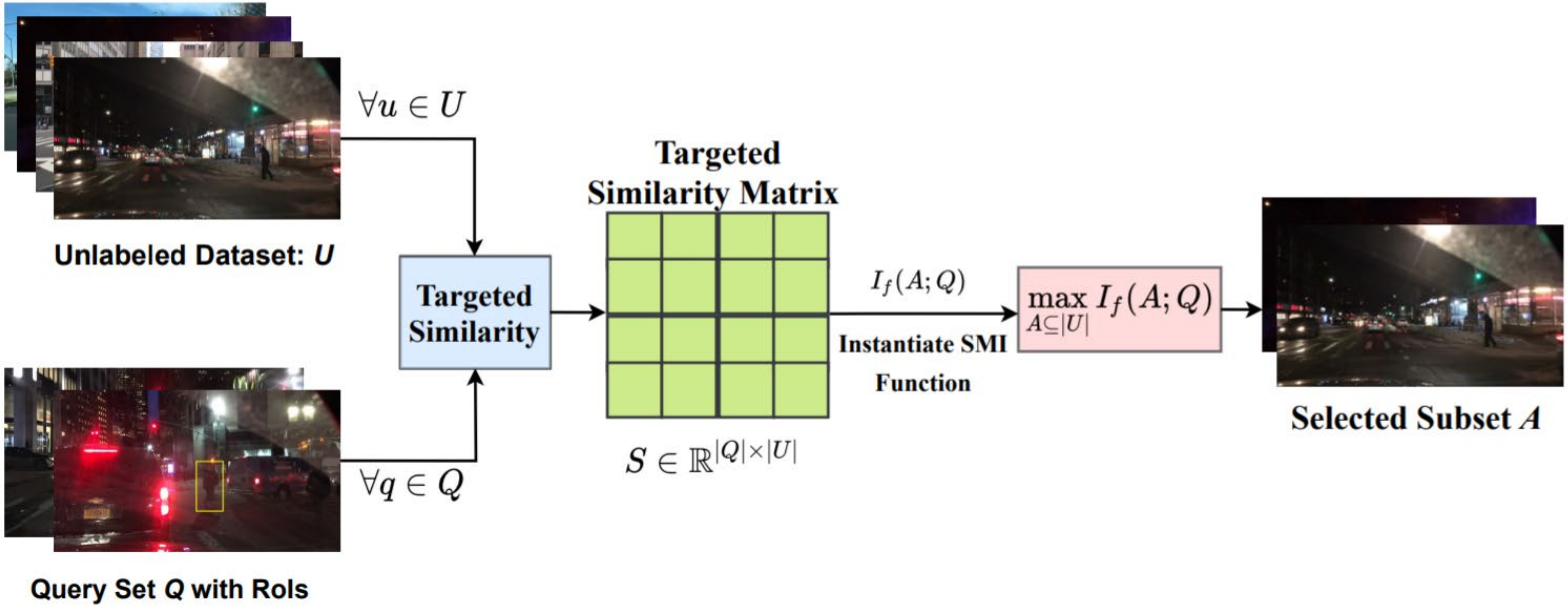
1. **for** selection round $i = 1 : N$ **do**
2. Train model $\mathcal{M}$ with loss $\mathcal{H}$ on the current labeled set $\mathcal{L}$ and obtain parameters θ_i
3. Compute $S \in R^{|Q| \times |U|}$ such that: $S_{qu} \leftarrow TargetedSim(\mathcal{M}_{\theta_i}, I_q, I_u), \forall q \in \mathcal{Q}, \forall u \in \mathcal{U}$
4. Instantiate a submodular function f based on S
5. $\mathcal{A}_i \leftarrow \mathrm{argmax}_{\mathcal{A} \subseteq \mathcal{U}, |\mathcal{A}| \leq B}\, I_f(\mathcal{A}; \mathcal{Q})$ (Greedy maximization of SMI to select a subset $\mathcal{A}$)
6. Get labels $L(\mathcal{A}_i)$ for batch $\mathcal{A}_i$ and $\mathcal{L} \leftarrow \mathcal{L} \cup L(\mathcal{A}_i), \mathcal{U} \leftarrow \mathcal{U} - \mathcal{A}_i$
7. **end for**
8. **return** trained model $\mathcal{M}$ and parameters θ_N

TALISMAN: Targeted Active Learning Framework for Object Detection

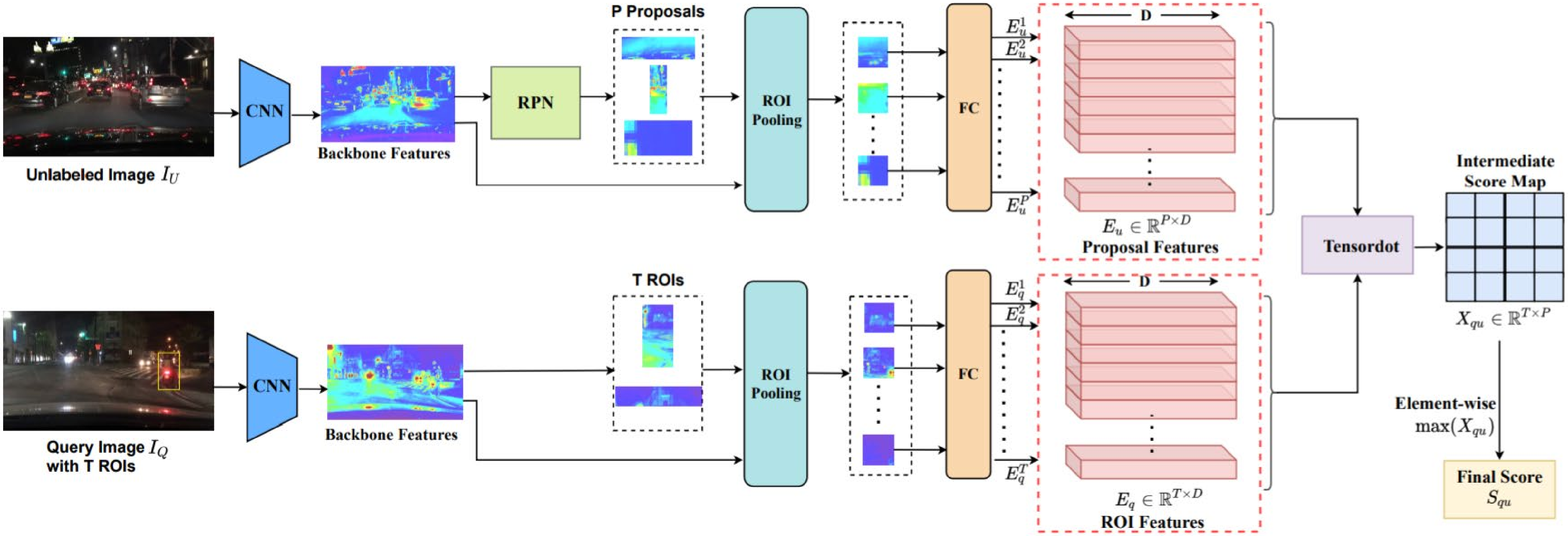
Require: Initial labeled set of data points: $\mathcal{L}$, large unlabeled dataset: $\mathcal{U}$, object detection model $\mathcal{H}$ for learning model $\mathcal{M}$, batch size: B, number of selection rounds: N

1. **for** selection round $i = 1 : N$ **do**
2. Train model $\mathcal{M}$ with loss $\mathcal{H}$ on the current labeled set $\mathcal{L}$ and obtain parameters θ_i
3. Compute $S \in R^{|Q| \times |U|}$ such that: $S_{qu} \leftarrow TargetedSim(\mathcal{M}_{\theta_i}, I_q, I_u), \forall q \in \mathcal{Q}, \forall u \in \mathcal{U}$
4. Instantiate a submodular function f based on S
5. $\mathcal{A}_i \leftarrow \mathrm{argmax}_{\mathcal{A} \subseteq \mathcal{U}, |\mathcal{A}| \leq B}\ I_f(\mathcal{A}; \mathcal{Q})$ (Greedy maximization of SMI to select a subset $\mathcal{A}$)
6. Get labels $L(\mathcal{A}_i)$ for batch $\mathcal{A}_i$ and $\mathcal{L} \leftarrow \mathcal{L} \cup L(\mathcal{A}_i), \mathcal{U} \leftarrow \mathcal{U} - \mathcal{A}_i$
7. **end for**
8. **return** trained model $\mathcal{M}$ and parameters θ_N

Architecture of TALISMAN



Targeted Similarity Computation

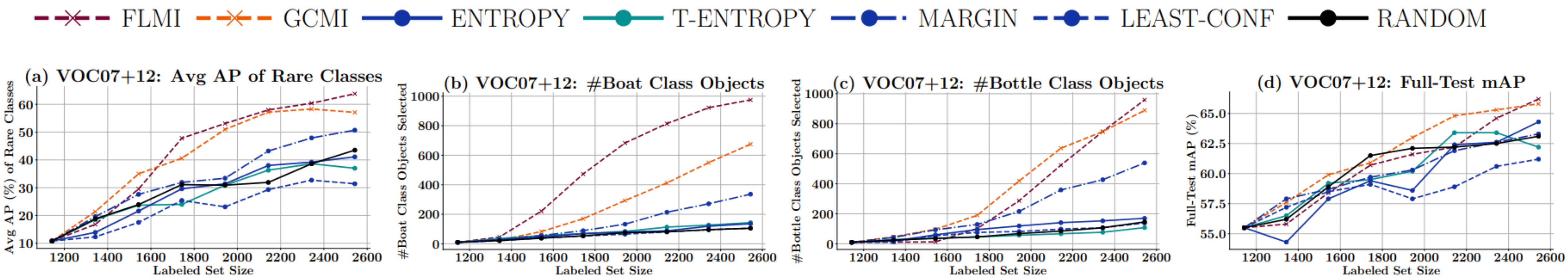


TALISMAN: Targeted Active Learning Framework for Object Detection

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

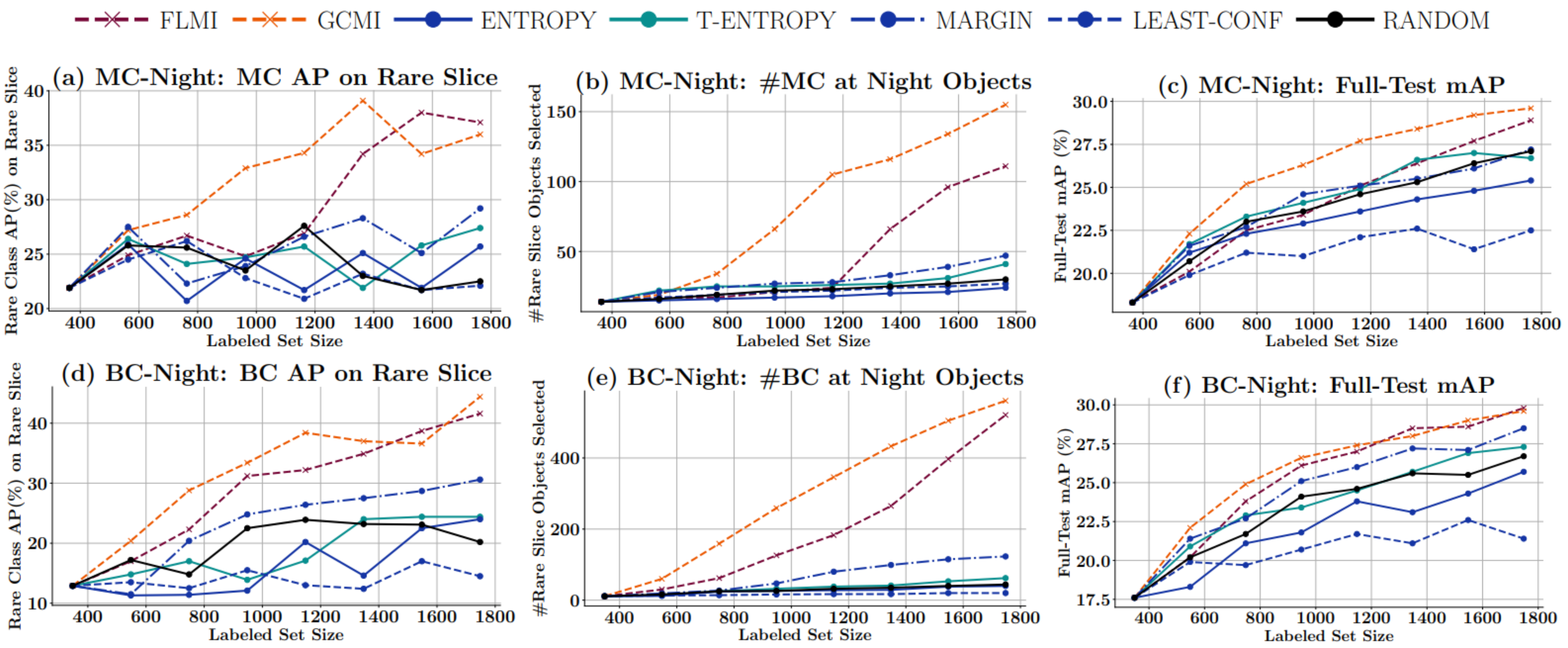
1. **for** selection round $i = 1 : N$ **do**
2. Train model $\mathcal{M}$ with loss $\mathcal{H}$ on the current labeled set $\mathcal{L}$ and obtain parameters θ_i
3. Compute $S \in R^{|Q|\times|U|}$ such that: $S_{qu} \leftarrow TargetedSim(\mathcal{M}_{\theta_i}, I_q, I_u), \forall q \in \mathcal{Q}, \forall u \in \mathcal{U}$
4. Instantiate a submodular function f based on S
5. $\mathcal{A}_i \leftarrow \text{argmax}_{\mathcal{A}\subseteq\mathcal{U},|\mathcal{A}|\leq B}\, I_f(\mathcal{A};\mathcal{Q})$ (Greedy maximization of SMI to select a subset $\mathcal{A}$)
6. Get labels $L(\mathcal{A}_i)$ for batch $\mathcal{A}_i$ and $\mathcal{L} \leftarrow \mathcal{L} \cup L(\mathcal{A}_i), \mathcal{U} \leftarrow \mathcal{U} - \mathcal{A}_i$
7. **end for**
8. **return** trained model $\mathcal{M}$ and parameters θ_N

Active Learning with Rare Classes on VOC07+12



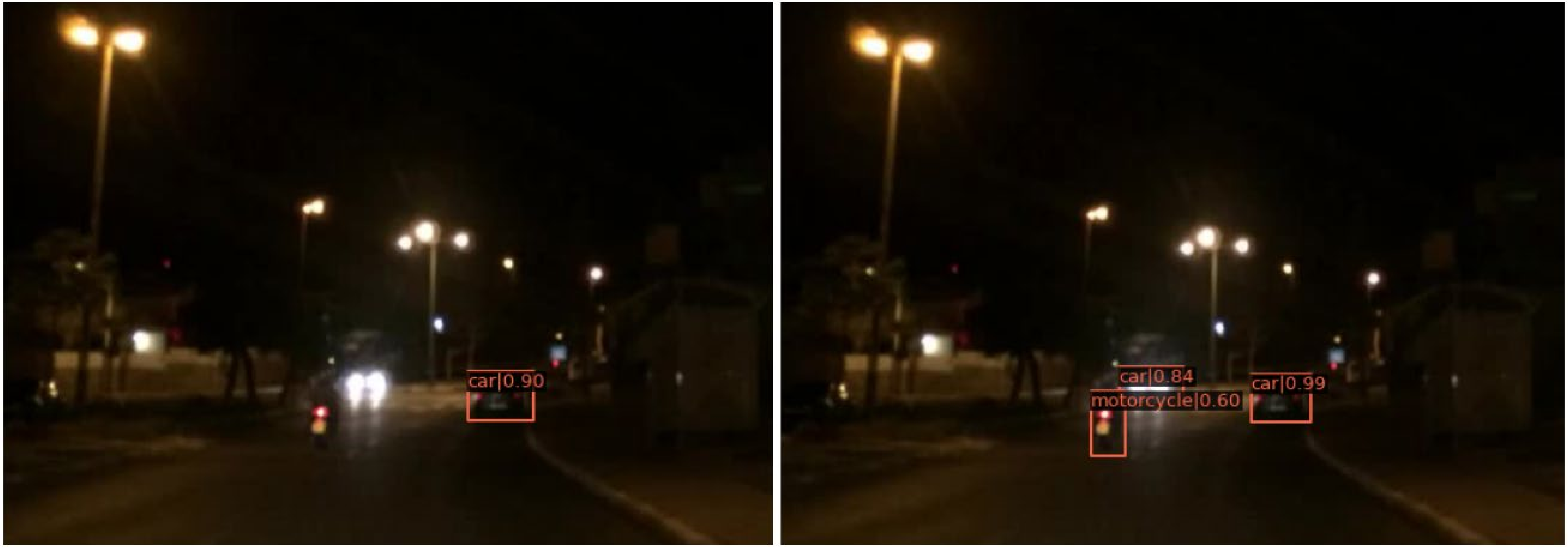
- Active Learning with rare classes on VOC07+12. Plot (a) shows the average AP of the rare classes, plots (b-c) show the number of boat and bottle objects selected respectively, plot (d) shows the mAP on the VOC07+12 test set.
- We observe that the SMI functions (FLMI, GCMI) outperform other baselines by ≈ 8% − 10% average AP of the rare classes.
- SMI functions are fair in selecting objects for both classes when the query set comprises of objects for both classes.

Active Learning with Rare Slices on BDD100K



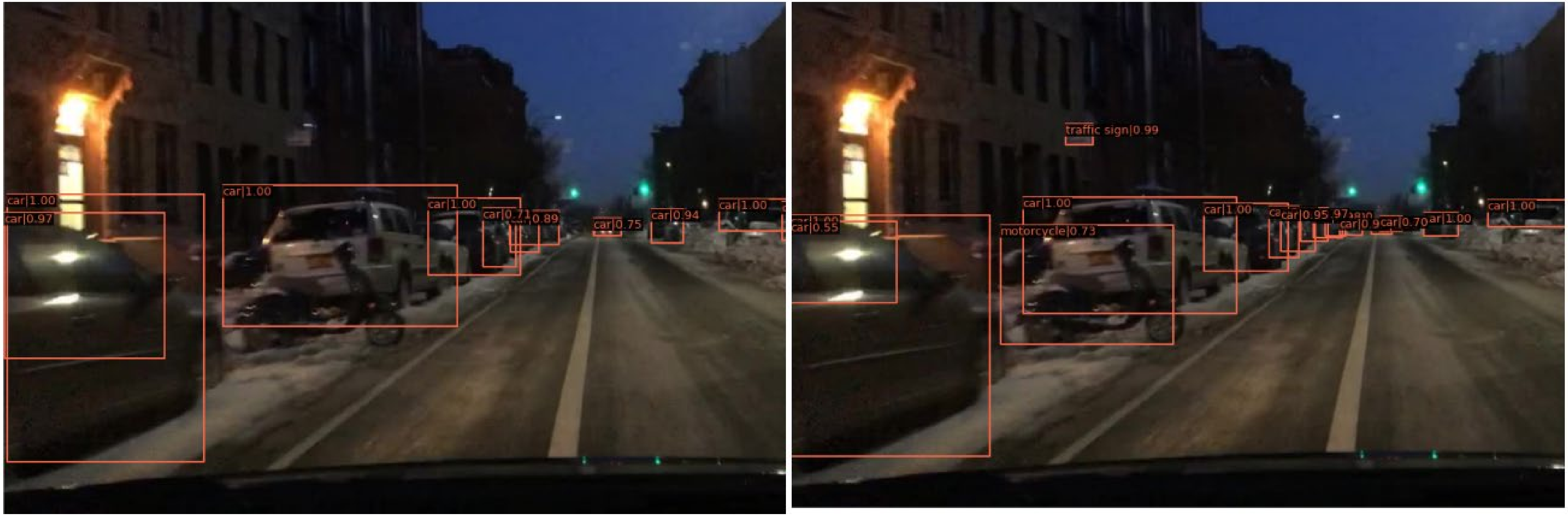
- We observe that the SMI functions (FLMI, GCMI) outperform other baselines by ≈ 5% − 14% AP of the rare class on the rare slice and by ≈ 2% − 3% in terms of mAP
- This improvement is because the SMI functions are able to target the rare slice region by using the ROI in the query image.

Qualitative Results



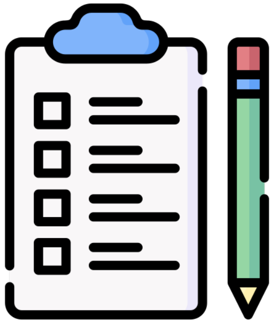
A false negative motorcycle at night on the road (left) fixed to a true positive detection (right) using TALISMAN.

Qualitative Results



A false negative motorcycle at night parked on the road (left) fixed to a true positive detection (right) using TALISMAN.

Conclusion



- In this paper, we present a targeted active learning framework TALISMAN that enables improving the performance of object detection models on rare classes and slices.

- We showed the utility of our framework across a variety of real-world scenarios of rare classes and slices on the PASCAL VOC07+12 and BDD100K driving dataset.

- Using the SMI functions, we observe a $\approx 5\% - 12\%$ gain compared to the existing baselines.

Thank You