Exemplar-Based Open-Set Panoptic Segmentation Network

Jaedong Hwang[1]* Seoung Wug Oh[2] Joon-Young Lee[2] Bohyung Han[1]

[1]Seoul National University [2]Adobe Research

(* work done during the internship program at Adobe Research)

Project Page

CVPR VIRTUAL JUNE 19-25

❑ Open-Set Panoptic Segmentation

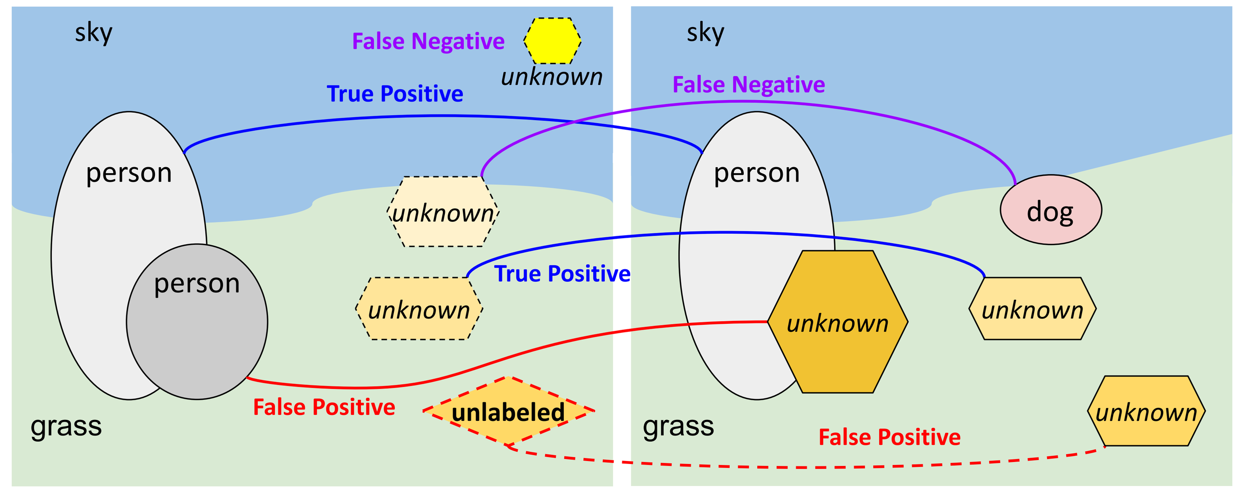
- **Labels**

  - *Class*
    - *Thing*
      - *Known*
      - *Unknown*
        - *Seen Unknown*
        - *Unseen Unknown*
    - *Stuff*
    - *Void*

- **Evaluation Metrics**

$$\mathrm{PQ} = \underbrace{\frac{\sum_{(p,g)\in TP} \mathrm{IoU}(p,g)}{|TP|}}_{\text{segmentation quality (SQ)}} \cdot \underbrace{\frac{|TP|}{|TP| + \frac{1}{2}|FP| + \frac{1}{2}|FN|}}_{\text{recognition quality (RQ)}}$$

- **Challenges**
  - Finding *unknown* instances is more difficult.
  - Annotating ∀ objects is almost impossible.
  - Find unlabeled *unknown* instances ⟹ PQ↓, RQ↓



Ground-truth Prediction

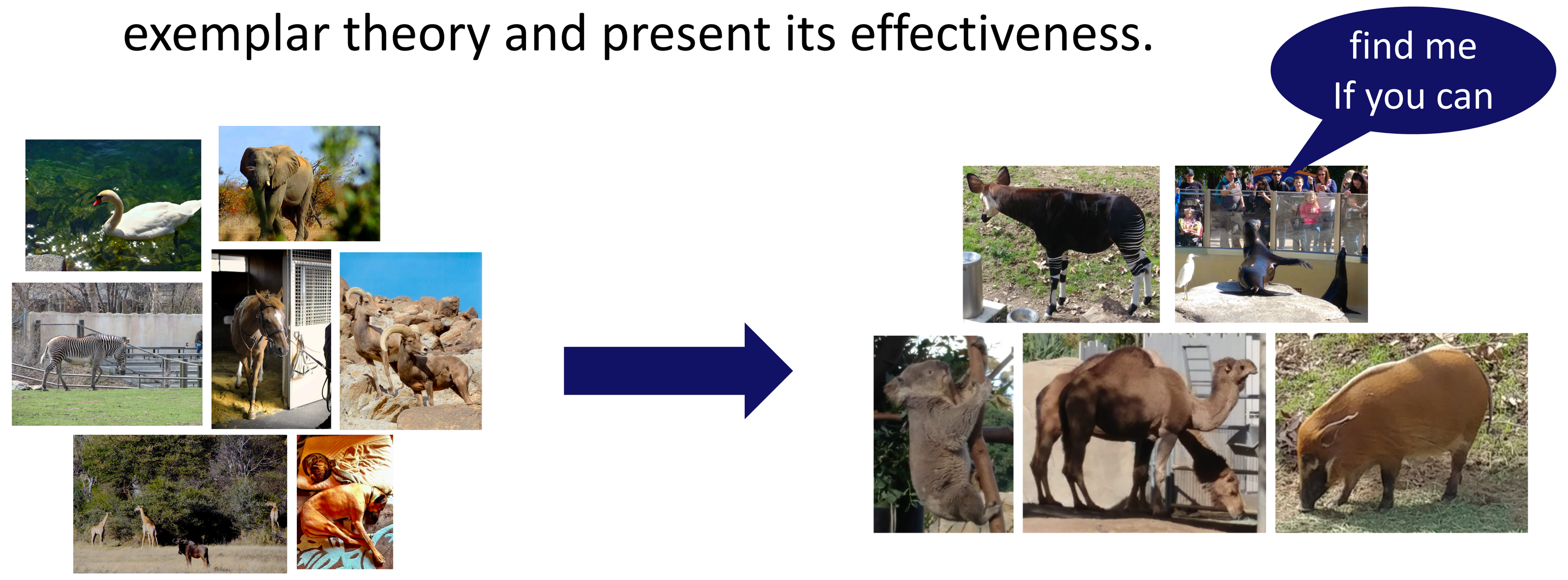
- **Assumptions for preventing label conflict**
  - ∀ *unknown* class ⊂ *thing* class.
  - Parts of *known objects* ≠ *unknown* class (e.g., tire in *car*).
  - *Unknown* classes in training images only appear on *void* regions.

- **Dataset based on COCO [1]**
  - We construct different splits with different ratio, *K* of unknown classes among *thing* classes.
    - **5%:** *car, cow, pizza, toilet*
    - **10%:** (5%) + *boat, tie, zebra, stop sign*
    - **20%:** (10%) + *dining table, banana, bicycle, cake, sink, cat, keyboard, bear*
  - We remove annotations for *unknown* classes during training.

❑ Introduction

- **Motivation**
  - Panoptic segmentation is widespread
  - **Real-world is an open-world** and annotating all instances and backgrounds are difficult
  - In exemplar theory, people categorize a new object by comparing exemplars of each class.

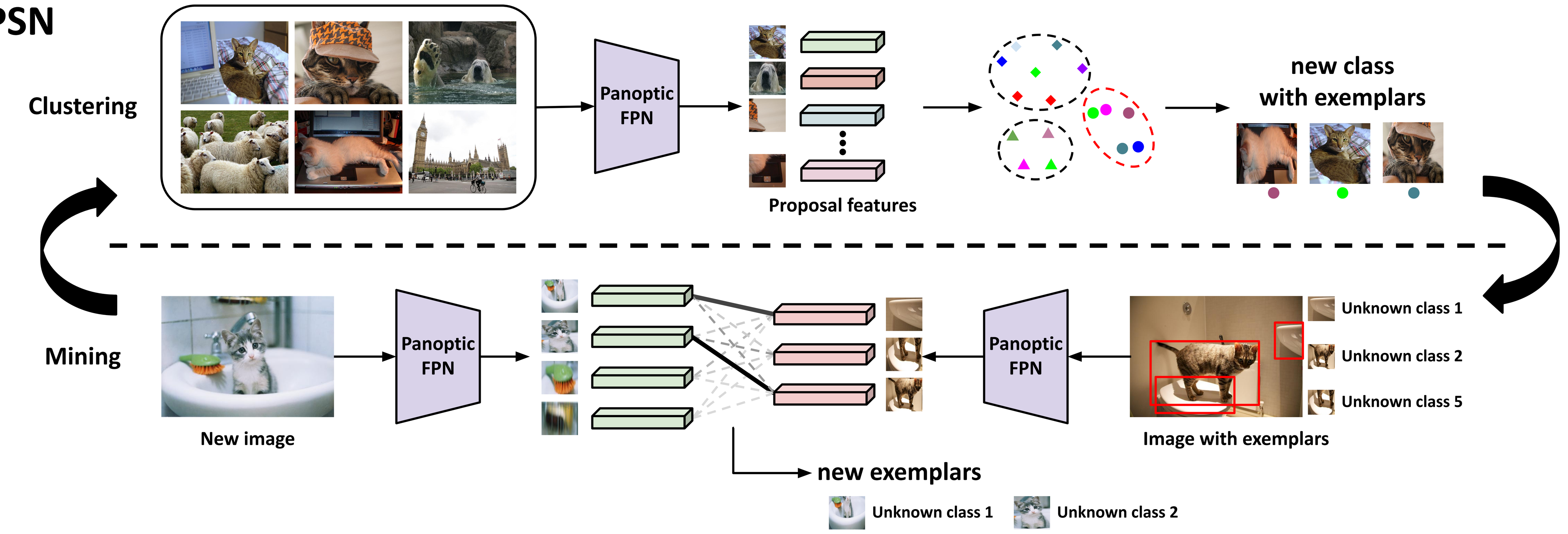
- **Our contributions**
  - We define **open-set panoptic segmentation (OPS)** task and make it feasible using reasonable assumptions through in-depth analysis of its inherent challenges.
  - We **construct a brand-new OPS benchmark** by reformatting COCO [1] and present performance of several baselines based on Panoptic FPN [2].
  - We propose a novel framework for OPS, **EOPSN** based on the exemplar theory and present its effectiveness.



Known classes Unknown classes

❑ Method

- **EOPSN**



- **Clustering – find unknown class and exemplars**
  - Store candidate proposals with object features.
  - K-means over-clustering and remove noisy clusters.
  - Remining clusters → *unknown* classes
  - Elements in the clusters → exemplars

- **Mining – discover more exemplars**
  - Extract features of proposals and exemplars
  - Compare features of proposals with exemplars.
  - **sim(**proposal, exemplar**)** > *threshold*

    ⟹ the proposal is a new exemplar

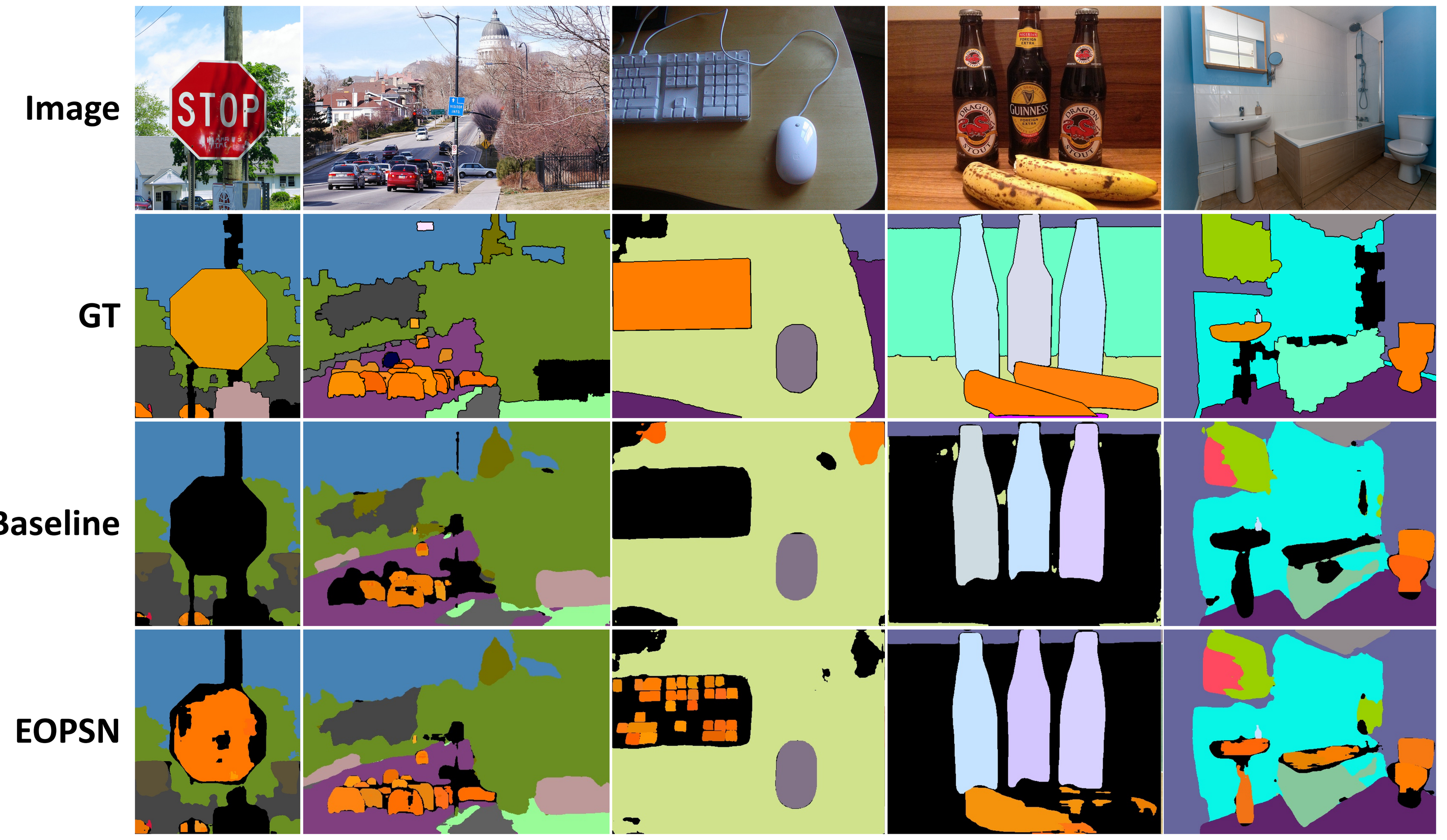
- **Baseline – Panoptic FPN [2]**
  - **Class-agnostic** regressor and mask generator.
  - Segment *known* first and then *unknowns* using objectness score from RPN.
  - 4 different baselines with utilization of boxes on *void* regions
    - Train as background (*void-background*)
    - Ignore during training (*void-ignorance*)
    - Avoid labeling as known (*void-suppression*)

      $$\mathcal{L}_{\text{void}} = \sum_{c\in\mathcal{C}^{\text{Th}}} -\log(1-p_c)$$

    - Train as a new class (*void-train*)

❑ Experiments

- **Qualitative Results (orange denotes *unknown* classes).**



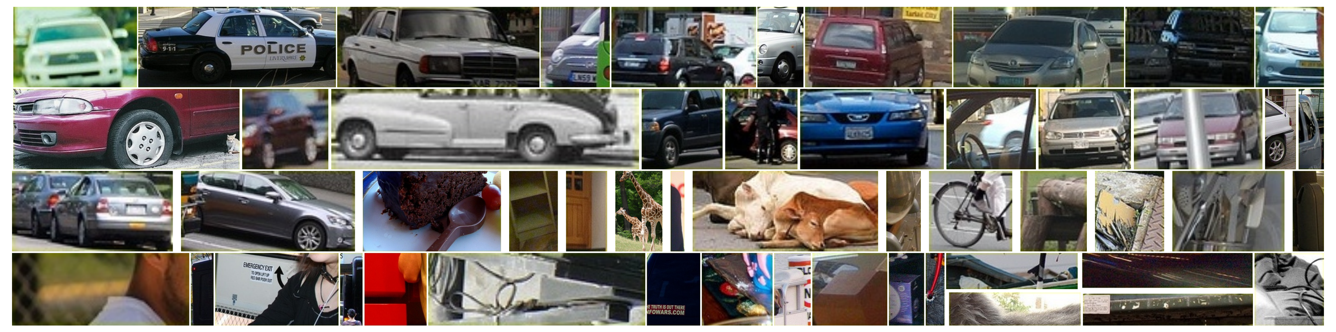
- **Baselines on COCO validation set with *K*=10%.**

Utilization of *void* regions	Known PQ	Known SQ	Known RQ	Known $\mathrm{PQ}^{\mathrm{Th}}$	Known $\mathrm{SQ}^{\mathrm{Th}}$	Known $\mathrm{RQ}^{\mathrm{Th}}$	Known $\mathrm{PQ}^{\mathrm{St}}$	Known $\mathrm{SQ}^{\mathrm{St}}$	Known $\mathrm{RQ}^{\mathrm{St}}$	Unknown PQ	Unknown SQ	Unknown RQ
Void-background	**37.7**	76.3	**46.6**	44.8	79.3	54.1	**29.2**	72.8	**37.5**	4.0	71.1	5.7
Void-ignorance	37.2	76.3	45.9	43.9	79.0	53.1	29.1	**73.0**	37.3	3.7	71.8	5.2
Void-suppression	37.5	75.9	46.1	**45.1**	**80.6**	**54.5**	28.2	70.2	36.1	7.2	**75.3**	9.6
Void-train	36.9	**76.4**	45.5	44.0	80.3	53.3	28.2	71.7	36.0	**7.8**	73.4	**10.7**

- **Comparison of Baseline and EOPSN with various *K*.**

K(%)	Model	Known PQ	Known SQ	Known RQ	Known $\mathrm{PQ}^{\mathrm{Th}}$	Known $\mathrm{SQ}^{\mathrm{Th}}$	Known $\mathrm{RQ}^{\mathrm{Th}}$	Known $\mathrm{PQ}^{\mathrm{St}}$	Known $\mathrm{SQ}^{\mathrm{St}}$	Known $\mathrm{RQ}^{\mathrm{St}}$	Unknown PQ	Unknown SQ	Unknown RQ
	Supervised	39.4	77.7	48.4	45.8	80.7	55.4	29.7	73.1	38.0	-	-	-
5	Baseline (*Void-train*)	37.7	76.7	46.4	44.2	80.4	53.5	28.3	71.3	36.2	10.0	73.8	13.5
	EOPSN	**38.0**	**76.9**	**46.8**	**44.8**	**80.5**	**54.2**	28.3	**71.9**	36.2	**23.1**	**74.7**	**30.9**
10	Baseline (*Void-train*)	36.9	75.4	45.5	43.2	79.0	52.4	28.3	70.4	36.2	8.5	73.2	11.6
	EOPSN	**37.7**	**76.8**	**46.3**	**44.5**	**80.6**	**53.8**	**28.4**	**71.8**	36.2	**17.9**	**76.8**	**23.3**
20	Baseline (*Void-train*)	36.9	**76.4**	45.5	44.0	80.3	53.3	28.2	**71.7**	36.0	7.8	73.4	10.7
	EOPSN	**37.4**	76.2	**46.2**	**45.0**	80.3	**54.5**	28.2	71.2	**36.2**	**11.3**	**73.8**	**15.3**

- **Exemplars in a found unknown class**



❑ Discussion

- Only tackle *seen unknown* **but** # of images↑ ⟹ # of *seen unknown*↑.
- New evaluation metrics are needed.

Reference

[1] Lin, Tsung-Yi, et al. "Microsoft coco: Common objects in context." In ECCV, 2014.

[2] Kirillov, Alexander, et al. "Panoptic feature pyramid networks." In CVPR. 2019.