Dual Decision Improves Open-Set Panoptic Segmentation

Hai-Ming Xu[1], Hao Chen[2,*], Lingqiao Liu[1,*], Yufei Yin[3]
[1]The University of Adelaide [2]Zhejiang University
[3]University of Science and Technology of China

BMVC 2022

190

Background

- ➢ Panoptic segmentation in closed-set may not generalize to open-set cases.
- ➢ Open-set panoptic segmentation (OPS) is challenging:
  - The appearance of unknown class object is diverse and hardly to directly modelling unknown classes from the given training images.
  - The given "void" area in training images is too noisy to provide effective supervisions.

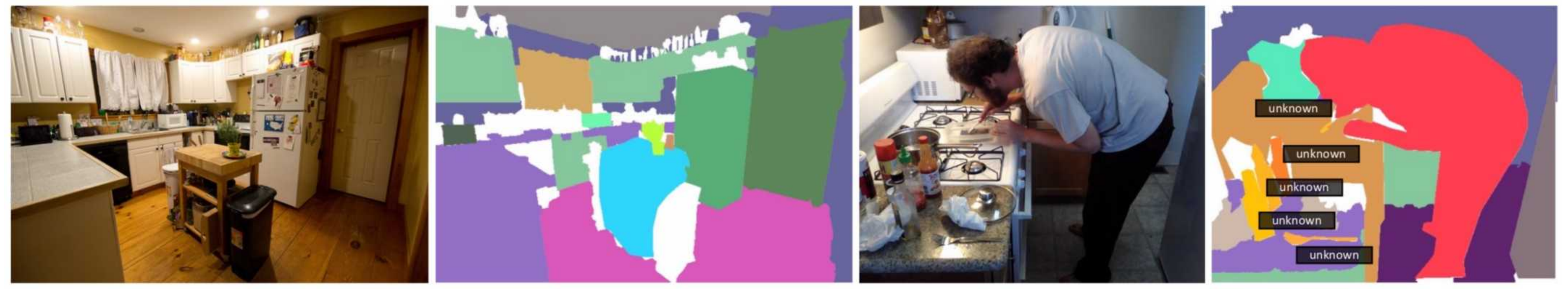


Image for Training | GT for Training | Image for Testing | GT for Testing

Comparisons of Existing OPS Methods

Model	PQ	SQ	RQ	Recall	Precision
Void-ignorance	3.7	71.8	5.2	11.0	3.4
Void-background	4.3	70.1	6.2	11.2	4.2
Void-suppression	7.2	75.3	9.5	27.6	5.8
Void-train	7.5	72.9	10.3	21.8	6.7
EOPSN	11.3	73.8	15.3	11.8	21.9

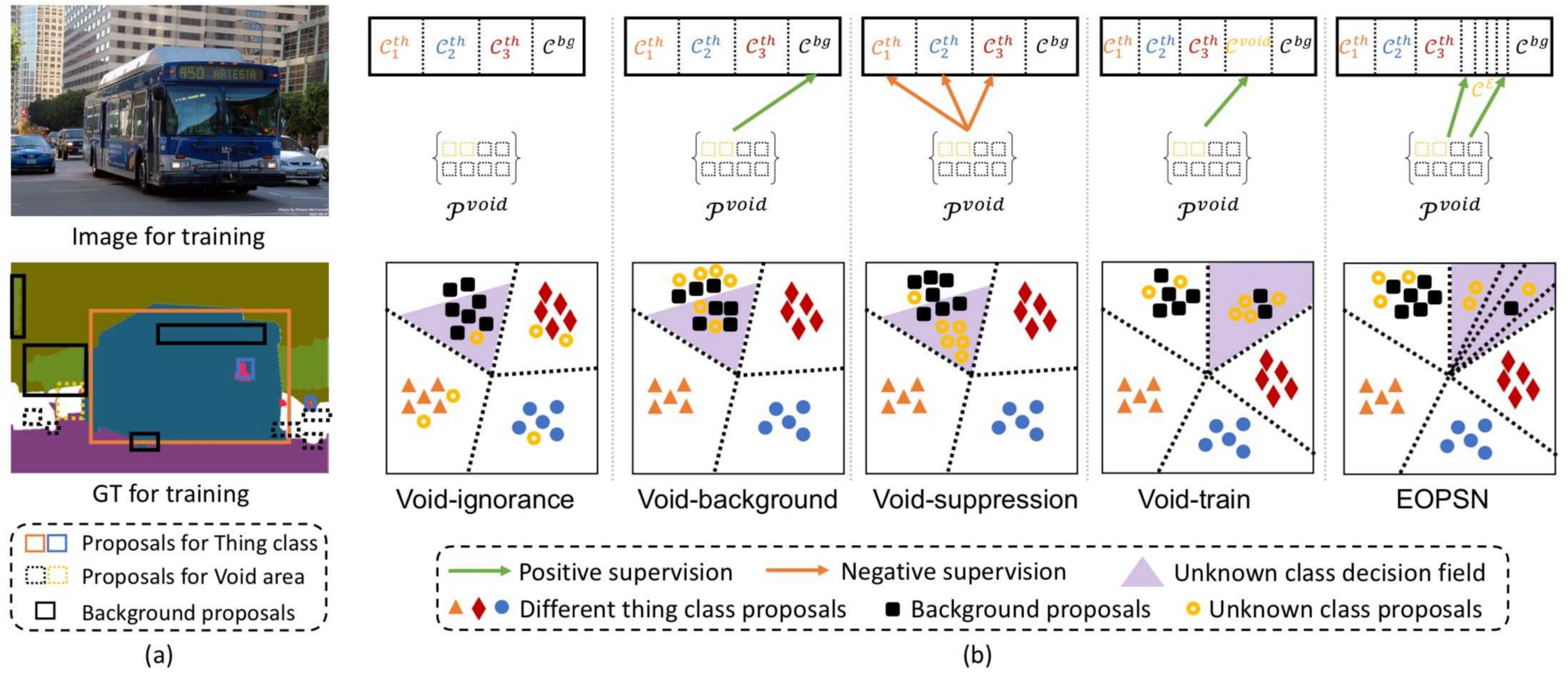
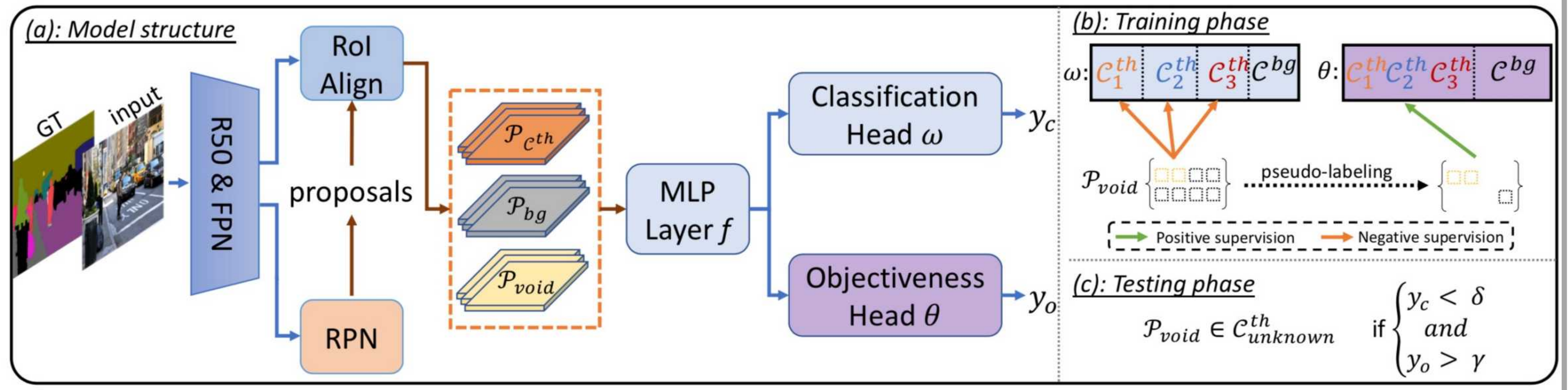


Figure 2: Illustration of existing OPS methods. (a) Proposal examples used in existing OPS methods. (b) Display of the usage of"void" class proposals in different OPS methods (top row) and how *unknown* class are predicted at testing phase (bottom row).

Our Approach



Objective functions:

➢ Classification head on known classes:

$$\min -\frac{1}{N_{\mathcal{P}_{void}}}\sum_{i\in\{\mathcal{C}^{th},bg\}}\sum_{k=1}^{N_{\mathcal{P}_i}}\log\frac{\exp\left(w_i^T f(\mathcal{P}_i^k)\right)}{\sum_{j\in\{\mathcal{C}^{Th},bg\}}\exp\left(w_j^T f(\mathcal{P}_i^k)\right)}$$

where w is the weight of classification head. $N_{\mathcal{P}_{void}}$ is the number of proposals except for those belonging to "void" class. $N_{\mathcal{P}_i}$ is the number of proposals in any specific *thing* class.

➢ Suppression of known class classifiers on "void" class proposals

$$\min -\frac{1}{N_{\mathcal{P}_{void}}}\sum_{i=1}^{N_{\mathcal{P}_{void}}}\sum_{k\in\mathcal{C}^{th}}\log\left(1-\frac{\exp\left(w_k^T f(\mathcal{P}^i_{void})\right)}{\sum_{\{\mathcal{C}^{th},bg\}}\exp\left(w_j^T f(\mathcal{P}^i_{void})\right)}\right)$$

where $N_{\mathcal{P}_{void}}$ means the number of "void" class proposals.

➢ Objectiveness head:

$$\min \frac{-1}{N_{\mathcal{P}_{void}}}\left(\sum_{i\in\mathcal{C}^{th}}\sum_{k=1}^{N_{\mathcal{P}_i}}\log\frac{\exp\left(\theta^T f(\mathcal{P}_i^k)\right)}{1+\exp\left(\theta^T f(\mathcal{P}_i^k)\right)}+\sum_{l=1}^{N_{\mathcal{P}_{bg}}}\log\frac{1}{1+\exp\left(\theta^T f(\mathcal{P}^l_{bg})\right)}\right)$$

where θ is the weight of objectiveness head and $N_{\mathcal{P}_{bg}}$ is the number of background proposals.

➢ Pseudo-labeling on objectiveness head:

$$\min \frac{-1}{N_{\mathcal{P}_{void}}}\sum_{i=1}^{N_{\mathcal{P}_{void}}}\mathbb{1}\left(\frac{\exp\left(\theta^T f(\mathcal{P}^i_{void})\right)}{1+\exp\left(\theta^T f(\mathcal{P}^i_{void})\right)}\geq\delta\right)\log\frac{\exp\left(\theta^T f(\mathcal{P}^i_{void})\right)}{1+\exp\left(\theta^T f(\mathcal{P}^i_{void})\right)}$$

where δ is the confidence threshold.

Experimental Results

K	Model	Known PQ	Known SQ	Known RQ	Known PQTh	Known SQTh	Known RQTh	Known PQSt	Known SQSt	Known RQSt	Unknown PQ	Unknown SQ	Unknown RQ	Unknown R	Unknown P
	Supervised	39.4	77.7	48.4	45.8	80.7	55.4	29.7	73.1	38.0	-	-	-	-	-
5	Void-supp.	38.0	77.0	46.7	44.8	80.6	54.1	28.3	71.7	36.1	6.7	76.2	8.8	39.9	4.9
	Void-train	37.3	76.7	45.9	43.6	80.4	52.8	28.2	71.5	36.0	8.6	72.7	11.8	29.8	7.3
	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	25.9	38.3
	Ours	38.1	77.7	46.6	45.1	80.9	54.3	28.1	73.1	35.7	**30.2**	**80.0**	**37.8**	32.8	44.5
10	Void-supp.	37.6	76.8	46.3	44.3	80.5	53.5	28.5	71.7	36.4	6.5	76.0	8.6	32.7	5.0
	Void-train	37.1	77.1	45.8	43.7	80.1	53.1	28.1	73.0	35.9	8.1	72.6	11.2	25.7	7.2
	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	19.0	30.2
	Ours	37.7	77.1	46.3	45.0	80.7	54.3	27.8	72.2	35.4	**24.5**	**79.9**	**30.7**	24.7	40.6
20	Void-supp.	37.5	75.9	46.1	45.1	80.6	54.5	28.2	70.2	36.1	7.2	75.3	9.5	27.6	5.8
	Void-train	36.8	76.3	45.4	44.1	80.1	53.5	27.9	71.6	35.6	7.5	72.9	10.3	21.8	6.7
	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	11.8	21.9
	Ours	37.1	75.8	45.7	45.0	80.6	54.3	27.6	70.1	35.3	**21.4**	**79.1**	**27.1**	21.9	35.4

Table 2: Comparisons of open-set panoptic segmentation performance against the state-of-the-art methods on MS-COCO *val* set with three *known-unknown* splits $K(\%)$ which denotes the ratio of *unknown* classes to all classes. Recall (R) and precision (P) of *unknown* objects are also presented for reference. The best results on *unknown* classes are bold highlighted.

Model	Known PQ	Known SQ	Known RQ	Known PQTh	Known SQTh	Known RQTh	Known PQSt	Known SQSt	Known RQSt	Unknown PQ	Unknown SQ	Unknown RQ	Unseen PQ	Unseen SQ	Unseen RQ
Void-supp.	35.8	76.7	44.5	43.0	81.2	52.5	27.7	71.6	35.4	7.6	75.5	10.1	4.5	75.9	6.0
Void-train	35.4	77.2	43.9	42.2	81.0	51.6	27.7	72.8	35.3	8.8	73.8	15.7	4.4	74.8	5.9
EOPSN	35.7	76.6	44.7	43.2	81.1	52.7	27.8	71.4	35.6	23.0	74.6	30.8	0.4	80.3	0.5
Ours	35.8	76.6	44.5	43.0	81.1	52.5	27.6	71.4	35.3	**30.2**	**80.2**	**37.7**	**9.3**	**82.5**	**11.2**

Table 3: Comparisons of OPS performance on MS-COCO *val* set under the newly proposed *zero-shot* setting. The best results on *unknown* and *unseen* classes are bold highlighted.

Class of prominent object:

- ➢ *Unknown*: cow
- ➢ *Unseen*: fire hydrant
- ➢ *Unseen*: refrigerator
- ➢ *Unseen*: parking meter