Appendix of "Dual Decision Improves Open-Set Panoptic Segmentation"

Hai-Ming Xu[1]
hai-ming.xu@adelaide.edu.au
Hao Chen[2]
stanzju@gmail.com
Lingqiao Liu[1]
lingqiao.liu@adelaide.edu.au
Yufei Yin[3]
yinyufei@mail.ustc.edu.cn

[1] The University of Adelaide, Adelaide, Australia
[2] Zhejiang University Zhejiang, China
[3] University of Science and Technology of China Anhui, China

In this appendix, we first present detail formulations of the panoptic segmentation evaluation metrics. We then list details about the dataset splits used in our paper. Finally, we present more experimental results on *zero-shot* setting with $K = 0\%$.

1 Formulation Details of Evaluation Metric

Three kinds of panoptic segmentation metrics are normally considered in the literature, i.e., panoptic quality (PQ), segmentation quality (SQ) and the recognition quality (RQ)

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

where $\text{IoU}(p,g)$ means intersection over union of predicted p and ground truth g segments. *TP*/*FP*/*FN* respectively denote the set of true positives, false positives and false negatives. For the open-set panoptic segmentation task, we use these three metrics to evaluate the performance of *known*, *unknown* and *unseen* classes.

2 Details of Dataset Splits

In our paper, we have conducted experiments on two kinds of open-set panoptic segmentation settings, i.e., *known-unknown* setting and *zero-shot* setting. For the *known-unknown* setting, three kinds of splits are evaluated following the previous work EOPSN [1], i.e., K% of classes are removed from the 80 *thing* classes of MS-COCO 2017 as *unknown* classes and $K = \{5, 10, 20\}$. We list the *unknown* classes as follows (the classes are removed cumulatively for these three settings respectively)

- car, cow, pizza, toilet

- boat, tie, zebra, stop sign
- dining table, banana, bicycle, cake, sink, cat, keyboard, bear

For the more realistic *zero-shot* setting, we build it upon the 5% split *known-unknown* setting mentioned above and further removes training images that contains instances belonging to the 20% tail *thing* class of MS-COCO. The removed tail classes has already been presented in the main paper.

3 Relationship between our method and K selection

Both of the classification and objectiveness heads are designed to be optimized on known class and background proposals in our work, and thus the effectiveness of our work does not depend on the setting of K. As results are shown in the following table, ours w/o PL performs much better than EOPSN on $K = 0\%$. Furthermore, our work can make use of the area originally labeled as void for better unseen recognition.

$K = 0\%$, unseen class performance	PQ	SQ	RQ
EOPSN	0.2	77.5	0.3
Ours w/o PL	5.0	78.3	6.4
Ours	**5.8**	**80.7**	**7.2**

References

[1] Jaedong Hwang, Seoung Wug Oh, Joon-Young Lee, and Bohyung Han. Exemplar-based open-set panoptic segmentation network. In *CVPR*, 2021.