Unlocking Constraints: Source-Free Occlusion-Aware Seamless Segmentation

(Supplementary Material)

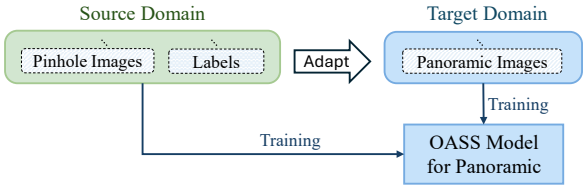
Yihong Cao[1,2,*] Jiaming Zhang[3,4,*] Xu Zheng[5,6] Hao Shi[7] Kunyu Peng[3] Hang Liu[1]
Kailun Yang[1,†] Hui Zhang[1,†]
[1]Hunan University [2]Hunan Normal University [3]Karlsruhe Institute of Technology [4]ETH Zürich
[5]HKUST(GZ) [6]INSAIT, Sofia University "St. Kliment Ohridski" [7]Zhejiang University

1. SFOASS *vs* OASS

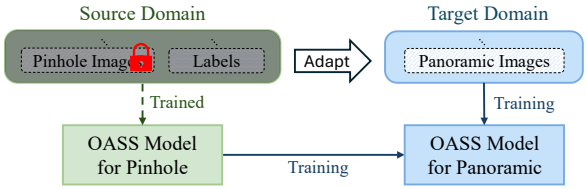
Occlusion-Aware Seamless Segmentation (OASS), introduced by [1], aims to address the issues of the narrow field of view, occlusion of perspective, and domain gaps in a seamless manner. As depicted in Fig. 1a, OASS requires the model to transfer from a labeled pinhole domain to the unlabeled panoramic domain, ultimately resulting in an OASS model that performs effectively within the panoramic domain, thereby enabling 360° panoramic perception. In addition, this task introduces amodal-level prediction, which means that the complete shape of the object needs to be segmented regardless of whether the object is occluded or not, in order to solve the problem of perspective occlusion. The OASS task encompasses five distinct segmentation tasks at once: semantic segmentation, instance segmentation, amodal instance segmentation, panoptic segmentation, and amodal panoptic segmentation.

Typically, the OASS model consists of a shared encoder and three branches: a semantic branch, an instance branch, and an amodal instance branch. As illustrated in Fig. 2, during the training phase, the outputs of each branch are supervised using the corresponding labels. During inference, in order to obtain all five segmentation maps (amodal panoptic, panoptic, semantic, amodal instance, and instance segmentation) in a single pass, the Occlusion-Aware Fusion module [1] combines the pixel-level class predictions from the semantic branch with the instance-level object predictions from the two instance-level branches, generating the final segmentation maps.

However, as a task based on Unsupervised Domain Adaptation (UDA), OASS requires simultaneous access to the data of both the source domain and the target domain during the adaptation process. This presents challenges in scenarios with data privacy and commercial restrictions,



(a) Occlusion-Aware Seamless Segmentation (OASS).



(b) Source-Free Occlusion-Aware Seamless Segmentation (SFOASS).

Figure 1. Comparison of Task Settings.

where sometimes the data of the source domain is usually prohibited from being accessed. To address these limitations, we further introduce Source-Free Occlusion-Aware Seamless Segmentation (SFOASS), as depicted in Fig. 1b. This more rigorous task extends the seamless segmentation capabilities of OASS while imposing an additional restriction: the source domain images and labels are inaccessible during domain adaptation. SFOASS relies solely on a pre-trained OASS model from the pinhole domain and unlabeled panoramic images from the target domain. The absence of source domain data during adaptation presents unique challenges, including the lack of explicit domain alignment and the inability to directly address domain-specific biases between the pre-trained source model and unlabeled target images. As such, SFOASS requires the development of innovative strategies to overcome these hurdles and ensure effective adaptation to the target domain.

To address these challenges, we propose UNconstrained Learning Omni-Context Knowledge (UNLOCK) frame-

*Equal contribution. †Corresponding authors (e-mail: kailun.yang@hnu.edu.cn, zhanghuihby@126.com).

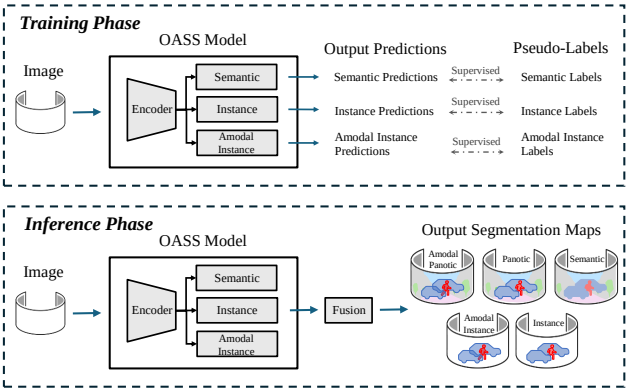


Figure 2. Workflow of the training and inference phases of the OASS model.

work, a novel framework that introduces Omni Pseudo-Labeling Learning (OPLL) to leverage the knowledge embedded in the pre-trained OASS model. At the same time, it incorporates an Amodal-Driven Contextual Learning mechanism (ADCL) to capture the intrinsic knowledge of the target domain, ensuring effective occlusion-aware segmentation performance in the target panoramic domain.

2. Benchmarks

In this work, we applied two benchmarks to evaluate the SFOASS task. These benchmarks are based on three datasets: KITTI360-APS [3], AmodalSynthDrive [5], and BlendPASS [1]. In accordance with the SFOASS task formulation, both KITTI360-APS and AmodalSynthDrive serve as source domains, with KITTI360-APS consisting of real-world pinhole data and AmodalSynthDrive consisting of synthetic pinhole data. BlendPASS serves as the target domain, composed of real-world panoramic data. Consequently, two **Pinhole-to-Panoramic** benchmarks are used: KITTI360-APS→BlendPASS and AmodalSynthDrive→BlendPASS. The former constitutes a Real-to-Real adaptation scenario, while the latter forms a Synthetic-to-Real adaptation scenario.

For the KITTI360-APS→BlendPASS benchmark, the annotated classes between the source and target domains are not fully aligned. As detailed in [1], the annotations of the BlendPASS dataset were pre-processed to match the source domain, KITTI360-APS, resulting in 18 categories: 11 *Stuff* classes (*road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, and sky*) and 7 *Thing* classes (*car, pedestrians, cyclists, two-wheelers, van, truck, and other vehicles*).

For the newly introduced AmodalSynthDrive→ BlendPASS benchmark, the class annotations are consistent with those of Cityscapes [2], with the exception that the AmodalSynthDrive dataset does not include the *Bus* class, which is present in BlendPASS. To address this discrepancy, we removed the *Bus* class from the annotations of BlendPASS, resulting in 18 categories: 11 *Stuff* classes (*road, sidewalk, building, wall, fence, pole, lights, sign, vegetation, terrain, and sky*) and 7 *Thing* classes (*person, rider, car, truck, bus, motorcycle, and bicycle*).

3. More Experiments Details

In the Amodal-Driven Contextual Learning (ADCL) of UNLOCK, only object samples with an overlapping area of the object region that is less than half of the full area of the object are stored in the amodal-driven object pool. This criterion ensures that excessive contextual information about the objects is not discarded under the spatial-aware mixing strategy. During the object mixing process, the occlusion order of pasted object samples is determined based on the order in which they are applied. Specifically, the object pasted later is placed in front of all previously pasted objects in the mixed image. If an object is fully occluded by other pasted objects, it is removed, as this scenario does not require segmentation in the SFOASS task. In this situation, fully occluded objects are outside the scope of the task, as they cannot be detected or even confirmed to exist, making their segmentation unnecessary for the objectives of SFOASS. For the final mixed omni pseudo-labels, semantic pseudo-labels are directly replaced by the semantic labels of the pasted objects. For instance pseudo-labels and amodal instance pseudo-labels, if the objects originally present in the current image (before any pasting) are completely occluded by the pasted objects, the corresponding instance-level pseudo-labels of those objects are removed (since these objects are entirely occluded in the mixed image after pasting). These operations are crucial to ensuring that the images mixed by the ADCL strategy retain reasonable contextual information. By carefully managing the occlusion and placement of objects, the proposed ADCL approach helps the model learn diverse contextual information, thereby facilitating adaptation to the unlabeled panoramic data of the target domain. Both OPLL and ADCL operate as data-level preprocessing strategies. OPLL selects informative pseudo-labeled samples based on object-level predictions, while ADCL manipulates object placement and occlusion relationships through spatial-aware mixing. These operations are designed to construct more effective training data and guide the model to adapt better to the unlabeled target domain without modifying the model architecture itself.

For the reproduction of the 360SFUDA++[9], we replicated the key innovations and extended the prototype-based approach, initially applied to the semantic segmentation, to the two instance-level branches commonly found in the OASS model, ensuring that both instance branches benefit from this technique. To ensure the robustness and reliability of our reproduction, we strictly followed the original ex-

perimental setup. Through multiple rounds of experimentation and fine-tuning, we determined the optimal configuration, achieving results that closely mirror those of the original work. For the retraining of existing UDA methods [1, 4, 6, 8] on the AmodalSynthDrive→ BlendPASS benchmark, we utilize the training protocols from [1] and adhere to the hyperparameters specified in the respective papers of each method. This approach ensures consistency with the original experimental setups, allowing for a fair comparison across the different methods.

4. Analysis for Hyper-parameters

In UNLOCK, OPLL serves as the key, and ADCL performs the unlocking action. We specifically analyze the effect of the hyperparameters in both OPLL and ADCL on the KITTI360-APS→BlendPASS benchmark. For the SFOASS task, performance is evaluated using five metrics: mAPQ for amodal panoramic segmentation, mPQ for panoramic segmentation, mIoU for semantic segmentation, mAAP for amodal instance segmentation, and mAP for instance segmentation. Since these metrics do not always vary consistently, we prioritize mAPQ, which evaluates amodal panoramic segmentation (including both *Stuff* and amodal-level *Thing* classes), while also taking into account the overall performance across all metrics. Furthermore, although the three branches are designed to be independent, they share the same encoder for feature extraction. As a result, adjustments made to the hyperparameters of a single branch will indirectly affect the other branches, thereby influencing the performance across all metrics.

τ^{fix}	τ^{per}	mAPQ	mPQ	mIoU	mAAP	mAP
0.1	0.9	24.37	23.78	38.84	10.41	10.43
0.2	0.7	24.45	23.76	39.10	10.45	10.43
0.3	0.5	**24.71**	**24.00**	39.03	**10.52**	10.52
0.4	0.3	24.58	23.93	**39.13**	10.42	10.49
0.5	0.2	24.28	23.78	38.92	10.47	**10.53**

Table 1. Performance analysis of using different values for τ^{fix} and τ^{per} in omni **amodal instance** pseudo-labels.

τ^{fix}	mAPQ	mPQ	mIoU	mAAP	mAP
0.1	24.52	23.88	38.95	**10.53**	10.45
0.2	24.34	23.74	38.77	10.31	10.51
0.3	**24.71**	**24.00**	39.03	10.52	**10.52**
0.4	24.44	23.81	**39.14**	10.39	10.35
0.5	24.31	23.70	38.93	10.50	10.38

Table 2. Performance analysis of using different values for τ^{fix} in omni **amodal instance** pseudo-labels.

OPLL. We systematically investigated the effects of these two parameters τ^{fix} and τ^{per} on the omni pseudo-labels from the amodal instance branch, the instance branch, and the semantic branch. Each experiment focuses solely on OPLL to clearly evaluate its individual contribution. As shown in Tables 1, 4, and 7, only the two thresholds associated with the branch under investigation were varied, while the thresholds for the other branches were held con-

τ^{per}	mAPQ	mPQ	mIoU	mAAP	mAP
0.9	24.54	23.79	38.71	**10.52**	10.36
0.7	24.43	23.79	38.80	10.46	**10.60**
0.5	**24.71**	**24.00**	**39.03**	**10.52**	10.52
0.3	24.67	23.71	38.86	10.49	10.57
0.2	24.56	23.90	39.02	10.47	10.35

Table 3. Performance analysis of using different values for τ^{per} in omni **amodal instance** pseudo-labels.

τ^{fix}	τ^{per}	mAPQ	mPQ	mIoU	mAAP	mAP
0.3	0.6	24.37	23.66	38.81	10.32	10.32
0.4	0.4	24.53	23.93	38.95	10.45	10.51
0.5	0.2	**24.71**	**24.00**	39.03	**10.52**	**10.52**
0.6	0.1	24.41	23.74	**39.12**	10.42	10.39
0.7	0.05	24.20	23.61	38.80	10.45	10.37

Table 4. Performance analysis of using different values for τ^{fix} and τ^{per} in omni **instance** pseudo-labels.

τ^{fix}	mAPQ	mPQ	mIoU	mAAP	mAP
0.3	24.06	23.50	38.60	10.28	10.39
0.4	24.20	23.52	38.65	10.22	10.43
0.5	**24.71**	**24.00**	39.03	**10.52**	**10.52**
0.6	24.54	23.89	**39.20**	10.46	10.37
0.7	24.12	23.47	38.69	10.21	10.44

Table 5. Performance analysis of using different values for τ^{fix} in omni **instance** pseudo-labels.

τ^{per}	mAPQ	mPQ	mIoU	mAAP	mAP
0.6	24.15	23.51	38.71	10.35	10.46
0.4	24.46	23.60	38.97	10.27	10.31
0.2	**24.71**	**24.00**	**39.03**	**10.52**	**10.52**
0.1	24.36	23.79	38.91	10.49	10.41
0.05	24.27	23.59	38.87	10.25	10.43

Table 6. Performance analysis of using different values for τ^{per} in omni **instance** pseudo-labels.

τ^{fix}	τ^{per}	mAPQ	mPQ	mIoU	mAAP	mAP
0.3	0.9	24.53	23.96	39.01	10.46	10.37
0.4	0.85	24.31	23.73	39.02	10.41	10.39
0.5	0.8	**24.71**	**24.00**	39.03	**10.52**	10.52
0.6	0.7	24.33	23.56	**39.05**	10.25	**10.58**
0.7	0.6	24.34	23.52	38.95	10.28	10.47

Table 7. Performance analysis of using different values for τ^{fix} and τ^{per} in omni **semantic** pseudo-labels.

τ^{fix}	mAPQ	mPQ	mIoU	mAAP	mAP
0.3	24.43	23.62	39.01	10.29	10.48
0.4	24.52	23.87	38.99	10.47	10.37
0.5	**24.71**	**24.00**	39.03	**10.52**	10.52
0.6	24.46	23.81	**39.05**	10.44	**10.63**
0.7	24.62	23.94	38.97	10.48	10.38

Table 8. Performance analysis of using different values for τ^{fix} in omni **semantic** pseudo-labels.

τ^{per}	mAPQ	mPQ	mIoU	mAAP	mAP
0.9	24.30	23.67	38.88	10.37	10.38
0.85	24.40	23.72	38.99	10.42	**10.59**
0.8	**24.71**	**24.00**	**39.03**	**10.52**	10.52
0.7	24.44	23.79	39.02	10.42	10.37
0.6	24.29	23.7	39.01	10.41	10.37

Table 9. Performance analysis of using different values for τ^{per} in omni **semantic** pseudo-labels.

stant at their final adopted values, as indicated by the gray background in the tables. Additionally, another set of experiments was conducted to examine the effect of varying individual thresholds, where one threshold was altered at a time while the others remained fixed, as shown in Tables 2, 3, 5, 6, 8, and 9. Overall, the performance results across the three tables demonstrate that our method exhibits low sensitivity to these hyperparameters. This robustness can be attributed to two aspects: (1) the omni pseudo-labels, which incorporate knowledge from all predictions while excluding low-quality predictions in the optimization process, and (2) the class-wise self-adjusting threshold mechanism, which dynamically maintains a balance between the number and accuracy of the generated labels for each class. For the amodal instance shown in Table 1, we observed that different values have a relatively minor impact on the mAP and mAAP of the instance-level branch. This is because the amodal instance branch focuses more on the true shape of the object and is less dependent on its appearance features. As shown in Tables 2 and 3, the increase in sample numbers due to lower thresholds or higher percentages improves the mAAP or mAP values. However, this increase in samples also slightly affects the accuracy of the *stuff* category, leading to a decrease in other metrics. For instance pseudo-labels, we observed that excessively low thresholds (as shown in the first row of Table 4 and 5) negatively affect the mAP and mAAP metrics. Lower thresholds lead to the inclusion of incorrect segmentation of the unoccluded regions of objects, which impairs the model's ability to correctly understand the scene. For the semantic pseudo-labels, a higher threshold (as shown in the last row of Table 7) can improve the accuracy of semantic labels. However, this also filters out more low-confidence pixels, which reduces the model's ability to capture rich contextual semantic information, ultimately leading to decreased performance. As shown in Table 9, although a higher percentage introduces more pixels, the noise labels it brings also lead to a decrease in performance.

τ'^{fix}	τ'^{per}	mAPQ	mPQ	mIoU	mAAP	mAP
0.99	0.05	24.80	24.00	38.57	10.47	10.68
0.95	0.10	24.95	24.05	39.62	10.84	11.25
0.90	0.15	**25.84**	**24.55**	**40.31**	**10.93**	**11.47**
0.85	0.20	25.74	**24.55**	39.27	10.81	11.20
0.80	0.30	25.39	24.39	39.23	10.62	10.91

Table 10. Performance analysis of varying thresholds τ'^{fix}, τ'^{per} for amodal-driven object pool of ADCL.

R	mAPQ	mPQ	mIoU	mAAP	mAP
5	24.83	24.02	39.19	10.44	10.78
8	25.34	24.42	39.24	10.51	10.67
10	**25.84**	**24.55**	40.29	**10.93**	**11.47**
12	25.34	24.49	**40.30**	10.85	11.25
15	25.40	24.47	39.05	10.70	10.76

Table 11. Performance analysis of using different numbers R of object samples in ADCL.

ADCL. We further analyzed the parameters τ'^{fix} and τ'^{per}, which control the quality of generated objects in the amodal-driven object pool, as well as the parameter R, which determines the number of objects pasted in the spatial-aware mixing strategy. These experiments are conducted independently of OPLL to isolate the effect of ADCL. As shown in Table 10, excessively low thresholds lead to a decline in object quality, resulting in the inclusion of background information or incomplete object shapes. Conversely, high thresholds improve the quality of individual objects but reduce the diversity of objects in the pool. For the number R of pasted objects, as shown in Table 11, too few objects limit the variety available for model training, while an overly large number results in overcrowding, which may obscure important contextual information from the *Stuff* class, thereby hindering the model's ability to capture the broader scene context.

5. More Visualization Results

5.1. More Visualization Results on SFOASS

As shown in Fig. 3, we further compare our proposed UNLOCK with existing UDA methods [1, 4, 6, 8] on KITTI360-APS→BlendPASS benchmark. The results demonstrate that, even without access to source domain images and labels, UNLOCK achieves comparable, or even surpassing, UDA methods. For example, in the middle column of Fig. 3, UNLOCK detects more cars accurately compared to UnmaskFormer [1]. In the right column, UNLOCK

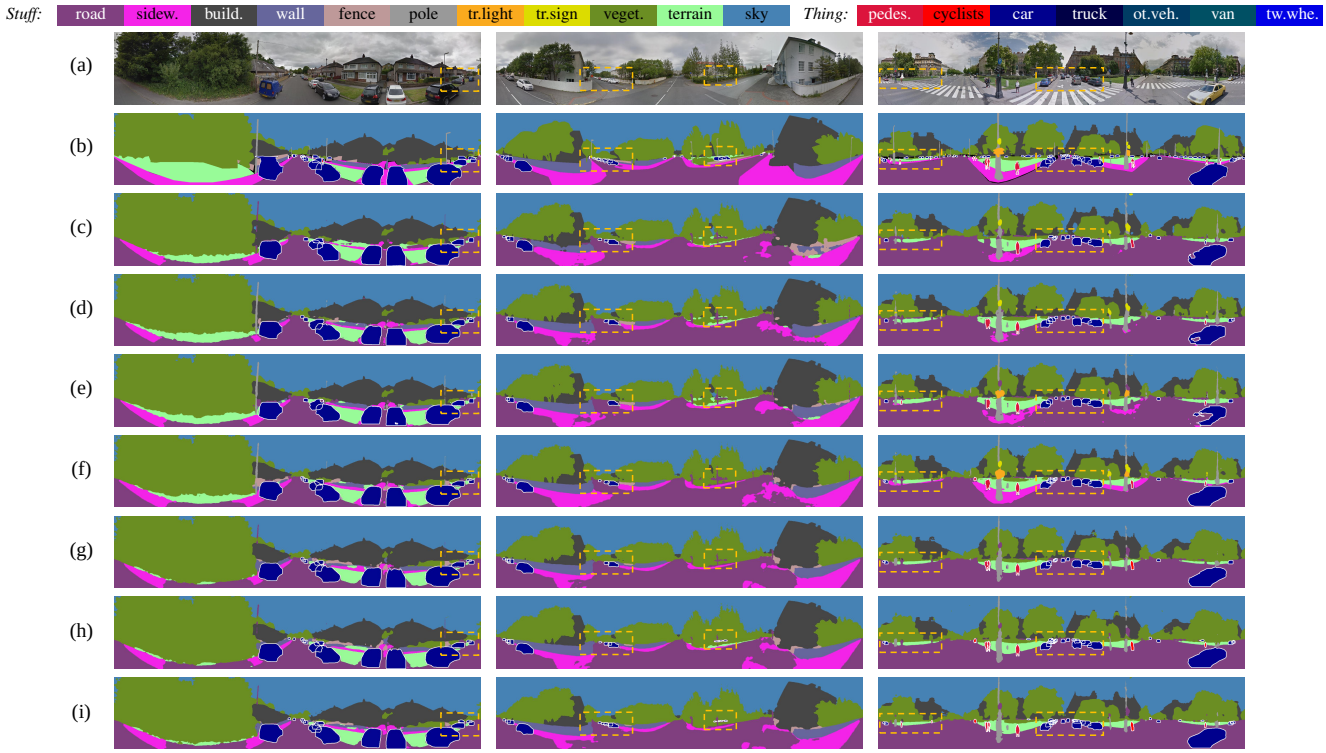


Figure 3. **Visualization for Amodal Panoptic Segmentation on KITTI360-APS→BlendPASS benchmark.** From top to bottom are (a) Image, (b) GT, (c) DATR [8], (d) Trans4PASS [6], (e) EDAPS [4], (f) UnmaskFormer [1], (g) Source-only, (h) 360SFUDA++ [9], and (i) UNLOCK (Ours).

excels in identifying and segmenting pedestrians, a result attributed to the designed ADCL method, which facilitates the learning of diverse object samples of the *Thing* class. We also provided the visualization results for amodal panoptic segmentation on AmodalSynthDrive→BlendPASS benchmark, as shown in Fig. 4. UNLOCK demonstrates exceptional segmentation performance, particularly when dealing with a high density of *Car* class. Our method identifies more cars with accurate shapes compared to other approaches. This successful adaptation from virtual to real environments not only demonstrates the robustness and generalization of our method but also provides a viable path for applying these advancements in real-world scenarios.

We also provide the visualization results of semantic segmentation and panoptic segmentation, as shown in Fig. 5. In the left part of Fig. 5, UNLOCK demonstrates superior performance in handling the *Stuff* class. For example, while many UDA methods misidentify elements like billboards on buildings as *Traffic Sign* class, UNLOCK correctly distinguishes these as *Building* class. In addition, as shown in the right part of Fig. 5, UNLOCK excels in the panoptic segmentation task, detecting more objects and providing more accurate predictions for each object's visible region compared to other methods (as seen in the third column).

5.2. Failure Case of UNLOCK

Figure 6 illustrates a failure case where vehicles are occluded by sparse fences, leading UNLOCK to miss their presence. Although the vehicles remain partially visible through the gaps, the repetitive vertical patterns of the fence interfere with the visual cues, making it challenging for the model to distinguish the objects from the background. This highlights a limitation of UNLOCK in dealing with structured, non-dense occlusions that disrupt spatial continuity and confuse contextual reasoning.

5.3. Qualitative analysis of ADCL

To ensure the reliability of the amodal-driven object pool, we apply stricter adaptive thresholds to filter out high-quality object samples. As shown in Fig.7, we visualize examples from two training images, with the selected high-quality amodal instances highlighted. For instance, the pedestrian in the first column of Fig.7 shows accurate segmentation with clear boundaries, while the car in the second column retains a reasonable and complete amodal shape despite being partially occluded by surrounding objects. These results demonstrate that the filtered object pool provides reliable samples, which in turn benefit the spatial-aware mixing strategy in ADCL.

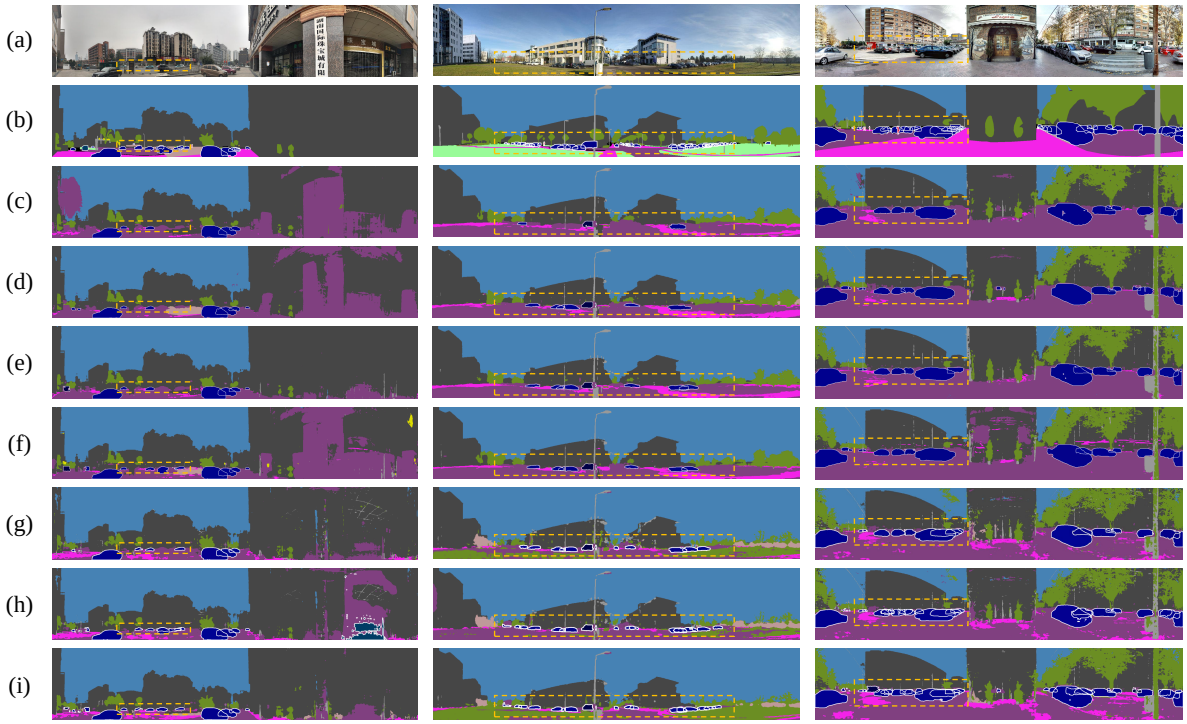


Figure 4. **Visualization for Amodal Panoptic Segmentation on AmodalSynthDrive→BlendPASS benchmark.** From top to bottom are (a) Image, (b) GT, (c) DATR [8], (d) Trans4PASS [6], (e) EDAPS [4], (f) UnmaskFormer [1], (g) Source-only, (h) 360SFUDA++ [9], and (i) UNLOCK (Ours).

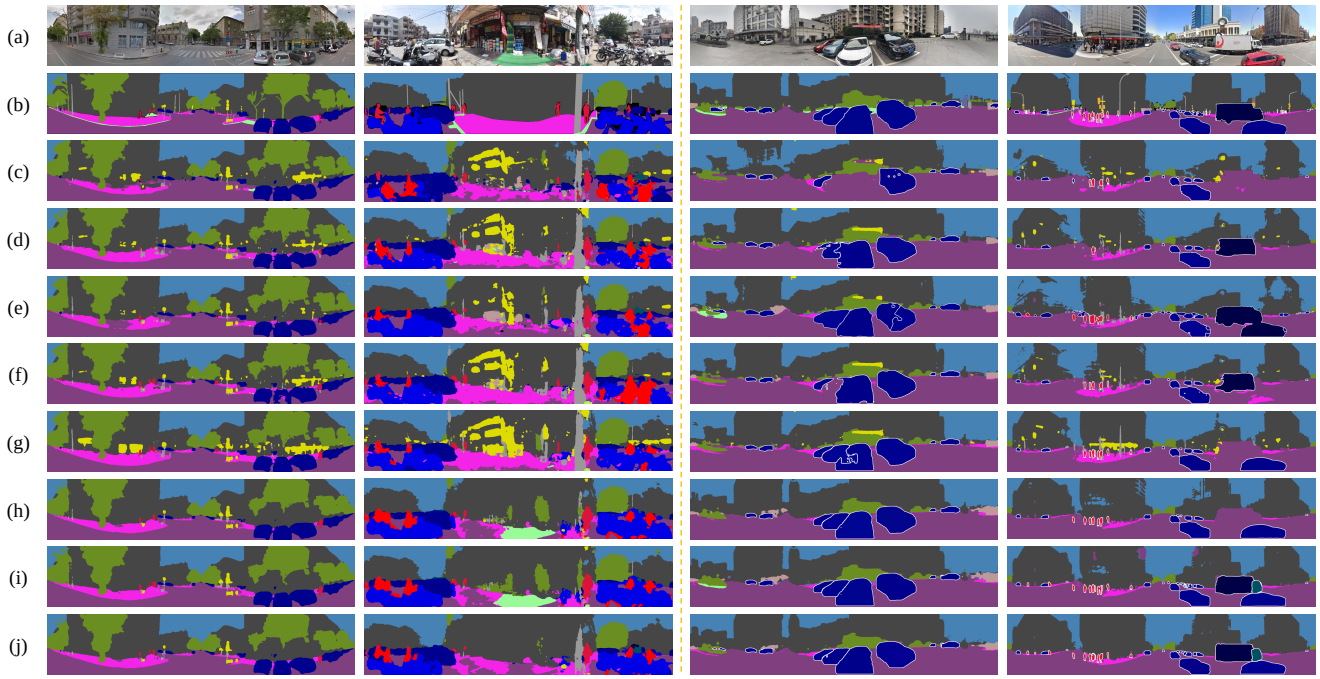


Figure 5. **Visualization for Semantic (Left) and Panoptic (Right) Segmentation on KITTI360-APS→BlendPASS benchmark.** From top to bottom are (a) Image, (b) GT, (c) DATR [8], (d) Trans4PASS [6], (e) UniDAPS [7], (f) EDAPS [4], (g) UnmaskFormer [1], (h) Source-only, (i) 360SFUDA++ [9], and (j) UNLOCK (Ours).

Figure 6. Example of a **failure case** where UNLOCK misses vehicles partially occluded by sparse fences.

Figure 7. Example of a **failure case** where UNLOCK misses vehicles partially occluded by sparse fences.

6. Discussion

6.1. Limitations and potential solutions

The 360° boundaries of panoramic images remain an underexplored aspect in the field of seamless segmentation, posing challenges for accurate perception at extreme viewing angles. Additionally, the interaction between the instance and amodal instance segmentation branches requires further investigation to fully leverage their complementary information.

Future research could explore strategies to mitigate label scarcity in the panoramic domain by incorporating semi-supervised learning techniques into amodal instance segmentation. Moreover, enhancing model robustness and generalizability across diverse panoramic environments could be achieved through domain generalization approaches, enabling improved adaptation to unseen real-world conditions.

In addition, the scalability of the proposed framework with respect to the number of object categories and the complexity of occlusion relationships remains a potential limitation. Since panoramic scenes often contain numerous instances with intricate occlusions, future work should further evaluate and improve the framework's performance in such large-scale, dense scenarios.

6.2. Societal Impacts

On the positive side, SFOASS enhances privacy handling by eliminating the need for access to source-domain images and labels during target-domain adaptation, safeguarding sensitive data. It also facilitates commercial deployment by mitigating data ownership restrictions. Additionally, by relying on a pre-trained model and unlabeled target-domain data, SFOASS helps avoid storage issues related to large training datasets, leading to more efficient resource utilization and supporting sustainable technological development.

However, challenges remain in handling heavily occluded objects and domain gaps, which may lead to misclassifications or biased predictions. In safety-critical applications like autonomous vehicles, such errors could result in accidents. Furthermore, the reliance on pre-trained models without access to source data can limit the adaptability of the system in unfamiliar environments, particularly when the target domain differs significantly from the conditions seen during training. This could lead to reduced robustness and performance in real-world scenarios, highlighting the need for ongoing validation and refinement.

Therefore, while SFOASS offers promising advancements, its deployment must be carefully managed to mitigate risks and ensure its robustness in real-world scenarios.

References

[1] Yihong Cao, Jiaming Zhang, Hao Shi, Kunyu Peng, Yuhongxuan Zhang, Hui Zhang, Rainer Stiefelhagen, and Kailun Yang. Occlusion-aware seamless segmentation. In *ECCV*, 2024. 1, 2, 3, 4, 5, 6

[2] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *CVPR*, 2016. 2

[3] Rohit Mohan and Abhinav Valada. Amodal panoptic segmentation. In *CVPR*, 2022. 2

[4] Suman Saha, Lukas Hoyer, Anton Obukhov, Dengxin Dai, and Luc Van Gool. EDAPS: Enhanced domain-adaptive panoptic segmentation. In *ICCV*, 2023. 3, 4, 5, 6

[5] Ahmed Rida Sekkat, Rohit Mohan, Oliver Sawade, Elmar Matthes, and Abhinav Valada. AmodalSynthDrive: A synthetic amodal perception dataset for autonomous driving. *IEEE Robotics and Automation Letters*, 2024. 2

[6] Jiaming Zhang, Kailun Yang, Chaoxiang Ma, Simon Reiß, Kunyu Peng, and Rainer Stiefelhagen. Bending reality: Distortion-aware transformers for adapting to panoramic semantic segmentation. In *CVPR*, 2022. 3, 4, 5, 6

[7] Jingyi Zhang, Jiaxing Huang, and Shijian Lu. Hierarchical mask calibration for unified domain adaptive panoptic segmentation. In *CVPR*, 2023. 6

[8] Xu Zheng, Tianbo Pan, Yunhao Luo, and Lin Wang. Look at the neighbor: Distortion-aware unsupervised domain adapta-

tion for panoramic semantic segmentation. In *ICCV*, 2023. 3, 4, 5, 6

[9] Xu Zheng, Pengyuan Zhou, Athanasios V. Vasilakos, and Lin Wang. 360SFUDA++: Towards source-free UDA for panoramic segmentation by learning reliable category prototypes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024. 2, 5, 6