

Prompt-driven Small Object Instance Segmentation in Earth Observation

Supplementary Material

Appendix A: Details of the ReSOS Dataset

A.1 Dataset Statistics

Table 7 summarizes the per-category statistics of the ReSOS dataset. The dataset comprises over 751k annotated instances across eight object categories, with a strong emphasis on small objects. We report both the mean instance area (in pixel squared) and the mean absolute size, defined as the larger dimension (height or width) of the instance’s bounding box. Across categories, the mean area remains below 800 px², confirming the small-object nature of the dataset. Categories such as *Swimming Pool*, *Large Vehicle*, and *Container* have relatively larger average sizes, likely due to their physical scale and less crowded spatial distributions. In contrast, *Small Vehicle* instances are both the most abundant and the smallest in absolute scale (mean max dimension of 17.3 pixels), posing significant challenges for instance segmentation models.

Table 7. Per-category statistics of the ReSOS dataset. The *Mean Absolute Size* is the average of max(height, width) per instance.

Category	Instance Number	Mean Area (px ²)	Mean Absolute Size (px)
Small Vehicle	458,105	143.8	17.3
Container	121,895	453.9	41.9
Ship	60,685	277.4	25.9
Storage Tank	35,163	294.6	18.3
Swimming Pool	29,500	729.0	37.5
Airplane	28,311	278.5	27.7
Large Vehicle	15,947	560.0	46.4
Helicopter	1,421	409.4	33.2

A.2 Visualization Examples

Fig 6 show visualization examples from the ReSOS dataset across eight object categories. Each row represents a different object class, with multiple sample patches showing various instances within image crops. The colored masks highlight the precise boundaries of each object instance, demonstrating the dataset’s annotation quality and the diversity of object appearances across different scenarios.

Appendix B: Visualization of Instance Segmentation Results on ReSOS

Fig 7 presents comparative instance segmentation results on the ReSOS dataset. The visualization compares per-

formance across different methods including SparseInst, Mask2Former, MaskDINO, and ground truth annotations, which generated by our SOPSeg. The colored segmentation masks reveal how each method performs on small object instance segmentation. This comparison highlights the varying capabilities of existing methods when handling dense small objects in complex remote sensing environments, particularly for challenging categories like vehicles and aircraft.

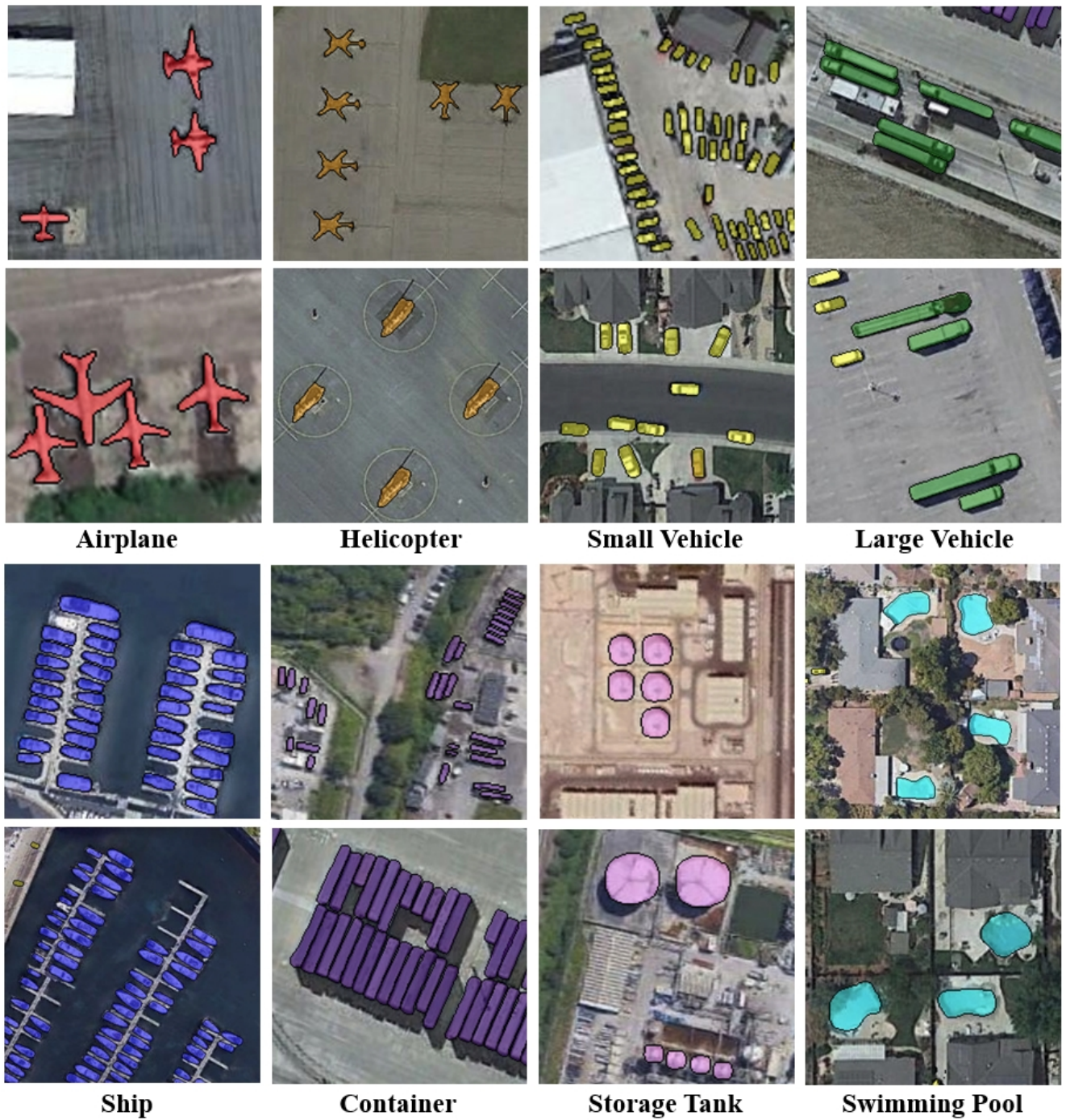


Figure 6. Visualization examples from the ReSOS dataset across eight object categories. Each row shows multiple sample patches with colored instance mask.

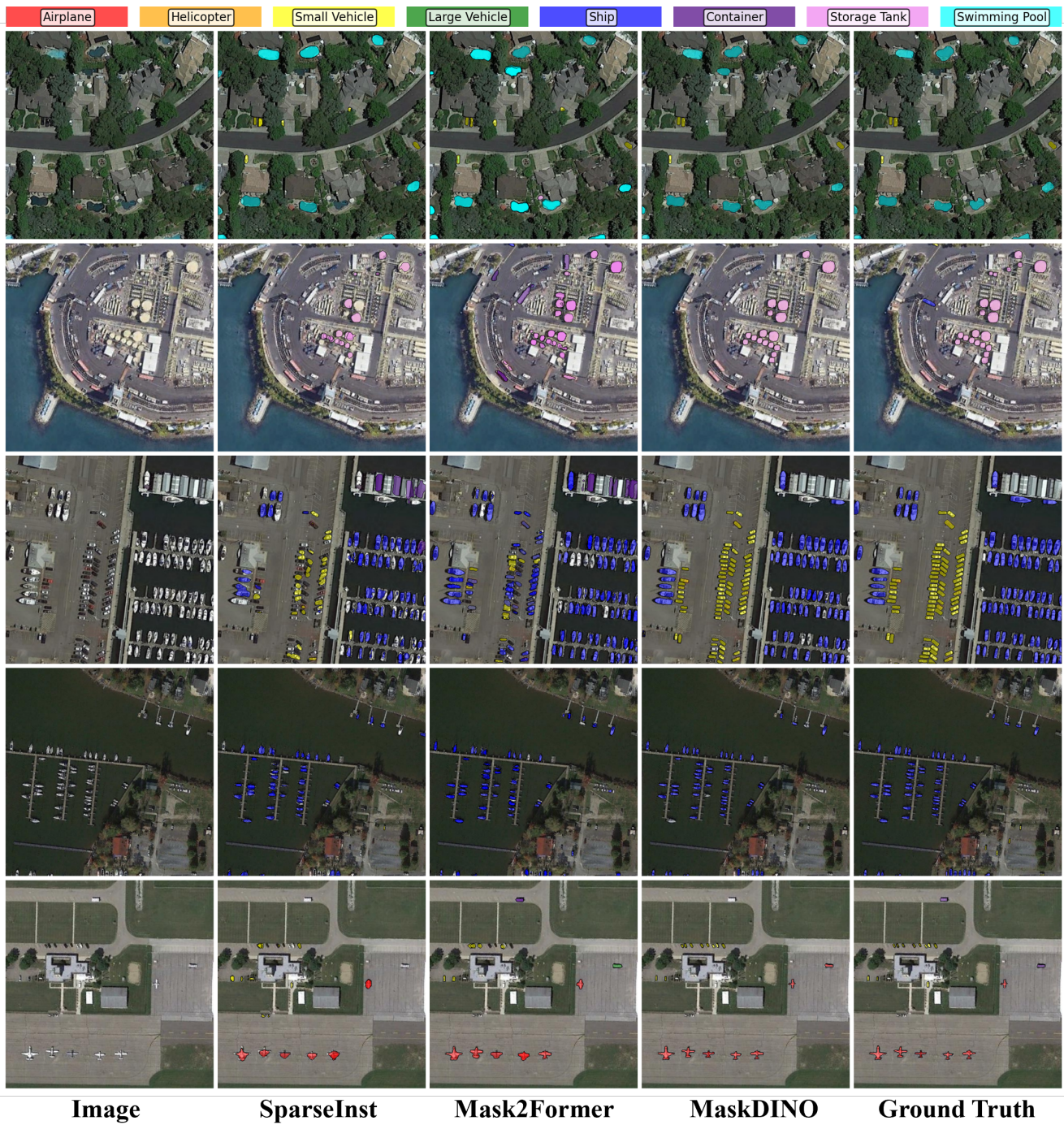


Figure 7. Comparative instance segmentation results on the ReSOS dataset. The ground truth annotations were generated by our SOPSeg.