SOS: Segment Object System for Open-World Instance Segmentation With Object Priors

Christian Wilms, Tim Rolff, Maris Hillemann, Robert Johanson, Simone Frintrop

University of Hamburg
Department of Informatics

ECCV 2024, Milan

Open-World Instance Segmentation (OWIS)

What is open-world instance segmentation?

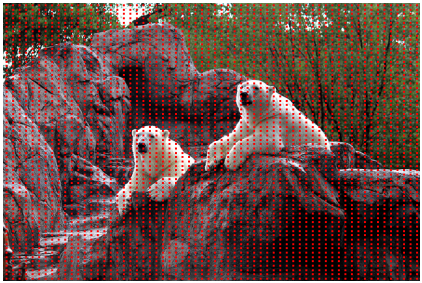
Annotation in training set

Annotation in test set

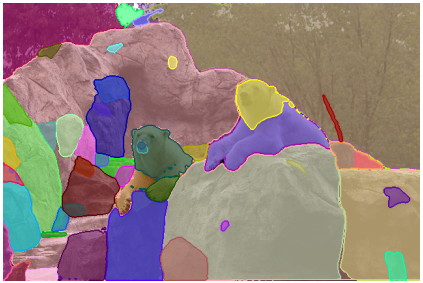
Task: Train on seen object class (`Person`,...),
test on unseen object classes (`Surf board`,...)

Segment Anything Model (SAM) for OWIS

What about solving OWIS with pre-trained SAM based on a grid of prompts [Kirillov et al., ICCV'23]?



Grid of prompts for SAM



Segments of SAM

Problem: SAM segments anything, **objects and stuff**

Segment Anything Model (SAM) for OWIS

What about solving OWIS with pre-trained SAM based on a grid of prompts [Kirillov et al., ICCV'23]?

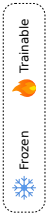
Our object focused prompts

Segments of SAM

Problem: SAM segments anything, **objects and stuff**

Our idea: Focus SAM prompts on objects, use segments as pseudo annotations

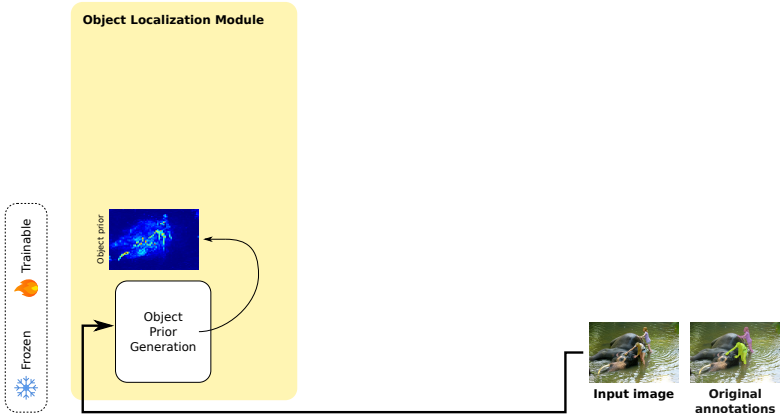
Segment Object System (SOS)



Input image

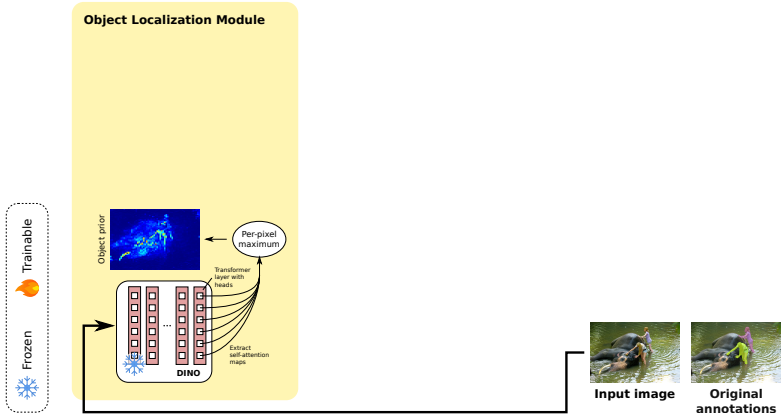
Original annotations

Segment Object System (SOS)



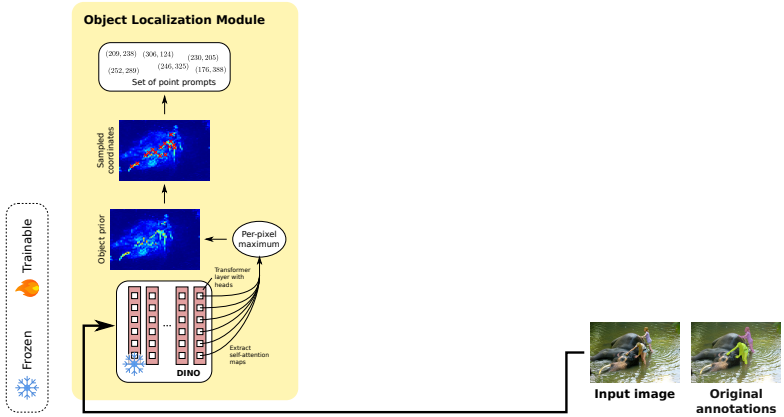
Rough localization of objects with object prior

Segment Object System (SOS)



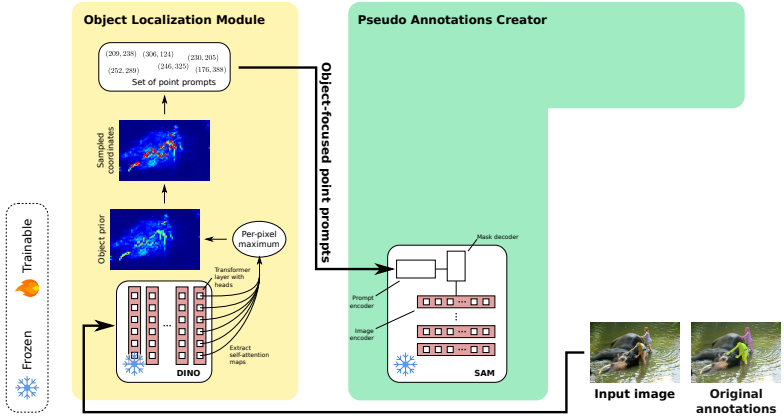
Self-attention maps from self-supervised DINO [Caron et al., ICCV'21] work best

Segment Object System (SOS)



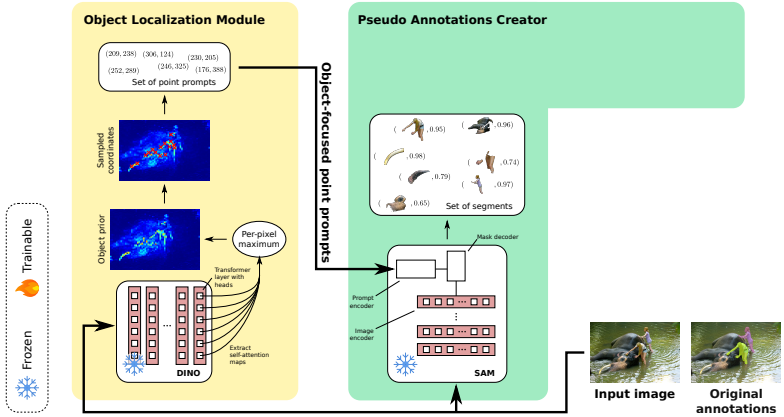
Sample object-focused prompts for SAM

Segment Object System (SOS)



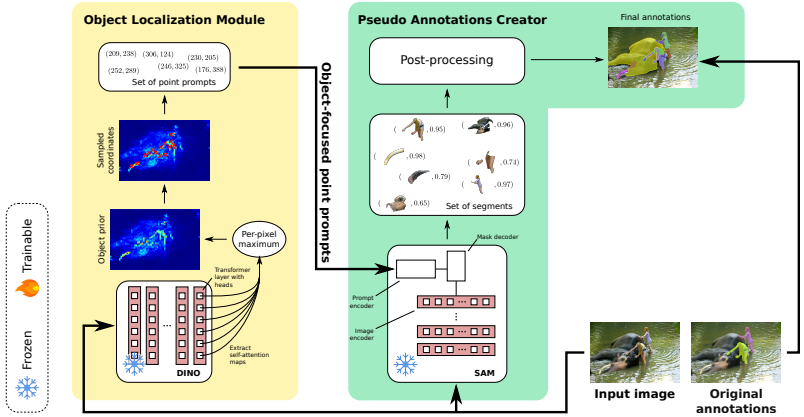
Sample object-focused prompts for SAM

Segment Object System (SOS)



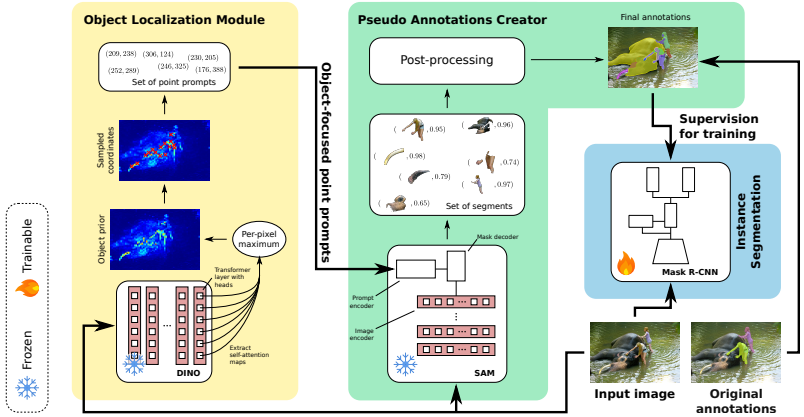
Generate segments with pre-trained SAM

Segment Object System (SOS)



Combine pseudo annotations with original annotations

Segment Object System (SOS)



Train vanilla, class-agnostic Mask R-CNN

SOS on Cross-Category OWIS: COCO (VOC) → COCO (non-VOC)

Evaluation on COCO dataset, training on VOC classes, evaluation on non-VOC classes

System	AP	AR	F_1
SAM [Kirillov et al., ICCV'23]	3.6	**48.1**	6.7
LDET [Saito et al., ECCV'22]	4.3	24.8	7.3
GGN [Wang et al., CVPR'22]	4.9	28.3	8.4
SWORD [Wu et al., ICCV'23]	4.8	30.2	8.3
SOS (ours)	**8.9**	39.3	**14.5**

AP Average precision
AR Average recall

Paper and code

Take-home Messages

1. We transform SAM into an **object-focused segmentation system w/o training**
2. SOS substantially improves precision in OWIS