MAPR 2025 THE 8th INTERNATIONAL CONFERENCE ON MULTIMEDIA ANALYSIS AND PATTERN RECOGNITION

Generative One-shot Camouflage Instance Segmentation

Thanh-Danh Nguyen^{1,2}, Vinh-Tiep Nguyen^{†1,2}, and Tam V. Nguyen³

¹University of Information Technology, Ho Chi Minh City, Vietnam, ²Vietnam National University, Ho Chi Minh City, Vietnam, ³University of Dayton, Dayton, OH 45469, United States {danhnt, tiepnv}@uit.edu.vn, tamnguyen@udayton.edu, †corresponding author



INTRODUCTION

Motivation: Addressing camouflage instance segmentation (CIS) efficiently given one-shot annotated samples.

Main contribution: CAMO-GenOS, a CIS framework leveraging one-shot annotated samples to drive a generative process for data enrichment.

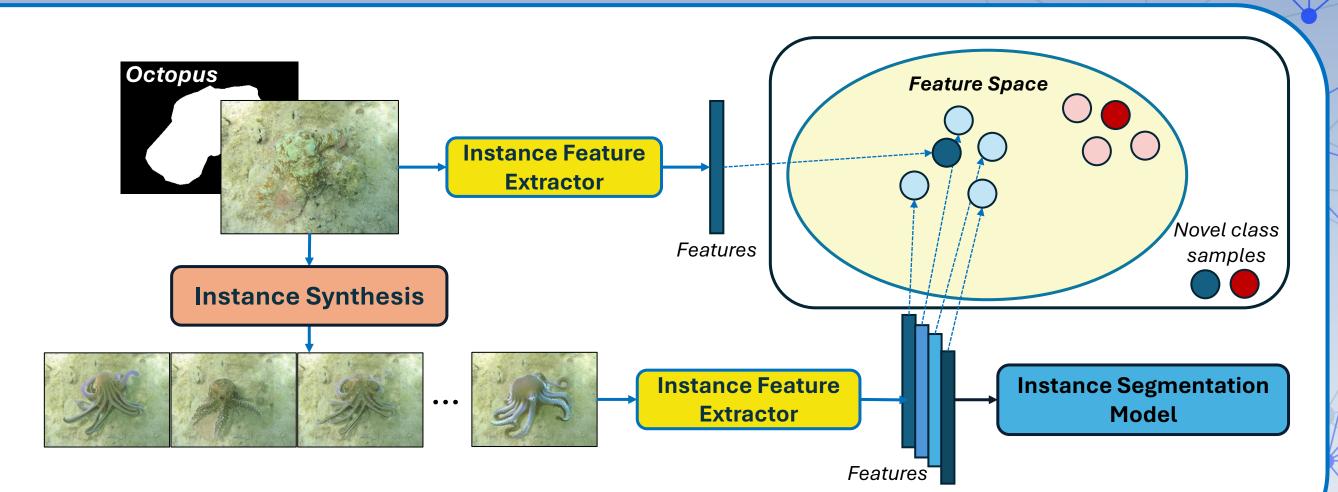
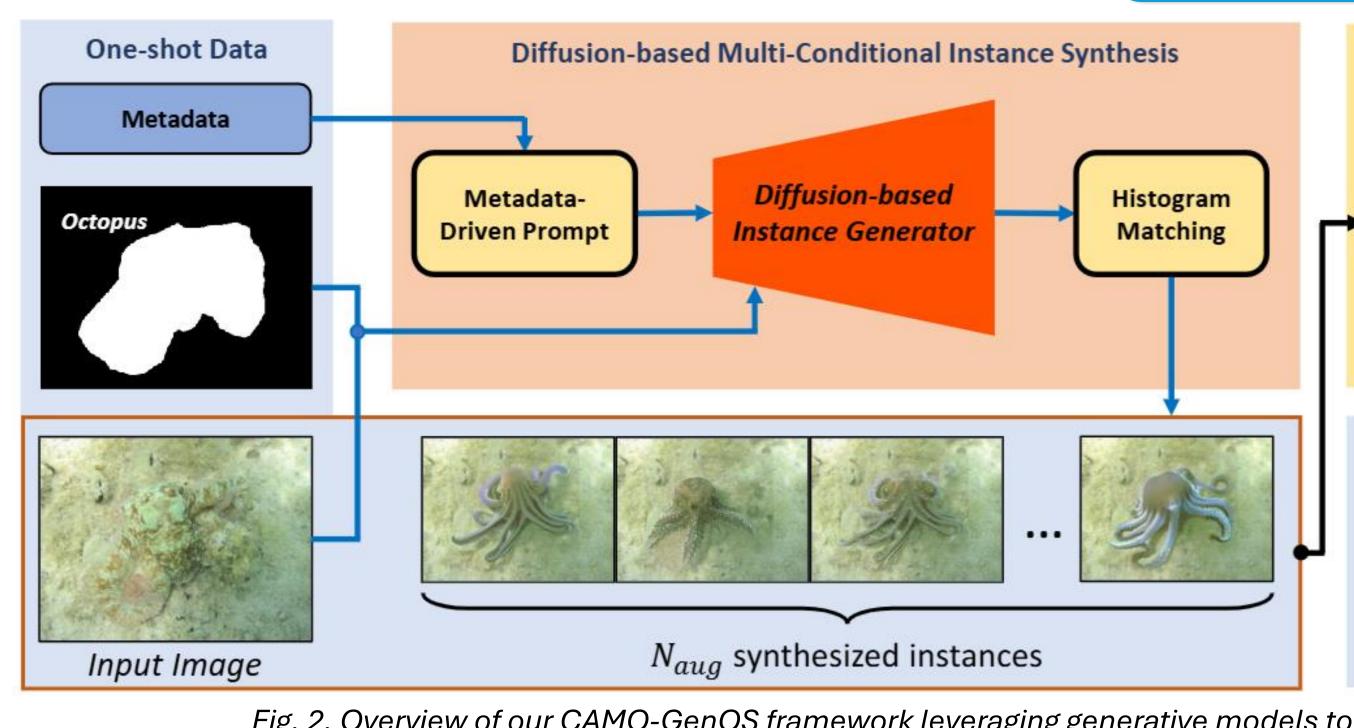
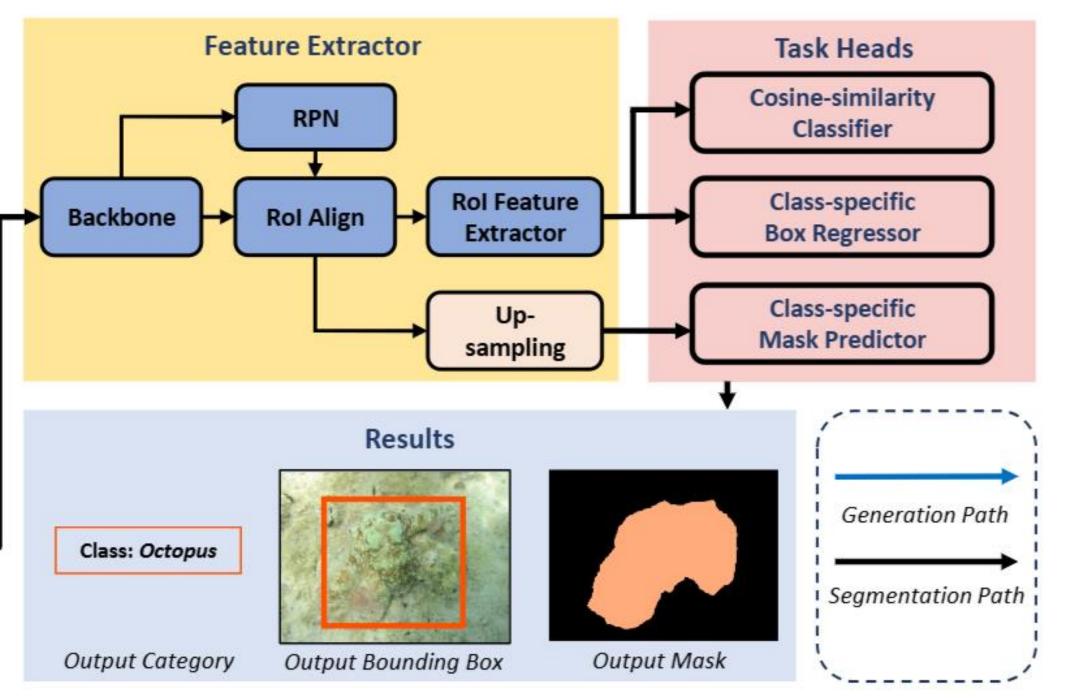


Fig. 1. Our CAMO-GenOS Concept Idea Presentation.

METHOD





Our **CAMO-GenOS** is a generative approach to exploit and enrich the one-shot CIS task.

Fig. 2. Overview of our CAMO-GenOS framework leveraging generative models to enhance one-shot camouflage instance segmentation.

One-shot Fine-tuning CIS: The CIS model is presented in two phases:

- Base phase: training with abundant annotated training data C_{base}
- Novel phase: fine-tuning on a disjoint set of novel classes C_{novel} containing one sample per each class of novel data

In our case, the one-shot fine-tuning relies on a single annotated sample, enriched by our proposed generative instance synthesis method.

Diffusion-based Multi-Conditional Instance Synthesis:

The diffusion-based generative model G(.) strictly takes the multiple conditions of the referenced query image $I_a \in \text{novel}$ training samples in C_{novel} , ground truth mask M_a , and guided text prompt P_q to return N_{auq} samples.

- Metadata-Driven Conditional Text Prompt follows the structure: "a photo of a/an [size] [meta-class] [instance class]"
- Histogram Matching Post-processing is to narrow the gap between the distribution of the original and the synthesized samples

RESULT

Tab. 1. SoTA comparison of our CAMO-GenOS evaluated on CAMO-FS benchmark. The backbones are COCO-80 FPN-ResNet-101

Method	Synthesis Base	nAP		nAP50	nAP75	nAPs	nAPm	nAPl	nAR1
Mask-RCNN [28] iMTFA [11] iFS-RCNN [10] FS-CDIS [7]		2.99 3.66 4.27 4.46	_	5.73 5.37 5.98 7.34	3.26 4.09 4.75 4.84	20.68 22.42 21.57 25.50	3.06 4.35 5.71 5.60	2.74 2.01 4.87 3.48	12.45 11.30 11.70 14.77
CAMO-GenOS (ours)	BlendedDiff [19] DiffInpainting [17] GLIGEN [18]	4.80 4.91 4.74	+0.34 +0.45 +0.28	7.79 7.84 7.53	5.37 5.47 5.31	28.59 26.54 28.10	5.67 5.06 4.79	3.32 4.02 5.28	17.85 17.18 17.65

Tab. 2. Ablation study on multiple instance generation-based methods

Instance Segmentation				Object Detection				
nAP		nAP50	nAP75	nAP		nAP50	nAP75	
4.46		7.34	4.84	3.88		7.71	3.21	
4.55		7.52	4.94	3.99		7.92	3.47	
3.94		7.44	3.64	4.01		8.05	3.44	
4.10		7.40	4.15	3.99		7.82	3.40	
4.80	+0.34	7.79	5.37	4.90	+1.02	8.09	4.78	
5.16	+0.61	8.25	5.73	4.97	+0.98	8.54	5.08	
4.19	+0.25	7.98	4.54	4.75	+0.74	8.38	5.16	
4.25	+0.15	7.36	4.71	4.79	+0.80	7.71	4.52	
4.91	+0.45	7.84	5.47	5.00	+1.12	8.33	5.26	
4.80	+0.25	7.90	5.32	4.97	+0.98	8.29	4.61	
4.04	+0.10	7.21	4.34	4.68	+0.69	7.84	4.84	
4.29	+0.19	7.30	4.60	4.70	+0.71	7.83	4.86	
4.74	+0.28	7.53	5.31	4.83	+0.95	7.94	4.85	
5.30	+0.75	8.26	6.02	5.23	+1.24	8.63	5.61	
4.39	+0.45	7.28	4.86	4.52	+0.51	7.87	4.52	
4.33	+0.23	7.28	4.74	4.75	+0.76	7.62	5.52	
	4.46 4.55 3.94 4.10 4.80 5.16 4.19 4.25 4.91 4.80 4.04 4.29 4.74 5.30 4.39	4.46 4.55 3.94 4.10 4.80 +0.34 5.16 +0.61 4.19 +0.25 4.25 +0.15 4.91 +0.45 4.80 +0.25 4.04 +0.10 4.29 +0.19 4.74 +0.28 5.30 +0.75 4.39 +0.45	4.46 7.34 4.55 7.52 3.94 7.44 4.10 7.40 4.80 +0.34 7.79 5.16 +0.61 8.25 4.19 +0.25 7.98 4.25 +0.15 7.36 4.91 +0.45 7.84 4.80 +0.25 7.90 4.04 +0.10 7.21 4.29 +0.19 7.30 4.74 +0.28 7.53 5.30 +0.75 8.26 4.39 +0.45 7.28	4.46 7.34 4.84 4.55 7.52 4.94 3.94 7.44 3.64 4.10 7.40 4.15 4.80 +0.34 7.79 5.37 5.16 +0.61 8.25 5.73 4.19 +0.25 7.98 4.54 4.25 +0.15 7.36 4.71 4.91 +0.45 7.84 5.47 4.80 +0.25 7.90 5.32 4.04 +0.10 7.21 4.34 4.29 +0.19 7.30 4.60 4.74 +0.28 7.53 5.31 5.30 +0.75 8.26 6.02 4.39 +0.45 7.28 4.86	4.46 7.34 4.84 3.88 4.55 7.52 4.94 3.99 3.94 7.44 3.64 4.01 4.10 7.40 4.15 3.99 4.80 +0.34 7.79 5.37 4.90 5.16 +0.61 8.25 5.73 4.97 4.19 +0.25 7.98 4.54 4.75 4.25 +0.15 7.36 4.71 4.79 4.91 +0.45 7.84 5.47 5.00 4.80 +0.25 7.90 5.32 4.97 4.04 +0.10 7.21 4.34 4.68 4.29 +0.19 7.30 4.60 4.70 4.74 +0.28 7.53 5.31 4.83 5.30 +0.75 8.26 6.02 5.23 4.39 +0.45 7.28 4.86 4.52	4.46 7.34 4.84 3.88 4.55 7.52 4.94 3.99 3.94 7.44 3.64 4.01 4.10 7.40 4.15 3.99 4.80 +0.34 7.79 5.37 4.90 +1.02 5.16 +0.61 8.25 5.73 4.97 +0.98 4.19 +0.25 7.98 4.54 4.75 +0.74 4.25 +0.15 7.36 4.71 4.79 +0.80 4.91 +0.45 7.84 5.47 5.00 +1.12 4.80 +0.25 7.90 5.32 4.97 +0.98 4.04 +0.10 7.21 4.34 4.68 +0.69 4.29 +0.19 7.30 4.60 4.70 +0.71 4.74 +0.28 7.53 5.31 4.83 +0.95 5.30 +0.75 8.26 6.02 5.23 +1.24 4.39 +0.45 7.28 4.86 4.52 +0.51	4.46 7.34 4.84 3.88 7.71 4.55 7.52 4.94 3.99 7.92 3.94 7.44 3.64 4.01 8.05 4.10 7.40 4.15 3.99 7.82 4.80 +0.34 7.79 5.37 4.90 +1.02 8.09 5.16 +0.61 8.25 5.73 4.97 +0.98 8.54 4.19 +0.25 7.98 4.54 4.75 +0.74 8.38 4.25 +0.15 7.36 4.71 4.79 +0.80 7.71 4.91 +0.45 7.84 5.47 5.00 +1.12 8.33 4.80 +0.25 7.90 5.32 4.97 +0.98 8.29 4.04 +0.10 7.21 4.34 4.68 +0.69 7.84 4.29 +0.19 7.30 4.60 4.70 +0.71 7.83 4.74 +0.28 7.53 5.31 4.83 +0.95 7.94 5.30 +0.75 8.26 <	

ORGANIZERS

















SPONSORS

