# Few-shot Camouflaged Animal Detection and Segmentation

Thanh-Danh Nguyen[1,4†], Anh-Khoa Nguyen Vu[1,4†], Nhat-Duy Nguyen[1,4†], Vinh-Tiep Nguyen[1,4], Thanh Duc Ngo[1,4], Thanh-Toan Do[5], Minh-Triet Tran[2,3,4] and Tam V. Nguyen[6*]

[1]University of Information Technology, Ho Chi Minh City, Vietnam. [2]University of Science, Ho Chi Minh City, Vietnam. [3]John von Neumann Institute, VNU-HCM, Vietnam. [4]Vietnam National University, Ho Chi Minh City, Vietnam. [5]Monash University, Clayton, VIC 3800, Australia. [6*]University of Dayton, Dayton, OH 45469, United States. ([†]equal contribution, *corresponding author)

CVPR JUNE 19-24 2022 NEW ORLEANS LOUISIANA

CV4Animals Workshop

## Introduction

**Overview:** Camouflage is a defense mechanism that animals use to conceal their appearance by blending in with their environment. Autonomously detecting camouflaged animals is helpful in various fields of computer vision: search-and-rescue mission; wild species discovery and preservation activities; media forensics, etc.

**Motivation:**

- Research on camouflaged animals suffers from the lack of data
- Camouflaged instances have their texture similar to the background

**Main contributions:**

1. A novel benchmark **CAMO-FS** for few-shot detection and segmentation on camouflaged animals
2. A **framework** to efficiently detect and segment camouflaged instances given a small number of training data for novel classes, thanks to **Instance Triplet Loss** and **Instance Memory Storage**

## CAMO-FS Benchmark Dataset

Our **CAMO-FS** benchmark contains:

- **2,852** images/ **3,342** camouflaged instances with fine-grained annotations for detection and instance segmentation (enhanced 163 images compared with CAMO++ dataset [1])
- **47** semantic classes, customized for few-shot learning with 1, 2, 3, and 5-shot

This dataset is among the firsts to address few-shot camouflaged detection and instance segmentation

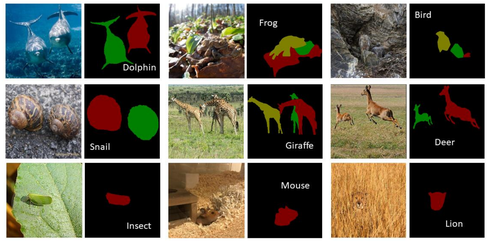


*Fig. 1 Exemplary camouflaged instances in CAMO-FS*

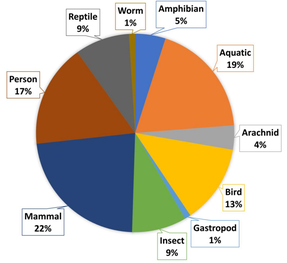


*Fig. 2 Distribution of camouflaged coarse-grained classes in CAMO-FS*

## General Framework

Originated from MTFA [2] model which is a two-stage training and fine-tuning mechanism, we introduce a framework for few-shot camouflaged detection and instance segmentation as below:

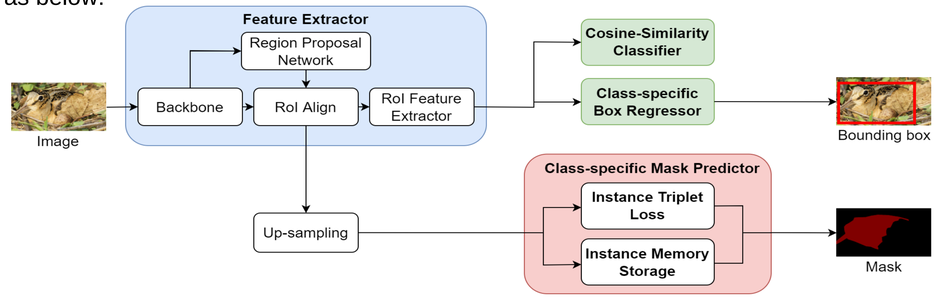


*Fig. 3 General framework for camouflaged animal detection and instance segmentation*

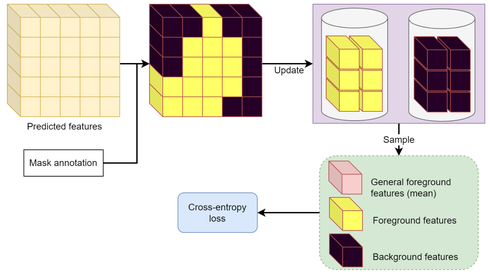


*Fig. 4 Instance Memory Storage*

**Instance Memory Storage**

Contains the background and foreground features per each class, and calculates the discrimination between areas of objects and no-objects in region proposals

**Instance Triplet Loss**

Minimizes/ maximizes the distances among foreground/ background representatives

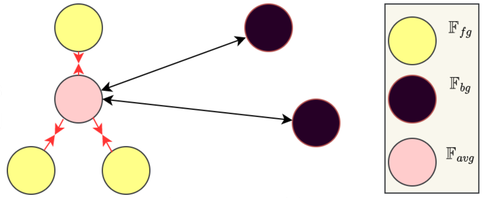


*Fig. 5 Instance Triplet Loss*

## Acknowledgement

This research was supported by the VNUHCM-University of Information Technology's Scientific Research Support Fund

## Experimental Results

*Tab. 1 Our improvement of Instance Triplet Loss and Memory Storage on MTFA model*

| # | Method | AP | AP50 | AP75 | APs | APm | APl | AR1 | AR10 | ARs | ARm | ARl |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Instance Segmentation | | | | | | | | | | | |
| 1 | Baseline | 3.66 | 5.37 | 4.09 | 22.42 | 4.35 | 2.01 | 11.30 | 13.58 | 25.97 | **12.96** | 12.53 |
| | + Triplet | 4.46 | 8.21 | 4.60 | 21.33 | 4.13 | **4.01** | 12.36 | 15.04 | 23.17 | 9.49 | 16.67 |
| | + Memory | **5.46** | **9.20** | **6.17** | **27.79** | **6.20** | **4.01** | **17.08** | **19.99** | **29.41** | 11.45 | **20.89** |
| 2 | Baseline | 6.21 | 8.92 | 7.28 | 32.64 | **7.75** | 3.50 | 18.88 | 21.12 | **35.82** | **15.49** | 20.14 |
| | + Triplet | 5.57 | 9.45 | 6.04 | 25.83 | 3.01 | 5.37 | 15.67 | 17.33 | 26.13 | 7.37 | 17.50 |
| | + Memory | **6.95** | **10.72** | **7.60** | **33.62** | 5.73 | **6.44** | **20.00** | **22.15** | 34.25 | 13.86 | **20.92** |
| 3 | Baseline | 6.16 | 8.95 | 6.68 | 33.74 | 6.19 | 5.08 | 20.25 | 22.95 | 36.83 | 16.31 | 21.63 |
| | + Triplet | 6.41 | 10.67 | 6.72 | 30.39 | 5.17 | 5.30 | 20.69 | 22.98 | 31.90 | 15.69 | 22.53 |
| | + Memory | **7.36** | **11.23** | **8.49** | **37.03** | **6.24** | **5.64** | **24.40** | **27.69** | **38.44** | **17.02** | **26.71** |
| 5 | Baseline | 5.95 | 8.67 | 6.94 | 34.71 | 6.25 | 4.85 | 21.29 | 24.42 | 36.86 | **14.51** | 24.83 |
| | + Triplet | 8.48 | 13.43 | 9.80 | 36.66 | 5.75 | 8.01 | 23.83 | 26.66 | 37.03 | 11.62 | 25.91 |
| | + Memory | **9.61** | **14.61** | **11.73** | **38.60** | **5.79** | **10.40** | **26.65** | **30.37** | **39.21** | 12.26 | **30.02** |
| | Detection | | | | | | | | | | | |
| 1 | Baseline | 2.93 | 5.86 | 2.20 | 20.95 | 4.18 | 2.03 | 9.25 | 10.84 | 21.74 | **11.49** | 8.77 |
| | + Triplet | 4.04 | 8.65 | 2.98 | 20.50 | 4.90 | **4.22** | 12.89 | **15.53** | 20.73 | 11.45 | **17.46** |
| | + Memory | **4.50** | **9.14** | **3.45** | **22.88** | **5.61** | 3.54 | **13.14** | 15.22 | **23.14** | 8.78 | 16.33 |
| 2 | Baseline | 5.90 | 8.87 | 6.83 | 33.04 | 9.74 | 3.10 | 17.26 | 19.25 | 34.04 | **15.74** | 19.61 |
| | + Triplet | **7.28** | **11.22** | **8.25** | 32.31 | **10.72** | **6.83** | **20.52** | **22.69** | 32.34 | 14.88 | **23.52** |
| | + Memory | 6.95 | 10.88 | 7.75 | **33.93** | 7.49 | 6.81 | 19.84 | 22.01 | **34.10** | 15.04 | 21.47 |
| 3 | Baseline | 5.84 | 8.98 | 6.29 | 34.56 | 7.78 | 4.31 | 19.13 | 21.83 | 35.80 | 15.93 | 21.09 |
| | + Triplet | 7.49 | **11.51** | 8.23 | **38.45** | 8.61 | **6.38** | **24.88** | **27.52** | **38.55** | 17.66 | 27.44 |
| | + Memory | **7.55** | 11.45 | **8.50** | 38.07 | **9.21** | 5.70 | 24.20 | 27.29 | 38.50 | **18.10** | **27.56** |
| 5 | Baseline | 5.84 | 9.13 | 6.04 | 35.44 | 8.17 | 4.22 | 19.67 | 22.96 | 35.94 | **14.16** | 22.58 |
| | + Triplet | 9.76 | 14.37 | 11.12 | **40.05** | **8.82** | 9.89 | 25.93 | 29.28 | **40.05** | 12.53 | 30.32 |
| | + Memory | **10.36** | **16.27** | **11.79** | 39.32 | 8.08 | **11.36** | **26.34** | **30.30** | 39.35 | 12.37 | **30.91** |

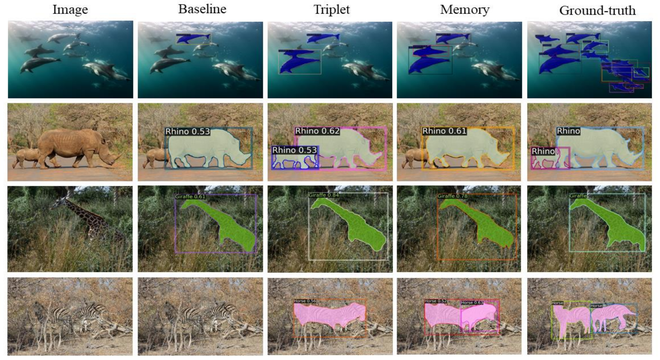


*Fig. 6 Qualitative comparison among the baseline MTFA and our proposed methods (5-shot)*

## References

[1] Le, T.-N. et al.: Camouflaged instance segmentation in-the-wild: Dataset, method, and benchmark suite. IEEE Transactions on Image Processing 31, 287-300 (2022)

[2] Ganea, D.A. et al.: Incremental few-shot instance segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 1185-1194 (2021)