INTERNATIONAL NEURAL NETWORK SOCIETY

IEEE WCCI 2024 Yokohama Japan

Few–Shot Camouflaged Object Segmentation

Ziqiu Wang[1], Yuying Li[1], Yang Yang[1], Yamin Li[2*], Gaoyang Liu[3]

[1] School of Artificial Intelligence, Hubei University

[2] School of Computer Science and Information Engineering, Hubei University

[3] Huazhong University of Science and Technology

Email: {yamin.li, yangyang}@hubu.edu.cn

* Indicates corresponding author

湖北大学 HUBEI UNIVERSITY

華中科技大學 HUAZHONG UNIVERSITY SCIENCE AND TECHNOLOGY

INTRODUCTION

Problem statement：

1. In the field of Camouflaged Object Segmentation (COS), traditional techniques heavily rely on supervised learning, which requires extensive labeled datasets. This dependency can be a limitation as these methods often struggle to accurately segment unseen classes.
2. Advanced unsupervised models (like SAM) face difficulties in effectively segmenting camouflaged objects when the distinction between foreground and background is unclear.

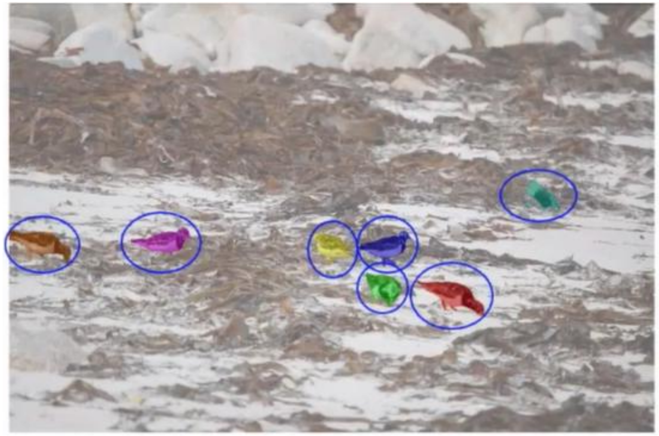
Can you find all the birds in the picture?

SAM fails to perceive the Camouflaged animals.

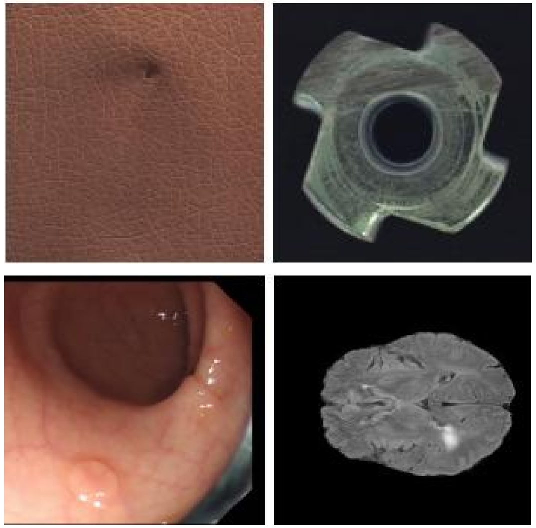
INTRODUCTION

Significance:

1. COS plays a critical role in various domains, such as medical diagnosis, military operations, agriculture applications, and industrial quality control.
2. The limited diversity in public COS datasets like COD10K restricts their applicability in complex real-world scenarios.



Answer,did you find them?

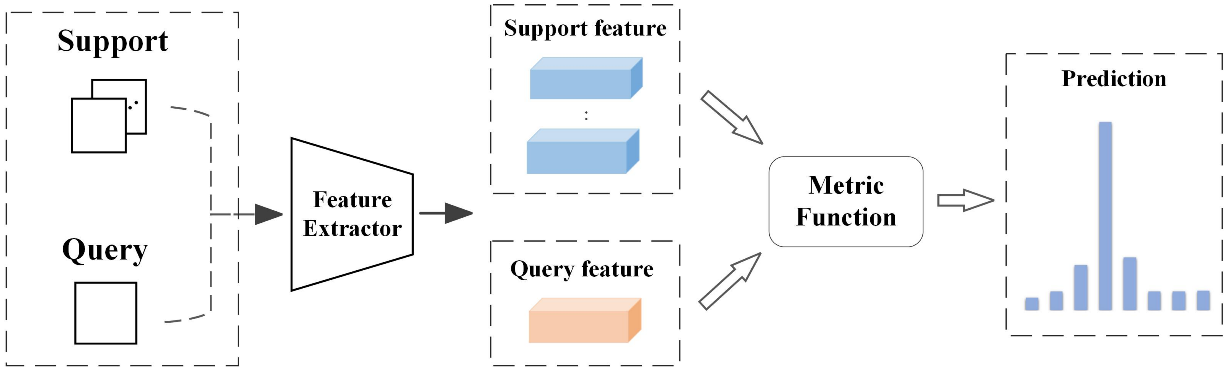


Camouflaged Object Segmentation in various domains.

RELATED WORK

Camouflaged Object Segmentation:

- COS is characterized by complex backgrounds and often indistinct object boundaries, requiring enhanced attention to edge details in learning models.



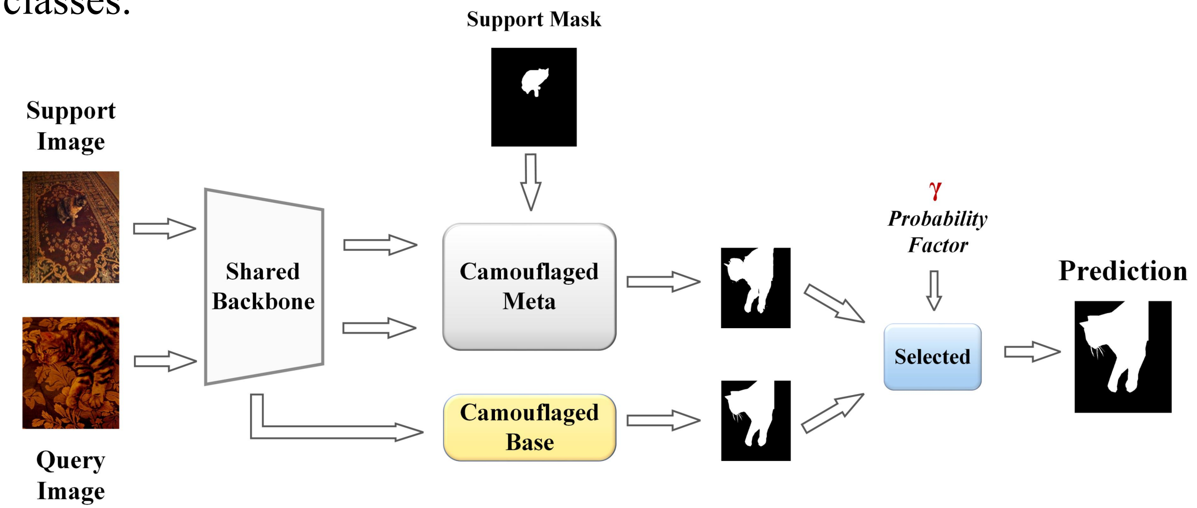
Few-Shot Learning:

- FSL, a subfield of meta-learning, is designed to replicate the human capability of quickly learning from limited examples.
- In FSL applications, metric-based methods are particularly prevalent, accounting for about 50% of the approaches.

METHODOLOGY

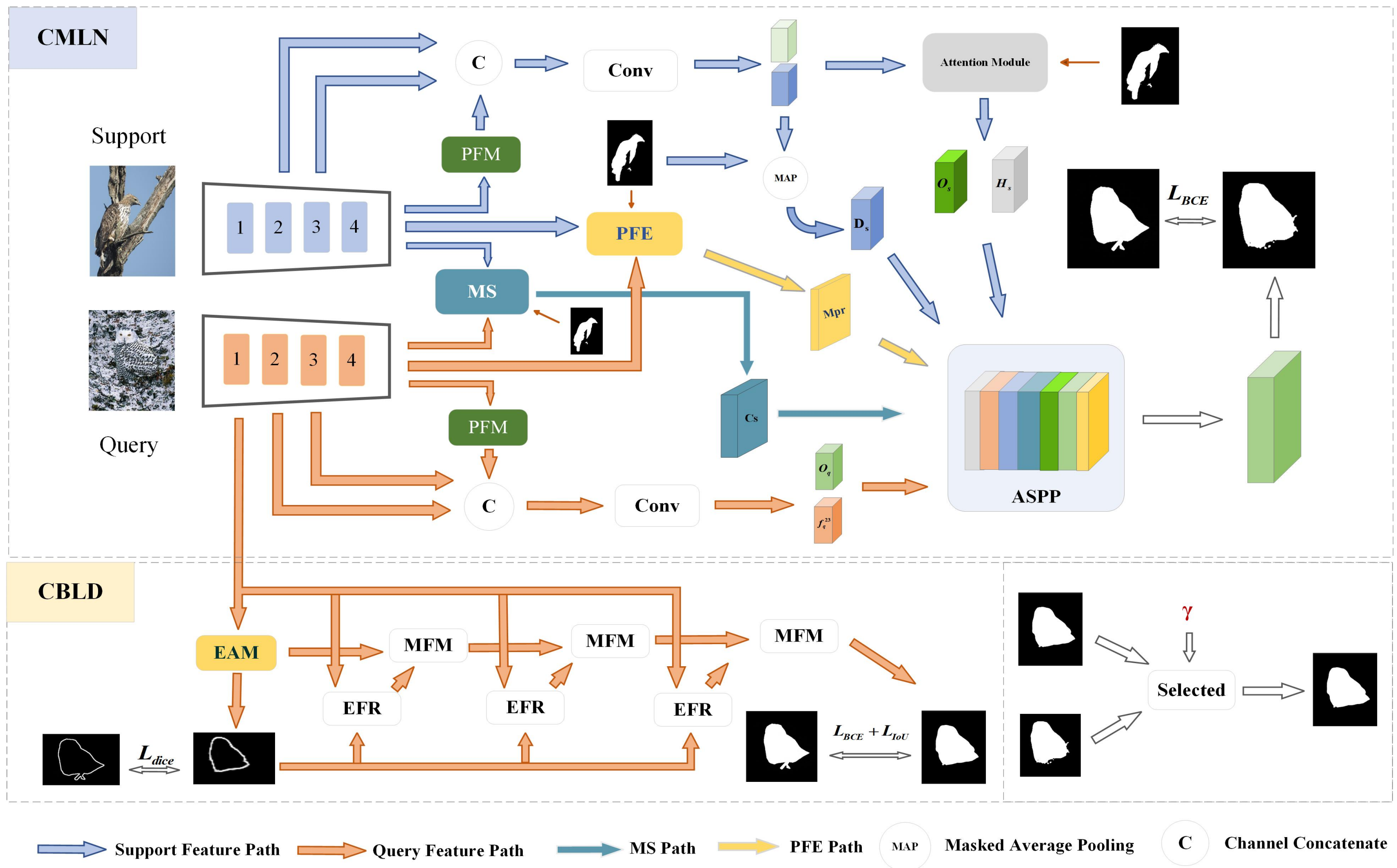
challenge:

- Challenges in segmenting camouflaged objects, especially in dealing with unseen classes.A key focus of our work is how to extract fine-grained information from the support set.
- The existing COS public datasets cannot be directly applied to few-shot scenarios due to the imbalance of classes.



An illustration of our CAMFS framework, which can segment camouflaged objects of unseen classes by transferring knowledge from support images information to query images.

METHODOLOGY



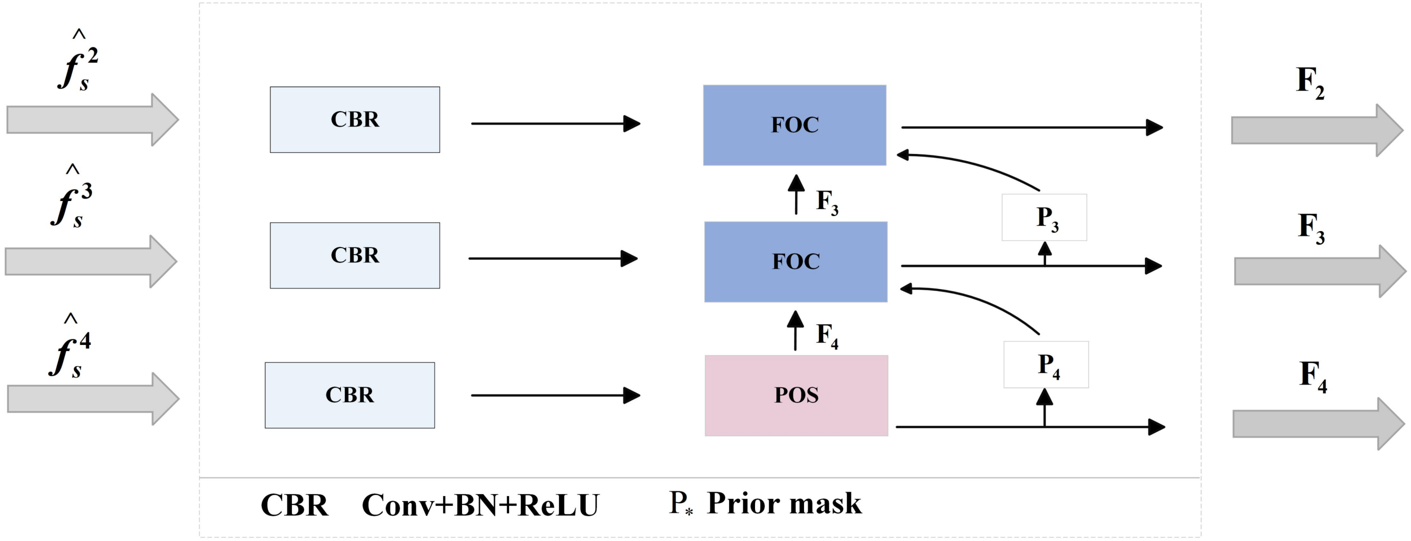
This illustration provides an overview of the CAMFS architecture, extracting pertinent information from the support set to effectively segment camouflaged object images in the query set.

METHODOLOGY

Addressing Challenges:

How to extract fine-grained information from the support set?

we employ a branching strategy for processing these features to effectively capture camouflage details while preserving the integrity of original features, The primary branch includes the Positioning and Focusing Module.



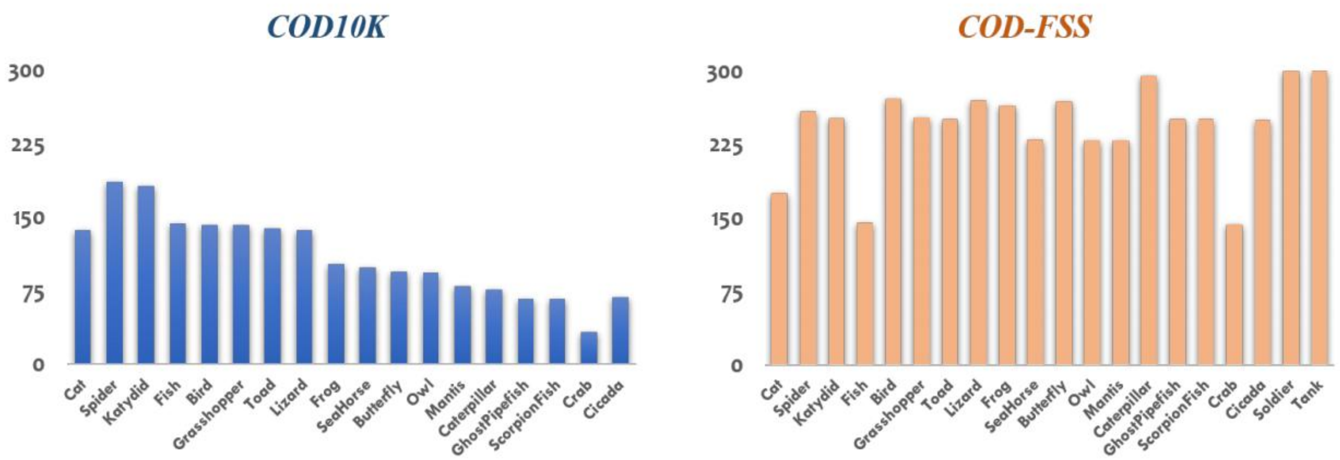
The PFM (Positioning and Focusing Module) enhances detailed feature extraction of camouflaged objects. It processes feature maps through a branching strategy, employs depthwise convolution to reduce channels, and uses iterative Positioning and Focusing techniques to enrich feature representation, boosting the detection of camouflaged targets.

METHODOLOGY

Addressing Challenges:

How to solve the issue that existing public datasets cannot be applied to the few-shot domain?

- We selected 18 categories with the highest image counts from COD10K and further augmented these categories to approximately 250 images each. Additionally, to incorporate diversity and address the lack of military camouflage images in COD10K, we included about 1000 images of military camouflage (tanks and soldiers).



Histogram left displays the distribution of samples within the 18 classes having the largest number of images in the COD10K dataset, while Histogram right illustrates the sample distribution across the 20 classes within our COS-FSS dataset.

EXPERIMENTS

Experimental Settings

- The CAMFS framework employs the ResNet-50 model,pre-trained on ImageNet, as its foundational backbone. For training, an NVIDIA GeForce RTX 4090 GPU was utilized.
- The optimization process was carried out using the Stochastic Gradient Descent (SGD) method, with the initial learning rate set at 0.004.
- The total loss for the CBLD stage is represented as $L_{\text{base}} = L_{\text{BCE}} + L_{\text{IoU}} + \lambda L_{\text{dice}}$ we set $\lambda = 3$.
- The second stage of training employs the episodic training method on the CMLN, using binary cross-entropy loss denoted by $L_{\text{meta}} = L_{\text{BCE}}(m_q, M_m)$

TABLE III: The result of comparative performance between CAMFS and five semantic segmentation methods on the COS-FSS benchmark dataset, with symbols ↑ and ↓ indicating, respectively, that higher and lower values of the corresponding metric signify better performance.

Method	F-IoU(%)	B-IoU(%)	FB-IoU(%)	$S_\alpha \uparrow$	$F_\beta^w \uparrow$	$M \downarrow$	$E_\phi^{ad} \uparrow$	$E_\phi^{mn} \uparrow$	$E_\phi^{mx} \uparrow$	$F_\beta^{ad} \uparrow$	$F_\beta^{mn} \uparrow$	$F_\beta^{mx} \uparrow$
FCN [49]	53.10	91.80	72.45	0.7230	0.5676	0.0748	0.7820	0.7800	0.7820	0.6110	0.6080	0.6100
PSPNe [50]	46.21	94.15	70.18	0.7421	0.4434	0.0625	0.8153	0.8129	0.8153	0.4810	0.4796	0.4810
Unet [51]	55.00	95.00	75.00	0.7630	0.5601	0.0538	0.8270	0.8246	0.8270	0.6020	0.6001	0.6020
DeepLab-v3+ [52]	55.20	94.60	74.90	0.7485	0.5640	0.0532	0.8296	0.8273	0.8290	0.6060	0.6048	0.6067
U2net [53]	-	-	-	0.7164	0.5192	0.0692	0.7928	0.7544	0.7834	0.5730	0.5667	0.5791
CAMFS(Ours)	**63.10**	**96.65**	**79.88**	**0.8404**	**0.7482**	**0.0330**	**0.8919**	**0.8819**	**0.8974**	**0.7720**	**0.7756**	**0.7865**

EXPERIMENTS

Few-Shot Setups:

- The CAMFS model's efficacy is evaluated on the complete COS-FSS dataset, consisting of 5322 images. The dataset is divided, allocating 4216 images from the latter 15 classes for training and 1106 images from the initial 5 classes for testing. The training follows the traditional episodic training paradigm, enabling the model to achieve effective segmentation on unseen data from the first 5 classes.Training is executed under a one-shot condition (K=1), where a single image from the support set guides the segmentation process for the query image.

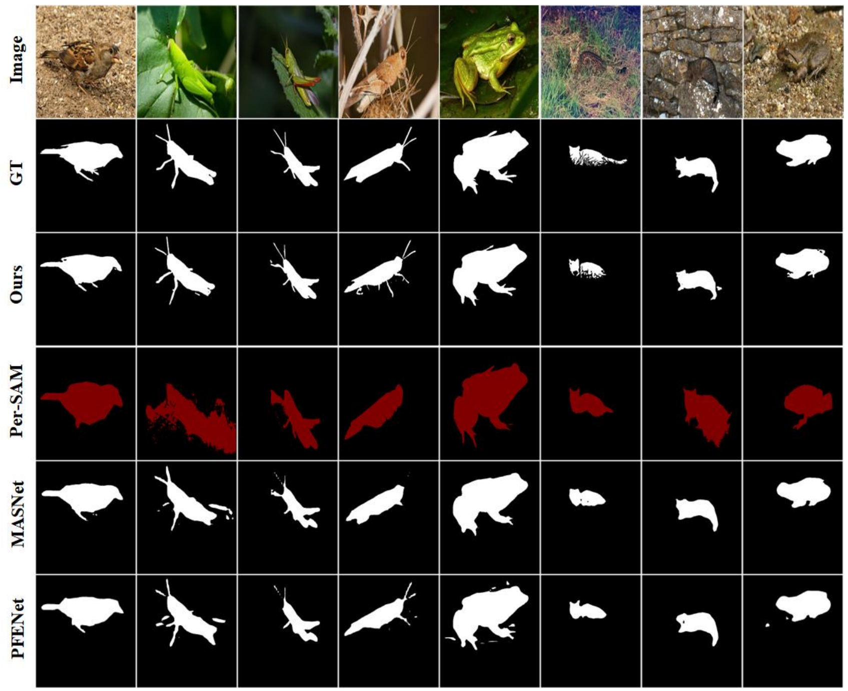
TABLE I: The result of CAMFS in comparison with other FSS on the COS-FSS dataset under one-shot (K=1) settings.

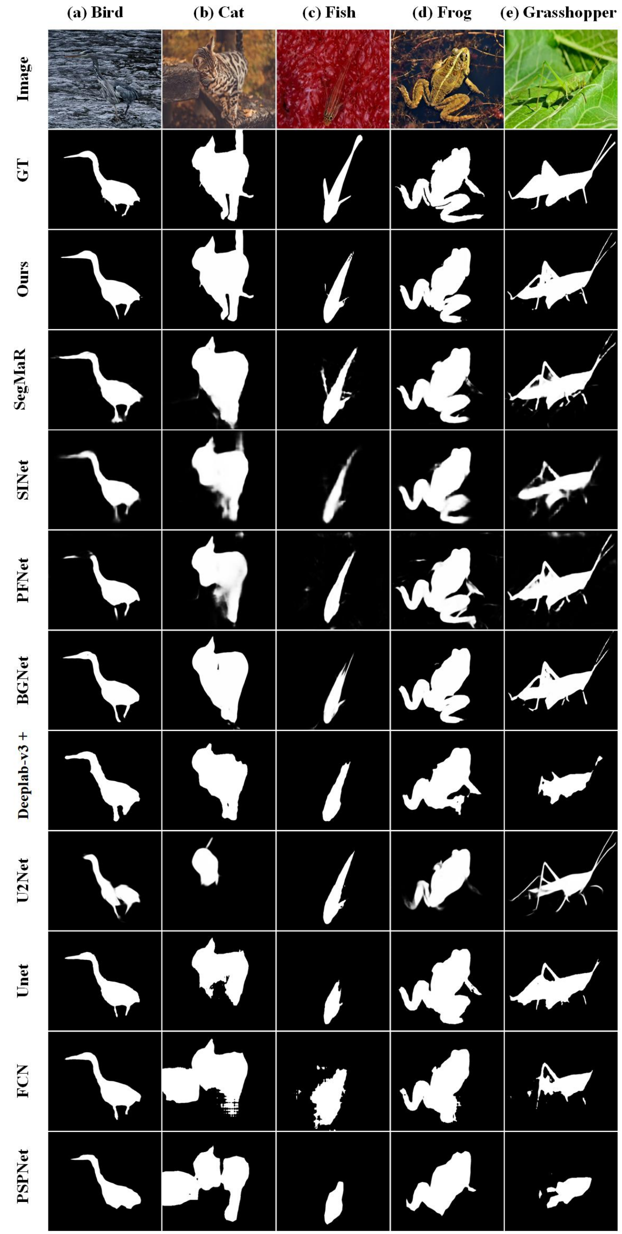
Method	Pub/Year	F-IoU(%)	B-IoU(%)	FB-IoU(%)	Class1	Class2	Class3	Class4	Class5	Mean-Iou(%)
PFENet [40]	CVPR_{2021}	57.36	93.24	75.30	51.00	60.15	63.90	51.75	60.02	57.36
MSANet [43]	arXiv_{2022}	60.94	96.08	78.51	**66.01**	40.77	**66.79**	70.67	**58.25**	60.39
Per-SAM [18]	arXiv_{2023}	23.77	81.92	52.90	-	-	-	-	-	-
CAMFS(Ours)	-	**63.10**	**96.65**	**79.88**	61.59	**62.73**	62.61	**71.03**	53.99	**62.39**

TABLE II: The result of evaluating CAMFS against five COS methods on the COS-FSS dataset.

Method	Pub/Year	$S_\alpha \uparrow$	$F_\beta^w \uparrow$	$M \downarrow$	$E_\phi^{ad} \uparrow$	$E_\phi^{mn} \uparrow$	$E_\phi^{mx} \uparrow$	$F_\beta^{ad} \uparrow$	$F_\beta^{mn} \uparrow$	$F_\beta^{mx} \uparrow$
SINet [12]	CVPR_{2020}	0.795	0.630	0.047	0.839	0.832	0.873	0.665	0.684	0.712
PFNet [13]	CVPR_{2021}	0.760	0.640	0.050	0.854	0.852	0.854	0.683	0.681	0.683
SegMaR [47]	CVPR_{2022}	0.805	0.661	0.046	0.851	0.857	0.870	0.680	0.703	0.729
BGNet [25]	IJCAI_{2022}	0.804	0.685	0.042	0.866	0.851	0.864	0.719	0.719	0.729
PFNet+ [48]	Ssis_{2023}	0.796	0.664	0.048	0.858	0.865	0.865	0.693	0.704	0.721
CAMFS(Ours)	-	**0.840**	**0.748**	**0.033**	**0.892**	**0.882**	**0.897**	**0.772**	**0.776**	**0.787**

RESULT





Comparative analysis of CAMFS model’s segmentation perfo rmance against various FSS methodologies, illustrated on the COS-FSS dataset under one-shot (K=1) conditions

CONCLUSION

Our primary contributions are summarized as follows:

- We conduct the first study to introduce few-shot learning into the field of COS.This pioneering approach significantly reduces the reliance on extensive labeled datasets typically required for COS tasks and enhances the model's capacity for generalization.
- We introduce a CAMFS framework, which includes two novel modules: Camouflaged-Meta and Camouflaged-Base. These modules are designed to ensure effective identification of unseen camouflaged samples using support images.
- We develop COS-FSS, a novel public benchmark dataset,representing the first few-shot camouflage object image segmentation dataset.
- Resource:The dataset and additional resources are available at (https://github.com/CAM-FSS/FSS-COD).

Future work:

- Future research directions include enhancing the model's ability for fine-grained segmentation of small-scale targets within images, further advancing the field of COS.

Thank you