Solving the Camouflage Object Segmentation Problem with Multitask-Learning and MAML Augmented SINet

William Cai[1], Xin Zhang[2]

[1]*Department of Engineering, Stanford University* [2]*Department of Physics, Stanford University*

Stanford Computer Science

Abstract

Motivation:

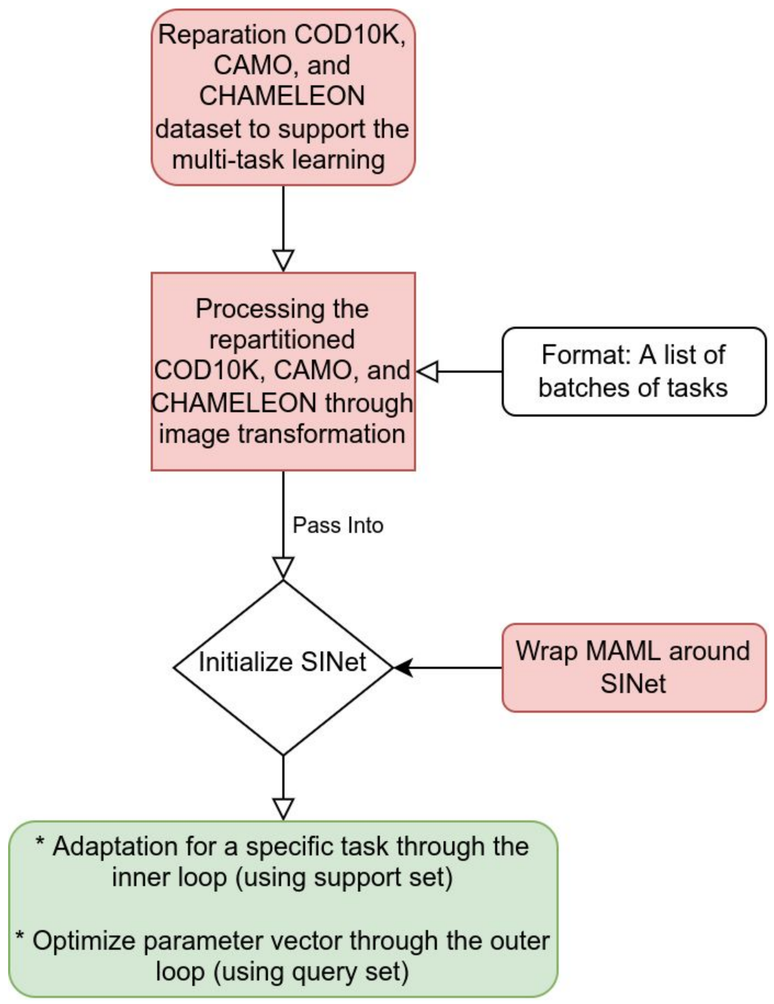
- Camouflage Object Segmentation (COS) about finding out the camouflaged objects that are “seamlessly” embedded in their surroundings.
- COS plays an important role in many areas including medical imaging diagnosis, detection of unqualified products on automatic production lines in the manufacturing industry, detection of personnel in search-and-rescue missions, etc.
- Current state-of-the-art model SINet can achieve a solid performance in solving COS but requires significant amount of computing resources and greatly relies on a huge dataset.

Project goal:

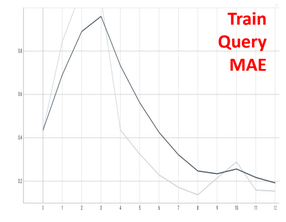
- By augmenting SINet with the Model-Agnostic Meta-learning (MAML) and multitask-learning, we aim to decrease the amount of computing resources and size of the dataset required for the training of SINet while preserving its original performance.

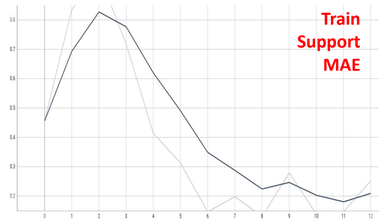
Methods

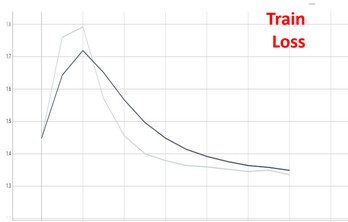
Model: Search Identification Network (SINet)
Datasets of animals: COD10K, CHAMELEON, CAMO
Types of photos: RGB (Img) and greyscale (GT)
Task: images of a given animal
Batch: contains a number of tasks
Allocation: Train : Val : Test = 7 : 2 : 1

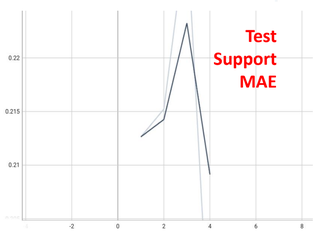


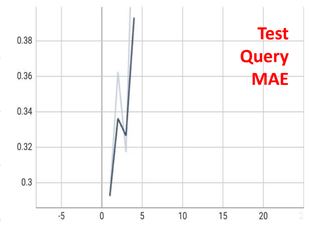
Results











Result 1: Performance of the model in the meta-training throughout with 12 bataches of animals in training. (Each batch contains 4 animal types)

Evaluation metric: MAE loss.
Abnormal peak: the curve manifests an interesting peak at the start before decay. It is likely that this increase of loss at the early stage of training is a consequence of a bad parameter initialization setting. The original author of the research does not use random initialization setting for SINet. Instead, every parameter is naively set to zero. For some reason the symmetry breaks down during training. We stick to this tradition in a good faith of ‘as is’.

Result 2: (a) Input image. One camouflaged object within each image. (b) Camouflaged object Highlight. (c) target identification by model.