Plankton Recognition with Domain Adaptation

Joona Kareinen
CVPRL

March 14, 2025

Motivation

Developing deep learning-based image analysis methods for real-world applications is challenging due to fragmented and heterogeneous datasets. Most image datasets are collected using specific imaging devices under controlled conditions, causing domain shifts when applied to new environments. As a result, models trained on one dataset can perform poorly on others. Domain adaptation (DA) techniques help address this issue by enabling models trained on one domain to generalize across different datasets.

One example application of domain adaptation is plankton recognition, which plays a key role in marine ecology. Plankton are a collection of microscopic organisms that serve as the foundation of aquatic ecosystems. Plankton are broadly classified into two primary groups: phytoplankton and zooplankton. Phytoplankton are primary producers at the base of the aquatic food chain, contributing approximately 50% of the world's oxygen production and around 40% of global carbon fixation [2, 3]. Similarly, zooplankton are critical in matter and nutrient cycling and transport carbon from surface to deep waters through feeding, daily migrations, and fecal pellets, playing a key role in global carbon cycles [1, 5].

Beyond their ecological importance, plankton also serve as indicators of ocean health. Plankton are sensitive to environmental changes: they respond rapidly to changes in temperature, nutrient availability, and water flows [4]. This makes them valuable for understanding the aquatic ecosystem dynamics, which in turn contributes to the prediction of environmental changes such as pollution and climate change [4].

Task description

In this hackathon task, your goal is to develop a deep learning model that can classify plankton species while generalizing to new imaging conditions. You are provided with a dataset containing images of three plankton classes, captured using three different imaging instruments (IFCB, CS, FC). Although all datasets are labeled, during training, you may only use:

- One or two labeled datasets (source domain).
- Unlabeled images from the test dataset (target domain).

This setup reflects a real-world scenario, where older labeled data is available, but newly collected data lacks annotations. You can download the dataset from here: onedrive-link.

Your challenge:

- Step 1: Plankton Recognition (Supervised Learning)
  - Train a model to classify plankton species using the labeled dataset.
  - Establish a baseline accuracy by testing the model on a different dataset (without adaptation).
- Step 2: Domain Adaptation
  - Improve model performance on the unlabeled target dataset by incorporating domain adaptation techniques.

Example approach can be: train a model on the IFCB dataset and test it on the CS dataset $\rightarrow$ This serves as your baseline. Then, train using both IFCB and unlabeled CS data to see if you can improve accuracy.

Hints to get started:

- Utilize pretrained models (e.g., ViT-Base, ResNet-18) using the PyTorch Image Models (timm) package.
- You do not have to use the full dataset for training, but can instead utilize a small subset (e.g., 100 images per class) to make training possible on your laptop.
- You can find more information from the results with full dataset from DA-Plankton paper.

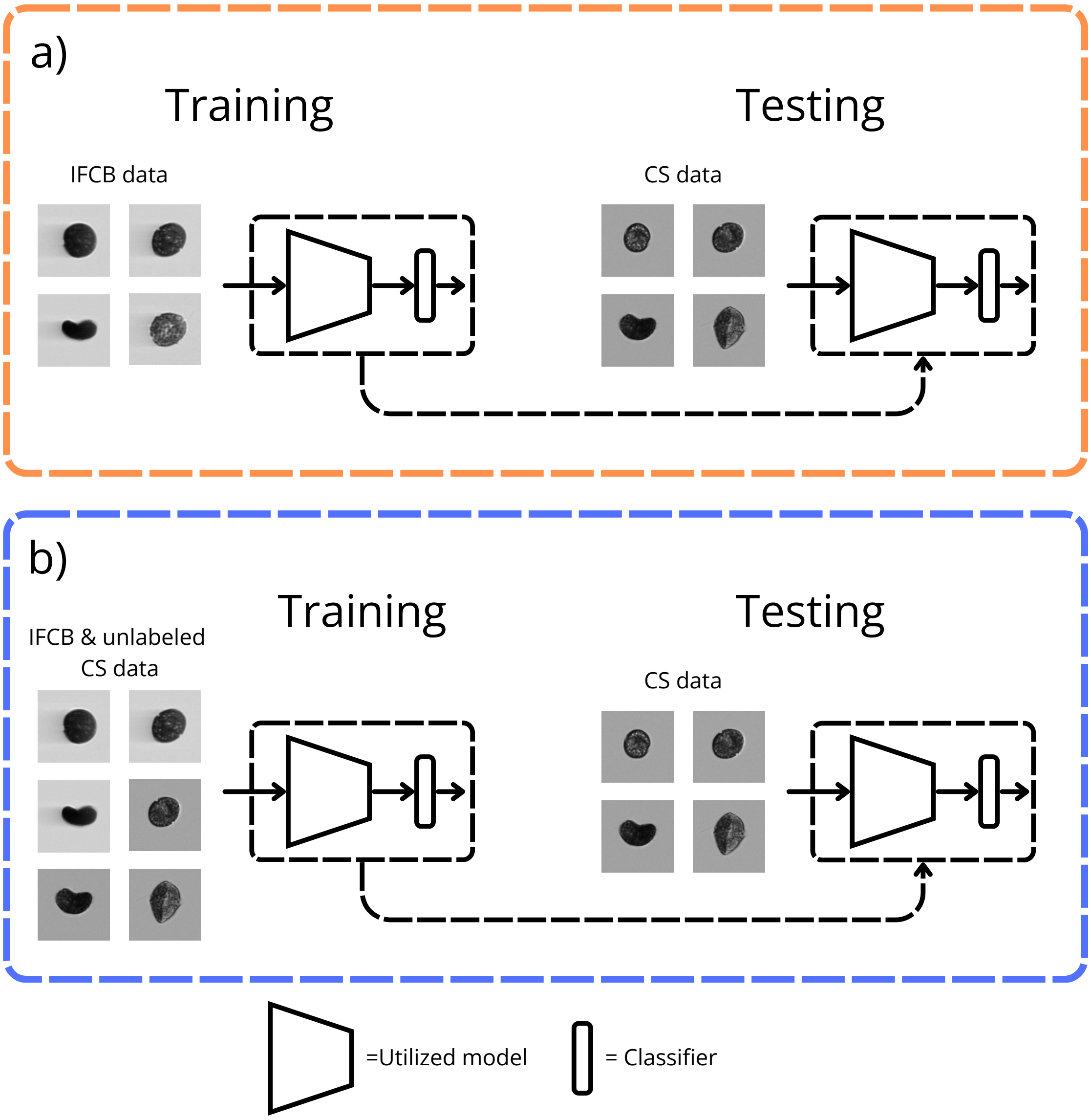

Figure 1: Example pipeline: (a) Supervised training for baseline performance, (b) Training a domain adaptation method using unlabeled target dataset images.

References

[1] Karl Banse. Zooplankton: Pivotal role in the control of ocean production: I. biomass and production. *ICES Journal of Marine Science*, 52:265–277, 1995.

[2] Paul G Falkowski. The role of phytoplankton photosynthesis in global biogeochemical cycles. *Photosynthesis research*, 39:235–258, 1994.

[3] Christopher B Field, Michael J Behrenfeld, James T Randerson, and Paul Falkowski. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, 281:237–240, 1998.

[4] Graeme Hays, Anthony Richardson, and Carol Robinson. Climate change and marine plankton. *Trends in Ecology & Evolution*, 20:337–344, 2005.

[5] Jefferson T. Turner. Zooplankton fecal pellets, marine snow, phytodetritus and the ocean's biological pump. *Progress in Oceanography*, 130:205–248, 2015.