



Gradient-Based Meta Learning for Morphologically Diverse Few-Shot Cell Segmentation

Rohan Sikand¹

¹CS 330 Project, Department of Computer Science, Stanford University

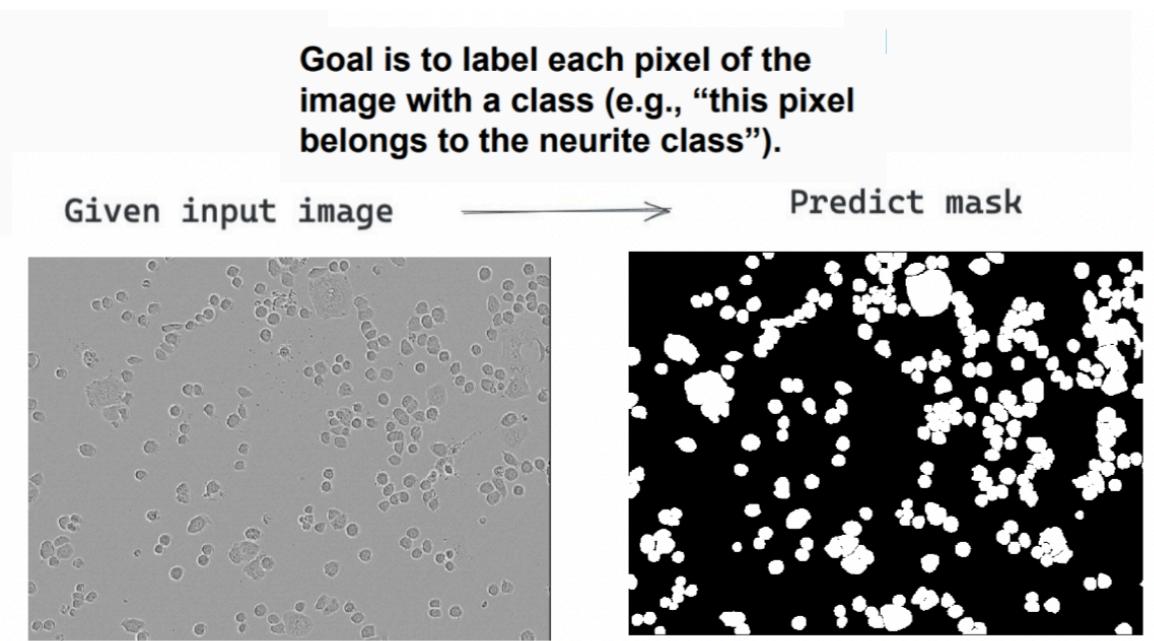
Stanford
Computer Science

Overview

Motivation:

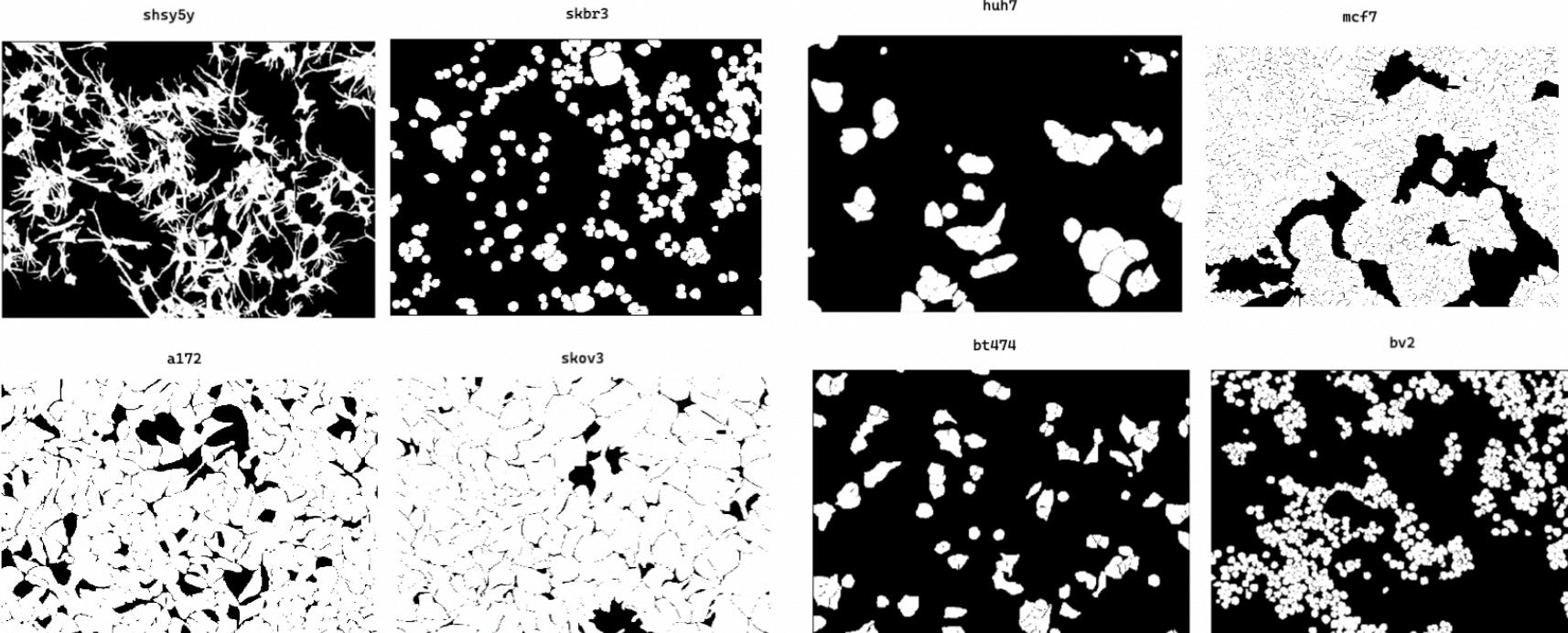
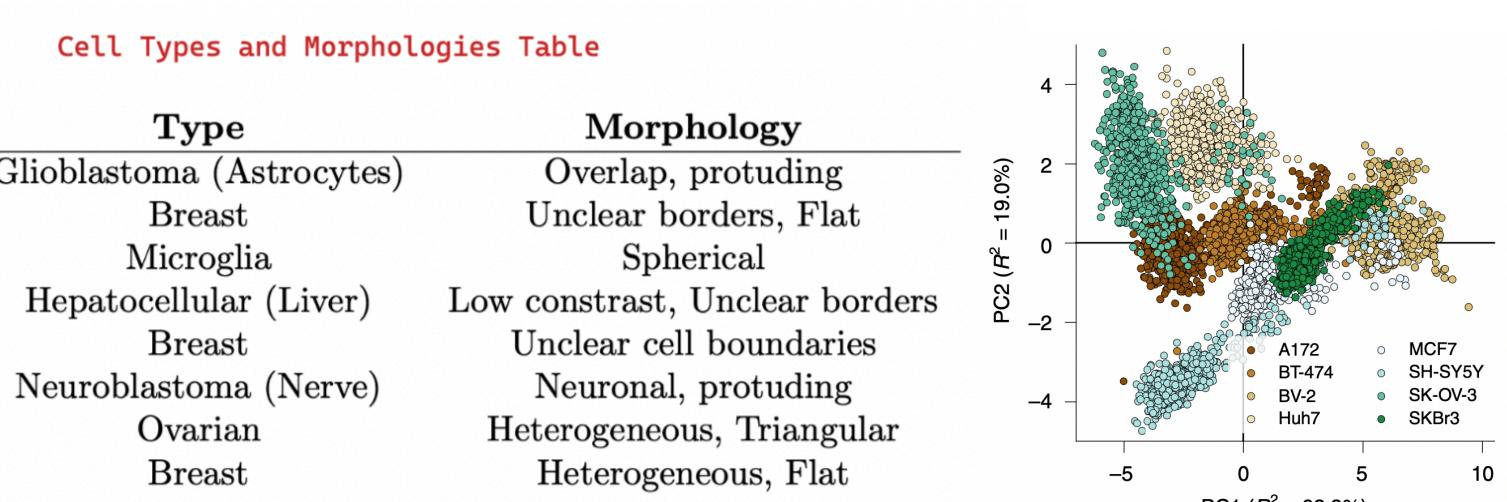
- Segmentation enables **phenotypic analysis** of cell structures: e.g., cell counting, measurements, identifying differences between healthy and diseased cells. Helpful for drug discovery.
- Manually segmenting many microscopic images is tedious** and expensive. Traditional deep learning models can help automate this process but **require large amounts of training data** for each new cell line.
- Here, we introduce a few-shot, meta learning model to segment **new cell types, of various morphologies, given only a few training samples** to resolve these issues.

Segmentation:



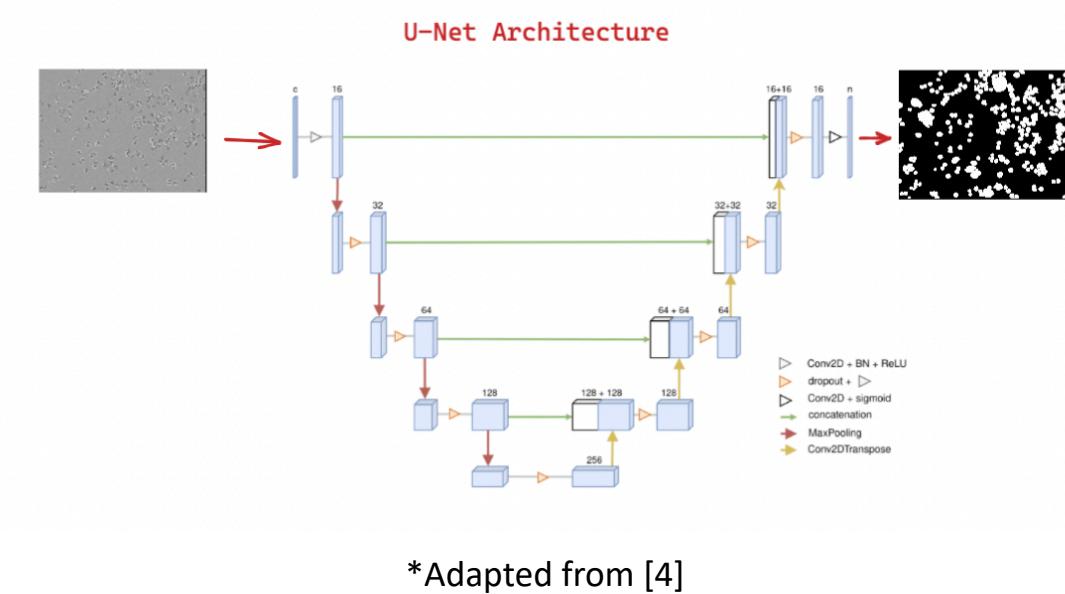
Dataset

- To assemble the tasks, we use the **LIVECell dataset** [2] which consists of quantitative phase contrasted images and masks for 8 different cell lines.
- Each cell type presents a unique morphology** which makes this a challenging problem to solve.



Methods

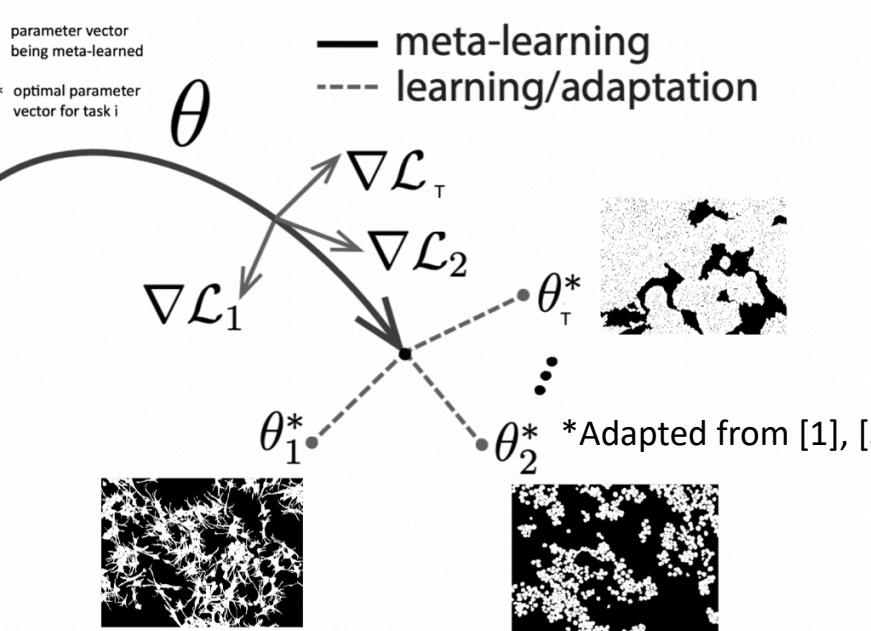
- We implement and adapt **Model-Agnostic Meta Learning (MAML)** for the segmentation task. We represent each cell type as a task.
- Let \mathbb{T} denote a **segmentation task corresponding to some cell type**. The meta-training process is done on $\mathbb{T}_{\text{train}} = \mathbb{T}_1, \dots, \mathbb{T}_n$, with the goal of adapting to \mathbb{T}_{test} more quickly. Each task \mathbb{T} consists of some data $\mathcal{D}_i : \{(x, y)_j\}$ where each $(x, y)_j$ tuple consists of the input cell image (x) and the associated segmentation mask (y). Each task \mathbb{T} is divided into $\{D_{\text{support}}, D_{\text{query}}\}$ where the **model is trained on k support samples (k -shot)** and tested on the query samples.
- For the model f , we use a standard **U-Net** neural network.



```

Algorithm 1 MAML for Few-Shot Cell Segmentation
t: segmentation task for specific cell type
f: model
p(T): distribution over tasks
α, β: inner learning rate, outer learning rate
randomly initialize θ
repeat
    sample batch of tasks τ ~ p(T);
    foreach τ_i ∈ τ do
        initialize φ_i with θ;
        sample {Dsupport, Dquery} ~ p(τi);
        evaluate g = ∇φi Lτi(fφi, Dsupport);
        adapt parameters φi ← φi - αg;
        evaluate test loss Lτi(fφi, Dquery);
    end
    update θ ← θ - β ∑τi ~ p(τ) ∇θ Lτi(fφi, Dquery);
until convergence;
  
```

Goal: meta-learn a parameter vector θ that can be fine-tuned to an optimal parameter vector ϕ_i for T_i using only k support samples and a few gradient steps.

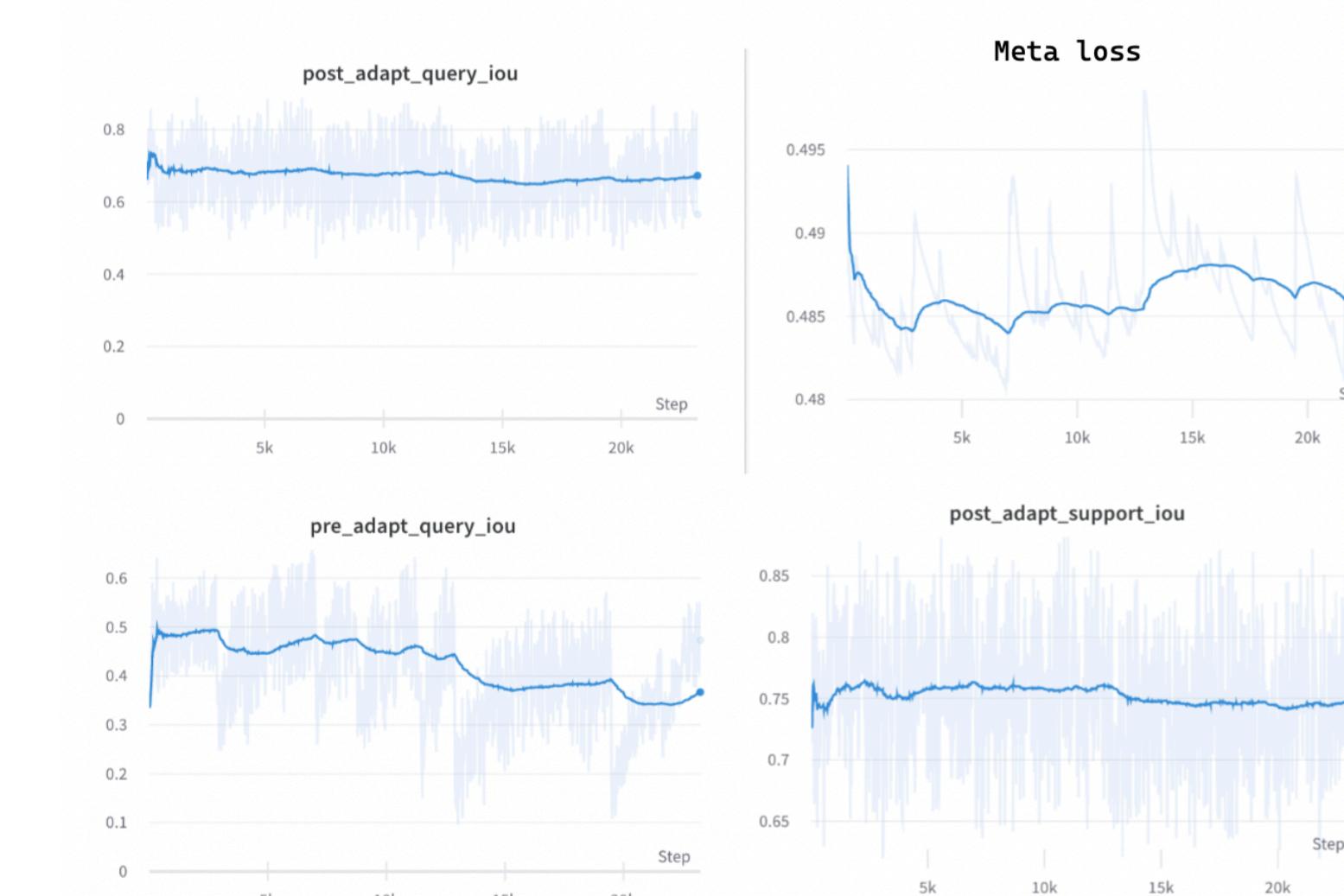


Experimental configuration:

Hyperparameter	Value
k (number of shots)	1, 3
Number of inner steps	10
Number of tasks	7 training, 1 test
Model	U-Net, VGG-11 encoder
Batch size	1, 3
Loss	Dice
Test metric	IoU (Jaccard)
Meta optimizer	Adam (LR = 0.001)
Inner optimizer	Gradient descent (LR = 0.4)

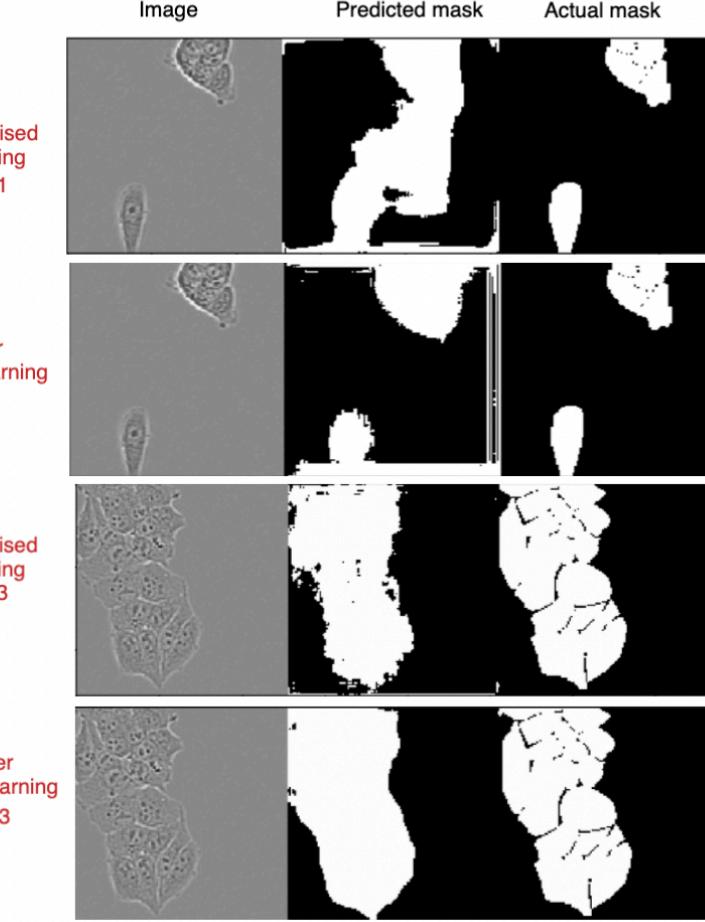
Experiments and Results

- To test our approach outlined in Methods, we conducted **meta-training on 7 of the 8 tasks**. We then tested the model on the 8th, unseen meta-test task as well as on the 7 training tasks.
- Meta-training set: [A172, BT-474, BV-2, Huh7, SH-SY5Y, SkBr3, SK-OV-3], Meta-testing set: [MCF7].
- We also evaluate the performance of a traditional supervised learning network in the few-shot and one-shot setting and use this as a baseline to compare to our meta learning approach.
- We adapted each model for 10 gradient steps after training during testing.
- We tested $k = 1$ and $k = 3$ to evaluate how each model performs in the one-shot setting as well as the few-shot setting.



Experiment	A172	BT-474	BV-2	Huh7	MCF7	SH-SY5Y	SkBr3	SK-OV-3
Supervised, $k = 1$	0.691	0.418	0.438	0.311	0.421	0.271	0.379	0.495
MAML, $k = 1$	0.628	0.418	0.348	0.303	0.526	0.265	0.488	0.843
Supervised, $k = 3$	0.685	0.513	0.288	0.602	0.696	0.565	0.616	0.894
MAML, $k = 3$	0.905	0.676	0.357	0.678	0.881	0.670	0.765	0.889

Predictions:



Conclusions

- The MAML model was able to increase the IoU on the meta test set (MCF7) over the supervised setting for both $k = 1, 3$.
- We hypothesize that, since **segmentation is a difficult task to learn quickly** (i.e., few-shot) due to the dense nature of the prediction, MAML was not able to perform as well as has shown on classification problems, but still performed reasonably well.
- Future work:**
 - Train for more steps (i.e., train longer).
 - Add more tasks for meta-training and/or perform task interpolation.
 - Test different neural network architectures.
 - Increase batch size and compute memory/power.
 - Test other gradient-based meta learning algorithms, in addition to MAML, such as Reptile [5].

Acknowledgements

- Thank you to the CS 330 teaching staff. I am especially grateful for the teaching assistants who were able to assist me during the course of the project in office hours.

References:

- [1] Finn, Chelsea, Pieter Abbeel, and Sergey Levine. "Model-agnostic meta-learning for fast adaptation of deep networks." International conference on machine learning. PMLR, 2017.
- [2] Edlund, Christoffer, et al. "LIVECell—A large-scale dataset for label-free live cell segmentation." Nature methods 18.9 (2021): 1038-1045.
- [3] Rußwurm, Marc, et al. "Meta-learning for few-shot land cover classification." Proceedings of the ieee/cvpr conference on computer vision and pattern recognition workshops. 2020.
- [4] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
- [5] Nichol, Alex, and John Schulman. "Reptile: a scalable metalearning algorithm." arXiv preprint arXiv:1803.02999 2.3 (2018): 4.