Detecting Aquaculture in the Brazilian Amazon using Deep Learning in a Low-Data Setting

Joshua Fan (jyf6@cornell.edu)[1], Laura Greenstreet[1], Felipe Pacheco[1], Yiwei Bai[1], Marta Eichemberger Ummus[2], Carolina Rodrigues da Costa Doria[3], Nathan Oliveira Barros[4], Bruce Forsberg[5], Jucilene Cavali[2], Xiangtao Xu[1], Alexander Flecker[1], and Carla Gomes[1]

(1) Cornell University, Ithaca, United States, (2) EMBRAPA Brazilian Agricultural Research Corporation, Campinas, Brazil, (3) Federal University of Rondonia, Porto Velho, Brazil, (4) Federal University of Juiz de Fora, Juiz De Fora, Brazil, (5) National Institute of Amazonian Research, Manaus, Brazil

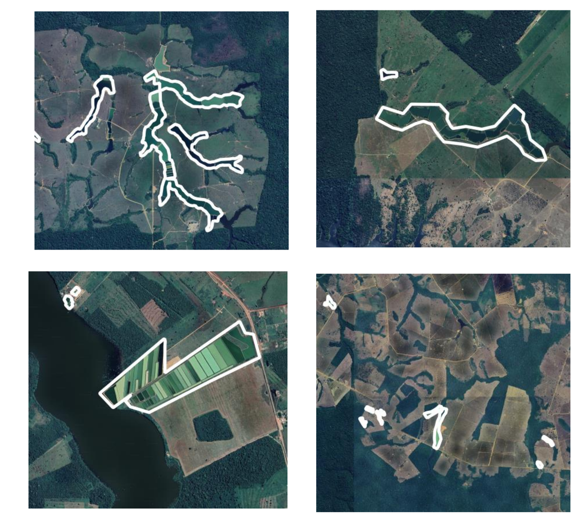
Paper GC21F-0969

Motivation

- Aquaculture is the farming or cultivation of aquatic organisms, including fish, mollusks, and crustaceans
- Aquaculture has the potential **to produce high-quality animal protein** with **lower environmental impacts** than traditional livestock like cattle
- However, adverse impacts such as freshwater use, eutrophication, and greenhouse gas emissions can vary by over two orders of magnitude based on factors such as species farmed, land-use change, and production practices [1]
- We currently lack basic information on the **number** and **area** of aquaculture ponds and their **locations** in regions such as the Amazon Basin
- Therefore, detailed spatial information is needed to understand impacts of aquaculture across the Amazon basin

Problem Formulation

- **Detecting aquaculture from remote sensing data offers the possibility of scalable, low-cost mapping**
- We want to segment and classify waterbodies as aquaculture or non-aquaculture
- Random forests can reliably segment waterbodies, making the problem primarily a **classification** task

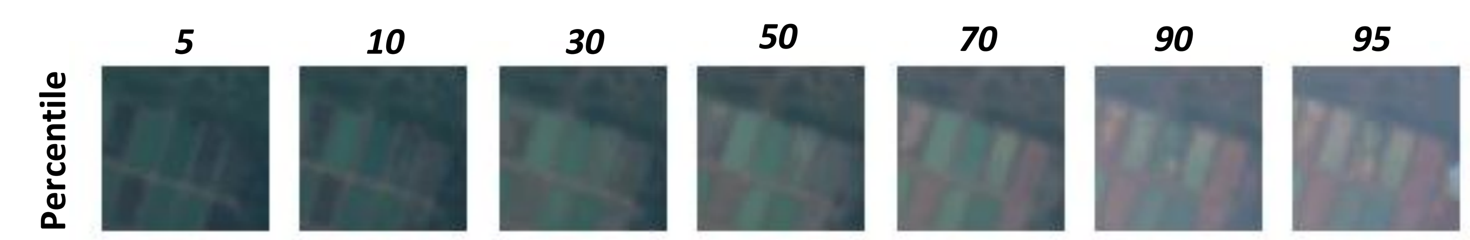


Example fish farms in Rondônia state

Can we train a model that performs well across the Amazon using a small number of examples from a few regions?

Data

- The Brazilian Agricultural Research Corporation, Embrapa, provided labels for over 2300 aquaculture operations in the Brazilian states of Rondônia and Amazonas, representing roughly 6000 waterbodies
- The National Water and Sanitation Agency (ANA) provided roughly 3000 negative examples (non-pond waterbodies) across both states
- We used Sentinel-1 and 2 data (14 spectral bands), taking images at 7 percentiles to capture temporal change, and processed the data into 32 x 32 tiles
- We trained the model on three subregions of Rondônia, each with 300-400 training examples, and tested the model's ability to generalize in both states

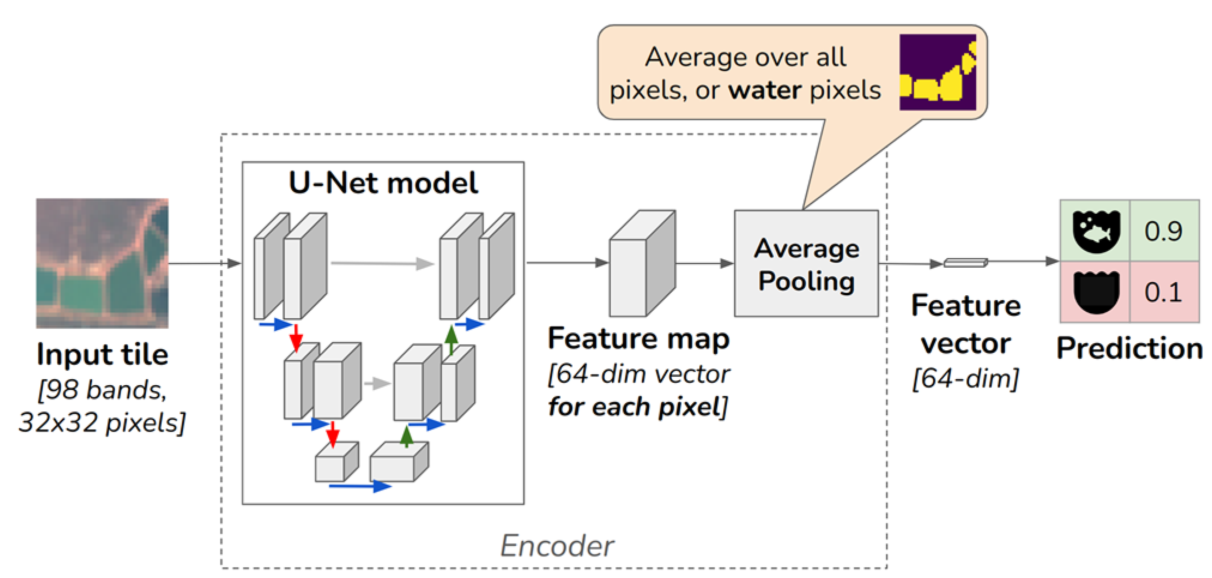


Deep Learning Methods

- Deep learning allows us to **learn** to extract predictive features from entire images, removing the need for feature engineering
- However, deep learning models often need large quantities of data to prevent overfitting
- **Supervised** models use **labeled** images to learn predictive features
- **Self-supervised** models use **unlabeled** images to learn features that separate images. A second classifier is then trained on these features to separate labeled images.
- Unsupervised models can make use of large quantities of unlabeled data

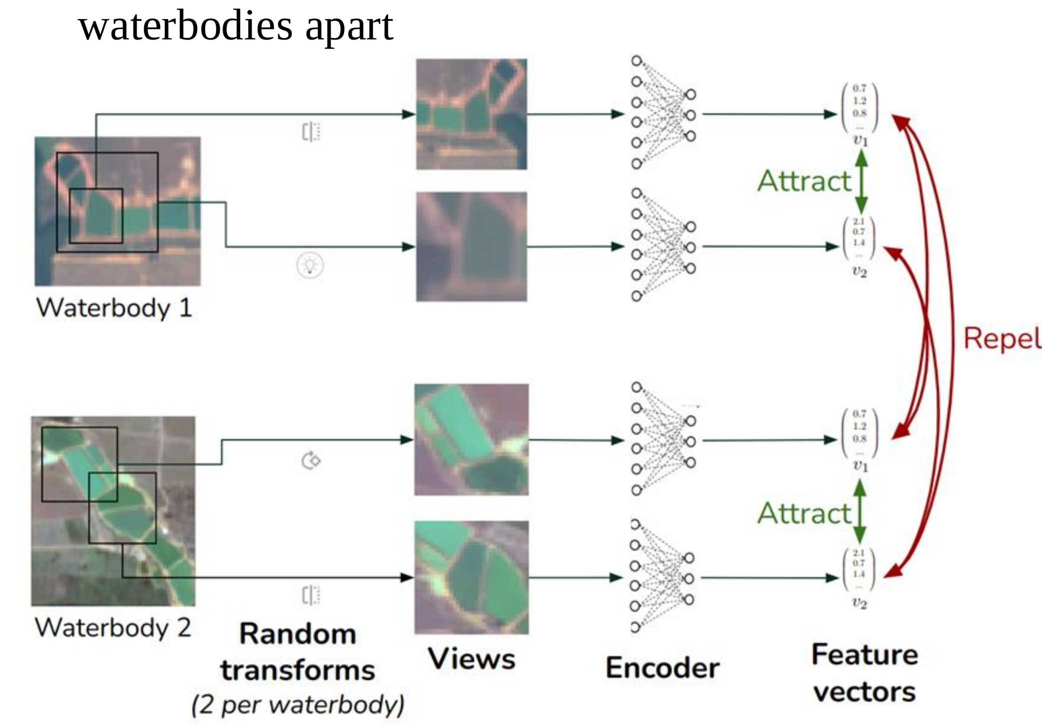
CNN/U-Net

- Supervised deep learning model
- Generates features for each pixel, uses **average pooling** to produce features for the full tile, then predicts waterbody type
- **Masked pooling:** only average over known water pixels to minimize overfitting on landscape features



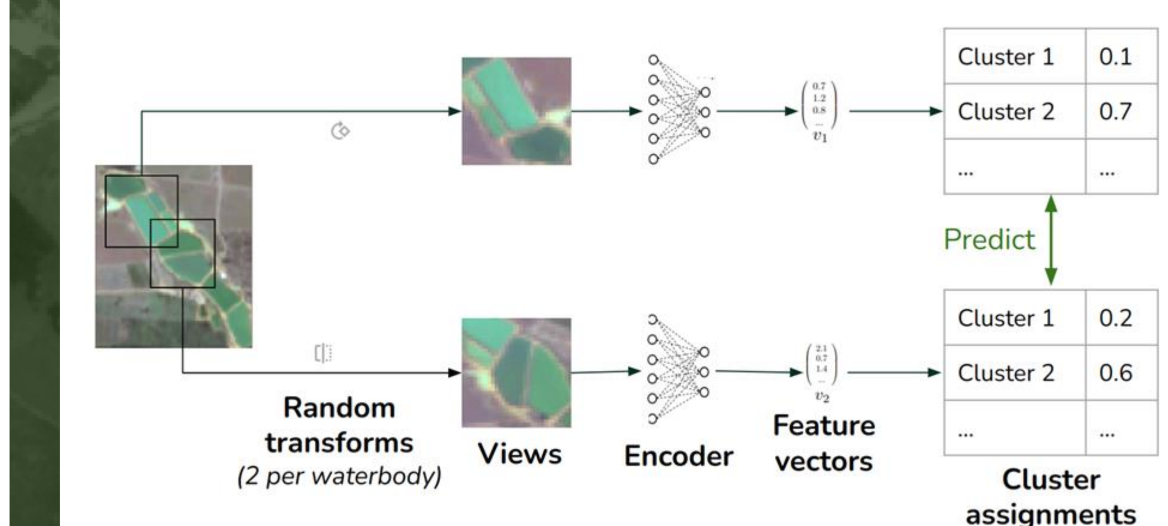
SimCLR

- Self-supervised contrastive deep learning model
- For each waterbody, generate two **views** by performing random transformations
- Learns representations by pushing views from the same waterbody closer and views from different waterbodies apart



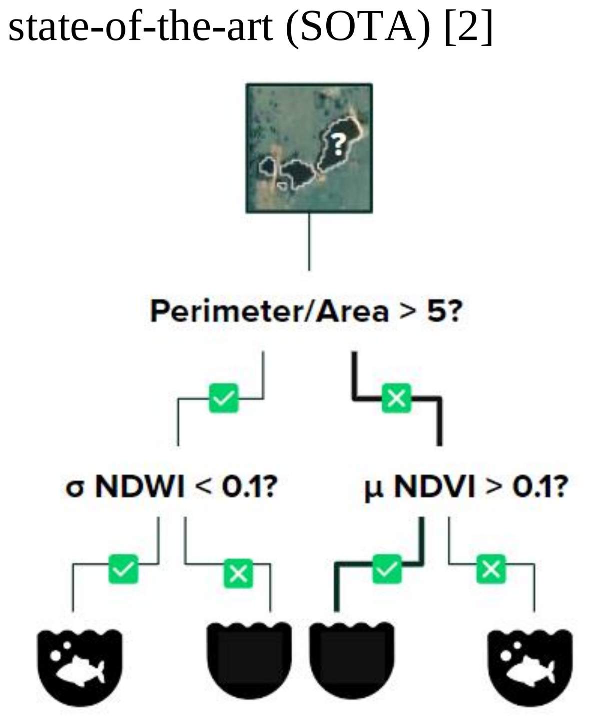
SWAV

- Self-supervised contrastive deep learning model
- Similar to SimCLR, except that images from different waterbodies are not necessarily pushed apart
- Instead, views from the same waterbody should be assigned to the same cluster



Baseline – Random Forests

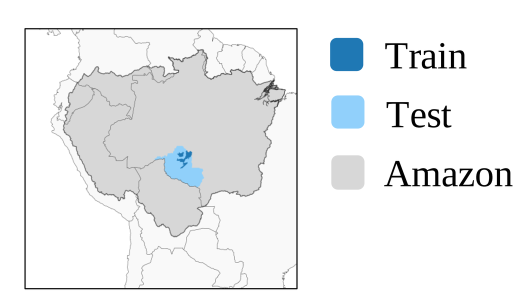
- Hand-engineer features, such as perimeter or mean NDWI, from NDWI mask of waterbody
- Then train a random forest to predict waterbody type
- Prior state-of-the-art (SOTA) [2]



Results

Rondônia → Rondônia (different parts)

	Random Forest	CNN	CNN + Masked Pooling	SimCLR[1]	SWAV[2]
F1	0.84 ± 0.01[3]	0.95 ± 0.01	0.95 ± 0.01	0.96 ± 0.01	**0.97 ± 0.00**
Accuracy	0.77 ± 0.01	0.93 ± 0.01	0.93 ± 0.02	0.94 ± 0.02	**0.96 ± 0.00**



Rondônia → Amazonas

	Random forest	CNN	CNN + Masked Pooling	SimCLR[1]	SWAV[2]
F1	0.35 ± 0.30	0.69 ± 0.04	0.87 ± 0.02	0.90 ± 0.03	**0.93 ± 0.00**
Accuracy	0.30 ± 0.29	0.78 ± 0.02	0.86 ± 0.03	0.89 ± 0.04	**0.93 ± 0.00**

[1]SimCLR results are with fine-tuning on labeled examples
[2]SWAV results are without fine-tuning, training a linear classifier on fixed features
[3]Format is mean ± standard deviation over 5 random seeds. For SimCLR and SWAV, seeds apply to the classifier.

See workshop paper for further results ➤

Conclusions

- Previous SOTA performed poorly in the Amazon basin
- Deep learning can achieve 90%+ accuracy only using hundreds of labeled examples, but generalization to new regions remains challenging
- Using contrastive learning to utilize large quantities of unlabeled data significantly improves generalization performance
- However, contrastive models are harder to train and sensitive to hyperparameters, initialization, and image transformations
- Future directions: investigate dataset **sampling bias** - we may be missing many challenging negative examples (e.g. small natural ponds)

References

[1] Jessica A Gephart, Patrik JG Henriksson,, et al. 2021. Environmental performance of blue foods. *Nature* 597, 7876 (2021), 360–365.
[2] Zilong Xia, Xiaona Guo, and Ruishan Chen. "Automatic extraction of aquaculture ponds based on Google Earth Engine." *Ocean & Coastal Management* 198 (2020): 105348.