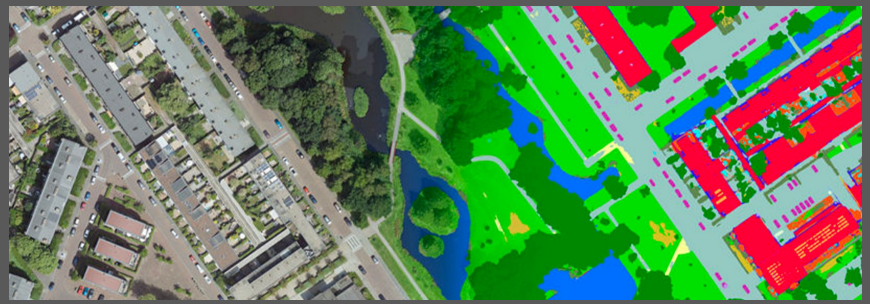
The Alan Turing Institute

Remote Sensing reading group



Conference round-up

18 July 2024

Outline

- Thijs van der Plas | CVPR (Earthvision)
- Anna Zanchetta & Barbara Metzler | ML4EO

https://arxiv.org/pdf/2404.06637

GeoSynth

Stable diffusion (tuned). Generate HR sat given OSM + gen captions + location. p(s | m, c, l)

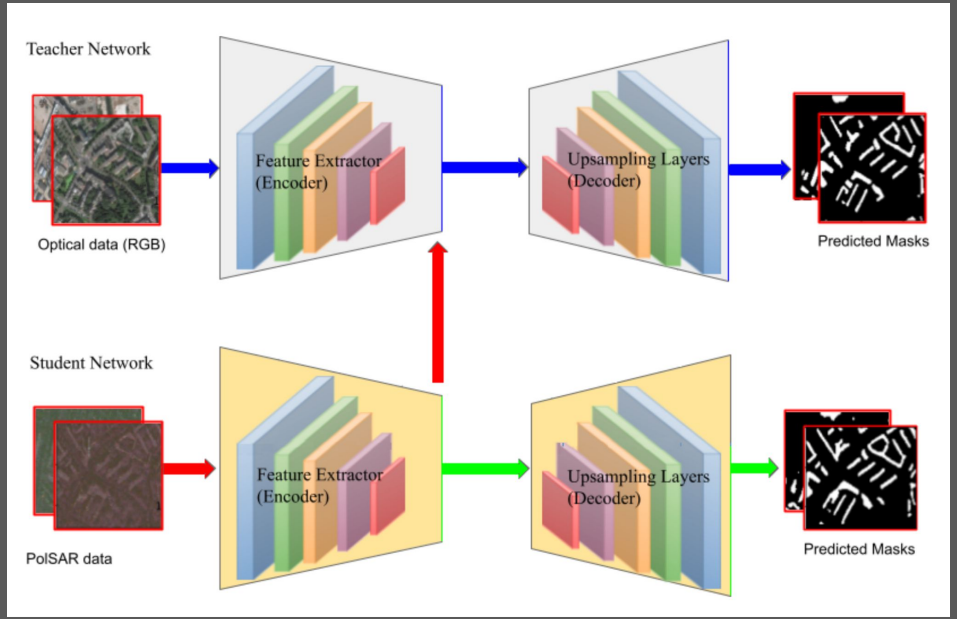
Encoders for im, text, loc. Diffusion in latent space.


Control
Seattle, Washington
San Diego, California
Des Moines, Iowa
New York City, New York
Boulder, Colorado
Dallas, Texas

Cross model Knowledge Distillation

1. Train optical supervised
2. KD SAR (student) and optical (teacher)
3. Fine tune SAR supervised

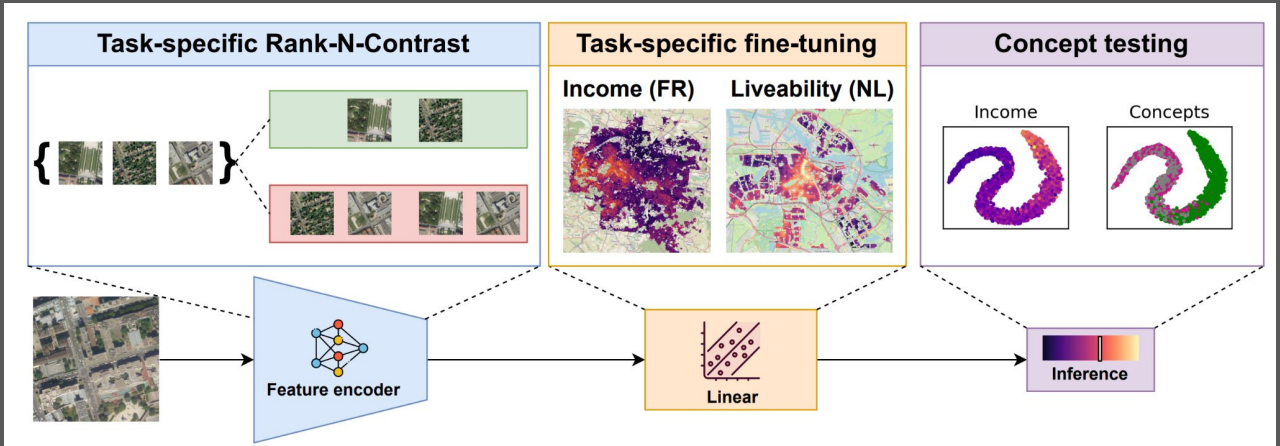
Improves SAR model (but does not exceed optical model)

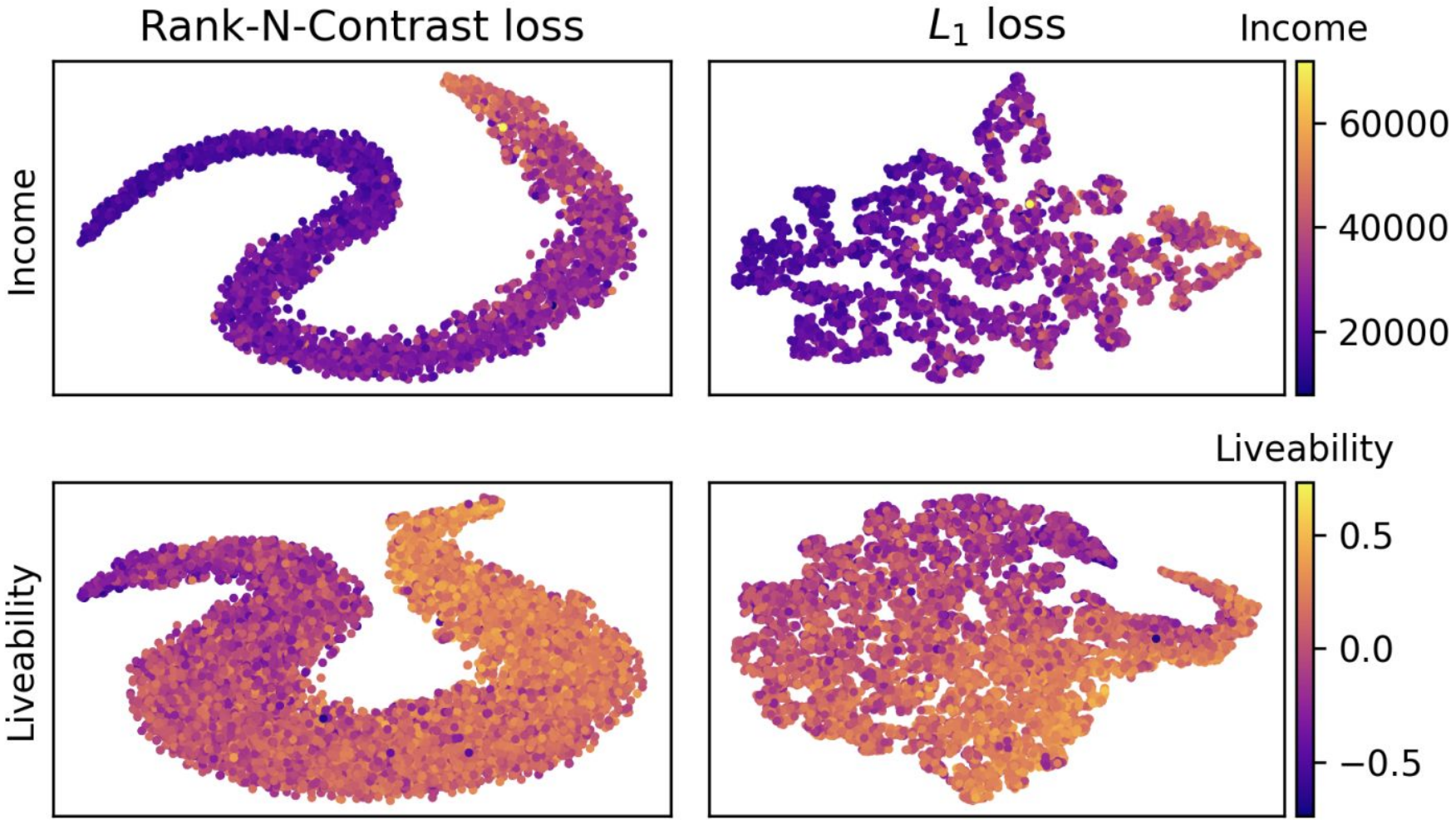


Contrastive pretraining for explainability

Rank-N-contrast (Neurips 2023)

1. Train Resnet with CL. 2. Freeze and train final layer on task.

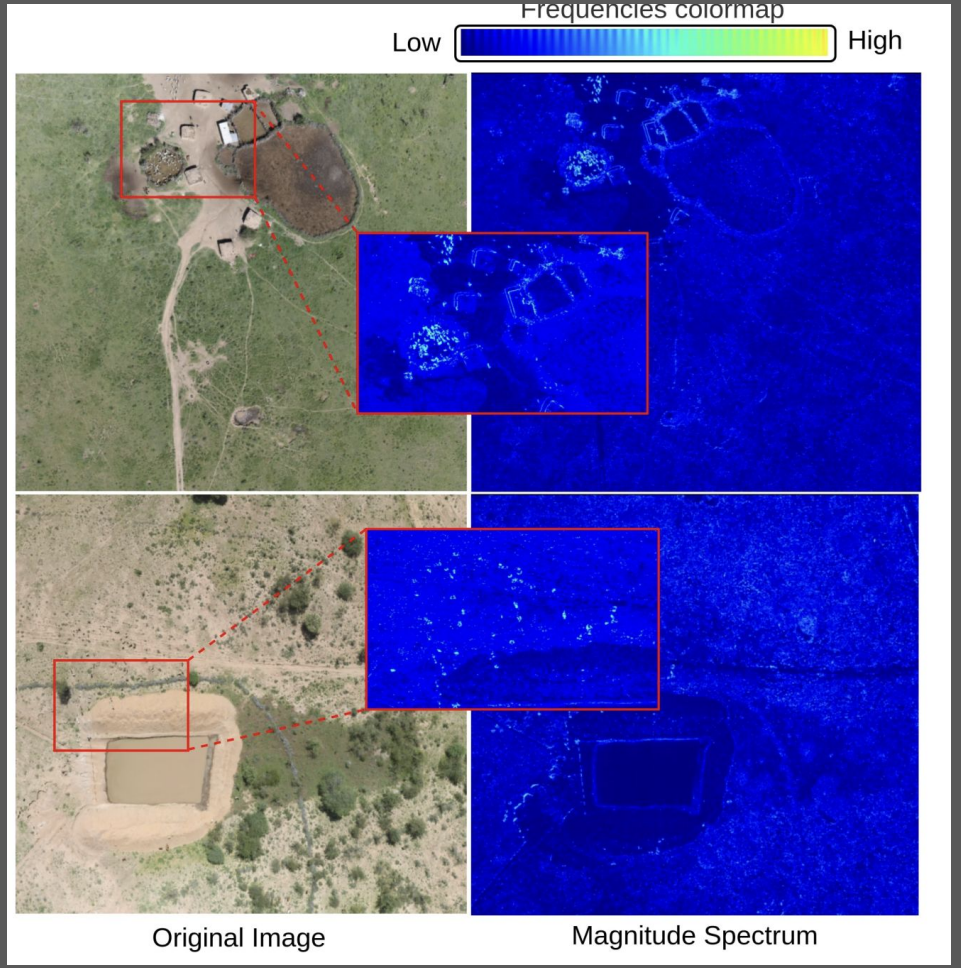


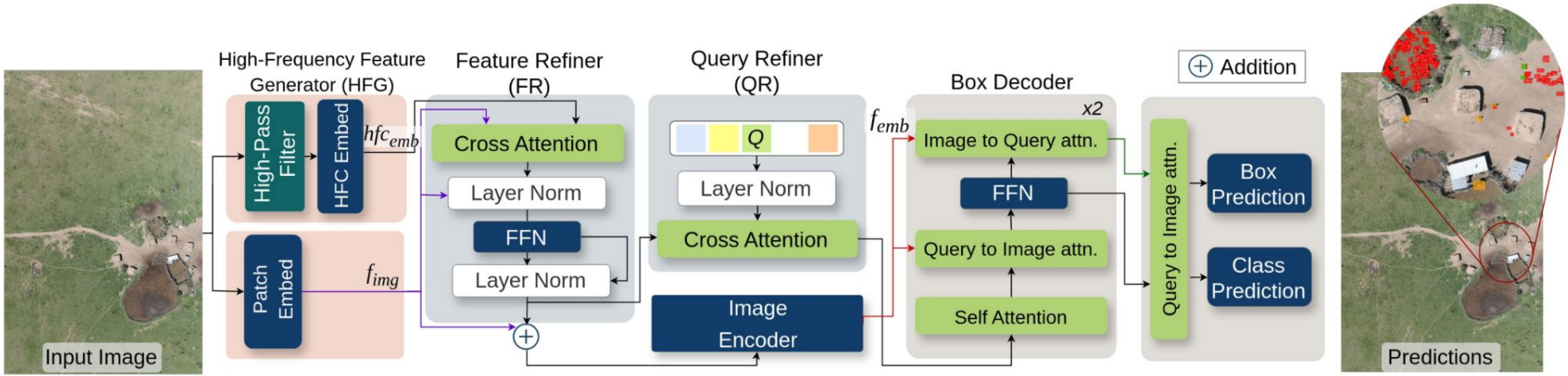

Rank-N-Contrast loss
L_1 loss
Income
60000
40000
20000
Income
Liveability
0.5
0.0
−0.5
Liveability

WildlifeMapper

Object detection (large mammals) -> very small

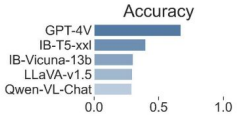
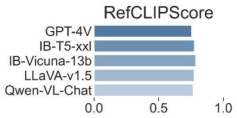
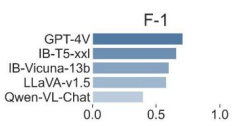
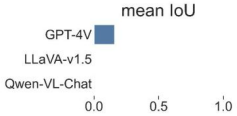
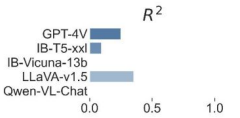
Generate high freq map via FT, use with high pass filter




Input Image
High-Frequency Feature Generator (HFG)
High-Pass Filter
HFC Embed
hfc_{emb}
Patch Embed
f_{img}
Feature Refiner (FR)
Cross Attention
Layer Norm
FFN
Layer Norm
Query Refiner (QR)
Q
Layer Norm
Cross Attention
Image Encoder
f_{emb}
Box Decoder
x2
Image to Query attn.
FFN
Query to Image attn.
Self Attention
Query to Image attn.
Addition
Box Prediction
Class Prediction
Predictions

Foundation models

[Benchmarking GPT4]

Scenario	Example	Results
Location Recognition [2.1]	👤: Make an educated guess about the name of the landmark shown in the image. 🤖: Based on the style, the dome, and layout of the grounds, my choice is: A. Nebraska State Capitol 	
Image Captioning [2.2]	👤: Generate a caption for the image in one sentence. 🤖: Aerial view of an airport terminal with adjacent aircraft, taxiways, and parking areas. 	
Land Use & Land Cover Classification [2.3]	👤: Classify the image into one of the following options. 🤖: The best option to describe the given image is: 28. Shipping Yard 	
Object Localization [3.1]	👤: Identify the extent of the object in the description below. Description: The gray windmill in the middle 🤖: [233, 383, 376, 542] 	
Object Counting [3.2]	👤: Count the number of trees in the given image to the best of your ability. 🤖: 134 	
Change Detection [4]	👤: Count the number of buildings in each damage category in JSON format. 🤖: ```json {"count_before": 75, "no_damage": 2, "minor_damage": 73, "major_damage": 0, "destroyed": 0} ```