Change-Aware Sampling and Contrastive Learning for Satellite Images

Utkarsh Mall   Bharath Hariharan   Kavita Bala
Cornell University

Paper Tag: TUE-PM-107

Cornell Bowers C·IS
College of Computing and Information Science

CVPR 2023
VANCOUVER, CANADA

Preview

Large-scale Challenges

Disaster monitoring

Climate change

Spatio-temporal satellite images

5 Terapixel information every week

Change-aware Contrastive Learning (CACo)

- Temporal information
- Change awareness
- Geographical information

Self-supervised representation

	EuroSat (Acc.)		BigEarthNet (mAP)		OSCD (F1)	Dynamic EarthNet (mIoU)
Pre-training (100k)	**ResNet-18**	**Resnet-50**	**ResNet-18**	**Resnet-50**	**ResNet-18**	**Resnet-18**
Random Init.	64.21	55.32	45.95	45.22	28.91	41.53
ImageNet	86.16	89.08	66.40	71.37	35.30	43.75
Moco v2	87.22	89.75	67.20	72.88	38.21	47.97
GSSL	87.74	90.19	67.36	72.86	44.06	46.77
SeCo	90.05	93.12	67.43	73.42	46.84	46.83
CACo	**93.08**	**94.48**	**69.43**	**73.63**	**50.29**	**50.20**

Change-Aware Sampling and Contrastive Learning for Satellite Images

Utkarsh Mall Bharath Hariharan Kavita Bala
Cornell University

Cornell Bowers C·IS
College of Computing and Information Science

CVPR 2023
VANCOUVER, CANADA

Introduction

Change Detection

Event Retrieval

Land Cover Classification

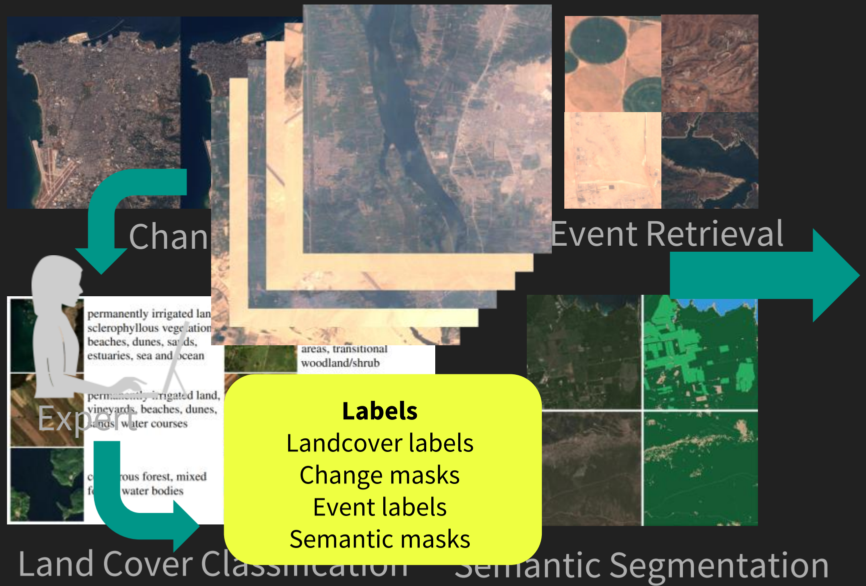
Semantic Segmentation

Climate change

Disaster monitoring

Applying computer vision to satellite images.

Supervision is expensive



Training these tasks require labeled data.

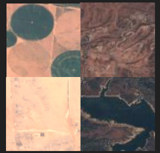
Labeling is even harder as labelling satellite images require experts.

Using self-supervision

Labels
Landcover labels
Change masks
Event labels
Semantic masks

Self-supervised representation

Change Detection



Event Retrieval

Land Cover Classification

Semantic Segmentation

With self-supervision we can learn a representation without labels.
Representation can used for downstream tasks with very few labels.

Change-Aware Constrastive (CACo) Learning

Goal: how can we best leverage the unique structure of satellite images for better self-supervised learning?

Temporal information

April 2016 July 2016 April 2022

Long-term temporal signal for self-supervision

Change Awareness

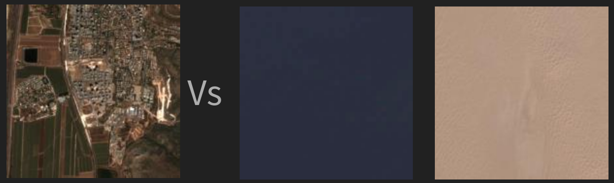
Change

No Change

Time difference does not imply change

Geographical Information

Informative Less Informative



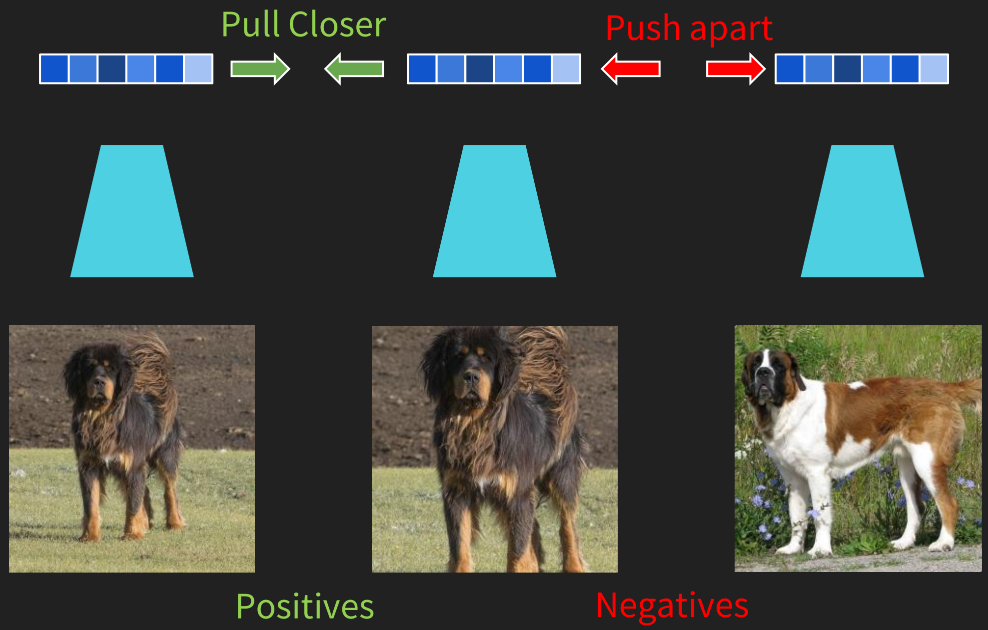
Some locations are more informative

NPID: Unsupervised Learning via Non-Parametric Instance Discrimination, CVPR 2018, Wu *et. al.*
MoCo: Momentum contrast for unsupervised visual representation learning, CVPR 2020, He *et. al.*
PIRL: Self-Supervised Learning of Pretext-Invariant Representations CVPR 2020, Misra *et. al.*
SimCLR: A simple framework for contrastive learning of visual representations, ICML 2020, Chen *et. al.*

Method

Contrastive Learning

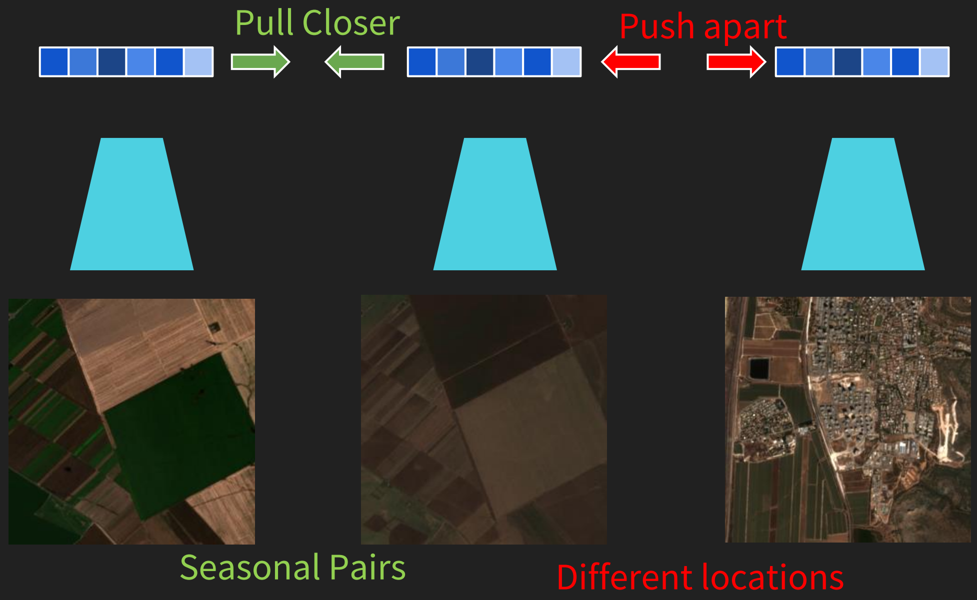
Learning embedding by using negative and positives examples.



NPID: Unsupervised Learning via Non-Parametric Instance Discrimination, CVPR 2018, Wu *et. al.*
MoCo: Momentum contrast for unsupervised visual representation learning, CVPR 2020, He *et. al.*
PIRL: Self-Supervised Learning of Pretext-Invariant Representations CVPR 2020, Misra *et. al.*
SimCLR: A simple framework for contrastive learning of visual representations, ICML 2020, Chen *et. al.*

Seasonal Contrast (SeCo)

Seco uses short-term temporal pairs as positives.

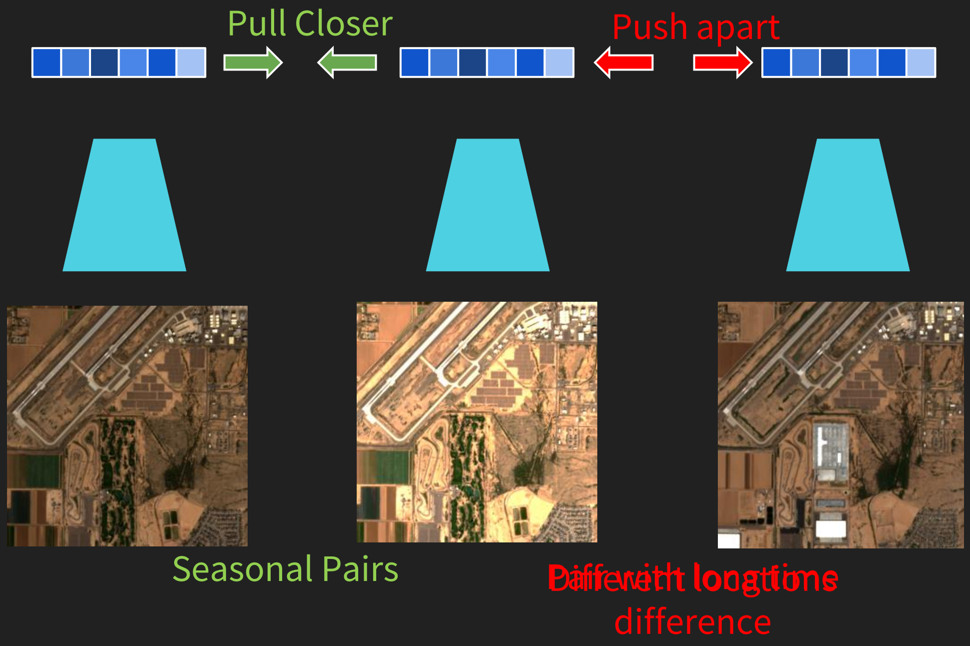


Mañas *et. al.*, Seasonal contrast: Unsupervised pre-training from uncurated remote sensing data. ICCV, 2021

Long-term Temporal Contrast

Short-term temporal difference: model seasonal variation and should be positive.

Long-term temporal difference: model actual changes and should be negative.

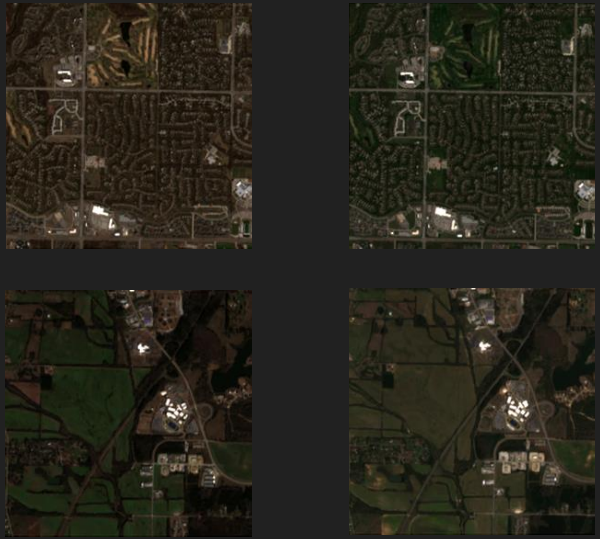


Change-awareness

Not all locations are equally likely to change in long-term.

Long-term pairs with changes

Long-term pairs with no changes

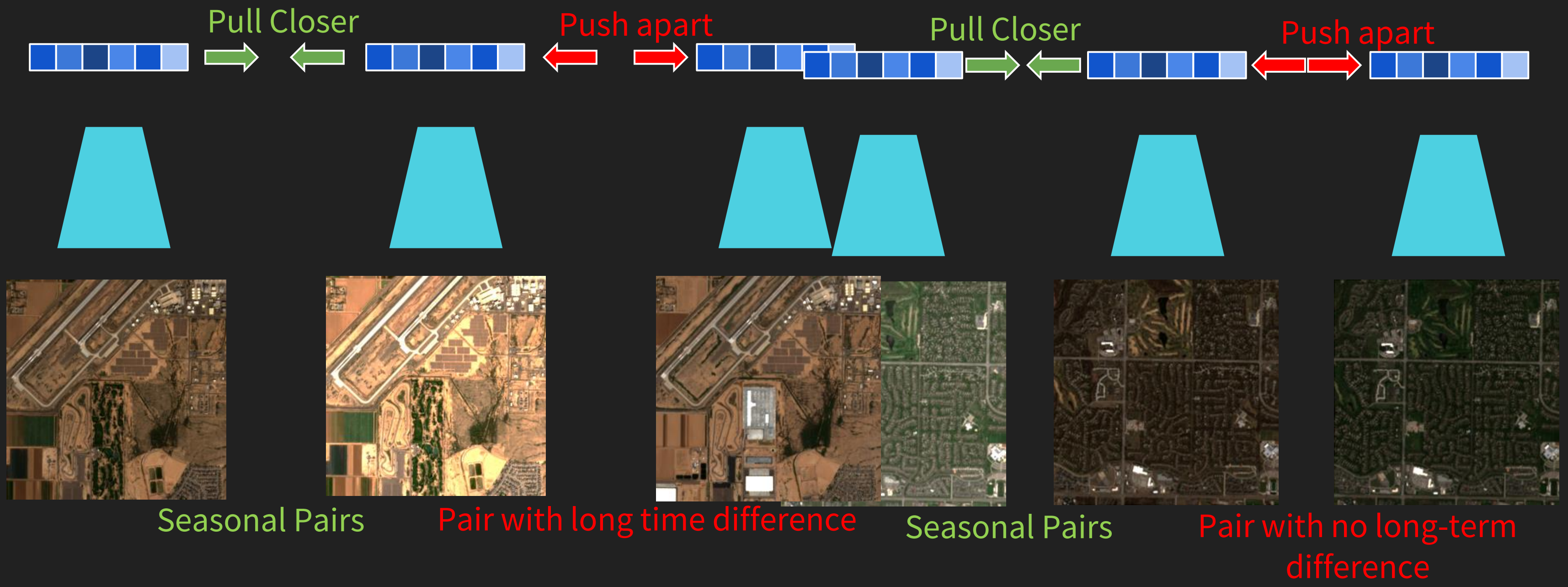


Some locations can change drastically in a few years.

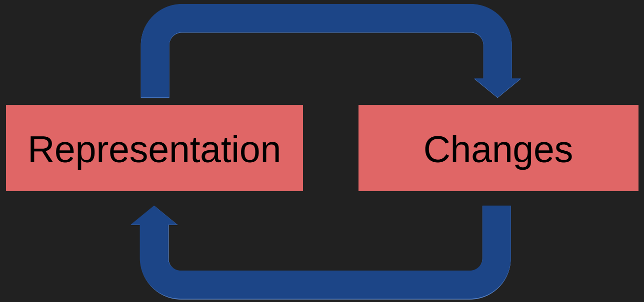
Whereas others do not.

Using Change Awareness

Push apart long-term temporal pairs if and only if there is a change. (CACo)



Finding Change using Representation

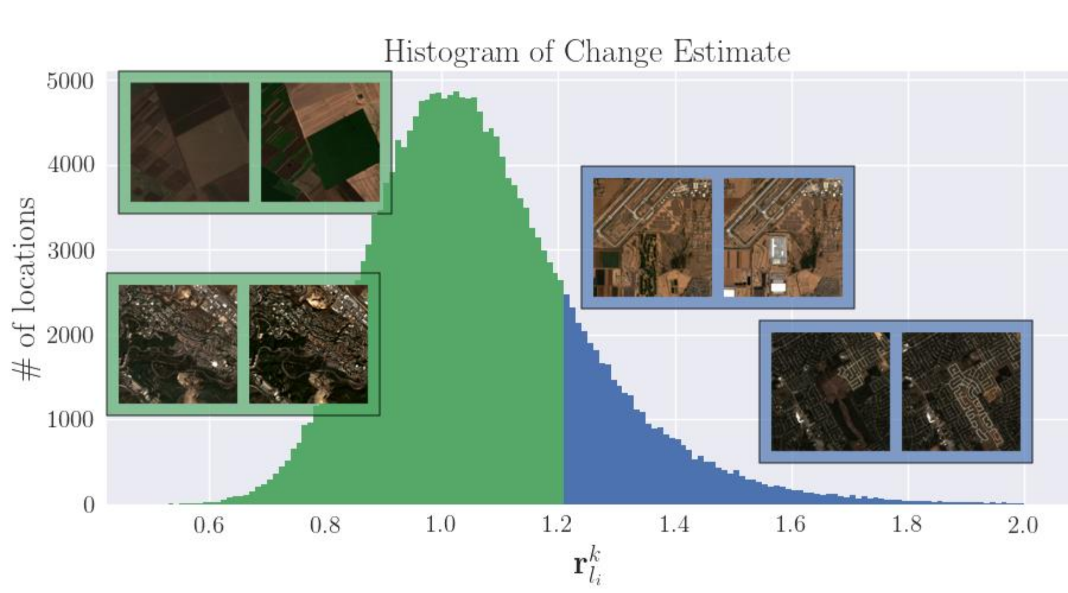


Bootstrapping representation

- Use representation to find change estimates
- Use changes to improve representation
- Repeat…

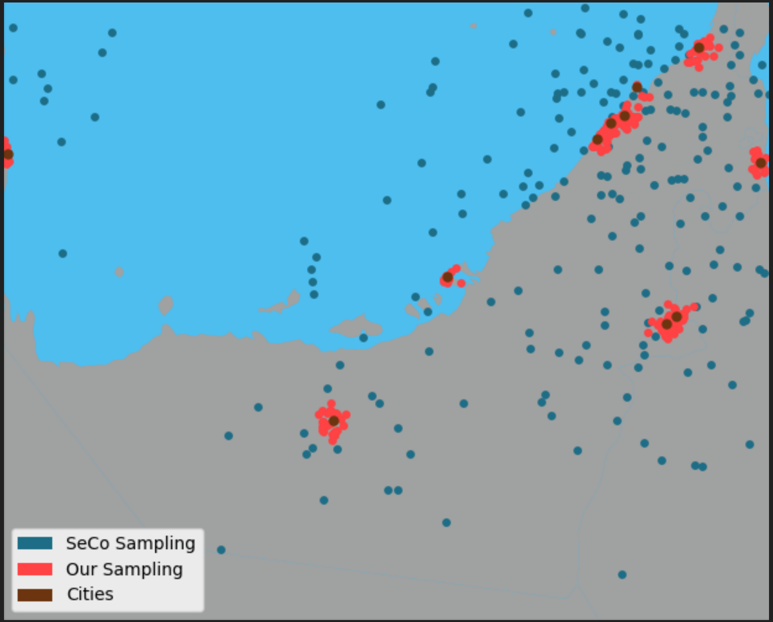
Temporal information | **Change Awareness** | Geographical Information

Change Estimate



Geographic sampling

Low information Samples



New Sampling Strategy:

- Sample using a stronger (σ= 5 km) Gaussian sampler around urban areas .
- Reject and resample if sample falls in ocean.

Results

Experimental setup

Using MoCo V2 Framework

Models:

- ResNet-18
- ResNet-50

Self-supervision Dataset

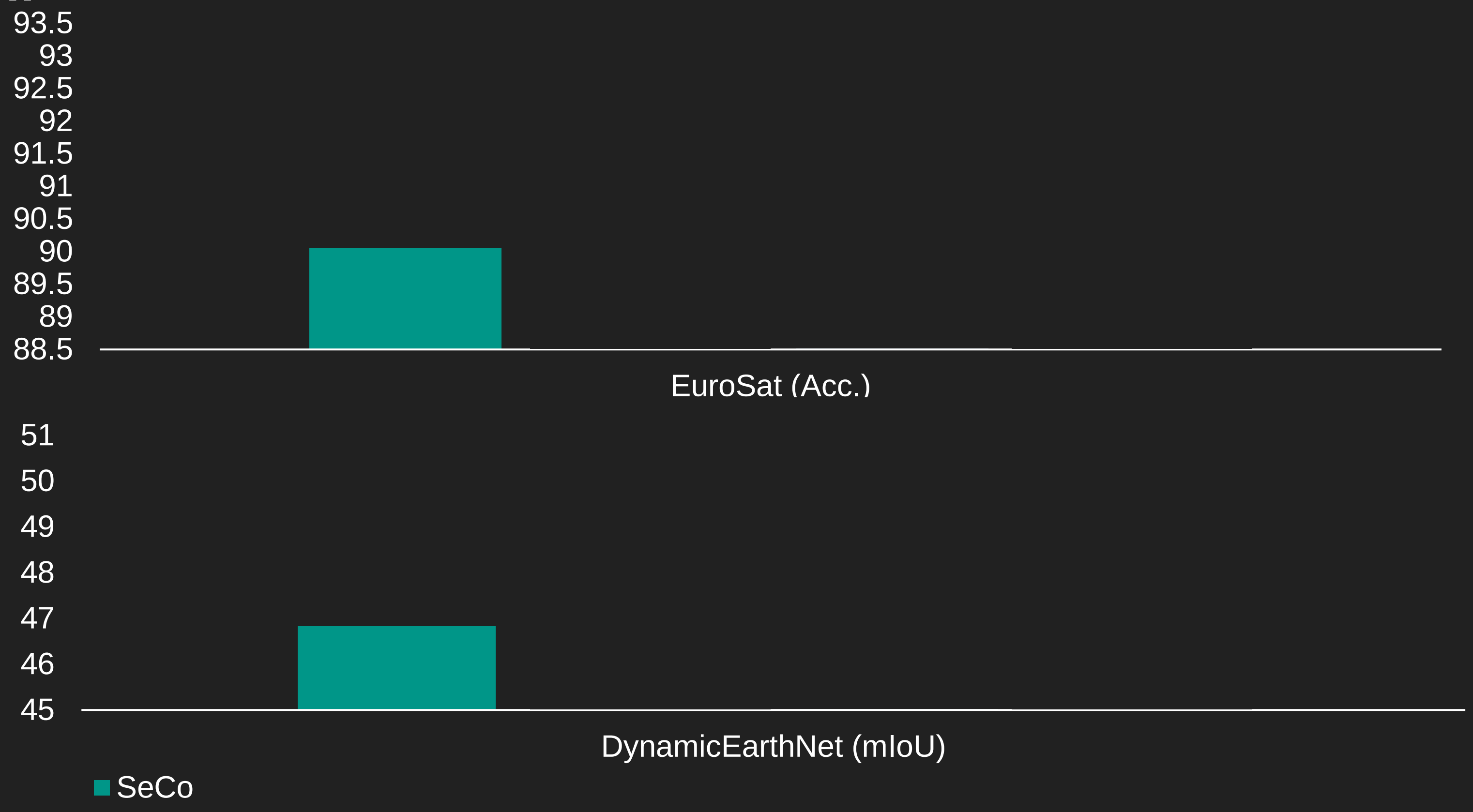
- 100k
- 1 Million

Downstream Tasks

Task	Benchmark	Metric
Landcover Classification	EuroSat	Accuracy
	BigEarthNet	mAP
Change Detection	OSCD	F1-Score
Semantic Segmentation	Dynamic EarthNet	mIoU
Event Retrieval	CaiRoad	AP@K
	CalFire	

MoCo V2:Improved Baselines with Momentum Contrastive Learning, CoRR 2020, He *et. al.*

Ablation


93.5
93
92.5
92
91.5
91
90.5
90
89.5
89
88.5
EuroSat (Acc.)
51
50
49
48
47
46
45
DynamicEarthNet (mIoU)
SeCo

Comparison to baselines

	EuroSat (Acc.)		BigEarthNet (mAP)		OSCD (F1)	Dynamic EarthNet (mIoU)
Pre-training (100k)	**ResNet-18**	**Resnet-50**	**ResNet-18**	**Resnet-50**	**ResNet-18**	**Resnet-18**
Random Init.	64.21	55.32	45.95	45.22	28.91	41.53
ImageNet	86.16	89.08	66.40	71.37	35.30	43.75
Moco v2	87.22	89.75	67.20	72.88	38.21	47.97
GSSL	87.74	90.19	67.36	72.86	44.06	46.77
SeCo	90.05	93.12	67.43	73.42	46.84	46.83
CACo	**93.08**	**94.48**	**69.43**	**73.63**	**50.29**	**50.20**

Conclusion

We present a novel self-supervised approach for contrastive learning on satellite images, leveraging three properties unique to them.

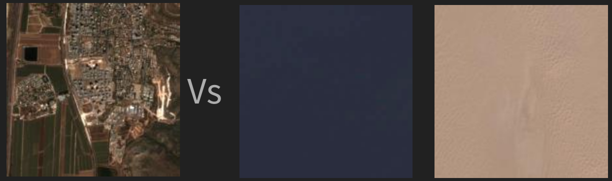
Temporal information

Long-term temporal signal for self-supervision

Change Awareness

Time difference does not imply change

Geographical Information



Some locations are more informative

Thank You!

https://research.cs.cornell.edu/caco/

Dataset and Code

Cornell Bowers C·IS
College of Computing and Information Science

CVPR 2023
VANCOUVER, CANADA