Geography-Aware Self-Supervised Learning

Kumar Ayush*, Burak Uzkent*, Chenlin Meng*, Kumar Tanmay, Marshall Burke, David Lobell, Stefano Ermon

2021 ICCV OCTOBER 11-17 VIRTUAL

Introduction

Self-supervised learning methods such as MoCo-v2 have shown promising results in capturing **good** representations.

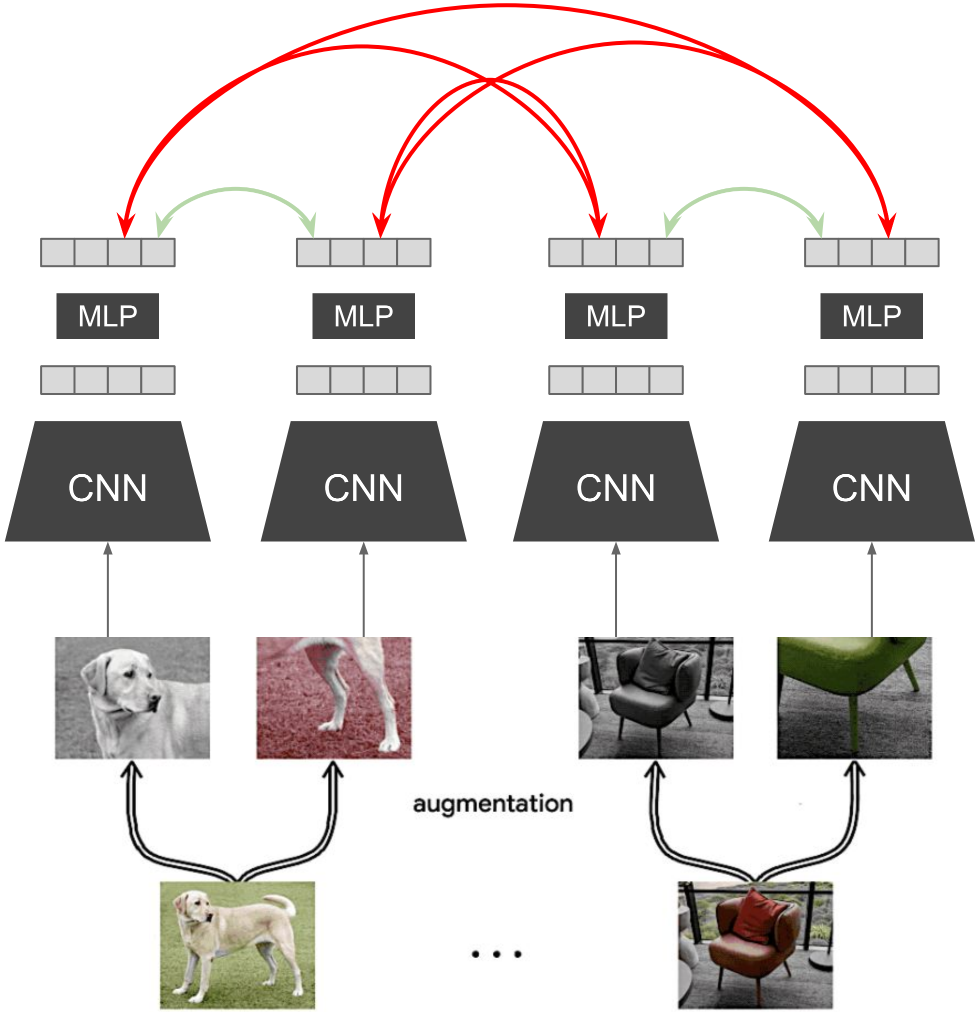
= positive pair

= negative pair

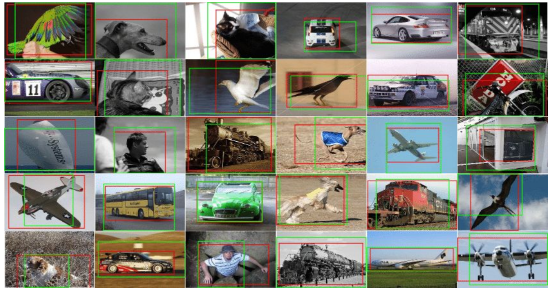
One criterion: how linearly separable -- classification accuracy

Introduction

Contrastive Learning based Self-Supervised Learning

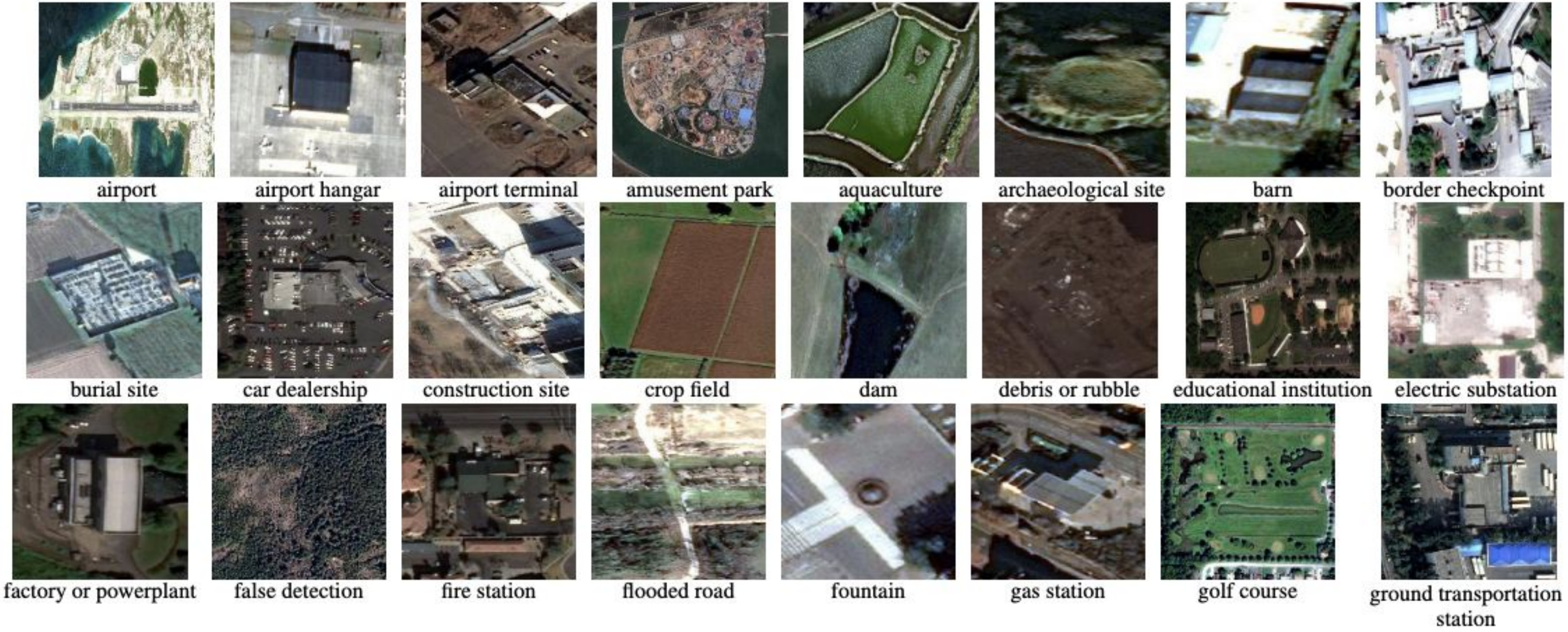


SOTA on various Computer Vision Tasks



PASCAL VOC

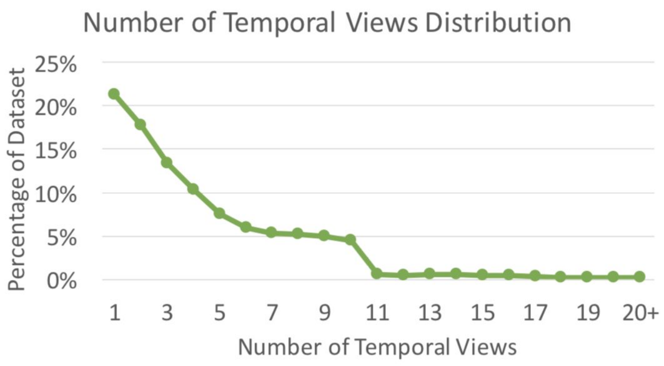
However!



Functional Map of the World Dataset (fMoW)

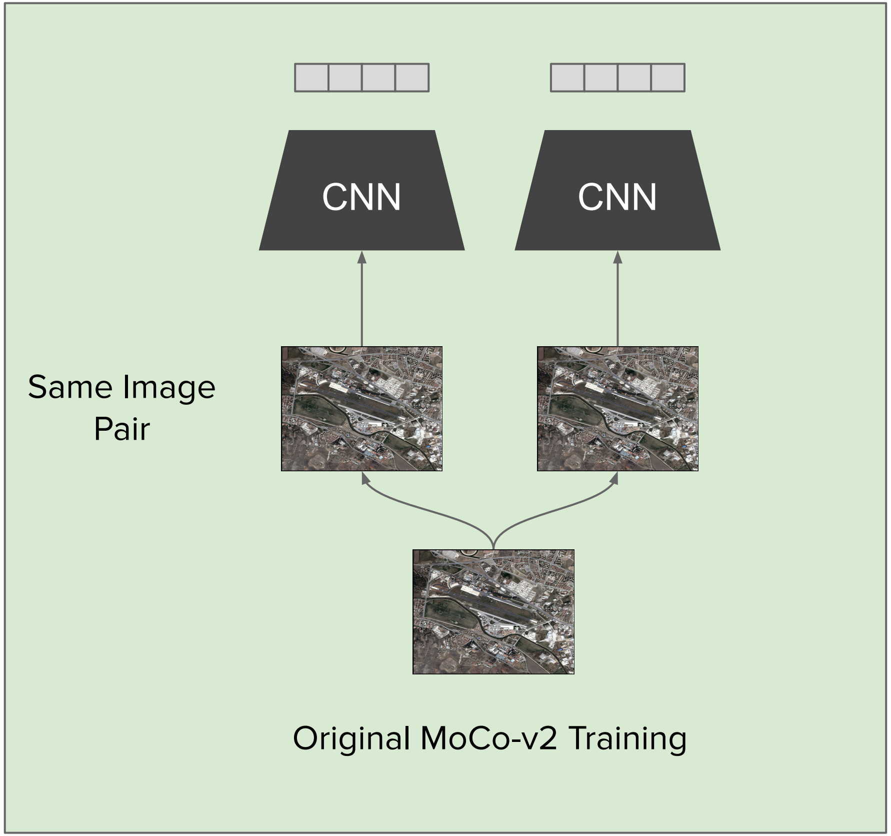
How can we address this?

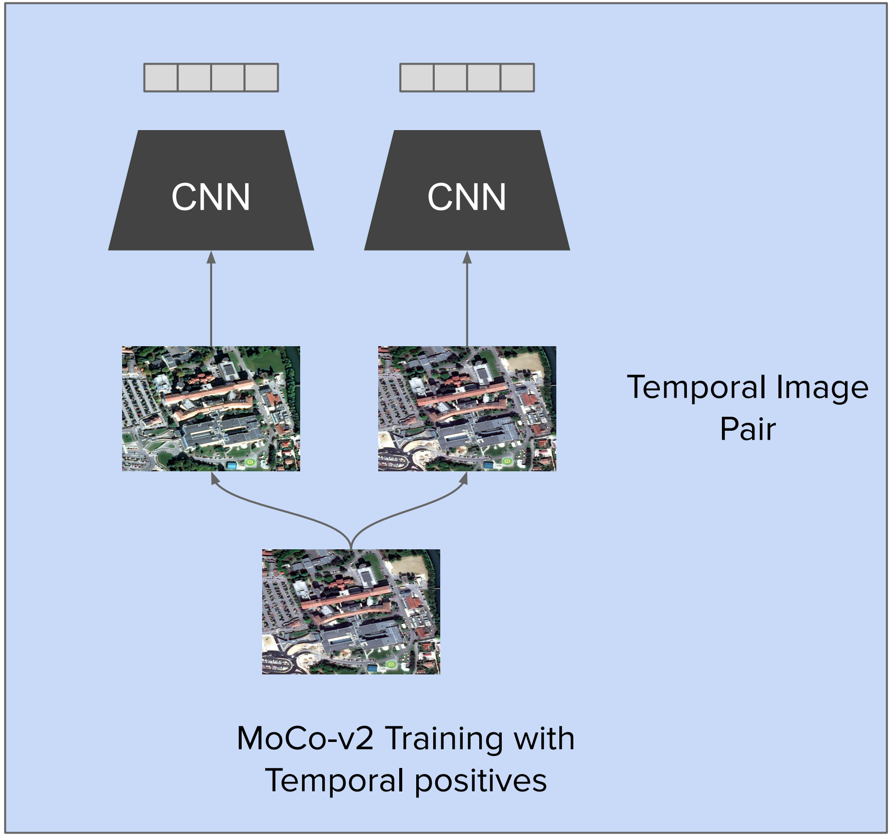
Spatially Aligned Temporal Data



Functional Map of the World (fMoW)

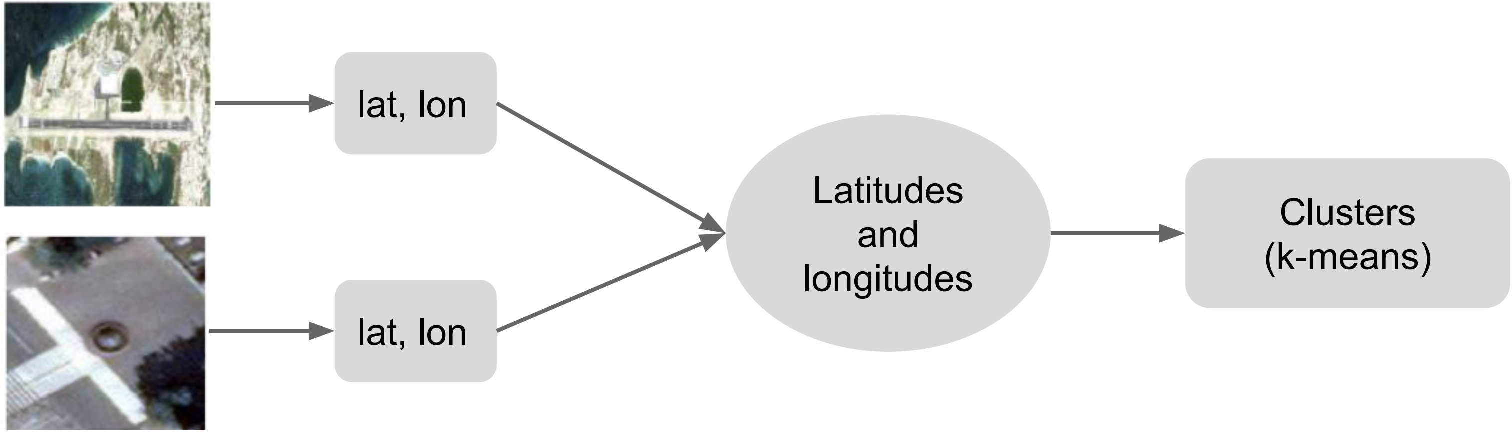
Contrastive Learning





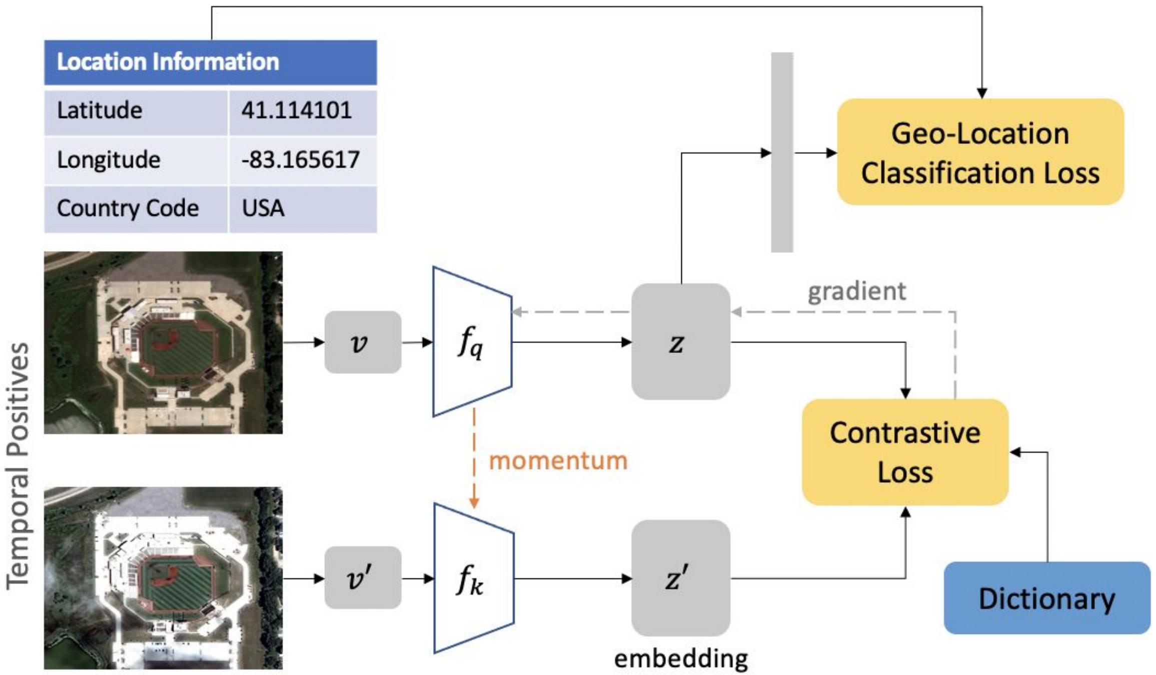
Distributions of Images in fMoW

Clustering



Clustering on fMoW

Proposed Framework



Combining Geo-location Classification and Temporal Contrastive Learning Loss

Results

	Backbone	F1-Score ↑ (Frozen/Finetune)	Accuracy ↑ (Frozen/Finetune)
Sup. Learning (IN wts. init.)*	ResNet50	-/64.72	-/69.07
Sup. Learning (Scratch)*	ResNet50	-/64.71	-/69.05
Geoloc. Learning*	ResNet50	48.96/52.23	52.40/56.59
MoCo-V2 (pre. on IN)	ResNet50	31.55/57.36	37.05/62.90
MoCo-V2	ResNet50	55.47/60.61	60.69/64.34
MoCo-V2+Geo	ResNet50	61.60/66.60	64.07/69.04
MoCo-V2+TP	ResNet50	**64.53/67.34**	**68.32/71.55**
MoCo-V2+Geo+TP	ResNet50	63.13/66.56	66.33/70.60

Experiments on fMoW on classifying single images. Frozen corresponds to linear classification on frozen features. Finetune corresponds to end-to-end finetuning results for the fMoW classification.

Results

pre-train	AP_{50} ↑
Random Init.	10.75
Sup. Learning (IN wts. init.)	14.44
Sup. Learning (Scratch)	14.42
MoCo-V2	15.45 (+4.70)
MoCo-V2-Geo	15.63 (+4.88)
MoCo-V2-TP	17.65 (+6.90)
MoCo-V2-Geo+TP	**17.74 (+6.99)**

Object Detection Results on xView

pre-train	mIOU ↑
Random Init.	74.93
Imagenet Init.	75.23
Sup. Learning (IN wts. init.)	75.61
Sup. Learning (Scratch)	75.57
MoCo-V2	78.05 (+3.12)
MoCo-V2-Geo	78.42 (+3.49)
MoCo-V2-TP	78.48 (+3.55)
MoCo-V2-Geo+TP	**78.51 (+3.58)**

Segmentation Results on SpaceNet

pre-train	Top-1 Accuracy ↑
Random Init.	51.89
Imagenet Init.	53.46
Sup. Learning (IN wts. init.)	54.67
Sup. Learning (Scratch)	54.46
MoCo-V2	55.18 (+3.29)
MoCo-V2-Geo	**58.23 (+6.34)**
MoCo-V2-TP	57.10 (+5.21)
MoCo-V2-Geo+TP	57.63 (+5.74)

Land Cover Classification on NAIP dataset

GeoImageNet

37.303518,
-121.897773,
San Jose,
United States,
California,
Santa Clara

10.595525,
76.041355,
Guruvayoor,
India,
Kerala,
Thrissur

-18.646245,
24.614868,
Botswana,
Chobe

21.881018,
-102.275360,
Aguascalientes,
México,
Aguascalientes,
Aguascalientes

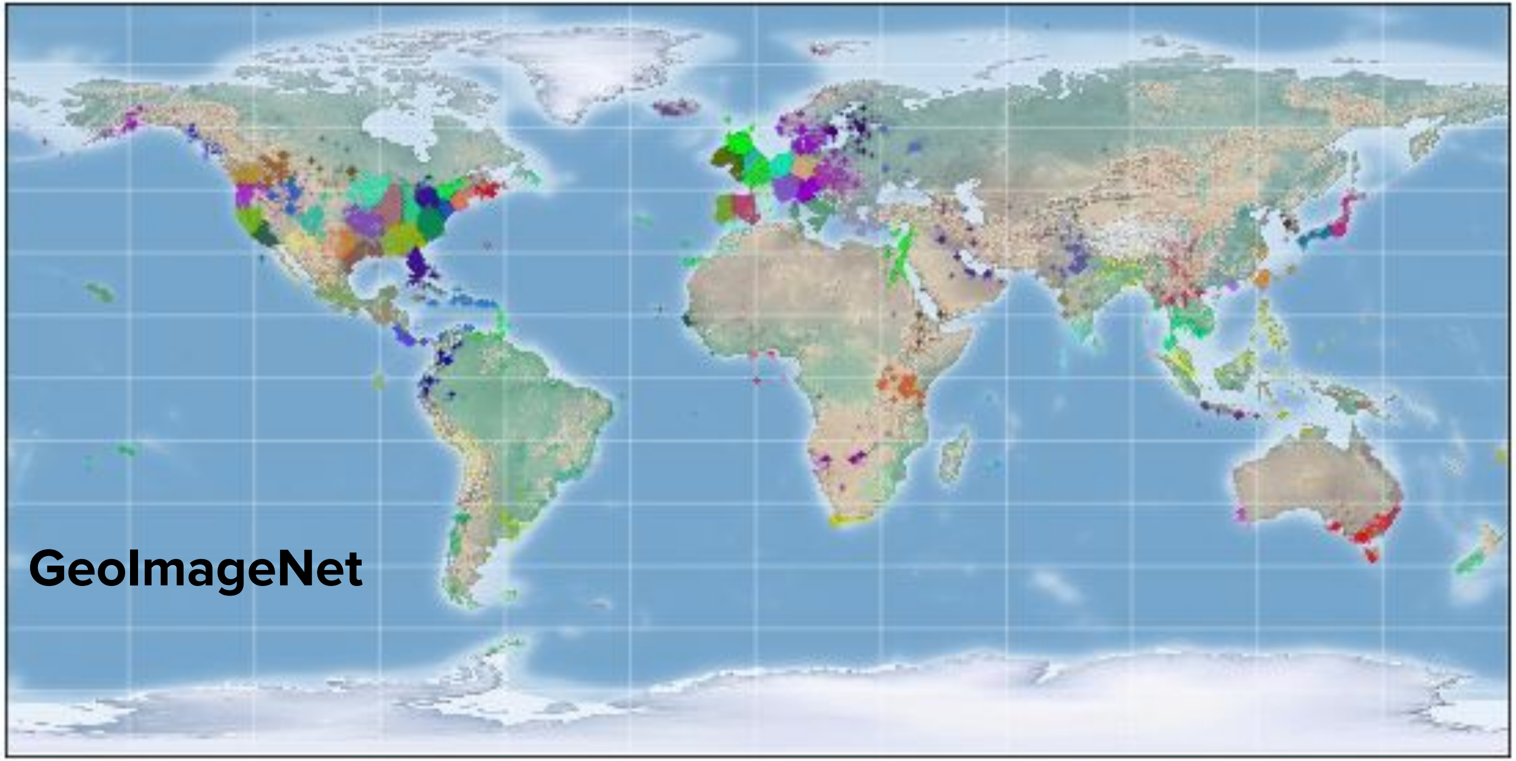
37.345989,
-77.645680,
Beach,
United States,
Virginia,
Chesterfield

-45.829337,
170.730800,
Te Matai,
New Zealand,
Otago

- We extract the **geo-location** information of **ImageNet** images using **FLICKR API**.
- We were able to find **543,435 images** with their **associated coordinates** across **5150 class** categories.
- This dataset is **more challenging** than ImageNet-1k as it is **highly imbalanced** and contains about **5× more classes.**

Distributions of Images in GeoImageNet

Clustering on GeoImageNet



Results

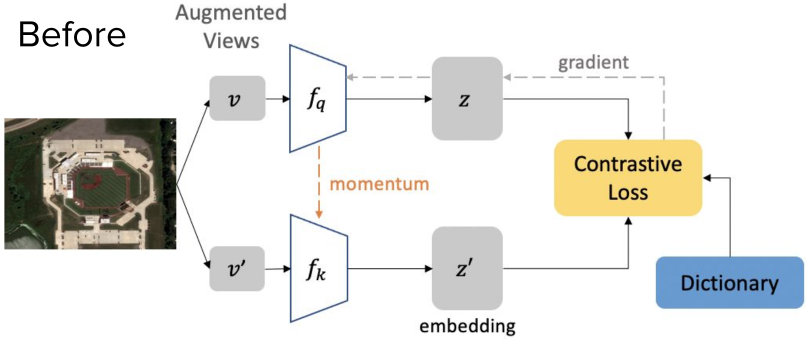
	Backbone	Top-1 (Accuracy) ↑	Top-5 (Accuracy) ↑
Sup. Learning (Scratch)	ResNet50	35.04	54.11
Geoloc. Learning	ResNet50	22.26	39.33
MoCo-V2	ResNet50	38.51	57.67
MoCo-V2+Geo	ResNet50	**39.96**	**58.71**

Experiments on GeoImageNet.

We divide the dataset into 443,435 training and 100,000 test images across 5150 classes. We train MoCo-V2 and MoCo-V2+Geo for 200 epochs whereas Sup. and Geoloc. Learning are trained until they converge.

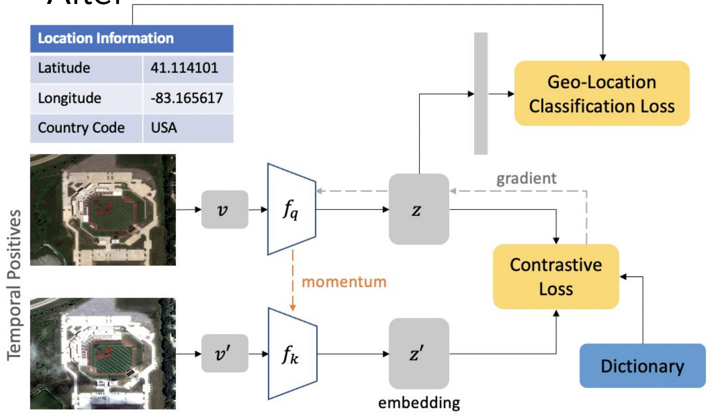
Conclusion

Before



We leverage **spatially aligned images** over time to construct **temporal positive pairs** in contrastive learning and **geo-location** to design pre-text tasks.

After



We **close the gap** between contrastive and supervised learning on various tasks for **remote sensing** and other **geo-tagged image datasets** like **GeoImagenet**.

Thank You!

- Website: https://geography-aware-ssl.github.io/
- Paper Link: https://arxiv.org/pdf/2011.09980.pdf
- Code: https://github.com/sustainlab-group/geography-aware-ssl
- Contact: {**kayush**, **buzkent**, **chenlin**}**@cs.stanford.edu**