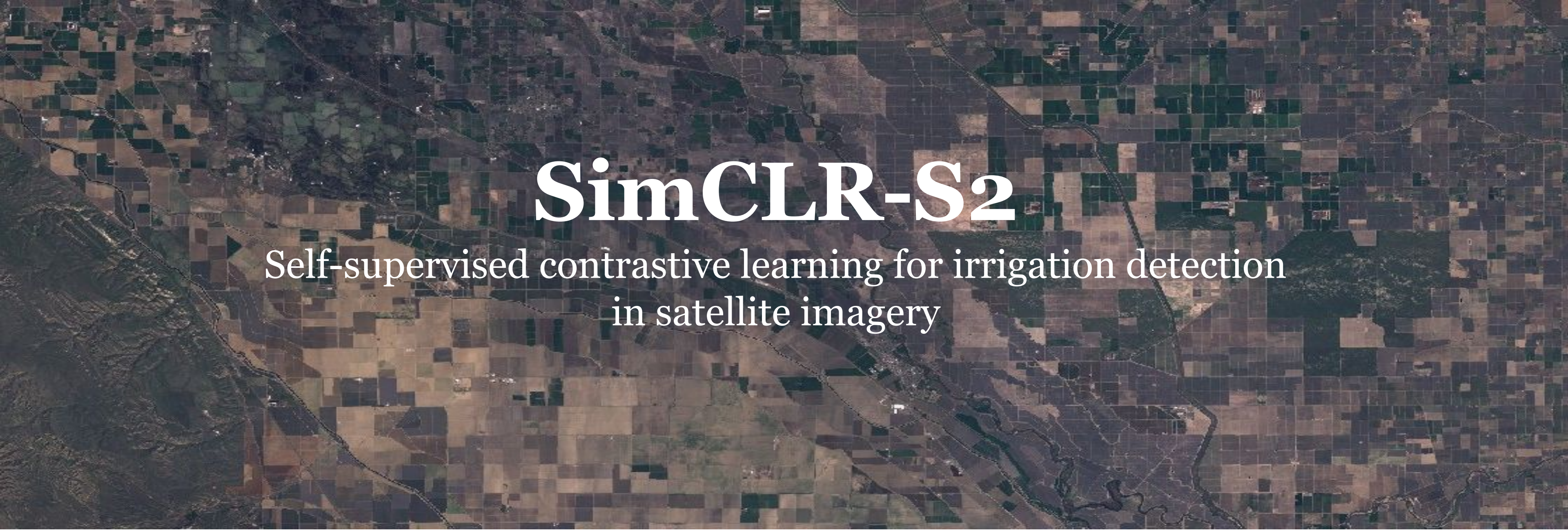


SimCLR-S2

Self-supervised contrastive learning for irrigation detection in satellite imagery

Detect presence of irrigation in uncurated and unlabelled data

Chitra Agastya, Ian Anderson, Sirak Ghebremusse

Berkeley
UNIVERSITY OF CALIFORNIA

https://mids-irrigation-capstone.github.io/

Sources

Sentinel-2 Images

Central Valley, CA
downloaded from Google Earth Engine

BigEarthNet-S2
(BEN)

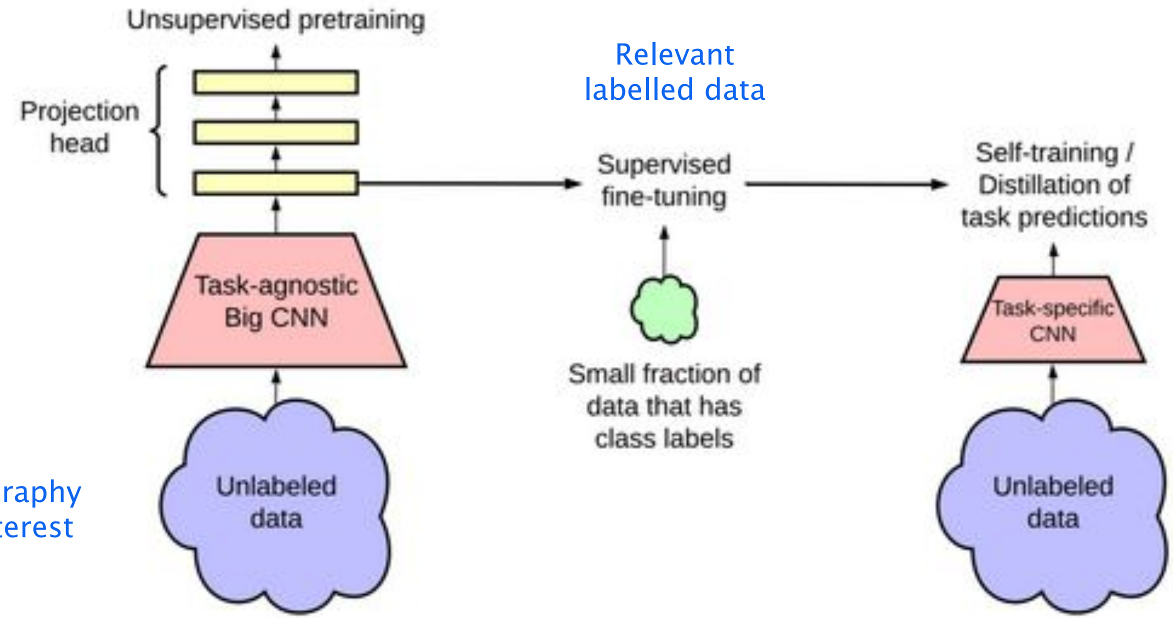
permanently irrigated land, sclerophyllous vegetation, beaches, dunes, sands, estuaries, sea and ocean

permanently irrigated land, vineyards, beaches, dunes, sands, water courses

- 590326 non overlapping images from 10 European countries
- Multi spectral images with 12 bands (i.e band 10 has been removed)
- Labelled images with multiple land cover classes, permanently irrigated land being one of them

Used data from croplands.org for generalization testing

SimCLR-S2



Relevant labelled data

Geography of interest

Geography of interest

1

2

3

Ref: [Big Self-Supervised Models are Strong Semi-Supervised Learners]

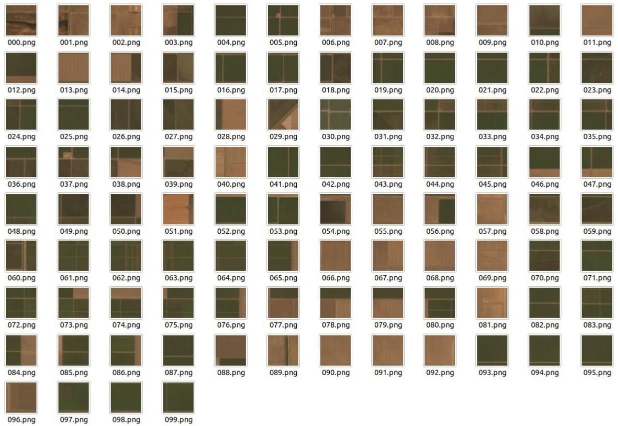
Case Study 1

Evaluate precision on geography of interest

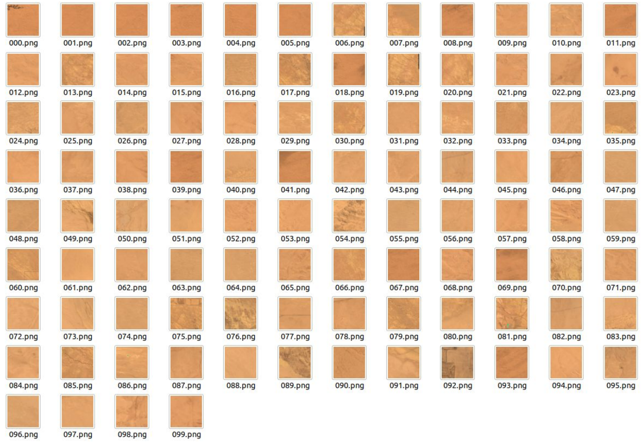
- Top 100 predictions of irrigated land with at least 99% confidence
- Supervised model's prediction confidence was far lower than that of SimCLR-S2 model
- Scoring done using Amazon Mturk and verified by visual inspection

Table 1. A comparison of precision scores from SimCLR-S2 and supervised baseline on unseen data from different geography.

Training data size (num records)	Precision (SimCLR-S2)	Precision (supervised)
190	0.99	0.11
570	1.00	0.2
1902	0.99	0.36
4756	0.98	0.95
9515	1.00	0.78
19024	1.00	0.47



SimCLR-S2: Top 100 predictions for irrigated land



Supervised: Top 100 predictions for irrigated land

Case Study 2

Evaluate recall on an unrelated geography

- Ground truth sourced from croplands.org for 6 different countries:
  Brazil, India, Indonesia, Myanmar, Tunisia and Vietnam
- Download 0.5km x 0.5km land for ground truth coordinates from croplands.org
- SimCLR-S2 generalized better than supervised baseline models

Table 2. A comparison of recall scores on SimCLR-S2 and supervised baseline for irrigated cropland from diverse geographies

Country	Training data (num records)	Recall (SimCLR-S2)	Recall (supervised)
Brazil	190	**0.75**	0.5
India	190	**0.9**	0.67
Indonesia	570	**0.76**	0.07
Tunisia	570	0.78	**0.91**
Vietnam, Myanmar	190	**0.9**	0.00

Berkeley
UNIVERSITY OF CALIFORNIA

Questions?

References

- G. Sumbul, M. Charfuelan, B. Demir, V. Markl, "*BigEarthNet: A Large-Scale Benchmark Archive for Remote Sensing Image Understanding*", *IEEE International Geoscience and Remote Sensing Symposium, pp. 5901-5904, Yokohama, Japan, 2019*.
- Chen et al., A Simple Framework for Contrastive Learning for Visual Representations, ICML, 2020
- Chen et al., Big Self-Supervised Models are Strong Semi-Supervised Learners, NeurIPS, 2020

Berkeley
UNIVERSITY OF CALIFORNIA

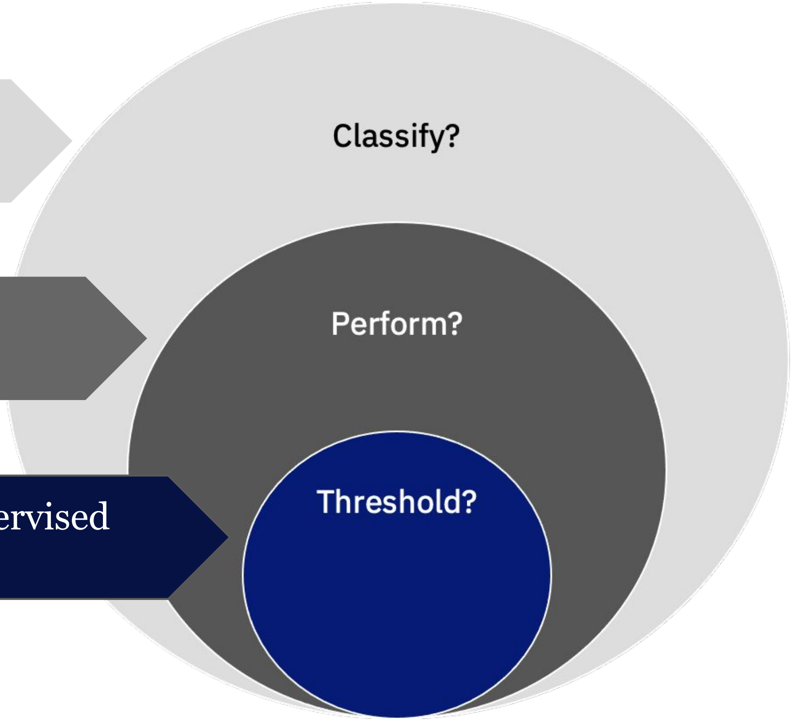
Appendix

Research Questions

Can we detect irrigation from satellite imagery with self-supervised contrastive learning?

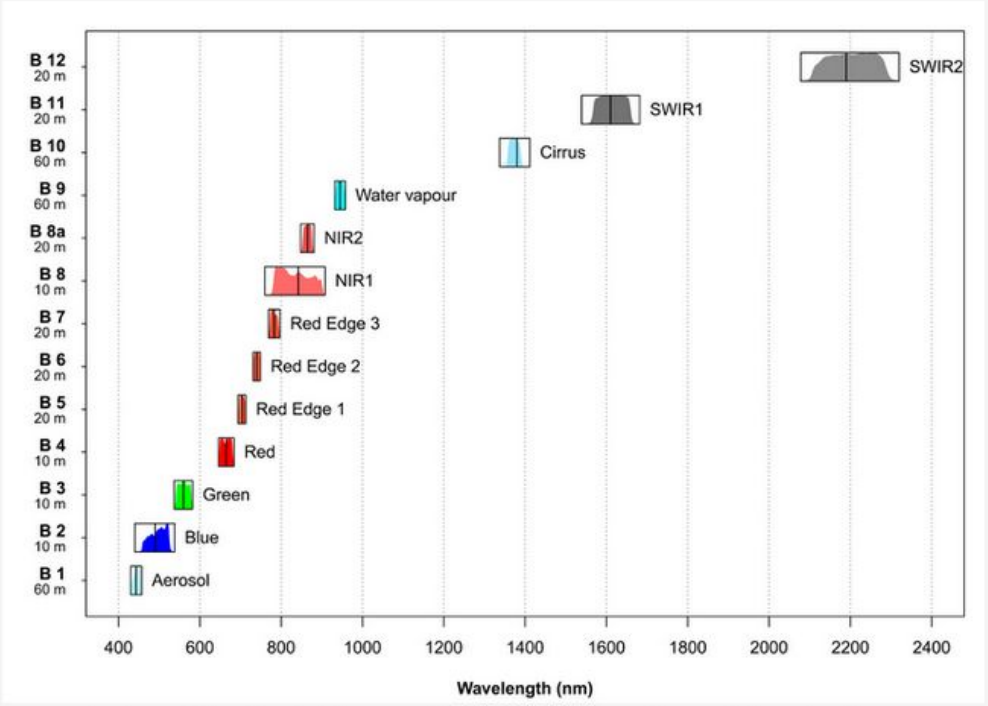
Can we achieve performance on par with supervised learning with a fraction of the dataset labelled?

At what thresholds of labelled data is the self-supervised technique still effective?



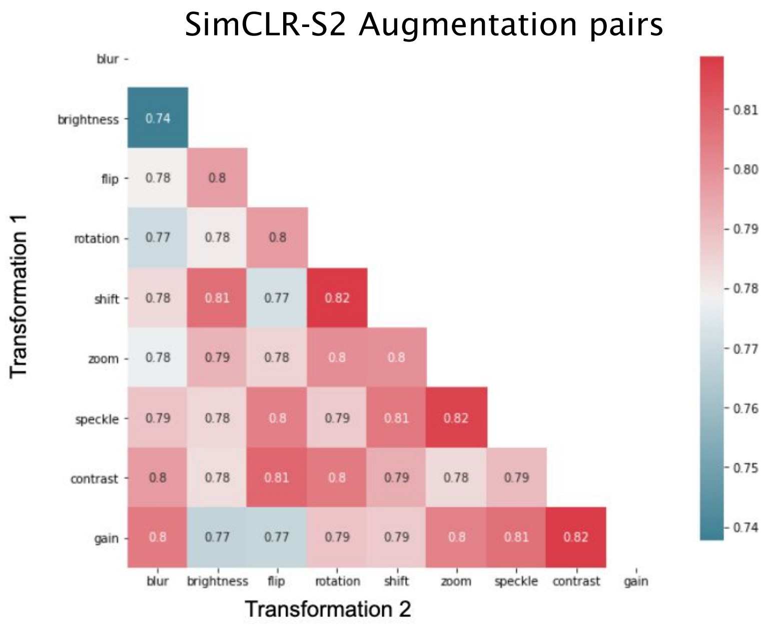
Spectral Bands - Spatial and Pixel Resolution

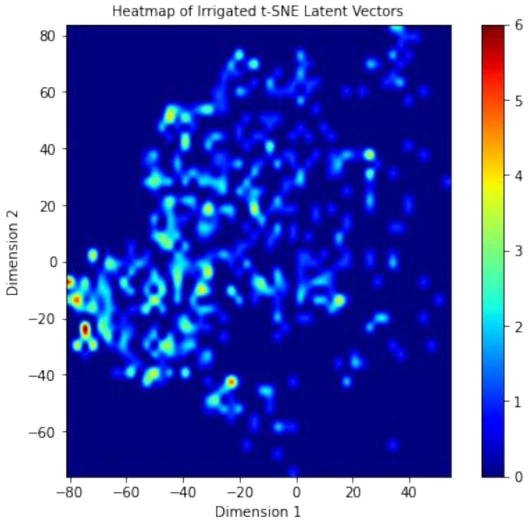
Band	Description	BEN Pixel Resolution
B12	SWIR 2	60x60
B11	SWIR 1	60x60
B09	Water Vapor	20x20
B8A	Vegetation red edge 4	60x60
B08	NIR	120x120
B07	Vegetation red edge 3	60x60
B06	Vegetation red edge 2	60x60
B05	Vegetation red edge 1	60x60
B04	Red	120x120
B03	Green	120x120
B02	Blue	120x120
B01	Coastal aerosol	20x20

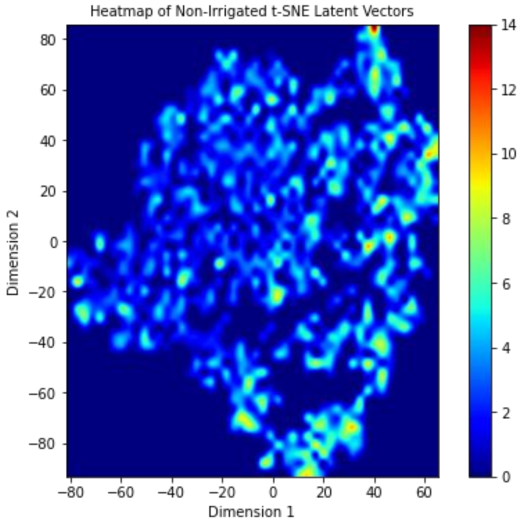


Berkeley
UNIVERSITY OF CALIFORNIA

Results: *Stage 1 SimCLR-S2 Pre-training*







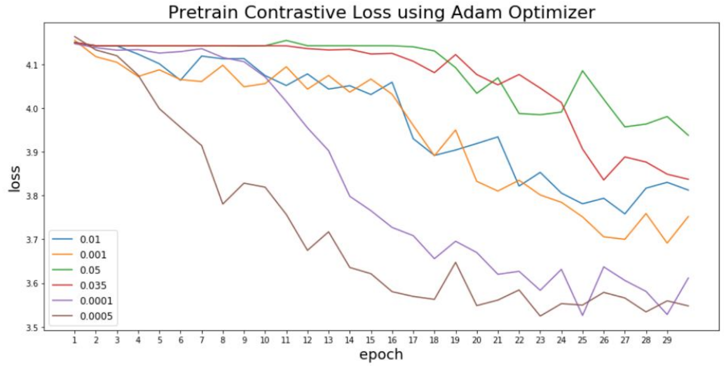
We drop the bottom 25%

Berkeley
UNIVERSITY OF CALIFORNIA

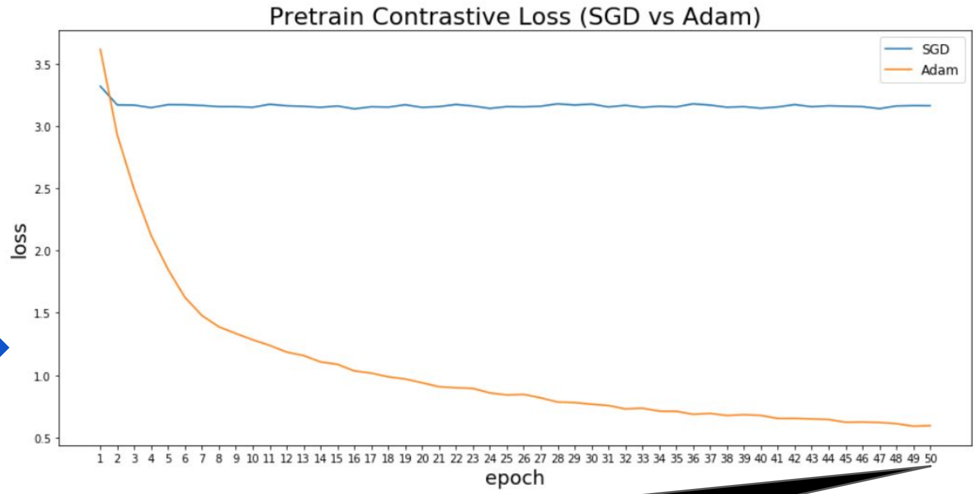
Results

Stage 1 SimCLR-S2: Pre-training Contrastive Loss

Hyper parameter tuning



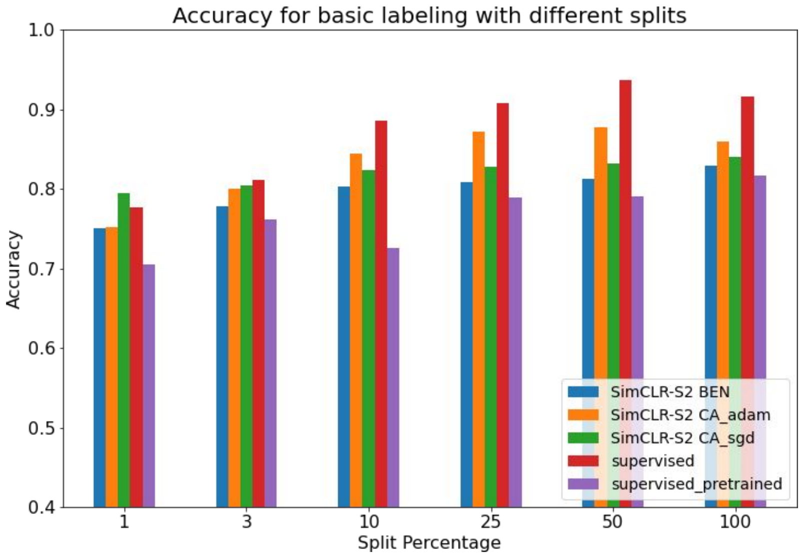
Tuning done on smaller sample for 30 epochs

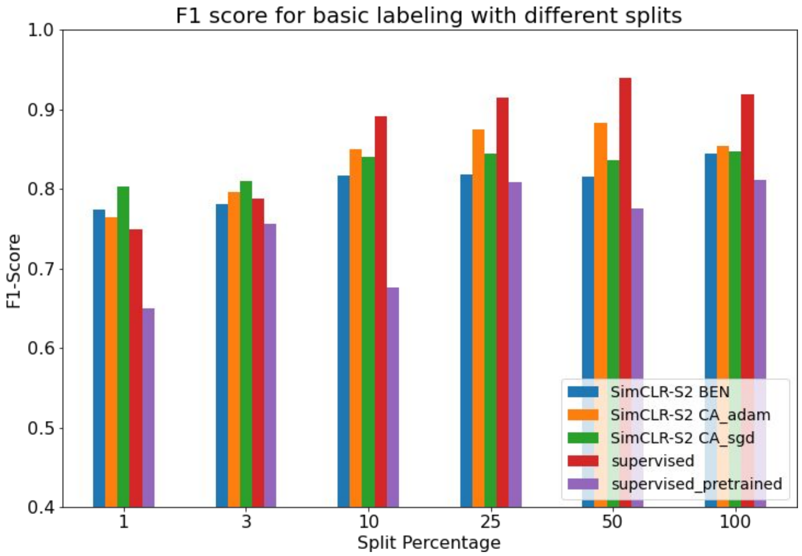


loss seems to continue on a downward trend

Results

Stage 2 SimCLR-S2 results vs supervised baseline



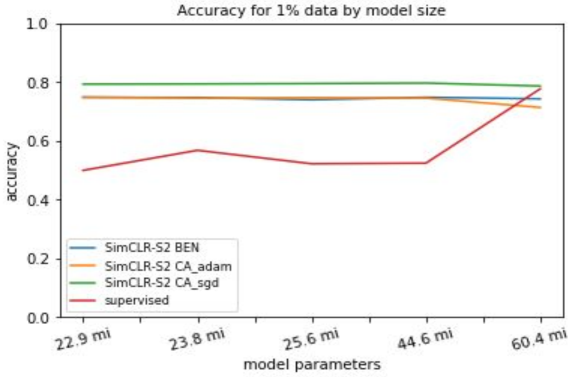


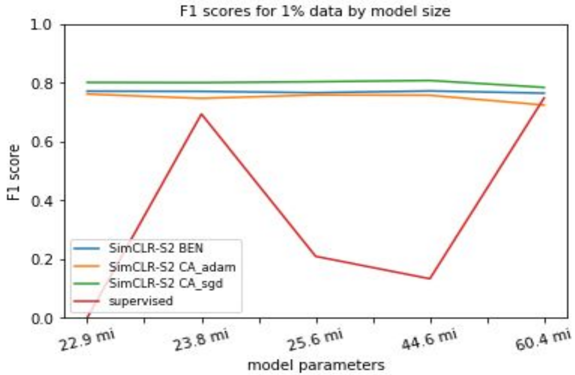
Performance of SimCLR-S2 fine-tuning results versus Supervised Baselines
Architecture: ResNet152

Berkeley
UNIVERSITY OF CALIFORNIA

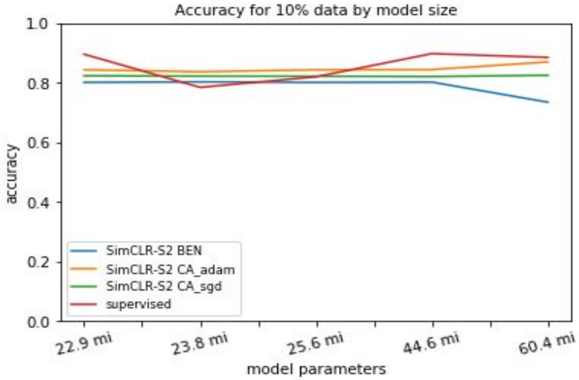
Results

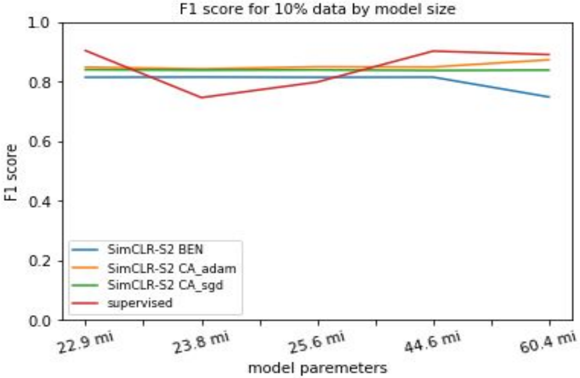
Stage 3 SimCLR-S2 results vs supervised baseline





1% data split





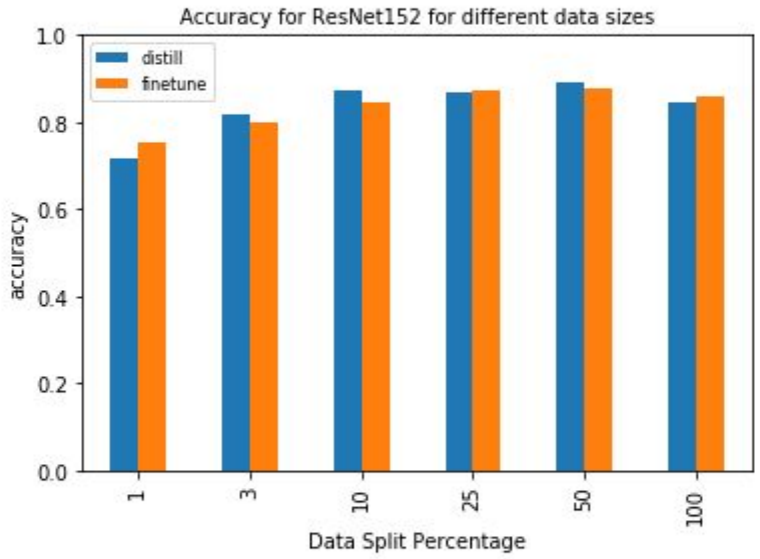
10% data split

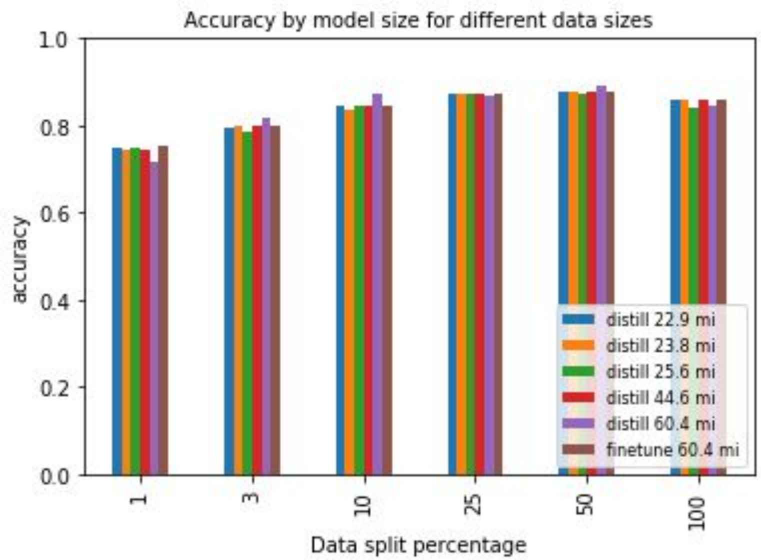
Comparison of SimCLR2 distillation scores with supervised baseline

Berkeley
UNIVERSITY OF CALIFORNIA

Results

Stage 3 SimCLR-S2 : Distill vs Fine-tune

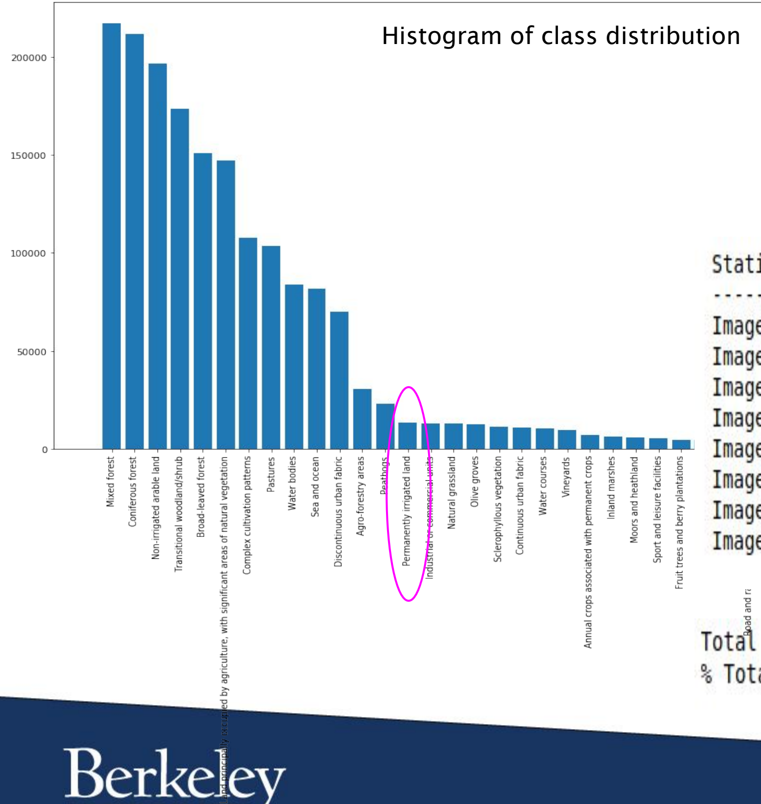




Distillation shows an improvement over finetune scores for many architectures and many data split percentages

Berkeley
UNIVERSITY OF CALIFORNIA

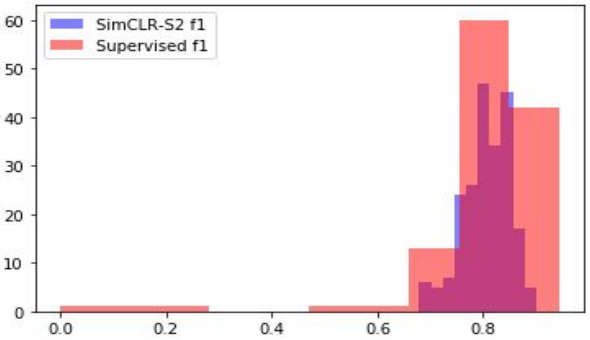
BigEarthNet EDA - True Class



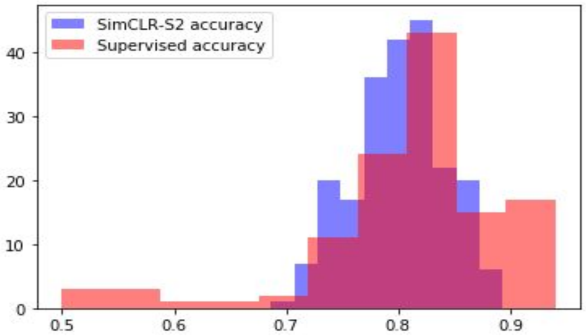

```
Statistics for other permanently irrigated cropland that are not included in "Permanently irrigated land"
------------------------------------------------------------------------------------------------------
Images with "Fruit trees and berry plantations" that are also "Permanently irrigated land": 445
Images with "Fruit trees and berry plantations" that are not "Permanently irrigated land": 4309
Images with "Rice fields" that are also "Permanently irrigated land": 1504
Images with "Rice fields" that are not "Permanently irrigated land": 2289
Images with "Vineyards" that are also "Permanently irrigated land": 1377
Images with "Vineyards" that are not "Permanently irrigated land": 8190
Images with "Olive groves" that are also "Permanently irrigated land": 2770
Images with "Olive groves" that are not "Permanently irrigated land": 9768

Total irrigated land: 36686
% Total irrigated land: 6.2145323092664055
```

Berkeley
UNIVERSITY OF CALIFORNIA

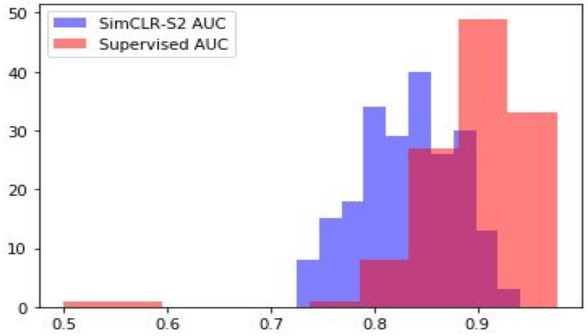
SimCLR-S2 - Impact on performance



(a) F1 Score



(b) Accuracy

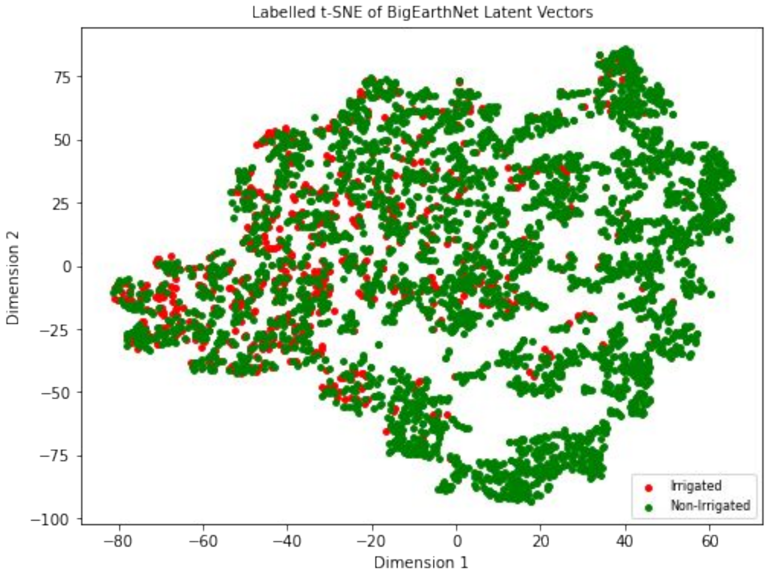


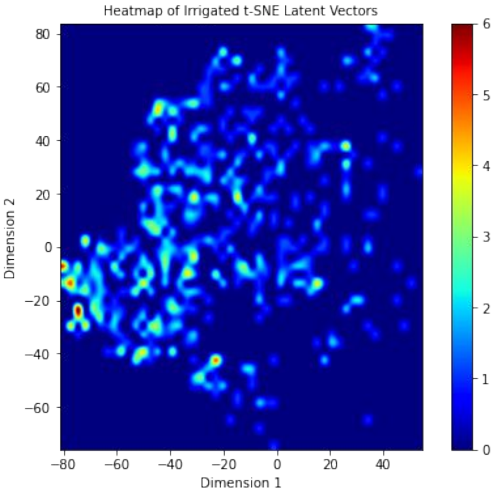
(c) AUC Score

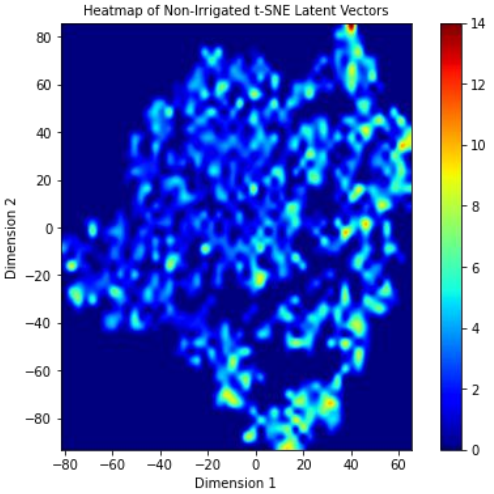
SimCLR-S2 seems to be bringing the scores closer. So smaller datasets see a big improvement in performance.

Berkeley
UNIVERSITY OF CALIFORNIA

t-SNE Results of Pretrained BigEarthNet Latent Vectors







Berkeley
UNIVERSITY OF CALIFORNIA