

Journal reference

A. Montanaro, D. Valsesia, G. Fracastoro, E. Magli, "Self-supervised learning for joint SAR and multispectral land cover classification", *arXiv preprint arXiv:2108.09075*, 2021

The Problem: Weakly Supervised Learning in Remote Sensing Imagery

- ▶ End-to-end DL models not able to outperform standard ML methods
- ▶ Lack of large labeled datasets
- ▶ Difficulty to fuse multiple imaging modalities
 - ▷ Some solutions:
- ▶ **Transfer Learning** from CV suffer from domain gaps
 - ▷ RGB vs Multichannel
 - ▷ Object-Centric vs Many Objects at different scale
 - ▷ Different Spatial Resolution among multisensor data
- ▶ **Self-Supervised Learning** is strictly related to CV:
 - ▷ Pretext Tasks involve geometric transformations
 - ▷ Contrastive Learning is not able to build HR feature maps
 - ▷ CV → Well separated clusters
 - ▷ RS → Overlapped clusters (metropolis and village)

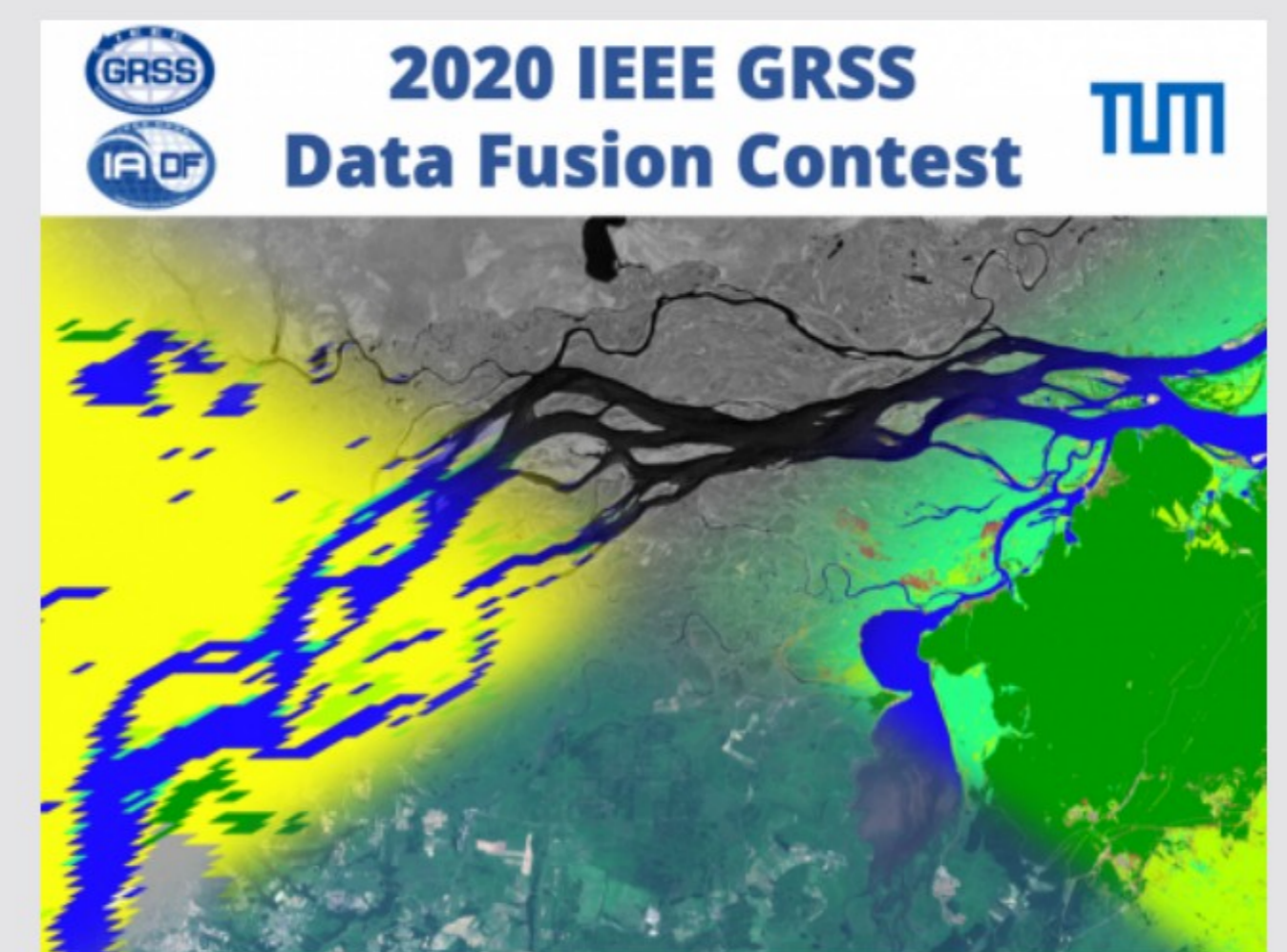


Spatial-Spectral Context Learning

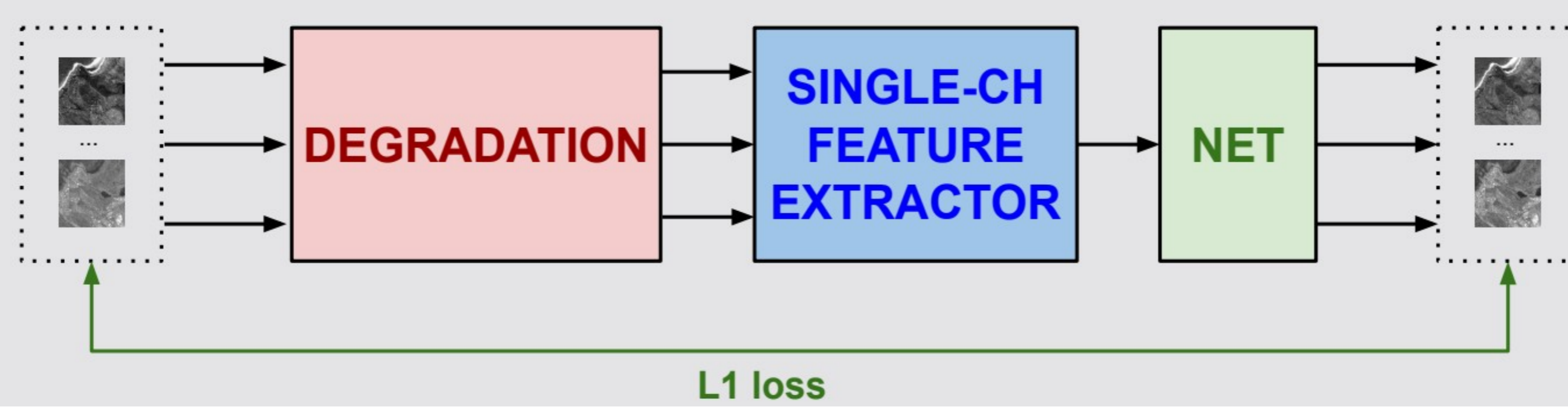
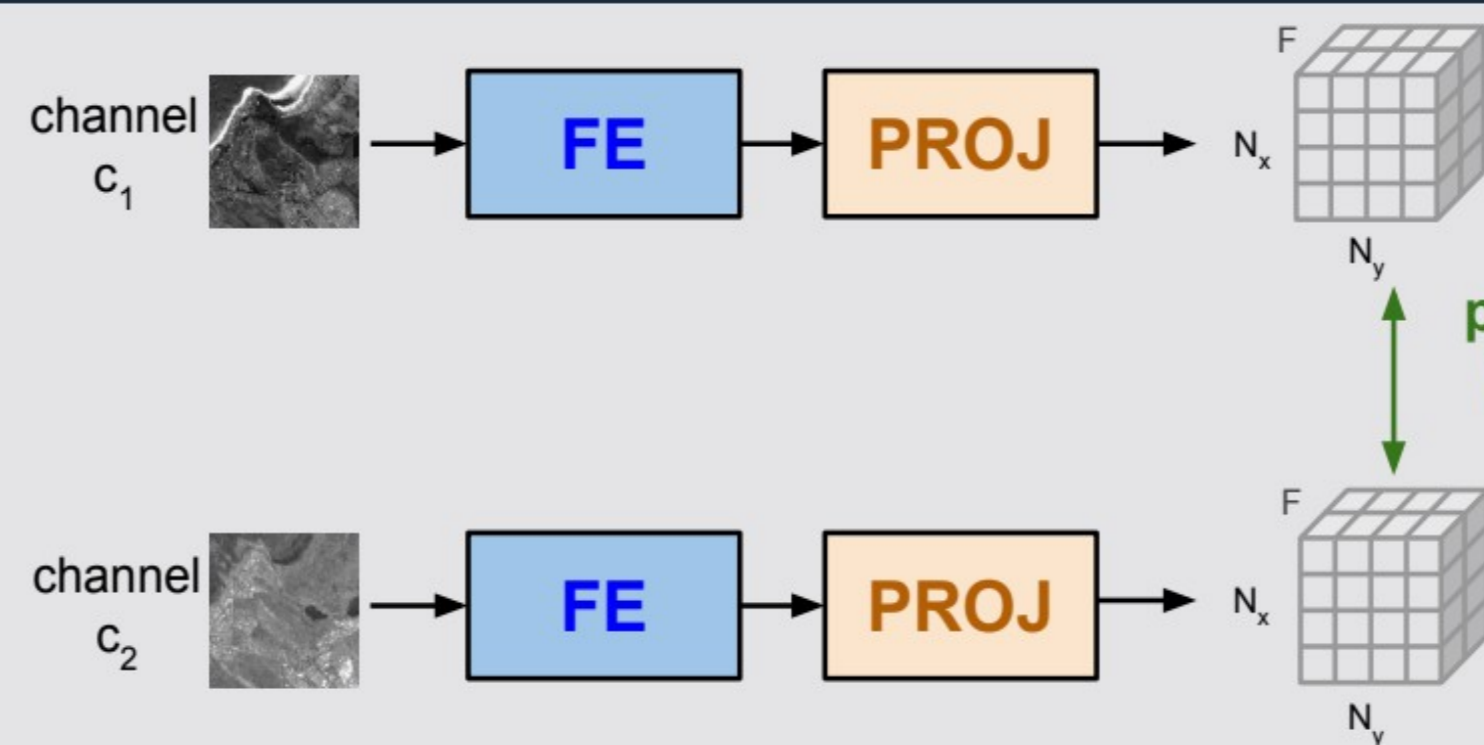
- ▶ Universal Framework for multichannel data
- ▶ Extracts Features able to capture material properties
- ▶ 2 Self-Supervised approaches trained sequentially
- ▶ **UniFeat**
 - ▷ Promoting pixel-wise similarity of low-level features between two different channels
 - ▷ Single-Channel Feature Extractor and Projection head
- ▶ **CoRe**
 - ▷ Reconstruct a degraded input to represent the spectrum highly informative for material properties
 - ▷ Degradation Process (Channel Dropout, Cutout, Blur)

Data Fusion Contest 2020 (Track 2)

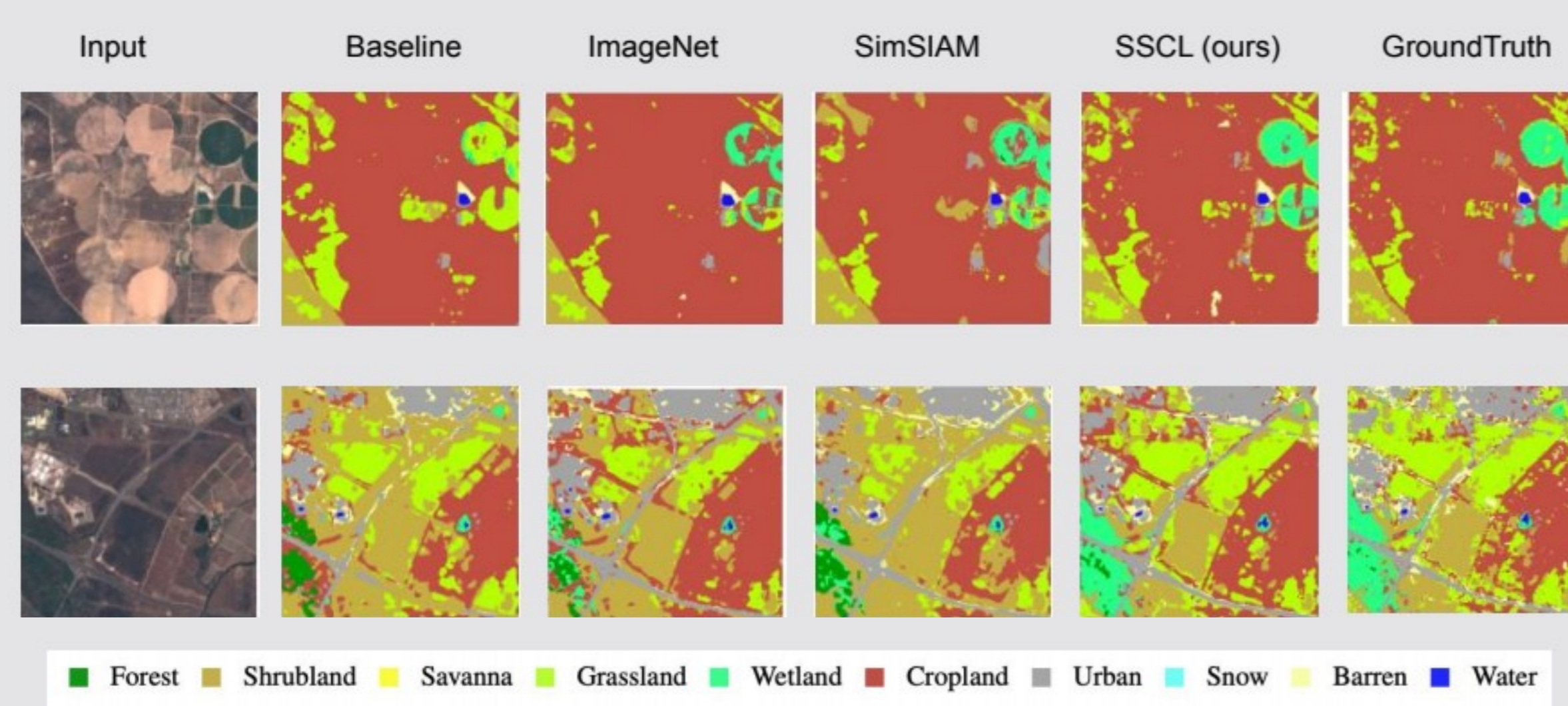
- ▶ **High Resolution** Land Cover Classification
- ▶ **Dataset:**
 - ▷ Train: 986 images
 - ▷ Test: 5128 images
 - ▷ Annotated HR (10m) maps 10 classes
 - ▷ S1 (2 polarization bands)
 - ▷ S2 (13 multispectral bands)
- ▶ **Experimental Setting**
 - ▷ Linear Protocol
 - ▷ Fine-Tuning



CNN Architecture



Results



	Baseline	ImageNet	SimSiam	SSCL
Forest	74.6 ± 5.7	77.7 ± 4.7	75.5 ± 1.3	83.8 ± 1.7
Shrubland	56.3 ± 1.9	50.7 ± 7.0	46.4 ± 7.1	60.1 ± 4.5
Grassland	33.3 ± 5.9	35.4 ± 8.1	37.4 ± 4.2	39.5 ± 8.0
Wetlands	11.0 ± 2.1	9.3 ± 3.7	13.4 ± 3.6	12.1 ± 2.4
Croplands	32.1 ± 5.3	36.8 ± 5.4	28.1 ± 4.1	29.4 ± 1.1
Urban	79.6 ± 3.9	77.1 ± 3.5	79.8 ± 6.9	78.9 ± 4.1
Barren	40.9 ± 5.2	42.2 ± 6.7	43.4 ± 10.3	44.8 ± 5.3
Water	99.2 ± 0.2	99.3 ± 0.1	99.3 ± 0.1	99.3 ± 0.1
AA	53.4 ± 1.3	53.6 ± 0.9	52.9 ± 1.1	56.0 ± 1.1
OA	65.1 ± 1.9	66.4 ± 0.8	64.6 ± 1.3	67.8 ± 0.9

Legend

- ▷ Baseline: Standard train procedure with randomly initialized model
- ▷ ImageNet: Train by using Transfer Learning from an architecture trained on ImageNet dataset.
- ▷ SimSiam: Train by finetuning the pretrained self-supervised model through SimSiam technique.
- ▷ SSCL: Train finetuning the proposed self-supervised model.

