

Multi-Label Noise Robust Methods in RS Image Classification (V. 2)

Mentor: Tom Burgert (t.burgert@tu-berlin.de)

Gliederung

1. Overview Research in Multi-Label Noise
2. Detailed Tasks

Research in Multi-Label Noise

Additive Noise: 0 -> 1

Subtractive Noise: 1 -> 0

Problem of other paper: uniform noise!

Label Matrix Y

		Classes				
Samples		1	0	0	1	0
		0	0	0	1	1
		0	1	0	0	0
		0	0	1	0	0
		1	0	1	1	0
		0	1	0	1	1



Label Matrix Y

		Classes				
Samples		1	0	0	1	0
		0	0	1	1	1
		0	1	0	0	0
		0	0	1	1	0
		0	0	1	1	0
		0	1	0	0	1

Research in Multi-Label Noise

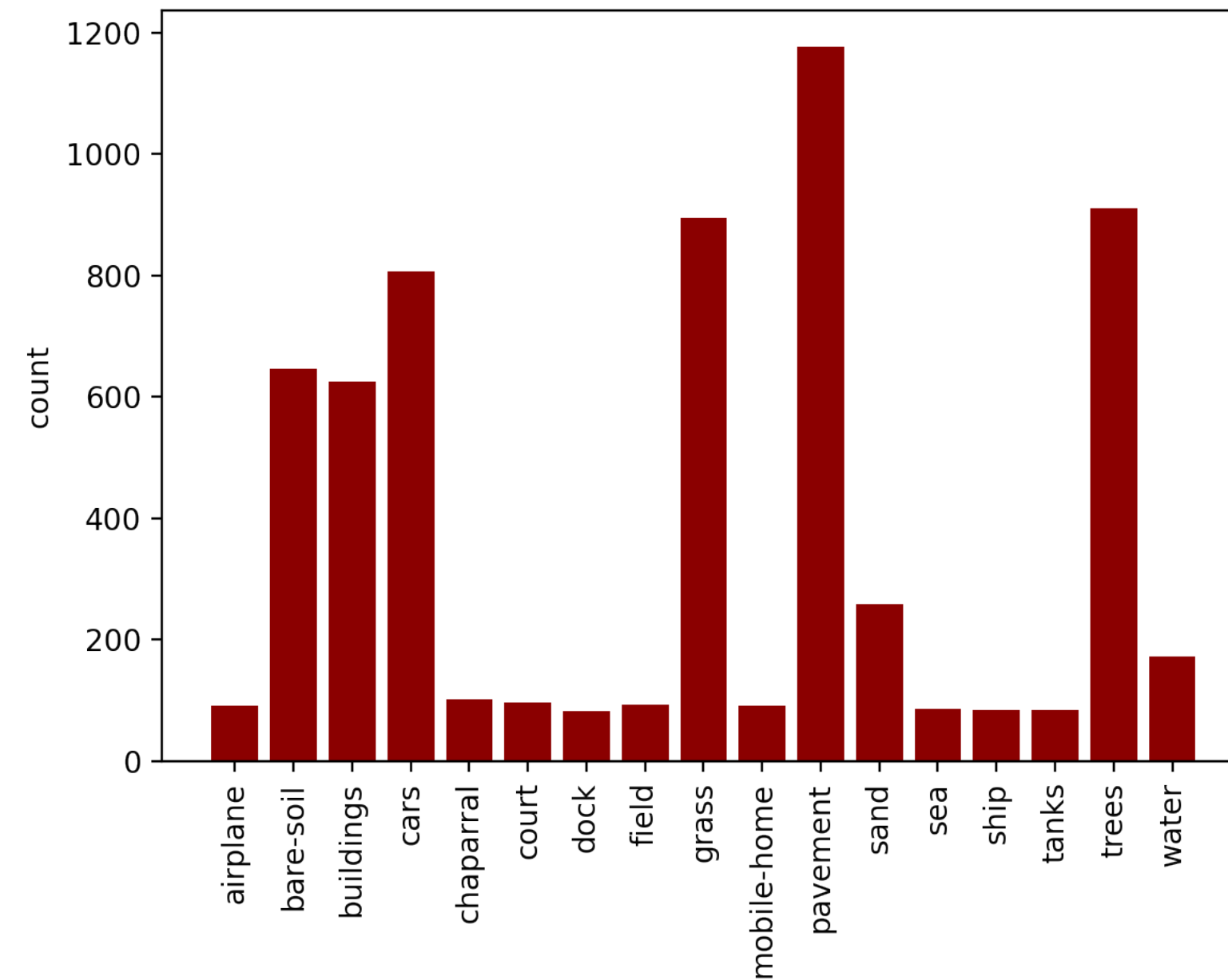
- % of noise is oriented at absolute number of present classes
- % of noise is injected per class, to preserve class distribution
- e.g. class 1 is annotated for 100 images and we inject 20% additive noise and 40% subtractive noise
 - => 20 Mal: 0 -> 1
 - => 40 Mal: 1 -> 0

Additive Noise: 0 -> 1

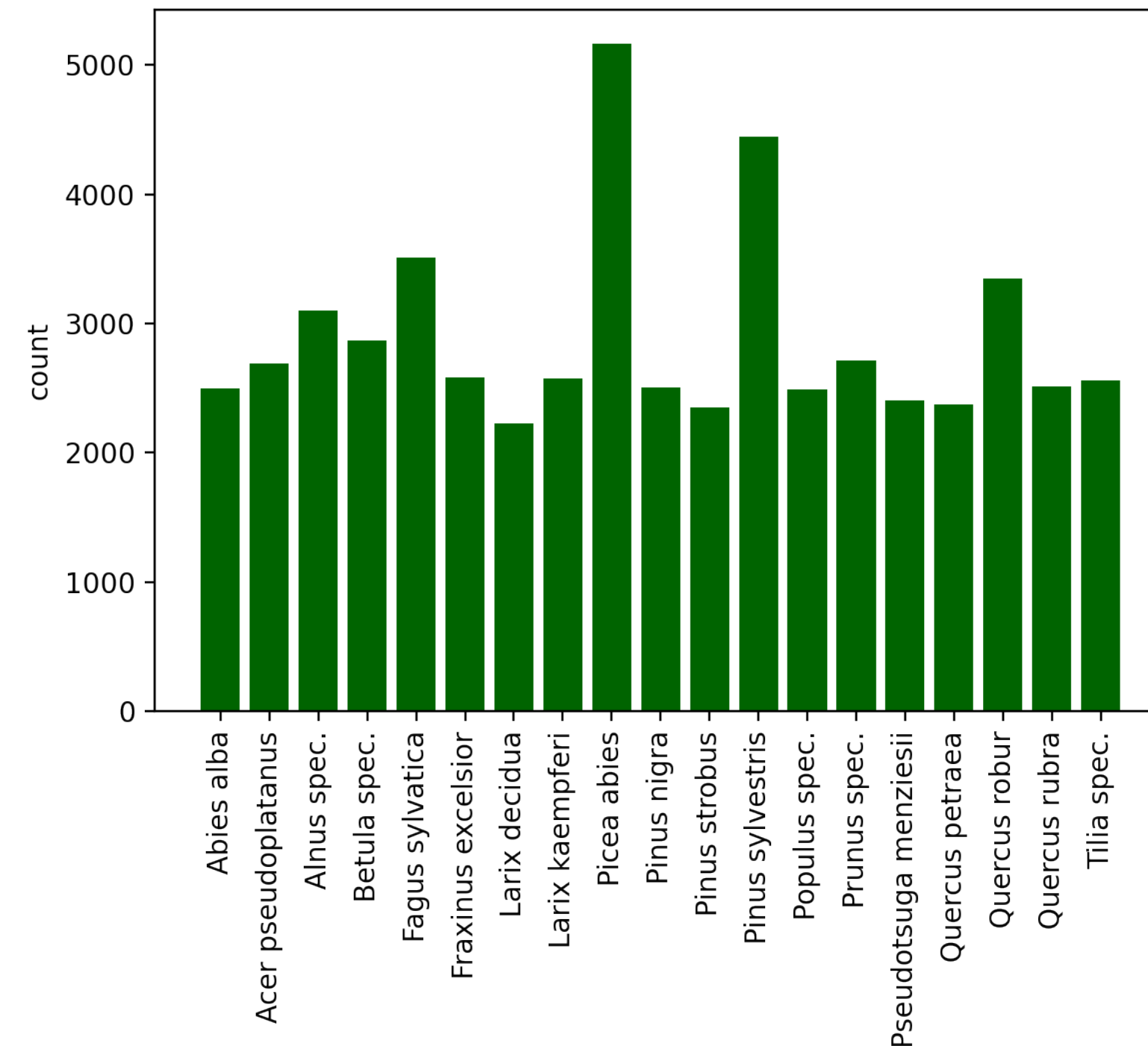
Subtractive Noise: 1 -> 0

Research in Multi-Label Noise

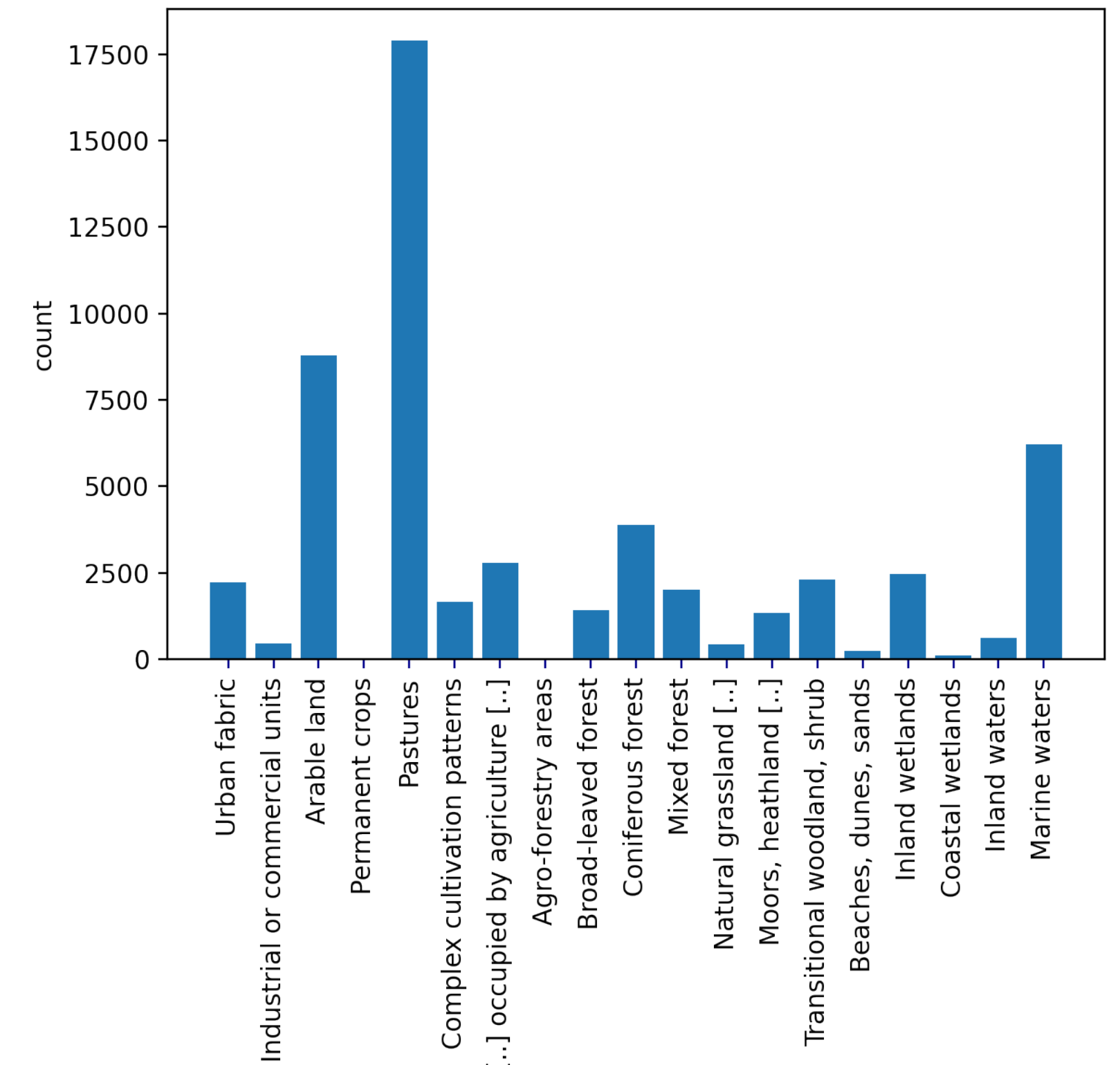
UCMerced



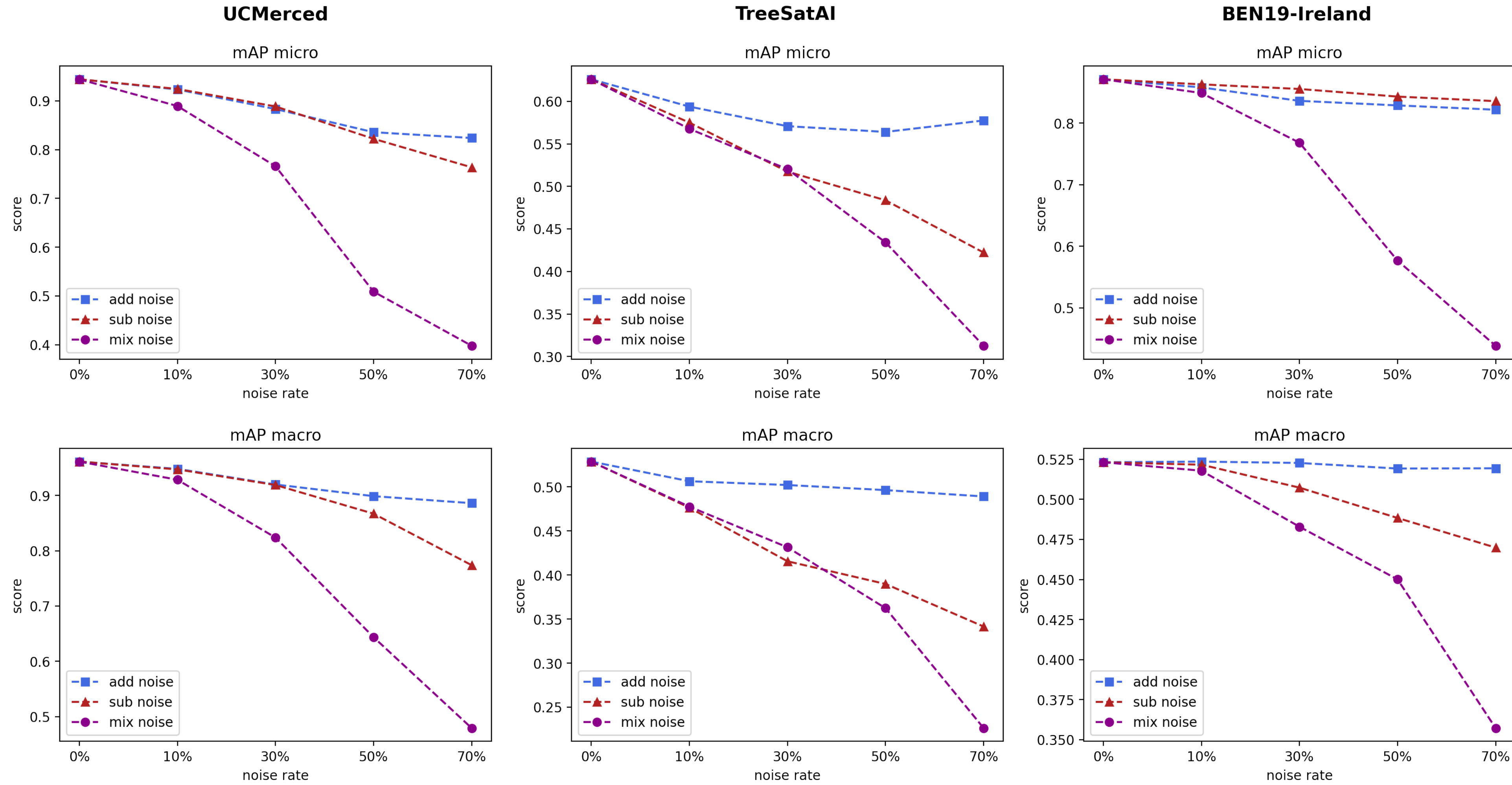
TreeSatAI



BEN19-Ireland



Research in Multi-Label Noise



Research in Multi-Label Noise

- Limitations from SLC noise robust methods for MLC:
 - => usually use DataAugmentation (limited for MLC+RS)
 - => SOTA all SSL (need of DA)
 - => include class-based decisions through prototyping, clustering or syntactic class label transfer
- Transferable:
 - => loss based methods (EMA predictions, prediction regularization)
 - => co-network learning (agreement, disagreement)

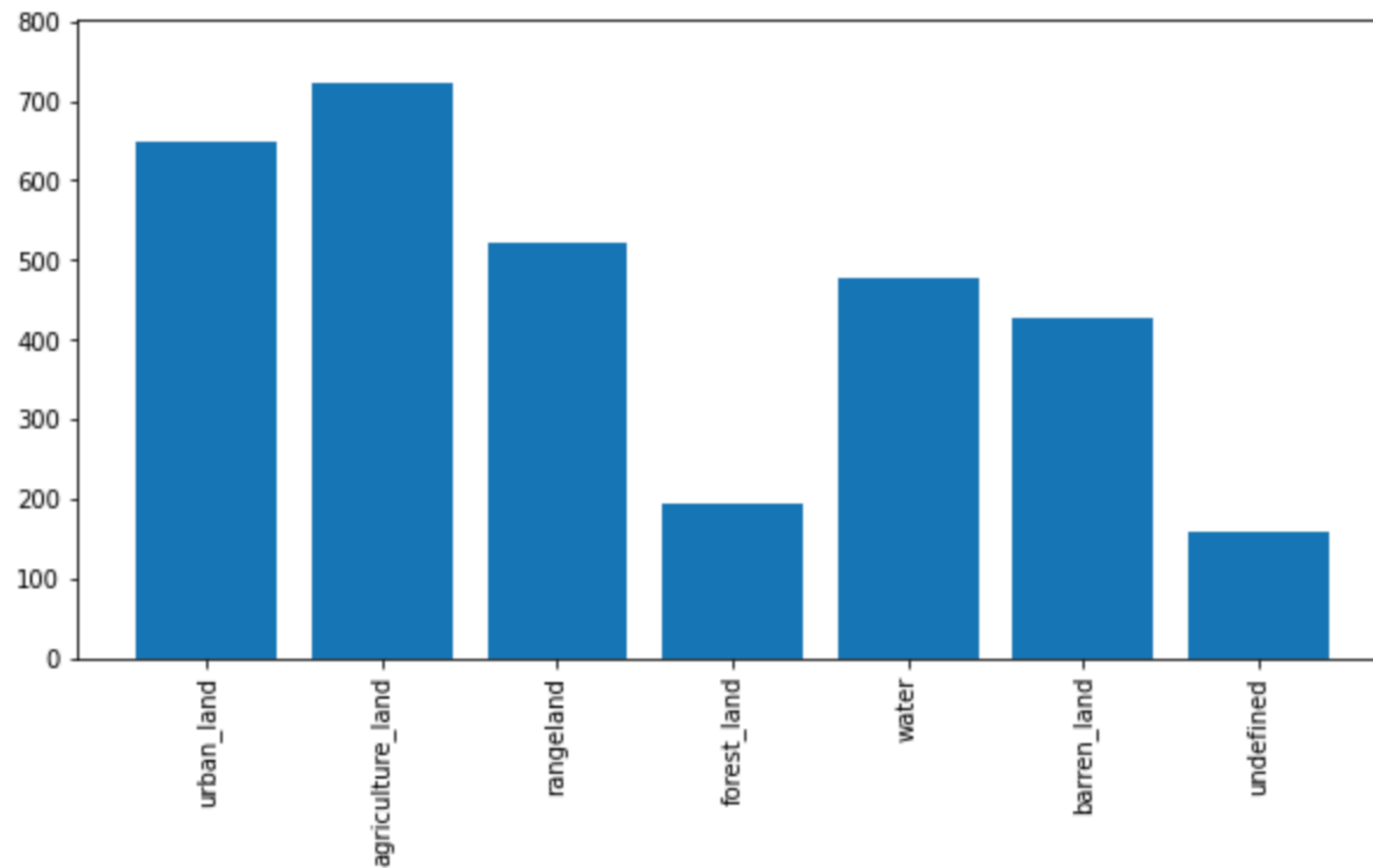
Tasks

1. Implement the CbMLC model in PyTorch

- Read-Up on:
 - => Graph Neural Networks (Message Passing Neural Network)
 - => Multi-head self attention
- Method very similar to LaMP (PyTorch Repo available)
 - => <https://github.com/QData/LaMP>
 - => <https://arxiv.org/pdf/1904.08049.pdf>

Tasks

2. Create a new MLC dataset from DeepGlobe



- Present/absent ratio: 0.56
- $E[\text{sum}(y_i)] == 3.9$
- Idea: sample 4-8 patches per image, s.t.
 $\Rightarrow E[\text{sum}(y_i)] < 2.5$
 \Rightarrow more balanced
 \Rightarrow no class in $> 50\%$ of samples

2. Create a new MLC dataset from DeepGlobe

- Link: <https://www.kaggle.com/datasets/balraj98/deepglobe-land-cover-classification-dataset>
- 6 classes | Image size: 2448p×2448p | 803 (train); 171 (val); 172 (test)
- Find smart patch sampling (256x256) procedure to build a sound MLC dataset
- Convert to LMDB files for quick training

3. Evaluate baseline/CbMLC for noise types individually/together

- Use RSiM Noise-Definition: https://git.tu-berlin.de/rsim/multi_label_noise/-/blob/master/noise_functions.py
- Use mAP micro/macro (not F1)
- Evaluate for additive/subtractive/mixed (balanced)
=> for 10%,30%,50%,70%
- Datasets: BigEarthNet-Ireland, MLC version of DeepGlobe

Tasks

4. Optional

- More datasets
- More base-line comparison Methods
- Customized CbMLC