RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

SONDRA
France-Singapore Lab

ONERA
THE FRENCH AEROSPACE LAB

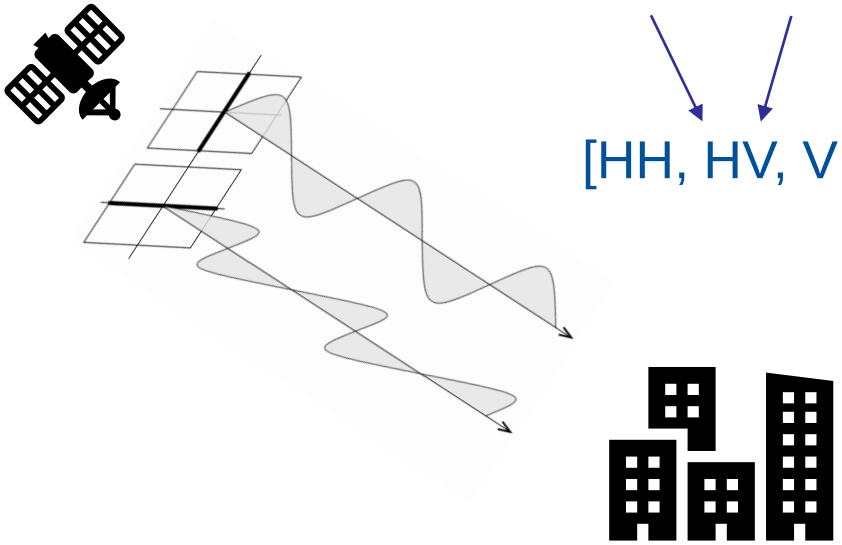
General Feature Extraction in SAR Target Classification: A Contrastive Learning Approach Across Sensor Types

Max MUZEAU[1,2], Joana FRONTERA-PONS[2], Chengfang REN[1], Jean-Philippe OVARLEZ[1,2]

[1]SONDRA, CentraleSupélec
[2]ONERA

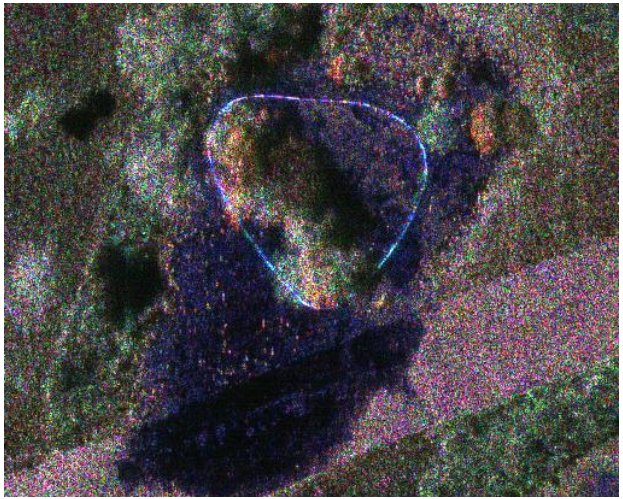
Synthetic Aperture Radar

Data type



horizontal vertical

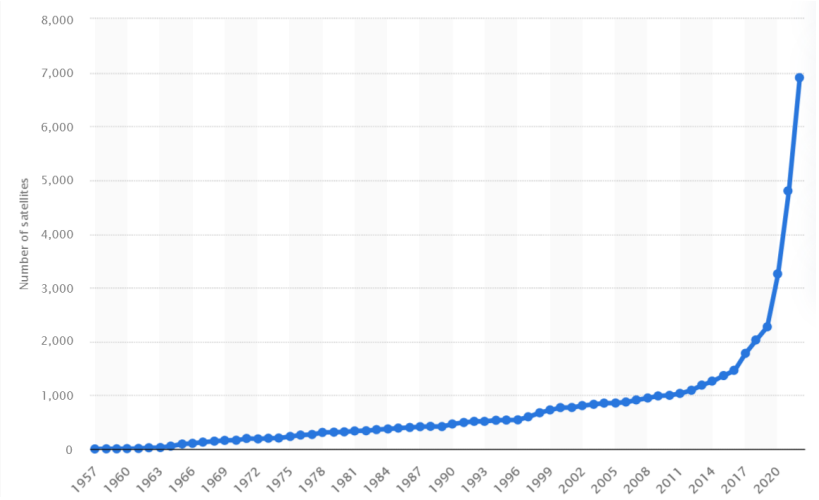
[HH, HV, VH, VV]



Sethi X-band image

Synthetic Aperture Radar

- Self-supervised learning:
  - Huge amount of data (TSX, Sentinel 1, UAVSAR, Biomass …)
  - Almost no ground truth - label



Number of satellites active from 1957 to 2022

Self-supervised learning

Objective

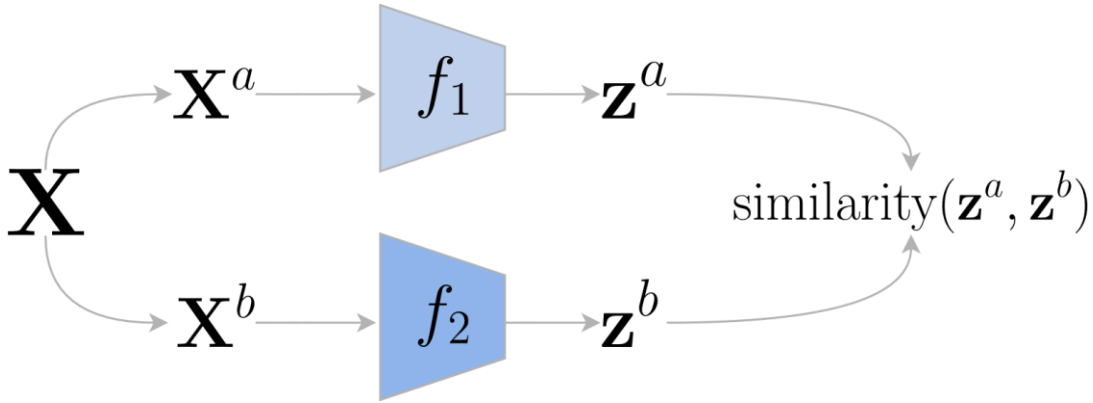
- Extract meaningfull features from image
- From unknown and known sensors

Dosovitskiy, Alexey, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." *arXiv preprint arXiv:2010.11929* (2020).
Caron, Mathilde, et al. "Emerging properties in self-supervised vision transformers." *Proceedings of the IEEE/CVF international conference on computer vision*. 2021.
Chiu, Li-Ling, and Shang-Hong Lai. "Self-Supervised Normalizing Flows for Image Anomaly Detection and Localization." *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.

Self-supervised learning

Siamese Networks

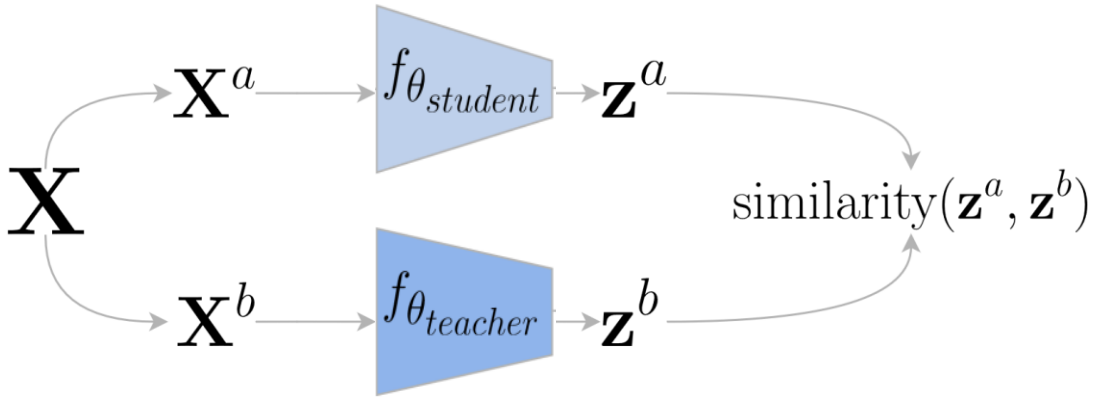
- Generate two augmented views of an image X
- Two networks encode in in a vector z
- These vector dimension have to be similar



Self-supervised learning

Siamese Networks

- One teacher and one student network
- Both trained from scratch



Self-supervised learning

Augmentations

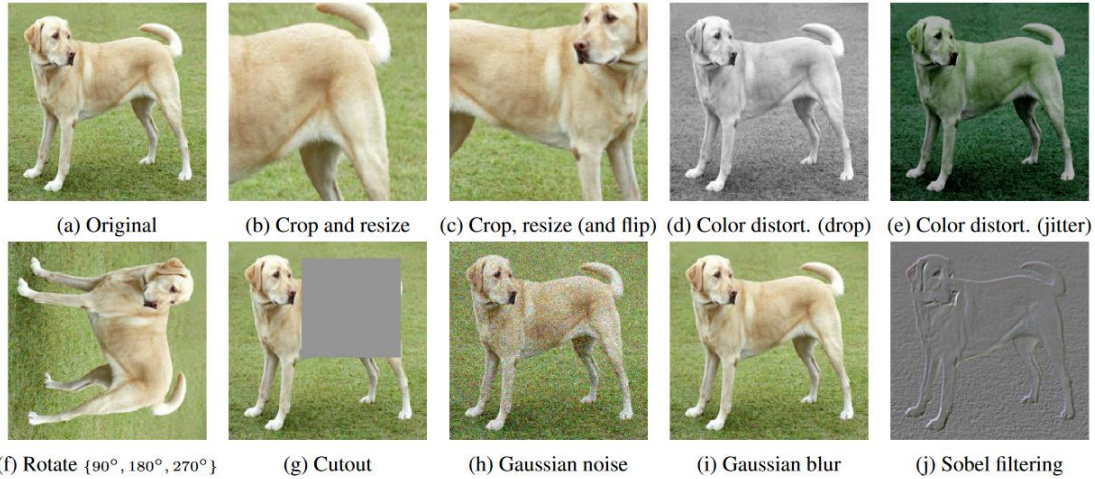


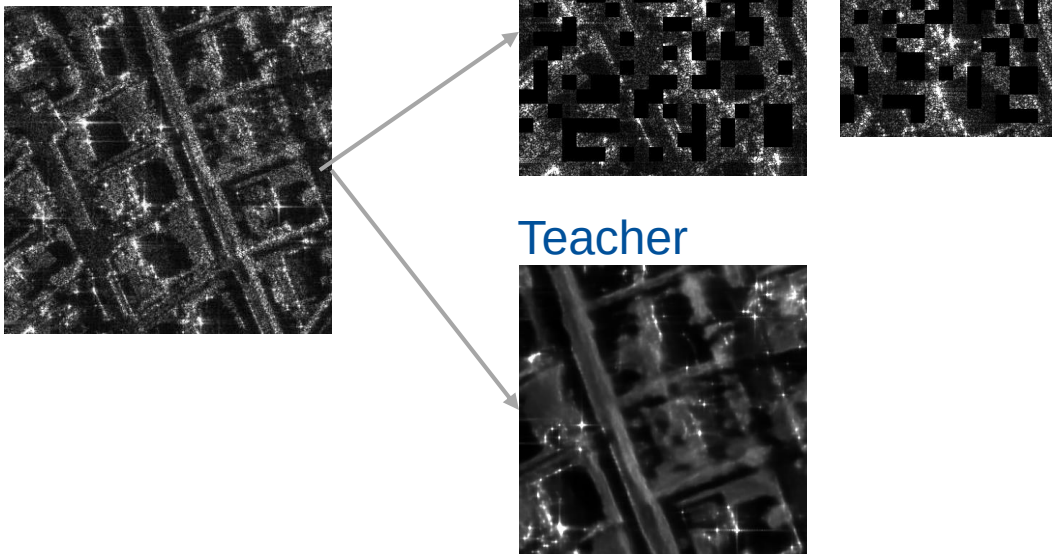
Figure 4. Illustrations of the studied data augmentation operators. Each augmentation can transform data stochastically with some internal parameters (e.g. rotation degree, noise level). Note that we *only* test these operators in ablation, the *augmentation policy used to train our models* only includes *random crop (with flip and resize)*, *color distortion*, and *Gaussian blur*. (Original image cc-by: Von.grzanka)

Chen, Ting, et al. "A simple framework for contrastive learning of visual representations." *International conference on machine learning*. PMLR, 2020.

Self-supervised learning

Augmentations

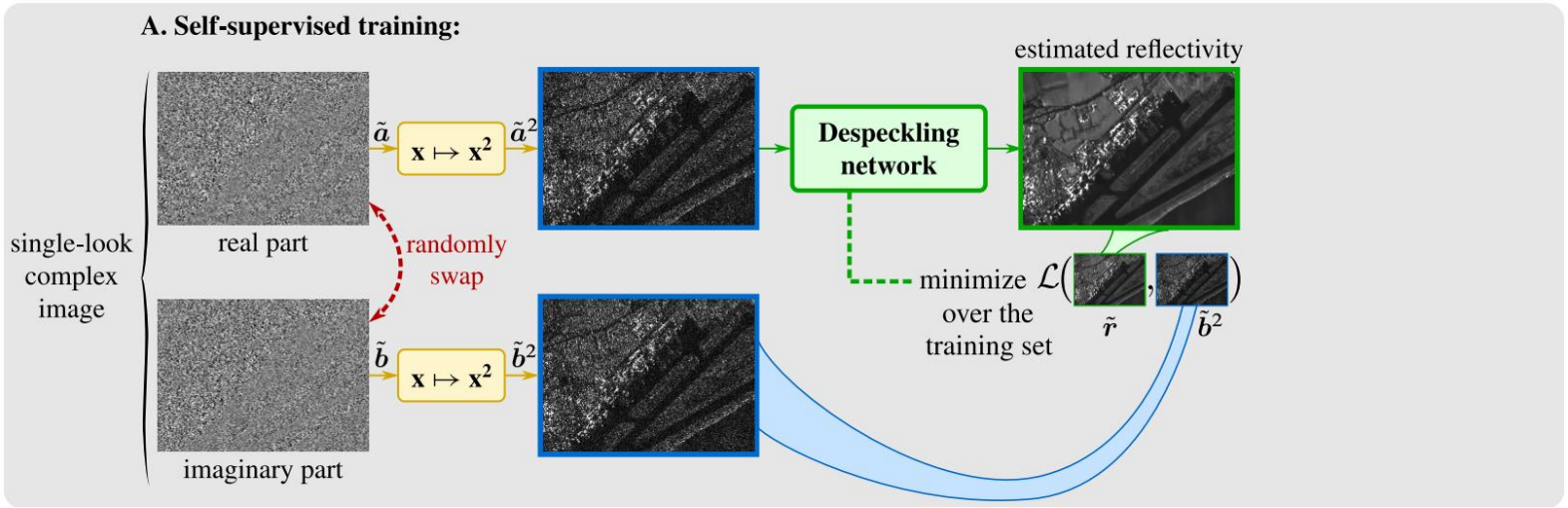
- Despeckling
- Global / local crop
- Masking
- Sub-bands/Sub-looks
- Intensity shift



- Trained only on SETHI data

Augmentations

Despeckling - MERLIN



$$\underset{z}{\operatorname{argmin}}\, \mathbb{E}_y\{L(z, y)\} \longrightarrow z = \mathbb{E}_y\{y\}$$

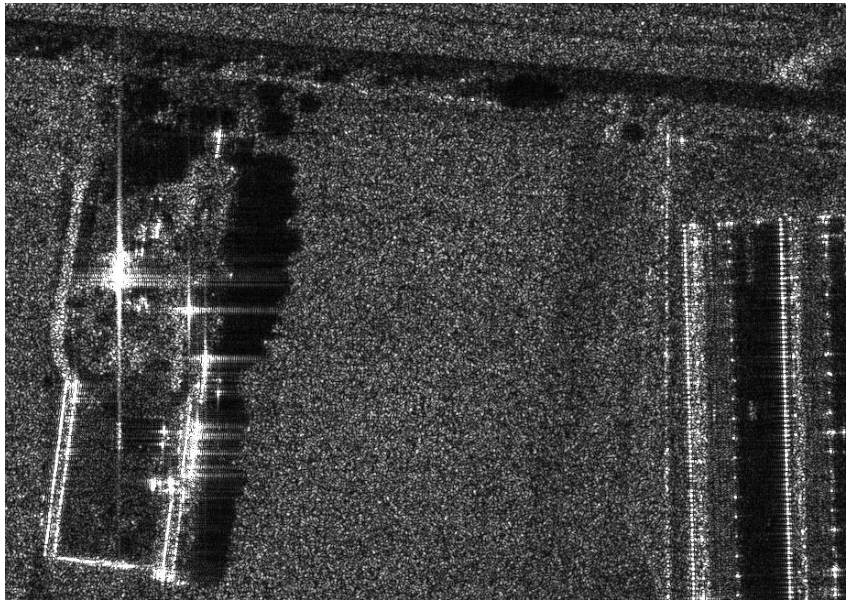
DALSASSO, Emanuele, DENIS, Loïc, et TUPIN, Florence. As if by magic: self-supervised training of deep despeckling networks with MERLIN. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, vol. 60, p. 1-13.

DALSASSO, Emanuele, DENIS, Loïc, MUZEAU, Max, *et al.* Self-supervised training strategies for SAR image despeckling with deep neural networks. In : *EUSAR 2022*

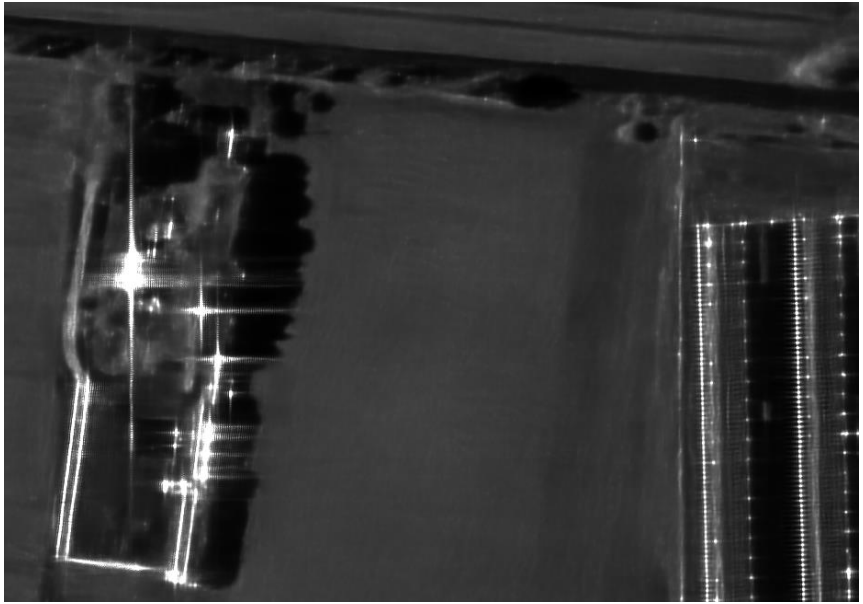
Augmentations

Despeckling - MERLIN

Noisy

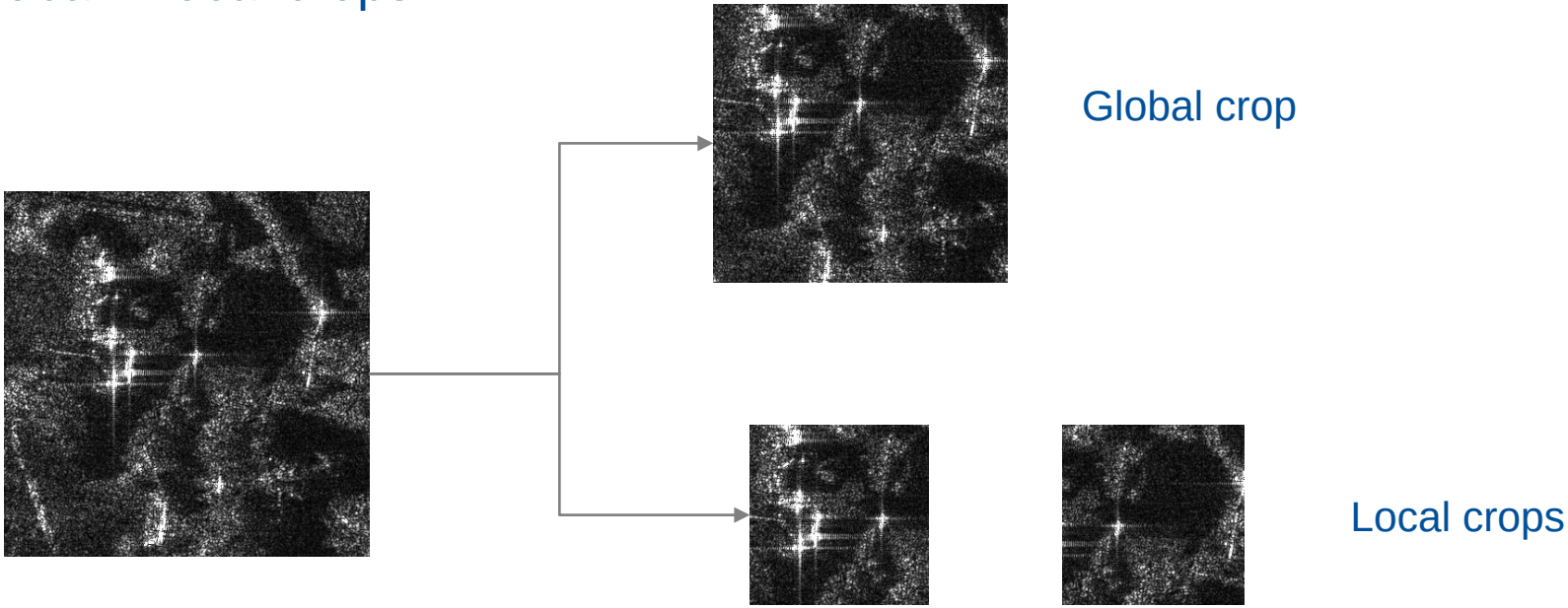


Denoised



Augmentations

Global – local crops

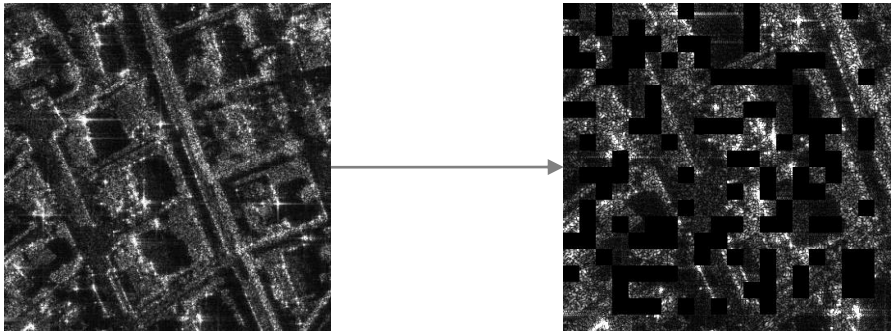


Augmentations

Masking tokens

- Random mask
- Transformer tokens
- Better scalability

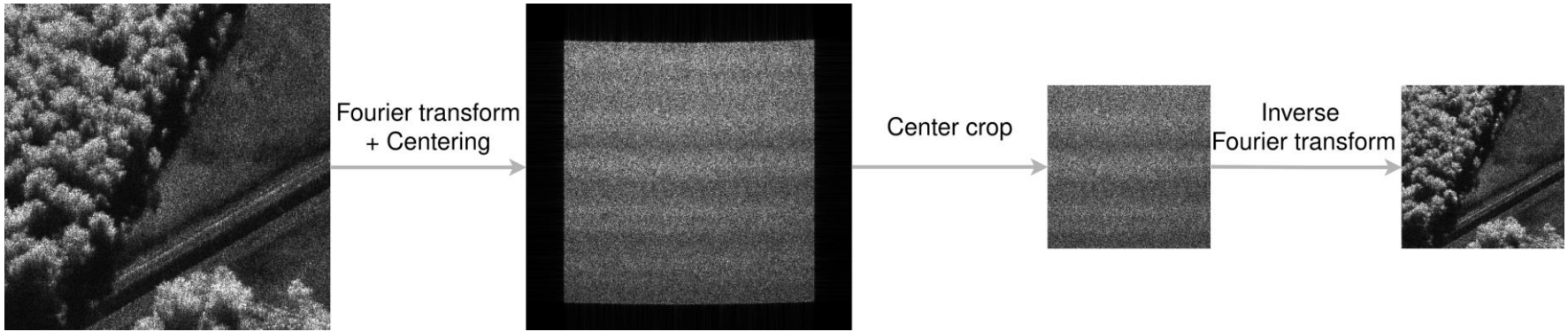
$[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n] \in \mathbb{R}^{d_f \times n}$ $[\mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_m] \in \mathbb{R}^{d_f \times m}$



Assran, Mahmoud, et al. "Masked siamese networks for label-efficient learning." *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 2022.

Augmentations

Sub-aperture



Augmentations

Intensity shift

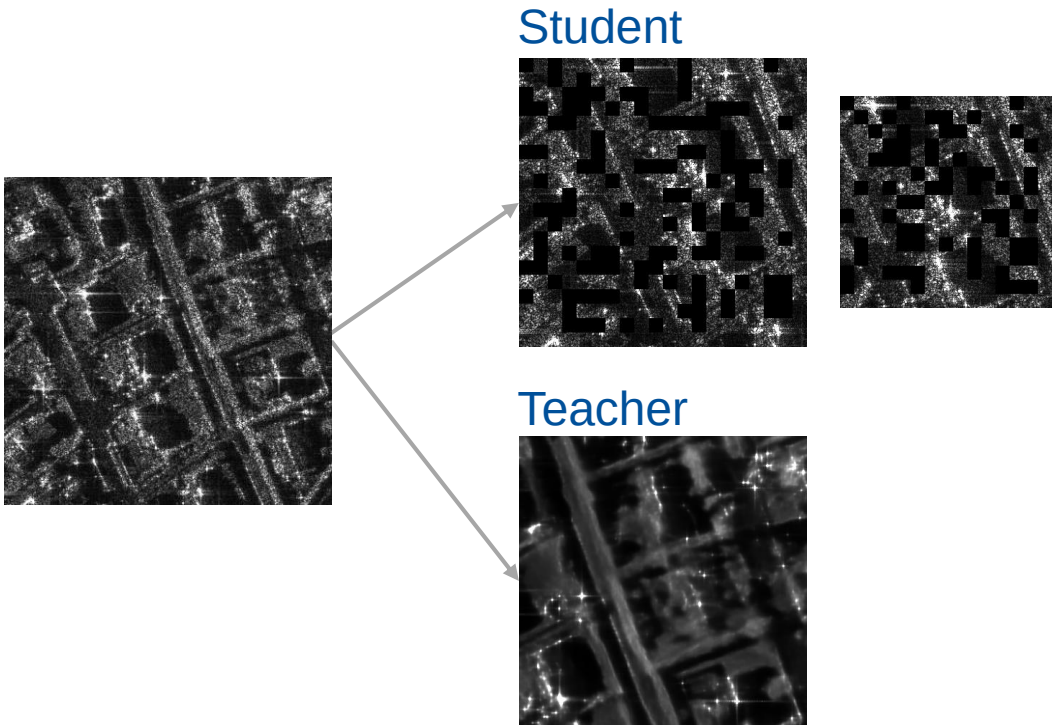
- Rely less on intensity
- More on structure information
- Parameter values depend on the use case

$$\mathbf{x}_{shifted} = \mathbf{x} + B \text{ ,where } \quad B \sim \mathcal{U}(a, b)$$

Self-supervised learning

Augmentations

- Despeckling
- Global / local crop
- Masking
- Sub-bands/Sub-looks
- Intensity shift

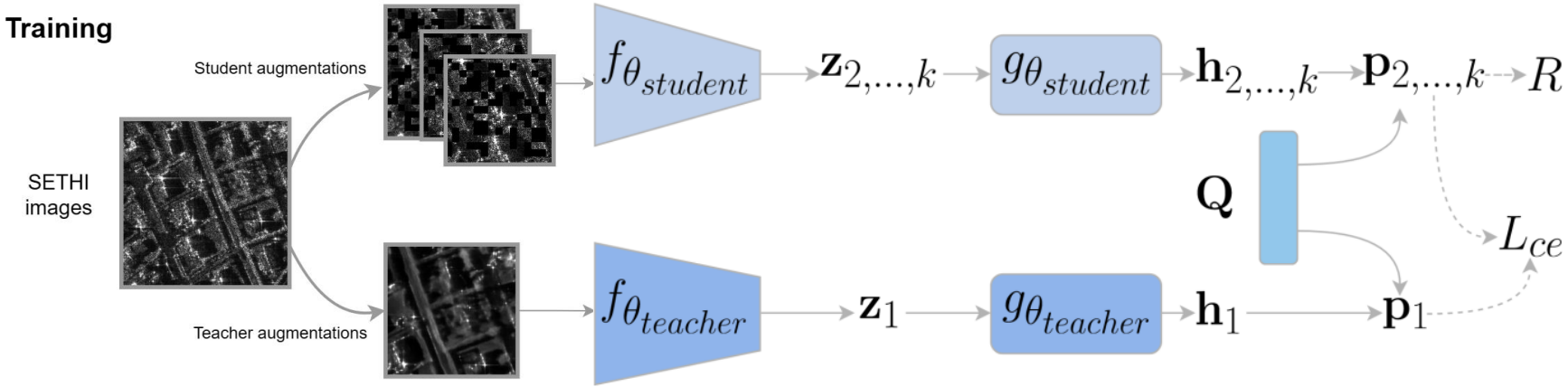


- Trained only on SETHI data

Self-supervised learning

Architecture

- Sethi X and L images for training

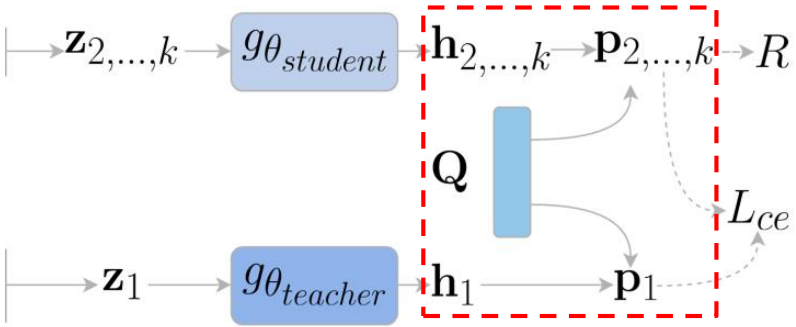


Self-supervised learning

Training

- Projection on a set of prototypes :
  - Force clustering
  - Few-shot learning

$$\mathbf{p} = \text{softmax}\left(\frac{\mathbf{s}}{\tau}\right) \text{ with } \left\{ \mathbf{s}_i = \frac{\mathbf{q}_i^T \mathbf{h}}{\|\mathbf{q}_i\|_2 \|\mathbf{h}\|_2} \right\}_{i \in [1,n]},$$

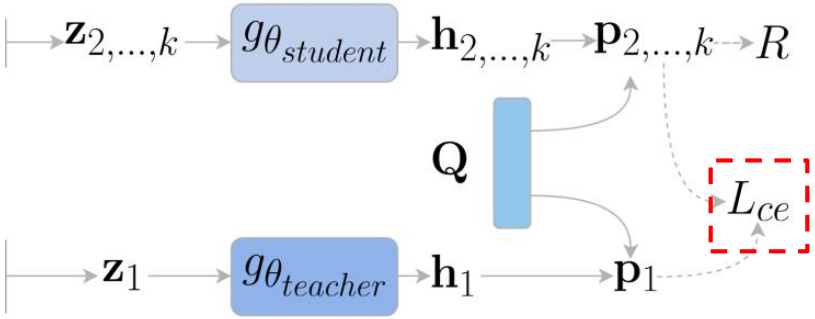


Self-supervised learning

Training

- Training losses:
  - Cross entropy

$$L_{ce} = \frac{1}{b\,(k-1)} \sum_{i=1}^{b} \sum_{j=2}^{k} \sum_{l=1}^{n} -\mathbf{p}_{i,1}^{l} \log \mathbf{p}_{i,j}^{l}\,,$$

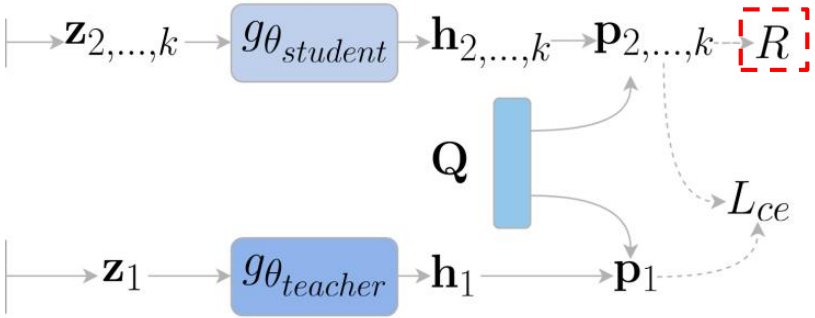


Self-supervised learning

Training

- Training losses:
  - Cross entropy
  - mean-entropy maximization regularizer

$$R = -\sum_{l=1}^{n} \overline{\mathbf{p}}^{l} \log \overline{\mathbf{p}}^{l} \text{ with } \overline{\mathbf{p}} = \frac{1}{b\,(k-1)} \sum_{i=1}^{b} \sum_{j=2}^{k} \mathbf{p}_{i,j}\,,$$



Self-supervised learning

Training

- Update network weights
  - Backpropagation for the student
  - Moving average update for the teacher

$$\theta_{teacher} \leftarrow m\,\theta_{teacher} + (1-m)\,\theta_{student}.$$

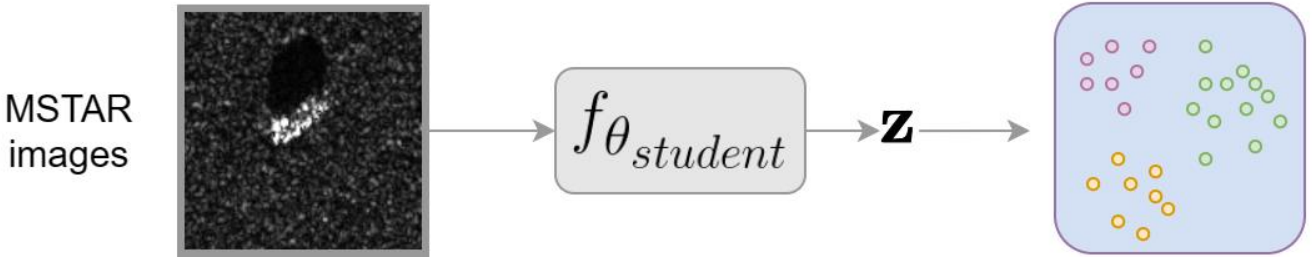
Experiences

Target classification

- Quantitative evaluation on MSTAR dataset
- k-NN on extracted features
- Comparison with a PCA and a ResNet-34

MSTAR dataset

Class	2S1	BRDM_2	BTR_60	D7
Number	1664	1282	451	573
Class	T62	ZIL131	ZSU_23_4	SLICY
Number	572	573	1401	2539

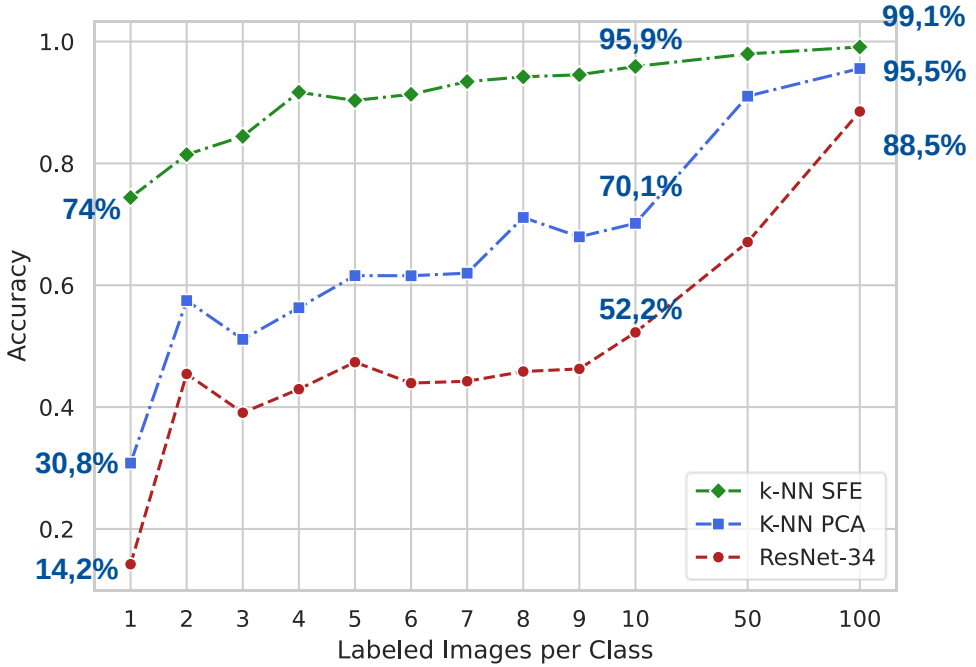


Experiences

Target classification

- Few-shot classification

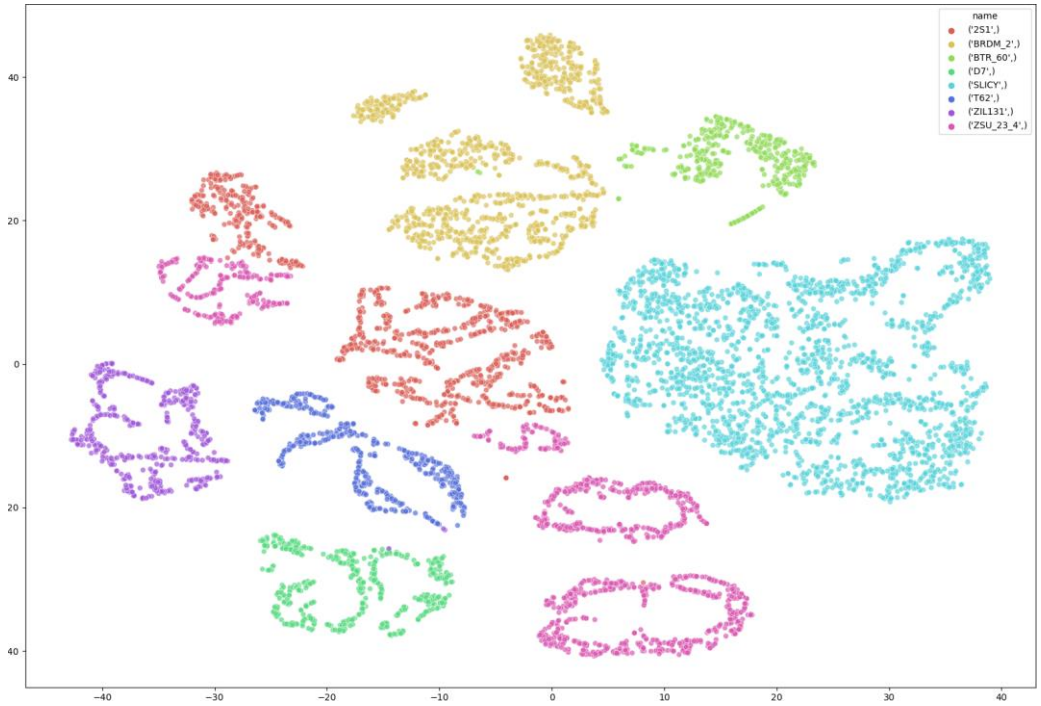
- For 10 labels/class the accuracy is:
  - 43,7% higher than a ResNet
  - 25,9% higher than a PCA



Experiences

Features visualisation

- Not trained on MSTAR
- t-SNE visualisation



Van der Maaten, Laurens, and Geoffrey Hinton. "Visualizing data using t-SNE." *Journal of machine learning research* 9.11 (2008)

Open source code

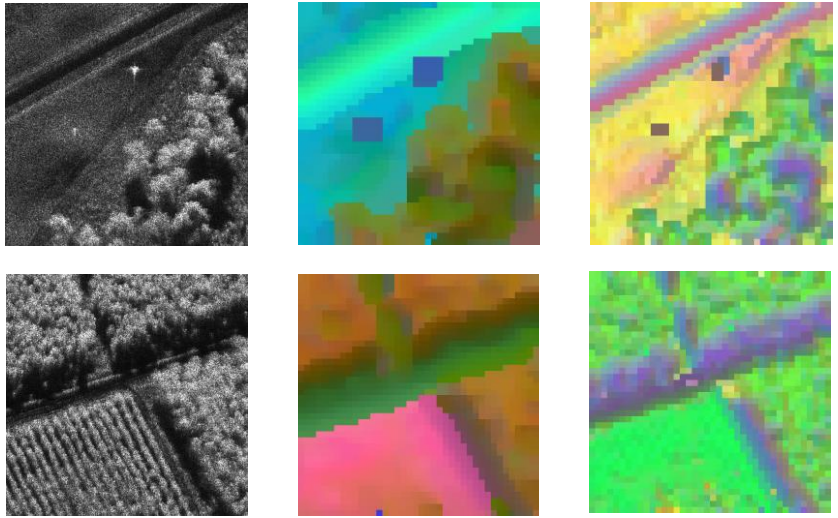
SAR feature extraction: https://github.com/muzmax/MSTAR_feature_extraction

Under apache 2.0 license

Journal paper

- Extended journal article (segmentation, visualisation, pattern detection)
  https://arxiv.org/abs/2407.00851

SAFE: a SAR Feature Extractor based on self-supervised learning and masked Siamese ViTs



RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

SONDRA
France-Singapore Lab

ONERA
THE FRENCH AEROSPACE LAB

Questions, comments, discussions?

Contact: max.muzeau@centralesupelec.fr

code

paper

Max MUZEAU, Jean-Philippe OVARLEZ, Chengfang REN, Joana FRONTERA-PONS