UNIVERSITÀ DI TRENTO

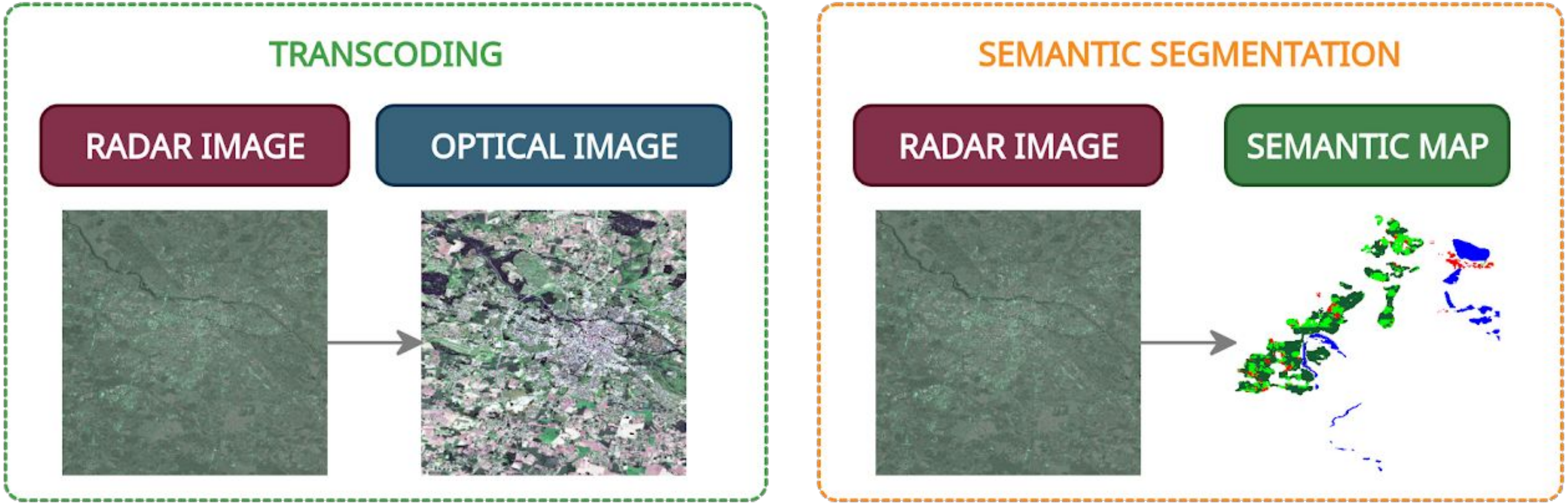
DLR

Self-Supervised Learning for Semantic Segmentation of Pol-SAR Images Via SAR-to-Optical Transcoding

Trento, 17/03/2021
Master's Degree in
Information and Communication Engineering

Student: Alessandro Cattoi
Co-Supervisor: Ronny Hänsch (DLR)
Supervisor: Lorenzo Bruzzone (UNITN)

Terminology

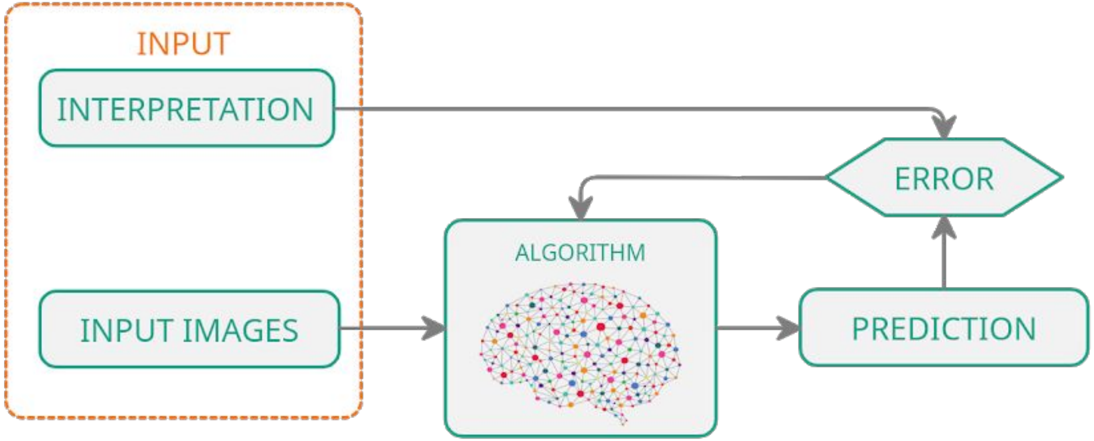


Outline

- Motivation
- Objective
- Methodology
- Transcoding Results
- Semantic Segmentation Results
- Label Sensitivity Analysis
- Takeaways

Automatic Images Interpretation

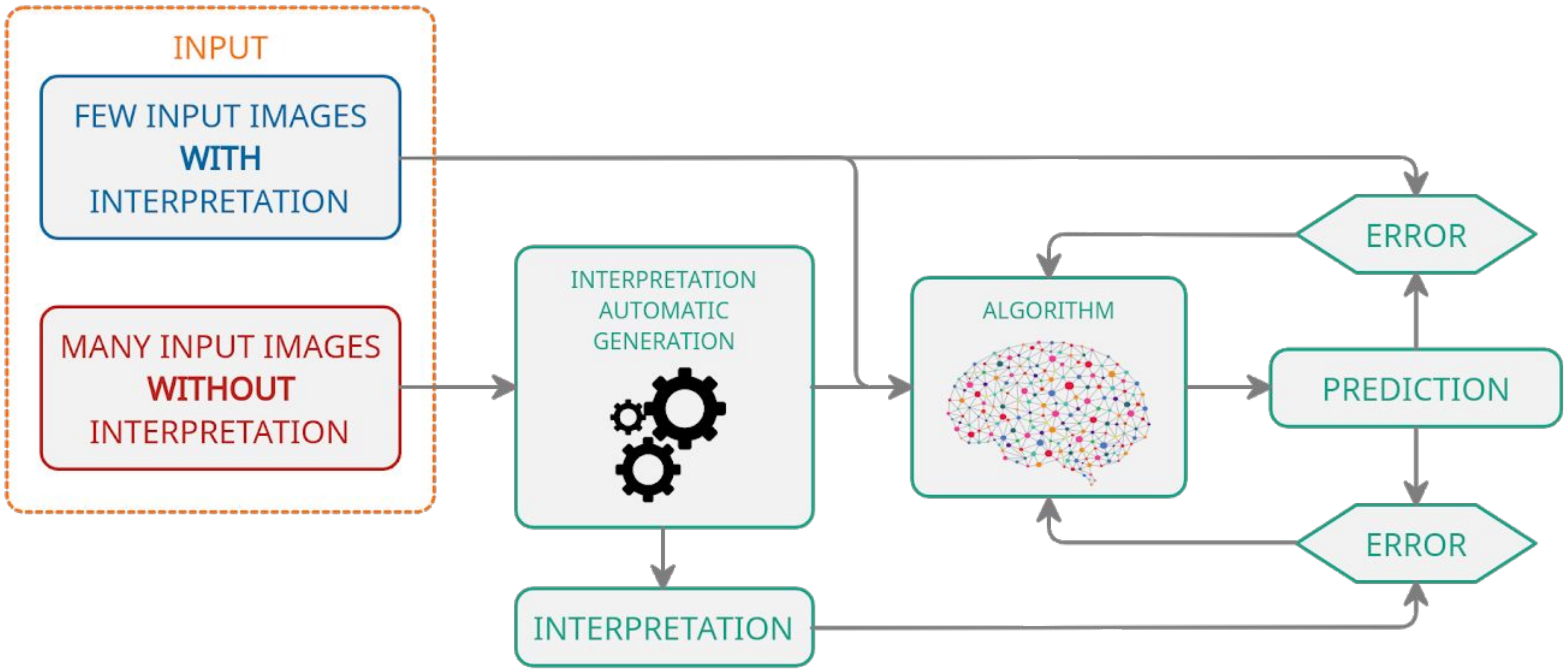
- **Supervised** Deep Artificial Neural Networks are the **state of the art** in the field of image classification.
- They require **many data** to achieve strong performances.
- The data have to be **accompanied with** their **interpretation**.



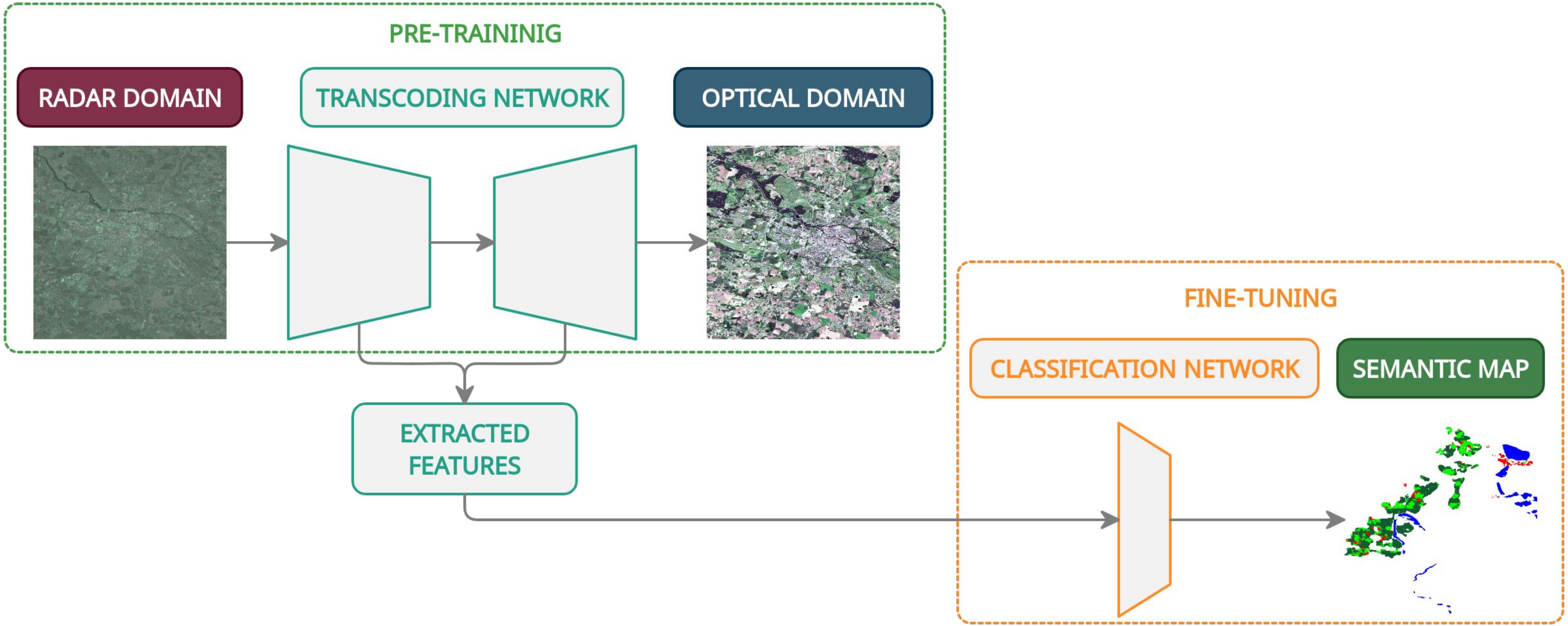
Thesis Context

- **Pol-SAR** images **annotations procurement** is particularly **complex**
- Interpreting a Pol-SAR images **requires expertise**
- **Automatic** radar images **classification requires a lot of data**
- This **thesis** investigates a technique to i**nterpret data** when **annotation** procurement is **costly**

Self-Supervised Approach



Objective


PRE-TRAININIG
RADAR DOMAIN
TRANSCODING NETWORK
OPTICAL DOMAIN
EXTRACTED FEATURES
FINE-TUNING
CLASSIFICATION NETWORK
SEMANTIC MAP

Objective

- Pol-SAR to Optical transcoding as features extractor.
- Self-supervised pre-trained models VS supervised.
- Training strategy comparison.
- Pre-trained models resilience to data scarcity.

Related Works

- **[Ley et al. 2018]** employ **self-supervision** to perform **semantic segmentation**
- **[Reyes et al. 2019]** employ **Cycle-GAN** to analyze **transcoding product quality**
- **[Saha et al. 2020]** employ Cycle-GAN **transcoding** for **change detection**

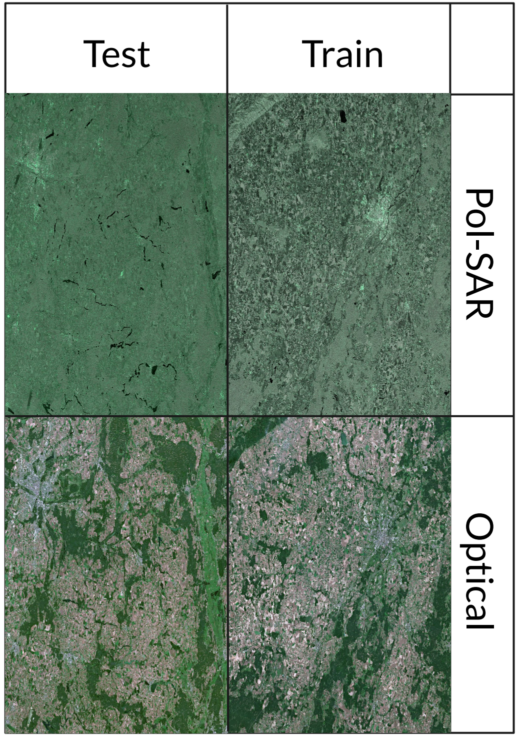
Data

Data origins:

- **Radar** Images come from **Sentinel-1.**
- **Optical** Images comes from **Sentinel-2**.
- All images have a resolution of **10 m x 10 m** x pixels.

Data Splitting:

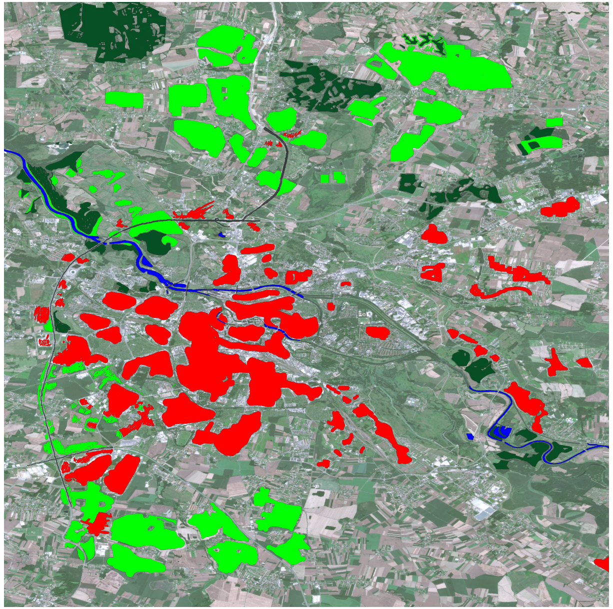
- **Train set** is acquired over the **Poznan** (Poland).
- **Test set** acquired over **Wroclaw** (Poland).
- All images have been **acquired in similar time** periods.



Labeled Data

Both test and train image are classified with the following categories:

- Highways ➡️ Grey
- Fields ➡️ Light green
- Urban ➡️ Red
- Water ➡️ Blue
- Forests ➡️ Dark Green



Pol-SAR Data Preprocessing

- **Pol-SAR** data are for their nature **complex-valued**.
- From polarimetric data a **real valued feature vector *v*** is extracted after computing the covariance matrix *C*:

$$v = \begin{pmatrix} log(C_{1,1}) \\ log(C_{2,2}) \\ log(|C_{1,2}|) \\ \frac{\mathbb{R}(C_{1,2})}{|(C_{1,2})|} \\ \frac{\mathbb{I}(C_{1,2})}{|(C_{1,2})|} \end{pmatrix}$$

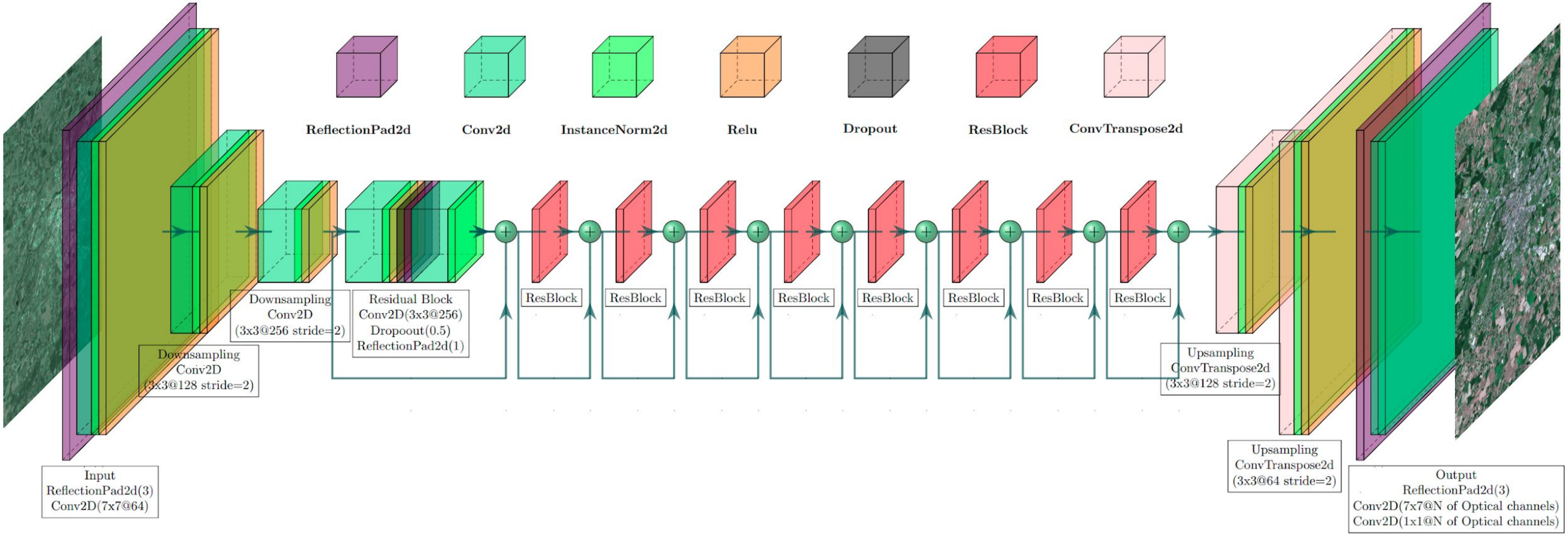
Transcoders Implementation

The **transcoding** phase has been performed using **three** different training **strategies**:

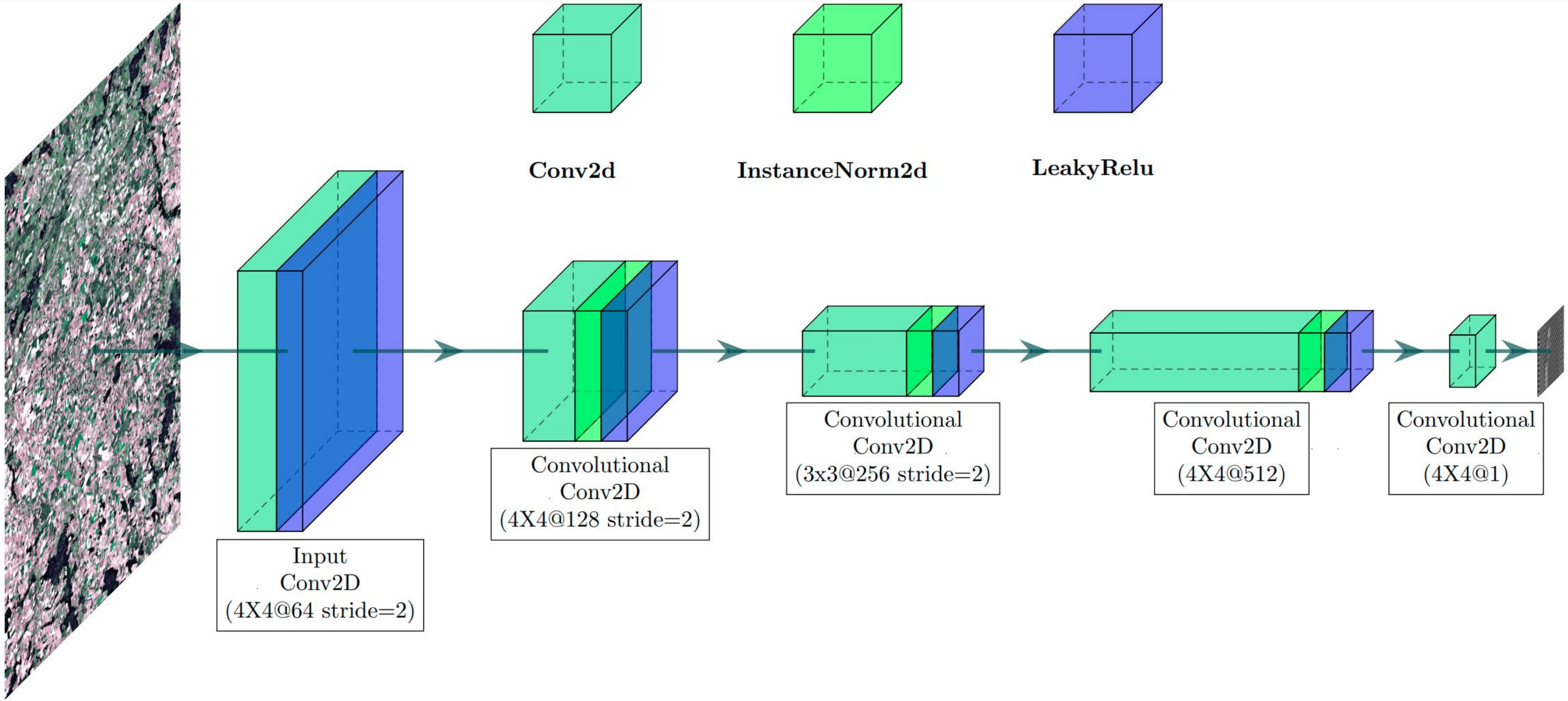
- **Regressive Transcoder (RT)**
- **Conditional Adversarial Transcoder (CAT)**
- **Cycle Consistent Adversarial Transcoder (Cycle-AT)**

All these implementations are based on the **same** network **architectures**, called generators and discriminators.

Generator Architecture



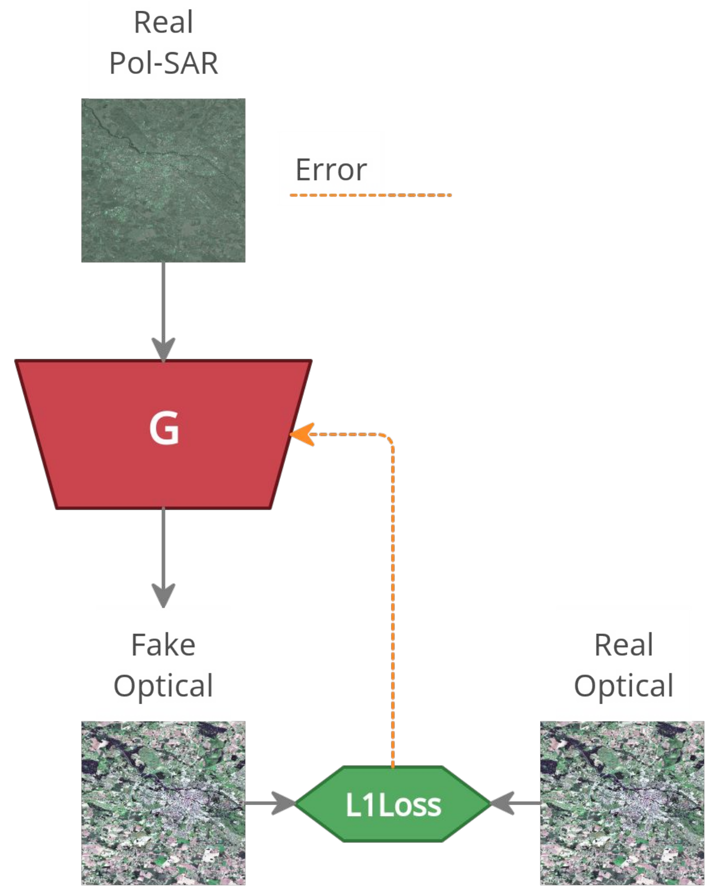
Discriminator Architecture



UNIVERSITÀ DI TRENTO

Regressive Transcoder (RT)

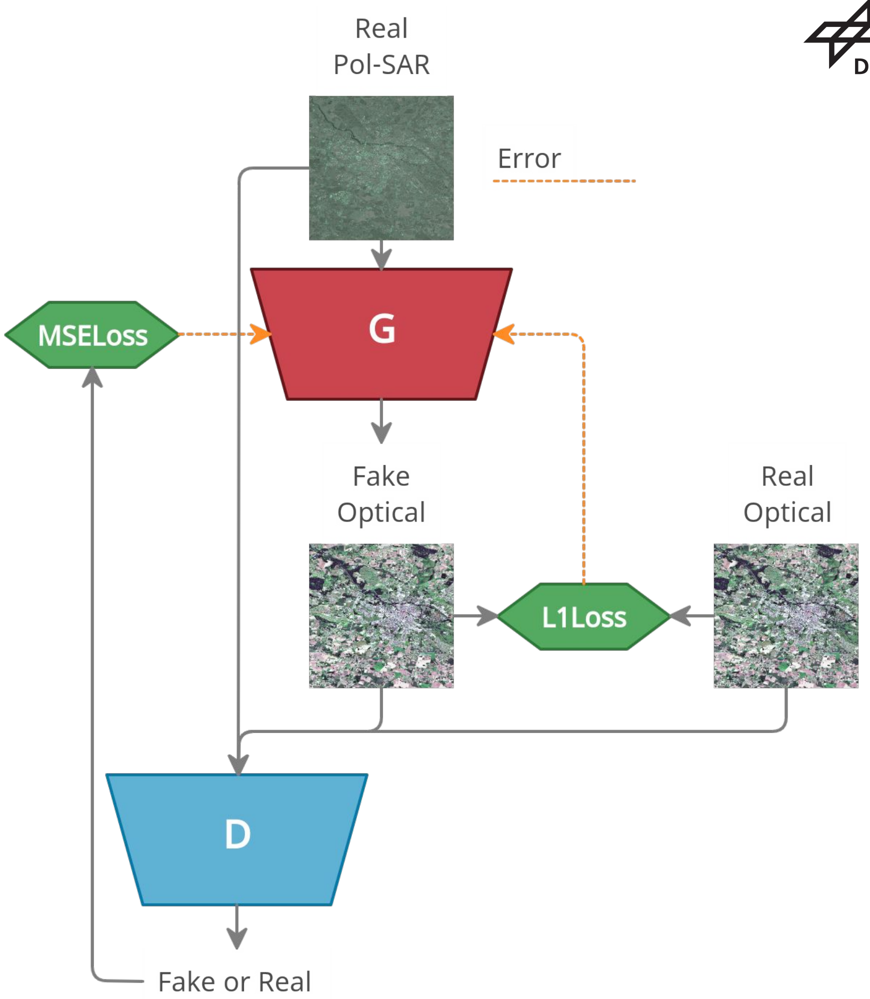
- **Simplest** training strategy, pure regressor.
- **Supervised**.
- **Based** on one **generator** architecture.
- **L1 loss** between the **generated sample** and the actual version of the **input radar image**.



UNIVERSITÀ DI TRENTO

Conditional Adversarial Transcoder (CAT)

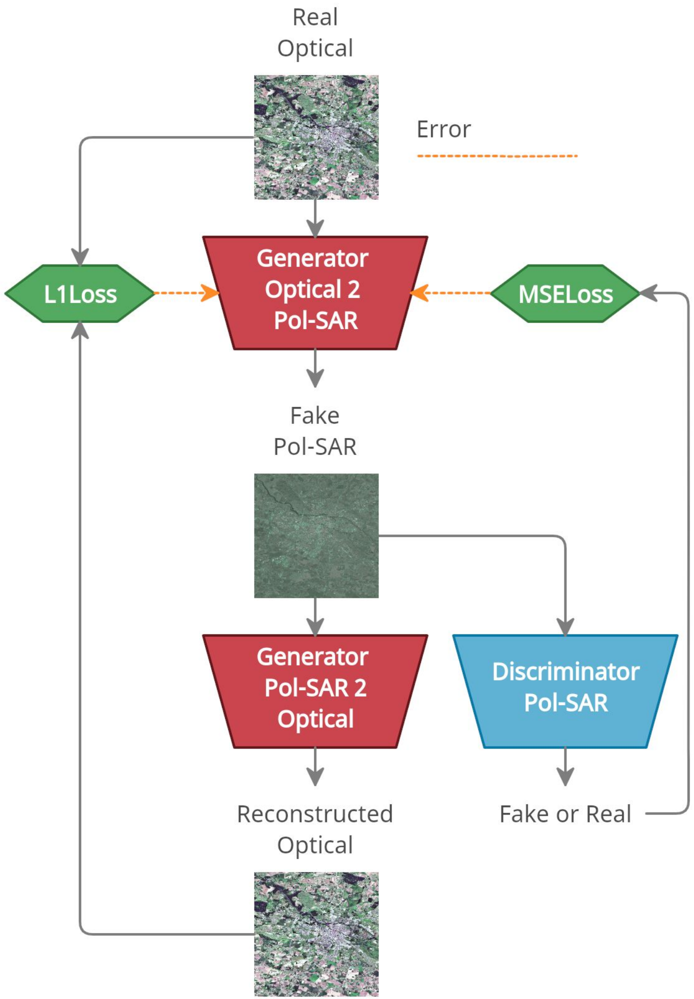
- **Conditional generative adversarial** training strategy.
- **Based** on one **generator** architecture and one **discriminator** architecture.
- **Supervised** algorithm.
- **L1 loss** between the **generated sample** and the actual version of the **input radar image.**
- **Adversarial loss** generated using discriminator D.



UNIVERSITÀ DI TRENTO

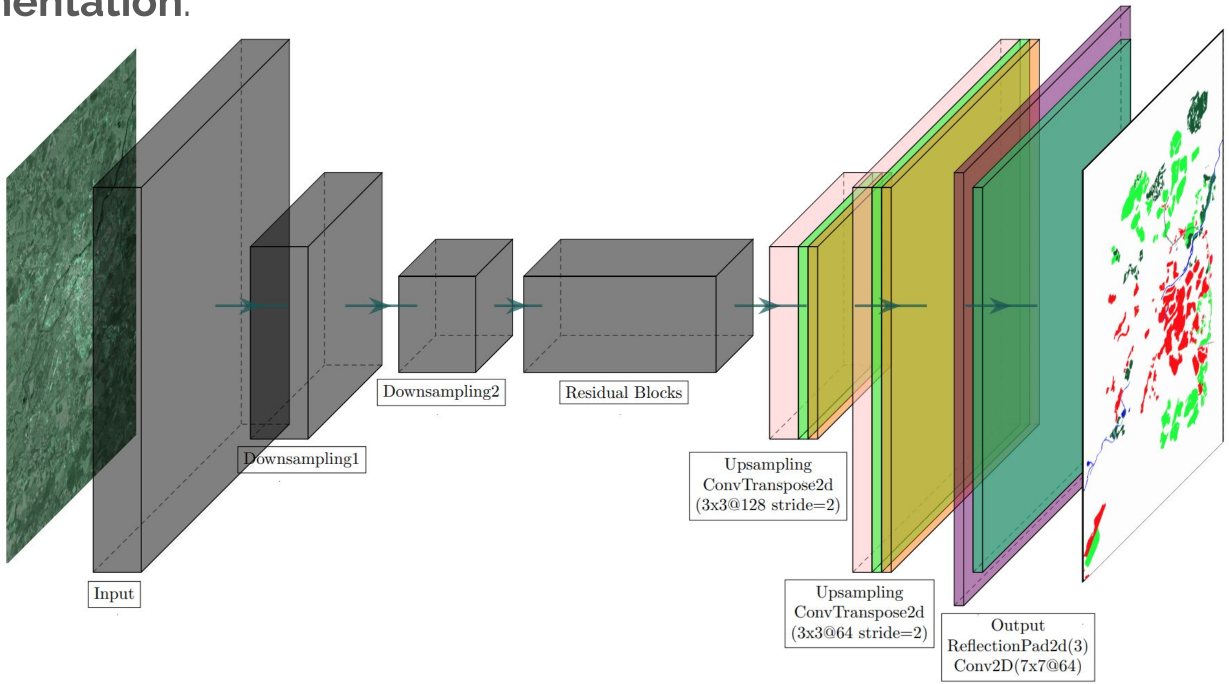
Cycle Consistent Adversarial Transcoder (Cycle-AT)

- **Cycle consistent generative adversarial** training strategy.
- **Based** on two **generator** and two **discriminator** architectures.
- **Unsupervised** algorithm.
- **L1 loss** between the **reconstructed sample** and the actual version of the **input radar image**.
- **Adversarial loss** generated using the two discriminators.
- Does **not** require **paired images**.



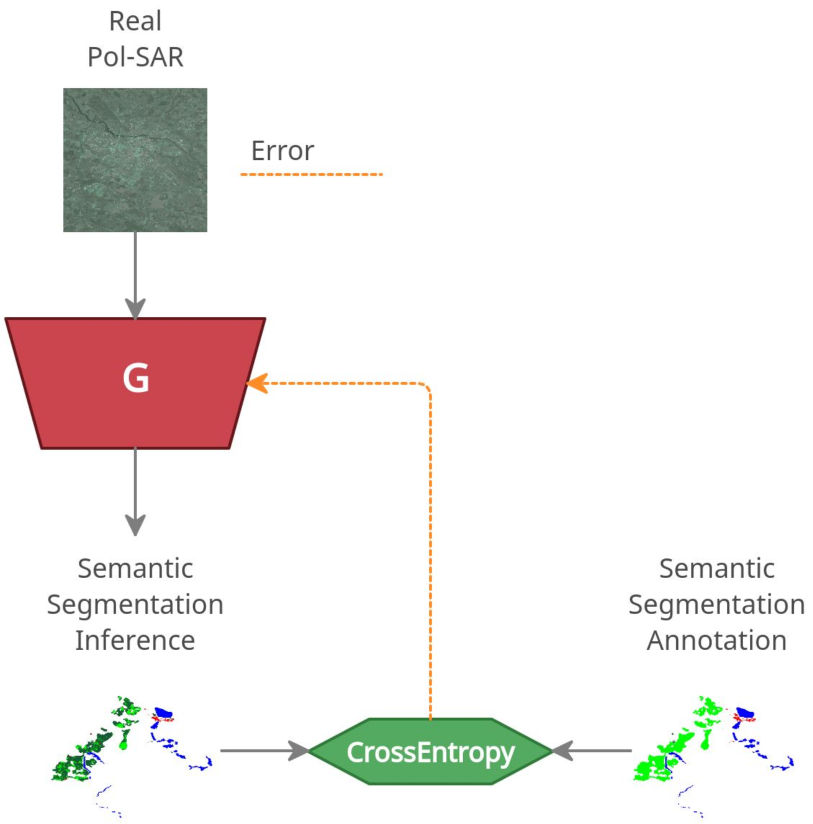
Semantic Segmentation Implementation

- Each **transcoder** after being trained has been **fine-tuned** to perform **semantic segmentation**.



Baseline (BL) / Fine-Tuning

- Fully supervised.
- **Based** on the **generator** architecture.
- **Cross entropy loss** between **predicted map** and **label map**.
- Require **fully annotated dataset**.



Networks Complexity

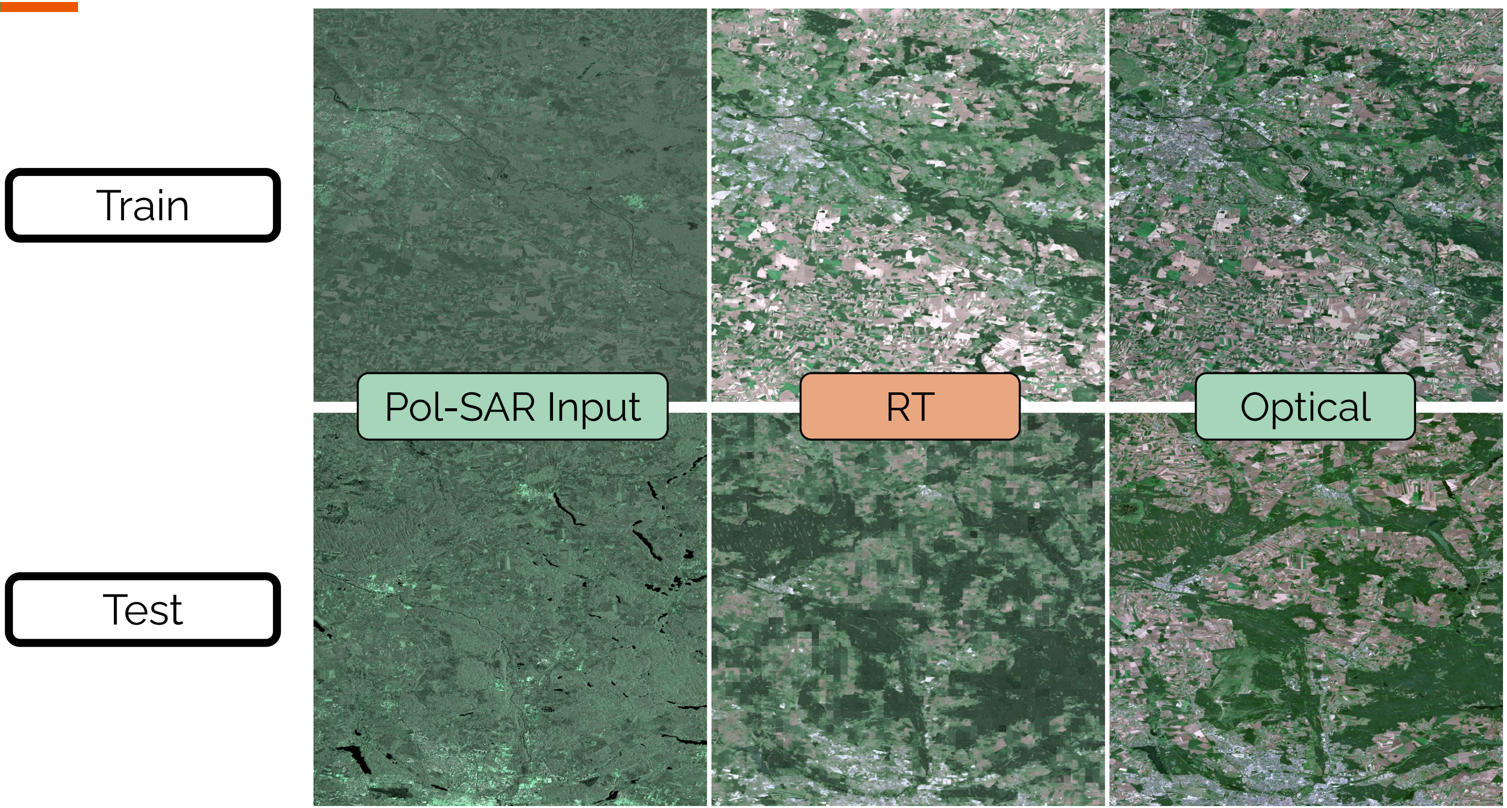
Network	Number of Generators	Number of Discriminators	Number of Parameters [M]
BL	1	-	10.8
RT	1	-	11.2
CAT	1	1	14.2
Cycle-AT	2	2	28.3

Optimal Configuration

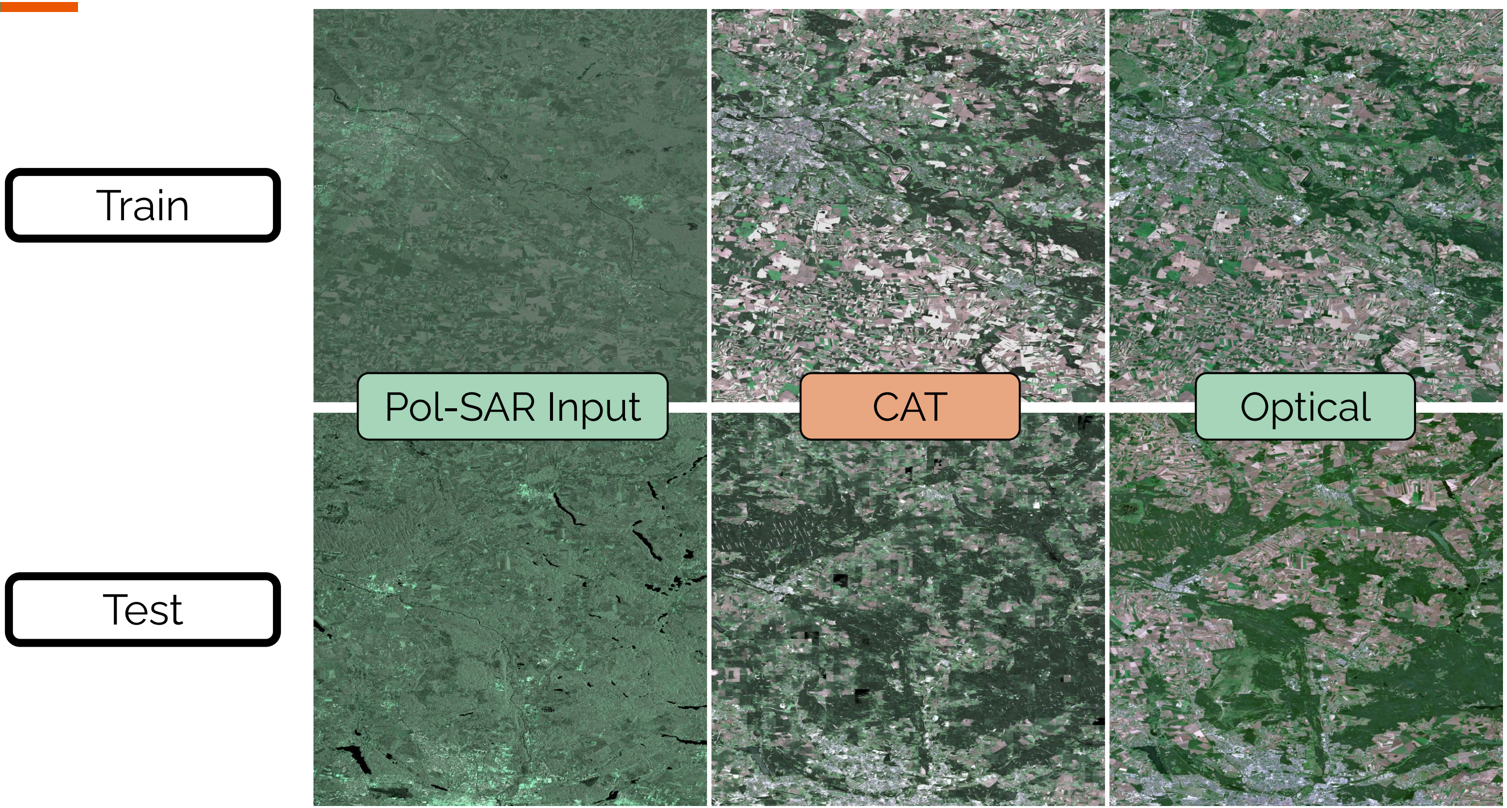
Implementation Name	Patch Size		Generator Deepness		Discriminator Deepness	
BL	128	256	9	6	X	
RT	128	256	9	6	X	
CAT	128	256	9	6	3	2
Cycle-AT	128	256	9	6	3	2
Final	***128***		***9***		***3***	

Performance Change
Worse
Equal
Better

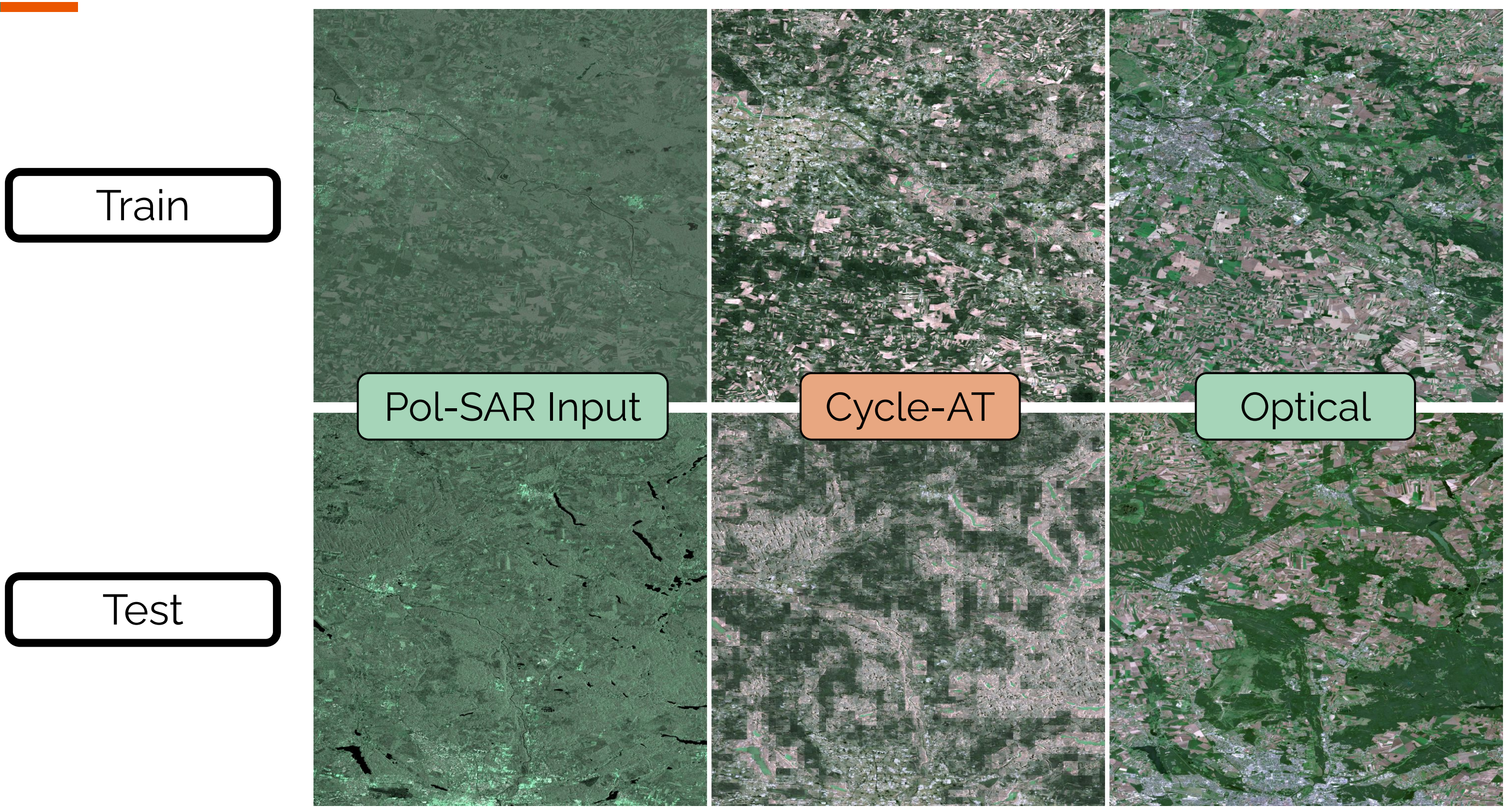
RT - Synthetic Optical Images Overview



CAT - Synthetic Optical Images Overview



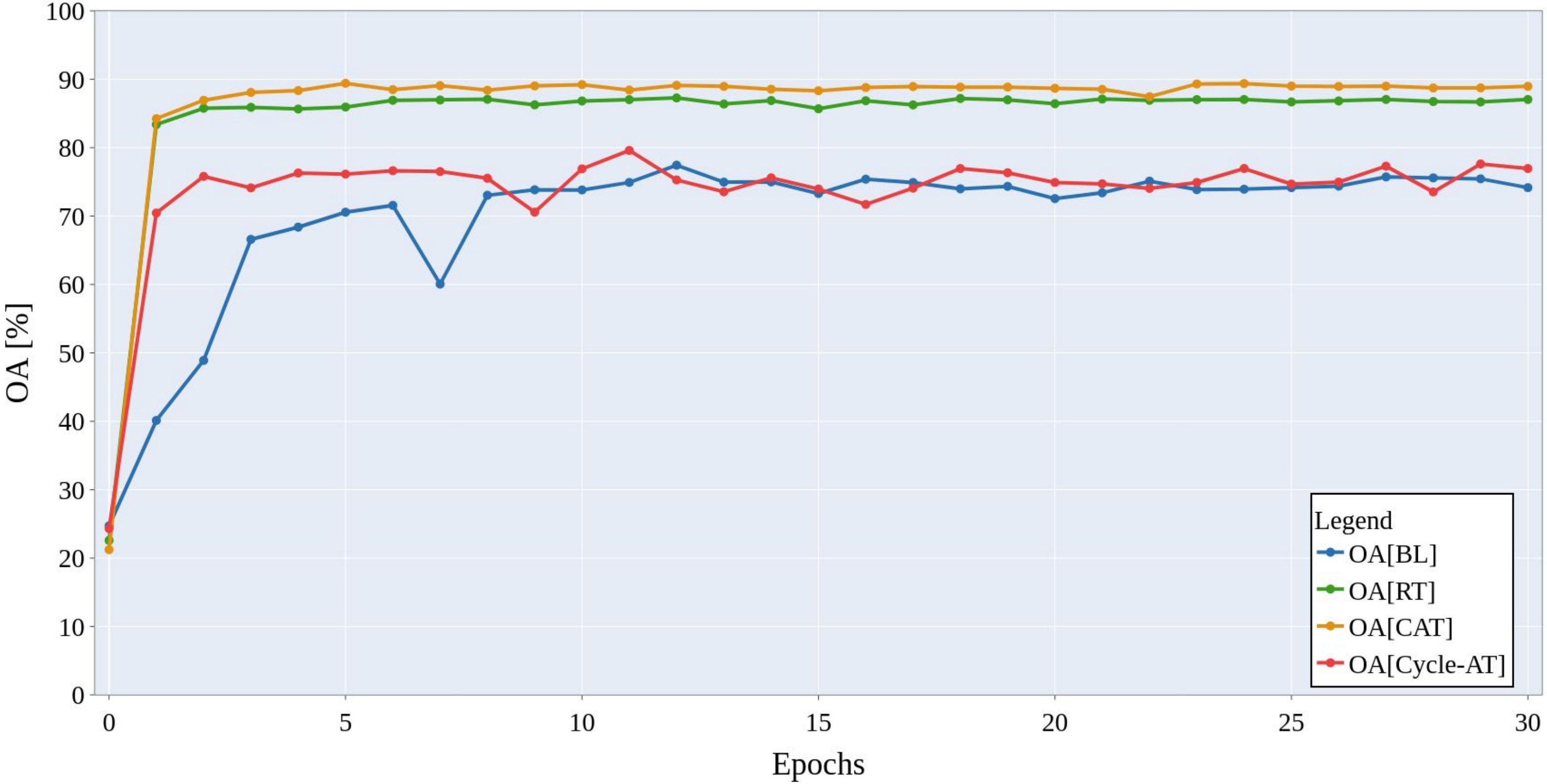
Cycle-AT - Synthetic Optical Images Overview



Transcoding Images Details - Test Set

Pol-SAR | RT | CAT | Cycle-AT | Optical

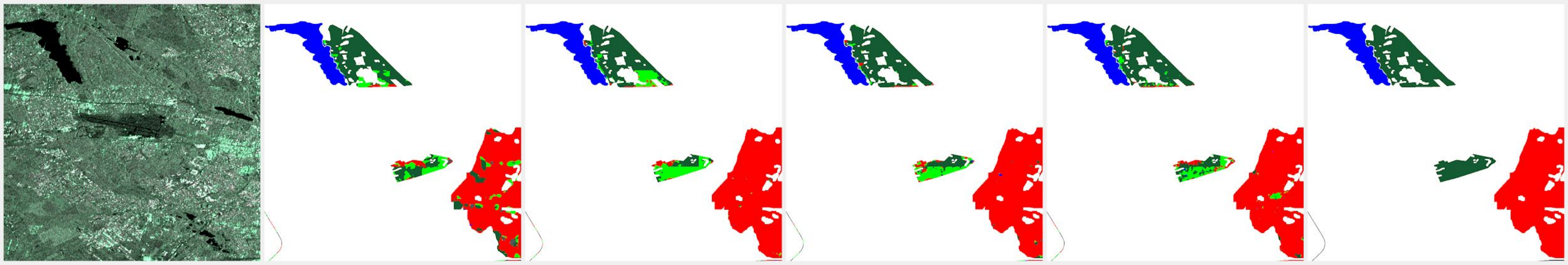
Semantic Segmentation Test Set Results



Classification Maps - Test Set Sample

Highways | Fields | Urban | Water | Forests

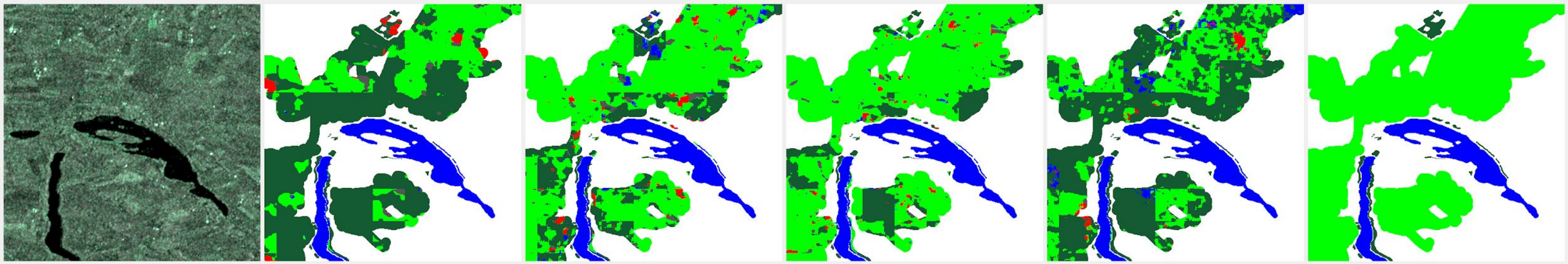
Pol-SAR Input | BL - **73%** | RT - **87%** | CAT - **89%** | Cycle-AT - **76%** | Label



Classification Maps - Test Set Sample 2

Highways | Fields | Urban | Water | Forests

Pol-SAR Input | BL - **73%** | RT - **87%** | CAT - **89%** | Cycle-AT - **76%** | Label



BL - Confusion Matrix

PREDICTIONS

LABELS	Highways	Fields	Urban	Water	Forests	Support
Highways	31.0 %	44.6 %	12.8 %	3.4 %	8.1 %	27768
Fields	2.8 %	37.2 %	1.2 %	0.0 %	58.8 %	414081
Urban	2.1 %	3.8 %	88.6 %	0.1 %	5.4 %	196265
Water	0.3 %	0.3 %	0.1 %	99.0 %	0.4 %	272931
Forests	1.5 %	7.0 %	1.3 %	0.3 %	89.9 %	376029
Support	30462	200782	187587	272612	595631	1287074

90 80 70 60 50 40 30 20 10 0

RT - Confusion Matrix

PREDICTIONS / **LABELS**

LABELS \ PREDICTIONS	Highways	Fields	Urban	Water	Forests	Support
Highways	44.0 %	35.2 %	14.7 %	3.8 %	2.3 %	27768
Fields	1.9 %	76.7 %	0.7 %	1.2 %	19.5 %	414081
Urban	0.7 %	1.1 %	97.8 %	0.0 %	0.3 %	196265
Water	0.1 %	0.2 %	0.1 %	99.2 %	0.4 %	272931
Forests	0.5 %	4.6 %	0.8 %	0.4 %	93.6 %	376029
Support	23291	347728	202269	278732	435054	1287074

90 80 70 60 50 40 30 20 10 0

CAT - Confusion Matrix

PREDICTIONS

LABELS	Highways	Fields	Urban	Water	Forests	Support
Highways	28.0 %	48.0 %	16.3 %	4.4 %	3.4 %	27768
Fields	1.2 %	80.8 %	1.4 %	0.3 %	16.3 %	414081
Urban	0.3 %	1.4 %	97.2 %	0.2 %	0.8 %	196265
Water	0.0 %	0.3 %	0.1 %	99.3 %	0.3 %	272931
Forests	0.2 %	2.7 %	0.8 %	0.3 %	96.0 %	376029
Support	14190	361943	204301	274851	431789	1287074

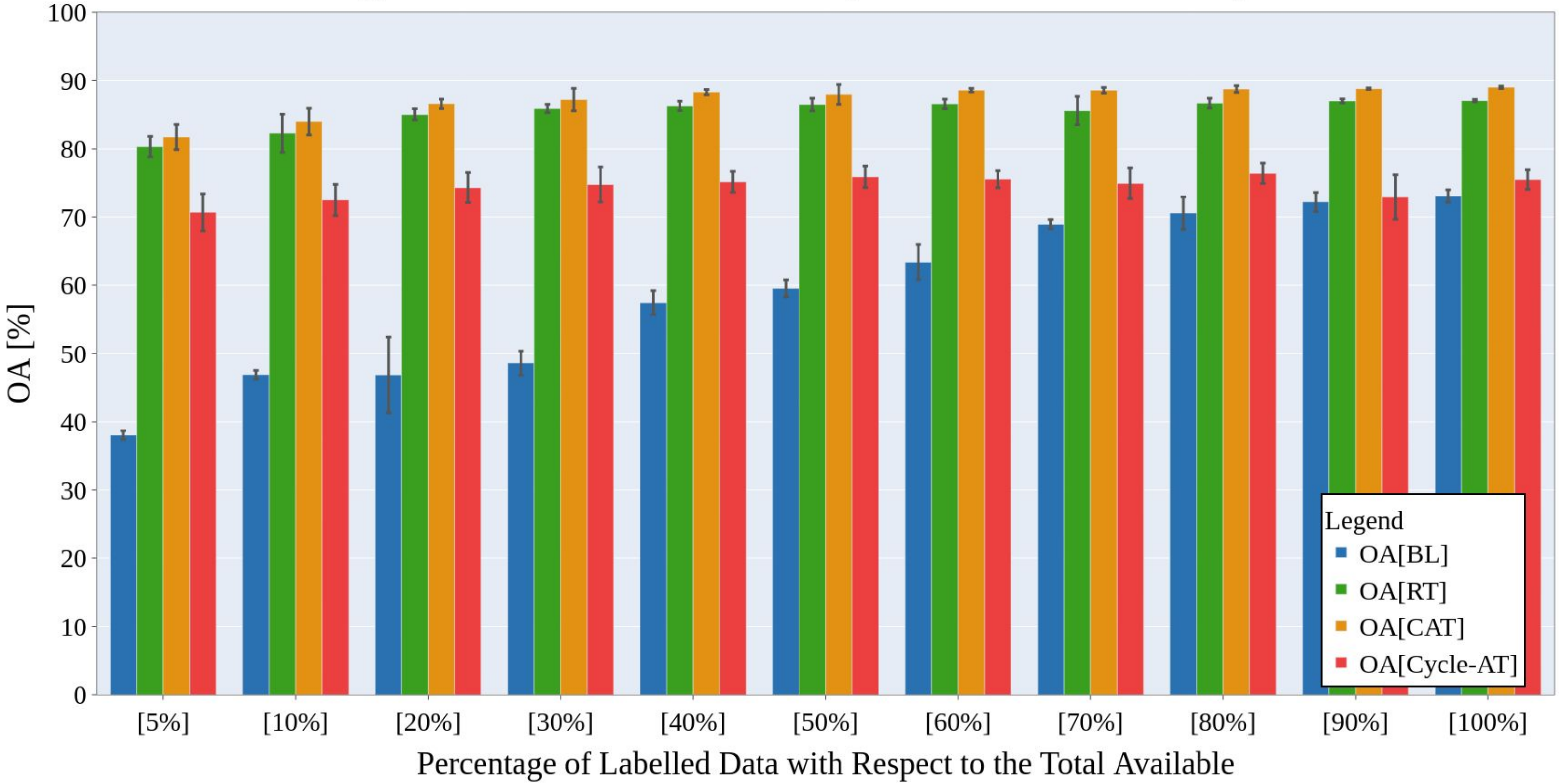
90 80 70 60 50 40 30 20 10 0

Cycle-AT - Confusion Matrix

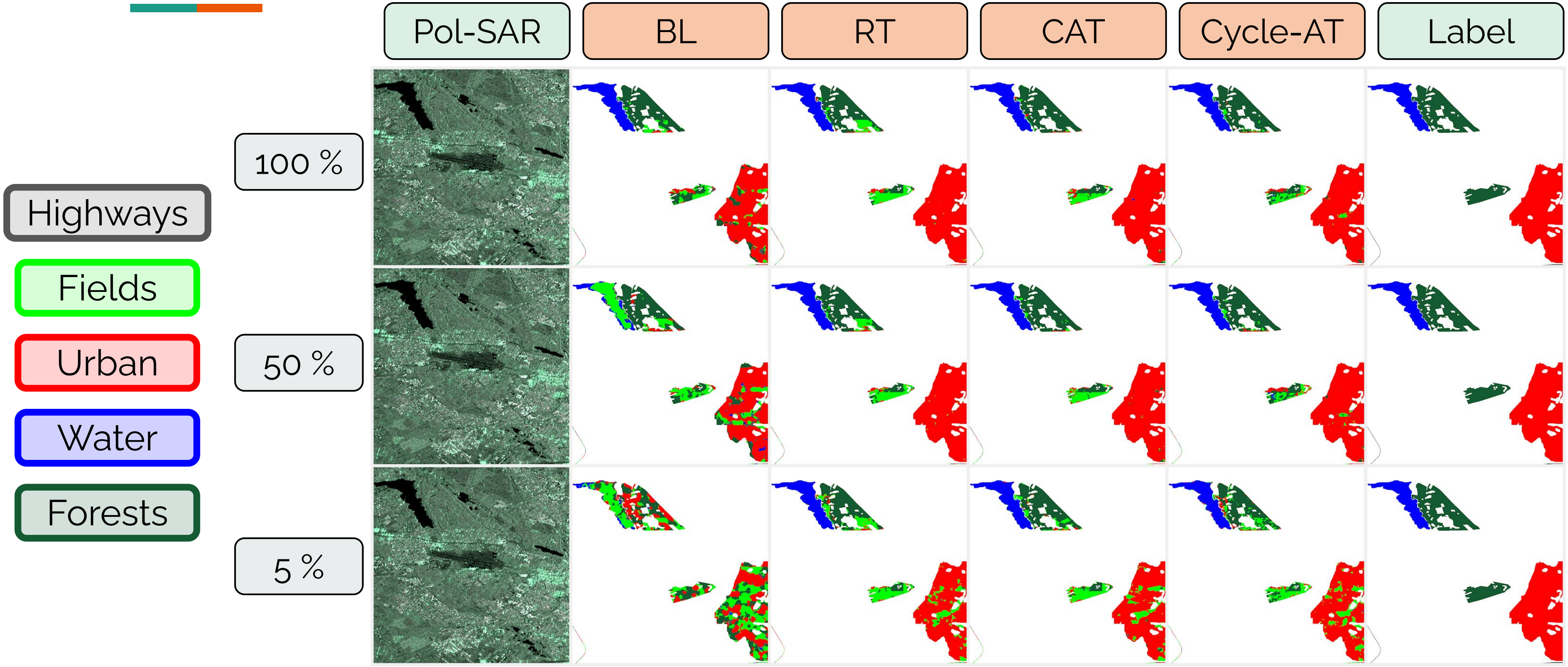
PREDICTIONS

LABELS	Highways	Fields	Urban	Water	Forests	Support
Highways	33.7 %	48.4 %	13.9 %	1.9 %	2.1 %	27768
Fields	1.3 %	73.9 %	0.6 %	0.7 %	23.6 %	414081
Urban	1.0 %	4.4 %	93.8 %	0.1 %	0.7 %	196265
Water	0.5 %	2.9 %	0.2 %	96.0 %	0.5 %	272931
Forests	0.6 %	23.3 %	0.6 %	0.5 %	74.9 %	376029
Support	20190	423631	193200	267342	382711	1287074

Labels Sensitivity Analysis



Classification Maps vs % of Training Data - Test Set Sample



Conclusions/Takeaways

- Pol-SAR to Optical **transcoding extracts effective features.**
- **Pre-trained** models **outperform** their **supervised** counterpart.
- **Pre-trained** models show **outstanding resilience** to data **scarcity.**
- **CAT** achieves **top** classification performances.
- **RT, the simplest,** can **compete** with the **CAT.**
- **Cycle-AT embeds the potential** capacity to **learn to transcode** employing **uncoupled data.**

Future Works

- Try a **dataset** using **balanced classes**.
- Try the network with more **complex target classes**.
- Verify the capacity of the network in **different climatic areas** and **seasons**.
- Train **Cycle-AT** using **unpaired** data.
- Train **Cycle-AT** using more data **leveraging** possibility to support **uncoupled resources**.

Greetings

THANK YOU FOR THE ATTENTION

Bibliography

- **[Ley et al. 2018]** A. Ley, O. Dhondt, S. Valade, R. Haensch and O. Hellwich. Exploiting GAN-Based SAR to Optical Image Transcoding for Improved Classification via Deep Learning. In EUSAR 2018; 12th European Conference on Synthetic Aperture Radar, pages 1–6, 2018. (Cited on pages 17, 21, 23, 27 and 45.)
- **[Reyes et al. 2019]** M. F. Reyes, S. Auer, N. Merkle, Corentin Henry and M. Schmitt. SAR-to-Optical Image Translation Based on Conditional Generative Adversarial Networks - Optimization, Opportunities and Limits. Remote. Sens., vol. 11, page 2067, 2019. (Cited on pages 22, 51 and 58.)
- **[Saha et al. 2020]** Sudipan Saha, F. Bovolo and L. Bruzzone. Building Change Detection in VHR SAR Images via Unsupervised Deep Transcoding. IEEE Transactions on Geoscience and Remote Sensing, pages 1–13, 2020. (Cited on pages 21 and 58.)

Image Cutting

Test and train Datasets have been cut into patches:

	Patch Size	Number of Patches for Transcoding	Number of Patches for Classification
Train	128	24702	2730
Test	128	25772	2340
Train	256	6341	1116
Test	256	6606	1012

Control the Adversarial MIN-MAX Game

CAT

Train the discriminator **every 100 generator optimisation steps**.

Cycle-AT

The discriminator is optimized using **samples randomly drawn** from a buffer which stores the **last 50 generated** synthetic samples. The drawn sample is with a probability of 50% the last generated or one of the last 50.

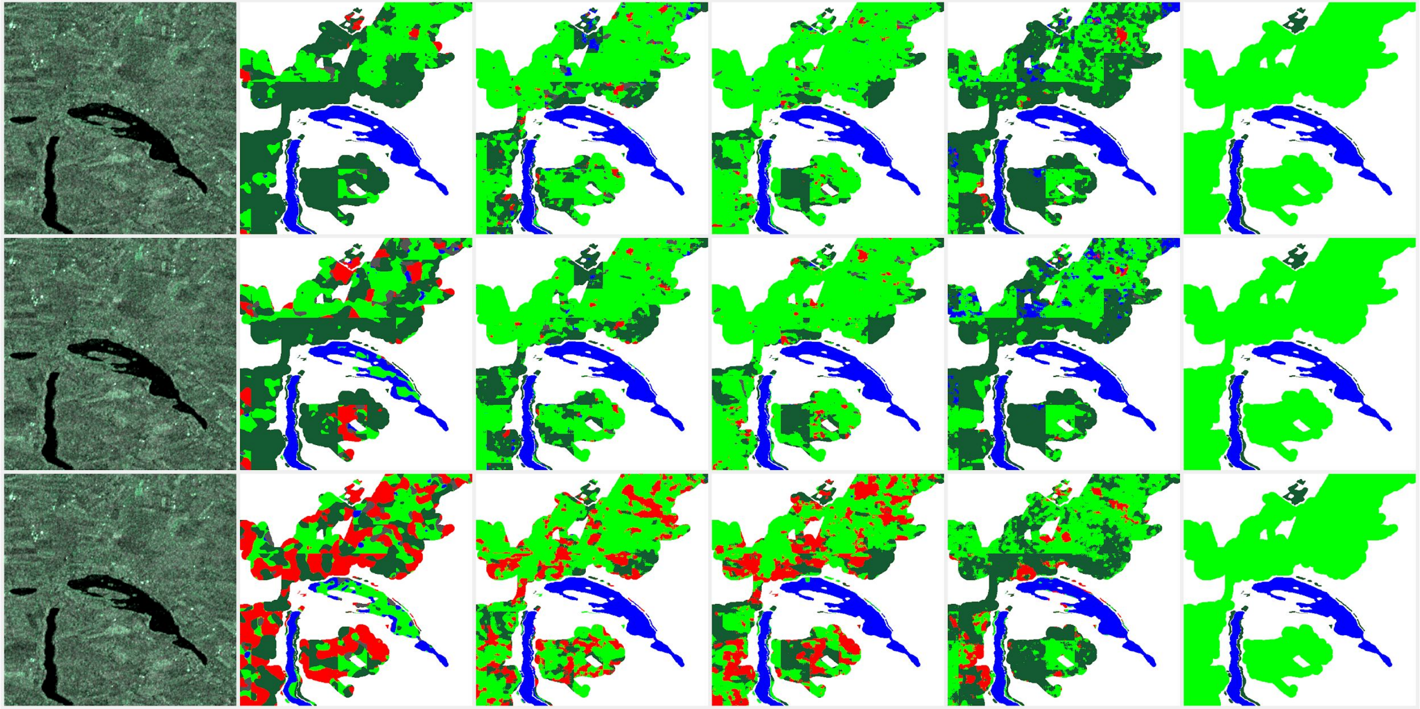
Classification Maps vs % of Training Data - Test Set Sample 2

Pol-SAR | BL | RT | CAT | Cycle-AT | Label

100 %

50 %

5 %



Highways

Fields

Urban

Water

Forests

Classification Maps vs % of Training Data - Test Set Sample 3

Highways

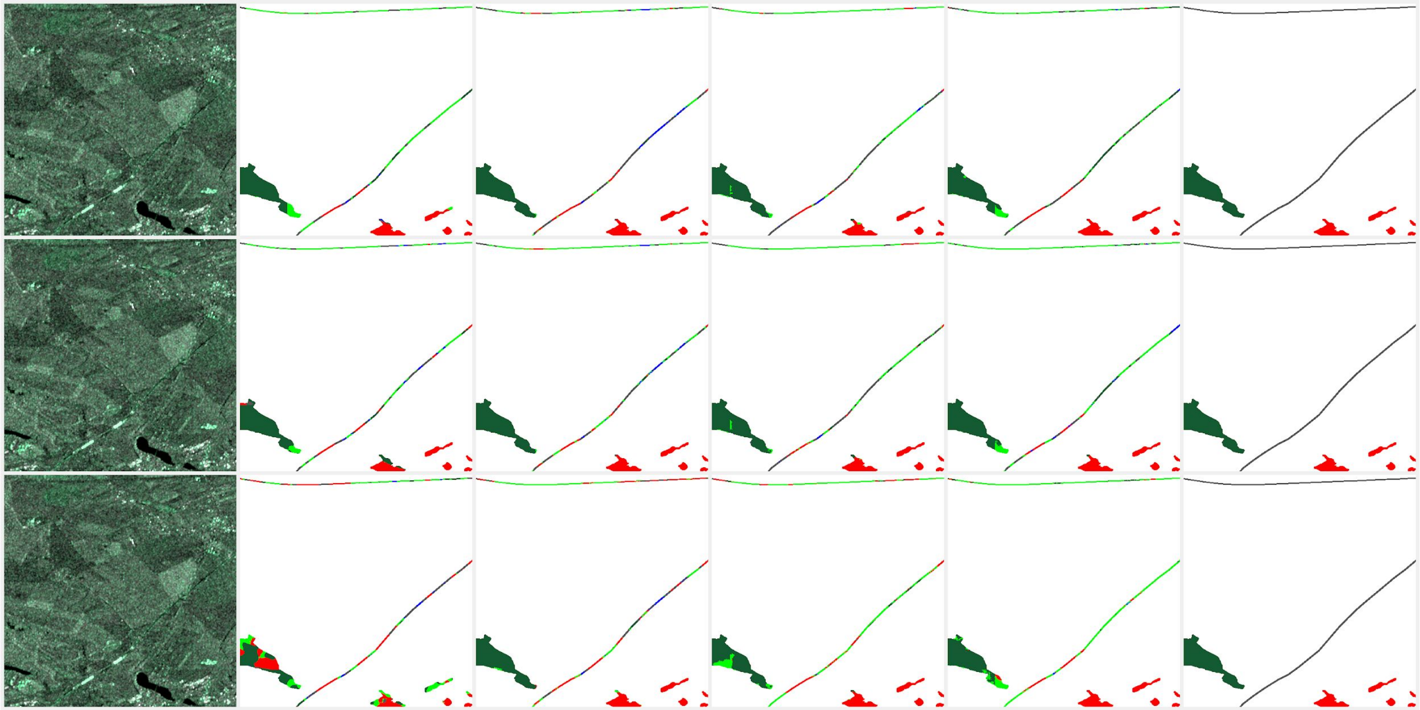
Fields

Urban

Water

Forests

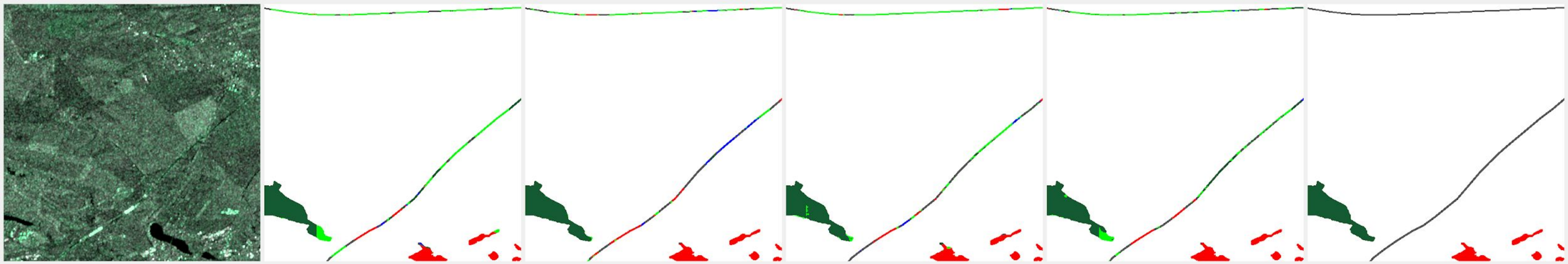
	Pol-SAR	BL	RT	CAT	Cycle-AT	Label
100 %						
50 %						
5 %						



Classification Maps - Test Set Sample 3

Highways | Fields | Urban | Water | Forests

Pol-SAR Input | BL - **73%** | RT - **87%** | CAT - **89%** | Cycle-AT - **76%** | Label



SEN-1 is a constellation of two polar-orbiting satellites, which performs C-band

SAR imaging and operates in dual polarization (HH+HV, VV+VH)