What Data are needed for Semantic Segmentation in Earth Observation?

Javiera Castillo-Navarro[1,2], Nicolas Audebert[1,2], Bertrand Le Saux[1], Alexandre Boulch[1], Sébastien Lefèvre[2]

[1]ONERA, Université Paris Saclay
[2]Université Bretagne-Sud/ IRISA

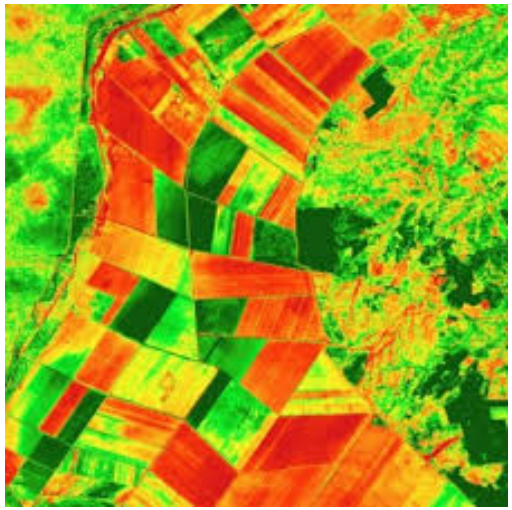
ONERA

UMR IRISA

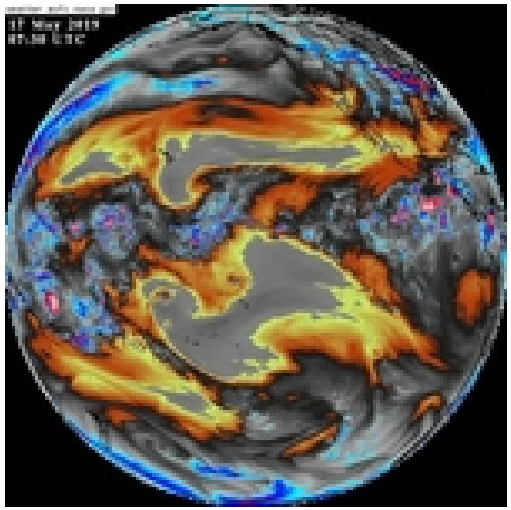
Motivation: Understanding EO data

- Multiple and useful applications for aerial or satellite images

Urban Planning



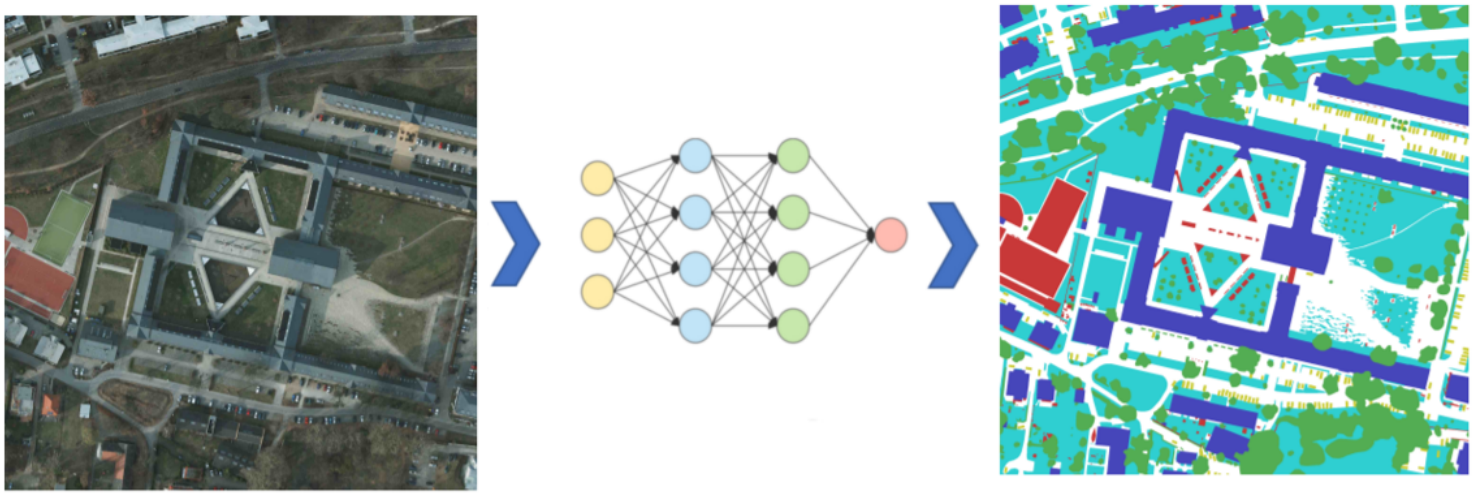
Agriculture



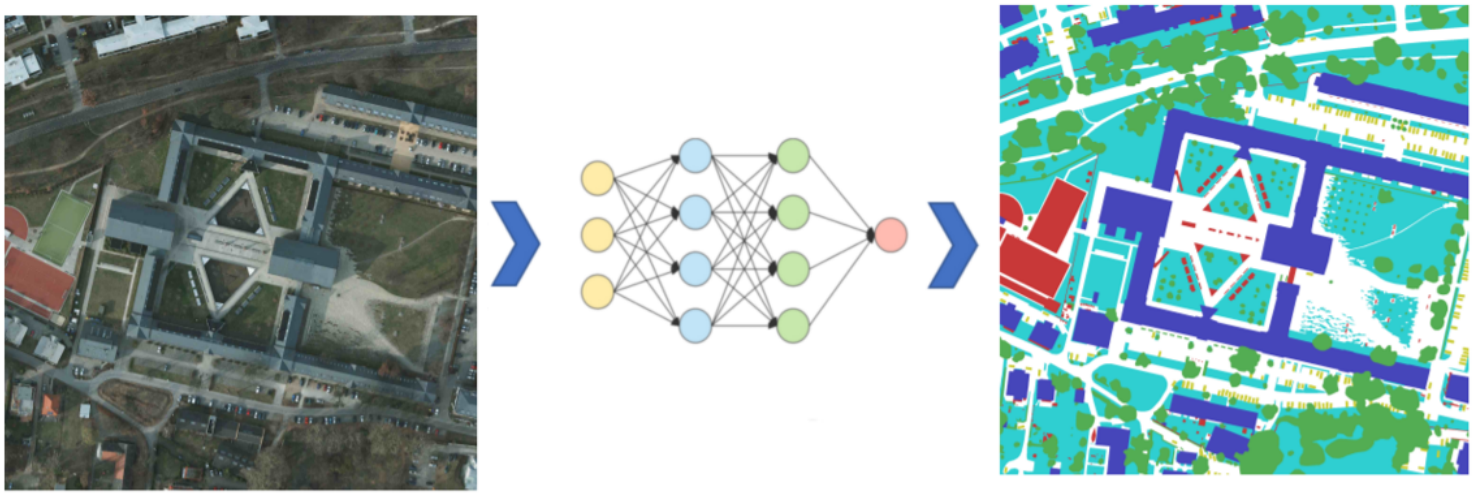
Meteorology

ONERA THE FRENCH AEROSPACE LAB

Automatizing the process of image mapping



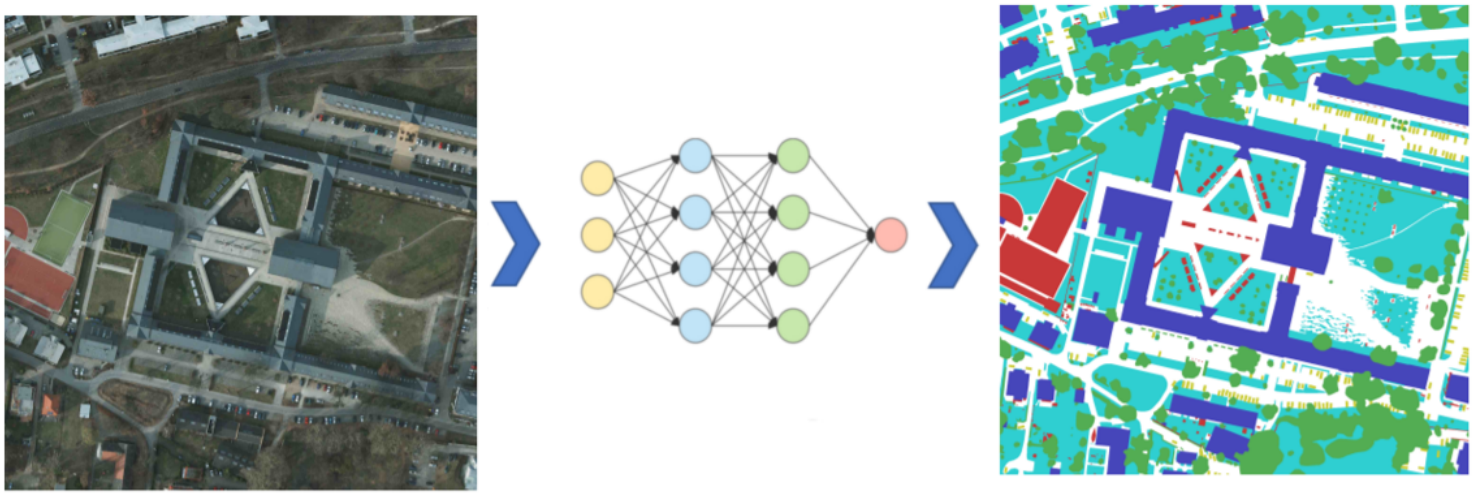
Automatizing the process of image mapping



- **Deep learning is the state of the art.**
- Our goal: semantic segmentation $\rightsquigarrow$ pixel-wise classification.
- **General knowledge:** We need a **big amount of annotated data!**

ONERA THE FRENCH AEROSPACE LAB

Automatizing the process of image mapping



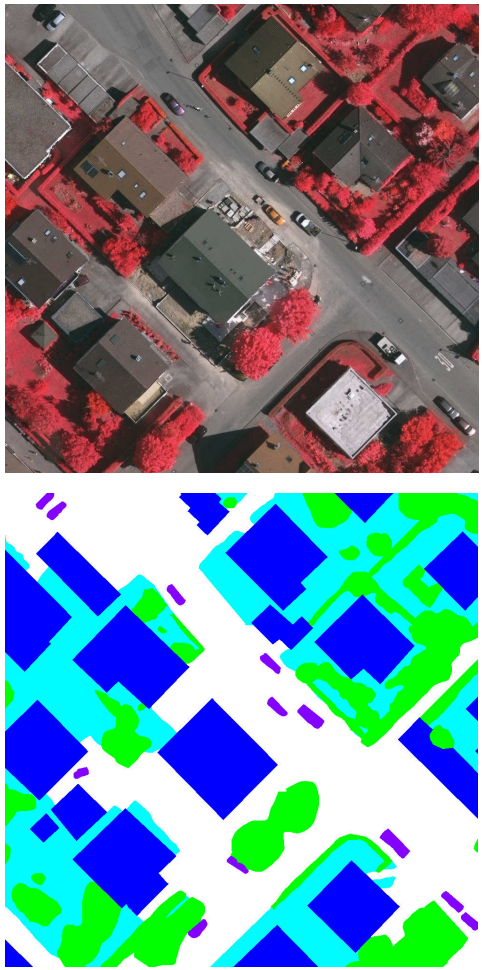
- **Deep learning is the state of the art.**
- Our goal: semantic segmentation $\rightsquigarrow$ pixel-wise classification.
- **General knowledge:** We need a **big amount of annotated data!**

... Do we?

ONERA THE FRENCH AEROSPACE LAB

How robust are Supervised Learning Networks for Semantic Segmentation with respect to Data Volume?

ISPRS Vaihingen Dataset



- Standard small-scale dataset for semantic segmentation in EO.
- Images over Vaihingen, Germany (single location, ~ 0.04 km^2 and ~ 0.8 GB).
- 9cm/pixel resolution. IRRG encoding.
- 6 visual categories.

Some experiments about robustness

Network architecture

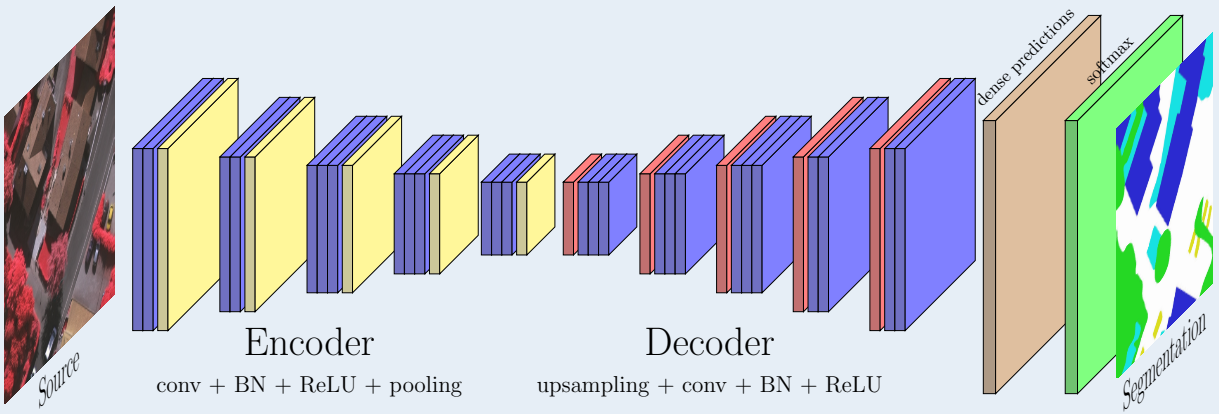


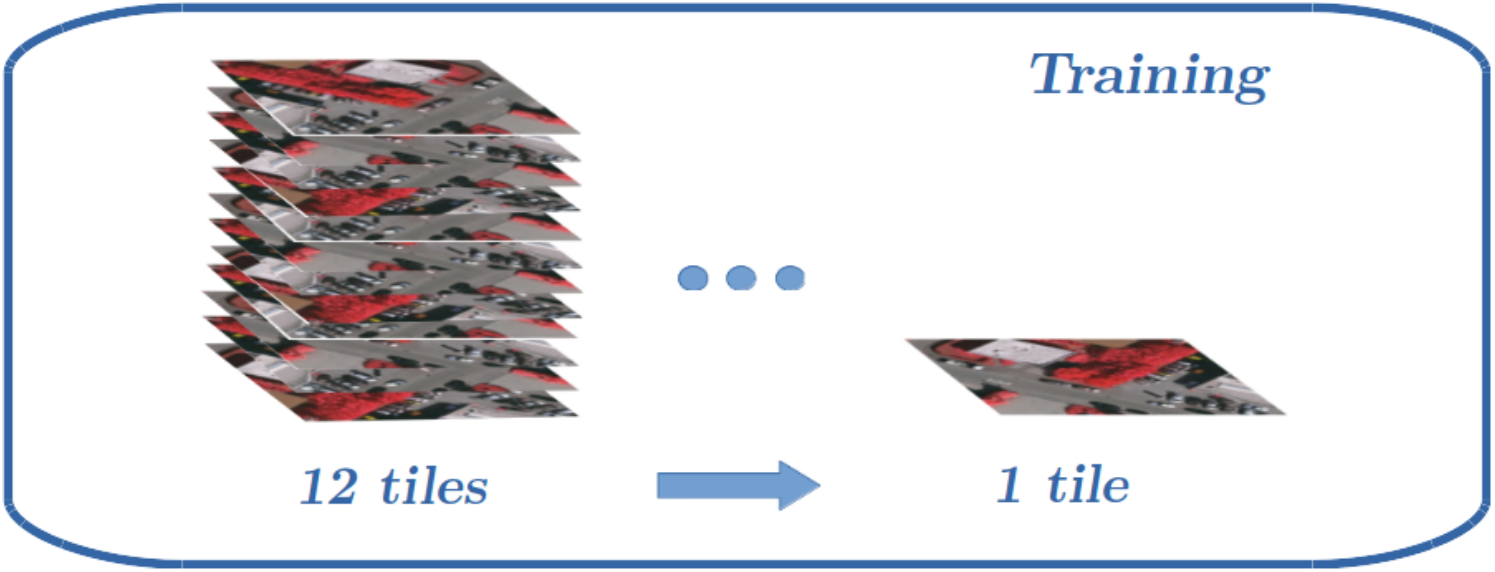
Figure: SegNet

- A well-known fully convolutional, encoder-decoder architecture for semantic segmentation.
- 89.8% Overall Accuracy over ISPRS Vaihingen[1].

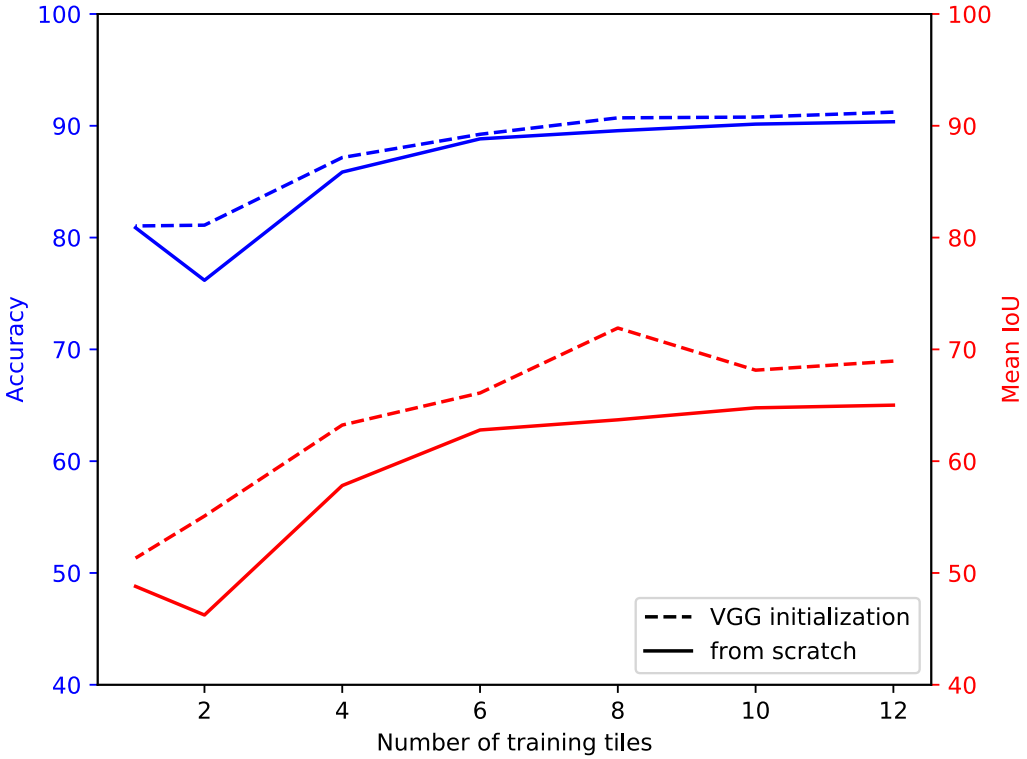
[1] Audebert, Le Saux and Lefèvre. "Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks", *ISPRS J. 140, 2018.*

Experimental settings

- Fixed test set (4 tiles).
- Decreasing number of tiles available for training.
- Good quality annotations.

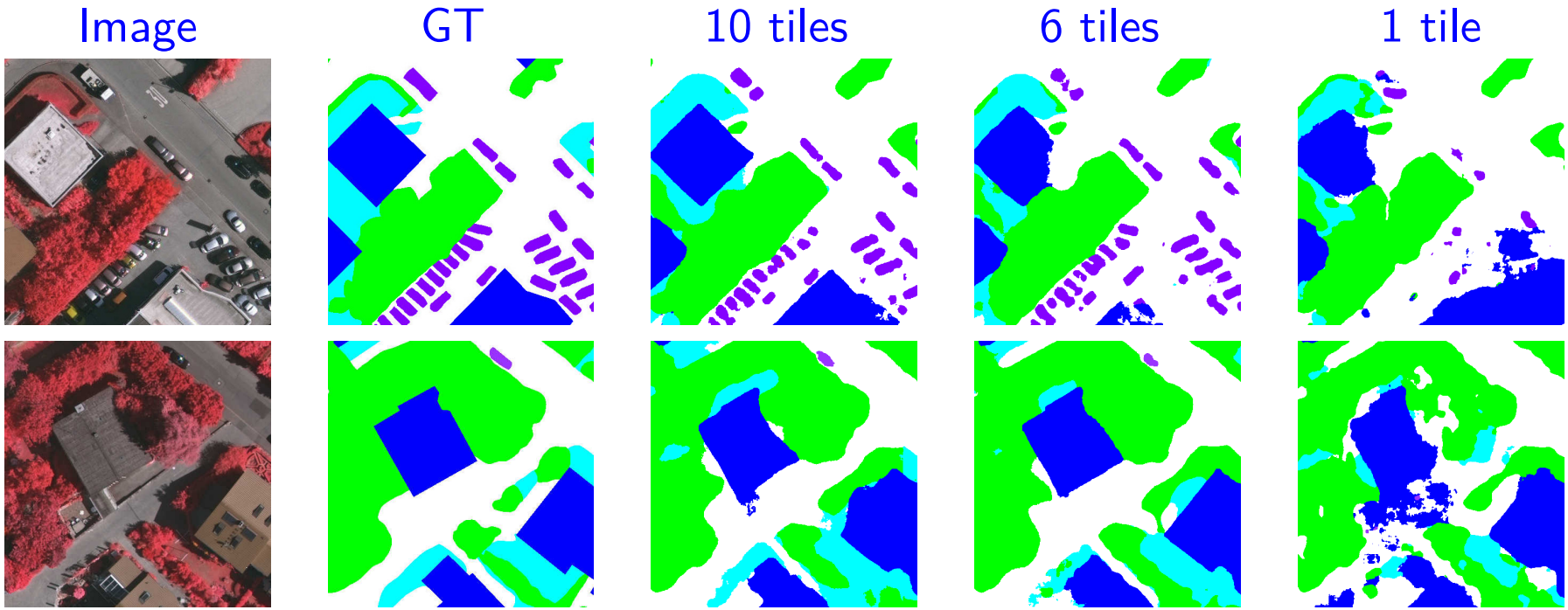


Results over ISPRS Vaihingen



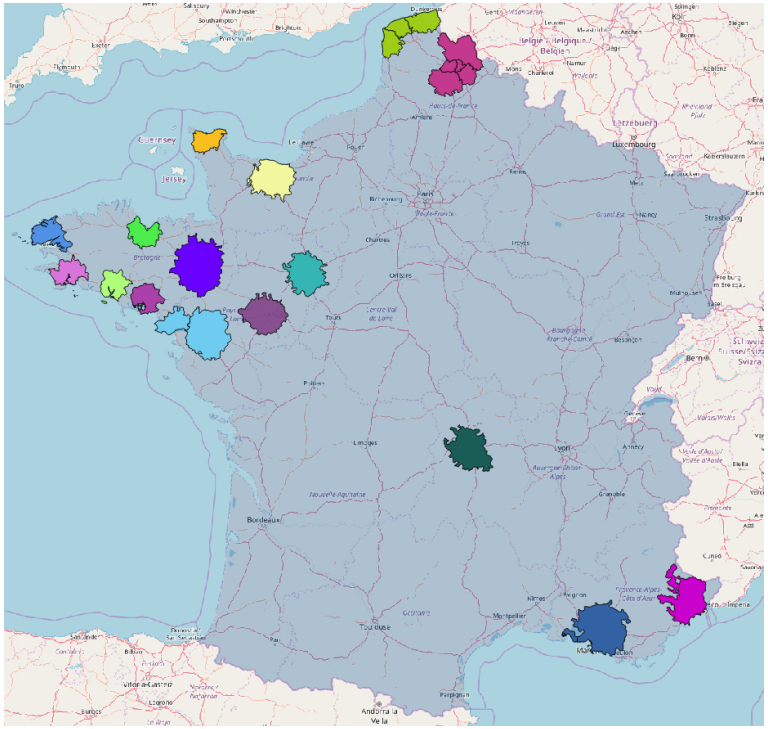
Nb. of tiles	VGG-16 init		Random init	
	OA	*mIoU*	*OA*	*mIoU*
12	91.2	69.0	90.4	65.0
10	90.8	68.1	90.2	64.8
8	90.7	71.9	89.6	63.7
6	89.2	66.1	88.8	62.8
4	87.2	63.2	85.9	57.8
2	81.1	55.1	76.2	46.2
1	81.0	51.3	80.9	48.8

Visually



How Can We Progress towards the ImageNet of Remote Sensing?

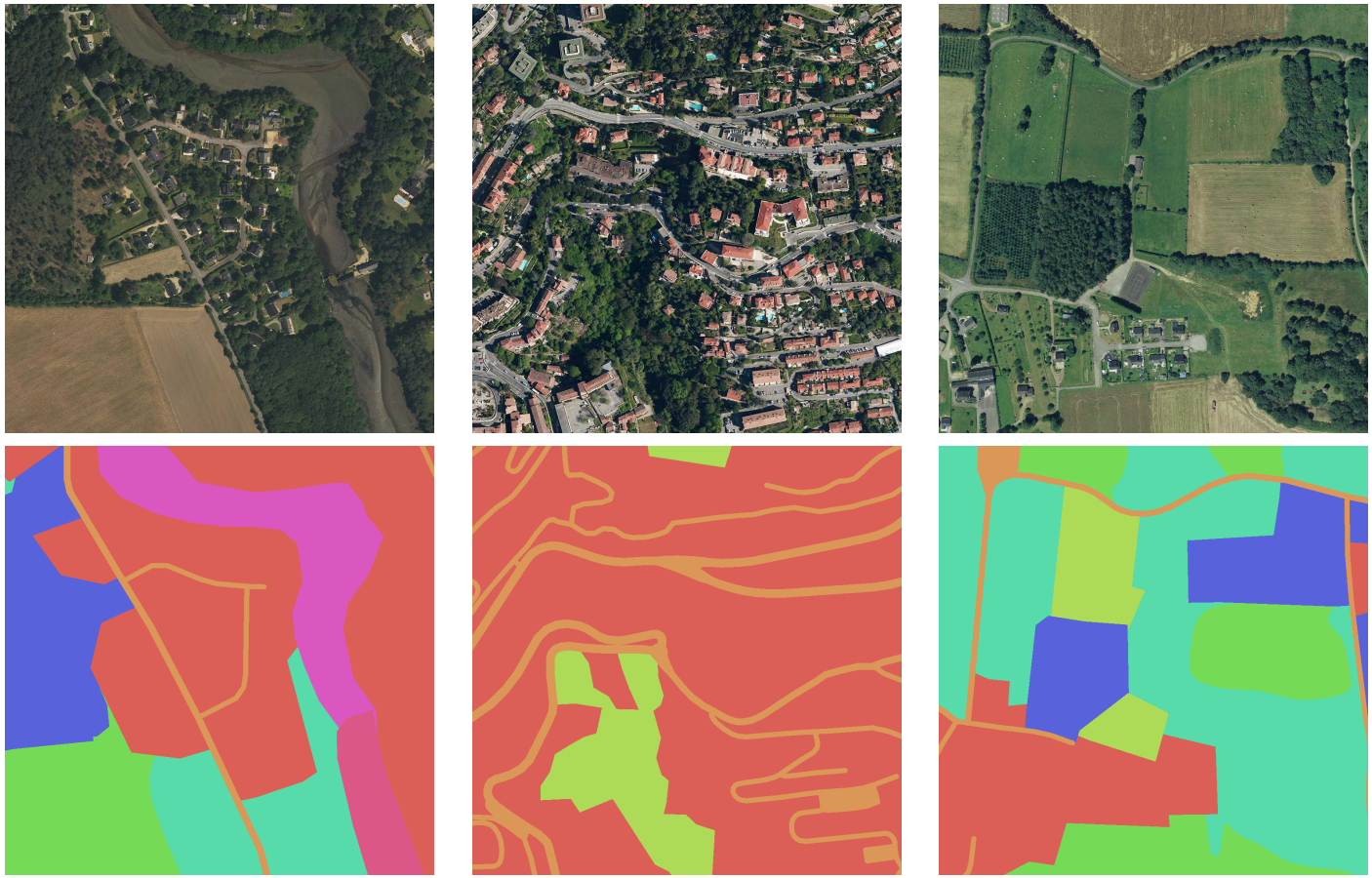
The MiniFrance Dataset



New dataset for large-scale cartography

- Very large dataset for semantic segmentation (> 53000 km^2 and ~ 150 GB.
- 16 conurbations all over France.
- Aerial images from BD ORTHO (IGN) at 50cm/pixel resolution and RGB encoding.
- 15 land-use clases from UrbanAtlas.

The Challenges of Large-Scale Data



- Quantity and variability of data.
- Higher semantics, not visual classes.

Results: the real life

Train set	*Overall accuracy*	*Mean IoU*

ONERA THE FRENCH AEROSPACE LAB

Results: the real life

Train set	*Overall accuracy*	*Mean IoU*
100 %	52.40	15.79

- SegNet accuracy decays considerably $\rightarrow$ **real life is hard**.

Results: the real life

Train set	*Overall accuracy*	*Mean IoU*
100 %	52.40	15.79
10 %	50.14	15.25

- SegNet accuracy decays considerably $\rightarrow$ **real life is hard**.
- The dataset is robust to uniform sampling.

Results: the real life

Train set	*Overall accuracy*	*Mean IoU*
100 %	52.40	15.79
10 %	50.14	15.25
Caen ($\sim$ 12.5%)	42.09	10.05

- SegNet accuracy decays considerably $\rightarrow$ **real life is hard**.
- The dataset is robust to uniform sampling.
- But **variability matters** $\rightarrow$ representativity of data $>$ quantity.

Results: Can the network generalize to every city?

Table: Results by conurbation, trained over the entire train set.

Score	Marseille	Rennes	Angers	Quimper	Vannes	Clermont	Lille	Cherbourg
OA	46.13	51.56	44.85	50.82	49.51	46.51	61.35	67.54
mIoU	12.77	15.05	13.15	13.93	12.66	11.40	16.93	15.82

Results: Can the network generalize to every city?

Table: Results by conurbation, trained over the entire train set.

Score	Marseille	Rennes	Angers	Quimper	Vannes	Clermont	Lille	Cherbourg
OA	46.13	51.56	44.85	50.82	49.51	46.51	61.35	67.54
mIoU	12.77	15.05	13.15	13.93	12.66	11.40	16.93	15.82

- How can we generalize equally to all locations?

Results: Can the network generalize to every city?

Table: Results by conurbation, trained over the entire train set.

Score	Marseille	Rennes	Angers	Quimper	Vannes	Clermont	Lille	Cherbourg
OA	46.13	51.56	44.85	50.82	49.51	46.51	61.35	67.54
mIoU	12.77	15.05	13.15	13.93	12.66	11.40	16.93	15.82

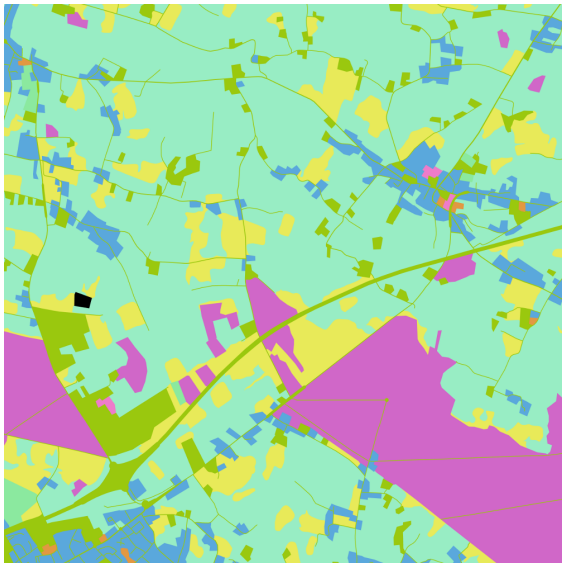
- How can we generalize equally to all locations?

  → Fine-tuning on cities of interest using interactive learning approaches?

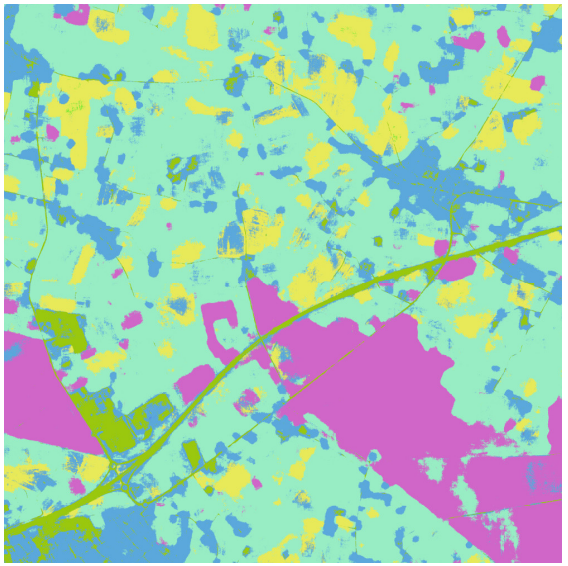
  → Semi-supervised strategies to leverage unlabeled data?

Visual results

Image



Ground-Truth



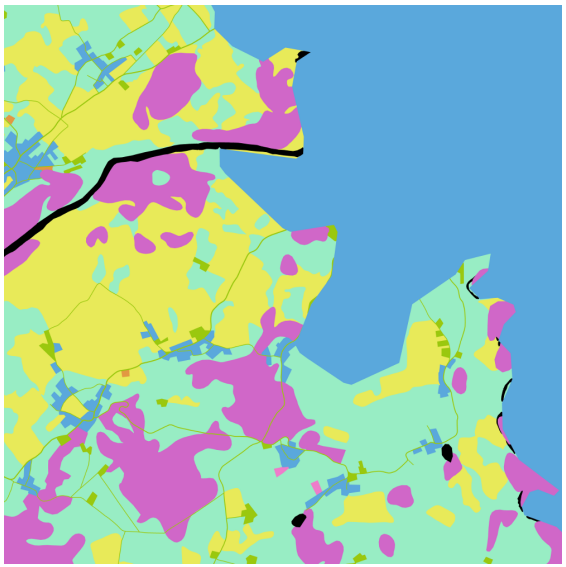
Prediction

- High-level semantics can be learned!

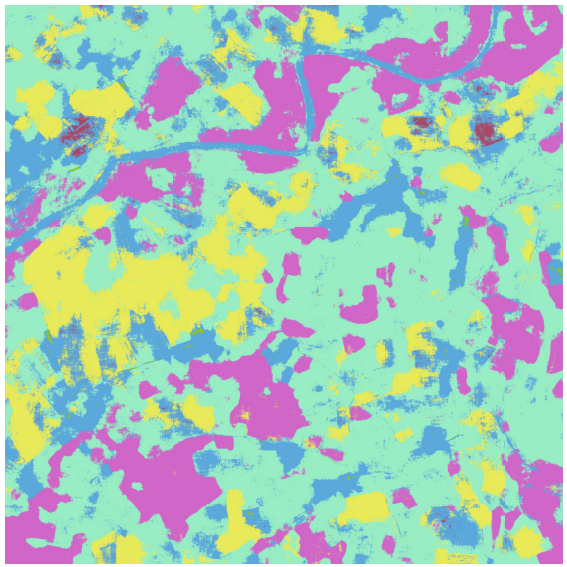
ONERA THE FRENCH AEROSPACE LAB

Visual results

Image



Ground-Truth



Prediction

- How can we learn from imperfect data?
  - → Weakly supervised learning approaches?

ONERA THE FRENCH AEROSPACE LAB

Conclusions

- On single-location datasets, current networks are robust to a drastic decrease of training data.

Conclusions

- On single-location datasets, current networks are robust to a drastic decrease of training data.
- Large-scale, real-life datasets raise new issues:

Conclusions

- On single-location datasets, current networks are robust to a drastic decrease of training data.
- Large-scale, real-life datasets raise new issues:
  $\rightarrow$ Images are more diverse.

Conclusions

- On single-location datasets, current networks are robust to a drastic decrease of training data.
- Large-scale, real-life datasets raise new issues:
  → Images are more diverse.
  → Available annotations are often approximate.

Conclusions

- On single-location datasets, current networks are robust to a drastic decrease of training data.
- Large-scale, real-life datasets raise new issues:
  → Images are more diverse.
  → Available annotations are often approximate.
  → Classes have a higher semantic level.

Conclusions

- On single-location datasets, current networks are robust to a drastic decrease of training data.
- Large-scale, real-life datasets raise new issues:
  → Images are more diverse.
  → Available annotations are often approximate.
  → Classes have a higher semantic level.
- MiniFrance challenges the potential of deep networks and provides lifelike use-cases.

Perspectives

A focus on generalization:

- Benefit from unlabeled data, through **semi-supervised** strategies.
- Consider weakly supervised methods to learn from imperfect annotations.

Coming soon...

The **MiniFrance dataset** for semantic segmentation (and other tasks) release...

Questions?

Thank you!

Acknowledgements

- Javiera Castillo-Navarro's work is partially funded by a cnes research grant.
- This work is part of MEDUSA and DELTA research projects at ONERA .