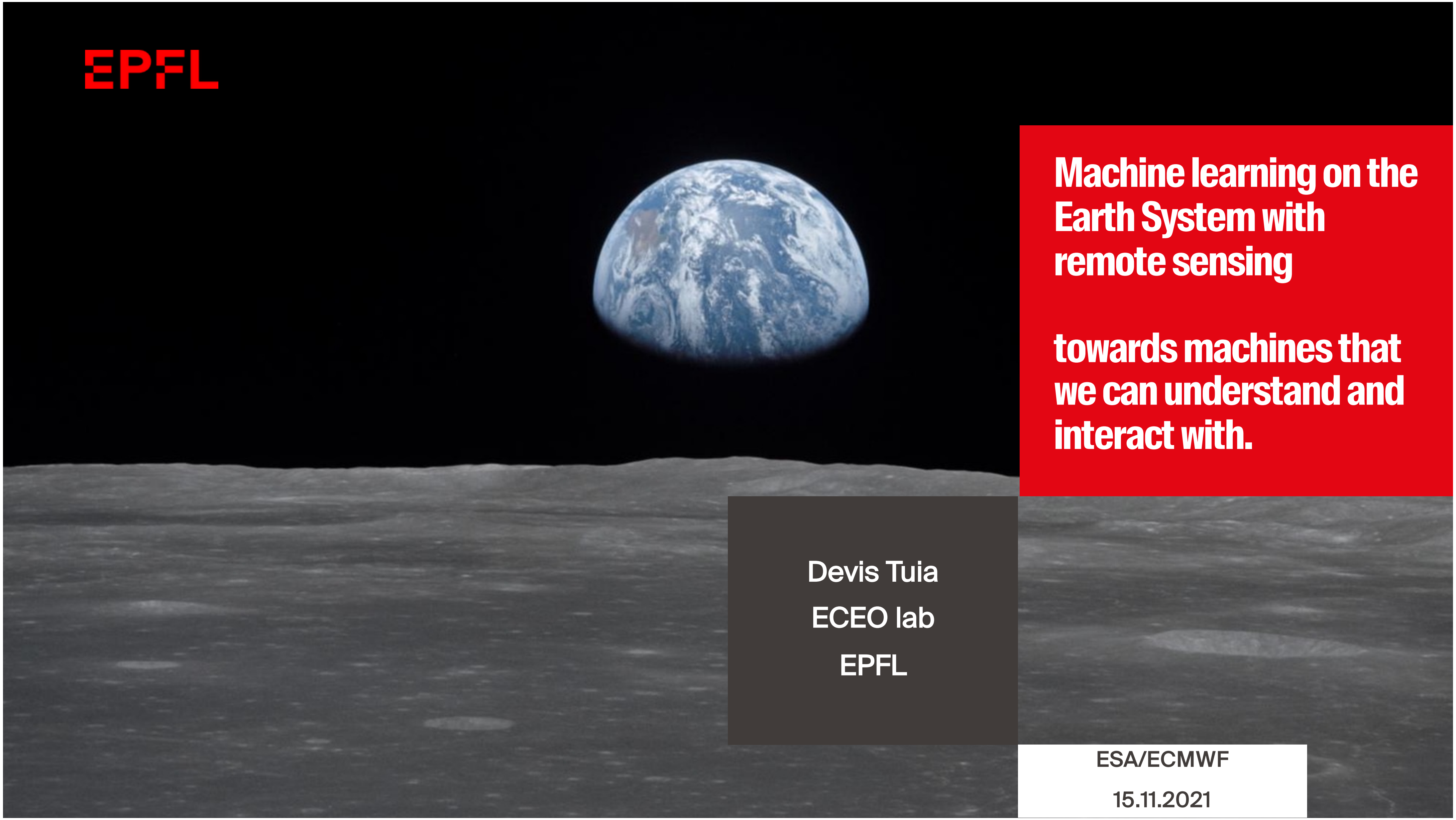


EPFL

Machine learning on the Earth System with remote sensing

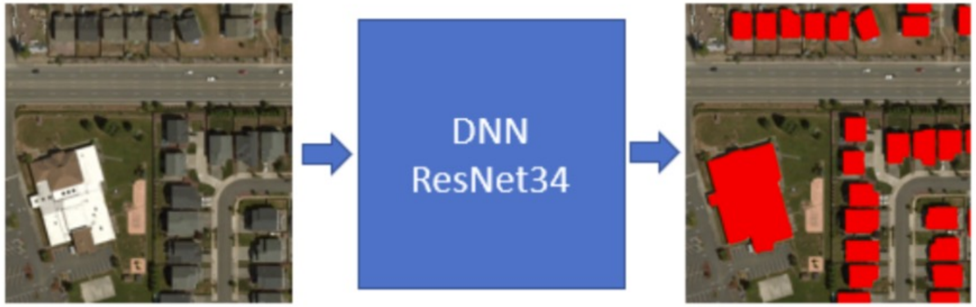
towards machines that we can understand and interact with.

Devis Tuia
ECEO lab
EPFL

ESA/ECMWF
15.11.2021

Applying deep learning with optical remote sensing data seems very easy

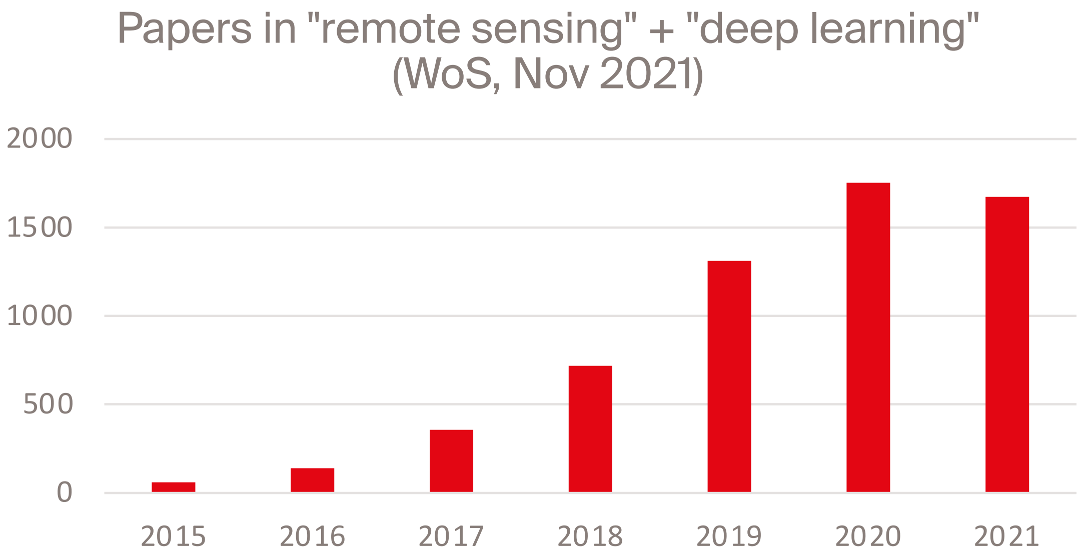
June 28th 2018: *Bing releases 125 million Building Footprints in the US as Open Data*
How?



Apply ResNet [He et al., 2015] + smart postprocessing

The low hanging fruit is a blessing...

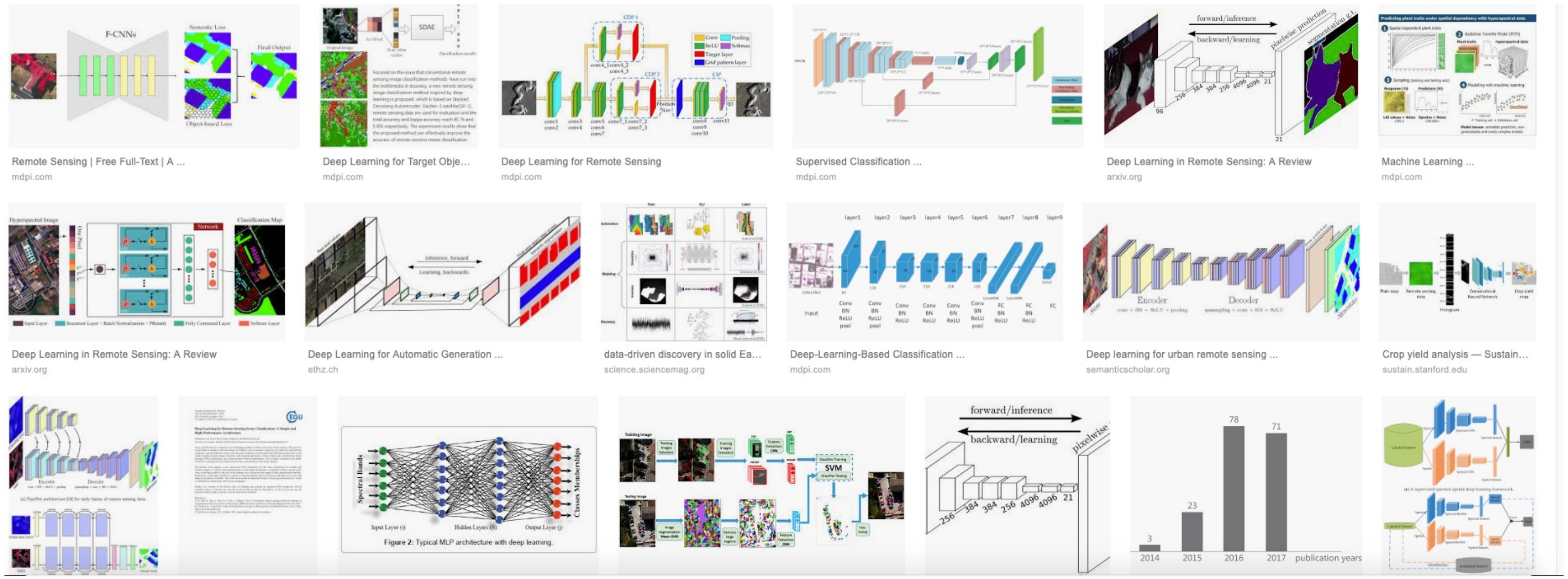
- We can advance several applications with this technology from CS
- Massive increase of "DL-in-RS" papers



Web of science, 4.11.2021

The low hanging fruit is a blessing... in disguise

- We can advance several applications with this technology from CS
- Massive increase of “DL-in-RS” papers
- Kind of DL-winter already.



EPFL

So we started thinking on what is exciting in AI4EO…

arXiv.org > cs > arXiv:2104.05107

Search...

Help | Advance

Computer Science > Computer Vision and Pattern Recognition

[Submitted on 11 Apr 2021]

Towards a Collective Agenda on AI for Earth Science Data Analysis

Devis Tuia, Ribana Roscher, Jan Dirk Wegner, Nathan Jacobs, Xiao Xiang Zhu, Gustau Camps-Valls

In the last years we have witnessed the fields of geosciences and remote sensing and artificial intelligence to become closer. Thanks to both the massive availability of observational data, improved simulations, and algorithmic advances, these disciplines have found common objectives and challenges to advance the modeling and understanding of the Earth system. Despite such great opportunities, we also observed a worrying tendency to remain in disciplinary comfort zones applying recent advances from artificial intelligence on well resolved remote sensing problems. Here we take a position on research directions where we think the interface between these fields will have the most impact and become potential game changers. In our declared agenda for AI on Earth sciences, we aim to inspire researchers, especially the younger generations, to tackle these challenges for a real advance of remote sensing and the geosciences.

Comments:	In press at IEEE Geoscience and Remote Sensing Magazine
Subjects:	**Computer Vision and Pattern Recognition (cs.CV)**; Signal Processing (eess.SP)
DOI:	10.1109/MGRS.2020.3043504
Cite as:	arXiv:2104.05107 **[cs.CV]**
	(or arXiv:2104.05107v1 **[cs.CV]** for this version)

Six exciting research directions

1. Reasoning, encoding spatial thinking
2. Massive multi-source (audio / video / social media)
3. Human machine interaction In natural language
4. Physics-based ML
5. Interpretable models
6. Discovering causal relations

EPFL

Six exciting research directions

Reasoning, encoding
spatial thinking

Physics-based ML

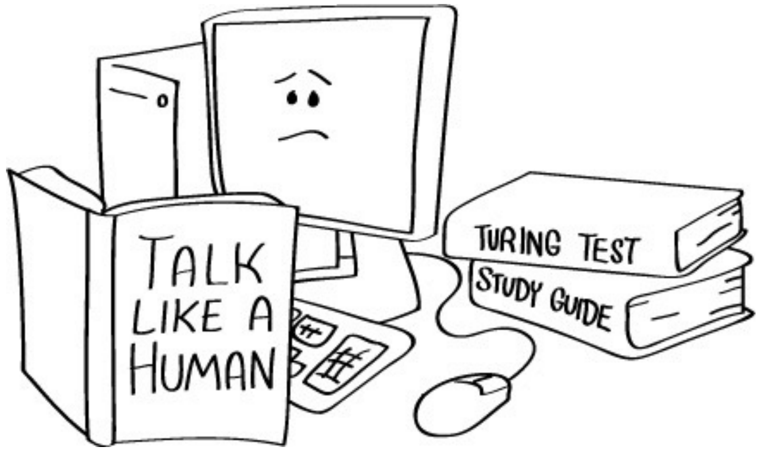
Massive multi-source
(audio / video / social media)

Interpretable models

Human machine interaction
In natural language

Discovering causal relations

Two (out of six) questions to move forward



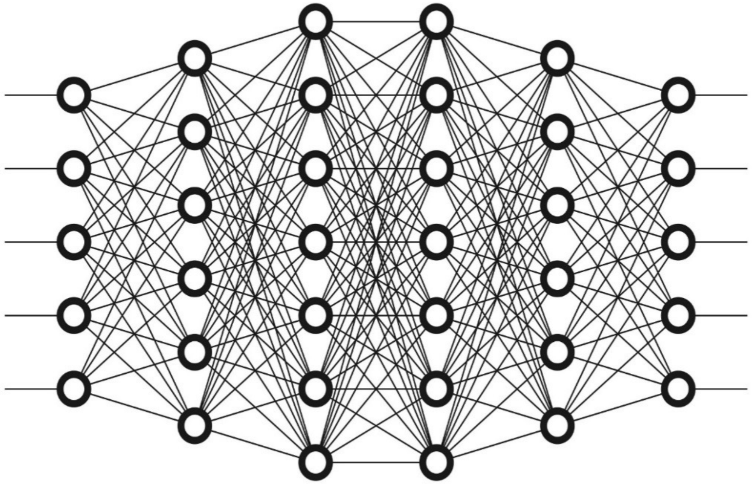
Understand what is going on within CNNs

XAI

Make them accessible to anyone

Interaction, NLP

What happens inside a CNN is a bit of a mystery...



0.74 Canal

0.23 River

0.02 Park

0.01 Station

0.01 Road

D. Tuia

What happens inside a CNN is a bit of a mystery...

0.74 Canal

0.23 River

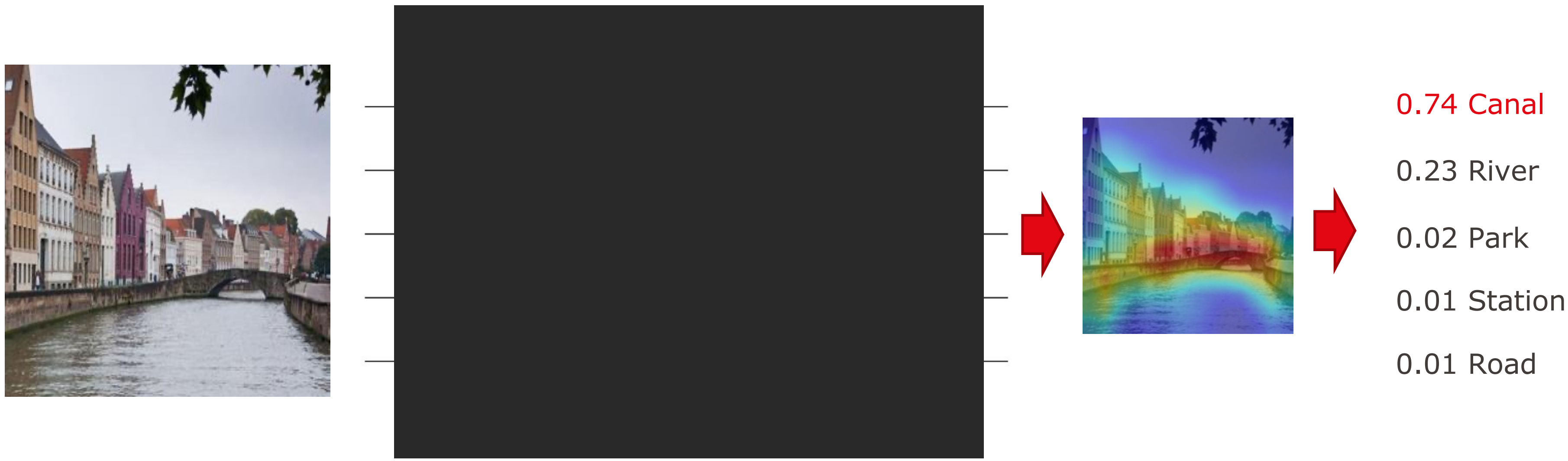
0.02 Park

0.01 Station

0.01 Road

D. Tuia

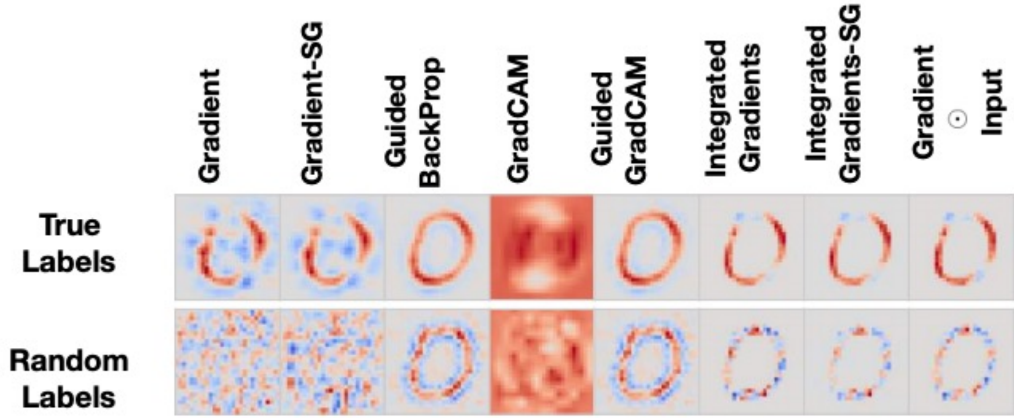
What happens inside a CNN is a bit of a mystery...



Zhou, Bolei, et al. "Learning deep features for discriminative localization.", *CVPR*. 2016.

Explaining what a neural network is doing is DIFFICULT

- Models can trick you by their capacity



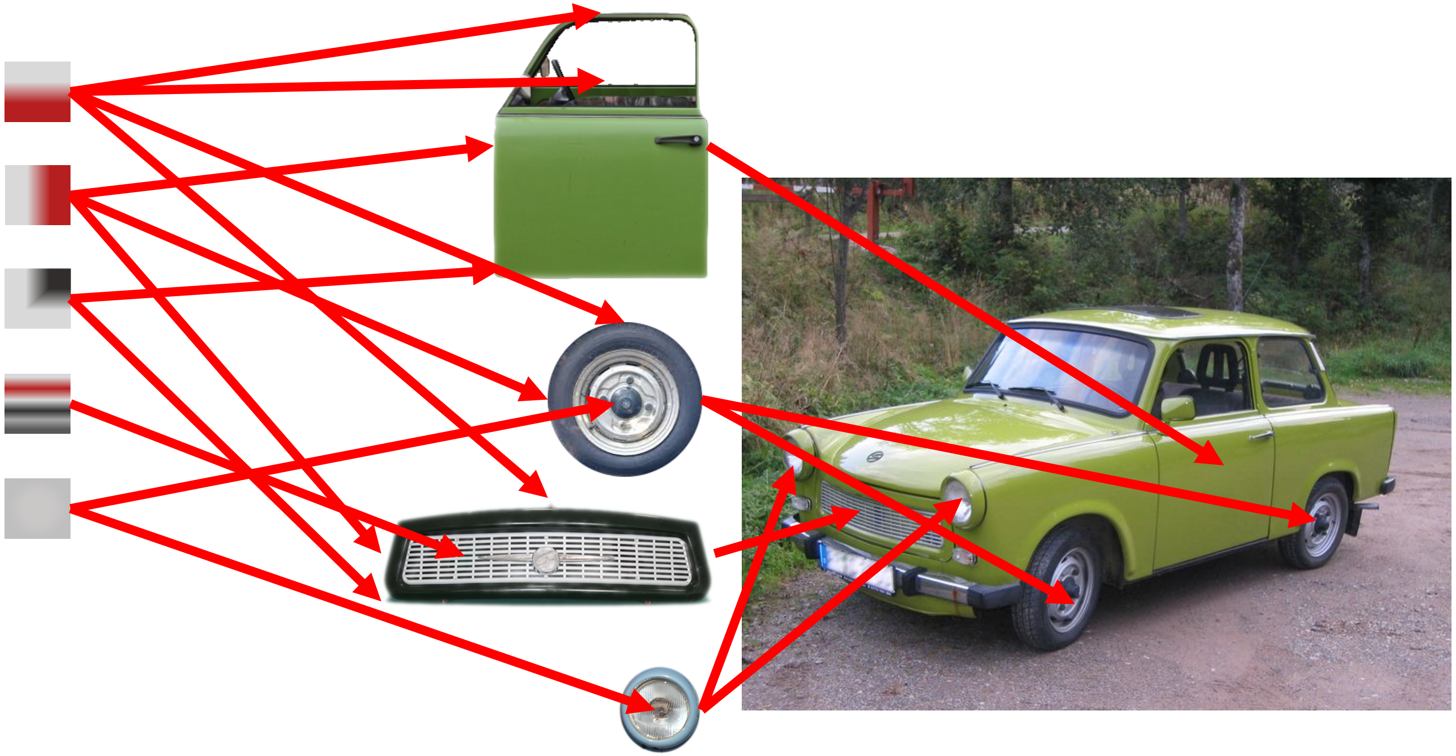
[Adebayo et al., *Sanity checks for saliency maps*, NeurIPS 2017]

- There are many ways to get to the same result
- Don't forget, we optimize for accuracy, not interpretability!

- Can we use simpler primitives to ensure explainability?

EPFL

We would like something like



Understanding landscape beauty (a quite subjective matter)

- Not all places are equally attractive
- Beauty is really subjective!

- However, there must be some kind of consensus

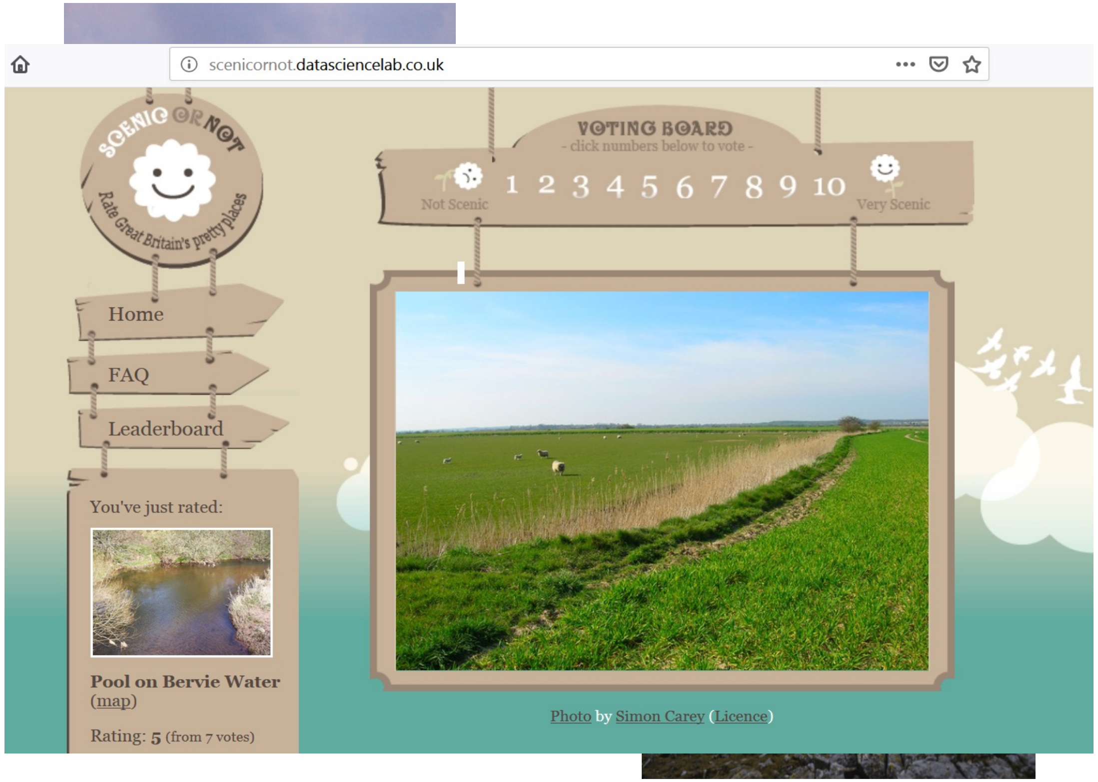
VS

Understanding landscape beauty (a quite subjective matter)

- Not all places are equally attractive
- Beauty is really subjective!

- However, there must be some kind of consensus

- Let's crowdsource that and understand factors related to beauty



EPFL

Understanding landscape beauty (a quite subjective matter)

- Not all places are equally attractive
- Beauty is really subjective!

- However, there must be some kind of consensus

- Let's crowdsource that and understand factors related to beauty

Photos © Sonse, Alex Healing, Tim Green, Jeremy Segrott (cc-by/2.0).

What are we aiming at

We want to go beyond saliency. We want to predict what makes a landscape scenic.

Scenic because of open fields

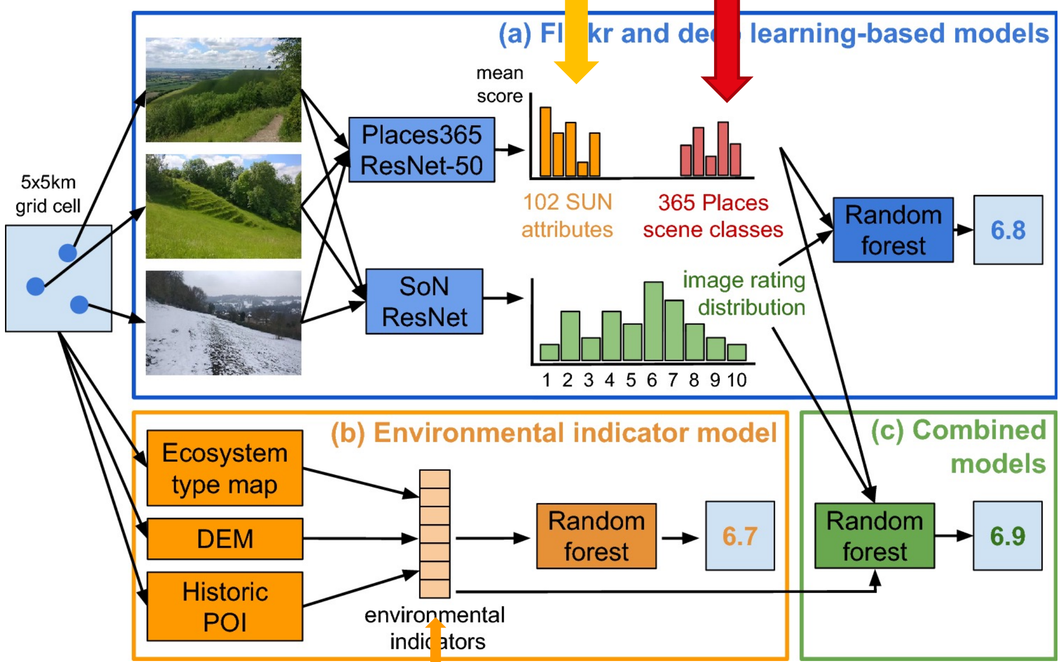
Not scenic because there are road signs

EPFL

Extracting interesting compositional factors

- Exercise
- Snow
- Swimming
- Rusty
- …

- Campsite
- Field
- Lagoon
- Mountain path
- …



- Naturalness
- Visual scale
- Complexity
- Uniqueness

- Object and scene characteristics from social media
- Indicators from remote sensing
- Fuse with random forest

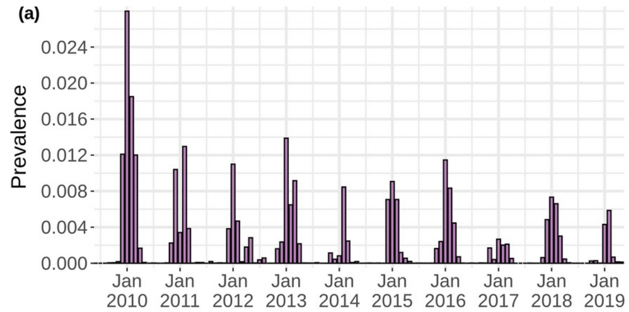
I. Havinga, D. Marcos, P. Bogaart, L. Hein, and D. Tuia. Computer vision a… ture the aesthetic quality of the landscape for ecosystem service assessments. *Scientific Reports*, 2021. h… articles/s41598-021-99282-0

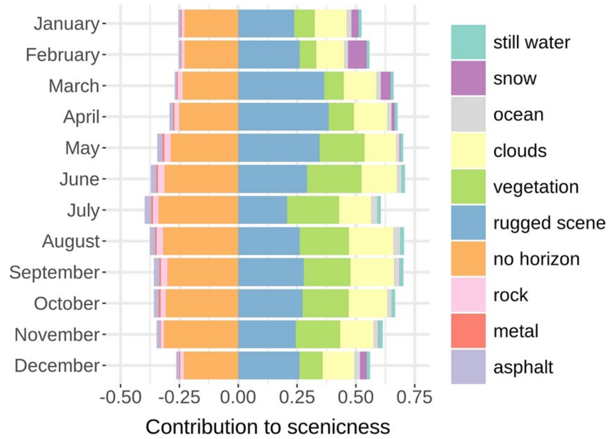
Extracting interesting compositional factors

- Studying evolution in time, thanks to Flickr data

flickr

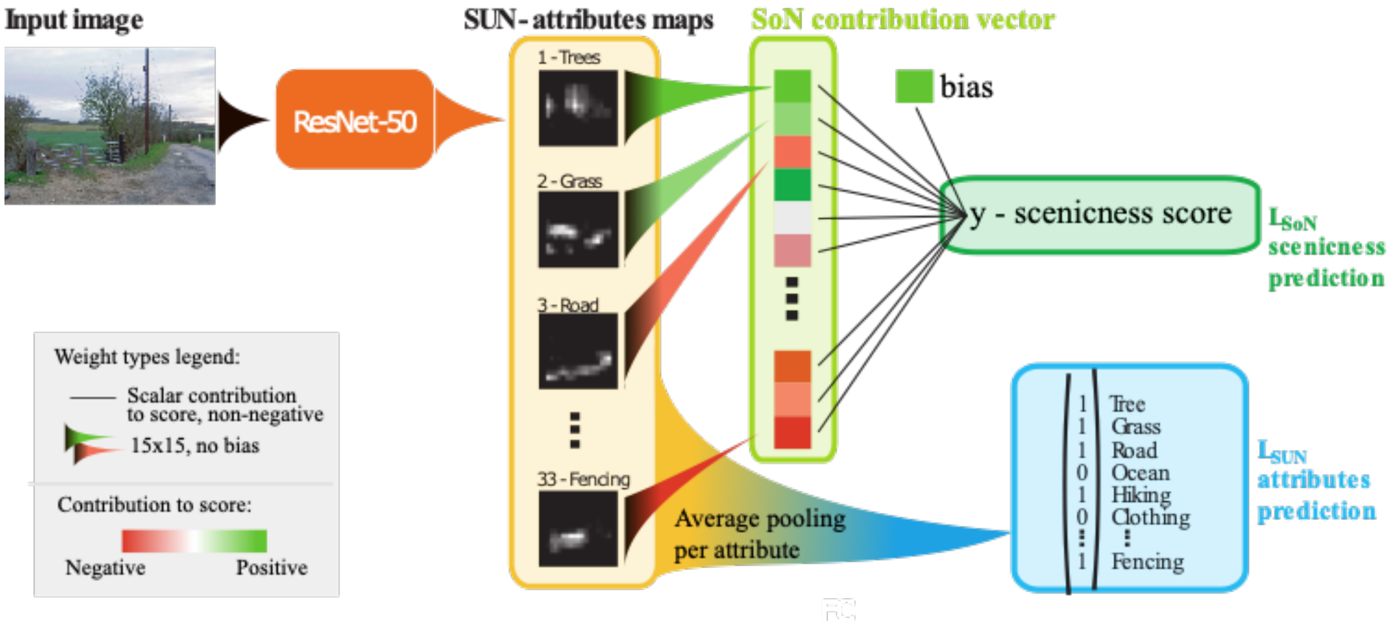
- Severeness of winters
- Concepts appearing at different times
- Different ecosystem services





I. Havinga, D. Marcos, P. Bogaart, L. Hein, and D. Tuia. Computer vision and social media data capture the aesthetic quality of the landscape for ecosystem service assessments. *Scientific Reports*, 2021. https://www.nature.com/articles/s41598-021-99282-0

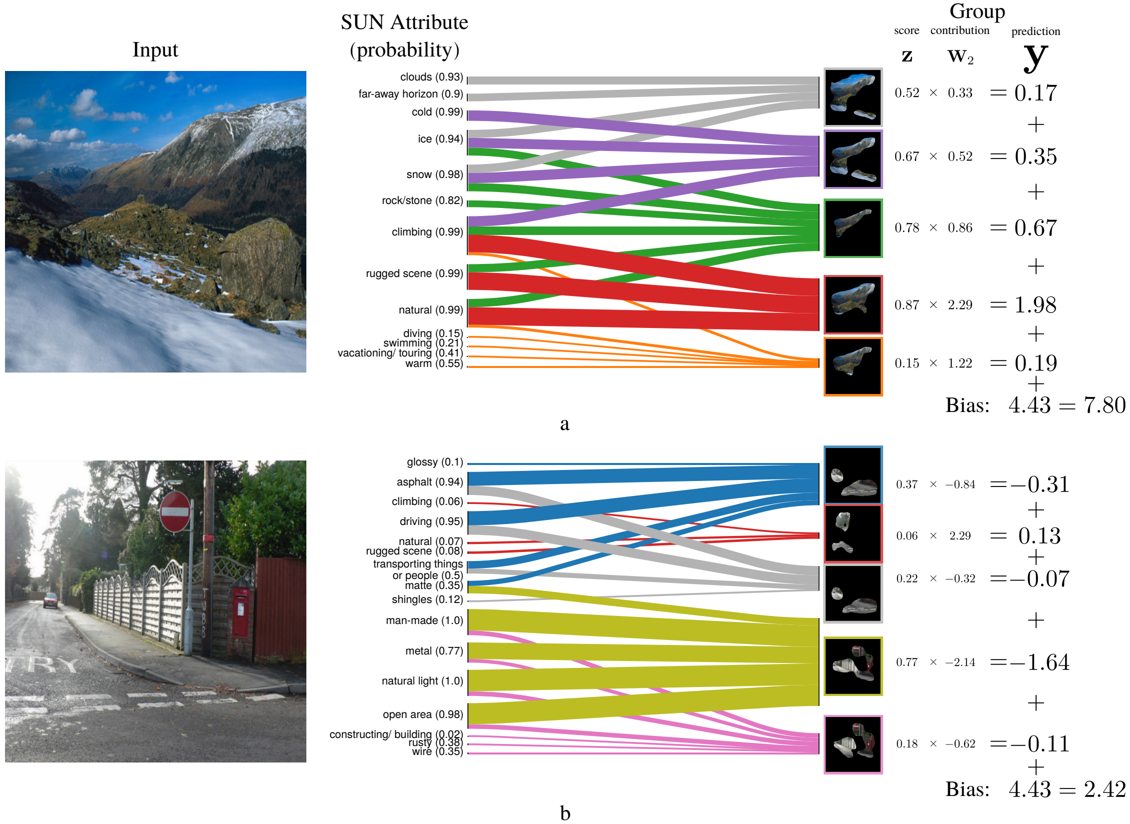
Semantic bottleneck: explaining via learned semantic concepts



Paper. Marcos, Fong, Lobry, Flamary, Courty, Tuia, Contextual semantic interpretability, *ACCV 2020*
https://arxiv.org/abs/2009.08720

Understand: explainable AI for subjective problems (and not only)

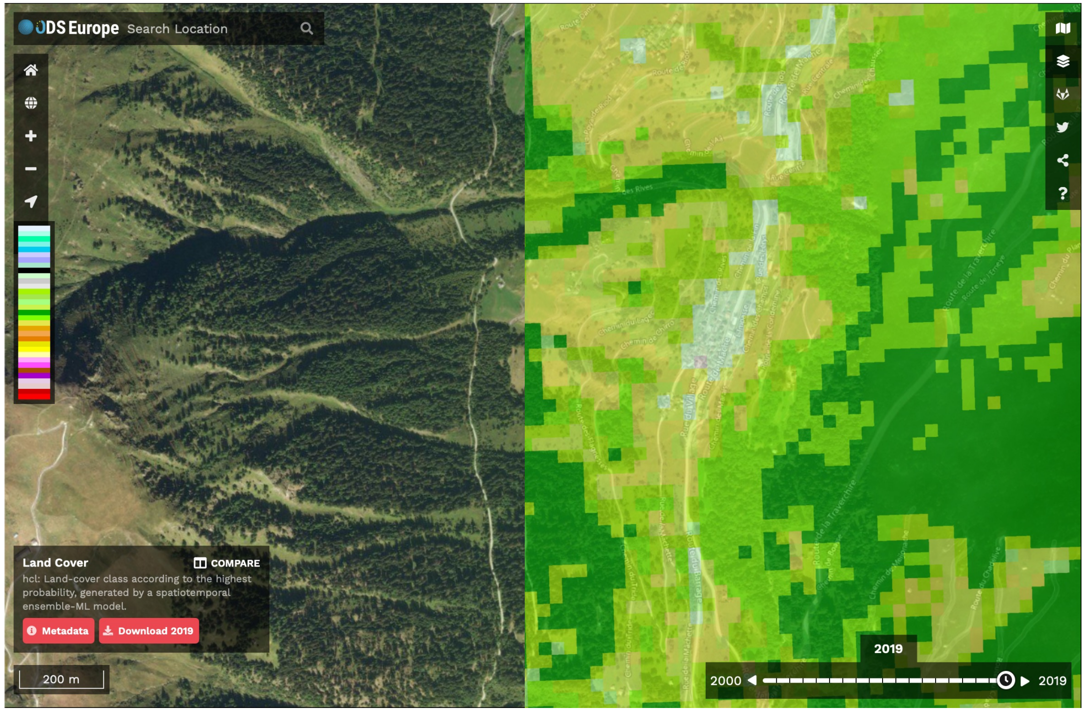
- We can now interpret scenicness in *"easy-to-understand and argue with"* terms



D. Marcos, S. Lobry, R. Fong, N. Courty, R. Flamary, and D. Tuia. Contextual semantic interpretability. In *Asian Conference on Computer Vision (ACCV)*, Kyoto, Kapan, 2020. https://arxiv.org/abs/2009.08720

Understand: explainable AI for subjective problems (and not only)

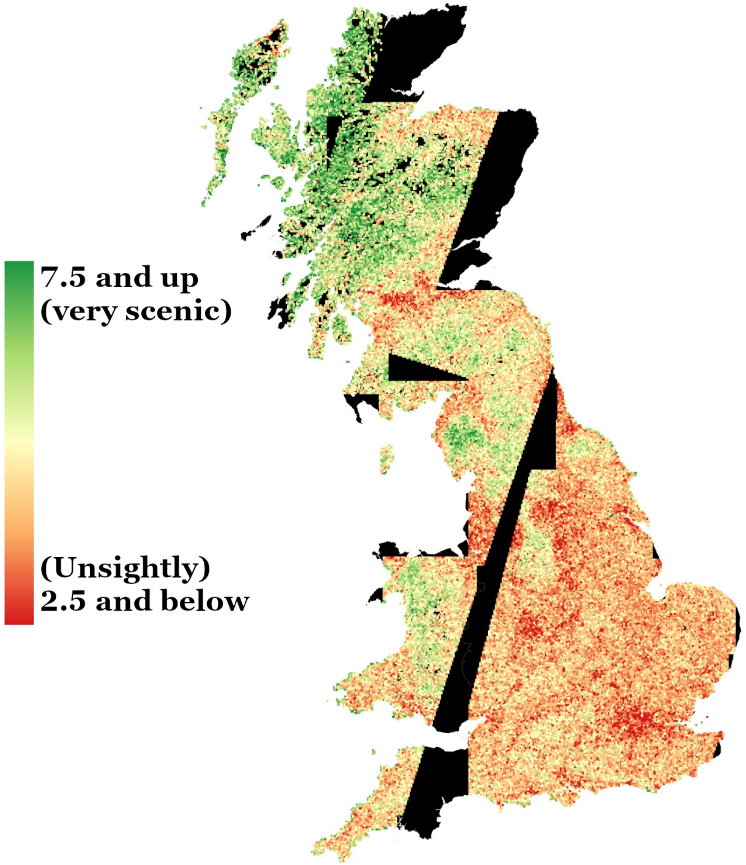
- We can study the same effects from satellite images
- And use land cover types as explanations
- For instance using Corine Land Cover
- Sentinel-2 images being available everywhere, we can scale findings!



A. Levering, D. Marcos, and D. Tuia. On the relation between landscape beauty and land cover: a case study in the U.K. at Sentinel-2 resolution with interpretable AI. *ISPRS J. Int. Soc. Photo. Remote Sens.*, 177:194–203, 2021. https://www.sciencedirect.com/science/article/pii/S0924271621001234

Understand: explainable AI for subjective problems (and not only)

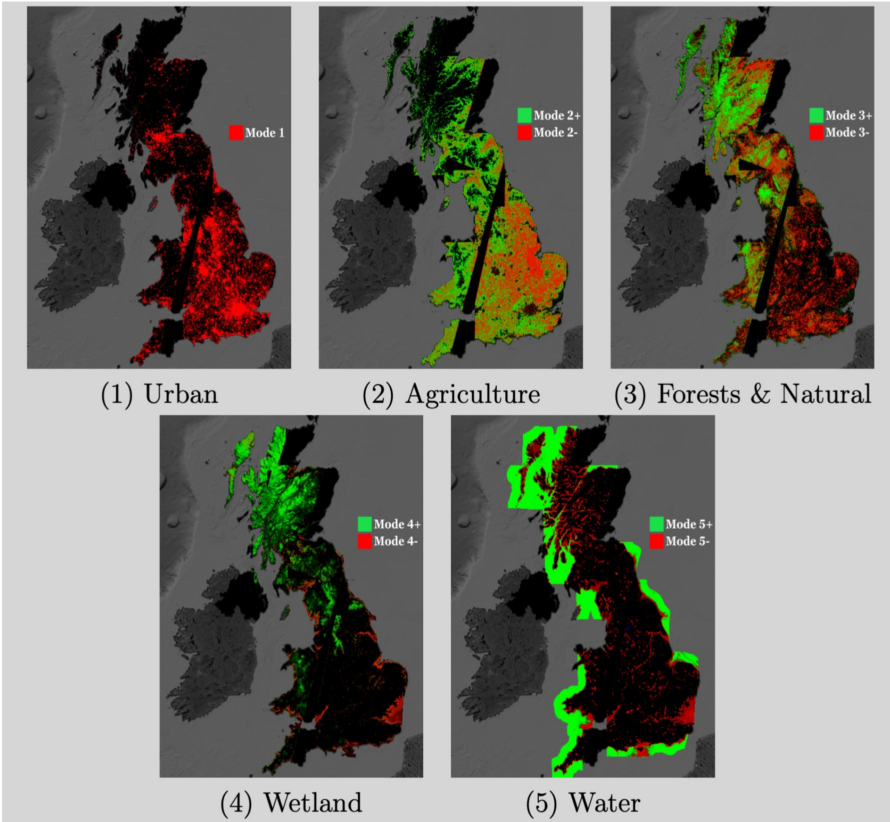
- We can study the same effects from satellite images
- And use land cover types as explanations
- For instance using Corine Land Cover
- Sentinel-2 images being available everywhere, we can scale findings!



A. Levering, D. Marcos, and D. Tuia. On the relation between landscape beauty and land cover: a case study in the U.K. at Sentinel-2 resolution with interpretable AI. *ISPRS J. Int. Soc. Photo. Remote Sens.*, 177:194–203, 2021. https://www.sciencedirect.com/science/article/pii/S0924271621001234

Understand: explainable AI for subjective problems (and not only)

- We can study the same effects from satellite images
- And use land cover types as explanations
- For instance using Corine Land Cover
- Sentinel-2 images being available everywhere, we can scale findings!

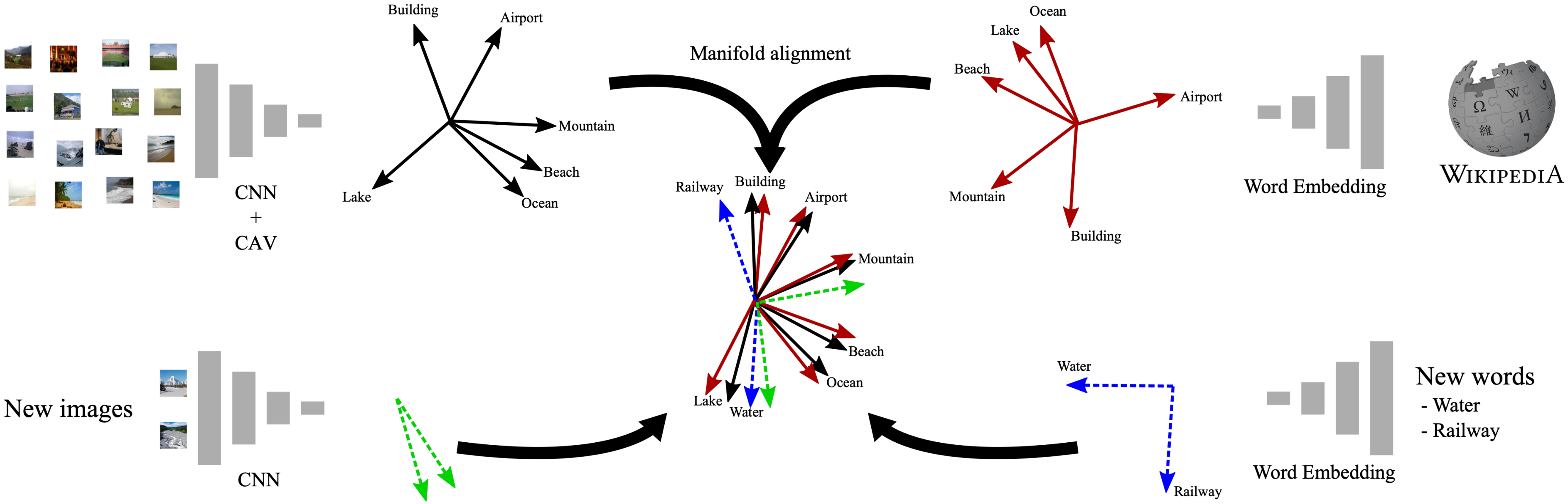


A. Levering, D. Marcos, and D. Tuia. On the relation between landscape beauty and land cover: a case study in the U.K. at Sentinel-2 resolution with interpretable AI. *ISPRS J. Int. Soc. Photo. Remote Sens.*, 177:194–203, 2021. https://www.sciencedirect.com/science/article/pii/S0924271621001234

EPFL

Moving forward: textual, open set interpretations

What we know

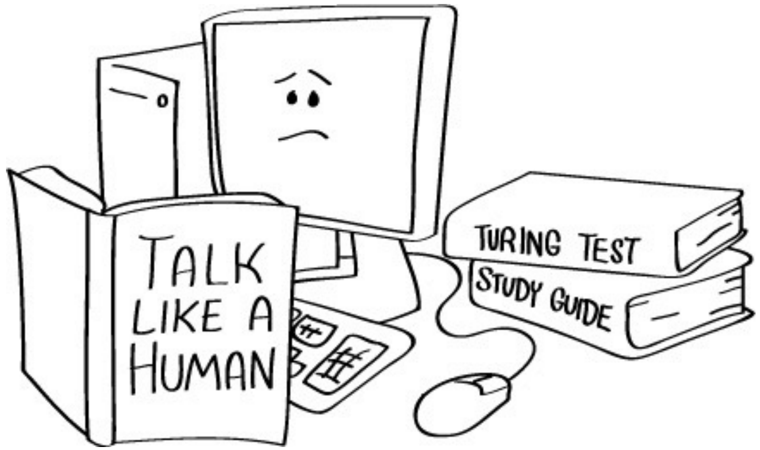


What we don't Know (yet)

P. Arendsen, D. Marcos, and D. Tuia. Concept discovery for the interpretation of landscape scenicness. *Mach. Learn. Knowledge Extraction*, 2(4):397–413, 2020. https://www.mdpi.com/2504-4990/2/4/22

Two (out of six) questions to move forward

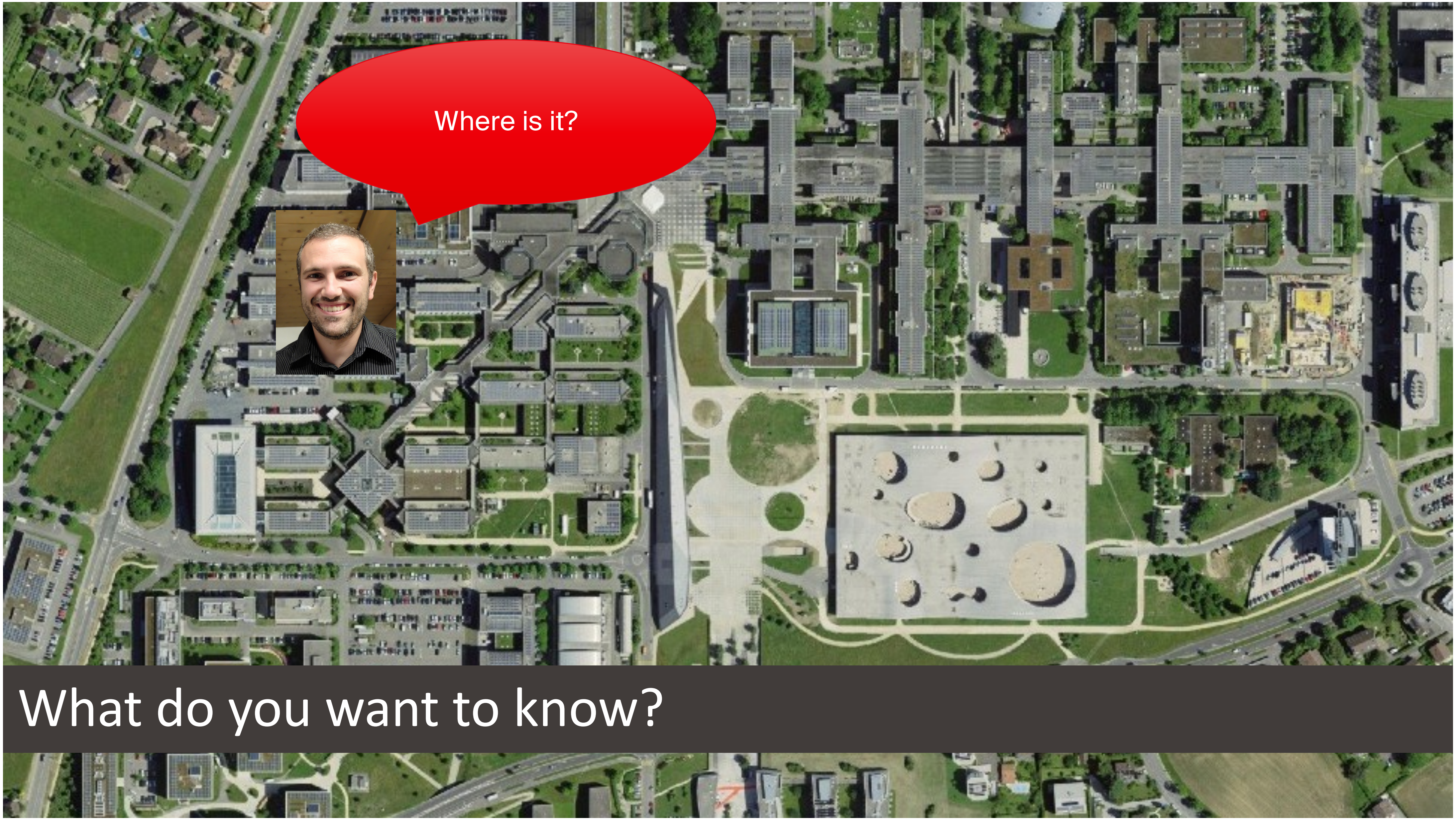
Understand what is going on within CNNs



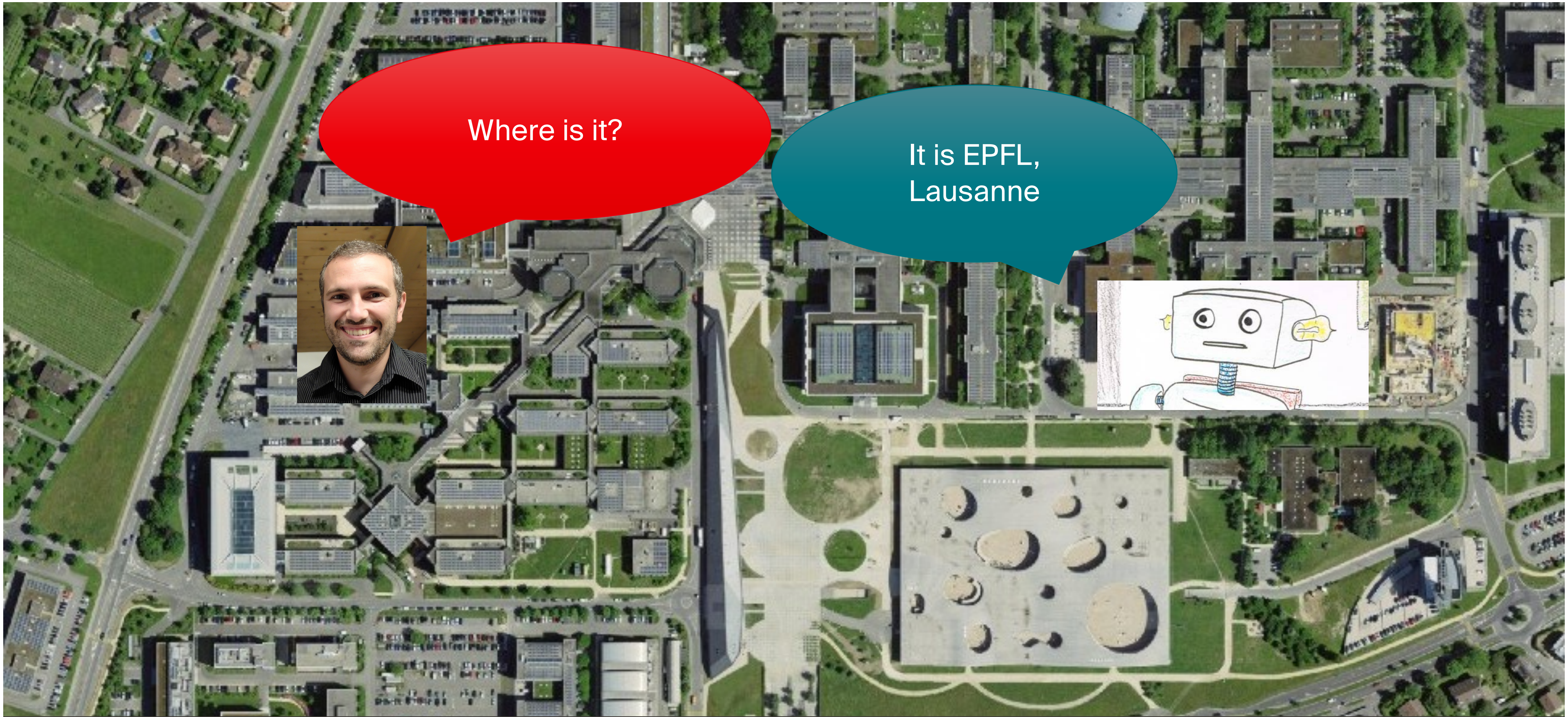
Make them accessible to anyone

Interaction, NLP

What do you want to know?

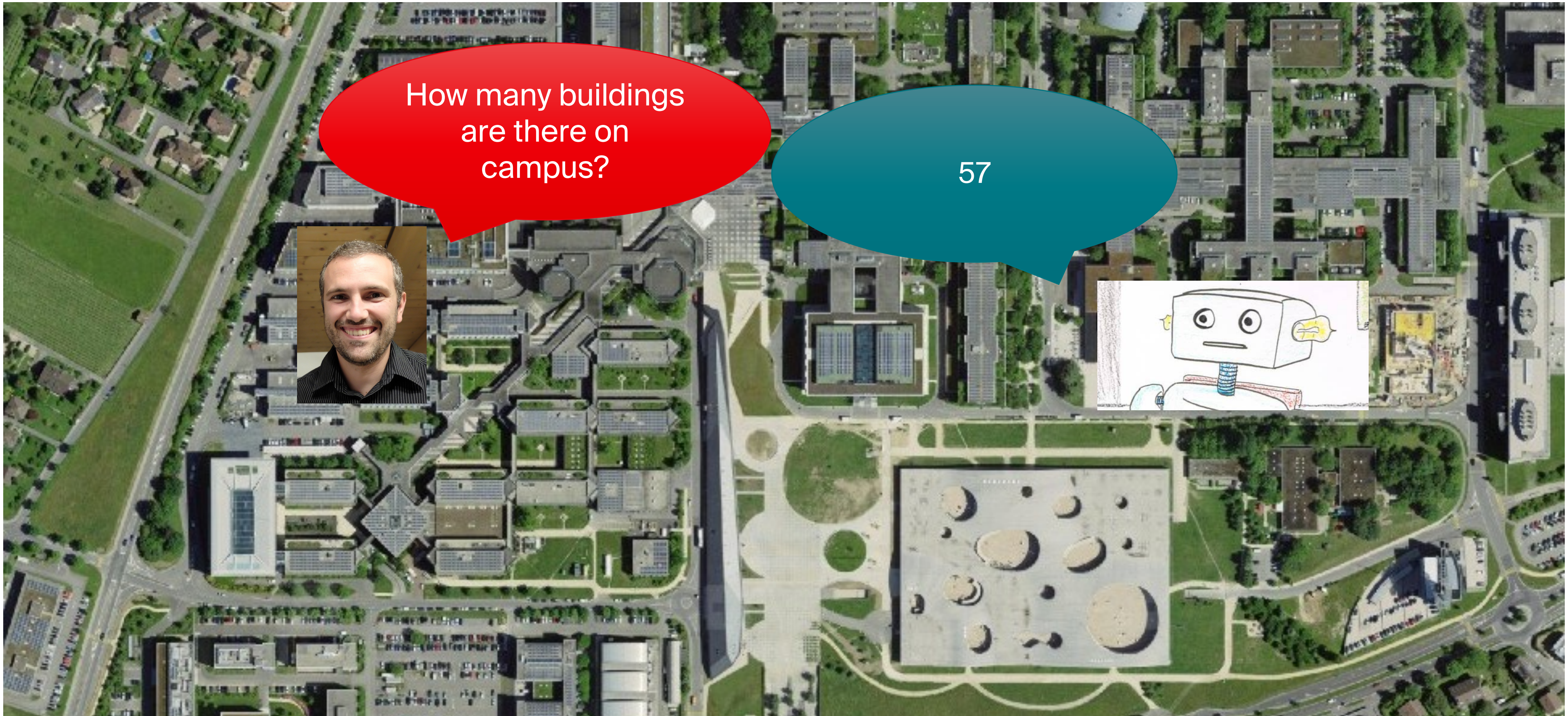


What do you want to know?



What do you want to know?

What do you want to know?



What do you want to know?

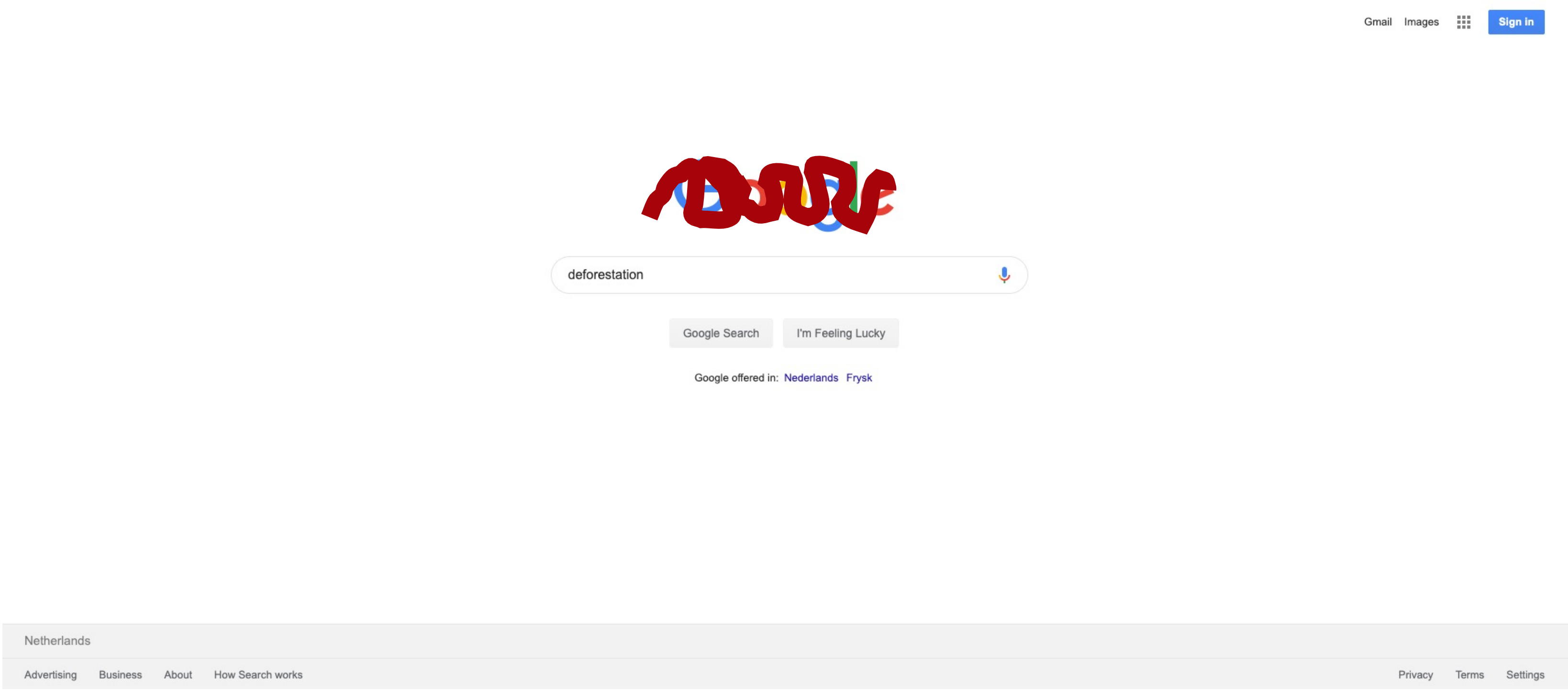
EPFL

This has great potential, but

- Non-experts are … **non technical** experts.
- Non-experts want answers to **specific questions**.
- Non-experts want to formulate questions as **sentences**.

What do we need?

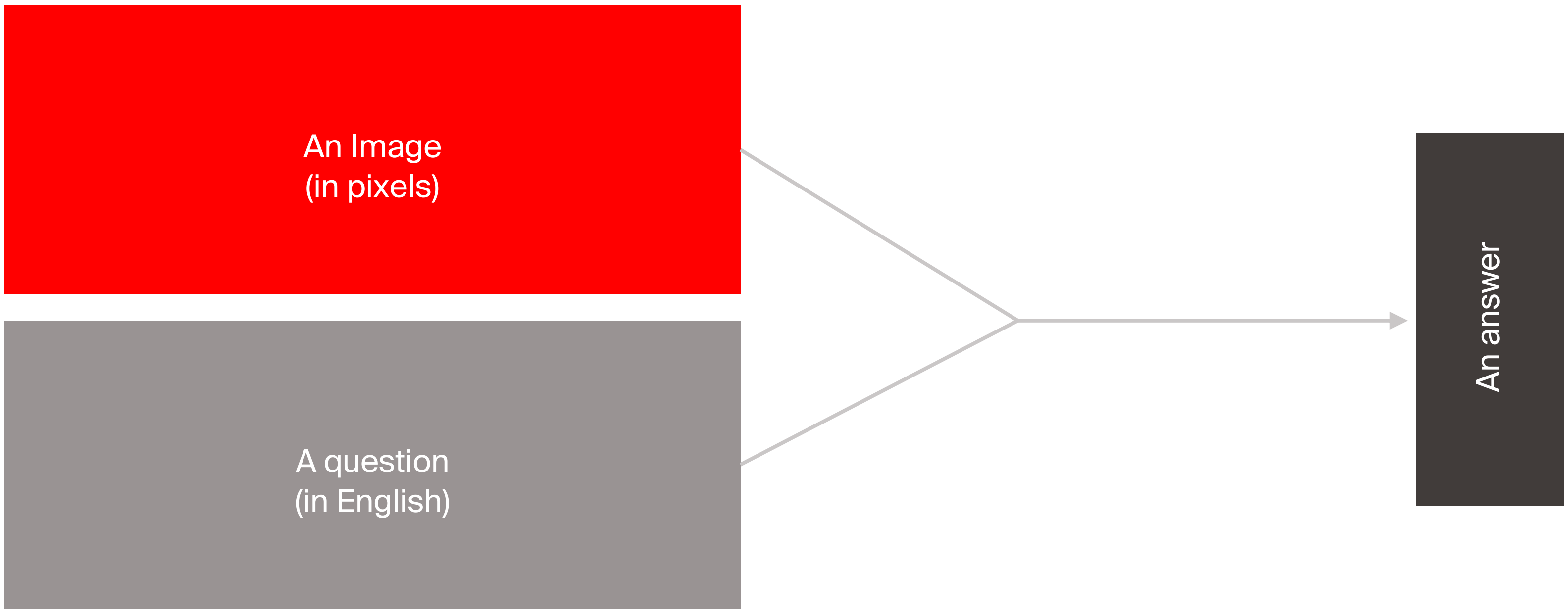
- For web-search it works a bit like that.



What do we need?

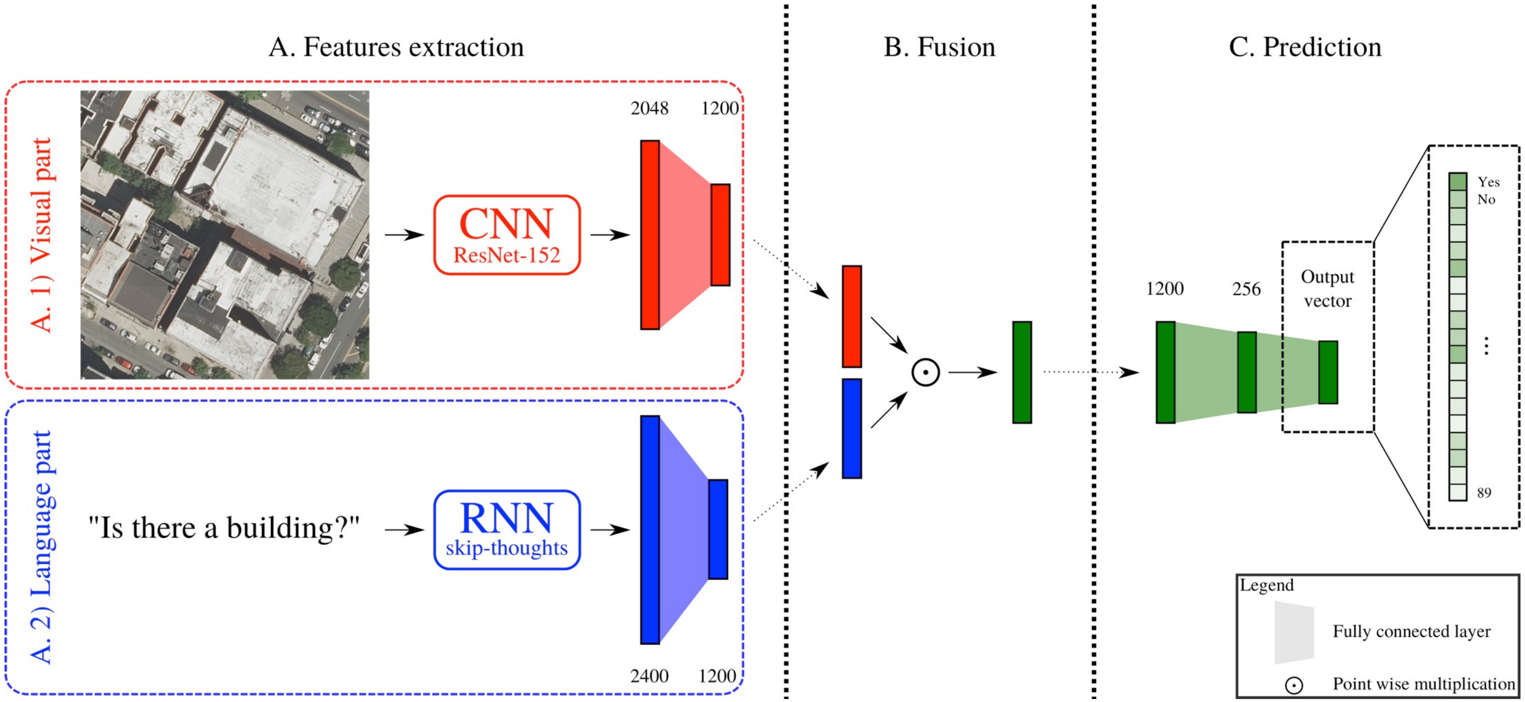
- For web-search it works a bit like that.
- With satellite images it just doesn't work (it's normal. It wasn't built for that)

Remote sensing visual question answering (RSVQA)



S. Lobry, D. Marcos, J. Murray, and **D. Tuia**. RSVQA: visual question answering for remote sensing data. *IEEE Trans. Geosci. Remote Sens.*, 58(12):8555–8566, 2020. https://arxiv.org/abs/2003.07333

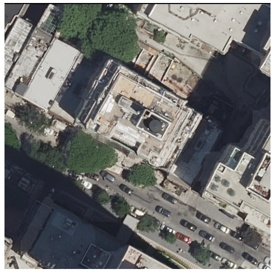
Remote sensing visual question answering (RSVQA)



S. Lobry, D. Marcos, J. Murray, and **D. Tuia**. RSVQA: visual question answering for remote sensing data. *IEEE Trans. Geosci. Remote Sens.*, 58(12):8555–8566, 2020. https://arxiv.org/abs/2003.07333

How do we train a model like this?

IMAGE



QUESTION
Is there a
cult place?

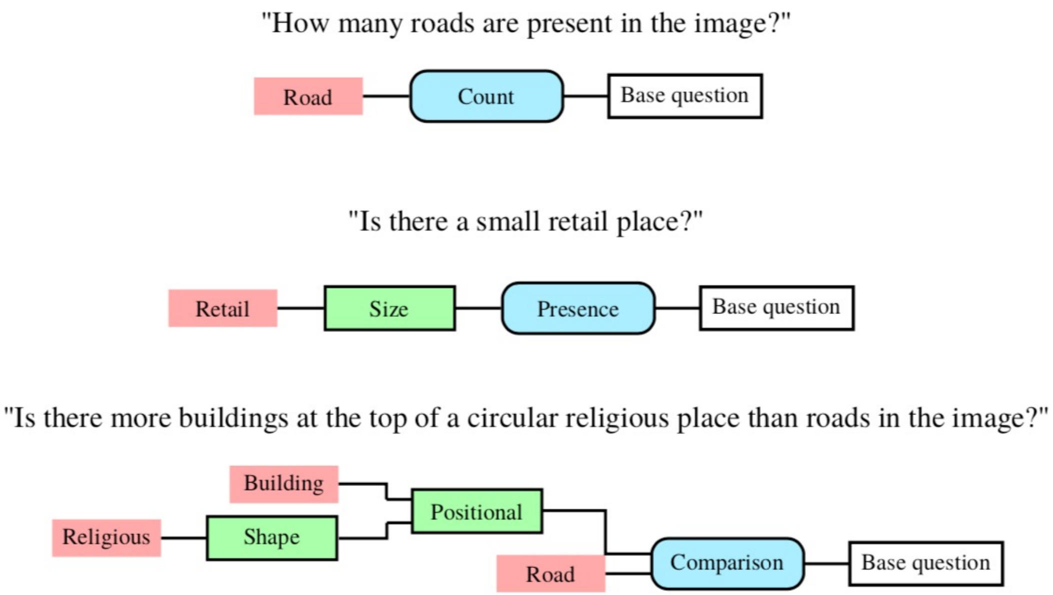
ANSWER
Yes!

We created two datasets:

- Sentinel-2 images (RGB)
  Entire Netherlands
  9 scenes
  772 image tiles
  77’200 triplets
- Aerial images (15cm)
  New York and Philadephia
  11’000 tiles (6000 m^2 each)
  1’110’000 triplets

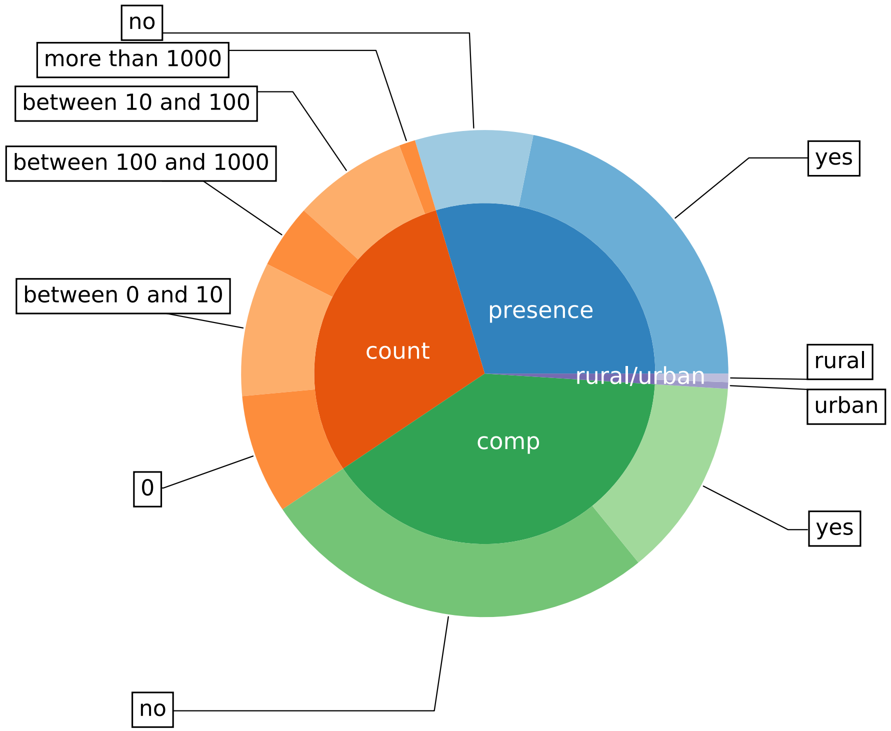
Generating text inputs / outputs

- We generated {image, **question**, answer} triplets with OSM



Generating text inputs / outputs

- We generated {image, question, **answer**} triplets
- We again use OSM



EPFL

Results

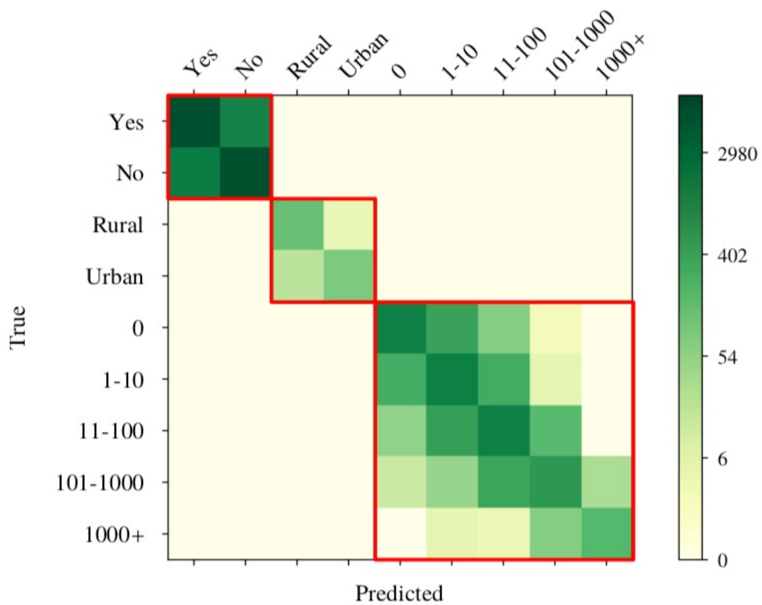
79% overall accuracy

73% if randomizing the image part

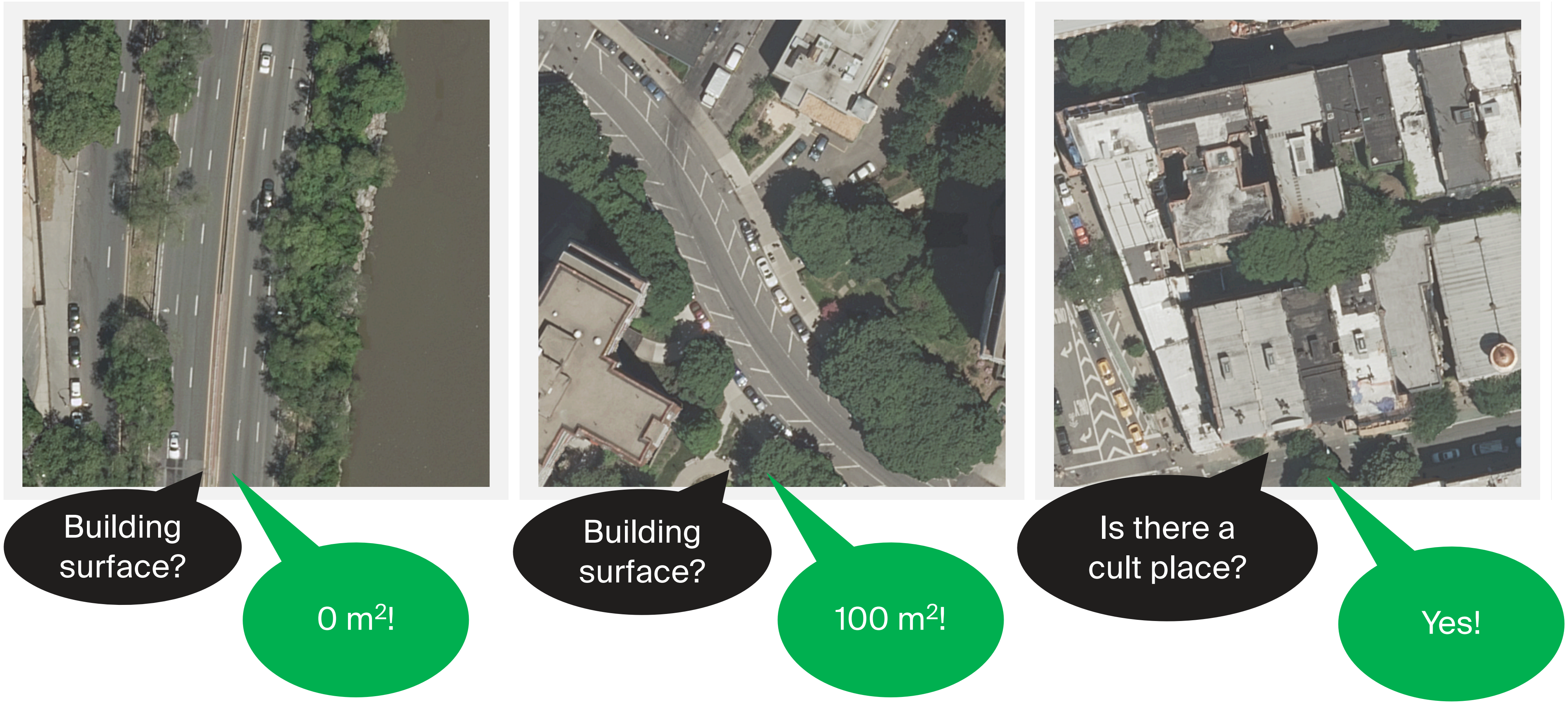
Count questions less accurate

Type	Accuracy
Count	67.01% (0.59%)
Presence	87.46% (0.06%)
Comparison	81.50% (0.03%)
Rural/Urban	90.00% (1.41%)
AA	81.49% (0.49%)
OA	79.08% (0.20%)

The model can make a good distinction between types of questions



Some results



From Lobry et al., IEEE TGRS 2020, https://arxiv.org/abs/2003.07333

Some results

How many buildings?

0!

How many bulidings?

1!

From Lobry et al., IEEE TGRS 2020, https://arxiv.org/abs/2003.07333

Get data and codes!

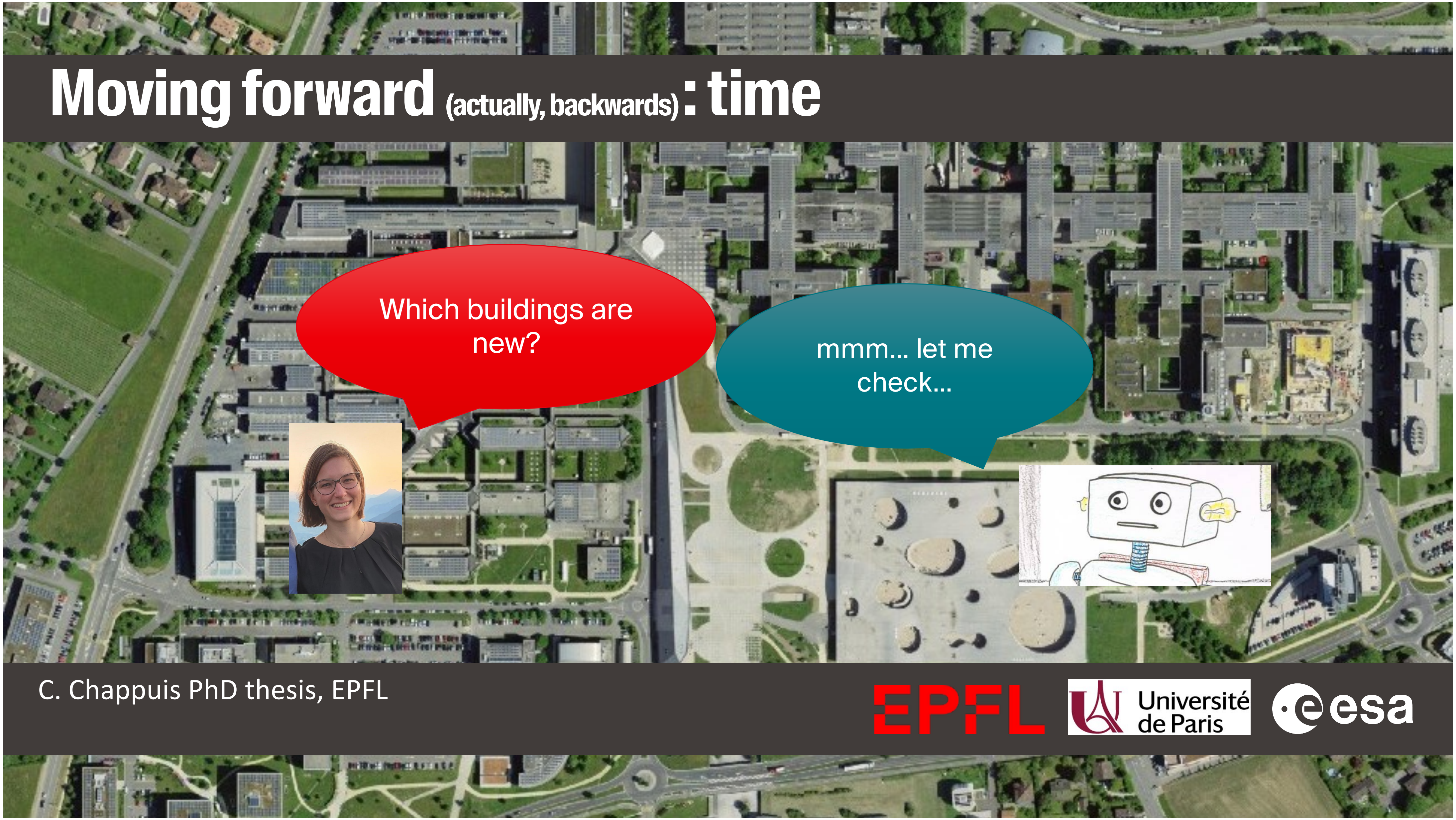
https://rsvqa.sylvainlobry.com/

DATASET

Very high resolution dataset	Low resolution dataset
10659	955664
Images	Questions

We have created a database using USGS' high resolution (15.24cm) orthorectified images and questions and answers derived from OSM. You can explore a subset of 50 images from this dataset here.

Moving forward (actually, backwards): time



Which buildings are new?

mmm… let me check…

C. Chappuis PhD thesis, EPFL

EPFL Université de Paris esa

In summary

EPFL

Earth observation and machine learning / computer vision are a match made in heaven.

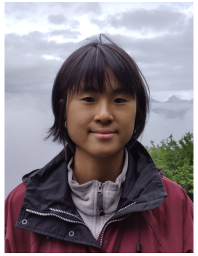
- There are many unsolved challenges for vision
- There is a lot of data out there

- and waaaaaaaay more importantly:

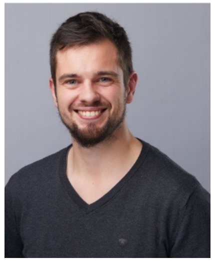
Making a difference with EO means making a difference for the future of our planet
For our air, water, food, biodiversity.

Thanks!

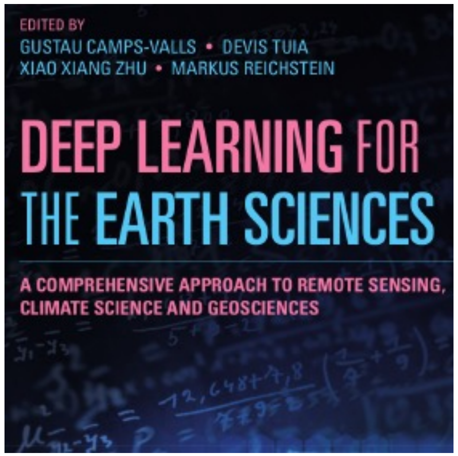
devis.tuia@epfl.ch, eceo.epfl.ch, @devistuia

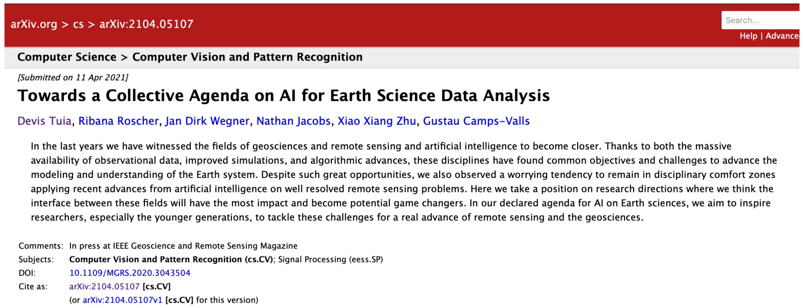


The book



The paper


arXiv.org > cs > arXiv:2104.05107

Search...

Help | Advance

Computer Science > Computer Vision and Pattern Recognition

[Submitted on 11 Apr 2021]

Towards a Collective Agenda on AI for Earth Science Data Analysis

Devis Tuia, Ribana Roscher, Jan Dirk Wegner, Nathan Jacobs, Xiao Xiang Zhu, Gustau Camps-Valls

In the last years we have witnessed the fields of geosciences and remote sensing and artificial intelligence to become closer. Thanks to both the massive availability of observational data, improved simulations, and algorithmic advances, these disciplines have found common objectives and challenges to advance the modeling and understanding of the Earth system. Despite such great opportunities, we also observed a worrying tendency to remain in disciplinary comfort zones applying recent advances from artificial intelligence on well resolved remote sensing problems. Here we take a position on research directions where we think the interface between these fields will have the most impact and become potential game changers. In our declared agenda for AI on Earth sciences, we aim to inspire researchers, especially the younger generations, to tackle these challenges for a real advance of remote sensing and the geosciences.

Comments: In press at IEEE Geoscience and Remote Sensing Magazine
Subjects: **Computer Vision and Pattern Recognition (cs.CV)**; Signal Processing (eess.SP)
DOI: 10.1109/MGRS.2020.3043504
Cite as: arXiv:2104.05107 **[cs.CV]**
(or arXiv:2104.05107v1 **[cs.CV]** for this version)

The network

ellis

European Laboratory for Learning and Intelligent Systems

Program “ML for Earth and Climate Science”