Meteors

Open Source Package for Explanations of Remotely-Sensed Images

DATA SCIENCE SUMMIT

Vladimir Zaigrajew,
Tymoteusz Kwieciński

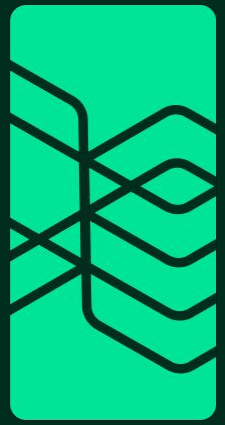
Warsaw University of Technology,
MI2.ai

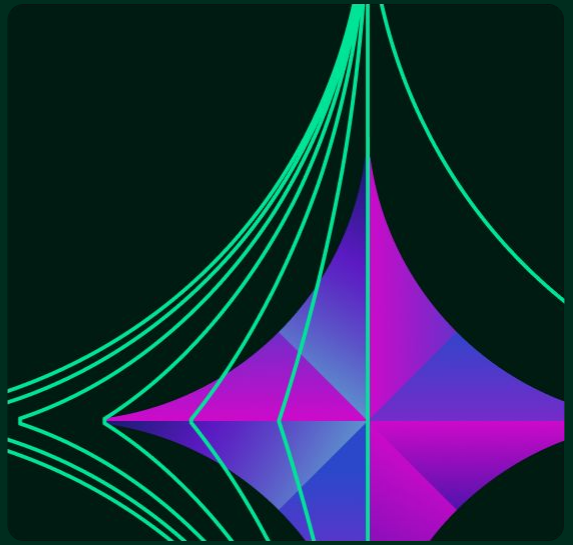
Overview

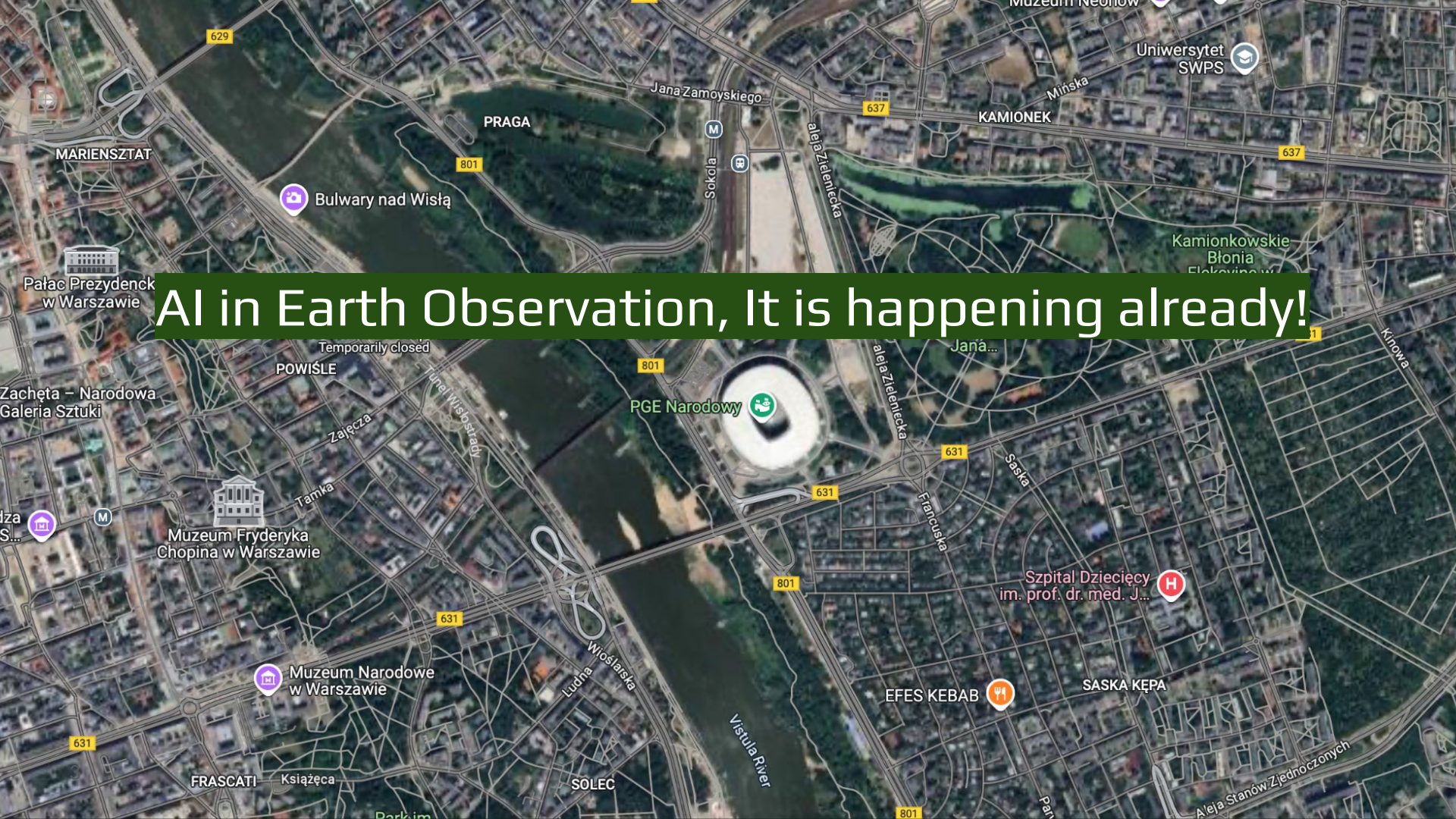
1. Explainable AI in Earth Observation, What's it all about?

2. From Research to Reality: Building XAI Tools

3. Explainable AI Open Source Package for Earth Observation - Meteors






AI in Earth Observation, It is happening already!
629
PRAGA
MARIENSZTAT
Bulwary nad Wisłą
Jana Zamoyskiego
637
KAMIONEK
Uniwersytet SWPS
Mińska
801
Sokola
aleja Zielenecka
Kamionkowskie Błonia
Pałac Prezydencki w Warszawie
Temporarily closed
POWIŚLE
Zachęta – Narodowa Galeria Sztuki
Tunel Wisłostrady
PGE Narodowy
Zajęcza
Tamka
Muzeum Fryderyka Chopina w Warszawie
631
Saska
Francuska
Szpital Dziecięcy im. prof. dr. med. J...
Muzeum Narodowe w Warszawie
Wioślarska
Ludna
EFES KEBAB
SASKA KĘPA
Vistula River
FRASCATI
Książęca
SOLEC
Aleja Stanów Zjednoczonych
Kinowa

TL;DR A lot!

Climate & Environment

- Optimize renewable energy systems
- Monitor deforestation and biodiversity
- Predict extreme weather events

Resource Management

- Optimize energy consumption
- Smart waste management
- Water conservation systems

Research & Modeling

- Climate change predictions
- Ecosystem modeling
- Species preservation

Data Analytics

- Carbon footprint tracking
- Environmental impact assessment
- Sustainability metrics

What Can AI Do From Space for Our Planet?

- Disaster Detection and Assessment

The Power of AI: New Algorithm for Building Detection

The war in Ukraine has caused widespread devastation, leaving countless buildings damaged or destroyed. Rebuilding efforts are crucial, and Azenzus, a company dedicated to innovative inspection solutions, is proud to play a role.

Our Machine Learning team has developed a groundbreaking algorithm that can effectively detect buildings on satellite images.

https://www.rebuilding-ua.org/the-power-of-ai-new-algorithm-for-building-detection/

MIT Technology Review

Featured Topics Newsletters Events Audio SIGN IN SUBSCRIBE

POLICY

How AI can actually be helpful in disaster response

Humanitarian teams in Turkey and Syria are using machine learning to quickly scope out earthquake damage and strategize rescue efforts

By Tate Ryan-Mosley February 20, 2023

https://www.technologyreview.com/2023/02/20/1068824/ai-actually-helpful-disaster-response-turkey-syria-earthquake

What Can AI Do From Space for Our Planet?

- Disaster Detection and Assessment
- Agricultural Monitoring


BBC

Home News US Election Sport Business Innovation Culture Arts Travel Earth Video Live

US farms are making an urgent push into AI. It could help feed the world

27 March 2024

Share Save

Sam Becker

Features correspondent

https://www.bbc.com/worklife/article/20240325-artificial-intelligence-ai-us-agriculture-farming


Forbes

INNOVATION > CLOUD

How Indian Farmers Are Using AI To Increase Crop Yield

By Janakiram MSV, Senior Contributor. I cover emerging technologies with a f... Follow Author

Feb 01, 2024 at 05:41am EST

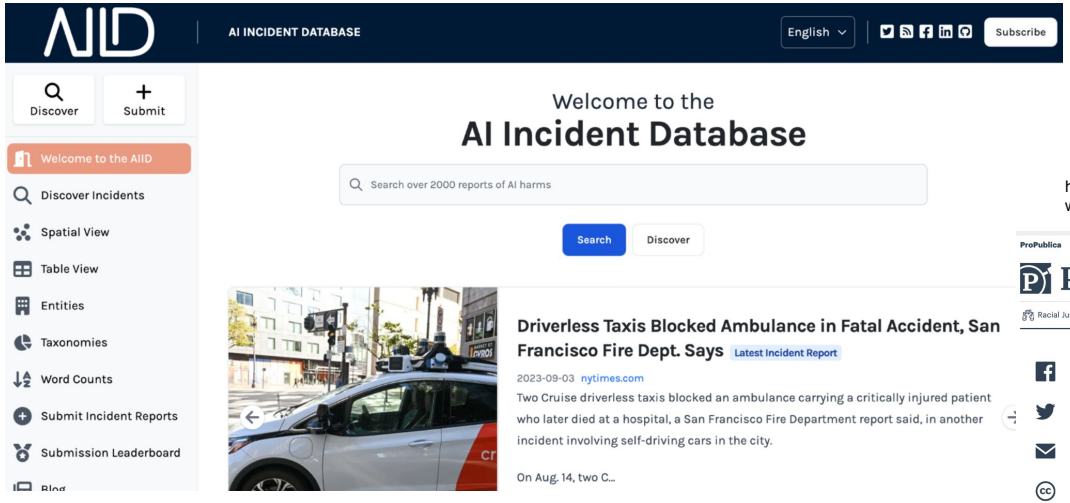
https://www.forbes.com/sites/janakirammsv/2024/02/01/how-indian-farmers-are-using-ai-to-increase-crop-yield/

“Farmers in the program grew 21 percent more plants per acre while using 9 percent less pesticide and 5 percent less fertilizer, according to the World Economic Forum.”

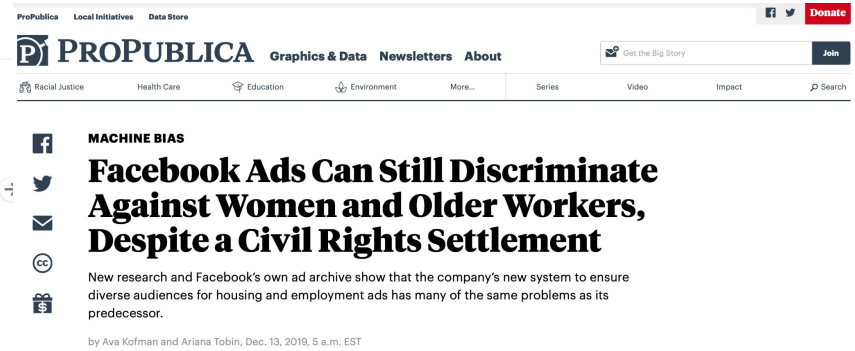
Explainable AI in Earth Observation, What's it all about?

Explainable AI

Read more at: https://incidentdatabase.ai/



https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/



https://www.propublica.org/article/facebook-ads-can-still-discriminate-against-women-and-older-workers-despite-a-civil-rights-settlement

MI2: Experts in Explainable AI

Explainable AI: Why It Matters

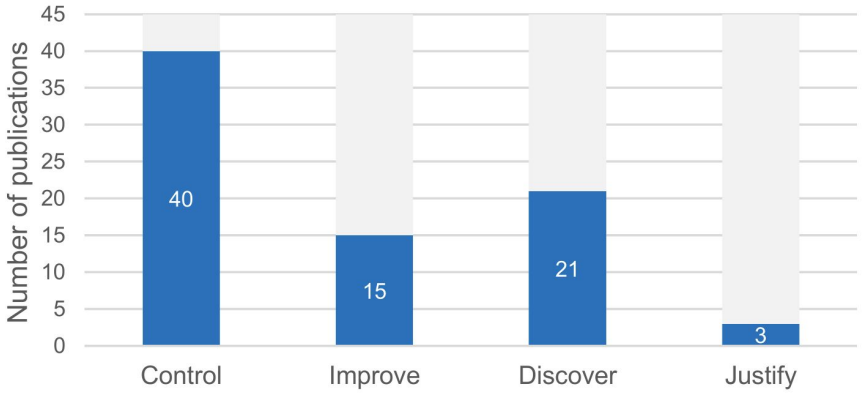


Fig. 1. Motivations underlying the integration of explainability into ML workflows for Earth Observation as perceived by published works.

Explainable AI for earth observation: A review including societal and regulatory perspectives, International Journal of Applied Earth Observations and Geoinformation, 12 June 2022

MI2: Advancing XAI for Earth Observation

Funding Organization

esa

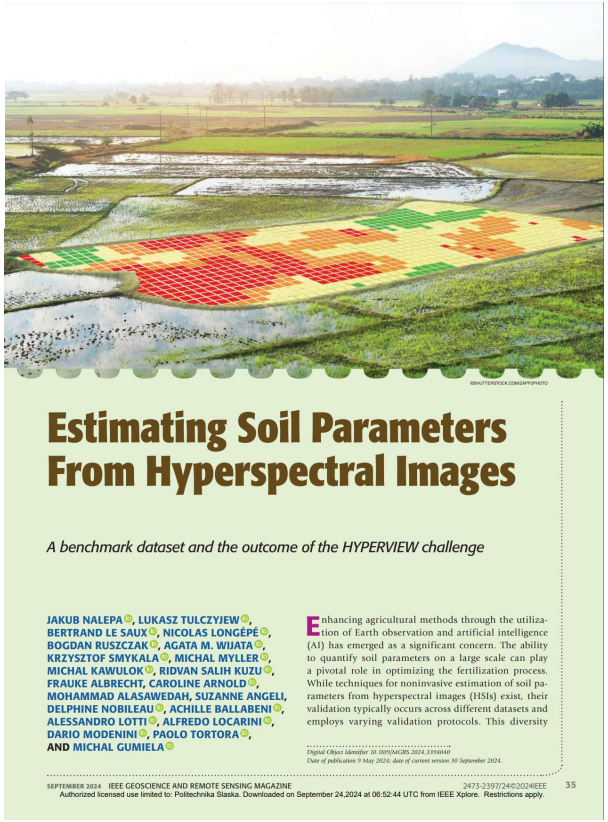
Synergy

MI

KP LABS
Mission Complete

Red Team:
- Analyse
- Explain
- Advise

Blue Team:
- Prepare Data
- Train AI Models
- Improve Model



Estimating Soil Parameters From Hyperspectral Images

A benchmark dataset and the outcome of the HYPERVIEW challenge

JAKUB NALEPA, LUKASZ TULCZYJEW, BERTRAND LE SAUX, NICOLAS LONGÉPÉ, BOGDAN RUSZCZAK, AGATA M. WIJATA, KRZYSZTOF SMYKALA, MICHAL MYLLER, MICHAL KAWULOK, RIDVAN SALIH KUZU, FRAUKE ALBRECHT, CAROLINE ARNOLD, MOHAMMAD ALASAWEDAH, SUZANNE ANGELI, DELPHINE NOBILEAU, ACHILLE BALLABENI, ALESSANDRO LOTTI, ALFREDO LOCARINI, DARIO MODENINI, PAOLO TORTORA, AND MICHAL GUMIELA

Enhancing agricultural methods through the utilization of Earth observation and artificial intelligence (AI) has emerged as a significant concern. The ability to quantify soil parameters on a large scale can play a pivotal role in optimizing the fertilization process. While techniques for nonimvasive estimation of soil parameters from hyperspectral images (HSIs) exist, their validation typically occurs across different datasets and employs varying validation protocols. This diversity

Digital Object Identifier 10.1109/MGRS.2024.3394040
Date of publication 9 May 2024; date of current version 30 September 2024.

TL;DR: HYPERVIEW Challenge

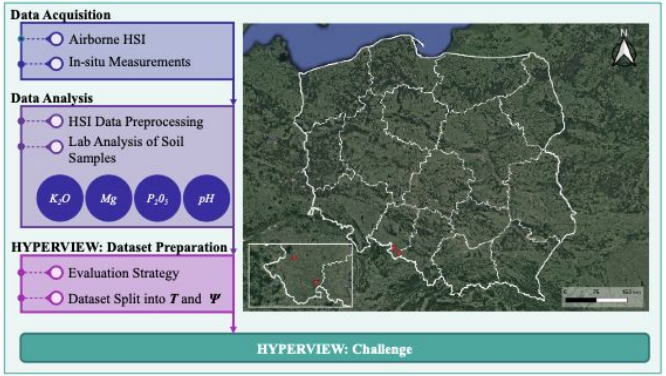
Goal: Enable rapid, automated soil analysis for agricultural fields.

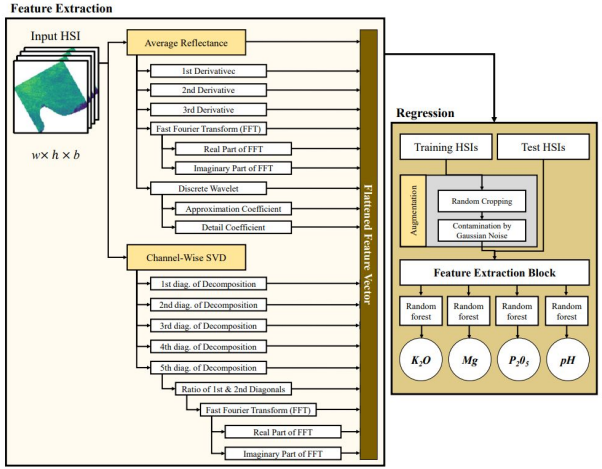
Dataset Overview:

- 1,732 training and 1,154 testing hyperspectral images
- Variable width/height ratios with 150-192 spectral bands (462-942 nm)
- Ground truth: 4 soil parameters (P, K, Mg, pH)

Models from HYPERVIEW Challenge:

1. EAGLEEYES (Random Forest)
2. ViT (Vision Transformer)
3. CNN (Convolutional Neural Network)





Red Teaming Models for Hyperspectral Image Analysis Using Explainable AI

Vladimir Zaigrajew, Hubert Baniecki, Łukasz Tulczyjew, Agata M. Wijata, Jakub Nalepa, Nicolas Longépé, Przemek Biecek







Warsaw University of Technology

UNIWERSYTET WARSZAWSKI

Silesian University of Technology

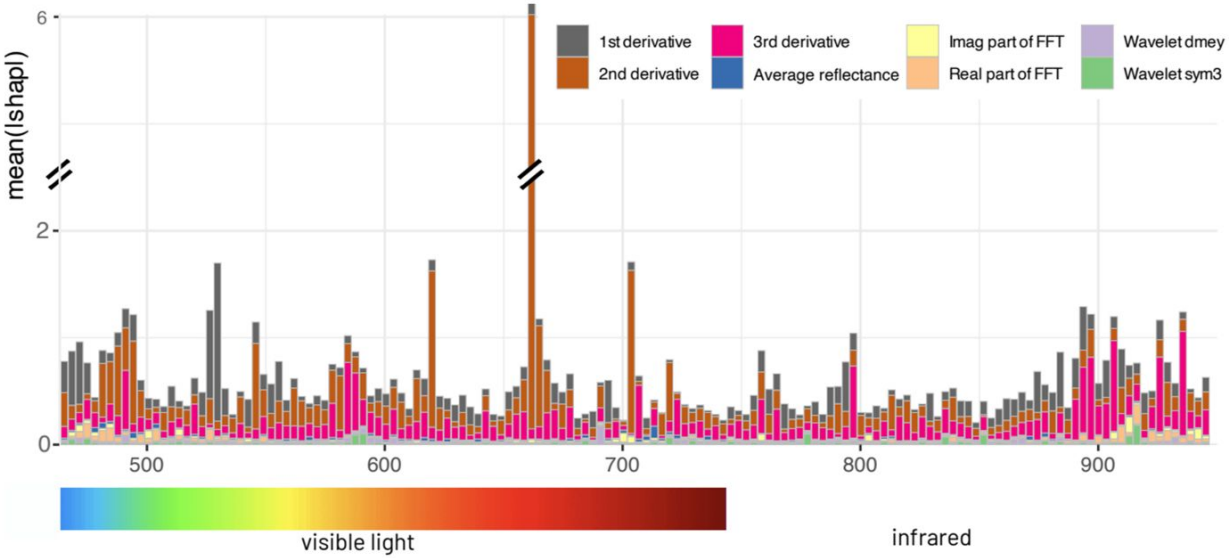
KP LABS Mission Complete

MI

esa

ICLR

Added Value: Wavelength Importance



Red Teaming Models for Hyperspectral Image Analysis Using Explainable AI, Arxiv, 2024

Do We Need All Features?

Dataset	F. selection	*P*	*K*	*Mg*	*pH*
HYPERVIEW	✗ (1200)	22.6	48.4	31.3	0.206
	✓ (3)	23.3 (+3%)	48.7 (+1%)	32.6 (+4%)	0.213 (+4%)
INTUITION-1	✗ (1490)	22.5	48.5	30.9	0.204
	✓ (3)	22.7 (+1%)	48.5 (0%)	32.5 (+5%)	0.214 (+5%)
HYPERVIEW (spatial)	✗ (2400)	22.4	47.8	31.0	0.205
	✓ (5)	23.0 (+3%)	48.9 (+2%)	32.5 (+5%)	0.212 (+3%)
INTUITION-1 (spatial)	✗ (3026)	22.3	47.8	31.1	0.204
	✓ (6)	23.2 (+4%)	50.1 (+5%)	32.4 (+4%)	0.209 (+2%)

Comparison of model performance on the test dataset using MAE (↓) for soil parameters. Reducing number of features to less than 1% performed similarly to baseline models, as confirmed by two-sample t-test, with minimal metric decline.

Deep Learning: CNN and Transformers

What worked ideally for Random Forest was not optimal for ViT and CNN. We needed more reliable, performant, and tailored methods for deep learning models, such as Integrated Gradients or LIME.

Explainable AI for Images

Concept-Based Explanation:

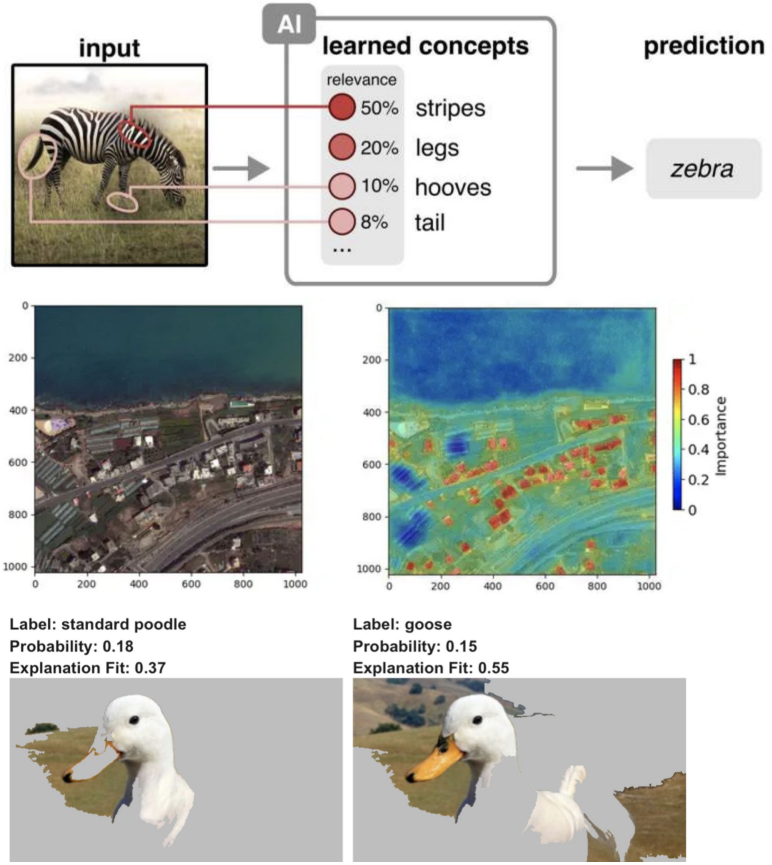
- Uses surrogate models to detect learned concepts within the main model
- Main model predicts classes while surrogate model identifies intermediate concepts

Gradient-Based Explanation:

- Highlights important input features by analyzing gradients
- Shows which pixels/regions influenced the model's decision

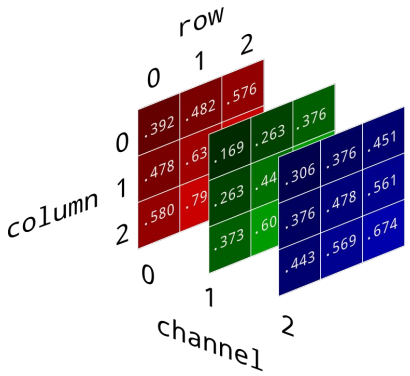
Perturbation-Based Explanation:

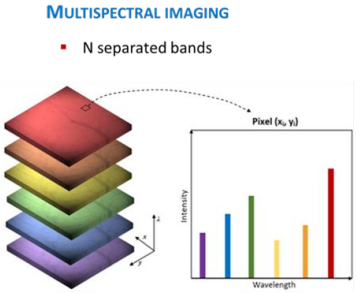
- Identifies critical regions by systematically masking parts of the image
- Reveals which areas must remain visible for classification

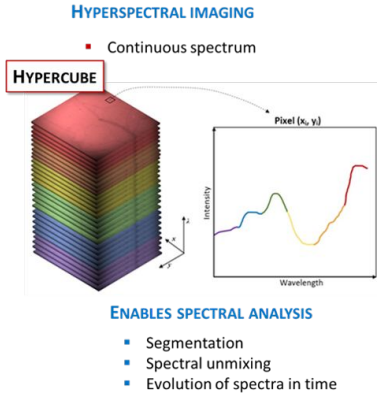


Explainable AI in Space - Why it is difficult?

- Earth observation data is 3D (two spatial dimensions + spectral bands)
- Traditional XAI methods are designed for 2D images

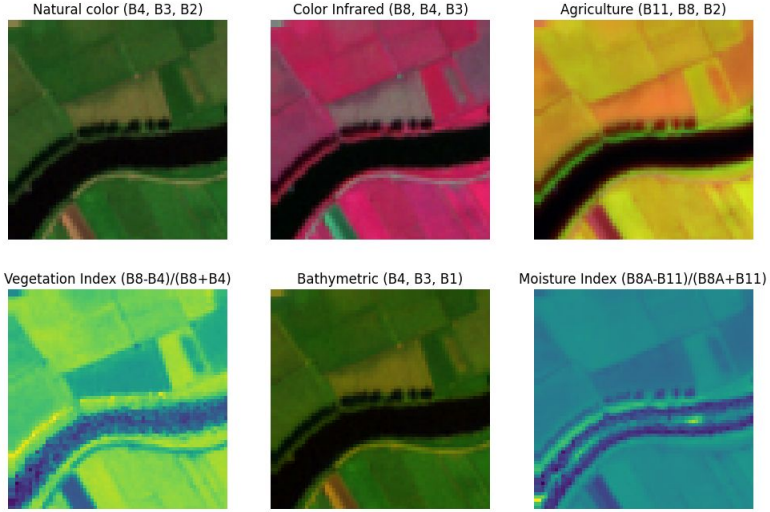






Explainable AI in Space - Why it is difficult?

- Earth observation data is 3D (two spatial dimensions + spectral bands)
- Traditional XAI methods are designed for 2D images
- Channels do matter, as each band captures unique surface information



Explainable AI for Earth Observation: Current Methods, Open Challenges, and Opportunities Arxiv, 8 Nov, 2023

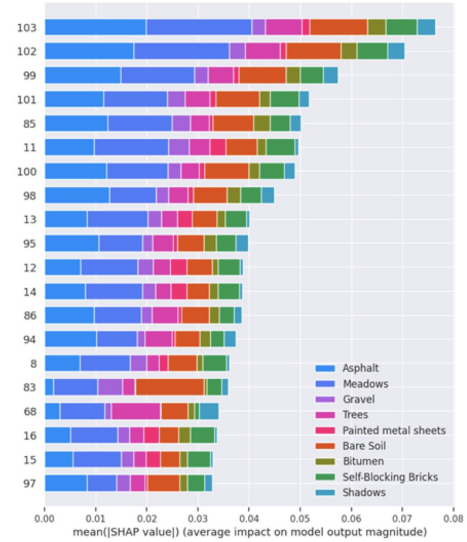
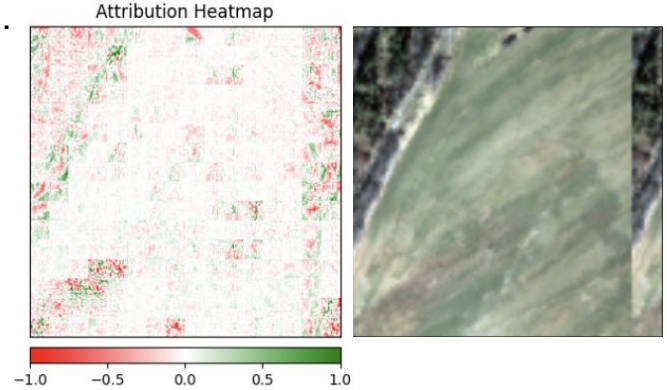


Figure 10: SHAP value plot for band-based interpretability, using Random Forest classifier on Pavia University hyperspectral image

Deep Learning: CNN and Transformers

What worked ideally for Random Forest was not optimal for ViT and CNN. We needed more reliable, performant, and tailored methods for deep learning models, such as Integrated Gradients or LIME.

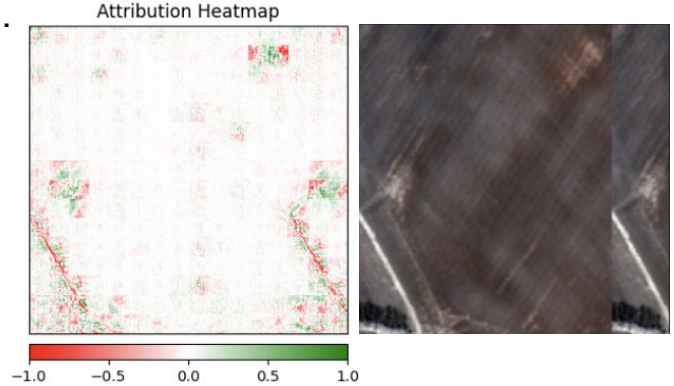


Key Findings:

- Without masking, methods focused excessively on forest and road artifacts

Deep Learning: CNN and Transformers

What worked ideally for Random Forest was not optimal for ViT and CNN. We needed more reliable, performant, and tailored methods for deep learning models, such as Integrated Gradients or LIME.

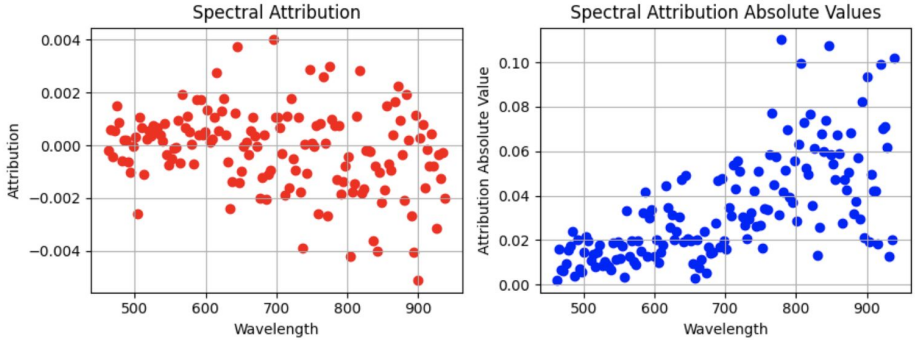


Key Findings:

- Without masking, methods focused excessively on forest and road artifacts
- Mirror padding caused models to double-focus on mirrored fields, compromising ground truth accuracy

Deep Learning: CNN and Transformers

What worked ideally for Random Forest was not optimal for ViT and CNN. We needed more reliable, performant, and tailored methods for deep learning models, such as Integrated Gradients or LIME.



Key Findings:

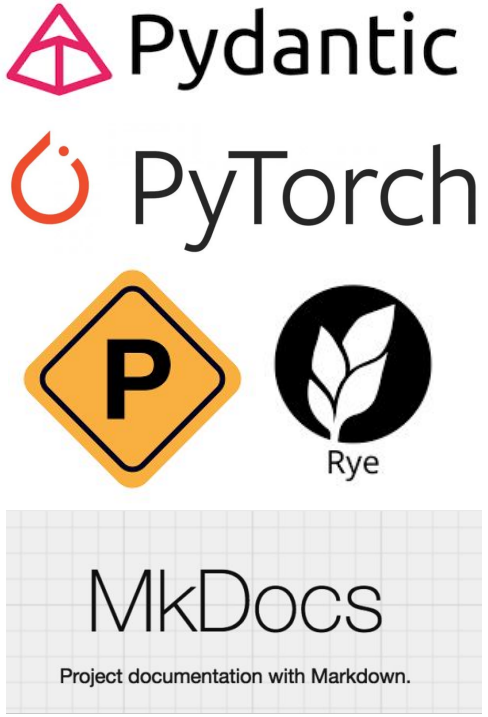
- Without masking, methods focused excessively on forest and road artifacts
- Mirror padding caused models to double-focus on mirrored fields, compromising ground truth accuracy
- Higher wavelength frequencies showed increased sensitivity and impact on model performance

Meteors package:
Making XAI for Remote Sensing Easy

Meteors objectives

- Knowledge base for XAI in remote sensing
- **easy XAI** framework for everyone
- intuitive and transparent
- well tested and **high-quality code**
- promoting the use of XAI in space

Package is for scientists and ML engineers who want to red-team their models used in space, but don't know how to do it.



few of innovative and well known tools used in package development

Meteors: Open Source Package for Explanations of Remotely-Sensed Images

unique XAI package for hyperspectral modeling

Emerged from our code used during the XAI analysis

Extensive **documentation** with real use-cases

7 different explanation methods, and even more yet to come!

supports **multiple problem types**:
regression, classification, segmentation

link to the package documentation

How can the package be used?
Few examples from the tutorials

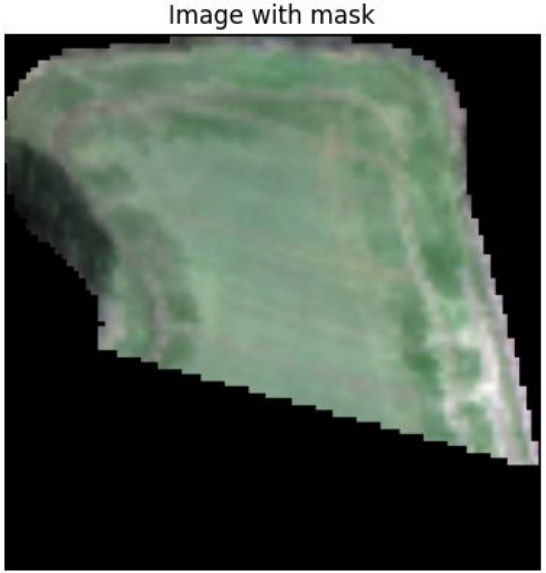
Simple visualization of an image

Numpy image from HYPERVIEW Challenge saved as `image`, wavelengths of the image saved as `wavelengths` and a `binary mask` covering the unwanted regions. How to visualize the image in RGB?

1. ensure correct rotation of the image
2. select proper bands corresponding to the RGB wavelengths
3. normalize the reflectance values
4. apply the binary mask
5. plot the image

Using plain numpy: 16 lines of code

```
import meteors as mt
hsi = mt.HSI(image = image, binary_mask=binary_mask,
    wavelengths=wavelengths, orientation="CWH")
ax = mt.visualize.visualize_hsi(hsi,
    title="Image with mask")
```



Mask importance analysis - problem description

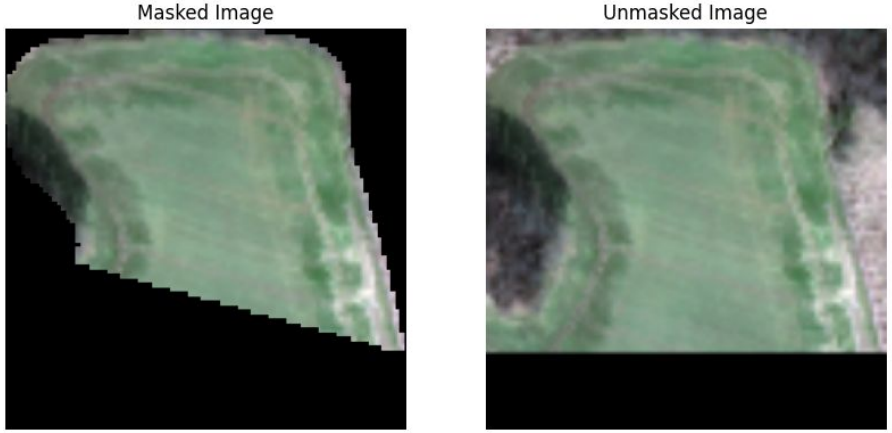
Hypothesis: *A model trained with a mask of a farmland area cannot accurately predict when the mask is not present.*

The HYPERVIEW dataset is masked by default - binary mask covers non-agricultural areas

Should the new data be masked as well?

How will model perform in real setting without the masks?

Sample image from the HYPERVIEW dataset

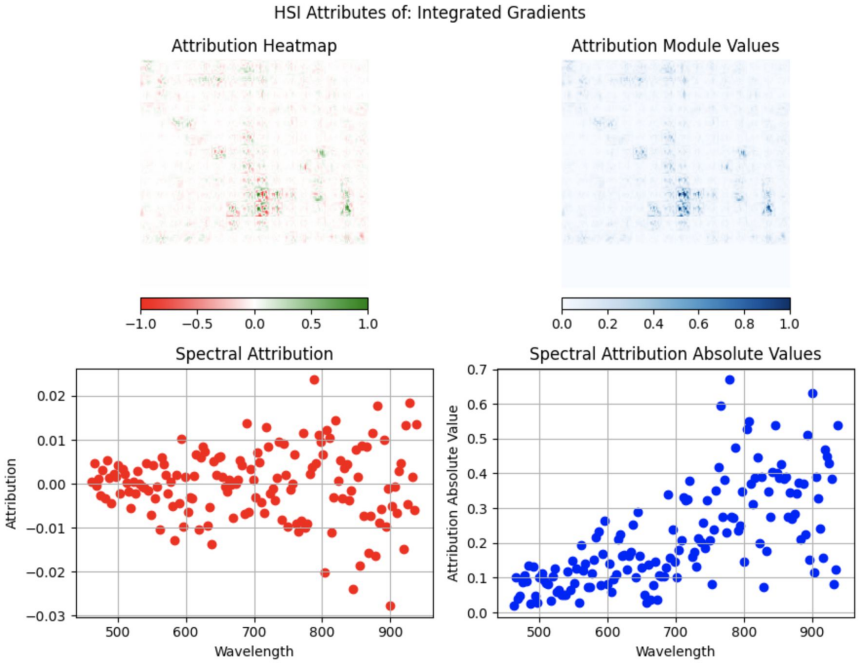


Mask importance analysis

For the mask importance analysis we used several gradient and perturbation methods

Interfaces of the methods are **simple** and **intuitive**, no broad knowledge of XAI is required to create the explanations

```python
# wrap the data - without the binary mask
not_masked_hsi = mt.HSI(image = image, wavelengths = wavelengths)
# wrap the model
explainable_model = mt.models.ExplainableModel(model,
    problem_type="regression")
ig = mt.attr.IntegratedGradients(explainable_model)
# create the explanation for target 0 - Phosphorus
explanation = ig.attribute(not_masked_hsi, target = 0)
# visualize the explanation
mt.visualize.visualize_attributes(explanation)
```



visualization of attribution - model focuses on all parts of the image, even on the parts that are not relevant

Bands importance for soil estimation

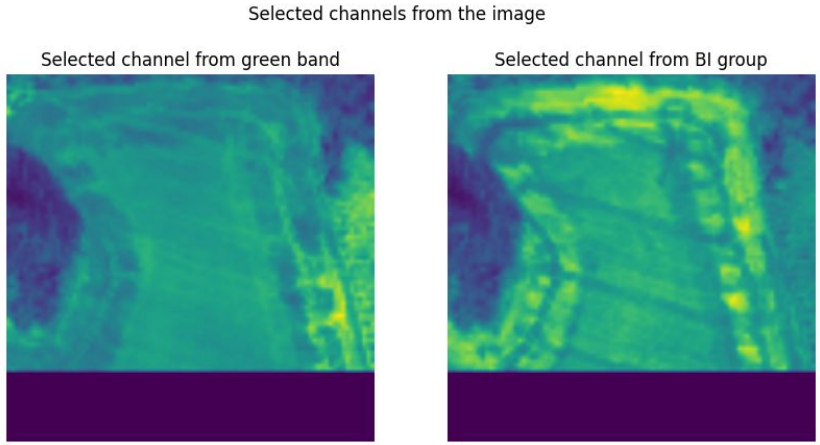
Spectral analysis is unique for multi- and hyperspectral imagery
It allows to inspect which of the bands are useful for predictions

Here using **LIME** explainer on the sample image we will verify which **bands** are useful for the model

Group of bands used in analysis:

BI - Bare Soil Index is set of wavelengths used to detect soil in the images;

G - Green visible light



visualization of what can be seen with different bands in the same image

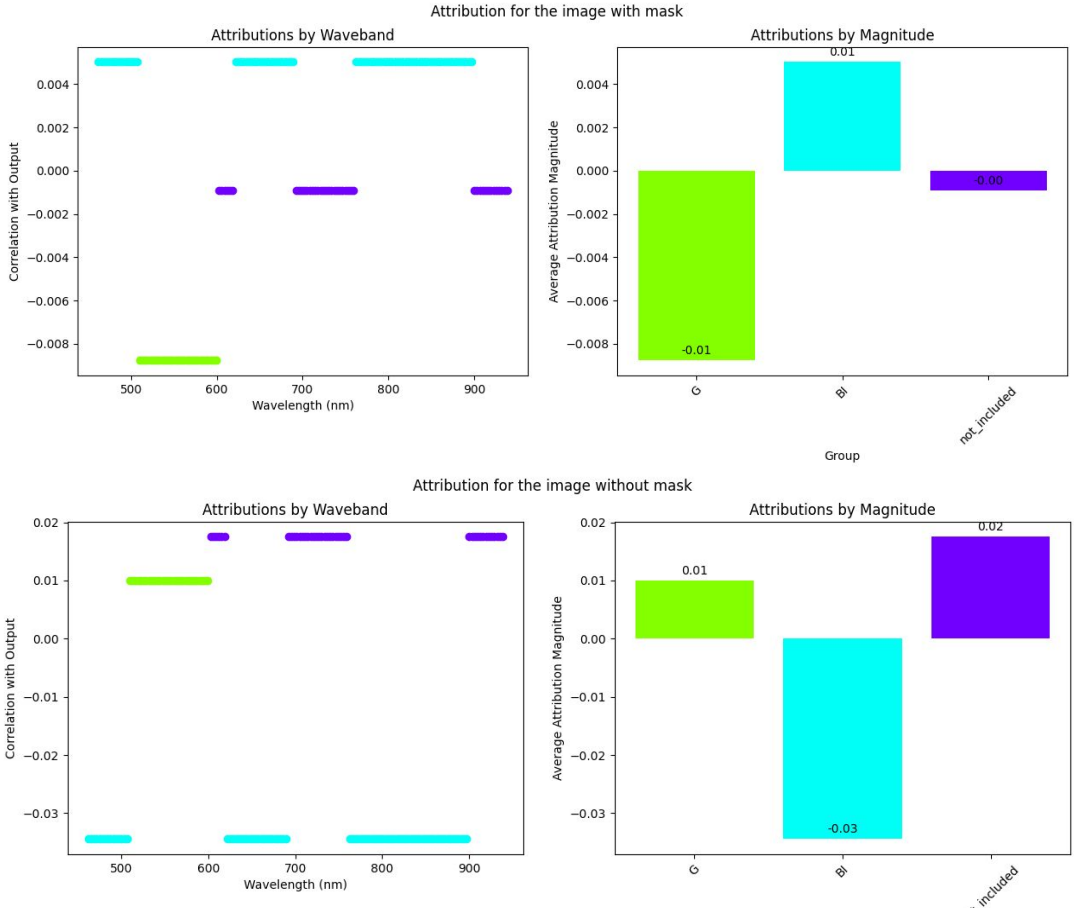
Bands importance for unmasked image

Charts on the left shows how **impact of the bands** varies for masked and unmasked images

For the image with mask the bands from *Bare Soil Index* are positively correlated with the output

For the unmasked image it is **the opposite**

This might again suggest that the model **incorrectly focuses on the unmasked region**



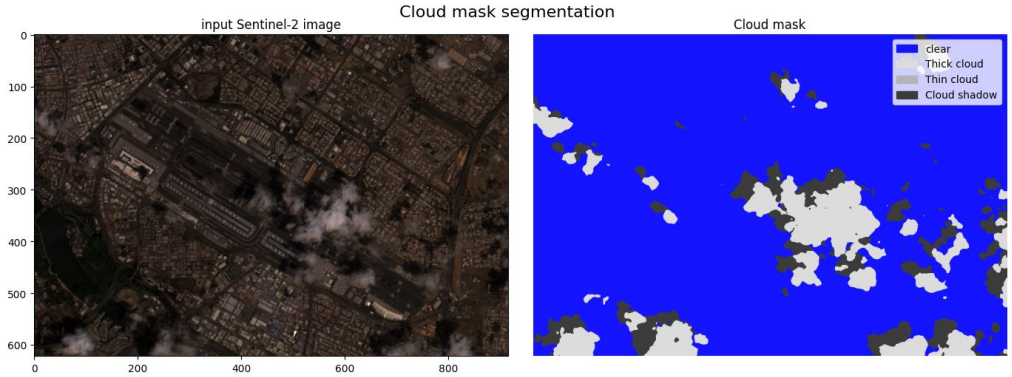
Cloud segmentation problem

Analysis of a top performing model **UNetMobV2_V2** trained on **CloudSEN12** dataset

Semantic segmentation of clouds on Sentinel-2 multispectral images

Segmentation is an **important field** in Remote Sensing - an initial step for further analysis

To analyze the Earth, we need to know if the ground is not covered



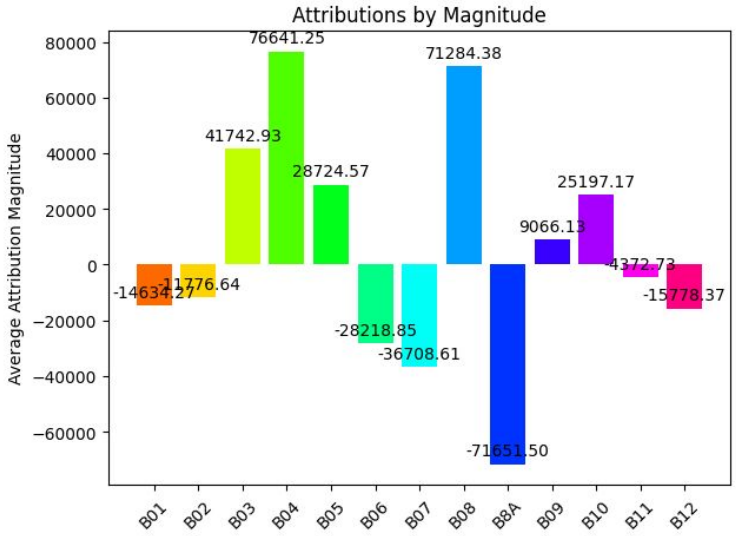
model performance on sample image

Bands importance for cloud segmentation

Understanding how bands influence segmenting *clear sky* class with **LIME** and our **novel approach** for explaining the **segmentation problem**

RGBI bands (B02, B03, B04, B08) were used to create *clear sky* class **ground-truth** for the training dataset

Our explanations hints that model utilizes the **RGBI** channels the most



attribution for each band - high attribution values suggest positive correlation with model's output, which means that more pixels will be classified as *clear sky*

Thank you for your attention!

Remember to leave your questions and rate the presentation in the section below!

Package

Feedback

DATA SCIENCE SUMMIT

ACADEMIC PARTNERS