

Vision-Language Models for Earth Observation

EarthPulse (Barcelona)

Description

Vision-language models are a type of model that can understand and generate natural language descriptions of visual content. These models have been shown to achieve state-of-the-art performance on a wide range of tasks, and have the **potential to unlock the full potential of Earth Observation data** by enabling new ways to interact with and understand the data.

Furthermore, **promptable computer vision**, pioneered by SAM¹, is a new paradigm in the field with the potential to unlock new applications in Earth Observation by enabling new ways to interact with and understand the data.

Research Topics

The following are a list of topics related to VLMs that the student can choose from to work during the project:

- Design, train and evaluate Vision-language models specific for EO data.
- Improve on the SAM architecture (promptable computer vision) with a model specific for EO data.

In both cases, the **evaluation** of the models on EO data is required.

Objectives

Due to the available time and resources, the research is focused on **small-scale demonstrations** that can be scaled in future works using existing datasets.

An important goal of the work is to **generate a scientific paper** describing the methodology and reporting results.

The student will implement their code in the open source PytorchEO library².

Required skills

The student should be comfortable working with:

- Python and Pytorch.
- Ubuntu (terminal) and Jupyter Notebooks.
- Satellite data and processing pipelines.

Willingness to work at our office at least 2 days a week and a startup mindset is a plus.

About us

EarthPulse is a startup based in Barcelona working in the intersection of Artificial Intelligence and Earth Observations. EarthPulse was funded in 2020 and, since then, it has been a leader in the

¹ <https://arxiv.org/abs/2304.02643>

² <https://github.com/earthpulse/pytorchEO>

European AI4EO community, with projects such as the AI4EO challenges platform³ and the EOTDL⁴ for the European Space Agency (ESA). The company is driven by the mission of making Earth Observation data more accessible and useful, leveraging AI as a key enabler, generating analytics that can be useful for monitoring, vulnerability, and impact assessment in different industrial sectors (such as utility managers, insurance companies, etc).

Our office is located at San Juan de la Salle 42, Barcelona. We are a team of 10 including software engineers and data scientists in the technical team from who you can learn from. We own a workstation with 2x3090 RTX NVIDIA GPUs that the student will use during the project. Additional cloud resources can also be leveraged if necessary. Our main interest with this collaboration is to generate quality scientific publications.

Learn more about EarthPulse at <https://earthpulse.ai/> or contact at juan@earthpulse.es

³ <https://platform.ai4eo.eu/>

⁴ <https://www.eotdl.com/>