Data Augmentation for Earth Observation

EarthPulse (Barcelona)

Description

Data augmentation techniques are used to artificially increase the size of a dataset by applying transformations to the original data. However, the transformations that are commonly used for natural images may not be suitable for EO data. The goal of this project is to comprehensively evaluate data augmentation techniques in different EO downstream tasks (image classification and segmentation) and design new augmentation techniques specific for EO data that may yield performance improvements.

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 library[1].

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 European AI4EO community, with projects such as the AI4EO challenges platform[2] and the EOTDL[3] 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

---

[1] https://github.com/earthpulse/pytorchEO
[2] https://platform.ai4eo.eu/
[3] https://www.eotdl.com/