M2 INTERNSHIP OFFER

Conditioned Multimodal Diffusion Model applied to Earth Observation

ANITI / ISAE-SUPAERO - Toulouse
Team: DISC/ADO

Duration: 5 to 6 month
Desired Start Date: February/March 2027
Localization: 10, avenue Marc Pélegrin, BP 54032 31055 Toulouse CEDEX 4 - France
Salary: $\sim$ 670€/month
Supervisors:

- Victor Sanchez, prenom.nom@isae-supaero.fr
- Thomas Oberlin, prenom.nom@isae-supaero.fr
- Mathieu Fauvel, prenom.nom@inrae.fr

Confidentiality: No
Possibility of Remote: No

1 Context

Remote Sensing (RS) data play a key role in applications such as land cover classification, field monitoring and climate change impact. These applications rely on multimodal measurements (e.g. optical, radar or thermal) from different sensors, which inherently have various spectral, spatial and temporal resolutions and are irregularly distributed in time and space. Tackling these inconsistencies is critical for many tasks. The internship will be within the ANITI project RELEO who aims to build a Foundation model for Earth Observation. This foundation model will be able to aggregate various measurable data (optical, radar, thermal) with external data (Meteorological data for instance). Additionally it will be able to generate land related physical variable [9] thanks to a hybrid architecture with physical models. Finally this model is going to be able to quantify uncertainties thanks to generative models.

Recent work demonstrates that diffusion models (DMs) [2] can effectively handle many RS tasks [6]. DMs creates data by reversing a noise diffusion process, and acquired popularity on image generation for their training stability [4] and sampling scalability [7].

Perceiver-IO [3] is a generic architecture that manage data from inconsistent settings while scaling linearly with the size of inputs and outputs. It is uniquely adapted for RS data application. Perceiver-IO can handle data inconsistency and resolution diversity as in RS (e.g., missing bands, varied resolutions) without architectural modifications. While [8] explore Perceiver-IO as backbone for molecule generation, its application to RS remains understudied.

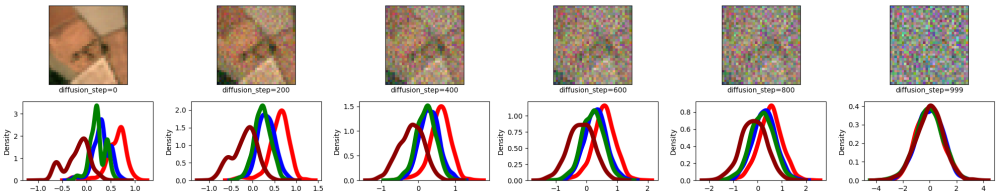


Figure 1: Example of forward Sampling of Sentinel 2 RGB image

2 Objectives

A diffusion model with a Perceiver-IO backbone has been proposed within the RELEO project. It amounts to learning a joint multimodal generative model tailored for EO data while enabling conditioning with respect to any modality. Yet, its use for several applications (e.g. climate change monitoring) remains limited because important external information (e.g. weather) is not included in the model. The main objective of the internship is to condition the generative model with respect to weather and related external data. To this end, the internship will rely on AGERA5 data, which is currently collected and integrated within the RELEO code base. Two open questions will be investigated: *how to condition the model (raw temporal variables or accumulated)?* and *what is the impact of each variable on the conditioning?*.

The developed model will be evaluated and compared against alternative methods from the literature, such as EcoMapper [1] or ClimSat [5].

3 Tasks

- Literature Review on the topic (multimodal generative models, PerceiverIO, diffusion models)
- Development and Experimentations in Python using Deep Learning Libraries (Pytorch, Lightning) and the RELEO code base
- Analysis of the results and comparisons with baselines

4 Desired Profile

- Last year Engineering School (or Master 2 equivalent) specialized in Deep Learning or Image Processing (ou école d'ingénieur équivalente)
- Required Skills : Python Programming, Deep Learning and Image Processing
- Appreciated Skills : Pytorch / Lightning coding, Knowledge in Remote Sensing data or application, Knowledge in Generative Models
- Quality : Autonomy, Aptitude for Teamwork
- Language : French and/or English

You don't need to meet every requirement to apply; we value diverse perspectives. We are committed to provide equal opportunities and encourage applications from candidates with diverse backgrounds and experiences

5 Environment and conditions

The intern will have an office at ISAE-SUPAERO reserach lab in the DISC team. The team member is composed of PhD student and researchers working on various application of deep learning.

- Weekly Meetings for internship supervision
- Participation to Project Meeting
- Collaborative code development for RELEO project

6 Perspectives

Depending on the outcome of the internship, a scientific publication will be considered. The internship could lead to a PhD thesis with the project researchers, but no founding is secured at this stage.

7 Applications

Applications (CV, cover letter, transcripts of records) should be sent to *Victor Sanchez :* `prenom.nom@isae-supaero.fr`. *Letters of recommendation are appreciated but remain optional.* **The application deadline is 01-12-2026.**

References

[1] Muhammed Goktepe, Amir hossein Shamseddin, Erencan Uysal, Javier Muinelo Monteagudo, Lukas Drees, Aysim Toker, Senthold Asseng, and Malte Von Bloh. Ecomapper: Generative modeling for climate-aware satellite imagery. In *Forty-second International Conference on Machine Learning*, 2025.

[2] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

[3] Andrew Jaegle, Sebastian Borgeaud, Jean-Baptiste Alayrac, Carl Doersch, Catalin Ionescu, David Ding, Skanda Koppula, Daniel Zoran, Andrew Brock, Evan Shelhamer, Olivier Hénaff, Matthew M. Botvinick, Andrew Zisserman, Oriol Vinyals, and João Carreira. Perceiver IO: A General Architecture for Structured Inputs & Outputs, March 2022. arXiv:2107.14795 [cs.LG].

[4] Tero Karras, Miika Aittala, Jaakko Lehtinen, Janne Hellsten, Timo Aila, and Samuli Laine. Analyzing and improving the training dynamics of diffusion models. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 24174–24184. IEEE, 2024.

[5] Johannes Leonhardt, Juergen Gall, and Ribana Roscher. Climsat–a diffusion autoencoder model for climate-conditional satellite image editing. *Science of Remote Sensing*, 11:100235, 2025.

[6] Yidan Liu, Jun Yue, Shaobo Xia, Pedram Ghamisi, Weiying Xie, and Leyuan Fang. Diffusion Models Meet Remote Sensing: Principles, Methods, and Perspectives. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–22, 2024. arXiv:2404.08926 [cs].

[7] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising Diffusion Implicit Models, October 2022. arXiv:2010.02502 [cs.LG].

[8] Yuyang Wang, Ahmed A. Elhag, Navdeep Jaitly, Joshua M. Susskind, and Miguel Angel Bautista. Swallowing the Bitter Pill: Simplified Scalable Conformer Generation, May 2024. arXiv:2311.17932 [physics].

[9] Yoël Zérah, Silvia Valero, and Jordi Inglada. Physics-constrained deep learning for biophysical parameter retrieval from sentinel-2 images: Inversion of the prosail model. *Remote Sensing of Environment*, 312:114309, 2024.