

Offre de stage : Apprentissage profond et modélisation physique pour l'extraction d'informations à partir de séries temporelles d'images satellite multi-modales

Contexte

Le projet ANR MAESTRIA vise à proposer des outils et méthodes de nouvelle génération pour la **production automatique de cartes d'occupation des sols** à partir d'imagerie satellite multi-temporelle à haute résolution.

Un des objectifs du projet est de produire des **cubes de données multi-modaux**, c'est à dire des variables échantillonnées de façon régulière dans l'espace et dans le temps intégrant des informations acquises par des capteurs de modalités différentes (optique, radar, etc.). Ces variables sont générées par des approches d'apprentissage automatique auto-supervisées, comme les auto-encodeurs variationnels [Doe16] [WDGH16].

Travail à réaliser

Afin que ces variables soient pertinentes et cohérentes d'un point de vue physique, on souhaite les **contraindre par des modèles bio/géophysiques des processus observés**. En effet, un nombre important de paramètres de cette nature peuvent être extraits des images de télédétection, comme par exemple l'humidité des sols [BMS⁺12, GZEB17, EHBZB17], l'indice foliaire, ou la teneur en eau et en chlorophylle [FFA⁺08], le rendement [BABM⁺16] ou la biomasse [CAMS17].

L'objectif du stage est d'intégrer certains de ces modèles bio/géophysiques dans un **système d'apprentissage profond** pour pouvoir les exploiter de façon conjointe avec l'architecture d'auto-encodeurs actuellement utilisée.

Les tâches à réaliser seront les suivantes :

1. Étude bibliographique des modèles.
2. Mise en oeuvre du modèle direct (la plupart des modèles envisagés sont disponibles sous forme de logiciel libre).
3. Apprentissage du modèle inverse à partir de données simulées.
4. Intégration dans l'architecture d'auto-encodeur existante.

Profil du candidat

- Formation en physique et/ou télédétection et/ou mathématiques appliquées/apprentissage/data science
- Connaissances en programmation, Python Data Science Stack de préférence

Conditions de recrutement

- Durée de 6 mois avec un début entre février et avril 2020
- Indemnité de 577 € par mois
- Candidatures : CV et lettre de motivation à jordi.inglada@cesbio.eu avant le 15/11/2019

Internship offer: Deep learning and physical models for information extraction from multi-modal satellite image time series

Context

The ANR MAESTRIA project aims at proposing new generation tools and methods for the **automatic production of land cover maps** from high resolution satellite image time series.

One of the project goals is the generation of **multi-modal data cubes**, that is, variables regularly sampled in space and time which integrate the information acquired by sensors of different modalities (optical, radar, etc.). These variables are generated using auto-supervised machine learning approaches, such as variational autoencoders [Doe16] [WDGH16].

Tasks

In order for these variables to be pertinent and coherent from a physical point of view, we want to **constrain them by bio/geophysical models of the observed processes**. Indeed, several of this kind of parameters can be extracted from remote sensing imagery, as for instance soil moisture [BMS⁺12, GZEB17, EHBZB17], leaf area index, water and chlorophyll content [FFA⁺08], yield [BABM⁺16] or biomass [CAMS17].

The goal of the internship is integrating some of these bio/geophysical models into a **deep learning system** to jointly use them with the autoencoder architecture currently used.

The tasks to be done are:

1. Study of the literature cited above
2. Implementation of the forward models (most of the above-mentioned models are available as free and open source software)
3. Inverse model training using data simulated by the forward models
4. Integration into the existing autoencoder architecture

Candidate profile

- Background on physics and/or remote sensing and/or applied mathematics, machine learning or data science
- Programming, knowledge of the Python Data Science stack is a plus

Hiring conditions

- Length: 6 months beginning between February and April 2020
- Gratification of 577 €/month
- Applications: CV and motivation letter to jordi.inglada@cesbio.eu before 15 November 2019

References

- [BABM⁺16] Marjorie Battude, Ahmad Al Bitar, David Morin, Jérôme Cros, Mireille Huc, Claire Marais Sicre, Valérie Le Dantec, and Valérie Demarez. Estimating maize biomass and yield over large areas using high spatial and temporal resolution sentinel-2 like remote sensing data. *Remote Sensing of Environment*, 184:668–681, Oct 2016.
- [BMS⁺12] Anna Balenzano, Francesco Mattia, Giuseppe Satalino, Valentijn Pauwels, and Paul Snoeij. Smosar algorithm for soil moisture retrieval using sentinel-1 data. *2012 IEEE International Geoscience and Remote Sensing Symposium*, Jul 2012.
- [CAMS17] Jose Alan A. Castillo, Armando A. Apan, Tek N. Maraseni, and Severino G. Salmo. Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the philippines using sentinel imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134:70–85, Dec 2017.
- [Doe16] Carl Doersch. Tutorial on variational autoencoders. *CoRR*, 2016.
- [EHBZB17] Mohammad El Hajj, Nicolas Baghdadi, Mehrez Zribi, and Hassan Bazzi. Synergic use of sentinel-1 and sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas. *Remote Sensing*, 9(12):1292, Dec 2017.
- [FFA⁺08] Jean-Baptiste Feret, Christophe François, Gregory P. Asner, Anatoly A. Gitelson, Roberta E. Martin, Luc P.R. Bidel, Susan L. Ustin, Gueric le Maire, and Stéphane Jacquemoud. Prospect-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment*, 112(6):3030–3043, 2008.
- [GZEB17] Qi Gao, Mehrez Zribi, Maria Escorihuela, and Nicolas Baghdadi. Synergetic use of sentinel-1 and sentinel-2 data for soil moisture mapping at 100 m resolution. *Sensors*, 17(9):1966, Aug 2017.
- [WDGH16] Jacob Walker, Carl Doersch, Abhinav Gupta, and Martial Hebert. An uncertain future: Forecasting from static images using variational autoencoders. *Lecture Notes in Computer Science*, pages 835–851, 2016.