PhD position: On the use of Deep learning for ocean SAR image semantic segmentation.

Context

There is an overwhelming amount of data from Copernicus Sentinel-1 satellites. For instance, thousands of images are produced every day representing a daily average of 3,45 TB SAR data published. A significant amount covers ocean surface, used for a wide range of applications involving public and private stakeholders. Notwithstanding the current great impact of Sentinel-1 images for the oceanographic community, we believe the imaging capabilities of these C-band SAR data acquired over ocean surface are not fully exploited.

Beyond wind field measurements, sea ice monitoring or oil spill detection, a set of metocean features are well observed by S-1 sensors. To name a few, atmospheric fronts, oceanic fronts, rain cells, micro convective cells, internal waves, gravity waves, biologic slicks, upwelling or wind streaks are phenomena of potential interest for many end-users while being generally discarded in the SAR images. For academic science and industry, new research perspectives and new potential applications/services could be triggered with the proposed advanced monitoring of these metocean mechanisms. In addition, flagging of these features could also benefit to current products/services providing better data quality. With recent progress in computer vision thanks to Deep Learning approaches in conjunction with the rise of large database and higher computing power, these flagging activities are now possible.

Assignment

For the very first time, our group (incl. IFREMER) has been testing DL framework in the aim to provide a semantic segmentation of Sentinel-1 scenes acquired over open ocean surface, with the final objective to tag all these metocean features which have been neglected for years.

More specifically, a database of 37k Wave Mode (WM) S-1 imagettes has been annotated with 10 different classes (Pure Ocean Swell, wind streaks, micro convective cells, rain cells, biological slicks, sea ice, icebergs/bright targets, low wind area, atmospheric fronts, oceanic fronts). First models were implemented and then validated. An assessment using an independent database with 10k WM imagettes has been carried out, revealing good classification accuracy of more than 85% for most classes, with acceptable false alarms rate (below 10-20%). Based on this classification work, a methodology has been proposed to provide a semantic segmentation of any ocean SAR image (including S-1 Wide swath images).

At this stage, our work has been focusing on:

- Preparing the SAR image to fed DL framework
- Ingesting solely SAR image (one layer corresponding to radar intensity), with no other input (metocean a priori, other derived SAR parameters...), and no contextual information (geographical specificities)

- Classification framework with semantic segmentation using a sliding local classification
- Limited active learning process
- Known on-the-shelf framework using Keras
- No multi-labelling approaches

The PhD applicant will help improving the aforementioned strategies (e.g. new data preparation process based on SLC SAR data, tackle co-occurring phenomena in our first mono-label training database, contextual information and multi-dimension inputs, direct semantic segmentation framework with active learning, custom framework with enhanced capabilities directly using Tensorflow, multi-scale and –labelling approaches...). The general idea of the PhD work will be to transpose and adapt the latest achievement of DL to the SAR community and propose real adaptation to SAR and metocean specificities. This may also include an in-depth revisit of SAR processing chains from a deep learning perspective.

The applicant will have access to the full S-1 archive via Copernicus DIAS (or more "limited" local archive @CLS or supercomputer Datarmor) and to high performance GPU (@CLS or IMT-A for test/implementation purposes, and cloud-based for larger training/inferences).

The PhD work will focus on the methodological approaches and will be carried out in coordination with application / thematic experts (at CLS, in collaboration with IFREMER and other actors such as meteorological agencies).

Candidate profile

- Skills:
    - Master of Science (or equivalent) in Applied Mathematics, Computer Science or Machine (Deep) Learning
    - Good programming skills (Python) with proven experience (i.e. github project)
    - Ideally experience with cloud computing (Dockers, Kubernetes...)
- Know-how: fluency in written communication (writing technical notes and scientific articles - in English in particular) and oral communication (presentation at contractual meetings or scientific conferences), work organization, scientific rigor.
- Soft skills: Dynamism, enthusiasm, good interpersonal skills, autonomy, capacity for innovation, taste for teamwork.

Application procedure

Candidates should send an e-mail to nlongepe@groupcls.com, ronan.fablet@imt-atlantique.fr and pierre.tandeo@imt-atlantique.fr containing:

1. Full CV

2. Letter of interest

3. Contact information for 2 references

Interview (in situ or webconf) will be planned in September/October.

PhD will be funded for 3 years from January 2020 (start earlier possible) and will be hosted at CLS Brest with supervision by IMT Atlantique (R. Fablet and P. Tandeo). The PhD scholarship will include teaching opportunities (Machine Learning, Data Science) in MSc./Eng. Degree at IMT Atlantique as well as visiting scholarships for short visits abroad in the framework of international partnerships.