SIOS hands on session

AI Sea ice surface classification techniques

Michel Tsamados + Weibin Chen

University College London

September 4, 2022

Table of Contents

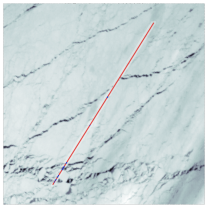
Dataset and general approach

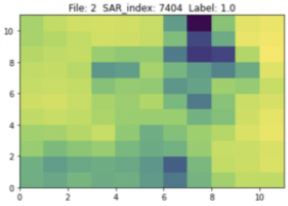
Dataset

- OLCI imagery:carried by Sentinel-3, provides data continuity for Envisat's Medium Resolution Imaging Spectrometer (MERIS). It features a ground spatial resolution of 300 m/pixel, a 1270 km wide swath coverage, and 21 spectral bands in the visible and near-infrared range 400-1020 nm. The directory of one OLCI image contains 21 radiances files, represents each spectral band.
- SAR: The Synthetic Aperture Radar Altimeter instrument of the Sentinel-3A and Sentinel-3B satellites provided the SAR altimetry level 1B (non-time critical) data used in this project.

Pre-processing of the data

- In the data pre-processing part, the training images come from finding the co-located points between SAR and OLCI points. All the SAR points are labelled using WMA (Waveform mixture algorithm) from Lee et al(2018).





Random forest

Random forest

- It is a traditional supervised machine learning algorithm.
- Essentially, it is a group of decision tree
- It uses bagging (allows each tree to take a random sample from the dataset with replacement, creating various trees as a consequence.)
- Voting algorithm

Random forest

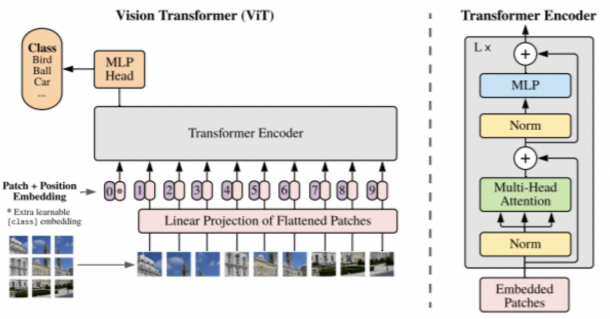
- It is available in scikit learning library.[1]
- It requires less computational resources than deep learning model
- More interpretable than deep learning model

Vision transformer

Vision transformer

- Vision transformer is adapted from 'Transformer', which is originally used in the "Natural Language Processing (NLP)" area, in the area of computer vision. [2]
- As sequential inputs for image recognition, Transformer separates the images into patches with positional embedding (which means the order of the patches matter. The vectors representing each patch are then subjected to a linear transformation. [2]
- It will send a class vector to the network's transform encoder. Vision transformer encoders have similar structure to traditional neural networks, with the exception being the addition of an attention component. [2]
- It examines the relationships between input token pairs.[2]

Vision transformer



- More details are in "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale"

Vision transformer

- Vision transformer is available in Keras (Tesnsorflow)[1]
- It requires less computational resources than other image classification algorithm like "Convolutional Neural Network (CNN)"
- Running on Google Colab (a useful cloud server) or Amazon AWS

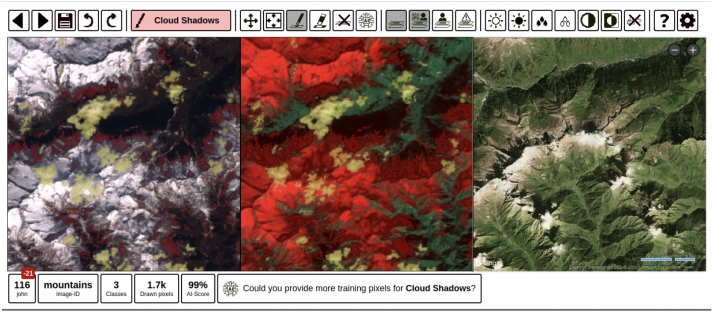
IRIS

Intelligently Reinforced Image Segmentation

- https://github.com/ESA-PhiLab/iris
- Support by AI (gradient boosted decision tree) when doing image segmentation
- Multiple and configurable views for multispectral imagery
- Simple setup with pip and one configuration file
- Platform independent app (runs on Linux, Windows and Mac OS)
- Multi-user support: work in a team on your dataset and merge the results

Intelligently Reinforced Image Segmentation

- Here is the interface of IRIS



Intelligently Reinforced Image Segmentation

Now let's have a look at a demo of IRIS to show how it works.

References I

- K. Salama, "Image classification with vision transformer," 2021.
- A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

Acknowledgment

This work was supported by various grants (ESA, NERC, UCL, etc) as well as by a generous invitation from SIOS to present at the Longyearbyen Sep 2022 workshop.