Deep Learning Projects

Introduction

The goal of these projects is to design a deep learning methodology to solve a task on a provided dataset. This involves:

- Implementing methods to load and preprocess the data
- Implementing and training several relevant models and/or one model with several parameter sets
- Computing relevant metrics and analyzing the results
- Visualizing results

Choose one of the projects proposed in the next sections.

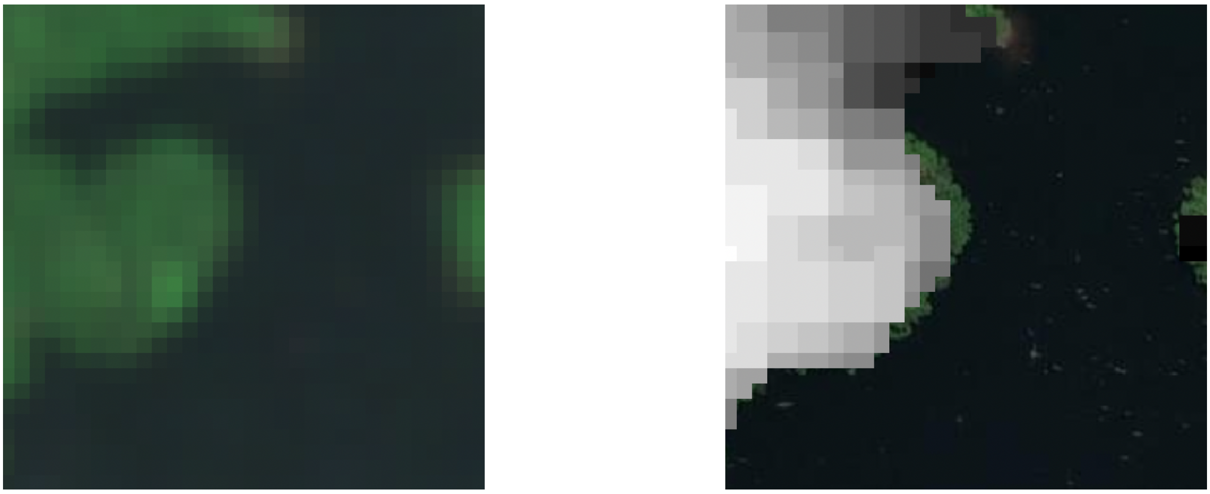
Then you can download the dataset you need for your project here.

1 Forest canopy height regression

Task. Vegetation height is a fundamental variable necessary for estimating carbon fluxes, understanding biodiversity, ecosystems services and many other applications. In this project, you will predict per-pixel vegetation height from multispectral Sentinel-2 images.

Data. A total of 9852 images (and the corresponding masks) will be used in this project. You can see an example of both in Figure 1. Each pair of image and mask is named with a unique id. In the data folder, you will find a CSV file where each id is assigned to either the train, validation or the test set.

- **Images.** Patches of Sentinel-2 multi-spectral images, all bands upsampled to a common 10m resolution. Each patch is of size 32×32.
- **Labels.** The labels are in the form of segmentation masks. There is one segmentation mask for each image with the identical size 32×32 pixels. The pixel values in the mask correspond to the canopy height in meters as modeled by [1]. Values of 255 correspond to 'nodata' values.



(a) Sentinel-2 image

(b) Canopy height displayed over a basemap

Figure 1: Example of the input image and the corresponding canopy height that is used as annotation

Challenges.

- You need to perform per-pixel regression (predict a continuous output for every pixel); your model architecture and loss function need to reflect that.
- Sentinel-2 images have 12 input bands; most standard deep learning models have been developed with natural 3-band images in mind; you may need to adapt an existing architecture to serve your purposes

References [1] Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R. and Wegner, J.D., 2022. Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. Remote sensing of environment, 268, p.112760.

2 Landcover mapping in Greenland with multi-temporal imagery

Task. Predict per-pixel landcover class from multi-temporal Landsat-8 imagery.

Data.

- **Images.** Patches of Landsat-8 multi-spectral images at 30m resolution. Each patch is of size 128×128. In the train dataset, the 1500 patches are from 2014, 2015 and 2016. In the test dataset, the 500 patches are from 2023.
- **Labels.** The labels are provided as segmentation masks of the same size as the corresponding images. They contain 6 different classes described in Figure 3.

Challenges.

- The semantic classes are imbalanced so choices will be made to achieve a reasonable trade-off.
- The train images are provided from 2014 to 2016 while the test images are from 2023 so the domain shift between train and test is both spatial and temporal. You are encouraged to test different ways of using the different dates during training.

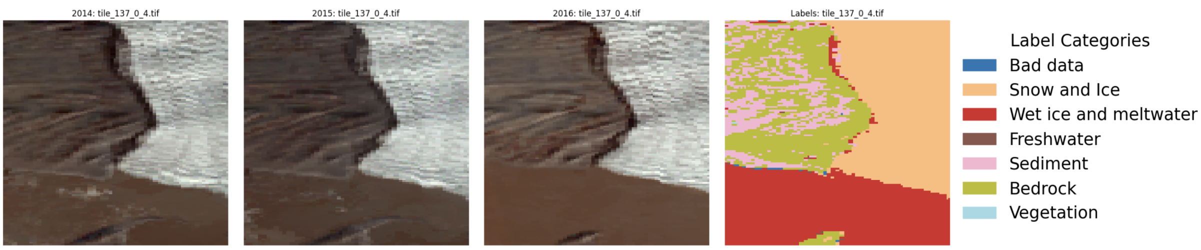


Figure 2: Example of a Landsat-8 RGB patch (2014, 2015, 2016) and associated labels

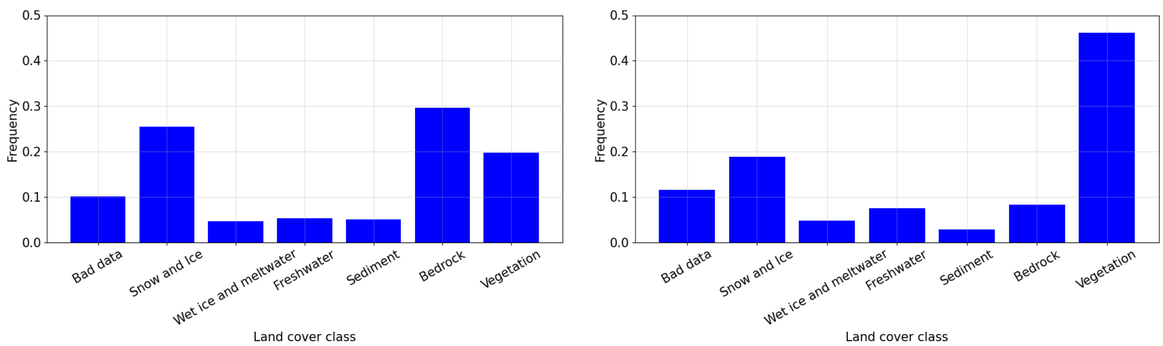


Figure 3: Label distribution in the train (left) and test (right) datasets

References [1] Grimes, Michael, et al. "Land cover changes across Greenland dominated by a doubling of vegetation in three decades." Scientific Reports 14.1 (2024): 3120.

3 Large Rocks Detection Dataset

Task The Federal Office for Topography swisstopo proceeds to the manual annotations of all large rocks in Switzerland to produce topographic maps. They are curious to observe what could be done with recent automatic methods. Thus this dataset is based on their annotations and your task will be to detect large rocks (over 5x5m) in Switzerland based on high-resolution RGB images and the digital surface model (DSM). You could compare the usage of one source of data (e.g., only DSM or only RGB) or try to combine them. Another possibility is to explore the differences between standard machine learning approaches (detecting local maximums, rugosity indices, etc.) and recent object detection models.

Data The study area is spread across Valais, Ticino and Graubunden. The tiles are geographically split into training and testing. The dataset includes :

- **Aerial images** at a 50cm resolution with RGB bands (swissIMAGE)
- **Digital surface model (DSM)** at 50cm resolution based on LiDAR data (swissSURFACE3D)
- **Hillshade raster** tiles derived from the DSM data, generated with QGIS with the hillshade function (Azimut 0, Vertical angle 0)
- Comprehensive **points annotations** of 2'625 large rocks from swisstopo annotators.

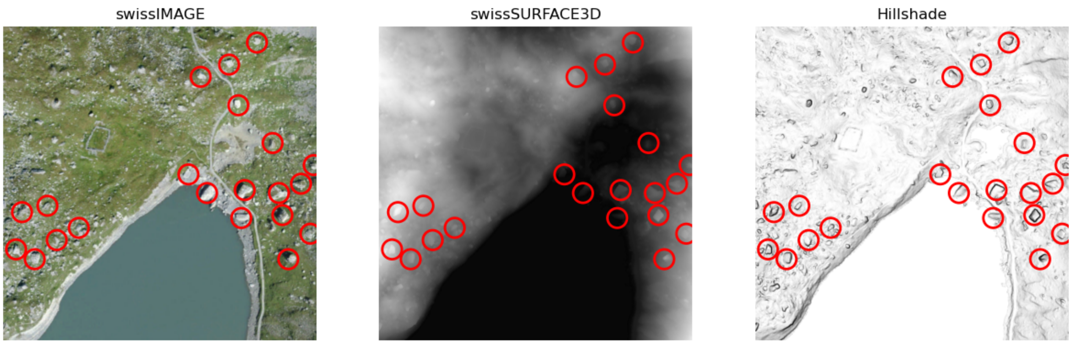


Figure 4: Aerial image from swissIMAGE, digital surface model from swissSURFACE3D and the corresponding hillshade raster with large rocks annotations from swisstopo annotators.

Challenges

- Building an object detection pipeline based on swisstopo annotations for aerial images and/or the DSM data will be your main challenge. You will also have to decide how to include the punctual rocks annotations (various shapes and sizes of bounding boxes, etc.).
- The precision-recall balance of the rock detection model will be an important point to observe and discuss. Too many false positive predictions lead to lengthy correction for the annotators while missing too many targets makes the predictions useless.
- To avoid over-fitting, you could investigate strategies such as data augmentation or starting for a pre-trained model.

Suggestions and references

- Make sure to have a look at the explanatory notebook ***Usefull_tips.ipynb*** provided with the dataset.
- For the deep learning approach, we strongly suggest you use the *YOLO-v8* object detection model with the *ultralytics* python library : `https://docs.ultralytics.com/`
- More information about swisstopo products such as swissSURFACE3d[1] or swissIMAGE[2] are available on their website.
- Link to download the dataset : `https://enacshare.epfl.ch/bY2wS5TcA4CefGks7NtXg`

[1] `https://www.swisstopo.admin.ch/de/hoehenmodell-swisssurface3d`
[2] `https://www.swisstopo.admin.ch/en/orthoimage-swissimage-10`

4 Change detection in Valais

Task Change detection in urban areas is crucial for monitoring the effects of urbanization, land use transformation, and environmental impacts. In Switzerland, aerial images of the entire territory are regularly captured, typically every few years, providing high-resolution data that allows for the continuous monitoring of changes both in natural and urban environments. In this project, you will focus on urban changes in the canton of Valais. You will perform per-pixel change detection by comparing pairs of aerial images from two different time periods (2017 and 2023). The task is to classify each pixel as either ”changed” or ”unchanged”.

Data The dataset is composed of 4851 pairs of RGB aerial images from Swisstopo, of size 256×256 and at 0.5m resolution. Each pair consists of coregistered images from two years (2017 and 2023), along with corresponding binary change detection masks (1 meaning ”changed” and 0 ”unchanged”). The images are paired by their unique ID and each pair is assigned to the train, validation, or test set, separated in three folders named `train`, `val` and `test`, respectively. The dataset is split geographically, as shown in Figure 6:

- **Train**: 3430 pairs around the city of Sion.
- **Test**: 980 pairs around the city of Martigny.
- **Validation**: 441 pairs around the city of Sierre.

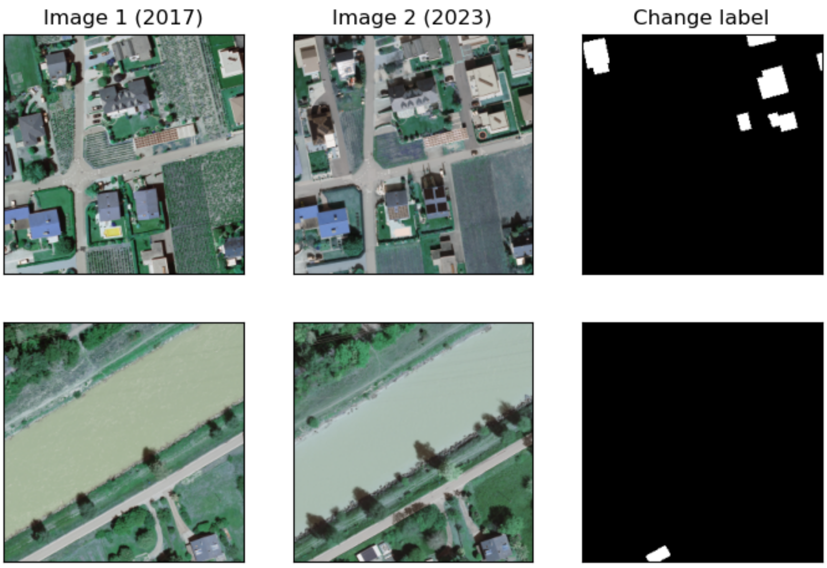


Figure 5: Examples of coregistered images and associated labels for change detection.

Challenges

- You will need to perform a per-pixel change detection to classify whether or not a pixel has changed. This is a pixel-wise classification task similar to segmentation, with the difference that there are two input images. The choice of architecture, loss function and evaluation metrics will need to take this specificity into account.
- The dataset is highly imbalanced, with a much smaller number of pixels labeled as ”changed” compared to ”unchanged”. This imbalance can affect model training and require the use of specific techniques to improve performance on the minority class.
- Change labels have been obtained automatically from the swissTLM3D landscape model of Switzerland. Temporal mismatches between the updates of this model and the images may result from time to time in incorrect labels. This can lead to undetected changes or false detections of change, introducing noise into the dataset. The results must therefore be critically evaluated.

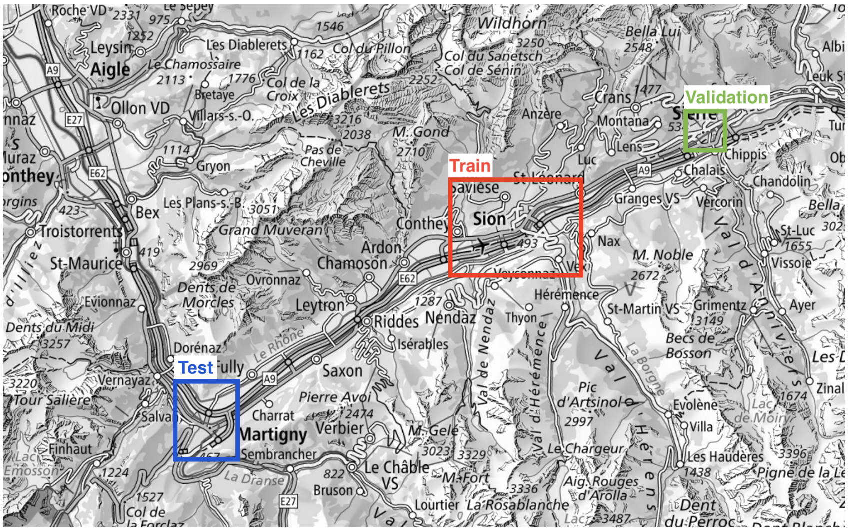


Figure 6: Geographic distribution of train, validation, and test sets in Valais, Switzerland.

5 Marine Debris Detection

Task Ocean plastic pollution is a threat for marine and coastal ecosystems and a major societal and environmental concern. Being able to monitor it with a timely detection is of big importance to coordinate cleaning efforts. Floating marine debris tends to create filaments referred to as *windrows*, generally used as proxies for plastic pollution. They can reach more than 50m of width and 500m of length, making them visible from optical satellite sensors such as Sentinel-2, having a resolution of 10 to 20m. The goal of the project is to develop an classifier that identifies the presence of floating marine debris in Sentinel-2 images.

Data Students will use a dataset composed of 12-bands Sentinel-2 image patches of size 32×32 pixels. Each image is associated with a binary label: "0" for negative labels (absence of floating debris), "1" for positive labels (presence of floating debris). The dataset is composed of 53535 training patches, 7436 validation patches and 13386 test patches. The samples for the validation and test set have been sampled to contain an equal number of positive and negative examples.

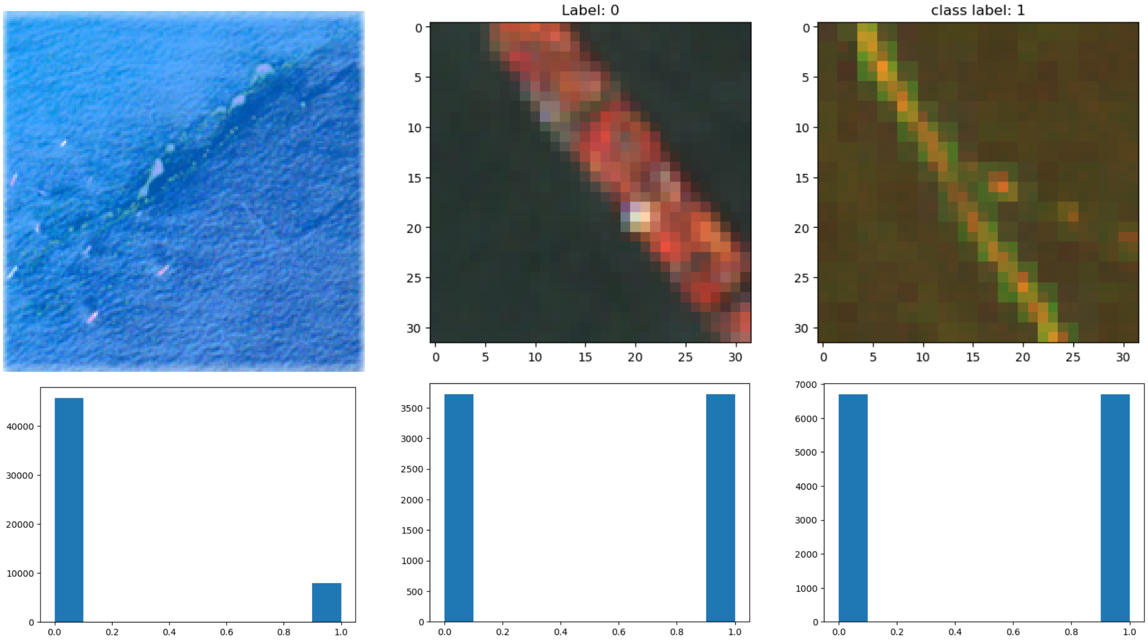


Figure 7: **Top**: example of floating material seen from Sentinel-2 on a 128×128 image patch, followed by an example of a negative (presence of a boat) and a positive (presence of floating debris) sample of 32×32 pixels. **Bottom**: label distribution within the training set (left), the validation set (center) and the test set (right)

Challenges

- The task is an image classification task, where the challenge is to associate a label to each image. You will have to study the performance of the classifier in three different scenarios: with the use of all the 12 spectral bands, with the use of only RGB+NIR bands, with the use of only the visible spectrum, to study the sensitivity of the model to the different spectral bands.
- The dataset is strongly imbalanced with only few positive examples in the training set: this needs to be taken into account for an effective training of the model.
- The dataset is obtained from a segmentation dataset with per-pixel annotations, such that if at least one positive pixel is present in the image, the image class is set to "1". Thus, for some images, it can be challenging to correctly classify them as only few pixels contain floating material.