Entry task

High-level description

Train a segmentation model that will output a binary mask identifying each pixel to the cloud (1)/no-cloud(0) class.

Assignment

Train and evaluate the segmentation model on the given dataset. The model should take a sentinel-2 image with 4 channels (RGB+NIR) of a given shape and output a binary mask of the same spatial shape.

Dataset

The input dataset is a publicly available dataset of Sentinel-2 image patches with corresponding masks available at zenodo. The dataset consists of 513 1022-by-1022 pixel sub-scenes, annotated to three classes CLEAR, CLOUD, CLOUD_SHADOW **but for this assignment classify just into two classes and consider CLOUD_SHADOW as CLEAR**.

We annotated some Setinel-2 acquisitions which are not included in the given dataset and will be used to test the resulting model.

Be aware that sentinel data consists of 13 bands (channels) and for this task, you should **use 4 channels as input RGB + NIR (band 8)**.

Requirements

- Model output shall be in ONNX format and **have size less than <250MB**.
- Provide information on how to run inference using ONNX runtime (ideally as a Colab jupyter notebook with all dependencies installed in the first cell).
- Model input shall be in shape NCHW, where N is batch_size, C is a number of input channels == 4, H is height == 224 and W is width == 224.
- Model output shall be in shape HW, where H is height == 224 and W is width == 224 with values {0,1}.
- Write a short report with training protocol (how was the dataset split, what were the hyperparameters, what is the validation strategy, metrics etc.)
- Model shall have input in the same resolution so tilling is necessary, do not use just resize.

Deliverables

- Trained ONNX model with documentation on how to run inference on CPU (ideally Colab notebook)
- Short report with experiment methodology

Deadline

The deadline for the task is 3 weeks from the assignment. Please be aware that finishing the assignment by the deadline is very important with all the deliverables. Focus on finishing all the requirements rather than tuning the model for the best accuracy. We understand that in three weeks you will not achieve state-of-the-art results (if yes then great!), but we are interested in the way you plan, execute and deliver your work.

Contact

If there is something unclear in the assignment, implement it in the way you understand it and we can discuss it further during the review. If you are not sure feel free to send us an email with your question at info@zaitra.io and we will reach back to you with clarification.

Copyright