**White paper for track 3 model**

1. Introduction

The track 3 classification model is an application which can download focused area Sentinel-2 images and classify it pixel by pixel automatically. It can find out all available data for the focus area and given range of date. After preprocessing for all data, it can classify each pixel in 6 categories: water, cloud, forest, low vegetation, bare land, and urban area.

This model can be used in flooding monitor. For example, when we know somewhere happened flooding in one day, this model can detect the pre- and post-event images to compare their water body area which can help us to find out the most flooded area.

This model is user friendly. Here are two ways to use it. One is that it can check the most severe flooding area, starting date, and ending date and save it as a csv file by itself. Then this model can search all downloadable data for classifying based on the csv file. The other way is that users can draw out their focused area and choose searching data range as model’s input. If there is no eligible data, model will pop up an alert to let user know.

1. Features
   1. Download and preprocessing data

We can download data with Sentinelsat API which can search for eligible data and download them. The conditions that can be selected here include data type, footfrint coverage, cloud coverage, time range and satellite orbit number, etc.

Firstly we need to sign up an account in Scihub and load a geojson file as footprint. Then we can search the eligible data and download them as zip files. Then transform them as GeoTIFF files.

After downloading data which overlap with the footprint, we need to combine them as one GeoTiff file. After that what we need to do is cutting the selected region from it. We can create a mask by given region and cut the overlapping area in the union file. But different file might use different coordinate reference system. Before creating the mask, we need to reproject the shape file with the coordinate reference system of Sentinel-2 data (EPSG:32615).

* 1. Classify image

Our cut GeoTiff file includes B2, B3, B4, and B8 band. At start, we need to generate NDVI (from (B8-B4/B8+B4)), and NDWI (from (B8-B3/B8+B3)) and add them as input. After that, this array with 6 channels can be inputted in the trained model which would output an 1 channel array with number 1 to 6 to show each pixel’s category. This output array would be save as a GeoTIFF file as input finally.

* 1. Visualization

We can also overlap the saved output on the Open Street Map or other maps and save it as a html file.

1. Disposition

This model can be deposited on a supercomputer remotely as a docker model. It not only can be called across systems, but also can automatically run 7\*24 hours.

1. Residual problem

Here are some problems need to be discussed. Due to the size of output is huge, how to storage them is a question for me. Secondly, do we need to save the original GeoTIFF images? More details about cloud storage might need to clear.