I used Unet with an EfficientNet backbone (ImageNet weights). For the reason of sparsity of target masks I have decided to take an advantage of [1] propositions and build two separate models: one to identify the very presence of erosion in the given image, another to do the segmentation itself (provided the first model classified the image as positive).

Without the combination of those two models, Unet could hardly reach 0.05 of IoU score.

I looked through the data samples and realized that I can hardly distinguish true erosion masked by the dataset creators and some other forms of soil imperfections.

Hence, I was not surprised when the EfNet classifier could not reach higher than 0.45 f1\_score.

Meanwhile, the segmentation model reached 0.45 of validation IoU score when trained on positive examples only.

As a result, upon combining classification and segmentation models 0.75 IoU score was received.

The bad thing is that predicting always uneroded soil gives 0.8+ IoU, which makes either the metrics or the modelling approach invalid.

To be honest, I have been agreeing more to the trained model than to the masks provided when looking at the data myself. That was the case since there are seemingly plenty of unmasked areas that have some form of soil erosion. It is hard to train a good model on a confusing data.

**Possible ways to improve the results:**

More high-quality data and targets.

As showed in [2], previously pretrained networks specifically on satellite images of high resolution (QuickBird in their case) appeared to perform better on Sentinel-2 datasets.

In study [3] some state of art were achieved in outlining soil degradation sites. They used UNet and some high resolution dataset, which was used to train the model from scratch for 300 epochs.

1. Xiaoling Xia et al 2019 J. Phys.: Conf. Ser. 1213 022003
2. ISPRS Journal of Photogrammetry and Remote Sensing 150 (2019) 59–69
3. AI for Earth Sciences Workshop at NeurIPS 2020.