

For the soil erosion detection task I tried to implement the same principle as described in paper "U-Net with VGG11 encoder pre-trained on ImageNet for image segmentation" (<https://arxiv.org/pdf/1801.05746.pdf>). I've used VGG16 encoder with the weights pretrained on ImageNet. I've trained only the decoder part of the network keeping the encoder weights frozen. This approach also can be extended by unfreezing encoder weights and training the whole model with low learning rate after the decoder was trained on the initial dataset.

The task of semantic segmentation is often hampered by a significant imbalance of classes in real world data. In most semantic segmentation tasks the background class will dominate the other classes so the network can quickly overfit the training dataset. In my implementation I used a binary cross-entropy loss function with additional DICE coefficient to deal with this problem. Intersection over Union (IoU) was chosen as the target metric.

Also the potential step towards the quality improvement of the Segmentation model can be to collect more high-resolution images of the fields with erosion as deep convolutional neural networks usually need a lot of data to generalize well and such architectures as U-Net are capable of performing segmentation very precisely. Therefore, increasing the amount of data can potentially improve the quality of the soil erosion detection models.