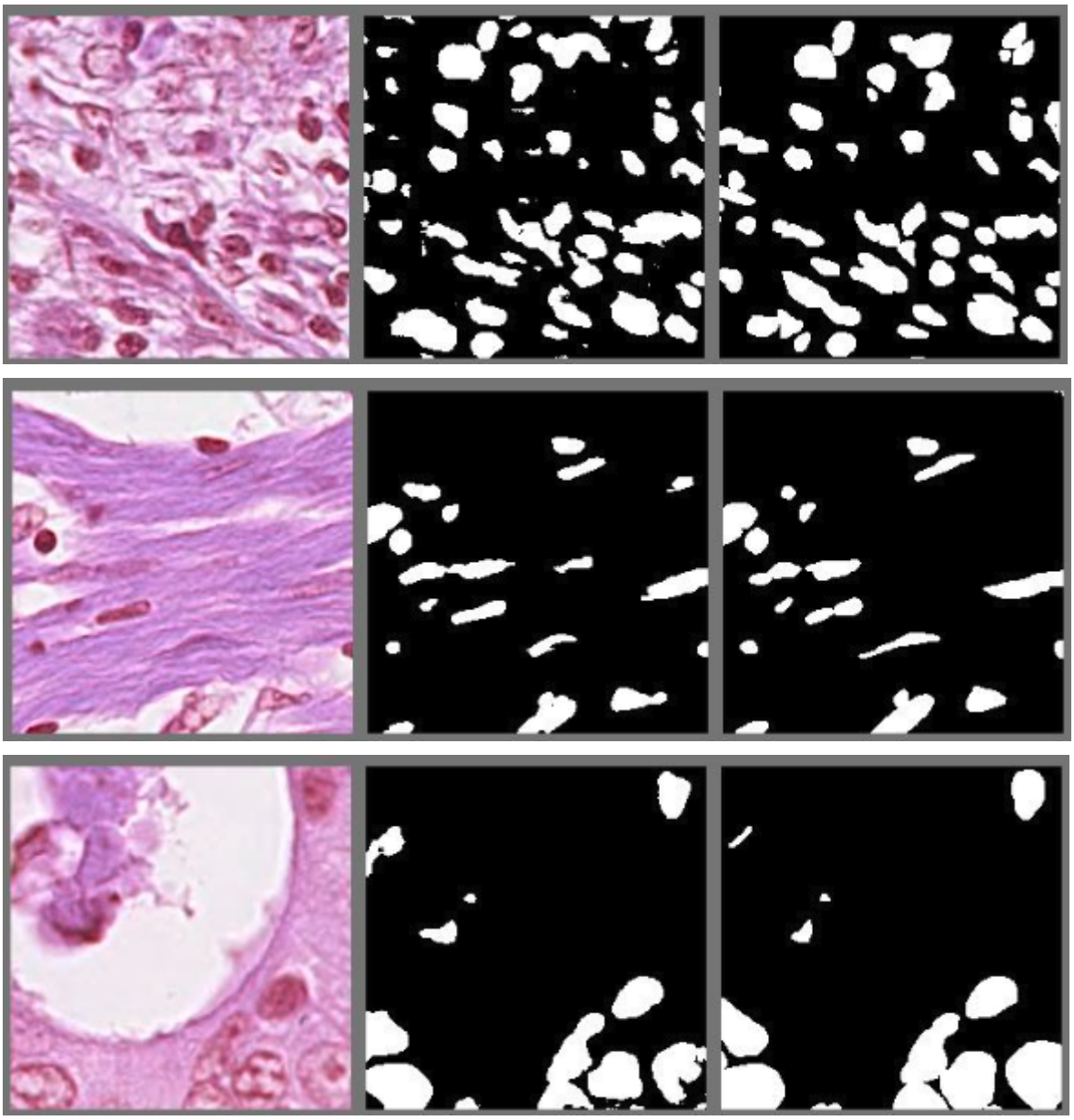
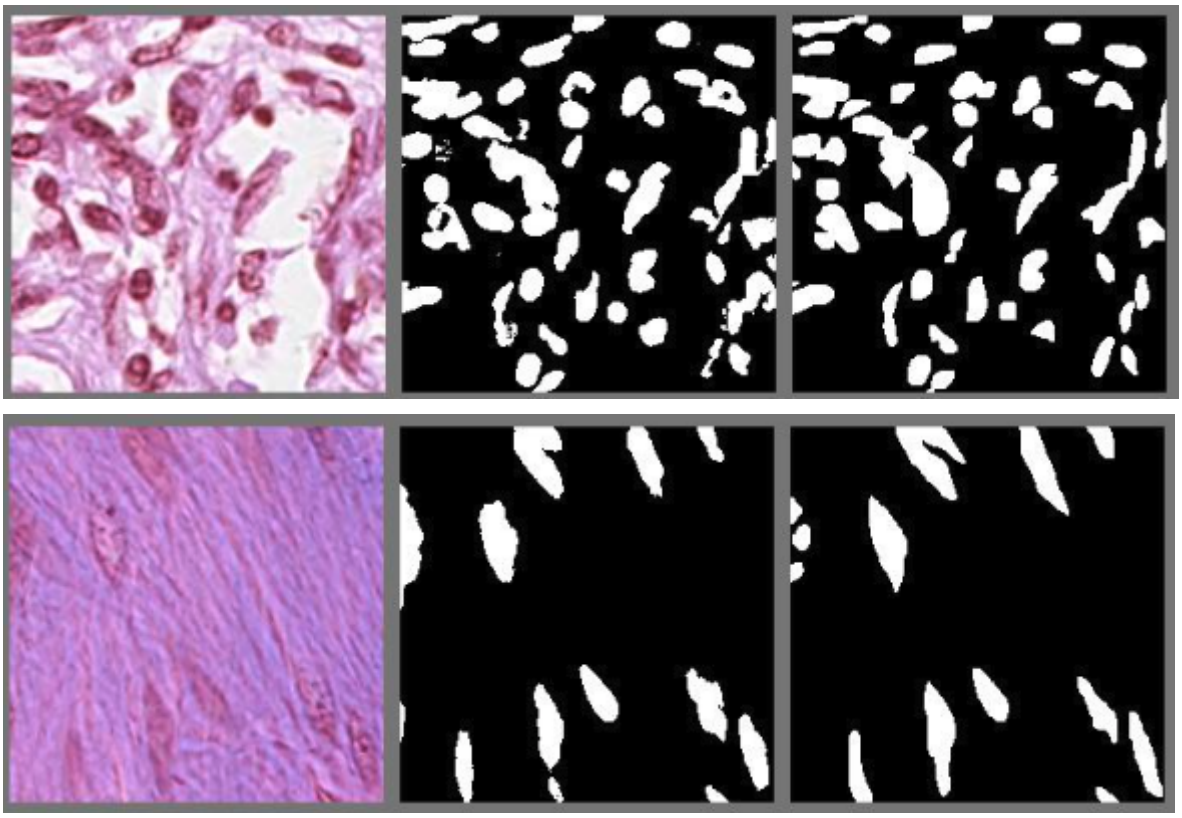
Mandatory 2 report

Good images

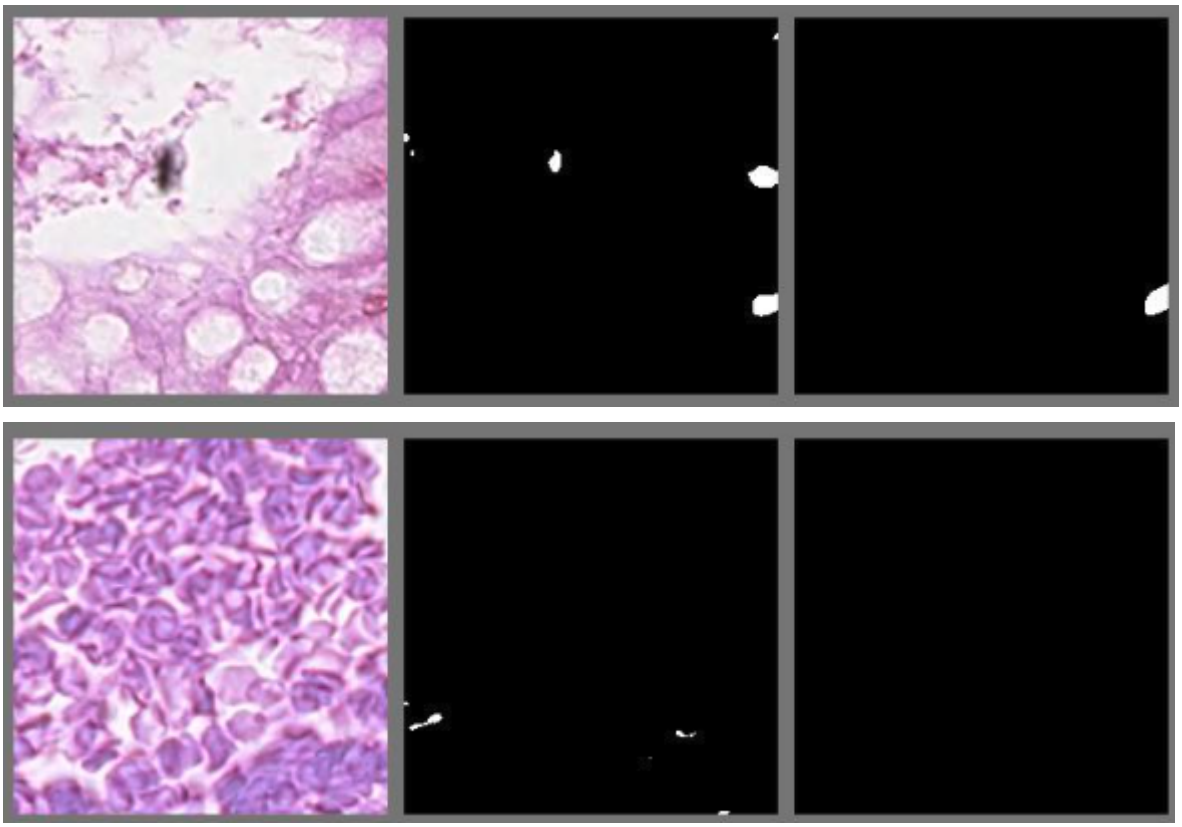
Most of the images were good. The model seems to predict the correct spots correctly, maybe with slightly different jaggedness of the border. But Most of the shapes seems correct. All the 5 images can be found in the folder `good_images`





Bad images

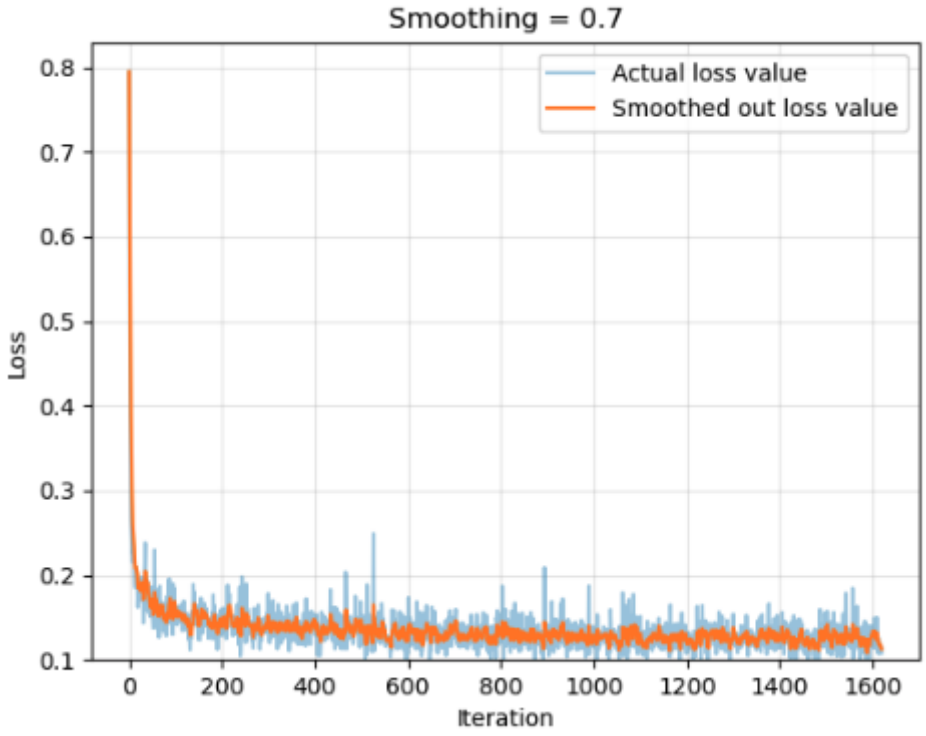
All of the images I checked seemed fine. The only exception is in the few images where the label is almost empty, then the prediction tends to classify a few spots which should not be classified. In other words, in empty images it can have a slight problem with TrueFalse predictions. But since this is a rare case, and it never predicts a lot it does not seem like a big problem. The two images can be found in the folder `bad_images`



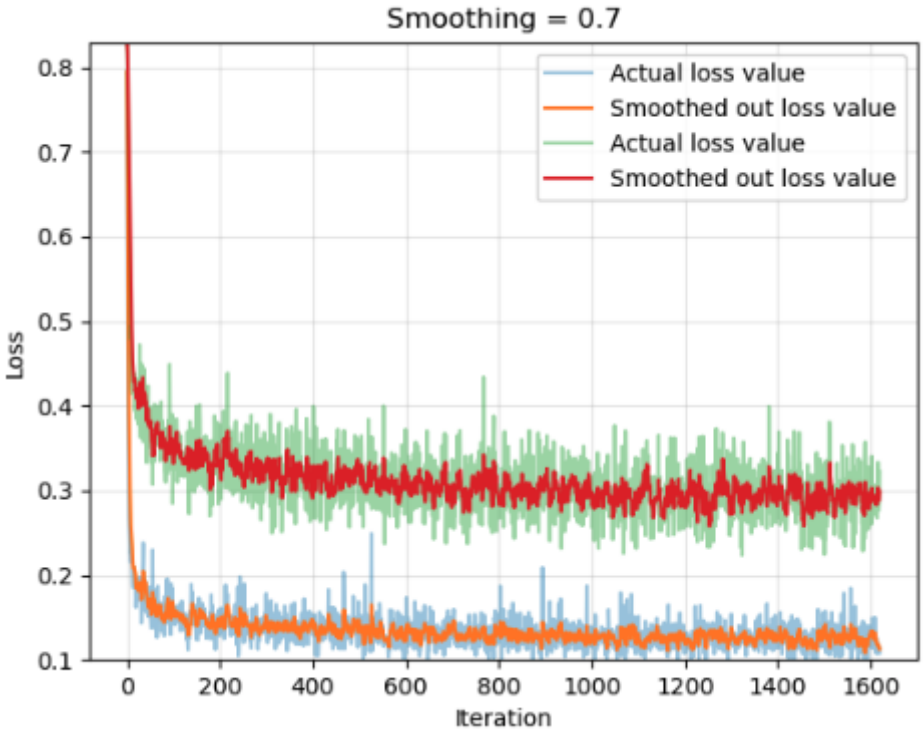
Model performance

The final model got a Dice score of **0.8553** on the validation set and **0.7948** on the test set. I ran it for **30 epochs**, with a **learning rate of 8e-4**. The original learning rate did not seem to converge, as it fluctuated a lot, especially around epoch 10-20. But when I lowered it is went down in a much more stable fashion.

BCE loss



Dice loss




```
Evaluating on testset:
dice score (the higher the better): 0.7948
```

```
Evaluation after epoch 25 took 0:00:01.779762
Epoch: 26
 * 26-005       0.120765 0.307658
 * 26-010       0.133851 0.263126
 * 26-015       0.131552 0.262992
 * 26-020       0.10008 0.342276
 * 26-025       0.107549 0.316904
 * 26-030       0.094655 0.279471
 * 26-035       0.11788 0.274511
 * 26-040       0.125228 0.272772
 * 26-045       0.116177 0.3323
 * 26-050       0.124837 0.240796
 * 26-053       0.116669 0.234524
Epoch 27 took 0:00:38.680668
EVALUATING dice score on validation set
dice score (the higher the better): 0.8460

Evaluation after epoch 26 took 0:00:01.852140
Epoch: 27
 * 27-005       0.113783 0.307229
 * 27-010       0.119499 0.278006
 * 27-015       0.124011 0.318798
 * 27-020       0.111849 0.316723
 * 27-025       0.120151 0.300237
 * 27-030       0.125856 0.265715
 * 27-035       0.116684 0.324661
 * 27-040       0.15522 0.283901
 * 27-045       0.157311 0.242266
 * 27-050       0.144964 0.265114
 * 27-053       0.122435 0.283211
Epoch 28 took 0:00:37.340373
EVALUATING dice score on validation set
dice score (the higher the better): 0.8305

Evaluation after epoch 27 took 0:00:01.946421
Epoch: 28
 * 28-005       0.155952 0.303498
 * 28-010       0.132346 0.248379
 * 28-015       0.116836 0.293945
 * 28-020       0.129778 0.275888
 * 28-025       0.120629 0.311187
 * 28-030       0.127329 0.291019
 * 28-035       0.126194 0.28681
 * 28-040       0.11435 0.321805
 * 28-045       0.127947 0.264161
 * 28-050       0.111562 0.354124
 * 28-053       0.122714 0.287241
Epoch 29 took 0:00:37.685606
EVALUATING dice score on validation set
dice score (the higher the better): 0.8403

Evaluation after epoch 28 took 0:00:01.828848
Epoch: 29
 * 29-005       0.117002 0.302684
 * 29-010       0.117696 0.311066
 * 29-015       0.118286 0.292724
 * 29-020       0.117102 0.287341
 * 29-025       0.137648 0.25426
 * 29-030       0.125235 0.291594
 * 29-035       0.138764 0.304264
 * 29-040       0.122581 0.279829
 * 29-045       0.109217 0.274885
 * 29-050       0.119095 0.333454
 * 29-053       0.115641 0.28461
Epoch 30 took 0:00:38.058635
EVALUATING dice score on validation set
dice score (the higher the better): 0.8553

Evaluation after epoch 29 took 0:00:01.951819
Saving model as a new best score has been achieve
Best dice score achieved on validation dataset wa
Saving plots
Plots saved
```