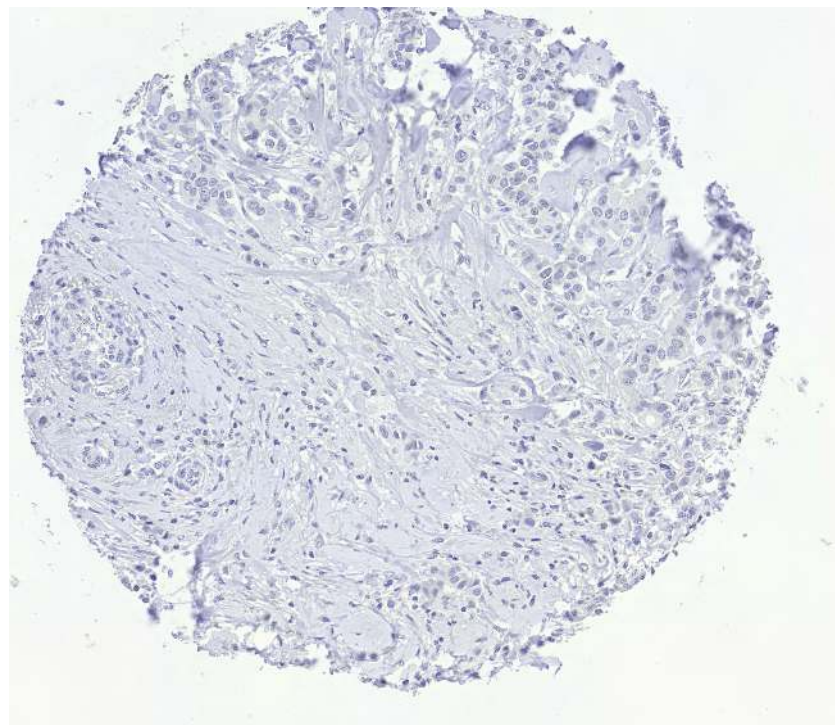


Immunohistochemistry stain detection in microscopy scans

Data



40-69%



0%

Labels

Before

- **null:** 156
- **70-100%:** 103
- **>0% and <5%:** 100
- **0%:** 89
- **>30%:** 71
- **5-15%:** 65
- **16-30%:** 56
- **40-69%:** 18
- **10-39%:** 11
- **1-9%:** 9
- **>0% and <1%:** 6

After

- **0%:** 95
- **0-5%:** 100
- **6-15%:** 74
- **16-39%:** 67
- **40-100%:** 121

Preprocessing

- Image trimming
- Image scaling
- Image normalization
- Data augmentation

Approaches

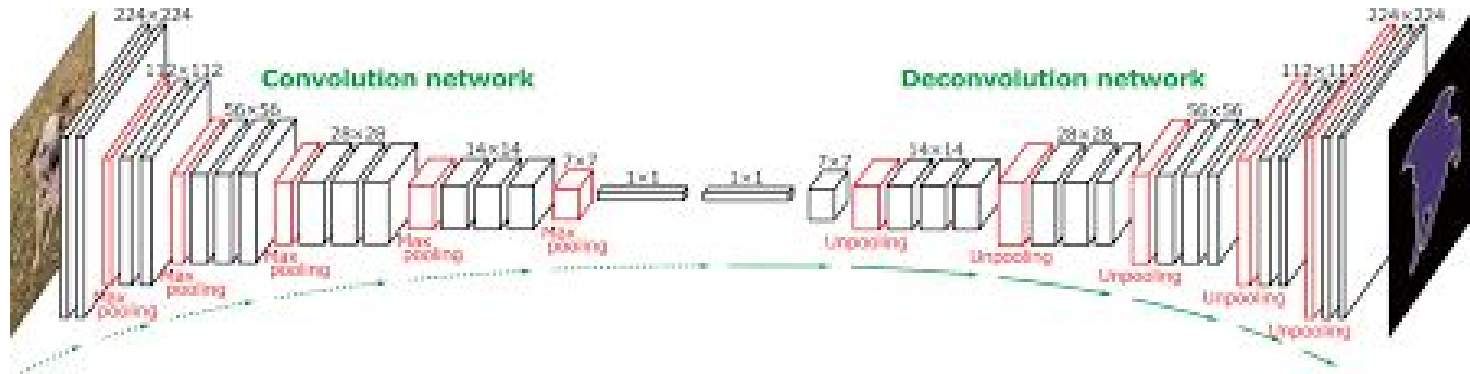
- Feature vectors -> PCA -> random forests
- Autoencoder -> PCA -> random forests
- Fully-convolutional neural networks

Results. Naive approach.

- 457 observations, $3 * 256 * 256$ features
- 100 features hold about 98% of the total variance
- Grid search with CV -> **56%** accuracy

Results. Autoencoder.

- (2 convolutions + maximum pooling) * 4 and reverse
- 75 features hold 95% of the total variance
- Grid search on XGBoost -> **30%** accuracy (near the random guess)



Results. CNN.

- NN God was in low spirits: (



WHY SO

SERIOUS?



ДА ВЫ ВСЕ ТУТ НАРКОМАНЫ

Results. A bit less naive approach.

- Take 50/100 bins from histogram of image as features.
- XGBoost
- ???
- PROFIT
- GridSearch -> **61%** accuracy (**68%** using 3 channels)

ASK

