

# Classification of breast tissue images

Руководитель проекта: Дима Фишман

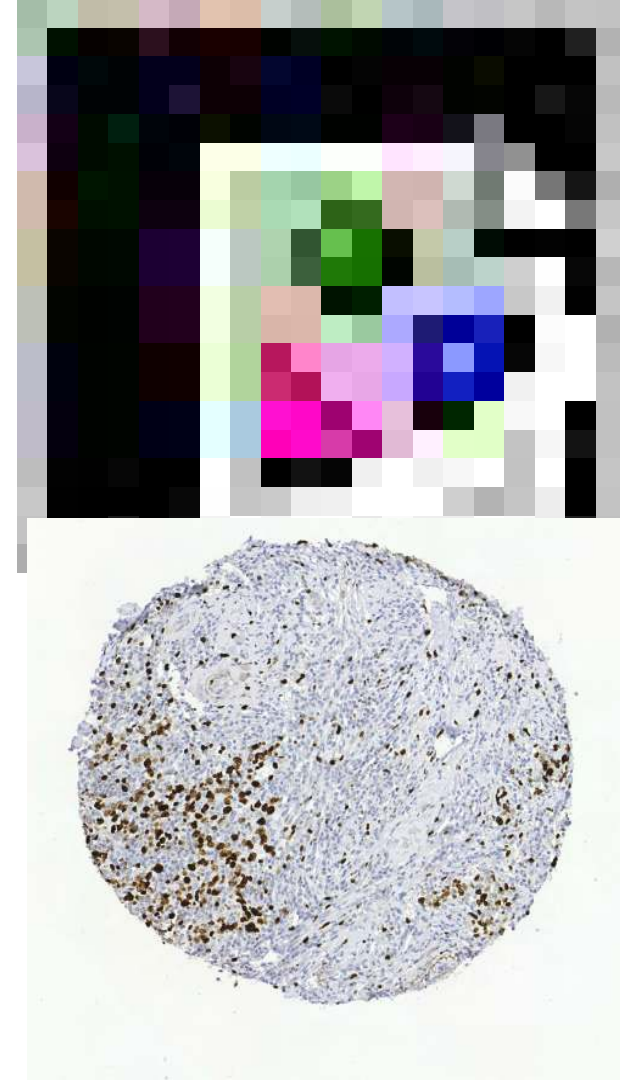
Команда: Михаил Папков, Всеволод Квачев, Василий Цветков,  
Желудкевич Анна

Санкт-Петербург  
2016

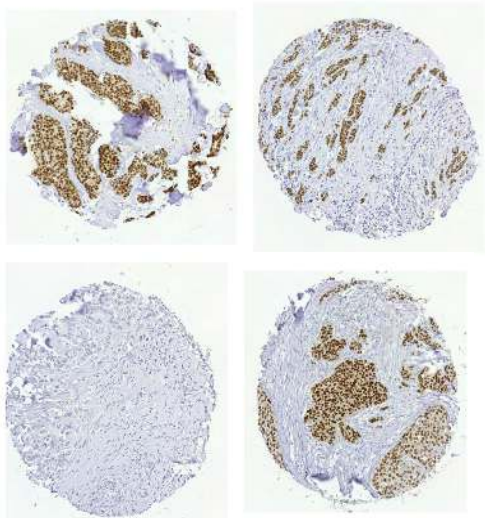
# A little bit of explaining

3 possible ways of analysis:

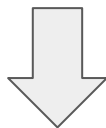
- ImageJ
- Commercial image analysis platforms (Aperio, TisuemorphKP, NuclearQuant)
- develop an algorithm which can analyse images with machine learning



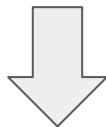
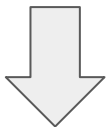
**Aim**



**Data description**



**Subdivision into classes**



**Preprocessing**



**Learning**



**Testing**



# Tools



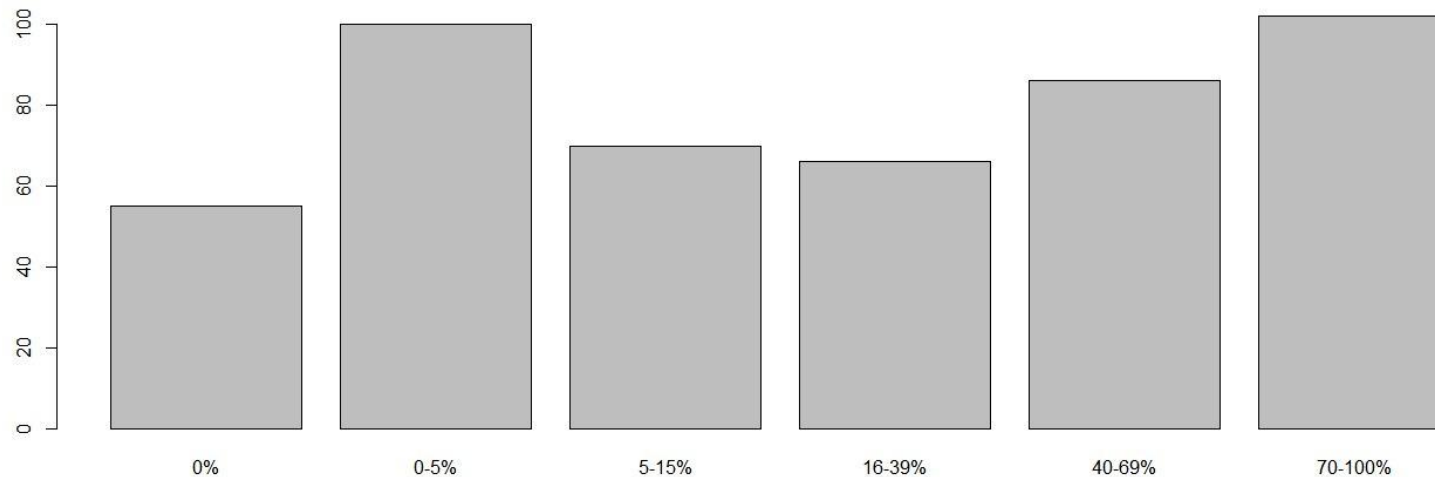
- EImage
- Caret



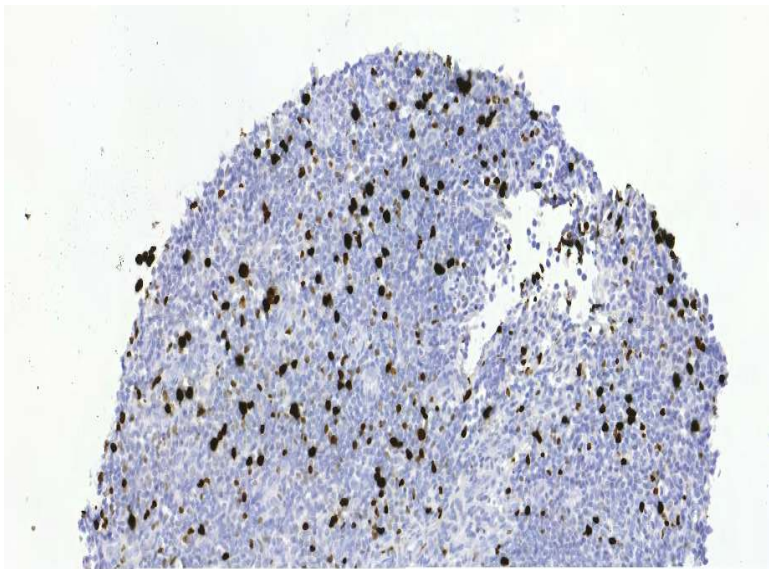
- PIL
- numpy
- scipy

# Preprocessing: classification

- tumor type
- staining and staining intensity
- percentage of positive tumor cells

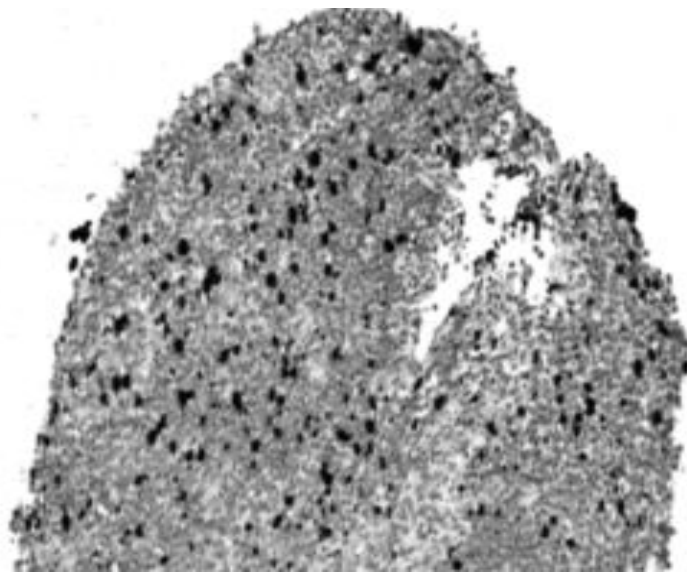
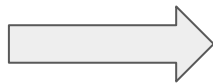


# Preprocessing: images



x 699

~4000x4000 px



x 480

256x256 px

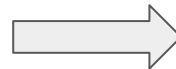
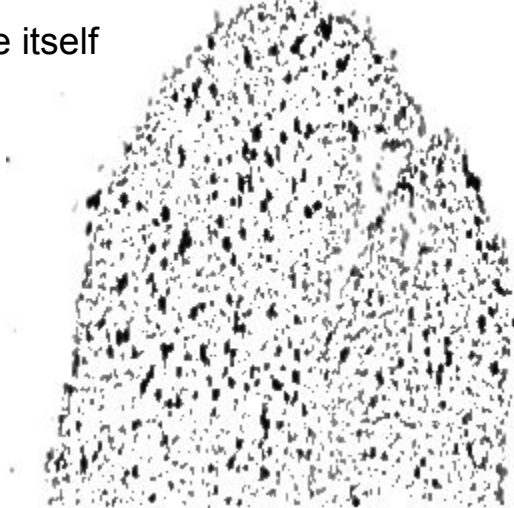
# Background removal



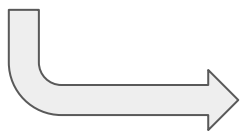
# Images



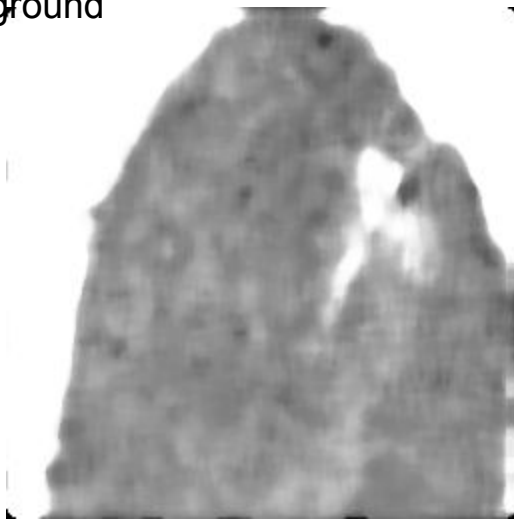
image itself



**1x65536  
vector**



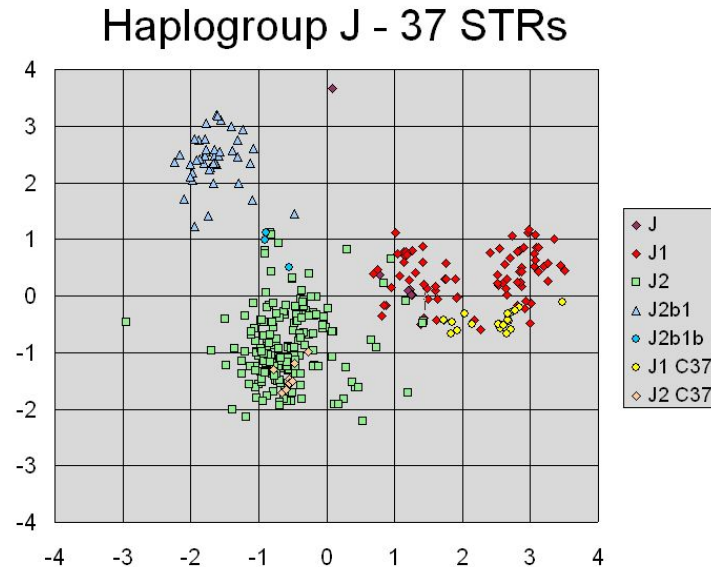
background





# PCA

- feature reduction: 65536 to 480
- add 5 features: mean intensity, median intensity, min, max, difference between mean and median



One last step



# Results and conclusions

- Accuracy: 40%

X0. X0.5. X16.39. X40.69. X5.15. X70.100.

X0. 8 2 1 4 1 3

X0.5. 6 21 7 3 13 1

X16.39. 0 2 1 1 2 0

X40.69. 2 2 3 5 3 5

X5.15. 0 1 2 3 1 0

X70.100. 0 2 5 9 1 21

**Questions?**

