Cell nuclei segmentation: support vector machine

Project proposal by
Michelle Emmert, Juan Hamdan, Laura Sanchis und Gloria Timm

Second Milestone

Implement Support Vector machine

First Milestone

Implement Dice score

Third Milestone

Implement Pre-processing

Our data

Timeline

Our data

28 images of nuclei

N2DH-GOWT1

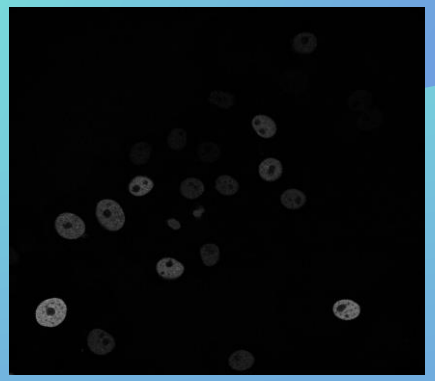
GFP transfected GOWT1 mouse embryonic stem cells

N2DL-HeLa

Histone 2B (H2B)-GFP expressing HeLa cells

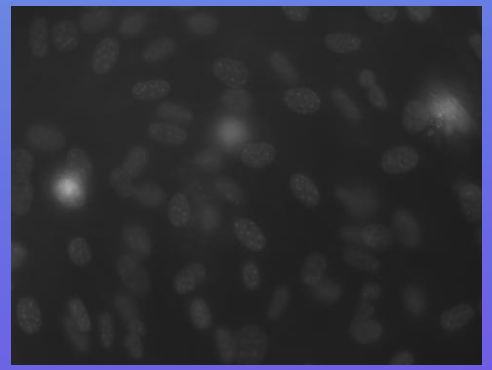
NIH3T3

mouse embyonic fibroblast – CD tagged (EGFP)



N2DH-GOWT1

N2DL-HeLa



NIH3T3

Our data

N2DH-GOWT1

- Medium difficulty
- Heterogeneous staining
- Prominent nucleoli + mitoses
- Cells entering and leaving the field of view and frequent cell collisions

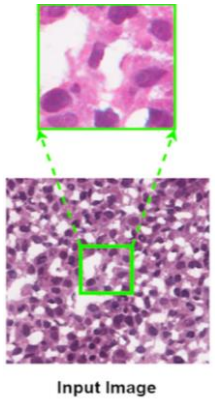
N2DL-HeLa

- High difficulty
- High cell density and low resolution
- Frequent mitoses
- Presence of colliding, entering and leaving cells with low fluorescence intensity

NIH3D3

- Medium difficulty
- Nuclei far apart + less clustering
- Nuclei homogeneous in shape and size
- But they also vary in brightness between images
- Images often contain visible debris.

Our goal



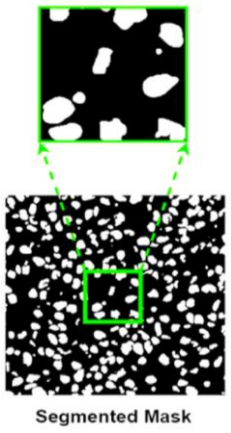
Input: Microscopic images

Preprocessing & filtering

Segmentation:

Support vector machine

Data mining: Counting nuclei



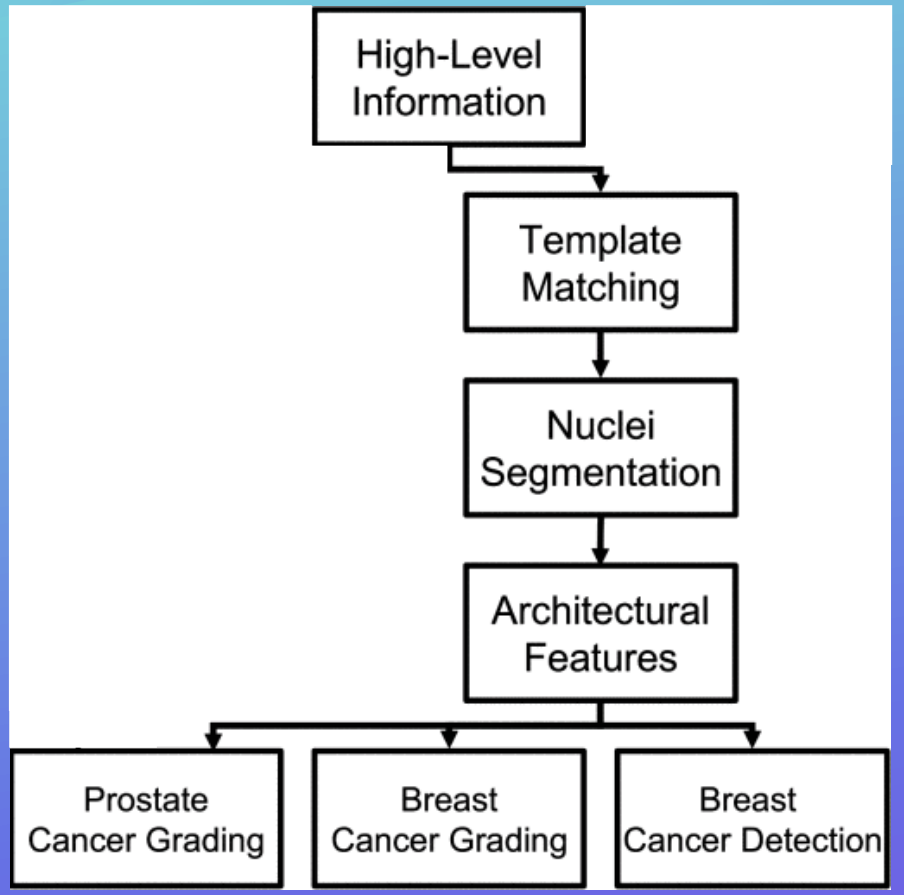
Output: Segmented image & number of nuclei

https://www.sciencedirect.com/science/article/pii/S0010482520304066

Our goal

Cell nuclei segmentation using support vector machine

Quantification of cell nuclei



quantify degree of malignancy (= grading)
→ key feature to predict patient prognosis and decide on treatment options

https://ieeexplore.ieee.org/abstract/document/4540988

Second Milestone

Implement Support Vector machine

First Milestone

Implement Dice score

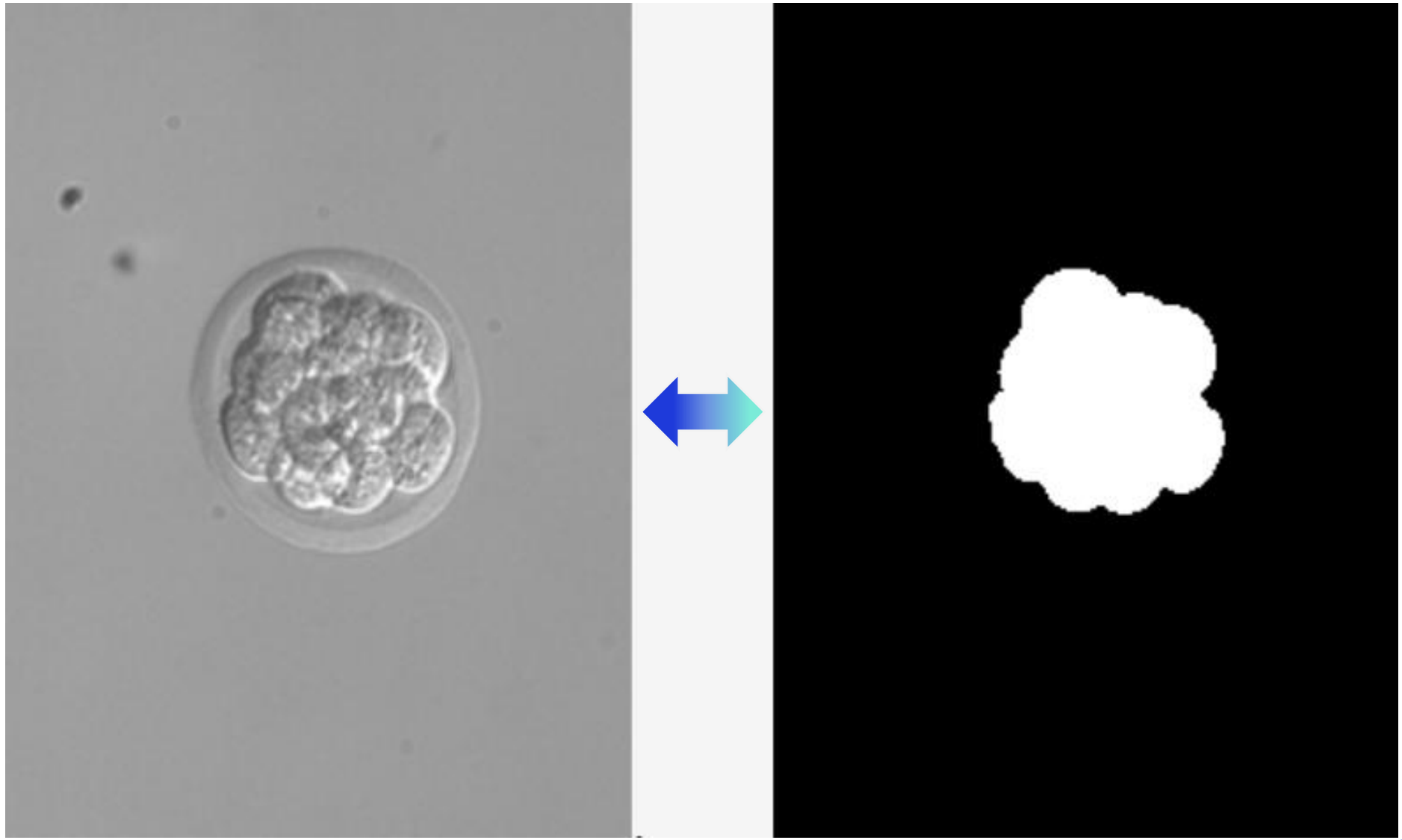
Third Milestone

Implement Pre-processing

Our data

Timeline

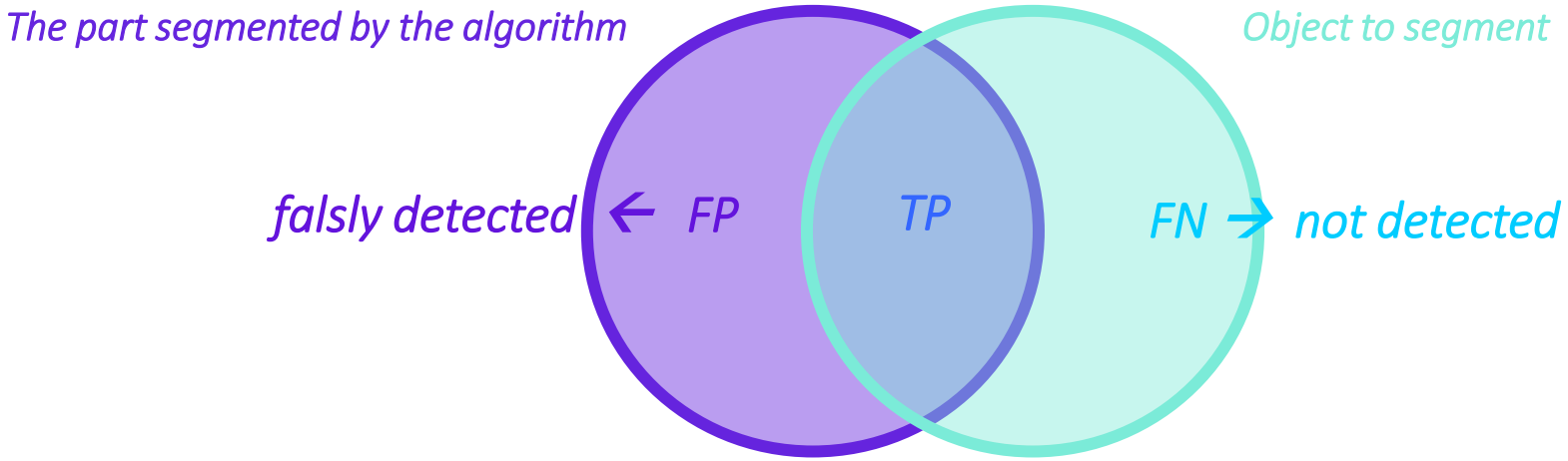
Dice Score



https://bbbc.broadinstitute.org/BBBC003

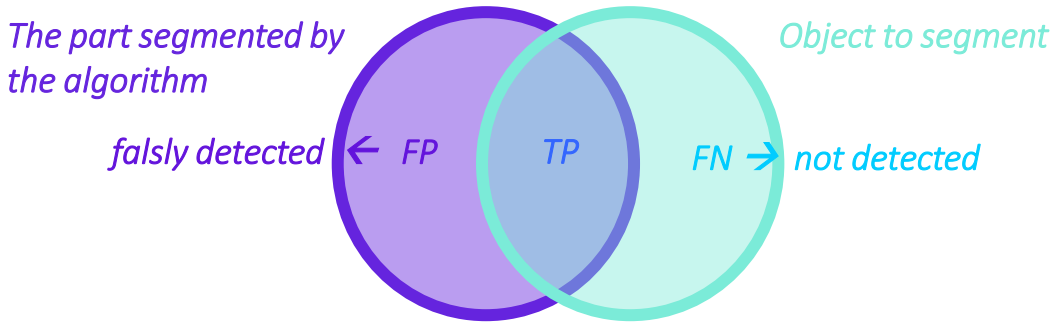
Dice Score

Evaluating the quality of the segmentation quantitatively



- $\text{IoU} = \frac{Intersection}{Union} = \frac{TP}{TP+FP+FN}$ ← not as well differentiable
- Better: $\text{Dice} = \frac{2*Intersection}{Union+Intersection} = F_1 = \frac{1}{\frac{1}{Prec}+\frac{1}{Recall}} = \frac{2\,TP}{2\,TP+FP+FN}$
- $\text{Loss} = 1 - Dice \quad \epsilon\,[0;1]$
- $\text{Loss} = 1 - \frac{2\sum bf}{\sum b+f}$

Dice Score



- $\text{IoU} = \frac{Intersection}{Union} = \frac{TP}{TP+FP+FN}$ ← not as well differentiable
- Better: $\text{Dice} = \frac{2*Intersection}{Union+Intersection} = F_1 = \frac{1}{\frac{1}{Prec}+\frac{1}{Recall}} = \frac{2\,TP}{2\,TP+FP+FN}$
- $\text{Loss} = 1 - Dice \quad \epsilon\,[0;1]$
- $\text{Loss} = 1 - \frac{2\sum bf}{\sum b+f}$

Planned analysis steps

- *Write code for Dice-Score function*
- *Unit-testing*
- *Write code for synthetic images*

Characteristics of the first Milestone

- *Measure for evaluating our model*

The purpose

- *Quantify the quality of our model*

Second Milestone

Implement Support Vector machine

First Milestone

Implement Dice score

Third Milestone

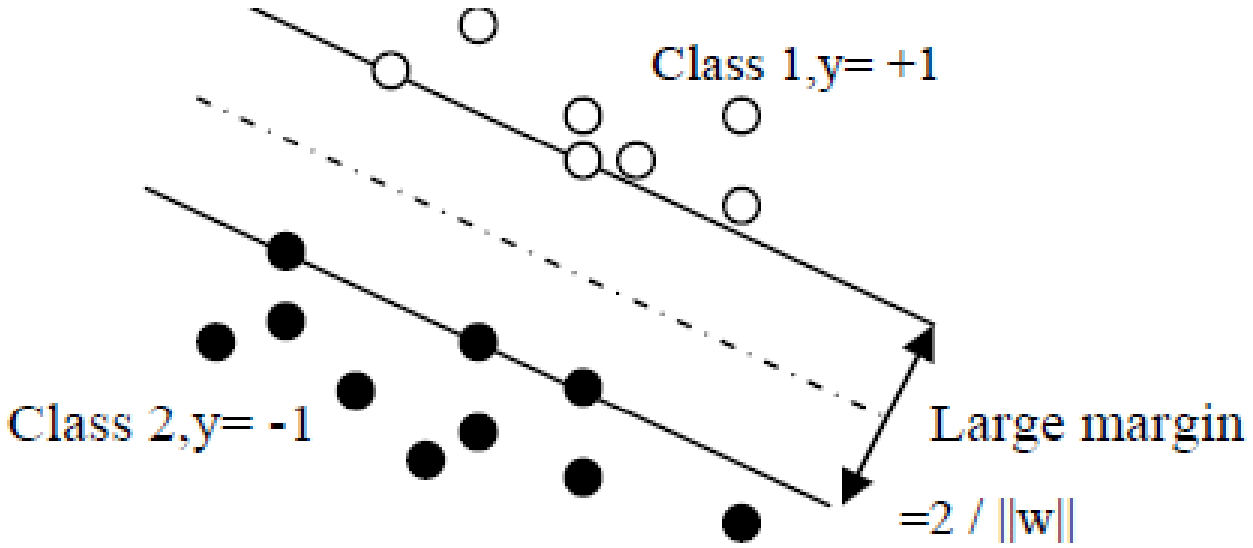
Implement Pre-processing

Our data

Timeline

Mathe – Michelle ☺☺☺

Support Vector machine

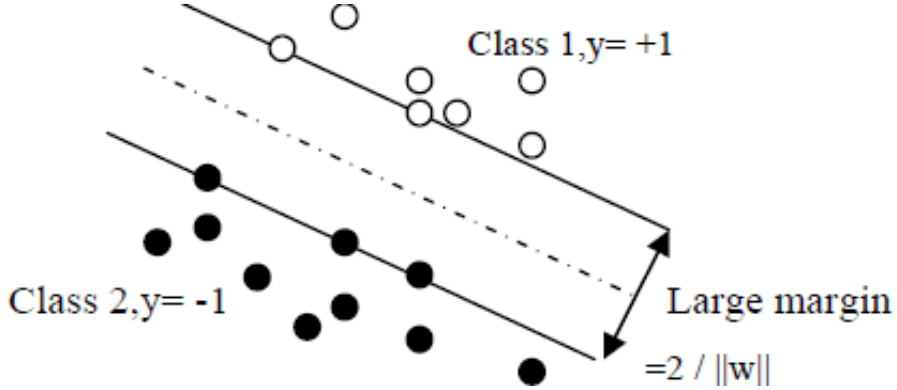


- based on statistical learning theory
- uses a decision function that defines a hyperplane
- hyperplane is the classifier, that categorizes data into two groups
- Phase 1: training phase

  Phase 2: generalization phase

https://ieeexplore.ieee.org/document/6525921

Support Vector machine



- based on statistical learning theory
  - uses a decision function that defines a hyperplane
    - hyperplane is the classifier, that categorizes data into two groups
      - Phase 1: training phase
      - Phase 2: generalization phase

Planned analysis steps

- *Implement a support vector machine*
- *Using the Dice-score-function to evaluate the performance of our SVM*

Characteristics of the second Milestone

- *Support vector machine labels pixels with ‚cell nucleus‘ or ‚background‘*

The purpose

- *Use trained model to automatically segment cell nuclei images*
- *Compare the segmented images to the ground truth*
  → *using Dice Score*

https://ieeexplore.ieee.org/document/6525921

Second Milestone

Implement Support Vector machine

First Milestone

Implement Dice score

Third Milestone

Implement Pre-processing

Our data

Timeline

Pre-processing

- Pre-processing steps:
  1. Noise reduction: reduce the noise while preserving the edges
  2. Super-pixel segmentation: join similar pixels together into a super-pixel region with a similar intensity value

- Desired effect:
  1. Average local pixel intensity values to reduce missing information from original picture
  2. Separate nuclei which appear fused for an improved segmentation

https://www.mathworks.com/help/images/ref/imgaussfilt.html

Pre-processing

- Pre-processing steps:
  1. Noise reduction: reduce the noise
  2. Super-pixel segmentation: join pixels into a super-pixel region
- Desired effect:
  1. Reduce missing information from picture
  2. Separate nuclei which appear fused

Planned analysis steps

- *Methods:*
  1. *Gaussian filter*
  2. *Gradient-ascend-based super pixel algorithms, e.g. Watershed*

Characteristics of the third Milestone

- *Improve Dice Score of segmentation method through better image quality*

The purpose

- *Achieve better registration results*

Second Milestone

Implement Support Vector machine

First Milestone

Implement Dice score

Third Milestone

Implement Pre-processing

Our data

Timeline

Timeline

Dice score

12.05 – 24.05.

25.05. – 14.06.

Support Vector Machine

Pre-processing

15.06. – 21.06.

from 22.06.

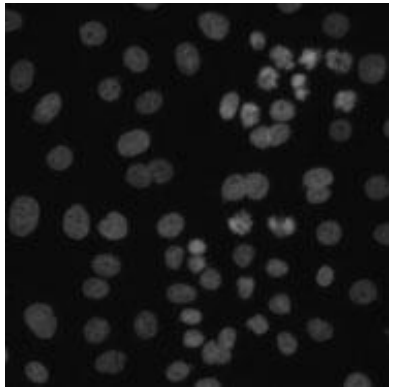
Explore other possibilities

Other options to be explored

- testing our algorithm **on other data**, on which other algorithms were already tested on
  → compare results

e.g. Broad Bioimage Benchmark Collection 001 which was already used in

Nosova SA, Turlapov VE (2019) Detection of Brain Cells in Optical Microscopy Based on Textural Features with Machine Learning Methods. Program Comput Soft 45, 171–179



- try out more advanced pre-processing

e.g. anisotropic filtering

https://bbbc.broadinstitute.org/BBBC001

Thank you for listening!