Introduction to Image Processing

Martin Schätz, PhD

@SchatzCZ martinschatz-cz

VMCF - Viničná Microscopy Core Facility

@IMCF_Vinicna

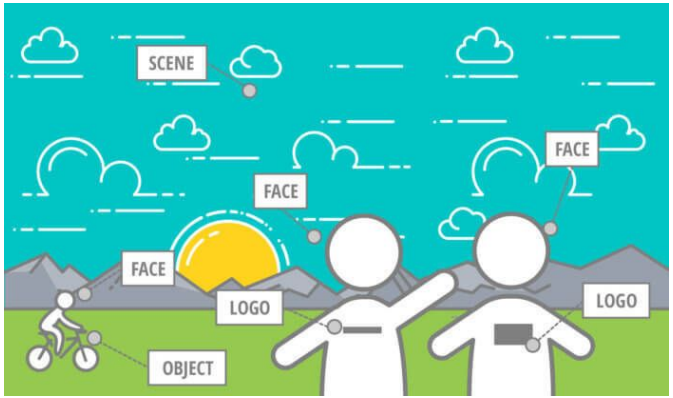
github.com/vmcf-konfmi

PŘÍRODOVĚDECKÁ FAKULTA
Univerzita Karlova

Term Definition

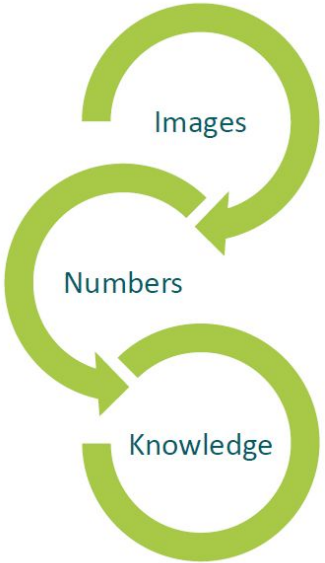
Image Analysis?

Image analysis (also known as “computer vision” or image recognition) is the ability of computers to recognize attributes within an image.



BioImage Analysis?

Understanding and quantifying microscopy & medical image data.



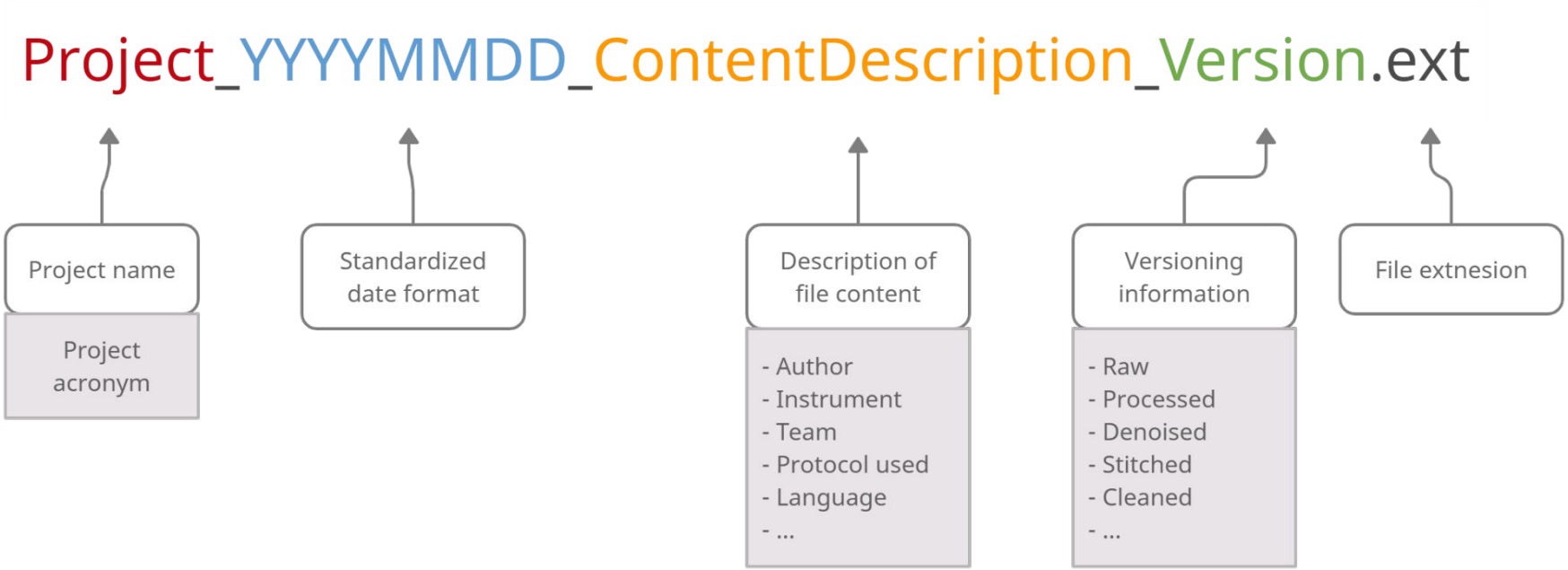
- Objective
- Reliable
- Reproducible
- Replicable
- Repeatable

Introduction to (Bio-)Image Analysis

Bio-image analysis is supposed to be

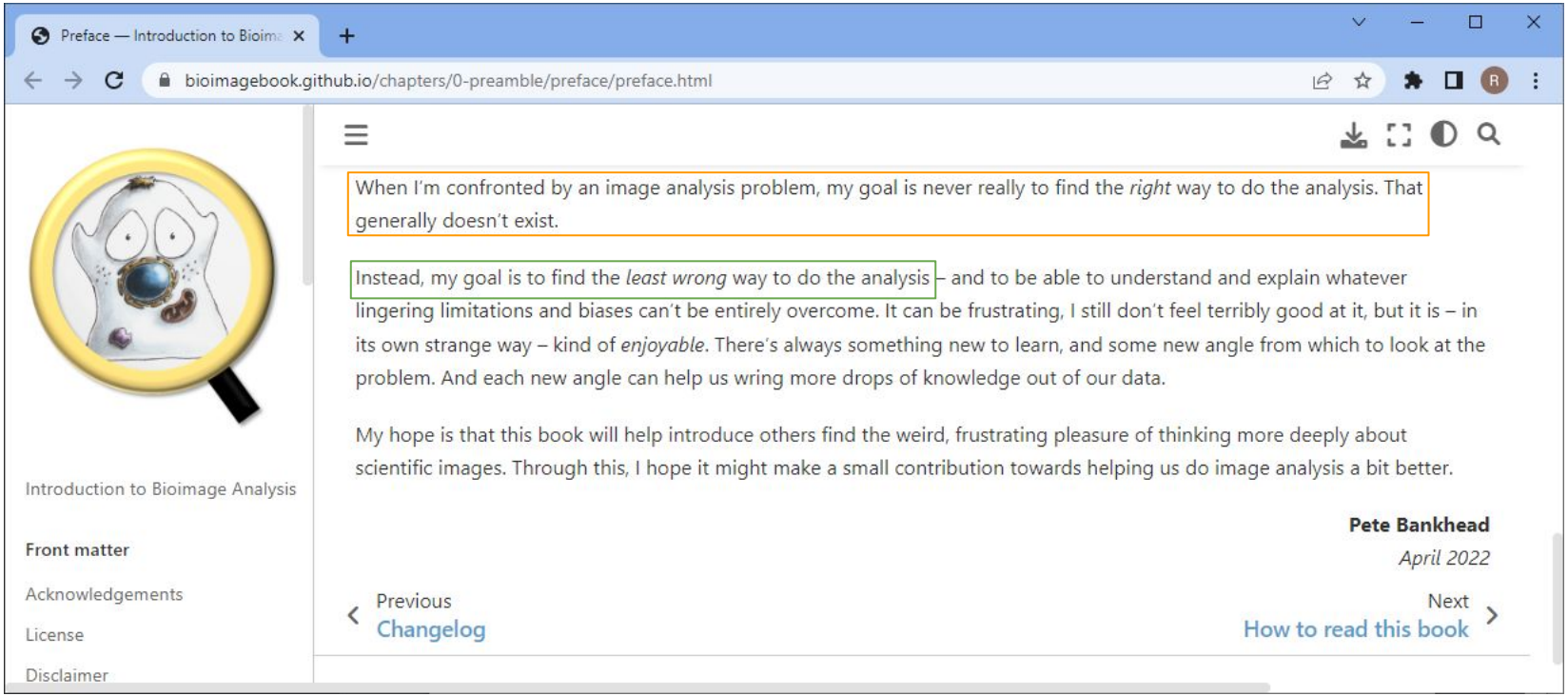
- Quantitative
  - We derive numbers from images which describe physical properties of the observed sample.
- Objective
  - The derived measurement does not depend on who did the measurement. The measurement is free of interpretation.
- Reliable (trustworthy / validated)
  - We are confident that the measurement is describing what it is supposed to describe.
- Reproducible
  - Enabling others to re-do the experiment. For this, documentation is crucial!
- Replicability
  - Others do execute the same analysis, potentially on other data, and see consistent results.
- Repeatable
  - We can do the same experiment twice under the same conditions and get the same measurements.

~~Harder,~~ Better, Faster, Stronger -> Setup naming convention



source:https://biblio.uottawa.ca/en/services/faculty/research-data-management/file-naming-and-organization-data

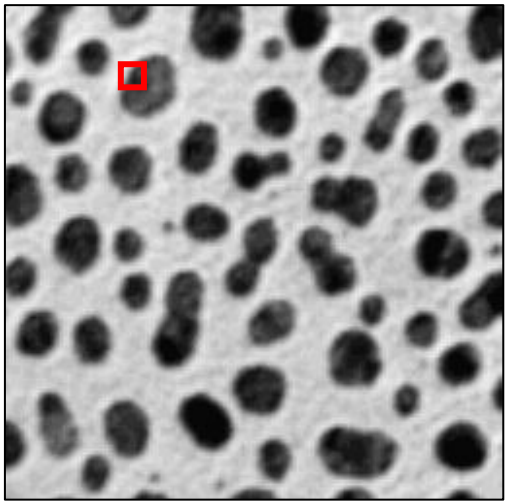
Bio-image Analysis: good scientific practice

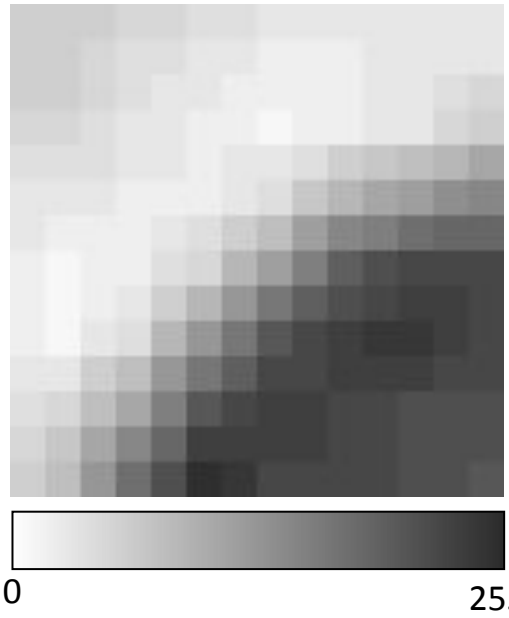


https://bioimagebook.github.io/chapters/0-preamble/preface/preface.html

Images and pixels

- An image is just a matrix of numbers
- Pixel: “picture element”
- The edges between pixels are an artefact of the imaging / digitization. They are not real!

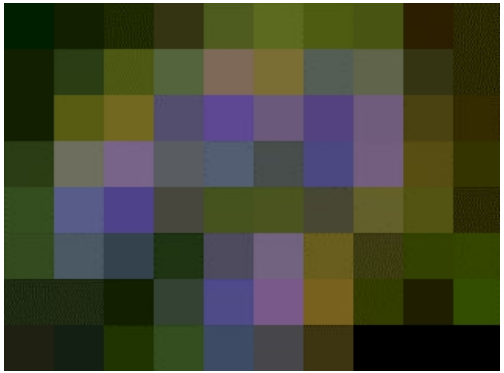




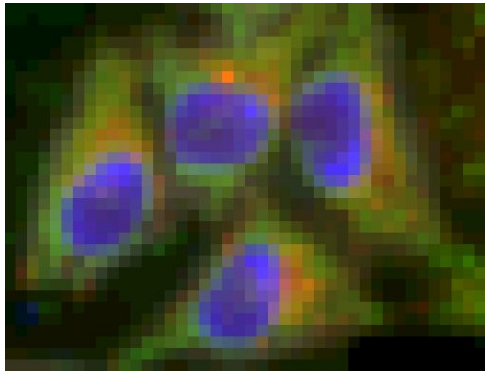
48	48	48	40	40	32	32	24	24	24	24	24	24	24
48	48	40	32	32	24	24	16	16	16	24	24	24	24
48	48	40	32	24	24	16	16	16	16	24	24	32	40
40	40	32	24	24	16	16	8	16	16	24	24	40	48
32	32	32	24	24	16	24	24	32	48	56	64	72	88
24	24	24	16	16	16	24	32	56	72	88	96	112	120
24	16	16	16	24	32	48	64	96	120	128	144	152	152
16	8	16	16	32	40	72	96	128	160	176	184	184	184
16	8	16	24	48	72	104	136	160	176	184	192	192	184
16	8	24	32	72	104	136	168	184	192	200	200	192	184
24	24	48	64	104	136	160	184	184	192	192	192	184	184
32	40	64	88	128	168	184	192	192	184	184	176	176	176
40	56	88	120	152	192	192	192	192	184	184	176	176	176
48	64	104	144	176	208	200	184	184	184	184	176	176	168

Pixel size versus resolution

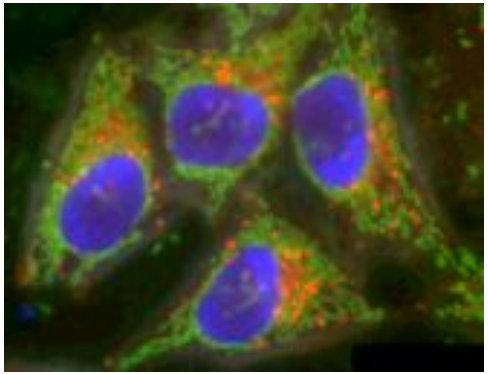
- Pixel size is a digital property of an image.
- You configure it during the imaging session at the microscope.



Pixel size: 3.3 µm



Pixel size: 0.8 µm

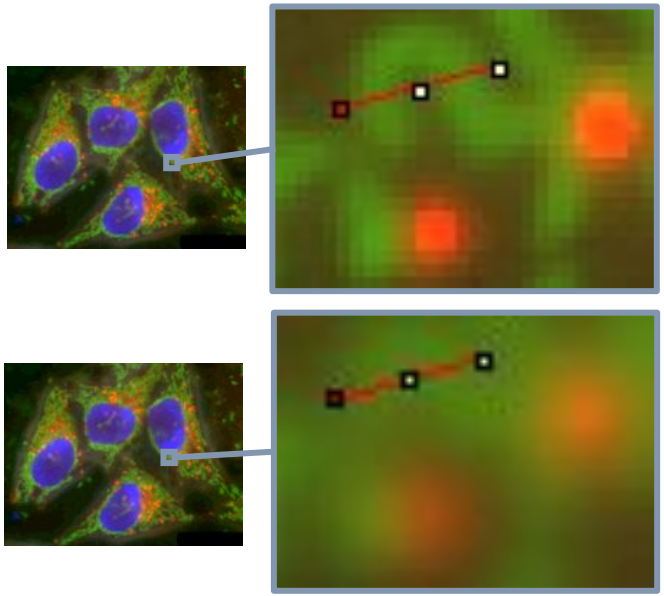


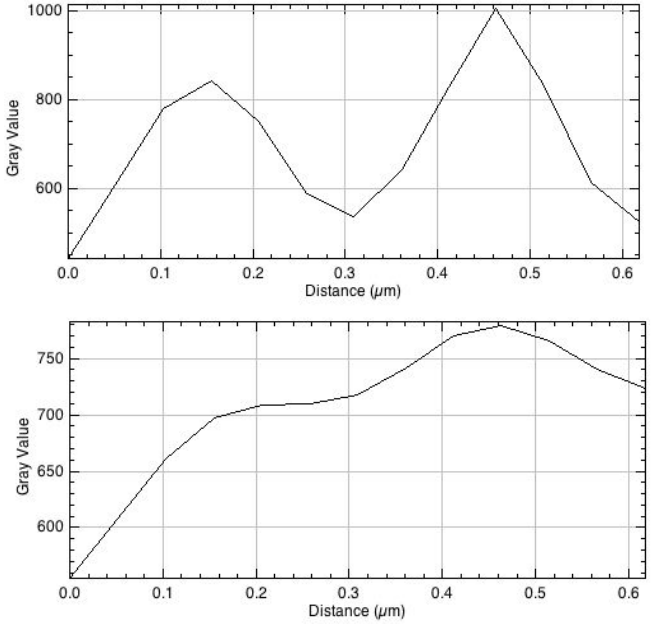
Pixel size: 0.05 µm

- We are not talking about resolution!

Pixel size versus resolution

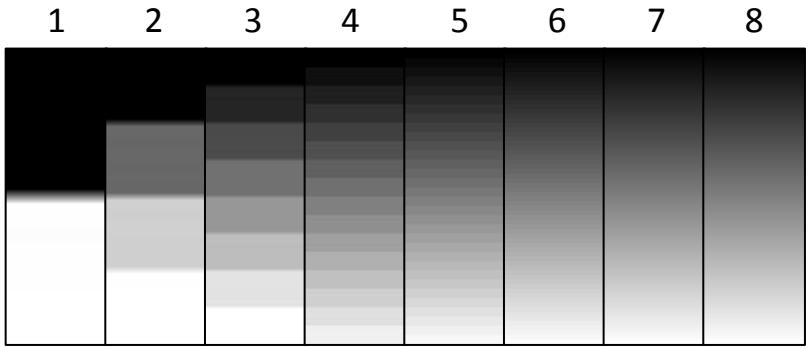
- Resolution is a property of your imaging system.
- The measure of how close object can be in an image while still being differentiable, is called spatial resolution.





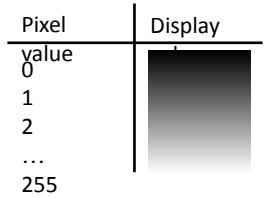
Bit-depth

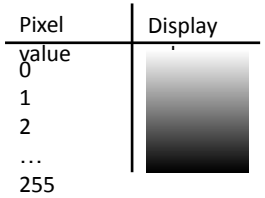
- A bits is the smallest memory unit in computers, *atomic data*.
- The bit-depth *n* enumerates how many different intensity values are present in an image:
  - 2^n grey values
- In microscopy, images are usually stored as 8, 12 or 16-bit images.

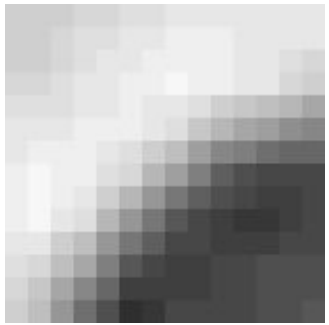


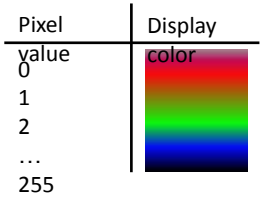
Colormaps / lookup tables

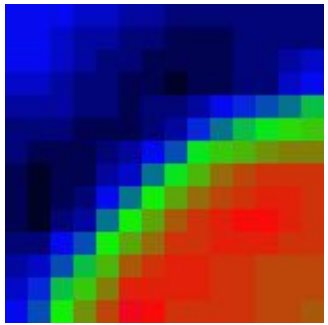
- The lookup table decides how the image is displayed on screen.
- Applying a different lookup table does not change the image. All pixel values stay the same, they just appear differently







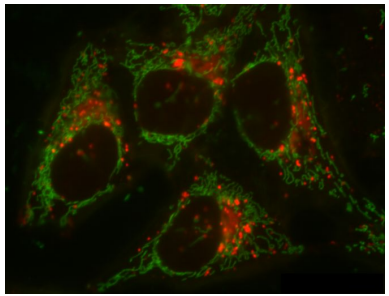




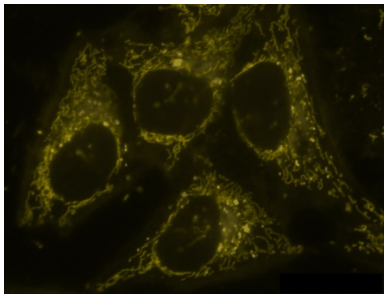
Colormaps / lookup tables

- Choose visualization of your color tables wisely!
- Think of people with red/green blindness!

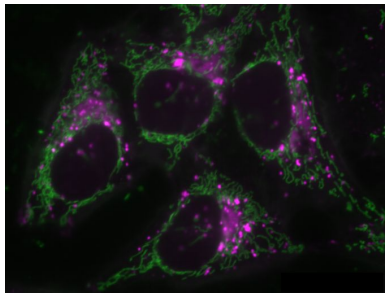
Common view

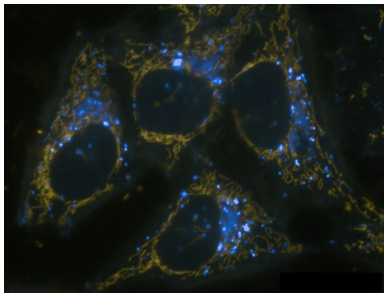


Red/green blind people may see it like this



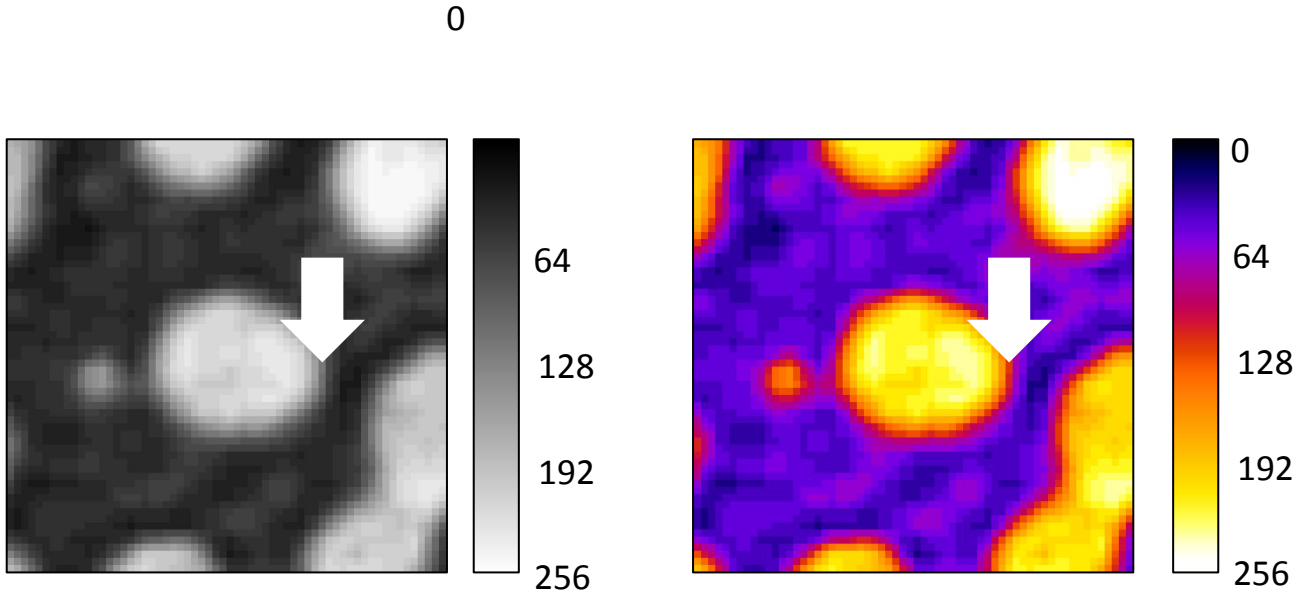
Replace red with magenta!





Colormaps / lookup tables

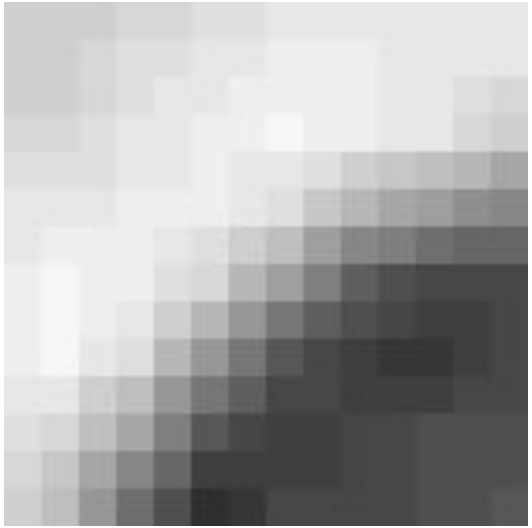
Which intensity does the marked pixel have?

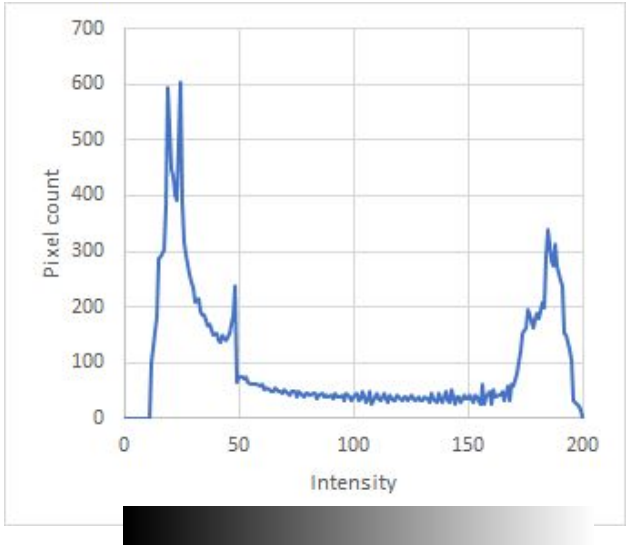


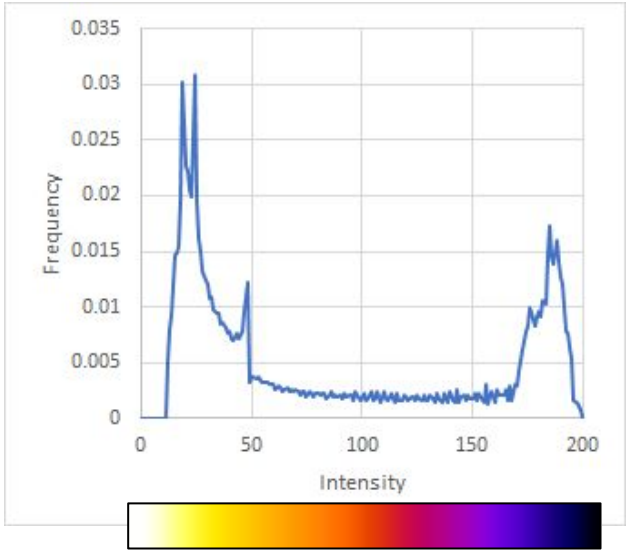
0 | 64 | 128 | 192

Histograms

- A histogram shows the probability distribution of pixel intensities.
- The probability of a pixel having a certain grey value can be measured by counting pixels and calculating the frequency of the given intensity.
- Whenever you see a histogram, try to imagine the lookup-table on the X-axis

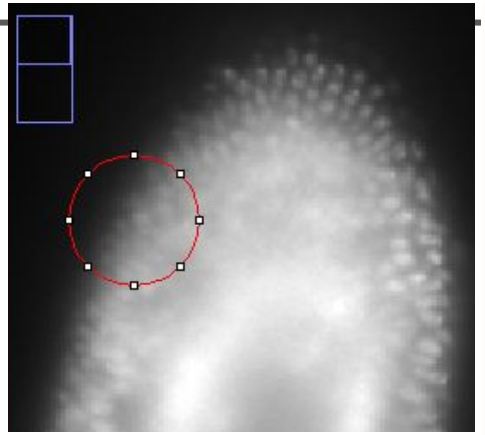


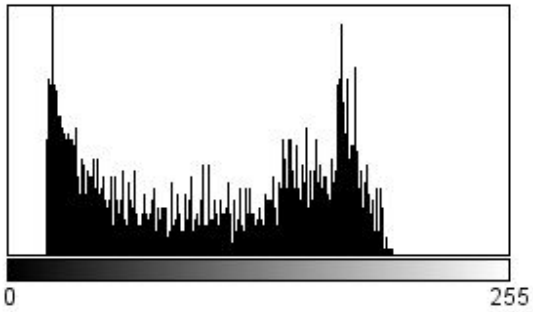




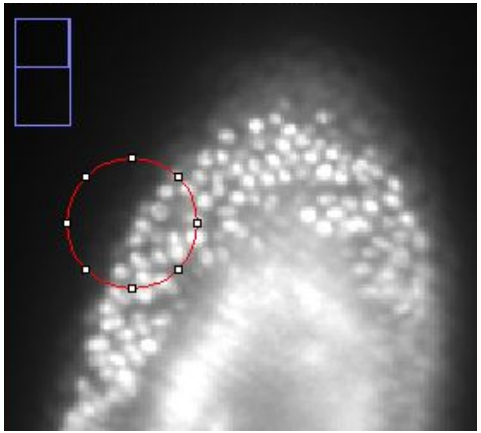
Histograms

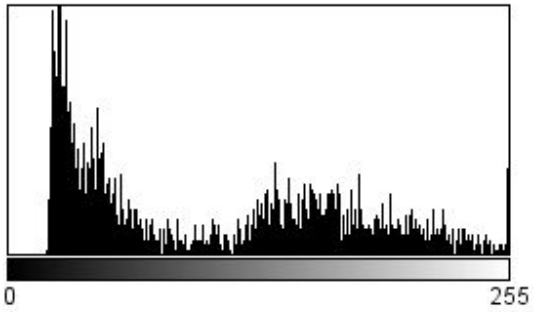
- Histograms are summaries of images
- Tell stories, e.g. about image quality





Count: 2053
Mean: 103.818
StdDev: 57.093
Min: 19
Max: 196
Mode: 22 (41)

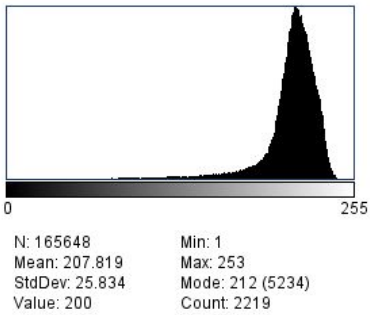


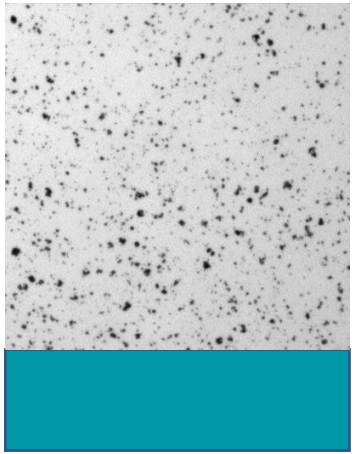


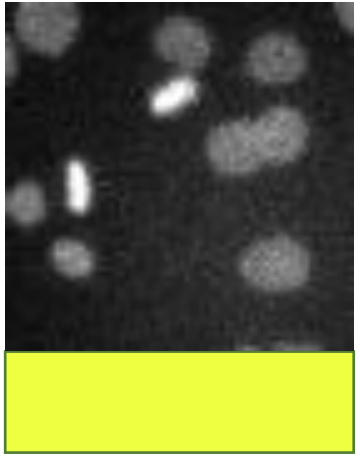
Count: 2053
Mean: 103.370
StdDev: 70.260
Min: 19
Max: 255
Mode: 25 (49)

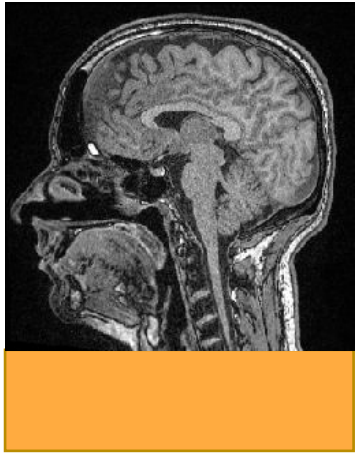
Histograms

To which of the three images does this histogram belong to?



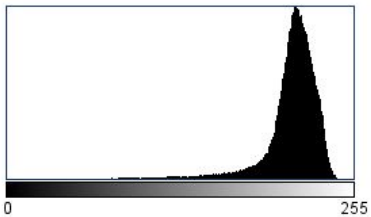






Histograms

To which of the three images does this histogram belong to?



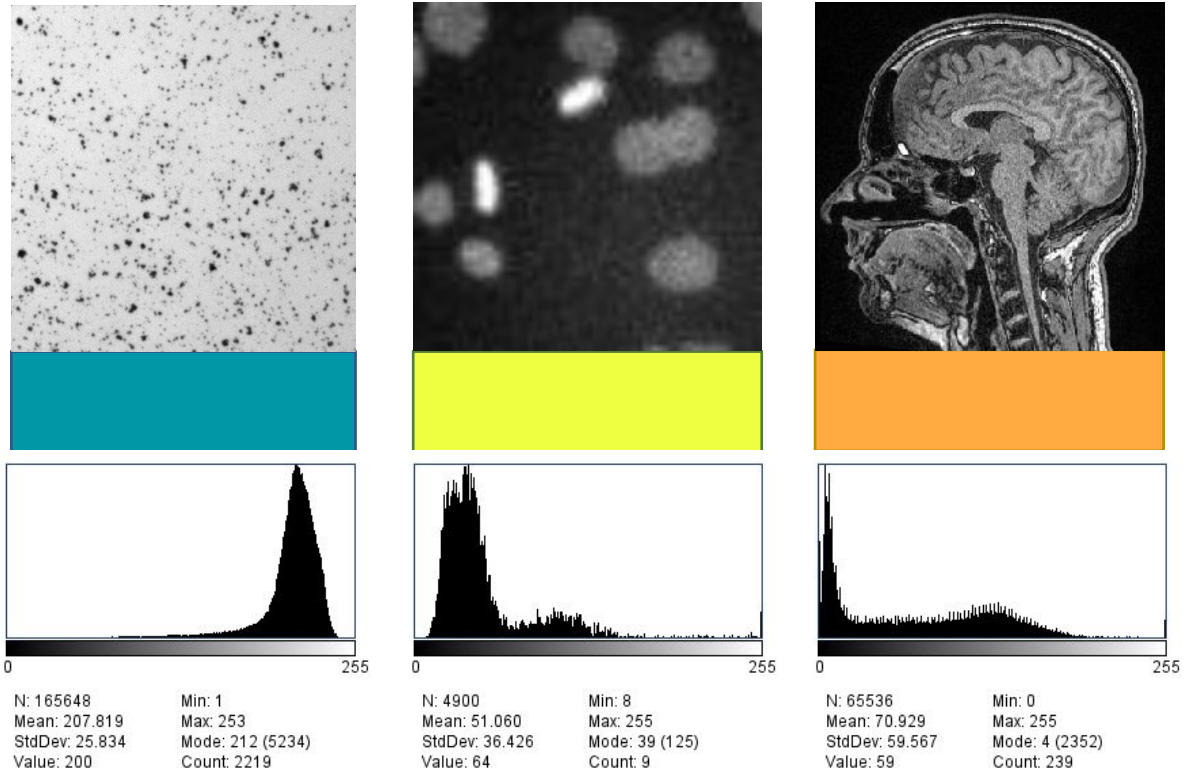


Image Processing and Filtering

Filters

- An image processing filter is an operation on an image.
- It takes an image and produces a new image out of it.
- Filters change pixel values.

- There is no “best” filter. Which filter fits your needs, depends on the context.
- Filters do not do magic. They can not make things visible which are not in the image.

- Application examples
  - Noise-reduction
  - Artefact-removal
  - Contrast enhancement
  - Correct uneven illumination

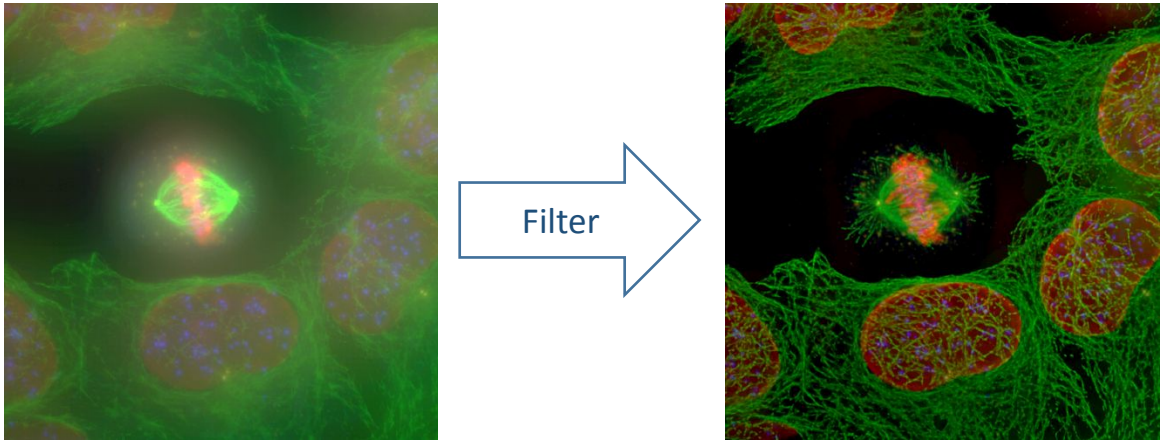
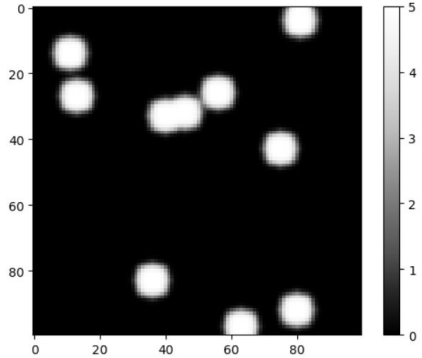


Image source: Alex Bird / Dan White MPI CBG

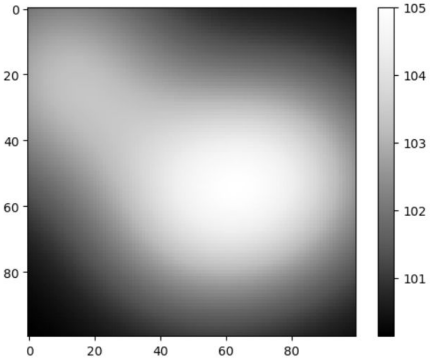
Effects harming image quality

- Image formation (simulated)

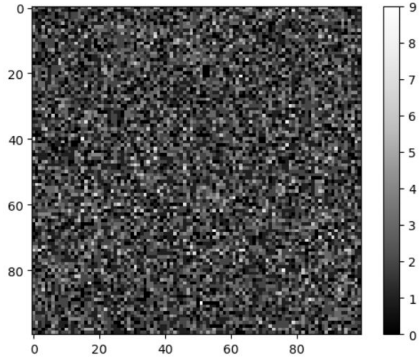
“nuclei”



“background”



“noise”



- Aberrations, defocus
- Motion blur

- Light from objects behind and in front of the scene (out-of-focus light)
- Dirt on the object slide
- Camera offset

- Shot noise (arriving photons)
- Dark noise (electrons made from photons)
- Read-out-noise (electronics)

Effects harming image quality

- Image formation (simulated)

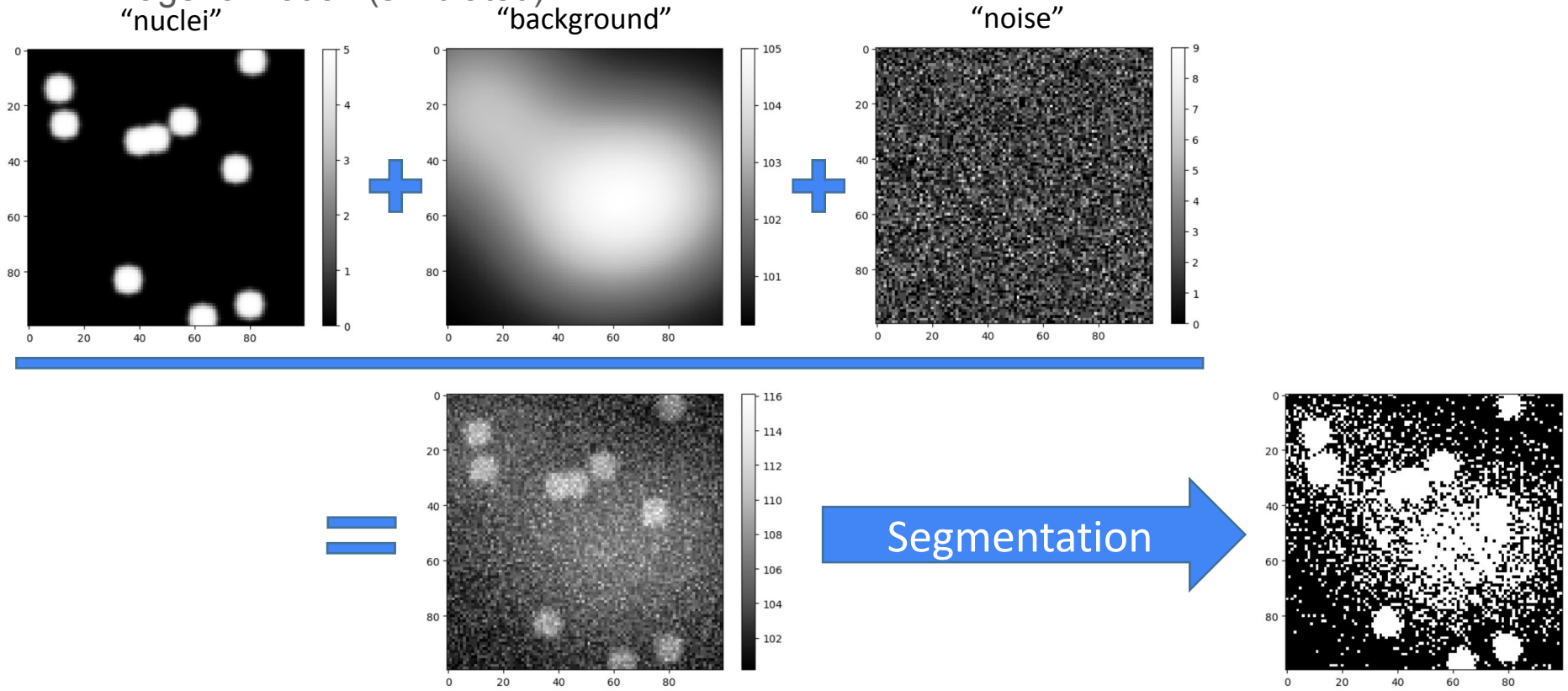
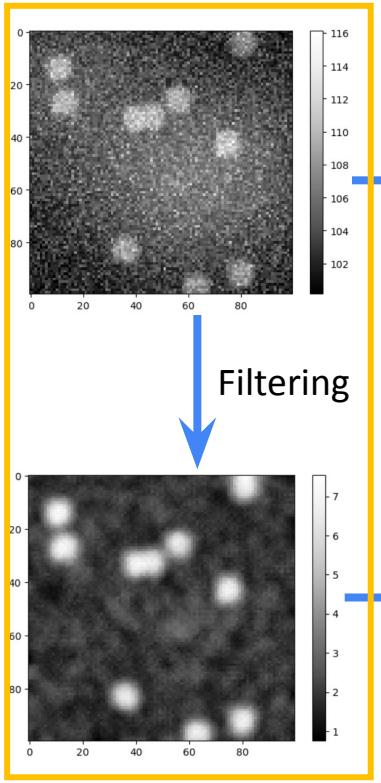
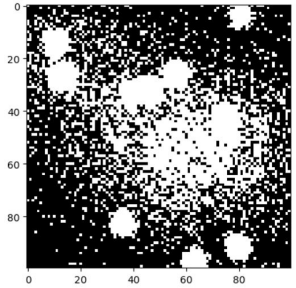


Image filtering

- We need to remove the noise to help the computer *interpreting* the image

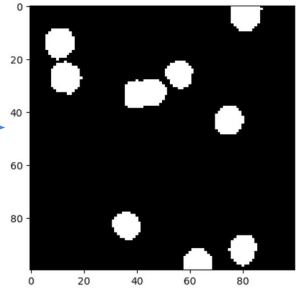


Segmentation



Filtering

Segmentation

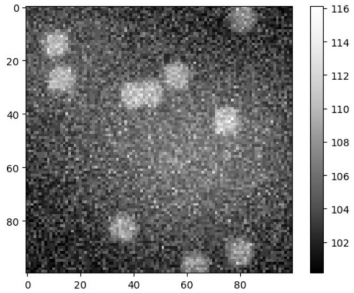


Oh no! I see thousands of tiny white objects!

Ok, it's just 9 objects.

Image filtering

- Attempt to invert / “undo” processes disturbing image quality



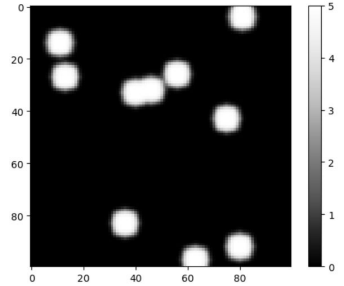
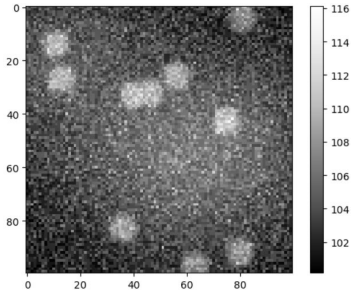
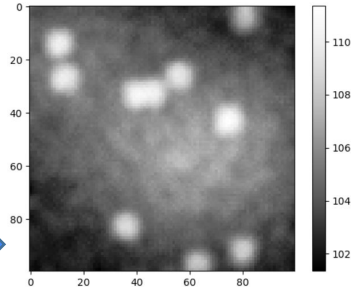


Image filtering

- Attempt to invert / “undo” processes disturbing image quality

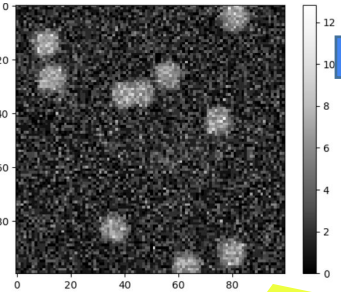


Denoise

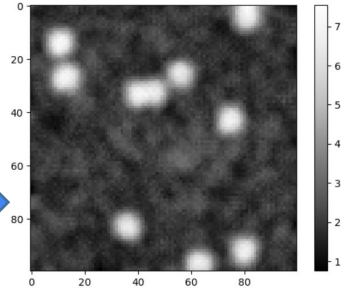


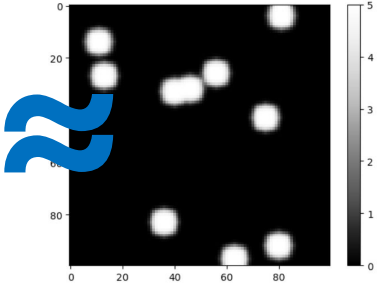
Low-pass filter: Allows low frequencies pass

Remove background



Denoise





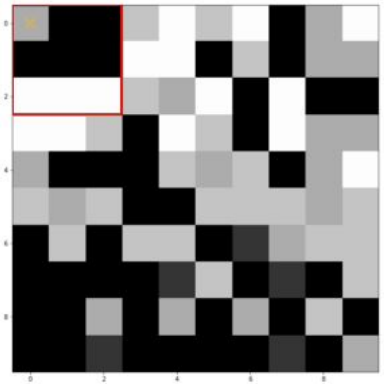
High-pass filter: Allows high frequencies pass

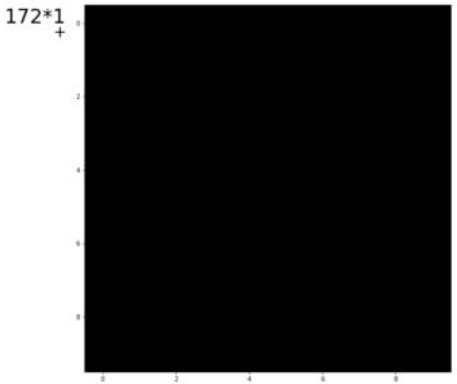
Band-pass filter: Allows a specific range of frequencies pass

https://github.com/BiAPoL/Bio-image_Analysis_with_Python/blob/49a787514a367829c3e0e1832f6cc533e96d549f/03_image_processing/simulated_dataset.i

Linear Filters

- *Linear filters* replace each pixel value with a weighted linear combination of surrounding pixels
- Filter *kernels* are matrices describing a linear filter
- This multiplication of surrounding pixels according to a matrix is called *convolution*





Mean filter, 3x3 kernel

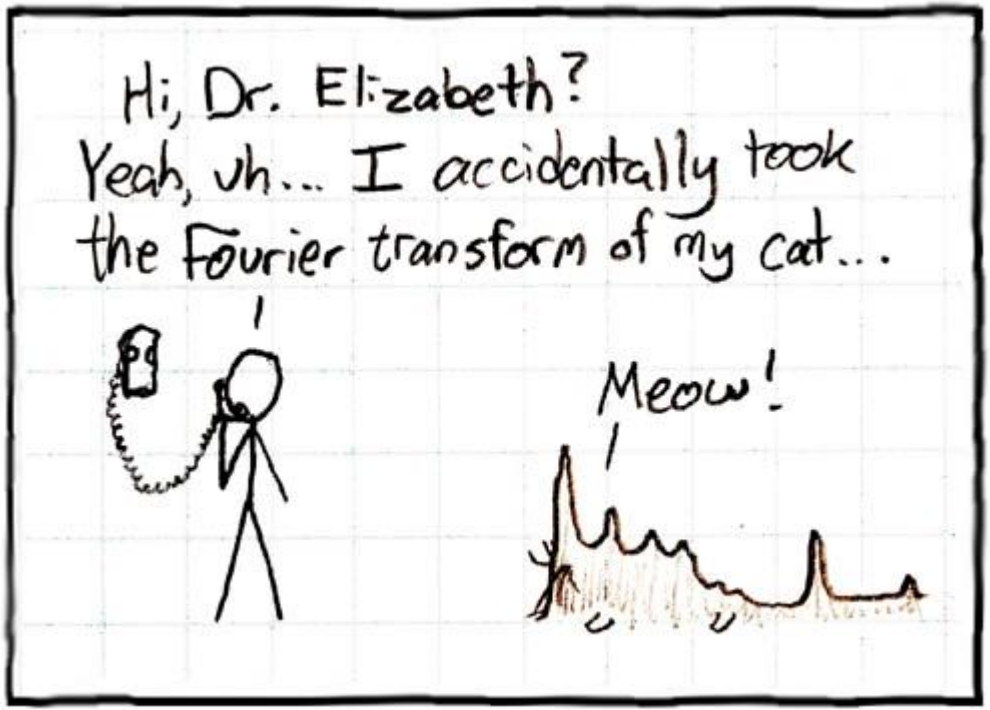
$$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

Animation source: Dominic Waithe, Oxford University
https://github.com/dwaithe/generalMacros/tree/master/convolution_ani

Linear Filters

Also we can apply them with Fourier transform

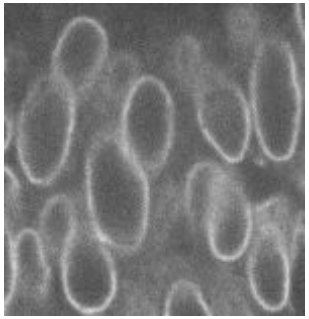
But… let's skip that for "safety reasons". And try to stick with using convolution in spatial domain.



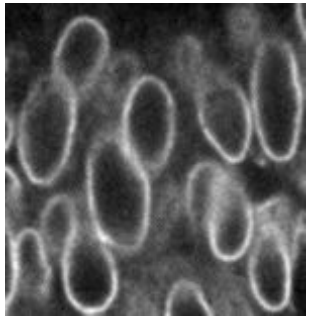
https://xkcd.com/26/

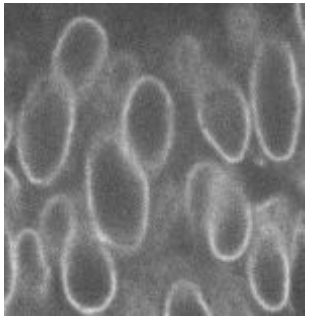
Linear filters

- Terminology:
  - “We convolve an image with a kernel.”
  - Convolution operator: *
    - Examples
      - Mean
      - Gaussian blur
      - Sobel-operator
      - Laplace-filter

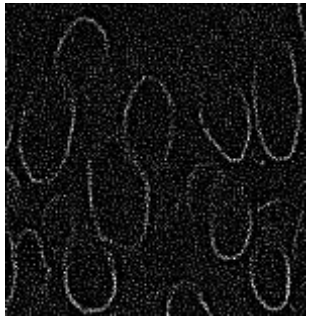


$$*\begin{bmatrix} 1 & 1 & 1 \\ 1 & 8 & 1 \\ 1 & 1 & 1 \end{bmatrix} =$$





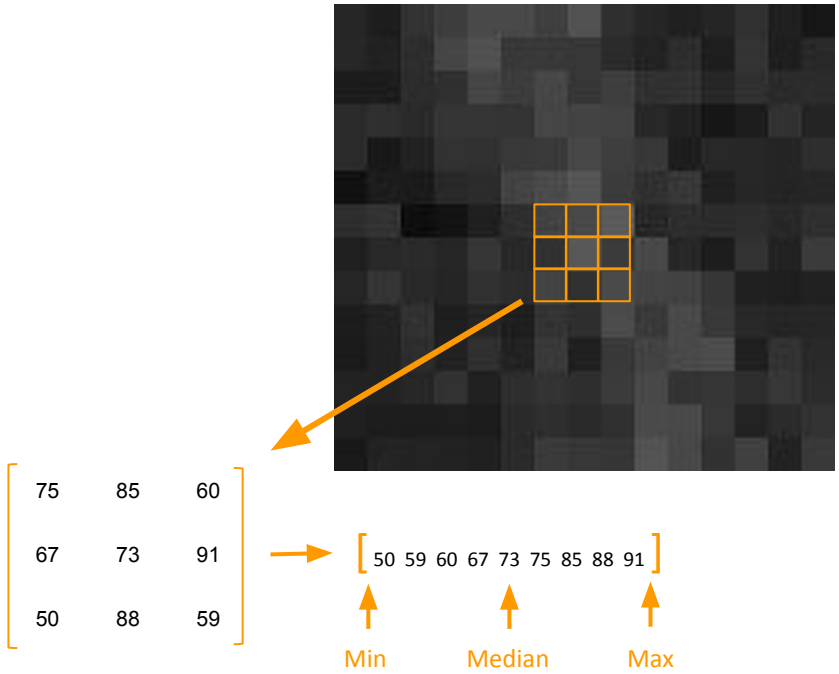
$$*\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} =$$



Nonlinear Filters

- Non linear filters also replace pixel value inside as rolling window but using a non-linear function.

- Examples: order statistics filters
  - Min
  - Median
  - Max
  - Variance
  - Standard deviation



Noise removal

- Gaussian filter
- Median filter (computationally expensive)

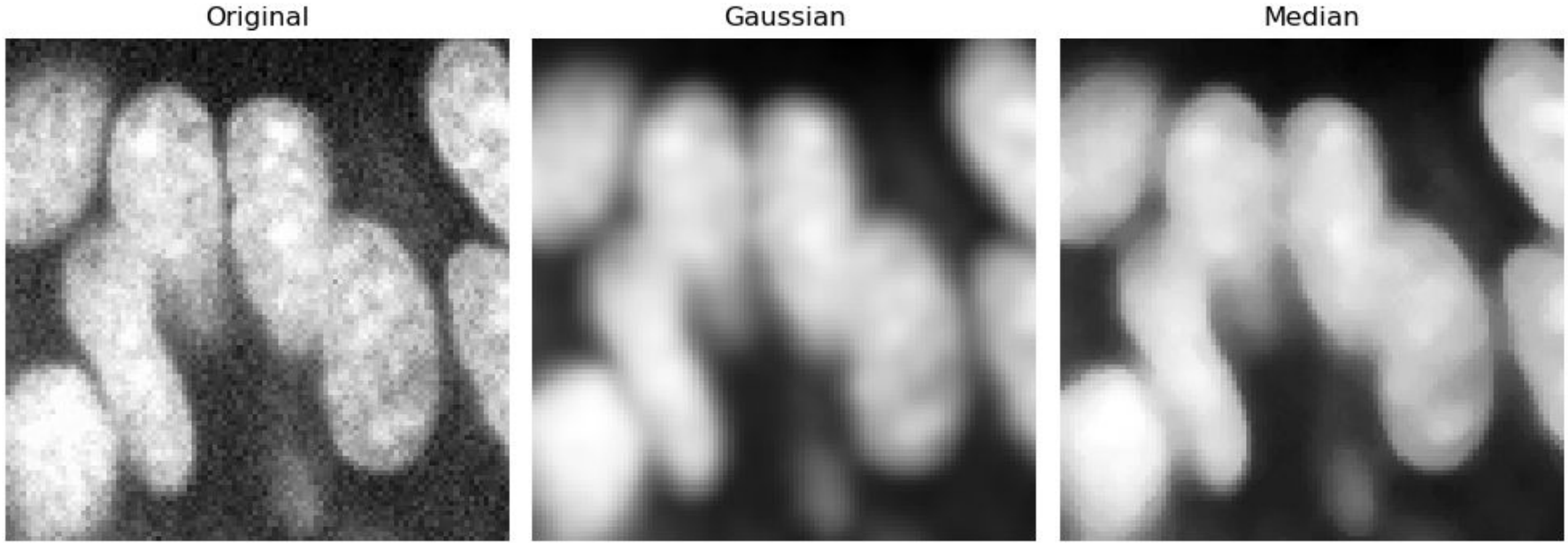
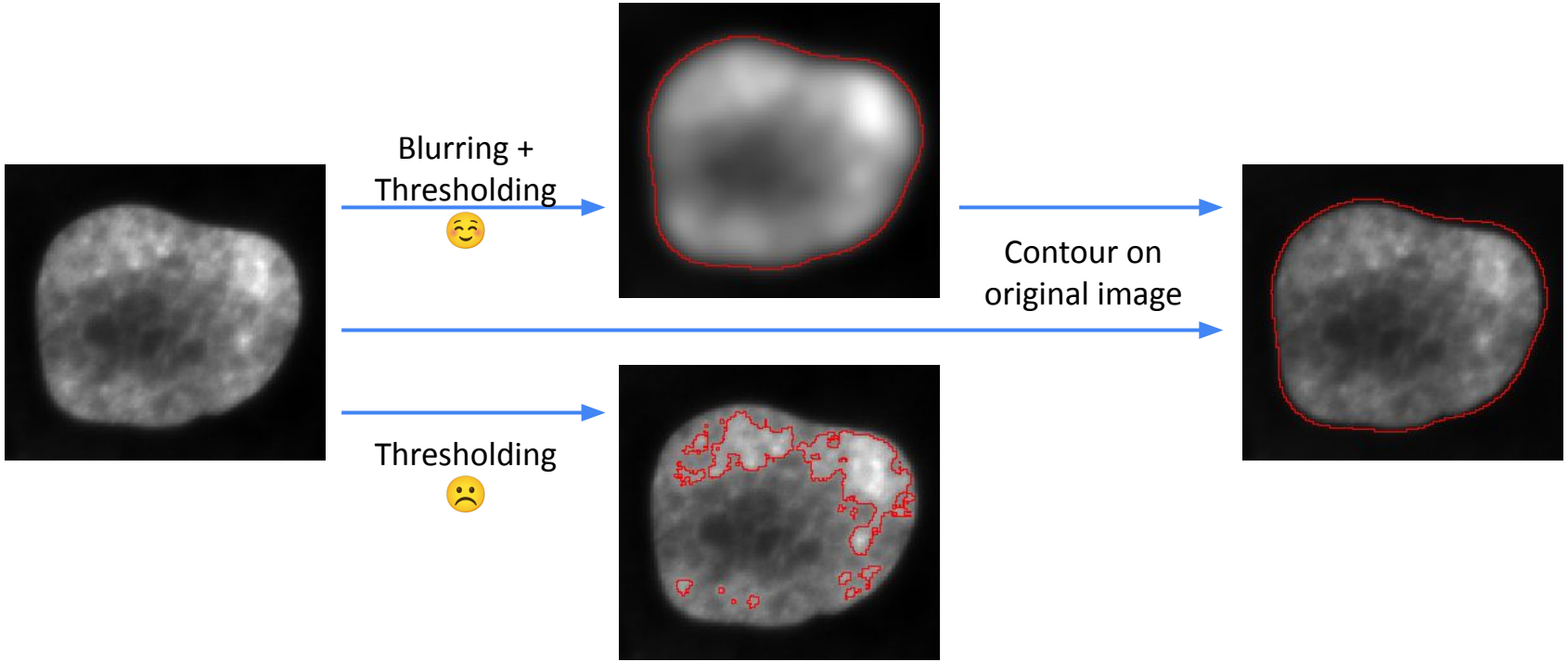


Image source: Mauricio Rocha Martins (Norden/Myers lab, MPI CBG)

Filtering for improving thresholding results

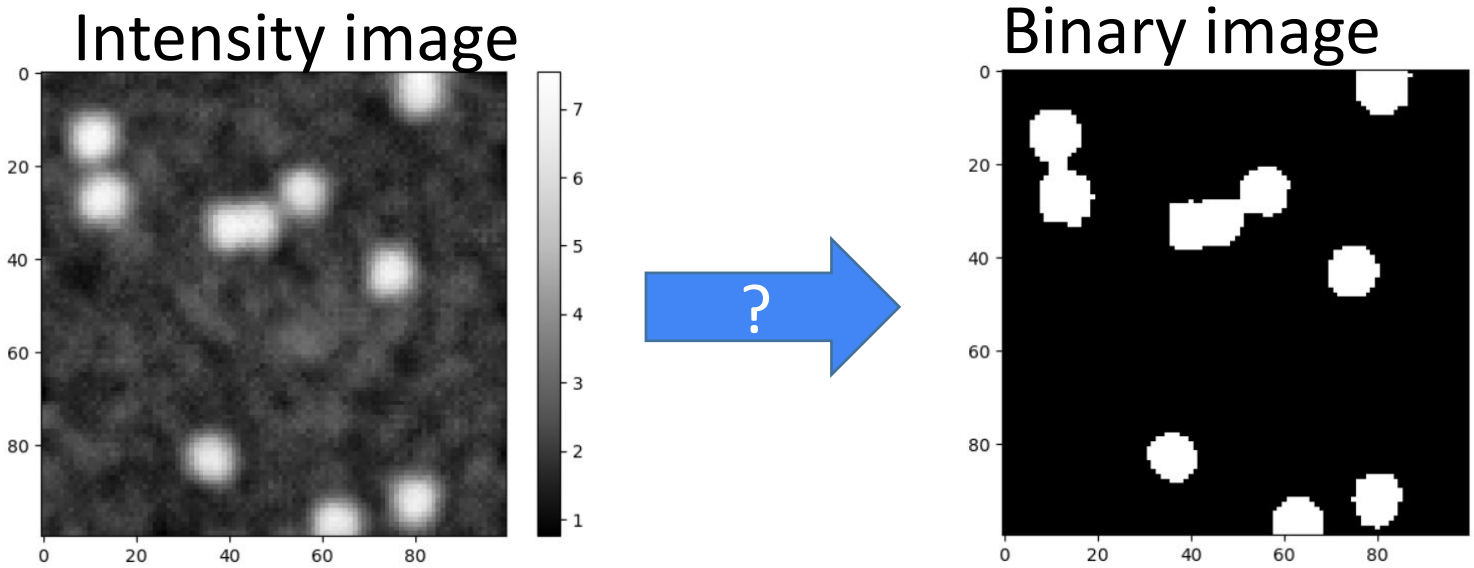
- In case thresholding algorithms outline the wrong structure, blurring in advance may help.
- However: **Do not** continue processing the blurred image, continue with the original!



Segmentation / binarization

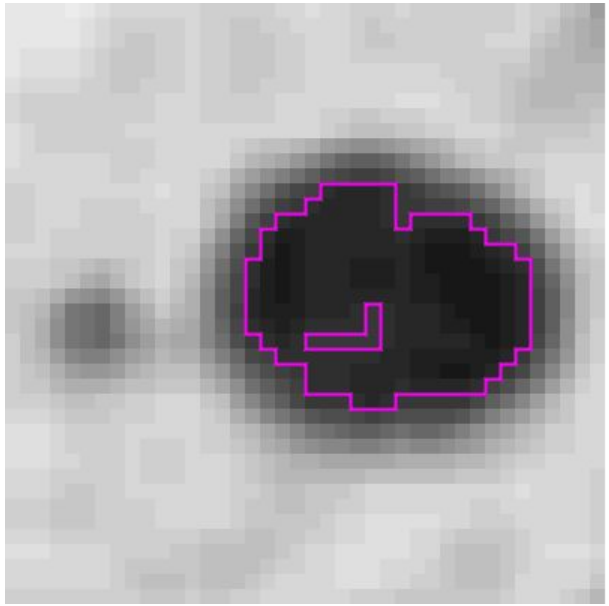
Thresholding

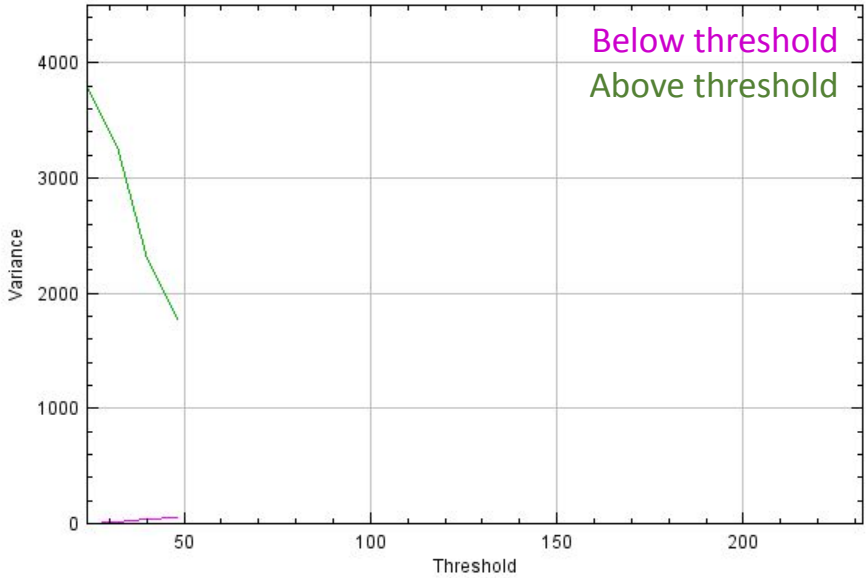
- Very basic and yet efficient segmentation technique
- Histogram based, to determine an intensity threshold
- Not state-of-the-art in many fields (anymore)



Thresholding: Otsu's method

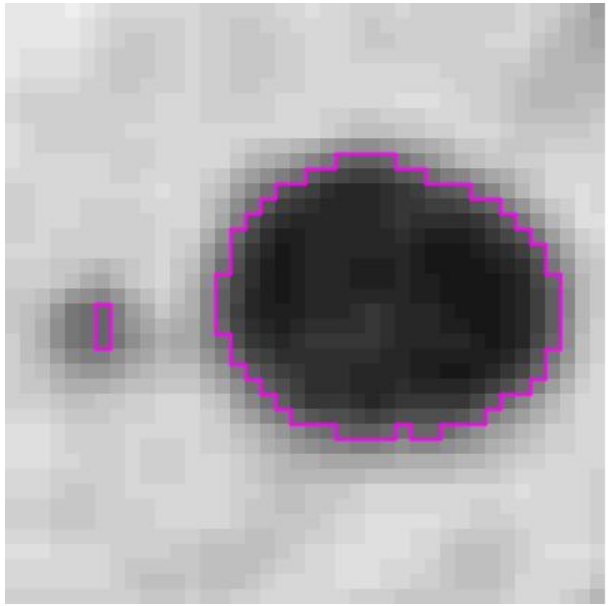
- Searching for a threshold where the variance in both classes (above/below threshold) becomes minimal.

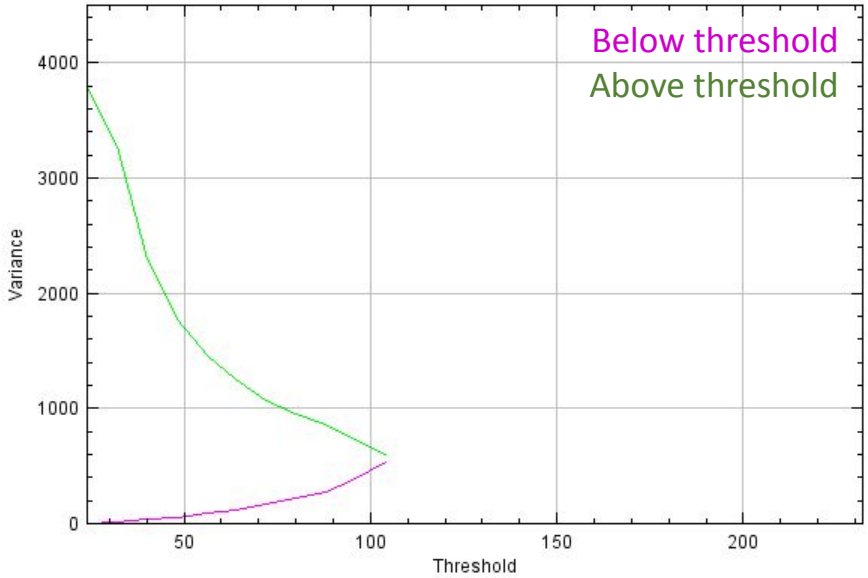




Thresholding: Otsu's method

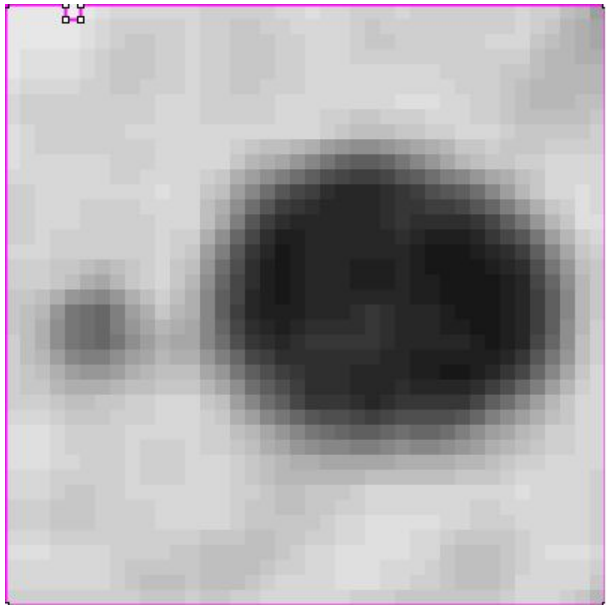
- Searching for a threshold where the variance in both classes (above/below threshold) becomes minimal.

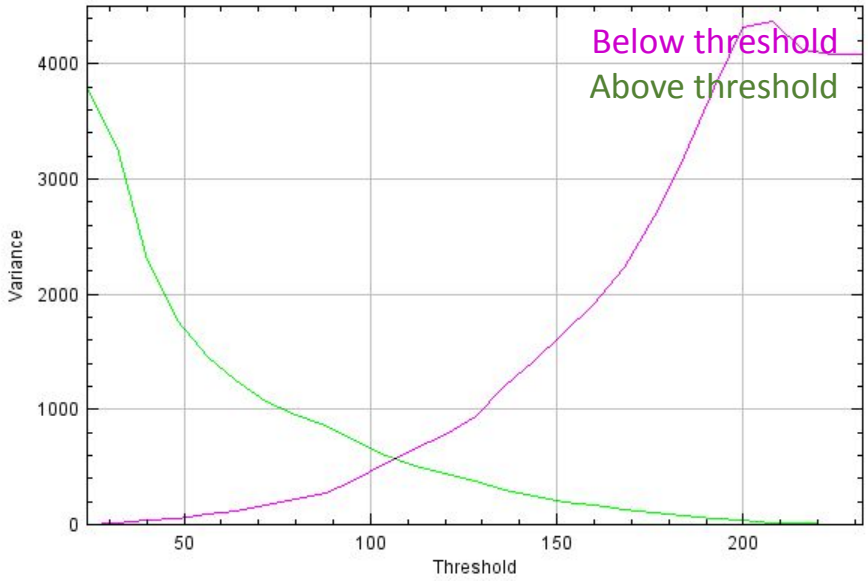




Thresholding: Otsu's method

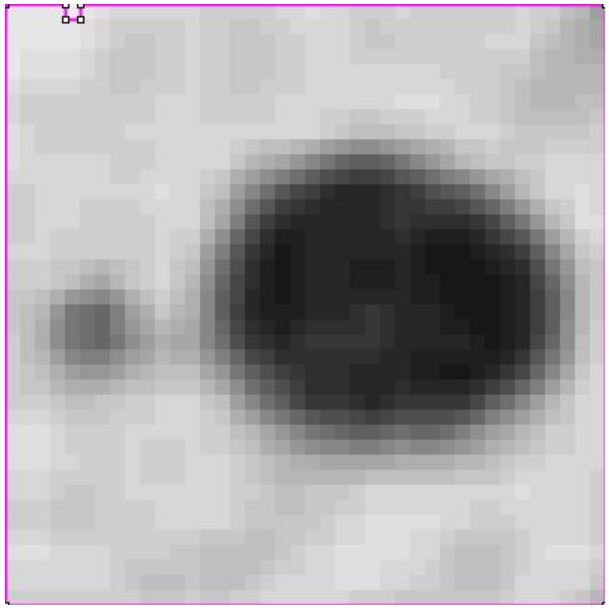
- Searching for a threshold where the variance in both classes (above/below threshold) becomes minimal.

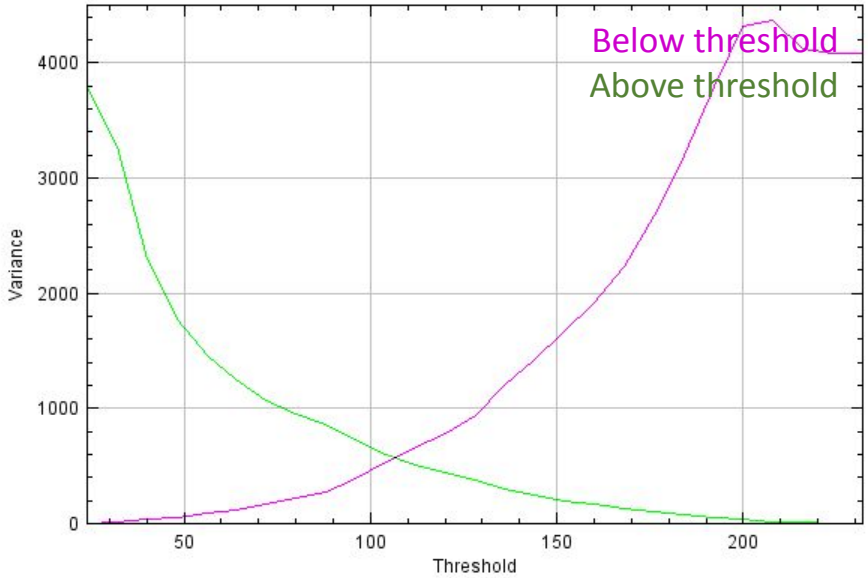




Thresholding: Otsu's method

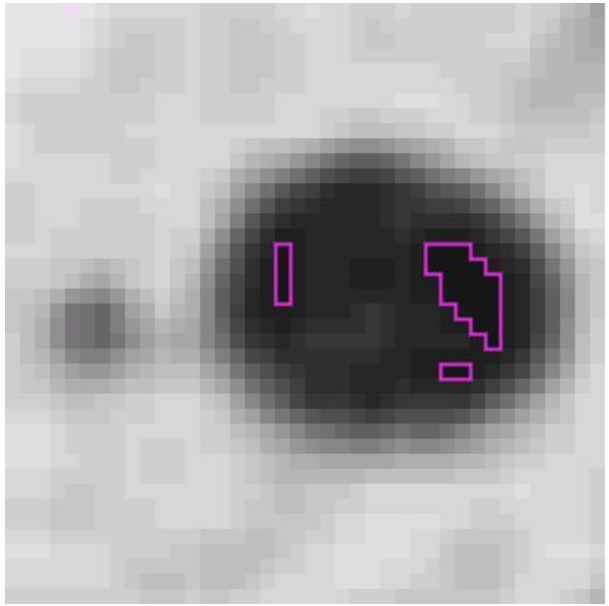
- Searching for a threshold where the variance in both classes (above/below threshold) becomes minimal.
- Weighted (!) sum variance

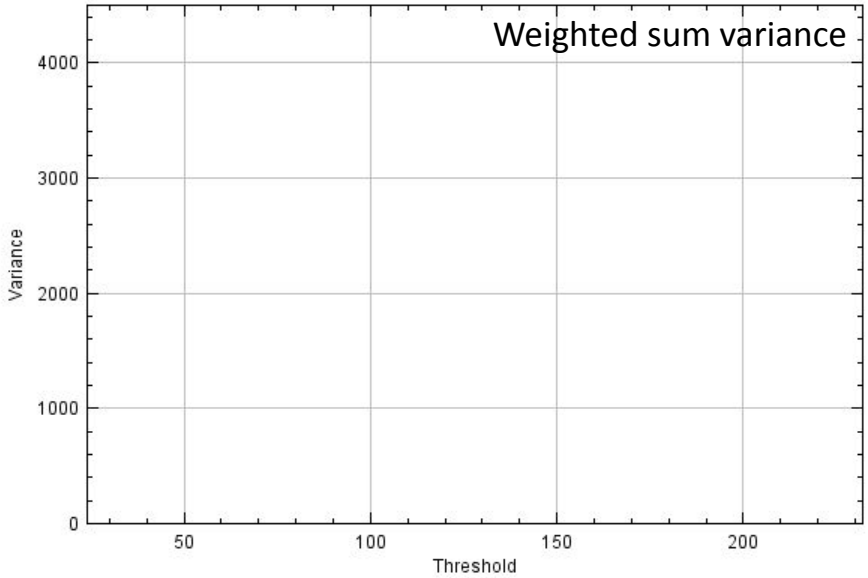




Thresholding: Otsu's method

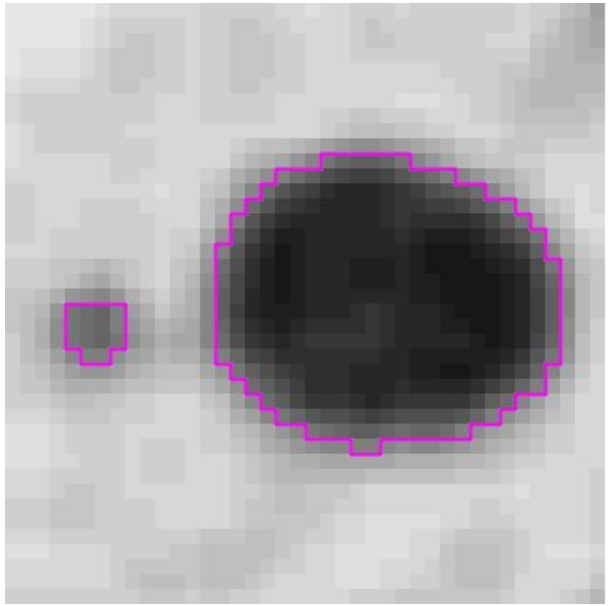
- Searching for a threshold where the variance in both classes (above/below threshold) becomes minimal.
- Weighted (!) sum variance

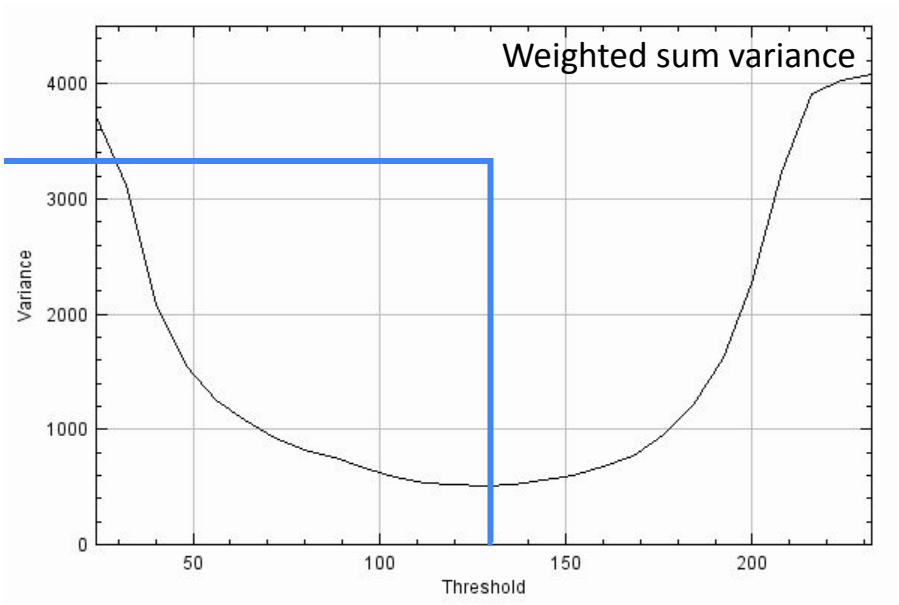




Thresholding: Otsu's method

- Searching for a threshold where the variance in both classes (above/below threshold) becomes minimal.
- Weighted (!) sum variance





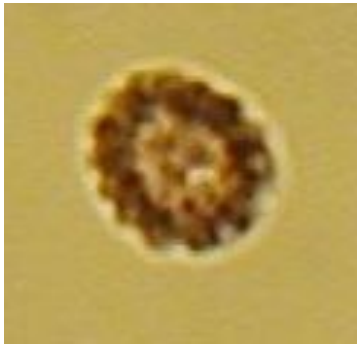
See also: http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.html

Morphological operations

Refining masks

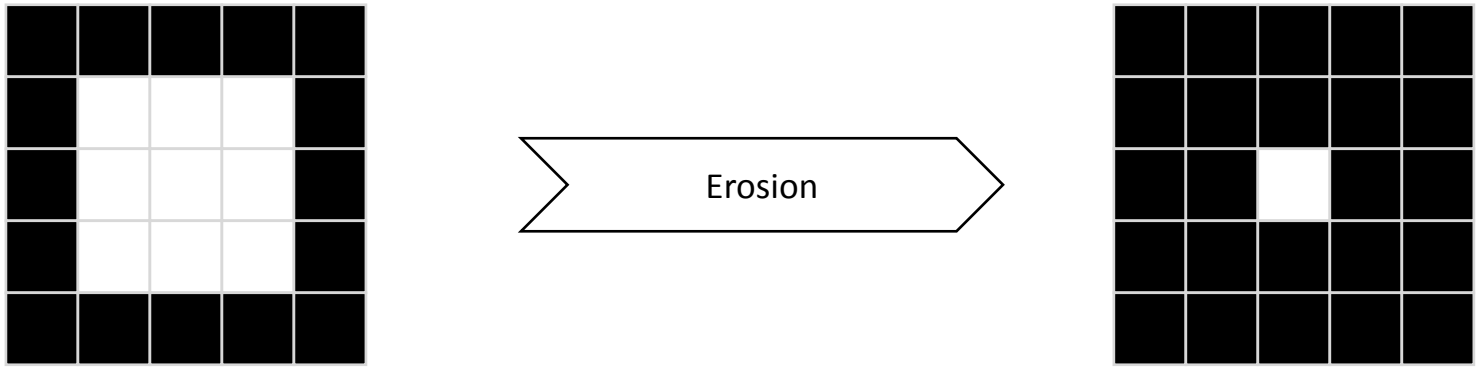
- Binary mask images may not be perfect immediately after thresholding.
- There are ways of refining them

Thresholding → Closing → Opening



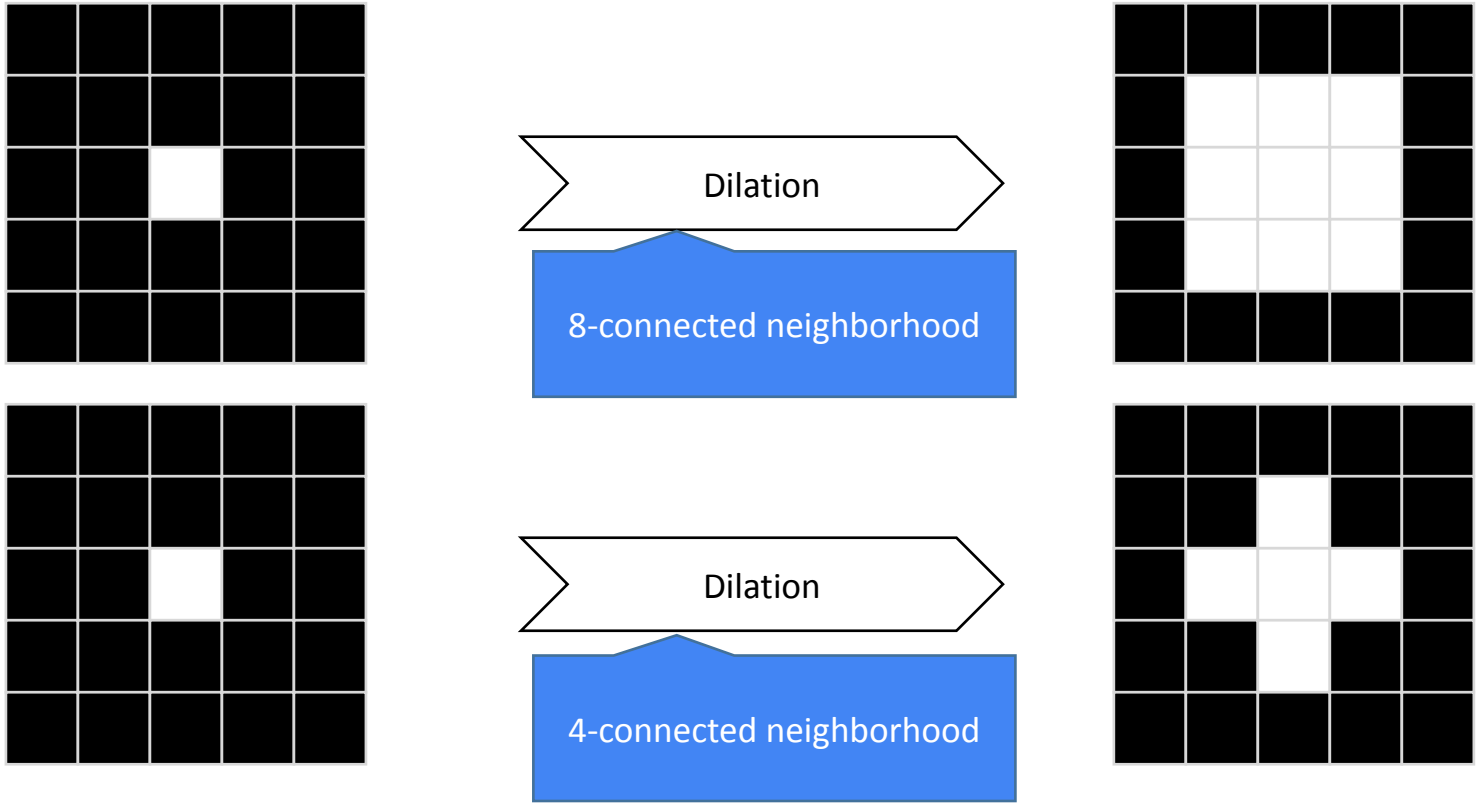
Erosion

- Erosion: Every pixel with at least one black neighbor becomes black.



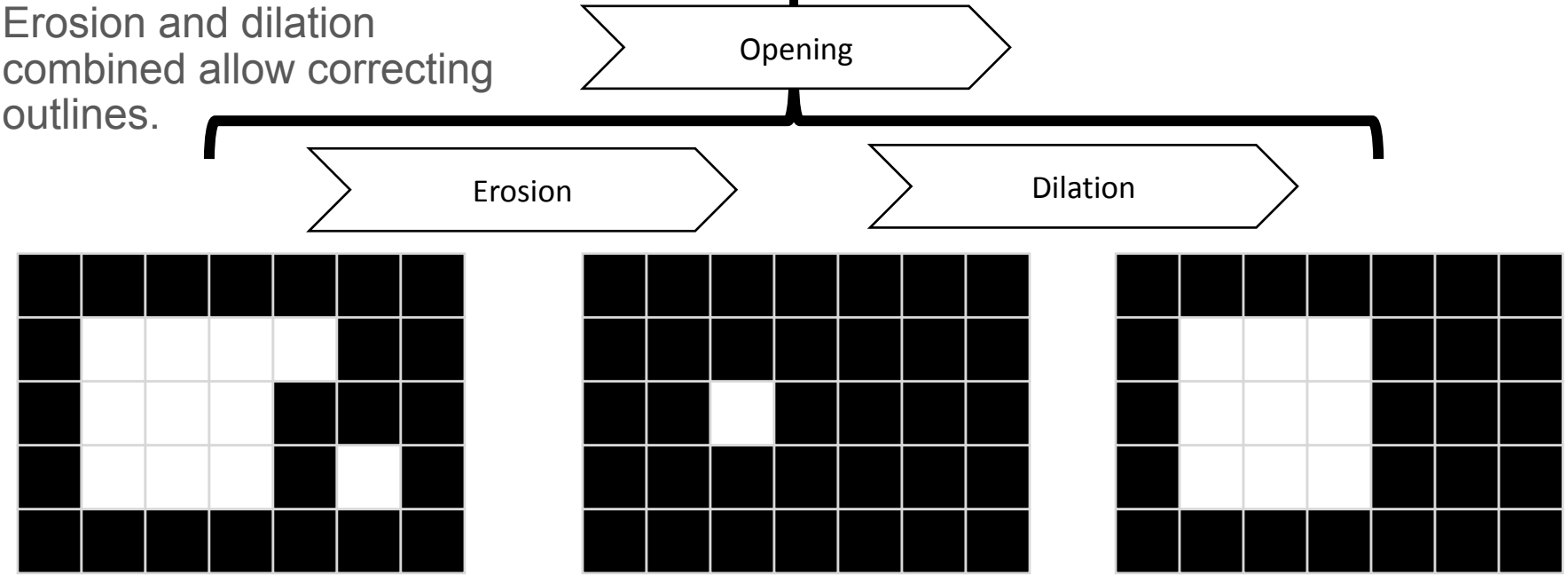
Dilation

- Dilation: Every pixel with at least one white neighbor becomes white.



Opening

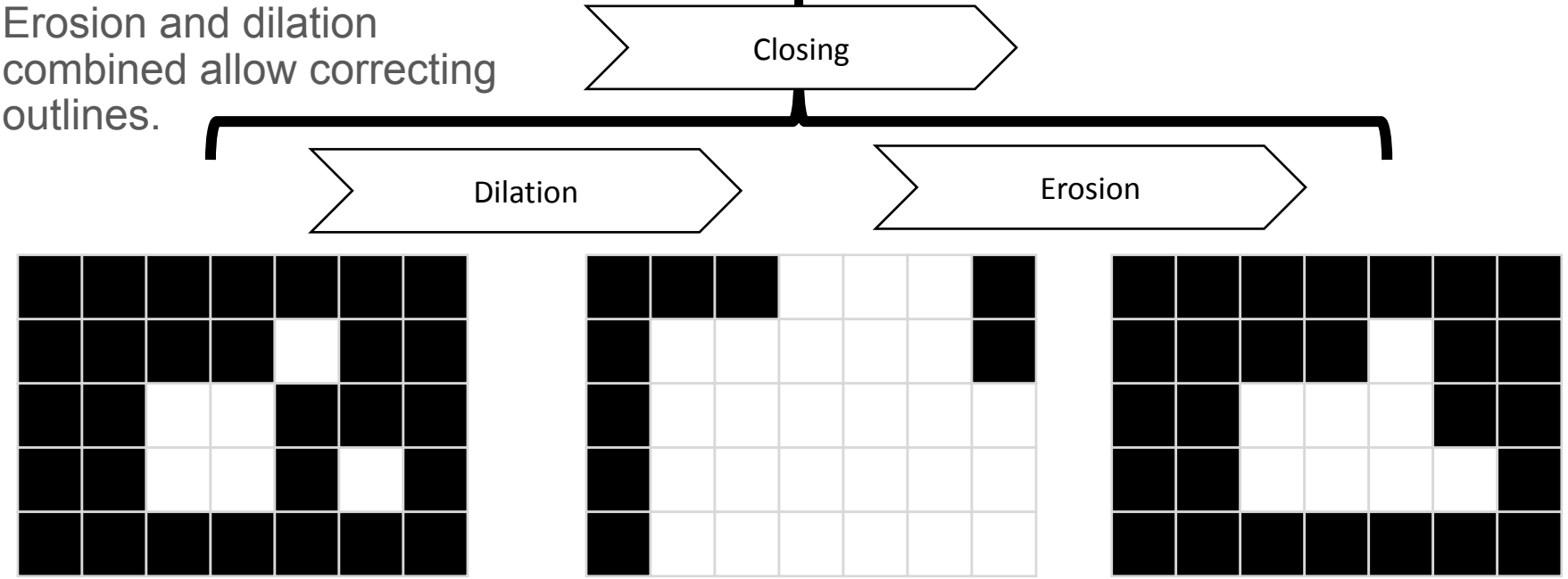
Erosion and dilation combined allow correcting outlines.



- It can separate white (high intensity) structures that are weakly connected
- It may erase small white structures

Closing

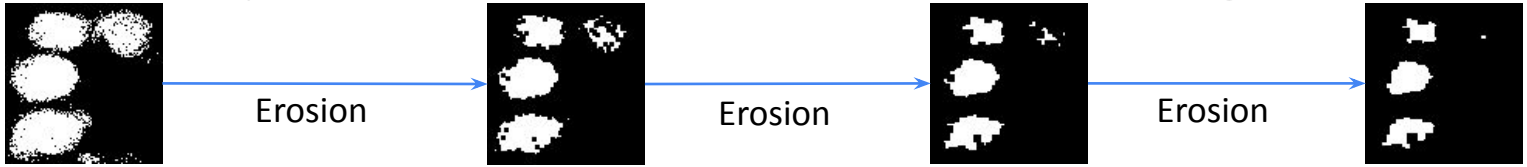
Erosion and dilation combined allow correcting outlines.



- It can connect white (high intensity) structures that are nearby
- It may close small holes inside structures

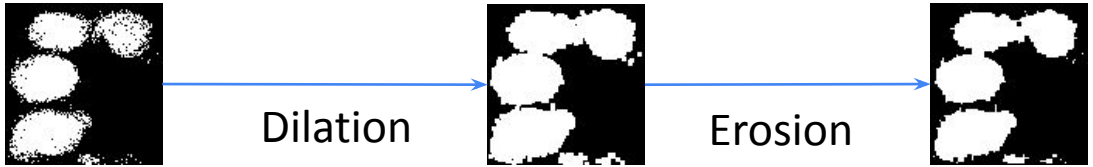
Chaining erosion and dilation

Erosion: Set all pixels to black which have at least one black neighbor.



Dilation: Set all pixels to white which have at least one white neighbor.

Closing: Dilation + Erosion

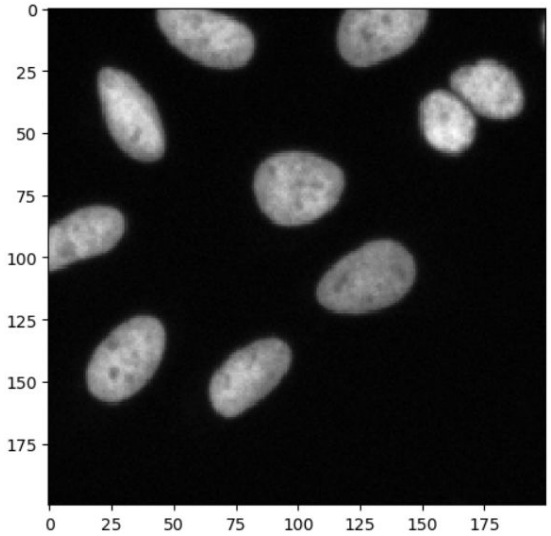


Opening: Erosion + Dilation

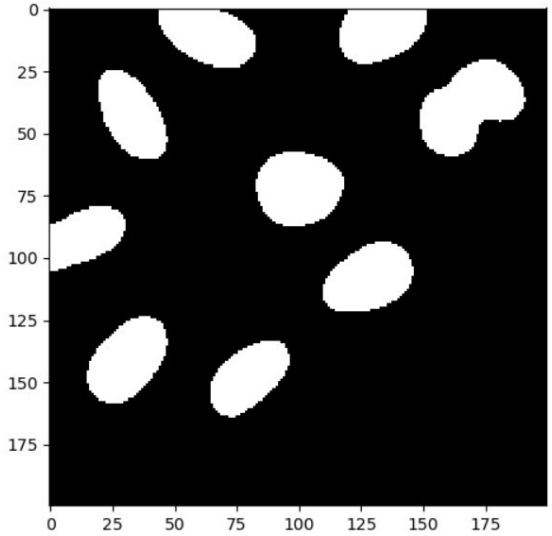
Slide adapted from Haase, Lombardot, Scientific Computing Facility, MPI CBG, licensed CC BY-NC 4.0

Terminology

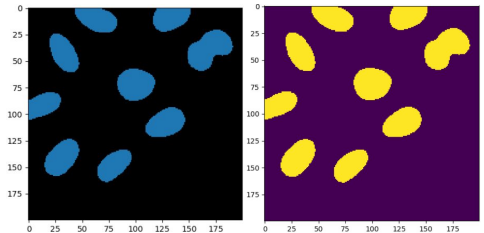
Intensity image



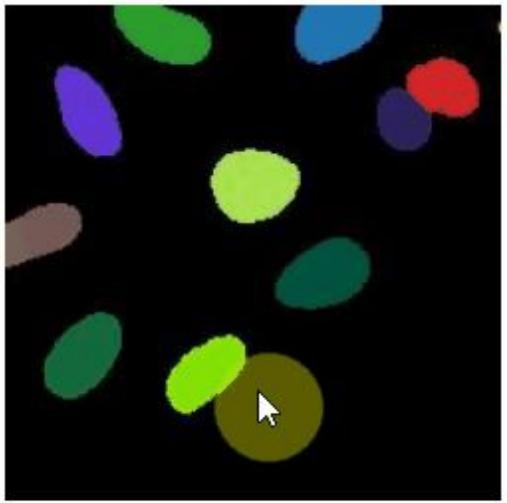
Binary image



No matter how they are displayed



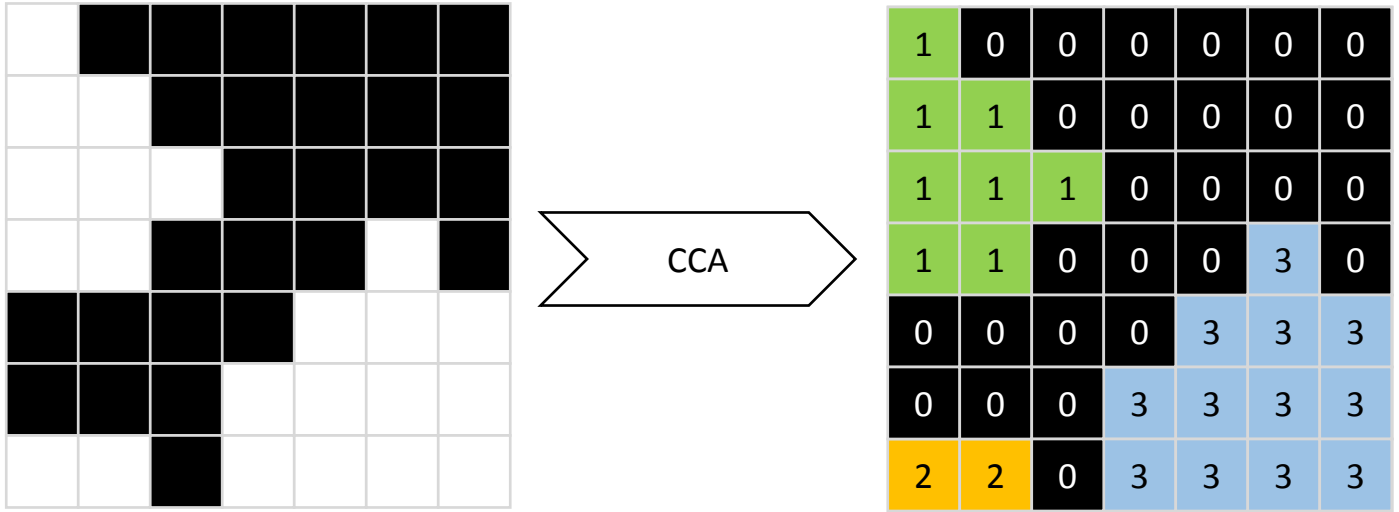
Label image



[y=152, x=92] = 0

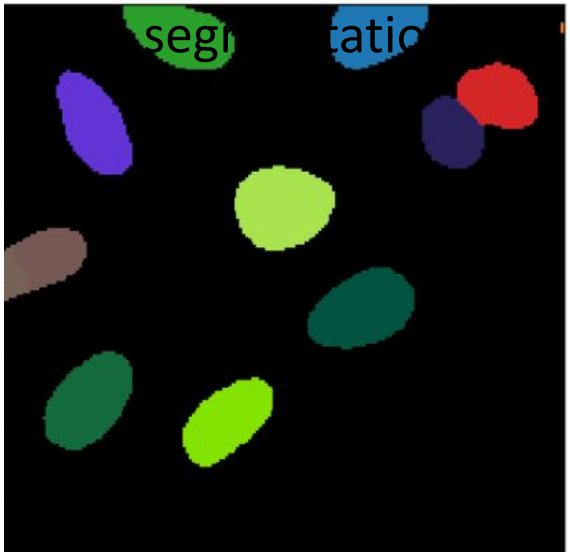
Connected component labelling

- In order to allow the computer differentiating objects, connected component analysis (CCA) is used to mark pixels belonging to different objects with different numbers
- Background pixels are marked with 0.
- The maximum intensity of a labelled map corresponds to the number of objects.



Terminology

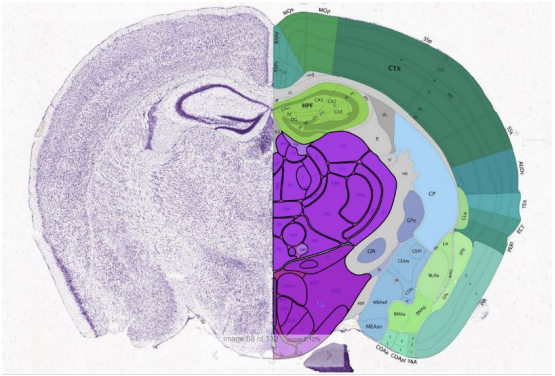
Instance

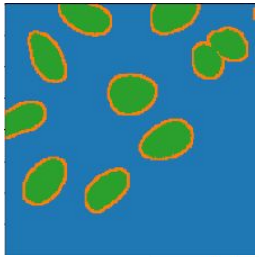


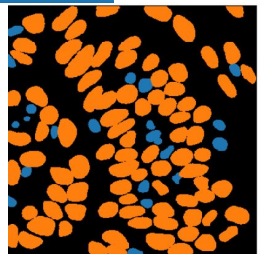
Instances:

- Cells, nuclei, cats, dogs, cars, trees

Semantic segmentation







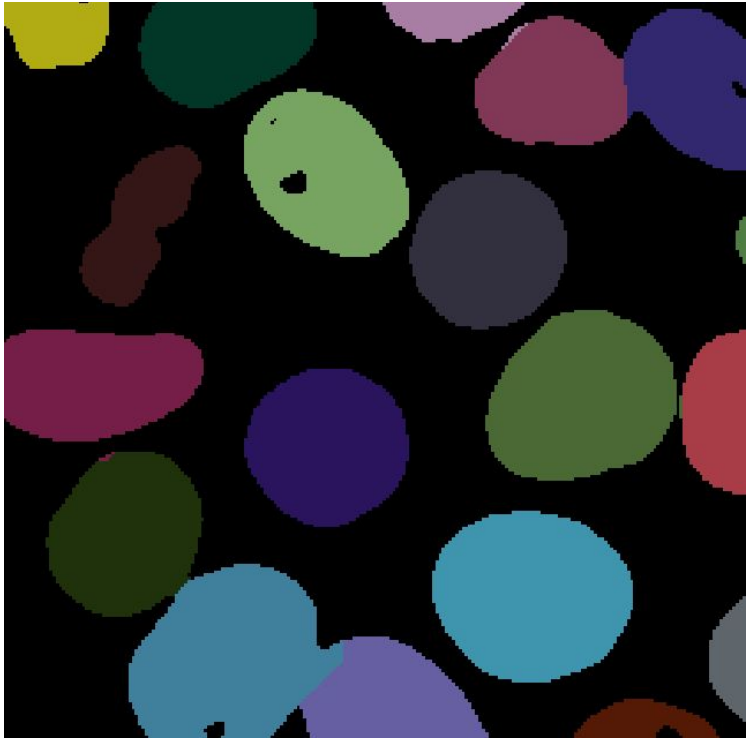
Regions:

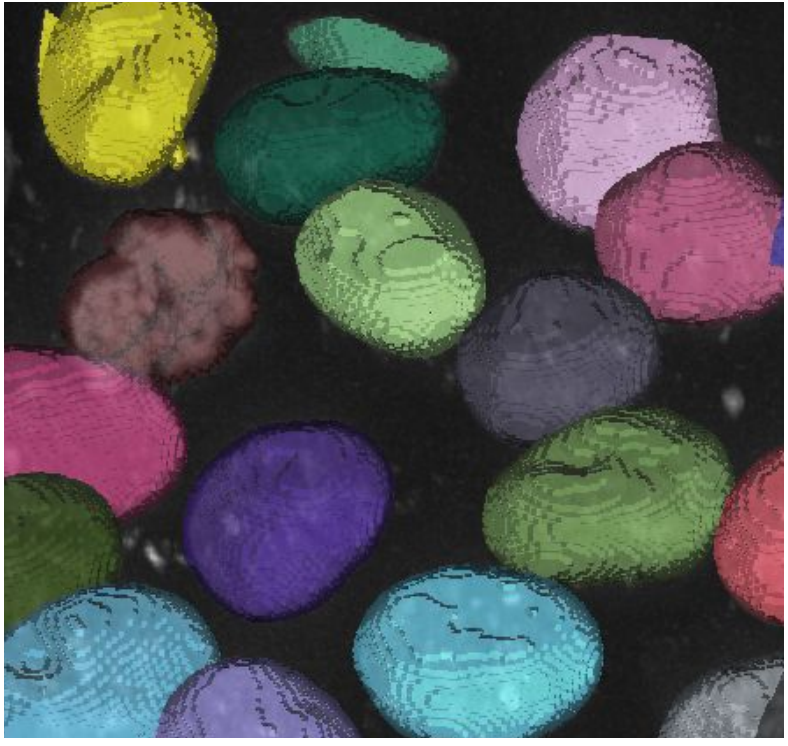
- Anatomical, geographical
- All pixels belonging to the same type of object have the same value

Source: Allen Mouse Brain Atlas [dataset]. Available from mouse.brain-map.org. Allen Institute for Brain Science (2011).

Motivation: Surface reconstruction

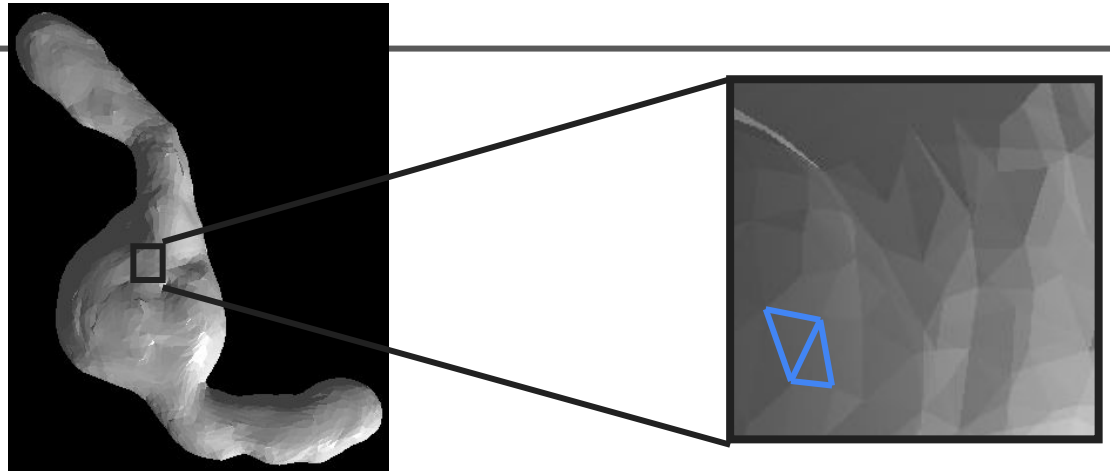
Pixel and voxel borders introduce artifacts, potentially problematic for measurements, e.g. surface area

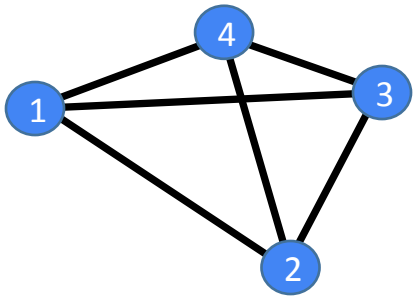




Surface meshes

Points on a surfaces connected by triangles forma a surface mesh





"Vertices" / points

Point x	Point y	Point z
x_1	y_1	z_1
X_2	Y_2	Z_2
X_3	Y_3	Z_3
X_4	Y_4	Z_4
…	…	…

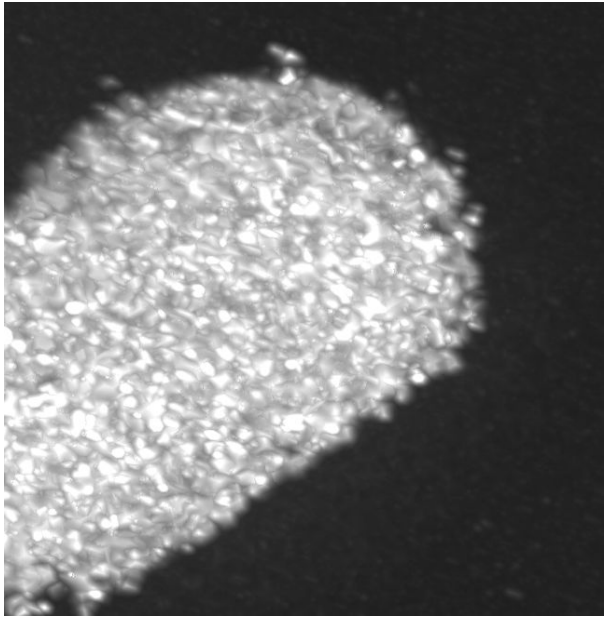
+

"Faces" / Triangles

Point 1	Point 2	Point 3
1	2	3
1	2	4
2	3	4
1	3	4

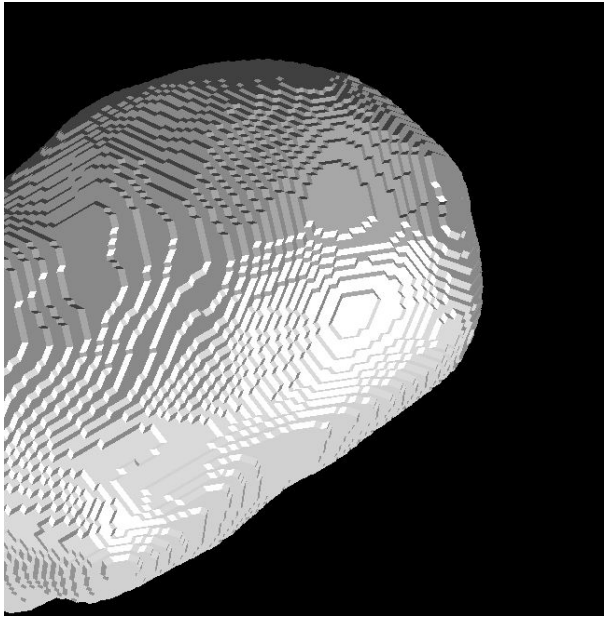
Data derived from AV Luque and JV Veenvliet (2023) licensed CC-BY : https://zenodo.org/record/7603081

Surface reconstruction



3D image of nuclei

Gaussian filtered

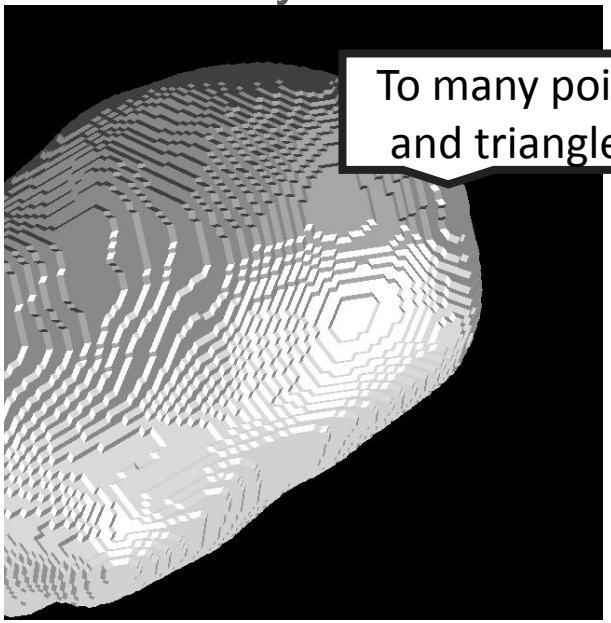


Binary 3D image
(visualized as surface mesh)

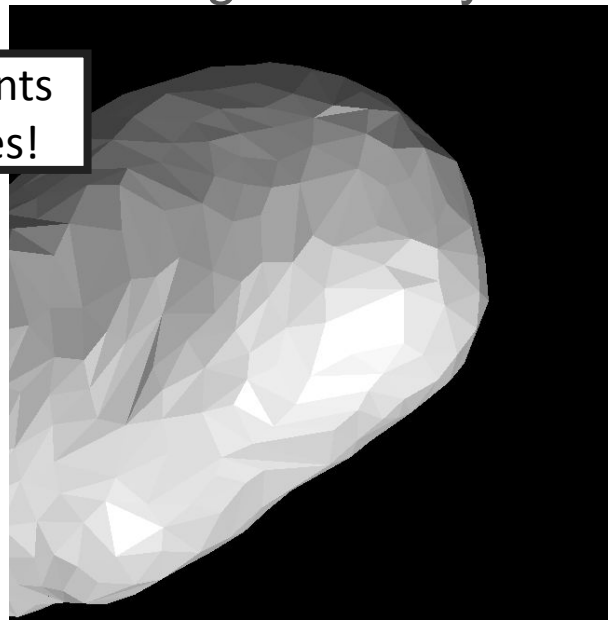
Data derived from AV Lugue and JV Veenvliet (2023)
licensed CC-BY : https://zenodo.org/record/7603081

Surface post-processing

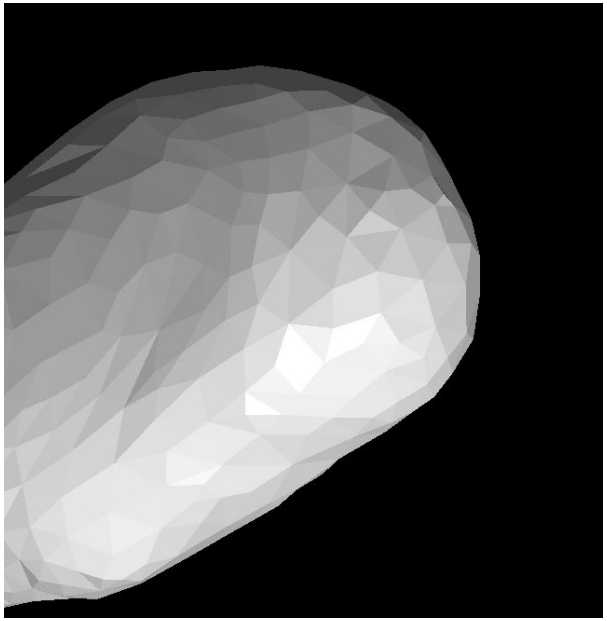
Necessary to better match biological reality.



Marching cubes result



Simplified mesh
(less points, locally averaged)



Smoothed mesh
(position locally planarized)

Data derived from AV Lugue and JV Veenvliet (2023)
licensed CC-BY : https://zenodo.org/record/7603081

Surface post-processing

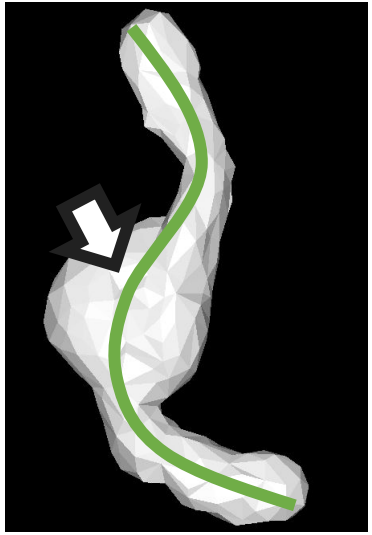
- Every processing step has consequences errors of later measurements
- Depends on desired measurement

Number of small concave regions

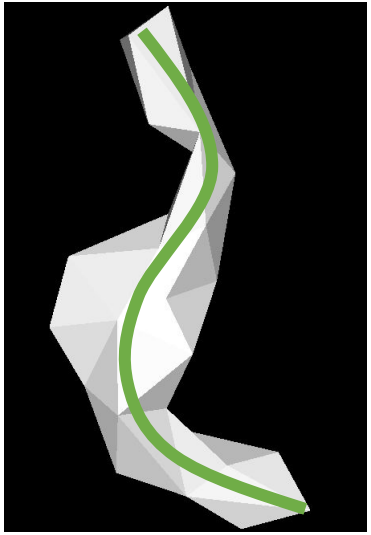
Total length

Surface mesh

Simplified by factor 0.5



Simplified by factor 0.05

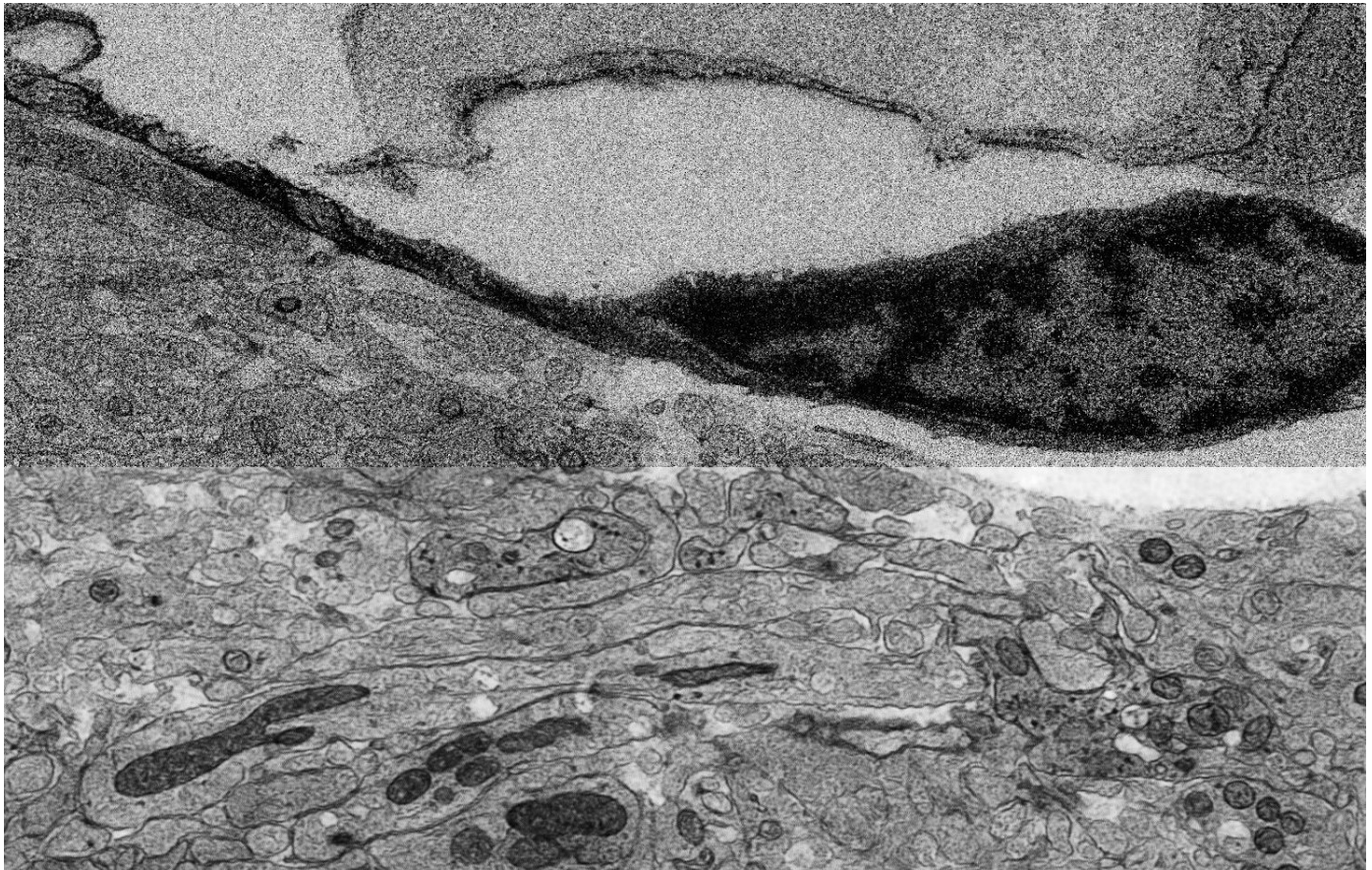


Simplified by factor 0.01

Data derived from AV Luque and JV Veenvliet (2023)
licensed CC-BY : https://zenodo.org/record/7603081

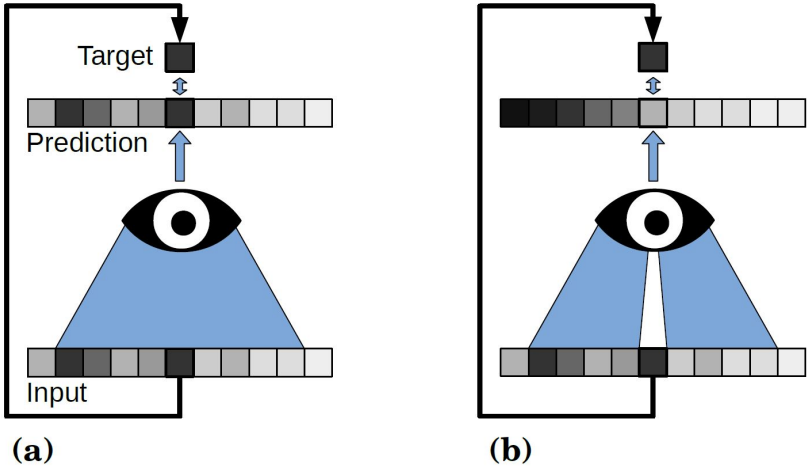
Bonus part

Deep Learning Denoising - Noise2Void



Deep Learning Denoising - Noise2Void

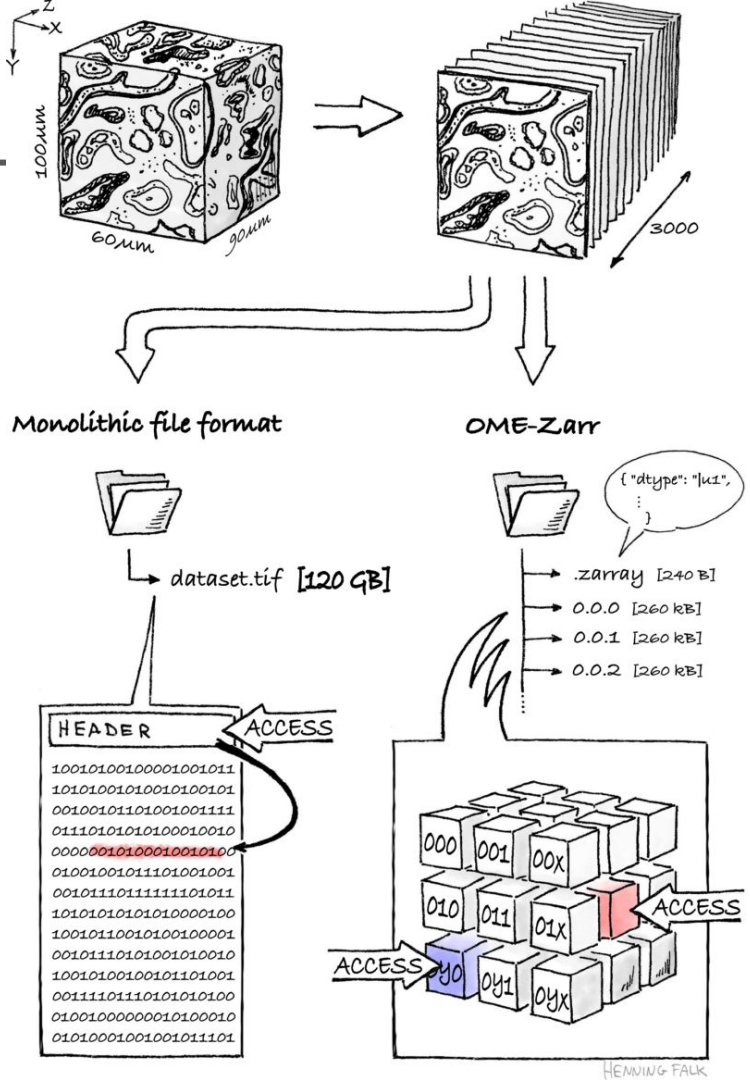
- Assumption: noise changes but image information remains.
  - The signal is statistically dependent on itself in the image.
  - Noise is statistically independent of the signal.
- Creating blind spots in the training image **(b)** will help the model recognize dependencies in the image.
- **Limitation**: Noise can be dependent based on digitalisation!



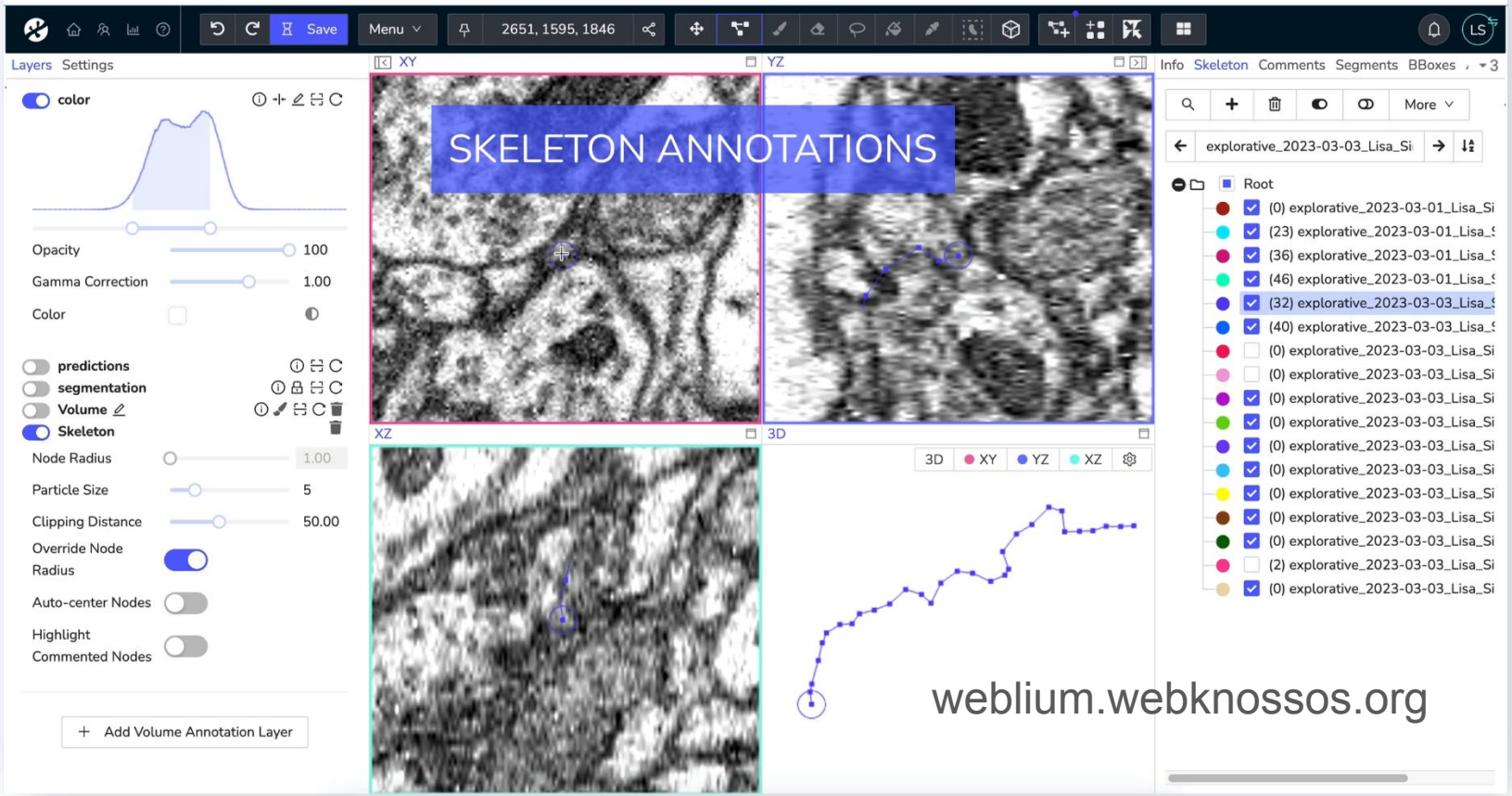
Zarr file format is structured into small, 3-dimensional chunks that can be directly accessed from the Zarr-directory. Hence, only the part of interest needs to be loaded instead of the full binary file.

This feature allows data to be accessed in sufficiently small pieces so that streaming the data of interest over a network instead of transferring the whole file becomes possible.

Falk, H. zarr-developers/zarr-illustrations-falk-2022 | Zenodo, 2022. URL https://doi.org/10.5281/zenodo.7037367

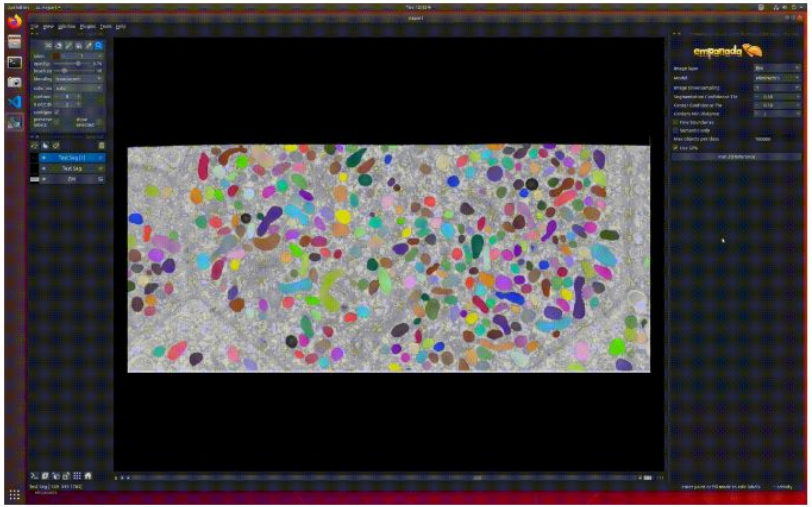


Visualize, share, and annotate your large 3D images online



empanada-napari

The empanada-napari plugin is built to democratize deep learning image segmentation for researchers in electron microscopy (EM). It ships with MitoNet, a generalist model for the instance segmentation of mitochondria. There are also tools to quickly build and annotate training datasets, train generic panoptic segmentation models, finetune existing models, and scalably run inference on 2D or 3D data. To make segmentation model training faster and more robust, CEM pre-trained weights are used by default. These weights were trained using an unsupervised learning algorithm on over 1.5 million EM images from hundreds of unique EM datasets making them remarkably general.



References

- Haase, R. (2021). Quantitative Bio-image Analysis. F1000 Research Limited. http://doi.org/10.7490/F1000RESEARCH.1118789.1
- Haase, R. (2023). Quantitative Bio-image Data Analysis in the age of Machine Learning. F1000 Research Limited. http://doi.org/10.7490/F1000RESEARCH.1119447.1
- Haase, R. (2023). Introduction to bio-image analysis I. F1000 Research Limited. http://doi.org/10.7490/F1000RESEARCH.1119427.1
- Haase, R. (2021). Interactive Data Visualization 101 with Fiji & Friends. F1000 Research Limited. http://doi.org/10.7490/F1000RESEARCH.1118801.1

BioImage Analysis Newsletter

CzechBIAS