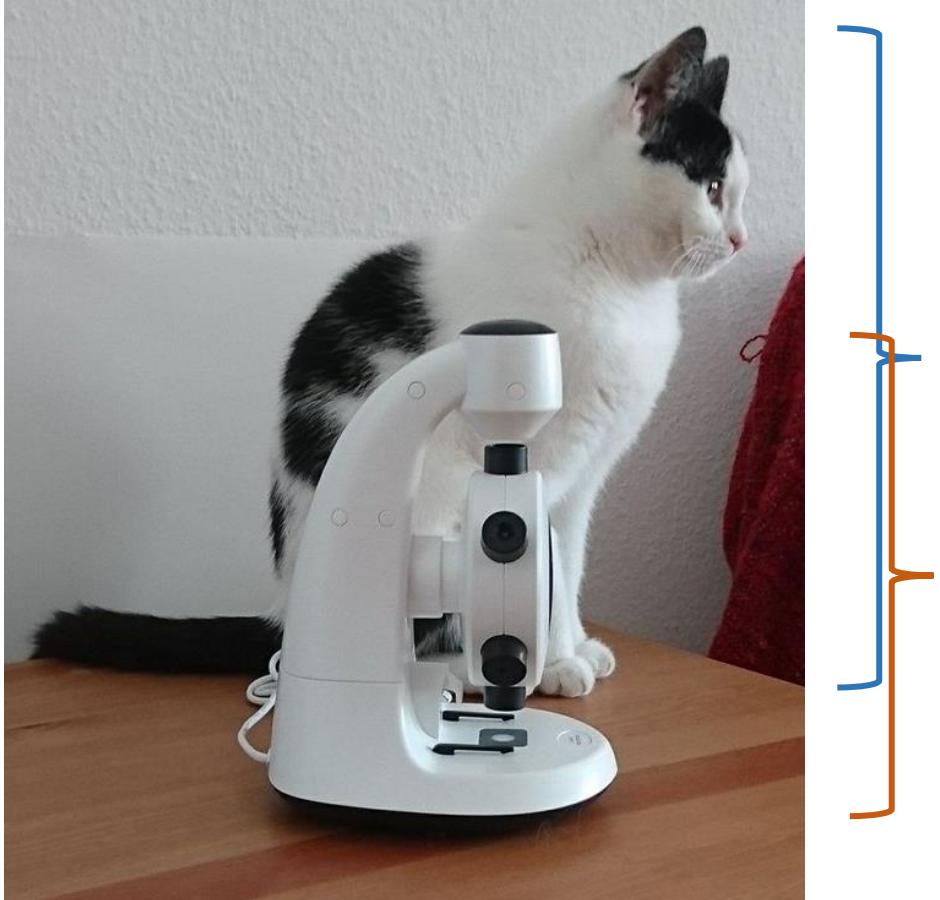


# Introduction to Bio-Image Analysis

Robert Haase

Reusing some materials made by Beth Cimini,  
Anne Carpenter & colleagues (Broad Institute)

- Deriving quantitative information from images of biological samples taken with microscopes



cat height = 1.5 x microscope height

- Deriving quantitative information from images of biological samples taken with microscopes + visualization

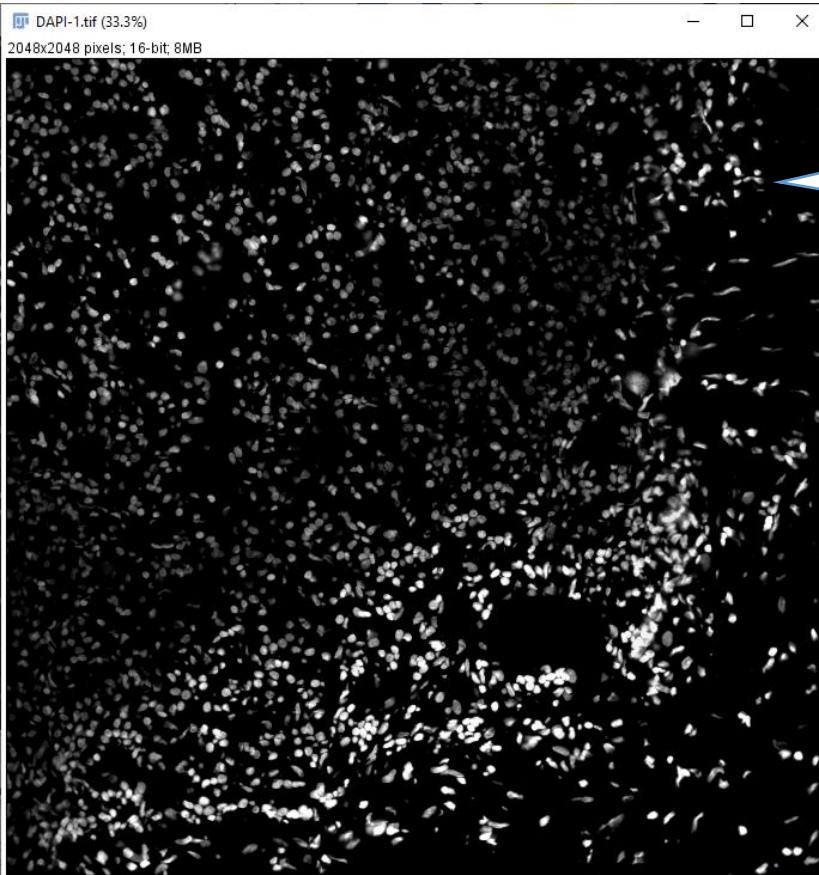
The figure displays three sequential steps in a quantitative bio-image analysis workflow using the napari software:

- Panel 1: Segmentation**  
Shows a grayscale image of Drosophila nuclei with a mitotic wave. The napari interface includes controls for opacity, contrast limits, gamma, colormap (gray), blending (translucent), and interpolation (nearest). The image is labeled "Drosophila, mitotic wave". A text box in the foreground asks: "How is the nuclei marker distributed?"
- Panel 2: Pixel count map**  
Shows the segmented nuclei as colored regions. The napari interface shows the "Result of connected" layer selected. A text box in the foreground asks: "How many nuclei are there?"
- Panel 3: Final output**  
Shows a final "Pixel count map" where each nucleus is assigned a color based on its size. The napari interface shows the "Result of label\_pixel\_count\_map [172 - 13]" layer selected. A text box in the foreground asks: "How large are these nuclei?"

The right side of the interface shows a sidebar with various processing tools: Noise removal, Background removal, Filter, Combine, Transform, Projection, Binarize, Label, Label processing, Label measurement, Map, and Mesh.

# Objective bio-image analysis

- Measurements should be objective, not influenced by human interpretation



Nuclei in this image are ...  
... more dense than in this image.

Use automation for less subjective analysis.

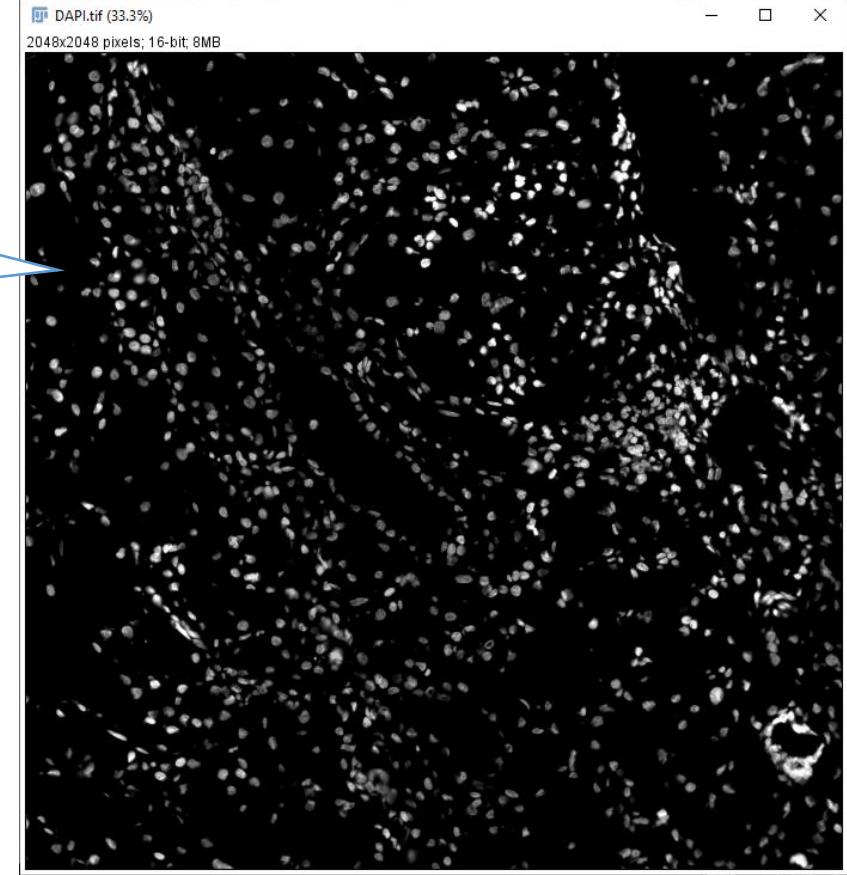
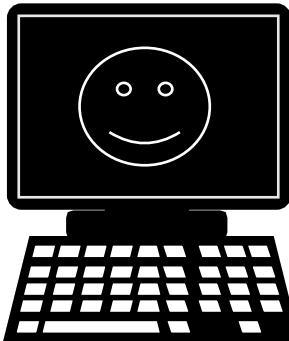
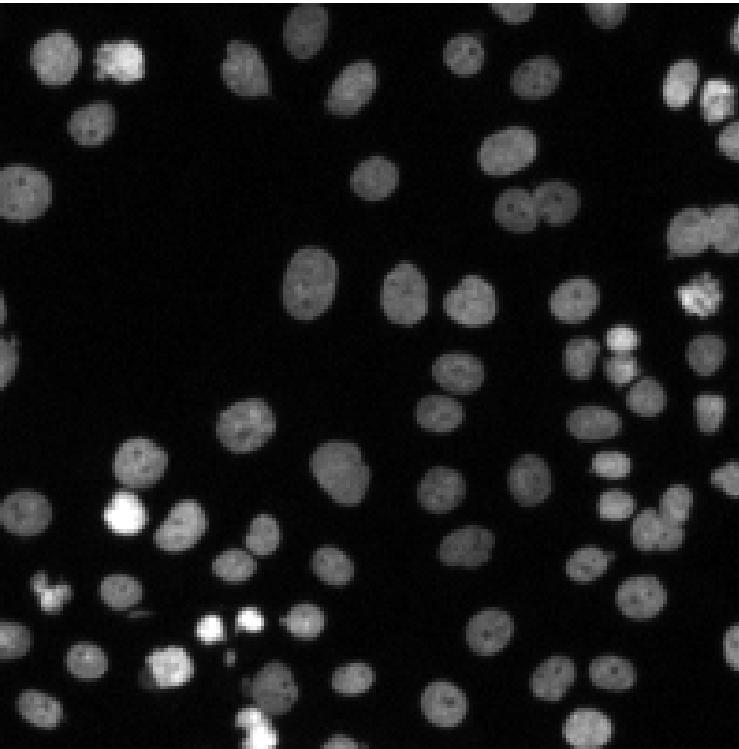


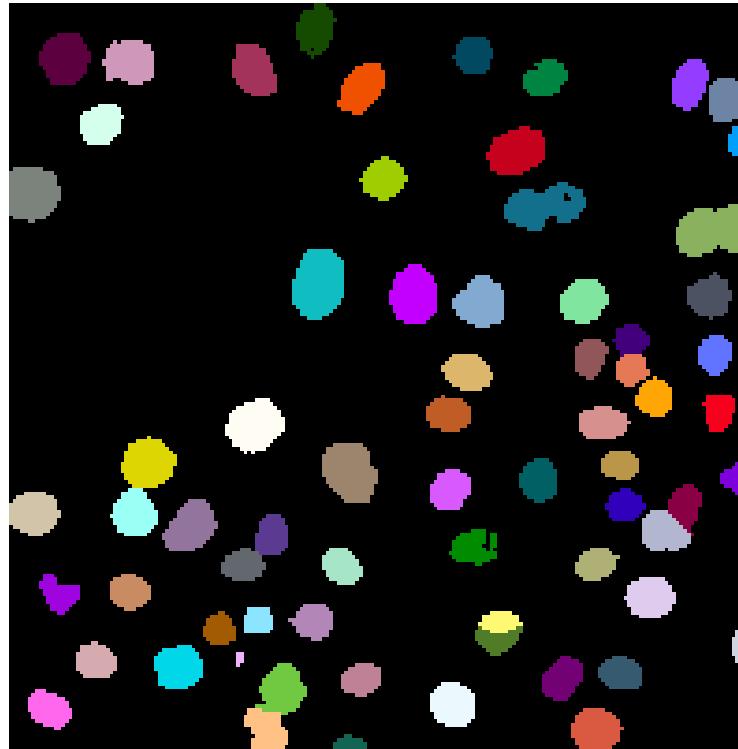
Image data source: Pascual-Reguant, Anna. (2021). Immunofluorescence staining of a human kidney (#2, peri-tumor area) obtained by MELC [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.4434462> licensed CC-BY 4.0

- Algorithms must be reliable (trustworthy). Visualization helps gaining trust in automated methods.

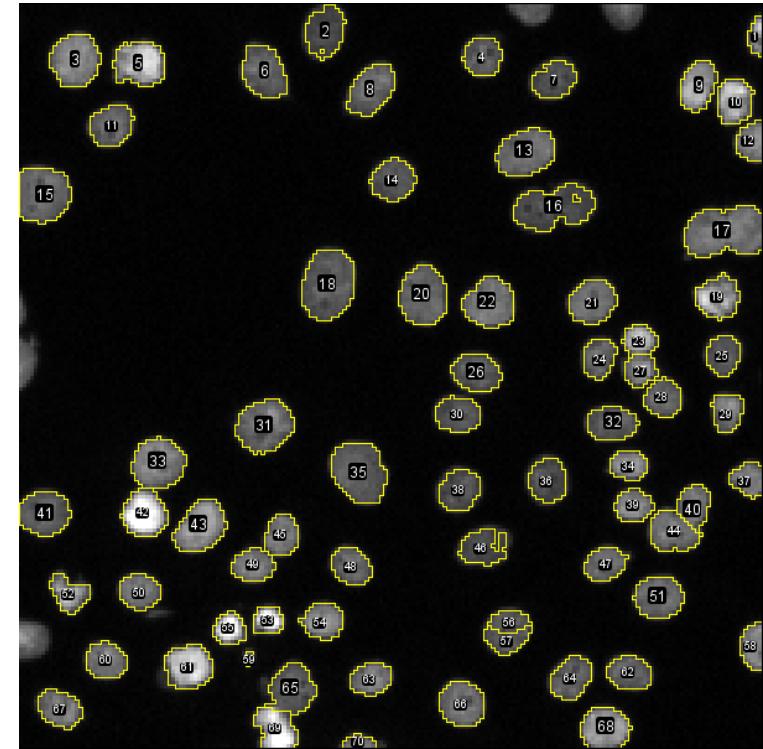
Original image



Label image



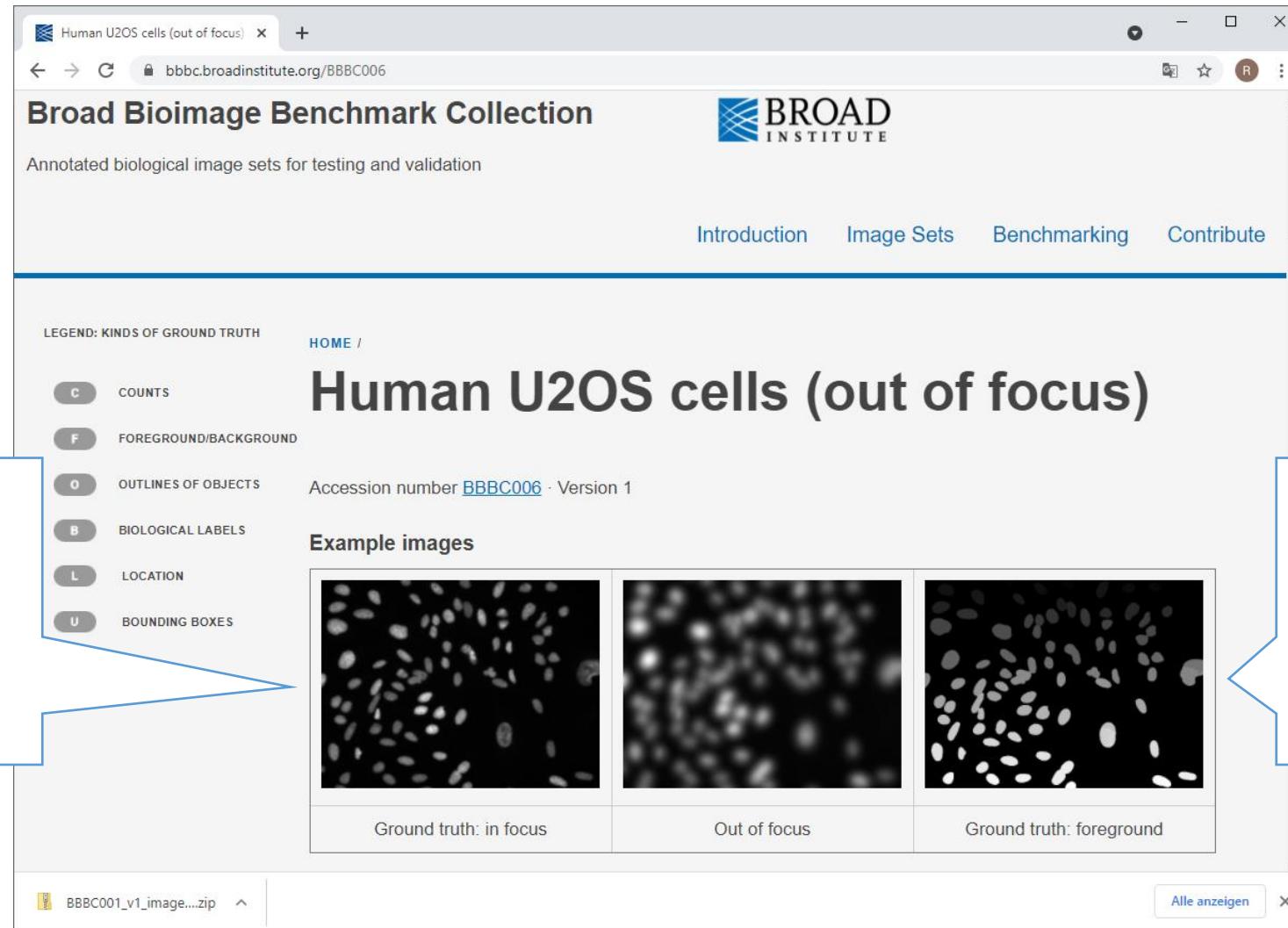
Overlay



There are 70 nuclei  
in this image.

# Reliable bio-image analysis

- Algorithms must be reliable (validated methods). Publicly available benchmark data sets allow to compare algorithms on common data.



Original image data

"Ground truth" label images

- “The image data was analyzed with ImageJ.”

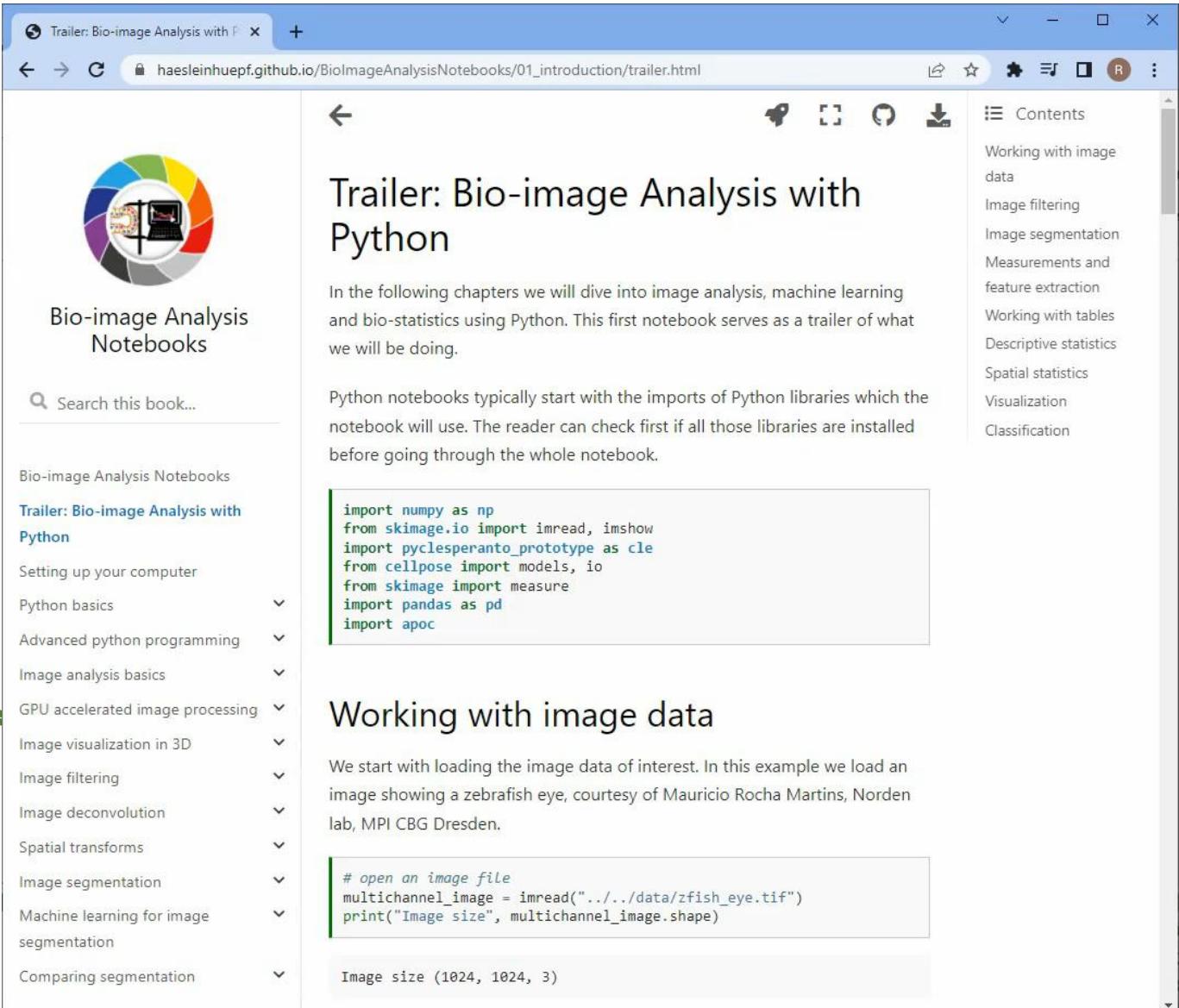
Can you reproduce  
what they did?

# Reproducible bio-image analysis

- “The image data was analyzed with ImageJ.”

Can you reproduce what they did?

Can you reproduce what they did?



The screenshot shows a web browser displaying a trailer notebook for "Bio-image Analysis with Python". The URL in the address bar is [haesleinhuepf.github.io/BioImageAnalysisNotebooks/01\\_introduction/trailer.html](https://haesleinhuepf.github.io/BioImageAnalysisNotebooks/01_introduction/trailer.html). The page title is "Trailer: Bio-image Analysis with Python". On the left, there's a sidebar with a circular logo and a search bar. Below the search bar is a list of topics: Bio-image Analysis Notebooks, Trailer: Bio-image Analysis with Python, Setting up your computer, Python basics, Advanced python programming, Image analysis basics, GPU accelerated image processing, Image visualization in 3D, Image filtering, Image deconvolution, Spatial transforms, Image segmentation, Machine learning for image segmentation, and Comparing segmentation. The main content area starts with a paragraph about the purpose of the trailer and imports for the first chapter. It then transitions to a section titled "Working with image data" with code for loading a zebrafish eye image and outputting its size.

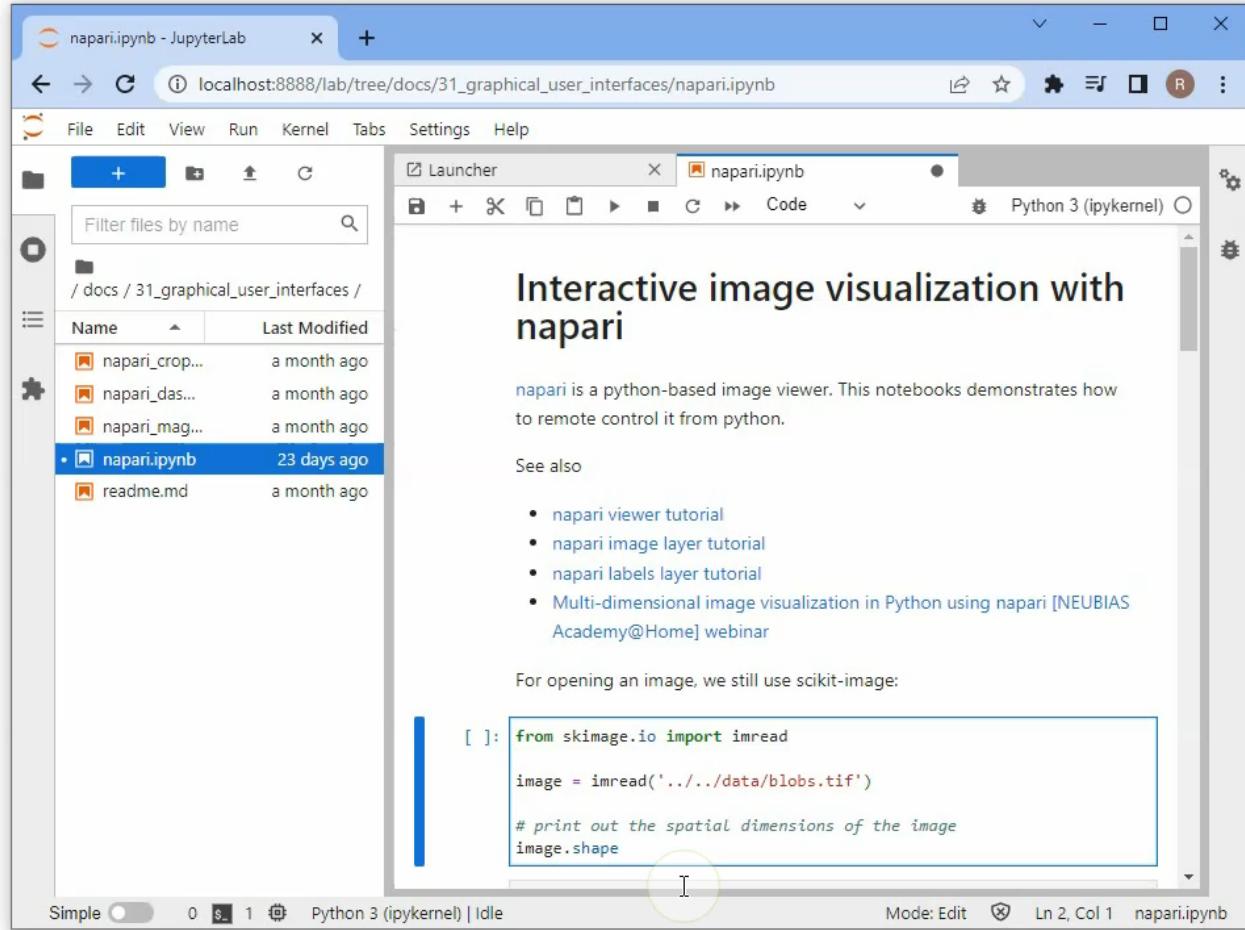
```
import numpy as np
from skimage.io import imread, imshow
import pyclesperanto_prototype as cle
from cellpose import models, io
from skimage import measure
import pandas as pd
import apoc
```

```
# open an image file
multichannel_image = imread("../data/zfish_eye.tif")
print("Image size", multichannel_image.shape)
```

Image size (1024, 1024, 3)

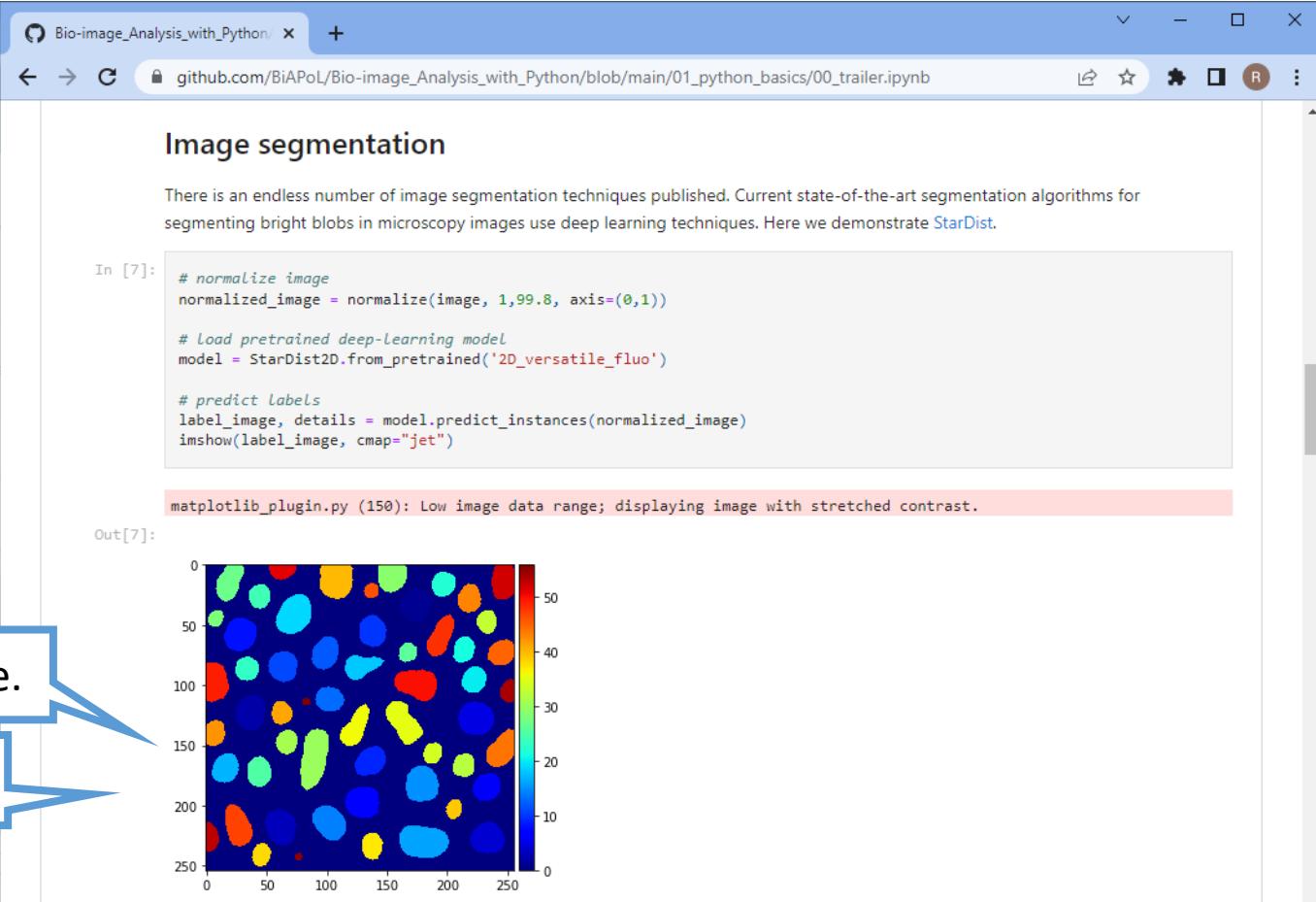
# Interactivity versus reproducibility

- Remote controlling graphical user interfaces



- Compared to wet-lab experiments, image analysis (in-silico) experiments are typically repeatable.
- In wet-lab experiments, samples may get destroyed while executing the experiment.
- Repeatability is a property of the experiment. You cannot improve repeatability by better documentation.

- However, you need to pay some extra attention to have repeatable image processing
  - Scripts need to be written in a way that their execution is repeatable.
  - Test it!



The screenshot shows a Jupyter Notebook cell titled "Image segmentation". The cell contains Python code for image segmentation using the StarDist2D library. The code includes normalization of the image, loading a pretrained model, and predicting labels. A warning message from matplotlib is displayed: "matplotlib\_plugin.py (150): Low image data range; displaying image with stretched contrast." Below the code, a heatmap visualizes the segmented blobs in various colors (red, orange, yellow, green, blue) against a dark background. A color bar on the right indicates intensity levels from 0 to 50. Three callout boxes with arrows point from the text below to the code and the resulting image:

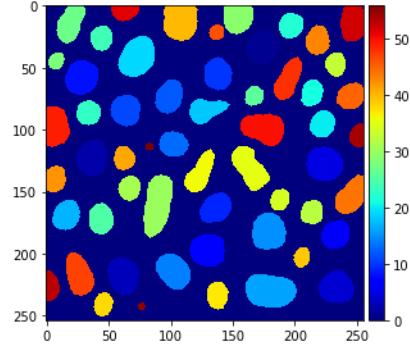
- "Run it once." points to the "# predict labels" line in the code.
- "Run it twice." points to the "imshow(label\_image, cmap="jet")" line in the code.
- "Are results identical?" points to the heatmap.

```
# normalize image
normalized_image = normalize(image, 1, 99.8, axis=(0,1))

# load pretrained deep-learning model
model = StarDist2D.from_pretrained('2D_versatile_fluo')

# predict labels
label_image, details = model.predict_instances(normalized_image)
imshow(label_image, cmap="jet")

matplotlib_plugin.py (150): Low image data range; displaying image with stretched contrast.
```



# 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**
    - Somebody else can do the experiment under *different conditions* and gets similar measurements. For this, documentation is decisive!
  - **Repeatable**
    - We can do the same experiment twice under the *same conditions* and get similar measurements.

# Image analysis is part of the experiment



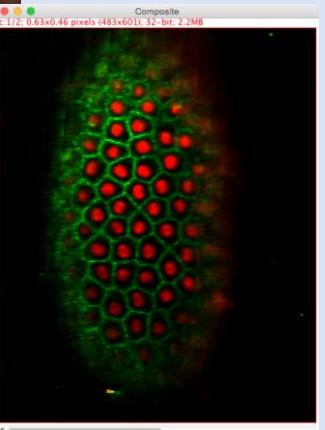
**PoL**  
Physics of Life  
TU Dresden



Observation

$$p_{ij}(t) = \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}]^\beta}{\sum_{j=1}^n [\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}]^\beta}$$

Modeling



Imaging

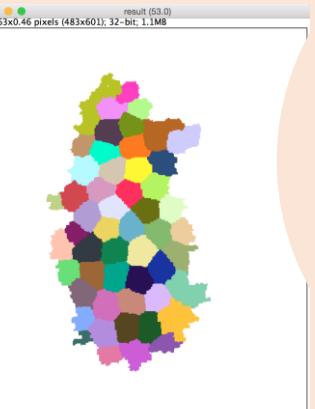


Image processing

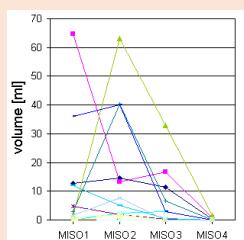
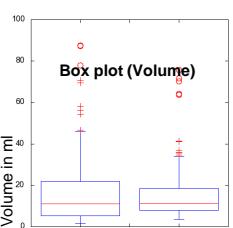


Image analysis  
Bio-statistics



- Think about how to analyze your images before starting the experiment.
  - Consider adapting your experiment so that quantitative image analysis can be performed easily.
  - Do small test-experiments.
- Think about controls, counter-proves, an easy to falsify null-hypothesis.
  - Be a lazy scientist. Do simple experiments.
- How can you exclude yourself from the experiment?
  - Think of blinding yourself or fully automate analysis.
- One experiment usually answers one question. Or less.

# Quiz

- To measure cell volume, we need cell segmentation.

True

False

- To *count cells*, we need cell segmentation.

True

False

- To count cells over time, we need *cell tracking*.

True

False

- For *cell tracking*, we need cell segmentation.

True

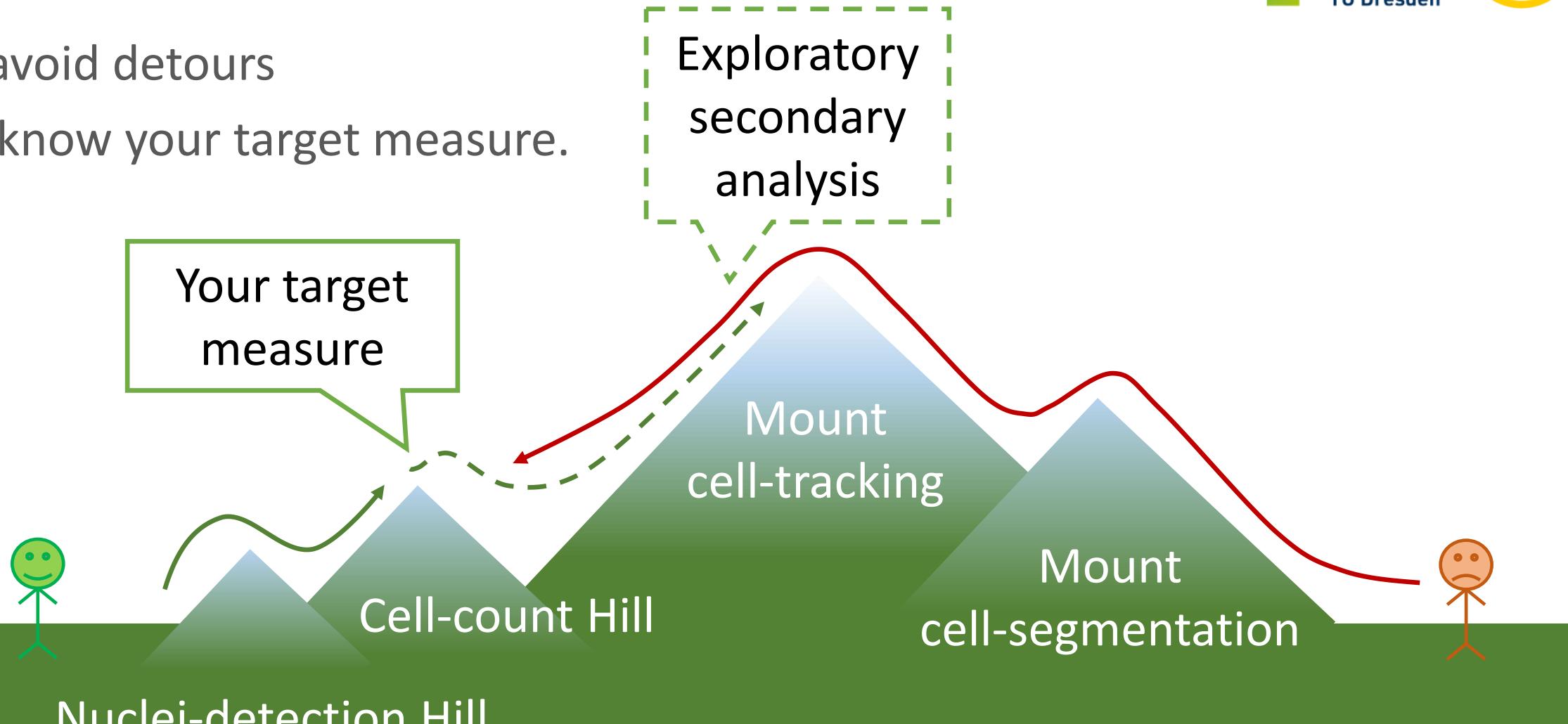
False

- A nuclei marker is best suited for cell tracking.

True

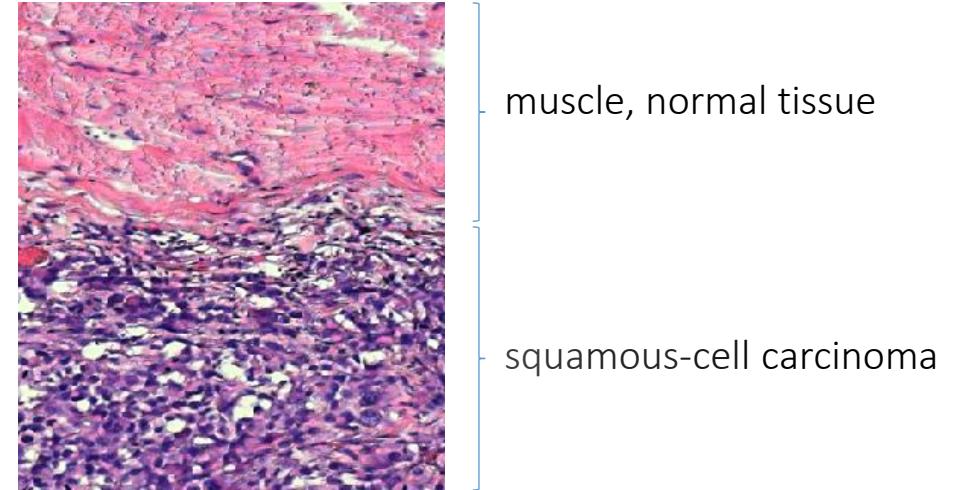
False

- ... to avoid detours
- Hint: know your target measure.

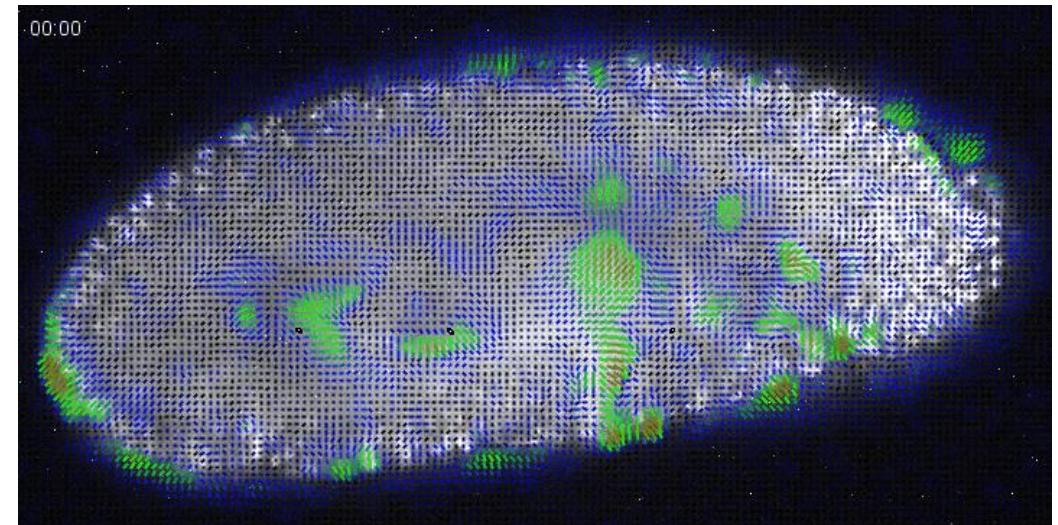


# Common questions

- Typical questions bio-image analysts deal with
  - Is signal intensity different under varying conditions?
  - How many cells are in my image?
  - How high is cell density?
    - Bio-statistics / medicine
  - How are different tissues characterized?
    - Machine learning



- Typical questions bio-image analysts struggle with
  - What force drives the observed processes?
  - What is the lineage tree of one particular cell?
  - Are observation A and observation B related?
  - Are structures observed in different color channels colocalized?



# Hypothesis-driven quantitative biology

- Hypothesis: Cell shape can be influenced by modifying X.

Should we use a different segmentation algorithm?

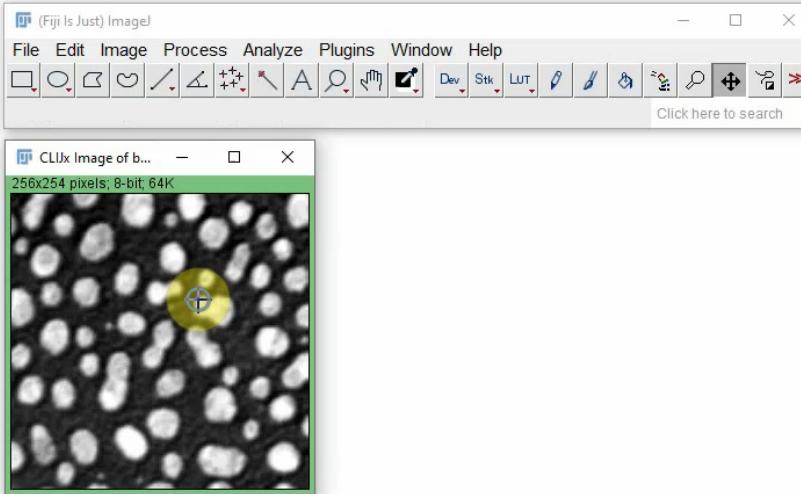
- Sample preparation
- Imaging
- Cell segmentation
- Circularity measurement
- Statistics

Shall we use a different microscope?

Is circularity the right parameter to measure?

# Hypothesis-driven quantitative biology

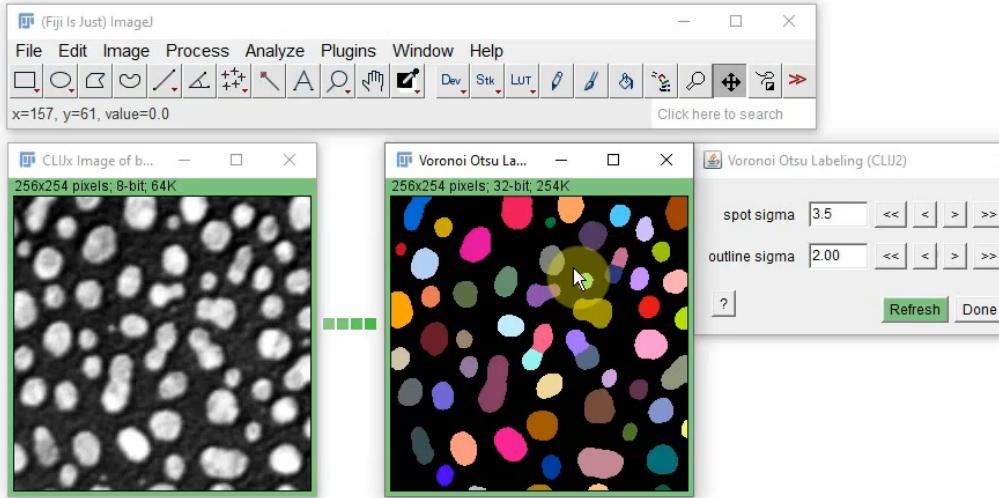
- Should we use a different segmentation algorithm?



<https://clij.github.io/>

# Hypothesis-driven quantitative biology

Is circularity the right parameter to measure?



<https://clij.github.io/>

# Hypothesis generating quantitative biology

- Hypothesis: Cell shape can be influenced by modifying X.
- Question: Which image-derived parameter is influenced when modifying X? Why?
  - Sample preparation
  - Imaging
  - Cell segmentation algorithm A, algorithm B, algorithm C
- Measurement of circularity, solidity, elongation, extend, texture, intensity, topology ...
  - Statistics

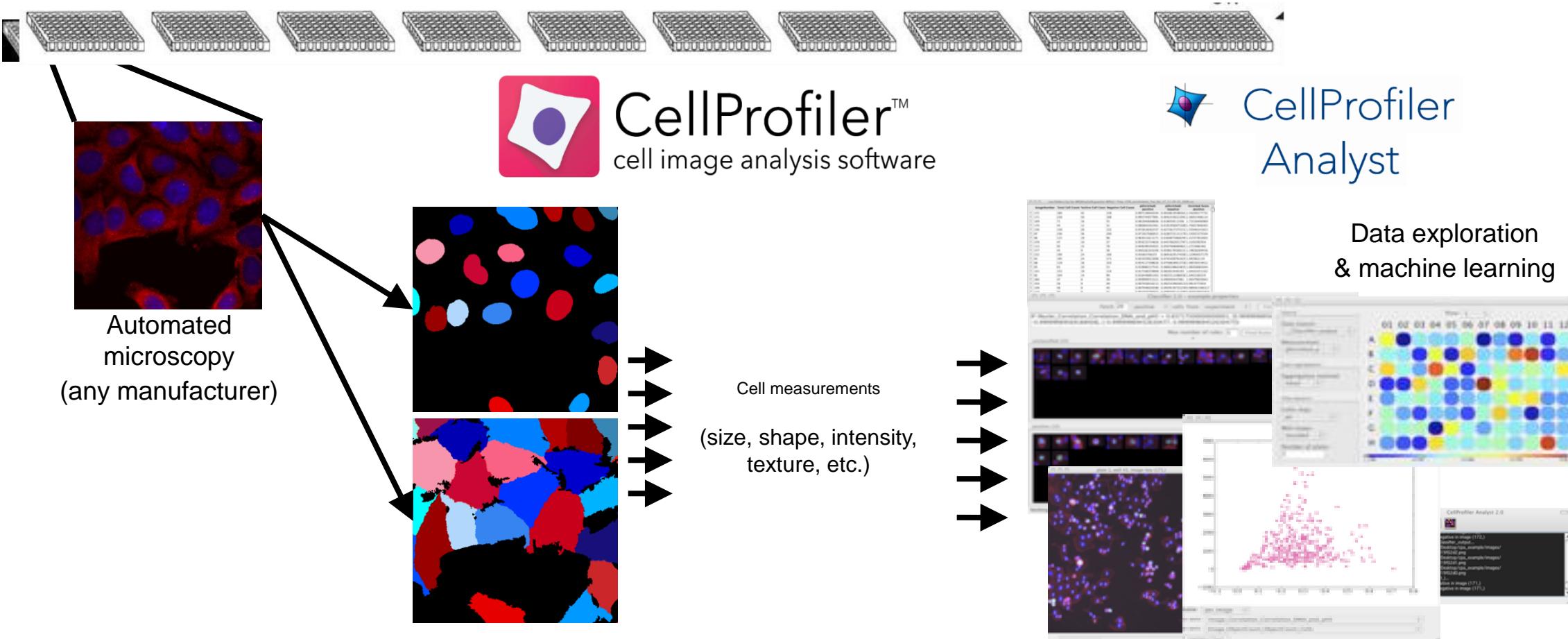
Which segmentation algorithms allow measurements that show a relationship with X?

Why?

Which parameter shows any relationship with X?

# Cell Profiler

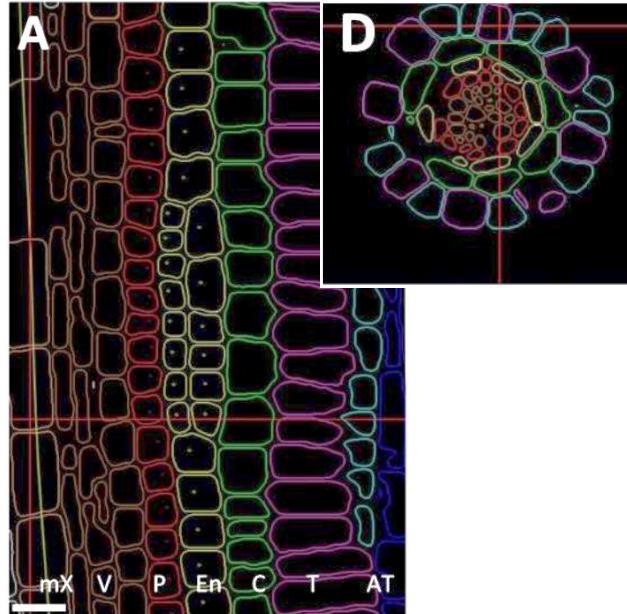
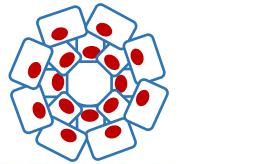
- Cells or organisms in multiwell plates, each well treated with a gene or chemical perturbant



# Image analysis beyond segmentation

- Spatial pattern / image analysis for cell biology

Membranes  
Nuclei



Structural biology of plants<sup>1</sup>

15 years ago!

[Joint Pattern Recognition Symposium](#)  
DAGM 2006: [Pattern Recognition](#) pp 182-191 | [Cite as](#)

Fast Scalar and Vectorial Grayscale Based Invariant Features for **3D Cell Nuclei Localization and Classification**

Authors

Authors and affiliations

Janina Schulz, Thorsten Schmidt, Olaf Ronneberger, Hans Burkhardt, Taras Pasternak, Alexander Dovzhenko, Klaus Palme

Conference paper

10

1.8k

Citations Downloads

Part of the [Lecture Notes in Computer Science](#) book series (LNCS, volume 4174)



1) Image data source: Pasternak et al. ([CC-BY 4.0](#)) <https://www.biorxiv.org/content/10.1101/2021.01.01.425043v2>

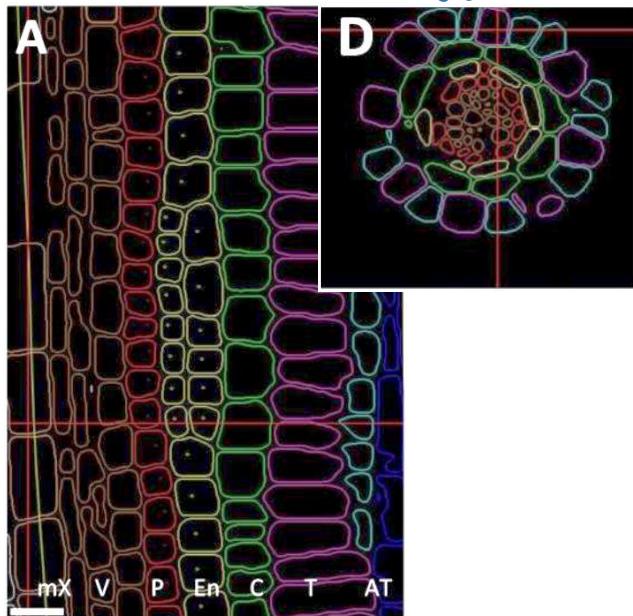
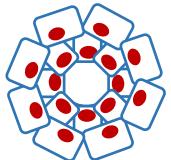
@Haesemeyer

@PoLDresden

# Image analysis beyond segmentation

- Spatial pattern / image analysis for cell biology

Membranes  
Nuclei



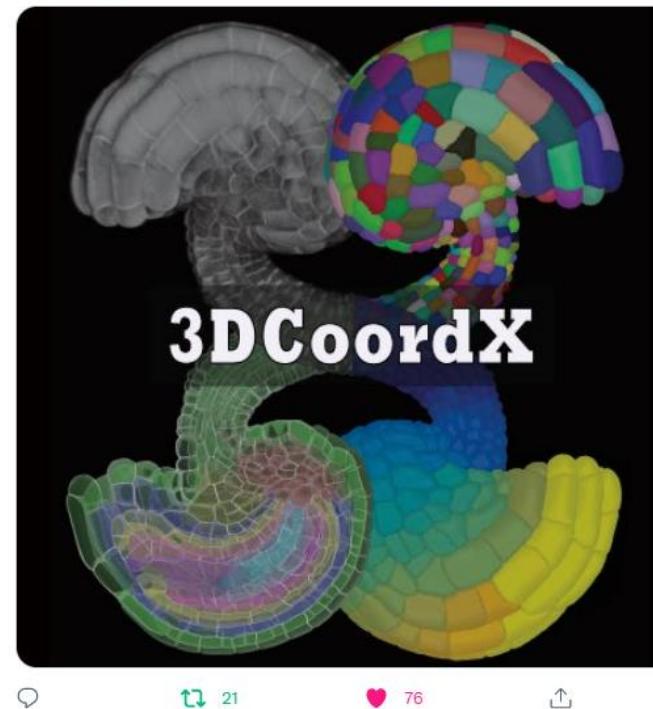
Structural biology of plants<sup>1</sup>



Kay Schneitz @PlantDevTUM · Mar 29

Just in case you hadn't yet noticed our supercool new toolbox for MorphoGraphX. Provides organ-centric spatial context to cellular features in 3D digital plant organs.  
[tinyurl.com/2n6rb6ds](http://tinyurl.com/2n6rb6ds)  
@athul\_r\_v @StraussSoe @tejasvinee\_m @kareninglelee @RichardSmithLab @PlantPhys

...



*Plant Physiology*®

Issues

Advance Articles

More Content ▾

Submit ▾

Purchase

About ▾

All Plant Physiology

Q Advanced Search

#### Article Contents

Abstract

Author notes

Supplementary data

#### ACCEPTED MANUSCRIPT

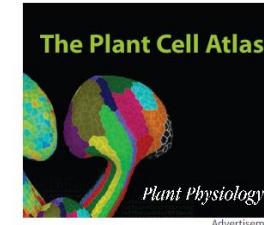
The annotation and analysis of complex 3D plant organs using 3DCoordX

Athul Vijayan, Soeren Strauss, Rachele Tofanelli, Tejasvinee Atul Mody, Karen Lee, Miltos Tsiantis, Richard S Smith, Kay Schneitz Author Notes

*Plant Physiology*, kiac145, <https://doi.org/10.1093/plphys/kiac145>

Published: 28 March 2022 Article history ▾

PDF Split View Cite Permissions Share ▾



American Society of Plant Biologists

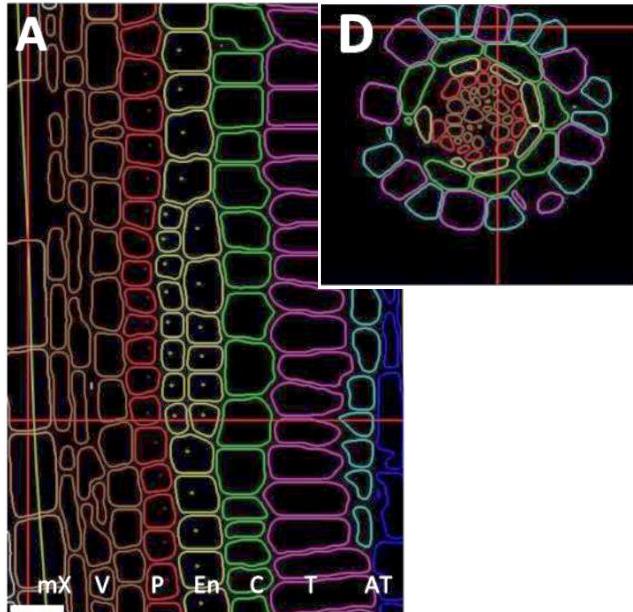
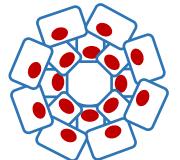


- 1) Image data source: Pasternak et al. (CC-BY 4.0) <https://www.biorxiv.org/content/10.1101/2021.01.01.425043v2>
- 2) Palla et al <https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiac145/6555043>

# Image analysis beyond segmentation

- Spatial pattern / image analysis for cell biology

Membranes  
Nuclei



Structural biology of plants<sup>1</sup>

Cell Reports  
Resource

**CytoMAP: A Spatial Analysis Toolbox Reveals Features of Myeloid Cell Organization in**

**nature methods**

Caleb R. Stoltzfus,<sup>1</sup> Jakub Filip,<sup>1</sup> Miranda R. Lyons-Cohen,<sup>1</sup> Jes...  
<sup>1</sup>Department of Immunology, University of Washington, Seattle, Washington, USA.  
<sup>2</sup>Seattle Children's Research Institute, Seattle, Washington, USA.  
<sup>3</sup>Department of Pediatrics, University of Washington, Seattle, Washington, USA.  
<sup>4</sup>School of Medicine, University of Washington, Seattle, Washington, USA.  
<sup>5</sup>Roche Innovation Center Munich, Penzberg, Germany.  
<sup>6</sup>Roche Innovation Center Zurich, Fällanden, Switzerland.  
<sup>7</sup>Lead Contact.  
\*Correspondence: geremy@uw.edu.  
<https://doi.org/10.1101/2021.01.01.425043v2>

Explore content ▾ About the journal ▾ Publish with us ▾

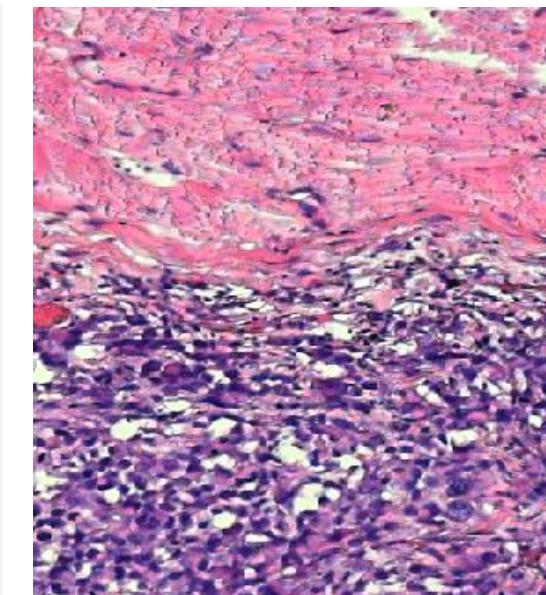
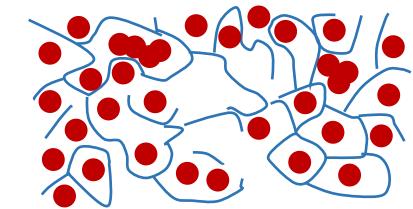
nature > nature methods > articles > article

Article | Open Access | Published: 31 January 2022

**Squidpy: a scalable framework for spatial omics analysis**

Giovanni Palla, Hannah Spitzer, Michal Klein, David Fischer, Anna Christina Schaar, Louis Benedikt Kuemmerle, Sergei Rybakov, Ignacio L. Ibarra, Olle Holmberg, Isaac Virshup, Mohammad Lotfollahi, Sabrina Richter & Fabian J. Theis

*Nature Methods* 19, 171–178 (2022) | [Cite this article](#)  
19k Accesses | 3 Citations | 375 Altmetric | [Metrics](#)



Cancer research / histology

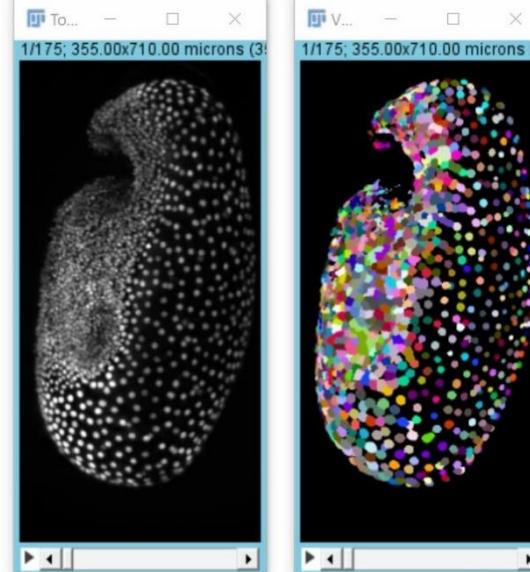
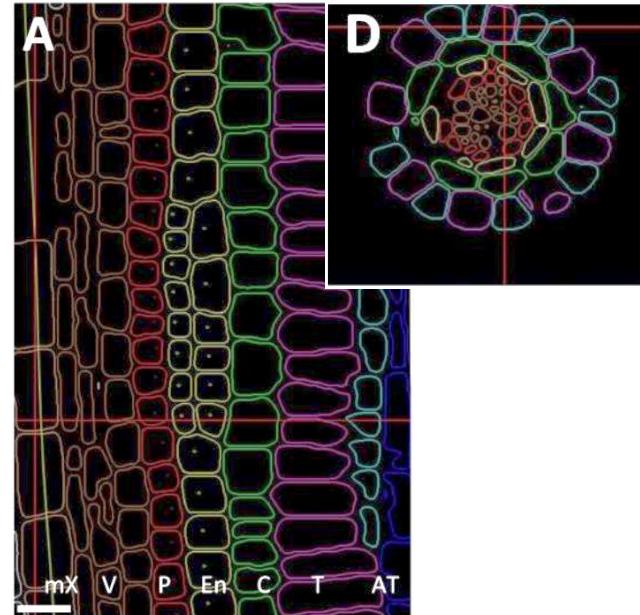
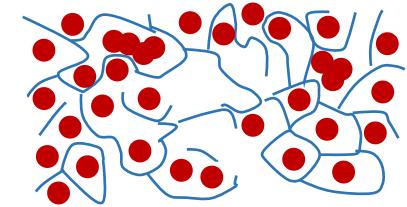
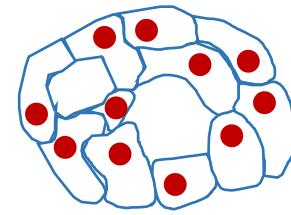
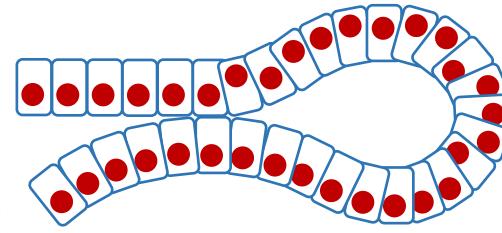
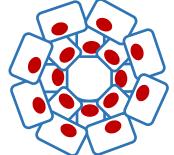


@haesleinhuepf 1) Image data source: Pasternak et al. (CC-BY 4.0) <https://www.biorxiv.org/content/10.1101/2021.01.01.425043v2>  
@PoLDresden 2) [https://www.cell.com/cell-reports/fulltext/S2211-1247\(20\)30423-X](https://www.cell.com/cell-reports/fulltext/S2211-1247(20)30423-X)  
3) Solorzano et al. <https://www.nature.com/articles/s41592-021-01358-2>

# Image analysis beyond segmentation

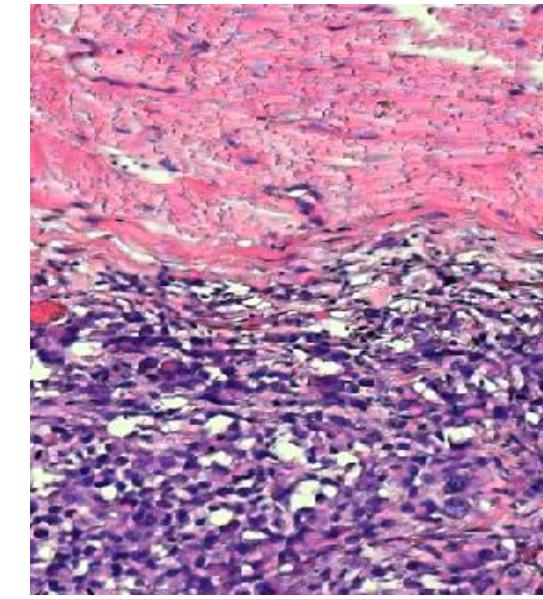
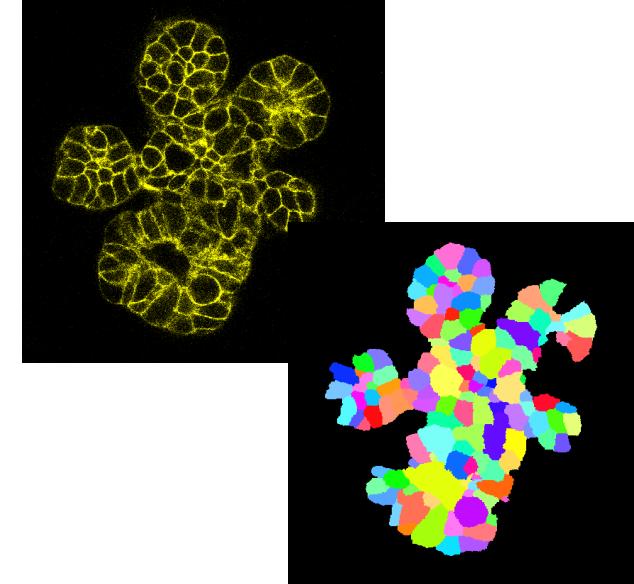
- Spatial pattern / image analysis for cell biology

Membranes  
Nuclei



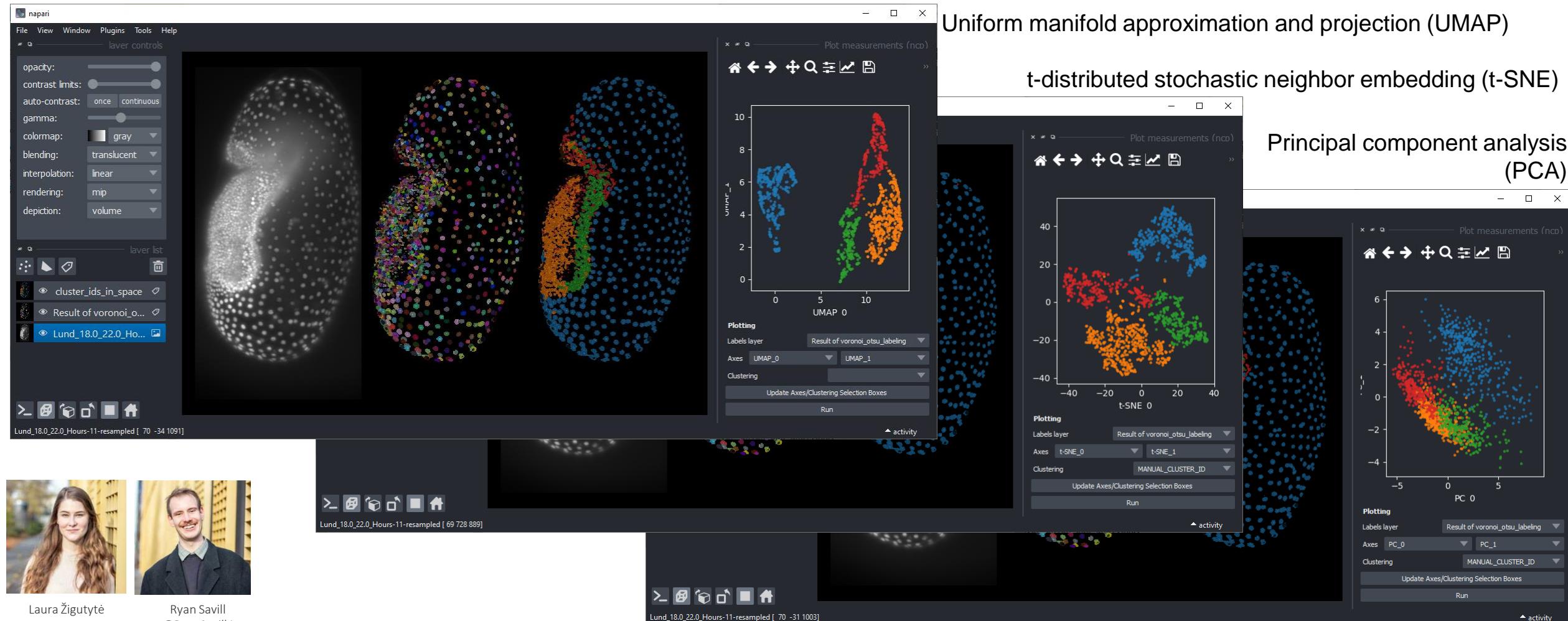
Structural biology of plants<sup>1</sup>

Early embryo development



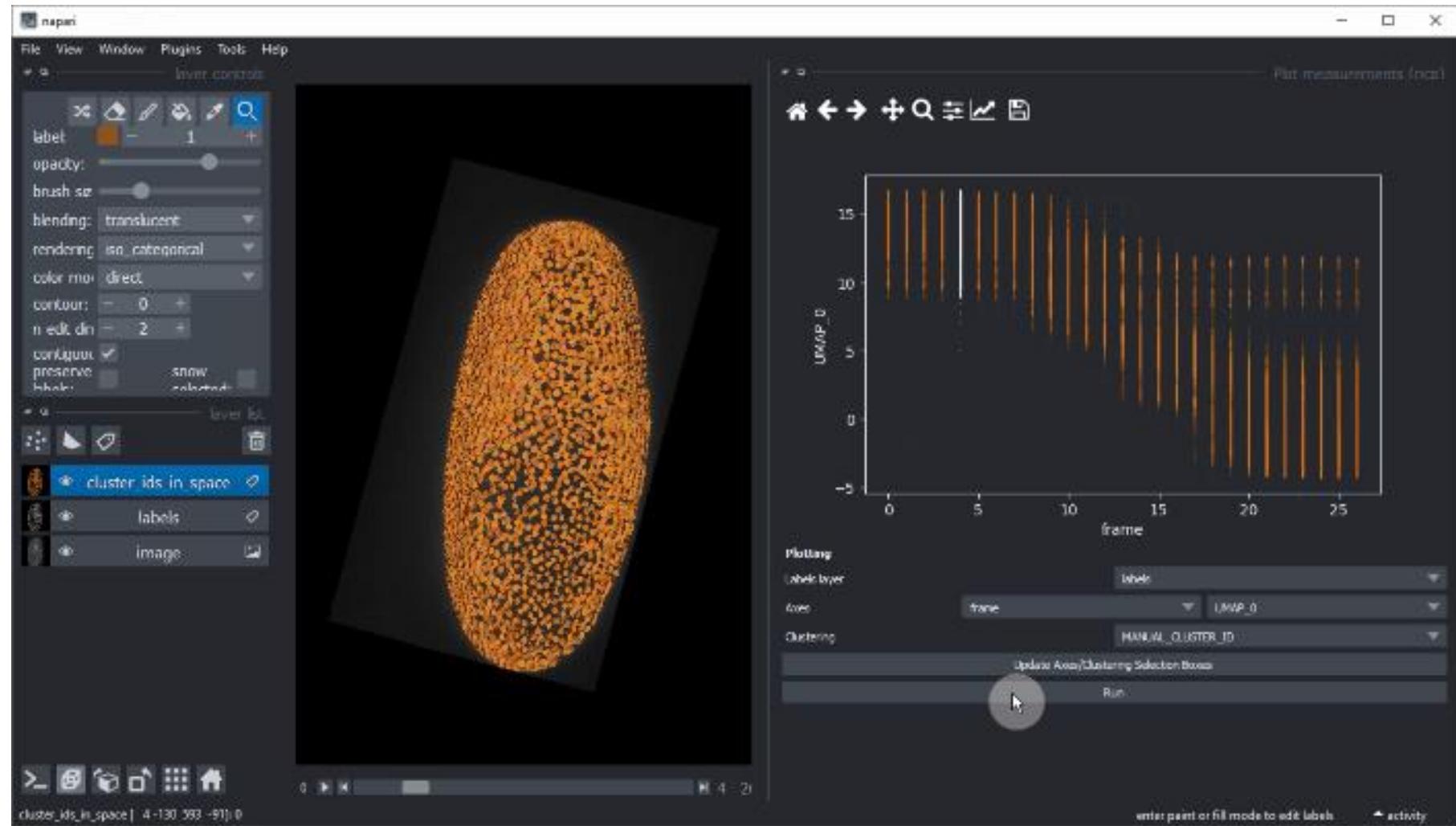
Cancer research / histology

# Dimensionality reduction



# Data exploration

- Manual clustering to gain deeper insights in relationships between measured parameters



Laura Žigutytė  
@zigutyte

Ryan Savill  
@RyanSavill4

Marcelo Zoccoler  
@zoccolermarcelo

Start talking to image-analysis / data-science experts  
early during your project.



[biapol@groups.tu-dresden.de](mailto:biapol@groups.tu-dresden.de)  
[scicomp@mpicbg.de](mailto:scicomp@mpicbg.de)

So far, you learned

- *Bio-image analysis*
  - Quantitative
  - Objective
  - Reproducible
  - Repeatable
  - Reliable
- When to talk to an expert

Coming up next

- Working with images from Python
- Image filtering