

Introduction to Bio-Image Analysis

Robert Haase

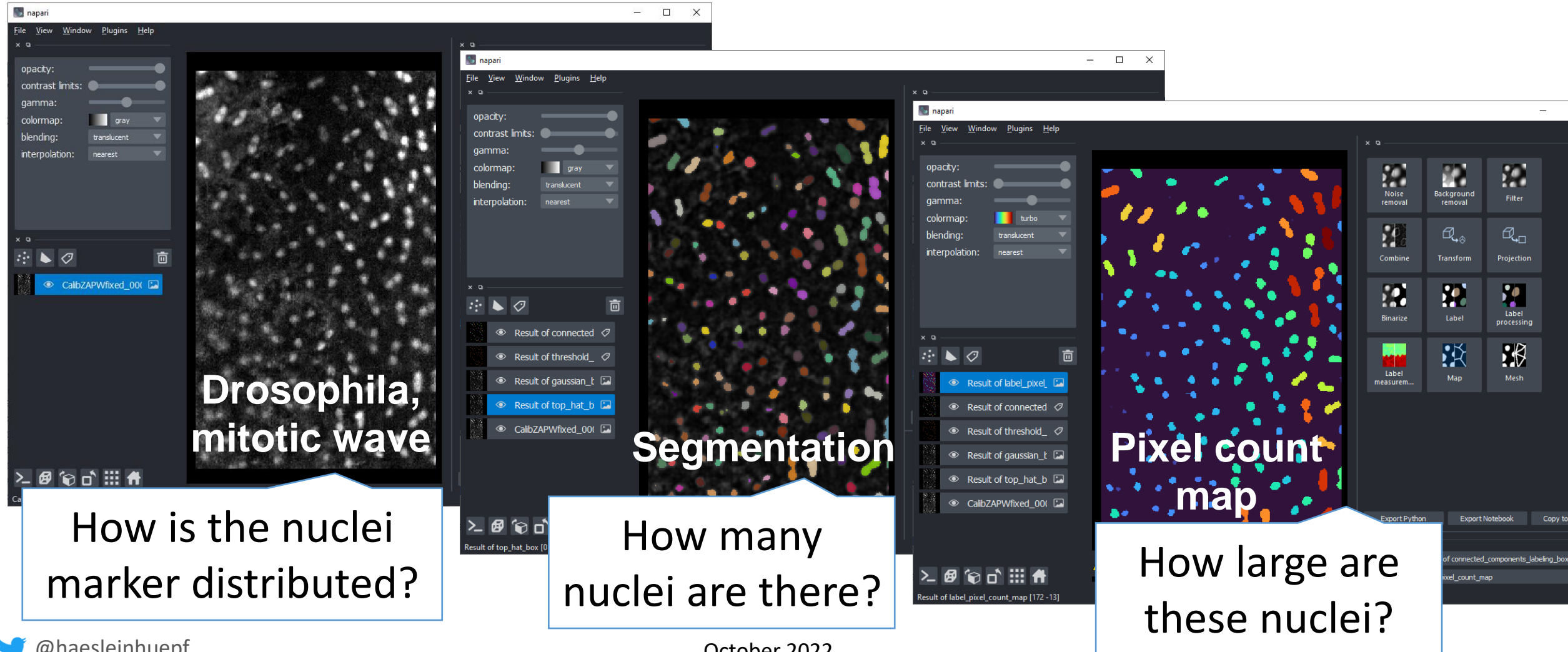
Reusing some materials made by Beth Cimini,
Anne Carpenter & colleagues (Broad Institute)

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

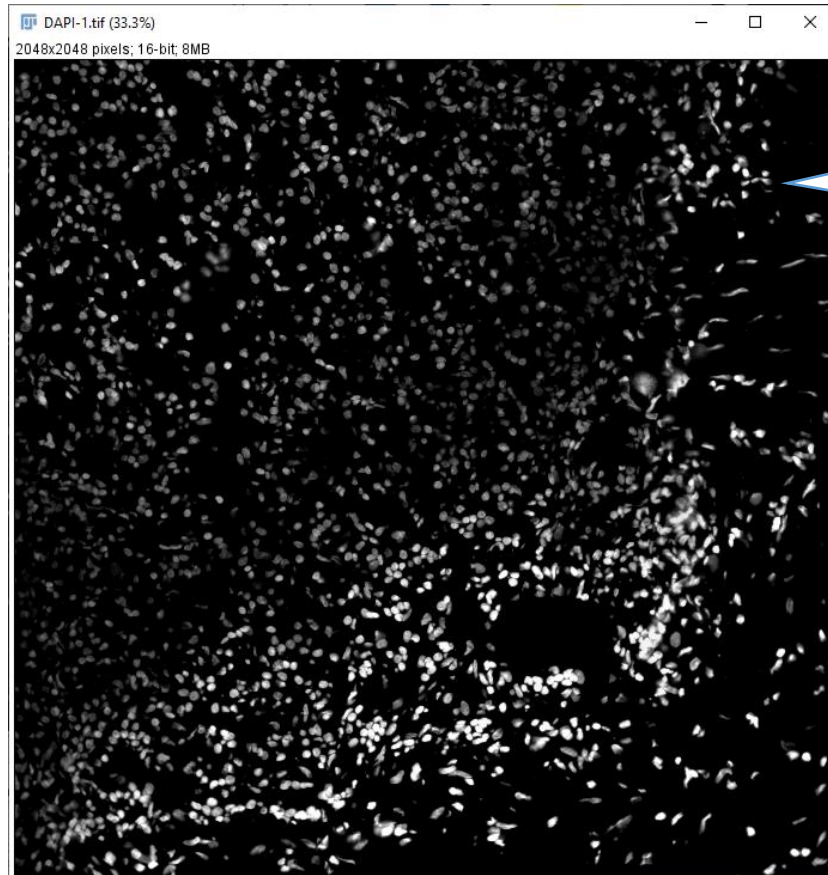


$$\text{cat height} = \underline{1.5} \times \text{microscope height}$$

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



- 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.

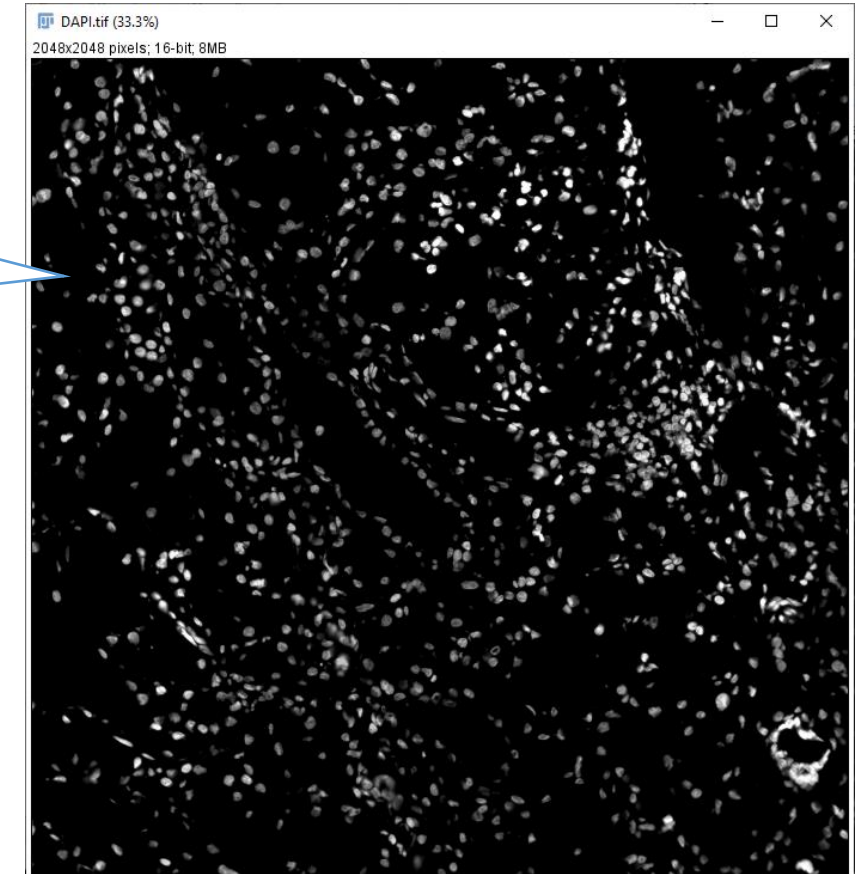
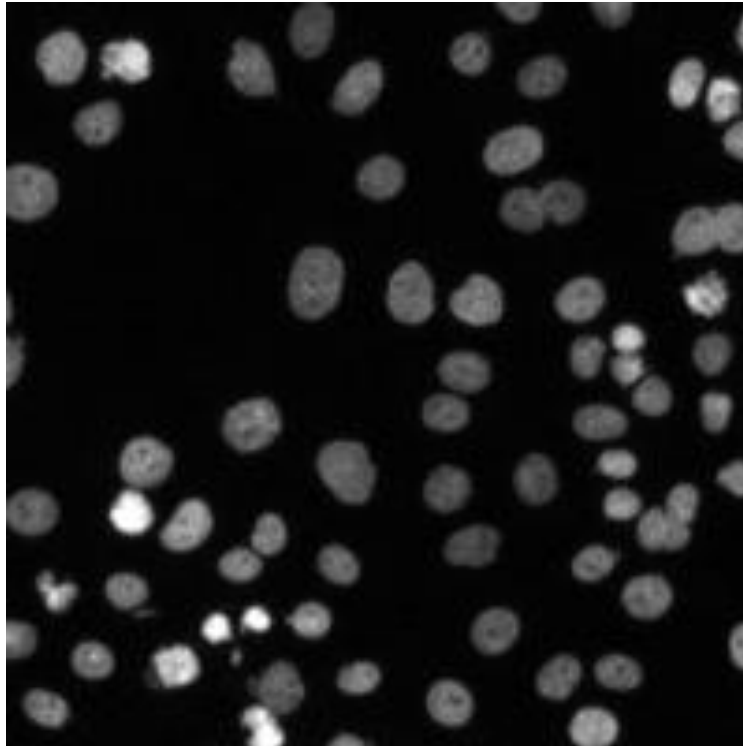


Image data source: Pascual-Reguant, Anna. (2021). Immunofluorescence staining of a human kidney (#2, peri-tumor area) obtained by MELC [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.4434462> licensed [CC-BY 4.0](#)

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

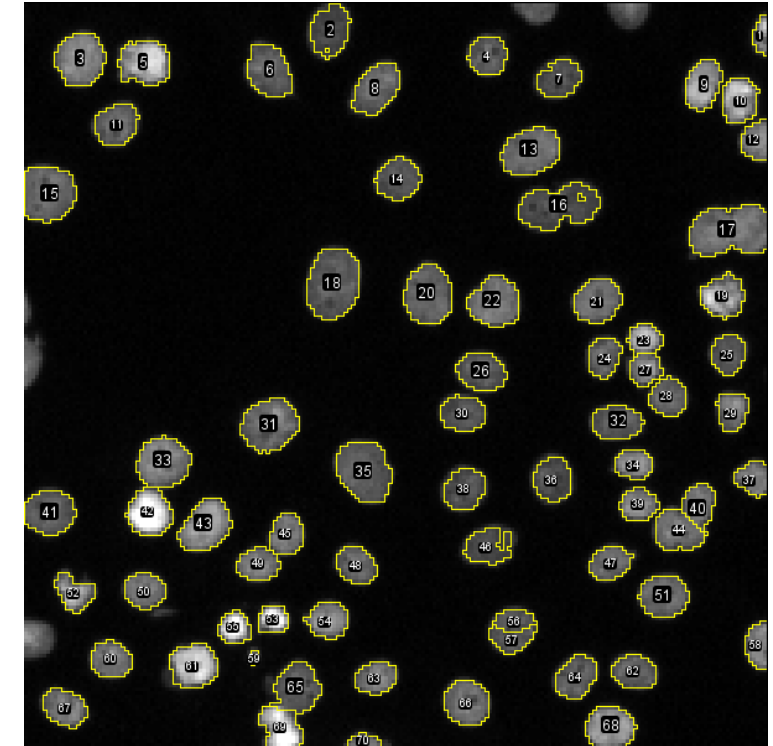
Original image



Label image

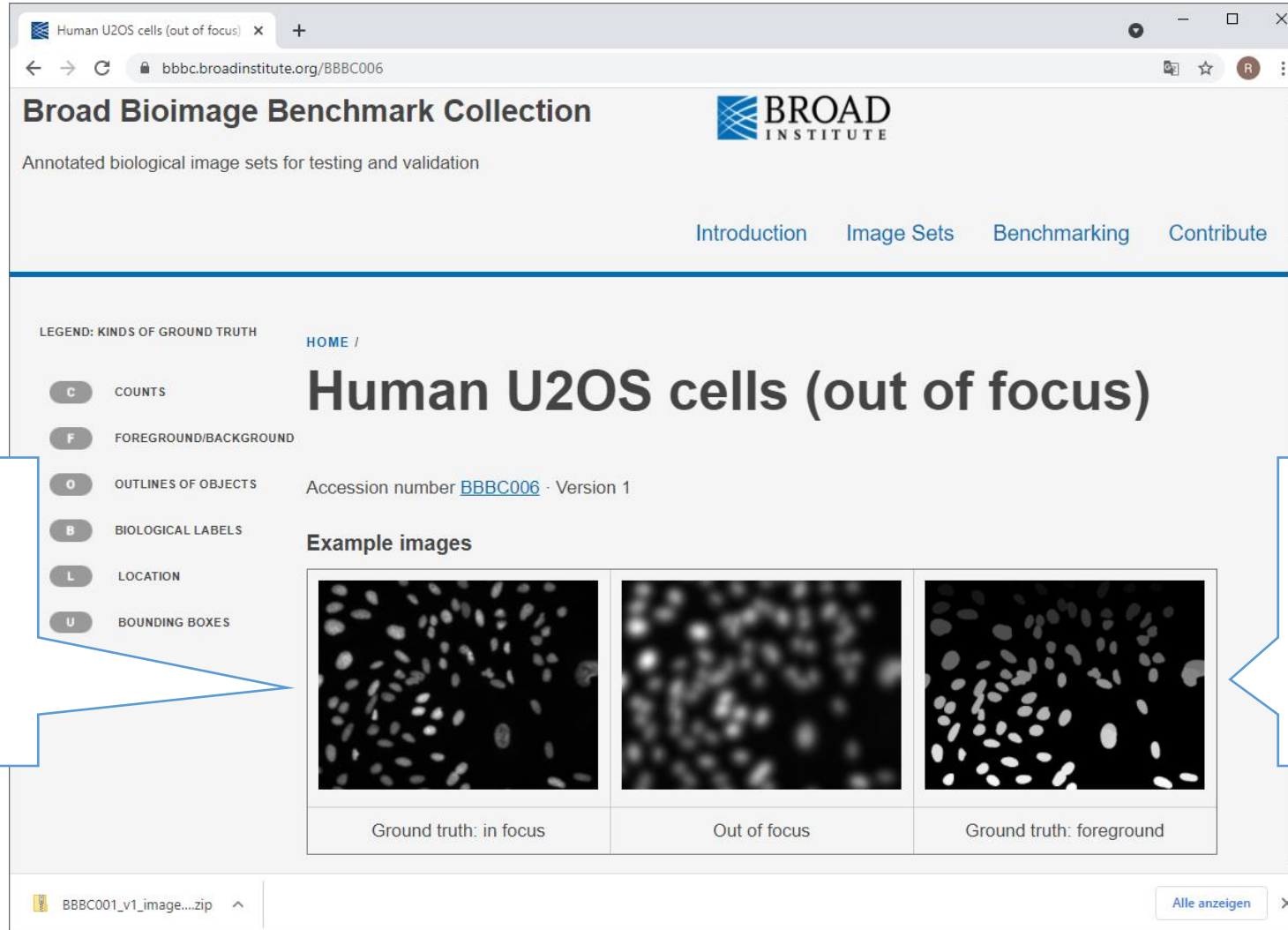


Overlay



There are 70 nuclei
in this image.

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



Human U2OS cells (out of focus)

Broad Bioimage Benchmark Collection

Annotated biological image sets for testing and validation

Introduction Image Sets Benchmarking Contribute

LEGEND: KINDS OF GROUND TRUTH

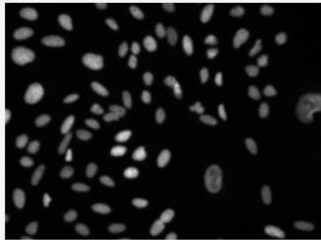
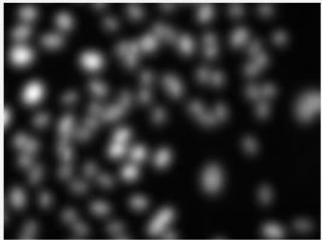
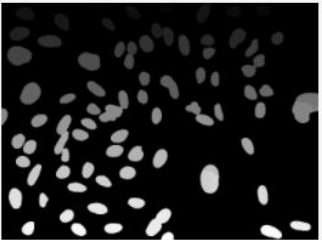
- C COUNTS
- F FOREGROUND/BACKGROUND
- O OUTLINES OF OBJECTS
- B BIOLOGICAL LABELS
- L LOCATION
- U BOUNDING BOXES

HOME /

Human U2OS cells (out of focus)

Accession number [BBBC006](#) · Version 1

Example images

		
Ground truth: in focus	Out of focus	Ground truth: foreground

BBBC001_v1_image....zip

Alle anzeigen

Original image
data

“Ground truth”
label images

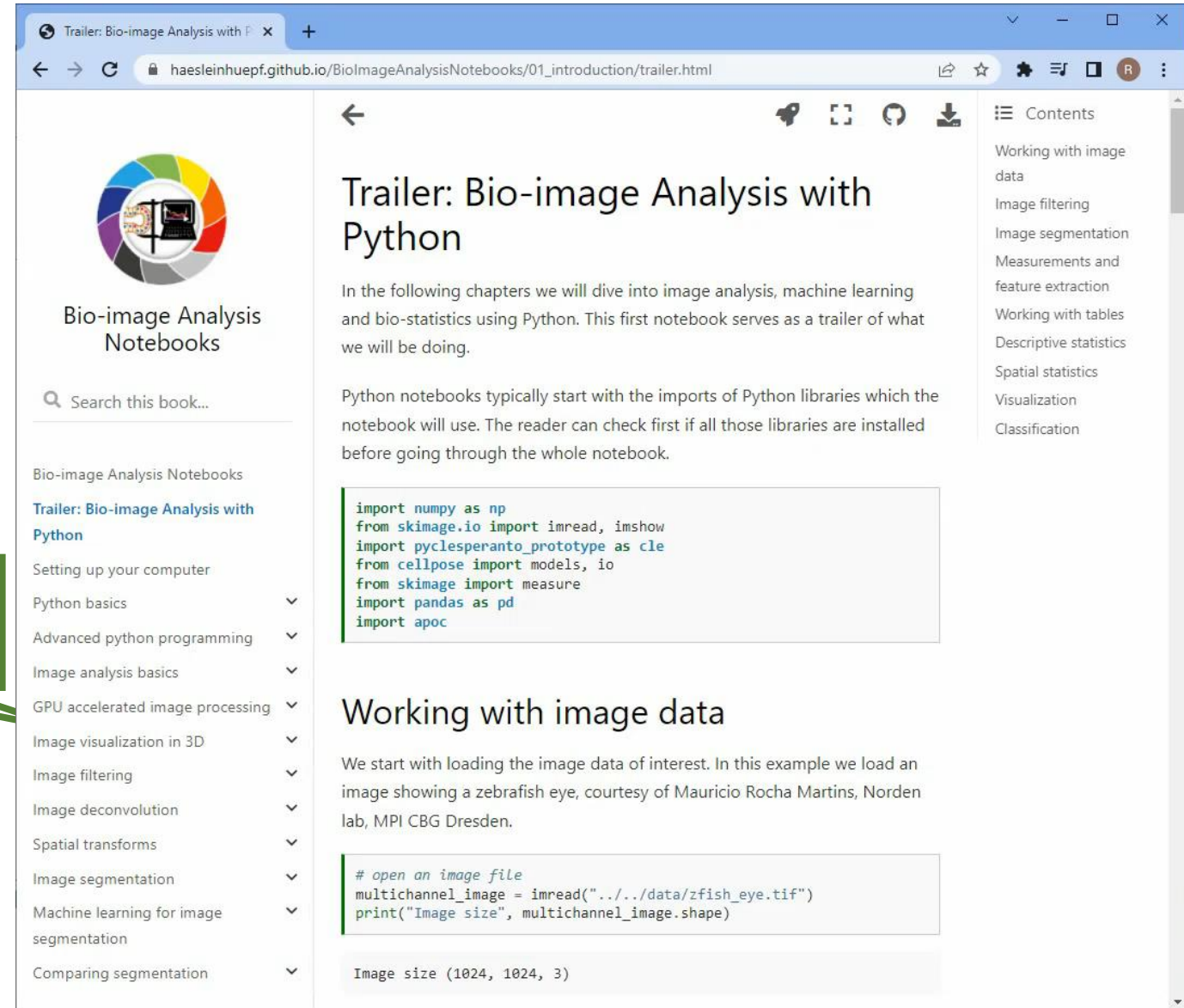
- “The image data was analyzed with ImageJ.”

Can you reproduce
what they did?

- “The image data was analyzed with ImageJ.”

Can you reproduce
what they did?

Can you reproduce
what they did?



Trailer: Bio-image Analysis with Python

In the following chapters we will dive into image analysis, machine learning and bio-statistics using Python. This first notebook serves as a trailer of what we will be doing.

Python notebooks typically start with the imports of Python libraries which the notebook will use. The reader can check first if all those libraries are installed before going through the whole notebook.

```
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
```

Working with image data

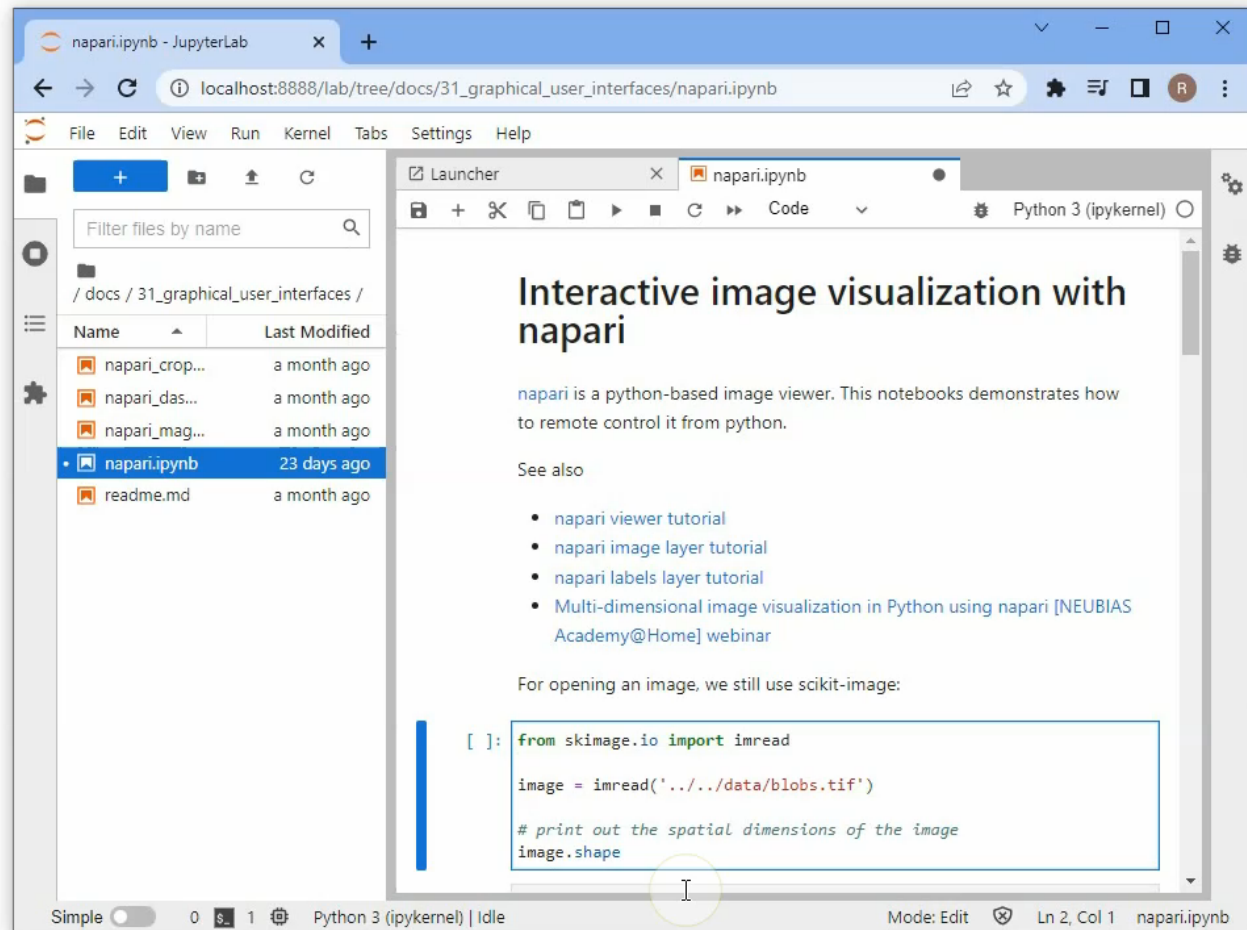
We start with loading the image data of interest. In this example we load an image showing a zebrafish eye, courtesy of Mauricio Rocha Martins, Norden lab, MPI CBG Dresden.

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

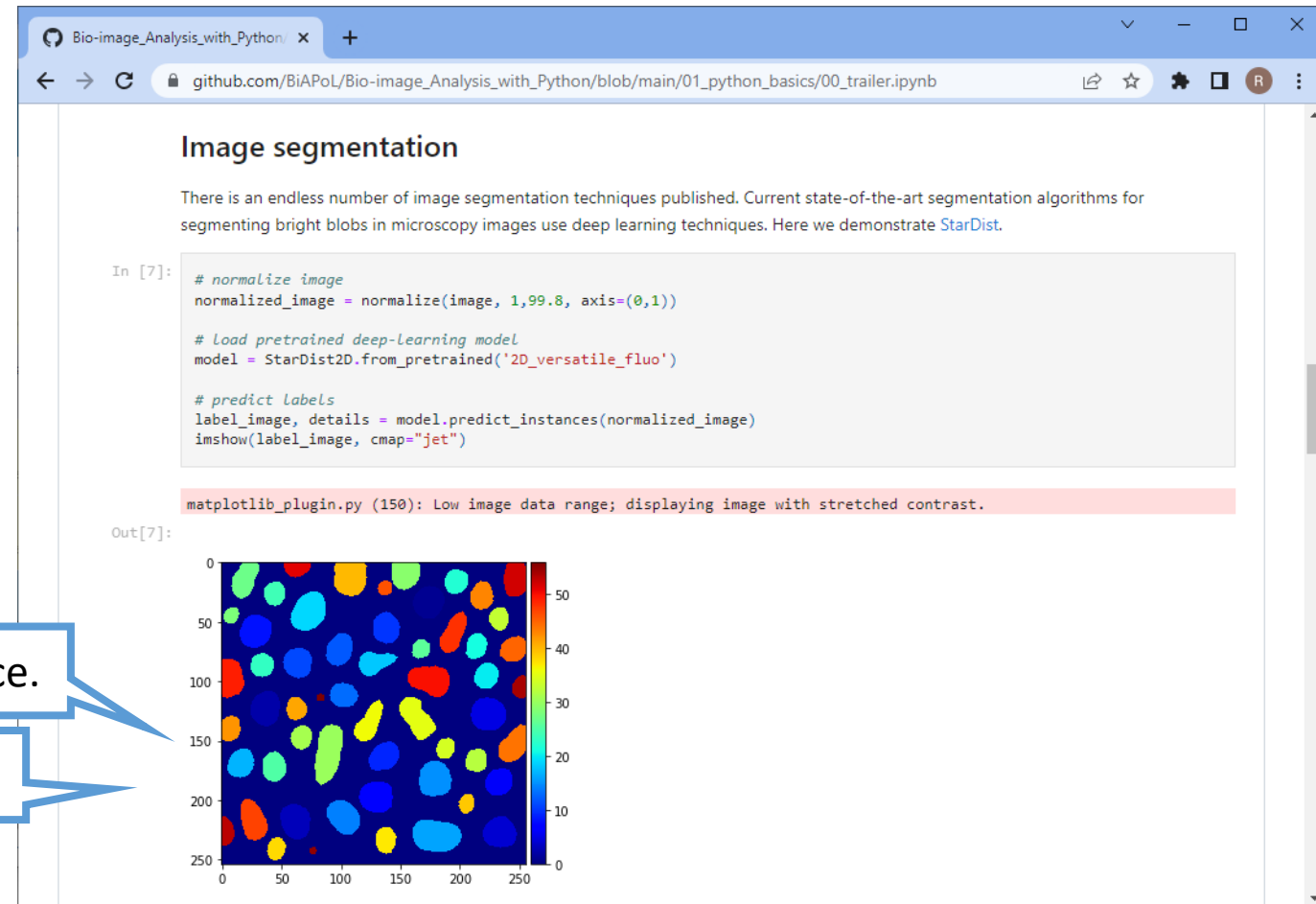
Image size (1024, 1024, 3)

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!



Run it once.

Run it twice.

Are results
identical?

- 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.

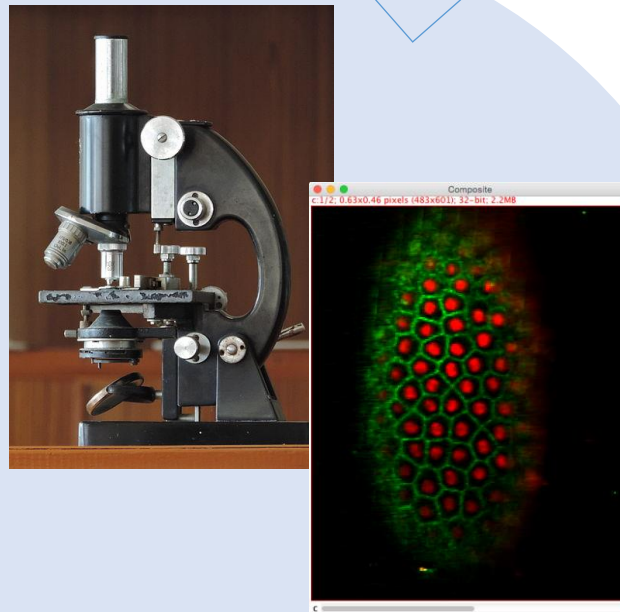
Image analysis is part of the experiment



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

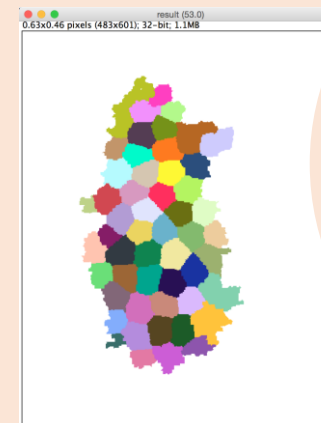


Image processing

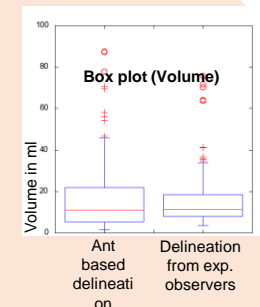
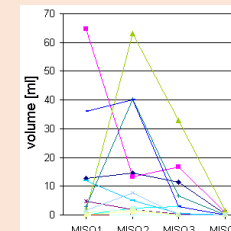


Image analysis
Bio-statistics

- 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.

- 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

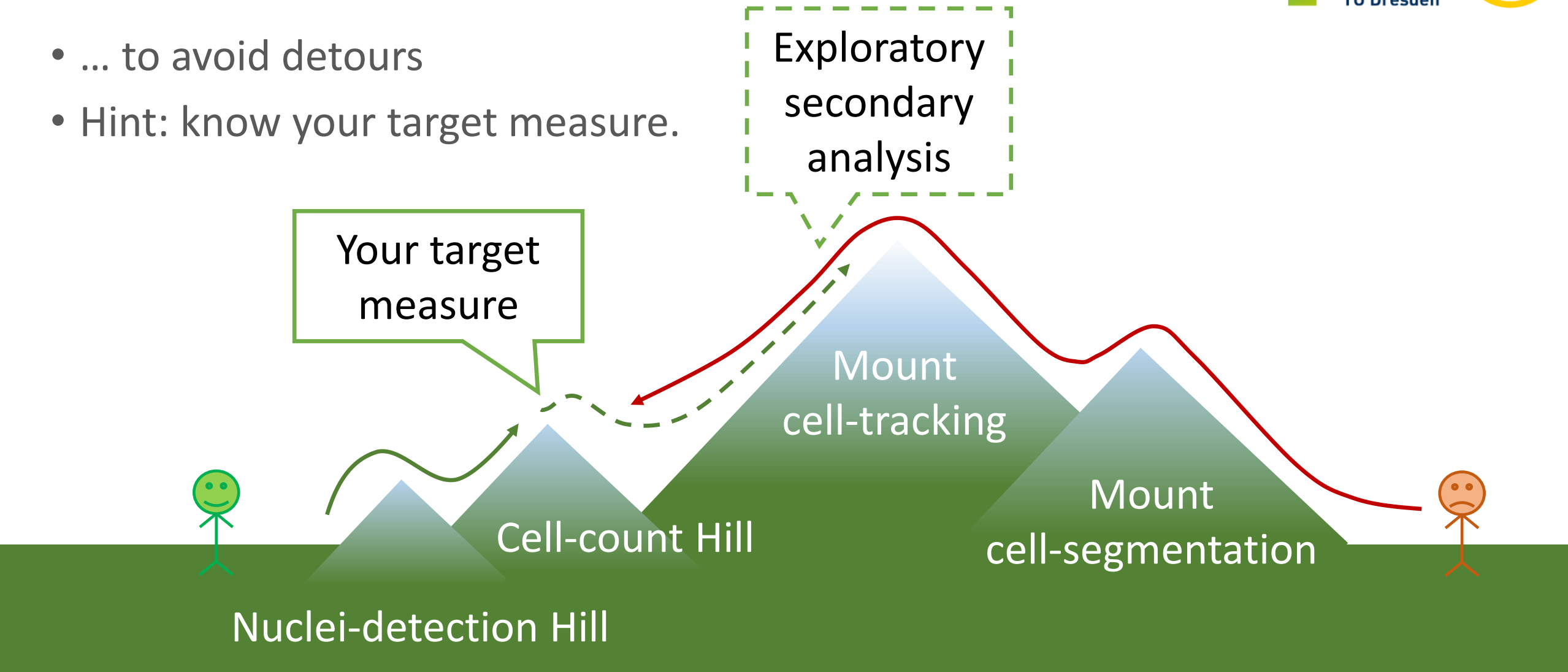
- A nuclei marker is best suited for cell tracking.

True

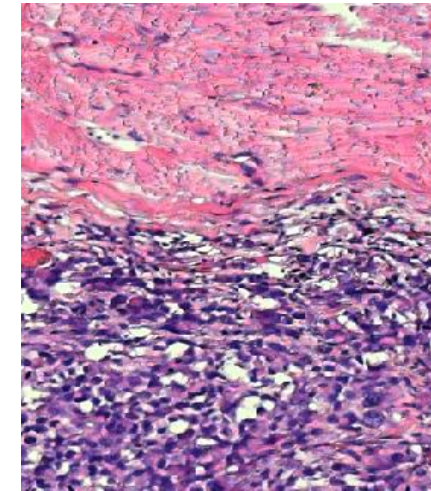
False

Choose your path wisely

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

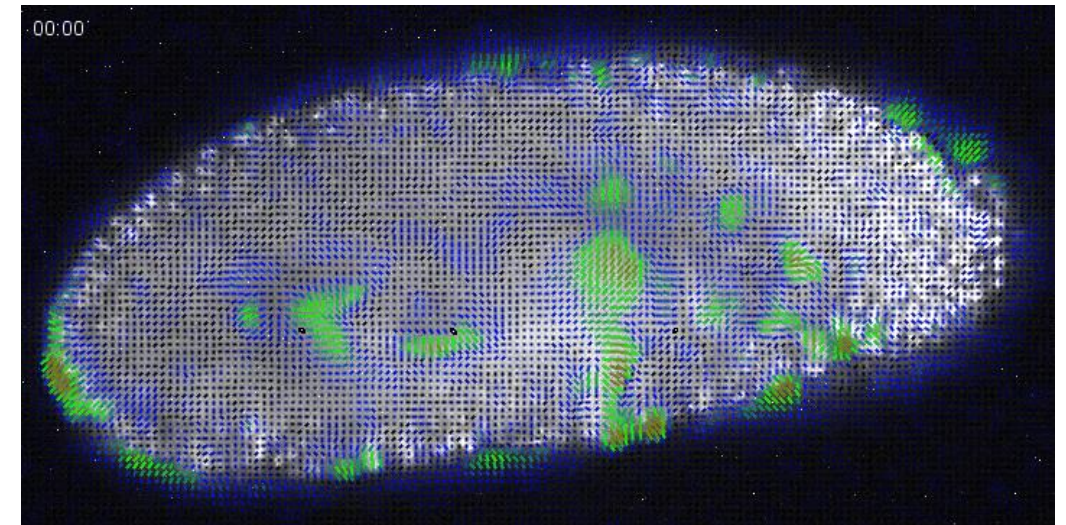


- Typical questions bio-image analysts deal with
 - Is signal intensity different under varying conditions?
 - How many cells are in my image?
 - How high is cell density?
 - Bio-statistics / medicine
 - How are different tissues characterized?
 - Machine learning
- Typical questions bio-image analysts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



muscle, normal tissue

squamous-cell carcinoma



Hypothesis-driven quantitative biology

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

Should we use a different segmentation algorithm?

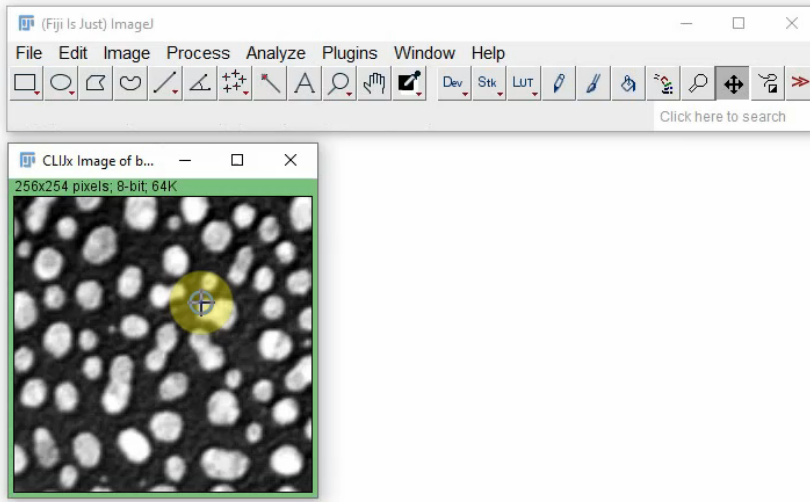
- Sample preparation
- Imaging
- Cell segmentation
- Circularity measurement
- Statistics

Shall we use a different microscope?

Is circularity the right parameter to measure?

Hypothesis-driven quantitative biology

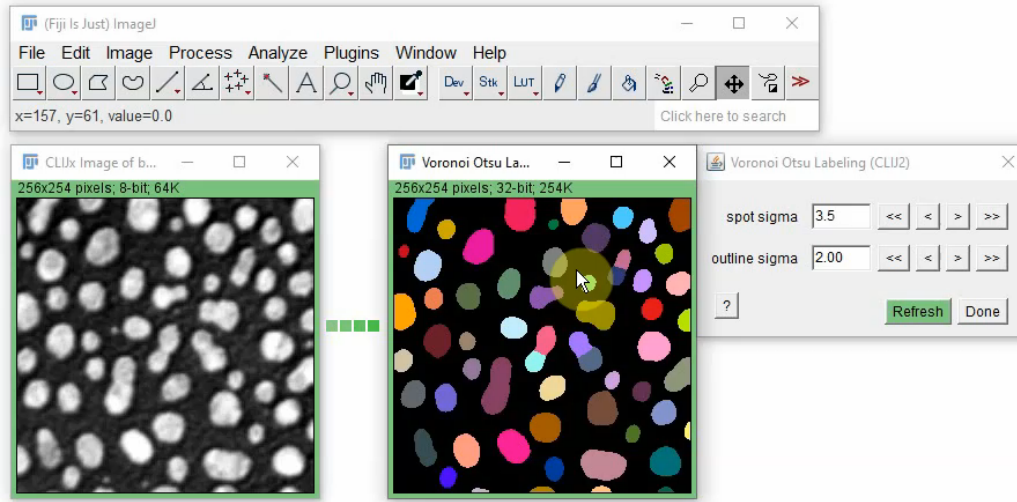
- Should we use a different segmentation algorithm?



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

Hypothesis-driven quantitative biology

Is circularity the right parameter to measure?



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

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

- Question: Which image-derived parameter is influenced when modifying X? Why?

- Sample preparation

- Imaging

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

- Cell segmentation algorithm A, algorithm B, algorithm C

Why?

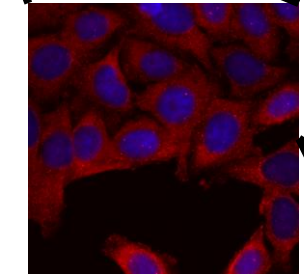
- Measurement of circularity, solidity, elongation, extend, texture, intensity, topology ...

- Statistics

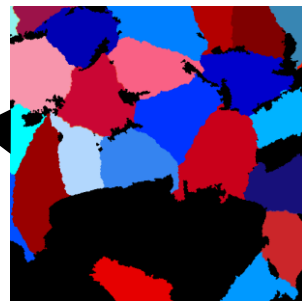
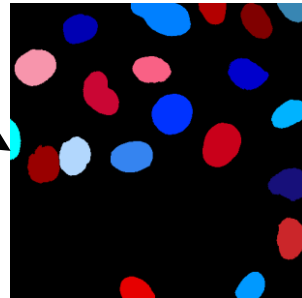
Which parameter shows any relationship with X?

Cell Profiler

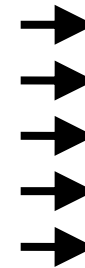
- Cells or organisms in multiwell plates, each well treated with a gene or chemical perturbant



Automated
microscopy
(any manufacturer)



Cell measurements
(size, shape, intensity,
texture, etc.)

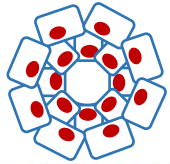


Data exploration
& machine learning

Image analysis beyond segmentation

- Spatial pattern / image analysis for cell biology

Membranes
Nuclei



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 Citations 1.8k Downloads

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

Structural biology of plants¹

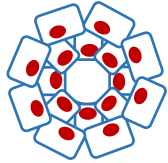


1) Image data source: Pasternak et al. ([CC-BY 4.0](#)) <https://www.biorxiv.org/content/10.1101/2021.01.01.425043v2>
@haesleinupei
@PoLDresden

Image analysis beyond segmentation

- Spatial pattern / image analysis for cell biology

Membranes
Nuclei

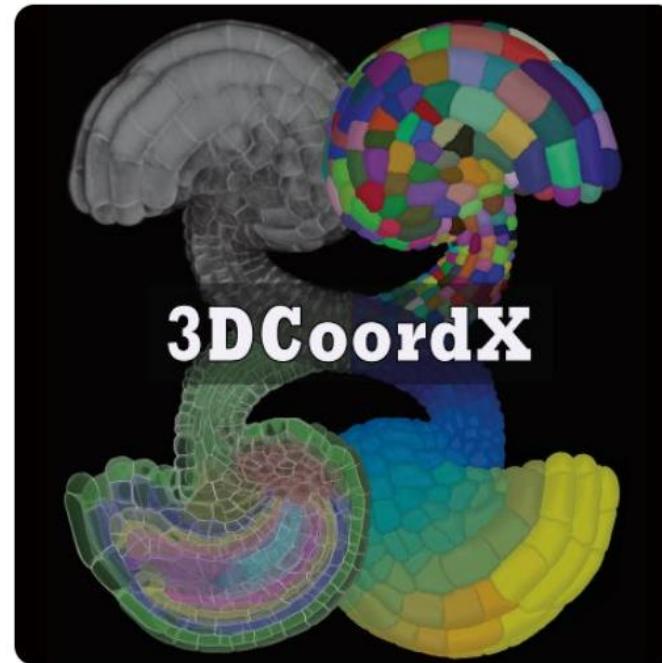


Kay Schneitz @PlantDevTUM · Mar 29

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

@athul_r_v @StraussSoe @tejasvinee_m @kareningleee @RichardSmithLab @PlantPhys



21 76

Plant Physiology[®] American Society of Plant Biologists

Issues Advance Articles More Content ▾ Submit ▾ Purchase About ▾ All Plant Physiology ▾ 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 ▾

The Plant Cell Atlas
Plant Physiology
Advertisement

Structural biology of plants¹

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



Cell Reports
Resource

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

Caleb R. Stoltzfus,¹ Jakub Filipiak,¹ Miranda R. Lyons-Cohen,¹ Jesse Kevin B. Urdahl,^{1,2,3} Mario Perle,¹ Department of Immunology, University of Washington, Seattle, Washington, USA
²Seattle Children's Research Institute, Seattle, Washington, USA
³Department of Pediatrics, University of Washington, Seattle, Washington, USA
⁴School of Medicine, University of Washington, Seattle, Washington, USA
⁵Roche Innovation Center Munich, Penzberg, Germany
⁶Roche Innovation Center Zurich, Zurich, Switzerland
⁷Lead Contact
*Correspondence: gemermy@uw.edu
<https://doi.org/10.1016/j.celrep.2021.10.016>

nature methods

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [nature methods](#) > [articles](#) > article

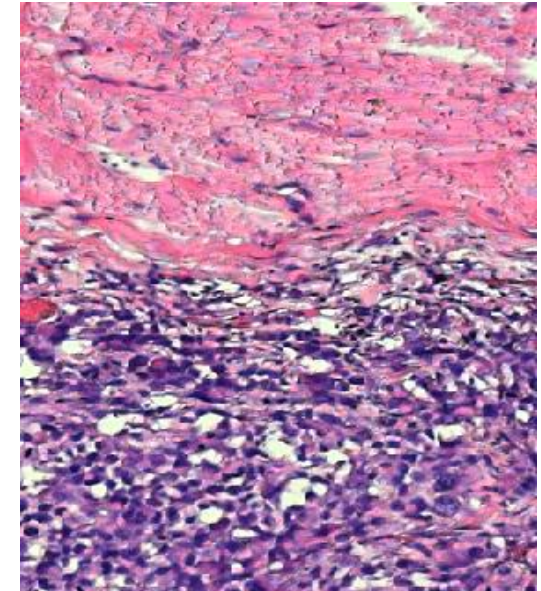
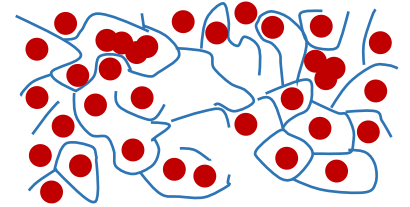
Article | Open Access | Published: 31 January 2022

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

Structural biology of plants¹

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

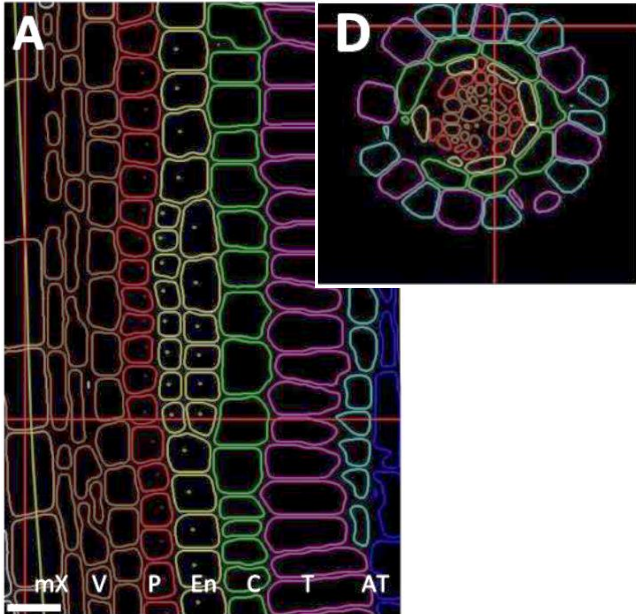
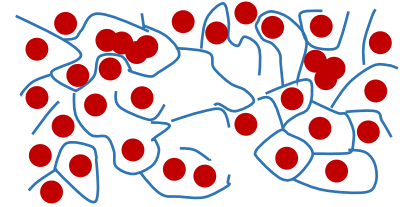
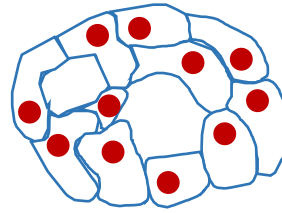
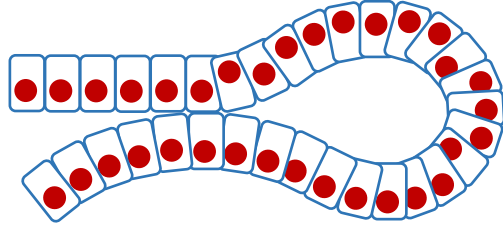
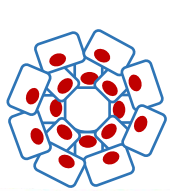
2) [https://www.cell.com/cell-reports/fulltext/S2211-1247\(20\)30423-X](https://www.cell.com/cell-reports/fulltext/S2211-1247(20)30423-X)

3) Solorzano et al. <https://www.nature.com/articles/s41592-021-01358-2>

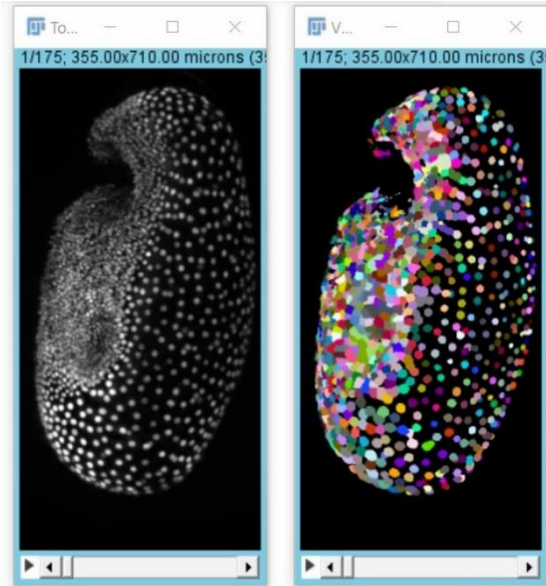
Image analysis beyond segmentation

- Spatial pattern / image analysis for cell biology

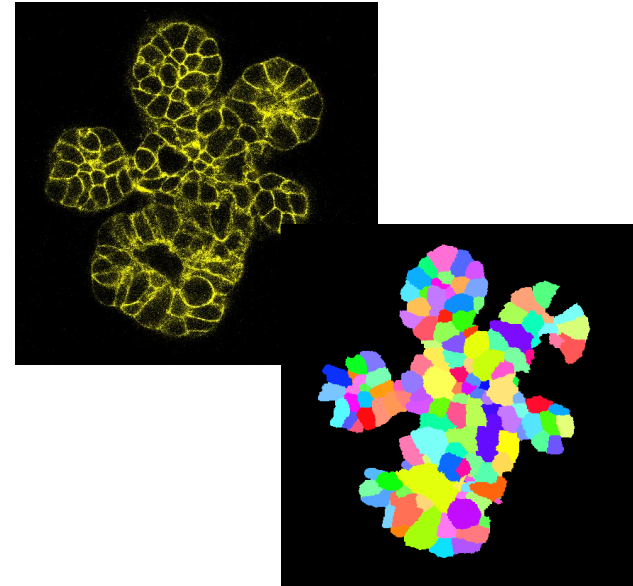
Membranes
Nuclei



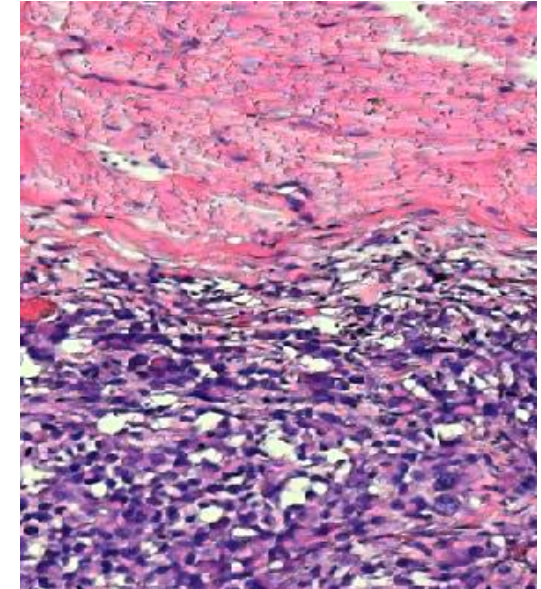
Structural biology of plants¹



Early embryo development



Organoid formation³

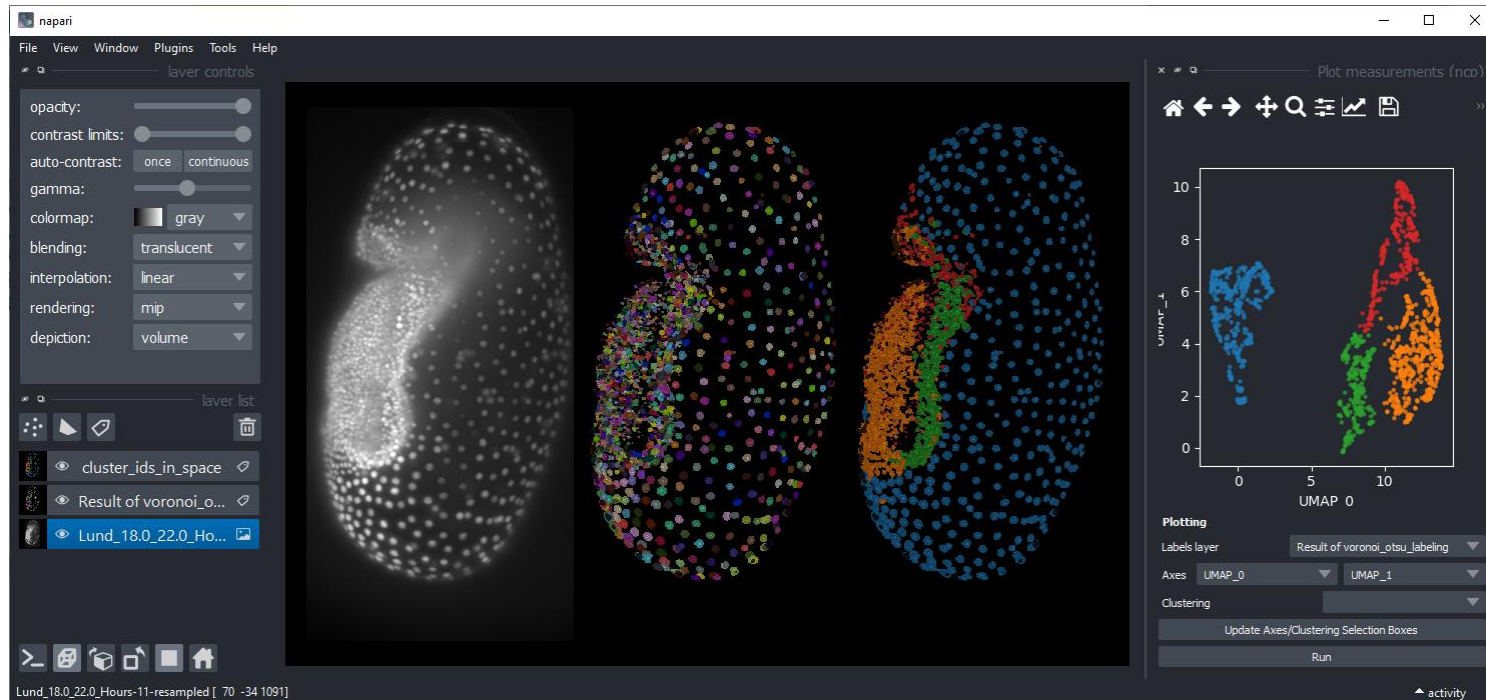


Cancer research / histology

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

3) Image data source: ByungHo Lee, Grapin-Botton Lab

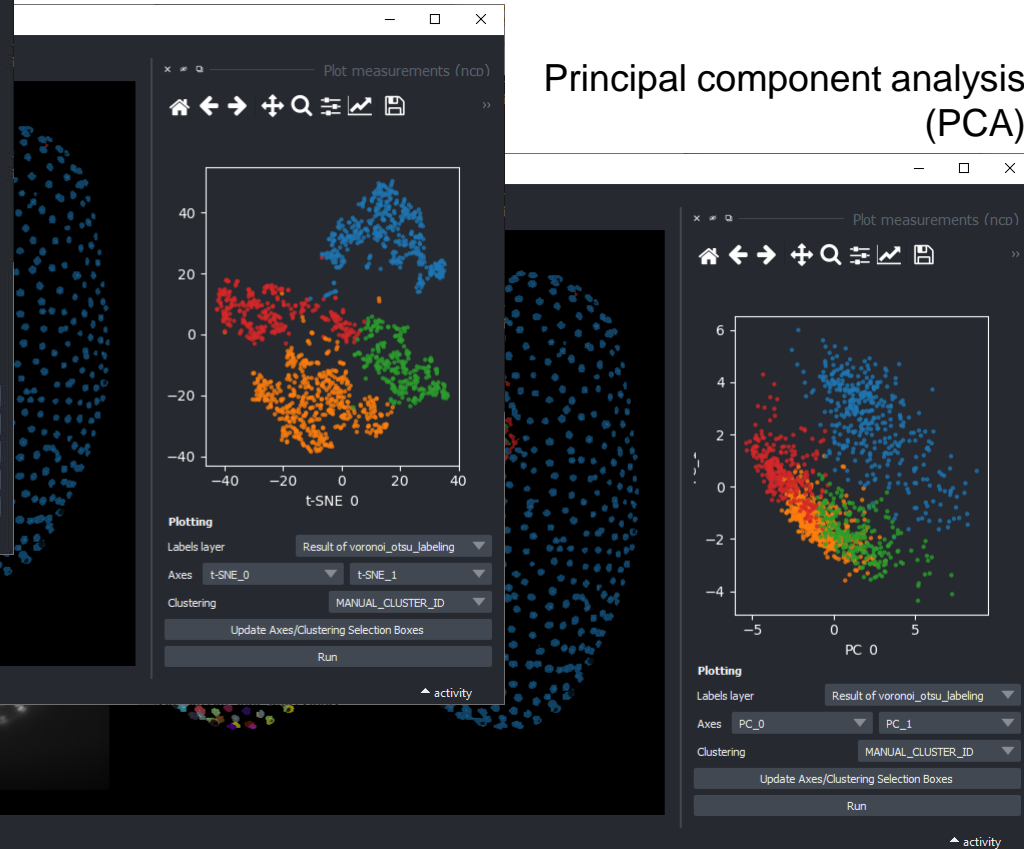
Dimensionality reduction



Uniform manifold approximation and projection (UMAP)

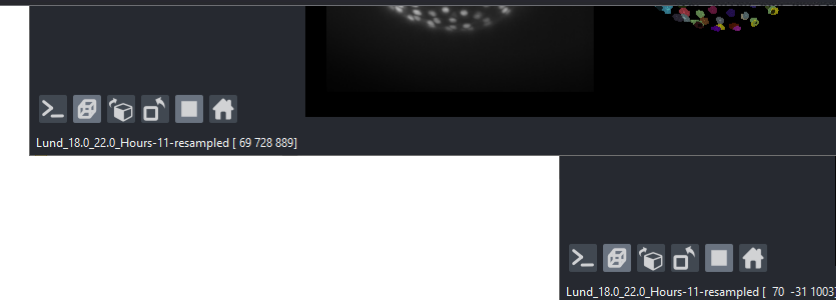
t-distributed stochastic neighbor embedding (t-SNE)

Principal component analysis (PCA)



Laura Žigutytė
@zigutyte

Ryan Savill
@RyanSavill4



Lund_18.0_22.0_Hours-11-resampled [70 -31 1003]

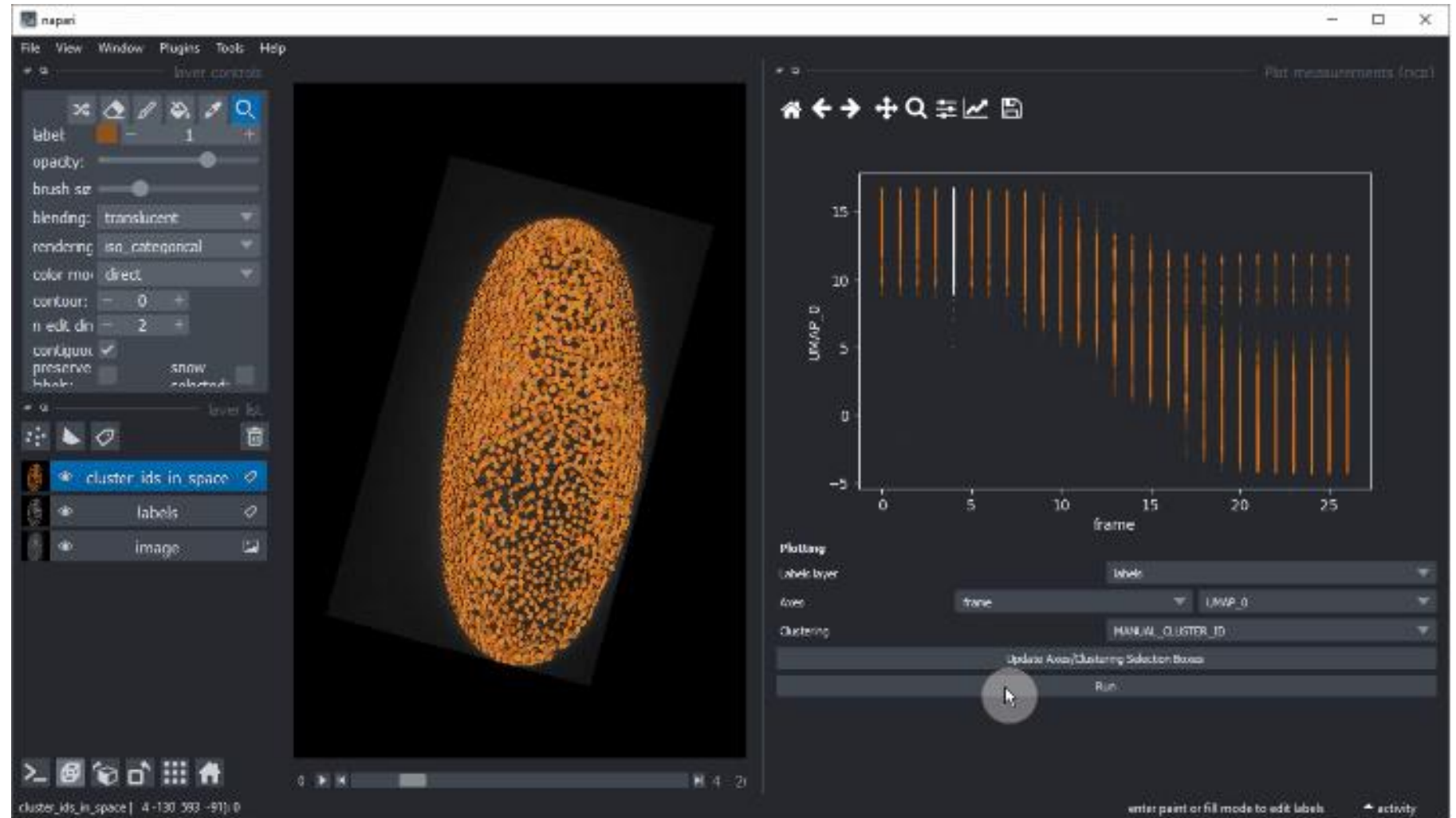


@haesleinhuepf
@PoLDresden

<https://github.com/BiAPoL/napari-clusters-plotter>

Data exploration

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



Laura Žigutytė
@zigutyte

Ryan Savill
@RyanSavill4

Marcelo Zoccoler
@zoccolermarcelo

<https://github.com/BiAPoL/napari-clusters-plotter>

Image data source: Daniela Vorkel, Myers lab, MPI-CBG/CSBD

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

biapol@groups.tu-dresden.de
ipf@mpi-cbg.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