Data science in cell imaging

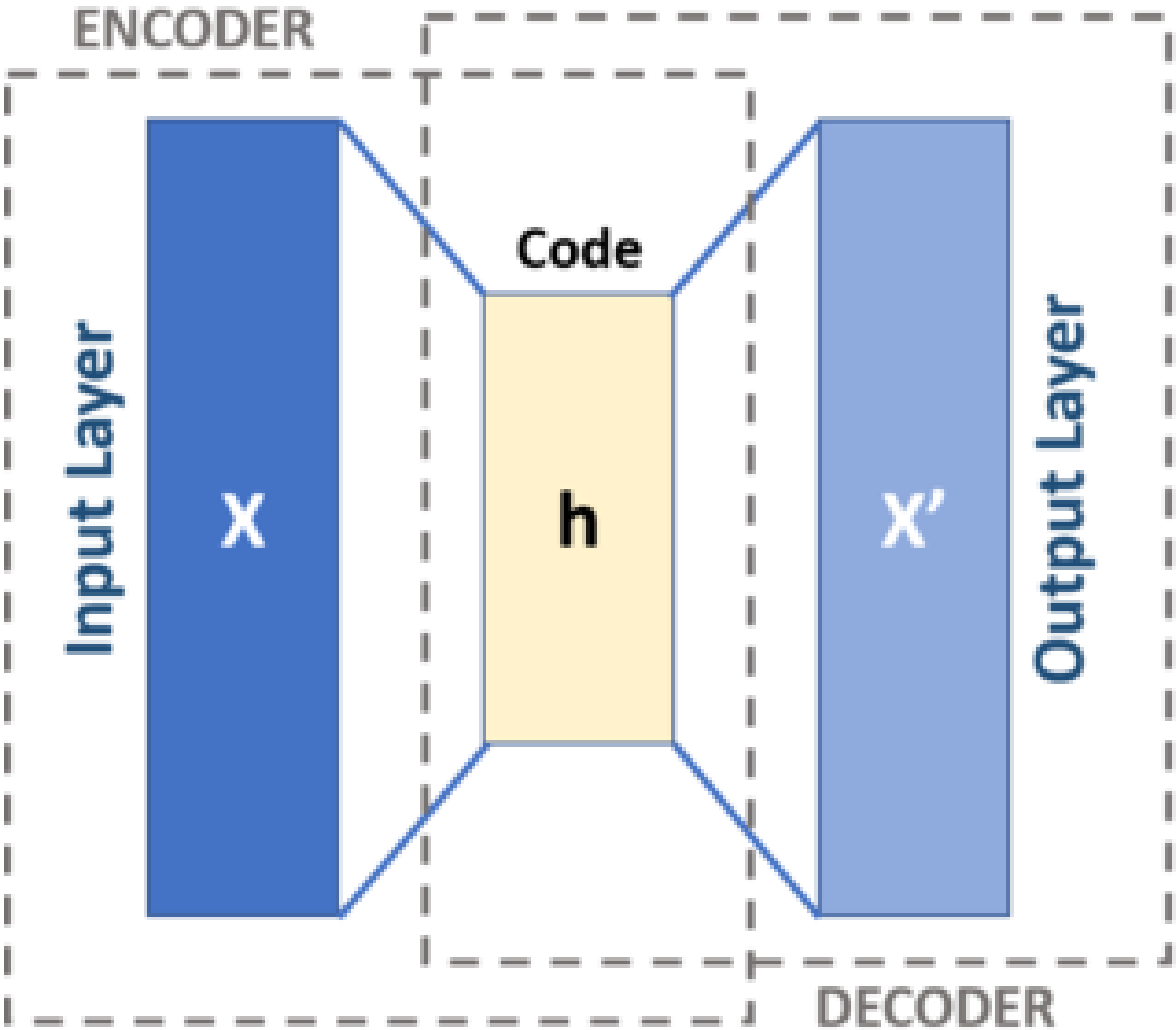
Lecture 6: deep learning in microscopy

“The Great Wave off Kanagawa”, by Hokusai, ~1830 (Source: Wikipedia)

Last time

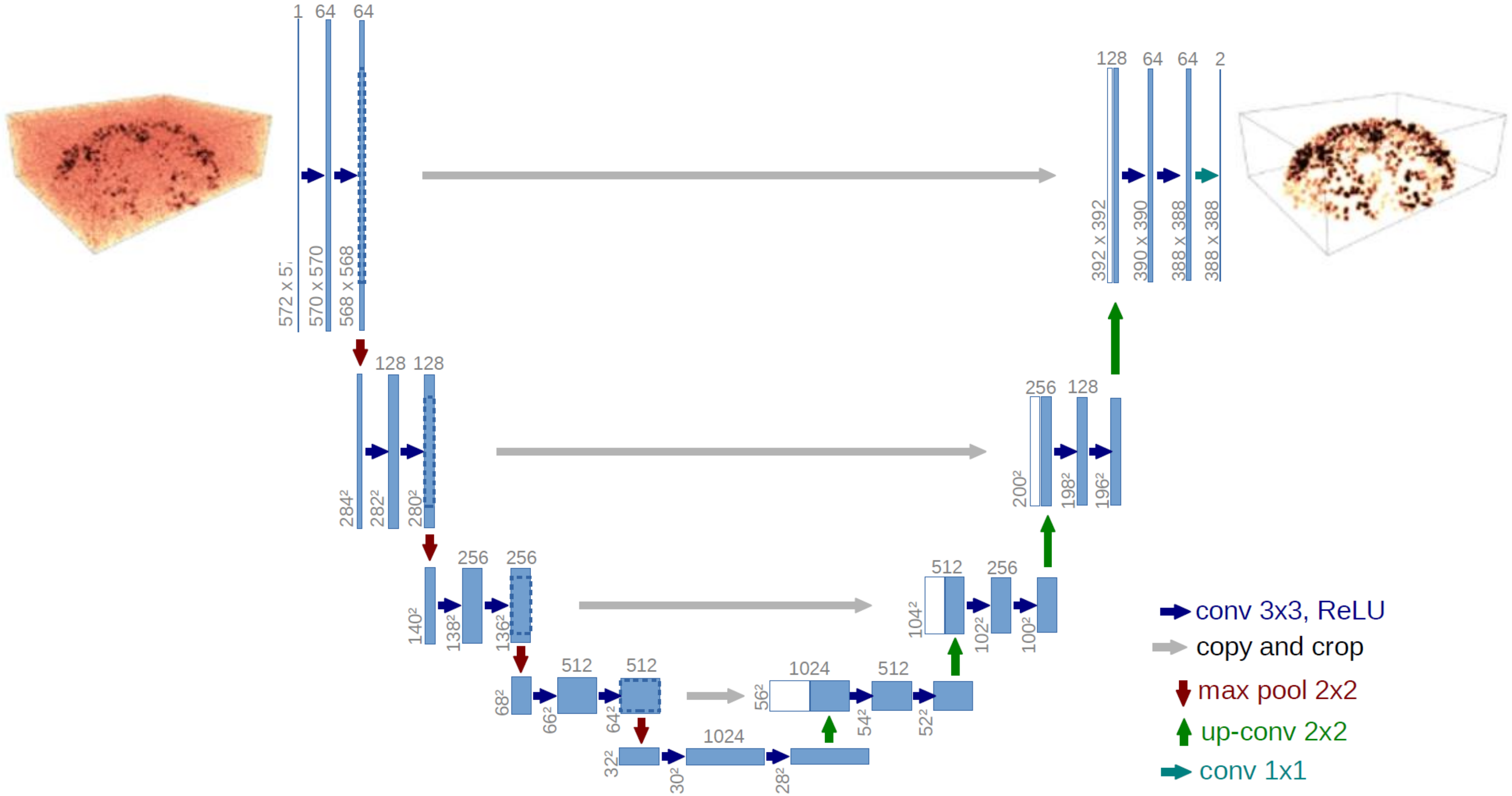
Enhancing cell image quality with deep learning

Autoencoder: unsupervised data encoding



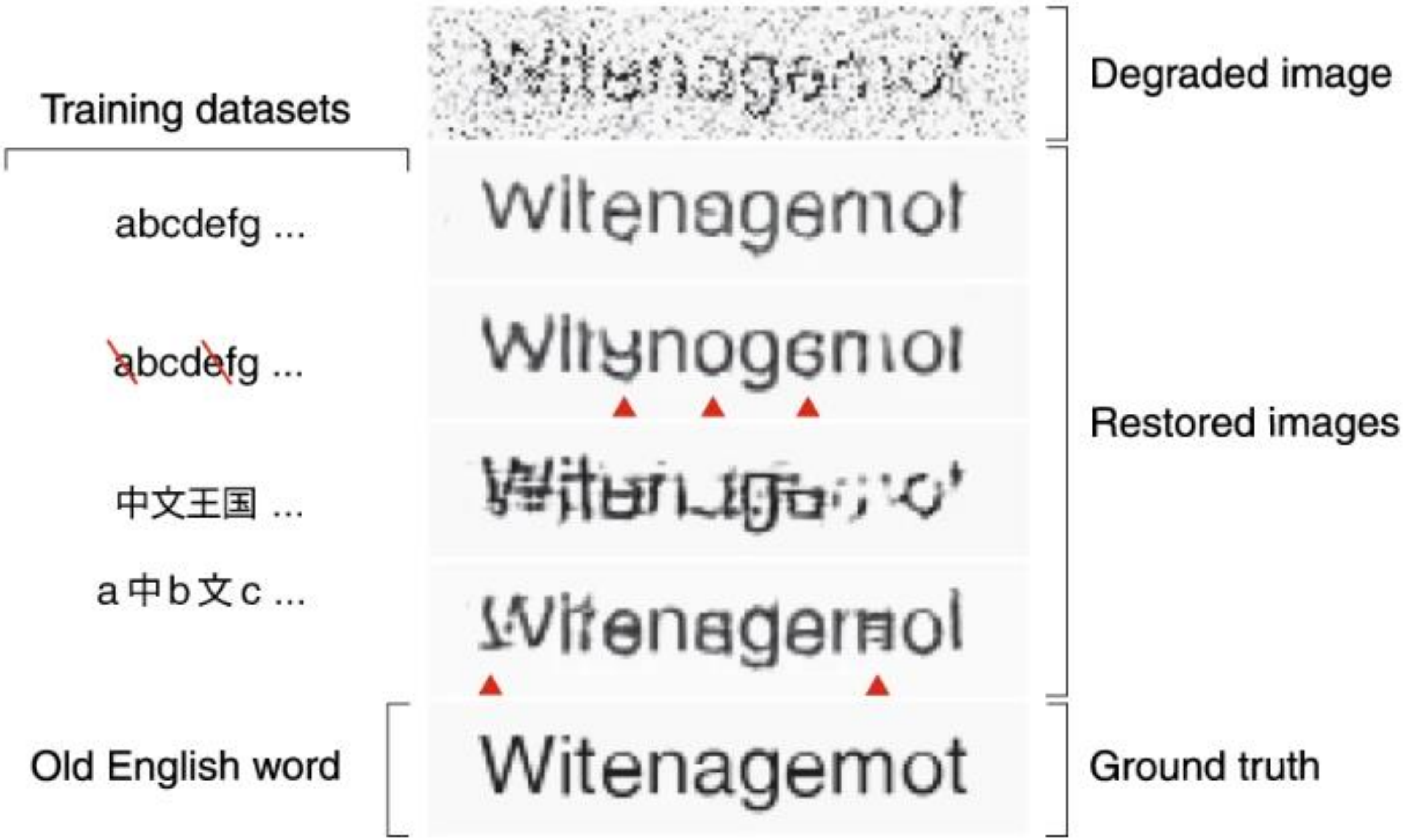
Wikipedia

The machinery: U-Net



Ronneberger et al. (2015)

The hallucination problem



Belthangady and Royer, 2019

Summary

- Leveraging experiment-specific information
- Matched images, semi-synthetic / synthetic training data
- High quality models can be trained without the availability of clean ground truth data
- Great for downstream analysis not for intensity-based measurements!
- Could work well “out of the box”
- CARE, ANNA-PALM and others..
- Similar ideas in computer vision (“super resolution”)
- Recommended Perspective, Belthangady & Royer (2019)

Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction

Chinmay Belthangady and Loic A. Royer*

Today and next week

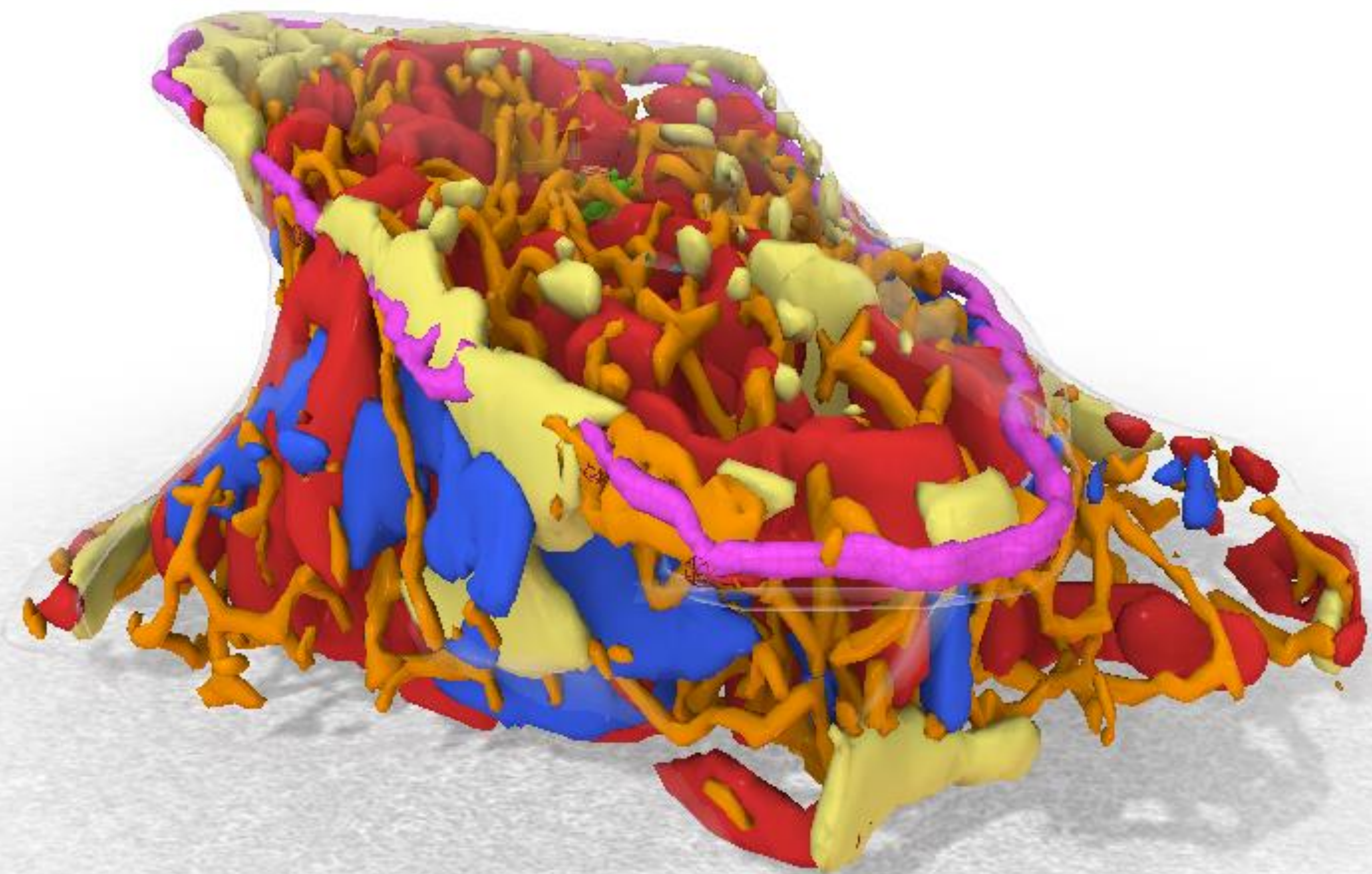
- Generative models for cell structure with deep learning
- Classifying cell state with deep learning

Guest lectures to come

- 13.5 Tammy Riklin Raviv, EE BGU, on computer vision in cell imaging
- 20.5 Kota Miura, NEUBIAS, on bioimage analysis (English)
- 27.5, Yaron Gurovich on “Identifying facial phenotypes of genetic disorders using deep learning”
- 3.6 Tal Shay, Life Science BGU, on systems biology

Look at a cell and know what it is doing

What it did



What it will do

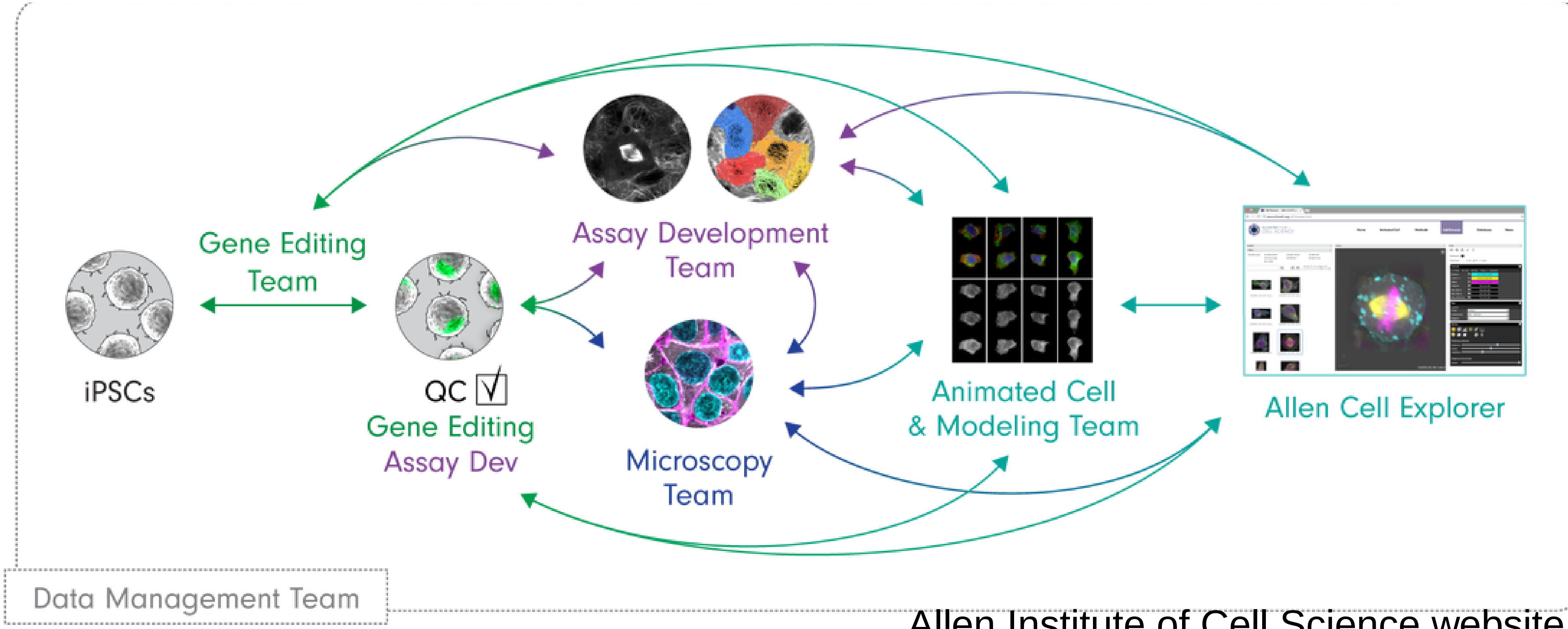
Structure ↔ Function

Slide adapted from Susanne Rafelski, Allen Institute of Cell Science

Allen institute of cell science

Our Mission

The mission of the Allen Institute for Cell Science is to create dynamic and multi-scale visual models of cell organization, dynamics and activities that capture experimental observation, theory and prediction to understand and predict cellular behavior in its normal, regenerative, and pathological contexts.



Allen Institute of Cell Science website

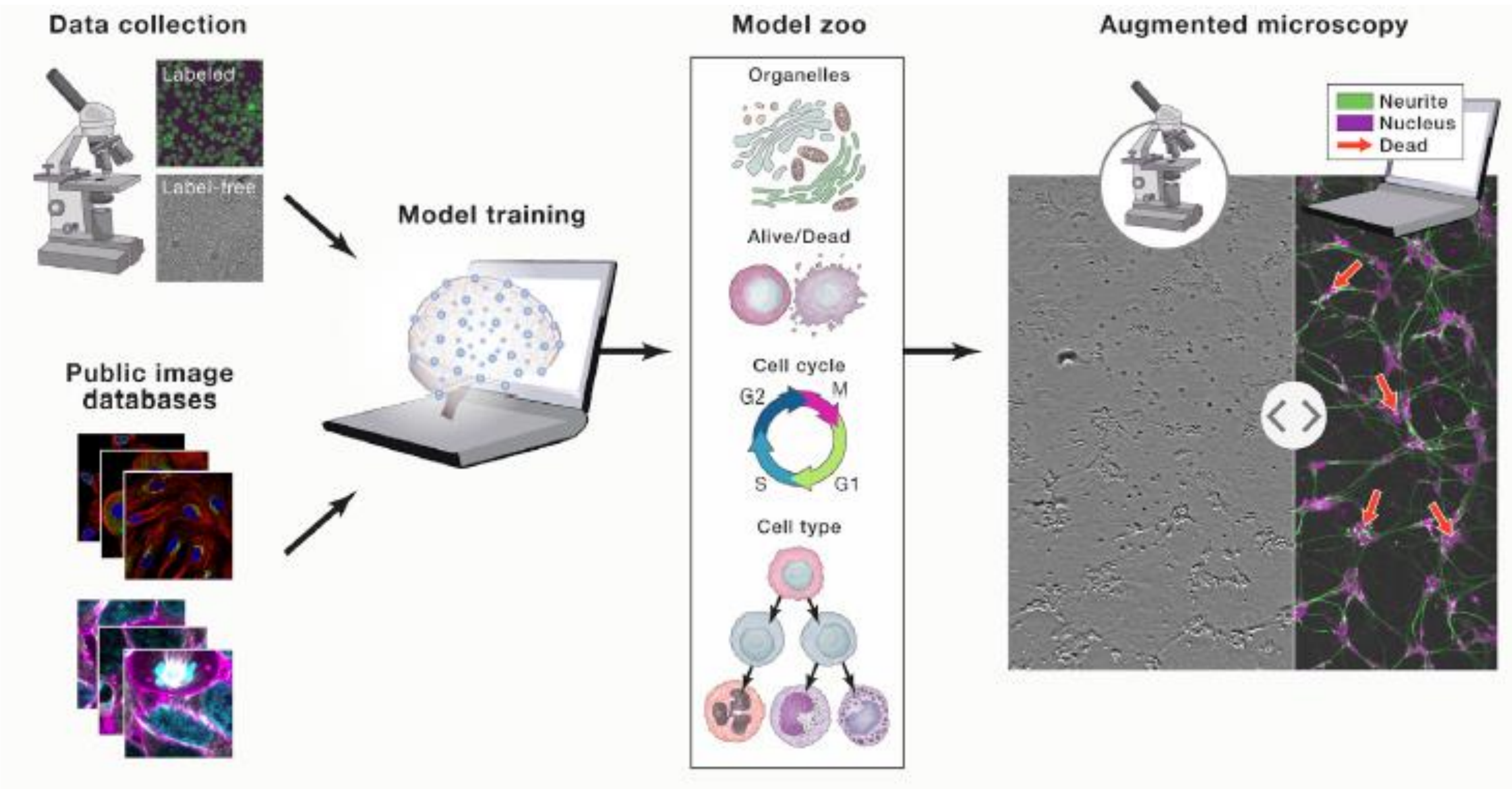
Allen institute of cell science

- Overview https://www.allencell.org/what-we-do.html
- Visualization https://www.allencell.org/visual-guide-to-human-cells.html
- 3D cell viewer https://www.allencell.org/3d-cell-viewer.html
- Cell feature explorer https://bit.ly/355Lq1j
- Publicly available cell lines, tools, data, code!
- Research projects https://www.allencell.org/cell-research-projects.html

The holy grail!

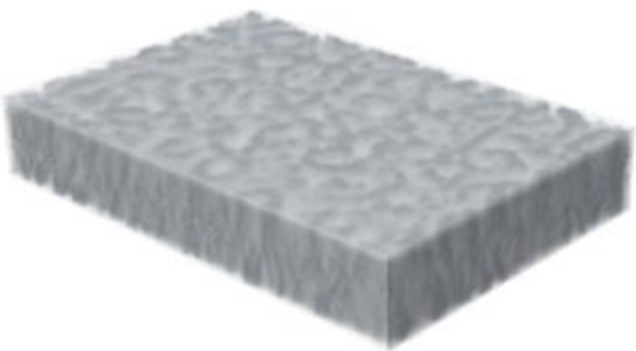
Can we train a generative model for accurate fluorescent imaging from label-free (transmitted light) imaging?

A future of augmented microscopy



Sullivan and Lundberg (2018)

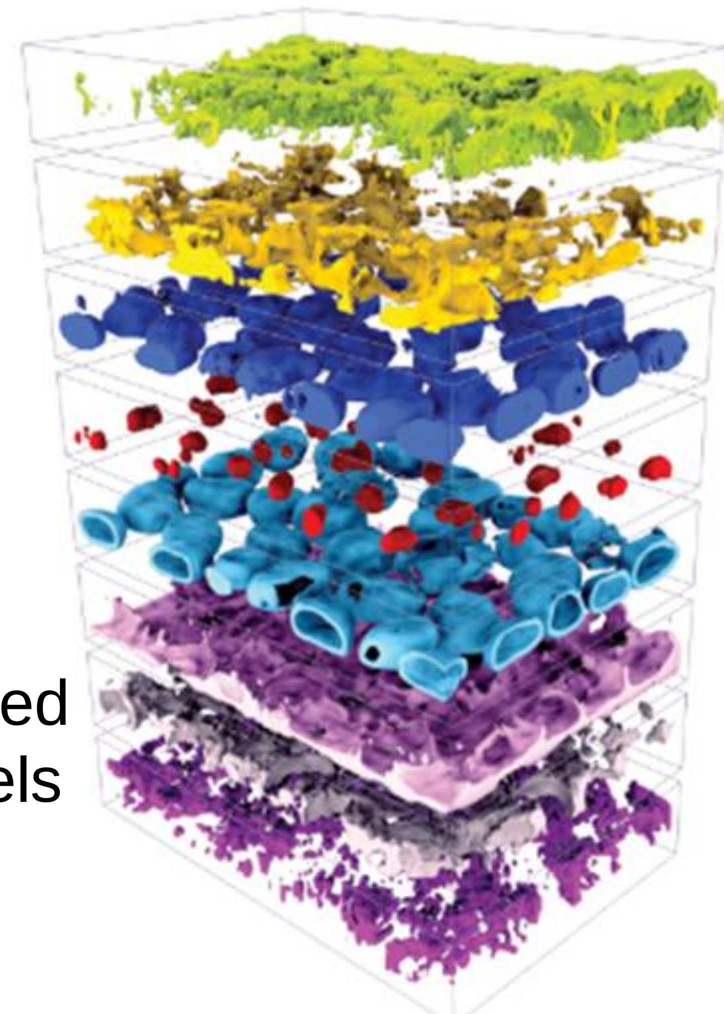
Label-free images contain information on the molecular organization of the cell!



Bright field

Predictive models

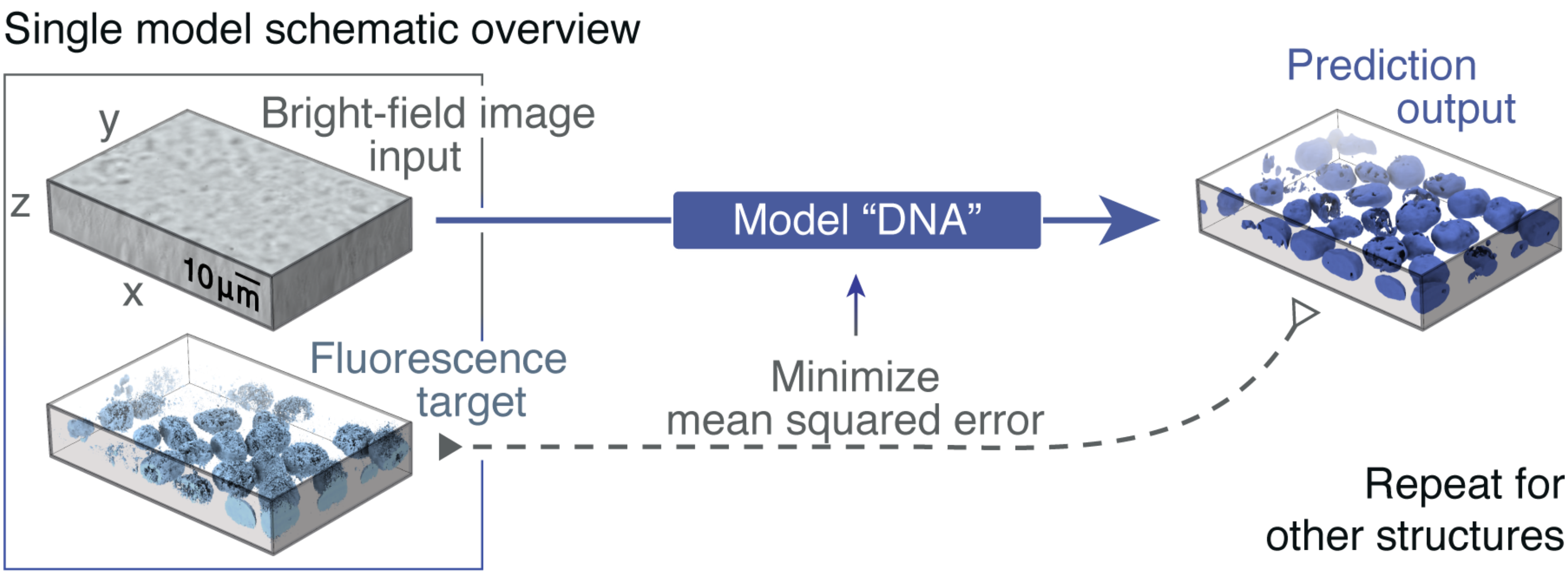
Predicted channels



Ounkomol et al. (2018)

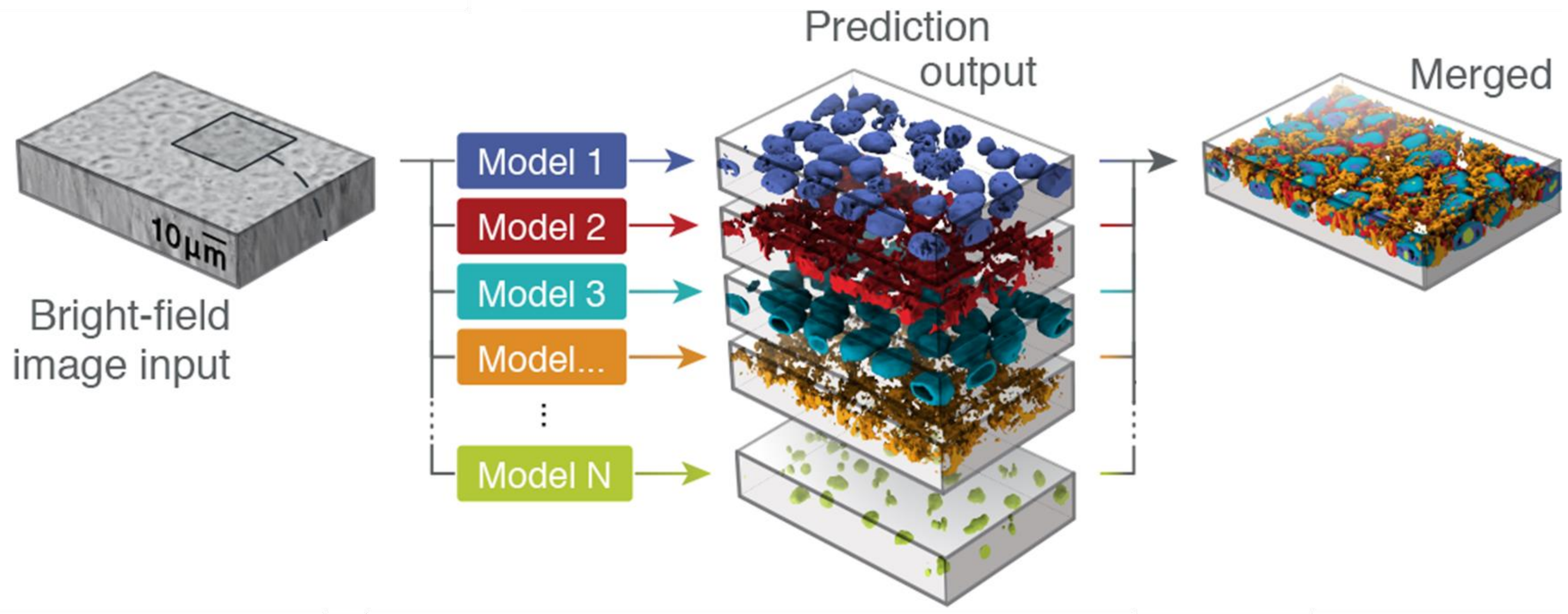
Christiansen et al. (2018)

Unstructured-to-structured information with supervised models



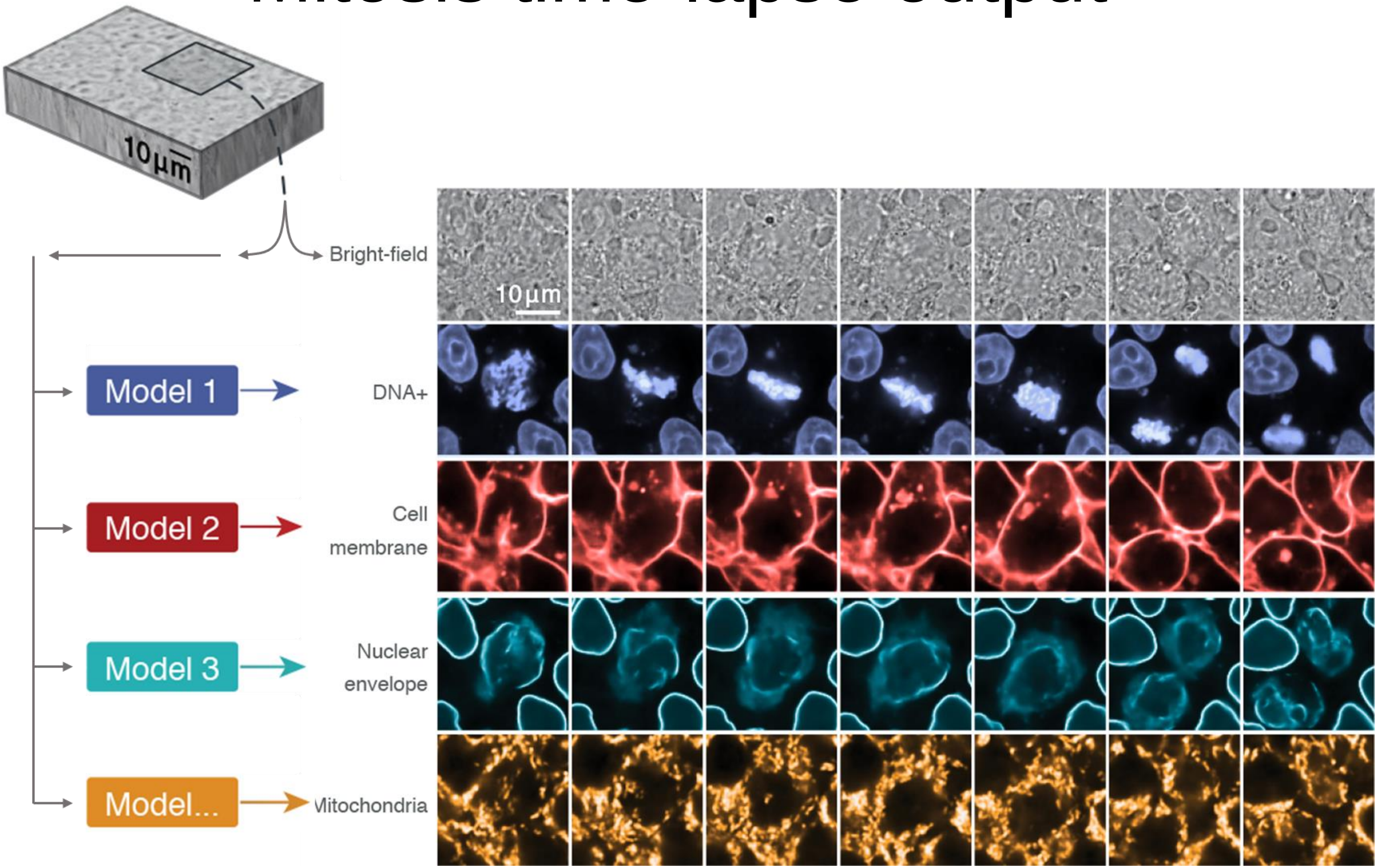
Ounkomol et al. (2018)

Combining multiple models



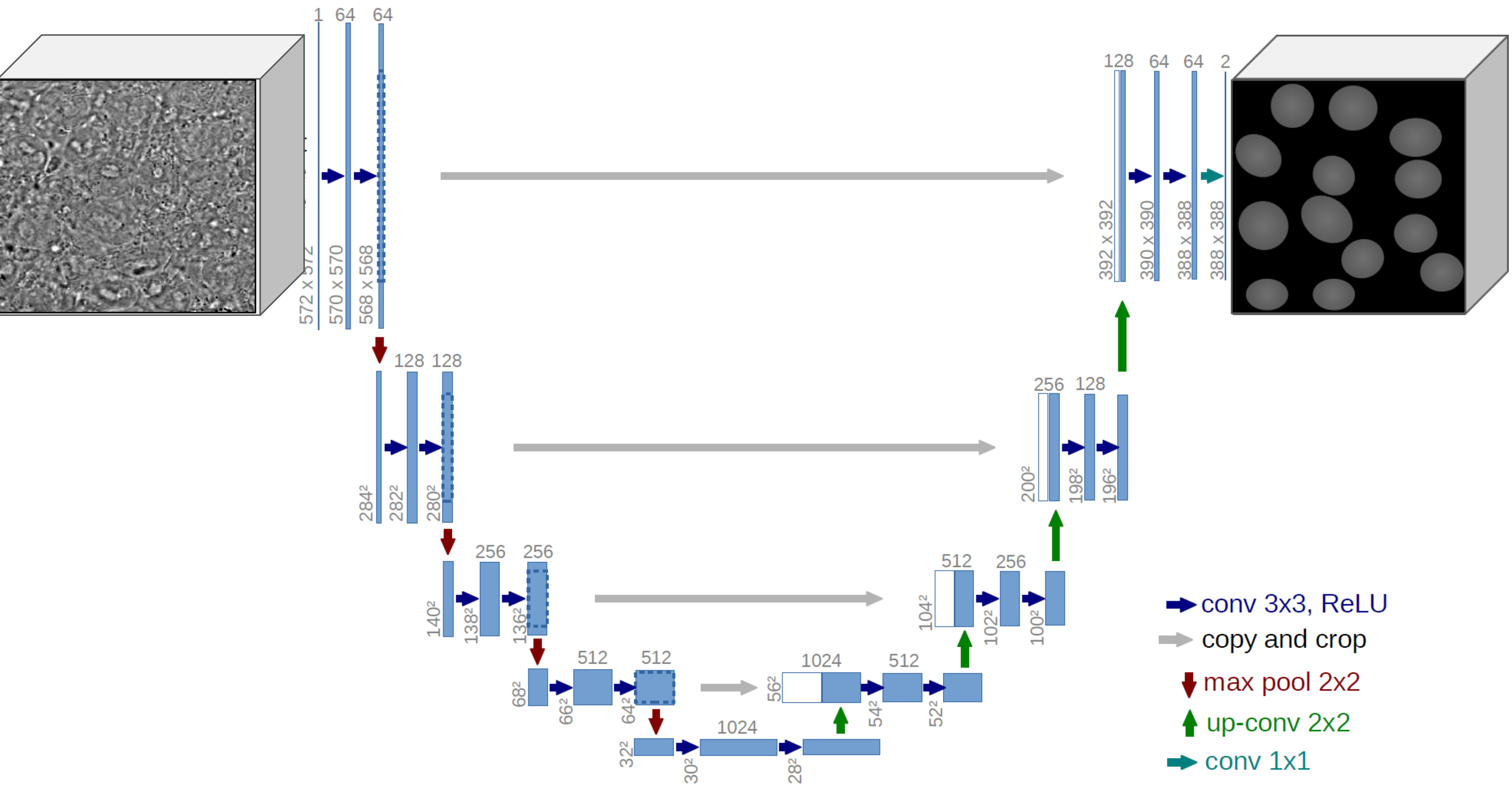
Ounkomol et al. (2018)

Mitosis time-lapse output



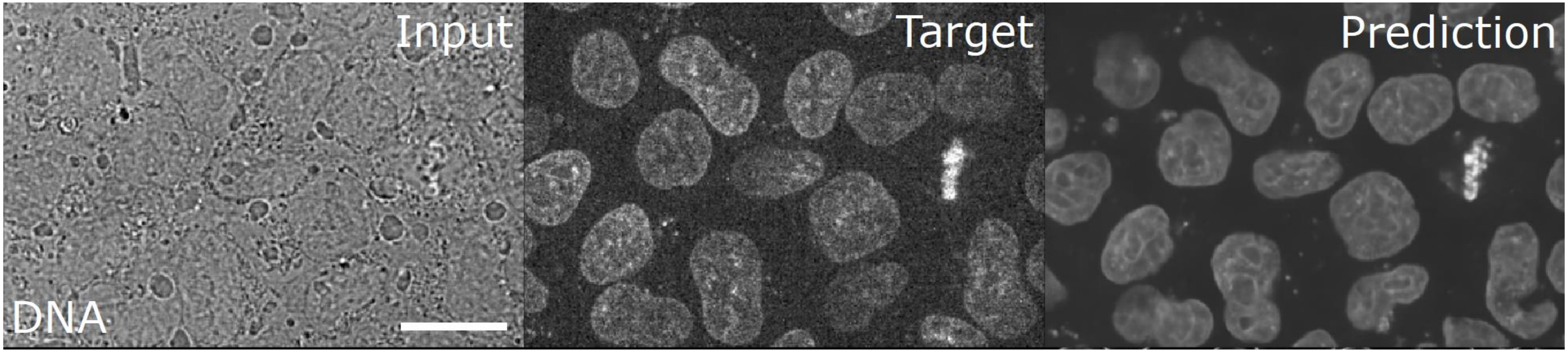
Ounkomol et al. (2018)

U-Net architecture



Greg Johnson, Ronneberger et al. (2015)

Predicting DNA localization from transmitted light

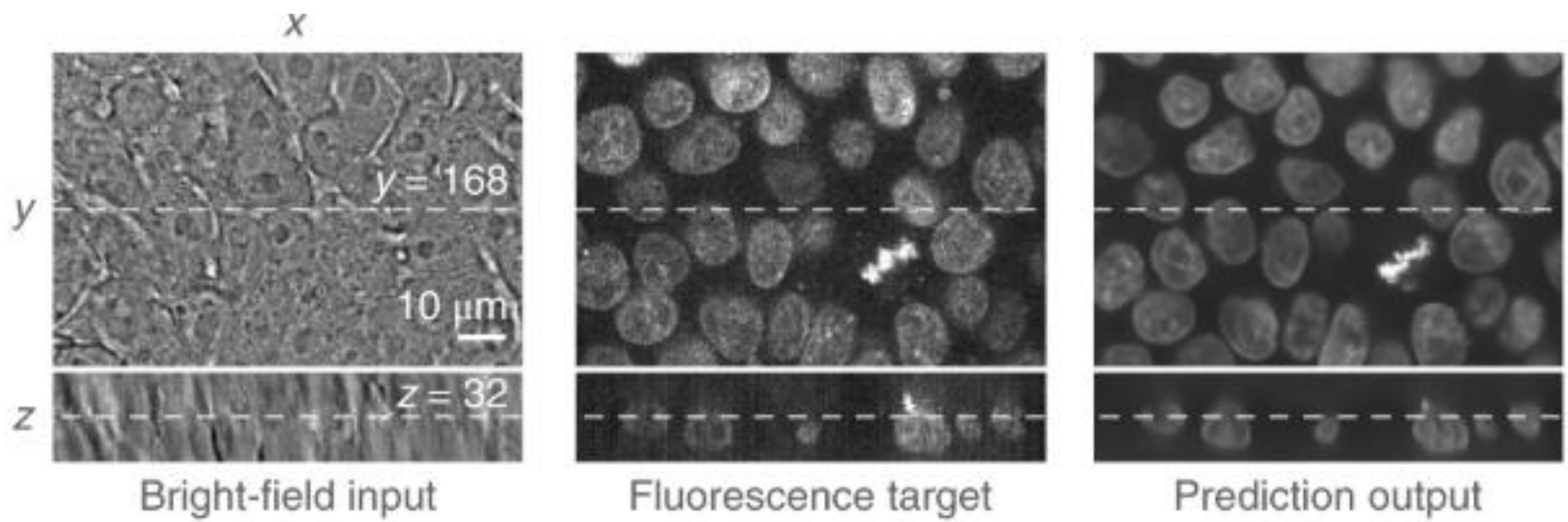


Good pixel-wise correlation between (3D) "true" and predictions. Less noise in predictions.

(assessing and improving similarity/distance measures could be a course project)

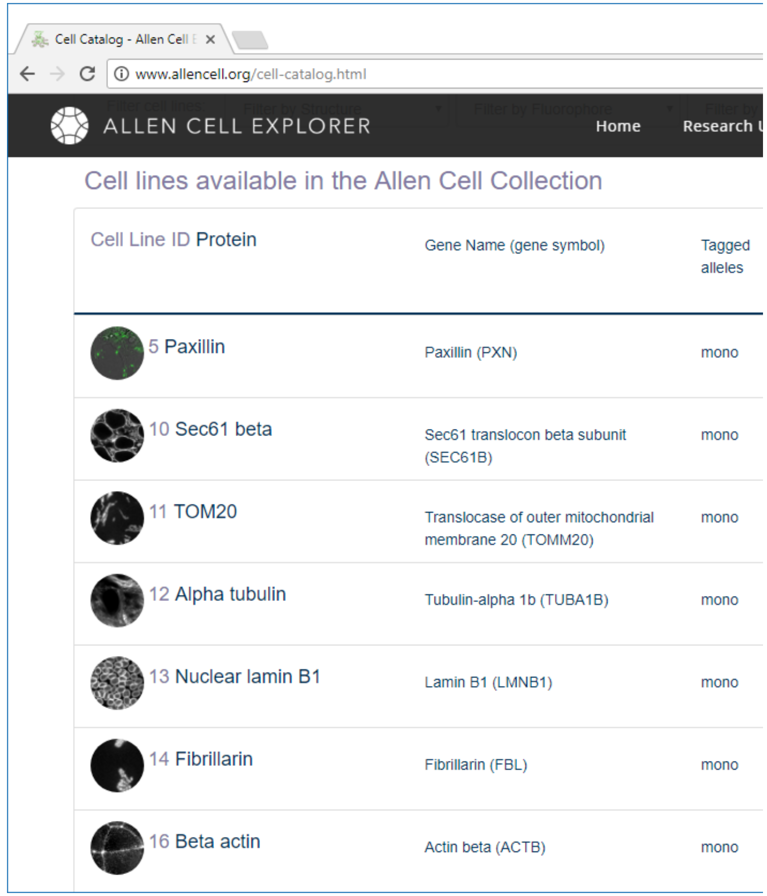
Greg Johnson

Predictions are 3D

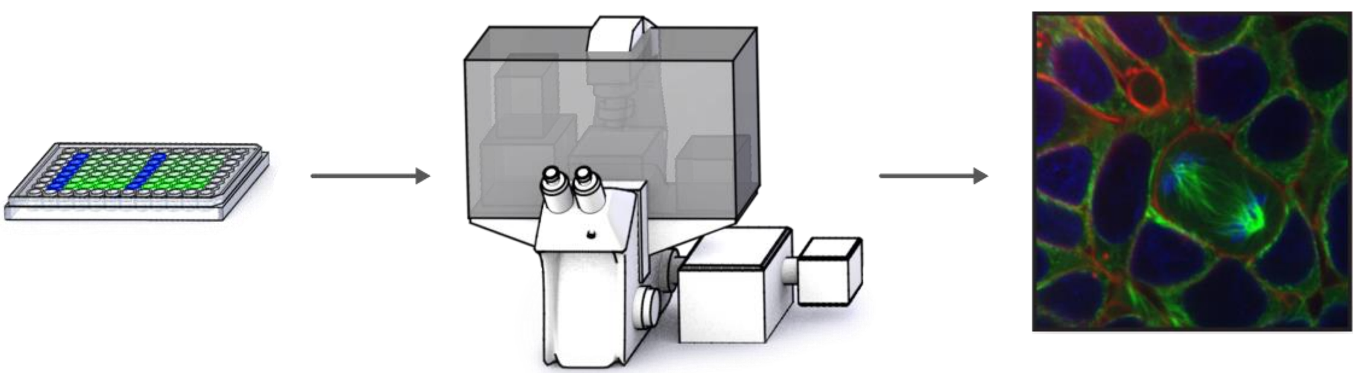


Greg Johnson

What about other structures?

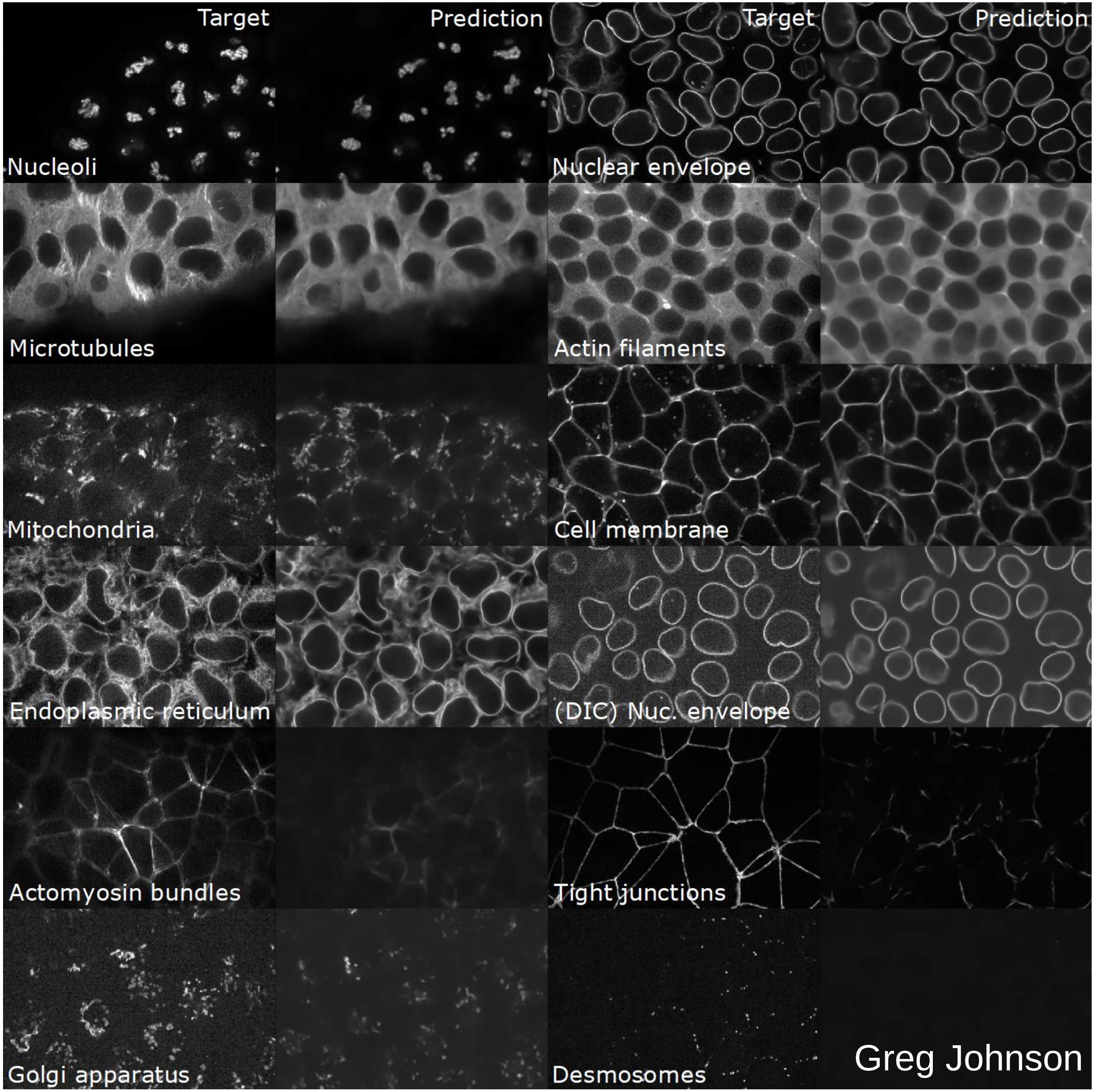


Cell Line ID Protein	Gene Name (gene symbol)	Tagged alleles
5 Paxillin	Paxillin (PXN)	mono
10 Sec61 beta	Sec61 translocon beta subunit (SEC61B)	mono
11 TOM20	Translocase of outer mitochondrial membrane 20 (TOMM20)	mono
12 Alpha tubulin	Tubulin-alpha 1b (TUBA1B)	mono
13 Nuclear lamin B1	Lamin B1 (LMNB1)	mono
14 Fibrillarin	Fibrillarin (FBL)	mono
16 Beta actin	Actin beta (ACTB)	mono



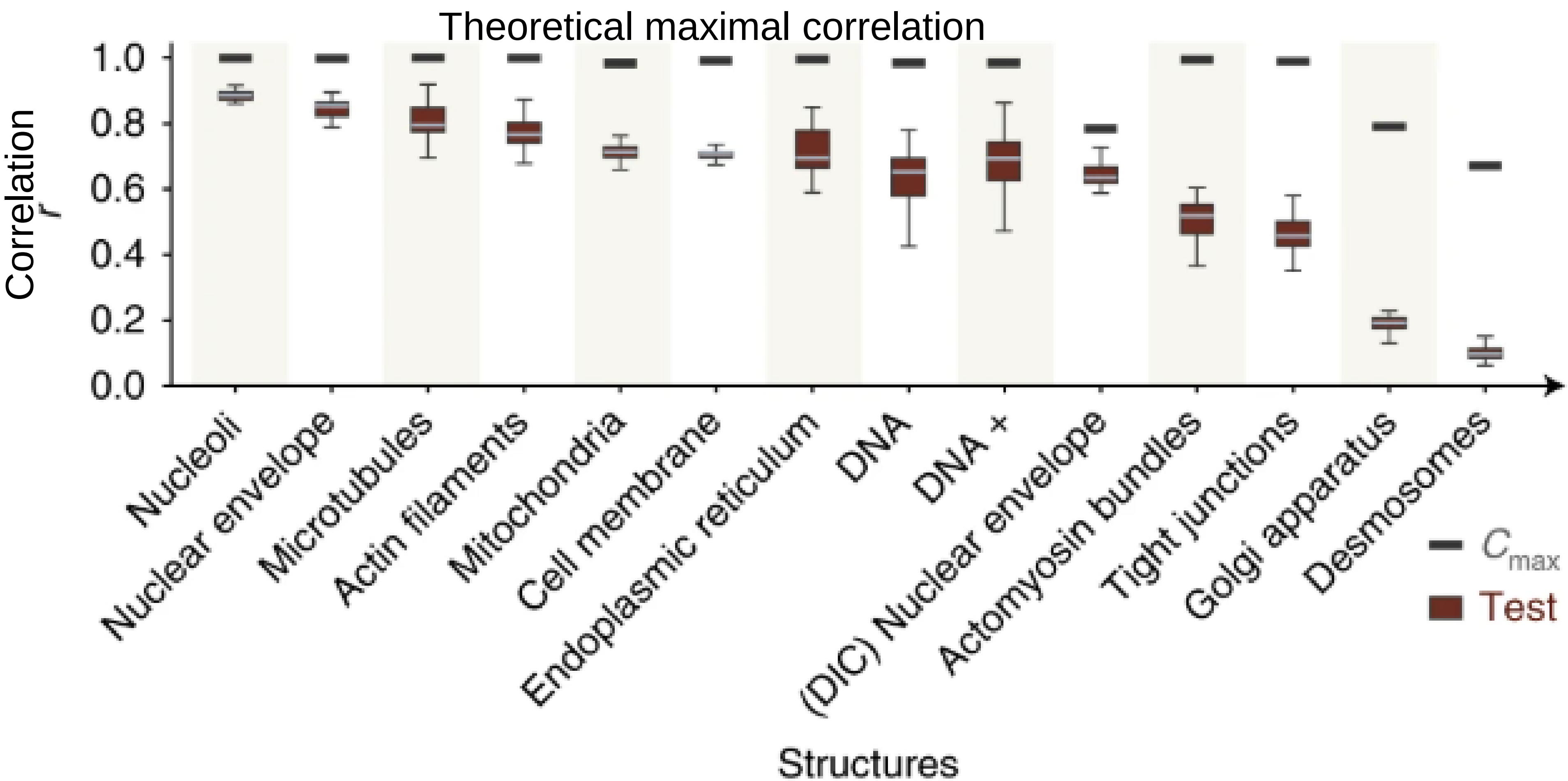
Greg Johnson

Results



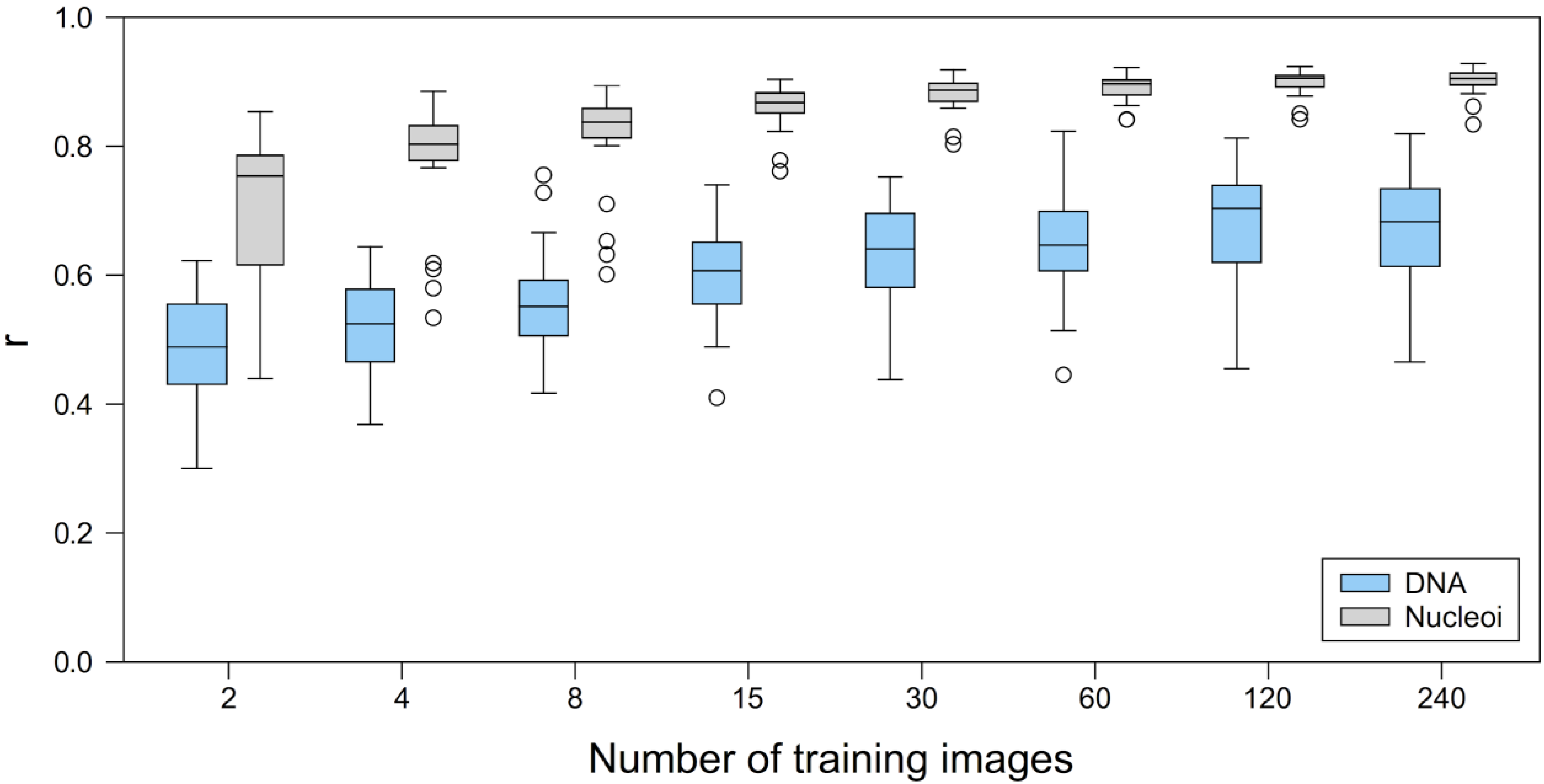
Greg Johnson

Predictions performance



Greg Johnson

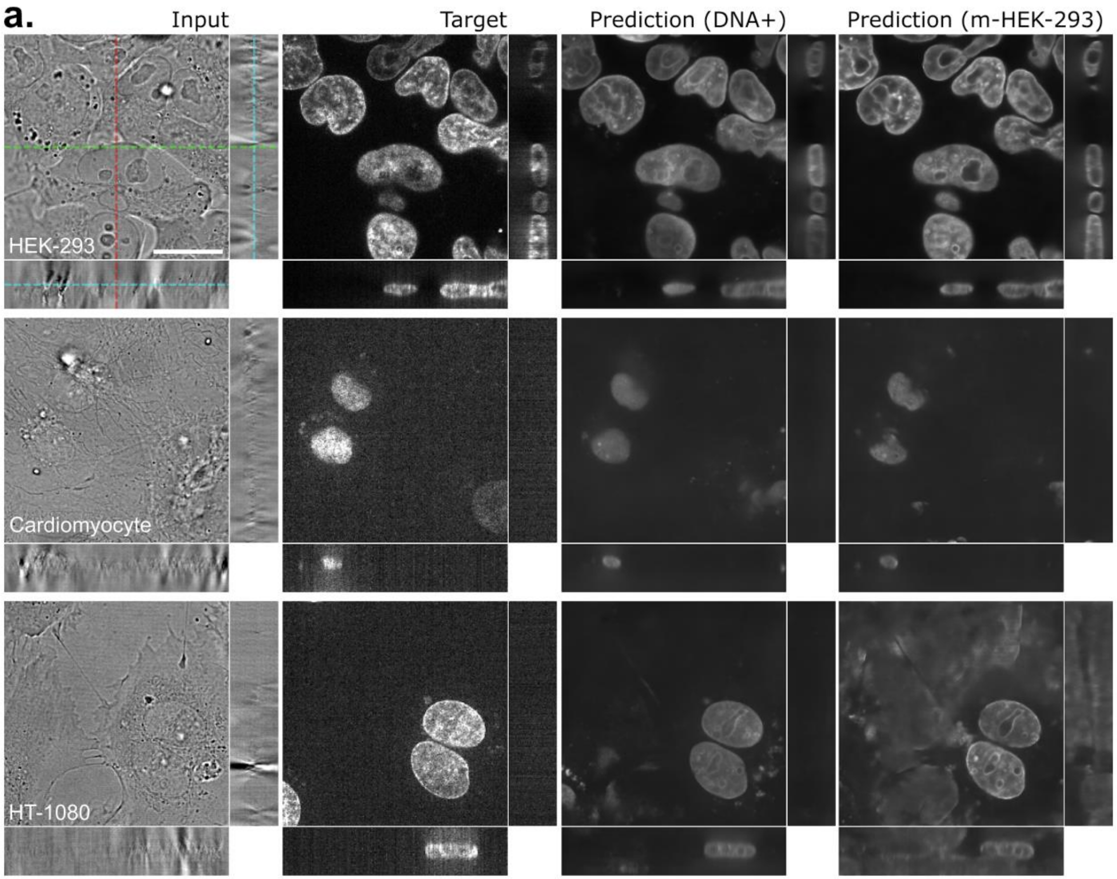
Models can be trained with fairly small datasets

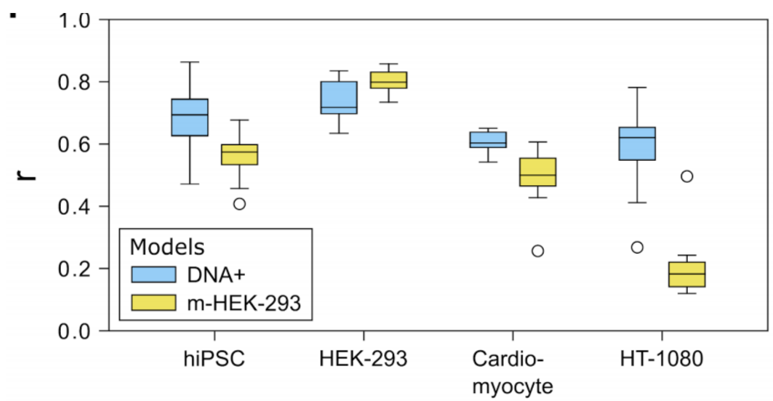


Greg Johnson

Generalization to other cell types

Always perform best for data similar to the training



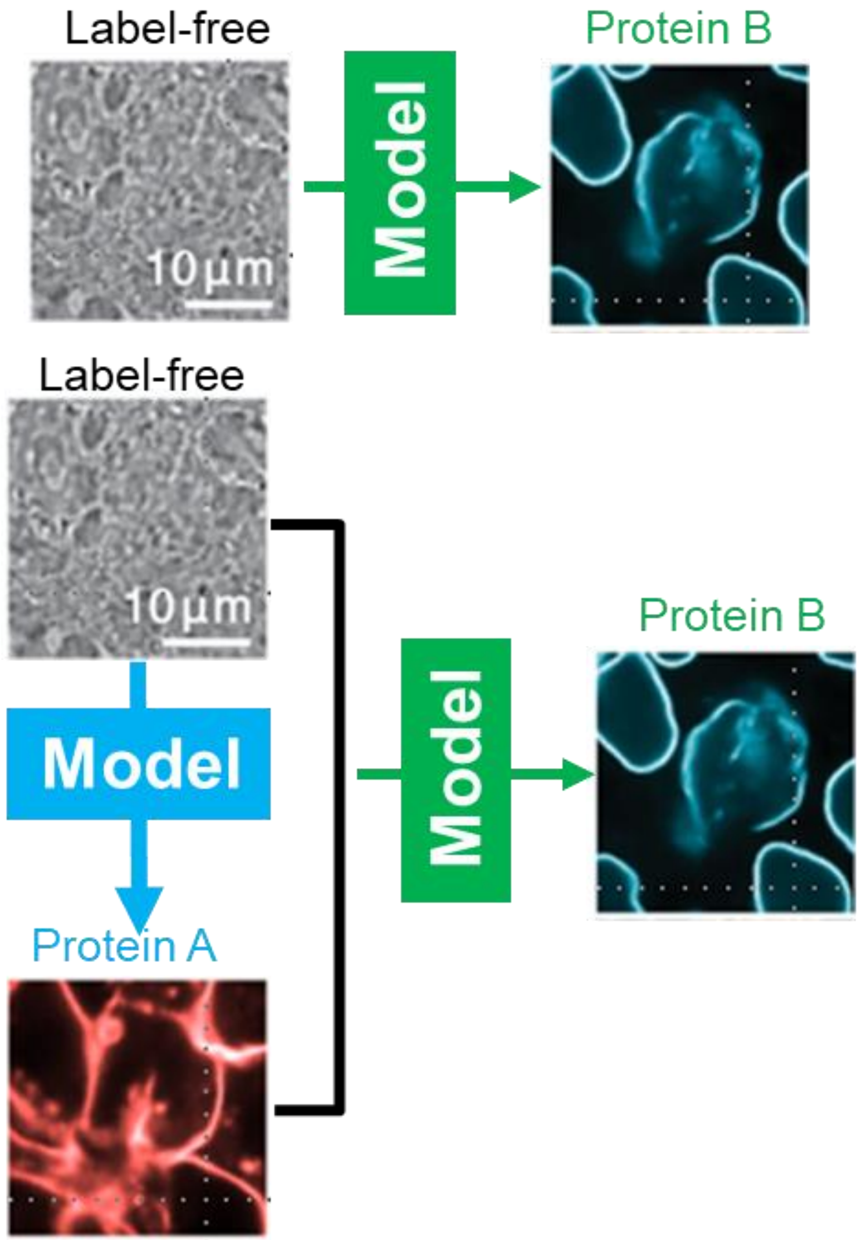


Greg Johnson

Summary

- Great for hypothesis generation!
- Be very careful regarding “hallucinated” pixel intensities
- Generalization? effects of perturbations / different cell systems / imaging

Follow-up project in the lab



Katya Smoliansky

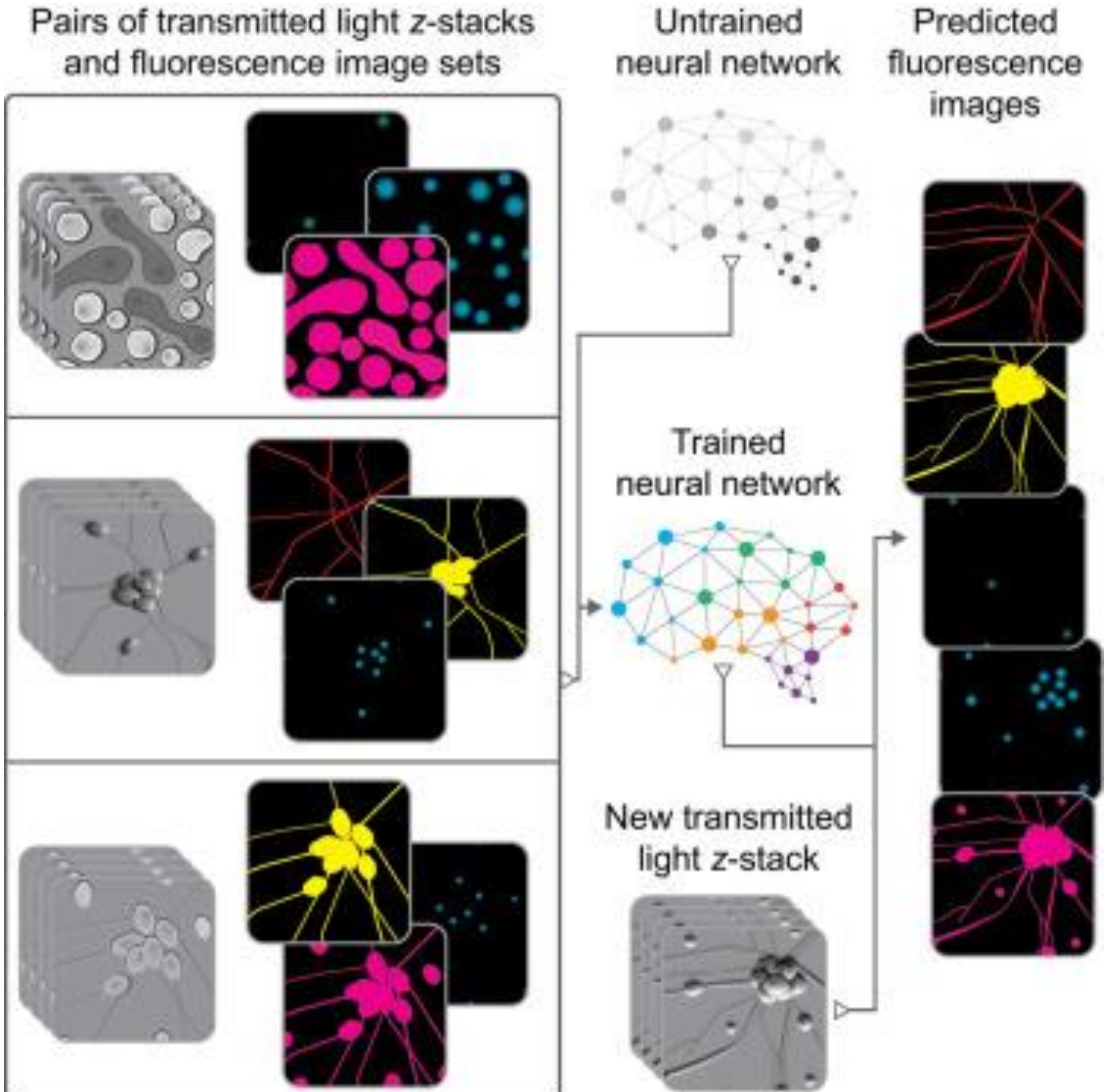
<u>Several relevant ideas for related course projects</u>

With the Allen Institute of Cell Science

In silico labeling: predicting fluorescent labels in unlabeled images

- Fluorescence microscopy images can be predicted from transmitted-light z stacks
  - Cell nuclei, live/dead, cell type, organelle type
- 7 fluorescent labels were validated across three labs, modalities, and cell types
- Multi-task learning
- Transfer learning: new labels can be predicted using minimal additional training data

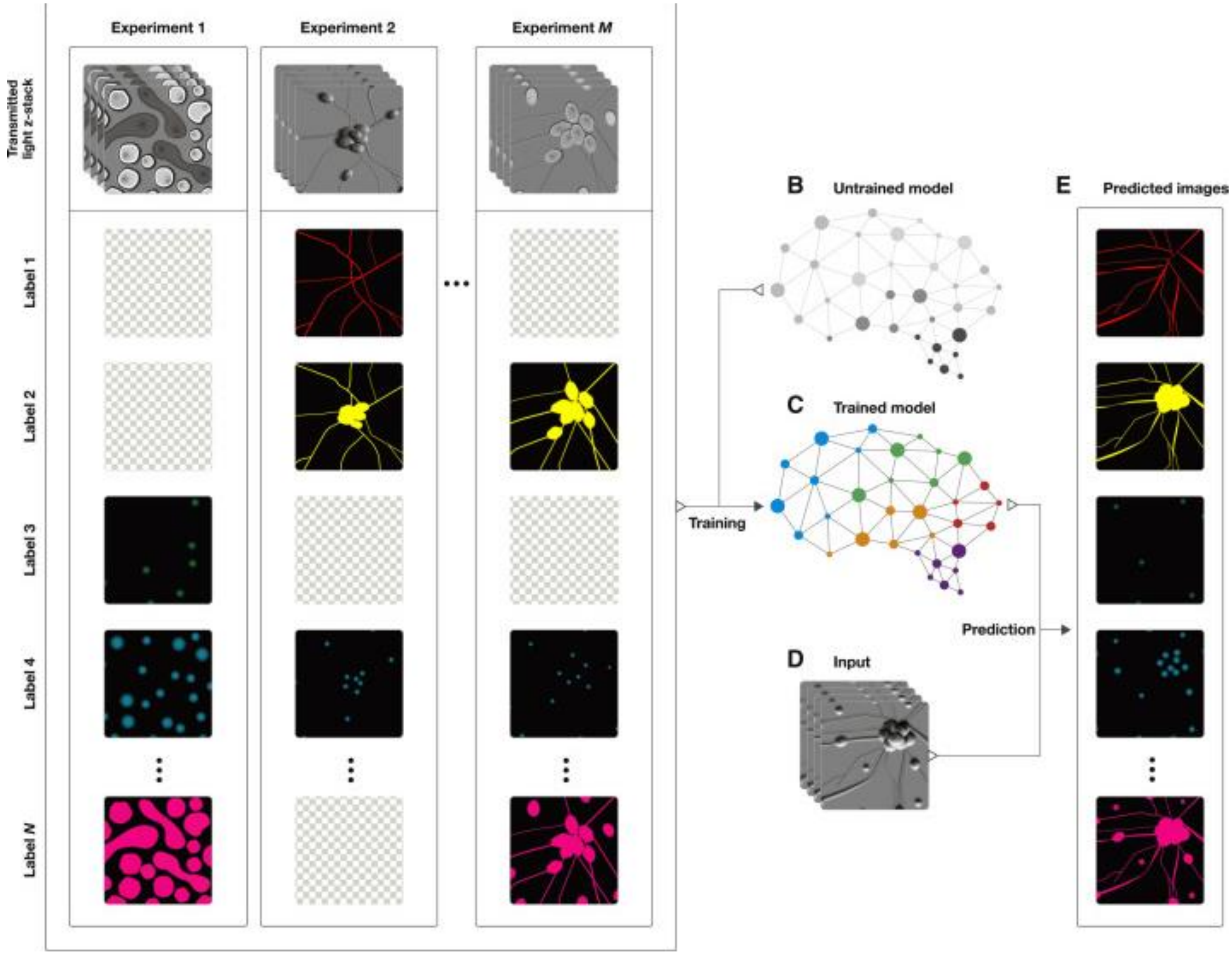
Christiansen et al. (2018)



Christiansen et al. (2018)

Multi-task learning

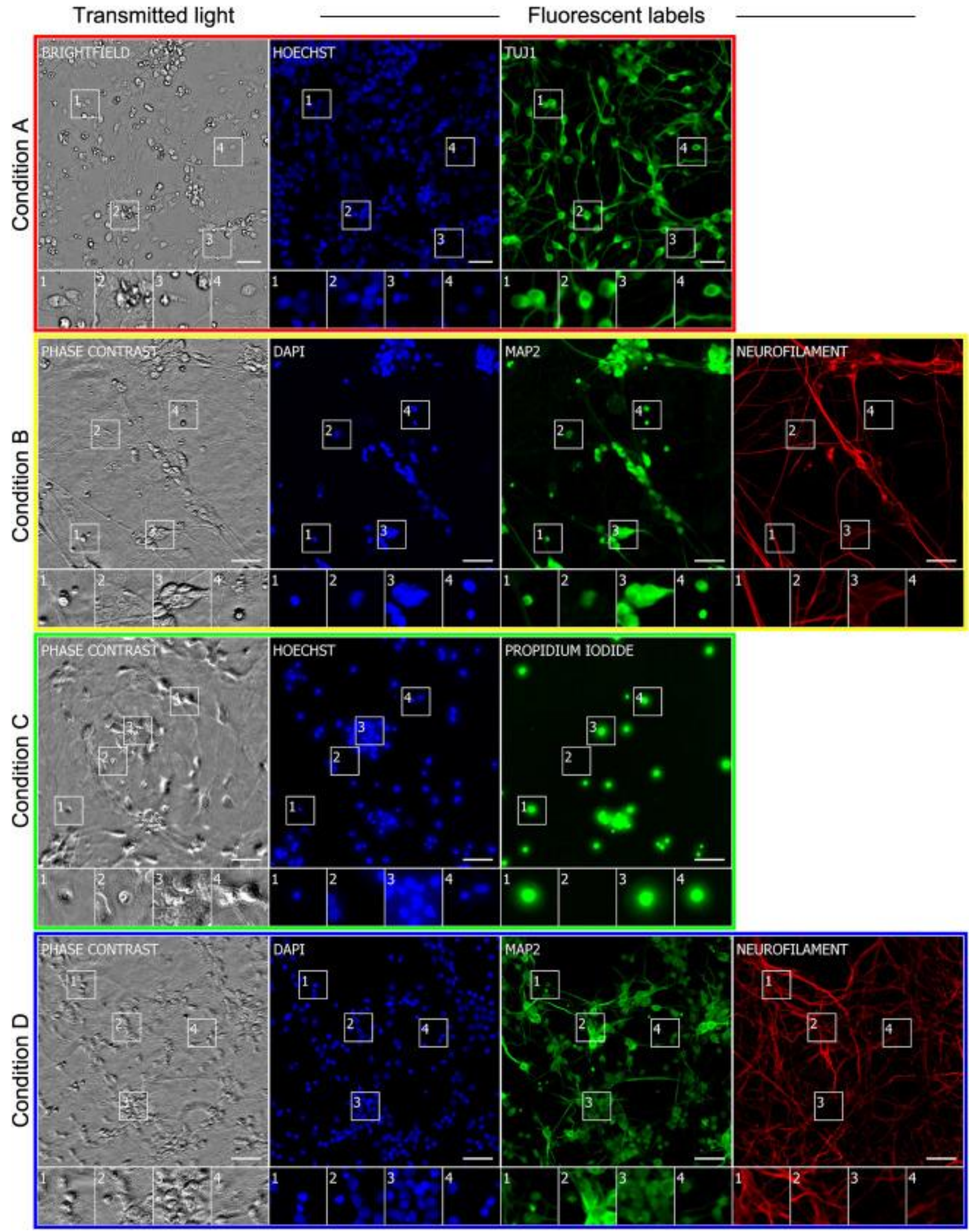
Learned abstractions can be reused across multiple tasks



Different cell types, labels, labs!

Christiansen et al. (2018)

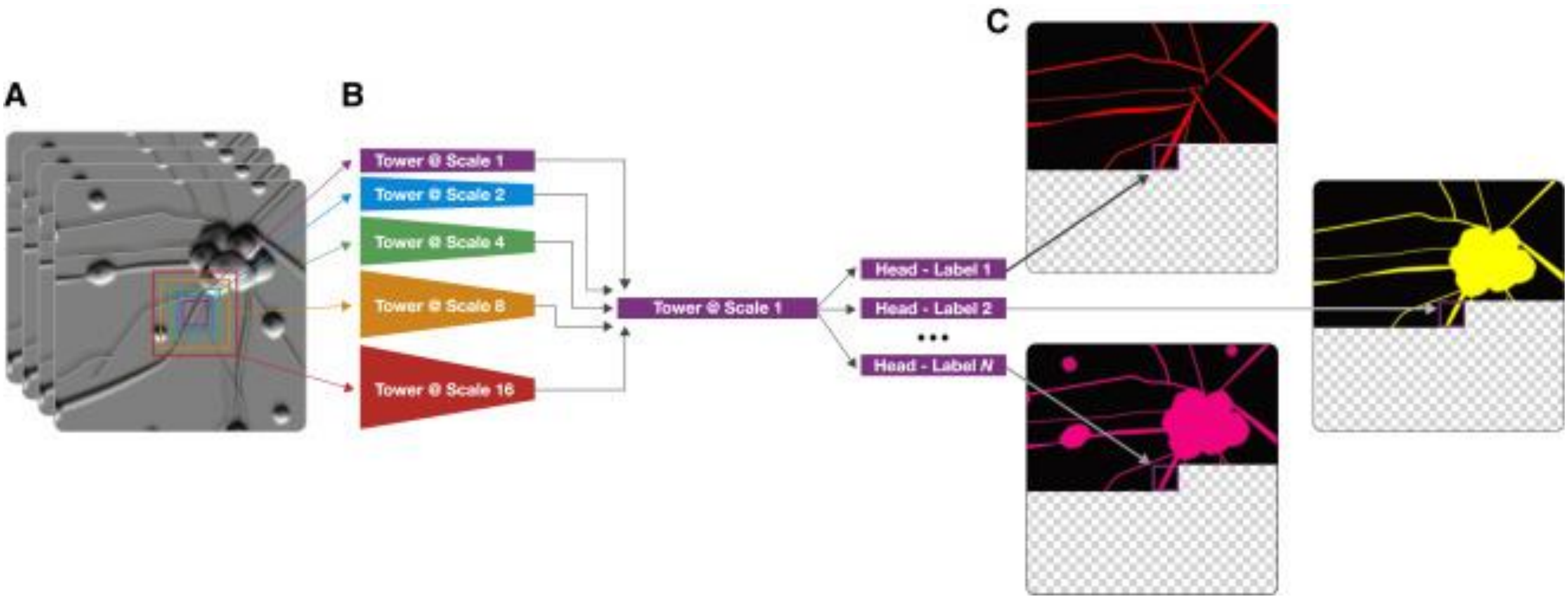
Training data



Christiansen et al. (2018)

Multi-scale learning

learning the spatial interpolation

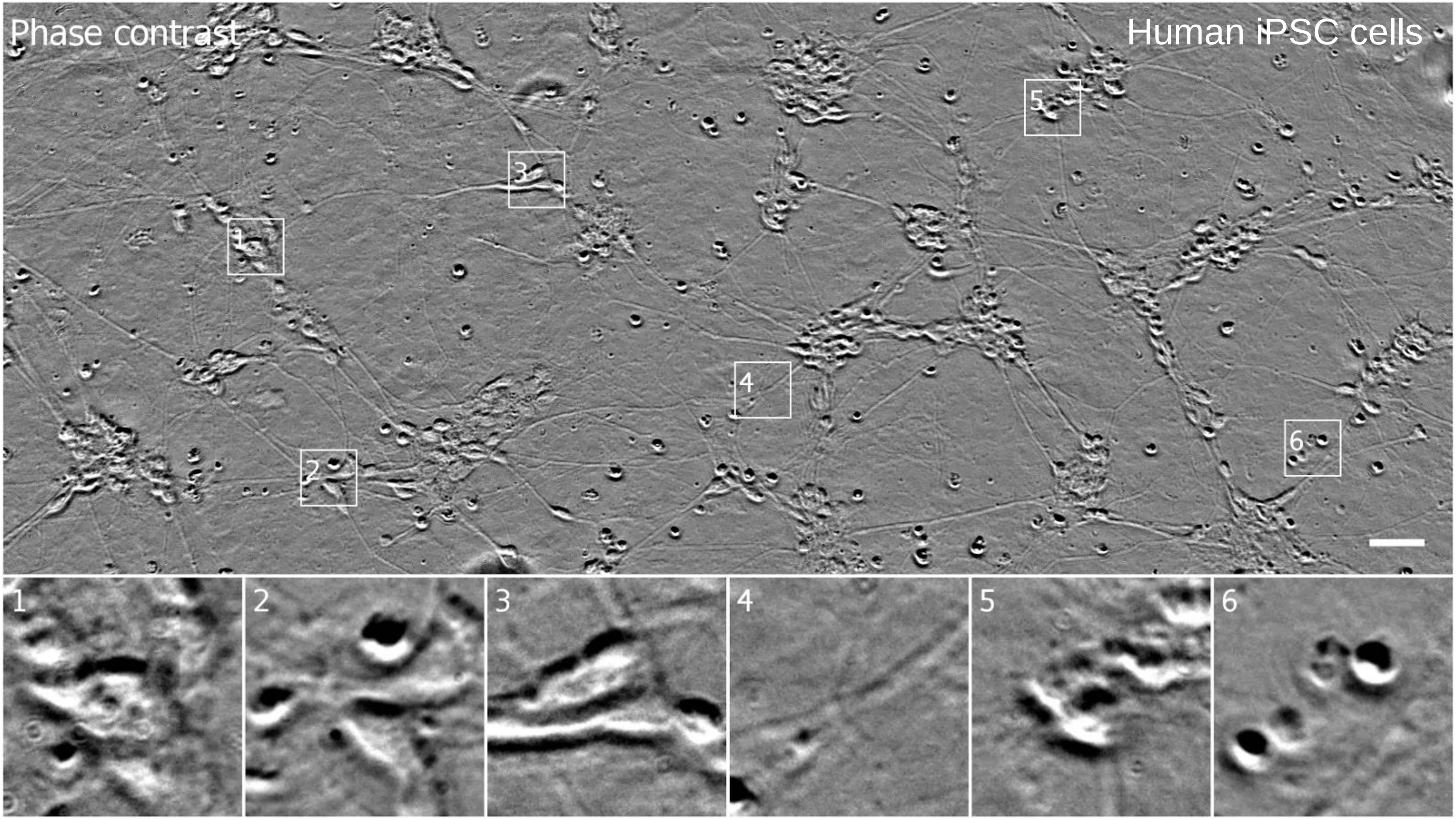


Inception modules (Szegedy et al., 2015) optimized with Google Hypertune (Golovin et al., 2017) to design the network architecture

Blogposts with simple explanations of inception networks [here](#) and [here](#)

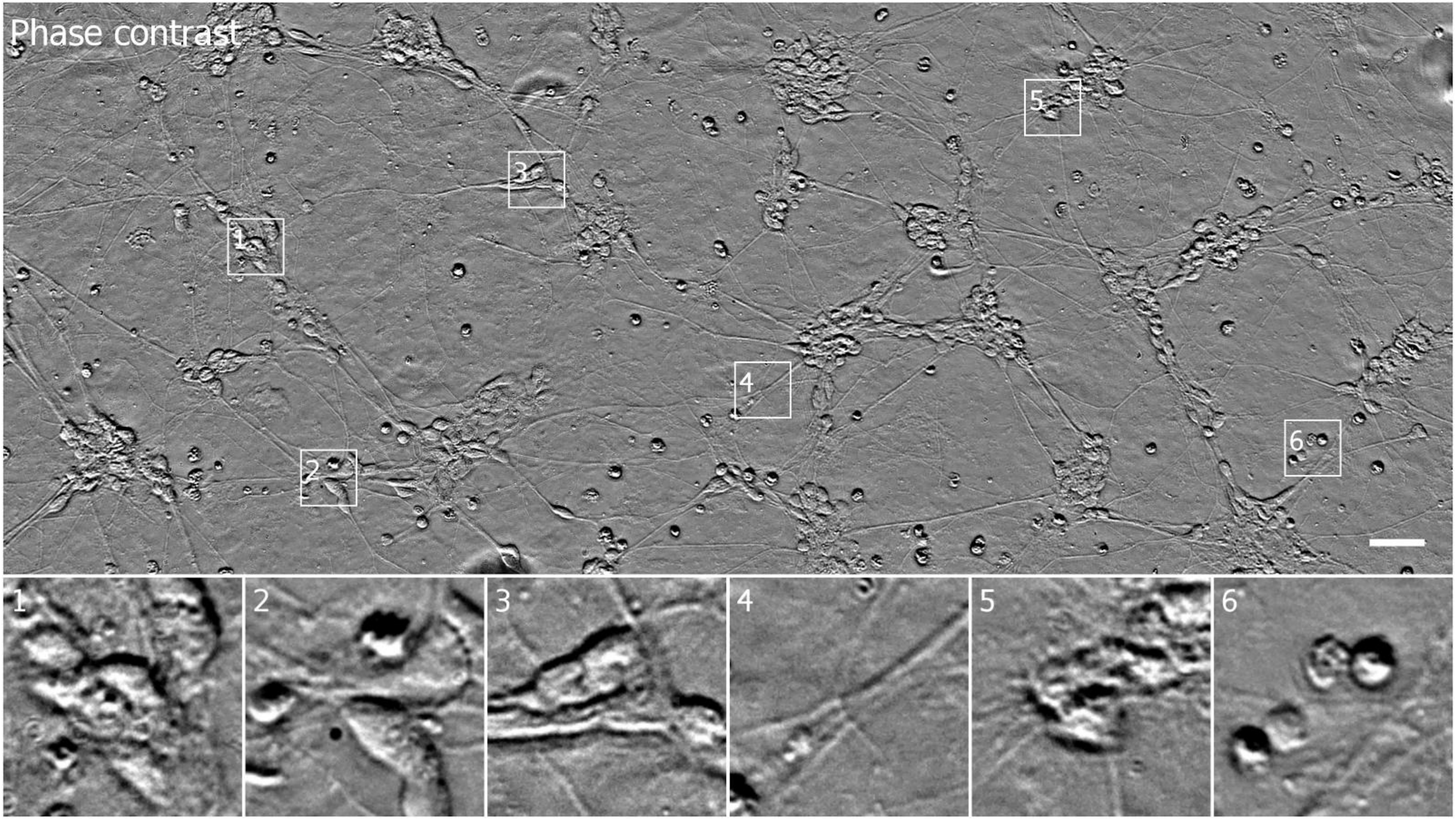
Christiansen et al. (2018)

Input: phase contrast imaging



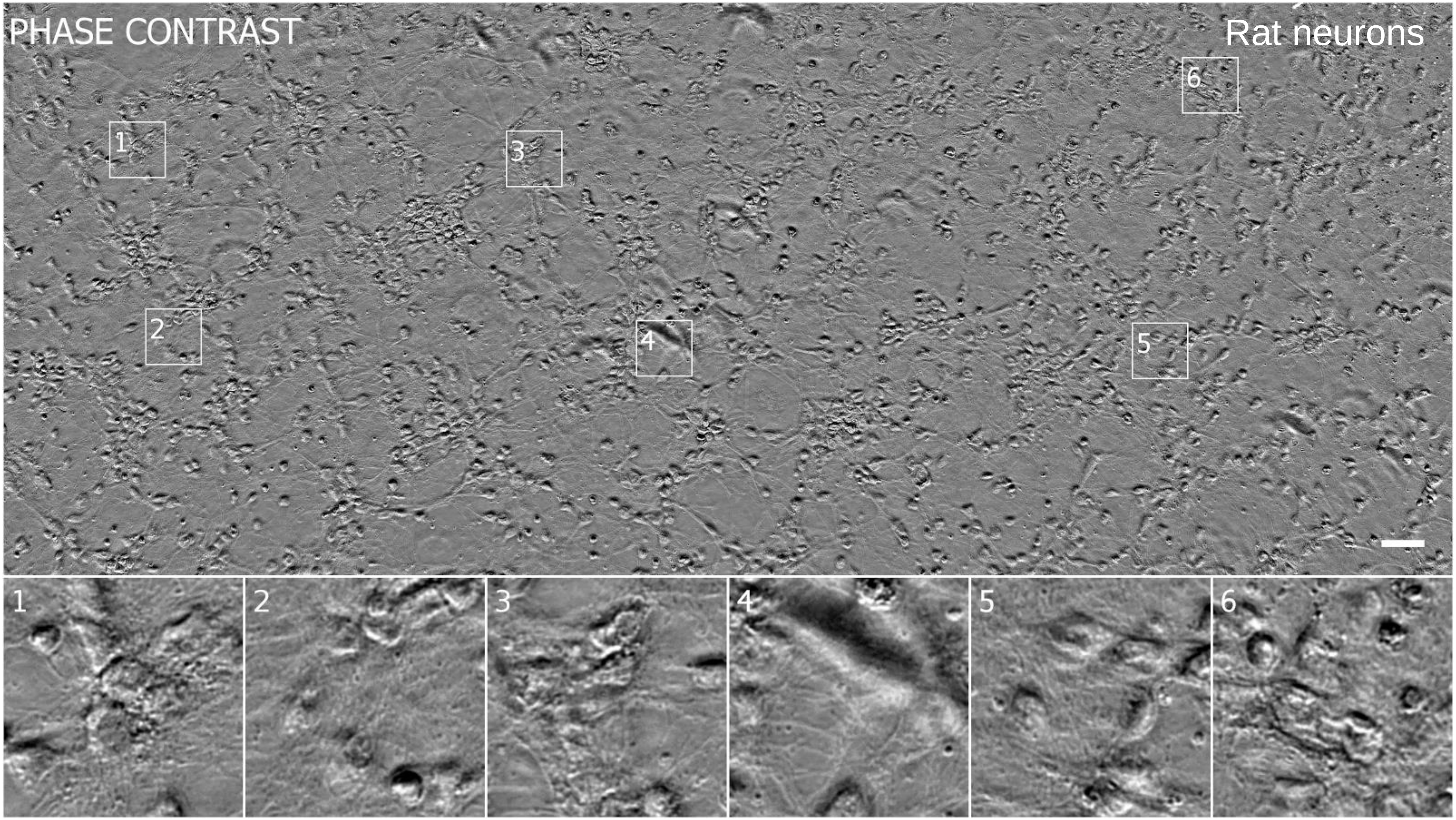
Eric Christiansen

Predictions



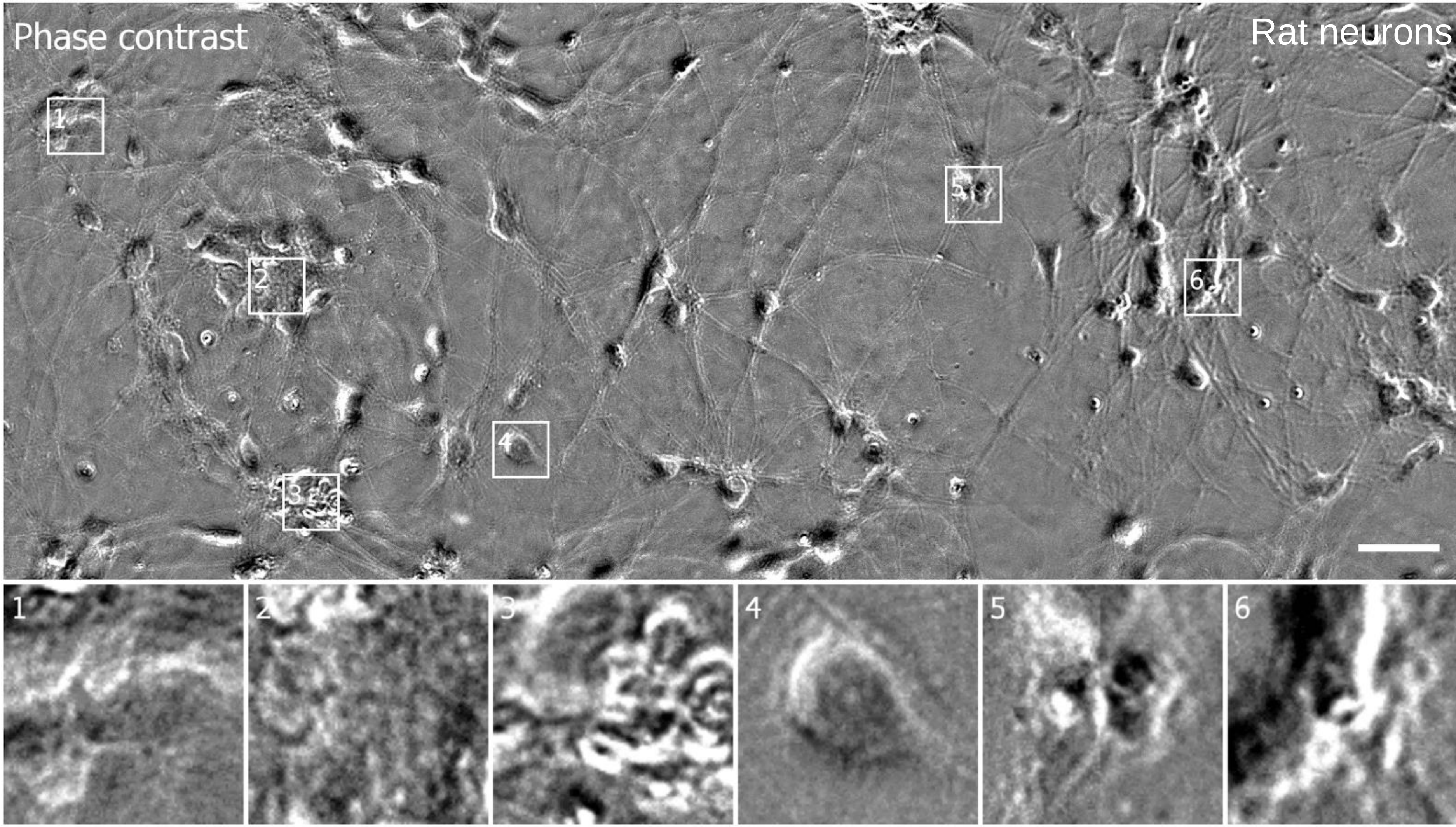
Eric Christiansen

Predictions



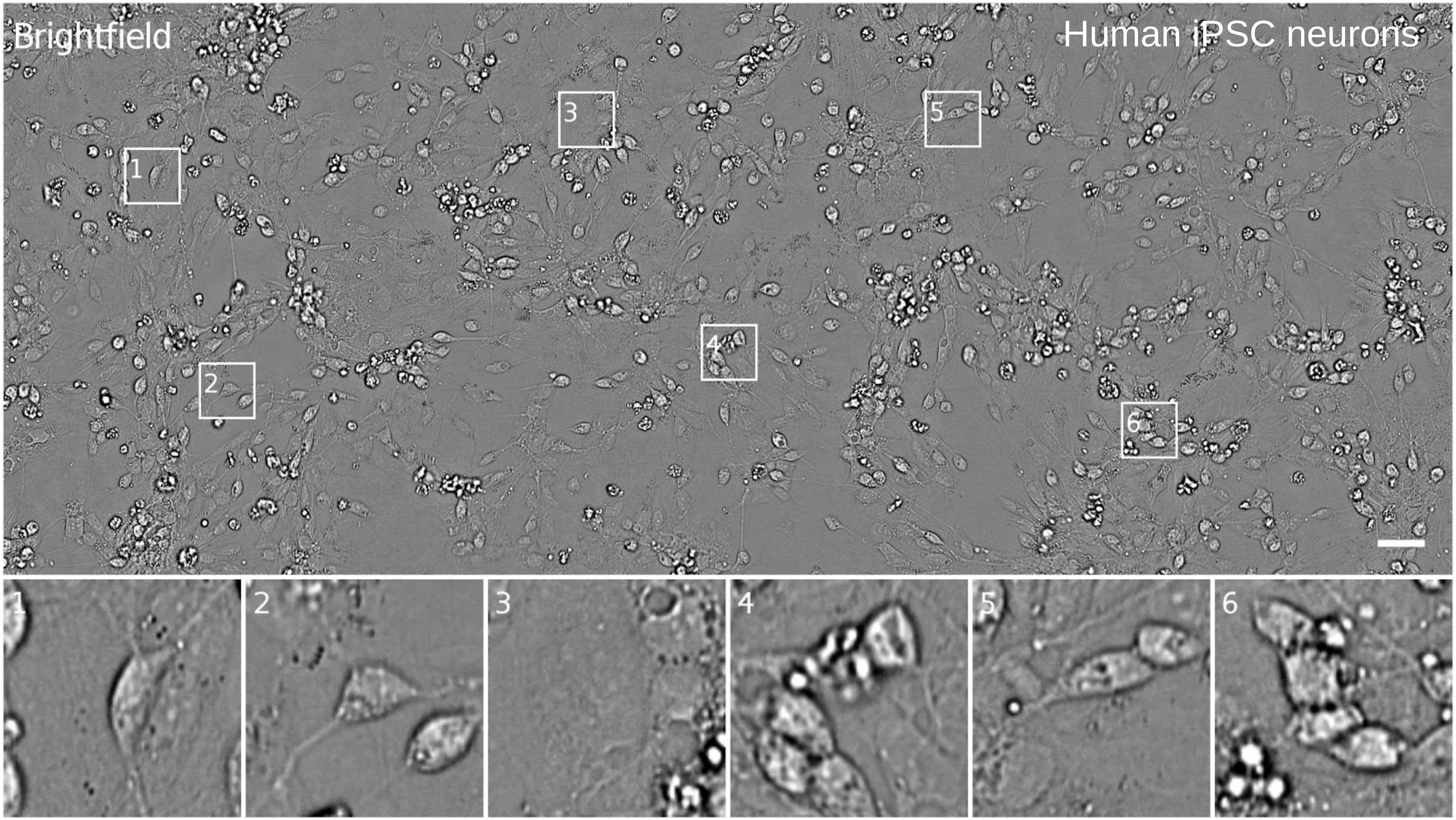
Eric Christiansen

Predictions



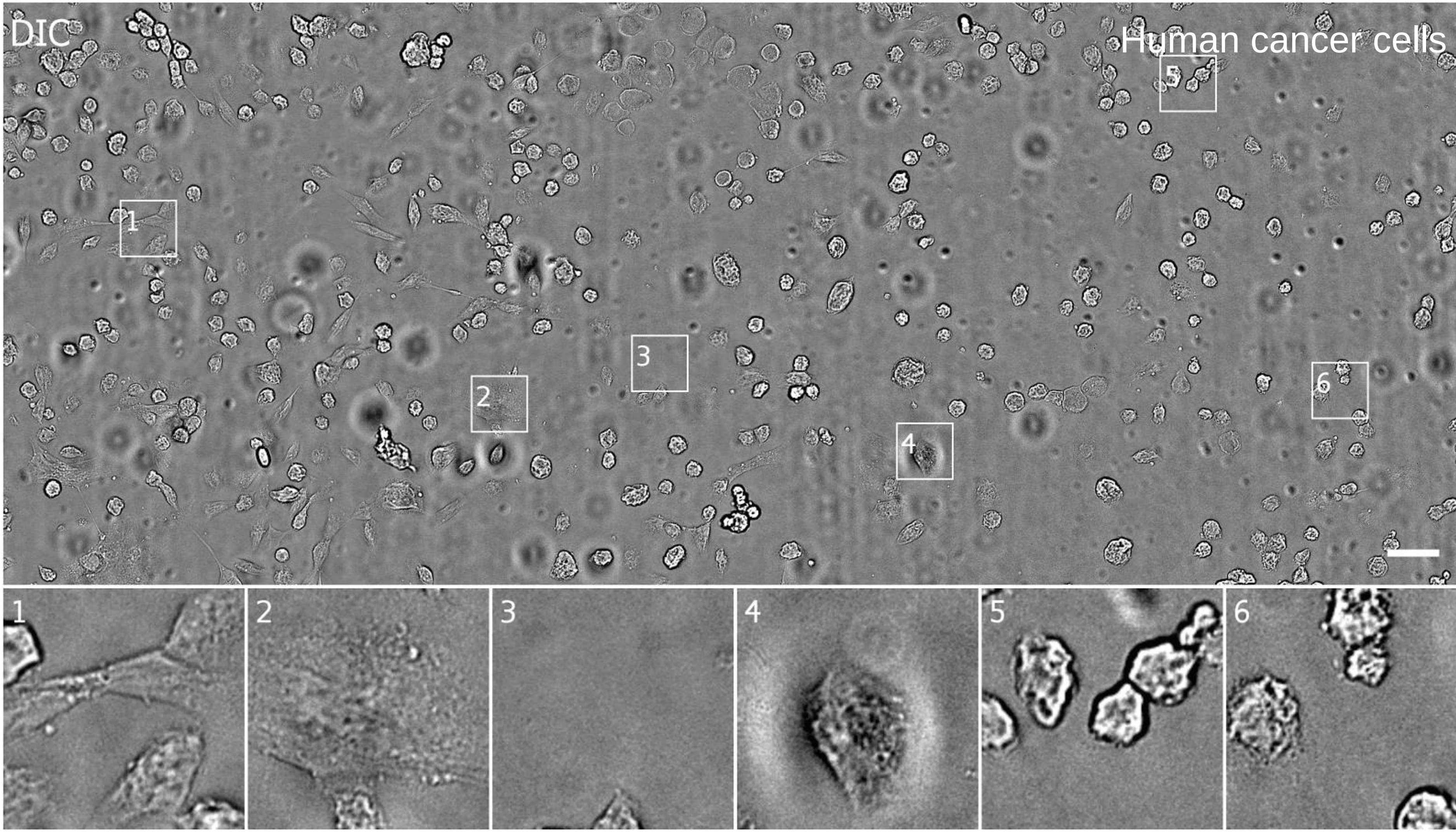
Eric Christiansen

Predictions



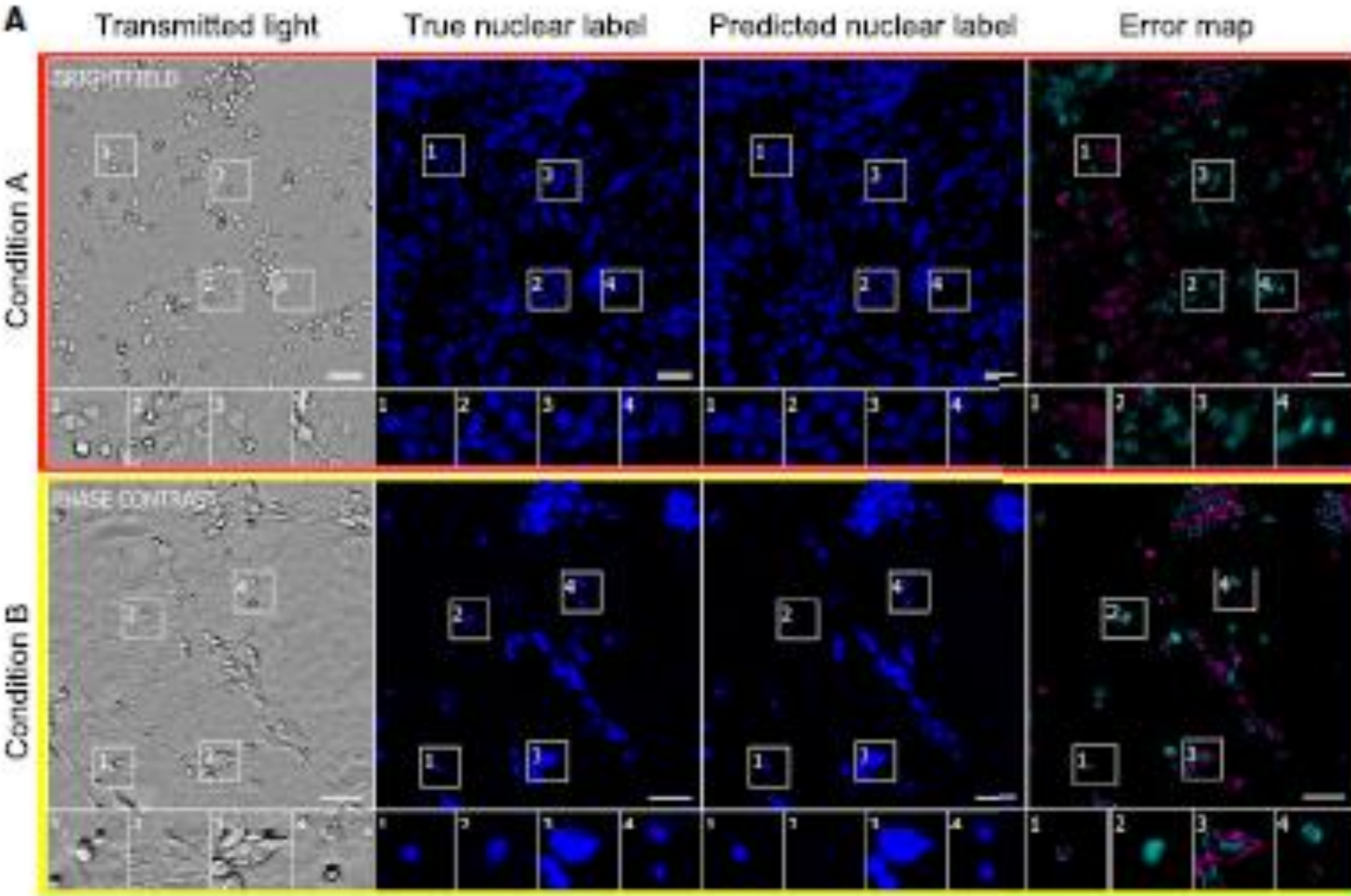
Eric Christiansen

Predictions



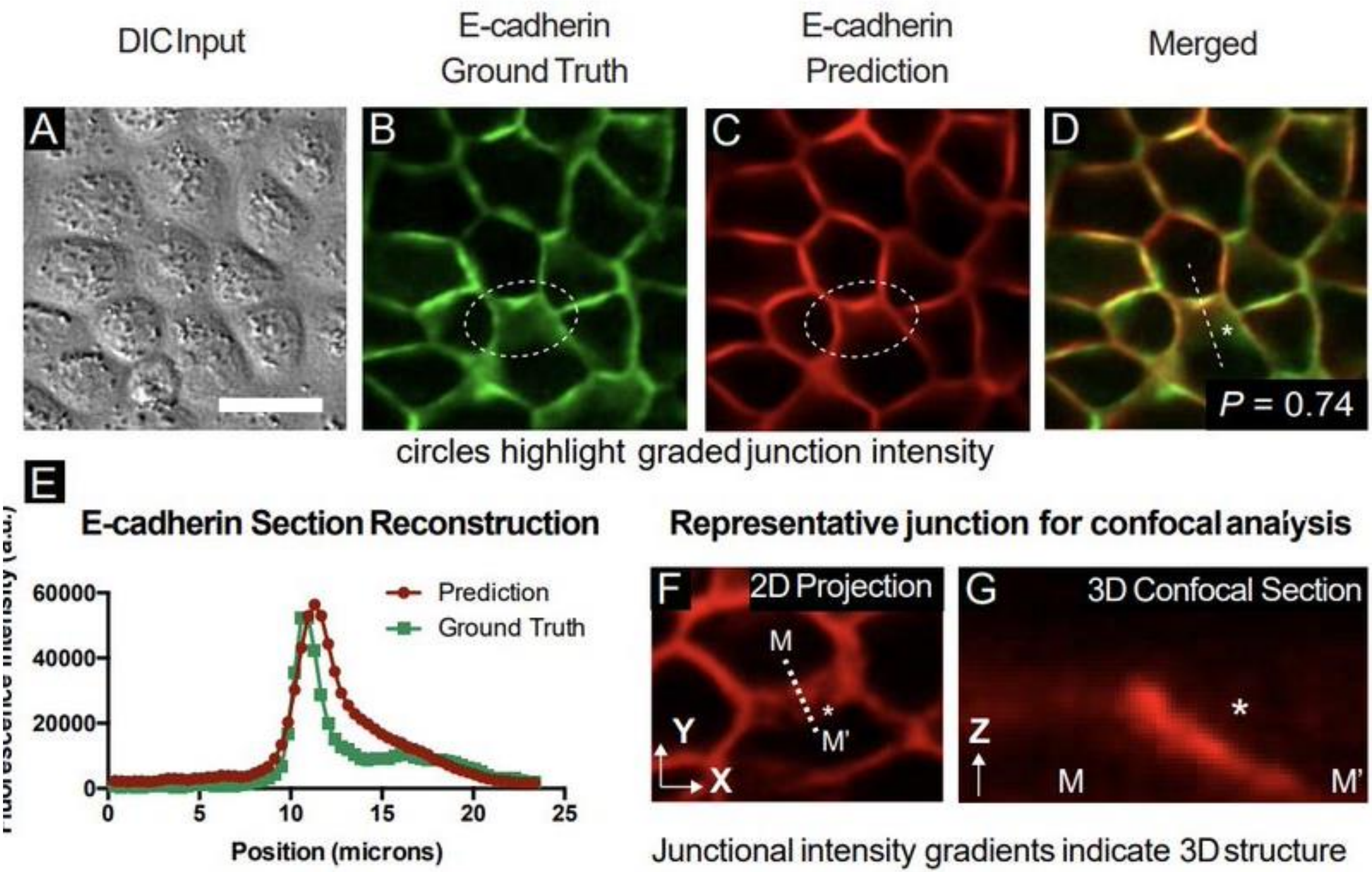
Eric Christiansen

Predicting nuclei (DAPI/Hoechst)



Christiansen et al. (2018)

Validations of downstream analysis



LaChance and Cohen (2020)

Examples of cross modality image mapping with deep learning

- PhaseStain: phase-to-histology (Rivenson, Liu, Wei, et al., 2019)
- Label free to physical cell properties (Guo, Yeh, Folkesson, at al., 2019)

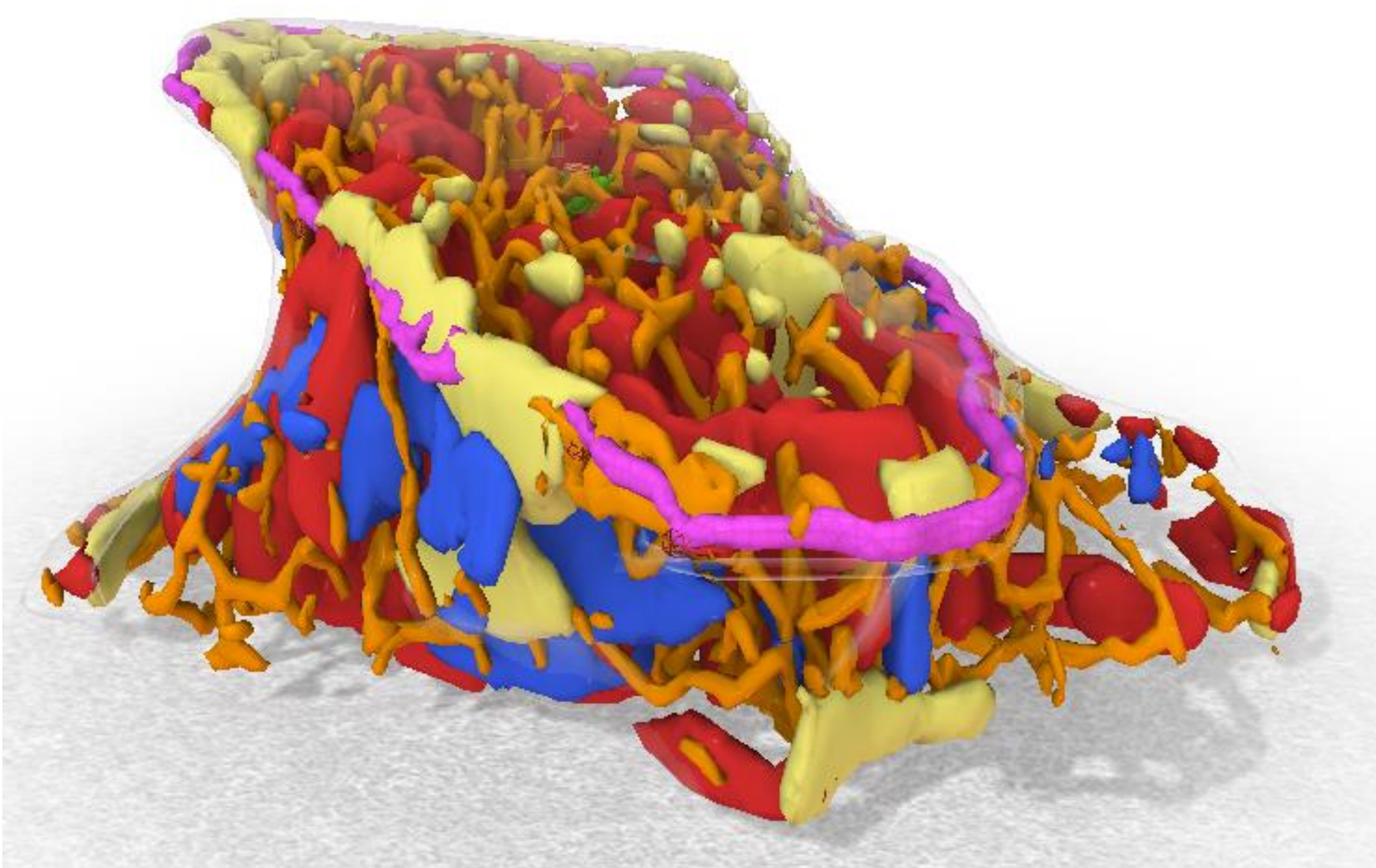
Could be picked as a student presentation

Other (non DL based) generative models for cell organization

- Robert Murphy’s lab, CMU, http://www.andrew.cmu.edu/user/murphy/
- Could be picked as a student/s presentation

Look at a cell and know what it is doing

What it did



Structure ↔ Function

What it will do

Slide adapted from Susanne Rafelski, Allen Institute of Cell Science

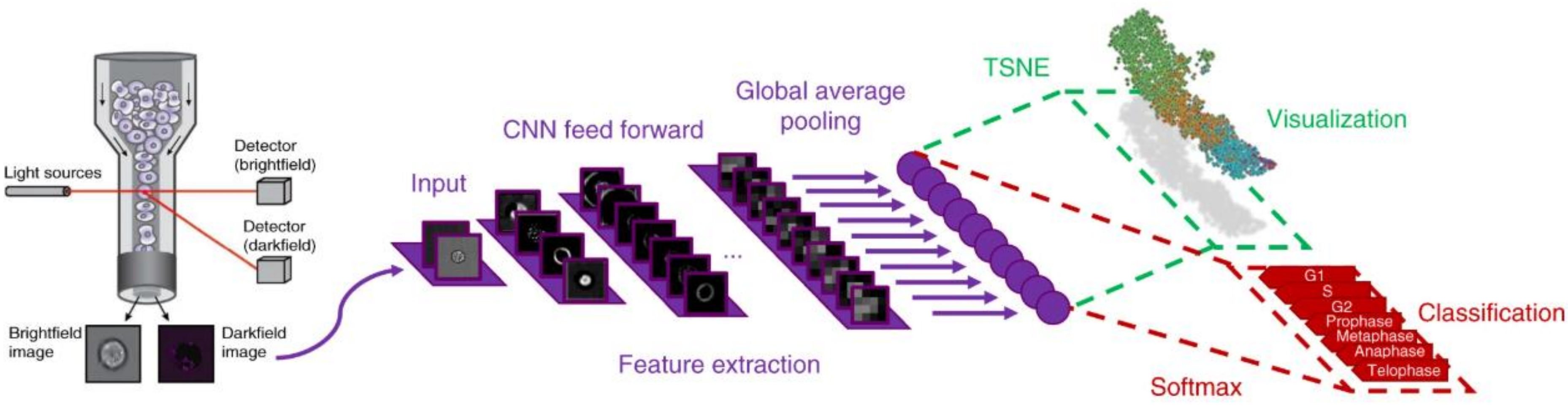
Classify cell state with deep learning

Predicting cell cycle / disease progression stage (“pseudo time”) with deep learning

- Combination of CNN and nonlinear dimension reduction enable reconstructing cell cycle from raw image data
- Unsupervised detection of a subpopulation of dead cells
- Supervised CNN classification surpassing boosting on image features
- Generalization of approach on diabetic retinopathy progression

Eulenberg (2017)

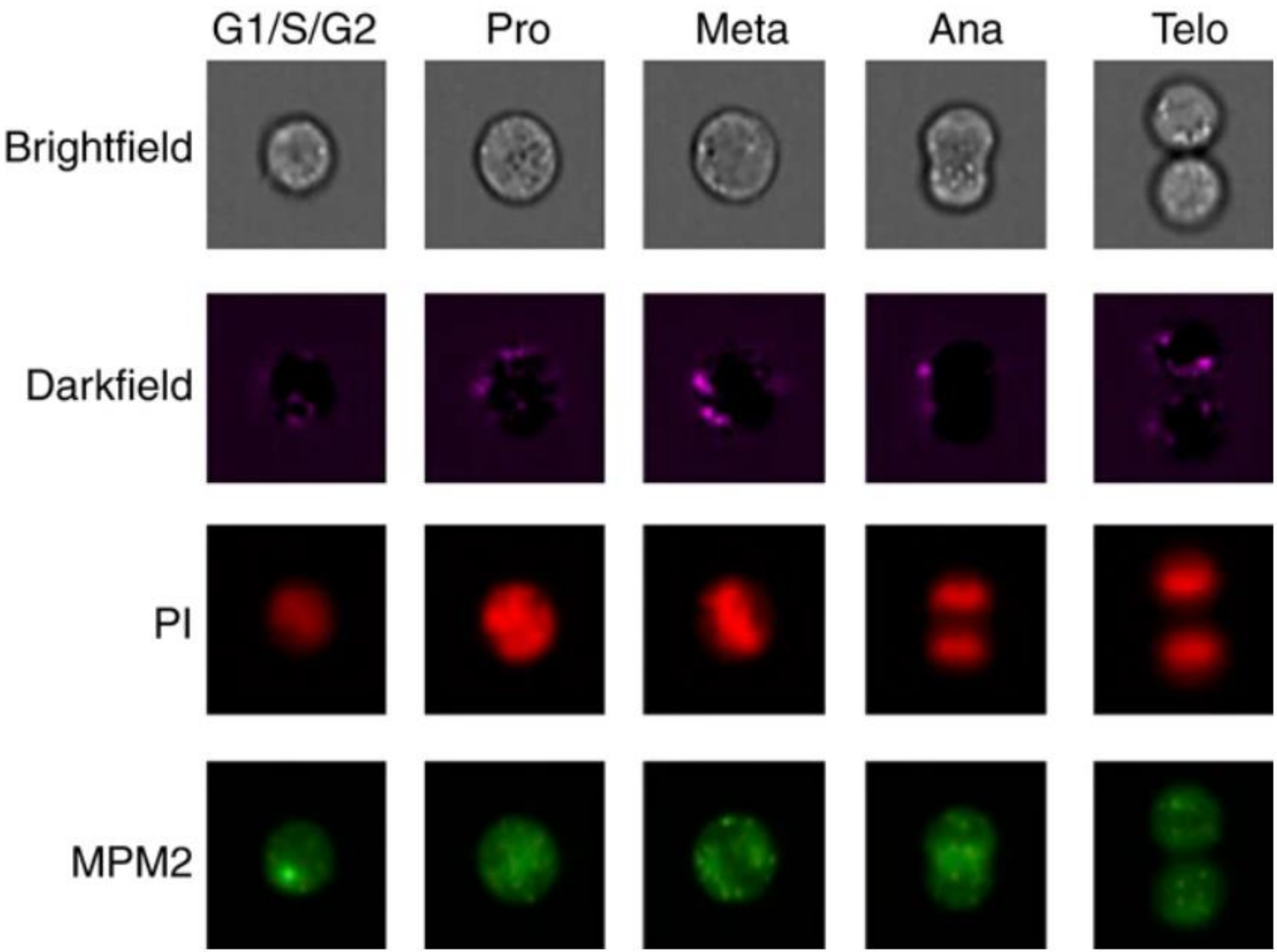
Experiment and analysis pipeline



- Imaging flow cytometry – one cell at a time (no need for segmentation), 37K human T-cells (Jurkat)
- Similarities between cells in terms of the features extracted by the network with tSNE (t-distributed stochastic neighbor embedding, Maaten and Hinton, 2008)

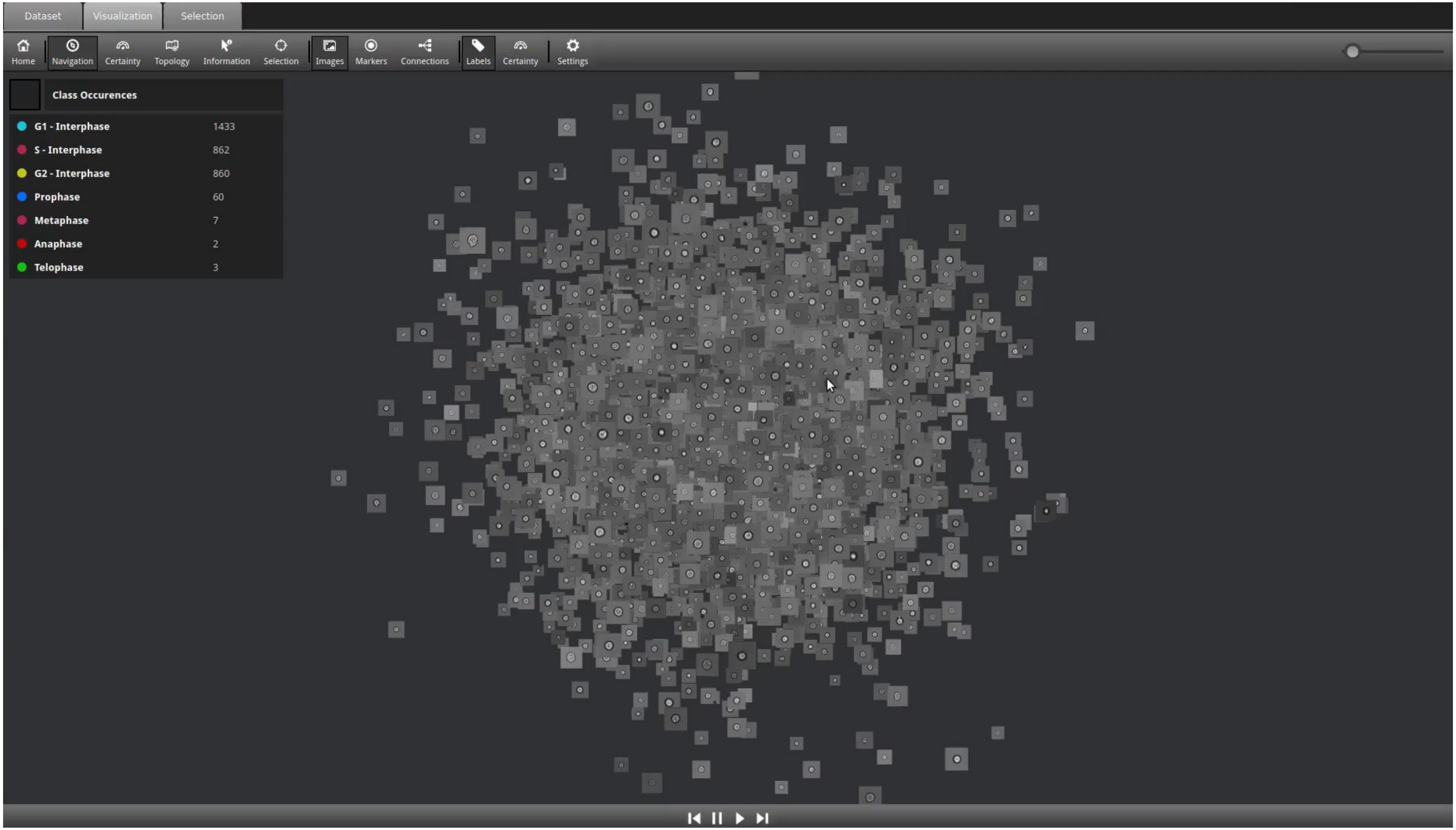
Eulenberg (2017)

Seven cell cycle stages



Eulenberg (2017)

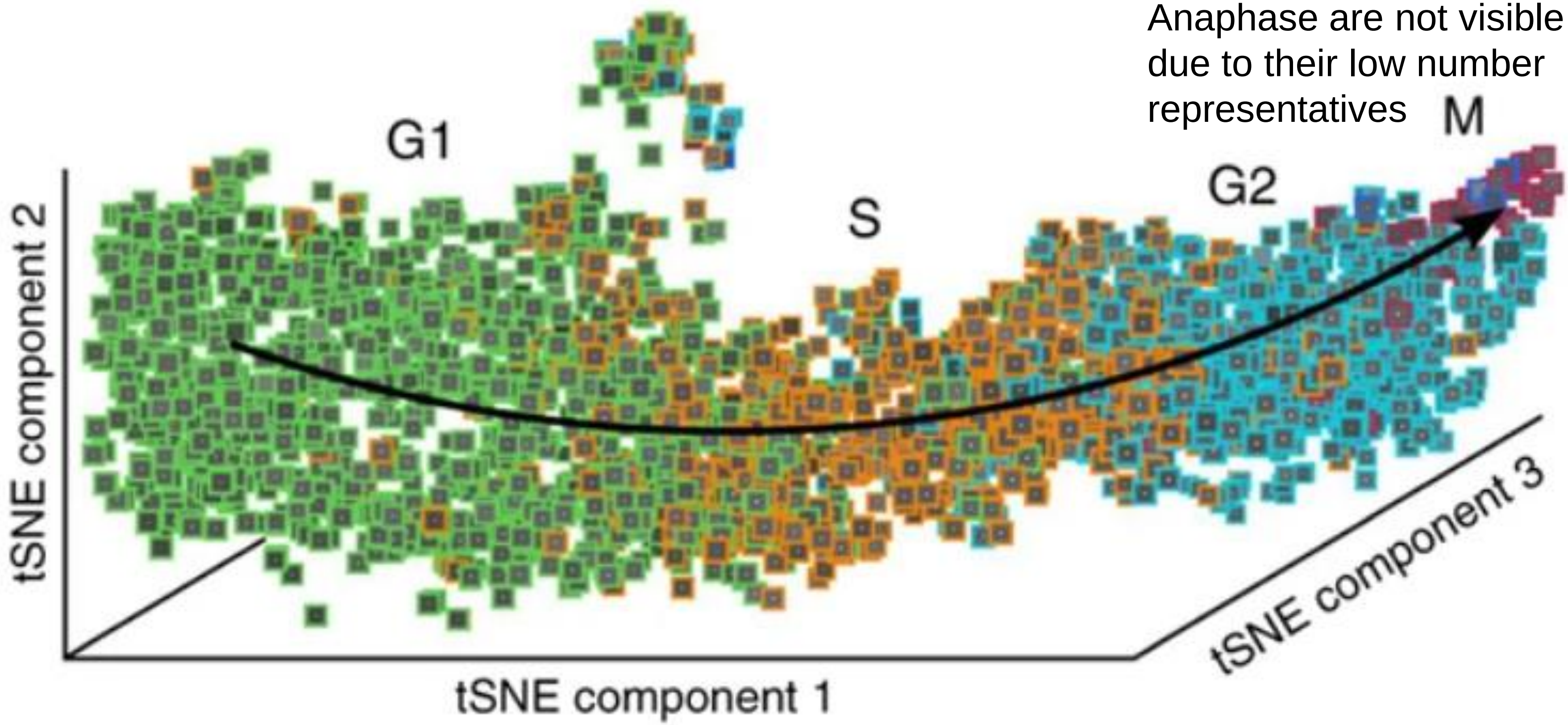
TSNE visualization



Eulenberg (2017)

Continuous temporal progression along cell cycle phases from the raw image data

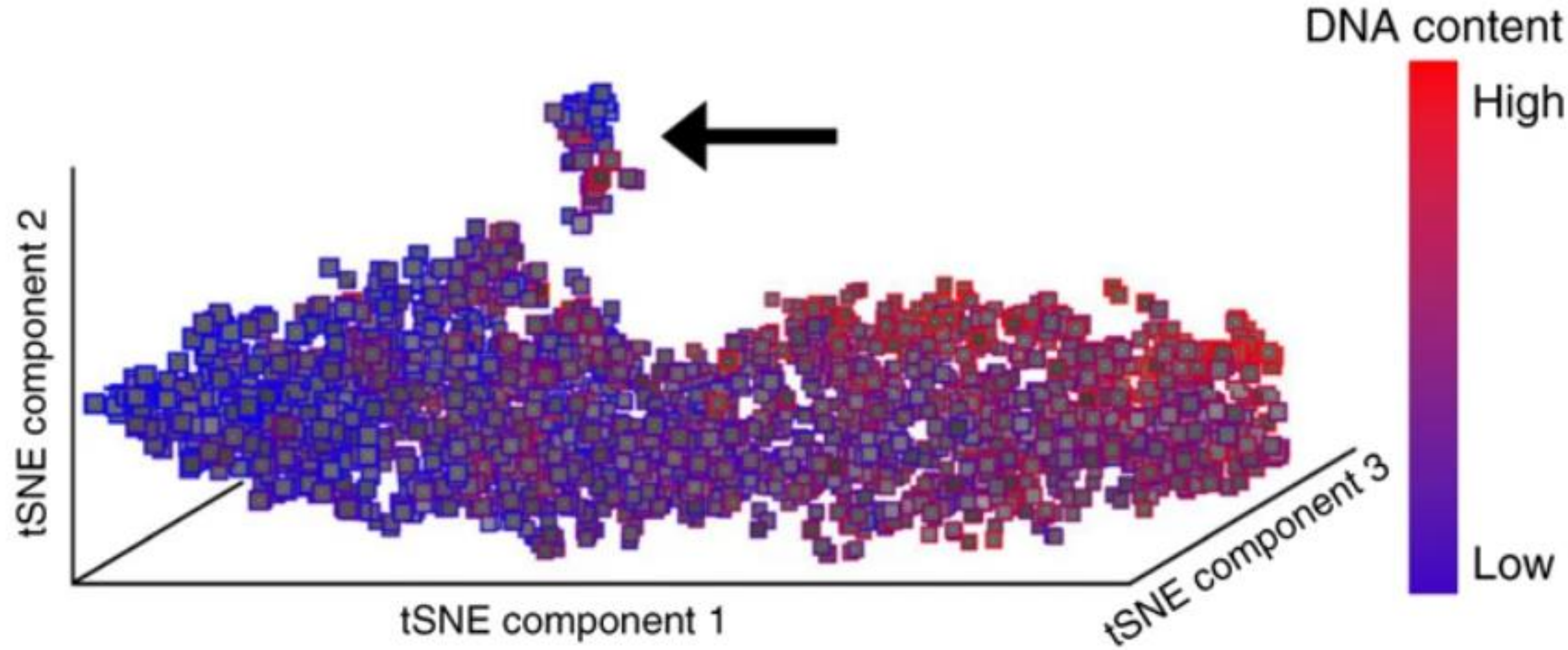
Prophase: red, Metaphase: blue. Telophase and Anaphase are not visible due to their low number representatives



Eulenberg (2017)

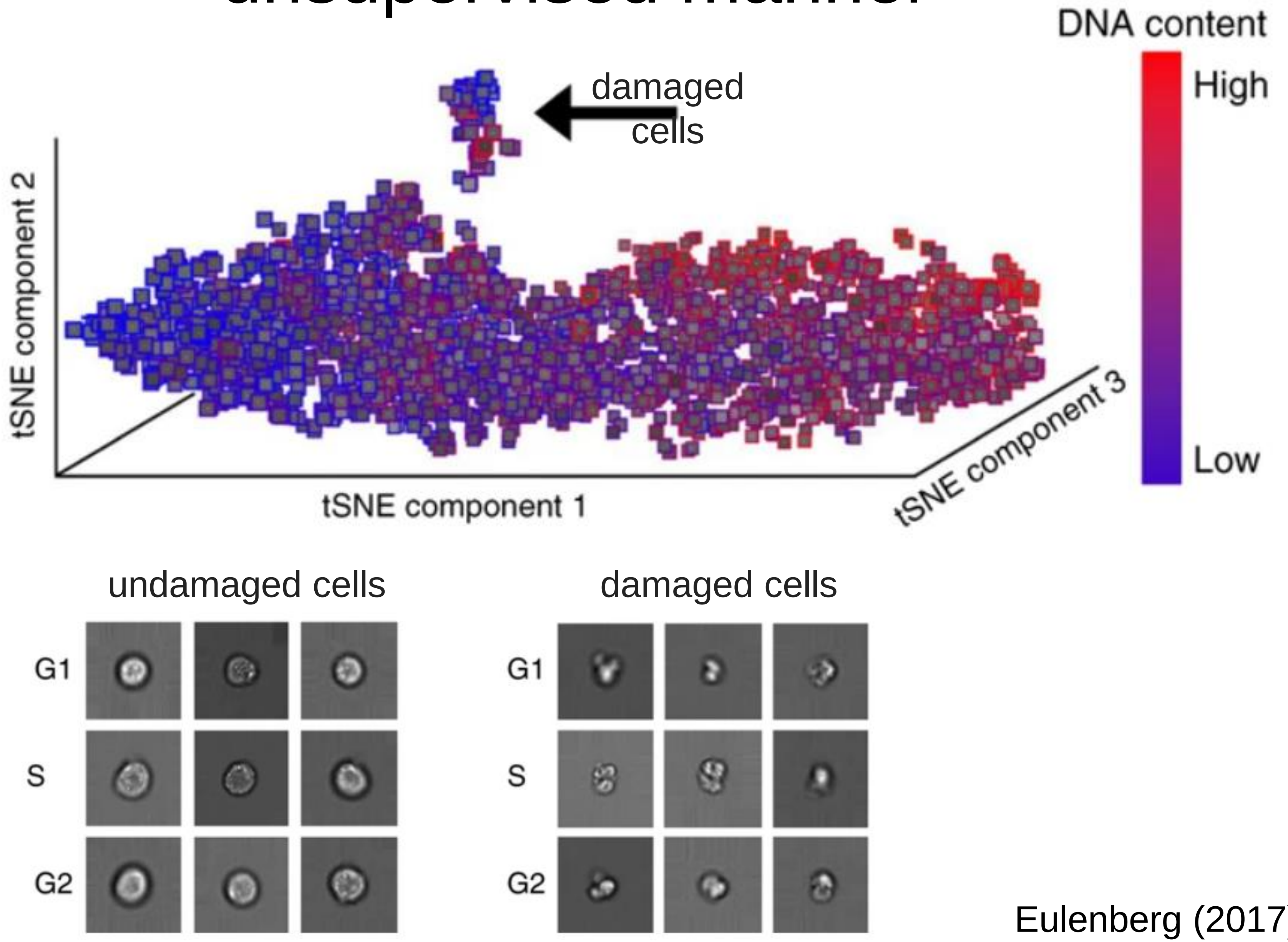
Visualization of interphase classes (G1, S, G2)

DNA content reflects the continuous progression



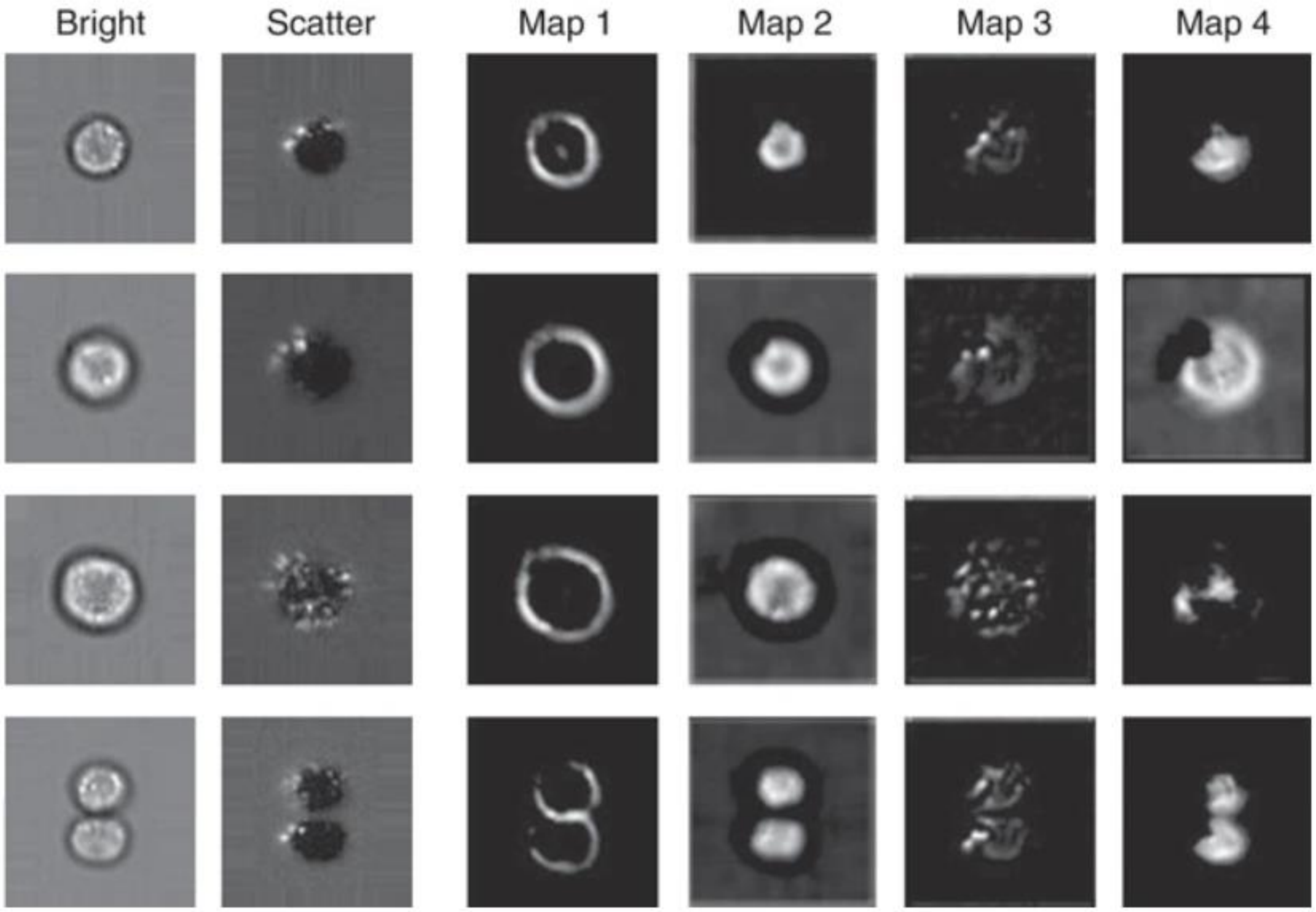
Eulenberg (2017)

Detecting abnormal cells in an unsupervised manner



Eulenberg (2017)

Activation patterns of intermediate layers

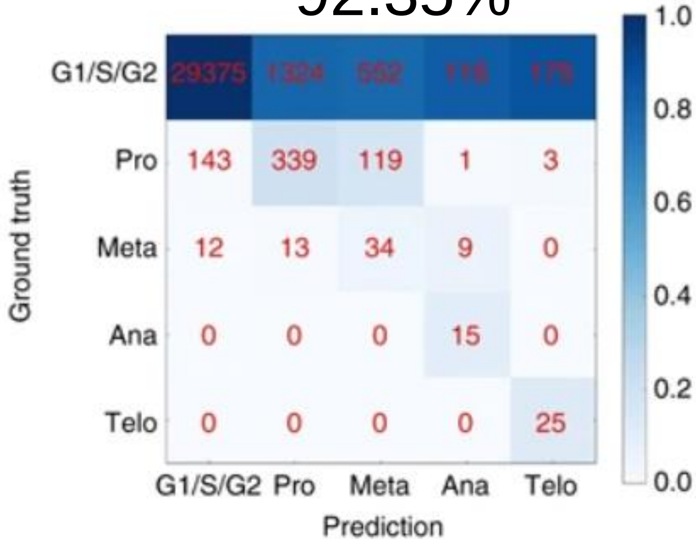


Eulenberg (2017)

Deep learning outperforms boosting for cell classification

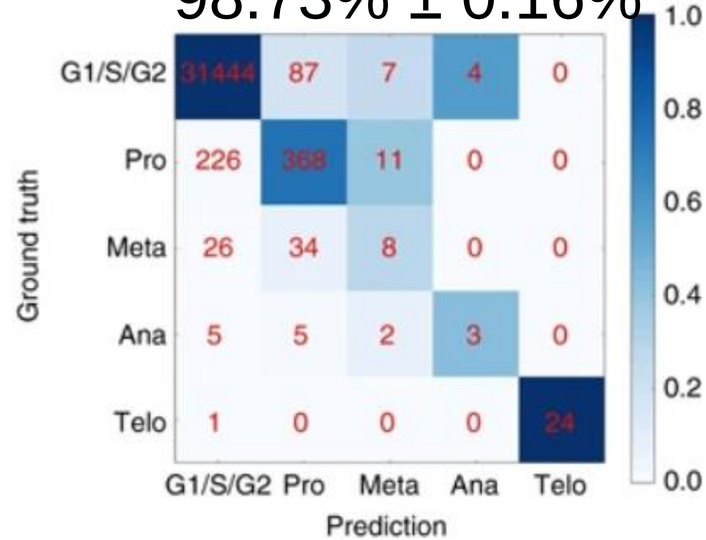
Boosting (Blasi et al. 2016)

92.35%

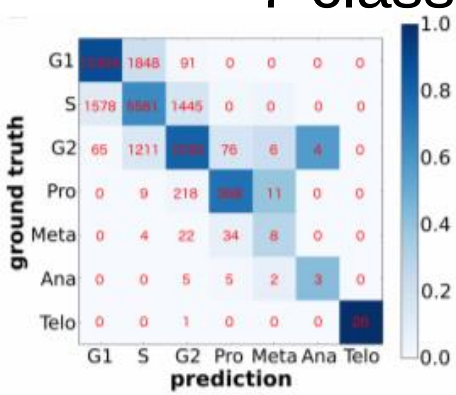


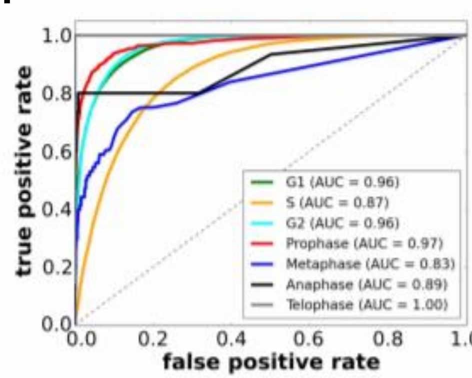
Deep learning

98.73% ± 0.16%



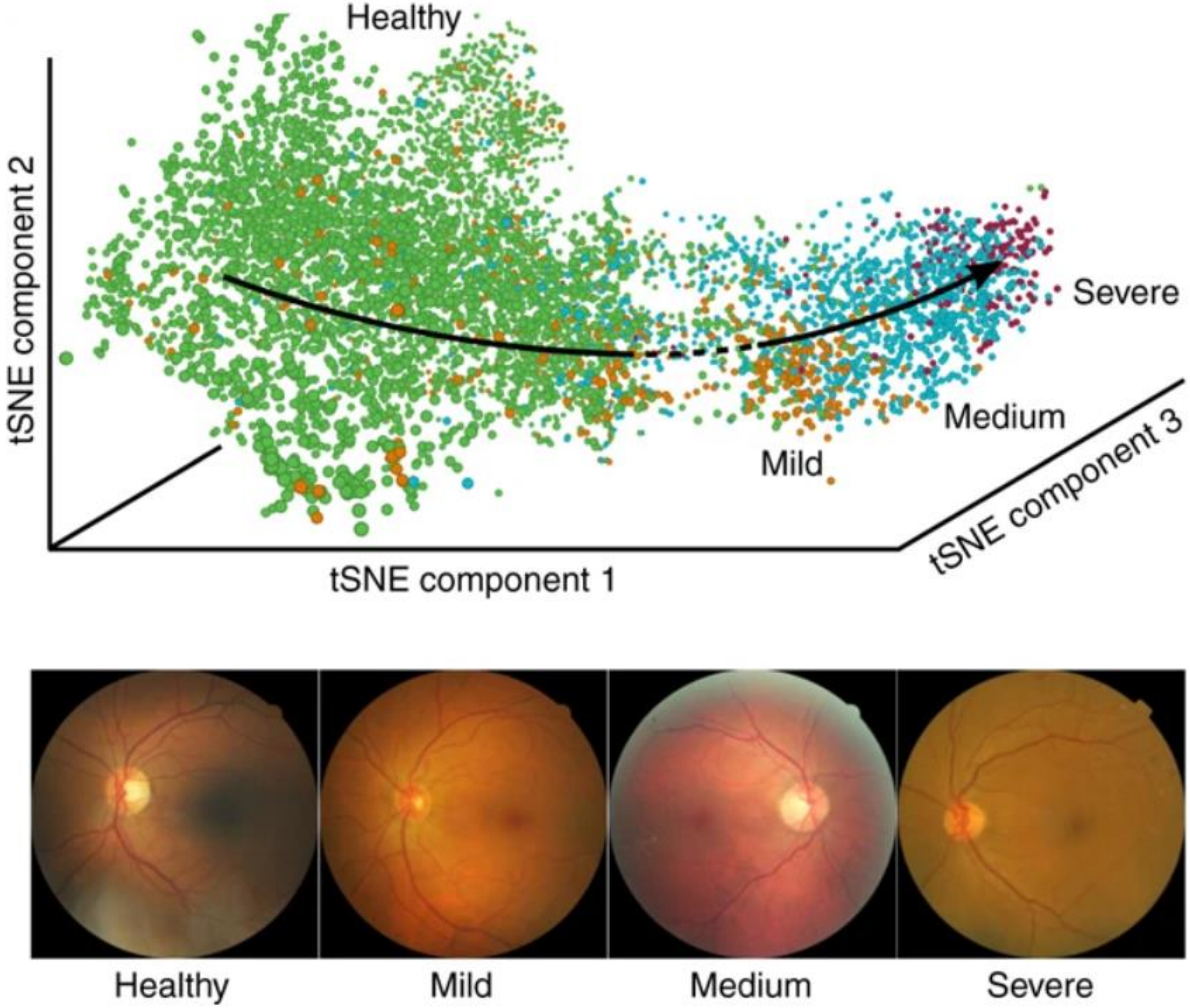
7 classes prediction





Eulenberg (2017)

Reconstructing disease progression: diabetic retinopathy

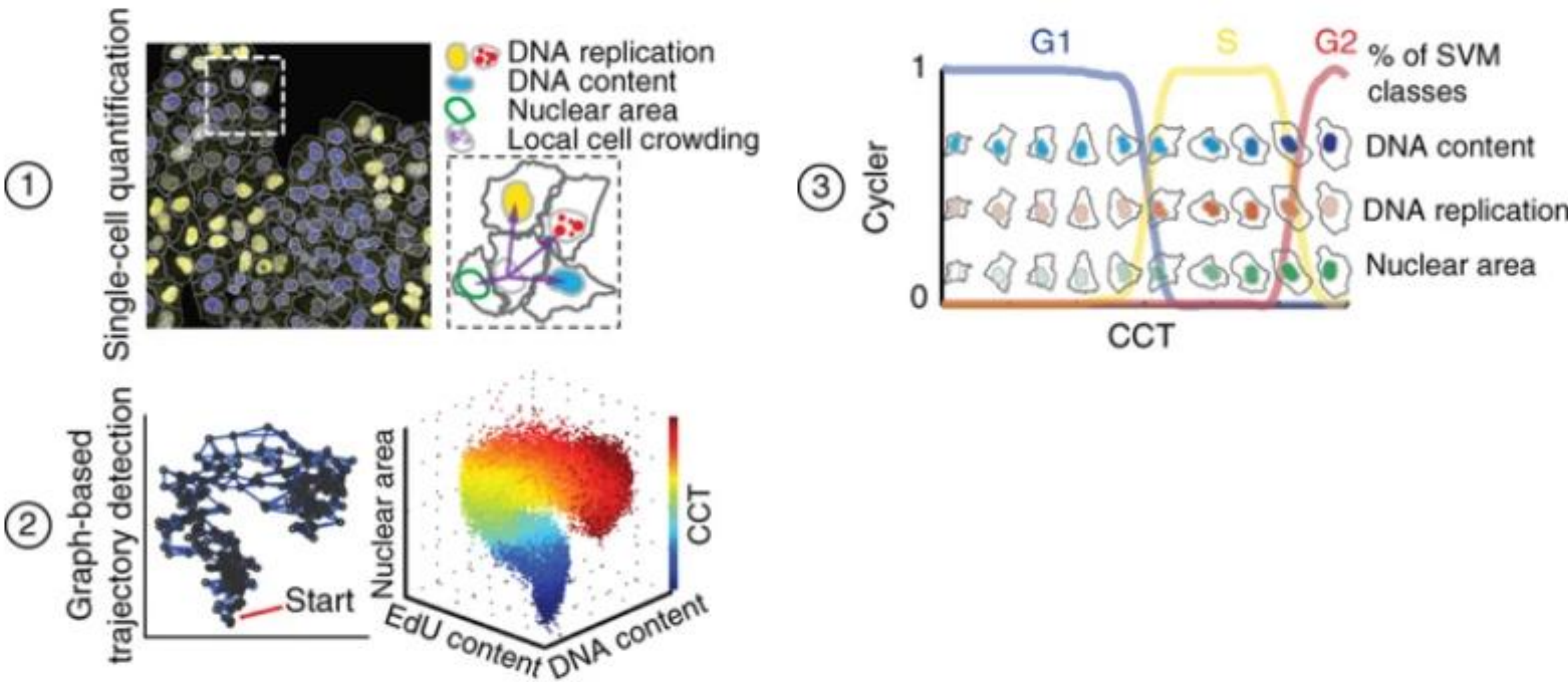


Eulenberg (2017)

Trajectories of cell-cycle progression from fixed cell populations

Including spatial context

(hopefully we'll go through it later in the course)



Gut, Tadmor, et al. (2014)

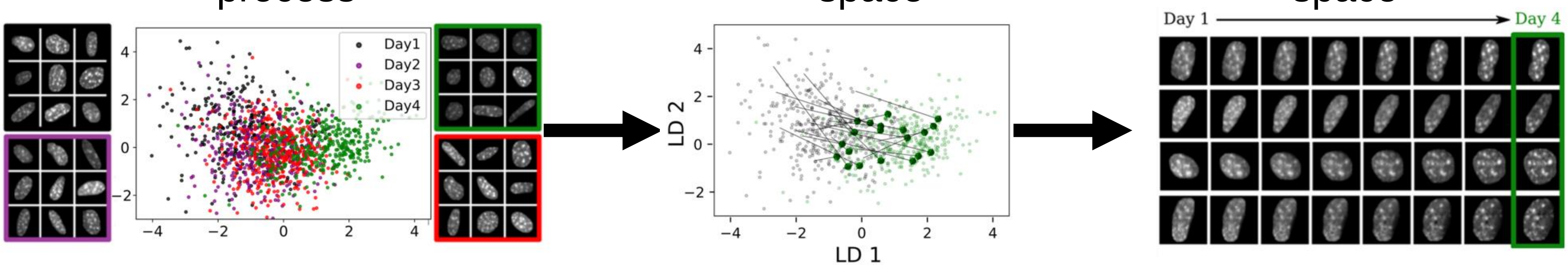
Predicting and visualizing pseudo time from snapshot images of single cells using autoencoders and optimal transport

Just came out a few days ago!

Fixed NIH3T3 nuclei imaged along a dynamic process

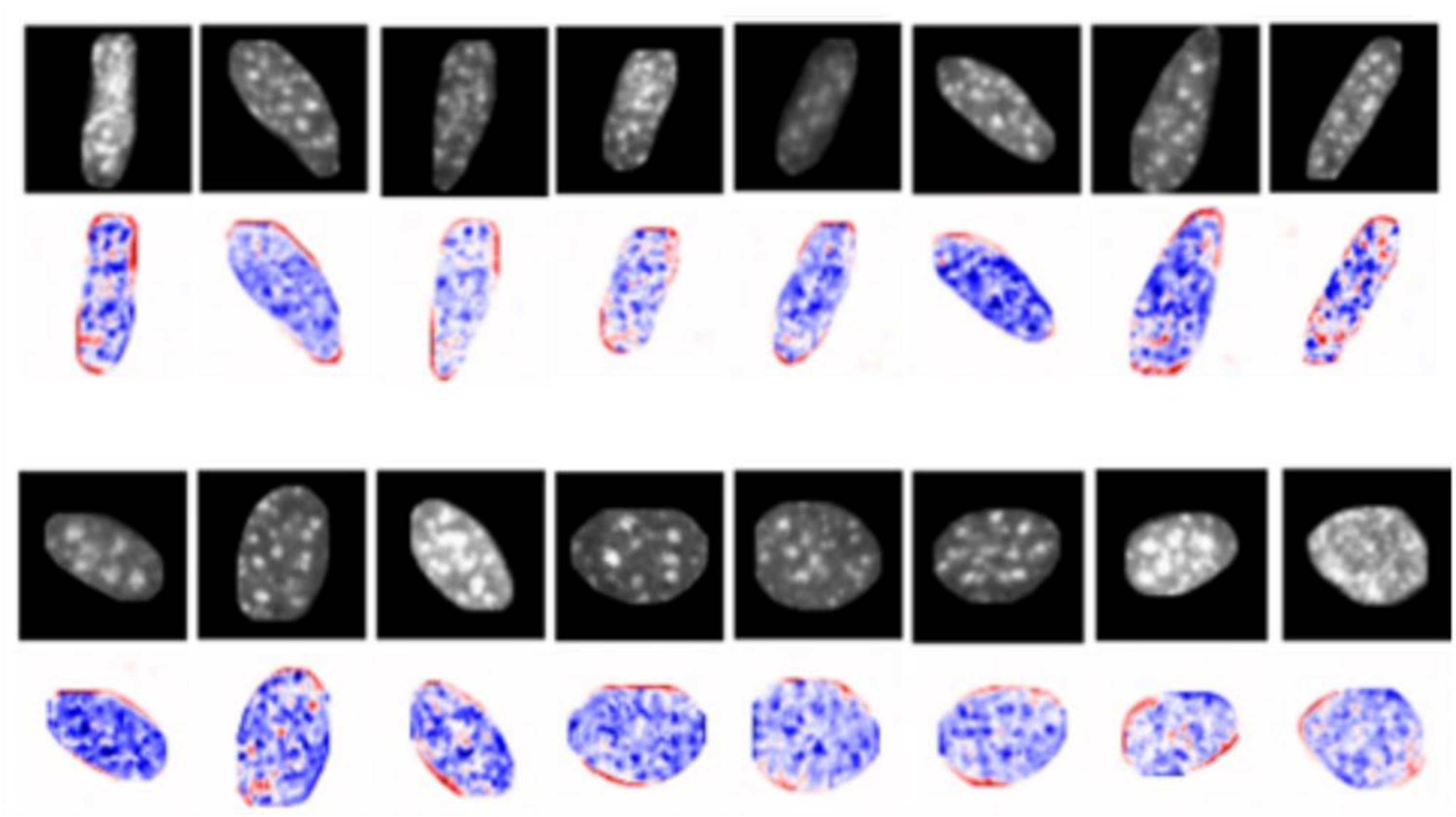
Predicted trajectories in the autoencoder feature space

Predicted trajectories mapped to the image space



Dai Yang et al. (2020)

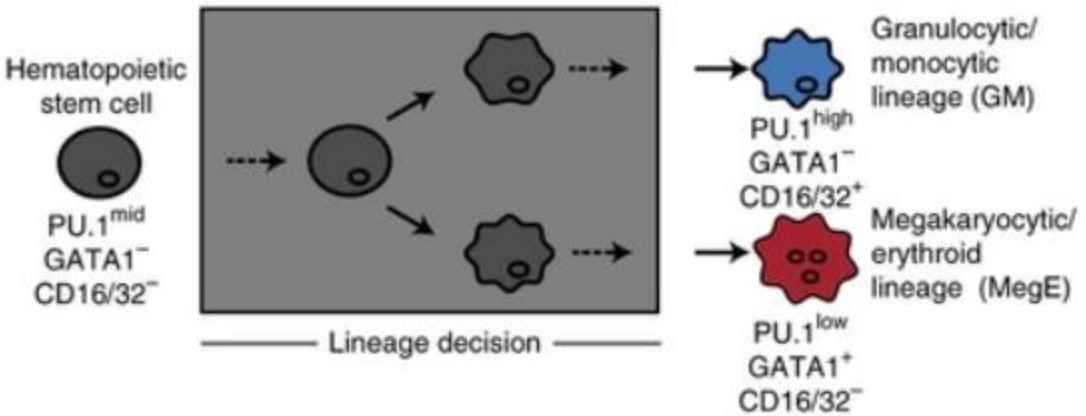
Interpretation of feature space by perturbing cell features and decoding the results to the image space



Dai Yang et al. (2020)

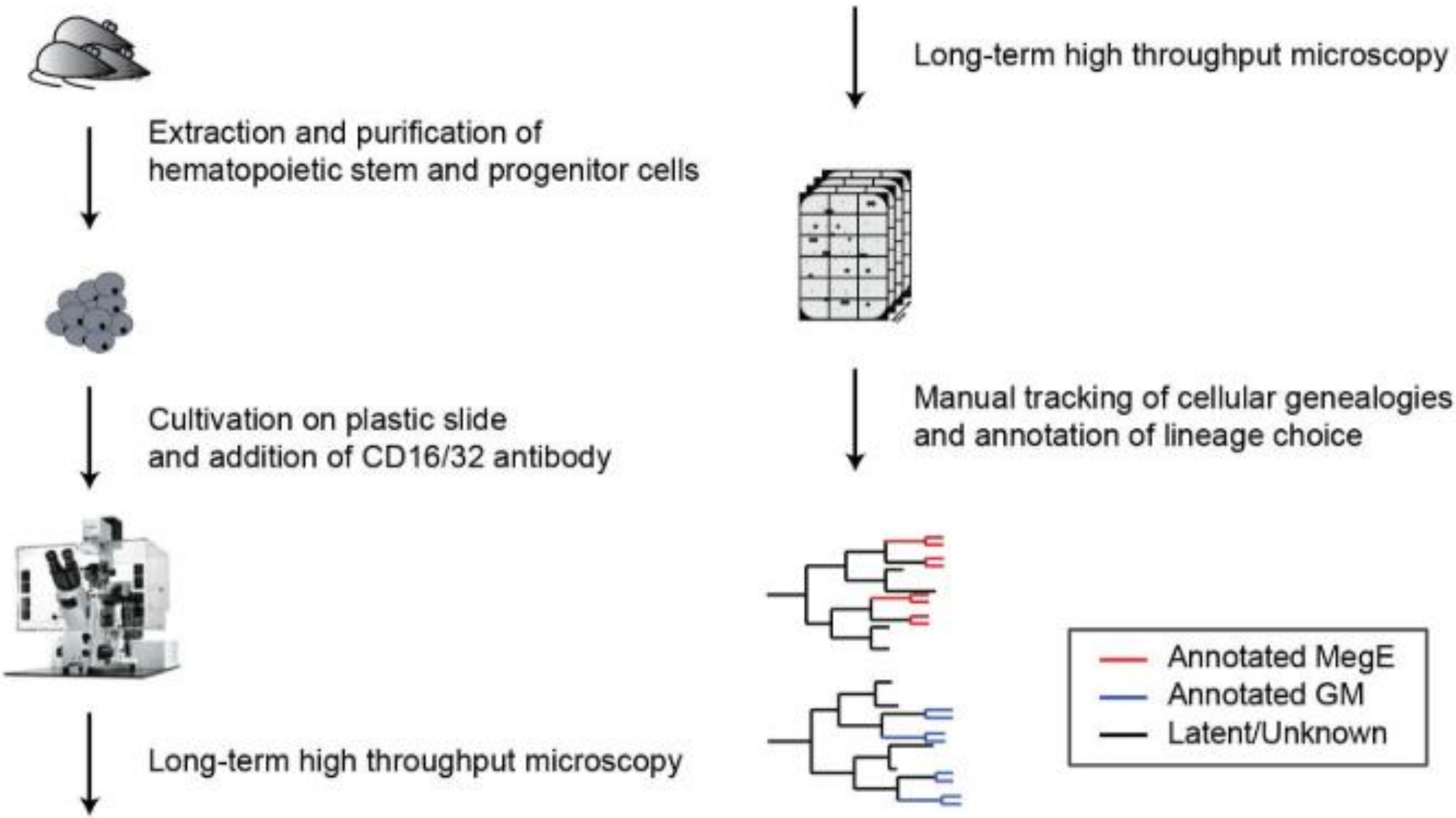
Predicting future differentiation

- Hematopoietic stem cells (HSCs) are progenitors to other blood cell types (GM or MegE lineage)
- Identification of cells with differentially expressed lineage-specifying genes without molecular labeling
- Lineage detected up to three generations before conventional molecular markers are observable
- 150 genealogies from 3 independent experiments, total 5,922 single cells manually tracked for up to 8 days (2.4M images)



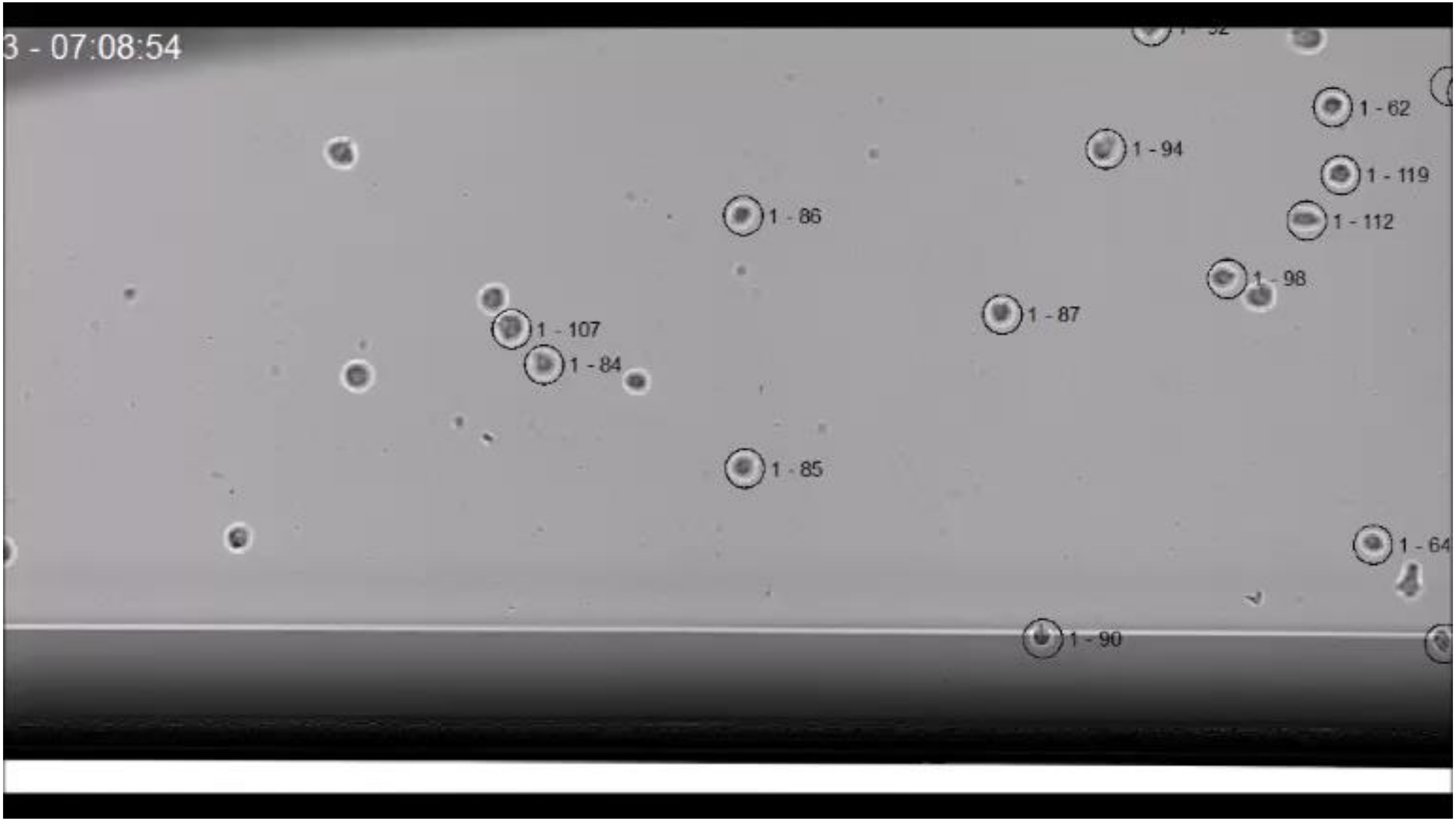
Buggenthin, Buettner, et al. (2017)

Experiment



Buggenthin, Buettner, et al. (2017)

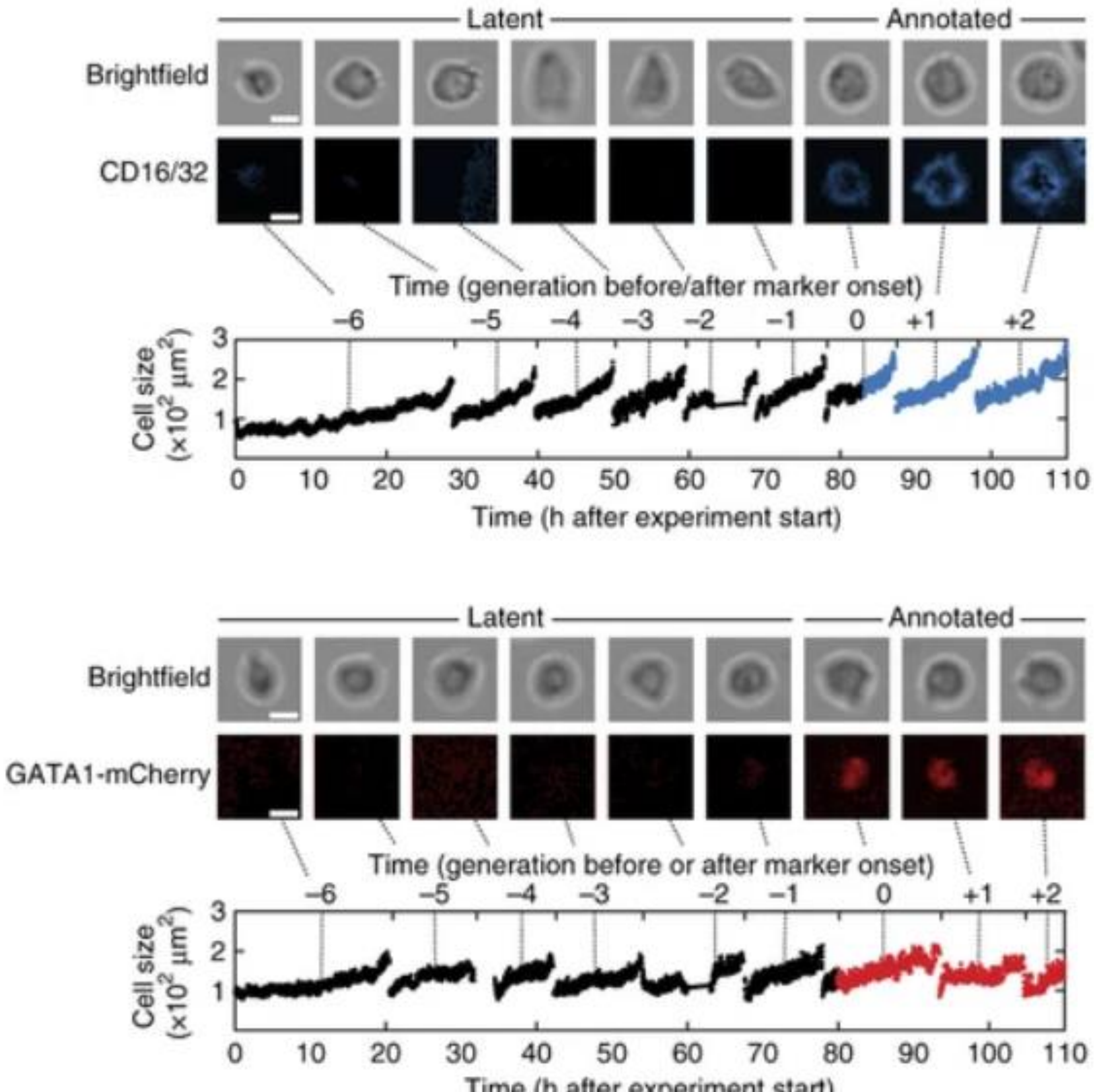
(Manual) cell tracking (26 hours)



Buggenthin, Buettner, et al. (2017)

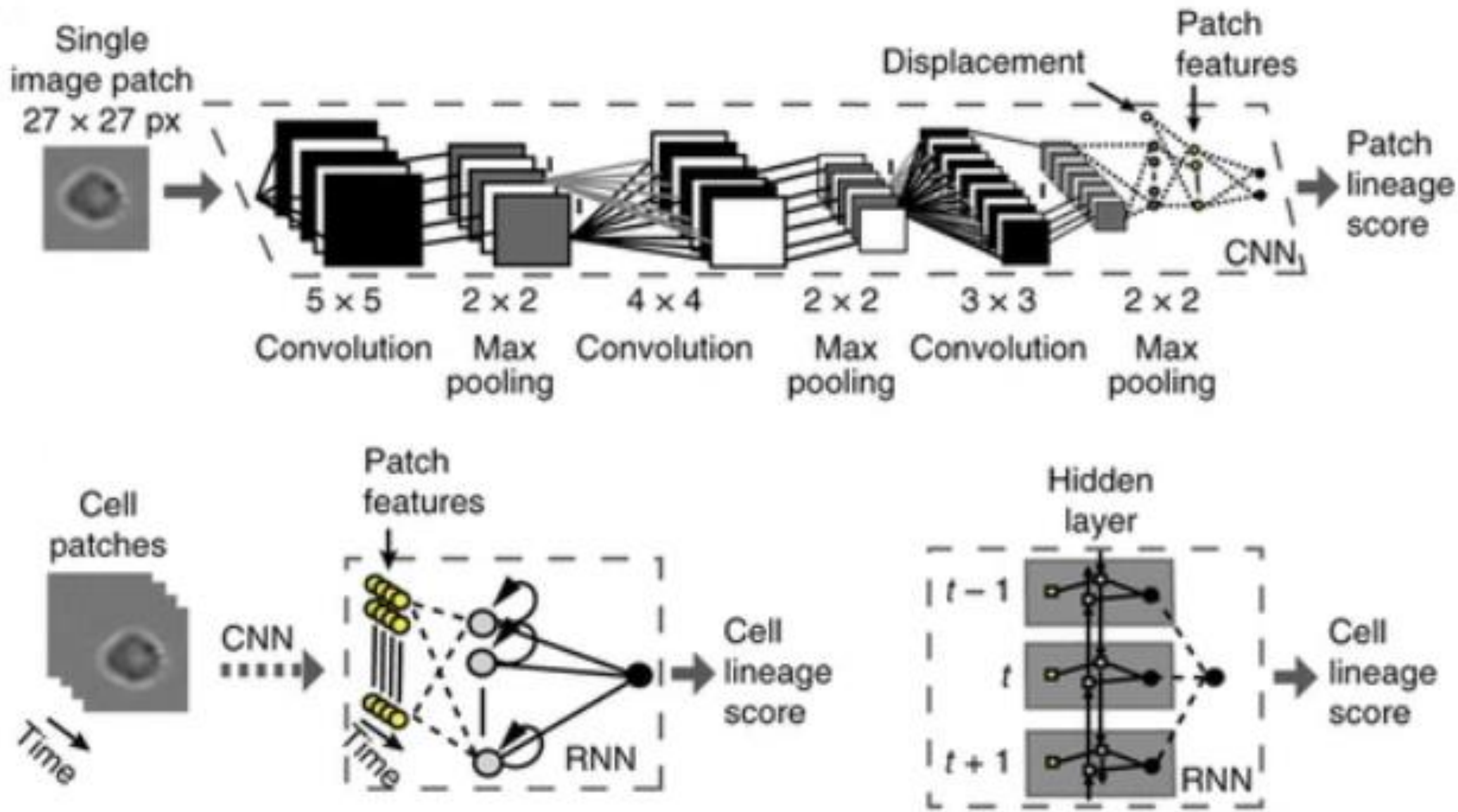
Single cells commitment

GM or MegE lineage



Buggenthin, Buettner, et al. (2017)

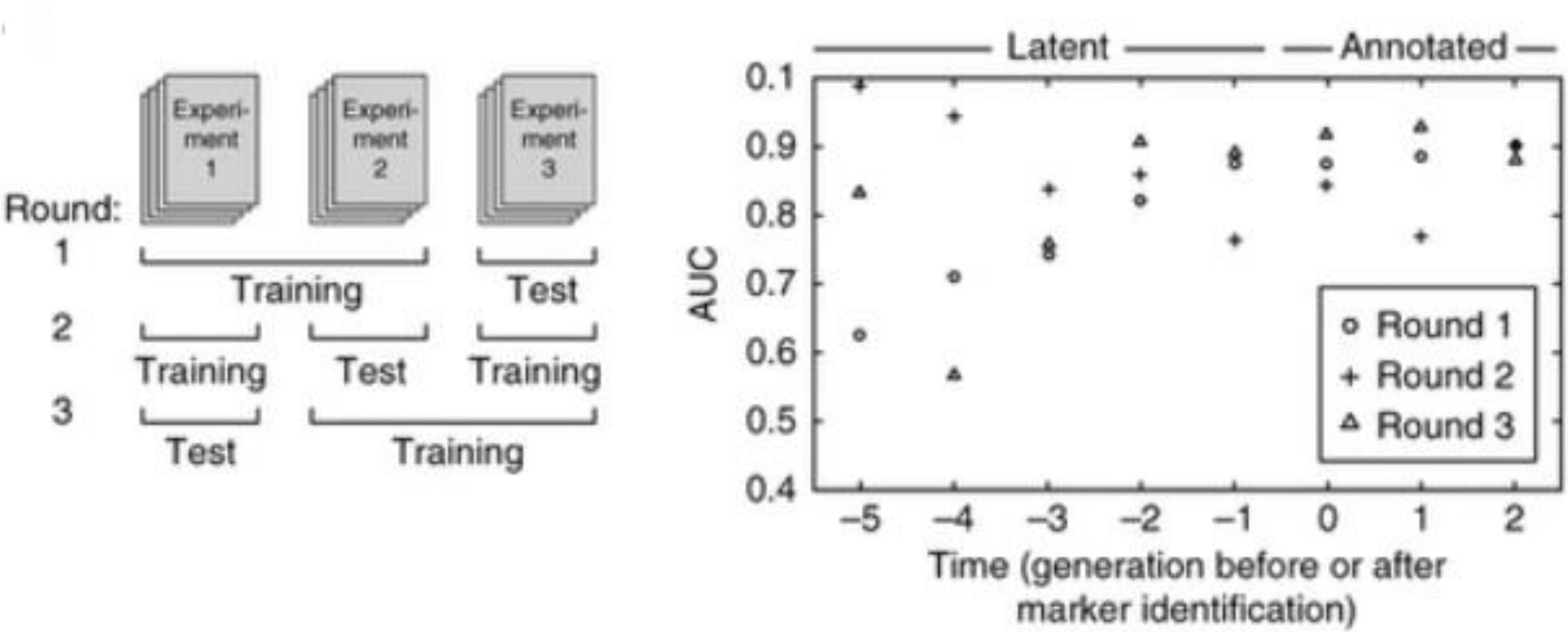
Network architecture



(There are better ways to handle temporal information,
blogposts of RNN, LSTM [here]() and their limitations [here]())

Buggenthin, Buettner, et al. (2017)

Identifying and predicting lineage from morphology and displacement



Buggenthin, Buettner, et al. (2017)