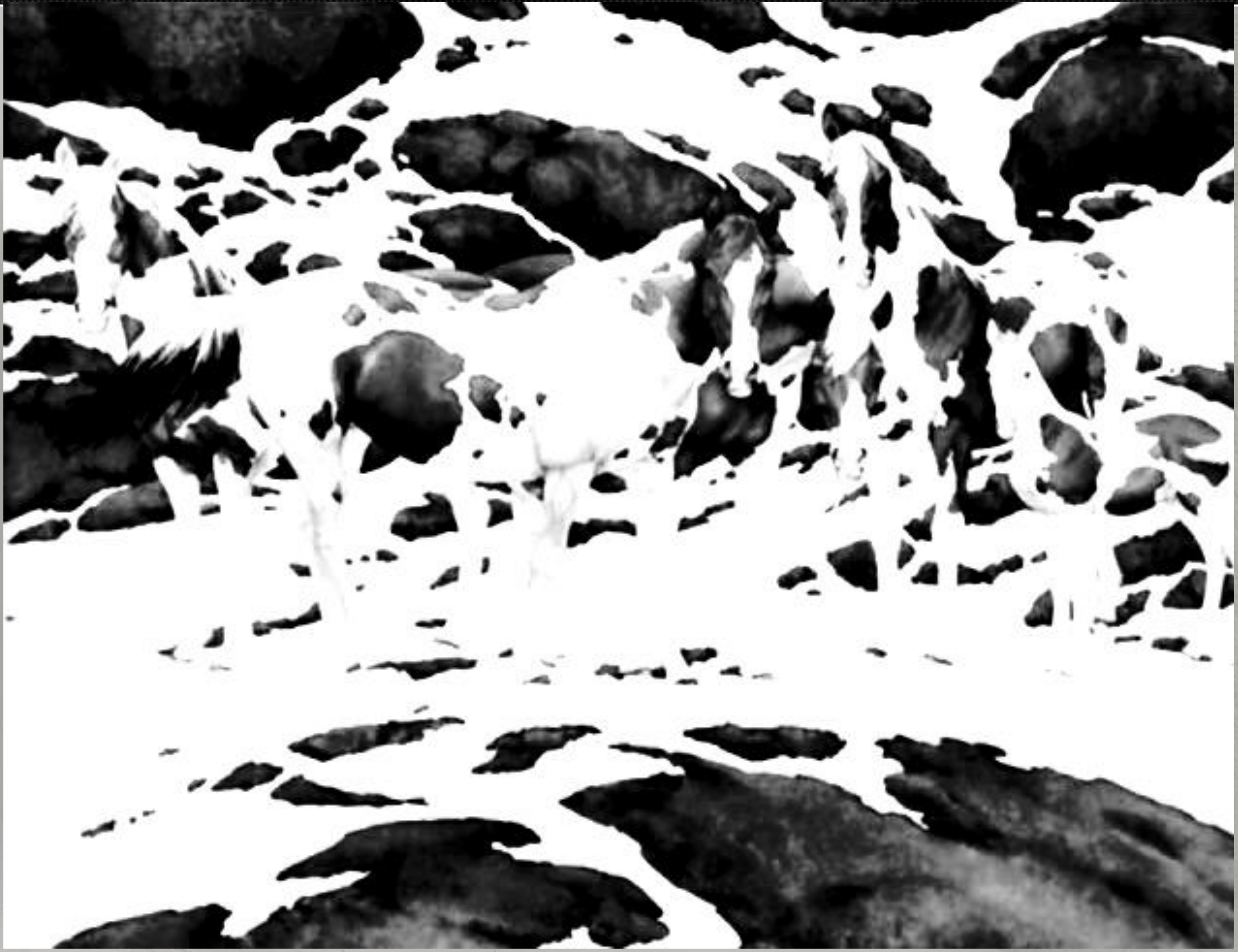
Life Beyond the Pixels

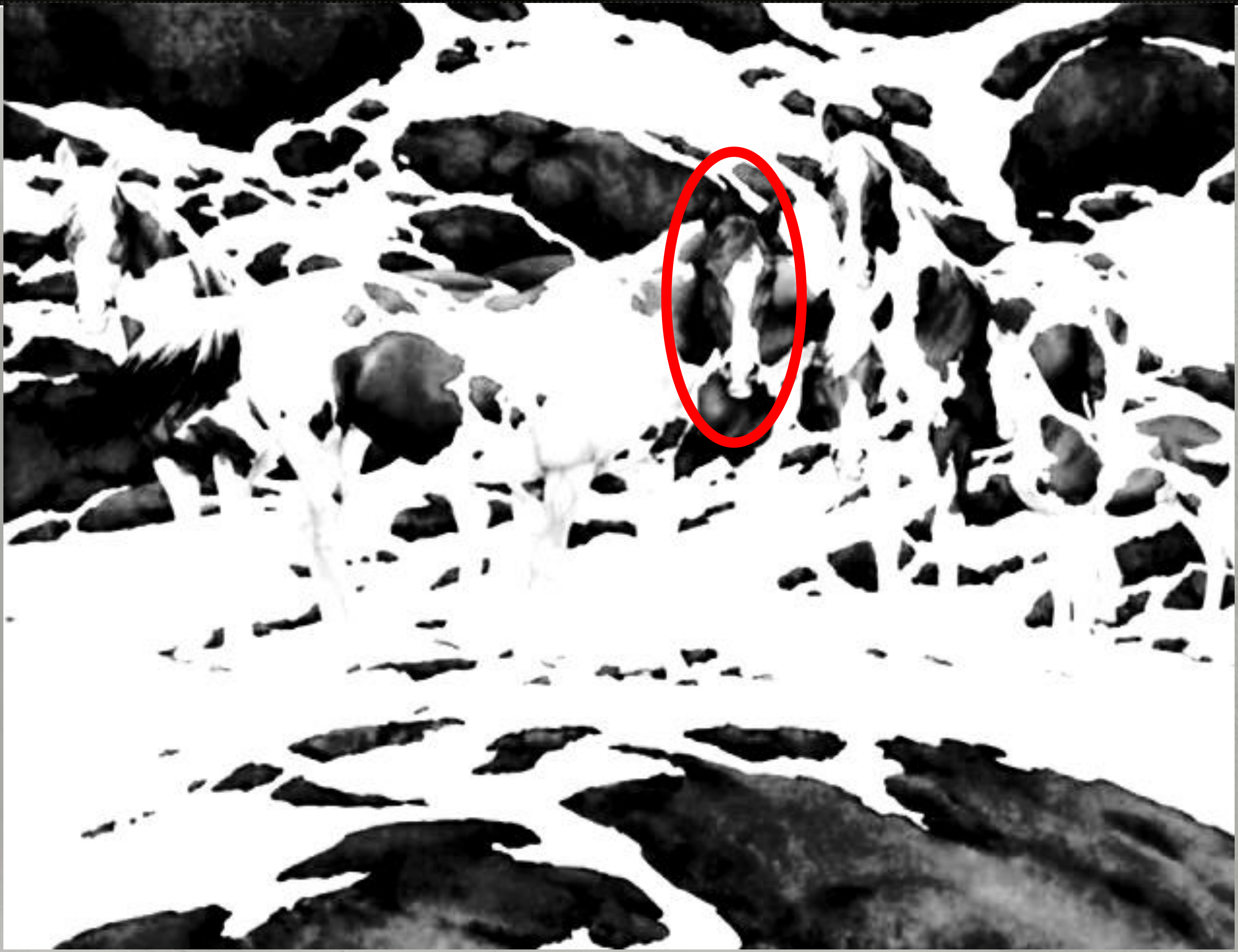
Peter Horvath
FIMM, Institute for Molecular Medicine Finland, Helsinki
BRC, Hungarian Academy of Sciences

Where…? What…? How many…?



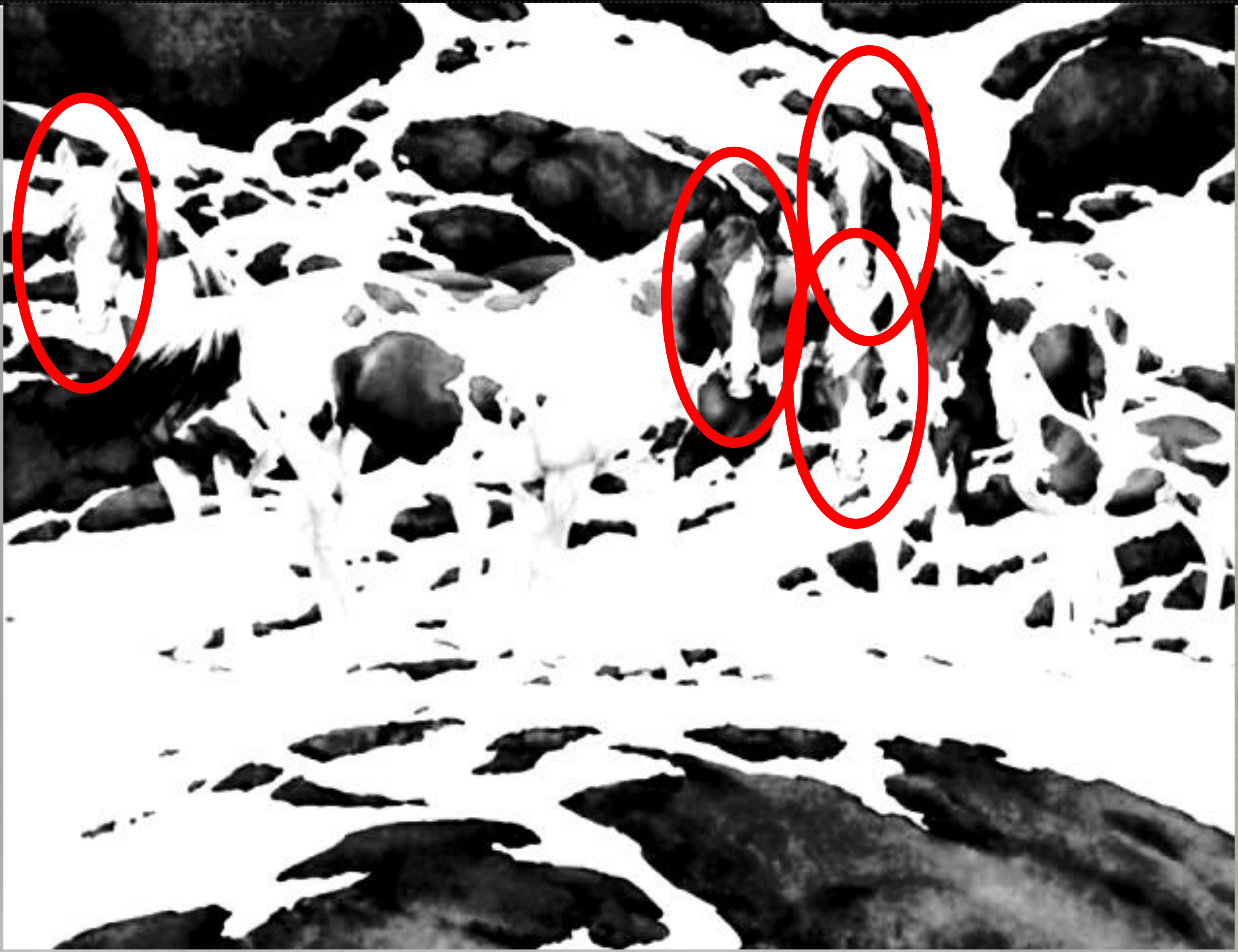
Bev Doolittle (1975)

Where…? What…? How many…?



Bev Doolittle (1975)

Where...? What...? How many...?



Bev Doolittle (1975)

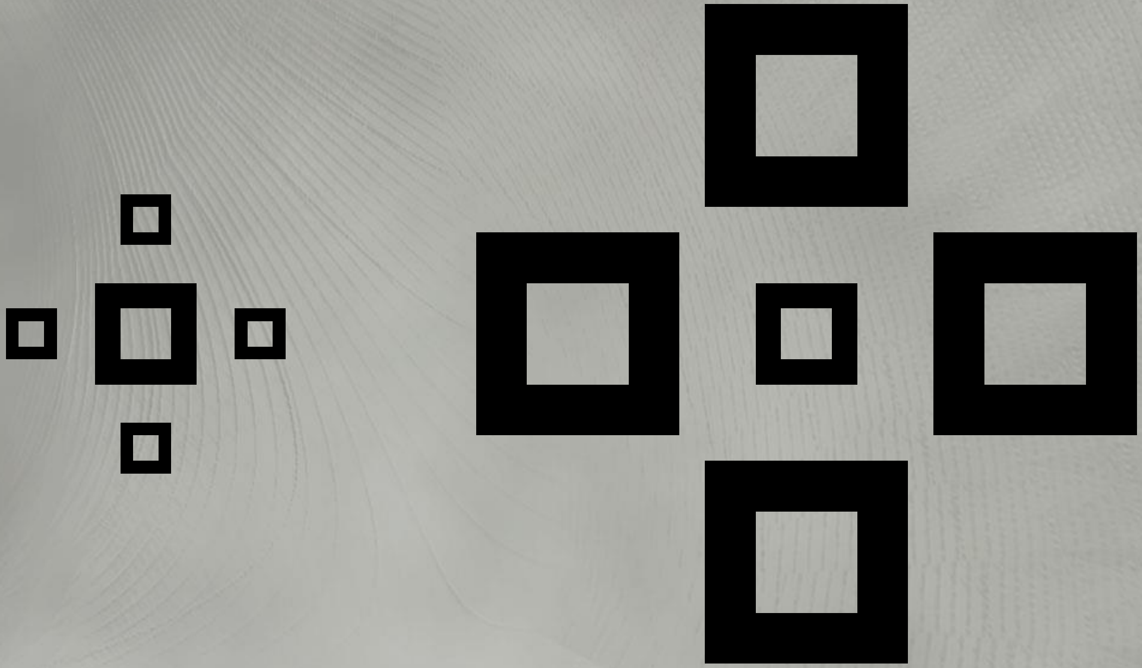
Where…? What…? How many…?

Bev Doolittle (1975)

Where...? What...? How many...?

Where…? What…? How many…?

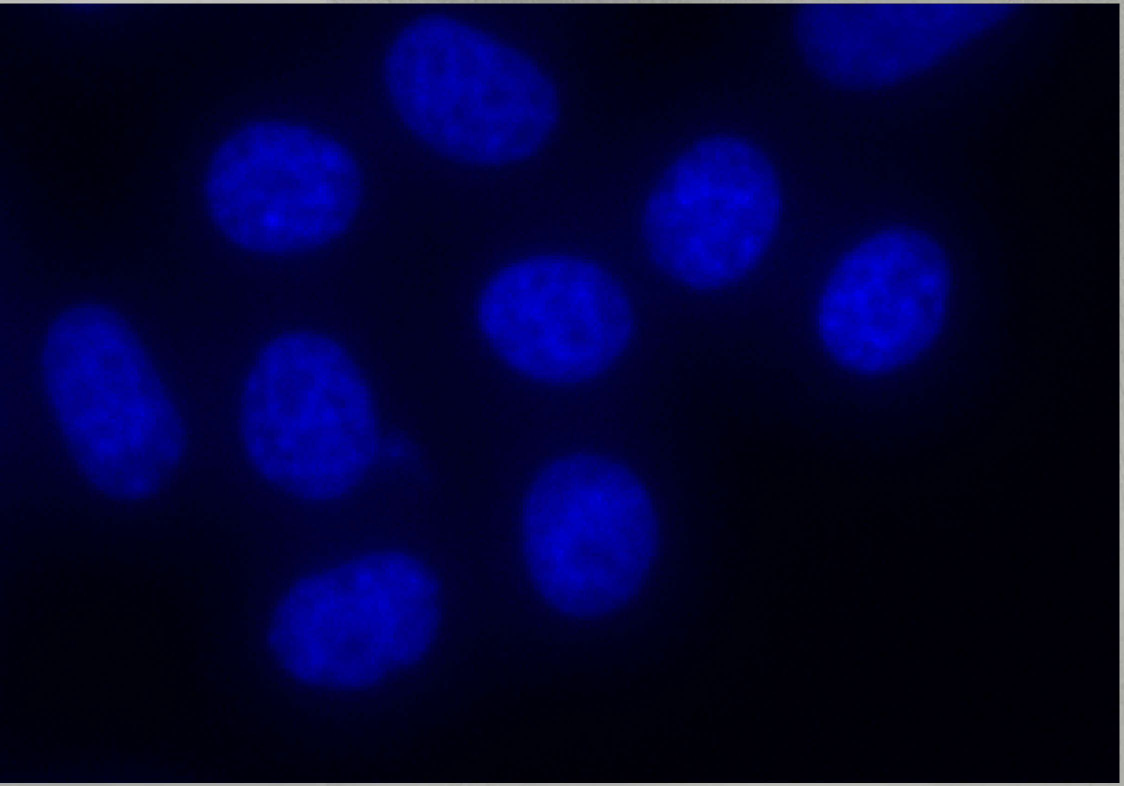
Where…? What…? How many…?



How fast can you count?

#	Human	Computer

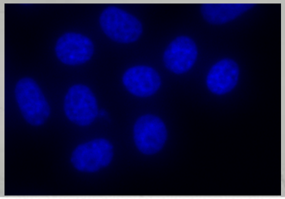
How fast can you count?



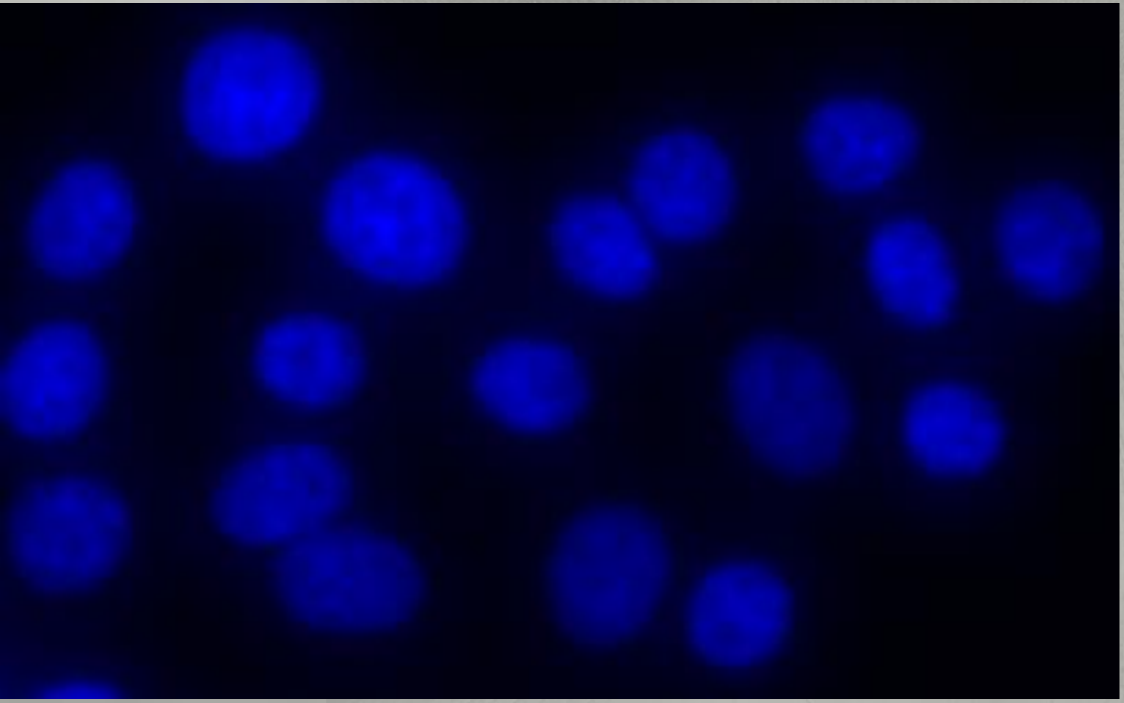
#	Human	Computer

How fast can you count?

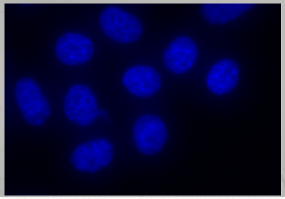
#	Human	Computer
9	4 mp	0.01 mp



How fast can you count?

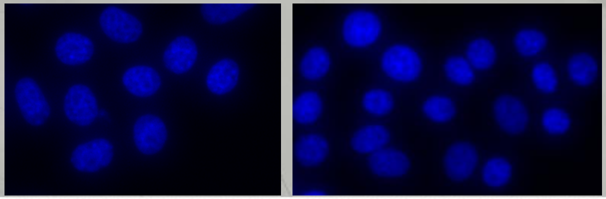


#	Human	Computer
9	4 mp	0.01 mp

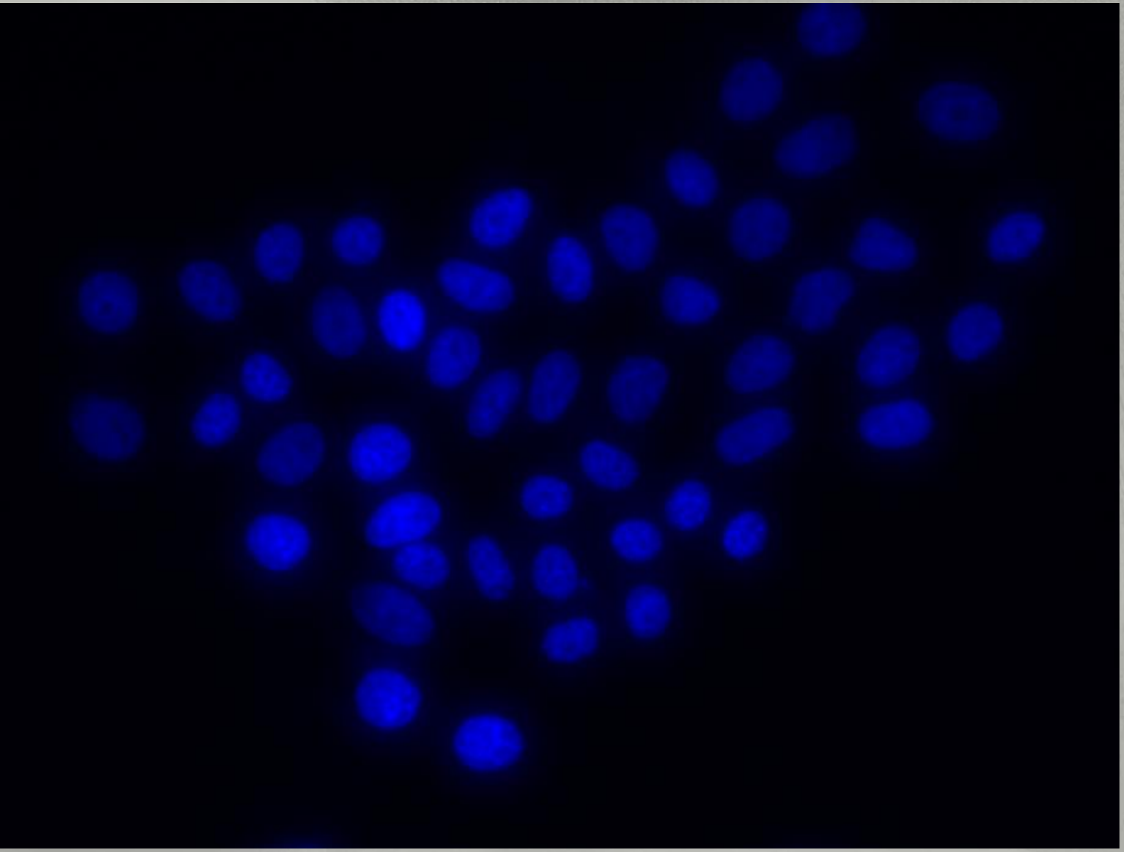


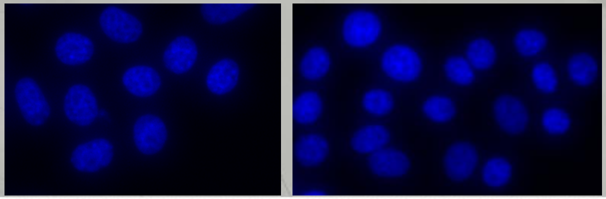
How fast can you count?

#	Human	Computer
9	4 mp	0.01 mp
18	7 mp	0.02 mp



How fast can you count?

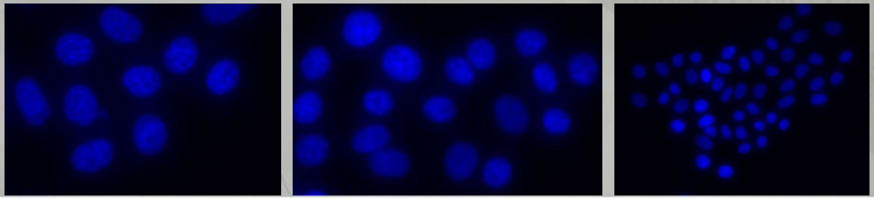




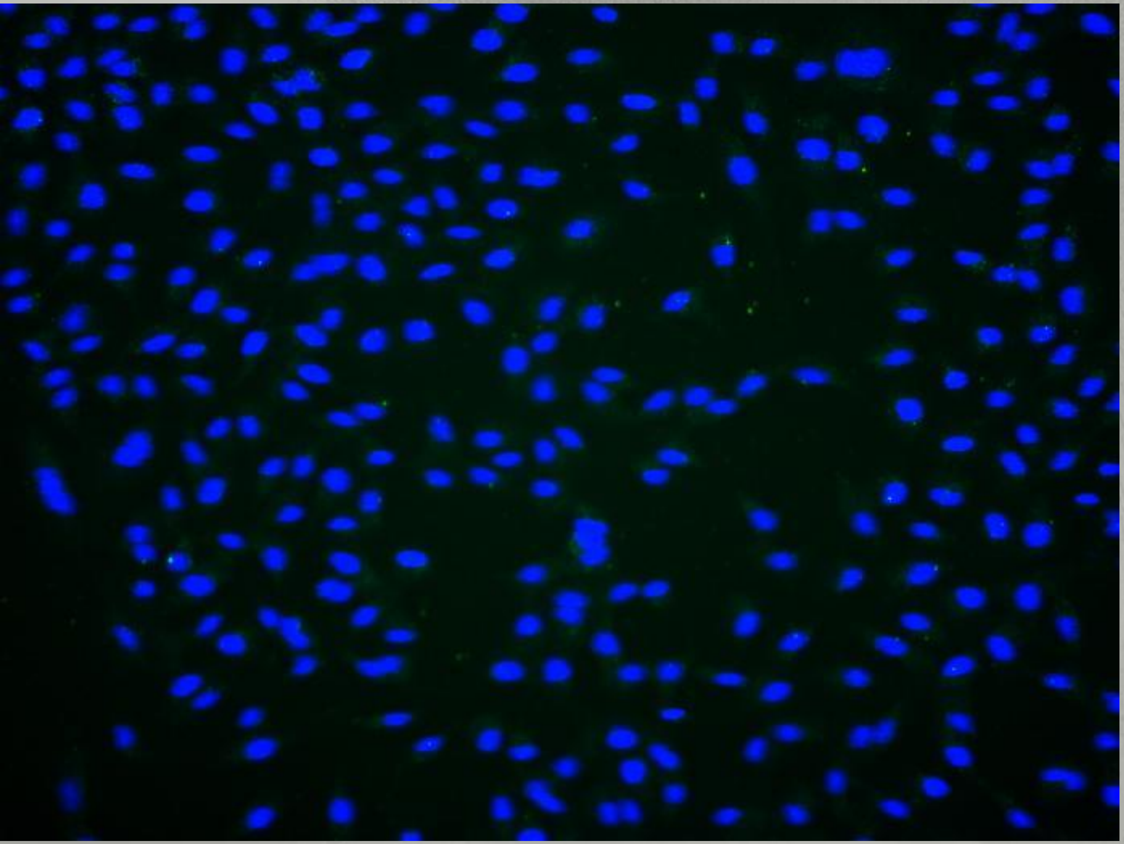
#	Human	Computer
9	4 mp	0.01 mp
18	7 mp	0.02 mp

How fast can you count?

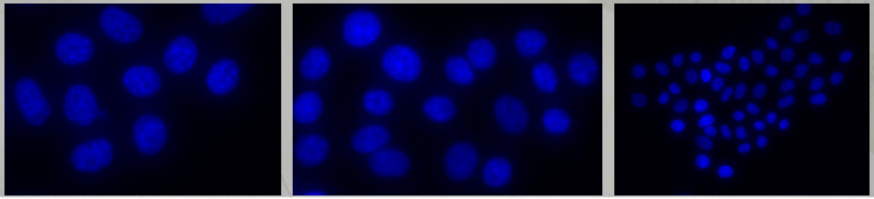
#	Human	Computer
9	4 mp	0.01 mp
18	7 mp	0.02 mp
51	17 mp	0.05 mp



How fast can you count?

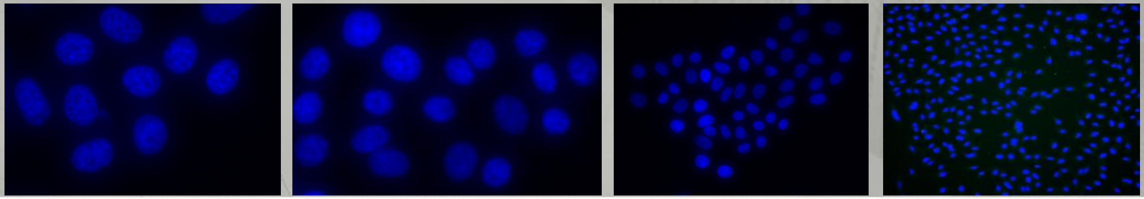


#	Human	Computer
9	4 mp	0.01 mp
18	7 mp	0.02 mp
51	17 mp	0.05 mp

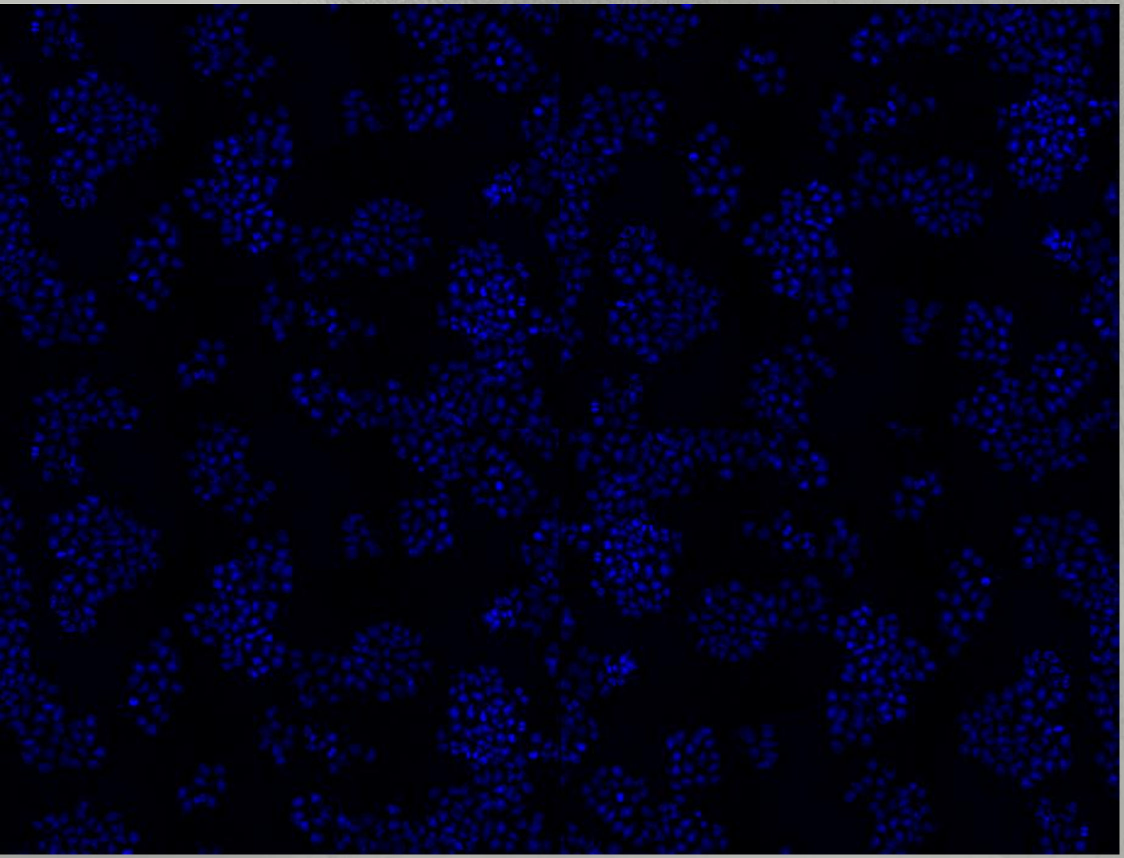


How fast can you count?

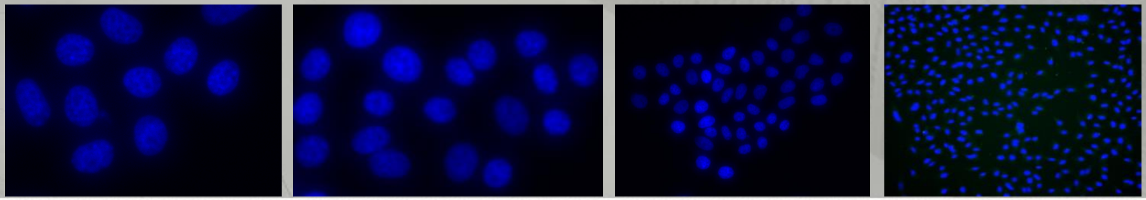
#	Human	Computer
9	4 mp	0.01 mp
18	7 mp	0.02 mp
51	17 mp	0.05 mp
210	60 mp	0.1 mp



How fast can you count?

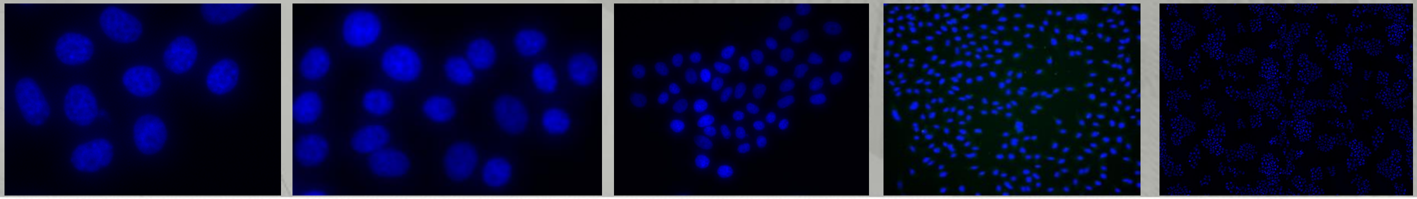


#	Human	Computer
9	4 mp	0.01 mp
18	7 mp	0.02 mp
51	17 mp	0.05 mp
210	60 mp	0.1 mp



How fast can you count?

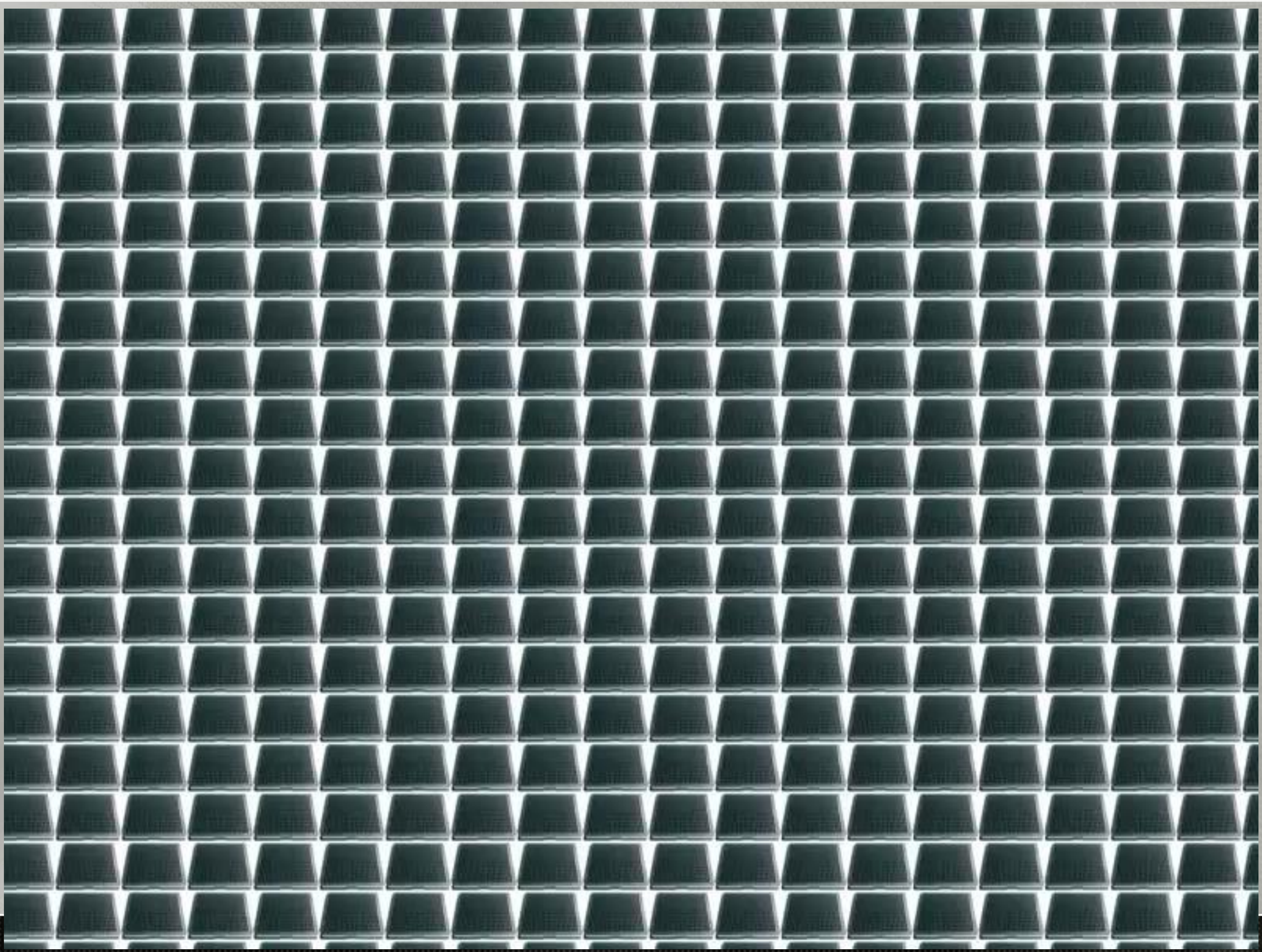
#	Human	Computer
9	4 mp	0.01 mp
18	7 mp	0.02 mp
51	17 mp	0.05 mp
210	60 mp	0.1 mp
3610	☺	1 mp



Human vs. Computer

Human:	***Computer:***
- Imprecise	***- Precise***
- Slow	***- Fast***
- Intelligent	***- Dummy*** ☺

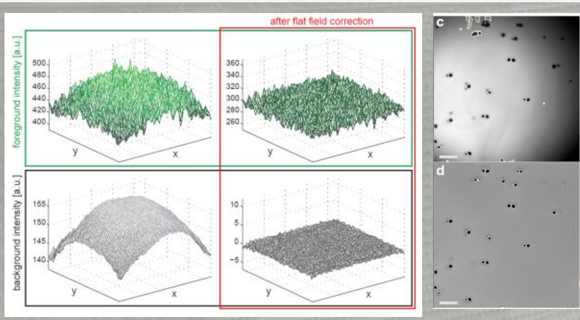
Precise, fast and intelligent



Research activities

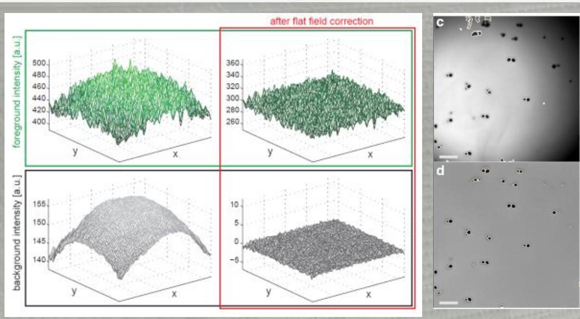
Research activities

- Image quality improvement
    - Flat field, photo bleaching



Research activities

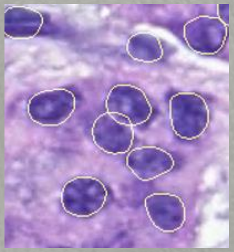
- Image quality improvement
    - Flat field, photo bleaching

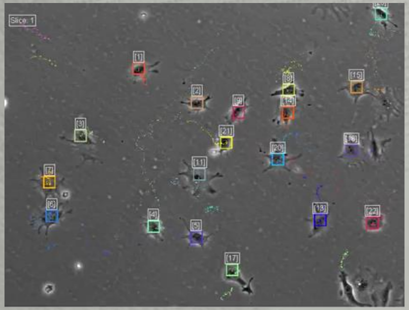


- Image segmentation and tracking

    www.celltracker.website

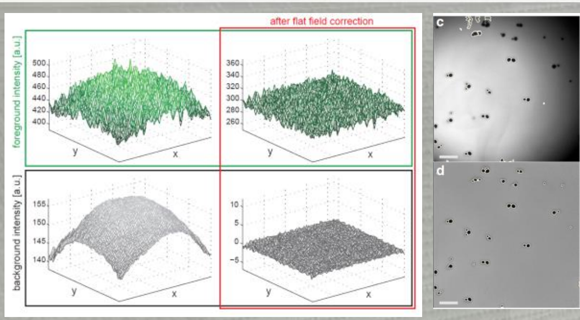
    Piccinini, Horvath. Bioinformatics 2016





Research activities

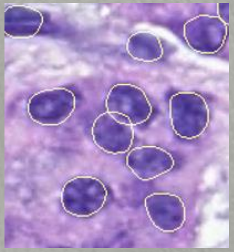
- Image quality improvement
  - Flat field, photo bleaching

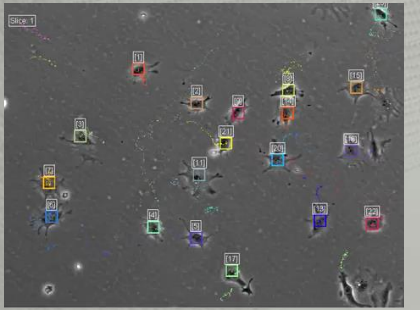


- Image segmentation and tracking

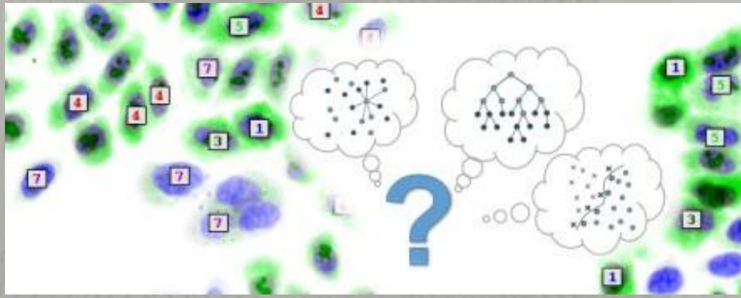
  www.celltracker.website

  Piccinini, Horvath. Bioinformatics 2016

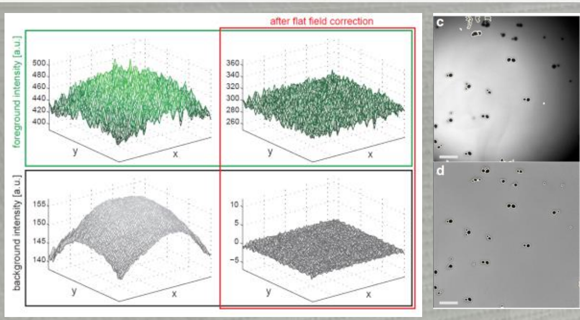




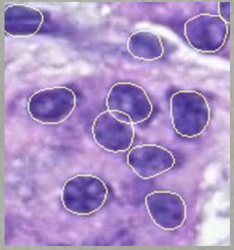
- Machine learning
  - Phenotypic profiling

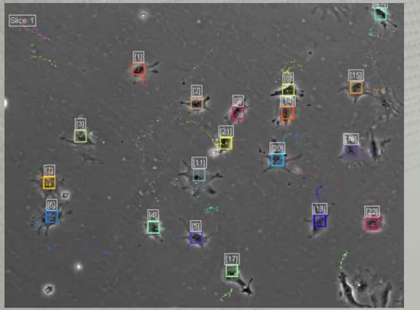


Research activities



- Image quality improvement
  - Flat field, photo bleaching

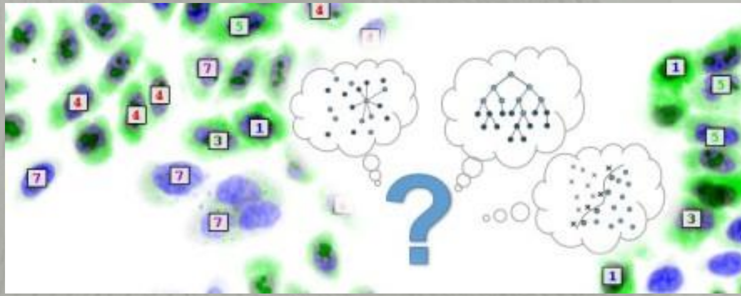




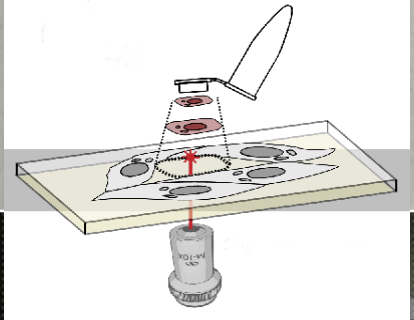
- Image segmentation and tracking

  www.celltracker.website

  Piccinini, Horvath. Bioinformatics 2016



- Machine learning
  - Phenotypic profiling

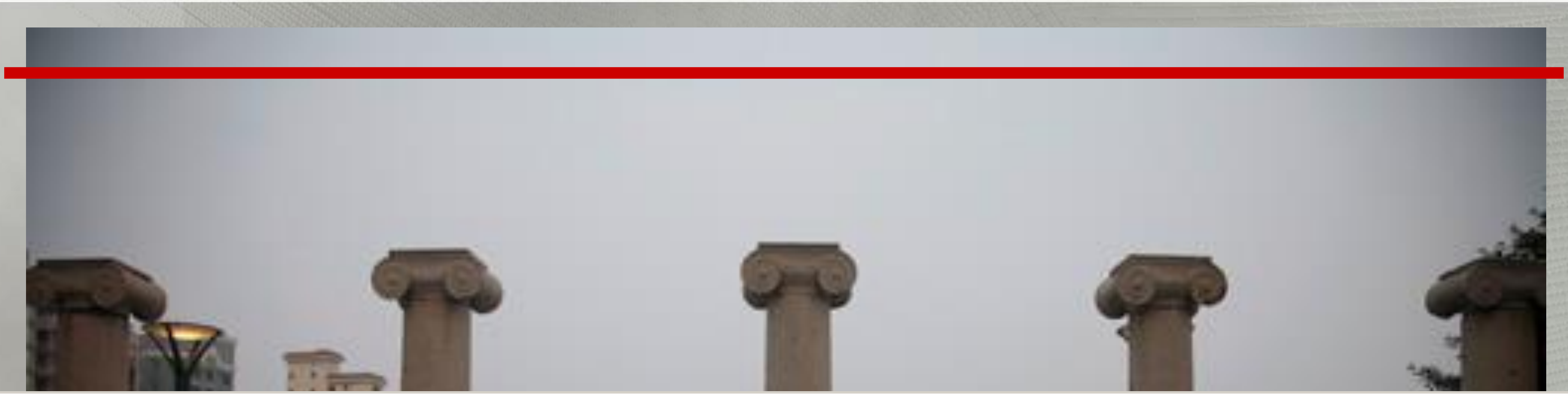


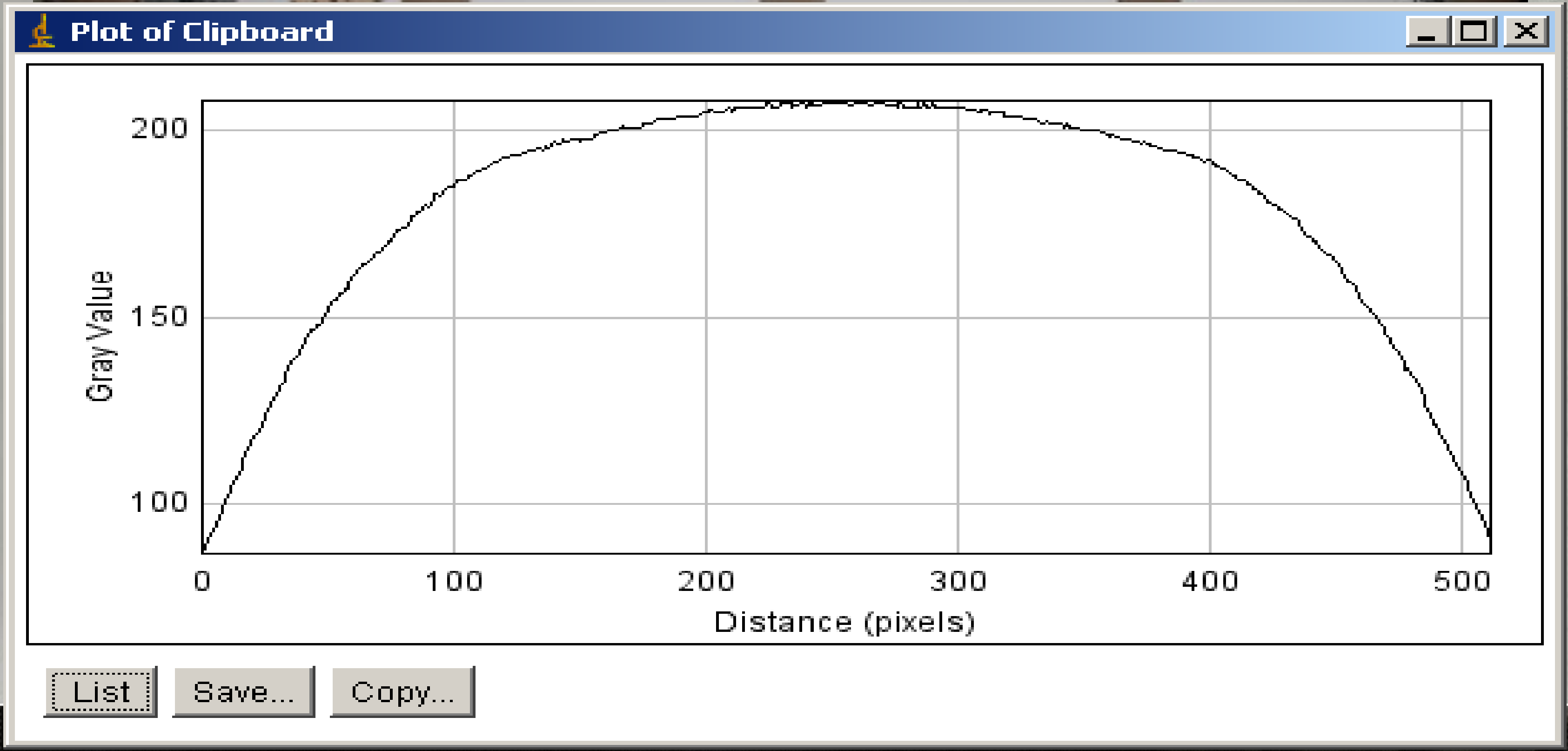
- CL^{2}M (correlative light-light microscopy)
  - High-throughput single cell isolation

Vignetting

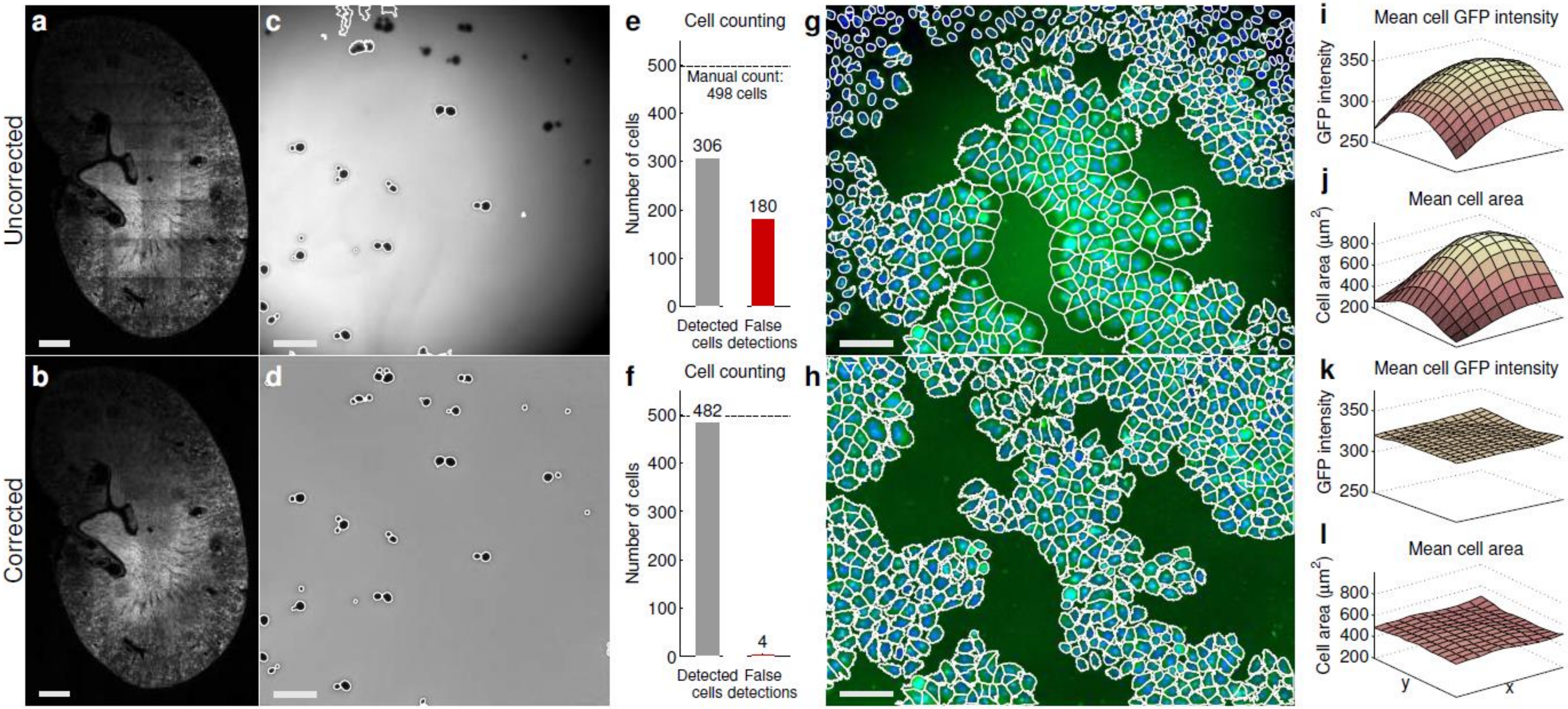
Vignetting

Vignetting

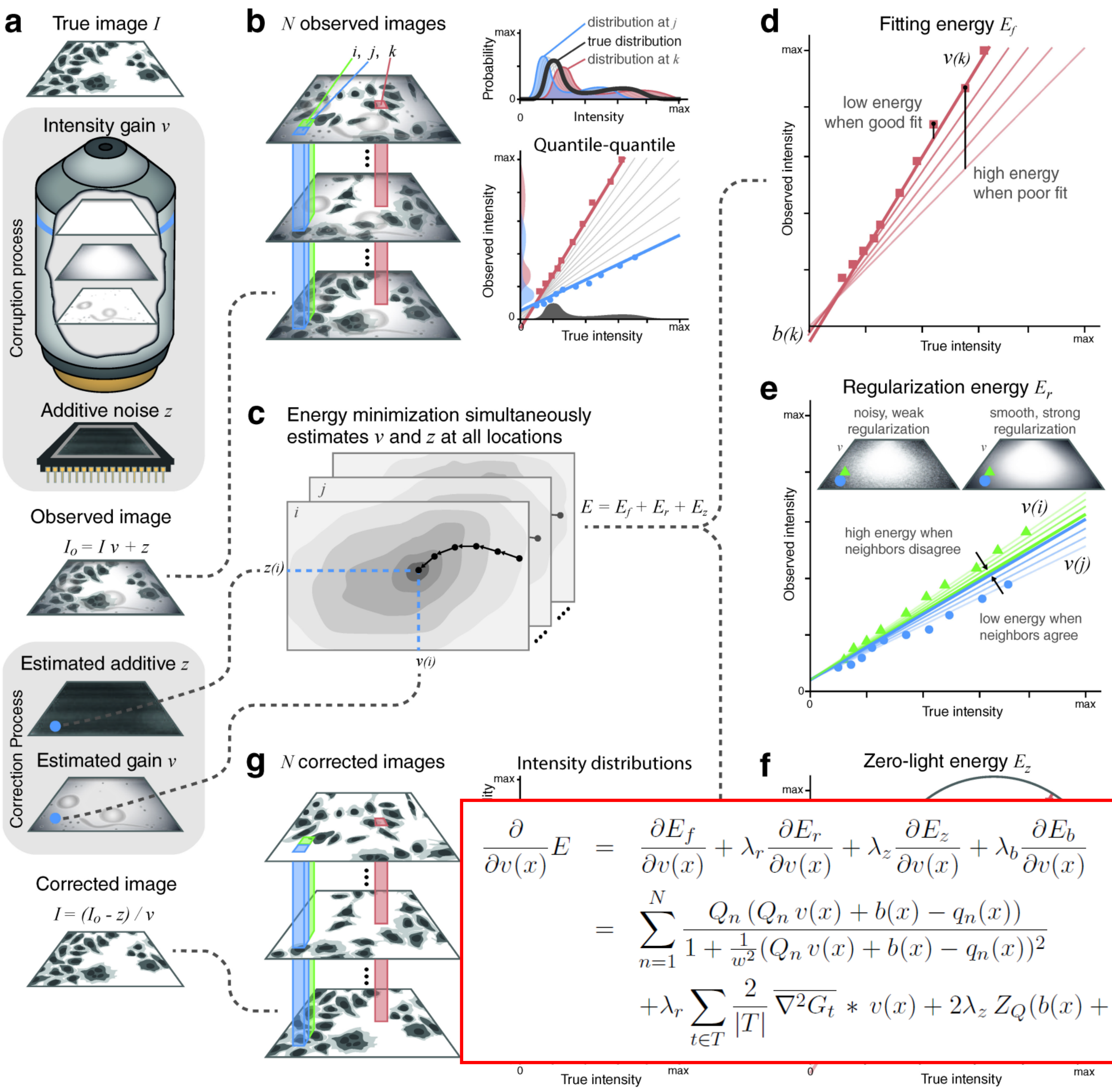



Plot of Clipboard
200
150
100
Gray Value
0
100
200
300
400
500
Distance (pixels)
List
Save...
Copy...

Illumination correction



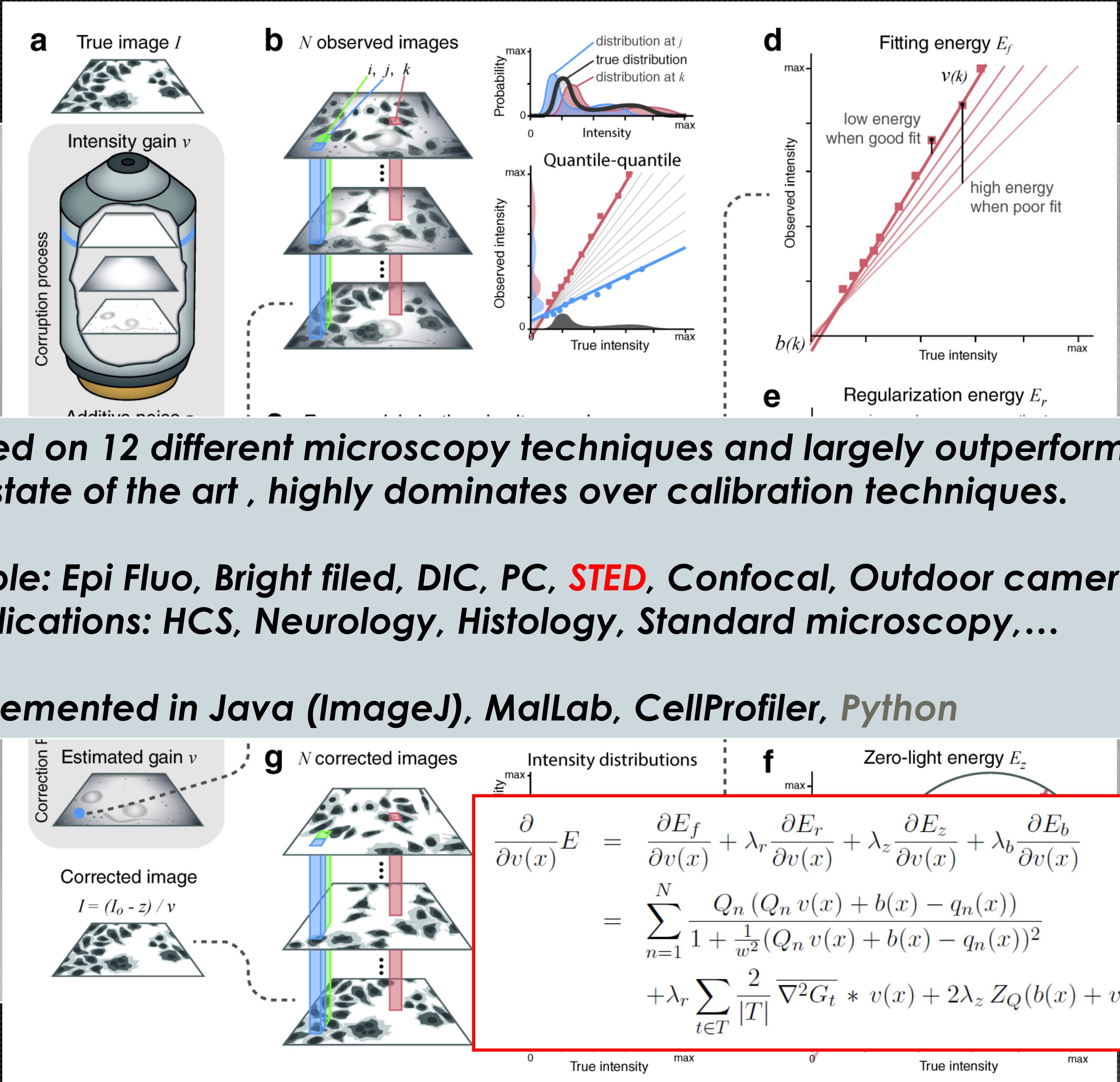
K. Smith, F. Piccinini, C. Balazs, G. Csucs, A. Bevilacqua, P. Horvath; *CIDRE: A general illumination correction method for optical microscopy*, ***Nature Methods 2015***



$$\frac{\partial}{\partial v(x)}E = \frac{\partial E_f}{\partial v(x)} + \lambda_r \frac{\partial E_r}{\partial v(x)} + \lambda_z \frac{\partial E_z}{\partial v(x)} + \lambda_b \frac{\partial E_b}{\partial v(x)}$$

$$= \sum_{n=1}^{N} \frac{Q_n \left(Q_n\, v(x) + b(x) - q_n(x)\right)}{1 + \frac{1}{w^2}\left(Q_n\, v(x) + b(x) - q_n(x)\right)^2}$$

$$+\lambda_r \sum_{t \in T} \frac{2}{|T|}\, \overline{\nabla^2 G_t} * v(x) + 2\lambda_z\, Z_Q\left(b(x) + v(x)\, Z_Q - Z_q\right).$$



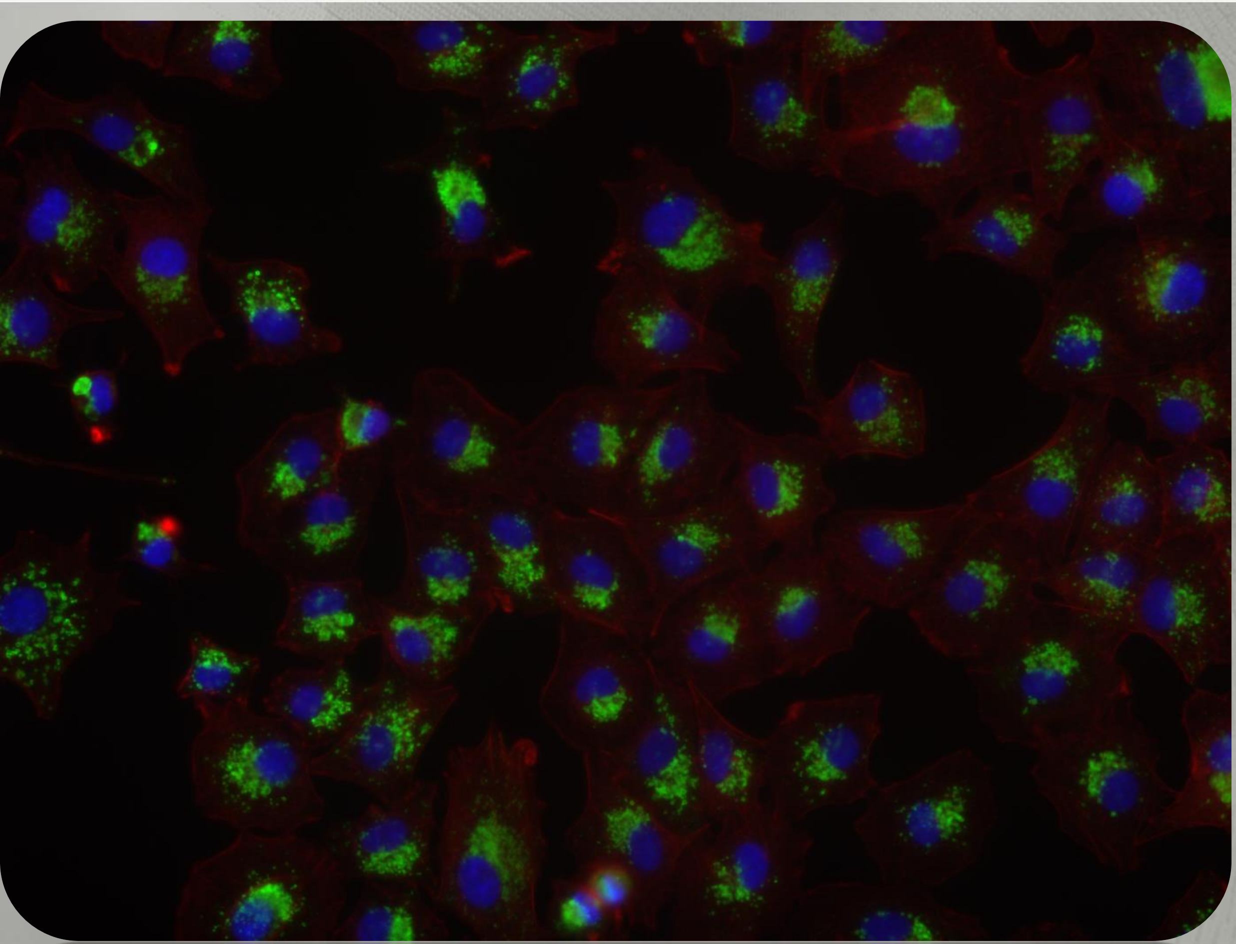
Tested on 12 different microscopy techniques and largely outperforms the state of the art , highly dominates over calibration techniques.

Usable: Epi Fluo, Bright filed, DIC, PC, STED, Confocal, Outdoor cameras
Applications: HCS, Neurology, Histology, Standard microscopy,…

Implemented in Java (ImageJ), MalLab, CellProfiler, Python

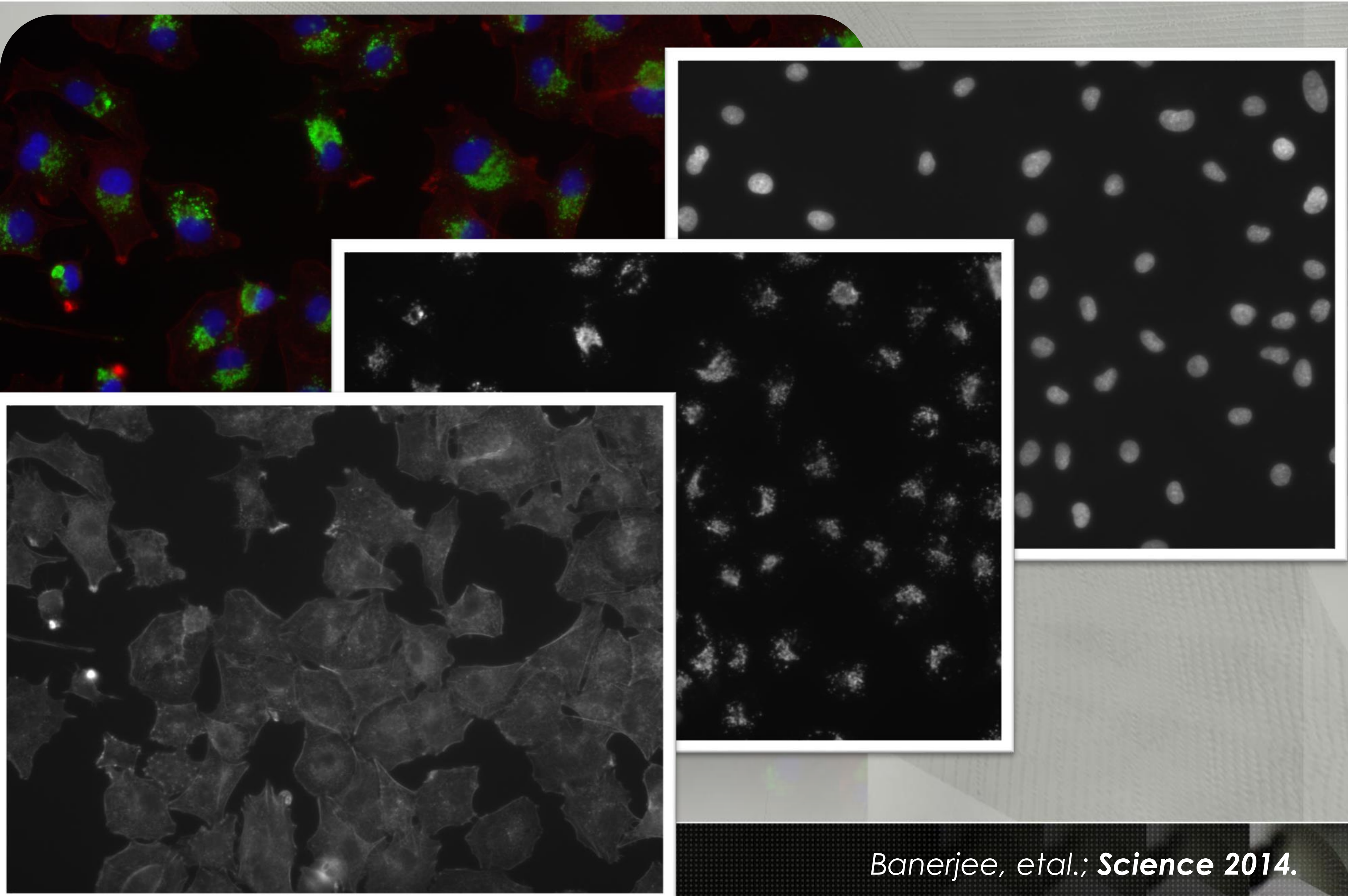
$$\begin{aligned}\frac{\partial}{\partial v(x)}E &= \frac{\partial E_f}{\partial v(x)} + \lambda_r \frac{\partial E_r}{\partial v(x)} + \lambda_z \frac{\partial E_z}{\partial v(x)} + \lambda_b \frac{\partial E_b}{\partial v(x)} \\ &= \sum_{n=1}^{N} \frac{Q_n\,(Q_n\,v(x) + b(x) - q_n(x))}{1 + \frac{1}{w^2}(Q_n\,v(x) + b(x) - q_n(x))^2} \\ &\quad + \lambda_r \sum_{t \in T} \frac{2}{|T|}\,\overline{\nabla^2 G_t} * v(x) + 2\lambda_z\, Z_Q(b(x) + v(x)\,Z_Q - Z_q).\end{aligned}$$

Image processing



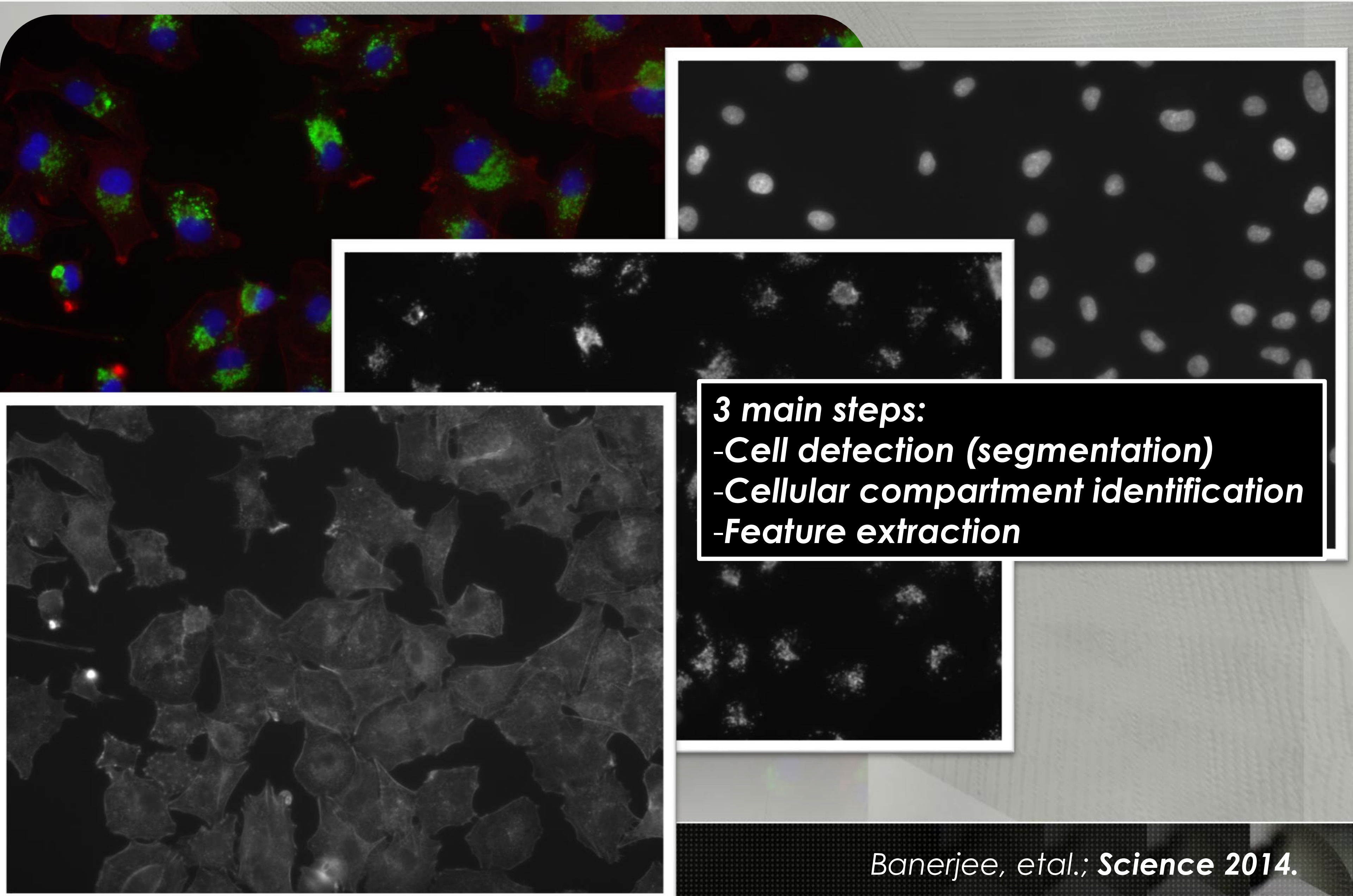
Banerjee, etal.; ***Science 2014.***

Image processing



Banerjee, etal.; ***Science 2014.***

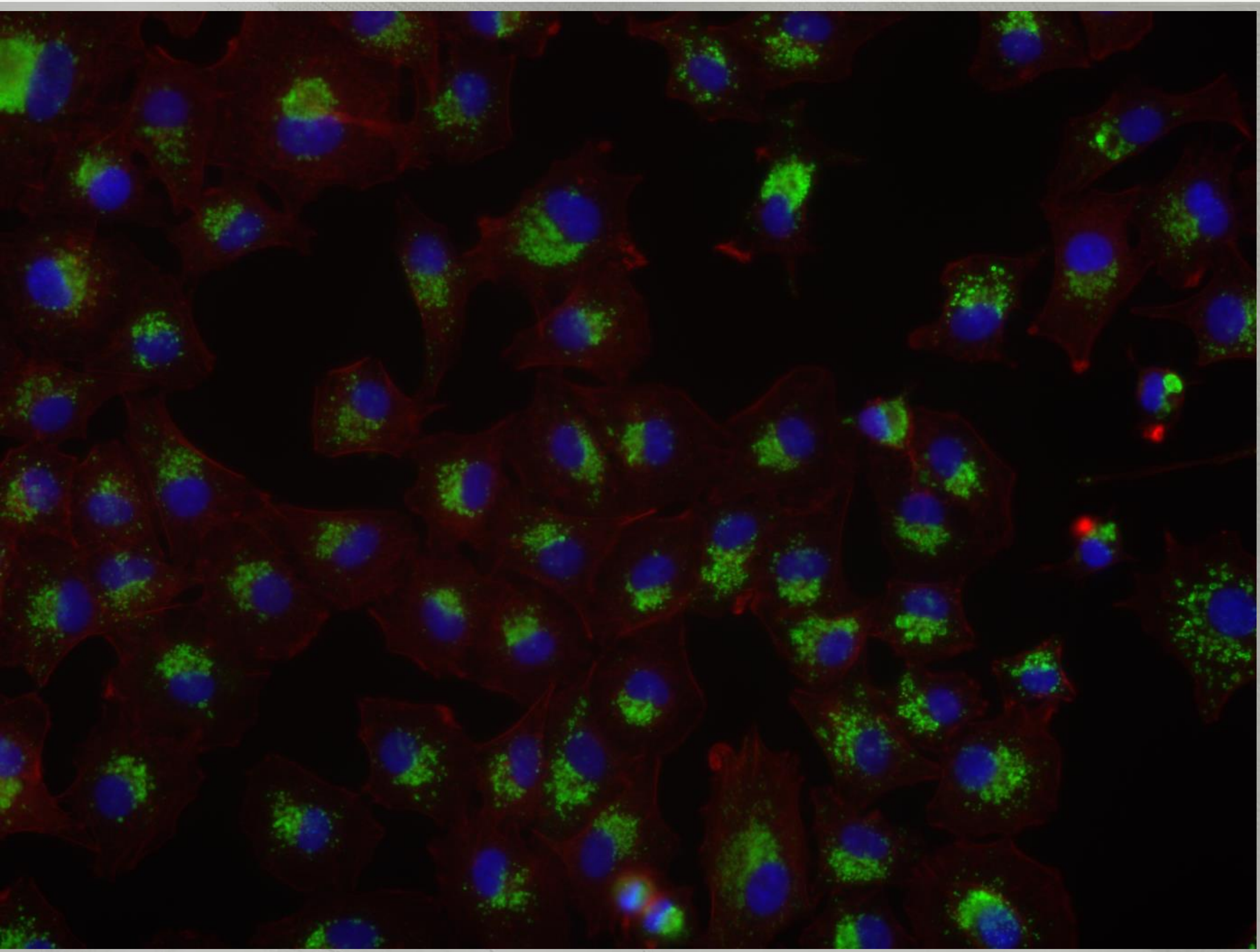
Image processing



3 main steps:
- *Cell detection (segmentation)*
- *Cellular compartment identification*
- *Feature extraction*

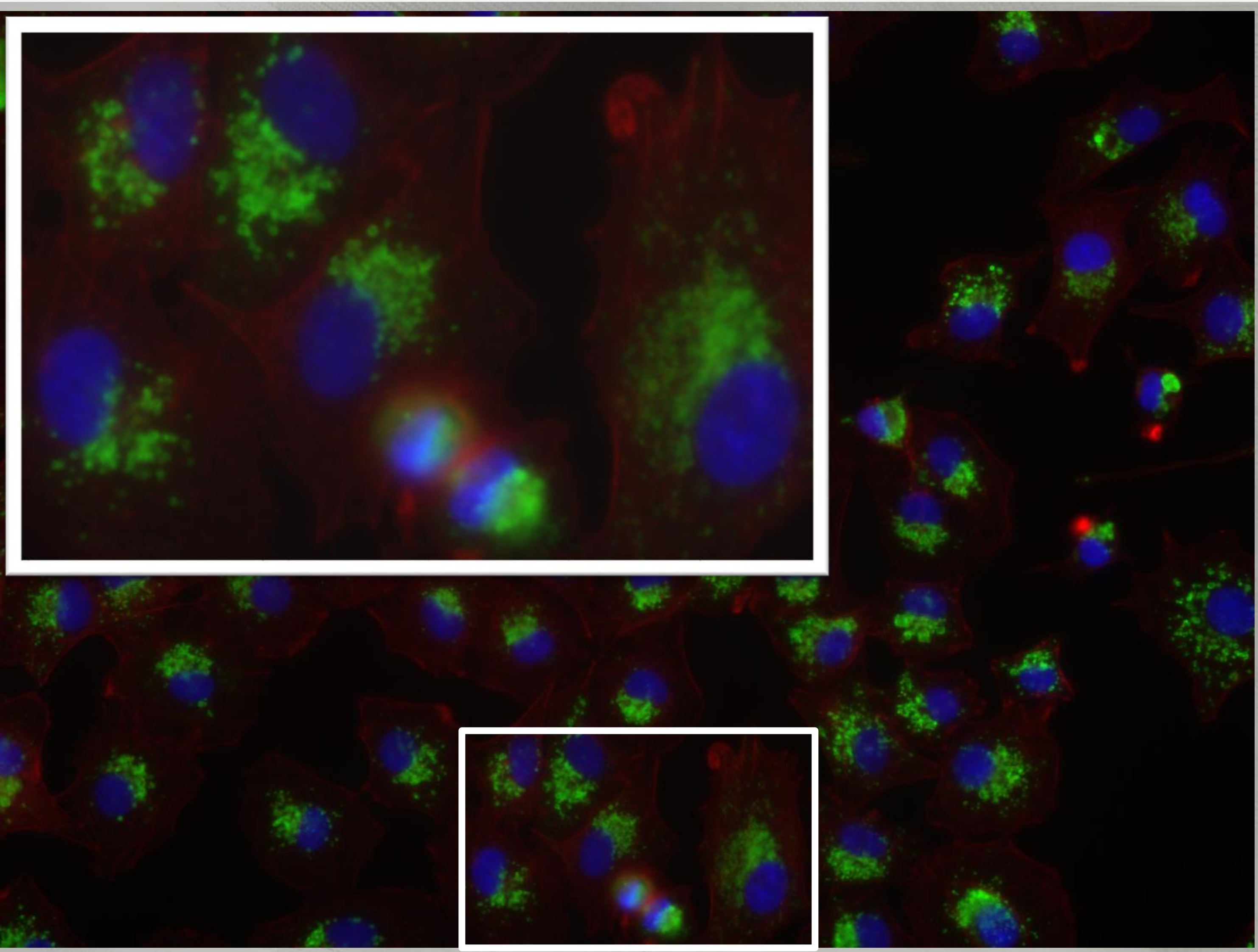
Banerjee, etal.; ***Science 2014.***

Image processing I. – cell detection



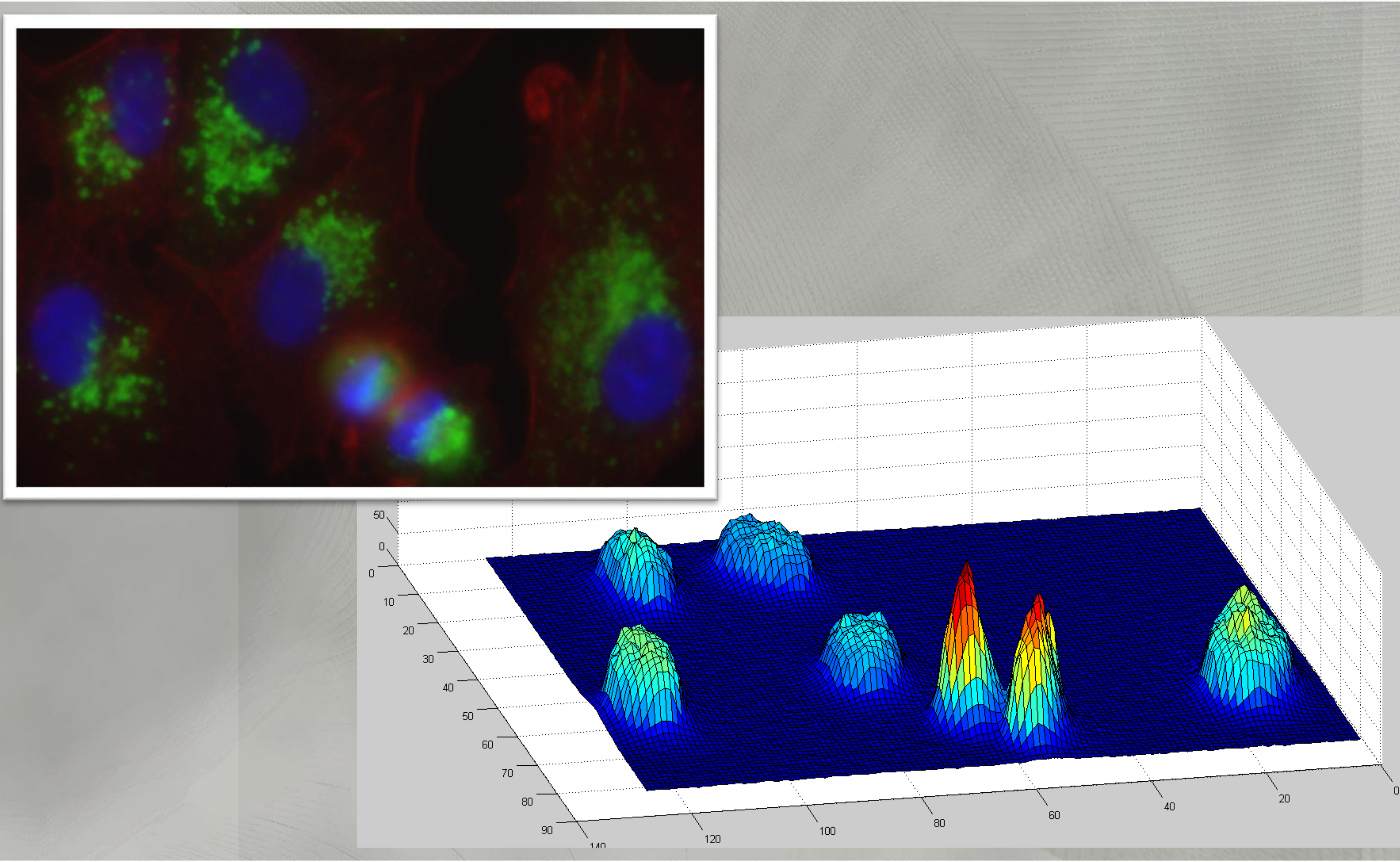
Banerjee, etal.; ***Science 2014.***

Image processing I. – cell detection



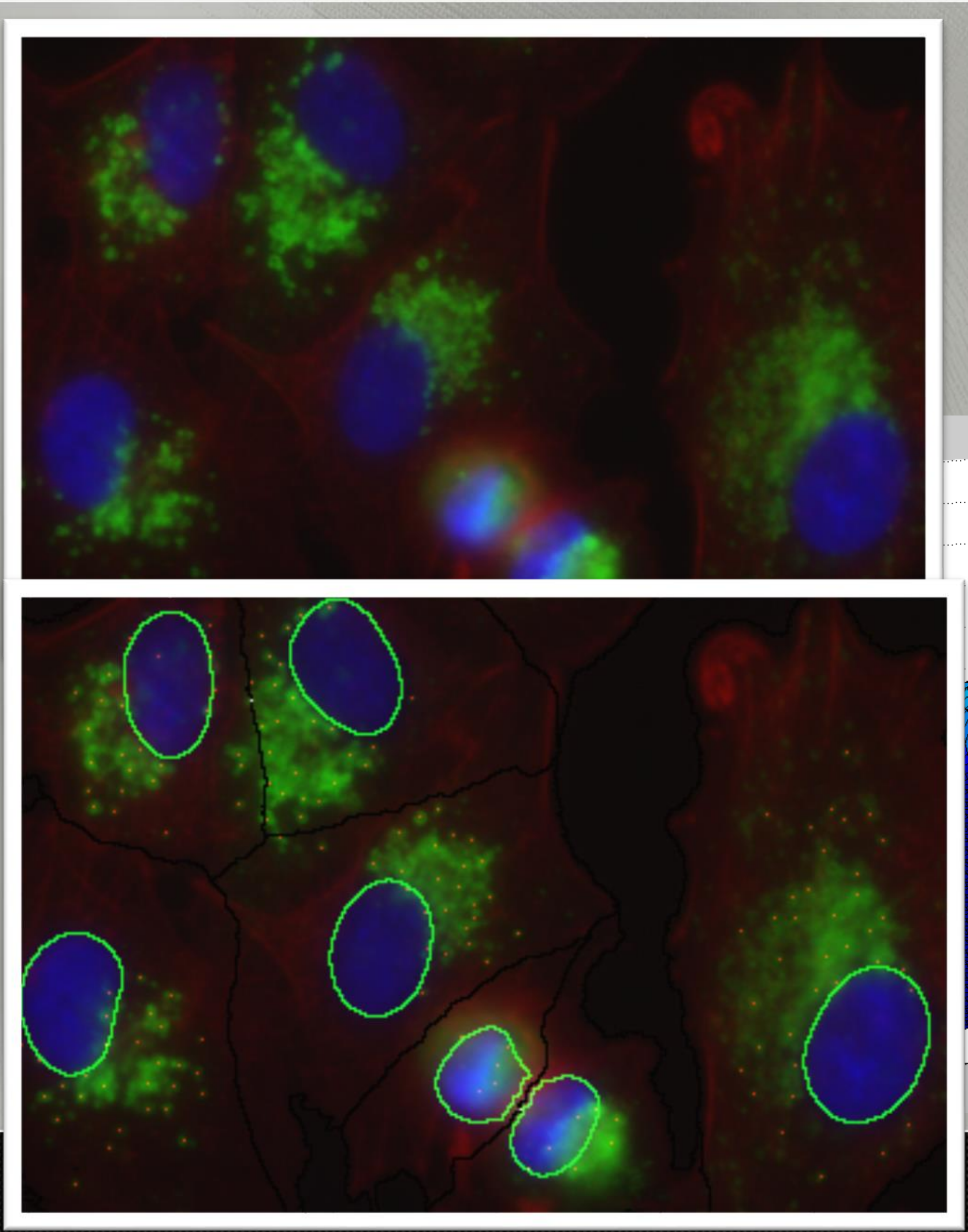
Banerjee, etal.; **Science 2014.**

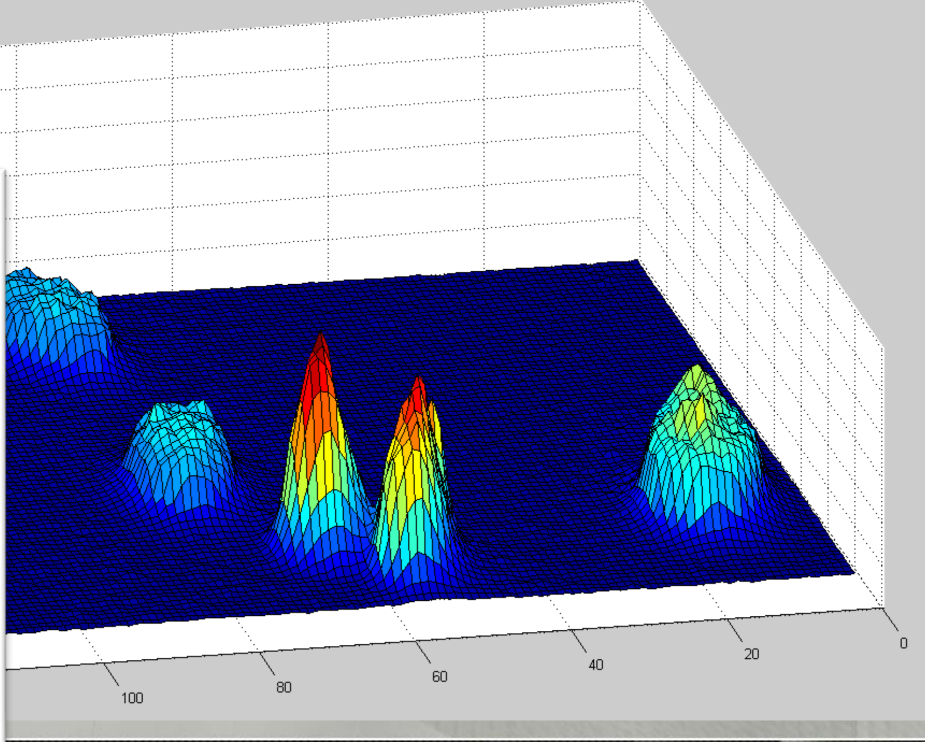
Image processing I. – cell detection



Banerjee, etal.; ***Science 2014.***

Image processing I. – cell detection





*Banerjee, etal.; **Science 2014.***

Image processing II. – compartments

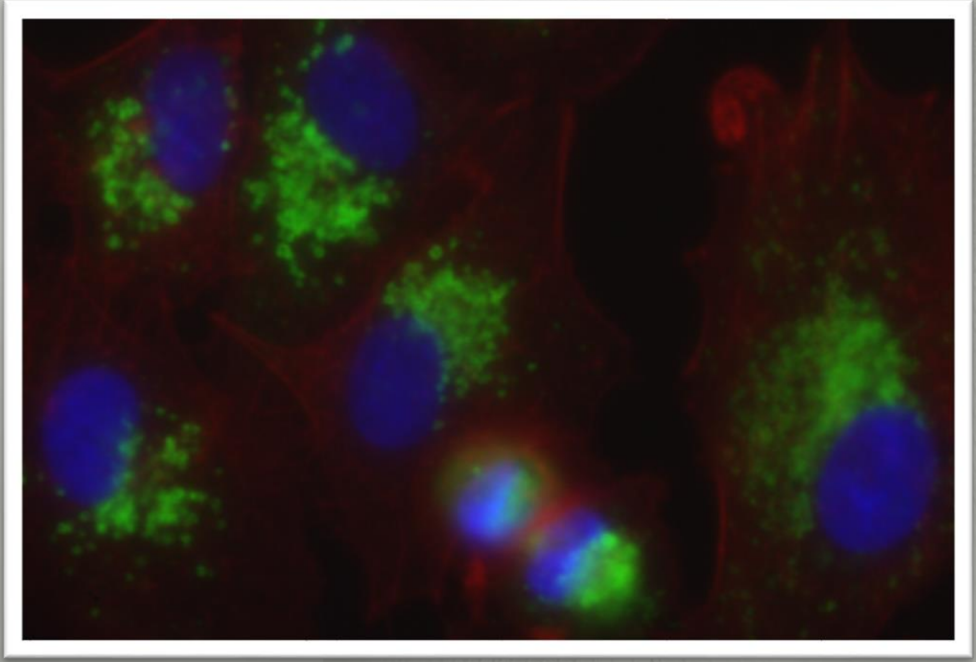


Image processing II. – compartments

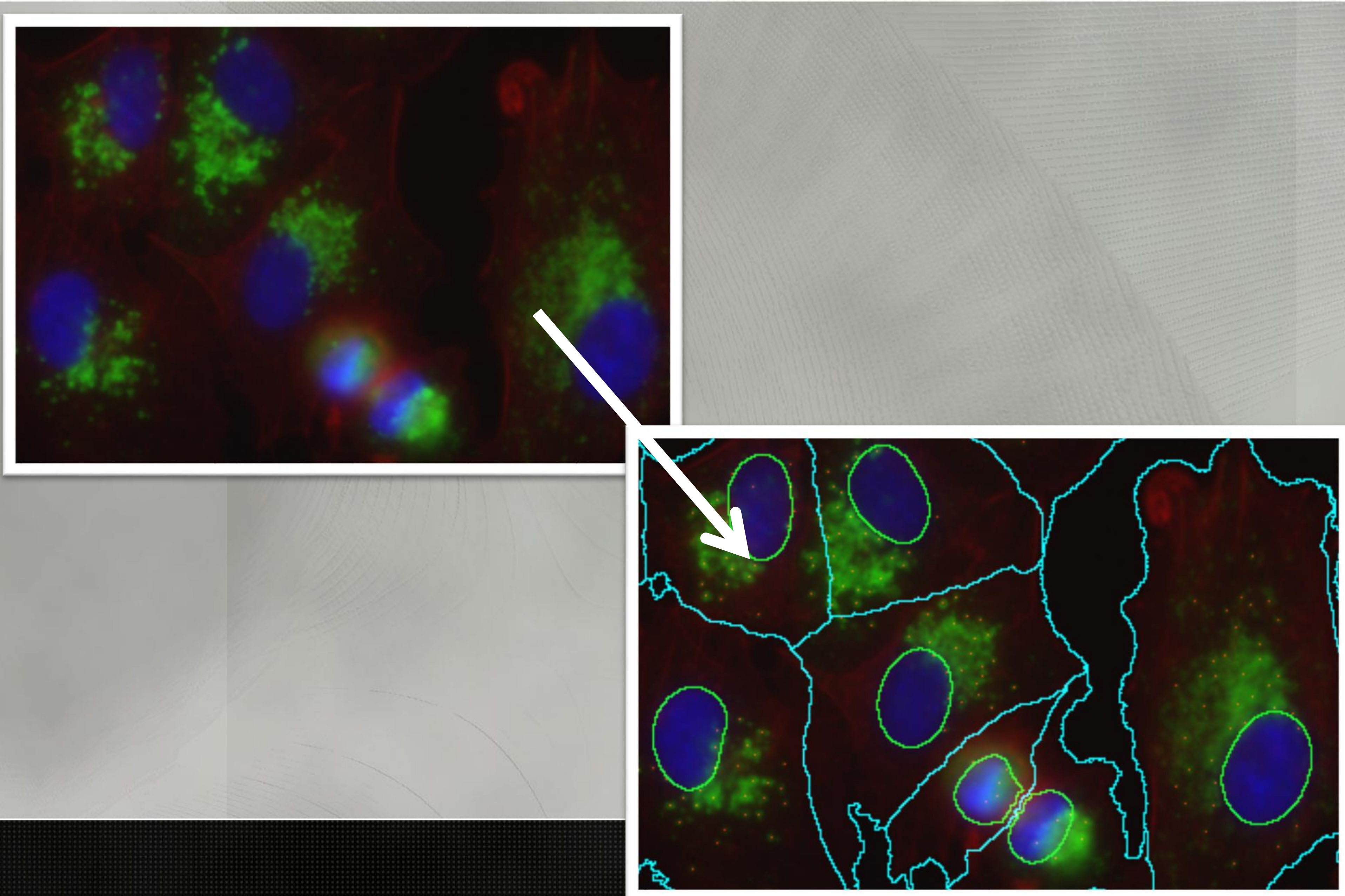


Image processing II. – compartments

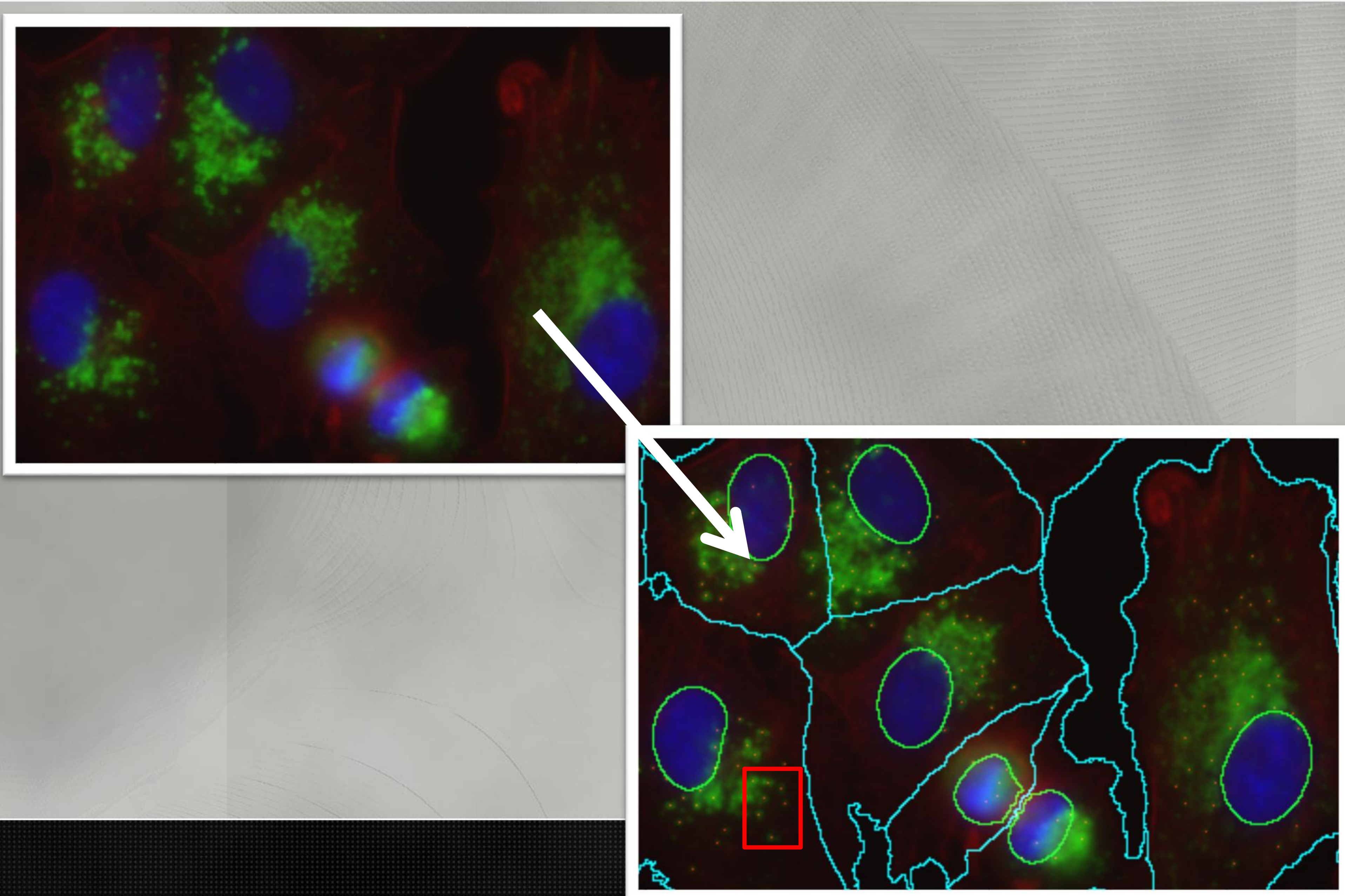


Image processing II. – compartments

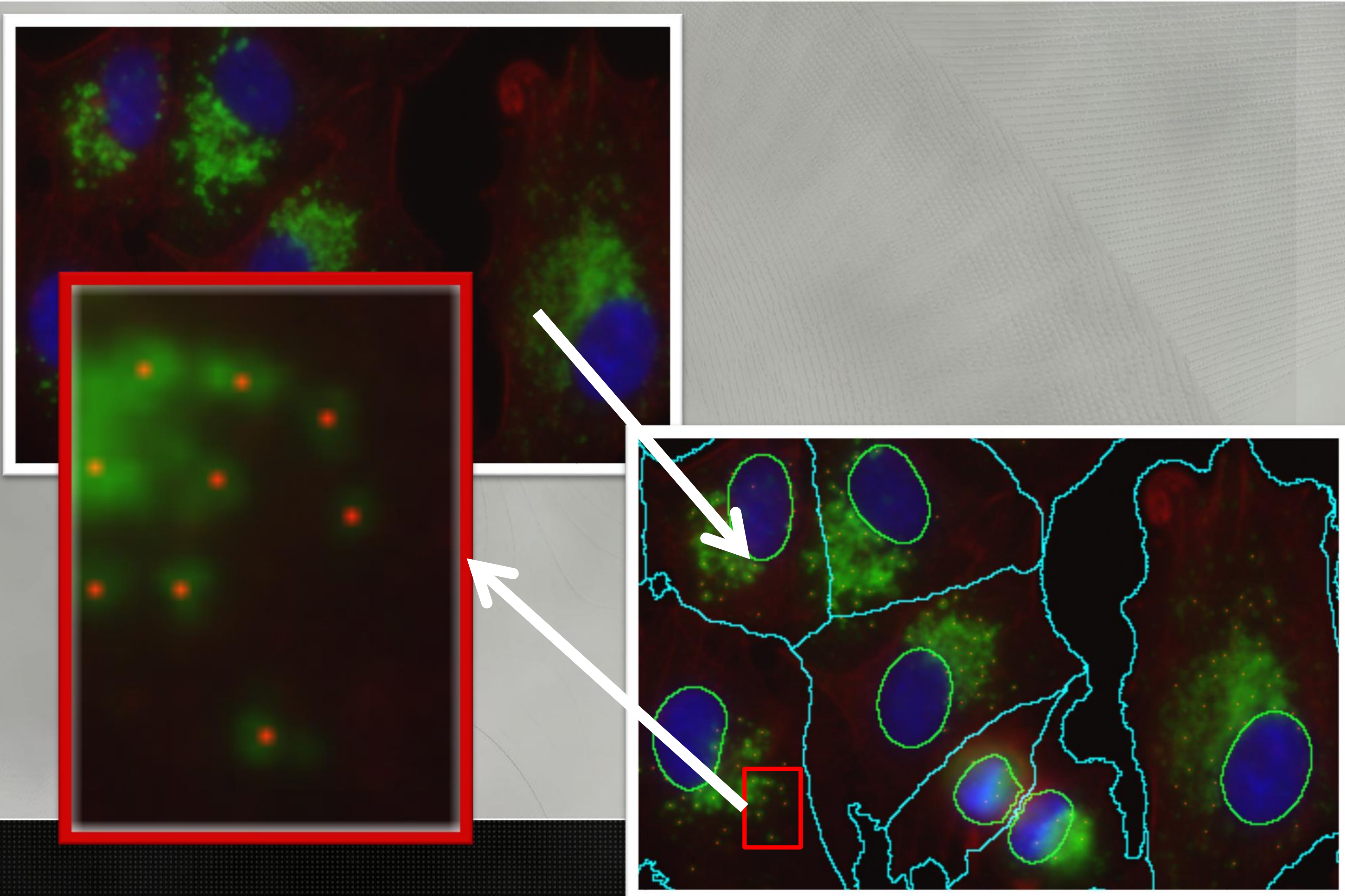


Image processing II. – compartments

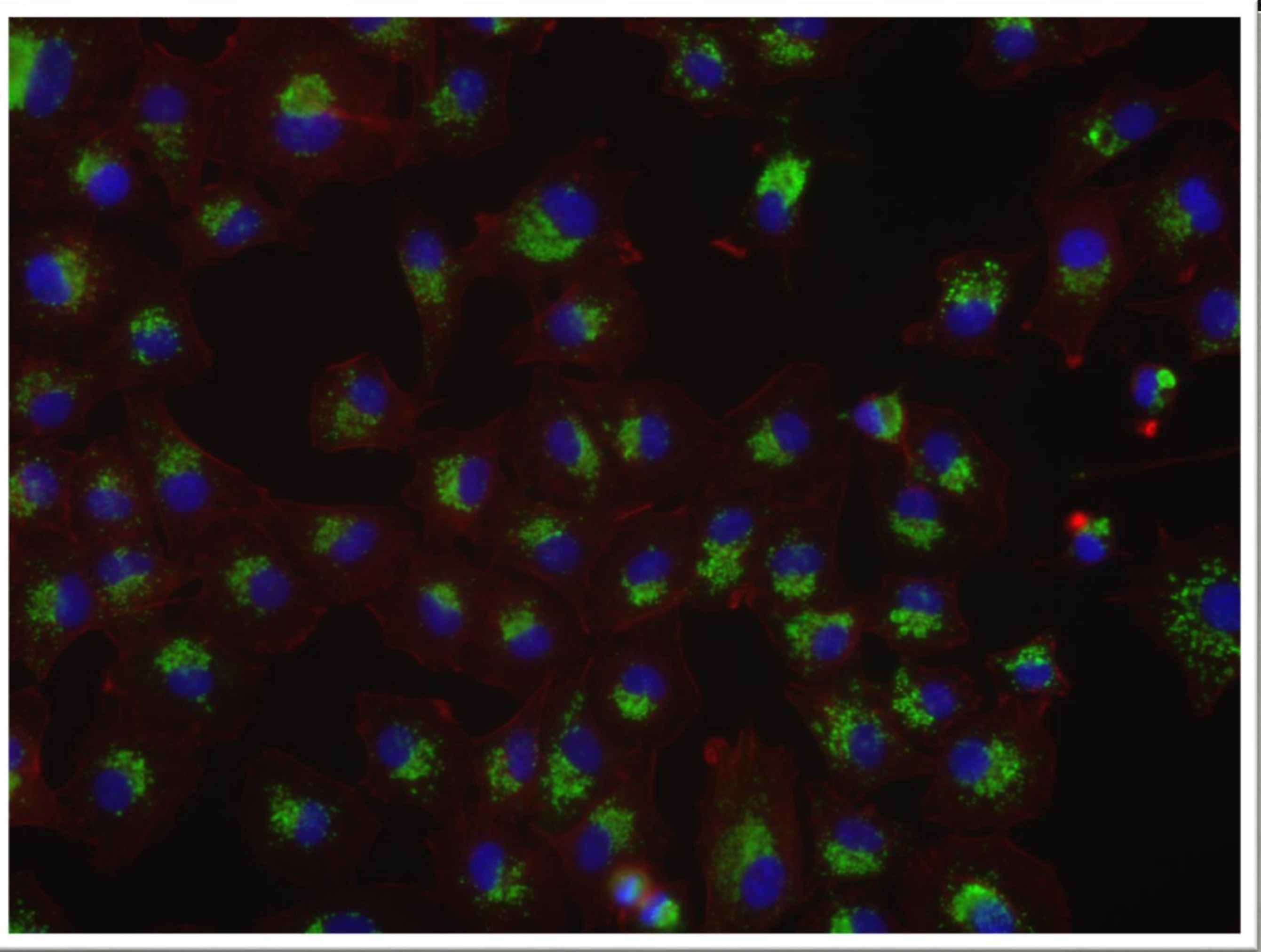


Image processing II. – compartments

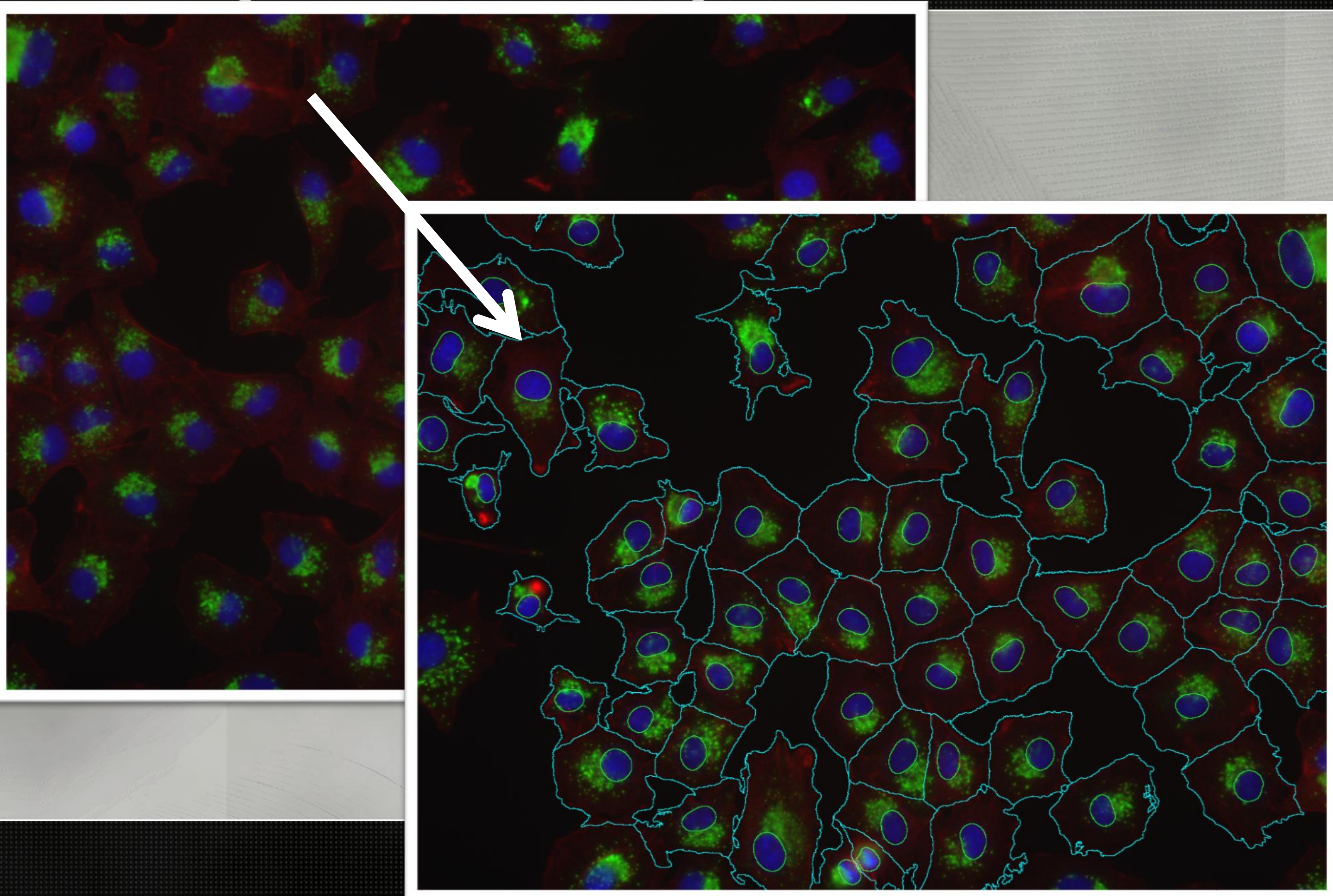
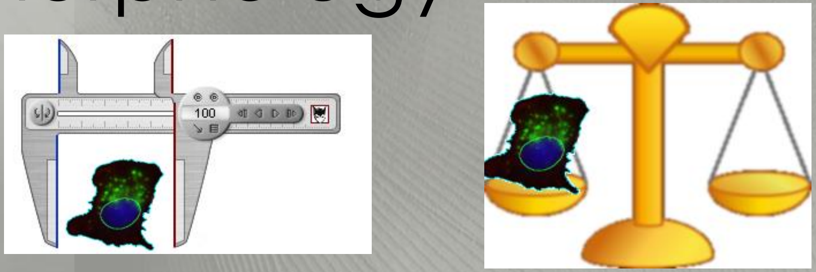
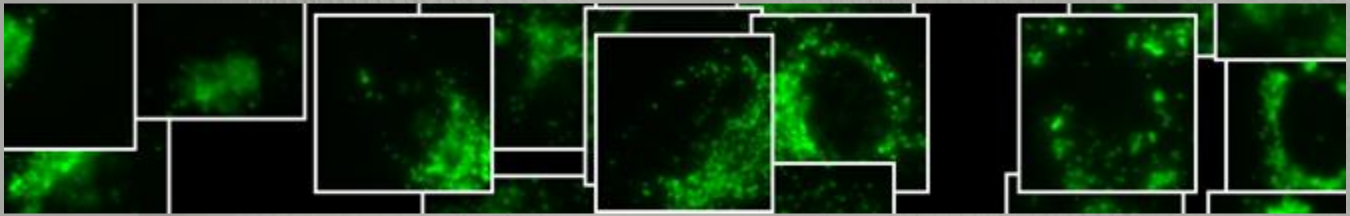


Image processing III. – feature extraction

- Morphology



- Intensity-based



- Texture

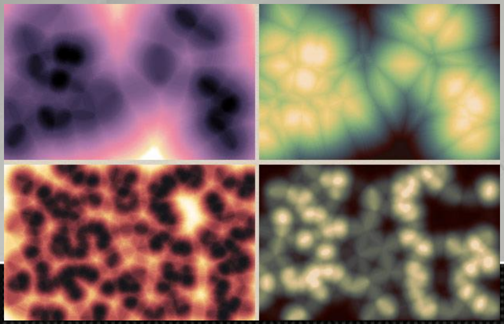


Image processing - Summary

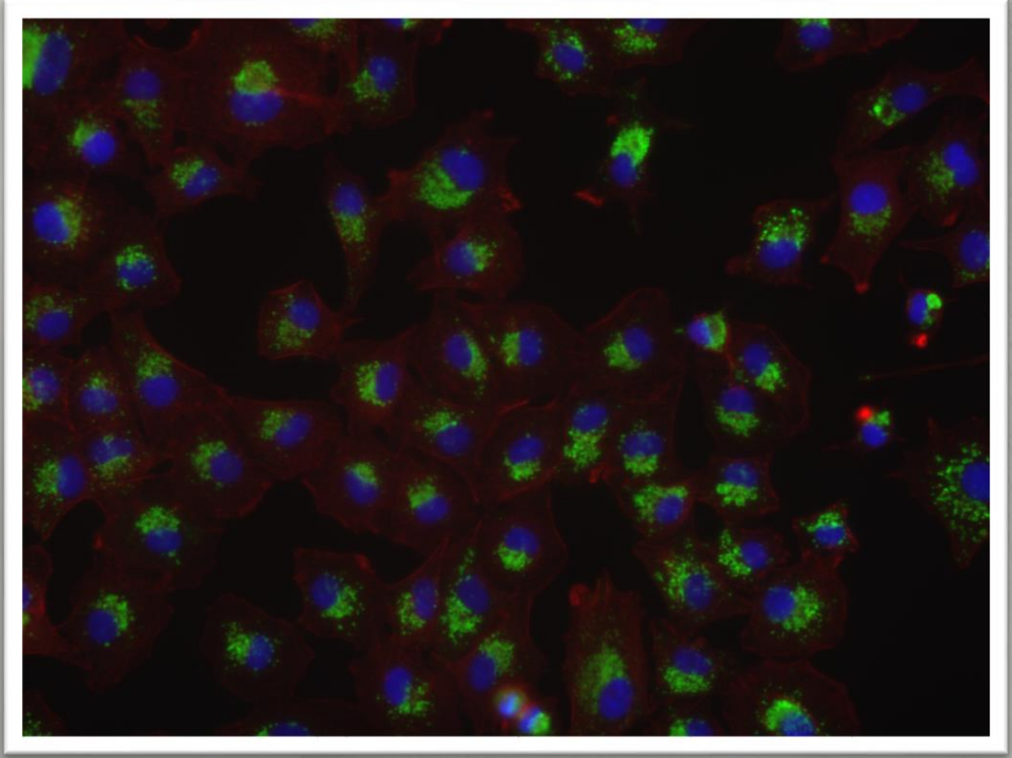
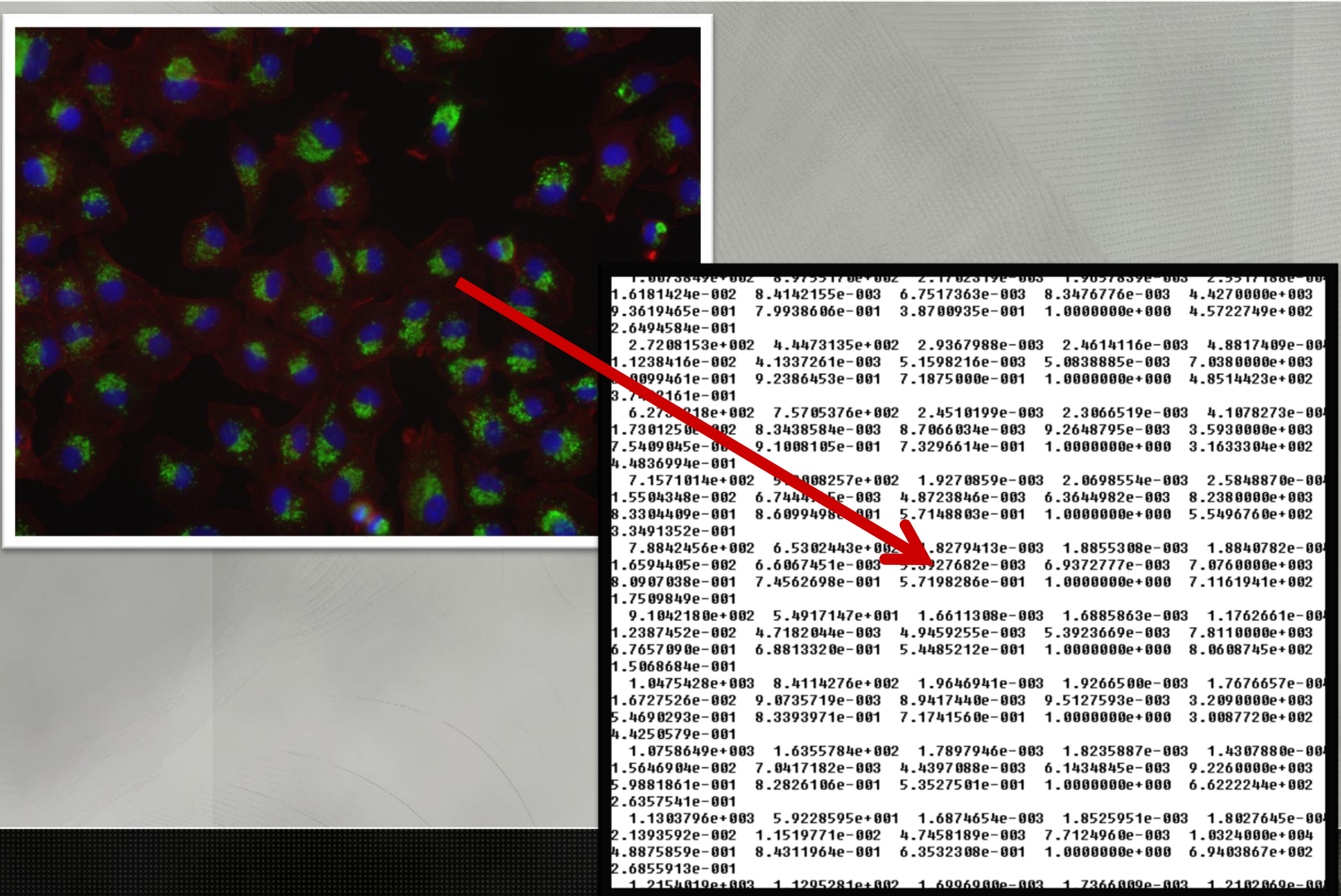


Image processing - Summary



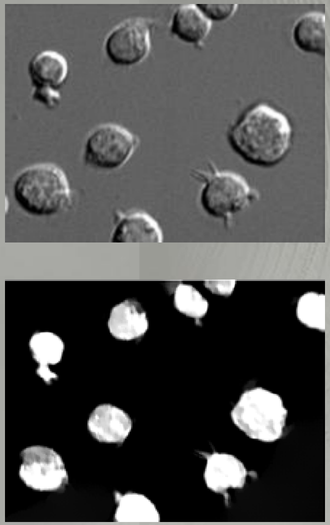
```
1.6181424e-002  8.4142155e-003  6.7517363e-003  8.3476776e-003  4.4270000e+003
9.3619465e-001  7.9938606e-001  3.8700935e-001  1.0000000e+000  4.5722749e+002
2.6494584e-001
  2.7208153e+002  4.4473135e+002  2.9367988e-003  2.4614116e-003  4.8817409e-00
1.1238416e-002  4.1337261e-003  5.1598216e-003  5.0838885e-003  7.0380000e+003
[illegible]0099461e-001  9.2386453e-001  7.1875000e-001  1.0000000e+000  4.8514423e+002
3.7[illegible]2161e-001
  6.27[illegible]218e+002  7.5705376e+002  2.4510199e-003  2.3066519e-003  4.1078273e-00
1.730125[illegible]02  8.3438584e-003  8.7066034e-003  9.2648795e-003  3.5930000e+003
7.5409045e-0[illegible]  9.1008105e-001  7.3296614e-001  1.0000000e+000  3.1633304e+002
4.4836994e-001
  7.1571014e+002  [illegible]008257e+002  1.9270859e-003  2.0698554e-003  2.5848870e-00
1.5504348e-002  6.744[illegible]5e-003  4.8723846e-003  6.3644982e-003  8.2380000e+003
8.3304409e-001  8.609949[illegible]001  5.7148803e-001  1.0000000e+000  5.5496760e+002
3.3491352e-001
  7.8842456e+002  6.5302443e+00[illegible].8279413e-003  1.8855308e-003  1.8840782e-00
1.6594405e-002  6.6067451e-003  [illegible]27682e-003  6.9372777e-003  7.0760000e+003
8.0907038e-001  7.4562698e-001  5.7198286e-001  1.0000000e+000  7.1161941e+002
1.7509849e-001
  9.1042180e+002  5.4917147e+001  1.6611308e-003  1.6885863e-003  1.1762661e-00
1.2387452e-002  4.7182044e-003  4.9459255e-003  5.3923669e-003  7.8110000e+003
6.7657090e-001  6.8813320e-001  5.4485212e-001  1.0000000e+000  8.0608745e+002
1.5068684e-001
  1.0475428e+003  8.4114276e+002  1.9646941e-003  1.9266500e-003  1.7676657e-00
1.6727526e-002  9.0735719e-003  8.9417440e-003  9.5127593e-003  3.2090000e+003
5.4690293e-001  8.3393971e-001  7.1741560e-001  1.0000000e+000  3.0087720e+002
4.4250579e-001
  1.0758649e+003  1.6355784e+002  1.7897946e-003  1.8235887e-003  1.4307880e-00
1.5646904e-002  7.0417182e-003  4.4397088e-003  6.1434845e-003  9.2260000e+003
5.9881861e-001  8.2826106e-001  5.3527501e-001  1.0000000e+000  6.6222244e+002
2.6357541e-001
  1.1303796e+003  5.9228595e+001  1.6874654e-003  1.8525951e-003  1.8027645e-00
2.1393592e-002  1.1519771e-002  4.7458189e-003  7.7124960e-003  1.0324000e+004
4.8875859e-001  8.4311964e-001  6.3532308e-001  1.0000000e+000  6.9403867e+002
2.6855913e-001
```

Challenges

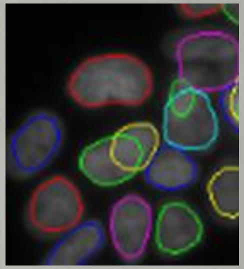
IMAGE SEGMENTATION

Label free image reconstruction

Complex object modeling

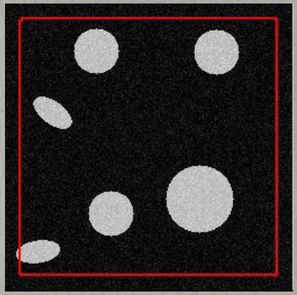


Koos, etal. ***DICTA 2015***
Koos, etal. ***Sci Rep. 2016.***
Koos, etal. ***SCIA 2017.***



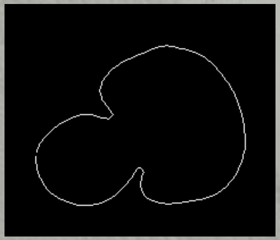
Overlapping

Horvath, etal. ***PattRec 2009***
Molnar, etal. ***Sci Rep. 2016.***
Brasko, etal. ***NatComm 2017.***



Modeling

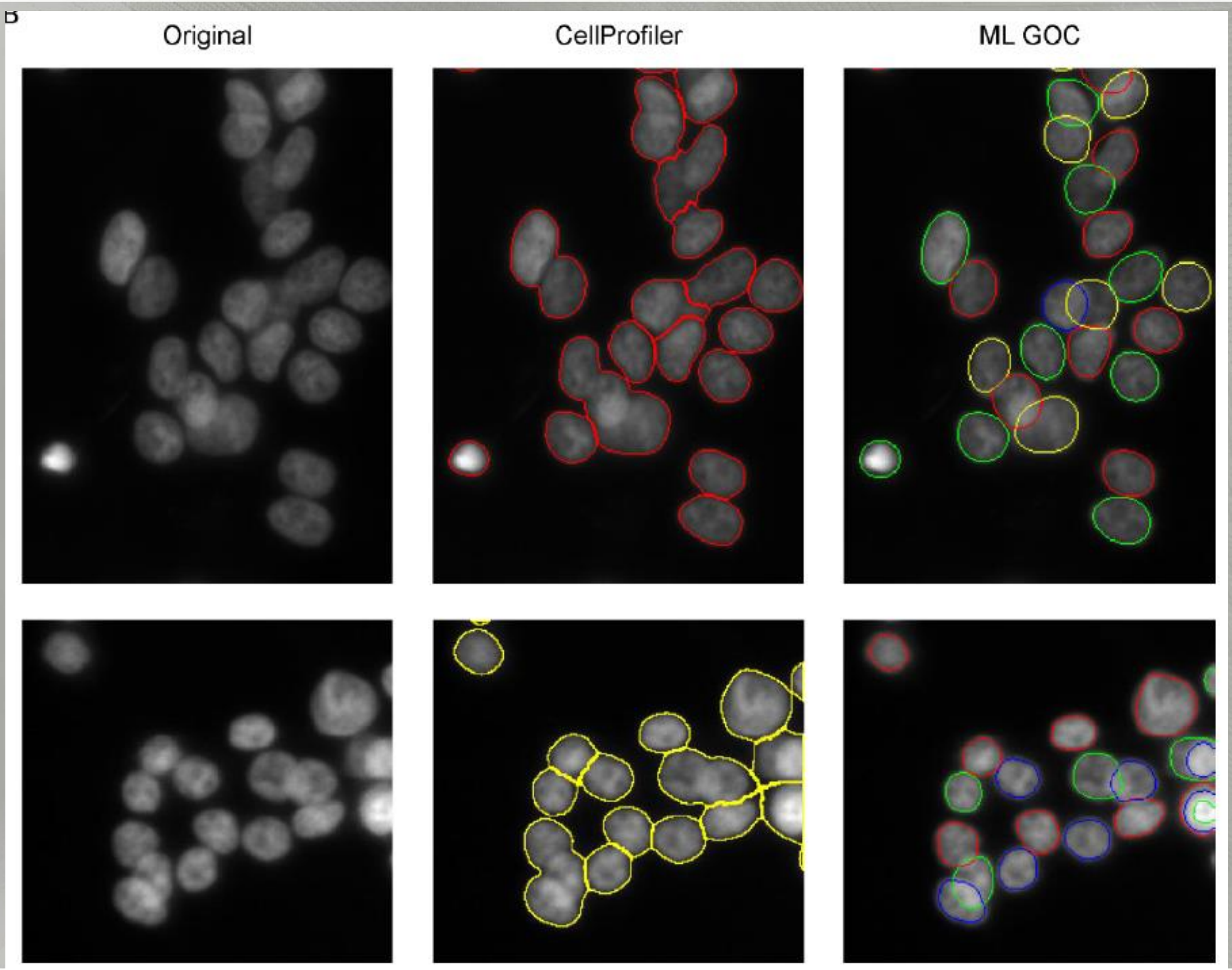
Molnar, etal. ***WAVC 2016.***
Molnar, etal. ***DICTA. 2017.***
Pfisterer, etal. ***NatComm 2017.***



Splitting

Molnar, etal. ***SCIA 2016.***

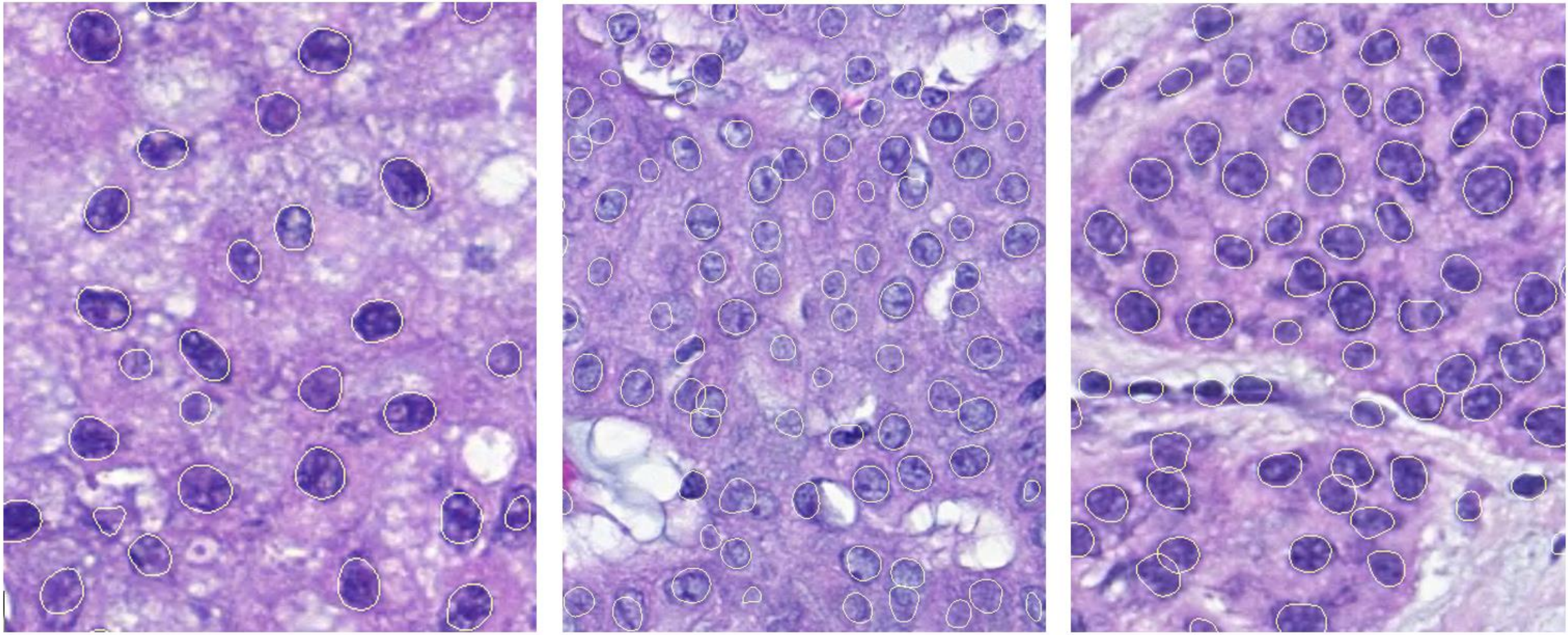
Image segmentation with priors



Molnar, etal. ***Sci. Rep. 2016.;*** Brasko, etal. ***Nature Comm. 2018.***

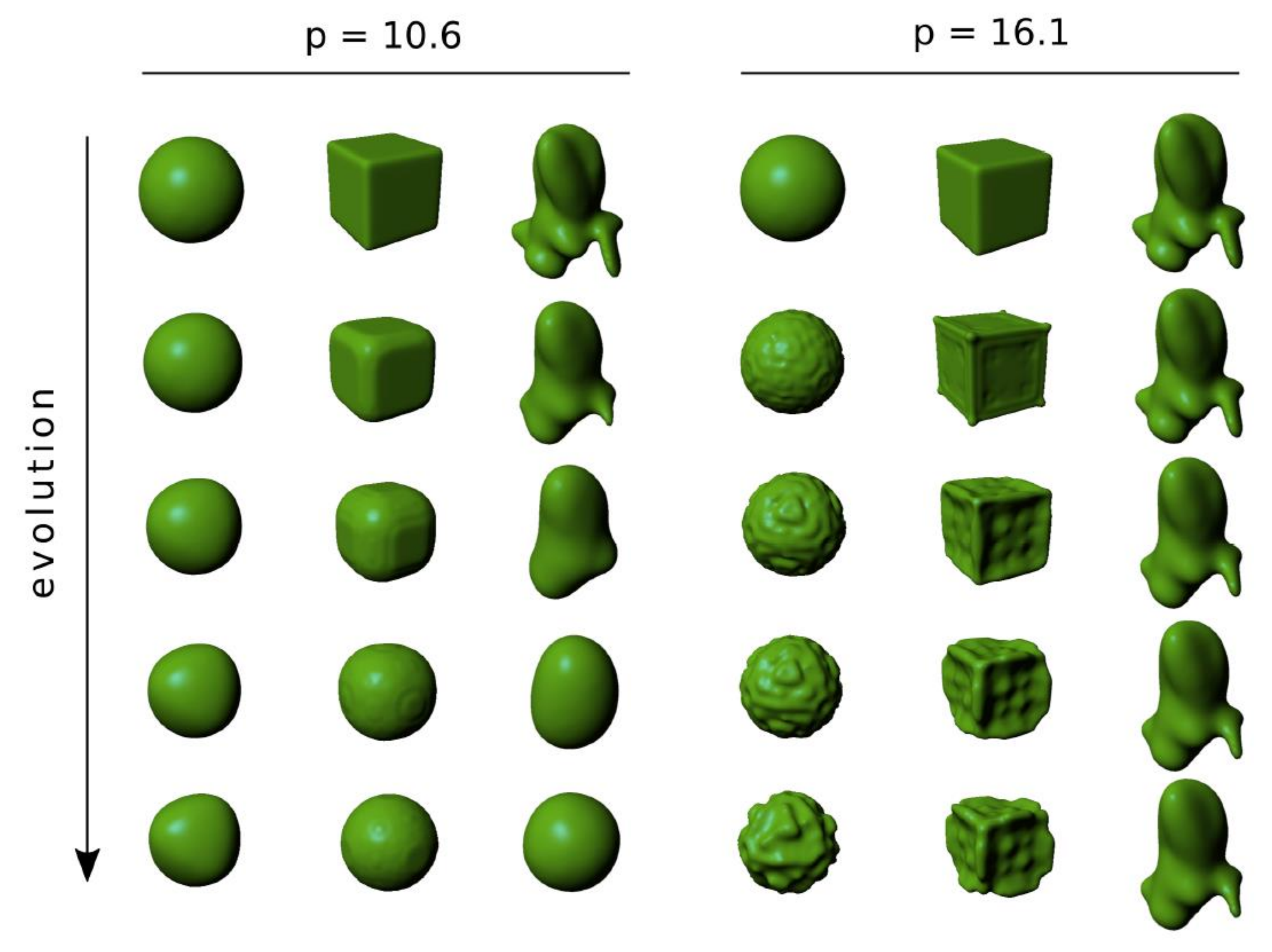
Single-cell tissue analysis

Hundreds of different breast cancer TMA spots



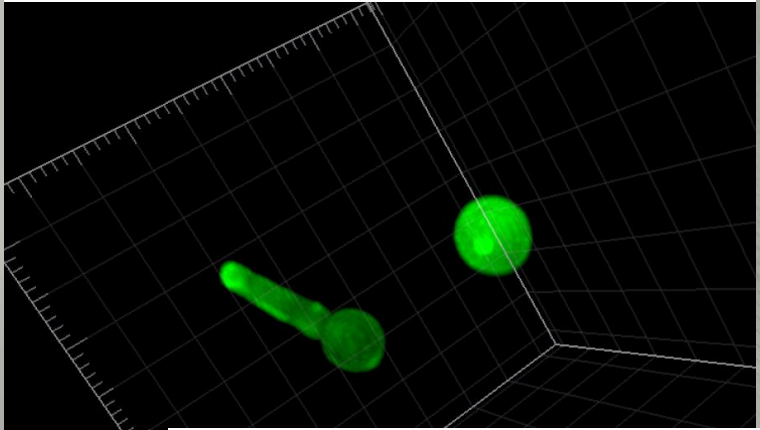
Data courtesy of Prof. J. Lundin

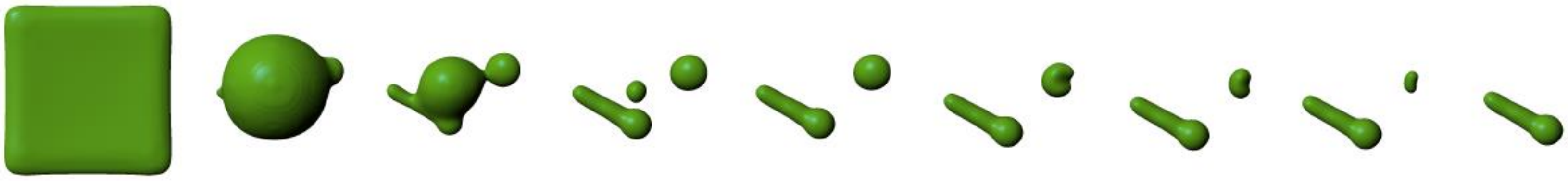
3D extension - GPU



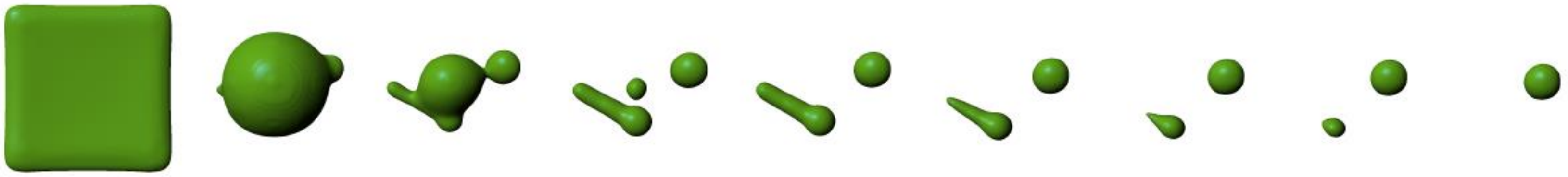
Molnar etal. ***DICTA 2017***

3D active surfaces





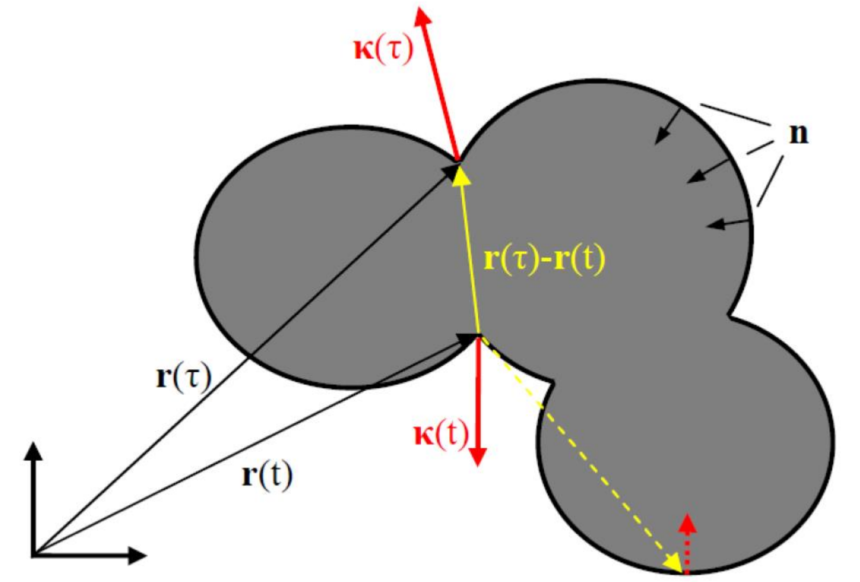
(a) Surface evolution while segmenting out the pseudohyphae form.

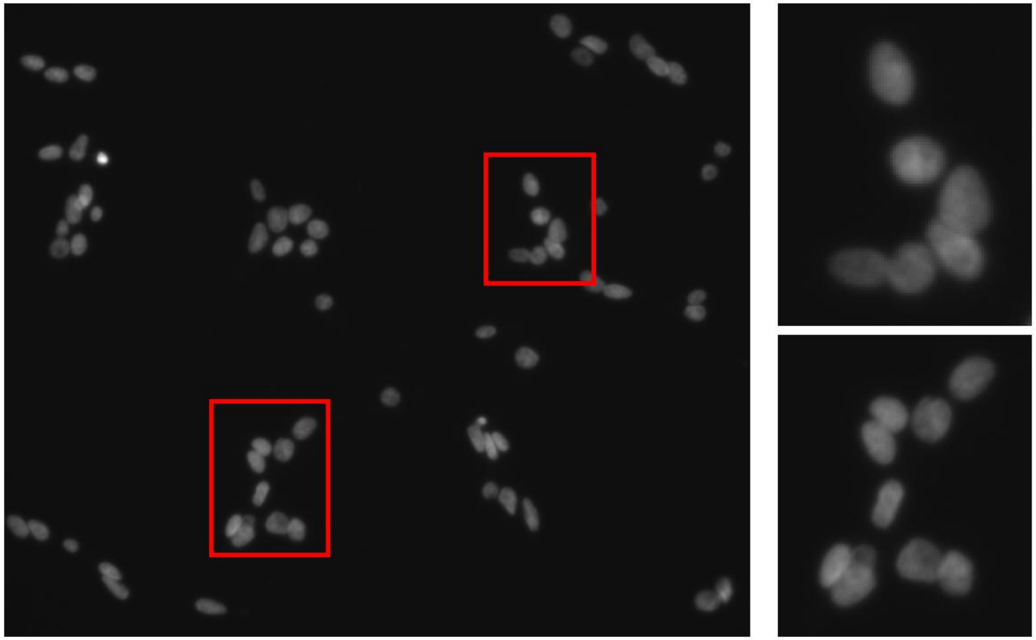


(b) Surface evolution while segmenting out the yeast cell in normal form.

Molnar etal. ***DICTA 2017***

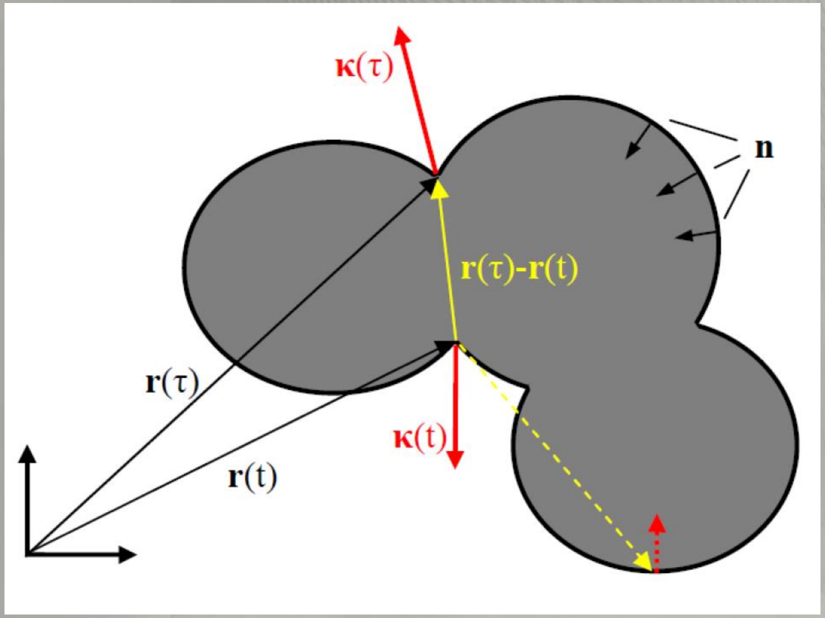
New active contour model

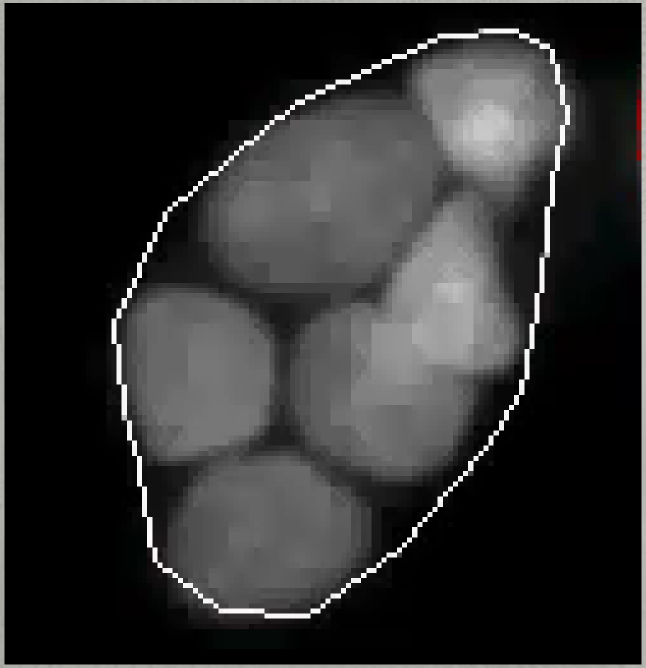




Molnar etal. ***IEEE ISVC 2016.***

New active contour model





Molnar etal. ***IEEE ISVC 2016.***

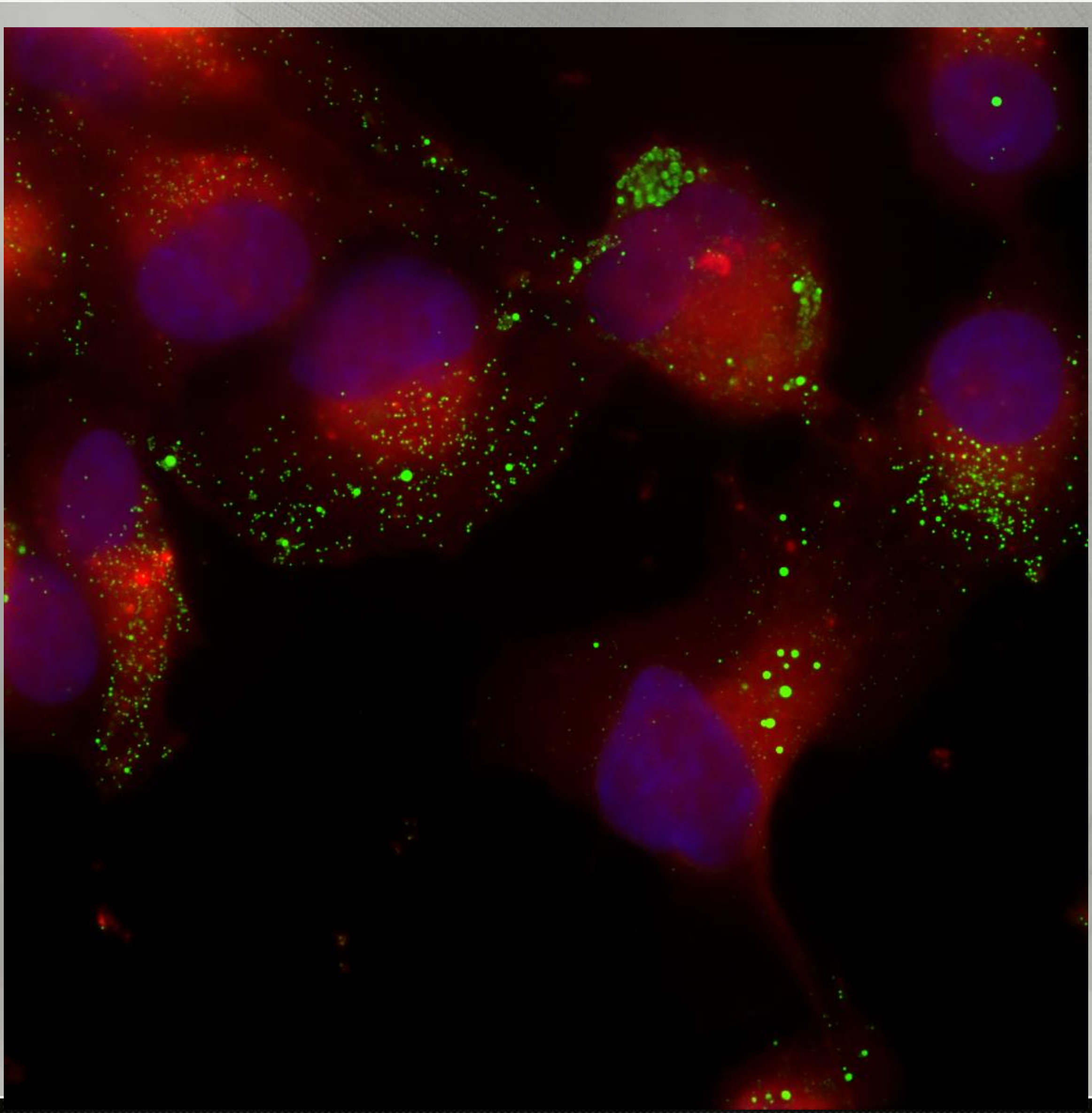
Prof. Elina Ikonen, Simon Pfisterer

LIPID DROPLET ANALYSIS

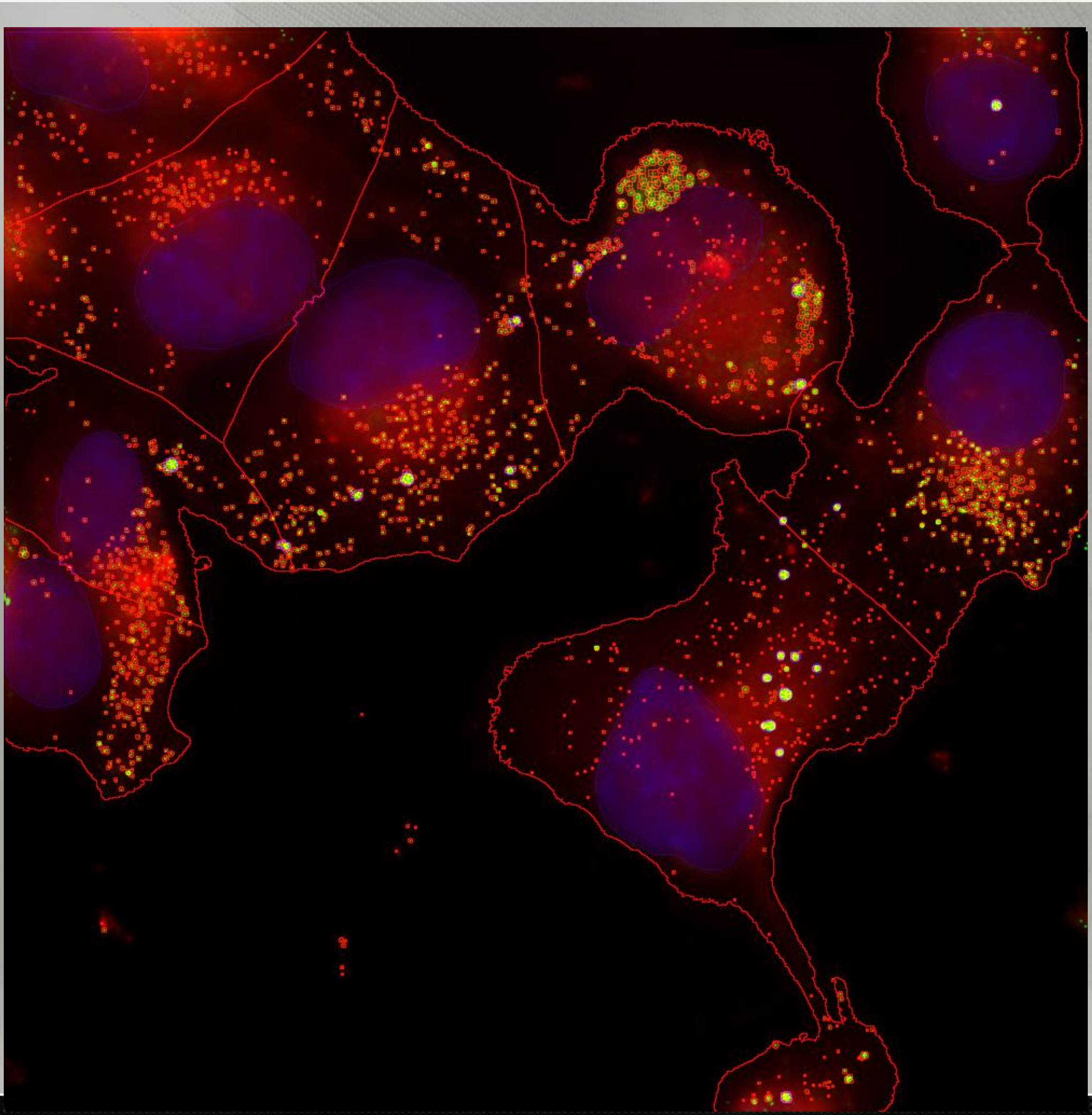
Pfisterer, S., ***etal;*** (2017) **Nature Communications**

Size selective droplet analysis

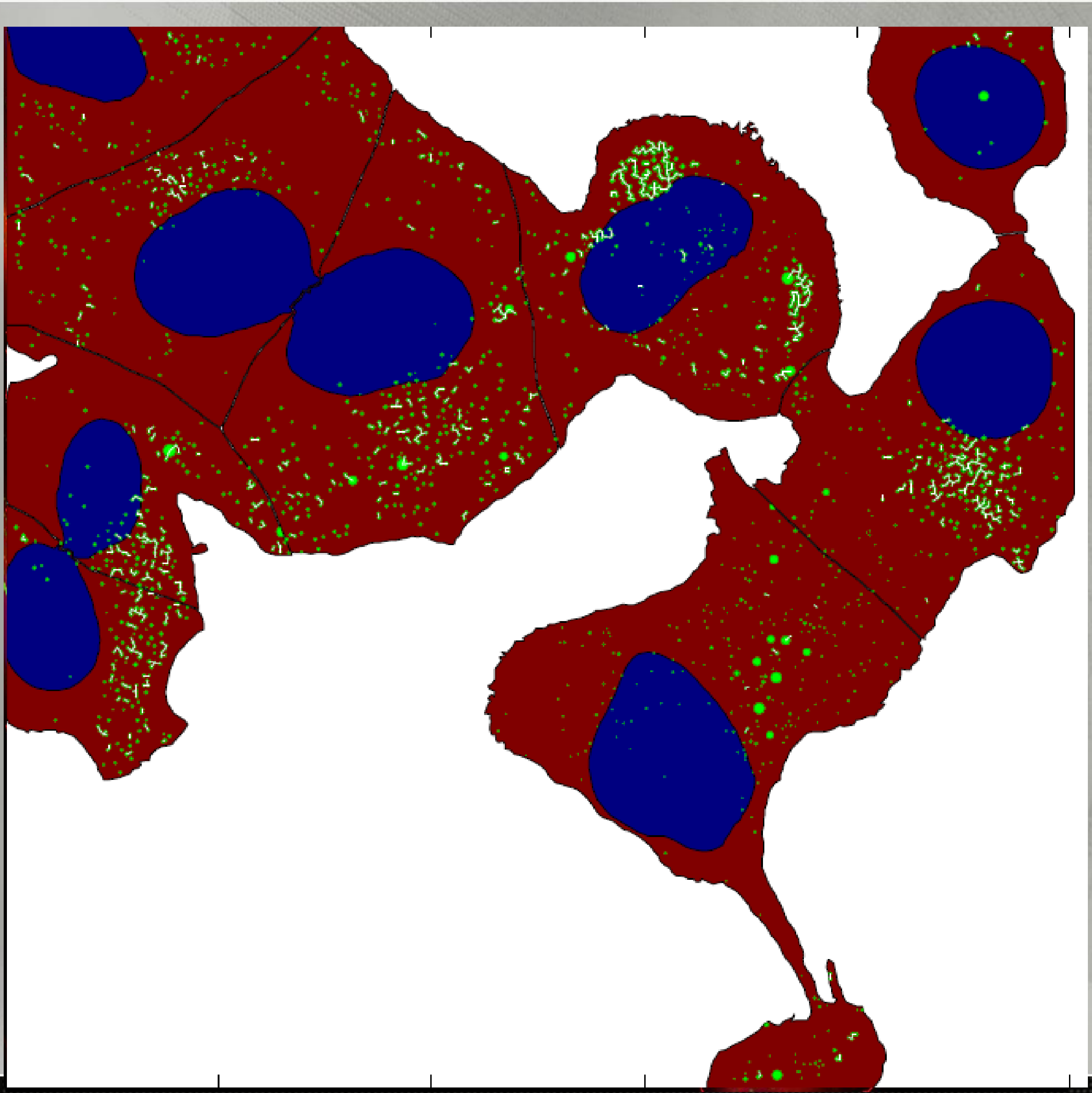
Size selective droplet analysis



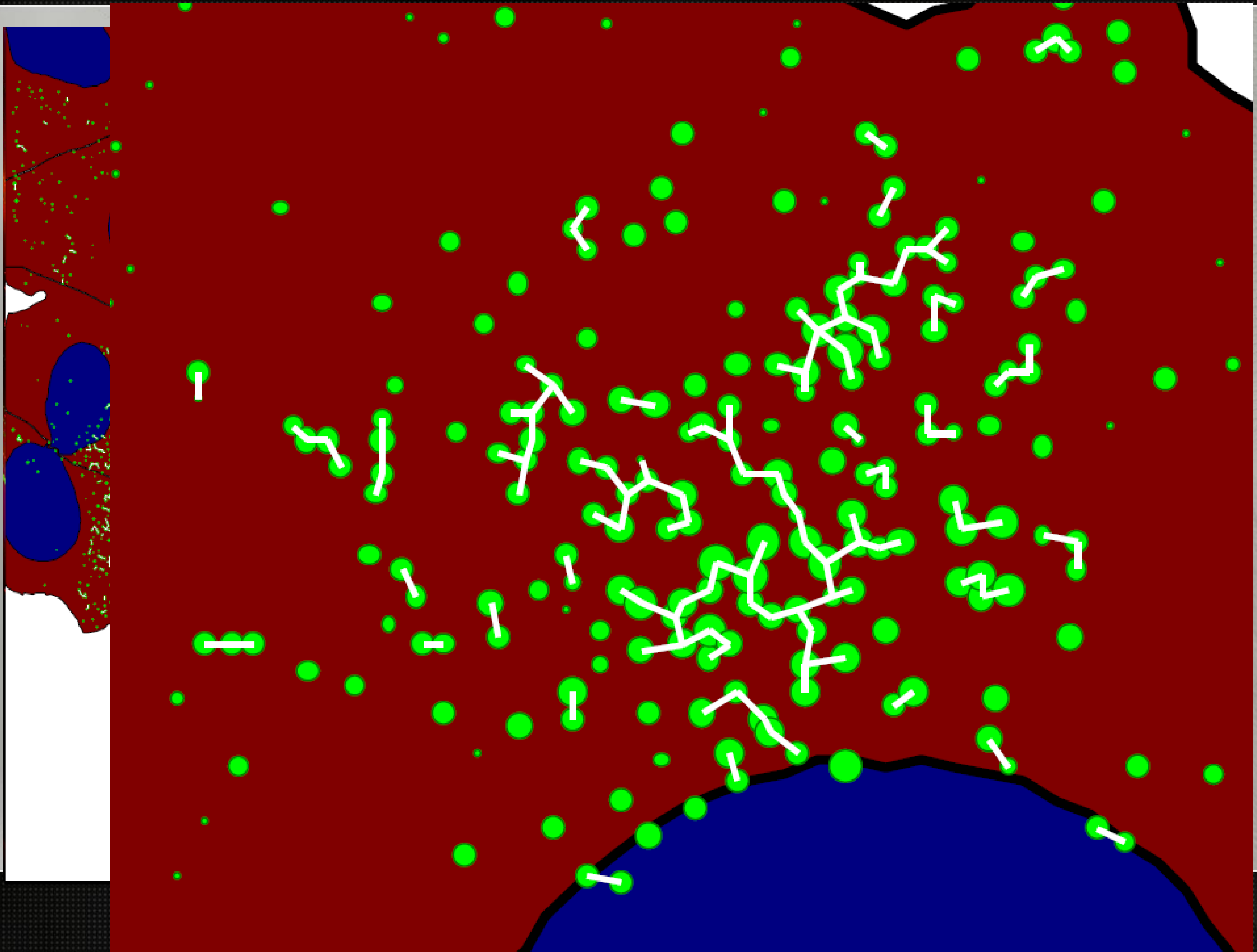
Size selective droplet analysis



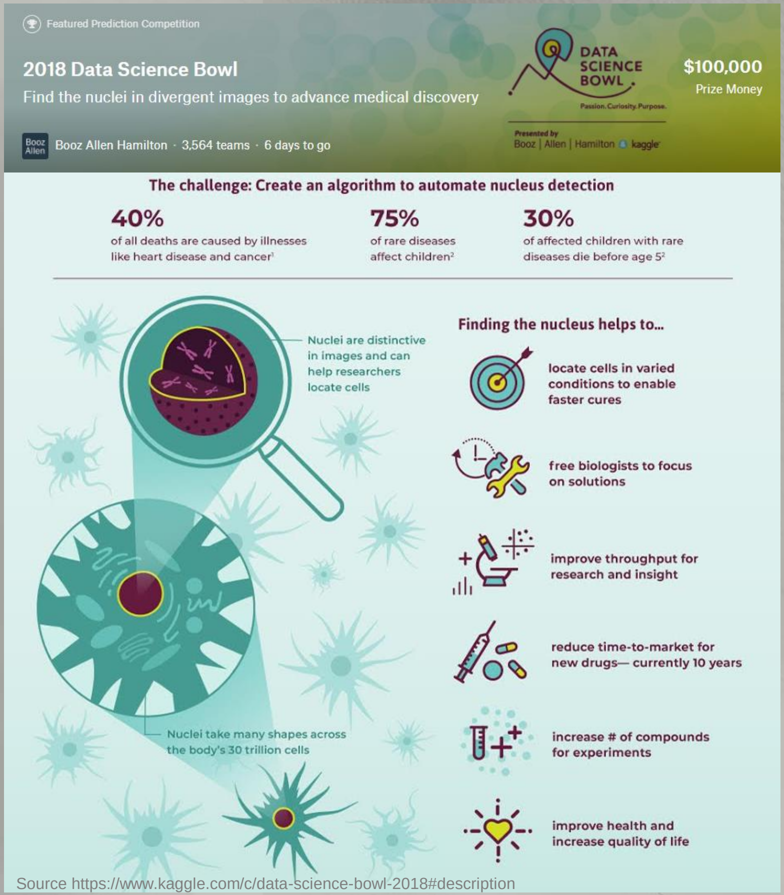
Size selective droplet analysis



Size selective droplet analysis



2018 Kaggle Data Science Bowl: find nuclei in diverse images


Featured Prediction Competition

2018 Data Science Bowl

Find the nuclei in divergent images to advance medical discovery

Booz Allen Hamilton · 3,564 teams · 6 days to go

DATA SCIENCE BOWL

Passion. Curiosity. Purpose.

Presented by Booz | Allen | Hamilton kaggle

$100,000

Prize Money

The challenge: Create an algorithm to automate nucleus detection

40% of all deaths are caused by illnesses like heart disease and cancer[1]

75% of rare diseases affect children[2]

30% of affected children with rare diseases die before age 5[2]

Nuclei are distinctive in images and can help researchers locate cells

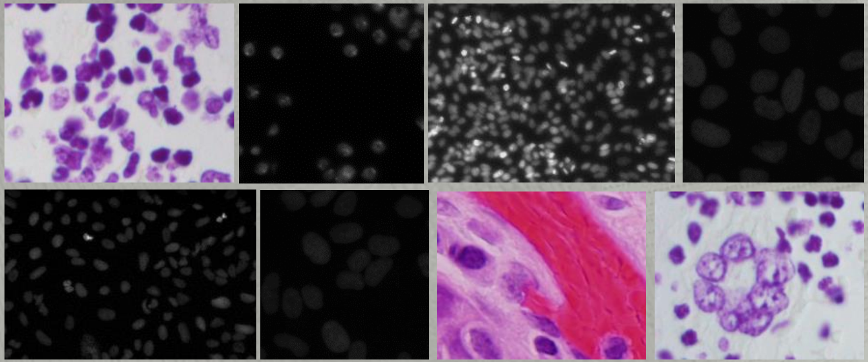
Nuclei take many shapes across the body's 30 trillion cells

Finding the nucleus helps to...

- locate cells in varied conditions to enable faster cures
- free biologists to focus on solutions
- improve throughput for research and insight
- reduce time-to-market for new drugs— currently 10 years
- increase # of compounds for experiments
- improve health and increase quality of life

Source https://www.kaggle.com/c/data-science-bowl-2018#description

Training data



Test data

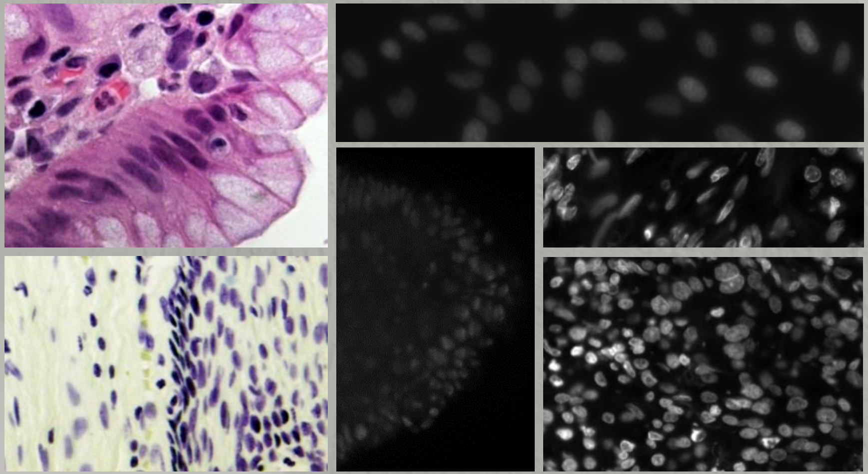
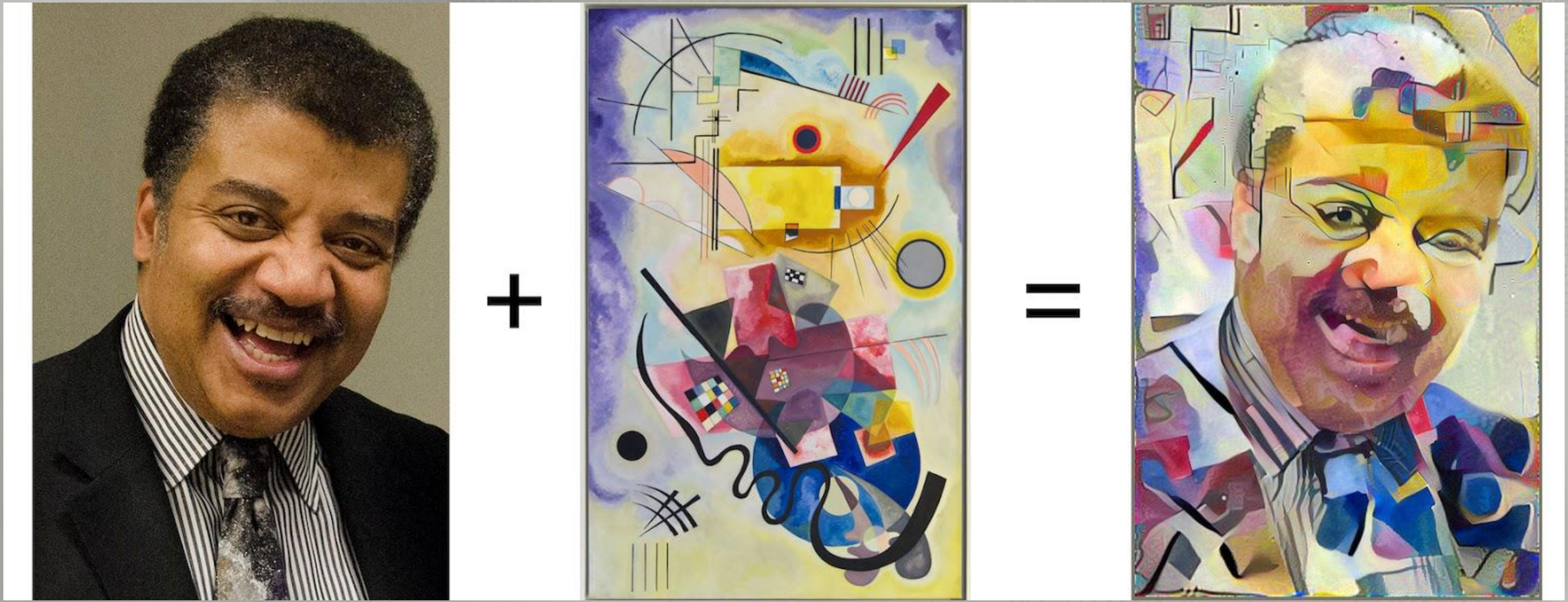
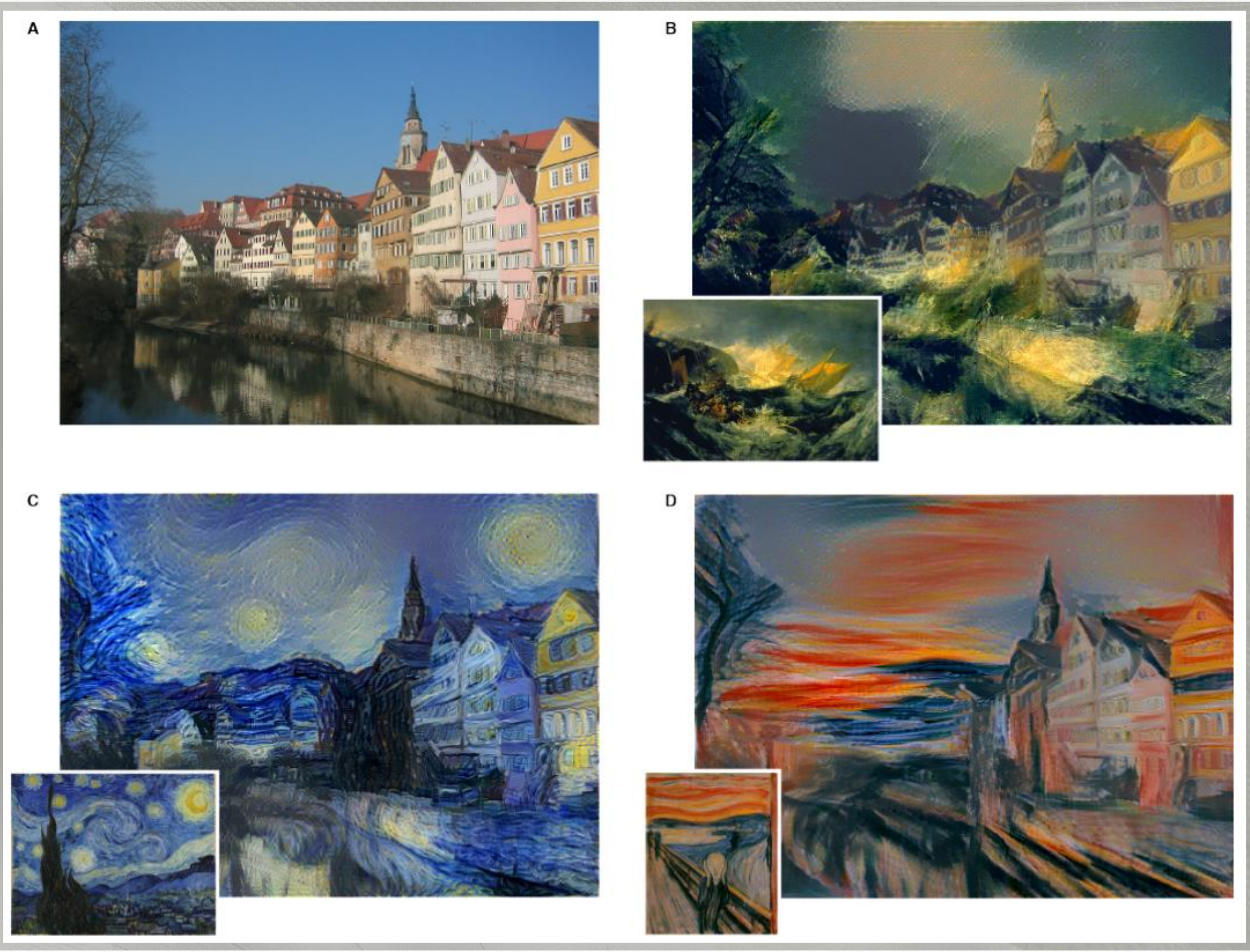


Image style transfer learning



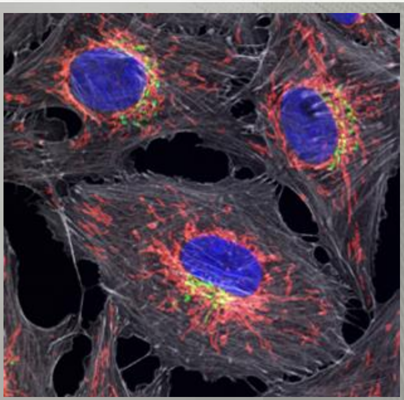
***Gatys etal:** A Neural Algorithm of Artistic Style; 2015.*

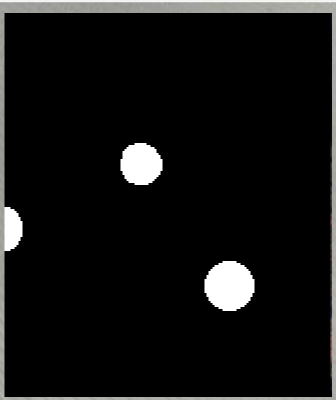
Image style transfer learning



***Gatys etal:** A Neural Algorithm of Artistic Style; 2015.*

Cell images

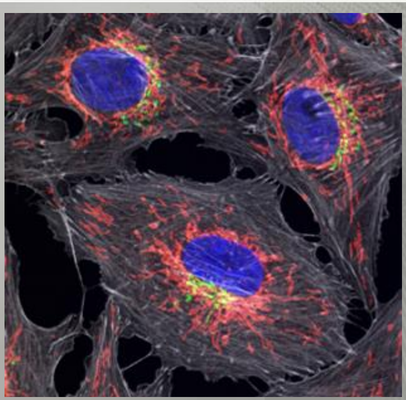


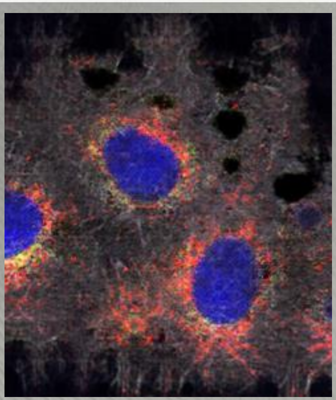


Original

Fake

Cell images

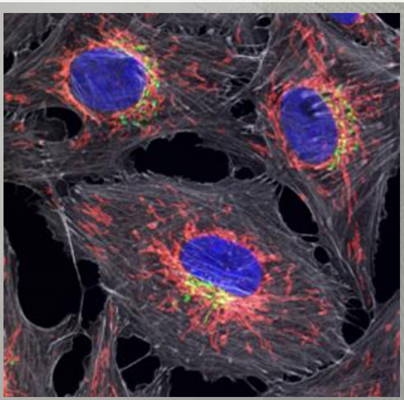




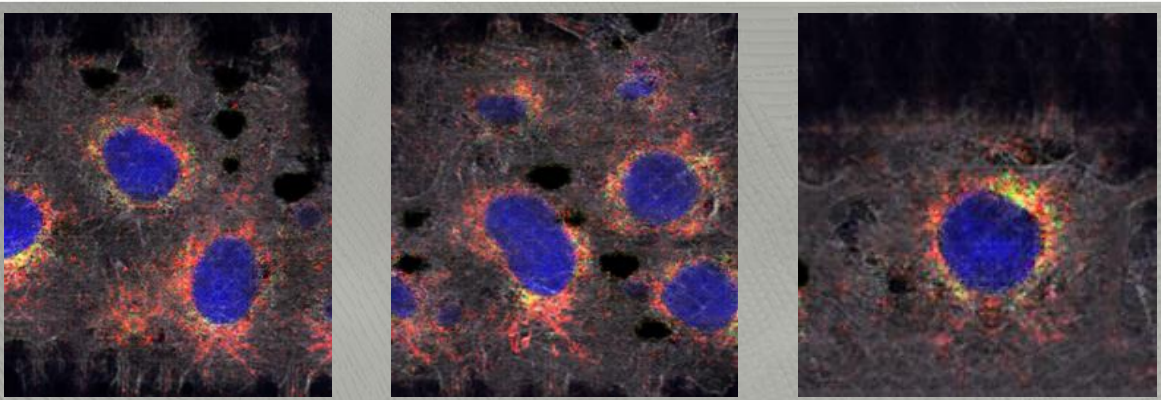
Original

Fake

Cell images

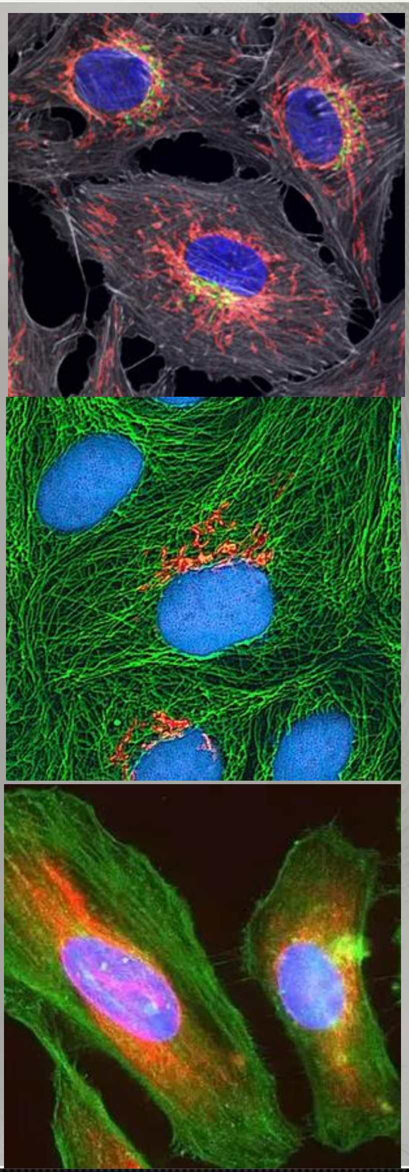


Original

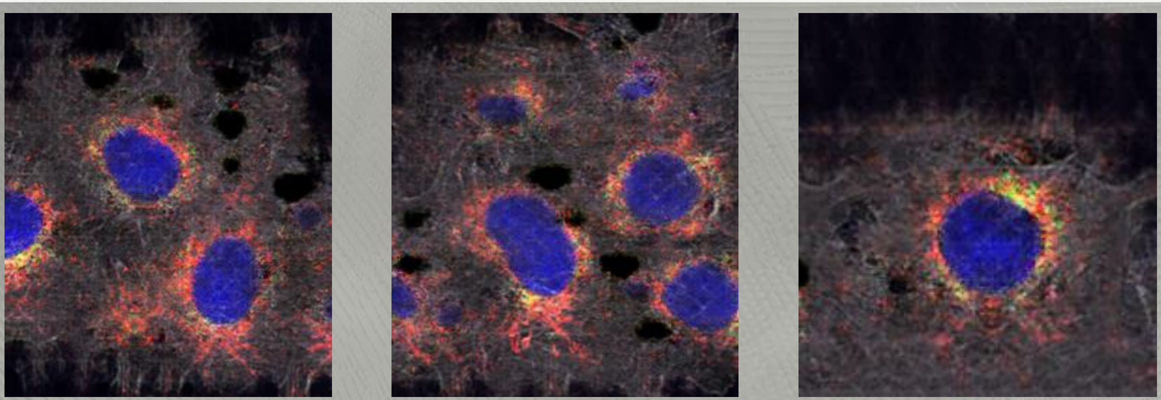


Fake

Cell images

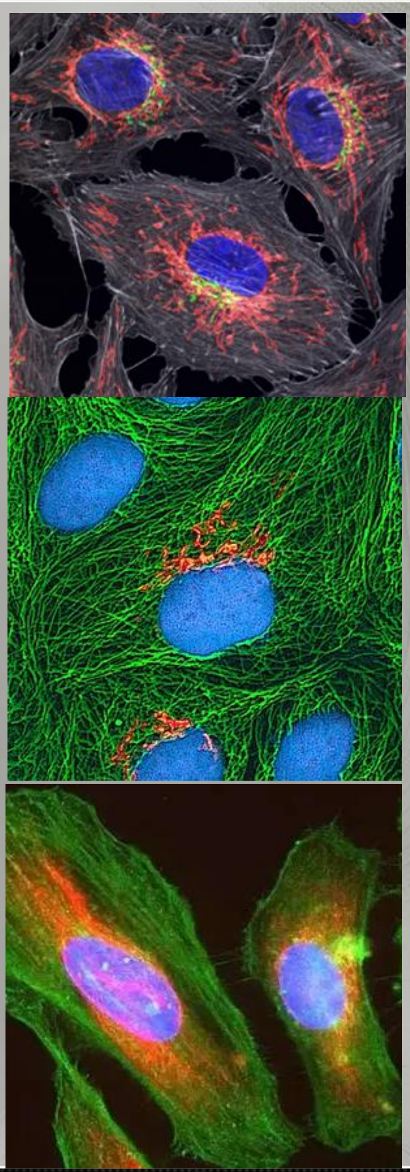


Original

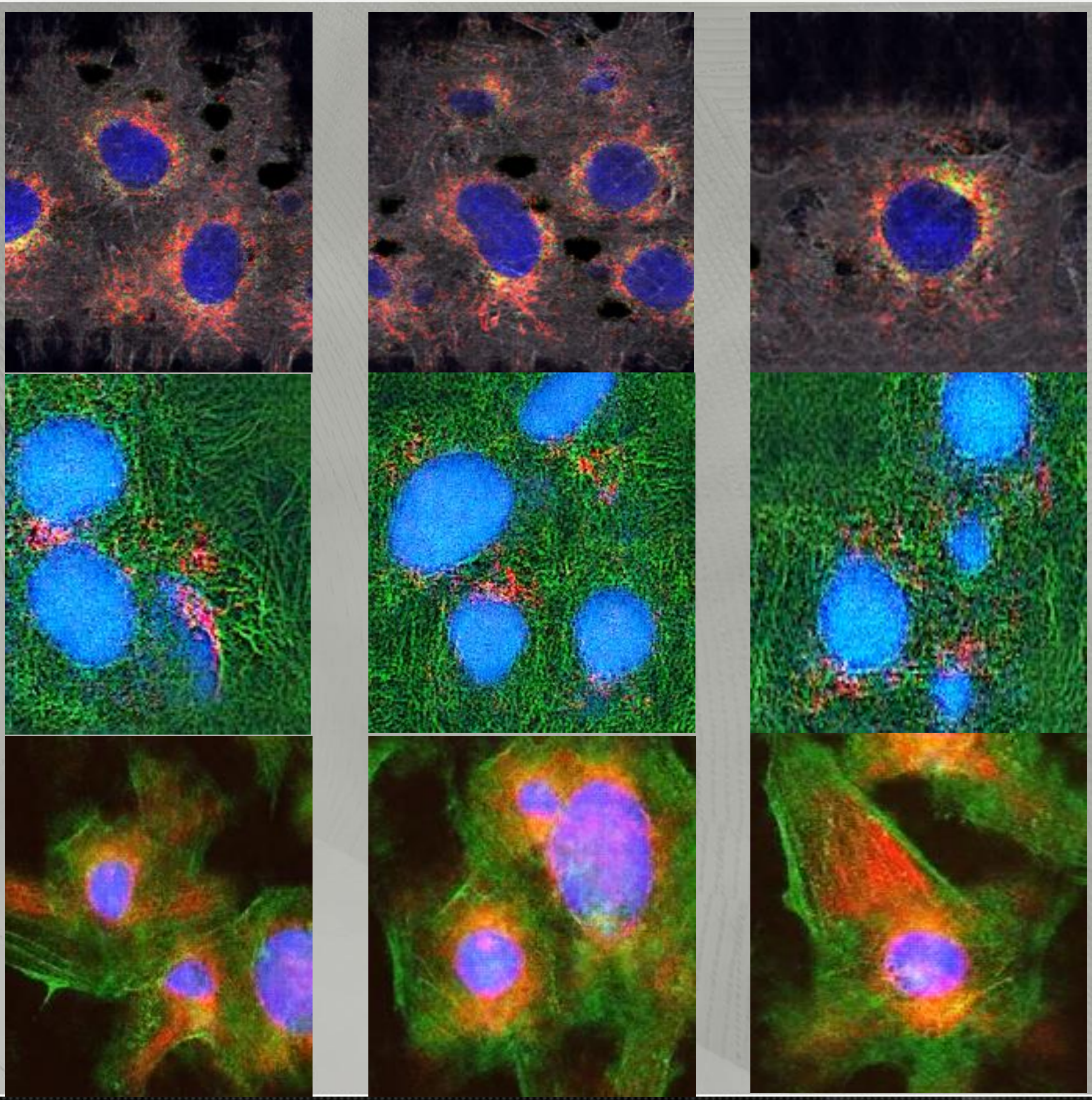


Fake

Cell images

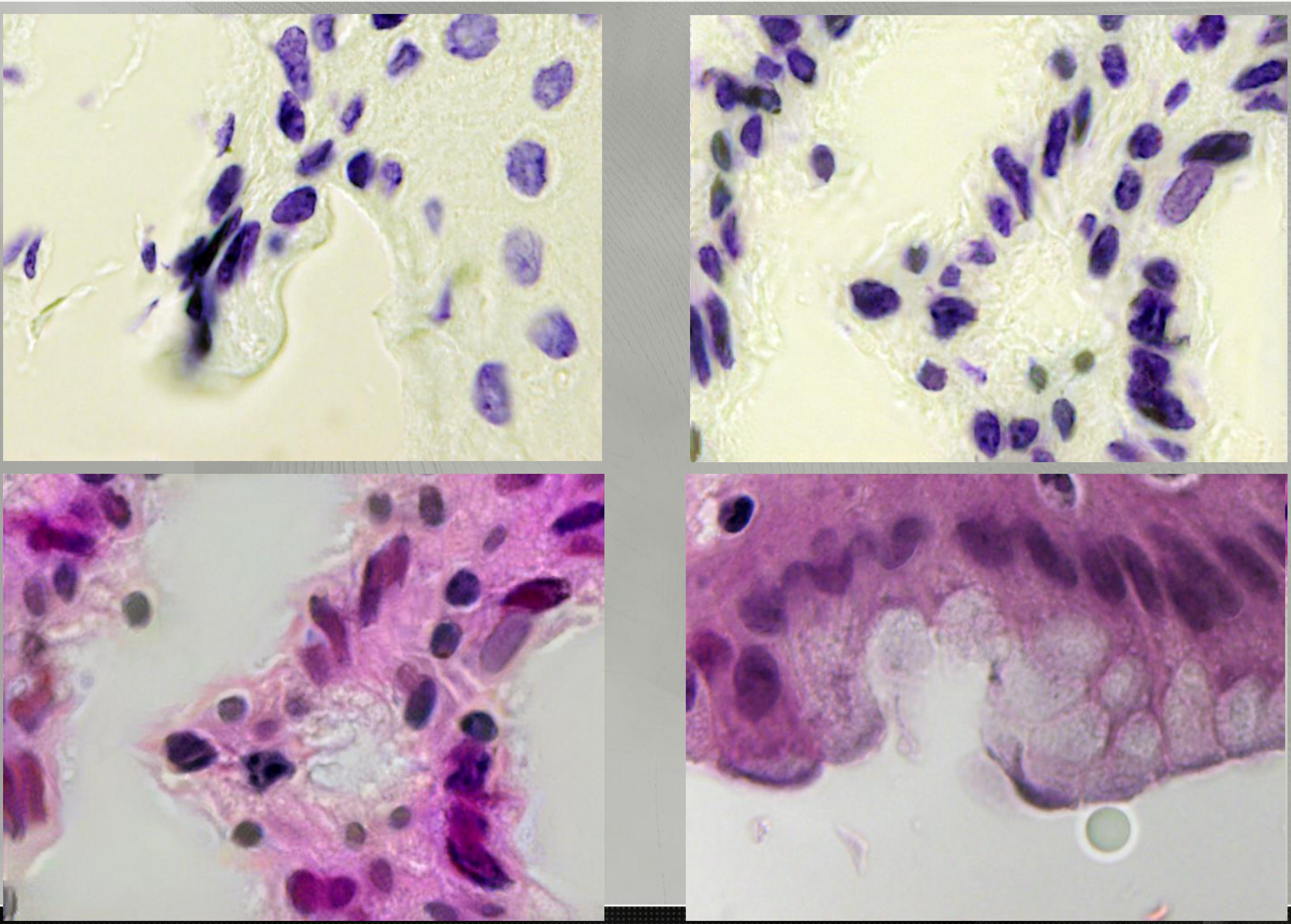


Original

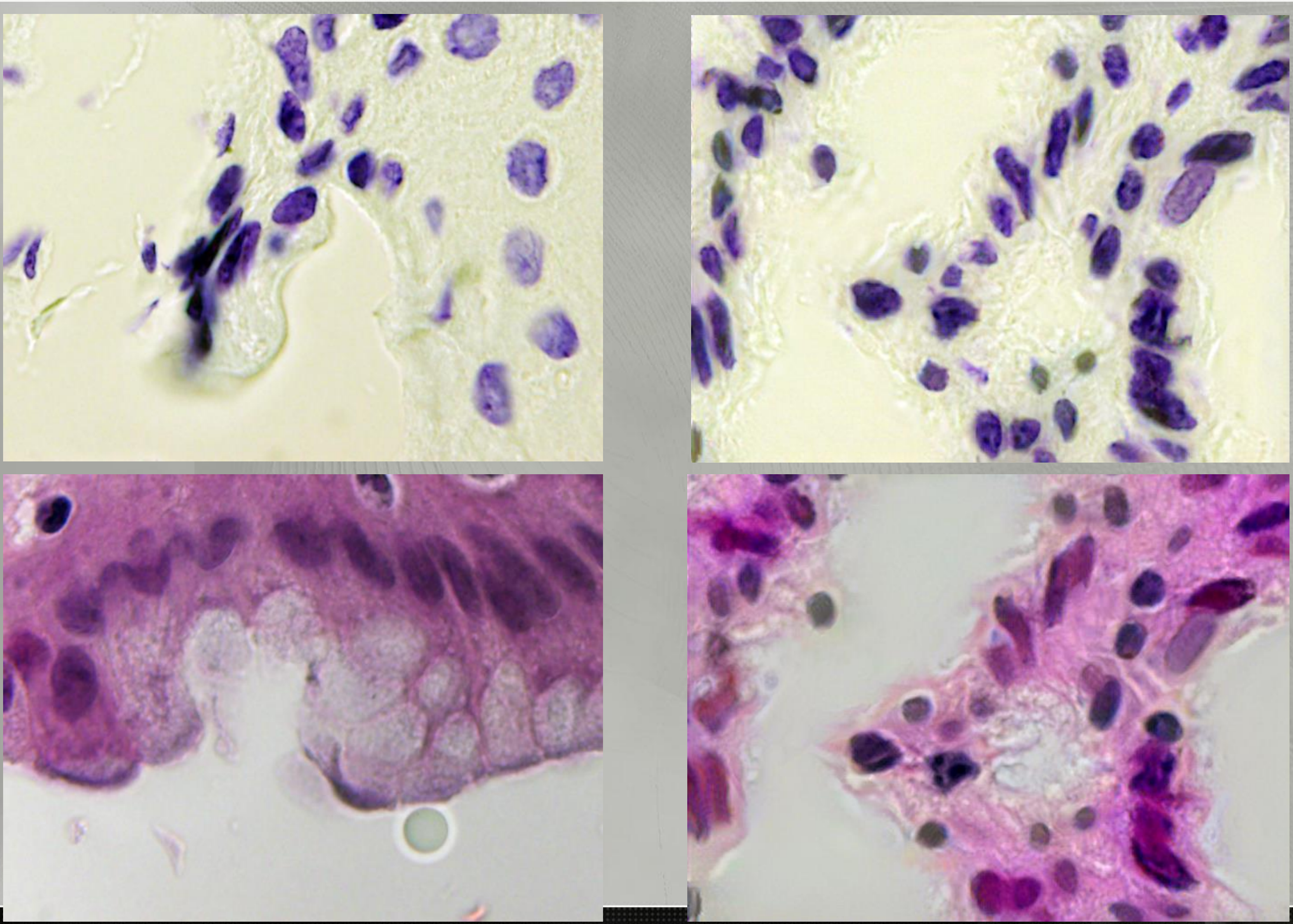


Fake

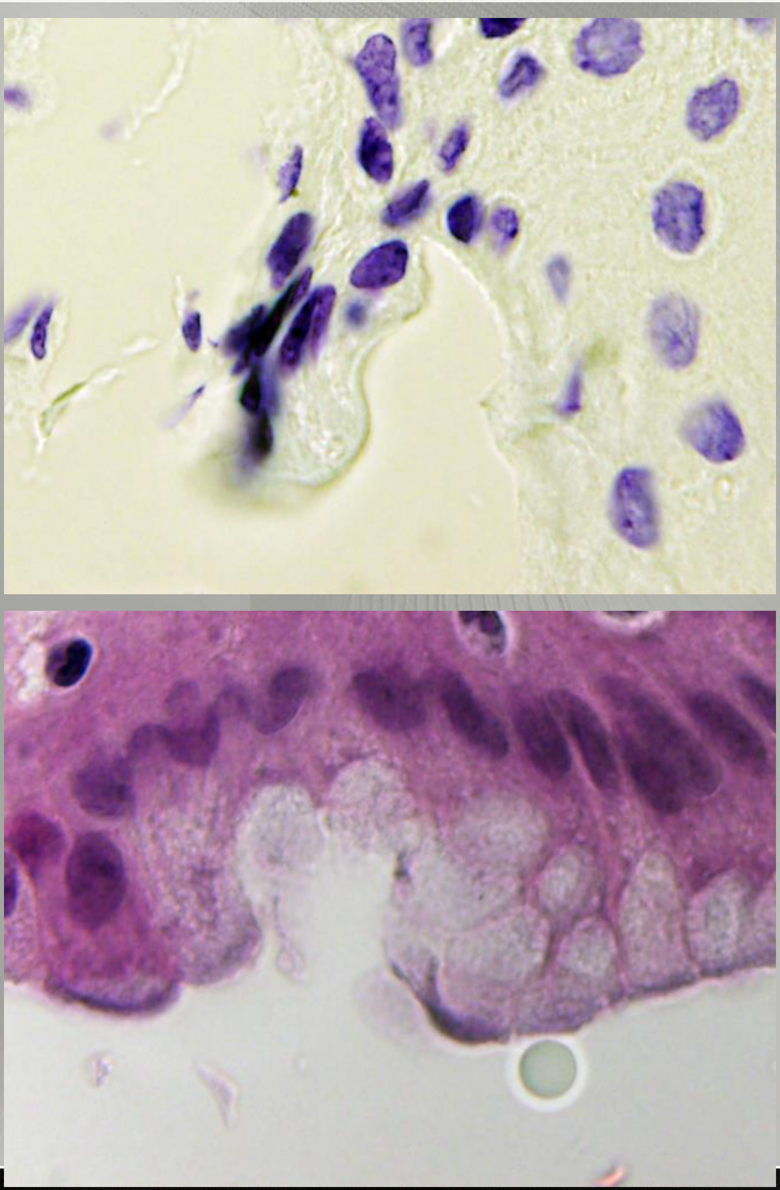
Tissue images



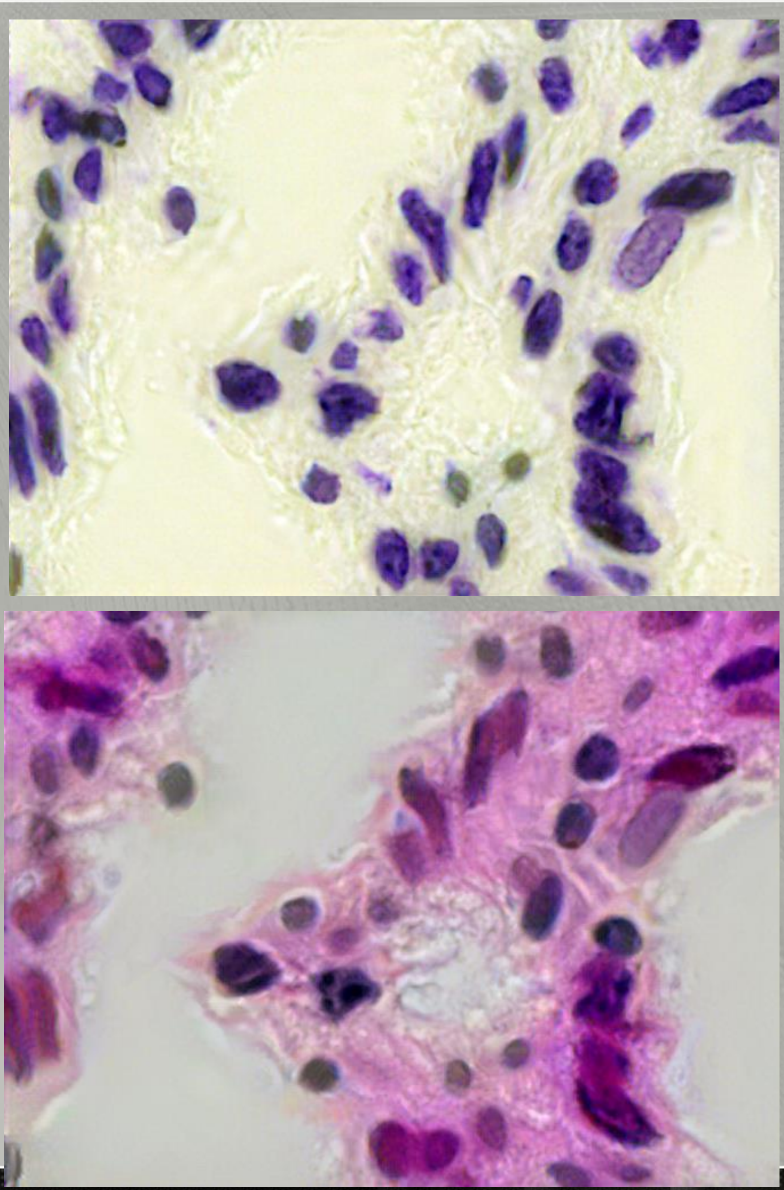
Tissue images



Tissue images

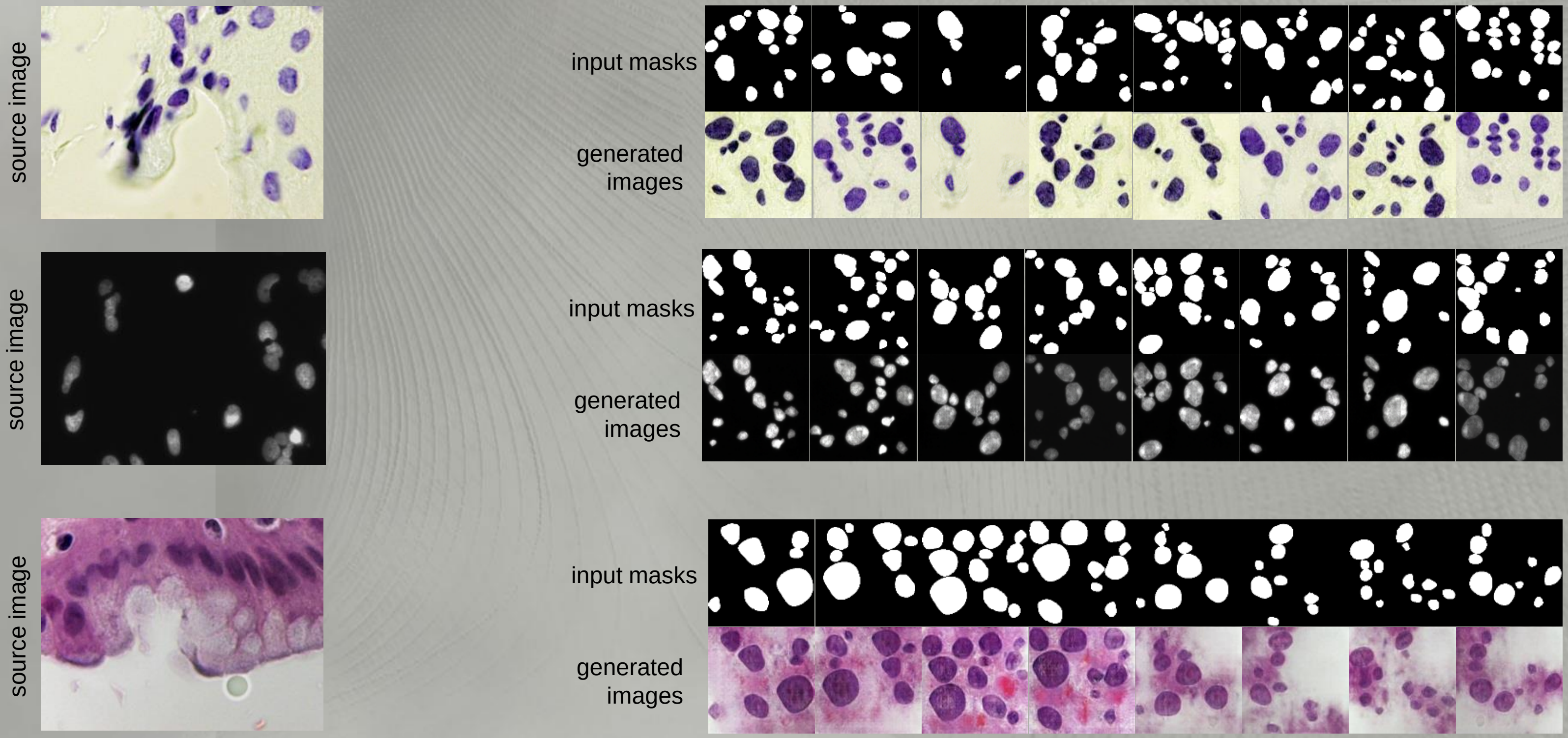


Original



Fake

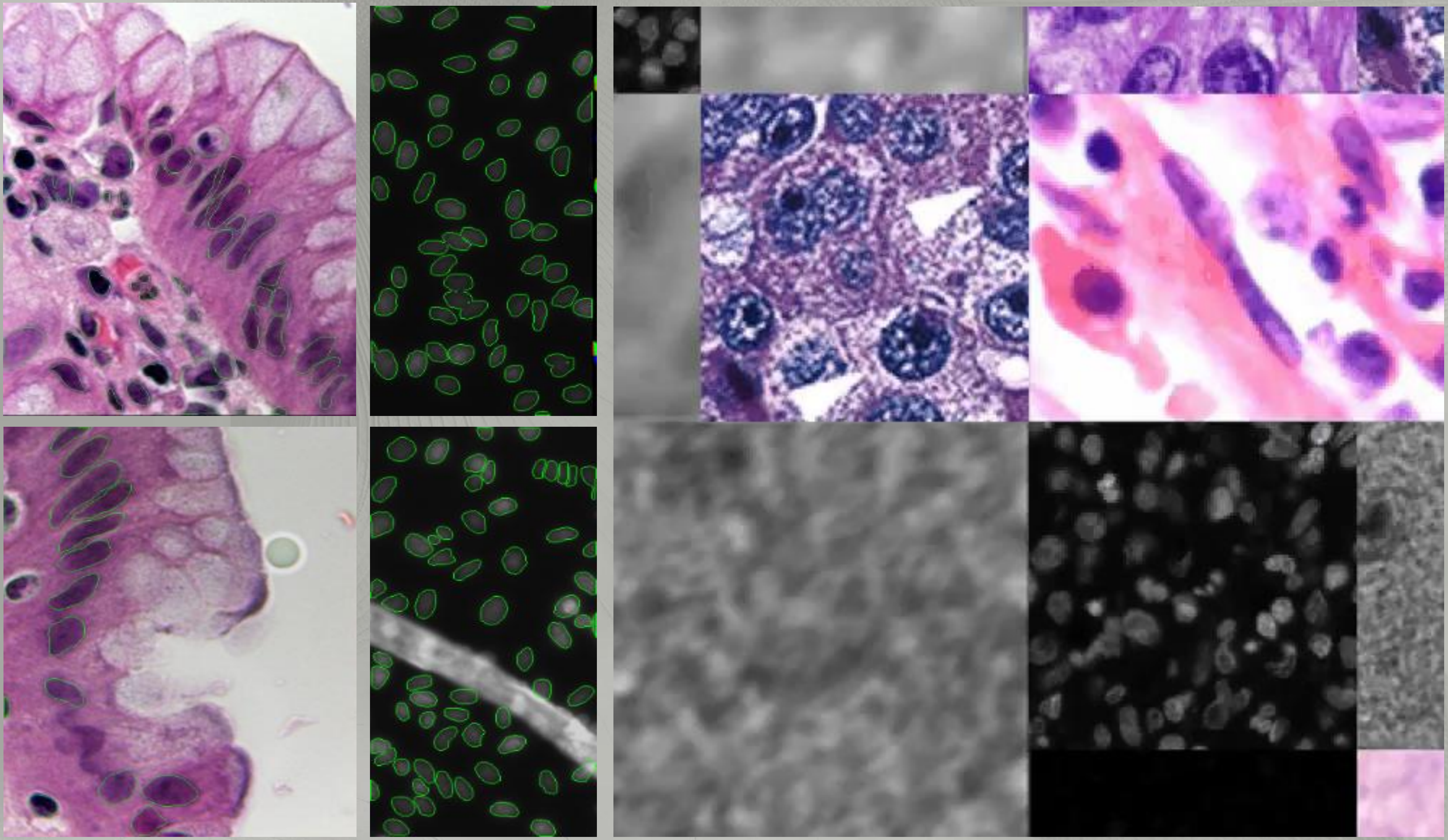
Image-to-image translation with conditional adversarial networks



Original

Fake

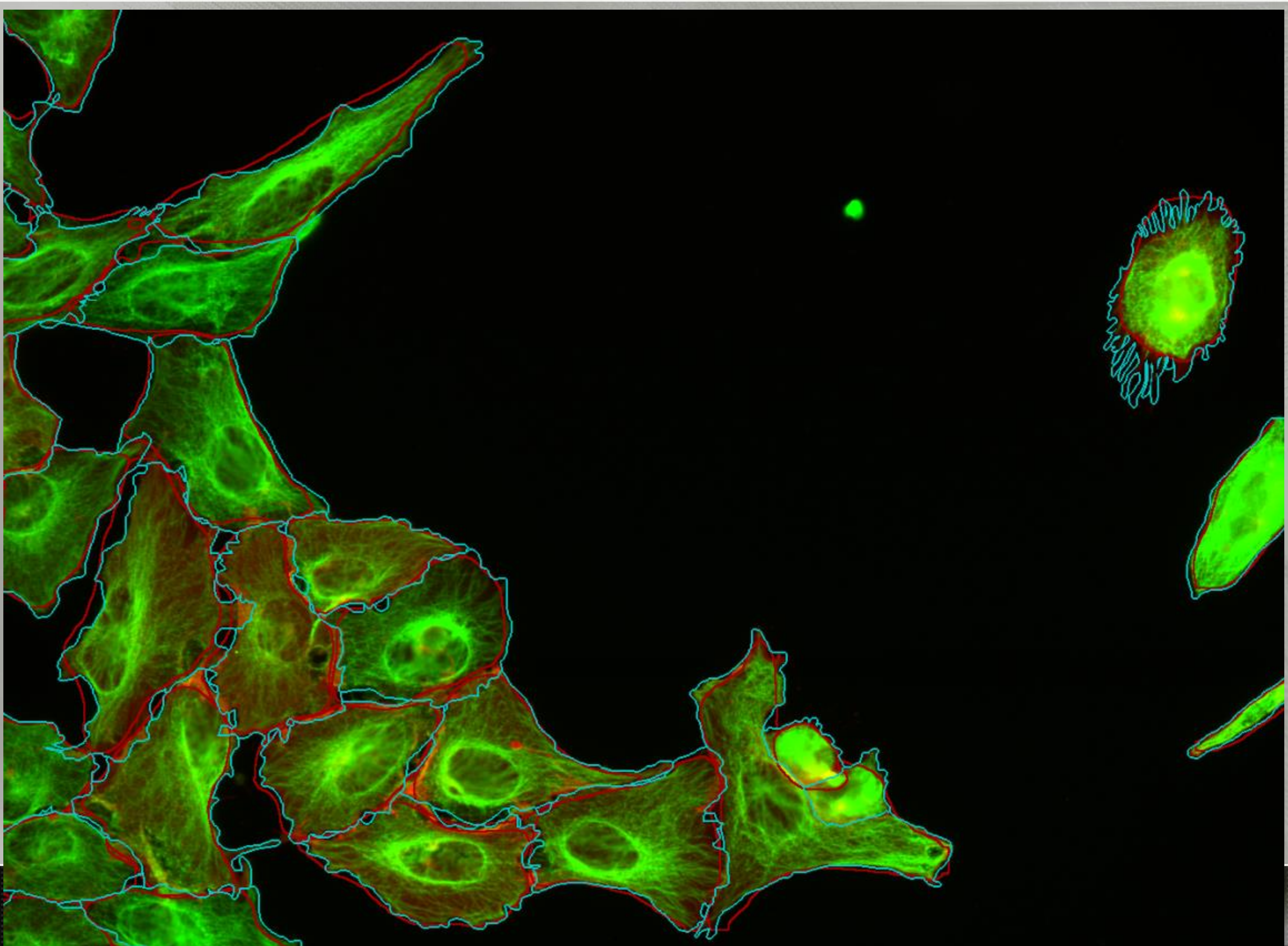
www.nucleAIzer.org



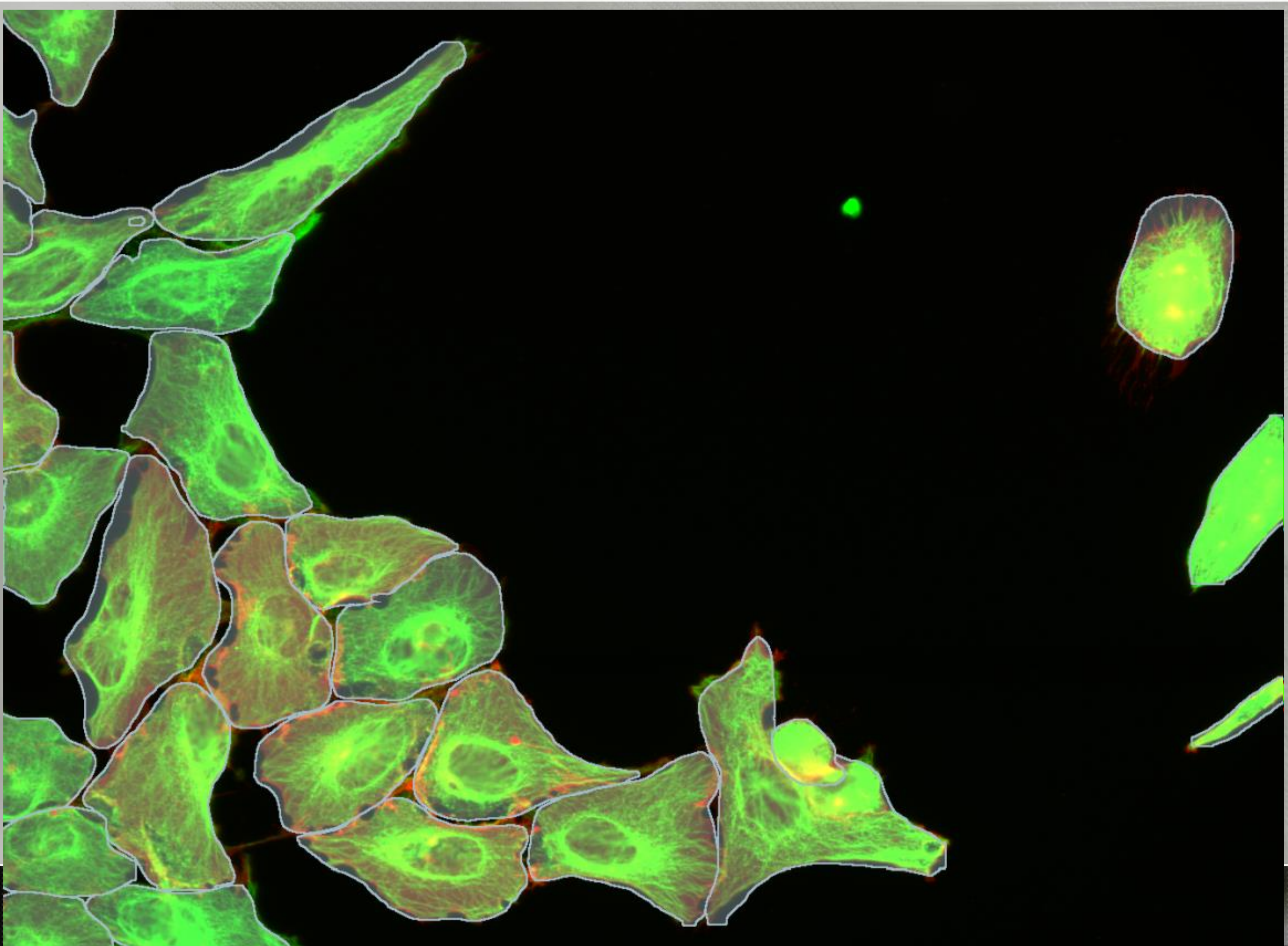
Hollandi, R., ***etal;*** under review

Cytoplasm segmentation using Mask-RCNN

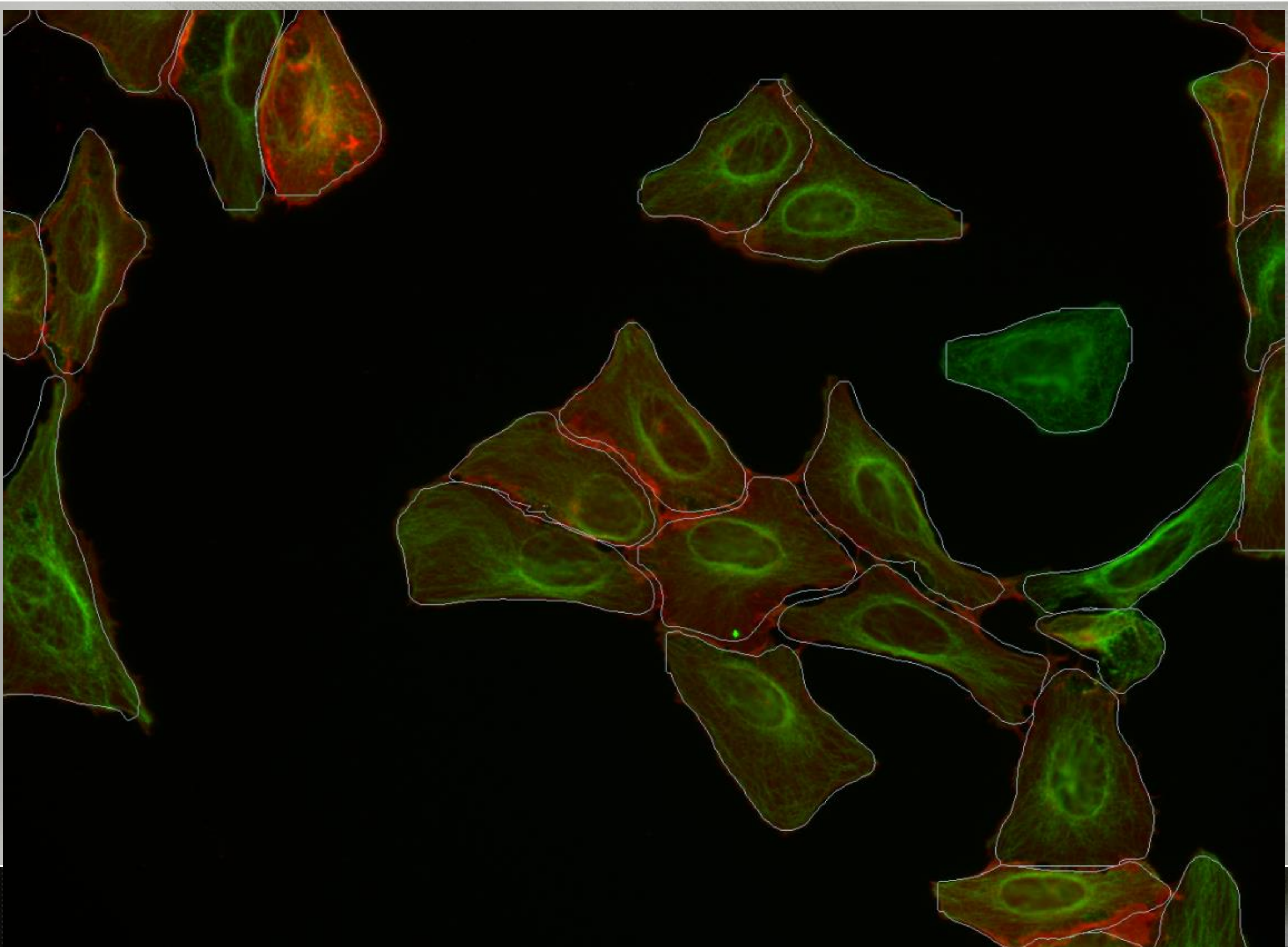
Cytoplasm segmentation using Mask-RCNN



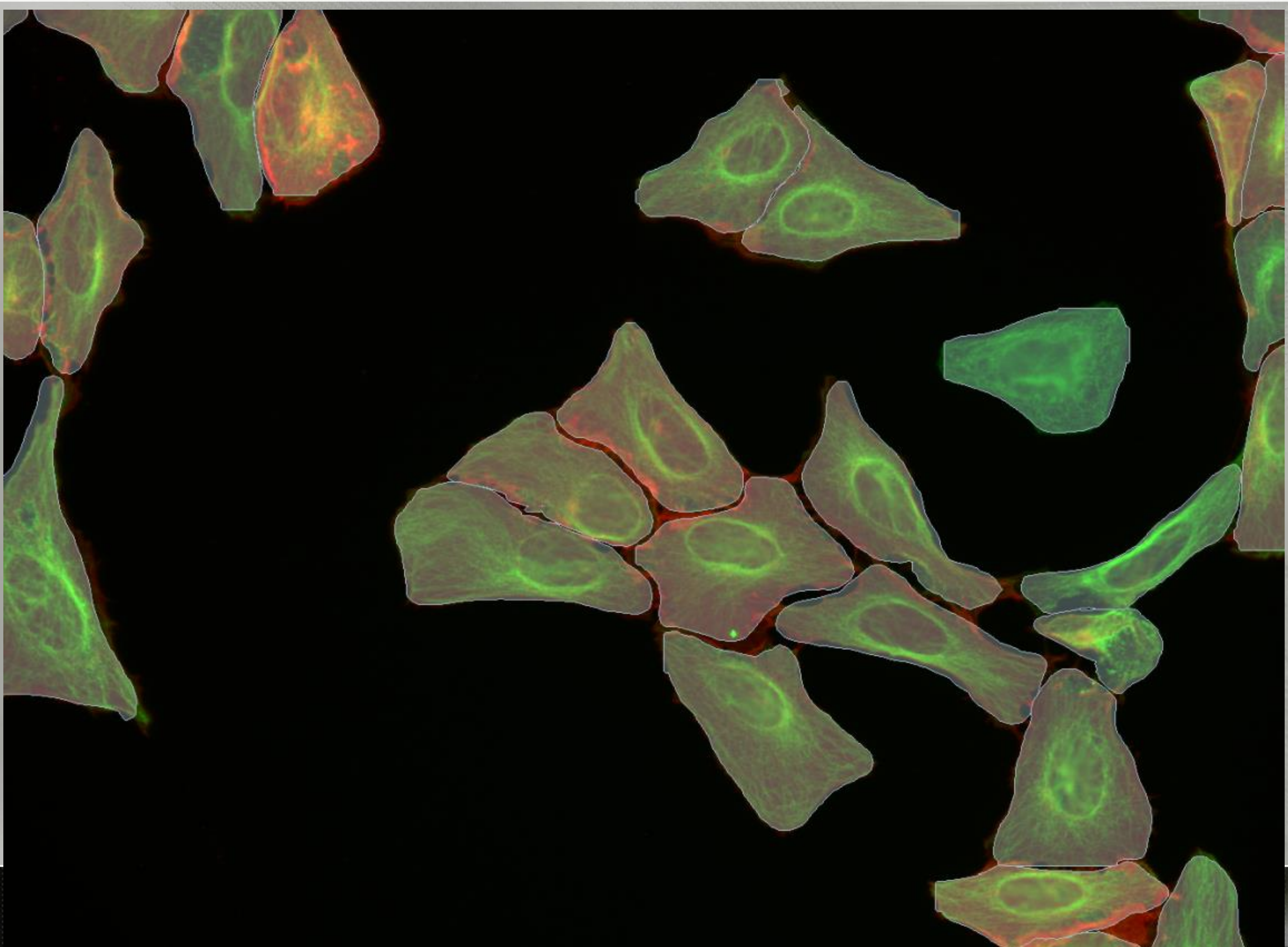
Cytoplasm segmentation using Mask-RCNN



Cytoplasm segmentation using Mask-RCNN

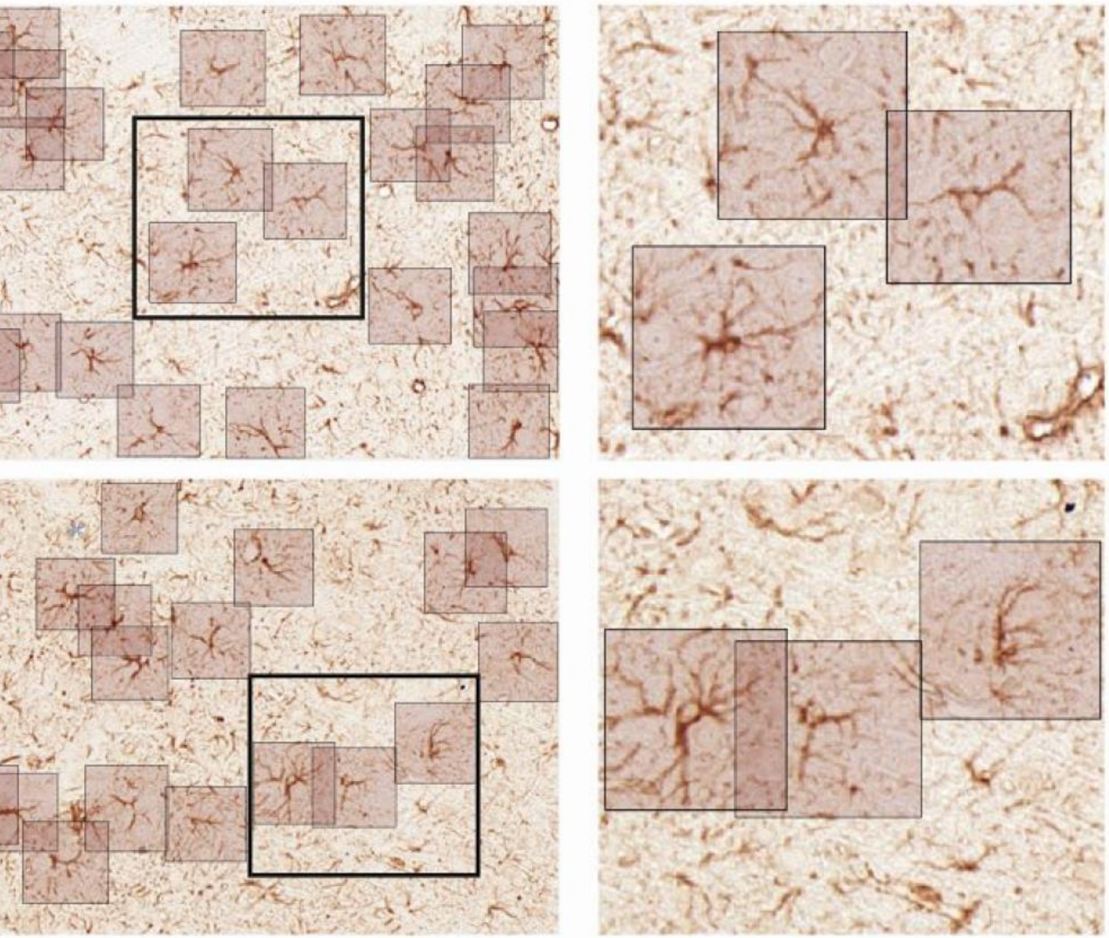


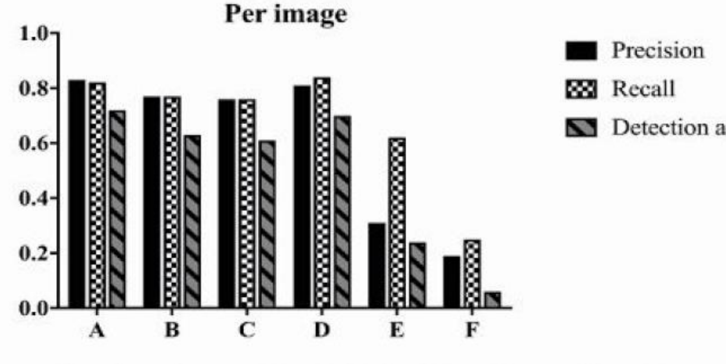
Cytoplasm segmentation using Mask-RCNN



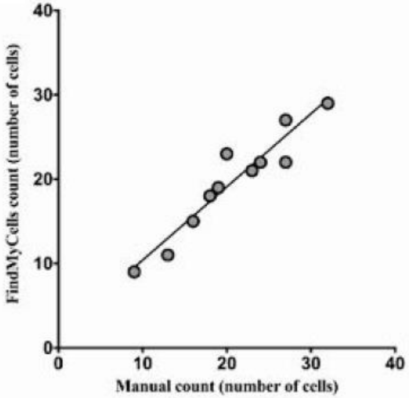
Astrocyte detection

at the level of human performance using deep learning



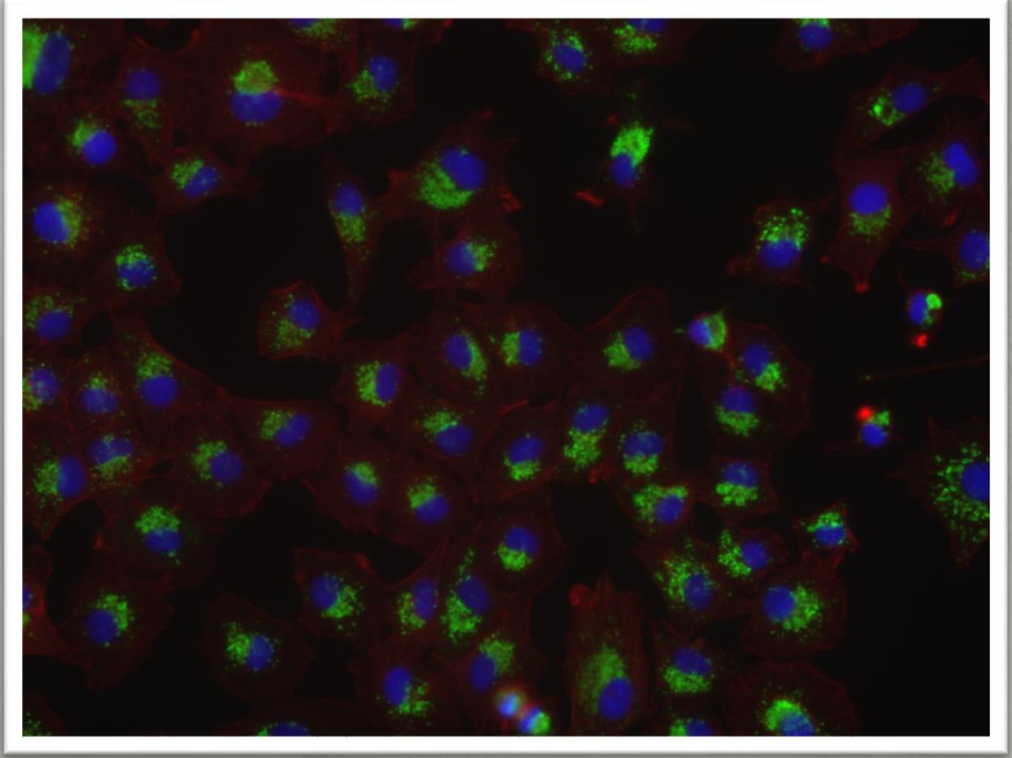


A - Intra-Expert Accuracy (Expert I), **B** - Intra-Expert Accuracy (Expert II),
C - Inter-expert Accuracy (Expert II vs Expert I), **D** - FindMyCells,
E - ilastik, **F** - Threshold based method

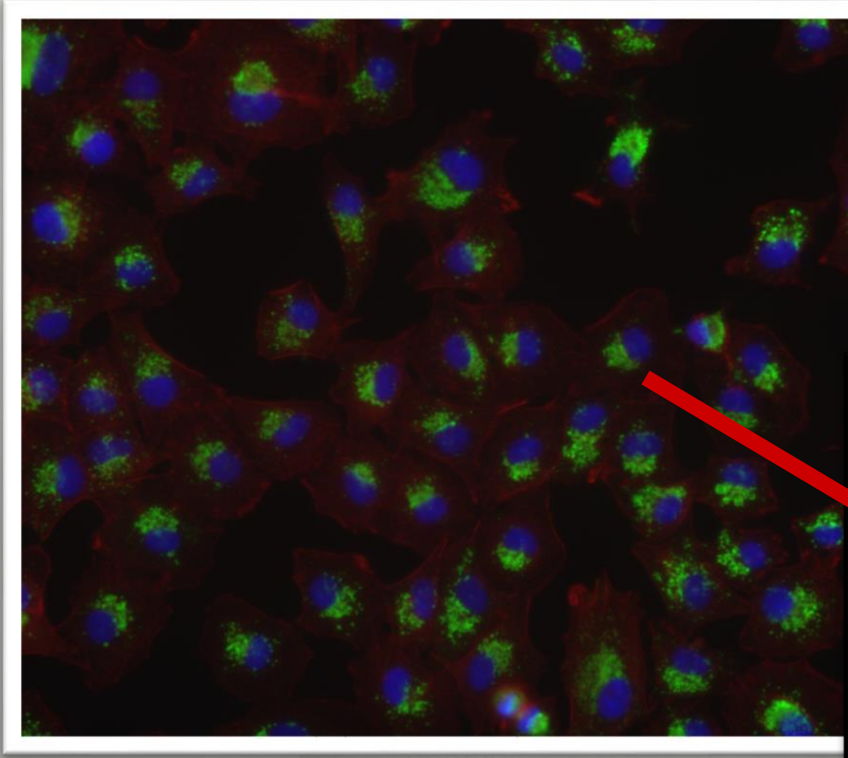


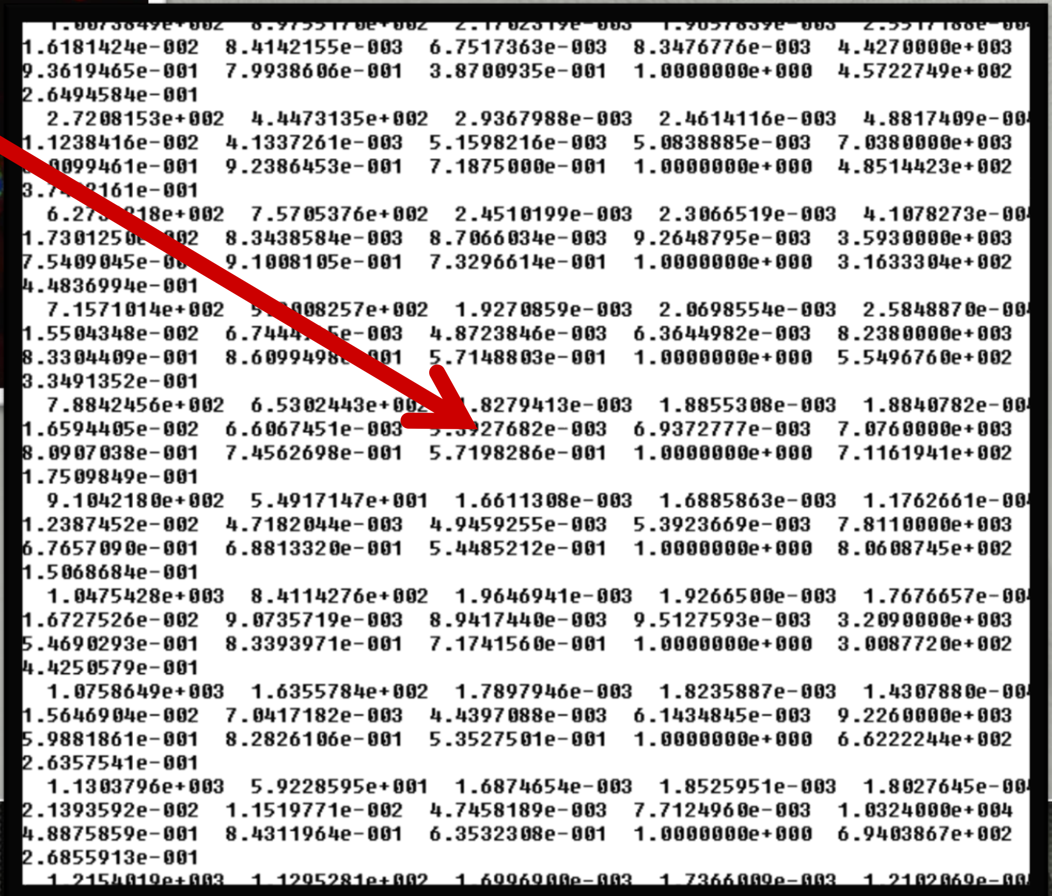
Suleymanova,Balassa, Saarma, Horvath: A deep convolutional neural network approach for astrocyte detection**. *Scientific Reports 2018***

Remember! Image processing. What next?



Remember! Image processing. What next?



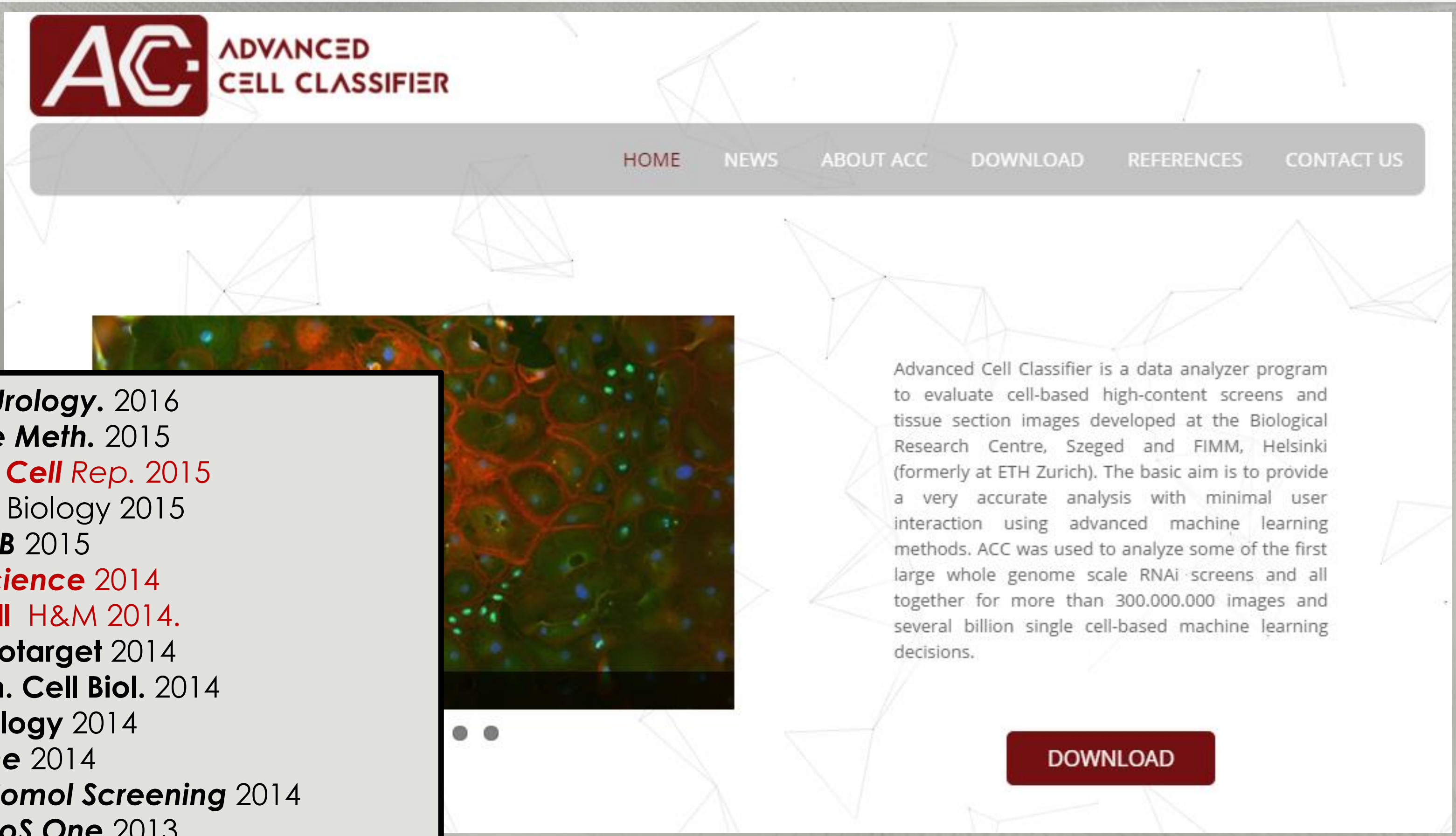


```
1.6181424e-002  8.4142155e-003  6.7517363e-003  8.3476776e-003  4.4270000e+003
9.3619465e-001  7.9938606e-001  3.8700935e-001  1.0000000e+000  4.5722749e+002
2.6494584e-001
  2.7208153e+002  4.4473135e+002  2.9367988e-003  2.4614116e-003  4.8817409e-00
1.1238416e-002  4.1337261e-003  5.1598216e-003  5.0838885e-003  7.0380000e+003
[illegible]0099461e-001  9.2386453e-001  7.1875000e-001  1.0000000e+000  4.8514423e+002
3.7[illegible]2161e-001
  6.27[illegible]218e+002  7.5705376e+002  2.4510199e-003  2.3066519e-003  4.1078273e-00
1.7301250[illegible]002  8.3438584e-003  8.7066034e-003  9.2648795e-003  3.5930000e+003
7.5409045e-00[illegible]  9.1008105e-001  7.3296614e-001  1.0000000e+000  3.1633304e+002
4.4836994e-001
  7.1571014e+002  5[illegible]008257e+002  1.9270859e-003  2.0698554e-003  2.5848870e-00
1.5504348e-002  6.744[illegible]5e-003  4.8723846e-003  6.3644982e-003  8.2380000e+003
8.3304409e-001  8.6099498[illegible]001  5.7148803e-001  1.0000000e+000  5.5496760e+002
3.3491352e-001
  7.8842456e+002  6.5302443e+00[illegible].8279413e-003  1.8855308e-003  1.8840782e-00
1.6594405e-002  6.6067451e-003  [illegible]927682e-003  6.9372777e-003  7.0760000e+003
8.0907038e-001  7.4562698e-001  5.7198286e-001  1.0000000e+000  7.1161941e+002
1.7509849e-001
  9.1042180e+002  5.4917147e+001  1.6611308e-003  1.6885863e-003  1.1762661e-00
1.2387452e-002  4.7182044e-003  4.9459255e-003  5.3923669e-003  7.8110000e+003
6.7657090e-001  6.8813320e-001  5.4485212e-001  1.0000000e+000  8.0608745e+002
1.5068684e-001
  1.0475428e+003  8.4114276e+002  1.9646941e-003  1.9266500e-003  1.7676657e-00
1.6727526e-002  9.0735719e-003  8.9417440e-003  9.5127593e-003  3.2090000e+003
5.4690293e-001  8.3393971e-001  7.1741560e-001  1.0000000e+000  3.0087720e+002
4.4250579e-001
  1.0758649e+003  1.6355784e+002  1.7897946e-003  1.8235887e-003  1.4307880e-00
1.5646904e-002  7.0417182e-003  4.4397088e-003  6.1434845e-003  9.2260000e+003
5.9881861e-001  8.2826106e-001  5.3527501e-001  1.0000000e+000  6.6222244e+002
2.6357541e-001
  1.1303796e+003  5.9228595e+001  1.6874654e-003  1.8525951e-003  1.8027645e-00
2.1393592e-002  1.1519771e-002  4.7458189e-003  7.7124960e-003  1.0324000e+004
4.8875859e-001  8.4311964e-001  6.3532308e-001  1.0000000e+000  6.9403867e+002
2.6855913e-001
  1.2154019e+003  1.1295281e+002  1.6996900e-003  1.7366009e-003  1.2102069e-00
```

Multi-parametric non-linear – Follow the fashion!

SINGLE CELL-BASED CLASSIFICATION FOR HCA

www.cellclassifier.org

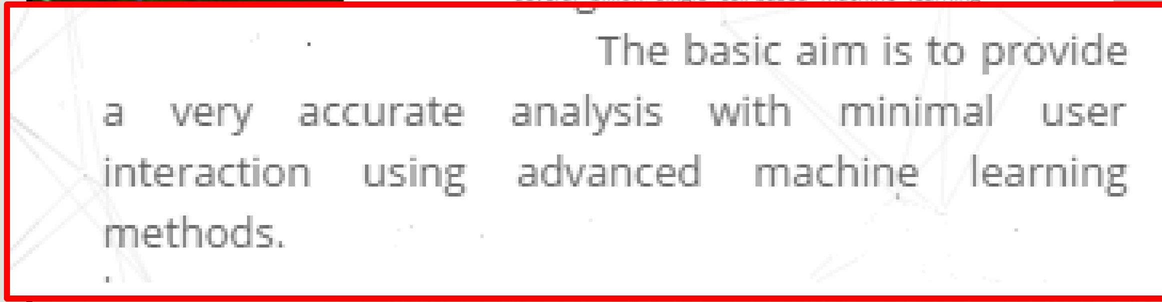


Advanced Cell Classifier is a data analyzer program to evaluate cell-based high-content screens and tissue section images developed at the Biological Research Centre, Szeged and FIMM, Helsinki (formerly at ETH Zurich). The basic aim is to provide a very accurate analysis with minimal user interaction using advanced machine learning methods. ACC was used to analyze some of the first large whole genome scale RNAi screens and all together for more than 300.000.000 images and several billion single cell-based machine learning decisions.

Saeed *etal.*: ***Eur. Urology.*** 2016
Smith *etal.*: ***Nature Meth.*** 2015
Badertscher etal.: ***Cell*** *Rep.* 2015
Kalapis etal.: **PLoS** Biology 2015
Ungricht *etal.*: ***JCB*** 2015
Banerjee, *etal.*; ***Science*** 2014
Balistreri, *etal.*: **Cell** H&M 2014.
Boutter *etal.*: **Oncotarget** 2014
Marino *etal.*: **Meth. Cell Biol.** 2014
Meier *etal.*: **J. Virology** 2014
Kiss *etal.*: ***PLoS One*** 2014
Smith *etal.*: ***J. of Biomol Screening*** 2014
Banerjee, *etal.*; ***PLoS One*** 2013
Palazzolo, Horvath Wong; ***PLoS One*** 2012
Misselwitz, Horvath *etal.*; ***PLoS Pathogens*** 2012
Huotari, Horvath *etal.*; ***PNAS*** 2012
Yamauchi, Horvath *etal.*; ***PLoS Pathogens*** 2011
Horvath, Wild *etal.*; ***J. of Biomol Screening*** 2011
Laurell, Beck *etal.*; ***Cell*** 2011
Wild, Horvath *etal.*; ***PLoS Biol*** 2010
Turgay, Ungricht *etal.*; ***EMBO journal*** 2010

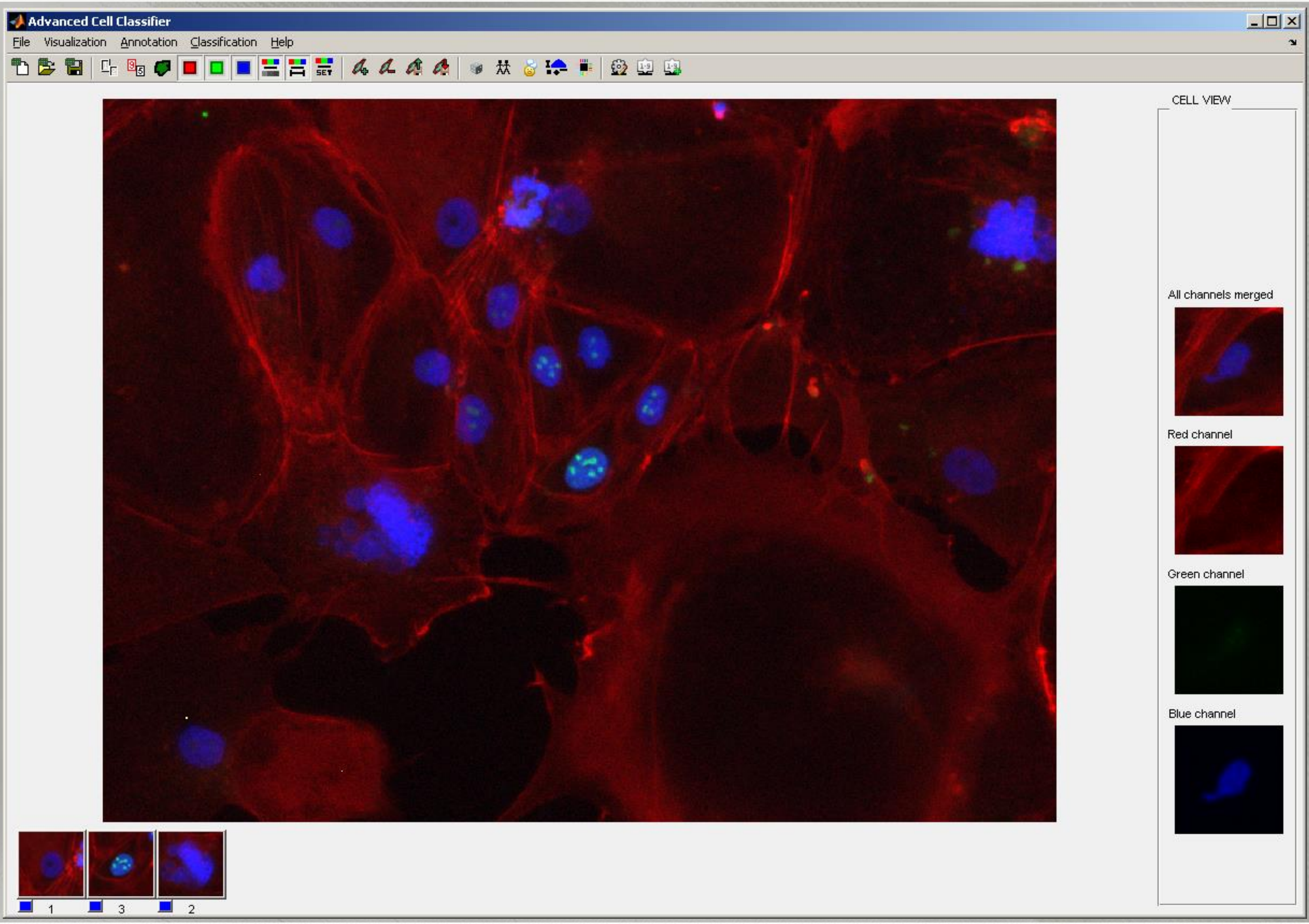
www.cellclassifier.org

Advanced Cell Classifier is a data analyzer program to evaluate cell-based high-content screens and tissue section images developed at the Biological Research Centre, Szeged and FIMM, Helsinki (formerly at ETH Zurich). The basic aim is to provide a very accurate analysis with minimal user interaction using advanced machine learning methods. ACC was used to analyze some of the first large whole genome scale RNAi screens and all together for more than 300.000.000 images and several billion single cell-based machine learning


The basic aim is to provide a very accurate analysis with minimal user interaction using advanced machine learning methods.

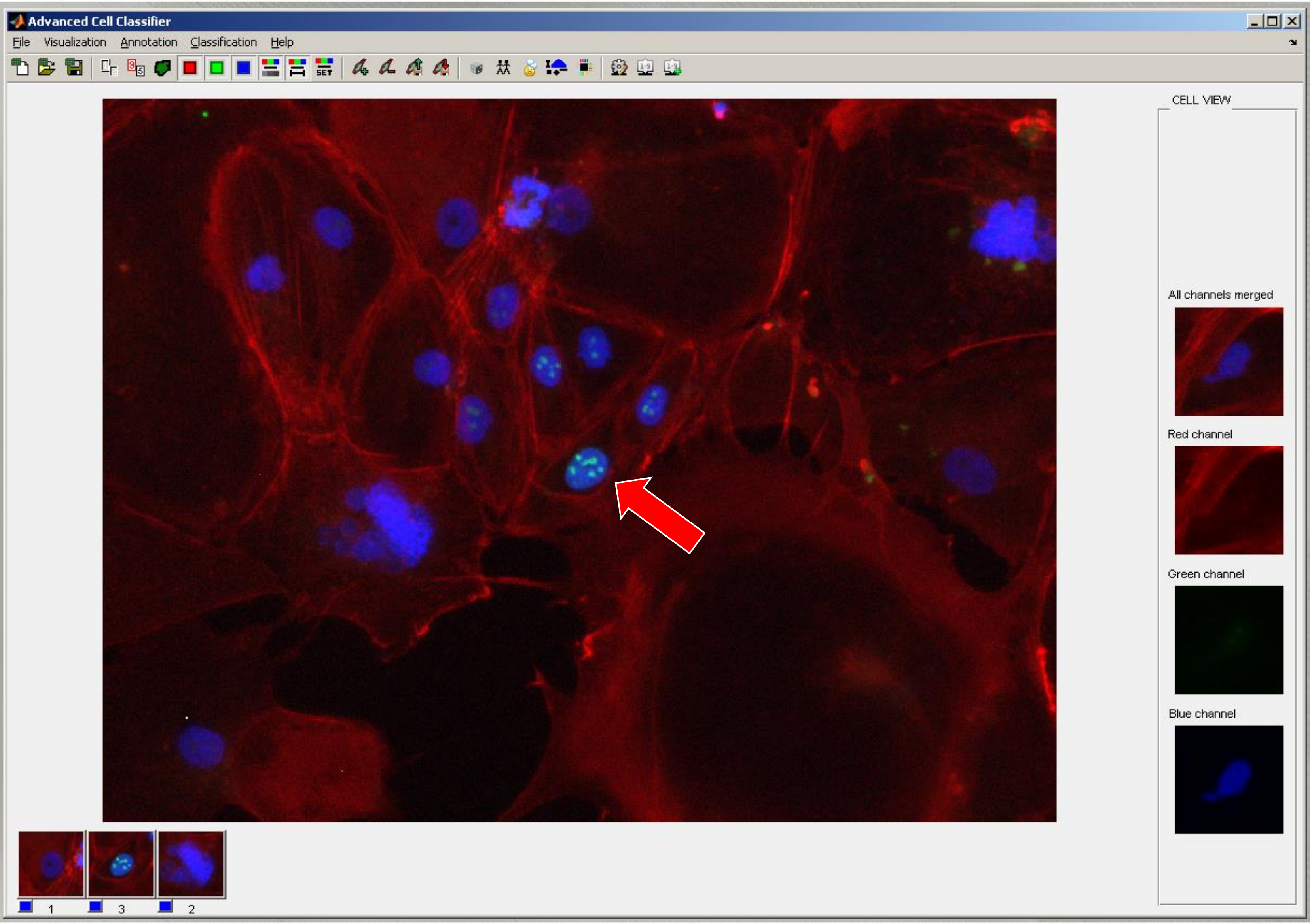
Saeed *etal.*: ***Eur. Urology.*** 2016
Smith *etal.*: ***Nature Meth.*** 2015
Badertscher etal.: ***Cell*** *Rep.* 2015
Kalapis etal.: **PLoS** Biology 2015
Ungricht *etal.*: ***JCB*** 2015
Banerjee, *etal.*; ***Science*** 2014
Balistreri, *etal.*: **Cell** H&M 2014.
Boutter *etal.*: **Oncotarget** 2014
Marino *etal.*: **Meth. Cell Biol.** 2014
Meier *etal.*: **J. Virology** 2014
Kiss *etal.*: ***PLoS One*** 2014
Smith *etal.*: ***J. of Biomol Screening*** 2014
Banerjee, *etal.*; ***PLoS One*** 2013
Palazzolo, Horvath Wong; ***PLoS One*** 2012
Misselwitz, Horvath *etal.*; ***PLoS Pathogens*** 2012
Huotari, Horvath *etal.*; ***PNAS*** 2012
Yamauchi, Horvath *etal.*; ***PLoS Pathogens*** 2011
Horvath, Wild *etal.*; ***J. of Biomol Screening*** 2011
Laurell, Beck *etal.*; ***Cell*** 2011
Wild, Horvath *etal.*; ***PLoS Biol*** 2010
Turgay, Ungricht *etal.*; ***EMBO journal*** 2010

Training the machine



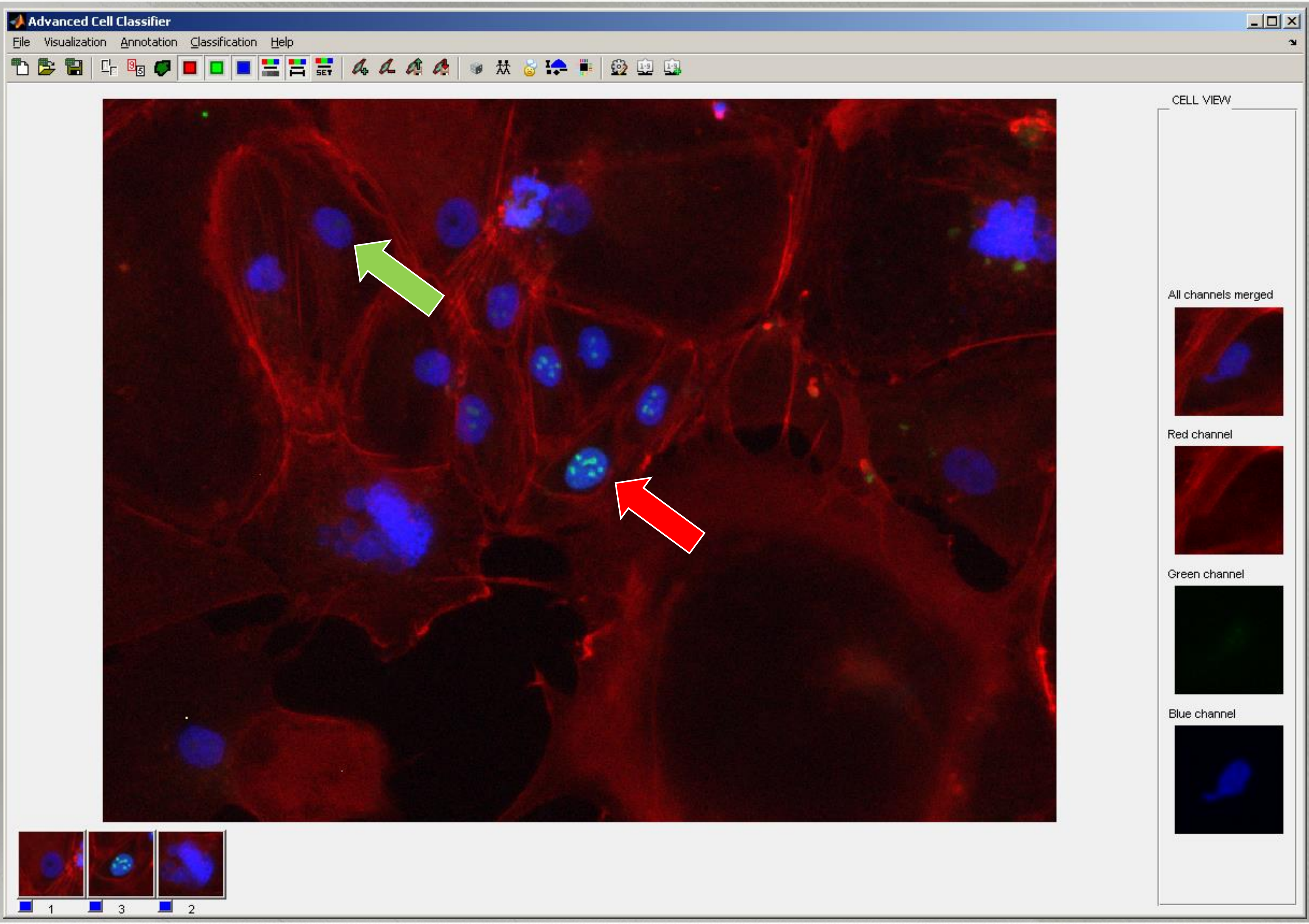
Piccinini F. and T. Balassa, *etal:. Cell Systems 2017*

Training the machine



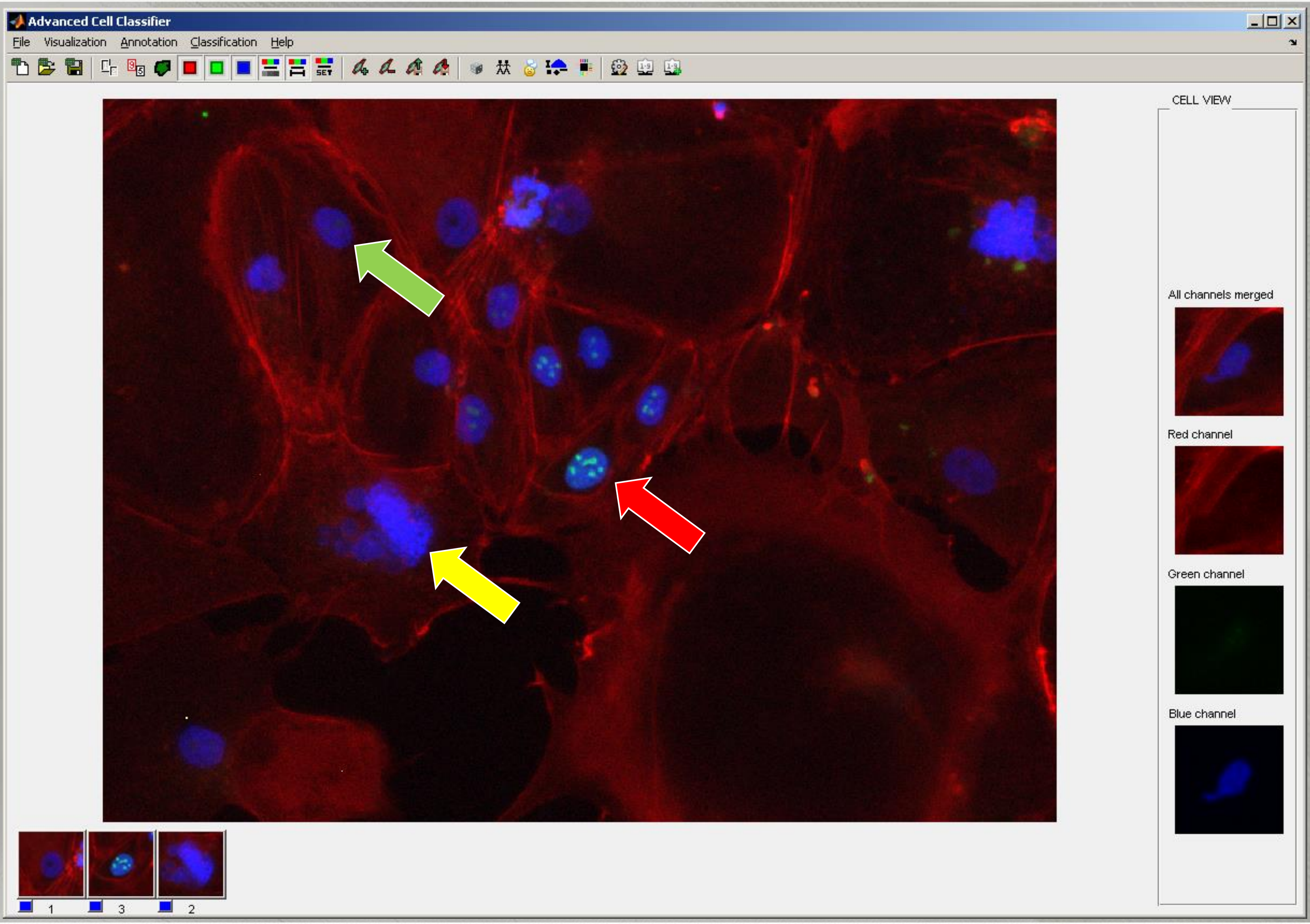
Piccinini F. and T. Balassa, *etal:. Cell Systems 2017*

Training the machine



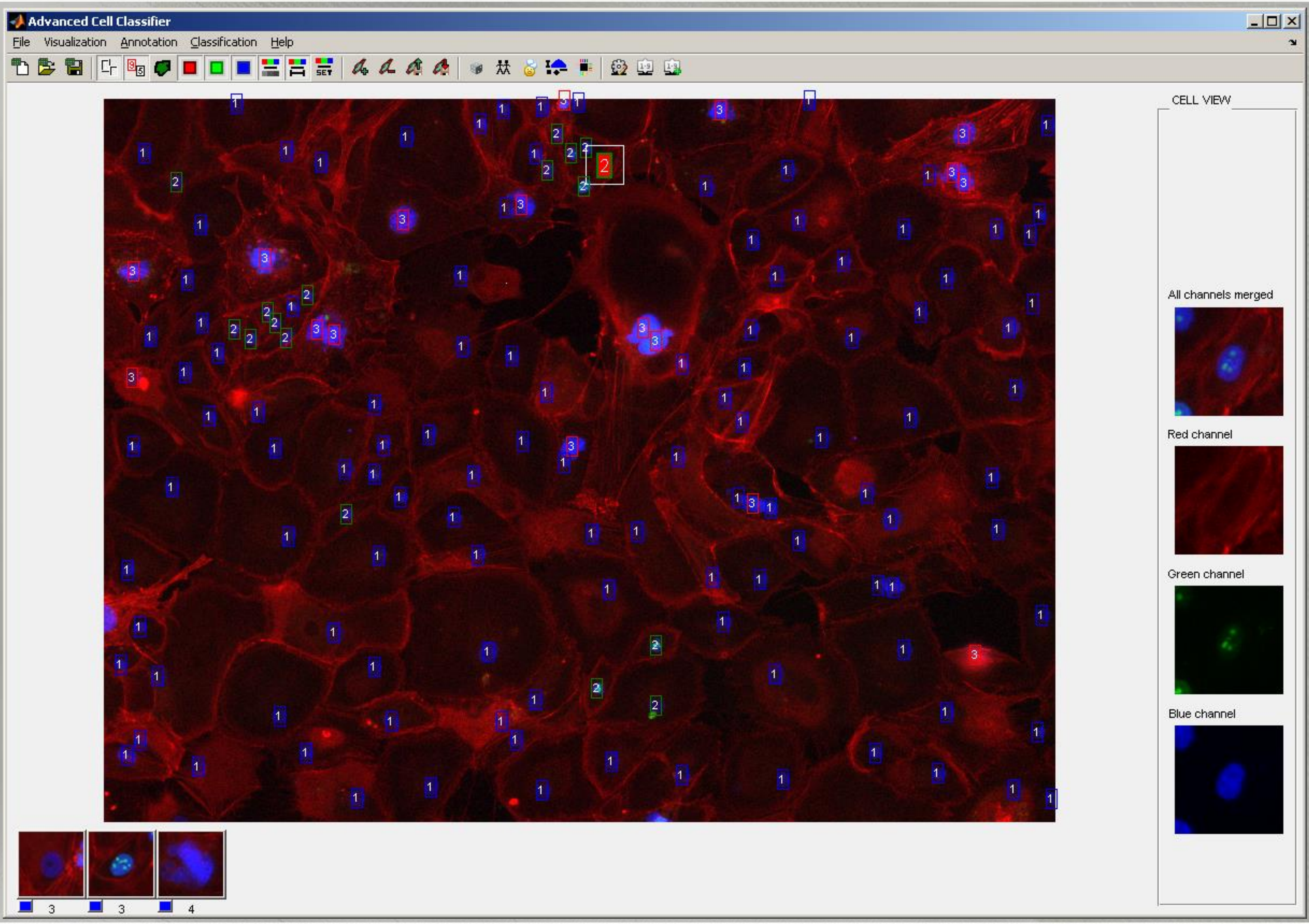
Piccinini F. and T. Balassa, *etal:. Cell Systems 2017*

Training the machine



Piccinini F. and T. Balassa, *etal:. Cell Systems 2017*

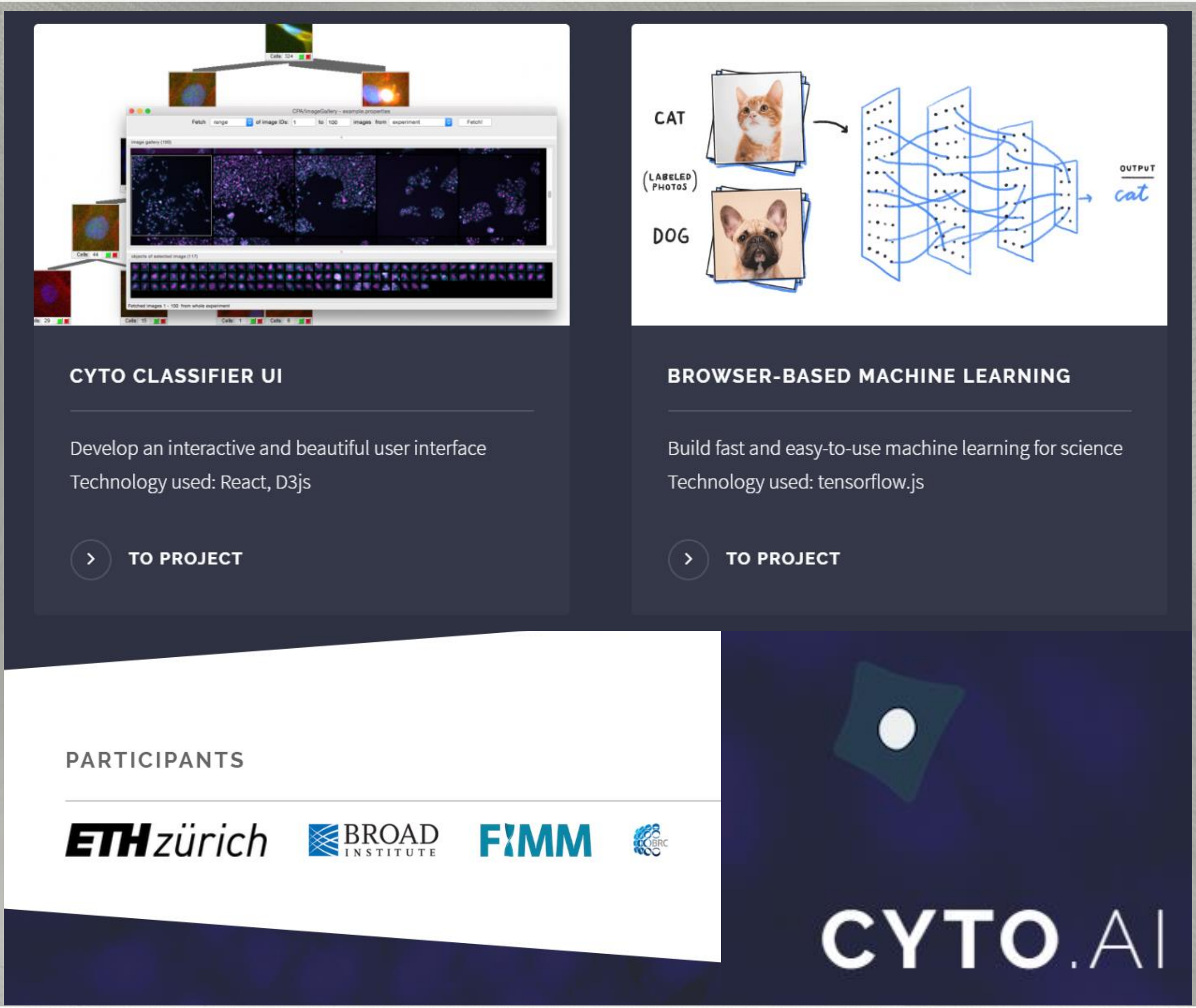
Automated classification



Piccinini F. and T. Balassa, *etal:. Cell Systems 2017*

Future: lets WEB and deeplearn

Anne Carpenter
Broad Institute



CYTO CLASSIFIER UI

Develop an interactive and beautiful user interface
Technology used: React, D3js

TO PROJECT

BROWSER-BASED MACHINE LEARNING

Build fast and easy-to-use machine learning for science
Technology used: tensorflow.js

TO PROJECT

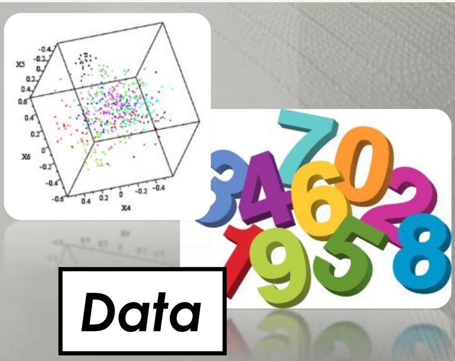
PARTICIPANTS

ETH zürich · BROAD INSTITUTE · FIMM · BRC

CYTO.AI

Classical data processing

Classical data processing



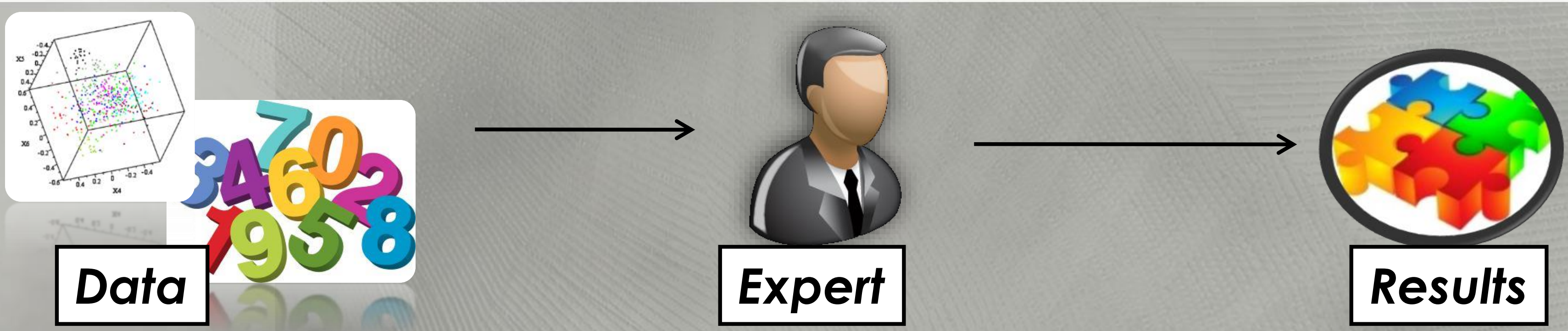
Data

Classical data processing

Data

Expert

Classical data processing

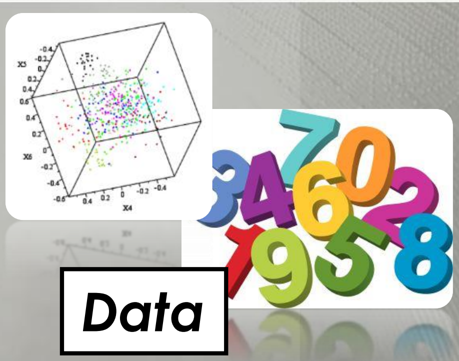


Data

Expert

Results

Classical data processing



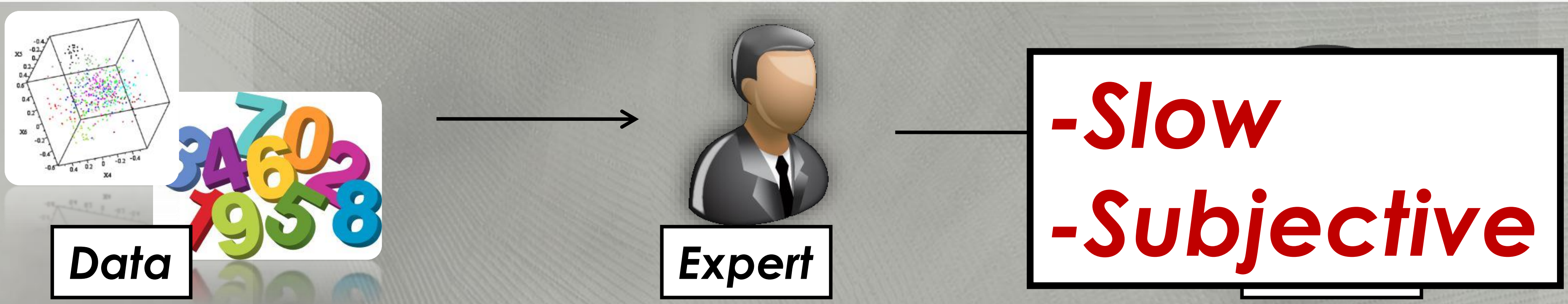
Data

Expert

-Slow

-Subjective

Classical data processing

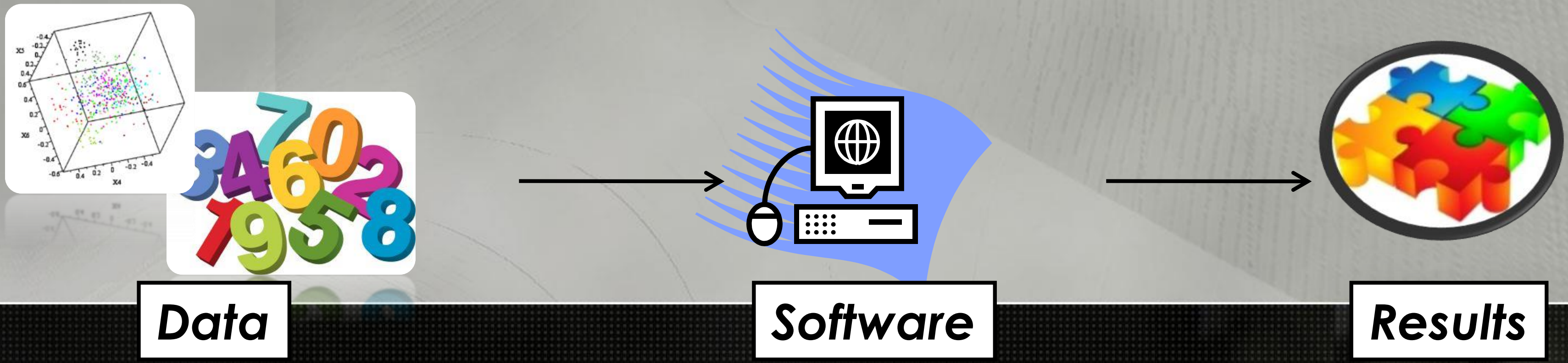


Data

Expert

-Slow

-Subjective

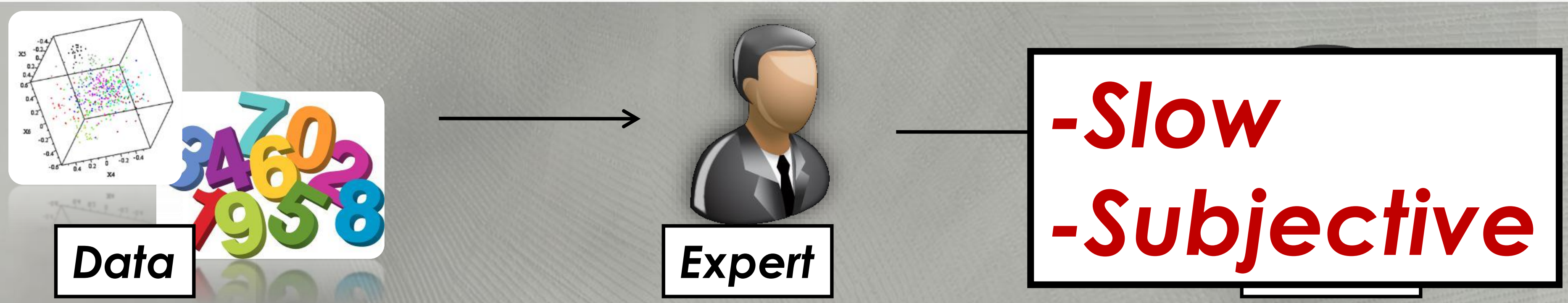


Data

Software

Results

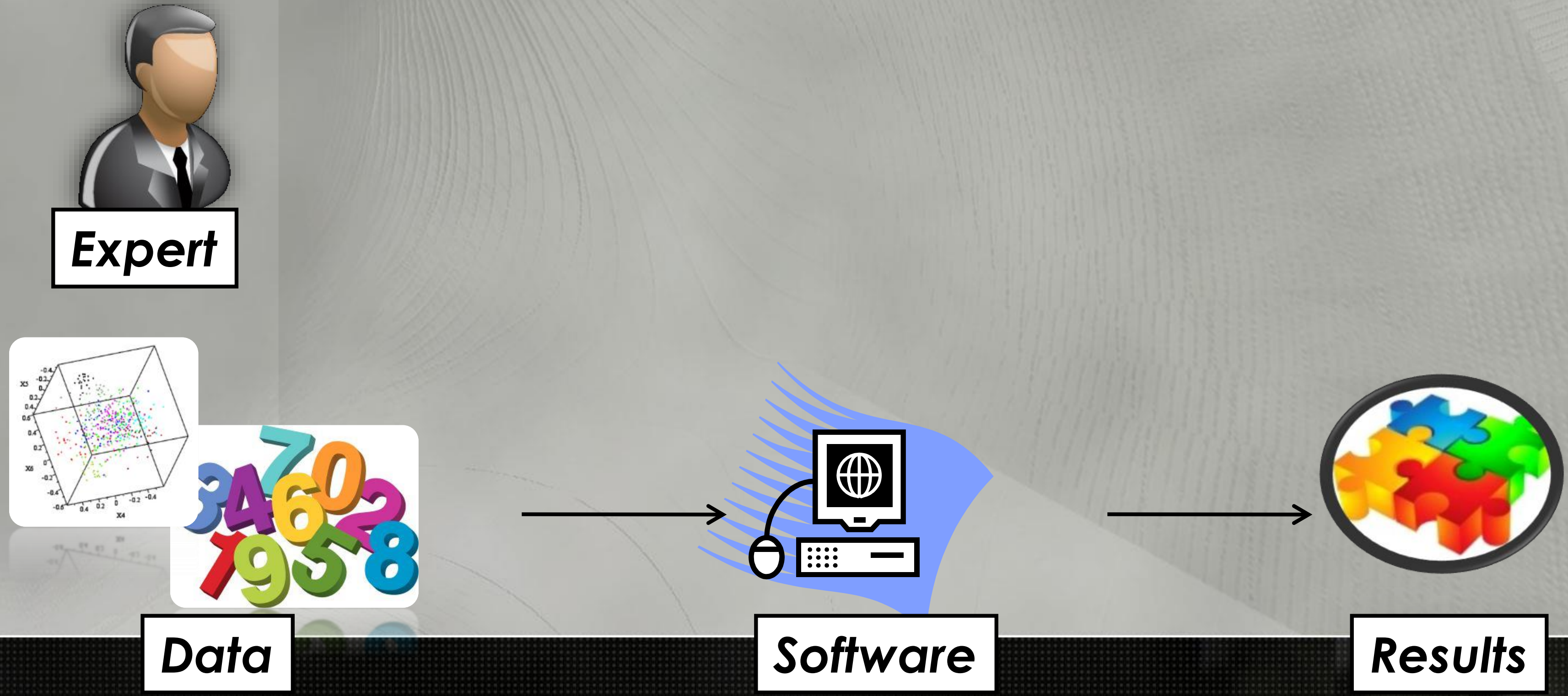
Classical data processing



Data

Expert

-Slow
-Subjective



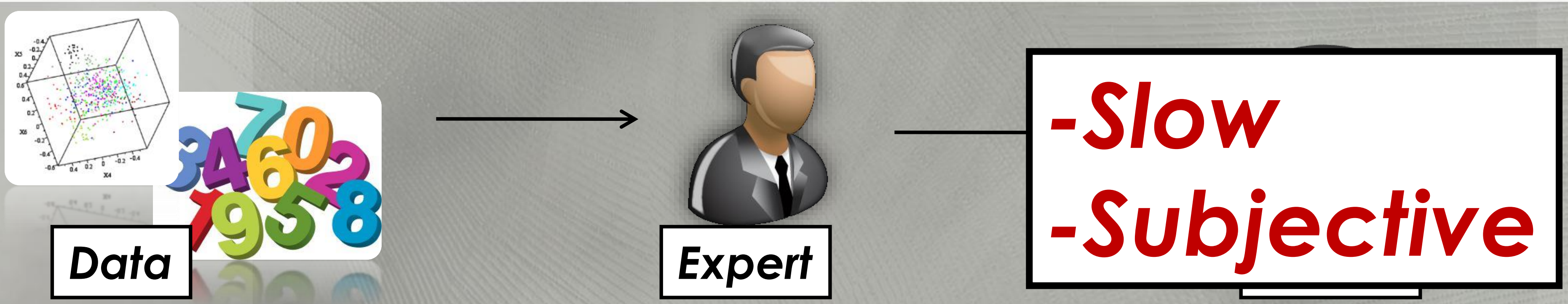
Expert

Data

Software

Results

Classical data processing

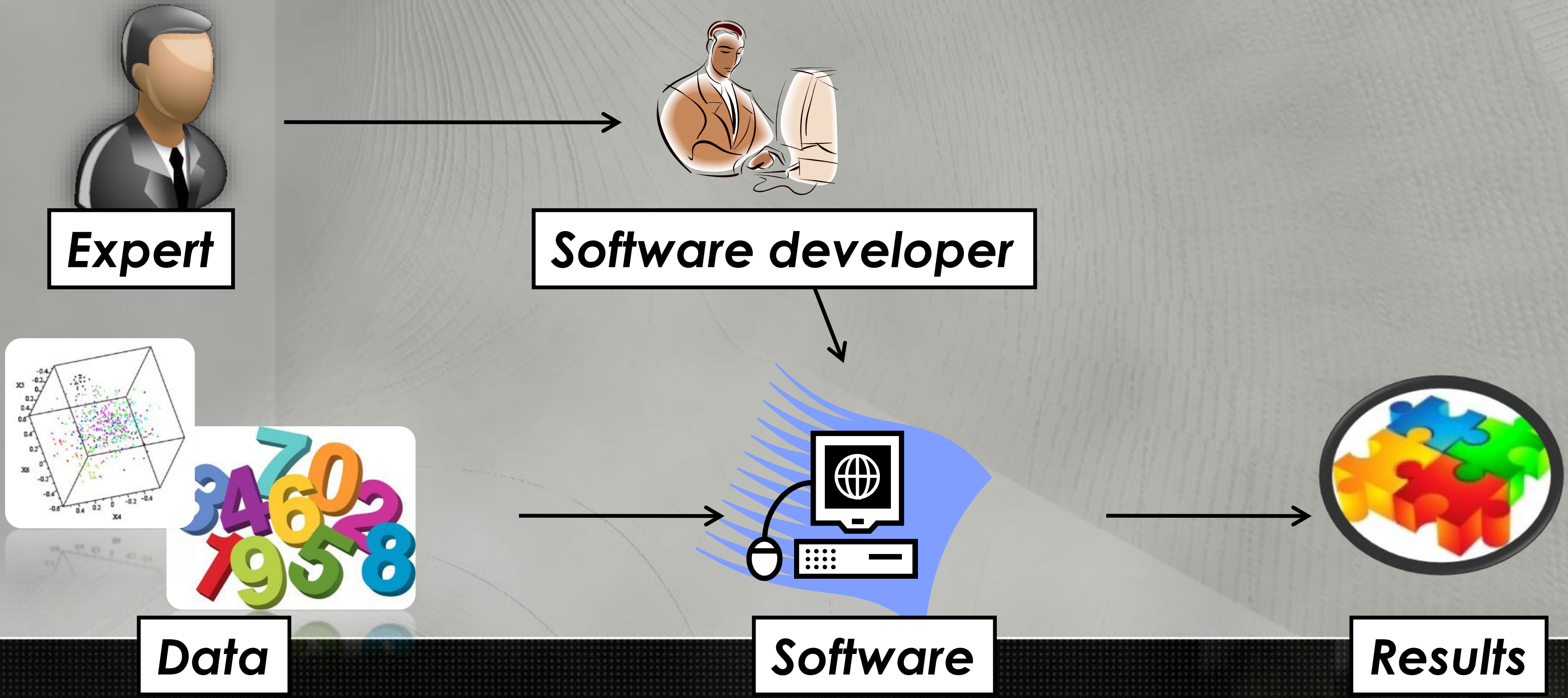


Data

Expert

-Slow

-Subjective



Expert

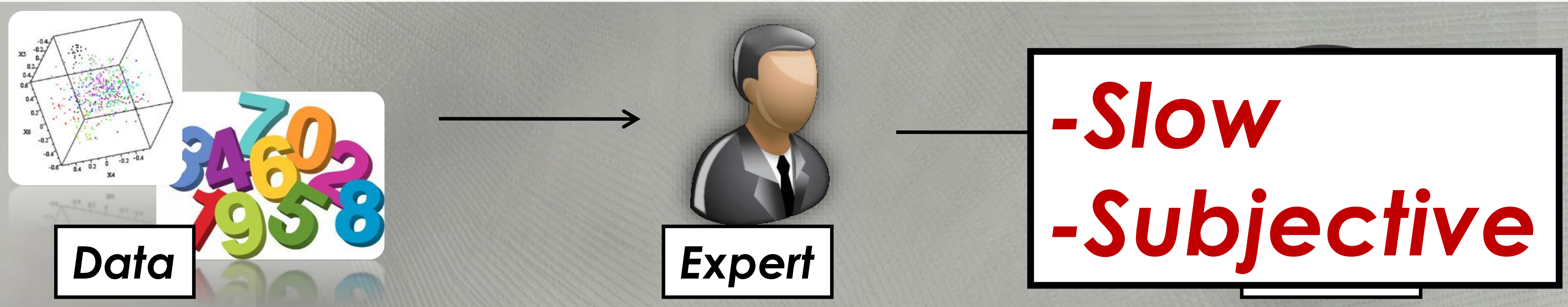
Software developer

Data

Software

Results

Classical data processing

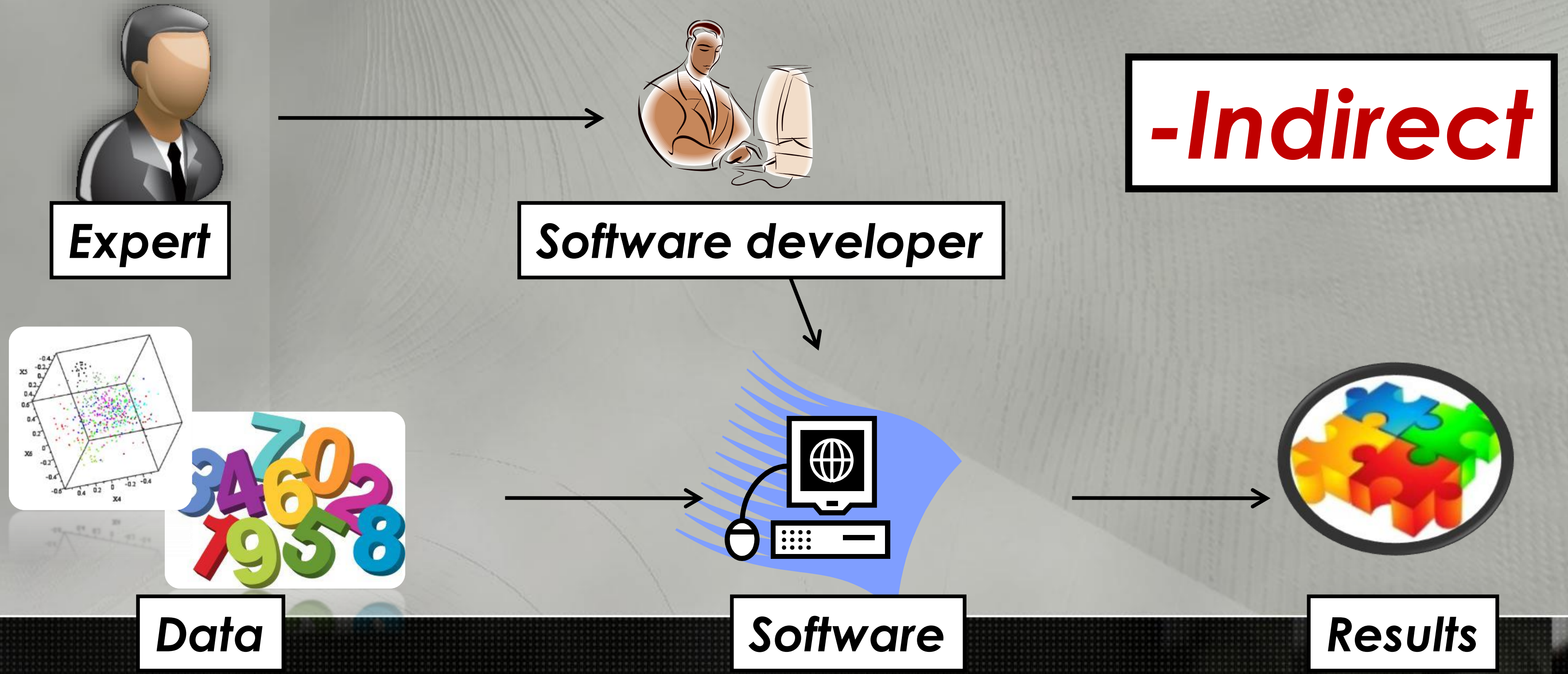


Data

Expert

-Slow

-Subjective



Expert

Software developer

-Indirect

Data

Software

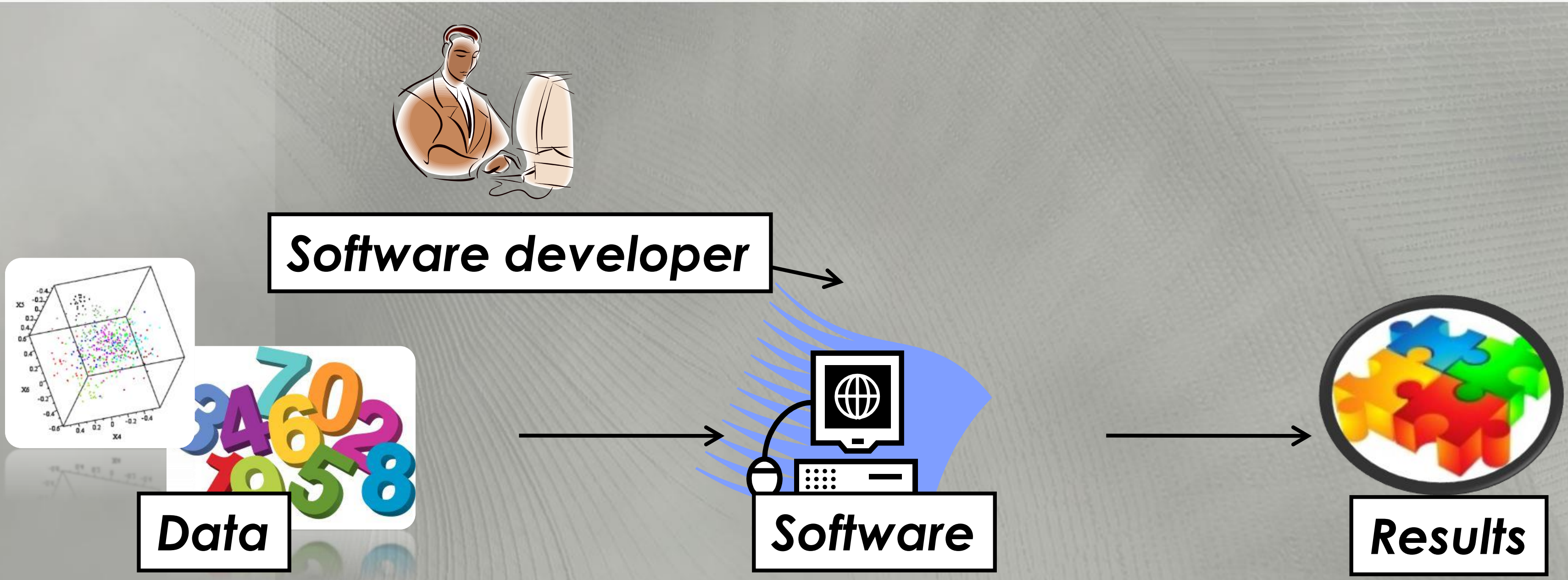
Results

Intelligent data processing/mining

Intelligent data processing/mining

Goal: Most effective combination and usage of available information and the knowledge of field expert

Intelligent data processing/mining



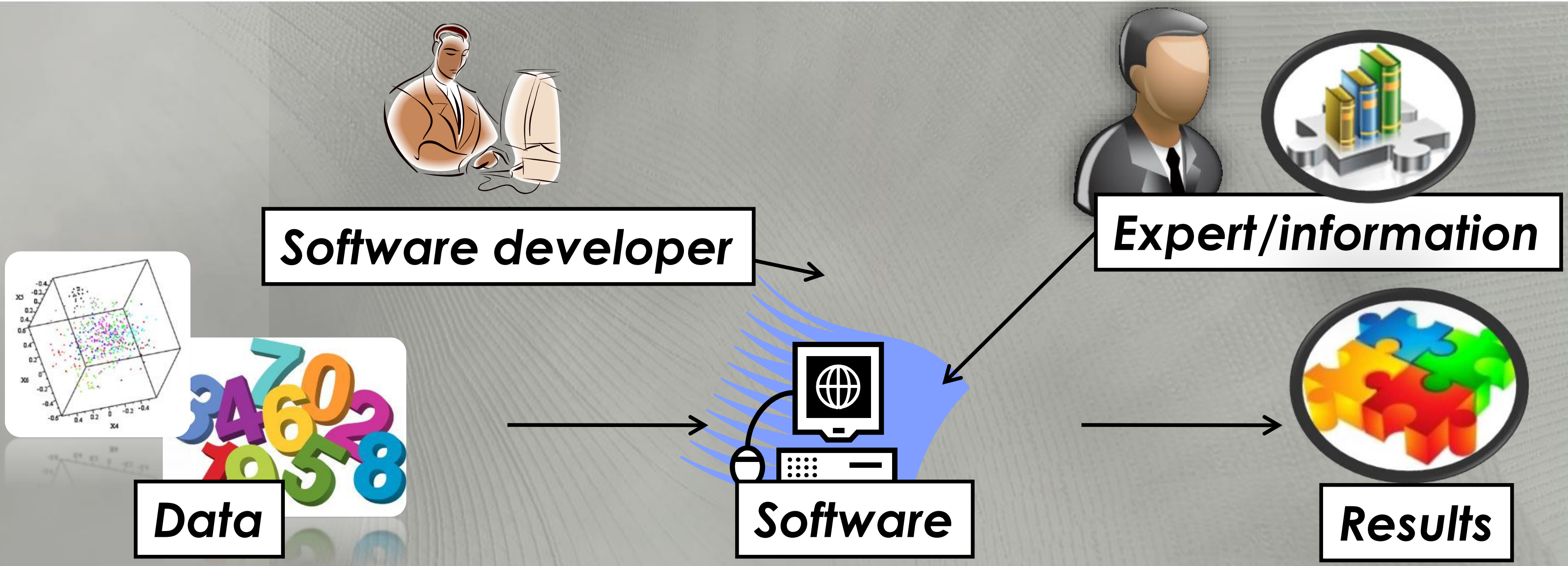
Software developer

Data

Software

Results

Intelligent data processing/mining



Software developer

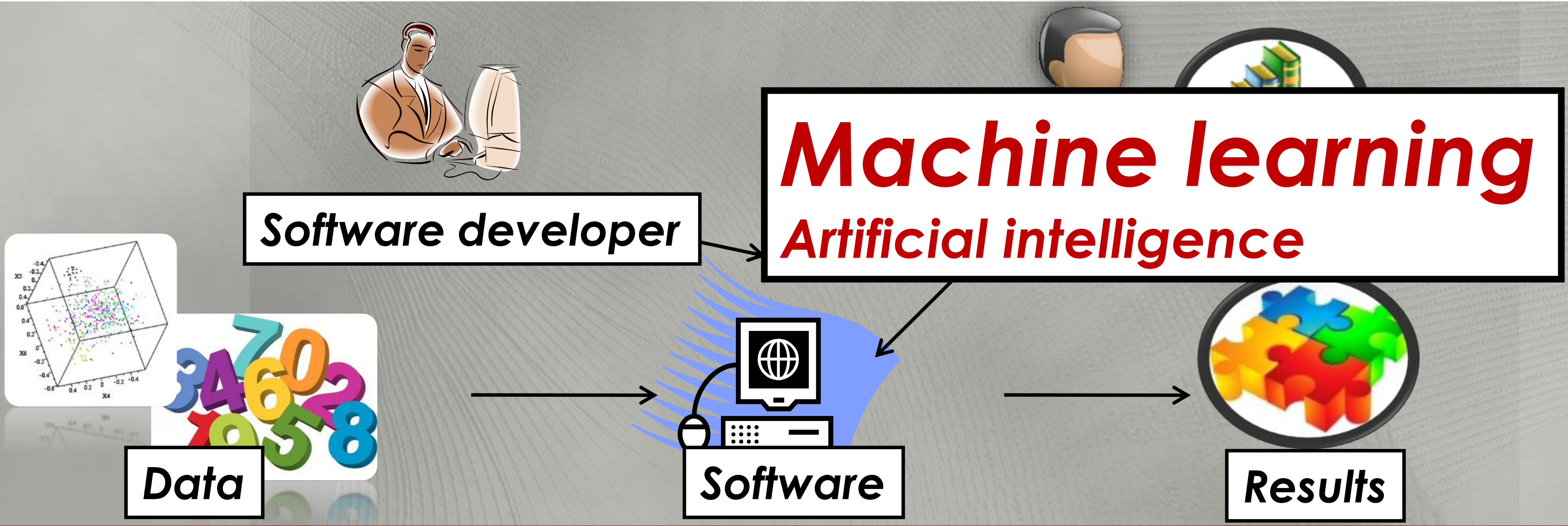
Expert/information

Data

Software

Results

Intelligent data processing/mining



Machine learning

Artificial intelligence

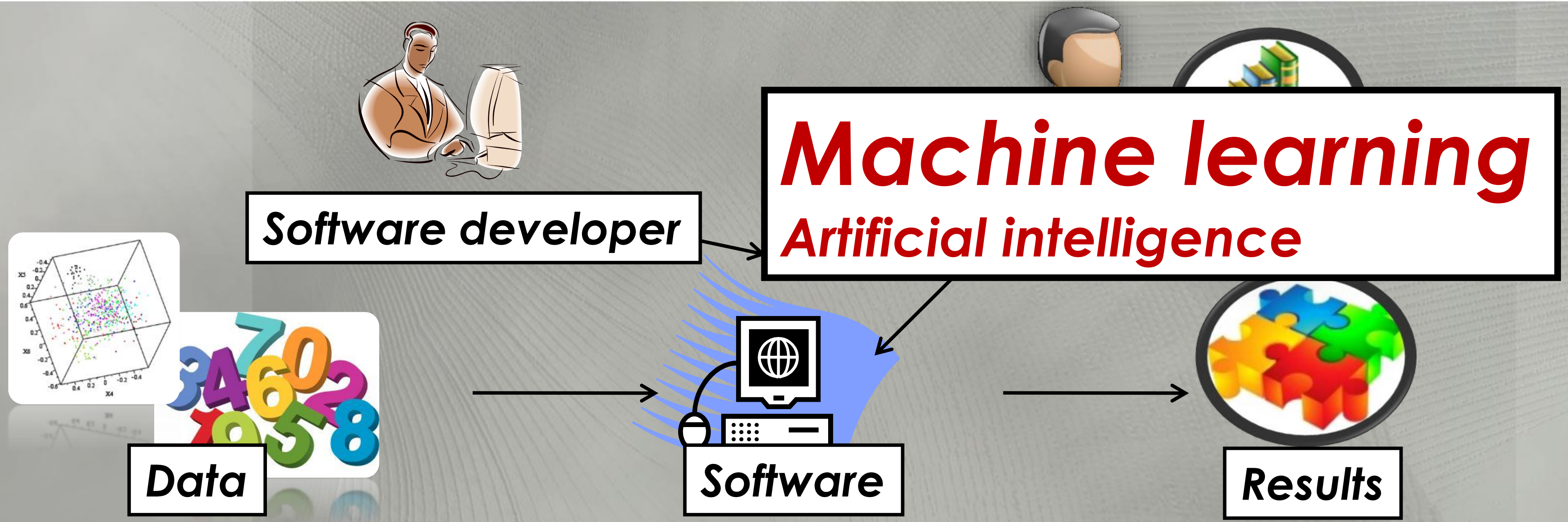
Software developer

Data

Software

Results

Intelligent data processing/mining



Software developer

Machine learning

Artificial intelligence

Data

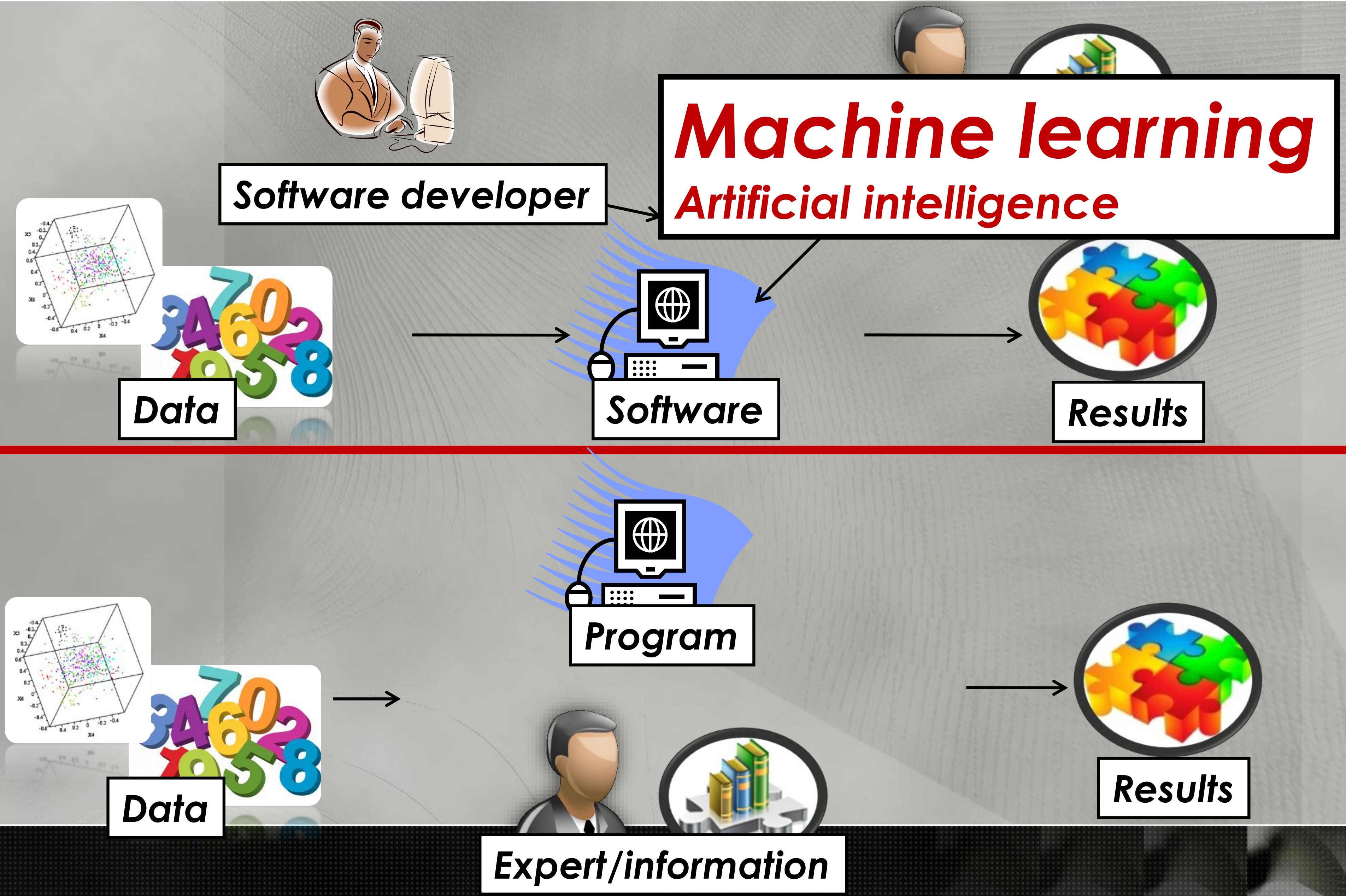
Software

Results

Data

Results

Intelligent data processing/mining



Software developer

Machine learning

Artificial intelligence

Data

Software

Results

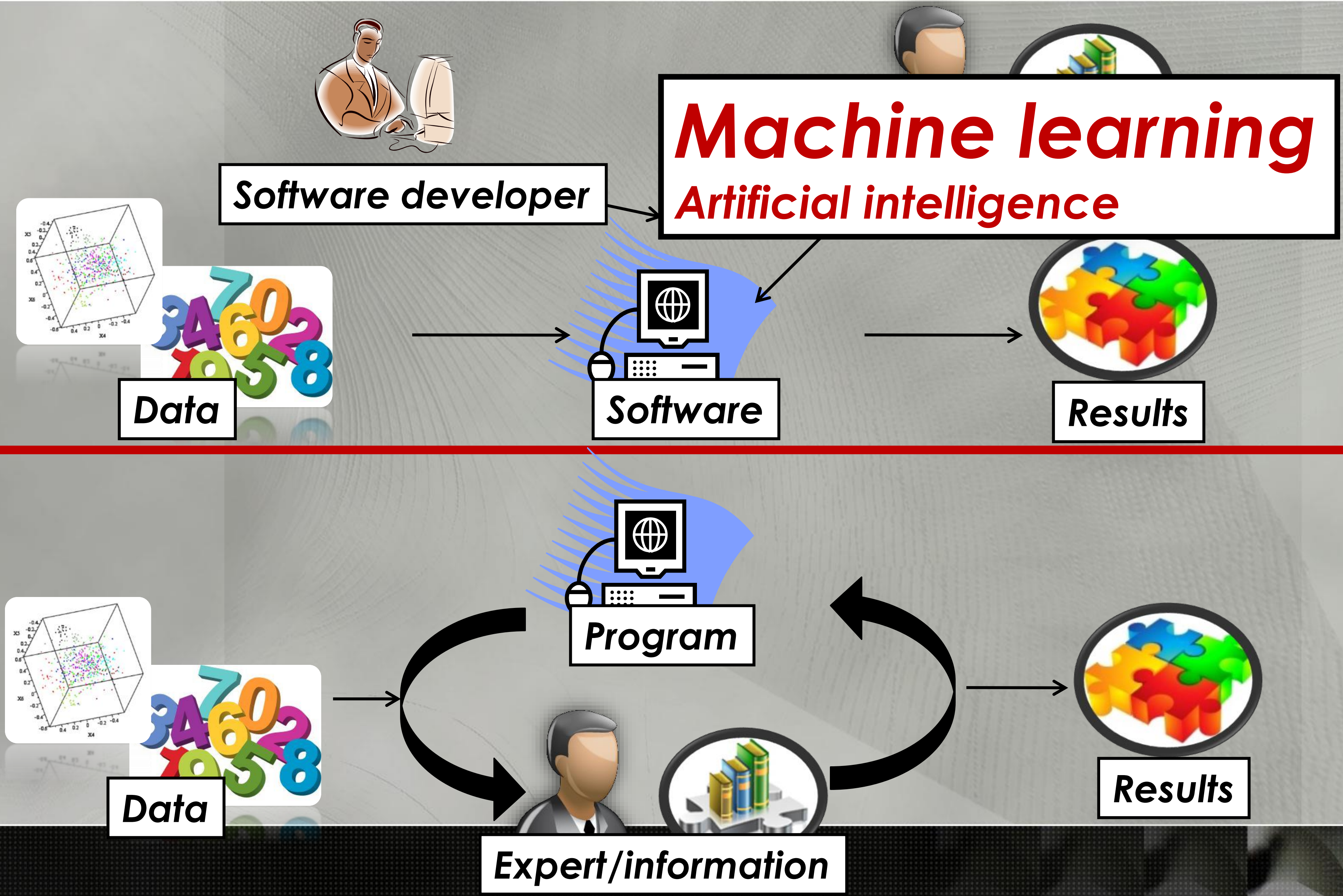
Program

Data

Results

Expert/information

Intelligent data processing/mining



Software developer

Machine learning

Artificial intelligence

Data

Software

Results

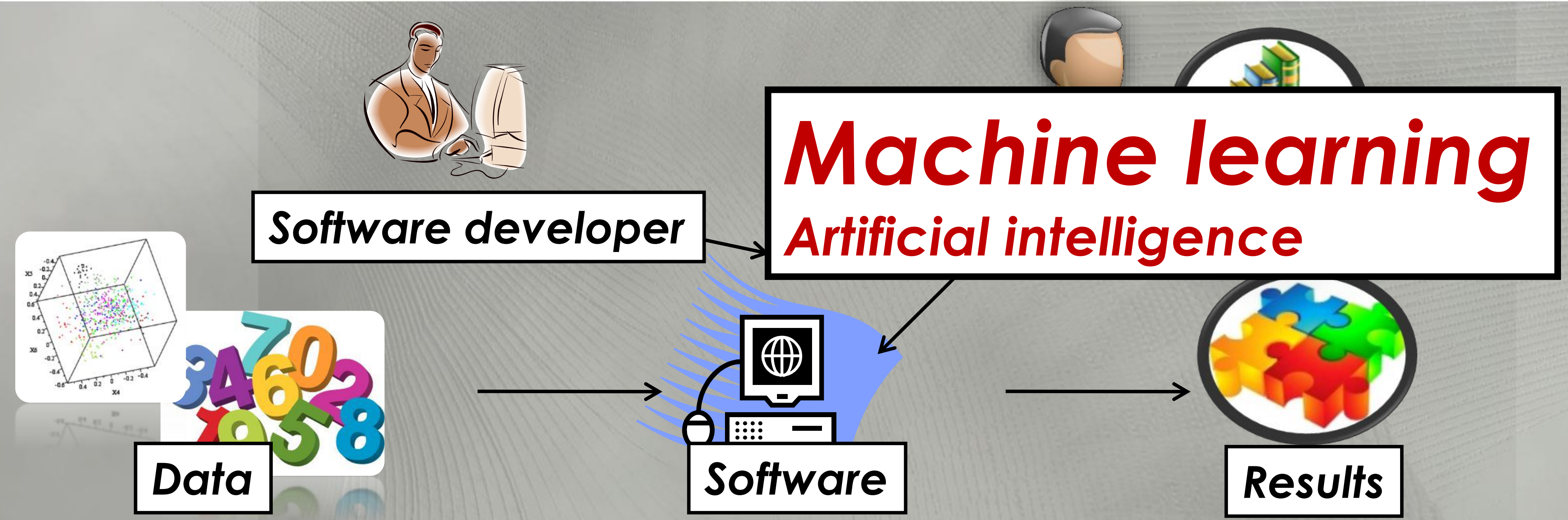
Program

Data

Expert/information

Results

Intelligent data processing/mining



Software developer

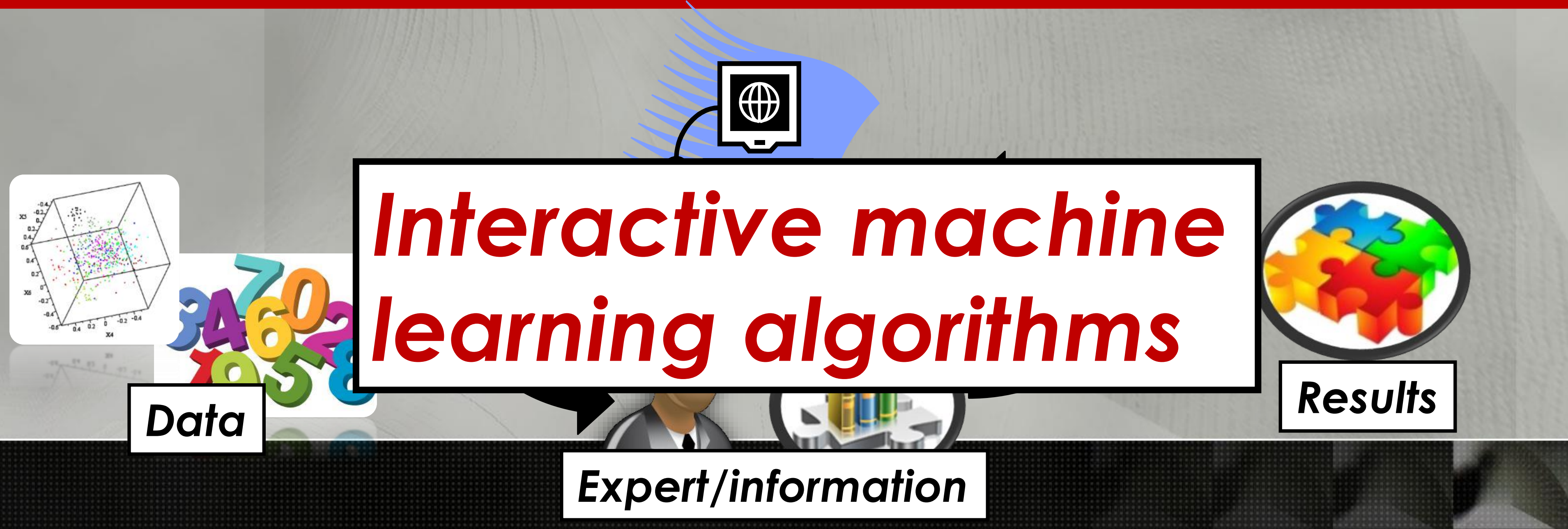
Machine learning

Artificial intelligence

Data

Software

Results



Interactive machine learning algorithms

Data

Results

Expert/information

TWO MAJOR QUESTIONS WE FORGOT TO ASK!

Is my analysis the most accurate?

Did I entirely discover my data? **Or partially?**

Active learning for HCS

Active learning for HCS

Aim

Active learning for HCS

Aim

- More precise and faster analysis

Active learning for HCS

Aim

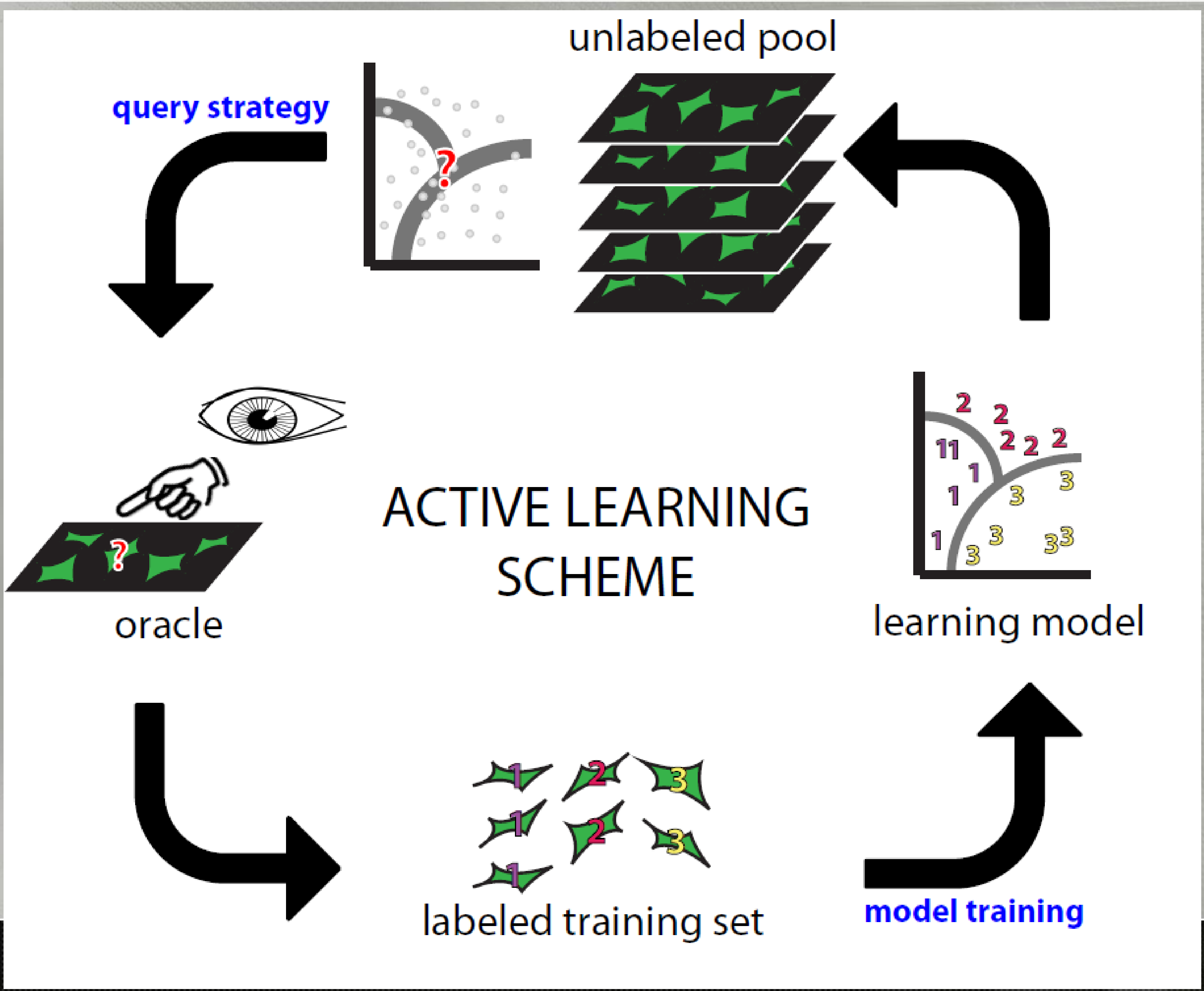
- More precise and faster analysis
- Using less human resources

Active learning for HCS

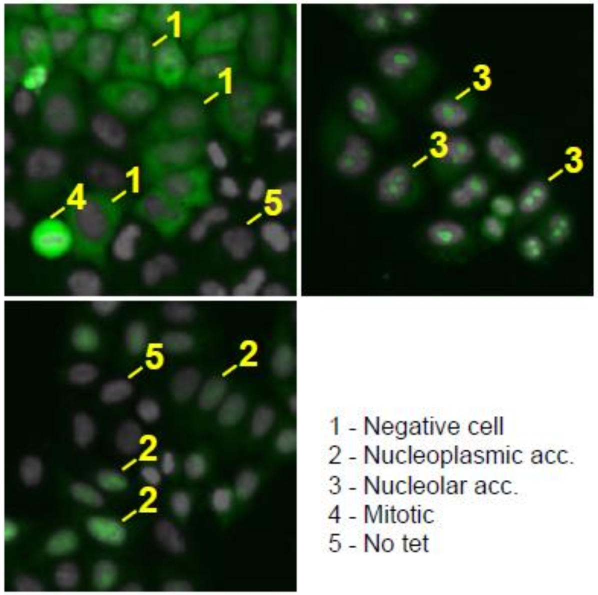
Aim

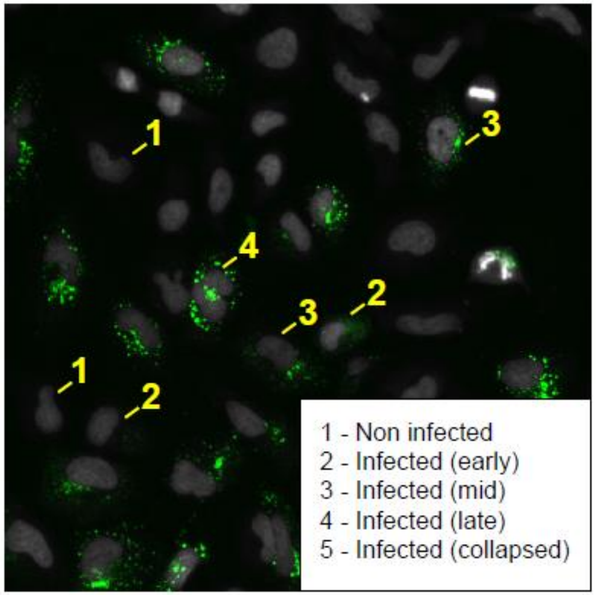
- More precise and faster analysis
- Using less human resources
- Asking more intelligent questions

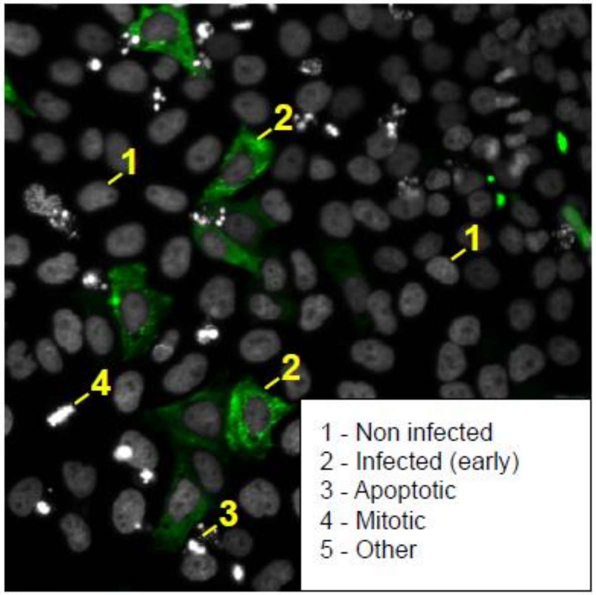
Active learning for HCS



Results

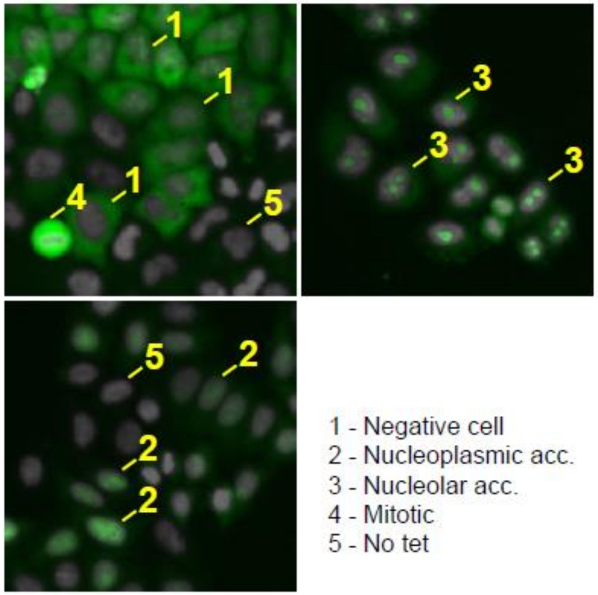




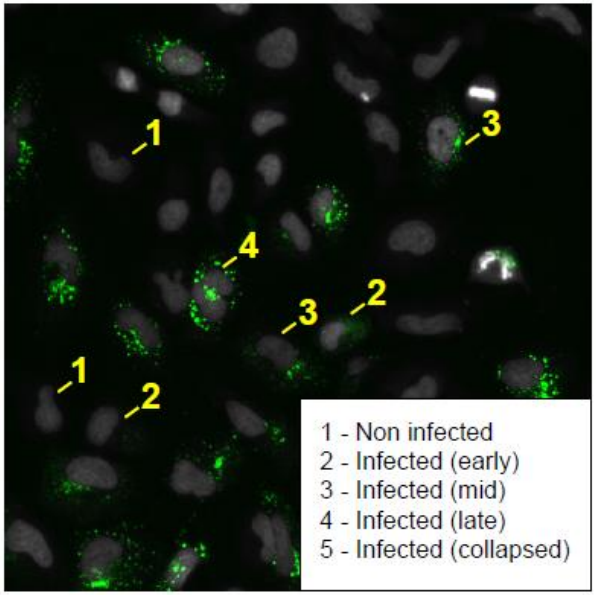


Smith K, Horvath P: Active Learning Strategies for Phenotypic Profiling of High-Content Screens, Journal of Biomolecular Screening, 2014.

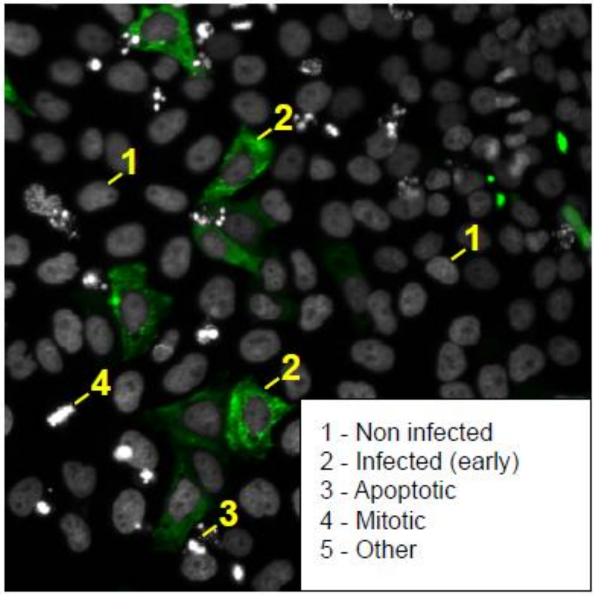
Results



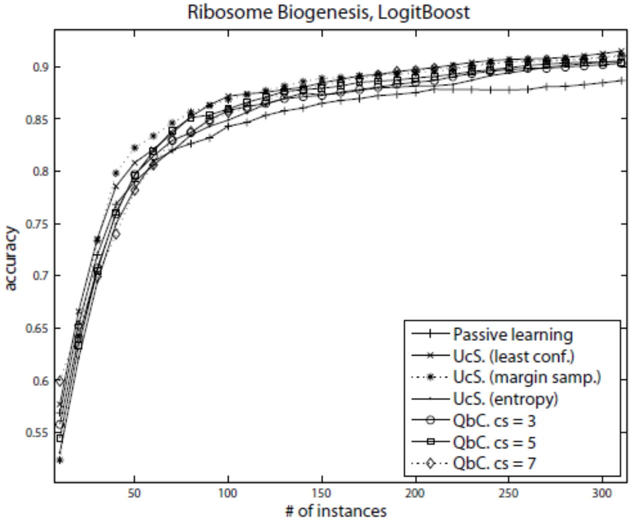
1 - Negative cell
2 - Nucleoplasmic acc.
3 - Nucleolar acc.
4 - Mitotic
5 - No tet

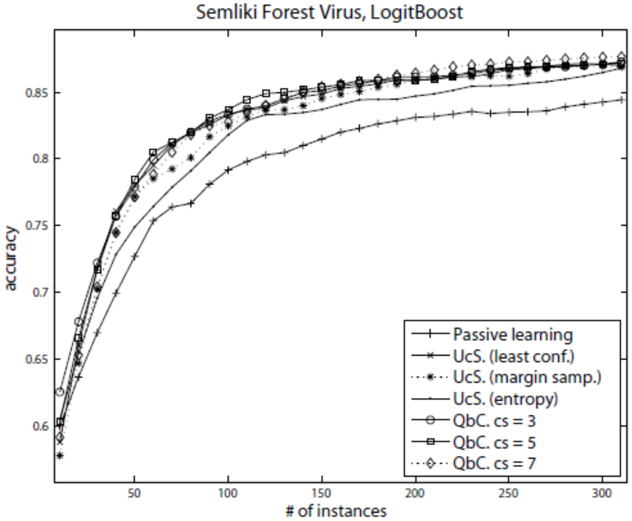


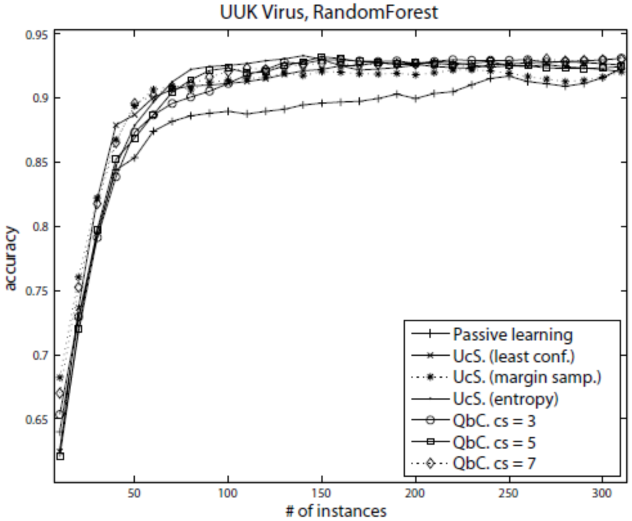
1 - Non infected
2 - Infected (early)
3 - Infected (mid)
4 - Infected (late)
5 - Infected (collapsed)



1 - Non infected
2 - Infected (early)
3 - Apoptotic
4 - Mitotic
5 - Other

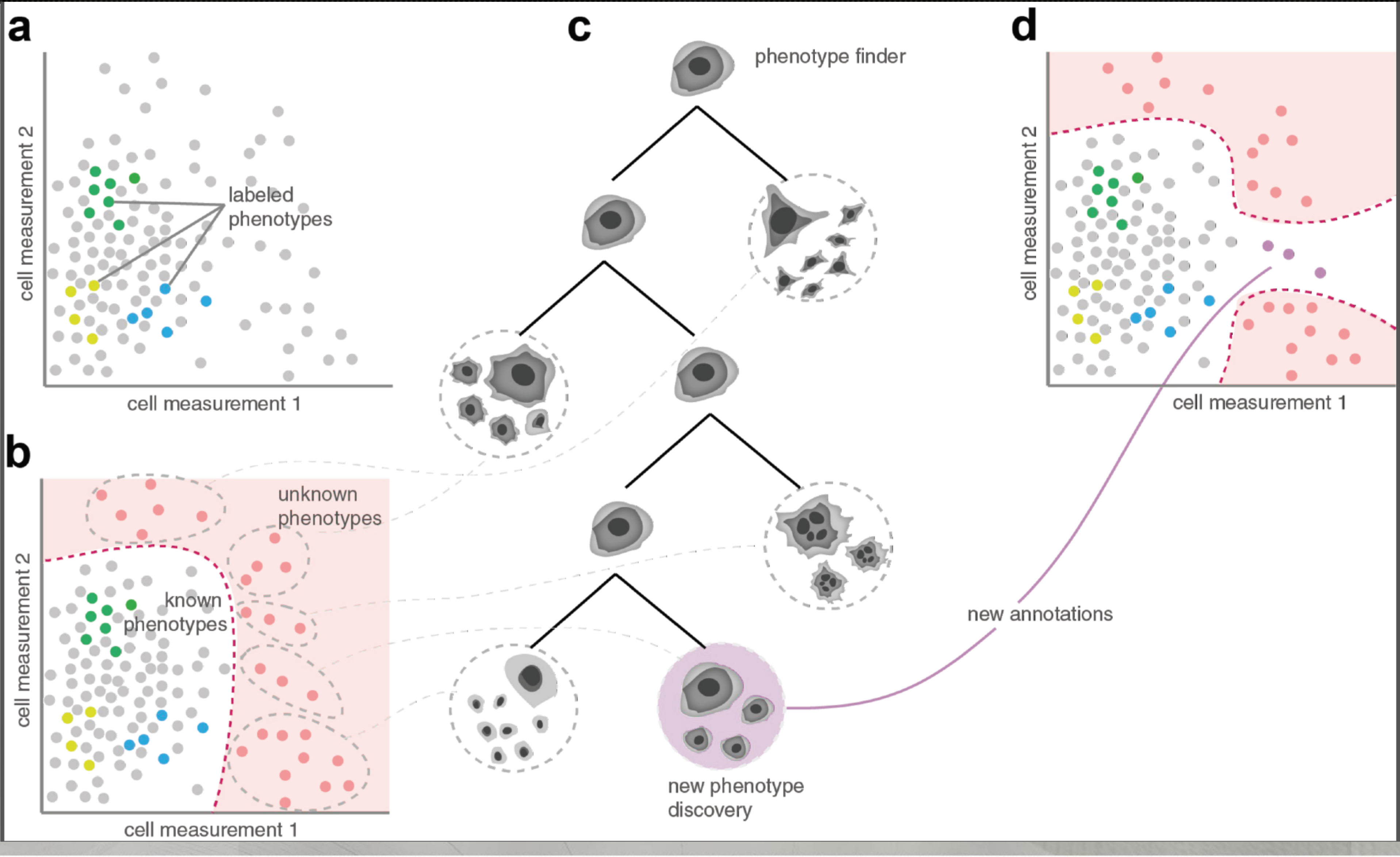

Ribosome Biogenesis, LogitBoost


Semliki Forest Virus, LogitBoost


UUK Virus, RandomForest

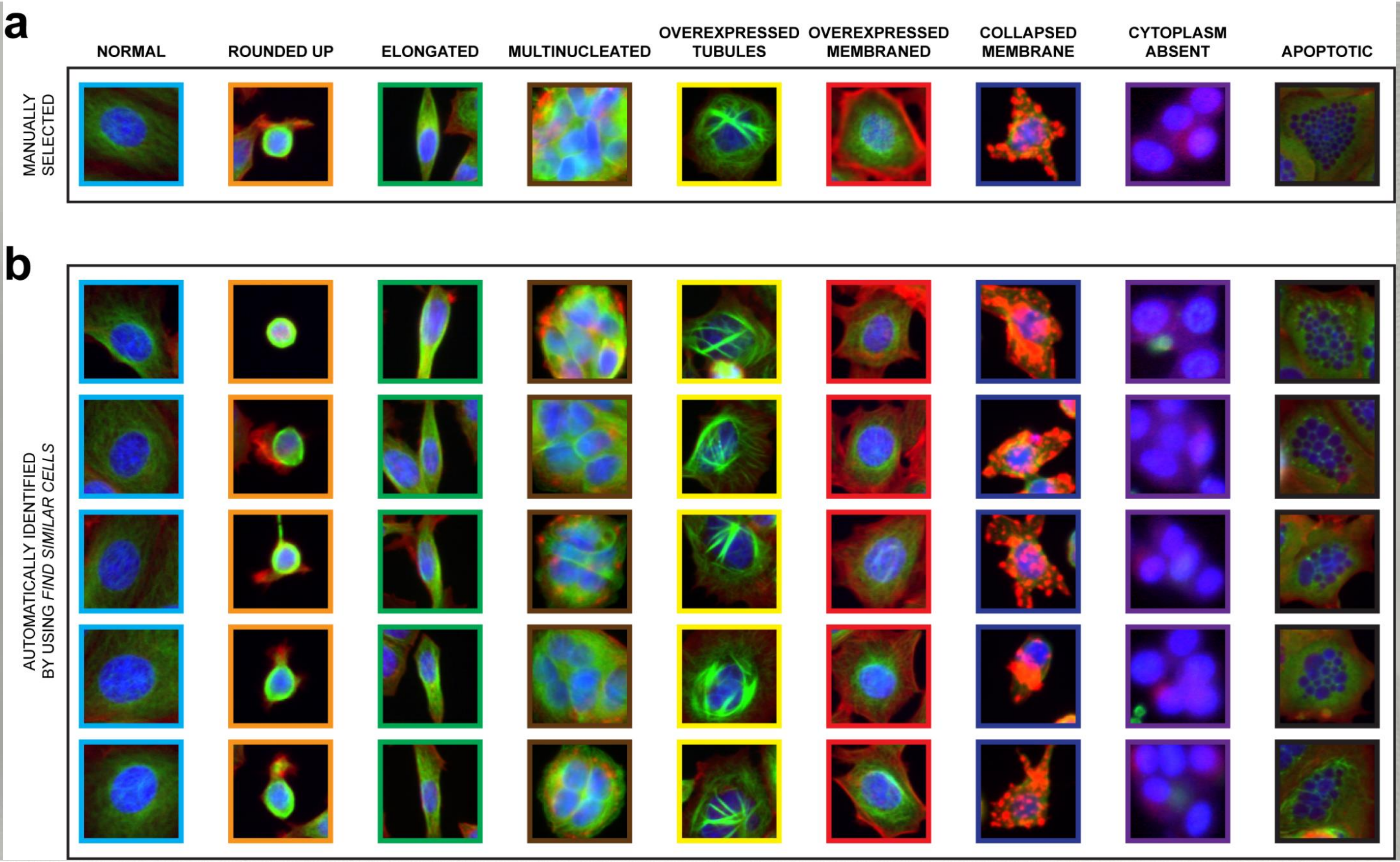
Smith K, Horvath P: Active Learning Strategies for Phenotypic Profiling of High-Content Screens, Journal of Biomolecular Screening, 2014.

Phenotype finder

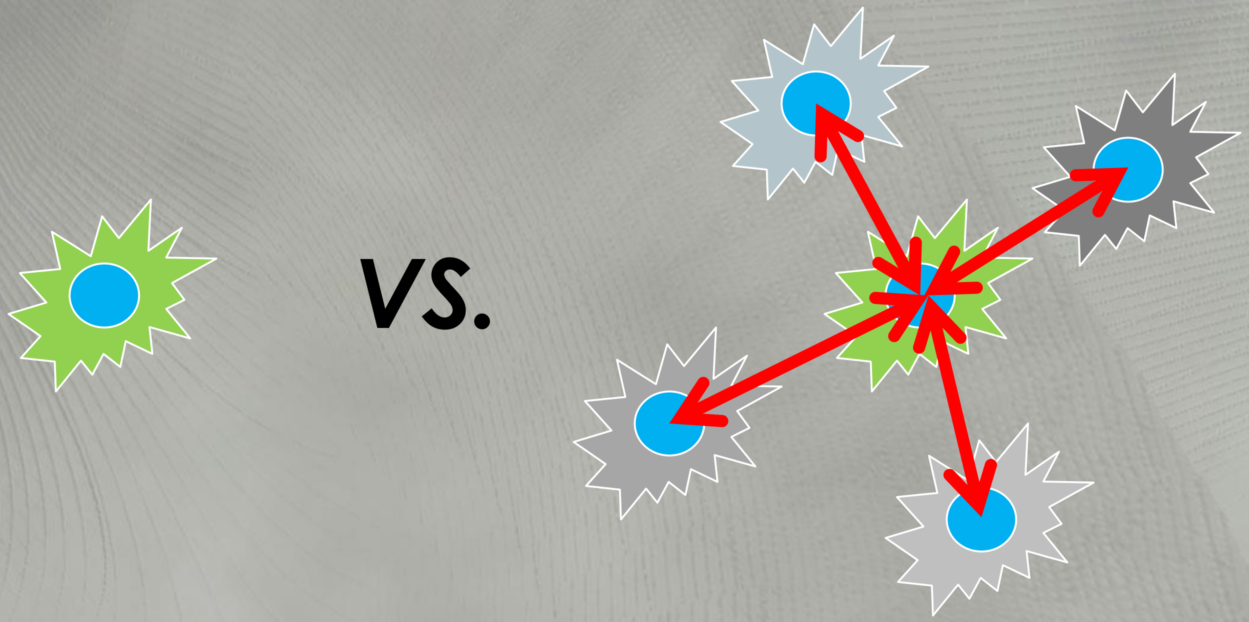


Ficcinini F. and T. Balassa, *etal:. Cell Systems 2017*

Results



Caie etal. High-Content Phenotypic Profiling of Drug Response Signatures across Distinct Cancer Cells

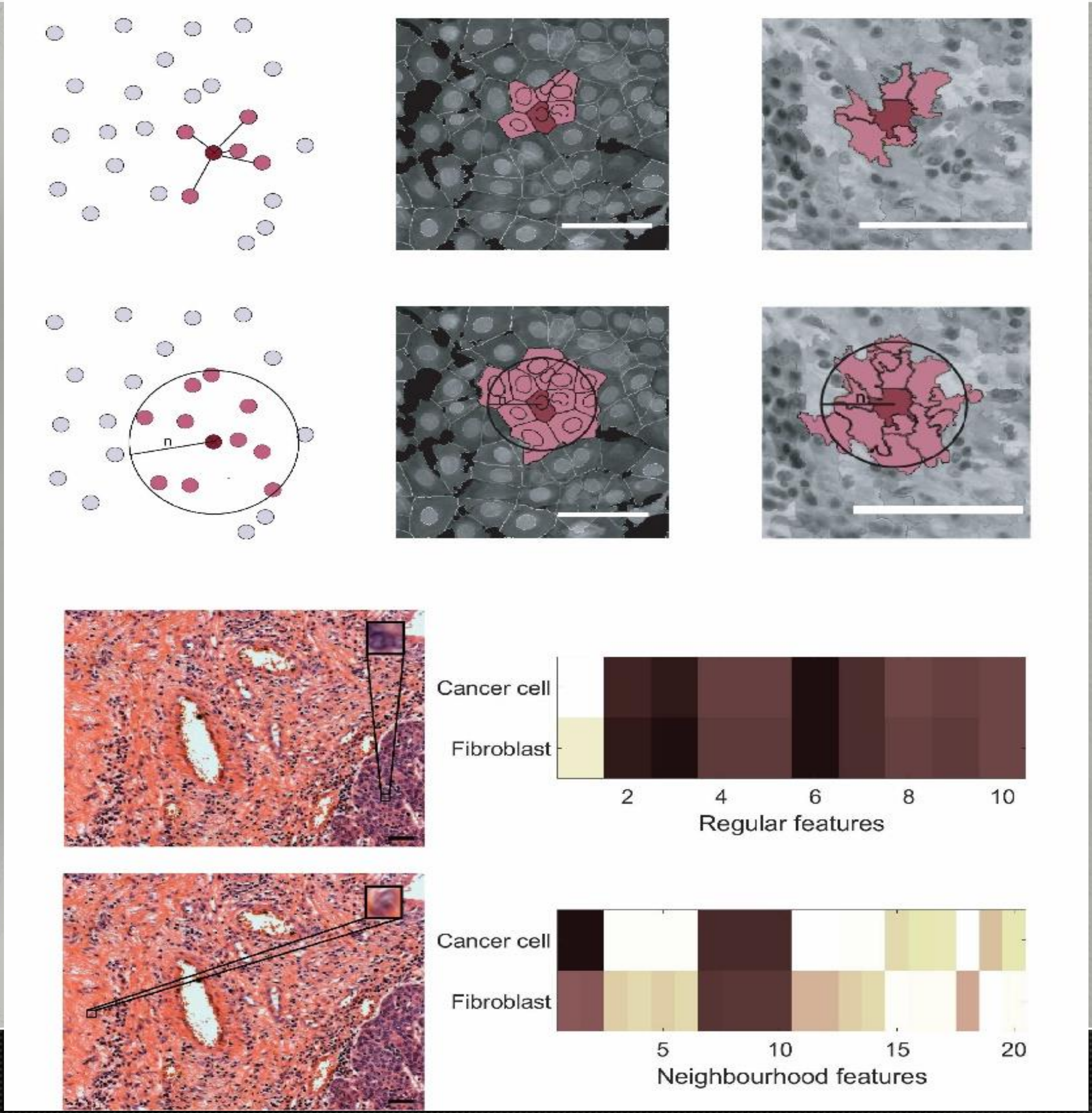


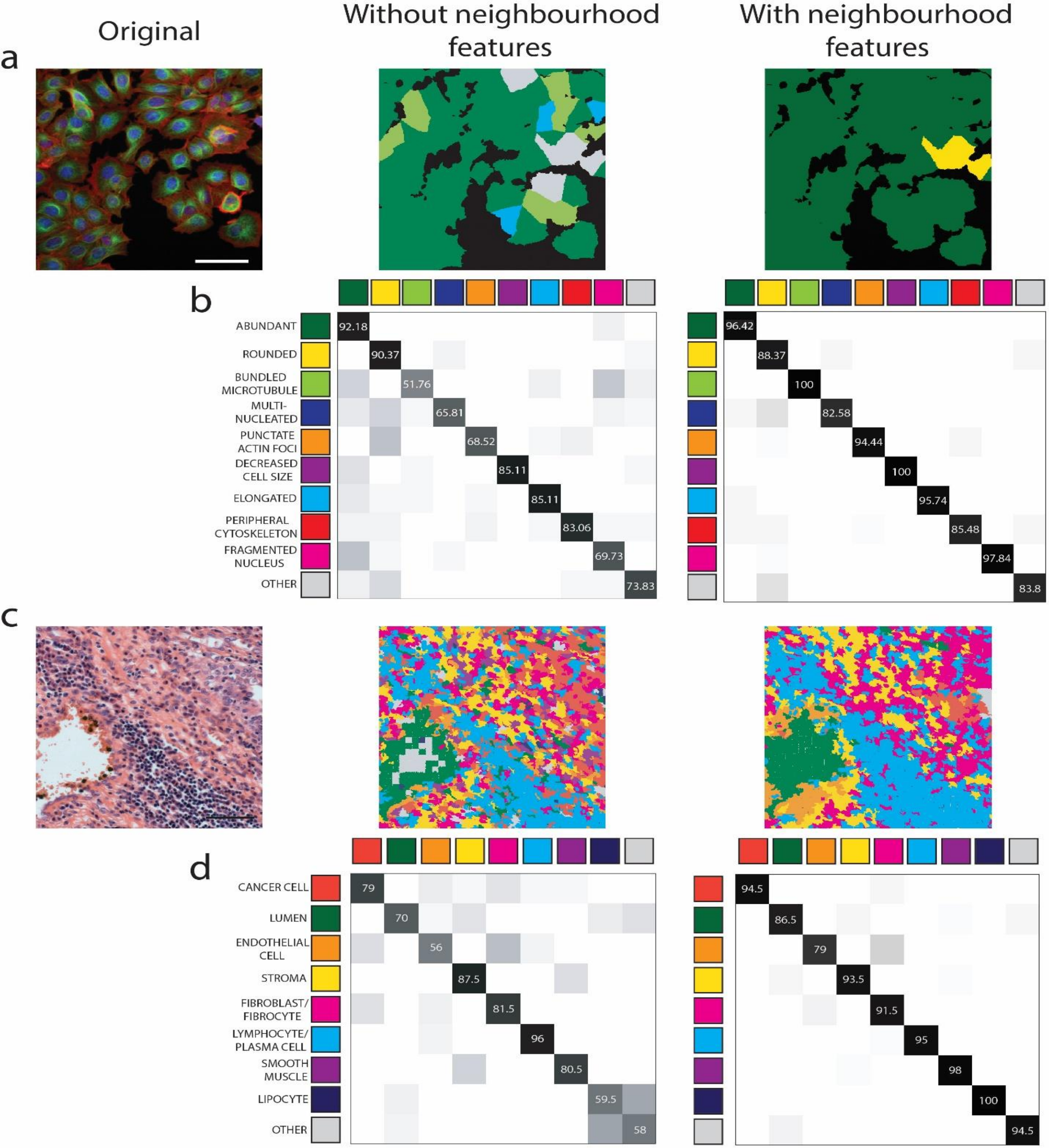
VS.

TELL ME YOUR NEIGHBOR I TELL WHAT YOU ARE

Toth etal. *Environmental properties of cells improve machine learning-based phenotype recognition accuracy,* ***2018. Sci. Rep.***

Major concept



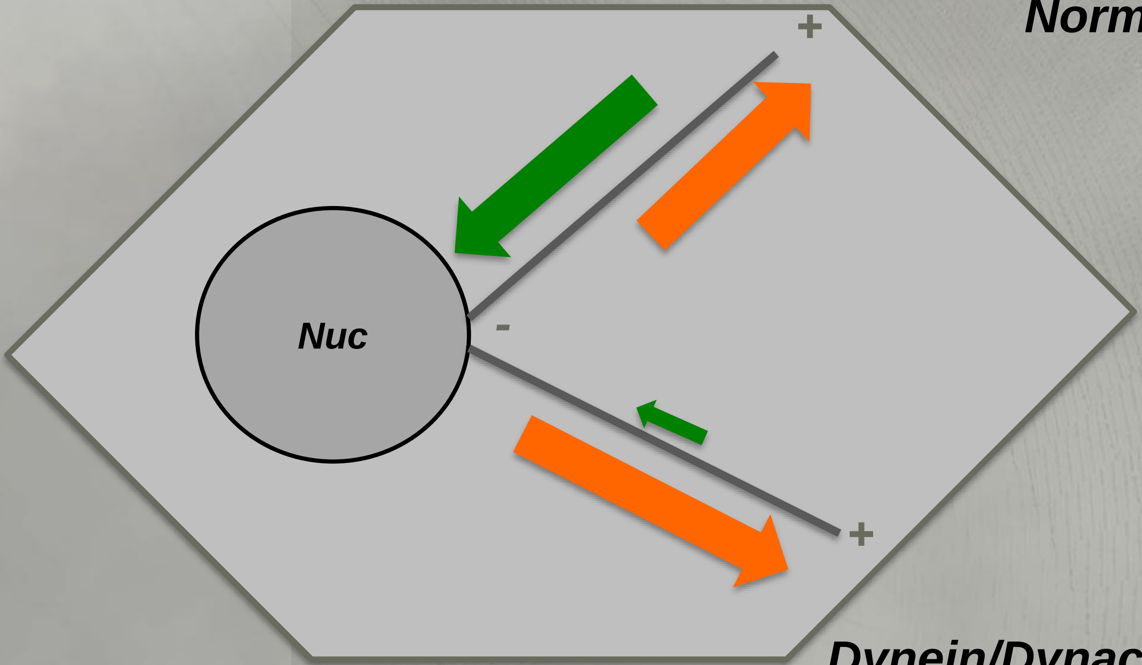

Original
Without neighbourhood features
With neighbourhood features
a
b
ABUNDANT
ROUNDED
BUNDLED MICROTUBULE
MULTI-NUCLEATED
PUNCTATE ACTIN FOCI
DECREASED CELL SIZE
ELONGATED
PERIPHERAL CYTOSKELETON
FRAGMENTED NUCLEUS
OTHER
92.18
90.37
51.76
65.81
68.52
85.11
85.11
83.06
69.73
73.83
96.42
88.37
100
82.58
94.44
100
95.74
85.48
97.84
83.8
c
d
CANCER CELL
LUMEN
ENDOTHELIAL CELL
STROMA
FIBROBLAST/FIBROCYTE
LYMPHOCYTE/PLASMA CELL
SMOOTH MUSCLE
LIPOCYTE
OTHER
79
70
56
87.5
81.5
96
80.5
59.5
58
94.5
86.5
79
93.5
91.5
95
98
100
94.5

Beyond single cell-based discrete decisions

REGRESSION MODELS FOR HCA

Localization of Late Endosome/Lysosomes

Helenius lab; ETH Zurich



Normal

Dynein/Dynactin kd

Lack of retrograde motility

Yamauchi, Boukari, Banerjee, Sbalzarini, Horvath, Helenius. Histone Deacetylase 8 is Required for Centrosome Cohesion and Influenza A Virus Entry. PLoS Pathogens. 2011.

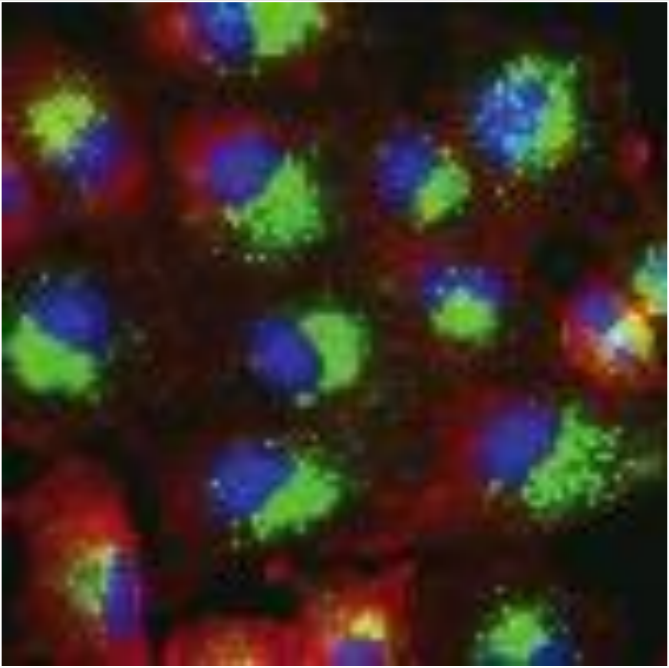
Localization of Late Endosome/Lysosomes

Helenius lab; ETH Zurich

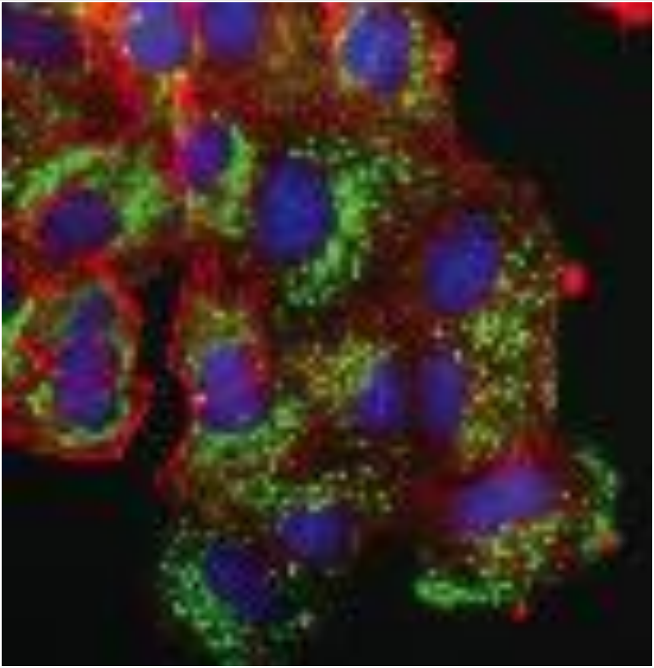
PACKED

SCATTERED

Control



Dynactin kd



Yamauchi, Boukari, Banerjee, Sbalzarini, Horvath, Helenius. Histone Deacetylase 8 is Required for Centrosome Cohesion and Influenza A Virus Entry. PLoS Pathogens. 2011.

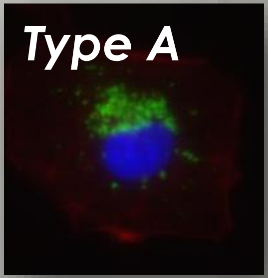
Regression models for HCA

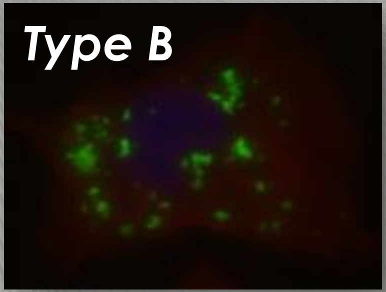
Regression models for HCA

- **Problem:** Discrete decisions

Regression models for HCA

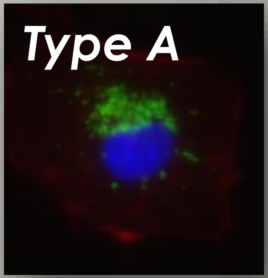
- **Problem:** Discrete decisions

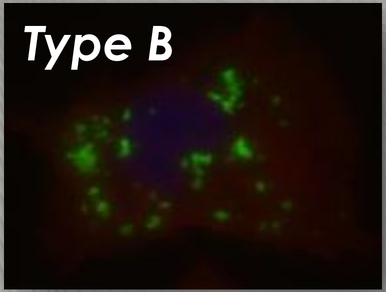




Regression models for HCA

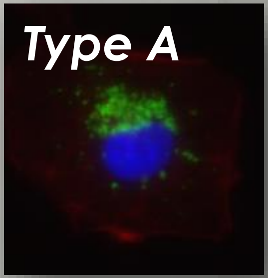
- Problem: Discrete decisions

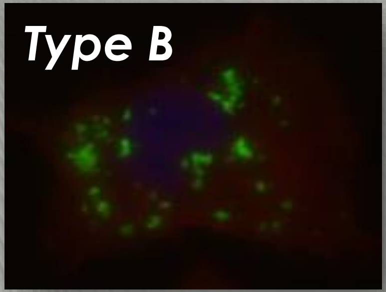




Regression models for HCA

- **Problem:** Discrete decisions

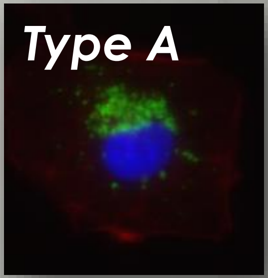


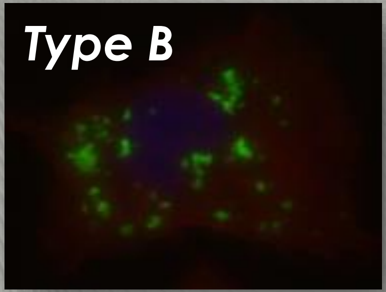


- **Solution:** Comparison

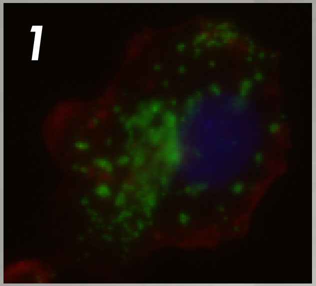
Regression models for HCA

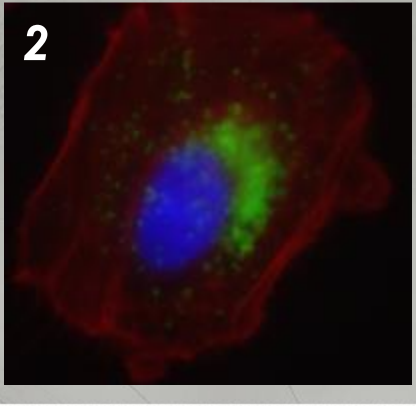
- **Problem:** Discrete decisions

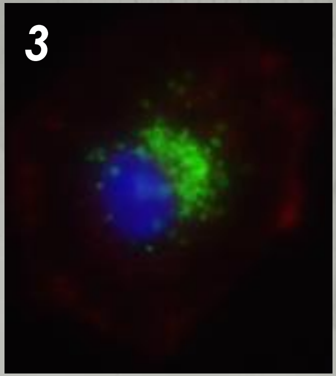


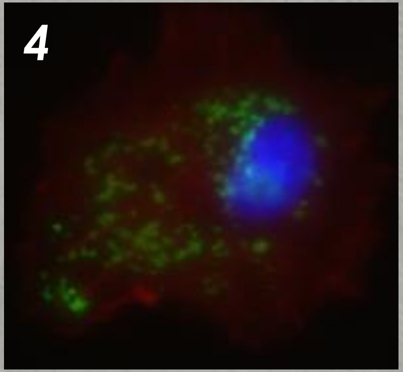


- **Solution:** Comparison



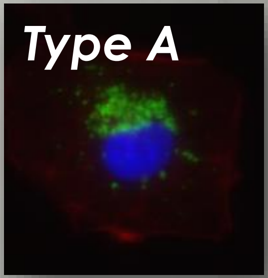


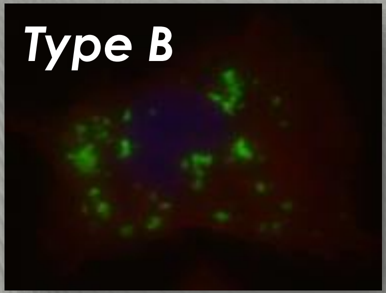




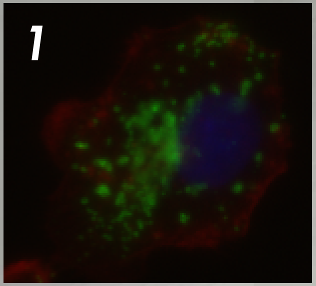
Regression models for HCA

- **Problem:** Discrete decisions





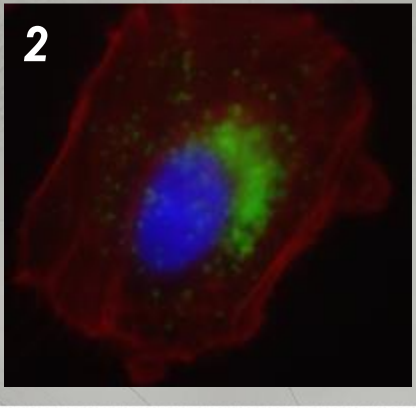
- **Solution:** Comparison



<

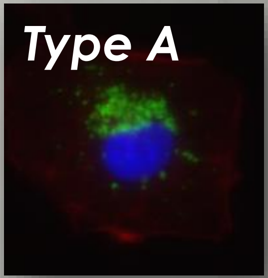
?

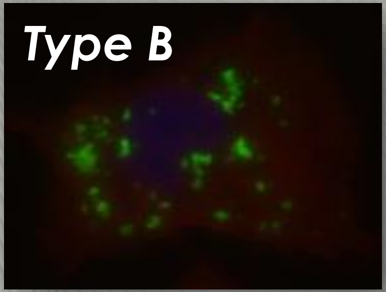
>



Regression models for HCA

- Problem: Discrete decisions

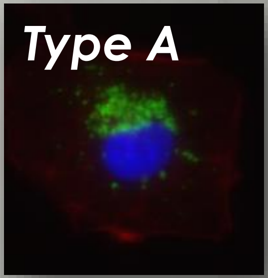


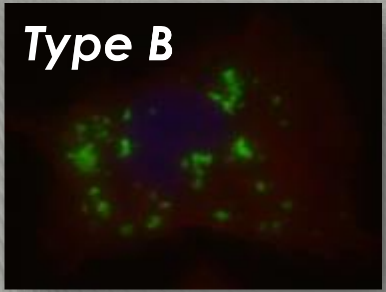


- Solution: Comparison

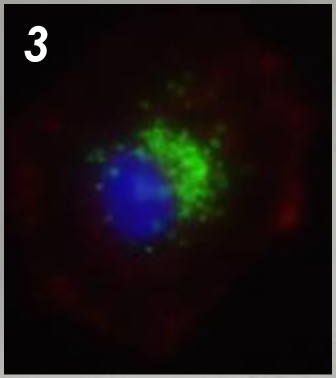
Regression models for HCA

- Problem: Discrete decisions

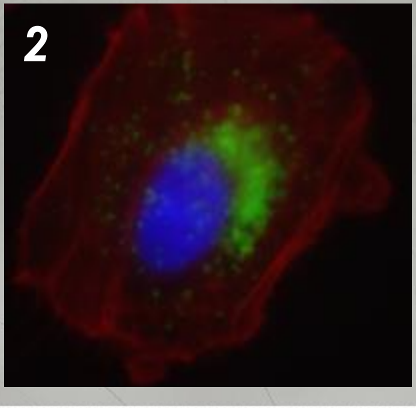




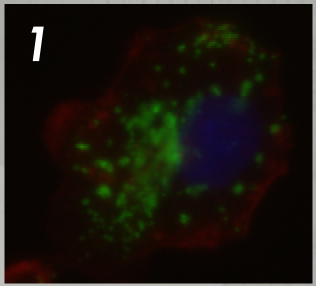
- Solution: Comparison



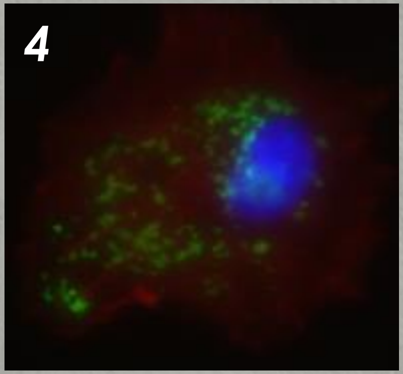
<



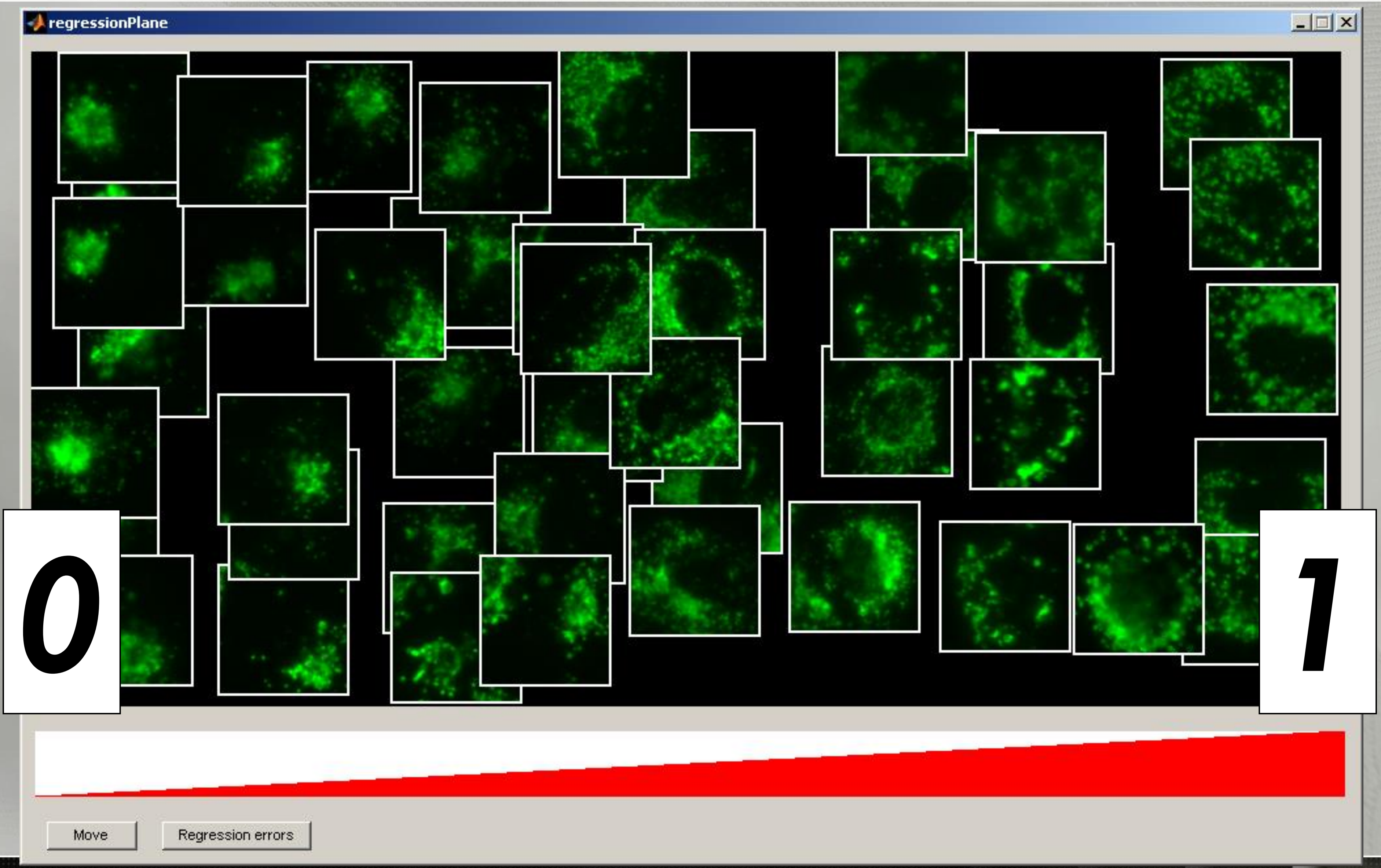
<

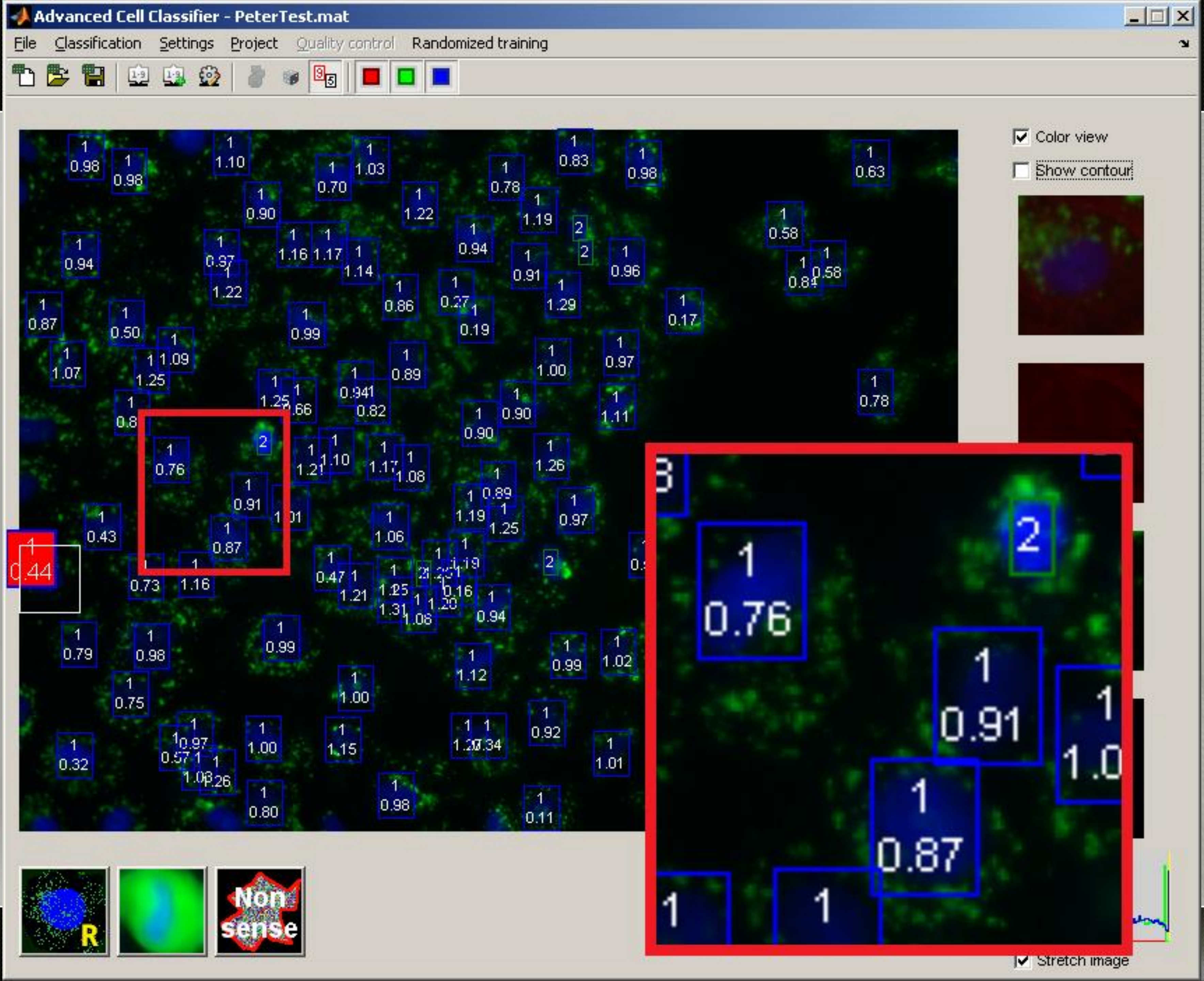


<

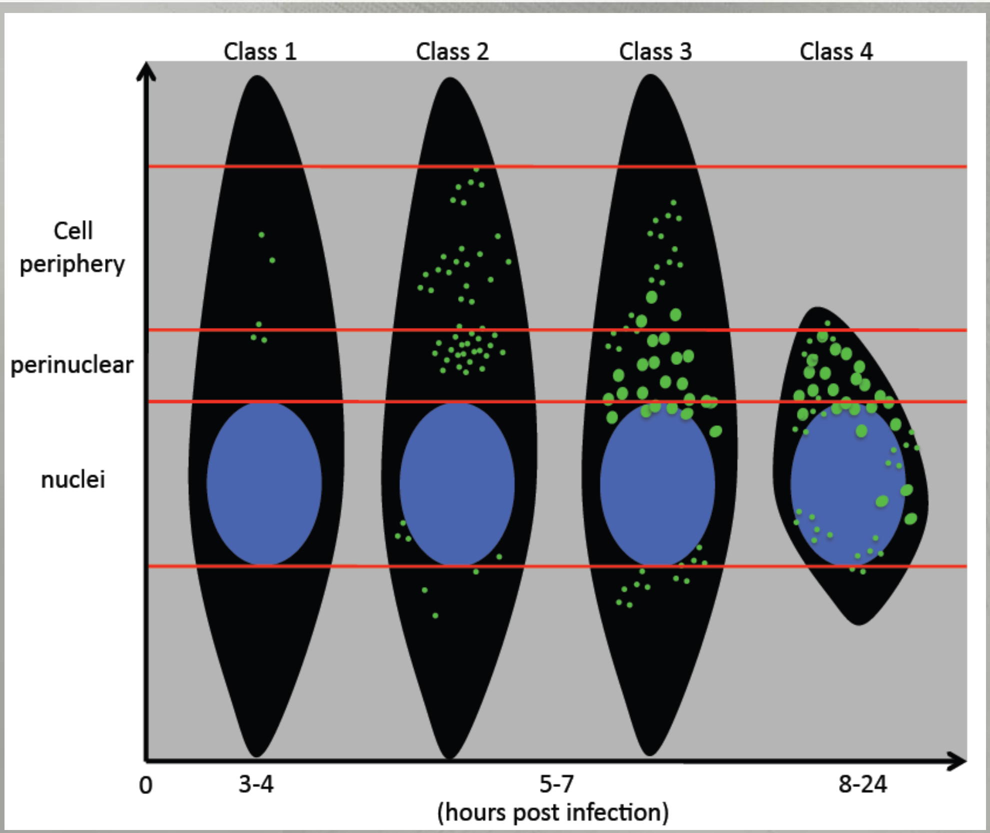


Regression plane concept



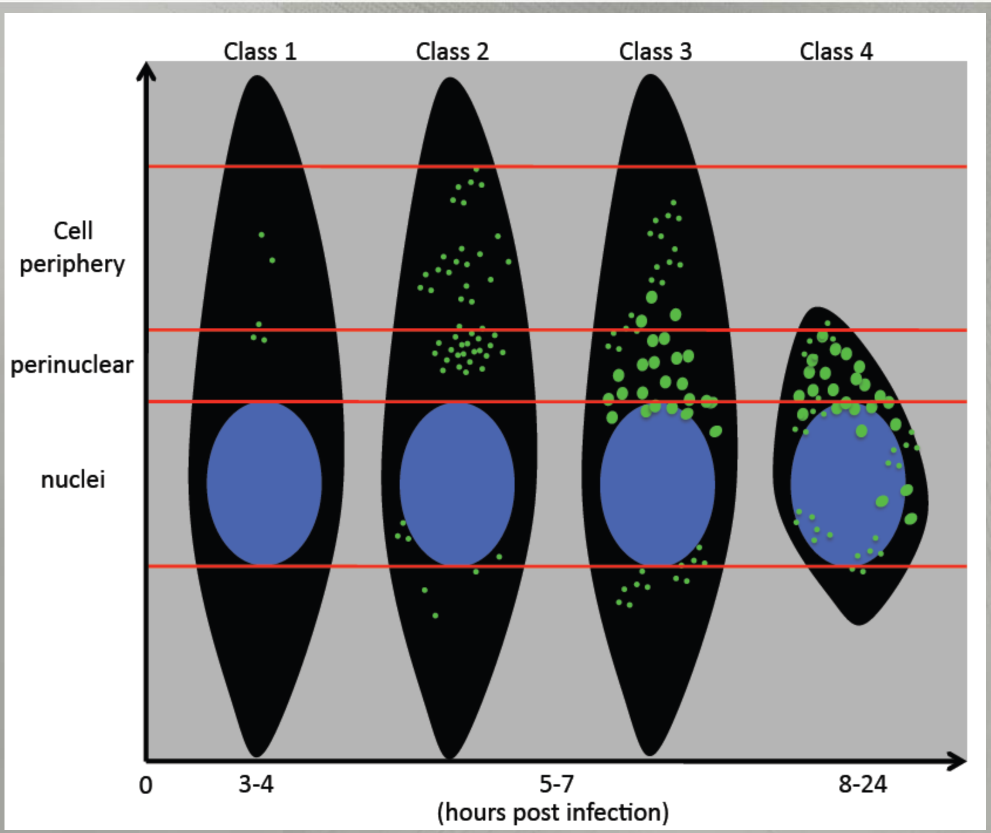

Advanced Cell Classifier - PeterTest.mat
File
Classification
Settings
Project
Quality control
Randomized training
☑ Color view
☐ Show contour
☑ Stretch image
Non sense
R
1
0.76
2
1
0.91
1
0.87
1
1.0
1
1
1
0.44

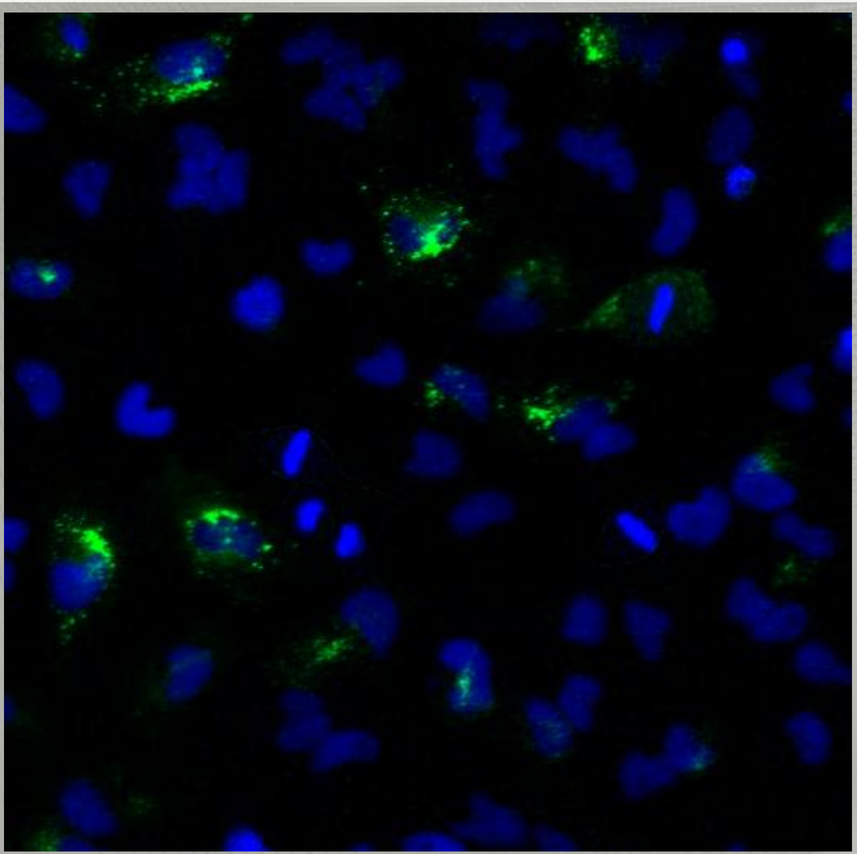
Further results - Semliki Forest virus genome-wide screen



Balistreri, Horvath, Helenius.
Cell Host and Microbe.

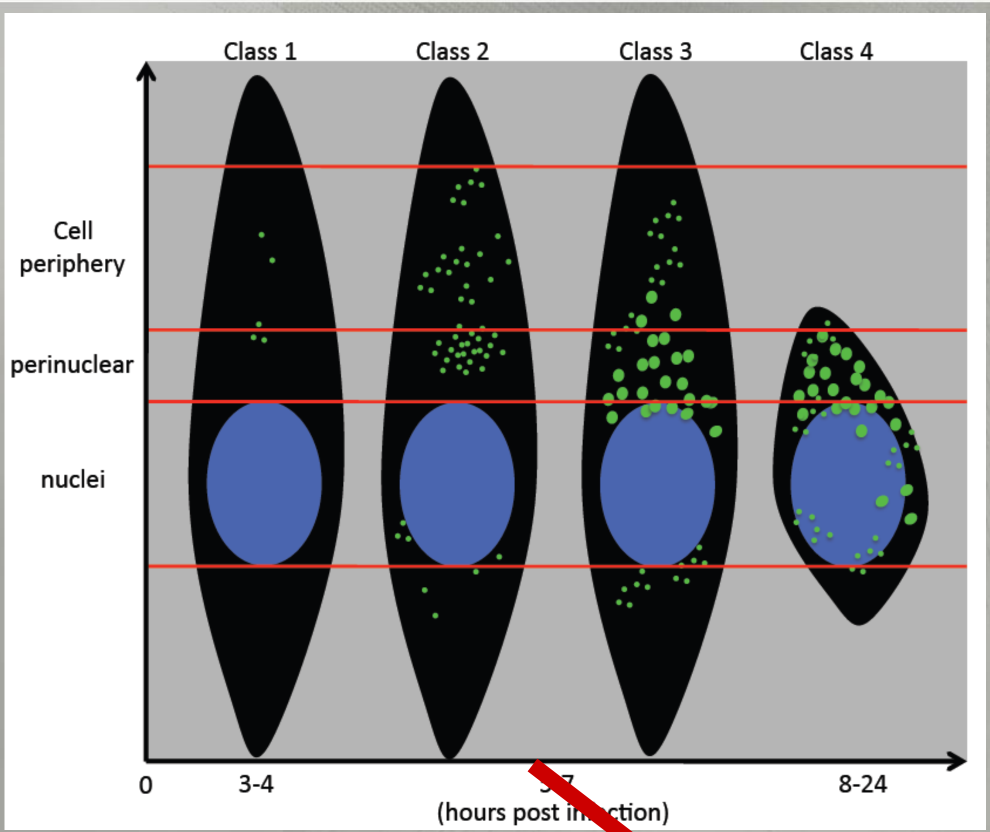
Further results - Semliki Forest virus genome-wide screen

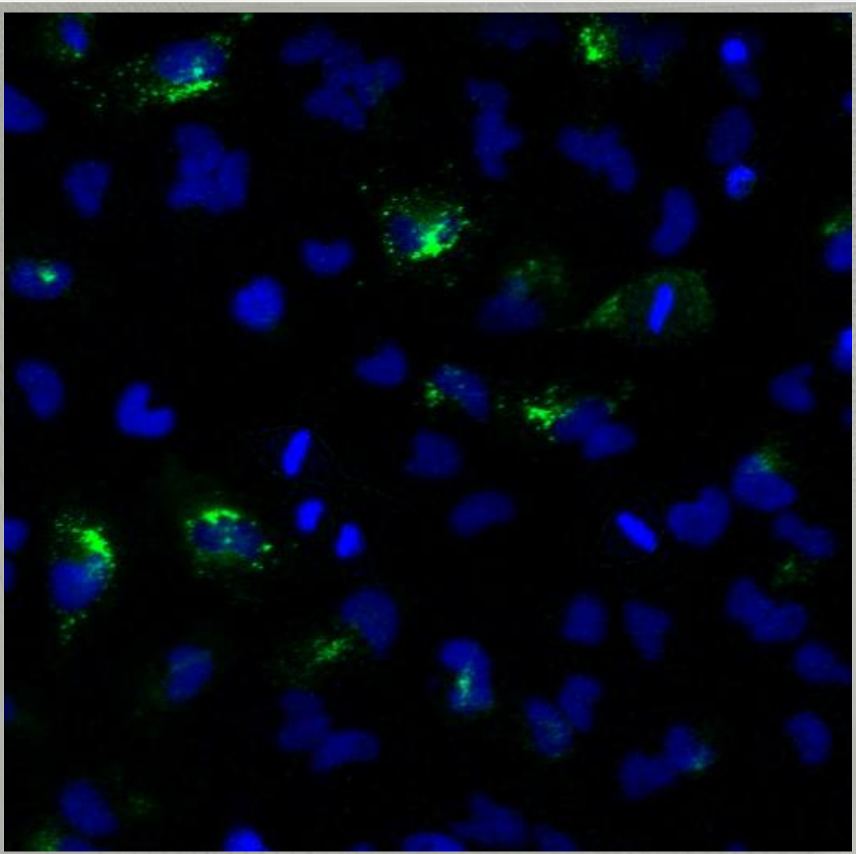


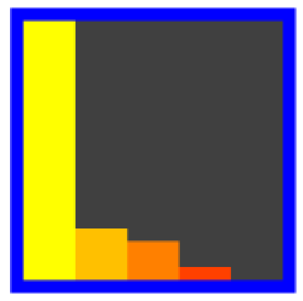


Balistreri, Horvath, Helenius.
Cell Host and Microbe.

Further results - Semliki Forest virus genome-wide screen

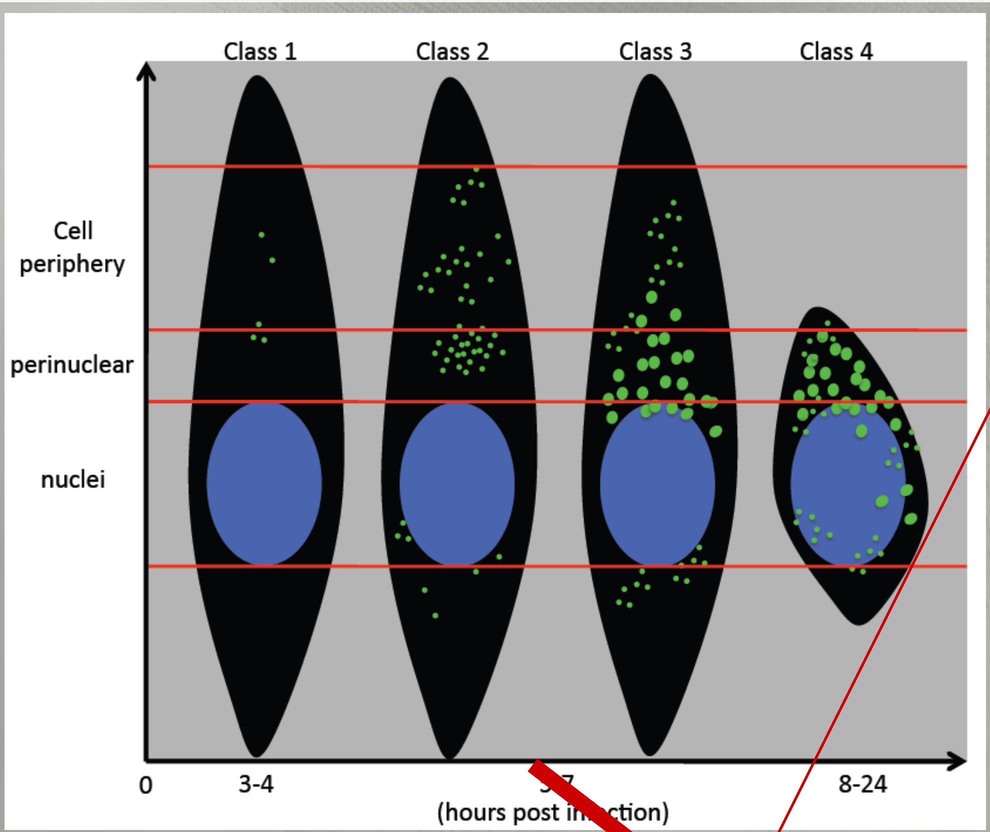


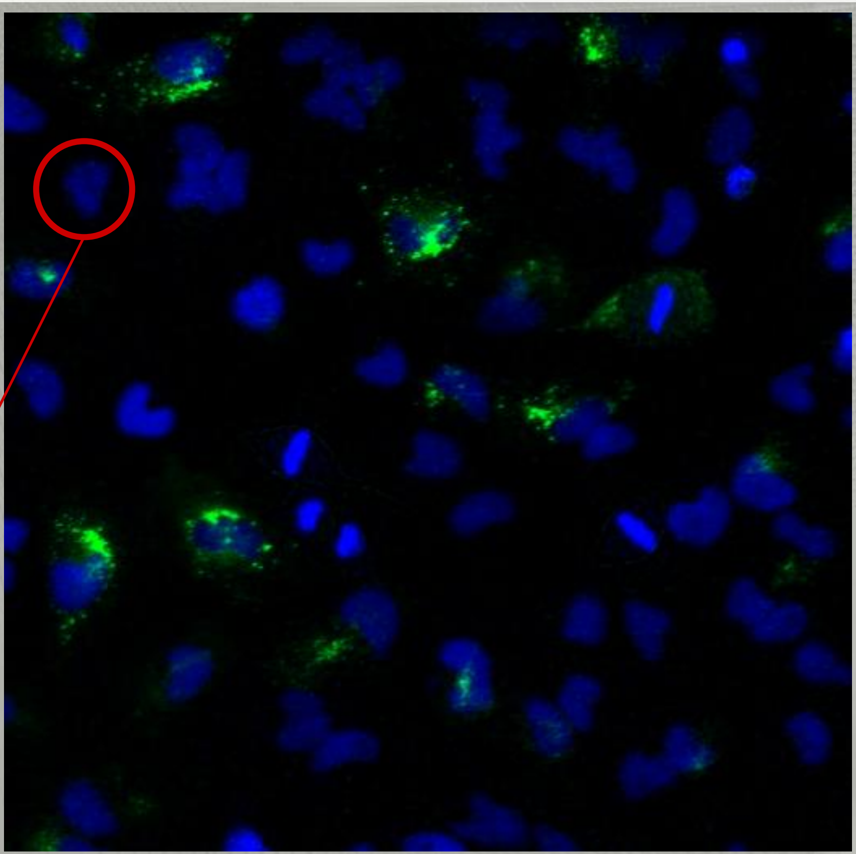


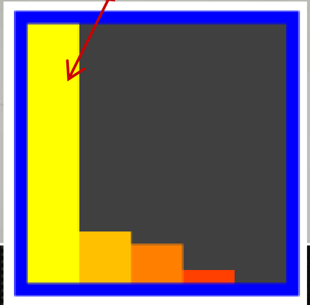


Balistreri, Horvath, Helenius.
Cell Host and Microbe.

Further results - Semliki Forest virus genome-wide screen

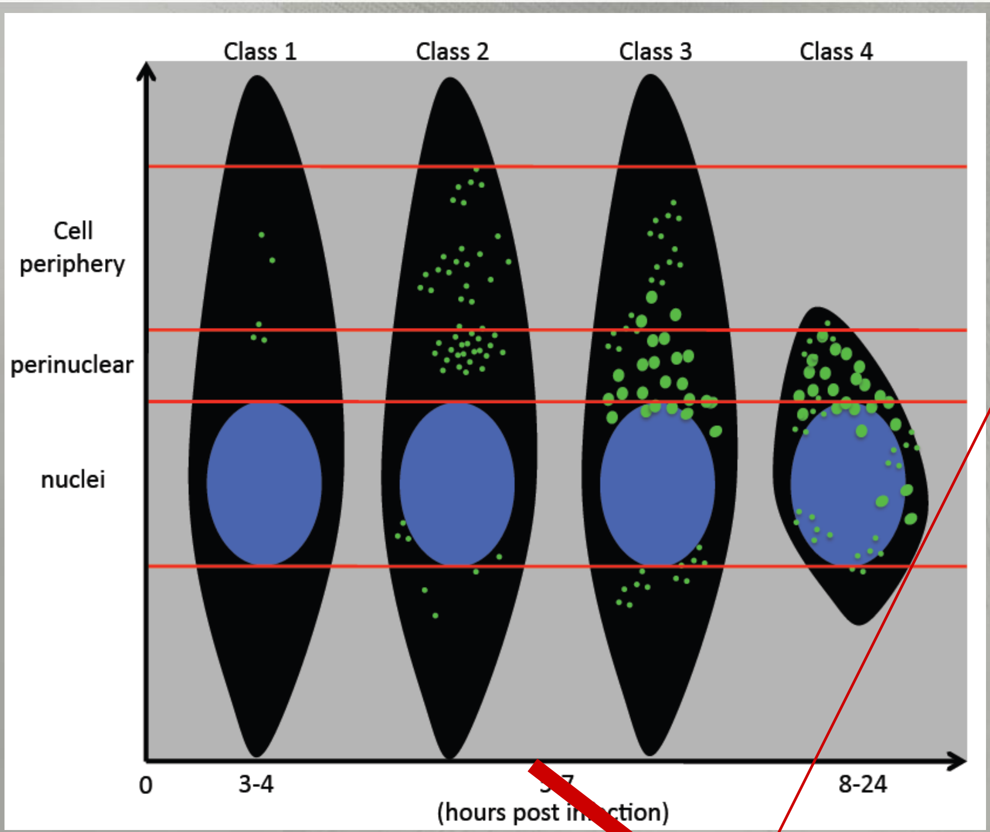


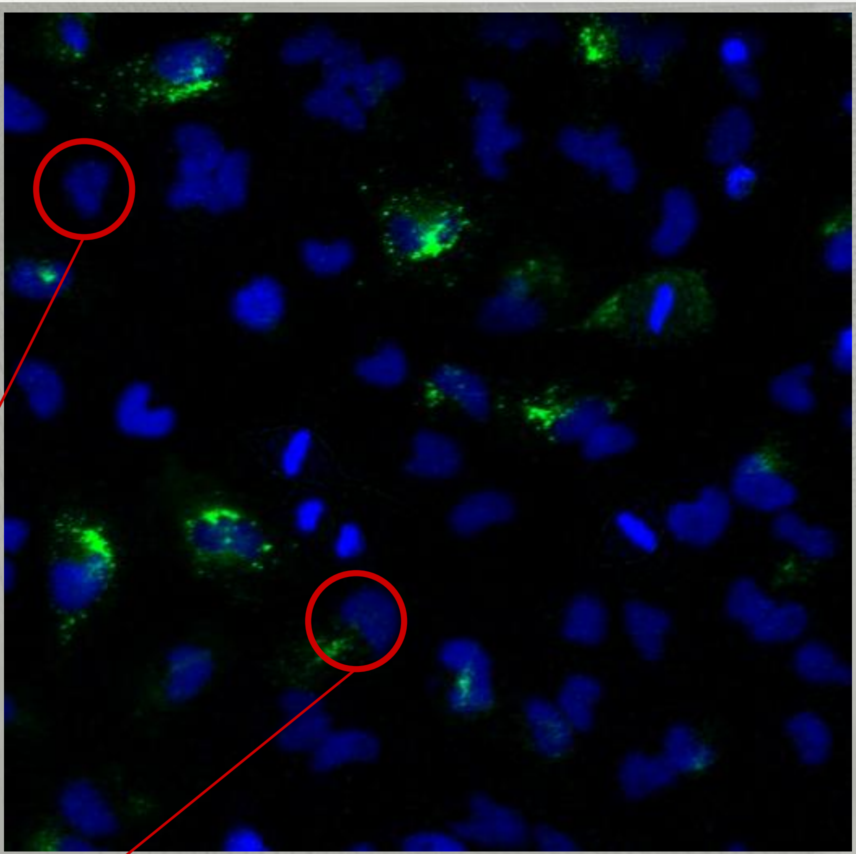


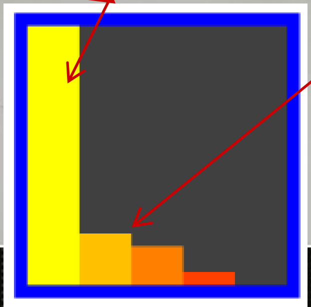


Balistreri, Horvath, Helenius.
Cell Host and Microbe.

Further results - Semliki Forest virus genome-wide screen

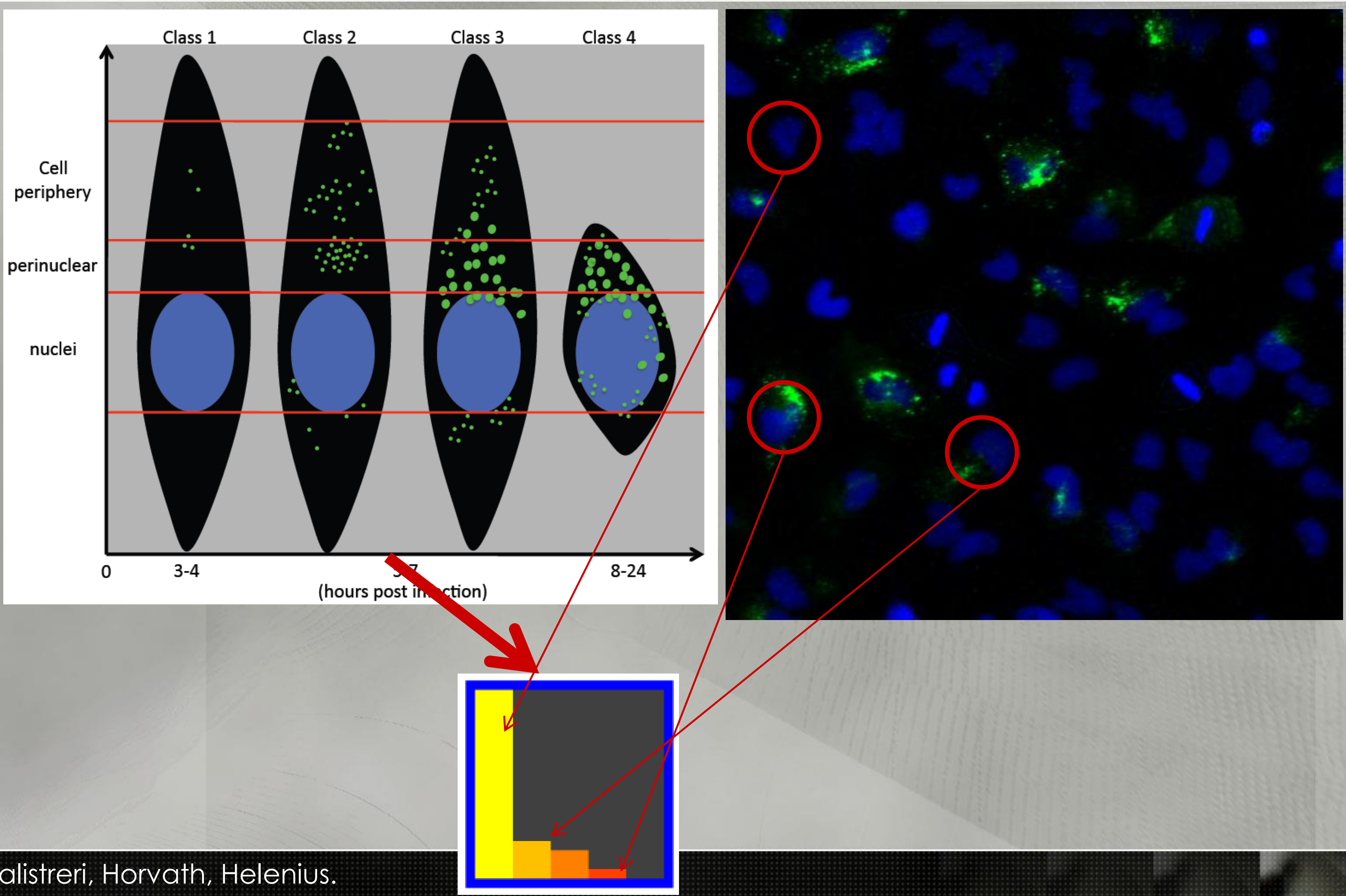






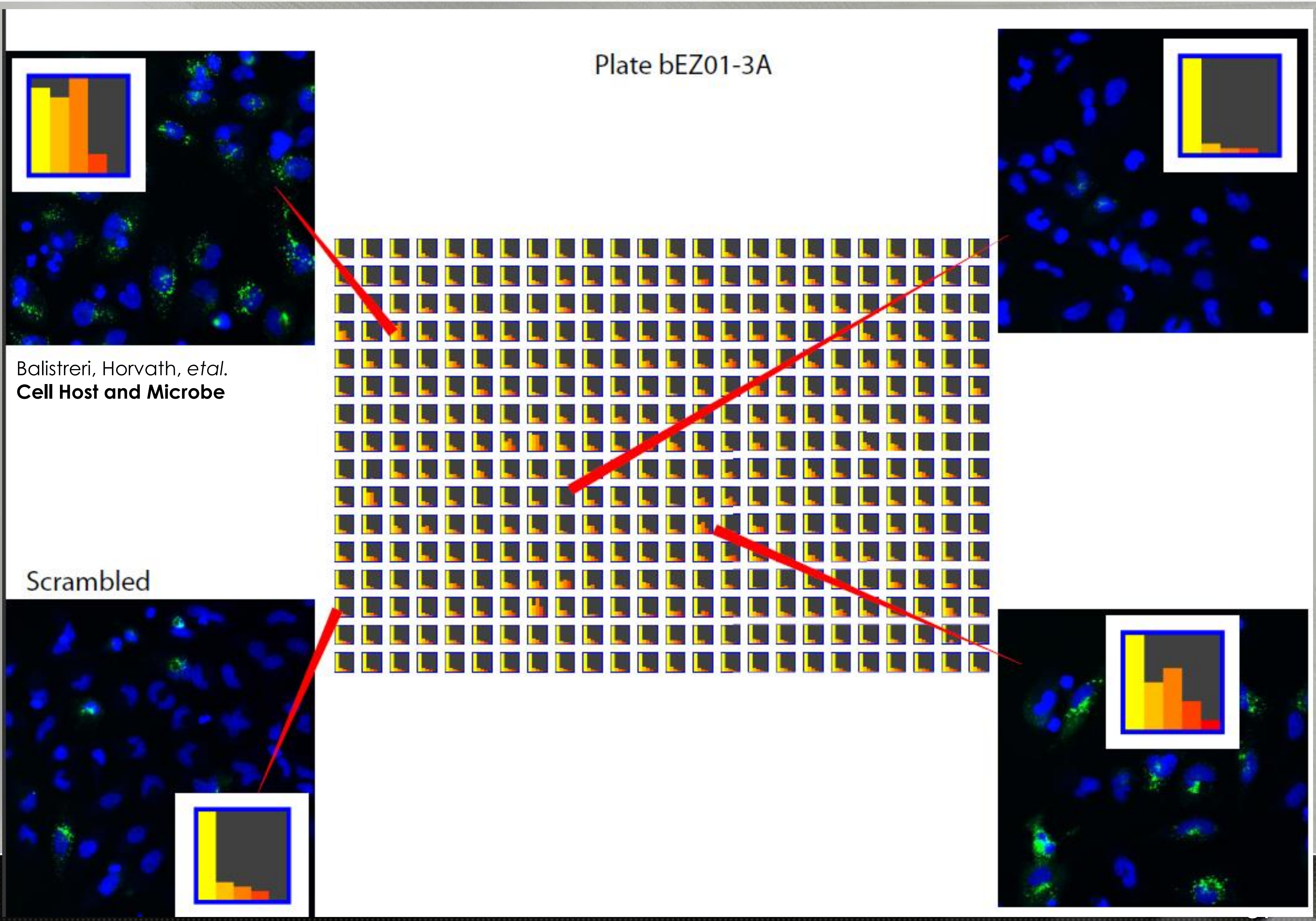
Balistreri, Horvath, Helenius.
Cell Host and Microbe.

Further results - Semliki Forest virus genome-wide screen



Balistreri, Horvath, Helenius.
Cell Host and Microbe.

Results

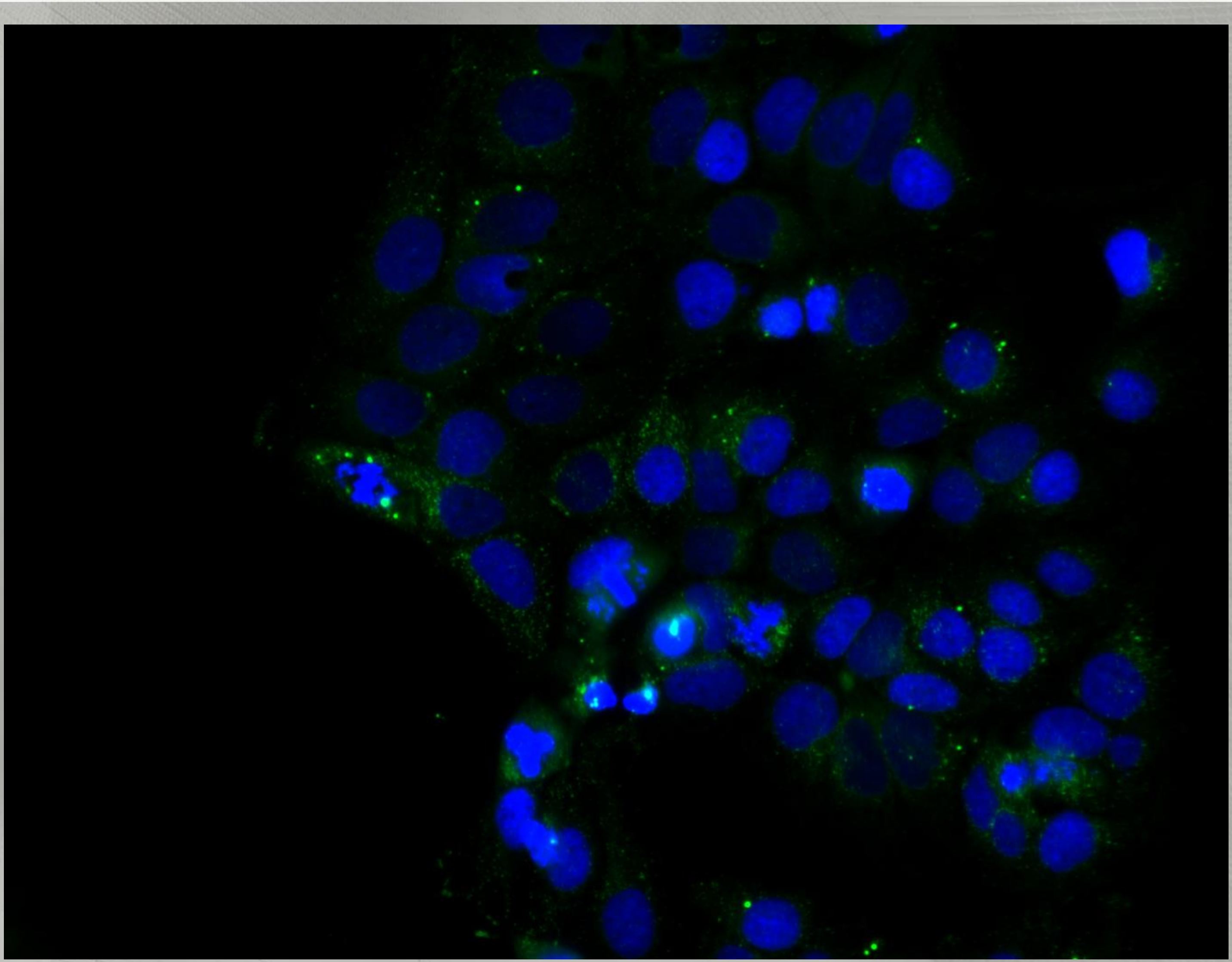


Balistreri, Horvath, *etal*.
Cell Host and Microbe

Lipid droplet phenotype analysis

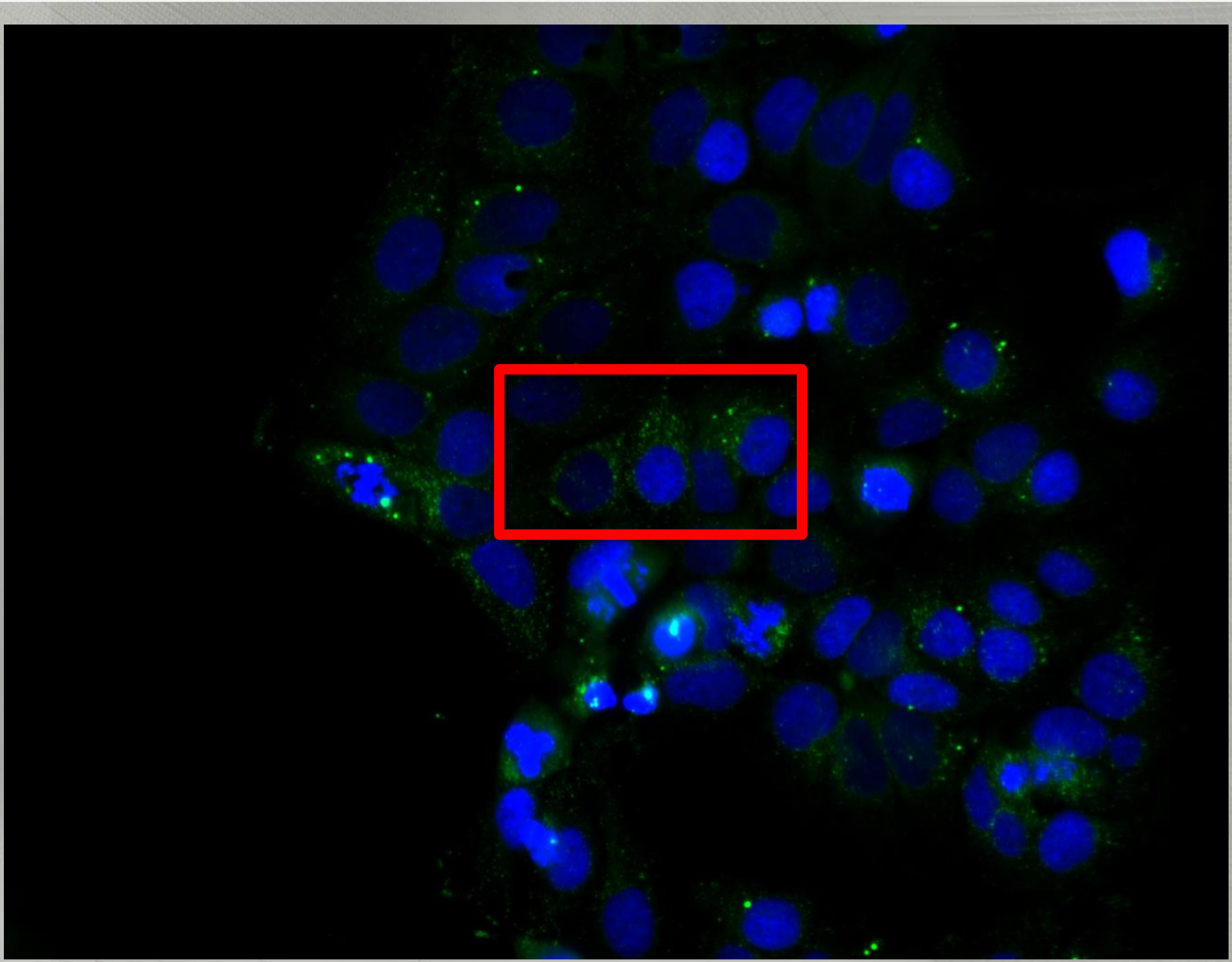
MORE COMPLEX DECISIONS

Lipid droplet phenotype analysis



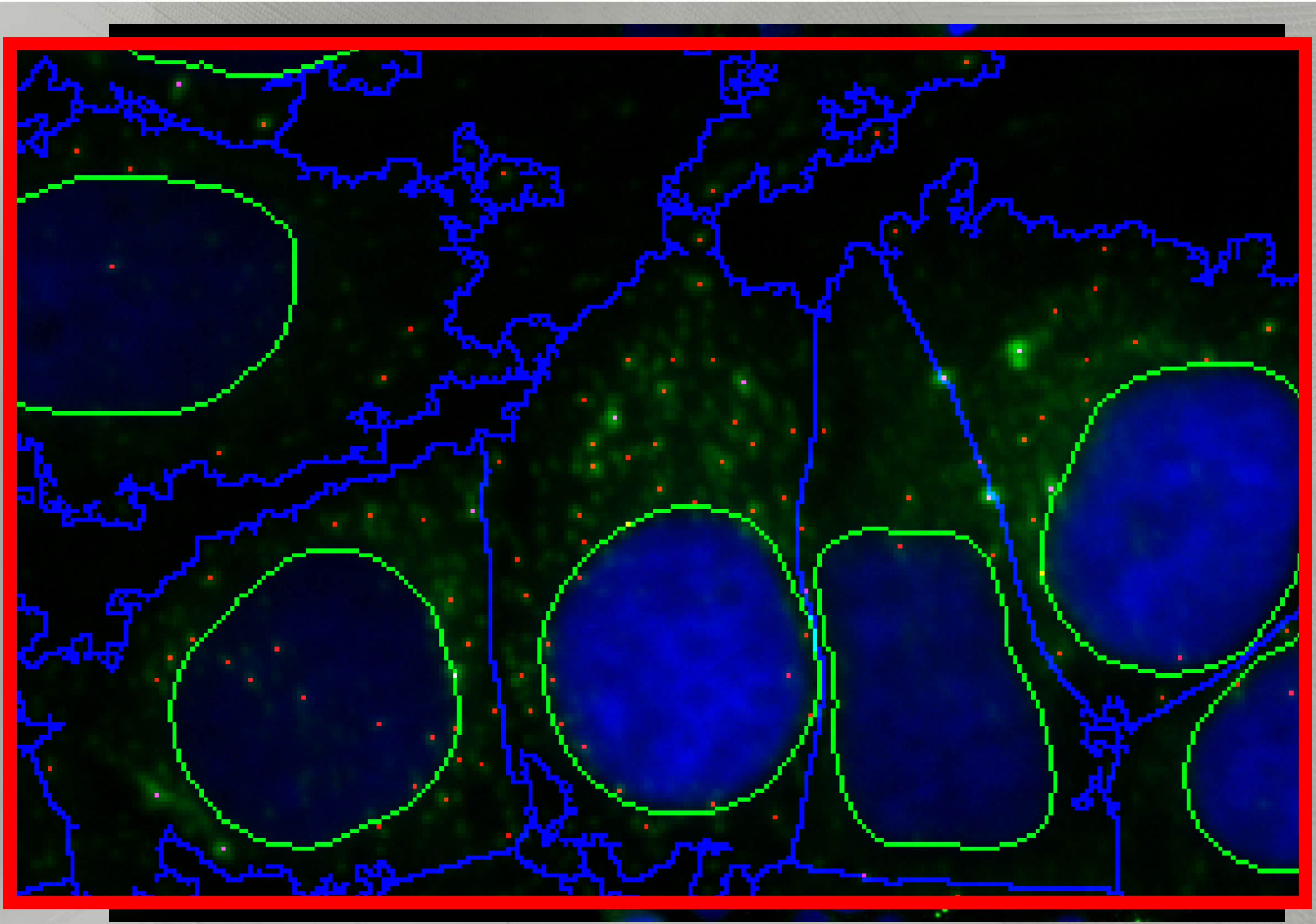
Based on: Surakka, ..., S. Timonen, V. Pietiäinen,… *etal*; The impact of low-frequency and rare variants on lipid levels. **Nature Genetics**

Lipid droplet phenotype analysis



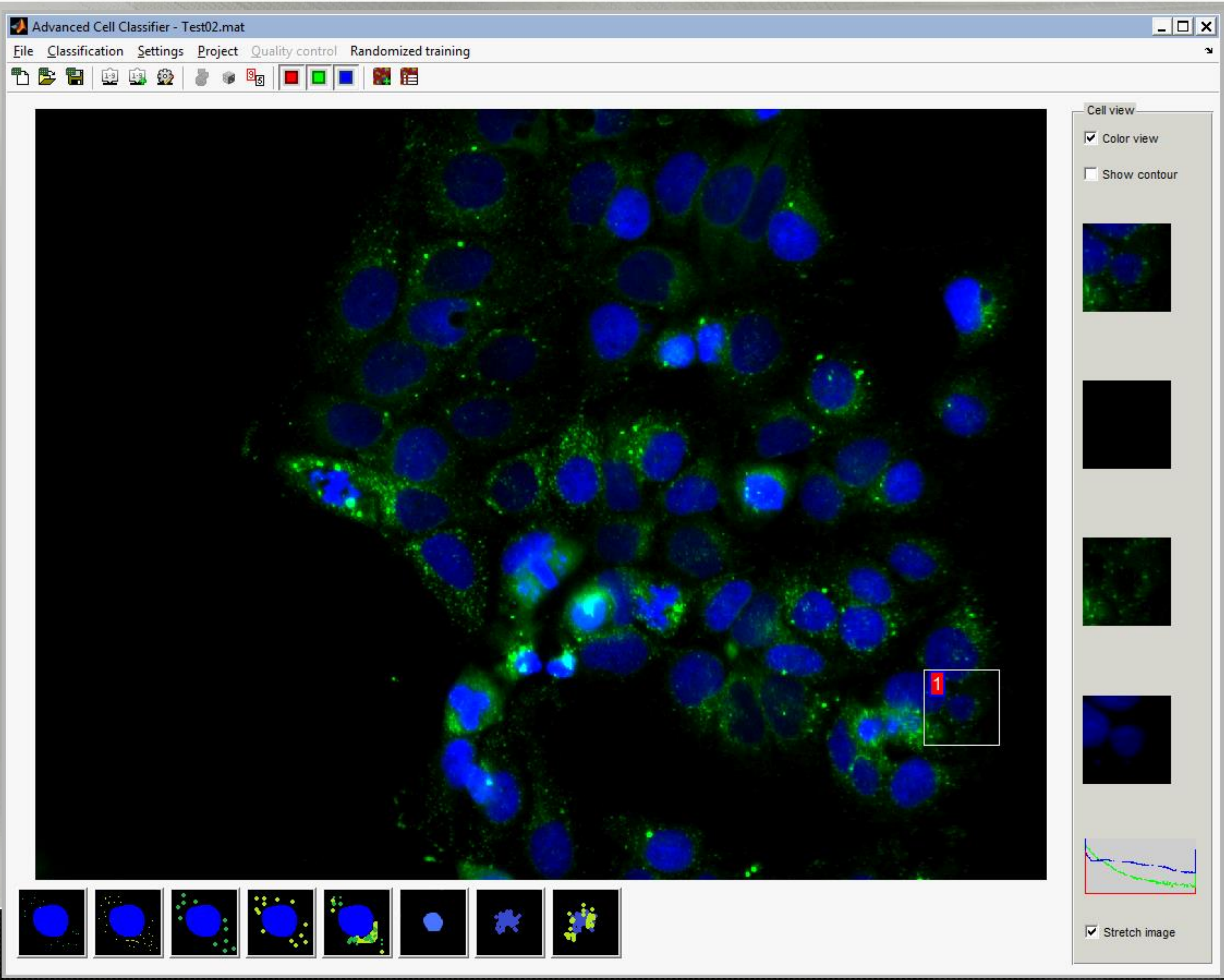
Based on: Surakka, ..., S. Timonen, V. Pietiäinen,... *etal*; The impact of low-frequency and rare variants on lipid levels. **Nature Genetics**

Lipid droplet phenotype analysis

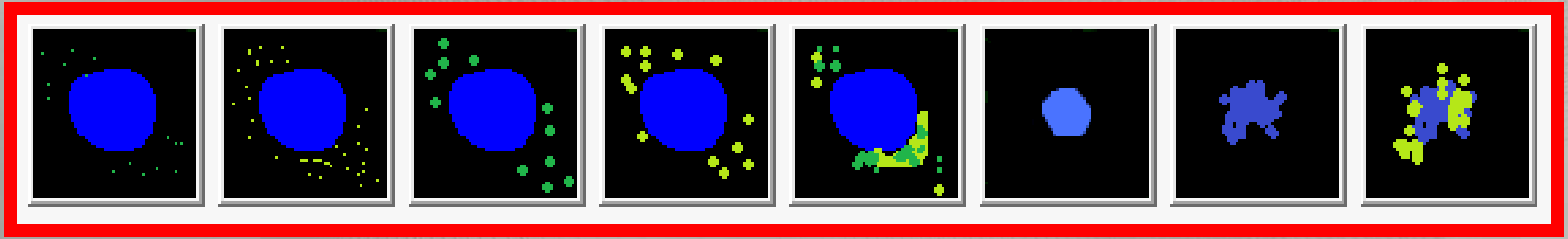


Based on: Surakka, ..., S. Timonen, V. Pietiäinen,... *etal*; The impact of low-frequency and rare variants on lipid levels. **Nature Genetics**

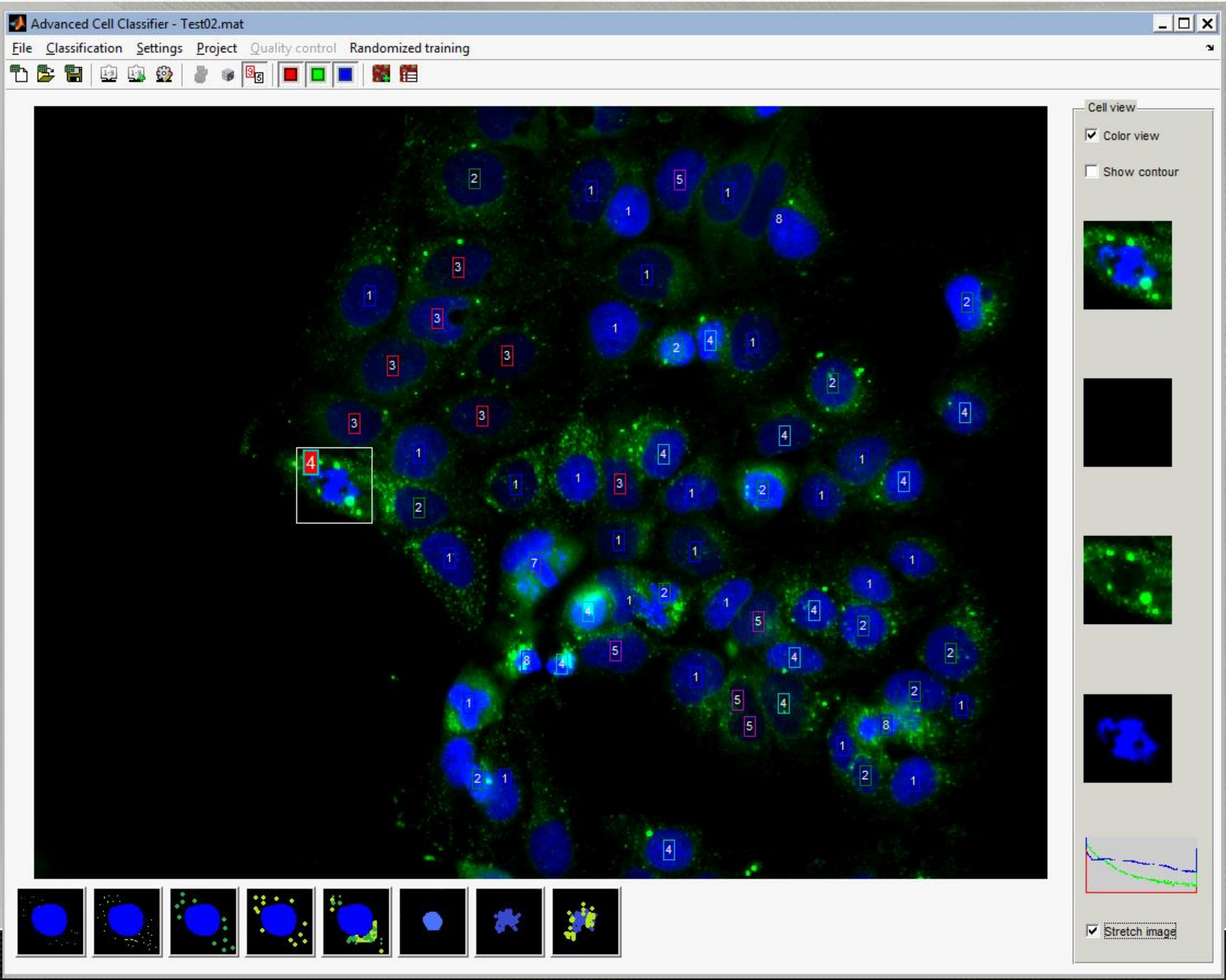
Using supervised classification



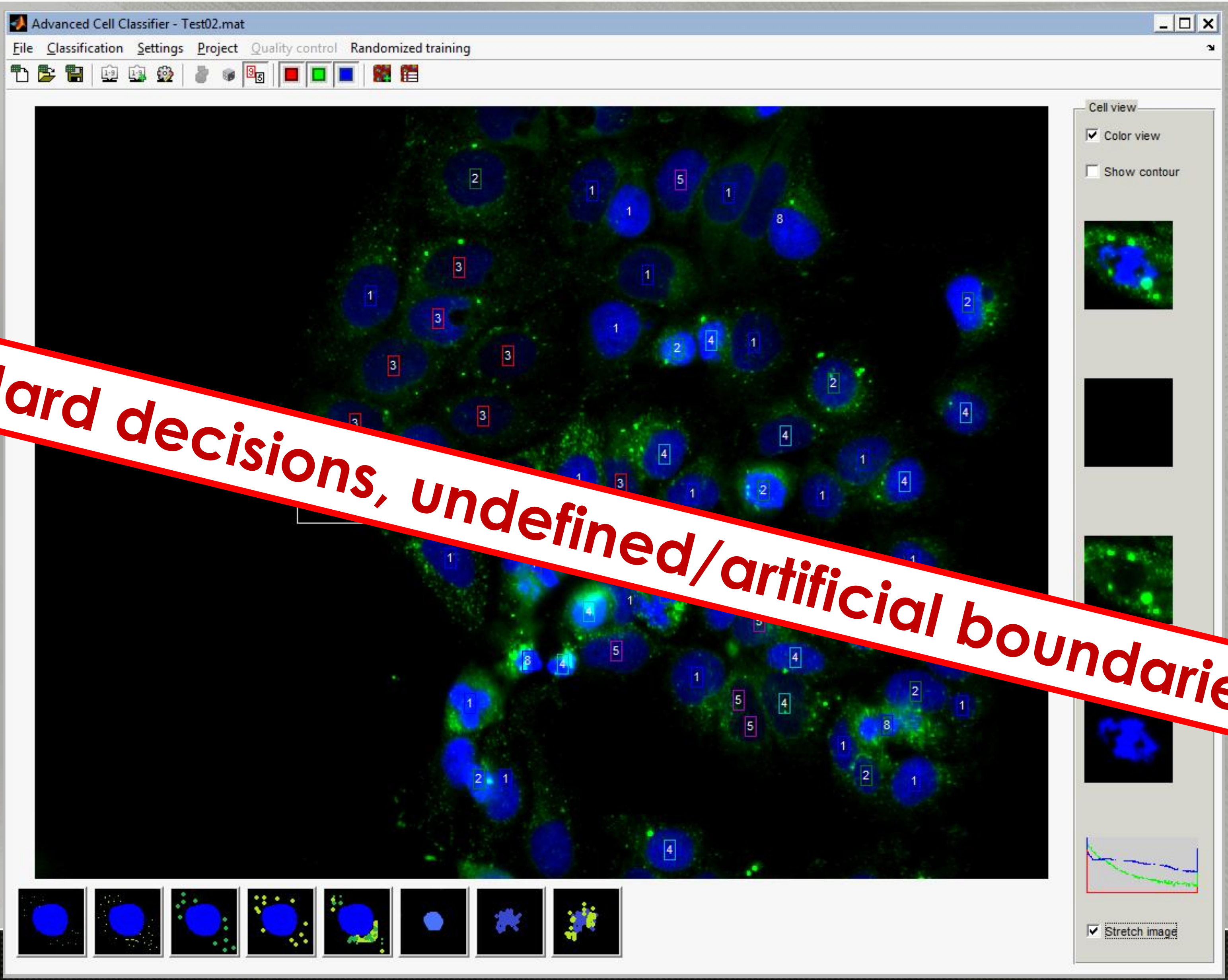
Using supervised classification



Using supervised classification



Using supervised classification



☑ Color view

☐ Show contour

☑ Stretch image

☹ Hard decisions, undefined/artificial boundaries ☹

2D regression plane concept

A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

2D regression plane concept

- Can we predict/learn more than one property of the cells?

A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

2D regression plane concept

- Can we predict/learn more than one property of the cells?
  - Yes, we can!

A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

2D regression plane concept

- Can we predict/learn more than one property of the cells?
  - Yes, we can!

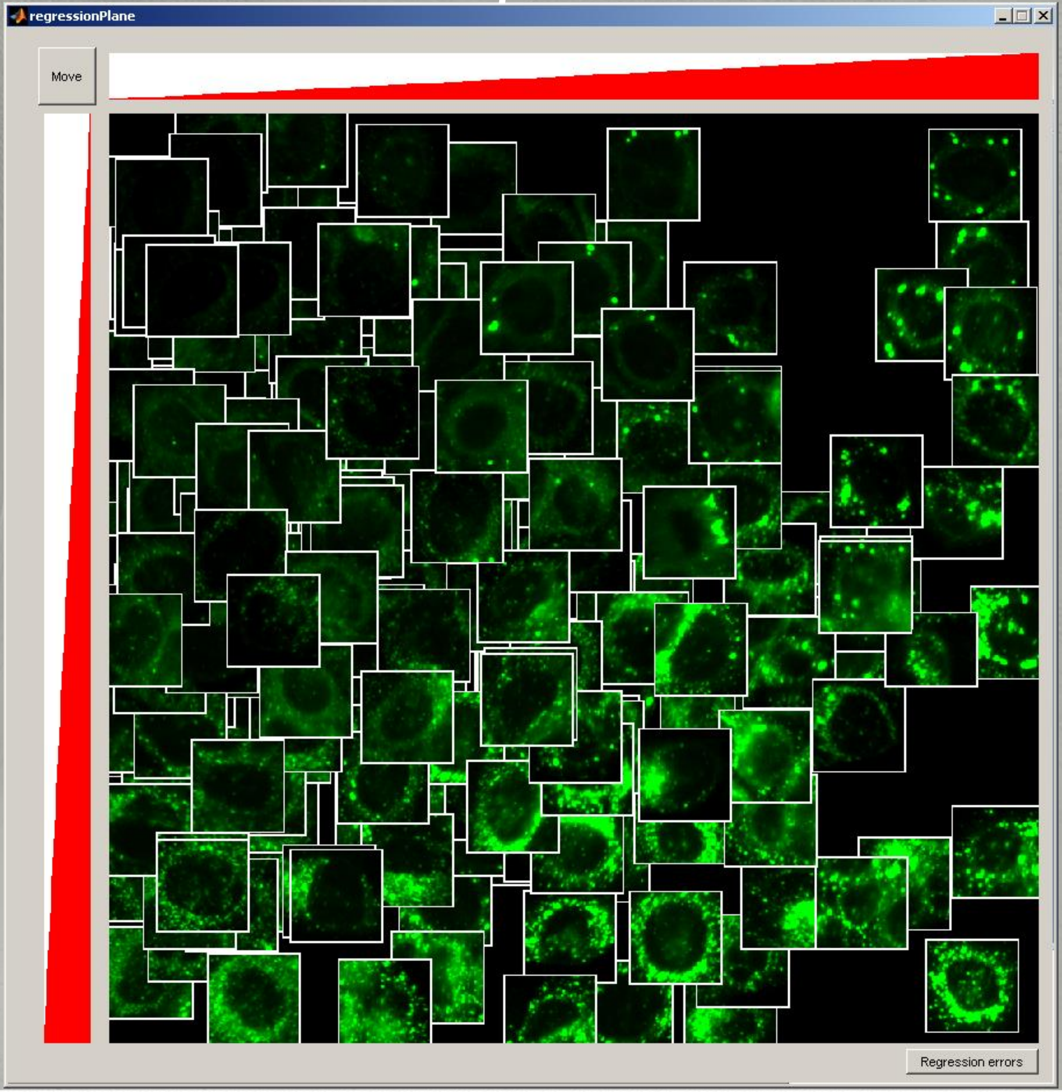
- I do not want to talk/analyze in terms of image processing quantities (jargon) but based on visual observations

A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

Example of an annotated plane

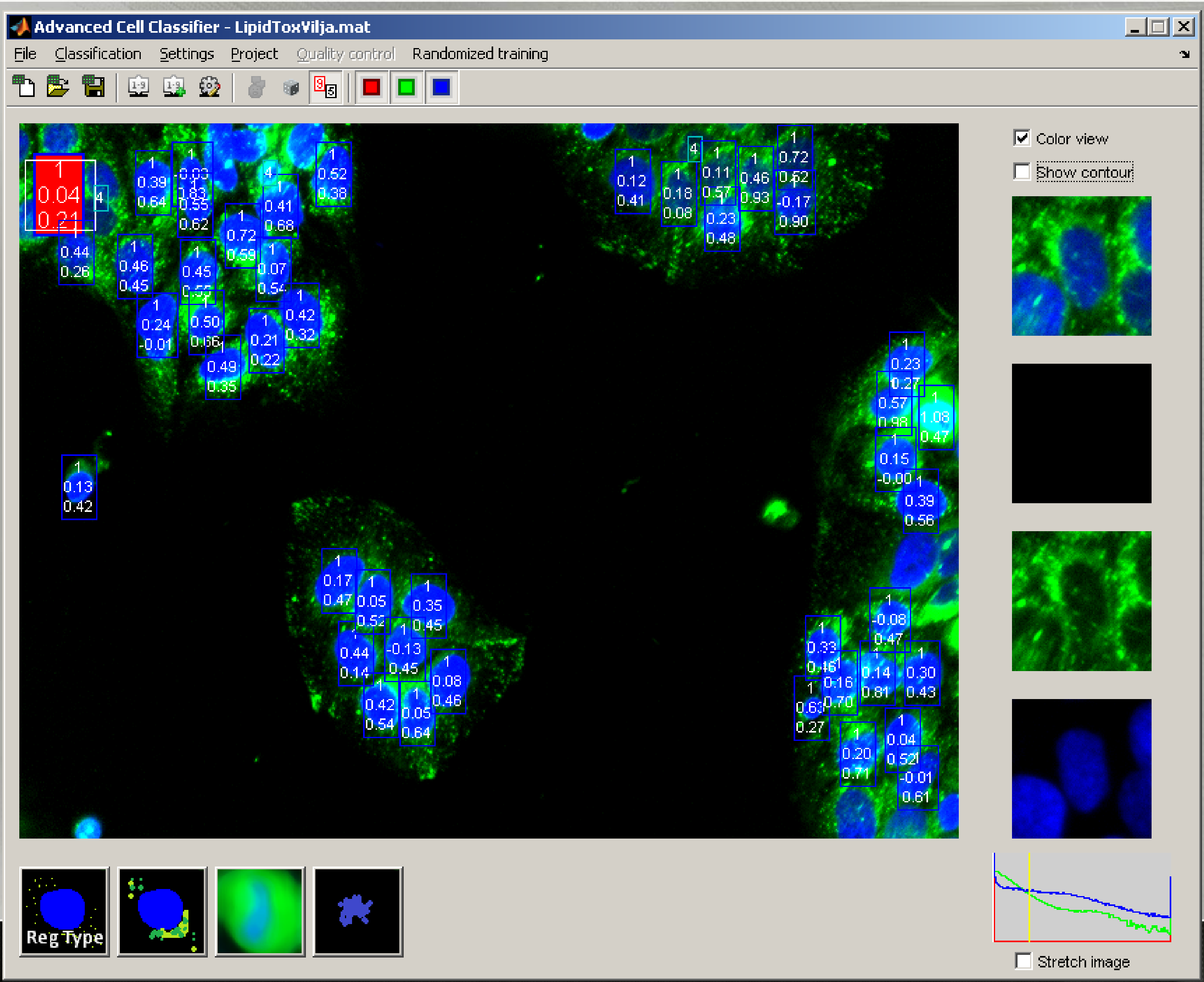
~Droplet size

~ Brightness



A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

Results

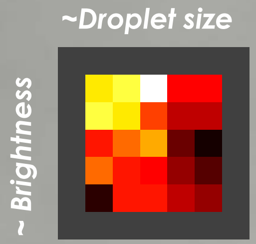


☑ Color view

☐ Show contour

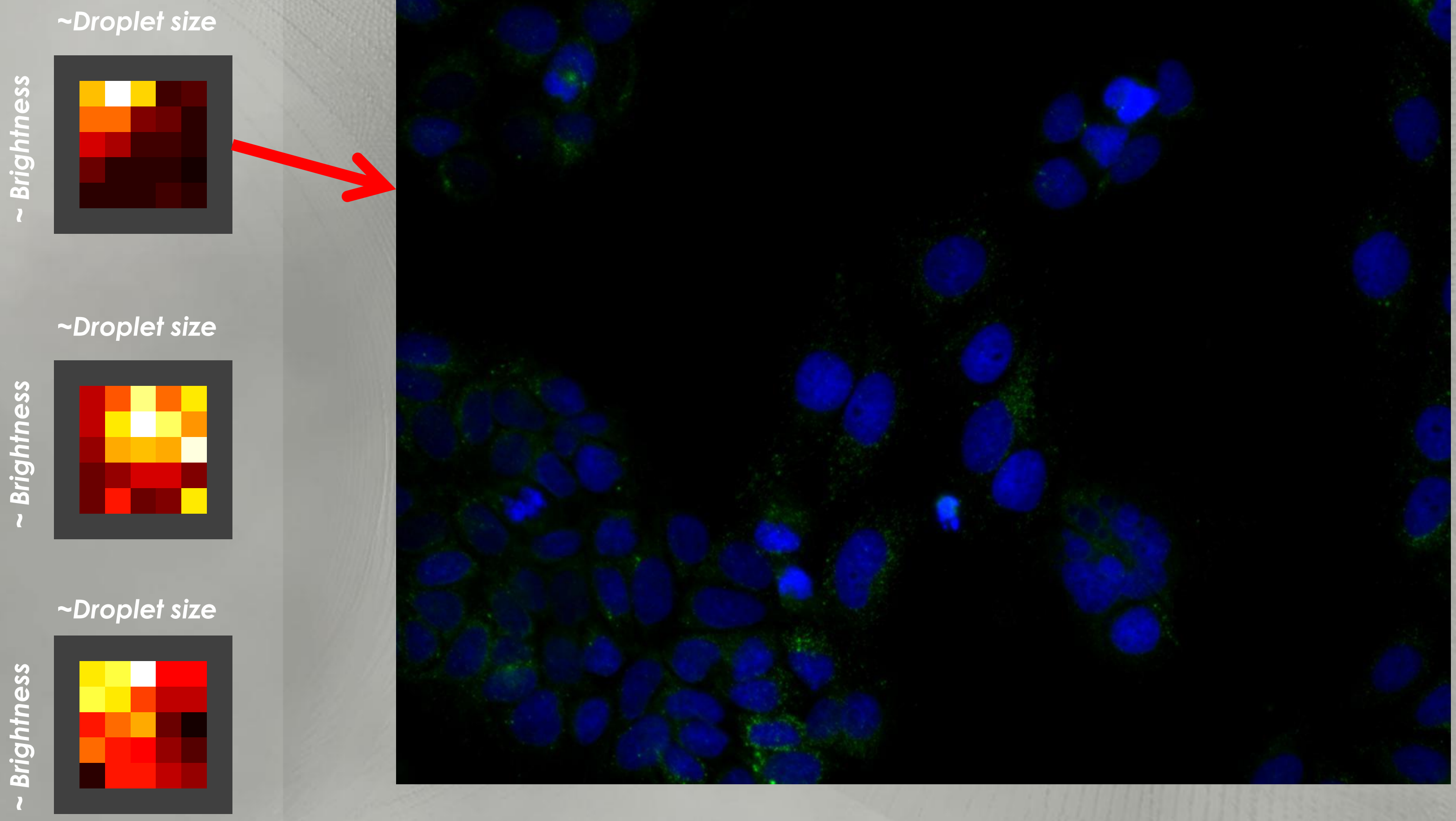
☐ Stretch image

Results



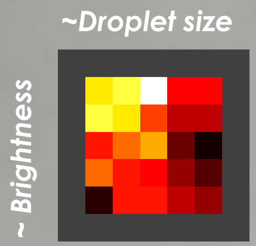
A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

Results



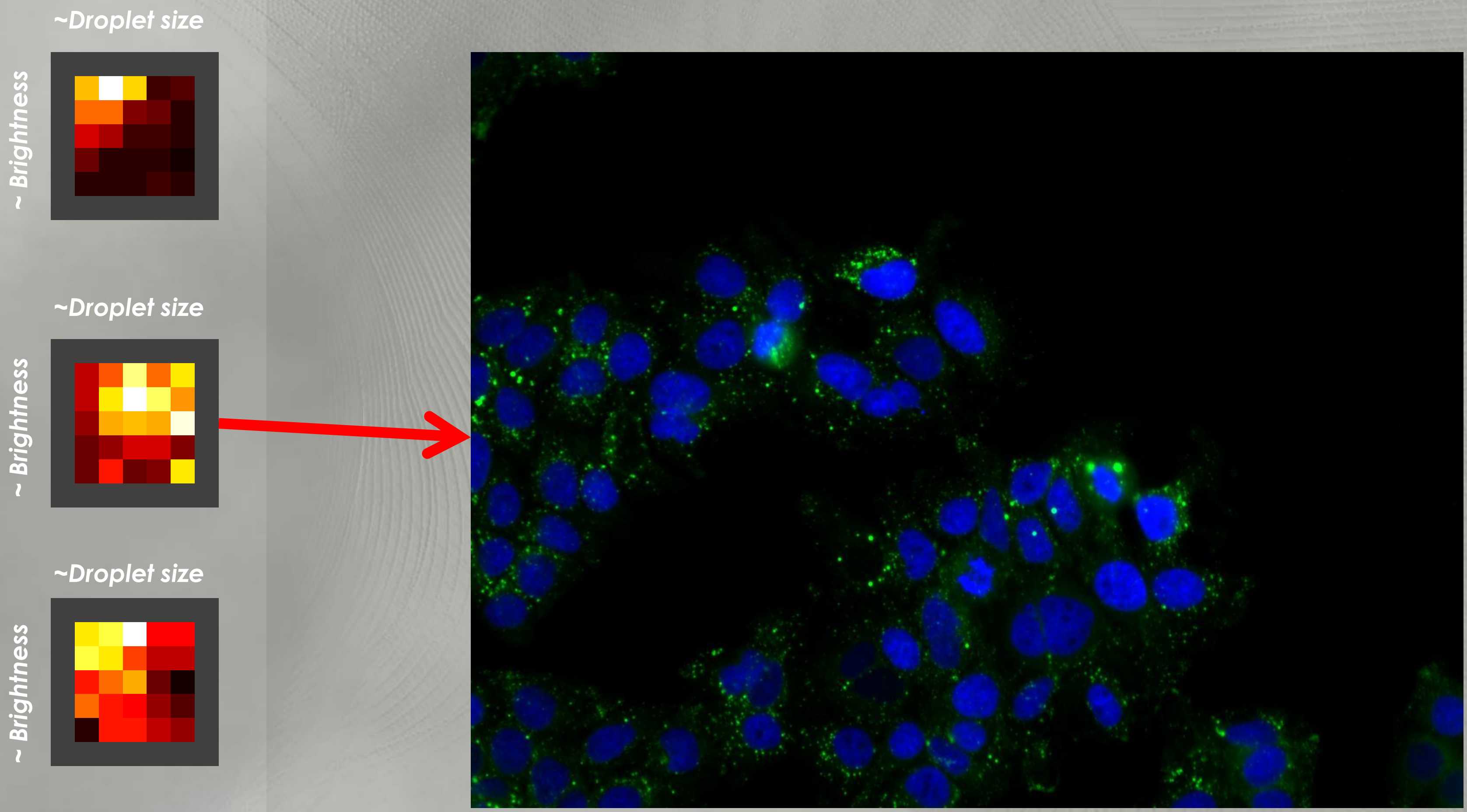
A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

Results



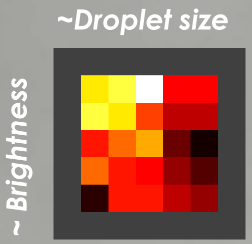
A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

Results



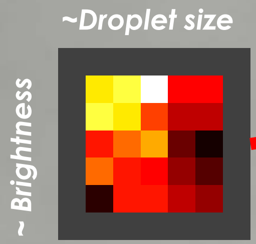
A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

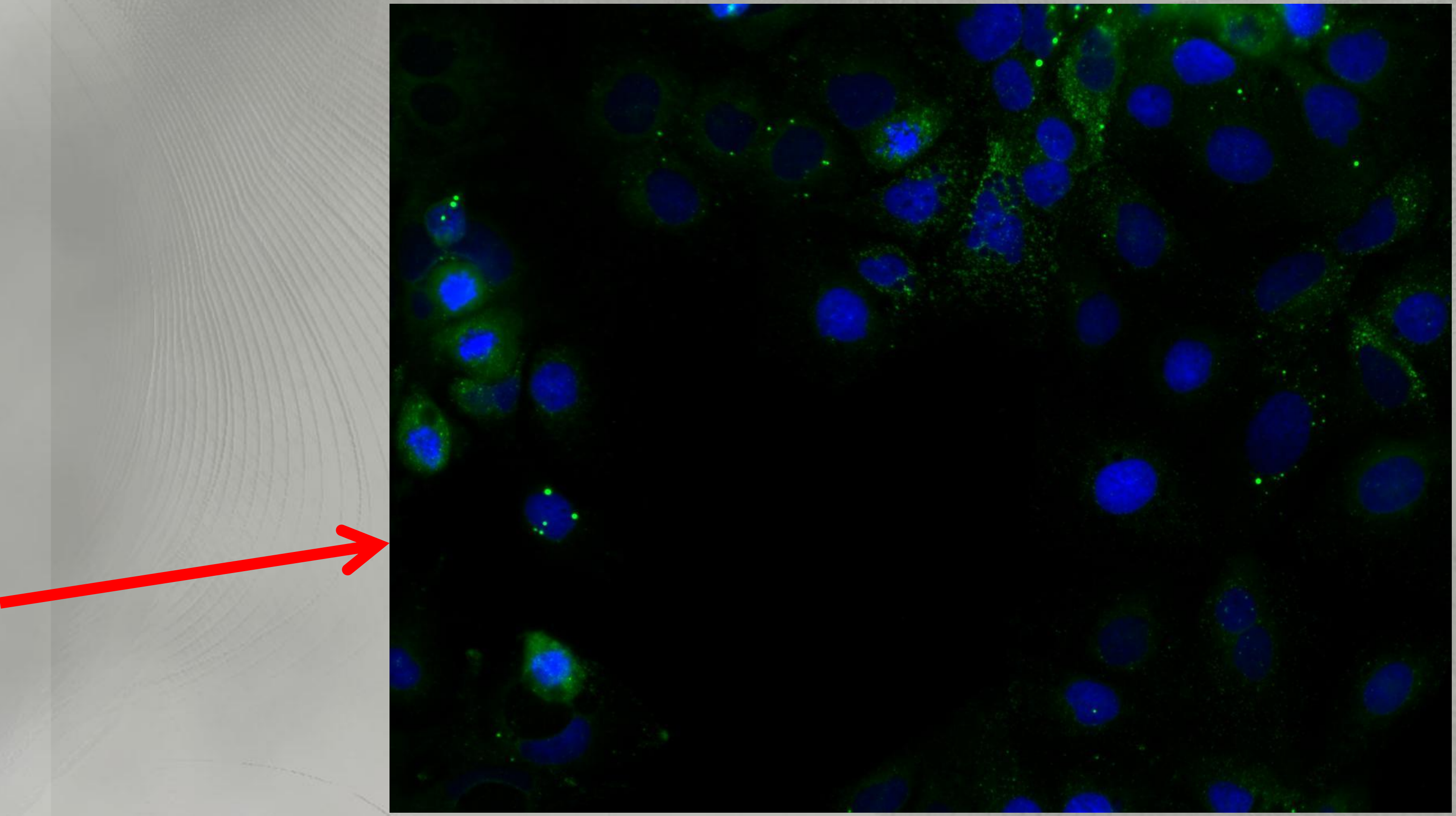
Results



A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

Results





A. Szkalisity, *etal:* A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes; KEPAF15

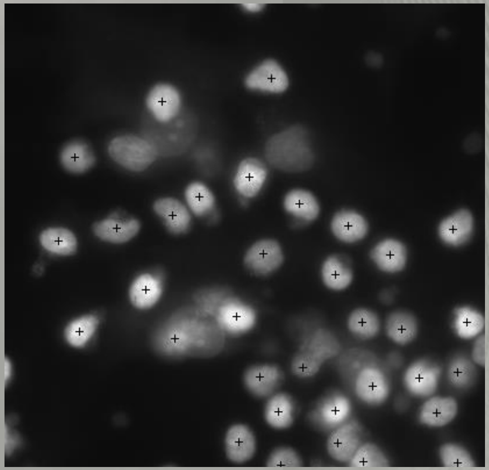
Childhood acute lymphoblastic leukemia

Prof. MD Bourquin Children's Hospital Zurich

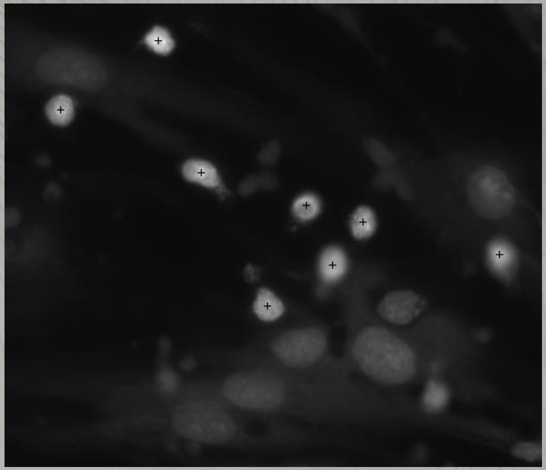
Drug and treatment strategy testing using high-content screening

Boutter, Horvath, Bourquin; ***Oncotarget 2014.***
Frismantas, etal. ***Blood 2017.***

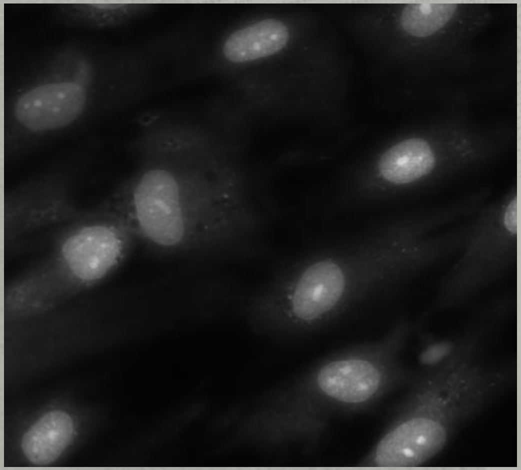
Biology idea



Untreated population



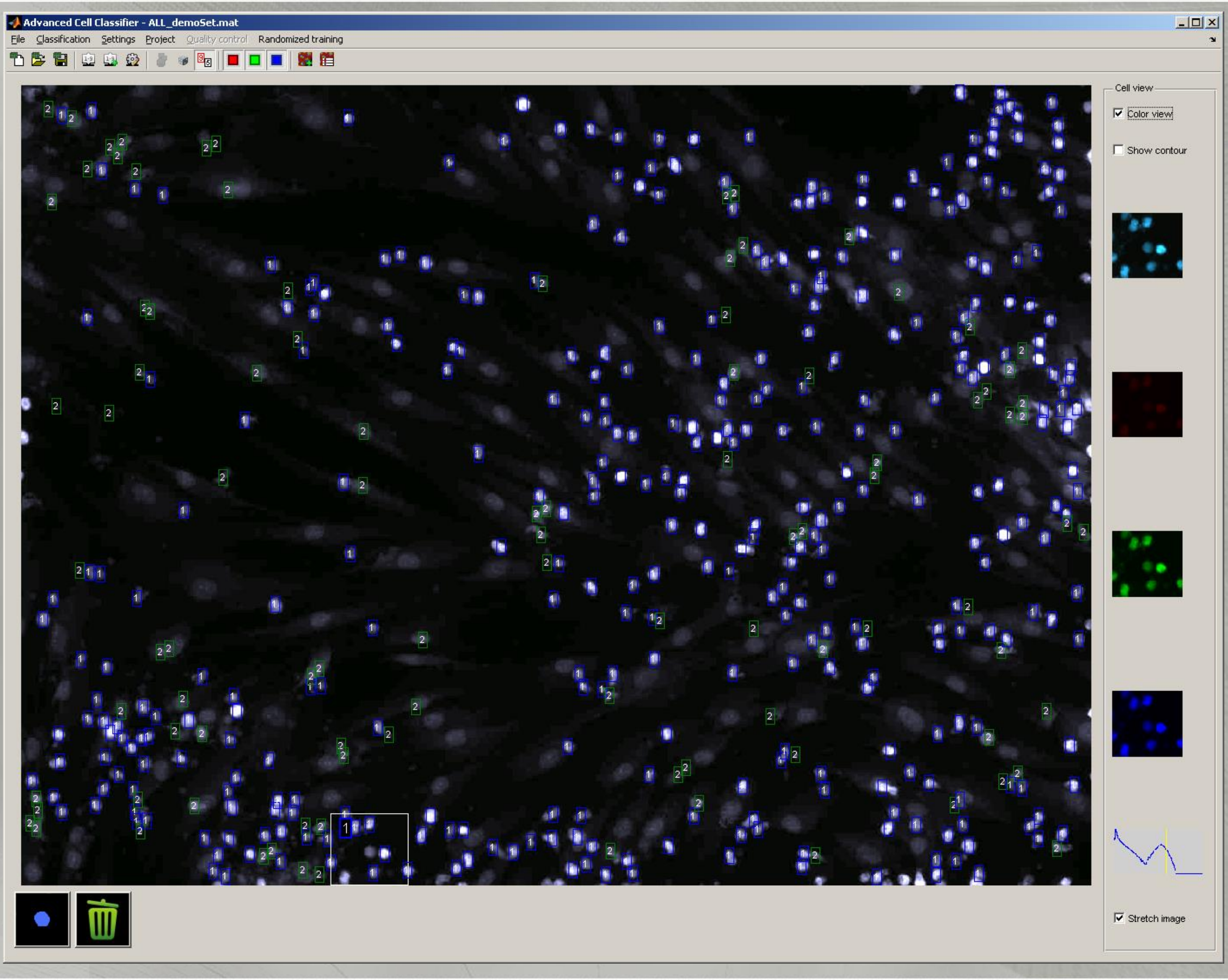
XY gene or drug



Bone marrow cells

Boutter, Horvath, Bourquin; ***Oncotarget 2014.***
Frismantas, etal. ***Blood 2017***

Detection of cells



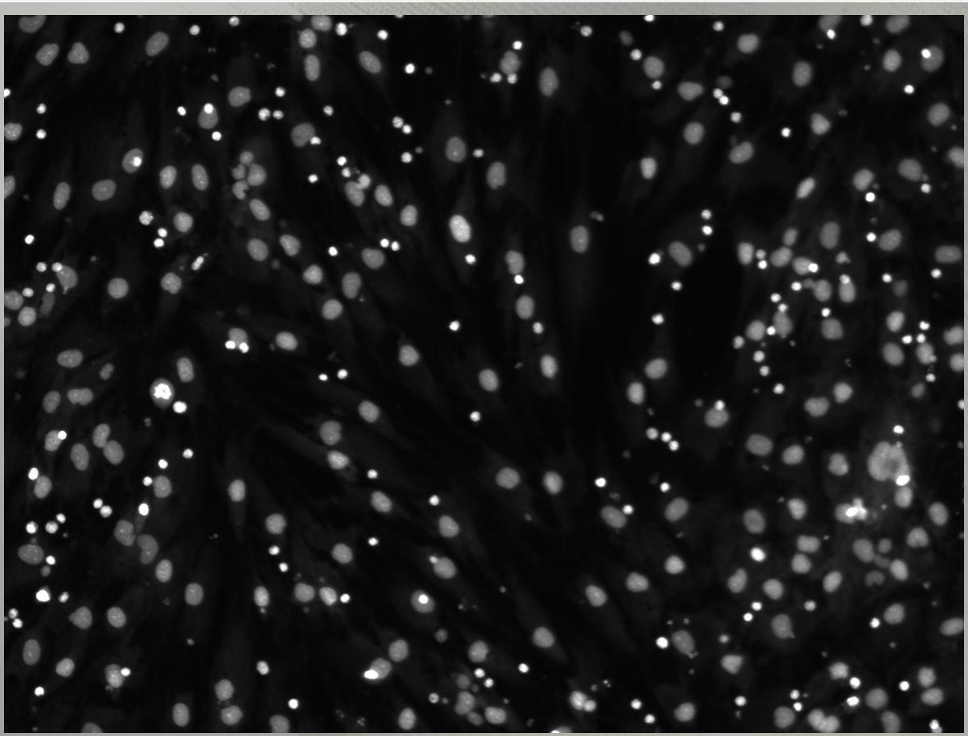
☑ Color view

☐ Show contour

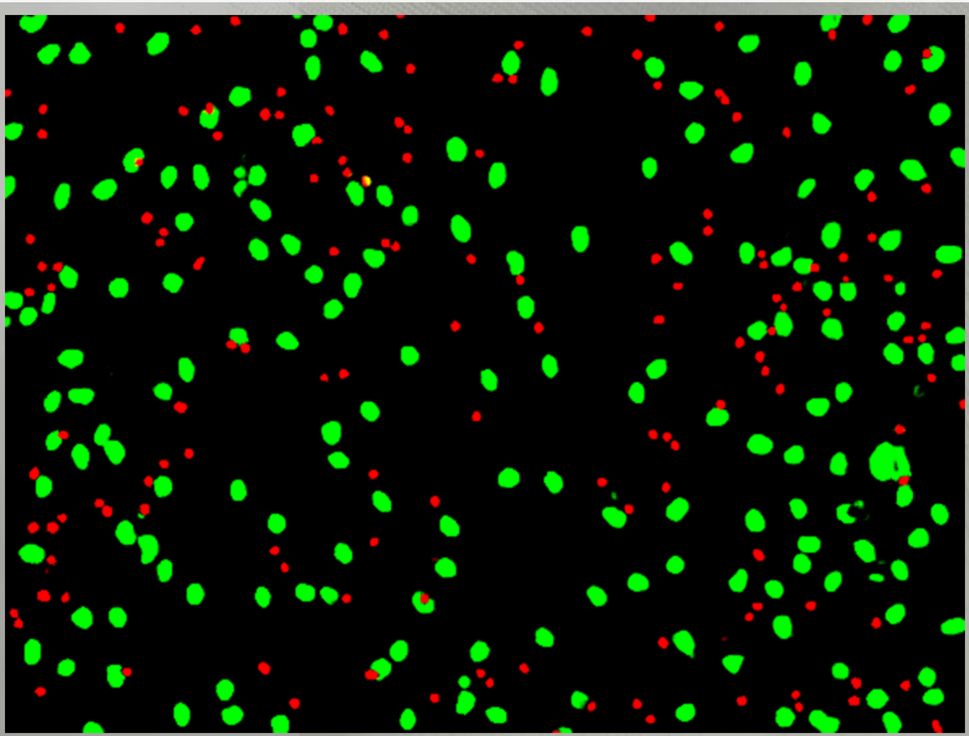
☑ Stretch image

Boutter, Horvath, Bourquin; ***Oncotarget 2014.***
Frismantas, etal. ***Blood 2017***

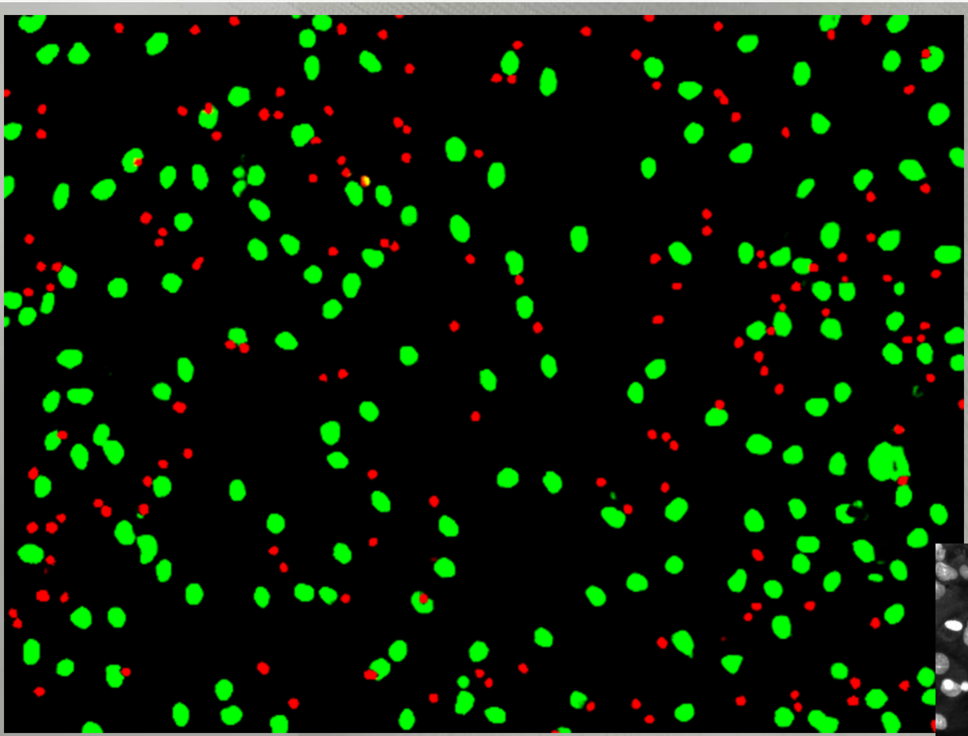
UNET deep learning (prelim)

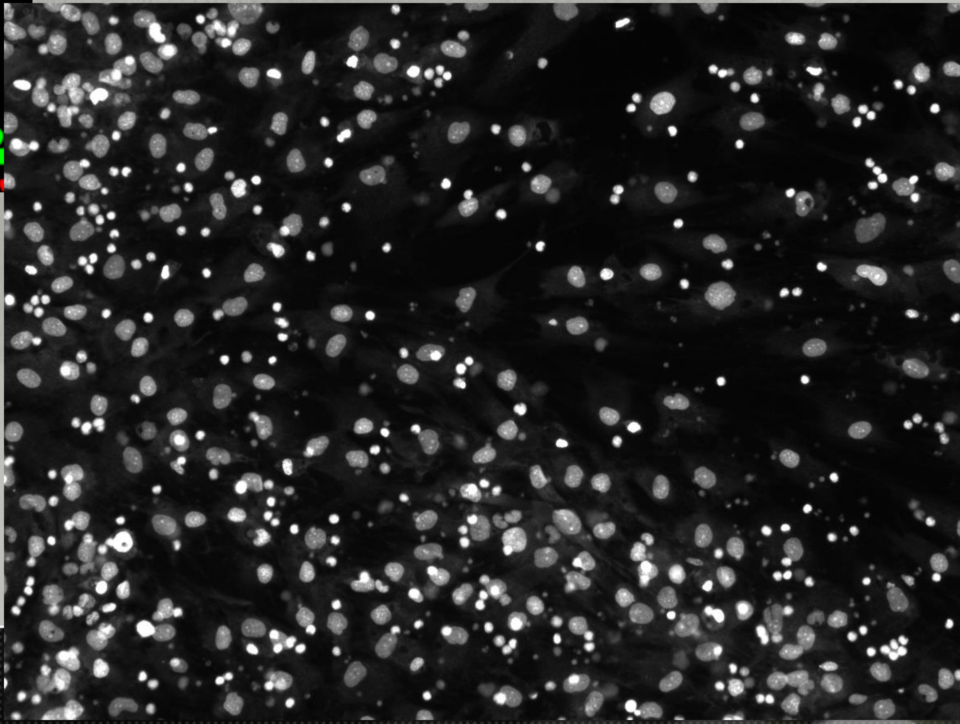


UNET deep learning (prelim)



UNET deep learning (prelim)





UNET deep learning (prelim)

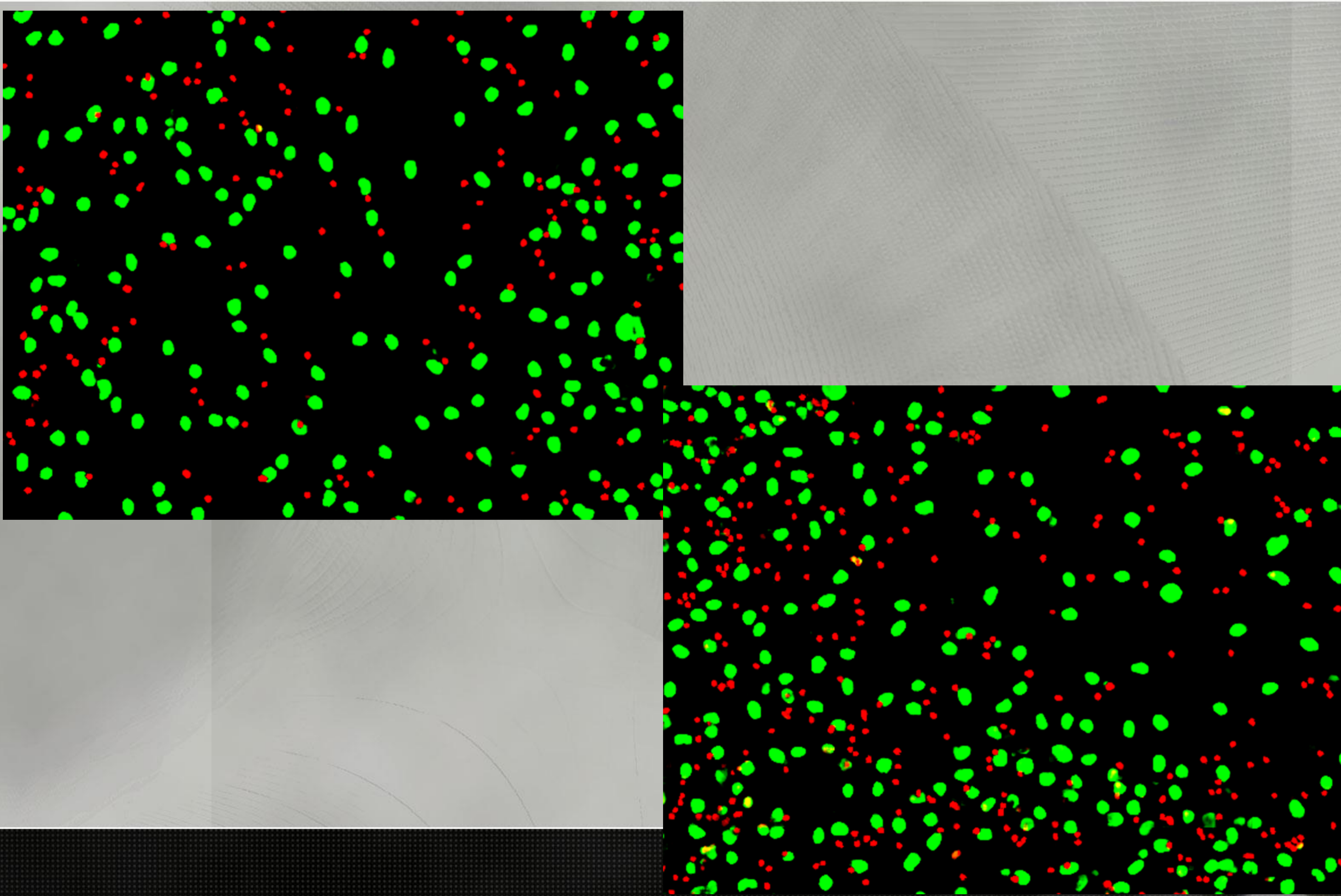
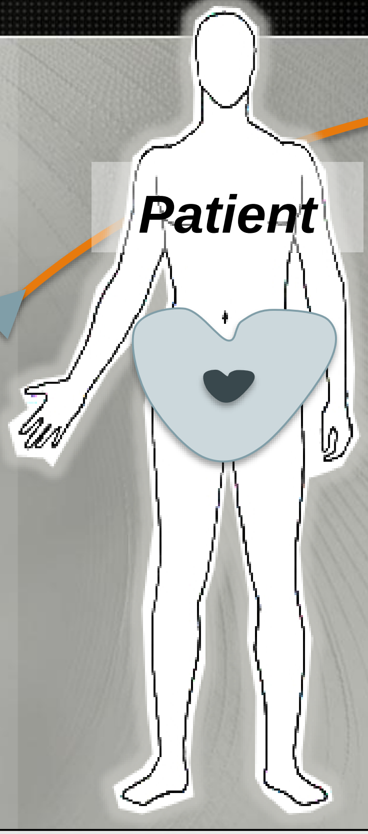


Image-based personalized and translational medicine

ERAPerMed

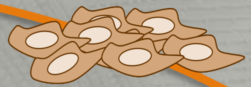
HORIZON 2020

Personalized **precision** medicine

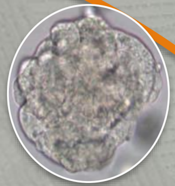


Patient

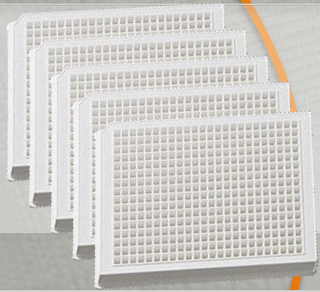
Oncology



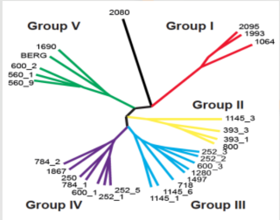
Cell model

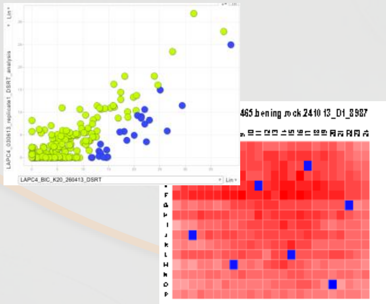


Primary drug testing
- *~550 oncology drugs*
- *Cell lines*
- *Patient samples*

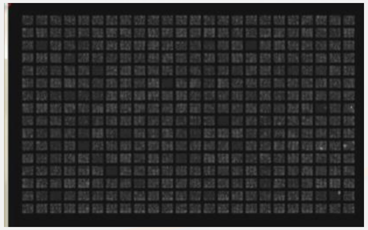


Bioinformatics





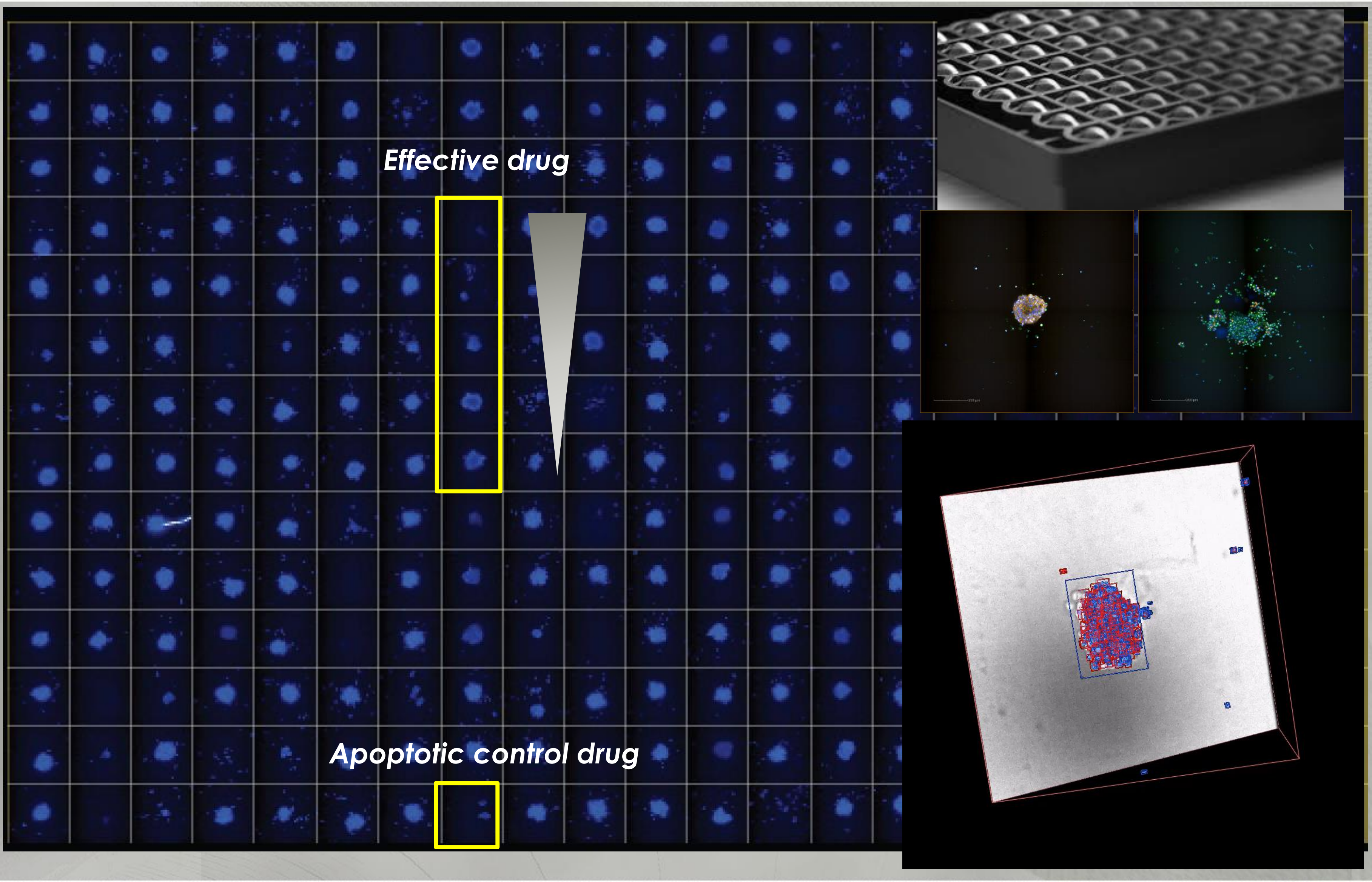
Imaging



Microscopy screening

Saeed etal. – European Urology (2017)

Pediatric brain tumors – 3D



Heike Peterziel, Olaf Witt, Sina Oppermann, KiTZ, Heidelberg, collaboration

Advancing 3D single-cell based screening

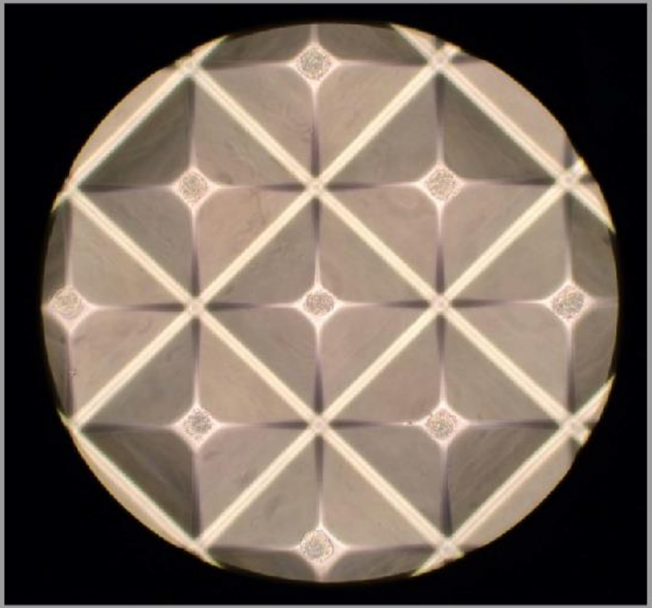
Pipeline:

- ➢Spheroid generation
- ➢Spheroid picker
- ➢Imaging
- ➢Image analysis

Spheroid generation

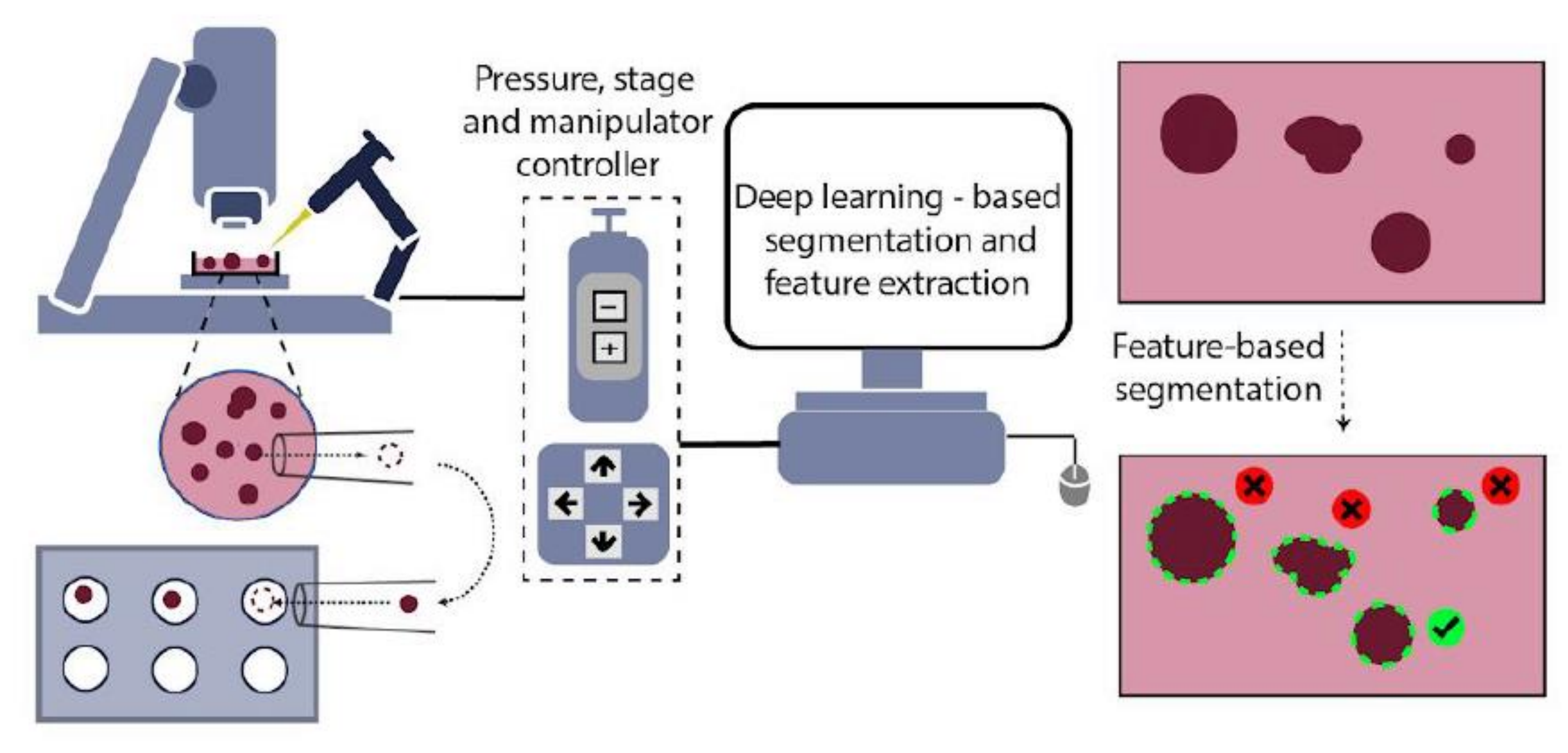
- **InSphero GravityPLUS Kit**
- **SphericalPlate5D**
- **Homogenous size (200-300 µm)**

Source: © InSphero AG

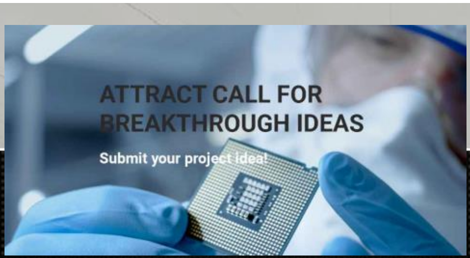


Source: © Kugelmeiers AG

Spheroid picker



HORIZON 2020

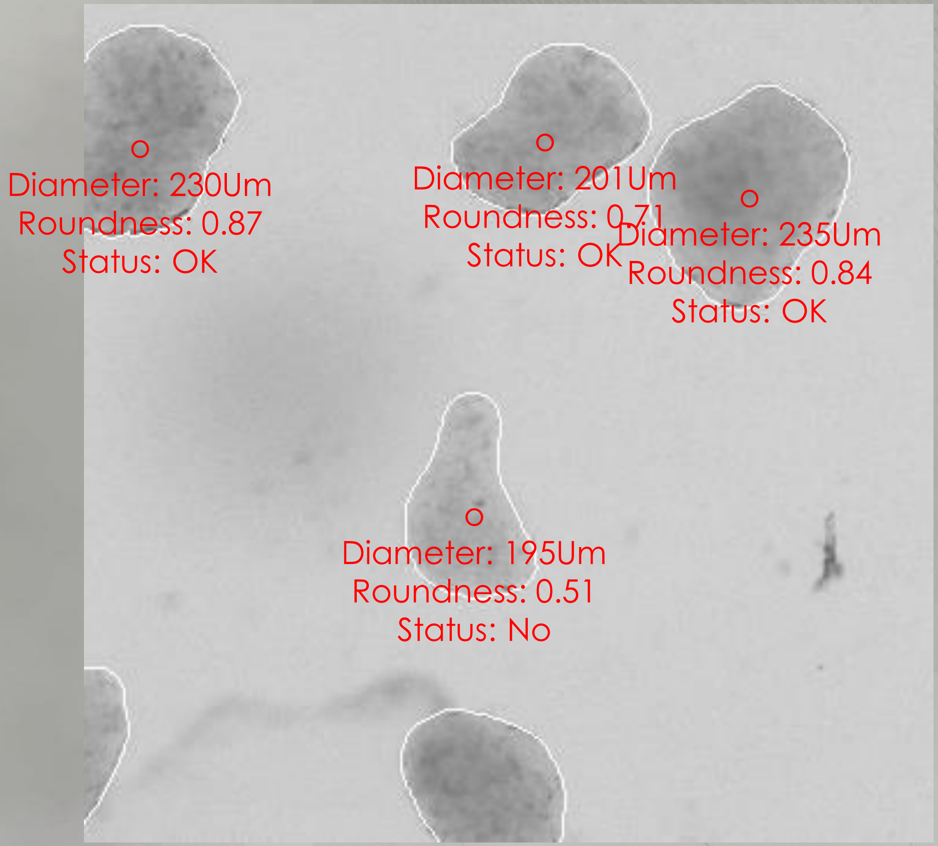

ATTRACT CALL FOR BREAKTHROUGH IDEAS

Submit your project idea!

ATTRACT

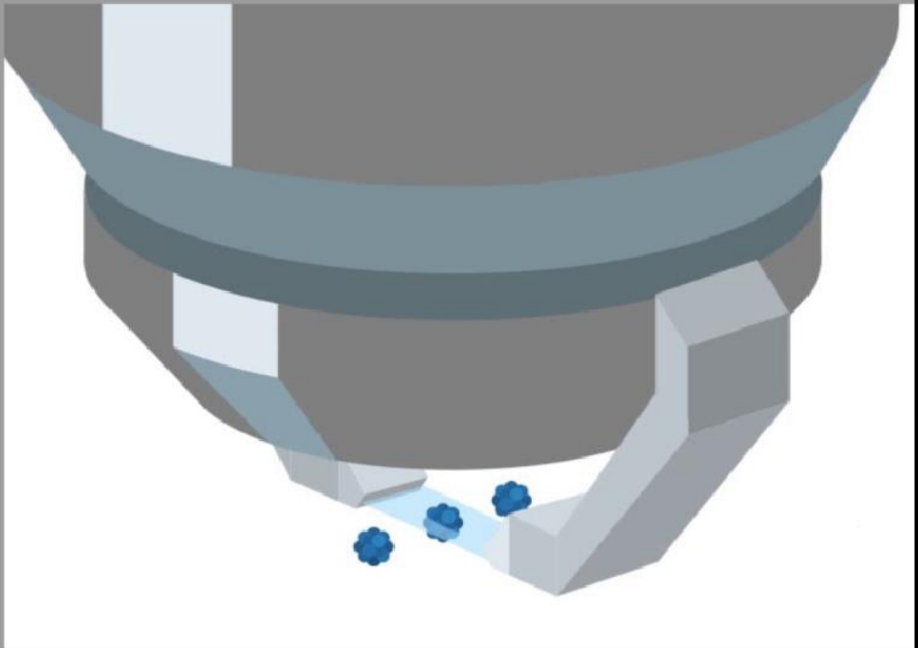
Spheroid picker

- **Deep learning detection**



3D large organoid imaging

- **SCALE tissue clearing**
- **Leica SP8 DLS**
- **3D imaging**



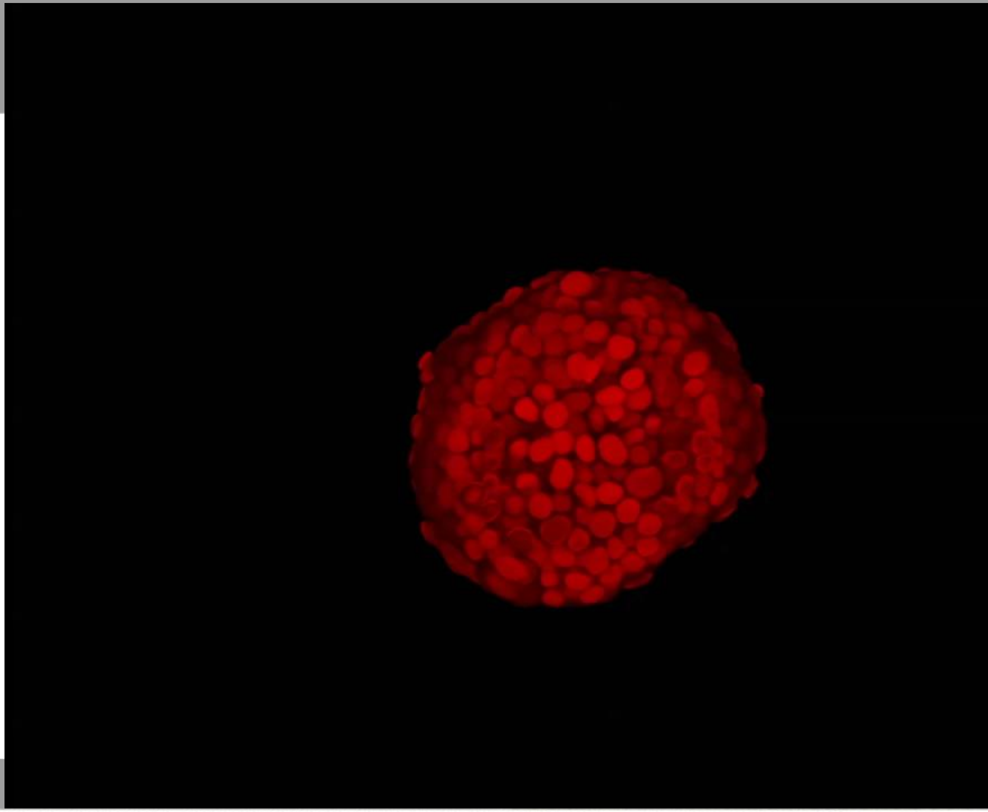
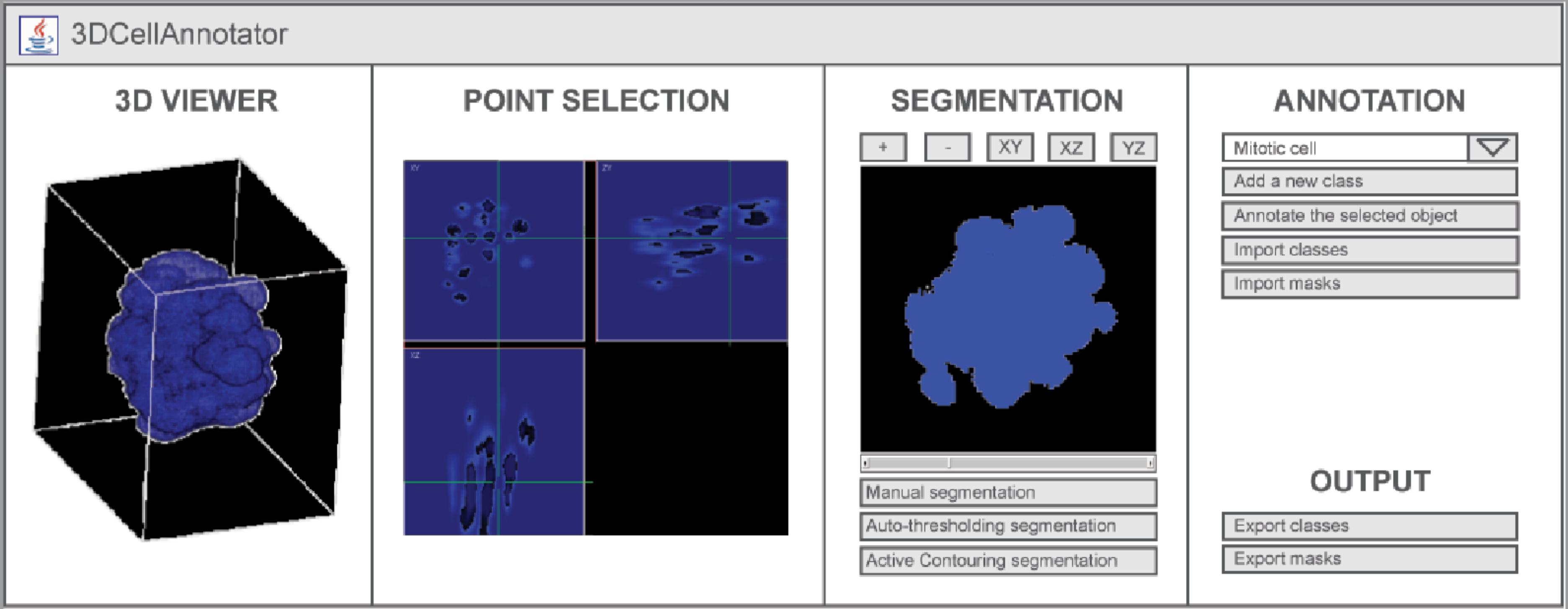


Image analysis

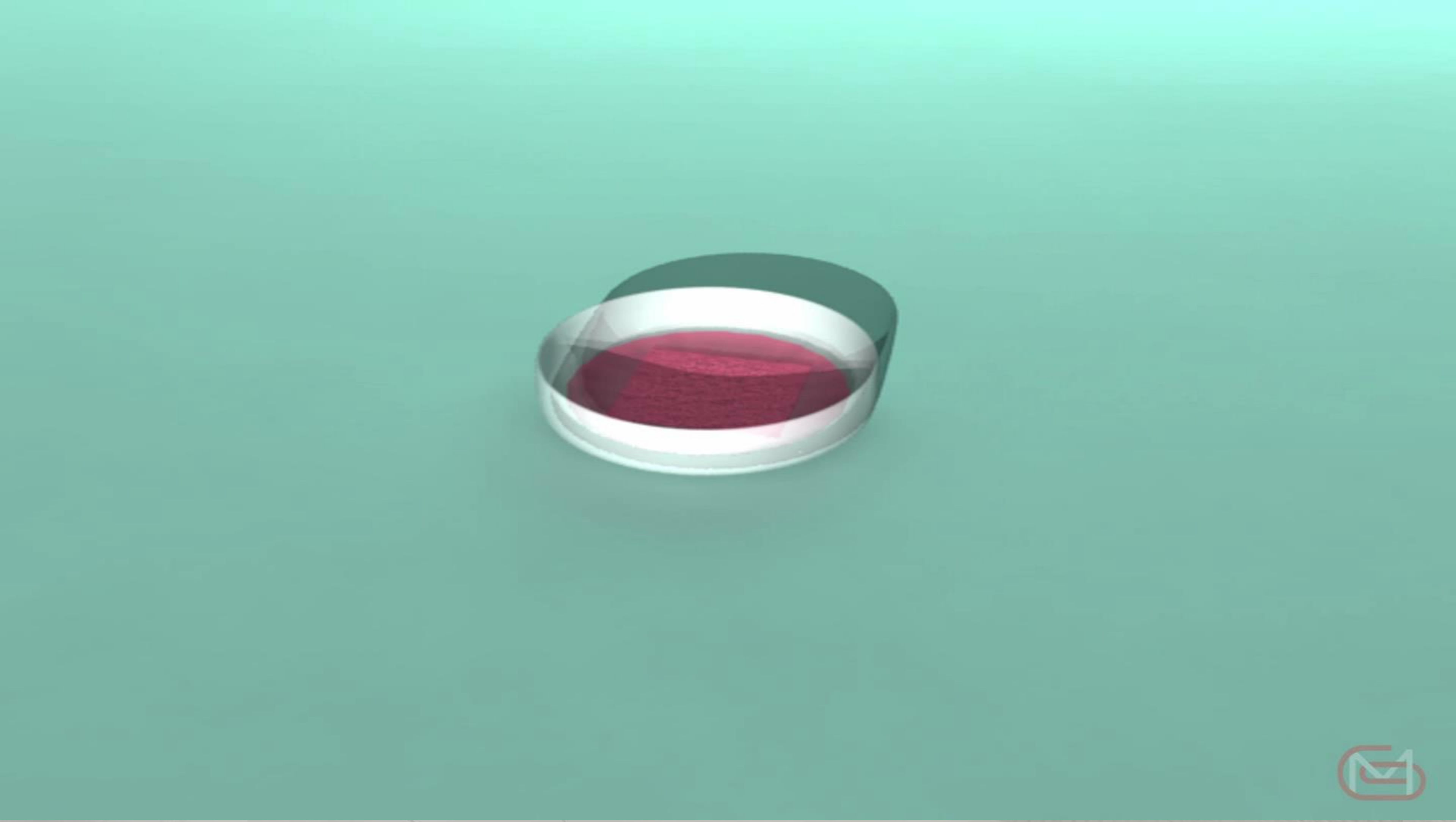
- ***3DCellAnnotator***
- **Active surface-based 3D compatible tool[2]**
- **Segmenting and phenotyping single cells**



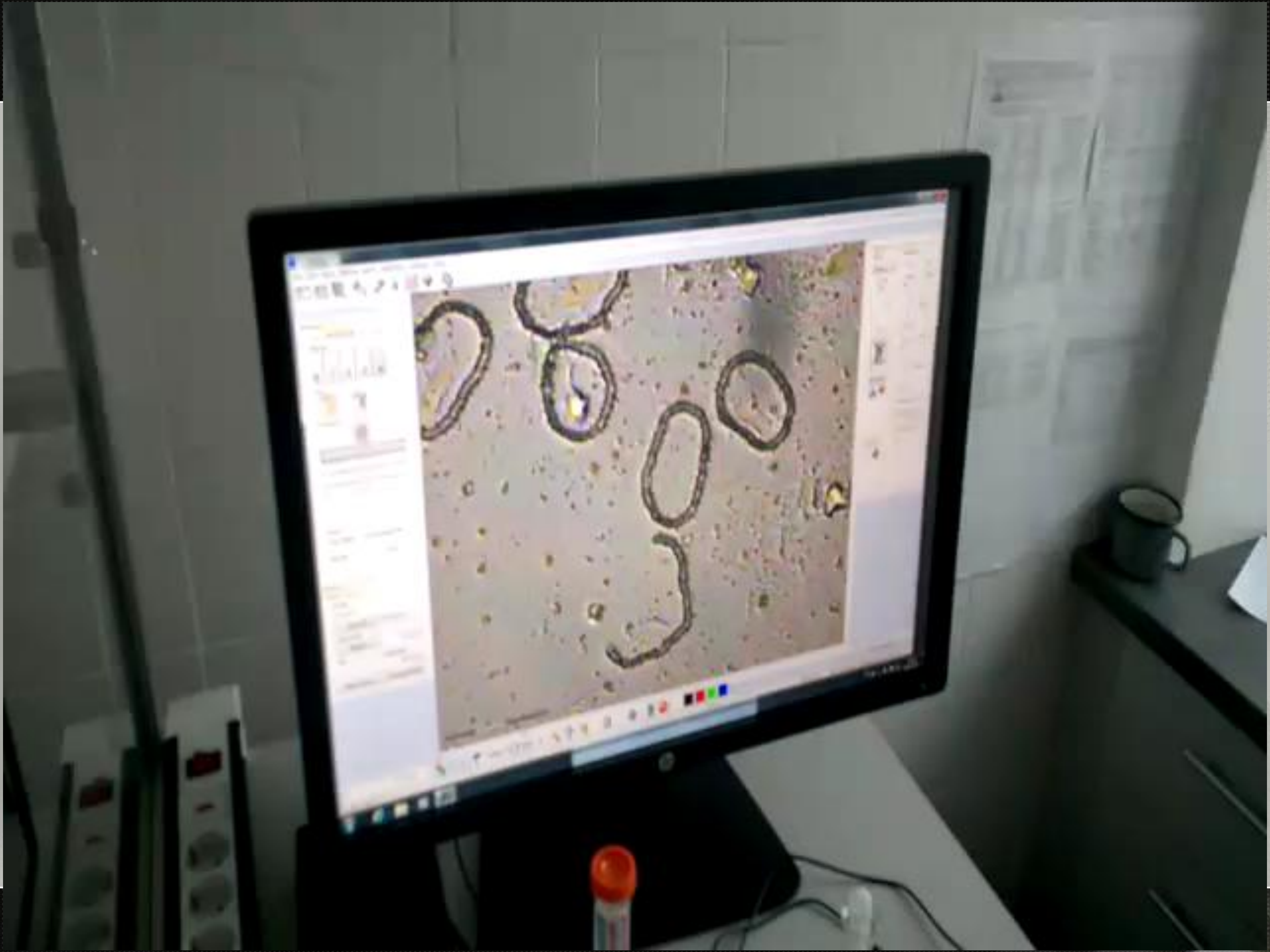
Can we learn more? How?

SINGLE-CELL

CAMI: Computer Aided Microscopy Isolation



Brasko, etal.; ***Nature Communications 2018.***

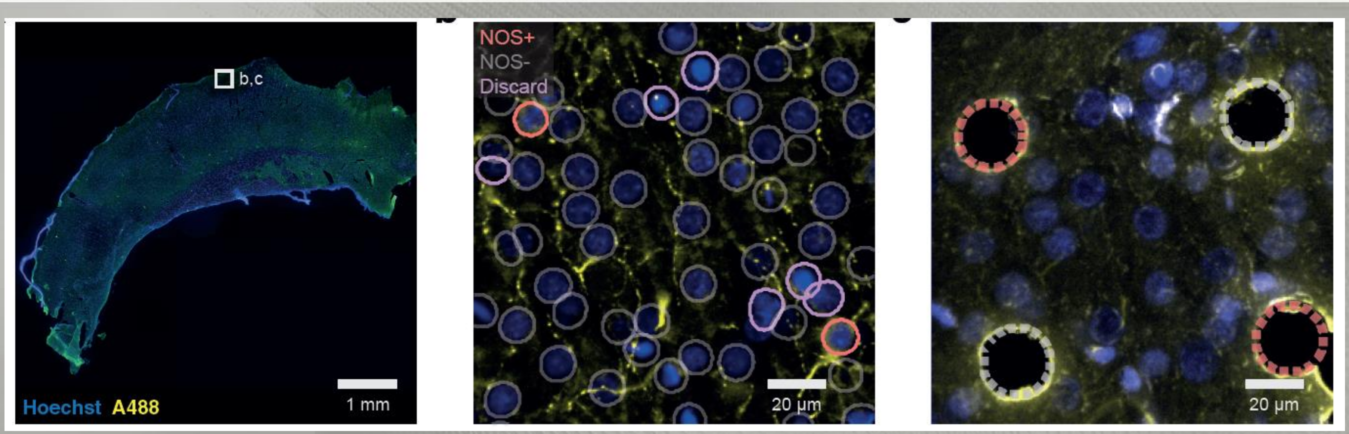


APPLICATIONS

1. Cerebral cortex

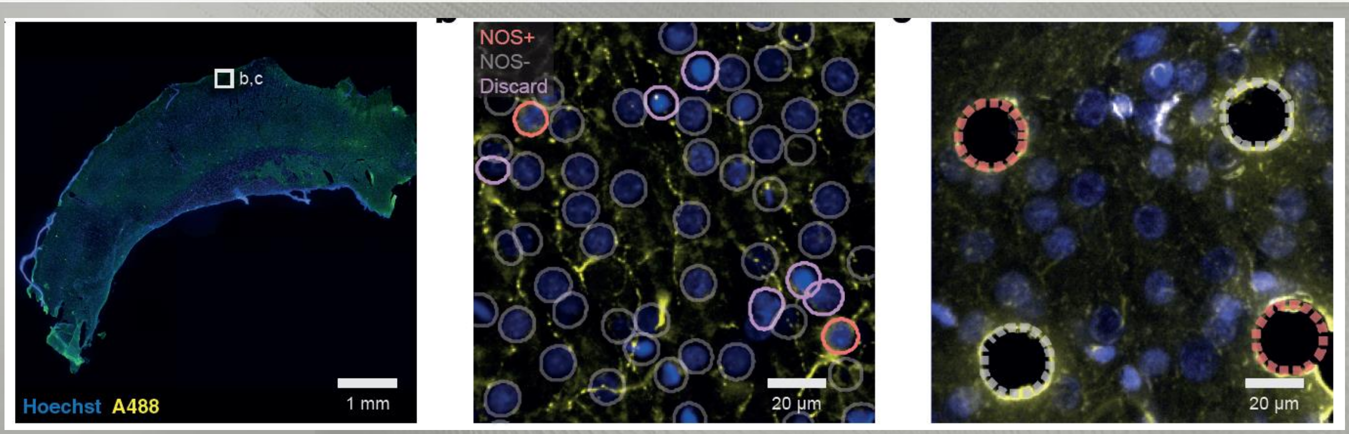
Brasko, etal.; ***Nature Communications 2018.***

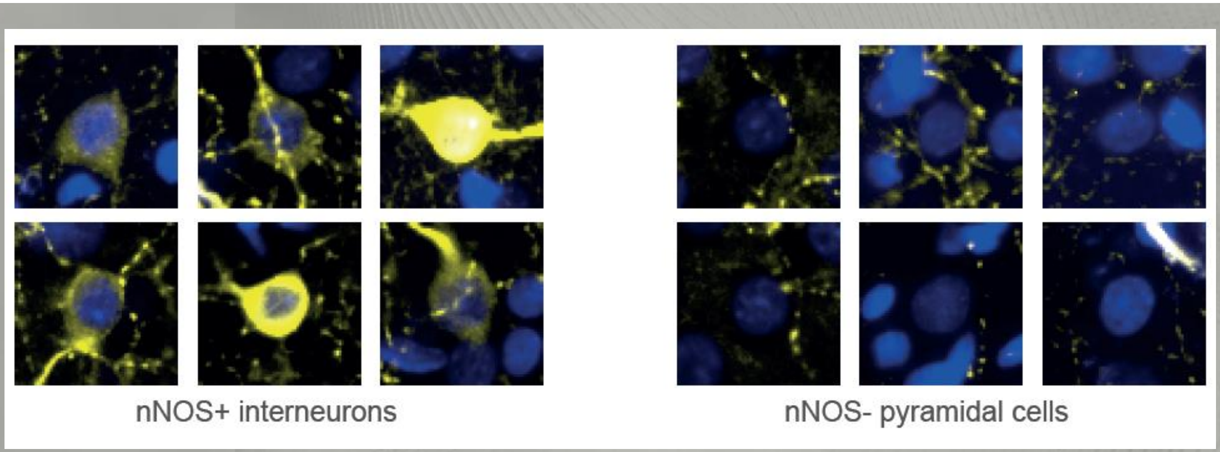
1. Cerebral cortex



Brasko, etal.; ***Nature Communications 2018.***

1. Cerebral cortex

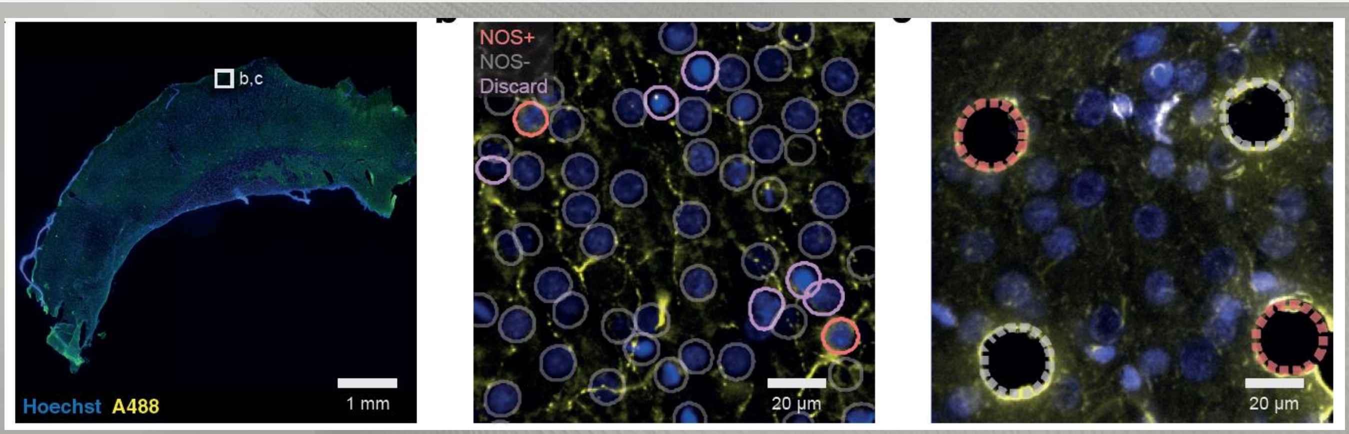


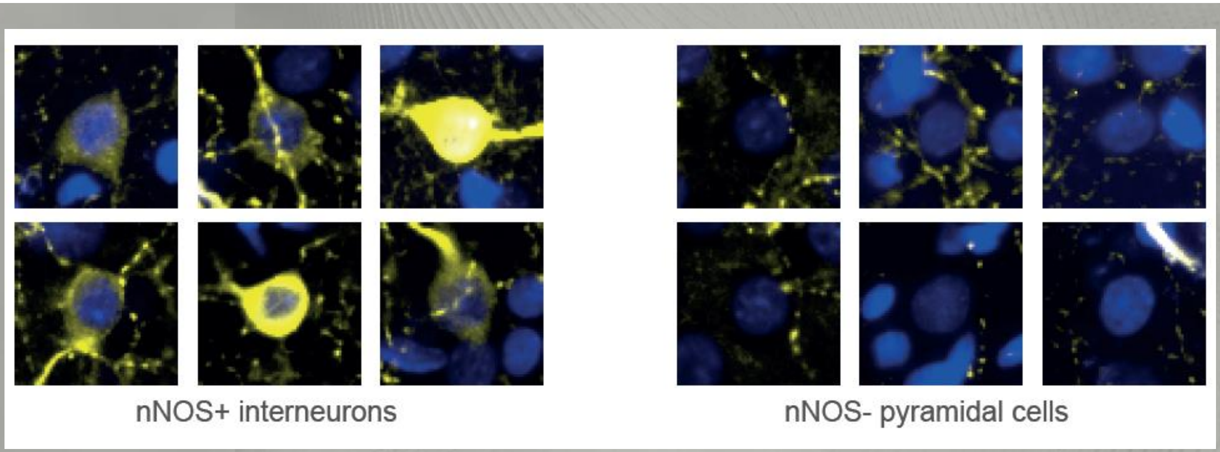

nNOS+ interneurons

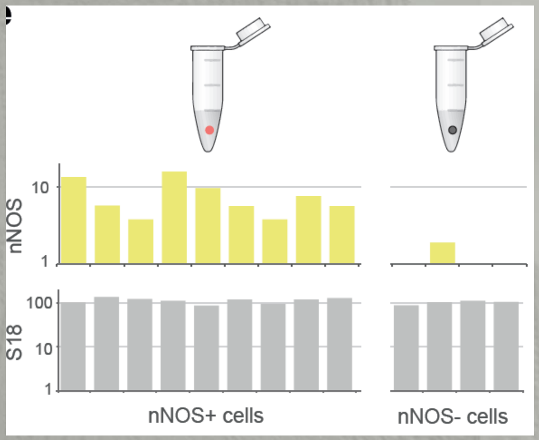
nNOS- pyramidal cells

Brasko, etal.; ***Nature Communications 2018.***

1. Cerebral cortex

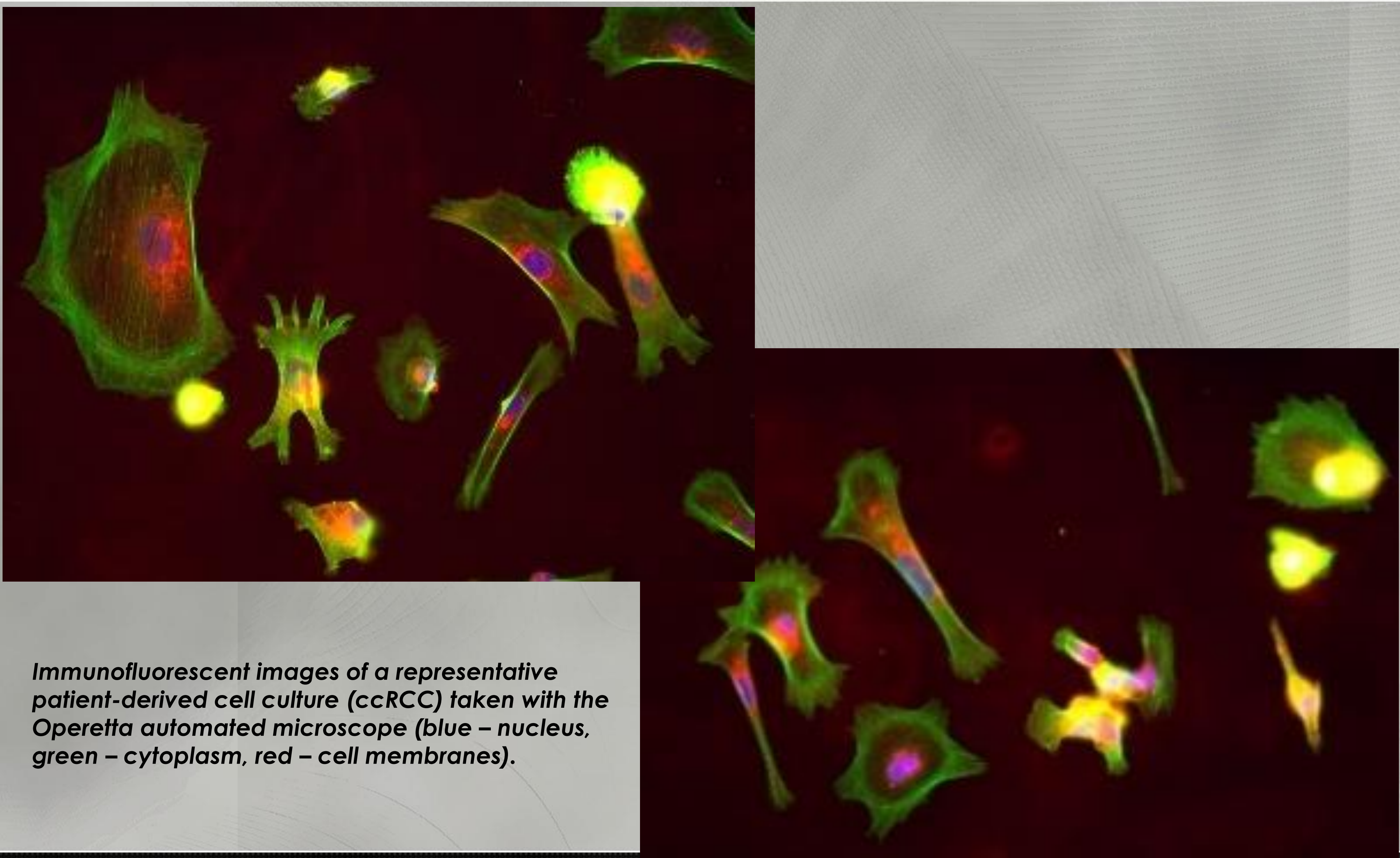






*Brasko, etal.; **Nature Communications 2018.***

3. Resolving ccRCC heterogeneity (report)

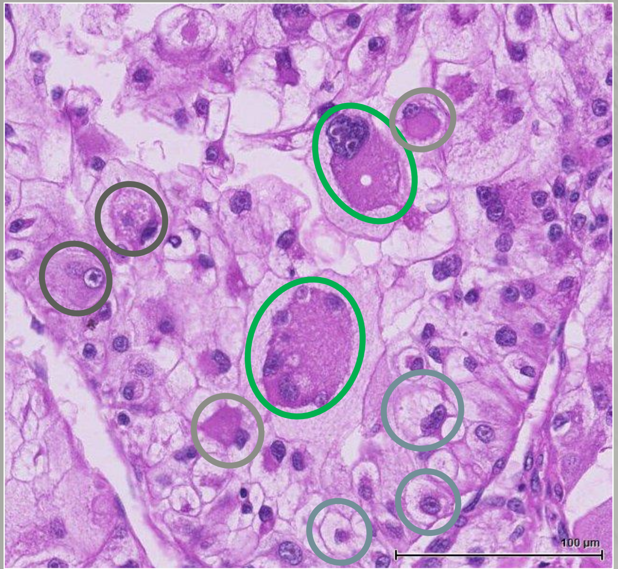


Immunofluorescent images of a representative patient-derived cell culture (ccRCC) taken with the Operetta automated microscope (blue – nucleus, green – cytoplasm, red – cell membranes).

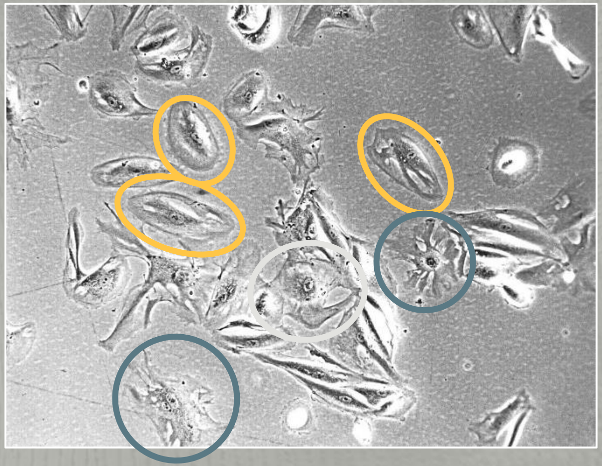
In collaboration with Hella Bolck, Peter Schraml, and Holger Moch; University of Zurich

Patient samples

Primary ccRCC with rhabdoid differentiation

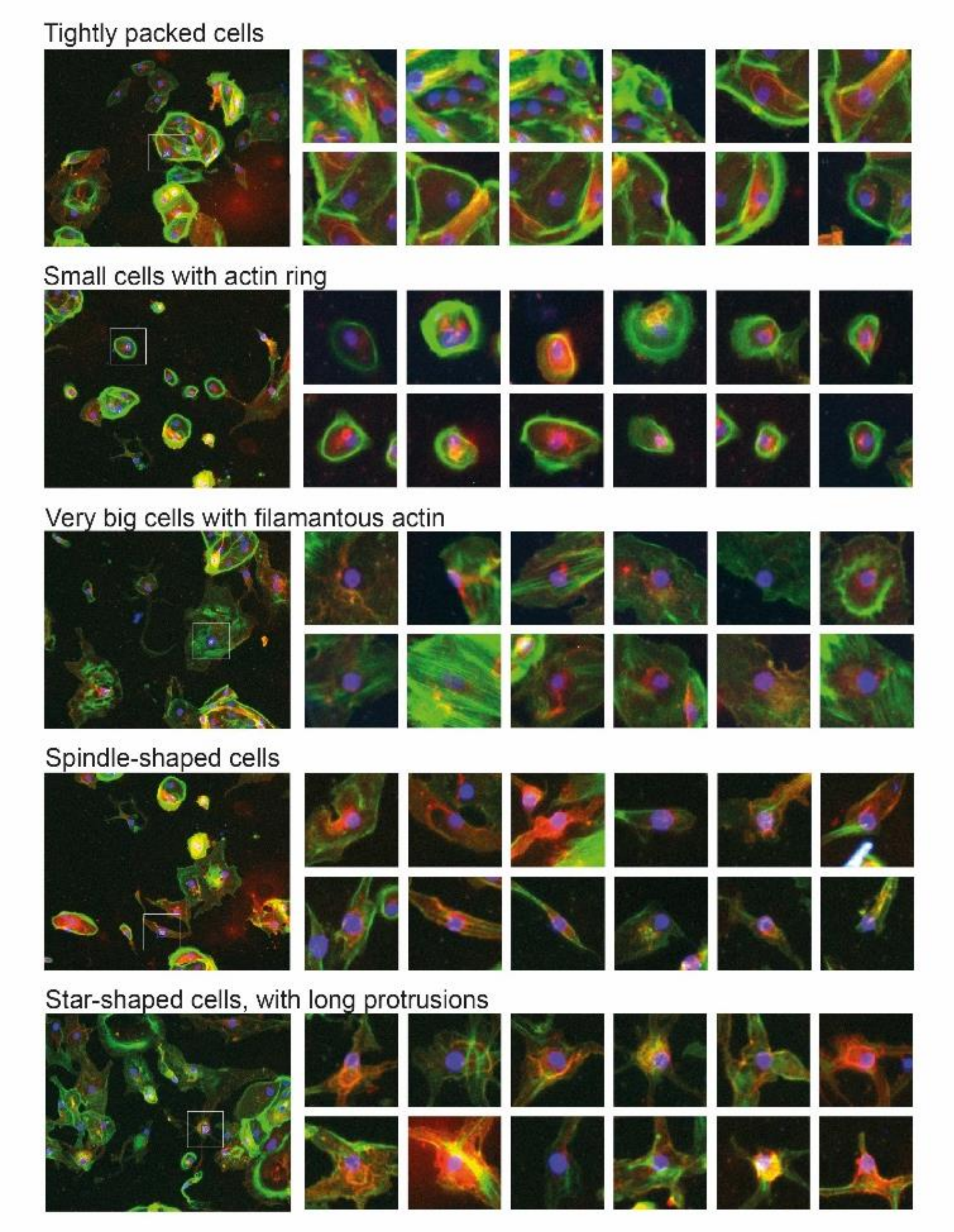


Corresponding patient-derived cell culture



Resolving ccRCC heterogeneity

Patient-derived cell culture (ccRCC) taken with automated microscope (blue – nucleus, green – cytoplasm, red – cell membranes). Five subclasses of morphologically distinct cells were defined and are depicted here.



Tightly packed cells

Small cells with actin ring

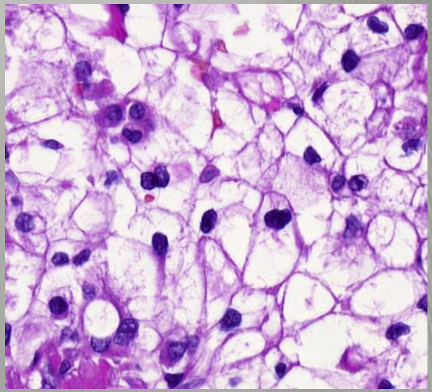
Very big cells with filamantous actin

Spindle-shaped cells

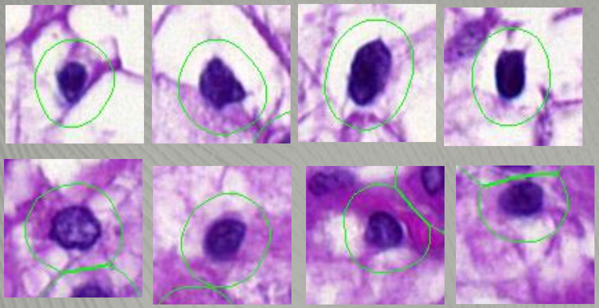
Star-shaped cells, with long protrusions

Definition of morphologically distinct cells in primary ccRCC

grade01_02

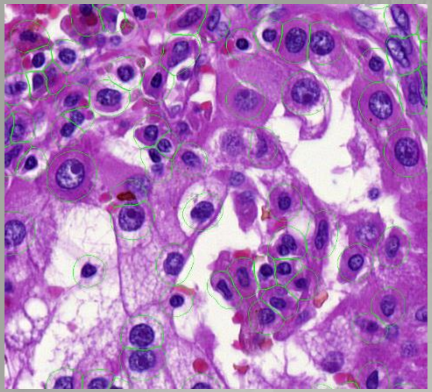


subclasses for grade01_02

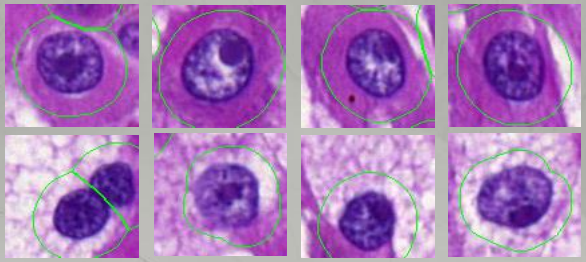


nucleus: grade01_02
cytoplasm: clear or slightly dense

grade03_04

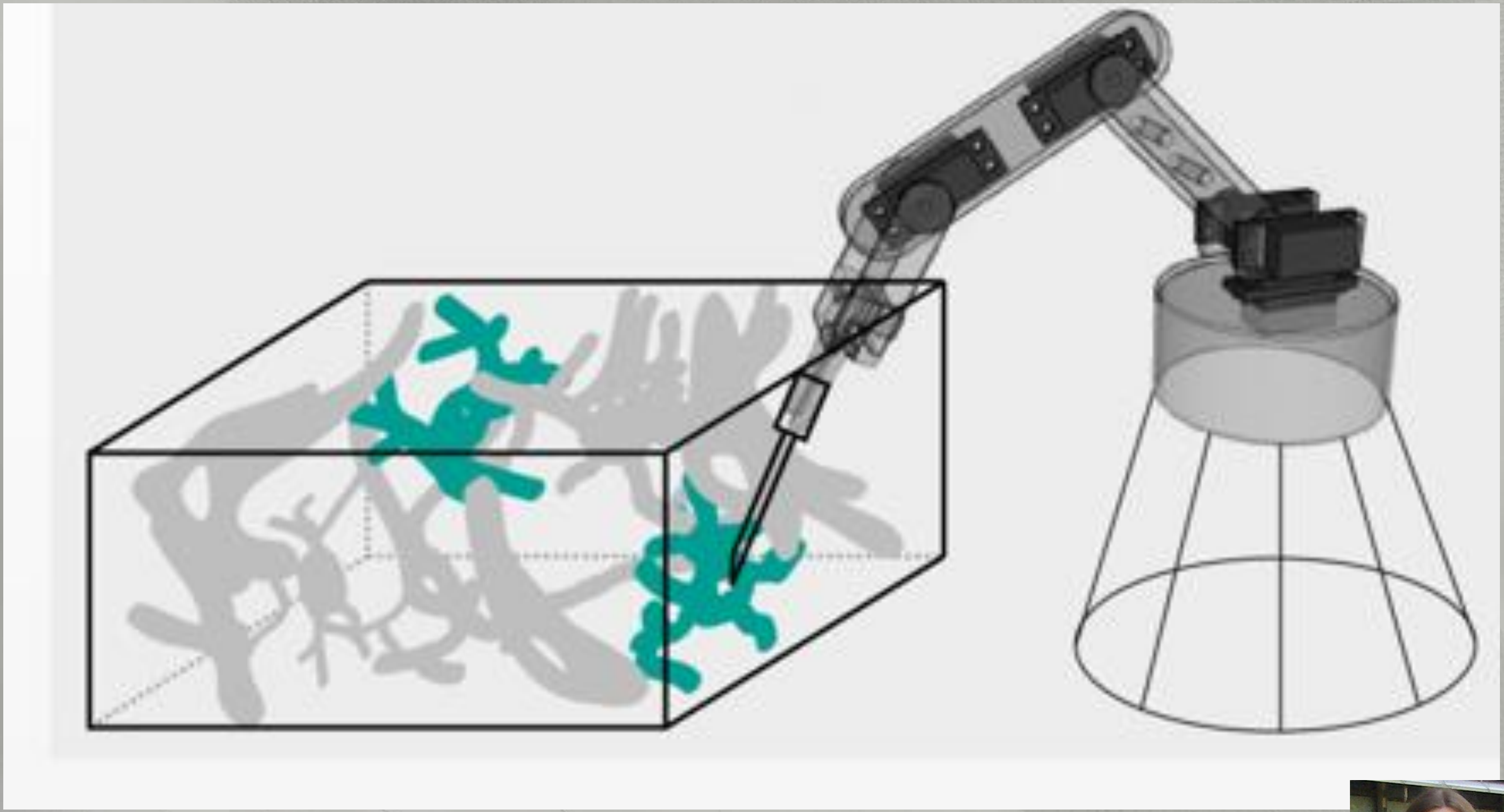


subclasses for grade03_04



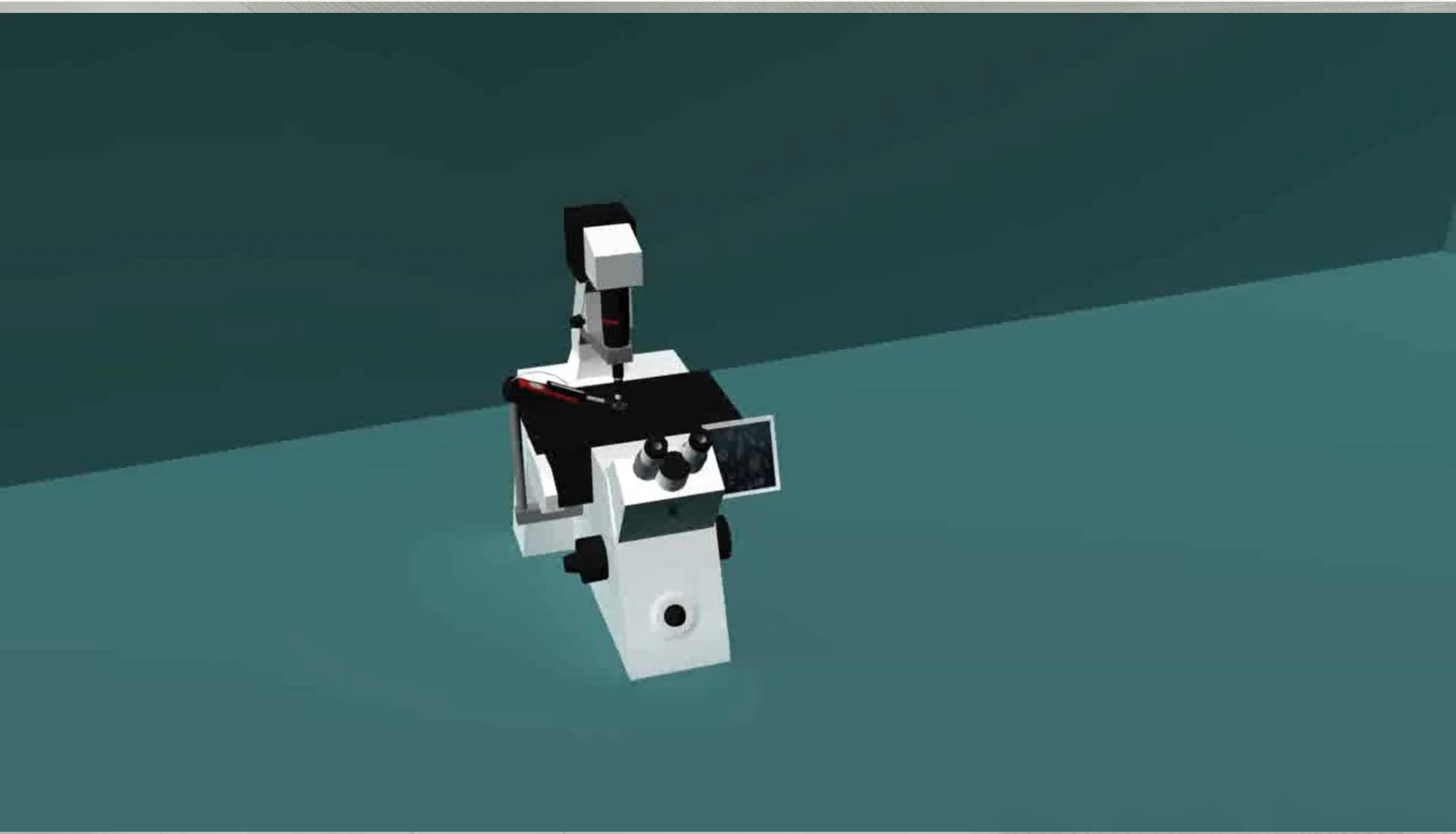
nucleus: grade03_04
cytoplasm: cytoplasms with different color intensities and textures

AutoPatch: single cell system



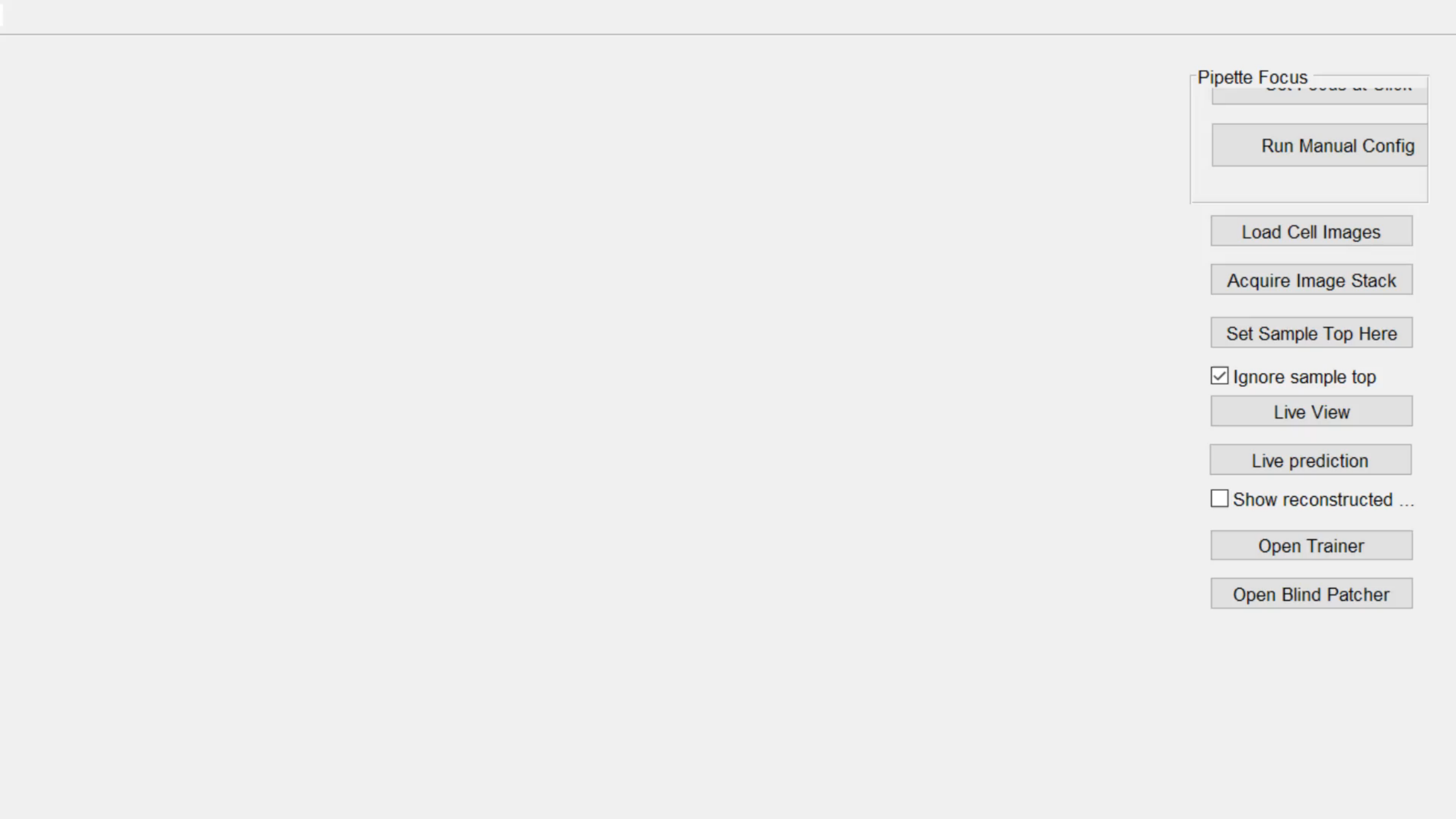
In collaboration with Tamás Gábor, Vígh László

AutoPatch: single cell system



Koós Krisztián etal.

Deep learning search



☑ Ignore sample top

☐ Show reconstructed ...

READ MORE ABOUT THE TOPIC!

Join us! European Cell-based Assays Interest Group

A GUIDE TO DRUG DISCOVERY NATURE REVIEWS | DRUG DISCOVERY

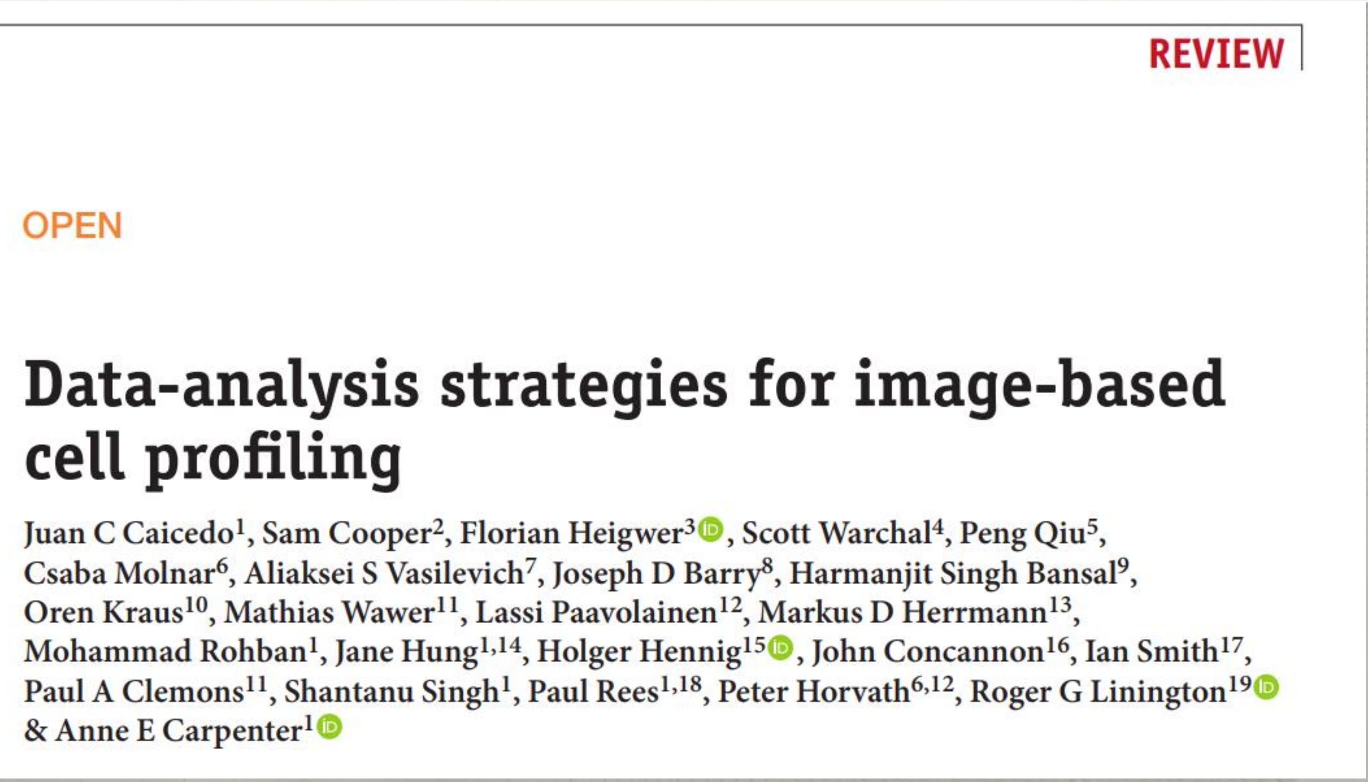
OPINION

Screening out irrelevant cell-based models of disease

Peter Horvath, Nathalie Aulner, Marc Bickle, Anthony Davies, Elaine Del Nery, Daniel Ebner, Maria C. Montoya, Paivi Östling, Vilja Pietiäinen, Leo Price, Spencer L. Shorte, Gerardo Turcatti, Carina von Schantz and Neil O. Carragher

www.eucai.org

Imaging phenomics

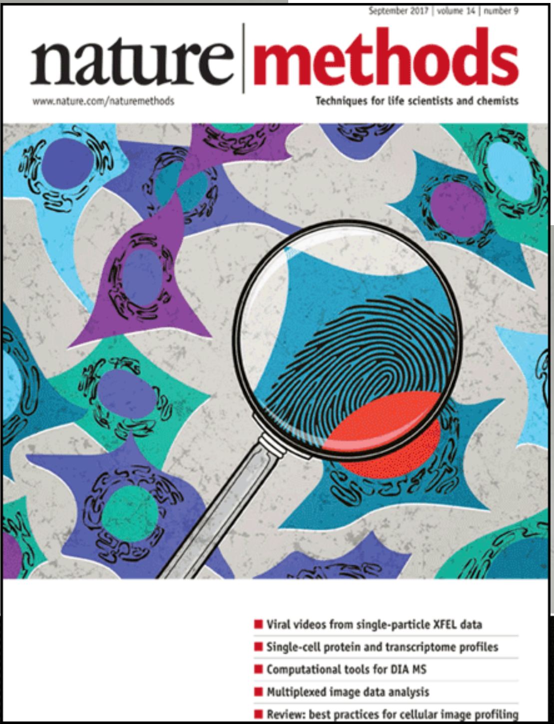


REVIEW

OPEN

Data-analysis strategies for image-based cell profiling

Juan C Caicedo[1], Sam Cooper[2], Florian Heigwer[3], Scott Warchal[4], Peng Qiu[5], Csaba Molnar[6], Aliaksei S Vasilevich[7], Joseph D Barry[8], Harmanjit Singh Bansal[9], Oren Kraus[10], Mathias Wawer[11], Lassi Paavolainen[12], Markus D Herrmann[13], Mohammad Rohban[1], Jane Hung[1,14], Holger Hennig[15], John Concannon[16], Ian Smith[17], Paul A Clemons[11], Shantanu Singh[1], Paul Rees[1,18], Peter Horvath[6,12], Roger G Linington[19] & Anne E Carpenter[1]



September 2017 | volume 14 | number 9

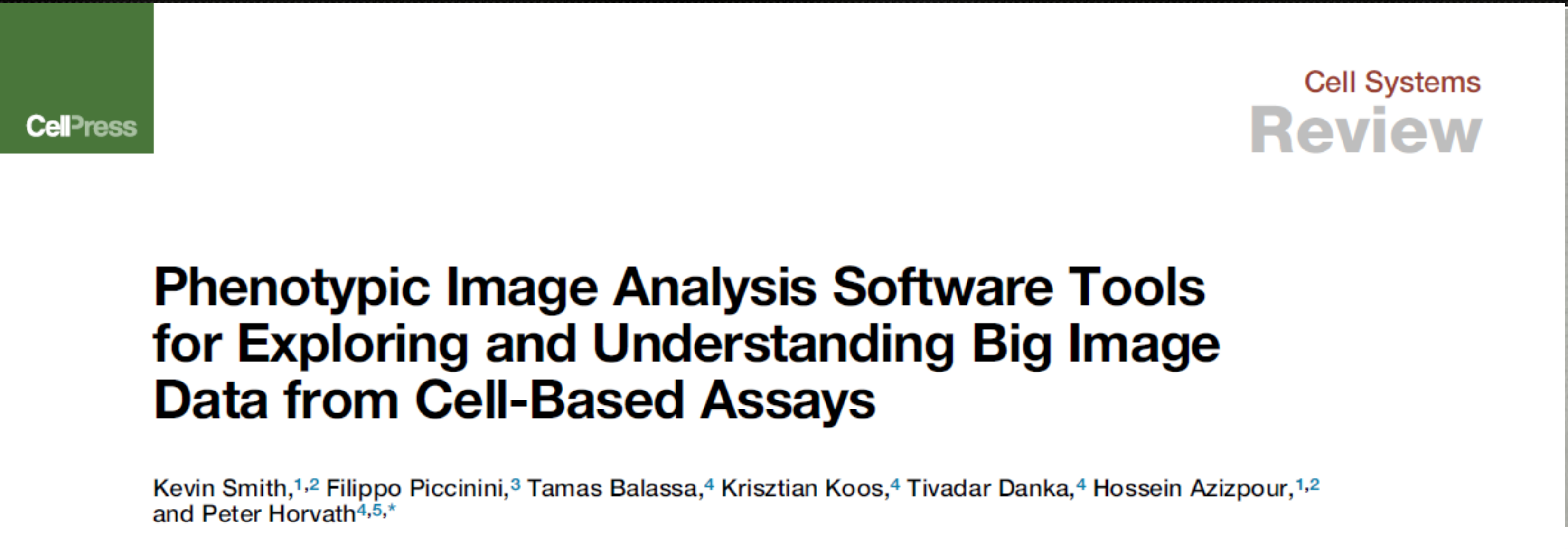
nature methods

www.nature.com/naturemethods

Techniques for life scientists and chemists

- Viral videos from single-particle XFEL data
- Single-cell protein and transcriptome profiles
- Computational tools for DIA MS
- Multiplexed image data analysis
- Review: best practices for cellular image profiling

Machine learning phenomics

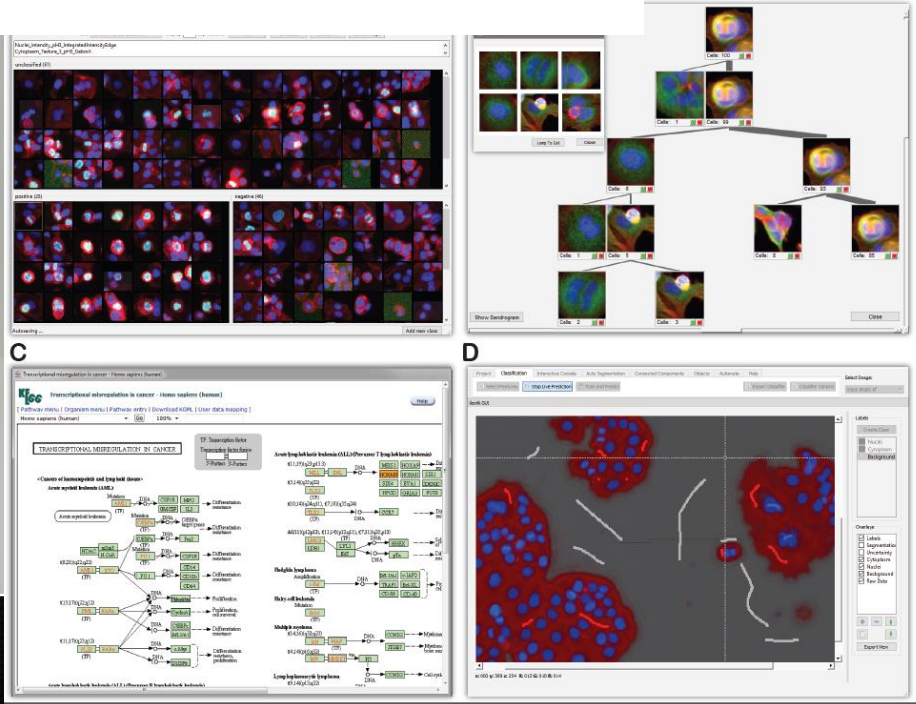


CellPress

Cell Systems
Review

Phenotypic Image Analysis Software Tools for Exploring and Understanding Big Image Data from Cell-Based Assays

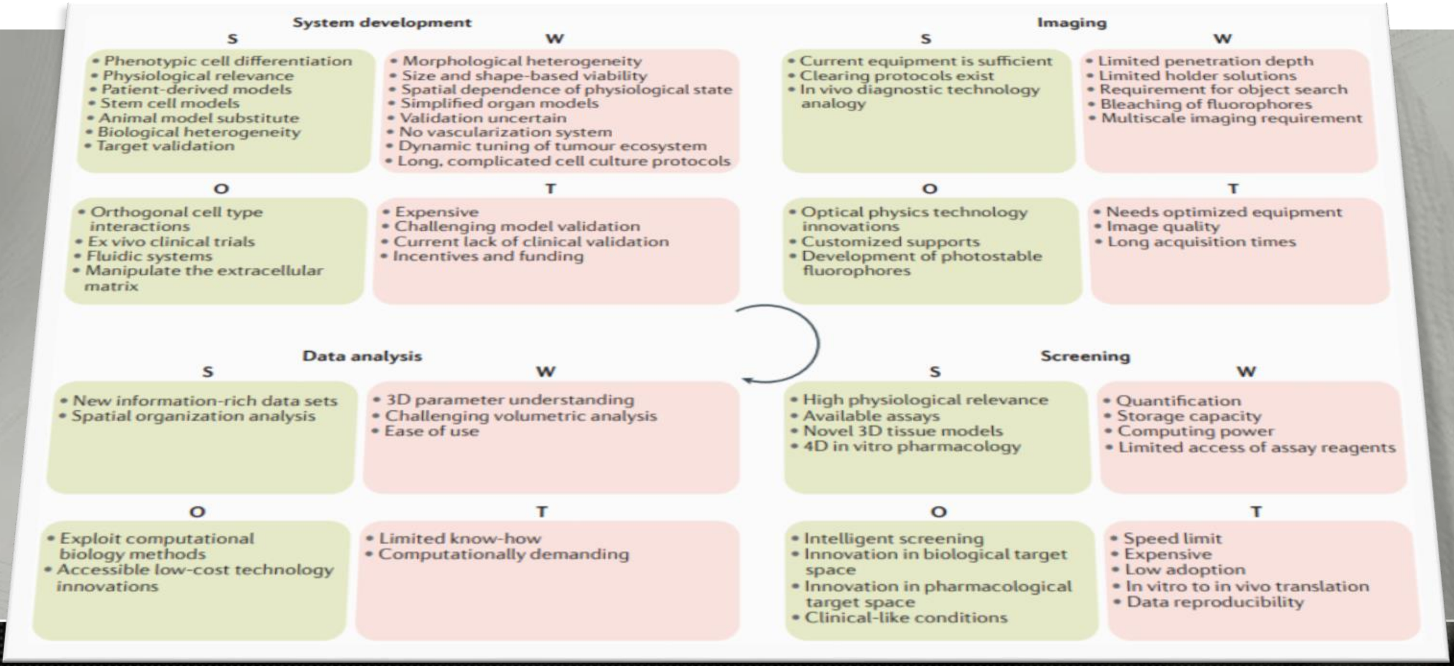
Kevin Smith,[1,2] Filippo Piccinini,[3] Tamas Balassa,[4] Krisztian Koos,[4] Tivadar Danka,[4] Hossein Azizpour,[1,2] and Peter Horvath[4,5,*]



3D high-content screening

Concerns, challenges and promises of high-content analysis of 3D cellular models

Neil Carragher, Filippo Piccinini, Anna Tesei, O. Joseph Trask Jr, Marc Bickle and Peter Horvath



System development

S
- Phenotypic cell differentiation
- Physiological relevance
- Patient-derived models
- Stem cell models
- Animal model substitute
- Biological heterogeneity
- Target validation

W
- Morphological heterogeneity
- Size and shape-based viability
- Spatial dependence of physiological state
- Simplified organ models
- Validation uncertain
- No vascularization system
- Dynamic tuning of tumour ecosystem
- Long, complicated cell culture protocols

O
- Orthogonal cell type interactions
- Ex vivo clinical trials
- Fluidic systems
- Manipulate the extracellular matrix

T
- Expensive
- Challenging model validation
- Current lack of clinical validation
- Incentives and funding

Imaging

S
- Current equipment is sufficient
- Clearing protocols exist
- In vivo diagnostic technology analogy

W
- Limited penetration depth
- Limited holder solutions
- Requirement for object search
- Bleaching of fluorophores
- Multiscale imaging requirement

O
- Optical physics technology innovations
- Customized supports
- Development of photostable fluorophores

T
- Needs optimized equipment
- Image quality
- Long acquisition times

Data analysis

S
- New information-rich data sets
- Spatial organization analysis

W
- 3D parameter understanding
- Challenging volumetric analysis
- Ease of use

O
- Exploit computational biology methods
- Accessible low-cost technology innovations

T
- Limited know-how
- Computationally demanding

Screening

S
- High physiological relevance
- Available assays
- Novel 3D tissue models
- 4D in vitro pharmacology

W
- Quantification
- Storage capacity
- Computing power
- Limited access of assay reagents

O
- Intelligent screening
- Innovation in biological target space
- Innovation in pharmacological target space
- Clinical-like conditions

T
- Speed limit
- Expensive
- Low adoption
- In vitro to in vivo translation
- Data reproducibility

Thank you

__BIOMAG – FIMM__

FiDiPro Host

__BIOMAG – HU__

Kutay Group
Smith Group
Helenius Group
Pal & Papp Groups
Lundin Group
Kallioniemi group
Ikonen Group
Moch/Schraml Group
Haracska Group
Tamas Group
Vigh Group

Lendület program

EMBO excellence in life sciences

SZÉCHENYI TERV
NAP
HUNGARIAN BRAIN RESEARCH PROGRAM
NEMZETI AGYKUTATÁSI PROGRAM

FEBS

SZÉCHENYI 2020

FIDIPRO Finland Distinguished Professor Programme

MAGYAR TUDOMÁNYOS AKADÉMIA 1825

MARIE CURIE ACTIONS