Analysing Cellular Microscopy Images

Dmytro Fishman

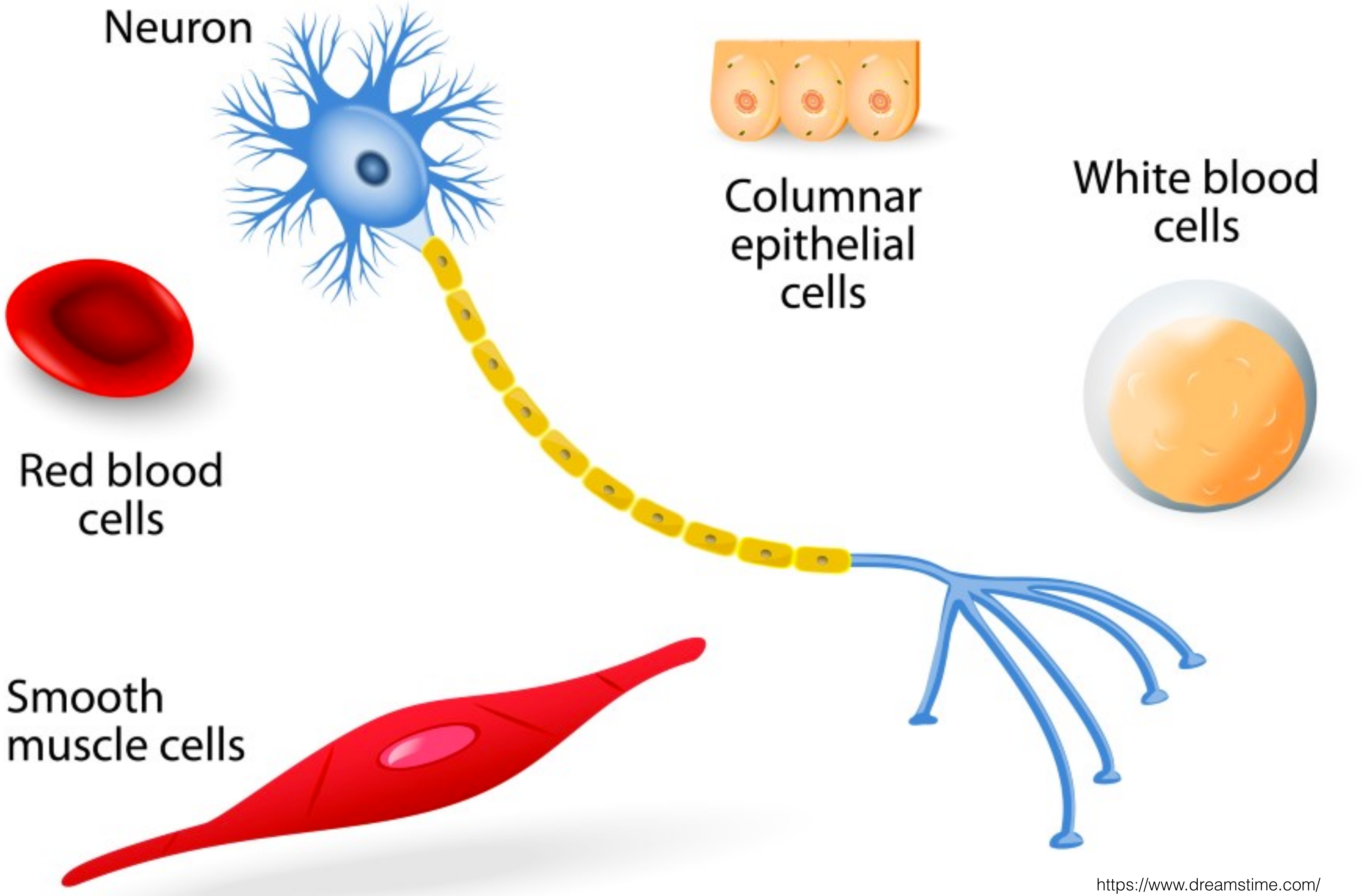
BIIT

COMPUTATIONAL NEUROSCIENCE
Research group at the University of Tartu

TARTU ÜLIKOOL · UNIVERSITAS TARTUENSIS · 1632

PerkinElmer®
For the Better

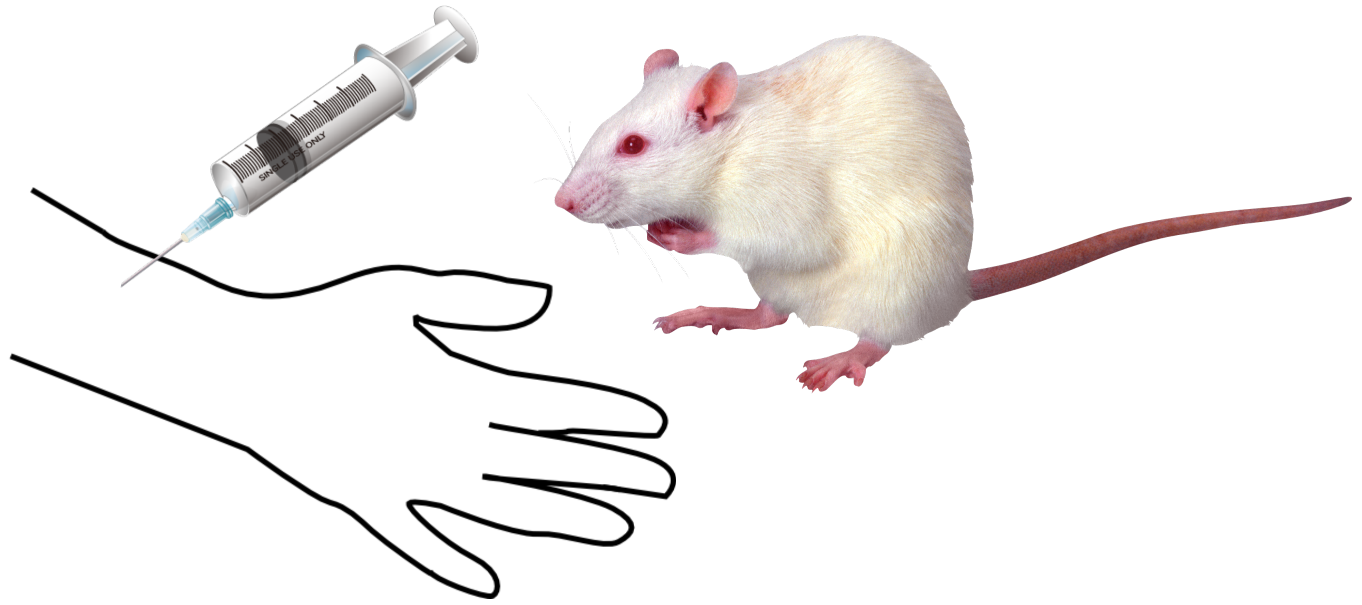
Human Cells



https://www.dreamstime.com/

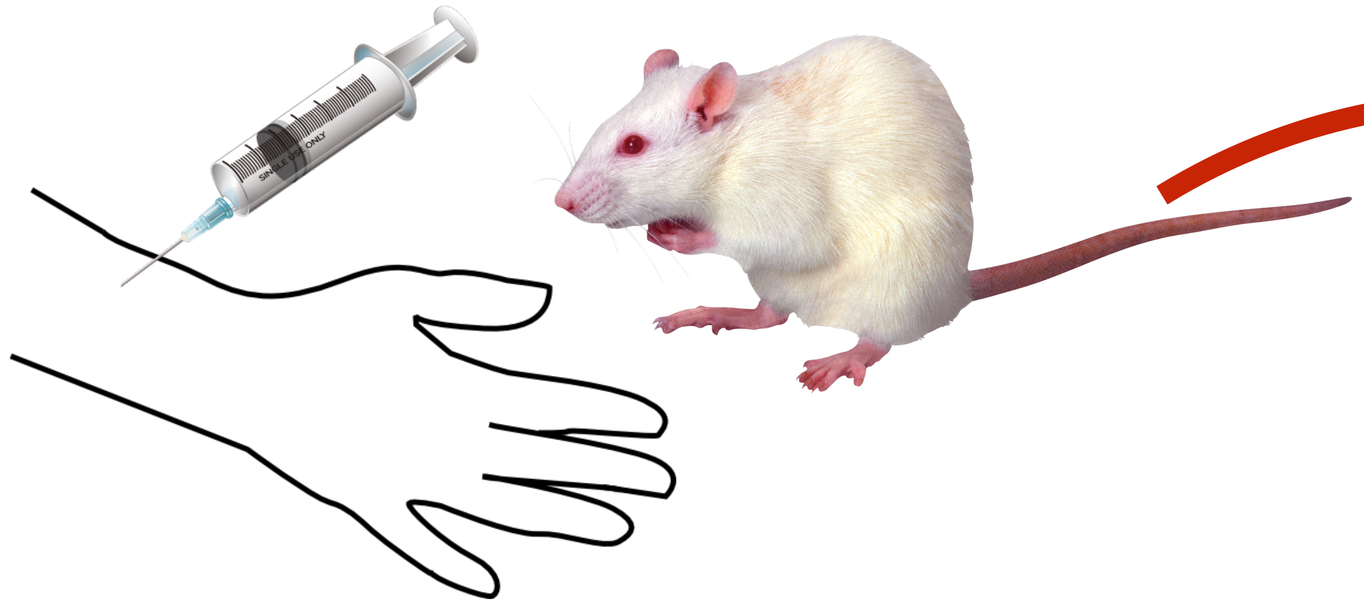
Producing Microscopy imaging

Producing Microscopy imaging



Extraction

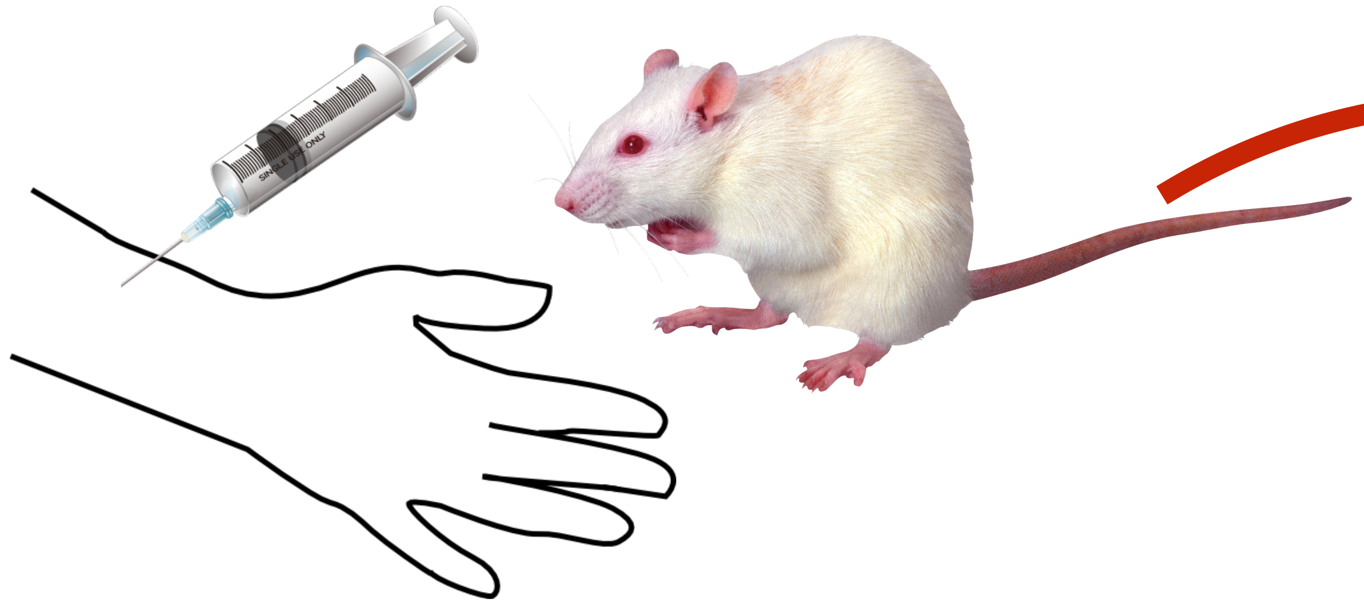
Producing Microscopy imaging



Extraction

Seeding

Producing Microscopy imaging

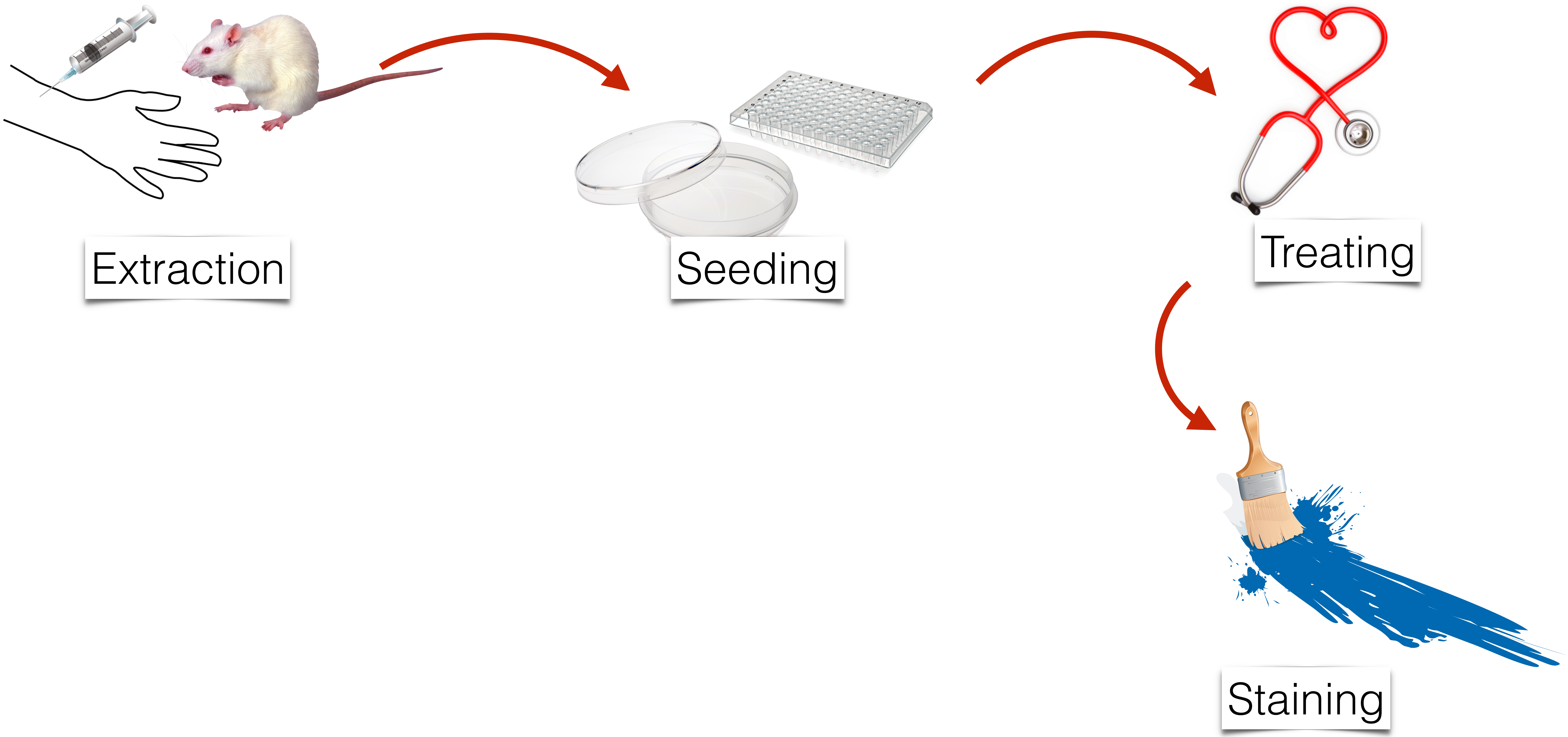


Extraction

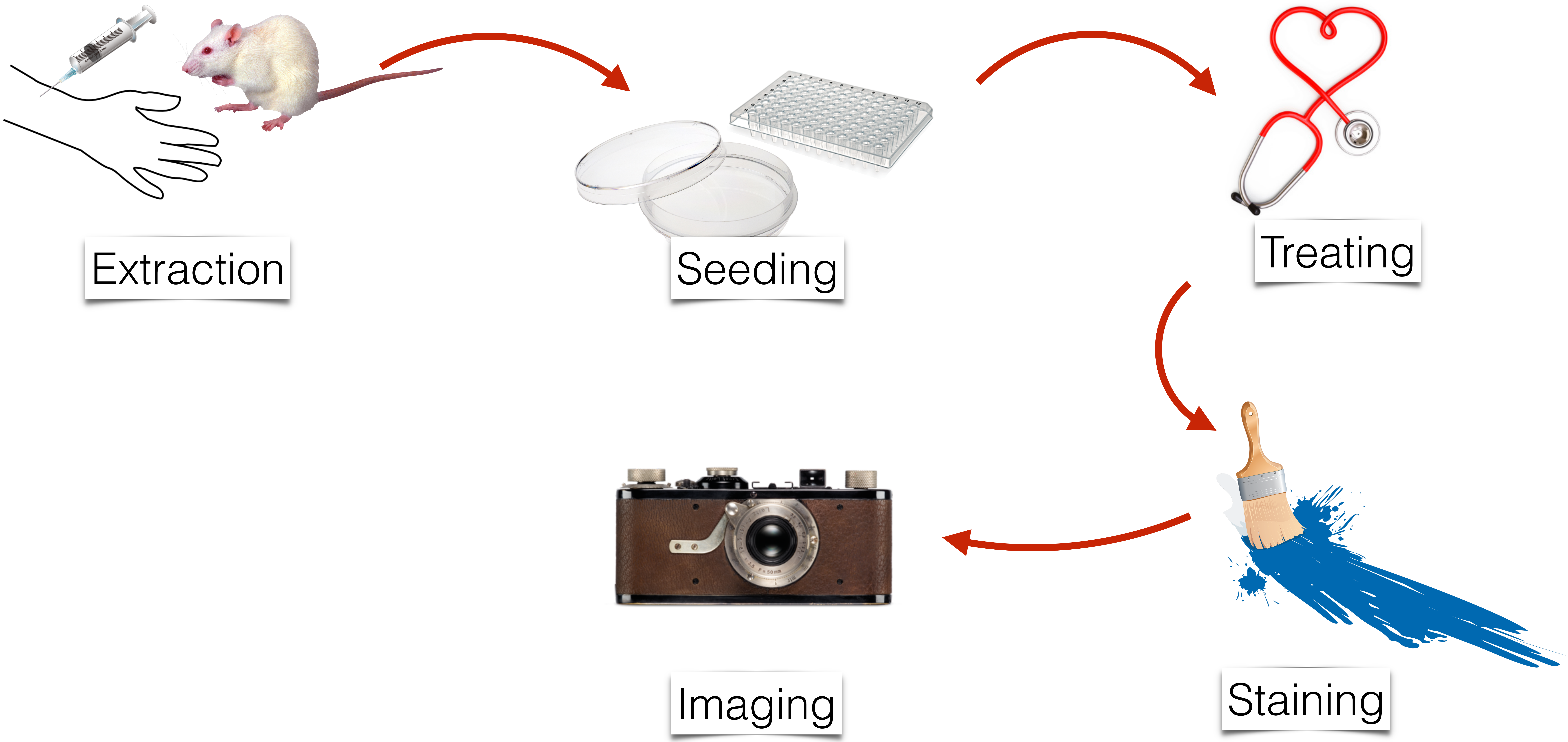
Seeding

Treating

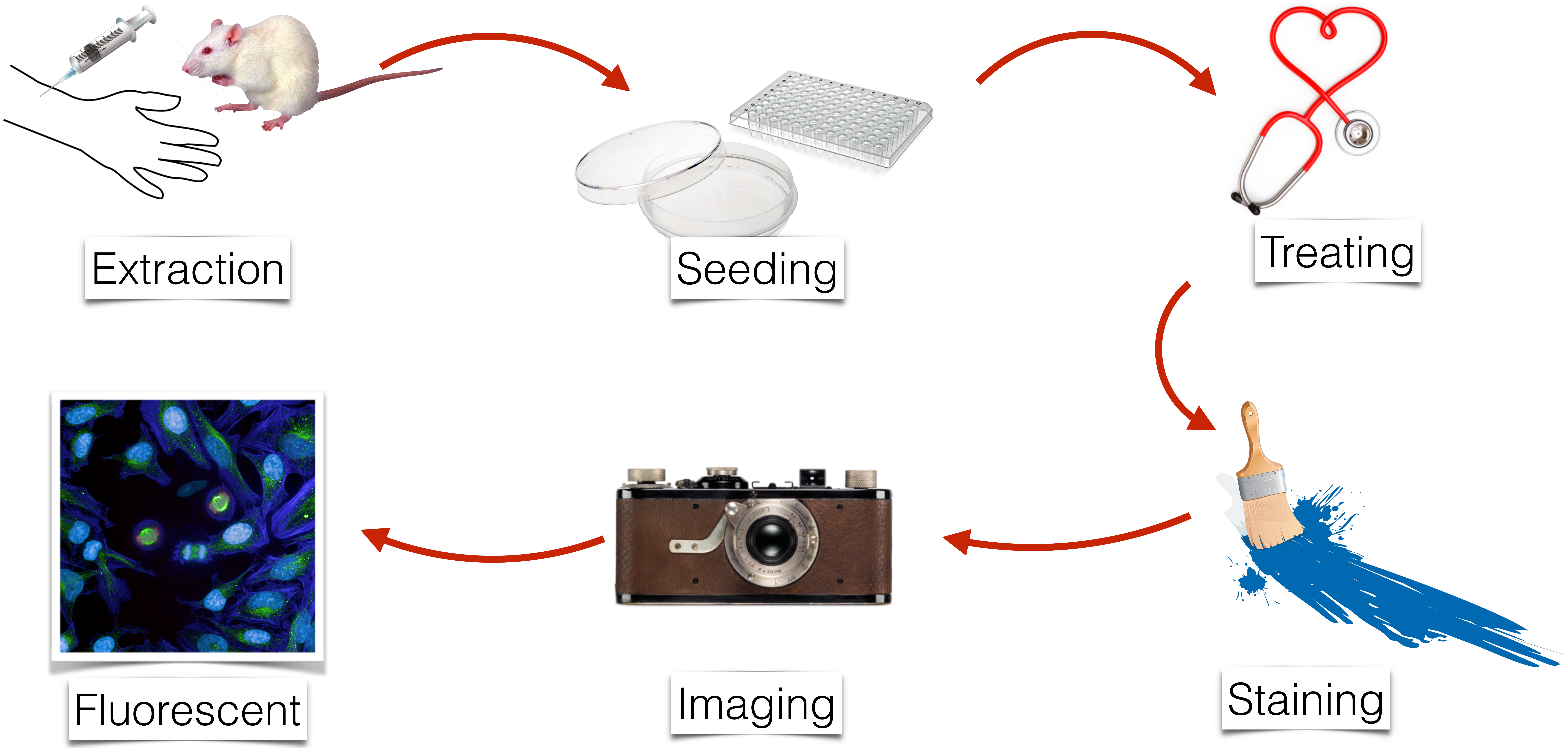
Producing Microscopy imaging



Producing Microscopy imaging

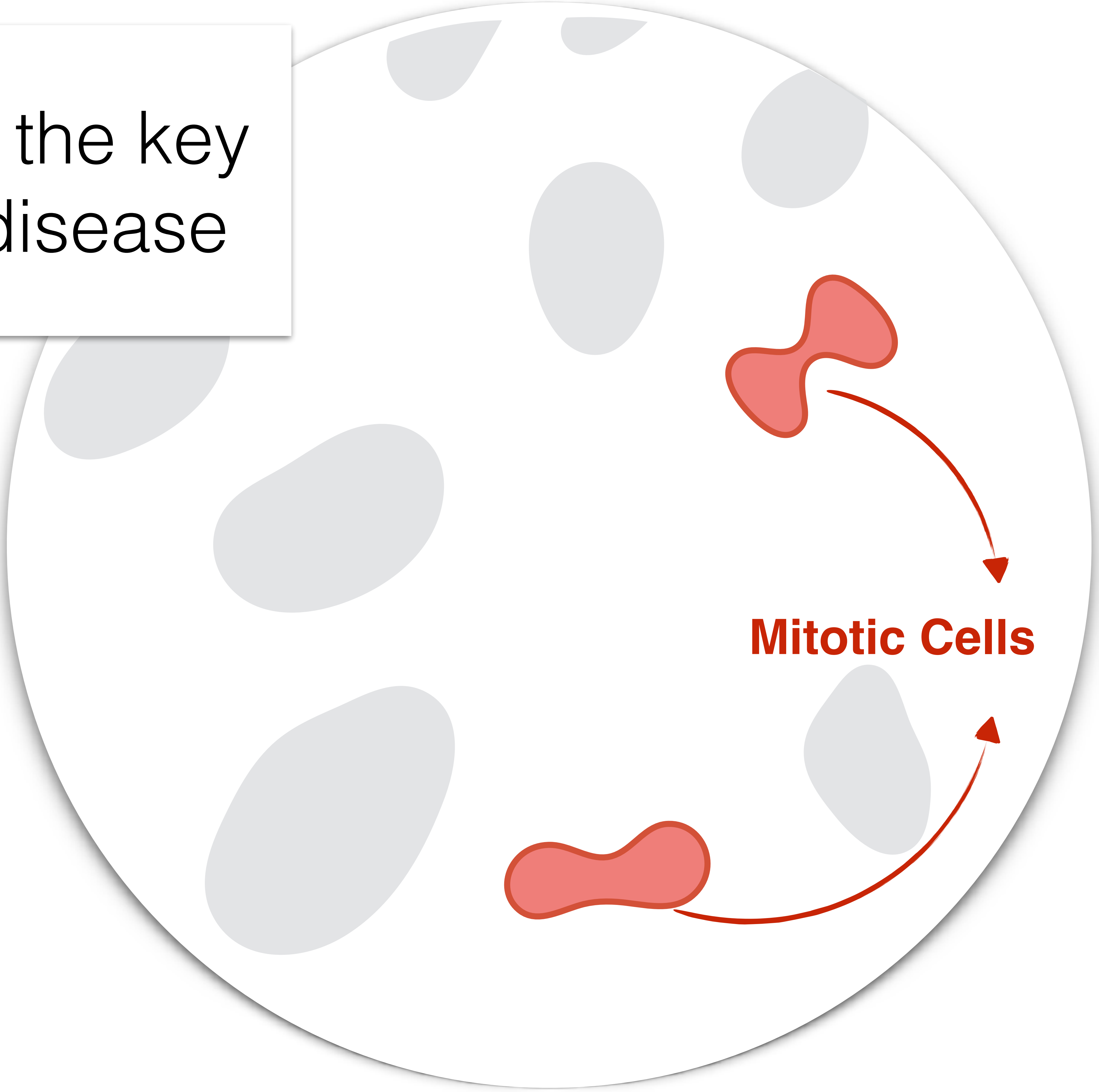


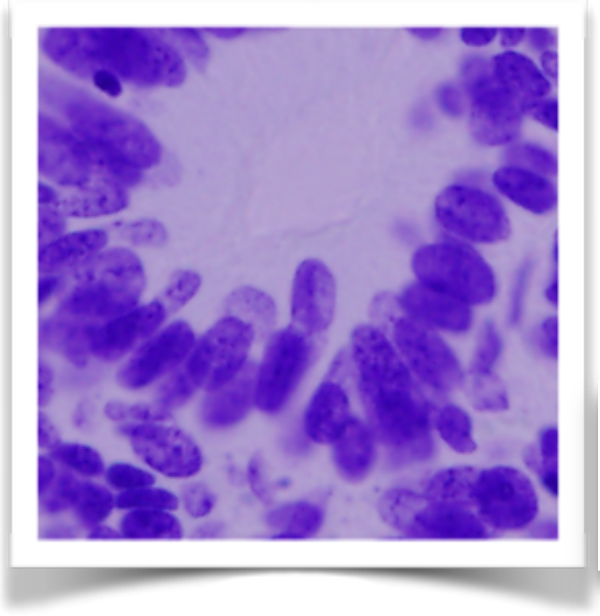
Producing Microscopy imaging



Breast cancer diagnostics

Mitotic cell count is one of the key diagnostic markers of the disease





Histology images

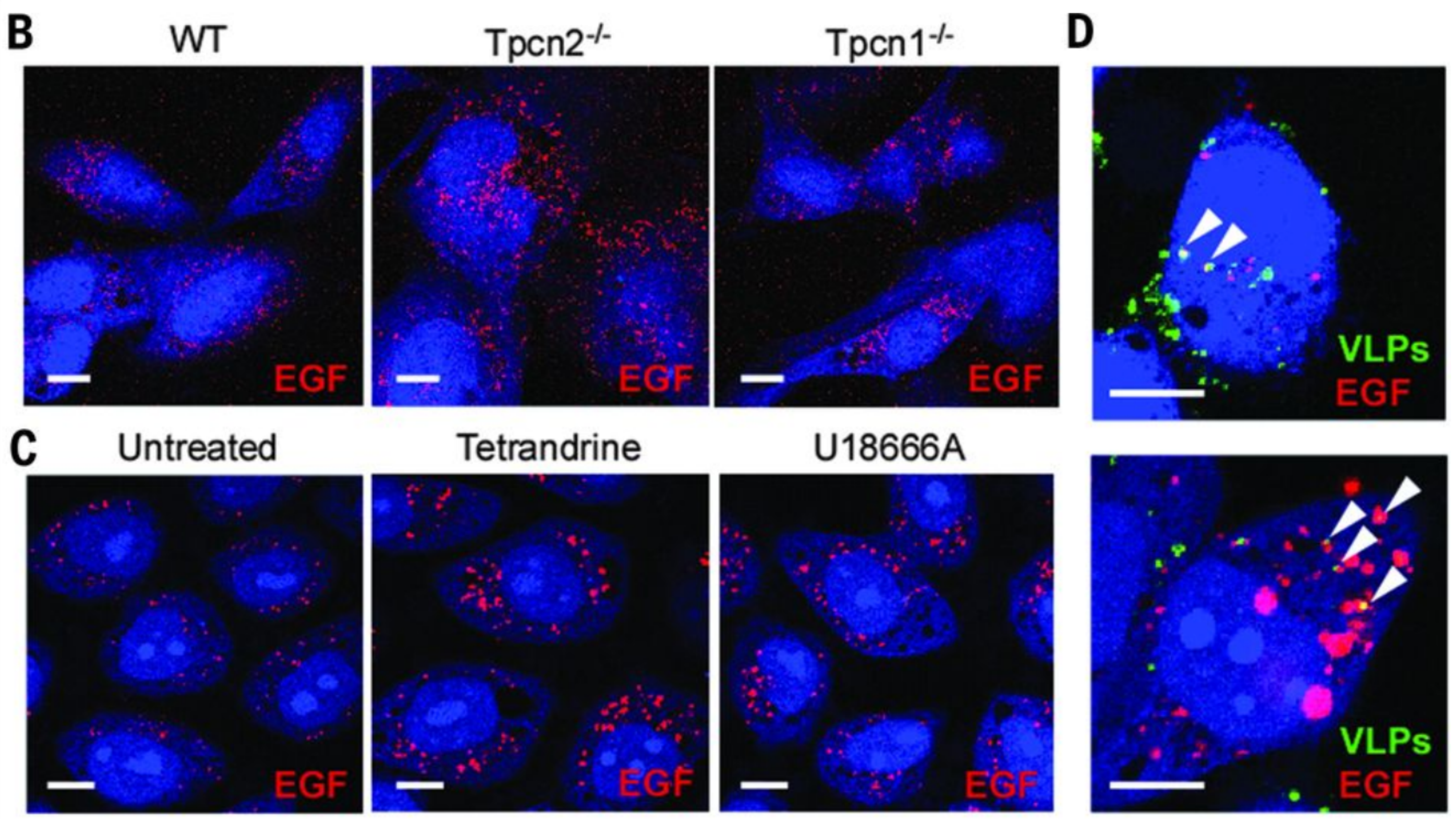
Ebola virus vaccine

REPORT

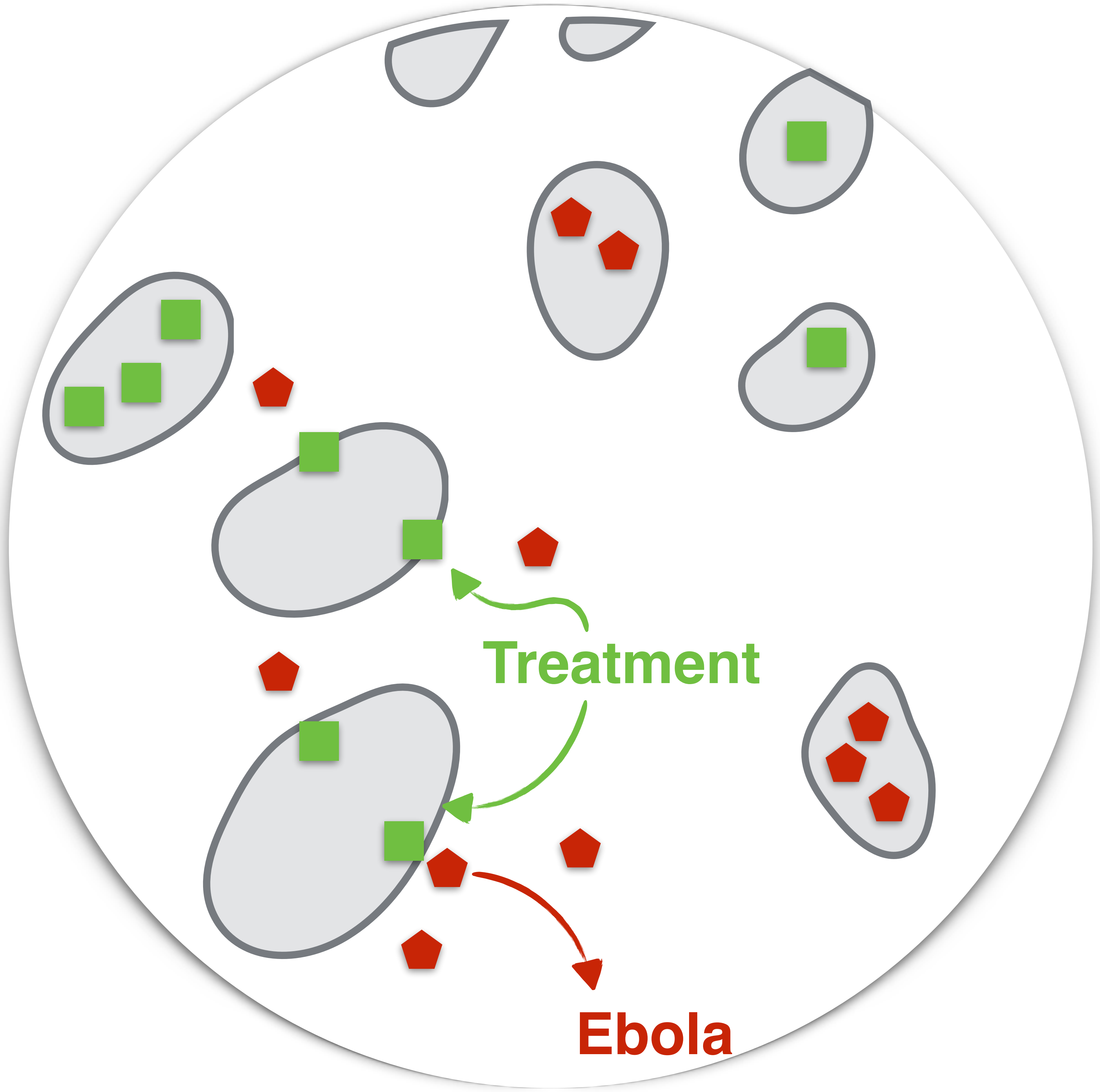
Two-pore channels control Ebola virus host cell entry and are drug targets for disease treatment

Yasuteru Sakurai[1], Andrey A. Kolokoltsov[2], Cheng-Chang Chen[3], Michael W. Tidwell[4], William E. Bauta[4], Norbert Klugbauer...
+ See all authors and affiliations

Science 27 Feb 2015:
Vol. 347, Issue 6225, pp. 995-998
DOI: 10.1126/science.1258758

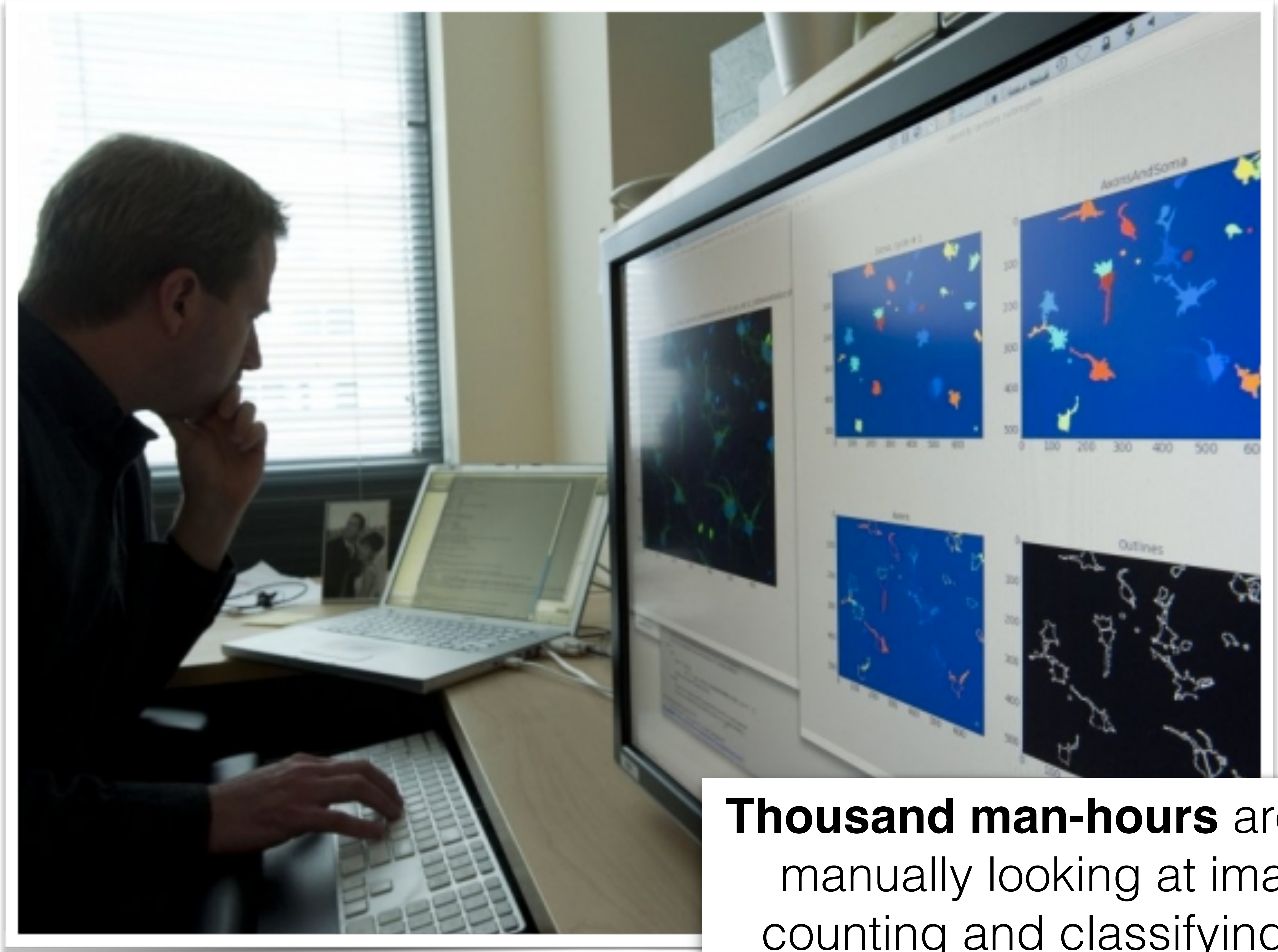


Fluorescent images



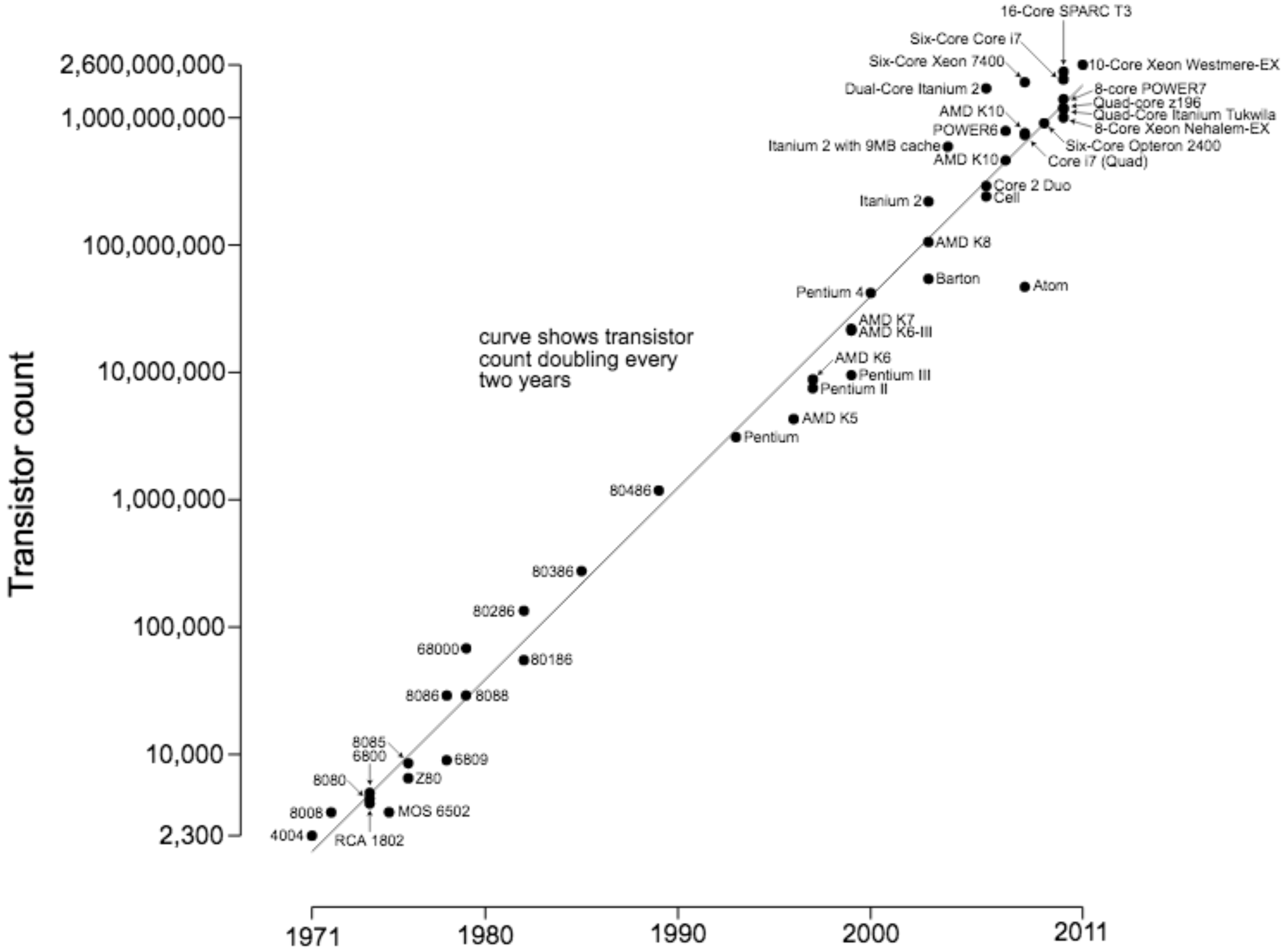
Current Best Method for Microscopy Image Analysis?

Current Best Method for Microscopy Image Analysis



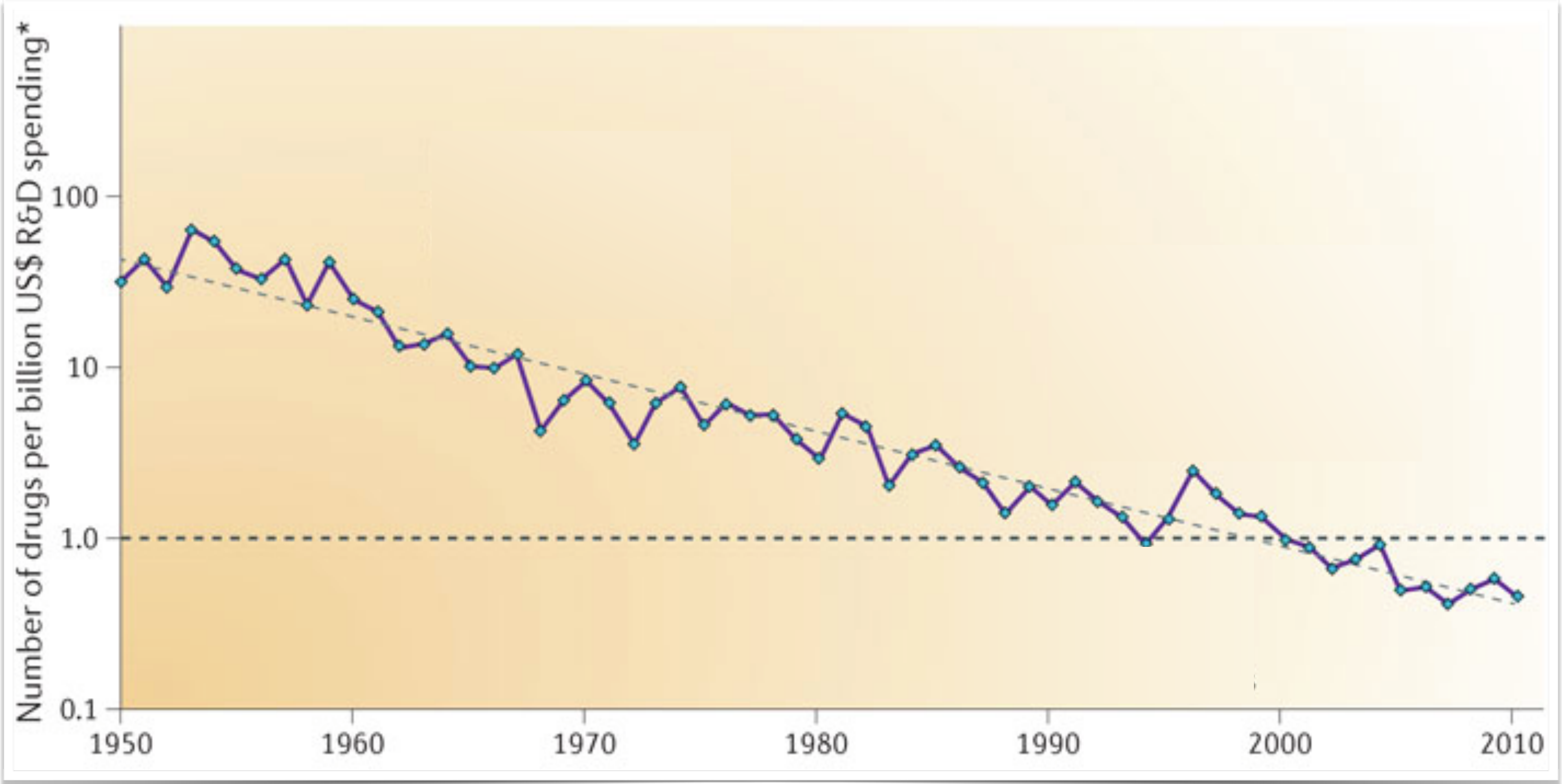
Thousand man-hours are spent manually looking at images, counting and classifying cells

Microprocessor Transistor Counts 1971-2011 & Moore's Law



https://phys.org/news/2015-04-silicon-valley-years-law.html

Eroom’s law

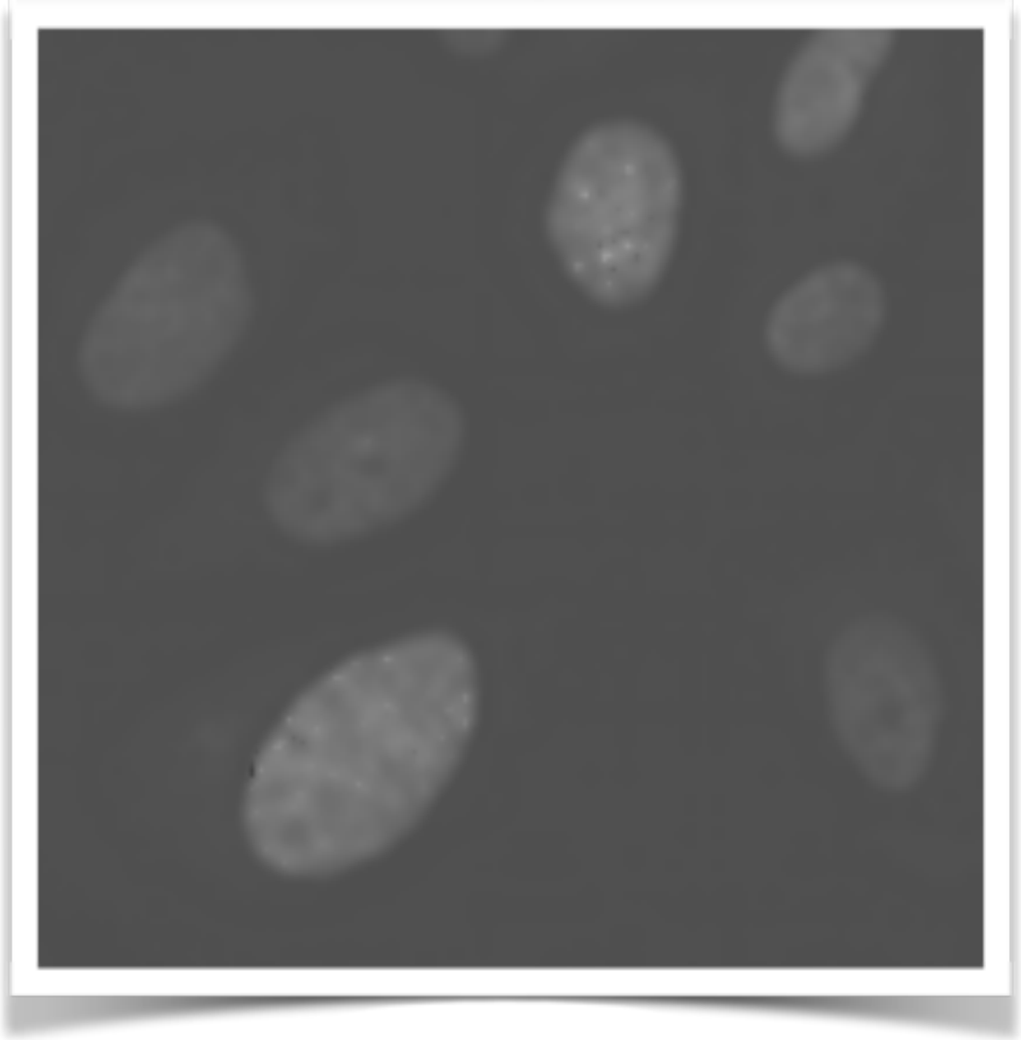


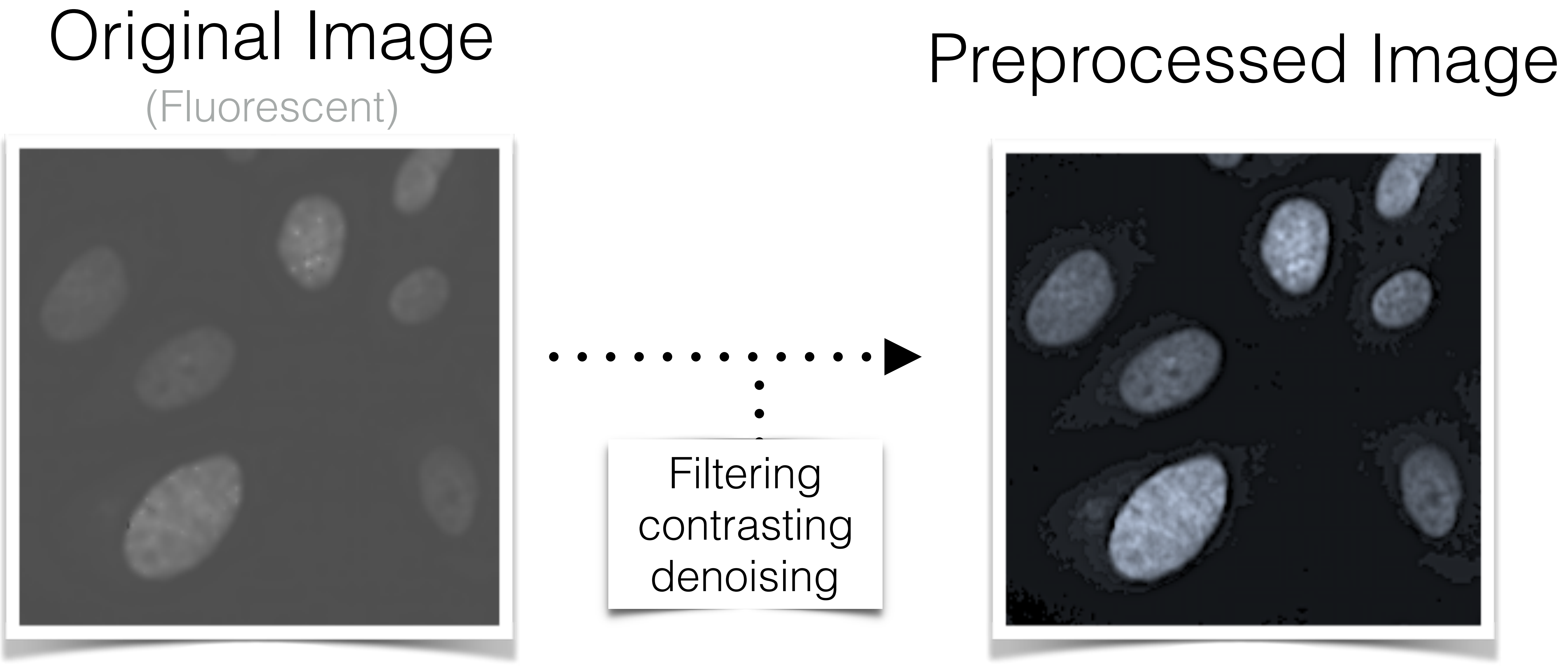
https://www.pharmacelera.com/publications/what-is-erooms-law/

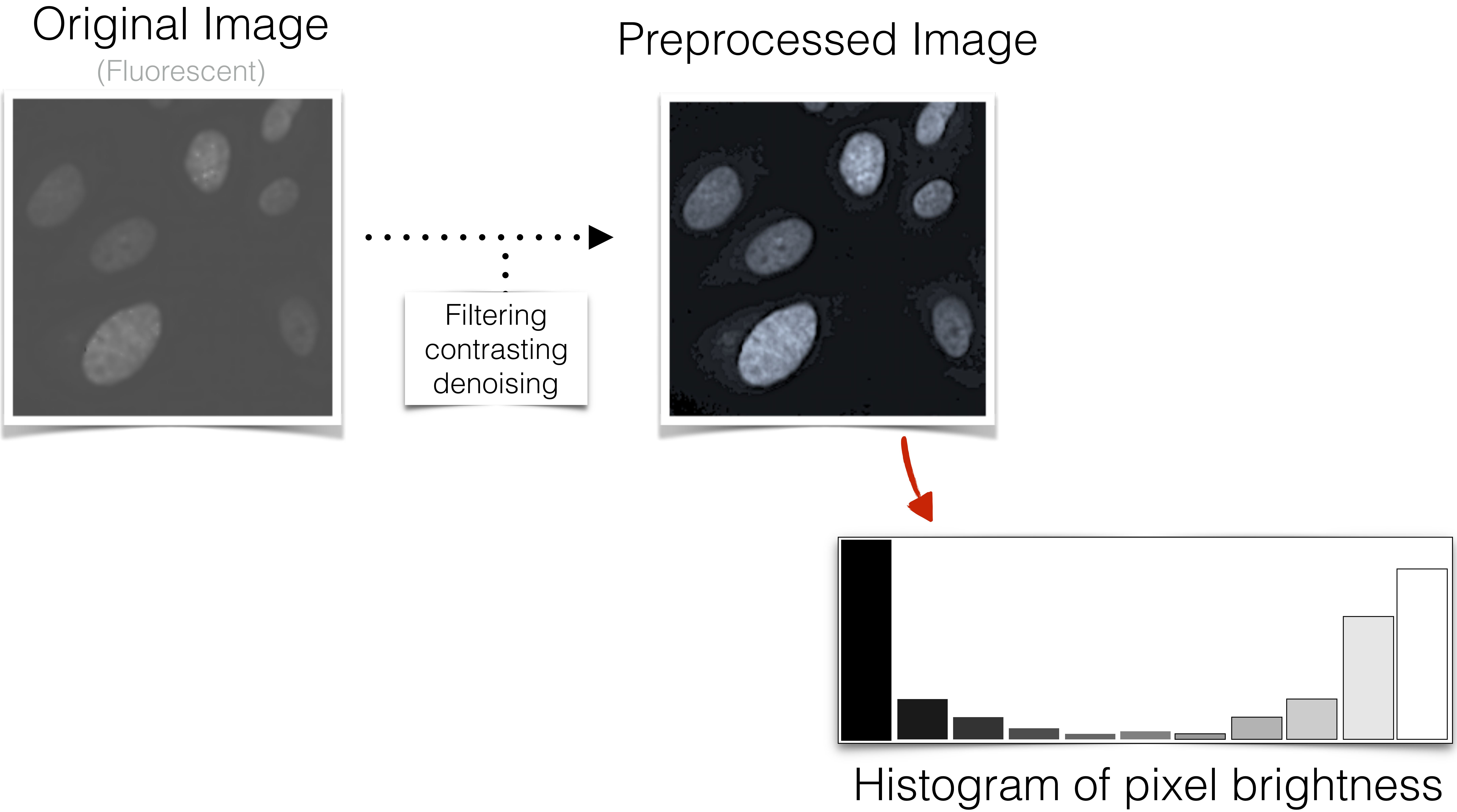
Classical Microscopy Image Analysis Pipeline

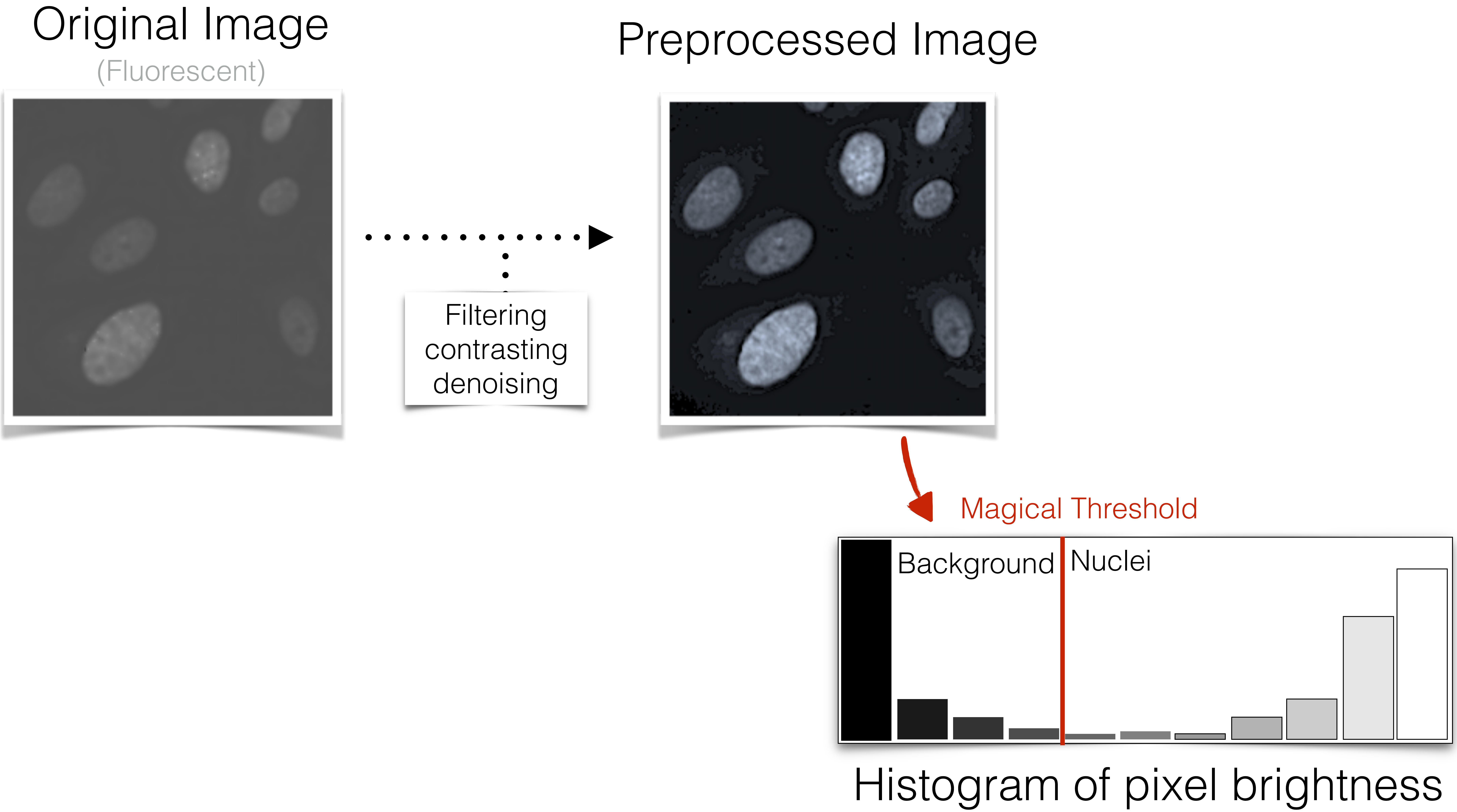
Original Image

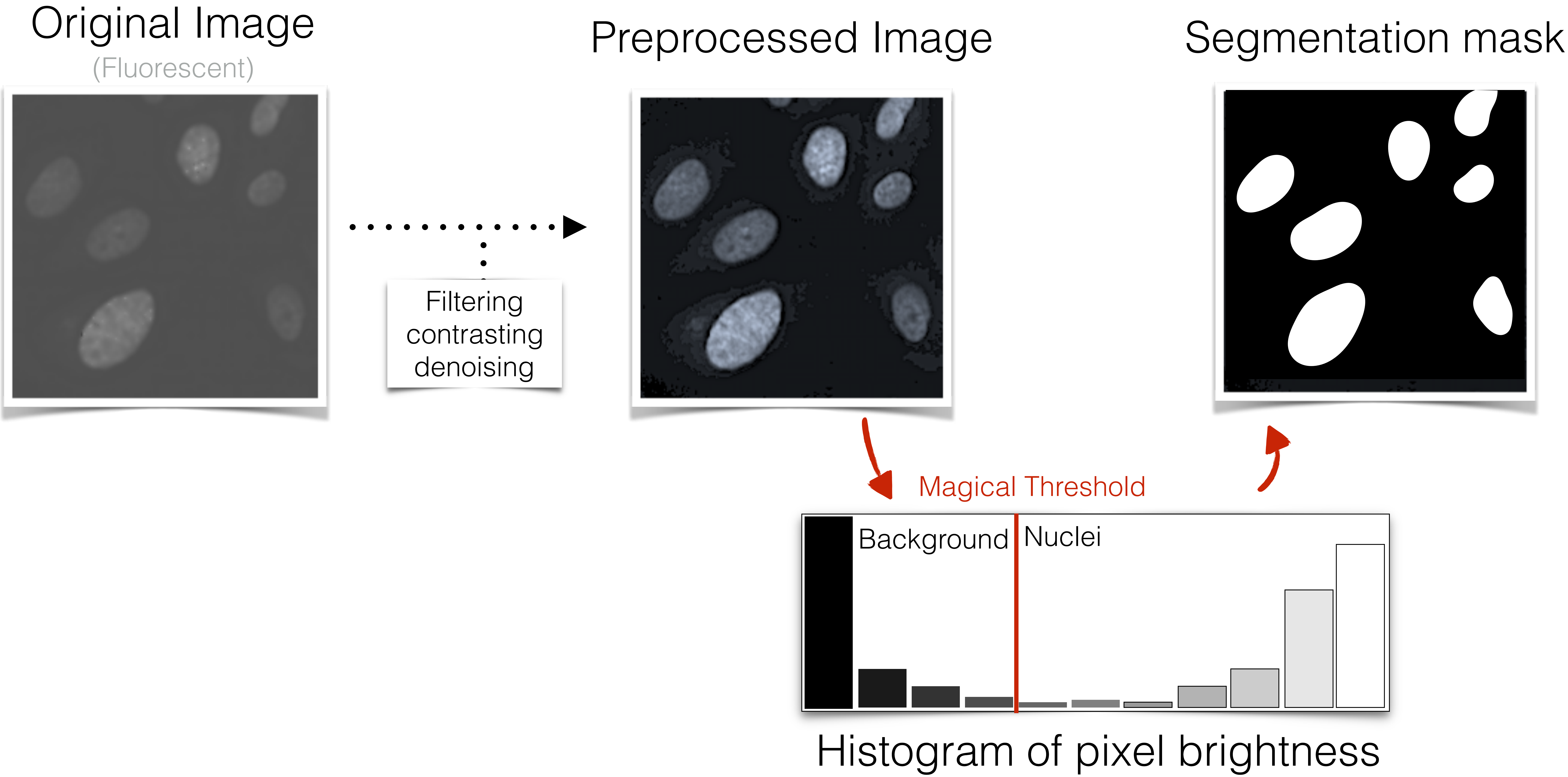
(Fluorescent)

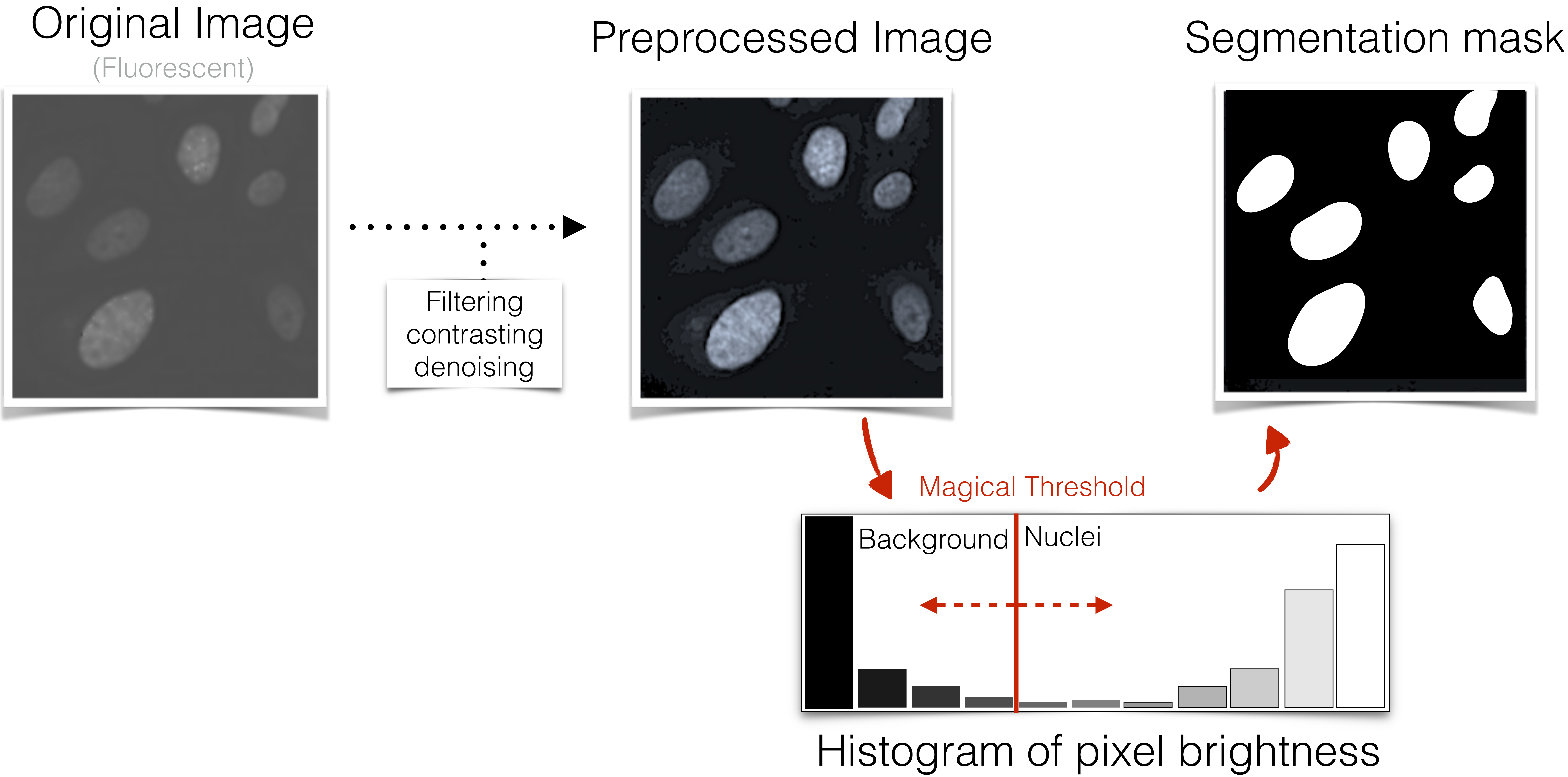


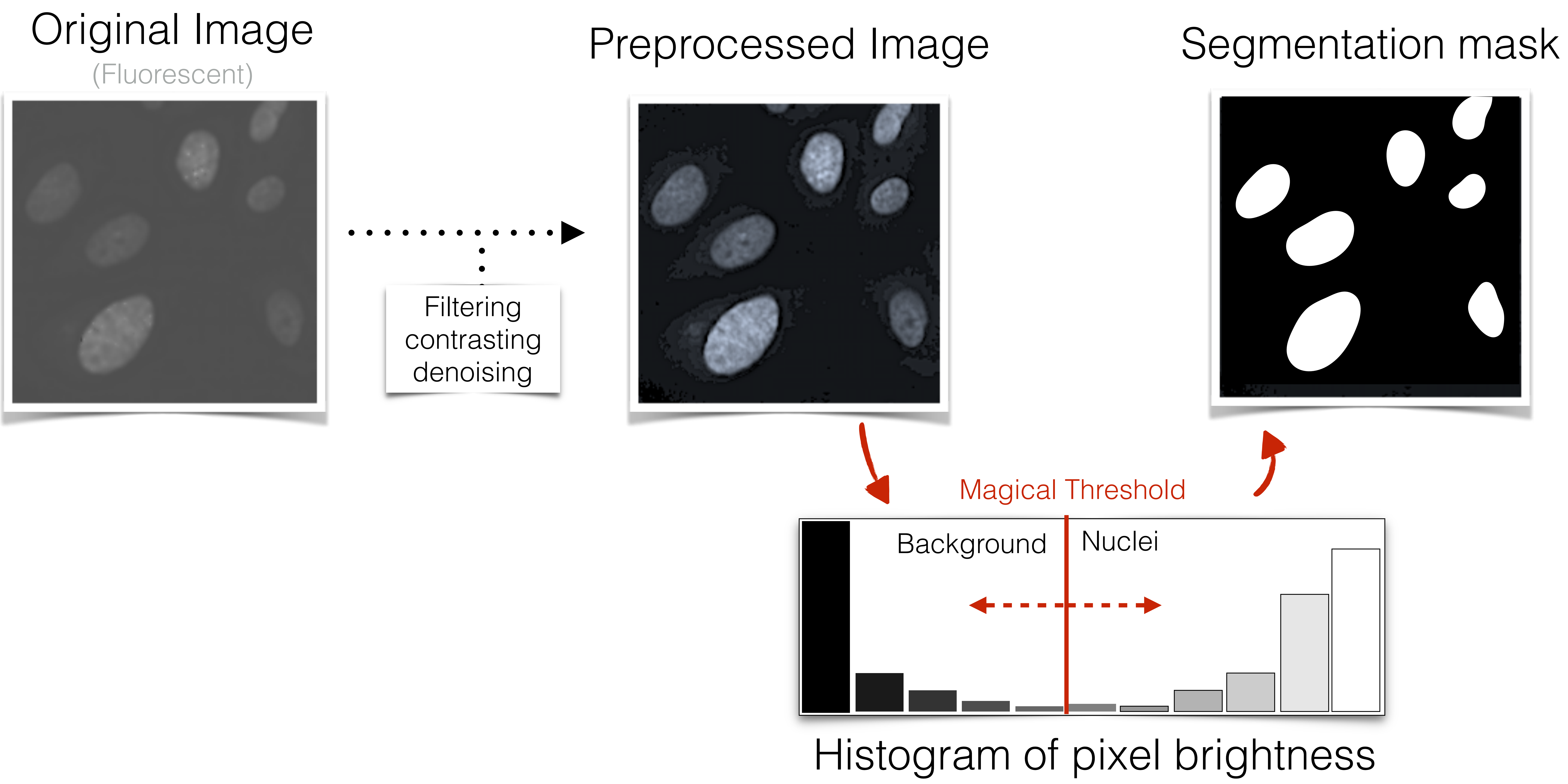

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image

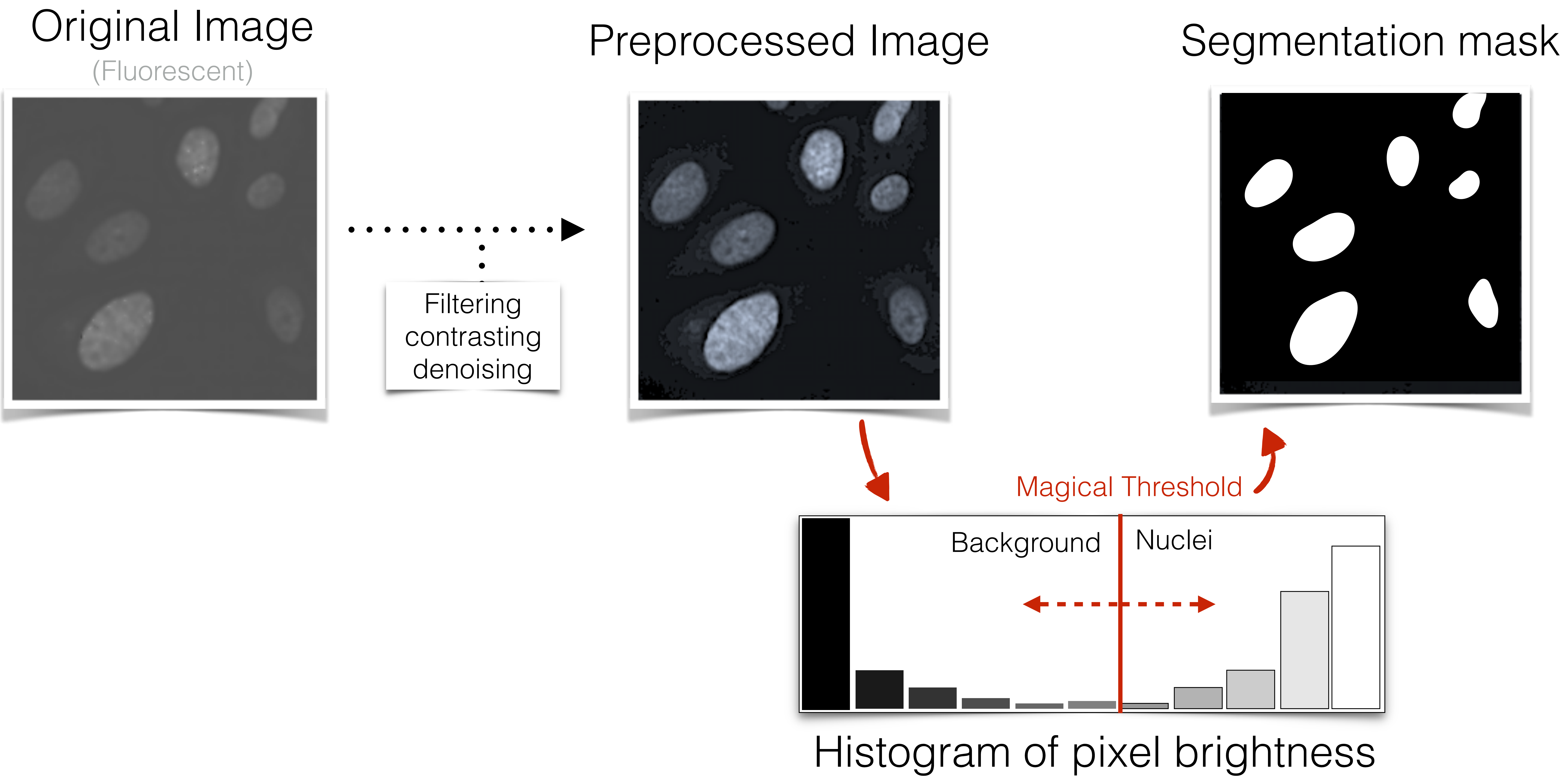

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Histogram of pixel brightness


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Magical Threshold
Background
Nuclei
Histogram of pixel brightness

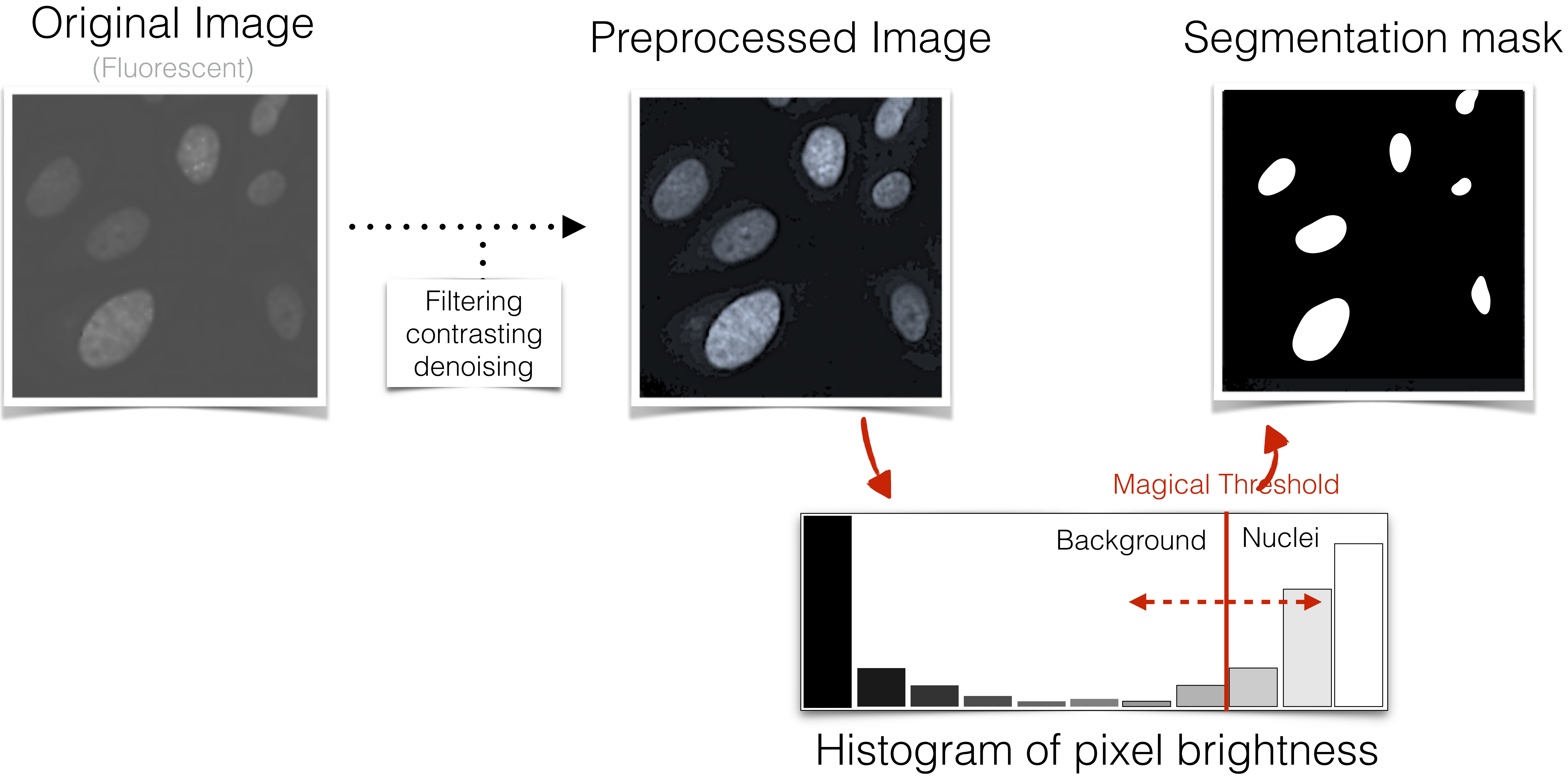

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Segmentation mask
Magical Threshold
Background
Nuclei
Histogram of pixel brightness

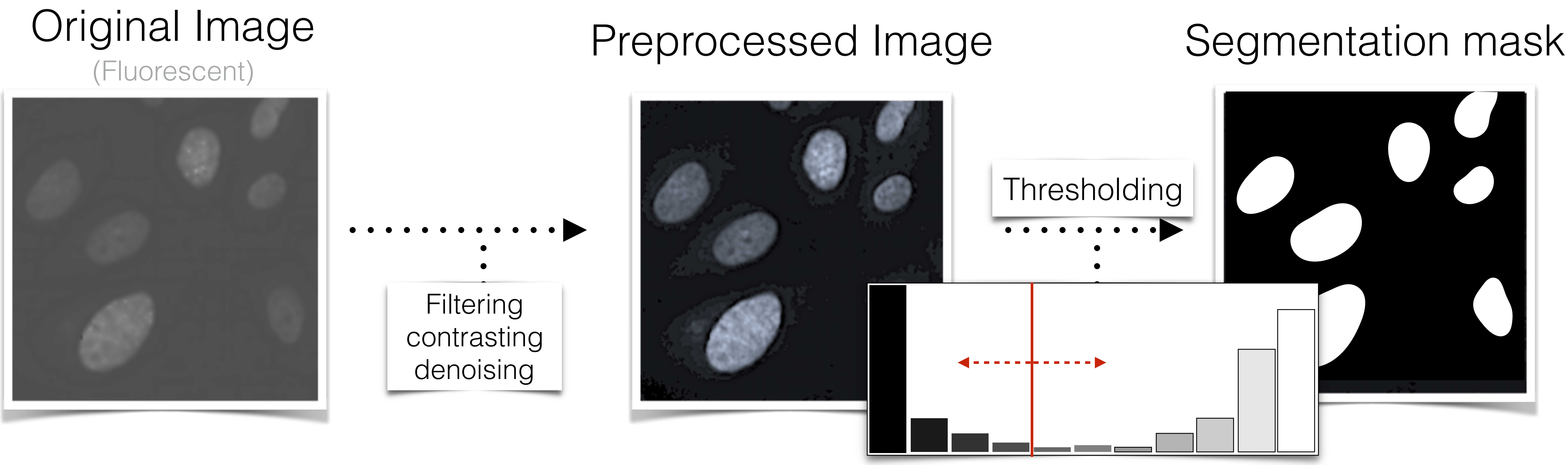

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Segmentation mask
Magical Threshold
Background
Nuclei
Histogram of pixel brightness

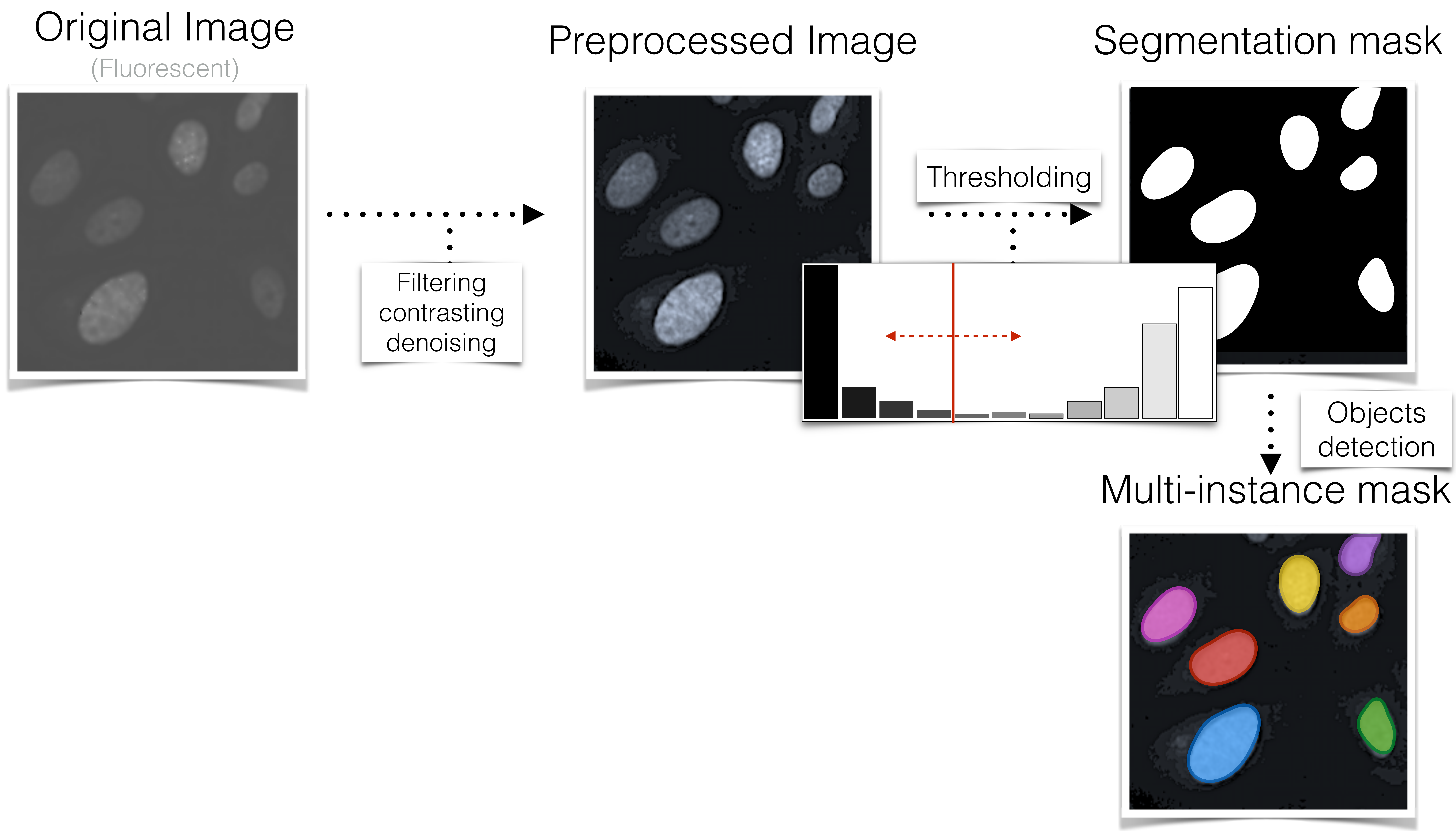

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Segmentation mask
Magical Threshold
Background
Nuclei
Histogram of pixel brightness

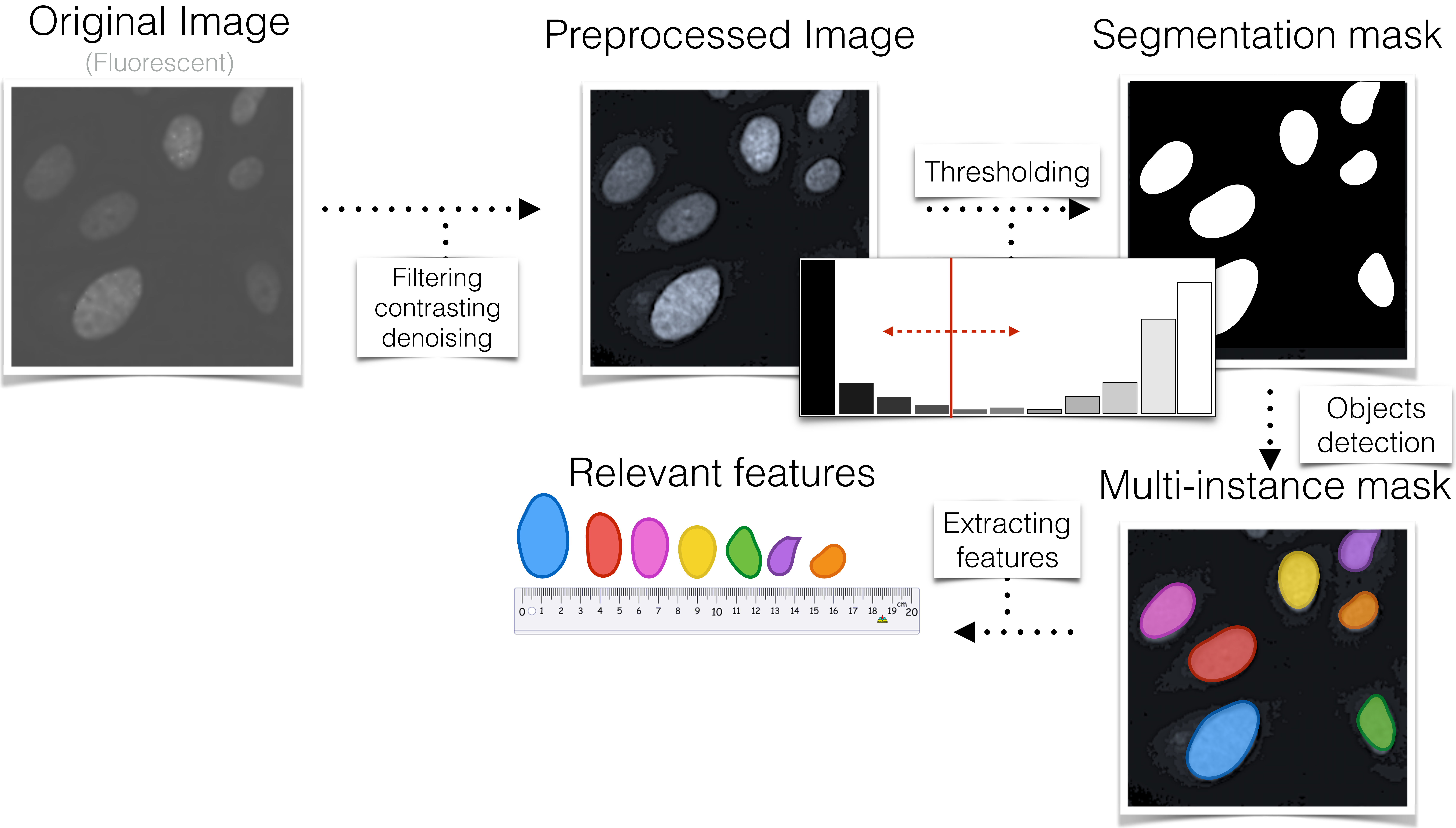


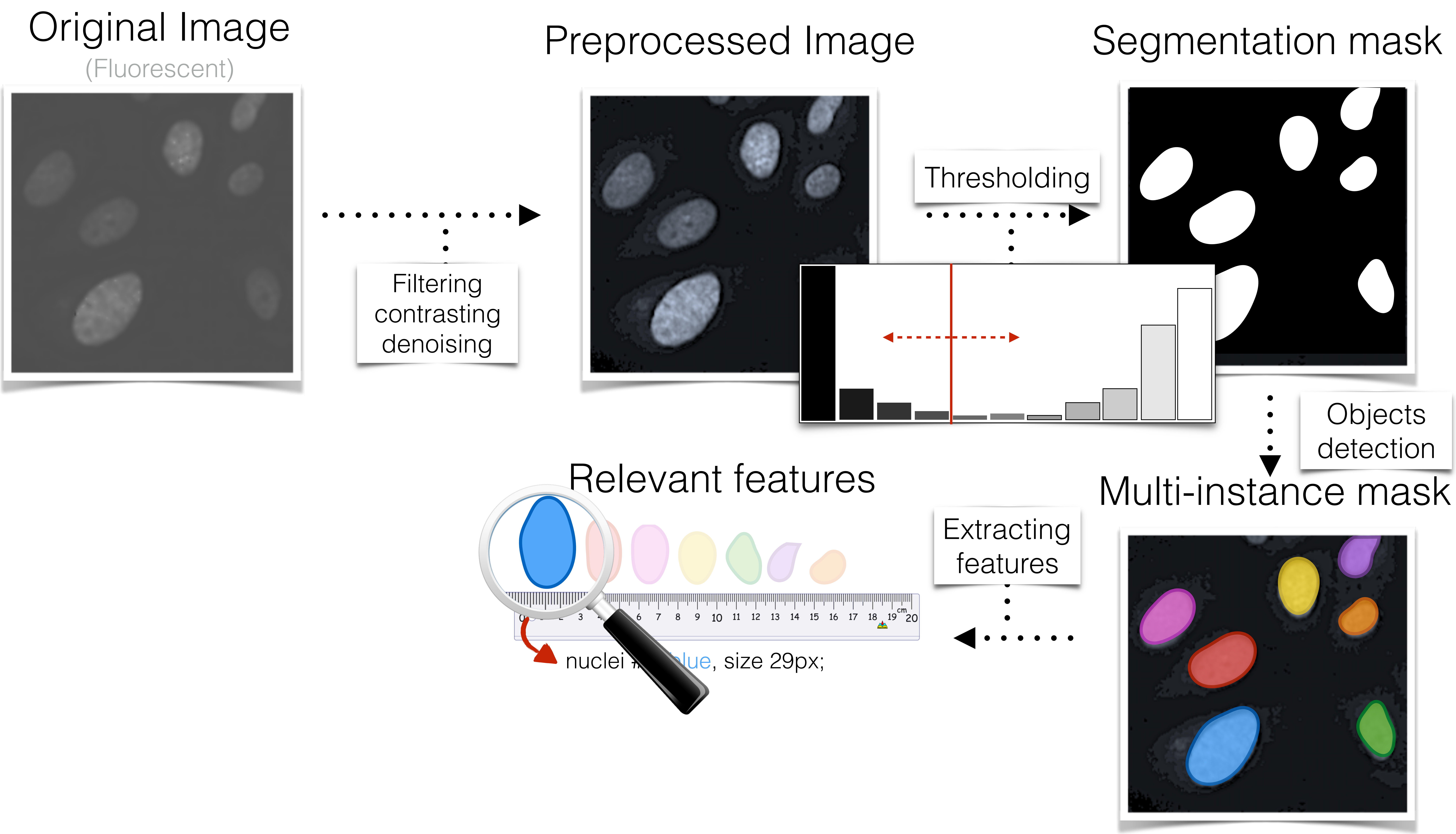
Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Segmentation mask
Magical Threshold
Background
Nuclei
Histogram of pixel brightness

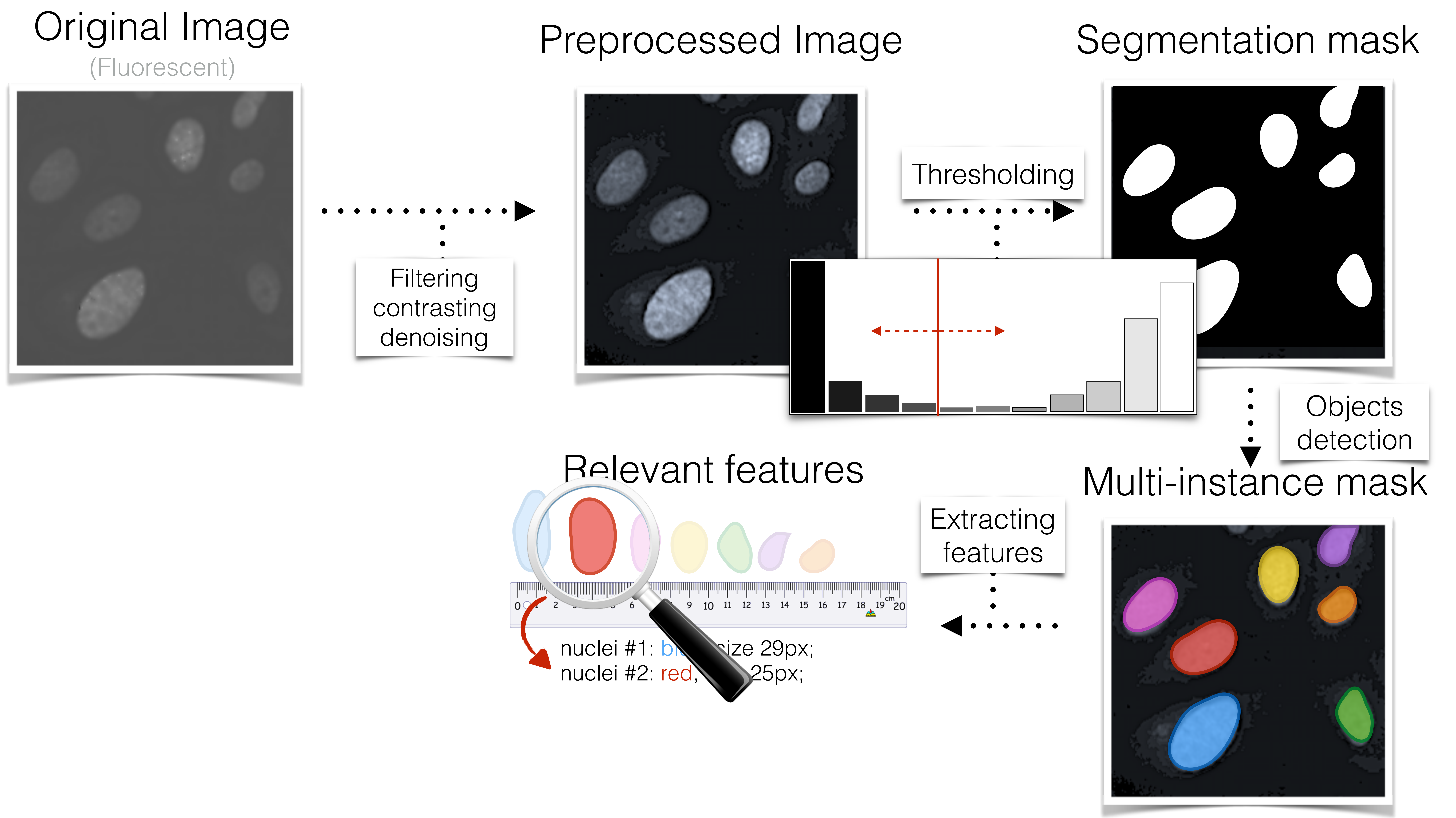

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Segmentation mask
Magical Threshold
Background
Nuclei
Histogram of pixel brightness


Original Image
(Fluorescent)
Preprocessed Image
Segmentation mask
Filtering
contrasting
denoising
Thresholding

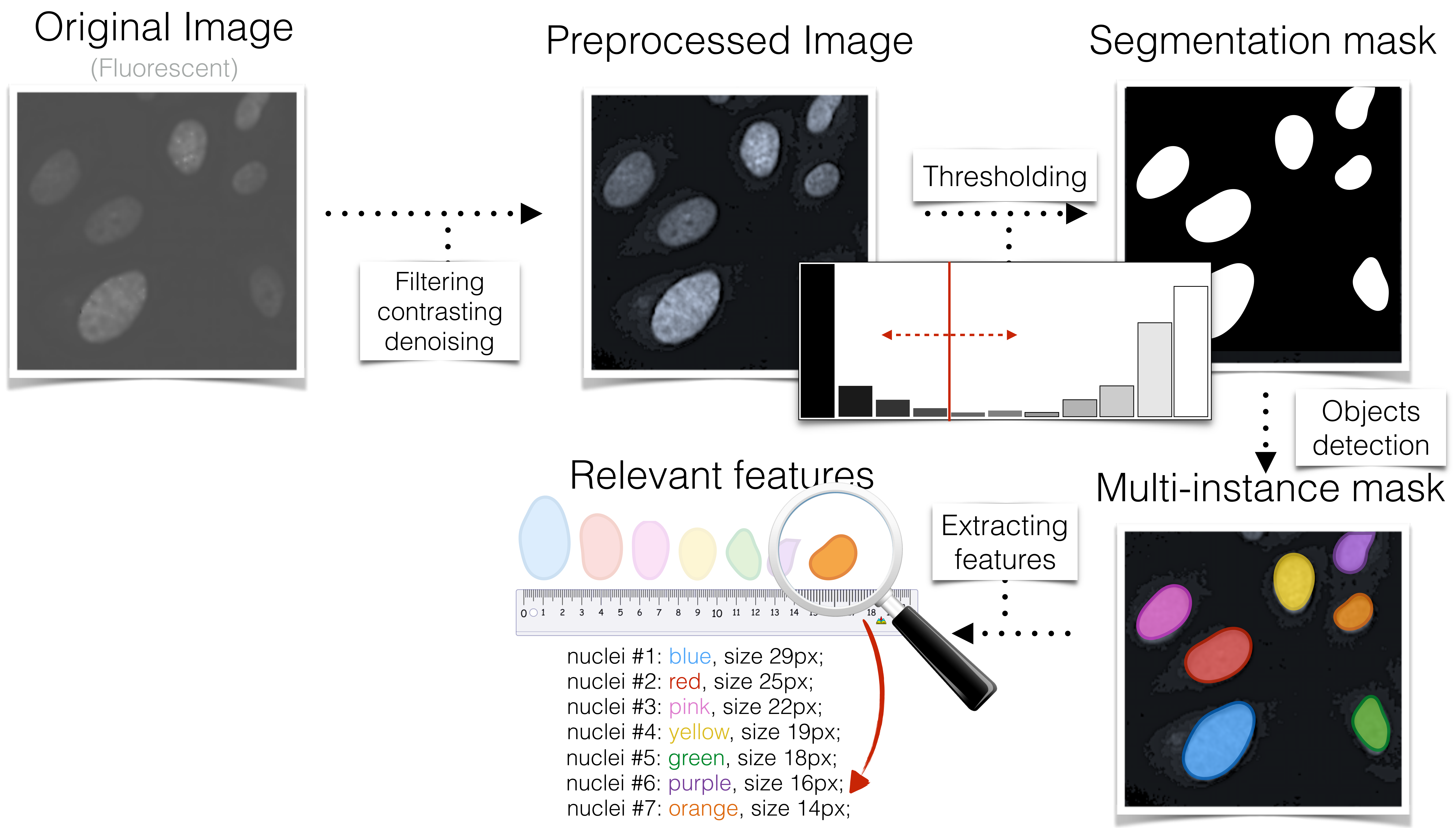

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Objects
detection
Multi-instance mask

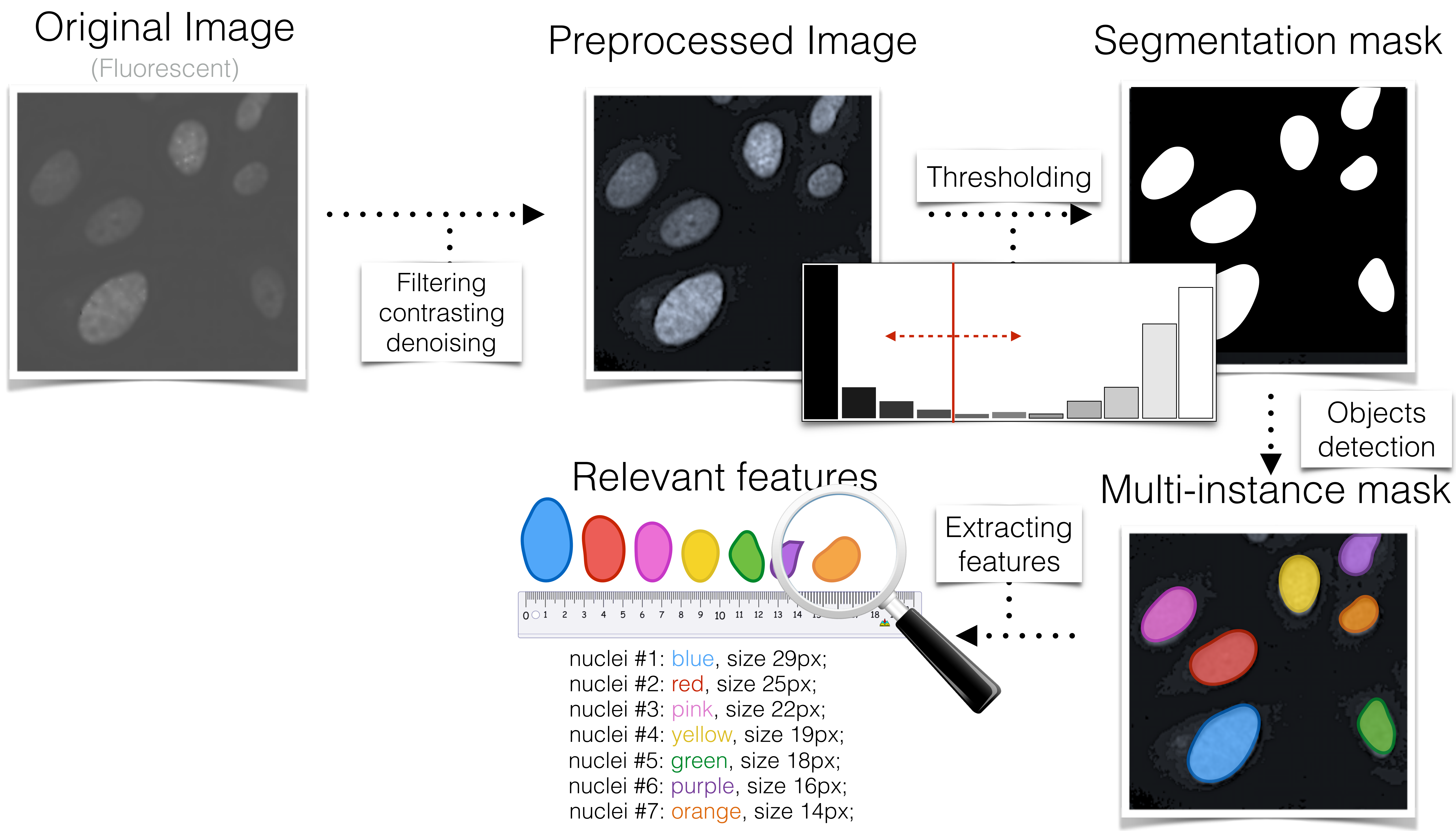

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
cm


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
nuclei #
lue, size 29px;

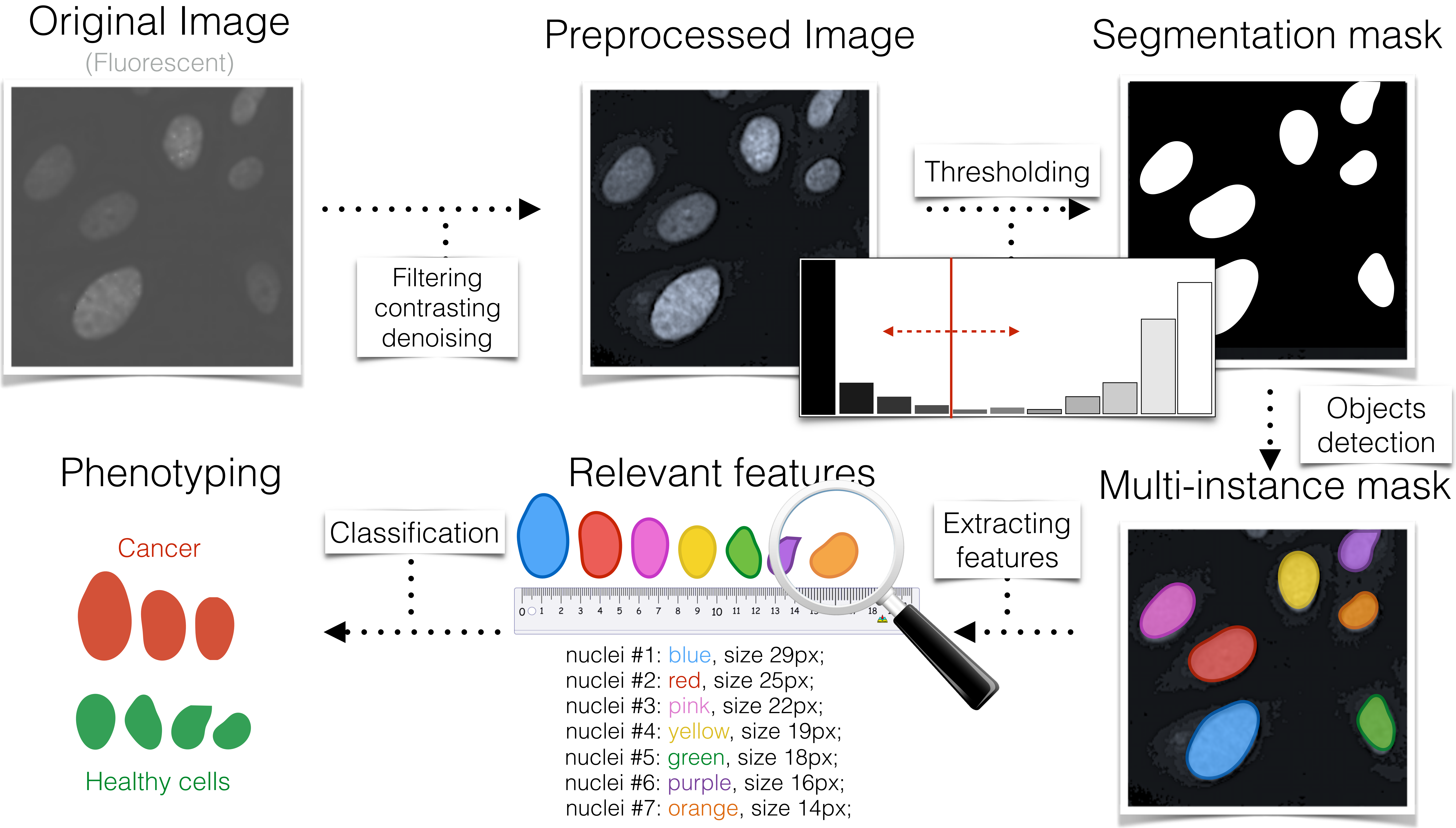


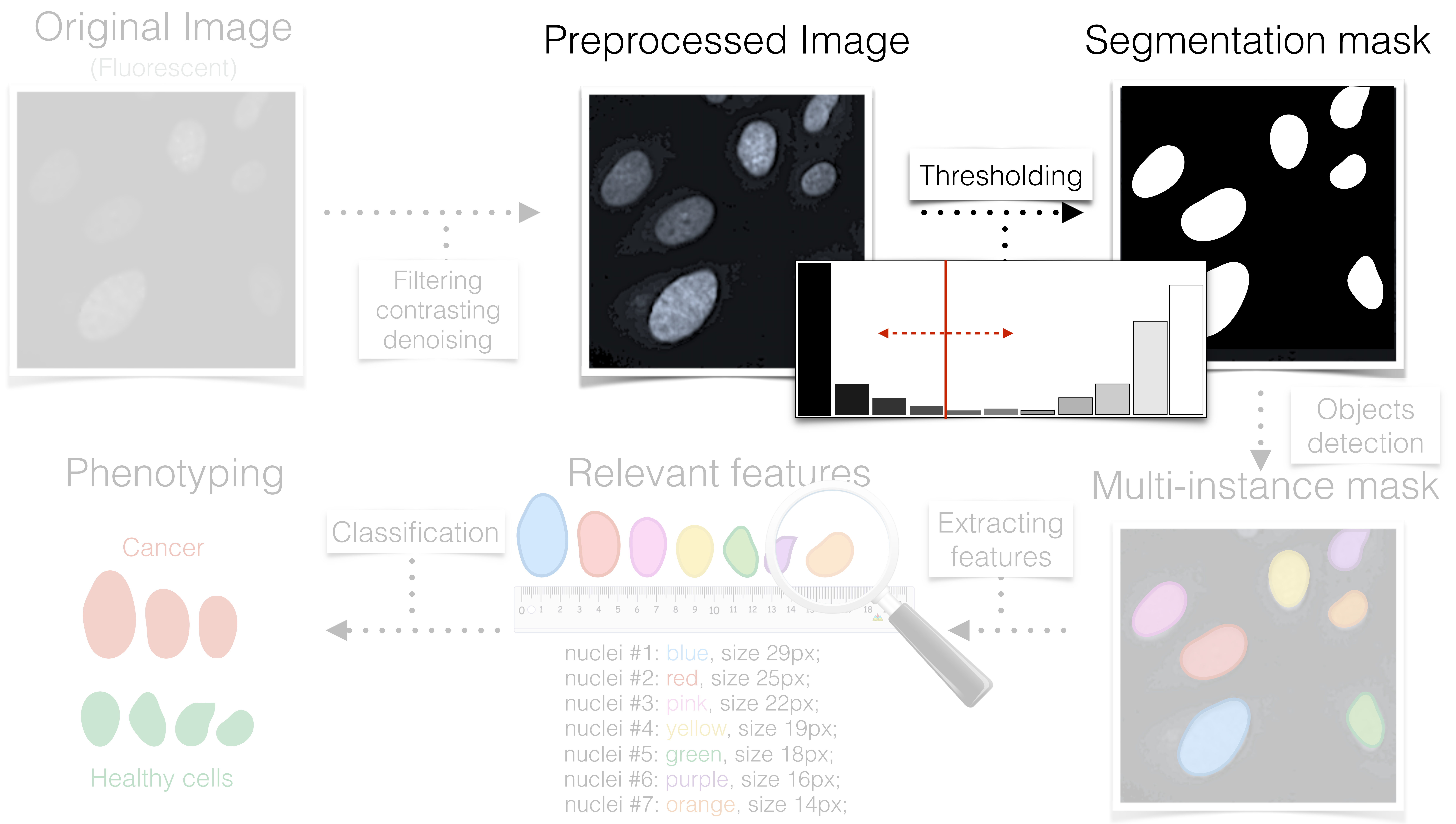
Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
nuclei #1: bl... size 29px;
nuclei #2: red, ... 25px;


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
nuclei #1: blue, size 29px;
nuclei #2: red, size 25px;
nuclei #3: pink, size 22px;
nuclei #4: yellow, size 19px;
nuclei #5: green, size 18px;
nuclei #6: purple, size 16px;
nuclei #7: orange, size 14px;



Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
nuclei #1: blue, size 29px;
nuclei #2: red, size 25px;
nuclei #3: pink, size 22px;
nuclei #4: yellow, size 19px;
nuclei #5: green, size 18px;
nuclei #6: purple, size 16px;
nuclei #7: orange, size 14px;


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
nuclei #1: blue, size 29px;
nuclei #2: red, size 25px;
nuclei #3: pink, size 22px;
nuclei #4: yellow, size 19px;
nuclei #5: green, size 18px;
nuclei #6: purple, size 16px;
nuclei #7: orange, size 14px;
Classification
Phenotyping
Cancer
Healthy cells



Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
Classification
Phenotyping
Cancer
Healthy cells
nuclei #1: blue, size 29px;
nuclei #2: red, size 25px;
nuclei #3: pink, size 22px;
nuclei #4: yellow, size 19px;
nuclei #5: green, size 18px;
nuclei #6: purple, size 16px;
nuclei #7: orange, size 14px;


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Can Deep Learning step in?
Objects detection
Multi-instance mask
Extracting features
Relevant features
Classification
Phenotyping
Cancer
Healthy cells
nuclei #1: blue, size 29px;
nuclei #2: red, size 25px;
nuclei #3: pink, size 22px;
nuclei #4: yellow, size 19px;
nuclei #5: green, size 18px;
nuclei #6: purple, size 16px;
nuclei #7: orange, size 14px;

Original Image
(Fluorescent)

Filtering
contrasting
denoising

Preprocessed Image

Thresholding

Segmentation mask

Can Deep
Learning step in?

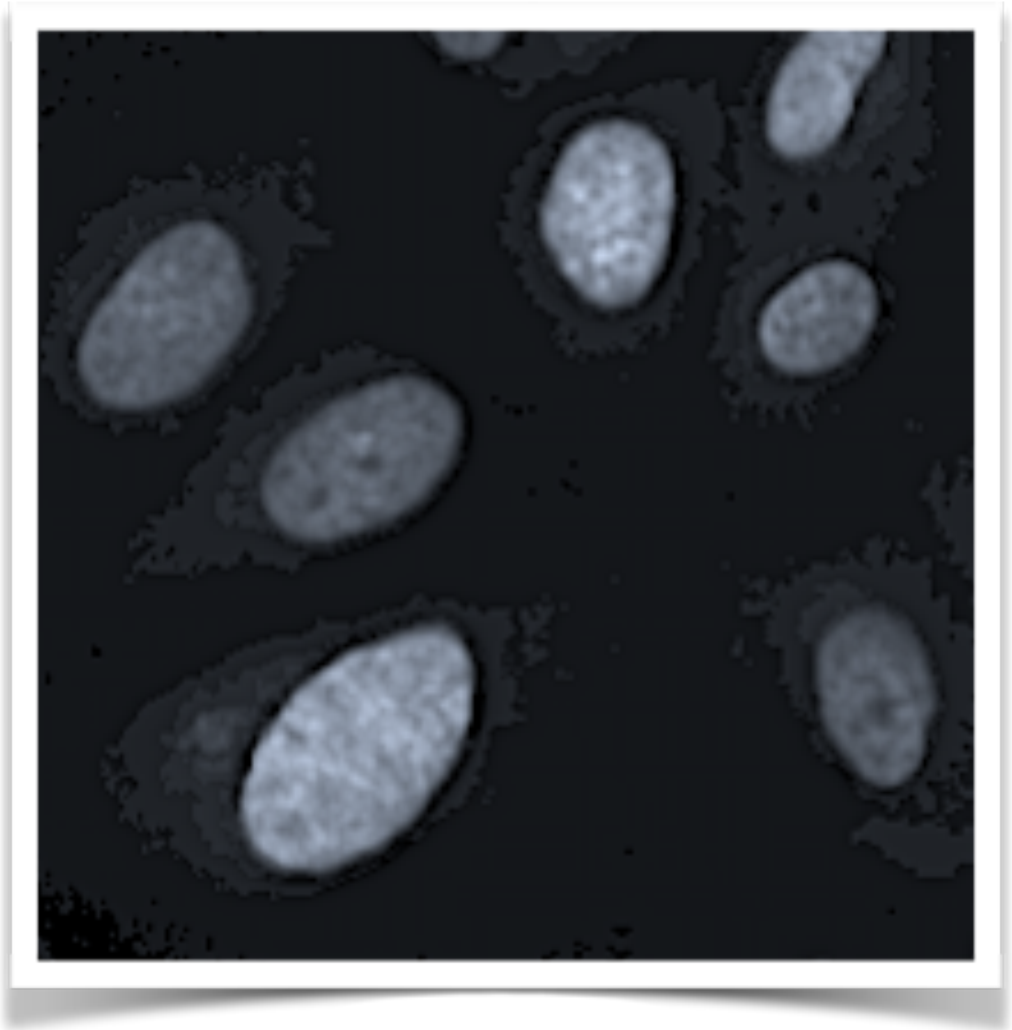
Approach I

Approach II

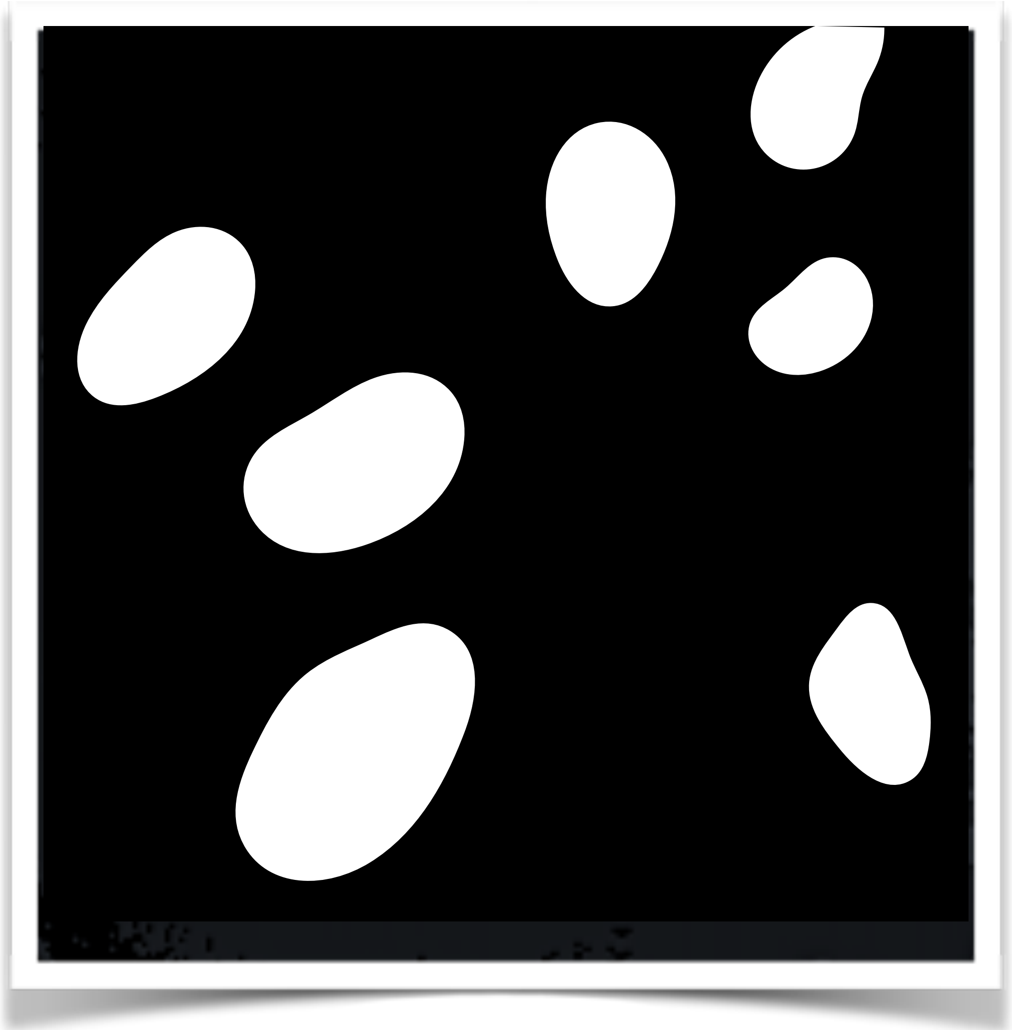
Approach III

Training DeepCell

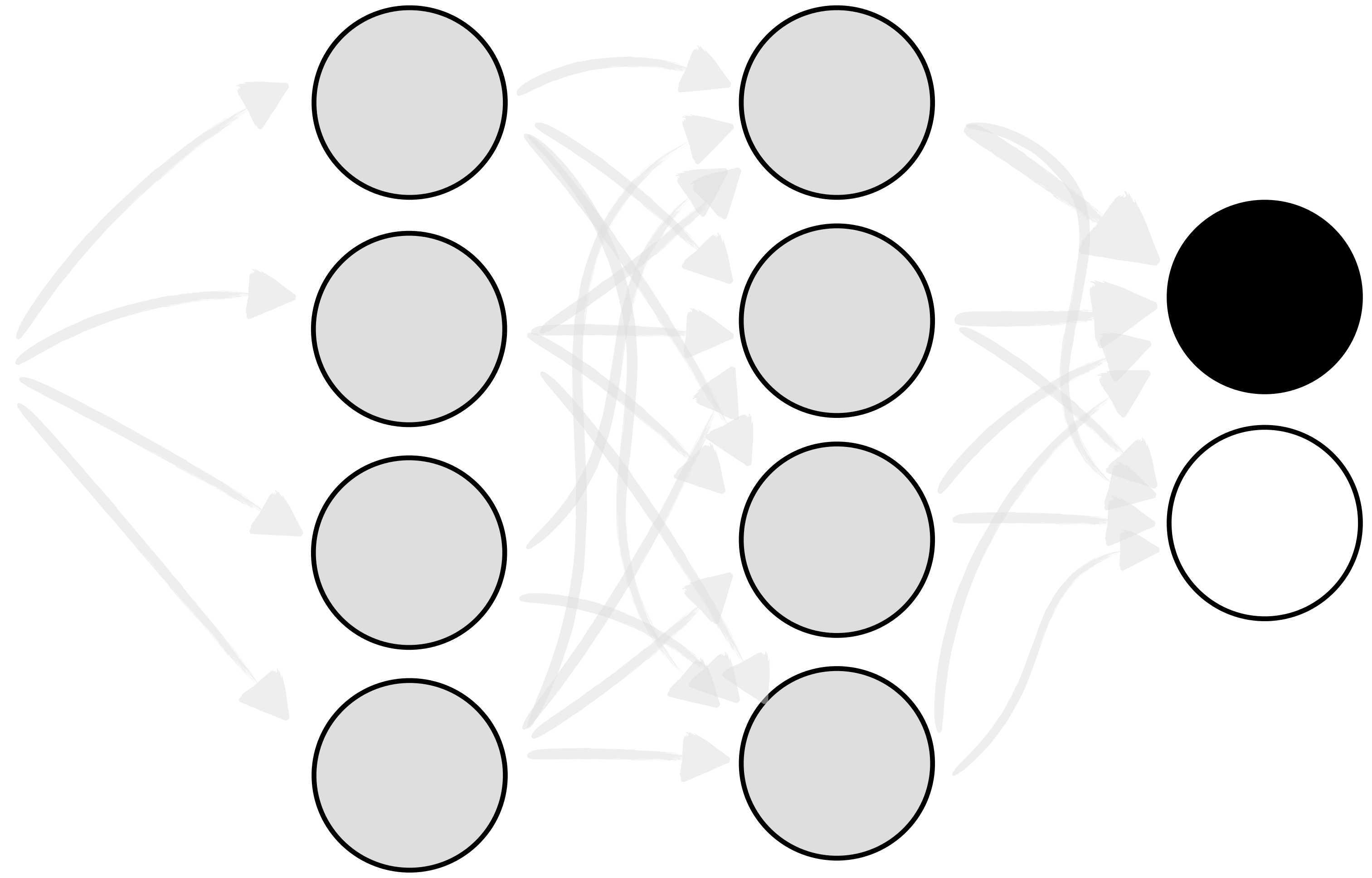
Original image



Segmentation



DeepCell

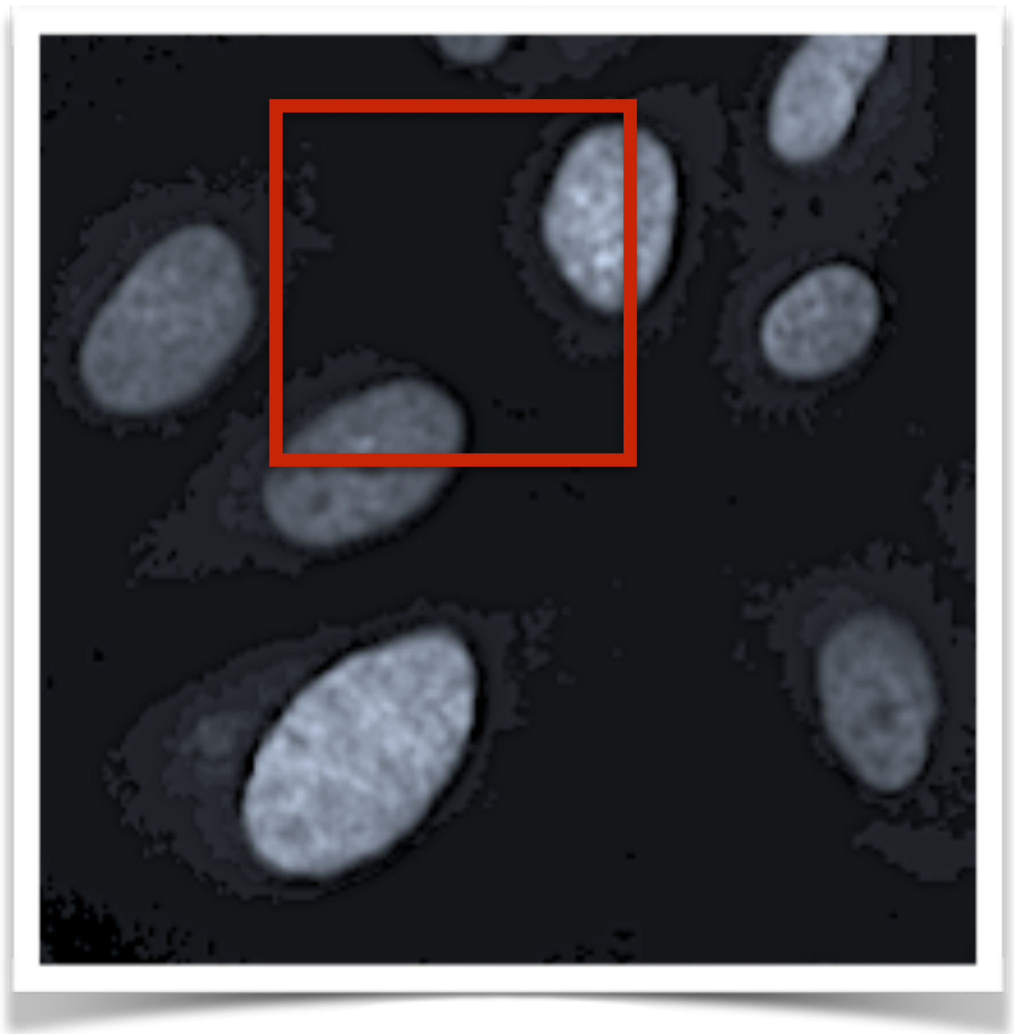


Training data

Training labels

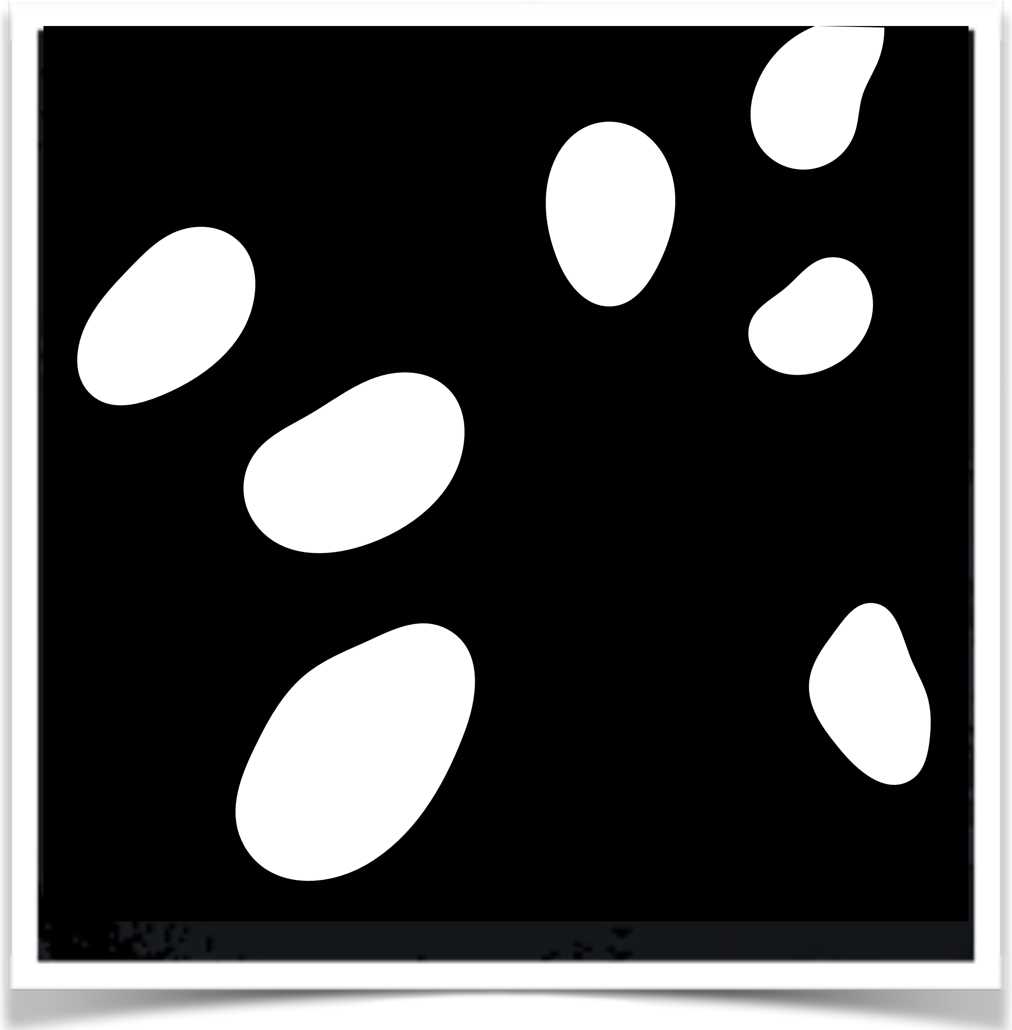
Training DeepCell

Original image

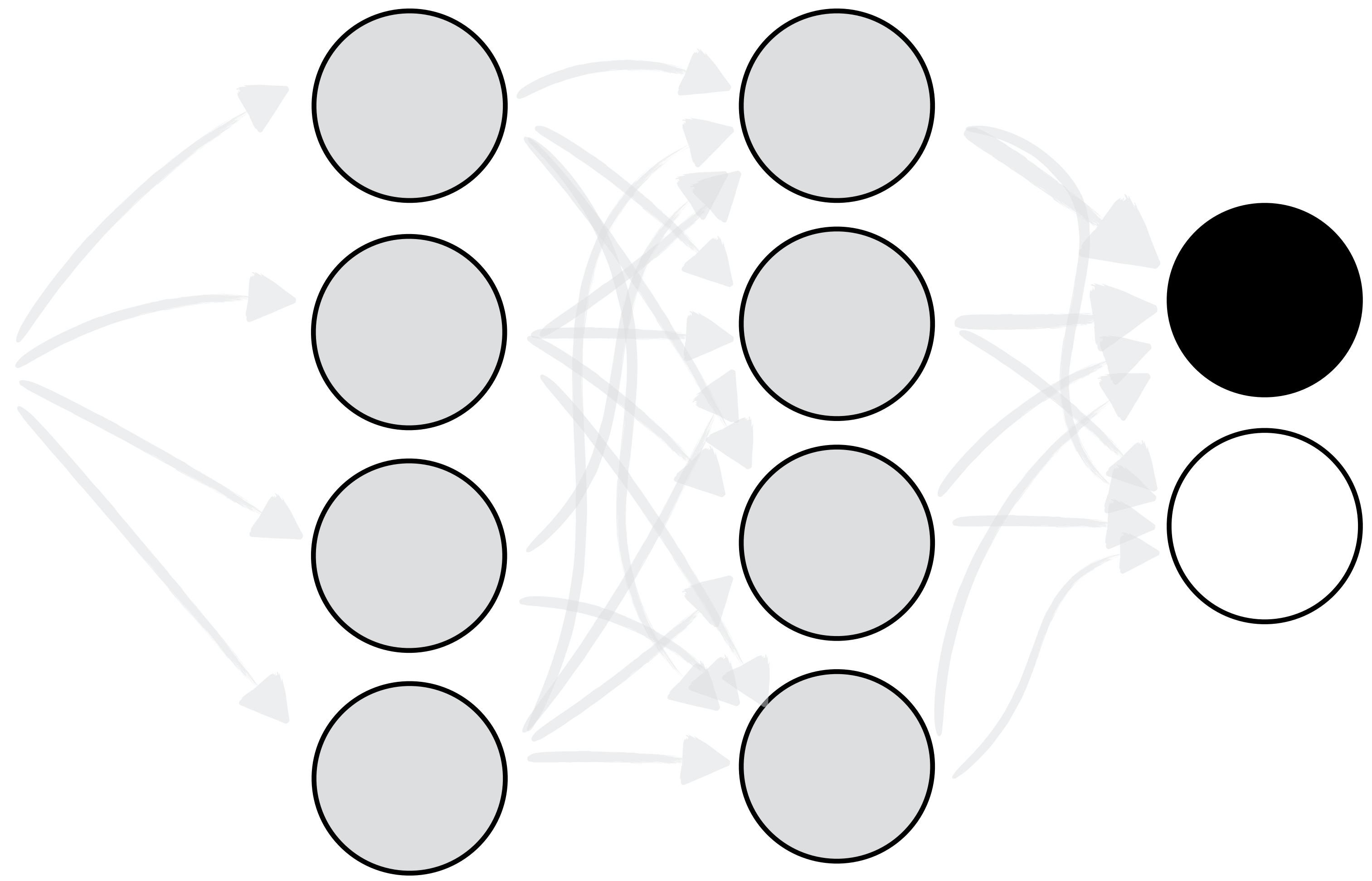


Extracted **patch** from original image used as input

Segmentation



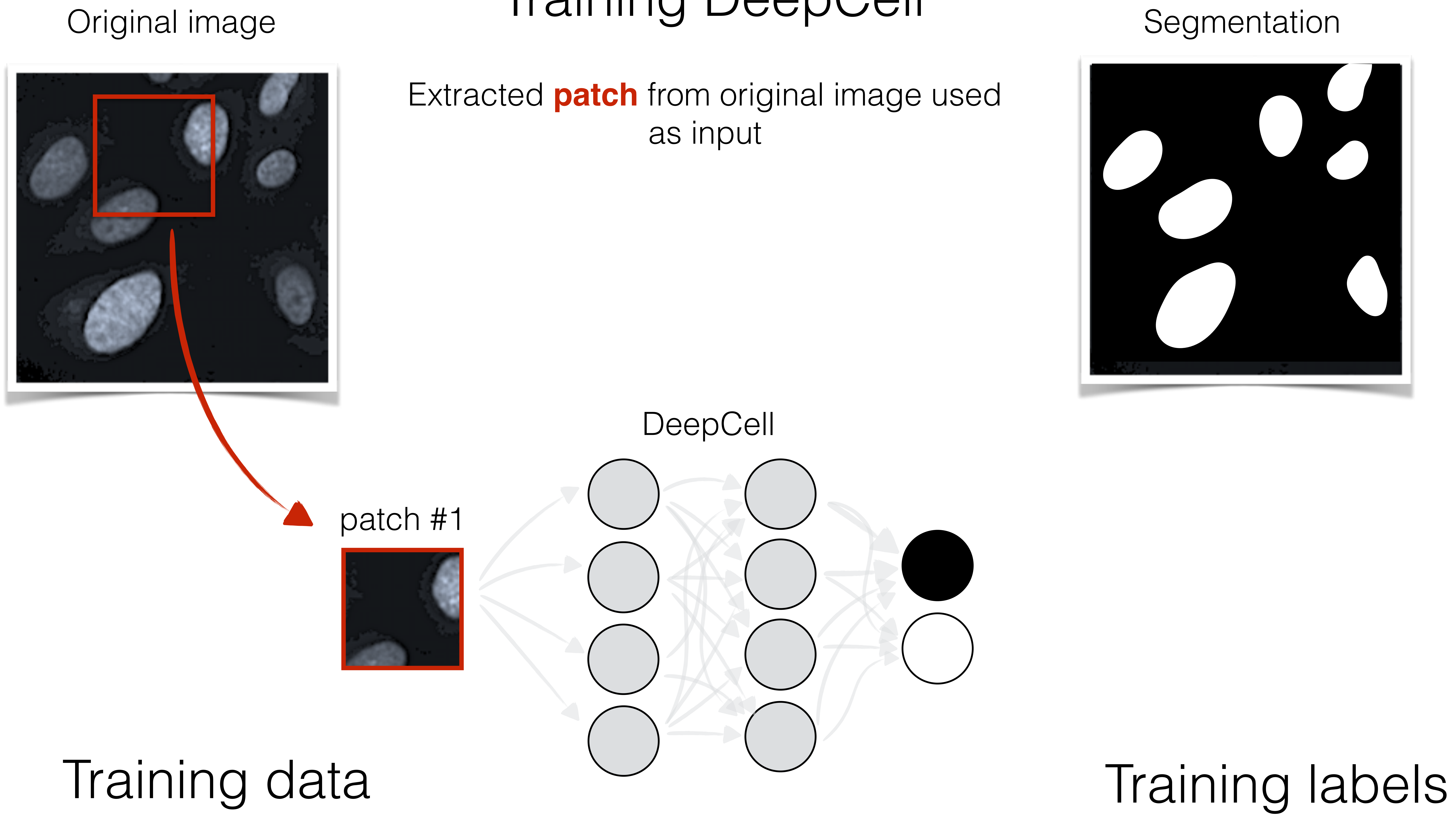
DeepCell

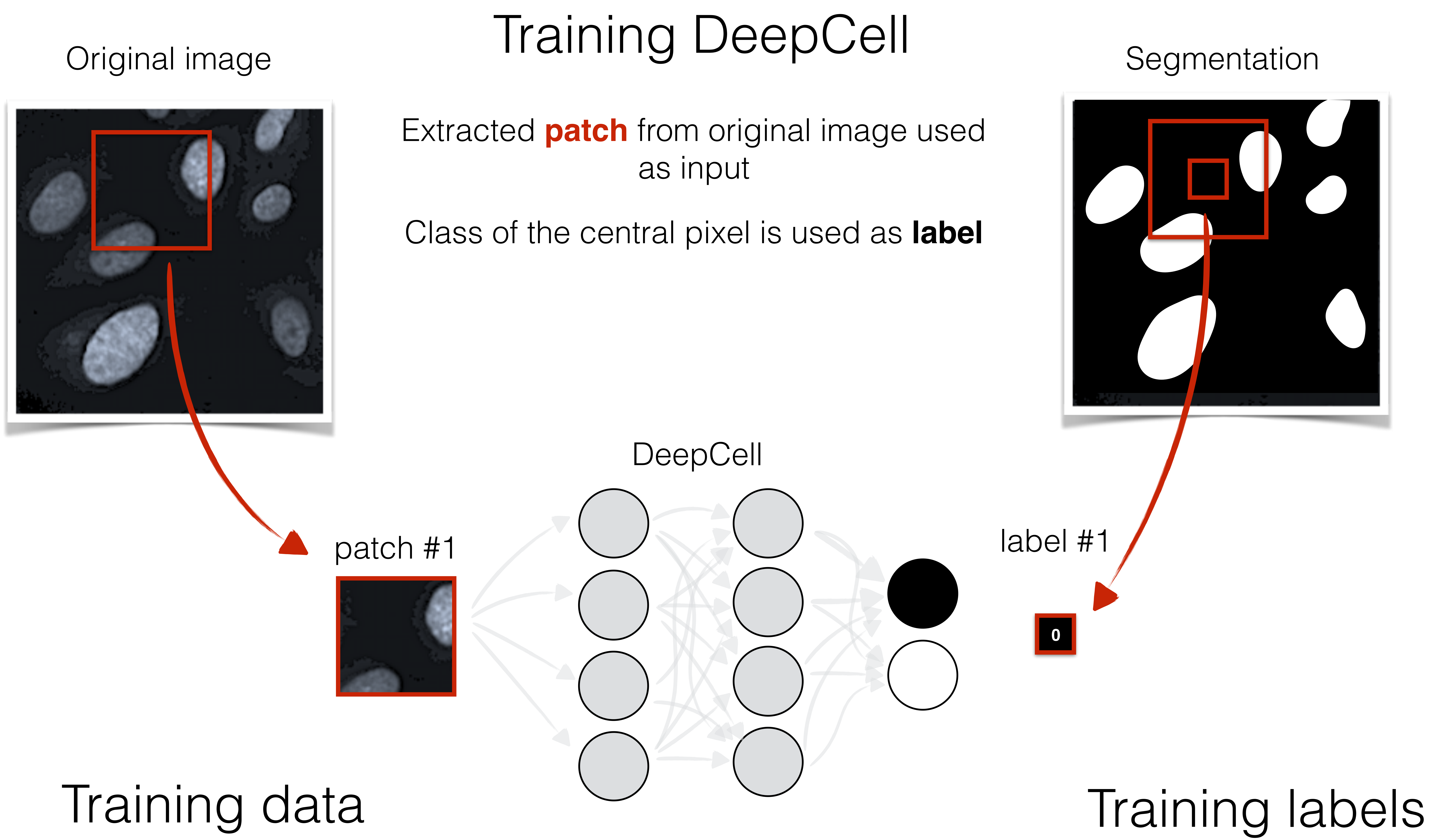


Training data

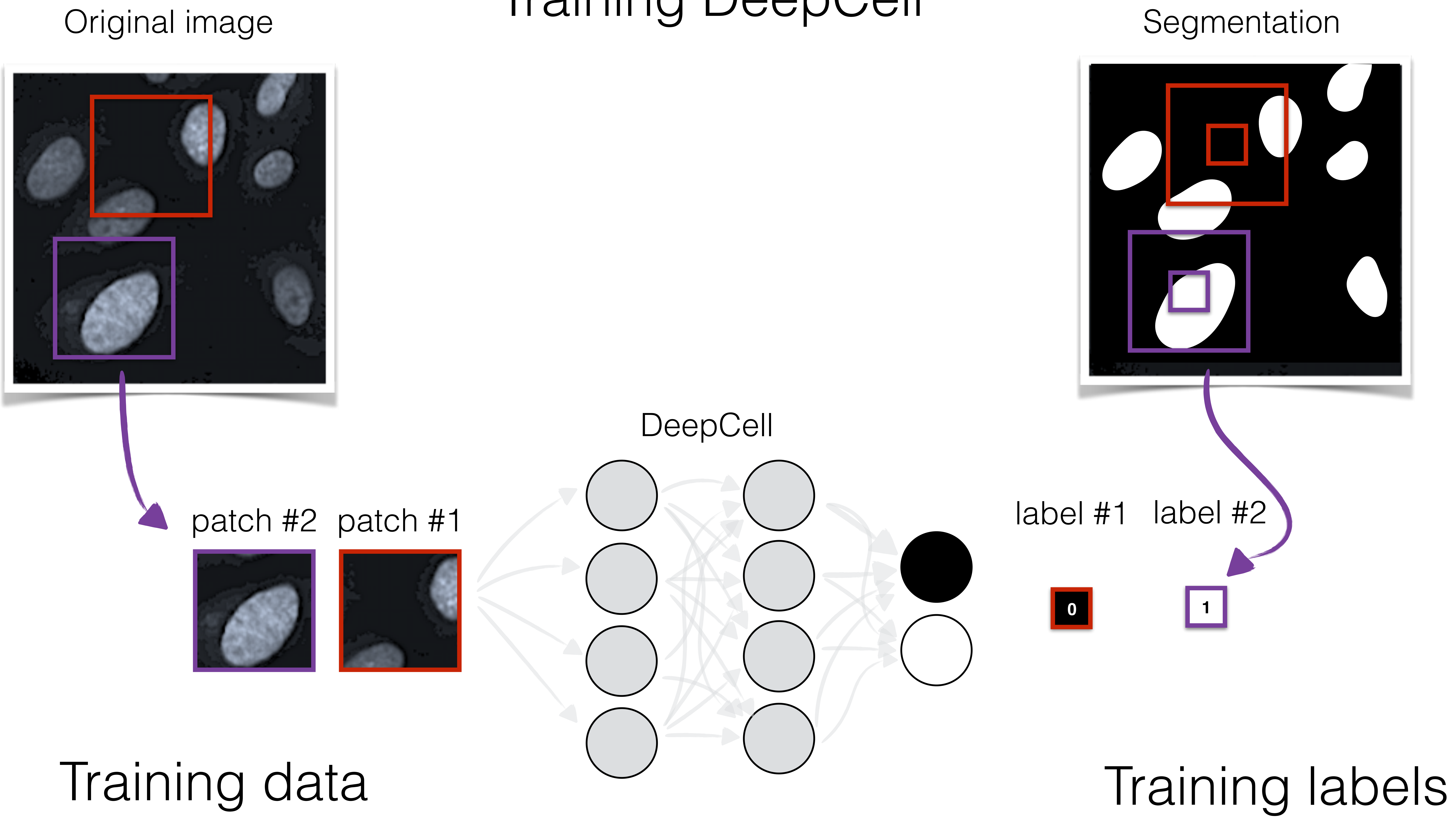
Training labels

Training DeepCell

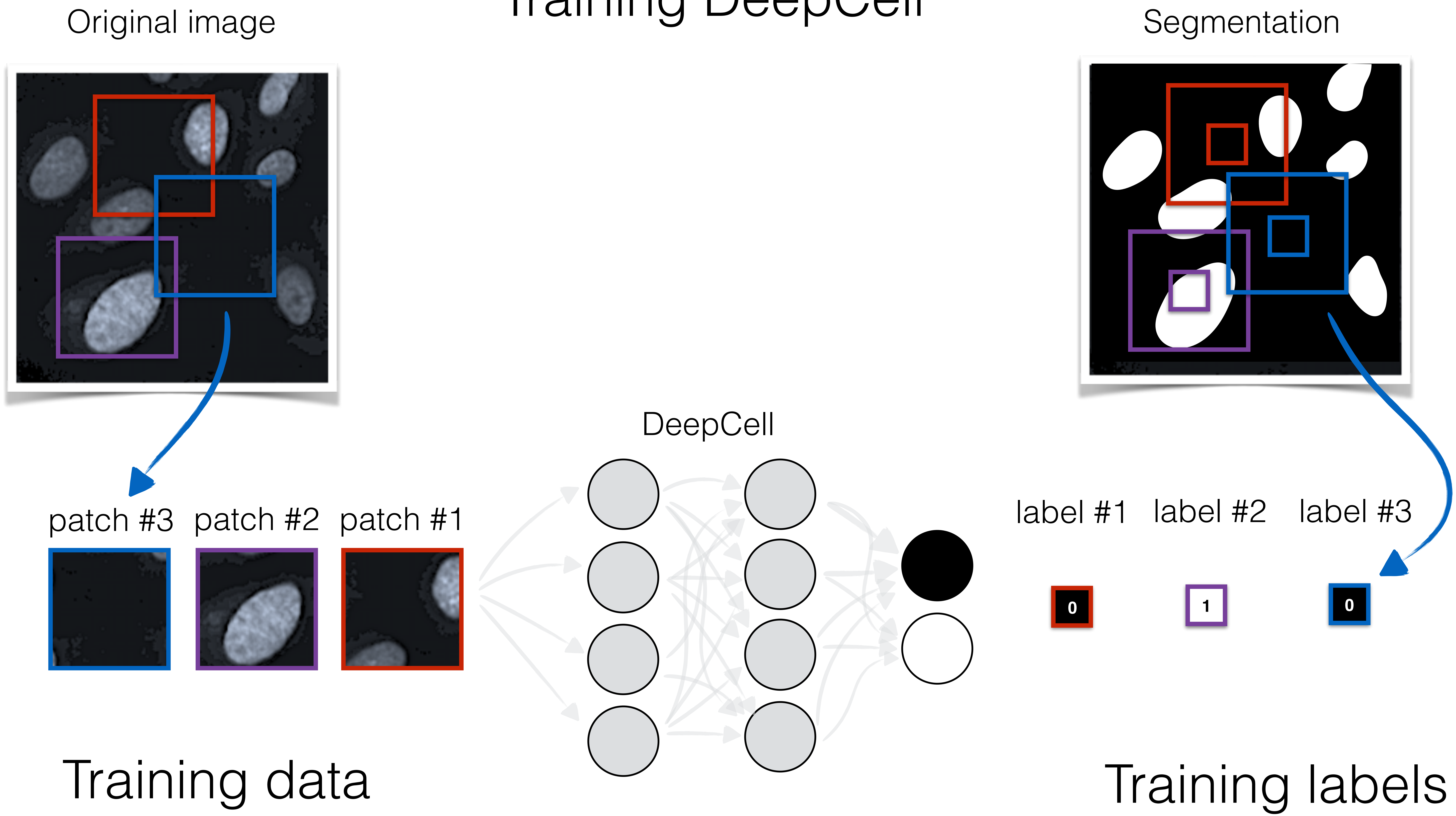



Training DeepCell
Original image
Segmentation
Extracted **patch** from original image used as input
Class of the central pixel is used as **label**
DeepCell
patch #1
label #1
0
Training data
Training labels

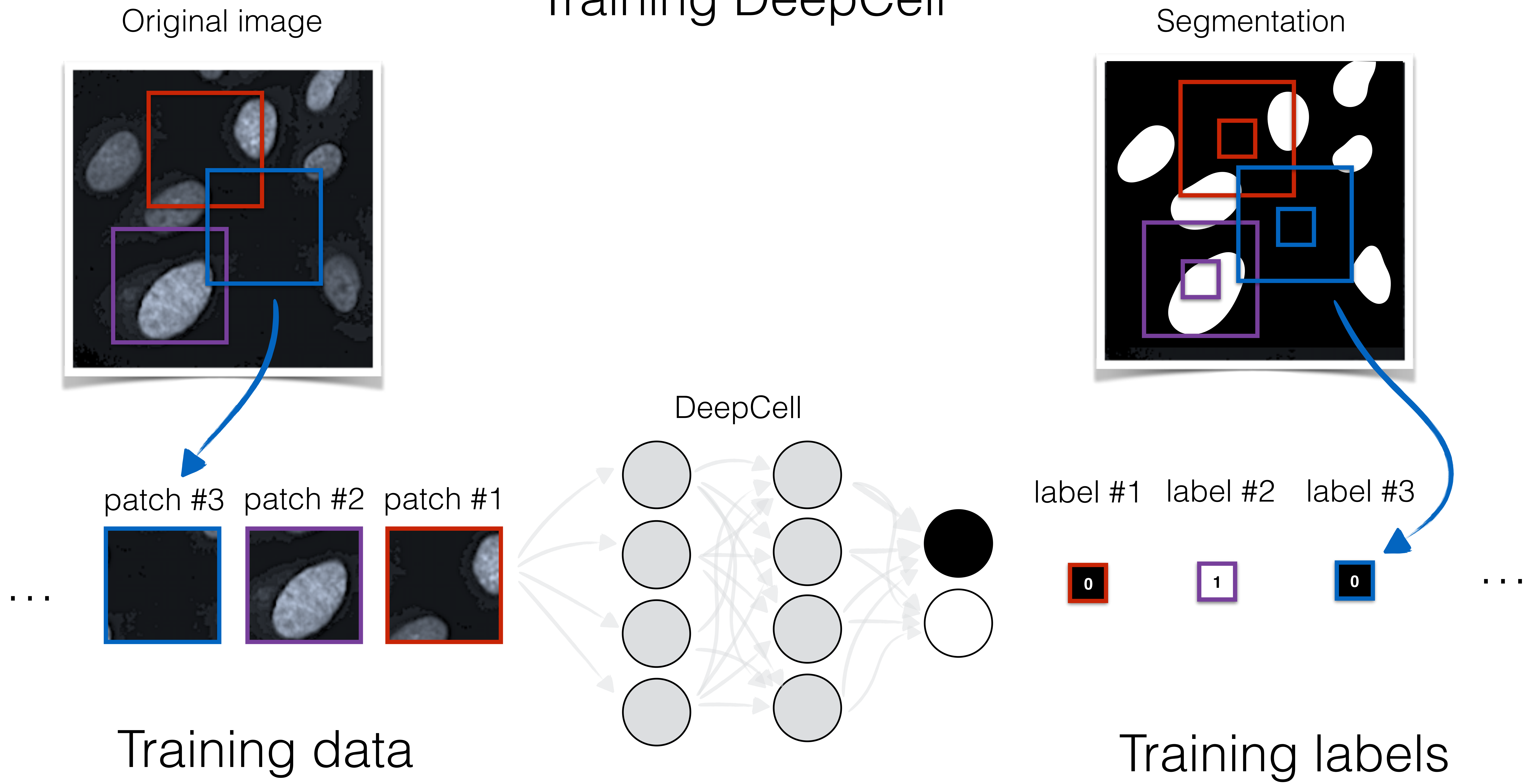
Training DeepCell



Training DeepCell

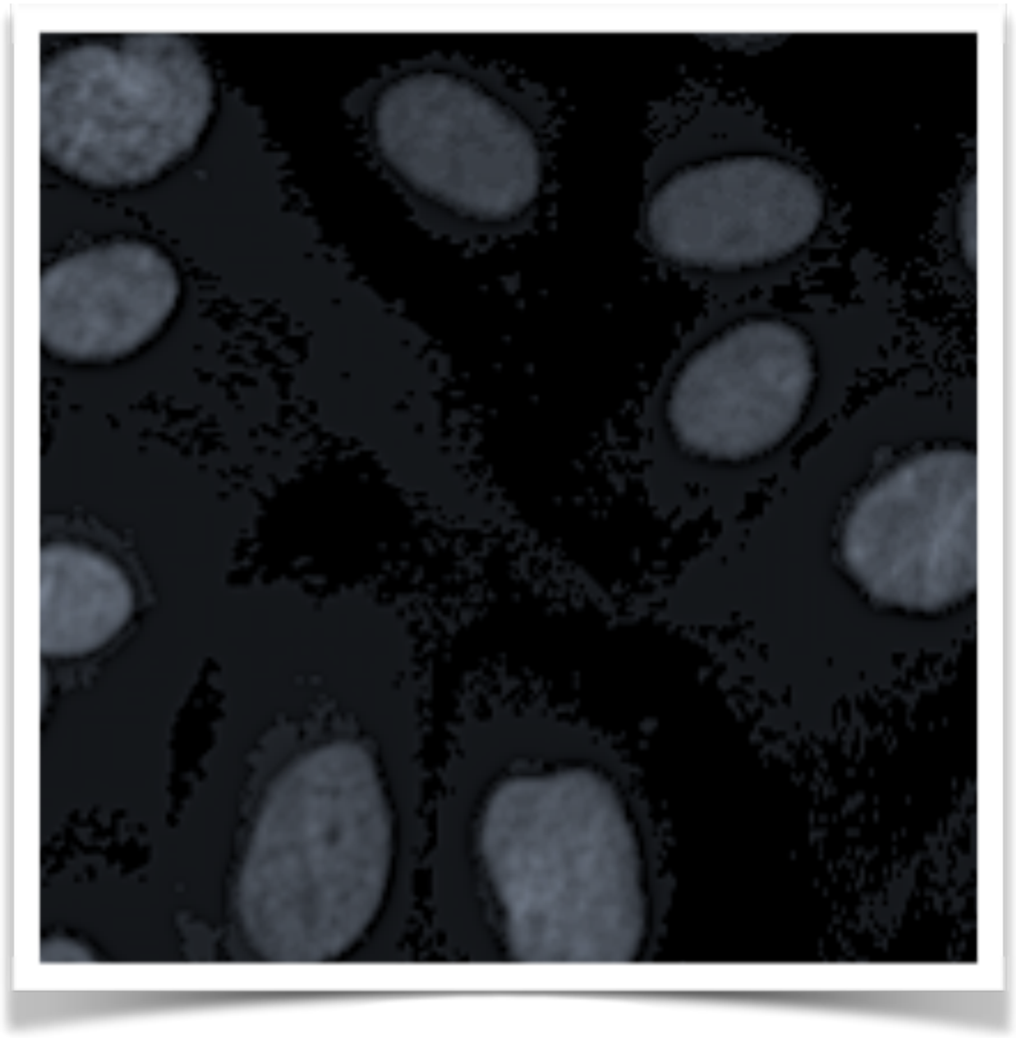


Training DeepCell

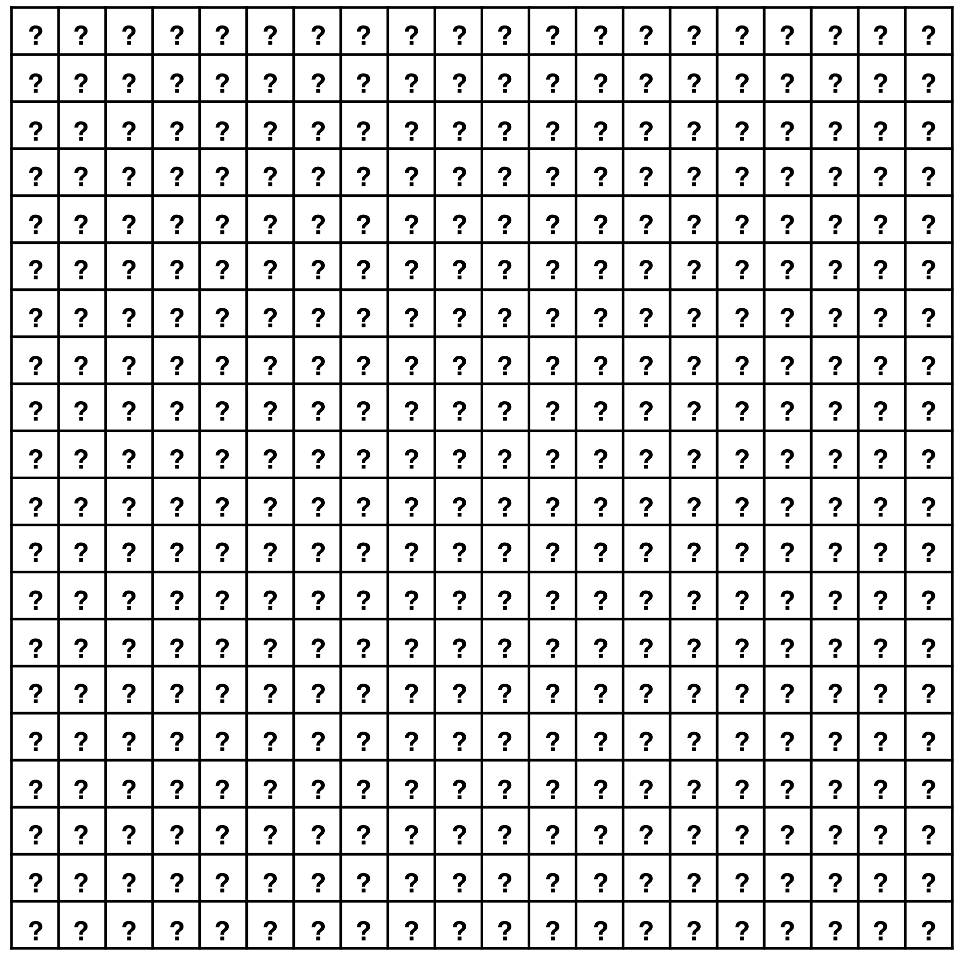


Predicting with DeepCell

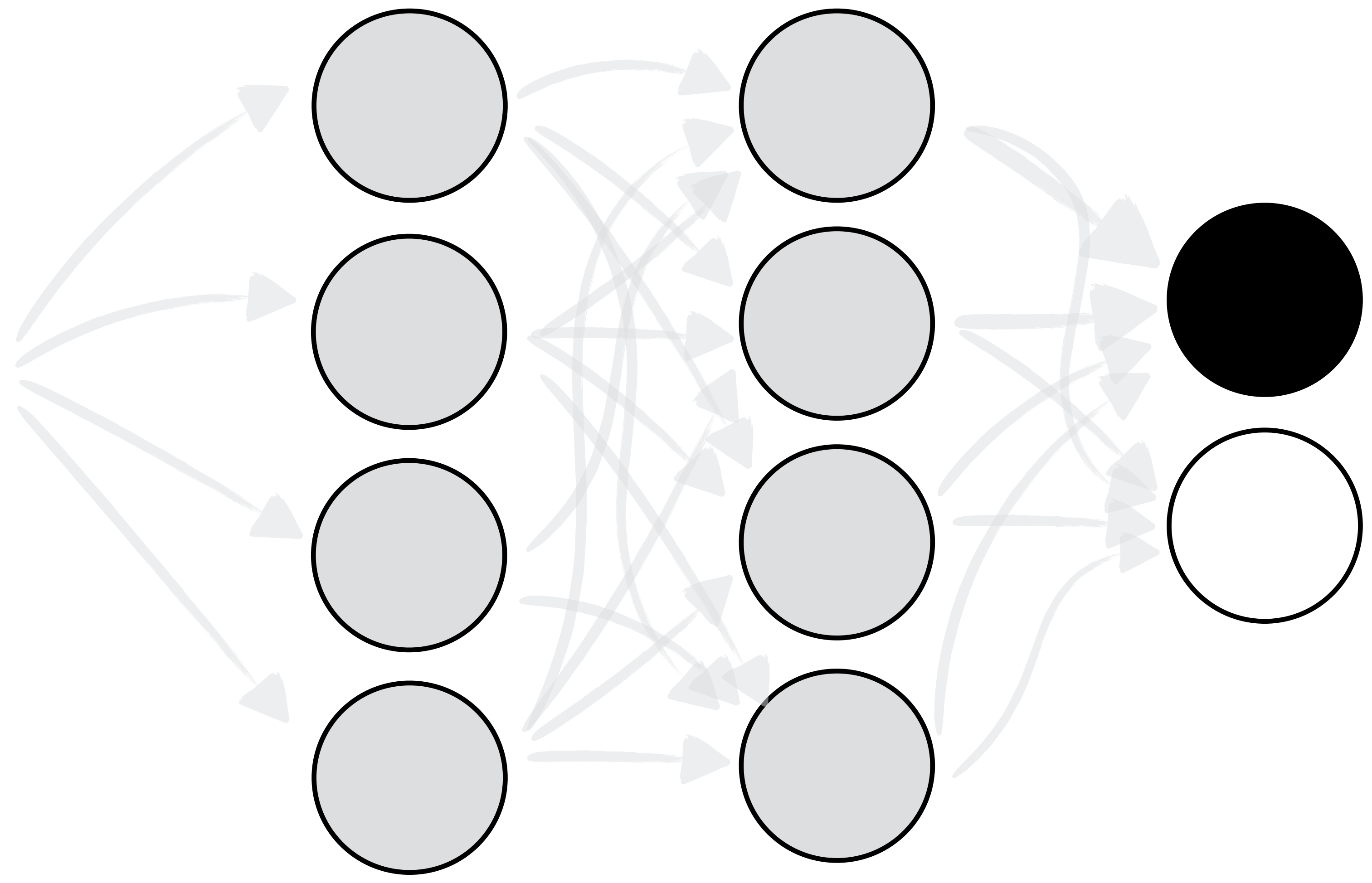
Test image



Predicted segmentation



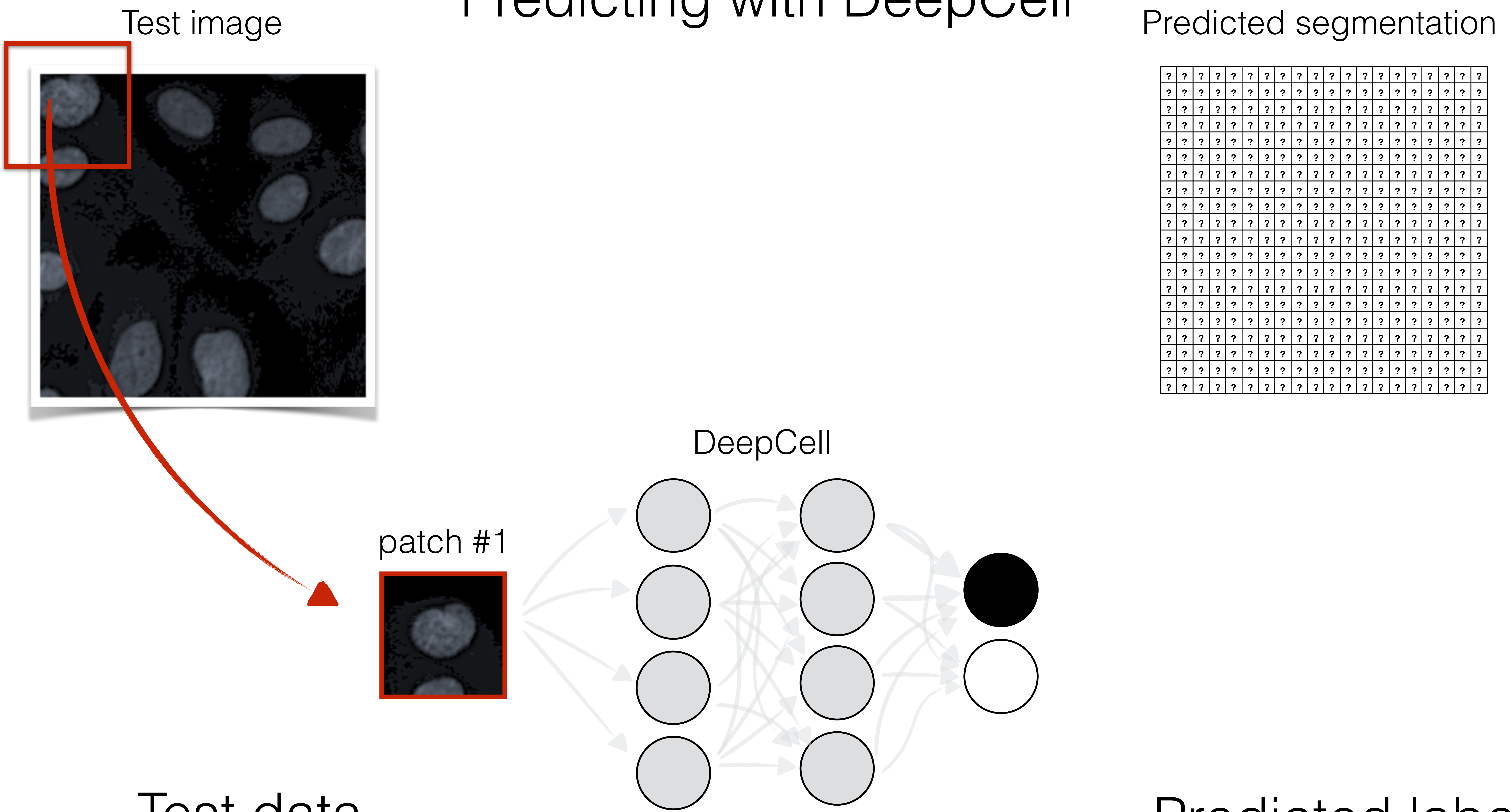
DeepCell



Test data

Predicted labels

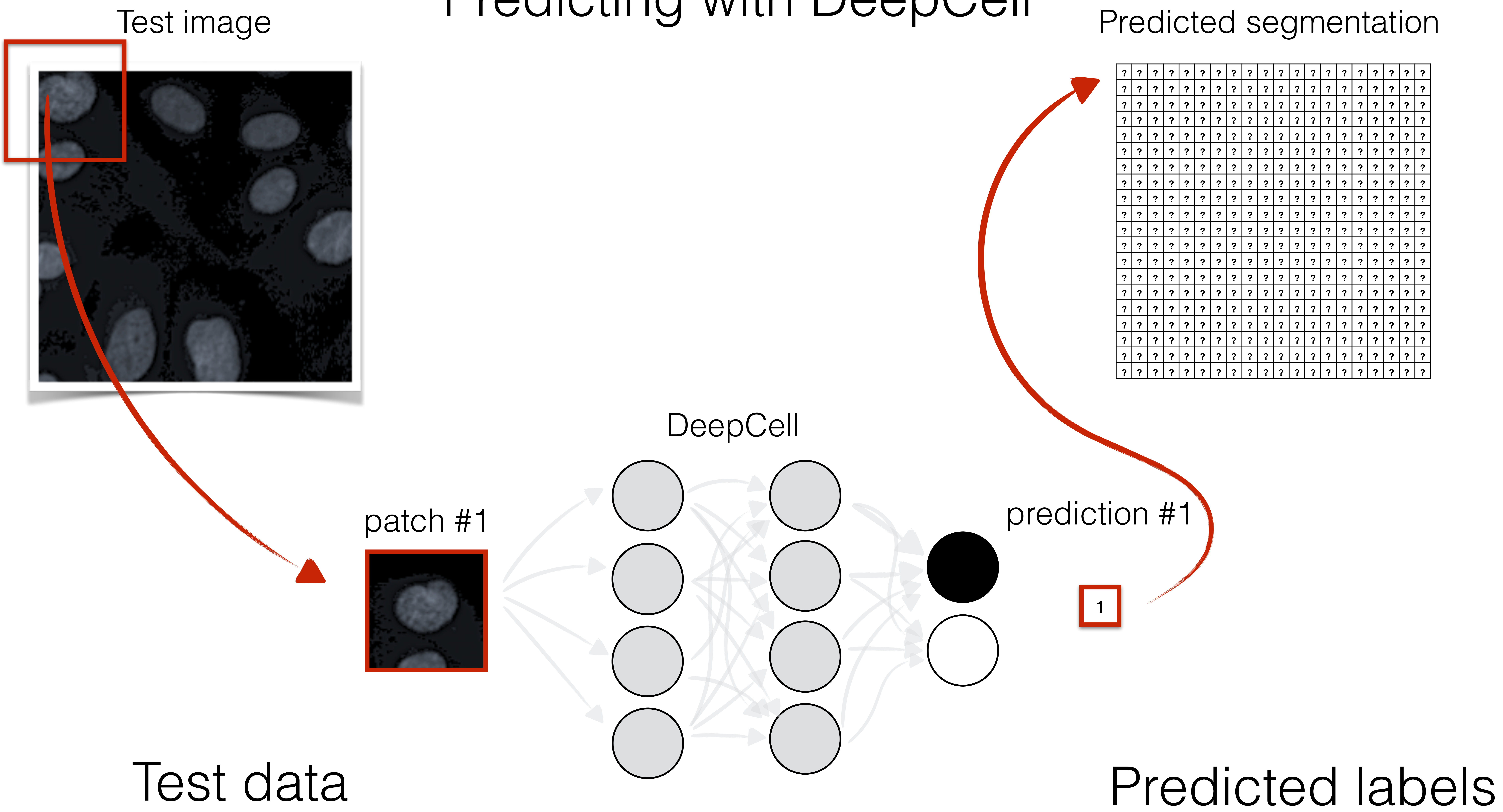
Predicting with DeepCell



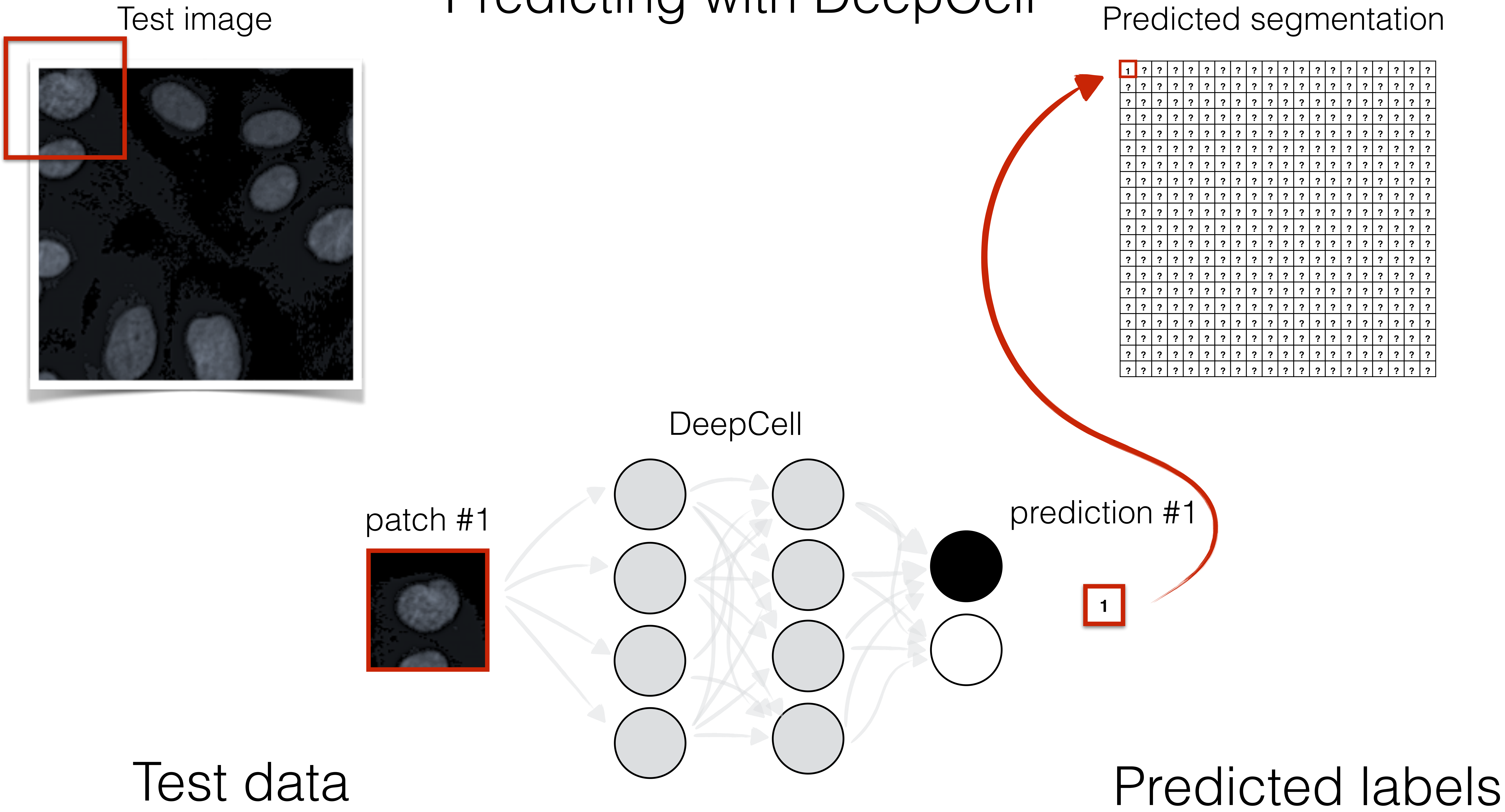
Test data

Predicted labels

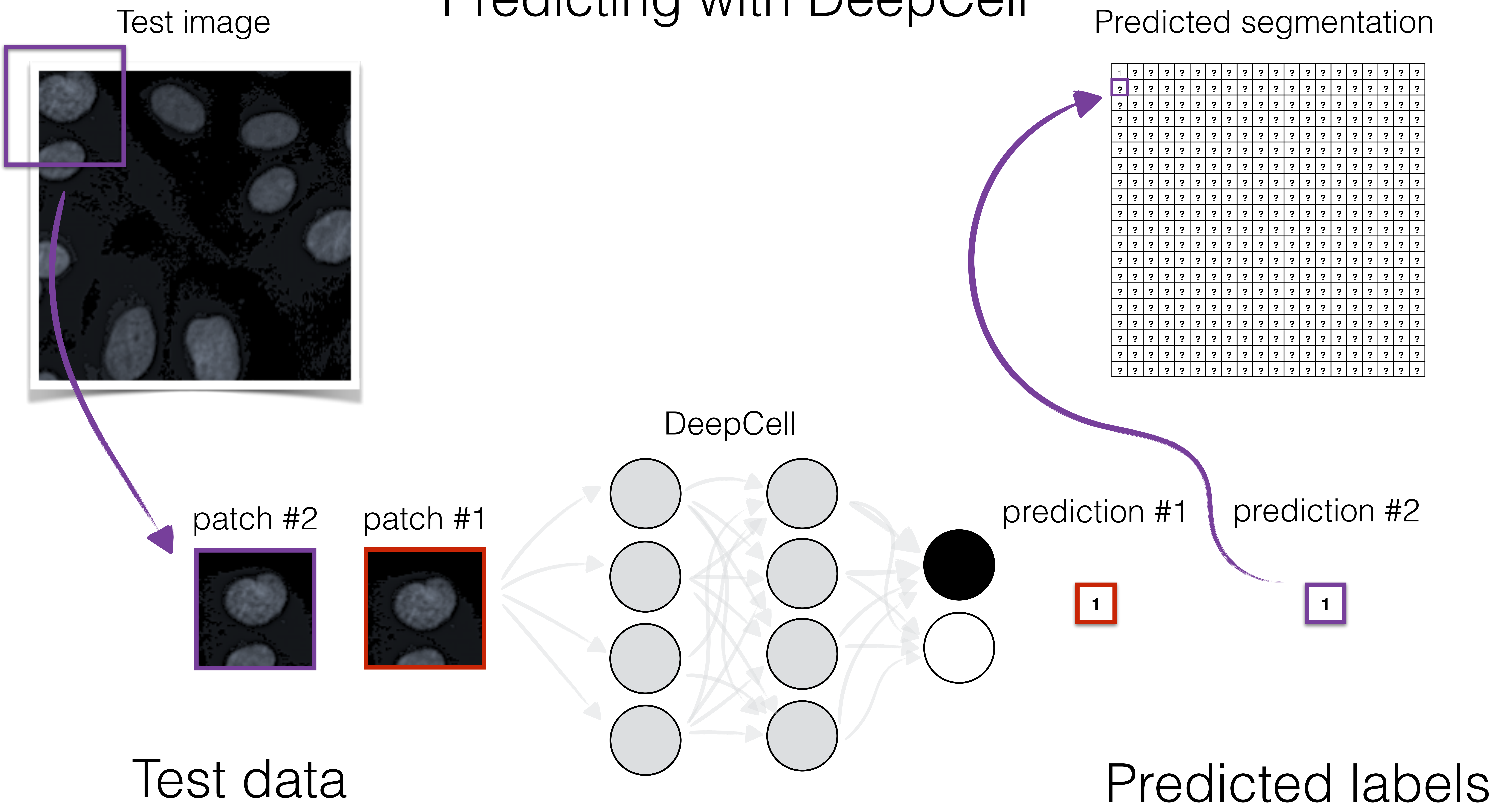
Predicting with DeepCell



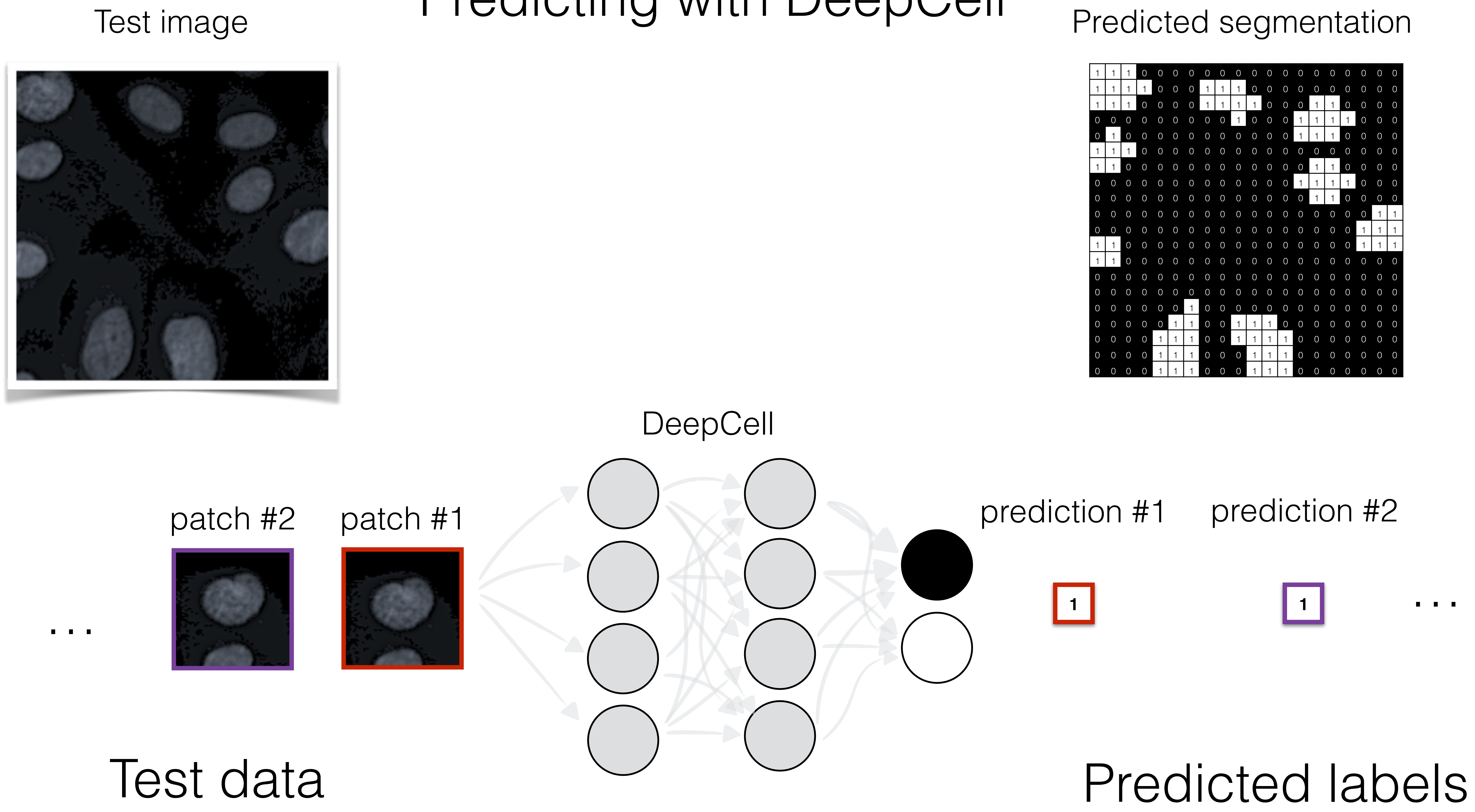
Predicting with DeepCell



Predicting with DeepCell



Predicting with DeepCell

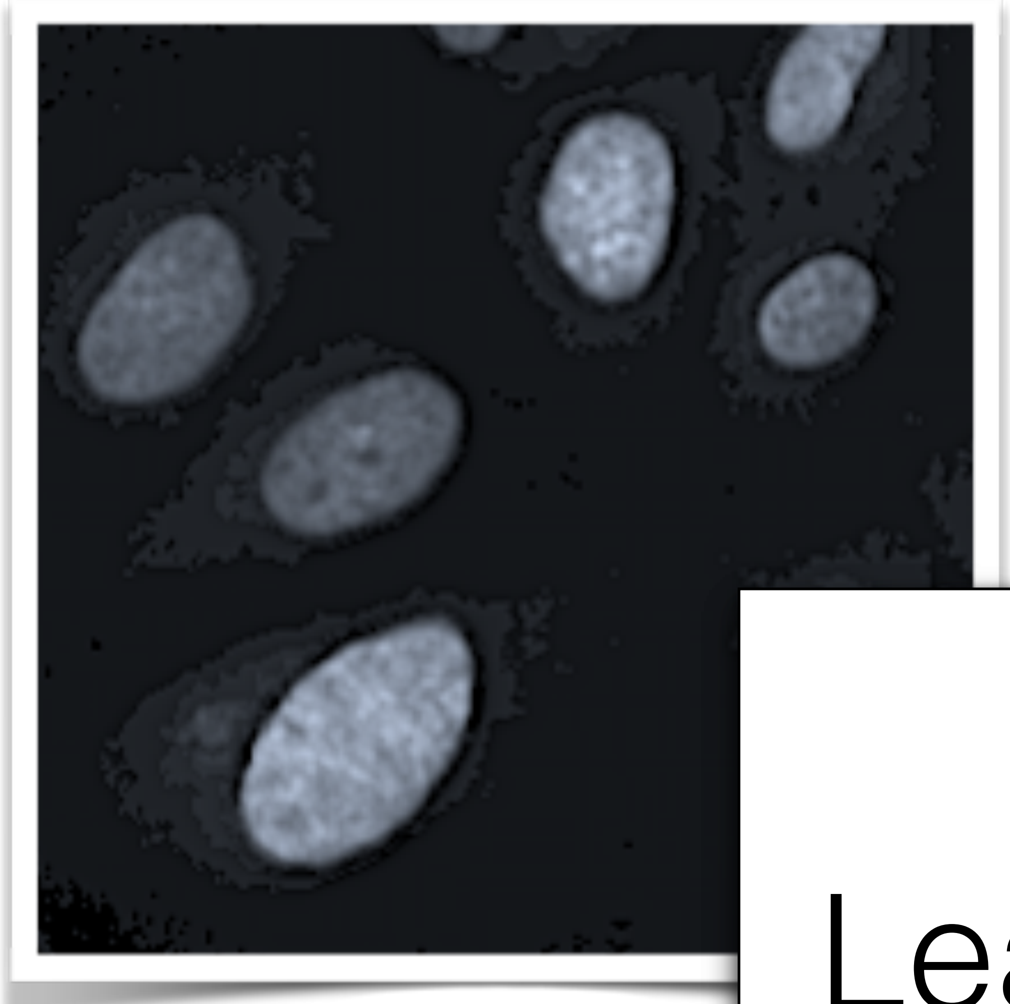


Original Image

(Fluorescent)

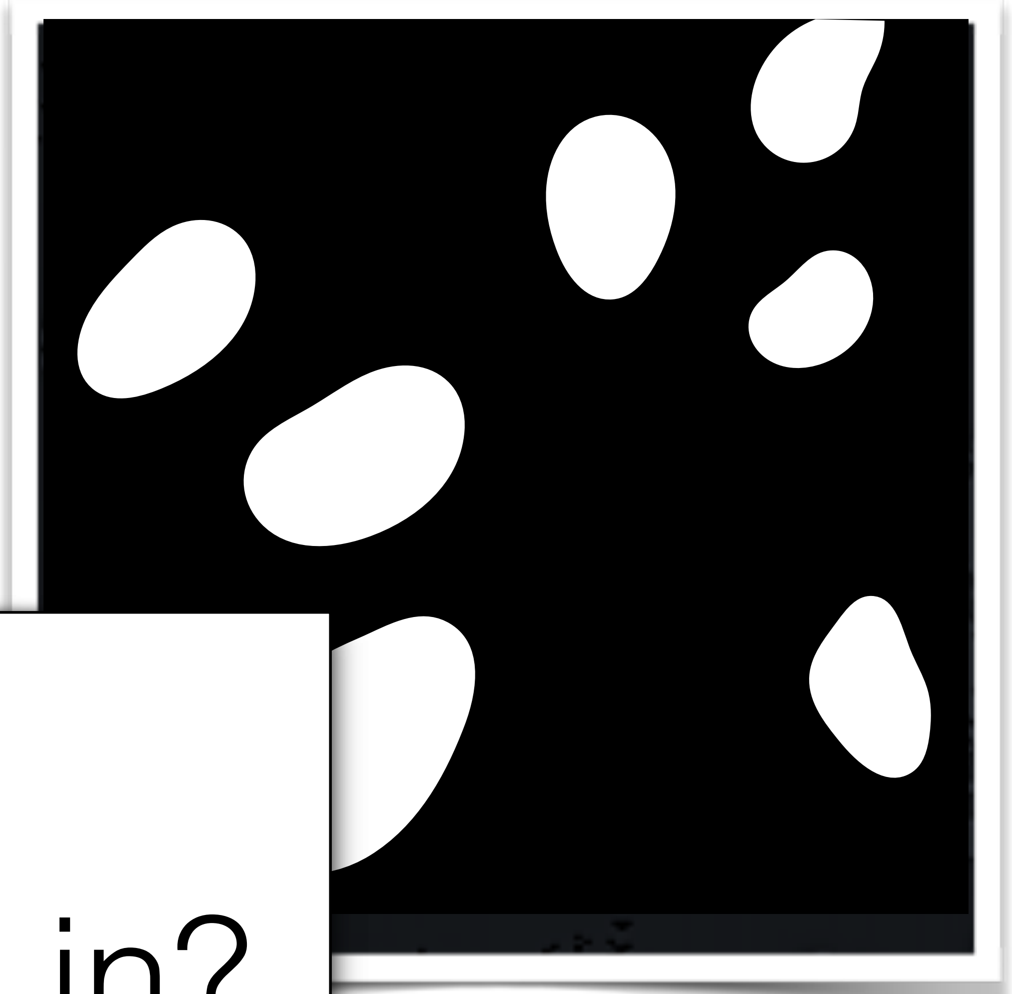
Filtering
contrasting
denoising

Preprocessed Image

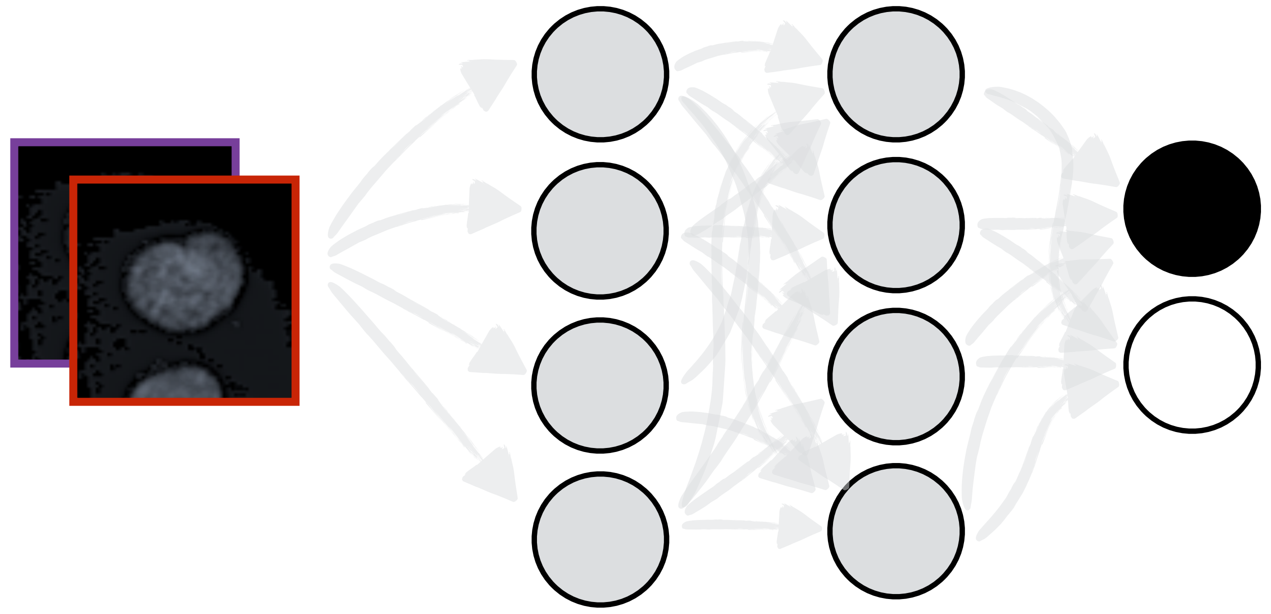


Thresholding

Segmentation mask



Can Deep Learning step in?



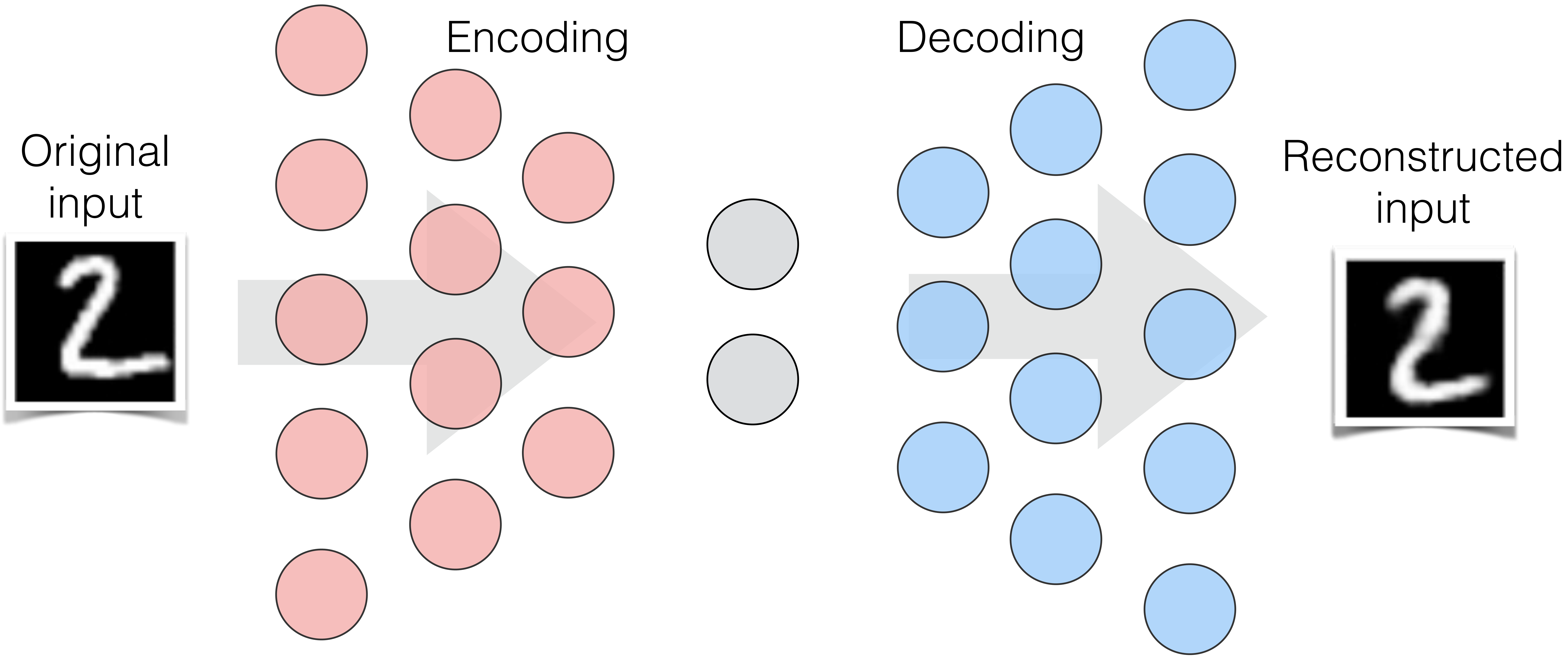
DeepCell (D. Van Valen et al.)

Approach I

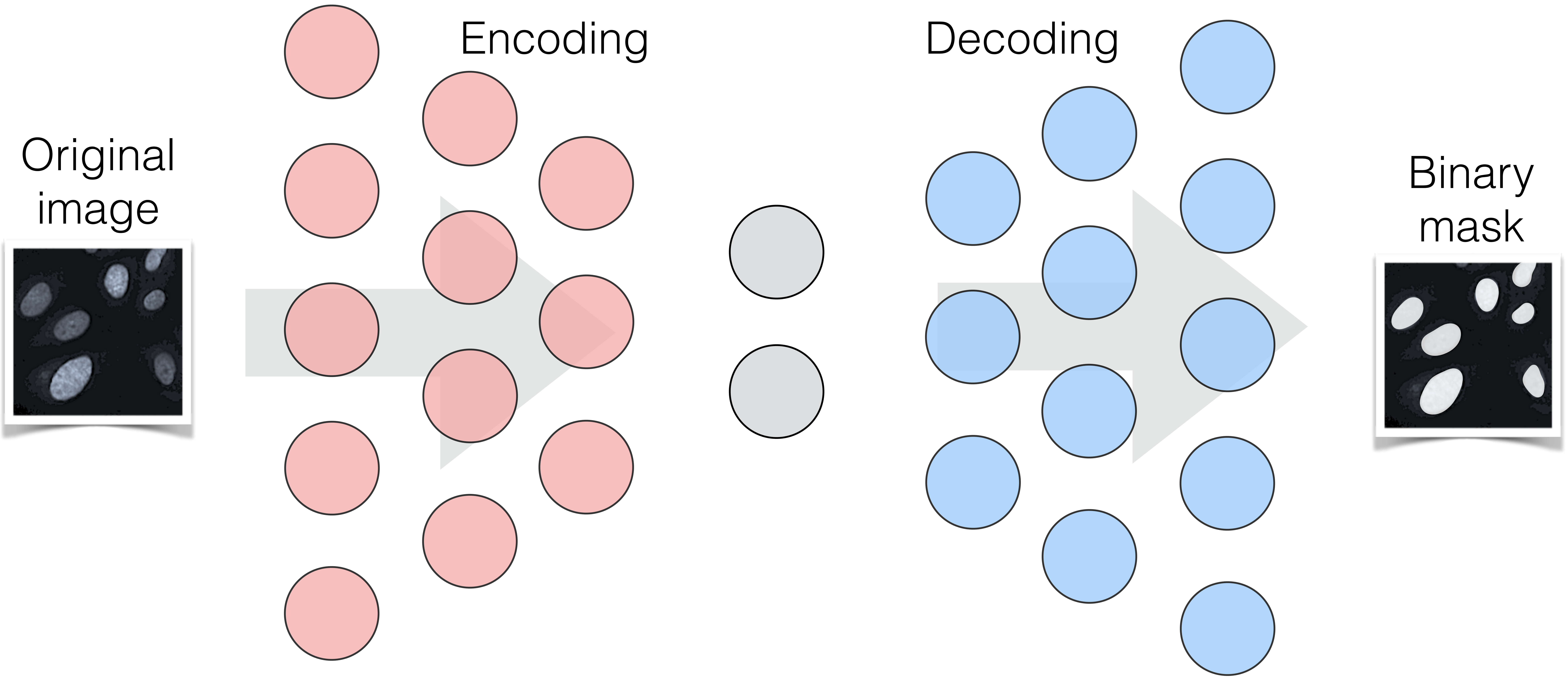
Approach II

Approach III

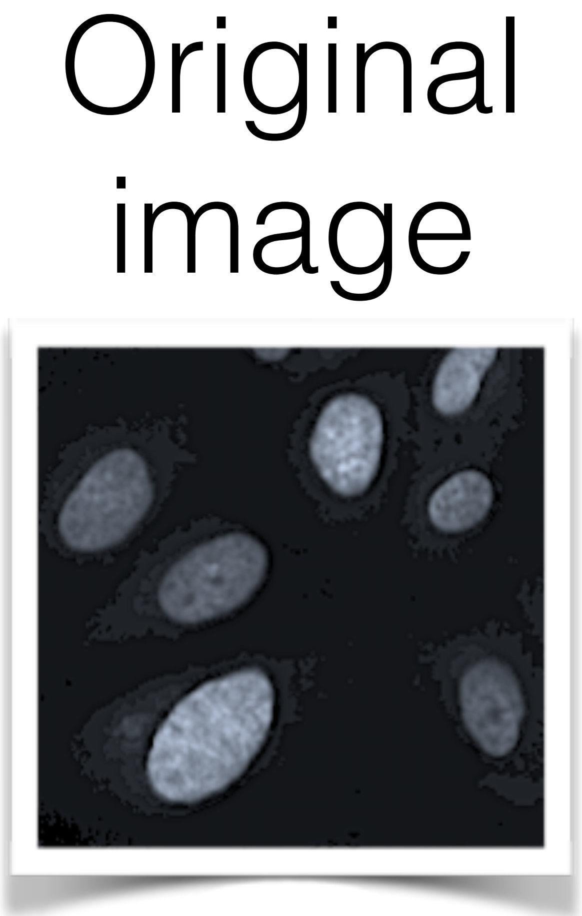
Autoencoder

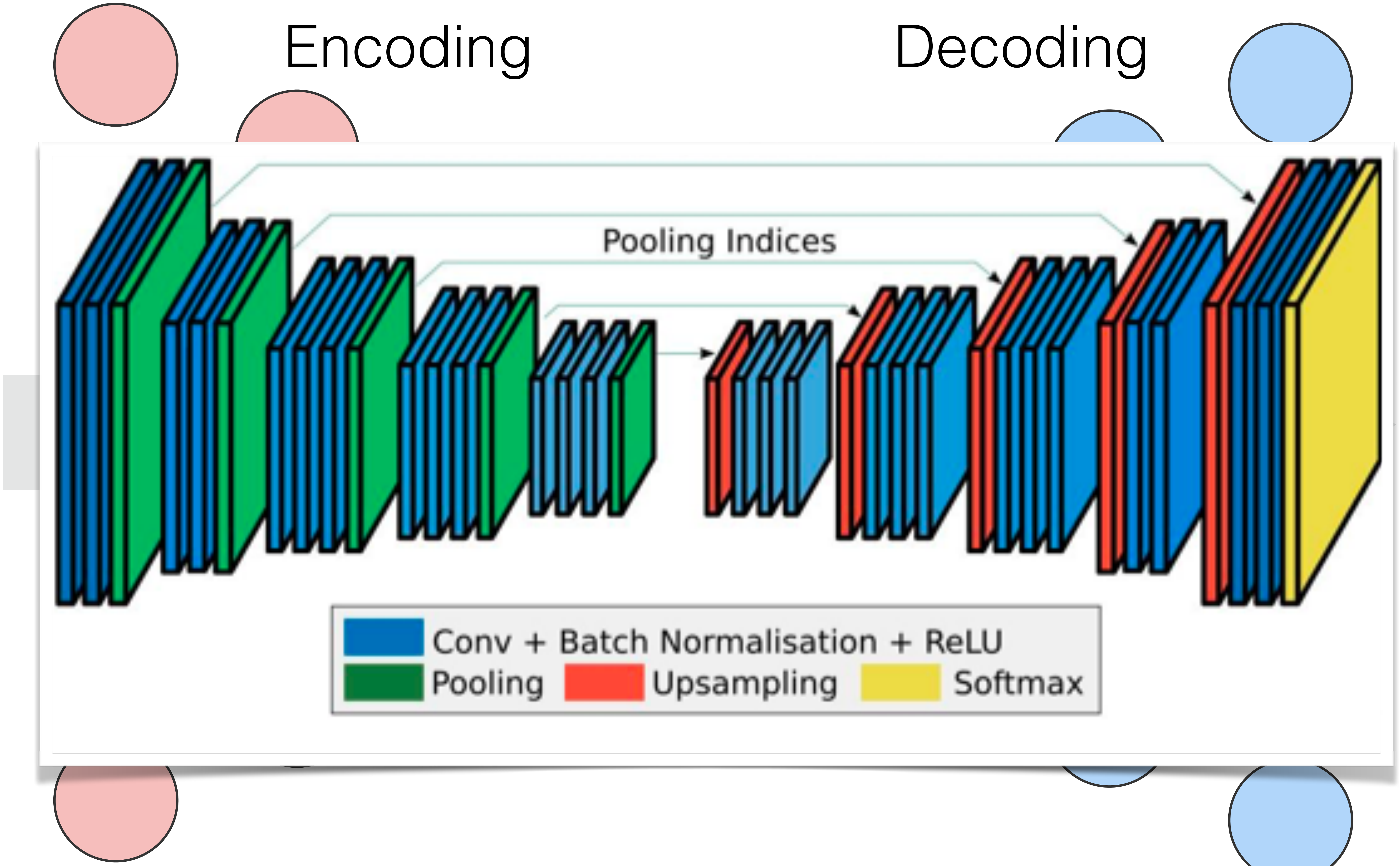


Autoencoder



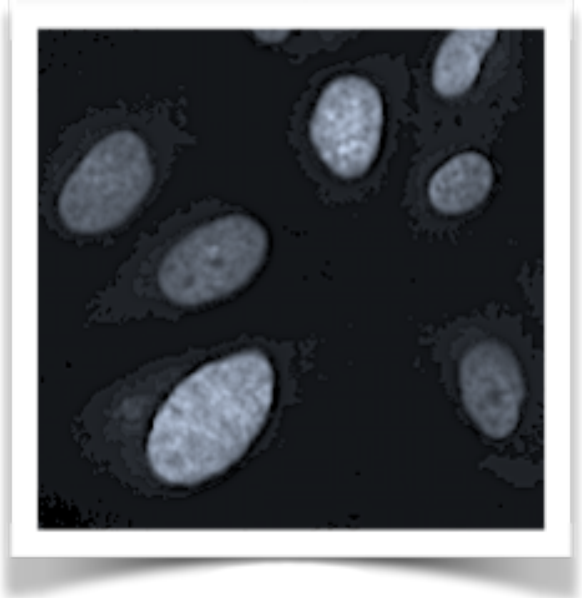
Segnet Architecture





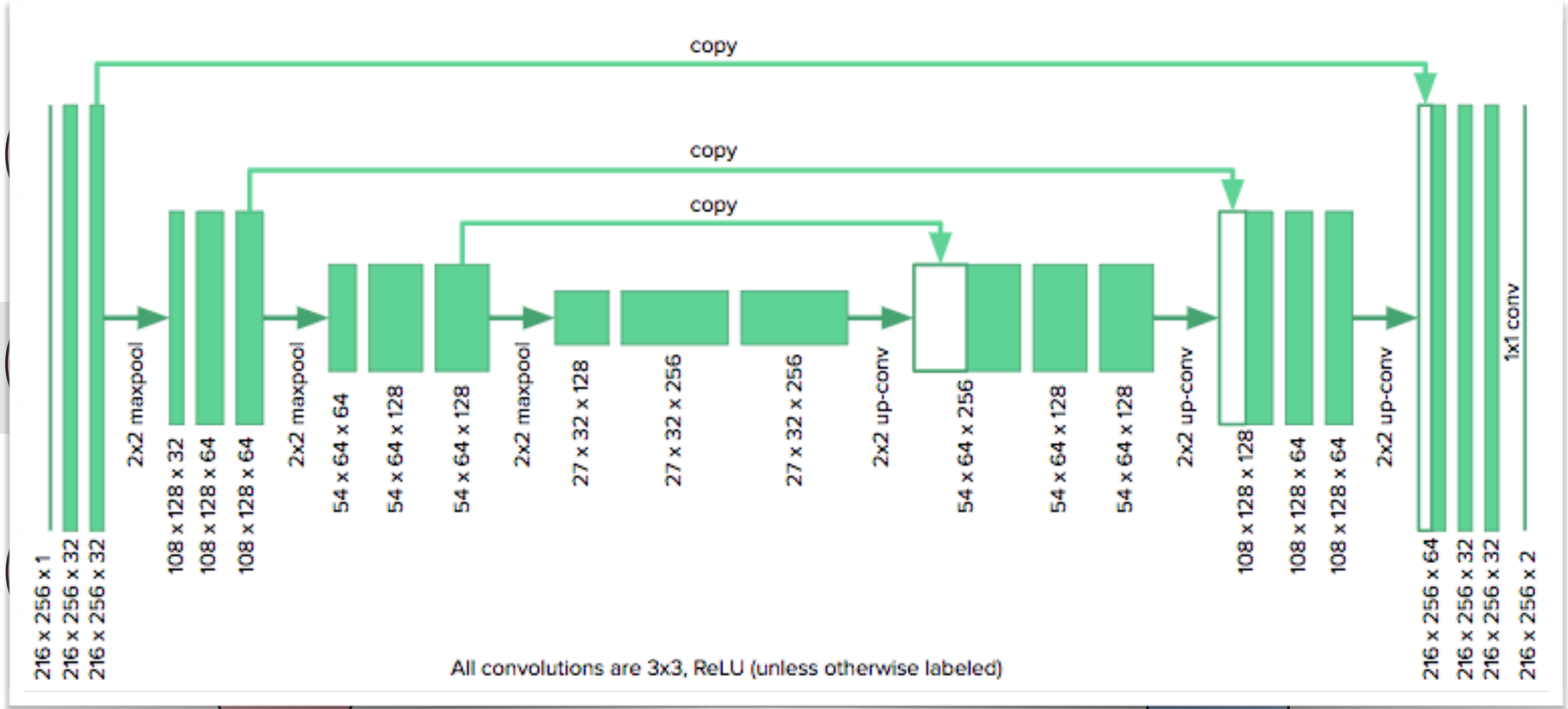
U-net Architecture

Original image



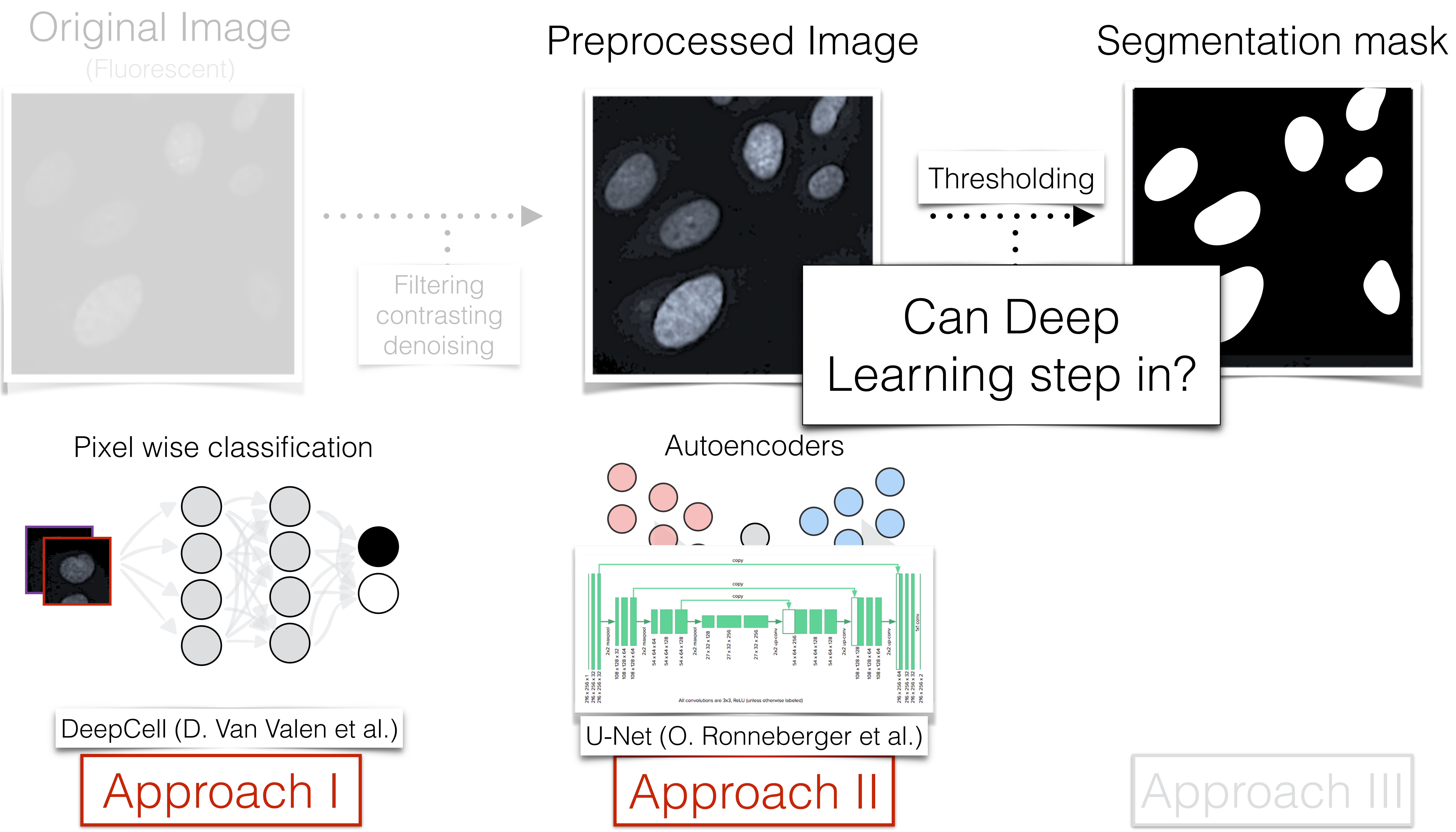
Encoding

Decoding



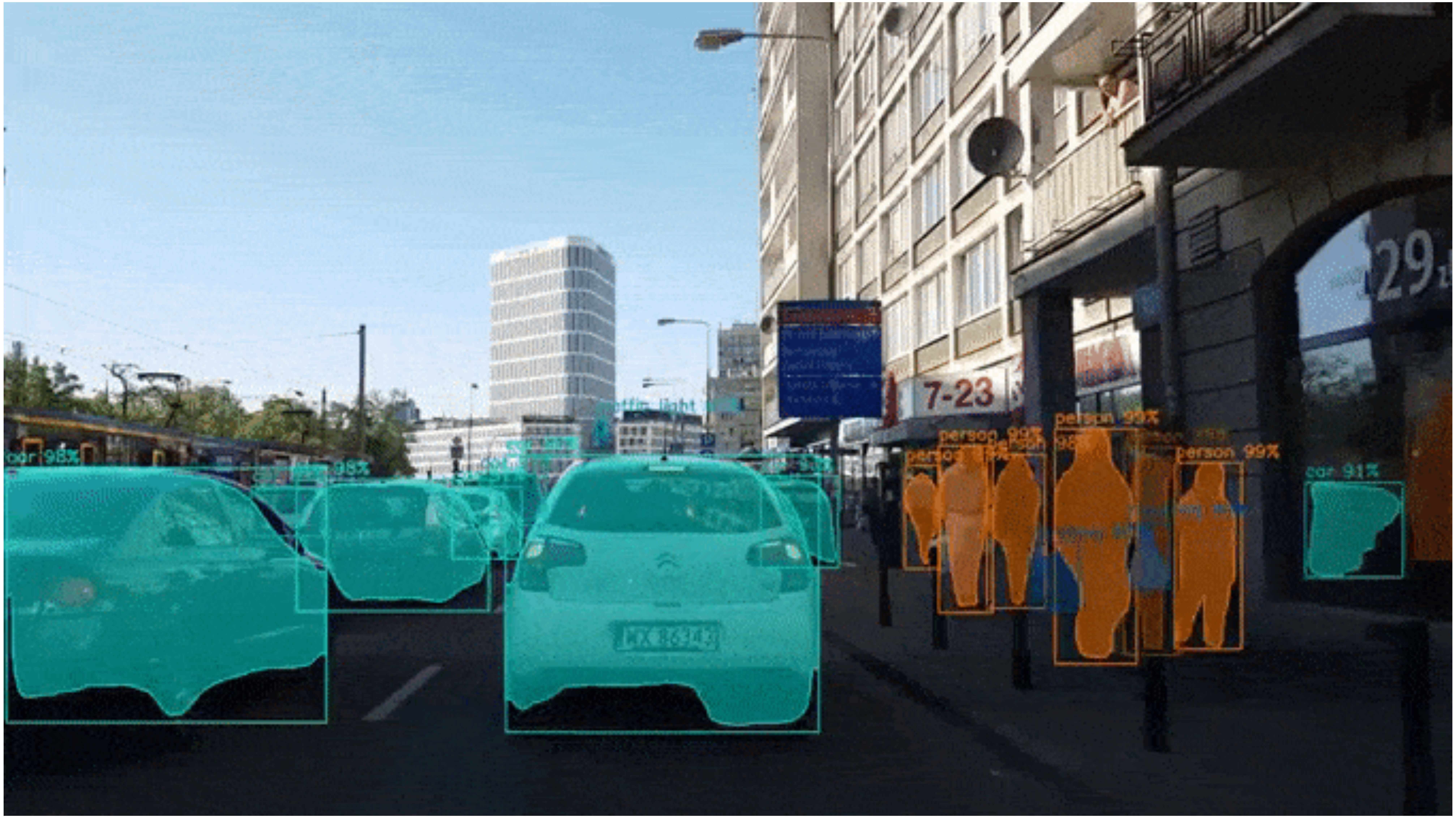
Binary mask

U-Net (O. Ronneberger et al.)



Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Can Deep
Learning step in?
Pixel wise classification
DeepCell (D. Van Valen et al.)
Approach I
Autoencoders
copy
All convolutions are 3x3, ReLU (unless otherwise labeled)
U-Net (O. Ronneberger et al.)
Approach II
Approach III

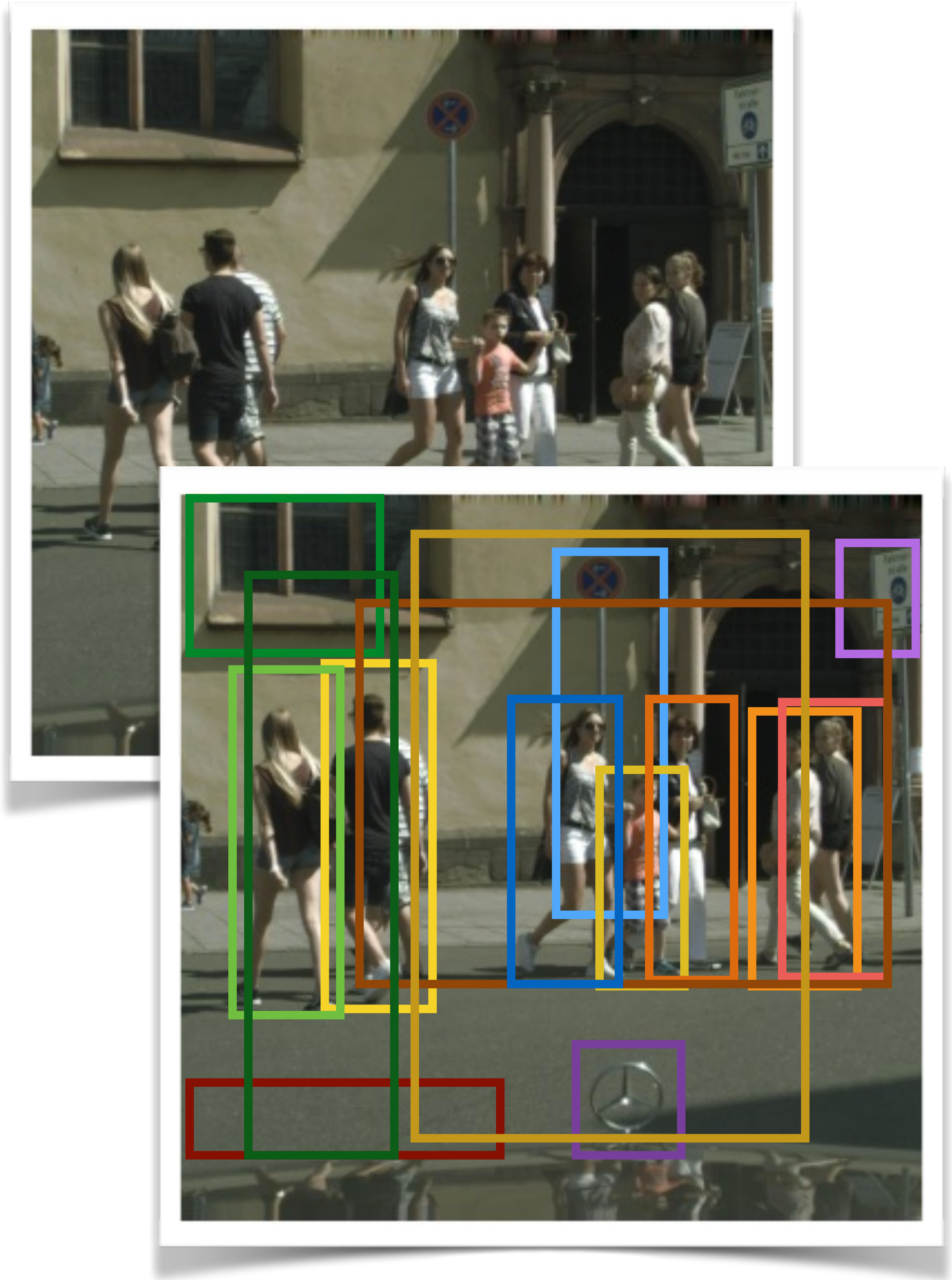
Mask R-CNN



He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r cnn. arXiv preprint arXiv:1703.06870.

Mask R-CNN

Mask R-CNN



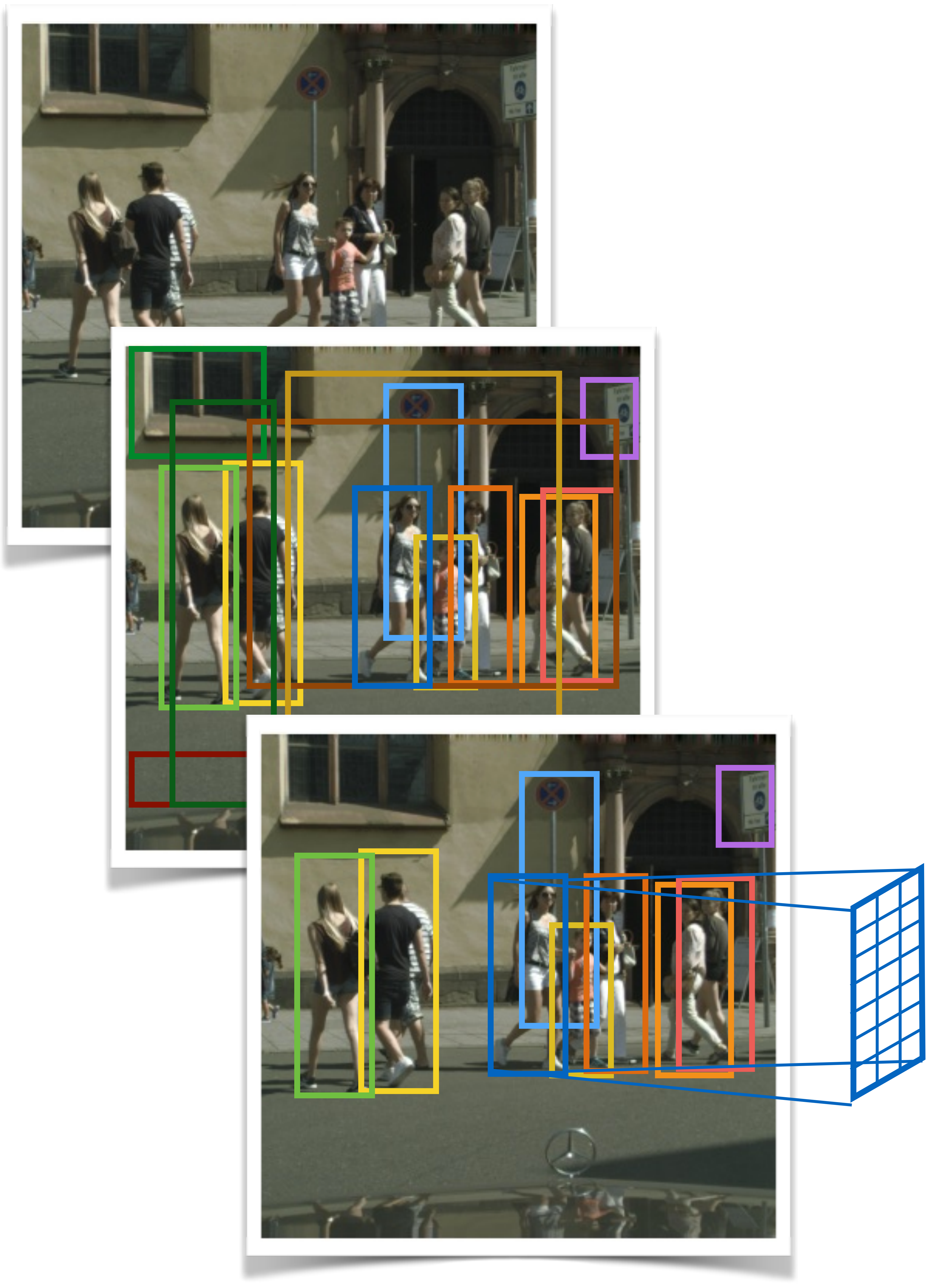
1. Proposes bounding boxes for objects (RoI)

Mask R-CNN

1. Proposes bounding boxes for objects (RoI)

2. Filters out bad RoIs

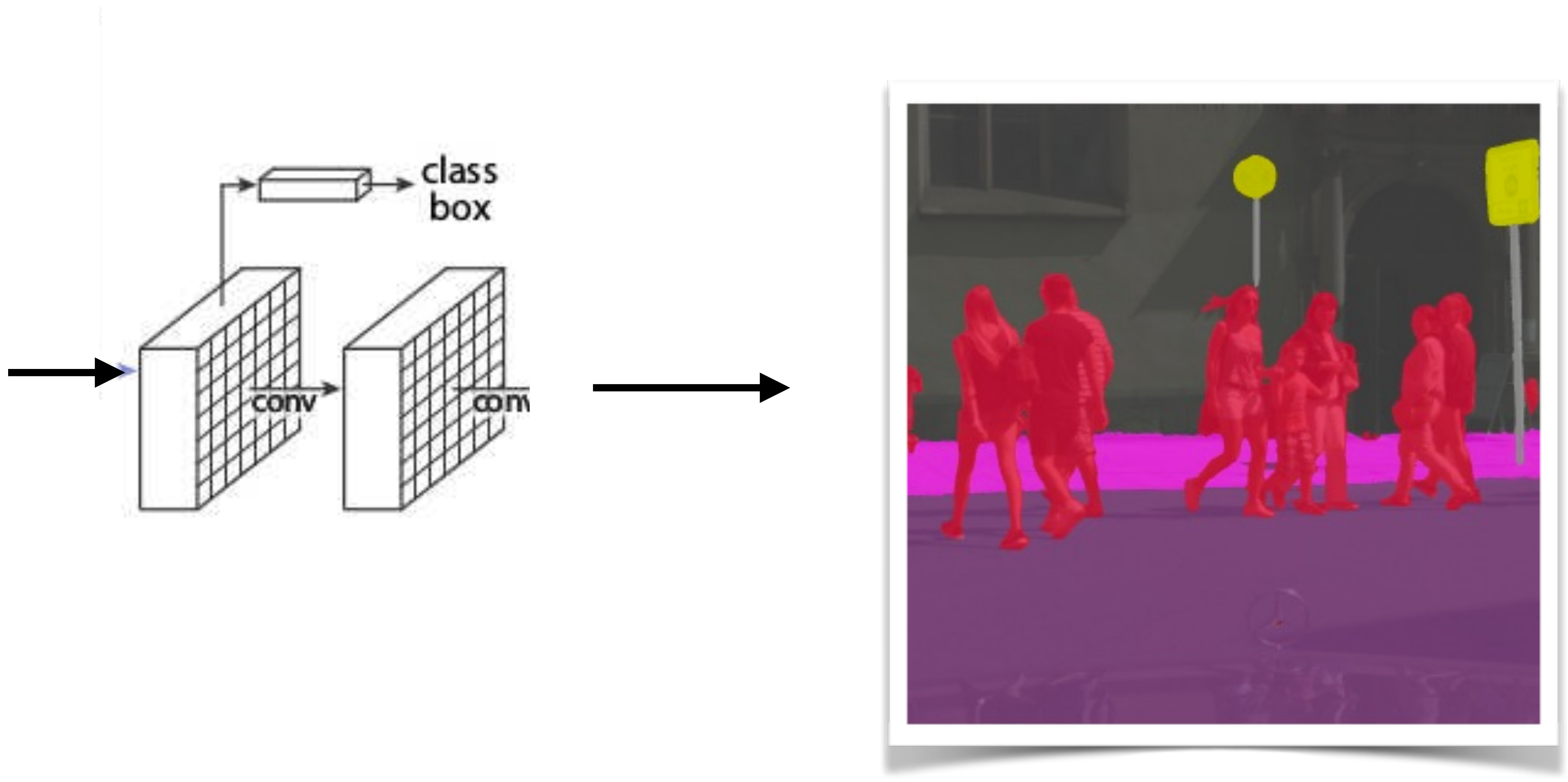
Mask R-CNN



1. Proposes bounding boxes for objects (RoI)

2. Filters out bad RoIs

3. For each RoI builds a mask

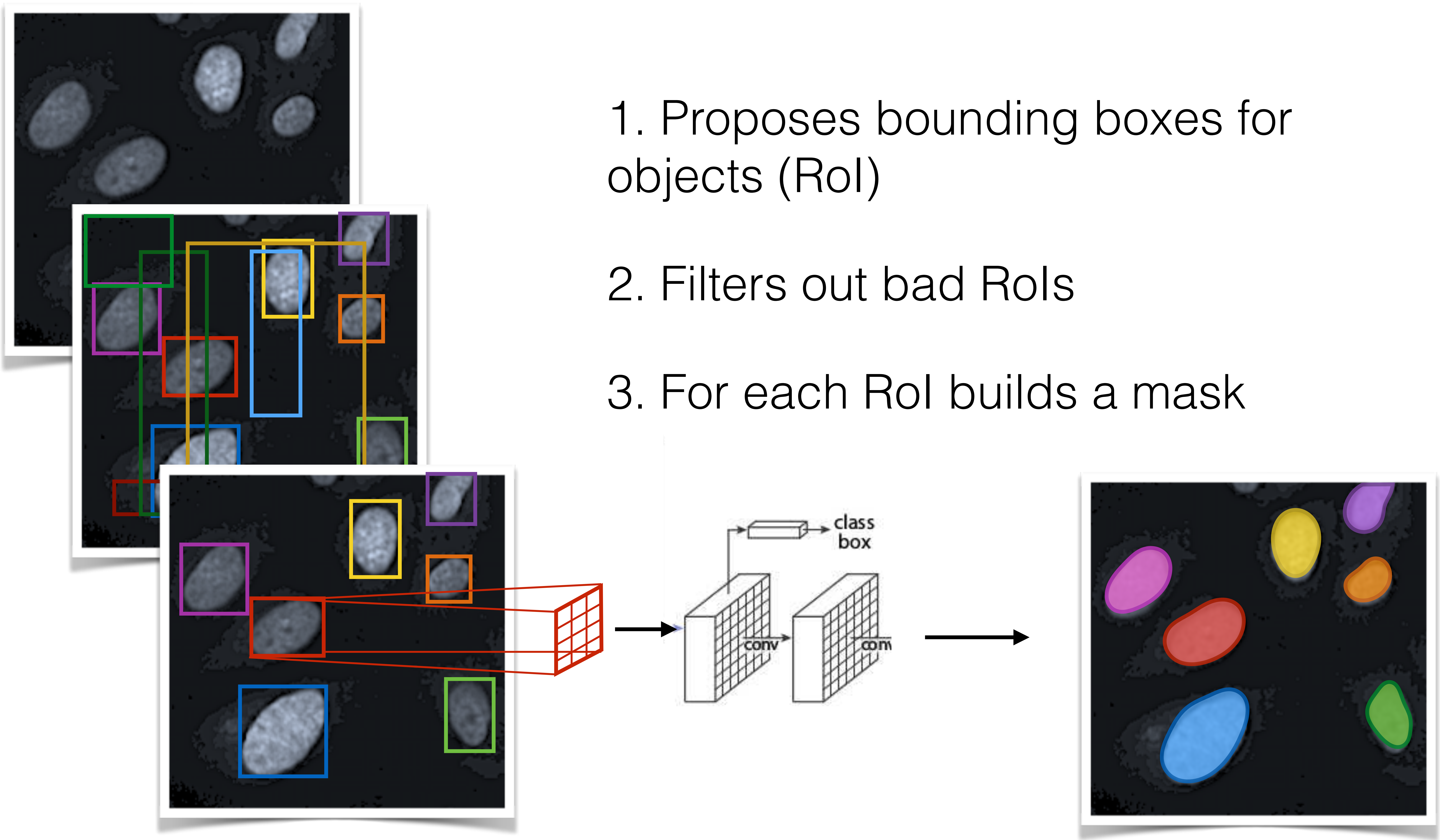


Mask R-CNN

1. Proposes bounding boxes for objects (RoI)

2. Filters out bad RoIs

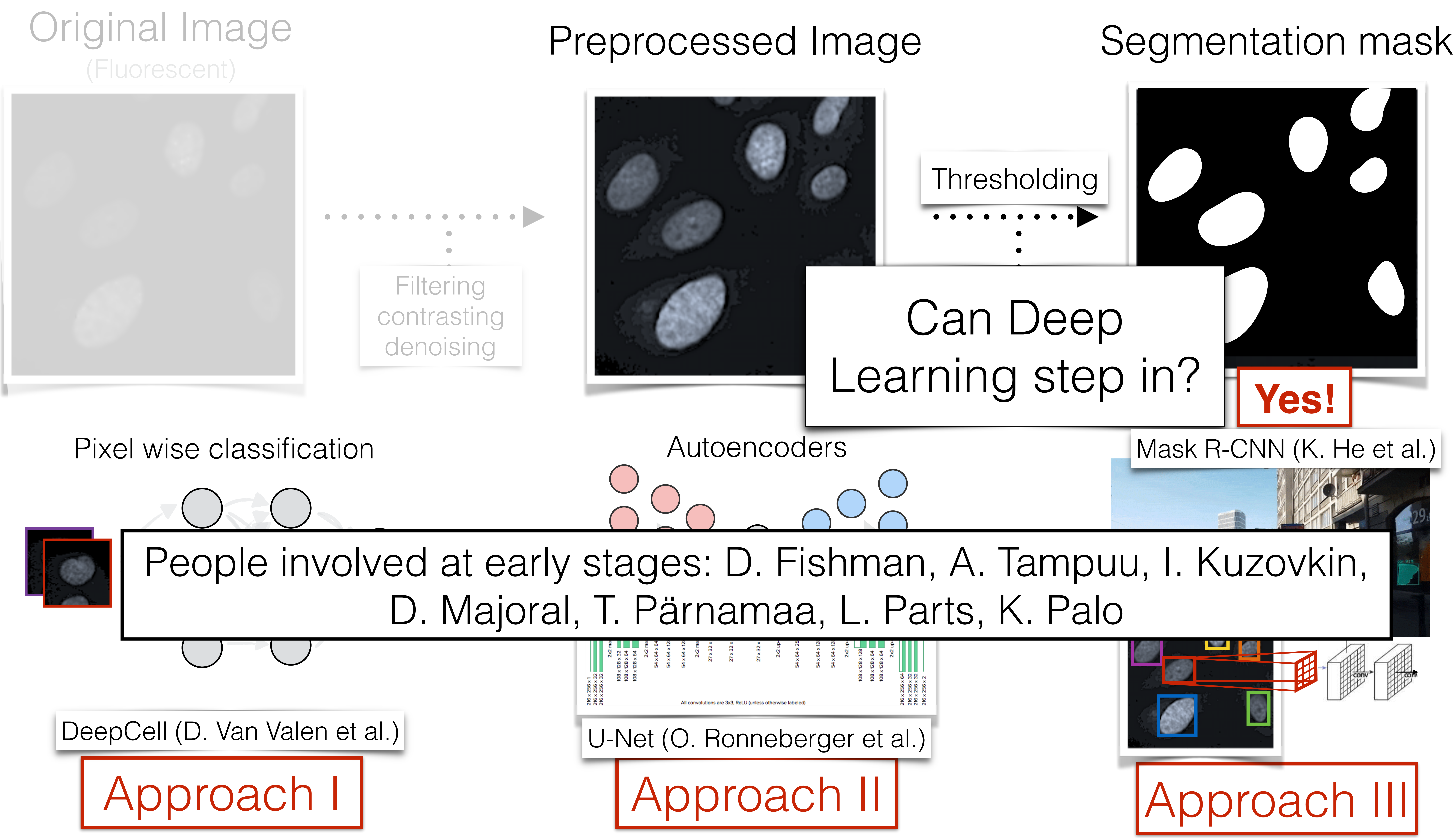
3. For each RoI builds a mask



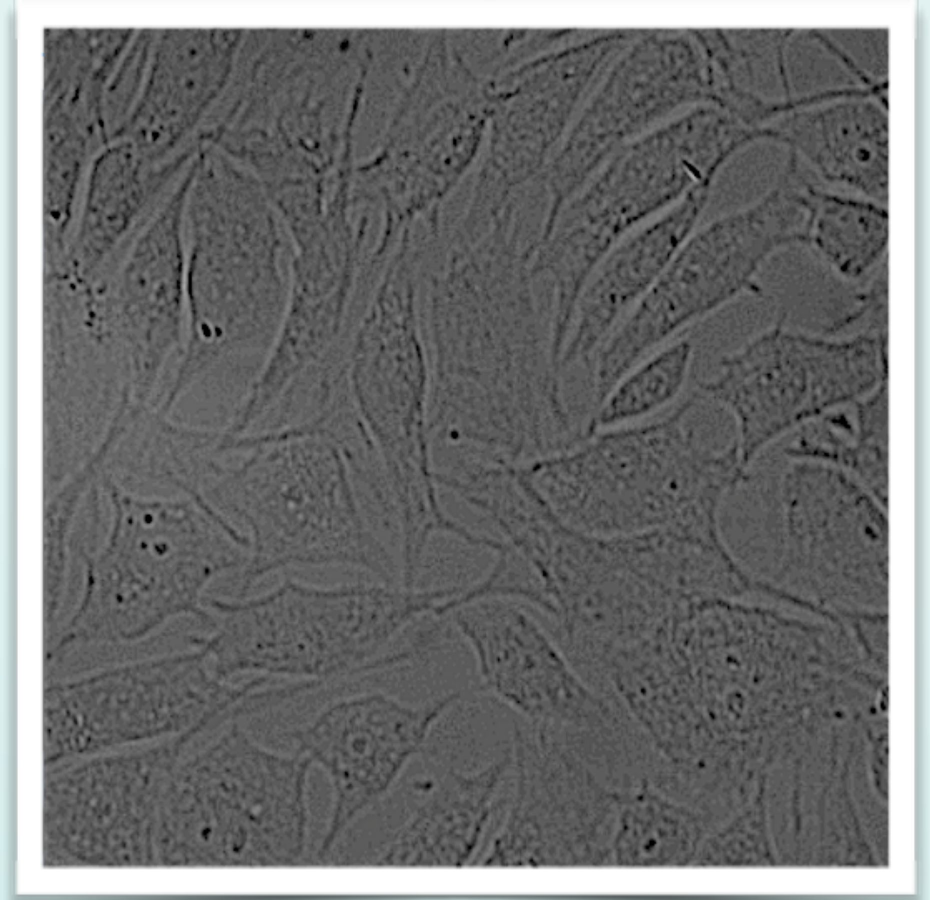
By Daniel Majoral

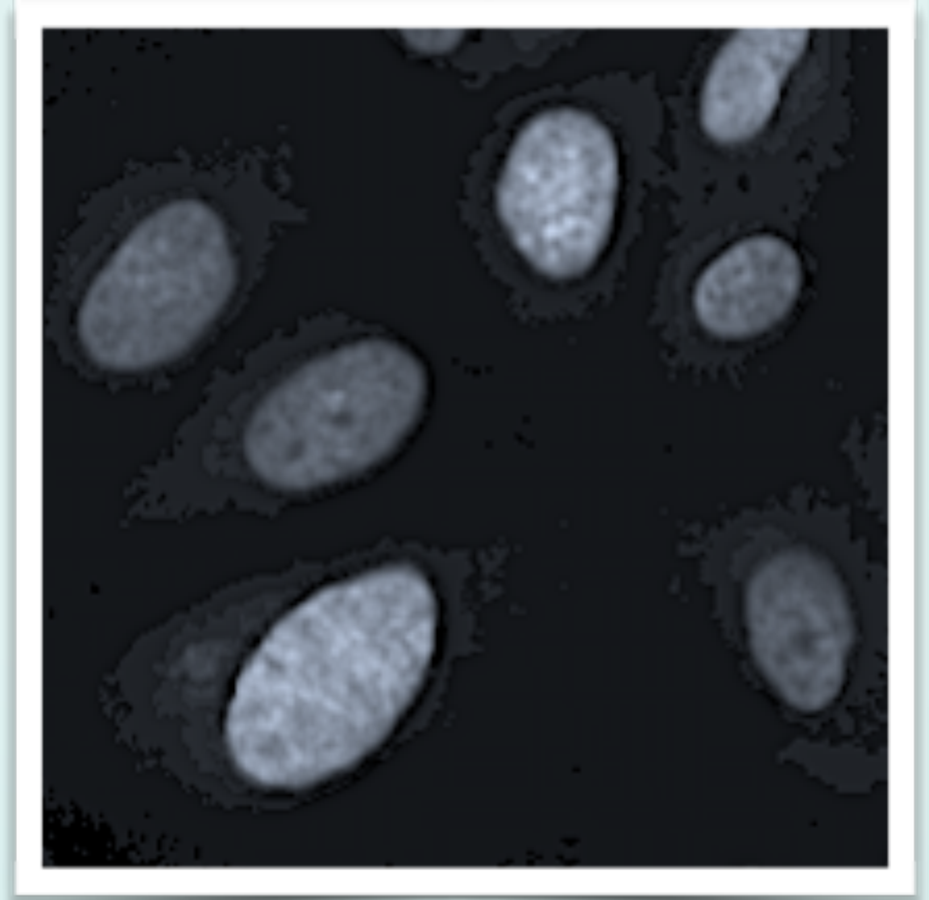

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Can Deep
Learning step in?
Pixel wise classification
DeepCell (D. Van Valen et al.)
Approach I
Autoencoders
U-Net (O. Ronneberger et al.)
Approach II
Mask R-CNN (K. He et al.)
Approach III

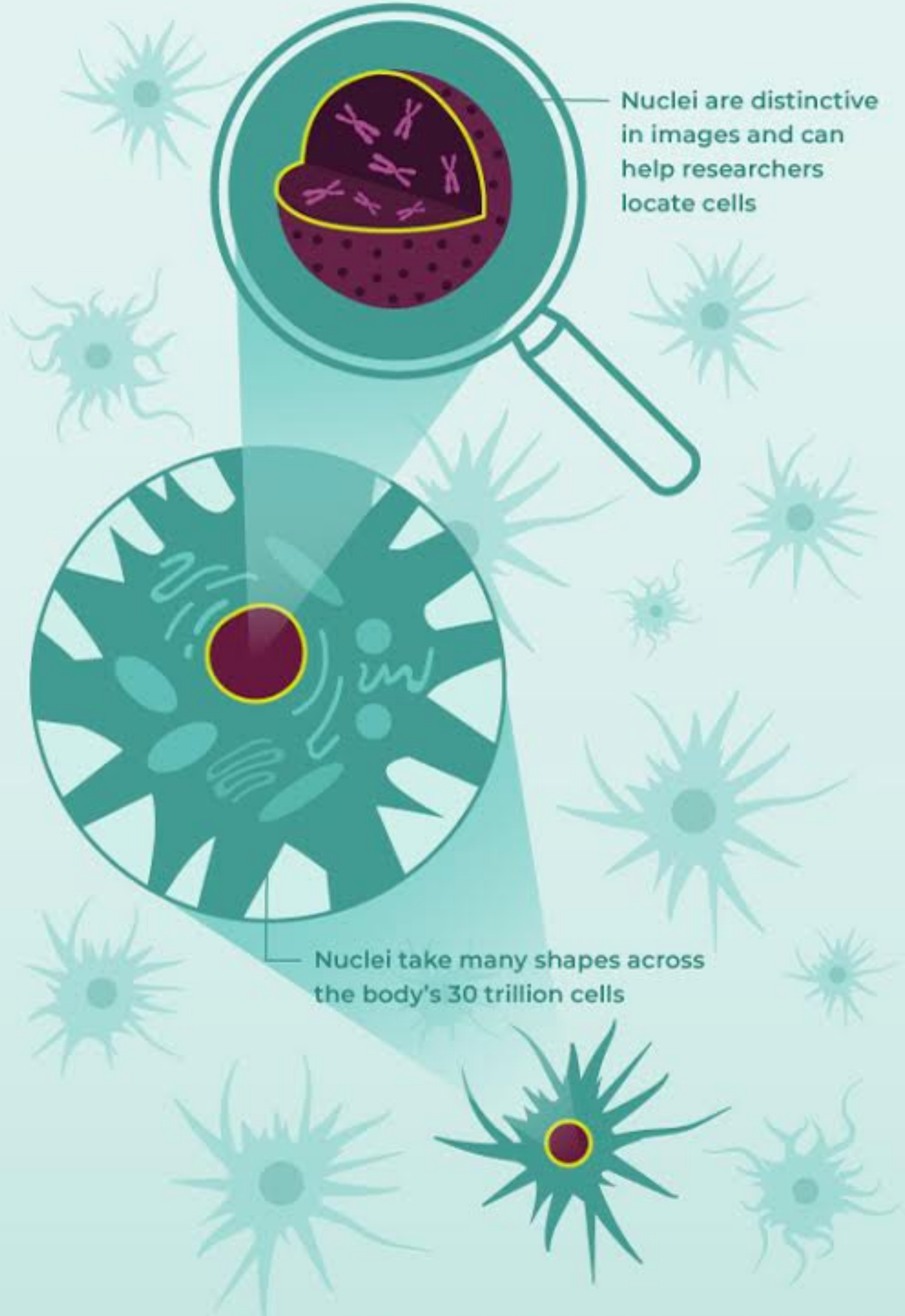

Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Can Deep
Learning step in?
Yes!
Pixel wise classification
Autoencoders
Mask R-CNN (K. He et al.)
DeepCell (D. Van Valen et al.)
U-Net (O. Ronneberger et al.)
Approach I
Approach II
Approach III


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Can Deep
Learning step in?
Yes!
Pixel wise classification
Autoencoders
Mask R-CNN (K. He et al.)
People involved at early stages: D. Fishman, A. Tampuu, I. Kuzovkin,
D. Majoral, T. Pärnamaa, L. Parts, K. Palo
DeepCell (D. Van Valen et al.)
U-Net (O. Ronneberger et al.)
Approach I
Approach II
Approach III

Spot Nuclei. Speed Cures.



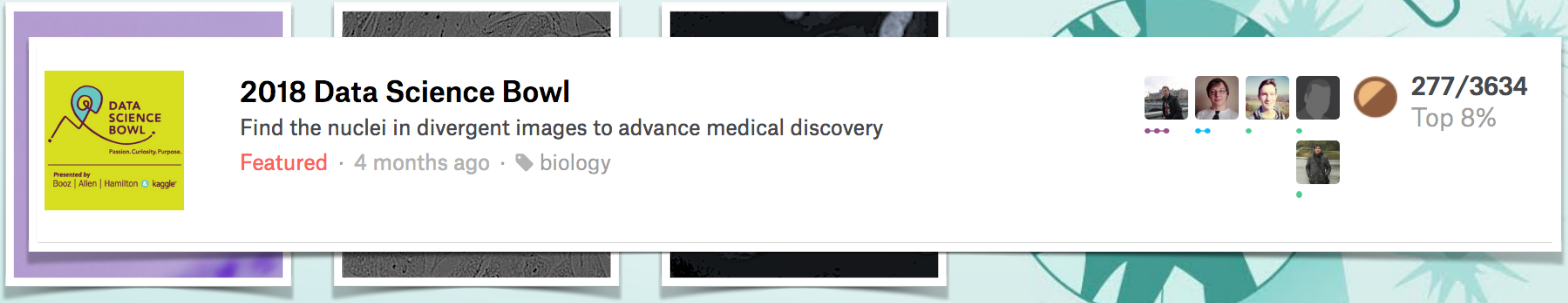


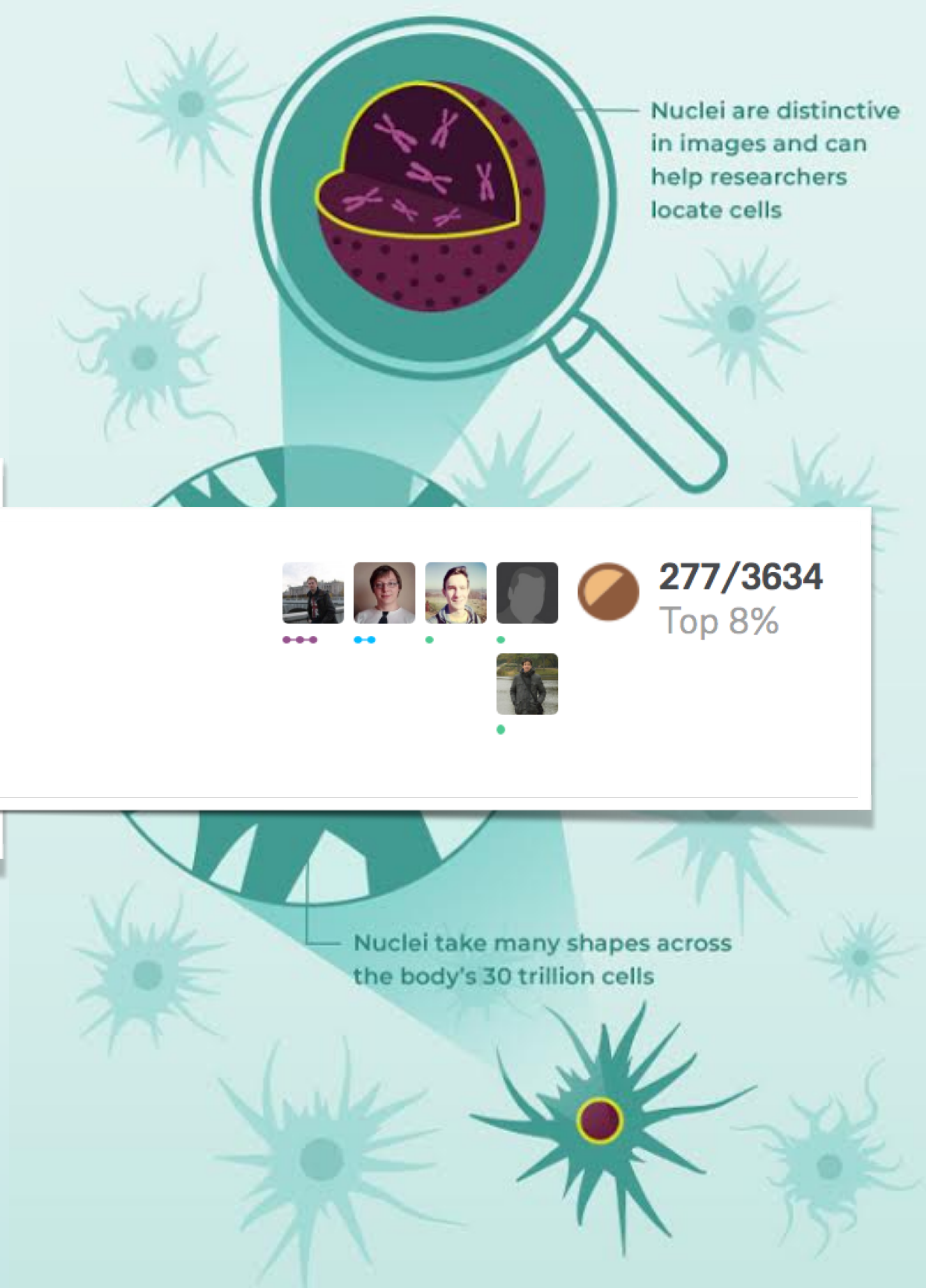


kaggle

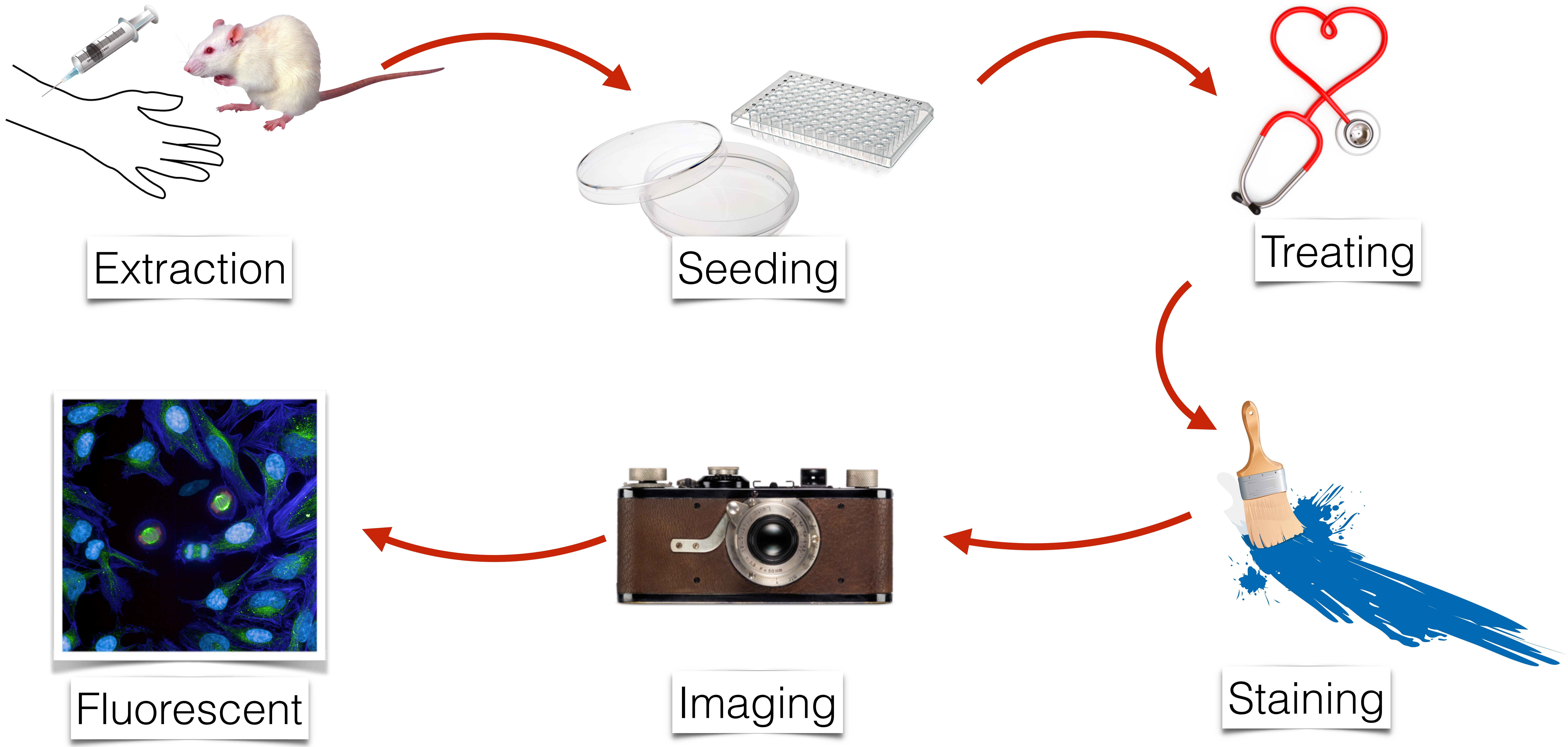
NATIONAL DATA SCIENCE BOWL

Spot Nuclei. Speed Cures.

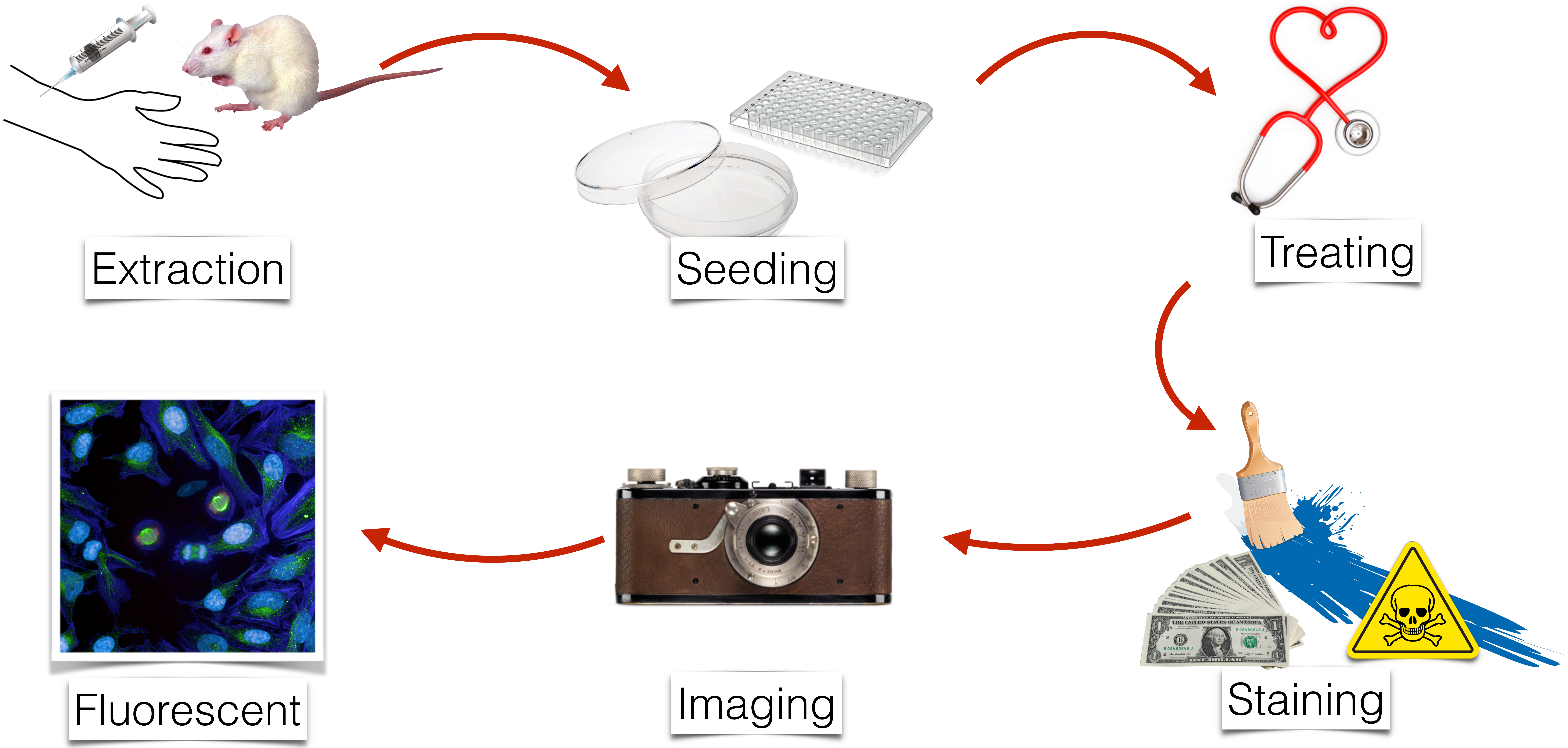




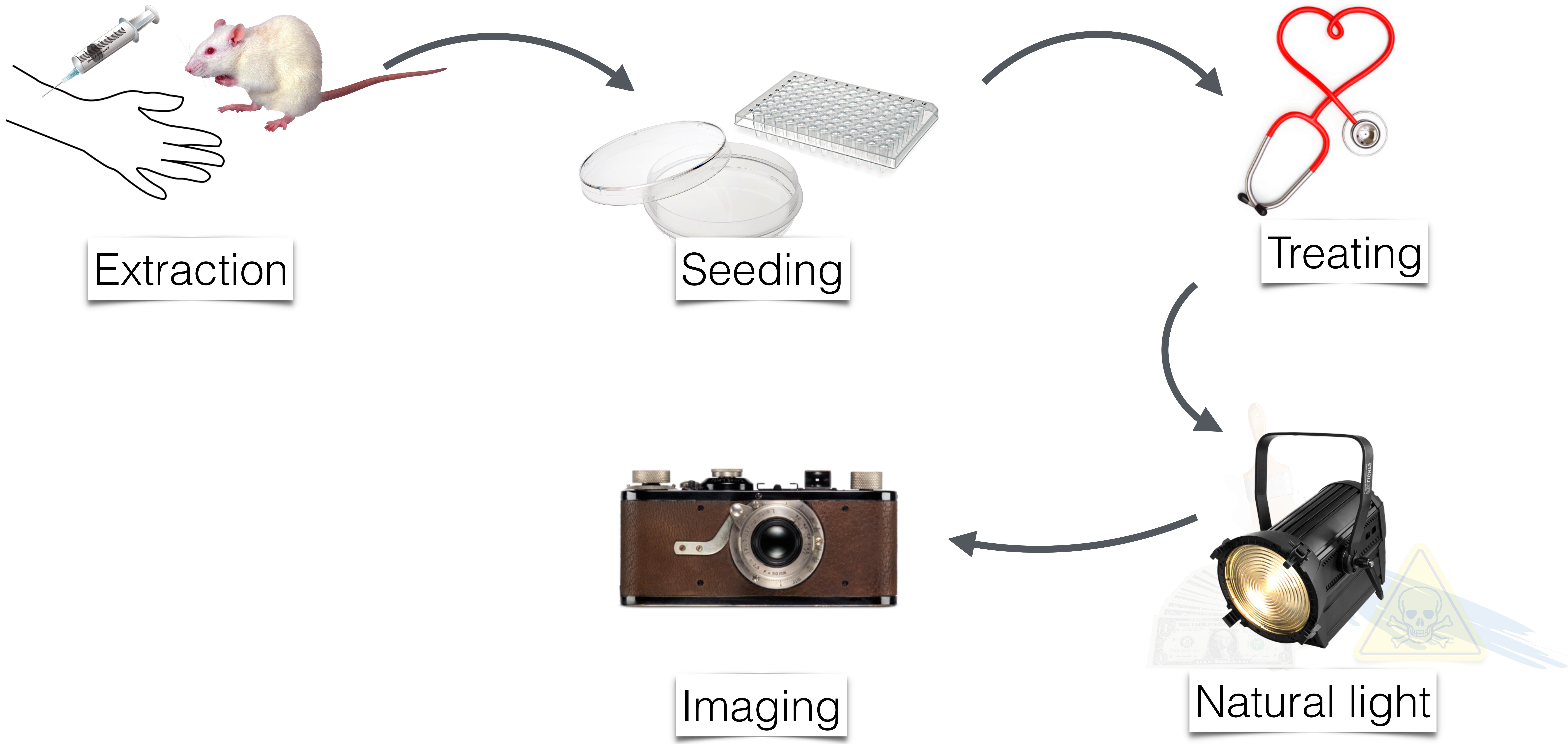
Producing Microscopy imaging



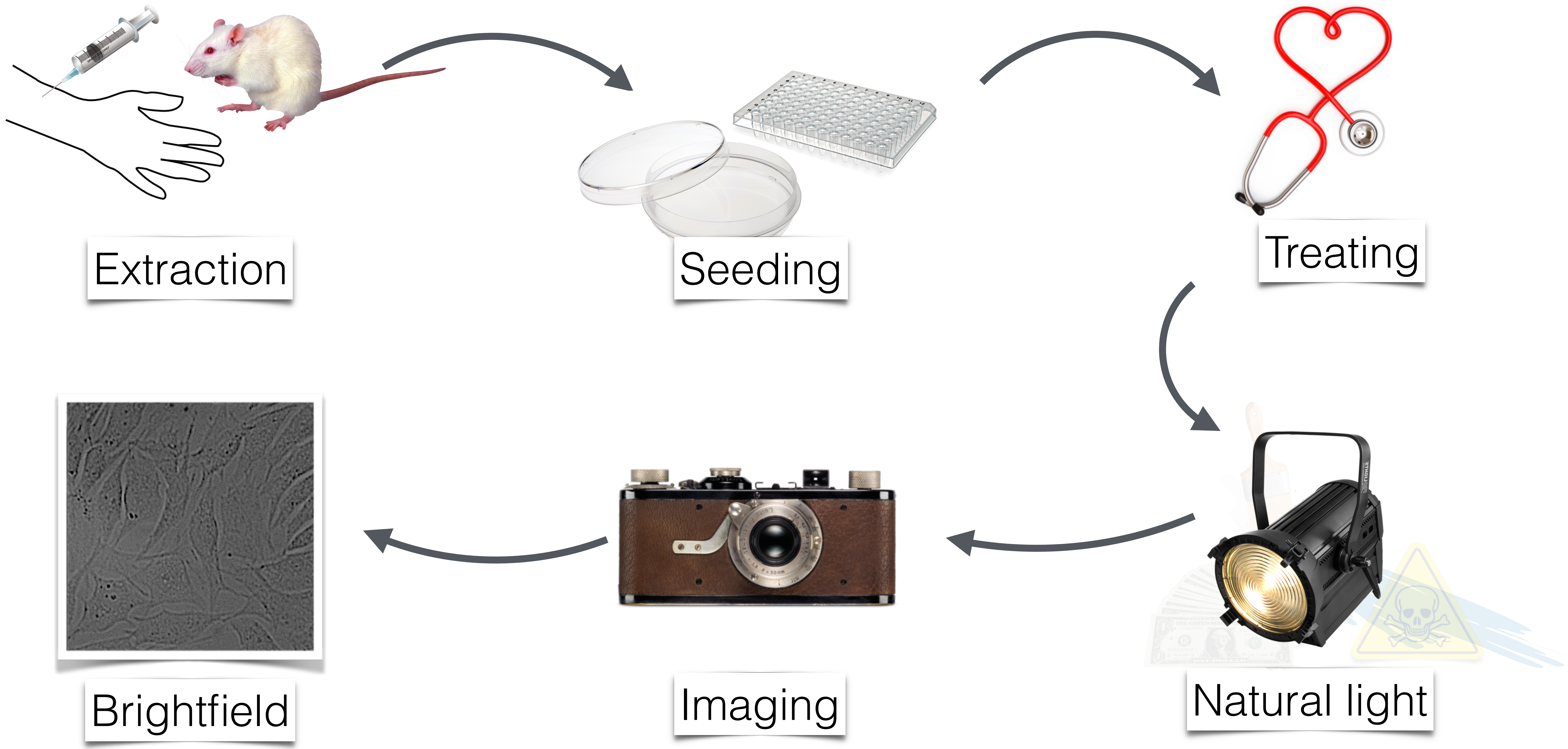
Producing Microscopy imaging

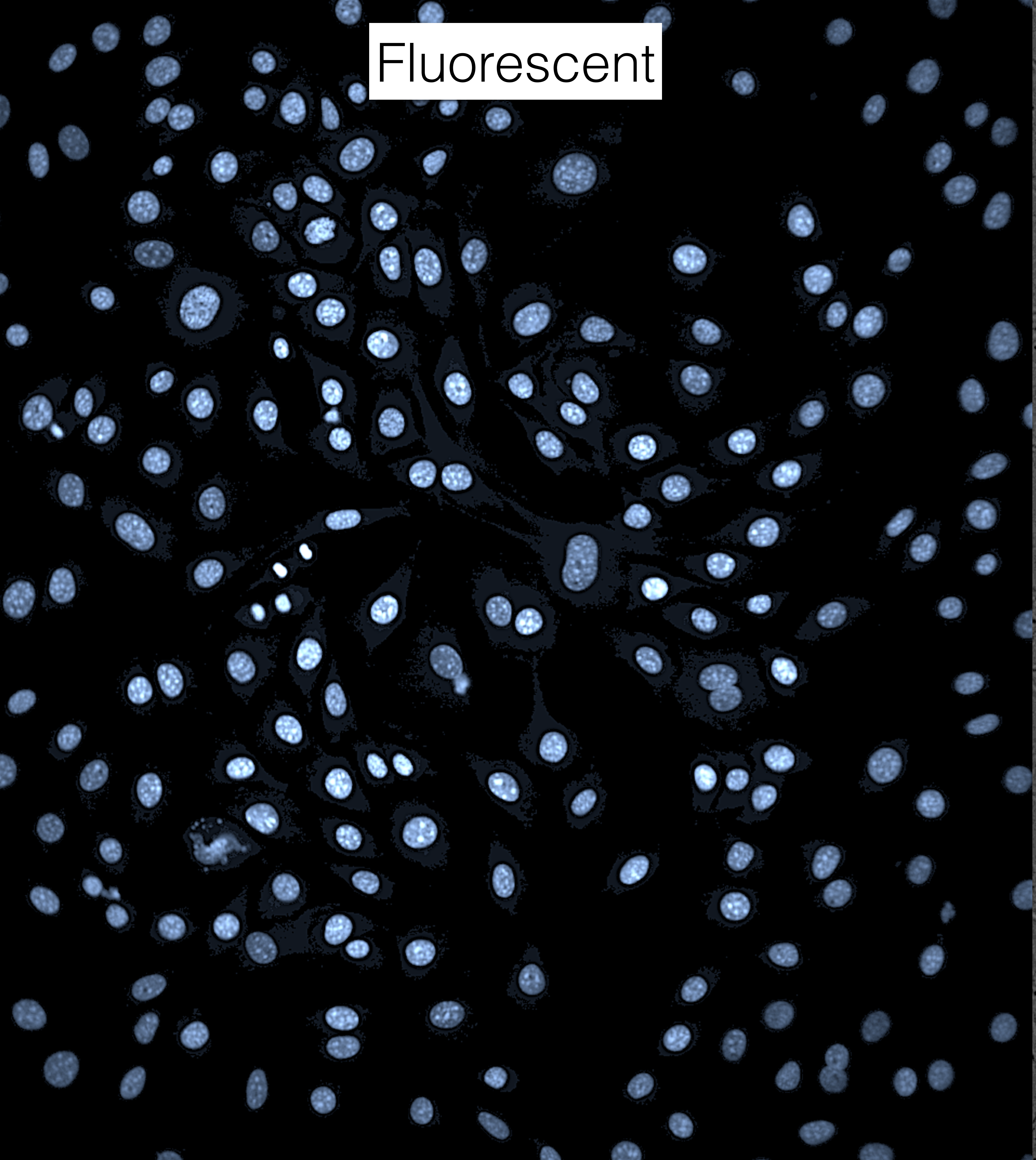
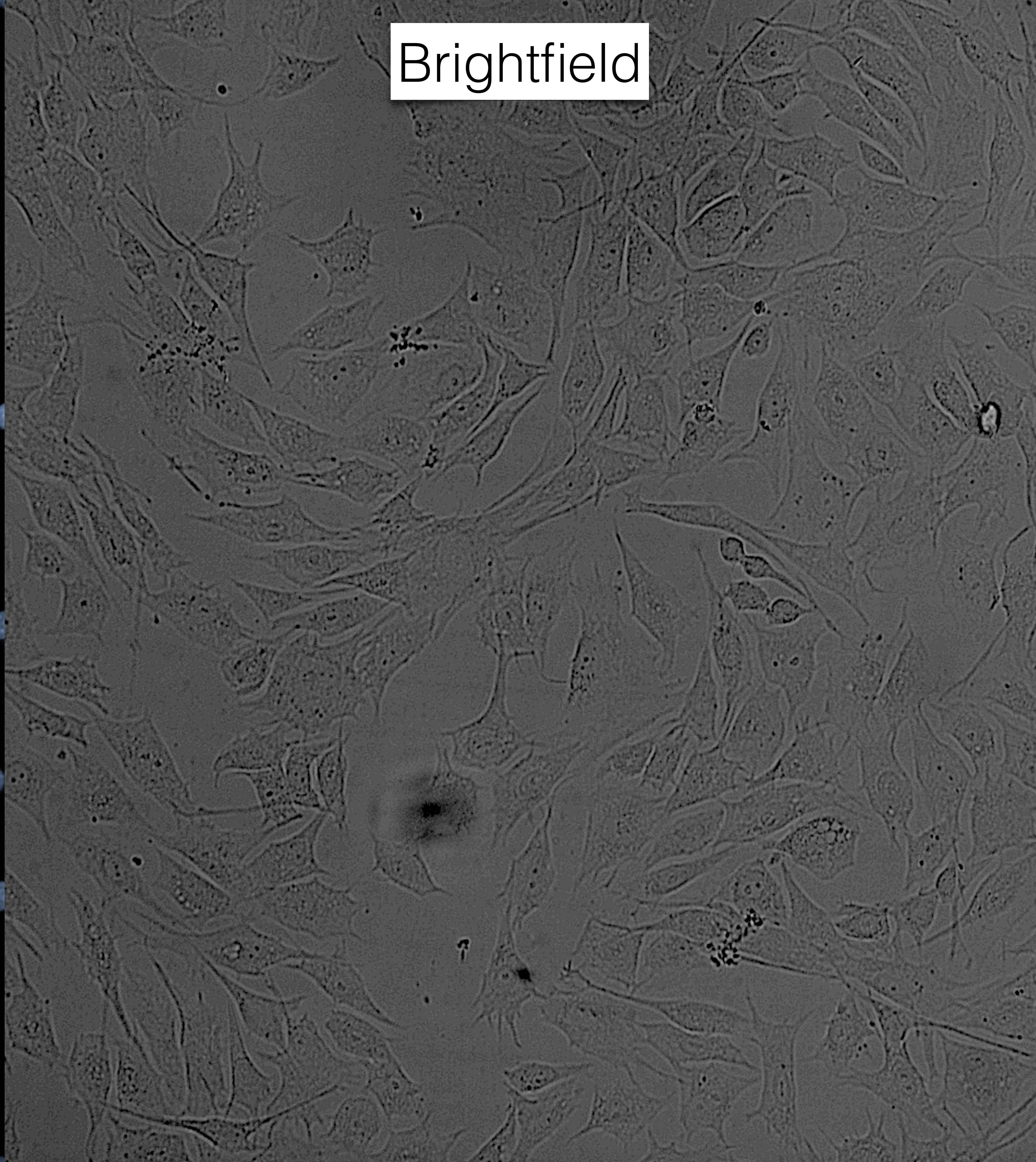


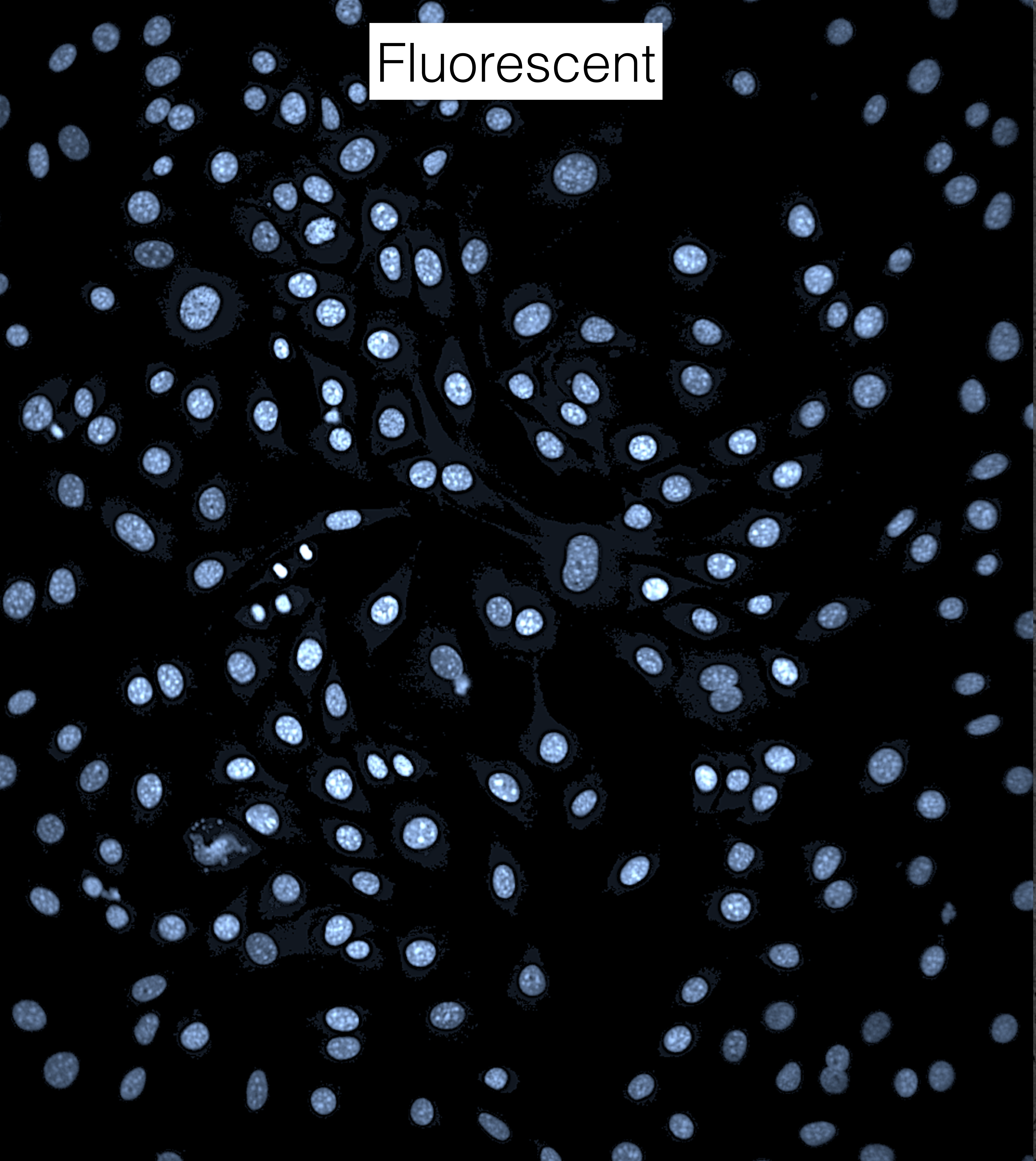
Producing Microscopy imaging

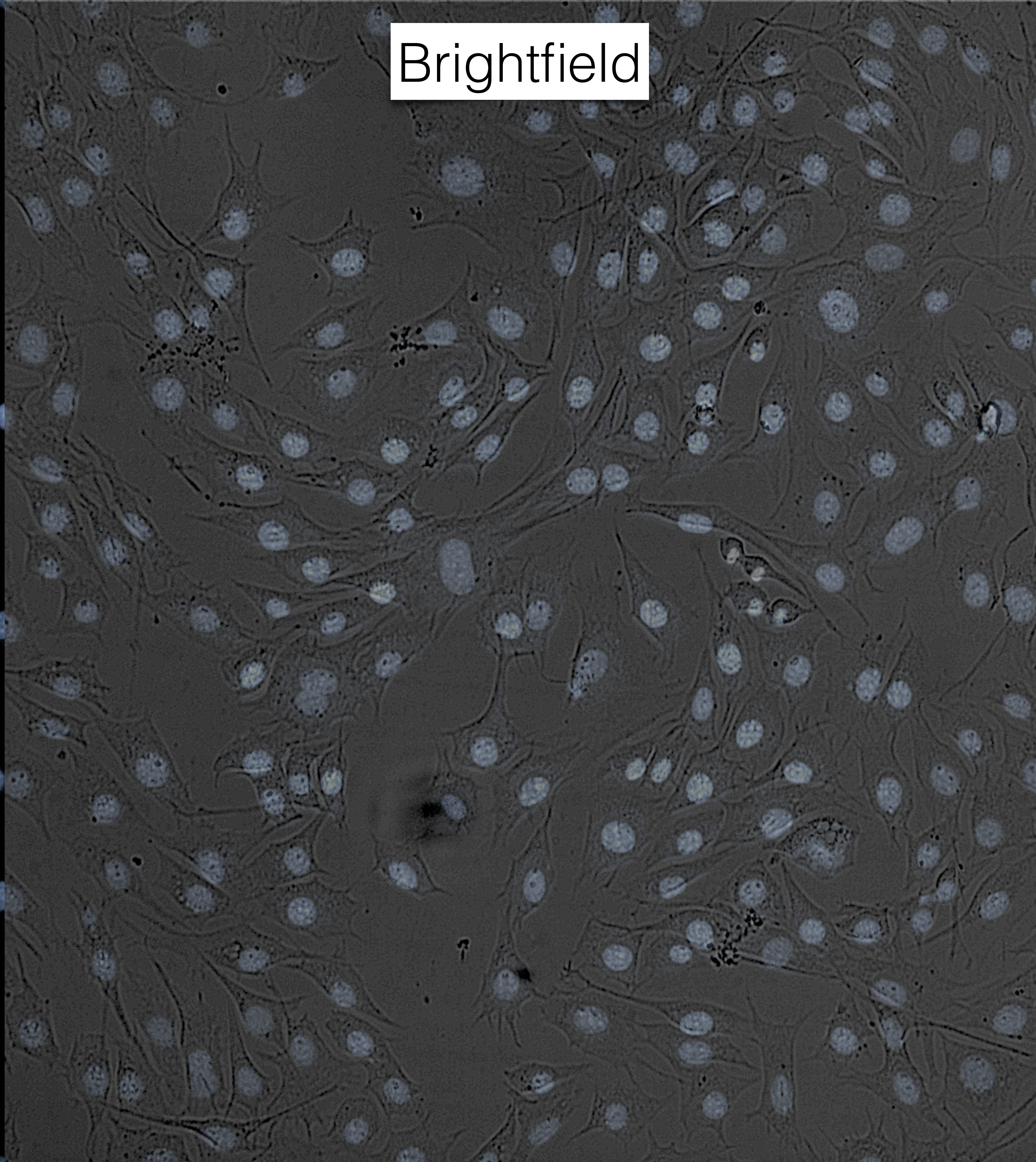


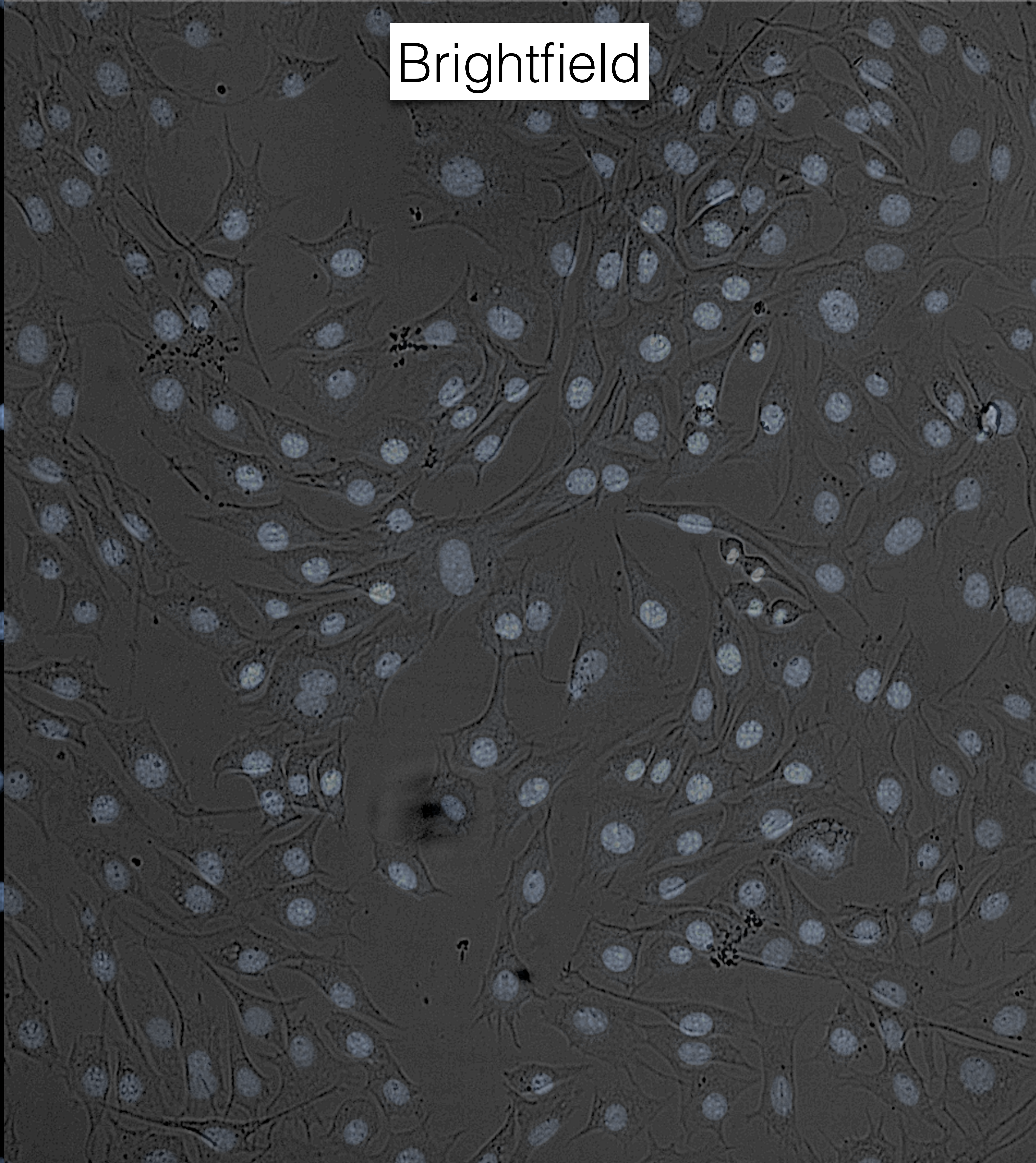
Producing Microscopy imaging




Fluorescent

Brightfield

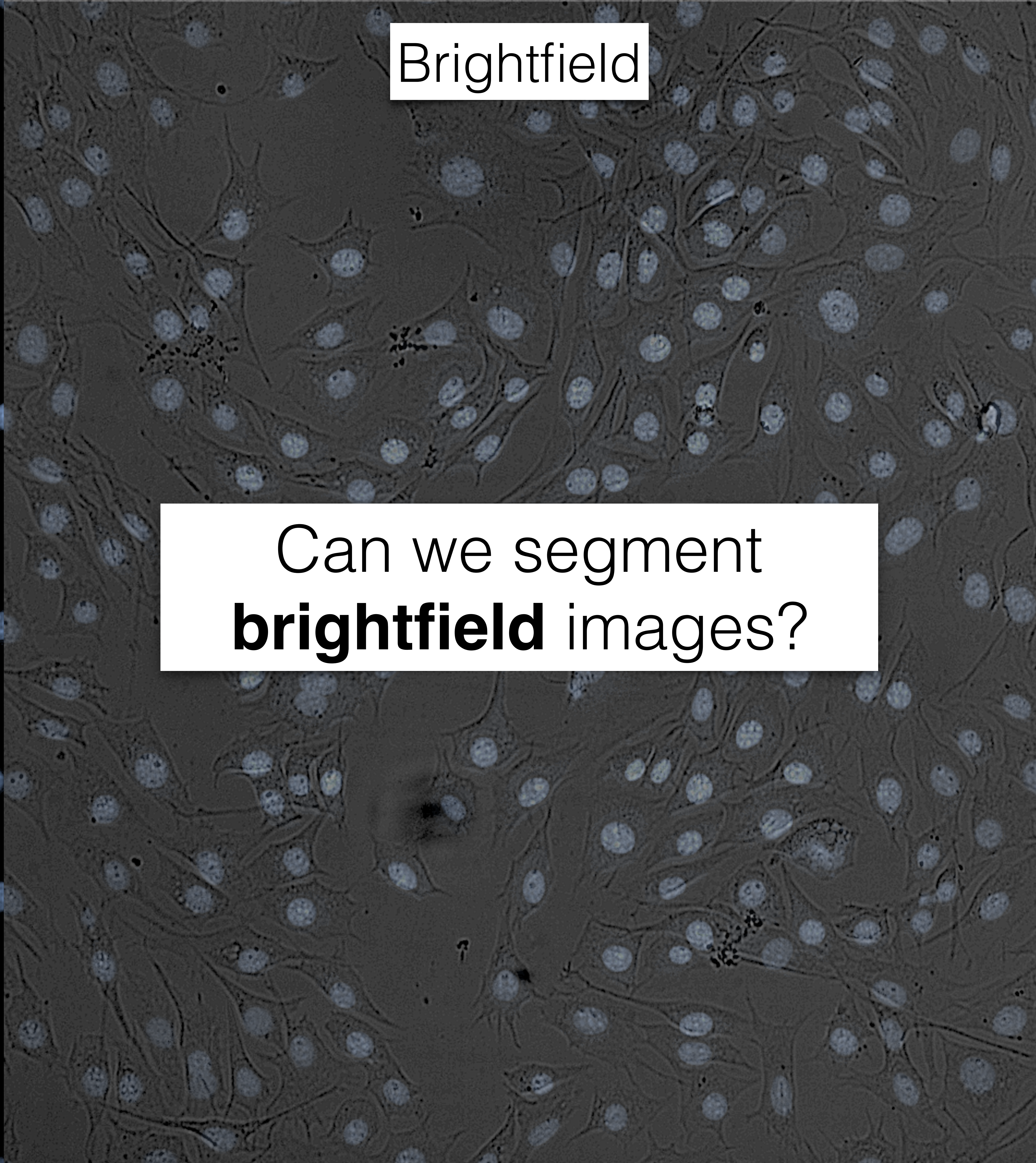

Fluorescent

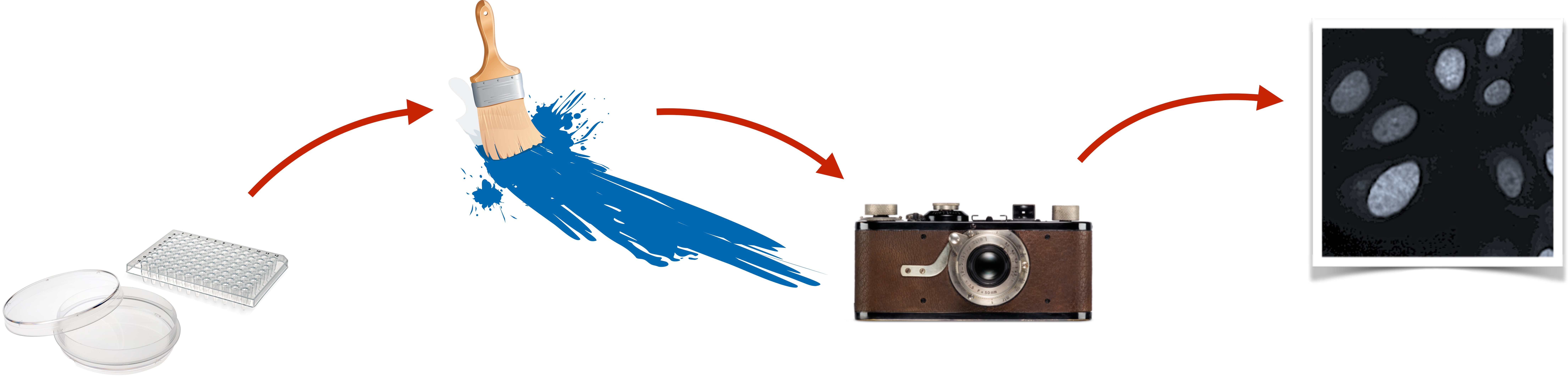

Brightfield

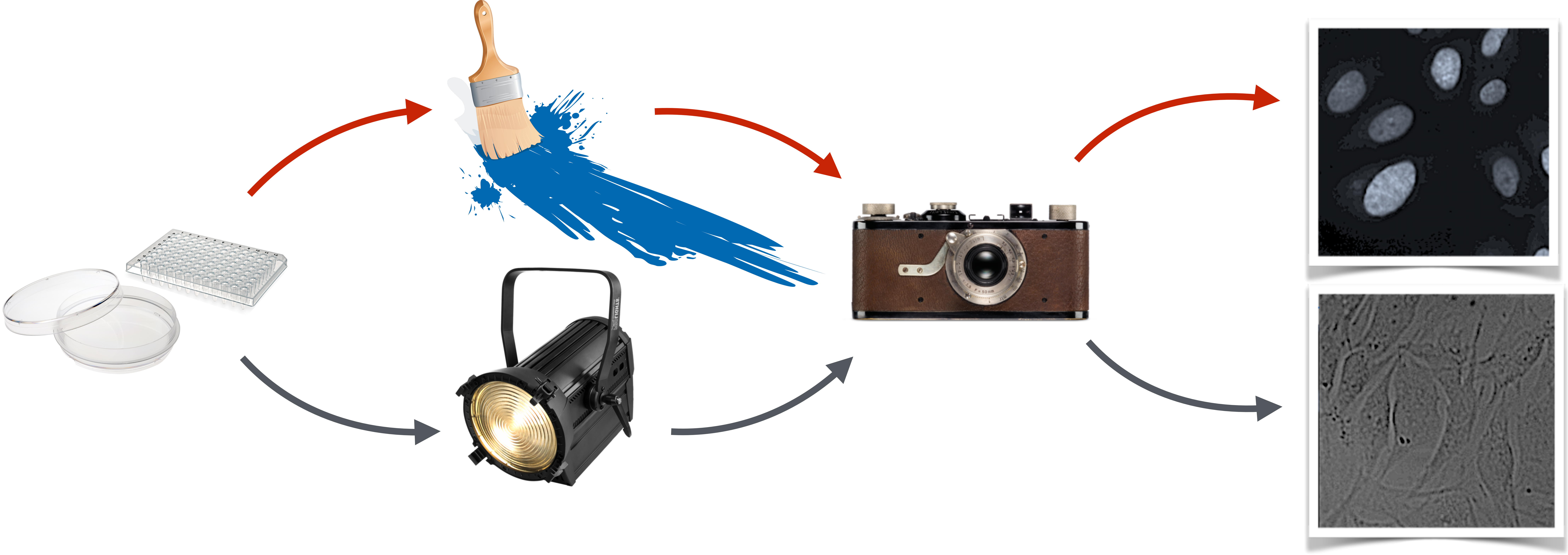


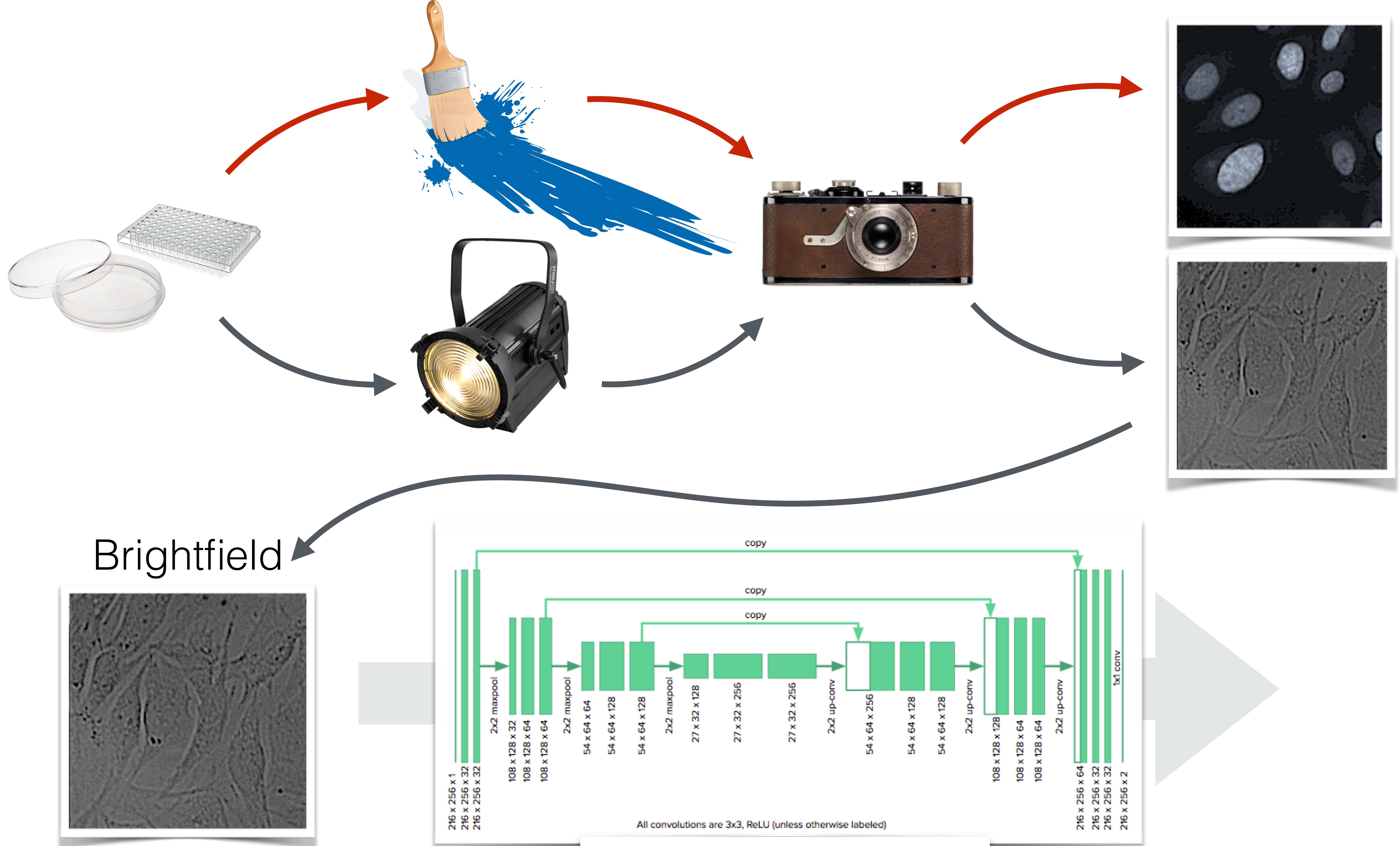
We know how to segment **fluorescent** images


Fluorescent
We know how to segment **fluorescent** images

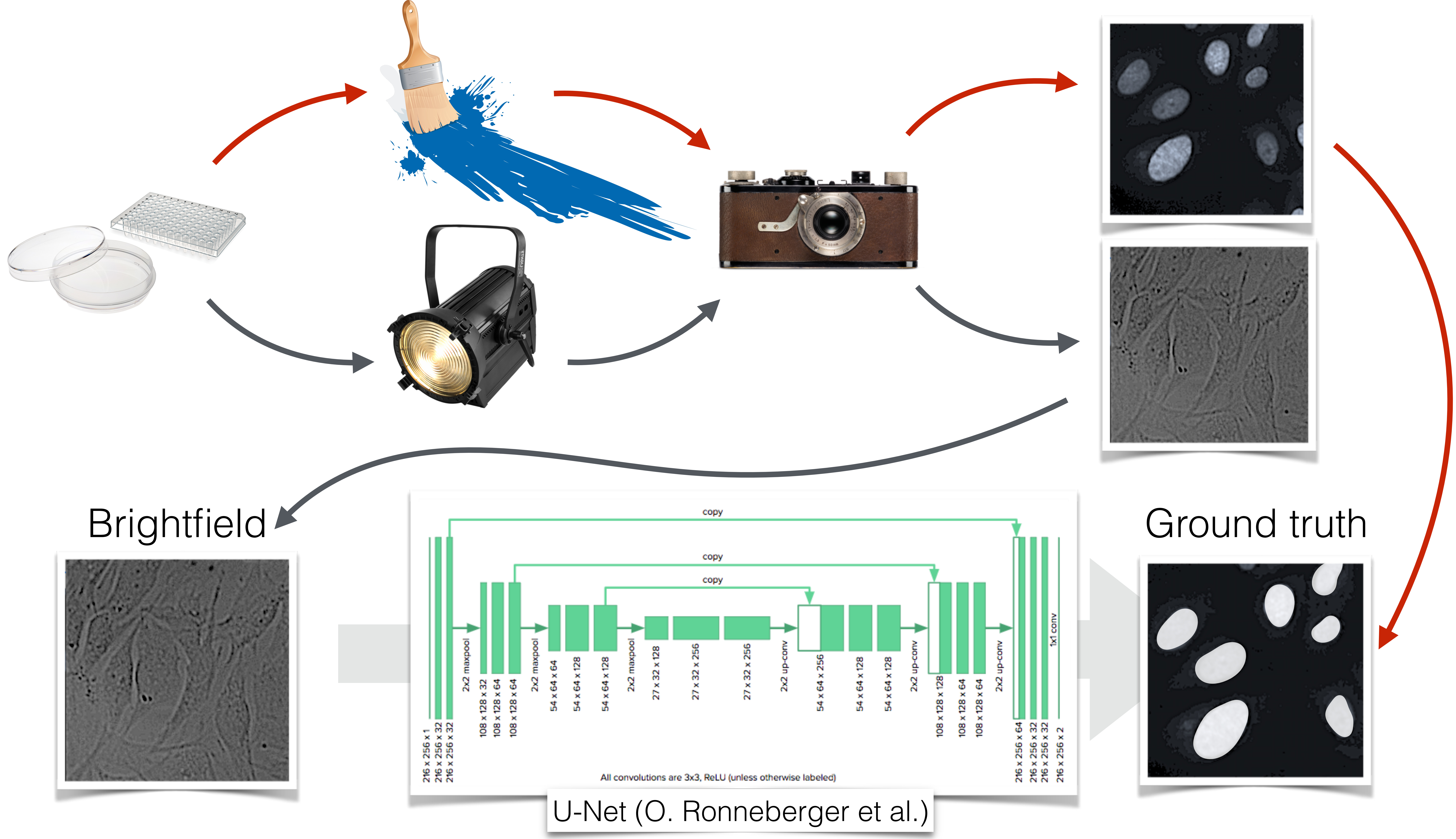

Brightfield
Can we segment **brightfield** images?



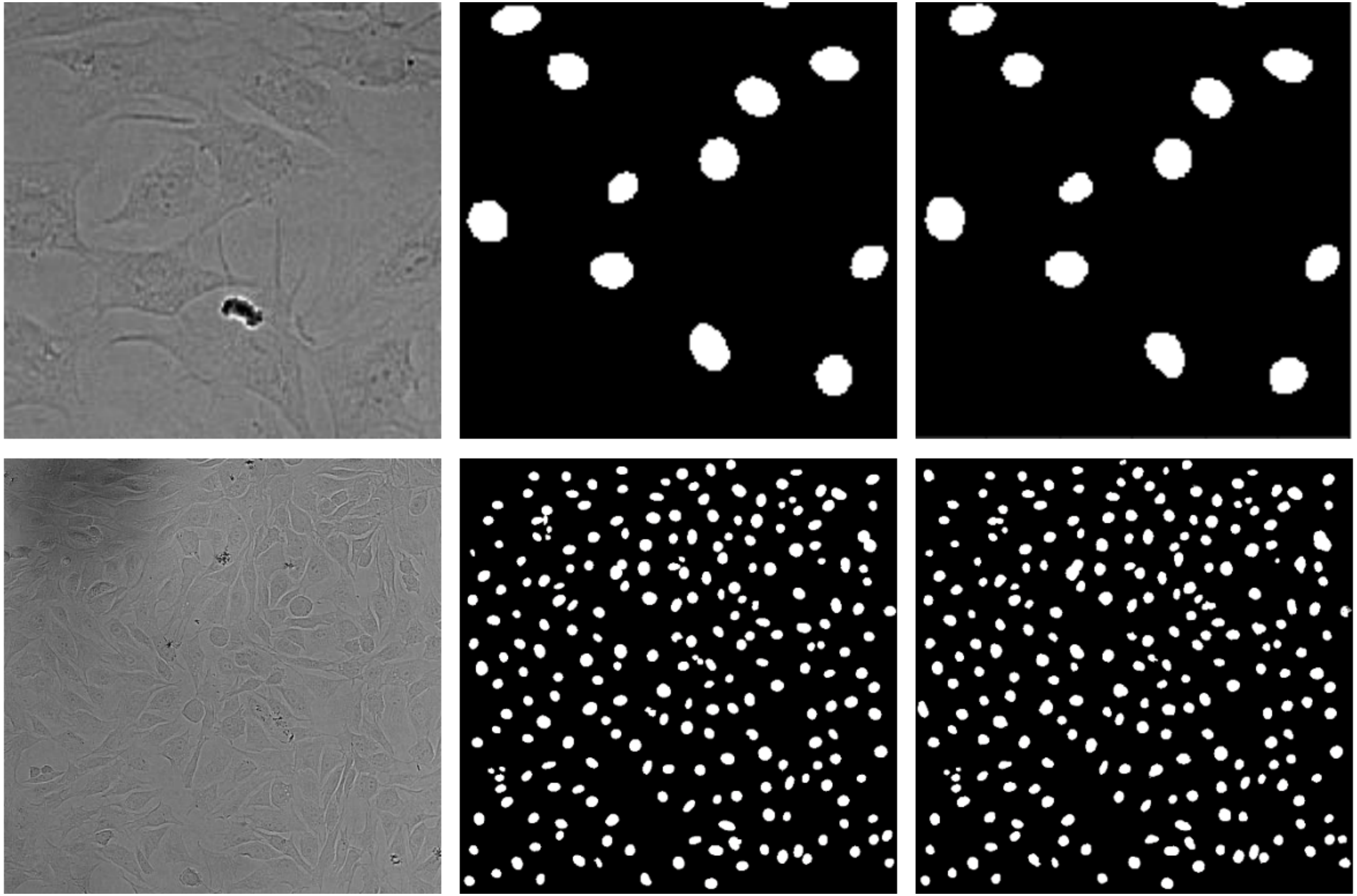




U-Net (O. Ronneberger et al.)



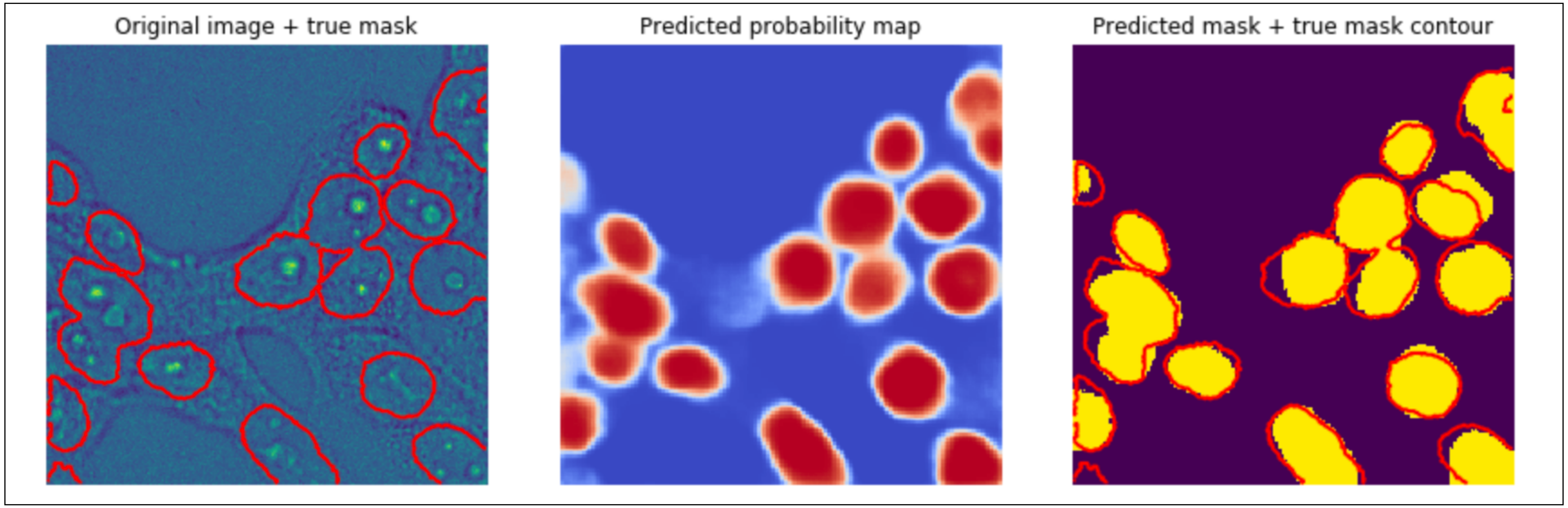
Brightfield
copy
copy
copy
216 x 256 x 1
216 x 256 x 32
216 x 256 x 32
2x2 maxpool
108 x 128 x 32
108 x 128 x 64
108 x 128 x 64
2x2 maxpool
54 x 64 x 64
54 x 64 x 128
54 x 64 x 128
2x2 maxpool
27 x 32 x 128
27 x 32 x 256
27 x 32 x 256
2x2 up-conv
54 x 64 x 256
54 x 64 x 128
54 x 64 x 128
2x2 up-conv
108 x 128 x 128
108 x 128 x 64
108 x 128 x 64
2x2 up-conv
216 x 256 x 64
216 x 256 x 32
216 x 256 x 32
1x1 conv
216 x 256 x 2
All convolutions are 3x3, ReLU (unless otherwise labeled)
U-Net (O. Ronneberger et al.)
Ground truth



Convolutional Neural Networks for Cellular Segmentation by Sten-Oliver Salumaa

Segmenting cell nuclei from brightfield images with deep convolutional neural networks

Sten-Oliver Salumaa[1,*], Dmytro Fishman[1,*], Daniel Majoral[1]?, …?, Kaupo Palo[2], Jaak Vilo[1], Leopold Parts[1,3,@]



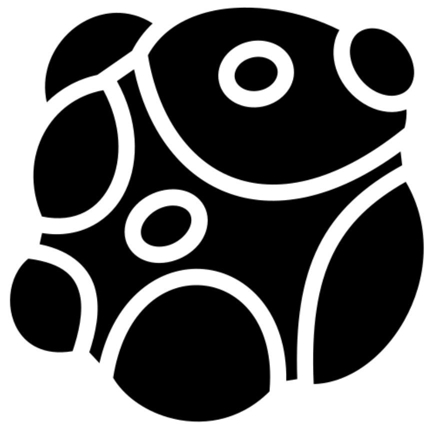
Will be added to the PerkinElmer **Harmony**


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Can be improved by Deep Learning
Objects detection
Multi-instance mask
Extracting features
Relevant features
Classification
Phenotyping
Cancer
Healthy cells
nuclei #1: blue, size 29px;
nuclei #2: red, size 25px;
nuclei #3: pink, size 22px;
nuclei #4: yellow, size 19px;
nuclei #5: green, size 18px;
nuclei #6: purple, size 16px;
nuclei #7: orange, size 14px;


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Can be improved
by Deep Learning
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
Classification
Phenotyping
Cancer
Healthy cells
nuclei #1: blue, size 29px;
nuclei #2: red, size 25px;
nuclei #3: pink, size 22px;
nuclei #4: yellow, size 19px;
nuclei #5: green, size 18px;
nuclei #6: purple, size 16px;
nuclei #7: orange, size 14px;

Cells can be of different types

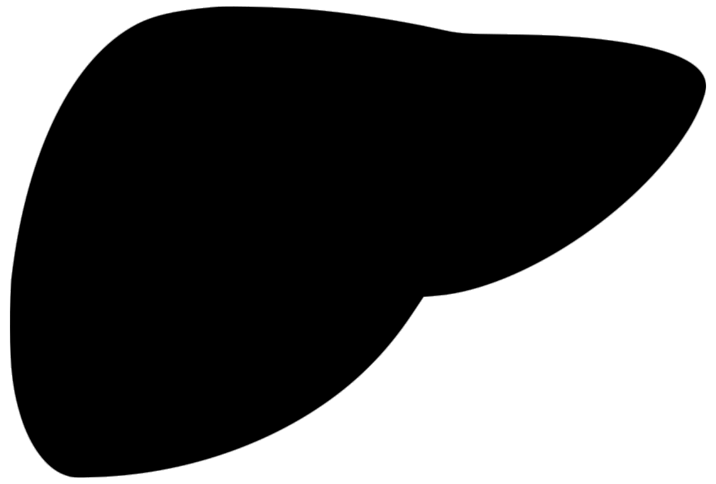
Cells can be of different types



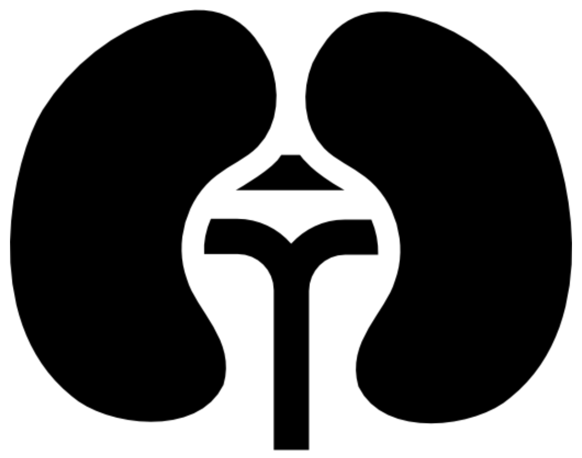
Fibrosarcoma (HT1080)

Breast cancer cell line (MCF7)

Mouse embryo tissue (NIH3T3)

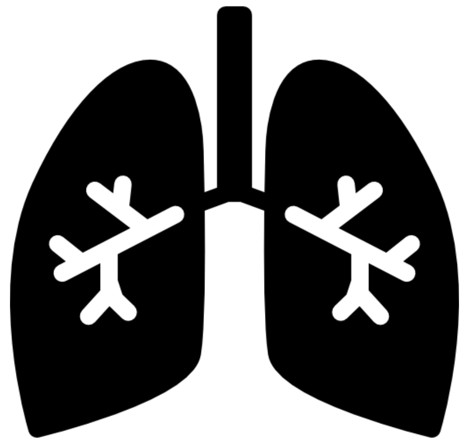


Liver cancer cell line (HepG2)



Madin-Darby Canine Kidney (MDCK) cells

Henrietta Lacks (HeLa) cells



Alveolar basal epithelial (A549) cells

Cells can be of different types

Fibrosarcoma (HT1080)

Breast cancer cell line (MCF7)

Mouse embryo tissue (NIH3T3)

Liver cancer cell line (HepG2)

Madin-Darby Canine Kidney (MDCK) cells

Alveolar basal epithelial (A549) cells

Henrietta Lacks (HeLa) cells

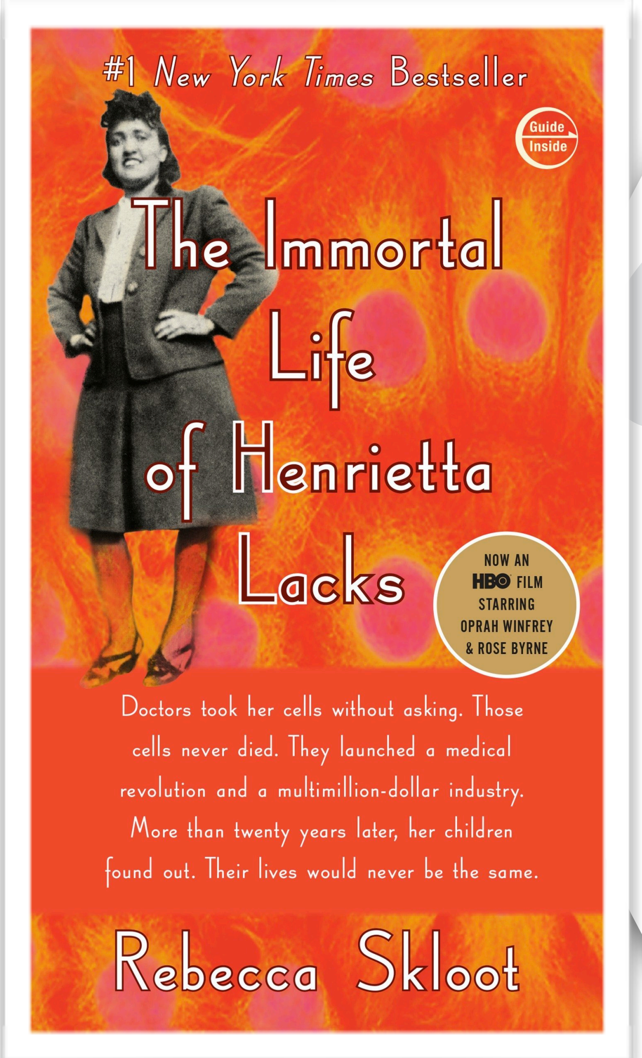
Cells can be of different types

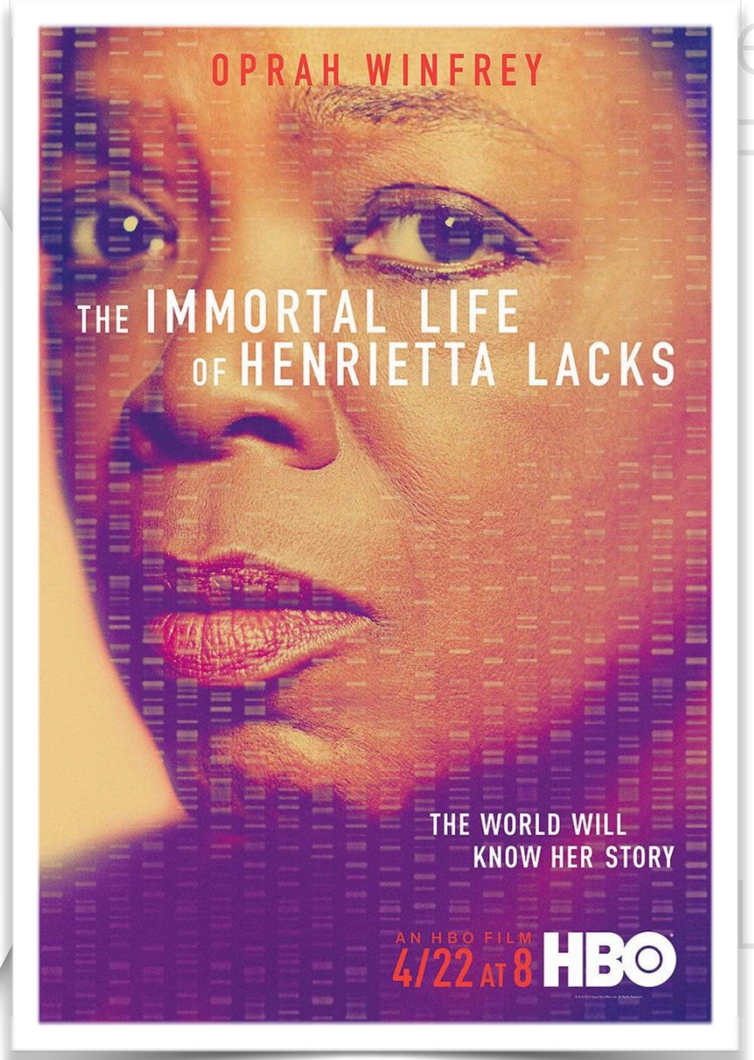
Fibrosarcom

er cell line

F7)

Mouse embryo tissue (NIH3T3)



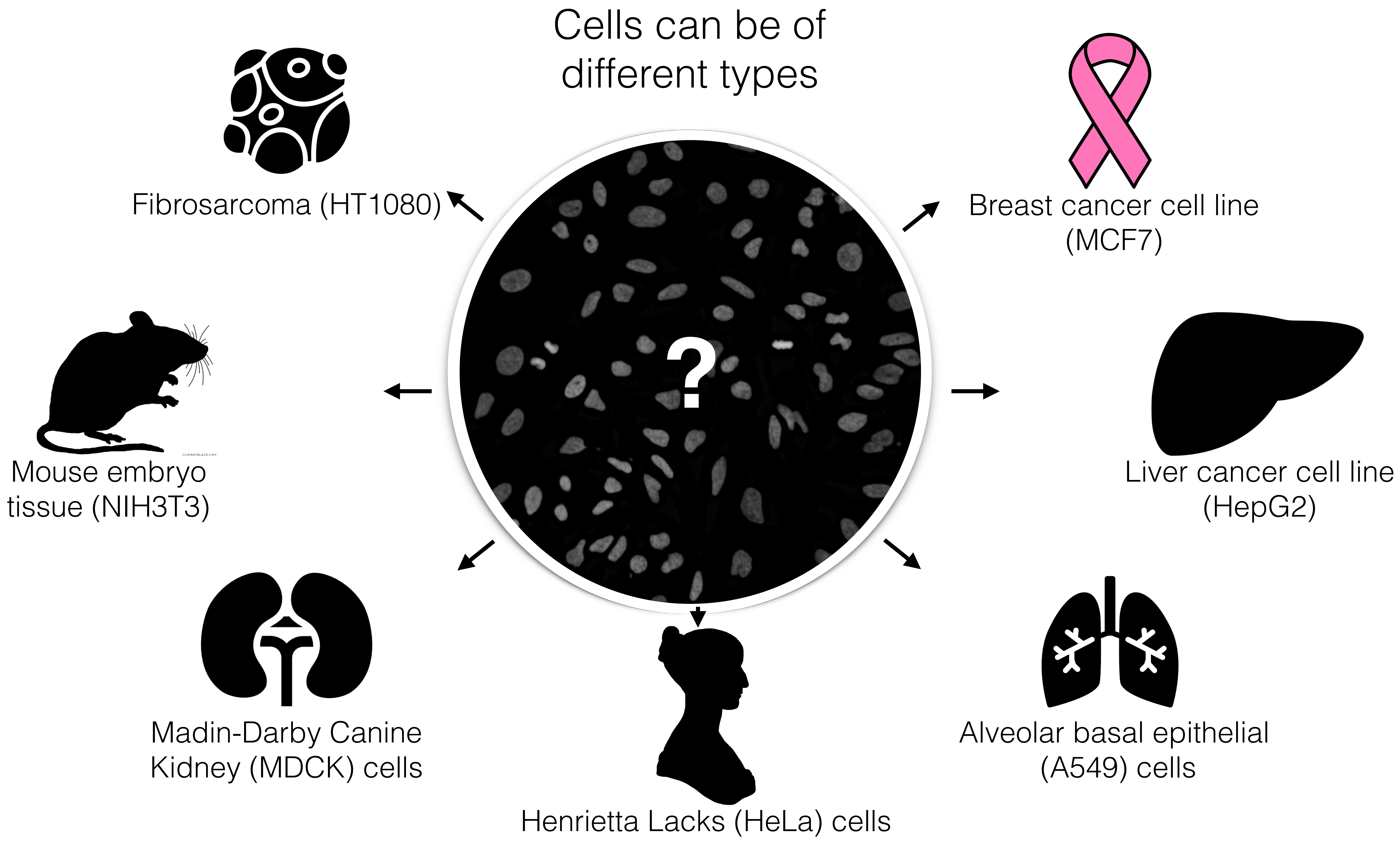


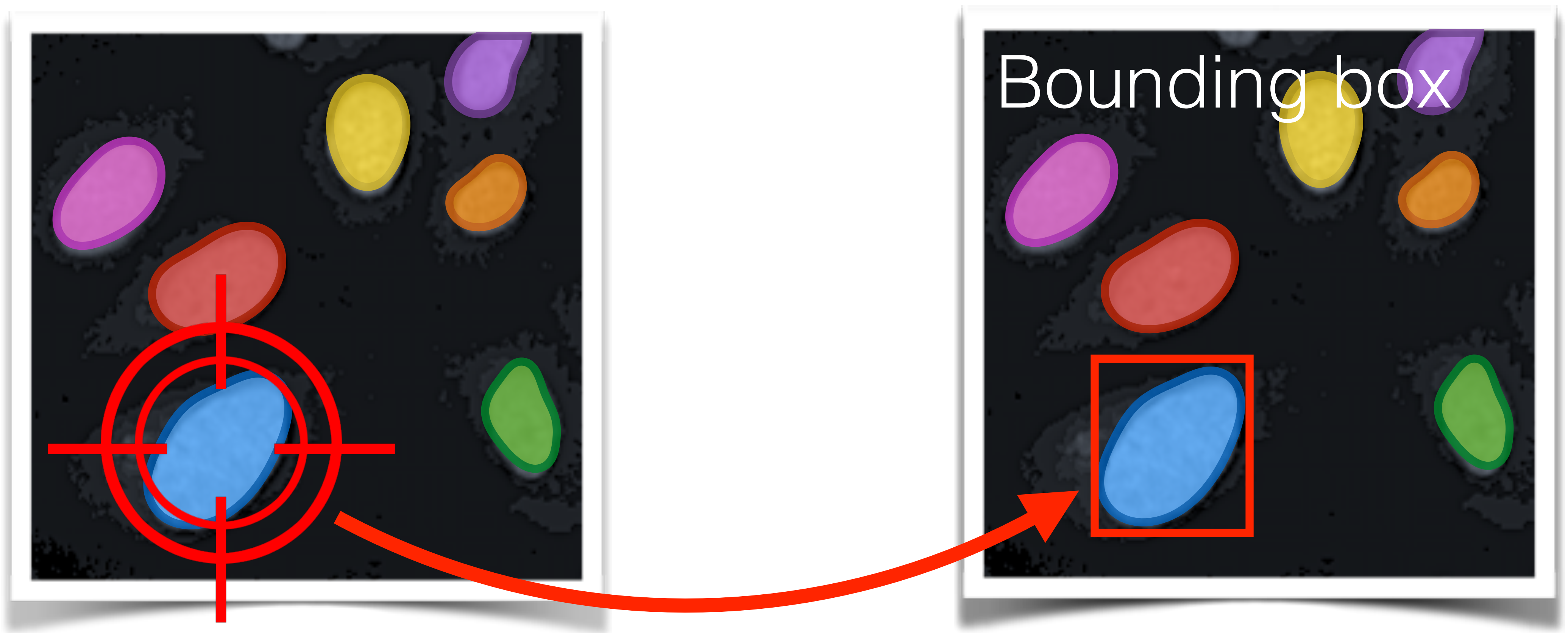
Liver cancer cell line (HepG2)

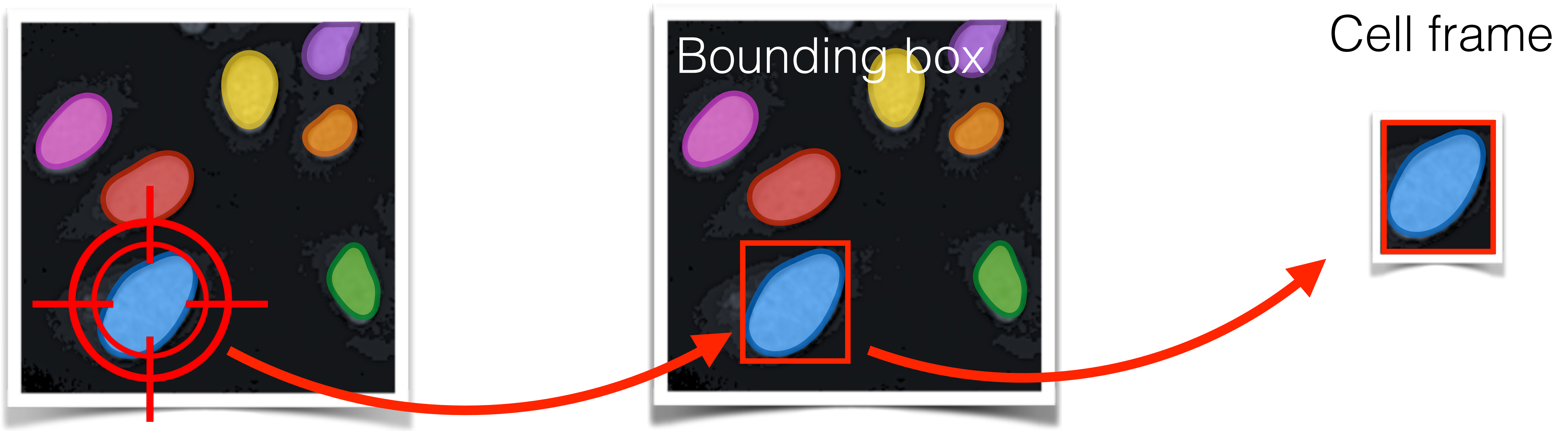
Madin-Darby Canine Kidney (MDCK) cells

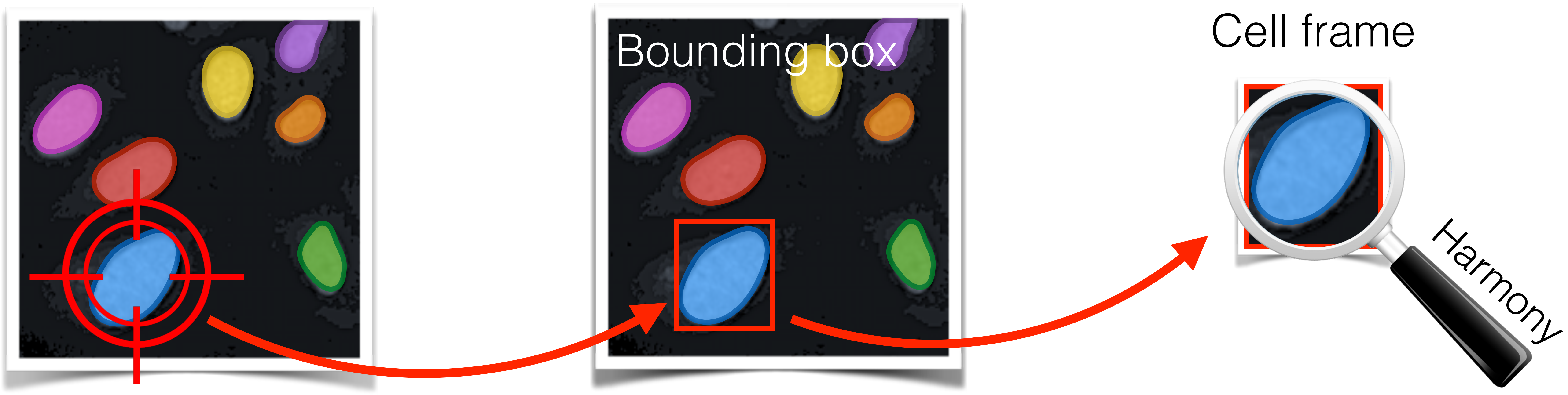
Alveolar basal epithelial (A549) cells

Henrietta Lacks (HeLa) cells


Cells can be of different types
?
Fibrosarcoma (HT1080)
Breast cancer cell line (MCF7)
Mouse embryo tissue (NIH3T3)
Liver cancer cell line (HepG2)
Madin-Darby Canine Kidney (MDCK) cells
Henrietta Lacks (HeLa) cells
Alveolar basal epithelial (A549) cells

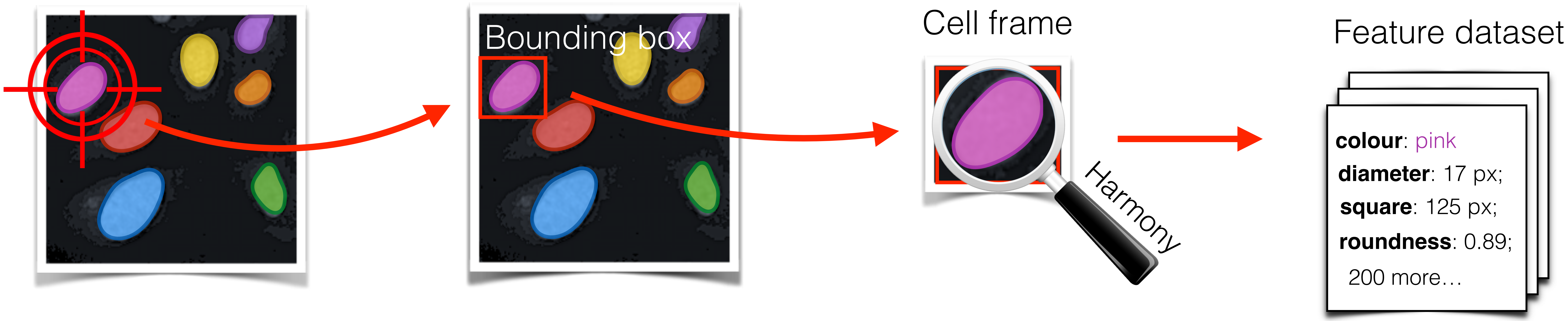

Bounding box

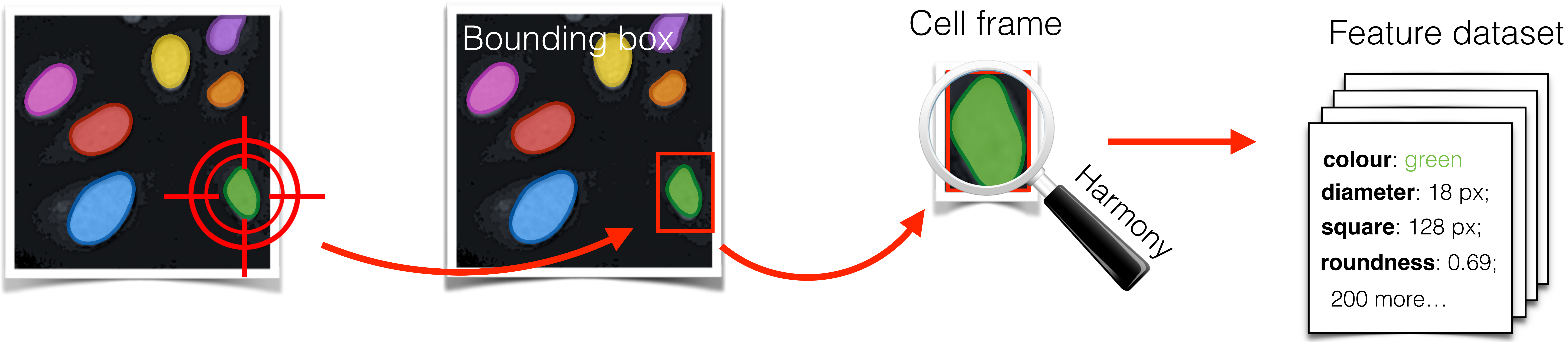

Bounding box
Cell frame


Bounding box
Cell frame
Harmony

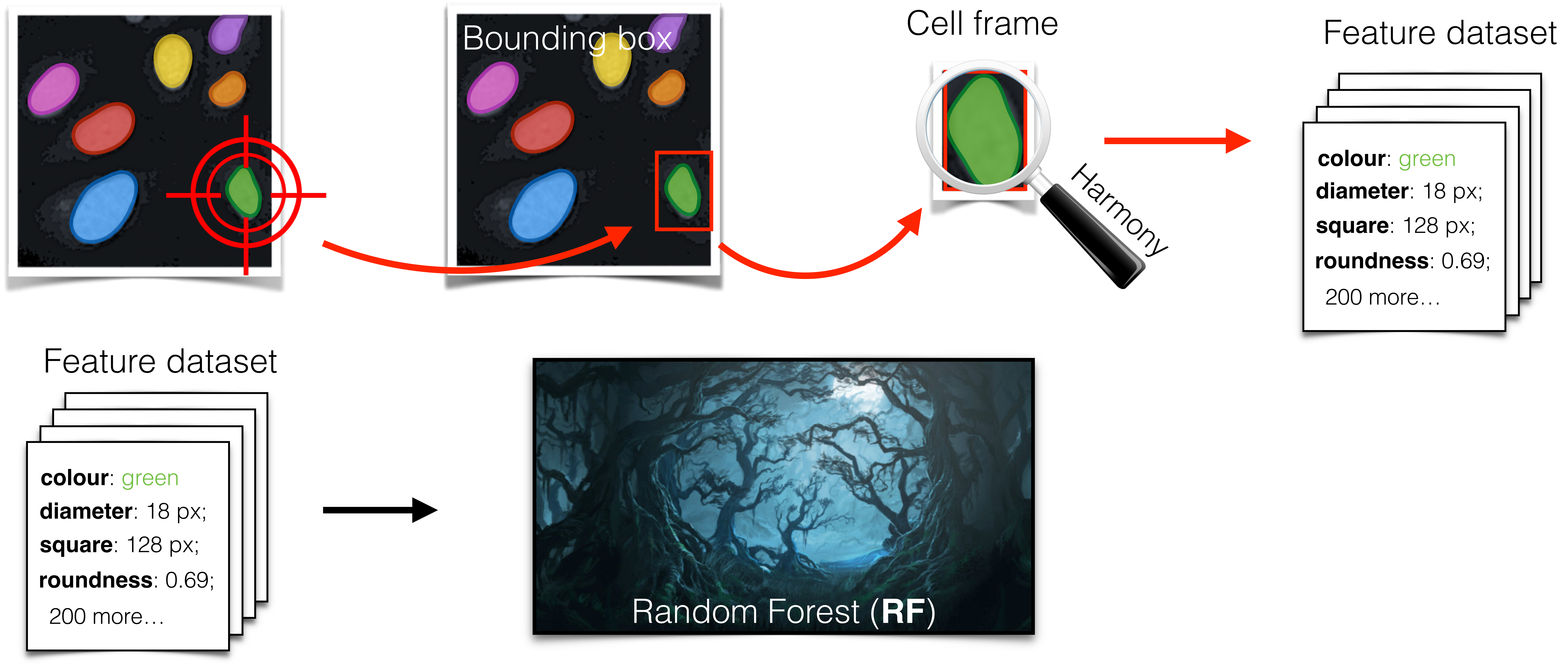

Bounding box
Cell frame
Harmony
Feature dataset
colour: blue
diameter: 25px;
square: 200 px;
roundness: 0.9;
200 more…

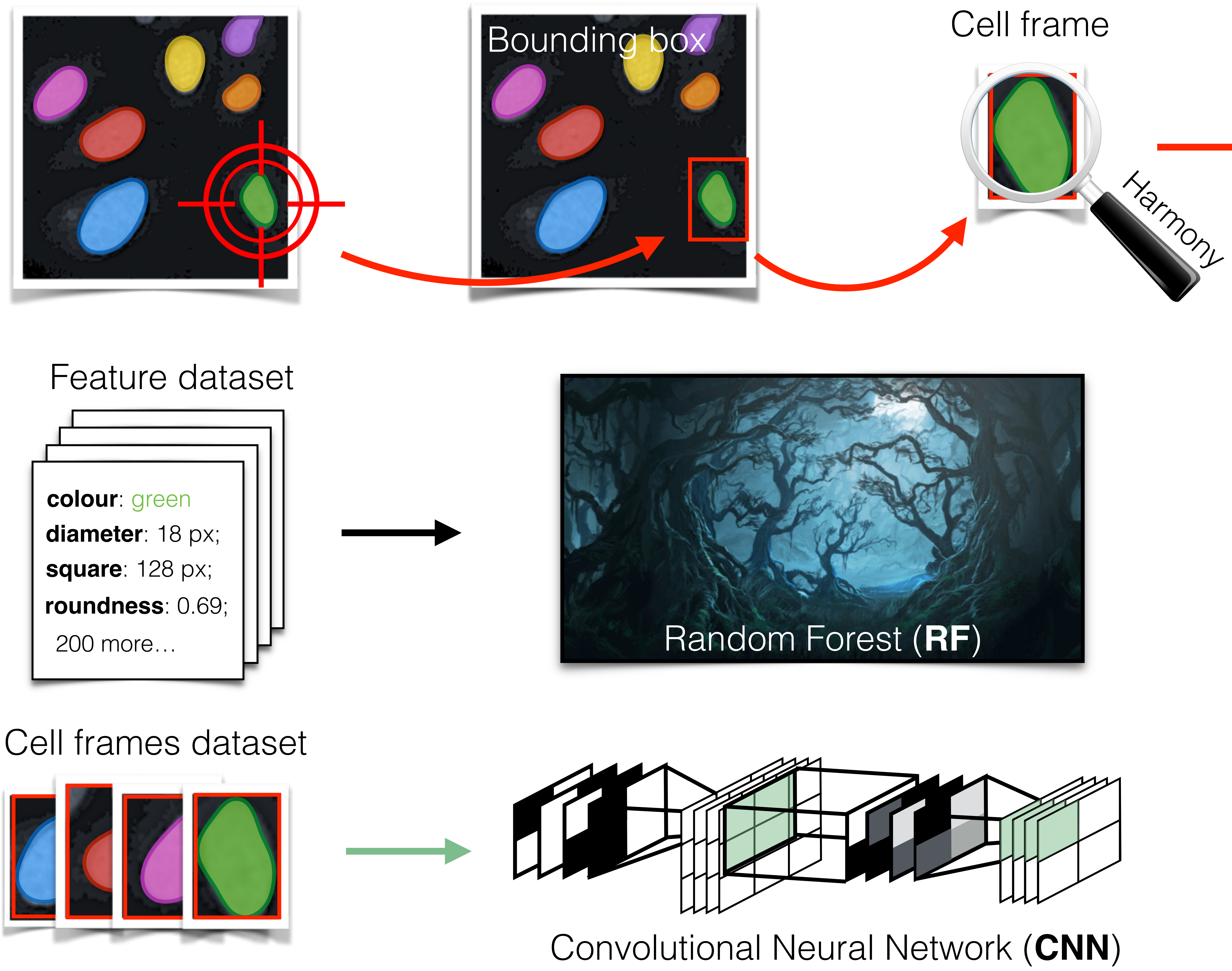

Bounding box
Cell frame
Harmony
Feature dataset
colour: red
diameter: 19 px;
square: 165 px;
roundness: 0.87;
200 more…

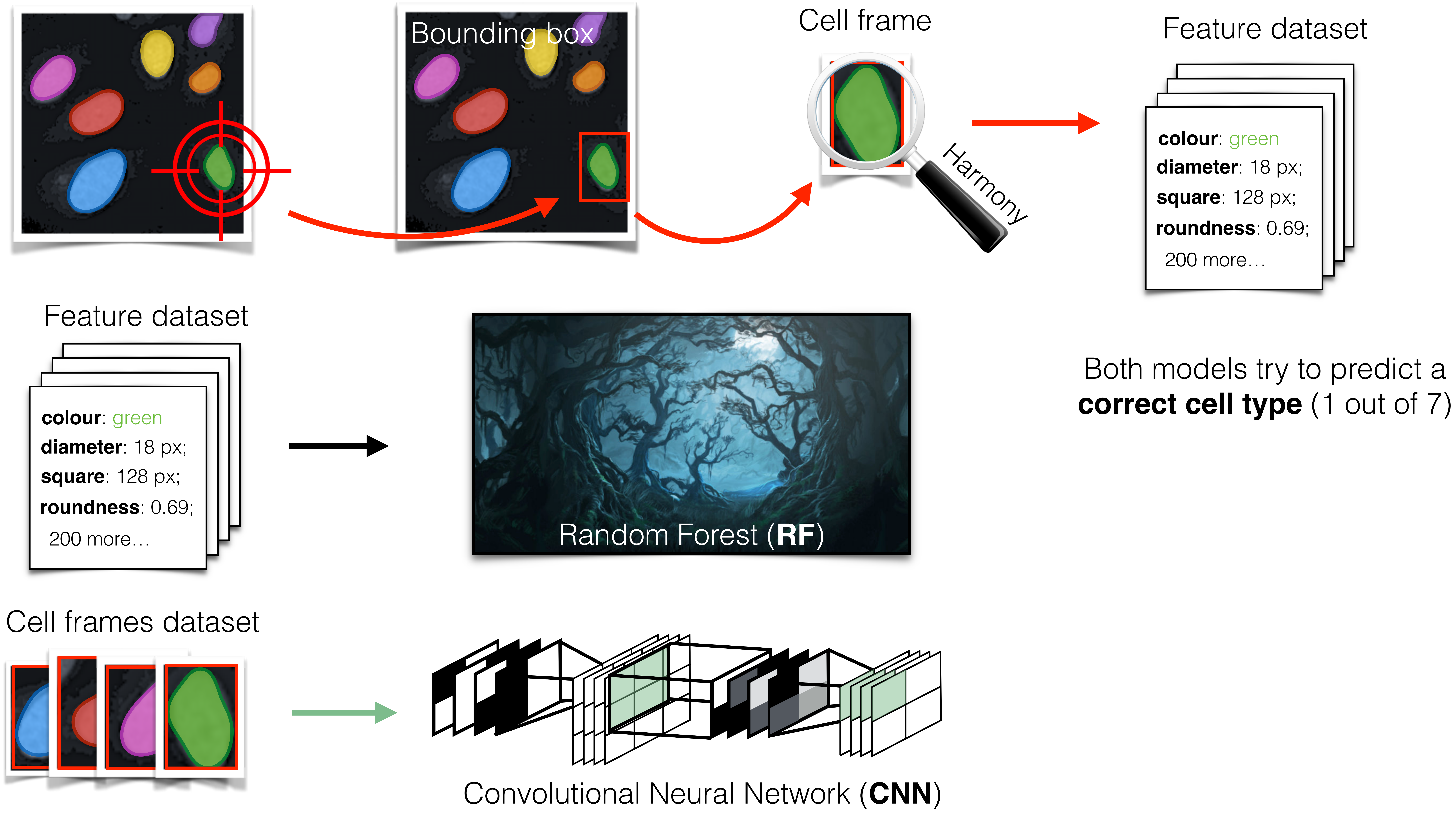

Bounding box
Cell frame
Harmony
Feature dataset
colour: pink
diameter: 17 px;
square: 125 px;
roundness: 0.89;
200 more…

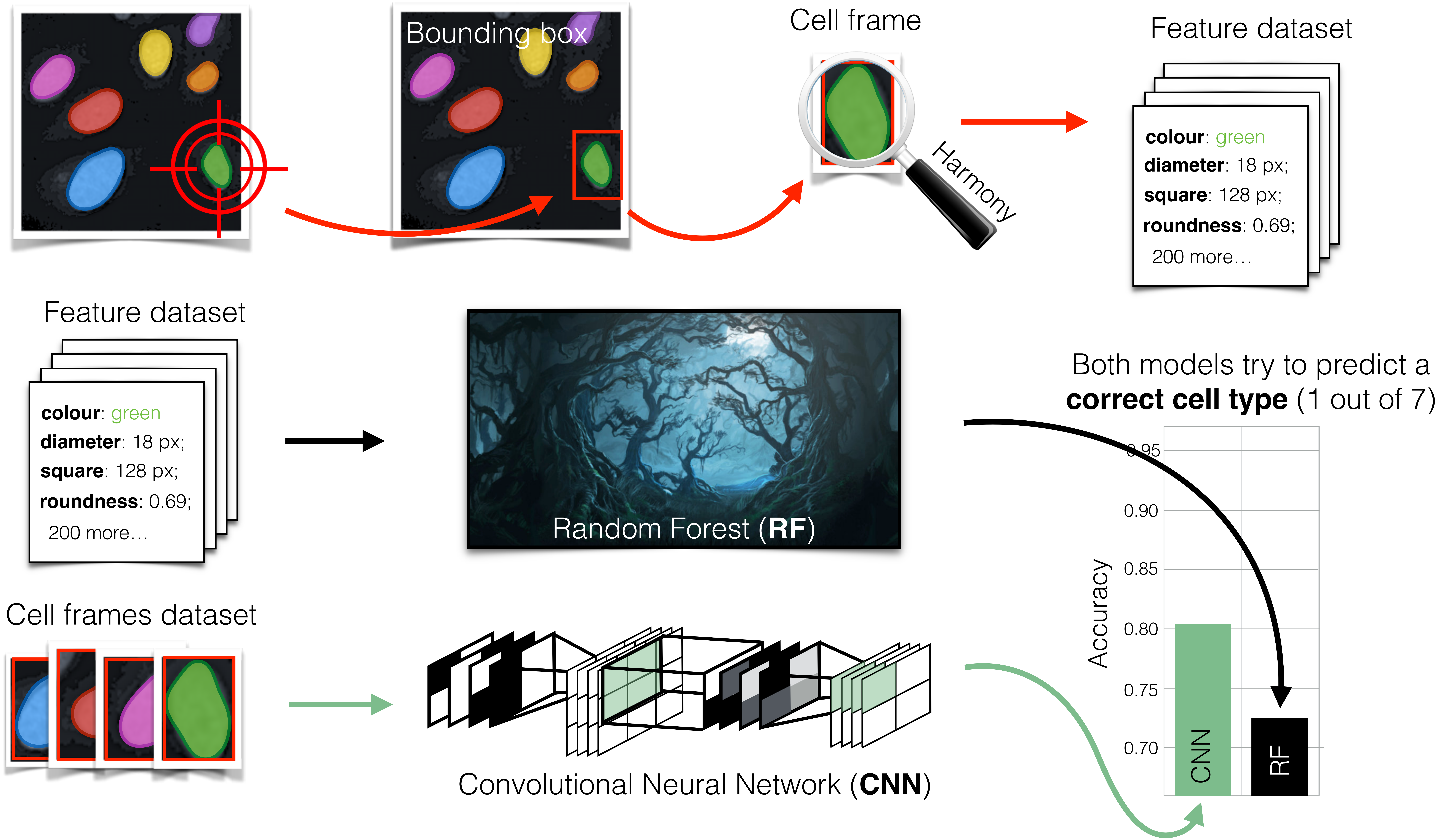


Bounding box
Cell frame
Harmony
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…

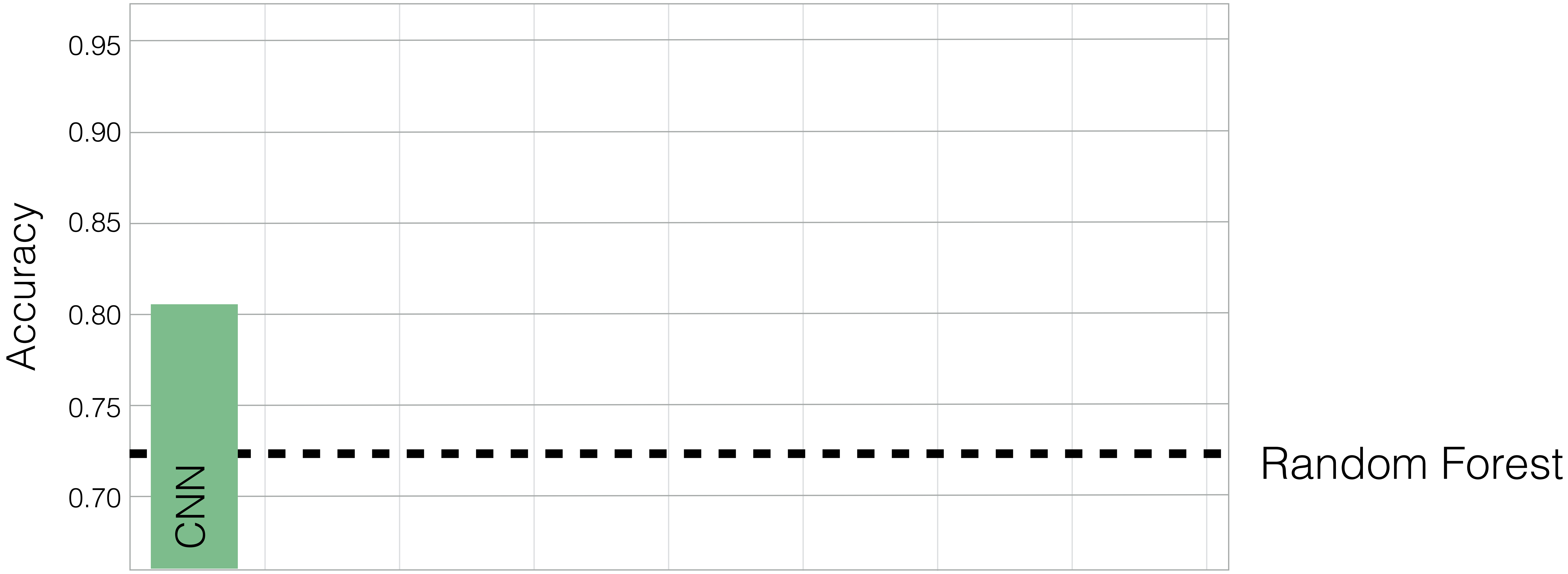

Bounding box
Cell frame
Harmony
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…
Random Forest (RF)


Bounding box
Cell frame
Harmony
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…
Random Forest (RF)
Cell frames dataset
Convolutional Neural Network (CNN)

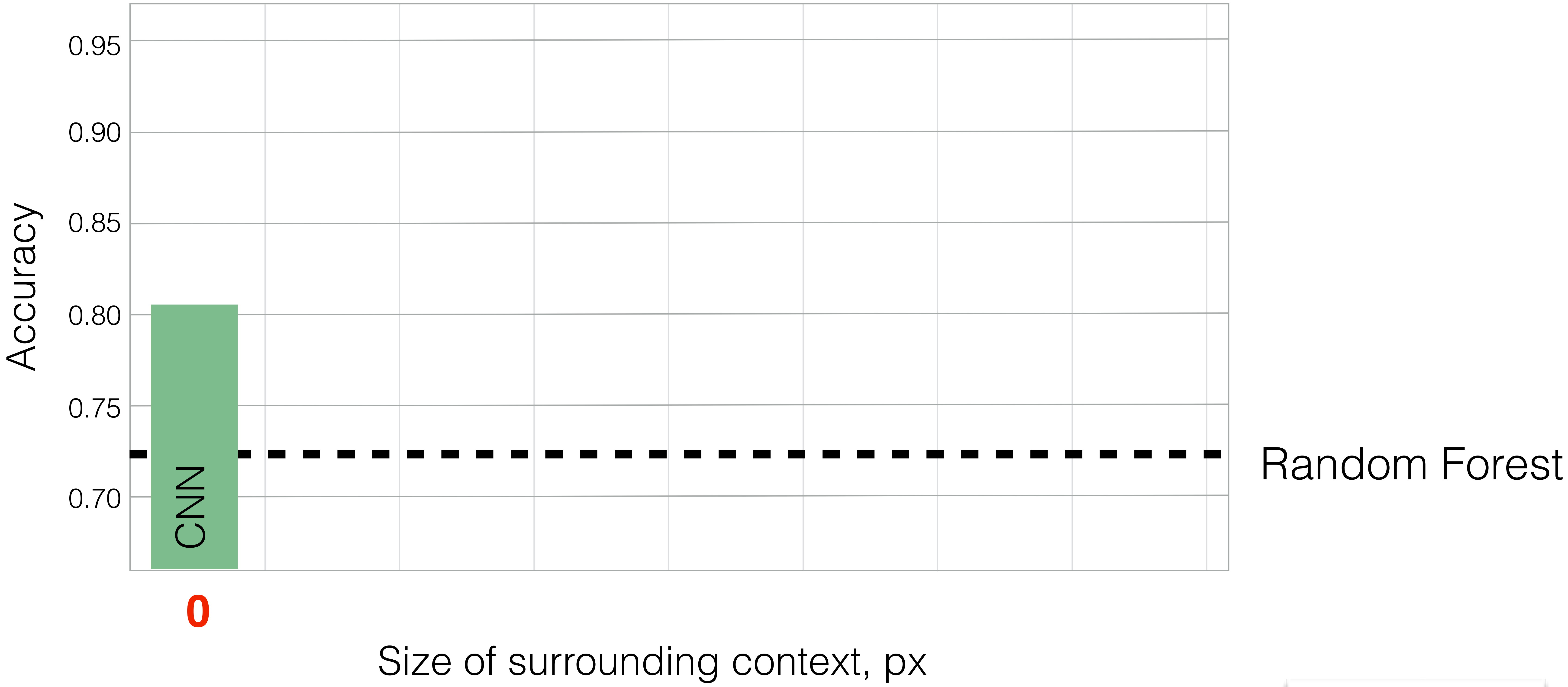

Bounding box
Cell frame
Harmony
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…
Random Forest (RF)
Both models try to predict a correct cell type (1 out of 7)
Cell frames dataset
Convolutional Neural Network (CNN)

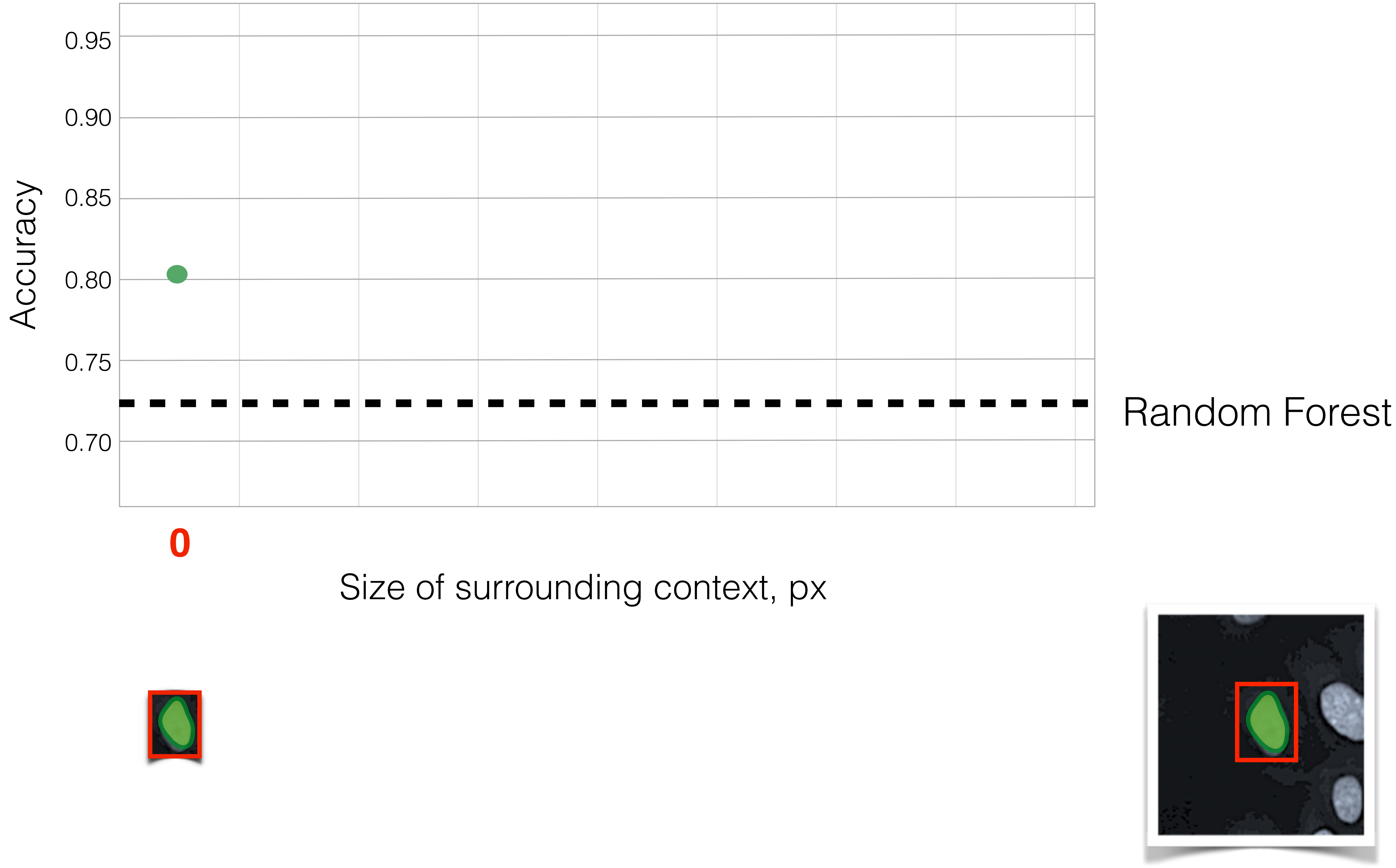


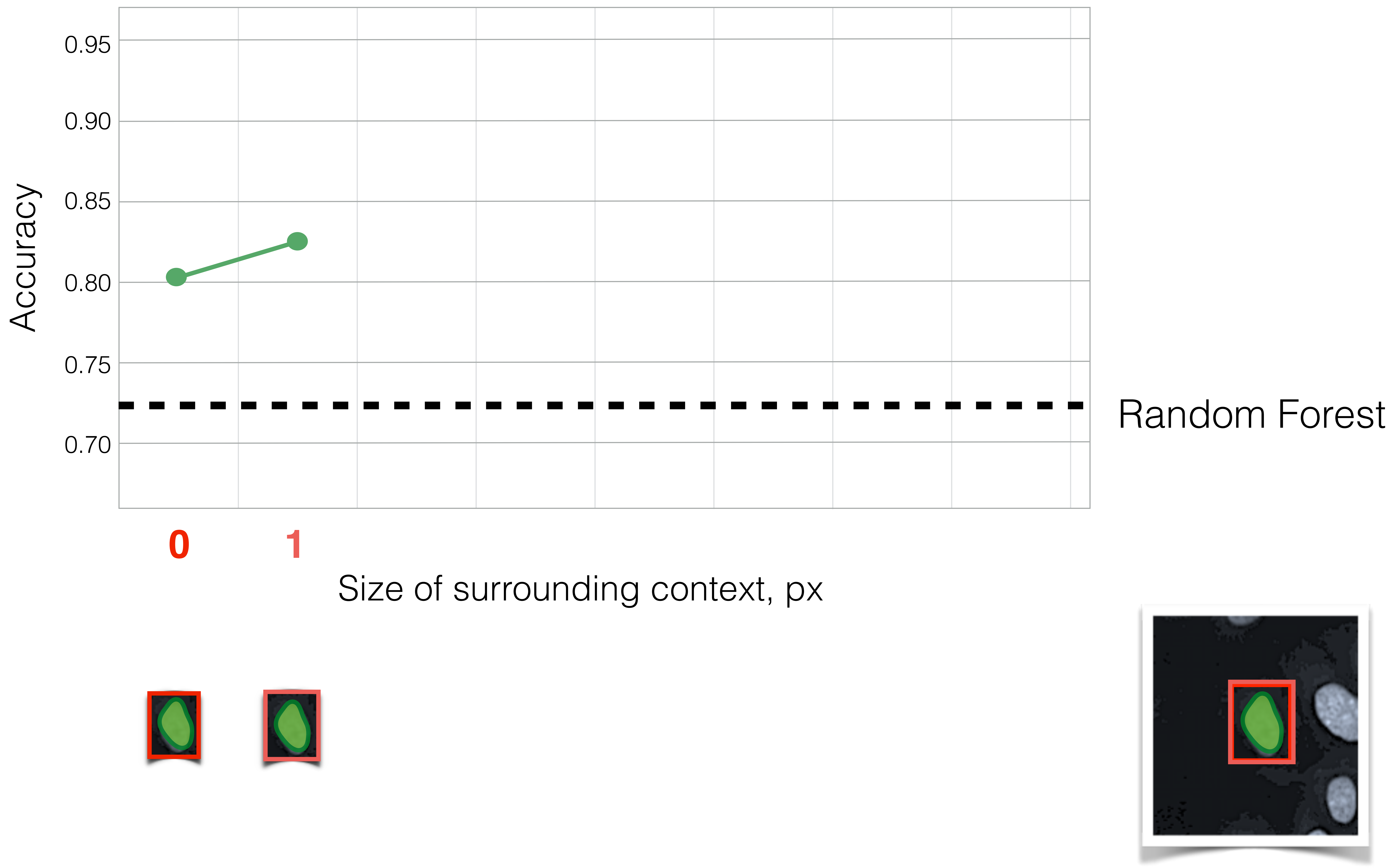
Bounding box
Cell frame
Harmony
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…
Feature dataset
colour: green
diameter: 18 px;
square: 128 px;
roundness: 0.69;
200 more…
Random Forest (RF)
Both models try to predict a
correct cell type (1 out of 7)
Accuracy
0.95
0.90
0.85
0.80
0.75
0.70
CNN
RF
Cell frames dataset
Convolutional Neural Network (CNN)

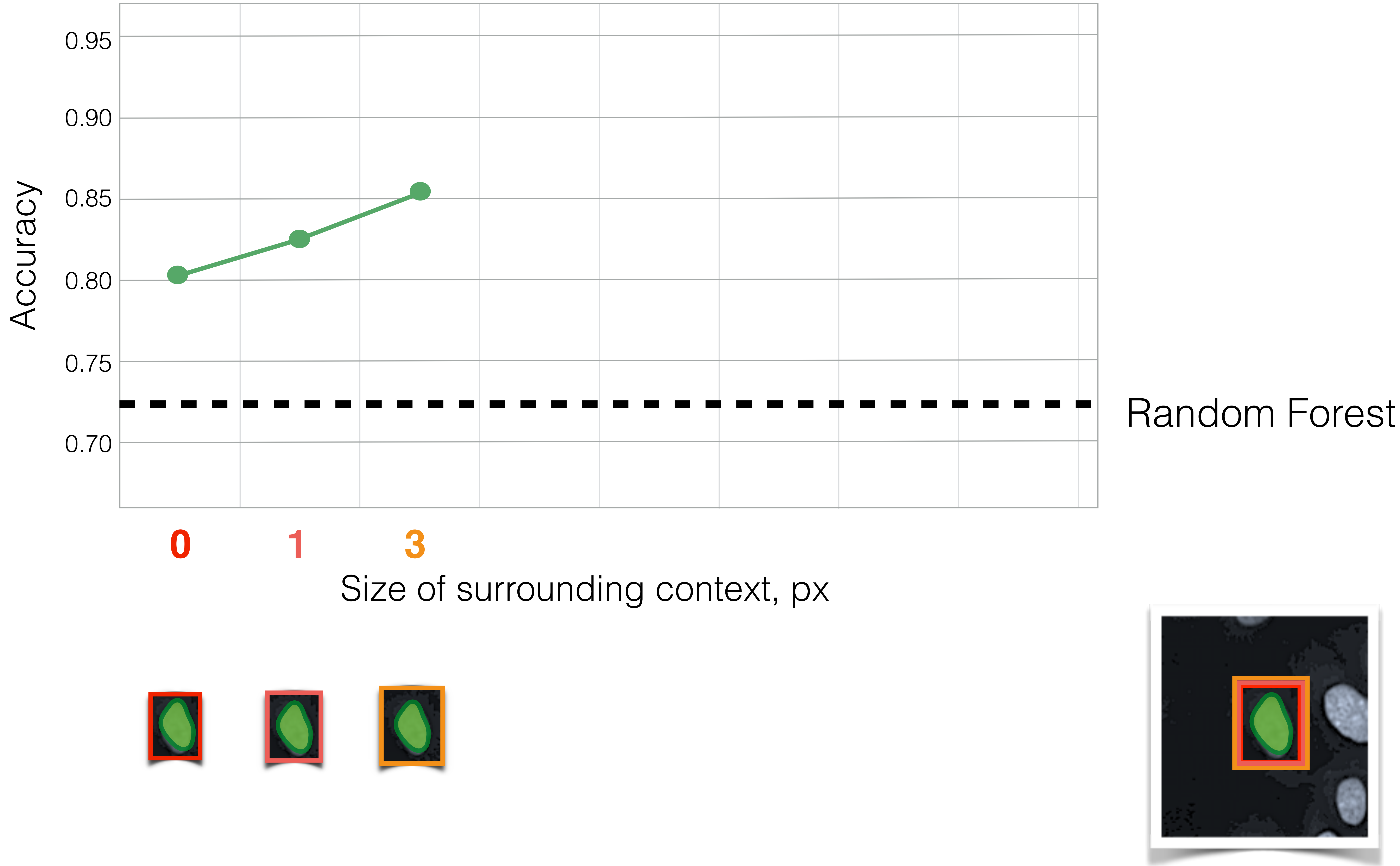


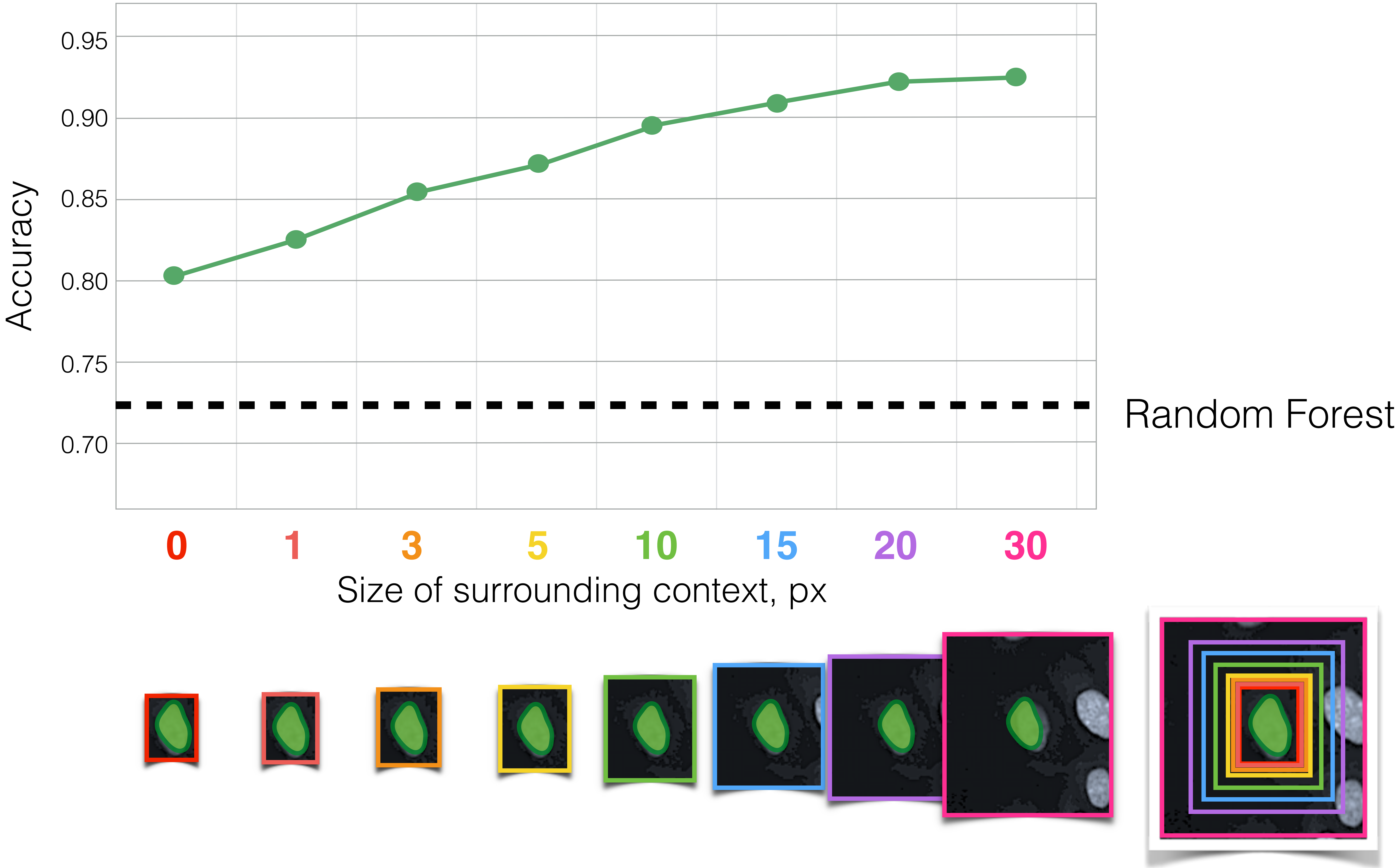
Accuracy
0.95
0.90
0.85
0.80
0.75
0.70
CNN
Random Forest


0.95
0.90
0.85
0.80
0.75
0.70
Accuracy
CNN
Random Forest
0
Size of surrounding context, px

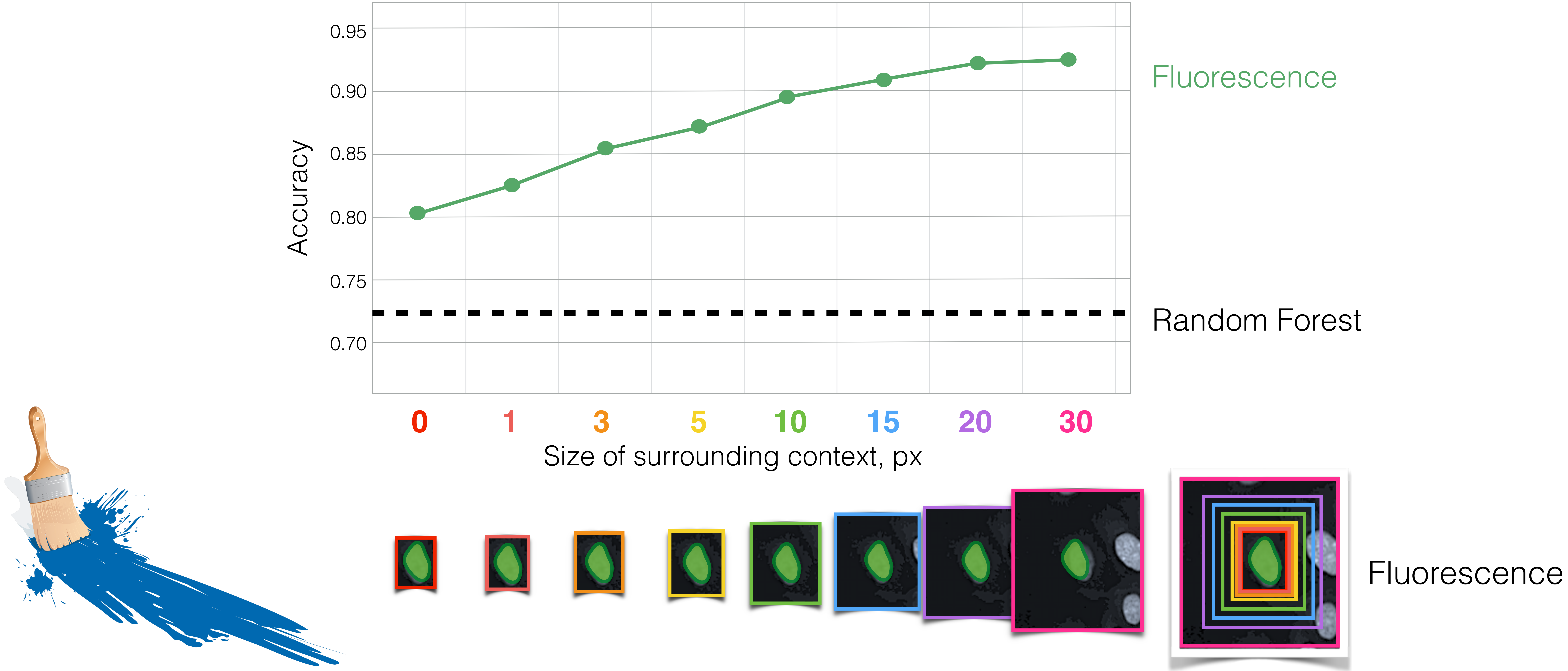

Accuracy
0.95
0.90
0.85
0.80
0.75
0.70
Random Forest
0
Size of surrounding context, px

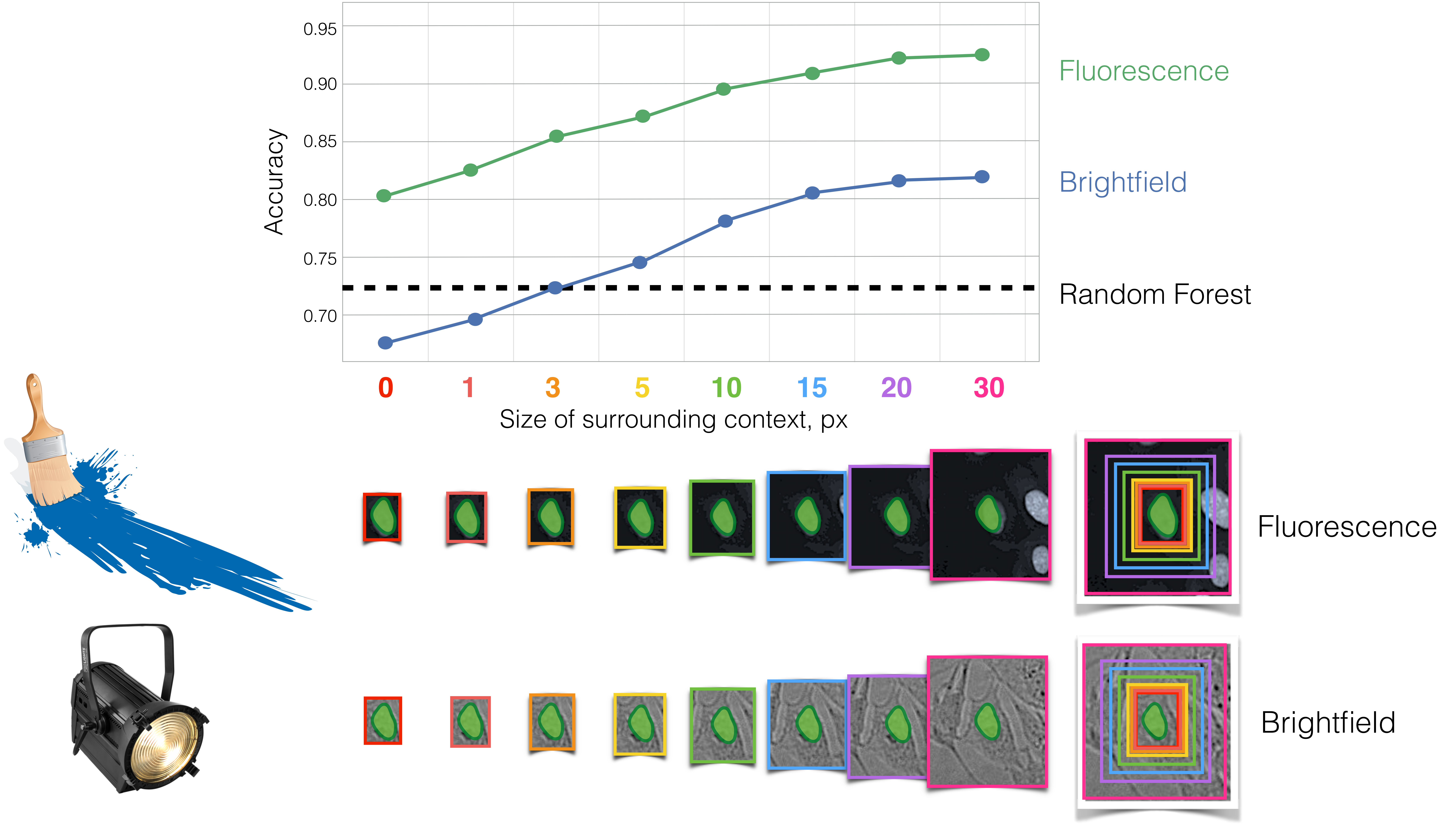

0.95
0.90
0.85
0.80
0.75
0.70
Accuracy
Random Forest
0
1
Size of surrounding context, px


0.95
0.90
0.85
0.80
0.75
0.70
Accuracy
Random Forest
0
1
3
Size of surrounding context, px

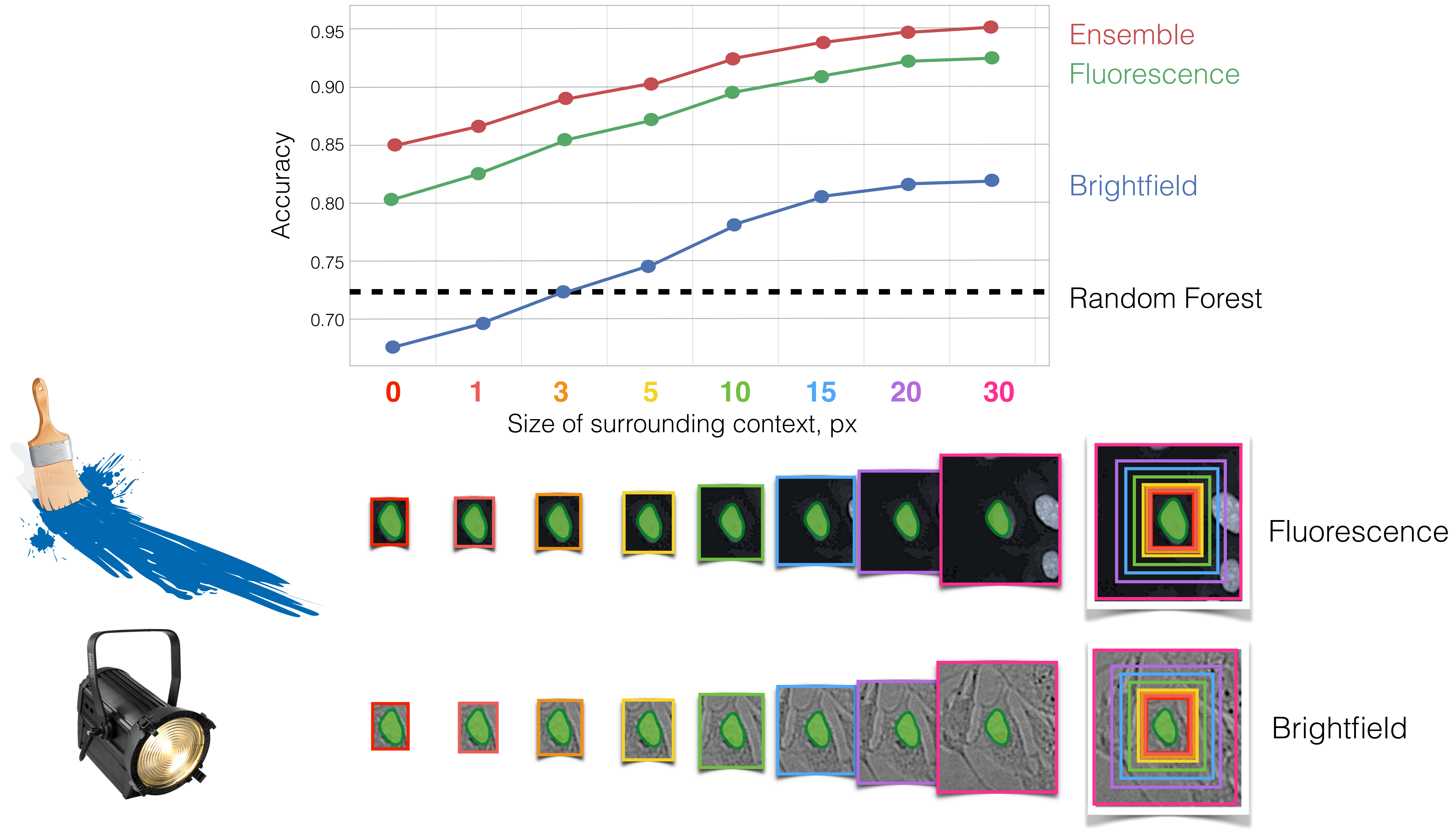


Accuracy
0.95
0.90
0.85
0.80
0.75
0.70
Random Forest
0
1
3
5
10
15
20
30
Size of surrounding context, px

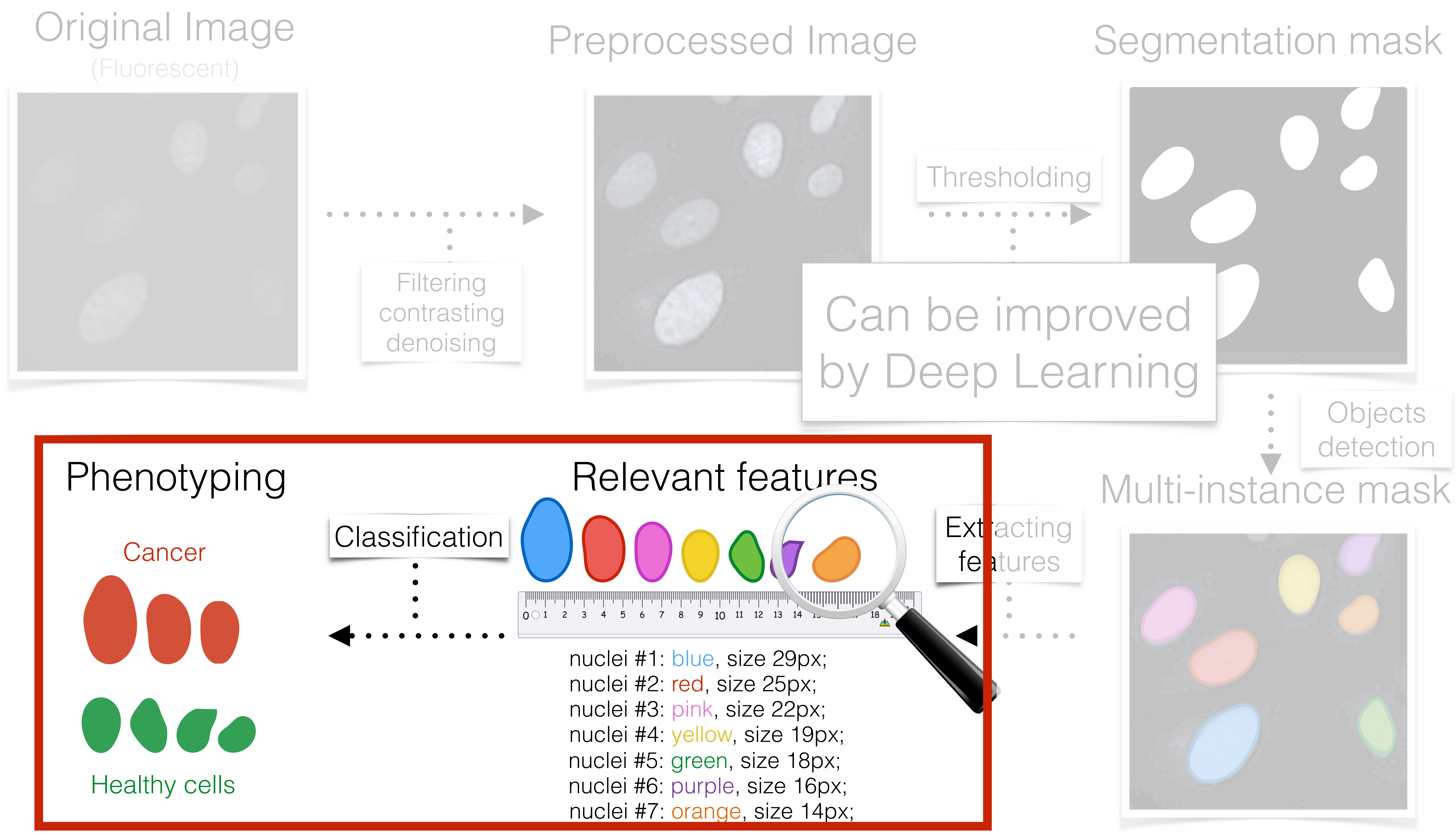

Accuracy
0.95
0.90
0.85
0.80
0.75
0.70
Fluorescence
Random Forest
0
1
3
5
10
15
20
30
Size of surrounding context, px
Fluorescence

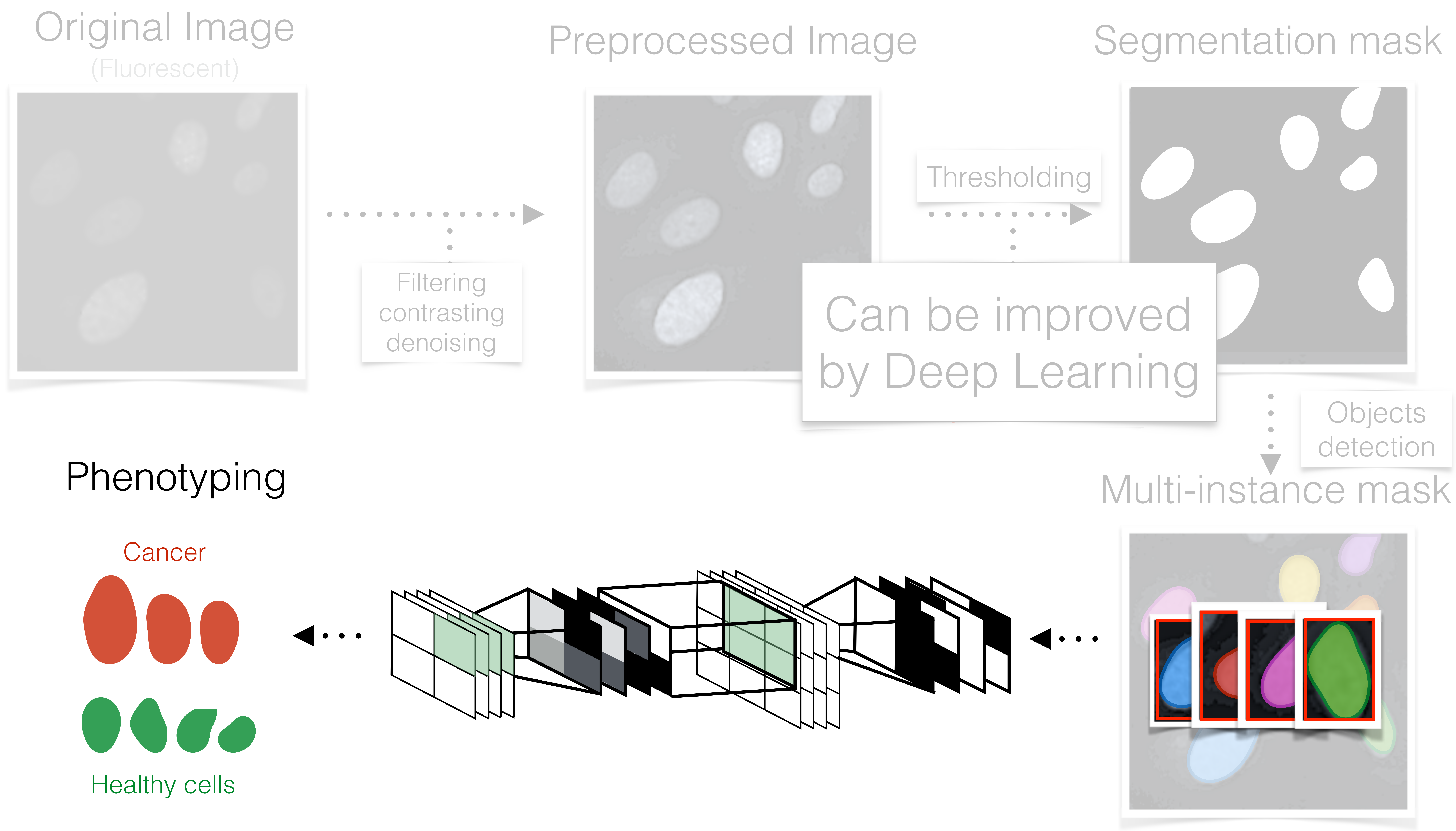


Accuracy
0.95
0.90
0.85
0.80
0.75
0.70
Fluorescence
Brightfield
Random Forest
0
1
3
5
10
15
20
30
Size of surrounding context, px
Fluorescence
Brightfield



Accuracy
0.95
0.90
0.85
0.80
0.75
0.70
Ensemble
Fluorescence
Brightfield
Random Forest
0
1
3
5
10
15
20
30
Size of surrounding context, px
Fluorescence
Brightfield


Original Image
(Fluorescent)
Filtering
contrasting
denoising
Preprocessed Image
Thresholding
Segmentation mask
Can be improved
by Deep Learning
Objects
detection
Multi-instance mask
Extracting
features
Relevant features
nuclei #1: blue, size 29px;
nuclei #2: red, size 25px;
nuclei #3: pink, size 22px;
nuclei #4: yellow, size 19px;
nuclei #5: green, size 18px;
nuclei #6: purple, size 16px;
nuclei #7: orange, size 14px;
Classification
Phenotyping
Cancer
Healthy cells


Original Image
(Fluorescent)
Preprocessed Image
Segmentation mask
Filtering
contrasting
denoising
Thresholding
Can be improved
by Deep Learning
Objects
detection
Multi-instance mask
Phenotyping
Cancer
Healthy cells

UNIVERSITY OF TARTU
Institute of Computer Science

Cell Phenotyping with Convolutional Neural Networks

Mikhail Papkov[1], Kaupo Palo[2], Leopold Parts[1,3], Dmytro Fishman[1,4]

[1] Institute of Computer Science, University of Tartu, [2] PerkinElmer, Inc., [3] The Wellcome Sanger Institute, [4] Quretec Ltd

Contacts:
mikhail.papkov@gmail.com
dmytro.fishman@gmail.com

Abstract

Cell phenotyping in microscopy images plays an important role in various biological and medical applications, e.g. cancer diagnostics. A vast variety of conditions, magnification and image modalities make this task a very challenging problem for classical image recognition methods. At the same time, deep learning has been shown to perform well under these conditions. Here we use a convolutional neural network to classify cell cultures. We show that deep learning outperforms traditional machine learning trained on handcrafted features extracted by the PerkinElmer software.

Introduction

Goals and Questions

1. How well can individual cells be classified into seven cell lines?
2. How do neural networks perform compare to «traditional» machine learning methods (Random Forest) trained on standard features dataset?
3. How does the same neural network architecture perform on different image modalities (fluorescence, brightfield)?
4. How important is the context around nuclei for classification?

The main motivation behind these questions is to help researchers who work with cell cultures. Automated cell image analysis could potentially reduce the amount of routine and speed up the studies.

Data Description

The dataset consisted of 3024 images 1080×1080 with 70 - 200 cells each in fluorescent and brightfield modalities. All cells on each image belong to one of the seven cell lines listed in Table 1. Examples of images are shown in Figure 1. For each

Table 1: Dataset description

Cell line	Description
A549	**human** adenocarcinomic alveolar basal epithelial
HT1080	**human** fibrosarcoma

Network architecture

Here we used altered Dürr and Sick architecture [2] proposed for single-cell phenotype classification. The number of dense layers was reduced compared to the original version of the architecture in order to prevent overfitting. The network structure is summarized in Table 2. The network and learning parameters are listed below:

- Implemented with Keras Python library using Tensorflow backend
- Batch normalization after convolutional and dense layers with batch size 8
- Scheduled learning rate (from 5×10^{-4} to 1.5×10^{-5})
- 25 epochs
- L2 regularization ($l = 5 \times 10^{-5}$)
- Adam optimizer

Table 2: Architecture of the Convolutional Neural Network [2]. Each row represents network layer in top-down order.

Layers	Output
Input	$70 \times 70 \times 1$
Conv 2D (3x3)	$68 \times 68 \times 32$
Conv 2D (3x3)	$66 \times 66 \times 32$
Max pool 2D (2x2)	$33 \times 33 \times 32$
Conv 2D (3x3)	$31 \times 31 \times 64$
Conv 2D (3x3)	$29 \times 29 \times 64$
Max pool 2D (2x2)	$14 \times 14 \times 64$
Conv 2D (3x3)	$12 \times 12 \times 128$
Conv 2D (3x3)	$10 \times 10 \times 128$
Max pool 2D (2x2)	$5 \times 5 \times 128$
Dense	100
Dropout (0.2)	100
Dense	50
Output	7

We compare the network to the Random Forest trained on features extracted from respective cells with the PerkinElmer software. Random Forest classifier was tuned with parameter grid search and recursive feature elimination.

Results

We found the network performance to be dependent on the size of context around the nuclei. Learn-

UNIVERSITY OF TARTU
Institute of Computer Science

Cell Phenotyping with Convolutional Neural Networks

Contacts:
mikhail.papkov@gmail.com
dmytro.fishman@gmail.com

Mikhail Papkov[1], Kaupo Palo[2], Leopold Parts[1,3], Dmytro Fishman[1,4]

[1] Institute of Computer Science, University of Tartu, [2] PerkinElmer, Inc., [3] The Wellcome Sanger Institute, [4] Quretec Ltd

Mikhail Papkov[1], Kaupo Palo[2], Leopold Parts[1,3], Dmytro Fishman[1,4]

e.g. ...
cha...
to perform well under these conditions. Here we use a convolutional neural network to classify cell cultures. We show that deep learning outperforms traditional machine learning trained on handcrafted features extracted by the PerkinElmer software.

Introduction

Goals and Questions

1. How well can individual cells be classified into seven cell lines?
2. How do neural networks perform compare to «traditional» machine learning methods (Random Forest) trained on standard features dataset?
3. How does the same neural network architecture perform on different image modalities (fluorescence, brightfield)?
4. How important is the context around nuclei for classification?

The main motivation behind these questions is to help researchers who work with cell cultures. Automated cell image analysis could potentially reduce the amount of routine and speed up the studies.

Data Description

The dataset consisted of 3024 images 1080×1080 with 70 - 200 cells each in fluorescent and brightfield modalities. All cells on each image belong to one of the seven cell lines listed in Table 1. Examples of images are shown in Figure 1. For each

Table 1: Dataset description

Cell line	Description
A549	**human** adenocarcinomic alveolar basal epithelial
HT1080	**human** fibrosarcoma

proposed for single-cell phenotype classification. The number of dense layers was reduced compared to the original version of the architecture in order to prevent overfitting. The network structure is summarized in Table 2. The network and learning parameters are listed below:

- Implemented with Keras Python library using Tensorflow backend
- Batch normalization after convolutional and dense layers with batch size 8
- Scheduled learning rate (from 5×10^{-4} to 1.5×10^{-5})
- 25 epochs
- L2 regularization ($l = 5 \times 10^{-5}$)
- Adam optimizer

Table 2: Architecture of the Convolutional Neural Network [2]. Each row represents network layer in top-down order.

Layers	Output
Input	$70 \times 70 \times 1$
Conv 2D (3x3)	$68 \times 68 \times 32$
Conv 2D (3x3)	$66 \times 66 \times 32$
Max pool 2D (2x2)	$33 \times 33 \times 32$
Conv 2D (3x3)	$31 \times 31 \times 64$
Conv 2D (3x3)	$29 \times 29 \times 64$
Max pool 2D (2x2)	$14 \times 14 \times 64$
Conv 2D (3x3)	$12 \times 12 \times 128$
Conv 2D (3x3)	$10 \times 10 \times 128$
Max pool 2D (2x2)	$5 \times 5 \times 128$
Dense	100
Dropout (0.2)	100
Dense	50
Output	7

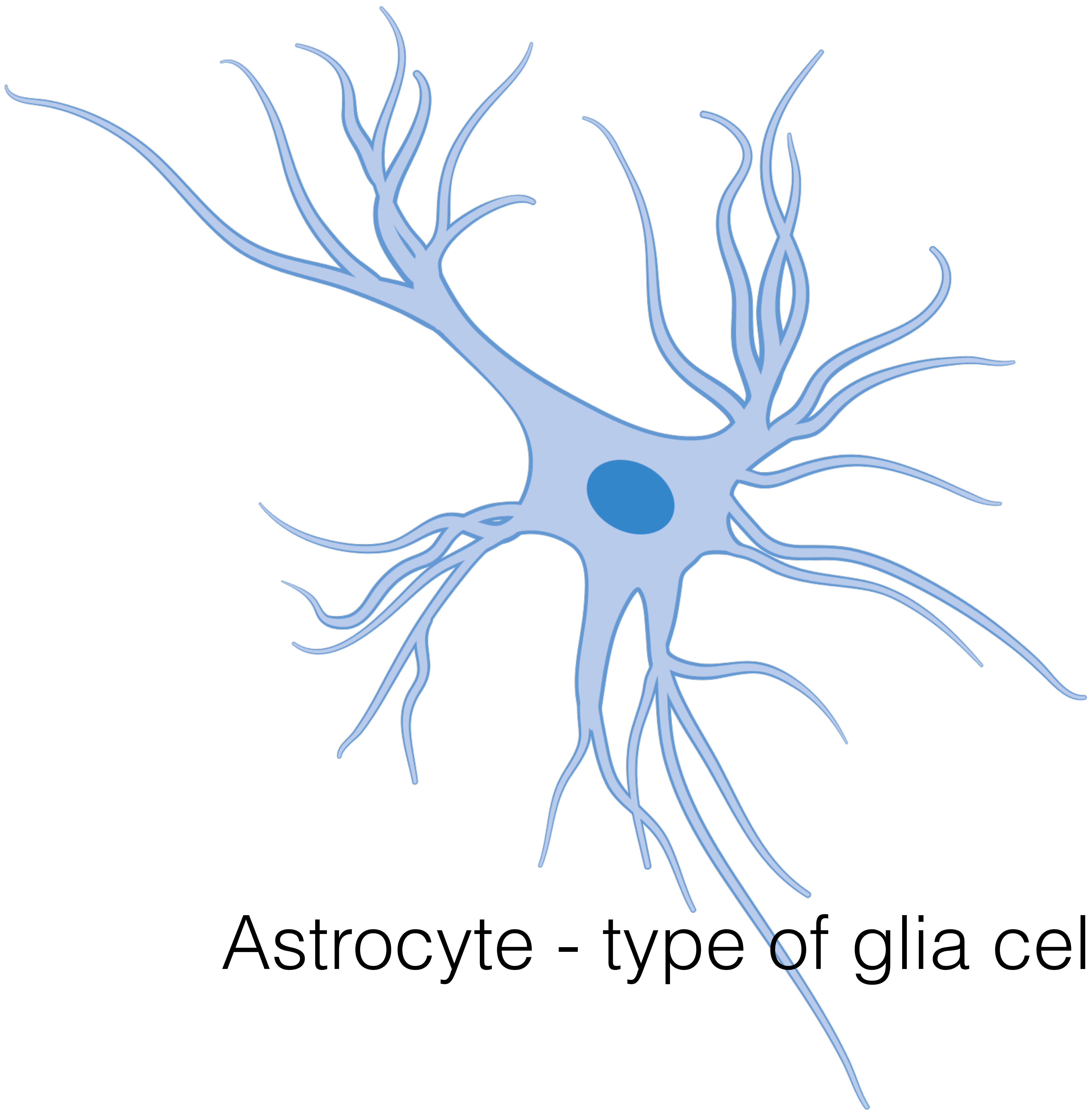
We compare the network to the Random Forest trained on features extracted from respective cells with the PerkinElmer software. Random Forest classifier was tuned with parameter grid search and recursive feature elimination.

Results

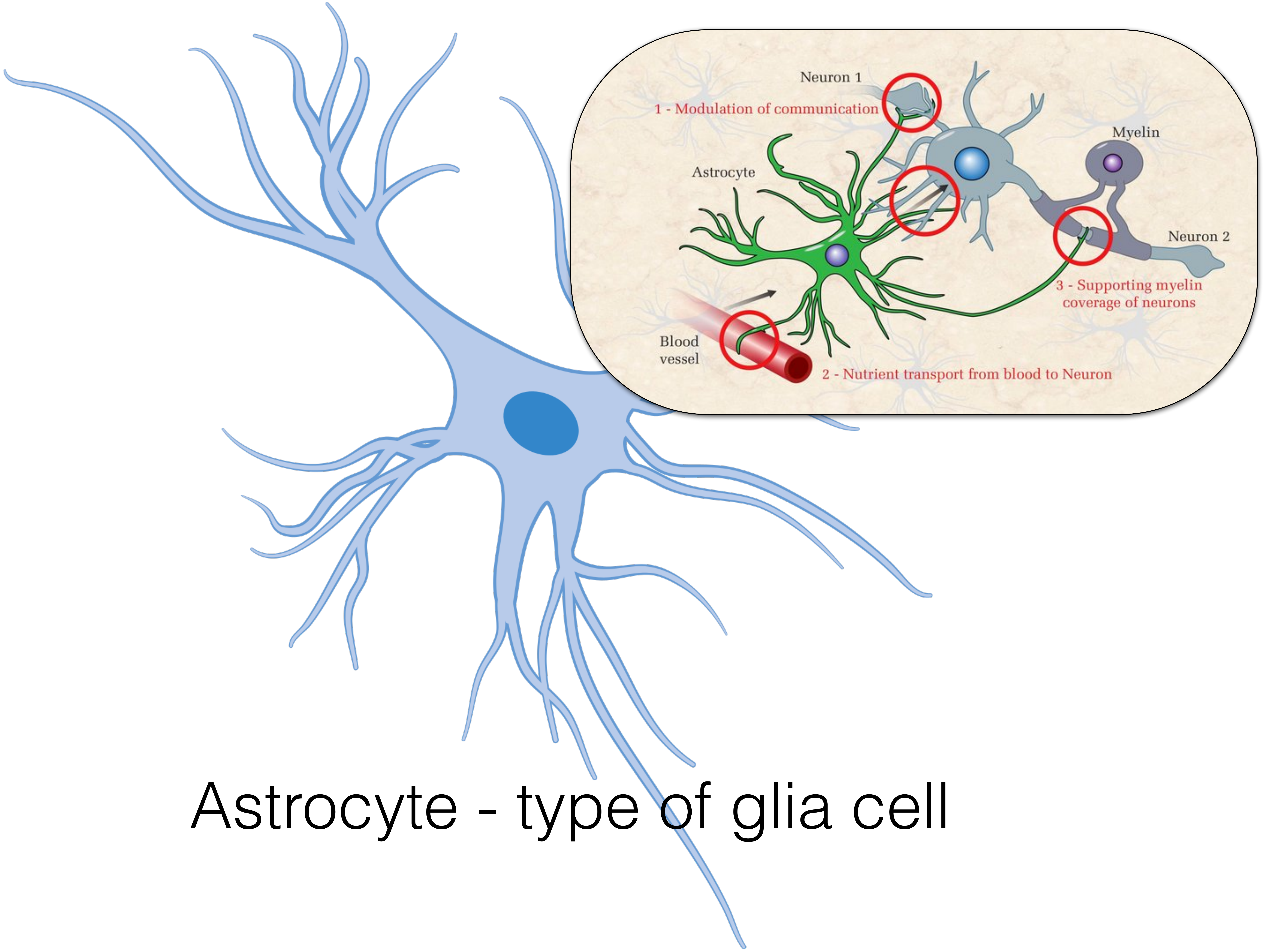
We found the network performance to be dependent on the size of context around the nuclei. Learn-

FANTASTIC BEASTS

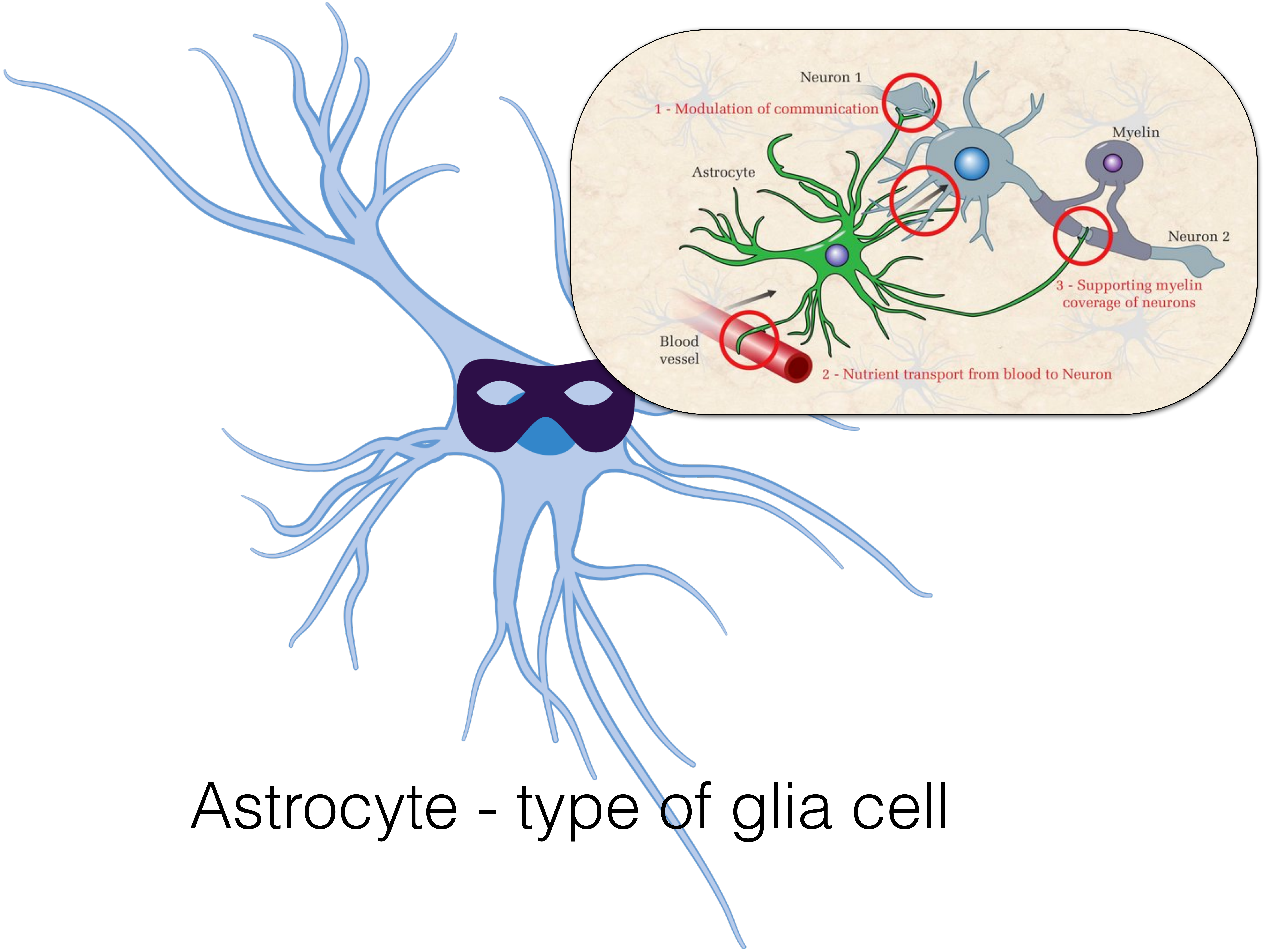
AND WHERE
TO FIND THEM



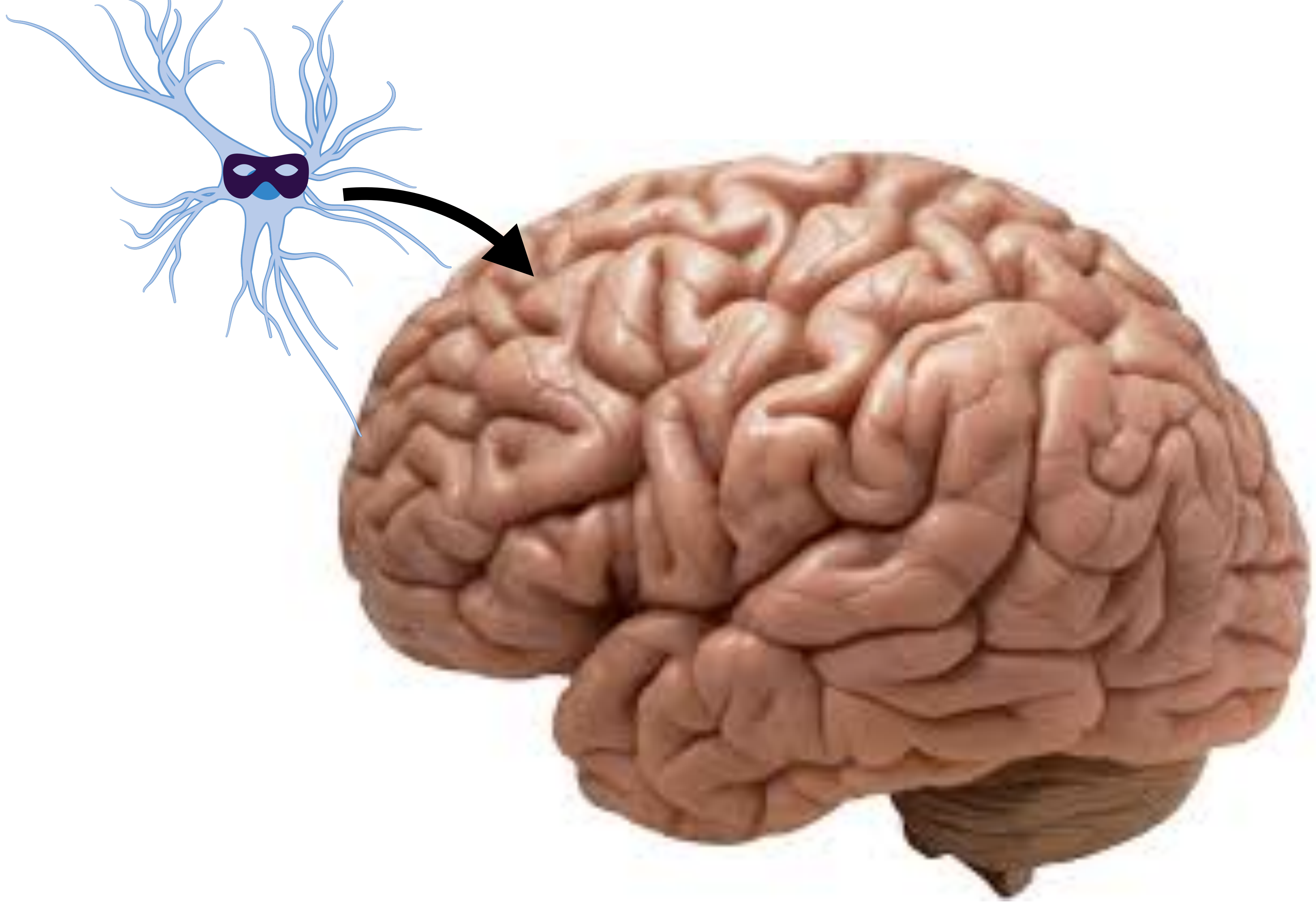
Astrocyte - type of glia cell



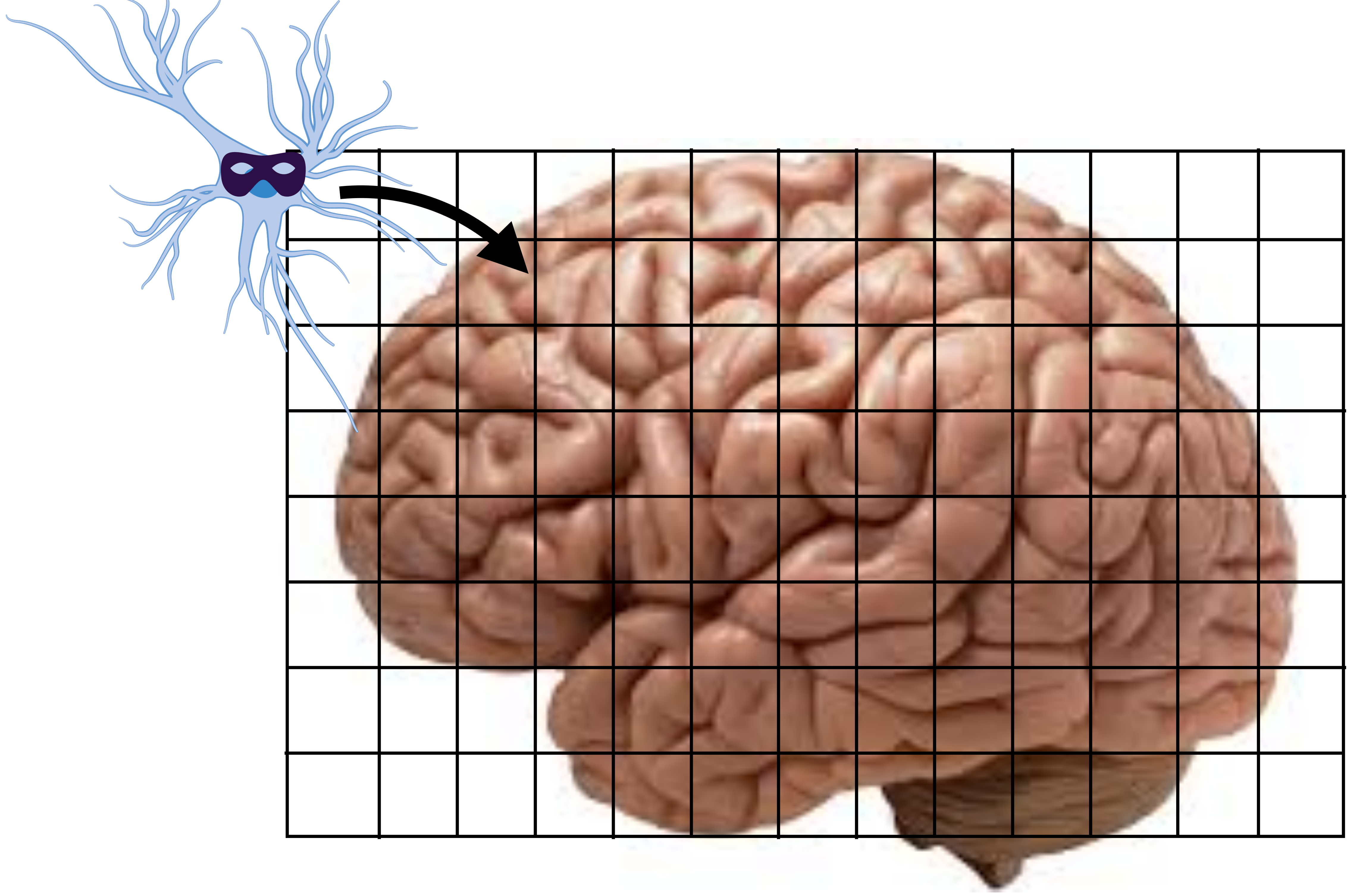
Astrocyte - type of glia cell



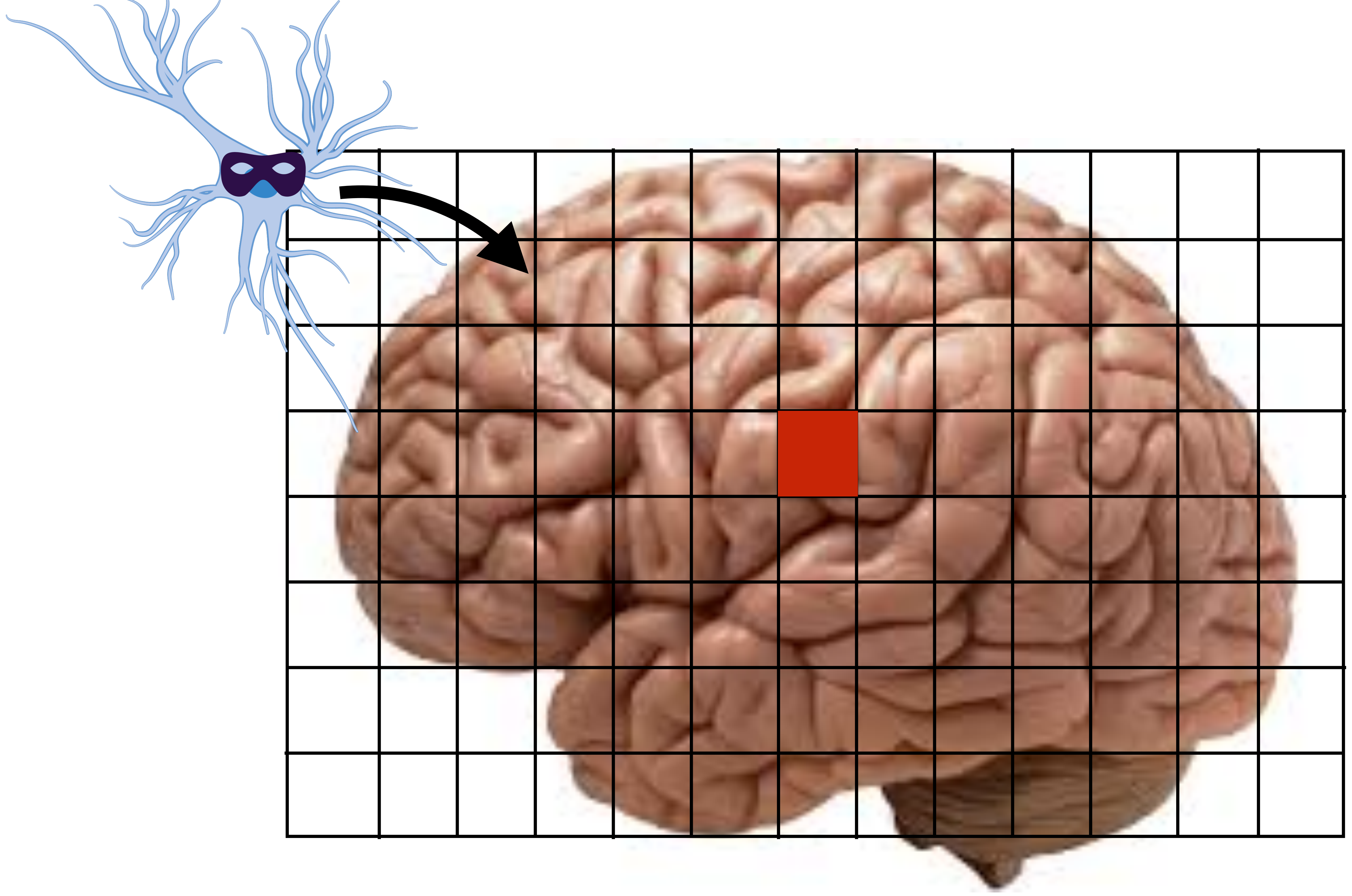
Astrocyte - type of glia cell



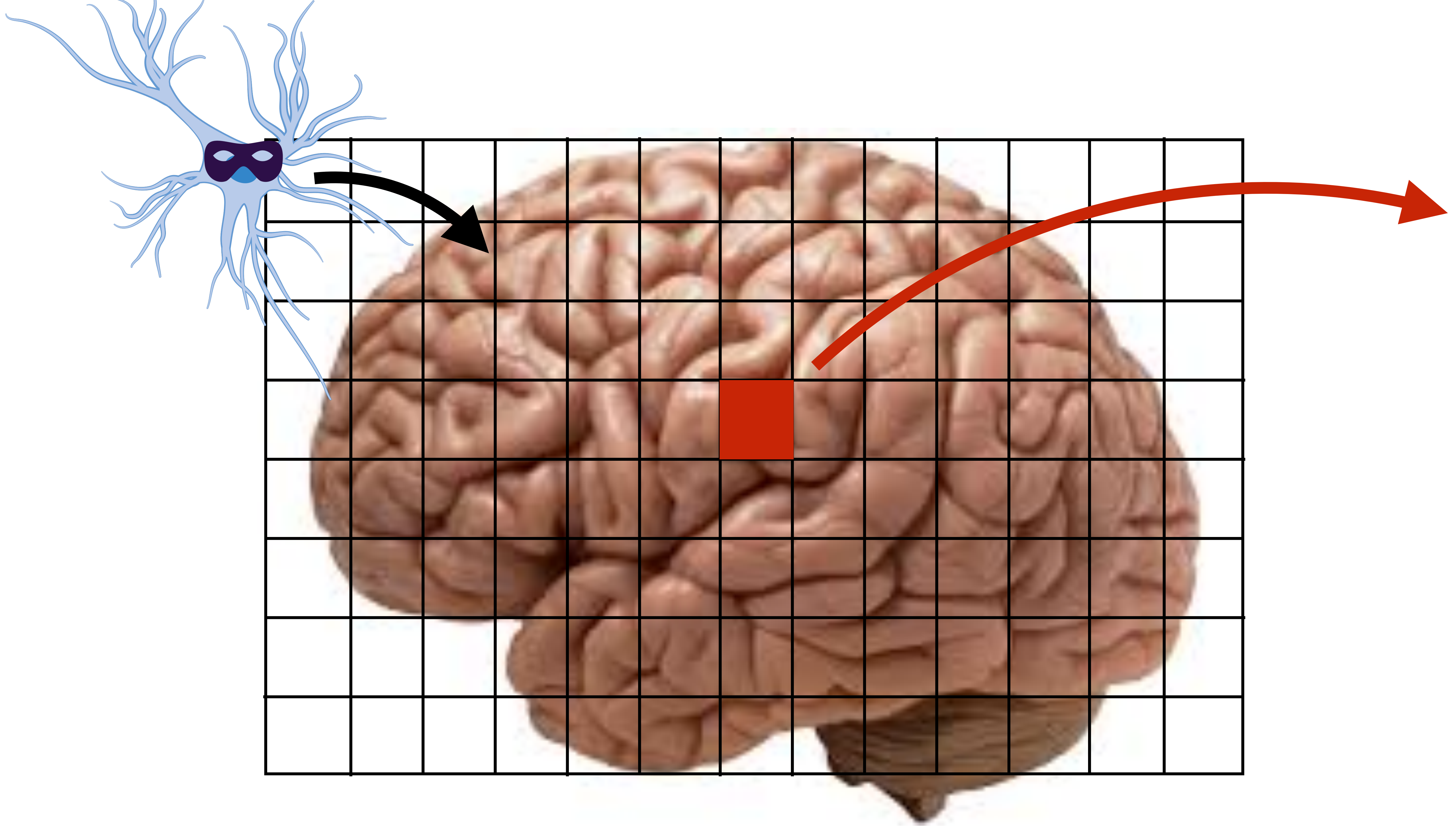
http://www.wingsforlife.com/en/latest/astrocytes-the-star-cells-1637/

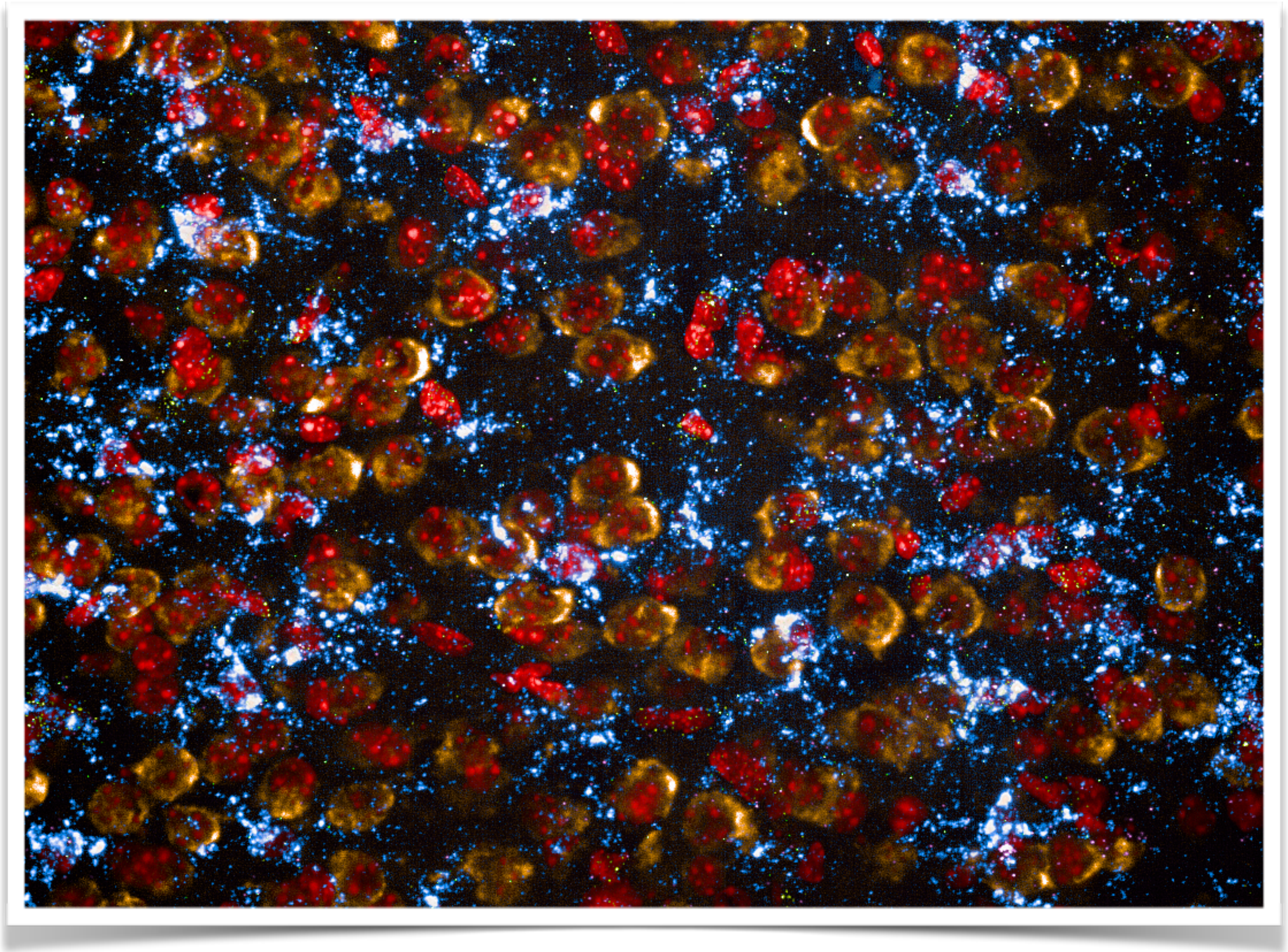


http://www.wingsforlife.com/en/latest/astrocytes-the-star-cells-1637/

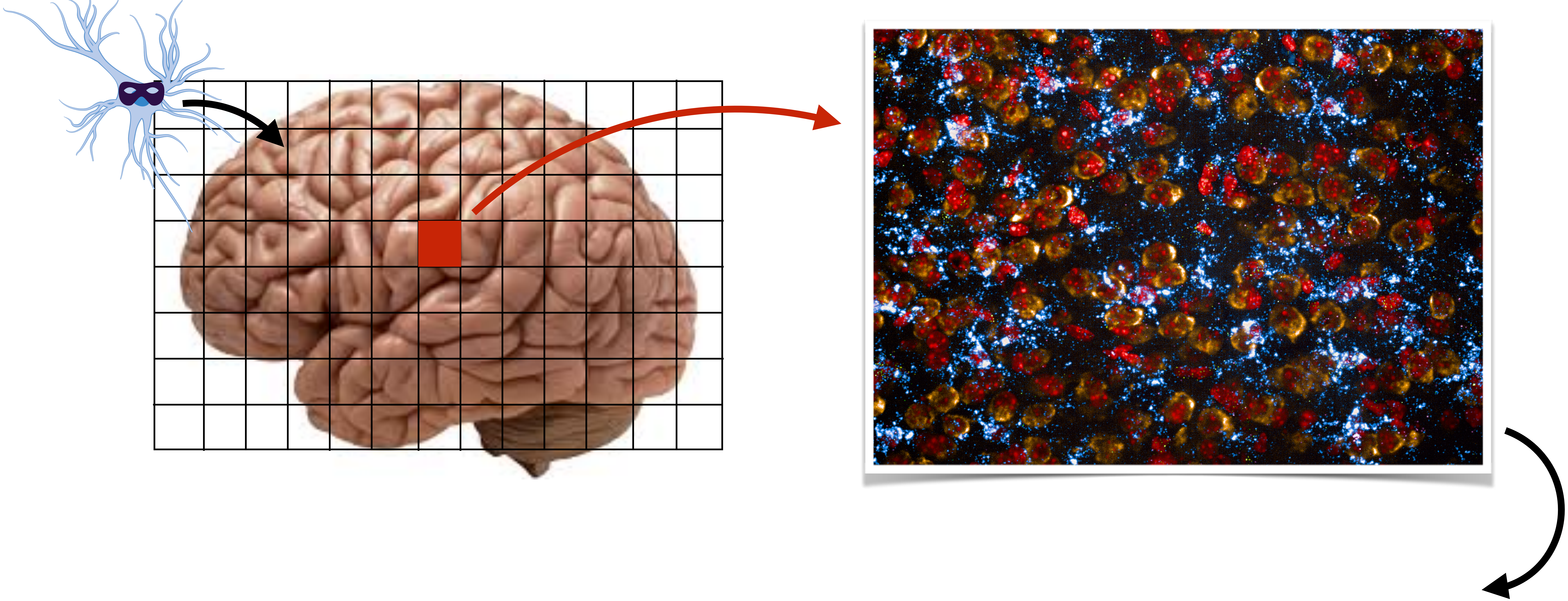


http://www.wingsforlife.com/en/latest/astrocytes-the-star-cells-1637/

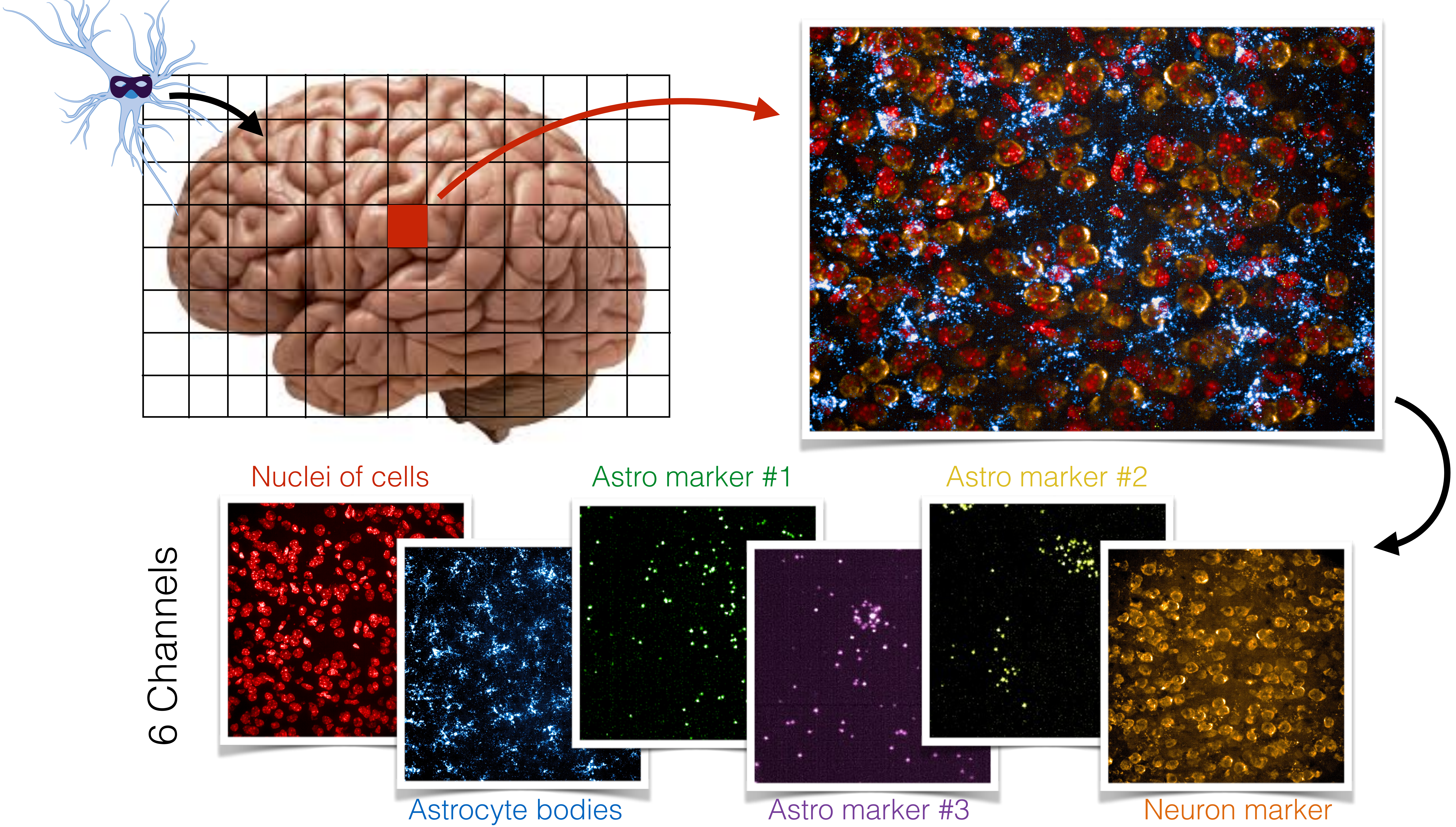




http://www.wingsforlife.com/en/latest/astrocytes-the-star-cells-1637/



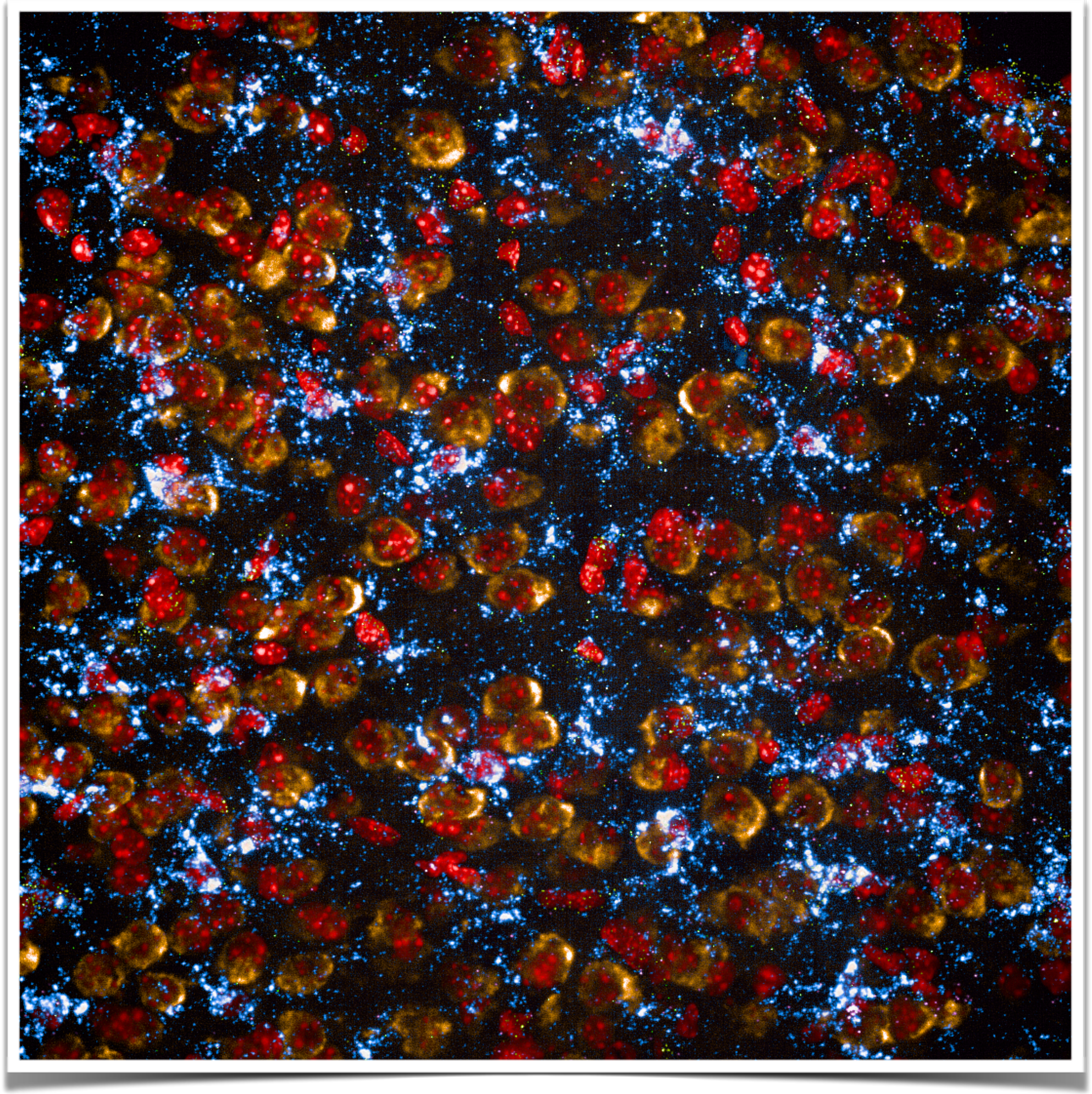
http://www.wingsforlife.com/en/latest/astrocytes-the-star-cells-1637/



http://www.wingsforlife.com/en/latest/astrocytes-the-star-cells-1637/

Current semi-manual pipeline

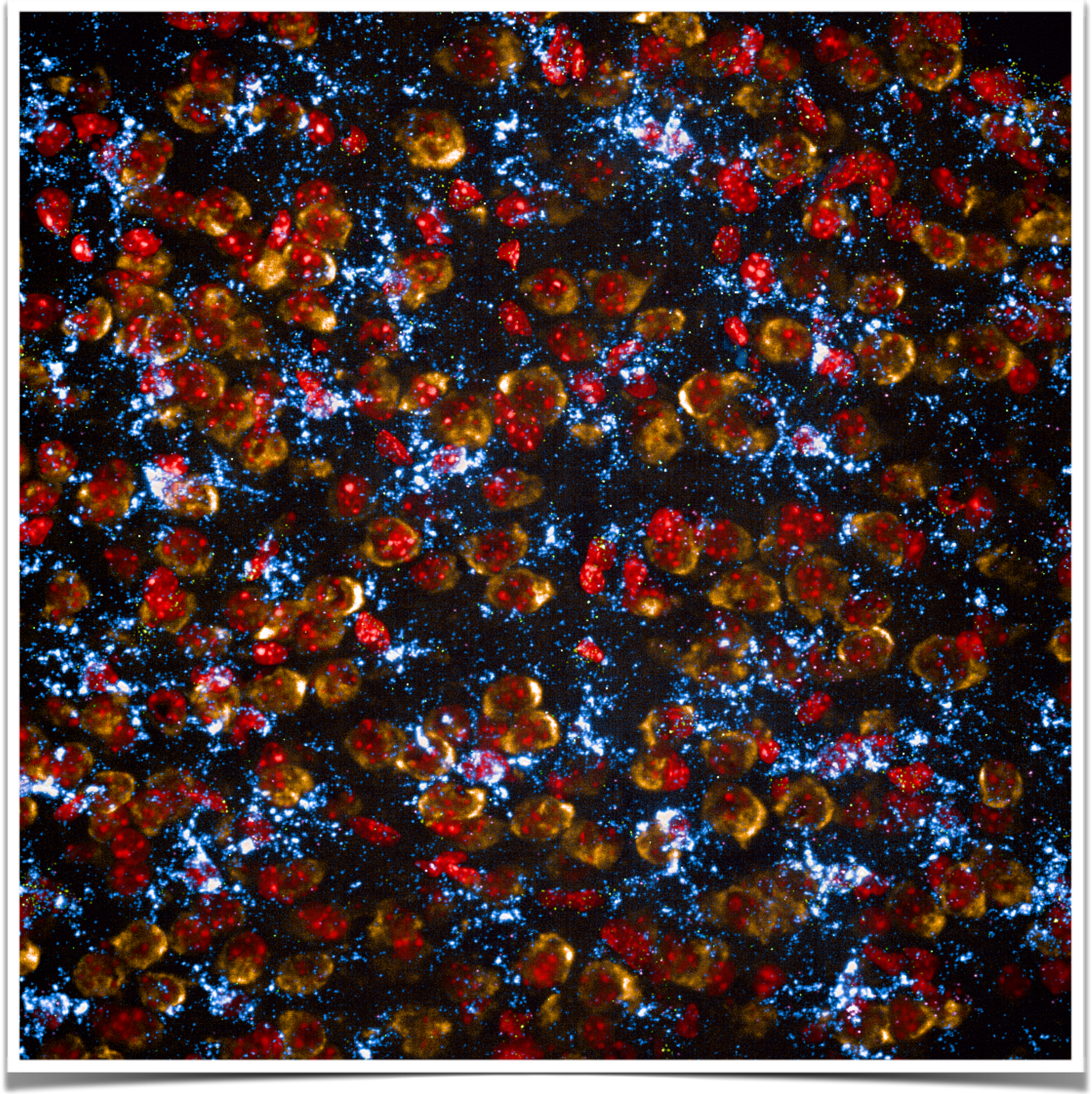
All astrocytes have…



Current semi-manual pipeline

All astrocytes have…

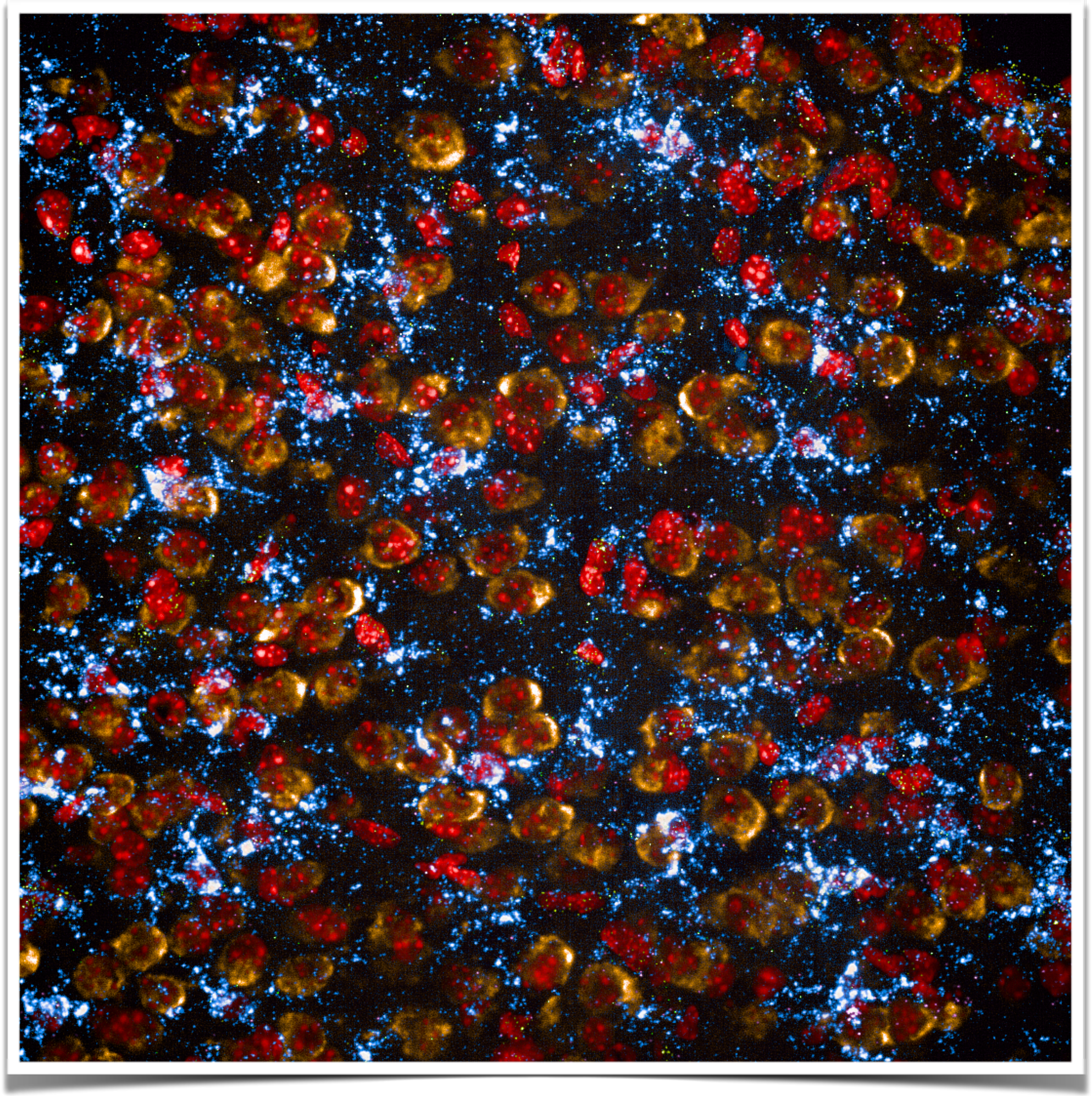
1 One nucleus (in **red**)



Current semi-manual pipeline

All astrocytes have…

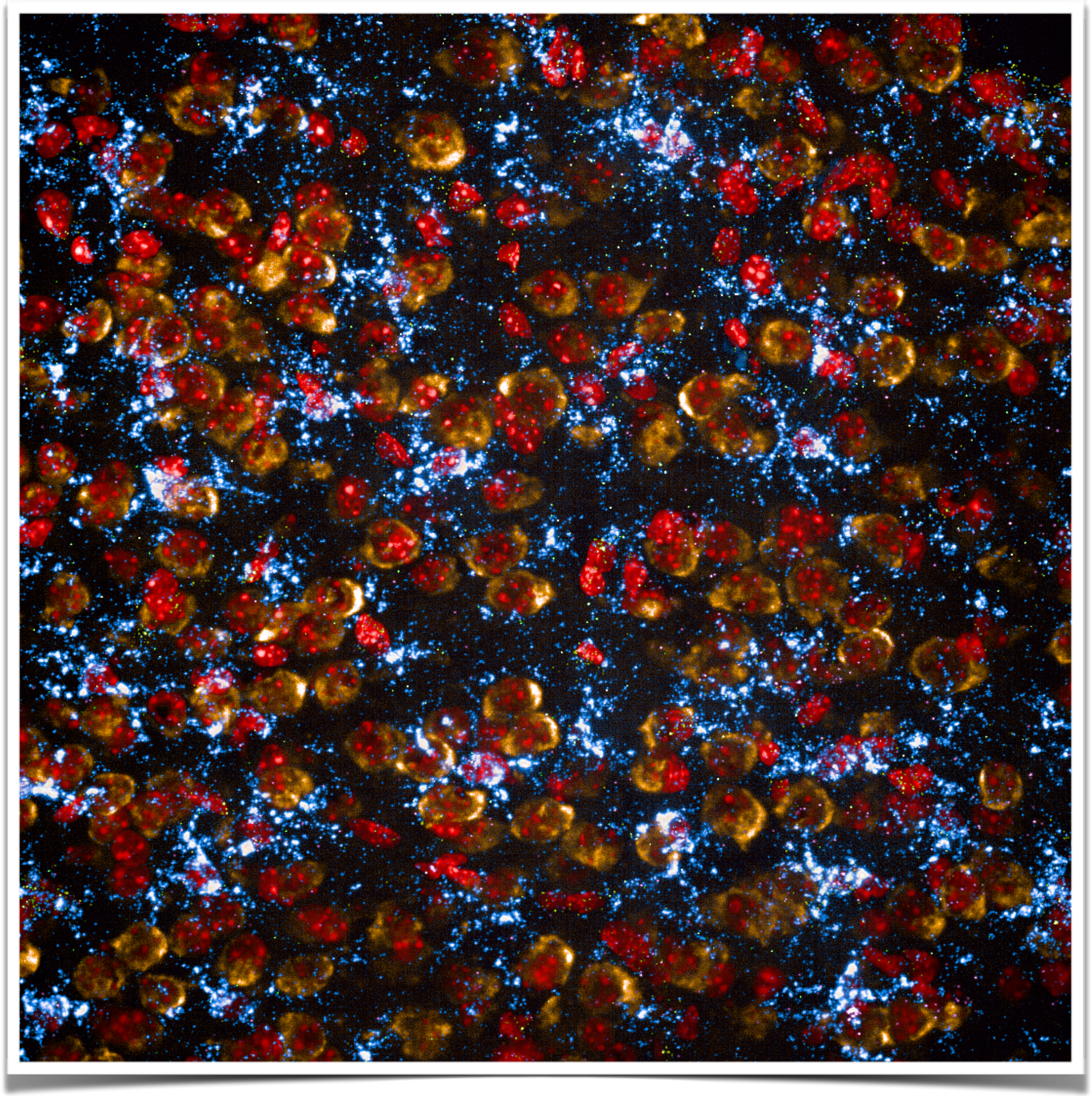
1. One nucleus (in **red**)
2. **Blue body**



Current semi-manual pipeline

All astrocytes have…

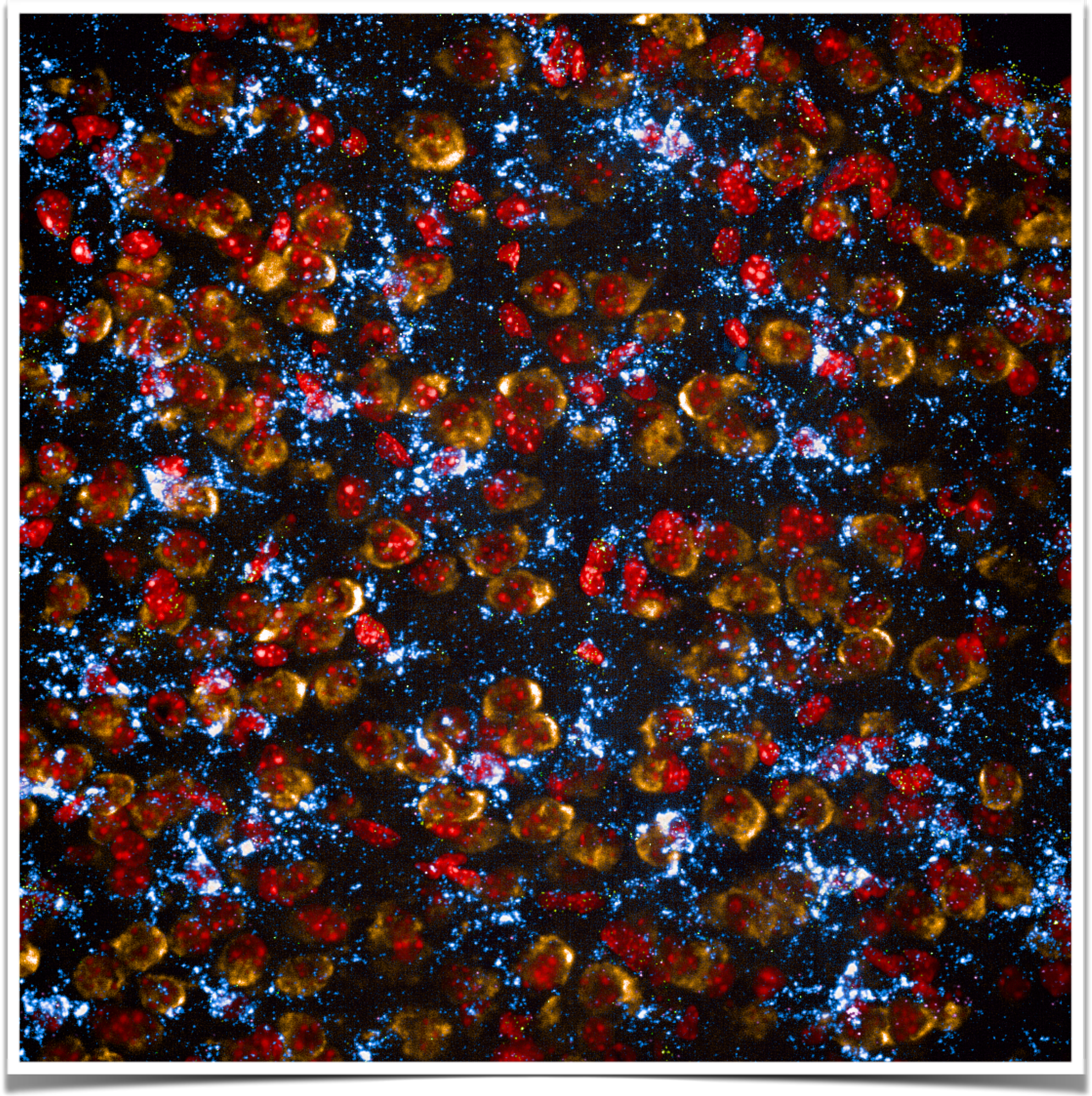
1. One nucleus (in **red**)
2. **Blue body**
3. No **neuron** marker



Current semi-manual pipeline

All astrocytes have…

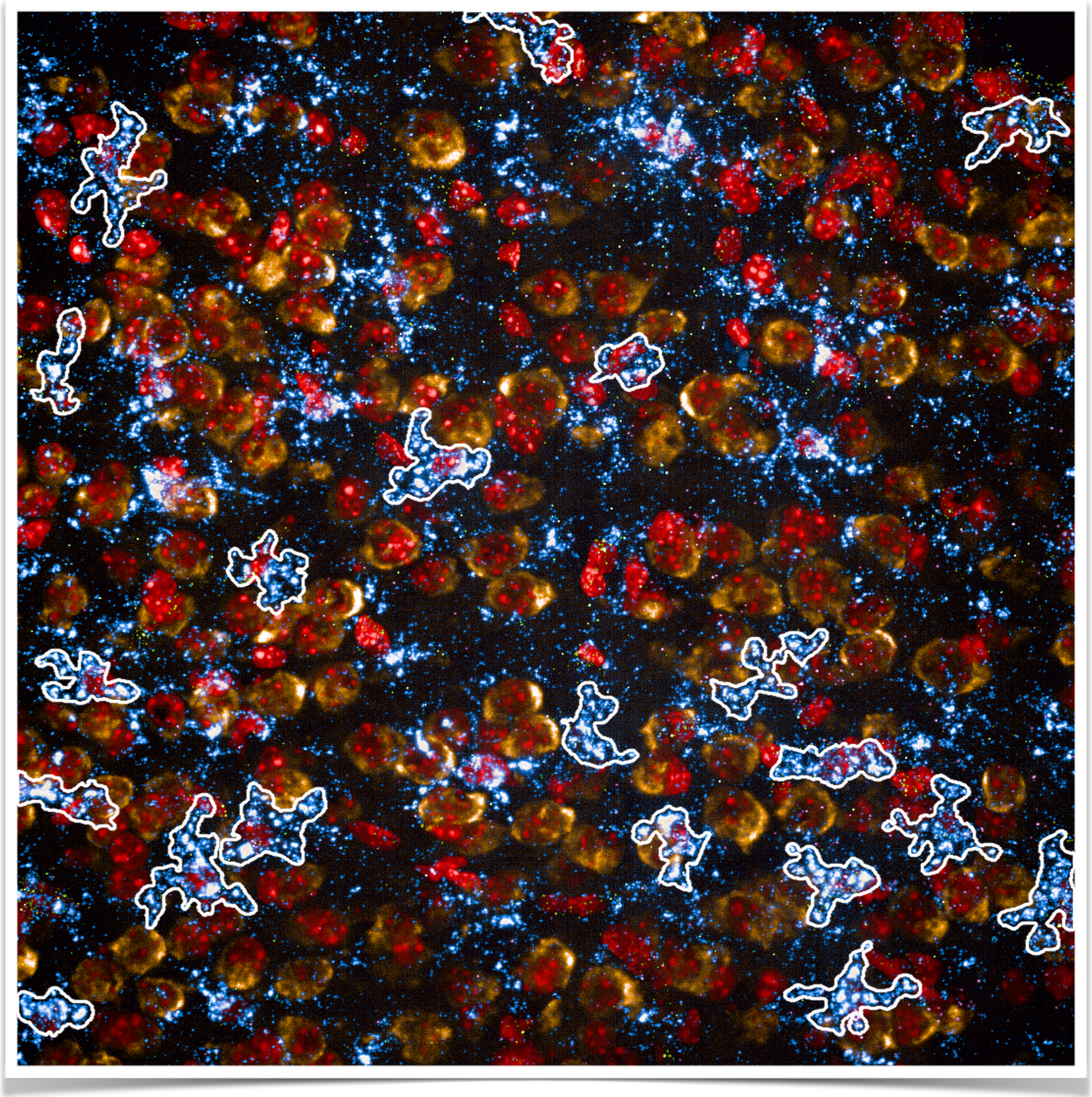
1. One nucleus (in **red**)
2. **Blue body**
3. No **neuron** marker
4. Other markers (**#1**, **#2**, **#3**) moderately present



Current semi-manual pipeline

All astrocytes have…

1. One nucleus (in **red**)
2. **Blue body**
3. No **neuron** marker
4. Other markers (**#1**, **#2**, **#3**) moderately present

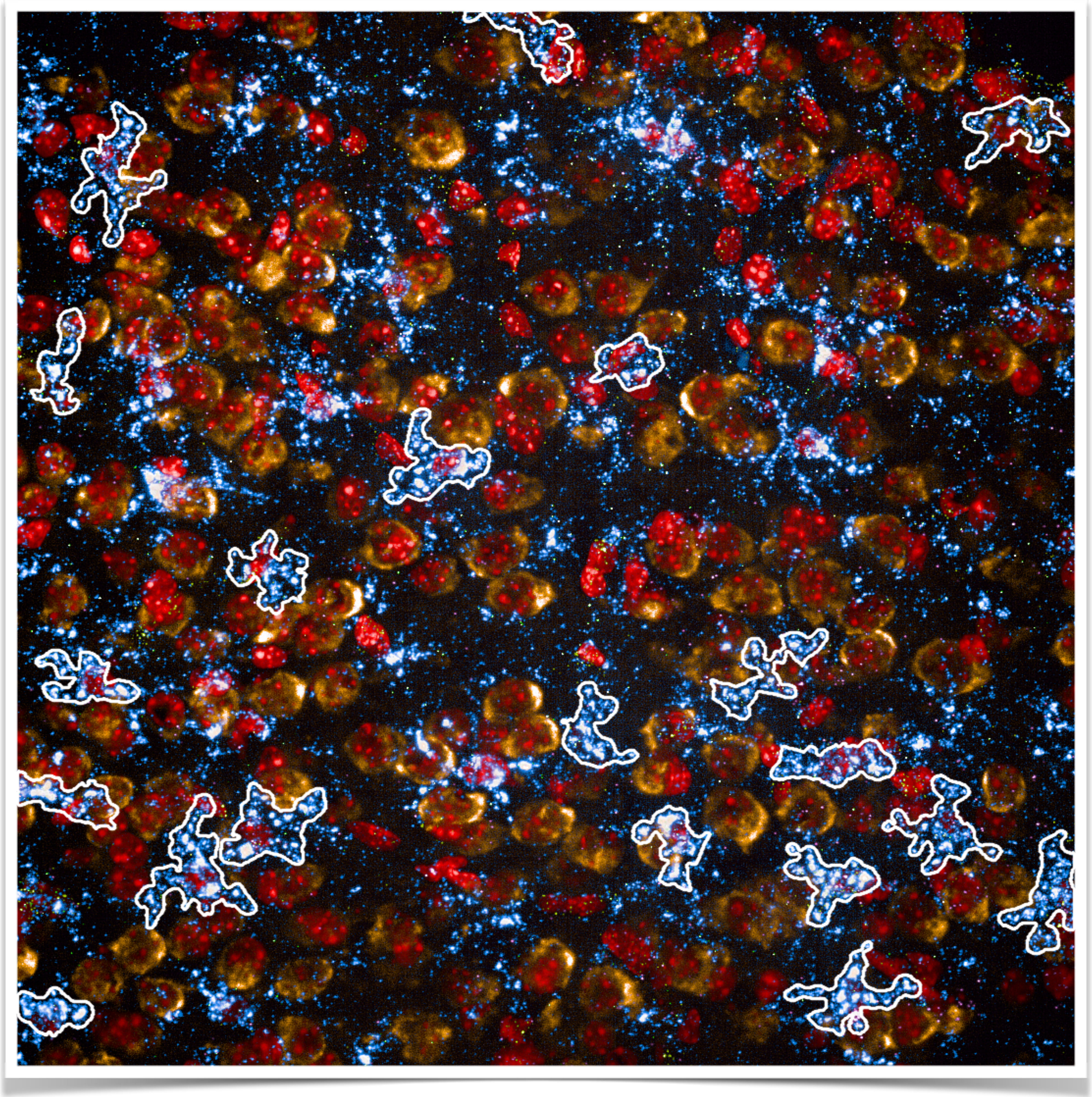


Current semi-manual pipeline

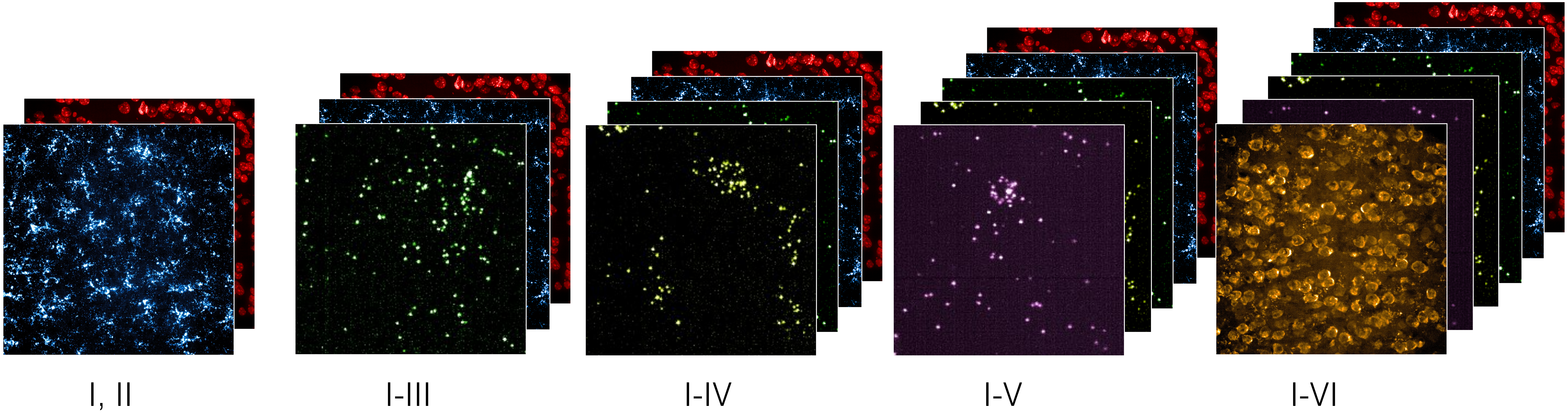
All astrocytes have…

1. One nucleus (in **red**)
2. **Blue body**
3. No **neuron** marker
4. Other markers (**#1**, **#2**, **#3**) moderately present

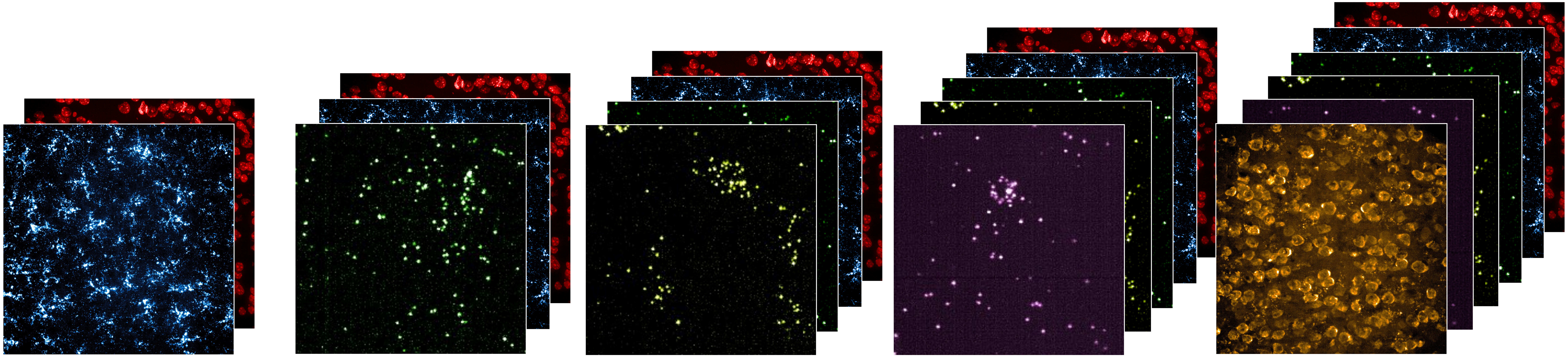
We attempted to **automate** the detection



Generated 5 datasets by combining original channels:

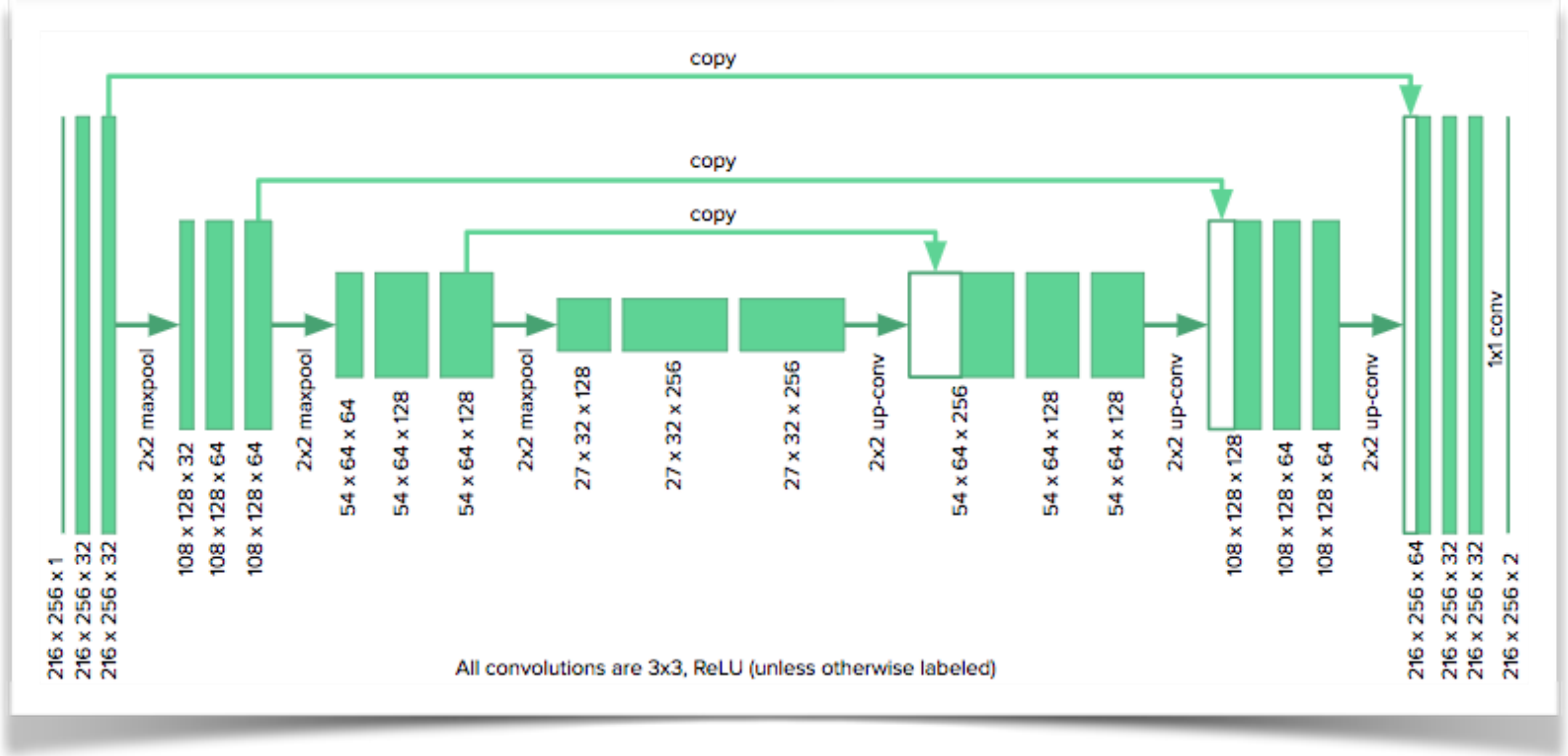


Generated 5 datasets by combining original channels:

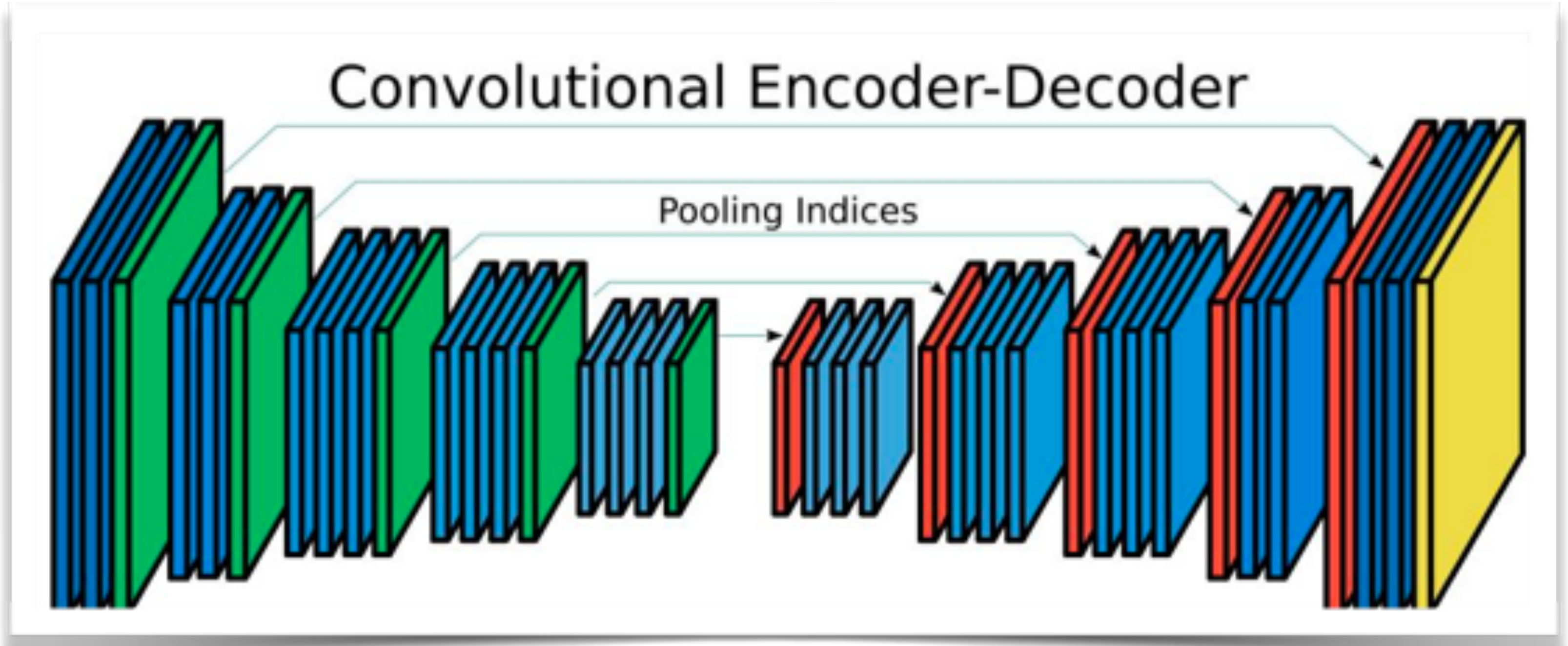


I, II　　I-III　　I-IV　　I-V　　I-VI

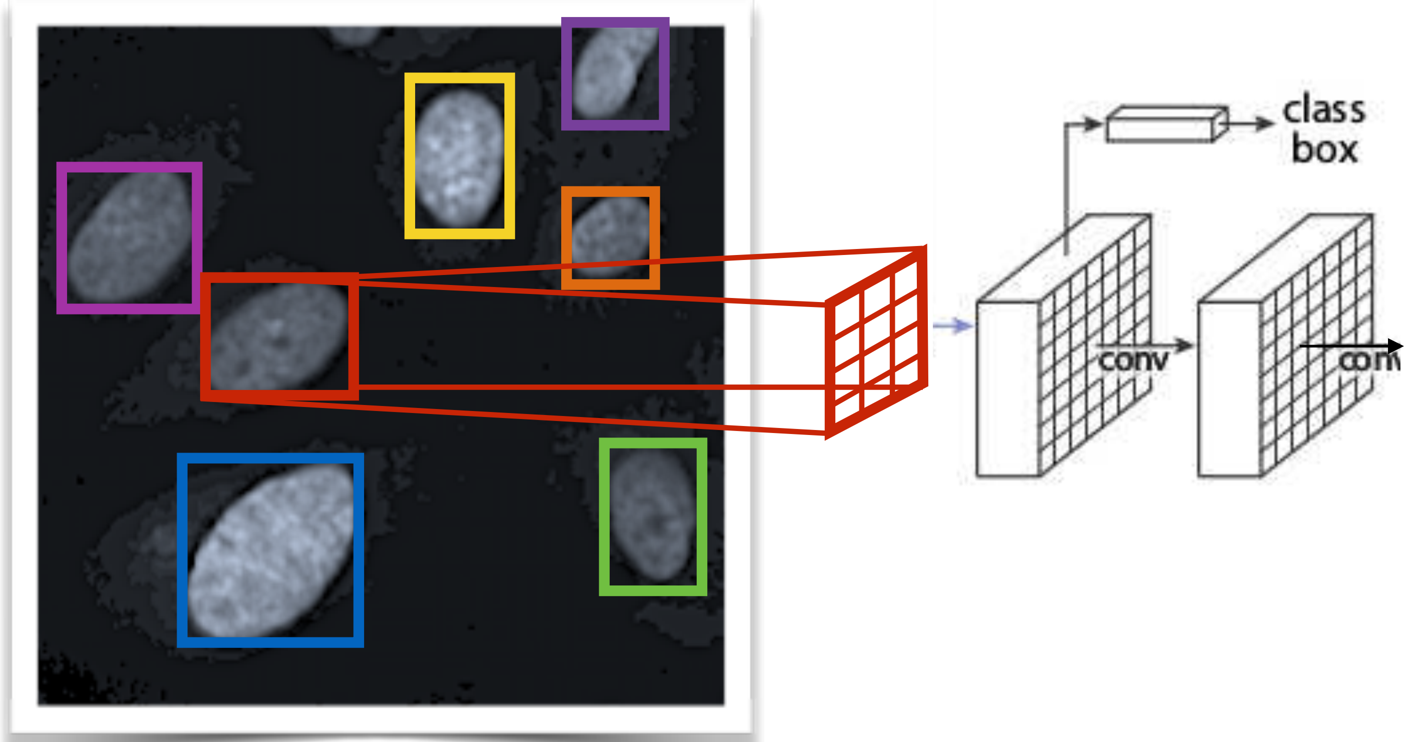
We tried several architectures that I have mentioned before



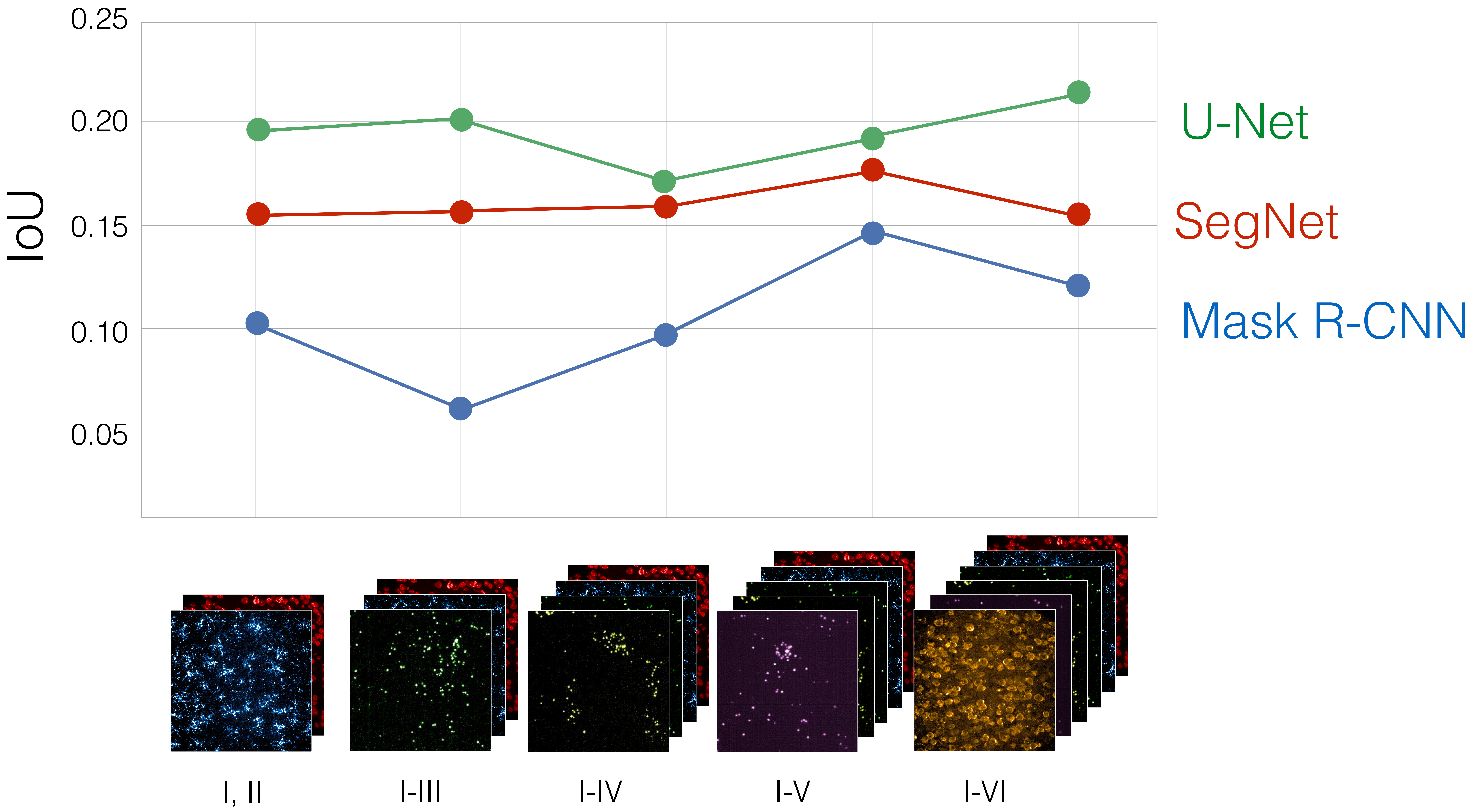
U-Net (O. Ronneberger et al.)



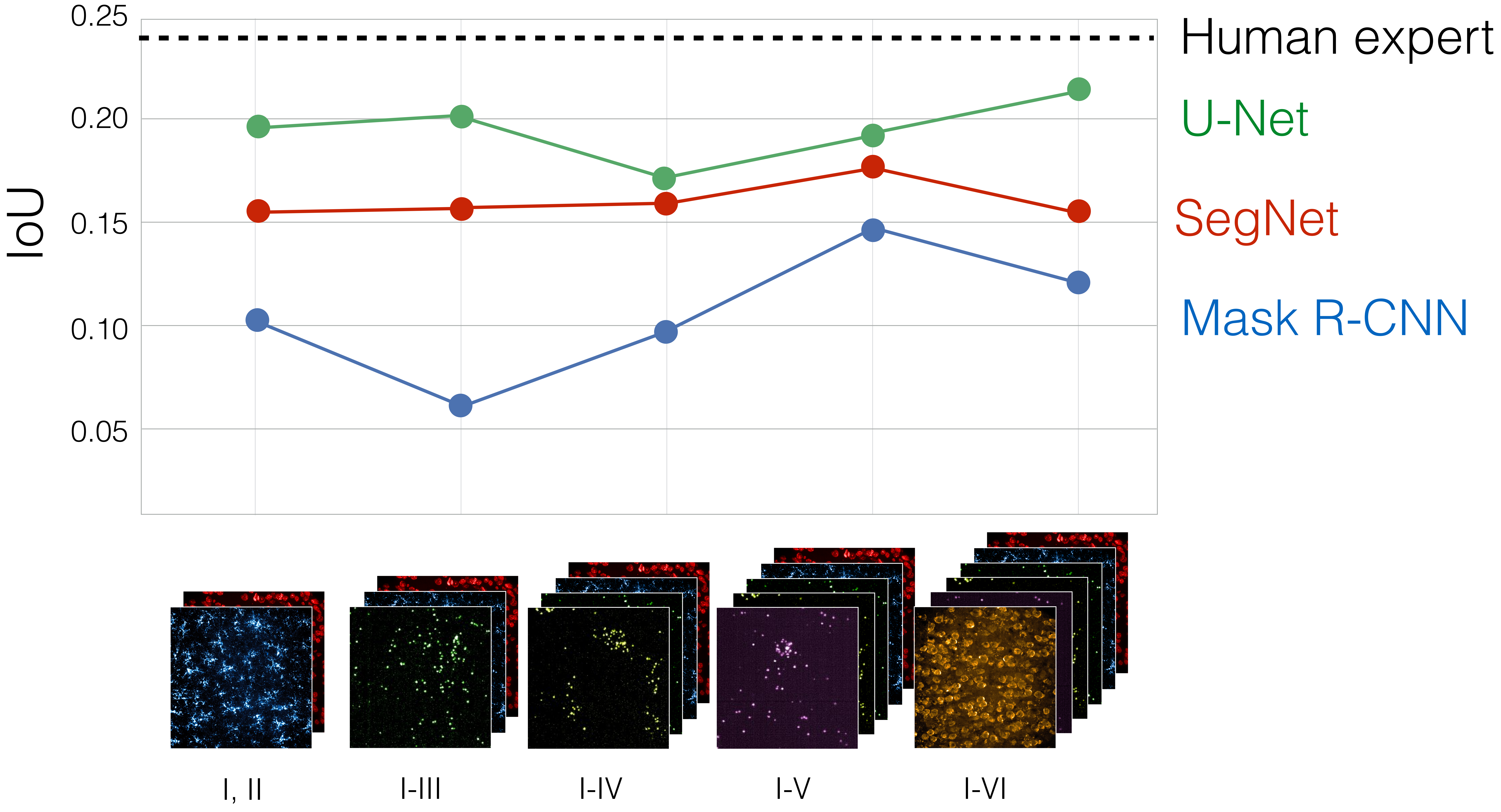
SegNet (V. Badrinarayanan et al.)



Mask R-CNN (K. He et al.)



0.25
0.20
0.15
0.10
0.05
IoU
U-Net
SegNet
Mask R-CNN
I, II
I-III
I-IV
I-V
I-VI


IoU
0.25
0.20
0.15
0.10
0.05
Human expert
U-Net
SegNet
Mask R-CNN
I, II
I-III
I-IV
I-V
I-VI


Original image
Mask R-CNN
U-Net
SegNet
Ground Truth - **white area**

UNIVERSITY OF TARTU
Institute of Computer Science

Astrocytes Segmentation in Brain Microscopy Images

Bohdan Petryshak[*,1,2], Oleksandr Pryhoda[*,1,2], Leopold Parts[2,3], Omer Bayraktar[3], Dmytro Fishman[2,4]

[1]Faculty of Applied Sciences , Ukrainian Catholic University, [2]Institute of Computer Science, University of Tartu, [3]The Wellcome Sanger Institute, [4]Quretec Ltd.

* These authors contributed equally to this work

Abstract

Brain is a vital part of all higher organisms, but mechanisms behind its functions remain poorly understood. It is a structured organ, with a variety of cell types ranging from neurons to immune cells non-uniformly distributed across space.

Localizing the different cell types and quantifying their gene expression patterns from microscopy images is the principal way to gain novel insights into the organization and inner workings of the brain. In particular, the challenging morphology of astrocytes makes them one of the most complex types of cells to identify. While a lot of manual work is currently needed to reliably segment astrocytes from images, the scale of data produced with modern microscopes renders this approach impractical. Here, we present an automated segmentation approach for brain microscopy images using deep learning.

We implemented and compared the performance of U-Net [1], Mask R-CNN [2] and SegNet [3] models on RNA fluorescence *in situ* hybridization images from mouse brain slices. The employed architectures are capable of reliably detecting and segmenting astrocytes, but have a high false positive rate, likely due to limitations of the training data.

Introduction

Semantic segmentation is one of the key problems in Computer Vision area. Identifying the different types of cells from brain images helps biologist to understand its inner mechanisms. Currently, it's done either via manual examination or semi-automated approaches that consume a lot of experts time and effort. Astrocytes are one of the most challenging types of the brain cells to segment due to their complex and heterogeneous structure. Here we present a fully-automated pipeline for segmenting astrocytes from microscopy images using SOTA CNN architectures: U-Net [1], Mask R-CNN [2], SegNet [3] (Figure 4).

Research questions are:

- Can Deep Learning help to segment astrocytes?
- Which neural network architecture works the best?
- How reliable produced segmentations are?

Data Description

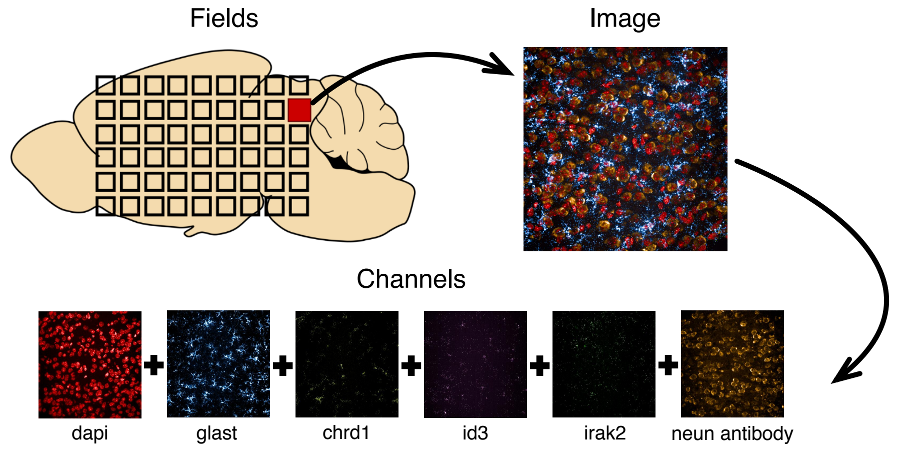


Figure 1: Structure of the data used in the study, one image in the dataset constitutes of 6 channels: dapi, glast, chrd1, id3, irak2 and neun antibody.

Channels description

dapi - marks the nuclei of astrocytes and other cells

glast - marker of astrocytes, to be used for segmentation

chrd1 - this marker is expressed as a spatial gradient across astrocytes, enriched in upper layers

id3 - this marker is expressed as a spatial gradient across astrocytes, enriched in deep layers and the most superficial layer.

irak2 - this marker is expressed in most astrocytes, slight spatial enrichment in upper layer astrocytes

neun antibody - marker of neurons

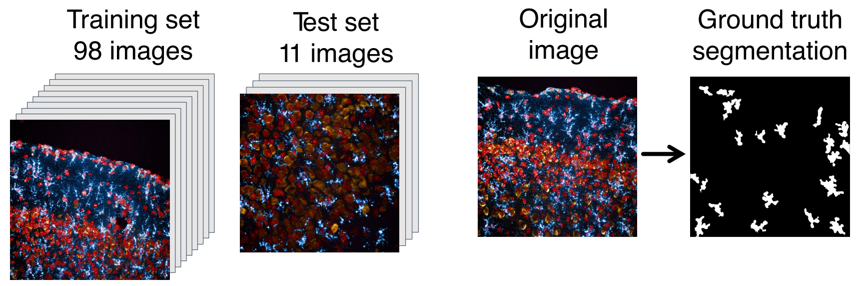


Figure 2: Example of train and test sets images with ground truth segmentation.

Methods

Data preprocessing

Different datasets were created by sequentially adding all 6 channels, starting with channels 1 and 2 and merging them into RGB using Harmony software (Figure 3).

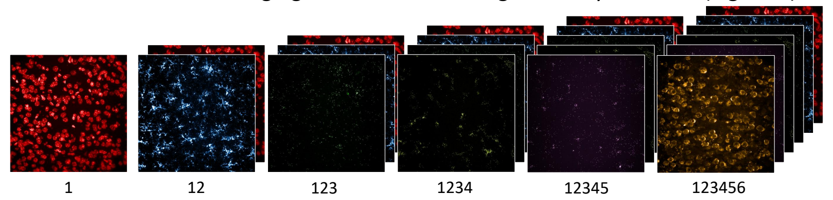


Figure 3: Datasets used for training. Number of images in each dataset corresponds to the number of merged channels. The datasets were named based on indexes of the channels that were used to generate them.

Models description

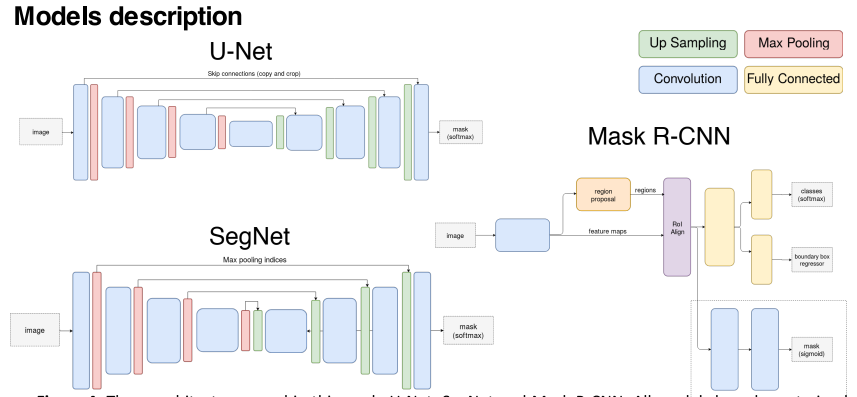


Figure 4: Three architectures used in this work, U-Net, SegNet and Mask R-CNN. All models have been trained using the same hyper-parameters, e.g. learning rate, regularization strength and optimization algorithm.

Results

Performance evaluation

Mean average precision at different intersection over union (IoU) thesholds metric was computed to assess the performance of used models

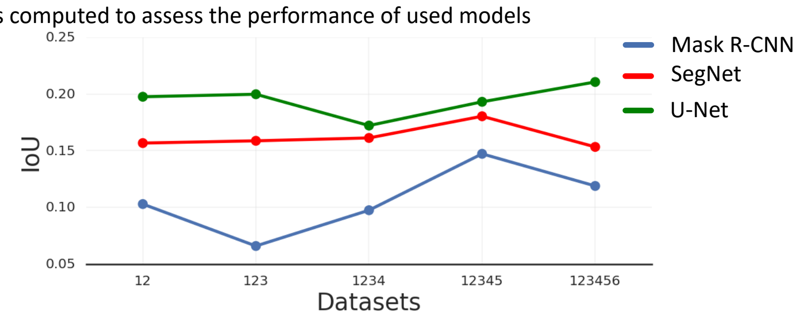


Figure 5: IoU score for U-Net, SegNet and Mask R-CNN for each dataset.

Segmentation example

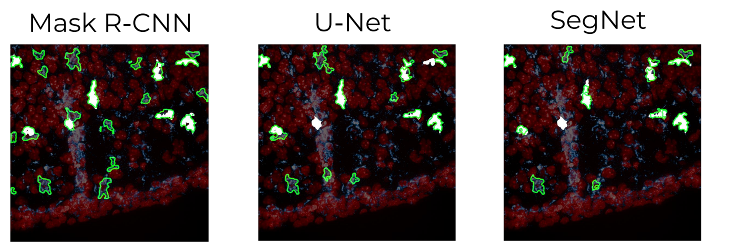


Figure 6: Segmentation example of U-Net, SegNet and Mask R-CNN. White color – ground truth segmentation. Green color – neural networks prediction. All models produce high number of false positive segmentations.

Interpretation of the results

Our experiments show that adding more channels into training data does not seem to significantly influence the model performance (Figure 5). Overall, U-Net has reached 0.21 IoU, SegNet – 0.18 and Mask R-CNN – 0.147.In general results proved CNN are able to segment the astrocytes in the microscopy images. However, all models produce high number of false positives that can be a result of insufficient quality and quantity of ground truth data.

Acknowledgements

University of Tartu ASTRA Project PER ASPERA Doctoral School of Information and Communication Technologies and High Performance Computing Center of the Institute of Computer Science at the University of Tartu.

References

[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox: U-Net. Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597v1 [cs.CV] 18 May 2015

[2] Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick. Mask R-CNN. arXiv:1703.06870v3 [cs.CV] 24 Jan 2018

[3] Vijay Badrinarayanan, Alex Kendall , Roberto Cipolla, Senior Member, IEEE. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. arXiv:1511.00561v3 [cs.CV] 10 Oct 2016

wellcome sanger institute

UCU. APPLIED SCIENCES FACULTY

UNIVERSITY of TARTU
Institute of Computer Science

Astrocytes Segmentation in Brain Microscopy Images

Bohdan Petryshak[*,1,2], Oleksandr Pryhoda[*,1,2], Leopold Parts[2,3], Omer Bayraktar[3], Dmytro Fishman[2,4]

[1]Faculty of Applied Sciences , Ukrainian Catholic University, [2]Institute of Computer Science, University of Tartu, [3]The Wellcome Sanger Institute, [4]Quretec Ltd.

* These authors contributed equally to this work

Bohdan Petryshak[*,1,2], Oleksandr Pryhoda[*,1,2], Leopold Parts[2,3], Omer Bayraktar[3], Dmytro Fishman[2,4]

and inner workings of the brain. In particular, the challenging morphology of astrocytes makes them one of the most complex types of cells to identify. While a lot of manual work is currently needed to reliably segment astrocytes from images, the scale of data produced with modern microscopes renders this approach impractical. Here, we present an automated segmentation approach for brain microscopy images using deep learning.

We implemented and compared the performance of U-Net [1], Mask R-CNN [2] and SegNet [3] models on RNA fluorescence *in situ* hybridization images from mouse brain slices. The employed architectures are capable of reliably detecting and segmenting astrocytes, but have a high false positive rate, likely due to limitations of the training data.

Introduction

Semantic segmentation is one of the key problems in Computer Vision area. Identifying the different types of cells from brain images helps biologist to understand its inner mechanisms. Currently, it's done either via manual examination or semi-automated approaches that consume a lot of experts time and effort. Astrocytes are one of the most challenging types of the brain cells to segment due to their complex and heterogeneous structure. Here we present a fully-automated pipeline for segmenting astrocytes from microscopy images using SOTA CNN architectures: U-Net [1], Mask R-CNN [2], SegNet [3] (Figure 4).

Research questions are:

- Can Deep Learning help to segment astrocytes?
- Which neural network architecture works the best?
- How reliable produced segmentations are?

Data Description

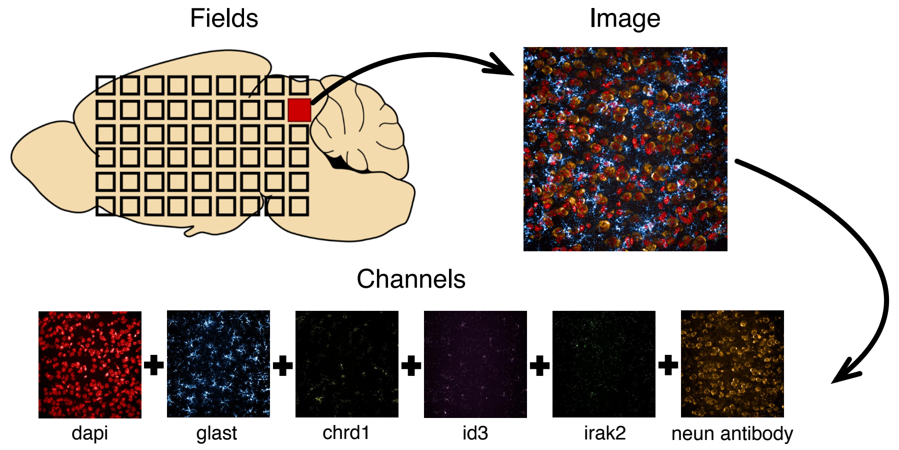


Figure 1: Structure of the data used in the study, one image in the dataset constitutes of 6 channels: dapi, glast, chrd1, id3, irak2 and neun antibody.

Channels description

dapi - marks the nuclei of astrocytes and other cells

glast - marker of astrocytes, to be used for segmentation

chrd1 - this marker is expressed as a spatial gradient across astrocytes, enriched in upper layers

id3 - this marker is expressed as a spatial gradient across astrocytes, enriched in deep layers and the most superficial layer.

irak2 - this marker is expressed in most astrocytes, slight spatial enrichment in upper layer astrocytes

neun antibody - marker of neurons

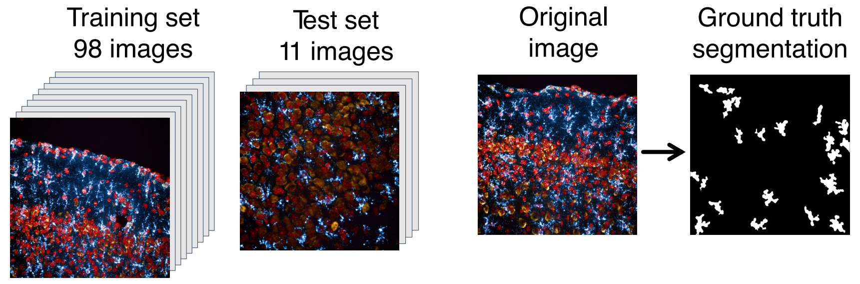


Figure 2: Example of train and test sets images with ground truth segmentation.

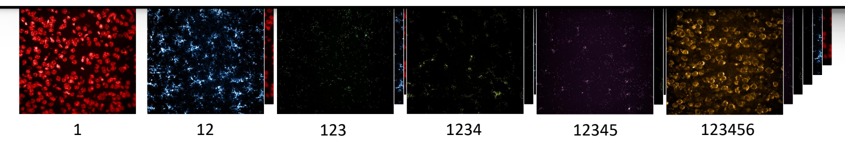


Figure 3: Datasets used for training. Number of images in each dataset corresponds to the number of merged channels. The datasets were named based on indexes of the channels that were used to generate them.

Models description

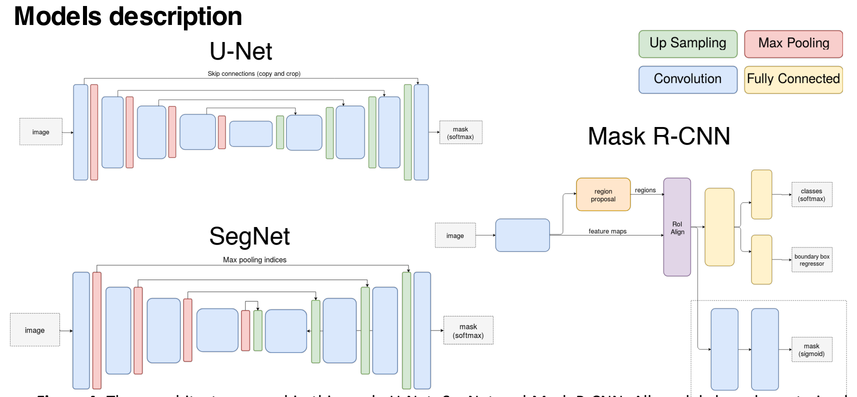


Figure 4: Three architectures used in this work, U-Net, SegNet and Mask R-CNN. All models have been trained using the same hyper-parameters, e.g. learning rate, regularization strength and optimization algorithm.

Results

Performance evaluation

Mean average precision at different intersection over union (IoU) thesholds metric was computed to assess the performance of used models

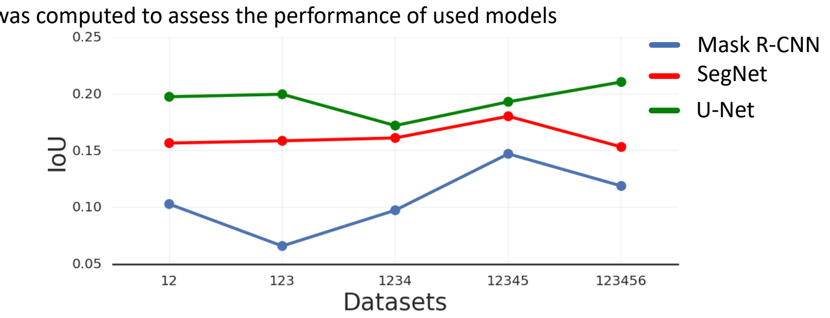


Figure 5: IoU score for U-Net, SegNet and Mask R-CNN for each dataset.

Segmentation example

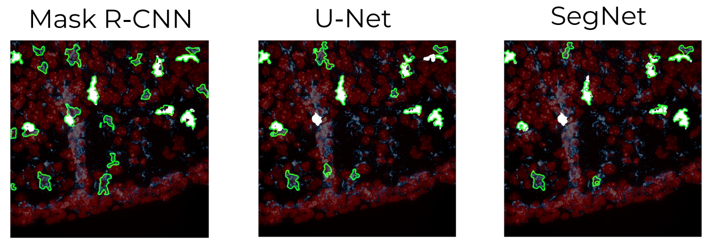


Figure 6: Segmentation example of U-Net, SegNet and Mask R-CNN. White color – ground truth segmentation. Green color – neural networks prediction. All models produce high number of false positive segmentations.

Interpretation of the results

Our experiments show that adding more channels into training data does not seem to significantly influence the model performance (Figure 5). Overall, U-Net has reached 0.21 IoU, SegNet – 0.18 and Mask R-CNN – 0.147.In general results proved CNN are able to segment the astrocytes in the microscopy images. However, all models produce high number of false positives that can be a result of insufficient quality and quantity of ground truth data.

Acknowledgments

University of Tartu ASTRA Project PER ASPERA Doctoral School of Information and Communication Technologies and High Performance Computing Center of the Institute of Computer Science at the University of Tartu.

References

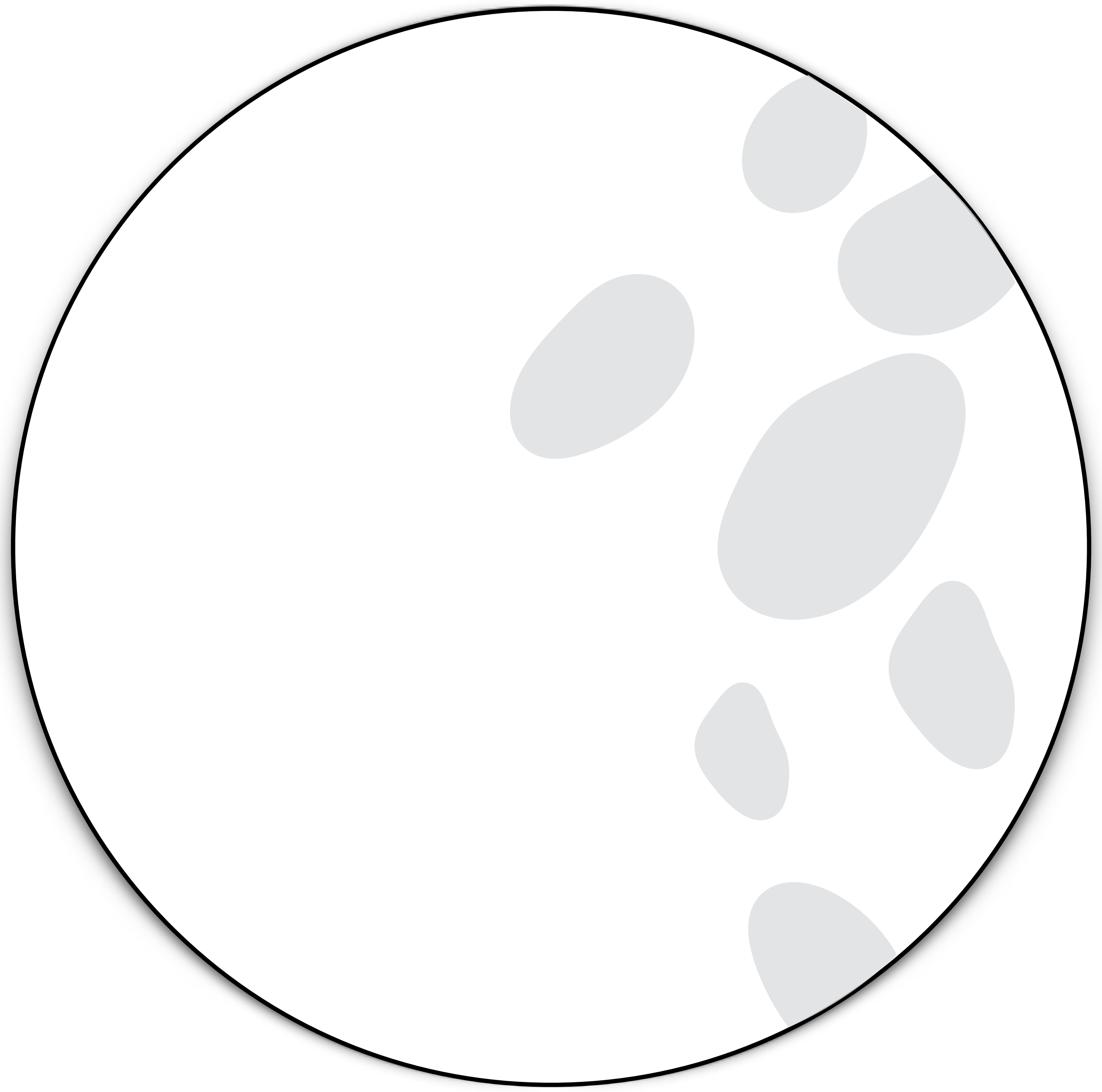
[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox: U-Net. Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597v1 [cs.CV] 18 May 2015
[2] Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick. Mask R-CNN. arXiv:1703.06870v3 [cs.CV] 24 Jan 2018
[3] Vijay Badrinarayanan, Alex Kendall , Roberto Cipolla, Senior Member, IEEE. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. arXiv:1511.00561v3 [cs.CV] 10 Oct 2016

wellcome sanger institute

UCU. APPLIED SCIENCES FACULTY

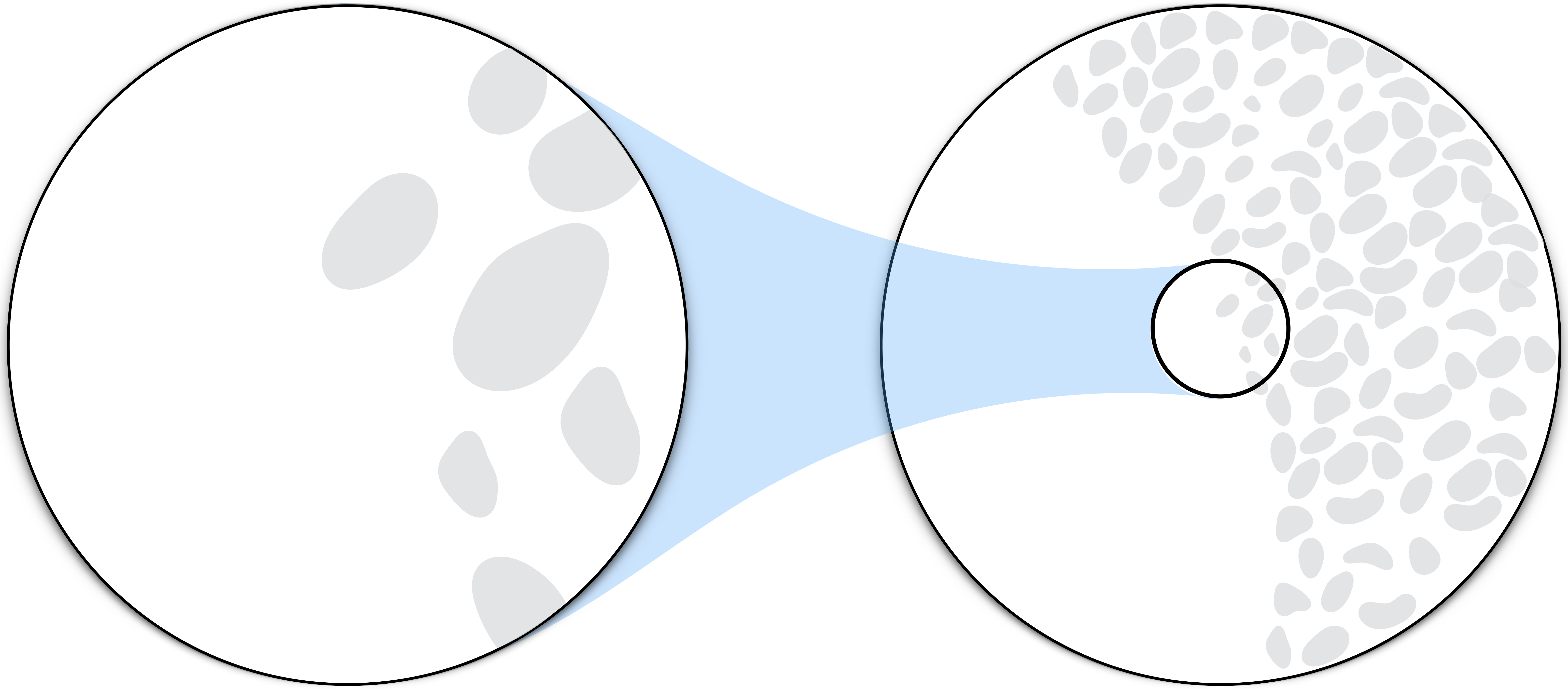
Let’s zoom out a bit…

Let’s zoom out a bit…



Cells

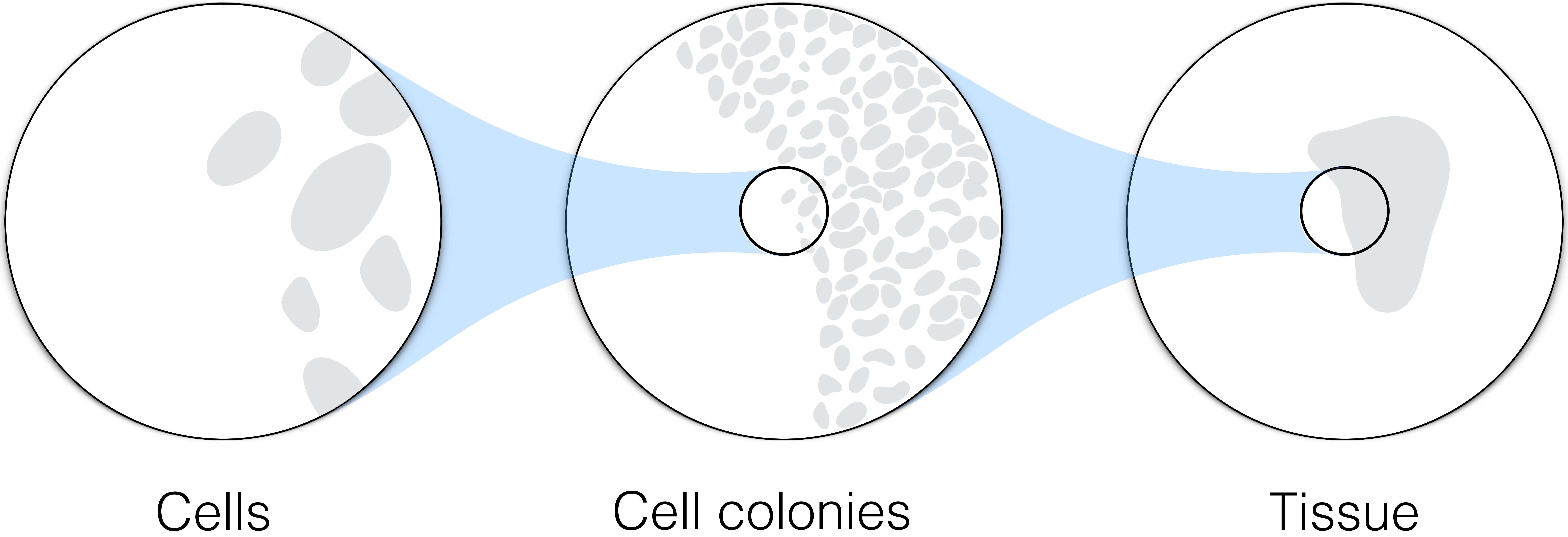
Let’s zoom out a bit…



Cells

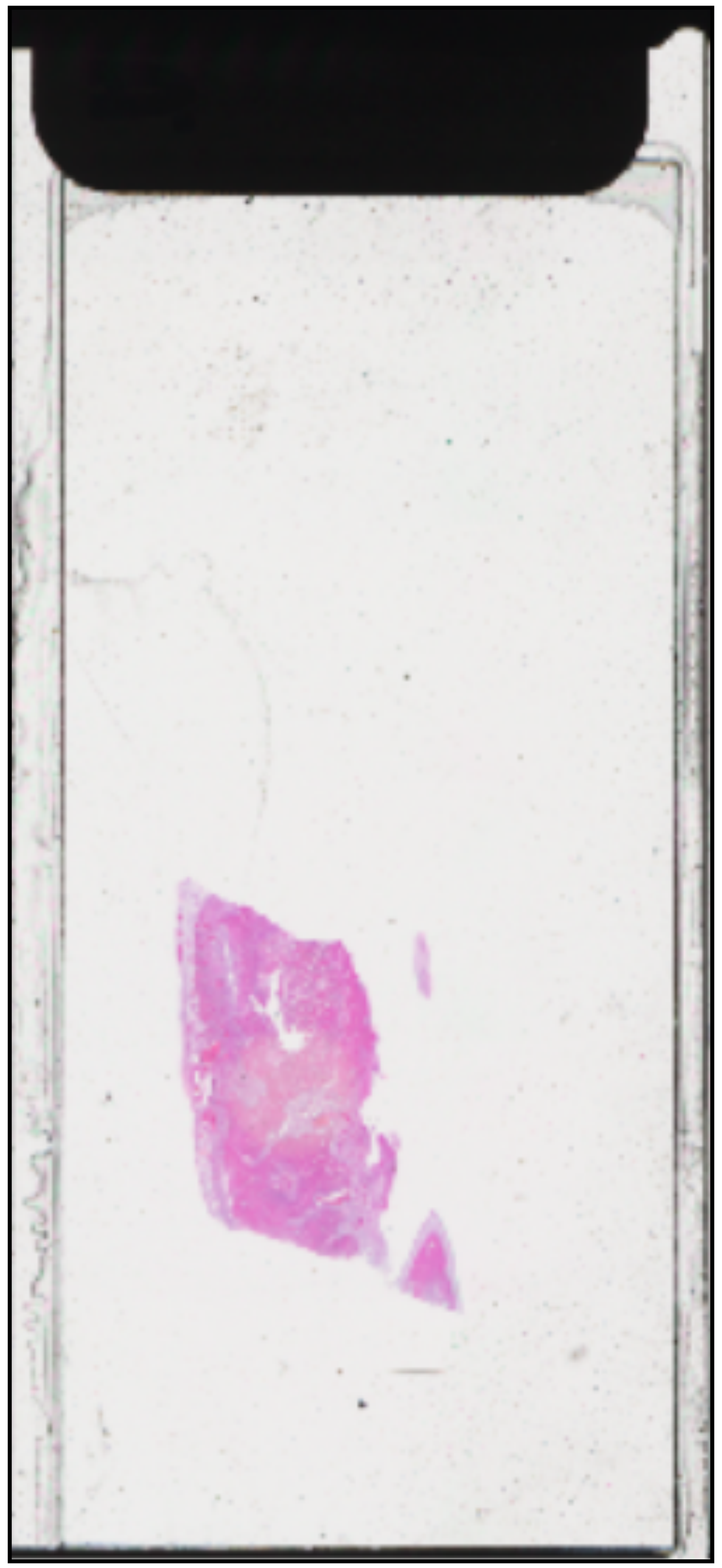
Cell colonies

Let’s zoom out a bit…

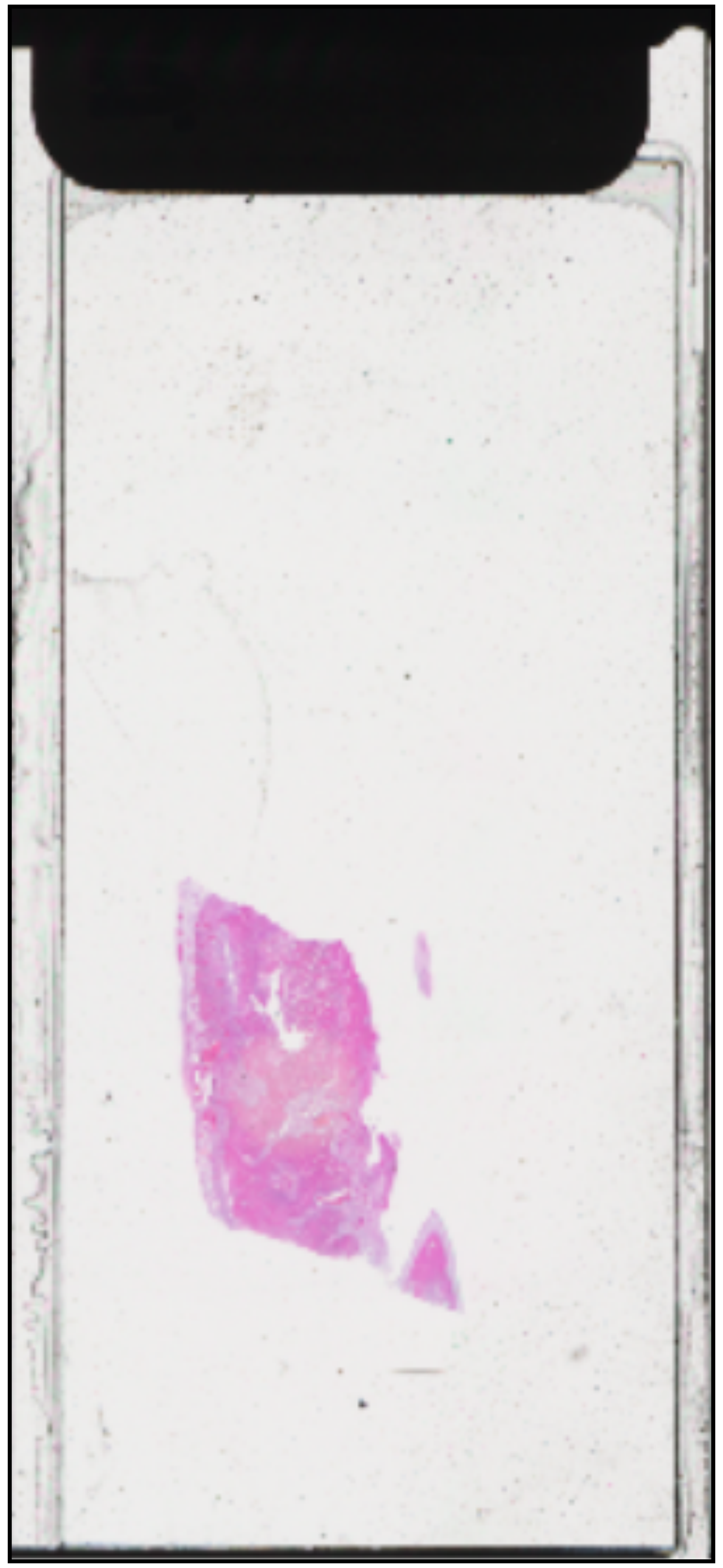


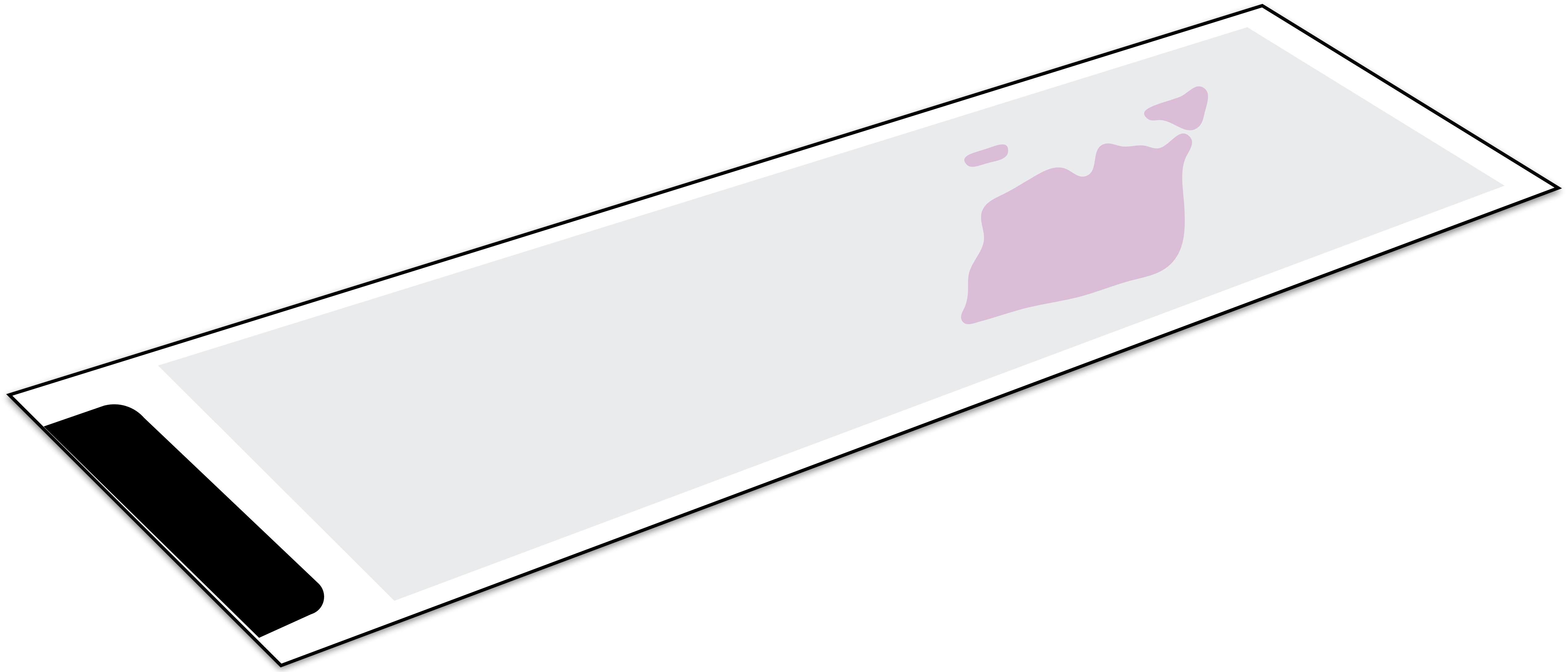
Tissue segmentation

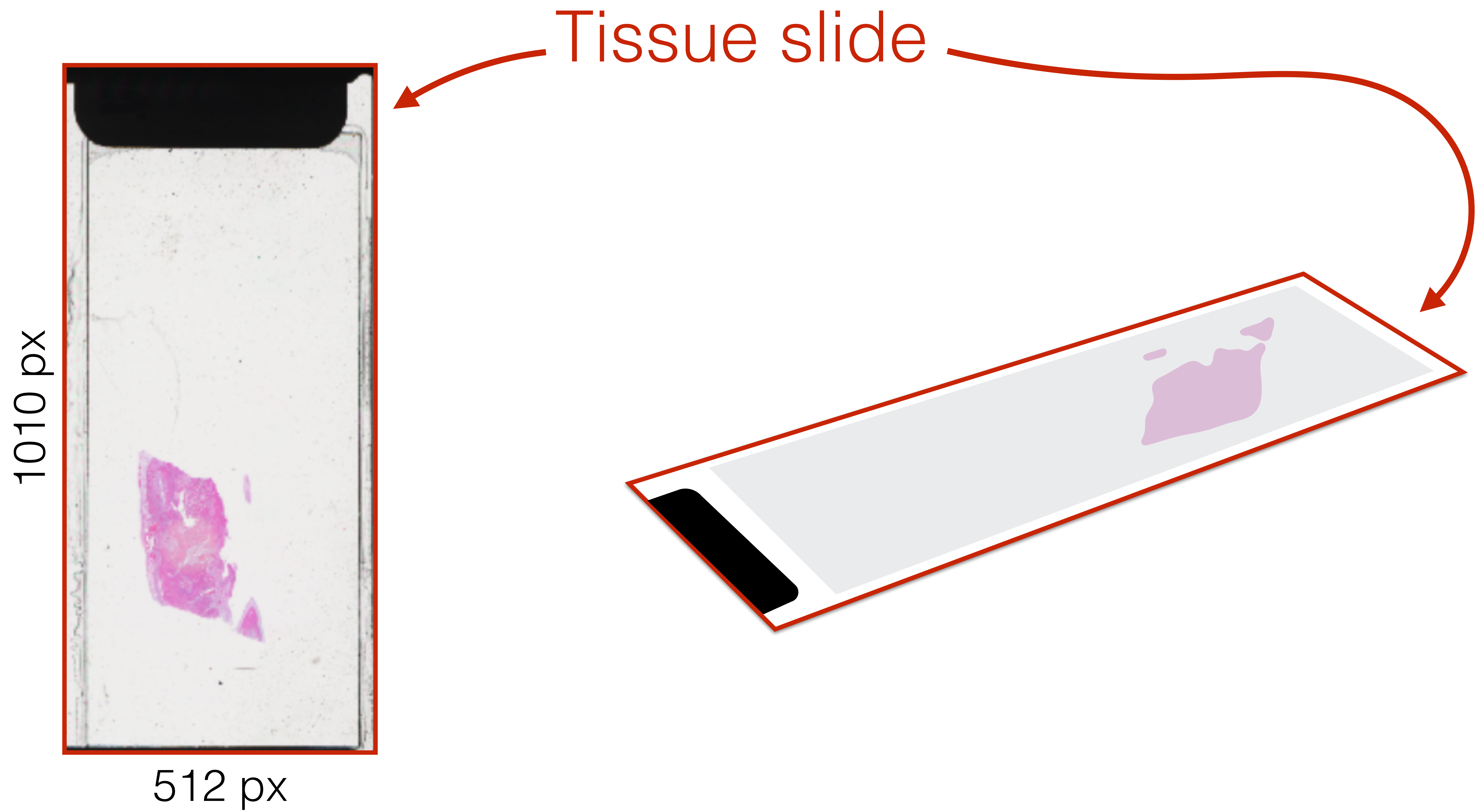
1010 px



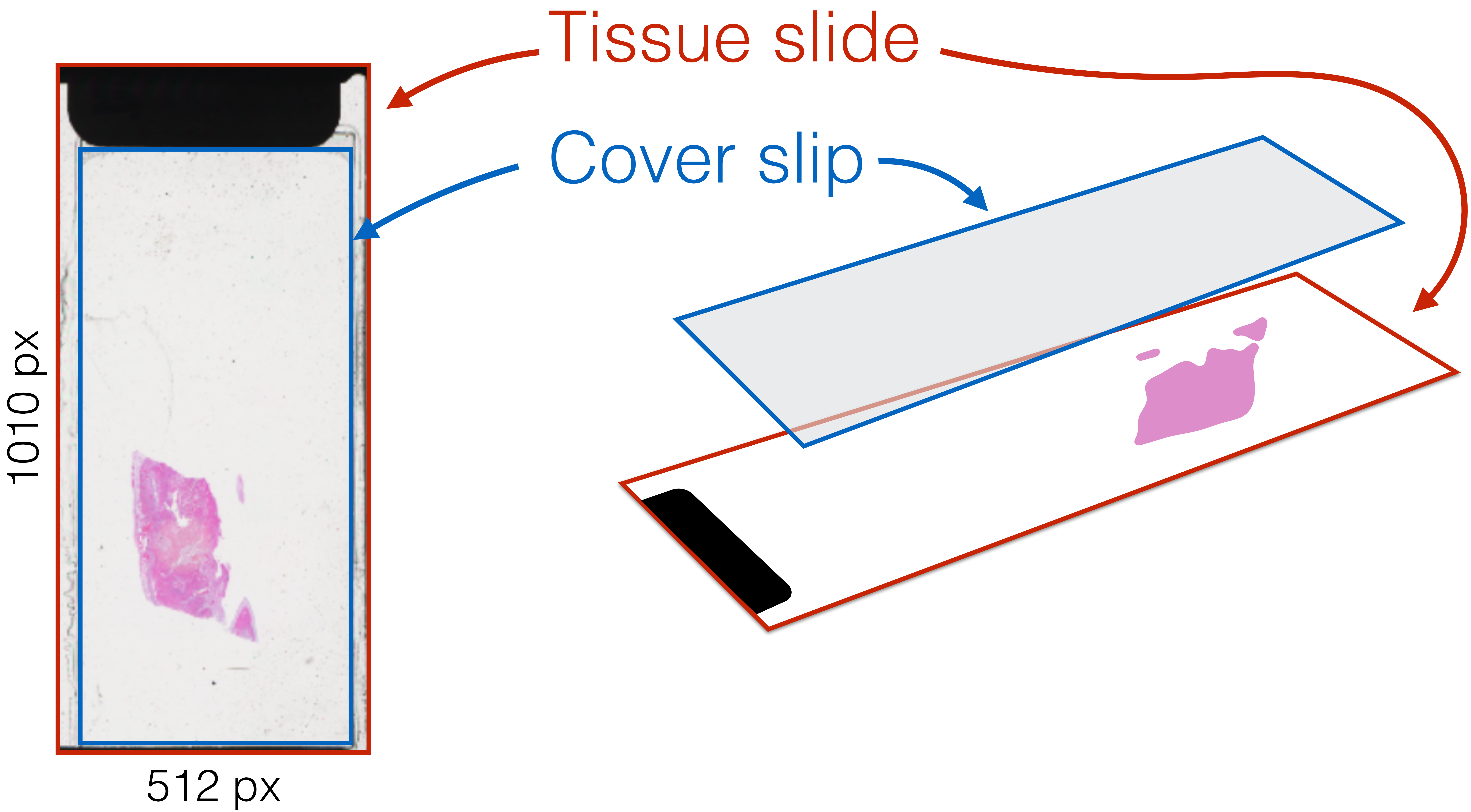
512 px


1010 px
512 px

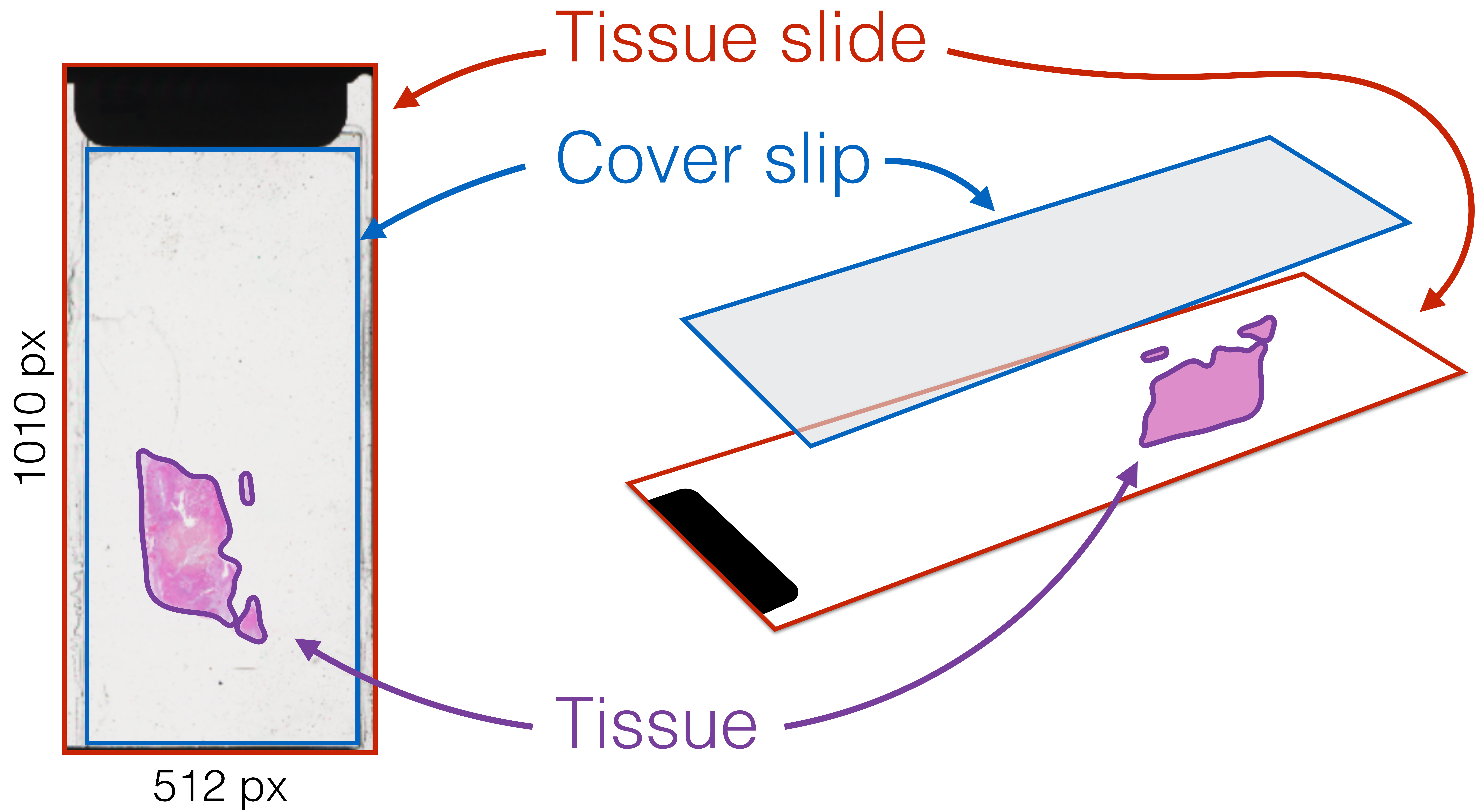




Tissue slide
1010 px
512 px



Tissue slide
Cover slip
1010 px
512 px


Tissue slide
Cover slip
1010 px
512 px
Tissue


APEX BOND
APEX BOND
Thermo Shandon
APEX BOND
APEX BOND
APEX BOND
APEX BOND
APEX BOND
APEX BOND
APEX BOND
APEX BOND

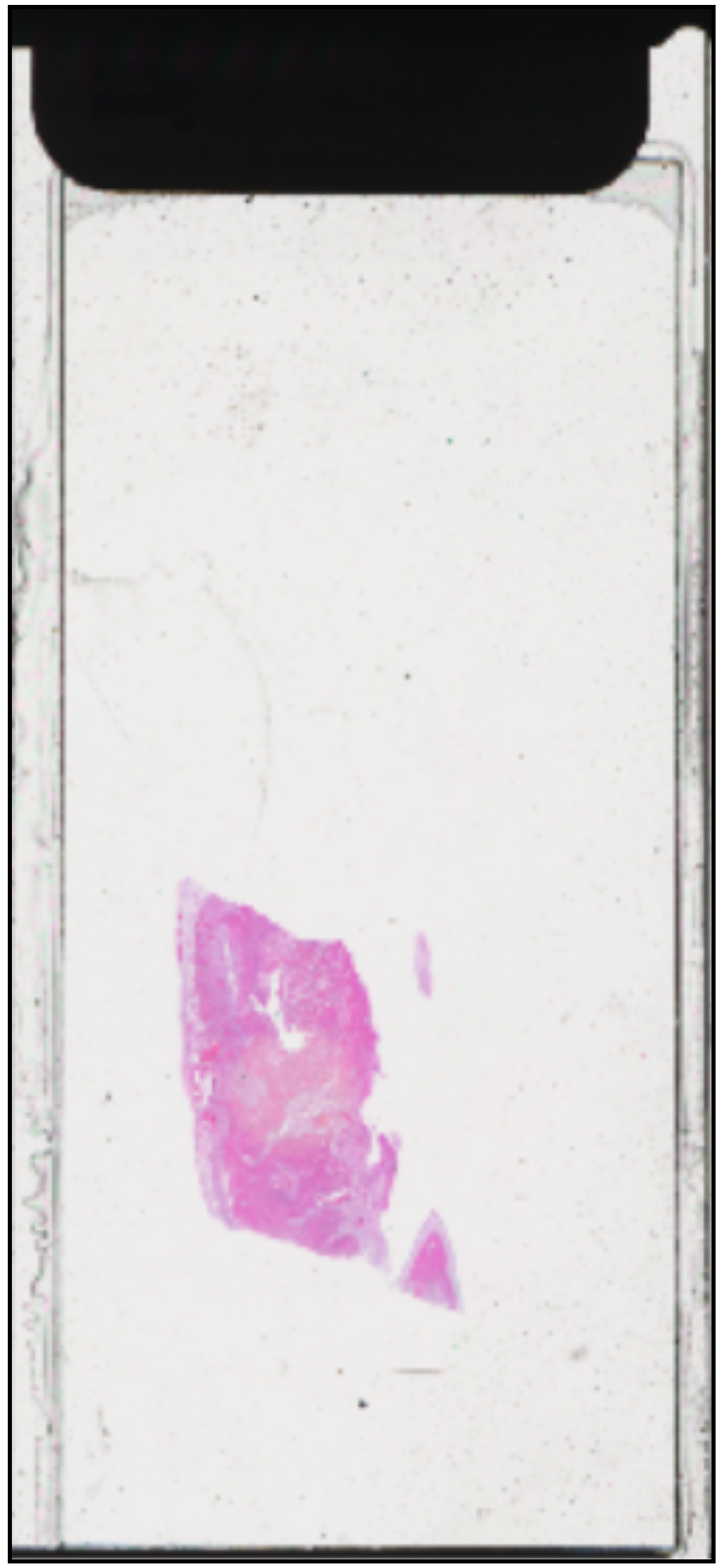
Cover slips are detectable by classical methods

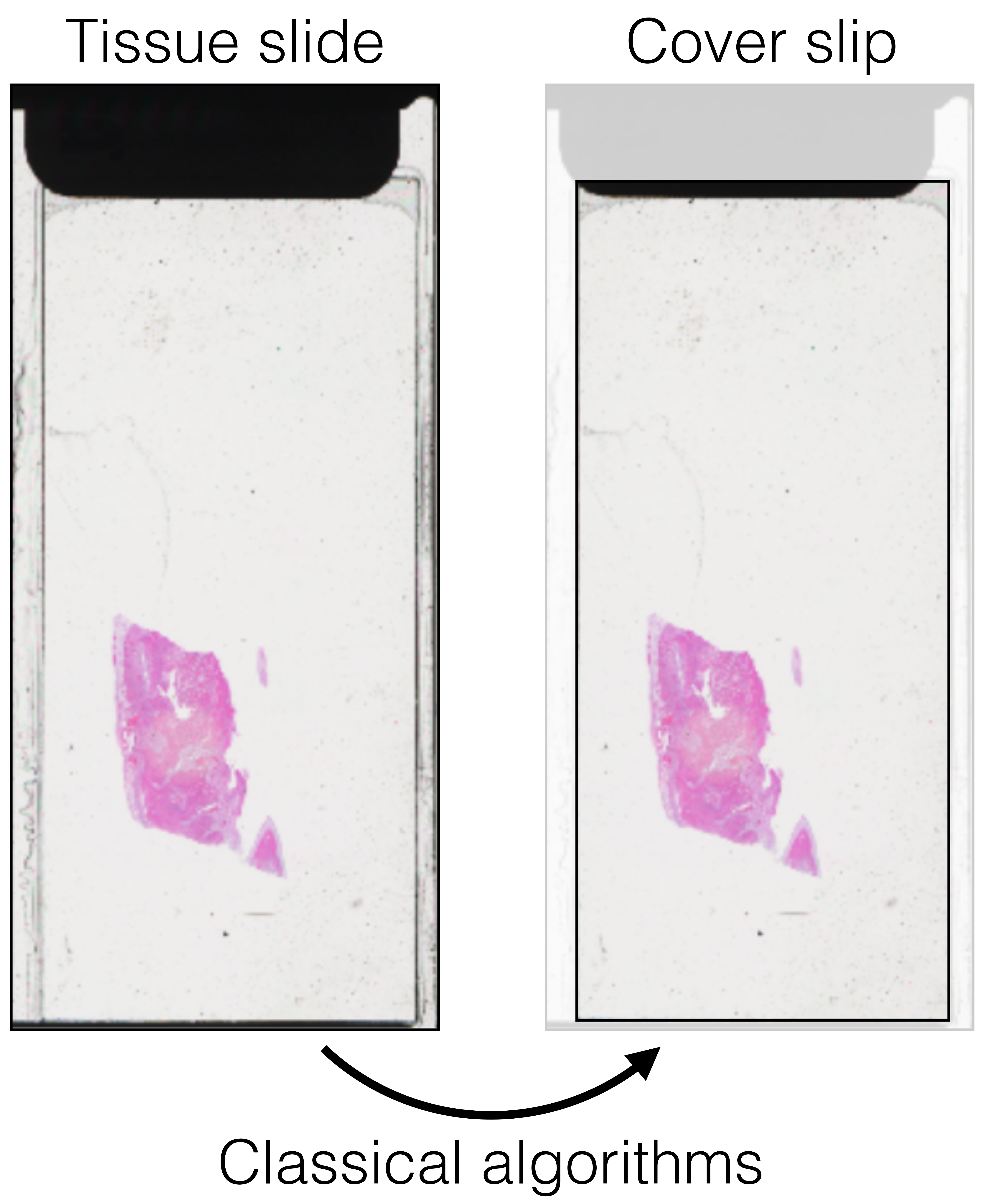
Cover slips are detectable by classical methods

Cover slips are detectable by classical methods

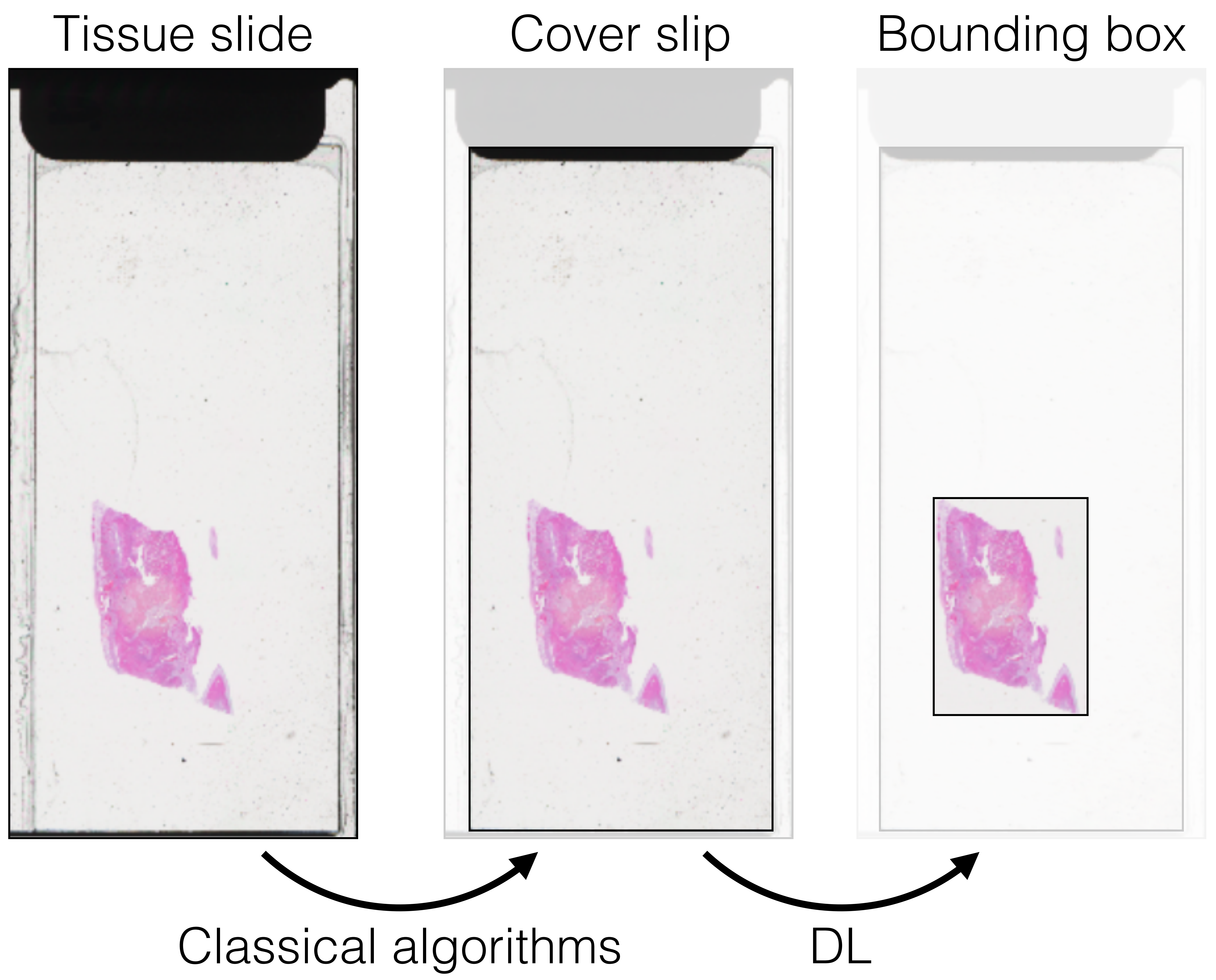
However **tissues** are much harder to identify

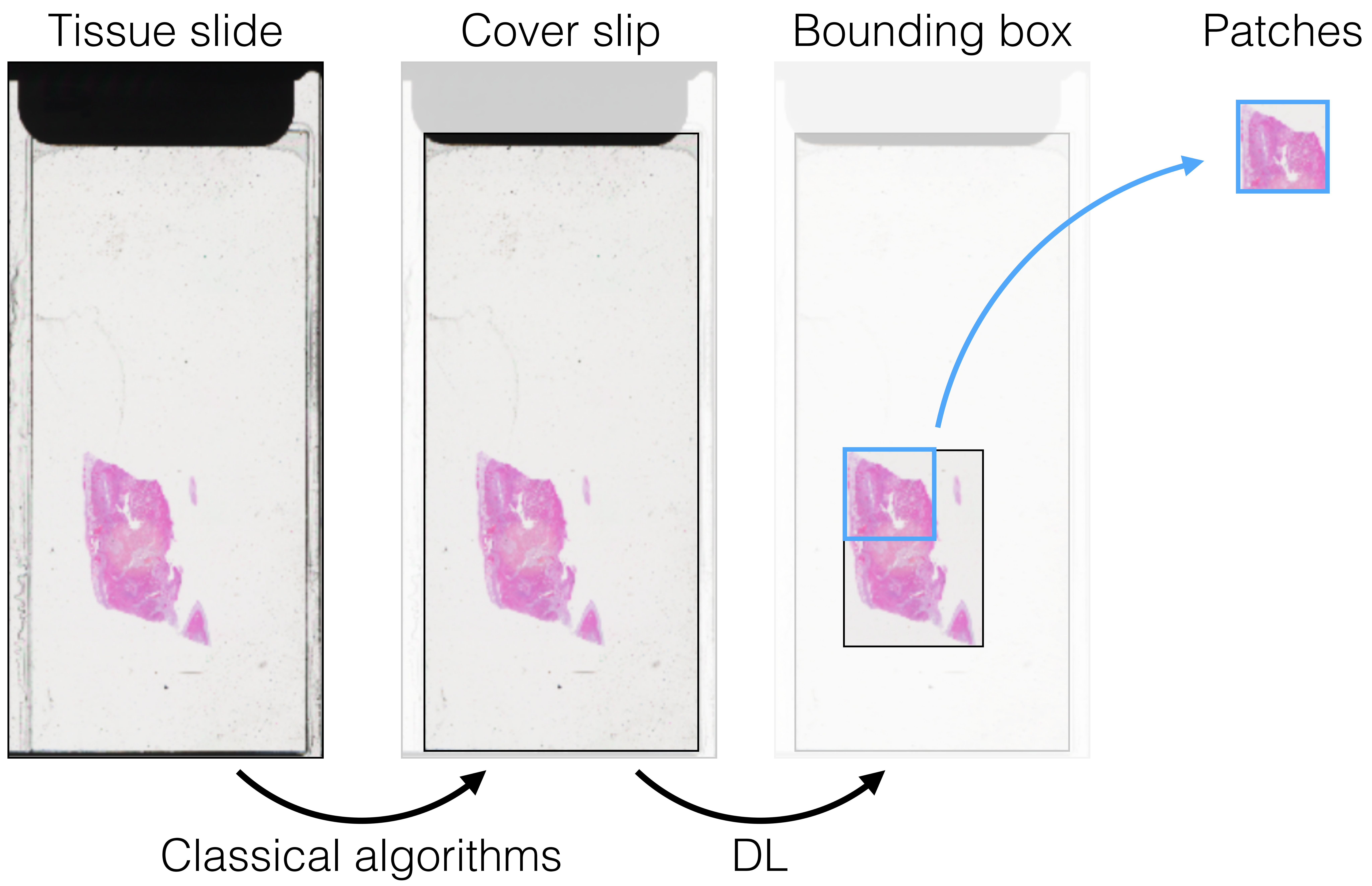
Tissue slide

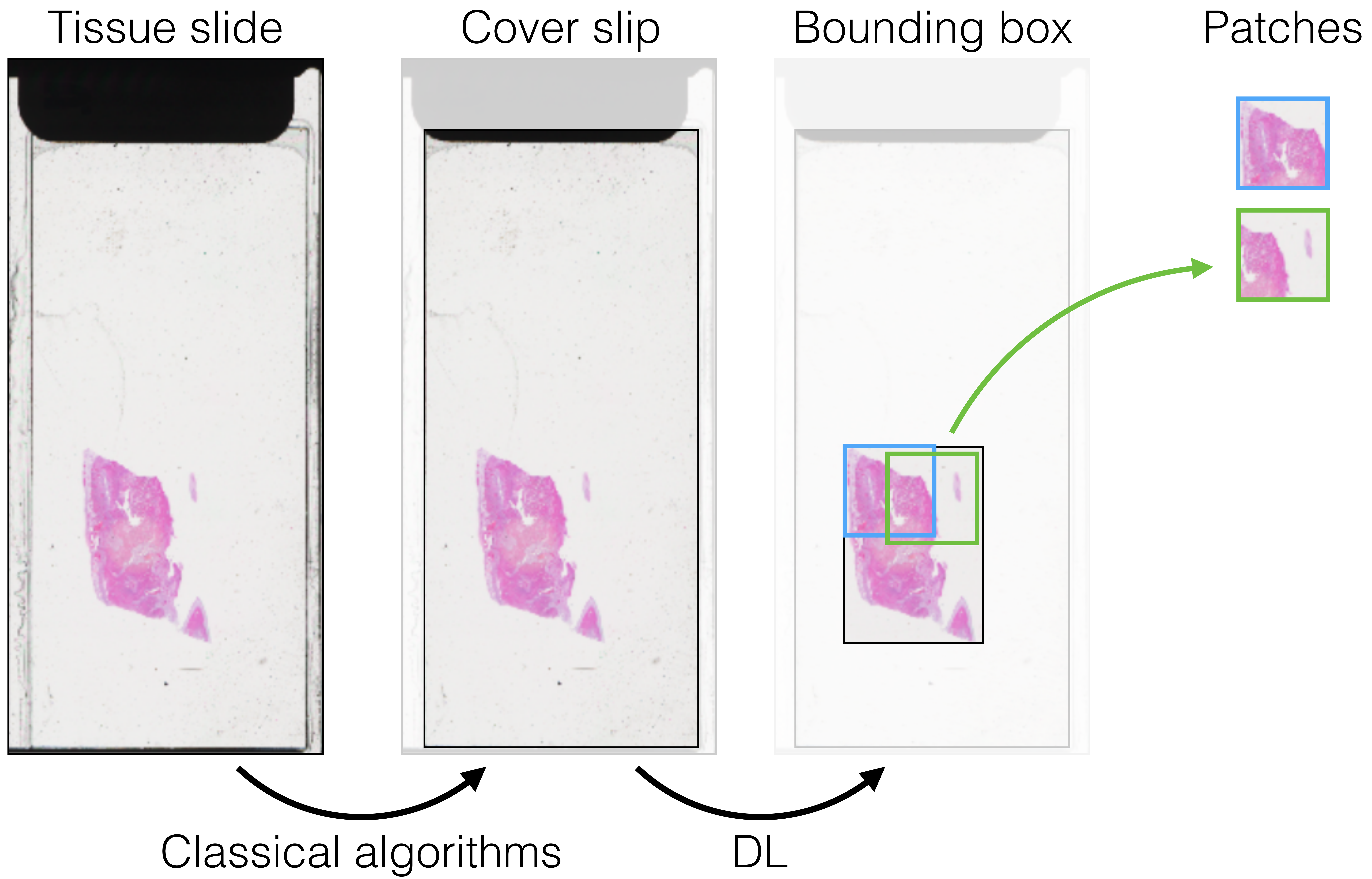


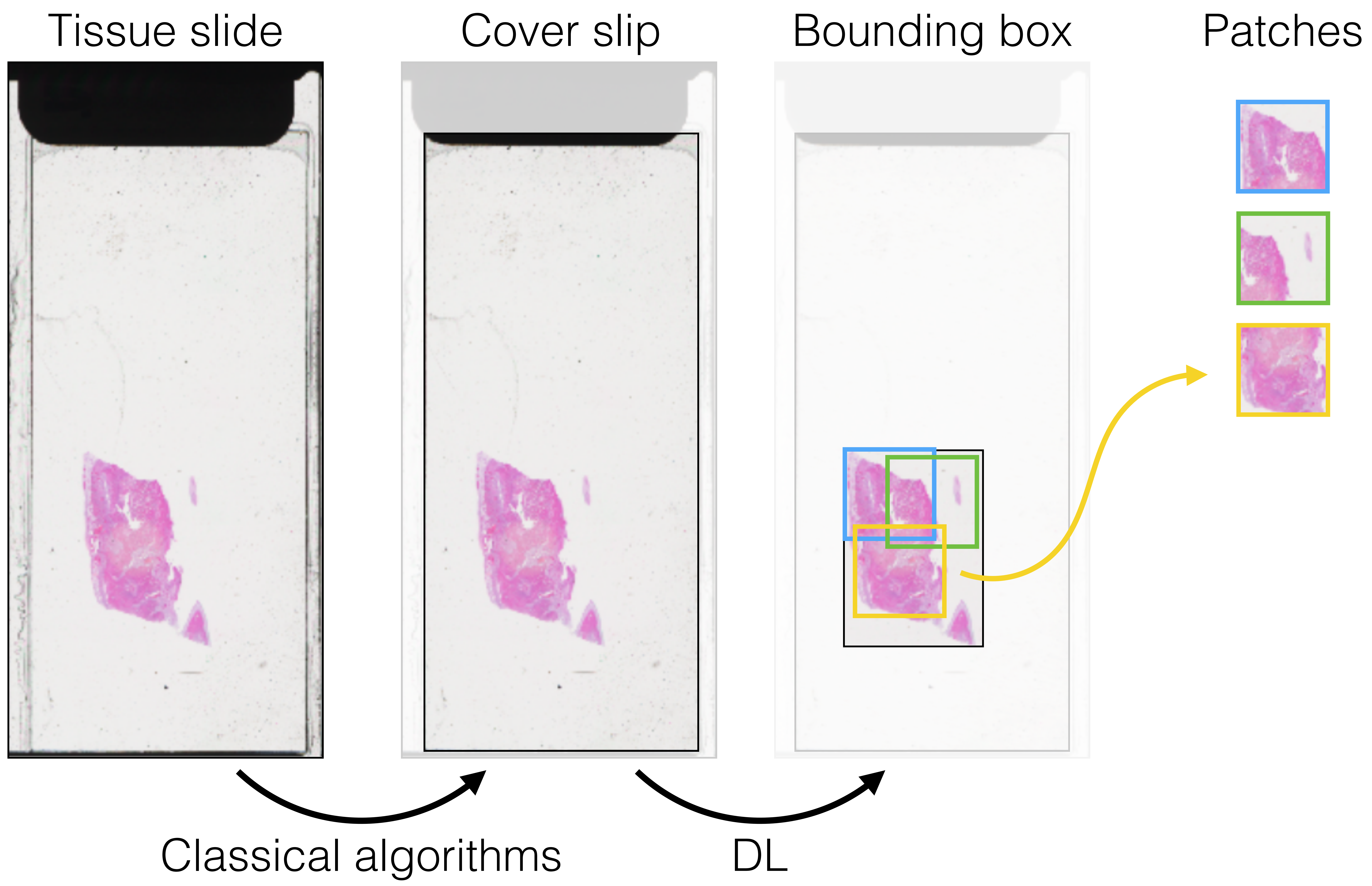


Tissue slide
Cover slip
Classical algorithms

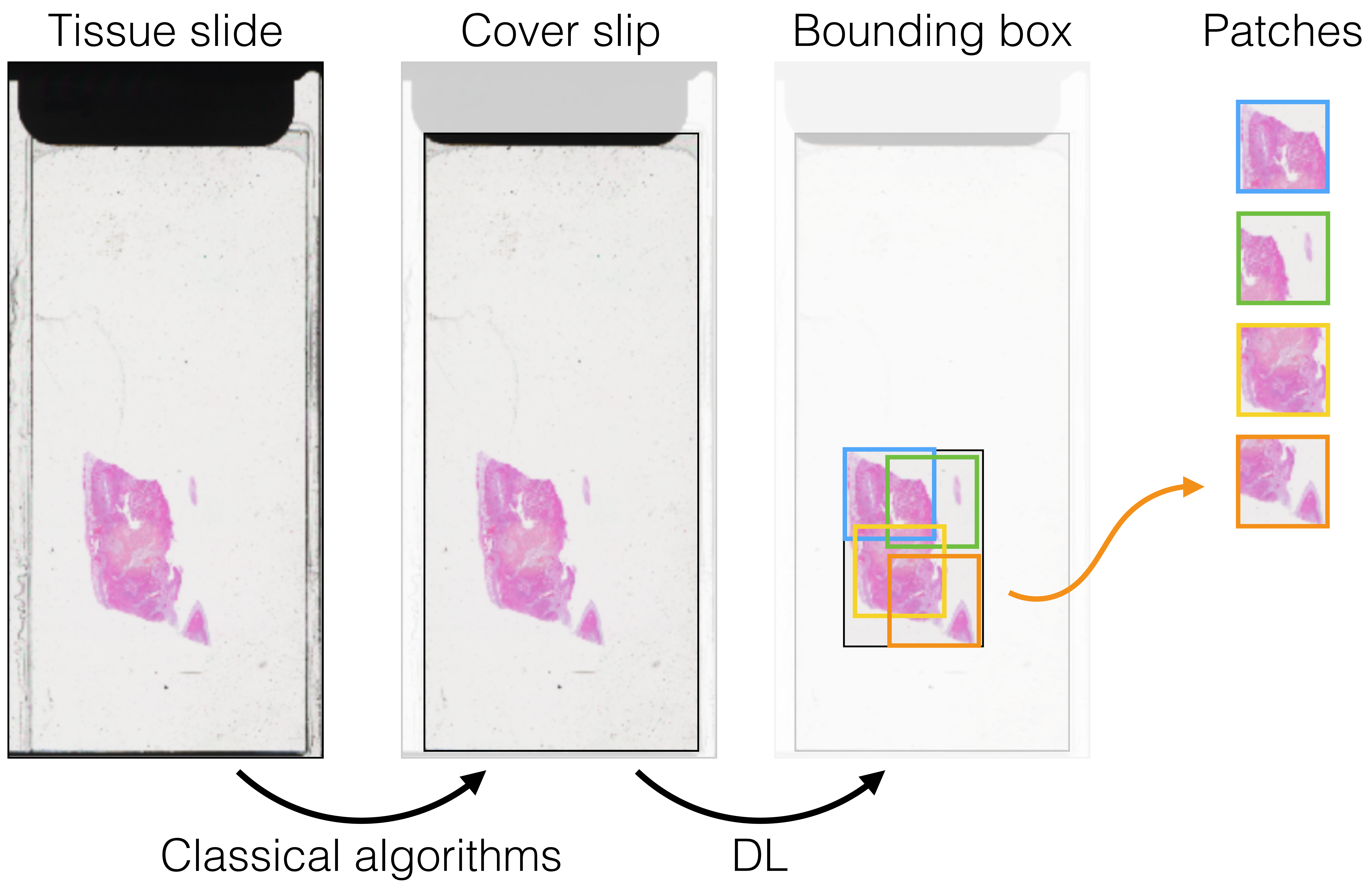

Tissue slide
Cover slip
Bounding box
Classical algorithms
DL

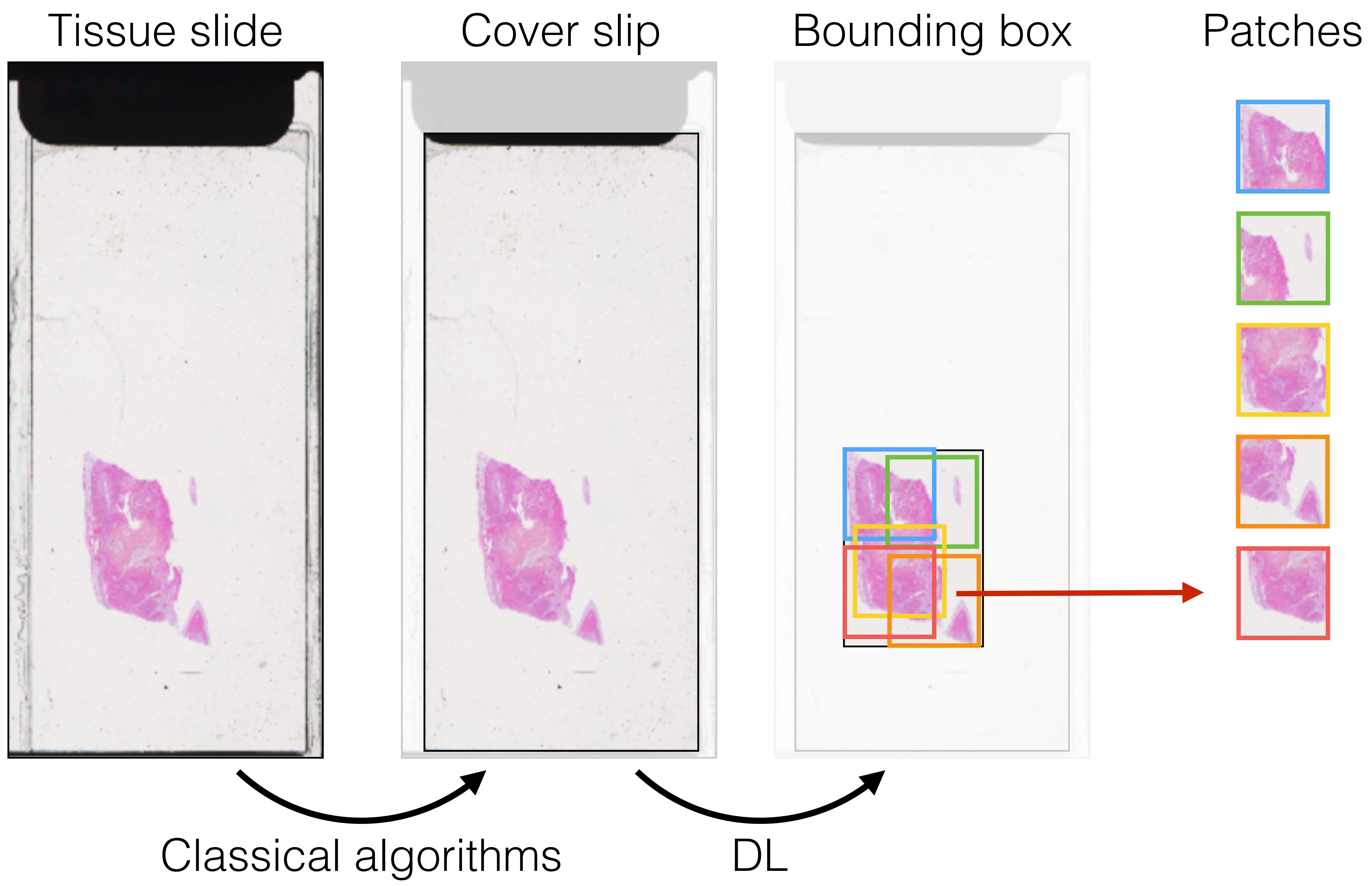

Tissue slide
Cover slip
Bounding box
Patches
Classical algorithms
DL

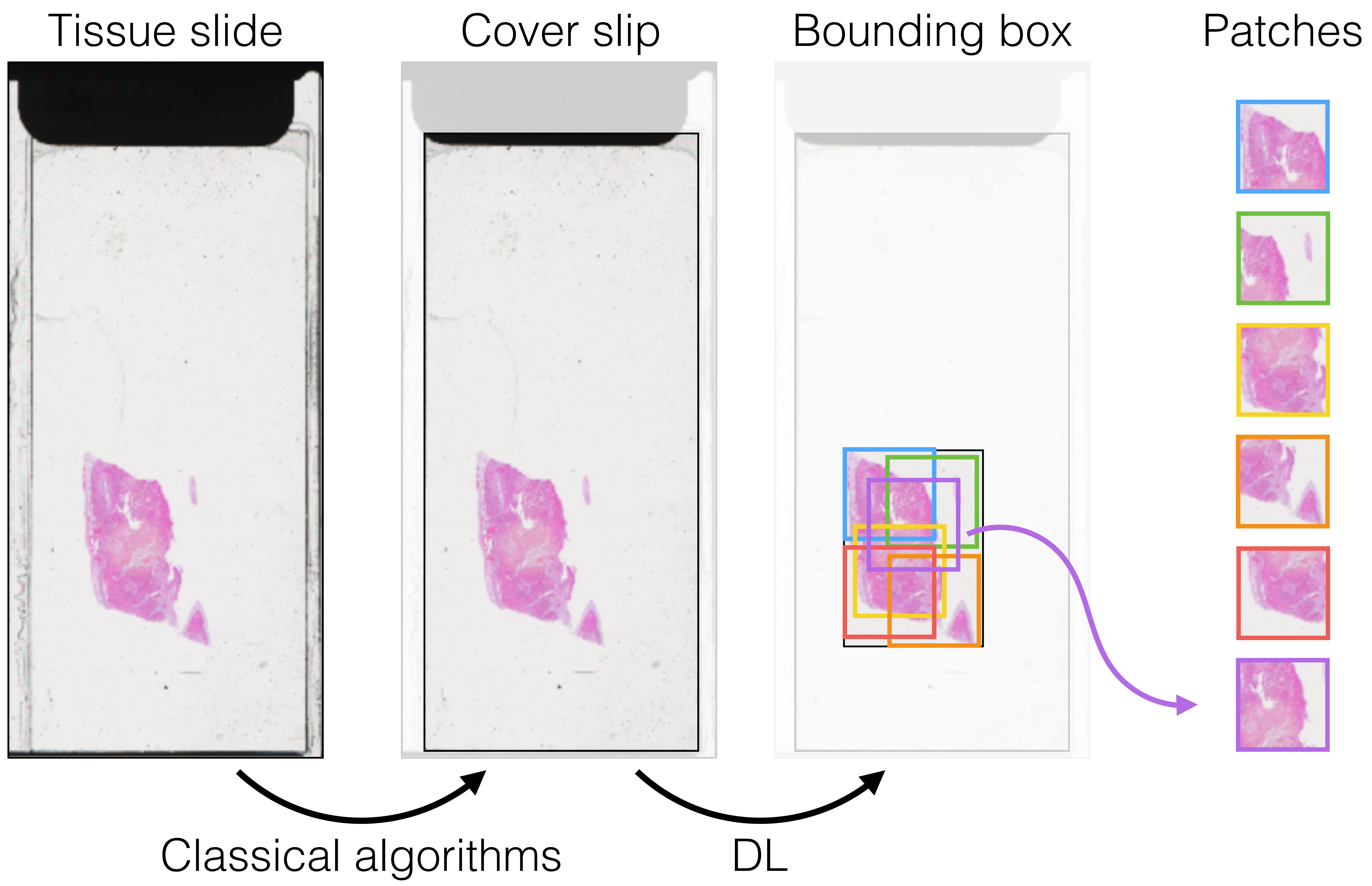

Tissue slide
Cover slip
Bounding box
Patches
Classical algorithms
DL

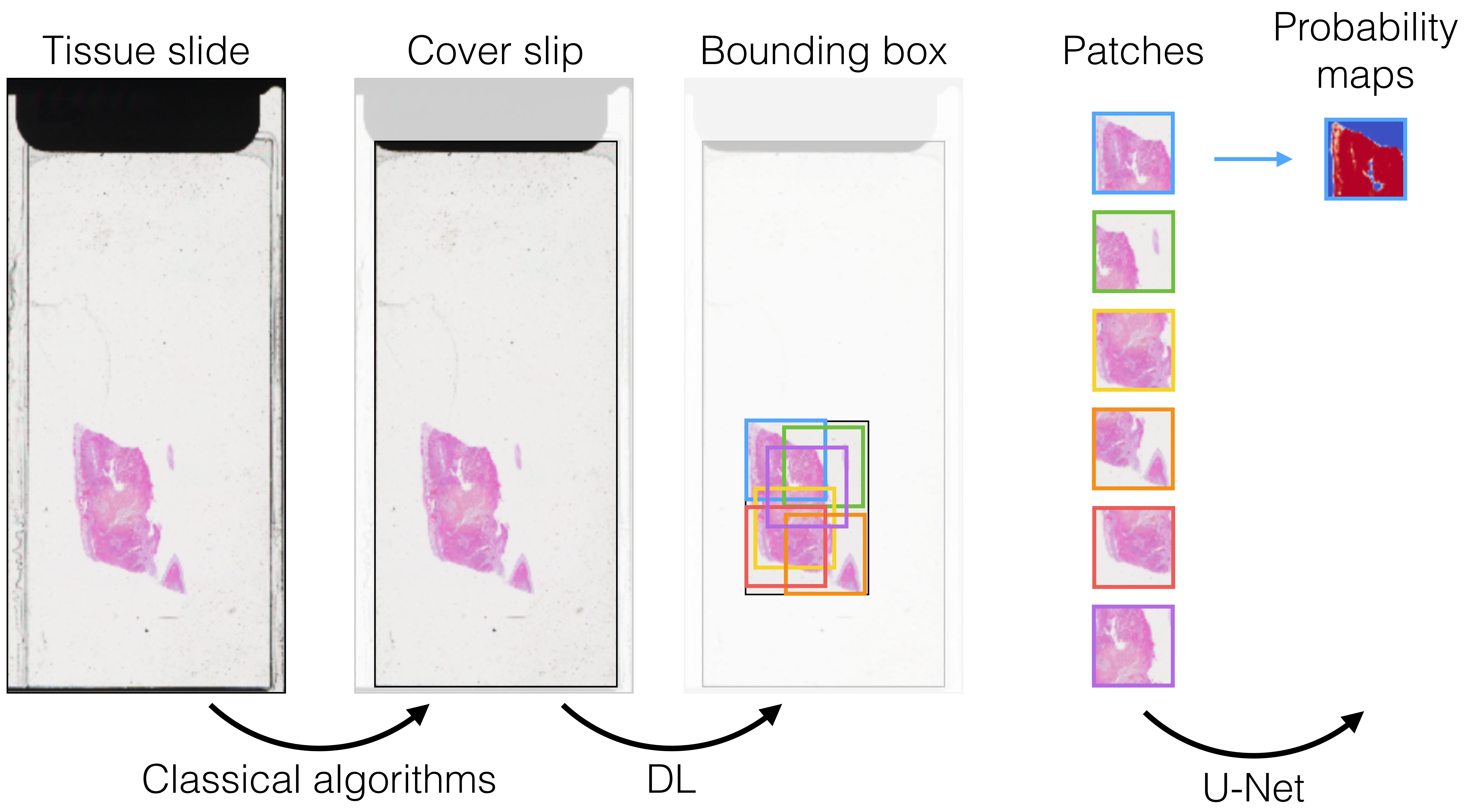


Tissue slide
Cover slip
Bounding box
Patches
Classical algorithms
DL

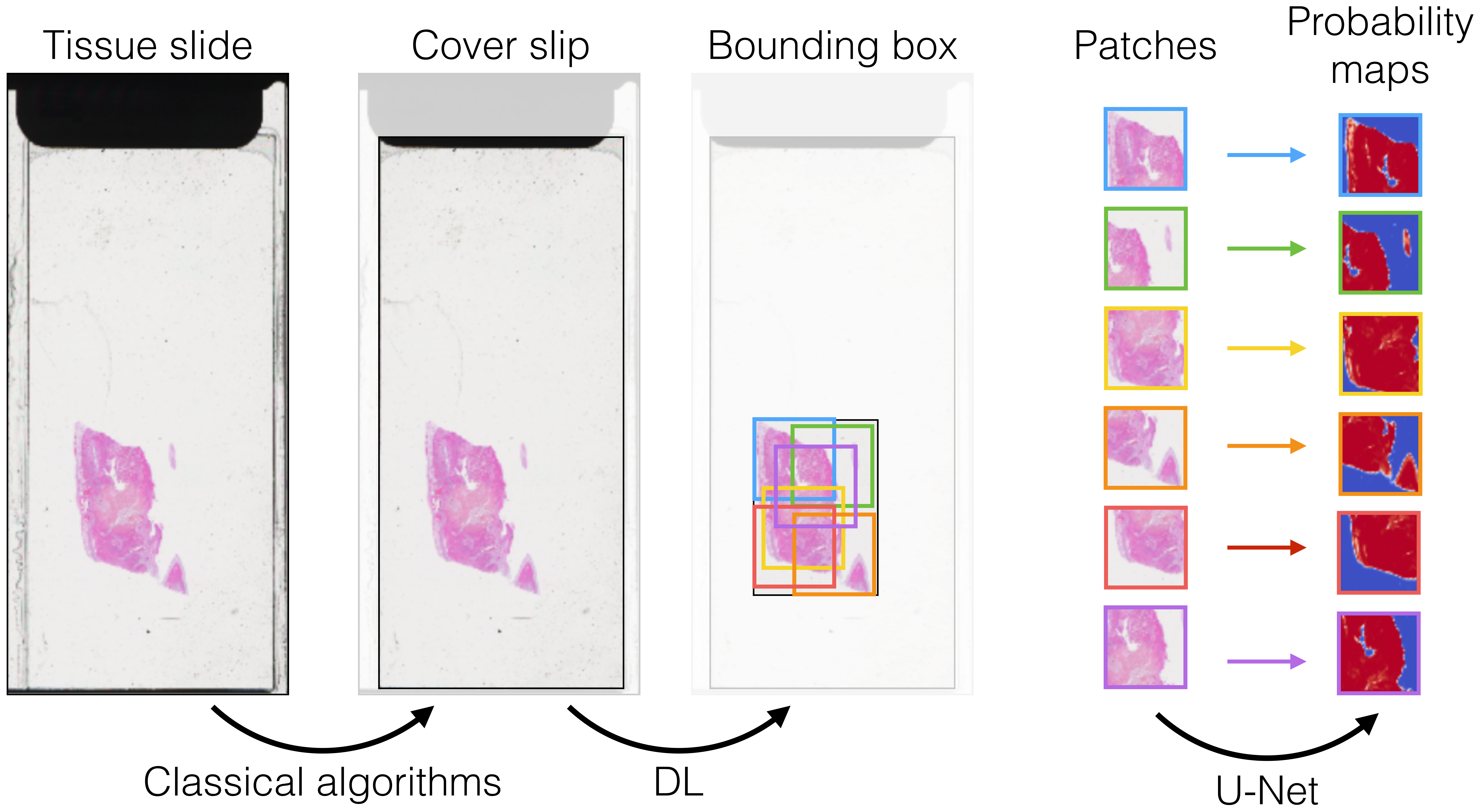

Tissue slide
Cover slip
Bounding box
Patches
Classical algorithms
DL

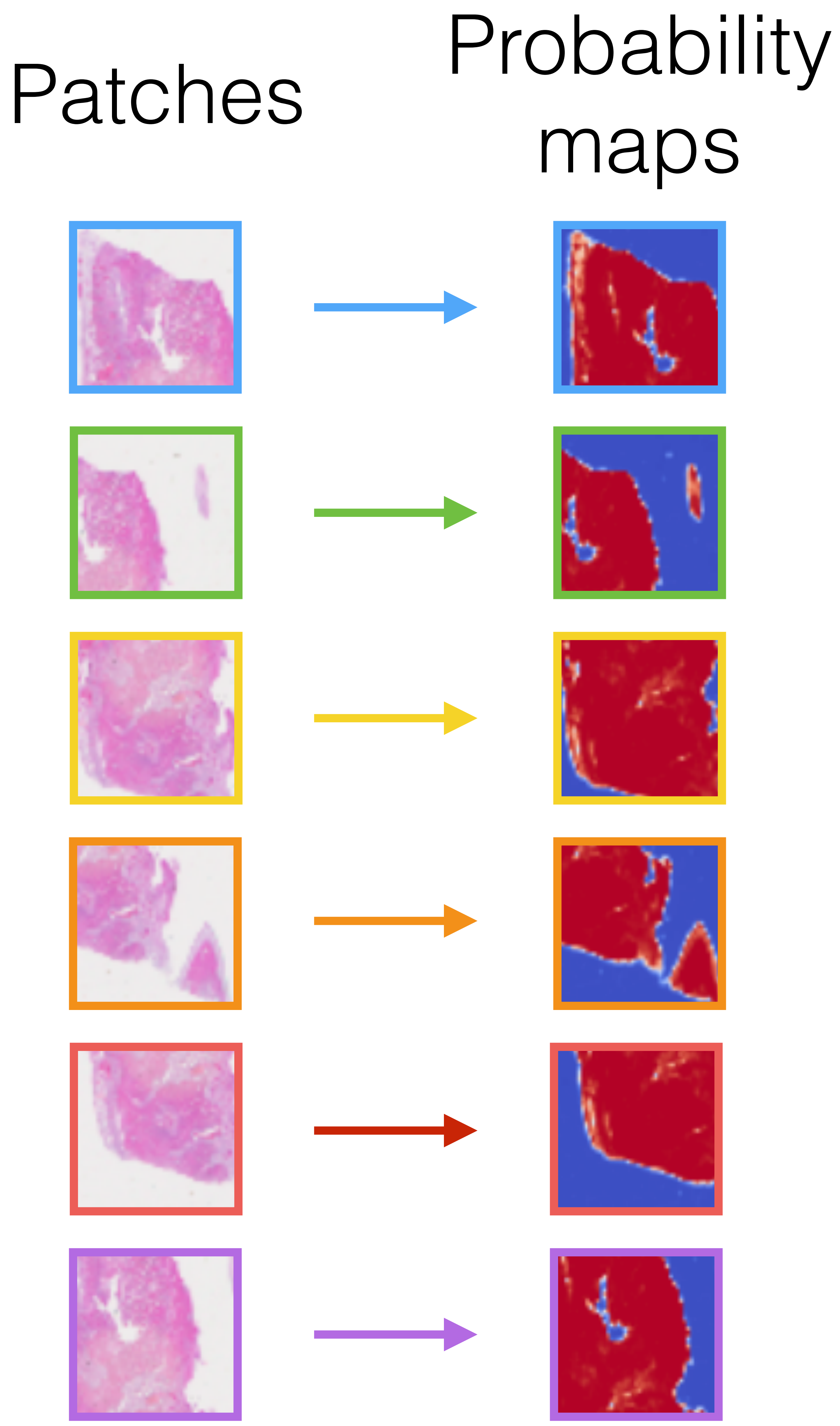

Tissue slide
Cover slip
Bounding box
Patches
Classical algorithms
DL

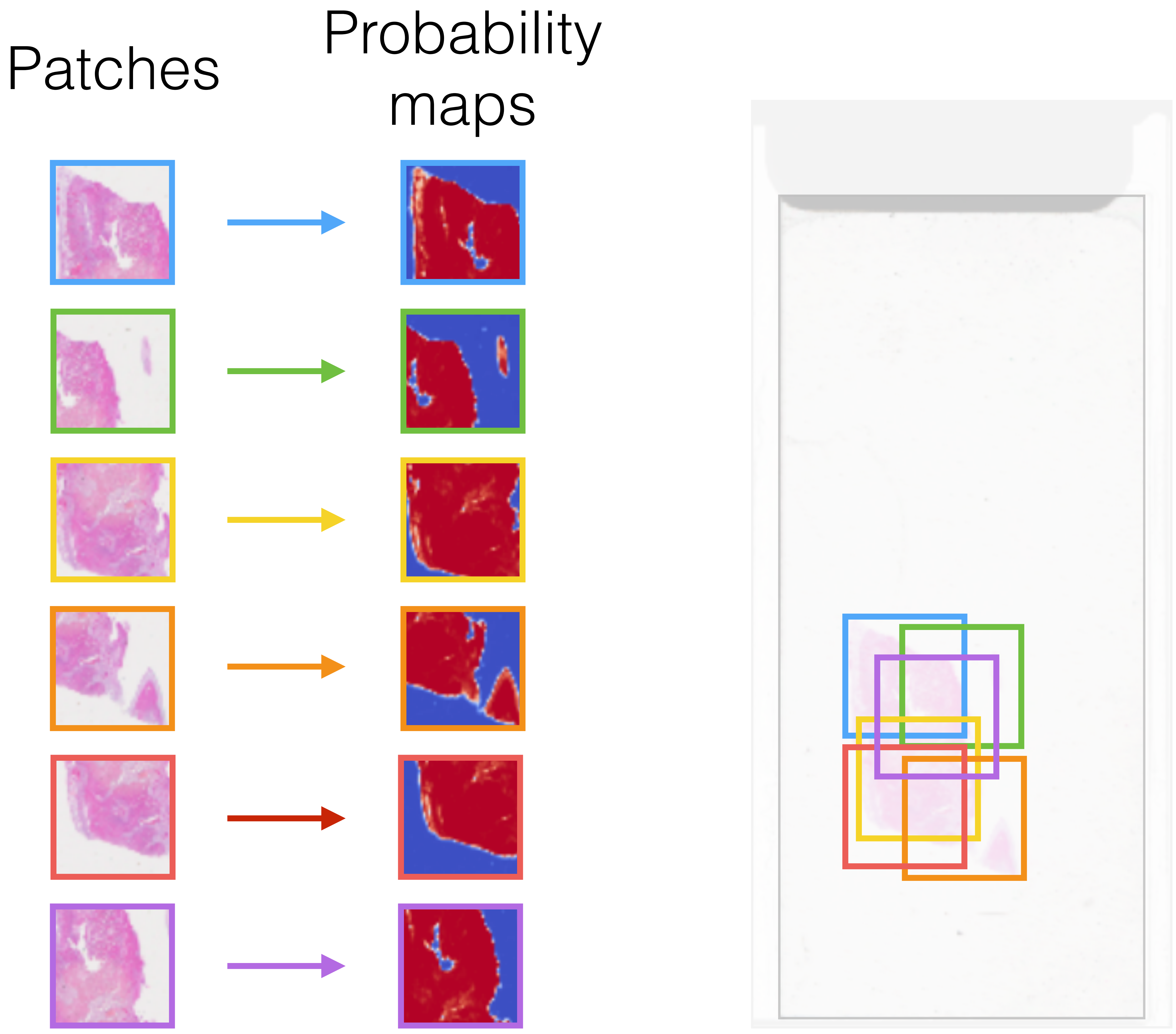

Tissue slide
Cover slip
Bounding box
Patches
Classical algorithms
DL

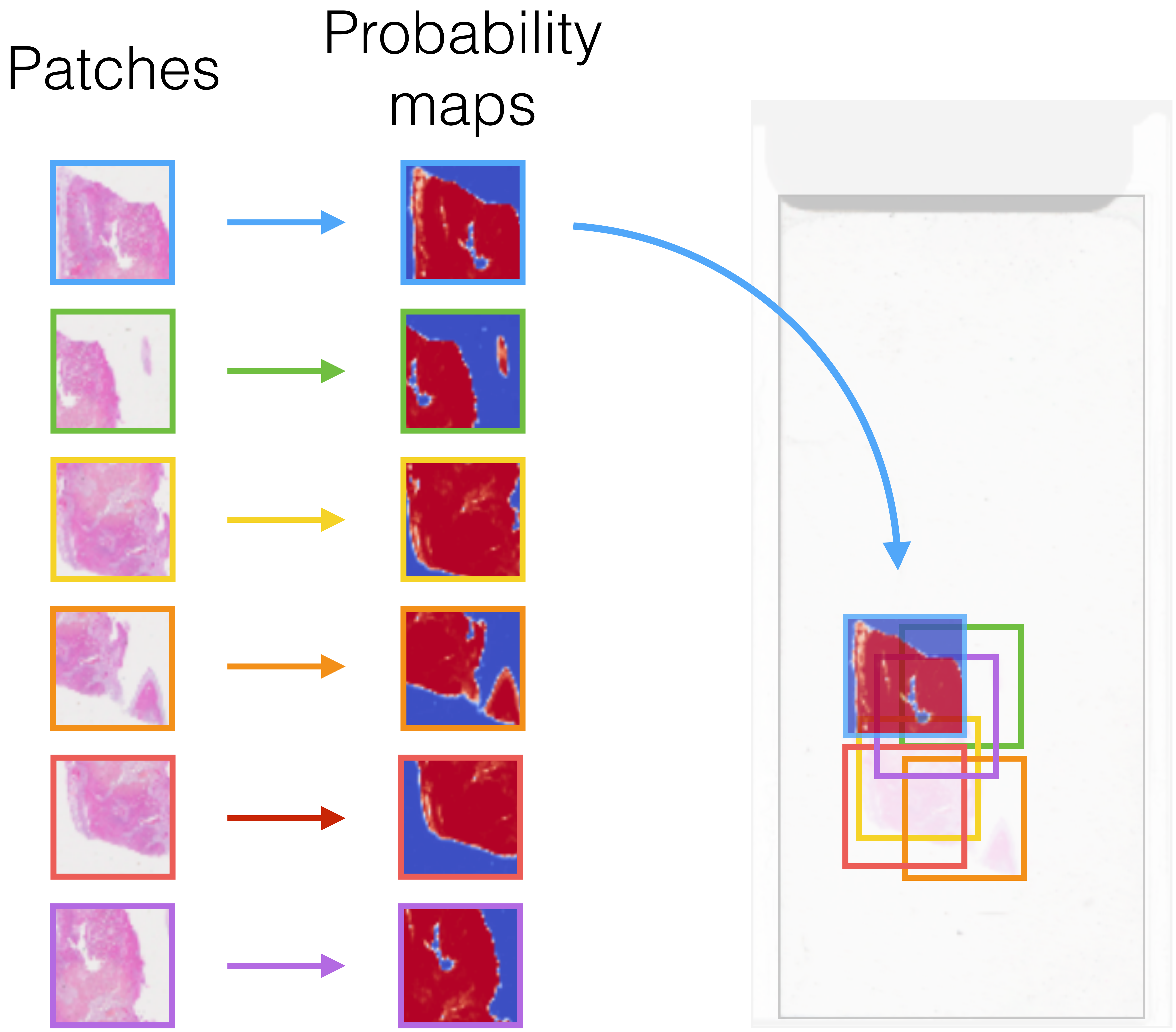


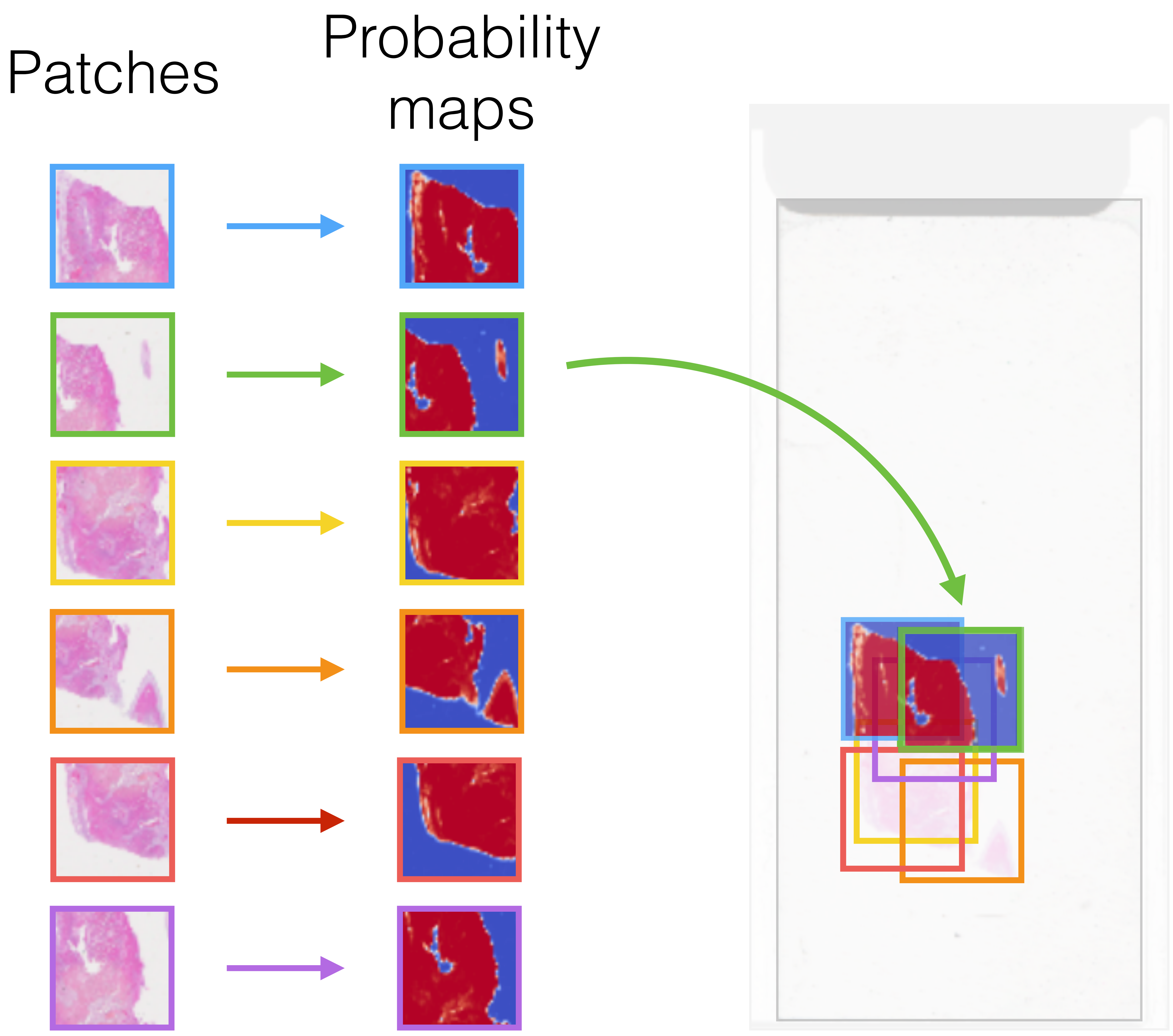
Tissue slide
Cover slip
Bounding box
Patches
Probability maps
Classical algorithms
DL
U-Net

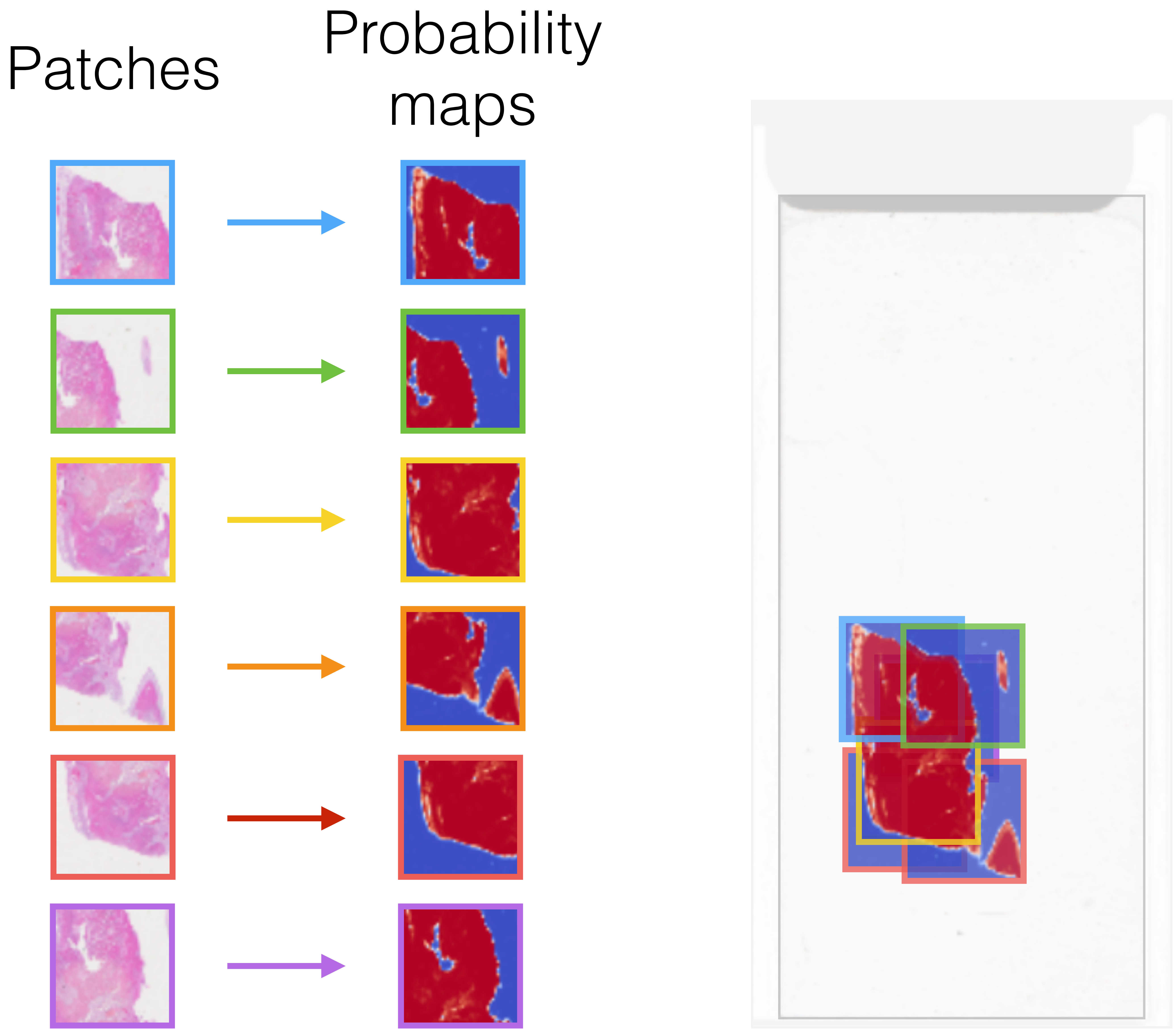

Tissue slide
Cover slip
Bounding box
Patches
Probability maps
Classical algorithms
DL
U-Net


Patches
Probability maps

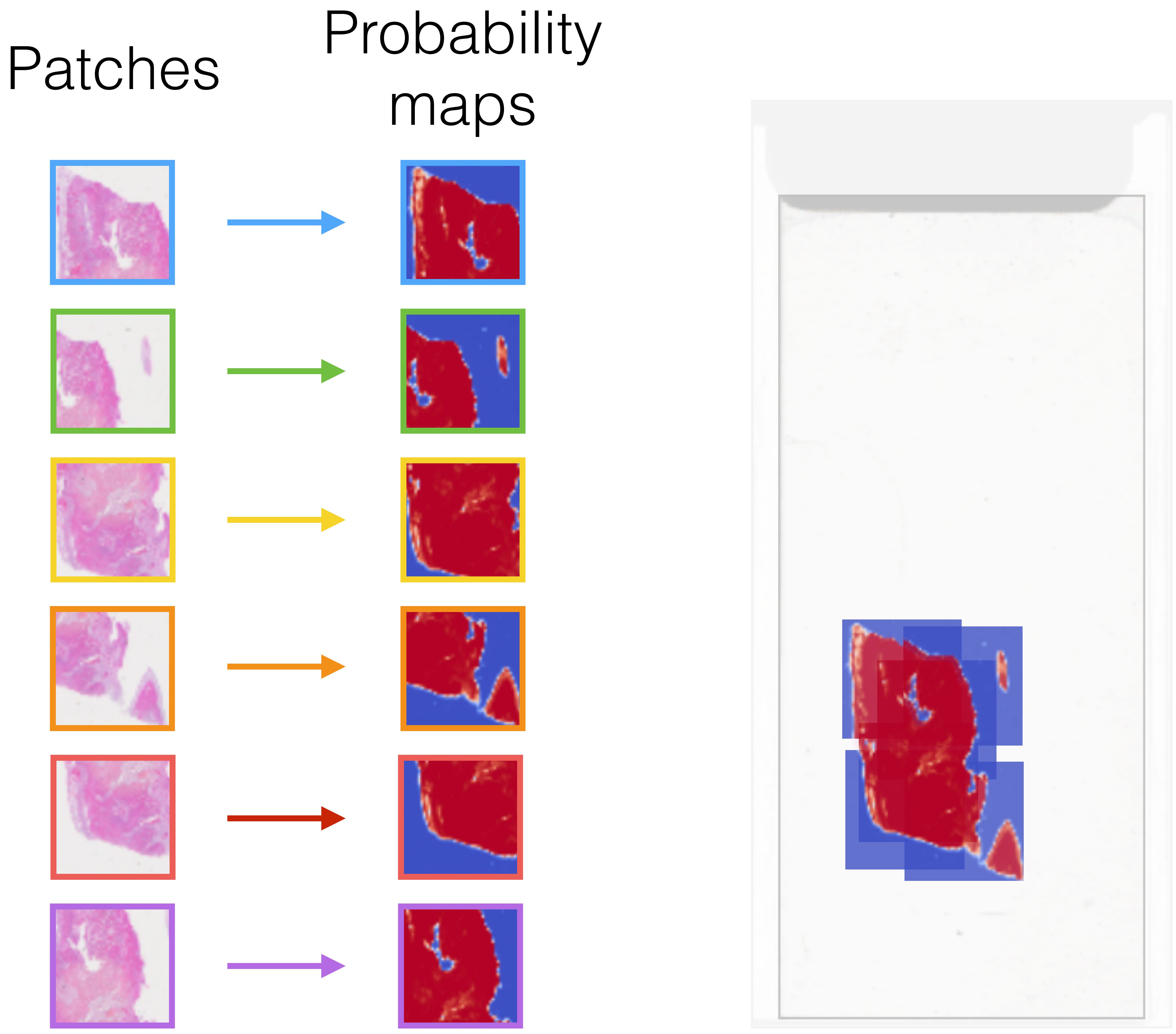

Patches
Probability maps


Patches
Probability maps


Patches
Probability maps

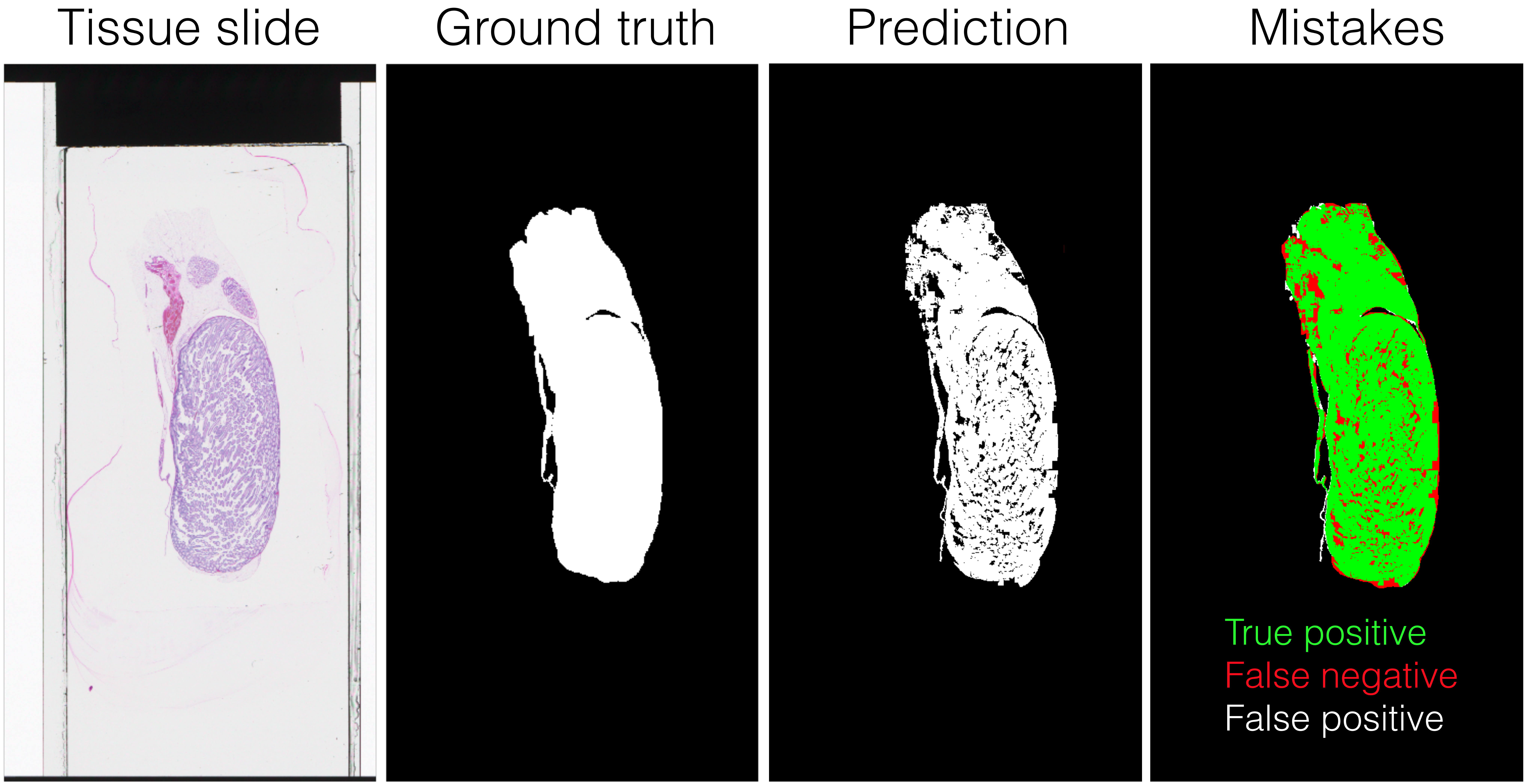


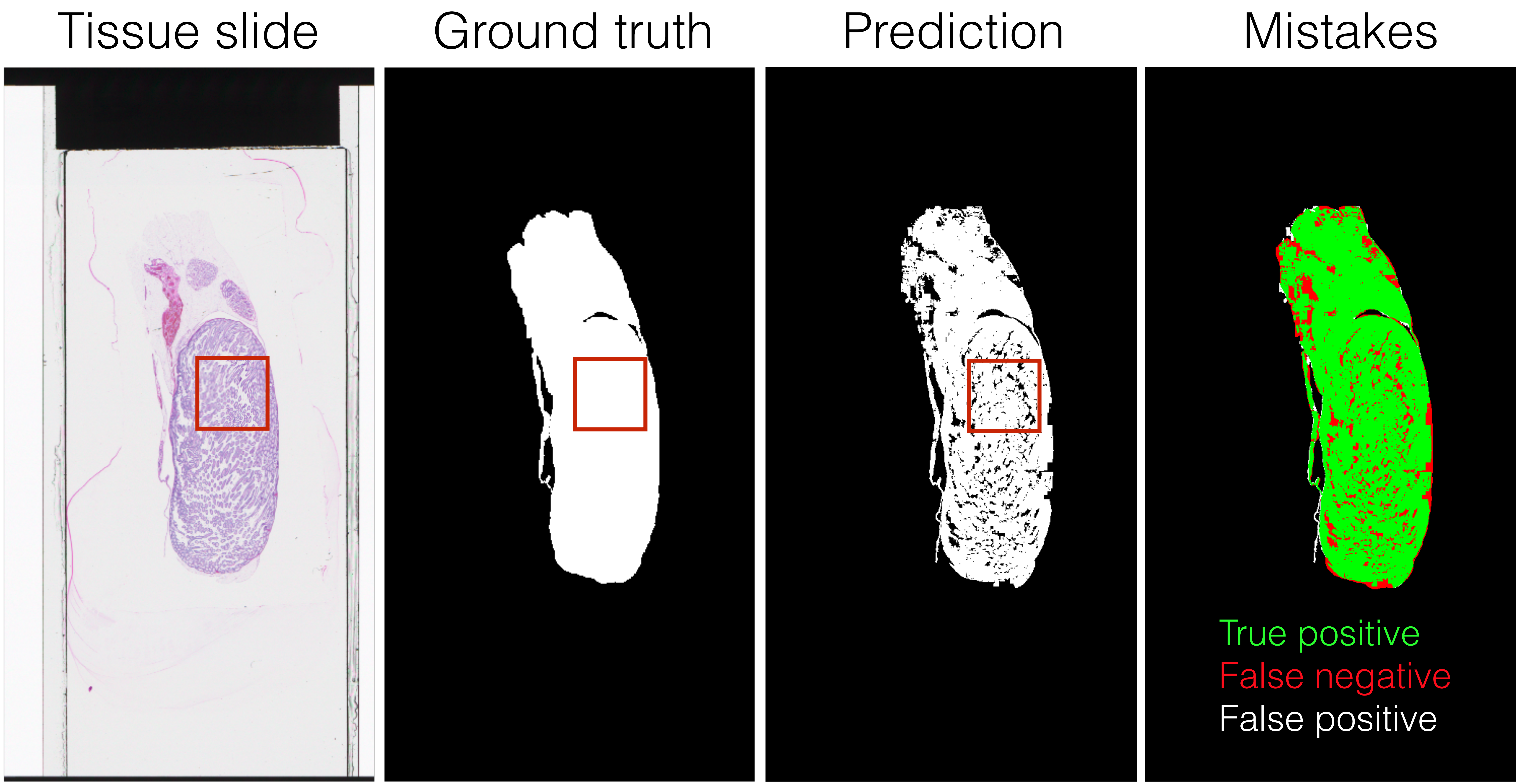
Patches
Probability maps

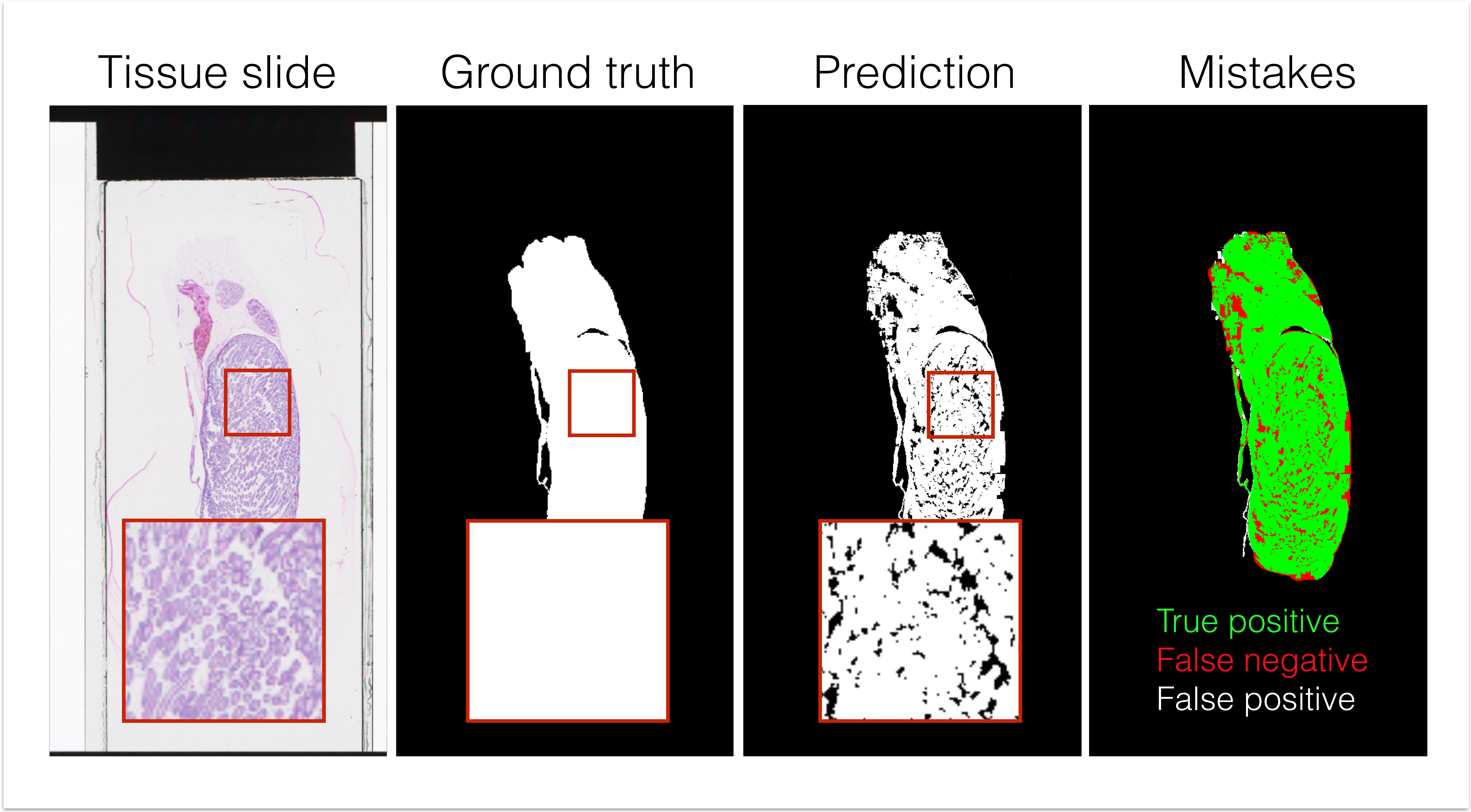

Patches
Probability maps

Patches
Probability maps
Thresholding & post-processing
Segmentation

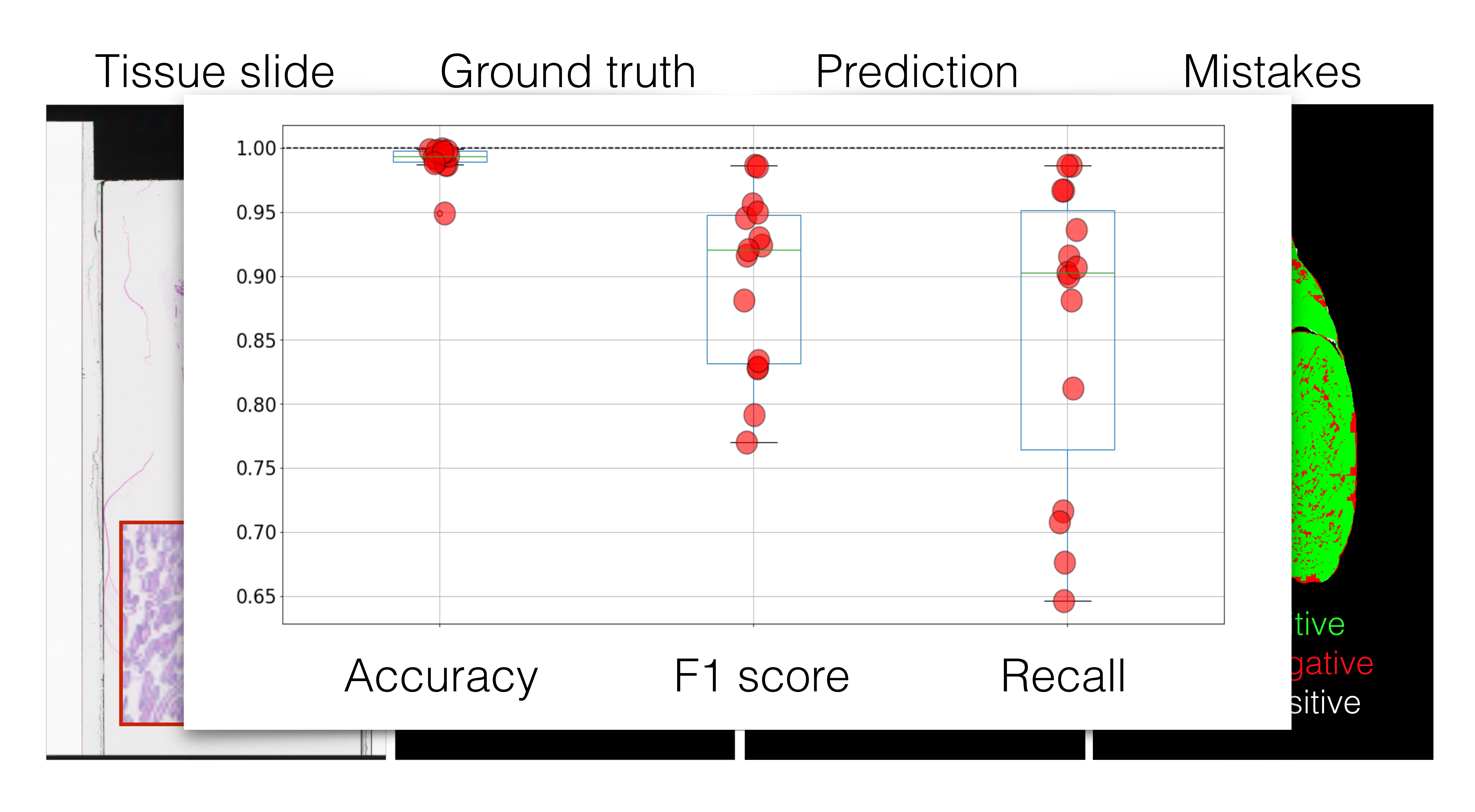

Patches
Probability maps
Thresholding &
post-processing
Segmentation


Tissue slide
Ground truth
Prediction
Mistakes
True positive
False negative
False positive

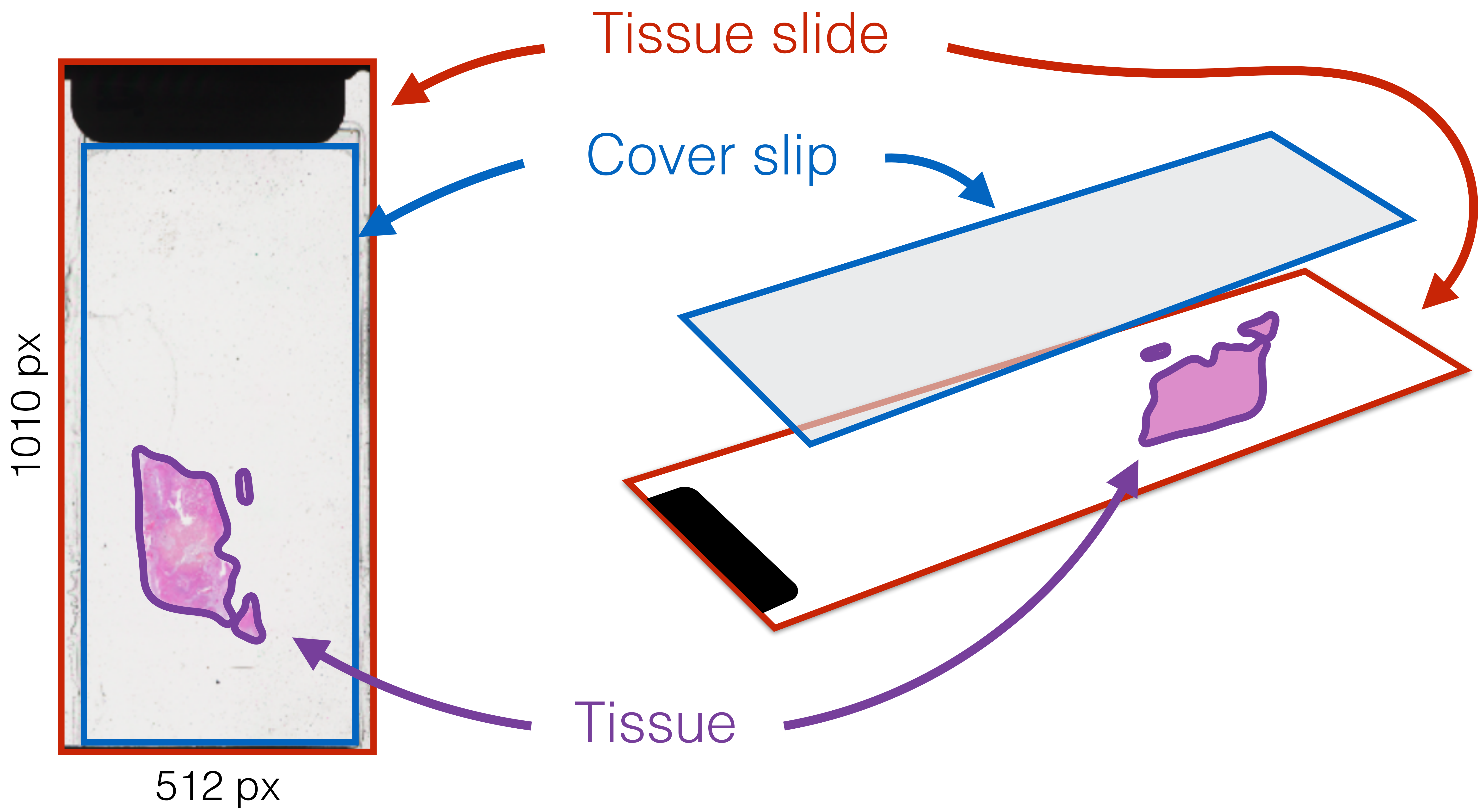

Tissue slide
Ground truth
Prediction
Mistakes
True positive
False negative
False positive



Tissue slide
Ground truth
Prediction
Mistakes
True positive
False negative
False positive


Tissue slide
Ground truth
Prediction
Mistakes
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
Accuracy
F1 score
Recall
tive
gative
sitive

PerkinElmer®
For the Better



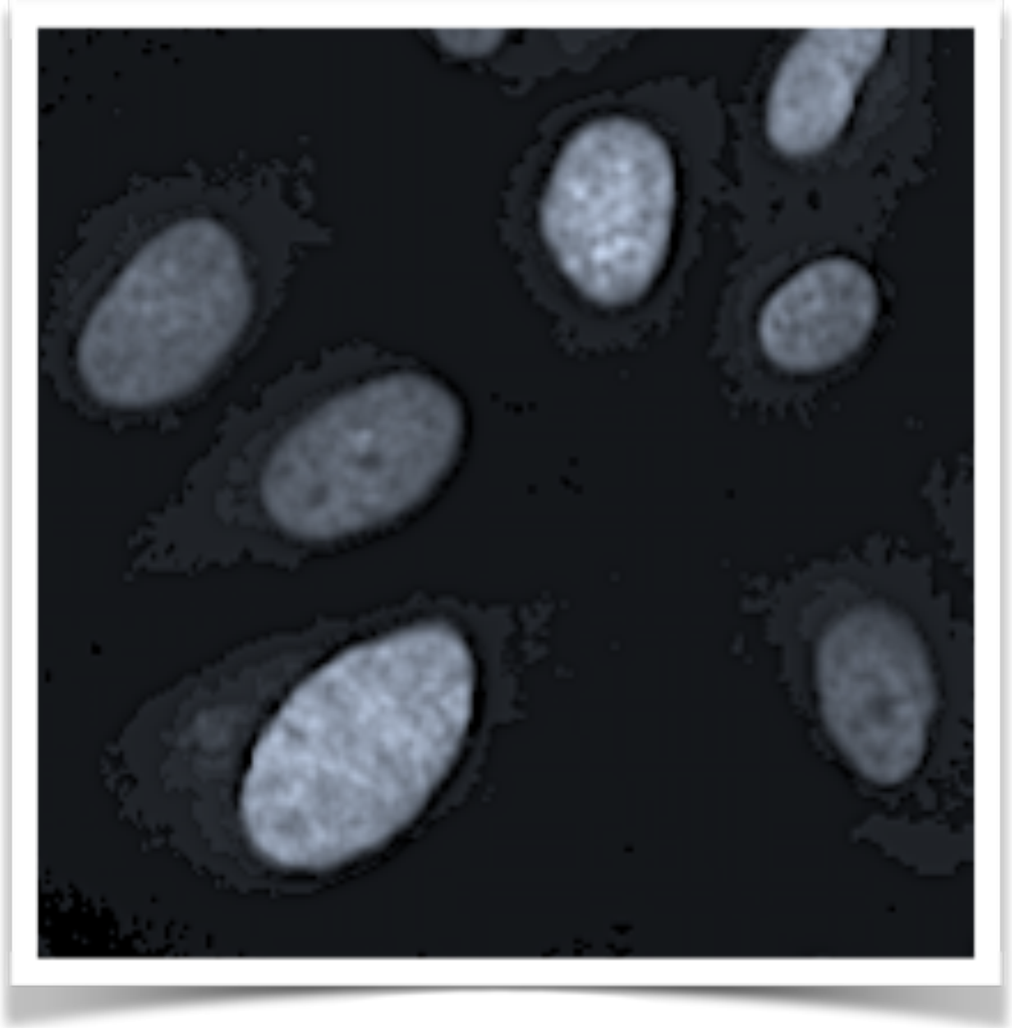
TARTU ÜLIKOOL · UNIVERSITAS TARTUENSIS · 1632

Dmytro Fishman

Oleksandr Pryhoda

UCU. APPLIED SCIENCES FACULTY

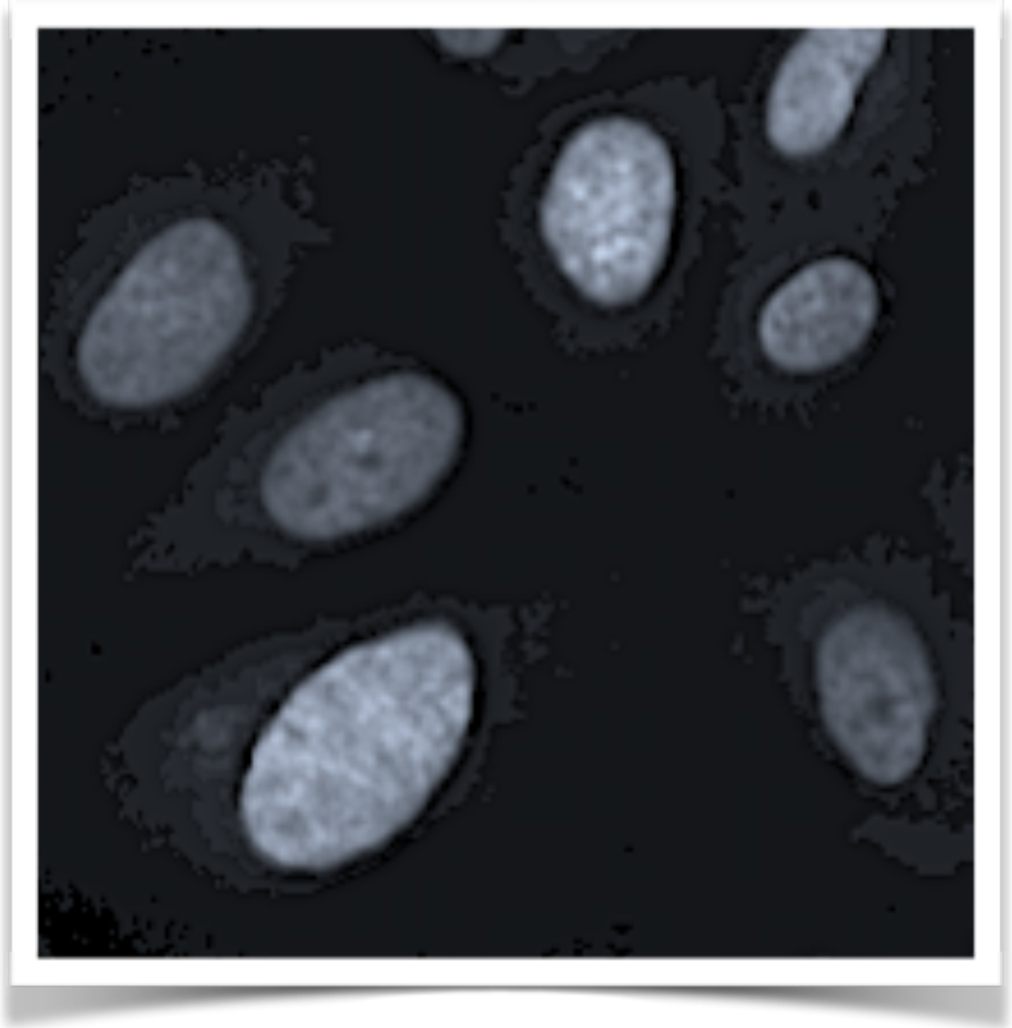
Fluorescence



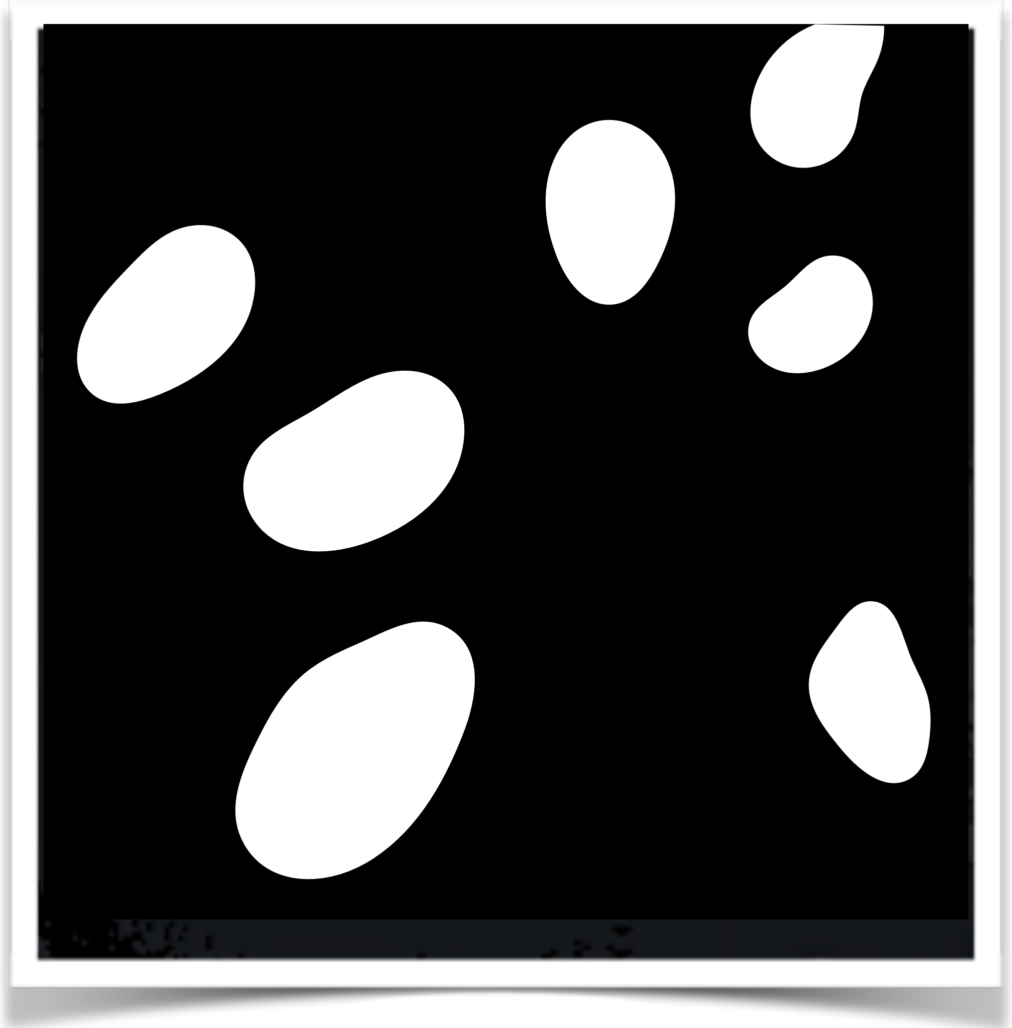
Summary

Summary

Fluorescence

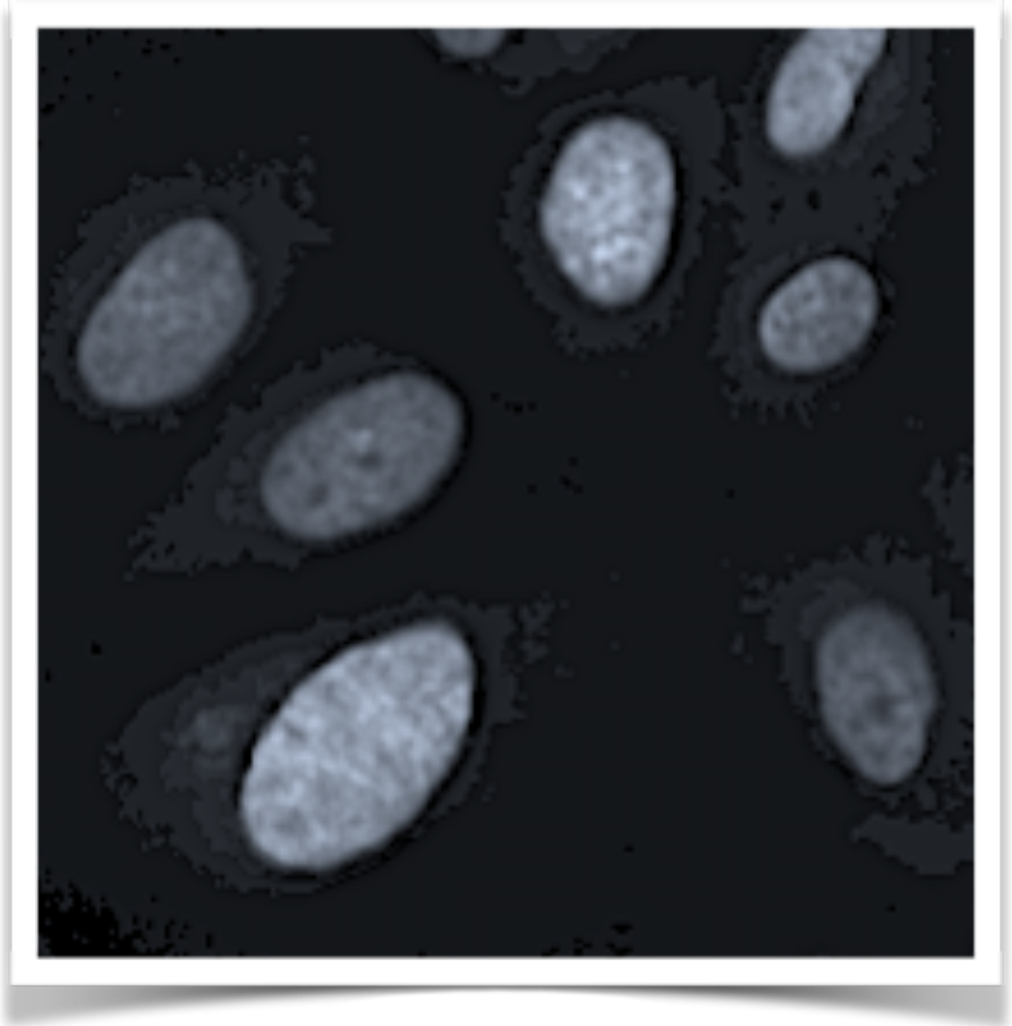


Segmentation

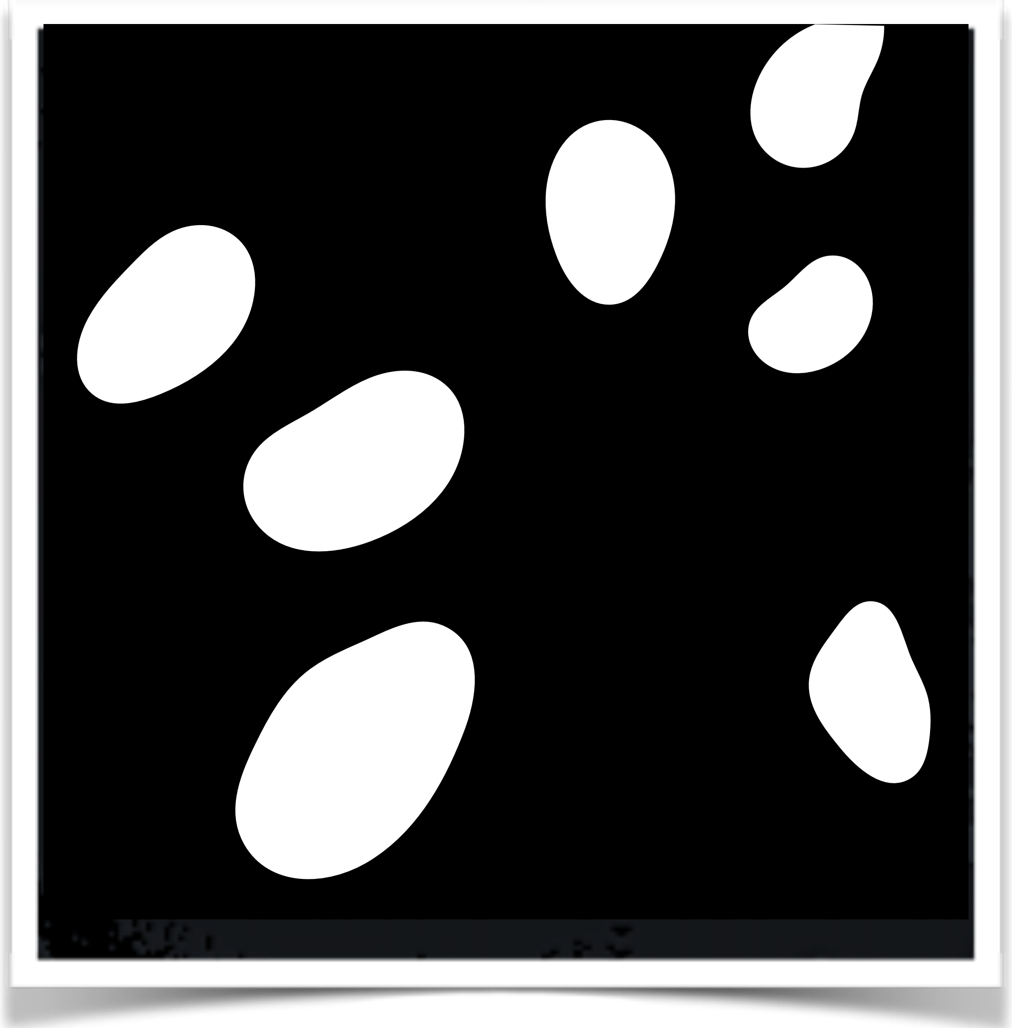


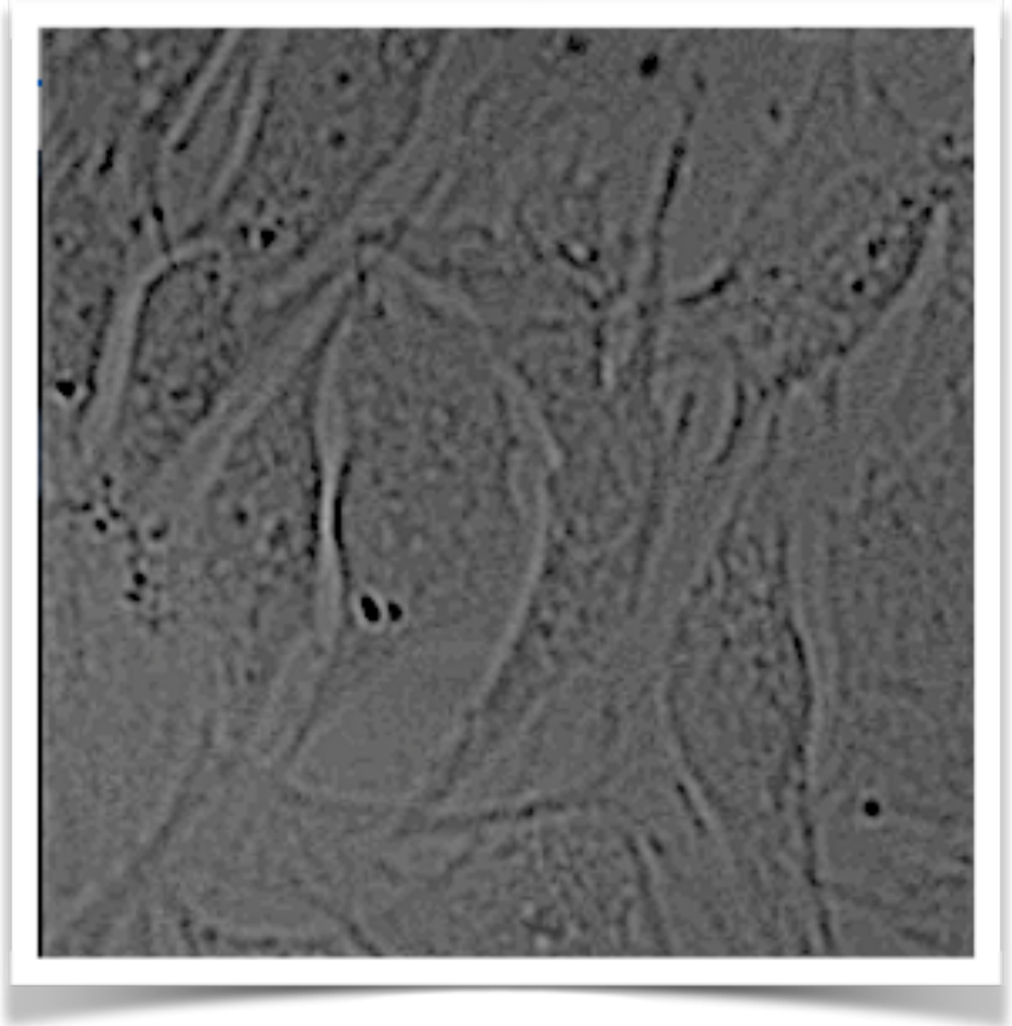
Summary

Fluorescence



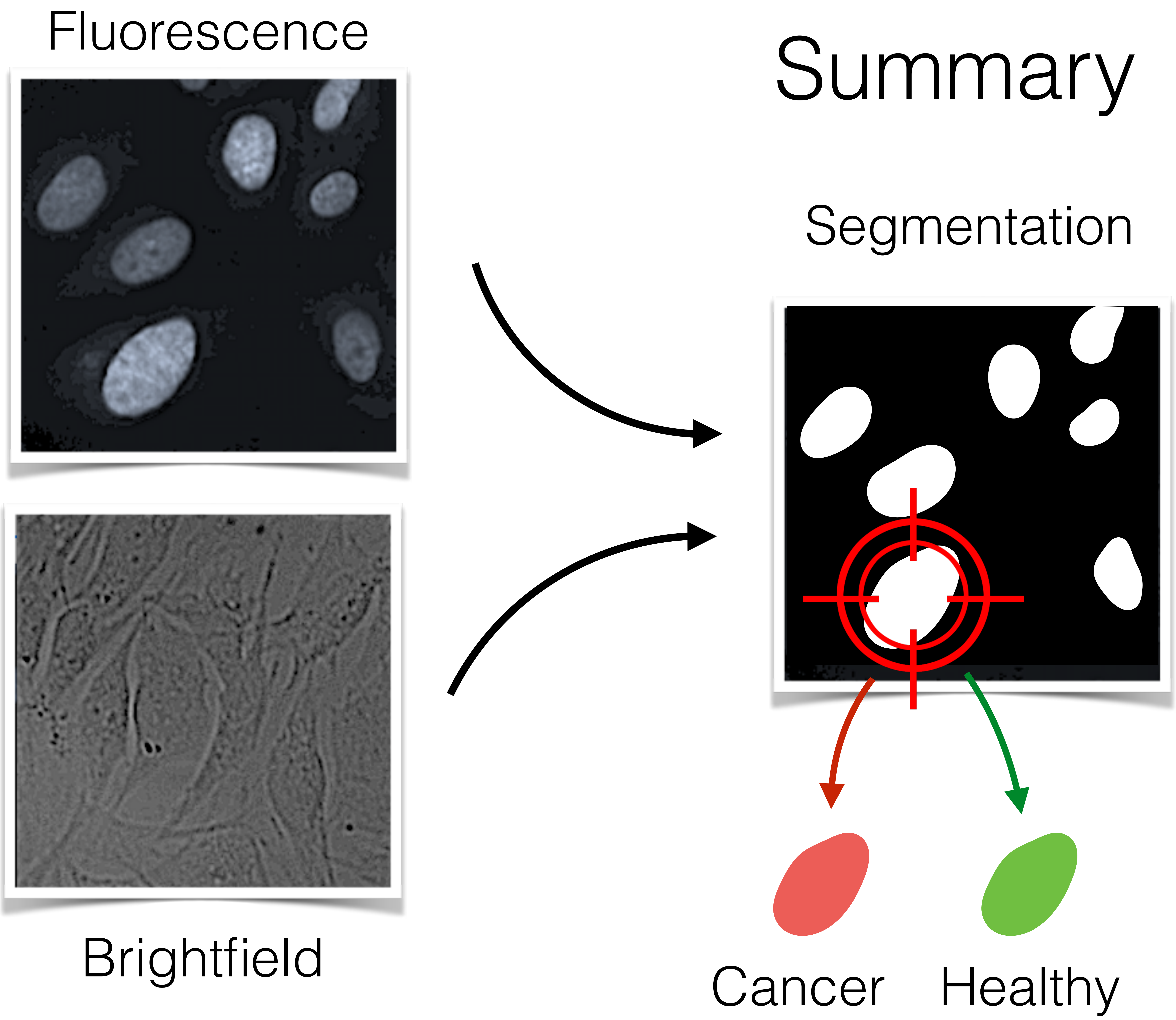
Segmentation

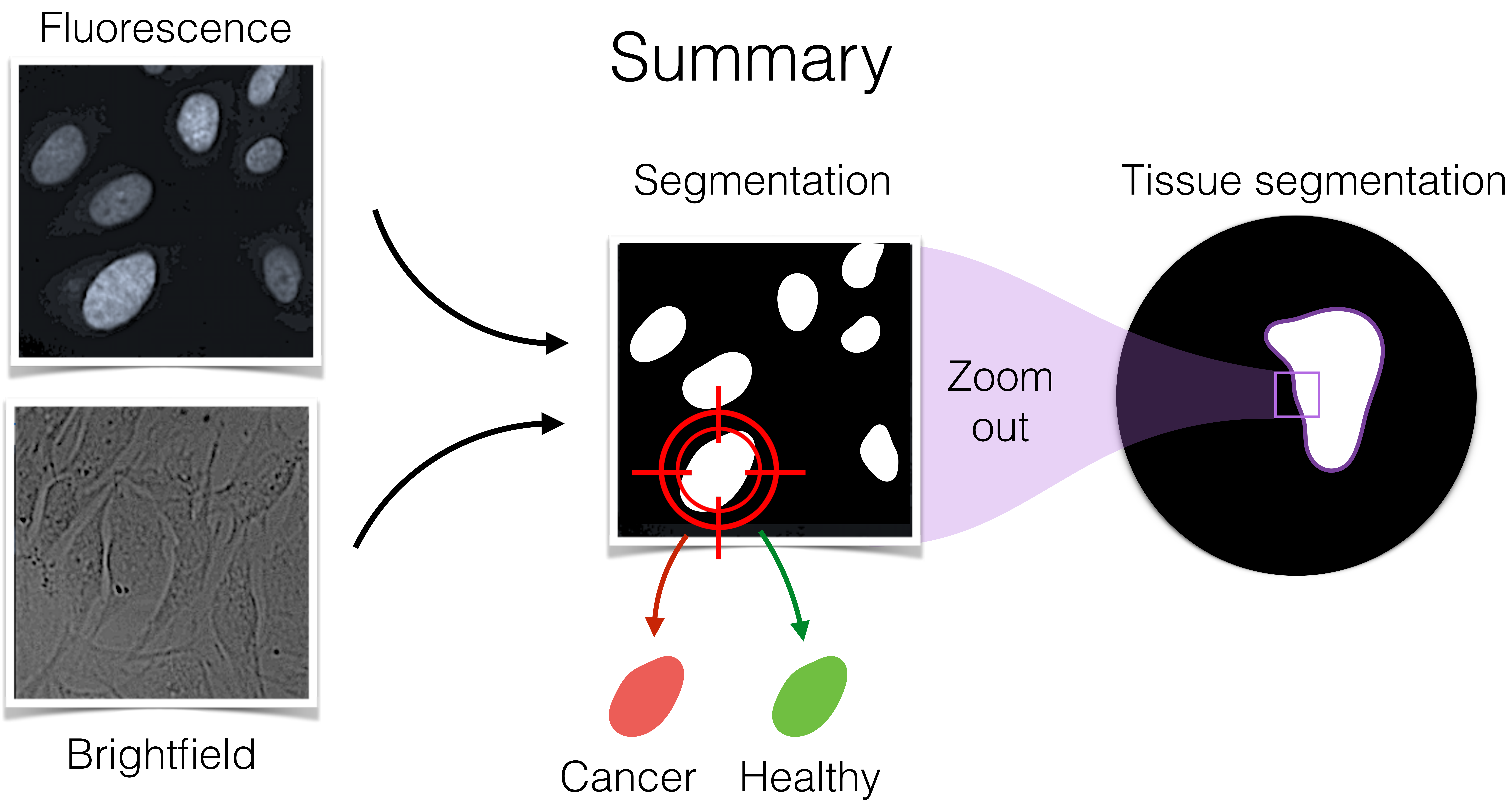




Brightfield

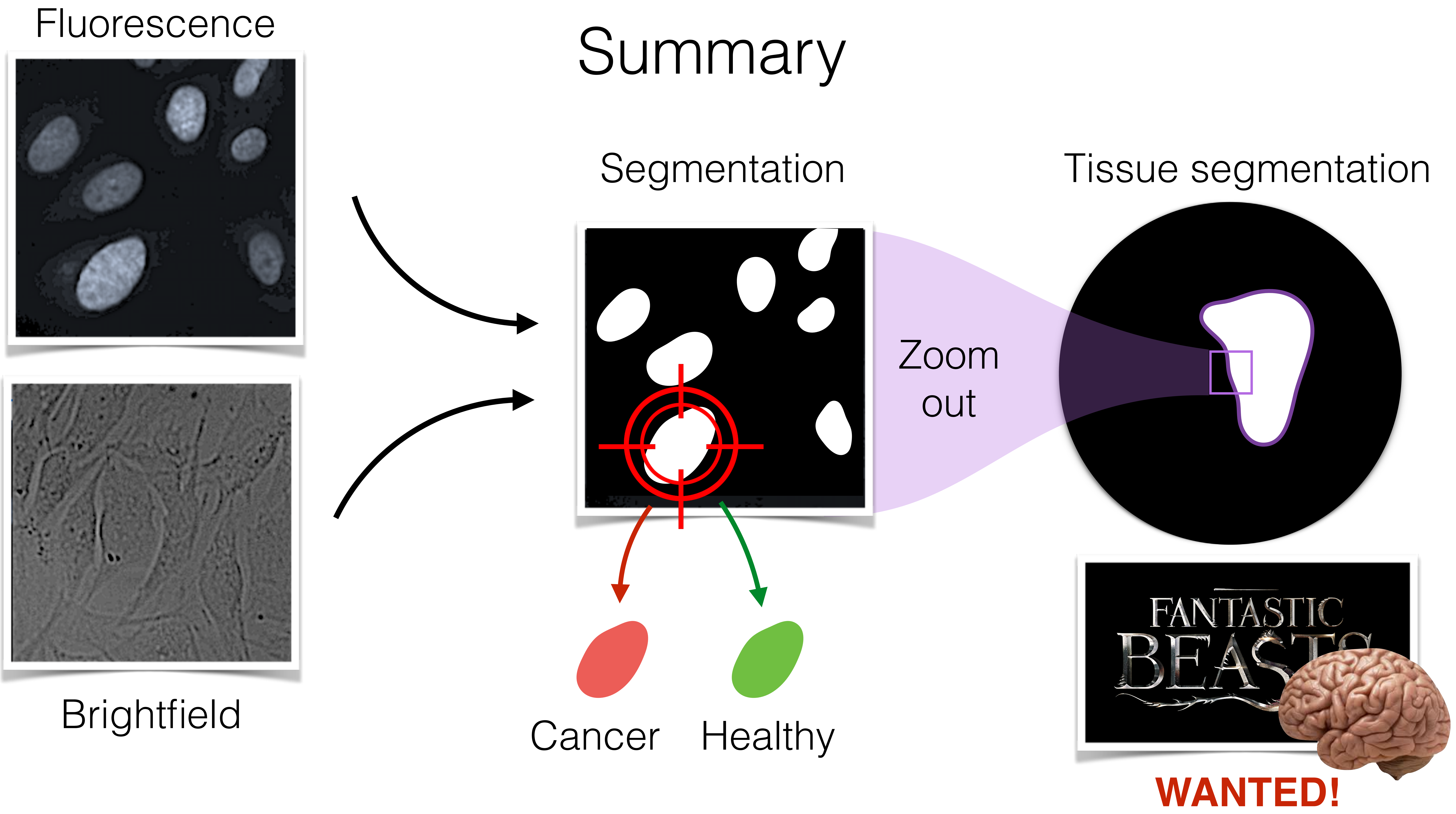
Summary





Summary
Fluorescence
Brightfield
Segmentation
Zoom
out
Tissue segmentation
Cancer
Healthy

Summary

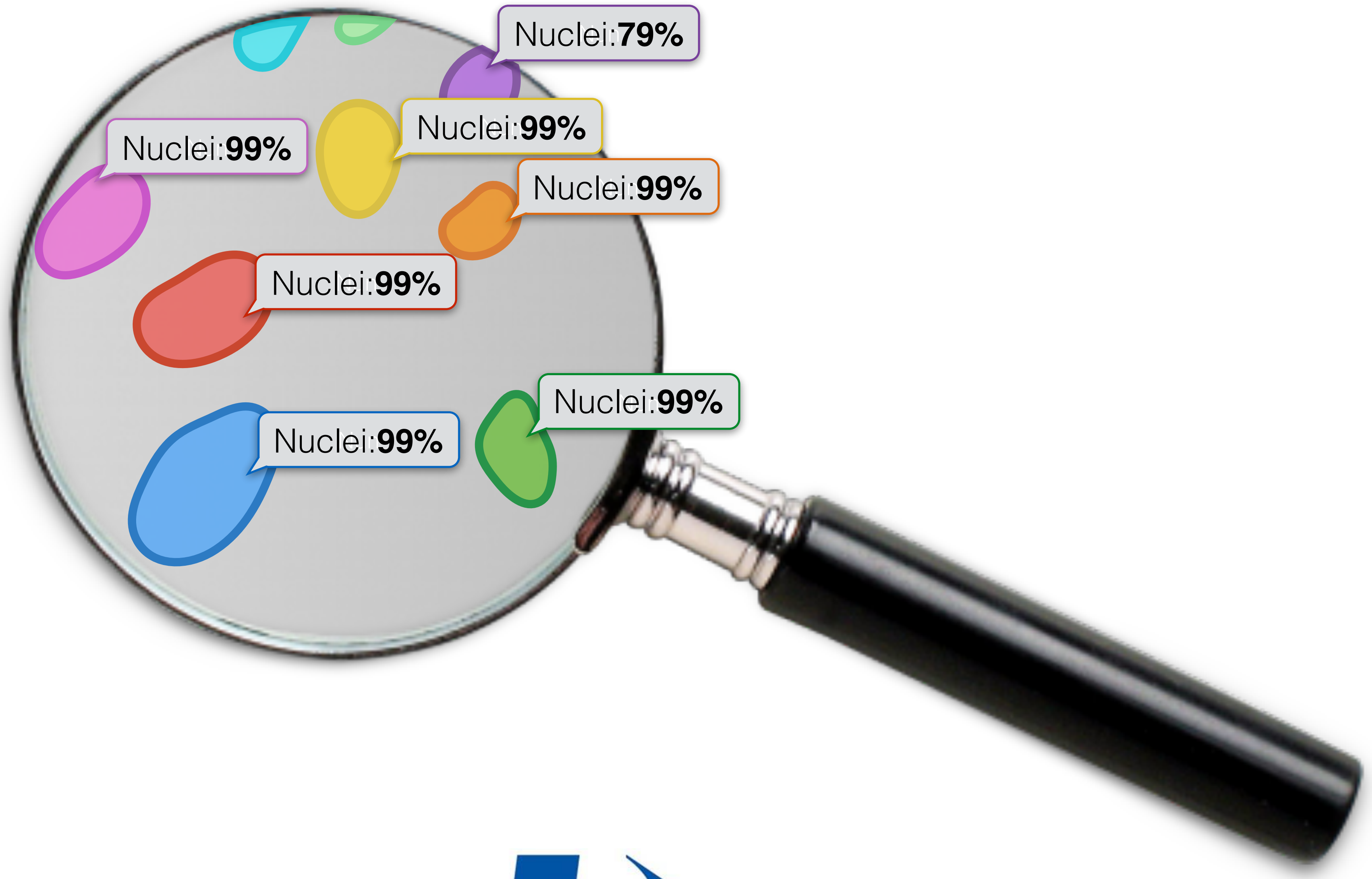

Fluorescence
Brightfield
Segmentation
Tissue segmentation
Zoom
out
Cancer
Healthy
FANTASTIC
BEASTS
WANTED!

Project team

TARTU ÜLIKOOL · UNIVERSITAS TARTUENSIS 1632

BIIT

Dmytro Fishman
Mikhail Papkov
Sten-Oliver Salumaa
Tanel Pärnamaa
Jaak Vilo



TARTU ÜLIKOOL · UNIVERSITAS TARTUENSIS 1632

COMPUTATIONAL NEUROSCIENCE
Research group at the University of Tartu

Daniel Majoral
Roman Ring
Ilya Kuzovkin
Ardi Tampuu
Raul Vicente

wellcome sanger institute

Leopold Parts
Omer Bayraktar
William Jones
Elizabeth Bell

PerkinElmer®
For the Better

Kaupo Palo
Martin Daffertshofer

TARTU ÜLIKOOL · UNIVERSITAS TARTUENSIS 1632

UCU. APPLIED SCIENCES FACULTY

Bohdan Petryshak
Oleksandr Pryhoda

dws4.me