# Histologie numérique sur MicroPICell : bénéficiez d'un service "deep learning friendly"

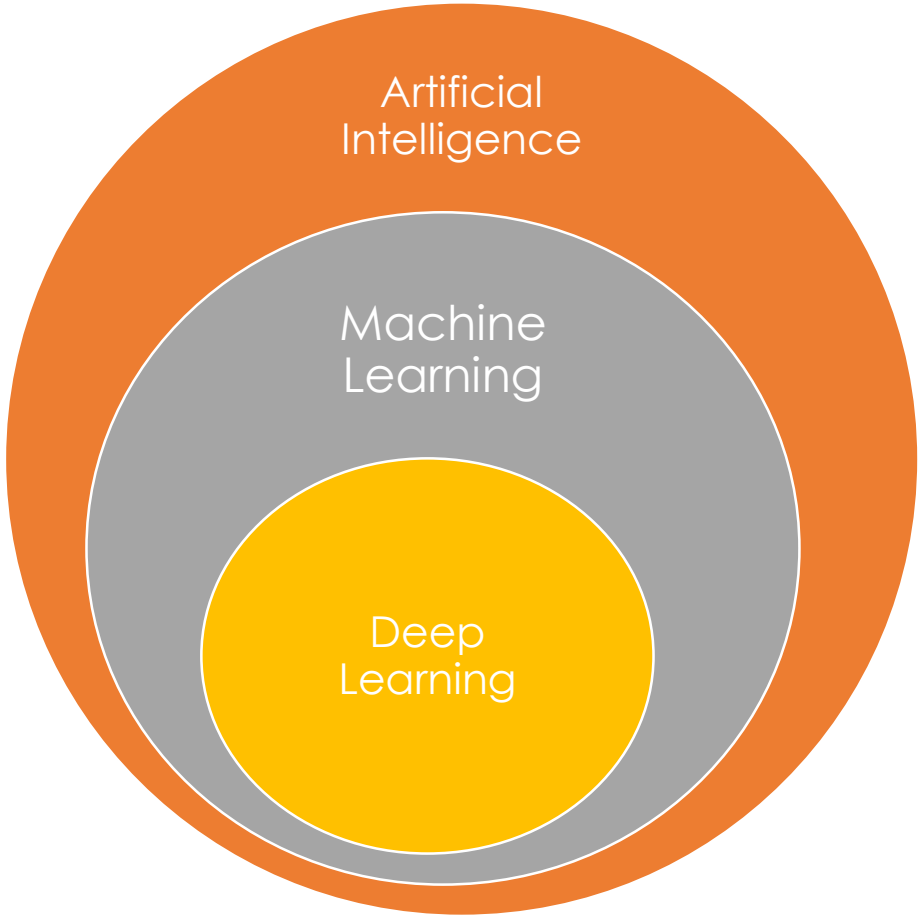
Perrine Paul-Gilloteaux

MicroPICell, BioCore

# Outline

- Qu'est-ce que le deep learning?
- Qu'est-ce que le deep learning peut vous apporter?
- Comment et avec quels moyens l'utiliser?
- Comment peut-on vous faciliter son usage?
- Présentation du Pôle numérique histologie sur MicroPICell

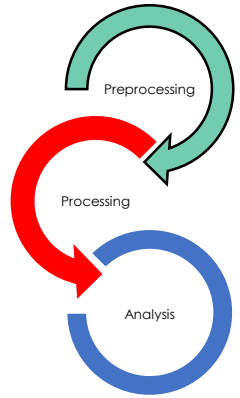
# Qu'est-ce que le deep learning?

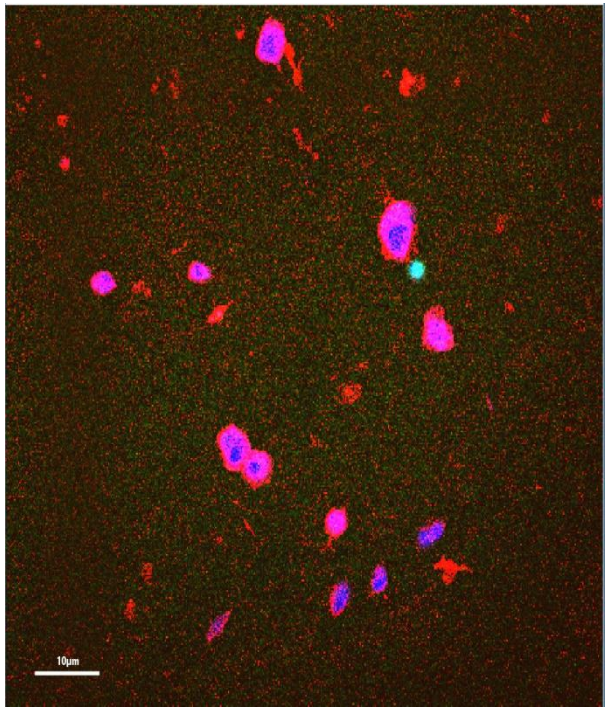


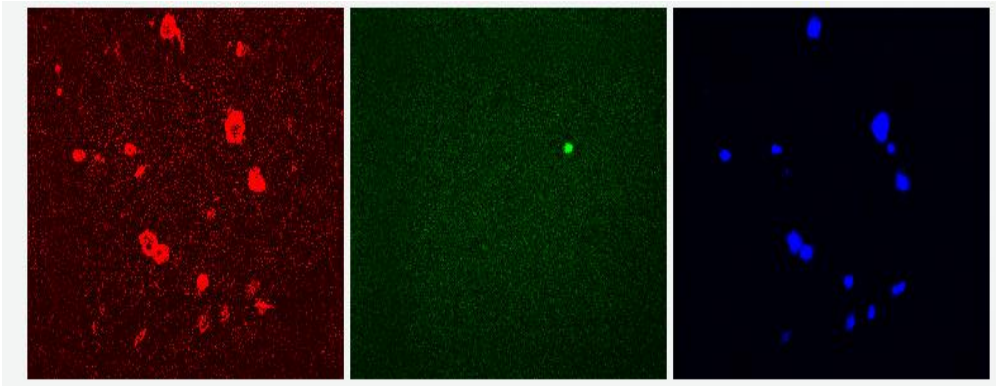

- Definition: Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine. (Larousse)

BioCore

# Segment cells: Ad hoc methods

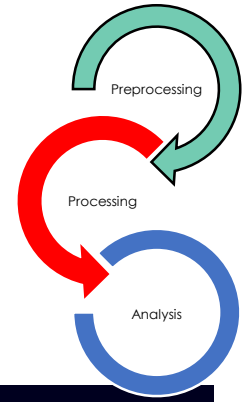


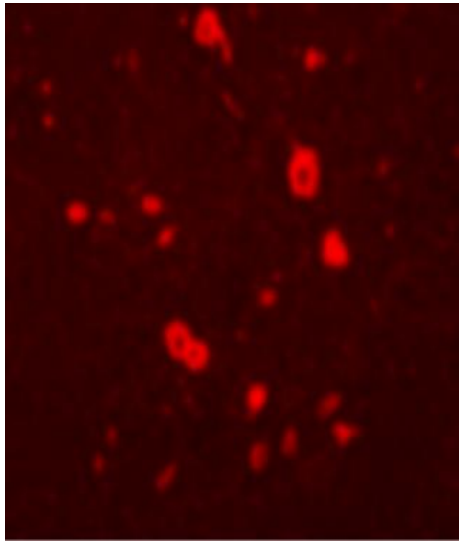



Select which channel to use

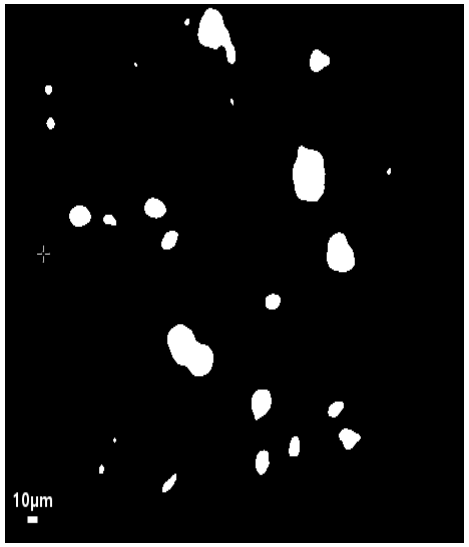
# Segment cells: Ad hoc methods



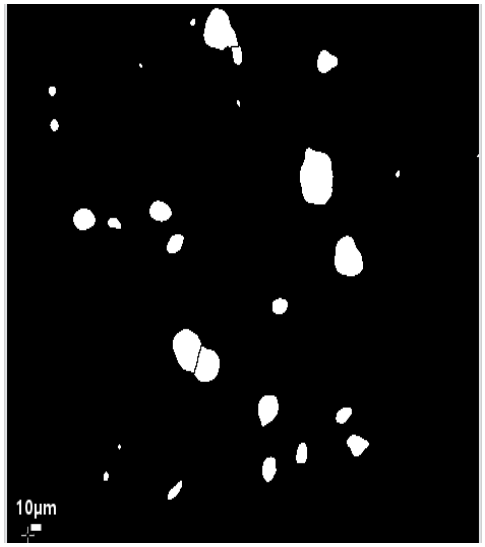

Example of workflow



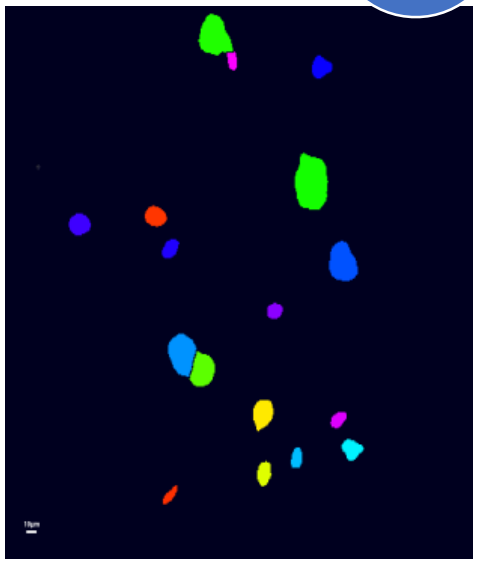
Smooth Image (Gaussian Filter Sigma 3 = convolution with gaussian kernel sigma 3)




Select a threshold or threshold method
Threshold : Kmeans 2 classes

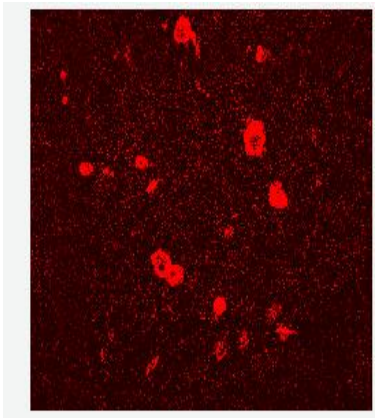


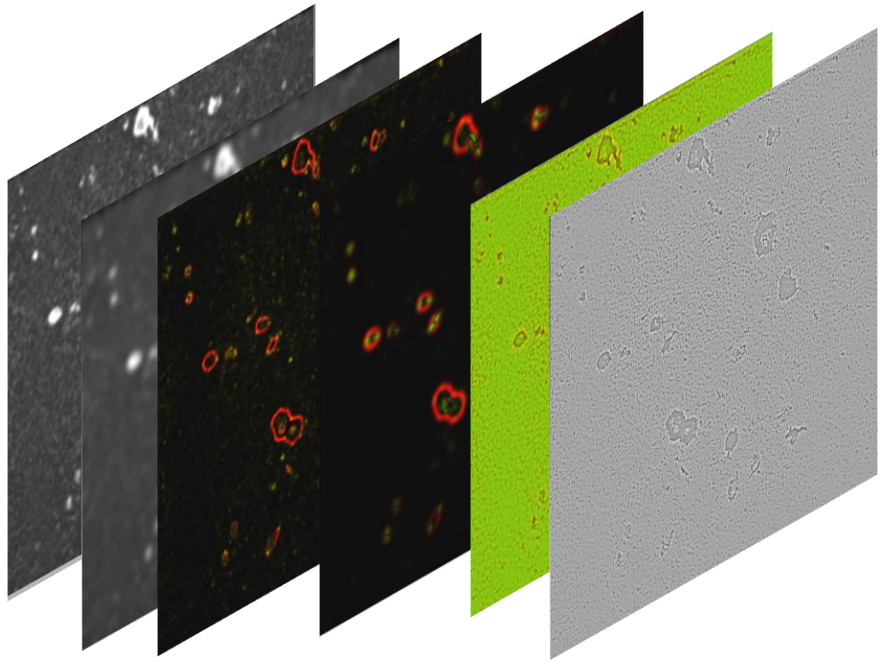

Separate touching objects: Watershed



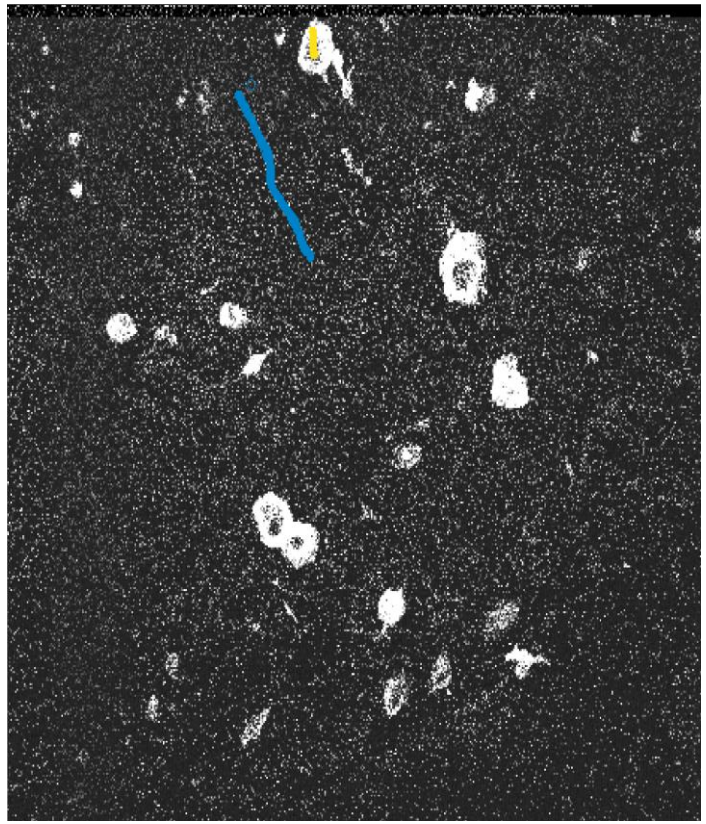
Remove small objects
Minimum size of object to keep: set threshold for size to 60 pixels

# Segment cells: Machine Learning

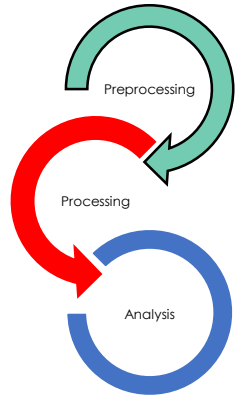




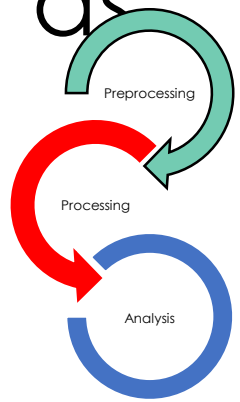
Compute predefined features (by convolution with different filters)



LABEL CLASS CELLS VS CLASS BACKGROUND

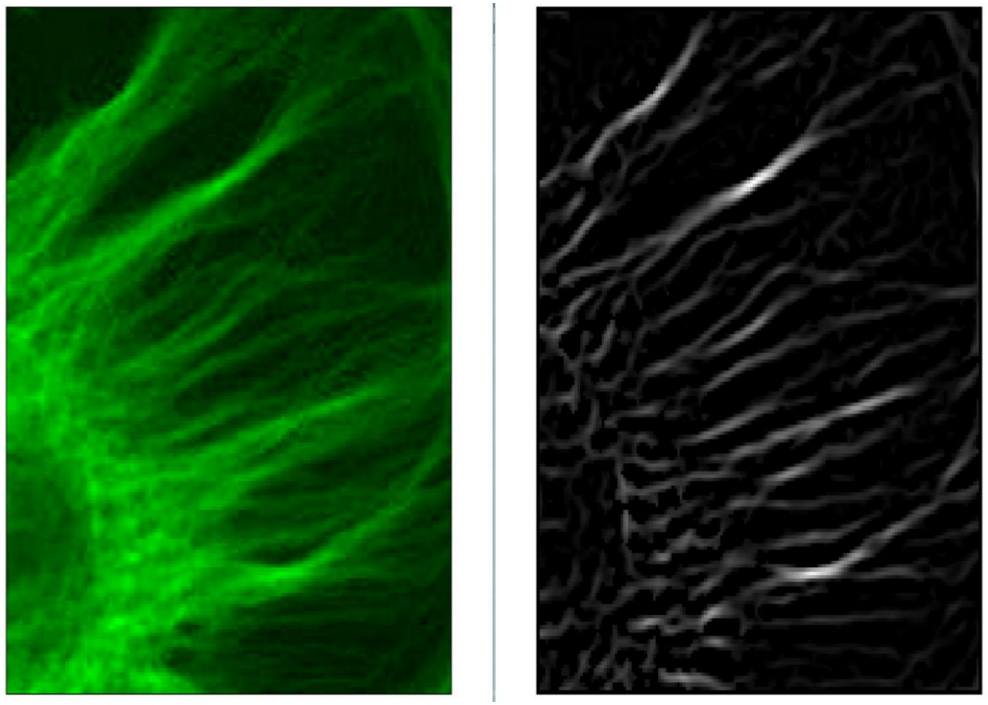


# Example with « tubes » enhancement, such as fibers, axons, cytoskeleton, mitochondrias, microtubules etc…

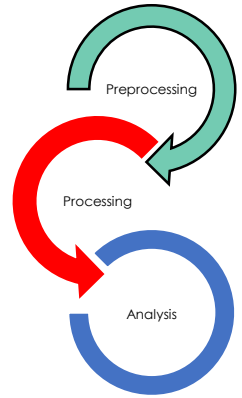


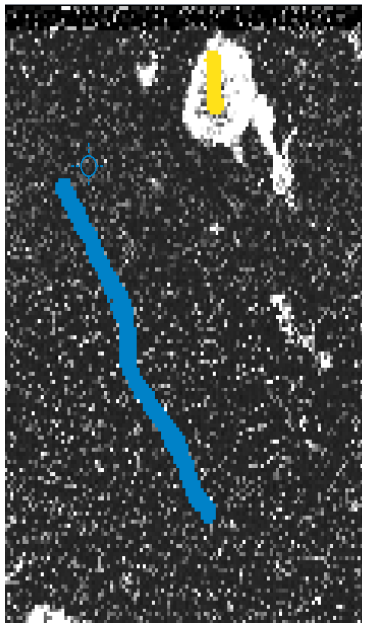

$$T(s) = \nabla^2 I(s) = \begin{bmatrix} I_{xx}(s) & I_{xy}(s) & I_{xz}(s) \\ I_{yx}(s) & I_{yy}(s) & I_{yz}(s) \\ I_{zx}(s) & I_{zy}(s) & I_{zz}(s) \end{bmatrix},$$

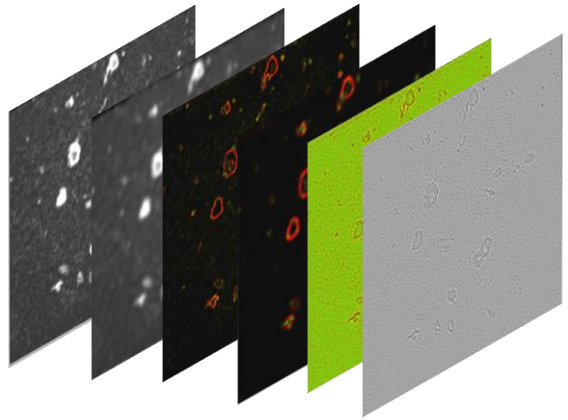
| Structure | $\lambda$ ratios | $\lambda$ constraints |
|---|---|---|
| Blob | $\lambda_1 \simeq \lambda_2 \simeq \lambda_3$ | $\lambda_1 > 0 \mid \lambda_2 > 0 \mid \lambda_3 > 0$ |
| Line | $\lambda_1 \simeq \lambda_2 \gg \lambda_3$ | $\lambda_2 > 0 \mid \lambda_3 > 0$ |
| wall | $\lambda_1 \gg \lambda_2 \simeq \lambda_3$ | $\lambda_3 > 0$ |



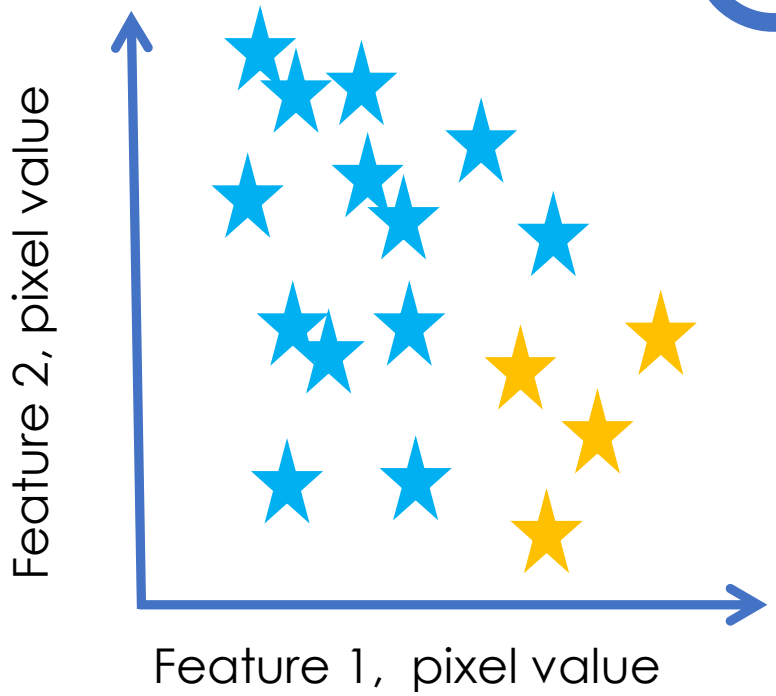
# Segment cells: Machine Learning



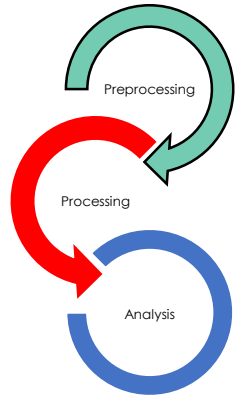



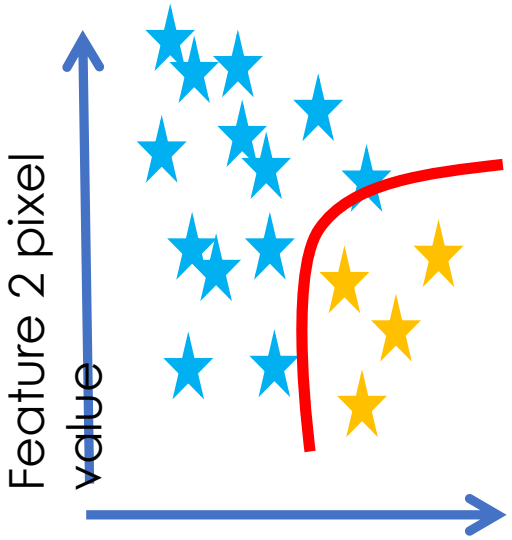


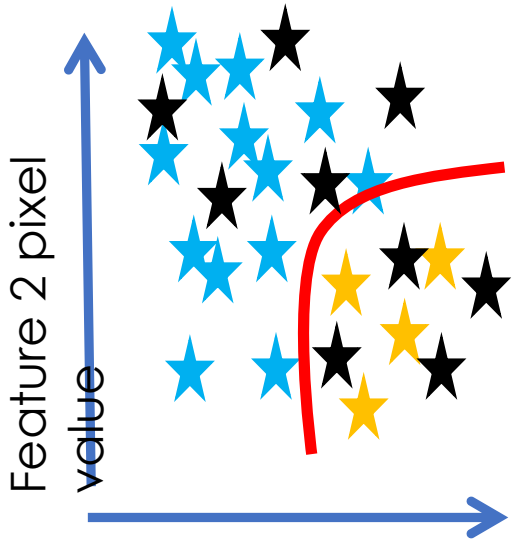
Compute predefined features (by convolution with different filters)



# Segment cells: Machine Learning

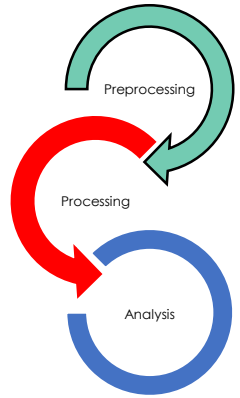


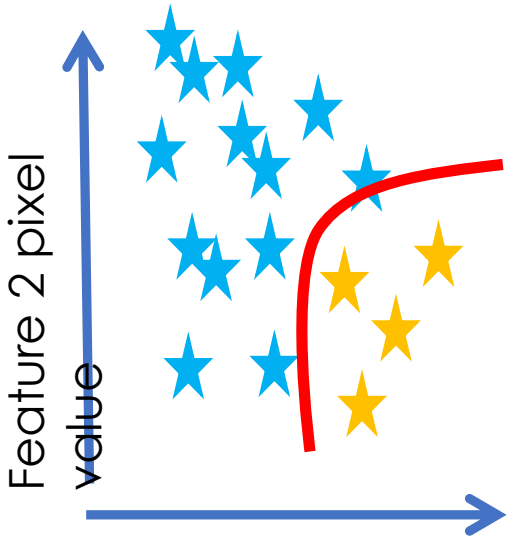



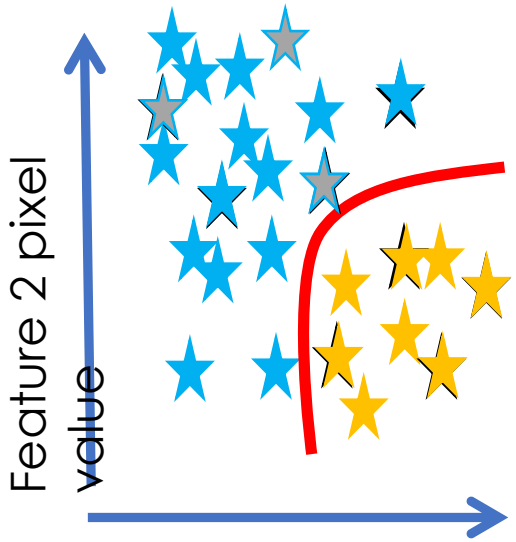

Classifier Model

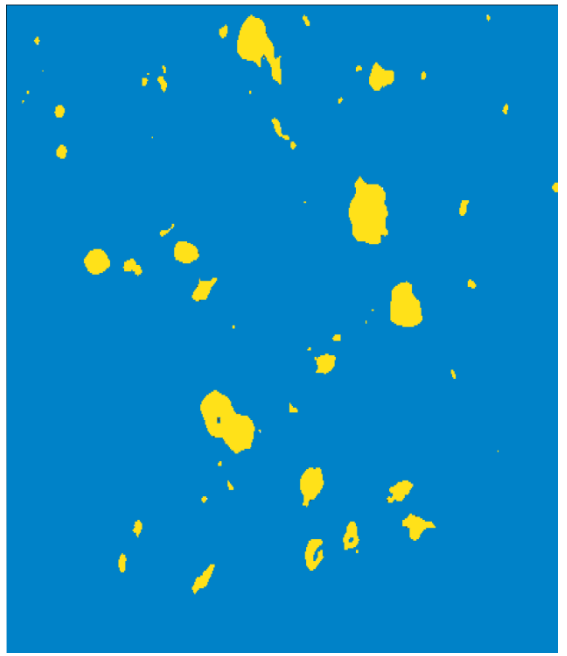
# Segment cells: Machine Learning



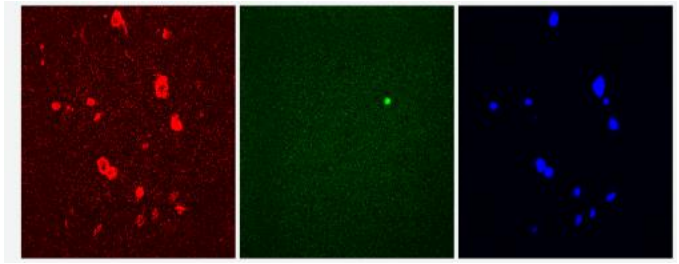



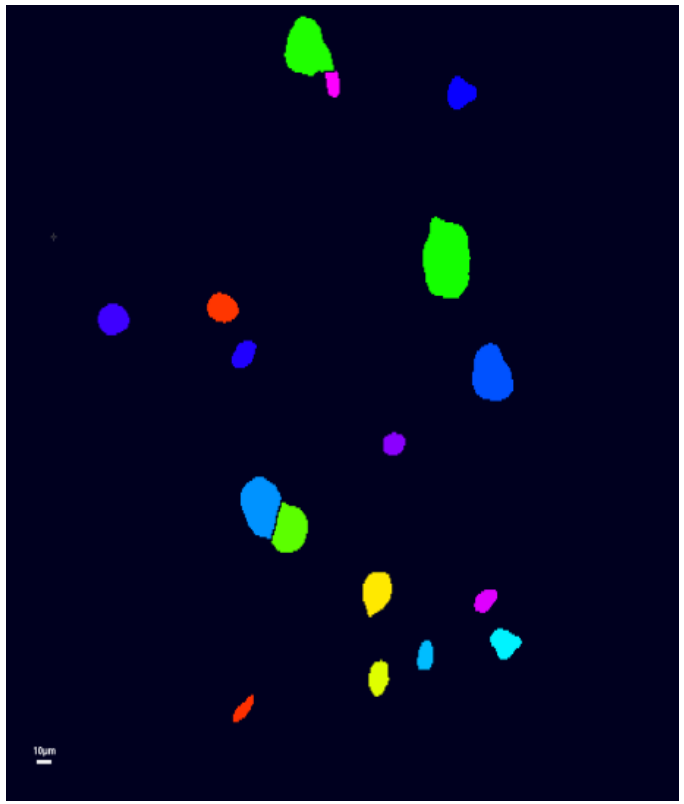

Classifier Model

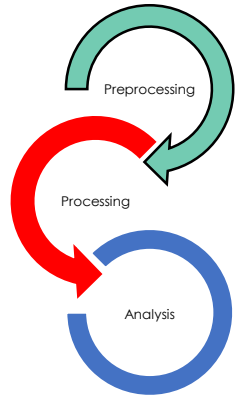


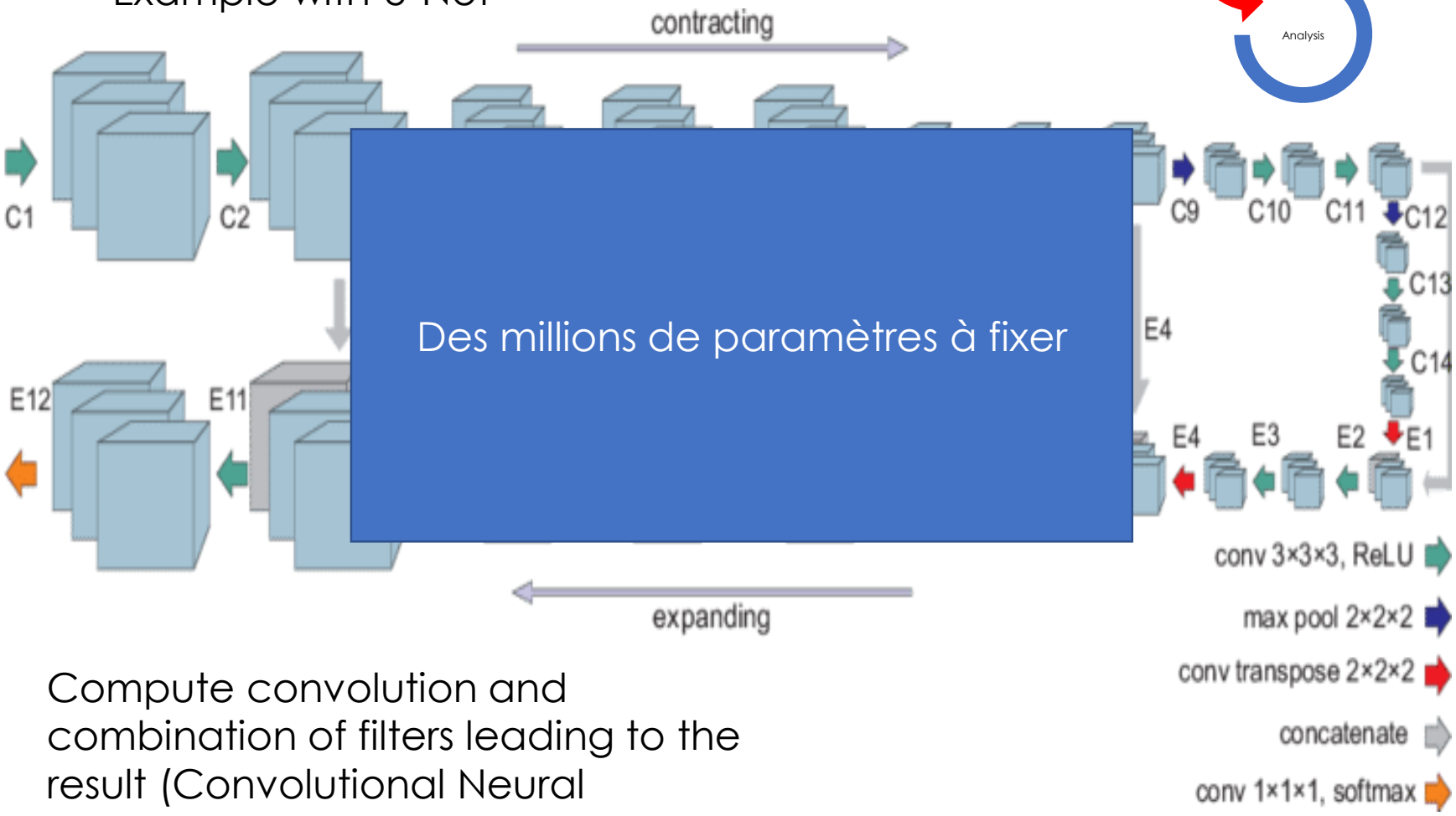
# Segment cells: Deep Learning





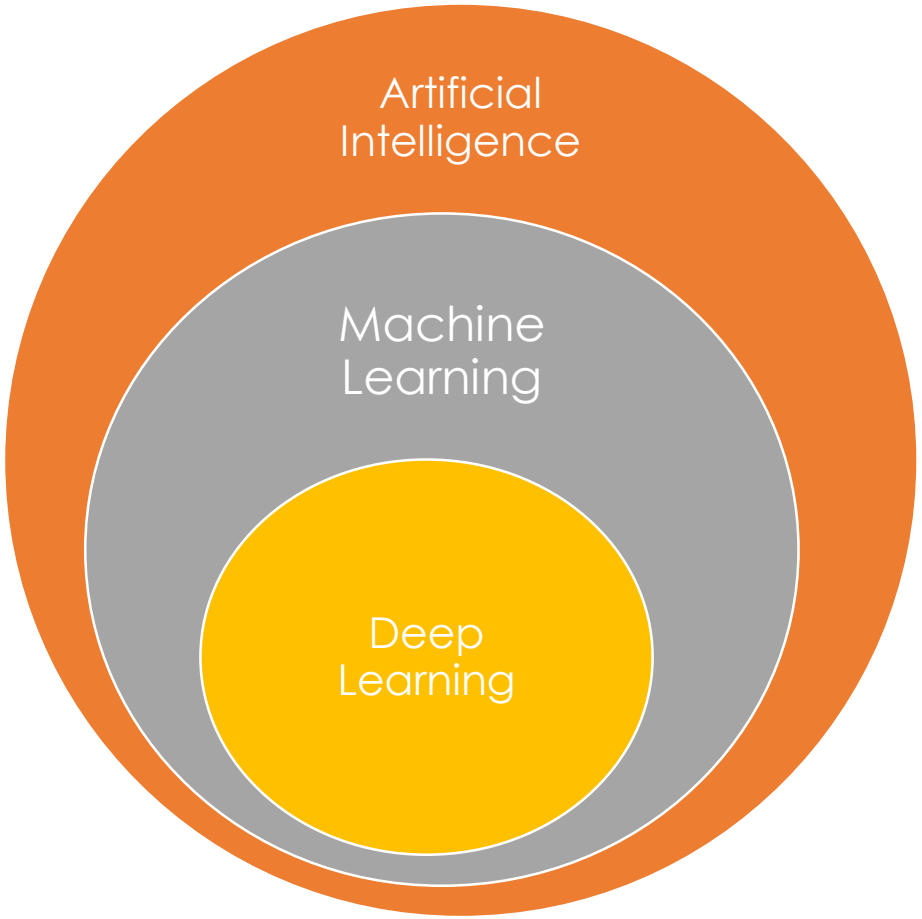



Example with U-Net

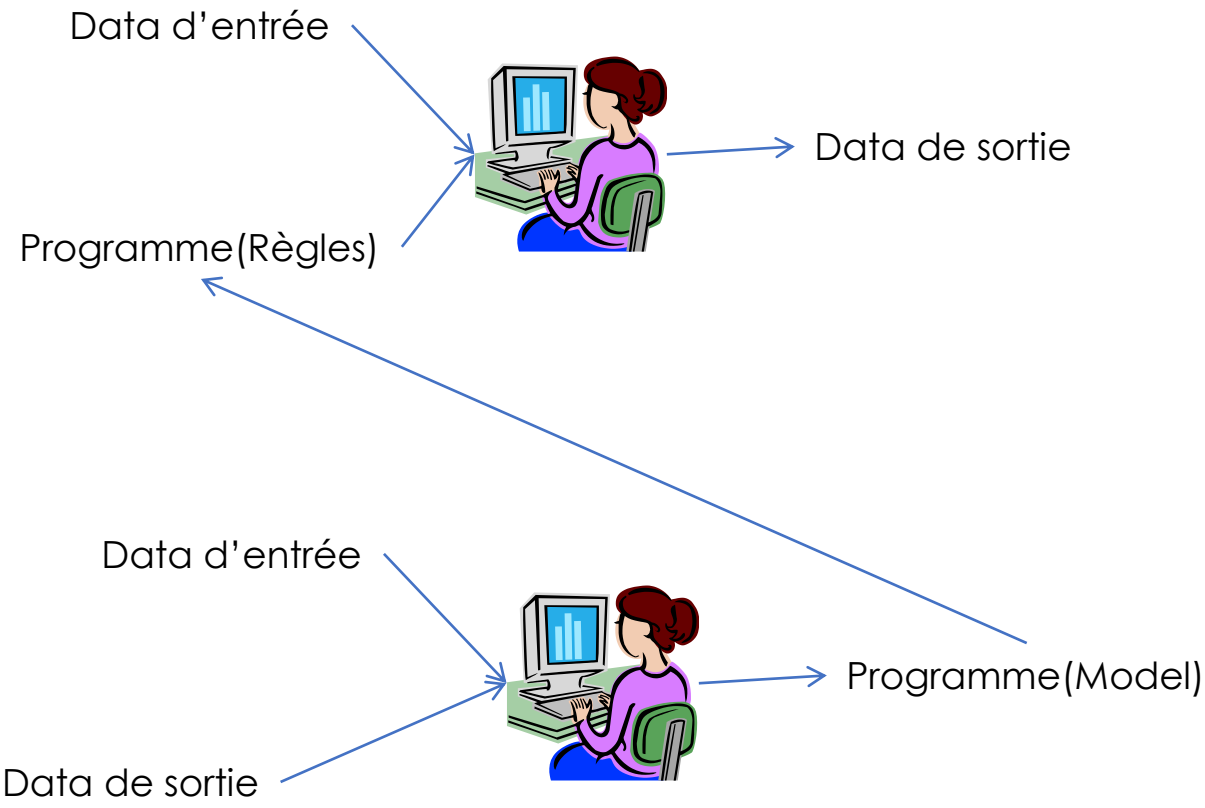



Compute convolution and combination of filters leading to the result (Convolutional Neural Networks)

# Qu'est-ce que le deep learning?

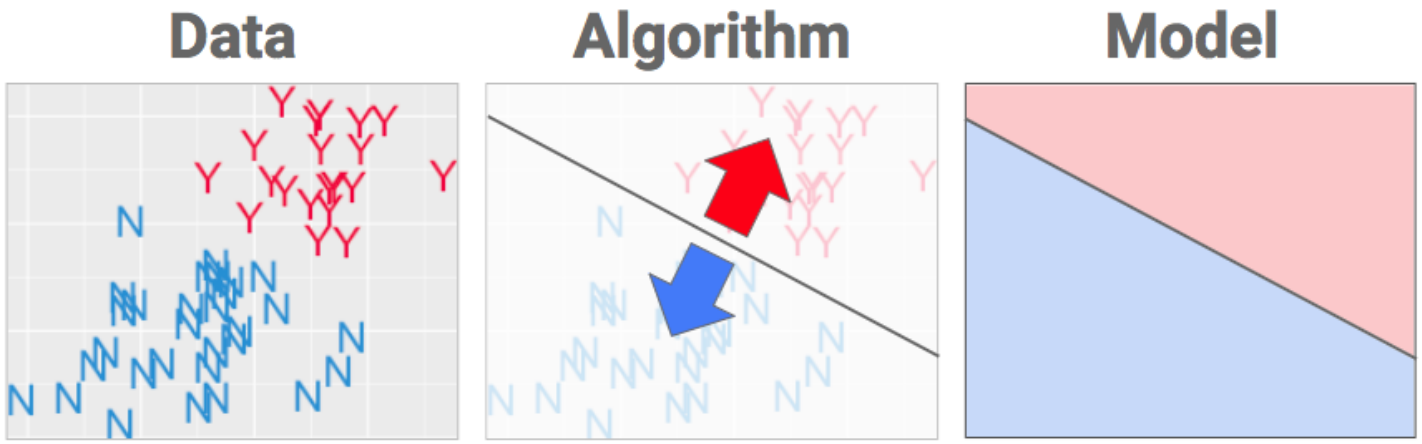



- Changement de paradigme

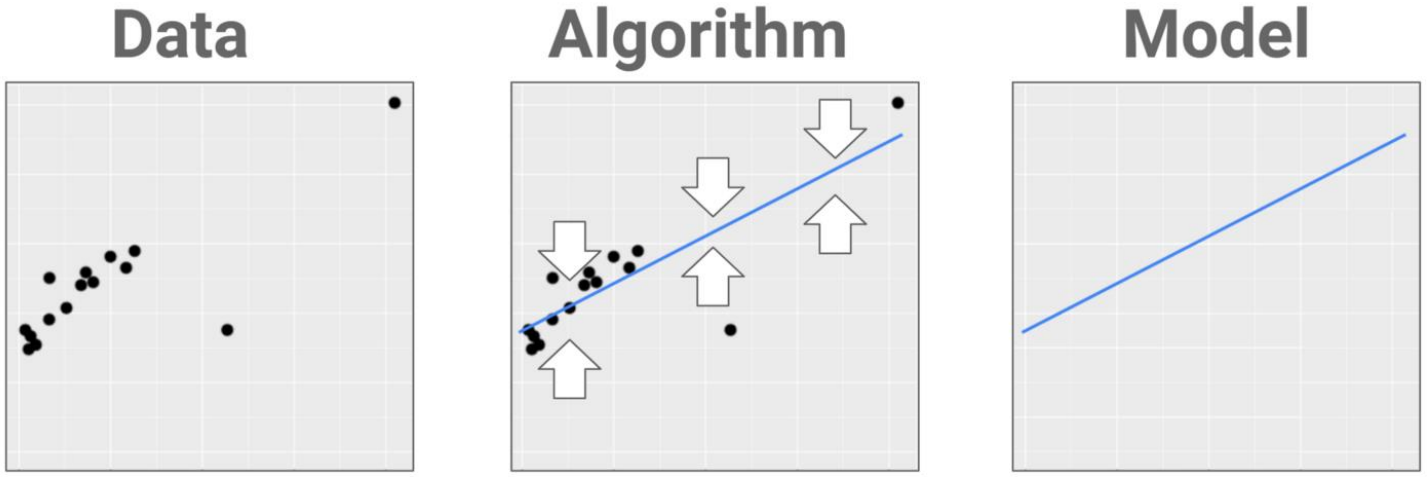



BioCore

# Classification vs Prediction

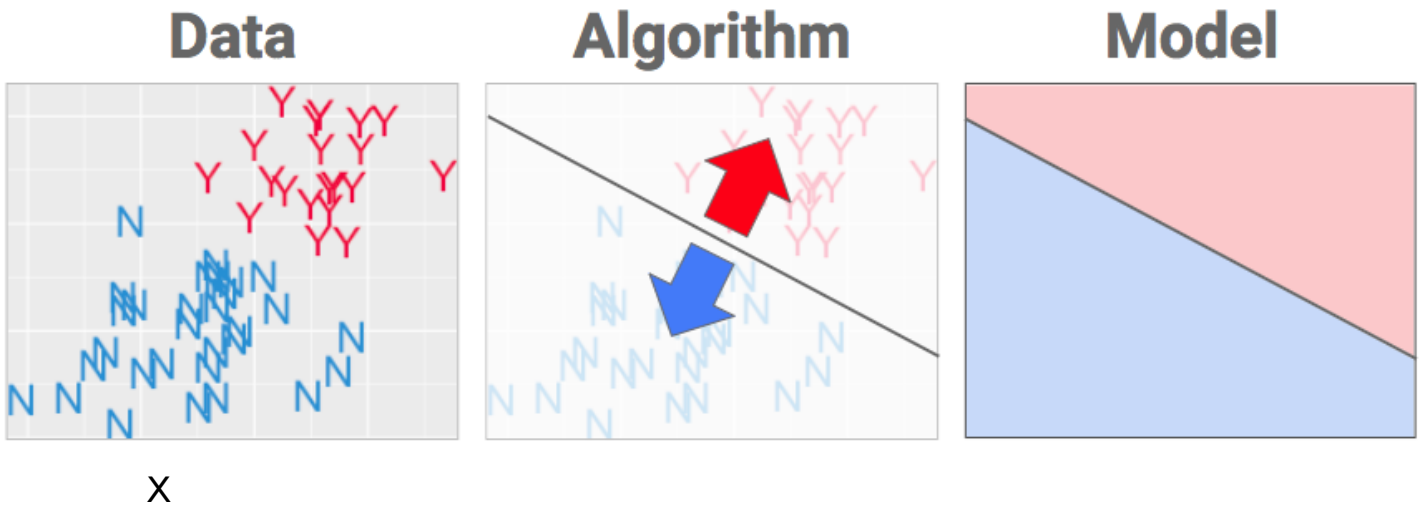



Classification

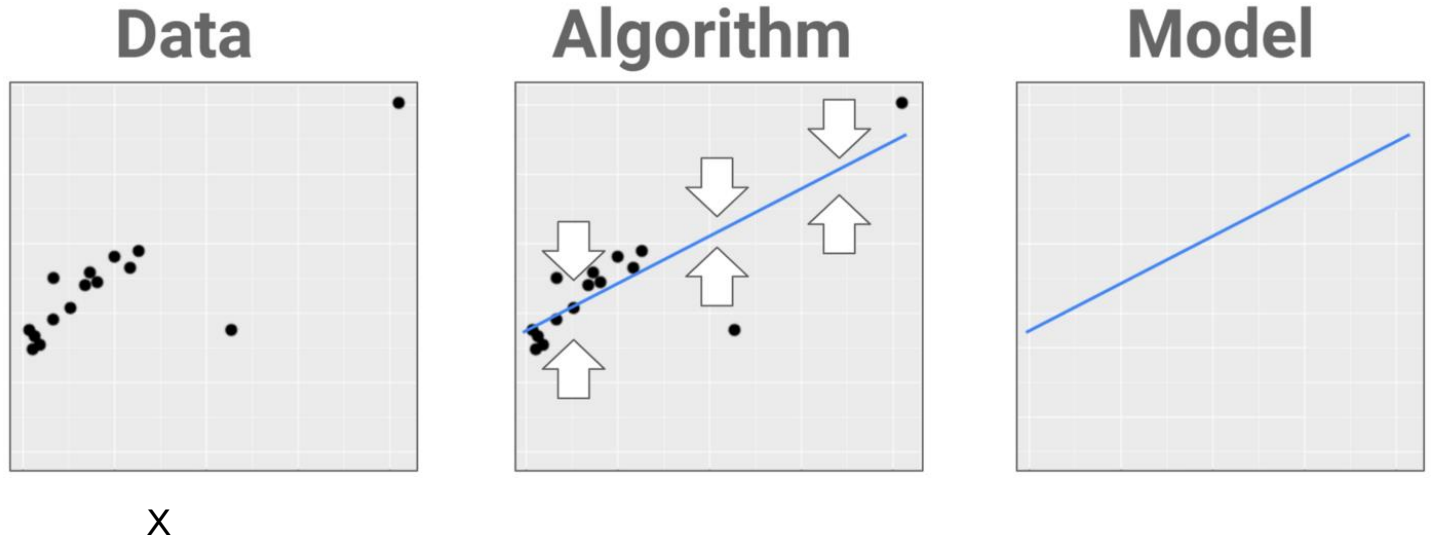



Prediction

# Features and Model




Feature 1=x
Feature 2=y
Output = Yes or No



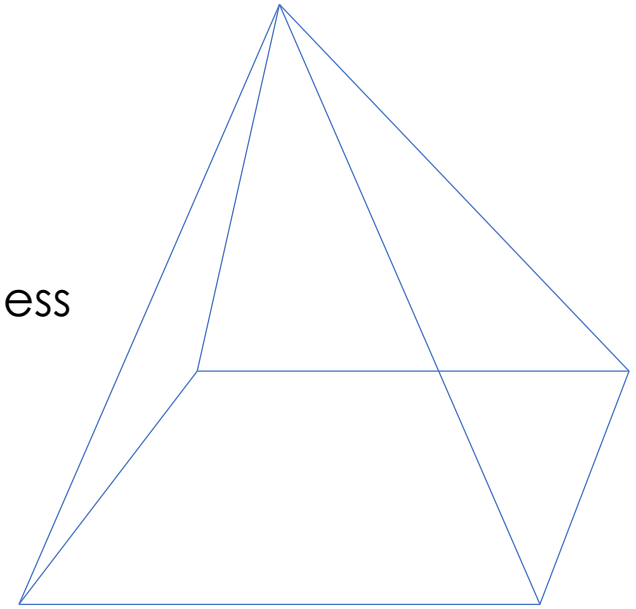

Feature
1=x
Output= y

# Plus d'info:

- Fidle.cnrs.fr

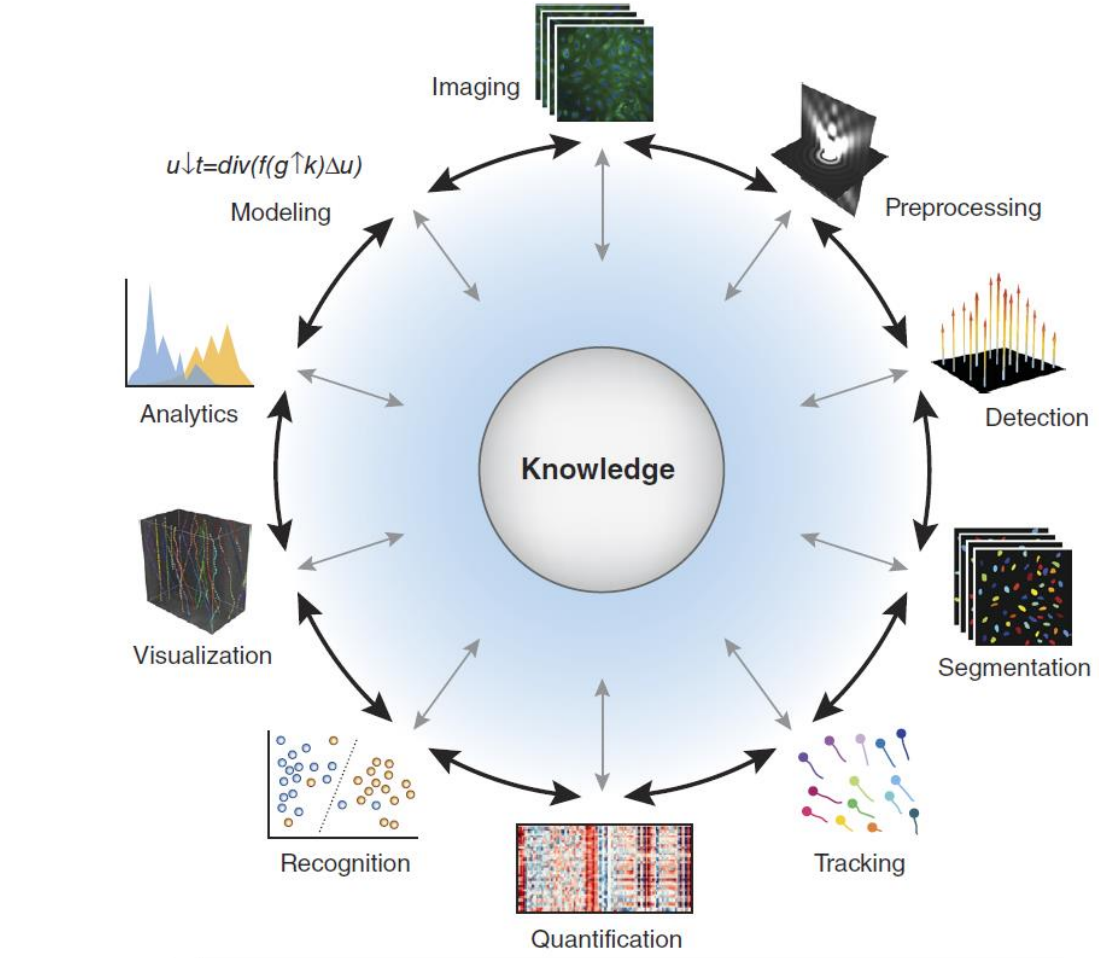

- Formation régionale CNRS INSERM 2022 deep learning pour la microscopie en biologie (Sensibilisation/appliquer un modèle/entrainer un modèle)
Vidéos sur demande, renouvellement probable 2024) T. Pécot, M. Feyeux, P. Paul-Gilloteaux

# Qu'est-ce que le deep learning peut vous apporter?




Less Toxicity, less invasiveness

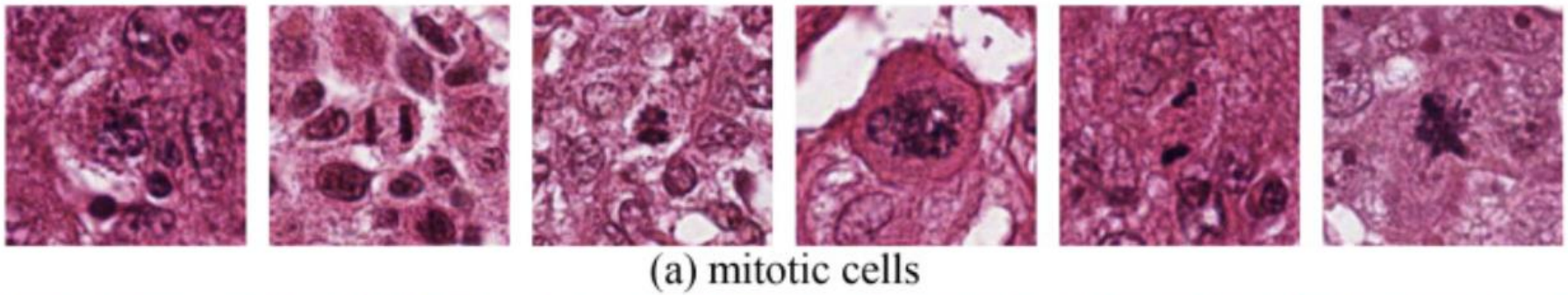
Field of view (size + in depth imaging)




BioCore

# Examples of classification tasks

- Image Classification:

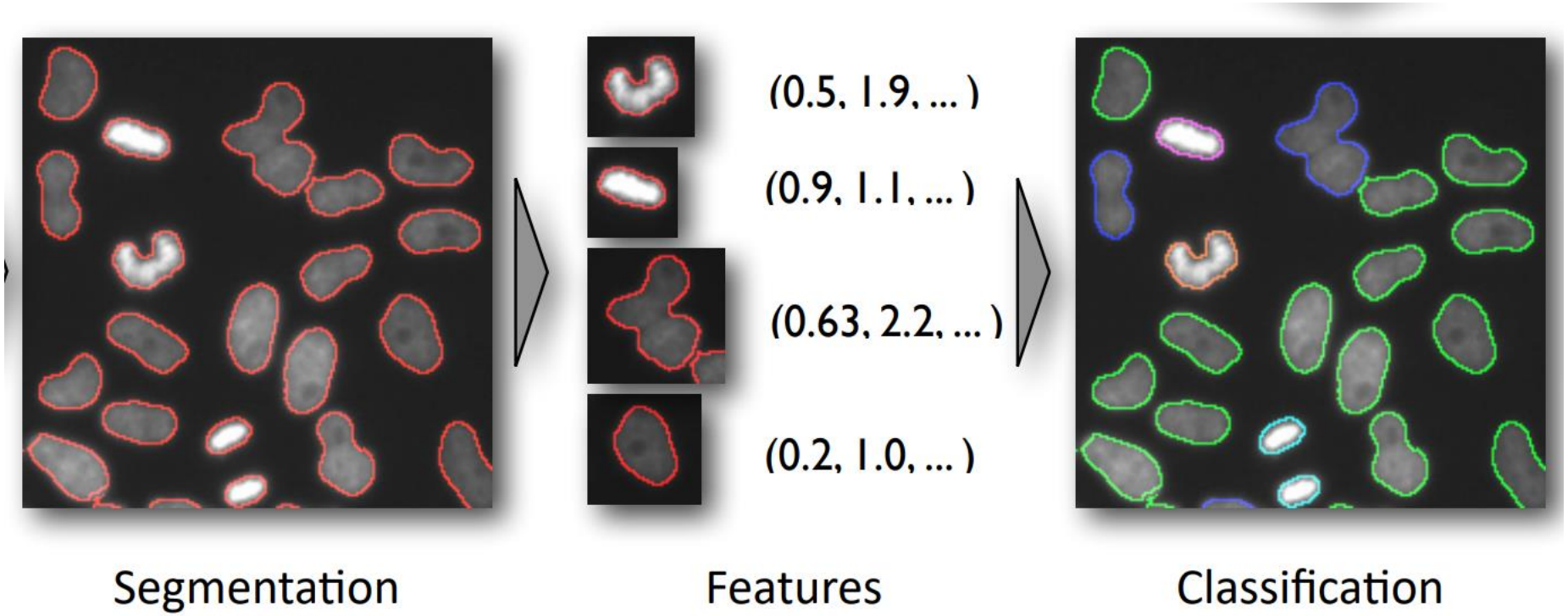


(a) mitotic cells

(b) non-mitotic cells

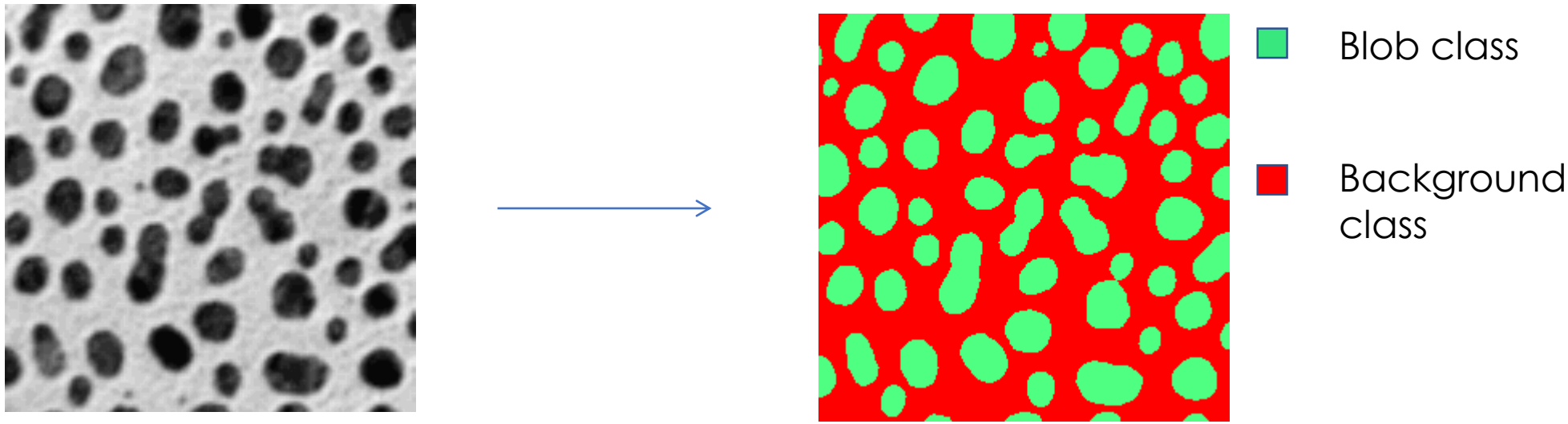
# Examples of classification tasks

- Object Classification:



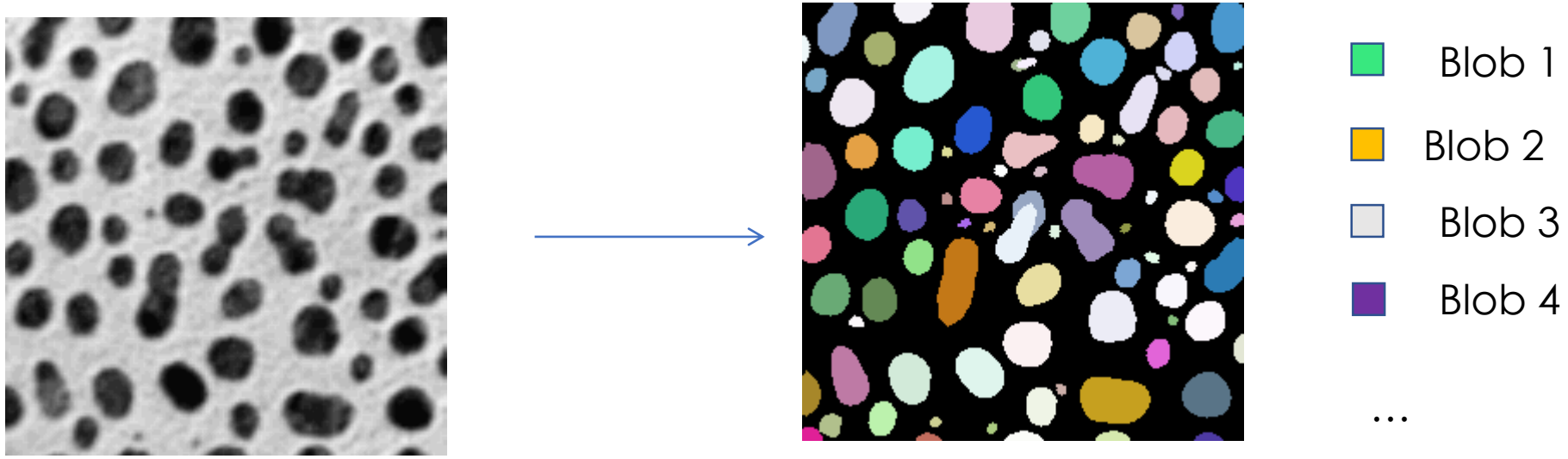
# Examples of classification tasks

- Semantic Segmentation: classify pixels in image in class



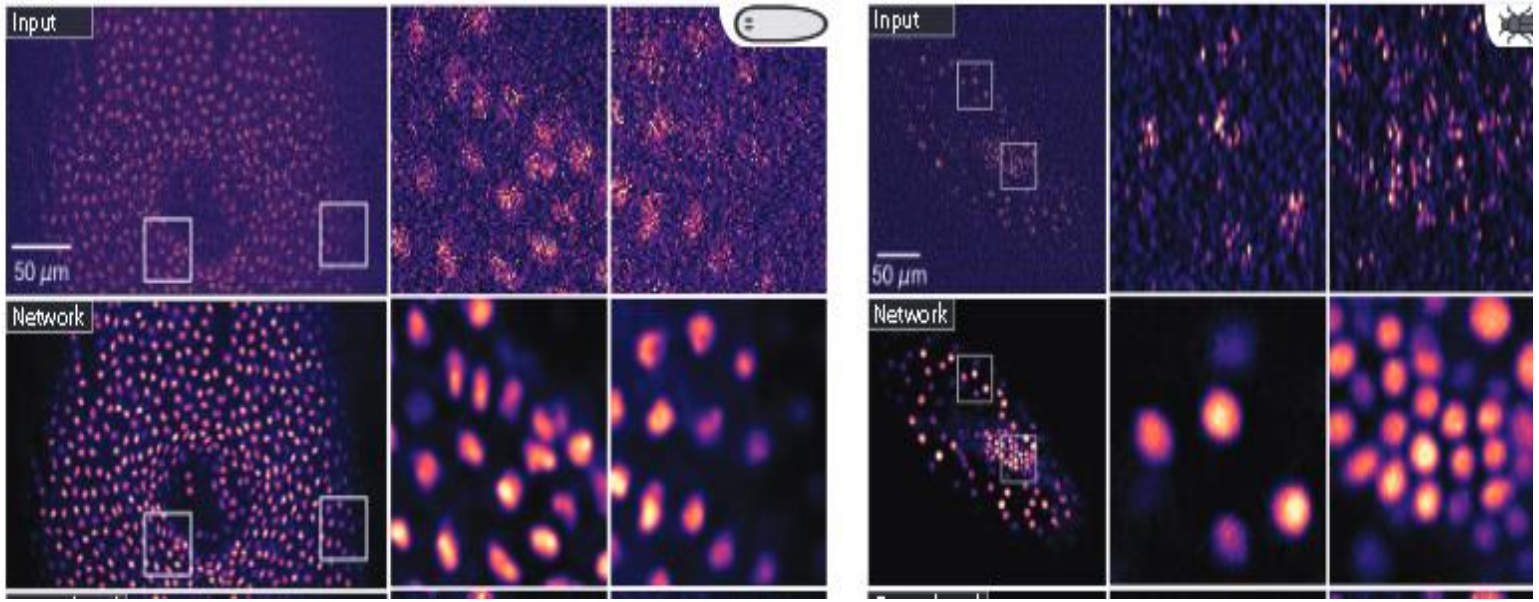
# Examples of classification tasks

- Instance Segmentation: classify pixels in image in object (instances)



# Examples of Prediction tasks

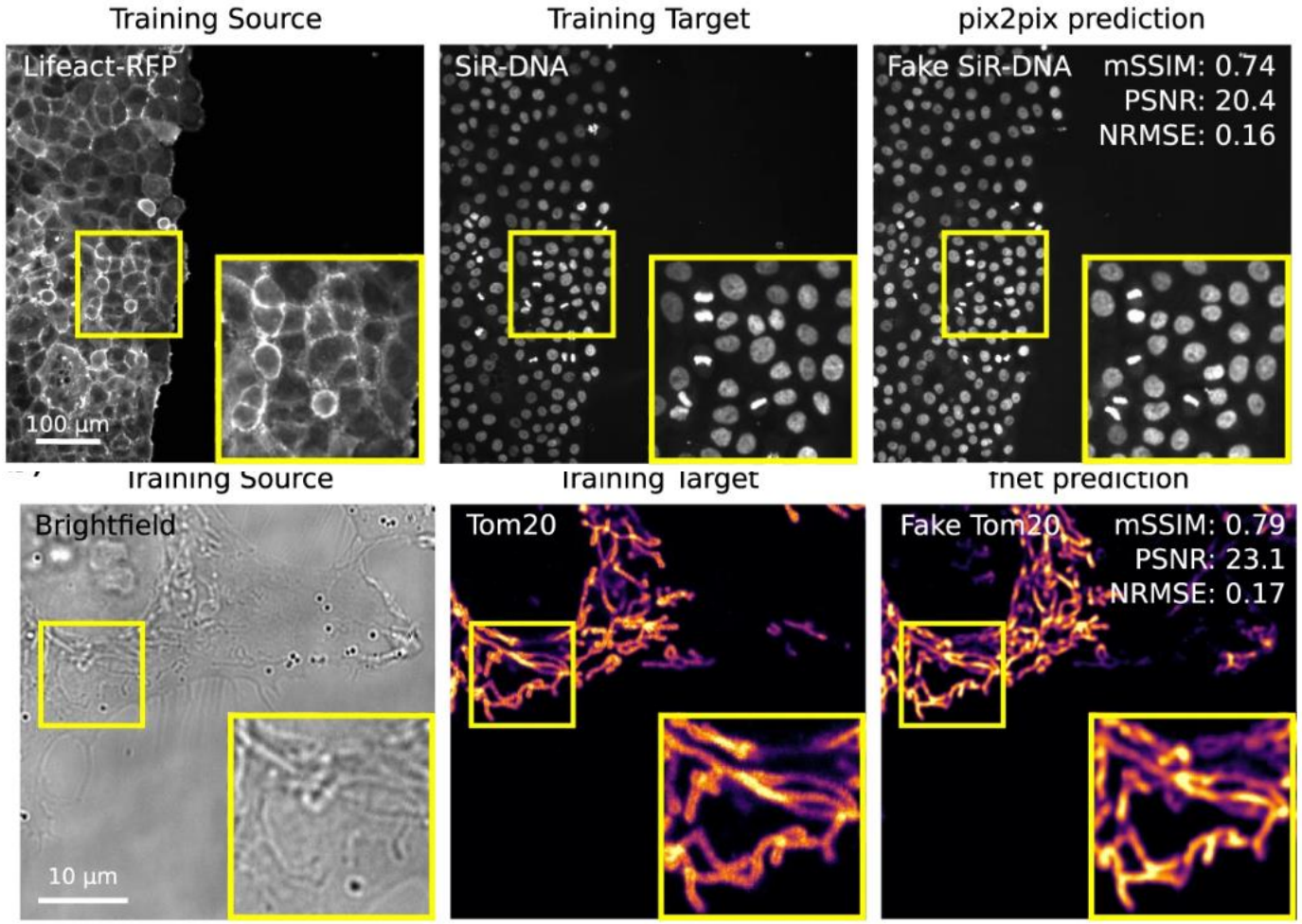
- Image restoration (denoising): predict the new value of pixels




Weigert, M., Schmidt, U., Boothe, T. *et al.* Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat Methods* **15,** 1090–1097 (2018). https://doi.org/10.1038/s41592-018-0216-7

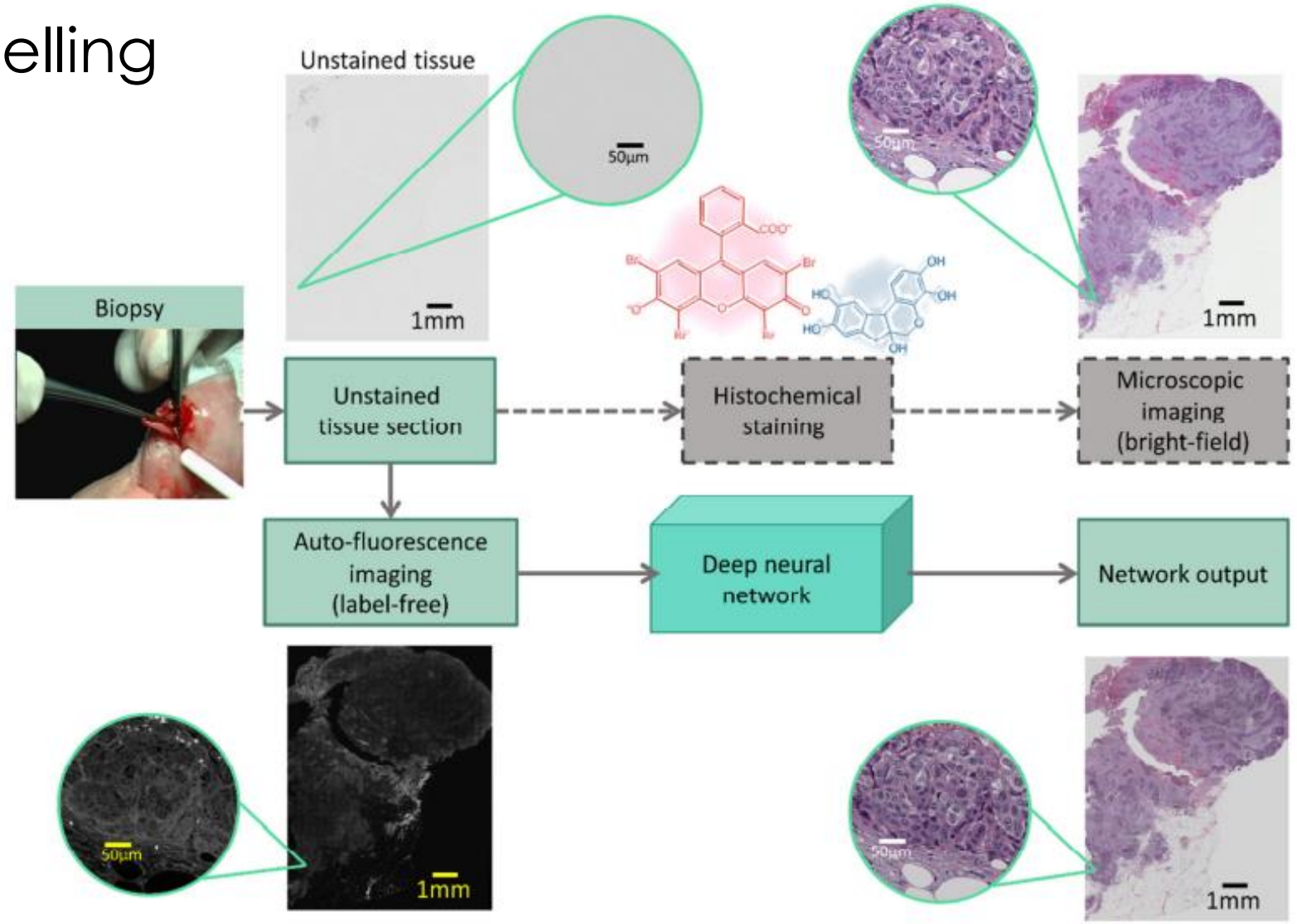
# Examples of Prediction tasks

- Image translation (predict another modality): predict the new value of pixels (style transfer)



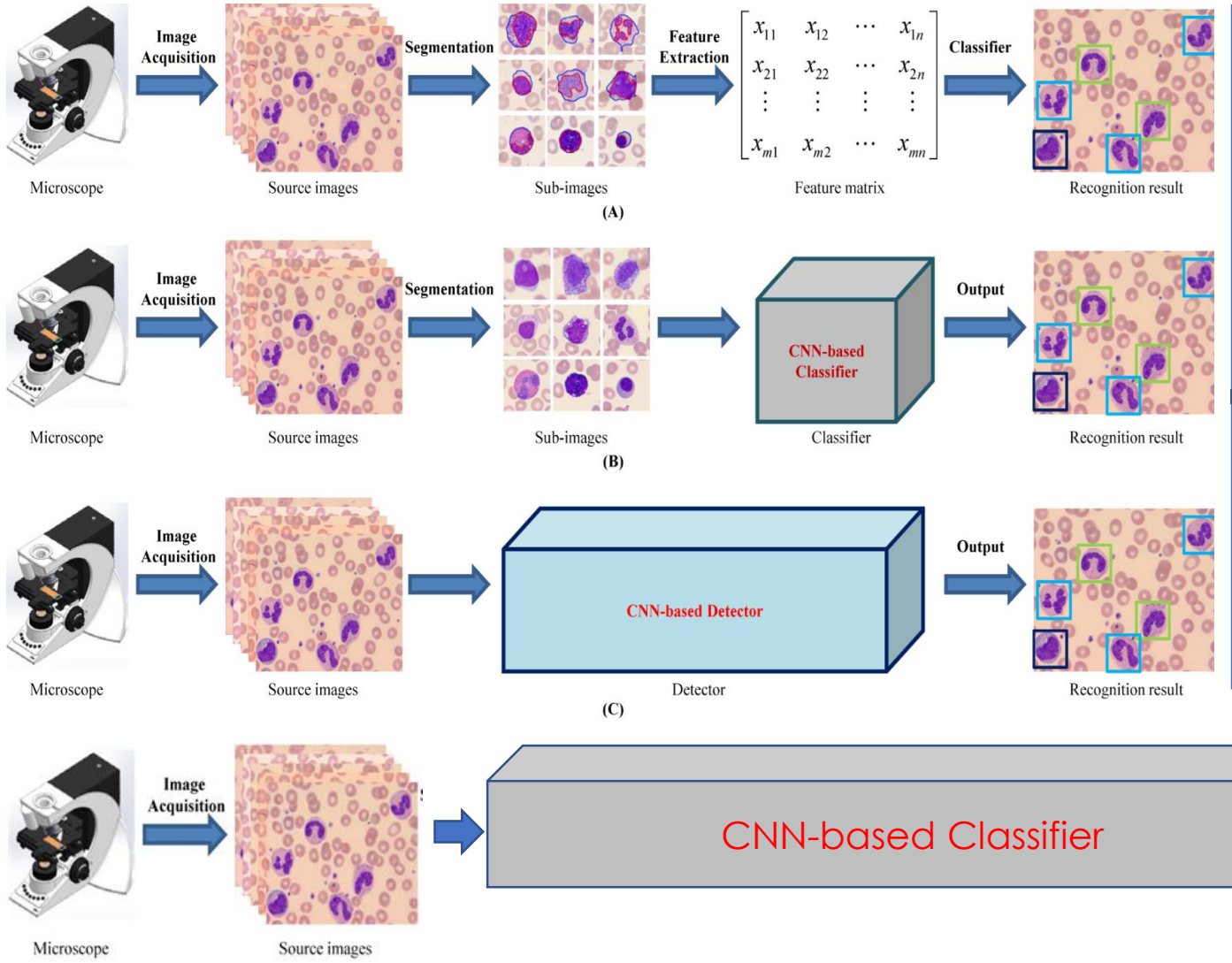

von Chamier, L., Laine, R.F., Jukkala, J. *et al.* Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nat Commun* **12,** 2276 (2021). https://doi-org.insb.bib.cnrs.fr/10.1038/s41467-021-22518-0

# In silico labelling




Rivenson et al. Nat Biomed Eng 3, 466–477 (2019). https://doi.org/10.1038/s41551-019-0362-y

# AI can be everywhere




Human defined algorithm

or

CNN-based Classifier

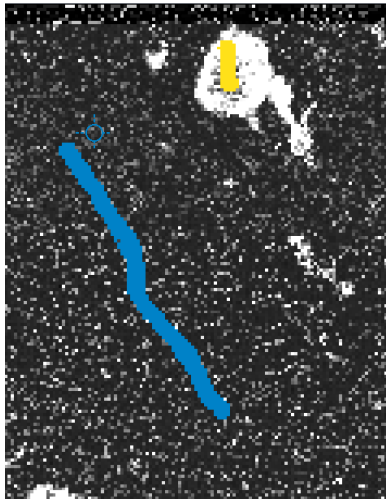
Sample classification (effect of drugs/longitudinal studies pathologies, biomarkers discovery..)

Sample classification (effect of drugs/longitudinal studies/ pathologies, biomarkers discovery..)

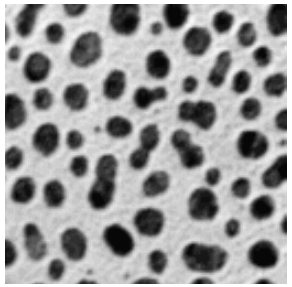
# Training datasets

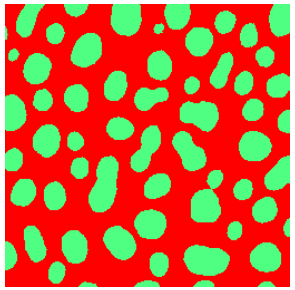
Manually or obtained (Manual annotation) or by acquisition strategy

- Dataset used to fit the algorithm
- Example:

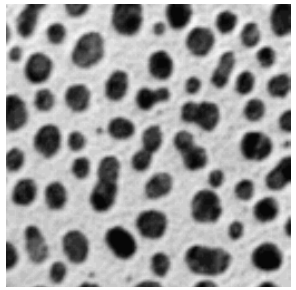


For pixel classification (semantic segmentation with machine learning)

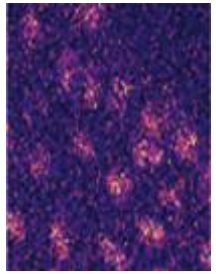


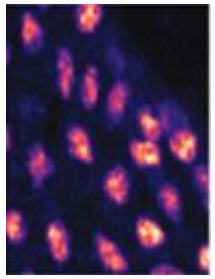


For pixel classification (semantic segmentation with Deep Learning)



For pixel classification (semantic segmentation with Deep Learning)



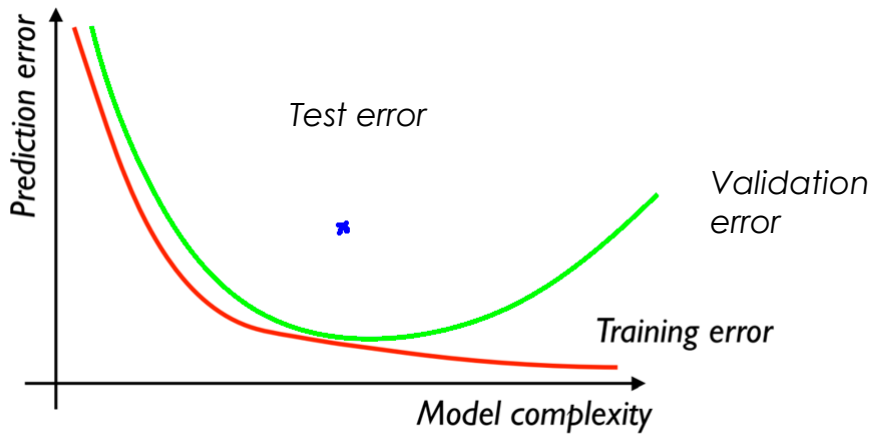


For denoising: Images acquired with different settings of the microscope

**If you test your algorithm/model on the training data set: you will get very good results. But what will happen on new dataset?**

# Cross validation

- Train or tune your algorithm on **training dataset**, and then test it on **test dataset : both are annotated/labelled.**




Test data

Labelled data 4

Validation data

Training data

| | | |
|---|---|---|
| Labelled data1 | Labelled data 2 | Labelled data 3 |
| Labelled data 1 | Labelled data 2 | Labelled data 3 |
| Labelled data 1 | Labelled data 2 | Labelled data 3 |

# Comment et avec quels moyens l'utiliser?

Créer une nouvelle architecture

Ré-Entrainer un modèle

Appliquer un modèle

Senior Data scientist

Thèse

Maîtrise avancée du python et de l'ecosysteme Keras/TensorFlow ou pytorch

Maitrise modérée python

Gros moyens de calculs

Des millions de data (>1000 images ANNOTEES)

BioCore

# Comment et avec quels moyens l'utiliser?

Créer une nouvelle architecture

Ré-Entrainer un modèle

Appliquer un modèle

Senior Data scientist

Thèse

Maîtrise avancée du python et de l'ecosysteme Keras/TensorFlow ou pytorch

Maitrise modérée python

Gros moyens de calculs

Des millions de data (>1000 images ANNOTEES)

BioCore

# Comment et avec quels moyens l'utiliser?

Créer une nouvelle architecture

Ré-Entrainer un modèle

Appliquer un modèle

Senior Data scientist

Thèse

Maîtrise avancée du python et de l'ecosysteme Keras/TensorFlow ou pytorch

Maitrise modérée python

Gros moyens de calculs

Des millions de data (>1000 images ANNOTEES)

BioCore

# Comment et avec quels moyens l'utiliser?

| | |
|---|---|
| Créer une nouvelle architecture | Senior Data scientist<br>Thèse |
| Ré-Entrainer un modèle<br>Transfer Learning | Maîtrise avancée du python et de l'ecosysteme Keras/TensorFlow ou pytorch<br>Maitrise modérée python<br>Gros moyens de calculs<br>~~Des millions de data (>1000 images ANNOTEES)~~ Quelques Images annotées |
| Appliquer un modèle | |

BioCore

# Comment et avec quels moyens l'utiliser?

Créer une nouvelle architecture

Ré-Entrainer un modèle

Appliquer un modèle

Senior Data scientist

Thèse

Maîtrise avancée du python et de l'ecosysteme Keras/TensorFlow ou pytorch

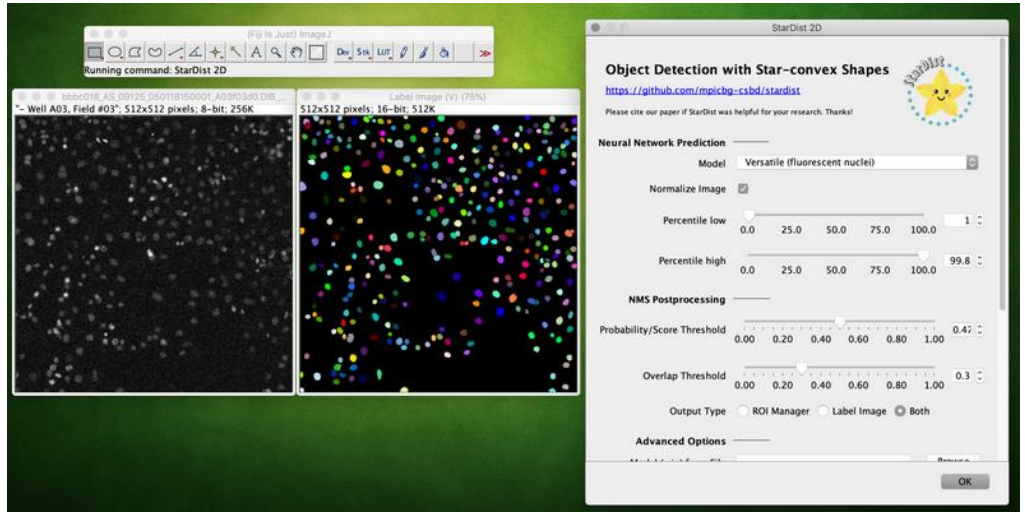
Maitrise modérée python

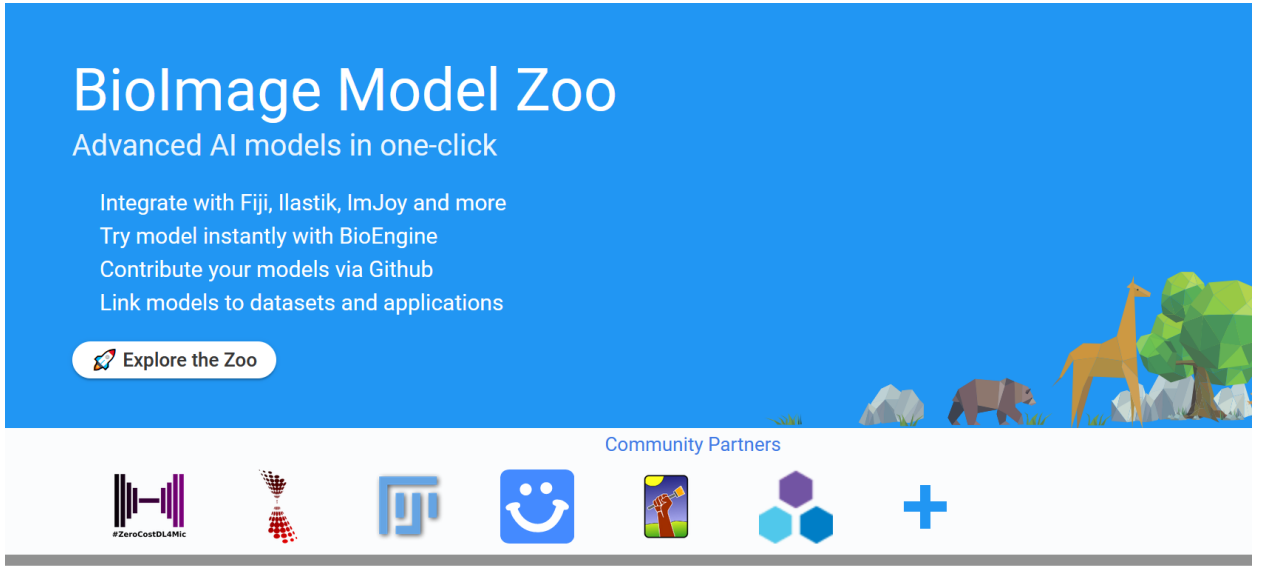
Gros moyens de calculs

Des millions de data (>1000 images ANNOTEES)

BioCore

# Pre-Trained Models



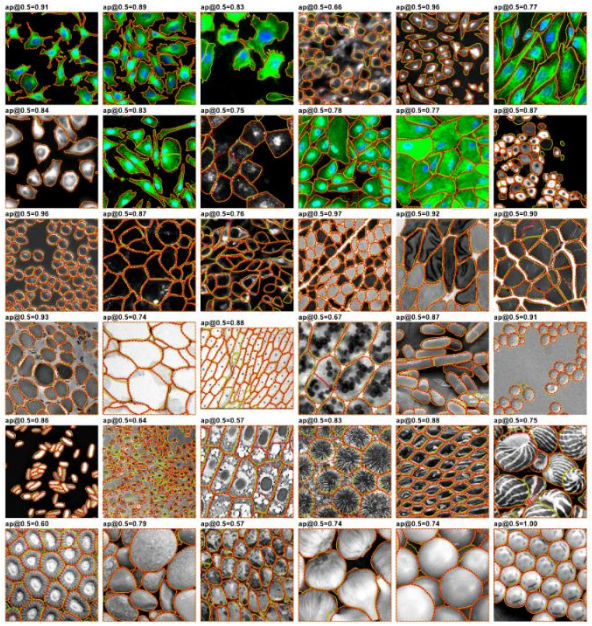
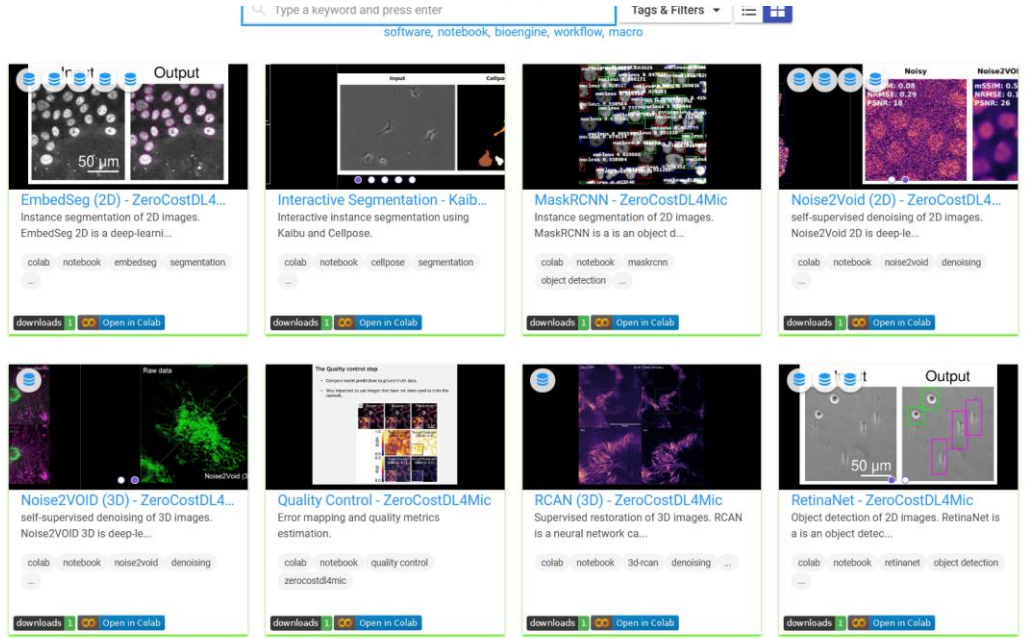



Figure 4: **Example cellpose segmentations for 36 test images.** The ground truth masks segmented by a human operator are shown in yellow, and the predicted masks are shown in dotted red line. Compare to Stardist and Mask-RCNN in Figure S3 and Figure S4.



# Comment peut-on faciliter son usage?

Créer une nouvelle architecture

Ré-Entrainer un modèle

- Faciliter le partage des données avec des collaborateurs (par exemple LS2N ou extérieur) au plus proche des ressources de calcul
- Faciliter la réalisation des annotations et les partager de manière normalisé (également open-science friendly)
- Assurer un haut débit et une uniformité des datas
- Proposer des solutions d'annotations « automatique »

BioCore

# Comment peut-on faciliter son usage?

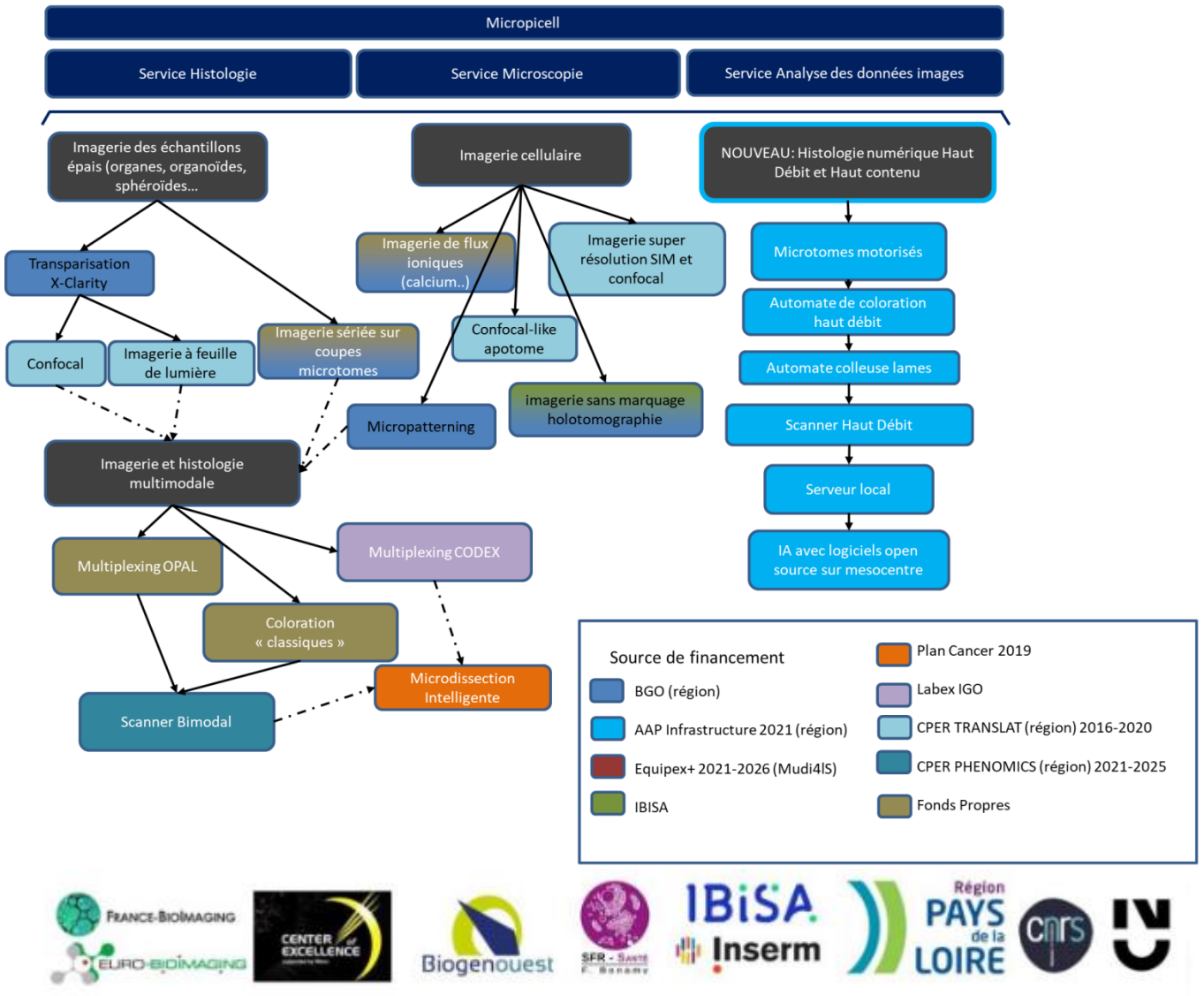
Appliquer un modèle

- Faciliter le partage des modèles

- Faciliter leur application sur les données sans installation d'environnement et en un click.

BioCore

# Le pôle histologie numérique MicroPICell

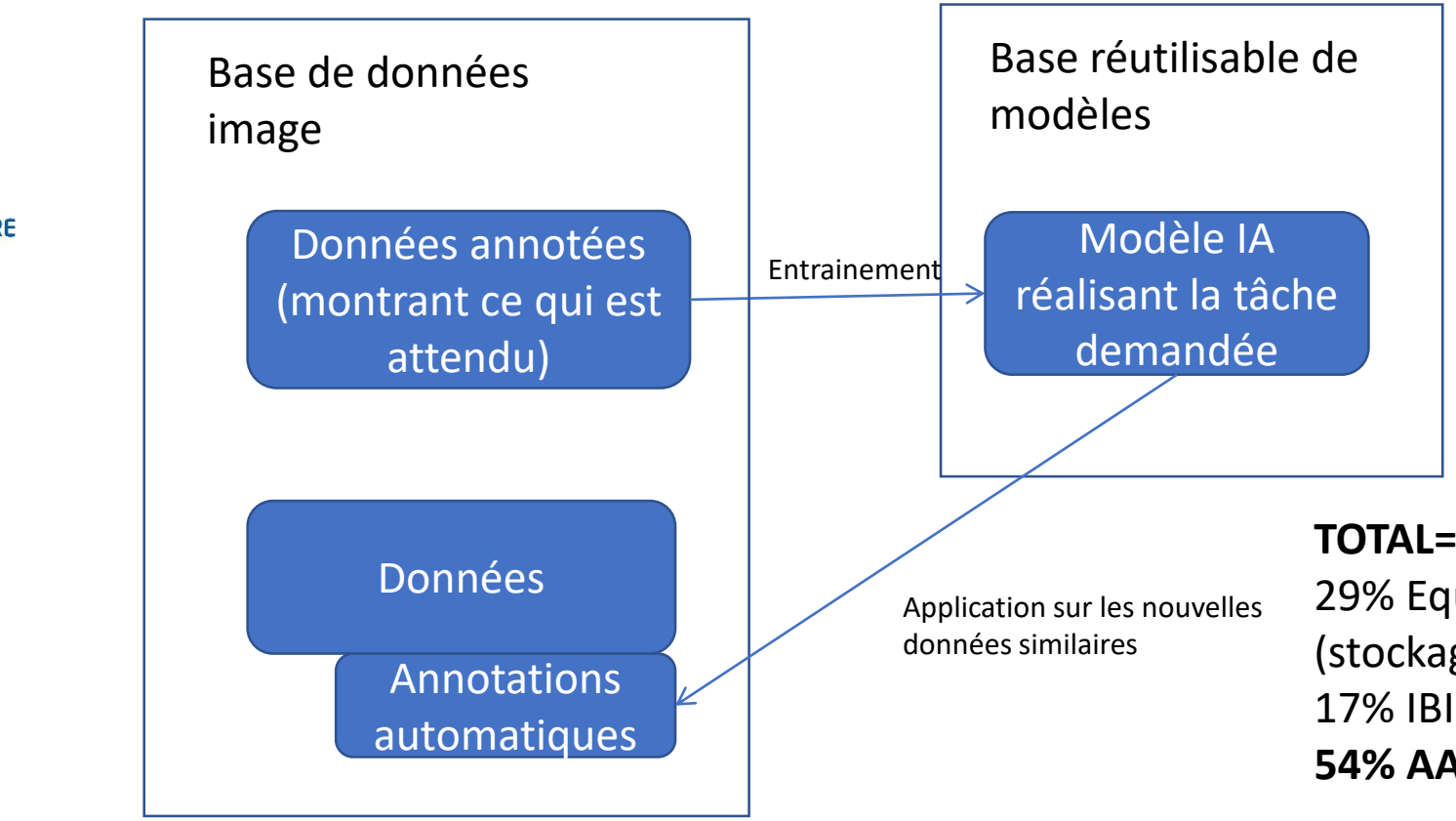



MicroPICell

BioCore

BioCore

# Un service unique de « big data » en histologie, allant de la production jusqu'à l'analyse




Exemple de tâches réalisables:
classification de tissus ou d'image, identification et comptage de différents types cellulaires, …

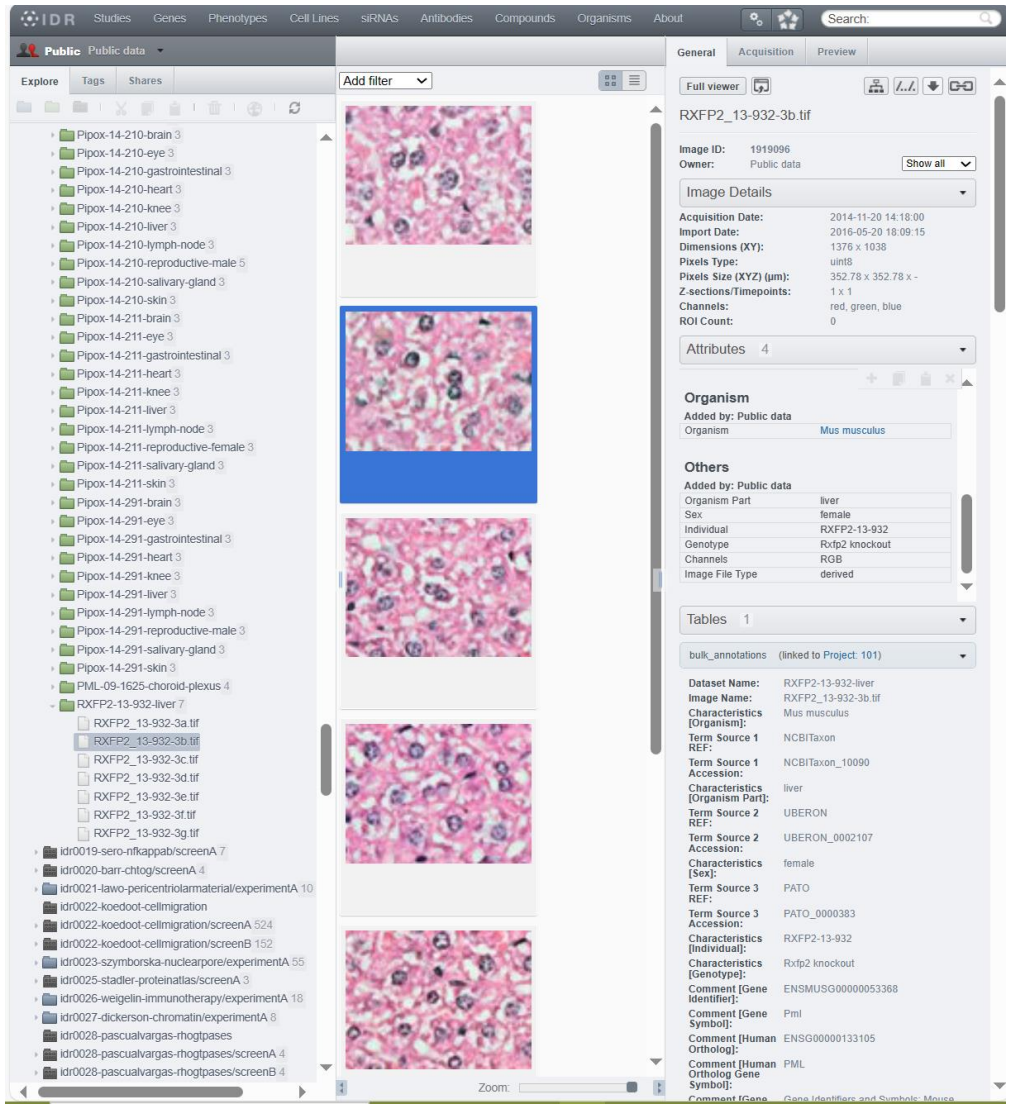
**TOTAL= 408keuros**
29% Equipex Mudi4LS (stockage et calcul Bird)
17% IBISA
**54% AAP Région**

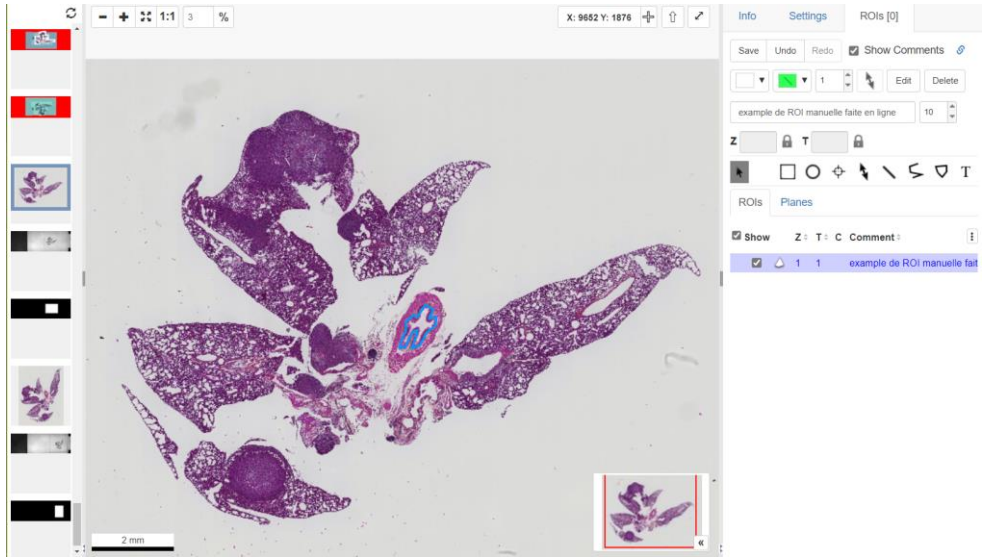
Préparation haut débit et en condition reproductible

Numérisation haut débit

Gestion et traitement des données, accès aux ressources de calcul au plus près des données

- Mise en place des outils Deep Learning: annotation et gestion image



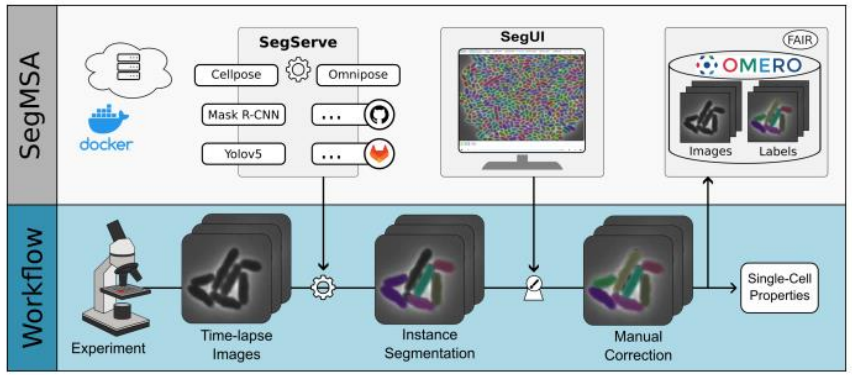




**Figure 1.** *ObiWan-Microbi* platform overview. Fully integrated end-to-end workflow for semi-automated image annotation and segmentation of time-lapse microscopy in the browser. *ObiWan-Microbi* consists of *SegServe*, *SegUI*, and OMERO jointly distributed in the microservice architecture *SegMSA*.

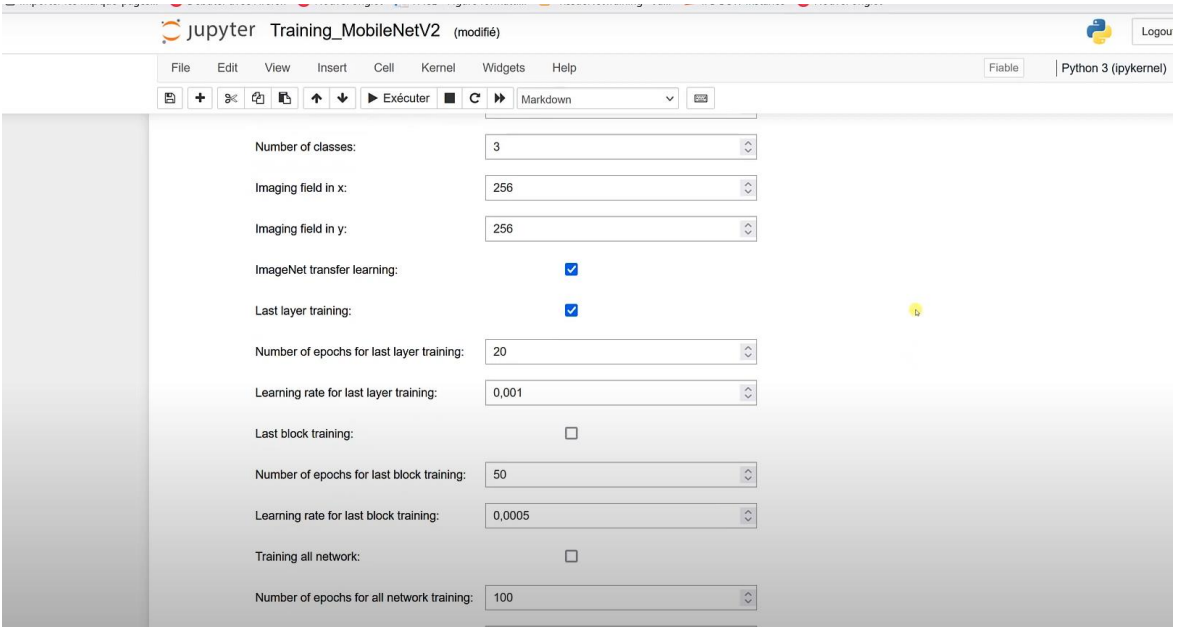
https://doi.org/10.1101/2022.08.01.502297

BioCore

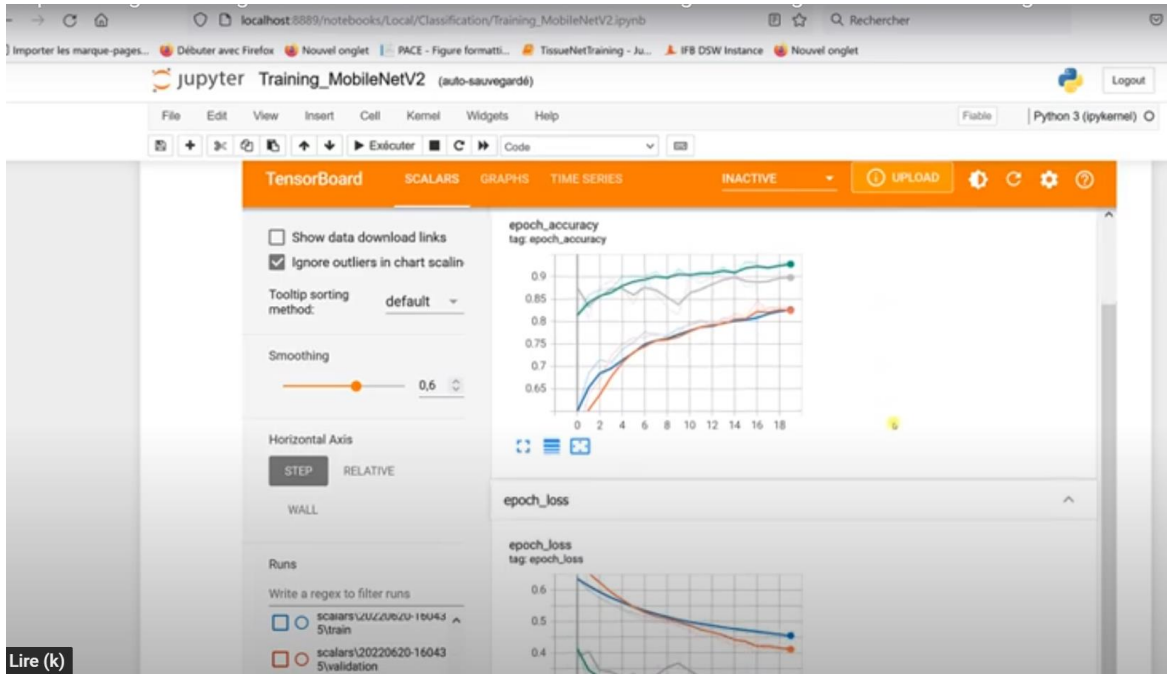
- Mise en place des outils Deep Learning: annotations et gestion image

Entrainer un modèle

Les data sont hébergées sur Glicid, au plus proches des ressources de calcul.
Elles sont annotées comme pour un dépôt dans une archive publique (BioImage Archive)




Exemple de NB T. Pécot Rennes




L'entrainement est accessible par accès en ligne de commande ou via des interfaces Jupyter Notebook (mise en place par M. Mongy FBI MicroPICell)

BioCore

- Installation des modèles sous omero par dépôt sous git

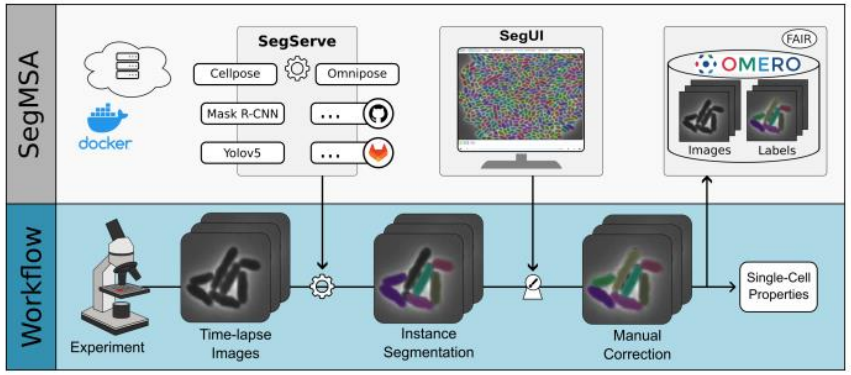
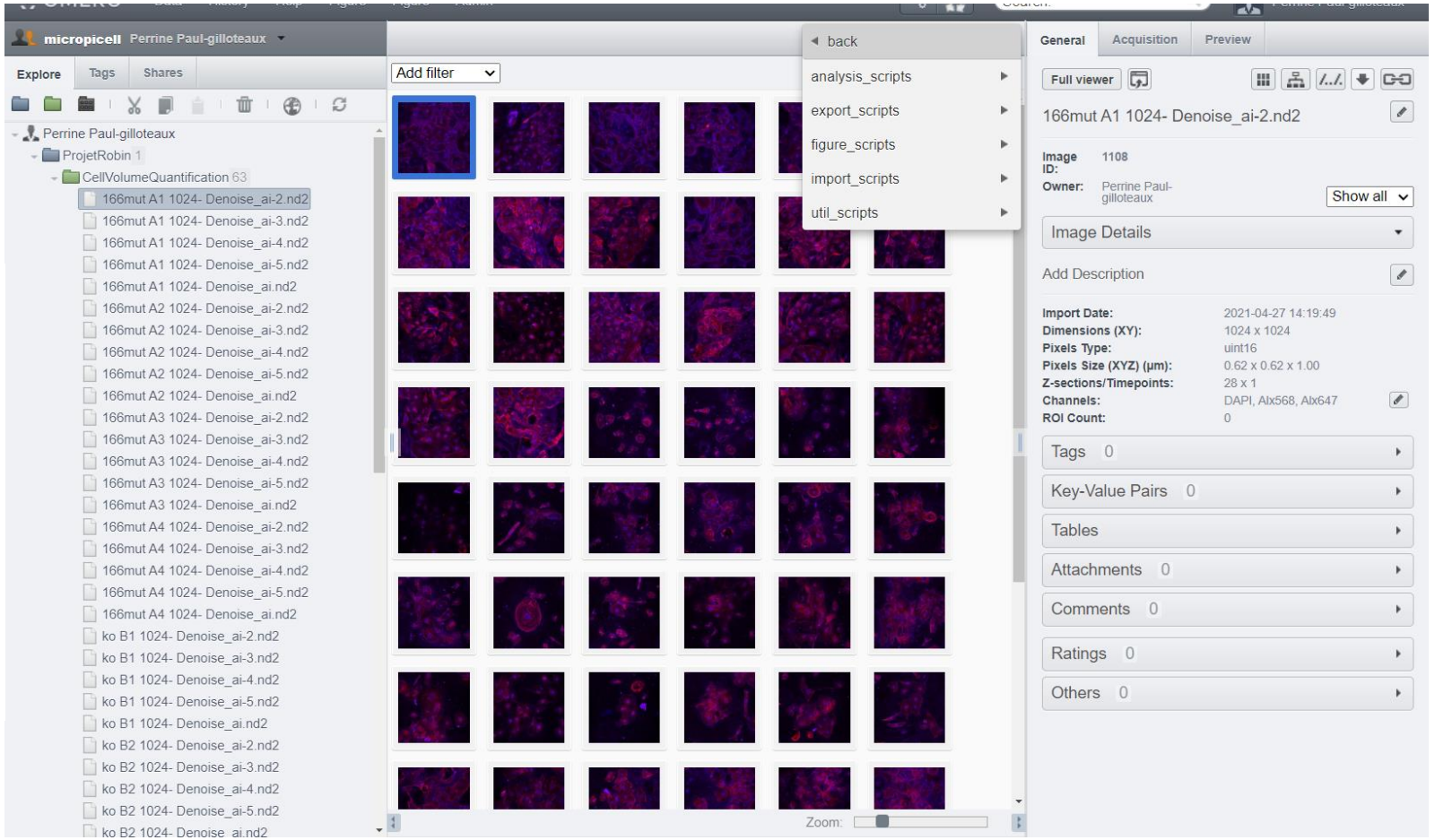
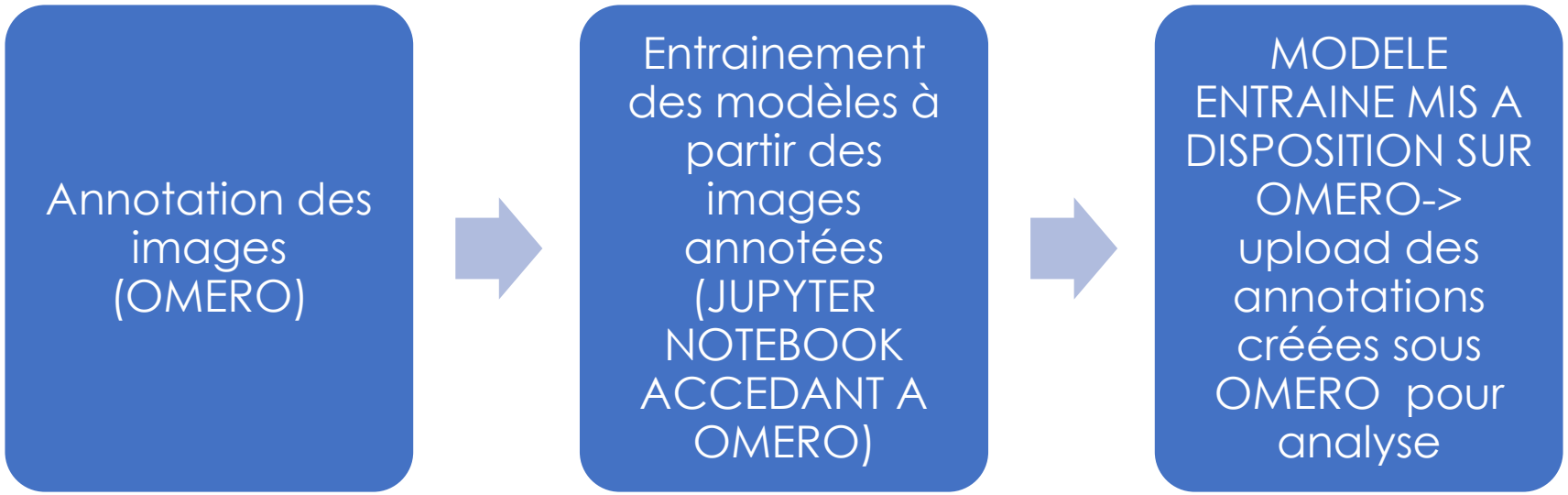



**Figure 1.** *ObiWan-Microbi* platform overview. Fully integrated end-to-end workflow for semi-automated image annotation and segmentation of time-lapse microscopy in the browser. *ObiWan-Microbi* consists of *SegServe*, *SegUI*, and OMERO jointly distributed in the microservice architecture *SegMSA*.

https://doi.org/10.1101/2022.08.01.502297

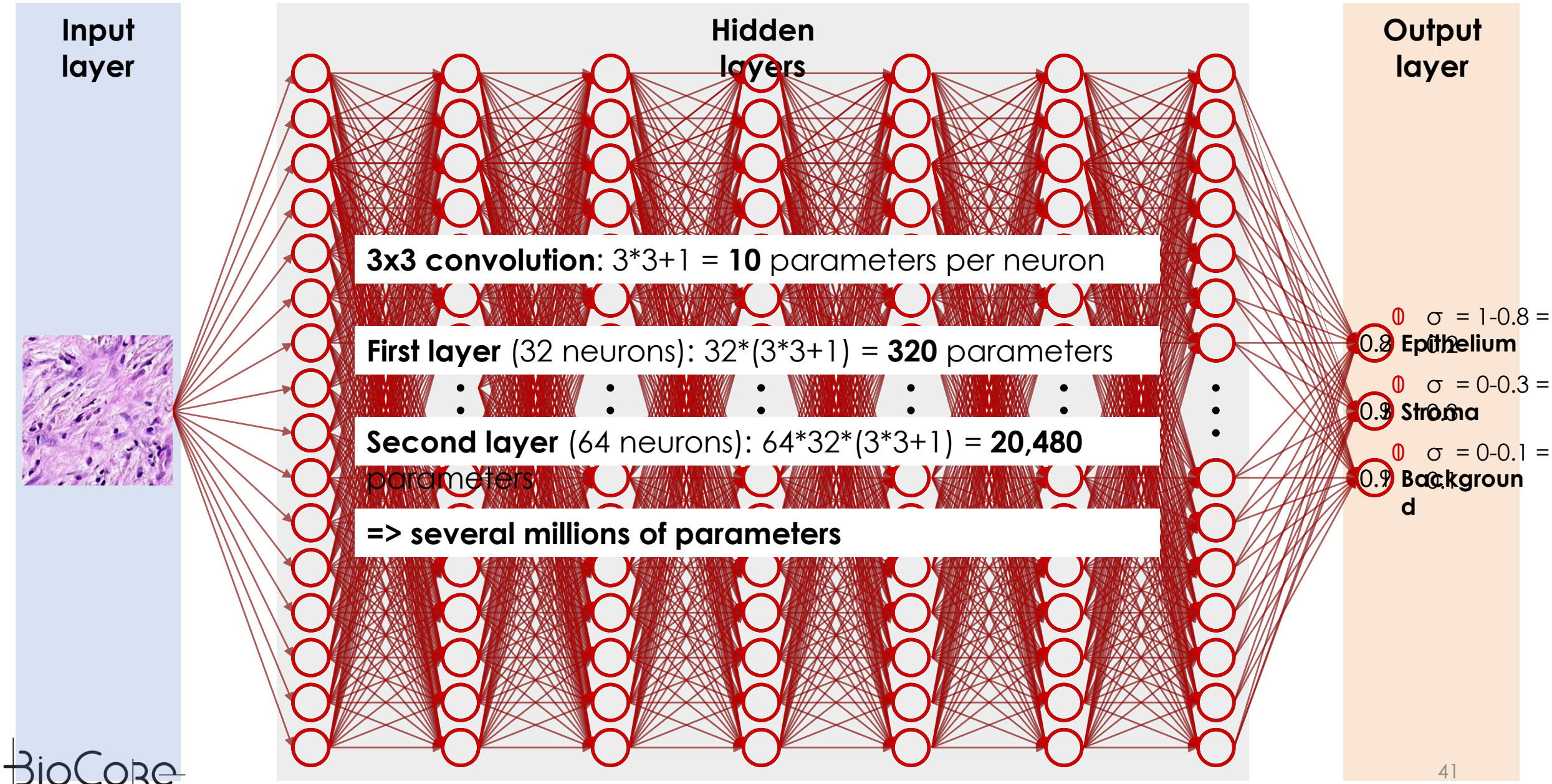


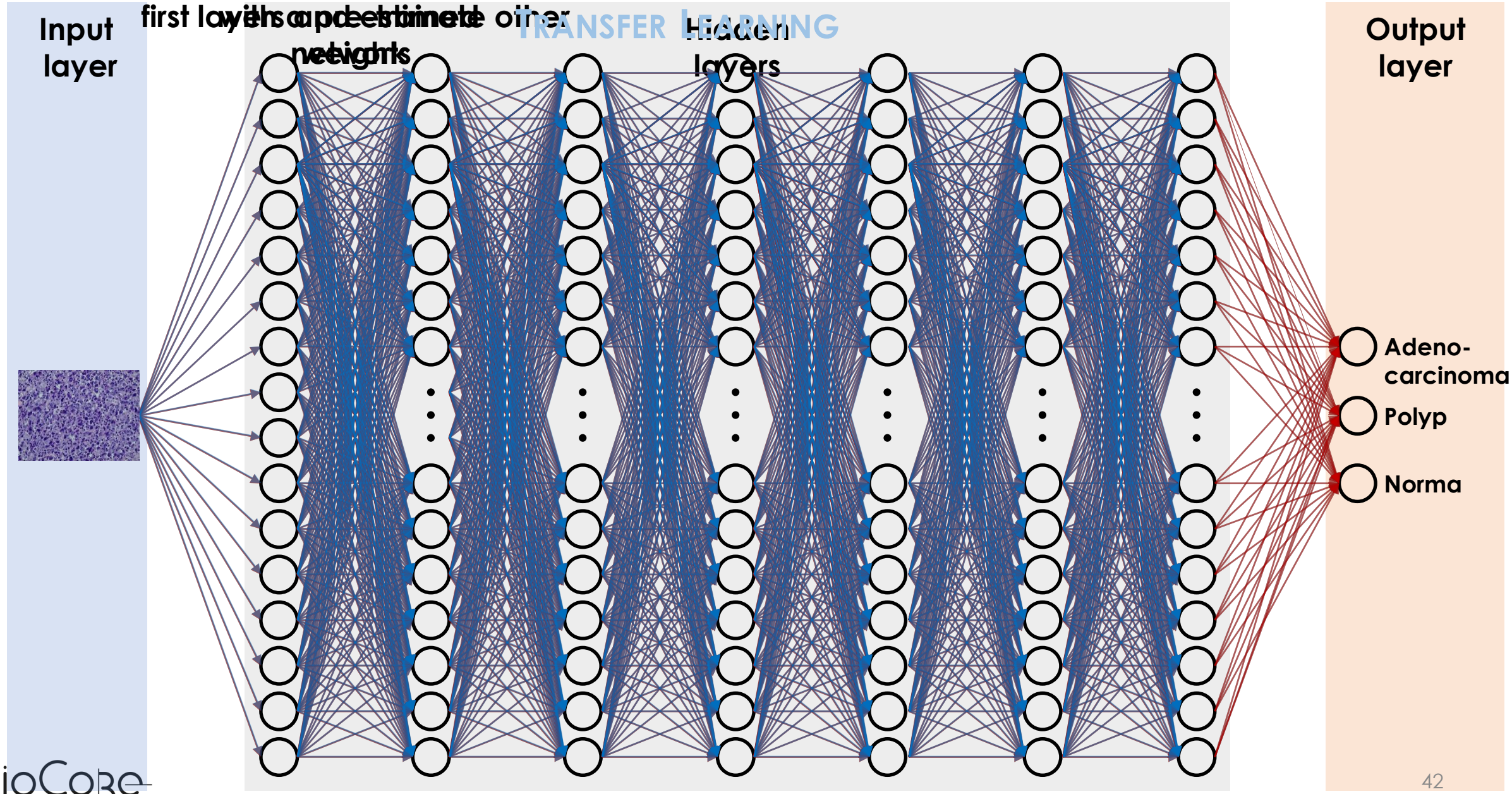

BioCore


Annotation des images (OMERO)
Entrainement des modèles à partir des images annotées (JUPYTER NOTEBOOK ACCEDANT A OMERO)
MODELE ENTRAINE MIS A DISPOSITION SUR OMERO-> upload des annotations créées sous OMERO pour analyse

BioCore

# Deep Convolutional Neural Networks




Input layer
Hidden layers
Output layer
Adeno-carcinoma
Polyp
Norma

# ACHAT D'EQUIPEMENT

High-throughput model

## NanoZoomer S360

S360

**Automated scanning up to 360 slides and high throughput of 82 slides/h***

- Scanning speed 20× mode (15 mm × 15 mm) : Approx. 30 s
  40× mode (15 mm × 15 mm) : Approx. 30 s
- Max. 360 slides

*In the case of 5 focus points

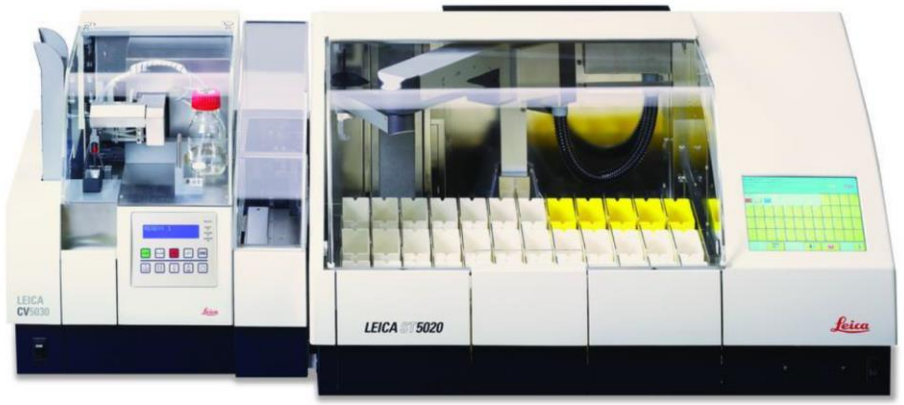
Janvier 2023

**Specifications**

| | | |
|---|---|---|
| Product name | | NEW NanoZoomer S360 |
| Part number | | C13220-01 |
| Scanning speed | 20× mode (15 mm ×15 mm) | Approx. 30 s |
| | 40× mode (15 mm ×15 mm) | Approx. 30 s |
| Objective lens | | |
| Compatible glass slide | | 26 mm × 76 mm<br>Thickness 0.9 mm to 1.2 mm |
| Slide loader | Standard size slide | 360 slides (30 slides × 12 cassettes) |
| | Double size slide | - |
| Scanning resolution | 20× mode | |
| | 40× mode | |
| Focusing method | | |
| Z-stack feature | | |
| Fluorescence imaging module | | No |
| Barcode reader | | |
| Slide format | | |
| Power supply | | |
| Power consumption (Scanner only) | | Approx. 200 VA |

# ACHAT D'EQUIPEMENT

AUTOMATE COLORATION et COLLEUSE:
Acquisition d'une station complète de coloration de lames histologiques, comprenant un colorateur et une colleuse de lamelles sur lames, avec transfert automatique des lames du premier équipement vers le deuxième.



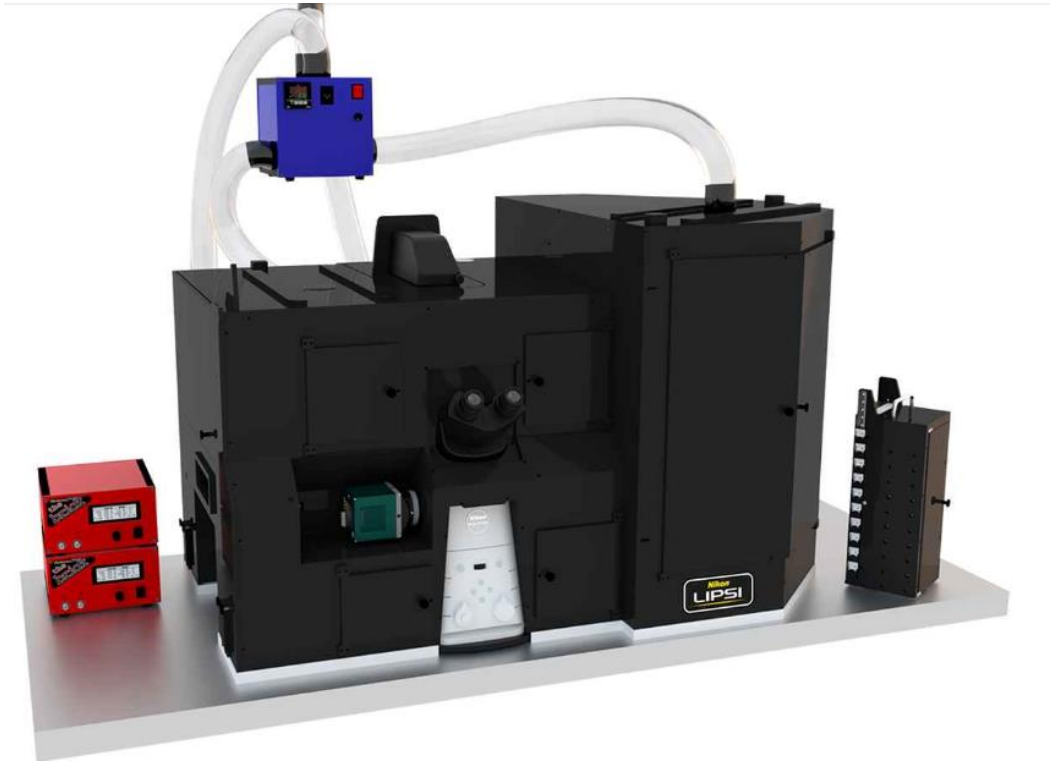

Avril 2023

MicroTome x2 + 1 Cryostat en cours de renouvellement.

# A NOTER POUR LES ANNOTATIONS

Possibilité d'utiliser des multimarquages en fluo, uniquement pour l'apprentissage (Fluo plus caméra couleur)

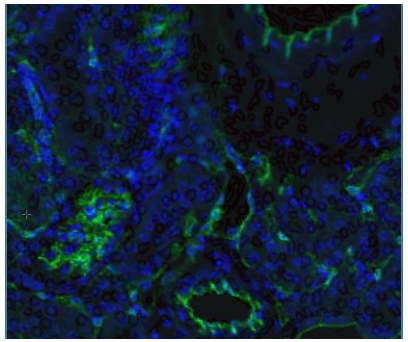
| | Main sya |
|---|---|
| Microscope | Motorized inverted microscope ECLIPSE Ti2 |
| Compatible Well Plate Types | 6-, 12-, 24-, 48-, 96-well plates, glass bottom plates & plastic bottom plates, slides and petri dishes |
| Multiple FOV Experiments | Center, Covering, Random, Random +Center and Regular pattern in each well |
| Illumination Methods | Brightfield, Phase contrast, Fluorescence |
| Image Acquisition | • Standard acquisition (Multichannel, Time lapse, multiple-positions, Z-stacks, large images, Multidimensional imaging)<br>• Intelligent acquisition |
| Imaging modalities | • Widefield (Up to two cameras simultaneously), Fast resonant confocal Nikon A1R HD, Yokogawa spinning-disk W1, Crest X-Light V2 spinning-disk<br>• TIRF & STORM (as stand-alone system while not using the robotic)<br>• Photoactivation and FRAP capabilities |
| Objectives | Up to 6 |
| Acquisition Speed | 96 well plate <1min30s (4 colors, one point per well, 18.8mm field of view camera) |
| Data Visualization | Plate view, Sample labeling, Heat map, Graphing |
| Field of View | up to 25mm |
| Capacity | Up to 20 well plates |

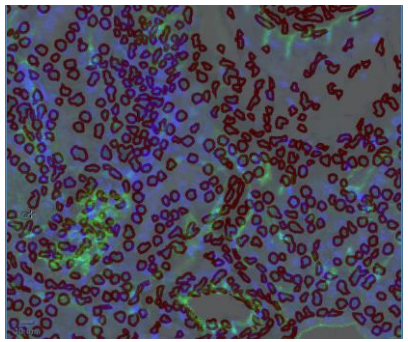


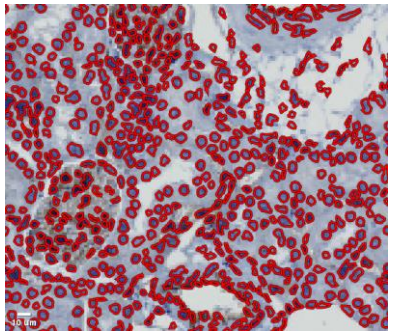

Juillet 2023

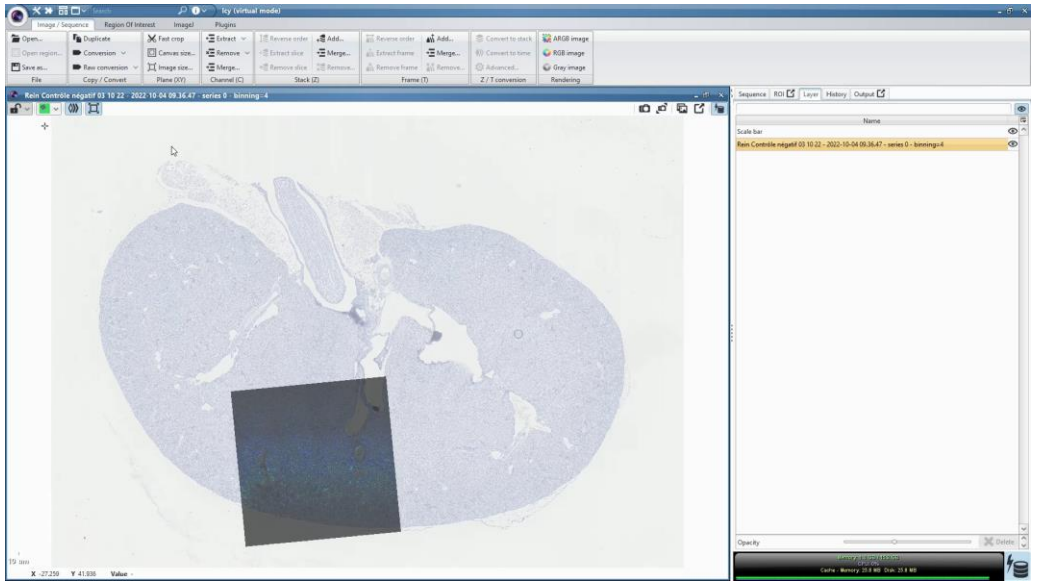
# Vers des annotations d'entrée « automatiques »

-> diminuer l'entrainement en utilisant l'imagerie multimodale (obtenir des annotations avec d'autres marqueurs plus compliqués à utiliser en routine par exemple plutôt que par annotation manuelle)



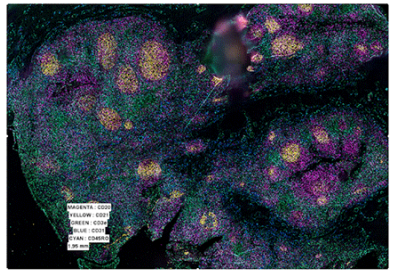


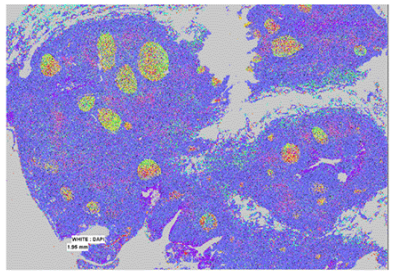


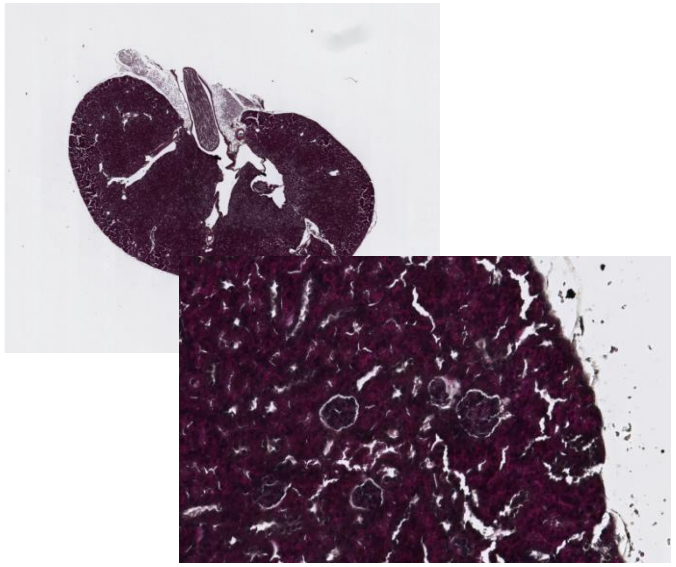


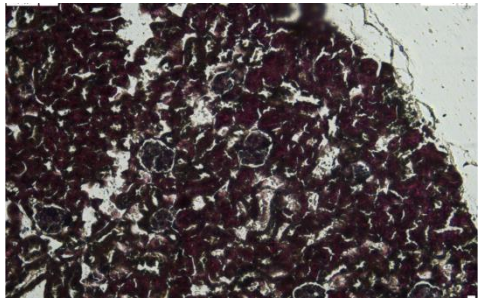
BioCore

# Vers des annotations d'entrée « automatiques »





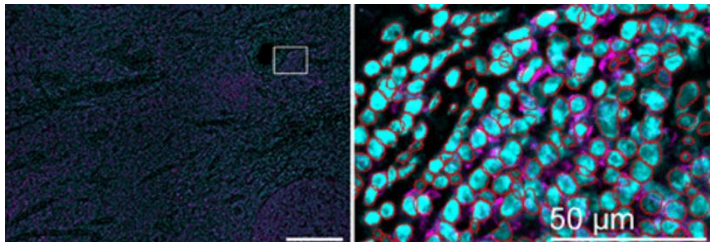




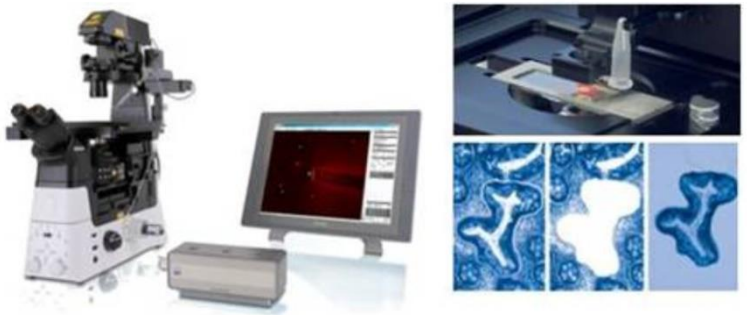
**PCR**

**Sequencing etc…**

10x, 20x or 40x (1 micrometer of cutting accuracy at 20x)







Phenocycler/Codex for fluorescence hyperplexing

Slide Scanner for fast full FOV

Microdissector (MMI cellcut)

# CALENDRIER PRÉVISIONNEL DU PROJET

| | 2021 | 2022 | | | | 2023 | | | | 2024 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | T4 | T1 | T2 | T3 | T4 | T1 | T2 | T3 | T4 | T1 | T2 | T3 | T4 |
| **Projet scientifique:** | | | | | | | | | | | | | |
| Choix et mise au point des logiciels open source de deep learning sur serveur non dédié (accès temporaire) | | | | | | | | | | | | | |
| Mise au point de l'organisation de la base de données et des protocoles d'annotations | | | | | | | | | | | | | |
| Rédaction et tests des protocoles standardisés sur automates et scanners | | | | | | | | | | | | | |
| Rapport intermédiaire | | | | | | | | | | | | | |
| Rédaction des protocoles d'utilisations de modèles entrainés et des protocoles d'entrainement des IA disponibles | | | | | | | | | | | | | |
| Mise à jour du Plan de Gestion des données pour les utilisateurs | | | | | | | | | | | | | |
| Passage sur l'architecture mesocentre définitive du projet | | | | | | | | | | | | | |
| Ouverture de l'infrastructure aux utilisateurs et journée Kick-off | | | | | | | | | | | | | |
| Utilisation de l'infrastructure pour premiers retours | | | | | | | | | | | | | |
| Rapport final du projet | | | | | | | | | | | | | |
| **Financement:** | | | | | | | | | | | | | |
| Scanner haut débit | | Rédaction CST | Appel d'offre | Choix et rapport de choix | Achat | | | | | | | | |
| Automates de coloration et colleuse (IBISA demande 2021) | Rédaction CST | Appel d'offre | Choix et rapport de choix | Achat et livraison | | | | | | | | | |
| Microtome 1 | | | | Rédaction CST | Appel d'offre | Choix et rapport de choix | Achat | | | | | | |
| Microtome 2 | | | | | Rédaction CST | Appel d'offre | Choix et rapport de choix | Achat | | | | | |
| Serveur tampon local (IBISA demande 2021) | | Rédaction CST | Appel d'offre | Choix et rapport de choix | Achat et livraison | | | | | | | | |
| Calcul GPU/CPU (EQUIPEX MUDI4LS) | | | | Rédaction CST | Appel d'offre | Choix et rapport de choix | Achat et livraison | | | | | | |
| Stockage CEPH ((EQUIPEX MUDI4LS) | | Rédaction CST | Appel d'offre | Choix et rapport de choix | Achat et livraison | | | | | | | | |

Ouverture Utilisateurs Pilotes
01 Février 2024

**Notre public cible:** les équipes de recherches fondamentales (pas de stockage d'information clinique) et les équipes de développement de nouvelles méthodes en IA

**Nos objectifs:**
-> Faciliter l'accès aux équipes de developpement de données annotées

-> Faciliter le réentrainement par les biologistes/chercheurs non informaticiens en diminuant le nombre de données à annoter grâce au transfer learning

-> Faciliter l'application d'algorithme et la gestion des données jusqu'au dépôt sur archives pérennes publiques en suivant les recommandations nationales et internationales.

BioCore

**MicroPICell Team**
Aurélie Doméné
Anne Gaignerie
Stéphanie Blandin
Steven Nedellec
Philippe Labrot
Philippe Hulin
Yevgeniya Simon
Marc Mongy (FBI.data)
*Annabelle Justin*
*Magalie Feyeux*

INCA Plan CANCER

LABEX IGO

AAP Région Pays de la Loire

BioCore

Perrine.paul-gilloteaux@univ-nantes.fr

UMS BioCore Inserm US16 / CNRS UAR 3556 / Nantes Université / CHU de Nantes
IRS UN - 8 Quai Moncousu – BP 70721
44007 Nantes cedex 1 - France