上課注意事項:

1. 教室內禁止攜帶食物飲料入內,僅允許“白開水”,請大家將食物飲料放置於教室外的桌上。
2. 請實體與線上學員掃描以下QR code進行線上簽到。
3. 請線上學員於課程開始前關閉自己的麥克風。
4. 線上學員若有問題,請先按下“舉手”,或於聊天室寫下問題,將於課程結束後在場地時間允許下,安排QA時間。
5. 現場學員發問時請使用麥克風才可進行收音。

線上簽到

課程材料與資訊連結

台大醫學院
National Taiwan University
College of Medicine

顯微影像核心 Imaging Core
分子影像重點技術平台
NTU Consortium of Molecular imaging Key Technology

AI 機器學習與深度學習介紹

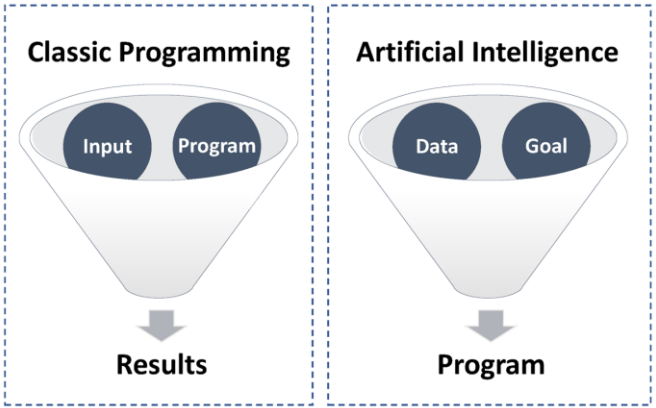
2025/4/11

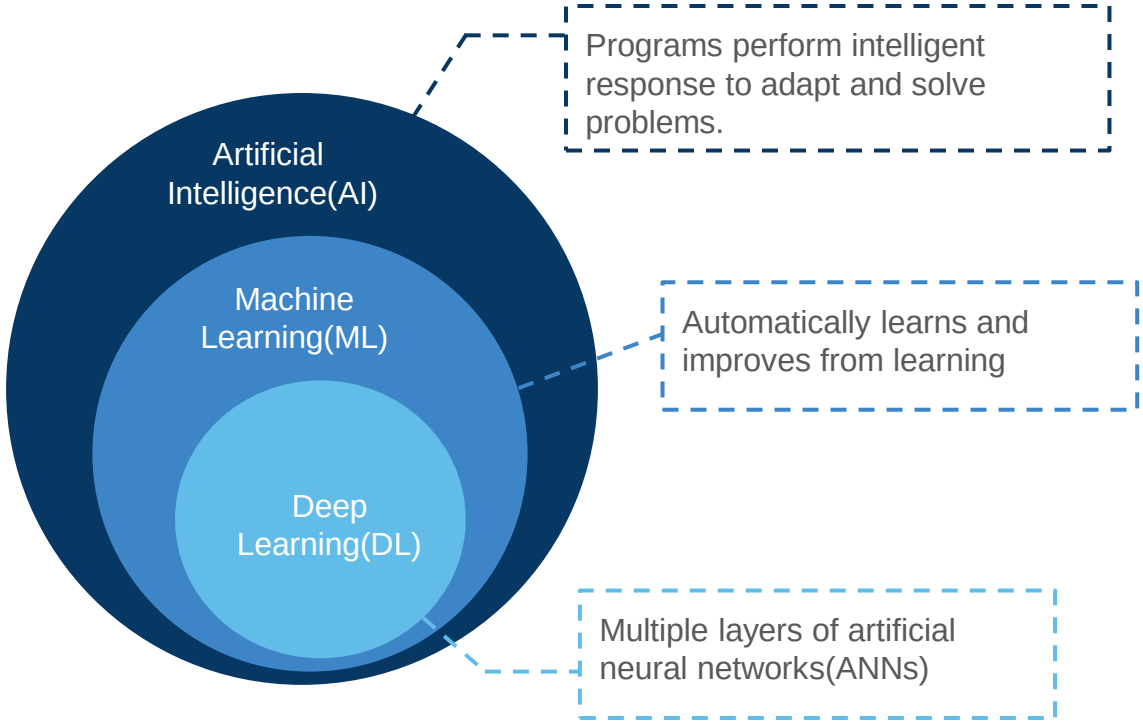
An-Chi, Luo
羅安琦

Outline

- About the Machine Learning (ML) & Deep Learning (DL)
  - Definitions and key differences
  - How models learn
- Relevance to Bio-Image Analysis (BIA)
  - Overview of commonly used tools
  - Platforms of ML and DL tools
- Three useful segmentation tools
  - Overview of Labkit, StarDist, and Cellpose

What is Artificial Intelligence(AI) ?

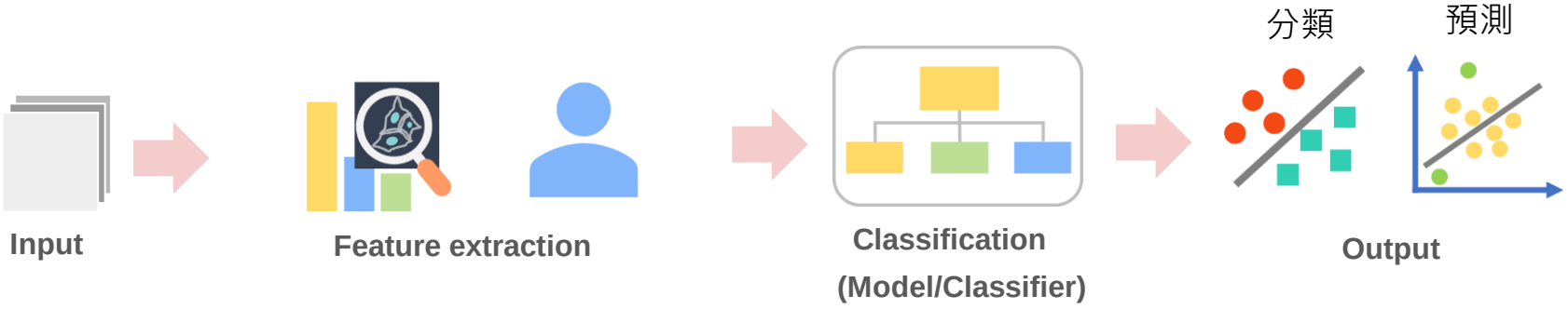


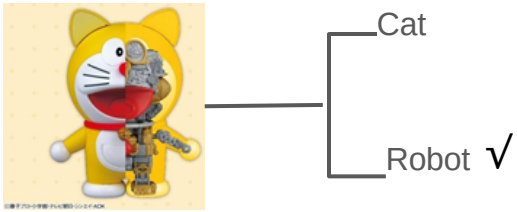


Definition of ML-Machine Learning

What is the Machine Learning…

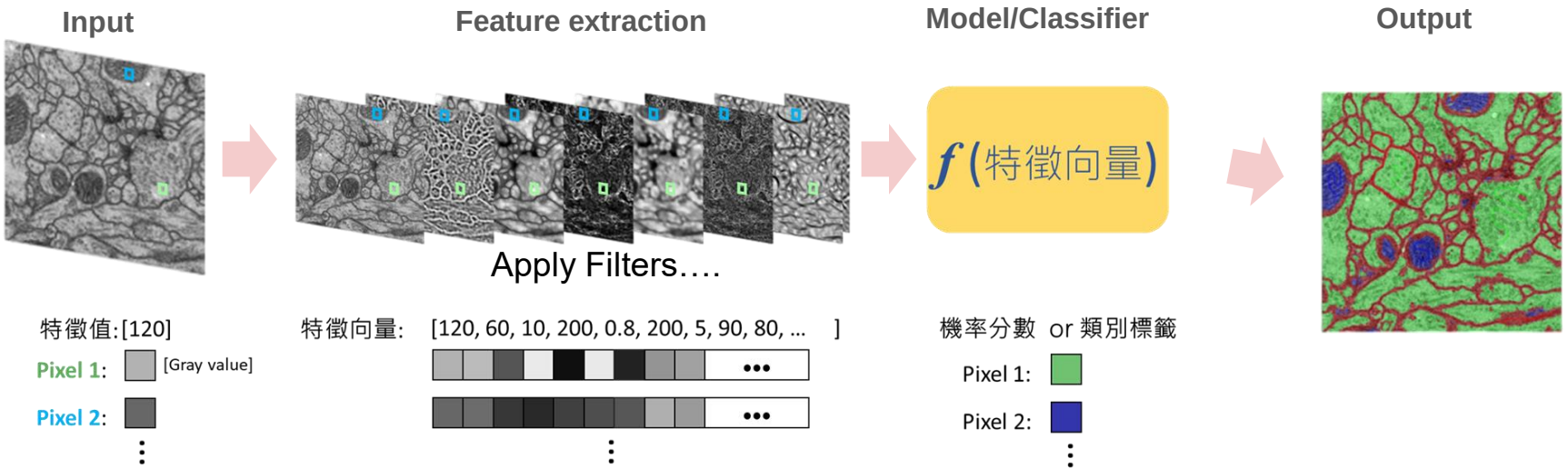
- ML uses algorithms to learn patterns from data for making predictions or decisions.
- ML allows systems to improve performance from experience without explicit programming.
- Feature extraction and classification are two independent steps.
- Feature extraction requires human participation.





Definition of ML-Machine Learning

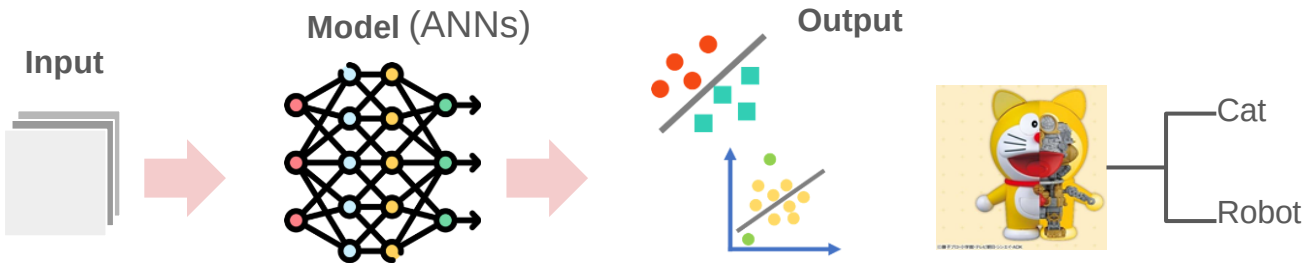
How Machine Learning actually works…

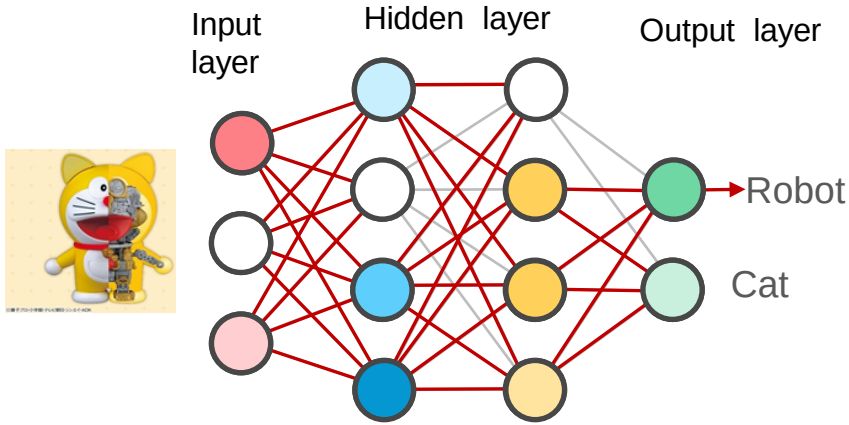


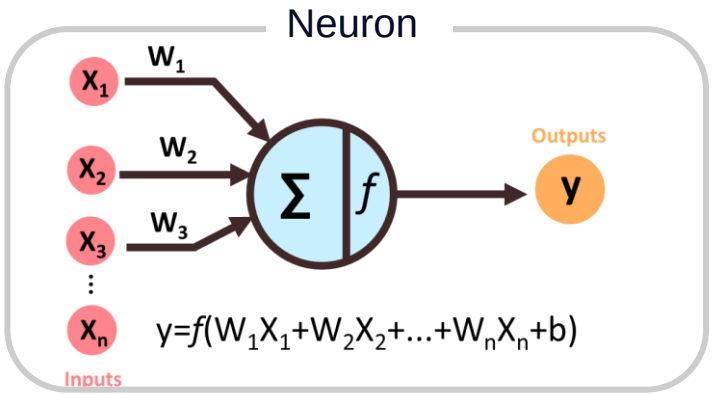
Machine learning is a great tool, but feature extraction heavily relies on domain expertise, making it less flexible for complex or large-scale data...

Definition of DL-Deep Learning

- DL is a subset of ML employing multiple layers of artificial neural networks (ANNs) to learn complex patterns from data and make predictions.
- Feature extraction is automatic and **simultaneous** with **classification.**

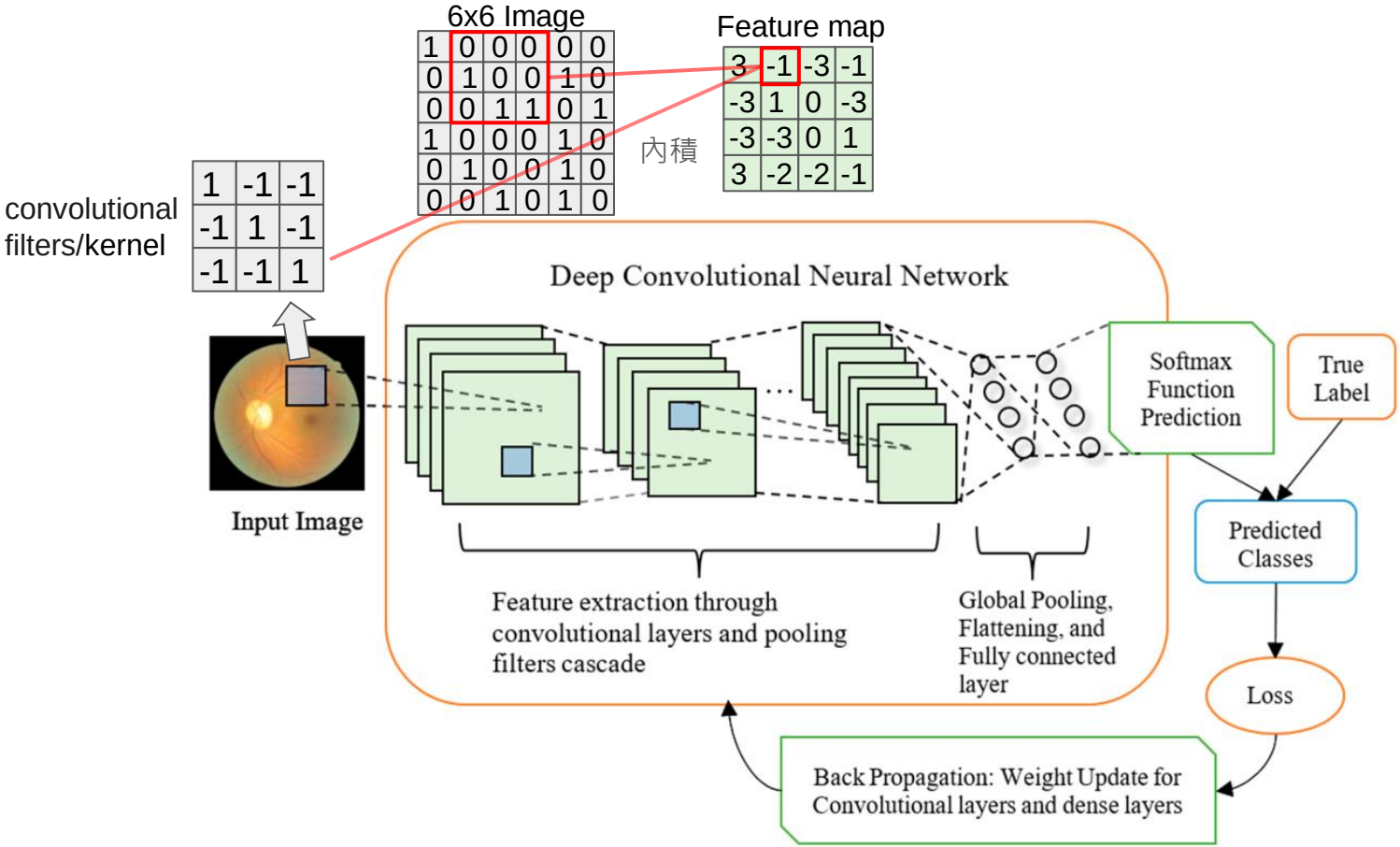






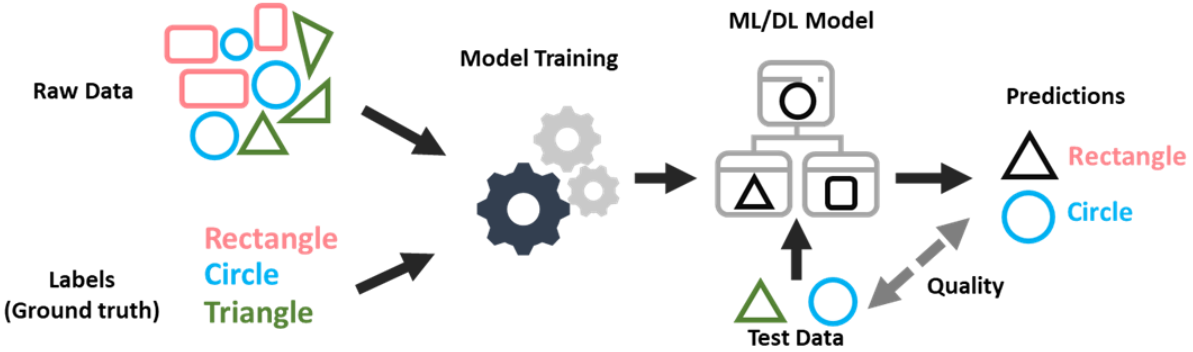
F: active function 啟動函數

How Convolutional Neural Networks Work(CNN) ?



Definition of AI, ML, and DL-Deep Learning

Supervised Learning



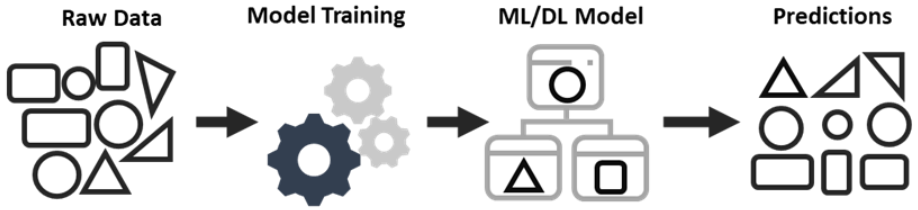
Classification

- SVM
- k-NN
- Random Forest
- CNN

Regression

- Liner
- Polynomial
- SVR
- CNN

Unsupervised Learning



Dimensionality Reduction

- PCA/SVD
- t-SNE
- UMAP
- Autoencoders

Clustering

- K-means
- Hierarchical
- DBSCAN
- OPTICS

Anomaly Detection

- One-class SVM
- GMM
- Isolation Forest
- Autoencoders

Relevance to BIA

From the Perspective of Bioimage Analysis workflow

Raw Image

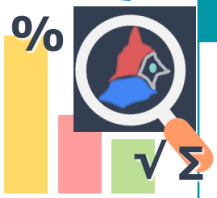
Image Preprocessing

Purpose	Tool	Type / Platforms	Images
Denoising	Content-Aware Image Restoration (CARE)	DL / Fiji & ZeroCost.	Any
	Noise2Void	DL / Fiji & ZeroCost.	Any
	DenoiSeg	DL/ Fiji	Any
Remove out-of-focus images	Microscope Image Focus Quality Classifier	DL/ Fiji & CellProfier	FL
Super-resolution	DeepSTORM	DL/ ZeroCost.	FL

Segmentation

Purpose	Tool	Type / Platforms	Images
Pixel level classification	Trainable WEKA segmentation	ML / Fiji	BF, FL, TEM, LS
	Labkit	ML / Fiji	Any
Instance segmentation	StarDist	DL/ Fiji & CellProfier & ZeroCost. & CellProfier & QuPath	BF, FL
	CellPose	DL/ Fiji & CellProfier & ZeroCost. & CellProfier & QuPath & GUI for Python	BF, FL
	CDeep3M	DL/ ZeroCost. & Web GUI	ET, FL, SEM

Relevance to BIA



Feature Extraction & Data Mining

Purpose	Tool	Type / Platforms	Images
Turn objects into numerical data for clustering and classification to capture explicit and latent features.	Trainable WEKA segmentation	ML / Fiji	BF, FL, TEM, LS
	Ilastik	ML / Ilastik & Fiji	BF, DIC, FL, PC, SEM, TEM
	CellProfiler Analyst	ML/ CellProfiler	Any
	Orange	ML & DL / Orange	FL
	Svetlana	DL/ Python-Napari	Any
	YAPiC	DL / Python & Fiji & Ilastik	Any
Object tracking	TrackMate + Cellpose or StarDist	DL / Fiji	Any
	DeepTrack2	DL / Python	FL, BF

Other

Purpose	Tool	Type / Platforms	Images
Image Cross-modality translation	Label-free prediction (fnet) 3D	DL / ZeroCost.	BF, EM
	CycleGAN	DL / ZeroCost	BF(H&E), Ph

ZeroCost: ZeroCostDL4Mic; BF: bright-field; FL: fluorescence; LS: light sheet; PC: phase contrast; SEM: scanning electron microscopy; TEM: transmission electron microscopy; Ph: phase contrast; DIC: differential interference contrast

Jan M, Spangaro A, Lenartowicz M, Mattiazzi Usaj M. From pixels to insights: Machine learning and deep learning for bioimage analysis. Bioessays. 2024;46(2):e2300114. doi:10.1002/bies.202300114

User-friendly Platforms for ML and DL

Local device

Ilastik

QuPath

- Built-in features(Both) for ML
- Extension(QuPath) for DL
- Bioimage Model Zoo for ML

- Plugins(Both)
- Native DL Inference Support(Fiji)
- Bioimage Model Zoo(Fiji)

CellProfier

Fiji

Cloud based

ZeroCostDL4Mic

Colab

- Adaptable Scripts
- Training with High-end GPUs
- Big Data Handling
- Model Diversity

- Pretrained DL/ML model repositories

Cloud-based training, local inference

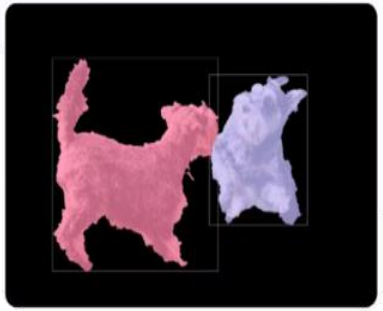
- Only supports TensorFlow SavedModel format

Powerful Segmentation Tools

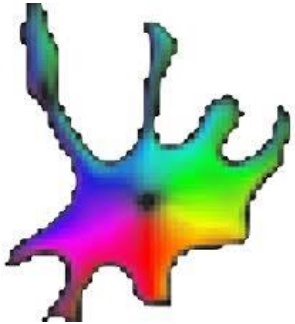
- LabKit
- StarDist
- CellPose

(a) Image

(b) Semantic Segmentation



(c) Instance Segmentation



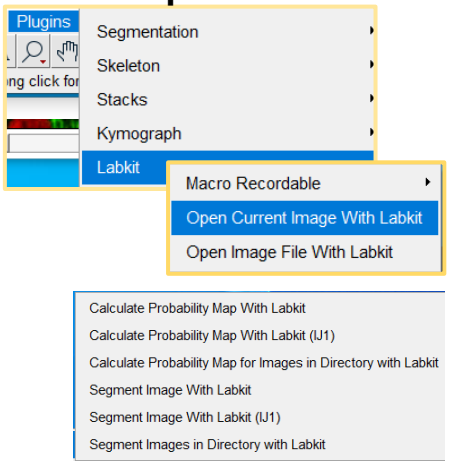
Encord. (2023, Month Day). *The ultimate guide to instance segmentation in computer vision*. Retrieved March 29, 2025, from https://encord.com/blog/instance-segmentation-guide-computer-vision/

Powerful Segmentation tool-LABKIT

A Machine learning-based segmentation tool

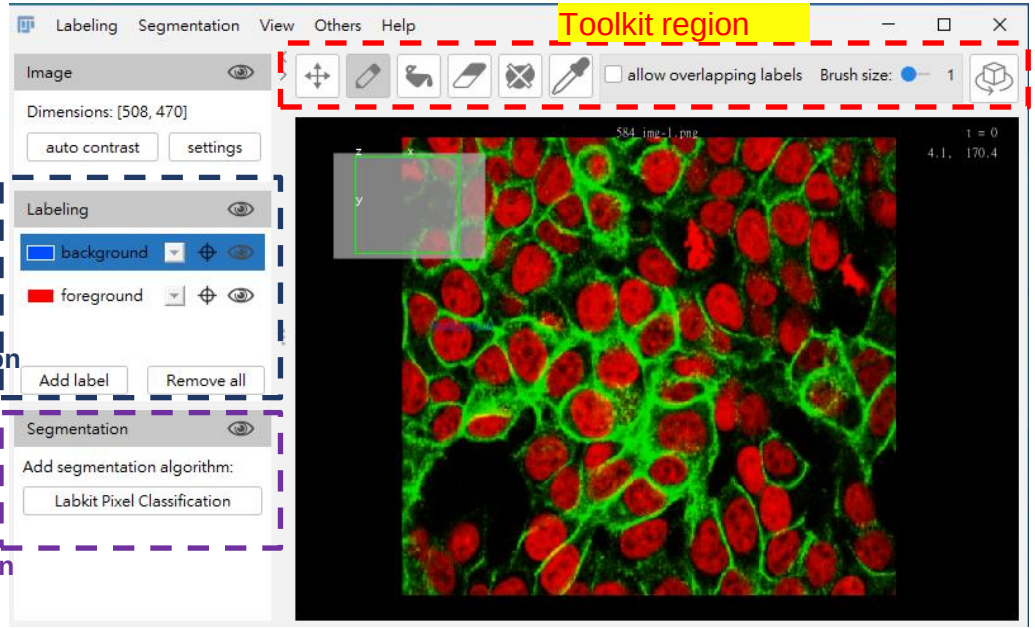
- Labkit's labeling toolkit trains a random forest classifier with minimal annotation
- Suitable for visualization and interaction with big-volume data
- Easy to manual labeling and Macro recordable

How to Open LABKIT

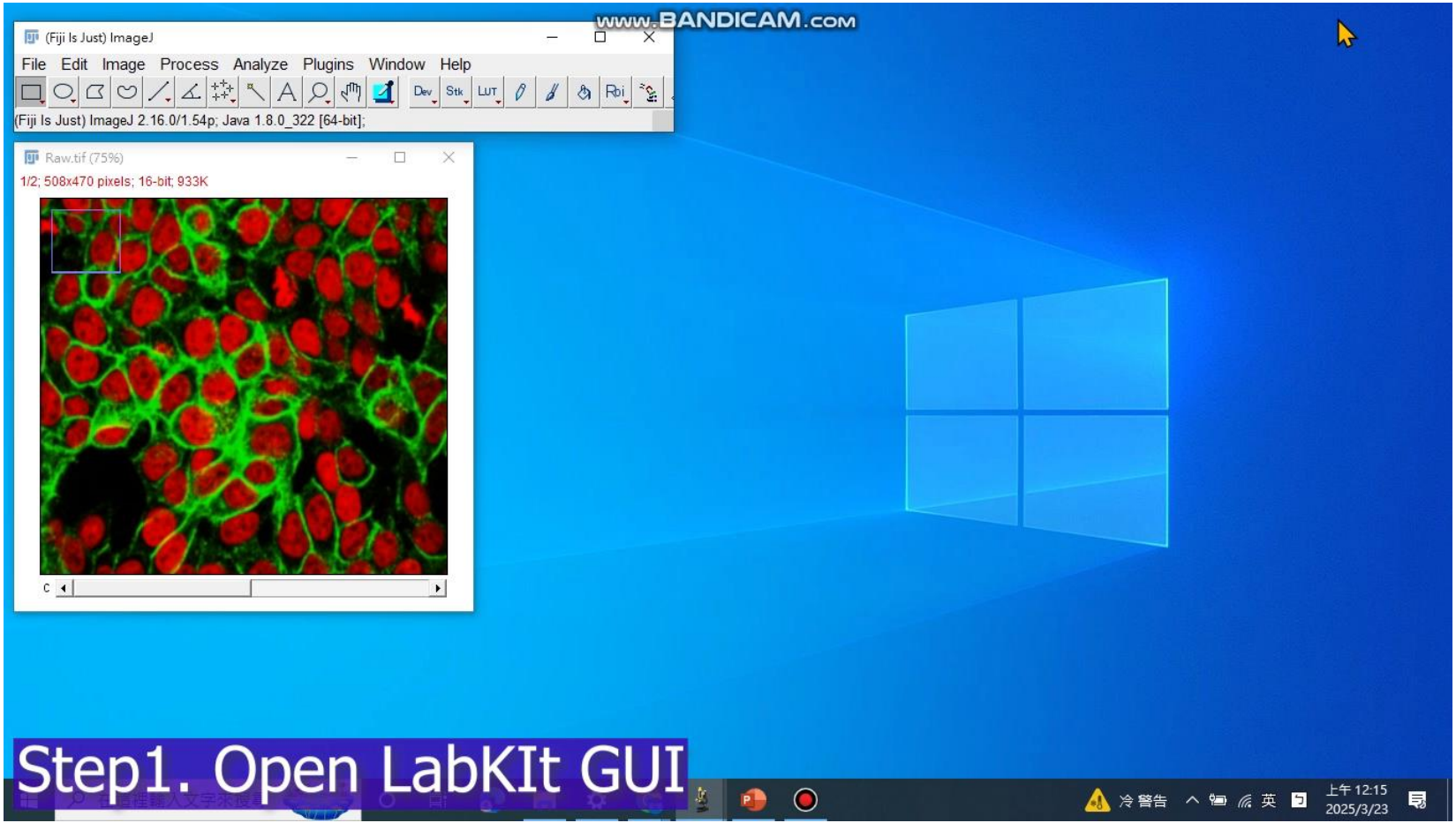


Labeling Tab
- **Manage Labels**
- **Assign new annotation**

Segmentation Tab
- **Manage classifier**
- **Execute classification**



How to run LabKit



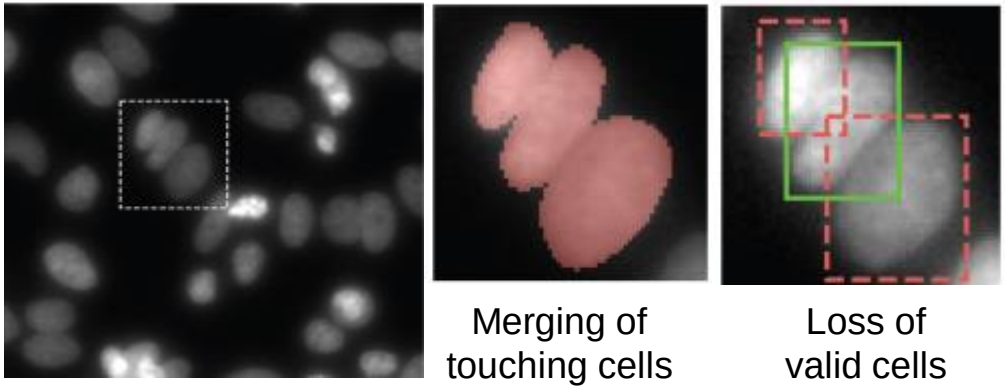
StarDist: Better Segmentation for Crowded Cells

A deep learning-based tool

- Work for fluorescence and H&E stain images
- Excel at segmenting crowded cells where traditional methods(e.g., Mask R-CNN) struggle

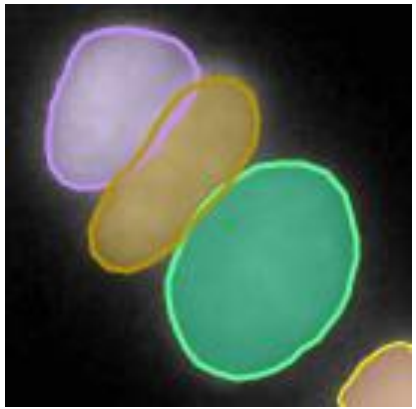
➢ **Core Technology: Detection cell with Star-convex Polygons**

Potential segmentation errors for Mask R-CNN



Merging of touching cells

Loss of valid cells

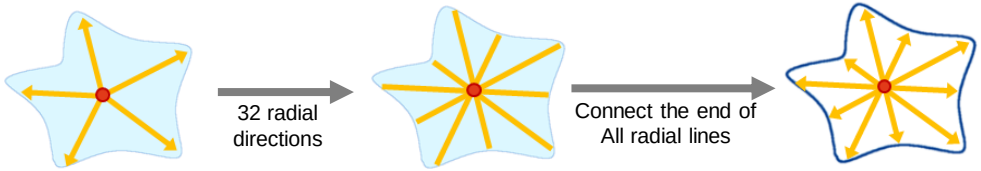


StarDist can do a better job!!

The concepts behind StarDist

➢ Core Technology: Detection object with Star-convex Polygons

Star-convex Polygons

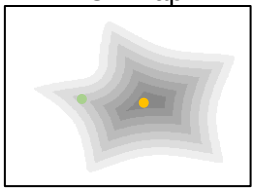


➢ How does it work?

- StarDist predicts the Object probability(OP) and Radial distance map …
- Object probability shows how likely a pixel is part of an object. (The higher values mean it closing to the center)
- Radial distance: the distance from each pixel to the edge of the target

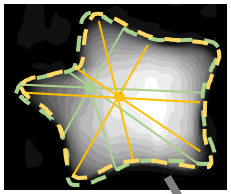
Based on the local maxima of OP maps, StarDist predicts polygons for each object

OP map

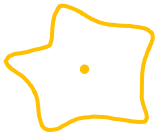


\+

Radial distance maps

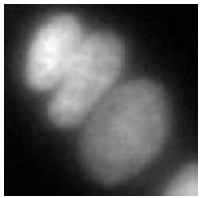


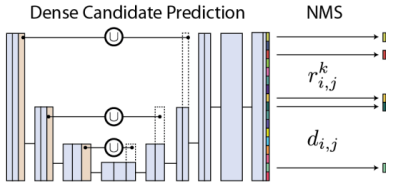
OP map is updated by comparing the mean OP of polygons.

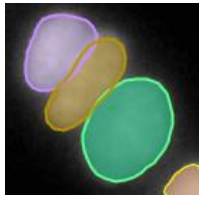


Define the real center and optimal polygons

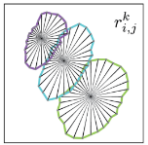
If crowded Cell



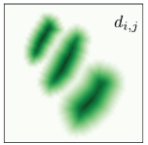




Radial distance

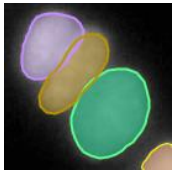


Object probability



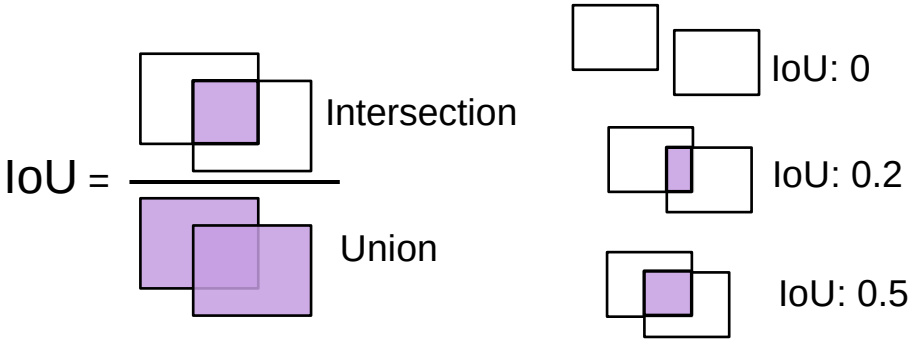
Need NMS selection…

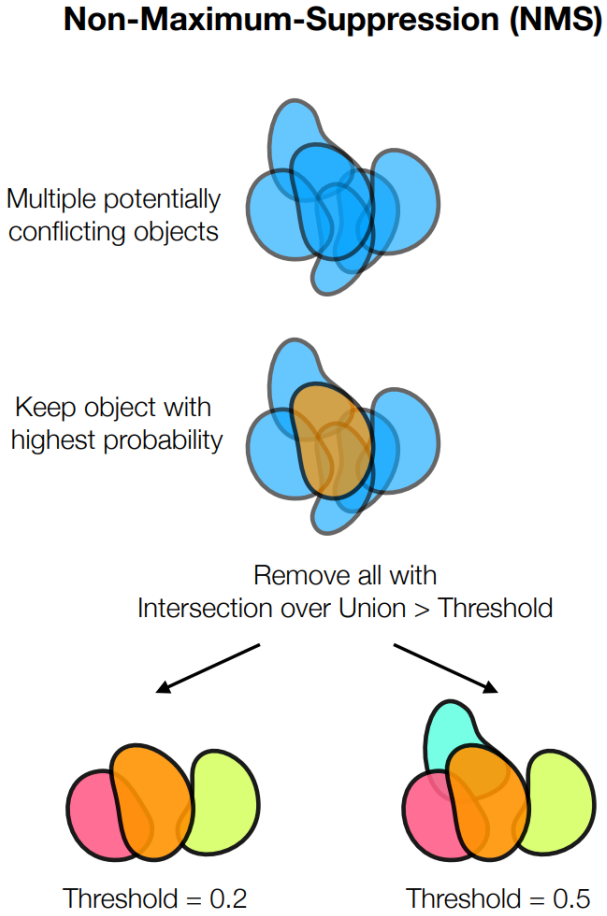
How NMS refines predictions

目標: 去除重疊物件（NMS 處理）：

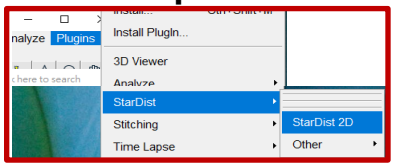
- 物件相鄰，重疊或過大可能導致模型不確定性，最終產生重疊的多邊形(overlapping polygons)。
- **非極大值抑制(NMS)** 依據 **重疊率(IoU) 的閾值** 來篩選 overlapping polygons ，保留置信度(mean object probability) 最高的 polygon，並移除其他 polygon。





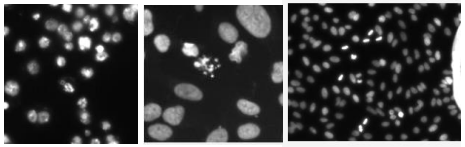
Segmentation Image with StarDist

How to Open StarDist

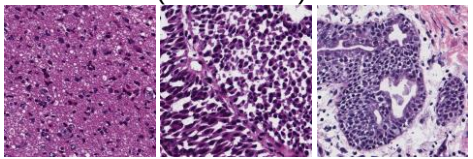


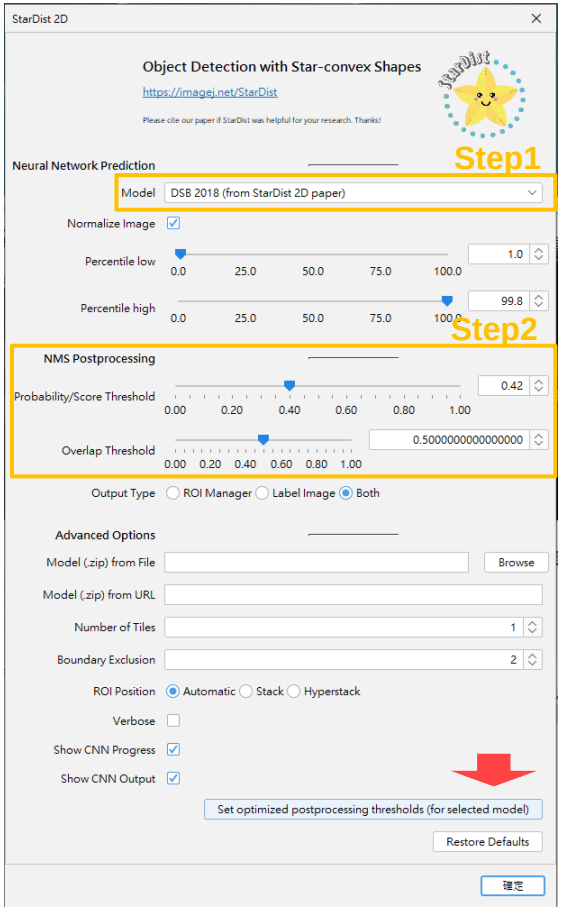
Step1. 選擇預訓練模型

- Versatile (fluorescent nuclei)
- DSB 2018 (from StarDist 2D paper)



- Versatile (H&E nuclei)

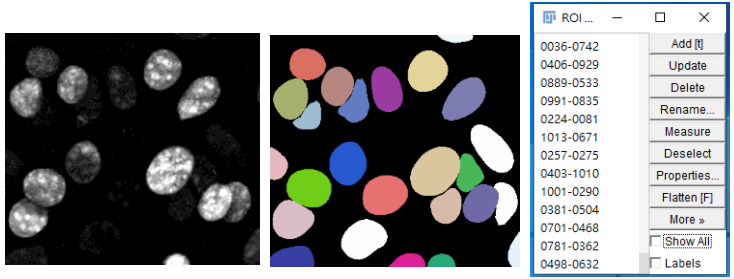




Step2. 調整NMS Postprocessing parameters

- Probability- Higher values lead to fewer segmented objects but will likely avoid false positives.
- Overlap Threshold- Higher values allow segmented objects to overlap substantially.

-Manually adjust or click the button to load the default setting…



Outputs

Things you need to remember…

1. **Probability and Overlap Threshold control number and position of objects.**

2. **The size and shape of objects cannot be controlled by probability and overlap threshold.**

3. **StarDist doesn't work out on non-star-convex objects.**

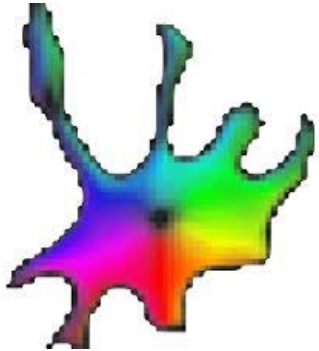
4. **Cite the paper, if you are using the plugin in your research.**

Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell Detection with Star-Convex Polygons. In *Lecture Notes in Computer Science* (pp. 265–273). Springer International Publishing. doi:10.1007/978-3-030-00934-2_30

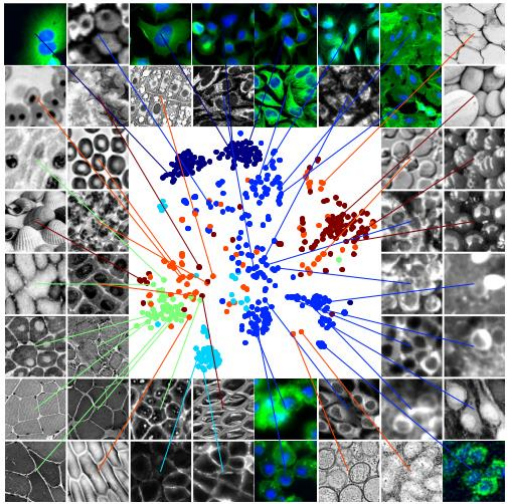
Powerful Segmentation tool-Cellpose3.0

A deep learning-based generalist segmentation tool

- Suitable for various image types (brightfield, fluorescence images, etc.)
- Full built-in models and other built-in models for different cell types
- Easy to do custom training: Fine-tune model with your own data
- Image restoration model for improving cellular segmentation

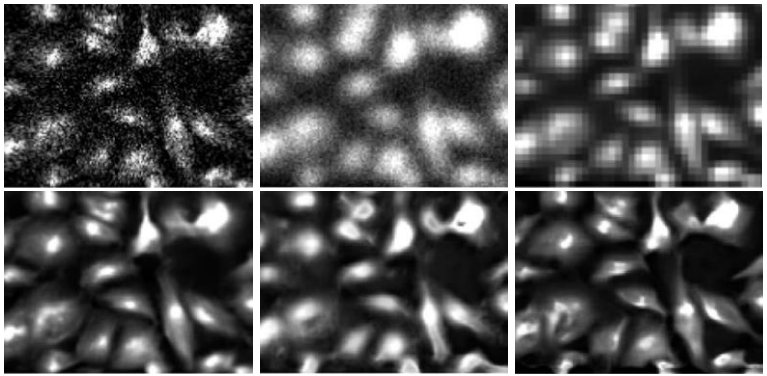


➢ **Core Technology: Flow-based segmentation**



Visualization of the diverse training dataset of Cellpose

The function of the restoration model



Denoising **Deblurred** **Upsampling**

The concepts behind CellPose

➢ Core Technology: Flow-based segmentation

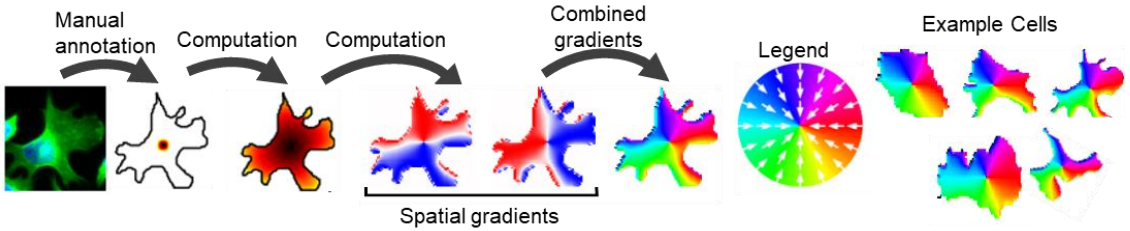
What is the Flow?

The flow is a vector that describes the direction and magnitude of each pixel's movement toward the cell center.

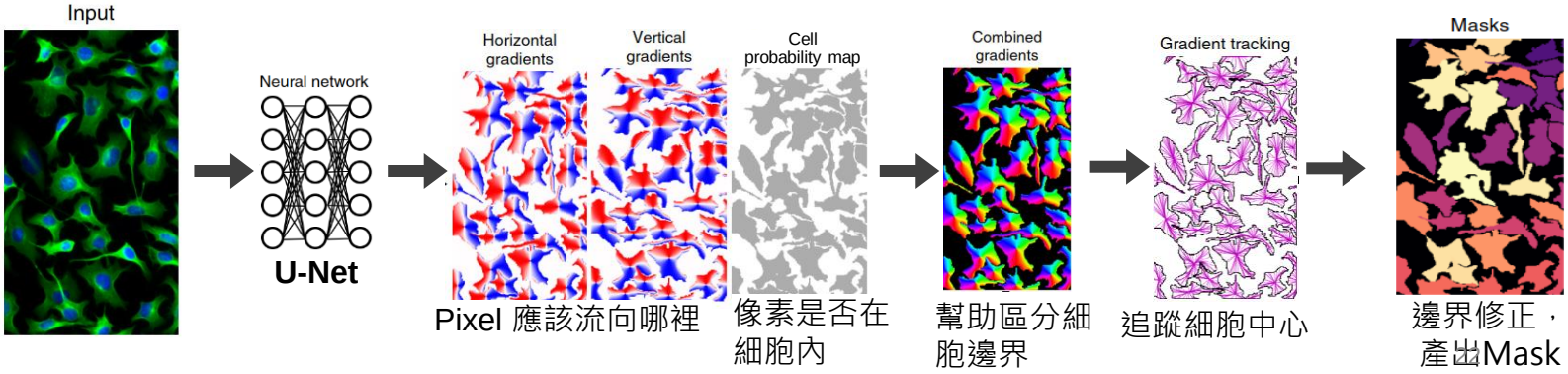
How does it work?

Cellpose predicts the cell probability map and flow fields, then traces back to the cell center to reconstruct the cell contours and generate a segmented cell mask.

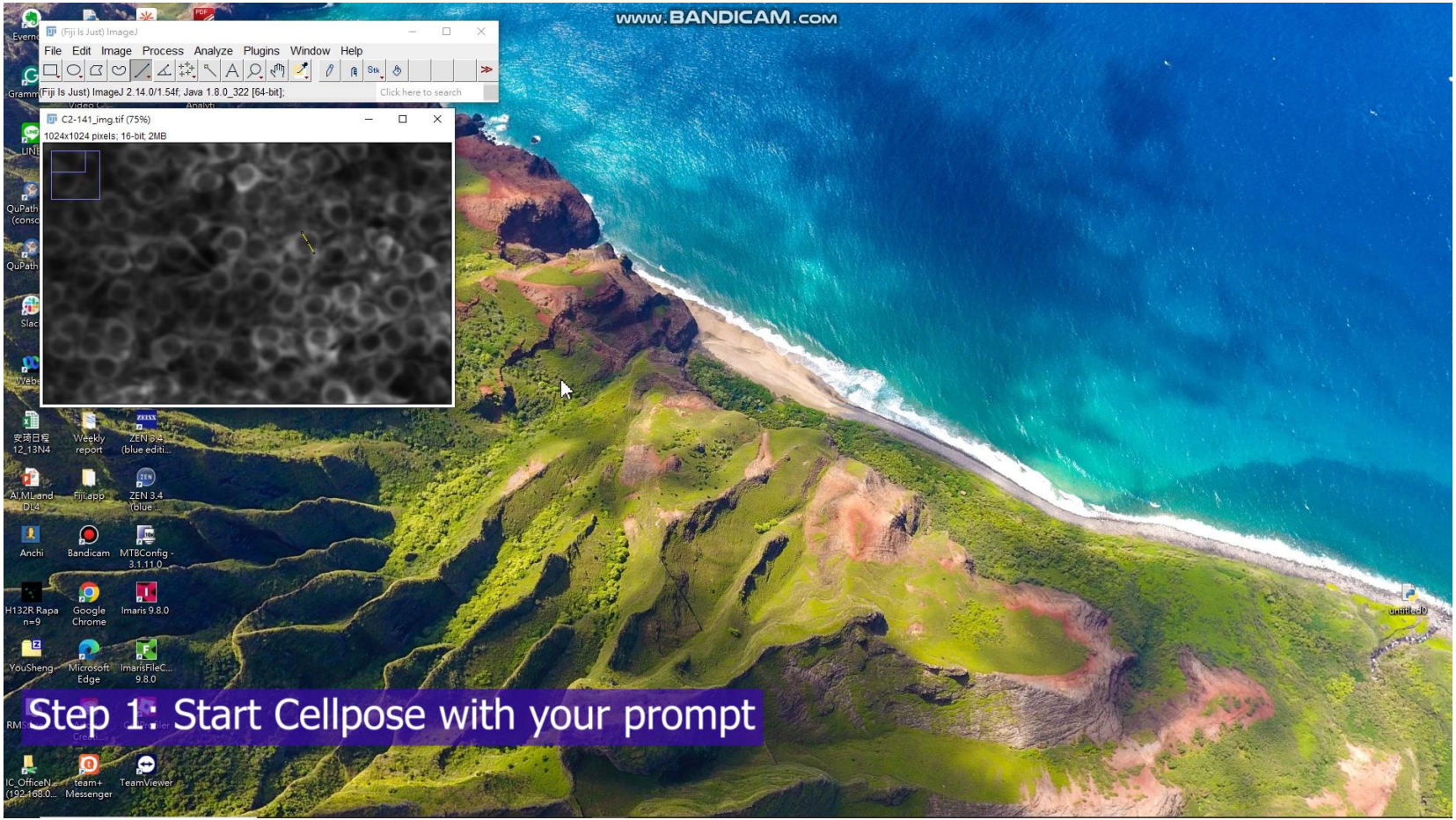
Training stage



Inference stage

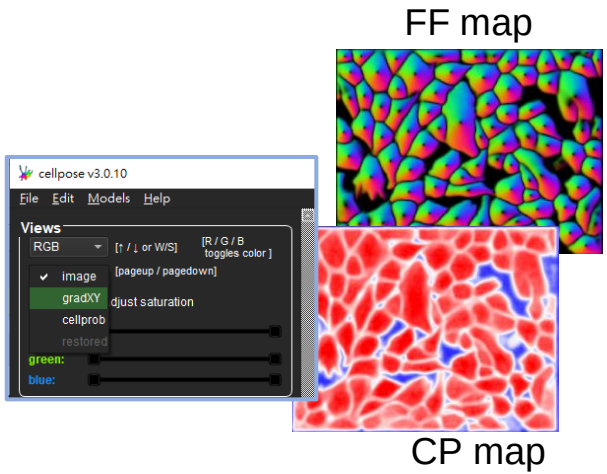


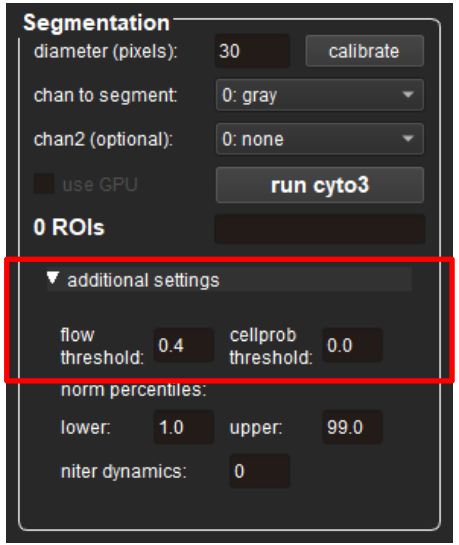
How to run CellPose



How to fine-tune the segmented Mask

- **Model Output: Cell probability(CP) & Flow Field(FF) map**
  - The model predicts Flow Field and Cell Probability.
  - The segmented Mask is generated based on these outputs.
- **Before Adjusting Parameters: Check FF& CP Maps!**
  - If cells are missing in both FF & CP maps, parameter tuning won't help!!
  - Preprocess your images or retrain the model.
- **Fine-tuning with Flow & Cellprob Threshold(FT amd CT)**
  - Adjusting FT & CT can improve segmentation.

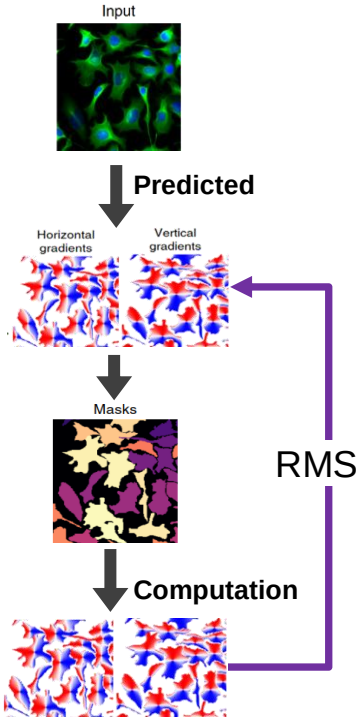


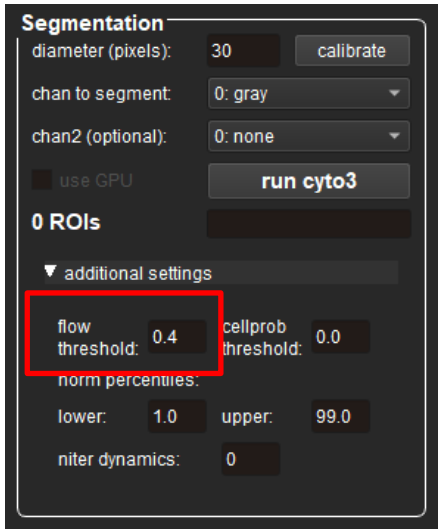


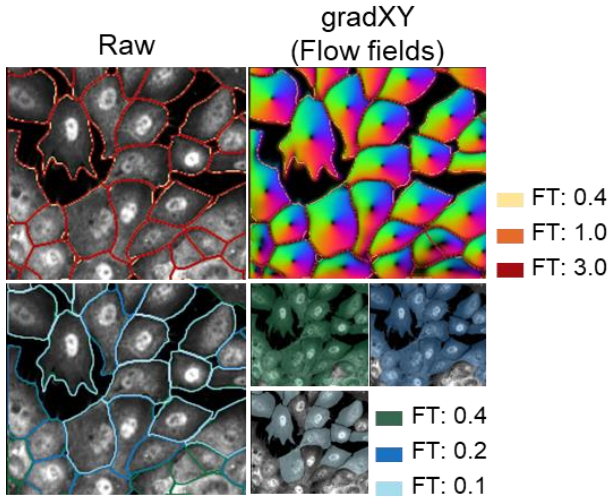
How to fine-tune the segmented Mask

What is the Flow threshold(FT)?

- A parameter can be used to filter out unreliable cell contours after segmentation.
- It evaluates the reliability of the segmentation results through Root Mean Square Error(RMSE)
- Higher RMS indicates a mismatch between the cell outline and predicted Flow Field, allowing removal with Flow_threshold(FT).







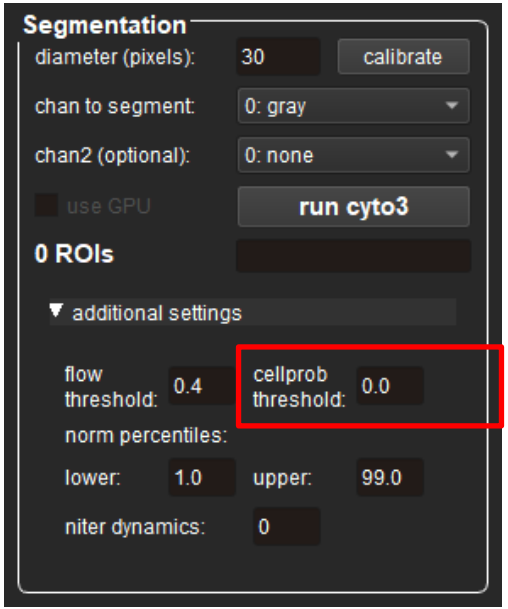
Range: 0.1to 3(Default: 0.4)

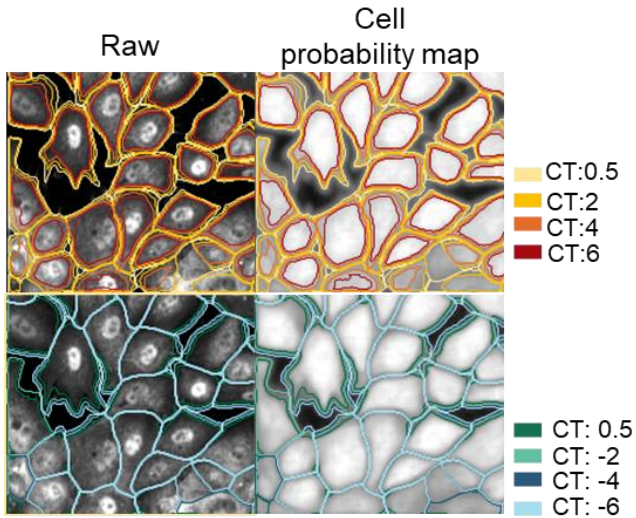
- FT 改變不太影響細胞輪廓，但是影響選到的細胞數。
- FT 越高，留下越多不穩定細胞，可能也保留部分錯誤偵測。
- FT 越低，僅保留高置信度的細胞，可能濾掉邊界模糊的真實細胞。

How to fine-tune the segmented Mask

What is the Cellprob threshold(CT)?

- Cell Probability is a value that indicates the possibility of a pixel belonging to a cell.
- Cellprob threshold is a standard. Only pixels with cell probability higher than this threshold will be considered as part of the cell.





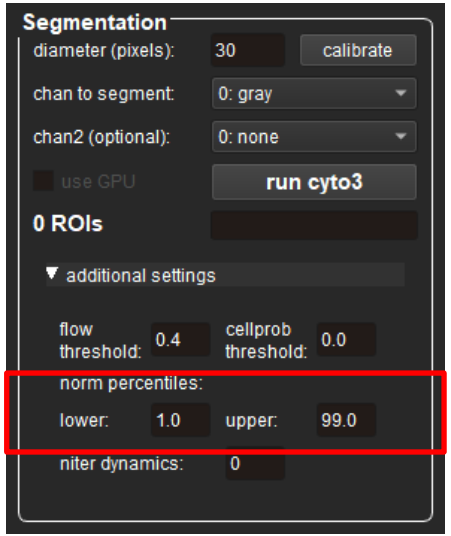
Range: -6 to 6(Default: 0)

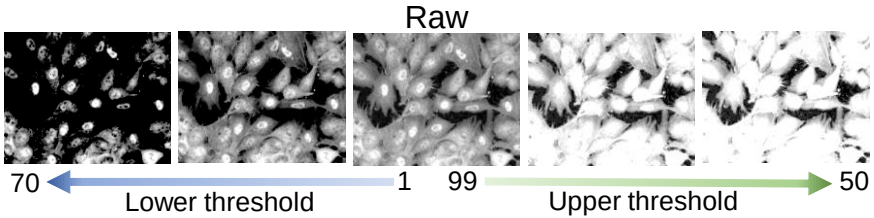
- CT 改變會影響細胞面積與數量。
- CT 越高，面積越小，可能會導致部分細胞無法被識別。
- CT 越低，面積越大，可能會選到更多背景，導致過度偵測。

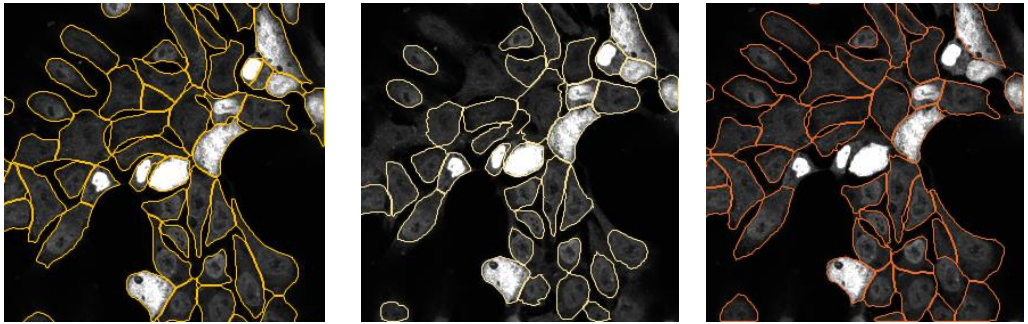
How to fine-tune the segmented Mask

What is the Norm percentile?

- Low contrast and High-intensity noise often cause problems for Cellpose.
- Norm percentile improves this issue by scaling the intensity values of the raw image to a specific range based on percentile thresholds.
- It improves the contrast and makes the segmented target area more obvious





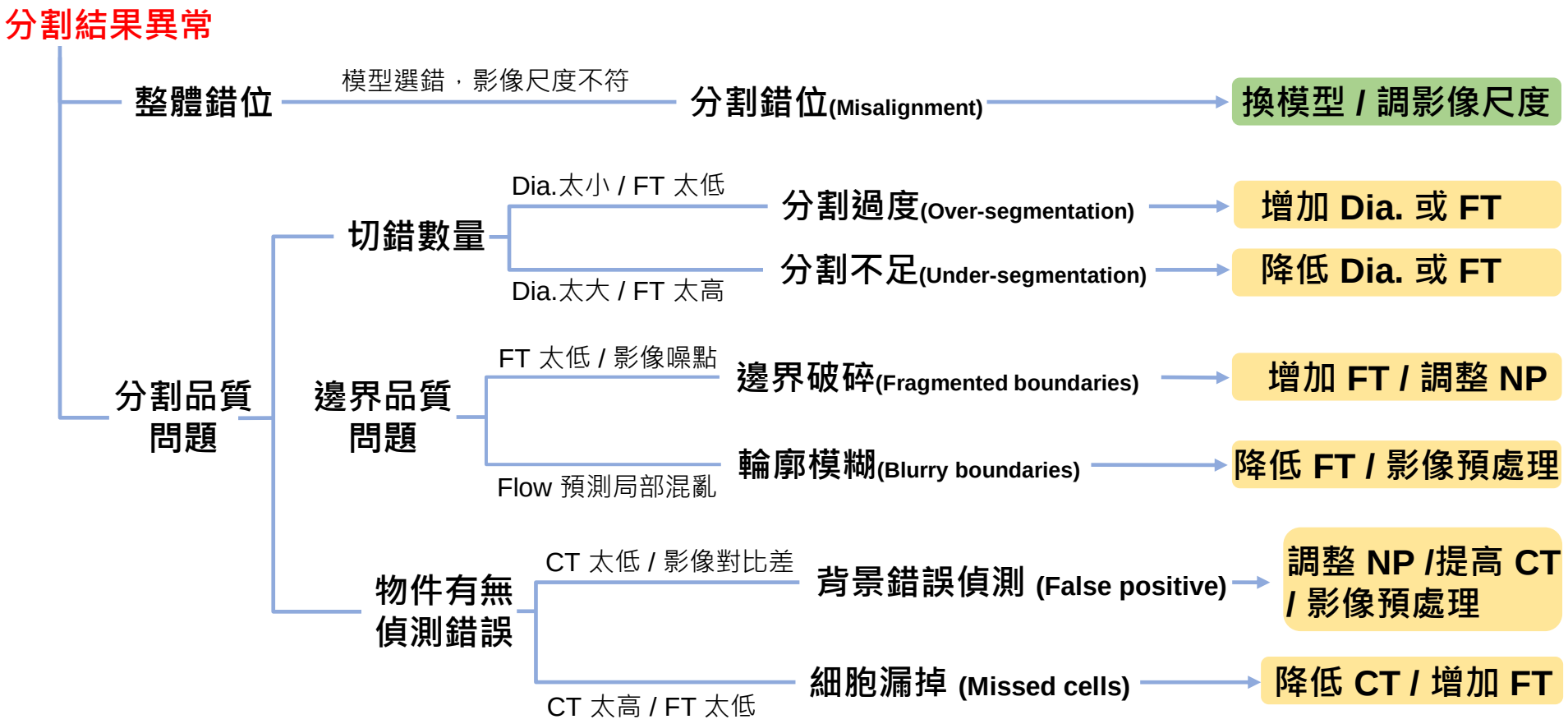


Default

Increase
Lower thr. to 65

decrease
Upper thr. to 78

Troubleshooting Cell Segmentation Errors

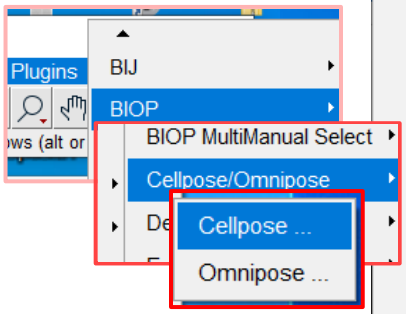


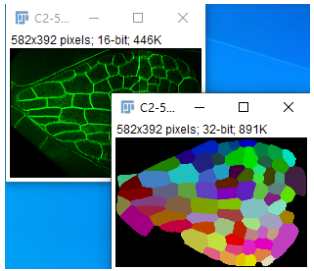
Embedding CellPose on your ImageJ analysis pipeline

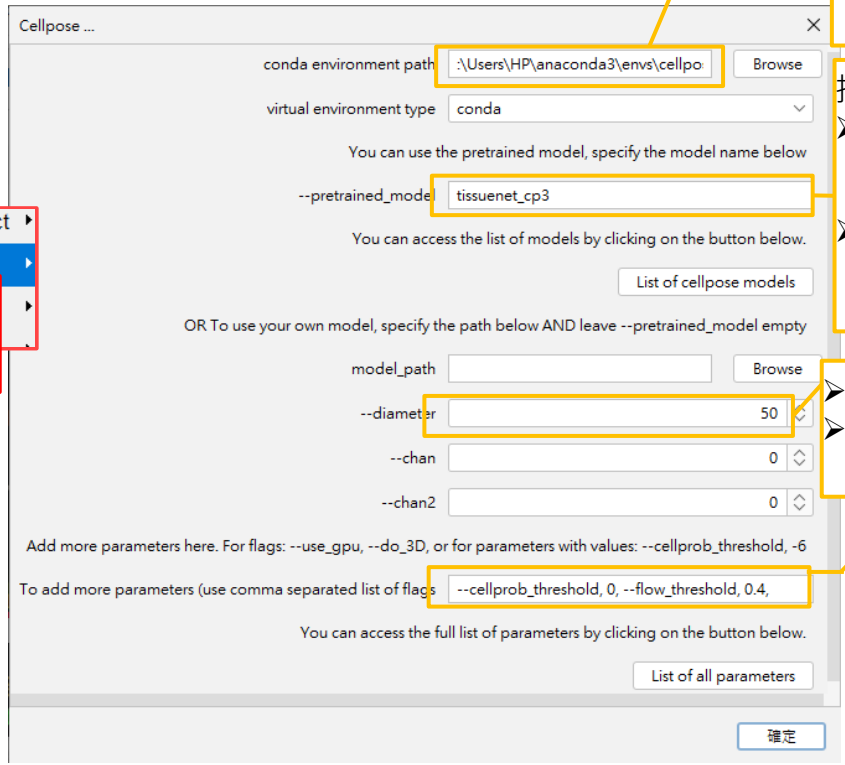
"BIOP: A plugin that integrates various tools into the Java environment.“

- Deep learning tools: Cellpose, Stardist(3D), Deepslice
- Registration program: Elastix, Transformix

Open BIOP







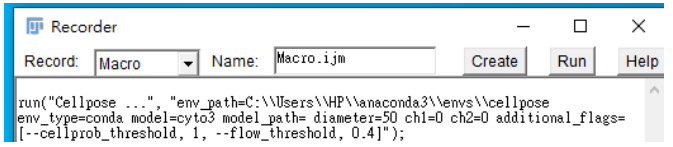
填入CellPose 安裝環境的路徑
e.g.
C:\Users\HP\anaconda3\envs\cellpose

指定你想要用的預訓練模型

- Full built-in model
  - For Cytoplasm : cyto3, cyto2, cyto (30 pixel)
  - For Nucleus: nuclei (17 pixel)
- Data specific model
  - tissuenet_cp3, livecell_cp3, bact_phase_cp3, yeast_PhC_cp3, bact_fluor_cp3, deepbacs_cp3

- 設定物件直徑，注意單位為 pixel
- 更改參數:

--cellprob_threshold, Value, --flow_threshold, Value,



Reference

➢ Part I & II

- **AI ML DL in Bioimage Analysis webinar:**
  https://www.youtube.com/watch?v=TJXNMIWtdac&t=120s
- **Machine Learning for Pixel and Object Classification:**
  https://www.youtube.com/watch?v=dstjhCPBDOY&t=181s
- **【機器學習2021】卷積神經網路 (Convolutional Neural Networks, CNN):**
  https://www.youtube.com/watch?v=OP5HcXJg2Aw&t=2899s
- Jan M, Spangaro A, Lenartowicz M, Mattiazzi Usaj M. From pixels to insights: Machine learning and deep learning for bioimage analysis. *Bioessays*. 2024;46(2):e2300114. doi:10.1002/bies.202300114
- Ali M, Benfante V, Basirinia G, et al. Applications of Artificial Intelligence, Deep Learning, and Machine Learning to Support the Analysis of Microscopic Images of Cells and Tissues. *J Imaging*. 2025;11(2):59. Published 2025 Feb 15. doi:10.3390/jimaging11020059

Reference

➢ Part III -LabKit

- **Labkit: Labeling and Segmentation Toolkit for Big Image Data:**

  https://www.youtube.com/watch?v=UZjZtmm7adU

- Arzt, M., Deschamps, J., Schmied, C., Pietzsch, T., Schmidt, D., Tomancak, P., … Jug, F. (2022). LABKIT: Labeling and Segmentation Toolkit for Big Image Data. Frontiers in Computer Science, 4. doi:10.3389/fcomp.2022.777728

➢ Part III-StarDist

- **Introduction to nuclei segmentation with StarDist - [NEUBIASAcademy@Home] Webinar:**

  https://www.youtube.com/watch?v=Amn_eHRGX5M&t=2030s

- Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell Detection with Star-Convex Polygons. In Lecture Notes in Computer Science (pp. 265–273). Springer International Publishing. doi:10.1007/978-3-030-00934-2_30

➢ Part III-CellPose

- Stringer C, Pachitariu M. Cellpose3: one-click image restoration for improved cellular segmentation. Nat Methods. 2025;22(3):592-599. doi:10.1038/s41592-025-02595-5
- Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods. 2021;18(1):100-106. doi:10.1038/s41592-020-01018-x

Reference

➢ **Part III-CellPose**

- **Cellpose2: human-in-the-loop model training (2x speed)**

  https://www.youtube.com/watch?v=3Y1VKcxjNy4

課後意見調查

附錄：Segmentation 更深的理解（選讀）

When Segmentation Cannot Be Fixed by Tuning….(1/2)

情境

- 因為影像本身對比度太差，導致一顆細胞被過度分割成許多小片段
- 直徑大小（Diameter）和影像強度縮放範圍（Normalization Percentile）都已經設定合理
- 在 CellPose 中，單靠調參數也無法修復分割(Segmentation)

狀況 1 - Over-segmentation

📌 CellPose 無法「重新合併」分散的細胞碎片

- **原因:**
  1. 細胞內部 flow 混亂，被預測出多個中心，導致分割破碎。
  2. FT 只能進行輪廓，也就是細胞層級篩選；而 CT 只能進行像素層級篩選。
  3. flow 無法透過 FT 與CT 改變，CellPose 亦無全局輪廓推理能力。
- **解法:**
  1. 提升影像品質(調整局部對比、去噪)
  2. 後處理合併小片段
  3. 換用更強的模型（如Omnipose）
  4. 接受影像的自然限制
  5. 模型再訓練

When Segmentation Cannot Be Fixed by Tuning….(2/2)

狀況 2 - Under-segmentation

📌 CellPose 無法「正確分開」連在一起的細胞。

- **原因:**
  1. 細胞內部 flow 混亂，兩顆顆細胞內部的 flow 無法形成清楚分開的收斂中心。
  2. FT 主要篩選不穩定 segmentation ，無法新增正確邊界；而 CT 只能調整那些像素被保留。
  3. flow 結構無法改變，預測錯誤是根本問題，CellPose 無法事後幫助兩顆細胞分離。
- **解法:**
  1. 提升影像品質(調整局部對比、去噪，讓模型能預測出兩個清楚中心)
  2. 換用更強的模型（如Omnipose）
  3. 後處理手動切割(針對少數錯誤合併物件)
  4. 接受影像的自然限制
  5. 模型再訓練

✳️ 總結:

- 大部分 segmentation 問題，都是「預測錯誤」+「參數篩選失當」雙重造成。
- 明確區分 pixel 層級錯誤 與 細胞層級篩選錯誤，是理解和改善結果的關鍵。
- 熟悉原理，才能真正理解模型的極限與可能。不只是調參，而是能夠自主判斷問題與選擇最佳解法。