Automatic Detection of Metastatic Breast Cancer Cells on Histopathological Slides of Lymph Nodes

Charles Boy de la Tour, Megan Fillion, Nicolas Grevet

CentraleSupélec

TABLE OF CONTENTS

Motivation and dataset

01

CentraleSupélec

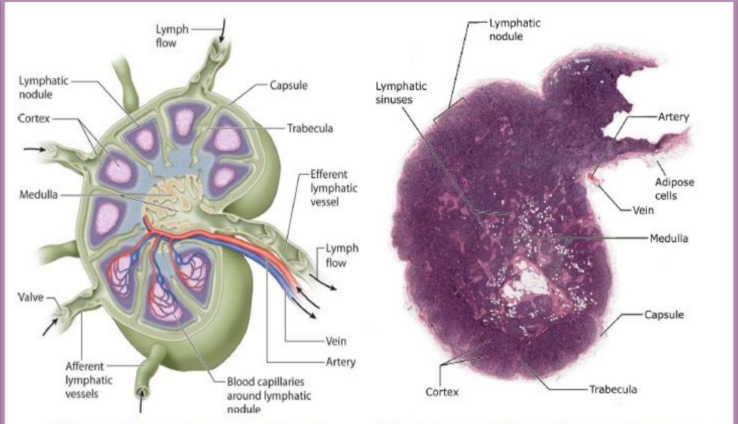
Motivation

- In 2020, an estimated **684,996** women across the world died from breast cancer.

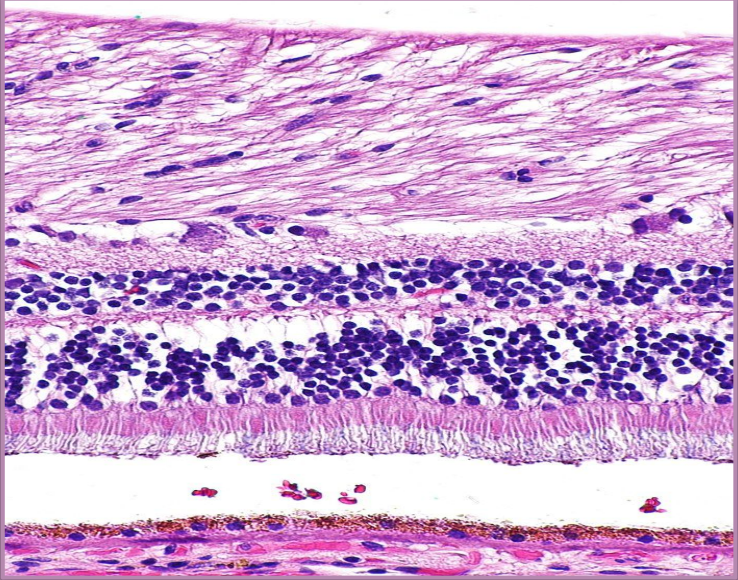
- If the breast cancer is located only in the breast, the 5-year survival rate of women with this disease is **99%**. If the cancer has spread to the regional lymph nodes, the 5-year survival rate is **86%**.

  ⇒ Early diagnosis is key.

- H&E stain is the combination of two histological stains: **hematoxylin (cell nuclei)** and **eosin (extracellular matrix and cytoplasm)**.



(a) : Illustrative lymph node image (b) : Histopathological lymph node image



CentraleSupélec

Deep Learning in Histopathology

Jeroen van der Laak, Geert Litjens, Francesco Ciompi

Advancements in CPATH are due to advancements in the field of Computer Vision (pretrained image models), improvements in microscopic scanning devices (WSIs) and public access to Datasets (CAMELYON)

Why deep Learning?

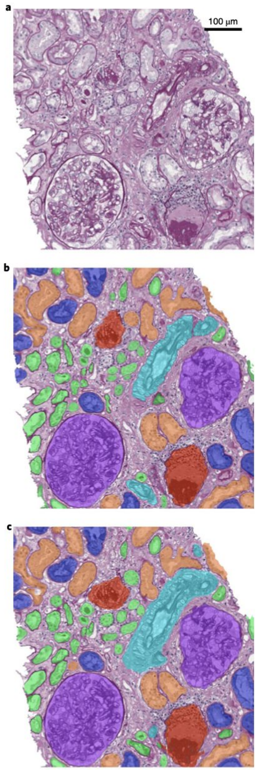
- Deep Learning networks learn meaningful and hierarchical features autonomously representations and outperform more traditional image analysis methods.

How to treat large or partially annotated data?

- Weakly supervised classification segment image with sparse/incomplete annotations
- End to end trained CNNs cannot be applied to full WSI, solution: tiling

Challenges:

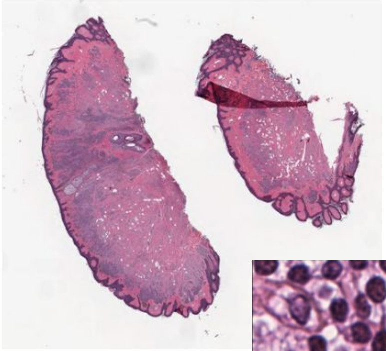
- Low performance on images with different staining / from different laboratory
    - Must add as much variation as possible: color shifting/blurring/random rotations
- Models only detect pathologies they were learned to recognize
    - Training models on all pathologies (costly)
    - Output class for segmentation “I don’t know”



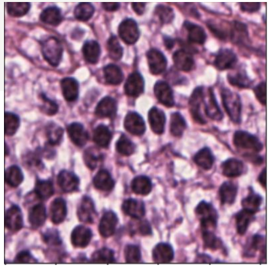
CPATH forkidney tissue segmentation

CentraleSupélec

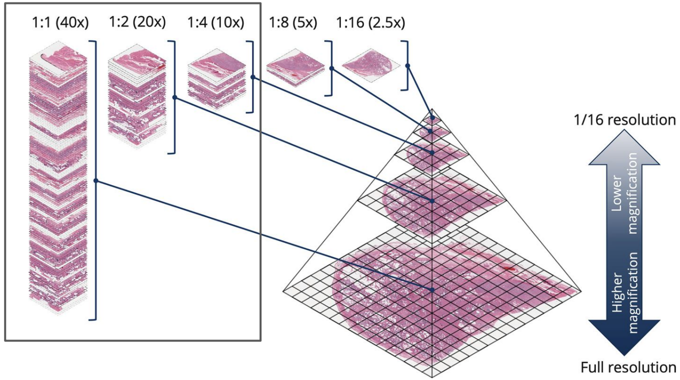
Datasets

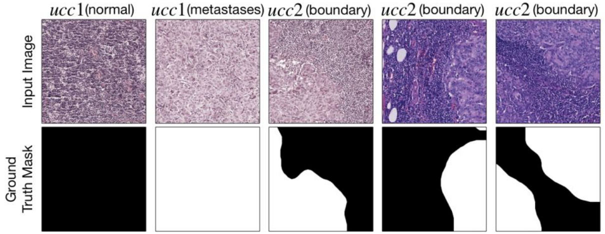


WSI



Patch size 256x256





Ground Truth Mask patch level

Gustave Roussy Cohort (GR)

- Private in-house dataset
- 236670 tiles
- 153 patients
- 70 positive WSIs
- HES and IHC stained slides together with **Metastatic Region Annotation**

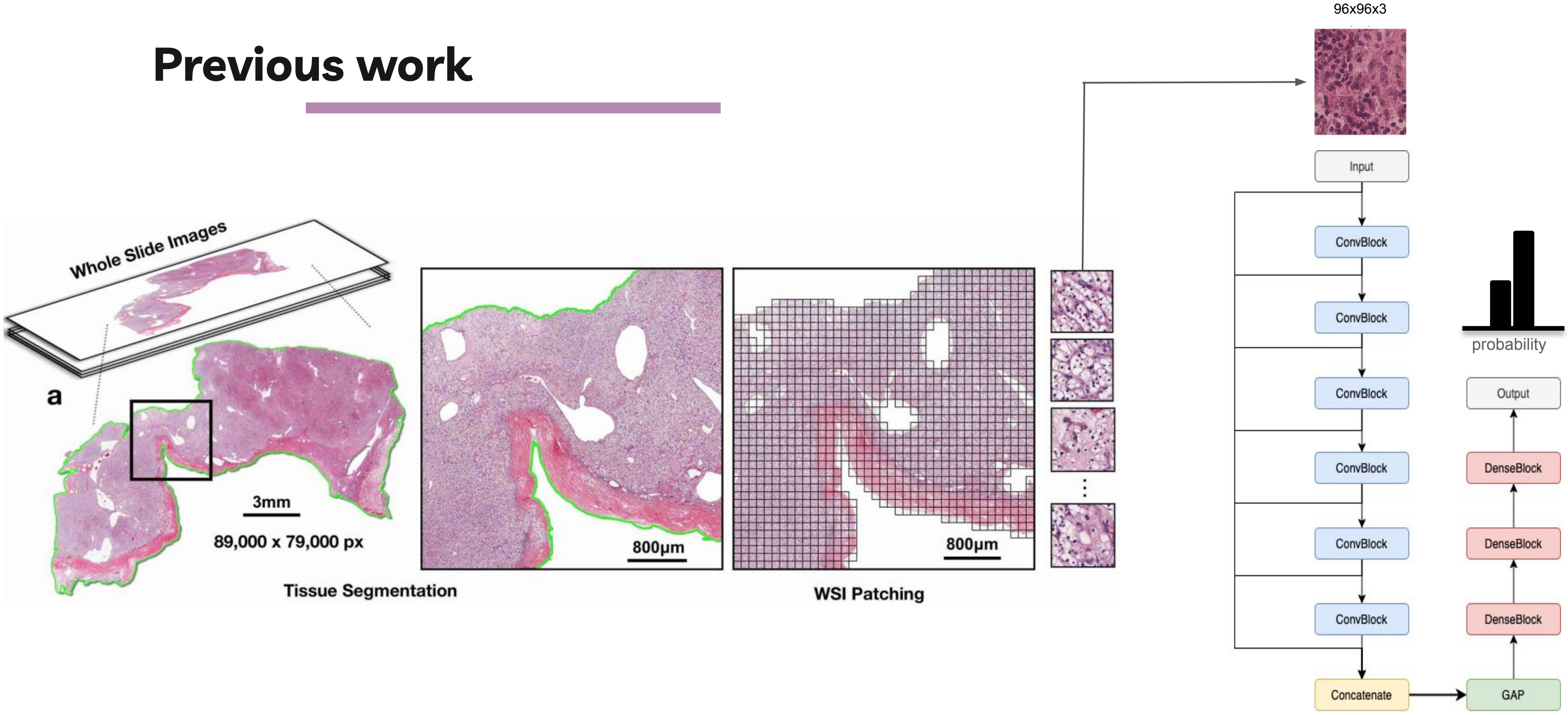
CentraleSupélec

Previous Work

02

CentraleSupélec

Previous work

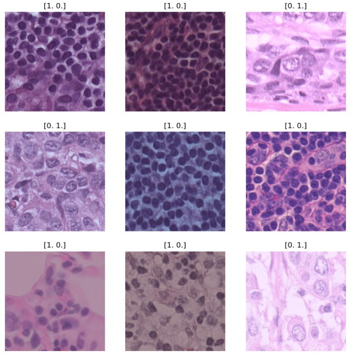


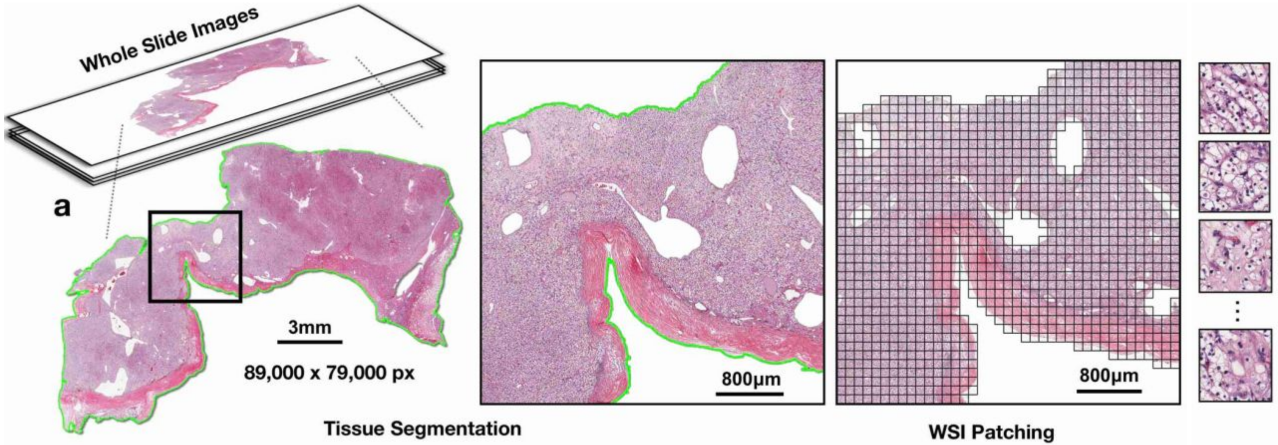
ResNet50	VGGLite	LNNet	ExtremeLungNet

Deep Learning Model
Architecture

CentraleSupélec

Limitations of the approach





Label
0
1
1
0

At least 1 pixel is classified as Tumor

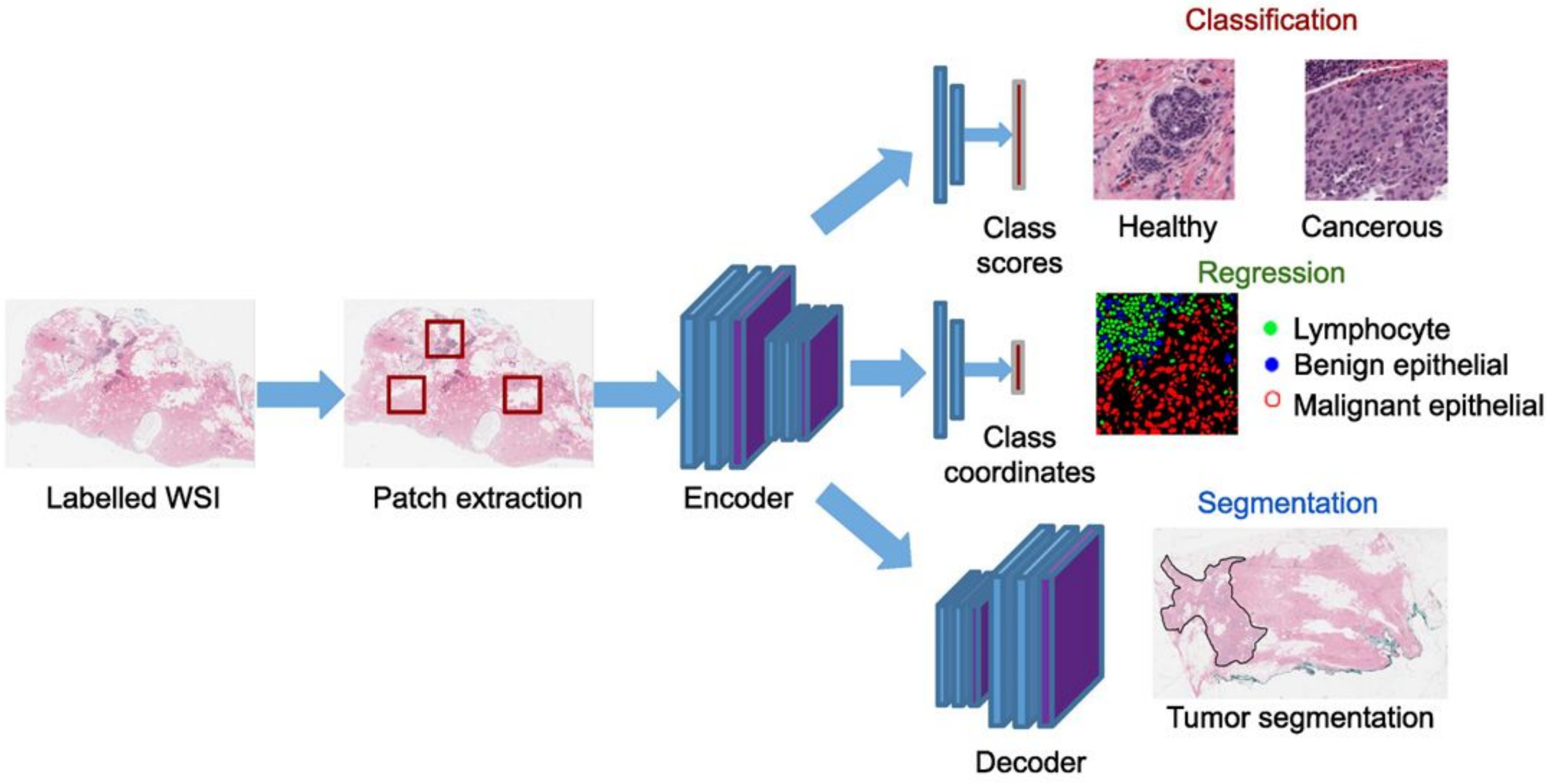
Lot of False Positives in the predictions

Segmentation approach

03

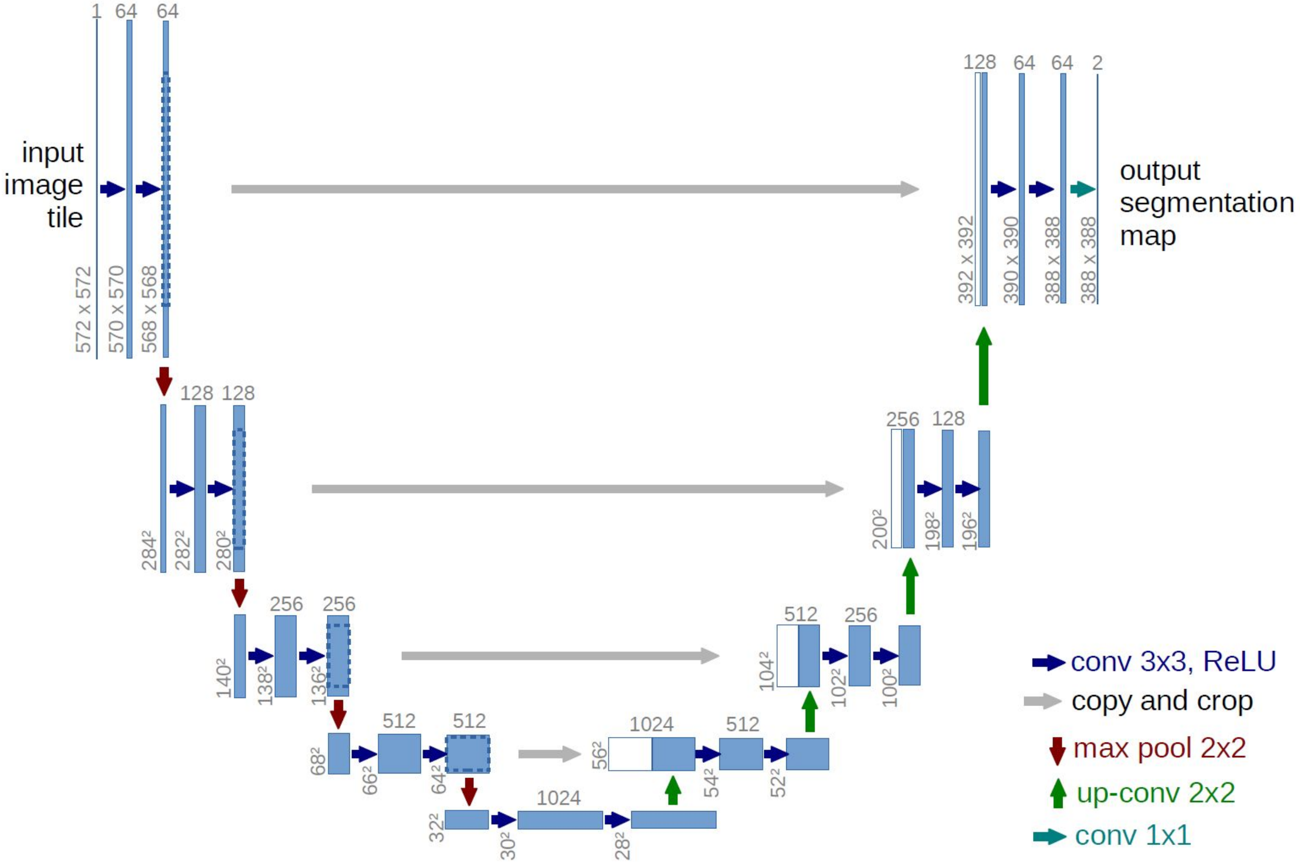
CentraleSupélec

From a classification to a segmentation problem



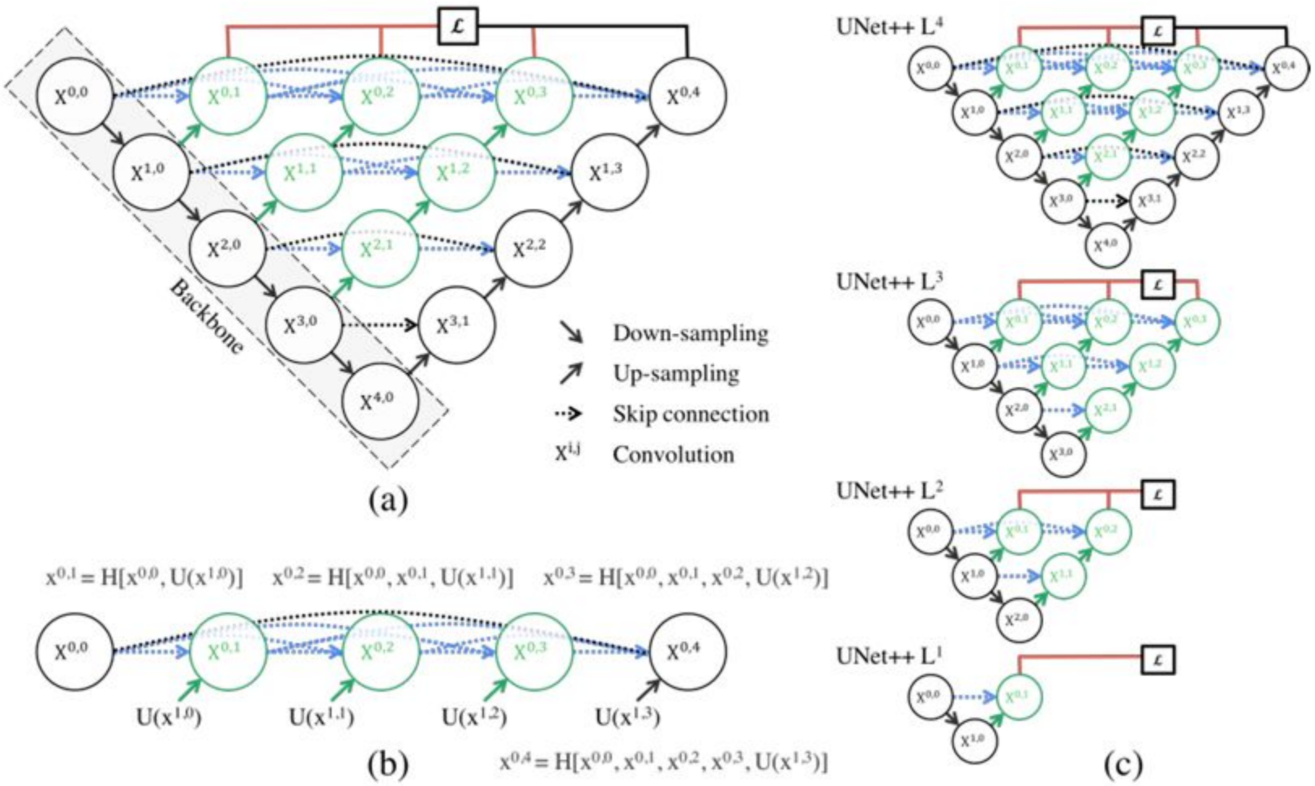
CentraleSupélec

UNet



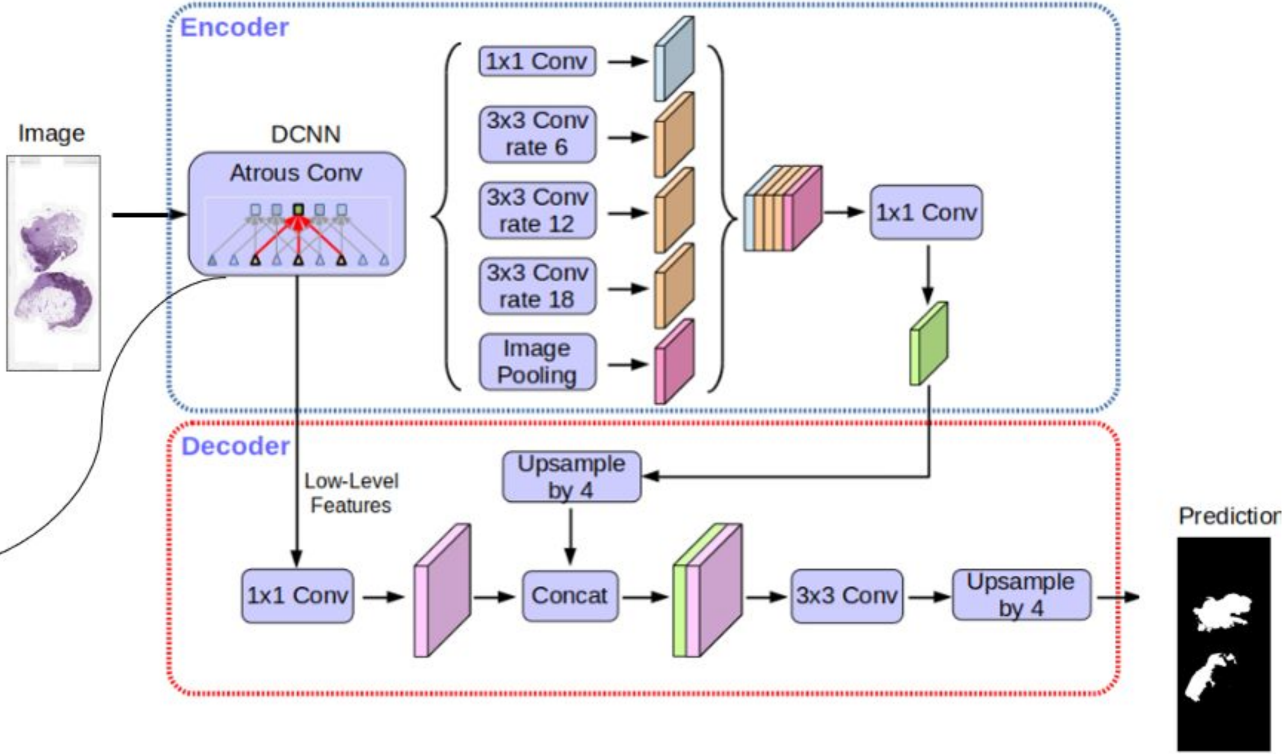
input image tile
1
64
64
572 x 572
570 x 570
568 x 568
128
128
284²
282²
280²
256
256
140²
138²
136²
512
512
68²
66²
64²
1024
32²
30²
28²
1024
512
56²
54²
52²
512
256
104²
102²
100²
256
128
200²
198²
196²
128
64
64
2
392 x 392
390 x 390
388 x 388
388 x 388
output segmentation map
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1

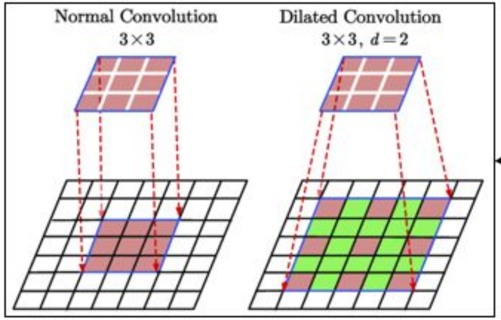
UNet++



CentraleSupélec

DeepLabV3+





CentraleSupélec

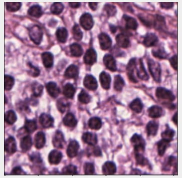
Results & Analysis

04

CentraleSupélec

Quantitative results

Magnification 10x:



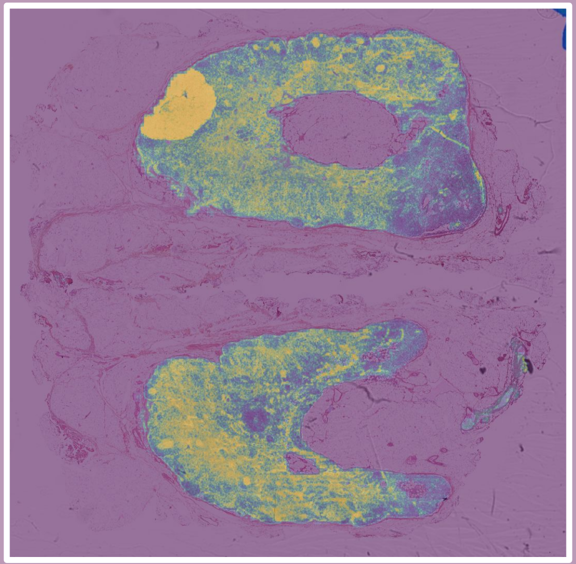
	Patches Accuracy	Pixel Accuracy	MIoU	Dice	False positives rate
UNet	0,842	0,800	0,664	0,835	0,202
UNet++	**0,882**	**0,873**	**0,760**	**0,884**	0,095
DeepLabV3+	0,855	0,850	0,720	0,852	**0,083**

Encoder (backbone) of models : Efficientnet-B0

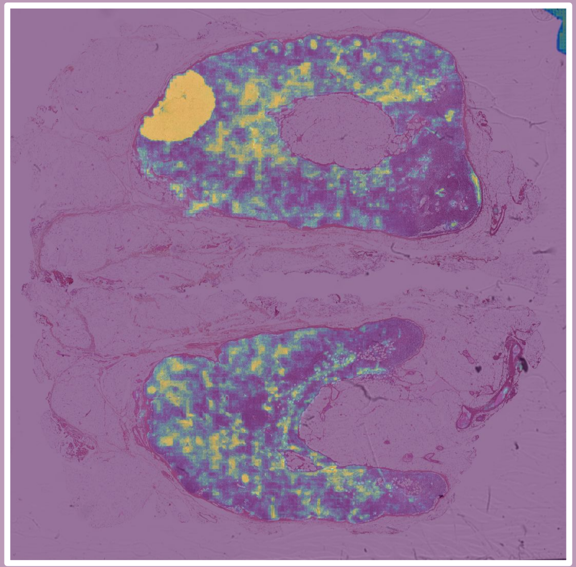
Evaluation of the trained models on a set of **+2100** annotated Gustave Roussy HES patches
Patches accuracy and **mean false positives rate** reported for patch-level evaluation
Means of **pixel accuracy**, **MIoU** and **dice** for the segmentation task

CentraleSupélec

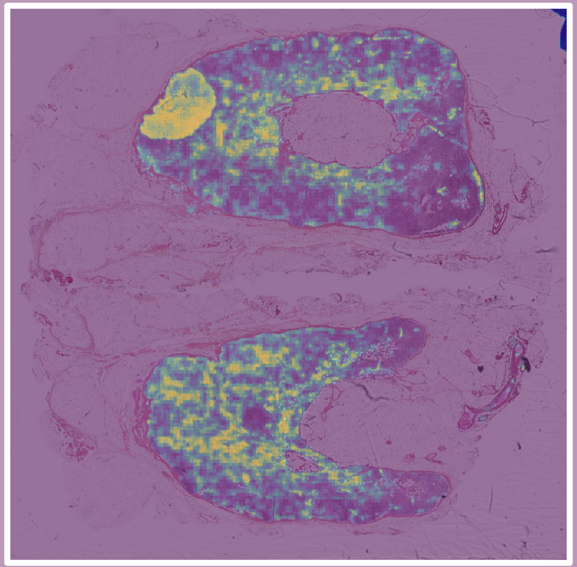
WSI level prediction



UNET



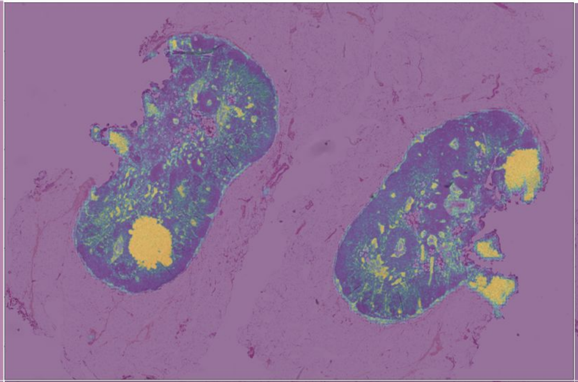
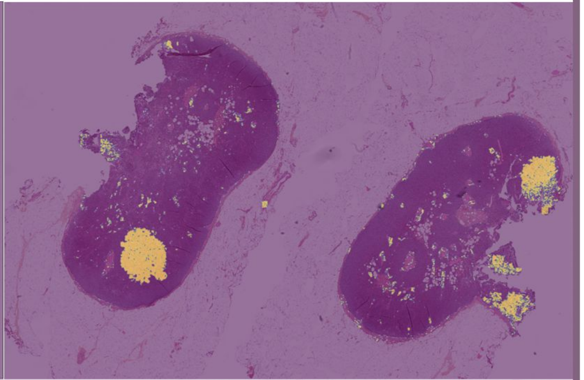
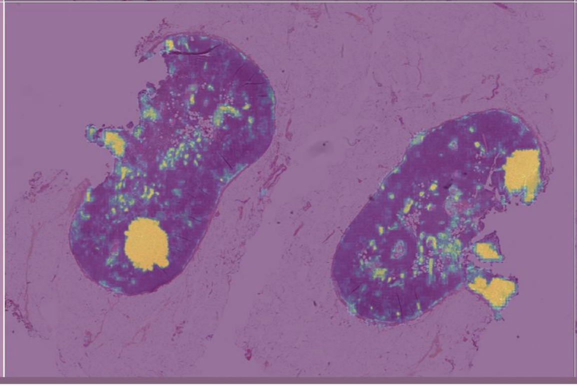
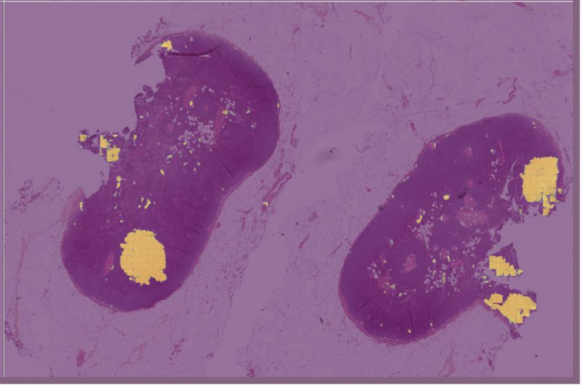
UNET++



DEEPLABV3+

CentraleSupélec

WSI level prediction

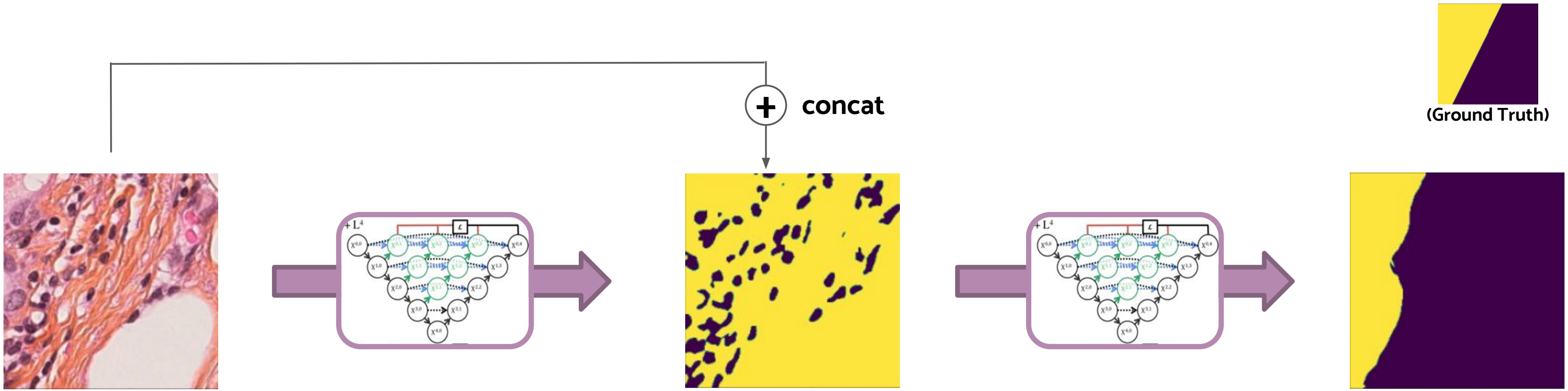
Models	Heat Map Probability	>90% Prediction
Unet		
Unet ++		

Going Further

05

CentraleSupélec

Additional Input: Nuclei Mask



Input Patch
(256x256x3)

UNET++
Task: Segment Nuclei

Input Patch + Nuclei Mask
(256x256x4)

UNET++
Task: Detect Metastasis

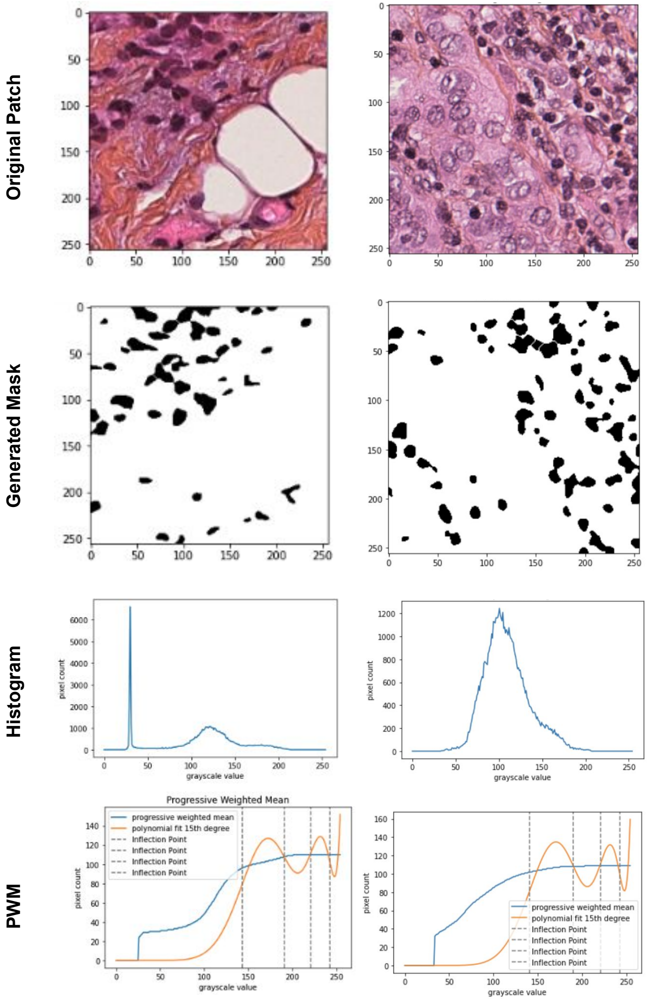
Output
Metastasis Detection Mask

Generating Nuclei Segmentation Mask: MANA

Automated nuclei detection algorithm that works on **different types of tissue** (bone, prostate, adrenal gland and thyroid) and **various magnifications** (10x, 20x, 40x)

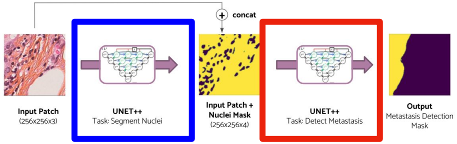
Procedure:

1. Preprocess input patch: Grayscale, invert, blur
2. Compute Grayscale Histogram
3. Compute Progressive Weighted Mean Curve & Fitted Polynomial Curve (15th degree)
4. Generate masks using inflection points as thresholds
5. Choose threshold that yields highest mean area of detected objects
6. Post Processing: remove small objects and resegment clusters



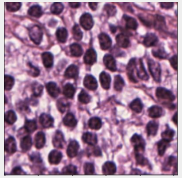
CentraleSupélec

Multi-tissue and multi-scale approach for Nuclei Segmentation in H&E stained Images - Salvi & Molinari

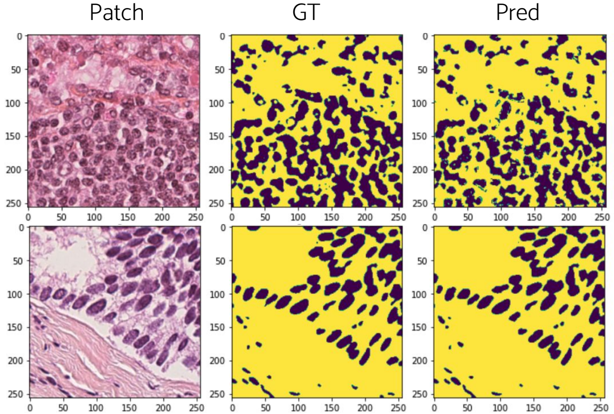


Results

Magnification 20x:

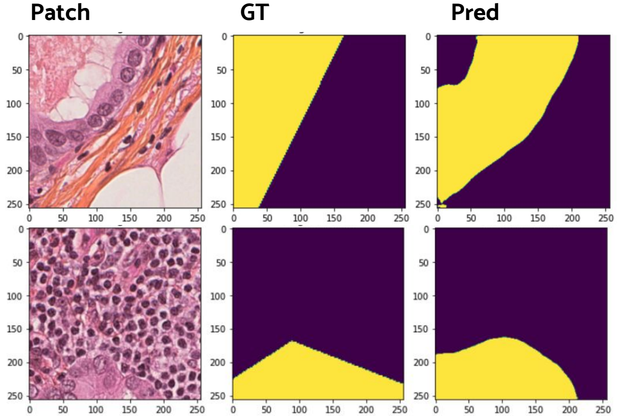


Nuclei Segmentation



	pixel accuracy	pixel recall	pixel precision	dice
nuclei segmentation	0.956	0.987	0.956	0.998

Metastasis Detection



model	**pixel accuracy**	**dice**	**patch accuracy**	**false negative rate**	**false positive rate**
Unet++	0.854	0.837	0.859	0.131	**0.146**
Unet++ with Nuclei Mask	**0.855**	**0.855**	**0.862**	**0.108**	0.1742

CentraleSupélec

Conclusion and Limitations

CentraleSupélec

Limitations

Magnification:

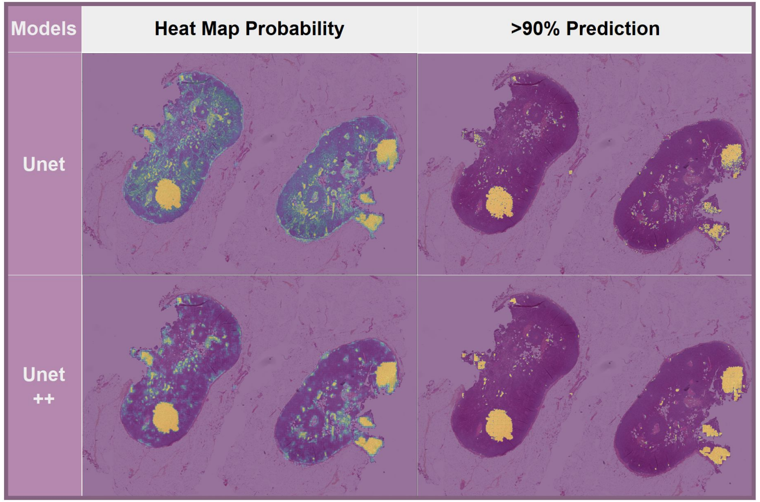
- 10x for all of the models presented in **segmentation approach**
- Metastasis detection using nuclei mask uses magnification 20x

Thresholding:

- Pixel level (pixel metrics, whole slide mask)
- Patch level (patch metrics)

Exhaustive Model Comparison/Complexity

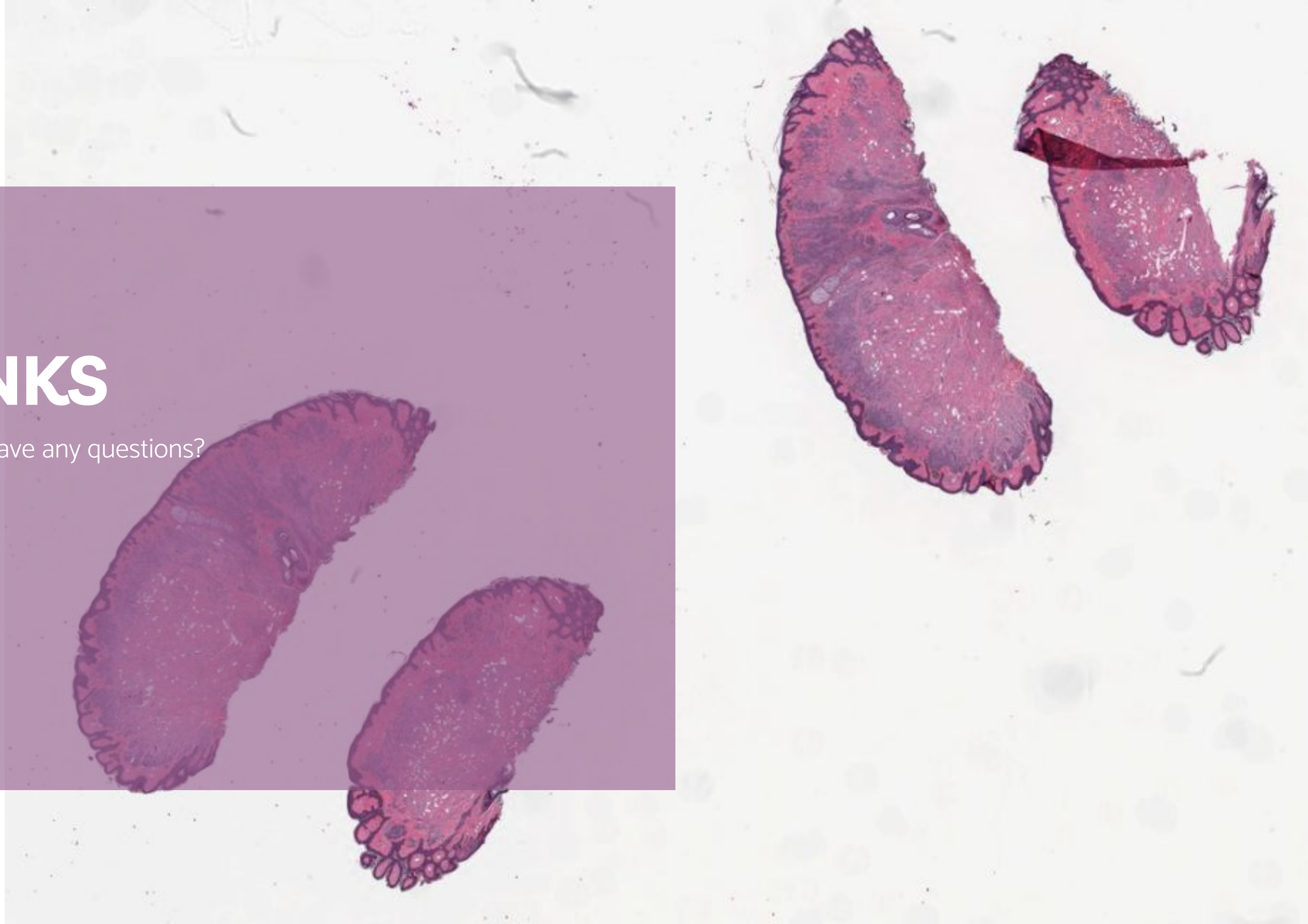
Only one dataset used (GR Seg)



CentraleSupélec

THANKS

Does anyone have any questions?



CentraleSupélec