Computational pathology: a new paradigm for studying the micro-tumoral environment

Prof. Nicolas Loménie

Context : the Whole Slide Images (WSI)

Figure 1. This image is a WSI. At a low magnification (x1.5).
At the highest resolution for instance 40x it weights between 5 and 10 GB of data just for one single patient examination.
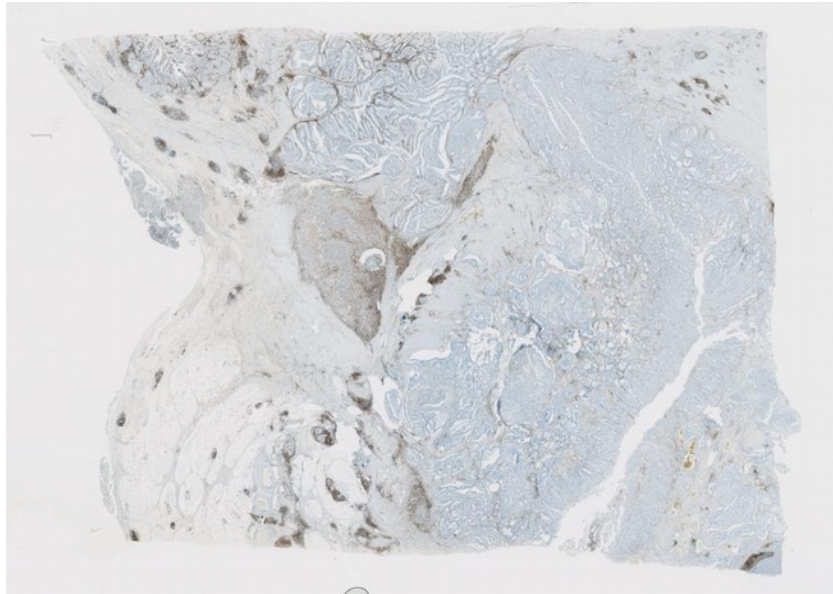
Figure 1. (follow.) By automatic analysis of thousands of them we can build up efficient models for immuno-oncology treatments for instance. Deep learning is a core mechanism to achieve this goal.
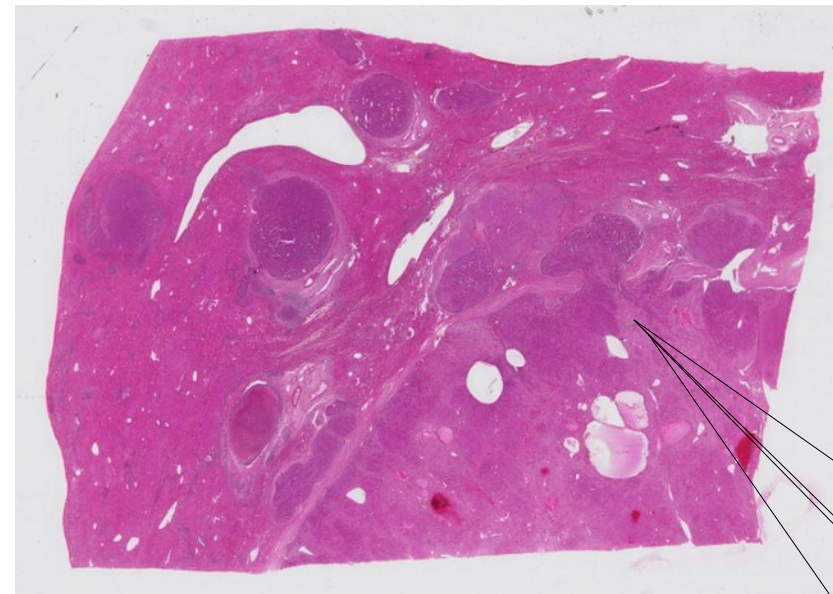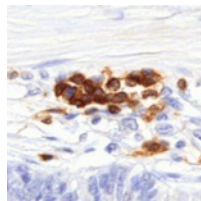
X1.5
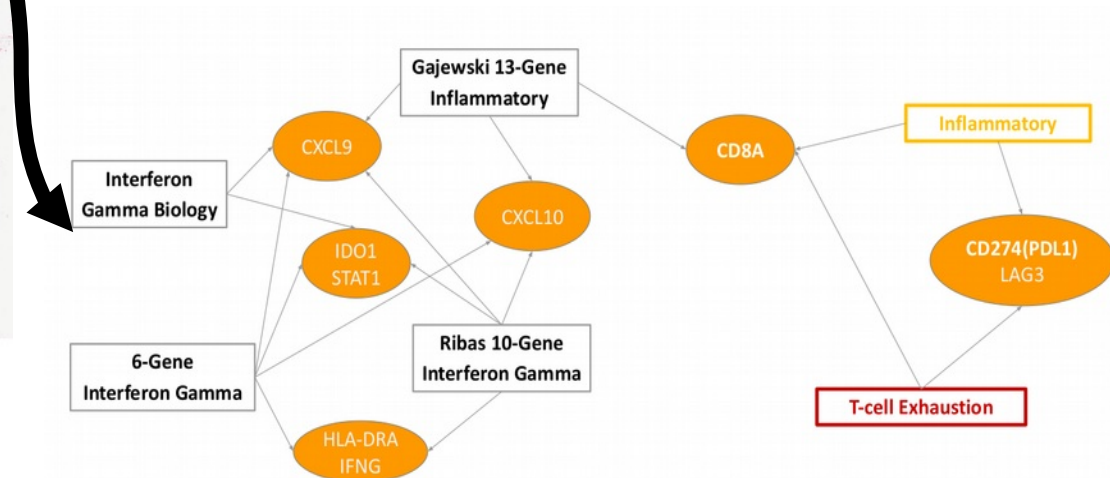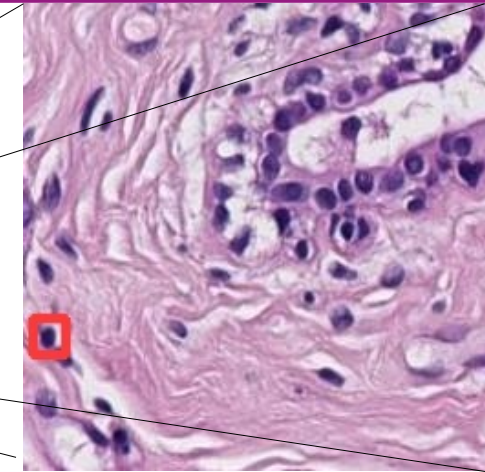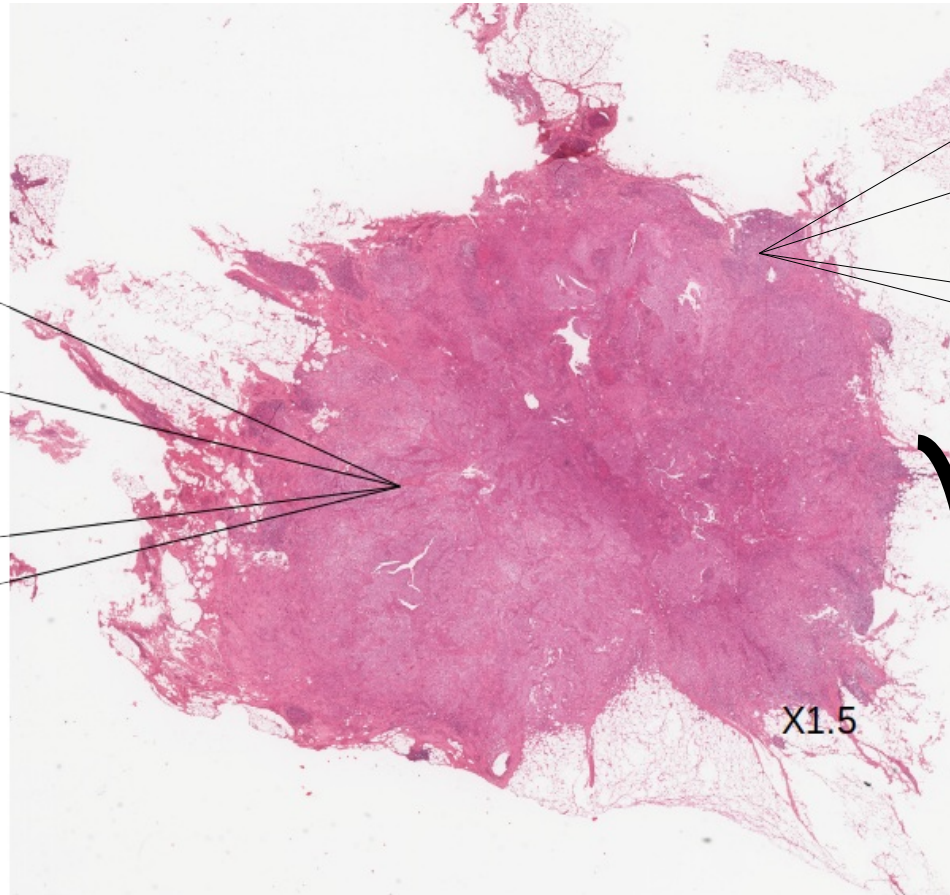
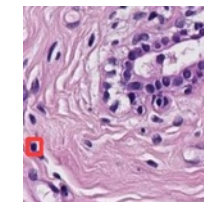# Context : the Whole Slide Images (WSI)

WSI : Whole Slide Image : digital version of a tissue specimen (x5, x20, x40)

Immunohistochemistry staining
69632pixels x 48384pixels, 9.41 GB uncompressed

Stained with hematoxylin and eosin (H&E)
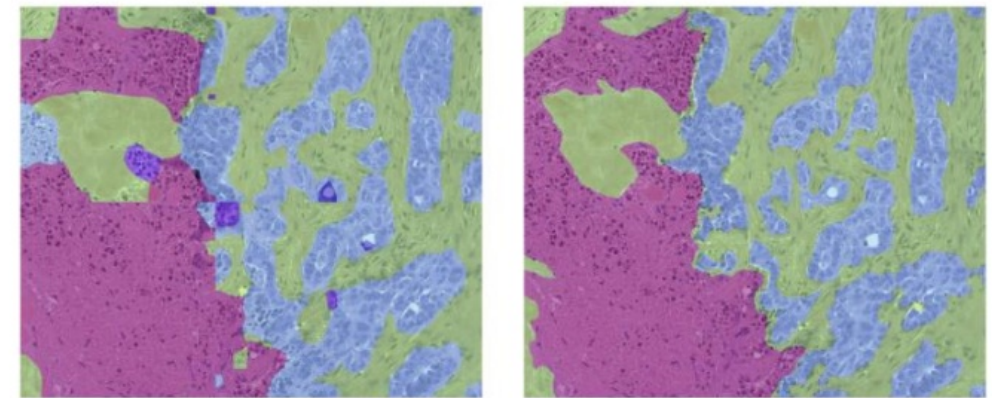59520pixels x 41216pixels, 6.85 GB uncompressed

2 current research programs and 1 translational one Image and omics for immunotherapy

Gene signature prediction based on WSI (CSC funding)

Contextual exploration of WSI (DiiP funding)

Highly predictive

Non informative

In 3 signatures:
- 6G Interferon Gamma
- Interferon Gamma biology
- Ribas 10G Interferon Gamma

Enrichment in **lymphocytes, plasma cells, and neutrophils.**
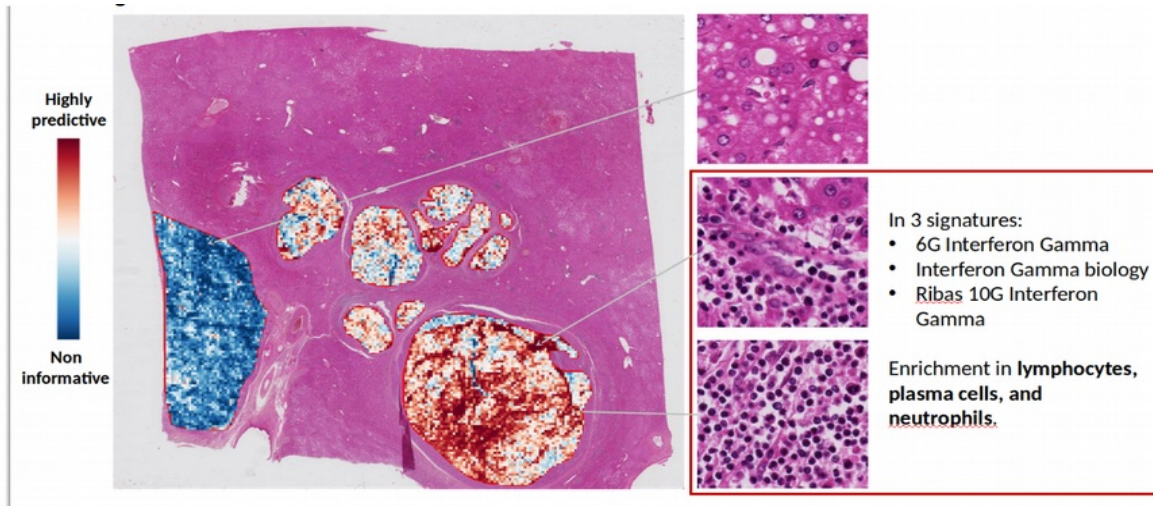
Prediction

Ground truth

Qinghe Zeng,
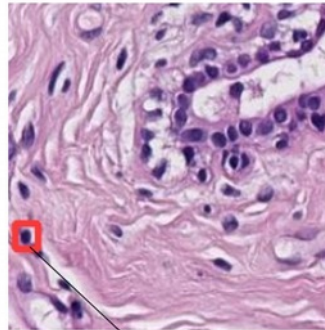PhD student, 3r^d year

Zhuxian Guo,
PhD student, 2nd year

2 current research programs and a translational one
Image and omics for immunotherapy

Skipping IHC and do-it-all with just H&E
PRT-K (INCa-DGOS) funding

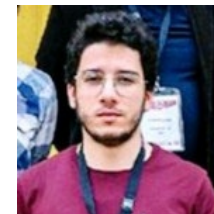Internal Image databases :

- Liver (HCC)

- Colon (CRC)

YOLO
Architecture
You Only Look Once
For **object detection**

Fine tuned for lymphocyte detection

Master 2 Informatique :
Parcours Vision Machine Intelligente
(VMI)
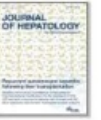
Amine Marzouki,
Research Engineer

Gene signature prediction based on WSI

[1]LIPADE, Université de Paris, Paris, France.
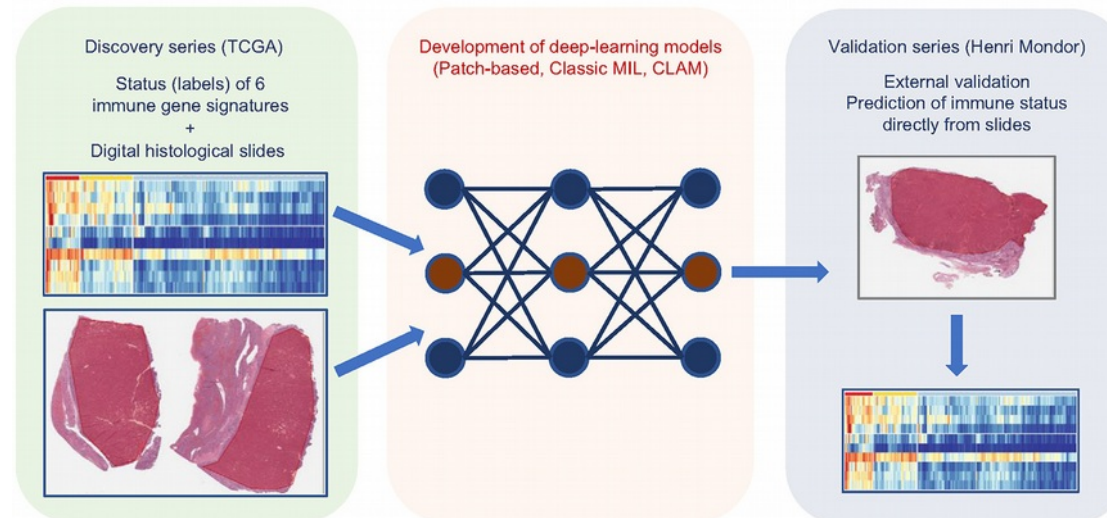[2]CHIC, INSERM U1138, Cordeliers Research Center, Paris, France.

Artificial intelligence predicts immune and inflammatory gene signatures directly from hepatocellular carcinoma histology

Qinghe Zeng [1, 2, †], Christophe Klein [1, †], Stefano Caruso [3], Pascale Maille [4, 5, 6], Narmin Ghaffari Laleh [7, 8], Daniele Sommacale [9], Alexis Laurent [9], Giuliana Amaddeo [10], David Gentien [11], Audrey Rapinat [11], Hélène Regnault [10], Cécile Charpy [4], Cong Trung Nguyen [5, 6], Christophe Tournigand [12], Raffaele Brustia [9], Jean Michel Pawlotsky [5, 6], Jakob Nikolas Kather [7, 8], Maria Chiara Maiuri [1] … Julien Calderaro [4, 5, 6] ♀ ⊞ ✉

Discovery series (TCGA)

Status (labels) of 6 immune gene signatures + Digital histological slides

Development of deep-learning models (Patch-based, Classic MIL, CLAM)

Validation series (Henri Mondor)

External validation Prediction of immune status directly from slides

From WSIs to gene signature prediction

- Medical objective : Stratification of HCC patients eligible to immunotherapy

- Fact : No standard of stratification in France

- **Research :** Some gene signature are predictive but costly

Sangro, Bruno, et al. "Association of inflammatory biomarkers with clinical outcomes in nivolumab-treated patients with advanced hepatocellular carcinoma." Journal of Hepatology 73.6 (2020): 1460-1469.

- **Opportunity** : histological specimen easily accessible in clinical routine (WSI)

- AI Objective : learn to predict gene signatures based on WSI

Omics data

	CXCL9	STAT1	CXCL10	IDO1	CD8A	HLA-DRA	IFNG	CD274 (PD-L1)	LAG3	CCL2	CCL4	GZMK	HLA-DMA	HLA-DMB	HLA-DOA	HLA-DOB	ICOS	IRF1	CCL5	CD27	CXCR6	CCR5	CXCL11	GZMA	PRF1	CD276	PDCD1LG2	TIGIT
6-Gene Interferon Gamma p(ORR) = 0.05, p(OS) = 0.009	■	■	■	■		■	■																					
Gajewski 13-Gene Inflammatory p(ORR) = 0.04, p(OS) = 0.05	■		■		■					■	■	■	■	■	■	■	■	■										
Inflammatory p(ORR) = 0.05, p(OS) = 0.01		■						■	■																			
Interferon Gamma Biology p(ORR) = 0.07, p(OS) = 0.008	■	■		■															■	■	■							
Ribas 10-Gene Interferon Gamma p(ORR) = 0.07, p(OS) = 0.02	■	■	■	■		■	■															■	■	■	■			
T-cell Exhaustion p(ORR) = 0.03, p(OS) = 0.04					■			■	■																	■	■	■

[1]Sangro, Bruno, et al. "Association of inflammatory biomarkers with clinical outcomes in nivolumab-treated patients with advanced hepatocellular carcinoma." Journal of Hepatology 73.6 (2020): 1460-1469.

Image Data

| Discovery series: TCGA (n=336 cases)	External validation series: Hospital Henri Mondor (n=139 cases)
https://portal.gdc.cancer.gov/	

- Gene expression data: RNA sequencing
- WSIs: n=349
  a) Scanner: Aperio scanner
  b) Magnification (20X): ~128x128 µm², ~0.5 µm/pixel
  c) Staining: Hematein-eosin (H&E)
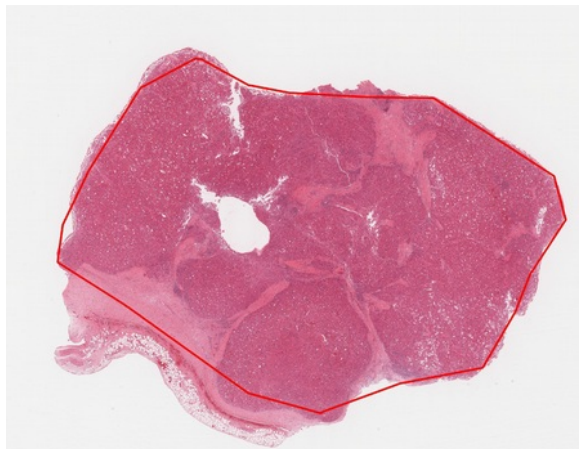
- Gene expression data: Nanostring (Panel IO360)
- WSIs: n=139 resections + 7 pre-operative biospies
  a) Scanner: Hamamatsu
  b) Magnification (20X): ~115x115 µm², ~0.45 µm/pixel
  c) Staining: Hematein-eosin-saffron (HES)
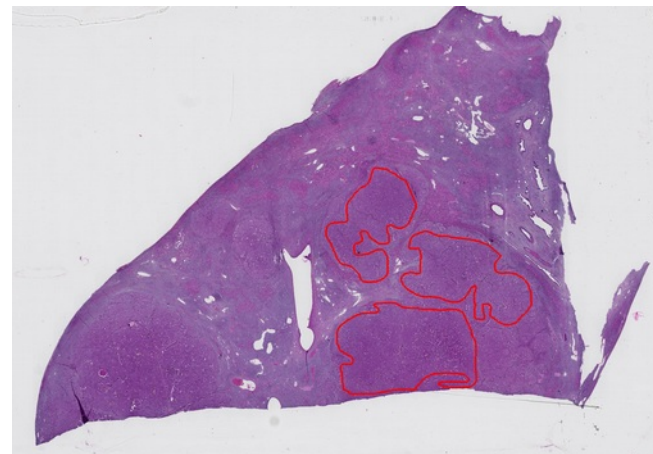
Annotations : tumoral areas annotated by PUPH Julien Calderaro

—— Tumor annotation

—— Tumor annotation

AI methodology for image-omics
Weakly Supervised Learning with Attention
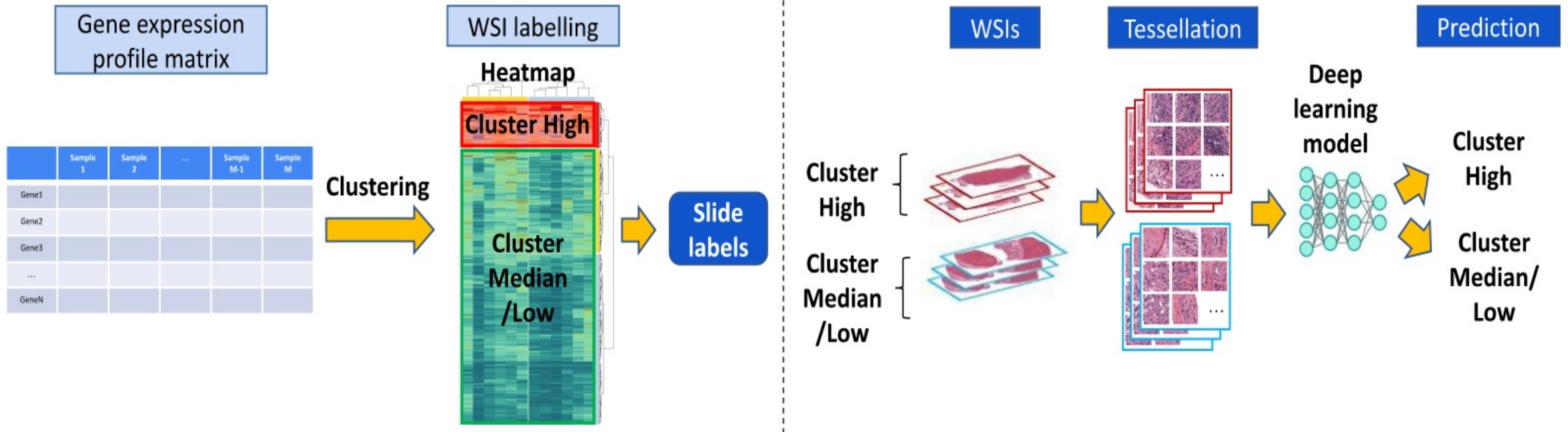
Discovery series n = 336
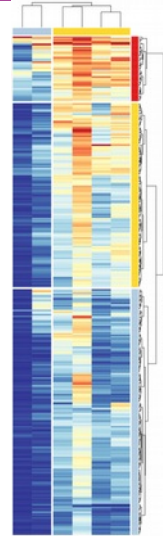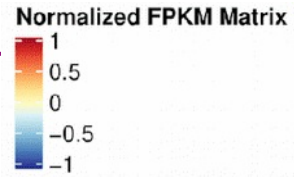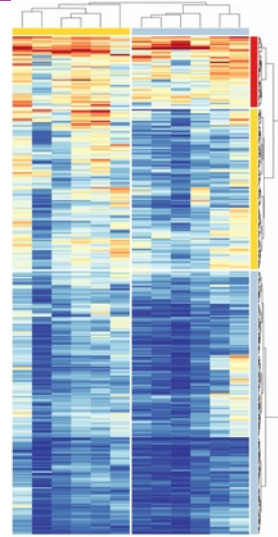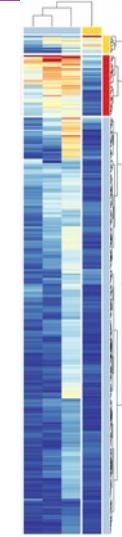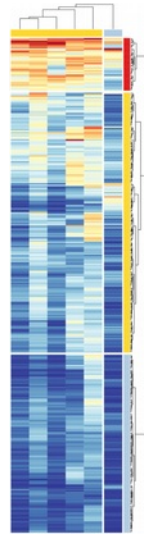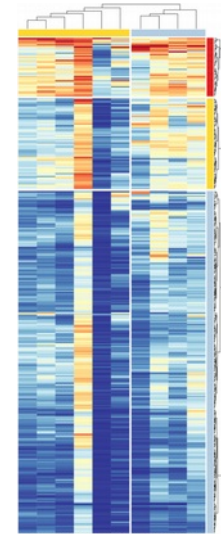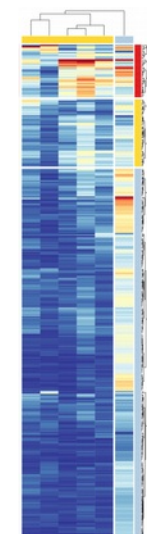
Hierarchical clustering : **rows** = samples/patients , **columns** = genes

1 patch = 256 x 256 pixels x 3
~ 200 000 values

WSI (bag)

N patches (instances)

Feature encoder

N * 1024 features
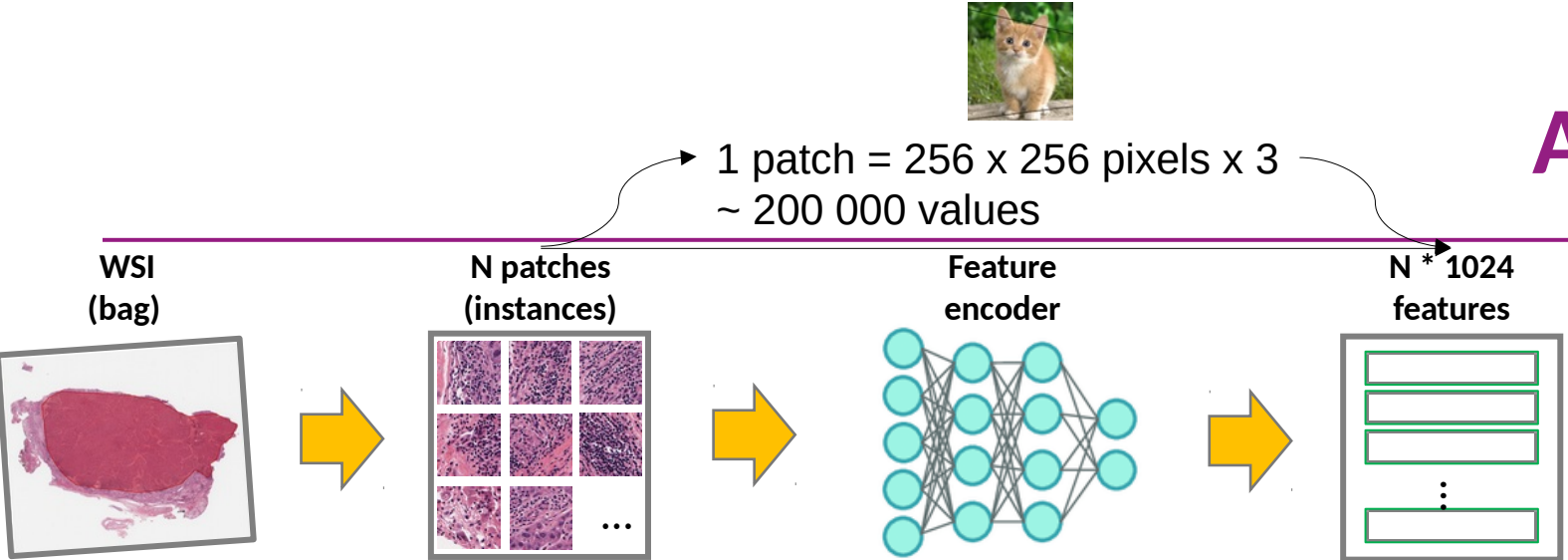
CLAM (clustering-constrained-attention multiple-instance learning)*
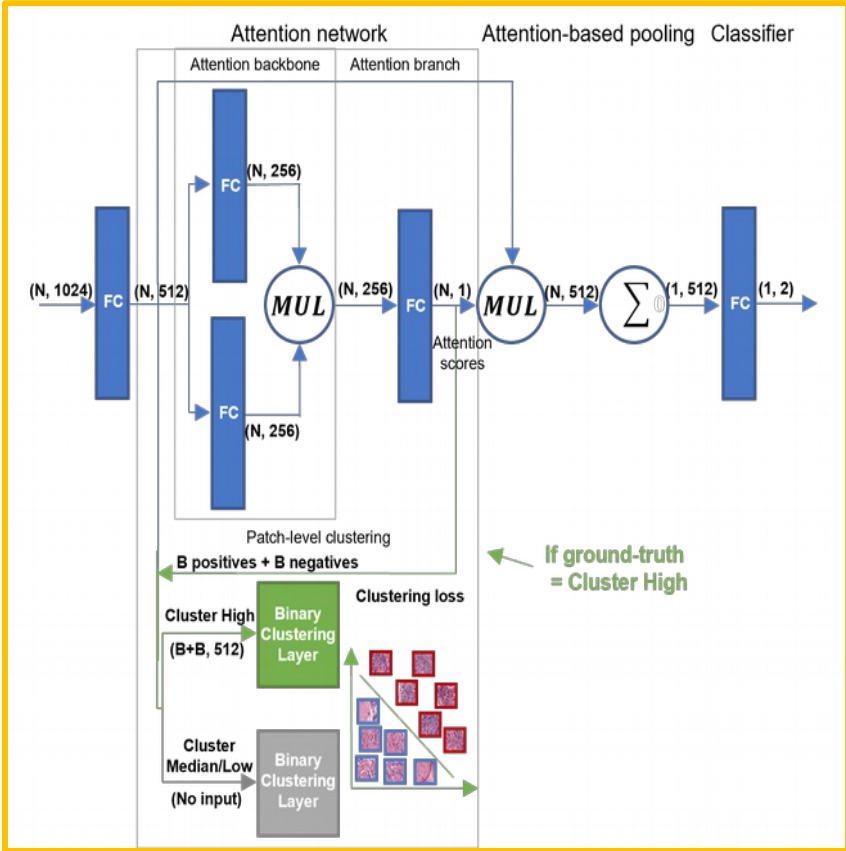
*Lu, Ming Y., et al. "Data-efficient and weakly supervised computational pathology on whole-slide images." Nature Biomedical Engineering 5.6 (2021): 555-570.

Kather, Jakob Nikolas, et al. "Pan-cancer image-based detection of clinically actionable genetic alterations." Nature Cancer 1.8 (2020): 789-799.

WSI prediction

Attention network   Attention-based pooling Classifier

Attention backbone | Attention branch

FC (N, 256)

(N, 1024) FC (N, 512) MUL (N, 256) FC (N, 1) MUL (N, 512) Σ (1, 512) FC (1, 2)

Attention scores

FC (N, 256)

Patch-level clustering

B positives + B negatives

Clustering loss

Cluster High   Binary Clustering Layer

(B+B, 512)

Cluster Median/Low   Binary Clustering Layer

(No input)

If ground-truth = Cluster High

Clustering-constrained Attention Multiple Instance Learning (CLAM)

https://github.com/mahmoodlab/CLAM

$$\mathbf{W}_1 \in \mathbb{R}^{512 \times 1024} \qquad \mathbf{h}_k = \mathbf{W}_1 \mathbf{z}_k^\top$$
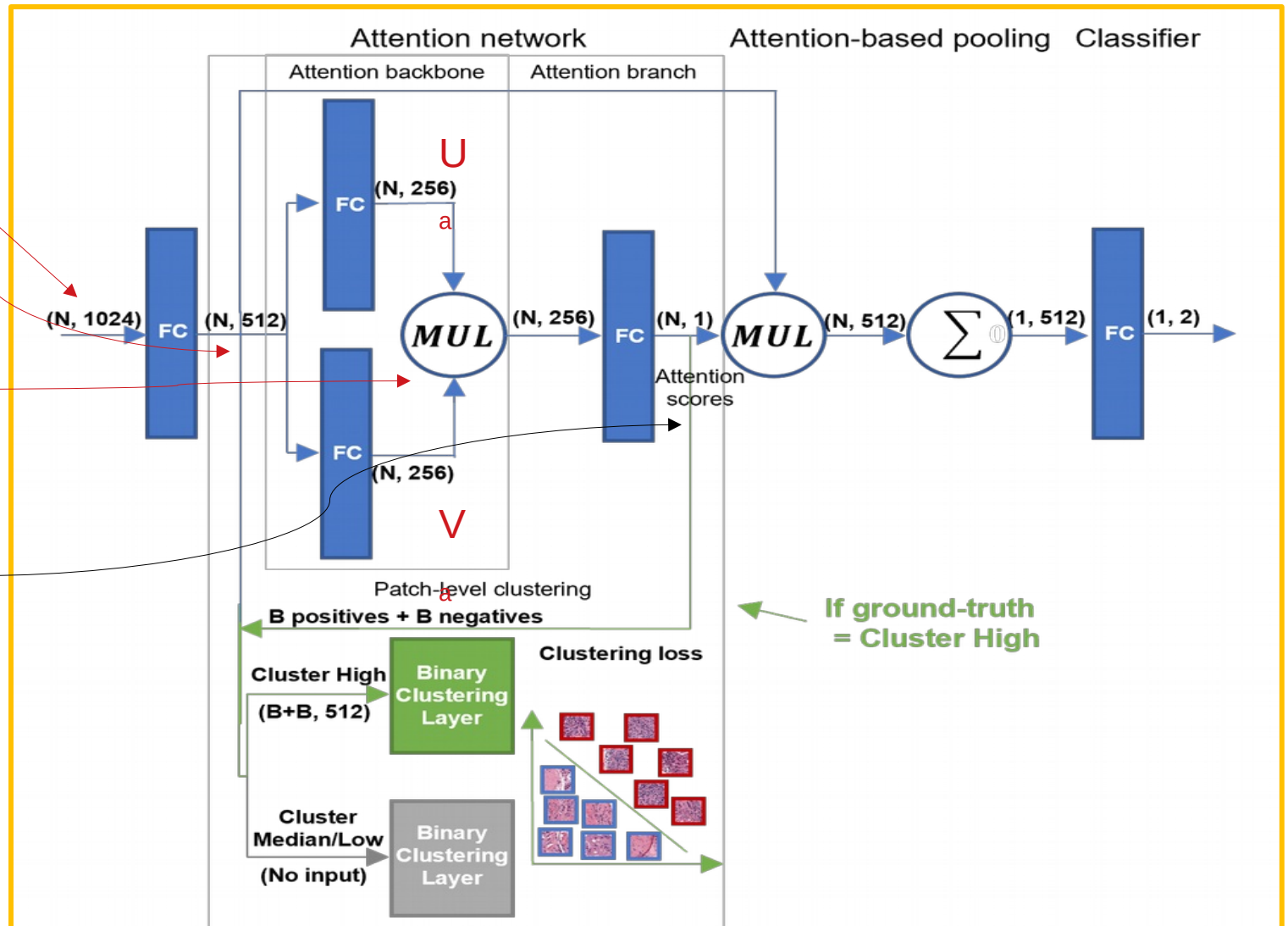
$$\mathbf{U}_a \in \mathbb{R}^{256 \times 512} \text{ and } \mathbf{V_a} \in \mathbb{R}^{256 \times 512}$$

$$a_{k,m} = \frac{\exp\left\{\mathbf{W}_{a,m}\left(\tanh\left(\mathbf{V}_a \mathbf{h}_k^\top\right) \odot \operatorname{sigm}\left(\mathbf{U}_a \mathbf{h}_k^\top\right)\right)\right\}}{\sum_{j=1}^{N} \exp\left\{\mathbf{W}_{a,m}\left(\tanh\left(\mathbf{V}_a \mathbf{h}_j^\top\right) \odot \operatorname{sigm}\left(\mathbf{U}_a \mathbf{h}_j^\top\right)\right)\right\}}$$

$$\mathbf{h}_{slide,m} \in \mathbb{R}^{1 \times 512} \qquad \mathbf{h}_{slide,m} = \sum_{k=1}^{N} a_{k,m} \mathbf{h}_k$$

$$\mathbf{W}_{c,m} \in \mathbb{R}^{1 \times 512} \qquad s_{slide,m} = \mathbf{W}_{c,m} \mathbf{h}_{slide,m}^\top$$

13

Validation data : Results (AUROC)
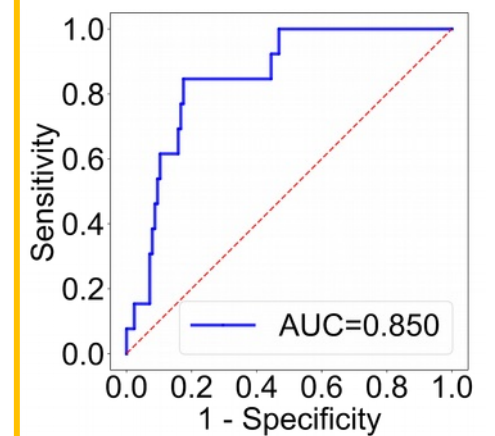
Best fold validation series (n=139 resections)

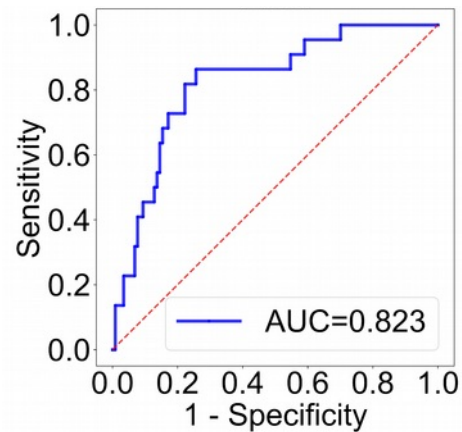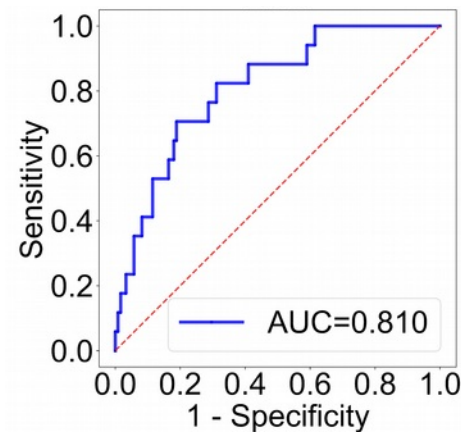### 6-Gene Interferon Gamma
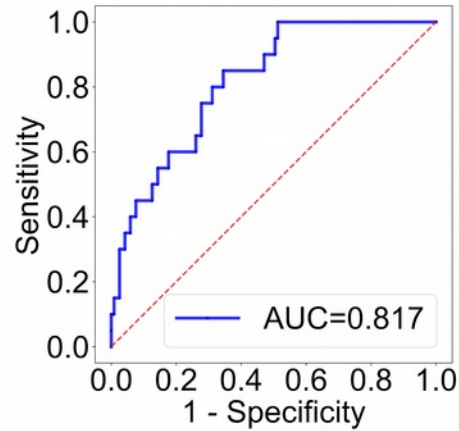
Gajewski 13-Gene Inflammatory
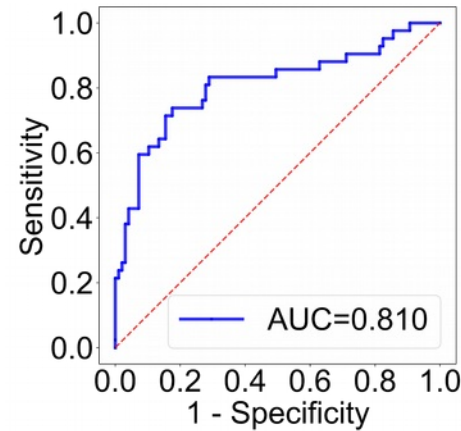
Inflammatory
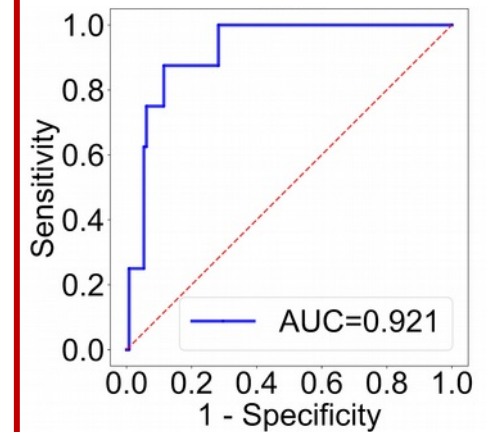

2nd best

Interferon Gamma Biology

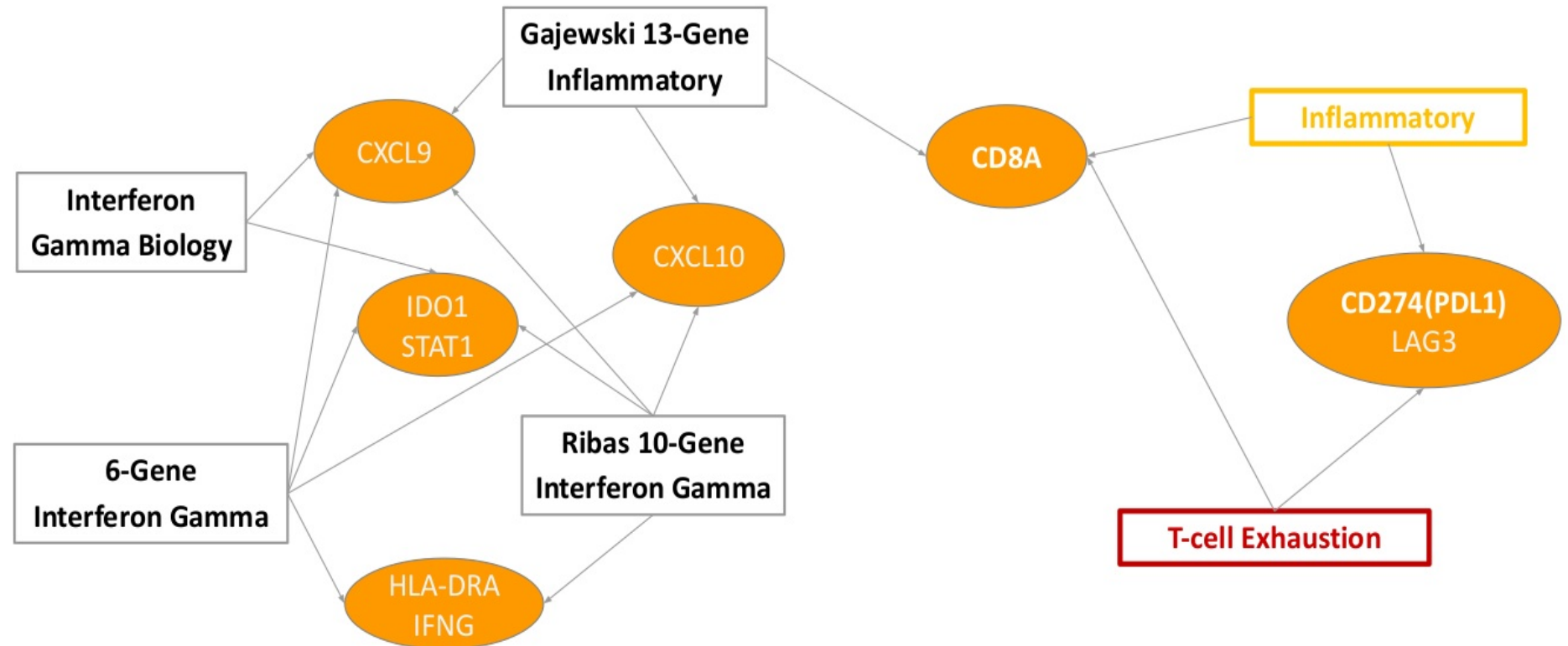### Ribas 10-Gene Interferon Gamma
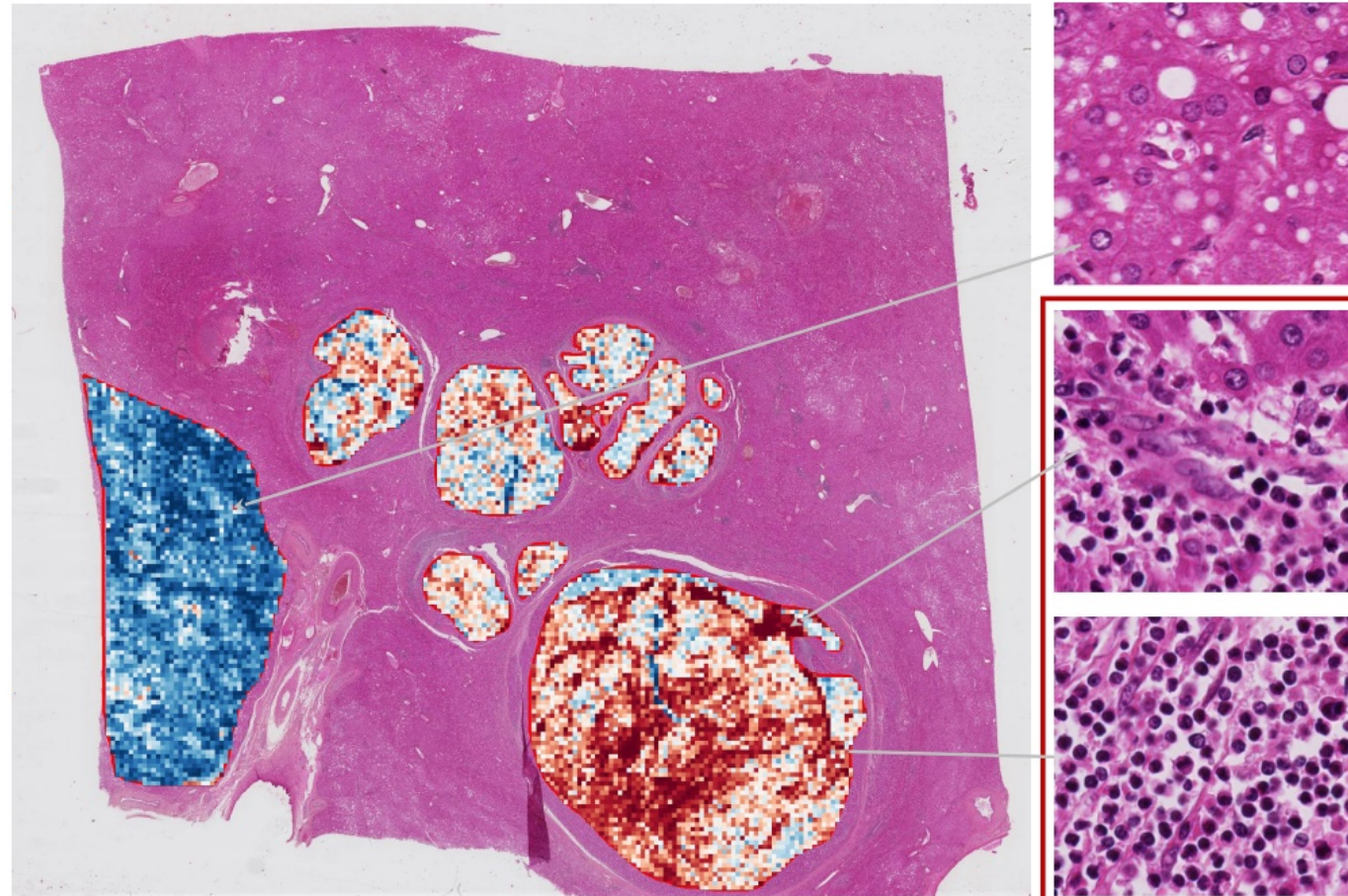

T-cell Exhaustion

Best

Back to the gene signatures

T-Cell Exhaustion : CD274/PD-L1, CD276, CD8A, LAG3, PDCD1LG2, TIGIT
Inflammatory : CD274/PD-L1, CD8A, LAG3, STAT1

Back to the gene signatures

Towards new
AI-based
**image-omics
biomarkers**

In 3 signatures:
- 6G Interferon Gamma
- Interferon Gamma biology
- Ribas 10G Interferon Gamma

Enrichment in **lymphocytes, plasma cells, and neutrophils**.

# Back to the gene signatures (see Poster session Qinghe :-))

Towards new
AI-based
**image-omics
biomarkers**

Lancet Oncology
Accepted 15
September 2023

Artificial intelligence based pathology as a biomarker of sensitivity to atezolizumab-bevacizumab in patients with hepatocellular carcinoma: a multicentric, retrospective study

Qinghe Zeng, Christophe Klein, Stefano Caruso, Pascale Maille, Daniela S Allende, Beatriz Mínguez, Massimo Iavarone, Massih Ningarhari, Andrea Casadei-Gardini, Federica Pedica, Margherita Rimini, Riccardo Perbellini, Camille Boulagnon-Rombi, Alexandra Heurgué, Marco Maggioni, Mohamed Rela, Mukul Vij, Sylvain Baulande, Patricia Legoix, Sonia Lameiras, Ismail Labgaa, Christine Sempoux, Antonia Digklia, Narmin Ghaffari-Laleh, Jakob Nikolas Kather, Omar S M El Nahhas, Pooja Navale, Callie Torres, Tung-Hung Su, Rondell P. Graham, María Teresa Salcedo, María Bermúdez Ramos, Nguyen H Tran, Jean-Michel Pawlotsky, Gontran Verset, Eric Trépo, Maria Varela, Andres Castano Garcia, Guillermo Mendoza-Pacas, Dominique Wendum, Giuliana Amaddeo, Hélène Regnault, Marie Lequoy, Alba Diaz, Maria Reig, Howard Ho Wai Leung, Pompilia Radu, Jean-François Dufour, Stephen Lam Chan, Juan Ignacio Marín-Zuluaga, Purva Gopal, Léa Bruges, Viviane Gnemmi, Jean-Charles Nault, Claudia Campani, Hyungjin Rhee, Young Nyun Park, Mercedes Iñarrairaegui, Guillermo Garcia Porrero, Josepmaria Argemi, Bruno Sangro, Antonio D'Alessio, Bernhard Scheiner, David James Pinato, Matthias Pinter, Valérie Paradis, Aurélie Beaufrère, Simon Peter, Lorenza Rimassa, Luca Di Tommaso, Arndt Vogel, Sophie Michalak, Jérôme Boursier, Nicolas Loménie, Marianne Ziol, Julien Calderaro.

Centre d'Histologie, d'Imagerie et de Cytométrie (CHIC), Centre de Recherche des Cordeliers, INSERM, Sorbonne Université, Université Paris Cité, Paris, France (Q Zeng Msc, C Klein Msc); **Laboratoire d'Informatique Paris Descartes (LIPADE), Université Paris Cité, Paris, France** (Q Zeng Msc, N Loménie PhD); **Université Paris Est Créteil, INSERM, IMRB, F-94010 Créteil, France** (S Caruso PhD, P Maille Msc, Prof J-M Pawlotsky MD, Prof J Calderaro MD); **Assistance Publique-Hôpitaux de Paris, Henri Mondor-Albert Chenevier University Hospital, Department of Pathology, Créteil, France** (S Caruso PhD,

Contextual exploration of WSI
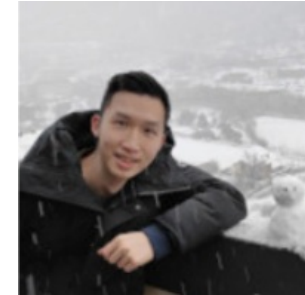
A HIERARCHICAL TRANSFORMER ENCODER TO IMPROVE ENTIRE NEOPLASM SEGMENTATION ON WHOLE SLIDE IMAGES OF HEPATOCELLULAR CARCINOMA

Zhuxian Guo[*†]   Qitong Wang[*]   Henning Müller[†]
Themis Palpanas[*‡]   Nicolas Loménie[*]   Camille Kurtz[*]

[*] Laboratory of Informatics Paris Descartes (LIPADE), Université Paris Cité, Paris, France
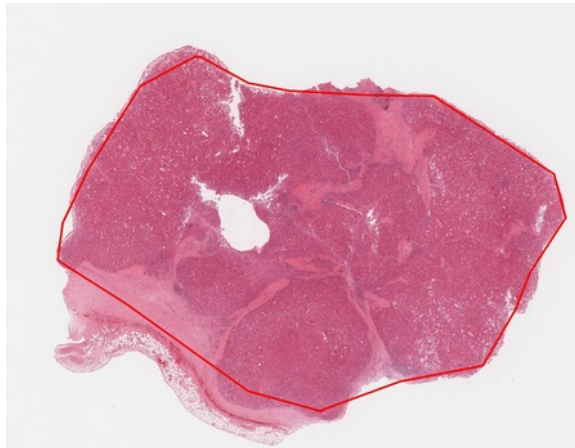[†] University of Applied Sciences of Western Switzerland (HES-SO Valais), Sierre, Switzerland
[‡] French University Institute (IUF), Paris, France

The annotation issue

Discovery series: TCGA (n=336 cases)	External validation series: Hospital Henri Mondor (n=139 cases)
https://portal.gdc.cancer.gov/	

• Gene expression data: RNA sequencing
• WSIs: n=349
  a) Scanner: Aperio scanner
  b) Field of view (20X): ~128x128 µm², ~0.5 µm/pixel
  c) Staining: Hematein-eosin (H&E)
	• Gene expression data: Nanostring (Panel IO360)
	• WSIs: n=139 resections + 7 pre-operative biospies
	  a) Scanner: Hamamatsu
	  b) Field of view (20X): ~115x115 µm², ~0.45 µm/pixel
	  c) Staining: Hematein-eosin-saffron (HES)
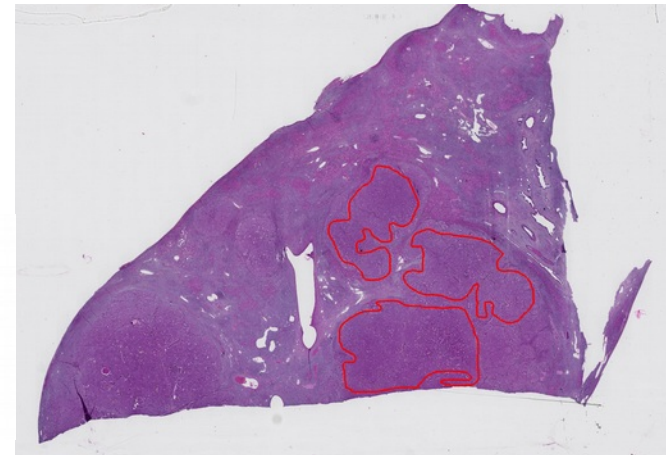
Tumoral areas annotated by Julien

—— Tumor annotation

—— Tumor annotation

HCC neoplasm segmentation: beyond the bag of patches, classification with context

Zoom	Pixel length	Morphological changes
1-10 µm	2-20 px	Aberrant and pleomorphic nuclei...
10-100 µm	20-200 px	Loss of relative cellular orientation...
100 µm	200 px	Aberrant tissue architecture...
1000 µm	2000 px	Pseudo-capsule...

Whole tumor area: yellow area ; 89343036µm²
Viable tumor area: green area : 42617872 µm²
Viable tumor burden : 42617872/89343036*100 = 47.7%
2D size : black line: 16 x 9.7 mm²

Liver Cancer Segmentation Challenge

2 mm
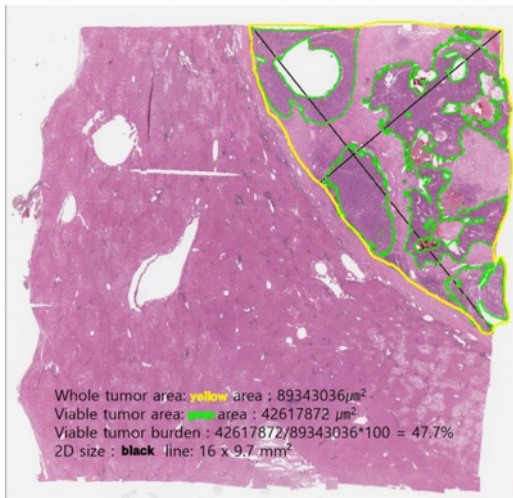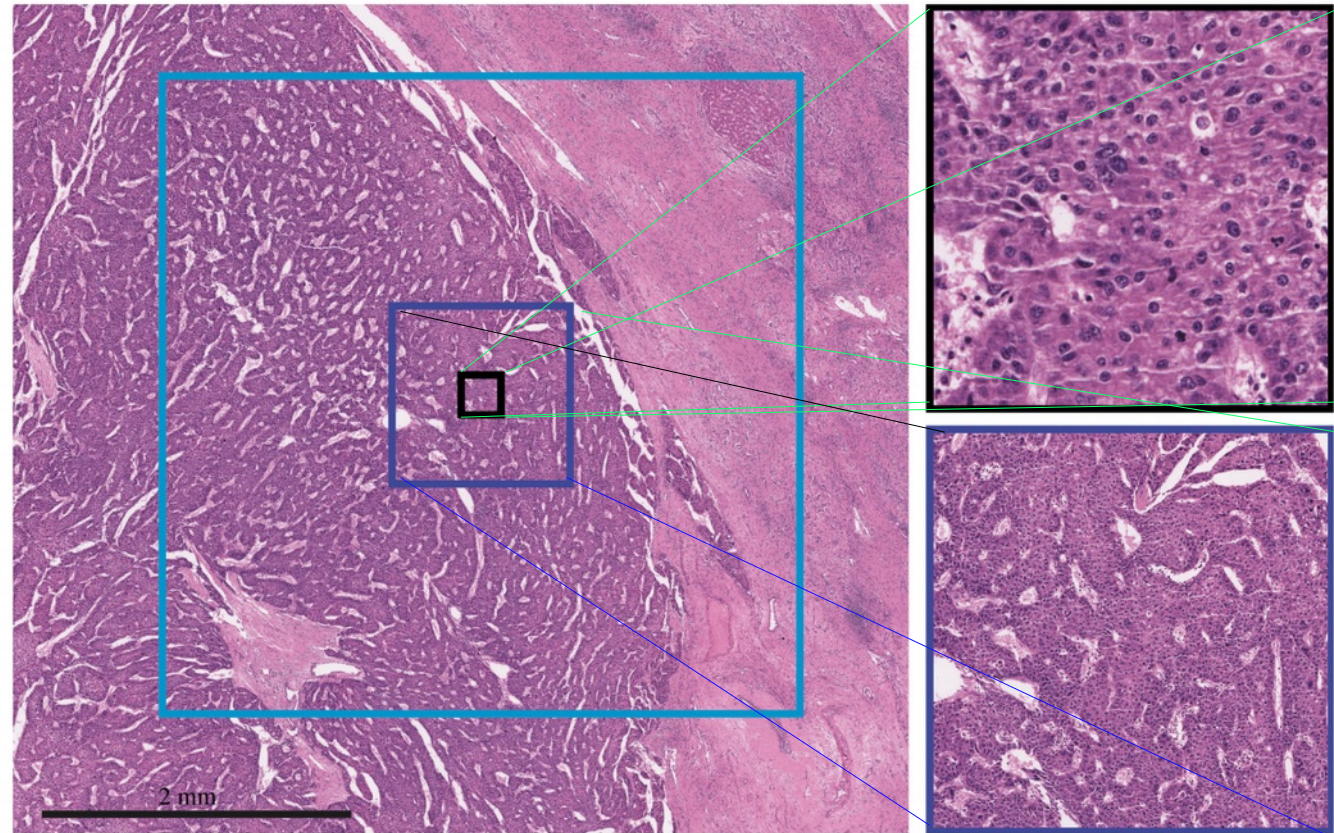
Rüdiger Schmitz et al., "Multi-scale fully convolutional neural networks for histopathology image segmentation: From nuclear aberrations to the global tissue architecture," *Medical Image Anal.*, vol. 70, pp. 101996, 2021.

https://paip2019.grand-challenge.org/Dataset/

Transformer-like architectures

Beyond CLAM (Attention mechanism),
ViT (for **Vision Transformer**) and **self-attention** mechanism

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Jacob Devlin   Ming-Wei Chang   Kenton Lee   Kristina Toutanova
Google AI Language
{jacobdevlin,mingweichang,kentonl,kristout}@google.com

Abstract

We introduce a new language representation model called **BERT**, which stands for **B**idirectional **E**ncoder **R**epresentations from **T**ransformers. Unlike recent language representation models (Peters et al., 2018a; Rad-

There are two existing strategies for applying pre-trained language representations to downstream tasks: *feature-based* and *fine-tuning*. The feature-based approach, such as ELMo (Peters et al., 2018a), uses task-specific architectures that include the pre-trained representations as addi-

"AI and medicine are working together to
v_1 v_2 v_3 .... $v_{6....}$
improve diagnosis and prognosis".
              .... v_{11}

1. Embedding : CNN or else in a vector

Each word like "AI" becomes a vector v_1 and so on.

2. Then the Transformer learn dependencies and context in a matrix

v_1 and v_3 are related to v_6 and so on
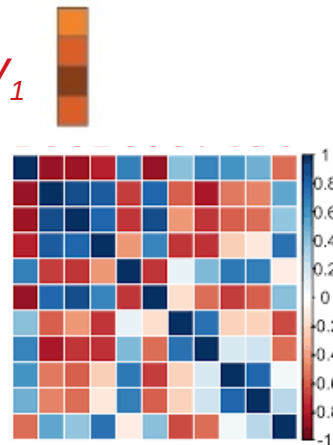
Transformer-like architecture
More and more spatial context

Visual Transformer

CLAM

4096 pixels

256 pixels

Introducing contextual information within expanded 4K patch
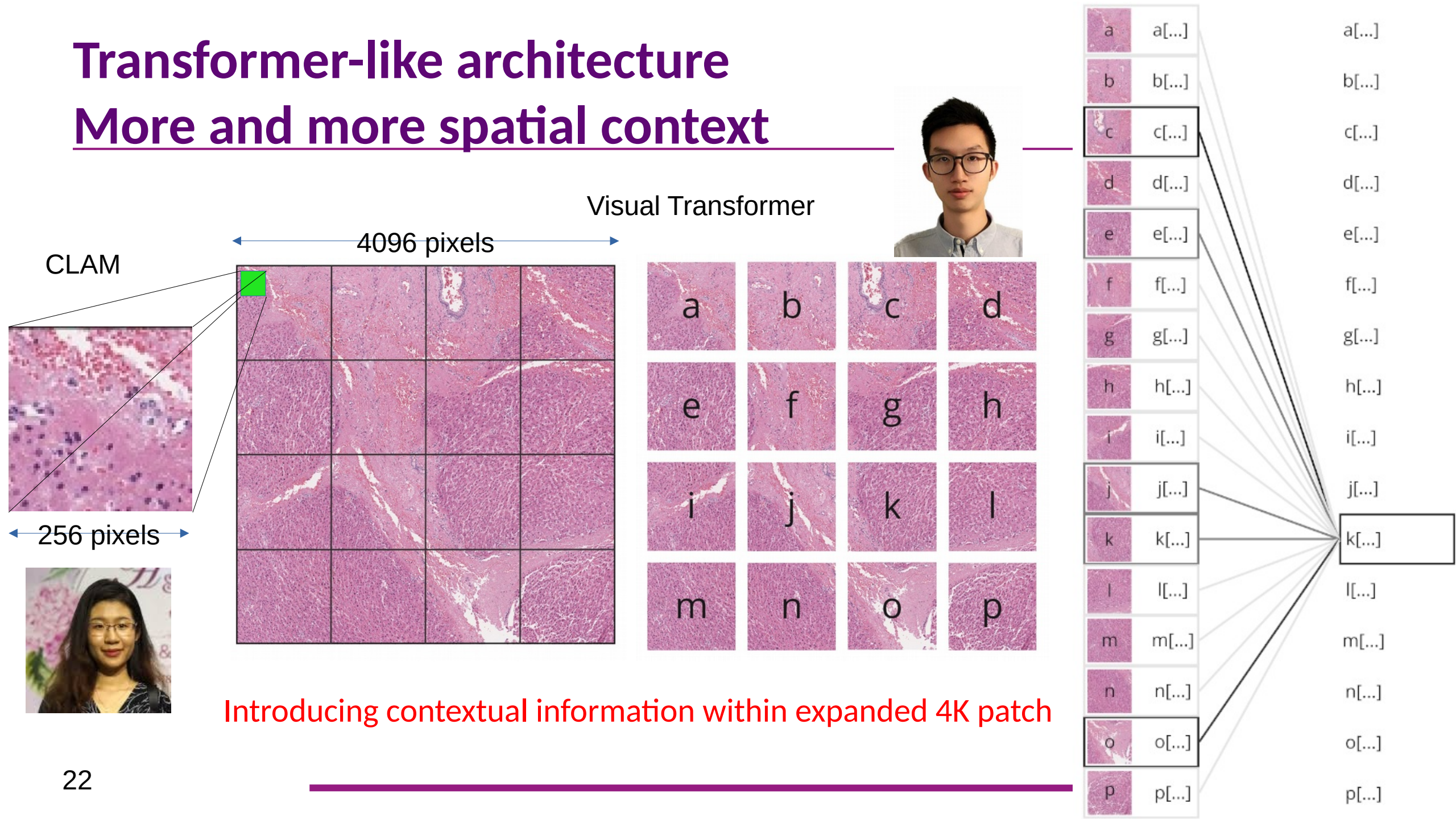
Transformer-like architecture
More and more spatial context

A new Architecture
HiTrans :

Hierarchy-based Transfomer Encoder
for entire HCC neoplasm
segmentation

HiTrans: Hierarchical Transformer, a Context-aware framework

(a) Ground truth

(b) R18_U-Net (4096)

(c) TR-I
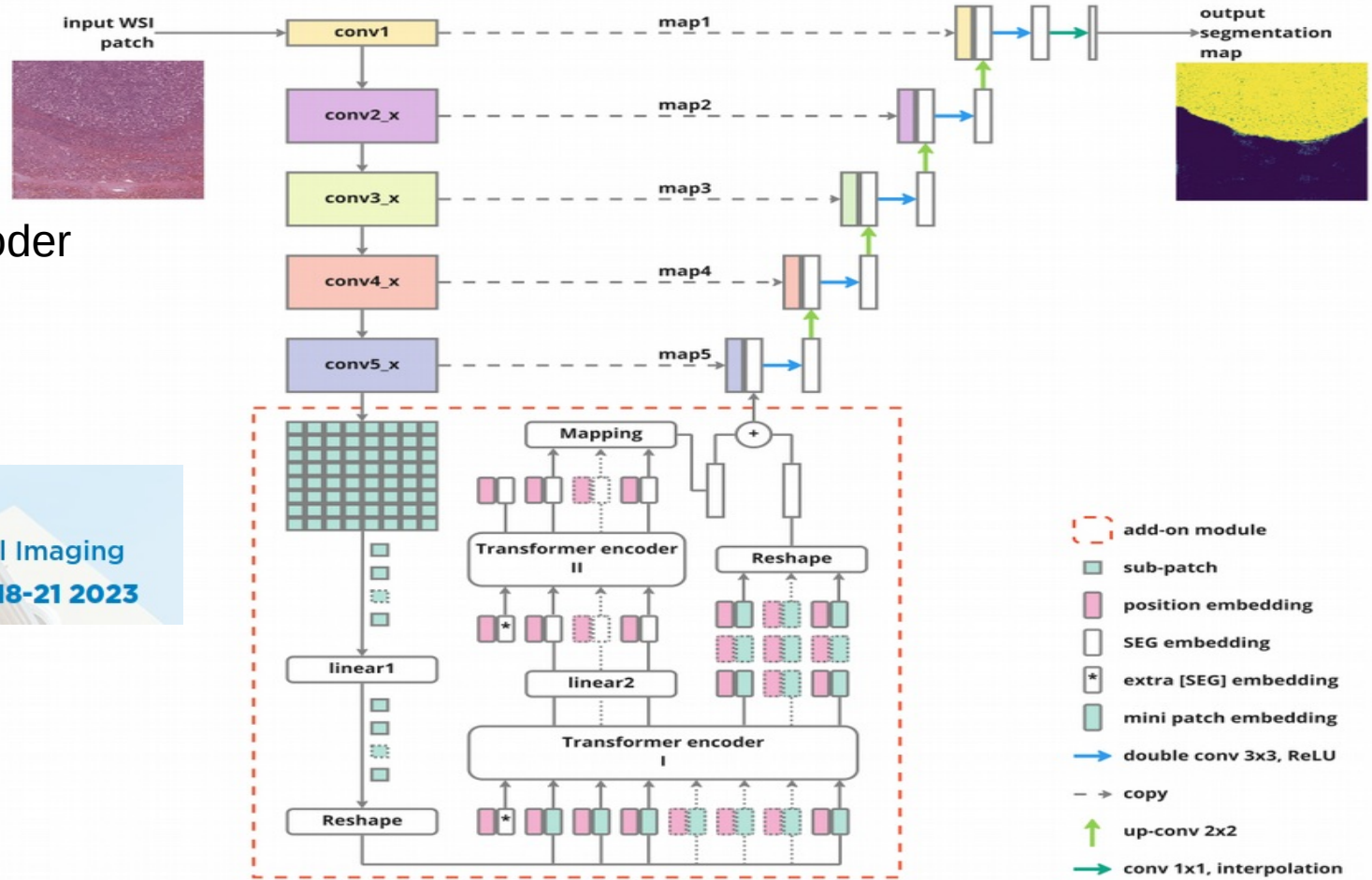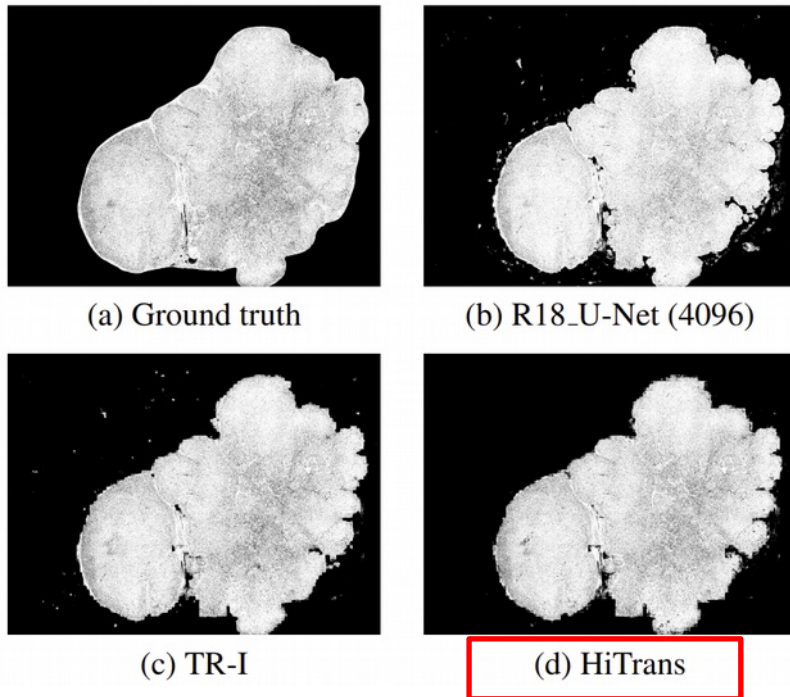
(d) HiTrans

Exp.	Method	Patch size	Avg. Jaccard
1	R18_U-Net	512	0.6609
2	R18_U-Net	4096	0.7202
3	TR-I	4096	0.7172
4	HiTrans	4096	**0.7513**

Deep Net

32 x 32 minipatch

4096 x 4096 patch

e_{Agg} $e_1 ... e_2$ e_3 e_4 e_5 ...

Transformer encoder I

512 x 512 sub-patch

Keep For 2nd Layer

e_{Agg}

$.e_{Agg1}$ e_{Agg2} e_{Agg3} ...

- 1 patch = 64 sub-patches
- 1 sub-patch = 256 minipatches

Transformer encoder II

24

HiTrans: Hierarchical Transformer,
a Context-aware framework

Hierarchical Transformer for Neoplasm Segmentation on WSI of HCC

Can ChatGPT Technologies Help Reading Our Slides?

Zhuxian GUO, PhD student zhuxian.guo@u-paris.fr
Supervisors: Nicolas LOMENIE[1], Camille KURTZ[1], Henning MULLER[2]

[1] Laboratory of Informatics Paris Descartes (LIPADE), Université Paris Cité, Paris, France

[2] University of Applied Sciences of Western Switzerland (HES-SO Valais), Sierre, Switzerland

PhD topic: Exploring the tissue and its micro-environment in whole slide imaging to maximize the impact of digital pathology in clinical practices"

Zhuxian GUO
Université Paris Cité

MOSAIC' DAY / JEUDI 06 AVRIL 2023 / 09H-18H / CAMPUS

https://youtu.be/Ne3ZVIpClsY?t=582

Translational projects : can we skip IHC ？

On going PRT-K project : Artificial intelligence to estimate the prognostic of colorectal cancer (CRC) in adjuvant setting using hematoxylin eosin safran histological slides
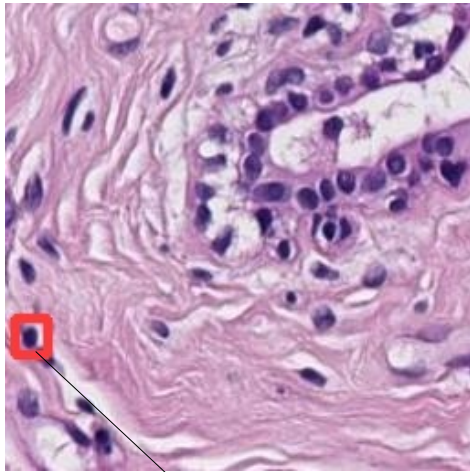
YOLO
Architecture

You Only Look Once
For **object detection**

Fine tuned for lymphocyte detection

Bounding boxes + confidence

S × S grid on input

Class probability map

Final detections

https://arxiv.org/abs/1506.02640

Segmentation and Detection (Final) Leaderboard

#	User (Team)	Algorithm	Created	Mean Position	ts_dice (Position)	lymph_froc (Position)
6th	amine.marzouki (FengHu)	UParisCite1	24 June 2022	6.5	0.6616 (10)	0.5437 (3)
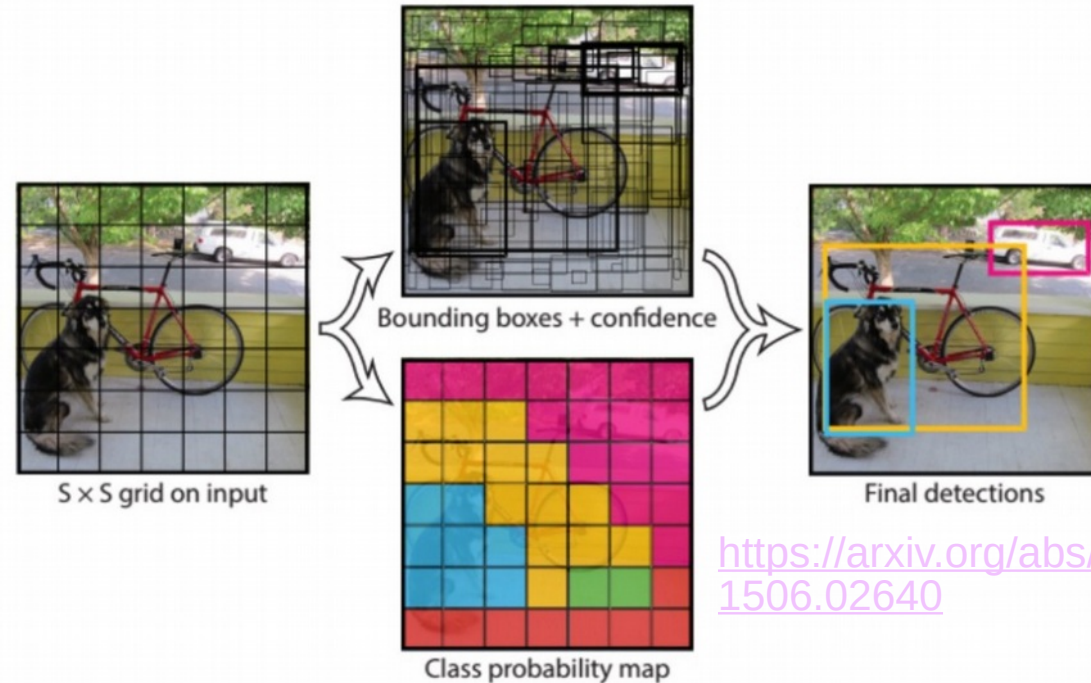
TIGER
Tumor InfiltratinG lymphocytes in breast cancER

Translational projects : a companion test ?

Lymphoid infiltration assessment in CRC

POCHI Project   - Collaboration avec PUPH JF. Emile – Hôpital Ambroise Paré.

**Colorectal Cancer
CRC**

IHC Coloration

Phase 2 clinical trial

Semi-automatic (IVD
Certification)
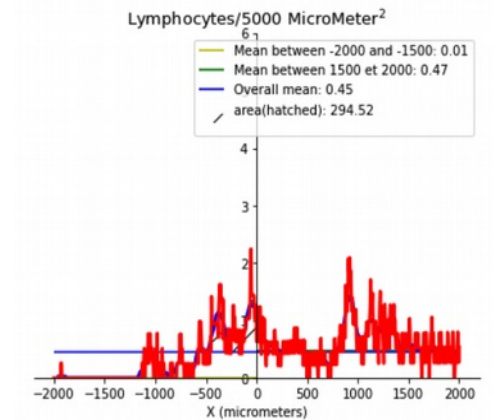
Translational projects : what is AI promise?

(a) Ground truth
(b) R18_U-Net (4096)
(c) TR-I
(d) HiTrans

HiTrans

(a) Ground truth
(b) Predicted lymphocytes / plasma cells
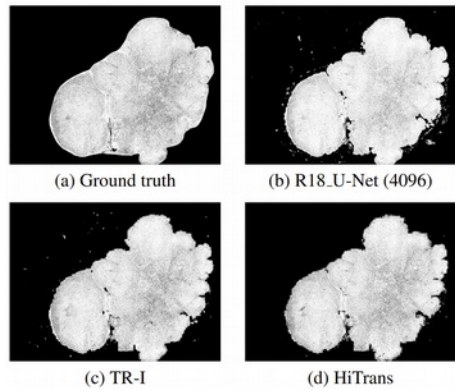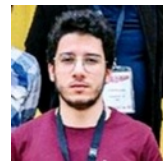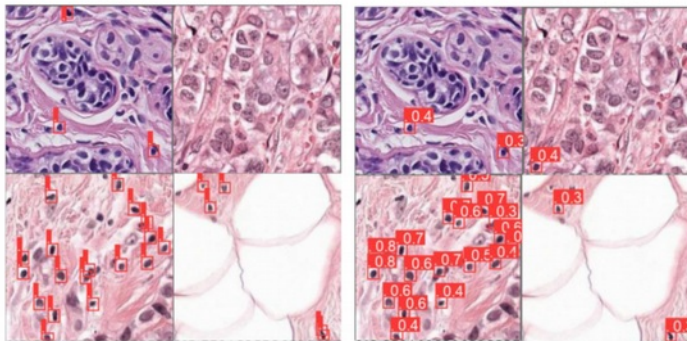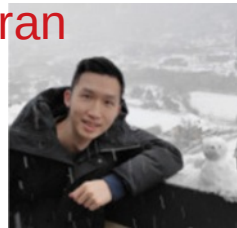
FROC Scores:
2 — 0.5731
1 — 0.6004
3 — 0.5718
Ranking team wise in the L1 leaderboard of TIGER challenge

Université Paris Cité
生物图腾 BIO-TOTEM
WARWICK

H&E Stained WSIs

Neoplasm Segmentation

TILs detection

Liver Cancer
PAIP challenge

Breast Cancer
TIGER challenge

Domain adaptation

HE Stained WSIs

COLON CANCER

POCHI
PRODIGE Cohort

https://w3.mi.parisdescartes.fr/sip-lab/
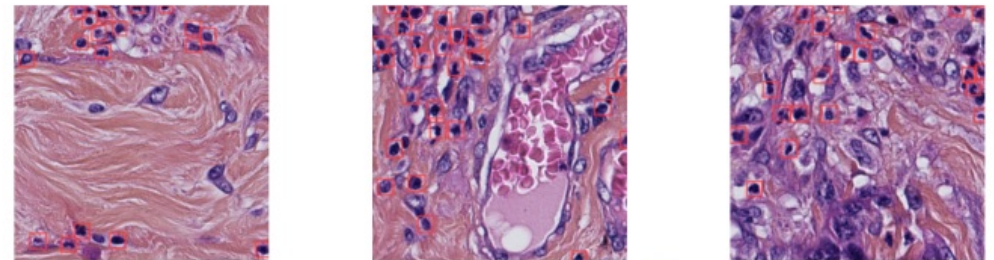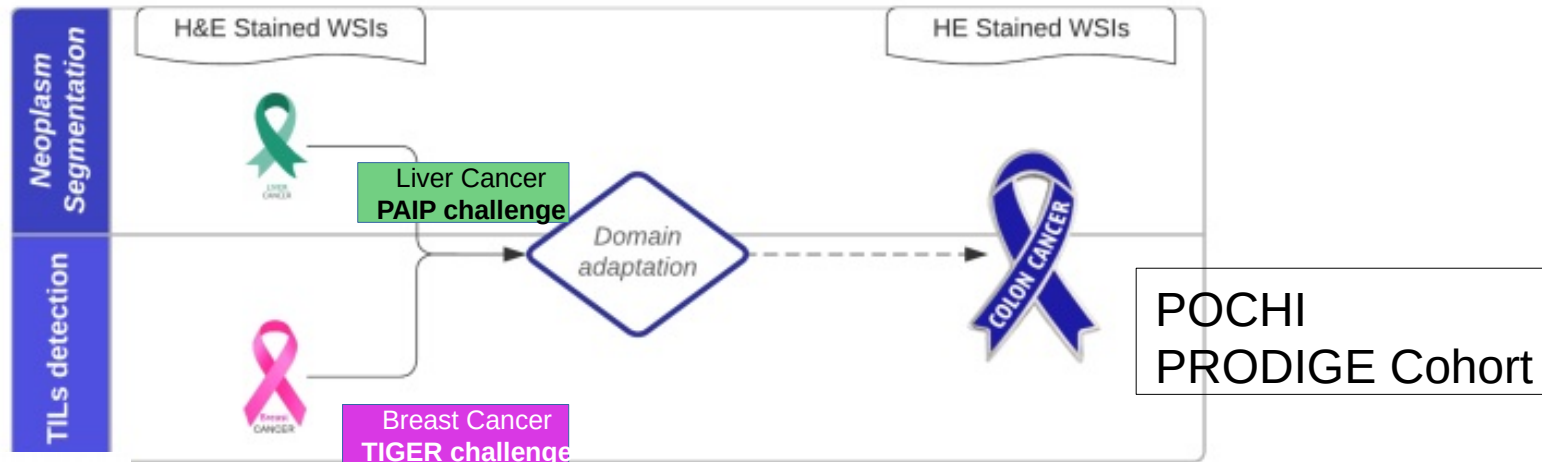
Digital Pathology Group / SIP / LIPADE

Camille Kurtz,   Ass. Prof.,   UPCité
Zhuxian Guo,   PhD student,   UPCité
Qinghe Zeng,   PhD student,   UPCité
Amine Marzouki,   IR,   UPCité
Christophe Klein,   IR,   CRC
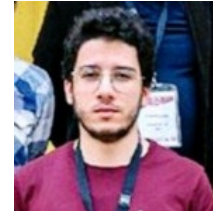
Jean-François Emile,PUPH, Ambroise-Paré
Julien Calderaro,   PUPH, Créteil
Philippe Bertheau,   PUPH, Saint-Louis
François Ghiringhelli,PUPH, Dijon

NICOLAS.LOMENIE@U-PARIS.FR

MERCI

data intelligence institute of Paris

Next...

Cluster1
Cluster10
Cluster0
Cluster2
Cluster3
Cluster4
Cluster6
Cluster9
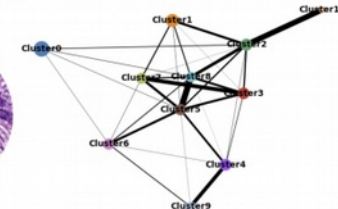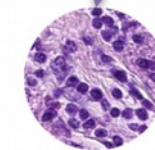
Christophe KLEIN
(Research Engineer)

Camille KURTZ
(Associate Professor)
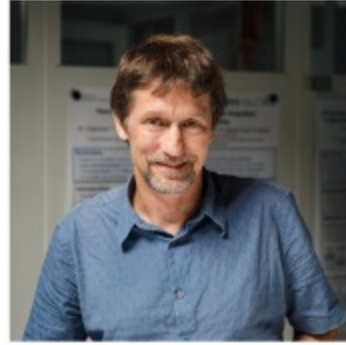
Henning MULLER
(Professor)

Julien CALDARERO
(PU-PH)

Jean-François EMILE
(PU-PH)

Themis PALPANAS
(Professor)

Amine MARZOUKI
(Research Engineer)

Qinghe ZENG
(PhD Student)
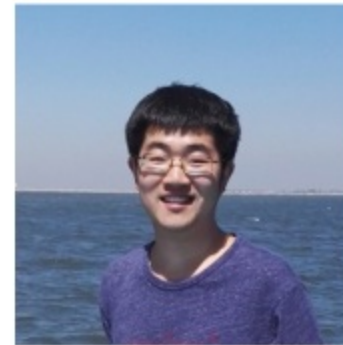
Qitong WANG
(PhD Student)

Zhuxian GUO
(PhD Student)

More and more spatial context (architectures)
CLAM → ViT → Graph neural networks

Medical Image Analysis
Volume 75, January 2022, 102264

Hierarchical graph representations in digital pathology
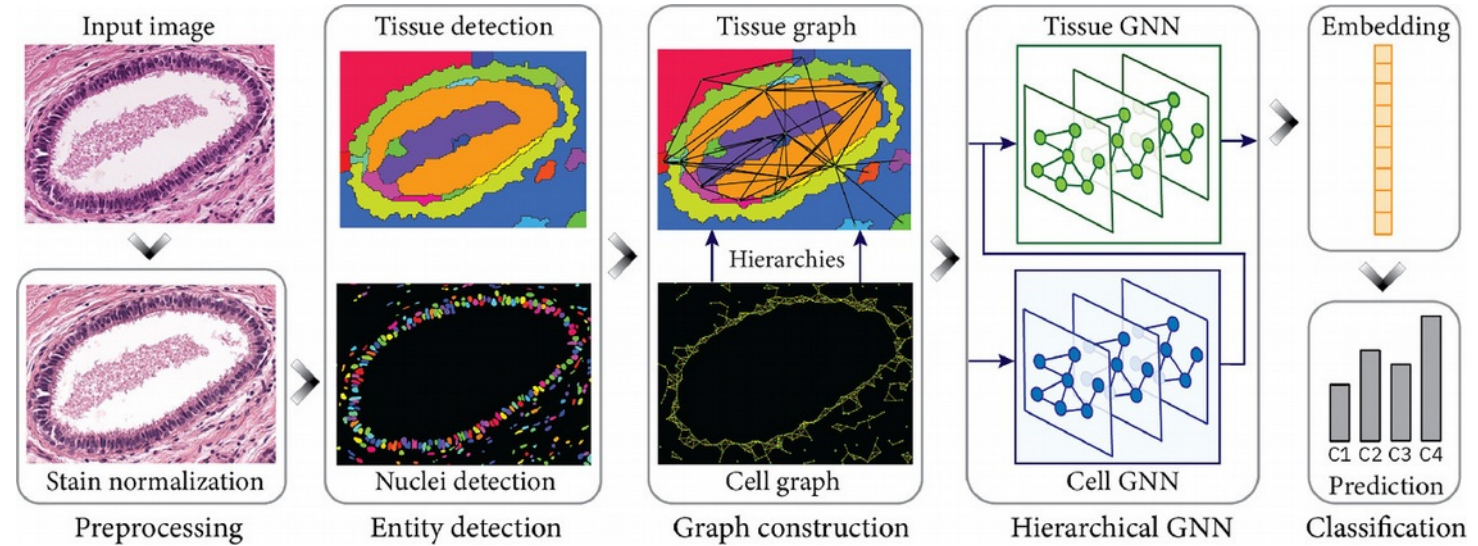
ARTICLES
https://doi.org/10.1038/s41551-022-00923-0

Derivation of prognostic contextual histopathological features from whole-slide images of tumours via graph deep learning

Yongju Lee[1,13], Jeong Hwan Park[2,3,13], Sohee Oh[4,13], Kyoungseob Shin[1,13], Jiyu Sun[4], Minsun Jung[2,5], Cheol Lee[2,6], Hyojin Kim[2,7], Jin-Haeng Chung[2,7], Kyung Chul Moon[2,6] and Sunghoon Kwon[1,8,9,10,11,12]
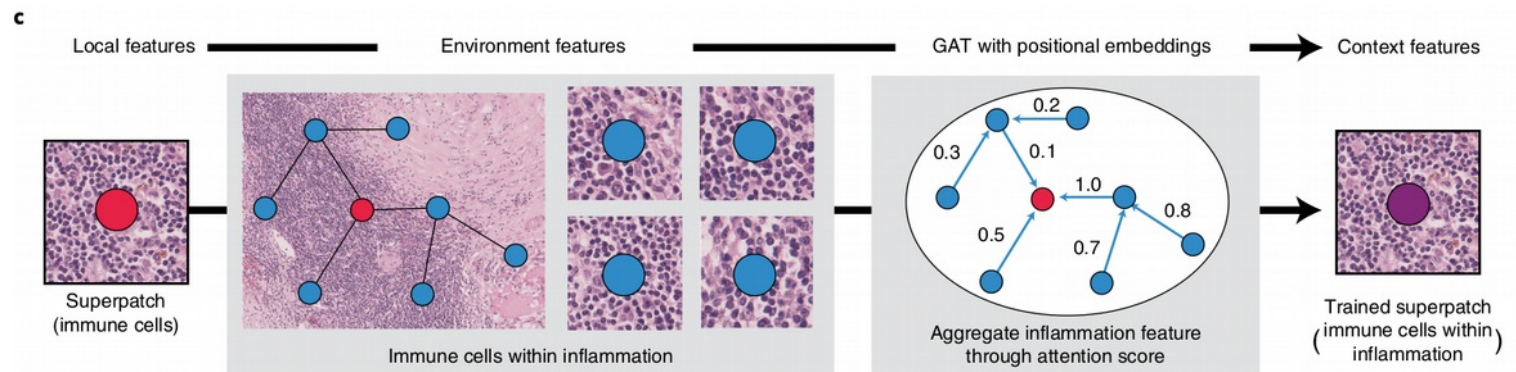
More and more spatial context (data)
CLAM → ViT → Graph neural networks
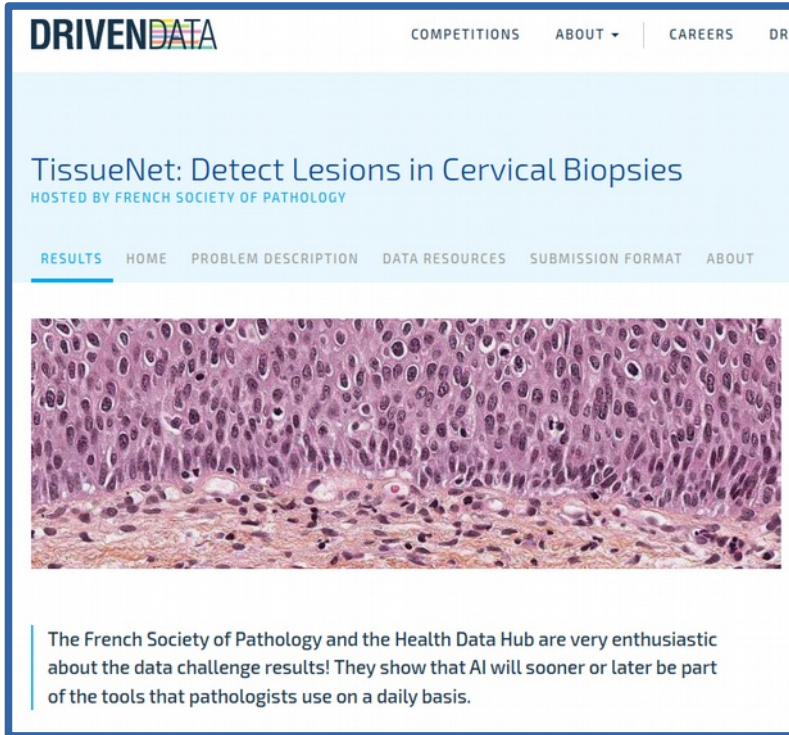
From WSIs to Spatial Transciptomics (ST) :

Through spatial interaction graphs between Tissue Modules (TM) discovered by machine learning
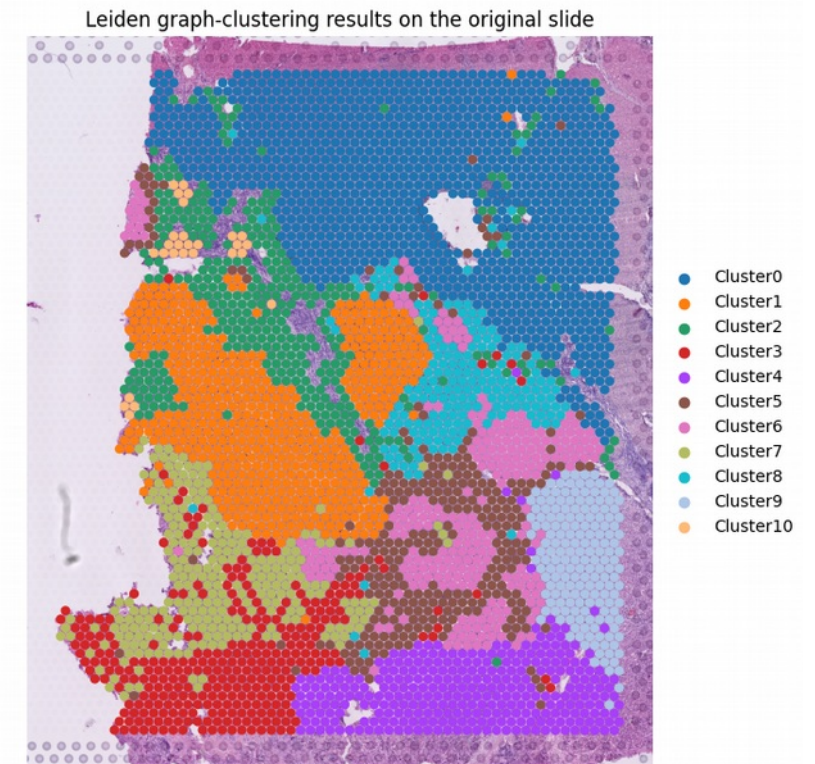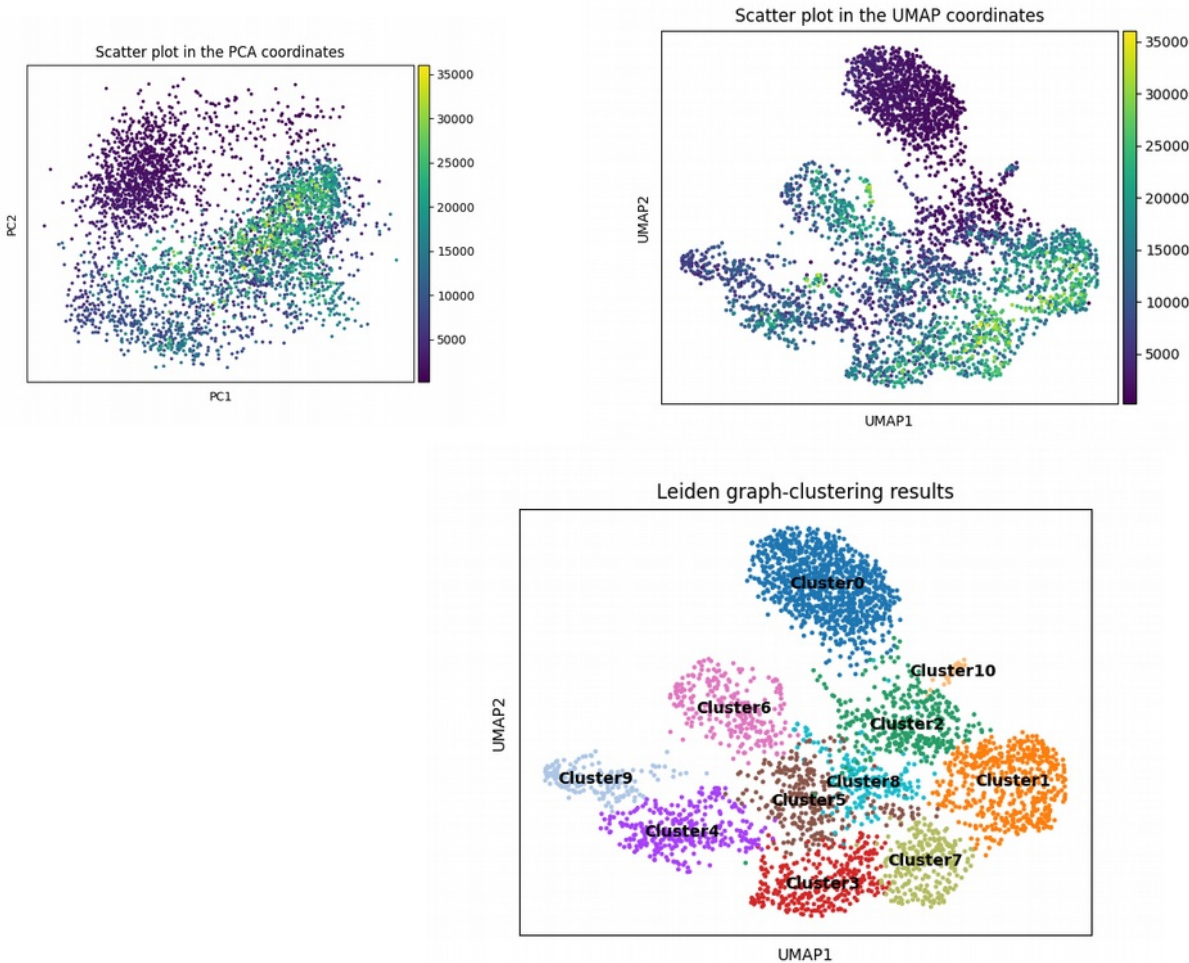
More and more challenges to learn, create a community, benchmark

https://www.health-data-hub.fr/data-challenge-visiomel

https://youtu.be/wwCS_-rjcxU

Liver Cancer Segmentation Challenge PAIP