Towards a General-Purpose Foundation Model for Computational Pathology

Mohammed Wassim **CHIKHI**   Hadia **AMJAD**

Université Paris Cité — M2 VMI

Supervisors: **Pr. Nicole Vincent**, **Pr. Nicolas Loménie**, **Pr. Camille Kurtz**

Nov 3, 2025

Université Paris Cité

nature medicine

Outline

Université Paris Cité

nature medicine

Outline

Problem Statement

Clinical/computational reality

- **Gigapixel WSI** $\Rightarrow$ patching, weak supervision, **domain shift**.
- Pathologists cover **thousands** of entities; one model per task is impractical.

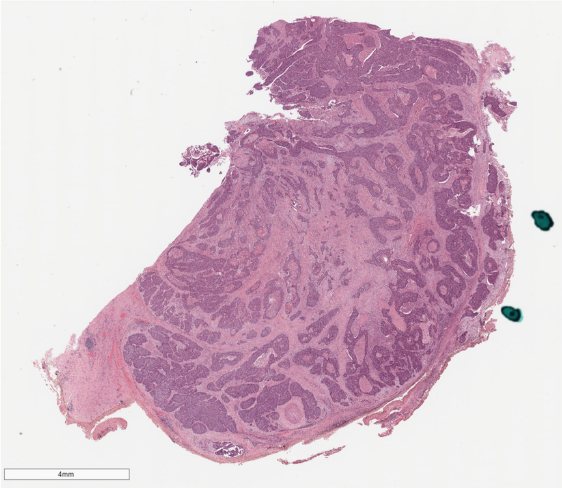
Limitations of prior practice

- **High annotation cost**; heterogeneous labels.
- **ImageNet transfer** mismatched to histology morphology.
- **TCGA** limited ($\sim$29k slides; mostly primary cancers).

Need

- A **general-purpose model** trained on **massive, diverse** pathology data.

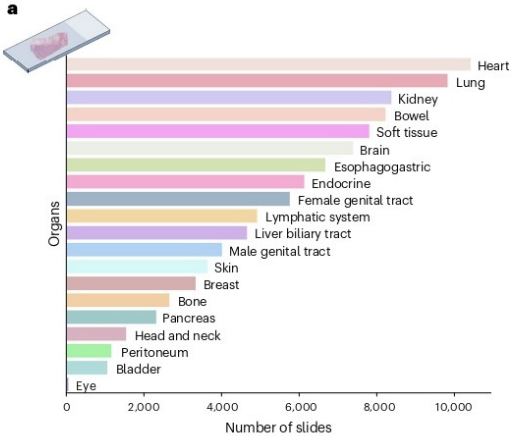
Example: Whole Slide Image (WSI)



Left: full slide (gigapixel). Right: zoom on tissue patches.

Each WSI can exceed 10 GB and contain millions of cells. Only small **patches** are used as inputs for training and inference.

Diversity matters (Fig. 1a)



Nature Medicine (2024) — Fig. 1a.

Outline

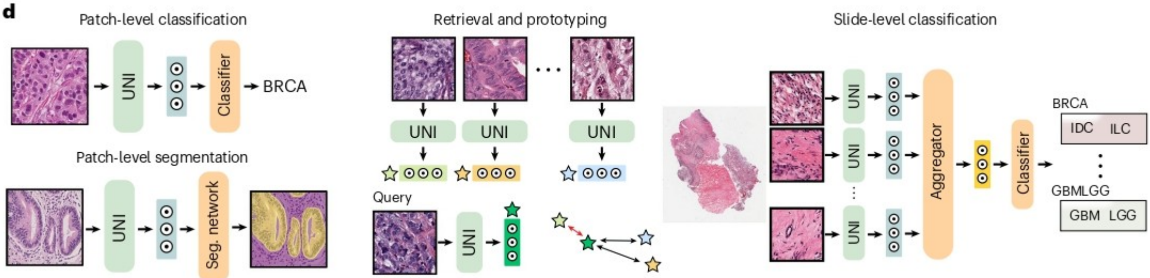
Université Paris Cité

nature medicine

What is UNI?

- **Backbone**: **ViT-L** (300M params, order of magnitude).
- **Pretraining scale**: $\gtrsim$100k WSI, 17 organs, 43 cancer families, 108 OncoTree codes.
- **Goal**: transferable **morphological invariants** for ROI/WSI tasks, retrieval, segmentation.
- **Baselines**: ResNet-IN, CTransPath, REMEDIS, etc.

UNI applications (Fig. 1d)



Nature Medicine (2024) — Fig. 1d.

Outline

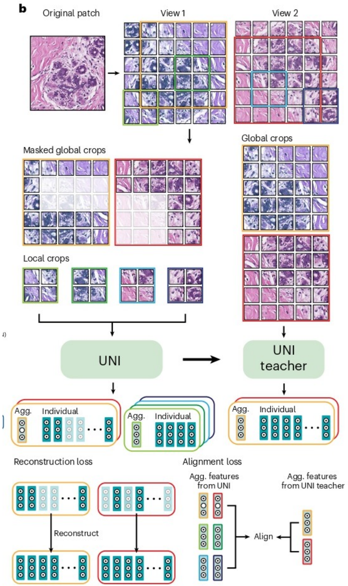
DINOv2 self-supervision (Fig. 1b)

Two objectives

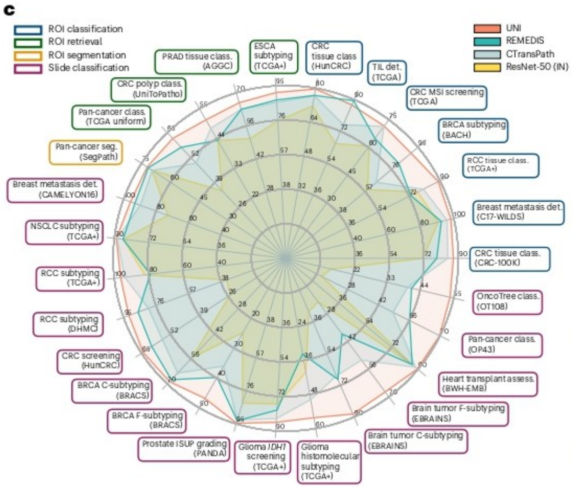
- **Self-distillation**: student matches **EMA teacher** across global/local crops.
- **Masked image modeling**: predict masked patches from context.

Design choices

- **L2-normalized head** ⇒ stable embedding space.
- **Ablation**: DINOv2 > MoCoV3 ⇒ algorithm **matters**.



Downstream task families (Fig. 1c)



Nature Medicine (2024) — Fig. 1c.

Outline

Motivation

UNI Overview

Method (Pretraining)

Benchmarks & Scale

Results

Critical analysis & Perspectives

Université Paris Cité

nature medicine

OT-43 / OT-108 (Fig. 2a)

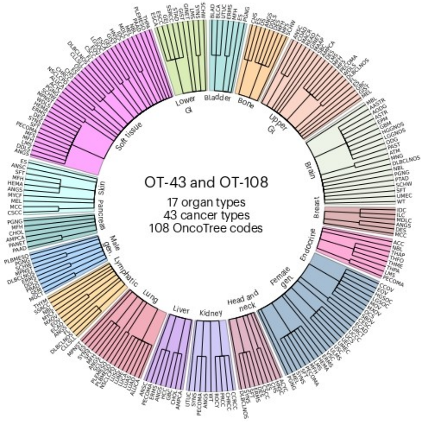
Why these sets?

- **Generalization** to **rare cancers** and cross-site diversity.
- Paper stats: ∼5.5k WSI; min ∼20 per OncoTree; 17 organs.

Message

- Hierarchy (organ → family → OncoTree) stresses **representation breadth**.

a



Outline

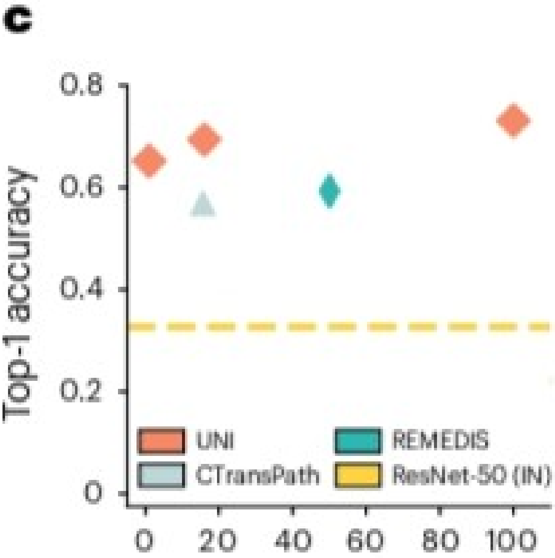
Scaling law (Fig. 2c & 2e)

Key numbers

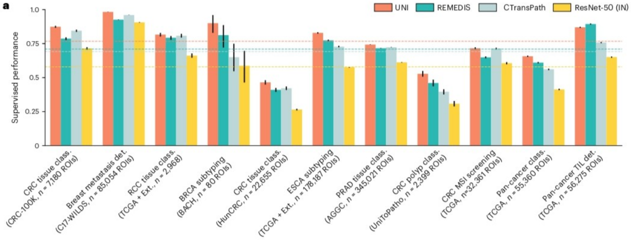
- Mass-1K → Mass-22K: **+4.2%**
- Mass-22K → Mass-100K: **+3.7%**
- **No plateau** observed.

Insight

- **DINOv2 > MoCoV3** even small-scale ⇒ **algorithmic choice** is crucial.



ROI-level performance (Fig. 3a)



Readable takeaways

- Focus: **Pan-cancer (32)**, **TIL**, **PRAD**, **CRC**.
- Average gain (paper): UNI > CTransPath by **+7.6%**.

Retrieval (Ext. Fig. 3)

Definition

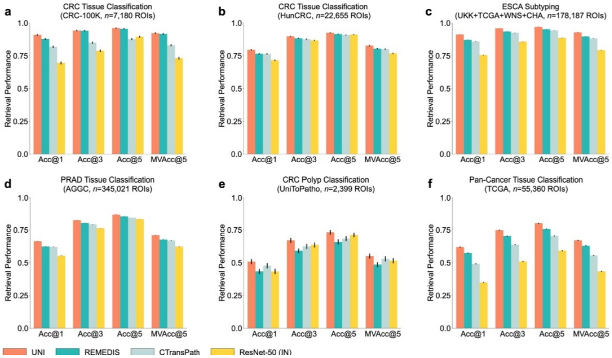
- Retrieve **morphologically similar** regions from a query patch.

Metrics

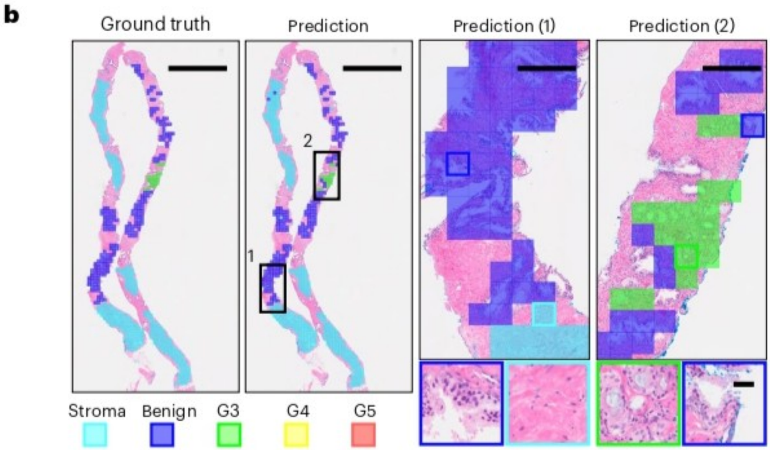
- Acc@K, **MVAcc@5** (majority vote among top-K).

Result

- Pan-cancer (32): **+4.6% Acc@1** vs REMEDIS (paper).



Qualitative ROI examples (Fig. 3b)



Nature Medicine (2024) — Fig. 3b.

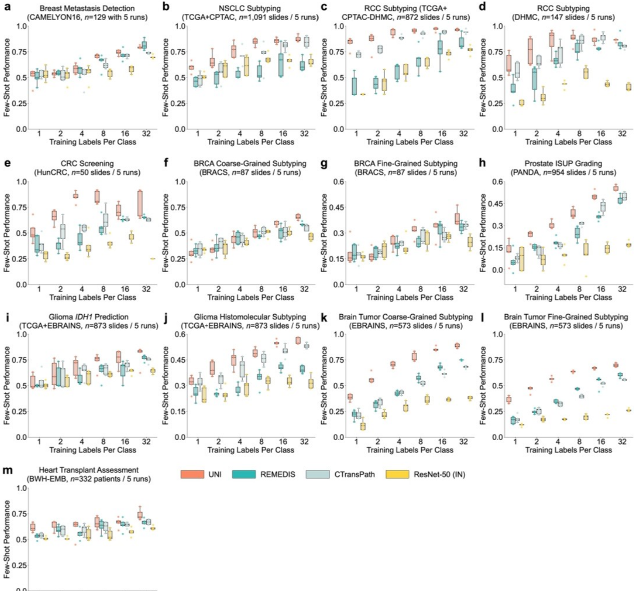
Few-shot slide classification (Ext. Fig. 1)

Label efficiency

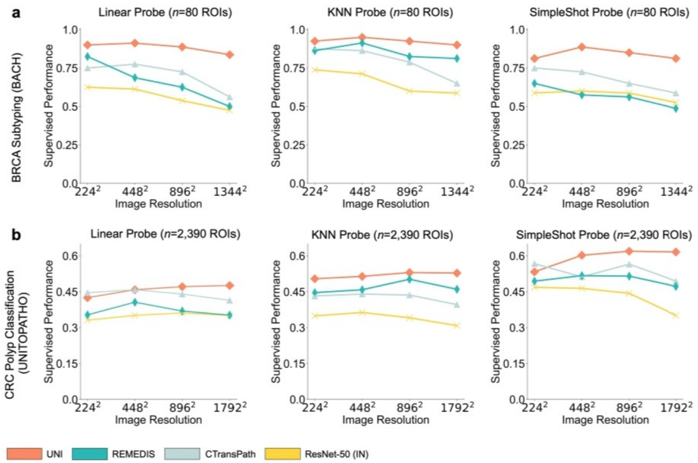
- 4 slides (UNI) ≈ 32 slides for others.
- ISUP grading: about 2× more label-efficient.

Examples

- Brain tumors (30 rare classes), RCC subtyping.

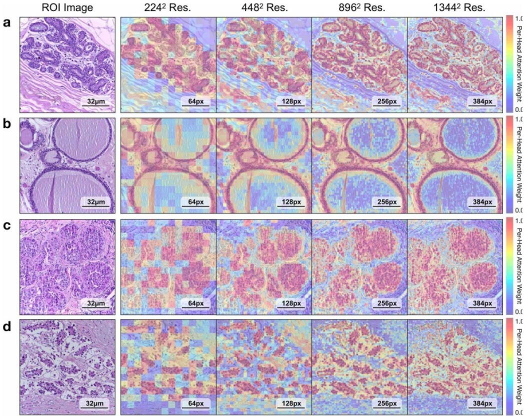


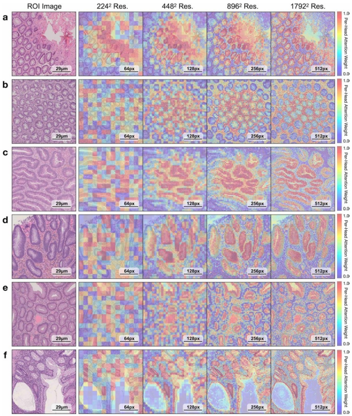
Resolution robustness (Ext. Fig. 4)



Nature Medicine (2024) — Ext. Fig. 4.

Interpretability — Attention heatmaps (Ext. Fig. 5–6)





Attention is plausible, not causal.

Outline

Motivation

UNI Overview

Method (Pretraining)

Benchmarks & Scale

Results

Critical analysis & Perspectives

Université Paris Cité

nature medicine

Limitations (critical view)

Architecture

ViT-L shows modest gains on dense segmentation (e.g., +0.005 dice reported); **CNNs** remain strong.

Modality

Vision-only (no reports/genomics) — multimodal directions exist.

Granularity

Pretraining at **ROI/patch** level $\Rightarrow$ WSI needs **MIL**; no direct WSI-level SSL yet.

Coverage

Mostly H&E FFPE; limited cytology / hematopathology.

Perspectives (concrete avenues)

- **Bigger backbones** (ViT-Giant), compute permitting.
- **Multimodal** (vision+language, vision+genomics).
- **WSI-level SSL** to reduce reliance on MIL.
- **Specialized domains**: cytology, hematology, IHC.
- Hybrids ViT/CNN or adapters for **dense segmentation**.

Key takeaways — contributions & limits

Contributions

- **State-of-the-art** across **34 tasks**; rare cancers up to **+19.6%**.
- **Few-shot**: **4** slides ≈ **32** (others); ∼**2×** label efficiency (ISUP).
- **Scaling**: +4.2% (1K→22K), +3.7% (22K→100K); **no plateau**.

Limitations

- ViT-L weaker on dense segmentation; unimodal; ROI-level pretraining; domain coverage gaps.

Université Paris Cité

nature medicine

References I

- Mahmood F. *et al.* Towards a general-purpose foundation model for computational pathology. *Nature Medicine*, 2024.
- Ilse M., Tomczak J., Welling M. Attention-based Deep Multiple Instance Learning. *ICML*, 2018.
- Chen R. *et al.* Scaling Vision Transformers for Whole Slide Image Analysis. *CVPR*, 2022.
- Lu M.Y. *et al.* Data-efficient and weakly supervised computational pathology on WSI. *Nat. Biomed. Eng.*, 2021.
- Wang H. *et al.* CTransPath: Transformers for computational pathology. *Med. Image Anal.*, 2023.
- Bioptimus Team. H-Optimus: A multimodal foundation model for biology. *arXiv*, 2025.

Université Paris Cité

nature medicine