

The background features a composite image. On the left and right, there are histological sections of tissue stained with hematoxylin and eosin (H&E), showing various cellular structures. A large magnifying glass is positioned over the center, focusing on a digital overlay. This overlay consists of a light blue grid with glowing cyan nodes and lines, and several semi-transparent, pinkish-purple circular shapes that resemble cells or data points.

From Promise to Practice: Preparing Histopathology for the Foundation Model Era

Heather Couture
Pixel Scientia Labs

Pathology Informatics: HIMA Imaging Science

May 18, 2026

Disclosure

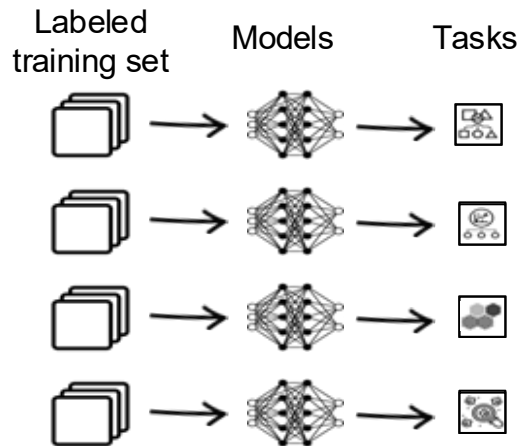
In the past 12 months, I have had a significant financial interest with the providers of the following services:

Owner and Principal Scientist of Pixel Scientia Labs, which provides advisory and fractional CV/ML services to science-driven teams.

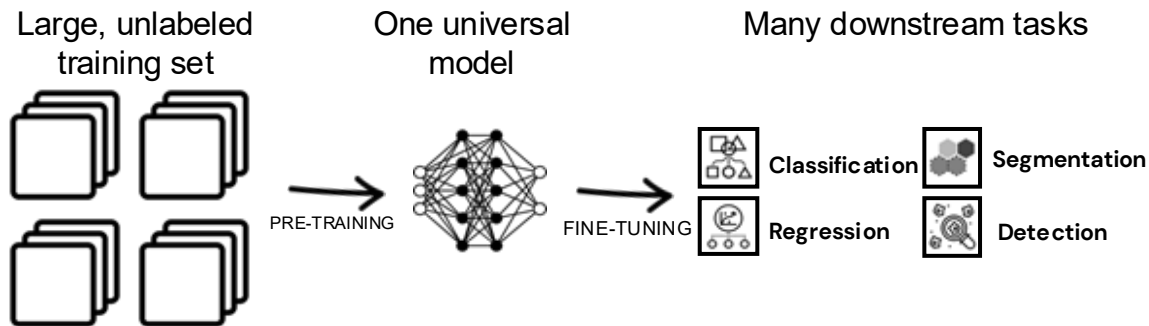
All relevant financial relationships have been mitigated. No product, service, or commercial entity will be promoted during this presentation.

The Paradigm Shift

Traditional ML



Foundation Models



For pathology:

- Qualitative image analysis
- Diagnostics
- Drug discovery
- Clinical trial planning

Goals of This Talk

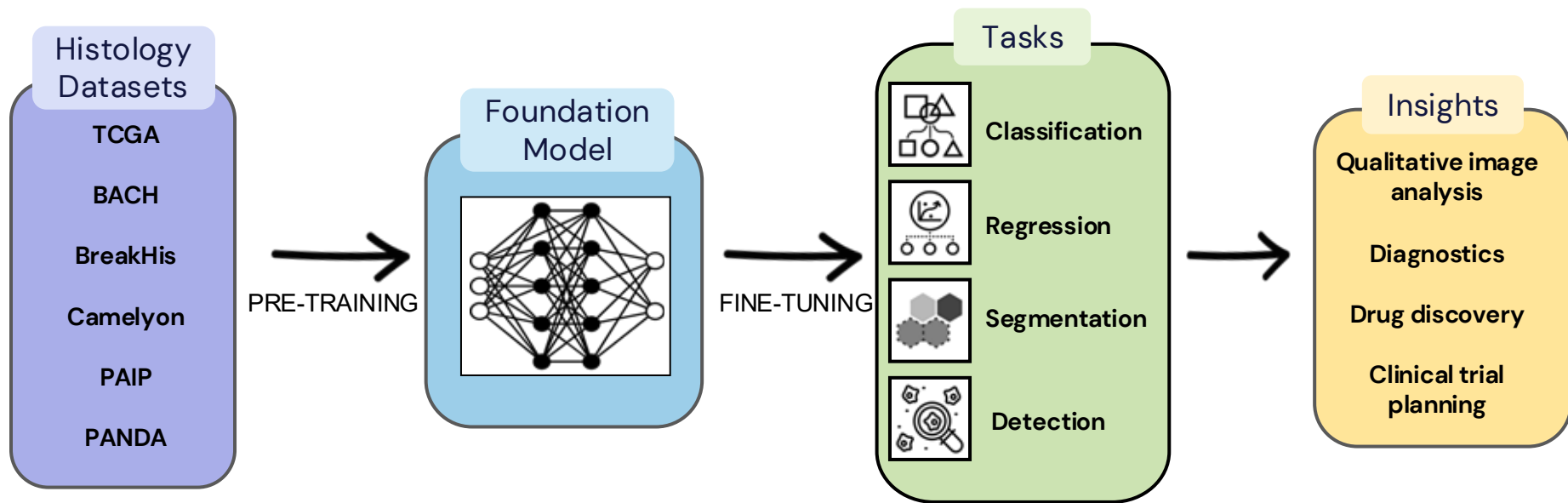
- What is a pathology foundation model?
- Core benefits
- Types of foundation models
- Three pillars: trustworthy, reliable, actionable
- Closing thoughts

“Foundation model” coined by Stanford Institute for Human-Centered Artificial Intelligence's Center for Research on Foundation Models

“A foundation model is any model that is trained on **broad data** (generally using **self-supervision** at scale) that can be **adapted** (e.g., fine-tuned) to a wide range of downstream tasks.”

Source: Bommasani, On the Opportunities and Risks of Foundation Models, 2021

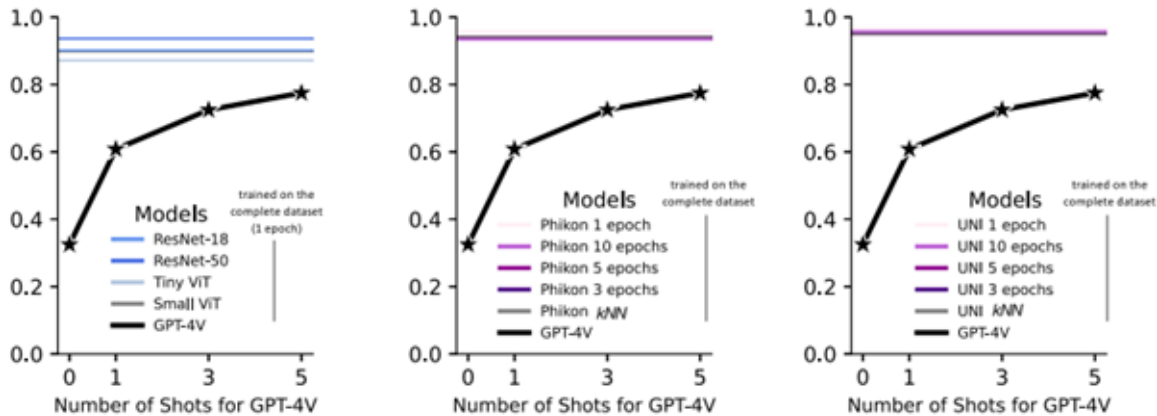
Foundation Models for Histology: Adaptable & Generalizable



First publicly available model ~2022
“Self-supervised learning”

Domain-Specificity is Critical

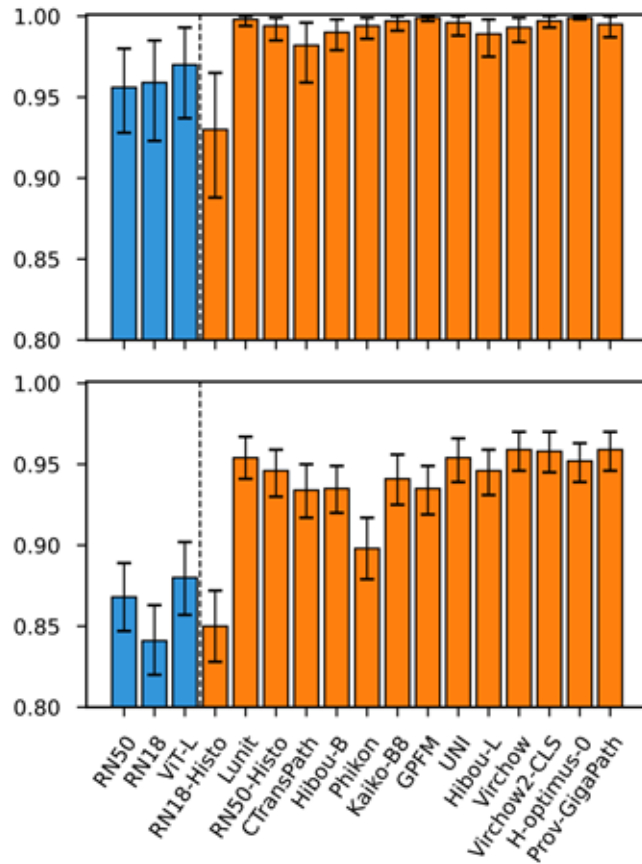
Colorectal Tissue Classification Accuracy



Source: Ferber, In-context learning enables multimodal large language models to classify cancer pathology images, 2024

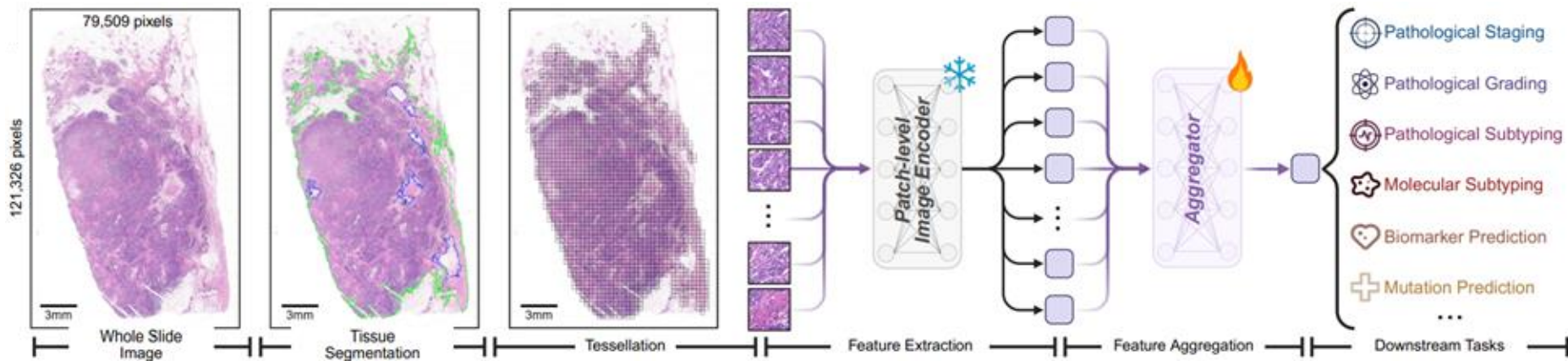
Ovarian Cancer Subtyping AUROC

Top: Transcandian Validation
Bottom: OCEAN Validation



Source: Breen, A Comprehensive Evaluation of Histopathology Foundation Models for Ovarian Cancer Subtype Classification, 2024

How to Use a Pathology Foundation Model



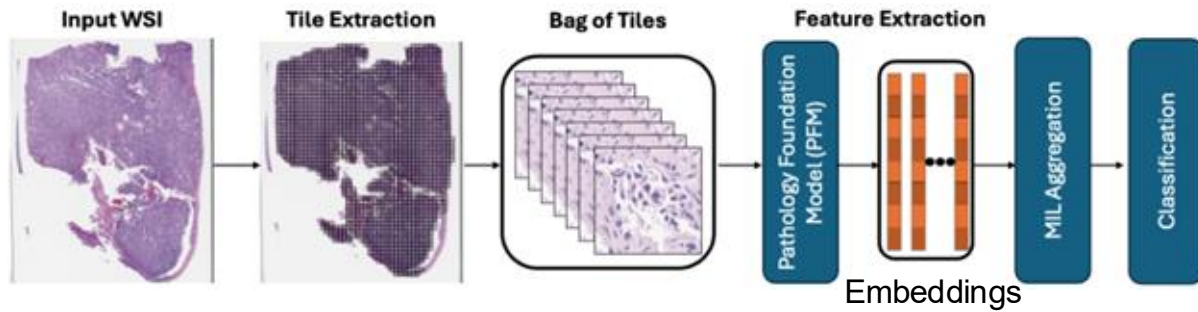
Source: Ma, PathBench: A comprehensive comparison benchmark for pathology foundation models towards precision oncology, 2025

Core Benefits of Foundation Models

- 1) Gigapixel images and weak supervision
- 2) Data efficiency
- 3) Rapid prototyping
- 4) Generalization
- 5) Integration with heterogeneous data

Core Benefits of Foundation Models

1) Gigapixel images and weak supervision



2) Data efficiency

3) Rapid prototyping

4) Generalization

5) Integration with heterogeneous data

Source: Kumar, Single GPU Task Adaptation of Pathology Foundation Models for Whole Slide Image Analysis, 2025

Core Benefits of Foundation Models

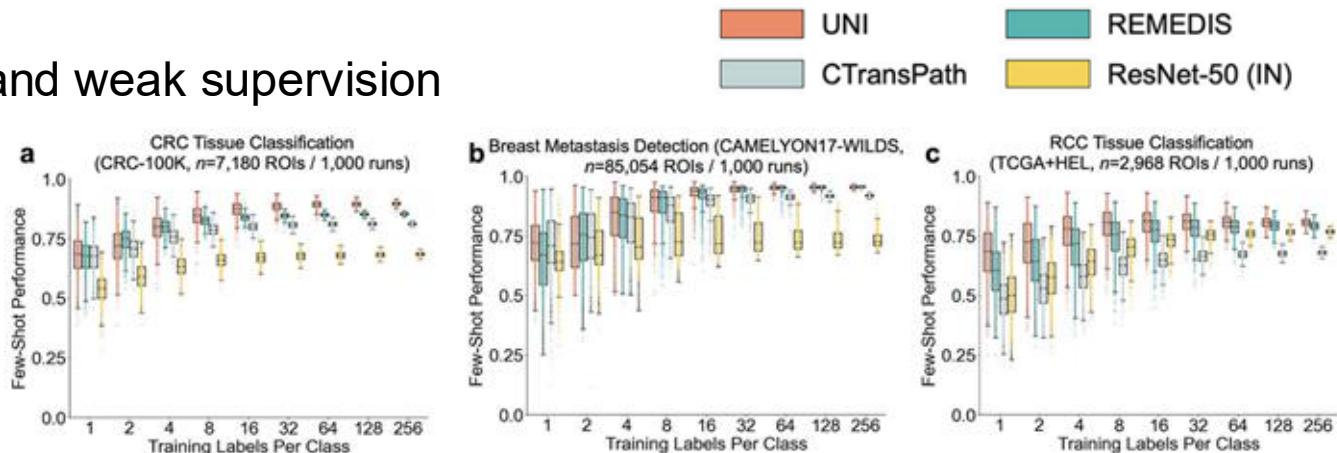
1) Gigapixel images and weak supervision

2) **Data efficiency**

3) Rapid prototyping

4) Generalization

5) Integration with heterogeneous data



Source: Chen, Towards a General-Purpose Foundation Model for Computational Pathology, 2024

Improved accuracy with fewer labels

Core Benefits of Foundation Models

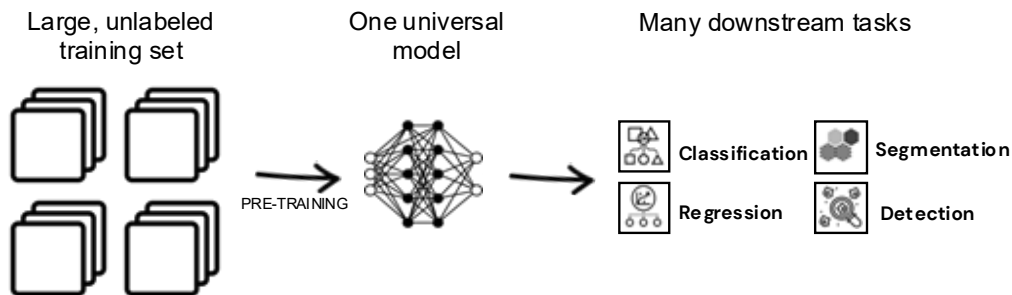
1) Gigapixel images and weak supervision

2) Data efficiency

3) **Rapid prototyping**

4) Generalization

5) Integration with heterogeneous data



Core Benefits of Foundation Models

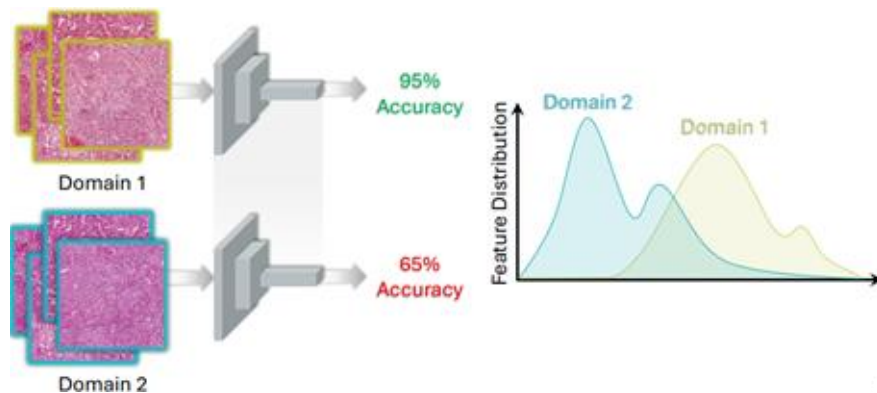
1) Gigapixel images and weak supervision

2) Data efficiency

3) Rapid prototyping

4) Generalization

5) Integration with heterogeneous data



Source: Zamanitajeddin, Benchmarking Domain Generalization Algorithms in Computational Pathology, 2024

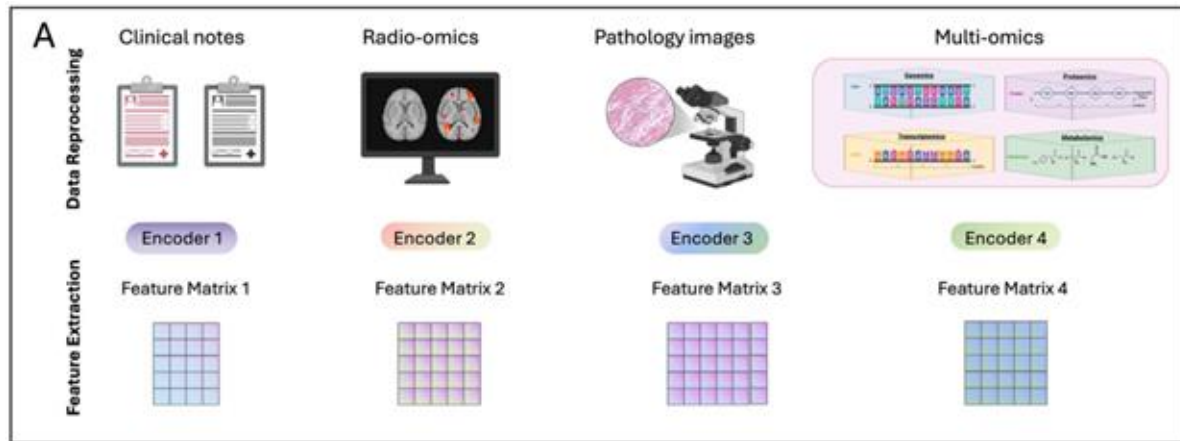
Core Benefits of Foundation Models

1) Gigapixel images and weak supervision

2) Data efficiency

3) Rapid prototyping

4) Generalization



5) Integration with heterogeneous data

Source: Zhang, Multimodal integration strategies for clinical application in oncology, 2025

Taxonomy of Pathology Foundation Models

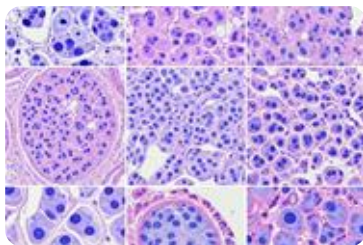
Single Modality (Images)

Tile-Level

(Patches)

Tile Features

Visual representations of localized morphology.



Multimodal (Images + Text/Data)

Tile + Text

Alignment of image patches with text descriptions.

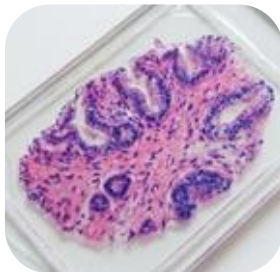


Slide-Level

(WSI)

Slide Reasoning

Global representations capturing spatial context across the entire tissue section.

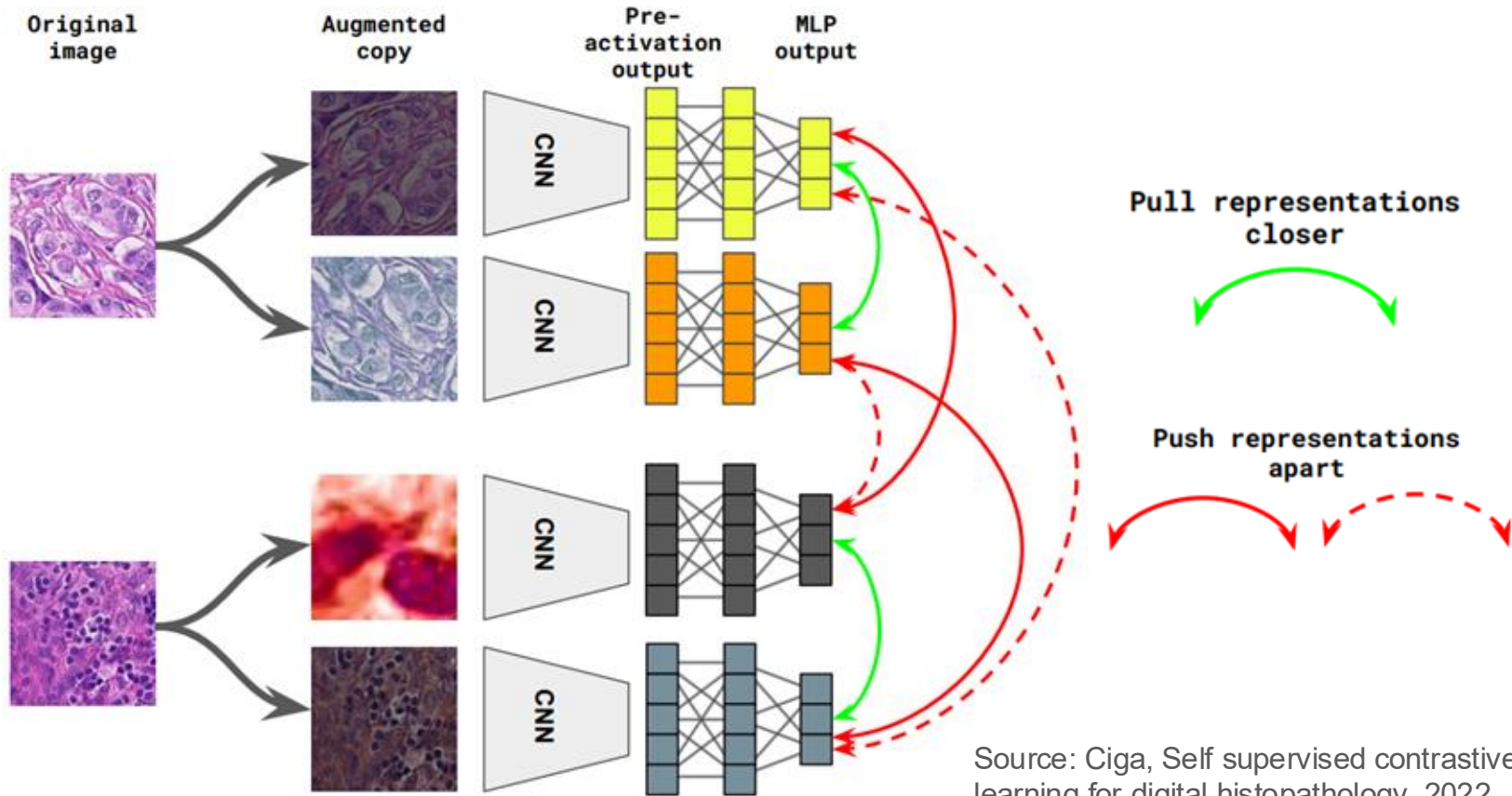


Integrated Insights

Combining WSIs with radiology, genomics, transcriptomics, and/or electronic health records.

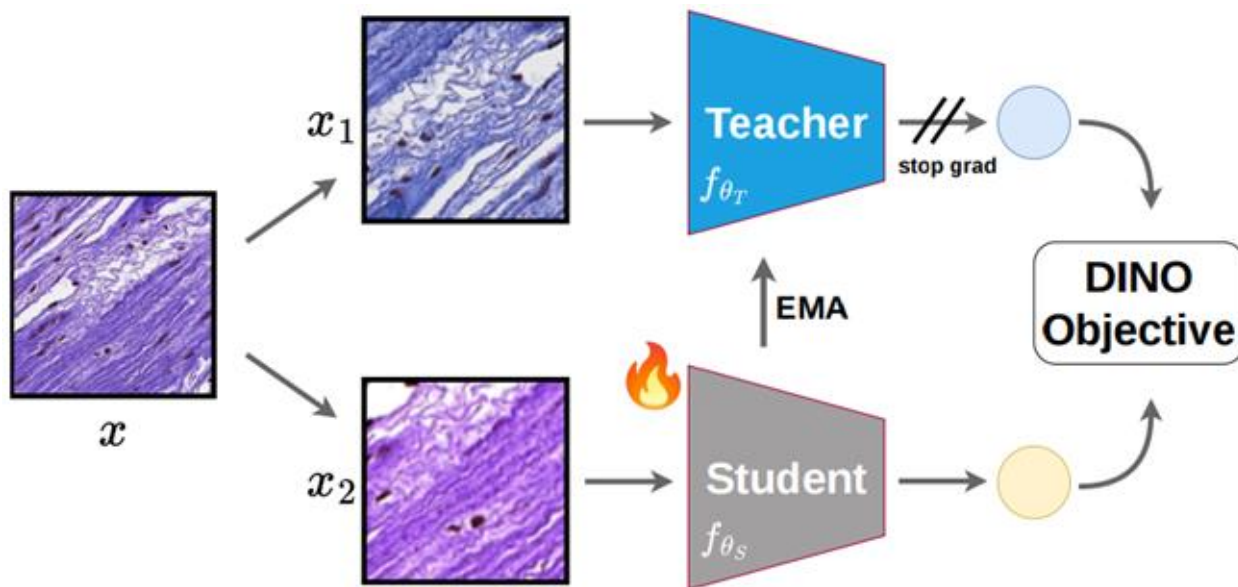


Self-Supervised Learning: Contrastive



Self-Supervised Learning: Self-Distillation

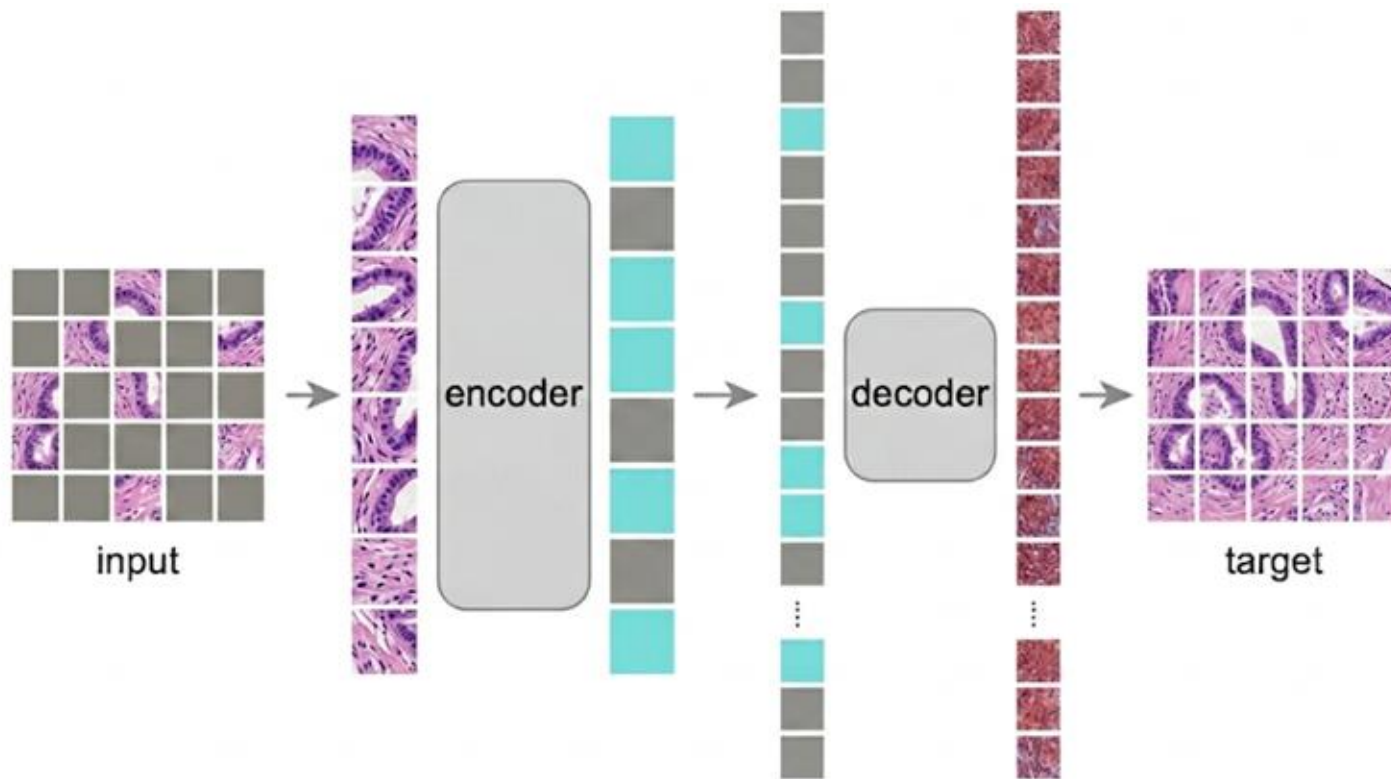
DINO:
distillation with
no labels



Source: Kapse, GenBio-PathFM: A State-of-the-Art Foundation Model for Histopathology, 2026

Self-Supervised Learning: Reconstruction

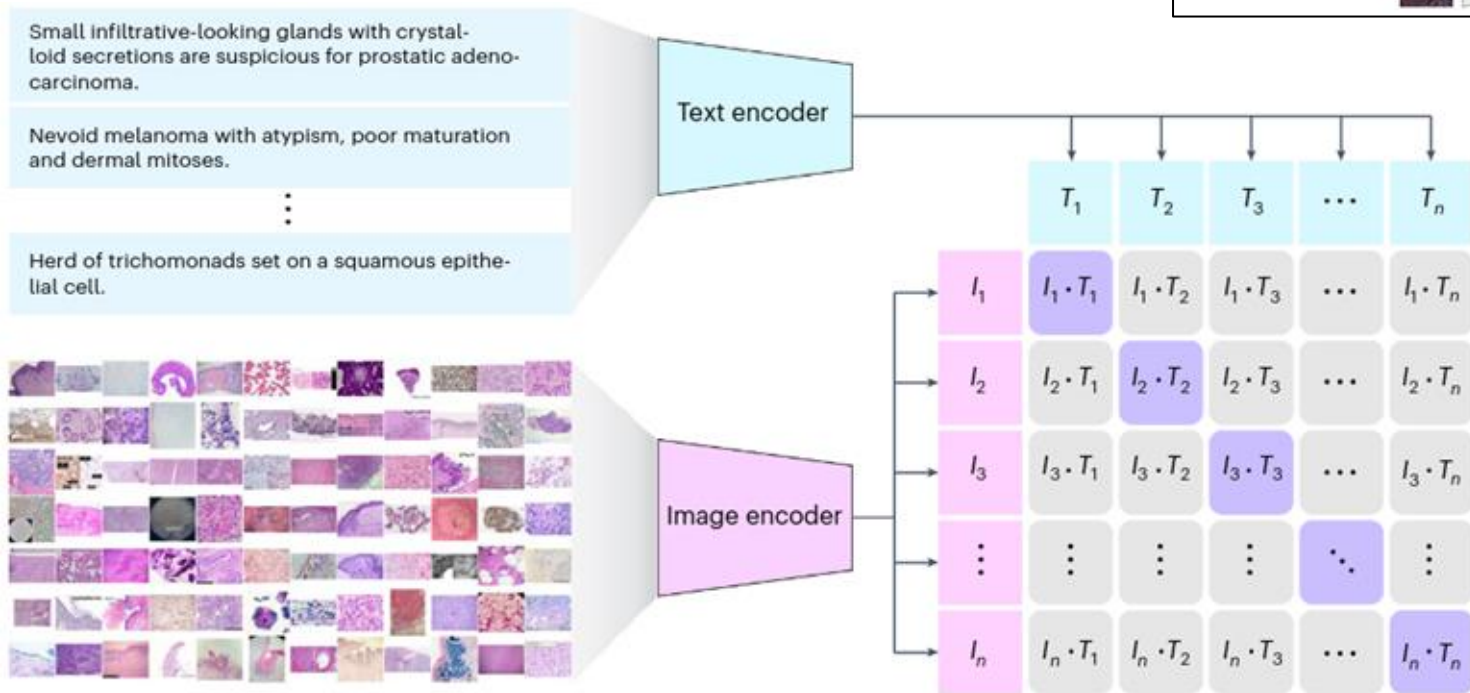
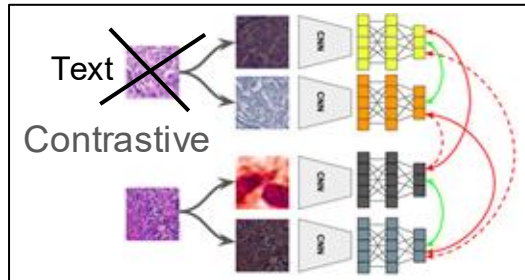
MAE: masked
autoencoder



Multimodal: H&E + Language

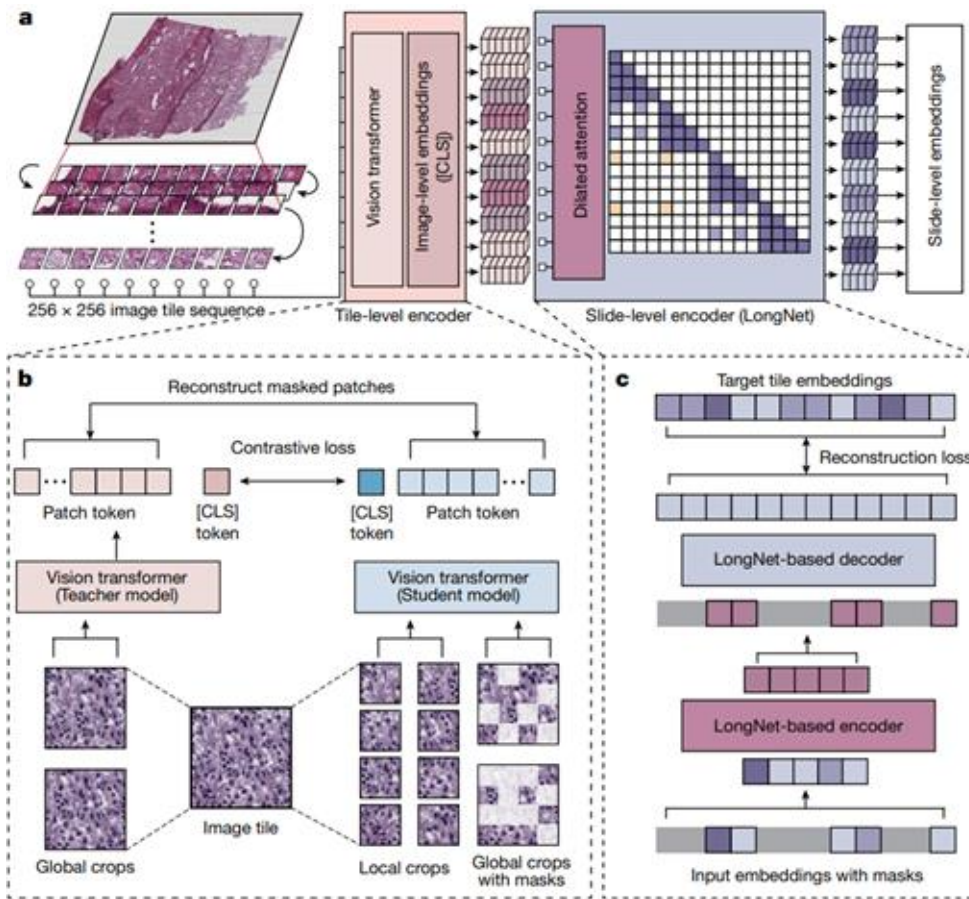
PLIP: Pathology Language–Image Pretraining

Fine-tuned CLIP on 208k pathology images + natural language description pairs



Source: Huang, A visual–language foundation model for pathology image analysis using medical Twitter, 2023

Slide-Level Foundation Model



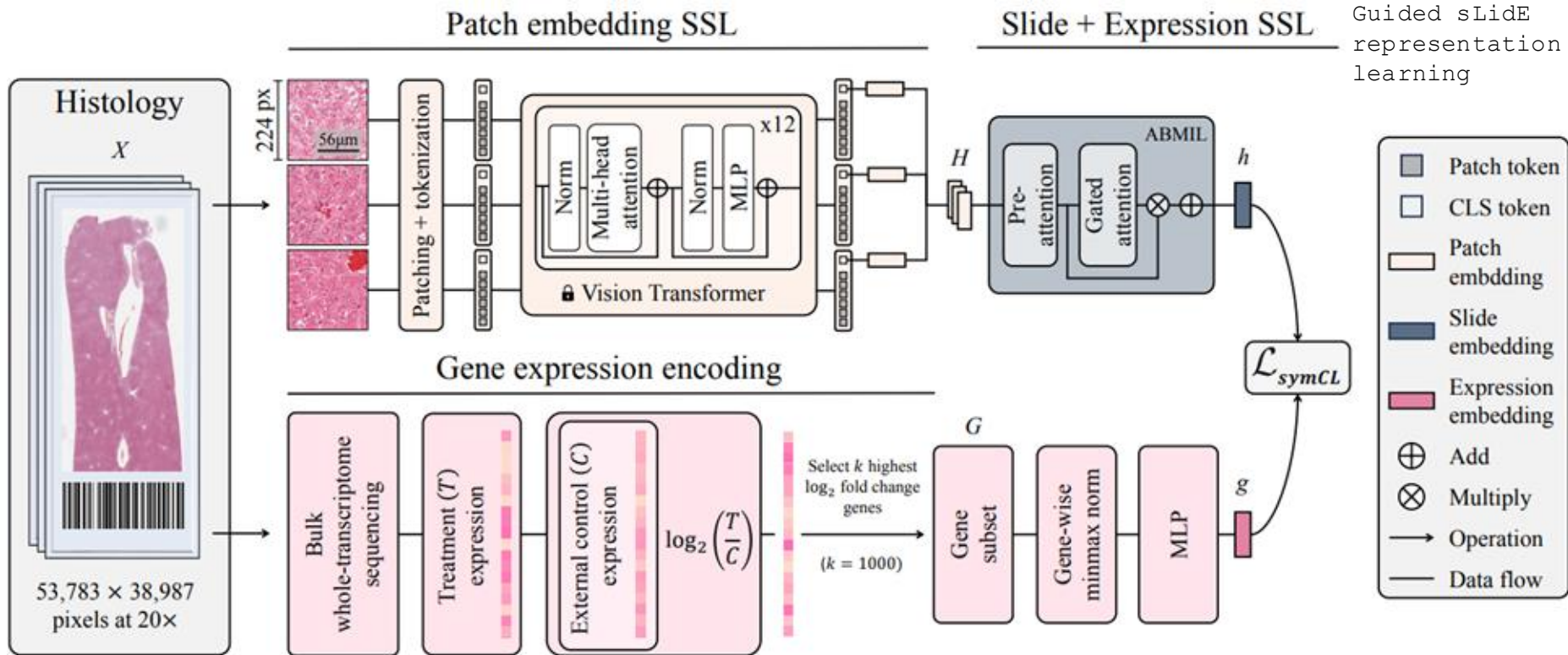
Prov-GigaPath

- Tile-encoder trained with DINOv2
- Slide-encoder trained with masked autoencoder

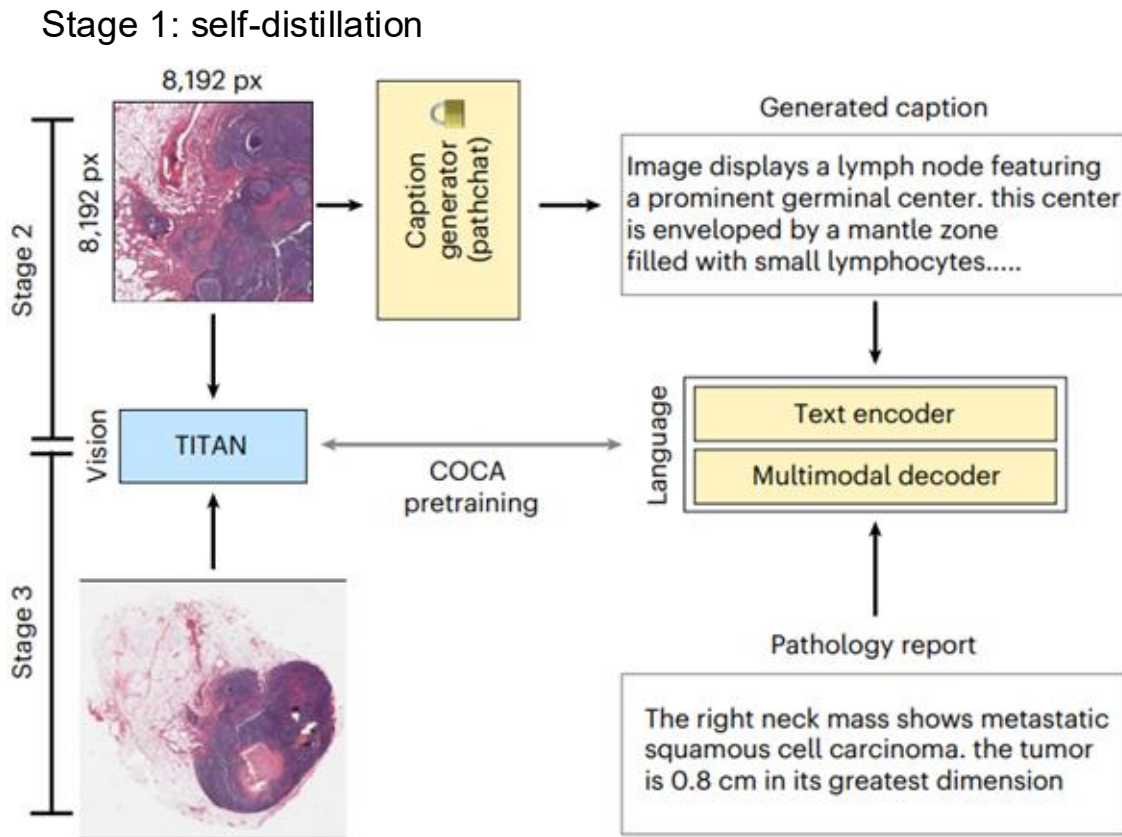
Source: Xu, A whole-slide foundation model for digital pathology from real-world data, 2024

Slide-Level Multimodal: H&E + Gene Expression

TANGLE:
TrANscriptomics-
Guided sLIDE
representation
learning



Slide-Level Multimodal: H&E + Language

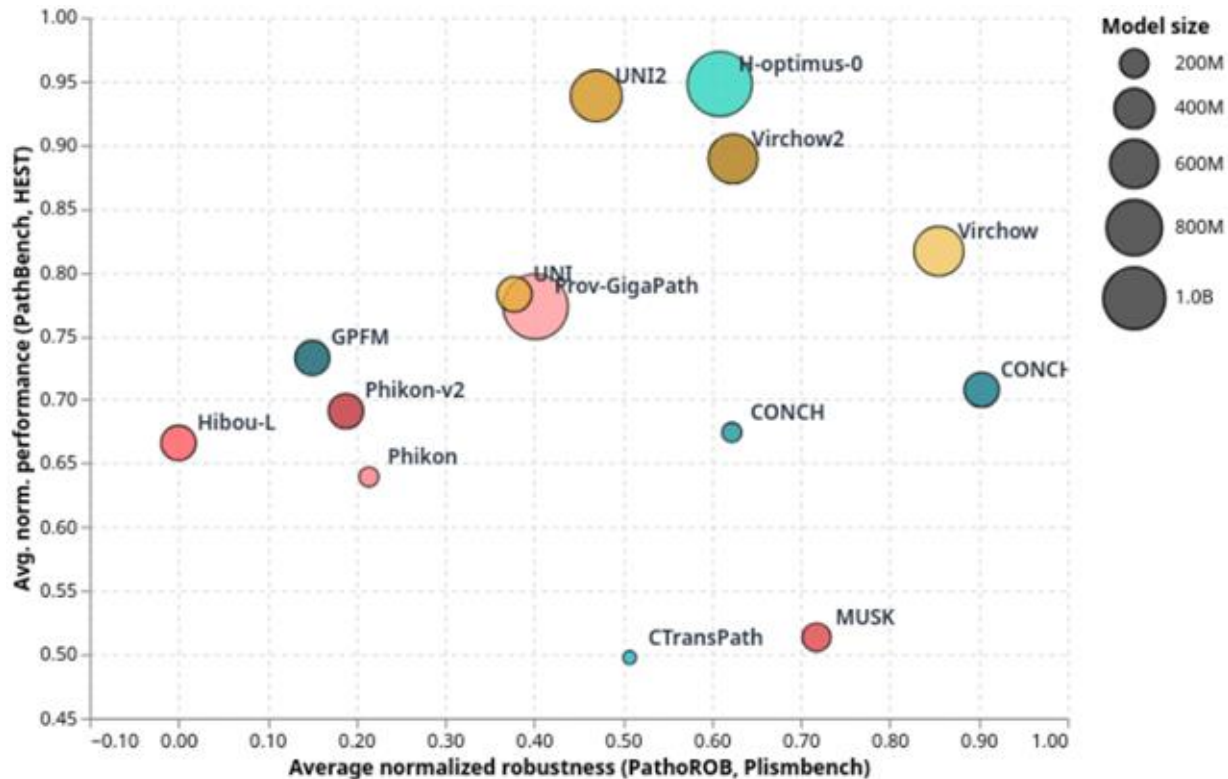


TITAN:
Transformer-based
pathology Image
and Text
Alignment Network

Source: Ding, A multimodal whole-slide foundation model for pathology, 2025

A Diverse Set of Tile-Based Foundation Models

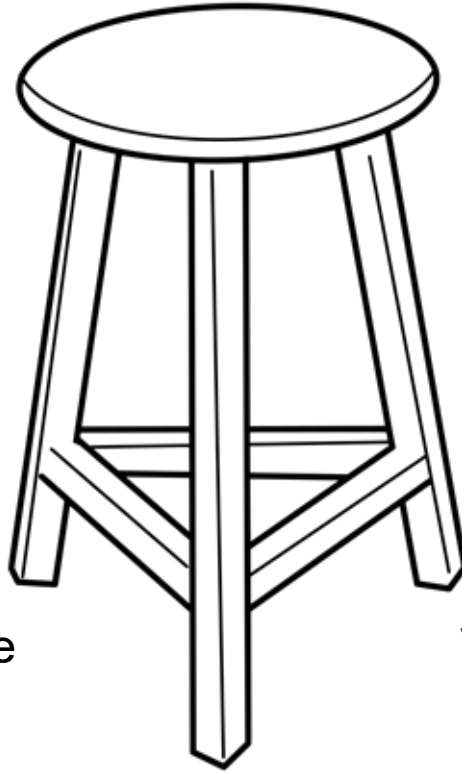
Robustness vs. Performance



Histoboard currently aggregates results from 10 published benchmarks covering:

- 411 evaluation tasks spanning classification, survival prediction, biomarker detection, and more
- 46 foundation models from academic and industry labs worldwide
- 20 organs including Bladder, Brain, Breast, Cervix, Colorectal, and others
- Robustness evaluation across domain shifts, scanners, and staining variations

From Promise to Practice: Three Pillars

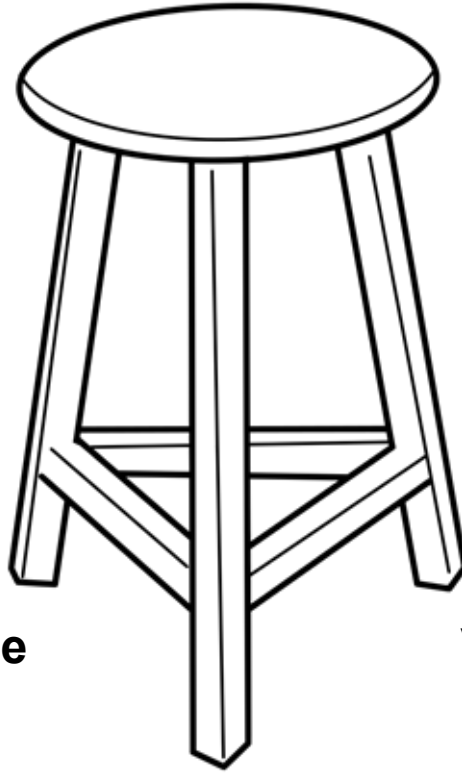


Trustworthy:
Robust & Reliable

Actionable:
Valuable & Sustainable

Responsible:
Fair & Transparent

From Promise to Practice: Three Pillars



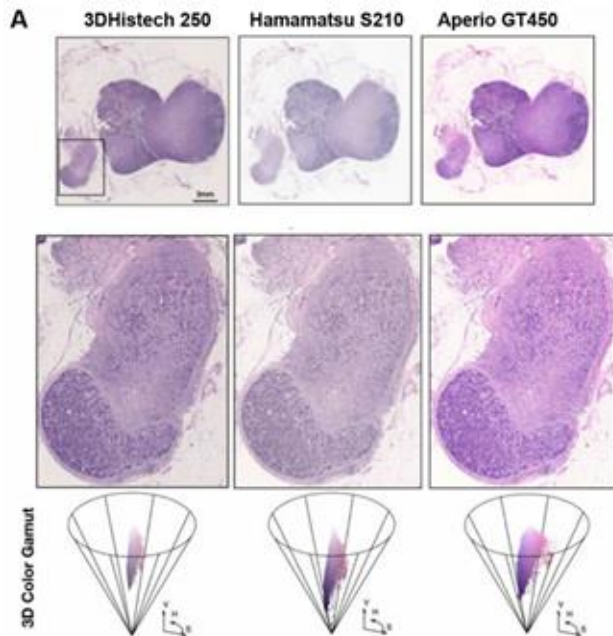
Trustworthy:
Robust & Reliable

Actionable:
Valuable & Sustainable

Responsible:
Fair & Transparent

Examples of Distribution Shift & Batch Effects

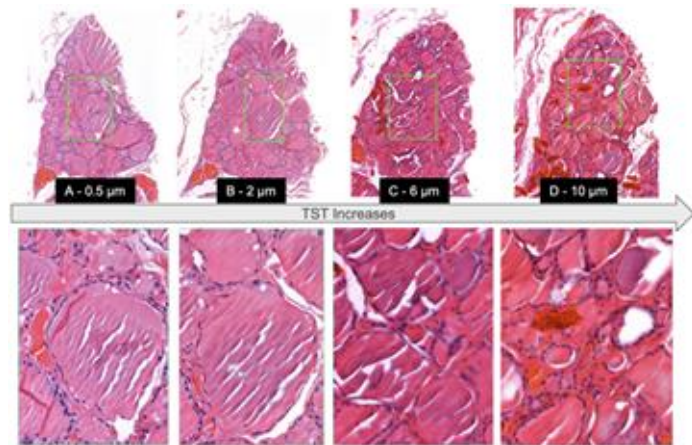
Scanner Variability



Source: Chen, Algorithm fairness in artificial intelligence for medicine and healthcare, 2023

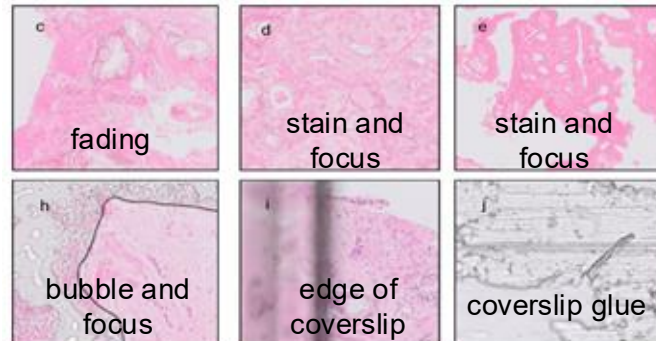
Tissue Thickness Variations

Source: Shah, Impact of Tissue Thickness on Computational Quantification of Features in Whole Slide Images for Diagnostic Pathology, 2025

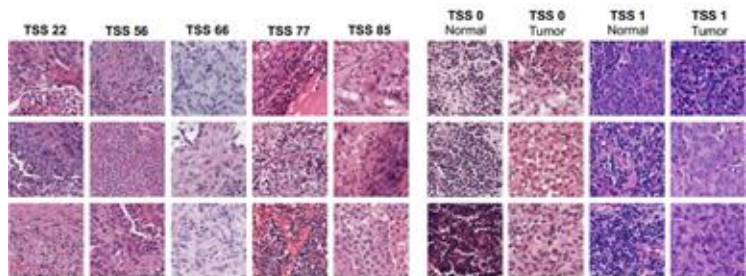


Slide Preparation and Image Acquisition Artifacts

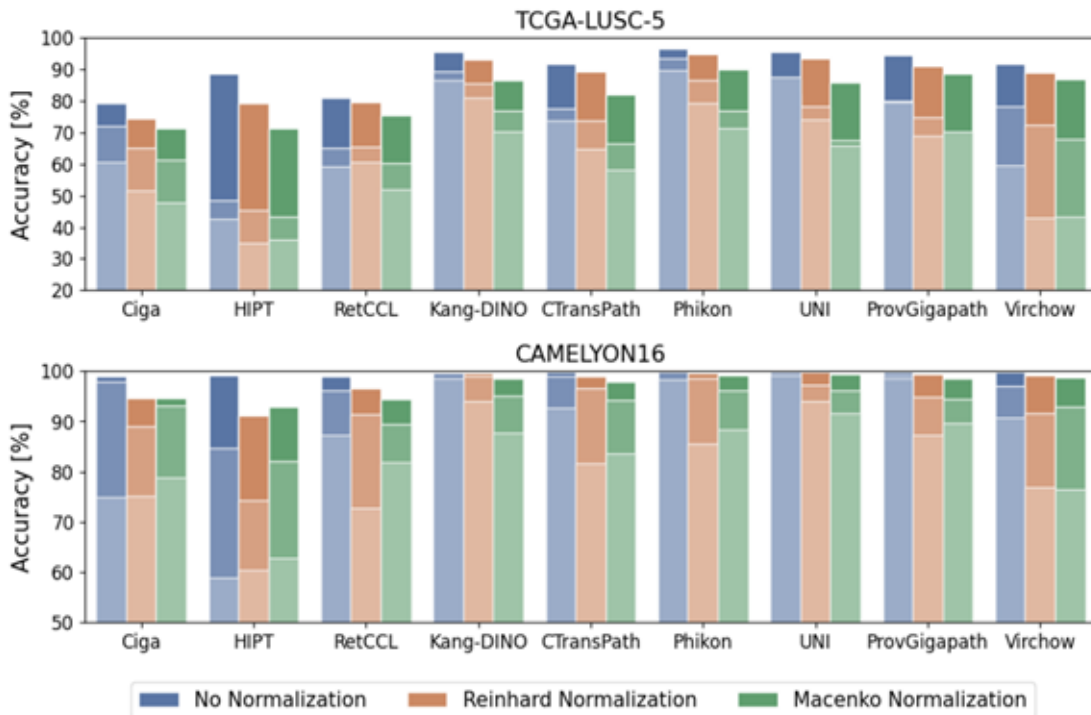
Source: Haghghat, PathProfiler: Automated Quality Assessment of Retrospective Histopathology Whole-Slide Image Cohorts by Artificial Intelligence – A Case Study for Prostate Cancer Research, 2021



Foundation Models Encode Batch Effects

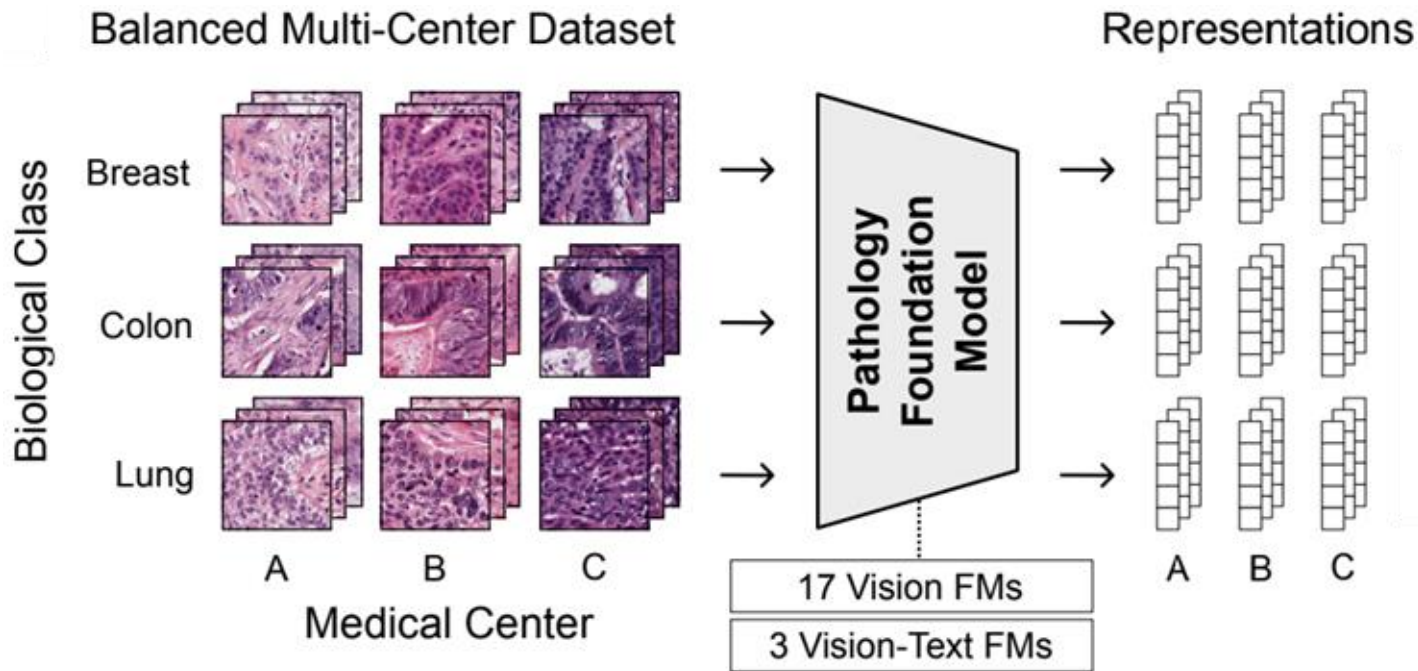


Tissue Source Site Prediction



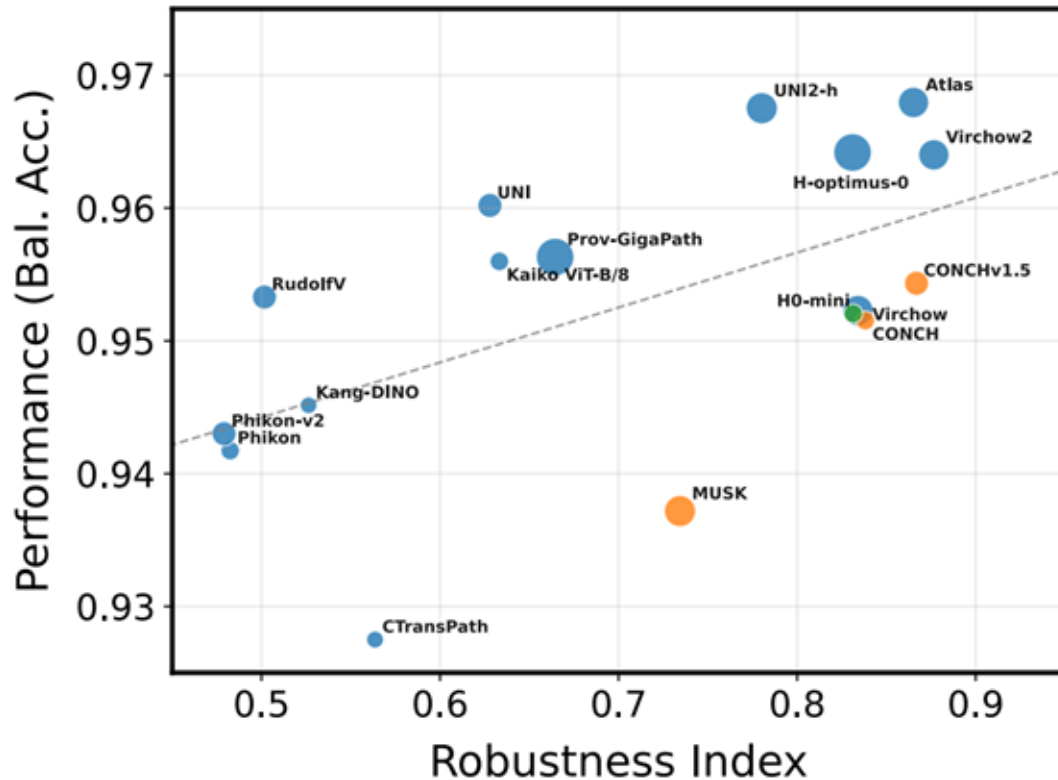
Benchmarking Robustness

PathoROB Benchmark



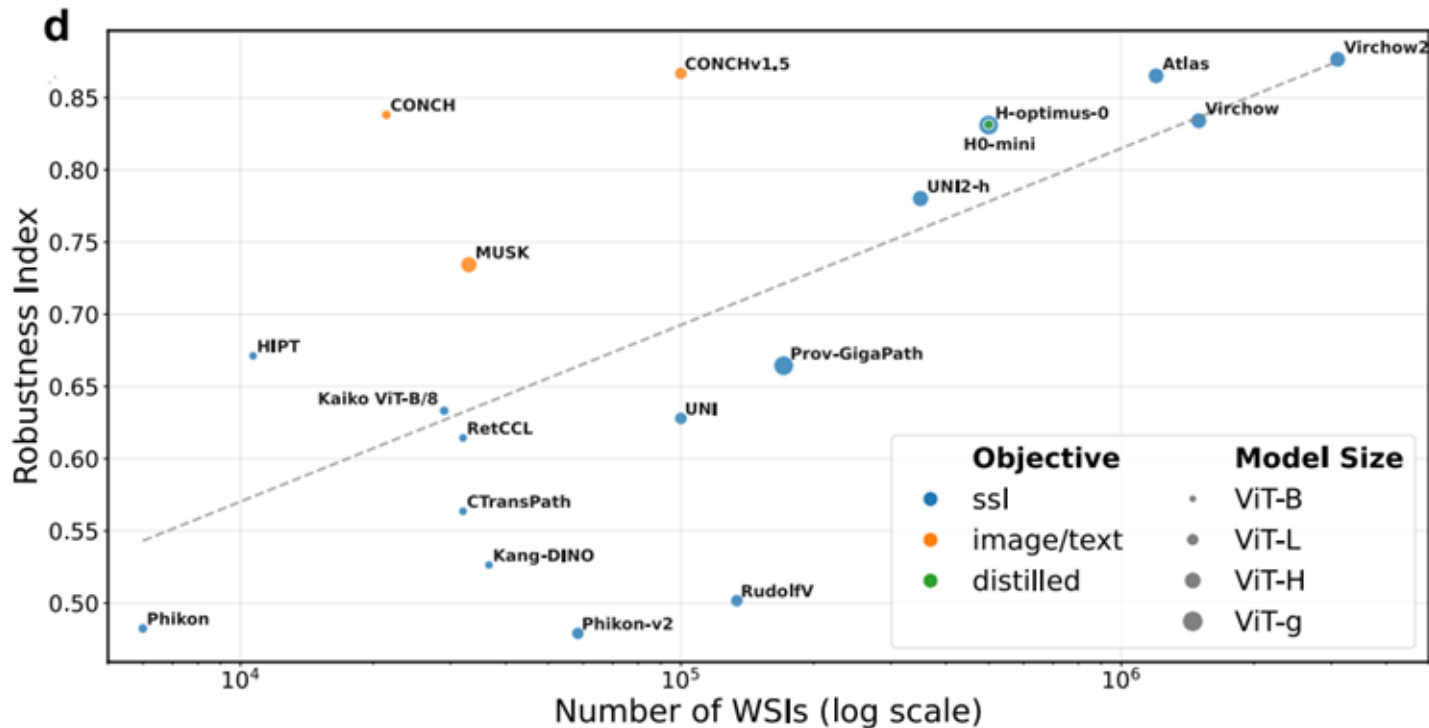
Source: Komen, Towards Robust Foundation Models for Digital Pathology, 2025

Some Models Are More Robust Than Others



Source: Komen, Towards Robust Foundation Models for Digital Pathology, 2025

What Makes a Model More Robust?



Larger training set?

More diverse training set?

Multimodal objective?

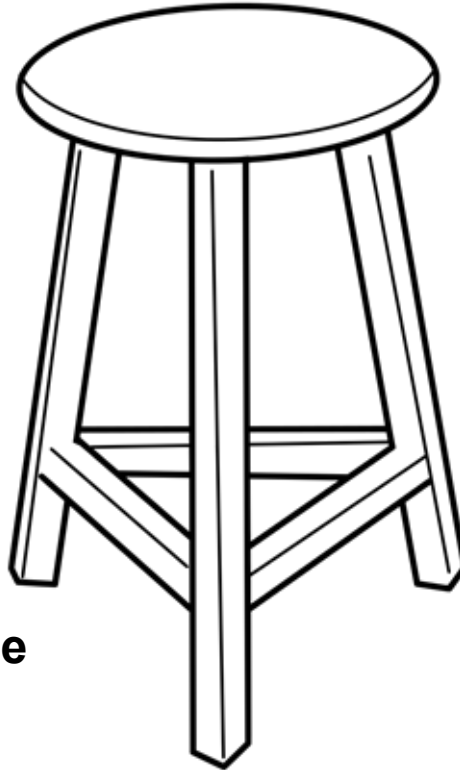
Source: Komen, Towards Robust Foundation Models for Digital Pathology, 2025

From Promise to Practice: Three Pillars

Beware of technical variations.

Some FMs are more robust than others.

All are susceptible to batch effects.

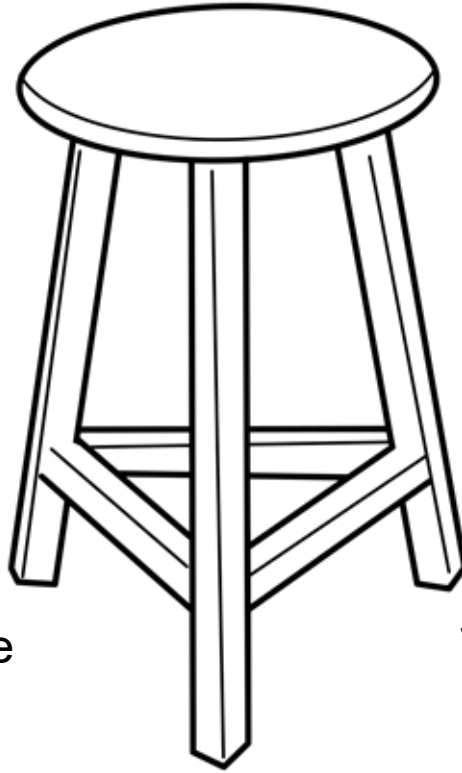


Trustworthy:
Robust & Reliable

Actionable:
Valuable & Sustainable

Responsible:
Fair & Transparent

From Promise to Practice: Three Pillars



Trustworthy:
Robust & Reliable

Actionable:
Valuable & Sustainable

Responsible:
Fair & Transparent

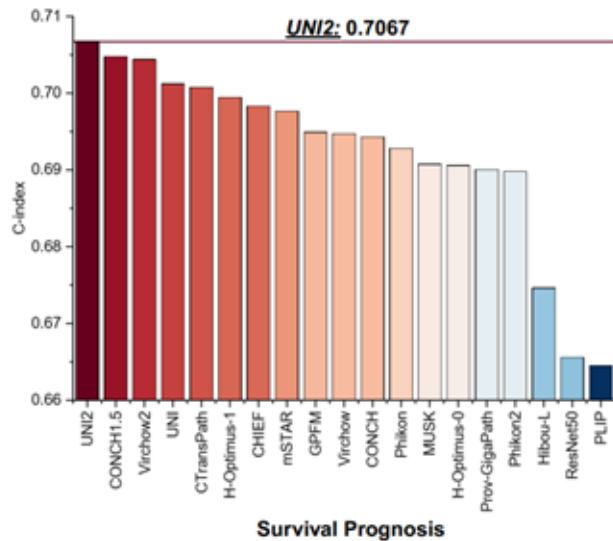
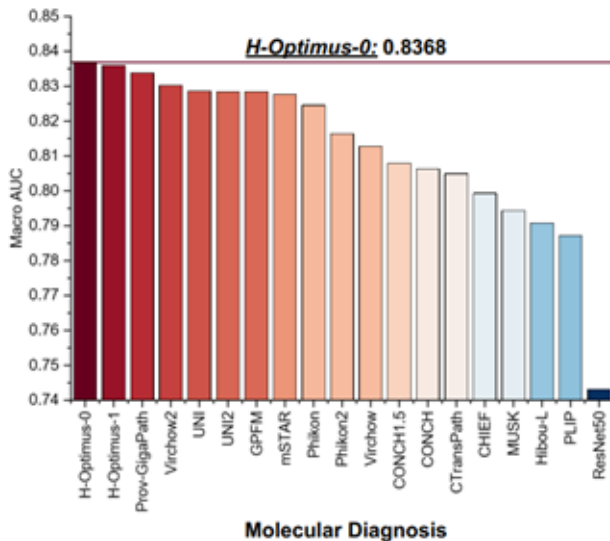
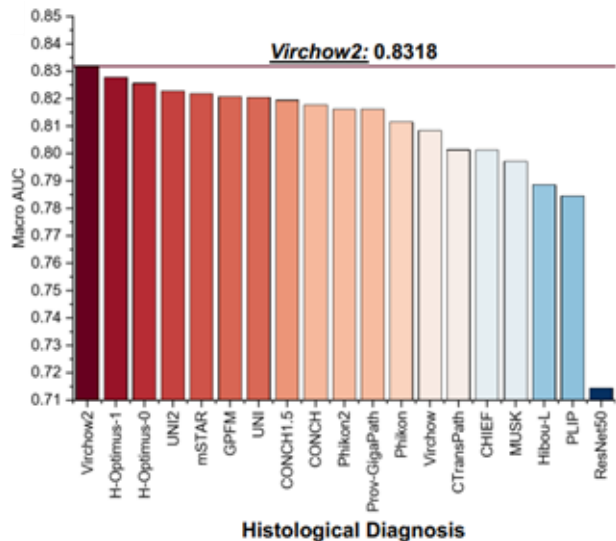
The Limitations of Benchmarks

Model	EVA	PathBench	Stanford	HEST	Patho-Bench	Sinal	STAMP	THUNDER	PathoROB
Virchow2 NC VIT-H/14 (632M)	🏆 1/5	🏆 1/8	🏆 1/8	🥈 2/9	🏆 1/2	🥈 2/4	🏆 1/4	🥈 2/7	🏆 1/7
Prov-GigaPath Open VIT-g/14 (1.1B)	4/5	4/8	🥈 2/8	5/9	-	🏆 1/4	🥈 2/4	4/7	5/7
Phikon-v2 NC VIT-L/14 (307M)	5/5	5/8	7/8	8/9	-	🥈 3/4	-	7/7	7/7
UNI2 NC VIT-H/14 (682M)	🥈 2/5	🥈 2/8	🥈 3/8	🏆 1/9	🥈 2/2	-	-	🏆 1/7	4/7
Hibou-L Open VIT-L/14 (307M)	🥈 3/5	8/8	6/8	4/9	-	-	4/4	5/7	-
CTransPath Open Swin-T/4 (28M)	-	7/8	8/8	9/9	-	4/4	🥈 3/4	-	6/7
HO-mini NC VIT-B/14 (86M)	-	-	4/8	🥈 3/9	-	-	-	🥈 3/7	🥈 3/7
CONCH 1.5 NC VIT-L/16 (307M)	-	6/8	-	7/9	-	-	-	6/7	🥈 2/7
GPFM NC VIT-L/14 (303M)	-	🥈 3/8	5/8	6/9	-	-	-	-	-

Source: <https://wearewaiv.github.io/histoboard>

Varied Performance Across Types of Tasks

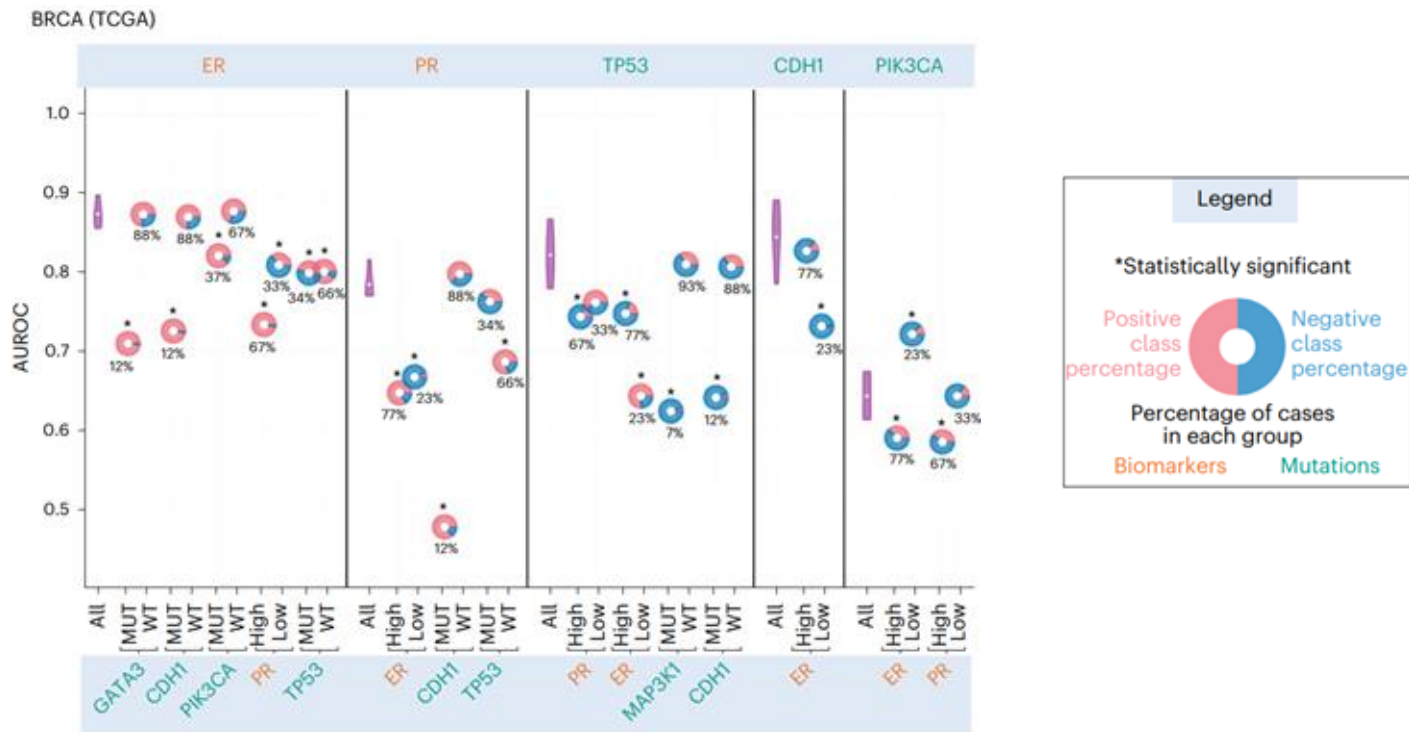
PathBench



Source: Ma, PathBench: A comprehensive comparison benchmark for pathology foundation models towards precision oncology, 2025

Beyond Global AUC

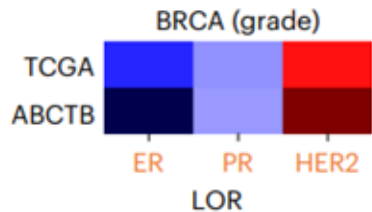
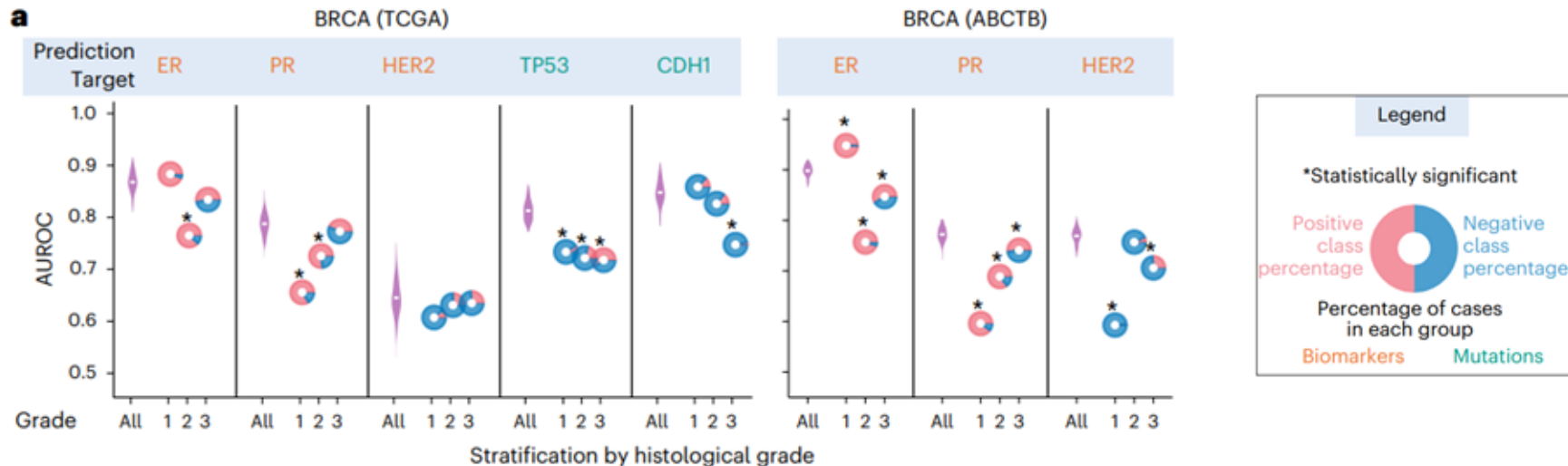
Stratified performance of WSI-based biomarker predictors



Source: Dawood, Confounding factors and biases abound when predicting molecular biomarkers from histological images, 2026

Checking for Bias

Biased performance of biomarker predictors across histological grades



Source: Dawood, Confounding factors and biases abound when predicting molecular biomarkers from histological images, 2026

Is the model capturing the relevant biological signal or relying on confounders?

mutual exclusivity -3 3 co-occurrence
log odds ratio

Where Does Pre-Training Data Come From?

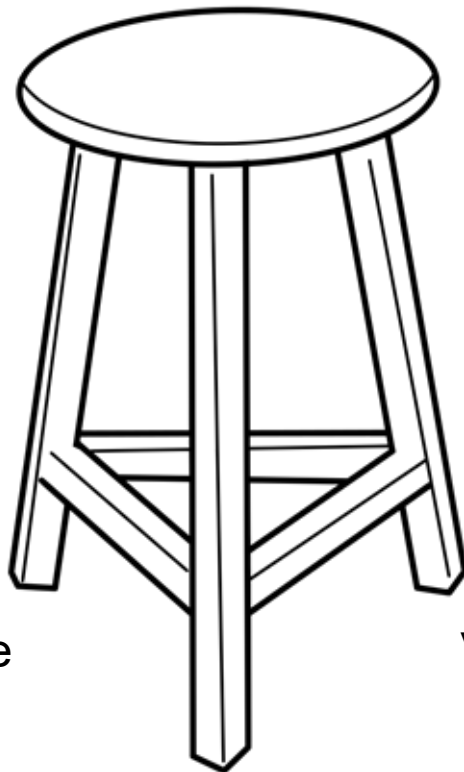
Model	Organization	Release date	# Cases	# WSIs	# Tiles	# Scanners	# Sites	Stains	Magnification	# Tissue types
mSTAR	HKUST	2024-03	26k	11k	116m	NA	>30	H&E	20x	32
Hibou-L	Hist.ai	2024-06	306k	1.1m	1.2b	NA	NA	H&E, non-H&E, cytology	NA	Human + vet
Prov-GigaPath	Microsoft	2024-06	30k+	171k	1.4b	NA	28	H&E, IHC	20x	31 tissue types
TITAN	Harvard	2024-07	NA	336k	NA	16	3	H&E, IHC, special	20	20 organ types
Virchow2	Paige	2024-08	225k	3.1m	1.9b	NA	diverse	H&E, IHC	5/10/20/40	200 tissue types
Phikon-v2	Owkin	2024-09	NA	58k	456m	>50	>100	NA	20x	>30 sites
CHIEF	Harvard	2024-09	NA	61k	15m	NA	NA	H&E	10x	19
CanvOI	Imagene	2024-09	NA	633k	70m	NA	100	Mostly H&E	20x	40
UNI2	Harvard	2025-01	NA	>350k	200m	NA	1	H&E, IHC	20x	>20 organ types
H-optimus-1	Biopitimus	2025-03	>800k	>1m	billions	3	4000	H&E	NA	50 organs
Midnight-12k	kaiko.ai	2025-04	NA	12k	NA	NA	NA	NA	5/10/20/40	32 cancer types
Atlas 2	Aignostics	2026-01	NA	5.5m	NA	NA	3	NA	5/10/20/40	NA
GenBio-PathFM	GenBio AI	2026-03	NA	177k	400m	NA	NA	H&E, IHC	5/10/20/40	7

From Promise to Practice: Three Pillars

Benchmarking helps, but what really matters is performance on your data.

Validate thoroughly with stratified metrics and check for confounders.

Ask questions about FM training data.

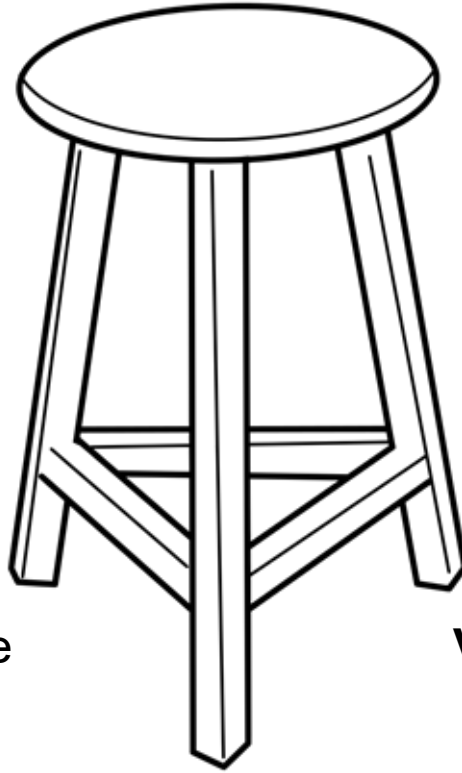


Trustworthy:
Robust & Reliable

Actionable:
Valuable & Sustainable

Responsible:
Fair & Transparent

From Promise to Practice: Three Pillars

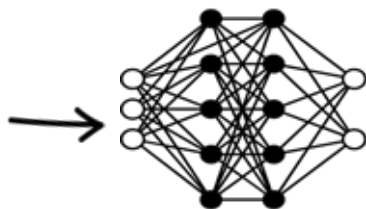
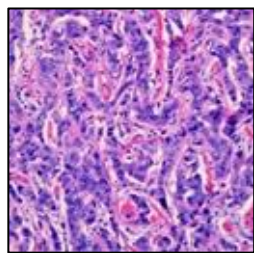


Trustworthy:
Robust & Reliable

Actionable:
Valuable & Sustainable

Responsible:
Fair & Transparent

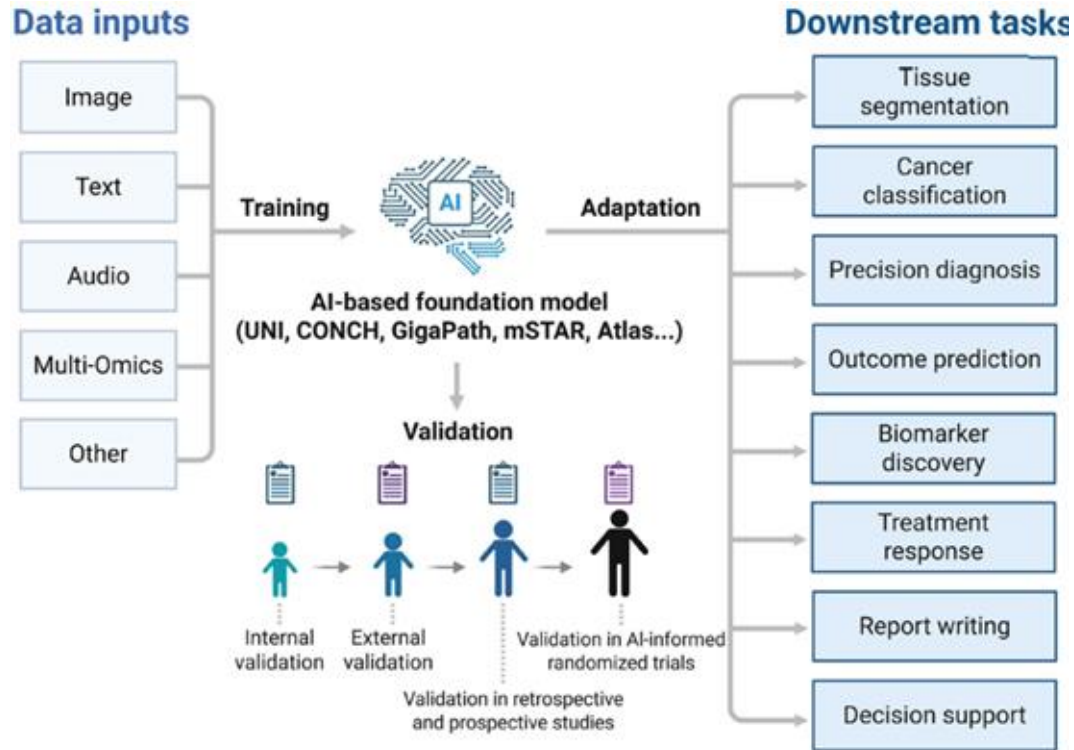
Embeddings Aren't Diagnoses



1024-D vector

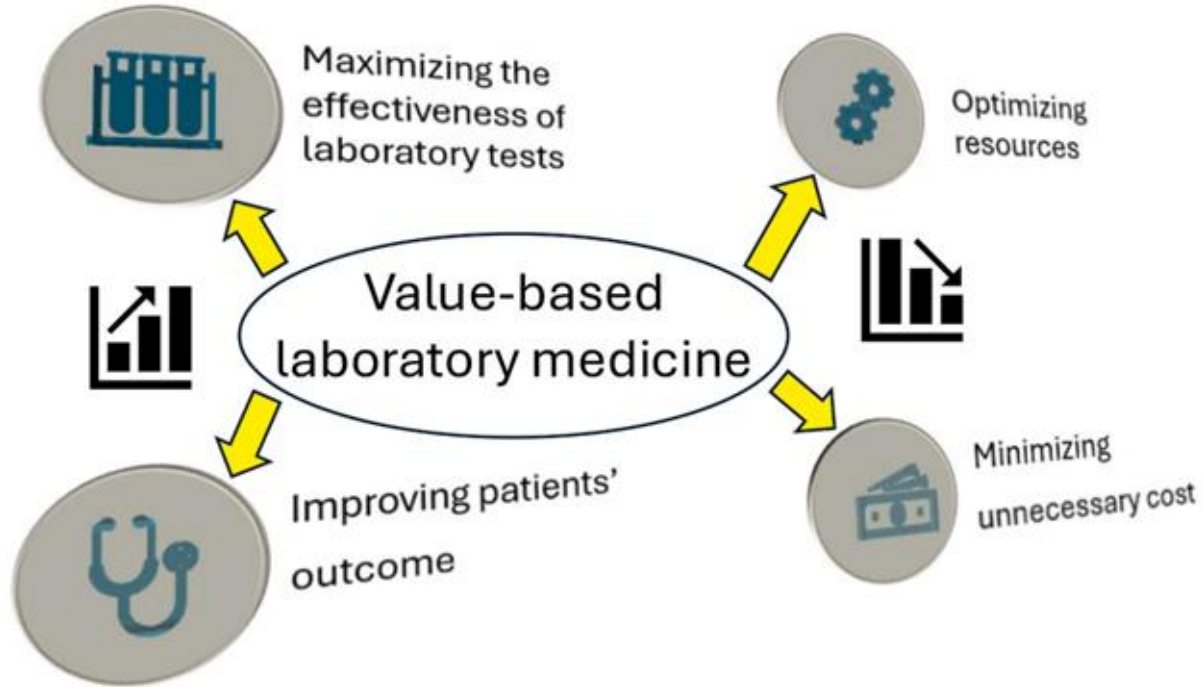
?

How Do We Use Foundation Models?



Source: Cheng, The role of artificial intelligence-based foundation models and “copilots” in cancer pathology: potential and challenges, 2026

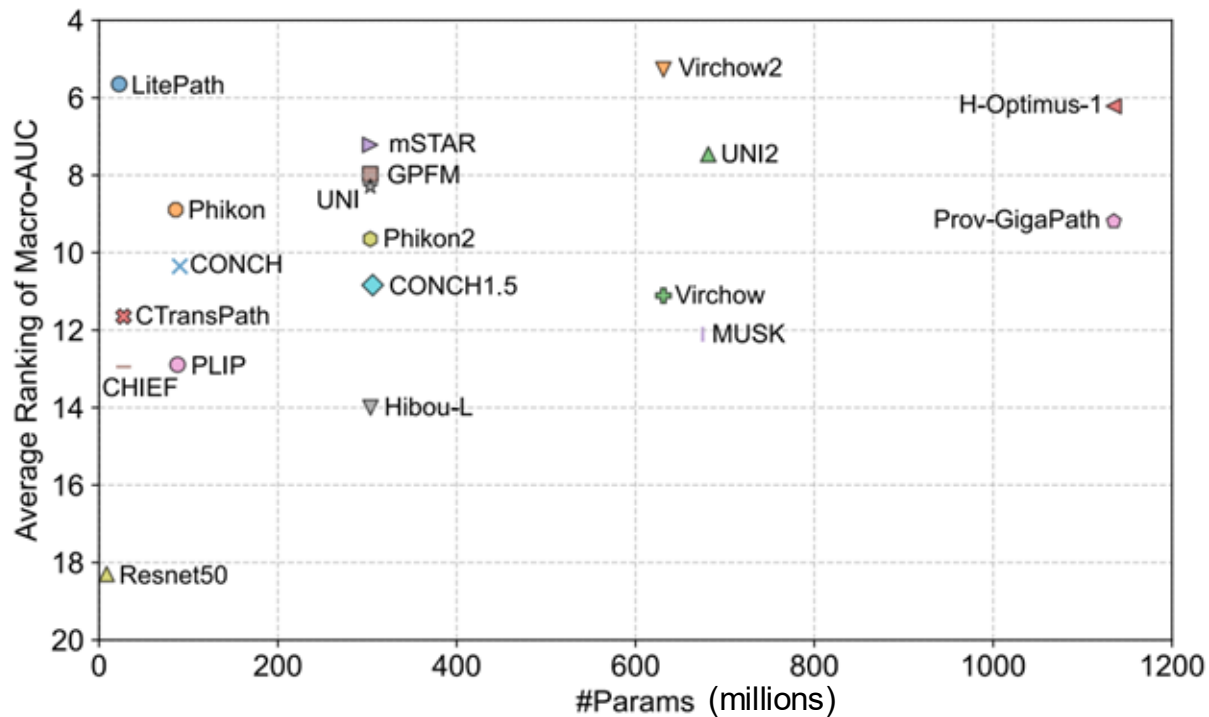
The Value Proposition



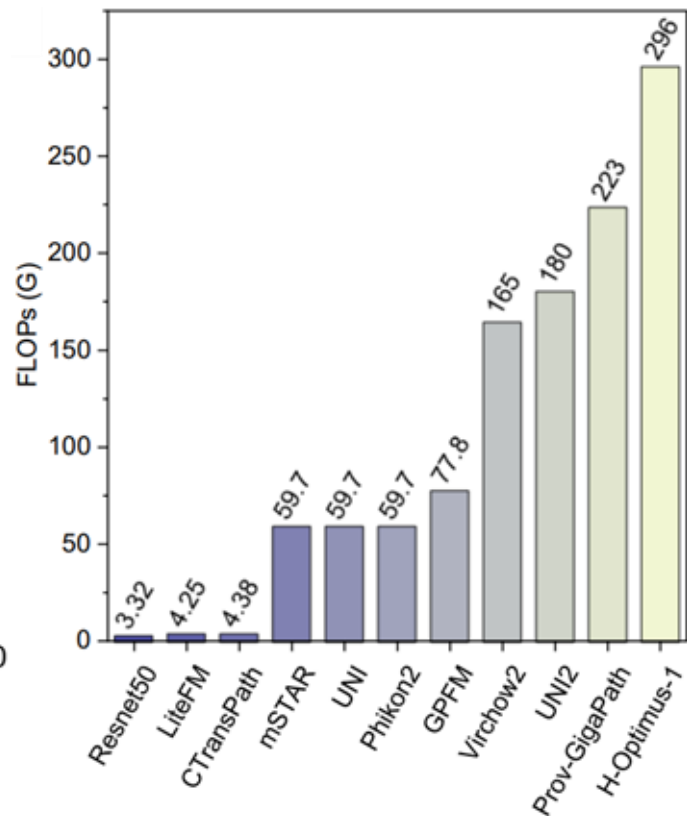
Source: Plebani, Promoting value-based laboratory medicine: Moving towards an innovative model of clinical laboratory, 2025

Does Model Size Matter?

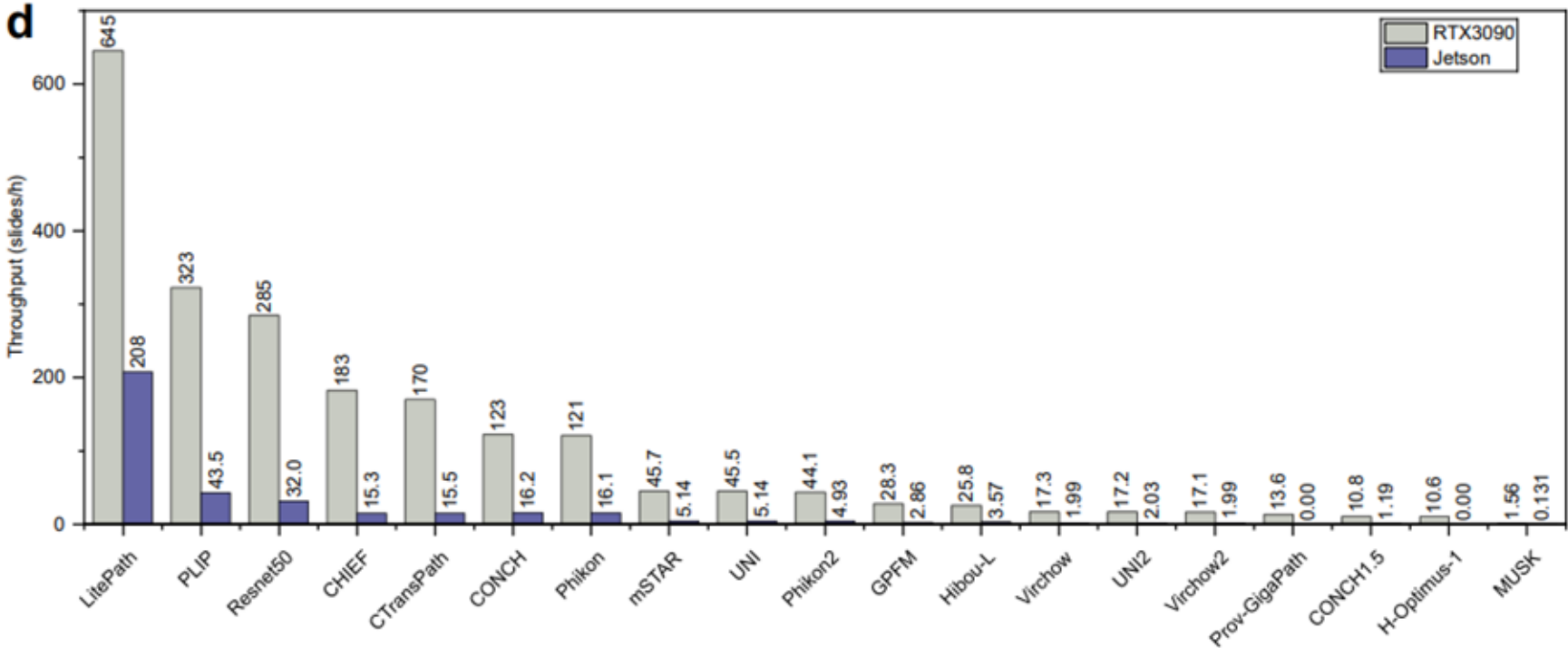
Model Size vs. Performance



Compute per Patch

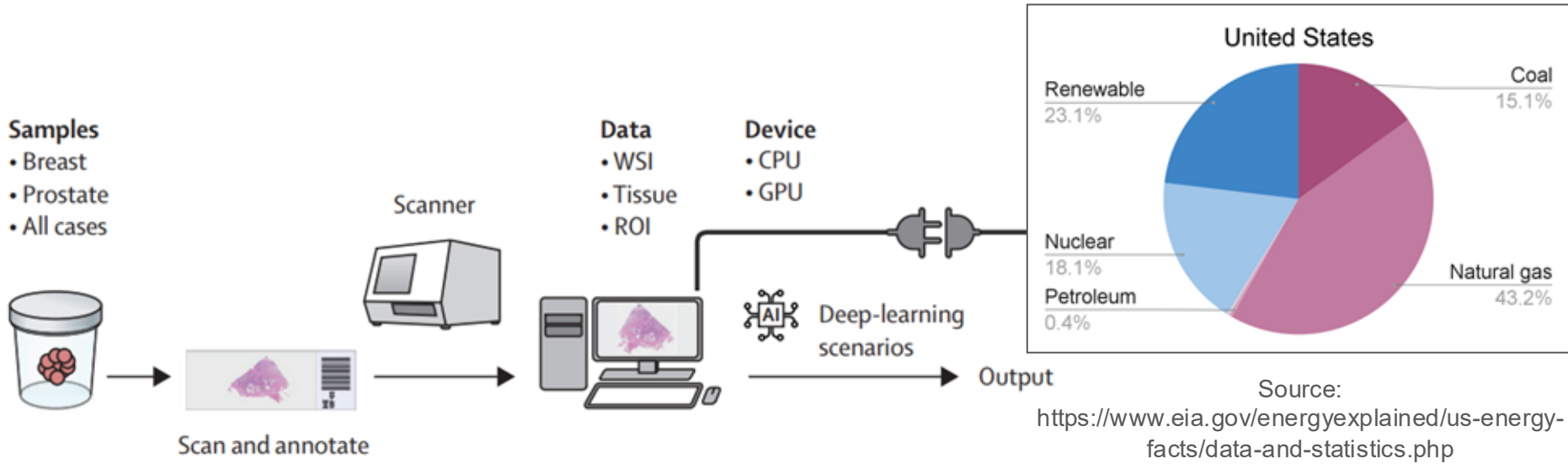


A Comparison of Slide Throughput



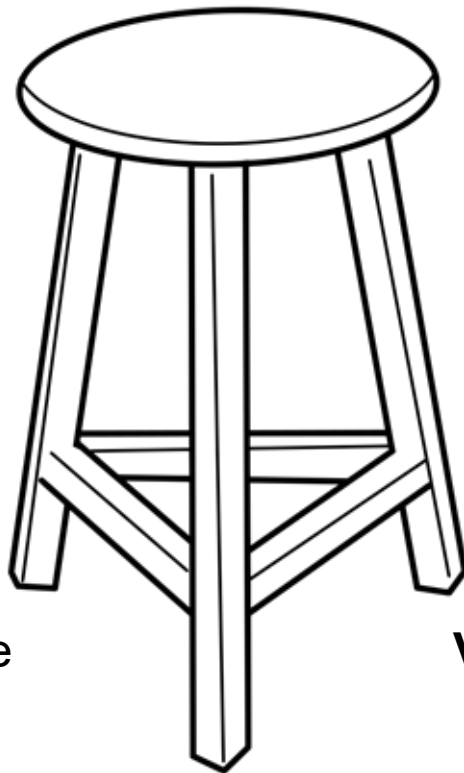
Source: Cai, A Deployment-Friendly Foundational Framework for Efficient Computational Pathology, 2026

Processing of Every WSI Uses Electricity & Produces Emissions



Source: Sadr, Operational greenhouse-gas emissions of deep learning in digital pathology: a modelling study, 2024

From Promise to Practice: Three Pillars



Trustworthy:
Robust & Reliable

Actionable:
Valuable & Sustainable

Responsible:
Fair & Transparent

Integrate into clinical workflow.
Have a clear value proposition.
Be mindful of compute costs.

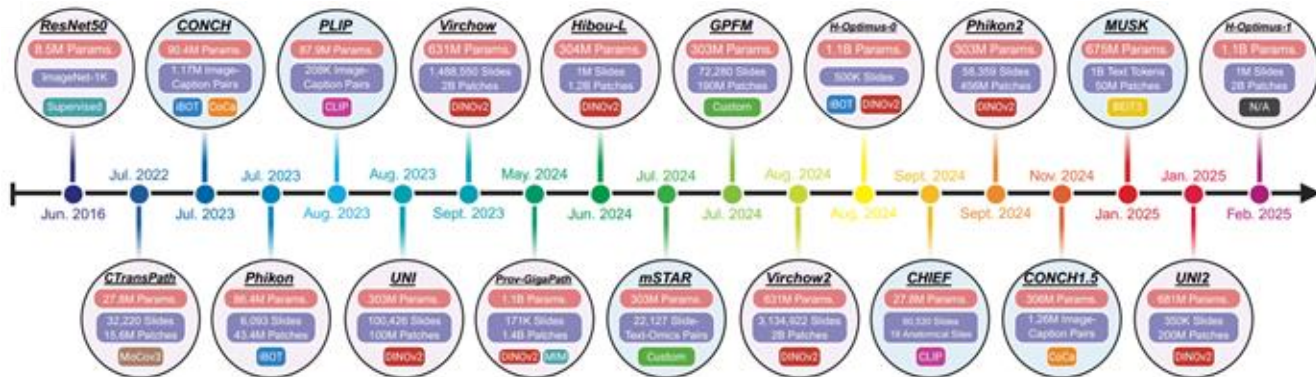
Closing Thoughts

Foundation models are the key components of future digital pathology tools.

They must be built on the pillars of trust, responsibility, and actionability.

Using more powerful and robust FM's doesn't change the need for thorough validation.

Whether you are developing FM's, building products with them, or simply an end user... Ask questions about the models, how they were trained, and how they were validated.



Source: Ma, PathBench: A comprehensive comparison benchmark for pathology foundation models towards precision oncology, 2025



Trustworthy Actionable
Responsible

Contact

**Heather D. Couture,
Pixel Scientia Labs**

<https://www.linkedin.com/in/hdcouture/>

heather@pixelscientia.com

