H-Optimus-1: A foundation model for computational histopathology

Marin Scalbert, Charlie Saillard, Thomas Peeters, Liam Gonzalez, Dasha Valter, Felipe Llinares-López, Zelda Mariet, Rodolphe Jenatton

LB174

Background

- AI has shown strong results in computational pathology across applications such as therapeutic response, molecular biomarkers, and prognosis.
- Foundation models (FM) have emerged as a promising approach, capable of addressing a wide range of downstream tasks simultaneously.
- We introduce H-Optimus-1, a histology FM achieving state-of-the-art (SOTA) performance across key downstream tasks.

Contributions

- Highly robust foundation model that learns complex, generic features from histology slides.
- SOTA performance on diverse and complex clinical tasks: biomarker prediction, mutation status classification, and spatial gene expression.
- Released cutting-edge tool to advance biomedical research, development, and diagnostics, freely available for academic use at: *https://huggingface.co/bioptimus/H-optimus-1*.

Method

Training:

H-Optimus-1 utilizes a 1.1-billion-parameters ViT and is pre-trained using SSL to learn robust, generalized visual representations from unannotated histology slides.

Dataset:

Feature extraction:

Input:
- 224x224 pixel histology tiles.
- MPP=0.5 (20x magnification).

Output:
- 1536- dimensional vectors.
- Used for diverse downstream tasks.

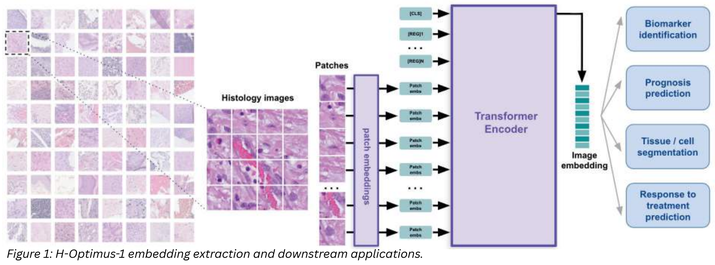

Figure 1: H-Optimus-1 embedding extraction and downstream applications.

Internal evaluations and benchmarks

Downstream evaluation methodology:

- 8 slide-level downstream tasks
- 9 histology FMs benchmarked with ABMIL [6]
- (learning rate, steps) selected via 5-fold cross-validation.

More details about the methodology:
H1 Model Benchmark

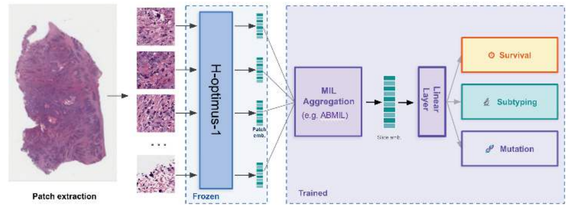

Figure 2: Multiple instance learning (MIL) for Whole Slide Image (WSI) analysis.

		UNI	GPFM	GigaPath	Virchow2	UNI-2H	H-optimus-0	H-optimus-1
Colorectal Cancer	MSI prediction	0.895	0.897	0.933	0.932	0.929	0.922	**0.948**
	KRAS prediction	0.674	0.663	0.637	**0.678**	0.662	0.627	0.641
	BRAF prediction	0.784	**0.796**	0.779	0.784	0.792	0.745	0.791
Breast Cancer	Metastasis detection	0.972	0.968	0.981	0.981	**0.990**	0.976	0.979
	HER2 prediction	0.749	0.727	0.796	0.776	0.799	0.814	**0.825**
	PR prediction	0.769	0.813	0.804	0.819	**0.841**	0.818	0.823
	ER prediction	0.843	0.846	0.835	0.841	0.869	0.844	**0.871**
Gastric Cancer	MSI prediction	0.907	0.892	0.863	**0.923**	0.903	0.899	0.915
	Average performance	0.824	0.825	0.829	0.842	0.848	0.831	**0.849**

Table 1: Performance comparison of H-Optimus-1 against other foundation models (AUC-ROC reported).

Use cases

MIT partners with Bioptimus for Clinical Cancer Prediction Study

Breast Cancer Recurrence Risk Prediction with H-Optimus-1 and STAMP

ICGI researchers build a winning pathology report generation model with H-Optimus-1

H-Optimus-1 for streamlining Spatial Transcriptomics workflows in IBD research

External evaluations

PathBench [1]:

- Includes 22 pathology FMs
- Diverse tasks: classification, grading, survival
- 41k+ cases
- 12 organs
- 26 hospitals

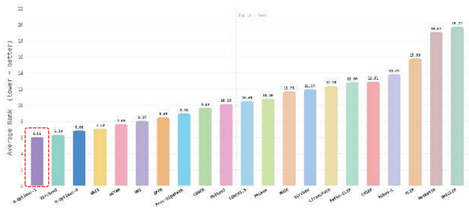

Figure 3: PathBench model ranking (overall).

→ SOTA Results: **H-Optimus-1 ranks first on average** across all evaluated methods and clinical tasks.

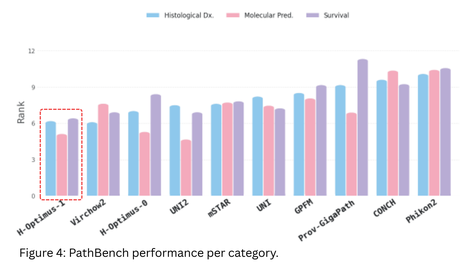

Figure 4: PathBench performance per category.

HEST [2]:

- Prediction of spatial transcriptomics (ST) gene expression signatures directly from H&E.
- Covers 9 organs.

→ **H-Optimus-1 ranks first**, achieving the highest accuracy in capturing localized gene expression patterns.

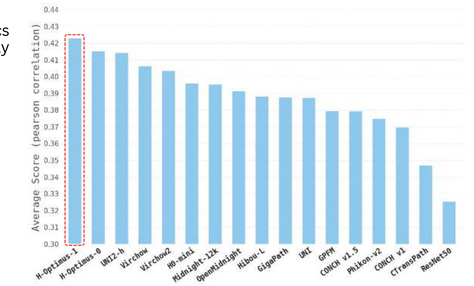

Figure 5: HEST-Benchmark results (03.04.26).

Thunder [3]:

- Tile-only evaluation of 23 foundation models across 16 datasets.
- Analyzes models across three axes: downstream performance, feature space organization, and robustness to noise.

→ **H-Optimus-1 ranks 3/23 in fewshot tasks & 2/23 in linear probing**, but 14/23 in segmentation.

References

[1] Ma, J., et al. PathBench: A Comprehensive Comparison Benchmark for Pathology Foundation Models towards Precision Oncology. arXiv preprint arXiv:2505.20202, May 2025.
[2] Jaume, G., et al. HEST-1k: A Dataset for Spatial Transcriptomics and Histology Image Analysis. Advances in Neural Information Processing Systems, December 2024.
[3] P. Marza et al., "THUNDER: Tile-level Histopathology image UNDERstanding benchmark," in Advances in Neural Information Processing Systems (NeurIPS), 2025.
[4] Zhou, P., et al. iBOT: Image BERT Pre-Training with Online Tokenizer. International Conference on Learning Representations (ICLR), April 2022.
[5] Oquab, M., et al. DINOv2: Learning Robust Visual Features without Supervision. Transactions on Machine Learning Research (TMLR), August 2023.
[6] Ilse, M., et al. Attention-based Deep Multiple Instance Learning. Proceedings of the 35th International Conference on Machine Learning (ICML), July 2018.
[7] Bejnordi, B. E., et al. "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer." JAMA, 318(22), 2199-2210 (2017).
[8] Campanella, G., et al. "Breast Metastases to Axillary Lymph Nodes [Data set]." The Cancer Imaging Archive (TCIA) (2019).
[9] Farahmand, S., et al. "HER2 and trastuzumab treatment response H&E slides with tumor ROI annotations [Data set]." The Cancer Imaging Archive (TCIA) (2022).
[10] Kim, K., et al. "PAIP 2020: Microsatellite instability prediction in colorectal cancer." Medical Image Analysis, 89, 102886 (2023).
[11] The Clinical Proteomic Tumor Analysis Consortium Colon Adenocarcinoma Collection (CPTAC-COAD) [Data set]." The Cancer Imaging Archive (TCIA) (2020).
[12] TCGA Research Network. Data generated by The Cancer Genome Atlas: https://www.cancer.gov/tcga.

Conclusion

H-Optimus-1 is one of the largest foundation models for pathology with over 1 billion parameters, allowing it to capture complex histological patterns and generate powerful embeddings for any use-case leveraging digital pathology.

- Trained on one of the largest & most diverse datasets to ensure robustness across various medical research applications.
- SOTA results across internal and independent external benchmarks.