PLUTO: Pathology-Universal Transformer

Dinkar Juyal*, Harshith Padigela*, Chintan Shah*, Daniel Shenker, Natalia Harguindeguy#, Yi Liu, Blake Martin, Yibo Zhang#, Michael Nercessian#, Miles Markey, Isaac Finberg, Kelsey Luu, Daniel Borders, Syed Ashar Javed, Emma Krause, Raymond Biju , Aashish Sood, Allen Ma, Jackson Nyman, John Shamshoian, Guillaume Chhor, Darpan Sanghavi, Marc Thibault, Limin Yu, Fedaa Najdawi, Jennifer A. Hipp, Darren Fahy, Benjamin Glass, Eric Walk, John Abel, Harsha Pokkalla#, Andrew H. Beck , and Sean Grullon

PathAI, Boston, Massachusetts; * Contibuted equally to study; # Employee of PathAI at the time of study

CONTACT: sean.grullon@pathai.com

PATHOLOGY DATA AND PROBLEM CHARACTERISTICS

- Histopathology is the diagnosis and study of diseases, involving microscopic examination of cells and tissue
- Pathology data exhibits multi-resolution characteristics
- Pathology slides vary in scanner type, stain, organ, and disease severity
- Prediction tasks span slide, tissue, and subcellular levels in pathology
- Diverse applications include classification and segmentation tasks

PLUTO: PATHOLOGY FOUNDATION MODEL DESIGN

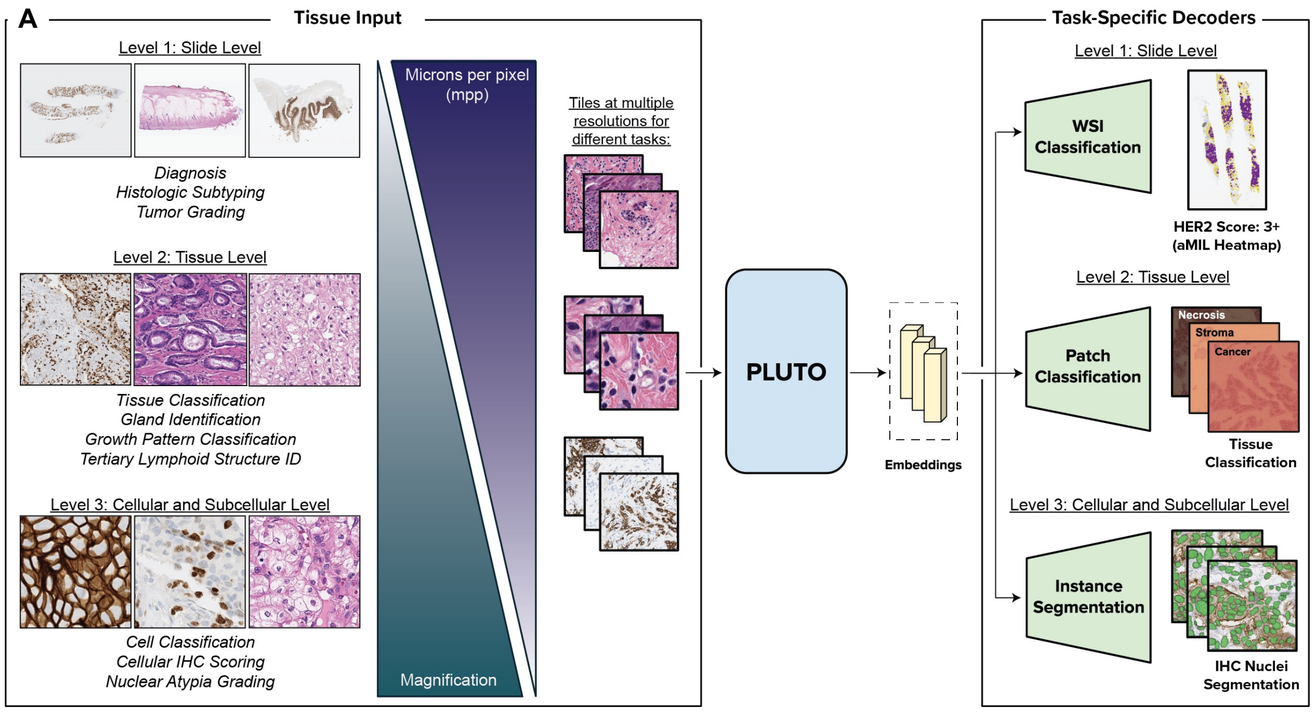

Figure 1. Overview of PLUTO. A) Outline of the PLUTO multi-resolution adaptation pipeline. Tiles are extracted from WSIs at multiple resolutions and correspond to scales that capture different biological contexts. We organize pathology tasks according to these biological contexts as slide level, tissue level, and cellular & subcellular level tasks, respectively. PLUTO generates embeddings that are task-agnostic and can be used in a variety of downstream tasks, where adaptation to WSI-level prediction, tile classification, and instance segmentation are shown. **B**) Detailed PLUTO architecture. WSI tiles at multiple resolutions are masked with varying patch sizes and passed to the backbone for self-supervised pre-training. The architecture is optimized for flexibility across multiple scales and patch sizes. In addition to DINO and iBOT losses, MAE and Fourier losses are applied across varying mask sizes to control the amount of low- and high-frequency information that is preserved.

PRETRAINING DATASET CHARACTERISTICS

Figure 2. Dataset characterization for the pre-training dataset. The distribution of the dataset by organ, disease, stain, scanner, and objective magnification is shown, as well as the distribution of cell point and tissue region annotations which augment the pre-training dataset. The number of biologically-meaningful substances (e.g. lymphocyte, blood vessel, Gleason pattern 3 prostate cancer, tumor bed). The large number of source sites (50+) guarantees large diversity during PLUTO self-supervised pre-training.

REPRESENTATIVE ADAPTATION APPROACH: SEGMENTATION

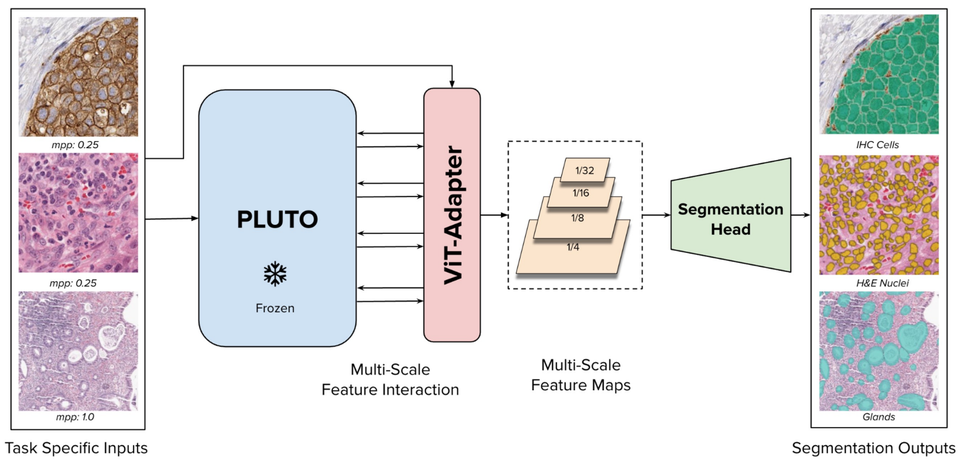

Figure 3. Instance segmentation adaptation with PLUTO. Example task-specific inputs and outputs using the frozen PLUTO backbone. An adapter is used that outputs maps at varying spatial and semantic resolutions, followed by a segmentation head to generate instance segmentation masks. This approach works across object scales from nuclei (top two images) to glands (bottom) and across stain types.

RESULTS

Key Results:

- Multiple instance learning (MIL) models using our frozen PLUTO model as a featurizer tend to outperform models with frozen and fine-tuned CNN backbones, Imagenet-pre-trained ViT backbones, and similar pathology pre-trained ViTs where applicable (Table 1).
- PLUTO achieves strong performance and is competitive with the best performing models, highlighting the effectiveness of diverse pre-training data for enhancing robustness (Figure 4).
- PLUTO achieves state-of-the-art performance on gland segmentation (Figure 4).
- PLUTO is performant, yet small (Figure 4).

Table 1. Performance of multiple instance learning (MIL) models with different ViT- and CNN-based featurizers on NSCLC subtyping and HER2 scoring tasks. The mean and standard deviation across 1,000 bootstrapped runs are reported.

Model	Dataset	Patch Size	Tuning	In-domain F1	In-domain AUROC	OOD F1	OOD AUROC
PLUTO	NSCLC	16	Frozen	**90.2(1.9)**	**94.0(1.6)**	**86.1(2.8)**	**91.2(2.5)**
Meta-DINOv2 ViT-S	NSCLC	14	Frozen	88.6(2.0)	92.0(1.9)	72.1(4.1)	81.9(3.8)
ShuffleNet	NSCLC	-	Frozen	83.6(2.4)	90.1(2.0)	72.2(4.2)	83.5(3.5)
ShuffleNet	NSCLC	-	Fine-tuned	88.1(2.2)	93.9(1.5)	42.5(8.0)	90.8(2.1)

Table 2. Summary of PLUTO performance across public datasets. The tile classification task was performed on CRC-100K and Camelyon17-WILDS datasets by linear probing the PLUTO embeddings while keeping the backbone frozen. The gland segmentation and nuclei segmentation tasks were performed on the GlaS and PanNuke datasets, respectively, by adapting PLUTO through multiple adaptation strategies while keeping the backbone frozen.

Model	Adaptation Head	Benchmark Name	Metrics	
			Acc.	Bal. Acc.
PLUTO	**Linear Head**	**CRC-100K**	**96.6**	**95.3**
ResNet50*	N/A	CRC-100K	94.7	N/A
			Acc.	Bal. Acc.
PLUTO	**Linear Head**	**C17-WILDS**	**96.2**	-
DenseNet-121*	N/A	C17-WILDS	70.3	-
			DICE	IoU
PLUTO	**Mask2Former**	**GlaS**	**91.2**	**84.5**
PLUTO	Mask R-CNN	GlaS	88.0	79.6
UNet*	N/A	GlaS	85.5	74.8
			bPQ	mPQ
PLUTO	**HoverNet**	**PanNuke**	**67.1**	**47.7**
PLUTO	Mask R-CNN	PanNuke	58.6	-
(Shui et al., 2023)*	N/A	PanNuke	55.3	36.9

*Fully Supervised Baseline Model

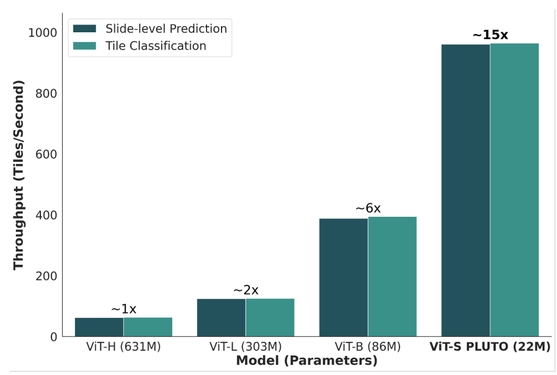

Figure 4. Throughput (tiles/sec) of models for tile-level and slide-level classification tasks. Tasks were performed with various backbones using patch size 16 with a tile size 224 × 224. Linear probes and Additive MIL classifiers were used as as adaptation heads for the tile and slide-level classification tasks, respectively.

KEY TAKEAWAYS

- PLUTO is a lightweight, performant, generalizable foundation model for pathology
- PLUTO is designed to take advantage of the multi-scale nature of WSIs and provide informative representations across biological scales
- PLUTO is trained on a diverse dataset of over 195M image tiles from 50+ distinct sites

ACKNOWLEDGMENTS

- We thank the members of PathAI Pathology Network who contributed towards this work.
- This poster template was developed by SciStories LLC. https://scistories.com/

PathAI