Self-Supervised Learning and Active Learning Pipeline for Cervical Cytology Clustering

Pathology, Levy Lab

Matthew Flowerhill[1], Gerardo Escoto[1], Esther Jin[1], I-Chuang Liao[1], Maya Sarkar[3], Xiaoying Liu[2], Louis Vaickus[2], Joshua Levy[1,2]

1. Cedars Sinai Medical Center, 2. Darmouth-Hitchcock Medical Center, 3. Mission San Jose High School

Abstract

Cervical cancer remains a major public health challenge, particularly in resource-limited settings where limited access to expert cytopathologists can delay diagnosis and reduce screening efficiency. We developed an end-to-end computational pipeline for cervical cytology whole-slide images that integrates YOLO-based cell detection, quality-controlled preprocessing, self-supervised teacher–student Vision Transformer learning (CytoFM), graph-based clustering, and active learning prioritization. Individual cell crops are transformed into high-dimensional feature embeddings that capture morphology and enable unsupervised graph-based community detection of morphologically similar cells, while an active learning framework prioritizes informative cells for expert review. Together, this pipeline provides a scalable workflow for AI-assisted digital pathology by reducing manual review effort and improving the efficiency of expert analysis.

Background

Honduras has one of the world's highest HPV burdens, and cervical cancer is the top cause of cancer death among women aged 15-44 (858 cases, 518 deaths annually).

With few pathologists and centralized care, diagnostic bottlenecks cause an average 232-day treatment delay, while poverty and distance limit screening in rural areas.

Current methods are poorly suited to low-resource settings: Pap smears need complex infrastructure, VIA has high false positives, and HPV tests require slow secondary triage.

Digital cytology can improve access by clustering cells by morphology so experts review prioritized, high-yield findings instead of entire slides. This reduces workload, speeds diagnosis, and supports scalable screening.

This study uses cervical cytology whole-slide images (WSI) from La Liga Medical Center (Honduras), reflecting real-world variability in staining, imaging, and cell density.

References

Ivezić, V., Radhachandran, A., Redekop, E., et al. (2025). CytoFM: The first cytology foundation model. arXiv. https://doi.org/10.48550/arxiv.2504.13402

Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8 (Version 8.0.0) [Computer software]. https://github.com/ultralytics/ultralytics

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Scientific Reports, 9. https://doi.org/10.1038/s41598-019-41695-z

Polański, K., Young, M. D., Miao, Z., Meyer, K. B., Teichmann, S. A., & Park, J.-E. (2019). BBKNN: fast batch alignment of single cell transcriptomes. Bioinformatics, 36, 964-965. https://doi.org/10.1093/bioinformatics/btz625

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: identifying density-based local outliers. Proceedings of the 2000 ACM SIGMOD international conference on Management of data, 93–104. https://doi.org/Q/335191.335388

Reinhard, E., Adhikhmin, M., Gooch, B., & Shirley, P. (2001). Color transfer between images. IEEE Computer Graphics and Applications, 21(5), 34–41. https://doi.org/10.1109/38.946629

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv preprint arXiv:1802.03426. https://doi.org/10.48550/arXiv.1802.03426

Methodology

Segmentation and Cell Detection

- YOLO cell detection and segmentation.
- Converts tile-local predictions into slide coordinates.

Morphology Aware QC and Domain Conditioned Policy Filtering

- Filters detections using morphological quality metrics.
- Removes low-quality or artifact-prone cell crops.

Self-Supervised Teacher-Student Adaption

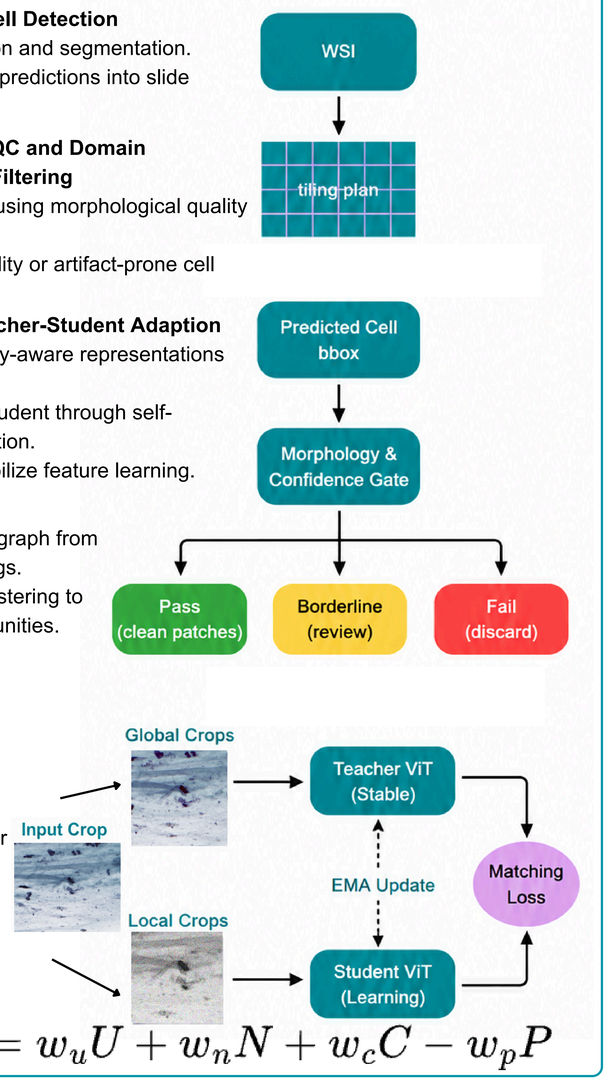
- Learns morphology-aware representations using CytoFM.
- Teacher guides student through self-supervised distillation.
- EMA updates stabilize feature learning.

Clustering

- Constructs a kNN graph from learned embeddings.
- Applies Leiden clustering to identify cell communities.

Active Learning Prioritization

- Scores cells using uncertainty, rarity, and diversity.
- Prioritizes informative cells for expert review.
- Produces a balanced review queue.



$$\text{Priority} = w_u U + w_n N + w_c C - w_p P$$

Visualization/Results

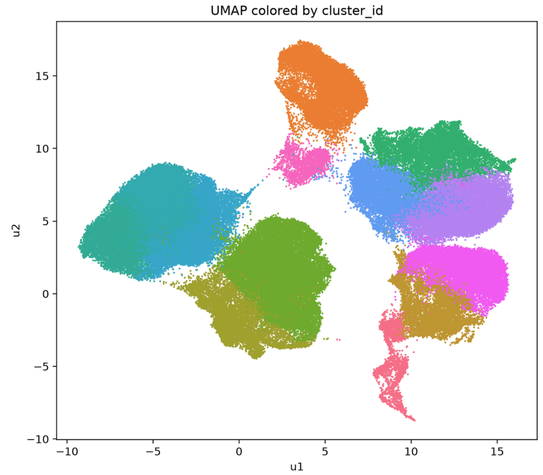


Figure 1:2D visualization of high-dimensional cell embeddings color-coded by graph-clustering assignment (cluster_id). Distinct spatial grouping confirms clear morphological community separation and validates embedding representation quality across the dataset.

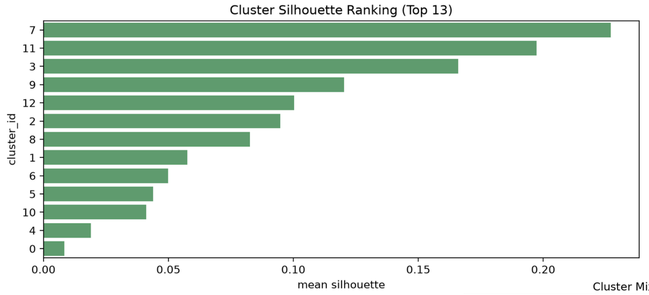


Figure 2: Cluster Separation Quality via Mean Silhouette Scores Ranking of the discovered cell clusters by mean silhouette coefficient, quantifying intra-cluster cohesion versus inter-cluster separability to validate graph clustering performance..

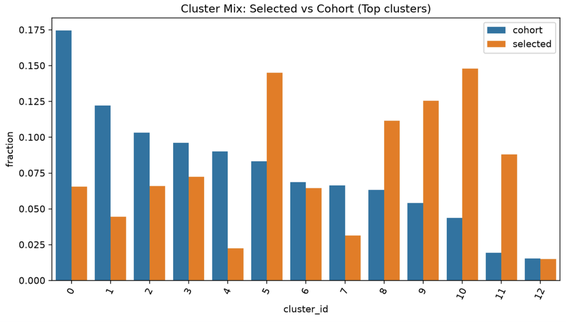


Figure 3: Cluster Mix Comparison (Selected Queue vs. Full Cohort) Comparison of cluster proportions between the baseline dataset cohort (blue) and the Step 8 prioritized review queue (orange). The decision layer down-samples dominant, redundant clusters (e.g., 0–2) and focuses on high-uncertainty or rare clusters (e.g., 5, 8–11) to maximize reviewer information efficiency.

Conclusions

Successfully established a fully deterministic, quality-controlled pipeline for large-scale cervical cytology data ingestion and preprocessing. Leveraged a Vision Transformer pre-trained via self-supervised distillation (CytoFM iBOT) that effectively captures rich cellular morphology without relying on scarce manual annotations. Graph-based community detection (kNN + Leiden) uncovers distinct, biologically meaningful cell phenotypes. Multi-signal priority scoring intelligently shifts review focus toward high-uncertainty and rare cell types, maximizing information yield per minute of expert review. Provides a scalable, reproducible foundation for digital pathology workflows that reduces annotation costs while accelerating diagnostic model development.

Future Work

Possible future work includes, but is not limited to, further fine-tuning the ViT, improving the YOLO model's accuracy in distiguishing between artifacts and cells and identifying cells, and diversifying the dataset with data from other hospitals for generalization.

Acknowledgements

Matthew Flowerhill expresses his sincere gratitude to Dr. Joshua Levy, Esther Jin and rest of Levy Lab at Cedars-Sinai Medical Center for their invaluable mentorship, guidance, and resources. Special thanks are extended to Dr. I-Chuang Liao for supplying the majority of the annotated data and the YOLO model. Additionally, he gratefully acknowledges collaborators from Dartmouth Health and the EDIT Cervical Screening Team for their support in providing the single-cell cervical images.