A Deep Learning Framework for Squamous Cell Segmentation in Whole Slide Images and Feature Extraction for Cervical Cancer Diagnostics

Maya Sarkar, Vivaan Mahajan, Lacie Zhou, Neha Shaik, Doniyor Babajanov, Neha Ranjith

Dartmouth Health

ABSTRACT

Cervical cancer (CC) remains a leading cause of preventable death among women, particularly in low-resource settings where access to early diagnostic tools is limited. Existing screening techniques such as Pap smears, colposcopy, and HPV testing require trained professionals and are often inaccessible in rural regions. In this study, a deep learning approach was trained and evaluated on a dataset of 1,038 whole slide images (WSIs) to enhance CC screening for automated cytopathological analysis. Using the YOLOv8 model, squamous cells were precisely segmented through high-resolution segmentation masks, which were then processed by a foundation model to generate 512-dimensional feature embeddings (FEs) from the second to last layer. These embeddings served as rich descriptors for downstream diagnostic tasks like clustering squamous cells based on morphological features. With the help of GPU acceleration and batch processing, our methodology was optimized for efficiency in order to facilitate the high-throughput analysis necessary for large WSI datasets. As a result of this work, a significant step towards augmenting the capabilities of pathologists has been made, improving the efficiency of CC screening and the potential to support equitable CC screening worldwide.

INTRODUCTION

CC is a major global health concern, ranking as the fourth leading cause of cancer-related death among women. It develops from precancerous changes in the cervical epithelium, with squamous cell carcinoma being the most common type and adenocarcinoma also contributing a significant portion of cases.

Early detection is critical for survival, as CC is often asymptomatic in its early stages. Screening methods like Pap smears and HPV testing have reduced mortality in high-income countries, but many low-resource regions, including Honduras, lack the proper infrastructure and trained personnel needed to implement regular screening programs.

Digital pathology and deep learning provide a promising solution, enabling automated analysis of digitized slides through Whole Slide Imaging (WSI). Using models like YOLOv8, this research develops a pipeline to accurately segment and analyze cervical cells, supporting early diagnosis and potentially improving screening access in underserved areas.

METHODS

- **YOLOv8 for Squamous Cell Segmentation:** The cornerstone of our approach is the use of the YOLOv8 model for the precise segmentation of squamous cells within WSIs. YOLOv8 is a highly efficient and accurate object detection and instance segmentation model. We have specifically trained this model on a large dataset of WSIs with detailed annotations of squamous cells. This training enables the model to learn the intricate morphological features of these cells and accurately delineate their boundaries with both bounding boxes and segmentation masks (polygons).
- **Foundation Model for Feature Extraction:** Following the segmentation of squamous cells by YOLOv8, the next critical step is to extract meaningful features from these individual cell images. For this purpose, we employ a large-scale, pre-trained foundation model. The segmented cell images, isolated using the bounding boxes generated by YOLOv8, are fed into this foundation model. Instead of using the final classification output of the foundation model, we tapped into its deeper layers to extract a rich feature representation of each cell. Specifically, we extracted the 512-dimensional embedding from the second-to-last layer of the model. This embedding serves as a dense feature vector that encapsulates the complex morphological characteristics of the squamous cell, such as nuclear size and shape, chromatin texture, and cytoplasmic properties. These high-dimensional features are then utilized for downstream diagnostic tasks.

RESULTS

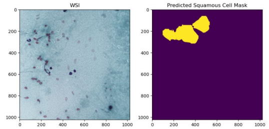

Figure 1. *Visualize YOLO Predicted masks for a 1024 by 1024 image patch*

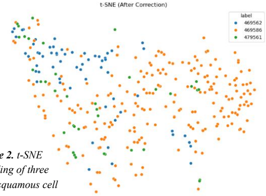

Figure 2. *t-SNE embedding of three separate squamous cell images*

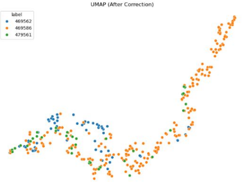

Figure 3. *UMAP Embedding of three Squamous Cell images*

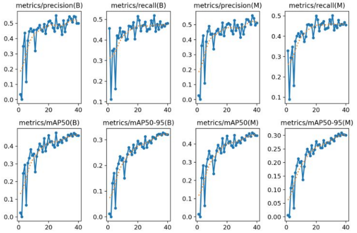

Figure 4. *Validation metric plots after every training epoch for 40 epoch*

RESULTS

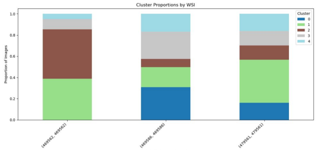

Figure 5. *The proportion of cell images for every cluster per WSI*

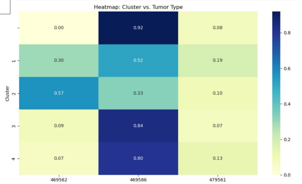

Figure 6. *Heatmap detailing the correlation between the clusters and WSI images*

- The YOLOv8 model achieved a high level of accuracy in both detecting and segmenting squamous cells. The training and validation loss curves show a consistent decrease over the epochs, indicating that the model was effectively learning the features of the target cells. The metrics for precision, recall, and mean Average Precision (mAP) on the validation set were strong, confirming the model's ability to generalize to unseen data.
- The FEs extracted from the segmented squamous cell images using the foundation model provided a rich representation of cellular morphology. To explore the utility of these features for downstream diagnostics, we performed clustering analysis on the embeddings, the results of this analysis revealing distinct clusters of cells. This demonstrates the potential of our approach to not only automate the detection of squamous cells but also to distinguish them based on their morphological characteristics, which could be invaluable for early cancer detection and diagnosis. The UMAP and t-SNE plots of the cell embeddings show clear separation between different WSI images, suggesting that the extracted features capture meaningful biological information.

CONCLUSION

This research has successfully demonstrated the development and application of a deep learning-based framework for the automated segmentation of squamous cells in CC WSIs and the subsequent extraction of high-dimensional feature embeddings. Our use of the YOLOv8 model, accelerated by GPU computing, provides an efficient and accurate solution for this challenging task. The insights gained from the analysis of the extracted cell features highlight the potential of our approach to significantly impact downstream diagnostic applications, from early cancer detection to the nuanced analysis of tumor morphology. Current methods such as Pap smears and VIA are often inconsistent and inaccessible, leading to late-stage diagnoses and higher mortality rates. This research provides a scalable and cost-effective alternative that reduces reliance on scarce specialists to strengthen Honduras' screening system by improving early detection, enabling prompt clinical decision-making, and ultimately reducing preventable CC deaths.