Automating Rapid On-Site Evaluation (ROSE) for Breast Fine Needle Aspiration Cytology in Sub-Saharan Africa

Krishna Tejaswini Sathi, Kim Hwang Yeo, Atul Zacharias, Krrish Thakkar

Johns Hopkins University | Whiting School of Engineering | Baltimore, MD

INTRODUCTION

Breast Cancer is the world's most diagnosed cancer, with incidence and mortality are rising at the steepest rate in Sub-Saharan Africa (SSA). 70% of breast cancer in SSA is diagnosed when the cancer progresses to late stages leading to 1 in 2 women succumbing to the disease.

Low accessibility to pathology services in SSA delays diagnosis

1 pathologist for every 2-5 million people

Diagnostic turn-around-time (TAT) is further prolonged when biopsy samples acquired at peripheral health centers are deemed inadequate for interpretation after transport over long distances to an urban center for interpretation.

Diagnostic delays prolonged to 11 months

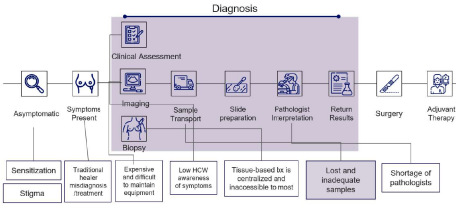

Figure 1— Challenges in breast cancer management in Uganda

Rapid on-site evaluation (ROSE) improves diagnostic yield but is not practiced in resource-limited environments. ROSE in breast cytology currently consists of subjective evaluation of cellularity. Literature shows that counting cell clusters in a cytology sample is a quantitative method of sample adequacy determination and reduces false-negative diagnoses.

OBJECTIVE

In this study, we propose and create a proof of concept of using machine learning (ML) to determine the sample adequacy of breast fine needle aspiration cytology (FNAC) in a resource-limited setting.

MATERIALS AND METHODS

A single de-identified breast FNAC slide was scanned using a whole slide imager. The resulting image was fragmented using a Python script into over 900 1024 X 1024 px images representing a light microscope's 16X magnification field of view.

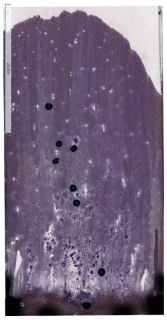

Figure 2A— Breast FNAC whole slide image

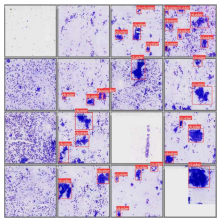

Figure 2B— 1024 X 1024 pixel patches of the whole slide image in Figure 1A

These images were manually annotated by creating bounding boxes around epithelial cell fragments on Roboflow and formed the training dataset of the machine learning model.

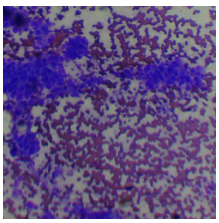

Figure 3A—Unannotated breast FNAC patch

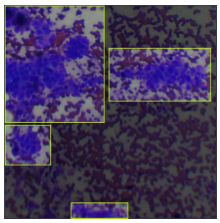

Figure 3B—Annotated breast FNAC patch

The convolutional neural network, You Only Learn Once (YOLO) v7 was used to train the annotated images on Roboflow. This resulted in test images with object detection bounding box predictions as shown in Figure 2B.

RESULTS

The model developed from 900 breast FNAC image fragments yielded a mean average precision (mAP) of 67%. A HAYEAR digital camera was installed on a light microscope and a one-minute video of a breast FNA slide moved along the microscope stage was recorded.

The generated model from this study was run on this video, resulting in bounding boxes overlaid on the video, successfully identifying epithelial cell fragments through a standard microscope.

Watch the machine learning model at work

CONCLUSION

The ML model developed in this study demonstrates technical feasibility of automated ROSE on a light microscope using artificial intelligence. Since this study, cluster detection model performance improved to a mAP of 93% and is currently being translated into a mobile application by the authors and collaborators in Uganda. This interrvention can improve the number and quality of slides reviewed by pathologists and reduce TAT in LMICs.

ACKNOWLEDGEMENTS

Mulago National Referral Hospital
Republic of Uganda
American Cancer Society
HOLOGIC®