A digital Pathology Platform for Automated Diagnosis and Classification of Cervical Cancer from pap-smear images (DigPath).

Wasswa William, Andrew Ware, Annabella Habinka Basaza-Ejiri, Johnes Obungoloch

Background: Globally, cervical cancer ranks as the fourth most frequent cancer in women with an estimated 570,000 new cases in 2018, representing 6.6% of all female cancers. Approximately 90% of deaths from cervical cancer occur in low and middle income countries with Uganda ranked 7th among the countries with highest incidences of cervical cancer in Africa. However, cervical cancer can be prevented through regular screening.
Objective: To develop a digital pathology platform for automated diagnosis and classification of cervical cancer from pap-smear images.
Methods: The digital pathology platform consists of three entities: (1) An automated low-cost digital microscope slide scanner producing quick, reliable and high-resolution cervical cell images from pap-smears. It is 3D printed and costs less than 500USD compared to the commercial-microscopes costing over 2000USD. It has an objective-lens from traditional microscopes; a Logitech web-camera for image capture placed at a distance from the lens calculated using the 4f principle, a motorized stage driven by 2 stepper motors for XY-movement and a third motor for focusing (Z-movement). Image capture is by a developed software in c++. Auto focusing is by an algorithm based on Fast-Fourier transform. The stage control is accomplished using the grbl library as used in the cnc machines to provide precise movements. Limit switches are used for position sensing, (2) An automated pap-smear analysis tool for diagnosis and classification of cervical cancer from pap-smear images. For image analysis pipeline, scene segmentation is achieved using a sequential elimination approach, image segmentation is achieved using trainable classifier, feature selection is with Simulated annealing coupled with a wrapper filter and classification is based on an enhanced fuzzy c-means algorithm and (3) Cervical cancer risk factors evaluation to automatically assess the likelihood of contacting cervical cancer given the risk factors implemented using Mamdani fuzzy logic based on the knowledge base provided by experts.
Results: The evaluation of the automated pap-smear analysis was carried out on three different datasets (single cell images, multiple cell images and pap-smear slide images from a pathology unit. Overall classification accuracy, sensitivity and specificity of '98.88%, 99.28% and 97.47%', '97.64%, 98.08% and 97.16%' and '95.00%, 100% and 90.00%' were obtained for each dataset respectively. The evaluation and testing conducted confirmed the rationale of the proposed approach that the selection of good features embeds sufficient discriminatory information that can increase the accuracy of cervical cancer classification. Evaluation of the fuzzy inference rules for cervical cancer risk assessment showed that the diagnosis produced by the tool were in agreement with the diagnosis from the cytopathologist. Evaluation of the microscope slide scanner showed that the developed microscope can provide high-resolution digital images. For a conventional pap-smear slide, the developed microscope can produce a digital image in less than three minutes with resolutions of 1.10µm and 0.42µm using a 10x lens and a 40x lens respectively. The image quality is comparable to high-end commercial microscopes at a cost of less than $500.
Conclusion: The major contribution of this platform in cervical cancer screening is that it reduces on the time required by the cytotechnician to screen very many slides by eliminating the obvious normal ones, hence more time can be put on the suspicious slides. The proposed

platform has the capability of analyzing a full-slide in less than four minutes as opposed to 5-10 minutes per slide in the manual analysis.

Recommendations: Despite the high performance of the approach, it, however, uses a chain of number of methods which makes it computationally heavy and this is a limitation of the proposed method. In the future deep learning, approaches will be explored to reduce the complexity of the approach for clinical use.

Funding: This work was funded by the African Development Bank- HEST project and the Commonwealth Scholarship Commission.