Classification of Cervical Cancer Tissues Using a Novel Low Cost Methodology for Effective Screening in Rural Settings

Suleiman Mustafa[#1], Steve Adeshina[*2], Mohammed Dauda[#3], Wole Soboyejo[##4]

[#]*National Agency for Science and Engineering Infrastructure (NASENI), Idu Industrial Area, PMB 391 Garki, Abuja, Nigeria*

[1]mustafa.suleiman@naseni.org, [3]mdsmatth@gmail.com

[*]*Nigerian Turkish Nile University Plot 681, Cadastral Zone C-OO, Research & Institution Area, Abuja, Nigeria*

[2]steve.adeshina@ntnu.edu.ng

[##]*African University of Science and Technology (AUST)/Princeton University Galadimawa Airport Road, PMB 681 Garki, Abuja, Nigeria*

[4]soboyejo@aol.com

Abstract— **Cervical cancer is the second highest killer of women in Sub-Sahara Africa. This is due to unavailability of adequate screening methods and inability to pay where existing methods are available. This work addresses the issue of identifying and classifying cervical cancerous tissues for images taken from a standard camera. A cost effective method is developed for the imaging of cervical cancer in developing countries by analysing images from a standard camera. The image is enhanced using noise reduction and a Canny edge detector algorithm to find the intensity of the edges. Preliminary results obtained from clinical settings and images obtained are analysed by identifying the edges of the tissues of the cervix and classifying them based on the frequency of the edges. The results show above 90% classification accuracy. This shows that image analysis algorithm has the potential to successfully diagnose cervix cancer.**

Keywords — **Canny edge detection, Cervical cancer, Classification, Cost-effectiveness, Image analysis.**

I. INTRODUCTION

Cervical cancer is known to be caused by the Human Pappiloma Virus (HPV). Cervical cancer is among the most common cancer affecting women worldwide with 528 000 new cases each year and 266 000 deaths in 2012. It is most notable in sub-Sahara Africa where there is lack of resources and awareness. In this region, 34.8 new cases of cervical cancer are diagnosed per 100 000 women annually, and 22.5 per 100 000 women die from the disease. These figures compare with 6.6 and 2.5 per 100 000 women, respectively, in North America. The drastic differences can be explained by lack of access to effective screening, early detection and treatment [1].

HPV is transmitted during sexual intercourse with an infected partner. Its incidence in Nigeria/Africa is on the rise due to poor screening. In contrast the incidence of HPV is declining rapidly in Western countries due to screening. In the developed countries the availability of resources has reduced the rate of cervical cancer through cytological screening for many years [2]. However, the colposcopy systems in the Western countries are expensive due to the optics and software [3]. While the Pap smear screening is the common choice for cervical screening, it is not always practical to administer, particularly in areas where the test is not available or is out of the price range of the patient.

Unaided visual inspection of cervix treated with 3-5% acetic acid is a common clinical approach. The acetic acid test also covers a wider area around the cervix and requires less training, equipment and infrastructure, not to mention fewer specialized medical personnel. This test works in the remotest of settings, bringing access to virtually everyone [4].

Aided visual inspection of the acetic acid treated cervix involves the use of a small, lightweight, low-powered (2-4X), monocular telescope to view acetic acid treated tissues. This technique is currently being evaluated in several Asian countries, but its sensitivity, specificity and any advantage that it confers in comparison with unaided visual inspection is yet to be determined.

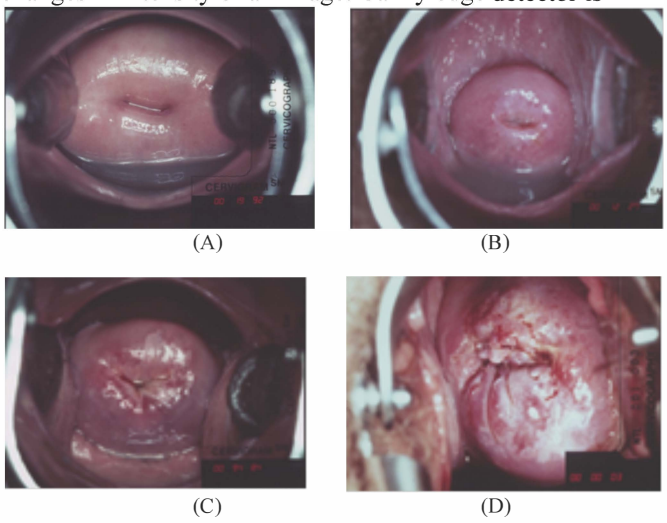
Low-tech and inexpensive screening tools exist and could significantly reduce the burden of cervical cancer deaths right now in less developed countries [1]. One of the major problems is early detection but with the use of innovative imaging modality this can be solved. There is therefore a need for low cost cervical cancer screening systems beyond subjective observation methods that can be used in the developing countries.

Recently, optical techniques have been investigated as an alternative detection method in a quantitative and objective manner. A multispectral digital colposcopy (MDC) was developed to acquire reflectance images of the entire cervix with white-light illumination [5].

Cervical screening procedures and the data collection confirm that HIV-infected women are at higher risk of presenting a positive cervical screening test as well as histological precancerous lesions compared to HIV negative women, highlighting the need to extend cervical cancer screening to all HIV care clinics in West Africa [3].

978-1-4799-4106-3/14/$31.00 © 2014 IEEE

Another research work for low cost cancer detection made use of the default edge detector features on ImageJ to Identify edges of tissue staining, categorize edge morphology, compare mean-grey values and categorize lesion [4]. From the online ImageJ manual the default edge detector used for that work is the Sobel operator. It is applied to highlight sharp changes in intensity of an image. Canny edge detector is



(A) (B) (C) (D)

Fig. 1 Images of Normal cervixes, slightly abnormal with early cancer and infiltrating cancer

known to be the optimal detector, which minimizes the probability of false positives as well as false negatives [6]. In this work, we aim to apply the Canny edge detector for the purpose of improving detection of cancerous tissues from normal tissues using standard low-resolution photographs.

II. METHODOLOGY

A. Dataset

The images used for this work are standard cervix images sourced from an online database hosted by the World Health Organization (WHO) [1]. A total of 15 images were used with resolutions of 696 x 490 pixels.

B. Image Enhancement

The sample regions to be examined were cut from the above images. To remove glare and reflection from the images the following features from the ImageJ software was applied:

I. Despeckle – To reduce the random distribution of light when it is scattered by a rough surface.
II. Noise reduction – To reduce noise.
III. Smoothen - Blurs the active image or selection.

[1] Cervix images that are used as a diagnosis and staging teaching module were sourced from an online database by the World Health Organization (WHO) available publicly at: www.gfmer.ch/Books/Cervical_cancer_modules/Aided_visual_inspection_atlas.htm

C. Edge Detection

Edge detection is a technique used to determine the presence of an edge or line in an image where the image brightness changes sharply or discontinuities and outlines them appropriately [7]. With the application of this technique we can identify the edges on cancer tissues and differentiate from normal tissues from the 2D images. Edges in an image are those points, which show sudden change of intensity, and with help of derivatives (1st order, 2nd order) we can find this change in an image [8]

Canny edge detector is a complex but accurate detector and it is based on the following 3 objectives:

I. Good Detection: The optimal detector must minimize the probability of false positives as well as false negatives.
II. Good Localization: The edges detected must be as close as possible to the true edges.
III. Single Response Constraint: The detector must return one point only for each edge point. [6]

The Canny Edge Detector Algorithm:

The procedures:

1. *Smoothen the image to reduce noise.*
   The image is smoothened first using a Gaussian filter with a specific standard deviation σ, to reduce noise [10].

2. *Compute intensity gradient.*
   The edges are marked where the gradients of the image has large magnitude. The algorithm finds where the grayscale intensity of the image changes the most [11]. Where (G_x) and (G_y) are values in horizontal and vertical direction respectively [6].

$$g(x,y) = [G_x^2 + G_y^2]^{1/2} \quad (1)$$

3. *Non-maximum suppression.* Calculate the edge direction. For each pixel compute the orientation of intensity gradient vector:

$$\theta = \tan^{-1}(G_y / G_x) \quad (2)$$

Digitize the edge direction. Once the edge direction is known, approximate it to an edge direction that can be traced in a digital image [11]. Transform angle θ to one of four directions: 0, 45, 90, 135 degrees.

Compute new array *N*: if

$$G(p_a) < G(p) < (p_b) \quad (3)$$

Where p is the current pixel, p_a and p_b are the two neighboring pixels in the direction of gradient, then $N(p) = G(p)$, otherwise $N(p) = 0$. Nonzero pixels in resulting array correspond to local maxima of G in direction $\theta(p)$.

4. *Tracing edges with hysteresis.* After completing the previous steps, the final step is to track along the remaining pixels that have not been suppressed and threshold the image to find the edge pixels [11] .At this stage two thresholds for the values of *G* are introduced:

Tmin and *Tmax*. Starting from pixels with $N(p) \geq T_{\max}$ find all paths of pixels with $N(p) \geq T_{\min}$ and put them to the resulting image [6].

III. EXPERIMENT RESULTS

Images were sourced from online database obtained from cervix cancer screening. The procedure involves positioning the patient in lithotomy position or supine with legs bent at knees. Good visualization is essential with directed light source to the genital area to get a clear view of the cervix. Wash the cervix with the acetic acid (3-5%) with the help of the syringe. Wait for approximately 1 minute. Inspect the cervix for aceto-white area(s). Using a low-resolution camera and the light source the image of the cervix is captured [4].

The images are loaded into ImageJ. Image enhancement is done using ImageJ features of 'despeckle', 'outlier removal' by applying a pixel radius of 25, threshold of 20 and a bright outlier, then finally smoothing for noise reduction. This smoothing filter replaces each pixel with the average of its 3 × 3 neighbourhood [10]. The values of pixel radius and threshold for outlier removal were obtained from experimentations carried out to determine suitable image enhancement.

The canny edge boundary plug-in feature is used to implement the Canny algorithm applied to find the edges. The plug-in is a Java implementation of the algorithm procedure introduced above with a pixel radius of 1.5, low threshold of 2.5 and high threshold of 5.0 for the smoothing. Experimentations were carried out to determine the most suitable pixel radius, low and high thresholds for the edge detection for the set of sampled images.

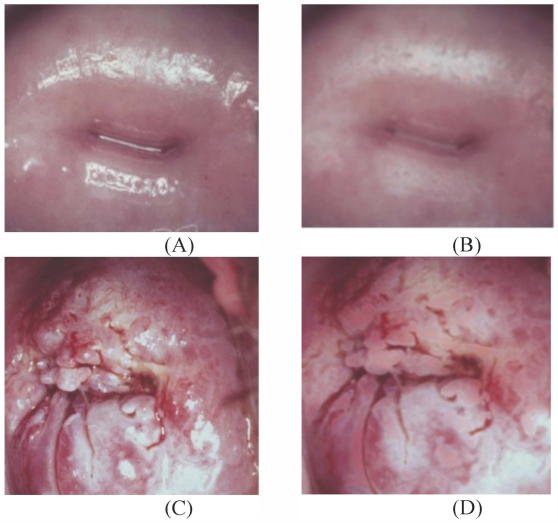
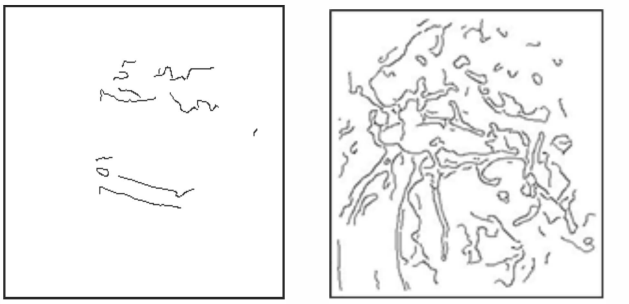


Fig. 2 Sample regions of images and the enhanced images

The above images Fig. 2, A and Fig. C show examples of the sampled sections of the images while B and D the image result after enhancement.



(A) Edge detection result image A (B) Edge detection result image H

Fig. 3 shows samples of the edge detection results.

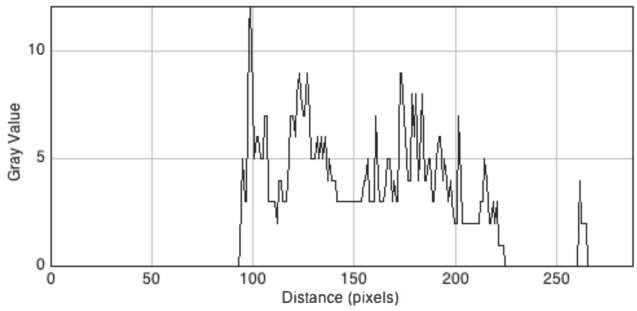


Fig. 4

Fig. 4 graph shows the intensity of the grey values for edge result (A) above in Fig 3.

To quantify the intensity of the edges the mean grey values is measured. For RGB images (colour images), the mean is calculated by converting each pixel to grey scale using the formula: Grey=0.299 * red + 0.587 * green + 0.114 * blue [9]. The images from the edge detection results were already in grey scale. Therefore by analysing the plot profile of the grey values against distance (pixels) and then calculating the mean grey value within the image. This is the sum of the grey values of all the pixels in the image divided by the number of pixels.

This enabled us to quantify the frequency of edges for classification as shown in the table below.

TABLE; ICANNY EDGE DETECTION RESULTS

Images	**Acetic Acid Test Visual Diagnosis**	**Mean Grey**	**StdDev**	**Area**
A	No medical intervention required.	1.895	2.521	78175
B	No medical intervention required.	2.427	2.572	140910
C	No medical intervention required.	2.693	3.224	84552
D	No medical intervention required.	2.166	3.2048	101738
E	No medical intervention required.	0.680	1.7188	67896
F	Refer the patient to Clinic.	7.189	8.674	154344
G	Refer the patient to Clinic.	6.555	6.702	92736
H	Refer the patient to Clinic.	4.626	5.859	103515
I	Refer the patient to Clinic.	11.09	12.128	76544
J	Refer the patient to Clinic.	7.547	8.44	116974
K	Refer the patient to Clinic.	12.409	13.447	96030
L	Refer the patient to	14.215	17.225	95040

	Oncology Centre.			
M	Refer the patient to Clinic.	7.226	8.822	53572
N	Refer the patient to Clinic.	4.204	5.2428	89323
O	Refer the patient to Oncology Centre.	6.864	7.586	106812

Table I above shows the Acetic Acid Test (AAT) visual result from the online data compared with the mean-grey values, the standard deviation (StdDev) and the Area of the sample regions.

The normal cervixes in images **A, B**, **C, D** and **E** have relatively lower mean values due to fewer edges. These results match with the doctor's visual observations where no treatment is required. Images **F, G, H, J, M, N and O** all have more edges and higher mean-grey values. Although from the acetic acid test **F** actually identified as a false positive. Images **I, K** and **L** have the highest mean values from the samples taken. L with the highest value is Infiltrating cancer which was referred to the Oncology centre after AAT.

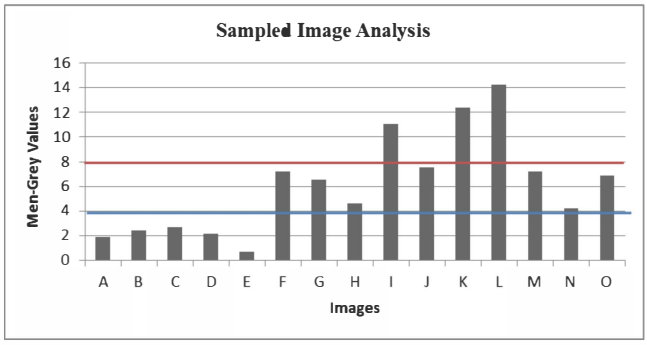


Fig. 5 Bar Chart of images and their mean-grey values.

From the above comparison we set a minimum classification of 4 and a maximum of 8, which indicates that the mean-grey values below 4 are normal, above 4, but below 8 are slightly abnormal and those above 8 highly abnormal from the images sampled **A-O**. We have successfully been able to identify normal cervixes and highly abnormal ones.

IV. DISCUSSION AND CONCLUSION

Medical imaging has become a major tool in clinical trials since it enables rapid diagnosis with visualization and quantitative assessment. It can be applied for low cost medical diagnosis to improve detection particularly in a case like cervical detection that can be tedious due to the process and is susceptible to human error. For cases with no significant indication of abnormal tissues and where no medical treatment is required this experiment will return few edges and low mean-grey value indicating healthy cervix. On the other hand in cases of invasive advanced cancers with proliferating, bulging, mushroom- or cauliflower-like growth, this will identify multiple edges and higher mean grey value therefore patients will require immediate attention. For cases in-between there are still challenges such as glare from the acid pool, camera reflection, and darkness from blood etc. that can affect the results.

Further research work is recommended with more image samples and other image processing techniques such as segmentation. The preliminary results suggest that low-resolution images from images can be used to image cervical cancers. This study is limited by the number of images used. Local efforts are being made to collect more images while observing all ethical issues. Furthermore, knowledge based approach will be applied to further investigate the subject. In addition a mobile-based application will be developed to translate this solution to be used in rural areas.

ACKNOWLEDGMENT

We wish to acknowledge Shivani Sud for her contribution to the development and preparation of this work. We wish to also acknowledge the staff and students of the Computational Biomedical Imaging and Modelling Centre (CBIM), Rutgers University, New Jersey for their support.

REFERENCES

[1] F. J. Soerjomataram et al, *Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11*, Lyon, France: International Agency for Research on Cancer, 2013. [Online]. Available: http://globocan.iarc.fr

[2] H.O Adami HO, J. Ponten, P. Sparen, R. Bergstrom, L. Gustafsson, L.G Friberg, *Survival trend after invasive cervical cancer diagnosis in Sweden before and after cytologic screening*. 1960-1984. Cancer 1994, 73(1):140-147

[3] A. Horo et al, *Cervical cancer screening by visual inspection in Côte d'Ivoire, operational and clinical aspects according to HIV status*. BMC Public Health 2012 12:237.

[4] S. Sud, *Women's Cancers in Rural India: Understanding limitations to health care access and the proposal of a portable imaging system for cervical cancer diagnosis*: Cervical Screening Component, 2013.

[5] S. Y. Park et al, *Automated image analysis of digital colposcopy for the detection of cervical neoplasia*, Journal of Biomedical Optics 13(1), 014029, January 2008

[6] J. Canny, *A Computational Approach to Edge detection IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. Pami-8, no. 6, November 1986

[7] W. Frei and C. Chen, "*Fast Boundary Detection: A Generalization and New Algorithm*," IEEE Trans. Computers, vol. C-26, no. 10, pages. 988-998, Oct. 1977.

[8] T. Lindeberg "*Edge detection and ridge detection with automatic scale selection*", International Journal of Computer Vision, 30, 2, pages 117—154, 1998.

[9] E. D. Selefichi1, O. G. Duliu, *Image Processing And Data Analysis In Computed Tomography*, Romanian. Journal. Physics. vol. 52, nos. 5–7, pages. 667–675, Bucharest, 2007

[10] R. C Gonzalez, R. E Woods and S.L Eddins, *Digital image Processing Using Matlab, Second Edition,* Gatesmark Publishing, USA, 2009.

[11] C. Solomon and T. Breckon, *Fundamentals of Digital Image processing: A practical approach with examples in Matlab,* Wiley-Blackwell, UK, 2011.

[12] F. Cui and L. Zou, "*Edge Feature Extraction Based on Digital Image Processing Techniques"* Proceedings of the IEEE International Conference on Automation and Logistics Qingdao, China September 2008.