An Artificial Intelligence Enabled Multimedia Tool for Rapid Screening of Cervical Cancer

Kumar Dron Shrivastav[a], Neha Taneja[b], Priyadarshini Arambam[c], Vandana Bhatia[d], Shelly Batra[c], Harpreet Singh[e], Eyad H. Abed[f], Priya Ranjan[g], Rajiv Janardhanan*[h]

[a]*Laboratory of Health Data Analytics and Visualization Environment, Amity Institute of Public Health, Amity University, Noida, India*
[b]*Laboratory of Disease Dynamics and Molecular Epidemiology Amity Institute of Public Health, Amity University, Noida, India*
[c]*Batra Hospital and Medical Research Centre, 1, Tughlakabad Institutional Area, Mehrauli Badarpur Road, New Delhi-110 062 (Near Saket Metro Station)*
[d]*Department of Computer Science and Engineering, Amity School of Engineering and Technology, AUUP, NOIDA, UP-201303, India*
[e]*Indian Council of Medical Research, V. Ramalingaswami Bhawan, P.O. Box No. 4911 Ansari Nagar, New Delhi - 110029, India*
[f]*Institute for Systems Research, University of Maryland, College Park, MD-20741, USA*
[g]*Department of Electronics and Communication Engineering, SRM University, Mangalagiri Neerukonda Tadikonda Rd, Mangalagiri, Mandal, Andhra Pradesh-522502, India*
[h]*Laboratory of Disease Dynamics and Molecular Epidemiology and Laboratory of Health Data Analytics and Visualization Environment Amity Institute of Public Health, Amity University, Noida, India*

Abstract

Cervical cancer is a major public health challenge. Further mitigation of cervical cancer can greatly benefit from development of innovative and disruptive technologies for its rapid screening and early detection. The primary objective of this study is to contribute to this aim through large scale screening by development of Artificial Intelligence enabled Intelligent Systems as they can support human cancer experts in making more precise and timely diagnosis. Our current study is focused on development of a robust and interactive algorithm for analysis of colposcope-derived images analysis and a diagnostic tool/scale namely the OM- The Onco-Meter. This tool was trained and tested on 300 In-

Email addresses: kdshrivastav@amity.edu (Kumar Dron Shrivastav), ntaneja@amity.edu (Neha Taneja), priyadarshini@batrahospitaldelhi.org (Priyadarshini Arambam), vbhatia2@amity.edu (Vandana Bhatia), shelly.batra@opasha.org (Shelly Batra), crbhmrc1@batrahospitaldelhi.org (Shelly Batra), hsingh@bmi.icmr.org.in (Harpreet Singh), abed@isr.umd.edu (Eyad H. Abed), ranjan.p@srmap.edu.in (Priya Ranjan), rjanardhanan@amity.edu (Rajiv Janardhanan*)

Preprint submitted to The Lancet Digital Health *June 1, 2020*

dian subjects/patients yielding 77% accuracy with a sensitivity of 83.56% and a specificity of 59.25%. OM-The Oncometer is capable of classifying cervigrams into cervical dysplasia, carcinoma $in-situ$ (CIS) and invasive cancer(IC). Programming language - R has been used to implement and compute earth mover distances (EMD) to characterize different diseases labels associated with cervical cancer, computationally. Deployment of automated tools will facilitate early diagnosis in a noninvasive manner leading to a timely clinical intervention for cervical cancer patients upon detection at a Primary Health Care (PHC).The tool developed in this study will aid clinicians to design timely intervention strategies aimed at improving the clinical prognosis of patients.

Keywords: Artificial Intelligence, Cervical Cancer, Cervigrams, Colposcopy, Early Detection

1. Introduction

Cervical cancer represents a major unmet clinical need and is a serious public health problem. Early detection is not widely available in developing countries, and standard detection method(Pap test) are not sufficiently sensitive or specific [1]. According to World Health Organization(WHO) estimates, cervical cancer is among the highest morbidity and mortality causing cancer, with 570,000 new cases annually. Report by WHO showed that 90% of cervical cancer affected women died in the recent past especially in Low and Middle Income Countries (LMICs) [2, 3].

India with a diverse genetic base and a resource-limited healthcare infrastructure, has a disproportionate burden and challenging disease pattern of cervical cancer, necessitating the need to develop robust diagnostic tools to facilitate large-scale screening in order to alleviate the clinical prognosis. The conventional approach for cervical cancer diagnosis is time-consuming, expert oriented, and requires a specific set of resources leading to misdiagnosis and missed diagnosis resulting into frequent false negatives and false positives [4].

Approximately 453 million women of Indian origin above the age of 15 years are at the risk of being afflicted by cervical cancer. Five percent of these Indian women are estimated to be infected by an Human Papilloma Virus (HPV) serotype (HPV-16/18) at any given instant, and an astounding 83.2% of invasive cervical carcinoma lesions are known to have either HPV 16 or 18 serotype [5].

The mortality rate from cervical cancer could be significantly reduced through large scale screening at both regional and global levels. A recent report suggested that unimodal screening of large scale patients using HPV vaccination is indeed insufficient to manage the burden of cervical cancer in a country like India endowed with wide genetic base and diverse geological relief structures [6].

Cervical cancer, a major challenge for public health professionals and clinicians, requires a reliable and stable technology to enable large-scale rapid screening to reduce its burden. This formed the premise for the development of an interactive and robust diagnostic tool endowed with attributes to efficiently document and analyze the morphological and clinical parameters to accurately detect and diagnose cervical cancer in resource limited settings [7]. Computational automation of digital colposcopy for its large scale screening can significantly enhance accuracy and minimize error rates in diagnosis enabling quicker and timely intervention strategies [8].

The development of an advanced automated diagnostic tool using image analytics would play central role in the use of non-invasive/minimally invasive technologies as an adjunct clinical aid for facilitating rapid and precision-oriented screening of cervical cancer[9, 10].

Om-The OncoMeter is an outcome of this targeted research, including an experiment to classify 1481 labeled cervigrams from the Intel Kaggle MobileODT repository, as well as 300 cervigrams of individual women collected from Batra Hospital and Medical Research Centre (BHMRC), New Delhi. We used R-studio (Integrated Development Environment for R programming language) to automatically demarcate the morphological features, colour intensity, sensitivity, lesion detection, contour formation, enhanced visual inspection and other clinical parameters as shown in Fig. 1. This will significantly improve rapid screening and early diagnosis in a noninvasive manner. We believe that not only can our open source software enable accurate disease labeling but can also facilitate triaging of cervical cancer lesions in order of their severity. To the best of our knowledge, this is the first report of an Earth Mover's Distace (EMD) based scale for ranking the lesions in order of severity, and can provide the rationale for the deployment of a robust and reliable triage system facilitating large-scale screening of cervical cancer in LMICs.

2. Methodology

In the present study, digitized cervigrams as well as socio-demographic data (Table 1) of 300 subjects were acquired from the Out Patient Department (OPD) of BHMRC after obtaining prior Institutional Ethical Clearance and Informed Patient Consent. R Version 3.6.1, an open source software tool equipped with highly advanced image analysis tools and techniques, has been used as the programming language. MySQL Version 14.14 Distrib 5.5.60 has been used for the development of the image/video repository for further processing of digitized cervigrams while LINUX (Ubuntu) version 14.04 has been used as the base computational platform. Digitized Cervigrams acquired using Digital Colposcope (Mobile ODT, Israel) from anonymized subjects presenting at OPD of BHMRC were subjected to preprocessing algorithms (specular removal algorithms) for noise removal before being analyzed for cervical cancer lesion detection using image processing algorithms as shown in Fig. 2. We made use of open source and robust algorithms like EMD using the data/cervigrams/digitized colposcopic images acquired from BHMRC as shown in Fig. 3. Apart from conventional statistical pedagogy, R has numerous dedicated features and capabilities in the area of advanced medical grade image analysis which helped ground the development process of OM-The OncoMeter (Fig.4). The salient features of the EMD(Earth Mover's Distance) algorithm which allowed utilization of rigorous quantitative approaches for colposcopic image comparison and classification follow as shown in Fig. 4 [11, 12, 13]:

- The first and foremost key feature of the algorithm is that it's free

- Image similarity-based detection
- Open source implementation in R
- Robust application to noisy images
- Cloud based storage accessibility and ubiquitous provisioning of services from a collected repository at local, regional and global levels.
- This cloud-based information provisioning will facilitate automated closed loop resource allocation which in turn will provide affordable and accessible health-care technology platform for effective management of disease burden.

Training of the algorithm used as input of 81 normal (control) cervigrams of Indian origin females obtained from the OPD of BHMRC (Fig. 2). This contributed to the choice of basic threshold parameters in the algorithm. EMD values between each two normal cervigrams from n1 to n81 shown in Fig. 3 were calculated and entered in a matrix. EMD was used to calculate the distances between cervigrams, and based on similarity indices, cervigrams were categorized to the nearest one. A matrix of 6561 EMD quantitative values was calculated from n1-n2, n1-n3,..., n1-n81, n2-n3, n2-n4,......,n2-n81 and continued until n80-n81(Fig. 3).

Calculation of the centroid of the normal cervigrams is crucial for designating the threshold values for normal cervigram. The threshold value of the centroid was taken as a reference value for normal cervigrams for facilitating the EMD-based Onco-Meter scale for ranking the cervical cancer lesions in order of severity (Table 2 and Fig. 4).

In addition to the calculation of the centroid for the normal cervigrams socio demographic data was analyzed using SPSS version 25.0 (SPSS Inc., Chicago, IL, USA). Descriptive statistics was applied and bivariate analysis was used to detect any significant difference between categorical variables. p value less than ($<$) 0.05 was considered to be statistically significant. The sensitivity and specificity of the software's capability to detect the cancer cervix lesions was done following the protocol of Coughin et al (1992) [14] and Lalkhen and McCluskey (2008) [15] (Table 3).

Calculation of the mean of EMD values was done in order to obtain the standard normal cervigram for further EMD calculation and comparison with other subsets of abnormal cervigrams (Fig. 4). Our calculation gave a mean value of 26.77. Cervigram number 52, was designated as the standard cervigram as all the normal cervigrams had least distances from it with respect to EMD values (Fig. 3). The image analytics algorithm-based automated processing of the cervigrams was programmed to obtain simultaneous generations of results so as to provide clinicians a valuable adjunct clinical decision-making aid.

The whole process from conceptualized to its demonstration has been depicted in Data Flow Diagram (DFD) (Fig. 1) with an emphasis on unhindered acquisition of cervigrams (Colposcope derived images) along with its processing and result generation in almost real-time using the optimal amount of computational resources.

3. Results

Socio-demographic data revealed that amongst the total of 300 married women enrolled for the current study, a vast majority (65.7%) were between 20-40 years age. 63% of the women were educated up to high-school level. Majority of these women were housewives (88%) with a family annual income between 2-10 lakhs (97%). Only 10.7% of the women enrolled in the study had attained menopause. 74% of the women presented with abnormal menstrual condition (Table I). A striking observation was that although 74.1% women with unhealthy cervix were between 20-40 years. Neither education ($p=0.692$) nor occupation ($p=0.362$) showed any statistically significant association with the health of the cervix (Table 1). Further, our analyses indicated that women from lower socio-economic strata belonging to a household with less than Rs 200,000 per annum (approximately USD 2700 per annum) had highly unhealthy cervix ($p=0.006$) indicating poor knowledge of adoptive reproductive and sexual health practices (Table 1), which is indeed responsible for increasing the vulnerability of women to cervical cancer particularly in LMICs such as India.

We believe that adoption and integration of Intelligent decision support systems involving the use of minimally invasive techniques to detect cancerous cervical lesions through automated segmentation of digitized cervigrams on a real time scale would not only form the rationale for development of effective triage methods towards the early diagnosis of lesions but also prioritizing treatment options.

To this end, An EMD based measurement scale-(OM-The OncoMeter) was developed with a threshold value of 26.77 based on cervigrams obtained from healthy subjects evaluated at BHMRC for being designated as normal. Beyond the threshold value of 26.77(EMD value), cervigrams were computationally tagged into different categories (Table 2 and Fig. 4). Development of an EMD based measurement scale for binning and categorization of cervigrams based on their EMD distance from Centroid of Normal in a quantitative manner is indeed a stellar achievement in the field of cervigram analysis. The scale was designed for large scale screening of cervical cancer and to act as an adjunct clinical aid for early and timely diagnosis by the clinicians. Illustration of the measurement scale was done as per their categorical EMD values obtained after the designation of Centroid value (26.77- EMD value based upon the cervigrams obtained from the OPD of BHMRC as well as its comparison with the abnormal cervigrams (Table 2 and Fig.4).

3.1. Sensitivity and Specificity

Current clinical regimens require large number of clinical tests to confirm or deny existence of a disease or to refer them for more advanced diagnosis in case of indecision. Unfortunately our clinical tests are unable to precisely and accurately delineate all patients with or without the disease. The sensitivity of a clinical test is its ability to identify patients afflicted with the disease. The specificity of a clinical test is its ability to identify patients who do not have the disease. Their joint relationship is expressed with RoC (Receiver operator characteristic) curves. The sensitivity and specificity of the software's capability to detect the cancer cervix lesions was also done following the protocols [14,15] to quantify the performance of our proposed algorithm for early detection of cervical cancer as per standard definitions. Table 3 essentially depicts the divergent data category included in our study which was tested by OM-The OncoMeter producing promising results with 83.56% sensitivity and 59.25% specificity with an accuracy of 77%.

Further research with an extensive amount of cervigrams from different regions of the Indian subcontinent encompassing populace with divergent genetic base and socio-cultural norms would form the rationale for identifying the vulnerable populations to the ravages of this disease. Processing of large numbers of cervigrams would remove the ambiguities/glitches associated within software. In other words, further inclusion of data points will make this scale more robust and trustworthy for deployment in the field as an indicator for large scale disease labeling at a community level.

We believe that with larger training data-set of higher resolution cervigrams, our algorithms have a potential to perform significantly better.

4. Discussions

An estimated 90% of the globally recorded cervical cancer-related deaths are in low-and middle-income countries (LMICs). Cervical cancer is a public health problem in LMICs like India, so much so that India alone accounts for one-quarter of the worldwide burden of cervical cancers [16]. It is estimated that cervical cancer will occur in approximately 1 in 53 Indian women during their lifetime compared with 1 in 100 women in more developed regions of the world [17] due to availability of efficient and accessible screening programs as well as diagnostic and treatment facilities. Apart from the preponderance of HPV infection, a variety of clinico-epidemiological risk factors such as early age of marriage, promiscuous sexual behaviour, multiple pregnancies, poor genital hygiene along with aberrant lifestyle choices such as smoking are known to be

Table 1: Socio Demographic Features of These Subjects Evaluated at the OPD at BHMRC

Category		Test Cases (N=219)				Control Cases (N=81)
		Women with cervicits	Women with Abnormal Cervix (Vaginitis, Nabothian Cyst, Cervical Erosion, Polyp)	Women with Precancerous	Women with Suspected Cancer	Women with Normal Cervix
Age	< 20	0 (0%)	1(0.46%)	0 (0%)	0 (0%)	0(0%)
	21-40	34(15.53%)	110(50.23%)	0(0%)	1(0.46%)	50(61.73%)
	41-60	21(9.59%)	47(21.46%)	2(0.91%)	1(0.46%)	25(30.86%)
	> 60	1(0.46%)	1(0.46%)	0(0%)	0(0%)	6(7.41%)
Education	Illiterate	8(3.65%)	13(5.94%)	0 (0%)	1(0.46%)	7(8.63%)
	Literate	23(10.5%)	63(28.76%)	0(0%)	0(0%)	27(33.33%)
	10th	8(3.65%)	21(9.59%)	0(0%)	0(0%)	11(13.58%)
	12th	4(1.83%)	20(9.13%)	1(0.46%)	0(0%)	11(13.58%)
	graduate	10(4.56%)	33(15.07%)	0(0%)	1(0.46%)	17(20.99%)
	postgraduate	3(1.37%)	9(4.11%)	1(0.46i%)	0(0%)	8(9.89%)
Occupation	Housewife	52(23.74%)	140(63.93%)	1(0.46%)	2(0.91%)	69(85.18%)
	Govt.-Job	0(0%)	2(0.91%)	1(0.46%)	0(0%)	2(2.47%)
	Private job	3(1.37%)	17(7.76%)	0(0%)	0(0%)	10(12.35%)
	Other	1(0.46%)	0(0%)	0(0%)		0(0%)
Economic Status (INR)	$> 10 Lakhs$	0(0%)	0(0%)	0(0%)	0(0%)	0(0%)
	2-10Lakhs	43(19.63%)	122(55.70%)	1(0.46%)	1(0.46%)	68(83.95%)
	$< 2 Lakhs$	13(5.93%)	34(15.53%)	1(0.46%)	1(0.46%)	13(16.05%)
	Don't know	0(0%)	3(1.37%)	0(0%)	0(0%)	0(0%)
Marital Status	Single	0(0%)	0(0%)	1(0.46%)	0(0%)	0(0%)
	Married	56(25.57%)	159(72.60%)	1(0.46%)	2(0.91%)	81(100%)
Menstrual-Health	Normal	1(0.46%)	11(5.02%)	0(0%)	0(0%)	34(41.98%)
	Abnormal	45(20.55%)	139(63.47%)	2(0.91%)	1(0.46%)	35(43.21%)
	Post-Menopausal	10(4.56%)	9(4.11%)	0(0%)	1(0.46%)	12(14.81%)
Total		Test Cases (N=219, 100%)				Control Cases (N=81, 100%)

Table 1: It depicts the socio-demographic features of the subjects/patients collected at BHMRC, New Delhi, India. These cervigrams were acquired from subjects evaluated at the OPD of BHMRC, New Delhi. The socio-economic demographic/ epidemiological data(table 1) of 300 subjects, was labeled by expert Gynecologist as normal, Cervicitis, precancerous, suspected cancerous and abnormal (Vaginitis, Nabothian cyst, Polyp and Cervical erosion) were taken to develop the novel OM- The Onco-Meter scale to rank the lesions in the order of their severity. Out of which 81 samples of subjects were found to be normal, 56(%) had Cervicitis, 159 subjects were found to have abnormal lesions, while 2 subjects each were suspected to have precancerous and cancerous lesions.it was observed that education(p=.692) and occupation(p=.362) did not show any statistically significant association with cervical condition. Although annual income showed statistically significant association with the health of the cervix.(p=0.006). Menstrual health was also significantly associated with the health of the cervix (p¡0.001)

Diagrammatic Representation of Steps Involved in Pre-processing and Processing of Digitized Cervigrams for detection of Abnormal

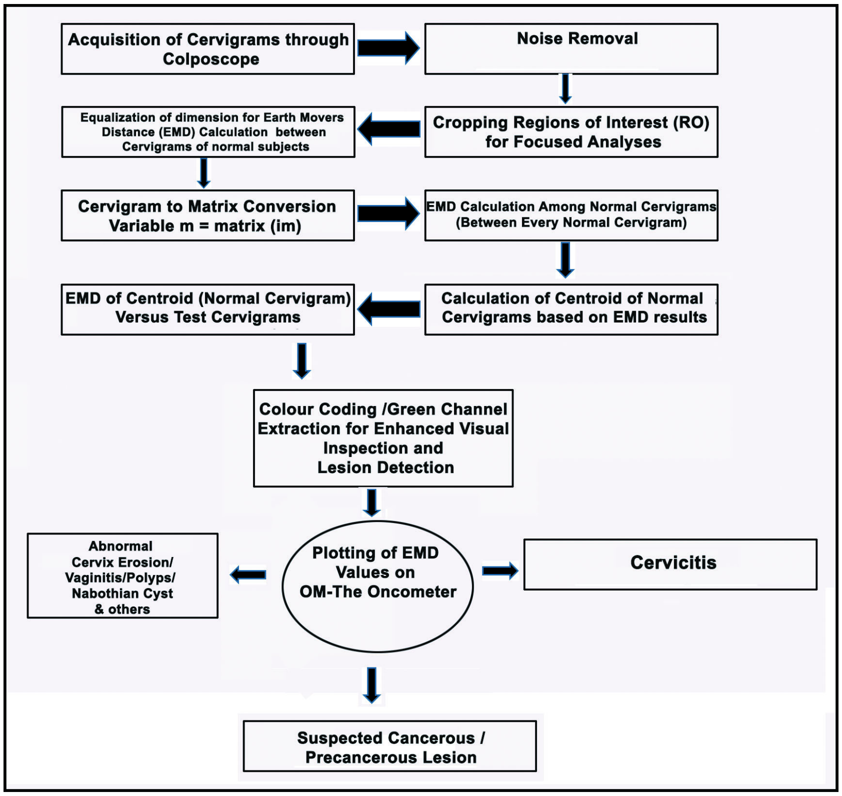

Figure 1: Our present study depicts a diagrammatic representation of the processing the digitized cervigrams obtained from the patients evaluated at the Out Patient Department of Batra Hospital and Medical Research Centre. The pre-processing algorithms include noise removal associated with the extraction of the image features (digitized cervigrams in this case) before cropping the raw image of cervigrams to define the regions of interest for equalization of the dimensions before being processed with Earth Movers Distance (EMD) to detect initially the variances between the cervigrams of normal subjects so as to set a threshold value for the cervigrams of the normal cervigrams. The regions of interest of the cropped regions were colour coded using green channel to enhance the sensitivity of visual inspection and lesion detection. Subsequently, the values obtained from the comparison of both normal and abnormal cervigrams were plotted on a scale virtually stack the cervigrams based upon the EMD value. This lead to the creation of OM The Oncometer, a scale used to rank the cervigrams in order of the disease severity.

Diagrammatic Representation of the Image Processing of Normal Cervigrams.

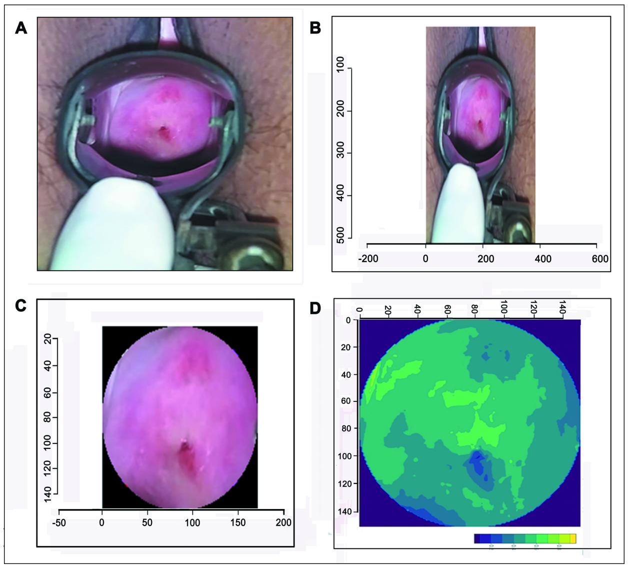

Fig. 2(a): It depicts the processing of the normal cervigrams by subjecting it to preprocessing algorithms to remove the noise potentially interfering with the extraction of the Image features before being processed with green channel colour coding algorithms and Earth Movers Distance (EMD) algorithm to facilitate the assignment of a numerical score to the normal cervigrams and also ranking them based upon their EMD score.

Diagrammatic Representation of the Image Processing of Abormal Cervigrams.

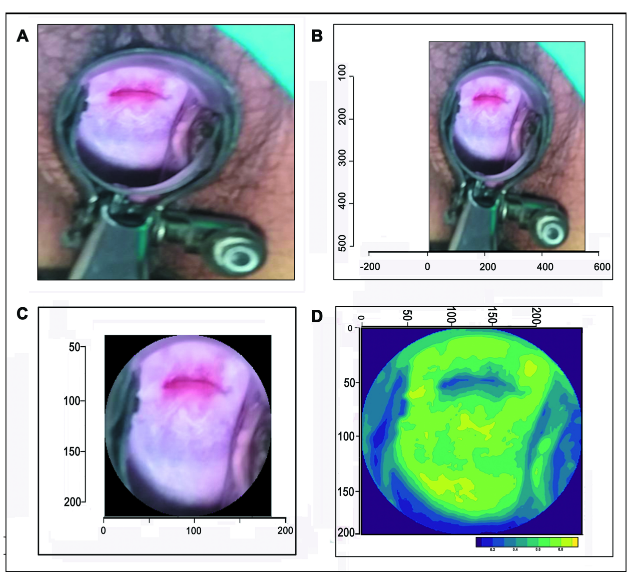

Fig. 2(b): It depicts the processing of the abnormal cervigrams by subjecting it to preprocessing algorithms to remove the noise potentially interfering with the extraction of the Image features before being processed with green channel colour coding algorithms and Earth Movers Distance (EMD) algorithm to facilitate the assignation of a numerical score to the abnormal cervigrams and also ranking them based upon their EMD score.

Figure 2: Cervigram Processing

Earth Movers Distance Based Calculation of Distances between Normal cervigrams among the subjects evaluated at Batra Hospital

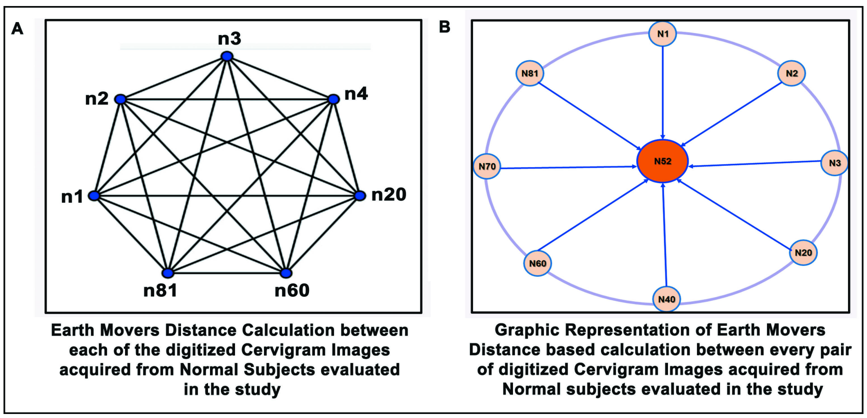

Figure 3: In the present study cervigrams of 84 normal subjects were acquired from the Out Patient Department of Batra Hospital and Medical Research Centre (BHMRC) after obtaining prior Institutional Ethical Clearances and Informed Patient Consent Forms. Each of the 84 cervigrams obtained from age matched normal subjects were compared with each other to calculate the variance between the cervigrams of the normal subjects. The Earth Movers Distance (EMD) enabled calculation of the variance within the cervigrams of the normal subjects enabled the calculation of the threshold value for the cervigrams of the normal subjects, which in this case was found to be 26.77 obtained from normal subjects evaluated the OPD of BHMRC, New Delhi. This threshold value was used to demarcate the normalness of the cervigrams along with its comparison with the abnormal cervigrams as observed in the patients evaluated in the Out-Patient Department of Batra Hospital and Medical Research Centre, Delhi as a part of the current study.

Development of OM- The OncoMeter for triaging of subjects/ patients with

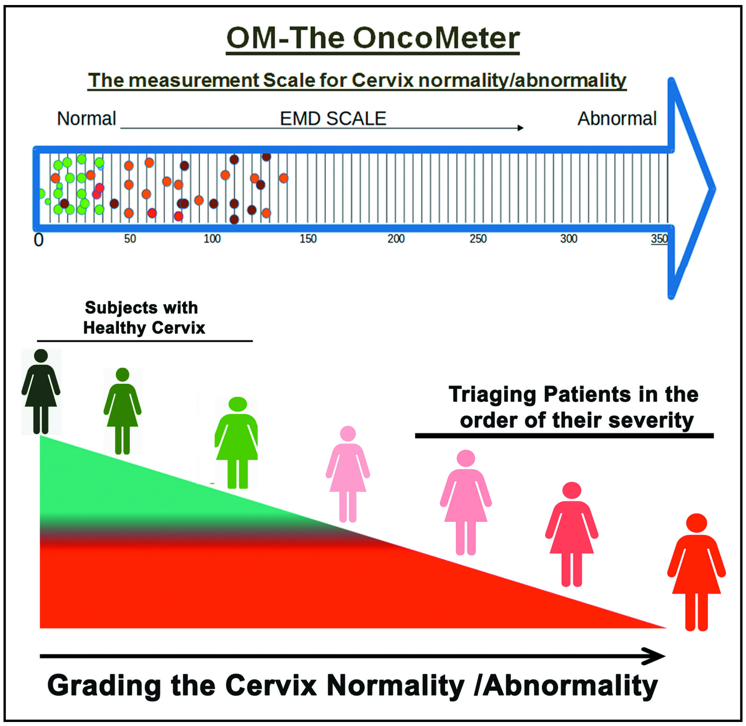

Figure 4: This figure summarizes the impact of our study where we have attempted to make a new scale to virtually rank and grade the cervigrams in the order of their normality or abnormality. This indeed forms the rationale to facilitate the rapid screening of the cervical cancer in the rural milieus of the Indian sub-continent besides ranking the cervigram lesions in the order of the severity of the disease progression. Such an automated processing of chores will ultimately help the clinicians to intervene the patients in need of intervention on a priority basis. The use of this technology will also help in augmenting the hospital based registry with community based data which will indeed provide a more realistic scenario of cervical cancer prevalence in resource limited healthcare systems prevalent in the Indian sub-continent and elsewhere in the world.

Table 2:Range measurement of different categorical cervigrams

Cervigram Types diagnosed by Oncologist	EMD Range on scale	Number of Subjects/Cervigrams
Normal	0-26.78	81
Abnormal(Nabothian Cyst, vaginitis, Cervical erosion, polyp)	8.162584-146.8711	159
Cervicitis	19.41418-140.2637	56
Precancerous	34.65366-36.1091	2
Suspected Cancer	67.77164-78.53794	2
	Total	300

Table 2: This table depicts the EMD value based computational tagging of cervigrams into different categories such as normal, abnormal, cervicitis, precancerous, precancerous/cancerous lesions. The EMD values obtained from the cervigrams form the rationale for developing om-the OnvoMenter, a scale used for computational tagging of the cervigrams in the order of their severity thereby facilitating the triaging of lesions

Table 3: Calculation of Sensitivity, Specificity and Accuracy of proposed algorithm.

Cases(Abnormal Cervigrams)-219		Controls(Normal Cervigrams)-81	
True Positive	183	True Positive	48
False Positive	36	False Positive	33
Sensitivity ($\frac{183}{219}$)	83.56%	Sensitivity ($\frac{48}{81}$)	59.25%
Accuracy ($\frac{231}{300}$)		77%	

Table 3: We use basic notions of sensitivity and specificity [15, 14] to quantify the performance of our proposed algorithm for early detection of cervical cancer as per standard definitions. This Table essentially depicts the divergent data category included in our study which was tested by OM-The OncoMeter and so far the scale has produced promising results with 83.56% sensitivity, 59.25% specificity and 77 % accuracy.

associated with increased risk of cervical cancer.

Results obtained from our data indicate that lack of adequate knowledge about the adoptive reproductive and sexual health practices is indeed responsible for increasing the vulnerability of women to cervical cancer.

A successful implementation of this strategy will have a positive impact on not only understanding the niche specific drivers for onset and pathological sequelae of cervical cancer but also provide the rationale for developing novel precision oriented niche specific non-linear predictive algorithms for early clinical diagnosis as well as prioritizing delivery of treatment options to high risk patients afflicted with cervical cancer. The adoption of low end mobile health based applications to propagate Internet and Communication Technology (ICT) based awareness about cervical cancer might not only contribute in overcoming region specific social stigmas and taboos associated with prevalence of cervical cancer but also facilitate remote connect of cancer afflicted patients with the treating physicians. The etiopathogenesis of Cervical Cancer is unique, as a consequence, it's characteristics pertaining to different symptoms and parameters necessary for disease labeling, tagging and further diagnosis are unique and community specific or niche specific. One of the biggest challenges is to increase the accuracy of the digitally acquired images which in turn pertains low resolution optics in prevailing colposcopes (In our case, MobileODT colposcope had only 13 Mega pixel resolution which seriously hampered our efforts to detect cervical cancer at early stage) and to the lack of optimal oncological image/video processing algorithms for outlining correct set of parameters [18, 19]. Some progress has been reported in histological image of cervical cancer by Taneja et.al [20] where multi-level set based image processing techniques and deep learning has been used to outline cell nuclei.

Our work should be looked as complementary or even advanced effort to analyze cervical cancer cervigrams at normal optical scale obviating the need of sophisticated microscopy equipment and hence making early cervical cancer detection accessible to economically weaker sections of the society. Modern efforts to leverage the artificial intelligence capabilities in resolving the cervical cancer issue are being reported continuously [21]. Disease-labeling of cancer is a data-intensive and rigorous task which requires accurate and specialized expertise. These skills are not readily available particularly in remote areas with limited healthcare infrastructure. The prevalence of misdiagnosis and missed diagnosis is high in absence of portable, accurate and robust diagnostic tools [19] where our proposed tool could aid the medical professionals to drastic improve the detectiion of cervical cancer at early stage. One of the strategies for ensuring early detection of a cervical cancer lesion pertains to large-scale rapid screening for the disease in PHCs. In the case of Cervical Cancer, screening is done through Liquid-based Cytology and Colposcopy. In Liquid based Cytology, the major hindrance is acquisition of a sample and the lack of specialized professionals to make the diagnosis. However, in the recent past there have been tremendous technological advancements with the advancements of AI (Artificial Intelligence) [22]. In a resource-limited country like India there is an urgent and unmet need to clearly identify reliable and robust visual cues in cervigrams. Further, these cues should be robust enough to remain invariant un-occluded under different optical transformations and even in the presence of unwanted objects in images like hair and other body parts. Such a process will eliminate artifacts and other irrelevant features to seamlessly segment structures for grading measurements of cervical cancer lesions. The automation of this feature ensures that the algorithm detects structural aberrations in cervigrams in a precise and timely manner to make Cervical Cancer detection even more reliable [23].

Optical clarity of image is paramount for visual inspection by a camera with direct human eye inspection. Optical aberration, colour misrepresentation, specular reflection in cervigrams are other challenges associated with conventional colposcope[24]. Our approach takes care of these problems facilitating the provisioning of much sharper and focused images, necessary for delineation and identification of cancerous lesions. Using image processing-based automated detection of cervical cancer can increase precision and minimize error. When detected and managed at an early stage, the clinical prognosis for cervical cancer is favorable [25].

This indeed necessitates the need for a technology-driven approach for developing an effective and precision oriented triage system for facilitating not only large-scale screening but also prioritizing patients for therapeutic intervention. This can minimize the recovery period through screening and follow-up. Taking into account the conditions, parameters, limitations and the knowledge sharing from manual to computational colposcopy, we tried to incorporate the ubiquitous information related to the labeling/tagging of the cervix and have developed an Artificial Intelligence enabled Intelligent Decision Support System- capable of serving as an adjunct clinical aid for facilitating rapid screening for detection of cancerous lesions. In this study, we have introduced an algorithm-based Colposcopic image (cervigrams) analysis and diagnostic technology which uses the attribution method to identify the cervix types of Indian subjects [26, 27]. R software libraries include advanced tools for medical image analysis, which deal with the quantification of the colposcope-derived cervigrams, thereby reducing the burden and time of medical professionals by representing images as numerical data for analysis apart from conventional image processing procedures [28, 29].

5. Future Directions

Development of multi-modal multi-sensor fusion technology by integrating signals from Artificial Intelligence enabled image analysis, HPV serotyping as well as Liquid Based Cytology to ensure precision oriented large scale rapid screening of subjects at community level to eliminate missed and mis-diagnosis of cervical cancer. Modules of health literacy about adoptive sexual and reproductive health practices will also be integrated in the software solution to alleviate the burden of this disease in vulnerable populations belonging to the lowest socioeconomic strata.

6. Targetted Audience

The core idea behind this software was to help the poor and needy women from LMICs but the vision is to implement it in developed countries also. It primarily targets patients as a self-diagnosing device and clinicians will use this system as an aide in their assessments.

7. Acknowledgements

We acknowledge the Indian Council of Medical Research, New Delhi for support as grant in aid to Dr Rajiv Janardhanan and Dr Priya Ranjan (Grant id no.-2029-0416-No.- ISRM/12(23)/2019). ICMR Senior Research Fellowship(ref. no- BIC/11(06/2016)) awarded to Mr. Kumar Dron Shrivastav and the support of mobile ODT for provision of cervigrams obtained through mobile ODT EVA system at BMHRC under the able supervision of Dr. Shelly Batra.

References

[1] Rao J, Escobar-Hoyos L and Shroyer KR, Unmet clinical needs in cervical cancer screening., MLO: medical laboratory observer 48 (2016) 8, 10, 14; quiz 15. doi:Retrievedfromhttps://www.mlo-online.com/.

[2] World Health Organization, Who. cervical cancer. early diagnosis and screening, WHOdoi:Retrievedfromhttps://www.who.int/cancer/prevention/diagnosis-screening/cervical-cancer/en/.Lastaccessedon14-01-2020.

[3] The Global Cancer Observatory, India factsheet, WHOdoi:Retrievedfromhttps://https://gco.iarc.fr/today/data/factsheets/populations/356-india-fact-sheets.pdf.Lastaccessedon14-01-2020.

[4] Shrivastav K.D., Mukherjee Das A., Singh H., Ranjan P., Janardhanan R., Classification of colposcopic cervigrams using emd in r., Advances in

Signal Processing and Intelligent Recognition Systems. SIRS 2018. Communications in Computer and Information Science 968 (2019) 298–308. doi:10.1007/978-981-13-5758-9_25.

[5] ICO/IARC, Ico/iarc information centre on hpv and cancer,india human papillomavirus and related cancers, fact sheet 2018., MLO: medical laboratory observerdoi:Retrievedfromhttps://hpvcentre.net/statistics/reports/IND_FS.pdf/.Lastaccessedon14-01-2020.

[6] Kumar P, Gupta S, Das AM and Das BC, Towards global elimination of cervical cancer in all groups of women., The Lancet. Oncology 20 (2019) 5. doi:10.1016/S1470-2045(19)30170-6.

[7] Tian R, Cui Z, He D, Tian X, Gao Q, Ma X, Yang JR, Wu J, Das BC, Severinov K, Hitzeroth II., Risk stratification of cervical lesions using capture sequencing and machine learning method based on hpv and human integrated genomic profiles., Carcinogenesis 40 (2019) 1220–1228. doi:10.1093/carcin/bgz094.

[8] Park SY, Follen M, Milbourne A, Rhodes H, Malpica A, MacKinnon NB, MacAulay CE, Markey MK, Richards-Kortum RR., Automated image analysis of digital colposcopy for the detection of cervical neoplasia., Journal of biomedical optics. 13 (2008) 1–10. doi:10.1117/1.2830654.

[9] Lange H, Wolters RH, Uterine cervical cancer computer-aided-diagnosis (cad)., Journal of biomedical optics. (2010) 1–10doi:Retrievedfromhttps://patents.google.com/patent/US7664300B2/en.Lastaccessedon14-01-2020.

[10] de Castro Hillmann E, Bacha OM, Roy M, Paris G, Berbiche D, Nizard V, Ramos JG., Cervical digital photography: An alternative method to colposcopy, Journal of Obstetrics and Gynaecology Canada 41 (2019) 1099–1107. doi:10.1016/j.jogc.2018.10.025.

[11] Rubner Y, Tomasi C, Guibas LJ., The earth mover's distance as a metric for image retrieval, International journal of computer vision 40 (2000) 99–121. doi:10.1023/A:1026543900054.

[12] Boltz S, Nielsen F, Soatto S., Earth mover distance on superpixels, 2010 IEEE International Conference on Image Processing 40 (2010) 4597–4600. doi:Retrievedfromhttps://ieeexplore.ieee.org/abstract/document/5651708/metrics#metricsLastaccessedon14-01-2020.

[13] Sun Y, Lei M., Method for optical coherence tomography image classification using local features and earth mover's distance., Journal of Biomedical Optics. 14 (2009) 054037. doi:10.1117/1.3251059.

[14] S. S. Coughlin, B. Trock, M. H. Criqui, L. W. Pickle, D. Browner, M. C. Tefft, The logistic modeling of sensitivity, specificity, and predictive value of a diagnostic test, Journal of clinical epidemiology 45 (1) (1992) 1–7.

[15] A. G. Lalkhen, A. McCluskey, Clinical tests: sensitivity and specificity, Continuing Education in Anaesthesia Critical Care & Pain 8 (6) (2008) 221–223.

[16] S. E. Forhan, C. C. Godfrey, D. H. Watts, C. L. Langley, A systematic review of the effects of visual inspection with acetic acid, cryotherapy, and loop electrosurgical excision procedures for cervical dysplasia in hiv-infected women in low-and middle-income countries, JAIDS Journal of Acquired Immune Deficiency Syndromes 68 (2015) S350–S356.

[17] Institute for Health Metrics and Evaluation, Institute for health metrics and evaluation. the challenge ahead: Progress in breast and cervical cancer. institute of health metrics and evaluation, IHMEdoi:http://www.healthmetricsandevaluation.org/publications/policyreport/challenge-ahead-progress-and-setbacksbreastand-cervical-cancer.Lastaccessedon15-04-2020.

[18] Huh WK, Papagiannakis E, Gold MA., Observed colposcopy practice in us community-based clinics: The retrospective control arm of the improve-colpo study., Journal of lower genital tract disease. 23 (2019) 110–115. doi:10.1097/LGT.0000000000000454.

[19] Hu L, Bell D, Antani S, Xue Z, Yu K, Horning MP, Gachuhi N, Wilson B, Jaiswal MS, Befano B, Long LR., An observational study of deep learning and automated evaluation of cervical images for cancer screening., Obstetrical Gynecological Survey 74 (2019) 343–344. doi:10.1097/OGX.0000000000000687.

[20] A. Taneja, P. Ranjan, A. Ujlayan, Multi-cell nuclei segmentation in cervical cancer images by integrated feature vectors, Multimedia Tools and Applications 77 (8) (2018) 9271–9290.

[21] L. J. Menezes, L. Vazquez, C. K. Mohan, C. Somboonwit, Eliminating cervical cancer: A role for artificial intelligence, in: Global Virology III: Virology in the 21st Century, Springer, 2019, pp. 405–422.

[22] CDC, Cervical cancer screening guidelines for average-risk women, CDC doi:{https://www.cdc.gov/cancer/cervical/pdf/guidelines.pdf}.

[23] Z. Xue, S. Antani, L. R. Long, J. Jeronimo, G. R. Thoma, Comparative performance analysis of cervix roi extraction and specular reflection removal algorithms for uterine cervix image analysis, in: Medical Imaging 2007: Image Processing, Vol. 6512, International Society for Optics and Photonics, 2007, p. 65124I.

[24] A. R. M. Al-shamasneh, U. H. B. Obaidellah, Artificial intelligence techniques for cancer detection and classification: review study, European Scientific Journal 13 (3) (2017) 342–370.

[25] H. Dvir, S. Gordon, H. Greenspan, Illumination correction for content analysis in uterine cervix images, in: 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06), IEEE, 2006, pp. 95–95.

[26] E. Kim, X. Huang, A data driven approach to cervigram image analysis and classification, in: Color Medical Image analysis, Springer, 2013, pp. 1–13.

[27] X. Huang, W. Wang, Z. Xue, S. Antani, L. R. Long, J. Jeronimo, Tissue classification using cluster features for lesion detection in digital cervigrams, in: Medical Imaging 2008: Image Processing, Vol. 6914, International Society for Optics and Photonics, 2008, p. 69141Z.

[28] J. Muschelli, A. Gherman, J.-P. Fortin, B. Avants, B. Whitcher, J. D. Clayden, B. S. Caffo, C. M. Crainiceanu, Neuroconductor: an r platform for medical imaging analysis, Biostatistics 20 (2) (2019) 218–239.

[29] T. Xu, H. Zhang, C. Xin, E. Kim, L. R. Long, Z. Xue, S. Antani, X. Huang, Multi-feature based benchmark for cervical dysplasia classification evaluation, Pattern recognition 63 (2017) 468–475.