IST-Africa 2019 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2019
ISBN: 978-1-905824-63-2

Automated Diagnosis and Classification of Cervical Cancer from pap-smear Images

Wasswa WILLIAM[1], Andrew WARE[2], Annabella Habinka BASAZA-EJIRI[3], Johnes OBUNGOLOCH[1]
[1]*Department of Biomedical Sciences and Engineering,*
Mbarara University of Science and Technology, Mbarara, 1410, Uganda
Tel: +256775046515, Email: wwasswa@must.ac.ug, jobungoloch@must.ac.ug
[2]*Faculty of Computing, Engineering and Science, University of South Wales, Prifysgol, UK*
Tel: 01443 4 82650, Email: andrew.ware@southwales.ac.uk
[3]*Department of Information Technology, Makerere University, Kampala, Uganda,*
Tel: +256772571444, Email: annabella.habinka@cit.ac.ug

Abstract: Globally, cervical cancer ranks as the fourth most prevalent cancer affecting women. However, cervical cancer can be treated if detected at an early stage. Pap-smear is a good tool for screening of cervical cancer but the manual analysis is error-prone, tedious and time-consuming. The objective of this study was to rule out these limitations by automating the process of cervical cancer classification from pap-smear images by using an enhanced fuzzy c-means algorithm. Simulated annealing coupled with a wrapper filter was used for feature selection. The evaluation results showed that our method outperforms many of previous algorithms in classification accuracy (99.35%), specificity (97.93%) and sensitivity (99.85%), when applied to the Herlev benchmark pap-smear dataset. The overall accuracy, sensitivity and specificity of the classifier on prepared pap-smear slides was 95.00%, 100% and 90.00% respectively. False Negative Rate (FNR), False Positive Rate (FPR) and classification error of 0.00%, 10.00% and 5.00% respectively were obtained. The major contribution of this tool in a cervical cancer screening workflow is that it reduces on the time required by the cytotechnician to screen very many pap-smears by eliminating the obvious normal ones, hence more time can be put on the suspicious slides. The proposed tool has the capability of analyzing 1-2 smears per minute as opposed to the 5-10 minutes per slide in the manual analysis.

Keywords: Pap-smear, Cervical Cancer, Automated Analysis

1. Introduction

Globally, cervical cancer ranks as the fourth most prevalent cancer affecting women with 527,624 women diagnosed with the disease and 265,672 dying from it every year [1]. In sub-Saharan Africa, 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, with over 80% of cervical cancer detected in its late stages [2]. Over 85 per cent of cervical cancer cases occur in less developed countries of which the highest incidences are in Africa, with Uganda being ranked 7th among the countries with the highest incidences of cervical cancer. Over 85% of those diagnosed with the disease in Uganda die from it [3]. Regular pap-smear screening is the most successful and effective attempt by medical science and practice to facilitate the early detection and screening of cervical cancer. However, the manual analysis of the pap-smear images is time-consuming, laborious and error-

prone as hundreds of sub-images within a single slide have to be examined under a microscope by a trained cytopathologist. To overcome the limitations associated with the manual analysis of pap-smear images, computer-assisted pap-smear analysis systems using image processing and machine-learning techniques have been proposed by several researchers [4-5].

1.1 Computer-Assisted Pap-Smear Analysis

Since the 1960's numerous projects have developed computer-assisted pap-smear analysis systems leading also to a couple of commercial products like AutoPap 300 QC (NeoPath, Redmond, WA, USA) [6] and the PapNet (Neuromedical Systems Inc., Suffern, NY, USA) [7] which were approved by the United States Food and Drug Administration (FDA). However, these have had limited impact on cervical cancer screening in most of the low developed countries. Ever since the first appearance of computers, significant development efforts have been aiming at supplementing or replacing the human visual inspection of pap-smears by computer-aided analysis, however, the problem turned out to be a lot harder than expected.

1.2 Review of Related Literature on Tools for Automated Analysis of Pap-Smear Images

Due to the fact that the manual visual screening of pap-smear images is very demanding, tedious, and expensive in terms of labour requirements, proposals for automating the pap-smear analysis have been proposed. It took more than 40 years before the first successful commercial systems appeared [8]. However, still, automated screening is not sufficiently cost-effective to completely replace the visual analysis from the relatively limited penetration of automated screening systems in the screening operations worldwide [8]. This section presents some of the systems that have been developed in this line.

The cytoanalyzer project in the US was the first attempt at building an automated screening device for pap-smears based on the concept of nuclear size and optical density [9]. Unfortunately, tests with the cytoanalyzer revealed that the device produced too many false rates (false negatives and false positives). There were numerous objects of a size similar to malignant cells present also on normal specimens. The project failed mainly because of this artefact rejection problem. The CYBEST was developed by Watanabe and coworkers at Toshiba in Japan based on initially special purpose electronic circuits and later with the advance in technology on digital computers [10]. They extracted 4 features from the cell; nucleus area, nucleus density, cytoplasmic area, and nuclear to cytoplasmic ratio. The prototypes were used in large field trials in the Japanese screening program and showed promising results but none of them became a product. A Precision Encoding and Pattern Recognition (PEPR) cathode ray tube (CRT) scanning system at Nijmegen was modified to study the feasibility of developing a fully automatic system for the prescreening of cervical smears. PEPR system was then renamed BioPEPR and developed into a general image analysis system for cervical cancer screening [11]. The system used nuclear area, nuclear optical density, nuclear texture, and nuclear to cytoplasmic ratio. There was no in-depth study made to assess the efficiency of BioPEPR system in detecting abnormalities of the endometrium and hence the product did not go to market. Another system that was developed was FAZYTAN [12], a system for fast automated cell segmentation, cell image analysis and feature extraction based on TV-image pickup and parallel processing. The system performed efficient and fast detection and segmentation of cells scanned in one TV frame within one second as well as the extraction of a large number of morphologic

features within a few seconds. However, FAZYTAN never reached the market, and an important reason for this was lack of cost-effectiveness.

In 1981 Cervical was developed and a trial was carried out to evaluate its performance for automated cytology prescreening system on cervical scrape specimens prepared by a suspension-polyline-hematoxylin preparation method [13]. The results were not too good and this was attributed to the number of false-positive signals caused by the hematoxylin stain and the number of specimens rejected as inadequate by the scanner. In 1998, FDA approved the first company (Tripath) with an automated cervical cancer screening product. The Tripath Company was later acquired by BD in 2006 and the system renamed BDFocalPoint Slide Profiler which is largely based on the AutoPap 300 system [6] with a new liquid-based specimen preparation technique called SurePath added to it to further improve the system performance. On the other hand, a number of academic researchers have proposed different algorithms for automated analysis of pap-smears as presented in [4].

This paper presents a tool for automated diagnosis and classification of cervical cancer from pap-smears using an enhanced fuzzy c-means algorithm.

2. Materials and Methods

2.1.1 Input Images

The cervical cancer classification was achieved through a sequential approach depicted in Figure 1. The approach was developed using Harlev dataset. The dataset consists of 917 single cells of pap-smear images prepared by Jantzen et al. [14]. The dataset contains pap-smear images that were obtained by skilled cytopathologists using a microscope connected to a frame grabber and taken with a resolution of 0.201µm/pixel. The images were segmented using CHAMP commercial software developed by DIMAC Imaging systems and then classified into 7 classes [14]. Furthermore, the performance of the classifier was evaluated on samples of pap-smears obtained from Mbarara Regional Referral hospital. Specimens were imaged using an Olympus BX51 bright-field microscope equipped with a 40×, 0.95 NA lens and a Hamamatsu ORCA-05G 1.4 Mpx monochrome camera, giving a pixel size of 0.25µm with 8-bit grey depth. Each image was then divided into 300 areas with each area containing between 200 to 400 cells. Based on the opinions of the cytopathologists, 10,000 cells in images derived from 98 different pap-smear slides were selected. These included cervical epithelial cells (250 normal and 250 abnormal cells), other types of cells and debris.

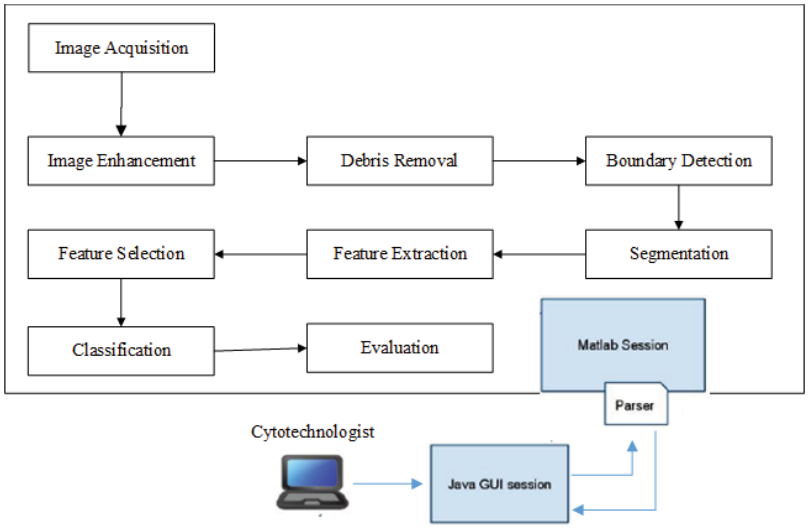

Figure 1: The Proposed Approach for Cervical Cancer Classification from Pap-Smear Images

2.1.2 Image Enhancement

Image contrast was to ensure that all the images had the same properties and for the purpose of stressing the regions of interest. A contrast local adaptive histogram equalization (CLAHE) [16] was applied to a grayscale image. Conversion to RGB channel was achieved using a grayscale technique implemented using Equation 1 as defined in [17].

$$New\ Grayscale\ Image = \big((0.3 * R) + (0.59 * G) + (0.11 * B)\big), \quad (1)$$

Where R=Red, G=Green and B=Blue color contributions to the new image.

2.1.3 Debris Removal

The main reason for the current limitations of many of the existing automated pap-smear analysis systems is that they fail to differentiate between cervical cells and debris. It has been shown that classifiers designed to differentiate between normal and pre-cancerous cells produce unpredictable results when artefacts exist in the pap-smear [18]. In this paper, a technique to identify cervix cells using a three-phase sequential elimination scheme (Figure 2) is presented.

Figure 2: Three-phase sequential elimination approach for debris rejection

The approach sequentially removes debris from the pap-smear if deemed unlikely to be a cervix cell. This is beneficial as it allows a lower-dimensional decision to be made at each stage.

Size Analysis: The area is one of the most basic features used in the field of automated cytology to separate cells from debris. One of the key changes with nucleus area assessment is that cancerous cells undergo a substantial increase in nuclear size [18]. Therefore, determining an upper size threshold that does not systematically exclude diagnostic cells is much harder, but has the advantage of reducing the search space. The method presented in this paper is based on a lower size and upper size threshold of the cervical cells as shown in Equation 2.

$$If\ Area_{min} \leq Area_{rel} \leq Area_{max}\ then < foreground > else < Background > \quad (2)$$

The objects in the background are regarded as debris and discarded from the image. Particles with area which fall in the rage of $Area_{min}$ and $Area_{max}$ discarded during the next stages.

Shape Analysis: The shape of the objects in a pap-smear is a key feature in the differentiation between cells and debris [14]. A region-based method (perimeter2/area (P2A)) has been used [19] for shape analysis. The P2A was chosen on the merit that it describes the similarity of an object to a circle. The P2A is sometimes referred to as shape compactness and is defined by Equation 3:

$$c = p^2 / A \quad (3)$$

Where c is the value of shape compactness, A is the area and p is the perimeter of the nucleus.

Texture analysis: Within a pap-smear, the distribution of average nuclear stain intensity is much narrower than the stain intensity variation among debris [18]. This fact was used as the basis to remove debris based on their image intensities and colour information using Zernike moments (ZM) [20]. The ZM of order n with repetition I of function $f(r,\theta)$, in polar cordinates inside a disk centered in square image $I(x,y)$ of size $m\ x\ m$ is given by Equation 4 as used in [18].

$$A_{nl} = \frac{n+1}{\pi} \sum_x \sum_y v^*_{nl}(r,\theta) I(x,y) \quad (4)$$

$v^*_{nl}(r,\theta)$ denotes the complex conjugate of the Zernike polynomial $v_{nl}(r,\theta)$. To produce a texture measure, magnitudes from A_{nl} centered at each pixel in the texture image are averaged.

2.1.4 Boundary Detection

Boundary detection is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness [21]. An approach based on a canny edge detection method was implemented to detect edges in an image by locating pixel locations where the gradient was higher than its neighbours [22]. This paper uses a 5x5 Gaussian filter with standard deviation = 1.5 to smooth the image in order to remove the noise.

2.1.5 Segmentation

Image segmentation was achieved through a locally adaptive thresholding and morphology.

(a) Adaptive Threshold

Due to the varying illumination in a pap-smear, a locally adaptive thresholding was used. A local adaptive thresholding technique developed by Niblack [23] was adopted where a local threshold *T(x, y)* at point *(x, y)* was calculated within a window of size *w*w* as shown in Equation 5.

$$T(x,y) = m(x,y) + k\delta(x,y) \quad (5)$$

Where $m(x,y)$ and $\delta(x,y)$ are the local mean and standard deviation of the pixels inside the local window and k is the bias which controls the level of adaptation varying the threshold value. Satisfactory segmentation results were obtained at k = -0.2 and w=15.

(b) Morphology

The two major techniques in morphology are dilation and erosion. Morphological techniques probe an image with a small template called a structuring element [49]. The dilation of an image f by a structuring element s produces a new binary image $g = f \oplus s$ with ones in all locations (x,

y) of a structuring element's origin at which that structuring element s hits the input image *f, i.e. g(x, y) = 1* if s *hits* f and 0 otherwise, repeating for all pixel coordinates (x, y). On the other hand, the erosion of a binary image *f* by a structuring element s produces a new binary image g = f⊖ s with ones in all locations *(x, y)* of a structuring element's origin at which that structuring element s fits the input image *f, i.e. g(x, y) = 1* if *s fits f* and *0* otherwise, repeating for all pixel coordinates *(x, y)* [24].

2.1.6 Feature Extraction

The features extracted from the images in this paper included the morphology features that have been extracted in [14] [25-26]. These features include: nucleus area, cytoplasm area, nucleus to cytoplasm ratio, nucleus gray level, cytoplasm gray level, nucleus shortest diameter, nucleus longest diameter, nucleus elongation, nucleus roundness, cytoplasm shortest diameter, cytoplasm longest diameter, cytoplasm elongation, cytoplasm roundness, nucleus perimeter, cytoplasm perimeter, nucleus relative position, maxima in nucleus, minima in nucleus, maxima in cytoplasm, minima in cytoplasm. Due to the biological significance of the nucleus in cancer classification, 3 more geometric (solidity, compactness and eccentricity) and 6 (mean, standard deviation, variance, smoothness, energy and entropy) textual features were extracted from the nucleus giving a total of 29 features. A method based on prior knowledge has been developed that extracts features from segmented images using pixel level information.

2.1.7 Feature Selection

Feature selection also called variable/attribute selection is the process of selecting subsets of the extracted features that give the best classification results. The method used in this tool combines the simulated annealing with a wrapper approach. This has also been proposed in [26] for feature selection. The performance of the feature selection is evaluated using a double-strategy random forest algorithm [27]. The search for the global optimum function is guided by a fitness value [28]. When simulated annealing is finished, all the different subsets of features accepted are saved, their fitness is then compared to find the one that performs the best. The fitness value search was obtained with a wrapper where K-fold cross-validation was used to calculate the error on the classification algorithm. The wrapper method considers the selection of a set of features as the search problem [29]. Different combinations from the extracted features are prepared, evaluated and compared to other combinations. A predictive model is then used to evaluate a combination of features and assign a score based on model accuracy. The features that give the least error are used to construct the classifier.

2.1.8 Classification

Cervical cancer classification is a complex task; therefore, classification models are also usually complex. However, the more complex the classification model, the less the chance to find a model that fits well to the data [36]. This was handled by dividing the problem into subproblems and tackling them one by one using the defuzzification method in future selection. This approach is referred to as hierarchical [30] and has been reported to yield better classification results [26].
The performance of a classifier was evaluated using accuracy, false positive, false negative, sensitivity, specificity and ROC Area metrics. Sensitivity measures the proportion of actual positives that are correctly identified as such whereas specificity measures the proportion of

actual negatives that are correctly identified as such. Sensitivity and specificity are given by Equation 6.

$$Sensitivity\ (TPR) = \frac{TP}{TP+FN},\ Specificity\ (TNR) = \frac{TN}{TN+FP} \quad (6)$$

Where TP=True positives, FN=False negatives, TN=True negatives and FP=False positives

3. Results

3.1 Pap-Smear Automated Analysis Tool Description

The methods described above are executed via a Java graphical user interface (GUI) shown in Figure 3. The GUI has a panel where a pap-smear is loaded and a cytotechnician selects an appropriate method for debris removal, boundary detection and segmentation, after which features are extracted using the extract features button. Features of one intermediate and superficial cells are displayed. Once the features have been extracted, the cytotechnician presses the classify button and the tool emits a diagnosis (normal/abnormal) and classifies it to one of the 7 classes/stages (normal superficial, normal intermediate, normal columnar, mild dysplastic, moderate dysplastic, severe dysplastic or carcinoma in situ).

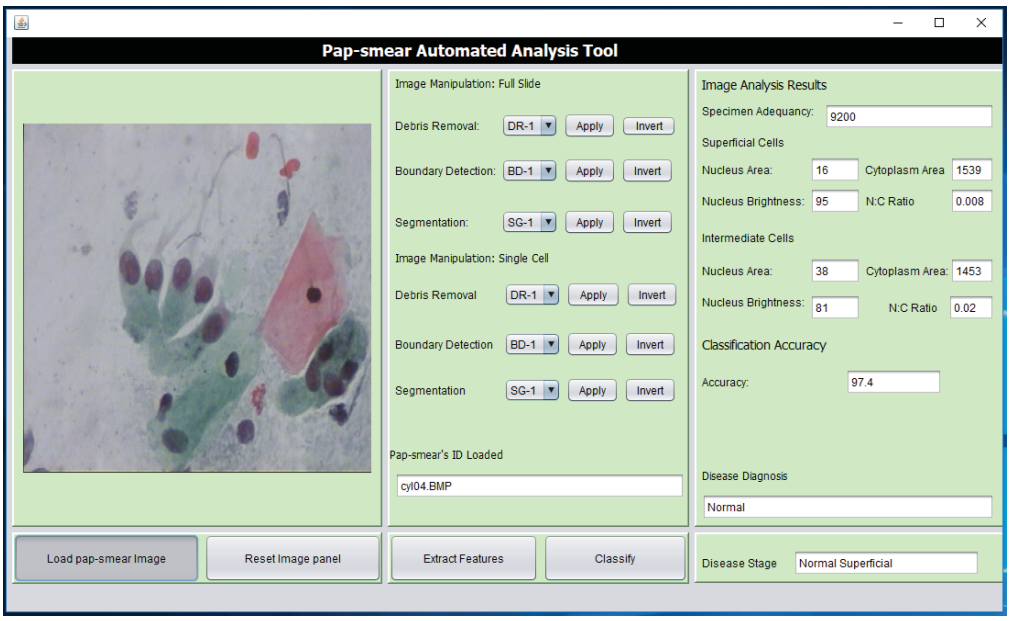

Figure 3: Graphical User Interface for the Pap-Smear Analysis Tool

3.2 Classification

A confusion matrix for the classification results on single cells is shown in Table 1. Of the 242 normal cells, 237 were correctly classified as normal and 5 were incorrectly classified as abnormal (1 normal intermediate and 4 normal columnar). Of the 675 abnormal cells, 674 were correctly classified as abnormal and 1 was incorrectly classified as normal. The overall accuracy, sensitivity and specificity of the classifier on this dataset was 99.35%, 99.85% and 97.93% respectively. A false negative rate (FNR), false positive rate (FPR) and classification error of 0.15%, 2.07% and 0.65% respectively were obtained.

Table 1: Cervical Cancer Classification Results from Single Cells

Abnormal		**Normal**	
False Negative	1	True Negative	237
True Positive	674	False Positive	5
Total	675	Total	242

Furthermore, the classifier was evaluated on a dataset of 60 full pap-smear images (30 normal and 30 abnormal pap-smear images) that had been prepared and classified by a cytotechnologist as normal or abnormal at Mbarara Regional Referral Hospital. Of the 30 normal pap-smears, 27 were correctly classified as normal and three were incorrectly classified as abnormal. All the 30 abnormal slides were correctly classified as abnormal. The overall accuracy, sensitivity and specificity of the tool on this dataset was 95.00%, 100% and 90.00%, respectively. A false negative rate, false positive rate and classification error of 0.00%, 10.00% and 5.00%, respectively were obtained as shown in the confusion matrix in Table 2.

Table 2: Cervical Cancer Classification Results from Pap-Smear Cells

Abnormal		**Normal**	
False Negative	0	True Negative	27
True Positive	30	False Positive	3
Total	30	Total	30

A Receiver Operating Characteristic (ROC) curve was plotted to analyze how the classifier can distinguish between true positives and negatives. This was necessary because the classifier needs to not only correctly predict a positive as a positive, but also a negative as a negative. This was obtained by plotting sensitivity (probability of predicting a real positive as positive), against 100-specificity (the probability of predicting a real negative as negative) as shown in Figure 4.

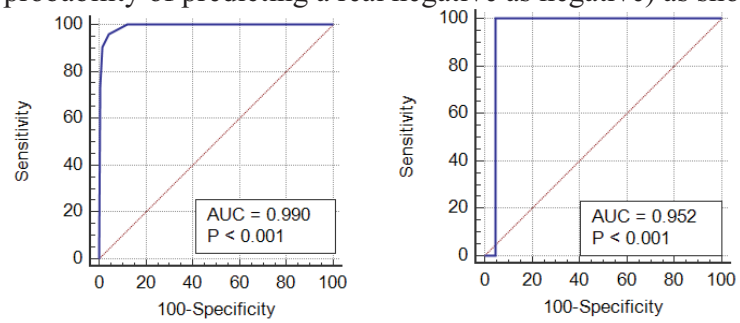

Figure 4: *ROC curve for the classifier performance on Herlev and Prepared pap-smear slide Datasets.*

The tool's performance was compared with state of art classification algorithms available in the literature. Results showed that our method outperforms some of the existing algorithms in classification accuracy (99.35%), specificity (97.93%) and sensitivity (99.85%), when applied to the Herlev benchmark pap-smear dataset as shown in Table 3.

Table 3: Comparison of the developed tool's performance with methods in [25], [31] and [32]

Method	***Method***	***Sensitivity***	***Specificity***	***Accuracy***
Zhang et al. [25]	Deep Convolutional Networks	98.2%	98.3%	98.3%
Bora et al. [32]	Ensemble Classifier	99.0%	89.7%	96.5%
Marinakis et al. [31]	Genetic Algorithm	98.5%	92.1%	96.8%
Our Approach	Enhanced Fuzzy C-means	99.85%	97.93%	99.35%,

4. Discussion

This research described the automated diagnosis and classification of cervical cancer from pap-smear images. The results in Table 1 are representative of the results that can be obtained from pre-processed smears, hence they provide a lower limit for the false positive and false negative rates on the cell level of 2.07% and 0.15% respectively. This implies that if the classier is presented with well-prepared slides then higher sensitivity values (>99%) can always be obtained as seen from the ROC curve in Figure 4. The classifier shows promising results in the classification of the cancerous cells with an overall accuracy of 99.35% on this dataset.

The results in Table 2 are representative of the results that can be obtained from a pap-smear slide from the pathology laboratory. A smear level false negative rate of 0.00% means that no abnormal cells were classified as normal, and therefore, the misclassification of an abnormal smear is unlikely. However, the 10.00% false positive rate means that some normal cells were classified as abnormal. However, confirmation tests are required to be carried out by the cytopathologist. The overall accuracy, sensitivity and specificity of the classifier on full pap-smear slides from the pathology lab was 95.00%, 100% and 90.00% respectively. The higher sensitivity of the classifier to cancerous cells could be attributed to the robustness of the feature selection method which selected strict nucleus constrained features that potentially indicate signs of malignancy.

5. Conclusion

In this paper, we have presented a tool for automated diagnosis and classification of cervical cancer from pap-smear images. The major contribution of this tool in a cervical cancer screening workflows is that it reduces on the time required by the cytotechnician to screen very many pap-smears by eliminating the obvious normal ones, hence more time can be put on the suspicious slides. Normally, a conventional pap-smear slide of size (5.7x2.5) mm obtained using a multi-head Olympia microscope may contain around 5,000-12,000 cells and it may take 5-10 minutes for manual analysis. The proposed tool has the capability of analyzing 1-2 smears per minute. With increased computer speed, efficiently written programs and implementation of this project using deep learning the potential to increase the number of smears analysed per minute. The evaluation and testing conducted with the Herlev database and pap-smear slides from Mbarara Regional Referral Hospital prove the validity of the tool. The tool currently uses a chain of methods and this makes it computationally expensive. In the future we shall implement the tool using deep learning. More so, in the future work, we plan to include a cervical cancer risk factors assessment and evaluation model into the tool.

Acknowledgement

The authors are grateful to the African Development Bank- HEST project for funding this research. The authors are also grateful to Mr Abraham Birungi from Pathology department of Mbarara University of Science and Technology, Uganda for providing support with pap-images.

Conflict of Interests

This paper has the assent of all co-authors and the authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] L.A. Torre, F. Bray, R.L. Siegel, J. Ferlay, J. Lortet-tieulent, A. Jemal, Global Cancer Statistics, 2012, CA a Cancer J. Clin. 65 (2015) 87–108. doi:10.3322/caac.21262.

[2] A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA Cancer J Clin. (2014). doi:10.3322/caac.20107.

[3] C. Nakisige, M. Schwartz, A.O. Ndira, Cervical cancer screening and treatment in Uganda, Gynecol. Oncol. Reports. 20 (2017) 37–40. doi:10.1016/j.gore.2017.01.009.

[4] W. William, A. Ware, A.H. Basaza-Ejiri, J. Obungoloch, A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images, Comput. Methods Programs Biomed. 164 (2018) 15–22. doi:10.1016/J.CMPB.2018.05.034.

[5] Y. Xue, S. Chen, J. Qin, Y. Liu, B. Huang, H. Chen, Application of deep learning in automated analysis of molecular images in cancer: A survey, Contrast Media Mol. Imaging. (2017). doi:10.1155/2017/9512370.

[6] S.F. Patten, J.S.J. Lee, D.C. Wilbur, T.A. Bonfiglio, T.J. Colgan, R.M. Richart, H. Cramer, S. Moinuddin, The AutoPap 300 QC system multicenter clinical trials for use in quality control rescreening of cervical smears: I. A prospective intended use study, Cancer. (1997). doi:10.1002/(SICI)1097-0142(19971225)81:6<337::AID-CNCR7>3.0.CO;2-I.

[7] T.J. O'Leary, M. Tellado, S.B. Buckner, I.S. Ali, A. Stevens, C.W. Ollayos, PAPNET-assisted rescreening of cervical smears: Cost and accuracy compared with a 100% manual rescreening strategy, J. Am. Med. Assoc. (1998). doi:10.1001/jama.279.3.235.

[8] E. Bengtsson, P. Malm, Screening for cervical cancer using automated analysis of PAP-smears, Comput. Math. Methods Med. (2014). doi:10.1155/2014/842037.

[9] W.E. Tolles, R.C. Bostrom, AUTOMATIC SCREENING OF CYTOLOGICAL SMEARS FOR CANCER: THE INSTRUMENTATION, Ann. N. Y. Acad. Sci. (1956). doi:10.1111/j.1749-6632.1956.tb32131.x.

[10] N. Tanaka, T. Ueno, H. Ikeda, A. Ishikawa, K. Yamauchi, Y. Okamoto, S. Hosoi, CYBEST Model 4. Automated cytologic screening system for uterine cancer utilizing image analysis processing., Anal. Quant. Cytol. Histol. (1987).

[11] D.J. Zahniser, P.S. Oud, M.C.T. Raaijmakers, G.P. Vooys, R.T. Van de Walle, Field test results using the bioPEPR cervical smear prescreening system, Cytometry. (1980). doi:10.1002/cyto.990010305.

[12] R. Erhardt, E.R. Reinhardt, W. Schlipf, W.H. Bloss, FAZYTAN: a system for fast automated cell segmentation, cell image analysis and feature extraction based on TV-image pickup and parallel processing., Anal. Quant. Cytol. (1980).

[13] J.H. Tucker, O.A. Husain, Trials with the cerviscan experimental prescreening device on polylysine-prepared slides, Anal Quant Cytol. (1981).

[14] J. Jantzen, J. Norup, G. Dounias, B. Bjerregaard, Pap-smear Benchmark Data For Pattern Classification, Proc. NiSIS 2005 Nat. Inspired Smart Inf. Syst. (2005) 1–9.

[15] P. Zamperoni, Image Enhancement, Adv. Imaging Electron Phys. (1995). doi:10.1016/S1076-5670(08)70006-5.

[16] K. Zuiderveld, Contrast Limited Adaptive Histogram Equalization, in: Graph. Gems, 1994. doi:10.1016/B978-0-12-336156-1.50061-6.

[17] C. Kanan, G.W. Cottrell, Color-to-grayscale: Does the method matter in image recognition?, PLoS One. (2012). doi:10.1371/journal.pone.0029740.

[18] P. Malm, B.N. Balakrishnan, V.K. Sujathan, R. Kumar, E. Bengtsson, Debris removal in Pap-smear images, Comput. Methods Programs Biomed. (2013). doi:10.1016/j.cmpb.2013.02.008.

[19] R. Montero, State of the art of compactness and circularity measures, Int. Math. Forum. (2009).

[20] W.Y. Kim, Y.S. Kim, Region-based shape descriptor using Zernike moments, Signal Process. Image Commun. (2000). doi:10.1016/S0923-5965(00)00019-9.

[21] E. Nadernejad, S. Sharifzadeh, H. Hassanpour, Edge Detection Techniques: Evaluations and Comparisons, Appl. Math. (2008). doi:10.1007/s10404-016-1785-3.

[22] B. Green, Canny edge detection, Retrieved: March. (2009). doi:10.1109/TPAMI.2005.173.

[23] W. Niblack, An introduction to digital image processing, Prentice-Hall Englewood Cliffs, 1986.

[24] R.M. Haralick, S.R. Sternberg, X. Zhuang, Image Analysis Using Mathematical Morphology, IEEE Trans. Pattern Anal. Mach. Intell. (1987). doi:10.1109/TPAMI.1987.4767941.

[25] L. Zhang, L. Lu, I. Nogues, R.M. Summers, S. Liu, J. Yao, DeepPap: Deep convolutional networks for cervical cell classification, IEEE J. Biomed. Heal. Informatics. (2017). doi:10.1109/JBHI.2017.2705583.

[26] E. Martin, J. Jantzen, Pap-Smear Classi cation, Master's Thesis, Tech. Univ. Denmark Oersted-DTU.

(2003).
[27] L. Breiman, Random forest, Mach. Learn. (2001). doi:10.1016/j.compbiomed.2011.03.001.
[28] A. Dekkers, E. Aarts, Global optimization and simulated annealing, Math. Program. (1991). doi:10.1007/BF01594945.
[29] S. Das, Filters, wrappers and a boosting-based hybrid for feature selection, Engineering. (2001).
[30] C. Vens, J. Struyf, L. Schietgat, S. Džeroski, H. Blockeel, Decision trees for hierarchical multi-label classification, Mach. Learn. (2008). doi:10.1007/s10994-008-5077-3.
[31] Y. Marinakis, G. Dounias, J. Jantzen, Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification, Comput. Biol. Med. 39 (2009) 69–78. doi:10.1016/j.compbiomed.2008.11.006.
[32] K. Bora, M. Chowdhury, L.B. Mahanta, M.K. Kundu, A. Kumar Das, A.K. Das, Automated Classification of Pap Smear Image to Detect Cervical Dysplasia, Comput. Methods Programs Biomed. (2017). doi:10.1016/j.cmpb.2016.10.001.