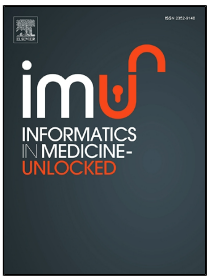
# Accepted Manuscript




Cervical Cancer Classification from Pap-smears using an Enhanced Fuzzy C-Means Algorithm

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

| | |
|---|---|
| PII: | S2352-9148(18)30247-8 |
| DOI: | 10.1016/j.imu.2019.02.001 |
| Reference: | IMU 157 |
| To appear in: | *Informatics in Medicine Unlocked* |
| Received Date: | 18 December 2018 |
| Accepted Date: | 11 February 2019 |

Please cite this article as: Wasswa William, Andrew Ware, Annabella Habinka Basaza-Ejiri, Johnes Obungoloch, Cervical Cancer Classification from Pap-smears using an Enhanced Fuzzy C-Means Algorithm, *Informatics in Medicine Unlocked* (2019), doi: 10.1016/j.imu.2019.02.001

# Cervical Cancer Classification from Pap-smears using an Enhanced Fuzzy C-Means Algorithm

Wasswa William[1], Andrew Ware (PhD)[2], Annabella Habinka Basaza-Ejiri (PhD)[3], Johnes Obungoloch (PhD)[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]College of Computing and Engineering, St. Augustine International University, Kampala, Uganda*
*Tel: +256772571444, Email: aejiri@saiut.ac.ug*

## Abstract

Globally, cervical cancer ranks as the fourth most prevalent cancer affecting women. However, it can be successfully treated if detected at an early stage. The Pap smear is a good tool for initial screening of cervical cancer, but there is the possibility of error due to human mistake. Moreover, the process is tedious and time-consuming. The objective of this study was to mitigate the risk of mistake by automating the process of cervical cancer classification from Pap smear images. In this research, contrast local adaptive histogram equalization was used for image enhancement. Cell segmentation was achieved through a Trainable Weka Segmentation classifier, and a sequential elimination approach was used for debris rejection. Feature selection was achieved using simulated annealing integrated with a wrapper filter, while classification was achieved using a fuzzy c-means algorithm.

The evaluation of the classifier was carried out on three different datasets (single cell images, multiple cell images and Pap smear slide images from a pathology unit). An overall classification accuracy, sensitivity and specificity of '98.88%, 99.28% and 97.47%', '97.64%, 98.08% and 97.16%' and '96.80%, 98.40% and 95.20%' were obtained for each dataset respectively. The higher accuracy and sensitivity of the classifier was attributed to the robustness of the feature selection method that was utilized to select cell features that would improve the classification performance, and the number of clusters used during defuzzification and classification. The evaluation and testing conducted confirmed the rationale of the approach taken, which is based on the premise that the selection of salient features embeds sufficient discriminatory information that leads to an increase in the accuracy of cervical cancer classification. Results show that the method outperforms many of the existing algorithms in terms of the false negative rate (0.72%), false positive rate (2.53%), and classification error (1.12%), when applied to the DTU/Herlev benchmark Pap smear dataset. The approach articulated in this paper is applicable to many Pap smear analysis systems, but is particularly pertinent to low-cost systems that should be of significant benefit to developing economies.

**Keywords:** *Pap-smear, Cervical Cancer, Fuzzy-C Means*

## 1.0 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 cancers detected in the late stages [2]. Over 85% of cervical cancer cases occur in less developed countries of which the highest incidences are in Africa, with Uganda being ranked seventh among the countries with the highest incidences of cervical cancer. Over 85% of those diagnosed with the disease in Uganda die from it [3]. This is attributed to lack of awareness of the disease aggravated by limited access to screening and health services. Regular Pap smear screening is the most successful and effective method in medical practice to facilitate the early detection and screening of cervical cancer. However, the manual analysis of 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 for each patient screened. Human visual grading of microscopic biopsy images tends to be subjective and inconsistent [4]. 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 [5-7].

### 1.1 Computer-assisted Pap smear analysis

Since the 1960's, numerous projects have developed computer-assisted Pap smear analysis systems leading to a number of commercial products, such as AutoPap 300 QC (NeoPath, Redmond, WA, USA) [8] and the PapNet (Neuromedical Systems Inc., Suffern, NY, USA) [9] which were approved by the United States Food and Drug Administration (FDA). However, these have had limited impact on cervical cancer screening in countries with less developed economies. Ever since the first appearance of computers, significant development efforts have been exerted to try to supplement or replace the human visual inspection of Pap smears with computer-aided analysis. However, the problem turned out to be more difficult than envisaged. Computer-assisted Pap smear analysis has proved to be a complex process, which comprises the following stages.

#### 1.1.1 Image acquisition

Digitization of Pap smear images is a complex task [10]. A Pap smear prepared using the conventional method covers a surface area of about 10 $cm^2$. A typical Pap smear contains between 10,000 and 100,000 cells with a typical cell having a diameter of about 35 μm [11]. This implies that under ideal conditions 100,000 cells could be packed in an area of 1 $cm^2$. To obtain reliable data for classification, the cells should be digitized at a maximum resolution (which is about 0.25 μm pixel size) [11]. This type of resolution is quite difficult to achieve with conventional imaging modalities. Hence, several Pap smear image acquisition techniques have been developed to provide efficient high-resolution digital Pap smear images. Three examples of such acquisition techniques are:

- Flying spot scanners. These measure one point at a time either by moving the entire sample in a raster fashion or by moving the measuring point or the illumination source over the sample [12].
- Continuous motion imaging. The single line integration systems use a line of a photosensor which is moved orthogonally to the line directly across the sample [13]. For better results, the speed with which the array is swept across the sample, or the sample moving below the array, needs to be synchronized with electronic scanning speed.
- TV-scanners. These integrate light from a whole rectangular area at once. The rectangular area is moved in order to capture the entire area of the sample. However, with this technique, for each captured image the microscope sensor has to be moved, the image has to stabilize and focused, and light has to be optimized [14].

#### 1.1.2 Preprocessing

Preprocessing is needed for background extraction, and for noise/debris removal in the Pap smear images. It also helps to define the regions containing cells, or the regions without cells, in order to reduce the area to be searched. Furthermore, preprocessing helps in determining the colour model to be used during image analysis. Preprocessing techniques include contrast stretching, noise filtering and histogram equalization [15]. Malm et al. [16] proposed a sequential classification scheme focused on removing debris in Pap smear images before segmentation. Other preprocessing approaches have been proposed by several researchers [17-19].

#### 1.1.3 Segmentation

Segmentation is needed for the definition of the regions of interest (ROI) in the image and is foundational to an automated cervical cancer screening system. Effective image segmentation facilitates the extraction of meaningful information and simplifies the image data for later analysis. Poor segmentation leads to poor results during image analysis [19]. Most of the time, due to the fundamentally important role of nuclei in a cervical cancer cell, cytopathologists are interested in the evaluation of the nucleus and cytoplasm parameters in order to facilitate cell-based diagnosis screening. Hence, accurate nucleus and cytoplasm segmentation are paramount. There are several segmentation methods which have been applied to Pap smear images and these include water immersion, active contour models, Hough transform, seed-based region growing algorithm and moving k-means clustering [5, 20-23]. All of these methods are solving a puzzling problem that images obtained from Pap smears are difficult to segment because of the diversity of cell structures, the intensity variation of background and overlapping of cell clusters; hence, an efficient preprocessing technique is paramount prior to segmentation. Recently, Lili et al. [24] proposed a superpixel-based Markov random field (MRF) segmentation framework to segment the nucleus, cytoplasm and image background of cervical cell images. Srikanth et al. [25] presented a method based on Gaussian mixture models (GMM) combined with shape-based identification of nucleus to segment the nucleus and cytoplasm from cervical cells. Song et al. [26] proposed a multiscale convolutional network (MSCN) and graph-partitioning-based method for accurate segmentation of cervical cytoplasm and nuclei. Specifically, deep learning via the MSCN was explored to extract scale-invariant features, and then, segment regions centered at each pixel were defined by a graph partitioning method. Zhi et al. [27] presented an algorithm for accurately segmenting the individual cytoplasm and nuclei from a clump of overlapping cervical cells using a joint level set optimization on all detected nuclei and cytoplasm pairs. The optimization is constrained by the length and area of each cell, a prior on cell shape, the amount of cell overlap, and the expected gray values within the overlapping regions.

#### 1.1.4 Feature extraction

In a Pap smear analysis, a feature can be defined as a piece of information that is relevant for solving the computational task related to the Pap smear image analysis. Features extracted from the Pap smear images can be broadly classified as structural or textural features [28] and these include:

- Size and shape: This contains morphometric features that express the overall size and shape of a cell. Examples include position and orientation dependent features, geometric features (area, perimeter, longest and shortest diameter), contour features (curvature, bending energy, convex hull, elliptic deviation, and Fourier descriptors) and invariant moment features [29].
- Intensity. These features use the absolute intensity values in the image. Some of the intensity features include the largest/lowest density and different region intensity features.
- Texture: Textural features help to obtain quantifiable measures of overall local density variability within an object of interest. Examples of texture measures are gradient image features, Laplace image features, flat texture features, topological gradients, run-length and co-occurrence features.
- Structure: With structural features, each chromatin particle in the cell is considered to be an object. Features are extracted by describing the relationships between these objects. Examples of structure features include the nearest neighbourhood graph, the minimum spanning tree graph and the convex hull.

Due to the importance of feature extraction to any automated cervical cancer screening system, a number of cell features including the nucleus area, nucleus perimeter, nucleus roundness, cytoplasm area, and nucleus to cytoplasm ratio, have previously been utilized to help facilitate cervical cancer classification [17,30-32].

#### 1.1.5 Feature selection

Feature selection involves evaluating and optimizing the feature space used for the actual classification. Adding more features to a set will not always lead to a better separability, but could instead introduce noise to the different classes; hence all features used in classification should add information, which increases the separability between the different classes [33]. Feature selection techniques include:

- Multivariate statistics. This encompasses a number of procedures that can be used to analyze more than one statistical variable at a time. Examples include principal component analysis (PCA) and linear discriminant analysis (LDA) [34].

- Genetic Algorithms. These use a model of genetic evolution to attempt to find an optimal solution to some kind of classification problem [35].
- Clustering methods. These involve unsupervised and supervised classification of samples into clusters. The two most common clustering methods are k-means and hierarchical clustering.
- Bayesian methods. These are based on Bayes' theorem [19].
- Artificial Neural Networks. These try to mimic the way the brain computes information [36].
- Support Vector Machines. These aim to separate multiple clusters with a set of unique hyperplanes that have the greatest margin to the edge of each cluster [37].

### 1.1.6 Classification

The aim of any automated cervical cancer screening system is to determine whether a sample contains any evidence of cancer. The most common method for classification involves analyzing all cells using selected features and then classifying each cell as normal or suspicious [38]. Another approach is to mimic the classification methodology used by cytotechnologists [39]. This involves analyzing the sample based on several factors such as patterns in cell distribution, the frequency of cells and cell clusters, the occurrence of degenerated cells and cytoplasm, and the abundance of bare nuclei. A number of researchers have developed classification techniques for automated diagnosis of cervical cancer from Pap smear images.

J. Su et al. [40] proposed a method for automatic detection of cervical cancer from Pap smear images using a two-level cascade integration system of two classifiers. The results showed that the recognition rates for abnormal cervical cells were 92.7% and 93.2% when C4.5 classifier or logical regression classifier was used individually; while the recognition rate was significantly higher (95.6%) when the two-level cascade integrated classifier system was used.
M. Sharma et al. [41] used the K-Nearest-Neighbors (KNN) method to classify the stage of cervical cancer from Pap smear images. A classification accuracy of 82.9% with 5-fold cross-validation was achieved. R. Kumar et al. [42] proposed a framework for automated detection and classification of cervical cancer from microscopic biopsy images using biologically interpretable features. The K-nearest neighbor method was used for cervical cancer classification. Performance measures for accuracy, specificity and sensitivity of 92%, 94% and 81% were obtained. T. Chankong et al. [43] presented a method for automatic cervical cancer cell segmentation and classification using the fuzzy C-means (FCM) clustering technique. Validation with Artificial Neural Networks (ANN) yielded accuracies of 93.78% and 99.27% for the 7-class and 2-class problems, respectively. J. Talukdar et al. [44] presented a fuzzy clustering based image segmentation of Pap smear images of cervical cancer cells using the Fuzzy C-Means (FCM) Algorithm. Two random numbers were utilized to form the membership matrix for each pixel to guide clustering. Promising results were obtained using the pixel level segmentation. M. Sreedevi et al. [45] presented an algorithm based on an iterative thresholding method for segmentation of Pap smear images and classification of cervical cells as normal or abnormal, based on the area parameter of the nucleus. The features of the nucleus were extracted using regional properties, and cells were classified as normal if the nucleus area was less than 1635mm, and classified as abnormal otherwise. A sensitivity of 95% and specificity of 90% was achieved.
This paper presents a potent approach for the automated classification of cervical cancer from Pap smears using an enhanced fuzzy c-means algorithm. A sequential elimination method is proposed for debris removal, Trainable Weka Segmentation for cell segmentation, and an efficient approach for feature selection to generate a feature subset that minimises the classification error.

## 2.0 MATERIALS AND METHODS

### 2.1.1 Input Images.

Cervical cancer classification was achieved in our study through a sequential approach (depicted in Figure 1). The approach was assessed using two DTU/Herlev datasets. Dataset 1 consists of 917 single cells of Pap smear images prepared by Jantzen et al. [46]. The dataset contains Pap smear images taken with a resolution of 0.201μm/pixel by skilled cytopathologists using a microscope connected to a frame grabber. The images were segmented using CHAMP commercial software (developed by DIMAC Imaging systems) and then classified into seven classes [46]. Of these, 200 images were used for training and 717 images for testing. Dataset 2 consists of 497 full slide Pap smear images prepared by Norup et al. [47]. Of these, 200 images were used for training and 297 images for testing. Furthermore, the performance of the classifier was evaluated on samples of 98 Pap smears (49 normal and 49 abnormal) obtained from Mbarara Regional Referral Hospital (MRRH). 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 and 400 cells. Based on the opinions of the cytopathologists, 10,000 objects in images derived from the 98 different Pap smear slides were selected of which 500 were free lying cervical epithelial cells (250 normal cells from normal smears and 250 abnormal cells from abnormal smears) and the remaining 9,500 were debris objects. This Pap smear segmentation was achieved using the Trainable Weka Segmentation toolkit.

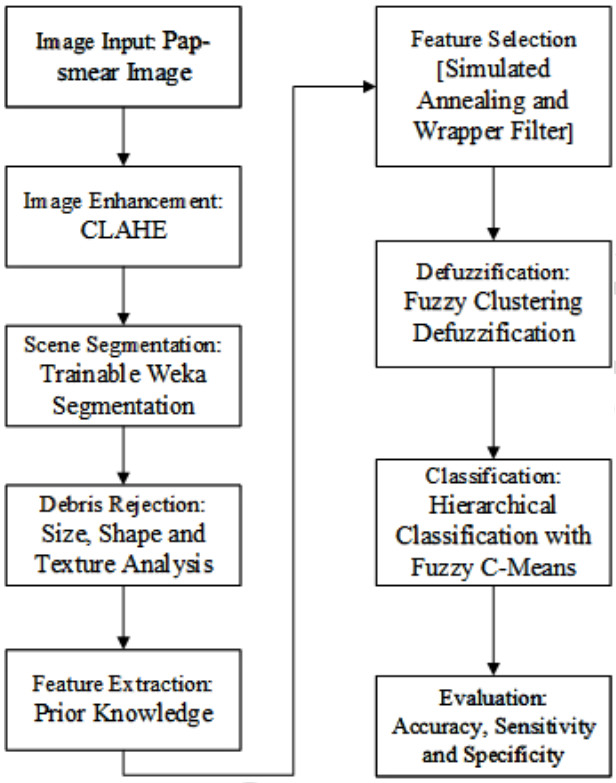



*Figure 1: The approach to achieve cervical cancer classification from Pap smear images*

### 2.1.2 Image Enhancement.

Image enhancement is very useful where the subjective quality of images is important for human and computer interpretation [48]. A contrast local adaptive histogram equalization (CLAHE) [49] was applied to the grayscale image. A clip limit value of 2.0 was determined to be appropriate for providing adequate image enhancement while preserving the dark features. Conversion to grayscale was achieved using a grayscale technique implemented using Equation 1 as defined in [50].

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

where R=Red, G=Green and B=Blue colour contributions to the new image.

A contrast local adaptive histogram equalization algorithm was implemented for image enhancement. This resulted in noticeable changes to the images (as shown in Figure 2) by adjusting image intensities where the darkening of the nucleus, as well as the cytoplasm boundaries, became easily identifiable using a clip limit of 2.0. CLAHE resulted in intensities in the images that were better distributed so as to facilitate further image analysis.

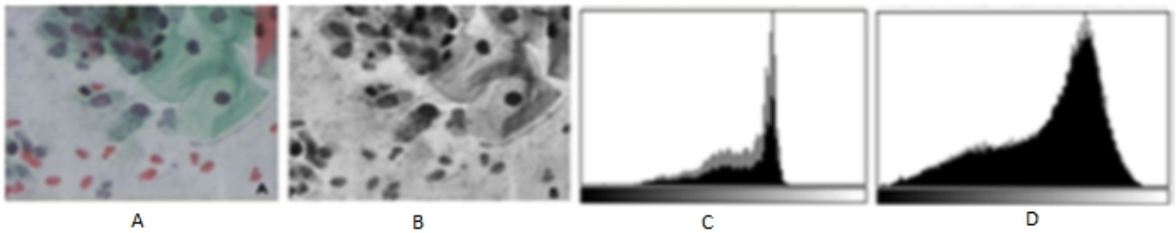



*Figure 2: Application of CLAHE (B) on the original Pap smear image (A). Original histogram (C) and Enhanced histogram (D)*

### 2.1.3 Pap smear segmentation.

The majority of cells observed in a Pap smear are, not surprisingly, cervical epithelial cells [39]. In addition, varying numbers of leukocytes, erythrocytes and bacteria are usually evident, while small numbers of other contaminating cells and microorganisms are sometimes observed. However, the Pap smear contains four major types of squamous cervical cells - superficial, intermediate, parabasal and basal - of which superficial and intermediate cells represent the overwhelming majority in a conventional smear; hence these two types are often used for a conventional Pap smear analysis [51]. A Trainable Weka Segmentation (TWS) was utilized to identify and segment the different objects on the slide. At this stage, a pixel level classifier was trained on cell nuclei, cytoplasm, background and debris identification with the help of a skilled cytopathologist, using the TWS toolkit [52]. This was achieved by drawing lines/selection through the areas of interest and assigning them to a particular class. The pixels under the lines were taken to be representative of the nuclei, cytoplasm, background and debris (depicted in Figure 3).

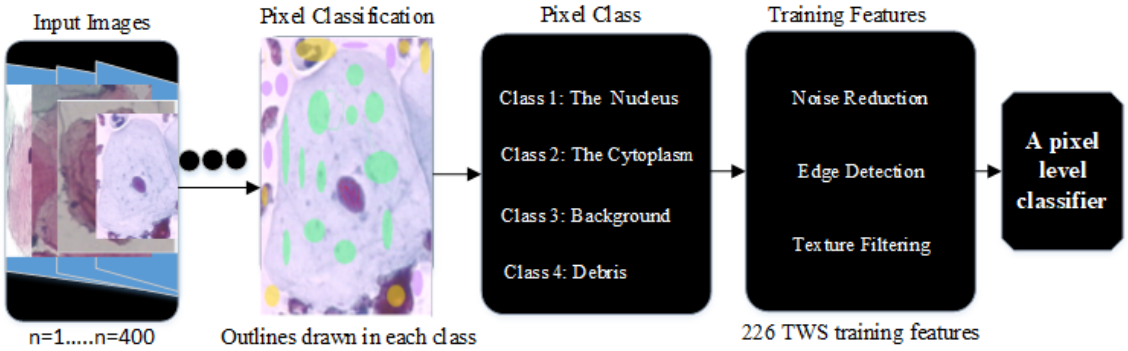



*Figure 3: Pap smear segmentation to achieve pixel level classification*

The outlines drawn within each class were used to generate a feature vector, $\vec{F}$ which was derived from the number of pixels belonging to each outline. The feature vector from each image (200 from Dataset 1 and 200 from Dataset 2) was defined by Equation 2.

$$\vec{F} = \begin{bmatrix} N_i \\ C_i \\ B_i \\ D_i \end{bmatrix}, \tag{2}$$

where $N_i$, $C_i$, $B_i$ and $D_i$ are the number of pixels from the nucleus, cytoplasm, background and debris of image $i$ as shown in Figure 4.

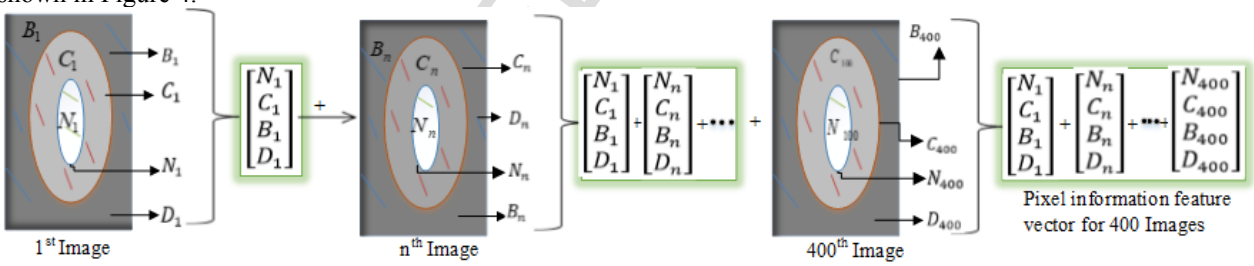



*Figure 4: Generation of the feature vector from the training images*

Each pixel extracted from the image represents not only its intensity but also a set of image features that contain much information including texture, borders, and colour, within a pixel area of 0.201µm$^2$. Choosing an appropriate feature vector for training the classifier was a great challenge and a novel task in the proposed approach. The pixel level classifier was trained using a total of 226 training features from TWS (as shown in Table 1).

*Table 1: Total number of pixels and training features for TWS segmentation*

| Class | Number of pixels | Number of training features |
|---|---|---|
| Nucleus | 6,538 | 226 |
| Cytoplasm | 9,668 | 226 |
| Background | 7,928 | 226 |
| Debris | 5,827 | 226 |

The classifier was trained using a set of TWS training features which included: (i) Noise Reduction: Kuwahara [53] and Bilateral filters [54] were used to train the classifier on noise removal. These have been reported to be excellent filters for removing noise whilst preserving the edges [54], (ii) Edge Detection: A Sobel filter [55], Hessian matrix [56] and Gabor filter [57] were used for training the classifier on boundary detection in an image, and (iii) Texture filtering: The mean, variance, median, maximum, minimum and entropy filters were used for texture filtering. The TWS performed extremely well in the segmentation of single cells and the full Pap smear image, as shown in Figure 5. This was very useful for the identification of cells and debris requiring further analysis.

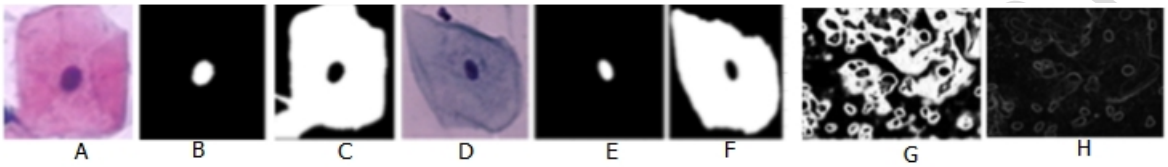



*Figure 5: Original single cell (A, D), Single cell nuclei segmentations (B, E), Single cell cytoplasm segmentations (C, F), original Pap smear image (G) and segmented Pap smear image (H)*

### 2.1.4 Debris Removal.

The main limitations of many of the existing automated Pap smear analysis systems is that they struggle to overcome the complexity of the Pap smear structures, by trying to analyze the slide as a whole, which often contain multiple cells and debris. This has the potential to cause the failure of the algorithm and requires higher computational power [58]. Samples are covered in artefacts - such as blood cells, overlapping and folded cells, and bacteria - that hamper the segmentation processes and generate a large number of suspicious objects. It has been shown that classifiers designed to differentiate between normal cells and pre-cancerous cells usually produce unpredictable results when artefacts exist in the Pap smear [16]. In this paper, a technique to identify cervix cells using a three-phase sequential elimination scheme (depicted in Figure 6) is presented.


*Figure 6: Three-phase sequential elimination approach for debris rejection*

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

**Size Analysis:** Size analysis is a set of procedures for determining a range of size measurements of particles [59]. The area is one of the most basic features used in the field of automated cytology to separate cells from debris. The Pap smear analysis is a well-studied field with much prior knowledge regarding cell properties [60]. However, one of the key changes with nucleus area assessment is that cancerous cells undergo a substantial increase in nuclear size [16]. Therefore, determining an upper size threshold that does not systematically exclude diagnostic cells is more difficult, 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. The pseudo code for the approach is shown in Equation 3.

$$If\ Area_{min} \leq Area_{roi} \leq Area_{max}\ then < foreground > else < Background >, \quad (3)$$

where $Area_{max} = 85{,}267\mu m^2$ and $Area_{min} = 625\mu m^2$ and $Area_{roi}$ is the area of the object being analysed.

The objects in the background are regarded as debris, and thus discarded from the image. Particles that fall between $Area_{min}$ and $Area_{max}$ are further analyzed during the next stages of texture and shape analysis (as shown in Figure 7).

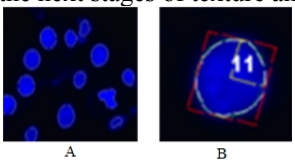



*Figure 7: Nuclei shape analysis (A) and Nucleus size analysis (B).*

**Shape Analysis:** The shape of the objects in a Pap smear is a key feature in differentiating between cells and debris [46]. There are a number of methods for shape description detection and these include region-based and contour-based approaches [61]. Region-based methods are less sensitive to noise but are more computationally intensive, whereas contour-based methods are relatively efficient to calculate but more sensitive to noise [16]. In this paper, a region-based method (perimeter2/area (P2A)) has been used [62]. The P2A descriptor was chosen on the merit that it describes the similarity of an object to a circle. This makes it well suited as a cell nucleus descriptor since nuclei are generally circular in their appearance. The P2A is also referred to as shape compactness and is defined by Equation 4.

$$c = \frac{p^2}{A}, \tag{4}$$

where $c$ is the value of shape compactness, $A$ is the area and $p$ is the perimeter of the nucleus. Debris were assumed to be objects with a P2A value (c) greater than 0.97 or less than 0.15.

**Texture analysis:** Texture is a very important characteristic feature that can differentiate between nuclei and debris. Image texture is a set of metrics designed to quantify the perceived texture of an image [63]. Within a Pap smear, the distribution of average nuclear stain intensity is much narrower than the stain intensity variation among debris objects [16]. This fact was used as the basis to remove debris from image intensities and colour information using Zernike moments (ZM) [64]. Zernike moments have utility for a variety of pattern recognition applications, are known to be robust with regard to noise, and have a good reconstruction power. In this work, the ZM as presented by Malm et al.[16] 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 $m\ x\ m$, given by Equation 5.

$$A_{nl} = \frac{n+1}{\pi}\sum_x\sum_y v^*_{nl}(r,\theta)I(x,y), \tag{5}$$

$v^*_{nl}(r,\theta)$ 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 [16].

### 2.1.5 Feature extraction

Feature extraction helps in converting the image to a format that is understandable to the classification algorithms. The success of the classification algorithm depends greatly on the correctness of the features extracted from the image. The cells in the Pap smears in the dataset used are split into seven classes based on characteristics such as size, area, shape and brightness of the nucleus and cytoplasm. The features extracted from the images included morphology features previously used by others [46,65]. 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, and minima in cytoplasm. Due to the biological significance of the nucleus in cancer classification, three geometric (solidity, compactness and eccentricity) and six textual features (mean, standard deviation, variance, smoothness, energy and entropy) were extracted from the nucleus, resulting in 29 features in total. A method based on prior knowledge has been implemented in MATLAB that extracts features from segmented images using pixel level information and mathematical functions.

### 2.1.6 Feature Selection

Feature selection (also called variable/attribute selection) is the process of selecting subsets of the extracted features that provide the best classification results. Among those features extracted, some might contain noise, while the chosen classifier may not utilize others. Hence, an optimum set of features has to be determined, possibly by trying all combinations. However, when there are many features, the possible combinations explode in number, and this increases the computational complexity of the algorithm. Feature selection algorithms are broadly classified into the filter, wrapper and embedded methods [66].

The method presented in this paper combines simulated annealing with a wrapper approach. This approach has been proposed elsewhere [65], but in this paper, the performance of the feature selection is evaluated using a double-strategy random forest algorithm [67]. Simulated annealing is a probabilistic technique for approximating the global optimum of a given function. The approach is well-suited for ensuring that the optimum set of features is selected. The search for the optimum set is guided by a fitness value [68]. When simulated annealing has completed, all of the different subsets of features are compared, and the fittest (that is, the one that performs the best) is selected. 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 implementation of simulated annealing is shown in Figure 8.

```
Begin simulated annealing
        T=chose starting point (typically around 0.2)
        N=number of iterations
        F=all features (Backward selection)
        Actual error=Error (F)
        For N=1 to Maximum iterations
                N=N-1
                Randomly chose to remove or add a feature to F
                New error=Error (F)
                If (e ((Actual error – New error)/T)>rand [0….1]
                        Accept and keep the new features in F
                        Actual error=New error
                Else
                        Keep the old features
                End if
        End for
End
```

*Figure 8: The Simulated Annealing Algorithm*

The wrapper method considers the selection of a set of features as the search problem [66]. 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 to assign a score based on model accuracy. The fitness error given by the wrapper is used as the error (F) by the simulated annealing algorithm shown in Figure 4. A fuzzy c-means algorithm was wrapped into a black box, from which an estimated error was obtained for the various feature combinations. Fuzzy c-means is a clustering algorithm that uses coefficients to describe how relevant a feature is to a cluster. The error estimate was obtained by *k*-fold cross validation as shown in Figure 9.

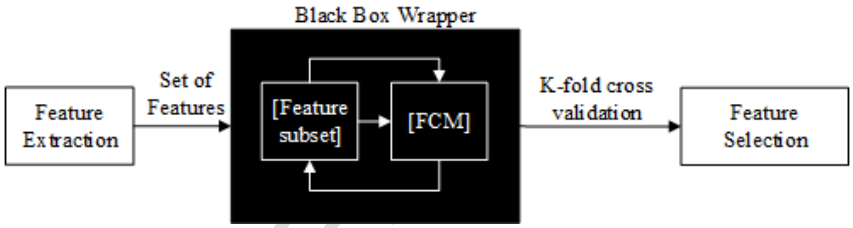



*Figure 9: The fuzzy c-means is wrapped into a black box from which an estimated error is obtained.*

### 2.1.7 The defuzzification

After feature selection, the fuzzy c-means algorithm does not tell us what information the clusters contain and how that information shall be used for classification. However, it defines how data points are assigned membership to the different clusters, and this fuzzy membership is used to predict the class of a data point [69]. A number of defuzzification methods exist [70-71]. However, in this paper, each cluster has a fuzzy membership (0-1) of all classes in the image. Training data are assigned to the cluster nearest to it. The percentage of training data of each class belonging to cluster *A* gives the cluster's membership, *cluster A = [i, j]* to the different classes, where *i* is the containment in cluster *A* and *j* in the other cluster. The intensity measure is added to the membership function for each cluster using a fuzzy clustering defuzzification algorithm. Fuzzy c-means allows data points in the dataset to belong to all of the clusters, with memberships in the interval (0 – 1) as shown in Equation 6.

$$m_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{2/(q-1)}}, \qquad (6)$$

where $m_{ik}$ is the membership for data point *k* to cluster center *i*, $d_{jk}$ is the distance from cluster center *j* to data point *k* and q $\in$[1…$\infty$] is an exponent that decides how strong the memberships should be. The FCM was implemented using the fuzzy toolbox in MATLAB.

A popular approach for defuzzification of the fuzzy partition is the application of the maximum membership degree principle, where data point *k* is assigned to class *m*, if and only if its membership degree $m_{ik}$ to cluster *i* is the largest. Genther et al. [72] proposed a defuzzification method using a fuzzy cluster partition in membership degree computation. Chuang et al. [73] proposed adjusting the membership status of every data point using the membership

status of neighbors. In the proposed approach, a defuzzification method based on Bayesian probability was used to generate a probabilistic model of the membership function for each data point, and the model was applied to the image to produce the classification information. The probabilistic model [74] is calculated as below:

1. Convert the possibility distributions in the partition matrix (clusters) into probability distributions.
2. Construct a probabilistic model of the data distributions as in [74].
3. Apply the model to produce the classification information for every data point using Equation 7.

$$\mathrm{P}(A_i \mid B_j) = \frac{P(B_j \mid A_i) * P(A_i)}{B_j}, \tag{7}$$

where $P(A_i), i = 0....c$ is the prior probability of $A_i$ which can be computed using the method in [74] and [75], where the prior probability is always proportional to the mass of each class.

The number of clusters to use was determined. This was necessary so that the built model can describe the data in the best possible way. If too many clusters are chosen, then there is a risk of overfitting the noise in the data. If too few clusters are chosen, then a poor classifier might be the result. Therefore, an analysis of the number of clusters against the cross-validation test error was performed. An optimal number of 25 clusters were attained and overtraining occurred above this number of clusters. Table 2 shows the results of the fuzziness exponent using different configurations.

*Table 2: Defuzzification fuzziness exponent calculation configurations*

| **2 Fold cross-validation with 60 reruns** | | **10 Fold cross-validation with 60 reruns** | |
|---|---|---|---|
| **Configuration** | **Fuzziness exponent** | **Configuration** | **Fuzziness exponent** |
| 05 Clusters | 1.2132 | 05 Clusters | 1.1402 |
| 10 Clusters | 1.2007 | 10 Clusters | 1.1263 |
| 15 Clusters | 1.1994 | 15 Clusters | 1.1036 |
| 20 Clusters | 1.1982 | 20 Clusters | 1.1010 |
| 25 Clusters | 1.1903 | 25 Clusters | 1.0930 |

The least defuzzification exponent (1.0930) from Table 2 was used to calculate the fitness error for feature selection. The errors were calculated using different cluster configurations as shown in Table 3.

*Table 3: Defuzzification fitness error calculation configurations*

| **2 Fold cross-validation with 60 reruns** | | **10 Fold cross-validation with 60 reruns** | |
|---|---|---|---|
| **Configuration** | **Fuzziness exponent** | **Configuration** | **Classification Error** |
| 05 Clusters | 12.1034 | 05 Clusters | 9.7322 |
| 10 Clusters | 11.1228 | 10 Clusters | 7.4561 |
| 15 Clusters | 10.1665 | 15 Clusters | 7.2189 |
| 20 Clusters | 9.8921 | 20 Clusters | 6.8923 |
| 25 Clusters | 9.6327 | 25 Clusters | 6.4210 |

The cluster configuration with the least error (as shown in Table 3) was used to select the features for classification. A total of 18 features out of the 29 features were selected for construction of the classifier. The selected features were: nucleus area (the actual number of pixels in nucleus; a pixel's area is 0.201μm$^2$); nucleus gray level (the average perceived brightness of the nucleus); nucleus shortest diameter (the shortest diameter a circle can have, when the circle is totally encircled around the nucleus); nucleus longest diameter (the longest diameter a circle can have, when the circle is totally encircled around the nucleus); nucleus perimeter (the length of the perimeter around the nucleus); maxima in nucleus (maximum value of number of pixels inside of a three pixel radius of the nucleus); minima in nucleus (minimum value of number of pixels inside of a three pixel radius of the nucleus); cytoplasm area (the actual number of pixels inside the cytoplasm); cytoplasm gray level (the average perceived brightness of the cytoplasm); cytoplasm perimeter (the length of the perimeter around the cytoplasm); nucleus to cytoplasm ratio (the relative size of the nucleus to the cytoplasm); nucleus eccentricity (the eccentricity of the ellipse that has the same second-moments as the nucleus region), nucleus standard deviation (the deviation of gray values of the nucleus region); nucleus variance (the variance value of the gray values inside the nucleus region); nucleus entropy (the entropy of gray values of the nucleus region); nucleus relative position (a measure of how well the nucleus is centred in the cytoplasm); nucleus mean (the mean gray values of the nucleus region) and nucleus energy (the energy of gray values of the nucleus region).

### 2.1.8 Classification evaluation methods

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 of finding a model that fits the data well [36]. This issue was handled by dividing the problem into subproblems and tackling them one-by-one using the defuzzification method described. This approach is referred to as the hierarchical approach [76] and has been reported to yield better classification results [65]. The performance of a classifier was evaluated using accuracy, false positive, false negative, sensitivity, specificity and ROC area metrics. Sensitivity (true positive rate) measures the proportion of actual positives that are correctly identified as such, whereas specificity (true negative rate) measures the proportion of actual negatives that are correctly identified as such. Sensitivity and specificity are given by Equation 8.

$$Sensitivity\ (TPR) = \frac{TP}{TP + FN}\,,\ pecificity\ (TNR) = \frac{TN}{TN + FP}\,, \tag{8}$$

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

## 3.0 RESULTS

### 3.1 Classification Accuracy

A confusion matrix for the classification results on the test single cells (Dataset 1 consisting of 717 single cells) is shown in Table 4. Of the 158 normal cells, 154 were correctly classified as normal and four were incorrectly classified as abnormal (one normal superficial, one intermediate and two normal columnar). Of the 559 abnormal cells, 555 were correctly classified as abnormal and four were incorrectly classified as normal (two carcinoma in situ cell, one moderate dysplastic and one mild dysplastic). The overall accuracy, sensitivity and specificity of the classifier on this dataset were 98.88%, 99.28% and 97.47% respectively. A False Negative Rate (FNR), False Positive Rate (FPR) and classification error of 0.72%, 2.53% and 1.12% respectively were obtained.

*Table 4: Cervical cancer classification results from single cells*

| **Abnormal** | | **Normal** | |
|---|---|---|---|
| False Negative | 4 | True Negative | 154 |
| True Positive | 555 | False Positive | 4 |
| Total | 559 | Total | 158 |

A Receiver Operating Characteristic (ROC) curve was plotted to analyze how the classifier can distinguish between the 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 ROC was obtained by plotting sensitivity (the probability of predicting a real positive as positive), against 100-specificity (the probability of predicting a real negative as negative) as shown in Figure 10.

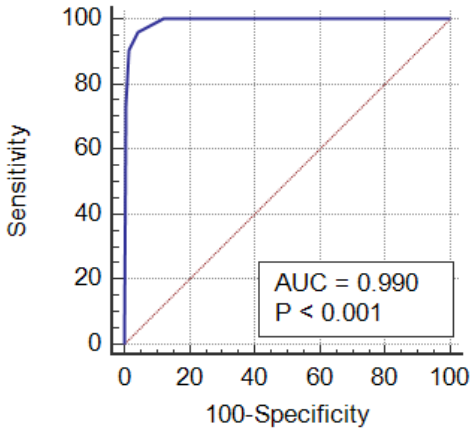



Figure 10: *ROC curve for the classifier performance on single cell images from DTU/Herlev dataset (Dataset 1).*

A confusion matrix for the classification results on test Pap smear slides (Dataset 2 of 297 full slide images) is shown in Table 5. Of the 141 normal slides, 137 were correctly classified as normal and four were incorrectly classified as abnormal. Of the 156 abnormal slides, 153 were correctly classified as abnormal and three were incorrectly classified

as normal. The overall accuracy, sensitivity and specificity of the classifier on this dataset were 97.64%, 98.08% and 97.16% respectively. A False Negative Rate (FNR), False Positive Rate (FPR) and classification error of 1.92%, 2.84% and 2.36% respectively were obtained.

*Table 5: Cervical cancer classification results from single cells*

| **Abnormal** | | **Normal** | |
|---|---|---|---|
| False Negative | 3 | True Negative | 137 |
| True Positive | 153 | False Positive | 4 |
| Total | 156 | Total | 141 |

Furthermore, the classifier was evaluated on a dataset of 500 single cell images (250 normal cells and 250 abnormal cells) that had been prepared and classified by a cytotechnologist as normal or abnormal from Mbarara Regional Referral Hospital. A confusion matrix for the classification results on this dataset is shown in Table 6. Of the 250 normal cells, 238 were correctly classified as normal and 12 were incorrectly classified as abnormal. Of the 250 abnormal cells, 246 were correctly classified as abnormal and four were incorrectly classified as normal. The overall accuracy, sensitivity and specificity of the classifier on this dataset were 96.80%, 98.40% and 95.20% respectively. A False Negative Rate (FNR), False Positive Rate (FPR) and classification error of 1.60%, 4.80% and 3.20% respectively were obtained.

*Table 6: Cervical cancer classification results from Pap smear cells*

| **Abnormal** | | **Normal** | |
|---|---|---|---|
| False Negative | 4 | True Negative | 238 |
| True Positive | 246 | False Positive | 12 |
| Total | 250 | Total | 250 |

The ROC curve analysis of the performance of the classifier on full slide Pap smears is shown in Figure 11.

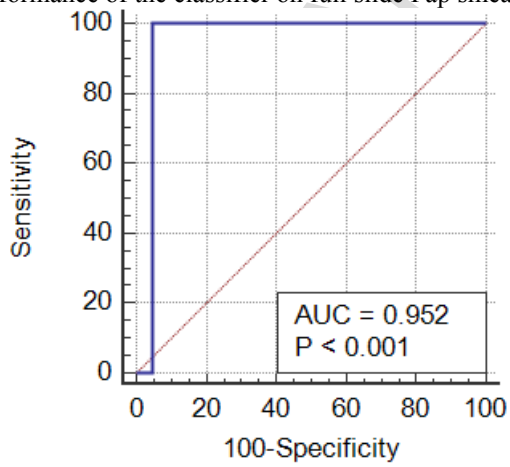



Figure 11: *ROC curve for the classifier performance on full slide Pap smears from* DTU/Herlev dataset (Dataset 2).

Furthermore, the classifier was evaluated on a dataset of 98 full Pap smear images (49 normal and 49 abnormal) that had been prepared and classified, as normal or abnormal, by a cytotechnologist at Mbarara Regional Referral Hospital. Of the 49 normal Pap smears, 45 were correctly classified as normal and four were incorrectly classified as abnormal. Of the 49 abnormal Pap smears, 47 were correctly classified as abnormal and two were incorrectly classified as normal. The overall accuracy, sensitivity and specificity of the classifier on this dataset were 93.88%, 95.92% and 91.84% respectively. A False Negative Rate (FNR), False Positive Rate (FPR) and classification error of 4.08%, 8.16% and 6.12% respectively were obtained.

The performance of the developed classifier was compared with results obtained by Martin et al. [65] and Norup et al. [47] on the same dataset (single cell dataset) using fuzzy based algorithms. Table 7 reports the performances of the two methods together with the results achieved by the proposed method. It was found that the proposed approach outperforms many of the existing fuzzy based classifiers in terms of FNR (0.15%), FPR (2.10%) and classification error (0.65%).

*Table 7: Comparison of the developed classifier's performance with methods in* [65] *and* [47]

| Method | *Method* | *FNR* | *FPR* | *Classification Error* |
|---|---|---|---|---|
| Martin et al. [65] | Supervised Hard C-means | 4.55% | 17.12% | 7.86% |
| Norup et al. [47] | Neuro-Fuzzy Inference Method | 3.85% | 10.74% | 5.67% |
| Our Approach | Enhanced Fuzzy C-means | 0.72% | 2.53% | 1.12% |

Furthermore, as shown in Table 8, the proposed approach was compared with contemporary classification algorithms documented in the relevant literature. Results show that the proposed method outperforms many of the documented algorithms in terms of classification cell level accuracy (98.88%), specificity (97.47%) and sensitivity (99.28%), when applied to the DTU/Herlev dataset benchmark Pap smear dataset (single cell dataset).

*Table 8: Comparison of the developed classifier's performance with methods in* [30], [32] and [77]

| Method | *Method* | *Sensitivity* | *Specificity* | *Accuracy* |
|---|---|---|---|---|
| Zhang et al. [30] | Deep Convolutional Networks | 98.2% | 98.3% | 98.3% |
| Bora et al. [77] | Ensemble Classifier | 99.0% | 89.7% | 96.5% |
| Marinakis et al. [32] | Genetic Algorithm | 98.5% | 92.1% | 96.8% |
| Proposed Approach | Enhanced Fuzzy C-means | 99.28% | 97.47% | 98.88%, |

### 3.2 Processing Time Analysis

This approach was tested on an Intel Core i5-6200U CPU@2.30GHz 8GB memory computer. Twenty randomly selected full Pap smear images were run through the algorithm, and the computational time was measured for both the individual steps and the overall duration. Average processing times for segmentation, debris removal, feature selection and classification were 38, 58, 23 and 42 seconds, respectively. Debris removal took longest (58 seconds), while feature selection was the shortest (23 seconds). The overall time taken per Pap smear image averaged 161 seconds, and was three minutes at most, demonstrating the feasibility for real-time diagnosis of the Pap smear.

## 4.0 DISCUSSION

This paper describes the automated analysis of Pap smear images to facilitate the classification of cervical cancer. Image enhancement using CLAHE makes the output of a processed image more suitable for image analysis. Unlike in many studies where CLAHE is applied on RGB Images [78-79], in the work documented here, CLAHE was applied to grayscale images as used in [80]. A Trainable Weka Segmentation (TWS) was utilized to provide a cheaper alternative to tools such as CHAMP. TWS produced excellent segmentations for the single images. However, segmentation results from full slide Pap smear images required more pre-processing before feature extraction. TWS has been used in many studies and its accuracy is largely dependent on the accuracy of training the pixel level classifier [81-82]. Increasing the training sample as reported by Maiora et al. [83] could improve the performance of the classifier. TWS's capability to produce good segmentation is due to its pixel level classification, where each pixel is assigned to a given class. However, the poor performance when segmenting the whole slide could be attributed to the small dataset used for building the segmentation classifier, as this was a manual process that involved annotation by an experienced cytopathologist.

Feature selection played an important role in this work, eliminating features that increased error in the classification algorithm. Eighteen out of the 29 extracted features were selected for classification purposes. It was noted that most of the features that added noise to the classifier were cytoplasmic features. This could be attributed to the difficulty in separating the cytoplasm from the background as opposed to the nucleus, which is darker [18]. Increasing the number of clusters during feature selection reduced the fuzziness exponent (Table 2). Similarly increasing the number of clusters using the fuzziness exponent of 1.0930 reduced the defuzzification fitness error to 6.4210 with 25 clusters and 10 fold cross validation (Table 3), less than that obtained by Martin et al. [65]. This implies that increasing the number of clusters reduces the defuzzification error computed by the defuzzification method presented in this paper, which is based on Bayesian probability to generate a probabilistic model of the membership function for each data point, and applying the model to the image to produce the classification information. An optimal number of 25 clusters was attained, and overtraining occurred when too many clusters (above 25) were used. A value of 25 clusters lower than 100 clusters could partly be because of the defuzzification method used. Its density measure works against overfitting by giving smaller clusters less influence as compared with larger clusters.

The results in Table 4 are representative of the results that can be obtained from pre-processed smears; hence, they provide a lower limit for the false negative and false positive rates on the cell level of 0.72% and 2.53% respectively. This implies that if the classifier is presented with well-prepared slides, then more sensitivity values (>99%) can always be obtained, as seen from the ROC curve in Figure 10. The classifier shows promising results in the classification of the cancerous cells, with an overall accuracy of 98.88% on this dataset (Dataset 1). The results in Table 5 are representative of the results that can be obtained from a Pap smear slide with many different types of cells. False negative and false positive rates on the smear level of 1.92% and 2.84% respectively were obtained; thus, the classifier still provided promising results on this dataset. The results in Table 6 are representative of the results that can be obtained from single cells from a Pap smear slide from the pathology laboratory. A false negative rate of 1.60% means that very few abnormal cells were classified as normal and, therefore, the misclassification of an abnormal smear is unlikely (specificity=95.20%). The 4.80% false positive rate means that a few normal slides were classified as abnormal (sensitivity=98.40%). The overall accuracy, sensitivity and specificity of the classifier on full Pap smear slides from the pathology lab were 93.88%, 95.92% and 91.84% respectively. The higher sensitivity of the classifier to cancerous cells could be attributed to the robustness of the feature selection method that selected strict nucleus constrained features that potentially indicate signs of malignancy. Despite the overall effectiveness of the approach, it does, however, involve many methods, making it computationally expensive. This in part curtails the full potency of the approach and therefore, in the near future, deep learning approaches will be explored to reduce the complexity.

## 5.0 CONCLUSION

The paper presents an approach for cervical cancer classification from Pap smears using an enhanced fuzzy c-means algorithm. A Trainable Weka Segmentation was proposed to achieve cell segmentation, and a three-step sequential elimination debris rejection approach was also proposed. Simulated annealing, coupled with a wrapper filter, was used for feature selection. The evaluation and testing, conducted with the DTU/Herlev datasets. and prepared pathological slides from Mbarara Regional Referral Hospital, confirmed the rationale of the approach of selecting 'good' features, to embed sufficient discriminatory information that can increase the accuracy of cervical cancer classification.

## Acknowledgement

The authors are grateful to the African Development Bank- HEST project for providing funds for this research, and the Commonwealth Scholarship Commission for the split-site scholarship to William Wasswa, in partnership with the University of Strathclyde. The authors are also grateful to Mr Abraham Birungi, from the Pathology Department at Mbarara University of Science and Technology, Uganda, for providing support with pap-images. Thanks also to Dr Mario Giardini, from the University of Strathclyde, for providing support with some image analysis.

## Conflict of Interests

This paper has the assent of all co-authors and the authors declare that there are no conflicts of interest. The authors declare that there are no conflict of interests 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. (2011). 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] H. Mabeya, K. Khozaim, T. Liu, O. Orango, D. Chumba, L. Pisharodi, J. Carter, S. Cu-Uvin, Comparison of conventional cervical cytology versus visual inspection with acetic acid among human immunodeficiency virus-infected women in Western Kenya, J. Low. Genit. Tract Dis. (2012). doi:10.1097/LGT.0b013e3182320f0c.

[5] 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.

[6] 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.

[7] A. Tareef, Y. Song, H. Huang, Y. Wang, D. Feng, M. Chen, W. Cai, Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling, Neurocomputing. (2017). doi:10.1016/j.neucom.2017.01.093.

[8] 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.

[9] 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.

[10] Y.F. Chen, P.C. Huang, K.C. Lin, H.H. Lin, L.E. Wang, C.C. Cheng, T.P. Chen, Y.K. Chan, J.Y. Chiang, Semi-automatic segmentation and classification of pap smear cells, IEEE J. Biomed. Heal. Informatics. 18 (2014). doi:10.1109/JBHI.2013.2250984.

[11] P. Malm, Automated cervical cancer screening through image analysis, (2014) 1–25.

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

[13] C. V. Biscotti, A.E. Dawson, B. Dziura, L. Galup, T. Darragh, A. Rahemtulla, L. Wills-Frank, Assisted primary screening using the automated ThinPrep Imaging System, Am. J. Clin. Pathol. (2005). doi:10.1309/AGB1MJ9H5N43MEGX.

[14] 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).

[15] S. Bhattacharyya, A Brief Survey of Color Image Preprocessing and Segmentation Techniques, J. Pattern Recognit. Res. (2011). doi:10.13176/11.191.

[16] 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.

[17] P. Sukumar, R.K. Gnanamurthy, Computer Aided Detection of Cervical Cancer Using Pap Smear Images Based on Adaptive Neuro Fuzzy Inference System Classifier, J. Med. Imaging Heal. Informatics. (2016). doi:10.1166/jmihi.2016.1690.

[18] Y.N. Dewi, D. Riana, T. Mantoro, Improving Naïve Bayes performance in single image pap smear using weighted principal component analysis (WPCA), in: 3rd Int. Conf. Comput. Eng. Des. ICCED 2017, 2018. doi:10.1109/CED.2017.8308130.

[19] G. Sun, S. Li, Y. Cao, F. Lang, Cervical cancer diagnosis based on random forest, Int. J. Performability Eng. (2017). doi:10.23940/ijpe.17.04.p12.446457.

[20] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nat. Methods. 9 (2012) 671–675. doi:10.1038/nmeth.2089.

[21] Y. Song, E.L. Tan, X. Jiang, J.Z. Cheng, D. Ni, S. Chen, B. Lei, T. Wang, Accurate cervical cell segmentation from overlapping clumps in pap smear images, IEEE Trans. Med. Imaging. 36 (2017) 288–300. doi:10.1109/TMI.2016.2606380.

[22] Y. Song, J.Z. Cheng, D. Ni, S. Chen, B. Lei, T. Wang, Segmenting overlapping cervical cell in Pap smear images, in: Proc. - Int. Symp. Biomed. Imaging, 2016. doi:10.1109/ISBI.2016.7493472.

[23] I. Muhimmah, R. Kurniawan, Indrayanti, Automated cervical cell nuclei segmentation using morphological operation and watershed transformation, in: Proceeding - 2012 IEEE Int. Conf. Comput. Intell. Cybern. Cybern. 2012, 2012: pp. 163–167. doi:10.1109/CyberneticsCom.2012.6381639.

[24] L. Zhao, K. Li, M. Wang, J. Yin, E. Zhu, C. Wu, S. Wang, C. Zhu, Automatic cytoplasm and nuclei segmentation for color cervical smear image using an efficient gap-search MRF, Comput. Biol. Med. (2016). doi:10.1016/j.compbiomed.2016.01.025.

[25] S. Ragothaman, S. Narasimhan, M.G. Basavaraj, R. Dewar, Unsupervised Segmentation of Cervical Cell Images Using Gaussian Mixture Model, in: IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., 2016. doi:10.1109/CVPRW.2016.173.

[26] Y. Song, L. Zhang, S. Chen, D. Ni, B. Lei, T. Wang, Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning, IEEE Trans. Biomed. Eng. 62 (2015) 2421–2433. doi:10.1109/TBME.2015.2430895.

[27] Z. Lu, G. Carneiro, A.P. Bradley, Automated Nucleus and Cytoplasm Segmentation of Overlapping Cervical Cells, in: K. Mori, I. Sakuma, Y. Sato, C. Barillot, N. Navab (Eds.), Med. Image Comput. Comput. Interv. -- MICCAI 2013, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013: pp. 452–460.

[28] I. Guyon, A. Elisseeff, Feature Extraction, Foundations and Applications: An introduction to feature extraction, Stud. Fuzziness Soft Comput. (2006). doi:10.1007/978-3-540-35488-8_1.

[29] Y. Mingqiang, K. Kidiyo, R. Joseph, A Survey of Shape Feature Extraction Techniques, in: Pattern Recognit. Tech. Technol. Appl., 2008. doi:10.5772/6237.

[30] 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.

[31] Y. Ramdhani, D. Riana, Hierarchical Decision Approach based on Neural Network and Genetic Algorithm method for single image classification of Pap smear, in: Proc. 2nd Int. Conf. Informatics Comput. ICIC 2017, 2018. doi:10.1109/IAC.2017.8280587.

[32] 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.

[33] G.K. Lakshmi, K. Krishnaveni, Feature extraction and feature set selection for cervical cancer diagnosis, Indian J. Sci. Technol. (2016). doi:10.17485/ijst/2016/v9i19/93881.

[34] P. Sukumar, S. Ravi, Computer aided detection and classification of Pap smear cell images using principal component analysis, Int. J. Bio-Inspired Comput. 11 (2018) 257–266. doi:10.1504/IJBIC.2018.092746.

[35] P. Mitra, S. Mitra, S.K. Pal, Staging of cervical cancer with soft computing, IEEE Trans. Biomed. Eng. (2000). doi:10.1109/10.846688.

[36] R.L. Becker, Applications of neural networks in histopathology., Pathologica. (1995).

[37] J.Y. Zhang, Y.X. Liu, Cervical cancer detection using SVM based feature screening, in: Med. IMAGE Comput. Comput. Interv. - MICCAI 2004, PT 2, Proc., 2004: pp. 873–880.

[38] N. Muñoz, F.X. Bosch, S. de Sanjosé, R. Herrero, X. Castellsagué, K. V. Shah, P.J.F. Snijders, C.J.L.M. Meijer, Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer, N. Engl. J. Med. (2003). doi:10.1056/NEJMoa021641.

[39] H. zur Hausen, Papillomaviruses and cancer: from basic studies to clinical application, Nat. Rev. Cancer. (2002). doi:10.1038/nrc798.

[40] J. Su, X. Xu, Y. He, J. Song, Automatic Detection of Cervical Cancer Cells by a Two-Level Cascade Classification System., Anal. Cell. Pathol. (Amst). 2016 (2016) 9535027. doi:10.1155/2016/9535027.

[41] M. Sharma, S. Kumar Singh, P. Agrawal, V. Madaan, Classification of Clinical Dataset of Cervical Cancer using KNN, Indian J. Sci. Technol. 9 (2016). doi:10.17485/ijst/2016/v9i28/98380.

[42] R. Kumar, R. Srivastava, S. Srivastava, Detection and Classification of Cancer from Microscopic Biopsy Images Using Clinically Significant and Biologically Interpretable Features, J. Med. Eng. 2015 (2015) 1–14. doi:10.1155/2015/457906.

[43] T. Chankong, N. Theera-Umpon, S. Auephanwiriyakul, Automatic cervical cell segmentation and classification in Pap smears, Comput. Methods Programs Biomed. 113 (2014) 539–556. doi:10.1016/j.cmpb.2013.12.012.

[44] J. Talukdar, C.K. Nath, P.H. Talukdar, 2013-Fuzzy Clustering Based Image Segmentation of Pap smear Images of Cervical Cancer Cell Using FCM Algorithm.pdf, 3 (2013) 460–462.

[45] I. Journal, C. Applications, T. Bangalore-, Papsmear Image based Detection of Cervical Cancer, 45 (2012) 35–40.

[46] 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.

[47] J. Norup, Classification of Pap-smear data by transductive neuro-fuzzy methods, Master's Thesis, Tech. Univ. Denmark Oersted-DTU. (2005) 71.

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

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

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

[51] N. Wentzensen, M. Von Knebel Doeberitz, Biomarkers in cervical cancer screening, Dis. Markers. (2007). doi:10.1155/2007/678793.

[52] I. Arganda-Carreras, V. Kaynig, C. Rueden, K.W. Eliceiri, J. Schindelin, A. Cardona, H.S. Seung, Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification, Bioinformatics. 33 (2017) 2424–2426. doi:10.1093/bioinformatics/btx180.

[53] K. Bartyzel, Adaptive Kuwahara filter, Signal, Image Video Process. 10 (2016) 663–670. doi:10.1007/s11760-015-0791-3.

[54] J.J. Francis, G. De Jager, The bilateral median filter, in: SAIEE Africa Res. J., 2005: pp. 106–111.

[55] N. Kanopoulos, N. Vasanthavada, R.L. Baker, Design of an Image Edge Detection Filter Using the Sobel Operator, IEEE J. Solid-State Circuits. 23 (1988) 358–367. doi:10.1109/4.996.

[56] O. Tankyevych, H. Talbot, P. Dokladal, Curvilinear morpho-Hessian filter, in: 2008 5th IEEE Int. Symp. Biomed. Imaging From Nano to Macro, Proceedings, ISBI, 2008: pp. 1011–1014. doi:10.1109/ISBI.2008.4541170.

[57] D. Dunn, W.E. Higgins, Optimal Gabor Filters for Texture Segmentation, IEEE Trans. Image Process. (1995). doi:10.1109/83.392336.

[58] J. Rissanen, Stochastic Complexity in Learning, J. Comput. Syst. Sci. (1997). doi:10.1006/jcss.1997.1501.

[59] T. Blaschke, Object based image analysis for remote sensing, ISPRS J. Photogramm. Remote Sens. (2010). doi:10.1016/j.isprsjprs.2009.06.004.

[60] J.M.W. Slack, Molecular Biology of the Cell, in: Princ. Tissue Eng. Fourth Ed., 2013. doi:10.1016/B978-0-12-398358-9.00007-0.

[61] D. Zhang, G. Lu, Generic Fourier descriptor for shape-based image retrieval, in: Proc. - 2002 IEEE Int. Conf. Multimed. Expo, ICME 2002, 2002. doi:10.1109/ICME.2002.1035809.

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

[63] W. Henry, Texture Analysis Methods for Medical Image Characterisation, in: Biomed. Imaging, 2010. doi:10.5772/8912.

[64] 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.

[65] E. Martin, J. Jantzen, Pap-Smear Classi cation, Master's Thesis, Tech. Univ. Denmark Oersted-DTU. (2003).

[66] S. Das, Filters, wrappers and a boosting-based hybrid for feature selection, Engineering. (2001).
[67] L. Breiman, Random forest, Mach. Learn. (2001). doi:10.1016/j.compbiomed.2011.03.001.
[68] A. Dekkers, E. Aarts, Global optimization and simulated annealing, Math. Program. (1991). doi:10.1007/BF01594945.
[69] T.P. Hong, C.Y. Lee, Induction of fuzzy rules and membership functions from training examples, Fuzzy Sets Syst. (1996). doi:10.1016/0165-0114(95)00305-3.
[70] S. Roychowdhury, W. Pedrycz, A survey of defuzzification strategies, Int. J. Intell. Syst. (2001). doi:10.1002/int.1030.
[71] P. (Institute for the S. of L. and E. Langley, Selection of Relevant Features in Machine Learning, Proc. AAAI Fall Symp. Relev. (1994). doi:10.1.1.43.4648.
[72] H. Genther, T.A. Runkler, M. Glesner, Defuzzification Based on Fuzzy Clustering, in: Proc. Third IEEE Conf. Fuzzy Syst. 1994. IEEE World Congr. Comput. Intell., 1994.
[73] M.A. Jaffar, N. Naveed, B. Ahmed, A. Hussain, A.M. Mirza, Fuzzy C-means clustering with spatial information for color image segmentation, in: 2009 3rd Int. Conf. Electr. Eng. ICEE 2009, 2009. doi:10.1109/ICEE.2009.5173186.
[74] T. Le, T. Altman, K.J. Gardiner, A probability based defuzzification method for fuzzy cluster partition, Proc. Intl'Conf. Artif. Intell. (2012) 1038–1043.
[75] J. Soto, A. Flores-Sintas, J. Palarea-Albaladejo, Improving probabilities in a fuzzy clustering partition, Fuzzy Sets Syst. (2008). doi:10.1016/j.fss.2007.08.016.
[76] 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.
[77] 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.
[78] K. Lakshmi, Automated Extraction of Cytoplasm and Nuclei from Cervical Cytology Images by Fuzzy Thresholding and Active Contours, Int. J. Comput. Appl. 73 (2013) 3–8.
[79] P.S. Maheswari, K. Jayasudha, R. Revathy, K. Yogalakshmi, Predicting the Severity of Cervical Cancer Using, Int. J. Res. Appl. Sci. Eng. Technol. 3 (2015) 141–144.
[80] J.J. de M. Sá Junior, A.R. Backes, A gravitational model for grayscale texture classification applied to the Pap smear database, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2015. doi:10.1007/978-3-319-23234-8_31.
[81] A.C. Dobens, L.L. Dobens, FijiWings: An Open Source Toolkit for Semiautomated Morphometric Analysis of Insect Wings, G3: Genes|Genomes|Genetics. 3 (2013) 1443–1449. doi:10.1534/g3.113.006676.
[82] M.A. Krueger, S.S. Huke, R.W. Glenny, Visualizing regional myocardial blood flow in the mouse, Circ. Res. 112 (2013). doi:10.1161/CIRCRESAHA.113.301162.
[83] J. Maiora, M. Graña, Abdominal CTA image analisys through active learning and decision random forests: Aplication to AAA segmentation, in: Proc. Int. Jt. Conf. Neural Networks, 2012. doi:10.1109/IJCNN.2012.6252801.

Authors declare that there is NO conflict of interest on this research.

This research has been approved by Mbarara University of Science and Technology Research Ethics Committee (MUREC) and supported by Mbarara Regional Referral Hospital Cancer Prevention Unit.

This research has been Funded by the African Development Bank and the Commonwealth Scholarship Commission and all consent to the publication of the Research Findings including on their websites.