which consists of 917 images of cropped single pap cells, was collected by the department of pathology at Herlev University Hospital and the department of automation at technical University of Denmark. The images were acquired at magnification of 0.201µm/pixel and average image size is 150x140 pixels.

3.1 Pre-processing

The Pap smear images usually have problems of inconsistence staining, poor contrast and unwanted noises. First the RGB images transformed into gray scale images, then the images pass through median filter to remove the unwanted noise, whereby the gray level of every pixel is replaced by the median of the intensity levels of the pixel neighborhood and preserving edge sharpness.

A. Linear Contrast Enhancement

In this work we used linear contrast enhancement algorithm to increase the contrast of the cervical cell. Linear contrast algorithm will spread out the distribution of the gray level over the whole possible range of the histogram (0 to 255). Significantly increasing their contrast whereby the biological changes of the nucleus and cytoplasm can be clearly seen also reduce the influence of the background on the cervical cell [9]

$$f_i(i,j) = 255 * [\frac{x_i(i,j)-b}{a-b}] \quad (1)$$

Where
$f_i(i,j)$ = new grey level of pixel i
$x_i(i,j)$ = Original grey level value
a=maximum grey level value of the image
b=minimum grey level value of the image

B. Background Segmentation

Background segmentations aims to extract or separate cells from background and remaining with cytoplasm and nucleus as or region of interest .we observed that cells and background regions have distinctive color in terms of brightness. Otsu's algorithm was used in segmenting cell from the background. The method is optimum in the sense that it maximizes between class variance, a well known measure used in statistical discriminate analysis. Otsu's methods has the important property that it is based entirely on computations performed on the histogram of an image, an easily obtainable 1-D array [10]
Otsu's algorithm may be summarized as follows

1. Compute the normalized histogram of the input image, Denote the components of the histogram by P_i i=0,1,2…..,L-1
2. Compute the cumulative sums, P_i(k) for K=0,1,2..L-1 using $P_i(k) = \sum_{i=0}^{L-1} P_i = 1 , P_i \geq 0$
3. Compute the cumulative means , m(k), for k=0,1,2….L-1 using $m(k) = \sum_{i=0}^{L-1} iP_i$
4. Compute the global intensity mean, m_G using $m_G = \sum_{i=0}^{L-1} iP_i$
5. Compute the between- class variance $\sigma_B^2(k)$ for k=0,1,2…,L-1 , using $\sigma_B^2(k) = \frac{[m_{GP_i(k)-m(k)]^2}}{P_i(k)[1-P_i(k)]}$
6. Obtain the Otsu threshold k^* as the value of k for which $\sigma_B^2(k)$ is maximum, if the maximum is not unique k^* by averaging the value of k corresponding to the various maxima detected.

   Obtain the separability measure η^* by evaluating $\eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2}$ at k=k^*

   Where σ_G^2= global variance.

3.2 Edge Detection

Edge detection is the approach used most frequently for segmenting images based on abrupt (local) change in intensity. A step edges involves a transition between two intensity levels occurring ideally over the distance of 1 pixel. The edge detection on cell is important since the shape and size of the cytoplasm and nucleus gives information about illness [3]. Edge detection based segmentation is frequently used technique which segments an image on basis of dissimilarity or heterogeneity with pixels or regions [4]. Edge detection steps which should be performed are

(1) Image smoothing for noise reduction
(2) Detection of the edge points
(3) Edge localization

In this paper we use Canny edge detector algorithm for the detection of the cytoplasm and nucleus boundary since the quality of the lines with regard to continuity, thinness and straightness is superior in canny image, also its only algorithm which capable of finding the best contours while eliminating all the edges associated with gray level matter in original image [10]. Canny edge algorithm consists of the following steps

1. Smoothing the input image with Gaussian filter
2. Compute the gradient magnitude and the angle of the images
3. Apply non-maxima suppression to the gradient magnitude image
4. Use double threshold and connectivity analysis to detect and link edges.

Let f(x, y) denote the input image and G(x, y) denote the Gaussian function

$$G(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

We form a smoothed image $f_s(x,y)$ by convolving G and f:

$$f_s(x,y) = G(x,y) * f(x,y) \quad (3)$$

The above equation is implemented using n x n Gaussian mask whose size must be specified. This operation is followed by computing the gradient magnitude and direction

$$M(x,y) = \sqrt{g_x^2 + g_y^2} \quad (4)$$

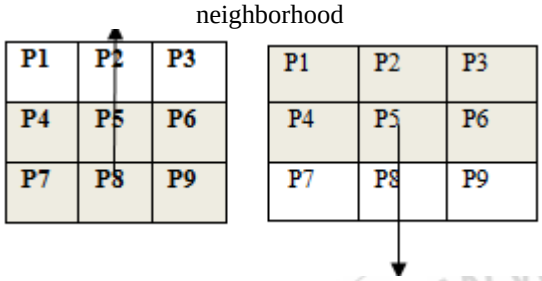
Any filter mask can be used to obtain g_x and g_y the filter mask also known as gradient operators or edge detectors.

$$\propto (x,y) = \tan^{-1}[\frac{g_y}{g_x}] \quad (5)$$

Whereby α (x, y) and M(x, y) are arrays of the same size as the image from which they are computed

Then non-maxima suppression is used to thin ridges from the images around local maxima by specify the number of discrete orientation of the edge normal (gradient vector). For example in 3 x 3 regions we can define four orientations. But we all know that every edge has two possible orientations.

Two possible orientations of horizontal edge in a 3x3 neighborhood



d_1, d_2, d_3 and d_4 Denote the four basic edge directions for 3 x 3 regions: horizontal $-45°$ and vertical $+45°$

We can formulate the following non-maxima suppression scheme for 3 x 3 region centered at every point (x, y) and α(x, y):

1. find the direction d_k that is close to α(x, y)
2. if the value of M(x, y) is less than the at least of its two neighbors along d_k

Let $g_N(x,y) = 0$ (suppression) otherwise

Let $g_N(x,y) = M(x,y)$

Where

$g_N(x,y)$ = The non-maxima suppression image.

Last step is to use double threshold where by Canny suggested that the ratio of the high to low threshold should be two or three to one

We can visualize the threshold operation as creating two additional images

$g_{NH}(x,y) = g_N(x,y) \geq T_H$ (6)

And

$g_{NL}(x,y) = g_N \geq T_L$ (7)

Where initially both $g_{NH}(x.y)$ and g_{NL} are set to zero, after thresholding g_{NH} will have fewer non-zero pixels than $g_{NL}(x,y)$because the letter image is formed with lower threshold.

We eliminate from $g_{NL}(x,y)$ the entire non-zero pixel from by letting

$g_{NL}(x,y) = g_{NL}(x,y) - g_{NH}(x,y)$ (8)

The non-zero pixels in $g_{NH}(x.y)$ and $g_{NL}(x,y)$ may be viewed as being strong and weak edge pixels respectively.

3.3. Nucleus Cytoplasm Contour Detection

NCC detection consists of three main steps which are gradient calculation, nuclear contour detection and last step is Cytoplasm detection. In this work we used canny detector to calculate gradient of the cell image since canny algorithm is capable of finding the best contour. Where the sensitivity was chosen to be 0.634 and the value of gamma was 6.56. Second, nuclear contour detection, since we know that for a normal cell nucleus is generally much smaller than its cytoplasm and the background of the cropped image. The method based on maximum gray gradient difference (MGLGD) was adopted for detection of the nucleus contour detection. The main idea behind the MGLGD is that the difference of the mean gray levels between the region inside and region outside the contour maximum at the contour of an object. In additional the gradient of the most pixels located on the contour are much larger than the gradients of the pixels. Finally the cytoplasm detection phase was used to determine the initial contour by moving ten pixels toward nucleus for detecting contour of the cell nucleus [11].

4. Results and Discussion

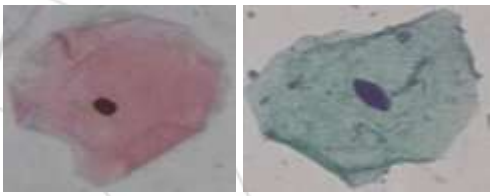


Figure 1: Pap smear normal cervical cells images

Figure 2: .Pap smear abnormal cell images

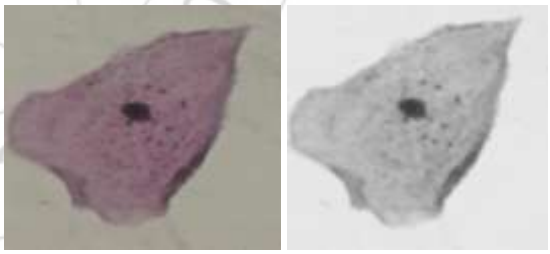


Figure 3: Original image and image after linear contrast enhancement respectively

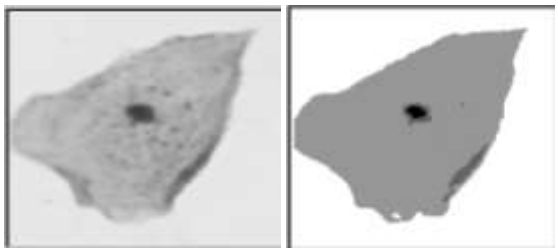


Figure 4: Median filtered cell image and background segmented cell image respectively.

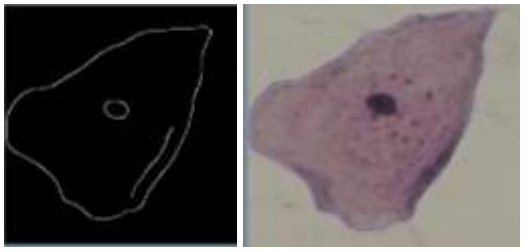


Figure 5: Gradient image obtained by applaying canny edge detector and detected NCCs respectively.

The first step in this proposed method of segmentation is removing of noise then segmenting cell from background by using Otsu's method which gives the best results by splitting the cells into three regions background,cytoplasm and nucleus in chinch boundary can be easly seen. Then canny algorithim was applied to the cell to obtain its smooth edge boundary of nucleus and cytoplasm. Lastly we use MGLGD method to detect and locate the boundary of the nucleus and cytoplasm.

From the figure 3 first row second column shows the contrast and sharpness of the image increase uniformly also the boundary edge was preserved since the original images shows poor contrast. Figure 4 shows image after passing through median filter for removing noise and unwanted content from the cell image, also shows the cell segmented from the background whereby background considered as white and cytoplasm region gray while the nucleus in black color this makes easy for the further analysis of the cell because boundary of the cytoplasm are well segmented from the background. The median filter gives us the super clean image compared to the original.

The figure 5 shows the gradient image of the cell after canny algorithm being applied on them, its shows smooth edge on both cells images which makes analysis of the cytoplasm area and nucleus area to the cell be easy task. Cytoplasm to nucleus ratio can be calculated also the shape feature of the cytoplasm and nucleus will be extracted for analysis, as we know the shape and size of the cytoplasm and nucleus has impact on the status of the cells. Size of a nucleus of the abnormal cells is always big and the shapes are not regular one. Hence it will results the N/C ratio to be large.

The figure 5 first row second column shows the nucleus and cytoplasm contours after applying MGLGD method for the detection of the nucleus and contour it. Results shows well contoured boundary of the cell cytoplasm and nucleus. Shortest and longest length of the cytoplasm can be calculated as well.

5. Conclusion

This work we have achieved to enhance and filter the images for the further analysis. We have seen how linear contrast enhancement works well in preserving the edge and sharpness of the cells images. We choose canny algorithm since it can gives smooth edge and boundary of the cytoplasm and nucleus. Then the boundary of the cytoplasm and nucleus was successfully detected and contoured. Further work is going on designing of best classifier for classification of cells into normal and abnormal with high accuracy.

References

[1] http://www.who.int
[2] http://screening.iarc.fr/cervicalindex.php
[3] M. Mohideen Fatima alias Niraimathi, Dr V. Seenivasagam, A hybrid Image Segmentation of a Cervical Cells by Bi-group enhancement and Scan line filling, International Journal of Computer Science and Information Technology and security,vol 2,No 2, April 2012,pp 368-375.
[4] M,-H. Tsai, Y,-K Chan, Z,-Z. Lin, S,-F. Yang-Mao, P, -C .Huang, Nucleus and Cytoplasm Contour Detector of cervical smear image, Pattern Recognition Letter 29(9) 2008, pp 1441-11453.
[5] Priyanka L.Ikhar, Mrs. Amolid D. Belsare, Determining cell nucleus and cytoplasm Boundaries Using Watershed Transform, International Conference on Electrical Engineering and Computer Science, 19th May 2013, Nagpur, pp 63-68
[6] S,-F. Yang-Mao, Y,-K. Chan, Y,-P, Chu ,Edge Enhancement Nucleus and Cytoplasm Detection of cervical smear images ,IEEE Transactions on systems, Man, and Cybernetics part B, Cybernetics 38(2),2008 pp 353-366.
[7] N.M Harand, S, Sadri, N. A. Moghaddam, R. Amirfattahi, An Automatic Method for Segmentation of epithelial cervical cells in Images of Thin Prep, Journal of medical systems 34(6)2010,pp1043-1054.
[8] M.E Plissiti, C. Nikou, A. Charchanti, Automated detection of cell Nuclei in Pap smear images Using Morphological Reconstruction and Clustering, IEEE Transaction on information technology in Biomedicine 15(2) 2011,pp 233-241
[9] Nor Ashidi Mat Isa, Nazahah Mustafa,Kamal Zuhairi Zamli, Mohd Yusoff Mashor, Improvement of Contrast Enhancement Technique for Cervical Cell of Pap Smear images by reducing the effects of unwanted background information.IEEE,2006..
[10] Rafael C. Gonzalez, Richard E. Woods Digital Image Processing, Third Edition, Pearson Education Asia ltd, Publishing House of Electronics Industry, Beijing, 2010
[11] Yung-Fu Chen, Po-Chi Huang, Ker-Cheng Lin, Hsuan-Huang Lin, Li-En Wang, Chung-Chuan Cheng, Tsung-Po Chen, Yung-Kuan Chan, and John Y. Chiang, Semiautomatic and Classification of Pap Smear Cells, IEEE, Journal of Biomedical and Health Informatics, 2013,pp1-15

Author Profile

Ayubu Hassan Mbaga received the B.E degree in Electronics and Communication Engineering from St Joseph University in Tanzania, Dar es salaam, Tanzania in 2010.He is currently working toward Msc degree in Signal and Information Processing at Tianjin University of Technology and Education, Tianjin, China. His research interests include Medical image processing, Pattern Recognition, Neural networks and learning machines.