Volume 2, Issue 11, November 2012 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Analysis of Malignant Cervical Cells Based on N/C Ratio Using Pap Smear Images

Dr. L. B. Mahanta*
Centre for Computational and Numerical studies
Institute of Advance Studies in Science and Technology

Kangkana Bora
Institute of Science and Technology
Gauhati University

Abstract*— This work presents an approach for the analysis of abnormal cervical cells based on nucleus/cytoplasmic (N/C) ratio which is one the most important morphological features based on which one can distinguish between normal and abnormal cervical cells. The proposed approach uses MATLAB to analyse the Pap smear images. Pap smear test is a simple harmless test that is used in diagnosis of cervical cancer. This may help pathologist in identification of cancer affected cells and help in prevention of cervical cancer.

***Keywords*—N/C ratio, Cervical cancer, Pap smear**

I. INTRODUCTION

Cervical cancer is a common cancer found among women after breast cancer. According to the survey, one woman dies every seven minutes of cervical cancer and by 2025, it is estimated to be one death in every 4.6 minutes. There has been a regular campaign against cervical cancer for the last 30 years in India, but this has had little impact on the morbidity and mortality from the disease, with India ranking fourth worldwide[1]. Government of India has undertaken several cancer control program but these measures have not been effective in reaching the rural regions due untrained manpower, lack of infrastructure, lack of awareness[2]. So, early detection of cervical cancer may take a very precious role in preventing this disease. One of the most important features of identifying the cancer affected cells is by analysing the N/C ratio. Manual processes in this field may not always satisfactory as it may be time consuming to visualize millions of cells and also it is error prone. So an automatic analysis may help pathologist to identify cancer affected cells correctly and with accuracy. In this approach 20 Pap images are taken for analysis which is collected from Dr B.Barooh Cancer Institute, Guwahati.

II. OVERVIEW OF CERVICAL CANCER

Cervical cancer is a cancer generally found in endocervix, ectocervix and transformation zone. The part close toward the uterus is called ectocervix and is covered by the squamous cells and the part next to the vagina is called endocervix and is covered by glandular cells. The place where these two cell meet is called the transformation zone. And most of the cervix cancer starts at this zone [3]. Fig 1gives an overview of the cervix part.

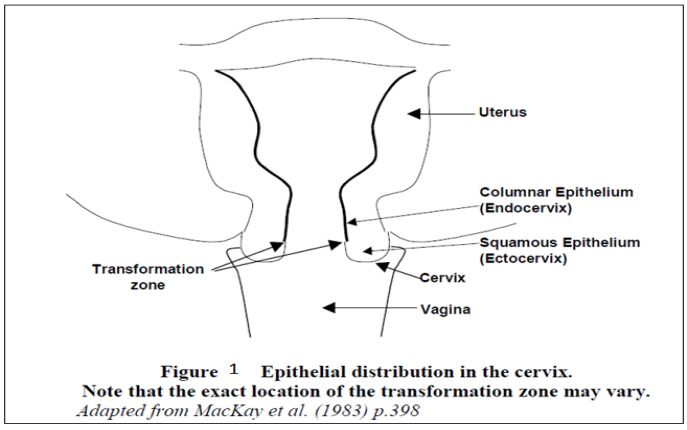


Figure 1 Epithelial distribution in the cervix.
Note that the exact location of the transformation zone may vary.
Adapted from MacKay et al. (1983) p.398

Fig. 1- overview of the cervix part.

Main causes of cervical cancer are –

- HPV (Human Papilloma Virus) infection
- Having multiple sex partners
- Giving birth before the age of 22
- Smoking
- Birth control pills.
- Low socio economic status
- Inadequate intake of folic acid

- Family history of cervical cancer

There are many morphological features are available based on which normal cells can be distinguished from abnormal cells. They may include nucleus area, cytoplasm area, nucleus/cytoplasm (N/C) ratio, perimeter of the nucleus, presence of hyperchromasia etc. Among them N/C ratio is one of the most important features.

III. N/C RATIO AND ABNORMALITY

The nuclear-cytoplasmic ratio (also variously known as the nucleus:cytoplasm ratio, nucleus-cytoplasm ratio, N:C ratio, or N/C) is a measurement used in cell biology. It is a ratio of the size (i.e., area) of the nucleus of a cell to the size of the cytoplasm of that cell [4].The N:C ratio indicates the maturity of a cell, because as a cell matures the size of its nucleus generally decreases.

So, one can find the N/C ratio as follows:

$$\text{N/C ratio} = (\text{Area of nucleus})/(\text{Area of the cytoplasm}) \quad (1)$$

That means if area of the nucleus and cytoplasm are calculated then from Eq (1) one can calculate the N/C ratio easily. In a normal cell, nucleus consumes very less amount of area than an abnormal cell's nucleus. And on the other hand cytoplasm of a normal cell consume larger amount of area than an abnormal cell. So if one find out the N/C ratio based on this concept then it can be explained that N/C ratio of a normal cell is less than N/C ratio of an abnormal cells. Fig 2 shows the area consumed by normal and abnormal cell's nucleus and cytoplasm. In our approach we tried to find out N/C ratio of nearly 100 individual cells obtained from the Pap smear images.

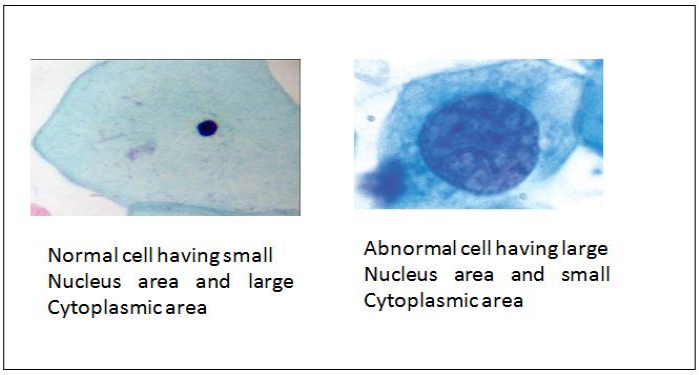


Fig 2 - Difference based area of nucleus and cytoplasm

IV. PROPOSED METHOD

In this approach firstly the individual cell is being cropped for further observations. There are three major methods are involved in our approach.

Step1 – Calculation of nucleus area
Step2 – Calculation of cytoplasm area
Step3 – Calculation of N/C ratio from the outputs obtained from step1 and step2 and its analysis

Step 1 – Calculation of nucleus area

The overview of the method used in this context is shown in fig3.

The step is to perform pre-processing operations on the input Pap smear image. The input image is a coloured image which is in RGB mode. As is known the colour images are difficult to process so we convert the image to greyscale image using rgb2gray function in matlab. Then we adjust the contrast of the image using imadjust function so that we can get a uniform background. It also helps in further processing of the image.

The step is to perform pre-processing operations on the input Pap smear image. The input image is a coloured image which is in RGB mode. As is known the colour images are difficult to process so we convert the image to greyscale image using rgb2gray function in matlab. Then we adjust the contrast of the image using imadjust function so that we can get a uniform background. It also helps in further processing of the image. Next Gaussian filter is applied on the image. After that the histogram of the image is obtained using imhist function, which helps in finding a suitable threshold for greyscale image to binary conversion using im2bw function of the image so that we can perform the morphological operation for feature extraction. Finally median filter is applied on the output image so that it can remove the salt and pepper noise present in the image.

The next step is to perform the morphological operations so that proper segmentation of the image can be done. Firstly the opening is performed using imopen function of matlab. This will remove the thin protrusions and also help in finding two separate nucleuses which are slightly connected. The complement operations are performed using imcomplement function. In matlab the object are set of white pixels that are connected to each other [5]. So to extract different features we have to give a value of 1 to all the pixels corresponding to a nucleus. But the main

difficulty for proper segmentation is presence of unwanted components like WBC and other cells. So we have to take good measures to remove those components. In our approach we performed morphological erosion using imerode function of matlab which is followed by Median filter of the image. Again all the areas are cleared using bwareaopen function to remove the entire object having less than 100 pixels. This will remove the WBC s present. Finally we have to fill up the holes using imfill function.

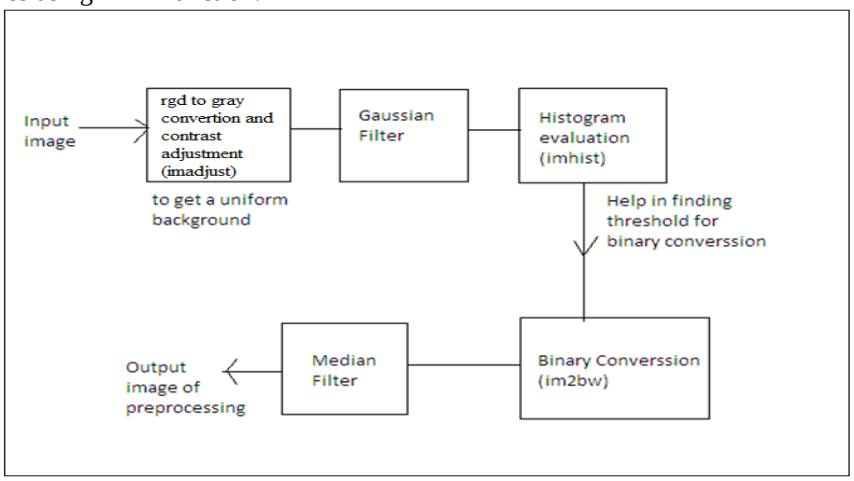


Fig 3(a) Pre-processing part

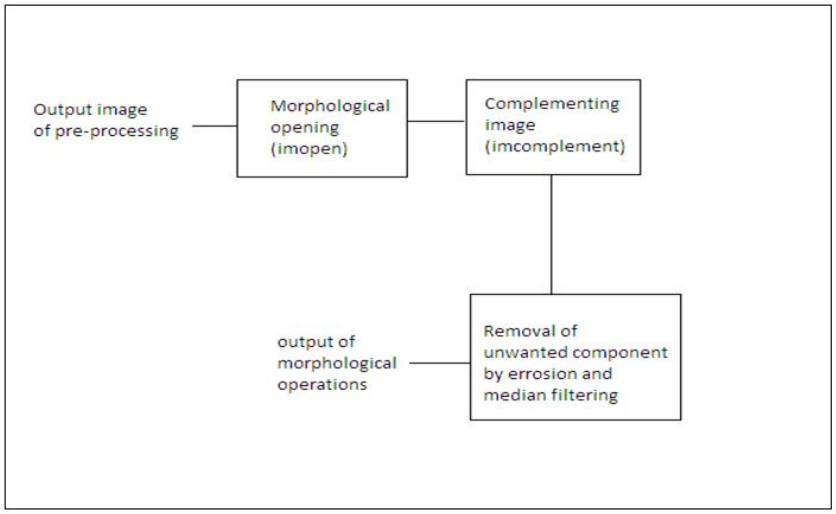


Fig 3(b) Morphological operations

After that we will get an output image which contains only the nucleus. Then using regionprops function area is calculated in pixels (let's say area_nucleus).

Step 2 – Calculation of cytoplasm area

To identify the cytoplasm level set contour algorithm is used [6] [7]. It will identify the cytoplasm including the nucleus. Then output image of this step is subtracted from the output image of step1. Then the cytoplasm will be found excluding the nucleus. After that using regionprops function of matlab the area is extracted which is expressed in no of pixels (let's say area is area_cytoplasm)

Step 3- Calculation of N/C ratio

The last step is to calculate the N/C ratio which can be obtained by

$$\text{N/C ratio} = \text{area\_nucleus}/\text{area\_cytoplasm}$$

where, area_nucleus is obtained from step1 and area_cytoplasm is obtained from step2.

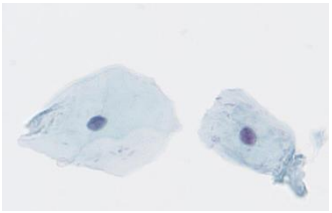
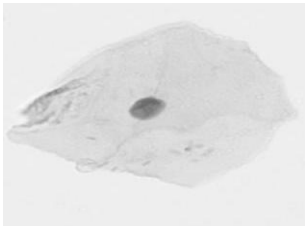
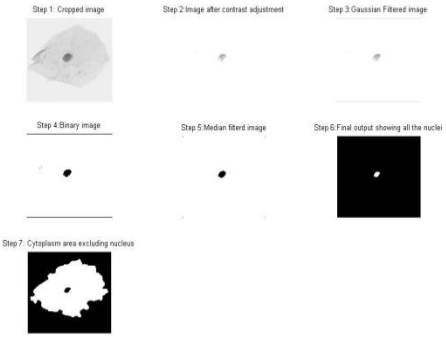
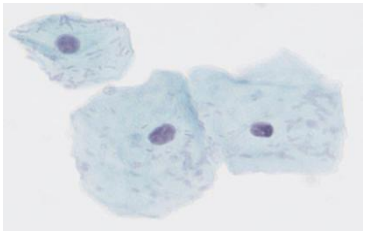
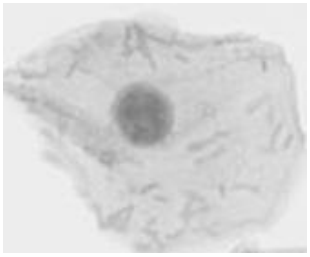
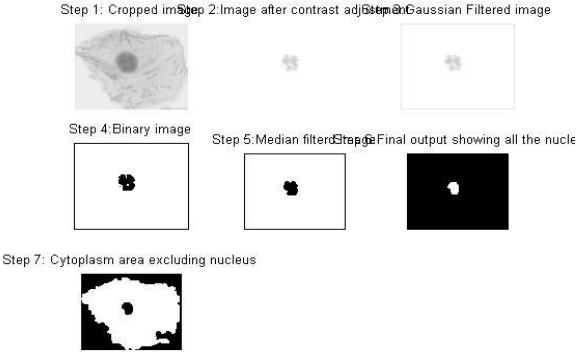
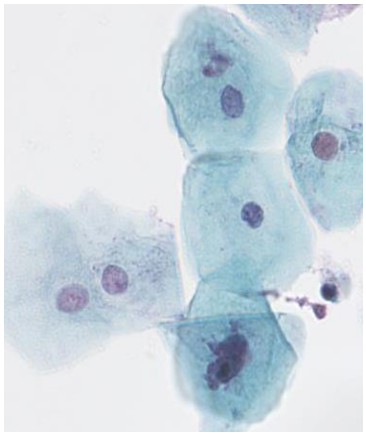
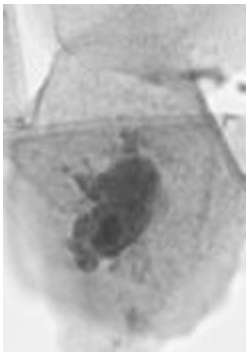
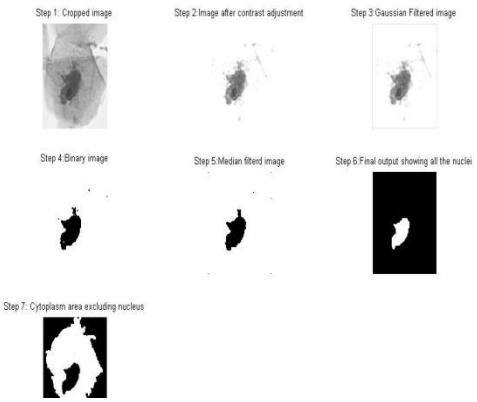
Bases on the output obtained the N/C ratio of a normal and an abnormal cell can be calculated. Our observations are shown in following explanations.

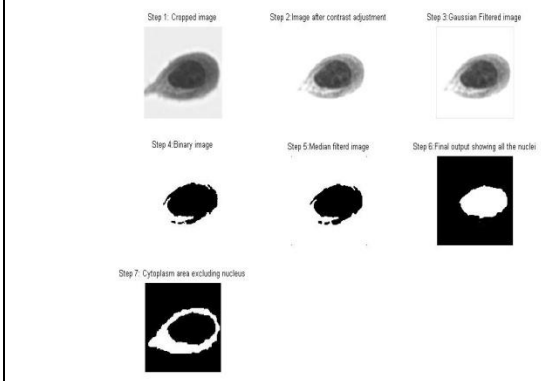
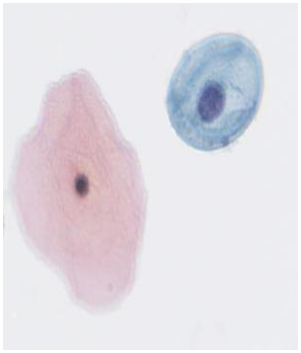
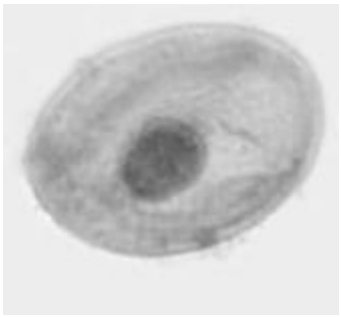
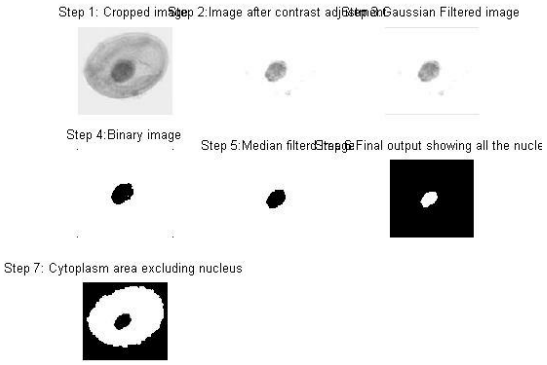
V. EXPERIMENTAL RESULTS

In our approach we analyse nearly 100 individual cervical cells 50 of them are normal and 50 are abnormal which are collected from 20 Pap smear images. Pap smear images are obtained from Dr B.Barooh Cancer Institute, Guwahati. To explain here 5 results are shown, among them 2 are normal and rest 3 are cancer affected. Table1 contains all 5 observations along with corresponding N/C ratio. In column 1 of Table1 the Pap smear image is shown, in column 2 the cropped cell is shown on which the method is used to find the N/C ratio and finally in column 3 the stepwise output of

the system is shown along with the N/C ratio of the corresponding cell. So, from this experimentation it can be concluded that the cell having N/C ratio less than 0.05 are can be considered as normal and cell having N/C ratio more than 0.05 can be considered as an abnormal cell. For the cell having two nucleuses, the N/C ratio need not to be calculated as they may either undergo division phase or it may be cancerous.

TABLE 1
OUTPUT OF THE PROPOSED APPROACH

Input image	Cropped image	Output
NORMAL CELL1 		**N/C ratio = 0.0165** 
NORMAL CELL 2 		**N/C RATIO = 0.0397** 
CANCER CELL1 		**N/C RATIO = 0.1781** 

<table>
<tr>
<td>  </td>
<td>  </td>
<td> 
</td>
</tr>
<tr>
<td>CANCER CELL 3
  </td>
<td>  </td>
<td>N/C RATIO = 0.0825
 
</td>
</tr>
</table>

VI. CONCLUSION

Medical images have various limitations such as low quality, presence of noise and human error in interpretation[8]. So this type of automatic analysis may help the pathologist to a great extent. But this system totally depends upon the quality and contrast of the image. Based on the contrast the output may be different for different users.

Some idea for further studies may include considering the whole Pap smear image for analysis at the same time instead of taking one cell at a time from that image**.**

ACKNOWLEDGMENT

We would like to offer the gratitude to Dr. J.D.Sharmah, Chief consultant, Pathology department of Dr. B. Barooah Cancer Institute for his valuable guidance and precious time and for providing the images for experimentation. This project has the consent of all co-authors and authorities of the institute, where this study has seen carried out and there exists no conflict of interest anywhere.

REFERENCES

[1] Sreedevi M T, Usha B S,Sandya S, "Papsmear Image based Detection of Cervical Cancer ", International Journal of Computer Applications (0975 – 8887) Volume 45– No.20, May 2012

[2] Dinshaw KA, Shastri SS, PatilSS,"Cancer Control Programme In India:Challenges For The New Millennium", Health Administrator Vol: XVII, Number 1: 10-13, pg.

[3] American standard society, "Cervical Cancer". http://www.cancer.org/healthy/findcancerearly/cancerscreeningguidelines/american-cancer-society-guidelines-for-the-early-detection-of-cancer, Last accessed onSeptember 24, 2012.

[4] Turgeon, Mary Louise (2005). Clinical hematology: theory and procedures. Hagerstwon, MD: Lippincott Williams & Wilkins. pp. 67. ISBN 0-7817-5007-5.

[5] Hao Yuan Kueh, Eugenio Marco, Mike Springer and Sivaraj Sivaramakrishnan, " Image analysis for biology", MBL Physiology Course 2008.

[6] Chunming Li, Chenyang Xu 2, Changfeng Gui and Martin D. Fox 1Level Set EvolutionWithout Re-initialization: A New Variational Formulation

[7] M. Kass, A.Witkin, and D. Terzopoulos, "Snakes: active con-tour models", Int'l J. Comp. Vis., vol. 1, pp. 321-331, 1987.

[8] Lipi B. Mahanta, Dilip Ch. Nath, Chandan Kr. Nath, "Cervix Cancer Diagnosis from Pap Smear Images Using Structure Based Segmentation and Shape Analysis", Journal Of Emerging Trends in Computing And Information Sciences,VOL. 3, NO. 2, February 2012

