

## Detection of Cervical Cancer at an Early Stage Using Hybrid Segmentation Techniques from PAP Smear Images

Asmita Ray<sup>1</sup>, Indra Kanta Maitra<sup>2</sup> and Debnath Bhattacharyya<sup>1</sup>

<sup>1</sup>*Dept of Computer Science & Engineering,  
Vignan's Institute of Information Technology, A.P.,*

<sup>2</sup>*Department of Information Technology,  
B.P. Poddar Institute of Management and Technology, Kolkata*

### Abstract

*Cervical cancer is the abnormal cell growth on cervix. The lower part of the uterus or womb that opens top of the birth canal is called cervix. Unlike other cancer cervical cancer has no telltale symptoms so cannot be detected at an early stage thus mortality rate is very among woman. PAP smear test is an efficient and popular screening procedure to detect the presence of abnormal cells in cervix. Manual screening is not always perfect, quantifying and analyzing the abnormal cells from the large number of PAP smear images is a tedious task. A novel approach has been proposed to detect and analyze the abnormal cells from PAP smear images at an early stage, that it will help for fast and proper treatment. The proposed method carried out four stages: preprocessing, pre-segmentation, segmentation and feature extraction. In preprocessing stage, unwanted noise and artifacts have been removed. In pre-segmentation stage, histogram has been computed first then morphological operation followed by edge detection techniques have been performed. In segmentation phase nucleus has been segmented by hybrid segmentation method which consists of thresholding and watershed segmentation. Two types of feature extractions like (a) morphological feature extraction, (b) texture feature extraction GLCM (Gray Level Co-Occurance Matrix) have been performed in this stage for accurate detection of abnormalities in cervical region. This paper presents a complete qualitative assessment which is used to develop a self-constructing algorithm for diagnosis the cancerous cell from PAP smear images.*

**Keywords:** *Cervical cancer, PAPSmear Images, Pre-Segmentation, Hybrid Segmentation, Feature Extraction, GLCM*

### 1. Introduction

Cervical cancer is the fourth most prevalent form of cancer after breast cancer that occurs in woman belongs to all age groups. Cervical cancer does not show any symptom at an initial stage until it is so advanced thus presently morbidity is increasing consistently. Real formation of cervical cancer is very slow, most cervical cancer takes more time to develop from normal to precancerous cells (dysplasia) [1]. Patient can feel pain in lower abdominal or back region at latest stage. Early detection and confirmation of this kind of cancer effectively treatable and preventable. Human Papilloma Virus (HPV) is the main cause cervical cancer [2]. This virus damage the squamous cell and grandular of the cervix, when it affects the squamous cell then it is called squamous cell carcinoma and grandular cell cancer is called adenocarcinoma. Squamous cell carcinoma is the most common form of cervix cancers arise in the epithelium cell of the cervical lining. Adenocarcinoma develops inside the uterus and, in the mucus-producing gland cells of

---

Received (October 15, 2017), Review Result (December 19, 2017), Accepted (January 17, 2018)

the endocervix [3]. PAP smear test is the preliminary manual screening method to identify the precancerous lesions of the cervix region.

Cervical cancer is one of the highest leading cause of death, worldwide, morbidity and mortality among the female is very high due to this cancer. Cervical cancer has the significant impact, according to World Health Organization (WHO) it is estimated that approximately 530,000 new cases in the world in 2012, more than 27,000 women died each year due to cervical cancer and around 80% of all cervical deaths occur in the developing country [4]. Approximately 70% of deaths occur in low- and middle-income countries [5]. Each year approximately half a million women develop cervical cancer and within this it is estimated that 85% from developing countries. Every year, 4648 women are diagnosed with cervical cancer and among this, 3,235 ends with death [6].

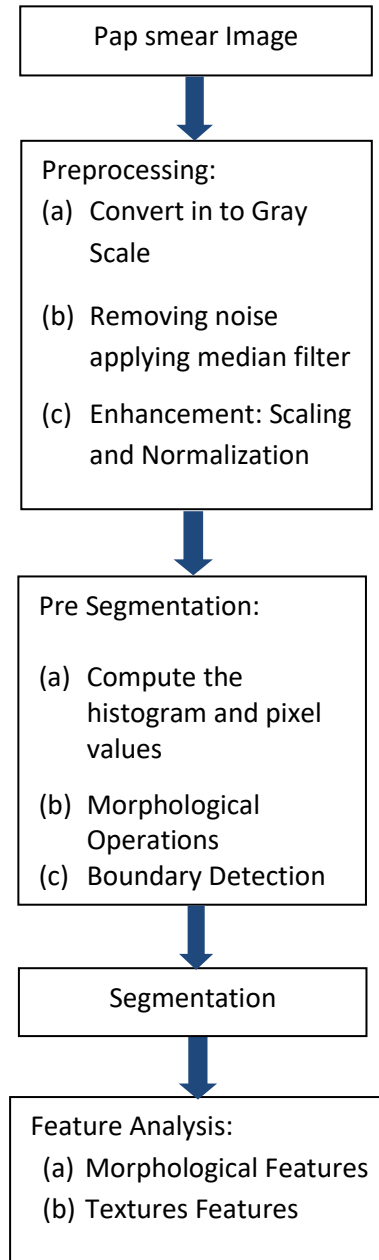
Several studies for segmentation of nuclei in cytological images have taken place [7-12]. Accurate detection and segmentation of cell nuclei from the PAP smear images is one of the challenging task due to of its high degree of overlapping with other cells, more than one nucleus in a cell, lack of homogeneity in image intensity. Nucleus is the most important cell structure for detecting the cervical cancer, Quantification of density and morphological changes in nucleus, detect the normal and abnormal cells from PAP smear images.

Current manual screening methods are very expensive and may have chances of producing inaccurate diagnosis result caused by human error. In this regard the computer assisted diagnosis system has a significant benefit which can reduce the financial cost as well as increases the screening accuracy.

In this paper, we propose a boundary detection algorithm followed by hybrid segmentation method for locating the contour of the nucleus which plays a key role for identifying the cervical cancer. In this paper we developed both abnormal lesion detection and feature quantification techniques. Development of accurate, robust segmentation approach and feature extraction techniques of this proposed algorithm are the two key factors for providing the facility of longitudinal monitoring and it will be very effective for analyzing the morphological and architectural alterations of pre-cancer and cancer's stage from PAP smear image.

## **2. Methodology**

Figure 1 shows the block diagram of proposed method. The figure represents the different steps of preprocessing, pre-segmentation, segmentation and feature extraction techniques to detect the cervical cancer from Pap smear images.



**Figure 1. Block Diagram of Proposed Method**

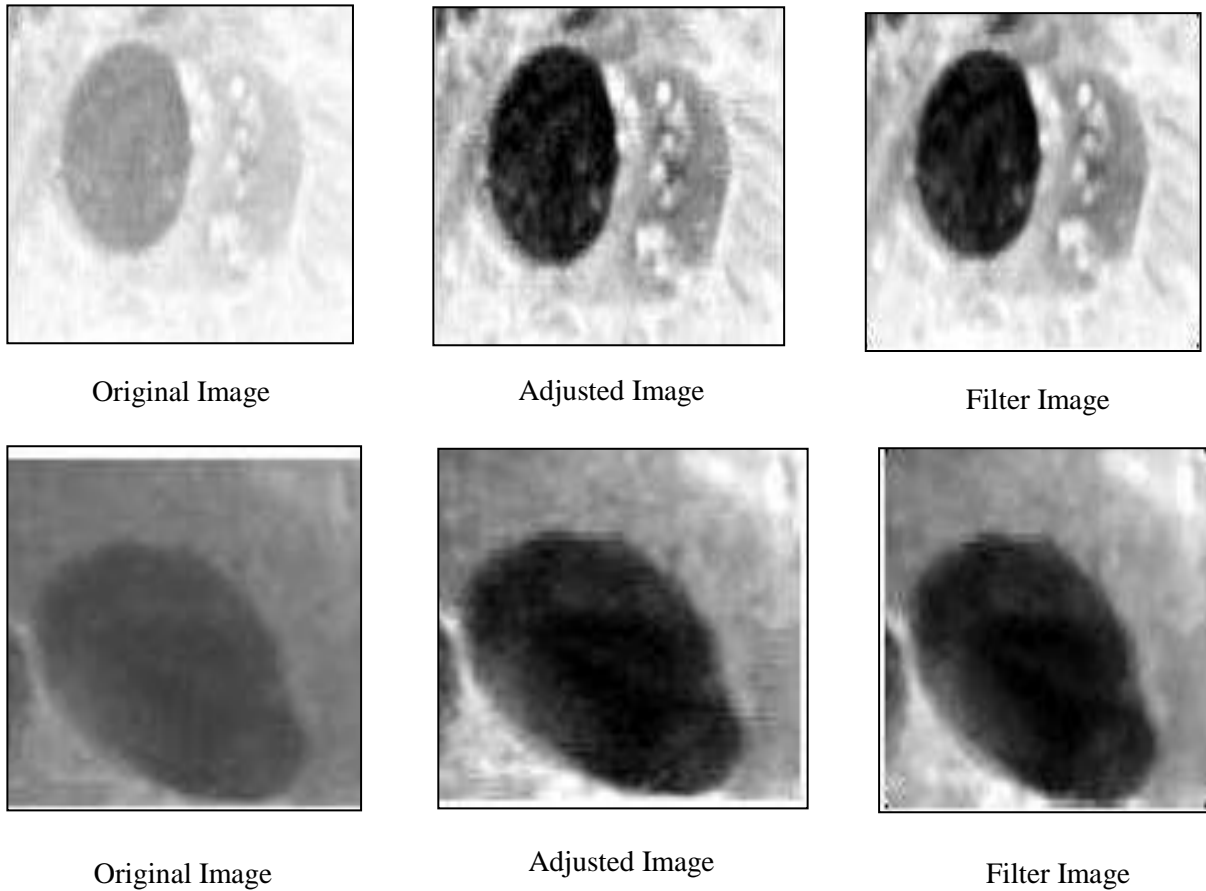
## 2.1. Detail Description of the Proposed Method

### **Preprocessing:**

Pap smear images play a crucial role for detection of cervical cancer from. Qualities of the Pap smear images are degraded due to various limitation and interference like variation of stain, uneven light across the field of view and noise during the transfer. So preprocessing of the image is the important step to assure the accurate analysis in segmentation step [13-14]. Pap smear images are the colored images, so the first and different phases of the preprocessing stage are as follows:

- (a) Covert it into gray scale image.

- (b) To enhance the contrast and remove the noise from the image, image adjustment and median filter of windows [3 X 3] have been applied, which are necessary to obtain stable boundaries and edges of nucleus.

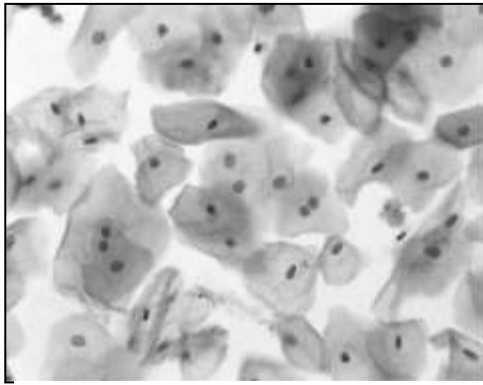


**Figure 2. Preprocess Images for Normal and Abnormal Cell**

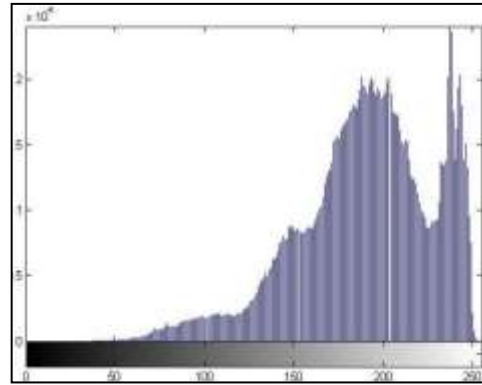
**Pre-Segmentation Stage:**

Pre-segmentation stage is the most important stage as in this stage, regions of interest has been detected effectively from the background. Steps for pre-segmentation of the proposed methods are as follows:

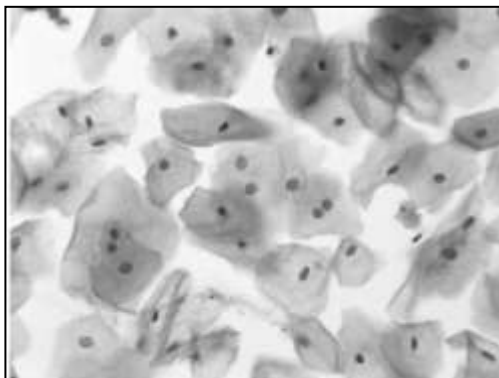
- a) Compute the gray –level histogram and pixel values.
- b) Apply morphological operations like erosion and dilation for removing small blemishes without effective the overall shapes of the object.
- c) Perform edge detection technique for defining regions of interest effectively.



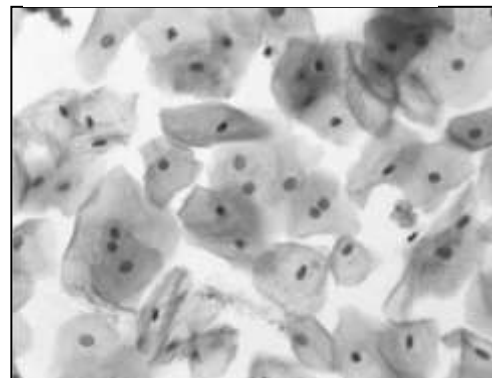
Original Pap Smear Cervical cytology



Histogram of Pap smear Image



Dilated Image

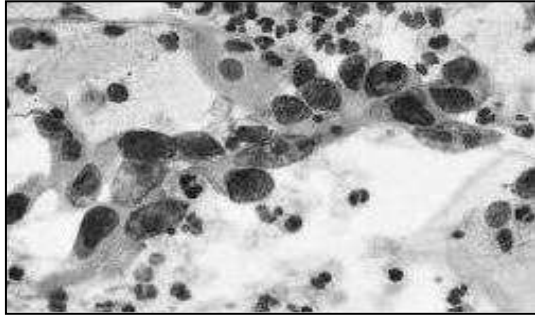


Eroded Image

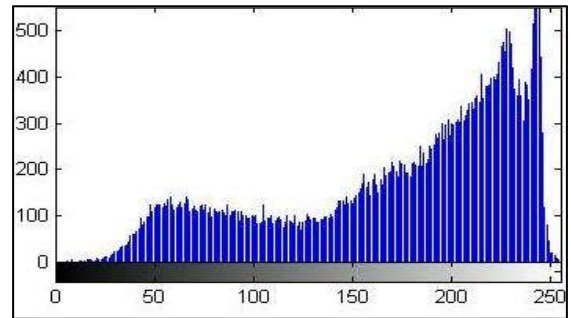


Edge Detected Image

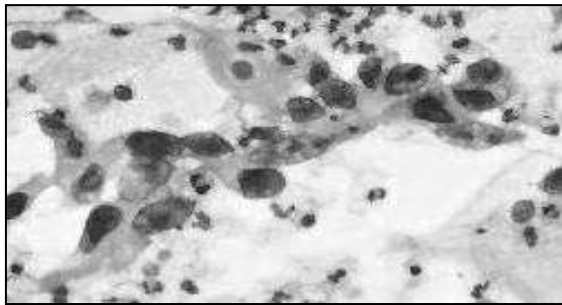
**Figure 3. Pre-Segmentation Images for Normal Pap Smear**



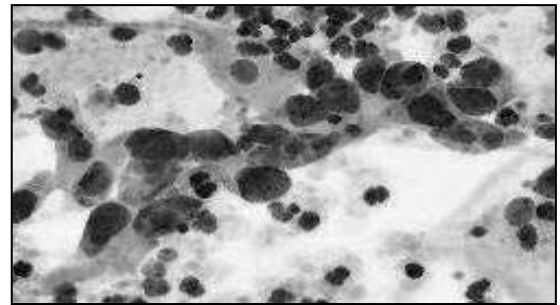
Original Pap Smear Cervical Cytology



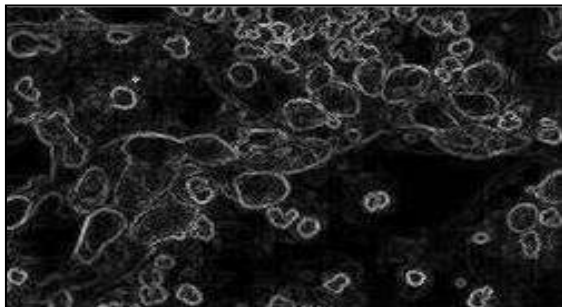
Histogram of Pap Smear Image



Dilated Image

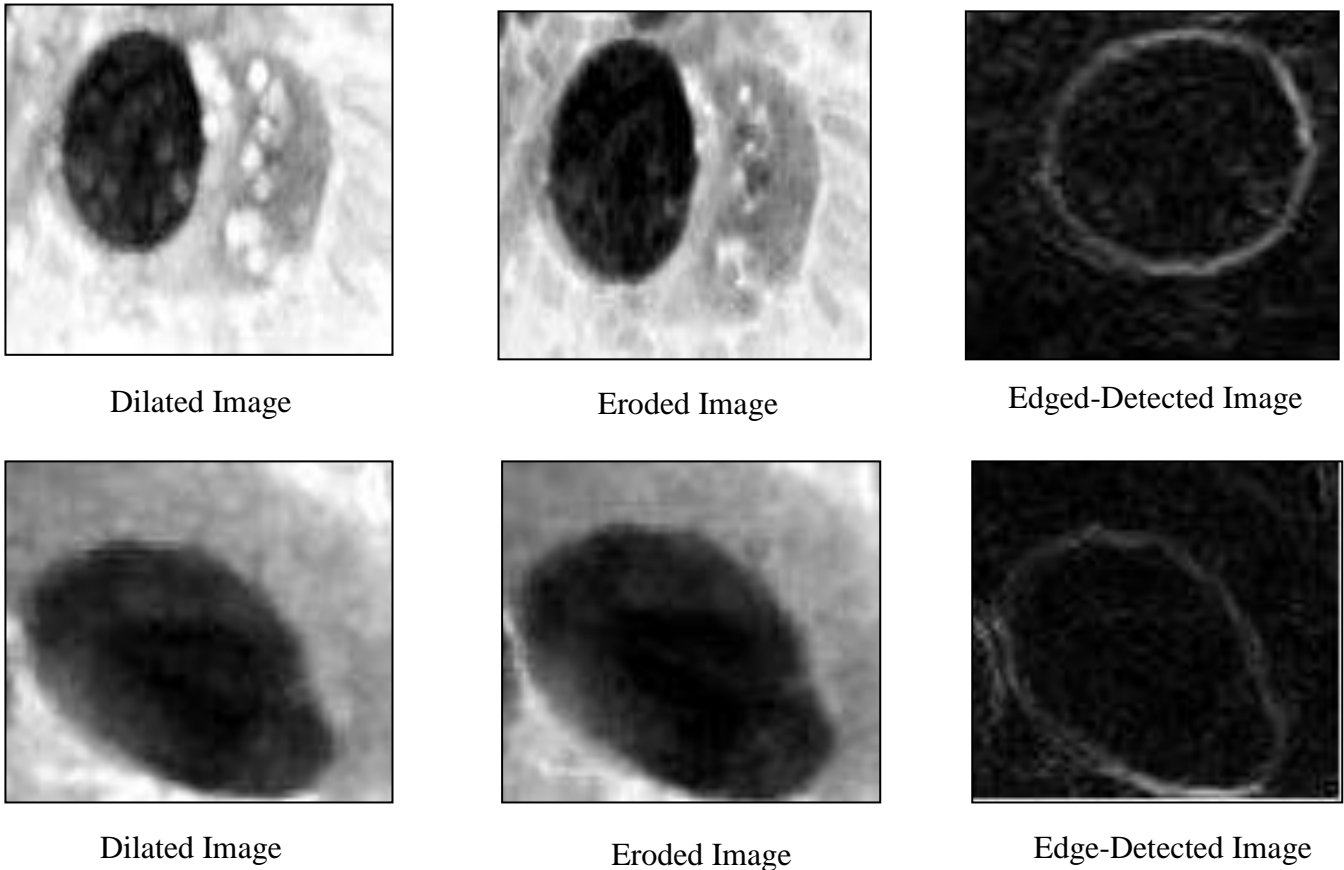


Eroded Image



Edge Detected Image

**Figure 4. Pre-Segmentation Images for Malignant Pap Smear**

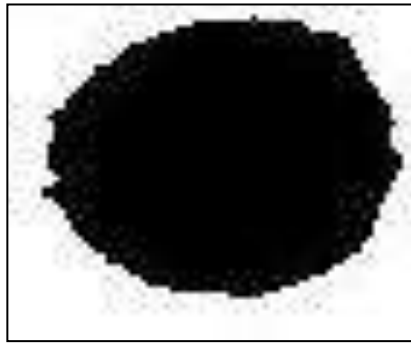


**Figure 5. Pre-Segmented Images for Normal and Abnormal Cell**

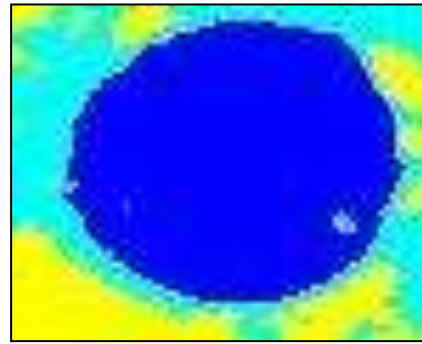
**Segmentation:**

**Watershed Segmentation:** Watershed segmentation is the powerful technique to delineate the object of interest. It has been used to determine the nucleus boundary which is very crucial step for detection of cervical cancer. Pixels have been determined either by the regional minima of the image or predefined set of markers. The main objective of this method is to determine the pixels of watershed lines which are used to isolate the neighboring catchment basin and different characteristics parts of image.

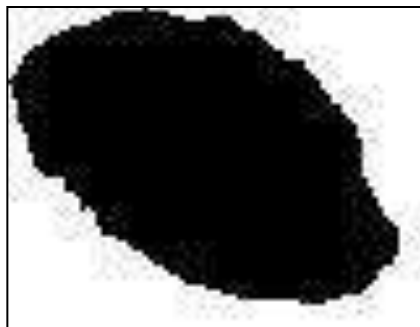
**Otsu's Thresholding Method:** Otsu is type of global thresholding method, it is very simplest, effective and fast segmentation technique thus it is widely used. In global thresholding the threshold (T) totally depend on gray level value and the character of pixel based on value of T. Gray-level histogram need to be performed before running the Otsu method. In this method, optimal gray-level threshold value has been selected for separating the object of interest from the background.



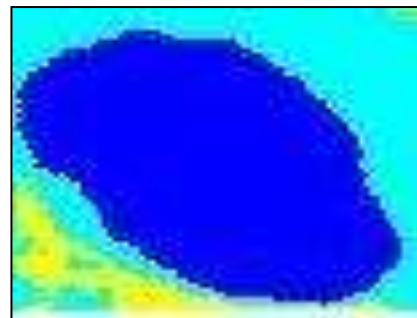
Threshold Segmentation



Watershed Segmentation



Threshold Segmentation



Watershed Segmentation

**Figure 6. Segmented Images for Normal and Abnormal Cell**

#### **Feature Extraction:**

Feature extraction is very effective method for making the decision, the aim of this process to transfer microscopically observed visual parameter to quantify process. Cervical cancer detection mainly depends on the two important information (a) cell structure abnormalities, (b) tissue level abnormal cell distribution. Thus, in case of cervical cancer screening, cell recognition *i.e.* cell features selection and calculation are the most essential step. Features are extracted to quantify the change in tissue or cell level as the abnormalities in size, shape, ratio, texture and color intensity can be detected. Cell nuclei play a key role for detection of cervical cancer. Two types of feature extraction have been performed (a) morphological feature extraction, (b) texture feature extraction GLCM (Gray Level Co-Occurance Matrix).

**Morphological Feature Extraction:** Morphological features are very essential as it describes the shape of the nucleus and this shape depicts tendency towards cancer. We extracted morphological features like area, diameter, roundness, nuclear-cytoplasmic ratio (N/C) for cancerous and non-cancerous cell.

**Area:** Area within the outlined nuclear perimeter is defined as nuclear area.

**Diameter:** Diameter is actually outlined nucleus.

**Roundness:** Number of chromosome in nucleus will change if the cells become malignant and as a result, it affects the circularity and roundness properties. Roundness describes the distance between the border and the centre of the area.



**Nuclear-Cytoplasmic Ratio (N/C):** In normal, cell nucleus occupies less amount of area compared to abnormal cell's nucleus and cytoplasm occupies larger area than an abnormal cell.

**Table 1. Morphological Feature**

Type	Area (in pixel)	Diameter	Roundness	NC Ration
Normal	18	4.2±0.70	0.796	0.262±0.006
Cancer	25	12.6±1.3	0.620	0.920±0.001

Table 1 describes the area, diameter roundness and NC ratio and also depicts how these morphological features are varying from normal to abnormal cell.

Texture Features - GLCM (Gray Level Co- Occurrence Matrix):

GLCM is the second order statistics which used for texture analysis and it is applied for extracting various textural features. The GLCM functions characterize the texture of an image by calculating the frequent occurrence of pairs of pixel with specific values and specified spatial relationship. The features which are used in this paper are:

**Contrast:** Contrast is the measurement of local variations in the gray level co-occurrence matrix.

**Correlation:** It measures how the reference pixel correlates with its neighbor over an image.

**Energy:** Energy is also called as angular second moment which defines the summation of squared elements in the GLCM. Energy is 1 for a constant.

**Homogeneity:** Homogeneity describes the distribution closeness of the element in the GLCM to GLCM diagonal.

**Table 2. GLCM Features**

Type	Contrast	Correlation	Energy	Homogeneity
Normal	0.6458	0.9514	0.1056	0.7995
Cancer	0.5335	0.9536	0.0928	0.8348

### 3. Conclusion

This paper proposes a computer-aided diagnosis system for cervical cancer detection with an efficient approach for segmentation and feature extraction method. This proposed algorithm has been represented a promising performance for segmenting the nucleus successfully. This algorithm has ability to suppress the noise and enhance the contrast effectively for clear distinguishable boundary of the nucleus. In our study we have used two different kinds of feature extraction method (a) morphological features, (b) texture features. Nuclear morphology study is the most useful objective tool for differentiating the malignant smear from the normal and which has been established the proposed algorithm robust and accurate with great facilities and capabilities of promising performance.

This study help the pathologists to start the treatment properly at an early stage and this algorithm exempted them from tedious time consuming sometimes difficult and error prone manual interpretation. Early detection of the cervical cancer can be treated successfully and save the life, in this respect our new proposed method is very useful for fast and proper treatment by accurate detection of cancerous cell.

## References

- [1] H. S. Cronj'e, "Screening for cervical cancer in the developing world", *Best Practice and Research: Clinical Obstetrics and Gynaecology*, vol. 19, no. 4, (2005), pp. 517–529.
- [2] NCCC, "Cervical cancer", <http://www.nccc-online.org/index.php/cervicalcancer>.
- [3] ACS, "What is cervical cancer?" 2011, American Cancer Society, <http://www.cancer.org/Cancer/CervicalCancer/DetailedGuide/cervical-cancer-what-is-cervical-cancer>.
- [4] <http://www.who.int/en/>.
- [5] G.K. Singh "Rural-urban trends and patterns in cervical cancer mortality, incidence, stage, and survival in the United States, 1950-2008", *J Community Health*, vol 3, no.1, (2012), pp.217-23.
- [6] The GLOBOCAN database (2012) Cancer incidence, mortality and prevalence worldwide. International Agency for Research on Cancer Lyon, France.
- [7] P. Bamford and B. Lovell, "Unsupervised cell nucleus segmentation with active contours", *Signal Processing*, vol.71, no.2, (1998), pp. 203-213.
- [8] P. Bamford and B. Lovell, "A water immersion algorithm for cytological image segmentation", *Proceedings of the APRS Image segmentation workshop*, University of Technology Sydney, Sydney, (1996).
- [9] [9] T. Mouroutis and S. J. Roberts, "Robust cell nuclei segmentation using statistical modelling", *IOP Bioimaging*, vol.6, (1998), pp. 79-91.
- [10] A. Garrido and N. Perez de la Blanca, "Applying deformable templates for cell image segmentation", *Pattern Recognition*, vol.33, (2000), pp. 821-832.
- [11] K.M. Lee and W.N. Street, "Learning shapes for automatic image segmentation", *Proceedings INFORMS-KORMS Conference*, Seoul, Korea, (2000).
- [12] G. Begelman, E. Gur, E. Rivlin, M. Rudzsky and Z. Zalevsky, "Cell nuclei segmentation using fuzzylogic engine", *International Conference on Image Processing*, vol. 5, (2004), pp. 2937-2940.
- [13] R. Chouhan, R. K. Jha and P. K. Biswas, "Enhancement of dark and low-contrast images using dynamic stochastic resonance", *IET Image Processing*, vol. 7, no. 2, (2013), pp. 174–184.
- [14] J. Mukherjee and S. K. Mitra, "Enhancement of color images by scaling the DCT coefficients", *IEEE Transactions on Image Processing*, vol. 17, no. 10, (2008), pp. 1783–1794.