

2010

## Classification cervical cancer using histology images

Rahmadwati

*University of Wollongong, rr516@uowmail.edu.au*

Golshah Naghdy

*University of Wollongong, golshah@uow.edu.au*

Montserrat Ros

*University of Wollongong, montse@uow.edu.au*

Catherine Todd

*University of Wollongong, cath@uow.edu.au*

E. Norachmawati

*University of Wollongong*

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>



Part of the [Physical Sciences and Mathematics Commons](#)

---

### Recommended Citation

Rahmadwati; Naghdy, Golshah; Ros, Montserrat; Todd, Catherine; and Norachmawati, E.: Classification cervical cancer using histology images 2010.  
<https://ro.uow.edu.au/infopapers/3592>

---

## Classification cervical cancer using histology images

### Abstract

This paper reports on methodologies and outcome of a study aiming at developing robust tool to evaluate and classify histology images of cervical cancer. Using the histology images acquired from the pathology laboratories in an Indonesian hospital, this study aims to classify cervical biopsy images based on four well known discriminatory features a) the ratio of nuclei to cytoplasm b) diameter of nuclei c) shape factor and d) roundness of nuclei. In this study, the cervical histology images are classified into three categories: 1) normal, 2) pre cancer and 3) malignant. The final system will take as input a biopsy image of the image of the cervix containing the epithelium layer and provide the classification using the new automated approach, to assist the pathologist in cervical cancer diagnosis.

### Disciplines

Physical Sciences and Mathematics

### Publication Details

Rahmadwati, G. Naghdy, M. Ros, C. Todd & E. Norachmawati, "Classification cervical cancer using histology images," in International Conference on Computer Engineering and Applications, 2010, pp. 515-519.

## Classification Cervical Cancer Using Histology Images

Rahmadwati, Golshah Naghdy, Montse Ross, Catherine Todd, E.Norachmawati

School of Electrical, Computer and Telecommunication Engineering

University of Wollongong

New South Wales Australia

[rr516@uow.edu.au](mailto:rr516@uow.edu.au), [golshah@uow.edu.au](mailto:golshah@uow.edu.au), [montse@uow.edu.au](mailto:montse@uow.edu.au), [catherinetodd@uowdubai.ac.ae](mailto:catherinetodd@uowdubai.ac.ae)

**Abstract**—This papers reports on methodologies and outcome of a study aiming at developing robust tool to evaluate and classify histology images of cervical cancer. Using the histology images acquired from the pathology laboratories in an Indonesian hospital, this study aims to classify cervical biopsy images based on four well known discriminatory features a) the ratio of nuclei to cytoplasm b) diameter of nuclei c) shape factor and d) roundness of nuclei. In this study, the cervical histology images are classified into three categories: 1) normal, 2) pre cancer and 3) malignant. The final system will take as input a biopsy image of the image of the cervix containing the epithelium layer and provide the classification using the new automated approach, to assist the pathologist in cervical cancer diagnosis.

**Keywords**-histology,cervical cancer, diagnosis,classification

### I. INTRODUCTION

In the United State, it is estimated that 11,070 cases will be diagnosed with *cervical* cancer this year (2008), and sadly to state about 3,870 women will die from it. [1]. The most common cancer in Indonesia is cervical cancer. Based on the data of female cancer program in Indonesia approximately 20% of 200.000 (40.000) of cancer patients is diagnosed with cervical cancer in every year [2]. This phenomenon is developed in developing country because this lack of experts combined with often insufficient lab facilities can make reading the slide [5]. Furthermore, this is caused by the competency and evaluation of skills of screeners and other health personnel involved in examination the cancer [6].

Generally, conventional histology is used by the pathologist to diagnose, however the diagnosis may vary in terms of interpretation, since the

examination is influenced by subjectivity factors. This condition is compounded by the heterogeneity features such as the experience and expertise of pathology. To reduce the subjectivity and assist the pathologist in diagnostic nprocess, it is important to develop a robust tool to evaluate and classify the histology slide. A computer assisted diagnosis to analysis cancer has been dedicated by many researchers with particular area such as: breast cancer, colon, prostate, cervix, lung etc.

A current image processing and analysis extracting the histological image are needed to provide prognostic features are used in this research such as nuclei cytoplasm ratio (NC ratio), nuclei area, nuclei diameter, shape factor of nuclei and pleimorphicity.

A few studies discuss analysis histology cervical images, although some have shown the grading of pre cancer [7]. In this paper, reports on the methodologies and outcome of study aimed at developing a robust tool to evaluate and classify histological cervical images into normal, pre cancer and malignant.

### II. MEDICAL BACKGROUND

#### A. Squamous Epithelium Structure

The cervical squamous epithelium is stratified, it consists of several layers i.e. basal layer (*stratum cylindricum*), parabasal layer (*stratum spinosum profundum*), intermediate layer (*stratum spinosum superficiae*) and superficial layer (*stratum corneum*). Each layer has different cell

structure and morphology characteristics. In normal tissue, the structure of cells takes place ordinary and controlled but it is possible to uncontrolled [8]. It means the cell is not in the right way, it becomes enlarge, change their shape going to abnormal characteristic. The ratio nuclei to cytoplasm for normal cervix is approximately 1:4 or 1:6. Fig 1 shows the structure of epithelium layer.

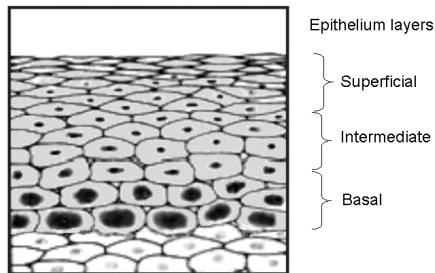


Figure 1. Structure of Epithelium layer

### B. Cervical Intraepithelial Neoplasia ( Pre cancerous)

*Cervical Dysplasia* is also called *Cervical Intraepithelial Neoplasia* (CIN) or precancer. It describes the presence of abnormal, precancerous cells on the surface of a cervix or its canal that take place in the squamous epithelium of the cervix as a precursor to invasive carcinoma. The grading of CIN will appear in continuous spectrum changes, which epithelium become more abnormal in appearance, or dysplastic. CIN is classified into three class i.e. CIN I, CIN II and CIN III. The classification of CIN correlates with the proportion of the total thickness affected by dysplasia.

In CIN, the nuclei are enlarged, the various in size (pleiomorphic), irregular structure, hyperchromatic, prominent nucleoli, little cytoplasm and loss of polarity.

### C. Invasive Cervical cancer

Invasive cervical cancer also is called cervical cancer, which is the most common cancer among woman in world wide. Cervical

cancer is detected by pathologist who examines the images under microscope.

Characteristically, cervical cancer arises from the squamo-columnar junction. The cell structure of the cancer tend to uncontrolled growth that ratio nuclei to cytoplasm is higher than normal[9]. Another characteristic of cancer is various in shape, normally, the shape of nuclei is circle but in abnormal the nuclei many various shape such as oval, tadpole, spindle etc. Fig 2 shown the Histology images of normal, pre cancer and malignant

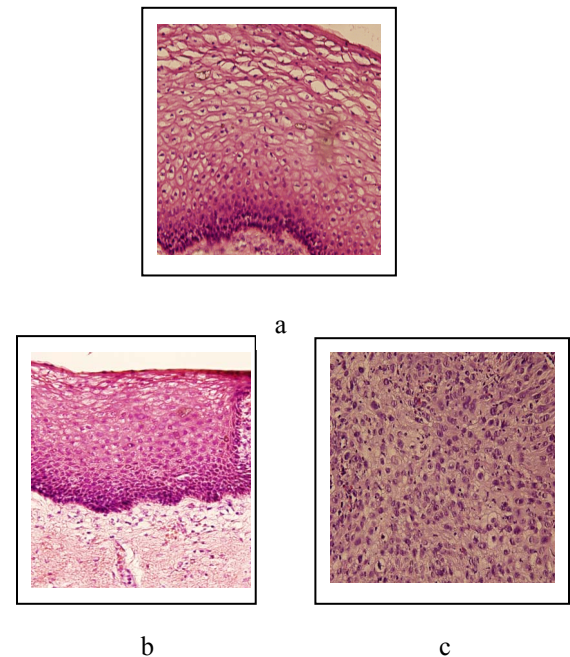


Figure 2. Histology images a. Normal cell  
b. Pre cancer c. cancer (malignant)

### III. METHODOLOGY

Using the histology images acquired from pathology laboratories in an Indonesian hospital, this study is aims to classify the cervical biopsy images based on the cellular structure and then analysis the characteristic each class. The k-means clustering is selected and used to segment the images. K-means algorithm separates the nuclei and cytoplasm from the background and connective tissue. In order to effectively calculate

the features, the nuclei and cytoplasm regions are shown and analysed in two separated images. This algorithm based on colour segmentation approach.

A. K-means Clustering

Histology images is processed using image processing step such as find the region of interest of the histology images, feature extraction and analyses the feature characteristic. Nuclei region is resulted by segmentation process. K-means clustering is one of simplest method and the fastest method for clustering. The number of clusters is predefined and selected points are used to define the centre of each clusters using the arithmetic mean. In histological slide, there are five dominant colour white (background), purple (nuclei), red (red blood cells), pale pink (cytoplasm in stroma), and dark pink ( cytoplasm in squamous epithelium).

Step 1: Images are processing using k-means clustering.

Table 1. shows the k-means clustering Processing. The segmentation results are presented as gray scale images, using different gray scales to represent tissue components.

Step 2: Improve the image quality using morphological operation

The morphological operations are used on both nuclei and cytoplasm images to remove pixels (noise) which do not belong to objects of interest. These operations result in much improved image quality needed for subsequent processing and feature calculation. Figure 3 shows the nuclei and cytoplasm area.

Step 3 : Matching Methods

Nuclei and cytoplasm are obtained in separate region. Each nuclei and cytoplasm has position then find the same position of each nuclei and cytoplasm. The areas of each corresponding nuclei and cytoplasm detected are calculated in order to calculate the feature referred to as nuclei to cytoplasm (N/C) ratio.

Table 1. K-means clustering

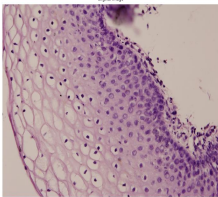
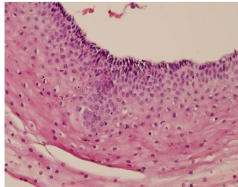
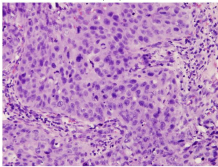
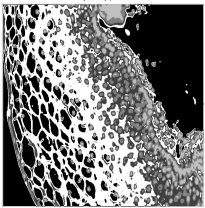
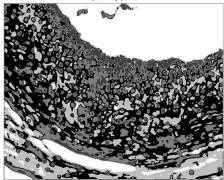
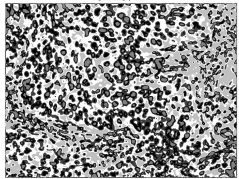
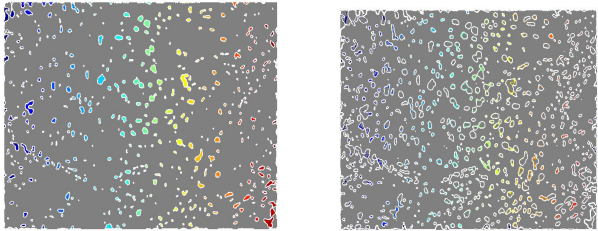
Original Images	k-means clustering Classification
 <p>Normal</p>  <p>Pre cancer</p>  <p>Malignant</p>	 <p>Normal</p>  <p>Pre cancer</p>  <p>Malignant</p>
 <div> <div>a</div> <div>b</div> </div>	

Figure 3. Nuclei and Cytoplasm have been improved using morphology operation.

Types	Nc ratio	diameter	Roundness	Shape factor
Normal	0.265±0.008	3.5+0.72	0.797	17.26
Pre cancer	0.675±0.002	5.4+0.65	0.765	22.35
Malignant	0.948±0.002	11.4+1.2	0.612	29.15

Step 4: Calculate the feature

$$N/C \text{ ratio} = \text{area nuclei} / \text{area cell} \quad (1)$$

Another process is calculated the roundness that is mean shape of nuclei still round or not. If normal nuclei has the smooth and circle shape but abnormal nuclei is contrary. The roundness factor of normal nuclei approximately 1

$$\text{Roundness} = \frac{4 \times \pi \times \text{Area}}{\text{perimeter}^2} \quad (2)$$

The compactness or various in shape

$$\text{Compactness} = \frac{\text{Perimeter}^2}{\text{Area}} \quad (3)$$

Abnormal cervix tends to has vary in term of shape and size nuclei which is called nuclear polymorphism, so from eq. (3) we calculate the level of variety nuclei. In malignant case, we meet the big nuclei, small nuclei, coarse and dark nuclei.

### B. Cell Morphology

A summary of the parameters of each of the tissue types is shown in table 2. Interestingly, these result shows that nuclear-cytoplasm ratio each types is indicate the high discriminatory power of this feature.

Another common criterion for malignancy is that nuclei vary size and shape includes roundness and shape factor. Two features are extracted and their differentiating capabilities are analysed. The roundness factor is near to 1, it means the shape of nuclei is circle. The shape factor is high, it means nuclei has irregular shape.

## V. CONCLUSION

Examination the morphological feature of histology slides trends to increasing nuclear cytoplasm ratio, shape factor and diameter, decreasing the roundness. Basic image processing techniques and matching methods were used to obtain the dimensions of nuclei and cytoplasm. Area, diameter and perimeter were then obtained and approximate nuclear cytoplasm calculated for normal, pre cancer and malignant.

We can conclude that the changes nuclear cytoplasm ratio and morphology features indicates each types of normal, pre cancer and malignant. It is known that higher N/C ratios are indicative of malignancy. In this study, the cervical histology images are classified into three categories: 1) normal 2) pre cancer 3) malignant. The results achieved using N/C feature is very promising and indicate the high discriminatory power of this feature.

Polymorphism are characterised by varying size and shape of the cells. The shape tend to irregular. Cells are often many times larger than their normal and other cells may be extremely small and primitive appearing.

This papers reports on the classification results achieved using each feature independently. The results indicate that N/C feature has the highest success rates and morphology features can differentiate each class.

## REFERENCES

- [1] National Cancer Institute, 2008
- [2] Data Cancer Female Program in Indonesia, 2008

Table 2. Summary Parameters

- [3] Wang, Yin Hai, Computer Assisted Diagnose of cervical intraepithelial neoplasia using histological virtual slide, Queen University of Belfast, 2008.
- [4] Thiran, J.-P., Morphological Feature Extraction for the Classification of Digital Images of Cancerous Tissues. IEEE, 1996.
- [5] Spotting Cervical Cancer Quicker, *British Journal of Cancer*.2007
- [6] Cervical cancer screening program in developing country, WHO Report, 2002
- [7] Kim, Kwang Baek, Nucleus Segmentation and Recognition of Uterine Cervical Pap-Smears, LNAI 4482, pp. 153–160, 2007. \_c Springer-Verlag Berlin Heidelberg 2007
- [8] Nauth, Hans Frederich, Gynecological Cytology, Thieme, 2007
- [9] Walker, DC, A Study of the Morphological Parameters of Cervical Squamous Epithelium, IOP, pp 1221-135, 2003