DOI: 10.21917/ijivp.2016.0181

COMPUTER AIDED DIAGNOSIS FOR DETECTION AND STAGE IDENTIFICATION OF CERVICAL CANCER BY USING PAP SMEAR SCREENING TEST IMAGES

S. Athinarayanan[1], M.V. Srinath[2] and R. Kavitha[3]

[1]*Department of Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, India*
E-mail: aathithee@gmail.com

[2]*Department of Computer Application, Sengamala Thayaar Educational Trust Women's College, India*
E-mail: drmvsri@gmail.com

[3]*Department of Computer Science and Engineering, PSN College of Engineering and Technology, India*
E-mail: mail2akavithaa@gmail.com

Abstract

The majority of the women of the world were affected by the disease of cervical cancer. As a result of this disease, their death rate was increase as hasty level. Hence so many number of research people was focused this notion as their research interest and also they have done so many number of solutions for finding this cancer by using some image processing technique and achieved a good results only in advanced and high cost techniques of LBC, biopsy or Colposcopy test Images. Therefore the reason, the authors have chosen this problem and also did not only to find whether the patient is affected by a cancer or not. In addition to the patient was affected by this cancer means and also to identify which severity stage of this disease the patient could be live. Then this work has done in based on the images of low cost pap smear screening test by using various image processing techniques with the help of Computerized Image Processing Software Interactive Data Language (IDL-Image Processing Language). Thus the final reports would be very useful to the pathologists for further analysis.

Keywords:

Cancer, Cervical Cancer, Nuclei Segmentation, Feature Identification, Classification

1. INTRODUCTION

More than three hundred thousand women die every year of this disease. Around 80% of these deaths occur in emerging nations like New York, China and India. There is no universal solution regarding the best sexual health intercession. Maximum fitness tutoring platforms commonly lead to a change in knowledge and attitudes, but have less effect on sexual and contraceptive behavior [1]. In developing countries like India, Cancer of the uterine cervix is the second largest cancer in women. In a country like India where the total women population is around 50 crores and the sexually active age group of 25 to 60 is around 30 crores, the need for an indigenous system for automated screening and storage and retrieval of cervical cyto images is the most necessity [1][2][3].

Pap Smear is the most popular screening method to detect cervical cancer from cervicted portion lesion portion. To sense the ration which is exaggerated by cervical cancer, it requires an recognized laboratory, trained cytologist and their repeatable visited reports used for evaluate the results [4][5]. In order to detect cervical cancer automatically, screening tests is used. The results of pap smear are common in our nation and there is no exact method to estimation, this causes the need for a number of visits and invasive procedures such as unnecessary biopsy which cause marvellous cost and unwanted impediment for patients. Consequently, if an another tactic than smear could be found it would help the patients appreciably [6][7]. LBC method was the choice from past 8 years in developed countries that static must not been capable to discover an appropriate residence in our country [8].

Image Segmentation is the one of the most important method of segmenting cell nuclei from the stained specimens. However, isolated nuclei of the cells in high-quality acquisitions provide difficult tricks in the segmentation is more number of nuclei with various characteristics under occurring different acquisition conditions in good-resolution scans of the complete microscope slide[9]. Image segmentation plays a very imperative role for analyzing the objects of an image [10]. Based on the result of segmentation reports features of the cervical cells could be extracted. From that features nuclei regions may be identified. Since the discovery of cervical cancer essentially be contingent on the outcomes of segmentation and each segmentation algorithm works well for assured class of images and not for all images [11]. Finally the Pap Test is compared to the new tests, it's a low cost technique and also by the use of image processing techniques assure to provide good percentage of cancer finding results.

The Bethesda system is used for reporting cervical or vaginal cytologic diagnoses the Pap smear results [12]. Abnormal results of this Bethesda system that is affected abnormal cervical atypical squamous cells are under this category this Low grade squamous intraepithelial lesion (LSIL), High grade squamous intraepithelial lesion (HSIL) [13][14]. LSIL usually indicates mild dysplasia. HSIL indicates moderate, severe or carcinoma in situ [15]. These categories represent the severity stage of the cervical cancer disease for the patients.

2. PROPOSED NOVEL ARCHITECTURE

The author proposed an overall design plan of a new novel Architecture is shown in given Fig.1.

The Fig.1 represents the details about overall proposed system architecture. It consists of many numbers of processes. The detailed descriptions of the process are explained in detail.

In First, Pap Smear Screening Test Slide will be taken and it was examined under the microscope namely Olympus BX51 Microscope with 100× Magnification mode then it captured as Pap Smear Image by using Olympus imaging software. These images are taken by the Pap Smear Public database, Hervlev Hospital, Denmark. The result of the imaging software image is color image. For the purpose of finding the cancer in a medical

image the given color images pap smear Images are converted into gray scale image. After conversion takes places, initial gray scale image contains some noises. Therefore, the reason to remove the noise from that image is by employing filter like anisotropic filter.

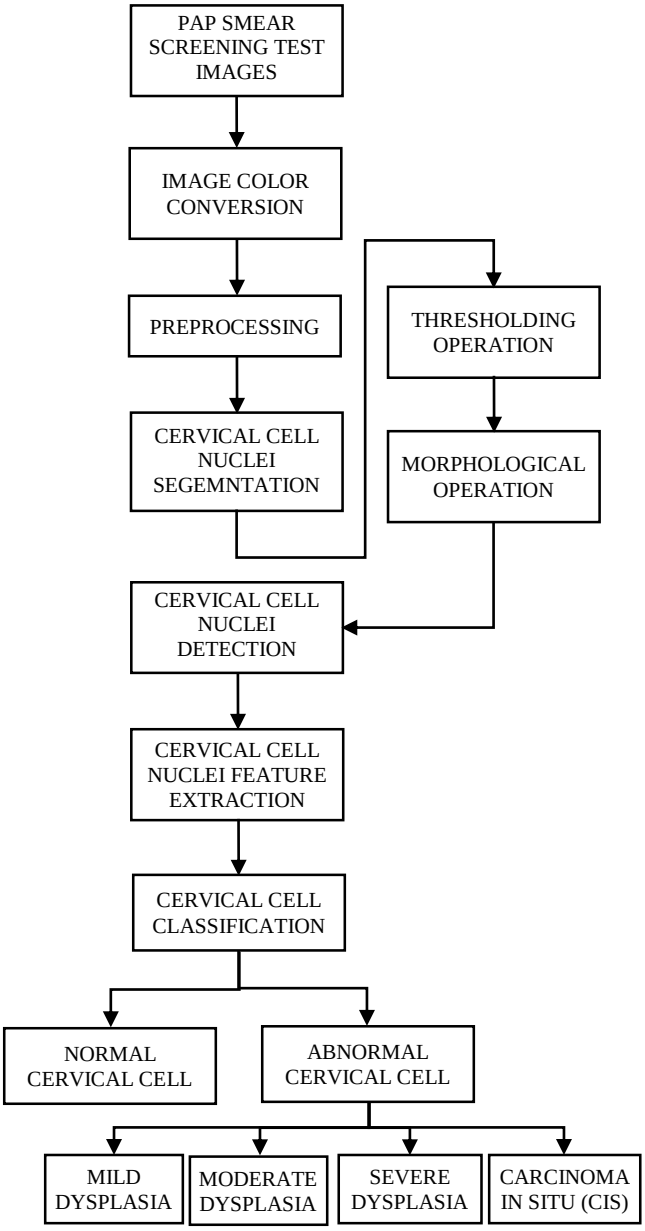

Fig.1. Proposed System Architecture

After this operation was over, got very clear noise removed Pap Smear image. Then this image contains a cervical cell nuclei i.e. neoplasm with the surrounding cytoplasm area. For the purpose of finding cervical cancer, need to segment the cervical cell nuclei from its surrounding cytoplasm. Why do segment the nuclei from its surrounding cytoplasm, because the nuclei comprises squamous intraepithelial lesion (SIL) and it growth may lead to progress low level to high level means the cervical cancer will be developed. Therefore the reason segmentation process to be done. This process is based on two operations like thresholding and morphological operation. In first to apply Local thresholding operation to the noise removed image and it's used for finding where the nuclei presence in an image. As, a result, will get somewhat cleared nucleus with some little mixture amount of unwanted background information. The unwanted background information such as cytoplasm area is removed from the result and threshold image. Morphological closing operation is applied for strengthening the portion of the nuclei and its shape. The operation is done to get cleared segmented cervical cell nuclei image.

From the last image finding the nuclei very clearly means to identify the edge of the nuclei as required. So for finding its edges, the nuclei segmented gray image is converted into a binary image. After that apply all the line scanning directions algorithms namely horizontal, vertical, +45 degree and -45 degree to that binary image, to acquire the edge of nuclei is obviously. This was shown as edge of the nuclei as white pixel remains all black pixels. As the resultant image was mapped into segmented cervical image means to get efficient edge based segmented cervical cell nuclei image. Then this final image, thus arrived, will be clearer to detect the nucleus with the edges than the initial pap smear gray scale image.

Followed by the nucleus detection process, features of the nucleus are found for feature extraction process. In this process, especially with the help of the XROI tool of the Interactive Data Language (IDL(Computerized Image Processing Software)), will be extracted the various features of the nucleus like area, perimeter, total no of pixels occupied by nucleus, minimum value, maximum value, mean value and standard deviation very accurately.

Then the last process of this proposed system is classification process. In this process is used for classified given cervical cell was normal or abnormal based on the extracted selective feature of the nucleus like area value. Then how to find this value in nucleus based on the decreased or increased area values of the nucleus. Then the results were represented a given cell was normal means we come to a conclusion, that cell having the particular patient was not affected by the cervical cancer. But the results represented a given cell was abnormal means the particular patient was affected by the cervical cancer. Not only to end my proposed system whether the patient is affected by this cervical cancer, but also it is used for finding the severity of the cancer based on the above said same feature (area) up, down tuned values of the nucleus. According to the cervical cancer Bethesda system these values based abnormal cell was further classified into mild, moderate, Severe or CIS dysplasia respectively. Hence finally a present status of the patient has been analyzed efficiently means the doctor is going to treat them as right way and cure from the disease.

3. IMPLEMENTATION AND RESULTS

The following diagram represents the details about the implementation methodology of the cervical cancer detection and classification.

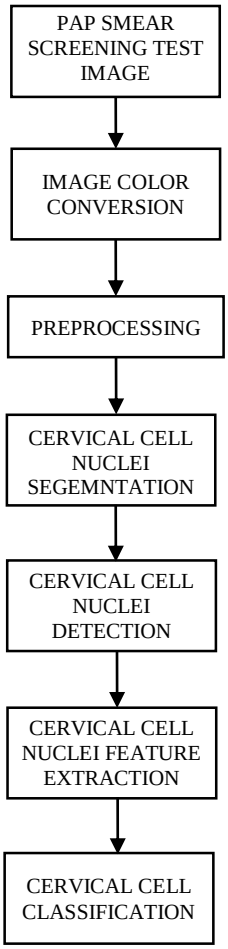

Fig.2. Main Process of the Proposed Work

Above, Fig.2 shows seven main processes of the proposed work, i.e. Pap Smear Image Collection, Pap Smear Image Color Conversion, Pre-Processing Stage, Cervical Cell Nuclei Segmentation, Cervical Cell Nuclei Detection, Cervical Cell Nuclei Feature Extraction, Cervical Cell Classification. Let us see one by one in detail.

3.1 PAP SMEAR IMAGE COLLECTION

This is the first process and the goal of this process is used for converting the Pap Smear Slides into Pap Smear Images. This process was done by mostly in Pathological Laboratory. For this purpose, experimental Pap smear images are acquired through a powerful micro scope by the skilled cyto-technicians. All images were captured with a resolution of 0.201 μm/pixel and it was taken from the open bench mark database of cervical cancer, Herlev University Hospital, Denmark. As a result of this process the author has taken the dataset 491 numbers of different dimensions of colored Pap Smear Images with the following class distribution:

- Normal dysplasia 48 images
- Mild dysplasia, 113 images
- Moderate dysplasia, 105 images
- Severe dysplasia, 110 images
- Carcinoma in situ, 114 images

Fig.3. Cervical Cyto Pap Smear Image

3.2 PAP SMEAR IMAGE COLOR CONVERSION

For the purpose of finding cancer from the color image, the author has converted these colored Pap Smear Images into gray scale images. This images are actual suitable to the additional processing like Pre-processing, segmentation, detection and classification. The Fig.4 shows a Pap Smear Images in Gray Scale Format [16].

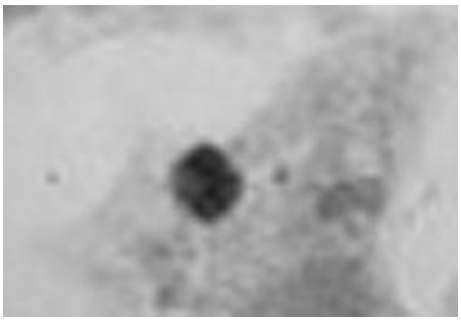

Fig.4. Grayscale Format

3.3 PRE-PROCESSING STAGE

This process is used for improve the appearance of an image. This procedure has done for before processing the original image. Mostly any images contain some unwanted noises initially. Therefore the reason the author just applied anisotropic filter for reducing the noise from the Pap Smear gray images. This filter gives the better results than the mean, median filter of preserving useful details in an image. Similar to the mean filter, this filter contains each pixels of an image look at its nearby neighbor pixels to make a decision, whether it's surrounding pixel or not. After the result of this process, we got much cleared anisotropic filter applied noise removed Pap Smear images and Fig.5 shows the Noise Filtered Image.

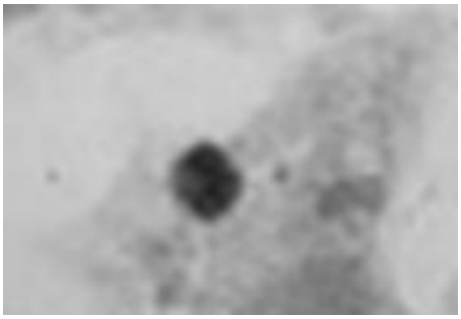

Fig.5. Noise Filtered Image

3.4 CERVICAL CELL NUCLEI SEGMENTATION

The next process of the implementation methodology is Cervical Cell Nuclei Segmentation process. After the process of the noise removal from the pap smear images. The author needs to segment the nucleus from its surrounding cytoplasm for cancer cell identification purpose. Why do we need to segment the nuclei from its surrounding cytoplasm? Because, the nuclei holds SIL and it growth may lead to progress low level to high level means the cervical cancer will be developed. Hence, the reason this segmentation process was done. This process comprises two kinds of the methods, 1) Thresholding Process 2) Morphological Closing Process.

3.4.1 Thresholding Process:

This process is based on two operations like thresholding and morphological operation. In first to apply Local thresholding operation to the noise removed image and it's used for finding where the nuclei presence in an image. As, a result, will get somewhat cleared nucleus with some little mixture amount of unwanted background information shown in given Fig.6.

Fig.6. Thresholding Image

3.4.2 Morphological Closing Process:

Hence, to remove the unwanted background information such as cytoplasm area from the result and thresholded image morphological closing operation is applied for strengthening the portion of the nuclei and its shape. Once the operation is done, clear segmented cervical cell nuclei image is obtained. The result of the above process, got the efficient shape of the nucleus is an image as shown in Fig.7.

Fig.7. Morphological Closing Image

3.5 CERVICAL CELL NUCLEI DETECTION

For getting the edge of the nuclei for the result of segmented image, this image is converted into binary image. The binary image is shown in given Fig.8. The four line masks and its application to the binary images are shown in Fig.9 to Fig.12. Then by combining all the ways of the line scanning yield to catch the edge of the nucleus in the binary image was done. This result is shown in Fig.13 and it has shown the results signified that all the boundaries of the nucleus and its color was white and remain pixels of an image represent black color. From this image, we have applied these edges into the segmented gray cervical cell image and we get the exact edges of the nucleus of the gray scale image by using the succeeding two conditions;

1) If the pixel value of the binary edge image is 1 means, there is no need to change the corresponding pixel value of the segmented gray image.
2) If the pixel value of the binary edge image is 0 means, the corresponding pixel value of the segmented gray image can be taken into the account.

The final image, thus arrived, will be clearer to detect the nucleus with the edges than the initial Pap smear gray scale image. The result of this image is shown in given Fig.14.

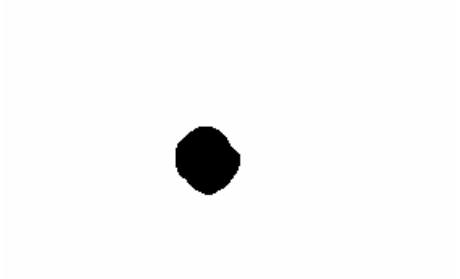

Fig.8. Binary image

-1	-1	-1
2	2	2
-1	-1	-1

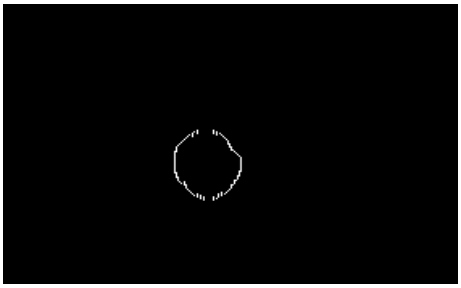

Fig.9. Horizontal Line Mask and Its Image

-1	2	-1
-1	2	-1
-1	2	-1

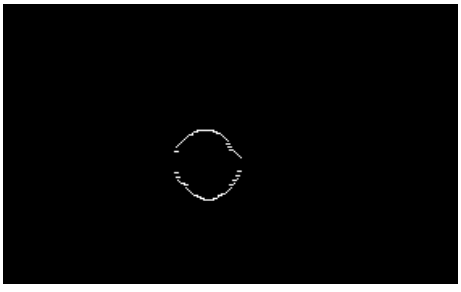

Fig.10. Vertical Line Mask and its Image

-1	-1	2
-1	2	-1
2	-1	-1

Fig.11. +45 Degree Line Mask and its Image

2	-1	-1
-1	2	-1
-1	-1	2

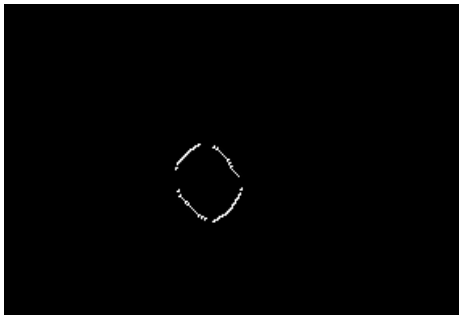

Fig.12. -45 Degree Line Mask and its Image

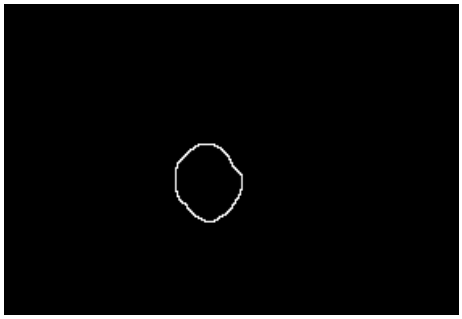

Fig.13. Edge of the Nucleus

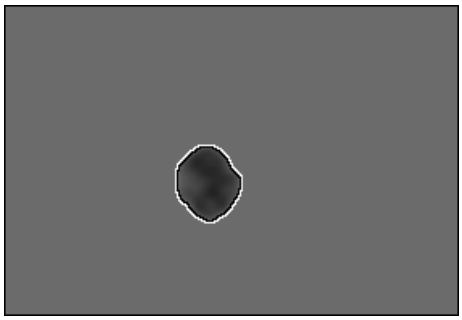

Fig.14. Nucleus Edge Detected Image

3.6 CERVICAL CELL NUCLEI FEATURE EXTRACTION

Followed by the nucleus detection process, features of the nucleus are found for cancer identification purpose. For this process, nucleus outline criteria is going to select based on its corresponding edges by using draw freehand ROI tool of the ROI window based on XROI Command of the Interactive Data Language IDL (Computerized Image Processing Software). ROI means Region of Interest. Simply it refers in this paper as a portion of the nucleus. The Fig.15 shows the nucleus criteria selection based on its edges. After selected criteria of the nucleus, ROI information window is appeared on the screen. This window displayed many features of nucleus like area, perimeter, pixels, minimum value, maximum value, mean and standard deviation and its corresponding values. The resultant ROI window is shown in given Fig.16.

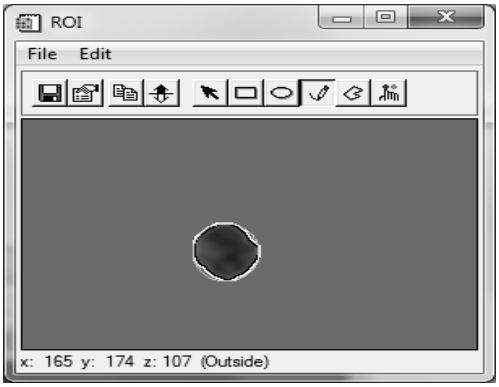

Fig.15. Nucleus Criteria Selection

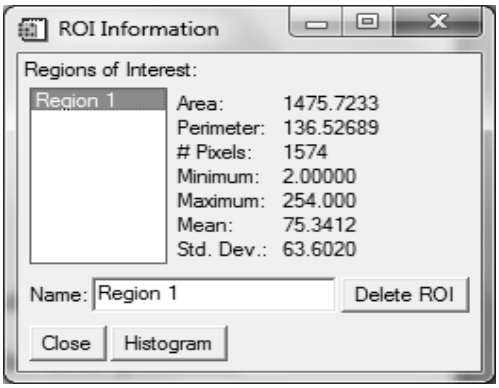

Fig.16. Resultant ROI Window

3.7 CLASSIFICATION OF CERVICAL CELL

Followed by the Extracted Features of the nucleus, based on selected feature like area value, further this cell is going to conclude as normal cell or abnormal cell (cancerous cell). Then if the cell was abnormal means in addition to find the cell is belongs to which stage of the cancer. This Classification process purely based on decision tree based classification system of figure and the table was shown in Fig.17 and Table.1. This table was formed according to the principles of Bethesda System of the cervical cancer and also this table contains predefined nucleus area values based classification and how these area values are calculated by means of tested 48 Normal Cells, and 113 Mild Dysplasia

cells,105 Moderate Dysplasia cells, 110 Severe Dysplasia cells, 115 Carcinoma In Situ Stages of cells of Pap Smear Bench Mark Data, Herlev Hospital, Denmark. So, totally 491 tested results based. This resultant rule based classification Table.1 based the given cervical cell image is classified. The results are given in Table.2.

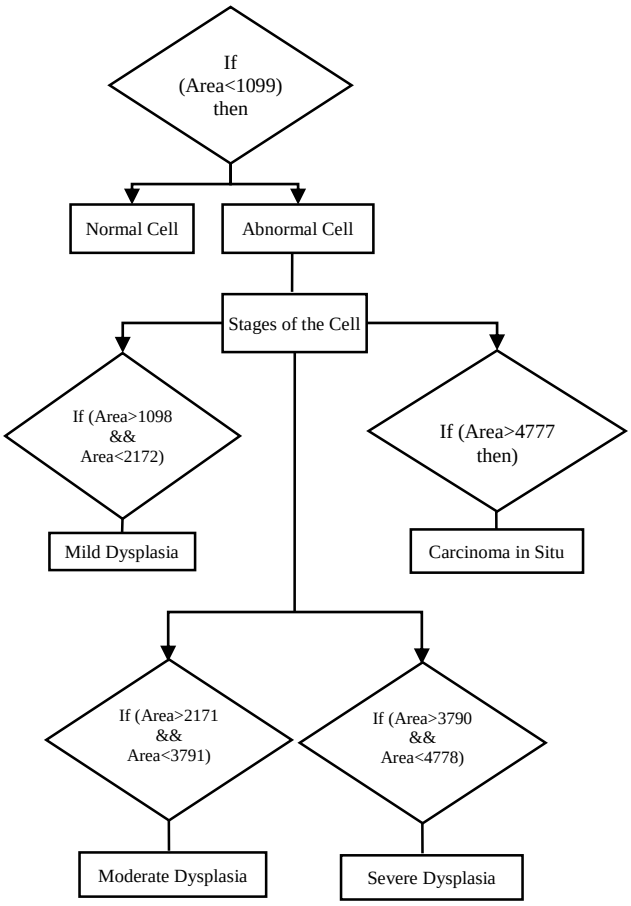

Fig.17. Decision Tree Classification

Table.1. Bethesda System and Cell Classification table

Cell Type (Bethesta System)		Area Value (Decision Tree Based)
Normal Cervical Cell		If (area < 1099.0000) then
Abnormal Cell	Mild Dysplasia	If (area > 1098.0000 && - Area < 2172.0000) then
	Moderate Dysplasia	If(area > 2171.0000 && - Area < 3791.0000) then
	Severe Dysplasia	If (area > 3790.0000&& - Area < 4778.0000) then
	Carcinoma in Situ	If(area >4777.0000) then

Table.2. The Pap Smear Images and its Results

Pap Smear Cell Image	
Area Value	1475.7233
Result	Abnormal Cell
Stage of the Cancer	Mild dysplasia

3.8 PERFORMANCE ANALYSIS

After the result for each cell has been gained through the automated detection, a binary classification set is applied to live the performance of the system, whether is normal or abnormal cell. And also the cell become abnormal cell means the severity classes further classified into four classes, that is Mild dysplasia, Moderate dysplasia, Severe Dysplasia and Carcinoma in Situ (CIS). This study was analyzed from a taken of out of 491 images (Normal dysplasia 48 images, Mild dysplasia, 113 images., Moderate dysplasia, 105 images, Severe dysplasia 110 images, Carcinoma in situ 115 images) this identification result done manually by the medical specialist and the images which was taken by Olympus BX-51magnifier with the resolution 0.201µm of the pap smear database of Herlev hospital, Denmark. The comparison between the results aimed from the output of the automated system of our proposed system (area of the nuclei based identification system) with automated system of existing color intensity based labelled system (mean value of the nuclei based identification system - Eko Supriyanto, 2011) shown in given Table.3 and corresponding results plotted graph is shown in Fig.18 to Fig.22.

Table.3. Comparison between Existing and Proposed System

Class Predicted	Existing/ Proposed	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
Normal	**Existing**	5	36	4	3	90.0%	55.5%	85.4%
Normal	**Proposed**	6	40	1	1	97.5%	85.7%	95.8%
Mild	**Existing**	89	6	8	10	42.8%	91.7%	84.0%
Mild	**Proposed**	93	7	7	6	50.0%	93.0%	88.4%
Moderate	**Existing**	81	10	7	7	58.8%	92,0%	86.6%
Moderate	**Proposed**	87	13	2	3	86.6%	96.6%	95.2%
Severe	**Existing**	90	2	8	10	20.0%	91.8%	83.6%
Severe	**Proposed**	98	4	4	4	50.0%	96.0%	92.7%
Carcinoma	**Existing**	86	4	12	13	25.0%	87.7%	78.2%
Carcinoma	**Proposed**	102	5	4	4	55.5%	96.2%	93.0%

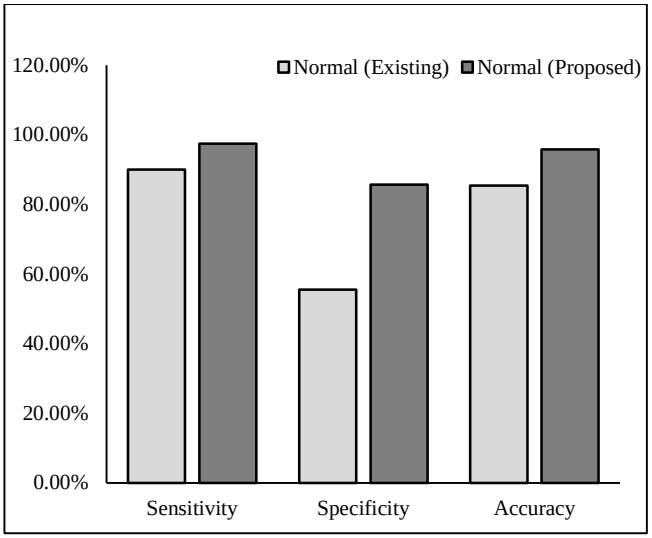

Fig.18. Normal dysplasia comparison graph

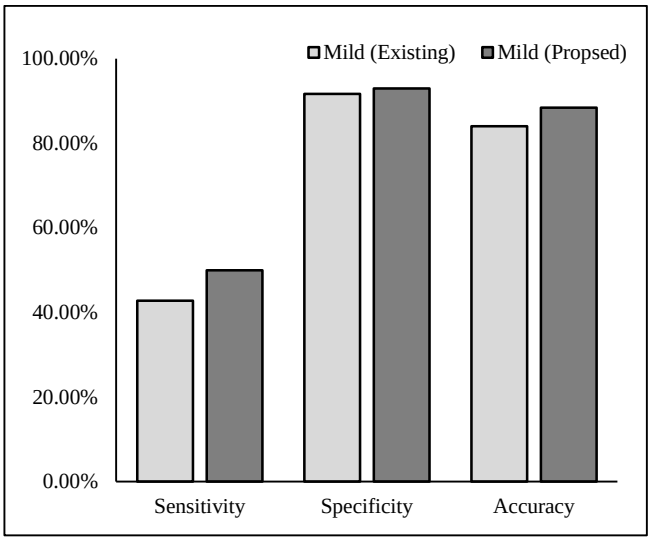

Fig.19. Mild Dysplasia comparison graph

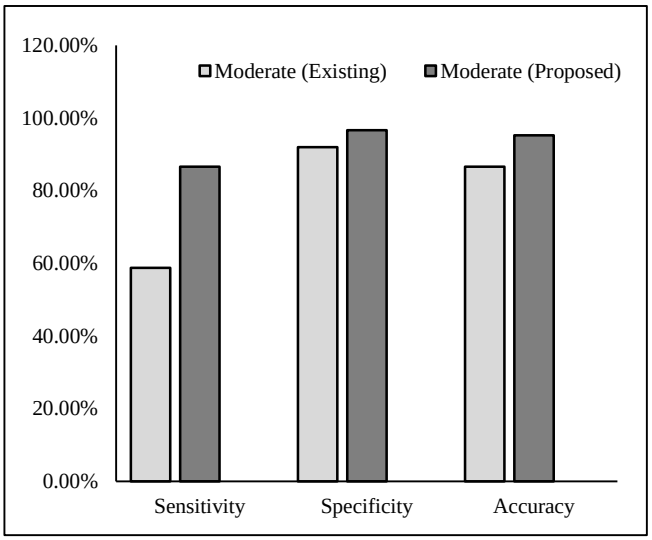

Fig.20. Moderate dysplasia comparison graph

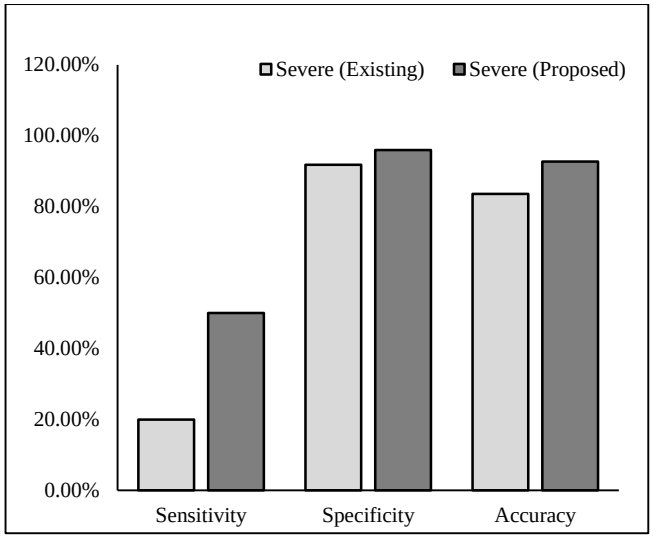

Fig.21. Severe Dysplasia comparison graph

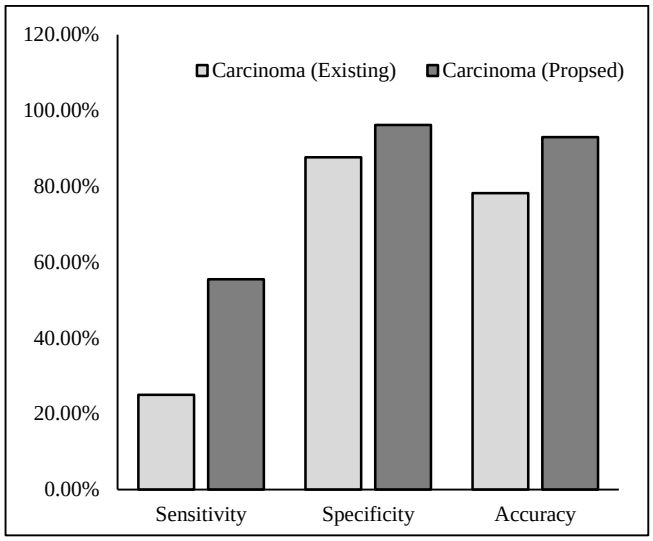

Fig.22. Carcinoma in Situ Dysplasia comparison graph

4. CONCLUSION

This work was focused for tried to save the women from the cervical cancer disease and reduce the death rate of it. Once the system were implemented globally means, if any of the women have gone to the clinic for taking pap smear screening test during their first visit, they must known the current status of this disease and according to the gossips of this tests, they will be diagnoses as the correct way.

REFERENCES

[1] Nadia Islam, Simona.C. Kwon, Ruby Senie and Navneetkathuria, "Breast and Cervical Cancer Screening among South Asian Women in New York City", *Journal of Immigrant and Minority Health*, Vol. 8, No. 3, pp. 211-221, 2006.

[2] Xingjie Wang, Carolyn Fang, Yin Tan, Andy Liu and Grace X. Ma, "Evidence-Based Intervention to Reduce Access Barriers to Cervical Cancer Screening Among Underserved

Chinese American Women", *Journal of Women's Health*, Vol. 19, No. 3, pp. 463-469, 2010.

[3] Shang Ying Hu, Ying Hong, Fang Hui Zhao, Adam K. Lewkowitz, Feng Chen, Wen Hua Zhang, Qin Jing Pan, Xun Zhang, Cindy Fei, Hui Li and You Lin Qiao, "Prevalence of HPV Infection and Cervical Intraepithelial Neoplasia and Attitudes towards HPV Vaccination among Chinese Women Aged 18-25 in Jiangsu Province", *Chinese Journal of Cancer Research*, Vol. 23, No. 1, pp. 25-32, 2011.

[4] Mustafa Guven and Caglar Cengizler, "Data Cluster Analysis-based Classification of Overlapping Nuclei in Pap Smear Samples", *Bio Medical Engineering OnLine*, Vol. 13, No. 1, pp. 1-19, 2014.

[5] Cao Feng, Chen Shuzhen and Zeng Libo, "New Abnormal Cervical Cell Detection method of Multi-Spectral Pap Smears", *Wuhan University Journal of Natural Sciences*, Vol. 12, No. 3, pp. 476-480, 2007.

[6] Patrik Malm, Byju N. Balakrishnan, Vilayil K. Sujathan, Rajesh Kumar and Ewert Bengtsson, "Debris Removal in Pap-Smear Images", *Computer Methods and Programs in Biomedicine*, Vol. 111, No. 1, pp. 128-138, 2013.

[7] Thanatip Chankong, Nipon Theera-Umpon and Sansanee Auephanwiriyakul, "Automatic Cervical Cell Segmentation and Classification in Pap Smears", *Computer Methods and Programs in Biomedicine*, Vol. 113, No. 2, pp. 539-556, 2014.

[8] Ritu Nayar and David Wilbur, "*The Bethesda System for Reporting Cervical Cytology*", 2nd Edition, Springer, 2015.

[9] M.C. Jobin Christ and R.M.S. Parvathi, "A Survey on MRI Brain Segmentation", *Advances in Computer Science, Engineering and Applications*, Vol. 166, pp. 167-177, 2012.

[10] Chitti Kokil Kumar, Arun Agarwal and Raghavendra Rao Chillarige, "Color and Texture Image Segmentation", *Multi-disciplinary Trends in Artificial Intelligence*, pp. 69-80, 2012.

[11] Ning Yu An and Chi Man Pun, "Color Image Segmentation using Adaptive Color Quantization and Multiresolution Texture Characterization", *Signal Image and Video Processing*, Vol. 8, No. 5, pp. 943-954, 2014.

[12] Edmund S. Cibas and Syed Z. Ali, "The Bethesda System for Reporting Thyroid Cytopathology", *American Society for Clinical Pathology*, Vol. 132, No. 5, pp. 658-665, 2009.

[13] Norma McFarlane Anderson, Patience E. Bazuaye, Maria D. Jackson, Monica Smikle and Horace M. Fletcher, "Cervical Dysplasia and Cancer and the use of Hormonal Contraceptives in Jamaican Women", *BMC Women's Health*, Vol. 8, No. 1, pp. 1-9, 2008

[14] Hong Xin Zhang, Yi Min Song, Su Hong Li, Yu Hui Yin, Dong ling Gao and Kui sheng Chen, "Quantitative Detection of Screening for Cervical Lesions with Thin Prep Cytology Test Clinical Oncology and Cancer Research", *Clinical Oncology and Cancer Research*, Vol. 7, No. 5, pp. 299-302, 2010.

[15] Naomi Jay, J. Michael Berry, Charissa J. Hogeboom, E.A. Holly, Teresa M. Darragh and Joel M. Palefsky, "Colposcopic Appearance of Anal Squamous Intraepithelial Lesions: Relationship to Histopathology", *Diseases of the Colon and Rectum*, Vol. 40, No. 8, pp. 919-928, 1997.

[16] Rana S. Hoda and Syed A. Hoda. "*Atypical Squamous Cells*", Fundamentals of Pap Test Cytology, pp. 81-94, 2007.