Features Classification for the Extracted ROIs of Microscopic Pap Smears Images

Loay E. George, Yusra A. Sroor, Haider G. Hussein

Abstract— Pap smears is a very effective screening test for cervical precancerous. However, hundreds of small windows have to be looked under microscope by a trained cytologist for a single slide from each patient. It makes this process very tedious and erroneous. The proposed system is designed to give a discrimination ability of the adopted features to do successful classification for the extracted ROI (i.e., nucleus of the cells). Two kinds of features have been used for discrimination purpose; the first kind is a set of geometrical features because many of the extracted Pap smear ROI are categorized according to their shapes (i.e., circular, oval, irregular, small, big). While the second kind of features is a set of texture features (local and global color features) because one of the main differences between cancerous cells and other types is the local changes in the brightness.

Keywords— *Pap smear, cervical precancerous, extracted ROI, geometrical features, texture features*

1. INTRODUCTION

Cervical screening using Microscopic Pap smear images is one of the most effective ways of detecting and diagnosing the disease even at an early pre-cancerous stage. During mass screening program there will be huge number of samples to be analyzed and diagnosed and the current manual screening methods are time consuming and restricts the capabilities of the cyto-technicians in diagnosing more samples in shorter time. Therefore there is a need for a support system for faster analysis of samples. In automated system for analysis of Microscope Pap Smears images there are three stages. The first stage is the pre-processing block which constitutes color conversion and enhancement to prepare the image for further processing. The next stage is the processing block whose objective is to segment the nucleus from a single cell. It aids in extracting regions of interest in an image. A good segmentation must be able to separate objects from the background to obtain the region of interest [1]

Following this stage is the feature extraction where the area of the nucleus is extracted for classification of the cervical cells as normal or abnormal.

The feature extraction where the area of nucleus is cytological images has been studied by several researchers [2-8].

*Loay E. George has Ph.D Degree from Physic Department in Digital Image Processing /University of Baghdad-Iraq. Email loayedwar57@yahoo.com
*Yusra A. Sroor, Master Student in Computer Science Department / University of Baghdad-Iraq. Email ysra.sroor@yahoo.com
*Haider G. Husse, FICMS(Pathology), received a fellowship degree in the field of pathology from the Iraqi Council of Medical Specialties and a membership at the international academy of cytopathology/ Central Public Health Laboratory/ Ministry of Health. Email haydergazi@live.com

2. PROPOSED SYSTEM

For the classification purpose of the objects found in Pap smear images, it is found that both types of features are required; that is, the geometrical features and the textural features.

In the following subsections the sets of features belong to both types are described in details.

I) Geometrical Features

The adopted geometrical features for classifying the nucleus regions and the surrounding cytoplasm for the cells appeared in Pap smear images are:

1. The cell area (A): it is determined as the number of pixels belongs to the nucleus area. Usually the cancerous cells are usually larger than the normal cells, while inflammatory cells are usually smaller.
2. The packing ratio (r): it is defined as the ratio,

$$r = P^2 / A \qquad \text{,...(1)}$$

   Where, (P) is the perimeter of the extracted region, and (A) represents the area of the region. The perimeter (P) is determined as the number of pixels lay at the external border line of the region.
   This parameter is very useful to decide whether the test region is circular (i.e., r=12.566) or not, if the region is far from being circular or has irregular boundary then the value or (r) will become larger.
3. The moments of the width and height variations of the region (Mom): this array of moments is determined using,

$$M_V(n) = \frac{1}{x_2 - x_1 + 1} \sum_{x=x_1}^{x_2} V(x) \left(\frac{x - p_{vx}}{p_{vx}} \right)^n \qquad \text{……..........(2a)}$$

$$M_H(n)=\frac{1}{y_2-y_1+1}\sum_{y=y_1}^{y_2}H(y)\left(\frac{y-p_{vy}}{p_{vy}}\right)^n \quad ,\dots\dots(2b)$$

Where,

$V(x)$ is the height of the region at column (x),

$V(y)$ is the width of the region at line (y),

y_1, y_2 are the upper and lower bound values for the vertical extent of the region,

x_1, x_2 are the left and right bound values for the horizontal extant of the region.

p_{vx} , p_{vy} are the vertical & horizontal pivot values. Two ways were used to determine the values of pivot values:

A. Geometrical mean:

$$p_{vx}=\tfrac{1}{2}(x_1+x_2)$$
$$p_{vy}=\tfrac{1}{2}(y_1+y_2) \quad ,\dots\dots(3a)$$

B. Mean or Average:

$$p_{vx}=\frac{1}{x_2-x_1+1}\sum_{x=x_1}^{x_2}x\,H(x)$$
$$p_{vy}=\frac{1}{y_2-y_1+1}\sum_{y=y_1}^{y_2}y\,V(y) \quad ,\dots\dots(3b)$$

4. The Extents Ratio (W_r): it is the ratio of the smallest extent (whether it is the width average or height average) to the longest extent:

$$W_r=\begin{cases}p_{vy}/p_{vx} & if\ \ p_{vy}<p_{vx}\\ p_{vx}/p_{vy} & if\ \ p_{vx}\le p_{vy}\end{cases} \quad ,\dots\dots(4)$$

II) Textural Features

The second type of features includes the set of features depending on the spatial variation in the color & intensity of the image signal within the extracted region of interest. Beside to the famous Haralick's (co-occurrence) texture features the following two sets of features are used:

1. The histogram moments (second, third & forth orders), for the three color components (Blue, Green, Red) and the intensity component.
2. The moments of local roughness in the color components and intensity within the extracted ROI.

In the following subsections some details about the used texture features are given.

$$p_x(i)=\frac{1}{N_g}\sum_{j=0}^{Ng-1}p(i,j),\quad p_y(j)=\frac{1}{N_g}\sum_{i=0}^{Ng-1}p(i,j) \quad ,\dots\dots(5a)$$

And μ_x, μ_y, σ_x and σ_y are the means and standard deviations of the partial probability functions:

$$\mu_x=\sum_{i=0}^{Ng-1}i\,p_x(i),\quad \sigma_x=\frac{1}{N_g}\sum_{i=0}^{Ng-1}(i-\mu_x)^2\,p_x(i) \quad ,\dots\dots(5b)$$

$$\mu_y=\sum_{j=0}^{Ng-1}j\,p_y(j),\quad \sigma_y=\sum_{j=0}^{Ng-1}(j-\mu_y)^2\,p_y(j) \quad ,\dots\dots(5c)$$

The co-concurrence matrix features are:

1. Angular Second Moment (ASM):

$$ASM=\sum_{j=0}^{Ng-1}\sum_{i=0}^{Ng-1}[p(i,j)]^2 \quad ,\dots\dots(6)$$

2. Contrast (C):

$$C=\sum_{n=0}^{Ng-1}n^2\left\{\sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1}p(i,j)\right\},\quad |i-j|=n \quad ,\dots\dots(7)$$

3. Correlation (Cor):

$$Cor=\frac{1}{\sigma_x\sigma_y}\sum_{i=0}^{Ng-1}\sum_{j}^{Ng-1}i\,j\,p(i,j)-\mu_x\mu_y \quad ,\dots\dots(8)$$

4. Variance (V):

$$V=\sum_{i=0}^{Ng-1}\sum_{j}^{Ng-1}(i-\mu_x)p(i,j)^2 \quad ,\dots\dots(9)$$

5. Inverse Difference Moment (IDM):

$$IDM=\sum_{i=0}^{Ng-1}\sum_{j}^{Ng-1}\frac{1}{1+(i-j)^2}p(i,j) \quad ,\dots\dots(10)$$

6. Sum Average (SA):

$$SA=\sum_{i=0}^{2Ng-2}i\,p_{x+y}(i) \quad ,\dots\dots(11)$$

Where x & y are the coordinates (row and column) of an entry in the co-occurrence matrix, and $p_{x+y}(i)$ is the probability of co-occurrence matrix coordinates summing $x+y$

7. Sum Variance (SV)

$$SV=\sum_{i=0}^{Ng-1}(i-SA)^2\,p_{x+y}(i) \quad ,\dots\dots(12)$$

8. Sum Entropy (SH):

$$SH=-\sum_{i=0}^{2Ng-2}p_{x+y}(i)\log_2\left(p_{x+y}(i)\right) \quad ,\dots\dots(13)$$

9. Entropy (H):

$$H=-\sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1}p(i,j)\log_2\left(p(i,j)\right) \quad ,\dots\dots(14)$$

10. Difference Variance (DV):

$$DV=\sum_{k=0}^{Ng-1}[k-\sum_{i=0}^{Ng-1}ip_{x-y}(i)]^2\,p_{|x-y|}(i) \quad ,\dots\dots(15)$$

11. Difference Entropy (DH):

$$DH=-\sum_{i=0}^{Ng-1}p_{|x-y|}(i)\log_2\{p_{|x-y|}(i)\} \quad ,\dots\dots(16)$$

12. Maximum Correlation Coefficient (MCC):

$$MCC=\underset{i,j}{Max}\left(\sum_{k=0}^{Ng-1}\frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}\right) \quad ,\dots\dots(17)$$

13. Information Measure Correlation:

(a) $$IMC_1=\frac{H_{XY}-H_{XY1}}{Max\{H_X,H_Y\}} \quad ,\dots\dots(18a)$$

(b) $IMC_2 = \sqrt{1 - exp(-2(H_{XY2} - H_{XY}))}$,...............(18b)

Where, H_X & H_Y are the entropies of p_x & p_y, respectively. While:

$$H_{XY} = -\sum_{j=0}^{Ng-1}\sum_{i=0}^{Ng-1} p(i,j) log_2(p(i,j)) \quad ,......................(19)$$

$$H_{XY1} = -\sum_{j=0}^{Ng-1}\sum_{i=0}^{Ng-1} p(i,j) log_2(p_x(i) p_y(j)) \quad ,......................(20)$$

$$H_{XY2} = -\sum_{j=0}^{Ng-1}\sum_{i=0}^{Ng-1} p_x(i) p_y(j) log_2(p_x(i) p_y(j)) \quad ,......................(21)$$

2. Histogram Moments of the Color Components

This set of texture features can be considered as global features, because the histogram for all the area of ROI, and the centralized moments, relative to the mean value, are determined using:

$$M(C,n) = \frac{1}{A}\sum_{i=0}^{255}(i - m_C)^n His(C,i) \quad ,......................(22)$$

Where,

C represents the color component (Red, Green, Blue or intensity).

$His(C,i)$ if the number of appearance of the pixel value (i) in the color component (C).

A is region area (i.e., number of pixels belong to extracted ROI).

m_C is the mean of the color component for the pixels belong to ROI:

$$m_C = \frac{1}{A}\sum_{i=0}^{255} i\, His(C,i) \quad ,....................(23)$$

3. Moments of the Roughness Histogram

This set of contrast features depends on the local variations in the pixel values relative to their local mean. The difference between each pixel value and the mean of its surrounding (lay within a window LxL) is considered as a local roughness indicator. The determined roughness values are determined, then the roughness histogram is computed; finally the moments' values are computed.

The roughness indicator is determined as:

$$R_C(x,y) = C(x,y) - mean_C(x,y,L) \quad , ...(24)$$

Where,

C represents the color component (Blue, Green, Red, or Intensity).

$mean(x,y,L)$ is a local operator applied on the area {x'∈ [x-L/2, x+L/2], y'∈ [y-L/2, y+L/2]}.

The roughness moments are determined using the following:

$$M_R(C,n) = \frac{1}{A}\sum_{i=-255}^{255} i\, His_R(C,i) \quad ,......................(25)$$

Where, $His_R(C,i)$ is the frequency number of occurrence of the roughness value (i) within the ROI.

3. FEATURE ANALYSIS & EVALUATION

This stage is required, exclusively, during the enrollment phase. Not all the adopted features can be guaranteed to have acceptable discrimination capability, so a set of analysis tests should be conducted to determine the scattering ratio of each feature in the feature space. Also, the best combination of discriminating features that leads to highest classification rate should found. The number of used features should be as small as possible, and they should lead to best classification rate.

The procedure followed to do the feature analysis task consists of the following steps:

1. Prepare a training set of samples, $\{f_{i,jk} \mid i \in [1, NoFeatures], j \in [1, NoClasses], k \in [1, NoSamlesPerClass]\}$ such that for each type of cells there is enough number of samples.
2. Apply a statistical analysis that consists of the following steps:
   A. For each set of samples belong to certain class of cells determine the mean $\{\mu_{ij} \mid i \in [1, NoFeatures], j \in [1, NoClasses]\}$ and standard deviation $\{\sigma_{ij}\}$.
   B. Test each sample value individually, and discard from the list of samples each feature value (f_{ijk}) doesn't satisfy the condition $|f_{ijk} - \mu_{ij}| < 3\sigma_{ij}$. In case of features discarding, repeat steps A & B till no feature is discarded or 20% of the overall features (belong to certain class and feature type) are discarded.
   C. From the remaining list of values determine the mean and standard deviation for the tested feature and for each class separately.
   D. For each class, determine the within-class scattering ratio for the feature (i):
   $$S_W(i,j) = \sigma_{ij} / \mu_{ij} \quad ,........................(26)$$
   If the values of $S_W(i,j)$ for certain feature (i) is mostly low over all classes then the feature is considered successful according to within scatter test.
   E. For the considered successful features, according to within scatter test, the between scatter ratio $\{S_B(i)\}$ is determined,
   $$S_B(i) = \frac{\sum_{j=1}^{NoClasses} \sigma_{ij}}{\frac{1}{n}\sum_{j=1}^{NoClasses-1}\sum_{k=j+1}^{NoClasses}|\mu_{ij} - \mu_{ik}|} \quad ,....................(27)$$
   Where,
   $$n = \frac{1}{2}(NoClasses+1)(NoClasses+2) \quad ,...............(28)$$
   F. For each feature, if its between-scatter ratio is low (<0.5) then the feature is considered suitable to be used for classification purpose.
   G. From the considered successful set of features, find the best combination of features that lead to highest classification rate. The classification rate is determined using:
   $$R = \frac{No.\ of\ \ Successful\ Hits}{Total\ \ No.\ of\ \ Classification\ Tests} \quad ,..........(29)$$
   H. For the successful features, store the mean values (i.e., template values) and the associated standard deviation for all classes are stored in a dedicated database.

4. MACHING AND CLASSIFICATION DECISION

This stage is the final stage in the classification (i.e., decision making) phase. In this work four the Euclidean distance measure had been used to determine the similarity degree between the feature vectors, extracted from the tested samples, with the template feature vectors, each representing certain class, which are stored in the database.

Four types of Euclidean distance measures have been tested, and the best one that leads to highest classification rate is adopted:

(i) $D(T_i,S_j)=\sum_k \left|T(i).f(k)-S(j).f(k)\right|$,................(30)

(ii) $D(T_i,S_j)=\sum_k \left[T(i).f(k)-S(j).f(k)\right]^2$,................(31)

(iii) $D(T_i,S_j)=\sum_k \left|\frac{T(i).f(k)-S(j).f(k)}{T(i).\sigma(k)}\right|$,...............(32)

(iv) $D(T_i,S_j)=\sum_k \left[\frac{T(i).f(k)-S(j).f(k)}{T(i).\sigma(k)}\right]^2$,................(33)

Where, *T(i).f(k)* is the template value of k^{th} feature belong to i^{th} class. *S(j).f(k)* is the value of k^{th} feature extracted from j^{th} sample.

5. FEATURE ANALYSIS AND RECOGNITION RESULTS

Table (1) presents the number of samples used for testing the proposed features to make the classification task in this project. The number of samples used for representing each cell type is not same; this is due to the frequency of occurrence of each cell type.

Table (1) The list of number of ROI samples used in the feature analysis & testing stage

Class Type	Class ID Number	No. of Samples
Abnormal (Cancerous)	1	74
Artificial	2	80
Normal: Endocervical	3	6
Normal: Metaplastic	4	11
Normal: Squamous	5	42

Table (2) lists the features, with their group index, used to make classification for the whole taken samples of ROI. These features are classified into five groups; two of depends on the geometry of the ROI, while the other three groups consist of texture features.

The discrimination power of the suggested features were evaluated by determining the within and between scatter ratio and it is found the "size" feature parameter plays good roll in discriminating a large percentage of cancerous cells and artificial objects.

Table (2) The list of Geometrical and Texture Features

Feature No	Group of Features	Feature Name
0	(1) Geometrical	Size
1		PackRatio
2		AspectRat
3-6	(2) Global Texture (Histogram Moments)	Avg, Std, Skw, Kor; Blue
7-10		Avg, Std, Skw, Kor; Green
11-14		Avg, Std, Skw, Kor; Red
15-18		Avg, Std, Skw, Kor; Intensity
19-20	(3) Geometrical {Moments, M(n)} of Width and Height Variations) Relative to a Pivot Value	M(2) & M(4); Width Variation; Average
21		M(3); Width Variation; Average
22		M(3); Absolute of Width Variation; Average
23		M(1); Width Variation; Mid
24		M(1); Absolute of Width Variation; Mid
25		M(2); Width Variation; Mid
26		M(2); Width Variation; Mid
27		M(3); Width Variation; Mid
28		M(3); Absolute of Width Variation; Mid
29		M(4); Width V; Mid
30		M(4); Absolute of Width, Mid
31-32		M(2) & M(4); Height Variation; Average
33		M(3); Height Variation; Average
34		M(3); Absolute of Height Variation; Average
35		M(1); Height Variation; Mid
36		M(1); Absolute of Height Variation; Mid
37		M(2); Height Variation; Mid
38		M(2); Height Variation; Mid
39		M(3); Height Variation; Mid
40		M(3); Absolute of Height Variation; Mid
41		M(4); Height Variation; Mid Value
42		M(4), Absolute of Height Variation; Mid
43-56	(4) Local Texture (Co-occurrence Matrix)	Blue (Quant=8)
57-70		Blue (Quant=16)
71-84		Green (Quant=8)
85-98		Green (Quant=16)
99-112		Red (Quant=8)
113-126		Red (Quant=16)
127-140		Intensity (Quant=8)
141-154		Intensity (Quant=16)
155-158	(5) Histogram Moments For Roughness (For Blue, Green, Red, Intensity)	M(2); Mask Size=3
159-162		M(3); Mask Size=3
163-167		M(4); Mask Size=3
167-170		M(2); Mask Size=5
171-174		M(3); Mask Size=5
175-178		M(4); Mask Size=5
179-183		M(2); Mask Size=7
183-186		M(3); Mask Size=7
187-190		M(4); Mask size=7

Tables (3) and (4) present the success ratio of each feature alone in signifying the cells type. It is clear that the local texture features (i.e., roughness and concurrence) are the most competitive features for discrimination the ROIs.

Table (5) presents the classification rates when two features are used in the similarity measure, it is obvious that the normalized similarity measures lead to better recognition rates (or hits ratio) in comparison with the associated non-normalized (absolute) measures.

Table (6) presents the classification hits when four features are used, the normalized city block similarity measure led to best classification success hits in comparison with other measures, here the first and second features were set (F_1=0, F_2=175) for the case of normalized city block similarity measure. While for other measure this set of features is taken (F_1=158, F_2=190, for non-normalized city block measure; F_1=156, F_2=190 for non-normalized measure; and F_1=106, F_2=182 for normalized Euclidean measure).

Tables (7) and (8) present the classification success hits when six and eight features, respectively taken. The set of the first part of features was set (F_1=0, F_2=175, F_3=41, F_4=62, F_5=34, F_6=139) for the normalized city block measure. As shown in both tables (7) and (8) the best success rate was always occurred when using normalized city block.

With the gradual increase of the number of features used in the similarity measure the classification hit rate was increased, and its increase was stopped after using:

1. Eight features for both kinds of normalized measures.
2. Six features for the non-normalized features.

As shown in the tables the texture features plays important role in the classification decision, while the geometrical features (except, the size and aspect ratio, i.e. F(0) and F(1)) plays a minor roll. Also, the local texture features perform better than the global texture features in the classification decision process

Table (3) The best found success rate of the features when using the similarity measure given by equations (30 or 31)

Feat No.	Feat. Group	Success%	Feat No.	Feat. Group	Success%
185	5	65.73	17	2	44.13
188	5	65.26	43	4	44.13
186	5	64.79	91	4	39.44
182	5	64.32	122	4	39.44
184	5	64.32	131	4	39.44
172	5	45.07	150	4	39.44
174	5	45.07	158	5	39.44
106	4	44.60			

Table (4) The best found success rate of the features when using the similarity measure given by equations (32 or 33)

Feat No.	Feat. Group	Success%	Feat No.	Feat. Group	Success%
115	4	66.20	151	4	46.01
143	4	66.20	0	1	45.54
59	4	61.50	38	3	42.72
87	4	61.50	63	4	42.72
101	4	59.15	64	4	42.72
65	4	46.01	91	4	42.72
93	4	46.01	92	4	42.72
123	4	46.01			

Table (5) The classification Success Hits when using Double Features

Abs, equation (30)		
F1	F2	Rate%
158	190	70.89
66	185	70.42
94	185	70.42
156	186	70.42
156	190	70.42

Square, equation (31)		
F1	F2	Rate%
156	190	71.36
156	186	70.42
157	186	70.42
160	190	70.42
162	190	70.42

Abs/Std, equation (32)		
F1	F2	Rate%
0	175	72.77
7	182	72.77
7	180	71.36
7	184	71.36
83	187	71.36

Square/Abs, equation (33)		
F1	F2	Rate%
106	182	71.83
169	184	71.83
106	185	71.36
106	186	71.36
106	190	71.36

Table (6) The classification Success Hits when using Four Features

Abs, equation (30), F1=158, F2=190		
F3	F4	Rate%
65	118	71.83
65	146	71.83
73	132	71.83
93	118	71.83
93	146	71.83

Square, equation (31), F1=156, F2=190		
F3	F4	Rate%
1	118	72.77
1	146	72.77
4	118	72.30
4	146	72.30
5	118	72.30

Abs/Std, equation (32), F1=0, F2=175		
F3	F4	Rate%
41	62	84.04
41	90	84.04
41	118	84.04
41	146	84.04
35	118	83.57
35	146	83.57

Square/Std, equation (33), F1=106, F2=182		
F3	F4	Rate%
0	7	79.34
0	15	79.34
0	62	79.34
0	90	79.34
0	118	79.34
0	146	79.34

Table (7) The classification Success Hits when using six Features

Abs, F1=158, F2=190, F3=65, F4=118		
F5	F6	Rate%
1	122	72.30
1	150	72.30
9	21	72.30
9	31	72.30
9	32	72.30
9	34	72.30
18	21	72.30

Square, F1=156, F2=190, F3=1, F4=118		
F5	F6	Rate%
1	1	72.77
1	2	72.77
1	4	72.77
1	5	72.77
1	6	72.77
1	8	72.77
1	9	72.77

Abs/Std, F1=0, F2=175, F3=41, F4=62		
F5	F6	Rate%
34	139	86.85
32	139	85.92
34	71	85.92
31	111	85.92
40	139	85.92
42	139	85.92
32	71	85.92

Square/Abs, F1=106, F2=182, F3=0, F4=7		
F5	F6	Rate%
33	76	83.57
1	36	83.10
1	38	83.10
11	33	83.10
19	138	83.10
22	138	83.10
24	138	83.10

Table (8) The classification Success Hits when using Eight Features

Abs, F1=158, F2=190, F3=65, F4=118, F5=1, F6=122			Square, F1=156, F2=190, F3=1, F4=118, F5=1, F6=1		
F7	F8	Rate%	F7	F8	Rate%
1	56	72.77	1	1	72.77
1	100	72.77	1	2	72.77
1	109	72.77	1	4	72.77
1	112	72.77	1	5	72.77
1	126	72.77	1	6	72.77
1	154	72.77	1	8	72.77

Abs/Std, F1=0, F2=175, F3=41, F4=62, F5=34, F6=139			Square/Std, F1=106, F2=182, F3=0, F4=7, F5=33, F6=76		
F7	F8	Rate%	F7	F8	Rate%
24	74	87.32	1	156	84.98
26	74	87.32	0	127	84.51
28	74	86.85	0	164	84.51
30	74	86.85	1	40	84.51
74	116	86.85	1	129	84.51
74	144	86.85	0	31	84.51

The results indicated that the performance of geometrical less than the texture features.

1. As an overall classification performance, the use of texture features is more suitable than depending on geometrical features.
2. The feature "Size" show good discrimination power for both the cancerous cells (because they mostly appear relatively large) and artificial cells (because they mostly appear small).
3. The classification performance of local based texture features (i.e., roughness & co-occurrence) is better than using global based features.
4. The use of concurrence features, only, leads to performance results which can be considered as promising results. Same encouraging results occurred when using roughness based features.
5. The artificial areas are mainly shows high correlation in their image signal due to its smoothness. Also, they show low roughness values in comparison with real types of cells area.
6. The cancerous cells show high degree of roughness in comparison with the normal types of cells.

REFERENCES

[1] Rafael C. Gonzalez, Richard E. Woods, ”***Digital Image Processing***”, Pearson Education, Inc. and Dorling Kindersley Publications, Inc.

[2] L. Yanxi, Z. Tong and Z. Jiayong, "***Learning Multispectral Texture Features for Cervical Cancer Detection***", the Robotics Institute, Carnegie Mellon University, N01-CO-07119, 2002.

[3] Z. Jiayong and L. Yanxi, "***Cervical Cancer Detection Using SVM Based Feature Screening_ The Robotics Institute***", Carnegie Mellon University, NIH award N01-CO-07119 and PA-DOH grant ME01-738, 2004.

[4] S. Jeremiah, A. Bernadetta, P. Maruli and S. Winfried, "***Pattern Recognition on 2D Cervical Cytological Digital Images for Early Detection of Cervix Cancer***", Dept. of Biomedical Engineering, Swiss German University, IEEE 978-1-4244-5612-3/09, 2009.

[5] P. Marina, N. Christophoros and C. Antonia, "***Combining Shape, Texture and Intensity Features for Cell Nuclei Extraction in Pap Smear Images***", IEEE Transactions on Information Technology in Biomedicine, Vol. 15, No. 2, March 2011.

[6] E. Supriyanto, N. Azureen, M. Pista, L. Ismail, B. Rosidi and T. Mengko, "***Automatic Detection System of Cervical Cancer Cells Using Color Intensity Classification***". Recent Researches in Computer Science, ISBN: 978-1-61804-019-0, 2011.

[7] B. Rosidi, N. Jalil, N. Pista, L. Ismail and E. Supriyanto, "***Classification of Cervical Cells Based on Labeled Colour Intensity Distribution***", International Journal of Biology and Biomedical Engineering, Issue 4, Vol. 5, No. 2/3, 2011.

[8] M. Lipi, N. Dilip and N. Chandan, "***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.

[9] Haralick, R. M., "***Statistical and structural approaches to Texture***", Proc. IEEE, Vol. 67, No. 5, PP. 786-804, 1979.