

Hyperchromasia and Texture as effective features for analysis of malignancy in Pap smear Images

Lipi B Mahanta¹ and Kangkana Bora²

Assistant Professor

¹*Central Computational and Numerical Studies
Institute of Advanced Study in Science and Technology
Guwahati, Assam*

²*Regional Institute of Science and Technology
9th Mile, Guwahati, Assam*

¹lipimahanta@yahoo.co.in, ²kangkana.bora89@gmail.com

Abstract

This work presents an approach for the analysis of abnormality in the cervical cells based on Texture and presence of Hyperchromasia, which are two important morphological features based on which one can distinguish between normal and abnormal cervical cells. The proposed approach is implemented in MATLAB®, a high level, interactive environment for data visualization/analysis/computation. This may help pathologist in identification of cervical cancer from Pap smear images and help in early diagnosis.

Keywords: *Pap smear, Cervical cancer, Histogram, Gray-level Co-occurrence Matrix, Hyperchromasia*

1. Introduction

Following breast cancer, cervical cancer is the second most prevalent cancers among women. According to a 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 and lack of awareness[2]. So, early detection of cervical cancer may take a very precious role in preventing this disease. Some of the features for abnormality identification are the texture analysis and to check the presence of Hyperchromatic Crowded Groups (HCG). In medical imaging, images are used to gain maximum information for diagnostic purpose and those images are digital in nature. In the study of malignancy in cervical cancer Pap smear images are used. Pap smear test is a quick, simple and essentially painless test to detect abnormal cervical cells. These Pap smear images are used as an input for our approach which will study the texture and HCGs of the cervical cells and based on the parameters obtained, given samples can be classified into two groups *i.e.*, normal class and abnormal classes.

2. Overview of Cervical Cancer

Cancer is a condition caused by uncontrollable growth of cells. Cervical cancer is a common cancer in women which is found in ectocervix, endocervix and in the transformation zone (Figure 1). Over 90% case of cervical cancer arises in transformation zone [7].

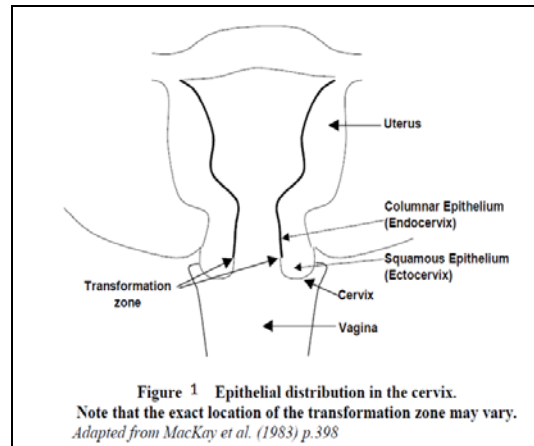


Figure 1. Epithelial Distribution of Cervix

There are many factors that lead to cancer:

- 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

3. Texture and Abnormality

Texture is an important characteristic for the analysis of many types of images. But a standard definition of texture doesn't exist. But for a simple definition Image texture is a function of the spatial variation in pixel intensities (gray values), is useful in a variety of application and has been a subject of intense study by many researchers [4]. Using texture analysis various features like homogeneity, Contrast, Energy, Correlation and frequency of images can be calculated. The texture of images also refers to the appearance, structure and arrangement of the parts of an object within the image [3]

It can be observed that cancer cells are different from normal cervical cells as normal cells pose a homogeneous texture and the frequency distribution of the gray levels is also homogeneous. So analysis of textures of Pap smear image can take an important role in cervical cancer detection.

For our approach we have combined two texture analysis approaches:

- a) Histogram analysis combined with frequency calculation of the image pixels.
- b) Gray level Co-occurrence matrix (GLCM).

Both of these techniques belong to Statistical approach methods which mainly represent texture based on distribution and relationship between gray levels.

4. Hyperchromasia and Abnormality

Another morphological feature which can indicate the presence of abnormality is the presence of Hyperchromatic Crowded Groups (HCG). Hyperchromasia is nothing but presence of a cluster of cells which contain more than 15 cells. Following figure shows the presence of Hyperchromatic Crowded Group (HCG).



Figure 2. Presence of HCG (Source: University College of London Site [5])

5. Proposed Method

In this approach abnormality analysis is performed on 976 images which are collected from Dr B.Baroo Cancer Institute, Guwahati , website of University college of London (they provide nano-zoomer using which one can visualize the original slides online) [5] and MDE laboratory[6].

Sample collected from Dr B.Baroo Cancer Institute – 24

Sample collected from University college of London site – 35

Sample collected from MDE Lab – 917

The Pap smear images are in .jpg and .bmp formats.

Snapshot of a part of Image Database is shown in Figure 3-



Figure 3.Sample Database Collected from MDE Lab

5.1. Texture Analysis

5.1.1. Histogram Approach

The histogram of an image is the count of how many pixels in the image possess a given grey-level value. For an 8 bits image, this may be represented by a graph with an x coordinate ranging from 0 to 255 and a y coordinate representing the respective pixel count. Figure 4.b shows the histogram of the image in Figure 4.a.

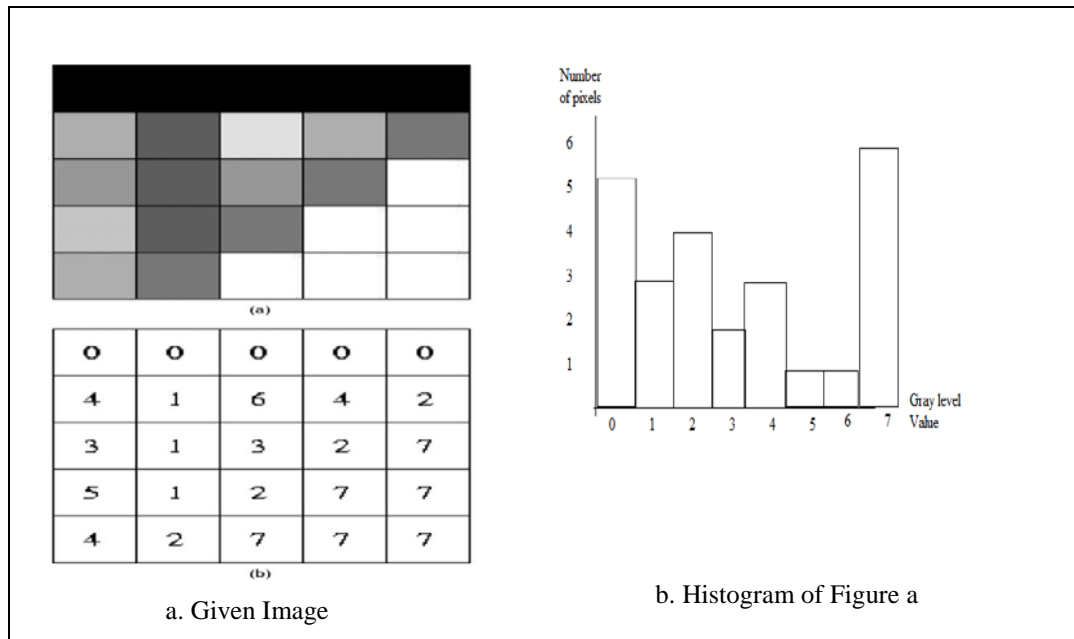


Figure 4. Histogram Representation of an Image

The following measures are taken in case of texture analysis –

1. The colored image is converted into gray scale image using rgb2gray function.
2. Histogram is evaluated using imhist function.
3. Frequency count is calculated from histogram which will give the occurrence of each pixel value having a particular gray level.
4. Then the frequency count is presented in a graph form where x co-ordinate represents the gray level and y co-ordinate represents the no of pixel having a particular gray level.

In case of an abnormal cervical cell the graph will have many maxima and minima points. But in case of normal cervical cell the graph will exhibit a smooth nature.

5.1.2 GLCM Approach

The gray-level co-occurrence matrix (GLCM) is a frequency matrix and a tabulation of how often different combinations of pixels brightness values (Gray levels) occur in an image. The GLCM indicates the frequency of a pair of pixels that are at “Exactly the same distance and direction of displacement vector” [4].

GLCM, c , is defined with respect to given (row, column) displacement h . And element (i, j) , denoted c_{ij} , is the number of times a point having gray level j occurs in position h relative to a point having gray level i . Let N_h be the total number of pairs, then $C_{ij}=c_{ij}/N_h$ is the elements of the normalized GLCM [8].

The co-occurrence probability provides a second-order method for generating texture features. These probabilities represent the conditional joint probabilities of all pair wise combinations of grey levels in the spatial window of interest given two parameters: Interpixel distance (δ) and orientation (θ). The probability measure can be defined as:

$$\Pr(x) = \{C_{ij} | (\delta, \theta)\}$$

Where C_{ij} (the co-occurrence probability between grey levels i and j) is defined as:

$$C_{ij} = \frac{P_{ij}}{\sum_{i,j=1}^G P_{ij}}$$

Where P_{ij} represents the number of occurrences of grey levels i and j within the given window, given a certain (δ, θ) pair;

Figure 5. Equation of Co-occurrence Probability [4]

G is the quantized number of gray levels. Orientation are of four types i.e. 0 degree, 45 degree, 90 degree and 135 degree. Figure 4 shows the orientation –

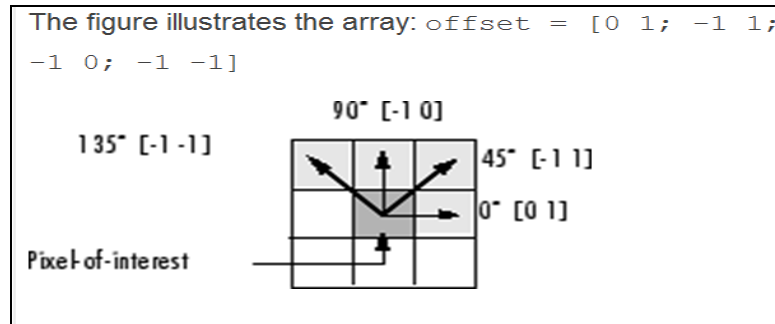


Figure 6. Orientation (θ) [8]

θ is represented as offset which is a p-by-2 array of integers specifying the distance between the pixel of interest and its neighbor. Each row in the array is a two-element vector, $[row_offset, col_offset]$, that specifies the relationship, or *offset*, of a pair of pixels. row_offset is the number of rows between the pixel-of-interest and its neighbor. col_offset is the number of columns between the pixel-of-interest and its neighbor. Because the offset is often expressed as an angle, the following table lists the offset values that specify common angles, given the pixel distance D .

Angle	Offset
0	[0 D]
45	[-D D]
90	[-D 0]
135	[-D -D]

For our approach we considered D value as 1.
One example of GLCM calculation is shown below:

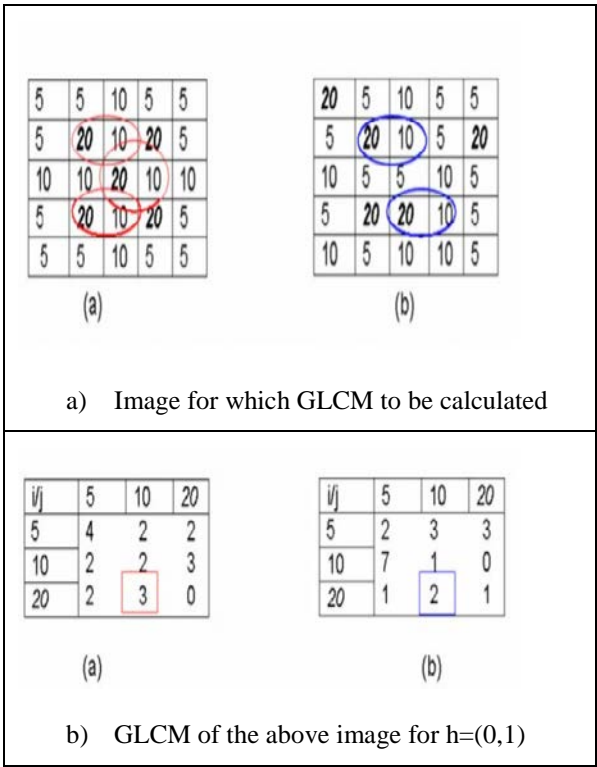


Figure 7. Example of GLCM Calculation [4]

Steps used to identify the texture properties using GLCM are as follows –

- a) Convert the rgb image to gray scale image using rgb2gray function.
- b) Construct the GLCM using function graycomatrix.
- c) Extracting properties like homogeneity, Contrast and Energy using function graycoprops.
- d) Analysis of output data and drawing a conclusion.

List of features extracted are shown in the following table –

Table 1. Feature that can be Extracted from GLCM

<i>PARAMETER</i>	<i>DESCRIPTION</i>	<i>FORMULA</i>
CONTRAST	Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. Range = $[0 \text{ (size(GLCM) - 1)}^2]$ The Contrast is expected to be low if the gray levels of each pixel pair are similar. Contrast is 0 for a constant image.	$\sum_{i,j} i - j ^2 P(i, j)$
HOMOGENEITY	Measures the local homogeneity of a pixel pair. The Homogeneity is expected to be large if the gray levels of each pixel pair are similar. Range = $[0 \text{ } 1]$	$\sum_{i,j} \frac{P(i, j)}{1 + i - j }$
CORRELATION	Returns a measure of how correlated a pixel is to its neighbor over the whole image. Range = $[-1 \text{ } 1]$ Correlation is 1 or -1 for a perfectly positively or negatively correlated image. The Correlation is expected to be high if the graylevels of the pixel pairs are highly correlated.	$\sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) P(i, j)}{\sigma_i \sigma_j}$
ENERGY	Measures the number of repeated pairs. The Energy is expected to be high if the occurrence of repeated pixel pairs is high. Energy is 1 for a constant image.	$\sum_{i,j} P(i, j)^2$

For our approach we are considering Homogeneity, Contrast and Energy.

5.2. To Check the Presence of HCG

For detecting the Hyperchromatic Crowded Groups following measures are taken into consideration.

1. Select the image for which HCG to be identified.
2. Crop the particular region of interest (ROI), i.e. the portion which is dark and crowded.
3. Find the number of nucleus inside the ROI.
4. If number_of_nucleus > 15 then “Hyperchromasia is present” and if < 15 then “Hyperchromasia is absent”.

Presence of Hyperchromasia leads to abnormality in cervical cells.

Step 3 of this approach comprise of two steps:

- a) Preprocessing - The step is to perform pre-processing operations on the input Pap smear image. The input image is a colored image which is in RGB mode. As is known that the color images are difficult to process so we convert the image to

grayscale 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 grayscale 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.

- b) Morphological Operations – 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 [10]. 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. Then no of objects are counted and this total count will give no of nucleus present in the image.

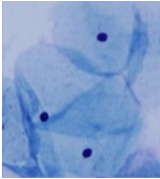
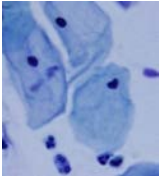
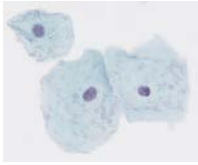
6. Results and Discussion

6.1. Texture Analysis

6.1.1 Histogram Approach

Texture of Normal Pap smear sample: The samples images taken as an input to the system (Only 6 images are shown here) are shown in the following table –

Table 2. Sample Normal Pap Smear Images taken for Frequency Calculation using Histogram Approach

 <p>Sample 1 (Source: Dr B.Barooch Cancer Institute)</p>	 <p>Sample 2 (Source: Dr B.Barooch Cancer Institute)</p>	 <p>Sample 3 (Source: University College of London Site [5])</p>
--	---	---

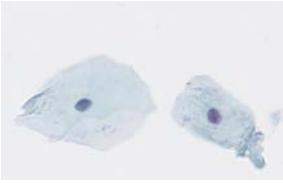
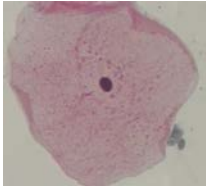
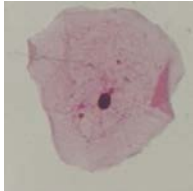
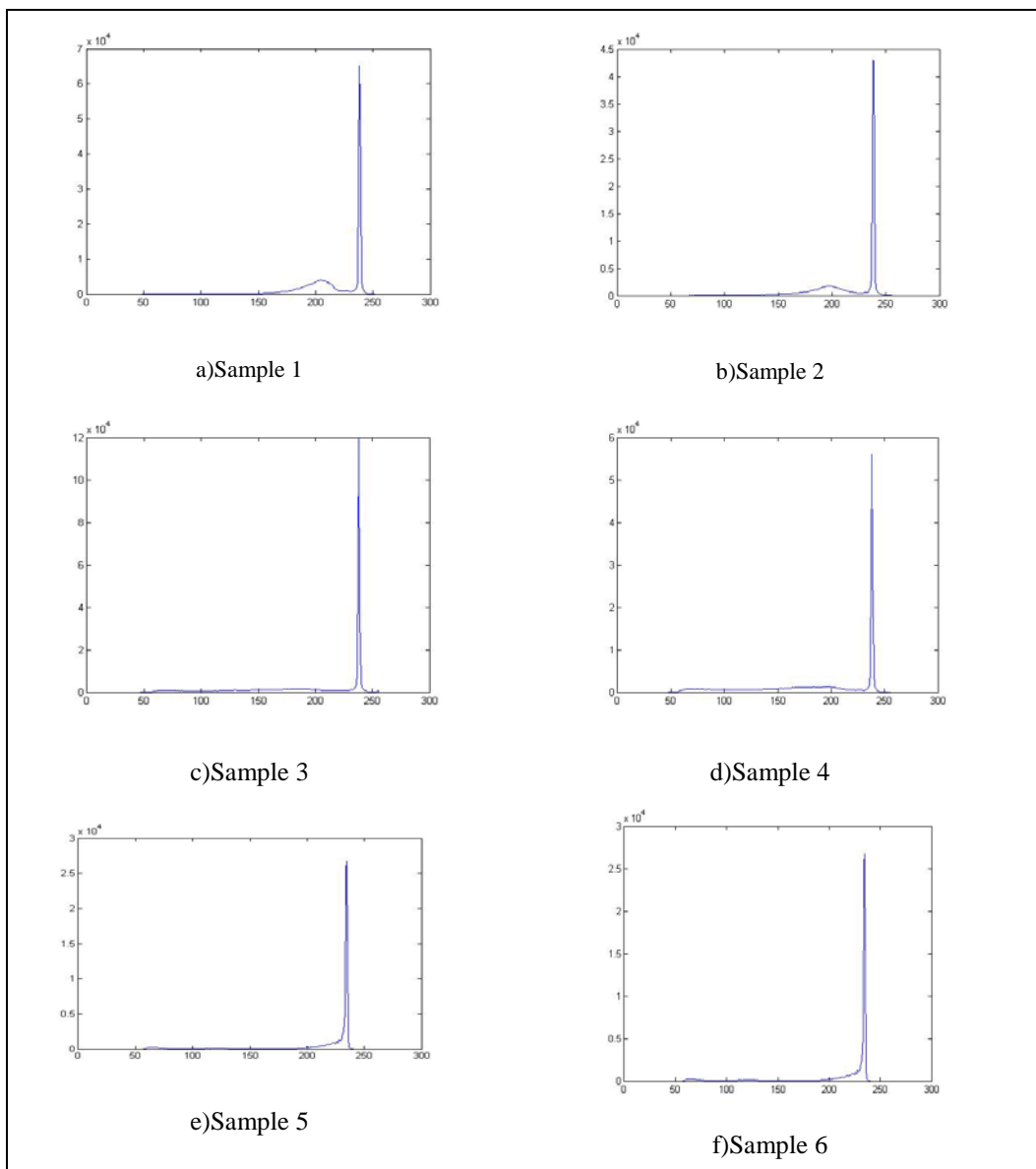
 <p>Sample 4 (Source :University College of London Site [5])</p>	 <p>Sample 5 (Source: MDE Lab[6])</p>	 <p>Sample 6 (Source MDE Lab[6])</p>
---	--	---

Table 3. Output of The System For The Images Of Table 1. In Frequency Graphs Of Normal Sample, X Co-Ordinate Shows The Gray Levels And Y Co-Ordinate Shows The Probability Of Pixels Having A Particular Gray Levels



Texture of Abnormal Pap smear sample

Table 4. Sample Abnormal Pap Smear Images taken for Frequency Calculation using Histogram Approach

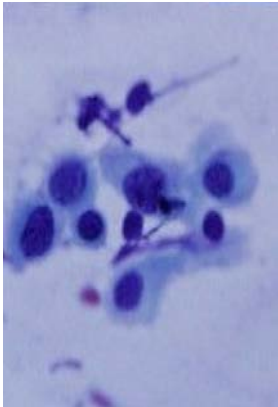
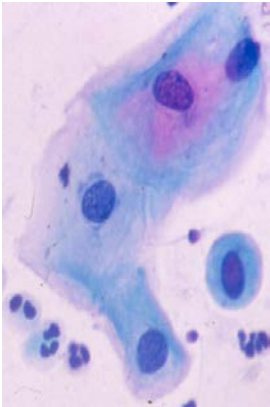
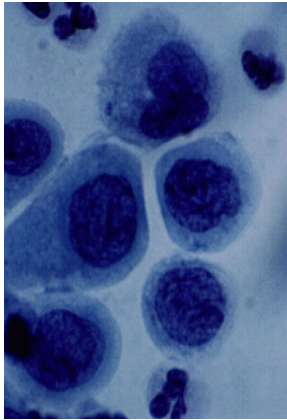

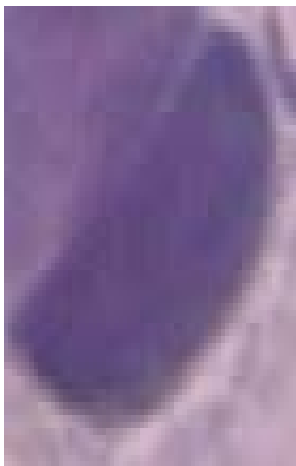
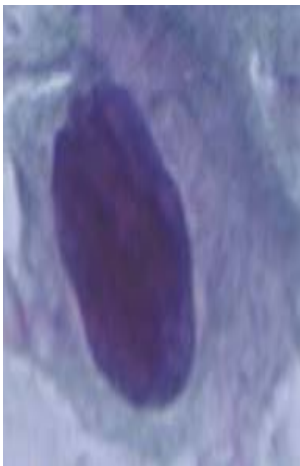
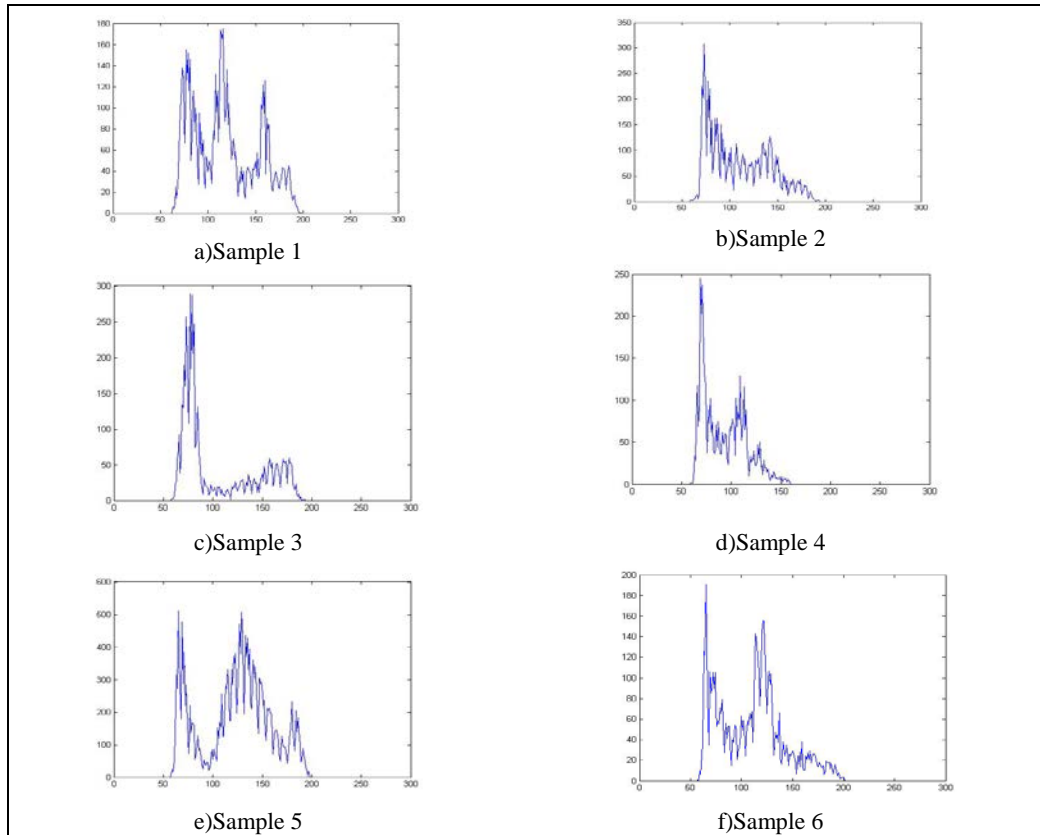
		
Sample 1 (Source: Dr B.Barooch Cancer Institute)	Sample 2 (Source: Dr B.Barooch Cancer Institute)	Sample 3 (Source: University College of London site [5])
		
Sample 4(Source: University College of London site [5])	Sample 5 (Source: MDE Lab [6])	Sample 6 (Source: MDE Lab [6])

Table 5. Output of the System for the Images of Table 4. In Frequency Graphs of Normal sample, X co-ordinate shows the Gray Levels and y co-ordinate shows the Probability of Pixels having a Particular Gray Levels



From our observations it can be concluded that Normal Pap Smear samples graphs are almost similar and have maxima points in the range from 200-250. And the graph is smooth in nature. But from abnormal Pap smear samples we can observe that the graph is having many maxima and minima points with lots of fluctuation having no specific range where it will have its maximum points. There is no similarity among the graphs.

6.1.2. GLCM Approach

Normal sample of 20 Pap smear images that are taken into consideration is in the following figure –



Figure 8. Normal Pap Smear Image Samples given as an Input To The System for Homogeneity, Contrast and Energy Calculation (Source: MDE lab [6])

Abnormal sample of 20 Pap smear images that are taken into consideration is in the following figure –

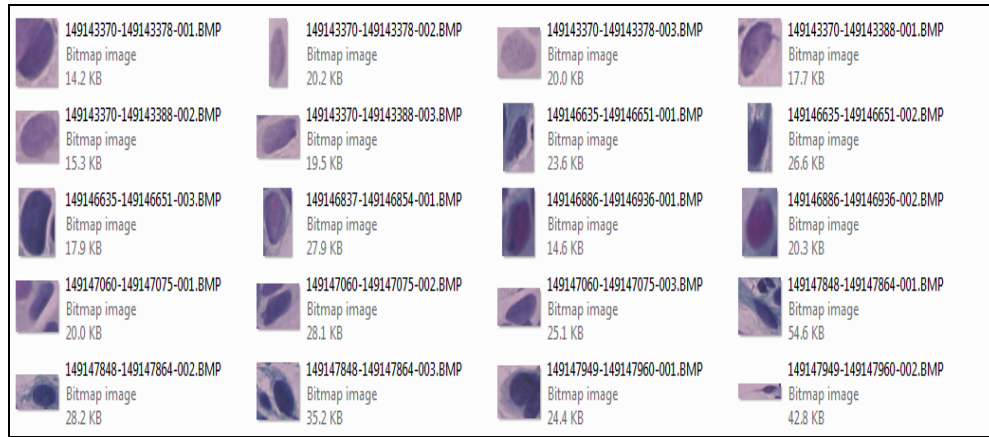


Figure 9. Abnormal Pap Smear Image Samples given as an Input to the System for Homogeneity, Contrast and Energy calculation (Source: MDE lab [6])

Homogeneity Measures: Only 40 samples are shown here where 20 are normal and 20 are abnormal samples.

Table 6. Homogeneity Value of 20 Normal and 20 Abnormal Samples

Normal	Sample no	Homogeneity				
		0	45	90	135	Sum
	1	0.9189	0.9093	0.9189	0.9093	3.6564
	2	0.945	0.9142	0.9215	0.9142	3.6949
	3	0.9664	0.9479	0.9533	0.9509	3.8185
	4	0.9501	0.922	0.9358	0.932	3.7399
	5	0.9489	0.9186	0.9275	0.9207	3.7157
	6	0.9442	0.9108	0.92	0.9164	3.6914
	7	0.935	0.9	0.9074	0.9018	3.6442
	8	0.9574	0.9402	0.9436	0.936	3.7772
	9	0.9533	0.9324	0.9395	0.928	3.7532
	10	0.9569	0.9316	0.9412	0.9355	3.7652
	11	0.96	0.9484	0.9508	0.9376	3.7968
	12	0.9629	0.9442	0.9404	0.943	3.7905
	13	0.9426	0.9119	0.9225	0.9133	3.6903
	14	0.9499	0.9215	0.9271	0.9234	3.7219
	15	0.9516	0.9257	0.9314	0.9221	3.7308
	16	0.9533	0.9324	0.9395	0.928	3.7532
	17	0.9569	0.9316	0.9412	0.9355	3.7652
	18	0.96	0.9484	0.9508	0.9376	3.7968
	19	0.9629	0.943	0.9442	0.9404	3.7905
	20	0.9536	0.9332	0.9326	0.926	3.7454
	Maximum	0.9664	0.9479	0.9533	0.9509	3.8185
	Minimum	0.9189	0.9	0.9074	0.9018	3.6442
Normal samples						
Abnormal Sample no		Homogeneity				
		0	45	90	135	sum
	1	0.8952	0.8844	0.8952	0.8752	3.55
	2	0.8907	0.8629	0.8907	0.8517	3.496
	3	0.8635	0.8483	0.8635	0.8422	3.4175
	4	0.8886	0.883	0.8886	0.8458	3.506
	5	0.8872	0.8778	0.882	0.8644	3.5114
	6	0.8651	0.8673	0.8131	0.8366	3.3821
	7	0.8651	0.8929	0.915	0.8669	3.5399
	8	0.8728	0.8899	0.9109	0.8998	3.5734
	9	0.888	0.8873	0.8866	0.8612	3.5231
	10	0.8392	0.8833	0.8919	0.8743	3.4887
	11	0.831	0.8811	0.8037	0.8962	3.412
	12	0.8202	0.8782	0.8921	0.8696	3.4601
	13	0.8056	0.83	0.878	0.8759	3.3895
	14	0.8168	0.8747	0.8937	0.8754	3.4606
	15	0.8208	0.8989	0.8226	0.892	3.4343
	16	0.8024	0.8809	0.8972	0.8675	3.448
	17	0.8156	0.8852	0.9039	0.8741	3.4788
	18	0.8966	0.8478	0.8759	0.8519	3.4722
	19	0.8234	0.895	0.8085	0.8858	3.4127
	20	0.842	0.8989	0.8186	0.812	3.3715
	Maximum	0.8952	0.8929	0.915	0.8998	3.5734
	Minimum	0.8024	0.83	0.8037	0.812	3.3715
Abnormal samples						

From the observation of Homogeneity it is observed that homogeneity of normal sample is in the range of 0.9-1.0 and sum generally fall in the range of 3.6-4.0. In case of abnormal sample the homogeneity value is in the range 0.7-0.9 and range of the sum is 3.2-3.6. Here sum indicates the summation of homogeneity values for all the orientations.

Contrast measures:

Table 7. Contrast Values of 20 Normal and 20 Abnormal Samples

Normal	Sample no	Contrast				
		0	45	90	135	SUM
	1	0.1205	0.1824	0.1625	0.1821	0.6475
	2	0.1103	0.1723	0.1574	0.1677	0.6077
	3	0.1072	0.1569	0.1489	0.168	0.581
	4	0.1197	0.1824	0.168	0.1868	0.6569
	5	0.0671	0.105	0.0938	0.0988	0.3647
	6	0.1	0.1604	0.1288	0.1378	0.527
	7	0.1027	0.1679	0.147	0.163	0.5806
	8	0.1117	0.1792	0.1605	0.1674	0.6188
	9	0.1047	0.1671	0.1481	0.1567	0.5766
	10	0.13	0.1221	0.1854	0.1972	0.6347
	11	0.0939	0.1395	0.1237	0.1524	0.5095
	12	0.1004	0.1581	0.1461	0.1539	0.5585
	13	0.0968	0.1501	0.1383	0.1586	0.5438
	14	0.0865	0.1431	0.1201	0.1316	0.4813
	15	0.0806	0.1051	0.1004	0.1349	0.421
	16	0.0742	0.1151	0.112	0.1202	0.4215
	17	0.1231	0.1861	0.1599	0.1852	0.6543
	18	0.0929	0.134	0.135	0.1483	0.5102
	19	0.1197	0.1883	0.1811	0.1996	0.6887
	20	0.0981	0.1212	0.0877	0.1235	0.4305
	MAXIMUM	0.13	0.1883	0.1854	0.1996	0.6887
	MINIMUM	0.0671	0.105	0.0877	0.0988	0.3647

a) Normal samples

Abnormal Sample no	Contrast				
	0	45	90	135	SUM
1	0.2792	0.2651	0.222	0.296	1.0623
2	0.2235	0.2838	0.2224	0.3205	1.0502
3	0.2074	0.3214	0.2818	0.3585	1.1691
4	0.2102	0.2736	0.2547	0.4298	1.1683
5	0.2733	0.2577	0.2538	0.3623	1.1471
6	0.2774	0.2829	0.298	0.4094	1.2677
7	0.2088	0.2328	0.2712	0.2063	0.9191
8	0.2694	0.323	0.2972	0.4019	1.2915
9	0.2752	0.2447	0.2723	0.2271	1.0193
10	0.2871	0.2137	0.2245	0.3726	1.0979
11	0.299	0.2277	0.2791	0.2771	1.0829
12	0.2892	0.2756	0.2333	0.4009	1.199
13	0.199	0.2227	0.2787	0.2617	0.9621
14	0.2124	0.2083	0.2866	0.2074	0.9147
15	0.2486	0.2068	0.2405	0.2096	0.9055
16	0.21	0.2343	0.3421	0.2345	1.0209
17	0.2775	0.3149	0.2712	0.3624	1.226
18	0.2887	0.3023	0.2392	0.2633	1.0935
19	0.1937	0.3341	0.3292	0.3193	1.1763
20	0.289	0.3214	0.299	0.41	1.3194
MAXIMUM	0.299	0.3341	0.3421	0.4298	1.3194
MINIMUM	0.1937	0.2068	0.222	0.2063	0.9055

b) Abnormal samples

After observing 976 images (20 of them are shown here) it can be concluded that the contrast values is in the range of 0.08-0.2 and sum will be in the range of 0.3-0.85 for normal samples. In case of abnormal samples the contrast range is 0.2-0.5. Sum will be the range of 0.85-1.5. Here sum indicates the summation of contrast values for all the orientations.

Energy measure:

Table 8. Energy Values of 20 Normal and 20 Abnormal Samples

Normal	Sample no	ENERGY				
		0	45	90	135	SUM
	1	0.2735	0.2494	0.2576	0.2496	1.0301
	2	0.2999	0.2767	0.2823	0.2789	1.1378
	3	0.2785	0.2597	0.2648	0.2628	1.0658
	4	0.3387	0.3123	0.3242	0.3208	1.296
	5	0.2514	0.2255	0.2321	0.2269	0.9359
	6	0.2844	0.2525	0.2597	0.2578	1.0544
	7	0.3426	0.3141	0.3223	0.3193	1.2983
	8	0.2609	0.2301	0.2367	0.2323	0.96
	9	0.2815	0.2667	0.2684	0.2625	1.0791
	10	0.2424	0.221	0.2254	0.2225	0.9113
	11	0.3357	0.3073	0.3129	0.3035	1.2594
	12	0.3152	0.2963	0.3019	0.2928	1.2062
	13	0.2991	0.2746	0.2831	0.2774	1.1342
	14	0.3977	0.3764	0.3772	0.3732	1.5245
	15	0.2952	0.2759	0.2744	0.27	1.1155
	16	0.2498	0.2245	0.2272	0.2205	0.922
	17	0.2987	0.2745	0.2826	0.2726	1.1284
	18	0.349	0.3286	0.3349	0.328	1.3405
	19	0.2909	0.2764	0.2838	0.2789	1.13
	20	0.3111	0.2941	0.298	0.2941	1.1973
	MAXIMUM	0.3977	0.3764	0.3772	0.3732	1.5245
	MINIMUM	0.2424	0.221	0.2254	0.2205	0.9113

Normal Sample

Abnormal Sample no	ENERGY				
	0	45	90	135	SUM
1	0.1815	0.1697	0.1717	0.164	0.6869
2	0.1276	0.1177	0.1276	0.1132	0.4861
3	0.1246	0.1039	0.11	0.1013	0.4398
4	0.1639	0.158	0.1403	0.1536	0.6158
5	0.1677	0.1514	0.154	0.1483	0.6214
6	0.1365	0.1138	0.1134	0.1027	0.4664
7	0.1275	0.1244	0.1342	0.1143	0.5004
8	0.1607	0.159	0.1702	0.1514	0.6413
9	0.1566	0.1436	0.1565	0.1459	0.6026
10	0.1496	0.1437	0.1549	0.1349	0.5831
11	0.1193	0.1133	0.1202	0.1086	0.4614
12	0.1169	0.1115	0.1084	0.0952	0.432
13	0.1676	0.1531	0.1613	0.1551	0.6371
14	0.1728	0.1622	0.1669	0.161	0.6629
15	0.1494	0.1351	0.1403	0.1287	0.5535
16	0.1891	0.1745	0.1788	0.1694	0.7118
17	0.1702	0.1506	0.1577	0.15	0.6285
18	0.1284	0.1109	0.1202	0.1148	0.4743
19	0.1146	0.1986	0.1976	0.1872	0.698
20	0.1946	0.1691	0.177	0.1741	0.7148
MAXIMUM	0.1946	0.1986	0.1976	0.1872	0.7148
MINIMUM	0.1146	0.1039	0.1084	0.0952	0.432

Abnormal Sample


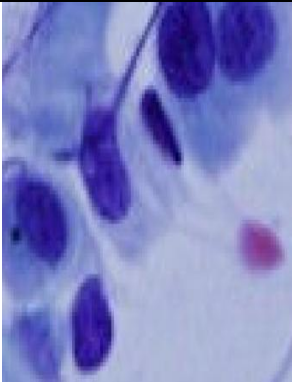
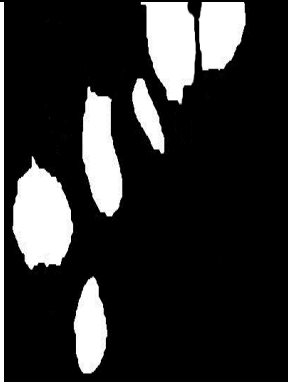
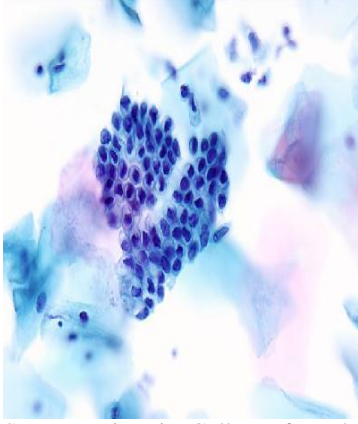
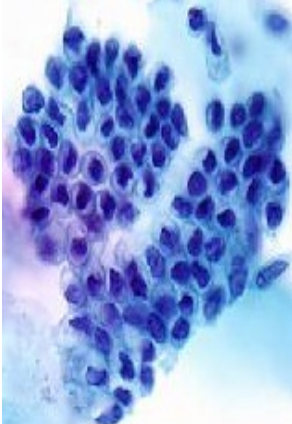

It can be concluded that Energy of normal sample is in the range of 0.2-0.4 and sum is in the range of 0.8-1.7. For Abnormal sample the energy is in the range of 0.08-2.0 and sum is in

the range of 0.3-0.8. Here sum indicates the summation of energy values for all the orientations.

6.2. To Check the Presence of HCG

The steps to check the presence of HCG is already explained above. Now the outputs are shown below (only two cases one for normal and another for abnormal sample is shown) –

Table 9. Snapshots of all the Output showing the Presence of Hyperchromasia

SL No	Input image	Cropped image	Output image showing all nuclei
1	 Source: Dr B.Baroo Cancer Institute		
2	 Source: University College of London		

Observation 1:

Discussion 1: (According to SL no 1 from above table)

Here number_of_nuclues = 6.

Hence, Hyperchromasia is absent.

Observation 2:

Discussion 2: (According to SL no 2 from above table)

Number_of_nuclues = 52

Hence, Hyperchromasia is present.

7. Conclusion

It can be concluded that texture analysis is one of the most important feature for abnormal and normal cervical cell classification. It can help in early detection of cervical cancer. Medical images have various limitations such as low quality, presence of noise and human error in interpretation [9]. 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. 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.

Acknowledgements

I would like to offer my 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. I would also like to thank Dr J. Mahanta, Director of Regional Medical research center, Dibrugarh; Dipankar Saikia, Executive of Infrasoftware Technology Ltd for his valuable instructions in many fields and Dr Deepjyoti Kalita for providing me all the medical books and help in understanding the medical terminologist used for this work. This project has the consent of all co-authors and authorities of the institute, where this study has been carried out and there exists no conflict of interest anywhere.

References

- [1] M. T. Sreedevi, B. S. Usha, S. Sandya, "Papsmear Image based Detection of Cervical Cancer", International Journal of Computer Applications (0975 – 8887), vol. 45, no. 20, (2012).
- [2] K. A. Dinshaw, S. S. Shastri, S. S. Patil, "Cancer Control Programme in India: Challenges for the New Millennium", Health Administrator, vol. XVII, no. 1, (2009), pp. 10-13.
- [3] G. Castellano, L. Bonilha, L. M. Li and F. Cendes, "Texture analysis of medical Images", Clinical Radiology, 59, (2004), pp. 1061-1069.
- [4] G. Saini, "Texture Analysis of Ct Scan Images," M.E final year thesis, Thapar University, Patiala, (2008), pp. 23-40.
- [5] University College of London, providing nano-zoomer, <http://www.eurocytology.eu/static/eurocytology/eng/cervical/LP1ContentGcont.html>, Accessed, (2012) October 25.
- [6] MDE lab, <http://labs.fme.aegean.gr/decision/downloads>, Accessed, (2013) January 2.
- [7] L. B. Mahanta and K. Bora, "Analysis of Malignant Cervical Cells Based on N/C Ratio Using Pap Smear Images", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, no. 11, ISSN: 2277 128X, (2012), pp. 341-346.
- [8] Mathwork Guide for all the matlab functions link: <http://mathwork.com/guide>, last visited on 19th Feb'13. Accessed, (2012) September.
- [9] L. B. Mahanta, D. C. Nath and C. K. 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, (2012), pp. 312-325.
- [10] M. E. Martinez, C. Vos, E. Santamaria, X. Jove and J. C. Socoro, "Non Linear Image Analysis for Fuzzy Classification of Breast Cancer", IEEE Proceedings of International Conference on Image Processing, vol. 2, (1996), pp. 943-946.

Authors



Dr Lipi B Mahanta

Assistant Professor - II

Central Computational and Numerical Studies

Institute of Advanced Study in Science and Technology

Paschim BoraGaon, Guwahati, Assam



Kangkana Bora
Assistant Professor - I
Regional Institute of Science and Technology
9th Mile, Guwahati, Assam