# TEXTURE ANALYSIS OF HISTOPATHOLOGICAL IMAGES TO IDENTIFY ANOMALOUS REGION

**K.Pratheesh Kannan***

**A.Ananthakumari****

**ABSTRACT**

The pathological image segmentation is important in cancer diagnosis and grading. In human body, tissues are characterized with the organization of their components. Cancer causes the changes in these organization. In order to diagnose the cancer disease, pathologist visually examine the changes in the tissue. This examination mainly relies on the visual interpretation. It may lead to considerable amount of observer variability. Hence, they may or may not identify the abnormal tissue. To avoid this problem robust algorithms are introduced for segmentation. Graph Run Length Method (GRLM), Gray Level Co-occurrence Matrix (GLCM) provides efficient way to segment the abnormal tissue. To a pathological image color graph was automatically generated by using Graph Run Length Method (GRLM). Gray Level Co-occurrence Matrix (GLCM) provides texture features of pathological image. The graph provides the arrangement of cells and structure of cells in a tissue. Based on the arrangement of cells, structure of cells, GLCM based texture features we can segment the abnormal tissue efficiently.

**Keywords**

Cancer diagnosis, Color graph, Observer variability, Pathological image segmentation, Visual interpretation.

* Department of Information Technology, PSN College of Engineering and Technology, Tirunelveli.

** Lecturer, IT Department, PSN College of Engineering and Technology, Tirunelveli.

## 1. INTRODUCTION

The imaging system was most widely, used in medical fields. The imaging system captures the digital images of human body and the image was not automatically analyzed or decision making. Human experts make analysis or decision by visually examining the images. The quantitative measures are compared with the visual assessment of experts. The reason behind the calculation is observer variability [6], [7]. Cancer grading was characterized by degree of distortion and irregularities observed during tissue testing. Gleason grading was most commonly used method to characterize the distortion and irregularities [3]. Gleason grading is based on the shape and size of the tissue. An advantage of Gleason grading is it's strong correlation with the biological behavior of the affected tissue and their spreading. At the beginning the degree of distortion was lower (Low grade cancerous). In this stage we can moderately differentiate the tissue components. Then, the distortion level becomes higher (High grade cancerous). In this stage we can not differentiate the tissue components.

Pathological image segmentation considered the color texture segmentation method [8]. Pixel based approaches divided the sample image into different groups. This is mainly based on color histograms such as k-mean clustering, fuzzy clustering, watershed transformation and thresholding. To segment the abnormal area region growing algorithm[5], split and merge algorithm and watershed transformations. Graph based method [1] consider the image as a weighted graph, it considered the node as a pixels. Region based method group the same pixels into cluster.

## 2. TEXTURAL FEATURES ANALYSIS FOR CANCEROUS IMAGES

The proposed method is an effective method that can be applied on cancerous images. The block diagram for this approach is shown below (see Figure1).

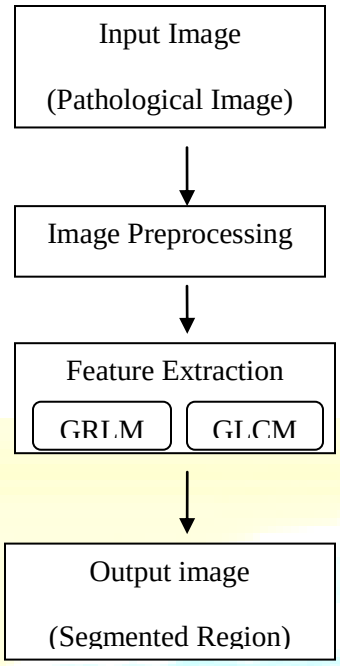



**Fig 1: Block Diagram**

## 2.1 Pathological Image

The pathological tissue is stained with hematoxylin and eosin. The input image is normally represented in RGB color model. A sample cancerous image is shown below (see Figure 2).

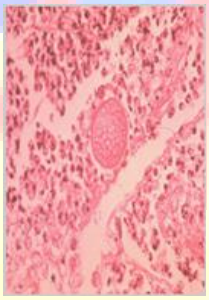
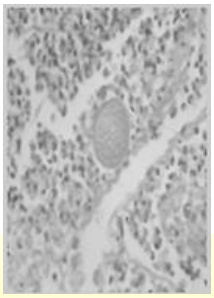


**Fig 2: Cancerous Image**

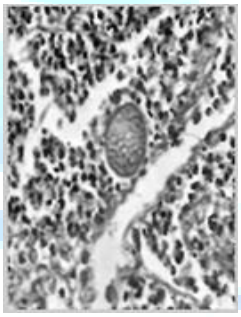
## 2.2 Image Preprocessing

The image preprocessing includes contrast enhancement and denoising. To improve the clarity of the pathological image contrast enhancement is done. During tissue staining, the

pathological image may contains unwanted noise. To remove the unwanted noise from cancerous image wiener filter was used. After Image Preprocessing we obtained the following images. (see Figure 3 (a) and (b) ).



**(a)**



**(b)**

**Fig 3: (a) Filtered Image (b) Contrast Enhancement Image**

## 2.3 Feature Extraction

The proposed approach uses, Graph Run Length Matrix (GRLM) along with Gray Level Co-occurrence Matrix (GLCM) for accurate segmentation.

### *2.3.1 Graph Run Length Matrix*

To extract the features color graph was generated on the cytological components and then, defines a run-length matrix using the edges of the generated graph. Finally extract the set of texture features from the graph run-length matrix.

#### *2.3.1.1 Color Graph Construction*

Color graphs for automated cancer diagnosis and grading generate the color graphs automatically [1]. In past, the structure of the tissue was quantified through the spatial distribution of cell nuclei. Here, accurate segmentation was not possible.To overcome, this problem the structure of the tissue is quantified through the spatial distribution of multiple components. To represent the color components circular primitives are used. The center point of the components are considered as centroids. To group the components K-mean clustering [4] is applied on the circular primitives. If the primitives are three, it's graph could consists of six different edges [1]. Based on the range of nearest primitives, the edge colors are calculated. The graph which is relevant to cancerous image is shown in figure 4.

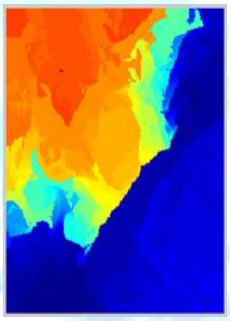


**Fig 4: Color Graph Image**

*2.3.1.2 Run Length Matrix Calculation*

In gray level images [2], [9], [10] run value was calculated on a set of consecutive pixels have the same color value. Similarly, in color images, a run value is defined as a set of consecutive primitives have the same color value.

*2.3.1.2 Texture Features for Graph Run Length Matrix*

*Short Run Emphasis*

The short path (or) run emphasis gives importance to shorter edge runs, which can be obtained by dividing the number of runs by the square of their lengths.

$$\text{Short Run Emphasis} = 1/n_r \sum_i \sum_j I(i,j) / j^2$$

*Long Run Emphasis*

The long path (or) run emphasis gives importance to longer graph edge runs, which can be obtained by multiplying the number of runs by the square of their lengths.

Long Run Emphasis=$1/n_r \sum_i \sum_j I(i,j) \times j^2$

*2.3.2 Gray Level Co-Occurrence Matrix*

GLCM creates a matrix with the directions and distances between neighbor pixels and then extract the meaningful statistical data from the matrix as a texture features.

*2.3.2.1 Texture Features for Gray Level Co-occurrence Matrix*

*Energy*

Energy is known as uniformity. It is a measure of homogeneity of an image. Homogeneity measures the similarity (or) difference between the pixels. It is expressed as,

Energy= $M_d (i,j)^2$

*Contrast*

Contrast is used to separate the image regions. Tumours are normally darker than normal cells. To detect these darker cells or abnormal tissue we must detect the density which is determined by this statistical data. The contrast value is high for cancerous tissue and low for normal tissue. It is expressed as

Contrast= $\sum_i \sum_j ( i-j)^2 M_d (i,j)$

*Homogeneity*

Homogeneity is to find out homogeneous property of each and every cell. If the cancer is in final stage homogeneous value is very high. It is expressed as,

Homogeneity= $\sum_i \sum_j M_d(i,j) / 1+ | i-j |$

The feature extracted image is shown in figure 5.

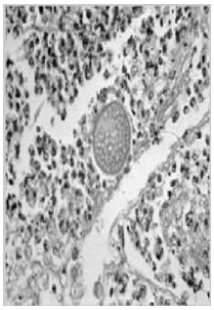


**Fig 5: Feature Extracted Image**

## 3. SEGMENTATION ALGORITHM

The proposed approach uses region growing algorithm for segmentation. In seed calculation step, seed regions are found by neighborhood relations [2]. The distance between every pair of adjacent primitives are computed and a pair is disconnected if the distance threshold is too high. Then, the small size connected components are eliminated and the remaining components are considered as initial seeds. In region growing step, the remaining primitives are iteratively assigned to the initial seed regions. In each iteration atleast one seed is connected with the adjacent components. Region growing continues until there were no unassigned primitives are left. Finally, we obtain the segmented image which is shown in figure 6.

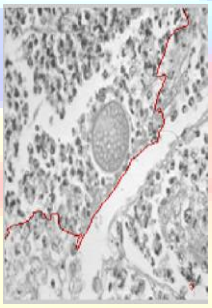


**Fig 6: Segmented Image**

## 4. EXPERIMENTAL RESULTS

In order to assess the feasibility result, the proposed feature extraction method was performed on various cancerous images. The extracted features will determine whether the region is normal or not. Using bench mark values of the cancer detection is to be charted.

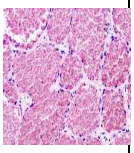
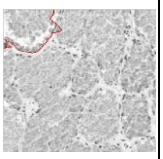
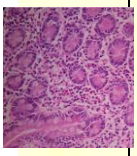
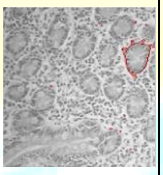
| Input Image | Features | Segmented Image |
|---|---|---|
|  | Energy=0.9996<br>Contrast=0.0088<br>Homogeneity=0.9998 |  |
|  | Energy=0.9998<br>Contrast=0.0034<br>Homogeneity=0.999 |  |

**Fig 7: Quantitative featutres and segmented regions**

This analysis clearly shows that, this best technique can be used to segment the abnormal tissue from normal tissue. In future, we can analyze more features so that an automated system can run and execute more quickly as well as accurately. The experimental results are shown in table.

### 4.1 Performance

In proposed system, the structure of the tissue was quantified through the spatial distribution of multiple components. Also, Gray Level Co-occurrence Matrix has been used along with Graph Run Length Matrix for feature extraction. But in an existing system, the structure of the tissue was quantified through the spatial distribution of cell nuclei. Also, it uses Graph Run Length Matrix for feature extraction.

The quantitative measures to assess the performance of the segmentation methods compared were sensitivity and specificity. In a ROC curve positive rate(sensitivity) is plotted on function of the false positive rate(100-specificity) for different cut-off points of a parameter. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold.

By comparing with existing algorithm the proposed method provides efficient segmentation. The following graph shows the performance of this approach.

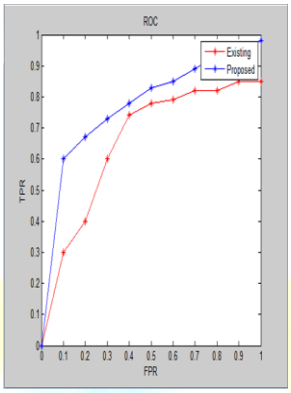



**Fig 8: Performance of the proposed system compared with existing system**

## 5. CONCLUSION

A new algorithm for the unsupervised segmentation of pathological images has been presented in this paper. The proposed algorithm was tested with simulated data. The result shows that, this method provides an efficient way in the segmentation of pathological images. In future, we can analyze more features so that an automated system can run and execute more quickly as well as accurately.

## 6. REFERENCES

[1] D. Altunbay, C. Cigir, C. Sokmensuer, and C. Gunduz-Demir, "Color graphs for automated cancer diagnosis and grading," *IEEE Trans.Biomed. Eng.*, vol. 57, no. 3, pp. 665–674, Mar. 2010.

[2] A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, and A. Murray, "Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa," *IEEE Trans. Inf. Technol. Biomed.*, vol. 2, no. 3, pp. 197–203, Sep. 1998.

[3] Claudia Wittke, Johannes Mayer, and Franz Schweiggert,"On the Classification of Prostate Carcinoma With Methods from Spatial Statistics"*IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY* IN BIOMEDICINE, VOL. 11, NO. 4, JULY 2007

[4] AkifBurakTosun, CenkSokmensuer, CigdemGunduz-Demir "Unsupervised Tissue Image Segmentation through Object-Oriented Texture",*2010 International Conference on PatternRecognition*

[5] Y. Shih*, Shouxian Cheng, " Automatic seeded region growing for color image segmentation", Computer Vision Laboratory, College of Computing Sciences, New Jersey Institute of Technology, Newark, NJ 07102, USA

[6] GDH Thomas, MF Dixon, NC Smeeton, NS Williams "Observer variation in the histological grading of rectal carcinoma"*J ClinPathol1983;36:385-391*

[7] CigdemDemir and B¨ulentYener ,"Automated cancer diagnosis based on histopathological images: a systematic survey", *TECHNICAL REPORT, RENSSELAER POLYTECHNIC INSTITUTE, DEPARTMENT OF COMPUTER SCIENCE, TR-05-09.*

[8] Tian Xia, Yizhou Yu University of Illinois at Urbana-Champaign and Jing HuaWayne State University "Automatic Detection of Malignant Prostatic Gland units in Cross-sectional Microscopic Images",*Proceedings of 2010 IEEE 17th International Conference on Image Processing September 26-29, 2010, Hong Kong*

[9] Wang xinlily,FritzAlbrtsen and Bent Foyn ,"Texture Features From Gray Level Gap Length Matrix",*MVA'94 IAPRWorkshop on Machine Vision Applications Dec. 13-1 5. 1994. Kawasaki*

[10] Horng-HaiLoh, Jia-GuuLeu "The Analysis of Natural Textures Using Run Length Features", *IEEE TRANSACTIONS ON INDUSTRIALELECTRONICS, VOL. 35, NO. **2,** MAY 1988*