International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 9, September- 2014

Generation of Colon Tissue using Structural and Statistical Methods

Rintu Mariya Thomas
PG Scholar, Communication Engineering
Amal Jyothi College of Engineering
Kanjirappally, India

K G Satheesh Kumar
HOD, ECE Department
Amal Jyothi College of Engineering
Kanjirappally, India

***Abstract*— Cell is the basic functional unit of an organism. The cells together form a tissue and tissues together form an organ. Cancer causes the cells in an area to change from its normal distribution. It affects the nearby cell in a multiplicative manner which is beyond the control of the cell generation, which changes the biological structure of human cell. The diagnosis depends upon the experience of the doctor. For proper diagnosis the use of computer is inevitable and currently lot of research is going on in this area. This potential of assisting pathologists in histopathological examination of tissues may lead to a considerable amount of subjectivity. These computer aided image analysis tools helps in reducing the subjectivity that provides quantitative information about tissues. A hybrid model technique is introduced here that use both statistical and structural methods. It is helpful in determining the malignancy level and early detection of cancers.**

Keywords*— statistical method, structural method, histopathological examination

I. INTRODUCTION

Digital pathology is the management of pathology information from virtual, digital microscopy. Compared to traditional pathology, slides can be analyzed using software is the major advantage. However, with the early detection of cancer increases the survival rates and the correct treatment options. Medical imaging methods provide effective diagnosis tools for screening, they may not be helpful in determining their malignancy level and early detection of cancers.

II. COLON CANCER

Colon is the first and the longest part of the large intestine which is formed of epithelial cells. It is one of the most common type of non-skin cancer in men and in women. It also called as colon cancer, colorectal cancer, bowel cancer, or rectal cancer.

Due to age increasing, living habits and genetic disorders most colorectal cancer occurs. Based on histopathological examination of biopsy tissue samples diagnosis and grading is related. By using the computational method which helps in decreasing the subjectivity level and provides quantitative measures. Rectal bleeding and anemia are symptoms associated with weight loss and changes in bowel habits. Certain factors that causes disease, includes age, personal history, family history, diet, exercise, smoking. Diagnosis of colorectal cancer is done via cancer biopsy during colonoscopy or sigmoidoscopy.

Fig. 1 and 2 shows an image of a colon tissue and microscopic image of a colon tissue section, in which it is stained in hematoxylin-and-eosin (H & E) technique. In this Fig 3, the components and the glandular structure of the colon tissue is indicated. An epithelial cell, marked with green circle, is formed of a nucleus and cytoplasm which represents dark purple and white colored areas in the image. Epithelial cells in the colon are lined up around a vacant region, called luminal area or lumen, forming the gland structure. Absorption of water and nutrients and secretion of mucus is performed in lumens. Lumens correspond to white colored regions. The other components in the tissue are stroma, which corresponds to pink colored regions. These components are responsible for holding the tissue together.

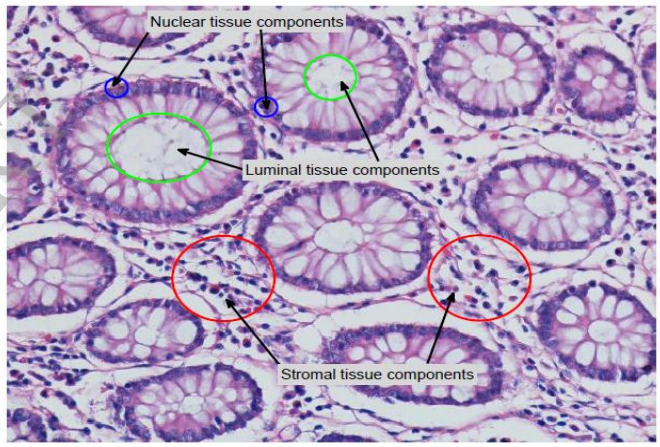

Fig. 1. Image of a colon tissue

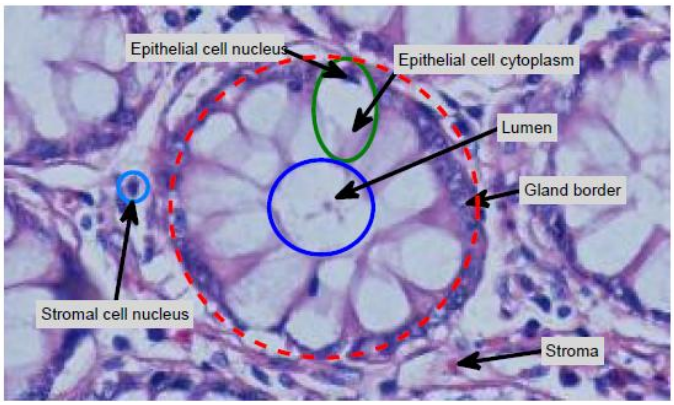

Fig. 2. Microscopic image of colon tissue section

C. C Agatay Bilgin, Cigdem Demir, [1] in their work, extracted nucleus and non nucleus pixels from the image by thresholding their intensities. Cigdem Gunduz-Demir, Melih Kandemir, Akif Burak Tosun, Cenk Sokmensuer, [2] in their work, made use of organizational properties to decompose the tissue image into a group of circular objects and non circular objects. In order to find the characterization, classify the pixels

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

into different classes (e.g., nucleus, non nucleus classes) and then by using these pixels information it forms gland regions. In another work of the same authors, based on their intensity values, in order to divide the pixels into nucleus and non nucleus a Bayesian classifer is used. Dogan Altunbay, Celal Cigir, Cenk Sokmensuer, and Cigdem Gunduz-Demir, [3] in their work says that there are four different approaches for tissue quantification. They are morphological, intensity-based, textural, and structural approaches.

III. METHODOLOGY

We introduce a new structural and statistical approach to generate a tissue image as tissue graph characterize the sub graphs with the distribution of their edges. Features are extracted on these graphs and use these for classify the tissue image. This method consists of segmenting the tissue images, generating the images as graphs and feature extraction from the graphs and finally classification. Each step is described in the following sections. Fig. 3 shows the overview of the proposed system.

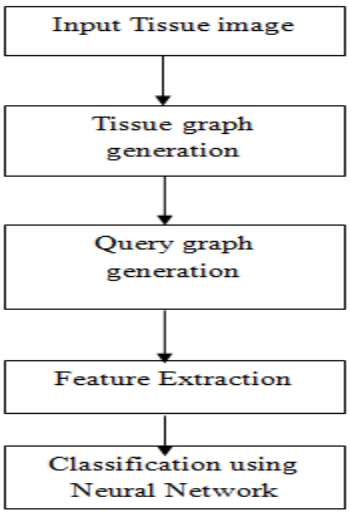

Fig. 3. Overview of the Proposed System

A. Tissue Graph Generation

The microscopic image of a colon tissue is stain in Hematoxylin and Eosin (H & E) technique. The tissue image is represented as graph, G = (V, E) in which V and E indicates the vertex and edge set of the graph. In this tissue graph generation, the main work is to locate the tissue components, classify them as the graph nodes and identifying the graph edges between theses nodes in the image. For this it decomposes a tissue image into a group of objects, which indicate different tissue components. It is very difficult to exactly locate the components. In order to define the objects the image is divided into a group of circular objects (nucleus and non nucleus objects). We separate H & E stain and threshold by using ostu's method. After that on each group of pixels, we trace a set of circular objects via Circle-fit algorithm. In this algorithm, we consider only the circles whose radii are greater than 9. The circular objects assigned as the graph nodes, then construct a tissue graph using Delaunay triangulation. Delaunay triangulation of a point set P in two dimensional space is defined as the set of triangles where each triangle conforms to the condition that there exist no point in the interior of the circle passing through the edge points px, py, pz of the triangle. It is used to build structural representations of various tissues for automated cancer diagnosis and grading.

B. Query graph generation

Query Graph is generated using Breadth First Search (BFS) Algorithm. At each level, node is expanded in a breadth wise direction. The elements are inserted in first-in first-out (FIFO) order. This algorithm is useful for small graphs, but for larger graphs it has many challenges. So the larger graph is divided into smaller graphs or sub graphs. From a given source vertex s, it finds all other vertices. This source vertex s, is represented as level 0 (L_0). At first, it visits all the vertices that are one edge distance away from s, named as level 1 (L_1). Then it visits all new vertices, which are two edge distance away from source vertex s, named as level 2 (L_2).By using the K–means clustering algorithm [8], pixels are divided into three classes (or clusters) based on their brightness.

C. Feature extraction

Gray-level co-occurrence matrix (GLCM) [6], [9] is used for the feature extraction. To reduce the complexity, 4 features are extracted from an image based on gray level intensities. Contrast, entropy, energy, correlation and homogeneity are the statistical parameters extracted. Contrast measures the amount of local variations present in an image and is difference moment of GLCM. It is the difference between the highest and the lowest values of a set of pixels. Energy measures the textural uniformity. So it is also called Uniformity or Angular second moment. It detects disorders in textures. Energy reaches a maximum value equal to one. Correlation measures the linear dependencies in the image. Homogeneity is also called as Inverse Difference Moment (IDM) and is defined as the quality or state of being homogeneous and also having a uniform structure. It achieves its largest value when most of the occurrences in GLCM are concentrated near the main diagonal.

D. Classification

The emergence of computers into day today life of human beings has helped us to save plenty of time. The classification problem is tough and many present algorithms are complex and slow. So here we go for Neural Network [10] for classification. Neural Network can be trained to solve problems which are difficult for human being. It classifies tissues as normal low grade and high grade cancerous colon tissues images. One of the most important applications is recognition.

IV. EXPERIMENTAL RESULTS

A. Data Set Preparation

This chapter describes the experimental methodology and provides the results of the proposed method. Images of normal, cancerous and non cancerous colon tissues are used as data set, which contains the tissues stained with hematoxylin

IJERT

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

and eosin technique. In this experiment, tissue samples used in the experiments were stained using H&E. 20x microscope objective lens are used for taking images and also another lens is placed at eye end. Each tissue image section data set has a thickness of 5 microns.

B. Results

In the proposed work the classification is done with the neural network classifier. It takes the input as an image and classifies whether it is normal, low grade or high grade cancerous colon tissue. For the correct classification firstly the neural network has to be trained with tissues which includes normal, low grade and high grade colon tissue. The input images of normal, low grade and high grade cancerous colon tissues are shown in Fig. 4, 6 and 8 respectively. Their corresponding classification results for a normal, low grade and a high grade tissue of these three images are shown in Fig. 5, 7 and 9 respectively

.

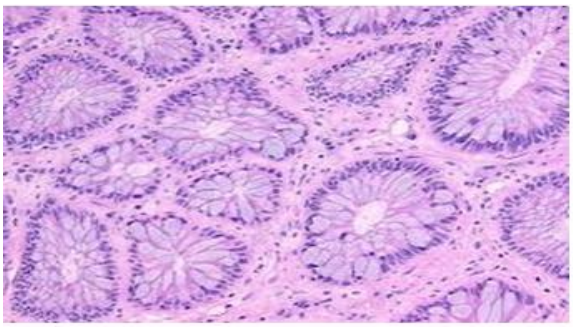

Fig. 4. Normal colon tissue

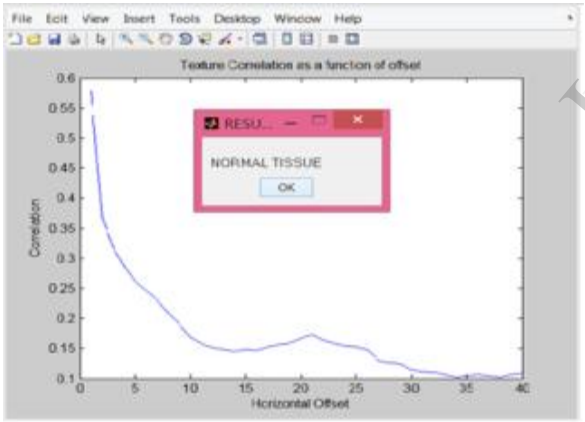

Fig. 5. Classification result for a normal tissue

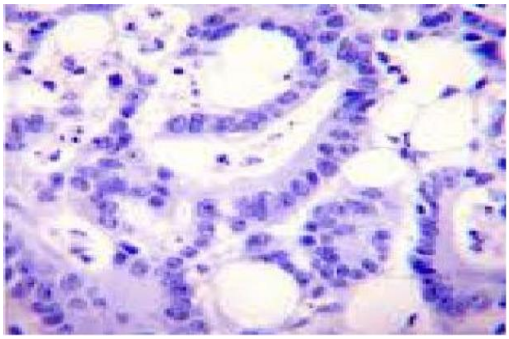

Fig. 6. High grade cancerous colon tissue

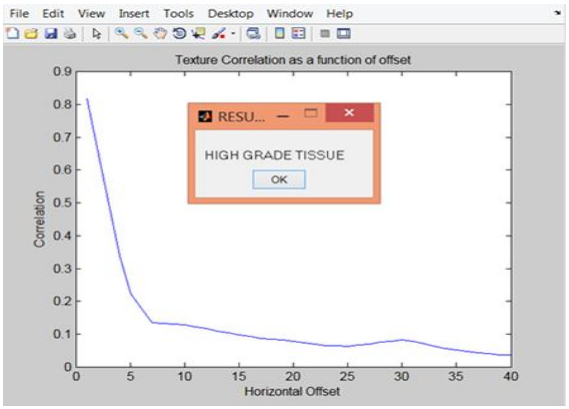

Fig. 7. Classification result for a high grade cancerous tissue

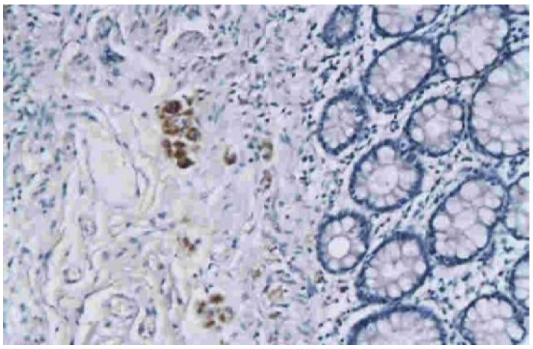

Fig. 8. Low grade cancerous colon tissue

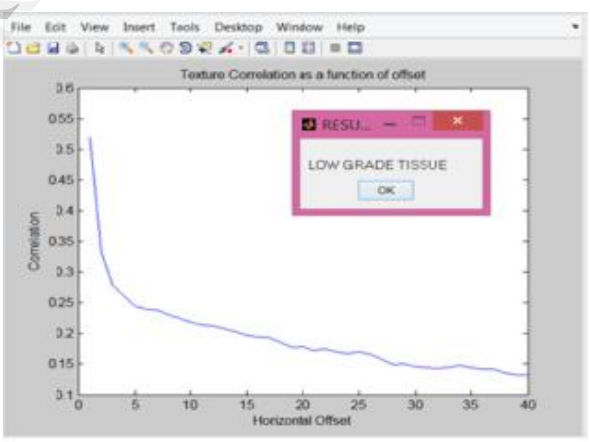

Fig.9. Classification Result for a Low Grade Cancerous Tissue

The proposed method also extracts statistical features. These approaches extract their features by identifying histological components of a tissue and quantifying their morphology and organization characteristics. The statistical features includes contrast, correlation, energy and homogenity. These features are extracted from a gray level co-occurrence matrix (GLCM). Fig. 10 shows the extracted features.

IJERT

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

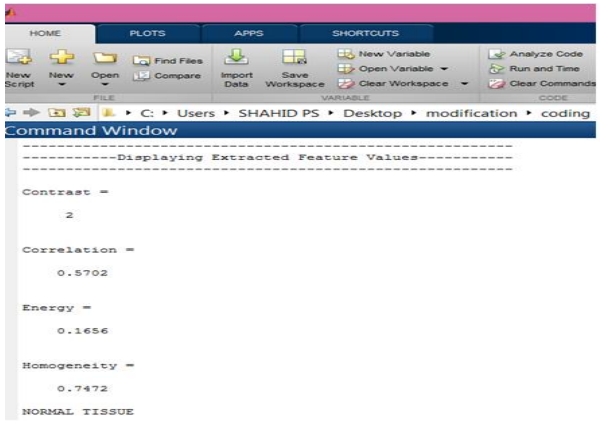

Fig.10. Extracted Features

For analysing the efficiency of the system we use accuracy, sensitivity and specificity. Accuracy is a measure of overall usefulness of the classification technique. Sensitivity and specificity, respectively, are used to calculate ability of a classifier to recognize patterns of positive and negative classes. They can be obtained using the following expressions:

Accuracy=TP+TN/TP+TN+FP+FN (5.1)

Sensitivity =TP/TP + FN (5.2)

Specificity =TN/TN + FP (5.3)

Where true positive (TP), true negative (TN) are the number of correctly classified positive and negative samples. False positive (FP) and false negative (FN) are incorrectly classified samples. Fig. 11 shows the results of sensitivity, specificity and accuracy values .

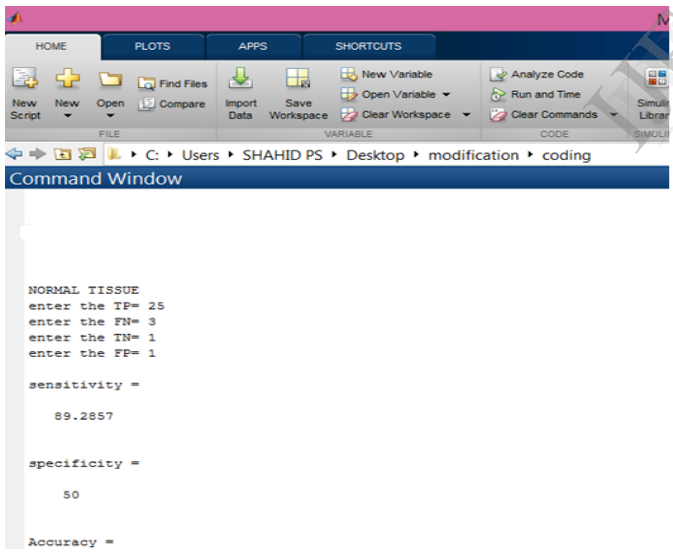

Fig. 11. Sensitivity, Specificity and Accuracy Values

V. CONCLUSION

The method is proposed to model spatial relationships of different tissue components. For that it defines the tissue images as objects to identify different tissue components. In order to find out the location of different tissue components, here use Delaunay triangulation for generating tissue graph. The purpose of graph generation is for checking the nodes. The proposed method represents an image as a tissue graph and a query graph is constructed by using BFS algorithm. These two graphs are compared by graph edit matrix and clustering is done by K-means algorithm. Classification is by using neural networks classifier and then use them for classifying tissues as normal low grade and high grade cancerous colon tissues images. It is an important tools for cancer diagnosis and grading are gaining importance in medicine and also helpful in determining the malignancy level and early detection of cancers.

REFERENCES

[1] C. C Agatay Bilgin, Cigdem Demir, Chandandeep Nagi, Bulent Yener, "Cell-Graph Mining for Breast Tissue Modelling and Classification", Pattern Recognition, vol. 38, pp. 1209-1223, Aug. 2005.

[2] Cigdem Gunduz-Demir, Melih Kandemir, Akif Burak Tosun, Cenk Sokmensuer, 'Automatic segmentation of colon glands using object-graphs", Cytometry, vol.64, pp. 101-109, Apr. 2005.

[3] Dogan Altunbay, Celal Cigir, Cenk Sokmensuer, and Cigdem Gunduz-Demir, "Color Graphs for Automated Cancer Diagnosis and Grading", IEEE transations on biomedical engineering, vol. 57, no. 3, pp. 665-686, Mar. 2010.

[4] H. Choi, A. C. Bovik, and K. R. Castleman, "A Resampling-Based Markovian Model for Automated Colon Cancer Diagnosis", IEEE transations on biomedical engineering, vol. 59, no. 1, pp. 281-299, Jan. 2012.

[5] Gulden Olgun, Cenk Sokmensuer, Cigdem Gunduz-Demir, "Graph Walk For Classification of Histopathological Images ", IEEE 10th International Symposium on Biomedical Imaging, April. 2013.

[6] Mari Partio, Bogdan Cramariuc, Moncef Gabbouj, and Ari Visa, "Rock texture retrieval using gray level co-occurrence matrix", Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland.

[7] F. Albregtsen, B. Nielsen, and H. E. Danielsen, "Adaptive gray level run length features from class distance matrices", Pattern Recognition, vol.3, no. 1, pp. 3746-3749, 2000.

[8] J. A. Hartigan, M. A. Wong, "A K-Means Clustering Algorithm ", Journal of the Royal Statistical Society. Series C (Applied Statistics), , vol.28, no. 1, pp. 100-108, 1979.

[9] Bino Sebastian V, A. Unnikrishnan and Kannan Balakrishnan,"Grey level co-occurrence matrices: Generalisation and some new features",International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), vol.2, no. 2, pp. 151-157, 2012.

[10] V.P. Plagianakosa, G.D. Magoulas, M.N. Vrahatis, "Distributed computing methodology for training neural networks in an image-guided diagnostic application ", Computer methods and programs in biomedicine, vol.8, no. 1, pp. 228-235, 2006.

IJERT

(This work is licensed under a Creative Commons Attribution 4.0 International License.)