International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE) Vol 3, Issue 5, May 2016

Colorectal Cancer Identification in CT Images based on Image Segmentation Techniques

[1] Sobiya Ambreen, [2] Narayana Swamy G
[1] PG Scholar, [2] Associate Professor
[1][2] Dept of ECE, Vidya Vikas Institute of Engineering and Technology, Mysuru
[1]sobiyaambreen92@gmail.com [2]nswamy1979@gmail.com.

***Abstract*— Colorectal cancer is a lethal disorder and a cancer, third most commonly found that will lead to death in both men and women. Computed Tomography (CT) imaging technique has become a primary method for detection and staging in those patients who are suffering from Colorectal cancer. In this paper CT images of the patients suffering from Colorectal cancer, who are in the beginning stages are used to detect the area and mean values of tumor area using Image segmentation techniques. This paper describes algorithms for preprocessing, clustering and post processing of CT images. The paper uses image enhancement techniques, clustering using adaptive k-means algorithm and it is implemented using MATLAB.**

***Keywords* – Colorectal Cancer, Computed Tomography, Adaptive K-Means Algorithm, Improved Histogram based Spatial Fuzzy C Mean (IHSFCM) Algorithm.**

I. INTRODUCTION

Colorectal cancer is a cancer due to the uncontrolled growth of cell in the colon or rectum. It is a cancer which is diagnosed as the third most common cancer in the world. Colorectal cancer has few alternative names namely, colon cancer, rectal cancer, bowel caner or colorectal adenocarcinoma. Most of the times, colorectal cancer appears because of the lifestyle and also due to increase in the age of both men and women. Hence, detection of cancer in the earlier stages is important. MRI scanning & CT scanning methods have key roles to play in the cancer assessment.

Image processing techniques are imparted to improve the contrast of the images and transform the obtained intensity levels into equivalent colors so that interpretation becomes easier. Any digital image processing application, basically has got few steps to be followed such as, preprocessing, segmentation and feature extraction etc. In the Preprocessing stage, median filtering method is used to get rid of the noise in an image, which folllows nonlinear digital filtering technique. The main aim of using median filtering method is to retain the features of the edges in the CT images while noise has to be removed. The contrast is increased along with the edges by using the technique sharpening of the image. A sudden transition from black to white will appear sharp, while slow or gradual transition from black to gray to white may sometimes look blurry. Hence sharpening methods must be used. Contrast limited adaptive histogram equalization can be used for the equalization of the contrast in the image.

Clustering techniques are used so that the differentiation of the cancer tissues is possible from other textures. Clustering methods are considered as the most important and unsupervised learning problem; since, every problem of this type deals with finding a structure in a given collection of unlabeled data. K-means clustering and Fuzzy C means clustering are used in this proposed work.

II. LITERATURE SURVEY

Colorectal cancer is a one of the primary disorder that causes death in the USA [1]. The colon and the rectum together form a cancerous malignant tumor or polyps are formed on the inner mucosal lining of the bowel. It can hamper the proper cell functioning. Fecal occult blood test (FOBT), colonoscopy, computed tomography colonography (CTC), and colon capsule endoscopy (CCE) are the methods used for diagnosis, and the treatment depends on the type and location of the cancer region. Colorectal carcinoma can affect variety types of cells that range from the adenomatous cells lining present in the colon to the WBC that protect it. Colorectal cancer is diverse, which can affect anyone. With innovations in technology, attempts to develop algorithms using Image processing are made to detect Colorectal Cancer in the earlier stages.

Abdominal computed tomography (CT) is very much valuable [2] to plan for a surgery for the removal of colon cancer since it demonstrates the regions where tumors are

extended and distant metastases. In CT, colon cancer appears as a discrete mass of soft-tissue, which narrows the lumen of colon. Wall thickening and luminal narrowing is also seen due to this cancer. Colonic malignancies like obstruction, perforation and fistula can also be identified using CT. In CT, any sort of extension of tumor region will appear as thickening and infiltration of pericolic fat outside the colon. Loss of fat planes that happen to be seen between the colon and adjacent organs can also be identified. The liver is the main organ that is affected from colorectal cancer. CT is very much critical in identification of recurrences, evaluation of anatomic relationships, documentation of post operative anatomy, and confirming the absence of new lesions during and prescribing the therapy later. Rapid advancement in the technology will improve the consistency & accuracy of CT.

Automatic colon segmentation carried out using CT images is one of the crucial step [3], which also includes the computer-aided detection (CAD) of colonpolyps and a virtual 3D transformation through of the colon. But, the accurate segmentation is very difficult because of the presence of the organs like lungs, stomach and small intestines, which are filled with air and the thinning of the colon area. Graphical inference method is proposed for the removal of extra-colonic components and to achieve very high quality segmentation. Each 3D air-filled object is represented as a set of 3D regions. A classifier is used, which is provided with region-level features to train the classifier. It separates the non-colon regions from colon regions. Experimental results show that the proposed method outperforms in detecting true colon regions, while a decrease in the extra-colonic components is also observed.

Multiple diagnostic modalities can be used [4] but it is expensive and time-consuming. Routine colorectal cancer (CRC) screening, detects CRC in the beginning stages and also reducing the incidence of CRC or the early CRC progression through polypectomy, and hence reducing the CRC mortality. Ultrasound (US), CT and MRI imaging modalities are used. FDG-PET and FDG-PET/CT, the nuclear medicine methods also contribute significantly in the cancer identification. Identification of liver metastases always remains as first choice using CT. Lung-metastases are identified by X-ray or chest CT. Improvements in CT, MRI and positron emission tomography (PET) have increased tha accuracy of the detection of occult lesions and provide better definition for the surgical anatomy.

Cancer staging is very much helpful [5] to choose the appropriate therapy. It is also helpful in studying the effects of different therapeutic methods and comparing their efficiencies. MX Category elimination is used when classification provides poor result. But the results of MX category show ambiguity (sometimes due to lack of information or uncertain in assigning specific category) both in application and in interpretation, it was abondoned. The accuracy of CT was to be assessed [6] with reconstructed coronal images in detection of the tumor status of colon cancer. The data of 154 patients with colon cancer were analyzed by using preoperative 64-slice multi-detected computed tomography (MDCT). Each colon tumor is staged using the RSROC's image classification, and the obtained results were compared with histopathology. The overall accuracy recorded for the tumor staging was found to be 91 of 154 (59.1%).

In order to provide an effective way for the detection of colon cancer in the beginning stages, digital images of colon polyps were investigated [7]. Improved Gradient Vector Flow (IGVF) was considered for appropriate segmentation. IGVF improves the ability of GVF snake model's to capture thin boundary from the cancer images. The segmented parts are characterized by features such as the shape of the polyp & its internal intensity distribution. They will become the input for the classifier. Support Vector Machine (SVM) classification method was used. The segmentation done manually and the segmentation obtained from algorithm are compared to test the accuracy. The proposed work produced very good results for segmentation of colonic polyps.

Recent improvements in the area of CT imaging have allowed an opportunity to provide a minimal invasive method with large bowel evaluation [8] with cost of the treatment sufficiently reduced and also reducing patient risk. A major requirement for this method is full distension of the lumen. The lumen should also be cleansed with proper separation of intestinal walls. The patients were provided with CT-WE (Water Enema) for bowel distension. CT-WE provided excellent results for the visualization of the colonic wall, along with providing good contrast between the wall, water-filled lumen, and the pericolic fat. Dual energy spectral CT acquisition cab be certainly a novel imaging technique, that generates monochromatic images with necessary material decomposition.

Cancer tissues obtained from in histopathology images show abnormal patterns [9], and hence the labeling of a histopathology image with cancerous regions is very important and carrying out corresponding image segmentation. Multiple Clustered Instance Learning (MCIL) was proposed to classify, segment and cluster the cancer cells in the colon histopathology images. MCIL method can simultaneously perform image-level classification, that is cancer image versus non-cancer image and pixel-level segmentation, that is cancer tissue versus non-cancer tissue along with patch-level clustering

which is nothing but cancer subclasses. Experimental results displayed the efficiency of MCIL in the analysis of colon cancers, which introduced an integrated learning framework for the purpose of classification and obtaining the cancer clusters using weakly supervised learning. The major advantage of MCIL is this method is evident over the state-of-the-art methods which can perform the individual tasks.

The accurate as well as reliable segmentation [10] of the colon is very important for 3D reconstruction, classification and automatic colonic polyp detection. An adaptive level is introduced to set a method for segmenting the colon when it is filled with air and opacified fluid in CT colonography. It was tested on 10 CT colongraphy datasets, containing approximate 4,500 slices. The obtained accuracy is 96.06% in comparison with manual segmentation..

III. DESIGN METHODOLOGY

The block diagram of the proposed work is shown in the above Figure 1. The method starts by reading a CT image in the JPEG form, followed by preprocessing. And the image is segmented using clustering technique. Finally the clustered sub image is selected and processed further. The proposed work uses Matlab 2014 to develop the program. 5 CT Images of the patients suffering from Colorectal cancer are collected from Radiologists.

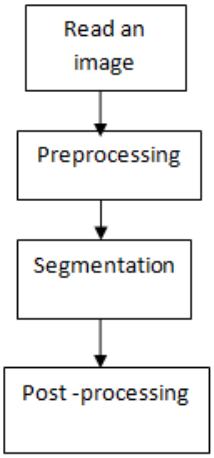

Figure 1: Block diagram of proposed method

Pre-processing:

The main objective of the Pre-processing of the image is enhancement. It is achieved by converting the given image into gray image. A two dimensional median filtering method is applied to get rid of unwanted noise. Median filtering is very effective compared to convolution especially when the main aim is to both reduce the noise as well preserving edges in the image are important. In Median filtering method all the image pixels are replaced with the median of image pixels, in a predefined (3-by-3) neighbourhood, given for every pixel. The filtered image is subjected to Unsharp masking for sharpening of the image. A blurred version of the image is considered and is subtracted from itself to obtain the result. In the final stage, contrast equalisation is carried out using Contrast Limited Adaptive Histogram Equalization (CLAHE) equalization. Figure 2 shows CLAHE distribution. The part of histogram which exceeds the clip limit as shown in the figure is not discarded. But instead, it is again distributed equally among all histogram bins. This pushes few bins over the limit of the clip again, and hence resulting in a very effective clip limit. The obtained clip limit is larger than the prescribed limit. The specific value of the limit is dependent on the image.

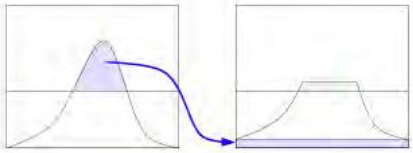

Figure 2: CLAHE distribution

Clustering algorithms:

K-means algorithm is one of the popular methods for Clustering algorithm. It has got four main disadvantages. It is a slow method with poor scaling on the time axis. It provides fixed number of clusters and hence sometimes it will be difficult to predict the value of K. The local optima computed will not be an optimal value and sometimes will give worse results. The different final clusters are obtained for different initial partitions. An Adaptive K-means clustering algorithm is used, to overcome the above mentioned disagnavtages.

Fuzzy C-means (FCM) algorithm is one of the most popular fuzzy methods used in image segmentation. It provides robust characteristics specifically for ambiguity and will retain more information compared to hard segmentation methods. The conventional FCM algorithm possesses a limitation, it will not incorporate that information regarding spatial context. It is also sensitive to imaging artifacts and noises. This limitation can be overcome by providing smoothing of the image just before segmentation. Conventional smoothing filtering methods can result in loss of important image details, specifically for image boundaries or edges. It is really a challenge and tough issue to have trade-off between smoothing and

clustering. Hence Improved Histogram based Spatial FCM (IHSFCM) clustering algorithm is used in the proposed work for the purpose of Image Segmentation which incorporates spatial neighborhood information present in the image, into the standard FCM clustering algorithm by using a priori probability. This feature can overcome the existing limitations of FCM. FCM algorithm which is based on histogram quickly converges, since it is going to cluster only the histogram and not the whole image. The main aim of the IHSFCM is to make use of the spatial information to take a decision on the class of a pixel for a given image.

Post Processing:

The Clustered image obtained close to the required data further processed. From the clustered image, the objects are selected by adding coordinates. The coordinates are selected using a mouse pointer. The binary image having the objects is returned. The boundaries are drawn for the objects selected. The properties like area and mean values are determined for the selected objects using region properties.

IV. IMPLEMENTATION AND RESULT

The implementation is done using Matlab 2014. The following images show the obtained results till date.

Figure 3 shows the CT image of a patient suffering from Colorectal cancer, which is considered as the input for the algorithm.

Figure 3: Input CT Image

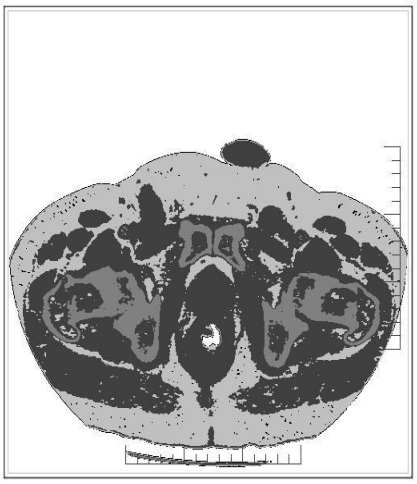

Figure 4: Histogram Segmented Image

Figure 5: Output of Adaptive K-mean Clustering algorithm

Figure 4 shows the Histogram Segmented Image, obtained after implementation of Improved Histogram based Spatial Fuzzy C Mean (IHSFCM) clustering algorithm Image Segmentation method.

Figure 5 shows the final result obtained after the implementation of Adaptive K-mean clustering algorithm.

V. CONCLUSION

The proposed algorithm provides a simple and efficient way for the diagnosis of Colorectal cancer in the beginning stages, by using Image Processing techniques. The algorithm is implemented on CT images which is not much expensive compared to MR Imaging. The proposed algorithm gave very good results for segmentation process and hence reducing the efforts of the Radiologist. The

algorithm can be enhanced in future by using advanced technologies, to develop an automatic segmentation algorithm as well as a module, which will further reduce the efforts of Radiologist and helps him for faster and much effective diagnosis of the Colorectal cancer in the beginning stages, which is difficult to carry out manually.

REFERENCES

[1] "*Colon Cancer and Gastrointestinal Imaging: A Case Report",*2013.

[2] Karen M. Horton, MD, Ross A. Abrams, MD, Elliot K. Fishman, MD, *"Spiral CT of Colon Cancer:Imaging Features and Role in Management"* 2000.

[3] Xiaoyun Yang, Xujiong Ye, Member, IEEE and Greg Slabaugh, Senior Member, IEEE, "*Multi-Label Region Classification and Semantic Linking for Colon Segmentation in CT Colonography*" 2014.

[4] Valentin Ignatov, Nikola Kolev, Anton Tonev, Shteryu Shterev, Elitsa Encheva, Tanya Kirilova, Tsvetelina Teneva and Krassimir Ivanov, "*Diagnostic Modalities in Colorectal Cancer –Endoscopy, CT and PET Scanning, Magnetic Resonance Imaging (MRI), Endoluminal Ultrasound and Intraoperative Ultrasound"* 2014.

[5] Giacomo Puppa, Graeme Poston, Per Jess, Guy F Nash, Kenneth Coenegrachts, Axel Stang, "*Staging colorectal cancer with the TNM 7th: The presumption of innocence when applying the M category"* 2013.

[6] I-Ha Lao1 Yen-Jen Wang2 Chee-Wai Mak1 Wen-Sheng Tzeng1 Reng-Hong Wu1, "*Accuracy of Computed Tomography in Staging the Tumor Invasion Status of Colon Cancer"* 2012.

[7] G. Shantha Lakshmi,MPhil, Prof. B. Mohamed FaizeBasha, MCA., M.Phil., "*Colorectal Cancer Detection using Image Processing based on IGVF Model"* 2014.

[8] Hong-xia Gong, Ke-bei Zhang, Lian-Ming Wu, Brian F. Baigorri, Yan Yin, Xiaochuan Geng, Jian-Rong Xu, Jiong Zhu, "*Dual Energy Spectral CT Imaging for Colorectal Cancer Grading: A Preliminary Study"* 2016.

[9] Yan Xu, Jun-Yan Zhu, Eric Chang and Zhuowen Tu, "*Multiple Clustered Instance Learning for Histopathology Cancer Image Classification, Segmentation and Clustering".*

[10] Dongqing Chen, Hossam Abd El Munim and Aly A. Farag, "*Segmentation of Colon in Tagged CT Colonography Using Adaptive Level Sets Method".*