Colorectal Cancer Detection using Image Processing based on IGVF Model

G. Shantha Lakshmi,MPhil,[1] Prof. B. Mohamed FaizeBasha, MCA., M.Phil.,[2]

[1]*(Department of Computer Science, Jamal Mohamed College, India.)*
[2]*(Department of Computer Science, Jamal Mohamed College, India.)*

Abstract: *Colorectal cancer is the third most common cancer in both men and women. Computed Tomography (CT) Colonography is a valid technique for detecting and screening colorectal cancers. In order to accomplish an effective way to identify colon cancer at an early stage, digital images of colon polyps have been investigated. In this paper, an improved gradient vector flow (IGVF) is considered as an essential method to segment an image appropriately. This new algorithm, IGVF can improve GVF snake model's ability to capture thin boundary indentation like the boundary of cancer images. The segmented candidates are typically characterized by features describing like the polyp shape and its internal intensity distribution. Such features will serve as input for the classification system. Classification is performed by Support Vector Machine (SVM).Finally, measure the overlap between the manual and the algorithm segmentation to test the accuracy of frame work. It is expected to obtain greater accuracy and to produce good segmentation results of colonic polyps.*
Keywords: *Colorectal Cancer, CT Colonography, IGVF, SVM*

I. Introduction

Colorectal cancer is cancer that starts in the colon or the rectum inside the large intestine. These cancers can also be referred to separately as colon cancer or rectal cancer, depending on where they start. It is the second leading cause of death among cancers and is the third most common form of cancer in the United States.
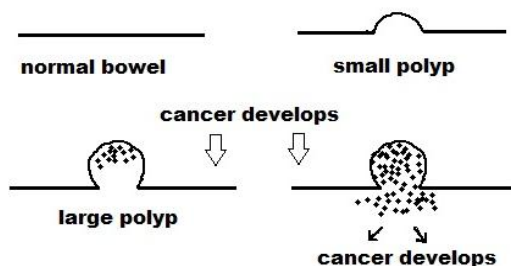

Figure 1 shows the Levels of cancer development

Given the fact that colorectal cancer is a largely preventable disease through routine detection and removal of adenomatous polyps, colon cancer prevention has now moved to the forefront. Levels of colon cancer development from the normal bowel are shown in figure 1.Most colorectal cancers begin as polyps. As polyps enlarge, they are more likely to develop into a cancer as shown in figure 2, which has the ability to disseminate through the body. The most important colorectal polyp is the adenoma, a small benign tumor growing to about 2 cm in size. Colonic adenomas are common and in the majority of patients there is no side effect on health. They are more common with increasing age. Appropriate evidence are shown that colonic adenomas are the early stage of colorectal cancer, although only a very small percentage of adenomas turn into malignant tumor and it may take nearly five to 15 years to change. The larger the polyp, the greater the probability that the polyp will have undergone malignant change and contain cancer (adenocarcinoma). Polyps range & its levels are shown in table 1. Therefore, polyp size is considered to be one of the most important factors in distinguishing benign polyps from cancerous ones.

Table 1 shows thepolyps range & its levels

Range	Levels
0cm	Normal
1cm-2cm	Small polyp
3cm-5cm	Large polyp
Above 5cm	Cancer develops

After detection and classification, accurate polyp segmentation could provide an easy way to measure polyp size, enhancing and improving the detection of significant lesion. Computer tomographic colonography is a minimally invasive technique and rapidly evolving diagnostic tool for the location, detection and identification of benign polyps on the colon wall on the early stage before their malignant transformation. Several methods were proposed for colonic polyp segmentation and detection.
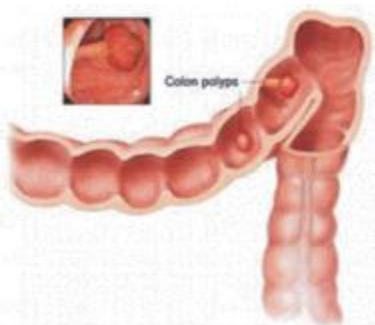
Figure 2 shows thepolyps situated on the colon wall

II.   Related Work
        Computed tomographic colonography (CTC) orvirtual colonoscopy is an evolving method fordetecting colon polyps which records the 3D image ofthe patient's abdomen [1]. About 95% of all cases of colon cancers arise from adenomatous polyps are initially benign, but at later stage they turn into malignant. The malignancy is normally referred with the size of polyps. The size of adetected polyp is an important criterion for further diagnosing and decision making [2]. Yoshida and N¨appi[3] firstly computed the geometric features to characterize polyps, folds and colonic walls at each voxel in the extracted colon. Then, the polyp candidates were detected by locating the voxels with low curvedness values and high shape index using hysteresis thresholding. Finally, they applied fuzzy clustering to the candidates to segment polyp. However, their algorithm may fail in obtaining the polyps with big masses and irregular shapes and voxels in the polyp region with different curvedness and shape index. Yao et al. [4] presented an automatic method to segment colonic polyps, which were based on a combination of knowledge-based intensity adjustment, fuzzy c-mean (FCM) clustering and deformable model. K. Gayathri Devi et al. [5] proposed CAD system, segmentation of the colon is performed by otsu's method of thresholding and clustered by k-means for extraction of the candidates. Classification is performed by SVM (Support Vector Machine) and ANN (Artificial Neural Network) from the training and testing images of candidate extraction.

III.   Methodology
        The overview of proposed work is shown in figure 3. The detection of polyps can be generally decomposed into four stages.The four phases are segmentation, candidate extraction, classification and overlap measurement. In the proposed system, segmentation of the colon is performed by an IGVF method. Classification is performed by SVM (Support Vector Machine) from the training and testing images of candidate extraction. The accuracy of the system is tested.
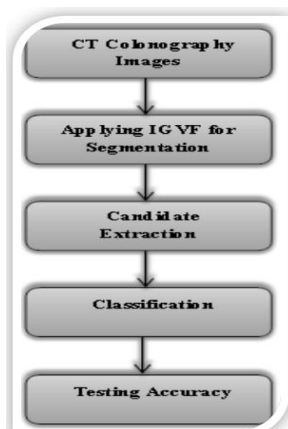
Figure 3 shows the overview of Proposed Work

Improved Gradient Vector Flow (IGVF)

This method is currently most efficient in the tracking cancer cells [9].The reason for using IGVF instead of traditional snake is that it cannot capture boundaries like "U" and "Ω" completely because of the counteraction of some external forces and the influence of the local minimum external forces. An improved algorithm is put forward by keeping the original GVF force and adding a new external force to solve the problems in GVF. The force balance equation of GVF snake model is:

$$F_{int} + F_{ext} = 0$$

In this improved GVFs, a new external force F_{ext} is presented and the force balance equation of new snake model is:

$$F_{int} + F_{ext1} + F_{ext2} = 0$$

Where F_{ext1} corresponds with F_{ext}.The additional force F_{ext2} is adaptive and it can confirm its size and direction timely to improve the performance of snake based on the position and force field of the control points. When the GVF snake energy is minimal, the result is not considered as final result. The estimation and calculation should be given according to the established standard. If the GVF snake is not final result after the estimation, the additional force F_{ext2} will work and will force the snake to the true boundary. And then the result can be considered as the final one. Or else the GVF snake result is the final result and the size of F_{ext2} is zero. That is to say, whether the additional force F_{ext2} will work is decided by the established standard and the force field of the control points. While the GVF snake is not final result, then the additional force F_{ext2} is usedto find the true boundary. At first, determine whether there are continuous points which have small gradient. Select two ends of the continuous contour points and connect the two points as a line. There are two vertical directions of the line. One vertical direction is where the snake comes from and the other vertical direction is the direction of the additional force F_{ext2}. The angle between abscissa and the direction of the additional force F is set to Θ. After the additional force F_{ext2} is added, the final force is marked as F_{new}. Here, the computational method of final force is not only one. So long as the new force can guide the snake to overcome the obstacles region, the computational method should be practicable. For example, as shown in figure 4, the angle value of F_{new} is set to Θ. Original GVF force is marked as F_{GVF}. And F_{GVF} is equal to F_{ext1}.

The calculating formula is:

$$F_{ext2} = F_{ext1} - F_{GVF} + F_{new}$$

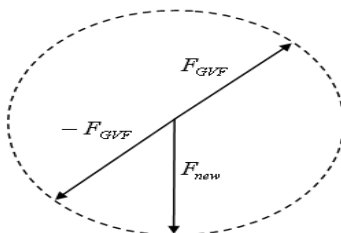The magnitude of F_{new} and F_{GVF} is equal:

$$|F_{new}| = |F_{GVF}|$$

Figure 4 shows the idea of new external field

When judging whether a control point is true or false, a specified gradient threshold is marked as T_1. When judging whether the new snake should stop approaching, another specified gradient threshold is marked as T_2. In fact, T_2 is used to judge if the snake reach true edge point. So, we can set the two thresholds T_1 and T_2 to be an equal value.

Candidate Extraction

Candidate extraction is used to identify and separate desired portions or shapes of an image. Colon lesion can be classified based on the parameters such as size, diameter of the growth, or according to their shapes, into common, fixed or flat. [10] Feature extraction phase the focus is on the extraction of three different parameters surface intensity characteristics, volumetric and surface shape characteristics and texture characteristics. The internal intensity of the candidates has been found to be a best feature to reduce a large number of false detections. The austral fold in the colon may be misinterpreted as polyp or growth in the identification of cancer. Texture is a powerful parameter that helps in the retrieval process. Texture on its own does not have the capability of finding similar images, but it can be used to classify textured images from non-textured ones and then be combined with another visual attribute like color to make the retrieval more effective.

Classification and Testing Accuracy

Classification of CT image data aims at discriminating multiple objects from each other within the image.Classification will be executed on the base of spectral or spectrally defined features, such as texture, volume and intensity in the feature space. It can be said that classification divides the feature space into several classes based on a decision rule. Candidate uniquely identified with the above associated features are then fed as input to a classifier for final evaluation with the help of training set or test set. The classification is performed by using Support Vector Machine (SVM) classification method. The numerical results were obtained by classifying with SVM using cost sensitive learning after normalizing the data. The goal of this classification is to separate normal images from abnormal images using three features which have been extracted in the feature extraction phase. The Segmentation accuracy S_A is calculated by:

$$S_A = \frac{\text{No. of correctly classified pixels}}{\text{Total number of pixels}} * 100$$

The accuracy of segmentation result by qualitatively computing the overlap between manual segmentation and the algorithm segmentation. The overlap for GVF and IGVF are shown in table 2 and its comparison is shown in figure 5. It can be computed by:

$$\Phi = \frac{2||S_c \cap S_m||}{||S_c|| + ||S_m||} \times 100\%$$

Where,

- ϕ is the overlap measurement
- S_c is the algorithm segmentation
- S_m is the manual segmentation
- $|| \cdot ||$ is the total number of pixels in a segmentation.

Table 2 shows the overlap for GVF & IGVF

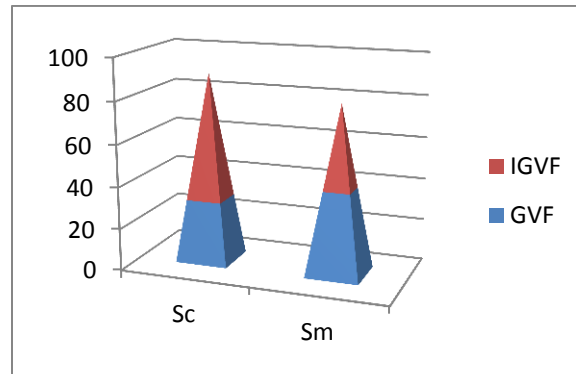| Model	Polyp 1	Poly2
S_c	40	45
S_m	30	20
Overlap for GVF(ϕ)	80.25%	85.17%
S_c	~60	~50
S_m	~30	~20
Overlap for IGVF(ϕ)	~92.00%	~90.25%

Figure 5 shows the comparison of GVF & IGVG

IV.   Conclusion

Colonic polyp segmentation is performed by IGVF with great false positive reduction. The variations in the segmentation and classification have been really effective in identifying the exact polyps. In earlier work, they attained the quantitative experiment results for the average overlap between the algorithm segmentation and manual segmentation is 85.17% ± 3.67%. Our preliminary results demonstrate that the proposed approach can produce good segmentation results of colonic polyps. It will obtain better accuracy in the overlap. Thus, this paper helps the radiologists to detect the exact polyp by distinguishing it (polyps) from the folds and false positives.

References

[1]   Abraham, H. Dachman and Hiro Yoshida, 2003.Virtual colonoscopy: past, present,and future.Elsevier Science (USA) RadiolClinnam 41, pp: 377–393. DOI:10.1016/S0033-8389(02)00124-0.
[2]   John, H. and Bond, 2000. Polyp Guideline: Diagnosis, Treatment and Surveillance for Patients with Colorectal Polyps, The American Journal Of Gastroenterology, 95: 11. Published by Elsevier Science, pp: 3053-3063.
[3]   Yoshida, H., N¨appi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. IEEE Transactions on Medical Imaging 20 (2001) 1261–1274.
[4]   Yao, J., Miller, M., Summers, R.M.: Automatic segmentation and detection of colonic polyps in CT colonography based on knowledge-guided deformable models. Volume 5031. (2003) 370–380.
[5]   Gayathri Devi. K., Radhakrishnan, R., Kumar Rajamani.: Computer Aided Diagnosis Scheme for Polyp Detection in CT Colonography Using K-Means Clustering and SVM. World Journal of Medical Sciences 9 (4): 273-281, 2013 ISSN 1817-3055.
[6]   Cohen, L.D., 1991. Note: On active contour models and balloons. CVGIP- Image understan.,53(2): 211-218.
[7]   Eviatar, H. and R. Soorjai, 1996. A fast, simple active contour algorithm for biomedical images pattern.Recogn.Lett., 55: 969-974.
[8]   Hou, Z. and C. Han, 2005. Force field analysis snake: An improved parametric active contour model. PatternRecogn.Lett., 26: 513-526.
[9]   Jinyong Cheng, Xiaoyun Sun.: Medical Image Segmentation with Improved Gradient Vector Flow Research Journal of Applied Sciences. 3951-3957, 2012 ISSN: 2040-7467
[10]   Farhan, Riaz, Mario   DinisRibeiro and Miguel Tavares Coimbra. A Review of Current Computer Aided Diagnosis Systems for Polyp Detection in Virtual Colonoscopy.