



Polyp Localization in Colonoscopy Images Using Vessel Density

Refika Sultan DOĞAN

Electrical and Computer Engineering Department
Abdullah Gül University
Kayseri, Türkiye
refikasultan.dogan@agu.edu.tr

Bülent YILMAZ

Electrical-Electronics Engineering Department
Abdullah Gül University
Kayseri, Türkiye
bulent.yilmaz@agu.edu.tr

Abstract—In this paper, we present a new approach for polyp localization in colonoscopy images. This approach is based on the determination of the polyp location using the vessel density in colon images. Primarily, we used pre-processing procedures on the colon images, and then blood vessel extraction techniques were employed. Later, segmentation of the vessel boundaries was performed. With the help of vessel boundaries we calculated the vessel density, and used this for the localization of the polyps. We tested the success of this approach using a publicly available image set (CVC-ClinicDB database). This database consisted of 612 images from 29 different polyps. This approach succeeds in correct detection of 24 out of 29 different polyps.

Keywords — vascularization; blood vessel; polyp localization; colonoscopy; image processing.

I. INTRODUCTION

Colorectal cancer is one of the leading causes of cancer death worldwide [1]. Colorectal cancer is the second most common cause of women and the third most common cause of men [2]. Colon cancer is the cancer of large intestine, the last part of the digestive system. Most colon cancer cases start with adenomatous polyps called small and benign cells group [3]. These polyps are divided into two main classes: hyperplasia, which are benign polyps, and adenomas are polyps that are strongly associated with malignancy. These growths, called polyps or adenomas, can be removed safely before they have progressed into malignant tumors. By searching for and removing polyps at an early stage, physicians hope to virtually eliminate this painful and deadly disease. Indeed, the death rate from colon cancer has dropped significantly over the past 15 years, and early detection and treatment is thought to play a major role in this decline [4].

Colonoscopy is known as the gold standard for detecting colon cancer or its pioneer lesions, colon polyps [5]. Colonoscopy allows examination of the colon via a light source and a camera, also allows the removal of the polyps with the aid of the catheter. However, colonoscopy is a practitioner-dependent procedure, so the attention deficit or fatigue of the practitioner sometimes causes the polyp to be missed. Polyps that are missed can turn into a cancerous structure and many years later the disease can be diagnosed as an advanced stage of cancer. In such cases the survival rate is lower than 10% [6]. Therefore, current studies aim to decrease polyp miss rates by

computer-aided polyp detection [7]. In recent years, a newly developed approach, wireless capsule endoscopy, is a method developed for human-independent computer-aided screening. In this method, the camera is placed in a vitamin-sized capsule. As this capsule travels along the digestive tract, thousands of images are recorded by the recorder connected to the waist of the patient. Although this procedure seems more advantageous for the patient, it is not very efficient in terms of examining the images. Because from one patient about 60,000 images are obtained in one procedure. Examination and processing of these data is not yet a preferred method due to its long assessment duration [3].

Different approaches have been proposed in the literature for computer-aided detection of colon cancer and polyps in endoscopic images. Various features have been used by these methods that analyze the shape, texture, and color of the colon. Neural networks (NN) and support vector machines (SVM) are the most common techniques used in classification [8]. Hwang et al. suggest watershed-based image segmentation and ellipse detection algorithm. New techniques have been proposed for separating non-polyp regions by matching the curvature orientation, the curvature, the edge distance and the density [9]. Alexandre et al. have provided a way to divide each view into smaller 40x40 pixels sub-images. The properties used are the RGB-values and the coordinates of each pixel [10]. In addition to these studies, Prasath et al. have shown that polyps can be isolated on a tissue using extraction of vascular characteristics in colonoscopy images [11].

In this study, we used the vascularization density feature on the tissue surface for the localization of the polyp site on colonoscopy images. In some approaches used in vascularization detection, we were inspired by the approach used in the vessel segmentation of images of retinal blood vessels [12]. In addition, we quantified the vessel density in the tissues. In this approach, we used image processing techniques to extract blood vessels of colon images. Firstly, we applied pre-processing on colon images, and then extracted blood vessels. The aim of the study is to determine the location of the polyps using vessel density.

II. METHODS

A. Image Database

We obtained 612 images with 29 different polyps from CVC ClinicDB [5]. This is a database of frames extracted from colonoscopy videos. These images include various types of polyps. In addition to the images, it provides the ground truth information for the location of polyps. This ground truth consists of a mask corresponding to the region covered by the polyp in the image. The database has been formed from 25 different studies in which at least one sequence containing a polyp has been extracted. This database is composed of images from 29 different sequences containing a polyp. These images have 288×384 size. Each image used in the study contains polyp, with different shapes, sizes, and number of polyps. Finally, a set of images were selected from each sequence, paying particular attention in showing several points of view of the polyp.

B. Pre-processing

Pre-processing is usually performed to remove noise from the background and enhance the image. There are different image enhancement approaches. Three main approaches suited for contrast enhancement in MATLAB: (1) image adjustment (imadjust), (2) histogram equalization (histeq), and (3) adaptive histogram equalization (adapthisteq). Adapthisteq function applies contrast-limited adaptive histogram equalization. Unlike histeq function, it takes small data regions (tiles) into account rather than the entire image. Each tile's contrast is enhanced so that the histogram of each output region roughly matches the specified histogram (uniform distribution by default). The contrast enhancement can be limited in order to refrain amplifying the noise which might be available in the image [13]. Some image processing techniques were used in this approach for the enhancement of colon tissue vessels. The Figure 3a shows an original image. Firstly, we converted into a size of 500 x 700 using bicubic interpolation to better analyze each of the images that we employed in the study before pre-processing.

We used 'rgb2hsv' function for the first step of pre-processing. HSV is a color space like RGB. It has three channels, instead of R, G, and B, it has H, S, and V, and the conversion formula is given in equation (1);

$$H = \cos^{-1} \left\{ \frac{1/2 [(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\}$$

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)] \quad (1)$$

$$V = \frac{(R + G + B)}{3}$$

Here R, G, and B represent red, green, blue and H, S and V represent hue, saturation, and value, respectively. Later, we used contrast-limited adaptive histogram equalization (CLAHE) only for V channel. After that, we recombined the new V channel with the original H and S channels. Subsequently, we converted this new HSV color space into

RGB color space back using 'hsv2rgb' function, as it is shown in Figure 1b. We extracted the green channel from this color image (RGB image), because the green channel has higher density when compared to red and blue channels.

C. Blood Vessels

We used the complementary function, which is a digital equivalent of obtaining a photographic negativity. This is especially useful for increasing white or gray detail embedded in a pre-dominantly dark area [14]. The negative of an image can be obtained by the 'imcomplement' function in MATLAB.

We then used histogram equalization function for enhancing the complementary image. Equation (2) defines the histogram equalization. The intensity levels are the continuous quantities normalized to the interval [0, 1] and p denotes the probability density function of intensity levels in an image used to distinguish between the functions of the input and output images. To obtain the output (processed) density levels, the following transformation is performed at the input levels,

$$s = \mathcal{T}(r) = \int_0^r p_r(\omega) d\omega \quad (2)$$

where ω is a dummy variable of an integration. The values of pixels, before and after processing, will be denoted by r and s, respectively. Here \mathcal{T} is a transformation that maps a pixel value r into a pixel value s [14].

We then used the morphological structure element called a morphological open function. There are two fundamental operations of gray-scale morphology, erosion and dilation, are explained in equations (3) and (4).

$$I_{dilated}(i, j) = \max_{f(n, m)=true} I(i + n, j + m) \quad (3)$$

$$I_{eroded}(i, j) = \min_{f(n, m)=true} I(i + n, j + m) \quad (4)$$

I represents the image and f(n,m) is defined as the digital image, and i, j, n, and m denotes pixel values.

$$A \circ B = (A \ominus B) \oplus B \quad (5)$$

$A \circ B$ is morphological opening, \ominus is erosion and \oplus is dilation in equation (5). Later, we used 2D median filter. After that, we used 'imadjust' function in MATLAB. This function maps the intensity values in gray-scale image, and the user determines the new intensity values of the images.

D. Detection of Boundaries Using Segmentation and Quantification of Vessel Density

We used the threshold function for extracting the blood vessels of the colon. We converted images into binary images by thresholding. This function produces binary images from indexed, intensity, or RGB images. To do this, it converts the input image to gray scale format, and then converts this gray scale image to binary by thresholding. The output binary image has values of 1 (white) for all pixels in the input image with luminance greater than the threshold level and 0 (black) for all other pixels.

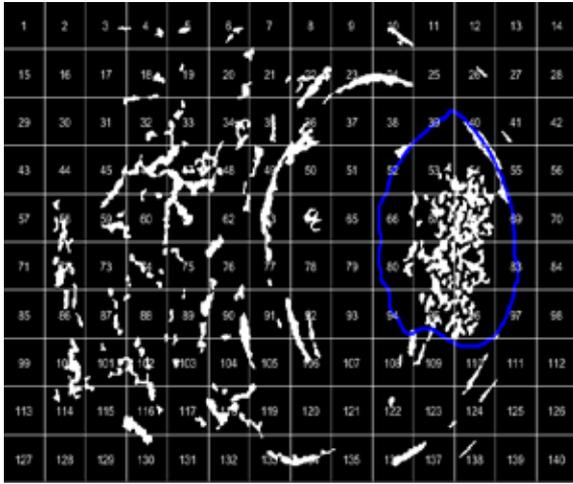


Figure 1. An example of subimages and the segmented vessels.

We applied area opening for very small and unwanted objects on the segmented image, and used quantification to determine the polyp location using the vessel images obtained from the segmentation section. While white parts represent vessels, black parts represent background of segmented image. The segmented images have 500x700 sizes, and we divided these images into subimages called tiles as 50x50 matrices. When we divided subimages of 50x50 sizes of 500x700 size images, we obtained new 14x10 sub-images as shown in Figure 1. The subdivision process generates a total of 140 sub-images for each original image. Then, we quantified for each subimage by computing the rate value showing the vessel density like as in equation (6). In addition to quantification, we visualized these rates for each image as 500x700 sizes images (Figure 2).

$$rate = \frac{\text{number of white pixels in the subimage}}{\text{number of subimage pixels}} * 100 \quad (6)$$

Finally, we compared the location of the subimage with the highest vessel density to the polyp boundaries of the ground truth automatically. For this, we have determined the boundaries of ground truth, and if the subimages with the highest vessel density were within the boundaries, we assumed that the polyp location was correctly determined.

III. RESULT

In order to explain and demonstrate our methods and findings visually we have prepared Figure 3. The original image is shown in Figure 3a. Firstly, we applied the pre-processing step to this image, converted into the HSV color space format, and applied the CLAHE only V channel for contrast enhancement. Figure 3b shows that vessels and other details became more distinct because the green channel of color images formatted as an RGB image gives the highest contrast between vessels and the background; this channel is used for extraction of the image edges. Therefore, we extracted

the green channel as shown in Figure 3c. In addition, we used the complementary function to enhance the visualization of the blood vessels of the colon and in the output image, dark areas became lighter and light areas became darker (Figure 3d). Later, we used again CLAHE but this time we applied on the whole image. Subsequently, we used the morphological operation to highlight the blood vessels of the colon and we used the morphological opening function for thinning the colon blood vessels. We used 2D median filter to enhance vessels and remove noise from morphologic operation applied images. Later we used 'imadjust' function and this operation increases the contrast of the output image it is shown in Figure 3e. We used the threshold function for segmentation of the blood vessels. This function converted the gray-scale image in Figure 3e to a binary image in Figure 3f, by replacing all pixels in the input image with luminance greater than level with the value 1 (white), and replacing all other pixels with the value 0 (black). Very small white objects are observed in Figure 3f. We used 'bwareopen' function, which removes all connected components that have fewer pixels than the desired pixels from the binary image, producing another binary image [15]. Figure 3g depicts blue lines that represent detected vessels on the original image. Figure 3h depicts the quantification results of each sub-image as a rate, and Figure 3i shows the comparison result of the sub-image with the highest vessel density to the blue line representing the boundaries of the ground truth. We made this comparison automatically in MATLAB environment. The last step is shown in detail in Figure 2. For example in that figure, 68th sub-image has the highest vessel density in the ground truth for this image. In addition, other sub-images having higher vessel densities in this image are also within the limits of the ground truth. This means that the polyp location we obtain from the ground truths matched the vessel density we computed.

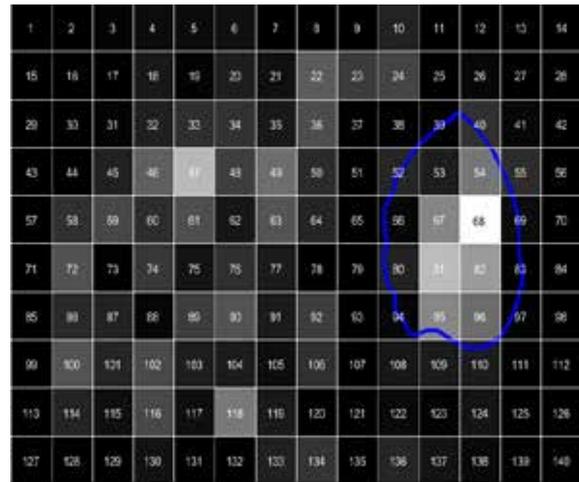


Figure 2. Quantificational 140 subimages as a 50x50 sizes from an image.

This study resulted in a total of 29 different polyps 24 of which matched the subimages with the highest vessel density.

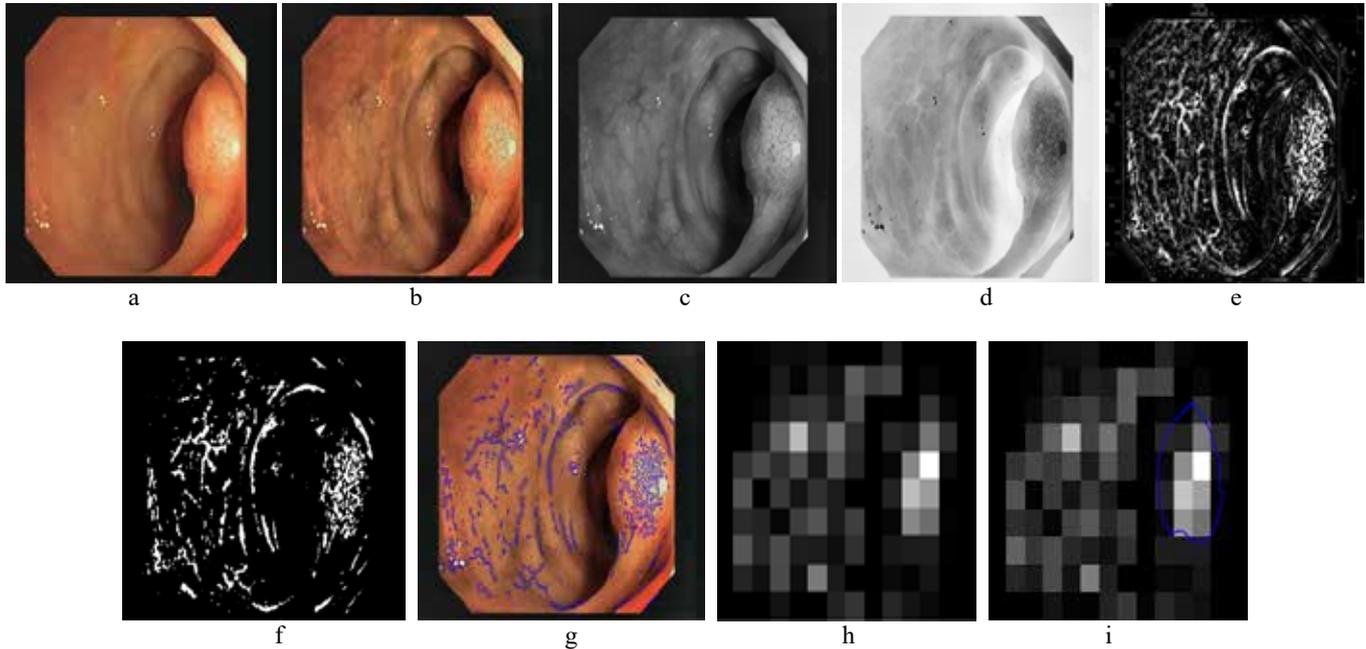


Figure 3. a) Original image, b) adaptive histogram equalization result, c) extracted green channel from RGB image, d) complementary image to enhance the visualization of the blood vessels, e) enhanced image to improve blood vessel visualization, f) blood vessels extracted using a threshold value, g) original colon image with vessel boundaries, h) vessel density image, and i) vessel density image matching the ground truth to determine the location of the polyp.

IV. DISCUSSION

Missed polyps are not always malignant, but they can turn into a malignant structure and cancer over time. Therefore, computer-aided automatic detecting has become a subject that is frequently studied recently. In order to make this distinction, it is necessary to differentiate the polyp and healthy tissue by applying the vascularization extraction methods obtained from the images. In this approach, we performed two main processes; the first step is the vessel extraction and second step is the quantification. We have experienced that these operations take 2.01 ± 0.3 seconds for each image. We can conclude that the vessel density based polyp localization is a suitable method.

Polyp position on the image, the camera angle of the colonoscopy, the number of polyps in an image, the distance and proximity of polyps to the camera, and the size of the polyp on the tissue were observed as the factors that affected the success of the polyp detection. We should also note that image segmentation is a complicated process, and unfortunately so far there is no one-size-fits-all solution.

Our future work will comprise the improvement of the vessel segmentation and quantification using more images including different polyps.

REFERENCES

- [1] World Health Organization. Fact sheet on cancer. WHO Media Centre Fact Sheets, [online], 4 July 2018.
- [2] A. Jemal, F. Bray, and J. Ferlay, "Global Cancer Statistics: 2011," *CA Cancer J Clin*, vol. 49, no. 2, p. 1,33-64, 1999.
- [3] <https://www.mayoclinic.org/diseases-conditions/colon-cancer/symptoms-causes/syc-20353669>, [online], 4 July 2018.

- [4] S. A. Kane, *Introduction to Physics in Modern Medicine*. 2009.
- [5] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, "WM-DOVA Maps for Accurate Polyp Highlighting in Colonoscopy: Validation vs. Saliency Maps from Physicians," *Comput. Med. Imaging Graph.*, vol. 43, pp. 99–111, 2015.
- [6] American Cancer Society. Cancer Facts & Figures 2015. Atlanta, GA: American Cancer Society; 2015.
- [7] S. Gurudu and J. Liang, "Information Processing in Medical Imaging," vol. 9123, no. December, 2015.
- [8] S. Ameling, S. Wirth, and D. Paulus, "Methods for Polyp Detection in Colonoscopy Videos : A Review," no. 14, 2008.
- [9] S. Hwang, J. Oh, W. Tavanapong, J. Wong, and P.C. De Groen, "Polyp Detection in Colonoscopy Video using Elliptical Shape Feature," *Proc. - Int. Conf. Image Process. ICIP*, vol. 2, pp. 465–468, 2007.
- [10] L.A. Alexandre, J. Casteleiro, and N. Nobre, "Polyp Detection in Endoscopic Video using SVMs," *KPKDD 2007 Proc. 11th Eur. Conf. Princ. Pract. Knowl. Discov. Databases*, vol. 4702, pp. 358–365, 2007.
- [11] V.B.S. Prasath and H. Kawanaka, "Vascularization Features for Polyp Localization in Capsule Endoscopy," *Proc. - 2015 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2015*, no. 2, pp. 1740–1742, 2015.
- [12] M. B. Patwari, "Classification and Calculation of Retinal Blood vessels Parameters," *IEEE's International Conferences For Convergence Of Technology, Pune, India*, no. April 2015, pp. 1–6, 2014.
- [13] <http://www.ee.columbia.edu/~sfchang/course/dip-S04/sample/Contrast-Enhancement/contrast-enhance-chang.htm>, *Contrast Enhancement Techniques*, [online], 4 June 2018.
- [14] R.C. Gonzalez and R. E. Woods, *Digital Image Processing, Nueva Jersey*. p. 976, 2008.
- [15] <https://www.mathworks.com/help/images/ref/bwareaopen.html>, *Bwareaopen*, [online], 4 July 2018.