

COLORECTAL POLYPS DETECTION IN VIRTUAL COLONOSCOPY USING 3D GEOMETRIC FEATURES AND DEEP LEARNING

Mohamed Yousuf^{*†} Samir Harb^{*††} Islam Alkabbany^{*} Asem Ali^{*} Salwa Elshazley[†] Aly Farag^{*}

^{*} Computer Vision and Image Processing Laboratory (CVIP), University of Louisville, Louisville, KY

[†] Kentucky Imaging Technologies, Louisville, KY

[‡] Faculty of Engineering, Ain Shams University, Cairo, Egypt

^{††} Higher Technological Institute, 10th of Ramadan City, Egypt

ABSTRACT

Early diagnosis of colorectal polyps, before they turn into cancer, is one of the main keys for treatment. In this work, we propose a framework to help radiologists in identifying polyp candidates using virtual colonoscopy. In the proposed approach, first a colon is segmented from a CT scan, then 3D reconstruction, to generate a surface representation of the colon, is performed. From the reconstructed 3D colon, 2D images are generated using the virtual colonoscopy, Fly-In approach. To enhance polyp detection, we fuse these 2D images and the 3D colon representation by generating 3D geometric feature maps, e.g. depth and curvature maps. CNN-based models are trained and validated to detect polyps using the generated feature maps, which are combined in multi-channel images. These images are successfully used to train a CNN-based model that detects polyps with mAP $\sim 97.1\%$.

Colorectal cancer, colon polyp, computerized tomography (CT), Detection, CNN, segmentation.

1. INTRODUCTION

Colorectal cancer (CRC) begins as small growths (polyps) that attach to the luminal wall of the colon or rectum. These growths must be diagnosed and treated promptly. If colon polyps are not diagnosed and treated at the right time, they may grow in size and become cancerous. That is why the American Cancer Society (ACS) recommends that people at average risk of colorectal cancer should start regular screening starting at age 45 [1]. The two screening procedures, which have the highest performance among other procedures, are the optical colonoscopy (OC) and Computed Tomography Colonography (CTC). However, one fourth of colorectal polyps are missed with performing screening [2]. Consequently, an automated system for the detection of polyps could be essential in reducing the number of missed polyps by providing the location and size of any potential polyps during the screening process. The Optical colonoscopy (OC) procedure, in which the colon and rectum are viewed using an endoscope, is the gold standard screening approach. The use of colonoscopy has been significant in terms of diagnosis, treatment, and screening. In addition, captured videos during OC can be analyzed to diagnose and detect colon polyps. Many studies have been introduced to analyze OC videos. Many of these algorithms have been discussed in different reviews, e.g., Hassan et al. [2] and Barua et al. [3]. Although, colonoscopy is a secure process, yet complications can arise, such as perforatio [4], bleeding [5] among other complications [6].

The second screening procedure of CRC is Computed Tomography Colonography (CTC), which can detect polyps with a similar

accuracy of OC but with lower cost. The American College of Radiology Imaging Network (ACRIN) [7] performed standard CTC and OC on 2531 patients at 15 study centers in the US. The study showed that per-patient sensitivity and specificity for CTC were: 0.9 ± 0.03 and 0.86 ± 0.02 , respectively, which are very close to OC. So, CTC is considered as a reliable and robust procedure to detect colon polyps. Therefore, there are several algorithms have been introduced to detect colon polyps within CT scans. Godkhindi and Gowda [8] trained a CNN classifier to identify the CT slices, which have polyps. However, their approach does not localize polyps in the CT slices. Uemura and Näppi [9] proposed a method to classify polyps using 3D CNN, however, the network needs high computational power in order to be trained on 3D data as shown in the experiments section. Furthermore, Tan's approach [10] classifies the polyp type if it was benign or malignant using a geometric feature called Gray-Level Co-Occurrence Matrix (GLCM) with AUC = 0.91. Zhang's work [11] was built on Tan's work by making a multi-model to extract the GLCM features with different sizes, which raised the AUC = 0.94 and raised the accuracy by 0.05 to be 0.964. However, these two methods require a polyp region as an inputs and they can't detect such regions. On the other hand, Virtual Colonoscopy (VC), which is a part of the Computed Tomography Colonography (CTC) procedure, is considered cheaper and safer alternative way for screening [12, 13]. Moreover, VC can be used to perform massive scale screening for polyp early detection and it also can help in performing the early detection in economic depressed regions and rural areas where there are limited availability of OC. Virtual colonoscopy mimics OC by visualizing the luminal surface of the colon using different visualization approaches. So many of the video analysis approaches that have been used in OC can be used with VC.

In this paper, we propose an approach, shown in Fig. 1, for colonic polyps detection in virtual colonoscopy. To develop an efficient approach that overcomes polyps detection and localization challenges, we exploit the 3D surface information in VC to generate different geometric features that discriminate between polyp and non-polyp regions. These features are used to train a CNN-based model for polyps detection. The main contributions of this work include developing:

1. a colon segmentation approach to accurately reconstruct 3D colon surface,
2. multiple approaches to extract 3D surface geometric features that increase detection performance, and
3. a CNN-based model, which is trained and validated using generated features that encode 3D surface geometry.

2. PROPOSED APPROACH

The Computed tomography colonography (CTC) framework consists of four-steps, as shown in Fig. 2-(a), starting from an abdominal

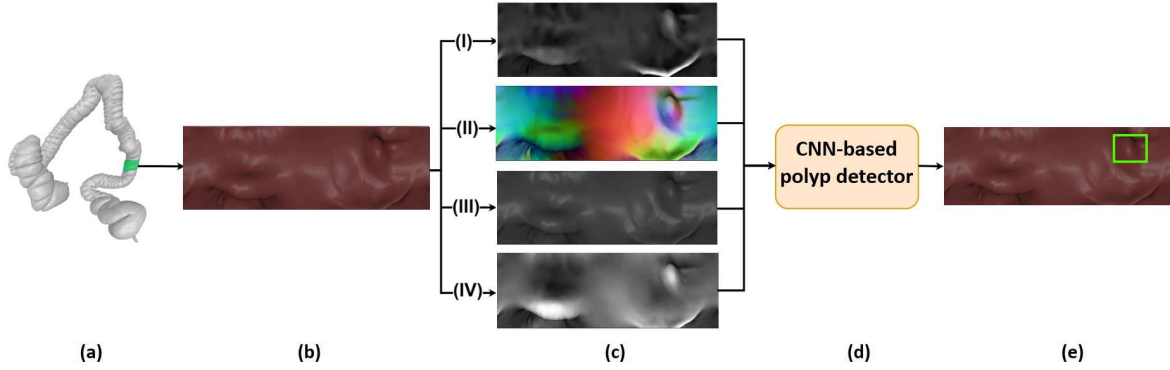


Fig. 1. The pipeline of the proposed polyp detection approach in VC: (a) The reconstructed 3D colon. (b) A desk-like rig of virtual cameras captures the inner wall of the region of interest. (c) Generated images: (I) The curvature map, (II) The normal map, (III) The classical Fly-In visualization albedo and lightning image, and (IV) The depth map. (d) A YOLO-based detector localizes polyps using the concatenated images in (c). (e) The locations and sizes of the detected polyp candidates.

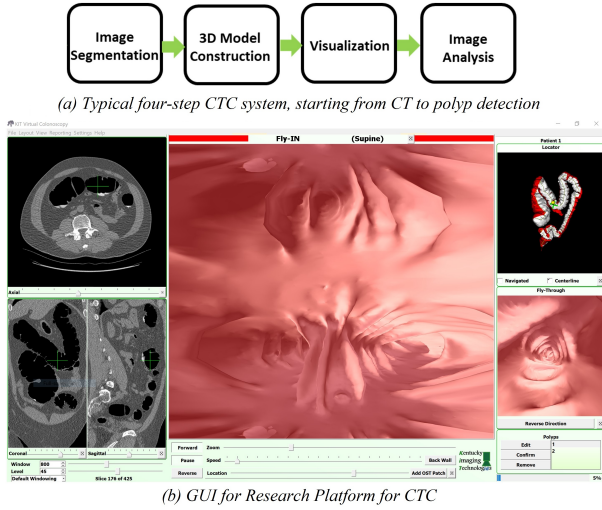


Fig. 2. CTC Platform: (a) Typical pipeline; (b) CTC components: 2D CT slices, 3D model reconstruction and visualizations: Fly-In and Fly-Through.

CT scan of a prepped patient. These steps are: 1) Image segmentation to isolate the lumen from the rest of tissues in the abdomen (e.g., the liver, lungs and small intestines, in addition to addressing uncertainties of CT acquisition); 2) 3D model building to construct the colon, which also involves extracting the centerline as a datum for visualization, and the registration of supine and prone CTC scans; 3) Visualization to present the lumen on radiology stations with details in 3D and corresponding 2D CT, in addition to functionalities for polyp editing; and 4) Analysis step which performs polyp detection, classification and archiving, and preparation of a full patient record. In this work, we focus on the first and last steps.

2.1. Images from Virtual Colonoscopy

To generate a sequence of images that visualize the luminal surface of a colon, as shown in Fig. 2-(b), the colon should be segmented from a CT scan then a 3D surface should be reconstructed. Finally, virtual camera(s) can be used for visualization, e.g. 1) Fly-Through [14] uses a virtual camera on the centerline to mimic OC, and 2) Fly-IN [15] uses a ring of virtual cameras around the centerline showing 360° field of view (FOV) of colon segments internally.

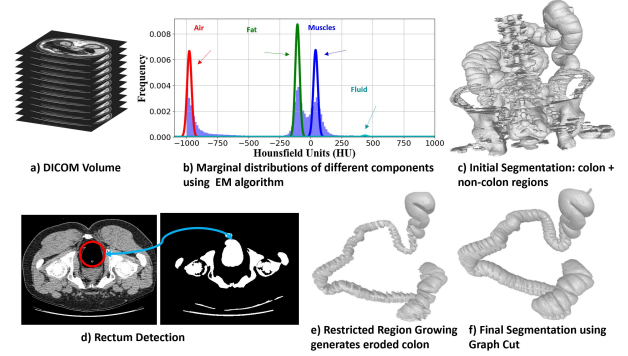


Fig. 3. The Proposed Colon Segmentation Approach

2.1.1. Colon Segmentation and 3-D reconstruction

Colon segmentation is a challenging problem due to the colon's asymmetric topology. Also, uncertainties appear due to the presence of Hounsfield (HU) intensity regions consisting of air, soft tissue, and high-attenuation structures like the bones. Also, there are other complications due to the presence of residual stool, parts of the diaphragm, lungs, and disconnected colon segments [16]. In this paper, we propose a segmentation approach that involves multiple steps as shown in Fig. 3. The first step is to calculate the empirical distributions of Hounsfield intensities in a DICOM volume, as shown in Fig. 3-b. The main components of a colon are the air, for which the characteristic peaks are almost at -1000 HU [17], and the fluid whose Hounsfield intensity is greater than 300 HU.

To extract the colon components, first, we estimate the marginal densities of air, fat, muscle, and fluid by fitting four Gaussian components using the Expectation Maximization (EM) algorithm, as shown in Fig. 3-b. Then, we identify colon regions using two thresholds. The first threshold, $< t_1$, is between air and fat (e.g., ~ -577 HU), and the second threshold, $< t_2$, is between muscle and fluid (e.g., ~ 305 HU). The HU values of colon regions should be $< t_1$ and $> t_2$.

An initial segmentation of both air and fluid could be generated using these two thresholds. However, this simple thresholding technique cannot isolate colon regions from non-colon regions because other tissues, e.g., lungs, have the same low HU as air. Also, there are non-colon regions, e.g., bones, that have a high HU as fluid, as

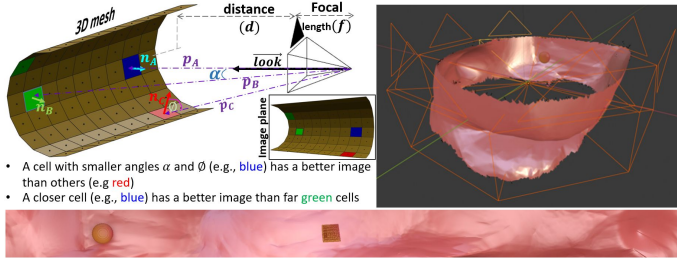


Fig. 4. Visualization of colon segment by FI. (Left) Camera geometry showing the cell visualization depends on: principal axis $\overrightarrow{look}$, projection direction p , surface normal n and distance d ; (right) Cameras configuration of a rig of 8 cameras over a ring; and (bottom) the rendered file.

shown in Fig. 3-c. Therefore, the initial segmentation is used to generate seeds from colon regions. These seeds are fed to a Graph Cut (GC) algorithm that formulates the segmentation problem as a graph partitioning problem. To extract seeds from the initial segmentation, the rectum region is identified in the first part of the DICOM volume. The rectum can be extracted as a disk-like region that has a low HU, as shown in Fig. 3-d. This region is used as a starting seed, from which other colon regions are extracted by region growing. However, since there are non-colon parts in the abdominal CT scan (e.g., small intestine, bones, etc.) that are interwind with colon parts, this yields errors in the region growing step. Therefore, restricted region growing is performed using the morphological operation to guarantee more separation between the two aforementioned classes (i.e., colon and non-colon). The output of the region growing step, shown in Fig. 3-e, is used as a seed for GC to generate the final segmentation, shown in Fig. 3-f.

2.1.2. Colon Visualization

After a colon is segmented, 3D reconstruction to generate a surface representation of the colon is performed and the centerline of the 3D colon is generated [18]. Then, the virtual colonoscopy Fly-In approach [15] is developed to provide state of the art accuracy in polyp detection. This method involves a desk-like rig of virtual cameras centered on the centerline of a reconstructed colon, providing 360° visualization of a cylindrical region of interest (ROI) which, when projected, provides a “filet”-like display of the internal surface of the ROI, as shown in Fig. 4. By moving the ROI along the centerline of the colon, radiologists would be able to examine the luminal surface and detect colonic polyps.

An optimal visualization approach projects most of the surface cells on the image plane without local deformations or loss. The visualization loss can be defined as a function of three factors: The angle (α) between the projection direction (p) and the camera’s principal axis ($\overrightarrow{look}$); the angle (Φ) between the projection direction (p) and the cell’s normal vector (n); and finally, the ratio between the camera focal length (f) and the cell’s distance (d) to the projection center on the direction of ($\overrightarrow{look}$). Figure 4 demonstrates the Fly-In on a rendered segment of the colon and shows a rig of 8 cameras over a ring. A near distortionless filet of the colon segment is shown as well. By adjusting the visualization frustum for the cameras, we can control the size of displayed ring and resolution.

2.2. Polyps Detection based on Modified Fly-In Approach

In the proposed polyp detection framework, Fly-In is used to generate 2D images of the internal surface of the colon for better visualization of the colon wall. In addition to the conventional images formed

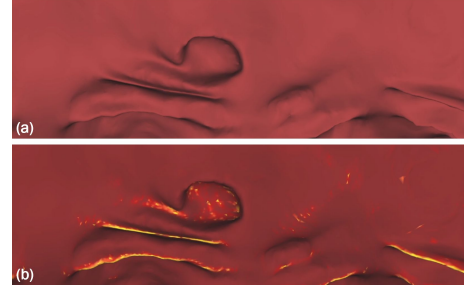


Fig. 5. a) An RGB image captured by virtual cameras. b) Adding curvature information to the image highlight the convex and concave regions.

by Fly-In (i.e., albedo and lightning), Fig. 1-c(III), we generate sets of images coding the geometric surface features:

1. The surface curvature map, as shown in Fig. 1-c(I), in which the curvature is calculated using the algebraic point set surface [19]. The curvature is based on moving least squares (MLS) fitting algebraic spheres to the surface.
2. Normal map which could be computed for each surface by getting two vectors on the surface. Then cross product these two vectors, for each vertex on the surface we can get two vectors by subtracting the 3d coordinates (x, y, z) of the vertex from two neighbors 3d coordinates as shown in Fig. 1-c(II). The normal map is represented as a 3-channel (RGB) image to represent the normal vector (N_x, N_y, N_z).
3. The depth map that reflects the smallest distance between each surface point and the centerline, as shown in Fig. 1-c(IV).

Our hypothesis is that fusing the 2D projections and the 3D colon representation in virtual colonoscopy can enhance polyp detection accuracy. The reason of this expected enhancement is that images formed by Fly-In can’t accurately encode convex (e.g., polyps) and concave (e.g., folds) surface regions if improper lightning is used, see Fig. 5. Therefore, the generated feature maps are combined as multi-channel image and fed into a deep learning model, which can be trained to extract polyps candidates in these images.

3. EXPERIMENTAL RESULTS

3.1. Dataset

Two sets of data are used to validate our proposed framework. For segmentation evaluation, we use a dataset provided by the American College of Radiology Imaging Network (ACRIN) [7] and Walter-Reed medical center. The polyps detection modules were trained and validated on a dataset of supine and prone scans of 49 patients, provided by CTC experts from the University of Wisconsin. The dataset contains 59 annotated polyps larger than 6 mm, detected and annotated by one of the three experienced radiologists.

3.2. Training and testing procedure

In order to assess the performance of the proposed automatic segmentation approach, the intersection over union metric, $IoU = \frac{g \cap s}{g \cup s}$, was used to compare the automatic segmentation S with the ground truth G . The results demonstrate that the proposed automatic segmentation approach was successful in segmenting 74% of the cases with an IoU of more than 90%.

For the polyp detection approach, we test each geometric feature separately, to know which one is the best to be used. Table 1 shows the comparison for each proposed geometric feature as an input for

Table 1. The mAP validation results for different approaches for the 3D-based detection model.

	YOLO-V7	Dynamic-RCNN R-50	Faster-RCNN R-50	Faster-RCNN R-101	RetinaNet Efficient-Net	Sparse-RCNN	Swin Transformer
Curvature (I)	97.1%	77.1%	76.2%	81.6%	83.6%	45.1%	94.0%
Normal (II)	79.4%	94.7%	85.8%	74.9%	93.0%	39.2%	90.3%
Original (III)	84.5%	91.5%	91.0%	92.0%	93.7%	64.37%	88.1%
Depth (IV)	82.9%	95.4%	88.6%	83.8%	92.5%	62.6%	89.2%

different state-of-the-art detection approaches like YOLO V7 [20], dynamic RCNN with ResNet 50 [21], faster RCNN with ResNet 50 or ResNet 101 as a backbone [22], Retina Net with Efficient Net backbone [23], Sparse RCNN [24], and Swin Transformer [25].

For the training procedure, all the detector's training batch size = 16 except the Swin transformer which has batch size = 8 due to the size of the model, but all other hyperparameters were the original hyperparameters from each model's original papers. All the training was performed on Nvidia TITAN RTX 24 Gb. Moreover, all models were trained to a high number of epochs (2000 epochs), and the best validation score for each model was chosen to be the tested model.

Since the main goal is to detect all polyps, Mean Average Precision (mAP) is the best metric to report in our application. As shown in Table 1, the curvature feature shows the highest mAP score with the use of YOLOv7, which is the most appearance feature that can distinguish between the curvature type if it is convex curvature or concave curvature. The colon contains a lot of concave curvatures which can deceive the detector, so a feature that can distinguish concave and convex surfaces is needed to make the detector work best since polyps have mostly a convex curvature shape. That is why the curvature feature is the most important. The second best feature is the depth feature with Dynamic RCNN network. Since depth can distinguish between the near and far tissue surfaces, it can give the model the needed feature for distinguishing between polyp and non-polyp parts. Then comes the normal feature with Dynamic RCNN network. We believe that Swin transformer has a big potential to achieve better results but it needs more polyps data than we have to train.

4. CONCLUSION

In this work, we proposed an approach to identify the locations and sizes of the polyps in virtual colonoscopy. The modified Fly-In method was developed to generate images encoding 3D surface geometry features. Then a CNN-based detector was trained using these images to extract potential polyp regions. High performance, i.e. mAP $\sim$ 97.1%, encourages radiologists to use the proposed approach in identifying polyp candidates using virtual colonoscopy.

5. COMPLIANCE WITH ETHICAL STANDARD

This research study was conducted retrospectively using human subject data made available in open access by ACRIN, Walter Reed Medical Center, and the University of Wisconsin. The research protocol is governed by the University of Louisville IRB No. 07.0252.

References

- [1] "The american cancer society," <https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/acs-recommendations.html>.
- [2] C Hassan and et. al., "Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis," *Gastrointestinal endoscopy*, 2021.
- [3] Ishita Barua and et. al., "Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis," *Endoscopy*, 2020.
- [4] Gaurav Arora et al., "Risk of perforation from a colonoscopy in adults: a large population-based study," *Gastrointestinal endoscopy*, 2009.
- [5] Ankie Reumkens et al., "Post-colonoscopy complications: a systematic review, time trends, and meta-analysis of population-based studies," *ACG*, 2016.
- [6] Gregory S Cooper et al., "Complications following colonoscopy with anesthesia assistance: a population-based analysis," *JAMA internal medicine*, 2013.
- [7] C D Johnson and et. al., "Accuracy of ct colonography for detection of large adenomas and cancers," *NEJM*, vol. 359, no. 12, 2008.
- [8] Akshay M Godkhindi and Rajaram M Gowda, "Automated detection of polyps in ct colonography images using deep learning algorithms in colon cancer diagnosis," in *ICECDS*.
- [9] Tomoki Uemura and et. al., "Comparative performance of 3d-densenet, 3d-resnet, and 3d-vgg models in polyp detection for ct colonography," in *Medical Imaging 2020: Computer-Aided Diagnosis*. SPIE, 2020.
- [10] Jiaxing Tan et al., "3d-glm-cnn: A 3-dimensional gray-level co-occurrence matrix-based cnn model for polyp classification via ct colonography," *IEEE transactions on medical imaging*, 2019.
- [11] Shu Zhang et al., "Mm-glm-cnn: A multi-scale and multi-level based glm-cnn for polyp classification," *CMIG*, 2023.
- [12] Perry J Pickhardt et al., "Colorectal cancer: Ct colonography and colonoscopy for detection—systematic review and meta-analysis," *Radiology*, 2011.
- [13] Thomas R De Wijkerslooth et al., "Reasons for participation and nonparticipation in colorectal cancer screening: a randomized trial of colonoscopy and ct colonography," *ACG*, 2012.
- [14] Koojoo Kwon et al., "Virtual anatomical and endoscopic exploration method of internal human body for training simulator," *Journal of Korean Medical Science*, 2020.
- [15] Mostafa Mohamed et al., "Fly-in visualization for virtual colonoscopy," in *ICIP*, 2018.
- [16] Zina Ravindran et al., "Automatic segmentation of colon using multilevel morphology and thresholding," in *ICCCI*, 2021.
- [17] Janne J Nappi et al., "Effect of knowledge-guided colon segmentation in automated detection of polyps in ct colonography," in *Medical Imaging 2002*, 2002.
- [18] M Sabry Hassouna and Aly A Farag, "Robust centerline extraction framework using level sets," in *CVPR*. IEEE, 2005.
- [19] Gaël Guennebaud and Markus Gross, "Algebraic point set surfaces," in *ACM siggraph 2007 papers*, pp. 23–es. 2007.
- [20] Chien-Yao Wang et al., "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv preprint arXiv:2207.02696*, 2022.
- [21] Hongkai Zhang et al., "Dynamic r-cnn: Towards high quality object detection via dynamic training," in *ECCV*, 2020.
- [22] Shaoqing Ren et al., "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, 2015.
- [23] Tsung-Yi Lin et al., "Focal loss for dense object detection," in *ICCV*, 2017.
- [24] Peize Sun et al., "Sparse r-cnn: End-to-end object detection with learnable proposals," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.
- [25] Ze Liu et al., "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021.