A Two-Level Approach Towards Semantic Colon Segmentation: Removing Extra-Colonic Findings

Le Lu[1], Matthias Wolf[1], Jianming Liang[1,2], Murat Dundar[1,3], Jinbo Bi[1], and Marcos Salganicoff[1]

[1] CAD & Knowledge Solutions, Siemens Healthcare, Malvern, PA 19355, USA
[2] Biomedical Informatics Dept., Arizona State University, AZ 85004, USA
[3] Computer Information Science Dept., IUPUI, IN 46202, USA

Abstract. Computer aided detection (CAD) of colonic polyps in computed tomographic colonography has tremendously impacted colorectal cancer diagnosis using 3D medical imaging. It is a prerequisite for all CAD systems to extract the air-distended colon segments from 3D abdomen computed tomography scans. In this paper, we present a two-level statistical approach of first separating colon segments from small intestine, stomach and other extra-colonic parts by classification on a new geometric feature set; then evaluating the overall performance confidence using distance and geometry statistics over patients. The proposed method is fully automatic and validated using both the classification results in the first level and its numerical impacts on false positive reduction of extra-colonic findings in a CAD system. It shows superior performance than the state-of-art knowledge or anatomy based colon segmentation algorithms [1,2,3].

1 Introduction

Colon cancer is the second leading deadly cancer for western population. Early detection and removal of colonic polyps is the critical step which helps to decrease the risk of colon cancer and improve survival. Computed tomographic colonography (CTC) or virtual colonoscopy has been widely used for detecting colorectal neoplasms via 3D computed tomography (CT) abdomen scans of the cleansed and air-distended colon. In the last decade, many computer aided detection and diagnosis systems [4,5,6,7] have been proposed and actively studied to improve the performance and reliability of human radiologists as second readers.

For any CAD system, the first processing is to identify all cleansed and air-distended colon segments, for detecting colonic polyps as next step, in 3D abdomen CT scans. Knowledge or anatomy based colon segmentation algorithms are described in past literature [1,2,3]. A two-step method is proposed, including "anatomy-based extraction" to find and remove non-digestive components, such as outer-air, bones and lung; and "colon-based analysis" to segment out extra-colonic parts (e.g., small bowel and stomach) by region growing from a seed in colon rectum. Removing extra-colonic lumen fragments is our main focus since non-digestive components are more isolated and can often be reliably excluded.

By considering the potential presence of partially or completely collapsed colon regions, multiple colon seeds [1,2] are needed to cover the decent portion of true colon volume. Numerous heuristic rules and thresholding parameters, e.g., when to stop region growing, how many seed points needed and how many slides apart to look for the next seed, are necessary to generate the final results. In [3], a centerline chaining based colon segmentation method is described. However the essential drawbacks of rule-based methods (as above) remains and the colon linking scheme [3] is only valid for well-distended or slightly collapsed colon cases.

In this paper, we present a two-level approach for colon segmentation. First all air segments (after pre-segmentation) are classified as either colon components or extra-colonic parts (small intestine, stomach and others) by statistical modeling on novel geometry feature sets per segment. Then we evaluate the overall segmentation quality using a new colon "daisy-chaining" tracking algorithm integrated with distance and geometry statistics over patients, which handles moderately or poorly distended colon significantly better than previous methods. The proposed method is validated using both the classification results in the first level and its numerical impact on false positive reduction of extra-colonic findings in a CAD system. It shows the superior performance than previous heuristic knowledge or anatomy based colon segmentation algorithms [1,2,3], on a 166 volume dataset. The remainder of this paper is organized as follows. Section 2 gives the details of the materials used and our taken approach. Experimental results and analysis are reported in section 3. Lastly, section 4 concludes the paper and directs future work.

2   Materials and Methods

An accurate segmentation of a specific organ (e.g., colon) is a very crucial part of a CAD system. Only if the organ can be accurately segmented it is possible to perform a reliable and thorough examination of that organ. A case of under-segmentation can lose important parts of an organ which may lead to miss potential lesions whereas an over-segmentation may generate avoidable false positives in areas outside of the organ. Our colon segmentation approach is composed of three steps: intensity-appearance based initial segmentation, independent colon component classification (level 1) and anatomical colon tracking (level 2).

Colon Pre-segmentation: Datasets of 166 prone and supine CT volumes from 92 patients, obtained from four hospitals and scanned using both Siemens and GE machines under different imaging protocols, are used to conduct our study. For pre-segmentation stage, a pixelwise thresholding is first performed on the entire volume based on a predefined air Hounsfield (HU) value (e.g., -800 HU). Pixels are then labeled as foreground (air) and background (non-air). Afterwards, adjacent foreground pixels are merged as fragments or components by running a standard connected component scheme. This step identifies all air-filled regions within the body and can often contain parts other than the colon, such as stomach, lung, small intestine and other random noises. Colon pre-segmentation is a general purpose process and handled similarly by other systems, thought more complicated thresholding and merging rules can be adopted in [1,5].

2.1 Colon Fragment Classification

Annotation & Data: We design and implement a user interface for marking each individual fragment after initial segmentation of potential colon areas. Each segmented component is labeled as one of the five anatomical categories: colon, small intestine, mixed (of colon and small intestine through opened ileo-cecal valve which happens in low frequency), stomach or lung, and noise. Since mixed class contains partial colon volume, both colon and mixed categories are treated as positive class. The rest of the categories are considered negative. It results 165 positive samples and 1665 negative samples in the training dataset of 110 volumes. There are 89 positives and 836 negatives based on a unseen testing dataset with 56 CT scans. Around 40% volumes have collapsed colon segments.

Feature Extraction: The success of colon and non-colon classification relies on discriminative features. In this paper, we explore a set of geometric (+ spatial) features, including volume, length, area, volume-length ratio and area-length ratio. For each potential colon component from pre-segmentation, the volume refers its total volume; the length is an estimation of the length computed by the *shortest-path*algorithm [8]; and the area means its surface area. The volume-length ratio and area-length ratio are the ratios of volume to length and area to length. We denote this as four feature set $\mathcal{F}_4$. Additionally, the normalized centroid position in 3D volume coordinates of each fragment can also be integrated as $\mathcal{F}_7$ where positions are normalized in the range of $[0, 1]$ (i.e., bounded by the extreme positions of initial segmented components).

Classification: Support vector machine (SVM) has been proven as an effective approach for classification. In particular, 1-norm SVM [9] can be used to construct classifiers and select important features simultaneously. The 1-norm SVM determines classifiers by minimizing a regularized training error $\lambda P(\cdot) + \sum_i \xi_i$ where ξ is a hinge loss occurred on the i-th training example and the regularization term $P(\cdot)$ is a 1-norm penalty on the weight vector of the linear classifier. The 1-norm SVM can be formulated as a linear program (LP) which, in general, can be solved efficiently by existing optimization tools. The classification accuracy is cost sensitive to the size or volume measurements of the segmented regions, in the context of colon segmentation task. It is treated as a more severe error if a large segmented volume (i.e., retaining more potential polyp candidates) is misclassified than a small region is missed. Hence we revise 1-norm SVM to associate the hinge loss occurred on each classification instance with a weight, ie. the volume size of the example. At last, we optimize the following problem to construct the classifier $\{w, b\}$

$$\min_{w,b} \lambda ||w||_1 + \sum_i \nu_i \xi_i$$
$$\text{s.t.} y_i(w^T x + b) \geq \xi_i, \quad \xi_i \geq 0. \tag{1}$$

where $x \in \mathcal{F}_4$ or $\mathcal{F}_7$. Given the trained linear classifier, receiver operating characteristic (ROC) curves can be drawn by thresholding the classifier output $(w^T x + b) \geq \tau$ at different operating points τ. The volume fragments classified as positive colon class by setting high detection rate $>= 99\%$ will be used for the next level processing.

2.2   Colon Fragment Tracing

Given the fragments as classified as colon by the previous step, we attempt to provide a global interpretation by incorporating anatomic colon structure and performing a "Daisy-chaining" colon fragments tracing. Algorithm 1 describes the processing details, and Figure 1 shows an illustrative example where a complete colon is composed of five collapsed segments.

> **foreach** *Colon CT Volume i* **do**
> > Identify the rectum R_i and the corresponding component $S(R_i)$;
> > Identify the cecum C_i and the corresponding component $S(C_i)$;
> > **if** $R_i == true$ *and* $C_i == true$ **then**
> > > **if** $S(R_i) == S(C_i)$ **then**
> > > > return component $S(R_i)$ or $S(C_i)$, exit;
> > >
> > > **end**
> > > **while** *True* **do**
> > > > Mark the rectum component $S(R_i)$ the first colon fragment;
> > > > Search the component S_j from unidentified colon segments that is "closest" (with distance D) to the lastly merged colon fragment;
> > > > **if** $D <= T$ **then**
> > > > > merge S_j with the already known colon fragments;
> > > > > **if** $S_j == S(C_i)$ **then**
> > > > > > return all current colon segments, exit;
> > > > >
> > > > > **end**
> > > >
> > > > **end**
> > > > **if** $D > T$ *or* $S_j == false$ **then**
> > > > > return all current colon segments, exit;
> > > >
> > > > **end**
> > >
> > > **end**
> >
> > **end**
>
> **end**

Algorithm 1. Colon "Daisy-chaining" Tracking Algorithm

Rectum and cecum are important landmarks as they describe the course of the entire colon, and they can be reliably located (100% for rectum; $\geq 85\%$ for cecum) via fusing the direct detection of themselves and other more extinctive landmarks on vertebrae and hip bones. In Algorithm 1, if there is no detection of cecum, all cecum relevant checking will be omitted. After colon tracking, a geodesic distance checking between rectum and cecum is used to reject cecum mislocation. The distance threshold T is calibrated using the distance statistics obtained from all valid collapsed colon connections of 110 training volumes. The mean distance is $29.78mm$ with standard deviation as $15.06mm$ and we set $T = 60mm$ for all following evaluations. Other richer descriptive geometry features (e.g., gradient orientations and curvatures on tubular shortest-paths) can be employed and are under study. Overall, $> 92\%$ volumes complete "Daisy-chaining" colon tracking and further remove more false positives. A probabilistic fusing formulation of fragment connectness modeling and verification, other than simple thresholding, can also be investigated in future work.
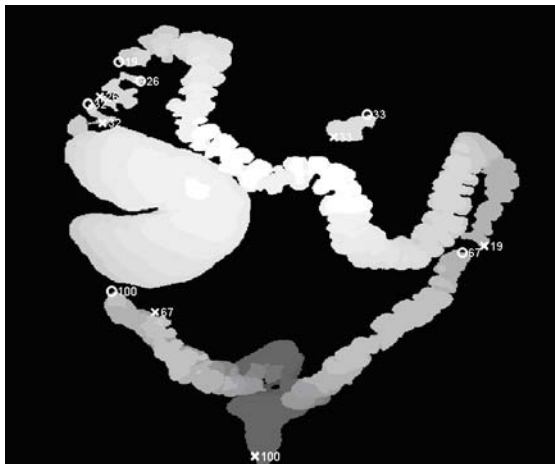
Fig. 1. This shows a colon (as maximum intensity projection image) that is broken into several pieces. The symbols "X" and "O" denote the start and end point of a component. The numbers indicate points belonging to the same component. In this example, the rectum component is #100 and the detected rectum landmark is close to "X100". The cecum component is #32. The complete colon consists of #100, #67, #19, #26, and #32 by our tracking scheme. Component #33 does not belong to the colon.

3   Experimental Results

In this section, we describe both the evaluation of colon segmentation using 1-norm SVM [9] as a single classification issue, and its integration and effectiveness on removing extra-colonic findings (ECR) in a CAD system.

Colon Segmentation Evaluation: We first analyze the performance characteristics of our classification based colon segmentation module. We demonstrate two types of ROC curves using per count or volume (mm^3) as measurement metrics, i.e., how many fragments (or how much volume) are correctly or incorrectly classified, which are plotted in Fig. 3. Our ECR module enables to remove $> 90\%$ or higher extra-colonic volumes (mm^3), at the detection rate of 99.5%, for training or testing dataset respectively. In [2], $83.2\% \sim 92.8\%$ colon volumes can be semi-automatically segmented or partially segmented. Manual seed placement is required in 16.8% CT scans. [1] detects 98% of the visible colonic walls, and approximately 50% of the extra-colonic components are removed, using 88 volumes. [3] reports sensitivity of 96% for colon volume and it is unclear, to what degree collapsed colons can be handled, in a smaller dataset of total 38 scans. An illustrative example of the MIP (Maximum Intensity Projection) image of a 3D colon volume, before and after ECR colon segmentation, is also shown in Fig. 2.
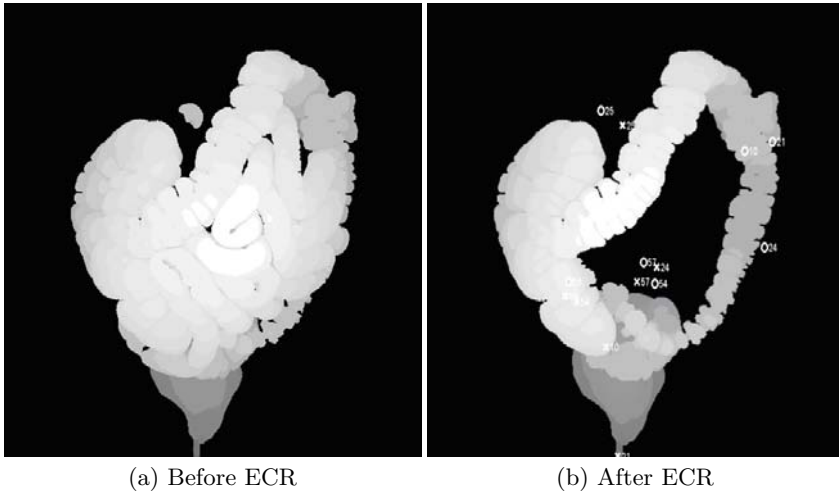
(a) Before ECR   (b) After ECR

Fig. 2. An illustrative example of showing the MIP (Maximum Intensity Projection) image of a 3D colon volume before and after ECR colon segmentation

ECR Impacts on CAD False Positive Reduction: We assign each true and false polyp candidate with the classification score $(w^T x + b)$, obtained by the colon fragment which contains itself. The empirical probability density functions (PDF) of positive (true) and negative (false) polyp candidates are plotted for both training and testing CT datasets in Fig. 4. It is clear that the blue peaks around $[0, -5]$ of FP distribution can be removed without significantly affecting the detection rate of true positives in red[1].

In another validation setting, we analyze and manually label 236 total false positive (FP) candidates, as the final outputs of our CAD system, after polyp classification but without applying ECR. 89 of them are verified as extra-colonic FPs and the rest are colonic. Our ECR colon segmentation (as a post-filter module) further removes 80 FP candidates where only 3 of them are actually colonic and 77 are true extra-colonic findings. It results in sensitivity of $77/89 = 86.5\%$ and specificity as $(147-3)/147 = 98\%$, which conforms the high detection rate setting in section 2. For true positive analysis, there is a very minimal but non-zero risk of losing polyps because of ECR colon removing, but we have not seen such cases in our extensive CAD pipeline study (out of 166 volumes used to build ECR). Finally, though our whole colon segmentation subsystem is trained using 110 tagged colon CT scans processed after electronic cleansing, it generalizes well and achieves similar performance on the clean-prep colon cases based on our further study.

---

[1] The kernel density estimation based PDF fitting smooths and slightly skews both distribution curves in Fig. 4. The minimum ECR score for positive polyp candidates is -2.1652. When ECR is used as pre-filtering for our colon CAD system, the total candidate number per volume drops from > 200 to 142, without missing any true polyp candidate.
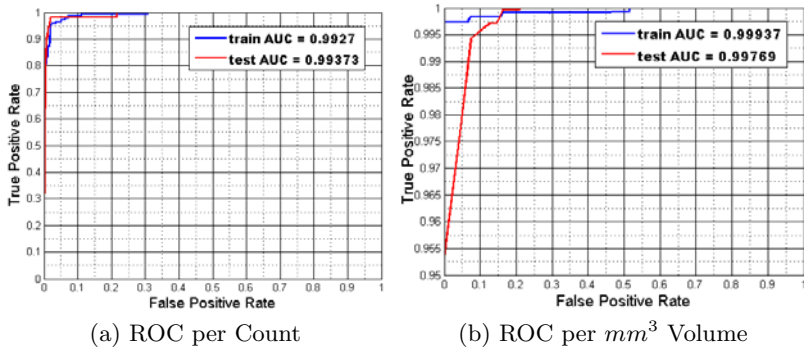
(a) ROC per Count   (b) ROC per mm^3 Volume

Fig. 3. ROC curves for colon fragment (air-filled connected component) classification using norm-1 SVM. (a) is measured as per fragment count, (b) is measured using absolute mm^3 volume. Red curves are for training and blue for testing.
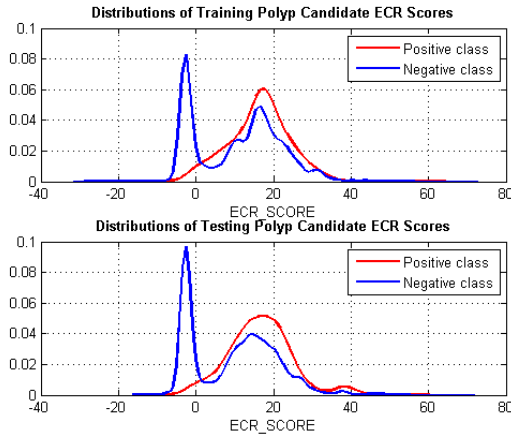

Fig. 4. The plot of ECR scores of positive/negative polyp candidates for both training (Top) and testing (Bottom) CT datasets, as outputs of the linear 1-norm SVM classifier on $\mathcal{F}_4$

4   Discussion

In this paper, we present a fully automatic, statistical based colon segmentation approach. After a generic pre-segmentation, it first classifies all air-filled fragments as binary colon or non-colon classes using continuous 1-norm SVM classifier on newly explored geometric features. Then a colon fragment tracking algorithm is employed to measure the overall colon segmentation confidence using anatomic information. The proposed method is validated using both the classification results on colon segmentation and its numerical impact on false positive reduction of extra-colonic findings in a CAD system. It shows superior

performance than previous knowledge or anatomy based colon segmentation work [1,2,3], using different but larger datasets.

The colon segmentation problem benefits from the general statistical models learned from a population, but it can also be considered as a binary testing/decision task of colon or non-colon components per volume case. In future work, the ideal colon classification process should be performed by considering all component information in one volume as a set, to make an optimal binary cut. Pairwise affinity based graph partitioning algorithm [10] can be employed as an alternative solution. Besides exploring more descriptive features for colon structure topology model (as in section 2.2), We also intend to address the model adaptation issue for groupwise or personalized colon segmentation using self-normalized tests [11], which can be capable of applying and adapting the trained model to datasets from different races (e.g., from Caucasian to Asian).

References

1. Näppi, J., MacEneaney, P., Dachman, A., Yoshida, H.: Knowledgeguided automated segmentation of colon for computer-aided detection of polyps in ct colonography. J. Comput. Assist. Tomogr. (2002)
2. Iordanescu, G., Pickhardt, P., Choi, J., Summers, R.: Automated seed placement for colon segmentation in computed tomography colonography. Academic Radiologyg 12, 182–190 (2005)
3. Frimmela, H., Näppi, J., Yoshida, H.: Centerline-based colon segmentation for ct colonographyy. Med. Phys., 2665–2672 (2005)
4. Baker, M., Bogoni, L., et al.: Computer-aided detection of colorectal polyps: Can it improve sensitivity of less-experienced readers? preliminary findings. Radiology 245, 140–149 (2007)
5. Yoshida, H., Näppi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. IEEE Trans. Medical Imaging 20, 1261–1274 (2001)
6. Summers, R., Jerebko, A., et al.: Colonic polyps: Complementary role of computer-aided detection in ct colonography. Radiology 225, 391–399 (2002)
7. Chowdhury, T., Whelan, P., Ghita, O.: A fully automatic cad-ctc system based on curvature analysis for standard and low dose ct data. IEEE Trans. Biomedical Engineering 55, 888–901 (2008)
8. Dijkstra, E.W.: A note on two problems in connections with graphs. Numerische Mathematic 1, 269–271 (1959)
9. Zhu, J., Rosset, S., Hastie, T., Tibshirani, R.: 1-norm support vector machines. In: Neural Information Processing System (2004)
10. Weiss, Y.: Segmentation using eigenvectors: A unifying view. In: Proc. of ICCV, pp. 975–982 (1999)
11. Gangaputra, S., Geman, D.: Self-normalized linear tests. In: Proc. of CVPR, pp. 616–622 (2004)