

# CAD for Polyp Detection: an Invaluable Tool to Meet the Increasing Need for Colon-Cancer Screening

P.Cathier<sup>a</sup>, S.Periaswamy<sup>a</sup>, A.Jerebko<sup>a</sup>, M.Dundar<sup>a</sup>, J.Liang<sup>a</sup>, G.Fung<sup>a</sup>,  
J.Stoeckel<sup>a</sup>, T. Venkata<sup>a</sup>, R.Amara<sup>a</sup>, A.Krishnan<sup>a</sup>, B.Rao<sup>a</sup>, A.Gupta<sup>a</sup>,  
E.Vega<sup>b</sup>, S.Laks<sup>b</sup>, A.Megibow<sup>b</sup>, M.Macari<sup>b</sup>, and L.Bogoni<sup>a\*</sup>

<sup>a</sup> *Computed Aided Diagnosis and Therapy, Siemens Medical Solutions*

<sup>b</sup> *NYU Medical Center*

---

**Abstract.** Colorectal cancer is the third most common cancer in both men and women and it was estimated that in 2003 nearly 150,000 cases would be diagnosed and 57,000 people would die. Screening has been accepted as a means for early detection of the disease, yet only a portion of the eligible population undergoes colorectal cancer screening, partially due to comfort issues of undergoing a full colonoscopic examination. Virtual colonoscopy has been demonstrated to be an effective means of performing screening and given the number of people who are candidate for screening better tools, such as CAD, are required to fulfill this increasing need. The developed CAD system, presented in this paper, has focused on the detection of polyps of sizes up to and including 20mm. The results have demonstrated that: the developed algorithm generalizes, the sensitivity and specificity for middle- to large-sized polyps is on the average 95% while the overall sensitivity is roughly 88% and the false positive has remained at a manageable 4 per volume.

*Keywords:* CAD, automated polyp detection, virtual colonography, colon screening.

---

## 1. Introduction

Colorectal cancer is the third most common cancer in both men and women. In [1] it was estimated that in 2003, nearly 150,000 cases of colon and rectal cancer would be diagnosed in the US, and more than 57,000 people would die from the disease, accounting for about 10% of all cancer deaths. While there is wide consensus that screening patients is effective in decreasing advanced disease, only 44% of the eligible population undergoes any colorectal cancer screening [2]. Multiple reasons have been identified for non-compliance, key being: patient comfort, bowel preparation and cost [3]. Virtual colonoscopy (VC) derived from computer tomographic (CT) images is gaining a broader acceptance as a screening method for colorectal neoplasia [4], and that better tools are needed to facilitate the efficiency of radiologists' efforts in detecting lesions.

Recent studies [4, 5] have demonstrated that radiologists' ability to detect polyps using CT-Colonography (CTC) or VC has sensitivities ranging between 80% and 90% with higher sensitivity for polyps larger than 10mm. The majority of these studies have

---

\* Corresponding author. *E-mail address:* Luca.Bogoni@siemens.com

involved experienced radiologists. However, in order to make CTC, and VC in general, a ubiquitous practice for screening, it is necessary that the average community radiologist to reproduce these results. Tools, such as Computer Aided Detection (CAD), may help in assuring a high level of sensitivity. In this paper, we present results suggesting that, in a rather near future, CAD could be employed as a second reader to aid in polyp detection.

The CAD method presented compares favorably with other CAD systems for polyp detection [6,7,8] in which reported sensitivity rate ranged between 83% and 100% and with the false positive rate ranged between 1 and 6 false positives per case. A direct comparison across the different CAD system is, however, difficult since the validation methods, patient selection and preparation, the data acquisition protocols and quality vary

## 2. Methods

The database of high-resolution CT images used in this study was obtained at NYU Medical Center. The cases were selected to develop and test the sensitivity and specificity of CAD as a tool to aid in polyp detection. 105 patients were selected so as to include positive cases (n=61) as well as negative cases (n=44). The sensitivity and specificity was established with respect to CTC by comparison to results from concurrent fiber-optic colonoscopy. The radiologist's findings, specified in a report, were converted to actual locations and measured by a medical student who, using Siemens' *syngo* Colonography, associated marks to locations (real coordinates). The medical student reviewed the CT data using both 2D and 3D endoluminal views. These markings were reviewed and confirmed by the experienced radiologist. Cases were anonymized; a masterfile was kept with the physician for any follow up clarifications, and then exported to CD. The locations and lesion's dimensions were then used in subsequent stages to automatically compute sensitivity and specificity with respect to polyp size.

The ability of the radiologist, and of the present CAD system, to detect polyps greatly depends on colon distension, insufflation and cleanliness. The protocols for bowel preparation and data acquisitions were those employed at NYU Medical Center. Oral phosphosoda was used for bowel cleansing, followed by evacuation prior to CT scans. Patients were scanned following manual insufflation with room air in both prone and supine positions. CT data was acquired using a Volume Zoom CT scanner (Siemens Medical Systems, Malvern, PA), 4x1mm slice detector, 120kV, 0.5 sec. gantry rotation and effective 50mAs. Pitch varied between 6 and 7 such that the entire abdomen and pelvis could cover within a 30 seconds breath-hold. Effective coverage speed varied between 12 and 14 mm /sec. CT images were reconstructed as 1.25 mm thick sections with a 1.0 mm reconstruction interval.

The proprietary algorithm (patents pending) included the following phases: data pre-processing, candidate generation, feature extraction, and classification. The data pre-processing phase includes colon segmentation and a transformation to render the volume isotropic. During candidate generation, based on a simple shape filter, loci of detection are identified. These are sequentially processed during the following phase in which multiple features are extracted. The features are based on moments of tissue intensity, volumetric and surface shape and texture characteristics. Each candidate, uniquely

identified, and the associated features are then fed to a classifier. Candidates are then evaluated and labelled as potential polyps. In the actual workflow, these would then be presented to the physician for review. The primary focus of the developed algorithms has been that of yielding high sensitivity and specificity. At present, the running time single volume, 375 slices - 512x512, was roughly 9 minutes on a Pentium IV 3.06GHz dual-processor machine with 2GB of memory. Major decrease computation times are expected once the desired combination of individual modules will have been refined.

In order to automatically process all the cases, a flexible framework was developed which would allow loading the cases (read as DICOM images), and push them through the various stages of the algorithm, outputting intermediary results. The modularity of the system has allowed assessing the effectiveness of each component and hence both improve and independently refine them.

### **3. Feature Selection, Classification and Results**

Training and Test Data: The 105 patients were randomly partitioned into two groups: training (n=64) and test (n=41). The test group was sequestered and only used to evaluate the performance of the final system. In order to automate the training and verification process, an Access™ database was developed to allow the software to automatically query the database and provide feedback as to whether the finding was a polyp or non-polyp and, in the case of polyps, also obtain its dimensions. The training group was used to design the classifier and to select the relevant features as described below.

Feature Selection: The feature selection stage is a key component of our approach. We use the “wrapper” method for feature selection [9] in which we let the classifier decide which features are useful. Essentially, the classifier is run iteratively on the training data using different feature sets – during each iteration, one or more features are added, until the cross-validation error no longer improves. The final feature subset contained 14 features, selected from an initial set of 66.

Test Results: We evaluated the system’s performance on the 64 patients in the training group using Leave-One Patient Out (LOPO) cross-validation. In this scheme, we leave-out both the supine and prone views of one patient from the training data. The classifier is trained using the volumes from the remaining 63 (i.e., 64-1) patients, and tested on both volumes of the “left-out” patient. This process is repeated 64 times, leaving out each of the 64 patients in the training group, and the resulting testing errors are averaged out to determine the LOPO error.

Results on Test Group: A classifier was trained using all 64 patients in the training data set using the 14 features selected in the “Feature Selection” phase. Only after this classifier was frozen, was the test group of 41 patients released for evaluation. This provides a very accurate estimate of system performance on completely unseen data – the only true test for a classification-based system. In the computation of the sensitivity, we counted a polyp as “found” if it was detected in at least one of the volumes (supine or prone) from the same patient.

#### *3.1 LOPO Results on training group*

Patient and Polyp Info: There were 64 patients with 125 volumes (some patients did not have both prone and supine studies). A total of 42 polyps were identified in this set: 12 small-size (1-5 mm), 21 mid-size (6-9mm), 9 large-size (10-20mm). Candidate Generation: The candidate generation stage generated an average of 28.5 candidates per volume while missing 2 small-size polyps. Classifier Results: We obtained a false positive rate of 4.1 per volume. The sensitivities obtained for different ranges of polyps are as follows: small-sized = 83%, mid-sized=95%, large sized=100% and overall=93%.

### 3.2 Results on Test Group (previously unseen)

Patient and Polyp Info: There were 41 patients with 82 volumes. A total of 35 polyps were identified in this set: 16 small-size (1-5 mm), 10 mid-size (6-9mm), 9 large-size (10-20mm). Candidate Generation: The candidate generation stage generated an average of 29.4 candidates per volume while missing 4 small polyps. Classifier Results: We obtained a false positive rate of 3.9 per volume. The sensitivities obtained for different ranges of polyps are as follows: small-sized = 69%, mid-sized=90%, large sized=100% and overall=83%.

## 4. Conclusions

The developed CAD system has focused on the detection of polyps of sizes up to and including 20mm. The results have demonstrated that: (a) the developed algorithm generalizes, when applied a completely unseen test set, albeit all the data is from a single institution; (b) the sensitivity and specificity for middle- to large-sized polyps is on the average 95% while the overall sensitivity is roughly 88%; (c) while the false positive rates could be improved, they have remained at a manageable 4 per volume often occurring in the small intestine and ileocecal valve – these could be reduced by more accurate pre-processing; however, these findings are readily eliminated by radiologists; and finally that (d) the reported sensitivity and specificity are well in the range of the reported sensitivities of expert radiologists.

Additionally, 3 polyps larger than 6 mm and 1 smaller than 5 mm, which were detected prospectively by the CAD system, were only found retrospectively by the radiologist following OC. CAD system also detected 2 polyps larger than 6 mm that were missed by conventional colonoscopy but found by radiologist. These observations further demonstrate the value of this CAD system. While the above population also included 10 masses, those have not been included as part of the statistics reported, since they are rarely missed by radiologists. Future studies will not only focus on improving the overall performance, but also evaluate the performance of the present algorithm when presented with data from various institutions and subject to different degrees of preparation.

---

## References

- [1] Martinez-Stoffel E. & Syngal S., "Colon Cancer Screening Strategies", *Curr Opin Gastroenterol* 18(5): 595-601, 2002.
- [2] "Trends in screening for colorectal cancer—United States, 1997 and 1999" *MMWR Morb Mortal Wkly Rep* 2001, 50:162-166.

- [3] Zack DL, DiBaise JK, Quigley EM, et al.: "Colorectal cancer screening compliance by medicine residents: perceived and actual". *Am J Gastroenterol* 2001,96:3002-3008.
- [4] Pickhardt P.J, Choi J.R., Hwang I., Butler J.A., Puckett J.L., Hildebrand H.A., (+et.al.), "Computer Tomographic Virtual Colonoscopy to Screen for Colorectal Neoplasia in Asymptomatic Adults", *NEJM*-2003, 349(23): 2191-2200.
- [5] Yee J., Geetanjali A.A., Hung R.K., Steinauer-Gebauer A.M., Wall A.D., McQuaid K.R., "Colorectal Neoplasia: Performance Characteristics of CT Colonography for Detection in 300 Patients", *Radiology*-2001, 219:685-692. Jerebko A.K., Malley J.D., Franaszek M., Summers R.M.. "Multinetwork classification scheme for detection of colonic polyps in CT colonography data sets", *Acad Radiol* 2003; 10:154-160.
- [7] Nappi. J., Yoshida Y., "Automated Detection of Polyps with CT Colonography", *Acad Radiol* 2002; 9:386-397.
- [8] Gokturk SB, Tomasi C, Acar B, Beaulieu CF, Paik DS, Jeffrey RB, Yee J, Napel S. "A statistical 3-D pattern processing method for computer-aided detection of polyps in CT colonography", *IEEE Trans Med Imaging* - Dec 2001 (Vol. 20, Issue 12).
- [9] John G.H., Kohavi R., Pfleger K., "Irrelevant Features and the Subset Selection Problem", *International Conference on Machine Learning*, 1994.