

An Investigation of Texture Features Based on Polyp Size for Computer-Aided Diagnosis of Colonic Polyps



Yeseul Choi, Alice Wei, David Wang, David Liang, Shu Zhang,
and Marc Pomeroy

1 Introduction

According to recent statistics published by the American Cancer Association, colorectal cancer is the second leading cause of deaths of men and women combined, with an estimated 101,420 new cases in 2019 (American Cancer [1, 9]). Early detection and removal of polyps significantly reduces the risk of death and screenings are recommended for those 50 or older. However, using clinical optical

Y. Choi

Department of Radiology, Stony Brook University, Stony Brook, NY, USA

Stuyvesant High School, New York, NY, USA

A. Wei

Department of Radiology, Stony Brook University, Stony Brook, NY, USA

Staten Island Technical High School, Staten Island, NY, USA

e-mail: alice.wei22@sitechhs.com

D. Wang (✉)

Department of Radiology, Stony Brook University, Stony Brook, NY, USA

Syosset High School, Syosset, NY, USA

D. Liang

Department of Radiology, Stony Brook University, Stony Brook, NY, USA

Ward Melville High School, East Setauket, NY, USA

S. Zhang

Department of Radiology, Stony Brook University, Stony Brook, NY, USA

M. Pomeroy

Department of Radiology, Stony Brook University, Stony Brook, NY, USA

Department of Biomedical Engineering, Stony Brook University, Stony Brook, NY, USA

© Springer Nature Switzerland AG 2021

H. R. Arabnia et al. (eds.), *Advances in Computer Vision and Computational Biology*, Transactions on Computational Science and Computational Intelligence,
https://doi.org/10.1007/978-3-030-71051-4_66

847

colonoscopy (OC), to screen the growing population of those over 50 is severely limited by our current resources. Alternatively, computed tomography colonography (CTC) has been a rising noninvasive solution to detect possible cancers [3, 5, 6, 11, 12], decreasing the amount of people being screened under OC. Furthermore, the majority of polyps are found to be nonneoplastic, named hyperplastic, which are abnormal growths with no risk, only taking up valuable resources when removed. Therefore, being able to distinguish between hyperplastic and adenomatous polyps through CTC screenings has become crucial.

In our previous research work, texture feature extraction techniques have been established and modified to improve the accuracy of diagnoses of polyps found through CTC screenings [4, 10]. Those texture features include intensity, gradient, and curvature. Pickhardt et al. suggests that the malignancy rates of polyps increase in correlation to their size [8]. Besides polyp's malignancy rates, general kernel approach for gradient and curvature texture extraction may also get influenced by the polyp size [7]. Thus, in this study, we design experiments to investigate the effect of separating polyps based on size on the performance of machine learning.

The remainder of this chapter is organized as follows. Section 2 depicts our method for conducting feature analysis in computer-aided diagnosis of colorectal polyps. In Sect. 3, experimental design and results are reported. Finally, discussion of our work and conclusions are given in Sect. 4.

2 Materials and Methods

The aim of this study is to investigate the performance of machine learning on texture features based on different polyp sizes. Figure 1 shows a 6-mm small-sized polyp and a 22-mm medium-sized polyp visualized via two-dimensional (2D) axial image.

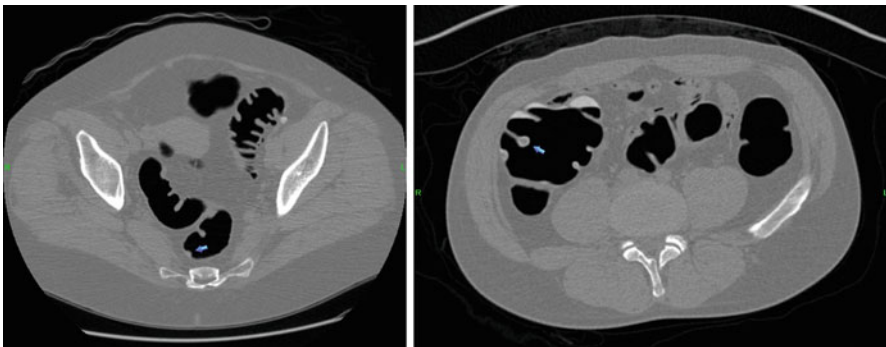


Fig. 1 A 6-mm polyp (left) and a 20-mm polyp (right) on 2D axial image

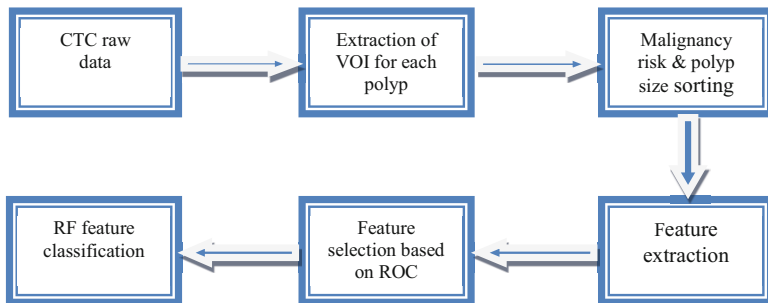


Fig. 2 Flowchart of our method for texture analysis

2.1 Flowchart of Our Method

The flowchart of our methods can be summarized in Fig. 2. First, from the original CT image, the volume of interest (VOI) of each polyp was extracted and further confirmed by the radiologists. Second, we applied an extended traditional Haralick model [2] with a total of 30 texture features [4], which were obtained from the VOIs of each polyp. In our study, the texture features of polyps consist of intensity, gradient, and curvature. Third, we performed a receiver operating characteristics (ROC)-based analysis to select the best feature sets. Finally, we employed a machine learning method, the Random Forest (RF) classifier, for feature classification.

2.2 Malignancy Risk of the Polyps

According to histopathology, polyps can be divided into two categories: nonneoplastic and neoplastic. Nonneoplastic polyps are benign polyps, including the subtypes of hyperplastic, mucosal polyp, and inflammatory, etc. On the other hand, neoplastic are malignancy risky, including serrated adenoma, tubular adenoma, tubulovillous adenoma, adenocarcinoma, etc. In this study, we labeled the nonneoplastic polyp as 0 and neoplastic polyps as 1, indicating their benign or malignancy risk accordingly.

2.3 Data Preparation

The dataset used in this study consists of a total number of 228 polyp masses found through a CTC database. Those polyp masses were confirmed by OC. The spatial resolution of the CT image is 0.7 by 0.7 by 1.0 mm³. All the polyp masses have a diameter size ranging from 6 to 30 mm. Polyps were grouped into three groups based on their sizes: 6–9 mm, 10–30 mm, and a combined group of 6–30 mm. For a

fair performance comparison, we have a balanced polyp pool of 114 polyps in each size group. Within each size group, all of the datasets had 57 benign polyps and 57 malignant polyps.

3 Results

We investigated the performance of the algorithm on the 6–9 mm, 10–30 mm, and 6–30 mm groups. The studied features include intensity, gradient, curvature, all combined features. We generated a measure of area under the curve (AUC) values of ROC curve. The performance was determined by assessing the sensitivity and specificity values. The averaged AUC information is illustrated in Table 1. Figure 3 shows the averaged ROC curve for the intensity feature. Figure 4 shows the averaged

Table 1 Averaged AUC information

Group	AUC information		
	6–9 mm	10–30 mm	6–30 mm
Intensity	0.7437	0.6942	0.7559
Gradient	0.6768	0.7226	0.6803
Curvature	0.5747	0.6493	0.6004
All features	0.7327	0.7915	0.7678

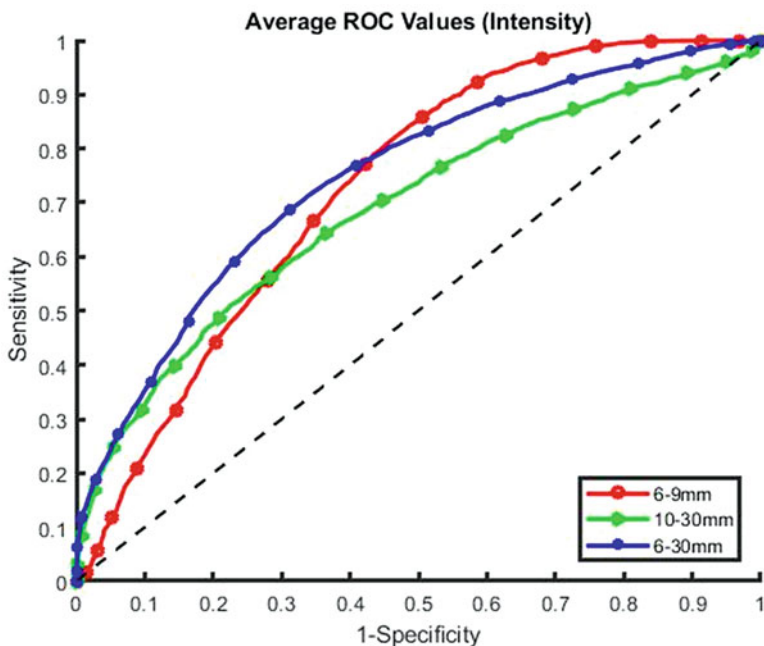


Fig. 3 The averaged ROC curves for the intensity feature

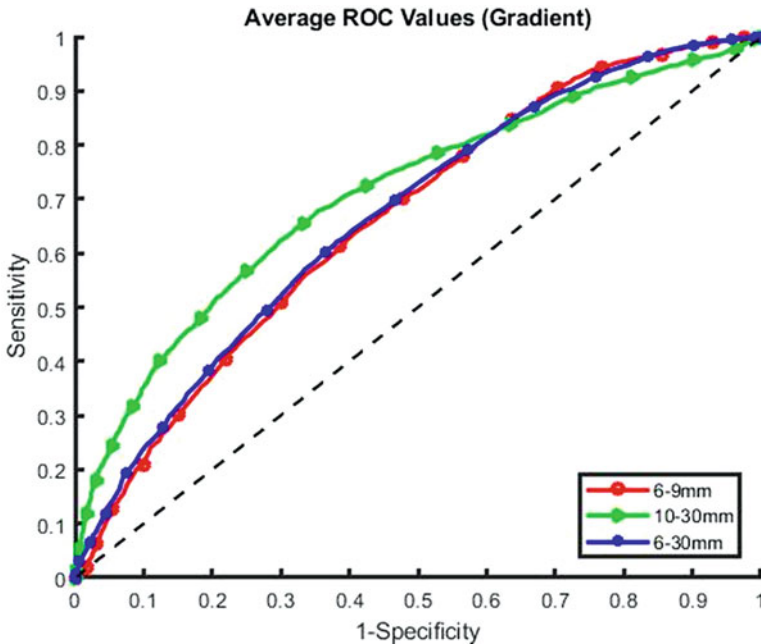


Fig. 4 The averaged ROC curves for the gradient feature

ROC curve for the gradient feature. Figure 5 shows the averaged ROC curve for the curvature feature. And Fig. 6 shows the averaged ROC curve for all combined features.

4 Discussion and Conclusions

Experimental results demonstrated that gradient and curvature were ideal distinguishing features for medium-sized polyps, whereas intensity was better for smaller-sized polyps. Due to their negligible proportions, the curvature and gradient features were flawed for the 6–9 mm polyps during the experiment. The opposite is true for the medium-sized polyp group.

When examining all features, the AUC value of the 10–30 mm polyps was greater than the AUC value of the 6–30 mm polyps, suggesting that separating the small and medium-sized polyps would be beneficial in identifying medium-sized polyps. However, the AUC value of the 6–9 mm polyps was lower than that of the 6–30 mm polyps, suggesting that this separation may not be ideal for identifying smaller polyps. The AUC value for all polyps is greater than that of the individual group of smaller polyps because it is averaged out by the higher performance of the identification of the 10–30 mm polyps. Furthermore, the smaller polyps have less identifiable features. This study shall facilitate computer-aided diagnosis of polyps

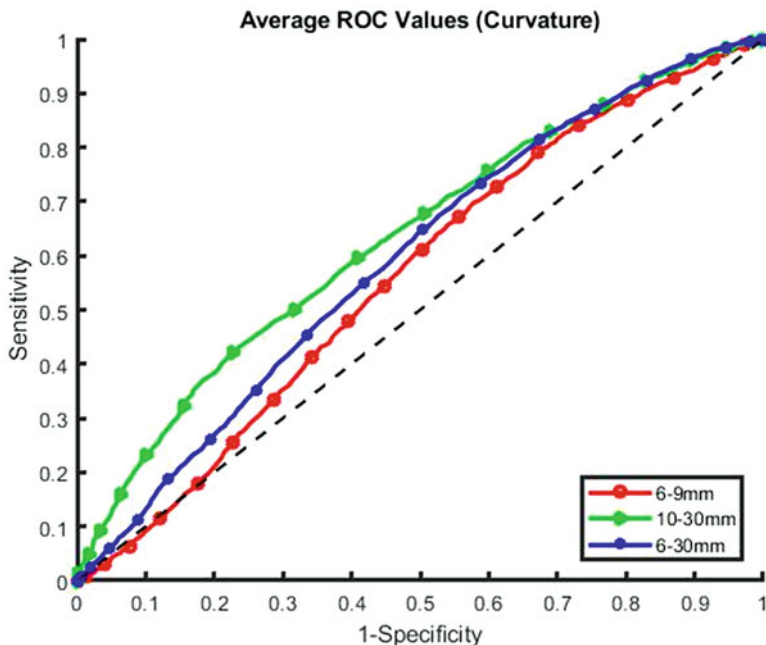


Fig. 5 The averaged ROC curves for the curvature feature

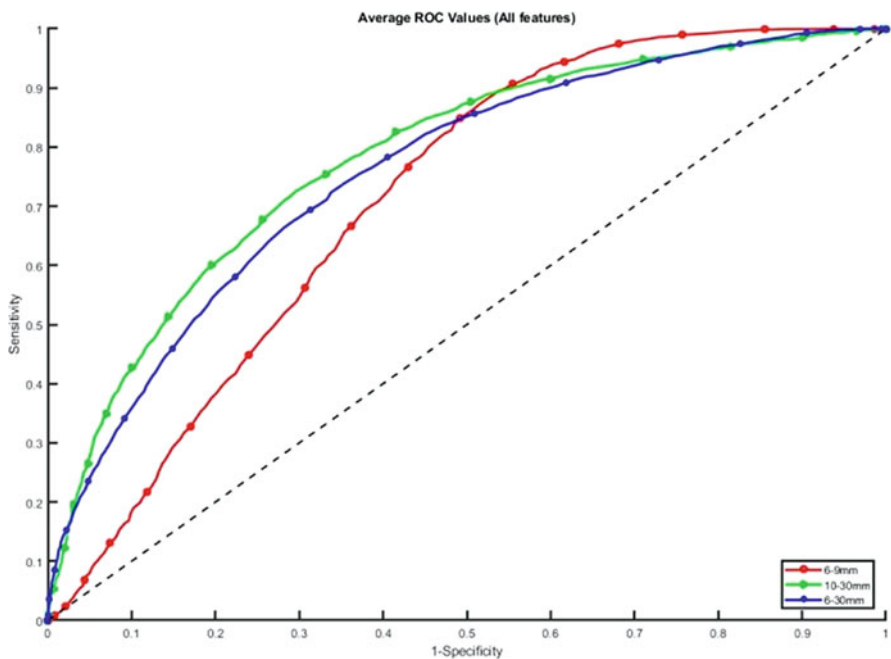


Fig. 6 The averaged ROC curves for all combined features

to achieve high performance by taking into account the contributions of different features among different polyp sizes.

Acknowledgments This work was partially supported by NIH grant #CA206171 of the National Cancer Institute and the Computer Science and Informatics Research Experience Program at Stony Brook University.

References

1. American Cancer Society, Cancer facts & figures 2019. Atlanta (2019)
2. R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **3**(6), 610–621 (1973)
3. L. Hong, A. Kaufman, Y. Wei, et al., *3D virtual colonoscopy. IEEE Biomedical Visualization Symposium* (IEEE CS Press, Los Alamitos, 1995), pp. 26–32
4. Y. Hu, Z. Liang, B. Song, et al., Texture feature extraction and analysis for polyp differentiation via computed tomography colonography. *IEEE Trans. Med. Imaging* **35**(6), 1522–1531 (2016)
5. Z. Liang, VC: An alternative approach to examination of the entire colon. *INNERVISION* **16**, 40–44 (2001)
6. Z. Liang, R. Richards, Virtual colonoscopy v.s. optical colonoscopy. *Expert Opin. Med. Diagn J* **4**(2), 149–158 (2010)
7. J. Liu, S. Kabadi, R.V. Uiter, et al., Improved computer-aided detection of small polyps in CT colonography using interpolation for curvature estimation. *Med. Phys.* **38**(7), 4276–4284 (2011)
8. P.J. Pickhardt, K.S. Hain, D.H. Kim, et al., Low rates of cancer or high-grade dysplasia in colorectal polyps collected from computed tomography colonography screening. *Clin. Gastroenterol. Hepatol.* **8**, 610–615 (2010)
9. R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2019. *CA Cancer J. Clin.* **69**(1), 7–34 (2019)
10. B. Song, G. Zhang, H. Lu, et al., Volumetric texture features from higher-order images for diagnosis of colon lesions via CTC. *Int. J. Comput. Assist. Radiol. Surg.* **9**(6), 1021–1032 (2014)
11. D. Vining, D. Gelfand, R. Bechtold et al., Technical feasibility of colon imaging with helical CT and virtual reality. Annual Meeting of American Roentgen Ray Society, New Orleans, pp. 104 (1994)
12. H. Zhu, Y. Fan, H. Lu, et al., Improved curvature estimation for computer-aided detection of colonic polyps in CT colonography. *Acad. Radiol.* **18**(8), 1024–1034 (2011)