

Chapter 4.7

Machine Learning for Automated Polyp Detection in Computed Tomography Colonography

Abhilash Alexander Miranda

Université Libre de Bruxelles, Belgium

Olivier Caelen

Université Libre de Bruxelles, Belgium

Gianluca Bontempi

Université Libre de Bruxelles, Belgium

ABSTRACT

This chapter presents a comprehensive scheme for automated detection of colorectal polyps in computed tomography colonography (CTC) with particular emphasis on robust learning algorithms that differentiate polyps from non-polyp shapes. The authors' automated CTC scheme introduces two orientation independent features which encode the shape characteristics that aid in classification of polyps and non-polyps with high accuracy, low false positive rate, and low computations making

the scheme suitable for colorectal cancer screening initiatives.

Experiments using state-of-the-art machine learning algorithms viz., lazy learning, support vector machines, and naïve Bayes classifiers reveal the robustness of the two features in detecting polyps at 100% sensitivity for polyps with diameter greater than 10 mm while attaining total low false positive rates, respectively, of 3.05, 3.47 and 0.71 per CTC dataset at specificities above 99% when tested on 58 CTC datasets. The results were validated using colonoscopy reports provided by expert radiologists.

DOI: 10.4018/978-1-60960-818-7.ch4.7

INTRODUCTION

According to the estimates provided by the American Cancer Society, there will be about 150000 new cases of colorectal cancer (CRC) and about 50000 deaths due to CRC in the United States in 2008 (American Cancer Society, 2008). For the year 2006, International Agency for Research on Cancer estimated that CRC constituted 12.8% of all cancer cases in Europe with incidence and mortality rates of 90 and 43.9, respectively, per 100000 people (Ferlay et al., 2007). The European Union acknowledges the CRC as the most common malignant disease and has put forward proposals to make CRC screening a priority for member states (Boyle et al., 2003; Council of the European Union, 2003). Screening can help detect early stage polyp growths in the colon and rectum which could be removed to prevent cancer from occurring. A polyp is an unwelcomed growth on the surface of the colon or rectum; some polyps are benign while others are cancerous. The most common screening methods for CRC are fecal occult blood testing, double contrast barium enema, colonoscopy, and computed tomography colonography (CTC) (Gluecker et al., 2003; Johnson et al., 2004; Rockey et al., 2005; Vernon, 1997).

Optimal Screening Test

Various studies conducted so far, as reviewed later, are yet to converge towards a polyp detection scheme that optimally attains high polyp detection and low false positives in real-time so that it could be recommended as a screening test for those at high risk of CRC. Though several studies have raised doubts regarding the reliability of the CTC in detecting the polyps (Levin et al., 2003; Mulhall et al., 2005; Ransohoff, 2004), it increasingly is becoming the method of choice for CRC screening. A proper understanding of the 3D shape features of polyps is vital in enhancing the accuracy of CTC (Kim et al., 2007; Shi et al., 2006; Summers et al., 2005; Yao & Summers, 2007).

Novelties in This Chapter

Two of the unique features of the presented research are: (i) design of two computationally efficient orientation independent colonic shape features per candidate polyp that offers excellent polyp - non-polyp discrimination, viz., the dominant Gaussian mass and the attenuation frequency of the shape distribution function, and (ii) state-of-the-art machine learning techniques with excellent generalization abilities are used to classify these features of polyps from haustral folds and other non-polyp tissues inside the colon.

Organization of the Chapter

The chapter is organized as follows. In Section Automated Polyp Detection in CTC we compare the various automated polyp detection schemes, particularly, the machine learning oriented techniques in general CTC literature. The practice of CTC and the data is explained in Section Materials and Methods. The process of analysis of the CT images is described in Section Feature extraction of candidate polyps, which includes colon segmentation, candidate surface extraction, and candidate shape feature extraction. The classification of the features based on three state-of-the-art machine learning techniques with emphasis on the Lazy Learning paradigm is explored and their performances are compared in Section Classification of candidate polyp surfaces.

AUTOMATED POLYP DETECTION IN CTC

CTC is the non-invasive method of detecting polyps in the colon and the rectum using a computed tomography scan. Complementing optical colonoscopy with the CTC is proposed as a viable screening test for CRC detection (Johnson & Dachman, 2000; Pickhardt et al., 2003). CTC is widely welcomed because it permits interactive

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/machine-learning-automated-polyp-detection/56177

Related Content

Machine Learning for Visual Navigation of Unmanned Ground Vehicles

Artem A. Lenskiy and Jong-Soo Lee (2012). *Machine Learning: Concepts, Methodologies, Tools and Applications* (pp. 1090-1107).

www.irma-international.org/chapter/machine-learning-visual-navigation-unmanned/56188

Construction of a Topic Database for Remote Conversation Support

Toru Homan, Masaya Morita, Yu Yamaguchi and Kazuhiro Kuwabara (2013). *International Journal of Software Science and Computational Intelligence* (pp. 77-91).

www.irma-international.org/article/construction-of-a-topic-database-for-remote-conversation-support/101319

Multilevel Image Segmentation by a Multiobjective Genetic Algorithm Based OptiMUSIG Activation Function

Sourav De, Siddhartha Bhattacharyya and Susanta Chakraborty (2013). *Handbook of Research on Computational Intelligence for Engineering, Science, and Business* (pp. 122-162).

www.irma-international.org/chapter/multilevel-image-segmentation-multiobjective-genetic/72491

On the System Algebra Foundations for Granular Computing

Yingxu Wang, Lotfi A. Zadeh and Yiyu Yao (2012). *Software and Intelligent Sciences: New Transdisciplinary Findings* (pp. 98-121).

www.irma-international.org/chapter/system-algebra-foundations-granular-computing/65125

The Formal Design Models of a Universal Array (UA) and its Implementation

Yingxu Wang, Jason Huang and Jingsheng Lei (2011). *International Journal of Software Science and Computational Intelligence* (pp. 69-89).

www.irma-international.org/article/formal-design-models-universal-array/60750