Computerized Medical Imaging and Graphics 41 (2015) 108–116

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

A general framework for wireless capsule endoscopy study synopsis

Qian Zhao [a,b,*,1], Gerard E. Mullin [b], Max Q.-H. Meng [a], Themistocles Dassopoulos [c], Rajesh Kumar [b]

[a] The Chinese University of Hong Kong, Shatin, Hong Kong
[b] Johns Hopkins University (JHU), Baltimore, MD 21218, USA
[c] Washington University School of Medicine, St. Louis, MO 63110, USA

ARTICLE INFO

Article history:
Received 4 January 2014
Received in revised form 2 April 2014
Accepted 29 May 2014

Keywords:
Computer-aided diagnosis
Wireless capsule endoscopy
Hidden Markov models
Support vector machines
Supervised classification
Study synopsis

ABSTRACT

We present a general framework for analysis of wireless capsule endoscopy (CE) studies. The current available workstations provide a time-consuming and labor-intense work-flow for clinicians which requires the inspection of the full-length video. The development of a computer-aided diagnosis (CAD) CE workstation will have a great potential to reduce the diagnostic time and improve the accuracy of assessment. We propose a general framework based on hidden Markov models (HMMs) for study synopsis that forms the computational engine of our CAD workstation. Color, edge and texture features are first extracted and analyzed by a Support Vector Machine classifier, and then encoded as the observations for the HMM, uniquely combining the temporal information during the assessment. Experiments were performed on 13 full-length CE studies, instead of selected images previously reported. The results (e.g. 0.933 accuracy with 0.933 recall for detection of polyps) show that our framework achieved promising performance for multiple classification. We also report the patient-level CAD assessment of complete CE studies for multiple abnormalities, and the patient-level validation demonstrates the effectiveness and robustness of our methods.

1. Introduction

Wireless capsule endoscopy (CE) [10,22] was invented to screen the gastrointestinal (GI) tract, especially the small bowel (previously not accessible non-invasively) using a simple outpatient test. It has significantly impacted the diagnostic approach for many diseases, such as bleeding, Crohn's and Celiac diseases, tumors, polyps, and other lesions [22].

Introduced by Given Imaging Inc. in 2000, over 1,000,000 Pillcam small bowel (SB) capsules alone have already been swallowed in the past 10 years since the device was first approved by the U.S. FDA. The size of the CE device is ϕ 11 mm × 26 mm. It consists of an imaging sensor, associated optics, and communication electronics. An outpatient examination typically produces more than 50,000 images, which are then manually and tediously examined by an expert reader by inspecting the full-length video. In addition to time constraints, this manual procedure cannot guarantee that every abnormality is detected due to their various sizes, positions and characteristics, and the experience of the clinicians. Development of a computer-aided diagnosis tool for CE assessment is therefore desirable and necessary. Fig. 1 shows some examples of CE images. The image resolution is 576 × 576. It can be seen that besides diagnostic relevant images (e.g. lesion (b) and polyp (c)), there are also a large amount of images with normal lumen (f), bile (a), air bubbles (d) and extraneous matter (e).

In this work, we propose a general framework to summarize CE videos into multiple classes. Fig. 2 shows the conceptual objective of CE study synopsis. A general hidden Markove models (HMM) is built based on statistical classifiers integrating multiple image appearance attributes (color, edge, and texture). The underlying support vector machine classifier [5] outputs are encoded as the binary observations of HMM. The proposed method is a generalized model instead of discriminative one, which can generate instances for clinical training and education. We have evaluated this composite framework by performing the video synopsis on a database of complete CE videos where study images were summarized into six most commonly seen classes – normal images, lesions, polyps,

* Corresponding author at: Johns Hopkins University (JHU), Baltimore, MD 21218, USA. Tel.: +1 2407227351; fax: +1 2024761270.
E-mail addresses: qzhao@alumni.ee.cuhk.edu.hk, kate.qzhao@gmail.com (Q. Zhao).

[1] Qian Zhao was a visiting graduate student at the Visual Imaging and Surgical Robotics Laboratory at JHU during this work. This work was supported by Dr. Kumar's discretionary funds and the HK ITF project #6902928 (PI: Max Q.-H. Meng).

http://dx.doi.org/10.1016/j.compmedimag.2014.05.011
0895-6111/

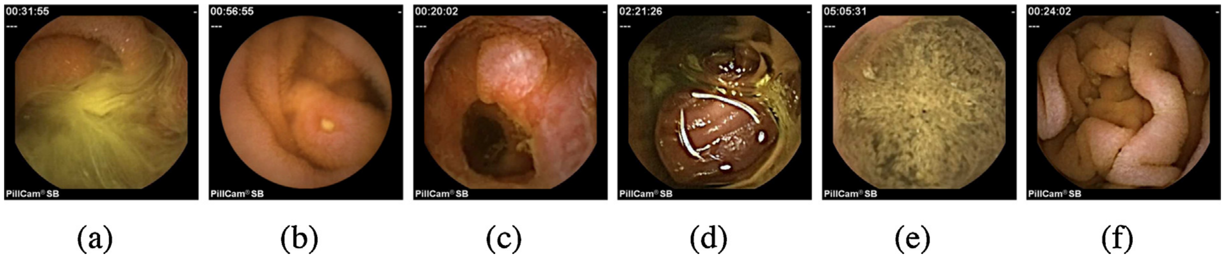

Fig. 1. Examples of CE image images – (a) bile, (b) lesion, (c) polyps, (d) large air bubble, (e) extraneous matter, and (f) normal lumen.

air bubbles, bile, and other extraneous matter. Compared to prior art reviewed above where image-based performance measures are reported, we describe patient-level validation of this general model to verify its effectiveness and robustness.

Our main contributions lie in: (a) the proposal of a general framework based on HMM integrating temporal information; (b) investigation of multi-class study synopsis; (c) development of a CAD workstation for automated CE assessment; and (d) validation on complete CE videos providing great potential for direct clinical application.

2. Related work

Prior related work mainly focused on statistical analysis for detection of various individual abnormalities, with bleeding [18,13,25] being the main focus. More recently, Yi et al. [25] introduced a clinically viable software for automated GI tract bleeding detection and classification. The major functional modules included a graph-based segmentation algorithm, specific feature selection and validation and cascade classification. The method focused on single abnormality, bleeding, while our framework is general for video synopsis with respect to multiple abnormalities. With wide application of CE, lesions [3,11,12], polyps [1,27], and non-informational frame detection [2] have also been explored in detail. Our group has also previously investigated several statistical methods for CE image analysis [12,3,20], primarily for Crohn's disease, and for patient level disease severity assessment.

However, all the above prior research (a) focuses on specific abnormalities, diseases or anatomy and (b) ignores the temporal information between neighboring images.

Some efforts have been made to analyze CE studies in the form of video streams instead of individual images, which mainly focused on applications such as small intestinal motility assessment [23], video summarization [7,15], and topographical segmentation [6]. These efforts are all limited to individual research goals or simple abnormality detection. Mackiewicz et al. [15] reported a scheme for color image analysis to discriminate between digestive organs. The method focused on detecting the boundary between the stomach and small intestine, and separately the small bowel and colon. Color, texture and motion features were used for SVM classification, and a multivariate Gaussian classifier built in a HMM framework. This topological segmentation using a simple left-to-right HMM is not sufficient to describe the complexity of abnormality sequences that may contain both normal and abnormal images. Our prior works investigated both supervised and unsupervised methods for CE video summarization [28,30,29]. All these works were validated using video clips instead of full-length videos.

More recently, Htwe et al. presented a CE video summarization method for bleeding detection [8]. The bleeding images were first detected based on color histogram and supervised classification. Then the image-level summarization was performed by using motion estimation and clustering. The color bar map was finally generated and served as the summary. The method was validated on 9 real-patient videos that were not described in detail. Furthermore, the method was currently designed specifically to detect bleeding, while our proposed method is applicable to multiple abnormalities. Iakovidis et al. attempted to reduce the number of frames in CE videos based on clustering [9]. The frames were represented by the SURF keypoints, followed by clustering using Homography estimation. The method was validated on short CE video clips and the selection of the number of cluster is manual and empirical. Their recent work stitched multiple CE images to generate a panoramic visual summaries for CE videos [21]. Liu et al. developed another study summarization method for CE videos,

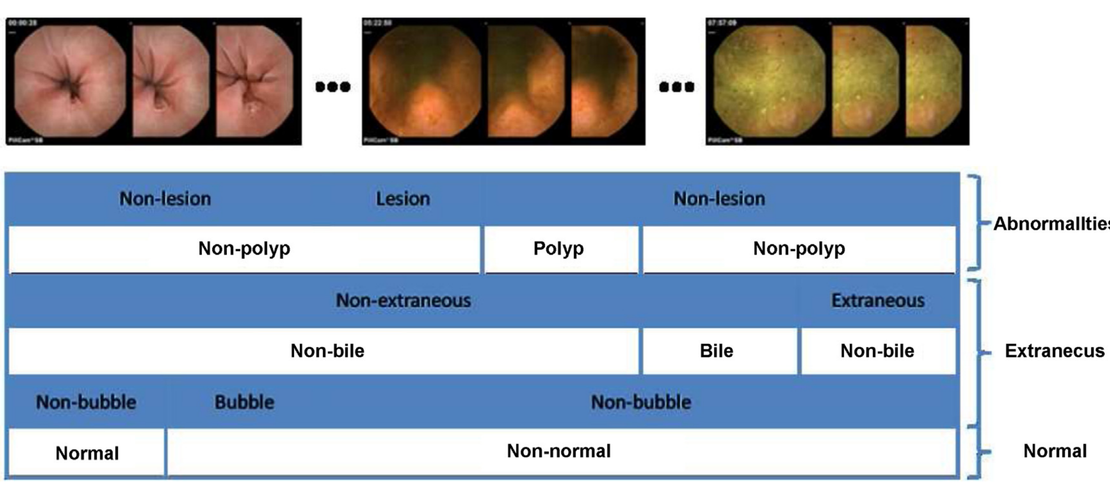

Fig. 2. The conceptual objective of CE study synopsis: a CE video is summarized into multiple classes including both abnormalities, extraneous matters and normal lumen.