7-2016

Automatic hookworm detection in wireless capsule endoscopy images

Xiao WU

Honghan CHEN

Tao GAN

Junzhou CHEN

Chong-wah NGO
Singapore Management University, cwngo@smu.edu.sg

See next page for additional authors

Automatic Hookworm Detection in Wireless Capsule Endoscopy Images

Xiao Wu*, *Member, IEEE*, Honghan Chen, Tao Gan, Junzhou Chen, *Member, IEEE*, Chong-Wah Ngo, *Member, IEEE*, and Qiang Peng

***Abstract*—Wireless capsule endoscopy (WCE) has become a widely used diagnostic technique to examine inflammatory bowel diseases and disorders. As one of the most common human helminths, hookworm is a kind of small tubular structure with grayish white or pinkish semi-transparent body, which is with a number of 600 million people infection around the world. Automatic hookworm detection is a challenging task due to poor quality of images, presence of extraneous matters, complex structure of gastrointestinal, and diverse appearances in terms of color and texture. This is the first few works to comprehensively explore the automatic hookworm detection for WCE images. To capture the properties of hookworms, the multi scale dual matched filter is first applied to detect the location of tubular structure. Piecewise parallel region detection method is then proposed to identify the potential regions having hookworm bodies. To discriminate the unique visual features for different components of gastrointestinal, the histogram of average intensity is proposed to represent their properties. In order to deal with the problem of imbalance data, Rusboost is deployed to classify WCE images. Experiments on a diverse and large scale dataset with 440 *K* WCE images demonstrate that the proposed approach achieves a promising performance and outperforms the state-of-the-art methods. Moreover, the high sensitivity in detecting hookworms indicates the potential of our approach for future clinical application.**

***Index Terms*—Computer-aided detection, hookworm detection, pattern recognition and classification, wireless capsule endoscopy.**

I. Introduction

WIRELESS capsule endoscopy (WCE) is a disposable, pill-shaped device for gastrointestinal (GI) diagnosis, which was first introduced in year 2001 to identify sources of obscure small bowel bleeding [1]. Due to its advanced visualization capability, the minimal invasion to patients, and without the need for air insufflation and sedation, WCE has rapidly become an important and wide-spread diagnostic technique, which has been used for several inflammatory bowel diseases and disorders, including obscure gastrointestinal tract bleeding ([2]–[4]), polyp ([5]–[7]), uler ([3], [5], [8]), tumor ([9]), Crohn's disease ([10]), and so on. It is reported that over one million patients worldwide have been examined with WCE.

As a leading cause of maternal and child morbidity in the developing countries of the tropics and subtropics, hookworm has affected over 600 million individuals globally [11]. Hookworm infection seriously threatens human health, causing intestinal inflammation, progressive iron/protein-deficiency anemia, mucosa damage, and malnutrition of human [12]. Hookworm infection in pregnancy can cause retarded growth of the fetus, premature birth and a low birth weight. Hookworm in children can cause intellectual, cognitive and growth problems [12]. Although efforts have been extensively conducted to automatic detect different pathologies, few works [13] have been explored for hookworm detection with WCE, which is the focus of this work.

WCE is mainly composed of lens, an imaging sensor, light sources, batteries and a radio transmitter system, which provides the inner visualization of the entire gastrointestinal tract. After swallowed by the patient, the WCE travels along the GI track with the physical peristalsis. It drops into stomach along esophagus and then passes through pylorus, duodenum, small intestine and colon. Finally, it arrives at the rectum and excretes from the anus. The WCE captures two or more color images of GI track per second. Usually, the entire process will last for around 8 hours until the batteries exhaust. Therefore, on average, it will produce over 50,000 images for each patient. These images are compressed and transmitted to a portable data recorder attached to the patient"s waist by radio frequency. The image data are then downloaded into a computer workstation, from which the trained endoscopists will manually examine these images, often frame by frame, to analyze various diseases of patients and identify areas with abnormal conditions. It will generally take two or three hours to evaluate the images of one patient, which is a time-consuming and laborious process. Therefore, it is greatly desired that intelligent approaches can be designed and implemented to provide support for endoscopists.

Unfortunately, automatic hookworm detection is a challenging task due to poor quality of images, presence of extraneous matters, complex structure of GI, and diverse appearances in color and texture. Fig. 1 shows several representative WCE images having hookworms. First, WCE images

Manuscript received January 12, 2016; accepted January 31, 2016. This work was supported by the National Natural Science Foundation of China (Nos. 61373121, 61036008 and 61272290), Program for Sichuan Provincial Science Fund for Distinguished Young Scholars (No. 13QNJJ0149), Program for Sichuan Provincial Key Technology Research and Development (No. 2012FZ0004), and the Fundamental Research Funds for the Central Universities. *Asterisk indicates corresponding author.*

*X. Wu is with the School of Information Science and Technology, Southwest Jiaotong University, Chengdu, 610031 China. (e-mail: wuxiaohk@home.swjtu.edu.cn).

H. Chen, J. Chen, and Q. Peng are with the School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China (e-mail: jzchen@home.swjtu.edu.cn; qpeng@home.swjtu.edu.cn; honghanchen@hotmail.com).

T. Gan is with the Endoscopy Center of West China Hospital, Sichuan University, Chengdu 610041, China(email:gantao-1@163.com).

C.-W. Ngo is with the City University of Hong Kong, Kowloon, Hong Kong (e-mail:cscwngo@cityu.edu.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2016.2527736

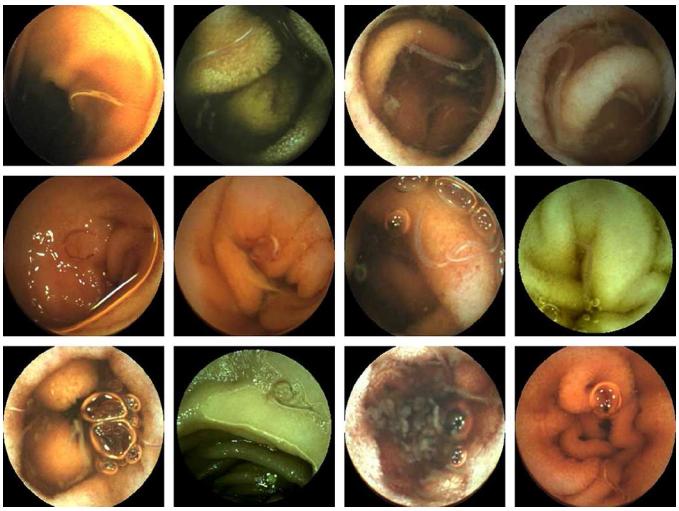

Fig. 1. Examples of WCE images having hookworms. Due to variable quality of images, presence of extraneous matters, complex structure of GI, and diverse appearances in terms of color and texture, automatic hookworm detection is a challenging task.

often suffer from noise and low visual quality because of complex circumstances in GI tract and poor illumination conditions. The quality of images is highly variable due to the uncontrolled peristalsis-driven motion of the capsule as moving through the GI tract. Second, there are many other contents in GI tract, such as food, stool, bile and bubbles, which influence the detection of parasitosis. Third, the free motion of the camera and the contractions that the gut undertakes produce various orientations and perspectives of the scene. Different parts of the intestinal tract (stomach, duodenum, jejunum-ileum, and cecum) present a variety of appearances with multiple colors and textures. Four, hookworms exhibit great variations in morphology and color. The hookworms attaching on mucosa demonstrate different shapes, widths and bend orientations. These challenges pose a great difficulty for automatic hookworm detection in WCE images.

To tackle the aforementioned challenges, a novel automatic hookworm detection approach is proposed for WCE images, which analyzes the morphology and geometrical characteristic of hookworms. The contributions of this work are as follow:

- To the best of our knowledge, this is the first few works to comprehensively explore the automatic hookworm detection for WCE images.
- The piecewise parallel region detection (PPRD) and the uncurled tubular region (UTR) are novelly proposed to detect the parallel regions and represent the extracted regions, respectively.
- To discriminate the unique features for different components of GI, such as hookworms, bubbles, and folds, the histogram of average intensity (HAI) is proposed to represent their properties.
- The experiments are performed on one of the largest datasets for WCE detection, which includes around 440 *K* WCE images of 11 patients. The results demonstrate that the proposed approach outperforms the state-of-the-art approaches.

This paper is organized as follows. Section II gives a brief overview of related work. Section III introduces the framework of the proposed hookworm detection. Sections IV – VI elaborate the detailed process of tubular region detection, parallel region detection, and pattern learning for hookworm detection, respectively. Section VII describes the experimental setup and performance evaluation. Finally, this paper is concluded with a summary.

II. Related Work

A. Pathological Abnormality Detection

In recent years, research on WCE becomes a hot topic. Computer aided detection systems based on different techniques have been extensively conducted, which bring endoscopists great convenience and efficiency. Comprehensive surveys on WCE research can be found in [14] and [15], which summarize the latest development from different aspects. The research on WCE covers a wide range of topics, including image enhancement [16], tissue segmentation or matching [3], [17], video segmentation [18], motility event detection [19], video summarization [20], and so on. A majority of the research on WCE focuses on pathological abnormality detection, which is the focus of this work.

The pathological abnormality detection research includes bleeding, polyp, ulcer, tumor, Crohn's disease, lymphangiectasia and other intestinal contents. The machine learning techniques frequently used for WCE lesion detection are Support Vector Machine (SVM) ([2], [5], [7]), Multi-Layer Perception neural network (MLP) ([4], [5]) and K Nearest Neighbor (KNN) ([21]). Among them, SVM is the most popular one, which has achieved good performance for many applications.

The detection of bleeding draws the most attention of researchers since bleeding is the main clinical pathology in GI tract. Superpixel is adopted in [2] to segment an image into small patches. The red ratio feature is proposed to classify patches as bleeding or not. The bleeding patches are then combined to form the final bleeding region. A color feature extraction method is presented in recent work [21] to discriminate the bleeding and normal frames. Each WCE frame is represented by words, where words are formed by clustering of color histograms. SVM and KNN are further applied for classifying and localization of bleeding regions.

A synergistic framework for automatic polyp and perforated ulcer detection is proposed in [5]. The images are first segmented to candidate regions. The geometric characteristics like curvature and eccentricity produce the final polyp candidates. The ulcer is classified with SVM based on color and texture features. Color texture feature is utilized for polyp detection and tumor detection [7], [9] using SVM, which integrates DWT (Discrete Wavelet Transform) and uniform LBP (Local Binary Pattern) to characterize WCE images. Polyps are detected in [6] based on the geometrical analysis and the texture content of WCE frames. The Fourier-Zernike moment features are computed and then the canonical discriminant analysis is used to automatically detect cancers [22]. The assessment of discrete

disease is explored in [10] to classify discrete lesions created by mucosal inflammation in Crohn's disease as well as quantitatively assess the lesion severity.

Although extensive works have been conducted on pathological abnormality detection for WCE images/videos, there are few works on automatic hookworm detection. In our early exploration [13], hybrid color gradient and contourlet transform are utilized to detect hookworms. However, this work is evaluated in a relative small and balanced dataset with 1,500 images, which is not applicable for practical scenarios with unbalance dataset. The characteristics of hookworms are quite different from bleeding, ulcers and polyps. The boundary and body of hookworms are also different from the patterns of existing pathologies. It remains unclear whether existing approaches for other lesion detection are also effective for hookworm detection, which motivates this research.

B. Features

Color and texture cues represented as global features are crucial for intestine content detection. In the cases of bleeding, bile and fecal detection, color cue is important because bleeding usually tends to be red while bile and fecal matters appear to be green. The color features include statistical features [4], moment [23], histogram [4], [21], red color ratio [2] or other descriptors [3], [10], [24]. The characteristics of polyp belong to the morphological and edge form. Edge and geometrical features are adopted in [5] to obtain polyp area and polyps are separated by curvature evaluation.

Texture features used for WCE pathological abnormality detection include LBP [7], [9], Gabor [5], Discreted Wavelet Transform [7], [9], versatile textures [5], MPEG-7 texture descriptors [10] and Contourlet transform [13]. Most aforementioned features are based on the combination of color and texture, showing good performance for the detection of mucosa lesions in intestine (e.g., ulcer or Crohn's disease) or other contents. In addition, spatial based texture features are extracted to detect ulcer in [3]. The round shape is exploited for polyp detection [5], [6], [25].

C. Machine Learning Techniques for Imbalanced Data

A critical factor that impacts the classification performance is the distribution of positive and negative samples in a dataset. For WCE images, the number of pathology samples is usually small, while there are a large number of normal images. Therefore, the data distribution is extremely imbalanced. Unfortunately, most machine learning approaches cannot deal well with highly imbalanced data. Traditional classifiers face the following two problems on imbalanced datasets [26]: (1) In order to achieve better accuracy, traditional approaches tend to favor classifying examples to majority class. Meanwhile, it will be ineffective at identifying examples of the minority class. (2) The performance of classifiers could be unreliable of operating on data where the distribution is different from training examples.

Techniques alleviating the problem of class imbalance are mainly divided into two categories: data sampling and boosting. Data sampling adjusts the imbalanced training data into balanced ones by under-sampling or over-sampling method. On the other hand, boosting repeatedly runs a weak learning algorithm on various distributions over the training data and produces several classifiers. These classifiers are then combined into a single composite classifier. Recently, two latest works are explored on imbalanced datasets [25], [27], which demonstrate that the ensemble machine learning methods perform better than SVM and classical ensemble learning methods. An effective feature learning method is proposed in [25] to detect the polyps in an imbalanced dataset. An imbalance machine learning method, Rusboost [26], has demonstrated excellent performance for classification on skewed data [28]. It has been successfully applied to reconstruction of cell lineages in biology [29], quantum measurement in physics [30], and mitosis detection of breast cancer [27]. In our work, the dataset is also highly imbalanced. Therefore, techniques for imbalanced data distribution are considered.

D. Datasets

The datasets for WCE pathological abnormality detection are relatively small, which are far from satisfactory to meet the requirement of real applications. Limited data samples will reduce the generalization ability of model training and affect the performance evaluation. Using a small number of sample frames can easily lead to over-tuning that may create an illusion of good performance, but these algorithms can easily fail in realistic conditions. For example, 533 images from 47 different patients are collected in [10]. For other works (e.g., [2]), a few hundreds or thousands of images from ten or twenty patients are used, which are rather small. The largest datasets so far are [3] and [24]. Although about 50 videos are used for evaluation in [3], there are only a small number of images (around 600 images). The main purpose of [24] is to segment intestine content, in which 95,148 images are used including commonly appeared bubble and turbid that are easier to obtain. In this paper, our dataset includes 11 patients with around 0.4 million WCE images. To the best of our knowledge, it is one of the largest datasets for automatic lesion detection for WCE images.

III. Framework

A. Properties of Hookworms

The characteristics of hookworms are quite different from bleeding, ulcer and polyps. The hookworm is a kind of small tubular structure with different intensities from mucosa and bubble edges, which has two major peculiarities.

1) Tubular structure: The hookworms have obvious boundary than other tubular structures. The edges of hookworm bodies usually in the form of nearly parallel curves.
2) Intensity property: The color of hookworms is usually grayish white or pinkish semi-transparent. Sometimes the hookworms have reddish or darker color compared to surrounding mucosa.

Motivated by these two properties of hookworms, we propose a novel framework for automatic hookworm detection.

B. Framework

The framework is illustrated in Fig. 2. Due to the reasons like low image resolution, movement of WCE, complex circumstance of GI tract and limited capability of lens [16], WCE

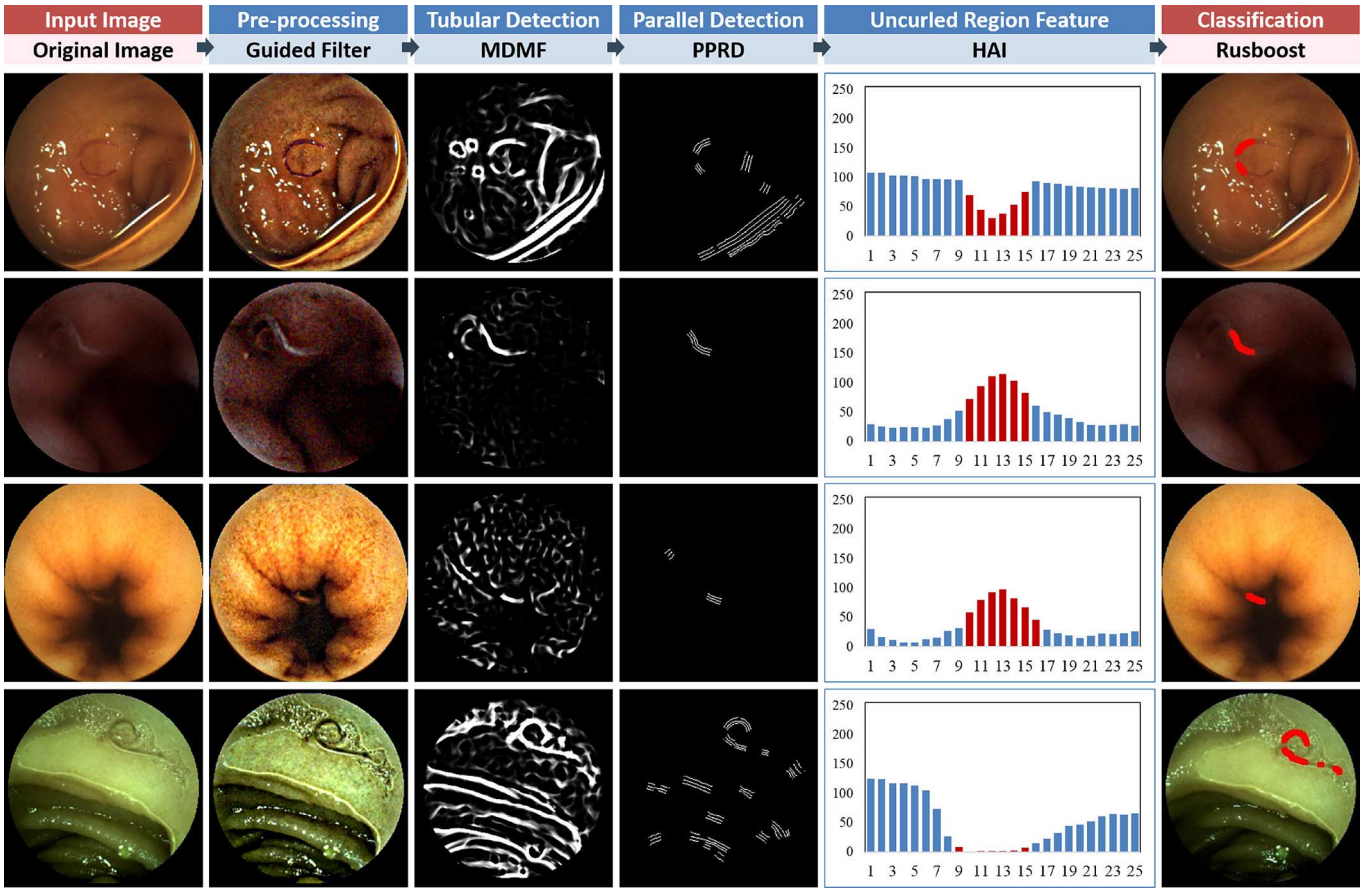

Fig. 2. The proposed framework of automatic hookworm detection and the detailed results for some representative examples.

images are usually with low contrast and the overall scenes are relatively dark, which are difficult to be recognized. These reasons significantly affect the performance of computer-aided diagnosis system. To have clear details and conspicuous texture information, the guided filter [31] is first adopted for WCE image enhancement due to its good performance and efficiency. Since the hookworm bodies demonstrate either lighter or darker patterns compared to surrounding mucosa, meanwhile, hookworms have various orientations and inconsistent widths, Multi-scale Dual Matched Filter (MDMF) is proposed to detect the tubular regions in WCE images. Based on the theory that a nonlinear system can be modeled by multiple piecewise linearization systems, Piecewise Parallel Region Detection (PPRD) is novelly proposed to detect the parallel edges. The edges have different lengths and appear with irregular shapes and diverse blending orientations. We rectify the parallel edges by expanding and stretching the curved regions into flat and regular regions, so as to ease feature extraction and pattern learning for later stages. The histogram of average intensity (HAI) is then proposed to represent the intensity patterns of different objects, including intensity contrast, intensity shape and central width. In order to deal with the problem of imbalance data, machine learning algorithm Rusboost is deployed to classify WCE images, which will output a decision for every image whether it contains hookworms or not. In the following sections, we will elaborate the detailed steps for automatic hookworm detection.

IV. Tubular Region Detection With Multi-Scale Dual Matched Filter

In medical domain, the extraction of tubular structure has been used in many applications, such as blood vessel, coronary artery, human pulmonary tree and abdominal aorta. Although the tubular structure of hookworms is similar to blood vessels, detection is much harder. First, hookworms have two kinds of patterns. The intensity of hookworms is either brighter or darker than the mucosa. Second, the hookworms and their different parts have diverse orientations. Third, the width of hookworms is inconsistent. To tackle these problems, the *Multi-scale Dual Matched Filter (MDMF)* is proposed to detect the tubular regions in WCE images.

A. Dual Matched Filter

Among different methods for tubular detection, the *matched filter* [32] is an effective approach, which is originally proposed to detect retinal vessels. The matched filter is a Gaussian-shaped template, which is based on prior information that the cross-section of a vessel is Gaussian-shaped. Since the vessels appear darker than other retinal surfaces, the tubular region will produce a higher response when the matched filter convolves with the image.

The intensity of hookworms is either brighter or darker than the mucosa, as shown in Fig. 3(a) and (d), respectively. To capture two different patterns of hookworms, the *Dual*

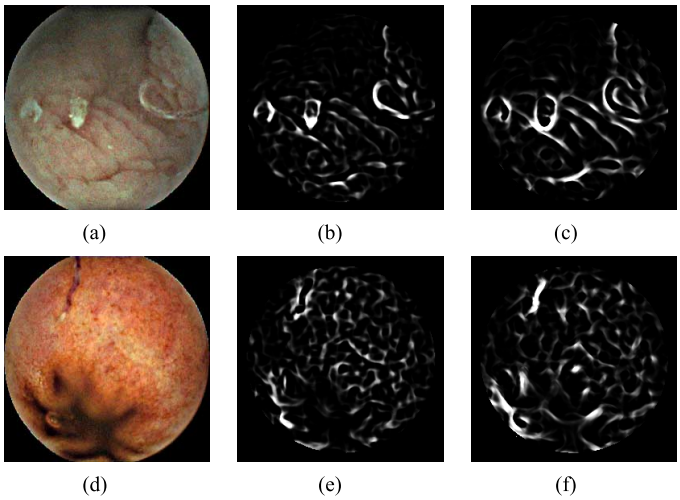

Fig. 3. Two patterns of hookworms and the corresponding DMF response results. The intensity of hookworms is either brighter (a) or darker (d) than the mucosa. DMF shows response in (b) for brighter hookworm and in (f) for darker hookworm. The classical matched filter for vessel detection considers only DMF $(-)$ response and hence fails in detecting brighter hookworms. (a) Enhanced image A. (b) DMF $(+)$ response. (c) DMF $(-)$ response. (d) Enhanced image B. (e) DMF $(+)$ response. (f) DMF $(-)$ response.

Matched Filter (DMF) kernel is proposed, which is defined as follows,

$$K_{\sigma}^{\pm}(x,y) = \pm e^{-(x^2/2\sigma^2)}, \quad \forall\, |y| \le \frac{L}{2} \tag{1}$$

where $K_{\sigma}^{\pm}(x,y)$ is the DMF kernel at a scale of σ. The symbol $\pm$ denotes two different hookworm patterns. The term $\exp(-x^2/2\sigma^2)$ denotes that the cross-section (x-axis direction) of a hookworm is Gaussian-shaped. From the perspective of two-dimensional space, the Gaussian term is repeated along y-axis with length L, which is set to 9. The two patterns of hookworms and their corresponding DMF responses are shown in Fig. 3. The matched filter response of a brighter hookworm, i.e., DMF $(+)$ exists in Fig. 3(b), while the darker hookworm response, i.e., DMF $(-)$ exists in Fig. 3(f). The classical matched filter for vessel detection only obtains the response results of Fig. 3(c) and (f), and the brighter hookworm cannot be detected. With the dual matched filter, it can detect both bright and dark hookworms.

B. Multi-Scale Dual Matched Filter

In order to adapt to different orientations, the kernel has to be rotated to all possible orientations, and a set of filter banks is obtained. The implementation of the filters is done by multiplying 0° filter with a rotation matrix, which is defined as follows,

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & tx \\ -\sin\theta & \cos\theta & ty \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{2}$$

where (x,y) is a point in the 0° filter. The point (x',y') is the rotated point around the pivot point (tx,ty) with degree θ. Here, the pivot point (tx,ty) is set to coordinate origin $(0,0)$. The filter bank is normalized to zero mean after rotation. Similar to [32], the number of filters is set to 12. The final response is the one with the maximum convolution from 12 filters in the filter bank.

In order to adapt the filter bank to different widths, multi-scale production with the filter bank is deployed to detect hookworms with different widths [33]. Let $f(x,y)$ be the image and $m_i(x,y)$ be the matched filter at scale σ_i. The response at scale σ_i can be expressed as

$$r_i(x,y) = m_i(x,y) * f(x,y) \tag{3}$$

where $*$ denotes the convolution operation. Assume that there is another response $r_j(x,y)$ at scale σ_j, the response production at scales σ_i and σ_j is defined as

$$P_{i,j}(x,y) = r_i(x,y) \cdot r_j(x,y). \tag{4}$$

To detect the tubular regions, the final response image of hookworms is the production of the widest, the middle and the thinnest scale responses. In our dataset, the width of the largest hookworm is less than 12 pixels. Therefore, the largest scale is empirically set as $\sigma = 2$ in this paper. Since the width of the thinnest hookworm is larger than 5 pixels, as a result, the smallest scale is set to 1.2. The appropriate value of middle scale is 1.6. Fortunately, with the production of multiple scales, the hookworm response is preserved, which highlights the potential regions of hookworms.

The final step for tubular structure detection is to get the binary regions from the response image. In order to determine the tubular area effectively, the Otsu's method is adopted, which can adaptively distinguish tubular structure and mucosa responses well.

V. Parallel Region Detection

After tubular region detection using multi-scale dual matched filter, the detection results contain potential tubular regions of hookworms as well as some non-tubular regions, such as bubbles and intestine folds, which have similar structure as hookworms. The edges of hookworm bodies usually demonstrate parallel shapes, which is a useful cue to distinguish tubular structure and non-tubular one. Therefore, *Piecewise Parallel Region Detection (PPRD)* is proposed to detect the parallel edges.

A. Piecewise Parallel Region Detection

The proposed method is based on the theory that a nonlinear system can be modeled by multiple piecewise linearization systems. Every small part of the edges can be seemed as straight lines (linear lines). Therefore, the whole edges are the combination of small straight lines.

First, for each binary tubular region, the Canny detector is adopted to detect the edges of tubular structure in the gray image. Short edges with small gaps are connected together to form a long edge by performing edge linking. The isolated dots are treated as noises and removed.

Second, piecewise linear detection method is proposed to detect parallel curves from edge pairs. As shown in Fig. 4, there are two parallel curves in a tubular region, $Curve_A$ and $Curve_B$, which are labeled in green and red, respectively. $Curve_A$ is treated as the reference curve, and the paralleled edge parts in $Curve_B$ will be detected. A *sliding window* moving along

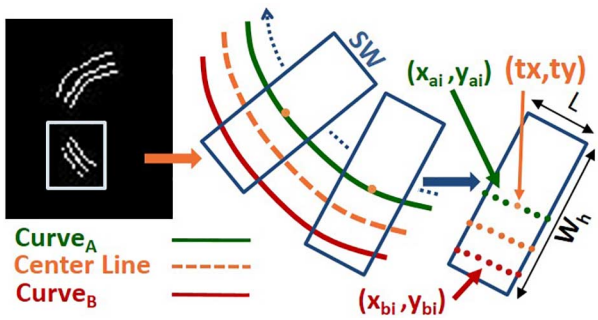

Fig. 4. Piecewise parallel region detection.

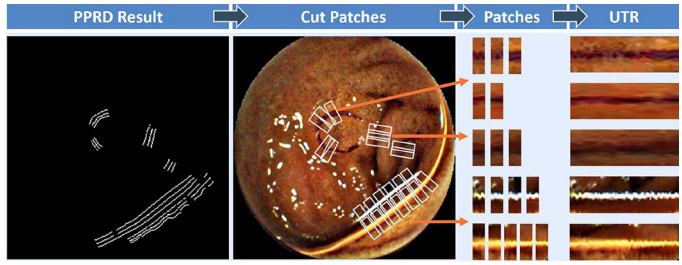

Fig. 5. The sliding window is moved along the central line, and the rectangle area is rotated based on its tangential direction. These rotated patches are catenated to form a flat tubular region.

$Curve_A$ is defined. The position of sliding window is determined by a central point (tx, ty). In this paper, the length of the sliding window is the same as the length of DMF (i.e., $L = 9$). Since the maximum width of a hookworm is 12, the width of sliding window is set to $W_h = 25$, which can detect parallel lines on both sides of the reference curve.

Let $(X_i, Y_i) = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ be a serial of points of $Curve_A$ within the sliding window. Assuming (X_i, Y_i) fall on a straight line, the point set matches the linear function (5) as follows,

$$y_i = a + bx_i \tag{5}$$

where a is a constant value and b is a slope parameter, which are unknown. The covariance matrix of X_i and Y_i can be calculated by (6):

$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n}(X_i - \bar{X})^2 & \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) \\ \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) & \sum_{i=1}^{n}(Y_i - \bar{Y})^2 \end{bmatrix} \tag{6}$$

where $\bar{X}$ and $\bar{Y}$ are mean values of X_i and Y_i, respectively. In order to determine whether the point set (X_i, Y_i) belongs to a straight line, the famous *Pearson product-moment correlation coefficient* [34] in statistics is employed as shown in (7).

$$r = \frac{\sigma_{xy}}{\sqrt{\sigma_{xx}}\sqrt{\sigma_{yy}}}, \quad \forall \sigma_{xx} \neq 0, \forall \sigma_{yy} \neq 0 \tag{7}$$

where the range of value r is $0 \leq |r| \leq 1$. When the point set falls exactly on a straight line, the strongest correlation occurs (i.e., $|r| = 1$). On the contrary, when the points fall in a random pattern, $|r|$ approaches to 0. Here, if $|r|$ is above a given threshold, say 0.9, we believe these points form a straight line, and they will be kept. Otherwise, these points will be discarded. Equation (5) can be converted to the following form by linear regression.

$$\hat{y} = \bar{y} + \hat{b}(x_i - \bar{x}) \tag{8}$$

where $\bar{x}$ and $\bar{y}$ are means of x_i and y_i, respectively. $\hat{y}$ is the regression value of corresponding x_i. $\hat{b}$ is the slope parameter and it can be solved by (9). The angle θ of the straight line formed by (X_i, Y_i) is solved by (10).

$$\hat{b} = \frac{\sigma_{xy}}{\sigma_{xx}} \tag{9}$$

$$\theta = arctan(\hat{b}) * 57.3. \tag{10}$$

Given the sliding window with angle 0°, the sliding window with θ around (tx, ty) can be calculated by (2). Let θ_i be the angle of reference curve falling into sliding window. The angle of the paralleled curve can be calculated in the same way, which is denoted as θ_j. Therefore, the difference between θ_i and θ_j is the criteria to measure the parallel relationship. If the difference is less than 15 degrees, i.e., $|\theta_i - \theta_j| < 15°$, the pair of lines will be treated as parallel. Some deviation is tolerable because the line pair is not strictly parallel. Since the number of filters in DMF bank is set to 12, the filter rotated within amount of 15° is tolerable. In this step, parallel curves are preserved. Non-parallel curves and parallel curves beyond the width of sliding window are removed.

Finally, the center line of parallel curves is extracted. Let (x_{ai}, y_{ai}) be a point of the reference curve. According to the geometric theory, the normal of (x_{ai}, y_{ai}) will intersect with a point (x_{bi}, y_{bi}) in the parallel curve. Hence, (x_{bi}, y_{bi}) in the parallel curve is the nearest point of (x_{ai}, y_{ai}). Therefore, the center line is a serial of center points between (x_{ai}, y_{ai}) and (x_{bi}, y_{bi}), which can be seen from Fig. 4. The small gap of the serial points are finally linked. With the aforementioned processing, the tubular structure with parallel edges is detected.

B. Uncurled Tubular Region

Since the hookworms attached on mucosa demonstrate different shapes, inconsistent lengths and diverse bending orientations, the detected tubular regions are usually curves with irregular shapes, which are not beneficial to discriminate hookworm and non-hookworm regions. In order to represent all candidate regions in the consistent format, therefore, it is necessary to expand and stretch the curved regions into flat and regular regions. The rectification can also facilitate the feature extraction and pattern learning for later stages.

The process of uncurled tubular regions is illustrated in Fig. 5. The sliding window is moved along the central line of the region, and the rectangle area is rotated with angle θ_c, which is calculated according to (2). Precisely, it is rotated based on its tangential direction. Then the rectangle area is mapped back to the original WCE image and the corresponding rectangle patch

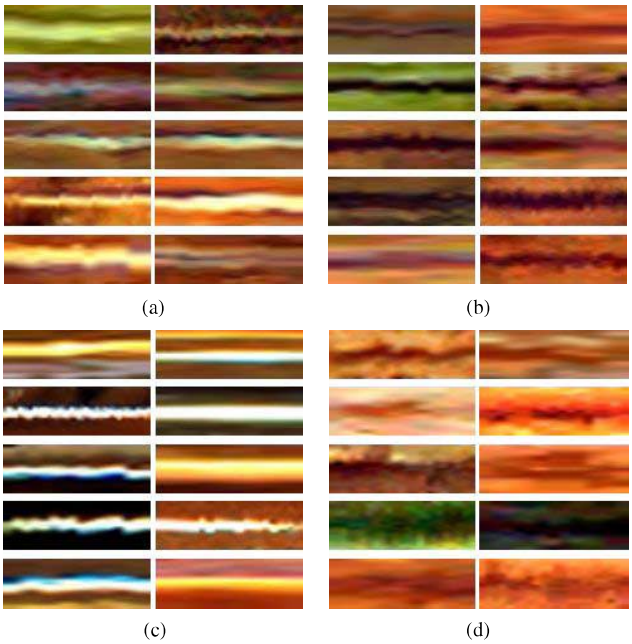

Fig. 6. The results of uncurled tubular regions for different types of objects, (a) bright hookworms, (b) dark hookworms, (c) bubbles, (d) folds. We can notice that their uncurled tubular regions are quite different.

is extracted from the color image. The patch includes the tubular regions as well as the surrounding mucosa texture. For each central line, serval patches will be extracted, and each patch is rotated to 0° so that it is arranged to horizontal direction. Finally, these rotated patches are catenated to form a flat tubular region. Since the sizes of tubular regions are inconsistent, the lengths of catenated regions are different. In order to represent parallel regions of different shapes, inconsistent lengths and diverse bending orientations with a unified format, these catenated regions are then normalized to a fixed size with 75×25. With this uniform structure, it becomes easier to represent the tubular regions for further processing.

With this step, the uncurled tubular regions consist of three parts, the central part and both sides surrounding it. The central part corresponds to the detected parallel regions, which may be hookworms, bubble edges, folds and so on. Both sides around the central region are the surrounding environment, such as mucosa. The results of uncurled tubular regions for different types of objects are demonstrated in Fig. 6.

VI. Pattern Learning for Hookworm Detection

A. Pattern Learning With Histogram of Average Intensity

From Fig. 6, we can notice that the uncurled tubular regions for hookworms, bubble edges and folds are quite different. Hookworms exhibit great variations in morphology and color. The color of hookworms is usually grayish white or pinkish semi-transparent. Sometimes the hookworms have reddish or darker color. Therefore, we classify the hookworms into two categories, bright hookworms and dark hookworms, which are illustrated in Fig. 6(a) and (b), respectively. Since different parts of the intestinal tract (e.g., duodenum, jejunum-ileum, and cecum) present a variety of appearances with multiple textures and colors, the surrounding background is quite diverse. The intensities of bubble edges and bright hookworms are brighter compared to the surrounding context, while the intensities of folds and dark hookworms are gloomy. Due to the reflection to light, the intensity of bubbles is much higher than bright hookworms, which can be seen from Fig. 6(c). In addition, the intensities of mucosa on both sides of hookworms have minor difference. But for the case of folds, the region of central part is not so conspicuous, and the difference between central part and two sides are relatively minor, which can be seen from Fig. 6(d). We can see that the patterns for different objects are quite unique.

As aforementioned, the body of a hookworm is a kind of tubular structure, in which the edges of a hookworm body are nearly parallel. Since the hookworm is attached to the intestine, the surrounding mucosa must be considered. From the viewpoint of intensity, the body of a hookworm is either brighter or darker than mucosa. Although the intensity of bubbles is also brighter than the mucosa, it is much higher than hookworms. The cross-section of hookworms is Gaussian-shaped (i.e., bell shape) and symmetric. Meanwhile, the width of hookworms is within a limited range, while most folds are wider than the hookworms. Based on these observations, the histogram of average intensity (HAI), which is the average intensity for corresponding rows of uncurled tubular regions, is proposed to represent the properties of intensity contrast, intensity shape and central width. The central part is the width of tubular regions, such as hookworms, bubbles or folds. The transition between the bins of central part and both sides provide the intensity contrast between tubular regions and surrounding mucosa. The average of intensity can eliminate the noises. If the edges of tubular regions are not parallel, the boundary of central part becomes vague and is not easy to be discerned. This information can help to remove folds with blurred edges.

The histogram of average intensity $\mathcal{H} = \{H_1, H_2, \ldots, H_j, \ldots, H_R\}$ is defined as follows,

$$H_j = \frac{1}{C} \sum_{i=1}^{C} I(i,j) \tag{11}$$

where $I(i,j)$ is the intensity value of point (i,j). The intensity of pixels in each row is averaged. Since the uncurled tubular regions has been normalized to 75×25, R is set to 25 and C is set to 75. Therefore, the histogram $\mathcal{H}$ is a 25 dimensional vector.

Four representative uncurled tubular regions and their corresponding histograms of average intensity are demonstrated in Fig. 7. As can be seen from Fig. 7(a), the histogram of average intensity for bright hookworms is symmetric bell curve shape. Meanwhile, the shape of histogram for dark hookworms is reverse bell shaped curve, which is illustrated in Fig. 7(b). The histogram for bubble edges is pretty similar to the one of bright hookworms, but its overall values are much higher than the bright hookworms. The peak value of intensity can be as high as 255. It can be seen from Fig. 7(c). In general, the histogram for folds is relatively flat and the boundary between the central part and surrounding background is vague, which is shown in Fig. 7(d). The central part is labeled in red and both sides are marked in blue, which correspond to the tubular objects and surrounding context, respectively. The region of central part is

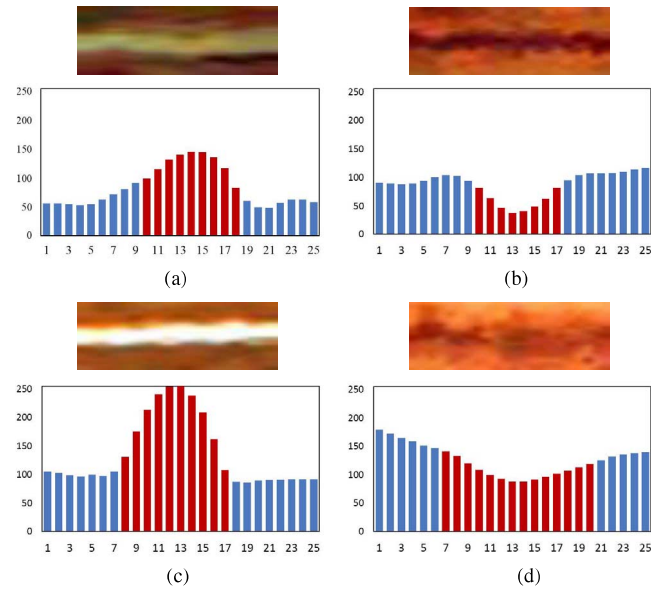

Fig. 7. The uncurled tubular regions and their corresponding histograms of intensity: (a) bright hookworm, (b) dark hookworm, (c) bubble, (d) fold.

TABLE I
DATASET INFORMATION AND OVERALL PERFORMANCE (%)

Patient	Dataset		Parallel Region		Classification	
ID	# HW	# Non-HW	HW	Non-HW	HW	Non-HW
1	41	41,571	85.4	83.8	78.1	18.5
2	55	43,539	67.3	84.2	61.8	20.3
3	13	34,956	69.2	89.0	69.2	28.4
4	70	45,058	80.0	89.7	55.7	7.3
5	110	12,850	58.2	85.6	53.6	12.7
6	43	42,462	65.1	98.8	60.5	26.9
7	32	53,330	90.6	67.7	78.1	5.8
8	1,770	35,573	93.8	98.5	93.5	44.8
9	2,549	37,017	92.8	87.4	92.6	25.1
10	42	44,066	66.7	88.6	59.5	24.3
11	103	46,374	72.8	89.2	69.9	27.7
Total	4,828	436,796	90.9	87.0	89.8	21.7

adaptively detected. Due to space limitation, we will not elaborate it.

B. Hookworm Classification With Rusboost

With the histograms of average intensity, it is easier to discern different types of objects. The machine learning classifier is adopted to distinguish hookworms from folds and bubbles. As mentioned before, the distribution of WCE images is pretty skewed. There are a huge number of negative samples while only a small portion of positive cases, as seen in Table I. More than 99% of images are non-hookworm, and hookworms take up less than 1%. In this work, we leverage the *Rusboost* [26] to deal with the classification problem on skewed data. The *Rusboost* combines *random under-sampling (RUS)* with *Adaboost*. RUS removes examples stochastically from the majority class until the desired balance is achieved.

In Rusboost, a training set of m examples $S = \{(x_1, y_1), \ldots, (x_m, y_m)\}$ is taken as input, where x_i is an instance in the feature space X, and $y_i \in Y$ is the class label corresponding to x_i. Rusboost repeatedly runs weak learners with T iterations. For each iteration, the majority class examples are randomly removed until the numbers of majority and minority class examples are balanced. Therefore, a new temporary training set S_t with distribution D_t is created. The weak learners compute a classifier or hypothesis $h_t(x_i, y_i) : X \rightarrow Y$ by minimizing the error ϵ_t in (12).

$$\epsilon_t = \sum_{(i,y):y_i \neq y} D_t(i)(1 - h_t(x_i, y_i) + h_t(x_i, y)). \tag{12}$$

Here, each weak leaner is a tree classifier fitted by S_t. The fitting process will grow a deep tree.

In the next iteration, the weight distribution D_{t+1} is updated according to (13) and normalized as (14).

$$D_{t+1}(i) = D_t(i)\alpha_t^{1/2(1+h_t(x_i,y_i)-h_t(x_i,y:y\neq y_i))} \tag{13}$$

$$D_{t+1}(i) = \frac{D_{t+1}(i)}{\sum_i D_{t+1}(i)} \tag{14}$$

where the updated parameter α_t is calculated as $\epsilon_t/(1-\epsilon_t)$. The training process continues for T iterations, and at last T weak hypotheses $h_1, \ldots, h_T$ are combined into a single hypothesis $H(x)$.

$$H(x) = \underset{y \in Y}{\operatorname{argmax}} \sum_{t=1}^{T} h_t(x, y) \log \frac{1}{\alpha_t}. \tag{15}$$

The histogram of average intensity is fed into Rusboost for binary classification, which will output the decision of either presence or absence detection. Note that there could be more than one uncurled tubular region detected in one image and each uncurled tubular region will be represented as one histogram of average intensity. Therefore, multiple histograms of average intensity will be obtained for one image and they will be classified by Rusboost separately. In order to guarantee that more potential hookworm images can be detected, we adopt a simple strategy to determine the final result. Only if any uncurled tubular region of this image is classified as having hookworm, the whole image will be treated as the hookworm image. This strategy can guarantee the sensitivity of hookworm detection.

VII. EXPERIMENTS

A. Dataset

To evaluate the performance of the proposed automatic hookworm detection, WCE images of 11 patients aged 14 – 74 years old are collected from the West China Hospital as the dataset. Each WCE image is a 24 bit RGB color image with the resolution of 256×256 pixels. All images are collected by the OMOM® WCE, which has been widely used in hospitals in China and other countries. Since the hookworms usually live in the upper small intestine, the images from pylorus to the end of videos are included in the experiments. Totally, there are around 440 *K* WCE images in the dataset, and on average each patient has around 40 *K* images. To the best of our knowledge, this dataset is one of the largest datasets for automatic disease detection for WCE images. Unlike other works that select only a few hundreds to thousands of WCE images for evaluation, in this work, the whole dataset is used. Among the eleven patients,

the eighth and ninth patients are heavily infected. The fifth patient has only 12,850 non-hookworm images, because the WCE stayed in his stomach about four hours and cannot enter into the duodenum. The detailed information of the dataset is listed in Table I.

In order to be consistent with the realistic condition and to evaluate the performance of different approaches applied to unseen patients, the experiments are carried out in 11-fold leave-one-patient-out cross-validation. Each time, the images of an "unseen" patient is treated as the testing set and the data of remaining ten patients are used as the training set. After eleven rounds of experiments, the results are averaged as the final performance score. To evaluate the performance, two experienced endoscopists examine and label all images as the ground truth. The images containing hookworms are labeled as positive samples, and negative otherwise.

B. Evaluation Criteria

To evaluate the performance of classification, *Sensitivity (SE)*, *Specificity (SP)* and *Accuracy (AC)* are adopted as the performance metric, which are widely used in medical image classification tasks (e.g., [2], [5], [9], [19], [24]). Ideally, we expect both sensitivity and specificity to be as high as possible. Since accuracy reflects sensitivity and specificity in relation to each other, it is also used to assess the overall performance of classification. High sensitivity means strong capability of detecting hookworm images. Due to the fact that an overwhelming majority of WCE images are non-hookworm, the detected "hookworm" images have to be manually inspected by a doctor. Thus, the specificity should be high to minimize the workload of the doctor. Overall, sensitivity is more important than specificity and accuracy. They are defined as follows,

$$\begin{aligned} Accuracy &= \frac{(TP+TN)}{N} \\ Sensitivity &= \frac{TP}{(TP+FN)} \\ Specificity &= \frac{TN}{(TN+FP)} \end{aligned} \tag{16}$$

where TP is the number of correctly classified images containing hookworms. FN is the number of hookworm images that are falsely classified as non-hookworm images. FP is the number of non-hookworm images that are falsely classified as hookworms. TN is the number of non-hookworm images that are correctly labeled as non-hookworms, and N is the total number of images.

C. Overall Performance

The overall performance is listed in Table I. We can see that 90.9% of hookworm images and 87% of non-hookworms images have parallel regions detected. Because there are a large number of WCE images containing intestinal folds and bubbles, most of them share the same characteristics of parallel edges of hookworms. Around 13% non-hookworm images are filtered by piecewise parallel region detection. After tubular region uncurling, pattern learning with histogram of average intensity, and classification using Rusboost, 89.8% of hookworm images are correctly detected, and 21.7% of non-hookworms are falsely detected as hookworms. Only 1% of hookworm images having parallel regions are treated as non-hookworms and falsely removed, which demonstrates the robustness of the proposed method. With the proposed approach, 65% of non-hookworm images having parallel regions are correctly filtered out, which significantly reduces the workload for endoscopists to evaluate.

TABLE II
PERFORMANCE COMPARISON OF SVM VS RUSBOOST FOR IMBALANCE DATA

Patient	# UTR		SVM			Rusboost		
ID	HW	Non-HW	AC	SE	SP	AC	SE	SP
1	69	153,266	99.1	4.4	99.1	86.4	37.7	86.4
2	53	170,423	99.4	17.0	99.5	90.7	79.3	90.7
3	16	212,292	98.6	6.3	98.7	83.1	56.3	83.1
4	139	143,734	99.4	15.1	99.5	92.2	48.9	92.3
5	185	42,265	98.3	16.8	98.6	89.4	69.2	89.5
6	58	236,449	99.4	17.2	99.4	91.6	55.2	91.7
7	51	121,144	98.7	3.1	98.7	85.2	51.0	85.2
8	2,838	231,321	97.6	8.1	98.7	83.5	67.0	84.0
9	7,309	165,589	95.3	7.8	99.1	84.4	60.0	85.5
10	41	250,716	98.1	19.5	98.1	85.3	65.9	85.3
11	167	202,331	99.0	6.6	99.1	88.9	50.9	89.0
Average			98.4	11.1	99.0	87.3	58.3	87.5

Some representative examples with detailed steps for automatic hookworm detection are illustrated in Fig. 2, from which we can see that the hookworm detection is pretty challenging. The image in the first row contains a hookworm, a bubble and folds. The images in the second and third rows have low contrast, in which the hookworms are difficult to be discerned from the original images. The overall color distribution of the fourth image is totaly different from other images. With the combination of proposed steps, hookworms can be correctly detected, which demonstrates the effectiveness and robustness of the proposed method.

D. Performance Comparison for SVM vs Rusboost Classification

For automatic hookworm detection, the data distribution is highly imbalanced. There are few positive hookworms and a huge amount of negative samples. Since most stat-of-the-art methods adopt SVM as the classifier, we will compare the performance of SVM and Rusboost under this imbalance dataset, which is listed in Table II. For a WCE image, several parallel regions will be detected, and each region is normalized to an uncurled tubular region (UTR). The total number of uncurled tubular regions for hookworm and non-hookworm images is listed Table II. In this experiment, the number of negative images we sampled for SVM training is 10 times the number of positives. Based on our observation, the performance of SVM classification becomes worse as a larger number of negative images are included. For Rusboost, we can feed all samples into the classifier. Rusboost can automatic down sample the negative images. It is easy to find that SVM achieves pretty high accuracy and specificity, which is the reason why SVM is popularly adopted by various tasks. Unfortunately, the sensitivity of SVM is poor when facing the imbalance data, which is only 11.1%. With the Rusboost, although the performance of accuracy and specificity drops slightly (around 10% reduction), the

TABLE III
PERFORMANCE COMPARISON FOR DIFFERENT FEATURES ON UNCURLED TUBULAR REGIONS

Methods	SVM			Rusboost			
	AC	SE	SP	AC	SE	SP	p(SE)
LBP_Gray	99.4	3.5	99.8	67.2 *	58.1	67.3 *	0.106
LBP_RGB	96.8	12.9	97.3	81.8 *	42.7	82.2 *	0.050
LBP_HSV	96.0	20.6	96.3	86.7 *	41.6	86.9 *	0.055
HOG	94.4	16.9	94.6	85.2 *	64.9	85.4 *	0.142
HAI_HSV	95.5	30.8	95.7	95.3 *	38.3	95.3 *	0.001
HAI	98.4	11.1	99.0	87.3	58.3	87.5	—

* stands for the proposed method (HAI) is statistically significant compared to the given method when $p<0.01$.

sensitively is significantly improved from 11.1% to 58.3%. It achieves a balance for all performance metrics.

E. Performance Comparison on Feature Extraction

In this work, histogram of average intensity (HAI) is proposed to capture the properties of uncurled tubular regions. We also evaluate the performance with other classic features like Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG), which have been widely used for many applications. The performance comparison on uncurled tubular regions is listed in Table III. Uniform LBP features on gray and color images of RGB and HSV color spaces are tested, respectively. The HOG feature is extracted on uncurled tubular regions with the cell size of 8 and the number of orientations of 9. The final feature is an 837 dimensional vector. To verify the effect of color information for the performance, we also test the HAI feature on HSV color space, which is extracted from three channels and combined into a 75 dimensional vector.

From Table III, we can see that the performance of LBP features is worse than HAI and HOG. LBP calculates the contrast pattern of each pixel with its neighboring pixels and forms the histogram of the whole image. Although LBP can describe the texture well, it loses the shape and spatial information. In other words, it cannot describe the width of the central part of uncurled tubular regions, and the information of intensity contrast between the central part and both sides is totally lost. Therefore, the performance of LBP is not satisfactory. For HOG, the appearance and shape of local objects can be characterized well by the distribution of local intensity gradients or edge direction. The whole image is divided into small cell regions, and a local histogram of gradient orientations for each region is then accumulated. These histograms are then combined together to form final features. As HOG describes the distribution of intensity gradients or edge directions well, the sensitivity is better than HAI. However, HOG with Rusboost model is hard to train, with high cost and low efficiency. The training process will cost 20 *GB* memory and spend around 70 hours per patient, which is prohibitively expensive for implementation and is unsuitable for the scenario of hookworm detection. On the contrary, HAI can be finished within 3 hours with at most 5 *GB* memory. The proposed method is a reasonable choice in terms of cost, efficiency and effectiveness.

When the color information is taken into account, we notice that accuracy and specificity are improved while sensitivity is dropped for LBP and HAI. Folds usually demonstrate similar dark colors (brown or black) as dark hookworms. Meanwhile, the edges of bubbles have white color like bright hookworms. It makes the discriminative capability of color not so prominent. More importantly, the body of hookworms is a little semi-transparent. Its color usually changes according to surrounding mucosa. The color of mucosa for different patients varies, and even different parts of a patient are different. Certain intensity information originally in gray scale is lost when a color image is separated into color channels. When the color information is integrated, a large portion of non-hookworm images will be correctly detected due to high color similarity. Therefore, accuracy and specificity are improved. At the same time, a small portion of hookworms are falsely filtered out and treated as non-hookworms, causing the drop of sensitivity. Although color is a critical feature for WCE pathological abnormality detection, such as bleeding and ulcer, it is not so useful for hookworm detection.

In order to evaluate the significance of the proposed feature compared to other traditional features, the paired t-test is conducted, which is listed in Table III. Since the number of non-hookworm samples is very large, the performance variations for different methods are pretty conspicuous. Therefore, the results of HAI compared to other methods are significant in terms of accuracy and specificity, which are marked with $*$ in Table III. For sensitivity test, the samples that are correctly detected as hookworms are labeled as 1, otherwise they are labeled as 0. We can see that the proposed feature is significant compared to HAI_HSV and Color LBP. However, the performance of the proposed feature is more or less similar to HOG, which is not statistically significant.

F. Performance Comparison With State-of-the-Art Approaches

In this section, we will compare the performance of the proposed method with the state-of-the-art approaches in WCE abnormality detection. Since this work is the first few works on hookworm detection for WCE images, there are limited works in the literature available for comparison. Therefore, we resort to the state-of-the-art approaches used for other abnormality detection in WCE images, and compare the proposed method with them. In [7], [9], a kind of color-texture is proposed for polyp and tumor detection and good performance has been achieved. Each color channel of a WCE image is decomposed by DWT. The texture histograms are generated by LBP, which are then classified by SVM. In our prior work [13], the hybrid color gradient (HCG) is proposed for hookworm detection, which has been proved its effectiveness. Four useful channels and corresponding gradient maps are calculated by oriented energy filters, which are then decomposed by Contourlet. A group of statistical values is fed into MLP to classify WCE images.

Previous researches have demonstrated that features, color models and classification methods will affect the performance greatly. In this work, the features we selected for comparison are color texture feature (DWT+LBP) and hybrid color gradient (HCG). The color models are RGB, HSV and CIE Lab. The popular classification tools include: SVM, KNN and MLP, which have been widely adopted in previous WCE image classification. Because there are many combinations, it is difficult to show all experimental results. We only list the results whose accuracy,

TABLE IV
PERFORMANCE COMPARISON WITH STATE-OF-THE-ART APPROACHES

Methods[1]	AC (%)	SE (%)	SP (%)	p(SE)	Missed
1. SVM-HSV-CT [7], [9]	50.9 *	53.4	51.2 *	0.005	1,3,5
2. SVM-RGB-CT	53.8 *	51.9	54.2 *	0.006	1,5,7
3. MLP-HCG [13]	53.2 *	55.0	53.5 *	0.017	1,7,11
4. KNN3-HSV-CT	69.4 *	52.3	69.4 *	0.021	7
5. KNN5-HSV-CT	68.6 *	53.1	68.6 *	0.066	7
6. Proposed method	**78.2**	**77.2**	**77.9**	—	None

[1] CT stands for the color texture feature (DWT+LBP).
* stands for the proposed method is statistically significant compared to the given method when $p<0.01$.

sensitivity and specificity are above 50%. For different parameter settings, the result with the best performance is selected. For SVM training, LIBSVM toolbox [35] is deployed and RBF and Linear kernel functions are tested. The suitable parameters c and g are obtained by grid search.

The comparison result is illustrated in Table IV. Although SVM can demonstrate good performance as shown in previous research works, SVM is not robust for hookworm image detection among the three classification methods. To achieve a higher accuracy, most hookworm images are falsely classified as non-hookworm images. Due to robust performance, better generalization ability and less training time, MLP is a commonly used classifier [4], [5]. However, in our experiment, MLP cannot perform well except if combined with the HCG feature. They have more or less the similar performance. KNN based methods have better performance in terms of accuracy and specificity. The performance is improved from 50% to 70%. Our proposed method outperforms other approaches for all three metrics. Accuracy, sensitivity and specificity reach around 78%. It is robust for imbalanced dataset. More importantly, the proposed method has the lowest miss diagnosis rate. There is no patient missed while other methods will miss some patients. The proposed approach achieves a promising performance, especially considering the fact that it is conducted on a diverse and large dataset. Moreover, the high sensitivity actually implies that it is a potential and effective way for future clinical application. Similarly, the significance test between the proposed method and state-of-the-art approaches is performed, which is listed in Table IV. We can see that the results of accuracy and specificity are significant, which are marked with $*$ in Table IV. Meanwhile, for sensitivity, the proposed solution significantly outperforms other methods, since all patients are correctly detected while a couple of patients are missed for other approaches.

Despite encouraging performance, there are still some images are falsely detected as hookworms. Although a serial of novel ideas have been proposed to capture the properties of hookworms, it is difficult to discriminate some regions from hookworms, which are visually similar to hookworms in terms of morphology, color and width. Examples of falsely detected images are shown in Fig. 8. From this figure, we can see that most of them are deep folds or regions with narrow width (e.g., Fig. 8(a)–(c)). There are also some feces matters with line shapes, such as Fig. 8(d) and (e). They demonstrate two kinds of intestine matters. Fig. 8(d) is brighter than mucosa but darker than bubble edges, so our method will falsely classify it as a bright hookworm. On the contrary, Fig. 8(e) is falsely recognized as a dark hookworm due to its high visual similarity. Finally, some intestine vessels (e.g., Fig. 8(f)) are classified as hookworms, because the vessel structure is pretty similar to hookworm body.

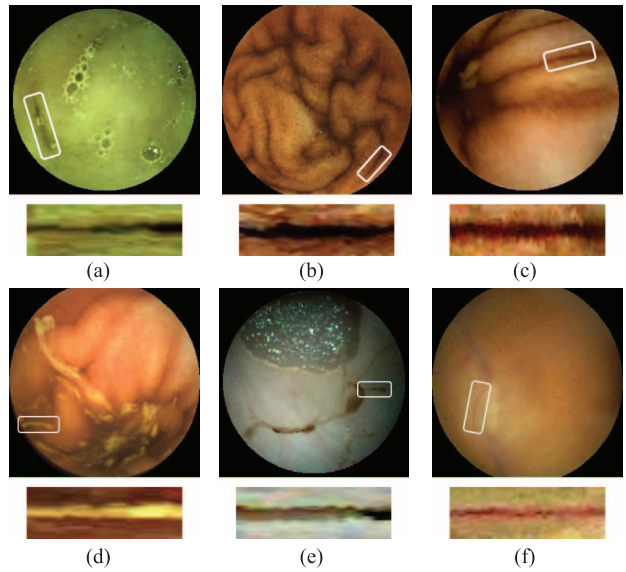

Fig. 8. Some false positive examples and their corresponding uncurled tubular regions. They are visually similar to hookworms, and difficult to be discriminated from hookworms.

VIII. CONCLUSION

Computer aided detection of hookworm for WCE images is a challenging task. By observing its unique properties, in this paper, we propose a serial of novel techniques to capture its characteristics, aiming to reduces the number of images a clinician needs to review. Experiments from different aspects demonstrate that the proposed method is a robust classification tool for hookworm detection, which achieves promising performance.

Although good performance has been achieved, there are still 23% of hookworm images cannot be correctly detected. In our future work, we will continue working on new solution to further improve the performance in our future work. The temporal and spatial relationship between consecutive images will be taken into consideration to boost the overall performance. In addition, due to its recent progress, deep learning approaches will be explored for hookworm detection. The ultimate goal is that automatic detection system can be used in a real condition to assist endoscopists, and can even obtain more accurate judgement than experienced endoscopists.

REFERENCES

[1] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature*, vol. 405, p. 417, 2000.
[2] Y. Fu, W. Zhang, M. Mandal, and M.-H. Meng, "Computer-aided bleeding detection in WCE video," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 2, pp. 636–642, Mar. 2014.
[3] P. Szczypiński, A. Klepaczko, M. Pazurek, and P. Daniel, "Texture and color based image segmentation and pathology detection in capsule endoscopy videos," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 396–411, 2014.
[4] S. Sainju, F. M. Bui, and K. A. Wahid, "Automated bleeding detection in capsule endoscopy videos using statistical features and region growing," *J. Med. Syst.*, vol. 38, no. 4, pp. 1–11, 2014.

[5] A. Karargyris and N. Bourbakis, "Detection of small bowel polyps and ulcers in wireless capsule endoscopy videos," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2777–2786, Oct. 2011.

[6] A. Mamonov, I. Figueiredo, P. Figueiredo, and Y. Tsai, "Automated polyp detection in colon capsule endoscopy," *IEEE Trans. Med. Imag.*, vol. 33, no. 6, pp. 1488–1502, Jun. 2013.

[7] B. Li and M. Q.-H. Meng, "Automatic polyp detection for wireless capsule endoscopy images," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 10 952–10 958, 2012.

[8] V. S. Charisis, L. J. Hadjileontiadis, C. N. Liatsos, C. C. Mavrogiannis, and G. D. Sergiadis, "Capsule endoscopy image analysis using texture information from various colour models," *Comput. Methods Progr. Biomed.*, vol. 107, no. 1, pp. 61–74, 2012.

[9] B. Li and M.-H. Meng, "Tumor recognition in wireless capsule endoscopy images using textural features and SVM-based feature selection," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 323–329, May 2012.

[10] R. Kumar *et al.*, "Assessment of Crohn's disease lesions in wireless capsule endoscopy images," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 355–362, Feb. 2012.

[11] A. Fenwick, "The global burden of neglected tropical diseases," *Public Health*, vol. 126, no. 3, pp. 233–236, 2012.

[12] Hookworm Wikipedia, 2014 [Online]. Available: http://en.wikipedia.org/wiki/Hookworm

[13] H. Chen, J. Chen, Q. Peng, G. Sun, and T. Gan, "Automatic hookworm image detection for wireless capsule endoscopy using hybrid color gradient and contourlet transform," in *Proc. 6th Int. Conf. Biomed. Eng. Informat.*, 2013, pp. 116–120.

[14] D. K. Iakovidis and A. Koulaouzidis, "Software for enhanced video capsule endoscopy: Challenges for essential progress," *Nature Rev. Gastroenterol. Hepatol.*, vol. 12, no. 3, pp. 172–186, 2015.

[15] A. Koulaouzidis, D. K. Iakovidis, A. Karargyris, and J. N. Plevris, "Optimizing lesion detection in small-bowel capsule endoscopy: From present problems to future solutions," *Expert Rev. Gastroenterol. Hepatol.*, vol. 9, no. 2, pp. 217–235, 2014.

[16] B. Li and M. Q.-H. Meng, "Wireless capsule endoscopy images enhancement via adaptive contrast diffusion," *J. Vis. Commun. Image Represent.*, vol. 23, no. 1, pp. 222–228, 2012.

[17] S. Seshamani, R. Kumar, G. Mullin, T. Dassopoulos, and G. D. Hager, "A meta method for image matching," *IEEE Trans. Med. Imag.*, vol. 30, no. 8, pp. 1468–1479, Aug. 2011.

[18] Y. Shen, P. Guturu, and B. P. Buckles, "Wireless capsule endoscopy video segmentation using an unsupervised learning approach based on probabilistic latent semantic analysis with scale invariant features," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 1, pp. 98–105, Jan. 2012.

[19] F. Vilarino *et al.*, "Intestinal motility assessment with video capsule endoscopy: Automatic annotation of phasic intestinal contractions," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 246–259, Feb. 2010.

[20] Q. Zhao, G. E. Mullin, M. Q.-H. Meng, T. Dassopoulos, and R. Kumar, "A general framework for wireless capsule endoscopy study synopsis," *Comput. Med. Imag. Graph.*, vol. 41, pp. 108–116, 2015.

[21] Y. Yuan, B. Li, and Q. Meng, "Bleeding frame and region detection in the wireless capsule endoscopy video," *IEEE J. Biomed. Health Informat.* , 2015, to be published.

[22] G. Chen, T. D. Bui, A. Krzyzak, and S. Krishnan, "Small bowel image classification based on Fourier-zernike moment features and canonical discriminant analysis," *Pattern Recognit. Image Anal.*, vol. 23, no. 2, pp. 211–216, 2013.

[23] L. Cui, C. Hu, Y. Zou, and M.-H. Meng, "Bleeding detection in wireless capsule endoscopy images by support vector classifier," in *Proc. IEEE Int. Conf. Informat. Automat.*, 2010, pp. 1746–1751.

[24] S. Segui *et al.*, "Categorization and segmentation of intestinal content frames for wireless capsule endoscopy," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 6, pp. 1341–1352, Nov. 2012.

[25] S. H. Bae and K.-J. Yoon, "Polyp detection via imbalanced learning and discriminative feature learning," *IEEE Trans. Med. Imag.*, vol. 34, no. 11, pp. 2379–2393, Nov. 2015.

[26] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "Rusboost: A hybrid approach to alleviating class imbalance," *IEEE Trans. Syst., Man Cybern., Part A, Syst. Humans*, vol. 40, no. 1, pp. 185–197, Jan. 2010.

[27] M. Veta *et al.*, "Assessment of algorithms for mitosis detection in breast cancer histopathology images," *Med. Image Anal.*, vol. 20, no. 1, pp. 237–248, 2015.

[28] J. F. Díez-Pastor, J. J. Rodríguez, C. I. García-Osorio, and L. I. Kuncheva, "Diversity techniques improve the performance of the best imbalance learning ensembles," *Inf. Sci.*, vol. 325, pp. 98–117, 2015.

[29] F. Amat *et al.*, "Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data," *Nature Methods*, vol. 11, no. 9, pp. 951–958, 2014.

[30] E. Magesan, J. M. Gambetta, A. D. Córcoles, and J. M. Chow, "Machine learning for discriminating quantum measurement trajectories and improving readout," *Phys. Rev. Lett.*, vol. 114, p. 200501, May 2015.

[31] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.

[32] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 263–269, Sep. 1989.

[33] Q. Li, J. You, and D. Zhang, "Vessel segmentation and width estimation in retinal images using multiscale production of matched filter responses," *Expert Syst. Appl.*, vol. 39, no. 9, pp. 7600–7610, 2012.

[34] S. M. Stigler, "Francis Galton's account of the invention of correlation," *Stat. Sci.*, pp. 73–79, 1989.

[35] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, p. 27, 2011.