2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Polyp Classification Based on Bag of Features and Saliency in Wireless Capsule Endoscopy*

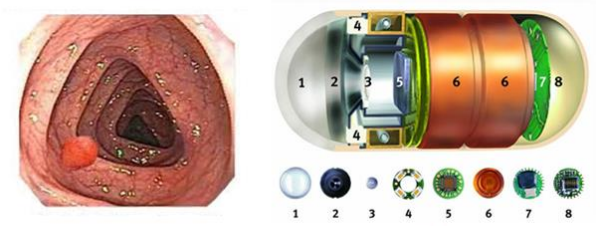
Yixuan Yuan and Max Q.-H. Meng, *Fellow, IEEE*

***Abstract*—Wireless capsule endoscopy (WCE) enables non-invasive visual inspection of the patients' digestive tract. However, the huge number of images from the WCE has been a hurdle for doctors to handle and thus it is urgent to develop computer-aided diagnosis systems to identify problematic images. To tackle this problem, an innovative algorithm based on the integration of the Bag of Features (BoF) method and the saliency map is proposed to detect polyps from the WCE images in this study. The algorithm constitutes of four steps. In the first step, by applying the BoF method, the visual words of all images are calculated by inputting the extracted Scale Invariant Feature Transformation (SIFT) feature vectors to the K-means clustering procedure. Then we calculate the saliency and non-saliency maps of the WCE images. Following that, the histogram of the visual words of each image is calculated by integrating histograms in both saliency and non-saliency maps with various weights to represent the WCE image. Finally, polyp classification of the WCE images is conducted by Support Vector Machine (SVM) classifier. Experiments on 436 polyp images and 436 normal images are carried out to validate the proposed algorithm. The proposed method with the weight 0.9 on the saliency region achieves a best polyp detection accuracy of 92%, sensitivity of 87.9% and specificity of 93%, demonstrating that the proposed method provides a good characterization and description for polyp classification.**

***Index Terms*—Wireless capsule endoscopy, bag of feature, saliency map, polyp classification**

I. Introduction

Colorectal cancers are among the great threats to human's health. According to the statistics of the Centre for Health Protection in Hong Kong for 2010, colorectal cancer is the second commonest cancer and it accounted for 16.6% of all new cancer cases [1]. Moreover colorectal cancer is the third leading cause of cancer deaths in males and second leading cause of cancer deaths in females in Hong Kong. In 2011, 1,904 persons died from colorectal cancer, accounting for 14.4% of all cancer deaths [1]. This severe situation inspires us to study the causes of the colorectal cancer. Actually the colorectal cancer always begins as a polyp [2], which is an abnormal growth of tissue projecting from a mucous membrane (Fig. 1(a)). Although most colorectal polyps do not become cancers, virtually all colon and rectal cancers start from them. Thus it is crucial for doctors to be able to detect polyps in their early stage and remove them before they deteriorate to cancer cells.



(a) (b)

Figure 1. (a) Illustration of polyps (b) Components diagram of a WCE. 1.Optical dome, 2.Lensholder,3-Lens, 4.Illuminating LEOs, 5.CMOS imager,6.Battery, 7.Transmitter, 8.Antenna.

Previous endoscopic imaging technologies for gastrointestinal (GI) tract includes upper endoscopy, colonoscopy [3, 4]. Although they demonstrated important value in diagnosing diseases of the digest tract, the traditional endoscopies can not reach to the small intestine due to technique limits. The advance of WCE [5] has dramatically changed the diagnosis and management of polyps inside the small intestine. It eliminates the pain of patients and more importantly provides direct visualization of the entire small intestine tract for the first time. After swallowing by a patient, the WCE travels through the entire GI tract propelled by peristalsis. Equipped with a micro-camera and wireless communication capability (Fig. 1(b)), the WCE sends out images of the GI tract to a data-recording device for doctors to examine and make diagnostic decisions. However this new technology is not perfect since it produces over 55,000 images for one patient during the eight-hour examination process [6]. This huge number of images creates a hurdle for doctors to process and it usually takes 2 hours to go through the images even by a well-trained professional clinician. Therefore, the computer-aided diagnosis systems are highly demanded.

Many studies have been carried out to classify polyps automatically in the WCE images. Alexandros Karargyris et al [7] applied Log Gabor filters and SUSAN edge detector to preprocess the images and then identified polyp candidates based on geometric information. Li et al [8] utilized a new texture feature that combines the advantages of wavelet transform and uniform local binary pattern with SVM [9] to discriminate between regions of normal and abnormal tissues. Hoda Eskandar et al [10] proposed to use region-based Active Contour Method (ACM) and geometric features for automatic detection of polyps. Majority of the published research work on the polyp classification adopted the approach of extracting the complete features of the whole images followed by a classification method.

*This project is partially supported by RGC GRF # 415709 awarded to Max Q.-H. Meng.

Yixuan Yuan is with the Department of Electronic Engineering, The Chinese University of Hong Kong, N.T., Hong Kong SAR, China (e-mail: yxyuan@ee.cuhk.edu.hk).

Max Q.-H. Meng is with the Department of Electronic Engineering, The Chinese University of Hong Kong, N.T., Hong Kong SAR, China (corresponding author, phone: 852-2609-8282; fax: 852-2603-5558; e-mail: qhmeng@ee.cuhk.edu.hk).

In recent years, the BoF method [11] derived from local key points has attracted much attention in the object recognition and categorization tasks. The main idea of the BoF is to treat a single image as an order-less collection of local key point features, quantize the descriptors into visual words and represent the image as a histogram of these words. The BoF method followed by a classification method is typically used to conduct the classification experiments. However, the traditional BoF method applies a uniform sampling strategy to the whole image and does not consider the situation when the object is not dominant or there are some other objects or background in the image, thus it tends to lead to an unsatisfactory results [12]. On the other hand, the saliency map that represents visual saliency of a corresponding visual scene, attracts more and more attentions of the researchers. Many new algorithms have been proposed to model the human vision and extract the important information of the images [13-16].

In this paper, we propose a novel polyp classification method based on the combination of the BoF method and the saliency map method. The proposed method not only takes advantage of the BoF method that represents images with effective histograms by the local SIFT features [17], but also emphasizes the SIFT features in the saliency map. Therefore, it can provide a good description of the WCE images for the polyp classification tasks. We first calculate the saliency map of the WCE images and outline the saliency points with relatively higher response than the average values. Then we extract the SIFT feature vector on the whole images and utilize the K-means clustering method on the features of the training images to obtain the visual words. After that, different weights for each SIFT feature are assigned based on whether the feature point is in the saliency map or not, thus we obtain the final integrated histogram of the image. Finally the integrated histograms are used as the input to a SVM classifier to carry out the polyp classification tasks.

Our main contributions can be summarized as follows:

- A framework integrating the BoF method and the saliency map method is proposed to detect the polyps. This method extracts the most important parts of images, thus it can portray the images effectively.
- We analyze the influence of weights assigned to the saliency and non-saliency maps in relation to the classification performance.

The rest of the paper is organized as follows: Section II presents the proposed method for polyp classification based on the BoF method and the saliency map method, including patch extraction and feature representation, formation of visual words, saliency map extraction, the integration method of visual words histograms, and SVM classification. The experimental results are discussed in Section III and we draw some conclusions at the end of this paper in Section IV.

II. METHODS

The flowchart of the proposed algorithm is illustrated in Fig. 2. The steps are described as follows: (1) Image patch extraction and feature description using the SIFT feature, (2) Visual words generation: quantizing the features by K-means clustering algorithm to obtain the visual words, (3) The saliency map extraction aims to represent the images effectively, (4) Construction of the integrated histogram: after extracting the saliency maps of the WCE images, the histograms of the saliency and non-saliency maps are computed and integrated with various weights to obtain the final combined histograms, and (5) SVM training and testing for polyps detection. These steps are discussed in details in the following sections.

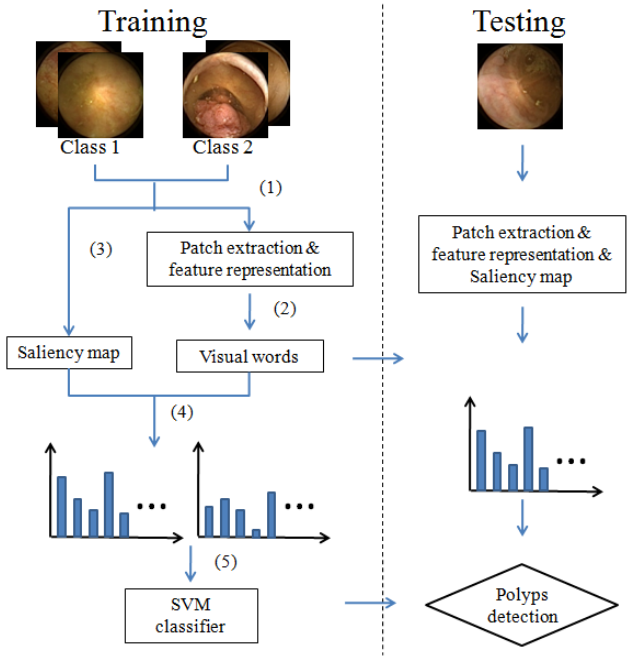


Figure 2. Flowchart of the proposed method.

A. Patch Extraction & Feature Representation

1) Region Detection

Different from the global feature describing a picture in a holistic way, one image is represented by the combination of the features of the key points in the BoF method. Here, we use the SIFT method [17] to detect the key points, which is one of the most powerful tools for extracting key points and insensitive to different conditions such as rotation, scale changes and noise. The scale spaces of an image are first obtained by convolving the image with multi-scale Gaussian functions. Then the differences of Gaussians (DoGs) at multiple scales are calculated. Following that, we compare each sample point to its 26 neighbors in 3×3 regions at the current and adjacent scales and select the key points if it is larger than all of these neighbors or smaller than all of them. Finally the selected local maxima and minima are set to be the detected key points.

2) Local Descriptors

The previous operations have assigned the SIFT key points in an image. The next step is to compute SIFT descriptors to describe local features. We sample a 16×16 region around each keypoint to calculate gradient magnitudes and orientations. These samples are then accumulated into orientation histograms over 4×4 sub-regions to summarize the contents, as shown in Fig. 3. In order to limit the descriptors' size, we reduce the bin size to 8 to represent the 360-degree range of orientations. Therefore, the dimension of the feature vector for each keypoint is 4×4×8 =128.

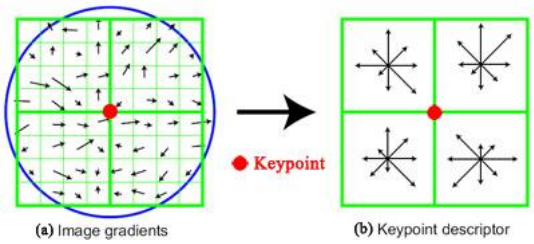


Figure 3. Illustration of the SIFT feature.

B. Formation of Visual Words

We use the K-means algorithm on all the SIFT feature vectors to generate the visual vocabulary. The resultant cluster centers serve as a vocabulary of visual words and the number of clusters determines the size of visual vocabulary. The visual words correspond to a discrete quantization of the descriptor space and further determine the representative features that are able to describe the input images. In this study, we set the number of visual words to 100.

C. Saliency Map Extraction

1) ROI Extraction

The original images obtained from the WCE video clips demonstrate the large dark background and obvious bounder as shown in Fig. 4(a, b). Thus we outline the maximum square inscribed in the circular image as region of interest (ROI) to remove the influence of these factors. The size of the obtained ROI is 180×180 from the original image of 256×256 and we find that the remaining part represents the main texture and shape characteristics very well and also makes the following procedures much easier.

Figure 4. Illustration of ROI extraction.

2) Saliency Map Extraction

The saliency map outlines the areas in images and videos that capture the attention of the human visual system. The region with greater value in the saliency map, the larger possibility the people tend to view. Since the saliency map extracts the important region of the original image, it can characterize the images effectively.

In this section, we utilize the method proposed by Murray et al [16] to extract the saliency map since it demonstrates excellent ability compared with state-of-the-art models. It can be summarized by the following pipeline:

$$I \xrightarrow{WT} \{w\} \xrightarrow{CS} \{z\} \xrightarrow{ECSF} \{\alpha\} \xrightarrow{WT^{-1}} s$$

Where WT and WT^{-1} are the wavelet transform and inverse of wavelet transform. CS represents the center surround mechanism and $ECSF$ is the extended contrast sensitivity function.

In the first stage, the original image is convolved with a bank of filters using a multi-resolution wavelet transformation to obtain the local oriented contrasts.

The second stage of this model consists of a simulation of the inhibition mechanisms in cells of the visual cortex, which effectively normalizes cortex cell responses to stimulus contrast. The sizes of the central and normalizing surround windows were learned by training a Gaussian Mixture Model (GMM) based on the eye-fixation data.

In the third stage, non-linear integration ECSF is calculated through a weighting function that is similar to the one proposed by Otazu et al. [18]. Moreover optimization is applied to fit psychophysical color matching data at different spatial scales. After that, an inverse wavelet transform is performed directly on the weights in the previous stage to integrate information at multiple scales.

After setting an appropriate threshold to the obtained saliency map, the points with higher response will formulate the saliency region.

D. Integration of Histogram

Once obtaining the 100 visual words, we map features of each key point in the saliency map to the nearest visual words and calculate the count of each visual word, yielding a histogram (w, d) of the saliency map, where w_i denotes the ith visual word in the 100-size dictionary and d_i counts the frequency of occurrence of it. Utilizing the same method, we obtain the histograms of the non-saliency map with the SIFT features.

Since the saliency map represents the important information of the image, we increase the weight α ($0.5 \le \alpha \le 1$) on the histograms of this part and decrease the weight $\beta = 1 - \alpha$ on the histograms of the non-saliency part. Let the histograms of visual words in saliency map and non-saliency map be h_1 and h_2, then the histograms integrating the BoF method and the saliency method can be obtained by $\alpha \times h_1 + \beta \times h_2$. The obtained integrated histograms highlight the features in the saliency map, thus it can characterize the images well. The workflow of the construction of the integrated histogram is shown in Fig. 5. In order to evaluate the influence of saliency map to the polyp classification, we gradually increase the weight α on the histograms in saliency region from 0.5 to 1 and evaluate the classification results under each setting.

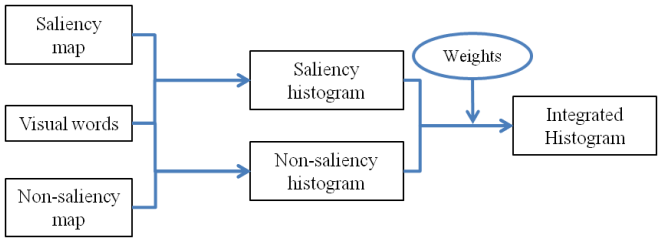


Figure 5. Construction of the integrated histogram.

E. Support Vector Machine

Support Vector Machine (SVM) [9] is a supervised machine learning method on the foundation of statistical learning. The basic idea of the SVM is to find the optimal hyper-plane that separates the points of different classes.

Considering a training set $(x_i, y_i), i = 1, 2 \ldots n$ and the associate output $y \in \{+1, -1\}$, the purpose of SVM is to find a following hyper-plane to classify the polyp images and the normal images:

$$w \cdot \Phi(x) + b = 0 \tag{1}$$

where $\Phi(x)$ is a nonlinear mapping from the input space to the feature space. We should maximize the margin of the separation plane:

$$w = \sum_{i=1}^{n} \lambda_i y_i \Phi^T(x) \tag{2}$$

where the Lagrange multipliers λ_i represent the parameters of the separating surface. It can be estimated through the maximization of L_D in the following:

$$L_D = \sum_{i=1}^{n} \lambda_i - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j K(x_i, x_j)$$
$$\sum_{i=1}^{N} \lambda_i y_i = 0 \tag{3}$$
$$0 \le \lambda_i \le C$$

where $K(x_i, x_j)$ is a kernel function and can be defined as the inner product $K(x_i, x_j) = \Phi^T(x_i)\Phi(x_j)$. C is the penalty parameter of the error term. In this paper, we use the linear kernel for the polyp classification problems.

Given an input of testing integrated histogram x, the trained SVM outputs the corresponding predict label of class by

$$s(x) = sign\{\sum_{i=1}^{n} \lambda_i y_i K(x_i, x) + w_0\} \tag{4}$$

III. EXPERIMENTAL RESULTS

To classify polyp in the WCE images, we proposed a new algorithm integrating the BoF method and the saliency map method. We conducted experiments using a set of 872 WCE images consisting of 436 polyp frames and 436 normal frames, extracted from different patients' WCE video clips. These images were all labeled by experienced clinicians. We randomly choose 300 normal samples and 300 polyp samples from the whole datasets as training sets, while the remaining 136 normal samples and 136 polyp samples are used as test sets. These procedures are repeated 10 times and the average recognition rates are used for the assessment of the classification performance. The obtained results are discussed as follows.

A. Results of Saliency Extraction

After obtaining the ROI (Fig. 6 (a, b)), we applied saliency map extraction method on the ROI images and the corresponding saliency maps are shown in Fig. 6 (c, d). The saliency map represents the human fixation behavior during free viewing, and people tend to focus on the region with greater values in the saliency map. As shown in Fig. 6 (c, d), the extracted saliency regions in the frames are satisfactory since they demonstrate the major information that interests the human visual system.

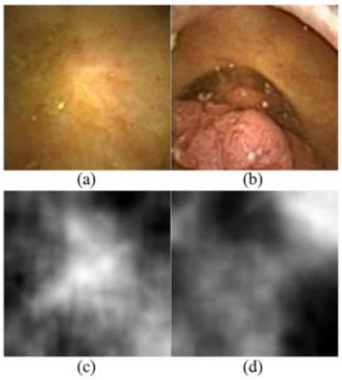


Figure 6. Illustration of saliency extraction.(a, b) Examples of WCE polyp image, (c, d)Corresponding saliency map of (a, b).

B. Results of Integrated Histogram

After the saliency map is computed, the points with relatively higher response than the mean value are obtained as the saliency region as shown in Fig. 7(c). In the proposed method, we first extracted the SIFT feature for each key point (Fig. 7(d)) and utilized the K-means clustering method on the obtained SIFT features of the training sets to obtain the visual words. Then we divided the SIFT patch into two parts: the key points in the saliency region (Fig. 7(e)) and that in non-saliency region (Fig. 7(f)). The SIFT features of these two regions are coded by hard assignments to the nearest visual words, yielding two histograms. As shown in Fig. 7(e), the SIFT key points in the saliency region provide sufficient and rich information about the whole image, thus we increased the weight on the histogram of the saliency region to emphasize the features in the saliency map and combined it with the weighted histogram of the non-saliency region to represent the complete information of the image.

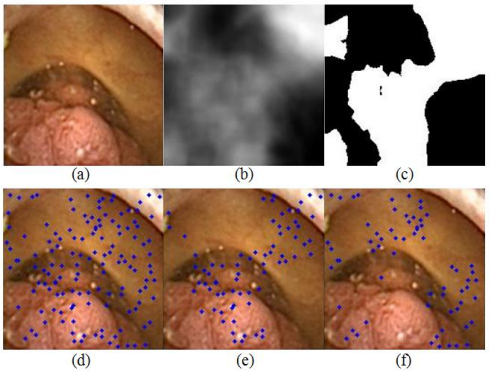


Figure 7. (a) Example of WCE polyp image, (b) Corresponding saliency map, (c) The saliency region, (d) The SIFT points in the whole image, (e) SIFT points in the saliency region, (f) SIFT points in the non-saliency region.

C. Results of Classification

We gradually increased the weights on the histograms of the saliency maps from 0.5 to 1 and applied *libsvm* [19] to classify the polyp images from the normal images. The classification performance in the WCE images is measured by the indexes of accuracy, specificity and sensitivity. These indexes are widely used to evaluate classification performance and can be obtained by the following formulas:

$$Accuracy = \frac{No.\ of\ correct\ predictions}{Total\ sample} \quad (5)$$

$$Sensitivity = \frac{No.\ of\ correct\ positives\ predictions}{No.\ of\ positives} \quad (6)$$

$$Specificity = \frac{No.\ of\ correct\ negative\ predictions}{No.\ of\ negatives} \quad (7)$$

As for the polyp classification problem in the WCE images, sensitivity shows the capability of detecting polyp images while specificity means the ability to avoid false detection. Moreover, there is a trade-off between these two criteria. The higher of the specificity, the lower the sensitivity is. The accuracy evaluates the overall classification results. In order to detect the polyp effectively, the proposed method should have high accuracy and sensitivity.

We listed the final results of 10 independent experiments in Table I. In the table, the mean, median, maximum, and minimum values of accuracy, specificity, and sensitivity under different saliency weights are shown respectively for comparison. The best classification result with the accuracy of 92% was achieved with weight of 0.9 on the saliency map. Moreover, when the weight of the saliency region is set to 1, the best result is not obtained. This result indicates that the non-saliency region still includes some useful information for polyp classification. When these two image regions are considered equally as in the traditional BoF method, i.e. the same weights for saliency map and non-saliency map, the accuracy is the lowest. It is noted that when the weight in the saliency map is increased from 0.5 to 0.9, the mean accuracy is improved by 4.85%.

TABLE I. COMPARISON OF CLASSIFICATION PERFORMANCE UNDER DIFFERENT WEIGHTS IN SALIENCY REGION

Saliency map weights		0.5	0.6	0.7	0.8	0.9	1
Non-saliency map weights		0.5	0.4	0.3	0.2	0.1	0
Accuracy (%)	**Mean**	**85.15**	85.6	85.85	88.05	**90**	89.35
	Median	84.75	85	85.5	87.5	90	89.5
	Max	88.5	89	89.5	91.5	**92**	91
	Min	82.5	83.5	84	85.5	88	86.5
Sensitivity (%)	**Mean**	74.5	75.5	76.1	81.4	**87.9**	88.8
	Median	75	75.5	76.5	81	87.5	88.5
	Max	81	81	82	90	94	95
	Min	67	70	71	75	84	85
Specificity (%)	**Mean**	97.8	97.7	97.6	95.7	**93**	90.7
	Median	98	98	98	95.5	93	91
	Max	99	99	99	99	94	96
	Min	93	93	93	92	91	86

Moreover, in order to view the results vividly, the average values of accuracy, specificity and sensitivity of 10 experiments are compared in the Fig. 8. It can be observed that the sensitivities under different weights in the saliency map are not as high as the corresponding specificities, demonstrating the fact that the variations of polyp images are much larger than those of the normal WCE images. The best classification results with the weight 0.9 in saliency map achieve the mean accuracy of 90%, sensitivity of 87.9% and specificity of 93%, showing that the proposed method provides a good characterization and description of the WCE images for polyp classification tasks.

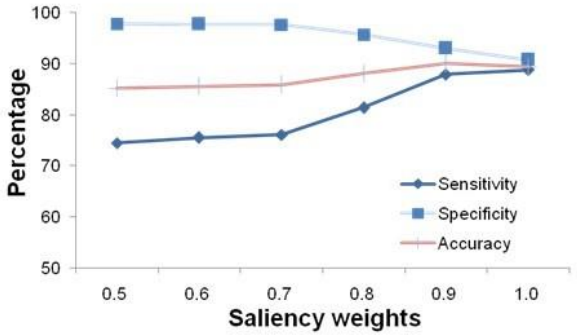


Figure 8. Classification performance with different weights in saliency regions.

D. Results of Comparison

There are many reported researches focusing on the polyp classification, but majority of them adopt the approach of extracting the complete features from the whole images followed by a classification method. To further evaluate the performance of the proposed method, we compared it with the state-of-art methods: the local binary pattern (LBP) feature [20], the Binary Gabor pattern feature (BGP) feature [21], the histogram of gradient feature (HOG) [22] and the method in [8], since these features all exhibit excellent capabilities in the image classification tasks.

The Local Binary Pattern (LBP) operator is a simple yet powerful gray-scale invariant texture feature, derived from a very simple binary coding method that compares neighboring pixels with the central pixel. The BGP feature is a new efficient and effective multi-resolution texture extraction approach. It can be obtained as follows: convolve images with Gabor filters with different orientations, binarize the obtained responses and assign "rotation invariant binary Gabor pattern (BGPri)" for each location. Finally images will be represented as the histogram of BGPri. The third algorithm we considered is the HOG feature. The basic idea of the HOG method is that the appearance and shape of the local object can be characterized well by the distribution of local intensity gradients or edge direction. In practice, it is implemented by dividing the complete images into small spatial regions, accumulating a local 1-D histogram of gradient orientations for each region and combining these histograms together to form final features. The final method we compared is the recent published method by Li et al [8] focused on the polyp detection in the WCE images. It is reported in his paper that the best detection accuracy is obtained with the combination of the uniform_LBP features extracted from all three different

levels of the wavelet transform in RGB color space. We implemented this algorithm under this setting and compared it with our proposed method. After that, we applied the SVM classifier to carry out experiments for these features and the corresponding results are showed in Table II.

TABLE II. AVERAGE CLASSIFICATION RESULTS USING DIFFERENT TEXTURAL FEATURES(100%)

	Proposed method	LBP	BGP	HOG	[8]
Acc.	90	68.38	71.32	87.13	72.06
Sen.	87.9	42.65	82.35	77.94	51.47
Spe.	93	94.12	60.29	94.32	92.65

It can be concluded from the Table II that the proposed method shows superior performance to those of the LBP method with an improvement of 21.62% and 45.25% in accuracy and sensitivity, respectively, while to those of BGP with an improvement 18.68% and 5.55%, respectively, and to those of HOG based features with an increase of 2.87% and 9.96%, respectively. This result validates that the integrated histograms emphasized the SIFT feature in the saliency map possess superior ability to characterize the WCE images. In addition, compared with Li's method, the proposed method shows improvement of 17.94% and 36.43% in accuracy and sensitivity, respectively, showing that the proposed method has better discrimination ability for polyp recognition in the WCE images.

IV. CONCLUSION

This paper presents a new algorithm for polyp detection in the WCE images based on the integration of the BoF method and the saliency map method. The BoF method characterizes the images with effective histograms using the local features, while the saliency map extraction method outlines the important parts of the images. Therefore, the hybrid method leads to a better description for the WCE images. In this method, the saliency map of the WCE image is calculated and the saliency regions are outlined via an appropriate threshold. Then the BoF visual words are obtained by inputting the extracted SIFT features into the K-means clustering method. The integrated histograms are calculated by assigning different weights on the features of the saliency and non-saliency regions. Finally, the polyp classification in WCE images is conducted by using the SVM classifier. The proposed scheme with the weight value of 0.9 achieved an accuracy of 90% as the best result, together with a sensitivity of 87.9% and specificity of 93%. Furthermore, the accuracy of our proposed method that emphasizes the SIFT feature in the saliency map, is 4.85% higher than that of the traditional BoF method, demonstrating the superiority of the proposed new algorithm. The performance of the proposed method has been compared with several state-of-the-art algorithms: LBP, BGP, HOG and the recent best published method proposed by Li et al. Experiment results show that the proposed method demonstrates superior discriminative ability for polyp detection compared with these methods.

ACKNOWLEDGMENT

I want to express my great gratitude to Jiaole Wang, who helped a lot for the preparation of this manuscript.

REFERENCES

[1] http://www.chp.gov.hk/en/content/9/25/51.html.

[2] http://en.wikipedia.org/wiki/Colorectal_polyp.

[3] G. Gay, M. Delvaux, and J.-F. Rey, "The role of video capsule endoscopy in the diagnosis of digestive diseases: a review of current possibilities," Endoscopy, vol. 36, pp. 913-920, 2004.

[4] M. Yu, "M2a (tm) capsule endoscopy: A breakthrough diagnostic tool for small intestine imaging," Gastroenterology Nursing, vol. 25, pp. 24-27, 2002.

[5] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," Nature, vol. 405, p. 417, 2000.

[6] R. Kumar, Q. Zhao, S. Seshamani, G. Mullin, G. Hager, and T. Dassopoulos, "Assessment of crohn's disease lesions in wireless capsule endoscopy images," Biomedical Engineering, IEEE Transactions on, vol. 59, pp. 355-362, 2012.

[7] A. Karargyris and N. Bourbakis, "Detection of small bowel polyps and ulcers in wireless capsule endoscopy videos," Biomedical Engineering, IEEE Transactions on, vol. 58, pp. 2777-2786, 2011.

[8] B. Li and M. Q.-H. Meng, "Automatic polyp detection for wireless capsule endoscopy images," Expert Systems with Applications, vol. 39, pp. 10952-10958, 2012.

[9] C. Cortes and V. Vapnik, "Support-vector networks," Machine learning, vol. 20, pp. 273-297, 1995.

[10] H. Eskandari, A. Talebpour, M. Alizadeh, and H. Soltanian-Zadeh, "Polyp detection in Wireless Capsule Endoscopy images by using region-based active contour model," in Biomedical Engineering (ICBME), 2012 19th Iranian Conference of, 2012, pp. 305-308.

[11] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, 2006, pp. 2169-2178.

[12] L. Yang, N. Zheng, J. Yang, M. Chen, and H. Chen, "A biased sampling strategy for object categorization," in Computer Vision, 2009 IEEE 12th International Conference on, 2009, pp. 1141-1148.

[13] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 20, pp. 1254-1259, 1998.

[14] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in Advances in neural information processing systems, 2006, pp. 545-552.

[15] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 34, pp. 1915-1926, 2012.

[16] N. Murray, M. Vanrell, X. Otazu, and C. A. Parraga, "Saliency estimation using a non-parametric low-level vision model," in Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 2011, pp. 433-440.

[17] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International journal of computer vision, vol. 60, pp. 91-110, 2004.

[18] X. Otazu, C. A. Parraga, and M. Vanrell, "Toward a unified chromatic induction model," Journal of Vision, vol. 10, 2010.

[19] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 2, p. 27, 2011.

[20] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern recognition, vol. 29, pp. 51-59, 1996.

[21] L. Zhang, Z. Zhou, and H. Li, "Binary Gabor pattern: An efficient and robust descriptor for texture classification," in Image Processing (ICIP), 2012 19th IEEE International Conference on, 2012, pp. 81-84.

[22] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, pp. 886-893.