# Automatic Polyp Detection in Endoscope Images Using a Hessian Filter

Yuji Iwahori, Takayuki Shinohara, Akira Hattori
Chubu University, 487-8501 Japan
iwahori@cs.chubu.ac.jp

Robert J. Woodham
UBC, Canada V6T 1Z4
woodham@cs.ubc.ca

Shinji Fukui
Aichi Univ. of Educ.
448-8542 Japan
sfukui@auecc.aichi-edu.ac.jp

M. K. Bhuyan
IIT Guwahati
781039 India
mkb@iitg.ernet.in

Kunio Kasugai
Aichi Med. Univ.
480-1195 Japan
kuku3487@aichi-med-u.ac.jp

## Abstract

*An endoscope is a medical instrument that acquires images inside the human body. This paper proposes a new approach for the automatic detection of polyp regions in an endoscope image using a Hessian filter and machine learning techniques. Previous approaches tried to detect candidate polyp regions based on rectangular patches. But, a purely patch-based approach can miss classify candidate regions because other information necessarily is included in each rectangular patch. Here, a Hessian filter is used to detect image regions corresponding to blob-like structures. Detailed color and edge features are extracted only for the detected candidate regions. SVMs (with Boosting) are constructed to classify candidate regions as polyps. The new approach is demonstrated experimentally. High accuracy is achieved.*

## 1 Introduction

Medicine is an important area of application for computer vision. Endoscopy allows medical practitioners to observe the interior of hollow organs and other body cavities in a minimally invasive way. Diagnosis involves both shape detection and the assessment of tissue state. For example, a polyp is a pathological condition directly related to geometrical shape. Diagnosis typically requires polyp removal and biopsy. Here, we consider a general purpose endoscope, of the sort still most widely used in medical practice.

Polyps are usually found via endoscopy but polyps can be missed. Automatic detection of polyps, with high accuracy, is an important aid to medical practice.

Previous work has used a patch-based approach [1][2][3]. In [1] and [2], patch features computed are the Color Wavelet Covariance (CWC) and the Local Binary Pattern (LBP). Candidate patches are classified using an SVM. In [3], higher dimensional features of the RGB color values and the XY position coordinates are used leading to improved classification performance.

Performance of previous patch-based approaches depends on the patch size. It is not straightforward to detect polyps with differing sizes in an image. Further, smaller polyps become quite sensitive to the features used for detection. It is difficult to imagine how to achieve robustness with a constant patch size.

In this paper, we detect candidate polyp regions based on the assumption that a polyp is convex. We compute color and edge features for each candidate region. By focusing on blob-like regions, the relevance of the features extracted is improved. SVM classification (with Boosting) is used to detect polyps with high accuracy.

## 2 Extraction of Candidate Polyp Regions

### 2.1 Candidate Polyp Regions from Convex and Concave Features

The convexity and concavity of the image intensity surface itself is used to select candidate polyp regions. A Hessian filter is applied to the endoscope image to determine regions of convex and concave image intensity. The image is segmented based on the results of Hessian filtering. Low pass filtering, via Gaussian smoothing, is included to remove the effects of noise and texture.

In [4], tube-like image features arising from blood vessels are enhanced according to the eigenvalues, $\lambda_1, \lambda_2$, of the associated Hessian matrix as shown in Table 1. Table 1 is the classification which we also use. Here, however, the goal is to extract blob-like features. The procedure to enhance blob-like structure is as follows:

Step 1. Generate a Gaussian filter, $G(x, y)$, with (width) scale $s$.

Step 2. Generate filters for the partial derivatives, $G_x(x, y)$, $G_y(x, y)$, and the second partial derivatives, $G_{xx}(x, y)$, $G_{xy}(x, y)$ and $G_{yy}(x, y)$.

Step 3. Compute $L_{xx}(x, y)$, $L_{xy}(x, y)$ and $L_{yy}(x, y)$ as the convolution of input image, $L(x, y)$, with $G_{xx}(x, y)$, $G_{xy}(x, y)$ and $G_{yy}(x, y)$ respectively. That is

$$L_{xx}(x, y) = G_{xx}(x, y) * L(x, y) \quad (1)$$
$$L_{xy}(x, y) = G_{xy}(x, y) * L(x, y) \quad (2)$$
$$L_{yy}(x, y) = G_{yy}(x, y) * L(x, y) \quad (3)$$

where $*$ denotes convolution.

Step 4. Compute the ($2 \times 2$) Hessian matrix image, $\mathbf{H}$, from Eqs.(1-3).

$$\mathbf{H} = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{bmatrix} \quad (4)$$

Table 1. Eigenvalues of the Hessian Matrix ($H$=high, $L$=low, $N$=noisy)

| $\lambda_1$ | $\lambda_2$ | pattern and structure |
|---|---|---|
| $N$ | $N$ | noisy, no preferred direction |
| $L$ | $H_-$ | tube structure (bright) |
| $L$ | $H_+$ | tube structure (dark) |
| $H_-$ | $H_-$ | blob-like structure (bright) |
| $H_+$ | $H_+$ | blob-like structure (dark) |

Step 5. Calculate the eigenvalues, $\lambda_1$, $\lambda_2$, $(\lambda_1 \leq \lambda_2)$, of $\mathbf{H}$

Step 6. Calculate the image, $I(x, y)$, where

$$ I(x,y) = \left( \lambda_1(x,y)^2 + \lambda_2(x,y)^2 \right) \left( 1 - \left( \frac{\lambda_1(x,y)}{\lambda_2(x,y)} \right)^2 \right) \tag{5} $$

Again, the classification table for pattern and structure based on the eigenvalues of the Hessian matrix is shown in Table 1. Here, $H$ corresponds to the high (largest in magnitude) eigenvalue and $L$ corresponds to the low (smallest in magnitude) eigenvalue. $H_+$ and $H_-$ denote the sign of the high eigenvalue, $H_+$ if $H > 0$ and $H_-$ if $H < 0$, respectively. Recall that $\lambda_1 \leq \lambda_2$.

An example of an original endoscope image is shown in Fig.1-(a). The corresponding image with blob-like structures enhanced according to Eq.(5) is shown in Fig.1-(b). A Voronoi diagram based segmentation [6] is used to generate candidate polyp regions. Segmentation with conservative (i.e., lower) threshold values is shown in Fig.1-(c). Red points correspond to points with highest values while points with blue points correspond to points with lowest value. Segmentation with higher threshold values is shown in Fig.1-(d). For this example, the segmentation shown in Fig.1-(d) determines the candidate polyp regions. False positives are acceptable at this stage. Features are extracted for each candidate region and the final classification whether or not each candidate region is a polyp is determined subsequently.
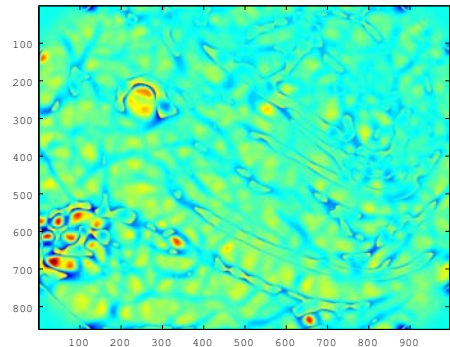
## 2.2 Feature Extraction

Features we use for classification are statistical measures obtained from the RGB color values and local edge-like descriptors obtained from the Histogram of Oriented Gradients (HOG) [7].

RGB color values for the candidate regions are converted also to the HSV, YUV, TSL, and Lab color systems. Maximum, minimum, mean, variance, mode, median, skewness, kurtosis, entropy and moment are computed for each component. There are 5 color systems, 3 components per system and 10 measures. The dimensionality of the resulting color intensity feature vector is $5 \times 3 \times 10 = 150$. Multiple color conversions and associated statistical measures are intended to provide robustness with respect to overall brightness changes in the endoscope images and to include color features not directly available in the RGB representation. HOG features for the candidate regions are computed for each component of the HSV representation. The window size is $3 \times 3$. Gradient direction is quantized into 8 directions. The dimensionality of the resulting HOG feature vector is $3 \times (3 \times 3) \times 8 = 216$.
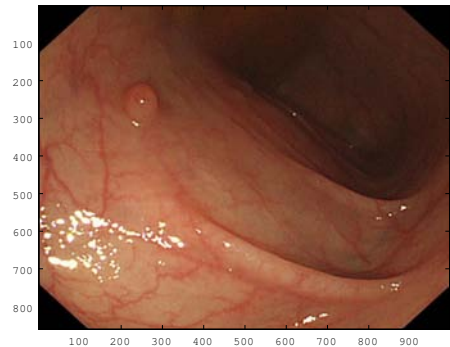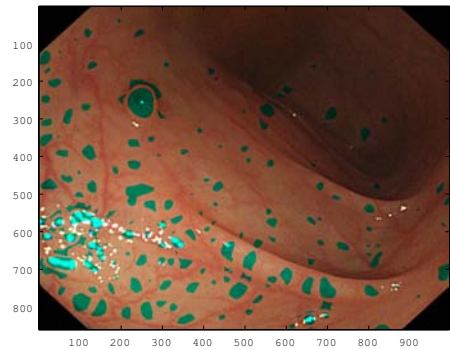


(a) Original Image



(b) Enhanced Image
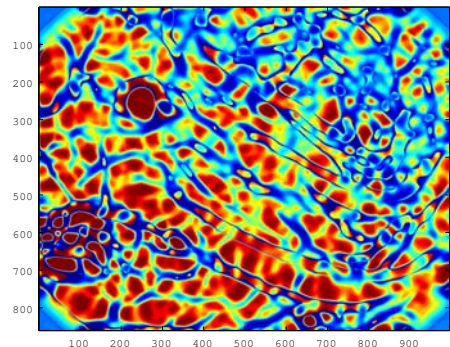


(c) Segmented Image



(d) Candidate Region

Figure 1. Extraction of Candidate Polyp Regions

High dimensionality is an issue. Feature selection is used to find useful combinations of these 150+216 features, as described in the next section.

# 3 SVM for Polyp Classification

We use an SVM with Gaussian kernel function. Determining a tractable, sub-combination of features to use plays an important practical role in classifier performance. Sequential Floating Forward Search (SFFS) [8], with incremental or decremental combination of features, is used to select features that give the best performance in training. Performance is evaluated by Leave One Out (LOO) cross-validation.

Boosting is also used to improve performance using multiple classifiers. Specifically, rounds of AdaBoost [9] are applied to each SVM trained with features selected for different combinations of SVM parameters. Each SVM trained is designated as a weak classifier and AdaBoost is used to improve classification performance.

# 4 Experiments

The labeled data set used in our experiments consisted of 128 real endoscope images, each $1000 \times 900$ pixels in size. The images all include regions with polyps. Labeling was done manually by occluding each polyp with a black mask. To illustrate, Fig.2-(a) shows an original image and Fig.2-(b) shows the same image with one polyp labeled. The full data set was labeled manually by marking different sized polyp regions with different sized masks. The histogram of labeled polyp sizes, given in pixels, is shown in Fig.3. This is the data set used in our experiments. Polyp sizes range from 1,000 to 13,000 pixels. Features are computed for the labeled polyp regions (positive examples) and for other regions (negative examples). Negative examples are chosen at random so that the number of true positives and the number of true negatives is equal in the labeled data set.

Our approach was compared with the approaches described in [3] and [5]. In [3], a patch based approach is described using as features the RGB intensity values and the XY positions of each pixel. An SVM is used for the classification. In our implementation of [3], we used PCA (Principal Component Analysis) to reduce the dimension of feature vectors before input to the SVM. In [5], the patch features used are Rotation Invariant Uniform Local Binary Pattern (RIULBP), standard derivation, skewness, kurtosis, entropy, energy and mean. In our implementation of [5], patch size is the same as that for our implemenation of [3]. SFFS is used for feature selection. Various combinations of distance parameter, $R$, (in SFFS) are used in the experiment. For comparison testing, grid search was used to determine SVM parameters for the methods in [3] and [5]. For our method, weak learners (each SVM) are constructed using SFFS to change the combinations of features selected. Rounds of AdaBoost on these weak learners were used to classify the data set. Ensemble learning results in better performance.

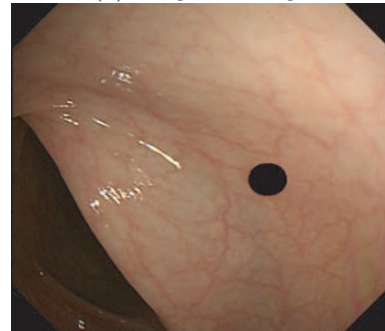Accuracy, specificity and sensitivity are defined by Eqs.(6–8), respectively.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Number\ of\ Positives + Number\ of\ Negatives}$$
(6)

$$Specificity = \frac{Number\ of\ Correct\ Negative\ Predictions}{Number\ of\ Negatives}$$
(7)

$$Sensitivity = \frac{Number\ of\ Correct\ Positive\ Predictions}{Number\ of\ Positives}$$
(8)
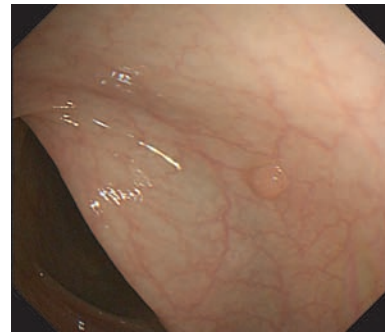


(a) Original Image



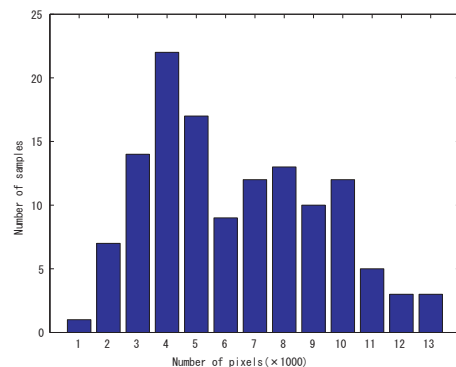(b) Labeled Image

Figure 2. Example Image



Figure 3. Sizes of Polyps in the Data Set

Results comparing our work to the methods described in [3] and [5] are shown in Table 2. Patch size and parameters chosen for use with the methods in [3] and [5] were those which gave the highest accuracy among the many combinations tried in our experiments. Results for three variants of our method also are shown in Table 2, using different combinations of statistical features obtained from the RGB intensity values and HOG. Features A used intensity+HOG, Features B used intensity alone. Features C used HOG alone. The best results from Table 2 are those for Features A, using the combination of statistical features from both the RGB intensity values and HOG. All three of our variants demonstrate better performance than the two previous methods considered.

Table 2. Classification Results (Comparison)

|  | Sens. [%] | Spec. [%] | Acc. [%] |
|---|---|---|---|
| Paper [3] | 76.8 | 65.7 | 71.3 |
| Paper [5] | 85.4 | 76.5 | 80.9 |
| Features | Sens. [%] | Spec. [%] | Acc. [%] |
| Features A | 93.9 | 94.3 | 94.1 |
| Features B | 91.6 | 90.2 | 91.0 |
| Features C | 86.2 | 85.4 | 85.8 |

Table 3. Result of Classification Using Boosting

| Rounds | Sens. [%] | Spec. [%] | Acc. [%] |
|---|---|---|---|
| 1 | 94.9 | 95.8 | 95.4 |
| 2 | 94.9 | 95.8 | 95.4 |
| 3 | 94.4 | 98.1 | 96.3 |
| 4 | 95.4 | 97.2 | 96.3 |
| 5 | 95.4 | 98.2 | 96.8 |
| 6 | 95.4 | 98.2 | 96.8 |
| 7 | 95.4 | 98.2 | 96.8 |
| 8 | 95.4 | 98.2 | 96.8 |
| 9 | 95.4 | 98.2 | 96.8 |
| 10 | 95.4 | 98.2 | 96.8 |
| 11 | 95.4 | 98.2 | 96.8 |
| 12 | 95.4 | 98.2 | 96.8 |
| 13 | 95.4 | 98.2 | 96.8 |
| 14 | 95.4 | 98.2 | 96.8 |
| 15 | 95.4 | 98.2 | 96.8 |

AdaBoost is used to improve accuracy further. At each round, we use the SVM which gives the highest accuracy under different combinations of SVM parameters. Table 3 show the result for 15 rounds of AdaBoost. With AdaBoost, 96.8% accuracy is achieved. Two actual results of our polyp detection with AdaBoost are shown in Fig.4-(a) and Fig.4-(b). Regions outlined by an ellipse-like shape are the regions classified as polyps.
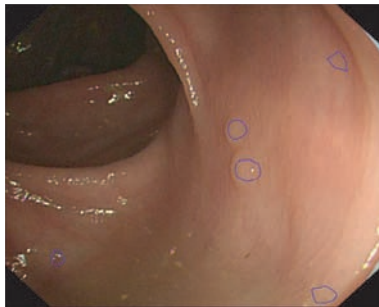


(a) Result 1



(b) Result 2

Figure 4. Examples of Polyp Detection

## 5 Conclusion

This paper demonstrates a new approach to automatic polyp detection. Candidate regions are extracted using a Hessian filter. We first improvement to detect the polyp candidate region. Judgment of polyp region or not is based on SVM classifier by changing combination of SVM parameters and feature selection. It was also suggested that Using each SVM as a weak classifiers and constructing Adaboost further improve the decision of polyp region. Based on the proposed approach, high accuracy is achieved, as demonstrated via experiments with real endoscope images.

Further improvement is remained for the small misjudged detections as the future subject.

## References

[1] D.K. Iakovidis, D.E. Maroulis, S.A. Karkanis, A. Brokos, "A Comparative Study of Texture Features for the Discrimination of Gastric Polyps in Endoscopic Video", 18th IEEE Symposium on Computer-Based Medical Systems, pp.575-580, IEEE Computer Society, 2005.

[2] S.A. Karkanis, D.K. Iakovidis, D.E. Maroulis, D.A. Karras, M. Tzivras, "Computer-Aided Tumor Detection in Endoscopic Video Using Color Wavelet Features", IEEE Trans. on Info. Technology in Biomedicine, Vol.7, No.3, pp.141-152, 2003.

[3] L.A. Alexandre, N. Nobre, J. Casteleiro, "Color and Position versus Texture Features for Endoscopic Polyp Detection", Int. Conf. on BioMedical Engineering and Informatics (BMEI), Vol.2, pp.38-42, 2008.

[4] A.F. Frangi, W.J. Niessen, K.L. Vincken, M.A. Viergever, "Multiscale vessel enhancement filtering", Medical Image Computing and Computer-Assisted Intervention (MICCAI), LNCS, Vol.1496, pp.130-137, 1998.

[5] B.P. Li, M.Q. Meng. "Comparison of Several Texture Features for Tumor Detection in CE Images", Journal of Medical Systems, 2011.

[6] A. Cheddad, D. Mohamad and A. Manaf, "Exploiting Voronoi diagram properties in face segmentation and feature extraction", Pattern Recognition, Vol.41, No.12, pp.3842-3859, 2008.

[7] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection", IEEE CVPR, Vol.1, pp.886-893, 2005.

[8] P. Pudil, J. Novovicova, J. Kittler, "Floating search methods in feature selection", Pattern Recognition Letters, Vol.15, No.11, pp.1119-1125, 1994.

[9] Y. Freund, R. E. Schapire, "Experiments with a New Boosting Algorithm", ICML, pp.148-156, 1996.