

Detecting Abnormal Regions in Colonoscopic Images by Patch-based Classifier Ensemble

Peng Li[†] Kap Luk Chan Shankar Muthu Krishnan Yan Gao
Biomedical Engineering Research Center
Nanyang Technological University
Research TechnoPlaza, 50 Nanyang Drive, XFrontiers Block, Singapore
[†] lipeng@pmail.ntu.edu.sg

Abstract

In this paper, a new method is proposed to detect abnormal regions in colonoscopic images by patch-based classifier ensemble. Through supervised learning from image patches of various sizes, a set of basic SVM classifiers is trained for each size. A diagnostic model can then be constructed based on the ensemble of basic classifiers which is then used to detect abnormal regions in colonoscopic images. The multiple sizes of patches provide multiple level representation of the image content, which can help improve detection results. Several fusion criteria are explored to determine the final output of the ensemble. Experimental results show promising performance of our proposed method.

1. Introduction

At present, the analysis of colonoscopic images and diagnosis of abnormalities that may manifest cancerous growth relies on the experiences and expertise of the medical experts, which need years of training. Therefore it is justified to develop a computer-assisted technique for screening of this potentially lethal disease. Previous research on colonoscopic image analysis focuses on the classification between normal tissues and tumors. Tjoa *et al* employed back-propagation neural network for colon status classification and proposed using principal component analysis to improve computation efficiency [9]. Since they used full-size colonoscopic images for the classification, classification error can be resulted when the size of abnormal region is small. Maroulis *et al* developed a detection system for colorectal lesions in endoscopy video-frames using multi-layer neural network as a classifier [5]. Recently Karkanis *et al* employed a linear discriminant analysis procedure to detect tumors in endoscopic video using wavelet features [3]. However, to our best knowledge, none of them attempt

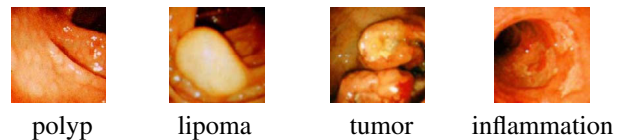


Figure 1. Colonoscopic images with different abnormal regions.

the discrimination between normal tissues and all different kinds of abnormalities. In fact, there are many kinds of abnormalities that can be seen in endoscopic images, such as polyps, tumors, inflammation, bleeding, ulceration and diverticula, etc.(Figure 1), and their image content show large variations. Furthermore, the normal regions of colonoscopic images also show large variations. The abnormal regions often do not occupy the whole image and are of various color, sizes and shapes. All these make the discrimination between normal and abnormal regions in colonoscopic images a tough job. The abnormal region detection problem in colonoscopic images can be regarded as a perceptual image segmentation problem. Many methods have been developed for computer-assisted segmentation of medical images, such as thresholding, region growing and classification based methods etc [6]. However, almost all of these methods partition the pixels in an image based on low level cues, such as illumination, color, texture, and the segmentation often disagree with the way human-beings partition the image. Therefore, these methods cannot produce a perceptually “good” segmented image. Recently, there are some attempts to segment the images by incorporating higher level knowledge. Shi *et al.* treated image segmentation as a graph partitioning problem and proposed a global criteria called “Normalized Cuts” for segmenting the graph [8]. Motivated by the Normalized Cuts, Ren *et al.* presented a binary classification model for segmentation [7]. Instead of using a single pixel for segmentation, they grouped pixels

into “superpixels”, which are roughly homogenous in shape and size. Then, these superpixels are grouped into different segments using a classifier. These superpixels provide a higher level cue for segmentation, which can produce better segmentation results than that of using single pixels. Jojic *et al.* proposed another intermediate appearance and shape model between pixel and the whole image called “epitomes” [2], which is a miniature, condensed version of an image containing the essence of the textural and shape properties of the image. They used some square patches as epitomes. The patch-based approach seems to be a good representation model for image segmentation, but choosing the size of the patches is still an open problem.

Patch-based approach has also been investigated by Karkanis *et al* for polyp and tumor detection in colonoscopic images [3]. They tried several patch sizes separately and chose the one with the least detection error rate to detect polyps and tumors in colonoscopic images. However, the sizes of the polyps and tumors are different and the shapes of the tumors are often irregular. Therefore, using one size patch may not be suitable for all types of abnormalities. Large variations in sizes and shapes of the abnormal regions makes current abnormal region detection problem even harder .

In this paper, we propose a new method of abnormal region detection using patch-based Support Vector Machine (SVM) classifier ensemble learning. Multiple sizes of patches provide multiple level cues of the abnormal regions, which can help produce more perceptually agreeable segmentation. The experimental results show that the multi-size patch-based approach can achieve good accuracy rate of abnormal region detection in colonoscopic images.

2. Learning Classification for Abnormal Region Detection

Figure 2 illustrates the flowchart of the proposed method for abnormal region detection in colonoscopic images. The details are as follows.

2.1. Image Preprocessing and Feature Extraction

The original endoscopic images obtained by the endoscopy system are *RGB* images with the resolution of 256×256 pixels. They are transformed into 3 bands in *CIELab* color space through which the color and luminance component can be processed and analyzed individually. Each band is scanned across and cropped into 3 fixed sizes of patches respectively. The patches are overlapped by 50% to ensure no abnormal region is missed. Here we investigated 3 sizes of the image patches for abnormal region detection in the colonoscopic images, namely, 48×48 , 32×32 and 16×16 (pixels). Feature can be extracted

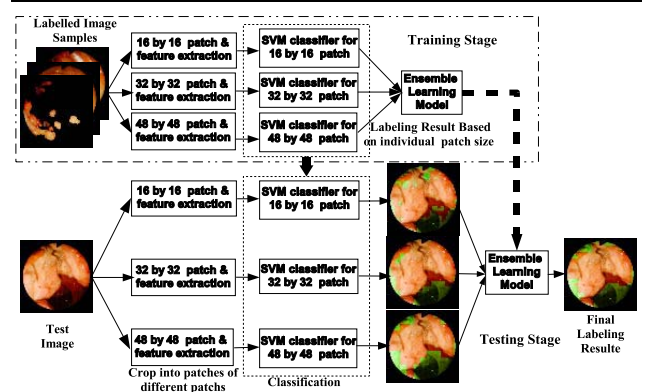


Figure 2. The flowchart of proposed method.

from these image patches for classification. The features extracted here include 1-dimensional histograms of luminance (*L*) and 2-dimensional histograms of the components *a* and *b* in *CIELab* color space. The number of bins of the histogram are chosen as 16 for 1-D histograms and 8 for 2-D histograms. Two-level Discrete Wavelet Transform (DWT)-based statistical features are employed as well, namely, the mean and standard deviation of the absolute value of the approximate and detailed coefficients from the two-level DWT decomposition of the image patches in the 3 channels of the *CIELab* color space. Altogether 128 features are extracted, giving rise to a feature vector of 128 dimension. Then the feature vectors from patches can be used to form the data set for classification. All the feature vectors, $X = \{x_i, i = 1, 2, \dots, N\}$ for *N* patches, are labelled as $y_i = \pm 1$ to indicate whether it is a normal patch or a patch containing abnormalities.

2.2. The Proposed Method

2.2.1. Classification using SVM SVM is increasingly used in many medical applications and has been shown to achieve higher performance than traditional classifiers [1]. SVM has good generalization by finding an optimal separating hyperplane which minimizes the classification errors made on the training set while maximize the “margin” between different classes. Given a training set *X* with *N* samples, the optimal separating hyperplane can be expressed as:

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x) + \rho \quad (1)$$

where $K(\cdot)$ is a kernel function, ρ is a bias, α_i are the solutions of a quadratic programming problem to find the maximum margin. There are only a few training samples whose α_i s are non-zero. They are called the support vectors, which

are either on or near the separating hyperplane. The decision boundary, i.e. the separating hyperplane is along these support vectors, whose decision values $f(x)$ (equation 1) approach zero. Compared with the support vectors, the decision value of positive samples have larger positive values and those of negative samples have larger negative values. Therefore, the magnitude of the decision value can also be regarded as the confidence of the SVM classifier. The larger the magnitude of $f(x)$ is, the more confident the classification is.

Using overlapped image patches, each pixel in the patch can be classified as normal or abnormal by a SVM classifier for the respective patch size and band. Thus each pixel in the original image can have at least one label. If a pixel gets different labels, the label of the patch that has the largest absolute value of decision value (confidence) is chosen as the final label of that pixel. Those pixels labelled as -1 are then detected as belonging to abnormal regions.

2.2.2. Patch-Based Classifier Ensemble We have mentioned in previous section that patch-based approach is suitable for abnormal region detection problem. However, the problem of what patch to use has to be solved first since one patch size is not suitable for representing all the abnormalities which show large variations in terms of colors, sizes and shapes. The use of multiple patch sizes seems to be a feasible solution because multiple patch sizes can provide multiple cues. Here, we use one SVM classifier for each patch size. If three patch sizes are used, there are 3 SVM classifiers. The key is to choose suitable criteria to fuse the information from these multiple patch sizes.

There are several ensemble approaches for combining these individual predictions, such as majority voting, max rule, etc. [4]. Given M ($M = 3$) individual classifiers with decision value f_i and label y_i ($y_i = \pm 1$), the goal is to find the final prediction y_{en} . We investigated 3 fusion rules in the experiment which are easy to implement because no further training is required. One is the weighted average of the decision value of M individual classifiers:

$$y_{en} = \text{sgn}\left(\sum_{i=1}^M w_i f_i\right) \quad (2)$$

where weight w_i is based on the average classification accuracy rate of the i th individual SVM classifier. The second is max rule: $y_{en} = y_j$, where

$$j = \arg \max_i |f_i| \quad (3)$$

The third rule is majority voting, in which the final label is the one that majority of the individual classifiers agrees. For an odd number M , the final decision is simply:

$$y_{en} = \text{sgn}\left(\sum_{j=1}^M y_i\right) \quad (4)$$

| Patch Size | 16×16 | 32×32 | 48×48 |
|------------|----------------|----------------|----------------|
| SVM | 76.3% | 81.7% | 86.1% |
| MFNN | 74.4% | 79.1% | 82.9% |

Table 1. Results of classification.

3. Experimental Results

3.1. Classification Based on Single Patch Size

We compared the performance of using SVM with Gaussian kernel and the widely used Multi-layer Feed-forward Neural Network (MFNN) [9][5] for discrimination between abnormal and normal regions in colonoscopic images. In the experiment, we used 58 colonoscopic images, including 12 normal ones and 46 with abnormal regions. The numbers of collected image patches of size 48×48 , 32×32 and 16×16 are 2000, 2097 and 2124 respectively. The pixels in the original image are manually labelled to provide ground truth. The patch containing mostly abnormal region is labelled as a negative sample, otherwise, it is labelled as a positive sample. A leave-one-out experiment was performed for the classification. Each time, the image patches from one of the original colonoscopic images were selected as a testing set and the patches from the other 57 images were used for training. The experiment was performed 58 times, and the average of the accuracy rates of total 58 results was taken as the final accuracy rate of the classification of patches. The results are listed in Table 1. This classification rate is also used to define w_i .

Obviously, SVM achieved higher classification accuracy rate than MFNN. In addition, we repeat the leave-one-out experiment 5 times for 16×16 patches, the average accuracy rate of MFNN is $73.5 \pm 7.4\%$, while that of SVM is $76.3 \pm 0\%$. It shows that the stability of SVM is also better than MFNN. Therefore, SVM is justified to be used as the basic classifier for the abnormal region detection in colonoscopic images.

3.2. Detecting Abnormal Region Using Learned Classifier Ensemble

A leave-one-out experiment is performed to detect abnormal regions in colonoscopic images using the proposed ensemble learning method. The detection results are compared with the manually labelled ground truth image. The average detection accuracy rate is reported in Table 2.

The majority voting scheme achieved slightly higher average classification rate than weighted voting and max rule. This may be due to the fact that the 3 basic SVM classifiers are based on different patch size or different level of representation of the model and they are not independent to each other. The maximum confidence level of 3 basic SVM clas-

| Ensemble Schemes | Average Detection Accuracy Rate |
|------------------|---------------------------------|
| Weighted Voting | $82.2 \pm 11.9\%$ |
| Max Rule | $81.4 \pm 11.9\%$ |
| Majority Voting | $83.4 \pm 12.0\%$ |

Table 2. Results of abnormal region detection.

sifier doesn't mean large confidence, because all the confidence level of 3 classifiers may be low. Therefore, max rule doesn't work well. The same explanation applies to the weighted voting scheme. However, if 2 out of 3 classifiers say "abnormal", the probability of the pixel belonging to "abnormal" is more probably greater than that of "normal". Therefore, the majority voting scheme shows better performance.

The detection results on 4 colonoscopic images are reported in Figure 3. Row A are the original images. Row B are the manually labelled abnormal regions, i.e. the ground truth. Row C~E are labelling results using single patch size 16×16 , 32×32 and 48×48 . Row F~H are the detection results by the proposed method using weighted voting, max rule and majority voting respectively. The regions in green are normal regions detected and the regions in original color are abnormal. The abnormal regions detected by using 3 patch sizes separately are different. By combining the 3 predictions, the final decision is better than the decision of using single patch size. Obviously, the abnormal regions detected using ensemble learning scheme are more agreeable with the ground truth than using single patch only, among which majority voting rule produces the best results.

4. Conclusion

A new method of abnormal region detection in colonoscopic images is proposed. Using labelled colonoscopic image patches of 3 sizes, 3 basic SVM classifiers are trained for each patch size. These classifiers can produce 3 decisions for each pixel in an colonoscopic image. The final decision is through the ensemble learning of these 3 basic SVM classifiers. Due to the use of multiple sizes of image patches, the proposed method can produce more perceptually agreeable detection results. Among all of the ensemble learning schemes, majority voting shows slightly better performance. The detection results show the promising performance of the proposed method.

References

- [1] N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines*. Cambridge University Press, 2000.
- [2] N. Jojic, B. Frey, and A. Kannan. Epitomic analysis of appearance and shape. In *Proceedings of Ninth IEEE International*

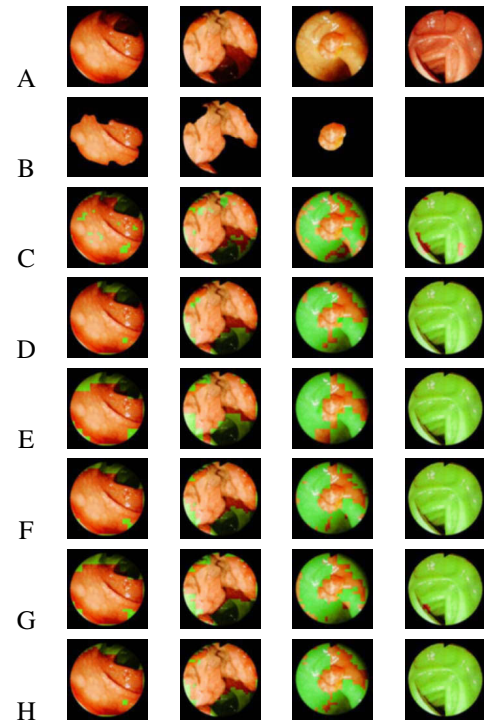


Figure 3. Detection results of 4 colonoscopic images.

- Conference on Computer Vision (ICCV 2003)*, pages 34–41, Nice, France, October 2003.
- [3] S. Karkanis, D. Iakovidis, D. Maroulis, D. Karras, and M. Tzivras. Computer aided tumor detection in endoscopic video using color wavelet features. *IEEE Transactions on Information Technology in Biomedicine*, 7(3):141–152, September 2003.
- [4] J. Kittler, M. Hatef, R. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, March 1998.
- [5] D. Maroulis, D. Iakovidis, S. Karkanis, and D. Karras. Cold: a versatile detection system for colorectal lesions in endoscopy video-frames. *Computer Methods and Programs in Biomedicine*, 70(2):151–166, February 2003.
- [6] D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2:315–337, 2000.
- [7] X. Ren and J. Malik. Learning a classification model for segmentation. In *Proceedings of Ninth IEEE International Conference on Computer Vision (ICCV 2003)*, pages 10–17, Nice, France, October 2003.
- [8] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, August 2000.
- [9] M. P. Tjoa and S. M. Krishnan. Feature extraction for the analysis of colon status from the endoscopic images. *BioMedical Engineering Online*, 2(9), April 2003.