Polyp Detection in Endoscopic Video Using SVMs

Luís A. Alexandre[1,2], João Casteleiro[1], and Nuno Nobre[1]

[1] Department of Informatics, Univ. Beira Interior, Portugal
[2] IT - Networks and Multimedia Group, Covilhã, Portugal

Abstract. Colon cancer is one of the most common cancers in developed countries. Most of these cancers start with a polyp. Polyps are easily detected by physicians. Our goal is to mimic this detection ability so that endoscopic videos can be pre-scanned with our algorithm before the physician analyses them. The method will indicate which part of the video needs attention (polyps were detected there) and hence can speedup the procedures. In this paper we present a method for polyp detection in endoscopic images that uses SVM for classification. Our experiments yielded a result of $93.16 \pm 0.09\%$ of area under the Receiver Operating Characteristic (ROC) curve on a database of 4620 images indicating that the approach proposed is well suited to the detection of polyps in endoscopic video.

1 Introduction

A polyp is an abnormal growth of tissue projecting from a mucous membrane.

In this paper we are concerned with polyps in the colon. An example is presented in figure 1. Polyps are important since they can, with time, turn into colon cancer. The cumulative risk of cancer developing in an unremoved polyp is 2.5% at 5 years, 8% at 10 years, and 24% at 20 years after the diagnosis [7]. If detected on an early stage these polyps can be easily removed.

In Portugal there are six thousands people per year diagnosed with the disease of colon cancer[1]. However, this disease is also the most tractable of all the digestive cancers when diagnosed at an early stage. This cancer is one of the most fatal illness all over the world.

Our interest in the video processing approach comes from the fact that there is a new medical examination, where the patient ingests a capsule (with the form of a pill) that films the digestive tube (video capsule endoscopy). The video is recorded in a device that the patient carries, usually in the belt. This video is then screened by the physician to search for polyps (and possibly other illnesses). Our goal is to develop a method that can be applied to the resultant video and avoid the time necessary to completely screen these videos. The idea is that if our systems detects polyps, the physician will focus on the signalled portions of the video with urgency. Videos where polyps aren't detected will be left for latter processing by the physician (have a lower priority).
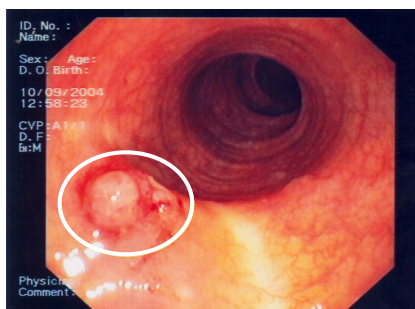
Fig. 1. An endoscopic image with a polyp marked by the white ellipse

Given the importance of colon polyps, some researchers have developed methods for its automatic detection. There are basically two approaches for this: video processing and CT (Computed Tomography) image processing [4,3,5]. Although the approaches based on CT images are able to produce a virtual representation of the colon which can speed up the visual analysis by the physician, they also have some disadvantages: the extensive amount of radiologist time (during CT scanning) involved in the process, the radiation that the patient is subjected to and the cost of such an exam. Given our motivation presented above, we are more interested in the video based approaches. We will now briefly describe some of the work done under this approach.

In [6] a comparative study of texture features for the detection of gastric polyps in endoscopic video was presented. Of the four approaches tested, texture spectrum histogram, texture spectrum and color histogram statistics, local binary pattern histogram and the color wavelet covariance, this last one presented the best results with an area under the ROC curve value of 88.6%.

In [8] the authors presented new approaches for extracting texture- and color-based features from colonoscopic images for the analysis of the colon status. Note that the abnormal status can be due to pathologies other than polyps. They used Principal Component Analysis (PCA) for feature selection and Backpropagation Neural Networks for classification. They found that using texture and color features improved classification results when compared to using only one type of feature.

In [9] the authors were also concerned with abnormality detection from endoscopic images. They use a fusion approach to reach a final decision from sub-decisions made based on associated component feature sets. They report that the overall detectability of abnormalities using the fusion approach is improved when compared with corresponding results from the individual methods.

In this paper we show that, given the SVM's ability to deal with high dimensional input spaces, we can produce very interesting results in terms of polyp detection in endoscopic video images by using only color and pixel position information, without any further feature extraction or selection technique.

The paper is organized as follows: the next section presents our method for polyp detection. Section 4 presents the experiments and the final section contains the conclusions.
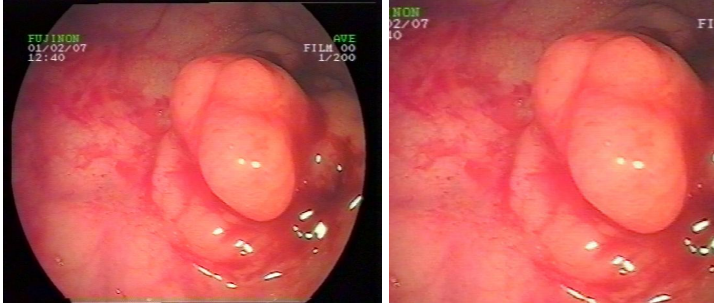
Fig. 2. Image before and after the operation that removes the black frame

2   Polyp Detection

The goal is to detect polyps on colonoscopic images similar to the one in figure 2. Our approach to this problem is the following: first we pre-processed the images to retain only the image portion that contains relevant information; then we subdivided each of the original images into sub-images of 40×40 pixels. Then we applied the feature extraction algorithms to these sub-images. Finally a classifier (SVM) was used to make the decision about the existence or not of a polyp in an image.

We will now describe these operations in more detail.

2.1   Pre-processing

The videos were captured with PAL (768×576) resolution. The frames have a black frame around the useful region of image as in figure 2. This black frame is removed leaving each image with a resolution of 514×469. This approach discards some of the useful area, but since we are working with video, we can recover the lost data from other video frames.

2.2   Image Division and Tagging

Our approach considered the division of the original images into smaller sub-images, that is, we will not classify directly an input image but, we subdivide it and classify each sub-image individually as containing a polyp or not. Then this information is used to classify the original image.

The idea consists in processing sub-images that can sometimes be completely contained within the polyp region. This means that ideally we should use sub-images of the size of a single pixel. Of course this would not produce enough data to have statistically significant results on the sub-image level. So, we define the sub-image area with dimensions of 40×40 pixels. This is small enough such that the sub-images are frequently completely contained in the polyp region but are also big enough to produce significant feature results.
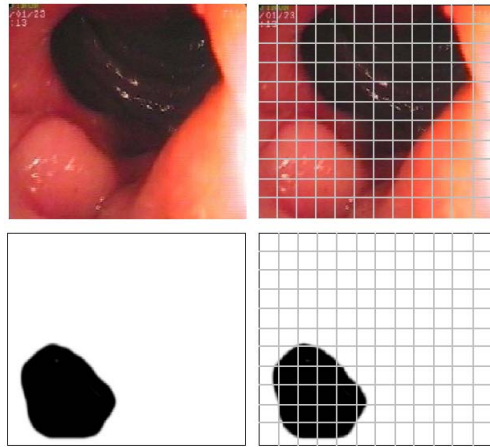
Fig. 3. From left to right and top to bottom: original image, its subdivision, the corresponding manual classification mask (black means polyp) and its subdivision

The sub-images were obtained by sliding a window with a 40 pixels step, both horizontally and vertically through the original image. This means that there is no overlap between the sub-images.

Given the dimensions of the input images after the black frame removal, the sub-division process generates 132 sub-images for each original image.

To simplify (automate) the manual classification of each of these sub-images we produced a binary image that was used as a classification mask. This mask is a manual painted image the size of the original image, that has the polyp region painted black and the remaining portion is white (an example is shown in figure 3). This painting was checked by a grastroentrologist.

The subdivision of the original image into sub-images is done also on the classification mask, yielding a sub-image that contains only black and white pixels. To assign a class label to the original sub-images we look at the corresponding classification mask sub-image and count the number of black pixels it contains. If this number is higher than a given threshold, λ, we consider that the sub-image 'contains' a polyp. This process allows the automatic classification of the sub-images.

2.3   Choosing the Sub-Image Classification Threshold λ

The value of λ can vary between 1 and 1600 (the total number of pixels of the sub-image). Naturally, as λ gets smaller, more sub-images are classified as containing polyps. This might look like a good option so that the system has a smaller false negative rate. But since these images are used for training the system, if we choose to classify a sub-image with few black pixels as a polyp, we are using very little real polyp information in that image to teach what a polyp is. In fact, if we choose λ smaller than 800, we may give more non-polyp than polyp information in a sub-image.

We studied the influence of λ in the classification results. These results are presented in the experiments section but we can say that the best results are obtained for a larger λ.

2.4   Feature Extraction

Out approach to feature extraction is quite simple and produced very interesting results. Given the capabilities of the SVMs in dealing with high-dimensional input data, we chose as features only color and position information for each pixel. Each pixel in a sub-image is represented by five values: its RGB components and its coordinates in the sub-image. So each sub-image is in fact represented by a total of 8000 features: 5 features for each of the 1600 pixels.

These features were reduced and centered in the corresponding training and test sets (see below).

3   Experiments

3.1   Dataset

The dataset consists of 35 video frames obtained with a Fuji 410 video endoscope system at the Hospital Cova da Beira, Portugal, during the year of 2007. The images were subdivided into smaller images after the pre-processing described in section 2.1. Each image produced 132 sub-images. Each sub-image was defined as polyp or not polyp according to the correspondent sub-image obtained from the manual generated classification mask described in section 2.2. The resulting data set contained 4620 images each with a dimension of 40×40 pixels.

The features were centered and reduced such that, for each feature, the mean value is 0 and the standard deviation is 1.

3.2   Classifier

The classifier used was a support vector machine (LIBSVM) [2]. The kernel type used was the radial basis function (RBF):

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma ||\mathbf{x}_i - \mathbf{x}_j||^2), \ \gamma > 0 \qquad (1)$$

where $\mathbf{x}_i$ is an input and γ is a parameter inversely proportional to the kernel width. The SVM with this kernel has two free parameters to be set: $C > 0$ that corresponds to the penalty parameter of the error and γ. The values for these parameters are discussed below.

A different weight can be assign to each class when the prior probabilities for each class are not equal. This is the case in our dataset given that there are more non-polyp images than polyp ones. The exact proportion depends on the threshold λ used. This weighting was done using the `svm-train` parameters w_0 and w_1. This is also discussed below.

Table 1. Number of images considered polyp for each threshold λ

λ	100	300	500	700	900	1100	1300	1500
N. of polyps	966	900	850	809	770	716	682	609
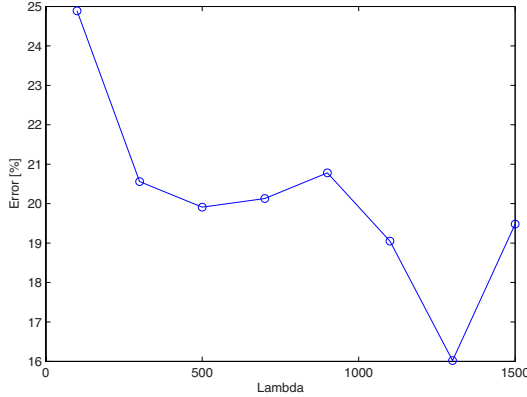
Fig. 4. Error for different values of λ

3.3   Evaluating the Effect of λ

In this section we present the results of experiments made to evaluate the effect of the value of λ used in the automatic classification of the sub-images, as discussed in section 2.3.

We varied λ from 100 to 1500 in steps of 200. Table 1 presents the number of sub-images that are classified as polyp given the value of λ considered.

For each value of λ a grid search was performed using half the dataset to find the best values of C and γ parameters for the SVM classifier. The search was done by varying C from 1 to 64 in integer powers of 2, and γ from 2^{-6} to 2^{-16} also in (negative) integer powers of 2.

We created a subset of with 10% of the images (462) randomly selected from the full image set. The error in this subset was evaluated with 2-fold cross-validation method. The results obtained are shown in figure 4. It can be seen that the smallest error, 16.02%, was obtained for $\lambda = 1300$. The correspondent value of C was 32 and $\gamma = 0.0001$. Given these results we decided to use $\lambda = 1300$ for our subsequent experiments.

3.4   Results on the Full Dataset

We evaluated the error using the 2-fold cross-validation method on the 4630 images, using $\lambda = 1300$ and $C = 32$. We experimentally found that good values for the class weights in this case are $w_0 = 1$ and $w_1 = 5$. The value of γ was varied to produce several points on the ROC curve. Figure 5 contains the ROC plot.
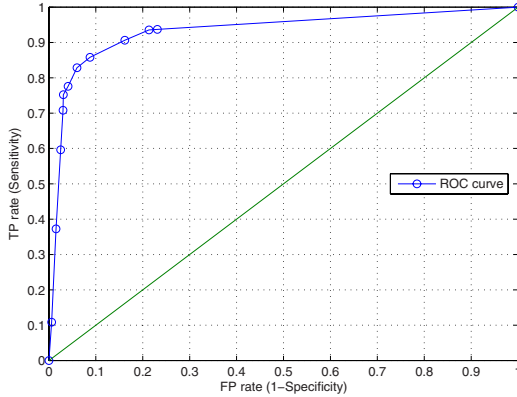
Fig. 5. ROC curve for $\lambda = 1300$. FP stands for False Positive and TP for True Positive.

The area under the Receiver Operating Characteristic (ROC) curve, (AUC), is $93.16 \pm 0.09\%$. The results show that we can get a value of False Negative Rate (FNR) of $6.31 \pm 1.04\%$ at $23.11 \pm 0.93\%$ of False Posite Rate (FPR).

4 Conclusions

In this paper we presented a method for polyp detection in endoscopic videos. The goal was to be able to develop a method that could do a first automatic screening of a endoscopic video and warn the physician of frames where attention is needed. Videos in which the method does not detect any polyp can perhaps be given a smaller priority then others where possible polyps are detected.

Our method subdivides each image into smaller images (with 40×40 pixels). These are the images that are searched for polyps. Of course if a subimage is considered a polyp, its parent image is also considered to have a polyp. (Other approaches can be used like the need for a number of detected subimages with polyp in order to consider that the parent image contains a polyp. Our approach is the most cautious: it perhaps implies the existence of some false positives but will minimize the false negatives). We did not analyse this aspect in this paper though. We focused on the correct classification of the subimages.

We used a very simple approach for feature extraction, relying only on color and pixel position. Although it creates many features per image (8000) the SVM was able to deal with this high dimensionality. It can take longer for the training phase, but may decrease the time taken when processing a new image with the trained system since the feature extraction does not involve many computations.

The results we obtained are quite satisfactory: in a database with 4620 images we were able to obtain an AUC value of $93.16 \pm 0.09\%$, a sensitivity of 93.69 ± 1.04 % at $23.11 \pm 0.93\%$ of FPR.

Future work will concern the application of this method to video images obtained from video capsule endoscopy instead of video from a colonoscope. Other future challenges include the identification of the type of polyp and its development stage.

Acknowledgements

References

1. Rui Cernadas. Algumas reflexões sobre o cancro colorrectal. In: Endonews (2004)
2. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines, Software (2001), available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`
3. Chowdhury, T.A., Ghita, O., Whelan, P.F.: A statistical approach for robust polyp detection in CT colonography. In: 27th Annual International Conference of the Engineering in Medicine and Biology Society, pp. 2523–2526. IEEE Computer Society Press, Los Alamitos (2005)
4. Gokturk, S.B., Tomasi, C., Paik, D., Beaulieu, C., Napel, S.: A learning method for automated polyp detection. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 85–92. Springer, Heidelberg (2001)
5. Huang, A., Summers, R.M., Hara, A.K.: Surface curvature estimation for automatic colonic polyp detection. In: Amini, A.A., Manduca, A. (eds.) Medical Imaging 2005: Physiology, Function, and Structure from Medical Images. Proceedings of the SPIE, vol. 5746, pp. 393–402 (2005)
6. Iakovidis, D.K., Maroulis, D.E., Karkanis, S.A., Brokos, A.: A comparative study of texture features for the discrimination of gastric polyps in endoscopic video. In: 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05), pp. 575–580. IEEE Computer Society, Los Alamitos (2005)
7. Stryker, S.J., Wolff, B.G., Culp, C.E., Libbe, S.D., Ilstrup, D.M., MacCarty, R.L.: Natural history of untreated colonic polyps. Gastroenterology 93(5), 1009–1013 (1987)
8. Tjoa, M.P., Krishnan, S.M.: Feature extraction for the analysis of colon status from the endoscopic images. BioMedical Engineering OnLine 2(9) (2003)
9. Zheng, M.M., Krishnan, S.M., Tjoa, M.P.: A fusion-based clinical decision support for disease diagnosis from endoscopic images. Computers in Biology and Medicine 35(3), 259–274 (2005)