

Texture-based Polyp Detection in Colonoscopy

Stefan Ameling¹, Stephan Wirth¹, Dietrich Paulus¹, Gerard Lacey²,
Fernando Vilarino²

¹Institute for Computational Visualistics, University of Koblenz-Landau, Germany

²Department of Computer Science, Trinity College Dublin, Ireland

stameling@uni-koblenz.de

Abstract. Colonoscopy is one of the best methods for screening colon cancer. A variety of research groups have proposed methods for automatic detection of polyps in colonoscopic images to support the doctors during examination. However, the problem can still not be assumed as solved. The major drawback of many approaches is the amount and quality of images used for classifier training and evaluation. Our database consists of more than four hours of high resolution video from colonoscopies which were examined and labeled by medical experts. We applied four methods of texture feature extraction based on Grey-Level-Co-occurrence and Local-Binary-Patterns. Using this data, we achieved classification results with an area under the ROC-curve of up to 0.96.

1 Introduction

1.1 Medical Background

Colon cancer is one of the leading causes of cancer death worldwide [1]. Although the interest on virtual endoscopy from computed tomography images is growing, colonoscopy is still the accepted gold standard for detecting colon cancer or its precursor lesions, colorectal polyps. The removal of adenomatous polyps has been shown to prevent most cancer. Although colonoscopy has become the preferred method of screening colon cancer, several studies showed that it contains an inherent miss rate of up to 25% for the detection of polyps and cancer, depending on its size and location (e.g. [2]).

Two main causes that lead to misses can be distinguished:

1. The endoscopist does not cover all parts of the colon (parts remain "unseen" by the camera),
2. the polyp appears on the screen but remains undetected by the doctor.

Most work – as well as ours – has been done on the second problem, the computer-assisted detection. Fig. 1 shows some colonoscopic images containing polyps. Artifacts like reflexions and non-uniform illumination complicate the computer vision task.

1.2 Related work

Polyp detection methods can be classified by the types of features that are used: shape features, texture features, color, or combined approaches.

For texture-based analysis, different sets of texture features are used which are extracted from the Grey-Level Co-occurrence Matrix (GLCM), the Local Binary Pattern (LBP), the Opponent Colour-Local Binary Pattern (OC-LBP), or Wavelet Transform (WT) (e.g. [3, 4]).

Wang et al. use a two-dimensional histogram for classification which combines LBP with average intensity [5].

Reviews of polyp detection methods have shown that many research results can be called into question due to the fact that the sample sets used for testing and training were relatively small, except for a few cases [6, 7].

A reliable classification system should be based on a huge set of high-resolution images containing many different types of abnormalities. We think that a classifier trained with less than hundred images at VHS-resolution or even lower cannot detect cancer and polyps in a way that meets the requirements of a clinical system.

Apart from the video data itself, ground truth information about the abnormalities present in the images is needed for training and evaluation of a classifier. Vilariño et al. proposed a promising architecture for automatic labeling of colonoscopy video for cancer detection using eye movement of the observer, where results from gaze tracking were used to select regions in the images [8].

2 Material and methods

2.1 Database

Due to the technological progress in terms of acquisition of high-resolution images more detailed texture structures of the colon mucosa become visible that can be utilized for the detection of colorectal lesions such as polyps.

More than four hours of almost complete colonoscopies were captured in Full HD resolution (1920×1080 pixels). After preselection, the videos were presented to five experts in colonoscopy to identify abnormalities. By clicking with the mouse on the relevant positions in the video, ground truth information was obtained. Afterwards, for every image a mask was created manually that labels abnormal regions (Fig. 1).

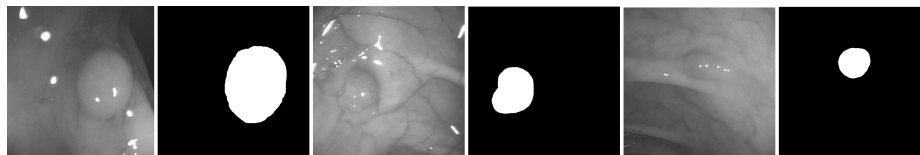


Fig. 1. Colonoscopy images with their corresponding masks labeling abnormal regions

2.2 Methods

We used two common texture analysis methods for texture feature extraction: grey-level co-occurrence matrices (GLCMs) [9] and local binary patterns (LBP) [10]. GLCMs describe how often different combinations of pixel values occur in an image. Each entry (i, j) of the matrix p represents the probability of going from one pixel value i to another pixel value j under a predefined angle α and distance d . Statistical measures on such a GLCM describe the texture and can be used as a feature vector. Some of the measures proposed in literature are: energy, homogeneity, entropy, correlation, inertia, cluster shade, and cluster prominence.

The LBP is extracted from a 3×3 neighbourhood as follows: The neighbourhood is binarised using the value of the center pixel as threshold. These values are then multiplied by certain weights and summed, leading to one single value for each neighbourhood. LBP feature vectors are formed by histogram bins of the distribution of the LBP values.

An extension of the LBP was proposed in [11], the Opponent Colour Local Binary Pattern (OC-LBP). The LBP operator is applied on each color channel separately. Additionally, each pair of colour channels is used, so that the center pixel of the neighbourhood and the neighbourhood itself are taken from different colour channels. All the histograms are concatenated to one feature vector.

Support vector machines (SVM) [12] are a commonly used classification technique because of many attractive features, such as the ability to deal with high-dimensional feature data. At first, a model is produced with a training set containing many data instances and then the model classifies the instances from a test set.

2.3 Experiments

Four different scenes of equal length were extracted from the database, each showing a different polyp. In total we used 1736 high-resolution images and every image contains a visible polyp from varying distance, angle and illumination.

Thus, it can be found out (1) how the texture features perform in general on this large database of high-resolution images and (2) if there is a relation between particular texture features and polyp types.

Since texture is a local property, we decided to divide each image into small image patches (50^2 and 70^2 pixels were used as patch sizes). We implemented four methods based on GLCM and LBP:

- GLCM 16: Four GLCM were computed for each patch with $d = 1$ and $\alpha \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. The following features have been computed on each matrix, leading to 16-dimensional feature vectors: energy, homogeneity, entropy, and correlation.
- GLCM 6: Four GLCM were computed as described above. The extracted features were: energy, homogeneity, entropy, inertia, cluster shade and cluster prominence. The means of the four values (from the four matrices) of each feature were computed leading to a 6-dimensional feature vector.

Table 1. AUC values of the different classifiers, best results are highlighted

Scene	Patch size	GLCM 6	GLCM 16	LBP	OC-LBP
1	70×70	0.83	0.95	0.89	0.94
2	70×70	0.70	0.75	0.76	0.87
3	70×70	0.89	0.88	0.95	0.96
4	70×70	0.65	0.68	0.80	0.91
1	50×50	0.80	0.89	0.86	0.89
2	50×50	0.71	0.75	0.71	0.84
3	50×50	0.80	0.83	0.88	0.95
4	50×50	0.63	0.65	0.77	0.89

- LBP: The LBP histogram was extracted as described in section 2.2. An amount of 64 histogram bins was used leading to a 64-dimensional feature vector.
- OC-LBP: The OC-LBP feature was extracted with an amount of 64 bins for each histogram, leading to a 576-dimensional feature vector.

With the extracted feature data an SVM-classifier was trained and tested using a radial basis function as kernel. We applied a stratified k -fold cross-validation with $k = 4$.

3 Results

Receiver Operating Characteristics (ROC) graphs were computed to analyze the performance of the classifiers based on the different texture features. To summarize the results, table 1 shows the area under the curve (AUC) values of the ROC graphs for the different scenes and patch sizes.

4 Discussion

As table 1 shows the best classification results were achieved by the OC-LBP method on almost every experiment and the GLCM 6 method performs worst.

Since the OC-LBP method is the only one of the tested methods that considers colour, colour seems to be an important feature for polyp detection in colonoscopy.

Comparing the two GLCM based methods there are two reasons for the poor results of the GLCM 6 method: (1) more features do not implicitly lead to better classification results and (2) by averaging the values from different GLCM too much information gets lost.

Another important result is, that every method itself achieves its best classification result either on scene 1 or scene 3 and its poor results either on scene 2 or scene 4. This leads to the assumption that there is no relation between one particular feature and polyp.

In general, texture seems to be a promising feature for automatic polyp detection. This reinforces results from previous works and indicates their validity for high-resolution images.

Future work with regard to an automatic polyp detection system will comprise the training of classifiers based on even more images showing different polyps, due to the high diversity of their appearances. Furthermore other methods for texture analysis will be examined.

Acknowledgement. Special thanks go to Beaumont Hospital Dublin, Ireland. This research was partly supported by Enterprise Ireland project endoview.

References

1. World Health Organization. Media Centre Fact Sheet No 297: Cancer; 2006 [cited 2008 Oct 04]. [online]. Available from: <http://www.who.int/mediacentre/factsheets/fs297/en/print.html>.
2. Thomson A, Ahnen D, Riopelle J. Intestinal polypoid adenomas. eMedicine, The Continually Updated Clinical Reference. 2007; Available from: <http://www.emedicine.com/med/TOPIC1175.HTM>.
3. Iakovidis DK, Maroulis DE, Karkanis SA. An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy. *Comput Biol Med.* 2006;36(10):1084–1103.
4. Karkanis SA, Iakovidis DK, Karras DA, et al. Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures. *Proc IEEE Conf Image Proc.* 2001; p. 833–836.
5. Wang P, Krishnan SM, Kugean C, et al. Classification of endoscopic images based on texture and neural network. *Proc IEEE EMBS.* 2001;4:3691–3695.
6. Ameling S, Wirth S, Paulus D. Methods for Polyp Detection in Colonoscopy Videos: A Review. University of Koblenz-Landau; 2008. 14/2008.
7. Shevchenko N, Mühldorfer S, Wittenberg T, et al. Untersuchung von Texturanalysemethoden zur automatischen Polypenerkennung. In: 7. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e. V., Tagungsband; 2008. p. 205–206.
8. Vilariño F, Lacey G, Zhou J, et al. Automatic labeling of colonoscopy video for cancer detection. *Proc Iberian Conf: Patt Recogn Image Anal.* 2007; p. 290–297.
9. Haralick RM, Dinstein I, Shanmugam K. Textural features for image classification. *IEEE Trans Syst Man Cybern.* 1973;3(6):610–621.
10. Pietikainen M, Ojala T. Nonparametric texture analysis with complementary spatial operators. *Texture Analysis in Machine Vision, Series in Machine Perception and Artificial Intelligence.* 2000;40.
11. Mäenpää T. The local binary pattern approach to texture analysis—extensions and applications; 2003. Ph.D. thesis, University of Oulu.
12. Vapnik V. *The Nature of Statistical Learning Theory.* Springer; 1999.