ONE-AGAINST-ONE CLASSIFICATION FOR ZOOM-ENDOSCOPY IMAGES

M. Häfner[1], R. Kwitt[2], F. Wrba[4], A. Gangl[1], A. Vécsei[3], and A. Uhl[2,*]

[1]Department of Gastroenterology and Hepatology, Medical University of Vienna, Austria
[2]Department of Computer Sciences, Salzburg University, Austria
[3]St. Anna Children's Hospital, Vienna, Austria
[4]Department of Clinical Pathology, Medical University of Vienna

[*]Corresponding author e-mail: uhl@cosy.sbg.ac.at

Keywords: pit-pattern classification, wavelet techniques, magnifying endoscopy, colon cancer

Abstract

In this paper, we present a novel approach for the classification of zoom-endoscopy images based on the pit-pattern classification scheme. Our feature generation step is based on the computation of a set of statistical features in the wavelet-domain. In the classification step, we employ a one-against-one approach using 1-Nearest Neighbor classifiers together with sequential forward feature selection. Our experimental results show that this classification approach drastically increases leave-one-out crossvalidation accuracy for our six-class problem, compared to already existing approaches in this research area.

1 Introduction

Recent statistics of the American Cancer Society reveal that colorectal cancer is the third most common cancer in men and women and the second most common cause of US cancer deaths. Since most colorectal cancers develop from polyps, a regular inspection of the colon is recommended, in order to detect lesions with a malignant potential or early cancer. A common medical procedure to examine the inside of the colon is colonoscopy, which is usually carried out with a conventional video-endoscope. A diagnostic benefit can be achieved by employing so called zoom-endoscopes, which achieve a magnification factor of up to 150 by means of an individually adjustable lens. In combination with dye-spraying to enhance the visual appearance (chromo-endoscopy) of the colon mucosa, zoom-endoscopy can reveal characteristic surface patterns, which can be interpreted by experienced physicians. Commonly used dyes are either methylene-blue, or indigo-carmine, which both lead to a plastic effect. In the research work of [13], the macroscopic appearance of colorectal polyps is systematically described and results in the so called *pit-pattern* classification scheme.

In this work, we try to tackle the problem of computer-assisted pit-pattern classification with a view to increase diagnostic accuracy. This is accomplished by computing a set of statistical image descriptors in the wavelet-domain and employing a novel classification approach based on a form of ensemble of specialists. Our approach is motivated by the work of [8], where the authors state that assessing the type of mucosal crypt patterns can actually predict the histological findings to a very high accuracy.

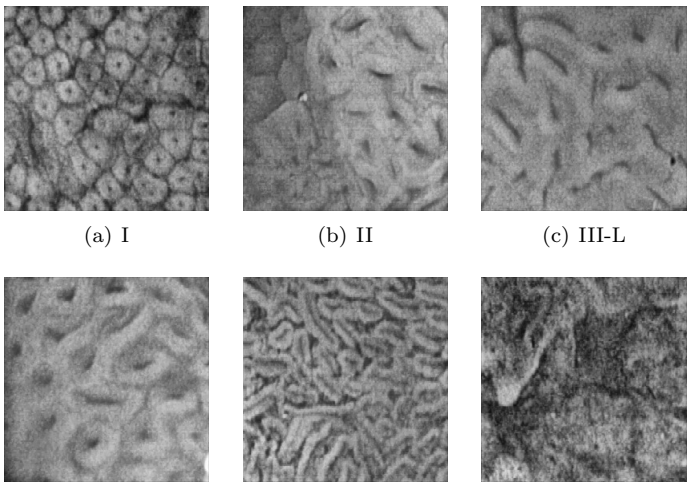
The remainder of the paper is structured as follows: in Section 2 we will briefly review the pit-pattern classification scheme and define our classification problem. In Section 3 we introduce the statistical image features. Section 4 then outlines the proposed classification approach followed by experimental results in Section 5. Finally, Section 6 concludes the paper with a short summary of our work and an outlook on further research.

2 Pit-Pattern Classification

Polyps of the colon are a frequent finding and are usually divided into metaplastic, adenomatous and malignant. Since the resection of all polyps is rather time-consuming, it is imperative that those polyps which warrant resection can be distinguished. Furthermore, polypectomy[1] of metaplastic lesions is unnecessary and removal of invasive cancer may be hazardous. The classification scheme of Kudo et al. [13] divides the mucosal crypt patterns into five groups. While types I and II are characteristic of benign lesions and represent normal colon mucosa or hyperplastic polyps, types III-L, III-S, IV and V represent neoplastic, adenomatous and carcinomatous structures. Using a magnifying colonoscope together with indigo carmine dye spraying, the mucosal crypt pattern on the surface of colonic lesions can be observed [14]. Several studies found a good correlation between the mucosal pit-pattern and the histological findings, where especially techniques using magnifying endoscopes (*zoom-endoscopes*) led to excellent results [7, 8, 11, 4]. Figure 1 shows six exemplary zoom-endoscopy images from our database for pit-patterns I to V.

Our classification problem now is to differentiate pit-types I,II,III-S,III-L,IV and V, which leads to a six-class problem since type III-L and III-S account for two classes. Note, that it is also possible to define a more coarse classification problem, where the aim is to differentiate types I and II from III-L, III-S, IV and V. However, in this work we focus on the more challenging six-class problem. Generally, pit-patterns I to IV can be characterized fairly well,

[1]the process of removing polyps



(a) I (b) II (c) III-L

(d) III-S (e) IV (f) V

Figure 1: Examples for pit-patterns I-V

while type V is a composition of unstructured pits. Table 1 provides a brief overview of the main characteristics of the different pit-patterns.

Pit-Pattern	Characteristics
I	roundish pits which designate a normal mucosa
II	stellar or papillary pits
III-S	small roundish or tubular pits, which are smaller than the pits of type I
III-L	roundish or tubular pits, which are larger than the pits of type I
IV	branch-like or gyrus-like pits
V	non-structured pits

Table 1: The characteristics of the different pit-patterns

Although at a first glance this classification scheme seems to be straightforward and easy to be applied, it needs some experience and exercising to achieve fairly good results. Correct diagnosis very much relies on the experience of the endoscopist as the interpretation of pit-patterns may be challenging [7]. Additionally, inter-observer variability of magnification chromoendoscopy has been described at least for Barrett's esophagus [16]. This inter-observer variability may to a lesser degree be also present in the interpretation of pit-patterns of colonic lesions. We therefore want to develop an approach for computer-based pit-pattern classification in order to enhance the quality of diagnostic results.

3 Statistical Image Features

In order to obtain a set of discriminative image features, we first compute a redundant image-representation using the Dual-Tree Complex Wavelet Transform (DT-CWT), originally proposed in [9]. The DT-CWT is an efficient realization of a complex wavelet transform which allows perfect reconstruction and has already been applied in texture-image retrieval [17] or even zoom-endoscopy image classification [15]. We have chosen the DT-CWT, since it is designed to overcome two commonly known shortcomings of the real, separable 2-D Discrete Wavelet Transform (DWT), which are lack of shift-invariance and poor directional selectivity. These properties, which come at the expense of a limited redundancy of 2^m in m dimensions, are important due to the following reasons: First, our images exhibit structures, oriented along directions other than horizontal, vertical or diagonal (see Figure 1), which are emphasized by the classic 2-D DWT. Second, the image acquisition process during colonoscopy is subject to several physical influences, which often cause image rotations and shifts. In 2-D, the implementation of the DT-CWT is straightforward by four parallel 2-D DWTs [18] and leads to to six complex subimages (denoted as complex subbands) per decomposition stage. The subbands capture image details oriented along approximately $\pm 15°, \pm 45°$ and $\pm 75°$.

In this work, we will use the statistical image features proposed in [15], where the marginal distributions of the wavelet coefficient magnitudes are modeled by two-parameter Weibull distributions. The probability density function (pdf) of the Weibull distribution with shape parameter c and scale parameter b is given by

$$p(x|c,b) = \frac{c}{b}\left(\frac{x}{b}\right)^{c-1} \exp\left\{-\left(\frac{x}{b}\right)^{c}\right\}, \qquad (1)$$

for $x > 0, b > 0$ and $c > 0$. The maximum-likelihood estimates (MLEs) $\hat{c}$ and $\hat{b}$ [12] of the Weibull distribution fitted to the subband coefficients then serve as two image features per subband. In a J-scale DT-CWT decomposition we thus obtain a total of $J \times 12$ features per image. Since the position of the subbands in the DT-CWT decomposition structure can be identified unambiguously by the tuple (s,k), we use these tuple to identify the MLEs for each subband (here, $s \in \{1,\ldots,J\}$ denotes the scale and $k \in \{1,\ldots,6\}$ denotes the orientation). For a J-scale DT-CWT an image feature vector $\mathbf{v}$ is constructed by simple feature concatenation as follows:

$$\mathbf{v} = \left[\hat{b}^{(1,1)}, \hat{c}^{(1,1)}, \hat{b}^{(1,2)}, \hat{c}^{(1,2)}, \ldots, \hat{b}^{(J,5)}, \hat{c}^{(J,6)}\right] \qquad (2)$$

We add an additional subscript i to the notation $\mathbf{v}_i$ to denote that the feature vector belongs to a particular image i. For comparative reasons, we will also compute the image features proposed in [1], which are simply the mean and standard deviation of the wavelet coefficient magnitudes per subband. The final image feature-vector for this feature set is then again constructed by concatenation.

4 One-Against-One Classification

Regarding the classification step of our work, we will use an approach that is commonly used to split multi-class problems into binary ones and is called *round robin binarization* (RRB) [6] (aka one-against-one classification). The idea of this approach is quite simple, namely to train one classifier for each possible pair of classes. Given a c-class problem, we thus have to train a total of $c(c-1)/2$ classifiers. In this setup, each binary classifier is trained using

only the examples of the two classes it has to discriminate. For example, let $V := \{\mathbf{v}_i\}_{1 \leq i \leq L}$ be the complete set of training vectors. Then, the classifier C_{ij}, which discriminates between class i and j is trained using the training set $S_{ij} := \{\mathbf{v}_n | \mathbf{v}_n \in i \vee \mathbf{v}_n \in j\}$.

Hence, we can view the RRB approach as some sort of expert system, where each classifier is an expert in discriminating only the samples from two classes. Given an arbitrary sample $\mathbf{v} \in \mathcal{F} \subset \mathbb{R}^d$, where $\mathcal{F}$ denotes the feature space and d denotes the feature space dimensionality, a classifier C_{ij} will output either i or j as the predicted class label. However, this prediction is a *qualified* one, if and only if the sample $\mathbf{v}$ actually belongs to either i or j. Otherwise, the prediction is termed an *unqualified* prediction. Hence, given that $C_{ij}(\mathbf{v}) = i$ we can at best conclude, that $\mathbf{v} \notin i$ [2]. This interpretation of the base classifier predictions is termed *voting against* and not *voting for*, which is actually wrong with regards to the aforementioned issues. The final class label prediction from the $c(c-1)/2$ classifiers is then determined by counting the votes against each class and selecting the very class with the smallest number of votes against. In [2] the voting against principle is discussed from a theoretical viewpoint.

In order to allow a comparative study to the approach in [15], where a single 1-Nearest Neighbor (NN) classifier is used to discriminate between the six classes, we will use 1-NN classifiers as our base classifiers as well. Furthermore, we propose an extension of the RRB procedure, namely to conduct an additional feature selection step in the training phase of the base classifiers. Specifically, we will use sequential forward feature selection [5] (SFFS) to select the very subset of features, which leads to the highest leave-one-out crossvalidation (LOOCV) accuracy [3]. Note, that we do not limit the number of selected features. Starting with a feature subset size of one, one feature is added in each iteration if this feature improves LOOCV accuracy. Given that $\mathcal{G} = \{1, \ldots, d\}$ denotes the set of all indices for all possible features (i.e. all shape and scale MLEs), we finally obtain $c(c-1)/2$ feature subsets $\mathcal{G}_{C_{ij}} \subset \mathcal{G}$ (one subset per base classifier) after the training procedure. Note, that in case of a J-scale DT-CWT the feature vector concatenation approach of Eq. (2) leads to a starting index set dimensionality of $|\mathcal{G}| = 12J$. The final classification accuracy of the RRB approach is then again estimated by the method of LOOCV.

5 Experimental Results

In this section, we present the experimental results of our work. Our image database contains 484 RGB images, acquired in 2005/2006 at the Department of Gastroenterology and Hepatology (Medical University of Vienna) using a magnification endoscope (Olympus Evis Exera CF-Q160ZI/L) with a magnification factor of 150x. To enhance visual appearance, dye-spraying with indigo-carmine was applied and biopsies or mucosal resections were taken to obtain a histopathological diagnosis (*our ground truth*). For pit-patterns I,II and V, biopsies were taken, since these types need not be removed. Lesions of pit-patterns III-S,III-L and IV have been removed endoscopically. Table 2 lists the number of image samples per class.

I	II	III-L	III-S	IV	V
126	72	62	18	146	60

Table 2: Number of image samples per pit-pattern

All images are first converted to the LUV color space and the (L)uminance channel information (grayscale) is retained. Before we decompose the images using the DT-CWT, we conduct two pre-processing steps in order to enhance the visual quality of the images: First, we use contrast-limited adaptive histogram equalization (CLAHE) [19] with 8×8 tiles and an uniform distribution for constructing the contrast transfer function. Second, we apply a Gaussian blur with $\sigma = 0.5$ using a 3×3 mask. Regarding the choice of filters for the DT-CWT, we use Kingsbury's Q-Shift (14,14)-tap filters (for decomposition levels ≥ 2) in combination with (13,19)-tap near-orthogonal filters (for decomposition level 1) [10]. The technical reasons for using different filter sets are explained in [18]. We vary the decomposition depth J from 4 to 6 in order to study the effect on the classification rates. In Table 3 we list the LOOCV results *with* and *without* RRB. The mean and standard deviation based image features of [1] are denoted as *Classic*.

	RRB + SFFS		[15] + SFFS	
Scales	Weibull	Classic	Weibull	Classic
4	91.74	90.70	84.50	81.40
5	95.04	93.39	88.02	83.47
6	**95.87**	93.39	88.84	83.88

Table 3: LOOCV accuracies with and without RRB using two different feature sets

As we can see, the highest LOOCV rates are achieved by employing a 6-scale DT-CWT decomposition together with the Weibull features and RRB. This is a plausible result, since increasing the feature-vector dimensionality equally increases the search-space of the SFFS. We further notice, that the RRB approach improves the classification rates is all cases. In order to see the effect on the per-class rates using the RRB approach, tables 4 and 5 provide the detailed confusion matrix results with and without RRB using the Weibull feature set. The diagonal elements of the confusion matrix contain the number of correct class predictions, whereas the other matrix cells contain the number of wrong class predictions. We notice the effect that RRB especially increases the LOOCV accuracy for the pit-pattern classes with very few samples (i.e. III-L and III-S). We suppose that this results from the redundancy in the base learners, which allows the correction of wrong votings and presumably simpler decision boundaries for the two-class problems.

6 Conclusion

In this paper, we have presented an approach for computer-assisted classification of zoom-endoscopy images based on

	I	II	III-S	III-L	IV	V
I	**118**	4	0	0	1	0
II	7	**68**	0	0	0	0
III-S	0	0	**62**	0	0	1
III-L	0	0	0	**18**	0	0
IV	0	0	0	0	**143**	4
V	1	0	0	0	2	**55**

Table 4: Confusion matrix **with** RRB using the Weibull features

	I	II	III-S	III-L	IV	V
I	**110**	9	1	0	1	0
II	13	**62**	0	2	4	0
III-S	1	0	**57**	0	3	2
III-L	0	0	3	**16**	1	0
IV	1	1	1	0	**131**	4
V	1	0	0	0	6	**54**

Table 5: Confusion matrix **without** RRB using the Weibull features

the pit-pattern classification scheme. By relying on a set of statistical image features, we have demonstrated that a one-against-one approach for splitting a multi-class problem into a set of two-class problems drastically increases classification performance. The detailed confusion-matrix results show a particular improvement in LOOCV accuracy for those classes with very few training samples. Further work will include the incorporation of color-information and the evaluation of classifiers other than K-NN.

Acknowledgments

This work is funded by the Austrian Science Fund (FWF) under Project No. L366-N15

REFERENCES

[1] W.Y. Ma B.S. Manjunath. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842, August 1996.

[2] F. Cutzu. How to do multi-way classification with two-way classifiers. In *Proceedings of the 3rd Joint International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP'03)*, pages 375–384, Istanbul, Turkey, 2003.

[3] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley & Sons, 2nd edition, November 2000.

[4] K.-I. Fu et al. Chromoendoscopy using indigo carmine dye spraying with magnifying observation is the most reliable method for differential diagnosis between non-neoplastic and neoplastic colorectal lesions: a prospective study. *Endoscopy*, 36(12):1089–1093, 2004.

[5] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Morgan Kaufmann, 2nd edition, 1990.

[6] J. Fürnkranz. Round robin rule learning. In *Proceedings of the 18th International Conference on Machine Learning (ICML'01)*, pages 146–153, Nashville, Tennessee, United States, 2001.

[7] D.P. Hurlstone. High-resolution magnification chromoendoscopy: Common problems encountered in "pit pattern" interpretation and correct classification of flat colorectal lesions. *American Journal of Gastroenterology*, 97:1069–1070, 2002.

[8] S. Kato et al. Assessment of colorectal lesions using magnifiying colonoscopy and mucosal dye spraying: Can significant lesions be distinguished? *Endoscopy*, 33:306–310, 2001.

[9] Nick G. Kingsbury. The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters. In *Proceedings of the IEEE Digital Signal Processing Workshop, DSP '98*, pages 9–12, Bryce Canyon, USA, August 1998.

[10] Nick G. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Applied and Computational Harmonic Analysis*, 10(3):234–253, May 2001.

[11] Kazuo Konishi, Kazuhiro Kaneko, Toshinori Kurahashi, Taikan Yamamoto, Miki Kushima, Akira Kanda, Hisao Tajiri, and Keiji Mitamura. A comparison of magnifying and nonmagnifying colonoscopy for diagnosis of colorectal polyps: a prospective. *Gastrointestinal Endoscopy*, 57:48–53, 2003.

[12] K. Krishnamoorthy. *Handbook of Statistical Distributions with Applications*. Chapman & Hall, 2006.

[13] S. Kudo, S. Hirota, T. Nakajima, S. Hosobe, H. Kusaka, T. Kobayashi, M. Himori, and A. Yagyuu. Colorectal tumorous and pit pattern. *Journal of Clinical Pathology*, 47:880–885, 1994.

[14] S. Kudo, S. Tamura, T. Nakajima, H. Yamano, H. Kusaka, and H. Watanabe. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointestinal Endoscopy*, 44(1):8–14, July 1996.

[15] R. Kwitt and A. Uhl. Modeling the marginal distributions of complex wavelet coefficient magnitudes for the classification of zoom-endoscopy images. In *Proceedings of the IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'07)*, pages 1–8, Rio de Janeiro, Brasil, 2007.

[16] A. Meining, T. Rösch, R. Kiesslich, M. Muders, F. Sax, and W. Heldwein. Inter- and intra-observer variability of magnification chromoendoscopy for detecting specialized intestinal metaplasis at the gastroesophageal junction. *Endoscopy*, 36:160–164, 2004.

[17] P. Rivaz and N. Kingsbury. Complex wavelet features for fast texture image retrieval. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'99)*, pages 109–113, Kobe, Japan, 1999.

[18] I. W. Selesnick, R. G. Baraniuk, and N. Kingsbury. The dual-tree complex wavelet transform - a coherent framework for multiscale signal and image processing. *IEEE Signal Processing Magazine*, 22(6):123–151, November 2005.

[19] K. Zuiderveld. Contrast limited adaptive histogram

equalization. In Paul S. Heckbert, editor, *Graphics Gems IV*, pages 474–485. Morgan Kaufmann, 1994.