



UNIVERSITÄT
KOBLENZ · LANDAU

Fachbereich 4: Informatik



Merkmale zur Polypenklassifikation in Koloskopie-Bildern

Bachelorarbeit
zur Erlangung des Grades
BACHELOR OF SCIENCE
im Studiengang Computervisualistik

vorgelegt von

Sandy Engelhardt

Betreuer: Dipl.-Inf. Stephan Wirth, Institut für Computervisualistik,
Fachbereich Informatik, Universität Koblenz-Landau

Erstgutachter: Dipl.-Inf. Stephan Wirth, Institut für Computervisualistik,
Fachbereich Informatik, Universität Koblenz-Landau

Zweitgutachter: Prof. Dr.-Ing. Dietrich Paulus, Institut für
Computervisualistik, Fachbereich Informatik, Universität Koblenz-Landau

Koblenz, im Oktober 2009

Kurzfassung

Die Koloskopie ist der Goldstandard zur Aufspürung von gefährlichen Darmpolypen, die sich zu Krebs entwickeln können. In einer solchen Untersuchung sucht der Arzt in den vom Endoskop gelieferten Bildern nach Polypen und kann diese gegebenenfalls entfernen. Um den Arzt bei der Suche zu unterstützen, erforscht die Universität Koblenz-Landau zur Zeit an Methoden, die zur automatischen Detektion von Polypen auf endoskopischen Bildern verwendet werden können. Wie auch bei anderen Systemen zur Mustererkennung werden hierzu zunächst Merkmale aus den Bildern extrahiert und mit diesen ein Klassifikator trainiert. Dieser Klassifikator kann dann für die Klassifikation von ihm unbekanntem Bildern eingesetzt werden. In dieser Bachelorarbeit wurde das vorhandene System zur Polypendetektion um Merkmalsdetektoren erweitert und mit den bereits vorhandenen verglichen. Implementiert wurden Merkmale basierend auf der Diskreten Wavelet Transformation, auf Grauwertübergangsmatrizen und auf Local Binary Patterns. Verschiedene Modifikationen dieser Merkmale wurden getestet und evaluiert.

Abstract

Colonoscopy is the gold standard for detection of colorectal polyps that can progress to cancer. In such an examination physicians search for polyps in endoscopic images. Thereby polyps can be removed. To support experts with a computer-aided diagnosis system, the University of Koblenz-Landau currently makes some efforts in research different methods for automatic detection. Comparable to traditional pattern recognition systems, features are initially extracted and a classifier is trained on such data. Afterwards, unknown endoscopic images can be classified with the previously trained classifier. This bachelor thesis concentrates on the extension of the feature extraction module in the existing system. New detection methods are compared to existing techniques. Several features are implemented, incorporating Graylevel Co-occurrence Matrices, Local Binary Patterns and Discrete Wavelet Transform. Different modifications on those features are applied and evaluated.

Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Vereinbarung der Arbeitsgruppe für Studien- und Abschlussarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Mit der Einstellung dieser Arbeit in die Bibliothek bin ich einverstanden. ja nein

Der Veröffentlichung dieser Arbeit im Internet stimme ich zu. ja nein

Koblenz, den 12. Oktober 2009

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 9 |
| 1.1 | Medical Background | 9 |
| 1.2 | Pattern Recognition Schemes | 10 |
| 1.3 | Content of this Work | 11 |
| 2 | Related Work | 13 |
| 2.1 | Form-based Detection | 13 |
| 2.2 | Texture-based Detection | 14 |
| 2.3 | Combined Methods | 16 |
| 2.4 | Discussion | 16 |
| 3 | Texture Features | 19 |
| 3.1 | Transform methods | 19 |
| 3.2 | Statistical methods | 23 |
| 3.2.1 | Co-occurrence Matrix (GLCM) | 23 |
| 3.2.2 | Local Binary Pattern (LBP) | 26 |
| 4 | Classification and Evaluation | 31 |
| 4.1 | Crossvalidation | 31 |
| 4.2 | Support Vector Machines | 31 |
| 4.3 | k -Nearest Neighbor Classifier | 32 |
| 4.4 | Evaluation of Results | 33 |
| 5 | System Description | 35 |
| 5.1 | Data | 35 |
| 5.2 | Patches | 36 |
| 5.3 | Experimental Flow | 38 |
| 6 | Feature Descriptions and Experiments | 41 |
| 6.1 | Existing features | 41 |
| 6.1.1 | GLCM6, GLCM16, LBP, OC-LBP | 42 |

| | | |
|----------|--|-----------|
| 6.1.2 | Discussion | 42 |
| 6.2 | Wavelet Features | 42 |
| 6.2.1 | Color Wavelet and Color Wavelet Covariance | 43 |
| 6.2.2 | Wavelet-Decomposition | 47 |
| 6.3 | GLCM features | 49 |
| 6.3.1 | ColorGLCM | 50 |
| 6.3.2 | OC-GLCM | 51 |
| 6.4 | LBP features | 53 |
| 6.4.1 | ColorLBP | 53 |
| 6.4.2 | Rotation-Invariant-LBP8 | 55 |
| 6.4.3 | Rotation-Invariant-Subset-LBP16 | 55 |
| 6.4.4 | Subset-LBP8 | 56 |
| 6.5 | Discussion of Results | 57 |
| 7 | Summary | 61 |
| 7.1 | Possible Improvements | 61 |
| 7.2 | Summary of this Work | 62 |

Chapter 1

Introduction

1.1 Medical Background

Cancer is a leading cause of death worldwide. It describes the transformation from a normal cell into a tumor; a progression from a pre-cancerous lesion to malignant tumours. According to the World Health Organisation (WHO) 655,000 people die from colon cancer per year [WHO06]. One third of the cancer burden could be cured if detected early and treated adequately. In Germany, the number of incidences mounts to 70,000 per year with approximately 30,000 fatalities [SPRS⁺08].

Colorectal or intestinal polyps are the most frequently occurring pre-cursor of colon cancer. Polyps can be broadly classified as neoplastic and nonneoplastic polyps. Among neoplastic polyps one differentiates between adenomatous and malignant. Approximately 95% of all colorectal carcinomas arise from adenomas, a fact that underlines the importance of treatment of colorectal polyps [TA07].

The abnormality of polyps is mainly detected when the surface of the lipoma is eroded or irregular in contrast to a smooth surface. The occurring forms can be classified in tubular, tubulovillous, or villous, primarily based on the overall percentage of villous component. The risk of progression to cancer of adenomas is related to their macroscopic appearance (size, villous components) as well as their microscopic architecture and degree of dysplasia. Considering the size, smaller adenomas (< 1 cm) have a lower risk of malignant potential [TA07].

Colonoscopy is the accepted gold standard for screening colon cancer or colorectal polyps. It allows diagnosis, therapy as well as biopsy. In most cases, the polyps are removed directly when detected. Nevertheless, there is a 6-12% miss rate for adenomas that are 1cm or larger; the miss rate for smaller adenomas is up to 25% [TA07]. This is due to the fact that the polyp can show up on the screen but is not identified by the physician because of non-attention or subjective

diagnosis. Furthermore, endoscopic analysis does not cover all parts of the colon. As a consequence, parts remain unseen by the camera.

It is recommended to have continuing surveillance of patients with previously removed adenomas. The interval between colonoscopies depends on the size, number, and histological type of polyp, as well as the patient's family history. Polyp recurrence rates are 20% at 5 years and 50% at 15 years [TA07].

A colonography is a visual recording of the colon obtained using computed tomography (CT) technology. This is a new proposed technique for detection of adenomas. Although this method produces a virtual representation of the colon which can speed up the visual analysis by the physician, it also holds some disadvantages: the extensive amount of radiologist working time during CT scanning, the costs of such an exam and of course the radiation that the patient is subjected to [ACN07]. When polyps are detected and must be removed, the patient undergoes colonoscopy analysis once more, which finally leads to a double intervention. Moreover, virtual colonoscopy has lower sensitivity than optical colonoscopy for small (<1 cm) adenomas [TA07]. Thus the potential to miss small polyps is higher.

The visual analysis of the endoscopic images has some drawbacks such as interpretational variation and non-suitability for comparative evaluation. Hence a computer-aided system for detection will help considerably in the quantitative characterization of abnormalities, thereby improving patient's care. It is desirable to develop a system that marks polyps reliably during the screening process leading to a significantly decreased miss-rate.

1.2 Pattern Recognition Schemes

Computer-aided systems for detection often incorporate the application of pattern recognition and classification. Traditionally, they consist of several moduli taking over specific tasks explained in the following.

In a first step sensor data is sampled and quantized, for example a video frame. Then a preprocessing might be applied to the image in order to improve the results of subsequent steps of processing. This either results in reduction of complexity or improvement of performance, or both. Additionally, features are extracted from the image leading to a noticeable reduction of representational space. They are now represented by vectors holding numeric or nominal values. After this, two disjunct sets of feature vectors have to be chosen, namely a training set for learning of the classifier and a test set for evaluation of the complete system.

Basically, one differentiates between supervised and unsupervised learning. Supervised learning deals with classes that are known before the training is applied. Features are extracted and mapped to these predefined classes. Otherwise, during the training phase disjunct classes have to be created, which involves clustering the

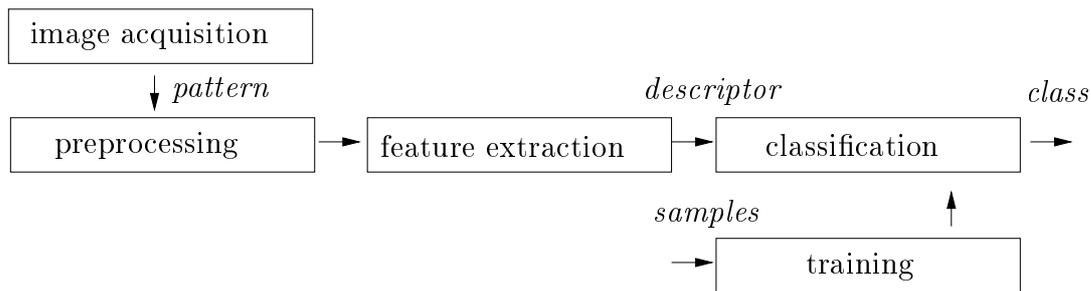


Figure 1.1: Traditional scheme for classification of patterns.

features in feature space. The trained classifier can then be used for classification of unknown patterns in the so called testing phase [Nie07].

100 percent of correct classified patterns is desirable but often hard to realize, depending on the complexity of the patterns. On this note, an analysis of the results is necessary to evaluate the performance and correctness of the system.

Computer-aided diagnosis concerning medical themes requires image acquisition, image processing, feature extraction and classification as depicted in the pipeline above (figure 1.1). Ameling [AWP⁺09] proposed a polyp detection system consisting of the above mentioned steps. Several feature extraction methods have been already implemented. Nevertheless, there is a potential to increase the performance of the system by examining the single modules.

1.3 Content of this Work

The approach of this work is to increase the quality of feature extraction on endoscopic images in the system of [AWP⁺09]. In figure 1.1 this module is depicted by the third step in the pipeline, which computes descriptors such as feature vectors from the images. The existing extraction methods will be customized to the task of polyp detection by testing different adjustments and versions of the features.

Since there is a high dependency between feature extraction and classification concerning the performance of the whole system, classification will be examined additionally.

To start with, a number of existing schemes for the detection of abnormalities in the colon are presented in the following chapter. In chapter 3, texture feature methods are introduced such as Graylevel Co-occurrence Matrices, Local Binary Patterns and the Wavelet Transform. Chapter 4 gives an overview about different

classifiers and the methods of evaluation. Subsequently, the system as well as the experimental flow is described in chapter 5. Chapter 6 explains the features that are implemented in this work and shows the results of the applied tests. Finally, chapter 7 gives a summary about this work and lists some aspects about possible further work.

Chapter 2

Related Work

This chapter gives an overview about the state of the art of computer aided detection systems in endoscopic images. It is especially focussed on the detection of colon cancer as well as intestinal polyps.

2.1 Form-based Detection

As described in 1.1, polyps can have tubular, tubulovillous or villous form. The following approaches take advantage of this, trying to find contours or segments on endoscopic images.

Krishnan et al. [KYC⁺98] describe a form-based approach for detection of abnormalities in the colon. First the image contours are extracted using the Canny Edge Detector, and the curvature of each contour is computed. Zero-crossings of curvature along the contour are detected then. Afterwards contour segments are filtered which are positioned between two zero-crossings. If this contour segment has the opposite curvature signs to those of its two neighboring contour segments, the presence of polyps or tumors will be rated as high.

The method of Hwang et al. [HOT⁺07] relies on the elliptical shape of colon polyps. In a first step a watershed-based image segmentation is applied to a frame. Then a binary edge map is constructed for each segmented region using a particular threshold. The map is used as input for an ellipse fitting algorithm. Ellipses are described as second order polynomials. They are mapped to the computed edge map using a least square fitting method. Among detected ellipses, those are filtered which do not represent actual polyps. The strength of the edge and the intensity value inside the ellipse serves as criteria for filtering.

Dhandra et al. [DHHM06] do not consider the edges like the methods mentioned previously. Their approach converts the endoscopic color image into HSI color space and then a watershed segmentation technique is applied. The classi-

fication of the image as abnormal is simply based on the number of watershed regions present in the image, compared to a certain threshold.

Another technique for detecting polyps was introduced by Kang et al. [KD03]. The endoscopic images obtain a contrast enhancement which is accomplished by performing a histogram stretching operation in RGB color space. In a next step, the Canny Edge Detection Algorithm is applied to each RGB color channel separately. The detected edges are combined afterwards to one result. Morphological operations such as dilation for edge thickening and connecting of disjoint edges are finally used. The resulting image segments are analyzed and filtered considering area, color and elliptical shape of the segment.

2.2 Texture-based Detection

Texture analysis is frequently used in image processing and pattern recognition for characterization of regions from digital images. Texture carries information about the micro-structure of the image regions and the occurring distribution of gray levels [Nie07].

Wang et al. [WKKT01] propose a feature extraction method called Local Binary Pattern (LBP) which is a local texture descriptor. The LBP can be combined with the image intensities to LBP/I. This LBP/I distribution is represented in a discrete two dimensional histogram. A log-likelihood-ratio called the G-statistic, which is a modification from Kullback's criterion, is used as a pseudometric for comparing LBP/I distributions. A Neural Network formed by Self-Organizing Maps (SOM) provides means for classification.

The concept of Li et al. [LCK05] is to transform the RGB endoscopic images into CIELab color space to analyze color and luminance separately. Patches of fixed size are used, which overlap 50% to each other. A two level Discrete Wavelet Transform decomposition is applied to each patch and CIELab channel separately. Afterwards, mean as well as standard deviations are calculated from the absolute values of approximation and detail coefficients of the second level of Wavelet decomposition. Another feature includes 1-dimensional histograms of the luminance channel L (with 16 bins) and 2-dimensional histograms of the a and b components in CIELab space (with 64 bins). Support Vector Machines serve as classification scheme.

Tjoa et al. [TK03] determine a so called texture spectra in the chromatic and achromatic domains in the colonoscopic image (HSI and RGB components). Therefore texture units and texture unit numbers are calculated, which are used to form a histogram. Six statistical measures are extracted from each texture spectrum: Energy, Mean, Standard Deviation, Skew, Kurtosis and Entropy. Principal Component Analysis (PCA) reduce the size of the features, which are afterwards

evaluated by a Backpropagation Neural Network (BPNN). They found that using texture and color features improves classification results when compared to using only one type of information.

The method of Alexandre et al. [ACN07] takes the color of a pixel (RGB components) and its position as feature vector dimensions. This implies a requirement of a high dimensional input space for SVM. In a preprocessing step they divide the original image into subimages of 40×40 pixels and classify each patch separately.

Karkanis et al. [KMGS99] propose a scheme which uses textural descriptors based on second order gray level statistics called Graylevel Co-occurrence Matrices (GLCM), initially proposed by Haralick [HDS73]. This method evaluates a series of matrices that describe the spatial variation of gray level values within a local area. In this approach four GLCM have been computed and four statistical features were determined: Angular Second Moment (Energy), Correlation, Inverse Difference Moment and Entropy. This forms a 16-dimensional feature vector, which is used as input for classification with Neural Networks.

In [KIM⁺00] a one dimensional Discrete Wavelet Transform (DWT) decomposition is performed on raster scanned images, resulting in four wavelet subimages for each patch. Then GLCMs are calculated on the wavelet domain and four statistical measurements are estimated. A Multilayer Feedforward Neural Network (MFNN) is employed for classification of the 16-dimensional feature vector. As experiments with both proposed methods indicate, the Wavelet Transform performs better than the simple GLCM method.

A one dimensional Discrete Wavelet Transform is performed by Karkanis et al. [KIKM01], resulting in approximation and detail components. The Daubechies wavelet basis is utilized due to their orthogonal property. Subsequently, GLCMs are extracted from the detail coefficients (without the lowpass-filtered channel) and four statistical measurements called Angular Second Moment, Correlation, Inverse Difference Moment and Entropy are calculated. 48-component feature vectors form the input to the Multilayer Feedforward Neural Network architecture. This proposal was implemented in CoLD (Colorectal Lesion Detector) [MIKK03] with incorporation of another classifier called Multilayer Perceptron Neural Networks (MLP).

Karkanis et al. [KIM⁺03] propose a new color feature extraction scheme named Color Wavelet Covariance (CWC) based on a fixed size sliding window. A three-level DWT decomposition is performed and GLCMs are extracted from the second wavelet level on each color channel separately. Afterwards, the aforementioned statistical measurements are computed. Covariance values of pairs of the estimated features constitute the 72-dimensional CWC feature vector. Linear Discriminant Analysis is used for classification of the features.

In [IMK06] different preprocessing methods and various feature extraction techniques are compared to each other. Color space transformations (e.g. RGB, K-L, CIE-Lab, HSV) are tested and incorporated with each extracted feature for instance Local Binary Patterns (LBP), Opponent-Color LBP (OC-LBP), Wavelet-Energy and CWC. Linear and non-linear classification modules are investigated.

Ameling et al. [AWP⁺09] compares existing feature extraction methods such as GLCM and LBP, exploiting the patch approach. Four different polyp scenes are chosen for testing. The GLCM6 feature as well as the GLCM16 feature computes four matrices on gray level patches. For GLCM6, six statistical measurements called Energy, Entropy, Inverse Difference Moment, Inertia, Cluster Shade, Cluster Prominence are extracted and the mean is computed from the extracted values. GLCM16 utilizes only four statistical measures, the same measures like [KMGS99] use in their application without averaging. LBP and OC-LBP features are also investigated. OC-LBP performed best on the preselected four scenes, combining texture and color information.

2.3 Combined Methods

As shown in the previous sections, there are many approaches for detection of lesions. However, there is not a single method to detect all kinds of lesions. Considering this fact, Zheng et al. [ZK01] combines multiple techniques. A multisensor data fusion technique based on Bayesian Inference is applied. This approach was further improved to an intelligent fusion-based clinical decision support in [ZKT05]. Subdecisions are estimated based on associated component feature sets ([TKK⁺01] [KWL⁺00] [WKHS02]) derived from the endoscopic images. Bayesian probability computations are employed to evaluate the accuracies of subdecisions and are utilized in estimating the probability of the fused decision.

2.4 Discussion

The results of the research groups are difficult to compare because of the usage of different data bases, which are beside this often too small to make reliable predictions. Additionally, the systems are not trained for all types of polyps.

Another aspect to consider is the resolution of the endoscopic images, which do not comply with technical progress. It is possible today to use full-high definition resolution instead of images of size 320×240 like in [IMK06, LCK05, KIM⁺03]. High resolution endoscopic images have the advantage to provide more concise information about the microstructures of the intestinal wall.

The micro-structure of the intestinal wall is defined by vasculature and mucosa, while the appearance of polyps is determined by the degree of dysplasia. On this note, texture features can be a discriminating aspect in detecting polyps. Graylevel Co-occurrence Matrices and Local Binary Patterns are popular methods as described in section 2.2. They estimate structural relationships between pixels which are necessary to provide adequate texture modelling.

The analysis of the form of polyps is often realized by finding edges or through region segmentation. It is difficult to predict whether this approach will lead to a reliable detection of polyps, because many similar shaped structures are found in the colon, for instance intussusceptions. Furthermore, there are different types and sizes of polyps having varying forms. These aspects result in a very complex segmentation task.

Other approaches are concerned with transform methods such as the Wavelet Transform, which has an advantageous effect on the representation and modelling of texture [CR95]. It is possible to perform multiresolutional analysis, which could have an enormous potential in examining endoscopic images. Due to the fact that the endoscope most likely has different distances to the intestinal wall during endoscopy scale variant features are computed when no adequate processing is applied. In this way the Wavelet Transform might be an appropriate mean.

Considering the classification techniques, two methods are frequently used in the presented computer-aided detection systems, namely SVMs and Neural Networks. A Support Vector Machine found also application in [AWP⁺09].

All in all the here discussed and positive evaluated approaches have a potential in detecting polyps reliably. The overall aim is to include color in the feature extraction techniques from [AWP⁺09], which seems to be one of the most promising information bases for polyp detection. A subset of the introduced texture methods such as Wavelet Transform, Graylevel Co-occurrence Matrices and Local Binary Patterns can be combined in different ways. Some of the combinations already have reference in literature, while others are never tested so far on endoscopic images.

Chapter 3

Texture Features

Texture can be seen as a rich source of visual information that is easily perceived by humans. Nevertheless, there is no strict definition of image texture. Generally speaking, textures are complex visual patterns composed of entities, or subpatterns that have special characteristics. Hence, texture can be regarded as a similarity grouping of such entities in an image [RK82].

3.1 Transform methods

Transform methods of texture analysis represent an image in a space whose coordinate system has an interpretation that is closely related to the characteristics of texture. Methods based on Fourier Transform perform poorly in practice, due to its lack of spatial location, while the Wavelet Transform method possesses a capability of time (space) location of signal spectral features [Mal89].

Several psycho-visual studies [RL93][Jul86] demonstrate that the human visual system processes images in a multi-scale manner. This knowledge motivates the use of multi-scale or multi-resolution approaches for texture analysis. Therefore, the scale is the most important parameter, which is determined by the size of the textural element or the considered neighborhood. The Wavelet Transform provides a formal technique for such an approach [CR95].

Wavelet Transform

There are two advantages to mention considering Wavelet Transform. It has been demonstrated that Discrete Wavelet Transform can lead to better texture modeling [Mey93]. Varying spatial resolution allows it to represent textures at the most suitable scale. Additionally, the wide range of choices for the wavelet basis function makes it easily adjustable.

The disadvantage of Wavelet Transform is that it is not translation-invariant [LCC97] and thus results in different coefficients as soon as the source signal is shifted.

The Wavelet Transform utilizes a basic function $\psi(t)$, the so-called ‘mother wavelet’, which is scaled with a factor α and shifted to the position τ of the time axis. The following integral describes the Wavelet Transform (WT) of a signal f

$$WT(\tau, \alpha) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{|\alpha|}} \psi\left(\frac{t - \tau}{\alpha}\right) dt = \int_{-\infty}^{\infty} f(t) \psi_{\alpha, \tau}(t) dt. \quad (3.1)$$

Limiting the range of α and τ to the following discrete values,

$$\alpha = 2^{-\mu}, \quad \tau = k\alpha, \quad \mu, k = \dots, 0, \pm 1, \pm 2, \dots \quad (3.2)$$

generates a family of Wavelets $\psi_{\mu, k}(t)$ from one basis function ψ by

$$\psi_{\mu, k}(t) = \frac{1}{\sqrt{|\alpha|}} \psi\left(\frac{t - \tau}{\alpha}\right) = \sqrt{2^\mu} \psi(2^\mu t - k), \quad \psi_{0, 0}(t) = \psi(t). \quad (3.3)$$

The orthogonal property of a wavelet family is of importance, since it maintains the textural structure along the different scales of the transform.

Two functions mutually orthonormal are initially adopted: the scaling function ϕ , which increasingly reduces the resolution of the function f and the mother wavelet function ψ . Scaling function and wavelet functions take over the general task of low- and highpass filtering. Other wavelets are then produced by translation of the scaling function ϕ and dilations of the mother wavelet ψ , according to the equations:

$$\phi_{\mu, k}(t) = \sqrt{2^\mu} \phi(2^\mu t - k) \quad (3.4)$$

$$\psi_{\mu, k}(t) = \sqrt{2^\mu} \psi(2^\mu t - k). \quad (3.5)$$

$\mu, k \in \mathbb{Z}$ are the scale and translation indices, respectively; the factor $\sqrt{2^\mu}$ is an inner product normalization.

The Wavelet Transform can easily be extended to multiple dimensions, because one can utilize a separable description. In the case of a two-dimensional image $\mathbf{f} = [f_{j,k}]$, the wavelet decomposition is obtained by separable filtering along the rows and columns of an image. The use of a pyramid-structured Wavelet Transform for texture analysis was first suggested in the pioneering work of Mallat [Mal89]. Pairs of wavelet filters including a lowpass filter g (scaling function) and a highpass filter h (wavelet function) are utilised to calculate the wavelet coefficients. In practice, the transform is computed by applying a separable filter bank to the image:

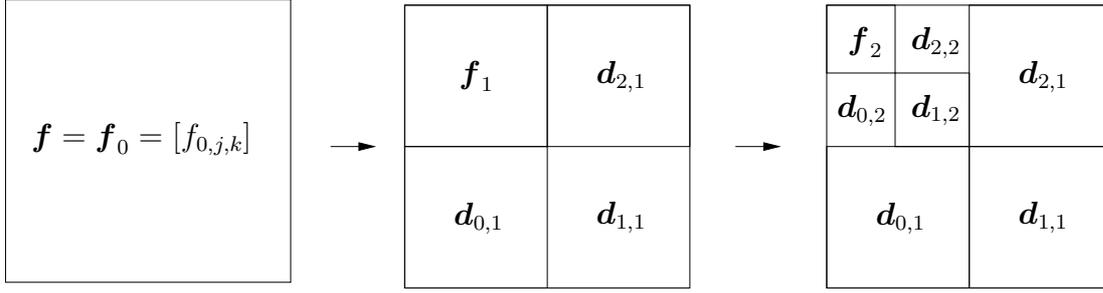


Figure 3.1: Twodimensional Wavelet Transform. f_μ is subsequently subdivided.

$$\mathbf{f}_\mu = \left[g_x * \left[g_y * \mathbf{f}_{\mu-1} \right]_{\downarrow 2,1} \right]_{\downarrow 1,2} \quad (3.6)$$

$$\mathbf{d}_{0,\mu} = \left[h_x * \left[g_y * \mathbf{f}_{\mu-1} \right]_{\downarrow 2,1} \right]_{\downarrow 1,2} \quad (3.7)$$

$$\mathbf{d}_{1,\mu} = \left[h_x * \left[h_y * \mathbf{f}_{\mu-1} \right]_{\downarrow 2,1} \right]_{\downarrow 1,2} \quad (3.8)$$

$$\mathbf{d}_{2,\mu} = \left[g_x * \left[h_y * \mathbf{f}_{\mu-1} \right]_{\downarrow 2,1} \right]_{\downarrow 1,2} \quad (3.9)$$

where $*$ denotes the convolution operator, $\downarrow 2,1(\downarrow 1,2)$ denotes the downsampling along the rows (columns) and \mathbf{f} is the original image.

Every subimage contains information of a specific scale and orientation of the coefficients. Spatial information is retained within the subimage. The original image \mathbf{f} is thus represented by a set of subimages at several scales at level μ . Subimage \mathbf{f}_μ is obtained by lowpass filtering and is referred to as the low resolution image. Its coefficients representing the approximation image, while the subbands labeled $\mathbf{d}_{0,\mu}, \mathbf{d}_{1,\mu}, \mathbf{d}_{2,\mu}$ represent the detail images at scale μ . The latter are obtained by bandpass filtering in a specific direction and thus contain directional information. Subimage $\mathbf{d}_{1,\mu}$ represents diagonal details while $\mathbf{d}_{2,\mu}$ gives horizontal high frequencies (vertical edges) and $\mathbf{d}_{0,\mu}$ contains vertical high frequencies (horizontal edges).

At the subsequent scale of analysis, the image \mathbf{f}_μ undergoes the decomposition using the same g and h filters, having always the lowest frequency component located in the upper left corner of the image as illustrated in figure 3.2. Each stage of the analysis produces four subimages whose size is reduced to the half compared to the previous scale.

A simple example for a scaling function as well as a wavelet basis function is the Haar function, defined as



Figure 3.2: Example of a three level Wavelet decomposition of an image with Haar basis [Nie07]

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \quad \psi(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2} \\ -1, & \frac{1}{2} \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \quad (3.10)$$

and applied in figure 3.2. The Haar scaling function and wavelet function is illustrated in figure 3.3.

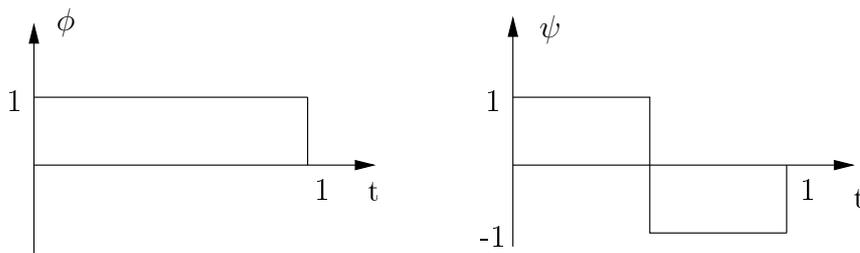


Figure 3.3: Haar functions $\phi(t)$ (scaling function) and $\psi(t)$ (wavelet function).

Other examples for wavelets are the Daubechies family [Dau92], a special form that was invented by Ingrid Daubechies or the Odegard wavelet [OB96].

3.2 Statistical methods

Statistical methods deal with observed combinations and relationships between the gray levels at specified positions. Statistics are classified into first-order, second-order and higher-order according to the level of pixel combination. The most popular second-order statistical features for texture analysis are derived from the so-called Co-occurrence Matrix [HDS73]. They are known to have a potential for effective texture discrimination in biomedical images [LSS⁺93]. Beside this, texture features provide measures of properties such as contrast, smoothness, coarseness, randomness, regularity, linearity, directionality, periodicity, and structural complexity. Local Binary Patterns (LBP) are another promising method for texture description.

3.2.1 Co-occurrence Matrix (GLCM)

A Graylevel Co-occurrence Matrix \mathbf{P} can be regarded as a second-order histogram of dimensions equal to the number of intensity levels, G , in the image. The matrix element $P_{\Delta x, \Delta y}(i, j)$ represents the absolute frequency with which two pixels with intensity i and j occur within a given neighborhood separated by a pixel distance Δx and Δy . Given a $M \times N$ image size of an input image \mathbf{I} containing G gray levels, let $I(m, n)$ be the intensity at image row m and column n .

$$P_{\Delta x, \Delta y}(i, j) = \sum_{n=1}^{N-\Delta y} \sum_{m=1}^{M-\Delta x} \begin{cases} 1, & \text{if } I(m, n) = i \wedge I(m + \Delta x, n + \Delta y) = j \\ 0, & \text{otherwise} \end{cases} \quad (3.11)$$

To compute the relative frequencies of each GLCM value, one must normalize the absolute values of $P_{\Delta x, \Delta y}(i, j)$ by using the following factor α

$$\alpha = \frac{1}{(M - \Delta x)(N - \Delta y)}. \quad (3.12)$$

A small 5×5 subimage with 4 gray levels and its corresponding GLCM $\mathbf{P}_{1,0}$ is illustrated in figure 3.4.

Another common notation of the GLCM is the usage of a distance-angle representation $\mathbf{P}_{a,\theta}$ as depicted in figure 3.5, where in most cases the computation is limited to the angles $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$, since the knowledge of $\mathbf{P}_{a,180}$, $\mathbf{P}_{a,225}$, $\mathbf{P}_{a,270}$, $\mathbf{P}_{a,315}$ adds nothing to the specification of the texture. For instance $\mathbf{P}_{a,180}$ can be regarded as the transpose of the matrix $\mathbf{P}_{a,0}$. Additionally, one can consider to compute a symmetric Co-occurrence Matrix out of this dependency.

If a rotation-invariant version of the GLCM wants to be achieved, one can calculate an average matrix out of the four matrices $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$.

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|-----------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Example Image | $P_{1,0}(i, j)$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1" style="border-collapse: collapse; text-align: center; width: 100px; height: 100px;"> <tr><td>0</td><td>2</td><td>3</td><td>3</td><td>3</td></tr> <tr style="border: 2px solid blue;"><td>3</td><td style="border: 2px solid blue;">1</td><td>2</td><td>3</td><td>3</td></tr> <tr><td>0</td><td>1</td><td>2</td><td>2</td><td>3</td></tr> <tr style="border: 2px solid blue;"><td>0</td><td style="border: 2px solid blue;">3</td><td style="border: 2px solid blue;">1</td><td>2</td><td>2</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>2</td></tr> </table> | 0 | 2 | 3 | 3 | 3 | 3 | 1 | 2 | 3 | 3 | 0 | 1 | 2 | 2 | 3 | 0 | 3 | 1 | 2 | 2 | 0 | 0 | 1 | 1 | 2 | <table border="1" style="border-collapse: collapse; text-align: center; width: 100px; height: 100px;"> <tr> <td style="border: none;">$i \backslash j$</td> <td style="border: none;">0</td> <td style="border: none;">1</td> <td style="border: none;">2</td> <td style="border: none;">3</td> </tr> <tr> <td style="border: none;">0</td> <td>1</td> <td>2</td> <td>1</td> <td>1</td> </tr> <tr> <td style="border: none;">1</td> <td>0</td> <td>1</td> <td>4</td> <td>0</td> </tr> <tr> <td style="border: none;">2</td> <td>0</td> <td>0</td> <td>2</td> <td>3</td> </tr> <tr> <td style="border: none;">3</td> <td>0</td> <td style="border: 2px solid blue;">2</td> <td>0</td> <td>3</td> </tr> </table> | $i \backslash j$ | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 1 | 1 | 1 | 0 | 1 | 4 | 0 | 2 | 0 | 0 | 2 | 3 | 3 | 0 | 2 | 0 | 3 |
| 0 | 2 | 3 | 3 | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1 | 2 | 3 | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 2 | 2 | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 3 | 1 | 2 | 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 1 | 1 | 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $i \backslash j$ | 0 | 1 | 2 | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 2 | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0 | 1 | 4 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | 0 | 0 | 2 | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 0 | 2 | 0 | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Figure 3.4: Example image and its Graylevel Co-occurrence Matrix with $\Delta x = 1$ and $\Delta y = 0$. The graylevel relation $i = 3$ and $j = 1$ is emphasized with an occurrence of 2.

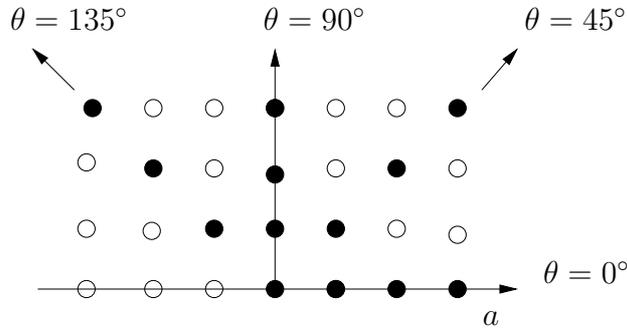


Figure 3.5: Geometry for angle representation for four distances a and four angles θ

It is obvious that coarser textures require larger values of the distance a , while it is recommended to choose $a = 1$ for finer resolutions to represent microstructures in the image.

A number of GLCM-based statistical features m can be calculated using the Co-occurrence Matrix for the purpose of texture discrimination; 14 of them were initially proposed by Haralick [HDS73]. The denotation of those measurements require the introduction of the following variables. The row and the column sums of the GLCM are described by P_x and P_y . Mean and standard deviation of those sums are denoted by μ_x , μ_y , σ_x , σ_y . $P_x(i)$ is the i th entry in the marginal-probability matrix obtained by summing the rows of $P(i, j)$.

$$P_x(i) = \sum_{j=0}^{G-1} P(i, j) \quad (3.13)$$

$$\mu_x = \sum_{j=0}^{G-1} i \sum_{j=0}^{G-1} P(i, j) = \sum_{j=0}^{G-1} iP_x(i) \quad (3.14)$$

$$\sigma_x^2 = \sum_{j=0}^{G-1} (i - \mu_x)^2 \sum_{j=0}^{G-1} P(i, j) = \sum_{j=0}^{G-1} (P_x(i) - \mu_x(i))^2 \quad (3.15)$$

This is equally defined for μ_y, σ_y . The following list contains the statistical measurements which can be extracted from the Co-occurrence Matrix.

- Angular Second Moment (ASM):

$$s_0 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j)^2 \quad (3.16)$$

Angular Second Moment can be regarded as a measure of homogeneity of an image. A homogeneous scene will contain only a few gray levels, resulting in a GLCM with only a few but high values of $P(i, j)$. Thus, the sum of squares of those values will be high.

- Correlation:

$$s_1 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{ijP(i, j) - \mu_x\mu_y}{\sigma_x\sigma_y} \quad (3.17)$$

Correlation is a statistical technique that shows whether and how strongly pairs of gray levels are related.

- Inverse Difference Moment (IDM):

$$s_2 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{P(i, j)}{1 + (i - j)^2} \quad (3.18)$$

IDM is influenced by the homogeneity of the image, for example IDM will get small contributions from inhomogeneous areas. The result is a low IDM value for inhomogeneous images and a high value for homogeneous images.

- Entropy:

$$s_3 = - \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j) \log_2 P(i, j) \quad (3.19)$$

Entropy describes the average information content and is a statistical measure of randomness.

- Cluster Shade:

$$s_4 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i + j - \mu_x - \mu_y)^3 P(i, j) \quad (3.20)$$

- Cluster Prominence:

$$s_5 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i + j - \mu_x - \mu_y)^4 P(i, j) \quad (3.21)$$

Cluster Shade and Cluster Prominence are measures of the skewness of the matrix, which can be seen as a lack of symmetry. When Cluster Shade and Cluster Prominence are high, the content of the image is not symmetric.

- Inertia (Contrast):

$$s_6 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - j)^2 P(i, j) \quad (3.22)$$

Inertia is a measure of local intensity variation that favours contributions away from the diagonal of the GLCM.

3.2.2 Local Binary Pattern (LBP)

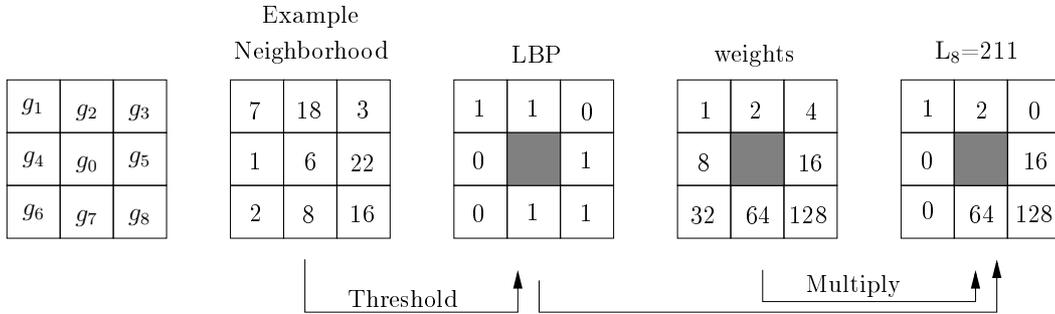


Figure 3.6: Calculation of the L_8 code from an example neighborhood

The Local Binary Pattern is defined as a grayscale invariant texture measure, derived from a general definition of texture in a local neighborhood. It was first introduced as a complementary measure for local image contrast in [OPH96] and had already found application in several polyp detection systems [KWL⁺00, WKKT01]. Several versions of this operator have been developed [POX00, OPM00, MOPS00] and will be shown here.

The original LBP method can be described as follows. Traditionally, the LBP works on a 3×3 neighborhood, therefore a subscript 8 for the eight neighbors is used in the following. The value of the center pixel g_0 is used as threshold for each of its eight neighbors as depicted in figure 3.6. If the value of the neighborhood pixel $g_i, i = 1, 2, \dots, 8$ is higher than the center pixel, 1 will be inserted at the corresponding position in the LBP, 0 otherwise. The LBP code for a neighborhood is then produced by multiplying the thresholded values with binomial weights given to the corresponding pixels, and summing up the result. The weights consist of power of two values. The following equation shows the computation of the LBP code L_8 .

$$L_8(g_0) = \sum_{i=1}^8 s(g_i - g_0) 2^{i-1} \quad s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (3.23)$$

A histogram can then be computed over the frequencies of the L_8 codes which occur in an image. The LBP is by definition invariant against any monotonic transformation of the gray scale. This means that only the contrast between the neighborhood and the center pixel is of importance instead of the gray value itself.

Opponent-Color LBP

An opponent-color version of the LBP was introduced [Mä03] and found application in [AWP⁺09]. The significant difference to the original LBP is the usage of single color channels of the RGB color space instead of the grayscale range. Furthermore, the feature consists of inter channel information, because the neighborhood is derived from a different color channel than the center pixel. All combinations of color channels result in six histograms. Additionally, three LBP-histograms are computed from each color channel separately. At the end, the frequencies held in nine histograms make up a feature vector.

Rotation-Invariant-LBP8

The original LBP code is not rotation invariant. It produces 256 different output values, corresponding to the $2^8 = 256$ different binary patterns that can be formed by the 3×3 neighborhood. Thus, rotating a particular binary pattern results in a different LBP code. From a local binary pattern, a 8-bit code can be clockwise generated as illustrated in 3.7 [POX00]. For the rotation invariant approach it is advisable to index the neighbor set in a way that they form a circular chain and interpolate the diagonal pixels as illustrated in figure 3.8.

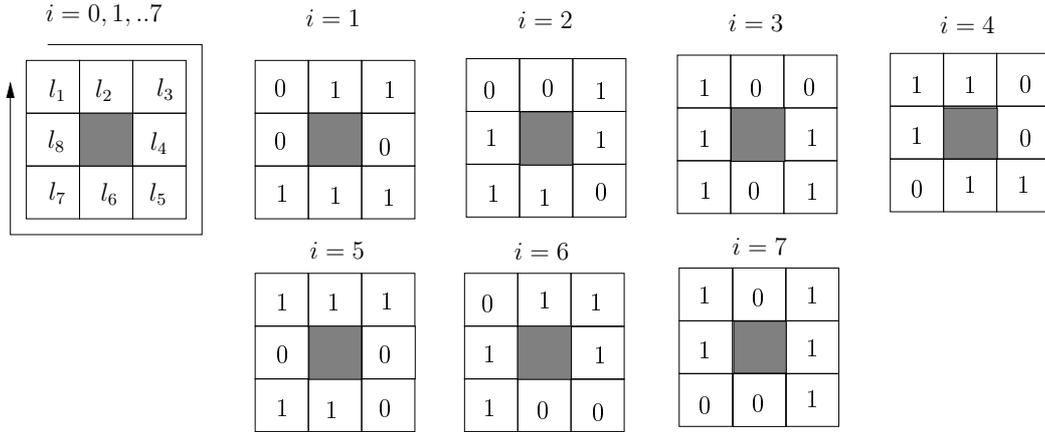


Figure 3.7: Rotating the LBP i -times clockwise. $i = 0$ is depicted in figure 3.6. The rotation-invariant bitcode for this pattern is obtained after three rotations ($i = 3$) with 001101111.

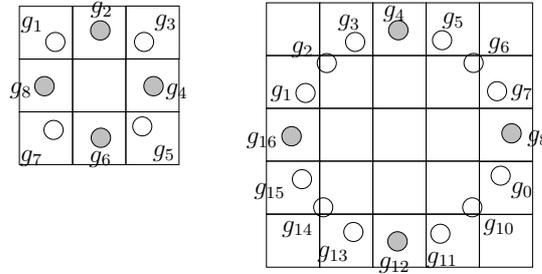


Figure 3.8: Interpolation of 3×3 neighborhood (left) and 5×5 neighborhood (right). Gray circles notate non-interpolated pixel, white circles are interpolated.

To remove the effect of rotation, an unique identifier is assigned to a pattern, that rotated has always the same bitsequence [OPM00].

$$L_8^{ri}(g_0) = \min \{ROR(L_8(g_0), i) \mid i = 0, 1, \dots, 7\} \quad (3.24)$$

Equation 3.24 defines a circular bit-wise right shift ROR on the 8-bit l_j i times. The superscript ri stands for rotation invariance. This corresponds to rotating the neighborhood clock-wise as many times as the maximal number of most significant bits is 0. Figure 3.9 shows the 36 possibilities for rotation invariant local binary patterns. Their frequencies are counted over the image and inserted into a histogram. It is superfluous to multiply the binomial weights with the LBP to compute the LBP code. Instead, the 8-bit pattern converted to a decimal number can be used as label.

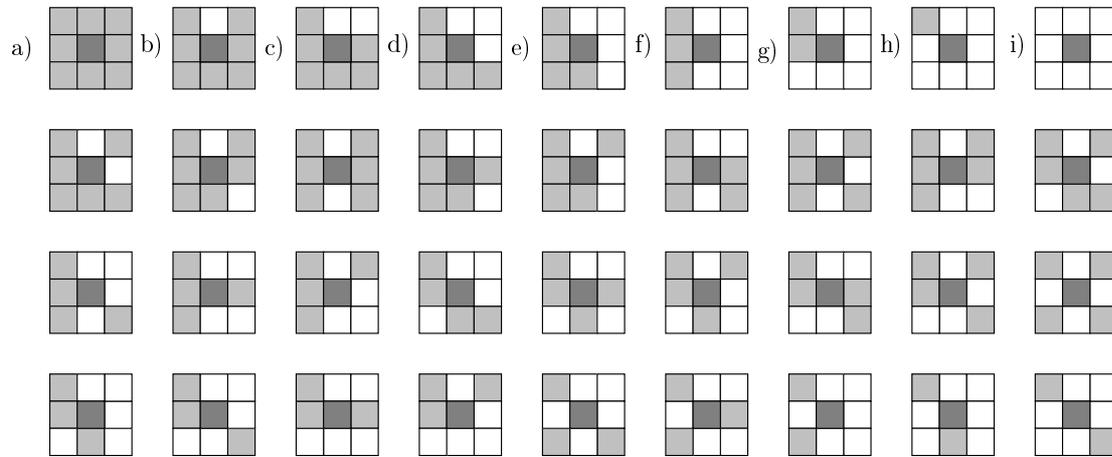


Figure 3.9: 36 unique rotation invariant LBP codes. Gray squares notate number 1, white squares 0.

Certain descriptions of microstructures can be expressed in L_8^{ri} considering the first row of figure 3.9. For example, pattern a) detects bright spots, pattern i) dark spots, while pattern c)-e) serves as edge description.

Subset-LBP8

Another LBP approach is to use subsets of the proposed schemes. Using all of the 36 patterns in figure 3.9 lead to a suboptimal result according to [MOPS00]. They state that some patterns sustain rotation better than other; the latter only confuse the analysis. They are more likely to be similar to a different structure upon rotation. To overcome this, only patterns are chosen for individual representation, that have a spatial transition (bitwise 0/1 changes) of at most two. For example, pattern 00000000 and 11111111 have 0 transitions, while the other seven patterns in the first row of figure 3.9 have two 0/1 transitions. Let U be a function that counts the 0/1 transitions in each pattern then the SubsetLBP8 is defined by the following equation:

$$L_8^{\text{subset}} = \begin{cases} L_8(g_0), & \text{if } L_8(g_0) \in S \\ 59, & \text{otherwise} \end{cases} \quad (3.25)$$

$$S = \{ x \mid 0 \leq x \leq 255 \wedge U(x) \leq 2 \} \quad (3.26)$$

Equation 3.25 assigns an unique label to the nine uniform patterns and their rotated versions, as illustrated in 3.9 a)-i) (11111111, 01111111, 00111111, 00011111, 00001111, 00000111, 00000011, 00000001, 00000000). This ends up in 58 frequency

bins. The 27 other rotation invariant patterns in figure 3.9 are being grouped under the miscellaneous label (59). Superscript set corresponds in this case to the use of rotation invariant uniform patterns as well as a subset of rotation variant pattern.

Rotation-Invariant-Subset-LBP16

A last LBP approach which is examined in this work is to use a bigger neighborhood than the L_8 operator does. The coarse 45° quantization level of the angular space through the 8 pixel neighborhood leads to a non-optimal representation [OPM00]. To address this, figure 3.8 shows a modification, where a clockwise defined neighborhood consisting of 16 pixel is presented. In this way, a finer resolution of 22.5° can be obtained. The gray values of neighbors which do not fall exactly in the center of pixels are computed by interpolation. The different spatial resolution can also be seen as advantageous when performing multiresolution analysis.

$$L_{16}(g_0) = \sum_{i=1}^{16} s(g_i - g_0)2^{i-1} \quad (3.27)$$

The L_{16} operator defined in equation 3.27 has $2^{16} = 65536$ output values and 243 rotation invariant patterns.

$$L_{16}^{\text{subset,ri}} = \begin{cases} \sum_{i=1}^{16} s(g_i - g_0) & , \quad \text{if } U(L_{16}) \leq 2 \\ 17 & , \quad \text{otherwise} \end{cases} \quad (3.28)$$

The first case shows that again only 17 patterns are used for individual labeling that have at most two 0/1 transitions. These correspond to the number of ones occurring in the bitcode, for example from 0 (pattern 0000000000000000) to 16 (pattern 1111111111111111). Label 17 groups the frequencies of all other patterns.

Chapter 4

Classification and Evaluation

After applying feature extraction methods on the image, learning a classifier is the subsequent step in a traditional pattern recognition system. A classification task usually involves training and testing data which consists of data samples. The disjunct separation into training and test sets can be done by crossvalidation. Each sample in the training set has one target value ω_i from a predefined set of class labels $\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}$ and values of the feature vector.

4.1 Crossvalidation

Crossvalidation is a commonly used technique to partition sample sets into complementary subsets. For instance, the k -fold crossvalidation separates the original sample into k subsets. $k - 1$ subsets are then used for training of the classifier, while the remaining data set serves as validation. This procedure is repeated k times, with each of the subsamples used once as testing data. The k results are then averaged over the folds. A stratified k -fold crossvalidation implies that the class distribution is retained in each subset.

4.2 Support Vector Machines

The Support Vector Machine (SVM) is a classifier that is already used in several approaches for the detection of polyps [IMK06] [LCK05] [KIM⁺03] [ACN07].

The goal of SVM is to produce a model which predicts classes of samples in the testing set, where only the features are given [CV95]. It can be seen as an extension of linear classifiers, where a linear decision function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ maps each feature to a positive or negative class. Linear classifiers are limited to certain applications, because not every problem is linearly separable as depicted in figure 4.1. To solve non-linear problems, a kernel function is utilized to project the

features into higher dimensional space. The approach for linear separable problems can be then applied in higher dimensions.

One representative kernel function is the radial basis function (RBF) that will commonly be used if the number of training objects is higher than the number of dimensions of the feature vectors [HCL08].

A hyperplane is used to separate the feature space in a way that features belonging to the same class are located on one side of a hyperplane. The most important property concerning hyperplanes is the distance between the feature and the hyperplane. The goal is to find the largest possible distance between the hyperplane and the features in the set.

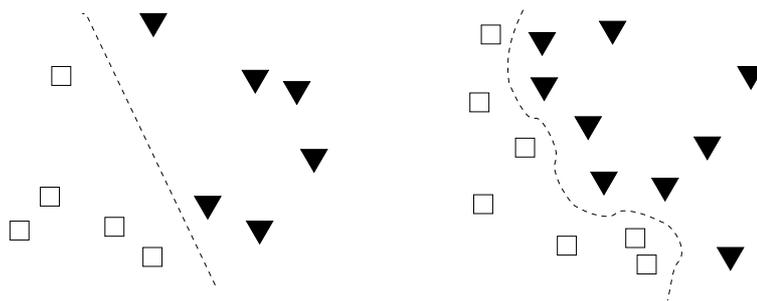


Figure 4.1: Example for linear separable feature space (left) and non-linear separable feature space (right).

4.3 k -Nearest Neighbor Classifier

The k -NN algorithm is a very simple approach for classification. It is a type of instance-based learning, or lazy learning, where all computation is deferred until classification. No explicit training step is required, because it consists only of storing the feature vectors and class labels of the training samples.

The multidimensional feature space is ideally partitioned into regions by locations and labels of the training samples. The objects, represented as vectors, are classified based on the closest training examples in the feature space. They are assigned to the class most common amongst its k nearest neighbors, based on a majority vote (figure 4.2).

On the one hand, large values of k reduce the effect of noise on the classification, on the other hand large values of k make boundaries between classes less distinct. It is recommended to choose k to be an odd number in two class classification problems to avoid ambiguous situations. Commonly, one uses Euclidean distance to determine the nearest neighbors, but in general any distance function could be applied [AKA91].

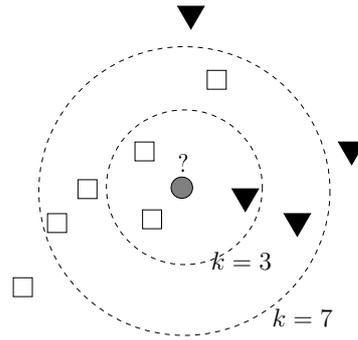


Figure 4.2: Example for k -Nearest Neighbor classifier.

| | sick person | healthy person |
|---------------|---------------------|---------------------|
| positive test | true positive (tp) | false positive (fp) |
| negative test | false negative (fn) | true negative (tn) |

Table 4.1: Medical confusion matrix for a two class classifier

4.4 Evaluation of Results

The results of the classification have to be evaluated. The motivation is to use a standard for evaluation such as a confusion matrix. Table 4.1 shows the confusion matrix for a two class classifier concerning medical issues.

The performance of a system is commonly evaluated using the data in the matrix for computing two statistical measures: specificity and sensitivity.

The specificity s measures the relation between determined healthy tissue and the real occurrence of healthy tissue. Hence it indicates the proportion of negatives which are correctly identified.

$$s = P(\text{negative test} | \text{healthy person}) = \frac{\text{tn}}{\text{tn} + \text{fp}} \quad (4.1)$$

The sensitivity t measures the proportion of actual positives which are correctly identified as such and is defined by the following conditional probability

$$t = P(\text{positive test} | \text{sick person}) = \frac{\text{tp}}{\text{tp} + \text{fn}}. \quad (4.2)$$

Considering classification of polyps, sensitivity describes the relation between actually detected polyps to the real number of polyps.

An optimal prediction can achieve 100% sensitivity (i.e. predict all people from the healthy group as healthy) and 100% specificity (i.e. predict all people from the sick group as sick).

sensitivity and specificity are closely related to the concepts of type I (α) and type II (β) errors. For example a false-positive result (healthy people wrongly identified as sick) is a type I error, while false-negative result (sick people wrongly identified as healthy) is a type II error.

Besides this, the receiver operating characteristic (ROC)-graph [GJ66] provides a further method for evaluation of classifiers. While sensitivity and specificity only rely on either positive or negative cases, the ROC-graph combines both. A threshold describing which object is assigned to each class can be varied by grouping pairs (s, t) of sensitivity and specificity together. Those pairs finally form the ROC-curve as illustrated in figure 4.3. The higher the true positive rate and smaller the false negative rate, the better is the prediction accuracy and the classifier.

The area under the ROC curve (AUC) can be computed from the ROC-graph as indication for the performance of the classification in one value. Considering (s_i, t_i) , $i = 1, 2..n$ with $s_0 = 0, t_0 = 0, s_n = 1$ and $t_n = 1$ as the points of the ROC-graph. The AUC a can be computed by the following equation

$$a = \sum_{i=1}^n \frac{1}{2} (t_i + t_{i-1}) (s_i - s_{i-1}) \quad (4.3)$$

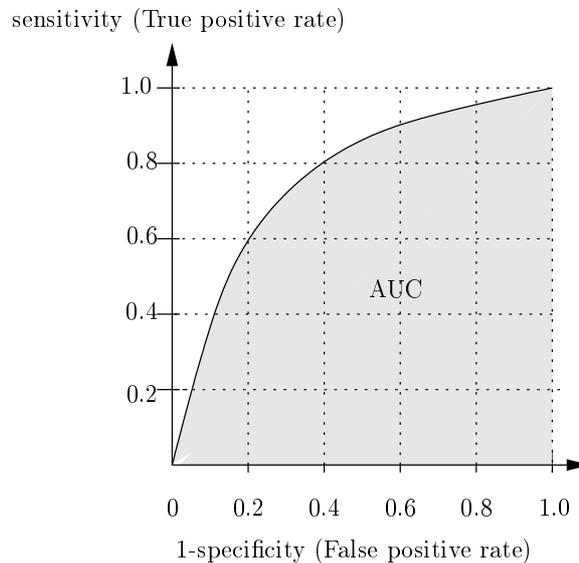


Figure 4.3: Example ROC Graph.

Chapter 5

System Description

This chapter introduces the single components of the system and gives an exact description of the classification scheme. Especially it is focussed on the changes with regard to the system of [AWP⁺09]. The software was programmed in C++. The following libraries are utilized:

- ITK - Segmentation and Registration Toolkit [ISNC05]
- QT - A cross-platform application and UI framework [QT]
- QWT - Qt Widgets [QWT]
- Wavelet - a Class Library for Wavelet Transforms on Images [Wav].

5.1 Data

The data base consists of four hours of video data from different colonoscopies initially used in [AWP⁺09]. The data has been evaluated by medical specialists from the Beaumon Hospital Dublin. The obtained ground-truth data was used to extract four scenes with polyps under varying illumination, view angle and distance. Each of the four scenes consists of approximately 400 single frames with a resolution of 800×800 pixel.

From the four scenes a heterogenous set of 130 frames is randomly chosen which is an important fact. In contrast to the test and training sets in the work from Ameling [AWP⁺09], this data set can be seen as quite heterogenous. [AWP⁺09] chose subsequent frames from the videos which are very similar leading to a duplication of nearly the same data in test and training set despite the use of crossvalidation.

To represent ground-truth data image masks are created as depicted in figure 5.1. The white region in the reference images describes the exact location, size and shape of a polyp.

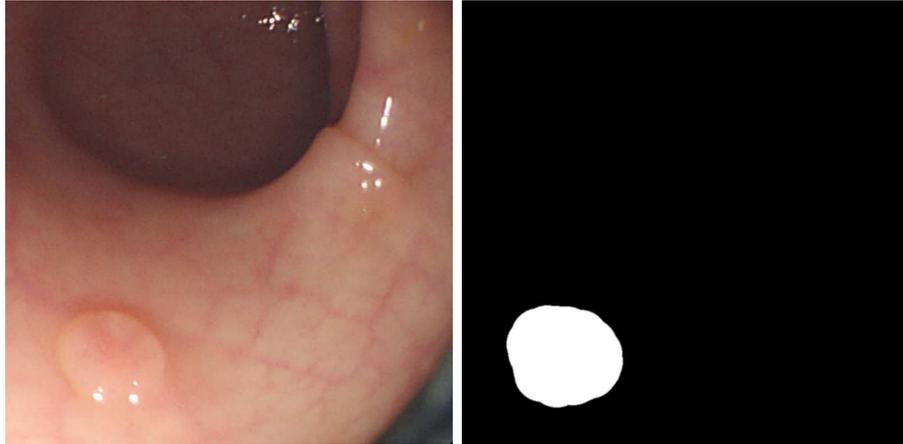


Figure 5.1: Endoscopic image and its reference mask

5.2 Patches

A patch approach is selected in the processing of each image. An image is subdivided into several square subimages, the so-called patches. Their size and degree of overlapping can be defined.

For each patch, a feature vector is computed and classified as *polyp* or *non-polyp*. This is a common technique in applications examining texture because texture features can be estimated on each of those subimages resulting in a local classification. The whole image was utilized for computation of texture features but the exact position of the polyp, if detected, would remain unknown.

Another technique is to apply region segmentation on endoscopic images such as Watershed Transform [VS91, DHHM06] or Region Growing [SC80]. Features can be computed separately from the detected regions instead of estimating them from patches. This method hardly depends on the performance of region segmentation and thus has not been considered here.

In [AWP⁺09] patches will be assigned to a class if the patch is completely filled with black or white pixels. Thereby the so called mask image serves as reference for class labeling, see figure 5.2. Two general remarks can be made on this approach. Patches which contain *polyp* and *non-polyp* information at the same time are not considered in the system. Thus, the border of a polyp remains completely unseen, although it may also contain important information for feature distinction and classification. Another aspect is that small polyps will not be detected, if the chosen patch size is bigger than the polyp. Consequently, the computed false-negative rate does not correspond to the real false-negative rate.

The approach in this work is to consider all patches. Patches that contain *polyp* as well as *non-polyp* information are classified corresponding to the following

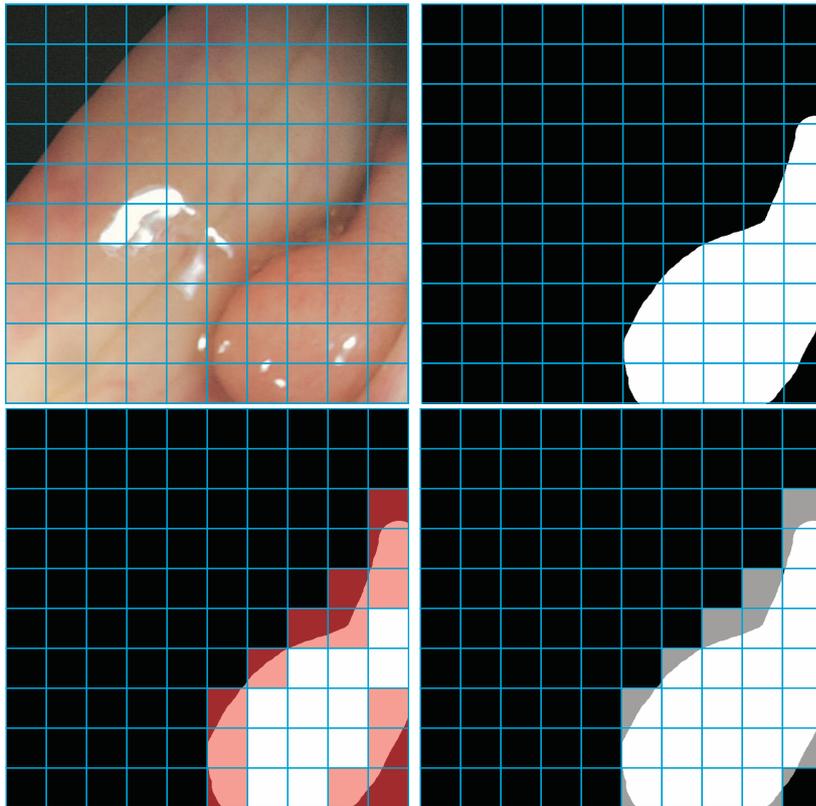


Figure 5.2: **up left:** endoscopic image divided in patches; **up right:** corresponding mask image **down left:** mask image - red colored patches are not considered in [AWP⁺09]; **down right:** mask image - the patches must contain over 625 of white pixels to be assigned to the *polyp* class

scheme: All intensity values in the mask image are counted. If the sum is equal or higher than value 625, the patch will be mapped to the *polyp* domain, otherwise it is a *non-polyp* patch.

Another aspect of the patch approach to consider is that some images are not fully subdivided depending on the endoscopic image size (800×800 pixel in this case) and the specified patch size as illustrated in figure 5.3 (left). Consequently, parts of the right and lower border remain unseen on each image due to the raster scan algorithm. To counteract this situation, a patch overlapping can be utilized, which is depicted in figure 5.3 (right). Neighboring patches overlap each other and all parts of the image are considered.

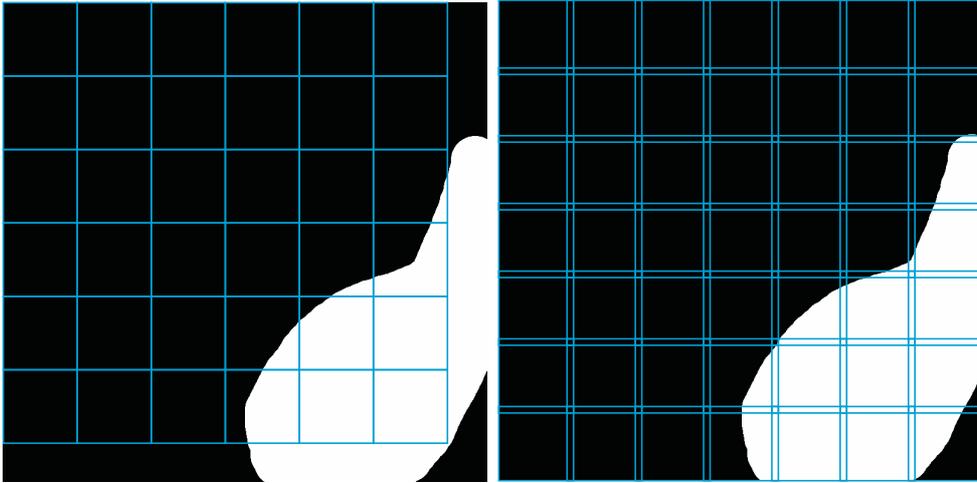


Figure 5.3: **left:** situation where the image size is not divisible without remainder by the patch size; **right:** applied patch overlapping

5.3 Experimental Flow

The extracted features must be classified and evaluated for instance through an AUC value. Therefore, a data mining software called WEKA [WEK], developed at the University of Waikato in New Zealand, provides adequate means. The software is written in Java and collects several machine learning algorithms. Tasks such as data pre-processing, classification, regression, clustering and visualization are manageable.

[AWP⁺09] utilizes in his work a library for Support Vector Machines called LibSVM [CL01]. The SVM is also available in WEKA which contains a wrapper class for LibSVM.

WEKA uses the attribute-relation file format (ARFF) as input. Such files have two distinct sections as shown in figure 5.4. The first section is the header information (1) which is followed by the data information (2). The header contains the name of the relation, e.g. the name of the feature and a list of the attributes which are the dimensions of the feature vector and their types. One attribute dimension is used as class assigner. In this case, the non-numeric classes *polyp* and *nonpolyp* serve as class labels. The second part consists of all data samples denoted by *@data*. Their attributes must correspond to the types defined in the header section. In figure 5.4 five data examples are given.

WEKA provides a knowledge flow interface for planning experiments with several classifiers at a time. In figure 5.5 the experimental flow of the tests and experiments presented in this work (chapter 6) is depicted. In (1) an ARFF-file is loaded, which is the output of the feature extraction module written in C++.

The *ClassAssigner* in (2) tells WEKA, which of the attributes describes the class. The *ClassValuePicker* (3) allows to choose the class label to be evaluated in the ROC. Step (4) produces a random subsample of the dataset and ensuring a special distribution spread of the classes. Different adjustments are possible in *SpreadSubsample* for instance a 1:1 or 1:2 class distribution of *polyp* samples to *non-polyp* samples can be chosen. [WP01] shows that a balanced training set produces better results. Thus, the class distribution in a training set can have a significant effect on the classification. Nevertheless, it is hard to predict which distribution is the best for a given problem. In the performed experiments, a fixed value of 1:1 has been chosen to overcome the problem that *non-polyp* patches occur more frequently than *polyp* patches. In step (5) all attributes are standardized to have zero mean and unit variance. Another preprocessing step was tested at this configuration step, called normalization. The *Standardization* resulted in a better performance and hence was integrated in the experimental flow.

A *CrossValidation* (6) is applied afterwards, which produces separated training and test sets for each of the n folds with $n = 4$. Subsequently, two different classifiers are trained on the generated set and both are validated on the test sets, called SVM (7a) and k -NN classifier (7b). For (7a) the same kernel type as in [AWP⁺09] is chosen which is a radial basis function (RBF). The RBF-kernel is useful for classification problems, where the number of training samples is much higher than the number of dimensions of the feature vector [HCL08]. (7b) is a k -nearest neighbor classifier with $k = 11$. The *PerformanceEvaluator* (8) evaluates the results of each classifier, given its results to a visualization tool (9a) as well as to a *TextViewer* (9b). In 5.6 an example output file illustrates the results of the classification with SVM. Different statistics can be analyzed, such as true positive rate, false positive rate, confusion matrix, ROC Area as well as correctly and incorrectly classified instances.

```

glcm6.arff
@relation glcm6
@attribute Mean_Energy numeric
@attribute Mean_Entropy numeric
@attribute Mean_IDM numeric
1 @attribute Mean_Inertia numeric
@attribute Mean_ClusterShade numeric
@attribute Mean_ClusterProminence numeric
@attribute class {polyp, nonpolyp}
@data
2 0.0085242, 7.57123, 0.379832, 16.773, 548.071, 43697.8, nonpolyp
0.000757848, 11.2385, 0.152345, 133.506, 373205, 1.03403e+08, nonpolyp
0.00283574, 9.92533, 0.242202, 98.1624, 314673, 5.72504e+07, nonpolyp
0.00608953, 8.69719, 0.349799, 64.3624, 453040, 9.54349e+07, nonpolyp
0.0064974, 8.06526, 0.387658, 6.7627, 11067.4, 662133, polyp

```

Figure 5.4: Example ARFF-File of the GLCM6-feature from [AWP⁺09].

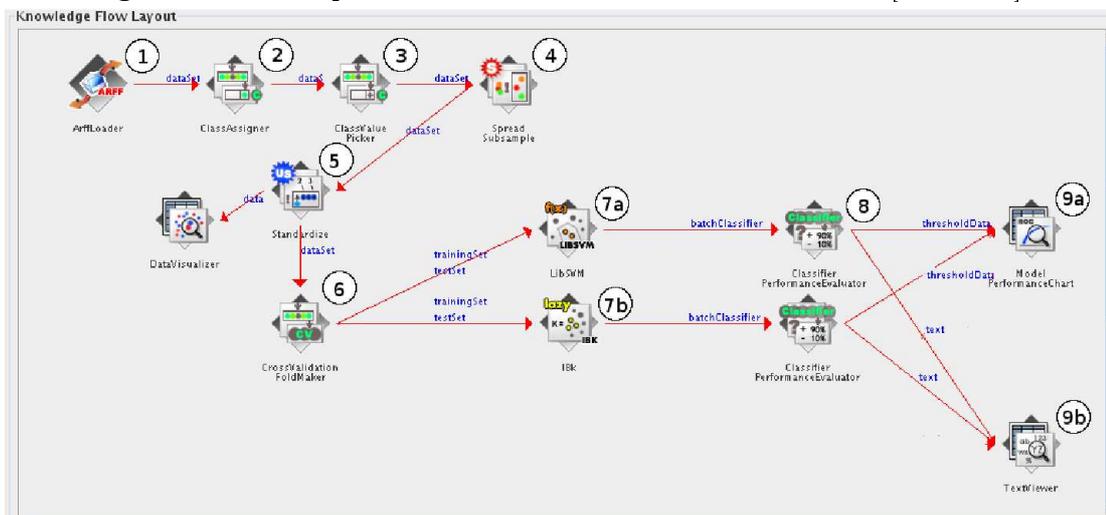


Figure 5.5: Experimental flow of the following tests with steps from (1)-(9a,9b)

```

classifier_LibSVM
=== Evaluation result ===

Scheme: LibSVM
Options: -S 0 -K 2 -D 3 -G 0.0 -R 0.0 -N 0.5 -M 40.0 -C 1.0 -E 0.0010 -P 0.1 -B -model /Users/engelhardt
Relation: glcm6-weka.filters.supervised.instance.SpreadSubsample-M1.0-X0.0-S1-
weka.filters.unsupervised.attribute.Standardize

Correctly Classified Instances      2487      68.5502 %
Incorrectly Classified Instances   1141      31.4498 %
Kappa statistic                    0.371
Mean absolute error                 0.4108
Root mean squared error             0.4532
Relative absolute error             82.1696 %
Root relative squared error         90.632 %
Total Number of Instances          3628

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
      0.729    0.358    0.671     0.729    0.699     0.74     polyp
      0.642    0.271    0.703     0.642    0.671     0.74     nonpolyp
Weighted Avg.   0.686    0.314    0.687     0.686    0.685     0.74

=== Confusion Matrix ===

  a  b  <-- classified as
1322 492 |  a = polyp
 649 1165 |  b = nonpolyp

```

Figure 5.6: Example output WEKA textfile of step (9b).

Chapter 6

Feature Descriptions and Experiments

This chapter gives an overview over the implemented features, starting with the features from [AWP⁺09] in section 6.1. They are tested on the new image set. Subsequently, new texture features are illustrated and evaluated. All of the applied methods such as Wavelet Transform, Graylevel Co-occurrence Matrix and Local Binary Patterns are already described in chapter 3, which provides the theoretical background. At the end of this chapter, a discussion about the performance of the features is attached.

In the following experiments these general aspects are examined considering the task of polyp detection:

- the performance of the classifiers
- the impact of color, color space and color channel
- the performance of texture features such as GLCM and LBP
- the impact of Wavelet Transform
- the impact of computing variances and covariances from the statistical measurements
- the patch size

6.1 Existing features

The polyp detection system of [AWP⁺09] implements four features, namely GLCM6, GLCM16, LBP and OC-LBP. For the purpose of a better comparison to the features implemented in this work, they are evaluated again by the new experimental flow described in section 5.3. The heterogeneous set of 130 endoscopic images is used as data base for feature extraction.

6.1.1 GLCM6, GLCM16, LBP, OC-LBP

The following table 6.1 shows the classification results for a patch size of 64×64 . They are specified by the area under the ROC curve (AUC) values of a k -Nearest-Neighbor classifier (k -NN) and a Support Vector Machine (SVM).

| Feature | RGB-Channel | | | Patch Size | AUC | |
|---------|-------------|---|---|----------------|---------|--------------|
| | R | G | B | | k -NN | LibSVM |
| GLCM6 | grayscale | | | 64×64 | 0.72 | 0.74 |
| GLCM16 | grayscale | | | 64×64 | 0.735 | 0.735 |
| LBP | grayscale | | | 64×64 | 0.75 | 0.76 |
| OC-LBP | × | × | × | 64×64 | 0.80 | 0.818 |

Table 6.1: Classification AUC values of the features implemented in [AWP⁺09]

6.1.2 Discussion

The AUC results from the tests applied in [AWP⁺09] are considerably higher than those depicted in table 6.1. The four homogeneous training and test sets used in [AWP⁺09] provide an easier way to classify *polyp* and *nonpolyp* patches. Some images hardly differ from each other, because successive frames are chosen and finally resulting in a loss of disjunction between training and test set.

Nevertheless, the overall essence of the tests from [AWP⁺09] corresponds to these results. The OC-LBP feature performs best by reason of using the RGB color space instead of grayscale images, since color has obviously a positive effect on classification of tissue images.

6.2 Wavelet Features

The Wavelet features introduced in the sections below are implemented due to the fact, that Wavelet Transform contributes to a better texture modelling [Mey93]. Varying spatial resolution allows it to represent textures at the most suitable scale.

It is examined, whether the Discrete Wavelet Transform has a positive effect on the classification of polyps. Therefore, different adjustments are tested such as decomposition level and basis function. This cannot be seen separately from the patch size, which determines the sizes of the subimages in subsequent levels of decomposition.

Additionally, the color approach has been further examined by testing different color spaces and color channels.

6.2.1 Color Wavelet and Color Wavelet Covariance

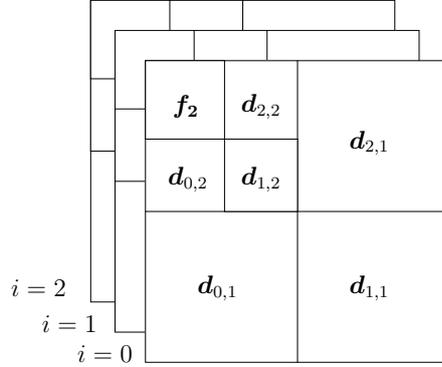


Figure 6.1: Wavelet detail components $\mathbf{d}_{l,\mu}^i$ with decomposition level μ and wavelet band l on color channel i and approximation component f_μ .

The Color Wavelet Covariance (CWC) feature, initially proposed by Karkanis et al. [KIM⁺03] was implemented and tested with different parameter adjustments. This method considers texture and color as information for discriminating polyps from normal tissue. In the following, several adjustments are described.

Considering the original image \mathbf{I} , one can obtain its color transformation from RGB in HSV, K-L, Cie-Lab space. Each of them results in three decomposed color channels $\mathbf{I}^i, i = 0, 1, 2$. Then a two level Discrete Wavelet Transform is applied on each color channel (\mathbf{I}^i) separately. The resulting nine subimages of the detail components $\mathbf{d}_{l,\mu}^i, l = 0, 1, 2$, from the second decomposition $\mu = 2$ are used for further processing (figure 3.1). Four GLCMs $\mathbf{P}_{a,\theta}$, with $a = 1, \theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ are computed on each of the nine subimages, resulting in 36 matrices.

$$\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^i) \quad i = 0, 1, 2, \quad l = 0, 1, 2, \quad a = 1, \quad \theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ \quad (6.1)$$

The number of intensities to compute GLCMs is reduced to 64 without any harmful implication in the resulted overall sensitivity, speeding up computation with only a minor loss of textural information. Four statistical measures s_m , called Energy ($m = 0$), Correlation ($m = 1$), Inverse Difference Moment ($m = 2$) and Entropy ($m = 3$) are extracted from these GLCMs, resulting in 144 texture values. They were initially proposed by Haralick [HDS73] from a set of 14 measures, defined in subsection 3.2.

$$s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^i)) \quad m = 0, 1, 2, 3 \quad (6.2)$$

In the proposed scheme, the $CWC_m^l(i, j)$ textural measure is finally estimated from covariances of the same statistical measure between color channels i, j at subimage $\mathbf{d}_{l,2}$.

$$\begin{aligned} & \text{cov}[s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^i)), s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^j))] \\ = & \sum_{\theta} [s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^i)) - E(s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^i)))] \times [s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^j)) - E(s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^j)))] \end{aligned} \quad (6.3)$$

This results in a 72 dimensional feature vector, consisting of 36 variances, as they relate features from the same color channel and 36 covariances from different channels.

$$CWC_m^l(i, j) = \begin{cases} \text{cov}[s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^i)), s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^j))], & \text{if } i < j \\ \text{var}[s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,2}^i))], & \text{if } i = j \end{cases} \quad (6.4)$$

For instance, detail component $\mathbf{d}_{0,2}^0$ of the second decomposition from the red channel is used to compute the GLCMs $\mathbf{P}_{1,0}$, $\mathbf{P}_{1,45}$, $\mathbf{P}_{1,90}$ and $\mathbf{P}_{1,135}$, where the measure Energy (s_0) is estimated. Those four measurements are used to compute relations between them (the variance).

| Feature | Basis | Patch Size | AUC | |
|--------------------------|-------|------------------|---------|--------------|
| | | | k -NN | LibSVM |
| Color Wavelet Covariance | Haar | 64×64 | 0.681 | 0.741 |
| | Haar | 128×128 | 0.748 | 0.773 |
| | Haar | 256×256 | 0.679 | 0.719 |
| | Daub8 | 128×128 | 0.724 | 0.77 |

Table 6.2: Color Wavelet Covariance Feature, test: patch size and basis function (Haar, Daubechies8) for second level of decomposition and RGB color space

Different modifications are performed on this feature, each time changing only one parameter. The first experiment reveals which basis and which patch size performs best. Only patch sizes of power of two are selected, as the author of [Wav] recommend, otherwise the Wavelet Transform suffers from a loss of precision. Table 6.2 shows the results. It turns out that the Haar basis, introduced in section 3.1, performs best with patch size 128×128 in contrast to the Daubechies-family (Daub8) and smaller or bigger patch sizes. Choosing subimages from the second decomposition level in the CWC features indicates, that it is probably not advisable to take a smaller patch size than 128×128 . This patch size leads to a size of

32×32 of the nine resulting subimages due to the wavelet decomposition, which might be a sufficient number of pixels to compute significant GLCMs.

A new feature called Color Wavelet was tested afterwards. This feature is similar to the Color Wavelet Covariance feature, but omitting to compute variances and covariances of the different color channels. The 144 statistical measures defined in equation 6.2 serve as input for the feature vector.

It turns out that this variant performs significantly better than the complete CWC feature. The results are depicted in table 6.3.

| Feature | Basis | Patch Size | AUC | |
|--------------------------|-------|------------------|---------|-------------|
| | | | k -NN | LibSVM |
| Color Wavelet Covariance | Haar | 128×128 | 0.748 | 0.773 |
| Color Wavelet | Haar | 128×128 | 0.793 | 0.82 |

Table 6.3: Color Wavelet Feature vs. Color Wavelet Covariance Feature for RGB color space

| Feature | Decomp | Patch Size | AUC | |
|---------------|--------|------------------|---------|-------------|
| | | | k -NN | LibSVM |
| Color Wavelet | 1 | 128×128 | 0.764 | 0.769 |
| | 2 | 128×128 | 0.793 | 0.82 |

Table 6.4: Color Wavelet Feature, test: decomposition for Haar basis and RGB color space

Considering the decomposition of the transform that is incorporated in the Color Wavelet feature, it is also possible to extract the features from the first level of decomposition. Table 6.4 holds the results and indicates that the second level is still a better choice for 128×128 pixel of patch size. Thus, the spatial resolution of the second scale level holds superior information for polyp detection.

Different basis functions are tested again for the Color Wavelet feature (table 6.5). The Haar basis performs best in comparison to the Daubechies-family (Daub4 and Daub8) and to a basis function called Odegard [OB96].

In a next step it is examined whether a certain color space is adequate for this feature. As proposed in [KIM⁺03] each patch is transformed from RGB in either HSV, K-L or CieLab color space before applying the Discrete Wavelet Transform on each channel. Table 6.6 shows that RGB color space is still the best choice for the Color Wavelet feature.

Afterwards it is evaluated whether a certain color channel or color channel combination of the RGB space is superior in detecting polyps. Table 6.7 shows that the best application for the Color Wavelet feature is to use all channels.

| Feature | Basis | Patch Size | AUC | |
|---------------|---------|------------------|---------|-------------|
| | | | k -NN | LibSVM |
| Color Wavelet | Haar | 128×128 | 0.793 | 0.82 |
| | Daub4 | 128×128 | 0.749 | 0.801 |
| | Daub8 | 128×128 | 0.759 | 0.798 |
| | Odegard | 128×128 | 0.678 | 0.727 |

Table 6.5: Color Wavelet Feature, test: basis function for RGB-colorspace

| Feature | Color Space | Patch Size | AUC | |
|---------------|-------------|------------------|---------|-------------|
| | | | k -NN | LibSVM |
| Color Wavelet | RGB | 128×128 | 0.793 | 0.82 |
| | K-L | 128×128 | 0.773 | 0.791 |
| | HSV | 128×128 | 0.759 | 0.781 |
| | CieLab | 128×128 | 0.791 | 0.801 |

Table 6.6: Color Wavelet Feature, test: color spaces for Haar-basis function

Nevertheless, the combination of red and blue channel performs slightly better than AUC values from other channels.

| Feature | RGB-Channel | | | AUC | |
|---------------|-------------|---|---|---------|--------------|
| | R | G | B | k -NN | LibSVM |
| Color Wavelet | × | | | 0.767 | 0.776 |
| | | × | | 0.728 | 0.758 |
| | | | × | 0.713 | 0.743 |
| | × | × | | 0.776 | 0.803 |
| | | × | × | 0.752 | 0.78 |
| | × | | × | 0.796 | 0.812 |
| | × | × | × | 0.793 | 0.82 |

Table 6.7: Color Wavelet Feature, test: combination of color channel for 128×128 patch size and Haar-basis

To outline the most essential results, the following box shows the best adjustments for the Color Wavelet feature, resulting from the applied tests.

| |
|--|
| Patch size: 128×128 Discrete Wavelet Transform: Haar-basis, 2nd level of decomposition Color Space: RGB, all color channels Measurements: Energy, Correlation, IDM, Entropy Best AUC: 0.82 |
|--|

6.2.2 Wavelet-Decomposition

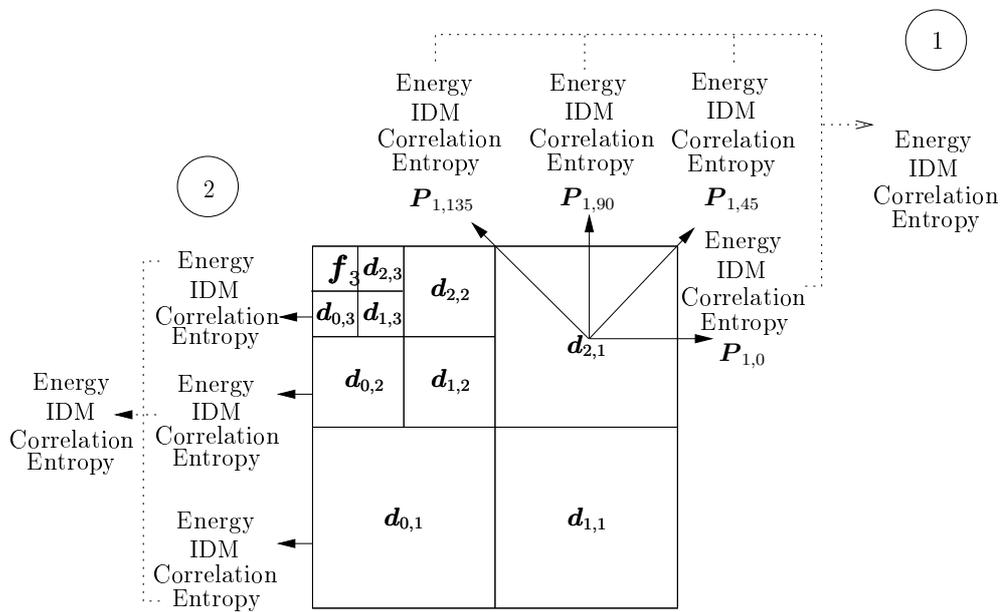


Figure 6.2: Example computation of $s_m^{l,\mu,i}$ for $l=2, \mu=1$ (step 1) and $s_m^{l,i}$ for $l=0$ (step 2) of the Wavelet-Decomposition feature, regarding the statistical measurements Energy ($m=0$), Correlation ($m=1$), Inverse Difference Moment ($m=2$), and Entropy ($m=3$) on color channel i .

The Wavelet-Decomposition feature is an approach for multiresolutional analysis. Subsequent scales of the Wavelet Transform are used to compute the feature. A three level DWT is applied to each color channel $i=0,1,2$ of the RGB color image. The detail components $d_{l,\mu}^i$, $\mu=1,2,3$ of each subband $l=0,1,2$

are utilized for further processing. Four GLCMs are computed on each of the 27 subimages resulting in $27 \cdot 4 = 108$ GLCMs.

$$\mathbf{P}_{a,\theta}(\mathbf{d}_{l,\mu}^i) \quad i = 0, 1, 2, \quad l = 0, 1, 2, \quad \mu = 1, 2, 3, \quad a = 1, \quad \theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ \quad (6.5)$$

For each GLCM four statistical measures s_m are extracted, namely Energy, Correlation, IDM and Entropy resulting in 432 measurements.

$$s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,\mu}^i)) \quad m = 0, 1, 2, 3 \quad (6.6)$$

Then for each subband, the measures of the four GLCMs are averaged over the angle value as shown exemplarily in figure 6.2.2, **step 1**. For instance, there is only one Energy measure per subimage left after averaging the Energy of GLCM 0° , 45° , 90° , and 135° .

$$s_m^{l,\mu,i} = \text{mean}_\theta[s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,\mu}^i))] = \frac{1}{4} \sum_\theta s_m(\mathbf{P}_{a,\theta}(\mathbf{d}_{l,\mu}^i)). \quad (6.7)$$

Subsequently, the mean is computed over the measurements from different decomposition levels, which is depicted in figure 6.2.2, **step 2**. For example considering color channel red, all Energy values from detail images $d_{0,1}$, $d_{0,2}$, $d_{0,3}$ are averaged.

$$s_m^{l,i} = \text{mean}_\mu[s_m^{l,\mu,i}] = \frac{1}{3} \sum_\mu s_m^{l,\mu,i}. \quad (6.8)$$

All in all the feature vector has 36 dimensions composed of 12 measures per color channel.

| Feature | Patch Size | Basis | AUC | |
|----------------------|------------------|-------|--------------|--------|
| | | | k -NN | LibSVM |
| WaveletDecomposition | 64×64 | Haar | 0.772 | 0.791 |
| | 128×128 | Haar | 0.796 | 0.795 |
| | 256×256 | Haar | 0.75 | 0.776 |

Table 6.8: Wavelet-Decomposition feature, test: patch size for RGB color space

In table 6.8 tests for the best patch size are shown. Only power of two values are applied. Patch size 128×128 performing best for this feature. A three level DWT implies, that the subimages on the third level have only a dimension of 16×16 using 128×128 patch size, and only 8×8 pixel using patch size 64×64 . In a subsequent step GLCMs are computed. Thus it would not make sense to decrease the patch size for this feature.

Subsequently it was tested whether one basis function is superior to the other choosing a patch size of 128×128 pixel. Table 6.9 implies that the Daubechies8 basis function works marginal better than the Haar basis function and its relative Daubechies4, considering the LibSVM results. For the k -NN classifier, the Haar basis is still the best choice.

| Feature | Patch Size | Basis | AUC | |
|----------------------|------------------|-------|---------|--------------|
| | | | k -NN | LibSVM |
| WaveletDecomposition | 128×128 | Haar | 0.796 | 0.795 |
| | 128×128 | Daub4 | 0.756 | 0.789 |
| | 128×128 | Daub8 | 0.773 | 0.799 |

Table 6.9: Wavelet-Decomposition feature, test: basis function for RGB color space

To summarize the results the following box shows the best adjustments for the Color Decomposition feature, resulting from the applied tests.

| |
|---|
| <p>Patch size: 128×128 Discrete Wavelet Transform: D8-basis, 3 level of decomposition Color Space: RGB Measurements: Energy, Correlation, IDM, Entropy Best AUC: 0.799</p> |
|---|

6.3 GLCM features

The Graylevel Co-occurrence Matrix features are implemented in order to compare them with the Wavelet features which also incorporate the computation of GLCMs. The purpose is to identify whether there is a benefit from the applied numeric transform.

Another aspect considered here is the impact of color. Reference values already exist in the grayscale features GLCM6 and GLCM16. Different color channels and color channel combinations are examined.

In GLCM6 and GLCM16 different numbers of statistical features are estimated from the Co-occurrence Matrices. These adjustments are also tested in the following.

Various patch sizes found application in the experiments. It is possible to choose much smaller sizes than in the Wavelet features because no decomposition is applied. Additionally, it is examined whether the patch overlapping has positive effects on the classification of polyps.

6.3.1 ColorGLCM

The ColorGLCM-feature is very similar to the Color Wavelet-feature, only omitting Discrete Wavelet Transformation. It utilizes different color channels of the RGB color space for the extraction of four GLCMs. The number of intensities to compute GLCMs is reduced to 64, which also holds true for all other features that incorporates GLCM computation.

$$\mathbf{P}_{a,\theta}(\mathbf{I}^i) \quad i = 0, 1, 2, \quad a = 1, \quad \theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ \quad (6.9)$$

Then for each of the twelve GLCMs four statistical measures s_m are extracted, namely Energy, Correlation, IDM, Entropy and additionally two values in a second test, namely Cluster Shade and Cluster Prominence are computed.

$$s_m(\mathbf{P}_{a,\theta}(\mathbf{I}^i)) \quad m = 0, 1, 2, 3, 4, 5 \quad (6.10)$$

| Feature | number of m | Patch Size | AUC | |
|-----------|---------------|----------------|---------|--------------|
| | | | k -NN | LibSVM |
| ColorGLCM | 4 | 64×64 | 0.823 | 0.83 |
| | 6 | 64×64 | 0.828 | 0.838 |

Table 6.10: ColorGLCM Feature, test: number of statistical measurements

Table 6.10 shows the marginal positive effect of using additionally Cluster Shade and Cluster Prominence as statistical measurements for GLCM description. The more measurements are extracted from GLCMs, the more dimensions the feature vector has. For $m = 0, 1, 2, 3$, the feature vector holds 48 values while for $m = 0, 1, 2, 3, 4, 5$ it is 72 dimensional. Both are acceptable values with regard to computational complexity. Haralick initially proposed 14 features in [HDS73], but it is not advisable to use all measurements. Some of them correlate to each other. In this case, Connors et al. [CTH84] propose using a set of the six features applied here.

In another test, it was examined whether it is advisable to use small patch sizes (table 6.11). A patch size of 32×32 pixel works very well for this feature, indicating that a maximum number of 399 pixel *non-polyp* information ($\approx 40\%$) is on a *polyp* patch. Remember that a minimum number of 625 pixel, which correspond to 25×25 pixel must belong to class *polyp* to classify the patch as *polyp*.

The best results for the ColorGLCM feature are shown below.

| Feature | number of m | Patch Size | AUC | |
|-----------|---------------|----------------|--------------|--------|
| | | | k -NN | LibSVM |
| ColorGLCM | 6 | 32×32 | 0.843 | 0.835 |
| | 6 | 64×64 | 0.828 | 0.838 |

Table 6.11: ColorGLCM Feature, test: patch size

| |
|---|
| <p>Patch size: 32×32 Color Space: RGB Measurements: Energy, Correlation, IDM, Entropy, ClusterShade, ClusterProminence Best AUC: 0.843</p> |
|---|

6.3.2 OC-GLCM

The Opponent-Color GLCM feature relates pairs of color channels by calculating GLCMs from the pixels of different color channels. Thus, it can be denoted as an inter-channel feature considering texture as well as color. Nine GLCMs are extracted from each combination of channels: red-green, red-blue and green-blue (without considering permutations). A total of 27 GLCMs are computed expressed by the following equation

$$\mathbf{P}_{\Delta x, \Delta y}(\mathbf{I}^{i,j}) \quad i, j = 0, 1, 2, \quad i \neq j, \quad \Delta x, \Delta y = -1, 0, 1. \quad (6.11)$$

Δx and Δy denote the distance to the center pixel from another color channel in x and y -direction as depicted in figure 6.3. Four statistical measures s_m , called Energy, Correlation, IDM and Entropy are estimated from this set of OC-GLCMs forming a feature vector of $9 \cdot 3 \cdot 4 = 108$ dimensions.

$$s_m(\mathbf{P}_{\Delta x, \Delta y}(\mathbf{I}^{i,j})) \quad m = 0, 1, 2, 3 \quad (6.12)$$

Several experiments are performed on this feature. Table 6.12 shows the results of running the OC-GLCM feature on different patch sizes. The feature works best with a patch size of 64×64 .

Considering the sizes of the endoscopic images (800×800) and the applied patch size, it is noticeable that the images are not fully subdivided. Parts of the right and lower border remain unseen on each image as figure 5.3 shows, which is undesirable and a loss of information as well. To solve this problem, patch overlapping can be applied. A patch size of 64×64 with 3 overlapping pixels fits

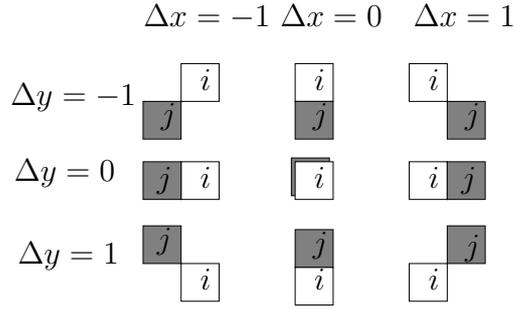


Figure 6.3: OC-GLCM feature: depiction of pixel relation from color channel i (white) and j (gray), where $i \neq j$. The white pixel represents the center pixel.

| Feature | Patch Size | Overlap | AUC | |
|---------|------------------|---------|--------------|--------------|
| | | | k -NN | LibSVM |
| OC-GLCM | 32×32 | 0 | 0.827 | 0.810 |
| | 64×64 | 0 | 0.831 | 0.832 |
| | 128×128 | 0 | 0.799 | 0.812 |
| | 64×64 | 3 | 0.832 | 0.848 |
| | 64×64 | 32 | 0.871 | 0.856 |
| | 128×128 | 16 | 0.801 | 0.795 |

Table 6.12: OC-GLCM Feature, test: patch size and overlapping

better in the given image size and leads to a higher AUC. Another interesting attempt is to use a patch size of 64×64 with 32 overlapping pixels, which can be seen as scanning the image in two different raster, performs best at all. This is also superior to applying a smaller patch size to the image.

It is possible that those very positive results constitute from the fact that certain information occurs twice or more in the training and test set, leading to a loss of disjunction. An indication for this is the superior performance of the k -NN classifier, which takes the most similar feature vectors as a basis for classification. A proposal is to strictly separate the endoscopic images that are used for training and testing. At the current stage, features are extracted first from the patches of all images and then the separation in training and test set is done.

A second test series evaluated whether a certain combination of color channels is more discriminating in feature space than other ones. A fixed patch size (64×64 pixel) and no overlapping is used for the experiments, shown in table 6.13. Involving all color channel combinations lead to a higher AUC (0.832 for LibSVM) than using only one channel combination or two combinations.

The most promising adjustments are illustrated in the following box.

| Feature | RGB-Channel | | | AUC | |
|---------|-------------|----|----|--------------|--------------|
| | RG | GB | RB | k -NN | LibSVM |
| OC-GLCM | × | | | 0.728 | 0.682 |
| | | × | | 0.775 | 0.783 |
| | | | × | 0.759 | 0.767 |
| | × | × | | 0.816 | 0.816 |
| | | × | × | 0.808 | 0.803 |
| | × | | × | 0.786 | 0.790 |

Table 6.13: OC-GLCM Feature, test: combination of color channel

Patch size: 64×64

Color Space: RGB, all color channels

Measurements: Energy, Correlation, IDM, Entropy

Best AUC: 0.832

6.4 LBP features

In the following, several variants of the Local Binary Patterns are tested. The aim is to compare the existing approaches OC-LBP and LBP with the new implemented features, evaluating the impact of color, inter- and intra-channel combinations. Also the different feature vector dimensions are examined. Due to the fast computation speed of LBP, it is not required to attempt a reduction of number of color channels.

It is focussed primarily on the various LBP approaches. They differ in the size of the neighborhood, in circular and non-circular representation (interpolation) and in the considered patterns. It should be evaluated which variant performs best.

6.4.1 ColorLBP

The ColorLBP feature is the application of the LBP-feature on the color channels of the RGB color space, which is defined in section 3.2.2. This feature can be seen as a subset of the OC-LBP-feature from [AWP⁺09], where the center pixel g_0 as well as the neighborhood $g_i, i = 1, 2, \dots, 8$ derive from the same channel. The number of bins in the LBP histogram is reduced to 64 due to computational efficiency, resulting in a feature vector of $3 \cdot 64 = 192$ dimensions.

Table 6.14 shows the results of the tests, comparing simple LBP-feature, OC-LBP-feature from [AWP⁺09] and ColorLBP-feature as well as different color channels to each other.

| Feature | RGB-Channel | | | Bins | Dimensions | AUC | |
|----------|-------------|---|---|------|------------|--------------|--------------|
| | R | G | B | | | <i>k</i> -NN | LibSVM |
| LBP | grayscale | | | 64 | 64 | 0.75 | 0.76 |
| | grayscale | | | 256 | 256 | 0.697 | 0.72 |
| OC-LBP | × | × | × | 64 | 576 | 0.880 | 0.818 |
| ColorLBP | × | × | × | 64 | 192 | 0.814 | 0.834 |
| | × | | | 64 | 64 | 0.785 | 0.756 |
| | | × | | 64 | 64 | 0.758 | 0.781 |
| | | | × | 64 | 64 | 0.739 | 0.764 |

Table 6.14: LBP vs. OC-LBP vs. ColorLBP Feature, test: color channel and histogram bins

The first two rows show that a reduced number of bins in the LBP histogram has overall advantageous effects on the classification. Considering the whole table 6.14, it can be observed that the ColorLBP-feature and the OC-LBP feature perform significantly better than the simple grayscale LBP feature. This indicates again that color plays an important role in classification of tissue.

Analyzing the area under the ROC curves of OC-LBP and its subset ColorLBP leads to the assumption that the feature vector of OC-LBP has too many dimensions in feature space, leading to a more complex classification task.

The last three rows of table 6.14 indicate that if one combines the histograms of all color channels will perform superior in comparison to the usage of only one single color channel. Each of them has nearly the same AUC rate as the grayscale LBP.

A summarization of the feature is given in the following.

Patch size: 64×64
LBP-Histogram bins: 64
Color Space: RGB, all color channels
Best AUC: 0.834

6.4.2 Rotation-Invariant-LBP8

Introduced in 3.2.2, the rotation invariant L_8^{ri} is implemented on a 3×3 neighborhood.

In a first experiment, the L_8^{ri} is tested on the single color channels of the RGB color space with patch size of 64×64 pixel. On each channel, the frequencies of 36 LBP-invariant patterns are counted, resulting from the shifted 8-bit codes as illustrated in figure 3.7 and figure 3.9. All in all a feature vector of $3 \cdot 36 = 108$ dimensions is obtained.

Additionally, it was tested whether an interpolation of the diagonal pixels (g_1, g_3, g_5, g_7) (figure 3.8, left) lead to a better circular representation, particularly with regard to rotational invariance.

| Feature | RGB-Channel | | | Interpolation | AUC | |
|-------------------|-------------|---|---|---------------|---------|--------------|
| | R | G | B | | k -NN | LibSVM |
| L_8^{ri} | × | × | × | | 0.78 | 0.79 |
| | × | × | × | × | 0.783 | 0.792 |

Table 6.15: Rotation-Invariant-LBP8-Feature, test: L_8^{ri} neighborhood diagonal interpolated vs. non-interpolated

Table 6.15 indicates that the interpolation of the diagonal pixels from the 3×3 neighborhood has slightly advantageous effects on the classification.

The most important facts of the feature are listed in the following.

Patch size: 64×64
Neighborhood: 3×3 , interpolated
LBP-Histogram bins: 36 per channel
Color Space: RGB, all color channels
Best AUC: 0.792

6.4.3 Rotation-Invariant-Subset-LBP16

Equation 3.28 defines a rotation invariant version of the LBP operator with a neighborhood consisting of 16 pixel. Only a subset of 17 patterns of the 243 rotation invariant patterns are used for assigning their frequencies to single bins. All other patterns are classified to a miscellaneous label. Hence, the feature vector consists of $18 \cdot 3 = 54$ dimensions, due to the fact that the feature is applied on each RGB color channel. A patchsize of 64×64 pixel is applied. The pixels

$g_1, g_2, g_3, g_5, g_6, g_7, g_9, g_{10}, g_{11}, g_{13}, g_{14}, g_{15}$ (figure 3.8, right) are estimated by interpolation concerning circular representation.

| Feature | RGB-Channel | | | Interpolation | AUC | |
|-----------------------------|-------------|---|---|---------------|---------|--------------|
| | R | G | B | | k -NN | LibSVM |
| $L_{16}^{\text{subset,ri}}$ | × | × | × | × | 0.78 | 0.799 |

Table 6.16: Rotation-Invariant-Subset-LBP16-feature

Comparing the L_8^{ri} with the L_{16}^{ri} feature leads to the conclusion that both features perform nearly the same. Thus, using a bigger neighborhood and a less number of rotational invariant patterns of the LBP does not have an impact on the classification of polyp images.

Significant properties of the feature are listed below.

Patch size: 64×64
Neighborhood: 16 pixel, interpolated
LBP-Histogram bins: 18 per channel
Color Space: RGB, all color channels
Best AUC: 0.799

6.4.4 Subset-LBP8

Section 3.2.2 describes a feature that combines rotation invariant and variant values called Subset-LBP8 LBP_8^{subset} . Only the local binary patterns 00000000, 00000001, 00000011, 00000111, 00001111, 00011111, 00111111, 01111111, 11111111 and their rotated versions are counted each of them separately in a histogram bin. The other occurring patterns are grouped under the miscellaneous label. All in all this feature forms a vector of 59 dimensions, which are again extracted from the single RGB color channels, resulting to 177 dimensions. Interpolation is performed on the provided 3×3 neighborhood as depicted in figure 3.8 (left).

| Feature | RGB-Channel | | | Interpolation | AUC | |
|-------------------------|-------------|---|---|---------------|---------|--------------|
| | R | G | B | | k -NN | LibSVM |
| LBP_8^{subset} | × | × | × | × | 0.816 | 0.835 |

Table 6.17: Subset8-LBP-feature

Comparing L_8^{subset} to the other LBP-features, this feature performs best. The AUC result is similar to the Color-LBP feature, probably due to the equal vector dimensions.

Significant properties of the feature are listed below.

Patch size: 64×64
Neighborhood: 3×3 pixel, interpolated
LBP-Histogram bins: 59 per channel
Color Space: RGB, all color channels
Best AUC: 0.835

6.5 Discussion of Results

- **Classifier:** Comparing the classification results from the Support Vector Machine with the k -NN classifier leads to the conclusion that with less exceptions the SVM has higher AUC values. There is no clear scheme recognizable in which cases the k -NN is better. For example, for the ColorGLCM feature and OC-GLCM feature the k -NN performs better when a small patch size was used. Testing a reduced number of color channels with OC-GLCM and ColorLBP features led to a higher AUC than the result of SVM in three times. Nevertheless, comparing the best classification results for each of the twelve features, the SVM holds the better results in eleven cases as shown in table 6.18.

Applying a patch overlapping to the endoscopic images yield in training and test sets to a duplication of the same data, leading to a better classification in favor of the k -NN classifier. The results are shown in table 6.12. Those findings are not considered in the overall evaluation, because of non-disjunction of test and training set.

The best AUC result of all applied tests resulted from the ColorGLCM feature. The k -NN classifier performed best in this case with an AUC of 0.843.

All in all, the classification differences between the two classifiers are small with a maximum difference of 0.06 AUC, which is quite acceptable.

- **Color versus gray:** Table 6.18 gives an overview about the presented features in this chapter. Only the best findings for each feature are illustrated. It is obvious that the color features are more discriminating in feature space than the gray level features LBP, GLCM6 and GLCM16 from [AWP⁺09].

For example the best color feature and the best grayscale feature have a difference of 0.083 in their AUC values. Thus, color plays an important role in classification of polyps.

| Feature | color | gray | AUC | |
|-----------------------------|-------|------|--------|---------|
| | | | LibSVM | k -NN |
| ColorGLCM | × | | | 0.843 |
| L_8^{subset} | × | | 0.835 | |
| ColorLBP | × | | 0.834 | |
| OC-GLCM | × | | 0.832 | |
| Color Wavelet | × | | 0.82 | |
| OC-LBP | × | | 0.818 | |
| WaveletDecomp | × | | 0.799 | |
| $L_{16}^{\text{subset,ri}}$ | × | | 0.799 | |
| L_8^{ri} | × | | 0.792 | |
| LBP | | × | 0.76 | |
| GLCM6 | | × | 0.74 | |
| GLCM16 | | × | 0.735 | |

Table 6.18: Overview over the presented features and their highest AUC

- **Color spaces:** Different color spaces are examined in the Color Wavelet Feature. RGB, K-L, HSV and CieLab color spaces are tested. The finding was that RGB color space holds the best color representation when wavelet transform is applied afterwards.
- **Reduction of color channels:** In the Color Wavelet, OC-GLCM and the ColorLBP feature, which is in each case one representative of the each group, a reduction of the number of color channels or color channel combinations was tested, resulting in a less dimensional feature vector and less computational complexity. It turns out that in each case it is the best choice to utilize all color channels. No clear superiority of a single color channel could be examined.
- **GLCM versus LBP:** Each feature implemented in this work incorporates the usage of either GLCM or LBP. Both are operating with similar performance. For instance, there is only a marginal difference between the AUC values from ColorGLCM, ColorLBP and the OC-GLCM feature. The features ColorGLCM and ColorLBP, which are the application of the GLCM and LBP on each color channel, perform surprisingly well reaching the best and the third best AUC values of all tests.

The main difference between these two texture features is the computational complexity. The computation of the GLCM is time-consuming in comparison to the Local Binary Pattern.

- **GLCM features:** ColorGLCM, OC-GLCM, Color Wavelet, WaveletDecomp, GLCM6, GLCM16 are the features involving Co-occurrence Matrices, listed in decreasing order of their AUC results. Those features have the widest range in classification results, providing the best and the worst feature.

A clear structure can be examined among the listed features. The best ones incorporate color into the feature estimation (ColorGLCM, OC-GLCM). The Wavelet features (Color Wavelet, WaveletDecomp) perform averagely in this field, even though color information is used in addition. The worst among the features are the gray level features GLCM6 and GLCM16.

- **LBP features:** L_8^{subset} , ColorLBP, OC-LBP, $L_{16}^{\text{subset,ri}}$, L_8^{ri} and LBP are the features involving computation of Local Binary Patterns. The subset version of LBP performs slightly better than the application of the LBP to each color channel (ColorLBP) or OC-LBP, due to the fact that only the most frequent LBP values are used for the computation of their frequencies. The rotation invariant versions of the LBP ($L_{16}^{\text{subset,ri}}$, L_8^{ri}) perform worse than their rotation variant relatives, but still better than the gray level LBP. The bigger neighborhood applied in $L_{16}^{\text{subset,ri}}$ did not contribute to a better feature space discrimination.
- **Discrete Wavelet Transform versus Non Numeric Transform:** The numeric transform represented by the Discrete Wavelet Transform applied in the Color Wavelet and Wavelet Decomposition feature does not contribute to a better polyp detection. Comparing the ColorGLCM feature with Color Wavelet feature, which is the same feature just without the numeric transform, shows the superiority of ColorGLCM. Furthermore computational advantages are apparent. The Wavelet Transform is variant with regard to translation, which is probably the main reason for this result.
- **Variiances and Covariances:** Karkanis et al. [KIM⁺03] presented an approach for computing variiances and covariances from the different color channels of the statistical measurements in the wavelet domain. In this work it is pointed out that estimating variiances and covariances from the measures is harmful for classification results as shown in the Color Wavelet and the Color Wavelet Covariance feature.

- **Patch sizes:** The best patch size clearly depends on the selected feature. Generally, the wavelet features require patch sizes with value power of two. They should be higher than the patch sizes of other features, due to the wavelet decomposition levels.

The patch size regulates the amount of *non-polyp* information when regarding patches that contain both domains (e.g. the polyp border). It can be significant, how much *non-polyp* information occurs in a patch that is actually classified as *polyp* patch, leading to a reduction of discriminance in feature space.

Chapter 7

Summary

7.1 Possible Improvements

A comprehensive polyp detection system should incorporate more than only a feature extraction module and a classification. The whole system must be customized to the task of polyp detection. Some proposals are listed in the following section.

- **Scale invariant features:** During colonoscopy an additional parameter describing the distance to the intestinal wall could be stored for each frame. This could be helpful for extraction of GLCM or LBP contributing to evaluation of the resolution level of the texture. The images can then be normalized and hence a scale invariance of features could be achieved.
- **Overcomplete Wavelet Transform:** In order to compute translation invariant features the overcomplete version of the Wavelet Transform (OCWT) [Bra03] could be applied. This version overcomes the main problem of DWT.
- **Disjunction of test and training set:** So far, Images are subdivided into patches and then features are extracted from each patch. The separation of the obtained feature vectors in training and test set is done via cross-validation afterwards. It is advisable to group the complete images first into training and test set and then extract features from the patches. Thus, it is possible to apply patch overlapping without the loss of disjunction. It is also possible to check which image belongs to which set in a straightforward way. An increased transparency of the training and testing step could be achieved thereby.
- **Preprocessing:** After image acquisition, an adequate preprocessing could be applied to each frame, removing artifacts from endoscopic images such as shifted RGB color channels or glossy spots.

- **Validation:** Let's assume that a polyp is detected on one patch and thus the location of the polyp is approximately known. It would be helpful to validate that result on the current as well as on the next frames. The position of the polyp could be further analyzed by applying feature extraction with patch overlapping or smaller patch sizes. A fixed threshold could be set indicating how many patches must be classified as polyps to get reliable results.

7.2 Summary of this Work

The detection of polyps in endoscopic images is a challenging task. The published methods in literature introduced in chapter 2 are hardly comparable. They work on different and very small data sets, often not considering different types of polyps. In this work, a very heterogenous set of images is chosen, containing frames from different scenes and different polyp types.

Several investigation are made to increase the feature extraction module of [AWP⁺09]. The overall aim was to include color in the feature extraction techniques, which is one of the most promising information base for polyp detection.

Primarily, different texture descriptors are combined to new features, incorporating Wavelet Transform, GLCMs and LBPs. The applied experiments produced the following results:

- Including color led to a significantly higher detection rate (+0.10 AUC for GLCM features). The single color methods performed equally well for the chosen data set. Only 0.05 AUC range lie between the best and the worst color method.
- The combination of all color channels of the RGB color space led to the best results.
- The Discrete Wavelet Transform does not have the expected positive impact on polyp detection.
- The Local Binary Pattern and the GLCM and their implemented variants perform equally well.
- The Support Vector Machine classifier holds superior results in comparison to k -NN, considering the number of higher classification results.

Bibliography

- [ACN07] ALEXANDRE, Luís A. ; CASTELEIRO, João. ; NOBRE, Nuno: Polyp Detection in Endoscopic Video Using SVMs. In: *11th European Conference on Principles and Practice of Knowledge Discovery in Databases*, Springer, 2007, 358-365
- [AKA91] AHA, David W. ; KIBLER, Dennis ; ALBERT, Marc K.: Instance-Based Learning Algorithms. In: *Machine Learning*, 1991, S. 37-66
- [AWP⁺09] AMELING, Stefan ; WIRTH, Stephan ; PAULUS, Dietrich ; LACEY, G. ; VILARINO, F.: Texture-based Polyp Detecion in Colonoscopy. In: *Bildverarbeitung für die Medizin 2009 - Algorithmen, Systeme, Anwendungen*, Springer Berlin Heidelberg New York, 2009, S. 346-350
- [Bra03] BRADLEY, Andrew P.: Shift-Invariance in the Discrete Wavelet Transform. In: *Digital Image Computing: Techniques and Applications*, 2003
- [CL01] CHANG, Chih-Chung ; LIN, Chih-Jen: *LIBSVM: a library for support vector machines*, 2001. – Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [CR95] COHEN, A. ; RYAN, R.D. ; COHEN, A. (Hrsg.) ; RYAN, R.D. (Hrsg.): *Wavelets and Multiscale Signal Processing*. London : Champman & Hall, 1995
- [CTH84] CONNERS, R.W. ; TRIVEDI, M.M. ; HARLOW, C.A.: Segmentation of a high-resolution urban scene using texture operators. In: *Computer vision, graphics, and image processing* 25 (1984), Nr. 3, S. 273-310
- [CV95] CORTES, Corinna ; VAPNIK, Vladimir: Support-Vector Networks. In: *Machine Learning* 20 (1995), Nr. 3, 273-297. citeseer.ist.psu.edu/cortes95supportvector.html
- [Dau92] DAUBECHIES, Ingrid: *Ten Lectures on Wavelets*. Society for Industrial & Applied Mathematics, U.S., 1992

- [DHHM06] DHANDRA, B. V. ; HEGADI, Ravindra ; HANGARGE, Mallikarjun ; MALEMATH, V. S.: Analysis of Abnormality in Endoscopic images using Combined HSI Color Space and Watershed Segmentation. In: *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)-Volume 04* (2006), S. 695–698
- [GJ66] GREEN, D.M. ; J.M., Swets: *Signal detection theory and psychophysics*. Peninsula Publishing, 1966
- [HCL08] HSU, C.W. ; CHANG, C.C. ; LIN, C.J.: A practical guide to support vector classification / National Taiwan University. 2008. – Forschungsbericht
- [HDS73] HARALICK, R. M. ; DINSTEIN, I. ; SHANMUGAM, K.: Textural Features for Image Classification. In: *IEEE Trans. Systems, Man, and Cybernetics* Bd. 3(6):610-621, 1973
- [HOT⁺07] HWANG, Sae ; OH, JungHwan ; TAVANAPONG, Wallapak ; WONG, Johnny ; GROEN, Piet C.: Polyp Detection in Colonoscopy Video using Elliptical Shape Feature. In: *International Conference on Image Processing*, 2007, 465-468
- [IMK06] IAKOVIDIS, D.K. ; MAROULIS, D.E. ; KARKANIS, S. A.: An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy. In: *Computers in Biology and Medicine* 36 (2006), Nr. 10, S. 1084–1103
- [ISNC05] IBANEZ, Luis ; SCHROEDER, Will ; NG, Lydia ; CATES, Josh: *The ITK Software Guide*. Second. Kitware Inc., 2005. – Updated for ITK version 2.4
- [Jul86] JULESZ, B.: Texton gradients: the texton theory revisited. In: *Biological Cybernetics*, 1986, S. 245–251
- [KD03] KANG, J. ; DORAISWAMI, R.: Real-time image processing system for endoscopic applications. In: *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on* 3 (2003), S. 1469–1472
- [KIKM01] KARKANIS, S. A. ; IAKOVIDIS, D. K. ; KARRAS, D. A. ; MAROULIS, D. E.: Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures. In: *IEEE International Conference on Image Processing* (2001), S. 833–836

- [KIM⁺00] KARKANIS, S.A. ; IAKOVIDIS, D.K. ; MAROULIS, D.E. ; THEOFANOUS, N.G. ; MAGOULAS, G.D.: Tumor recognition in endoscopic video images using artificial neural network architectures. In: *Proceedings of the 26th Euromicro Conference 2*, 2000
- [KIM⁺03] KARKANIS, S. A. ; IAKOVIDIS, Dimitrios K. ; MAROULIS, Dimitrios E. ; KARRAS, Dimitris A. ; TZIVRAS, M.: Computer-aided tumor detection in endoscopic video using color wavelet features. In: *IEEE Transactions on Information Technology in Biomedicine* 7 (2003), Nr. 3, 141-152. <http://citeseer.ist.psu.edu/718632.html>
- [KMGS99] KARKANIS, S. A. ; MAGOULAS, G. D. ; GRIGORIADOU, M. ; SCHURR, M.: Detecting abnormalities in colonoscopic images by textural description and neural networks. In: *Proc. of Work. on Mach. Learn. in Med. Appl., Advance Course in Artif. Intell.-ACAI99* (1999), S. 59–62
- [KWL⁺00] KRISHNAN, S.M. ; WANG, P. ; LIN, Z. ; VIKRAM, N. ; XUE, Z.: Development of texture analysis method for medical endoscopic colour images. In: *Proceedings of IEEE-EMBS Asia-Pacific Conference on Biomedical Engineering*, 2000
- [KYC⁺98] KRISHNAN, S. M. ; YANG, X. ; CHAN, K. L. ; KUMAR, S. ; GOH, P. M. Y.: Intestinal abnormality detection from endoscopic images. In: *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE Bd. 2*, 1998, S. 895–898
- [LCC97] LU, C. ; CHUNG, P. ; CHEN, C.: Unsupervised Texture Segmentation via Wavelet Transform. In: *Pattern Recognition*, 1997, S. 729–742
- [LCK05] LI, P. ; CHAN, KL ; KRISHNAN, SM: Learning a multi-size patch-based hybrid kernel machine ensemble for abnormal region detection in colonoscopic images. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on 2* (2005)
- [LSS⁺93] LERSKI, R. ; STRAUGHAN, K. ; SHAD, L. ; BOYCE, D. ; BLUML, S. ; ZUNA, I.: MR Image Texture Analysis - An Approach to Tissue Characterisation. In: *Magnetic Resonance Imaging Bd. 11*, 1993, S. 873–887
- [Mä03] MÄENPÄÄ, T.: *The local binary pattern approach to texture analysis - extensions and applications*, University of Oulu, Diss., 2003

- [Mal89] MALLAT, Stephane: A theory of multiresolution signal decomposition: the wavelet representation. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11 (1989), S. 674–693
- [Mey93] MEYER, Y.: *Wavelets: Algorithms and Applications*. Philadelphia : Society for Industrial and Applied Mathematic, 1993
- [MIKK03] MAROULIS, D. E. ; IAKOVIDIS, D. K. ; KARKANIS, S. A. ; KARRAS, D. A.: CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames. In: *Computer Methods and Programs in Biomedicine* 70 (2003), Nr. 2, S. 151–166
- [MOPS00] MÄENPÄÄ, T. ; OJALA, T. ; PIETIKÄINEN, M. ; SORIANO, M.: Robust Texture Classification By Subsets Of Local Binary Patterns. In: *Proc. 15th International Conference on Pattern Recognition*, 2000
- [Nie07] NIEMANN, Heinrich: *Klassifikation von Mustern, 2. überarbeitete und erweiterte Auflage im Internet*. 2007 <http://www5.informatik.uni-erlangen.de/Personen/niemann/klassifikation-von-mustern/m00links.html>
- [OB96] ODEGARD, J.E. ; BURRUS, C.S.: Smooth biorthogonal wavelets for applications in image compression. In: *Proceedings of DSP Workshop*, 1996
- [OPH96] OJALA, T. ; PIETIKÄINEN, M. ; HARWOOD, D.: A comparative study of texture measures with classification based on feature distributions. In: *Pattern Recognition* Bd. 29, 1996, S. 51–59
- [OPM00] OJALA, T. ; PIETIKÄINEN, M. ; MÄENPÄÄ, T.: Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns. In: *Proc. ECCV 2000*, 2000
- [POX00] PIETIKÄINEN, M. ; OJALA, T. ; XU, Z.: Rotation-Invariant Texture Classification Using Feature Distributions. In: *Pattern Recognition* Bd. 33, 2000, S. 43–52
- [QT] *Qt Cross-Platform Application Framework*. – Software available at <http://trolltech.com/products/qt>
- [QWT] *Qwt - Qt Widgets for Technical Applications*. – Software available at <http://qwt.sourceforge.net>
- [RK82] ROSENFELD, A. ; KAK, A.: *Digital Picture Processing*. New York : Academic Press, 1982

- [RL93] RAO, A. ; LOHSE, G.L.: Identifying high level features of texture perception. In: *CVGIP. Graphical models and image processing*, Academic Press, <http://www.elsevier.com>, 1993, S. 262–271
- [SC80] SNYDER, W.E. ; COWART, A.E.: An Iterative Approach to Region Growing. In: *ICPR80*, 1980
- [SPRS⁺08] *Kapitel S3-Leitlinie ‘Kolorektales Karzinom’*. In: SCHMIEGEL, W. ; POX, C. ; REINACHER-SCHICK, A. ; ADLER, G. ; FLEIG, W. ; FÖLSCH, U.R. ; FRÜHMORGEN, P. ; GRAEVEN, U.: *Gastroenterol.* Stuttgart, New York : Georg Thieme Verlag KG, 2008, 1-73
- [TA07] THOMSON, Alan B. ; AHNEN, Dennis J.: Intestinal Polypoid Adenomas / eMedicine, The Continually Updated Clinical Reference. Version: 2007. <http://emedicine.medscape.com/article/179284-overview>. 2007. – Forschungsbericht
- [TK03] TJOA, Marta ; KRISHNAN, Shankar: Feature extraction for the analysis of colon status from the endoscopic images. In: *BioMedical Engineering OnLine* 2 (2003), 4, Nr. 1, 9. <http://www.biomedical-engineering-online.com/content/2/1/9>. ISBN 1475–925X
- [TKK⁺01] TJOA, M.P. ; KRISHNAN, S.M. ; KUGEAN, C. ; WANG, P. ; DORAIWAMI, R.: Segmentation of clinical endoscopic image based on homogeneity and hue. In: *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 3, 2001, S. 2665–2668
- [VS91] VINCENT, Luc ; SOILLE, Pierre: Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (1991), Nr. 6
- [Wav] *Wavelet and Image Class Library*. – Software available at <http://sourceforge.net/projects/wavelet/>, Reference at <http://herbert.the-little-red-haired-girl.org/en/software/wavelet/docs/Wave.pdf>
- [WEK] *WEKA Classification Algorithms*. – Software available at <http://sourceforge.net/projects/wekaalgorithms/>

- [WHO06] WHO: Fact sheet no 297: cancer / WHO Media Centre. Version:2006. <http://www.who.int/mediacentre/factsheets/fs297/en/print.html>. 2006. – Fact Sheet
- [WKHS02] WANG, P. ; KRISHNAN, S.M. ; HUANG, Y. ; SRINIVASAN, N.: An adaptive segmentation technique for clinical endoscopic image processing. In: *Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering 2, 2002*, S. 1084–1085
- [WKKT01] WANG, P. ; KRISHNAN, S. M. ; KUGEAN, C. ; TJOA, M. P.: Classification of endoscopic images based on texture and neural network. In: *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE* Bd. 4, 2001, S. 3691–3695
- [WP01] WEISS, Gary M. ; PROVOST, Foster: The Effect of Class Distribution on Classifier Learning. In: *Technical Report ML-TR-43, Department of Computer Science, Rutgers University* (2001), 1
- [ZK01] ZHENG, M.M. ; KRISHNAN, S.M.: Decision support by fusion in endoscopic diagnosis. In: *The Seventh Australian and New Zealand Intelligent Information Systems Conference, 2001*, S. 107–110
- [ZKT05] ZHENG, M. M. ; KRISHNAN, S. M. ; TJOA, M. P.: A fusion-based clinical decision support for disease diagnosis from endoscopic images. In: *Computers in Biology and Medicine* 35 (2005), Nr. 3, S. 259–274