

# Evaluation of Different Distortion Correction Methods and Interpolation Techniques for an Automated Classification of Celiac Disease

M. Gadermayr<sup>a,\*</sup>, M. Liedlgruber<sup>a</sup>, A. Uhl<sup>a</sup>, A. Vécsei<sup>b</sup>

<sup>a</sup>*Department of Computer Sciences  
University of Salzburg, Austria*

<sup>b</sup>*St. Anna Children's Hospital, Department of Pediatrics  
Medical University, Vienna, Austria*

---

## Abstract

Due to the optics used in endoscopes, a typical degradation observed in endoscopic images are barrel-type distortions. In this work we investigate the impact of methods used to correct such distortions in images on the classification accuracy in the context of automated celiac disease classification.

For this purpose we compare various different distortion correction methods and apply them to endoscopic images, which are subsequently classified. Since the interpolation used in such methods is also assumed to have an influence on the resulting classification accuracies, we also investigate different interpolation methods and their impact on the classification performance. In order to be able to make solid statements about the benefit of distortion correction we use various different feature extraction methods used to obtain features for the classification.

Our experiments show that it is not possible to make a clear statement about the usefulness of distortion correction methods in the context of an automated diagnosis of celiac disease. This is mainly due to the fact that an eventual benefit of distortion correction highly depends on the feature extraction method used for the classification.

*Keywords:* Endoscopy, Celiac disease, barrel-type distortion, distortion correction, medical image classification

---

## 1. Introduction

Today, medical endoscopy is a widely used procedure to inspect the inner cavities of the human body. As a consequence different medical fields exist for

---

\*Corresponding author

*Email addresses:* mgadermayr@cosy.sbg.ac.at (M. Gadermayr),  
mliedl@cosy.sbg.ac.at (M. Liedlgruber), uhl@cosy.sbg.ac.at (A. Uhl)

which automated decision-support systems based on endoscopic images have been developed [1]. But since images taken with endoscopes often suffer from various kinds of degradations, a pre-processing of the images is often necessary in order to make them useful for an automated analysis [2] (to cope with e.g. sensor noise, focus and motion blur, and specular reflections [1]).

A different type of degradation, present in all endoscopic images, is a barrel-type distortion. This type of degradation is caused by the wide-angle (fish eye) nature of the optics used in endoscopes (although the strength of the distortion varies depending on the endoscope used). Such a distortion is also claimed to affect diagnosis since it introduces non-linear changes in the image, due to which the outer areas of the image appear significantly smaller than they actually are [3]. As a consequence the estimation of area or perimeter of observed lesions can be significantly incorrect depending on the position in the image [4, 5]. With respect to an automated classification such distortions are also suspected to lead to corrupted features due to an inhomogeneous magnification [6]. It has also been mentioned in literature that barrel-type distortions might lead to “complications using token matching techniques for pattern recognition” [4]. Since the seminal work on distortion correction (DC) for endoscopic images [6] several distortion correction procedures have been developed to overcome the problems caused by such distortions [4, 5, 7, 8].

To the best of our knowledge up to now there are only a few studies available, which investigate the impact of barrel-type distortions and distortion correction on the accuracy of automated classification systems for endoscopic images [9, 10, 11, 12].

In [9] the impact of distortion correction on the classification accuracy regarding celiac disease images has been investigated. Gschwandtner et al. showed that most feature extraction methods evaluated failed to take advantage of applying distortion correction as a pre-processing step to the endoscopic images, resulting in a decreased classification accuracy. In the follow-up work in [11] the authors evaluated another distortion correction method with a similar outcome. In [12] the authors evaluated the distortion correction methods from [9, 11] with additional feature extraction methods. In this work the authors showed that only a few feature extraction methods are able to take a slight advantage of distortion correction (with respect to the overall classification rates). The work presented in [10] investigated the impact of distortions and interpolation artifacts caused by distortion correction methods on the accuracy of a classification of celiac disease images. The clear outcome was that there is indeed a negative impact of barrel-type distortions on the classification accuracy. Moreover, this negative effect gets more apparent the farther away from the center of distortion (CoD) features are extracted. But it has also been shown that this also accounts to distortion corrected images due to interpolation artifacts. The bottom-line result of all four studies was that the interpolation artifacts caused by the distortion correction are very likely one reason why the classification results have not been improved.

To facilitate a more thorough investigation of the effect of distortion correction on the accuracy of an automated diagnosis of celiac disease, the present

work compares four different distortion correction methods (instead of two methods at most, like in earlier studies). In addition, this work is the first study which also compares various different interpolation methods used for distortion correction (previous studies were based on bilinear interpolation only). To be able to make solid statements about the usability of the different distortion correction methods in the context of endoscopic image classification, we use various different feature extraction methods, which are evaluated on distortion corrected images.

The remaining part of this work is organized as follows: In Section 2 we briefly describe the medical background of celiac disease, the staging system commonly used to diagnose this disorder, and the motivation behind automated classification systems targeted at this disorder. After a discussion of barrel-type distortions and the problems inherent to this type of degradations in Section 3, we describe the methods for distortion correction evaluated in this work in Section 4. In Section 5 we then briefly summarize the feature extraction methods used throughout our experiments. Experimental results and configuration details for our experiments are presented in Section 6, followed by a discussion in Section 7 and concluding remarks in Section 8.

## 2. Automated Classification for Celiac Disease Diagnosis

Celiac disease, commonly known as gluten intolerance, is a complex autoimmune disorder that affects the small bowel in genetically predisposed individuals of all age groups after introduction of food containing gluten. Characteristic for the disease is an inflammatory reaction in the mucosa of the small intestine. During the course of the disease the mucosa loses its absorptive villi and hyperplasia of the enteric crypts occurs, leading to a diminished ability to absorb nutrients.

Endoscopy with biopsy is currently considered the gold standard for the diagnosis of celiac disease. During standard upper endoscopy at least four duodenal biopsies are taken. Microscopic changes within these specimen are then classified in a histological analysis according to the Marsh classification proposed in 1992 [13]. Subsequently, Oberhuber et al. proposed the modified Marsh classification [14] which distinguishes between classes Marsh-0 to Marsh-3, with subclasses Marsh-3a, Marsh-3b, and Marsh-3c, resulting in a total number of six classes. According to the modified Marsh classification Marsh-0 denotes a healthy mucosa (without visible changes of the villous structure) and Marsh-3c designates a complete absence of villi (villous atrophy).

In accordance to the work in [9, 11, 10], we consider the four classes Marsh-0 and Marsh-3a to Marsh-3c only throughout this work, since visible changes in the villi structure can be observed only for classes Marsh-3a to Marsh-3c (in case of the classes Marsh-1 the number of intra-epithelial lymphocytes is increased and in case of Marsh-2 the crypts of Lieberkuhn are proliferated). In addition, in this work we focus on the 2-classes case only (i.e. Marsh-0 and Marsh-3) since in this case the image database available is fairly well balanced with respect to

the images in each class. In the remaining part of this work we denote these classes by “No-celiac” and “Celiac”.

Throughout the past decades automated classification systems got an emerging field of research for endoscopic diagnosis and treatments [15]. An automated system identifying areas affected by celiac disease in the duodenum would offer the following benefits (among other):

- Methods helping to locate specific areas for biopsies might improve the reliability of celiac disease diagnosis. Especially when considering that biopsies are invasive and thus the number of biopsies taken should kept small, assisted biopsy site targeting is desirable. Such a targeting can be supported by an automated system for the detection of areas affected by celiac disease.
- The whole diagnostic work-up of celiac disease, including duodenoscopy with biopsies, is time-consuming and cost-intensive. To save costs, time, and manpower and simultaneously increase the safety of the procedure it would be desirable to develop a less invasive approach avoiding biopsies. Recent studies [16] investigating such endoscopic techniques report reliable results. These could be further improved by analysis of the acquired visual data (digital images and video sequences) with the assistance of computers.
- The (human) interpretation of the video material captured during capsule endoscopy [17] is an extremely time consuming process. Automated identification of suspicious areas in the video would significantly enhance the applicability and reduce the costs of this technique for the diagnosis of celiac disease.

### 3. Barrel-type Distortions

As already mentioned earlier, barrel-type distortions are typical to endoscopic images. The reason for this is the optics used in endoscope devices. In endoscopy it is usually desirable to have a wide field-of-view (FOV) in order to be able to inspect as much as possible without the necessity of continuously adjusting the direction of the tip of the endoscope. But the enlarged visible area comes at the price of noticeable distortions, especially at very wide FOV values as commonly used in endoscopes (typical values range from  $100^\circ$  to  $170^\circ$ ).

In addition, the level of distortion is also dependent on the distance of the endoscope to the tissue in focus. Since endoscopes usually have a very short focal depth (a few millimeters only), barrel-type distortions get even more amplified. While moving the endoscope farther away from the tissue of interest would theoretically help to reduce the barrel-type distortions, the resulting images most likely will be out of focus since endoscopes usually also have a very narrow depth of field (i.e. the range before and after the focal plane inside which objects appear sharp).

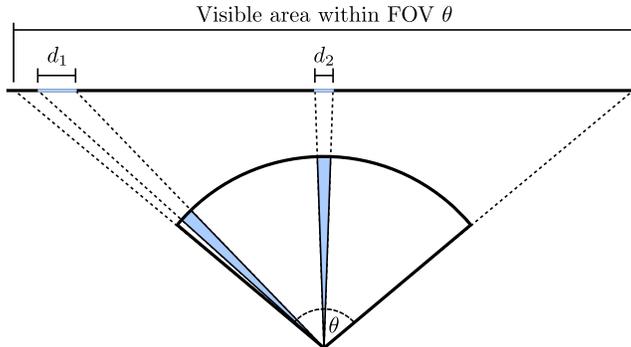


Figure 1: Illustration of the differently large object areas  $d_1$  and  $d_2$  imaged within two angular slices of equal angular width (denoted by the blue slices).

Hence, depending on the FOV and the focal depth, barrel-type distortions may vary significantly in strength between the CoD (which corresponds to the optical axis) and the outer regions of an image (i.e. the image magnification decreases with the distance to the CoD).

Figure 1 shows how the distance to the CoD affects the distortion. While the object area imaged by the angular slice pointing toward the CoD is narrow, we notice that the angular slice farther away from the CoD covers a much wider object area. In other words, the object magnification is much higher in the proximity of the CoD as compared to outer image areas.

An example for barrel-type distortions is shown in Fig. 2. This image shows a checkerboard pattern captured with an endoscope. We immediately notice that lines, which are usually straight in a checkerboard pattern, are not straight anymore due to severe distortions. We also notice that – as already indicated above – the tiles farther away from the image center are considerably smaller as compared to the tiles close to the center of the image.

#### 4. Correction of Barrel-type Distortions

The correction of barrel-type distortions consists of two integral parts, which influence the appearance of the distortion-corrected images: the distortion correction method and the interpolation of missing information, once the distortion has been estimated correctly.

##### 4.1. Distortion Correction

Barrel-type distortions are radial distortions. That is, by neglecting possible small non-radial defects in the optical system, the lenses in question exhibit rotational symmetry. As a consequence distortion correction usually consists of two steps: first, based on one or more distorted images, the camera parameters (intrinsic and extrinsic, including the CoD) and the parameters for the assumed radial distortion model are estimated. Then, according to the parameters found,



Figure 2: An example for barrel-type distortions showing a checkerboard pattern captured with an endoscope.

the distortion is corrected by applying a pincushion distortion, which is the inverse to the barrel-type distortion and therefore exhibits rotational symmetry too.

Due to the rotational symmetry of the pincushion distortion the correction simply corresponds to shifting pixels farther away from the CoD. The direction of this shift follows the direction from the CoD to the original pixel position. However, the strength of the shift varies, depending on the distance of the original pixel position to the CoD (see Fig. 3).

Since the pixels in an undistorted image lie on a regular grid, we need to compute the distorted coordinates and use them to look up pixel values for the undistorted image in the distorted image. Hence, assuming that the CoD  $\hat{\mathbf{x}}_{\mathbf{c}} = (x_c, y_c)^T$  has been estimated, the computation of the coordinates for each distorted point  $\mathbf{x}' = (x', y')^T$  inside a distorted image can be written in its most general form as

$$\mathbf{x}' = \hat{\mathbf{x}}_{\mathbf{c}} + \frac{(\mathbf{x} - \hat{\mathbf{x}}_{\mathbf{c}})}{\|\mathbf{x} - \hat{\mathbf{x}}_{\mathbf{c}}\|_2} \cdot f(r), \quad (1)$$

where  $\mathbf{x} = (x, y)^T$  denotes the respective undistorted point (i.e. the pixel coordinates in the undistorted image),  $f(\cdot)$  denotes a distortion function which returns the strength of the distortion depending on the points  $\mathbf{x}$  and  $\hat{\mathbf{x}}_{\mathbf{c}}$ , and  $r$  denotes a parameter defining the strength of the distortion (e.g. the distance of an undistorted pixel to the CoD).

In the past various different methods to eliminate barrel-type distortions in images have been proposed. In the following we briefly describe each of the methods we used for our experiments. The set of methods described below has been chosen since in [18] it has been shown that the underlying models of these methods are able to cope well with barrel-type distortions in endoscopic imagery.

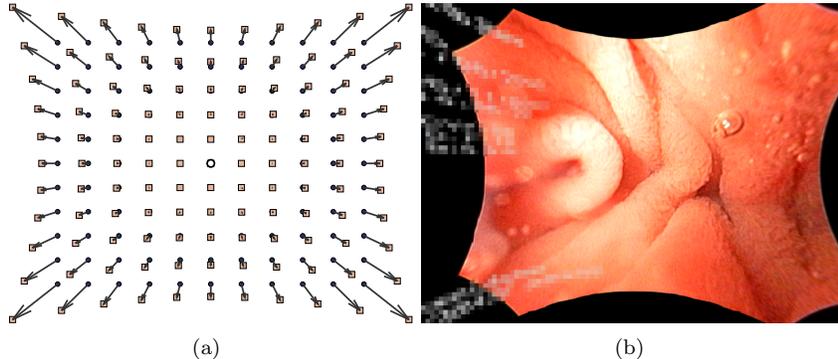


Figure 3: (a) Illustration of the shifts applied to distorted pixels (denoted by blue dots) to obtain corrected pixel positions (denoted by squares) during distortion correction (the CoD is denoted by the white circle in the image center). (b) An example of an endoscopic image after applying distortion correction.

#### 4.1.1. Even-order Polynomial Model

A very popular camera calibration method has been introduced by Zhang [19]. In this work the camera parameters are estimated using a homography estimation. It is common practice to model the lens distortion function by a Taylor expansion of the form

$$L(r) = \alpha_0 + \alpha_1 \cdot r + \alpha_2 \cdot r^2 + \alpha_3 \cdot r^3 + \dots \quad (2)$$

In [19],  $r = \|\bar{\mathbf{x}}\|_2$  with  $\bar{\mathbf{x}}$  being the normalized image coordinates of  $\mathbf{x}$ . Zhang also limits this expansion to second-order and fourth-order terms, yielding the even order polynomial:

$$f(r) = k_1 \cdot r^2 + k_2 \cdot r^4, \quad (3)$$

where  $k_i$  denote the parameters of this model (i.e.  $k_1 = \alpha_2$  and  $k_2 = \alpha_4$ ). It must be noted that, while the original work in [19] is based on Equ. (3), the implementation used throughout our experiments also uses an additional 6-th order term<sup>1</sup>, which results in

$$f(r) = k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6. \quad (4)$$

#### 4.1.2. Polynomial Model

In the work, proposed by Alvarez et al. [20], the lens distortion function is modeled in a way similar to Zhang. But the authors do not restrict the polynomial to terms of an even-order degree. Their lens distortion function is modeled as a polynomial of order four:

$$f(r) = k_0 + k_1 \cdot r + k_2 \cdot r^2 + k_3 \cdot r^3 + k_4 \cdot r^4, \quad (5)$$

<sup>1</sup>In this work we used the implementation provided by the OpenCV library, which can be downloaded from <http://sourceforge.net/projects/opencvlibrary/>

with  $r = \|\mathbf{x} - \mathbf{x}_c\|_2$  being the distance between the CoD and an undistorted point  $\mathbf{x}$ . The implementation used throughout our experiments can be downloaded freely<sup>2</sup>.

#### 4.1.3. Parameter-free Model

In [21] a parameter-free distortion correction method is proposed. The authors motivate this by the fact that the model based on the Taylor expansion does not work well for strong distortions. This especially accounts to distortions caused by fish-eye or wide-angle lenses since the approximate infinite distortions in the periphery of these images can not be modeled well by a polynomial.

As a consequence the authors propose a parameter-free model which uses a high number of point correspondences between distorted points  $\mathbf{x}'$  and undistorted points  $\mathbf{x}$ . Based on these correspondences a scatter plot is generated, which plots the radial distance between  $\mathbf{x}'$  and  $\mathbf{x}$  (i.e.  $\|\mathbf{x}' - \mathbf{x}_c\|_2 - \|\mathbf{x} - \mathbf{x}_c\|_2$ ) against the distance of  $\mathbf{x}'$  to the CoD (i.e.  $\|\mathbf{x}' - \mathbf{x}_c\|_2$ ). Then for each point in the scatter plot the right scale factor is computed for each available  $\mathbf{x}'$ . In case of points  $\mathbf{x}'$ , which are not available, the respective values are interpolated to some precision. The values computed this way are stored in a lookup table  $T$  for the distortion correction process. Using  $T$  we can also compute a second table  $T'$ , to obtain distortion strengths for computing the points  $\mathbf{x}'$  from the undistorted points  $\mathbf{x}$ . The distortion function can then be formulated as

$$f(r) = L(r), \quad (6)$$

with  $r = \|\mathbf{x} - \mathbf{x}_c\|_2$  and  $L(\cdot)$  being a function which uses  $T'$  to find an appropriate distortion strength for the given  $r$ . The implementation used for this method is a custom software, which has already been used in [11, 12].

#### 4.1.4. Division Model

The work proposed in [22] is based on the so-called division model, which has been proposed in [23]. While this model is not an approximation to traditional models, it is a good approximation to radial distortion. In contrast to the polynomial models it only needs one parameter  $\xi$  to be estimated in order to correct radial distortions. In this model the relationship between the radius of a distorted point  $\mathbf{x}'$  (i.e.  $r_d = \|\mathbf{x}' - \mathbf{x}_c\|_2$ ) and the radius of an undistorted point  $\mathbf{x}$  (i.e.  $r_u = \|\mathbf{x} - \mathbf{x}_c\|_2$ ) is formulated by the following simple equation:

$$r_u = \frac{r_d}{1 + \xi \cdot r_d^2}. \quad (7)$$

After some reordering of the terms we end up with the following distortion function

$$f(r) = r_d = \frac{2 \cdot r}{1 + \sqrt{1 - 4 \cdot \xi \cdot r^2}}, \quad (8)$$

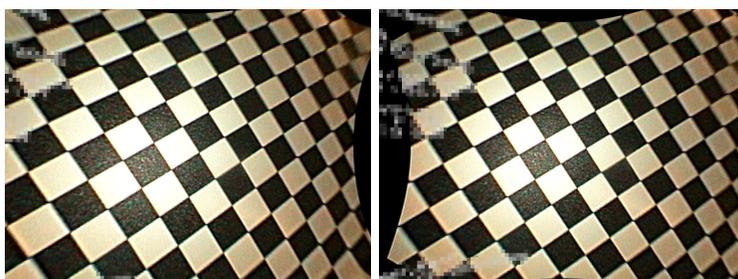
with  $r = r_u$ .

---

<sup>2</sup>The software can be downloaded from <http://www.ipol.im/pub/art/2010/ags-alde/>

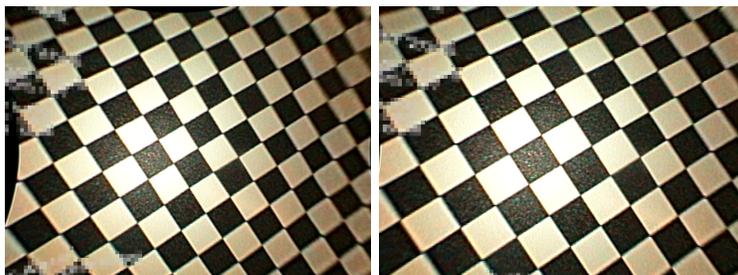


(a)



(b)

(c)



(d)

(e)

Figure 4: Result of the different distortion correction methods: (a) original checkerboard image, (b) DC from [19] (see Section 4.1.1), (c) DC from [20] (see Section 4.1.2), (d) DC from [21] (see Section 4.1.3), and (e) DC from [22] (see Section 4.1.4).

The implementation used throughout our experiments can be downloaded freely<sup>3</sup>.

#### 4.2. Brief Comparison of the Distortion Correction Methods

Figure 4(a) shows an example checkerboard pattern, recorded with an endoscope. In Figs. 4(b)-(e) we show the distortion correction results as produced

---

<sup>3</sup>The software can be downloaded from <http://arthronav.isr.uc.pt/easycamcalib/>

by the different distortion correction methods. From these example images we notice that the distortion correction works roughly equally well in case of each method. Especially in the center region the lines of the checkerboard pattern are rectified accurately. But towards the image borders the results yielded by the methods differ slightly. In these regions we notice that the lines in the undistorted images are not perfectly straight.

### 4.3. Interpolation Methods

During correction of a barrel-type distortion pixel positions, which are on a regular grid in the distorted image, get moved farther away from the CoD. As a consequence, there is information missing between the undistorted points and some sort of interpolation must be performed.

Since the choice of the interpolation method used for distortion correction also affects the texture properties of the correction result, we evaluate different interpolation methods throughout our experiments. While we investigate a small set of interpolation methods only there also exist other methods (for a comprehensive overview see [24]).

The five methods we compare in our experiments are:

- **Nearest-neighbor interpolation**

This interpolation scheme is the simplest. In order to select the value for a specific pixel it just assigns the color of the nearest neighboring pixel (i.e. with the smallest distance in the sense of the Euclidean distance). While being very simple, this method is also prone to introducing artificial high frequency content in an image due to aliasing artifacts. Despite the visually less appealing results, such artifacts may also get problematic when it comes to the extraction of features, which are then used for classification.

- **Bilinear interpolation**

Compared to nearest-neighbor interpolation, this method is a little bit more involved but also results in visually more appealing results without introducing too much artifacts. In our implementation the value for some pixel is computed by taking the weighted average of the pixels in a  $2 \times 2$  neighborhood.

- **Bicubic interpolation**

Another interpolation method used in our experiments is bicubic interpolation. It is similar to bilinear interpolation but uses a  $4 \times 4$  pixel neighborhood for the computation of the weighted average. While being computationally more complex, the results of this interpolation method in general exhibit less interpolation artifacts as compared to bilinear interpolation. This method also performs better when it comes to preserving edges.

- **Lanczos interpolation**

The Lanczos filter [25] of size  $8 \times 8$  is known to imitate the perfect low pass filter (sinc) quite effectively and reduces ringing artifacts. With this filter, the available frequencies are retained, whereas aliasing is avoided.

- **Edge-preserving interpolation**

The Lanczos filter is known to retain the available frequencies. However, in this case due to the stretching of the image in peripheral regions, high frequencies (sharp edges) are missing after undistortion. Consequently, maintaining the available frequencies is not enough to reconstruct the real edge information.

Therefore, an edge-preserving interpolation method by Wang et al. [26] has been implemented. In contrast to usual (linear) interpolation kernels, in this nonlinear approach the behavior of the interpolation depends on the image properties. Near edges high frequencies are encouraged (similar to nearest neighbor interpolation), whereas smooth regions are retained smooth (similar to linear interpolation).

In Fig. 5 we show image regions cut out from an undistorted checkerboard pattern, using the different interpolation methods (undistorted using the method by Melo et al. [22]; the region used for the patches shown in Fig. 5(c)-(g) is indicated by a red rectangle in Fig. 5(b)). As expected, we notice the strongest artifacts in case of the nearest-neighbor interpolation (blocky appearance of the interpolated pixels), while the other methods produce visually more appealing results. When comparing the bilinear and the bicubic interpolation we hardly notice any difference apart from a slightly sharper appearance in case of the latter method (may be attributed to the edge-preserving property in case of this method). The Lanczos interpolation is visually very similar to the bicubic interpolation but produces ringing artifacts. The edge-preserving method, proposed in [26], produces very sharp images. While being visually similar to the nearest-neighbor interpolation, this method produces visually more appealing results.

## 5. Feature Extraction Methods Evaluated

To facilitate an investigation of the impact of different distortion correction methods on the classification accuracy of an automated classification system, we use various different feature extraction methods to obtain features. In addition, using a diverse set of features also allows us to investigate the effect of the different interpolation methods on the resulting classification rates.

The features evaluated can be roughly divided in methods operating in the spatial domain and those operating in the transform domain. In the following we briefly sketch the basic methodology behind each feature extraction method used. For a more detailed explanation we refer the reader to the original works (given as reference).

### 5.1. Spatial Domain Features

#### 5.1.1. Local Binary Patterns (LBP)

Local binary patterns (LBP) [27] are a powerful method to describe local texture properties within an image. In its simplest form, this method compares

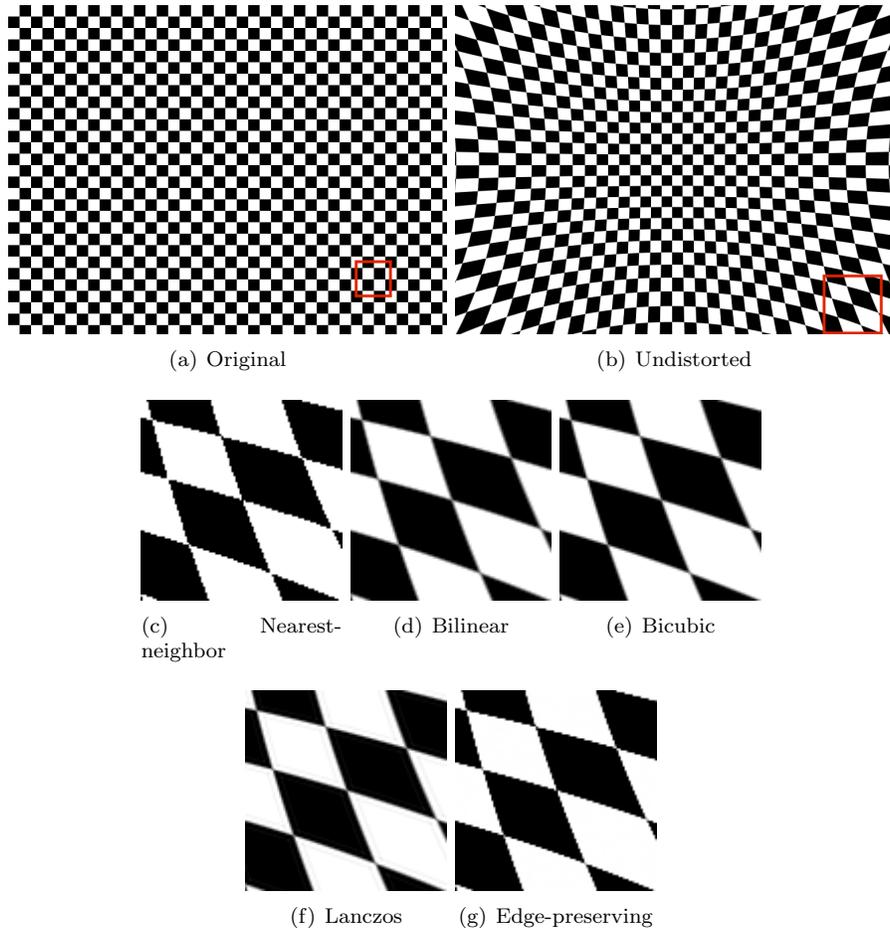


Figure 5: Comparison of the different interpolation methods when used for a distortion correction of a checkerboard pattern: (a) original image, (b) the original image after applying undistortion, (c) nearest-neighbor interpolation, (d) bilinear interpolation, (e) bicubic interpolation, (f) Lanczos interpolation, and (g) edge-preserving interpolation. The red rectangles in (a) and (b) indicate the position of the patches used for (c)-(g).

the grayscale value of a pixel to the values of the eight nearest neighbors. If the value of a neighbor exceeds the center pixel value, the respective neighborhood position is set to one. The number resulting from the neighborhood bit sequence corresponds to the LBP number. In other words, the neighbors of each pixel are thresholded by the respective center pixel and the resulting binary sequence is used to obtain the final LBP number. Based on the LBP numbers computed for all pixels in the source image a histogram is generated, which then serves as the feature vector.

In a subsequent work, Ojala et al. proposed a generalization of this idea [28]. In contrast to the original LBP operator, this new operator allows to

use more than eight neighbors. In addition, these neighbors are no more the nearest neighbors, but circularly distributed around the center pixel. Since the respective neighbor positions do not necessarily correspond to integer pixel positions, interpolation is needed in order to obtain the respective grayscale values. To investigate spatial dependencies beyond the nearest neighbors, this operator also allows to enlarge the radius for the neighbors (in this case the source image is usually blurred prior applying the LBP operator).

Throughout the remaining work we use the abbreviations LBP-NN and LBP-I to distinguish between the nearest-neighbor variant and the subsequently proposed LBP operator, respectively.

#### 5.1.2. Local Ternary Patterns (LTP)

Following the idea of the LBP-NN operator, Tan and Triggs proposed an extension which aims at an enhanced robustness in terms of image noise [29]. To achieve this, the authors introduce a threshold value below which the absolute pixel difference (between the center pixel and a neighbor pixel) yields zero. If the neighbor pixel has a value above (below) the center pixel value plus (minus) the threshold, the LTP operator yields 1 (-1) for this comparison. The resulting pattern is a ternary one from which the histograms are computed.

#### 5.1.3. Gray Level Co-occurrence Matrix (GLCM)

The GLCM, originally proposed in [30], is a 2D-histogram, which describes the spatial relationship between neighboring pixels. The matrix is created based on the co-occurring values of pixels across an image (for some fixed pixel offset). In other words, for each possible combination of two pixel values the GLCM stores the number of co-occurrences within an image for a given displacement between the pixels. While the displacement is fixed for a single GLCM, it can be adjusted with respect to the pixel distance and the direction.

To obtain features for the classification, we compute a GLCM for four different directions (up, down, left, and right) and compute a subset of the statistical features proposed in [30] (i.e. contrast, correlation, energy, and homogeneity) on each GLCM. The final features used are composed by concatenating the Haralick features extracted. In order to obtain features at multiple scales we also carry out experiments with different distance values for the displacement and concatenate the resulting feature vectors.

#### 5.1.4. Shape-Curvature Histogram (SCH)

SCH [31] is a shape feature, originally developed in the context of automated celiac disease diagnosis. The feature is based on the extraction of the geometric property "curvature" from previously detected edges. In order to compute the feature, first an edge map is computed using the Canny-Edge detector. Based on this map, for each edge pixel the orientation is estimated. Then, by investigating the orientation values of neighboring pixels, the curvature is computed. Finally, a histogram, based on the curvature values within the image, is created. This histogram is then used as the feature vector for classification.

#### 5.1.5. Edge Co-occurrence Matrix (ECM)

The principle of ECM has been proposed in [32]. After applying eight differently orientated directional filters (rotated Sobel filters) on the source image, a gradient magnitude image is constructed for each direction. Based on these the direction with the maximum response is determined for each pixel, followed by masking out pixels with a gradient magnitude below some threshold. Then the methodology of GLCM is used to obtain the ECM for one specific displacement. As suggested in [32], we compute the element-wise sum of eight ECMs (one for each direction) to obtain the final ECM for one specific displacement distance. For the subsequent classification the whole ECM is used as the feature vector. In order to obtain features at multiple scales we also carry out experiments with different distance values for the displacement and concatenate the resulting feature vectors.

#### 5.1.6. Spatial Size Distribution (SSD)

The spatial size distribution as an image processing feature has been originally proposed in [33]. The features are computed as an autocorrelation of the difference between the original image and an opened version (morphological opening) of this image. The operator can be configured to meet certain requirements by changing the type and size of the structuring element used for the opening. According to the authors SSD combines properties of texture (fine-grained) as well as shape features (coarse-grained), since different resolutions can be considered by adjusting the size of the structuring element. For simplicity we use circular structuring elements in different sizes only.

#### 5.1.7. Gaussian Markov Random Fields (GMRF)

The features of this method, originally proposed in [34], are based on the so-called Markov parameters. To obtain these features each color channel of an image is subject to a parameter estimation of a GMRF. The neighborhoods employed are of Geman-type [35]. The final feature vector is obtained by concatenating the Markov parameters estimated for each color channel along with the respective approximation errors.

#### 5.1.8. Joint-color LBP (JC-MB-LBP)

Due to the noise-sensitive nature of the LBP operator we developed an operator which is less sensitive to noise [36]. This is achieved by computing LBP numbers from averaged pixel blocks instead of single pixels (actually an extension of the LBP-NN operator). In addition, the resulting features also consider inter-channel relationships between different color channels by constructing joint-histograms across multiple color channels. More precisely, this operator uses two different color channels, resulting in 2D-histograms which are used as features.

#### 5.1.9. Local Color Vector Patterns (LCVP)

The LCVP operator [37] is based on the idea of the JC-MB-LBP operator. But instead of computing the LBP transform for each color channel separately

or computing joint-histograms, this operator treats an image as a color vector field. Based on suitable similarity measures between two color vectors (e.g. length of difference vector between two color vectors in CIELAB color space), a compact histogram descriptor, incorporating all color information available, is computed for an input image. The histograms are used as features for the classification.

#### 5.1.10. Fractal Features (*FRACTAL*)

The method proposed in [38] is based on the computation of fractal features. First an image is pre-filtered using the MR8 filter bank. Subsequently the local fractal dimension for each pixel of the eight filter responses is computed. Then a bag-of-visual-words approach is used to generate a histogram for an image. These histograms are then used as features for the classification.

### 5.2. Transform Domain Features

#### 5.2.1. Pyramidal Wavelet Transform (*WPC*)

This method [39] applies the pyramidal DWT to the input images. Based on the resulting wavelet decomposition structures, we compute the energy or the  $l$ -norm from the coefficients of a subband to obtain the feature for the respective subband. This method selects a subset of subbands for feature extraction (chosen based on the  $l$ -norm computed from the coefficients within a subband). The feature values computed from these subbands make up the feature vector.

When applied to color images, this operator is applied to each color channel separately and the final feature vector for an image is obtained by concatenating the feature vectors from the single channels.

#### 5.2.2. Best-Basis Centroids Base (*WT-BBC*)

The Best Basis Centroids Base method [40] uses the Best-Basis algorithm [41] to find an optimal basis for each image in a training set and computes a centroid over all resulting wavelet packet decomposition structures. After transforming all images into this basis, the most informative subset of the resulting subbands is selected to compute a feature from all coefficients within each subband (energy, variance, or  $l$ -norm). The selection of the subset of subbands is based on a cost-function computed on all coefficients within a subband (log-energy, entropy, or  $l$ -norm). The feature values computed from the most informative subbands form the feature vector.

When applied to color images, this operator is applied to each color channel separately and the final feature vector for an image is obtained by concatenating the feature vectors from the single channels.

#### 5.2.3. Local Discriminant Basis (*WT-LDB*)

This method [39] is based on the Local Discriminant Basis algorithm (LDB), originally proposed in [42]. The LDB algorithm facilitates finding an optimal wavelet packet decomposition basis with respect to discrimination between different classes of images. The LDB algorithm is therefore used to find an optimal

basis into which all images are transformed to. Based on the resulting wavelet decompositions for each subband we use the variance of the coefficients contained within a subband as feature. Instead of using all subbands available for feature extraction we select a subset of subbands which exhibit the highest discriminative power. The feature values computed from these subbands make up the feature vector. Similar to WT-BBC, the operator is applied to each color channel separately and the final feature vector is obtained by concatenating the feature vectors from the single channels.

#### 5.2.4. Wavelet-based GMRF (WT-GMRF)

This method, proposed in [34], is an extension to the GMRF method. While the GMRF method operates in the spatial domain, the WT-GMRF method is based on wavelet-transformed color channels in order to be able to capture texture details at different resolutions. For this purpose each color channel of an input image is transformed using the pyramidal DWT. Then the Markov parameters are estimated for each resulting subband (except for the approximation subband). In order to obtain the final feature vector for classification, the parameter vectors (again containing the approximation errors too) from all subbands and color channels are concatenated. The rather time-consuming feature selection used in [34] is not carried out in this work.

#### 5.2.5. WT-GMRF With Custom Neighborhoods (WT-GMRF-CNH)

This method is an extension to the WT-GMRF method [34]. In order to be able to better capture details from the different wavelet subbands (horizontal, vertical, and diagonal) this method uses neighborhoods specifically tailored to the differently orientated subbands instead of the Geman-type neighborhoods used earlier. Similar to WT-GMRF we carry out no feature selection, as done in [34].

## 6. Experiments

### 6.1. Image Database Used

The image test set used contains images taken during duodenoscopies at the St. Anna Children’s Hospital using pediatric gastroscopes without magnification (two GIF-Q165 and one GIF-N180, Olympus, with resolution  $768 \times 576$  and  $528 \times 522$  pixels, respectively). The main indications for endoscopy were the diagnostic evaluation of dyspeptic symptoms, positive celiac serology, anemia, malabsorption syndromes, inflammatory bowel disease, and gastrointestinal bleeding. The mean age of the patients undergoing endoscopy was 11.1 years (range 0.8-20.9 years). The female to male ratio was 1.43:1. Images were recorded by using the modified immersion technique, which is based on the instillation of water into the duodenal lumen for better visibility of the villi. The tip of the gastroscope is inserted into the water and images of interesting areas are taken. A study [43] shows that the visualization of villi with the immersion technique has a higher positive predictive value. Previous work [44] also found

Image class			
	$\mathbf{N}_O$	$\mathbf{N}_E$	$\mathbf{N}_P$
No-celiac	100	163	59
Celiac	67	124	23
<b>Total</b>	<b>167</b>	<b>287</b>	<b>82</b>

Table 1: The detailed ground truth information for the celiac disease image database used throughout our experiments.

that the modified immersion technique is more suitable for automated classification purposes as compared to the classical image capturing technique. Images from a single patient were recorded during a single endoscopic session and differ by the presented duodenal region only.

There are two duodenal regions with completely different geometric properties, i.e. the Duodenal Bulb and the Pars Descendens. Since the bowel resembles a tube, the chosen perspective considerably changes among images. Textures within images from the Bulbus region lie in the tangent plane to the surface, therefore, the most important distortion is that caused by the endoscopes’ optics. The mucosa texture seen within the Pars Descendens region varies between a tangential orientation to a perspective that points out of the surface of the image. Consequently, distortions with respect to texture homogeneity are also caused by differences in perspective in addition to the optics’ distortion. Thus, we concentrate on image material taken from the Duodenal Bulb in this work.

We have created a set of textured image patches with optimal quality to assess if the required classification is feasible under “idealistic” conditions. Thus, the captured data was inspected and filtered by several qualitative factors (sharpness, distortions, visibility of features). In the next step, texture patches with a fixed size of  $128 \times 128$  pixels were extracted in a manual fashion (the size turned out to be optimally suited in earlier experiments on automated celiac disease diagnosis [44]).

To generate the ground truth for the texture patches used, the condition of the mucosal areas covered by the images was determined by histological examination of biopsies from the corresponding regions. Severity of villous atrophy was classified according to the modified Marsh classification in [14].

Table 1 shows the number of images available per considered image class, where  $\mathbf{N}_O$ ,  $\mathbf{N}_E$ , and  $\mathbf{N}_P$  denote the number of original endoscopic images, the number of patches extracted from these images, and the number of patients, respectively.

If we only use patches close to the CoD during the classification process, the effect of distortion correction with respect to feature sizes together with its potential benefits for texture analysis would be negligible, only artifacts caused by interpolation might degrade the classification result. In order to clarify this, we plot the center position of all texture patches used along with the boundary all patches fall into in Fig. 6. The gray background in these images denotes the area which actually shows the image as taken by the endoscope.

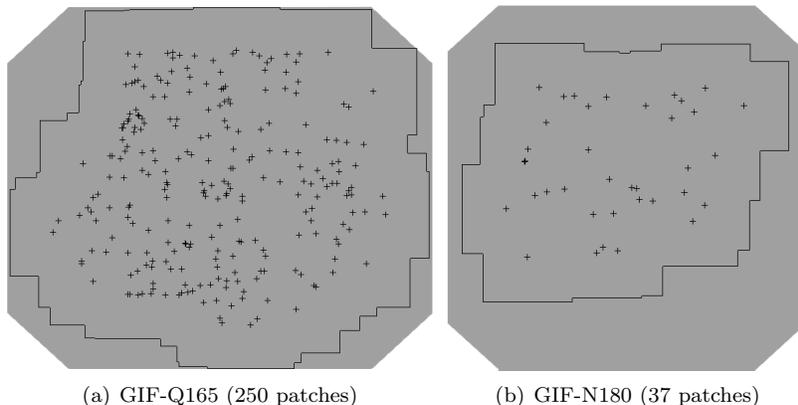


Figure 6: Distribution of patch centers along with the boundary all used patches fall into.

We notice that especially for the GIF-Q165 endoscope, the patches are well distributed (except for the areas where patient related information is overlayed in the upper and lower left corner). In addition, from the boundary shown, we notice that the patches used also cover the marginal parts of the visible endoscope region fairly well. For the GIF-N180 endoscope, a rather small number of patches is used overall, but still these are sufficiently well distributed. Obviously it turns out that it is not the case that only patches close to the CoD are involved in our application, so there should be room for classification improvements by distortion correction.

### 6.2. Classification Setup

In order to evaluate the different combinations of DC methods and interpolation methods in conjunction with the feature extraction methods, we use the nearest neighbor (NN) classifier for the classification process. This rather weak classifier has been chosen to emphasize more on the effect of different combinations of DC methods and interpolation methods and not on highest possible overall classification rates.

To estimate the classification accuracies we use the leave-one-patient-out cross-validation (LOPO-CV) [45] in order to avoid any bias during classification. In this setup one image out of the database is considered as an unknown image. The remaining images are used to train the classifier (omitting those images which originate from the same patient as the image left out). The class of the unknown image is then predicted by the system. These steps (training and prediction) are repeated for each image, yielding an estimate of the overall classification rate.

### 6.3. Distortion Correction Calibration and Setup

For the distortion calibration, performed prior to the distortion correction, we used a set of images containing calibration patterns (i.e. checkerboard pat-

terns, see Fig. 4(a)). The calibration has been carried out as follows (for each endoscope used):

- **Even-order polynomial model [19] (see Section 4.1.1)**  
Between 140 and 144 calibration points have been extracted out of four different barrel-type distorted images. These points have then been fed into the calibration software used.
- **Polynomial model [20] (see Section 4.1.2)**  
This method requires sets of points which lie on a straight line in the undistorted image. To this end we manually extracted three lines from an undistorted image, each consisting of five points. This has been repeated for four different calibration images. For the calibration process we used the software provided by the authors.
- **Parameter-free model [21] (see Section 4.1.3)**  
For an appropriate calibration of this method a significantly higher number of manually extracted points is necessary, as compared to the other methods. Hence, we extracted between 326 and 429 points out of 9 to 14 different images. For the calibration as well as for the distortion correction we used our own implementation [11, 12].
- **Division model [22] (see Section 4.1.4)**  
In order to carry out the distortion calibration for this method, we used the software as provided by the authors. This software allows an automatic recognition of the calibration points. However, to ensure an accurate calibration, we manually corrected the points in case of recognition failures. The calibration is based on 72 to 97 points extracted from three different calibration images.

It must be noted, that for each DC method the number of calibration points has been chosen such that the calibration was as accurate as possible (i.e. the different numbers are no arbitrary choices).

As explained in the previous section, the texture patches as counted in Table 1 have been obtained by manually selecting  $128 \times 128$  pixels sized squares. Since after distortion correction these data do no longer correspond to squares these cannot be used immediately for subsequent classification (most evaluated feature extraction techniques implicitly assume at least a rectangularly shaped texture patch).

Therefore we apply the following technique to generate square-shaped texture from distortion corrected image material: based on the original (distorted) endoscopic images, we record the coordinates of the center of the extracted  $128 \times 128$  pixels. Subsequently, distortion correction is applied to the entire original images and the recorded center coordinates are mapped into the distortion corrected image. Using these coordinates, a  $128 \times 128$  pixels texture square is extracted from the distortion corrected image which is then used for classification. This technique is also illustrated in Fig. 7.

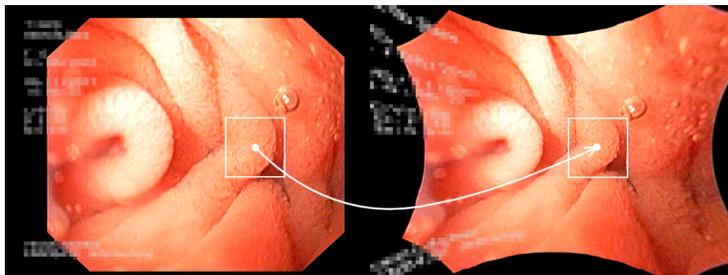


Figure 7: Generation of distortion corrected texture patches

#### 6.4. Experimental Results

Figure 8 shows the minimum and maximum overall classification rates obtained by our feature extraction methods, for all different combinations of DC methods and interpolation methods. In this figure the black solid line shows the results obtained on the original images (i.e. without any distortion correction applied). The green areas denote rates which are higher as compared to the results with the original images, whereas the red area shows rates below the results of the original images.

From this figure we immediately notice that there is only one method for which the results after distortion correction are always higher or at least comparable to the original rates (WT-GMRF-CNH). In case of this method we consistently achieve higher accuracies, no matter which combination of DC and interpolation method is used. For many other methods, applying distortion correction sometimes improves and sometimes lowers the classification rates (depending on the combination of the DC and interpolation method used). In case of quite a few methods we even observe consistently lower results when applying distortion correction (WT-LDB, WT-BBC, LCVP, ECM, SSD, SCH, and FRACTAL).

Regarding the type of the features (i.e. spatial domain or transform domain) there is no clear trend visible. Some of the feature extraction methods, which are able to yield a result improvement by applying a DC method, analyze an image at multiple resolutions (e.g. WT-GMRF and WT-GMRF-CNH). But this also applies to methods, which perform rather poor when a DC method is applied (e.g. WT-LDB and WT-BBC).

From Fig. 8 we also notice that the results of the different combinations of DC methods and interpolation methods may exhibit a high variation from feature to feature. Some features seem to be very sensitive to the right choice of such a combination (e.g. WT-LDB, GMRF, and SCH). Other methods, on the other hand, deliver roughly the same overall classification rates, no matter which combination of DC method and interpolation method is used (e.g. JC-MB-LBP, LBP, LTP, and GLCM).

The detailed overall classification rates from our experiments are given in Table 2. In this table result improvements by distortion correction (as compared

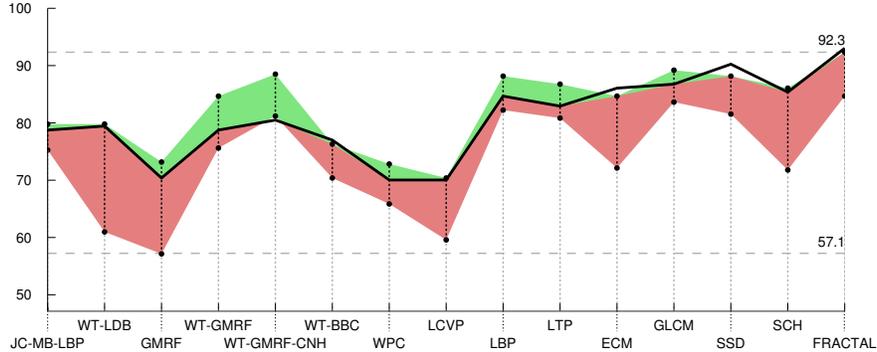


Figure 8: Minimum and maximum overall classification rates obtained after applying different combinations of DC methods and interpolation methods (shown by the black dots) and the results obtained by classifying the distorted images (solid black line).

to the original rates) are shown in bold face.

A more detailed view on the results from Fig. 8 is given in Fig. 9. This figure shows the results obtained by the different feature extraction methods after applying distortion correction. The differently colored bars show the minimum, maximum, and mean classification rates for all DC methods among the different interpolation methods used. For comparison purposes we also show the results obtained by the different feature extraction methods when applied to the original images (wide gray bars in the background). From this figure we notice that the DC method by Melo et al. in most cases delivers the highest overall classification rates among all DC methods tested (this applies to the minimum, maximum, and mean rates). We also notice that in case of the WT-GMRF method the different DC methods deliver roughly the same classification rates.

A similar view on the results is provided by Fig. 10. But, in contrast to Fig. 9, this figure compares the different interpolation methods evaluated. Hence, the colored bars show the minimum, maximum, and mean classification rates for all DC methods among the different DC methods used. Again, we also show the results obtained by the different feature extraction methods when applied to the original images. From this figure we notice that none of the interpolation methods is able to consistently improve the classification rates when used for distortion correction.

## 7. Discussion

The results presented in Section 6.4 showed that the classification rates obtained after a distortion correction highly depend on the features used. In this section we therefore aim at finding reasonable explanations for this outcome.

### 7.1. Impact of Distortions on the Features

Barrel-type distortions as well as a distortion correction induce distortions to the endoscopic images (i.e. downscaling of image portions at the image borders

	Original	Alvarez					Melo				
		Cubic	Lanczos	Linear	Nearest	Wang	Cubic	Lanczos	Linear	Nearest	Wang
JC-MB-LBP	78.7	78.4	78.4	78.7	76.0	77.7	75.3	78.4	78.0	78.0	75.3
WT-LDB	79.4	68.3	69.0	68.6	69.0	67.9	76.7	71.4	76.0	<b>79.8</b>	75.6
GMRF	70.4	62.4	62.4	64.8	64.1	69.0	65.5	62.0	<b>73.2</b>	68.3	<b>71.4</b>
WT-GMRF	78.7	<b>84.7</b>	<b>84.0</b>	<b>83.3</b>	<b>82.2</b>	<b>83.6</b>	<b>84.7</b>	<b>84.3</b>	<b>84.7</b>	<b>83.3</b>	<b>84.3</b>
WT-GMRF-CNH	80.5	<b>83.6</b>	<b>83.3</b>	<b>88.2</b>	<b>85.4</b>	<b>87.5</b>	<b>87.1</b>	<b>82.6</b>	<b>88.2</b>	<b>86.1</b>	<b>87.5</b>
WT-BBC	77.0	76.0	75.6	72.8	75.3	72.8	73.9	76.0	70.7	76.3	70.7
WPC	70.0	<b>70.7</b>	67.6	69.0	<b>71.4</b>	68.3	70.0	<b>70.4</b>	67.9	67.6	67.6
LCVP	70.0	63.1	63.8	62.7	<b>70.4</b>	63.4	67.6	66.6	68.6	69.3	66.2
LBP	84.7	<b>85.7</b>	<b>87.1</b>	<b>87.5</b>	84.3	<b>88.2</b>	<b>87.5</b>	<b>87.5</b>	<b>85.4</b>	<b>86.1</b>	<b>87.8</b>
LTP	82.9	<b>84.0</b>	<b>86.8</b>	<b>85.7</b>	<b>83.6</b>	<b>84.7</b>	<b>84.7</b>	<b>83.3</b>	<b>86.4</b>	82.2	<b>84.7</b>
ECM	86.1	77.4	79.1	75.3	80.5	79.1	83.6	84.7	82.9	82.9	81.5
GLCM	86.8	<b>87.1</b>	<b>87.1</b>	85.4	85.7	86.1	<b>89.2</b>	<b>88.9</b>	<b>88.9</b>	<b>88.9</b>	<b>88.5</b>
SSD	90.2	85.0	83.3	86.8	88.2	87.1	84.0	84.3	86.1	84.7	85.4
SCH	85.4	79.8	81.2	81.9	79.1	79.1	81.2	<b>86.1</b>	78.0	80.8	79.4
FRACTAL	93.0	84.7	88.5	87.1	88.5	88.9	89.5	87.1	90.2	87.1	90.9

	Original	Zhang					Hartley				
		Cubic	Lanczos	Linear	Nearest	Wang	Cubic	Lanczos	Linear	Nearest	Wang
JC-MB-LBP	78.7	<b>79.8</b>	78.0	78.4	<b>79.4</b>	78.0	77.0	76.0	77.7	76.7	76.7
WT-LDB	79.4	64.5	64.5	69.0	76.7	71.8	74.2	61.0	68.6	72.1	72.8
GMRF	70.4	57.1	57.1	67.2	67.9	69.0	57.5	59.2	57.5	64.8	63.4
WT-GMRF	78.7	<b>81.2</b>	<b>81.2</b>	<b>83.3</b>	<b>79.8</b>	<b>81.9</b>	<b>84.0</b>	75.6	<b>82.6</b>	<b>80.8</b>	<b>79.8</b>
WT-GMRF-CNH	80.5	<b>85.0</b>	<b>81.2</b>	<b>88.5</b>	<b>81.9</b>	<b>85.7</b>	<b>83.3</b>	<b>81.5</b>	<b>85.4</b>	<b>82.2</b>	<b>84.3</b>
WT-BBC	77.0	73.2	75.3	70.4	75.3	70.7	71.8	74.6	70.7	71.4	72.1
WPC	70.0	66.6	66.2	67.2	69.7	65.9	<b>71.8</b>	<b>72.8</b>	<b>71.4</b>	<b>72.5</b>	<b>72.8</b>
LCVP	70.0	66.9	64.8	62.7	62.4	59.6	64.1	66.9	61.3	62.4	63.1
LBP	84.7	82.2	82.2	82.2	<b>85.0</b>	82.9	<b>85.7</b>	<b>86.8</b>	<b>85.0</b>	<b>85.4</b>	<b>85.4</b>
LTP	82.9	<b>84.0</b>	81.9	<b>85.0</b>	81.5	<b>84.0</b>	<b>85.4</b>	<b>84.3</b>	<b>85.0</b>	80.8	<b>85.0</b>
ECM	86.1	79.4	83.3	78.7	80.1	79.8	79.1	77.7	72.1	77.0	76.3
GLCM	86.8	85.7	85.4	83.6	85.7	83.6	85.0	84.7	85.4	86.4	84.3
SSD	90.2	82.6	83.3	82.9	85.4	84.3	83.3	85.0	83.3	83.3	81.5
SCH	85.4	77.7	78.4	77.0	82.2	78.0	75.6	77.0	73.5	78.4	71.8
FRACTAL	93.0	92.0	88.2	92.3	90.9	91.3	89.5	88.9	90.6	88.5	88.9

Table 2: Detailed overall classification rates for all combinations of feature extraction methods, DC methods, and interpolation techniques evaluated.

in case of barrel-type distortions and interpolation artifacts in case of distortion correction). Depending on the sensitivity of the feature extraction method used, such distortions may also influence the resulting features. As a consequence patches may potentially get misclassified since the distance between two feature vectors is changing too in case of sensitive features. We therefore conducted an analysis which aims at providing an answer to the following question: Does the distance of a patch to the CoD influence the features (i.e. which features are sensitive to different levels of distortions)? If this is the case we also want to know whether applying a DC method is able to lower this effect.

Since all our images are suffering from distortions we had to simulate the

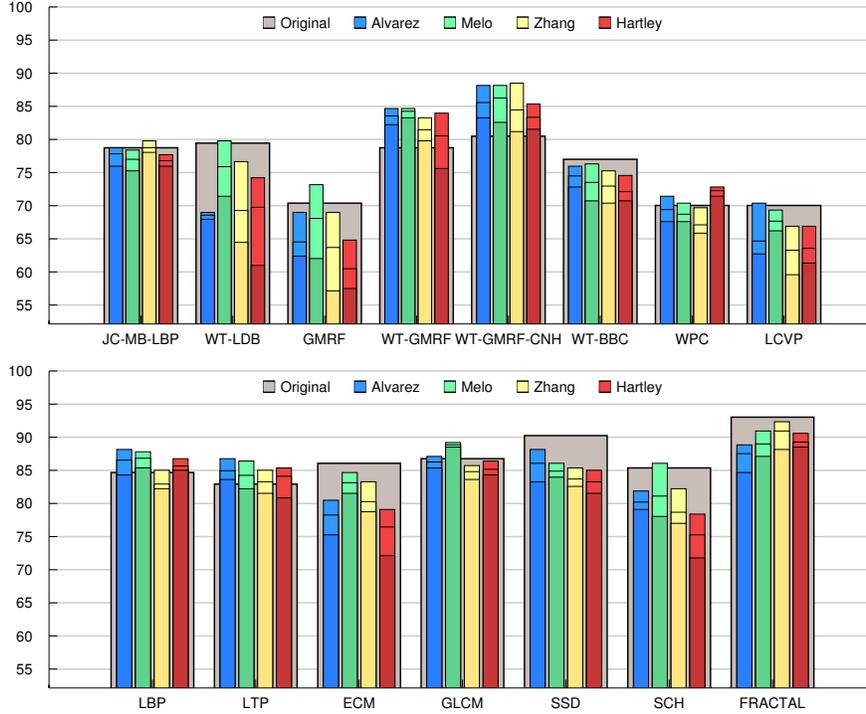


Figure 9: Comparison of the overall classification rates for the different feature extraction methods evaluated after applying the DC methods. The lines at the top of the bars indicate the minimum, maximum, and mean rates over all interpolation techniques used.

different levels of distortion for our analysis. For this purpose we extracted a patch with a size of  $128 \times 128$  pixels centered at the CoD from an originally sized distorted image (denoted here as  $P^{(c)}$ ). Since the distortion is rather small around the CoD, it is a reasonable assumption that this patch is nearly free of distortions.

Based on the division model (using bilinear interpolation) we then distort the patch  $P^{(c)}$  to obtain a distorted patch  $P_{x,y}^{(d)}$  centered at  $(x, y)$ . Depending on the distance of the patch  $P_{x,y}^{(d)}$  to the CoD we therefore obtain patches with different levels of distortions. We carry out the distortion simulation for all positions located on a regular grid across the image plane, which is centered at the CoD. This way it is assured that the location of  $P_{0,0}^{(d)}$  exactly corresponds to the location of  $P^{(c)}$ . To obtain a high number of different patches  $P_{x,y}^{(d)}$  we have chosen a grid spacing of two pixels.

Having computed the patches  $P_{x,y}^{(d)}$  for all grid points, we compute the feature distance between the feature vectors extracted from the patches  $P_{0,0}^{(d)}$  and  $P_{x,y}^{(d)}$  (for one specific feature extraction method). For a patch  $P_{x,y}^{(d)}$  we denote this feature distance as  $D_{x,y}$ . Since we are interested in the feature distances

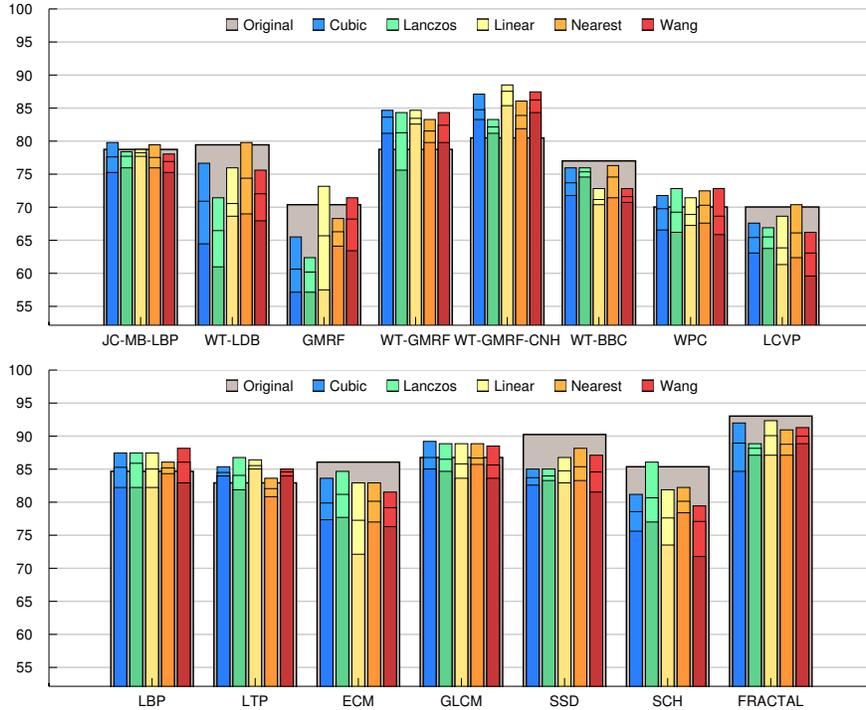


Figure 10: Comparison of the overall classification rates for the different feature extraction methods evaluated after applying the DC methods with different interpolation techniques. The lines at the top of the bars indicate the minimum, maximum, and mean rates over all DC methods used.

depending on the distance to the CoD we compute the mean feature distance for different annuli of same width (i.e. concentric rings centered at the CoD) based on  $D_{x,y}$ . As a consequence, the maximum possible Euclidean distance from the CoD to the edges of the source image ( $\approx 360$  pixels) is divided into 40 concentric rings. The number of rings used has been determined experimentally and seems to be an appropriate choice to capture enough details without being sensitive to noise.

This analysis has been carried out for all feature extraction methods used. To be able to investigate the influence of a DC method on the feature distances, we also carried out such an analysis for distortion corrected images. This has been done by applying the DC method (division model) to the patches  $P_{x,y}^{(d)}$  and again computing the mean feature distances for the different annuli. Since images from healthy patients usually contain more high-frequency details we carried out the analysis for a no-celiac image as well as for a celiac image.

The resulting plots for all feature extraction methods are shown in Fig. 11. In these plots the results for the image class “No-celiac” are shown in black, whereas the feature distances for the class “Celiac” are shown in green. The

dotted lines show the results for the distorted patches  $P_{x,y}^{(d)}$ , while the solid lines show the results after applying a DC method to the patches  $P_{x,y}^{(d)}$ . The x-axis in these plots indicates the Euclidean distance of the patches  $P_{x,y}^{(d)}$  in a certain annulus to the CoD. The y-axis shows the normalized mean feature distance between the patches  $P_{x,y}^{(d)}$  and the center patch  $P^{(c)}$ . The normalization can be justified by the fact that we are only interested in the relative differences. In addition, the ranges for feature differences vary significantly depending on the feature used. While in endoscopic images we usually have black borders of a certain width, restricting the locations for meaningful patches, in our analysis we investigated all possible patch positions within the image bounds. As a consequence, our distance computations are also carried out for image regions which are usually not used to extract patches. The respective range of distances to the CoD (approximately distances greater than 295 pixels) is indicated by a gray rectangle on the right of the plots in Fig. 11. It must be noted that the gray regions are left out from the discussion below.

From Fig. 11 we immediately notice that the distance of a distorted patch (i.e. dashed lines) to the CoD indeed has an impact on the feature distances. For almost all methods the feature distances get significantly higher the farther away the patches are located from the CoD (i.e. the more distorted the patches are). While this accounts to celiac patches as well as to no-celiac patches, for more than half the methods (LBP, LTP, GMRF, WT-GMRF, WT-GMRF-CNH, SSD, SCH, GLCM, and WT-BBC) the feature distances are at least slightly higher for no-celiac images as compared to celiac images (especially for higher levels of distortion). An explanation for this phenomenon is the fact that these images contain more high-frequency details as compared to the images from the class ‘‘Celiac’’. Since, as already mentioned earlier, areas of the image appear significantly smaller than they actually are at the outer regions of distorted images, the details are shifted towards higher frequencies in the frequency spectrum, resulting in a change of the features. Images from the class ‘‘Celiac’’, on the other hand, do not contain many high-frequency details, thus the impact of the distortions is not that pronounced.

If distortion correction is applied to the patches (i.e. the solid lines) we again notice that the feature distances get higher for patches farther away from the CoD. This is most probably caused by the interpolation artifacts induced by the DC method as these get stronger the farther away a patch is from the CoD. But for most features this effect is not as dramatic as observed for the distorted patches. Now for almost half of the methods (JC-MB-LBP, LCVP, WT-GMRF, WT-GMRF-CNH, ECM, and FRACTAL) the feature distances are at least slightly higher for celiac images as compared to no-celiac images (especially for higher levels of distortion). But for the remaining methods the distances are now roughly equal when comparing the different image classes. This may be explained by the fact that applying a DC method introduces interpolation artifacts which potentially cancel out high-frequency details in no-celiac images. In case of celiac images this effect is not that pronounced since these mostly contain low-frequency details only.

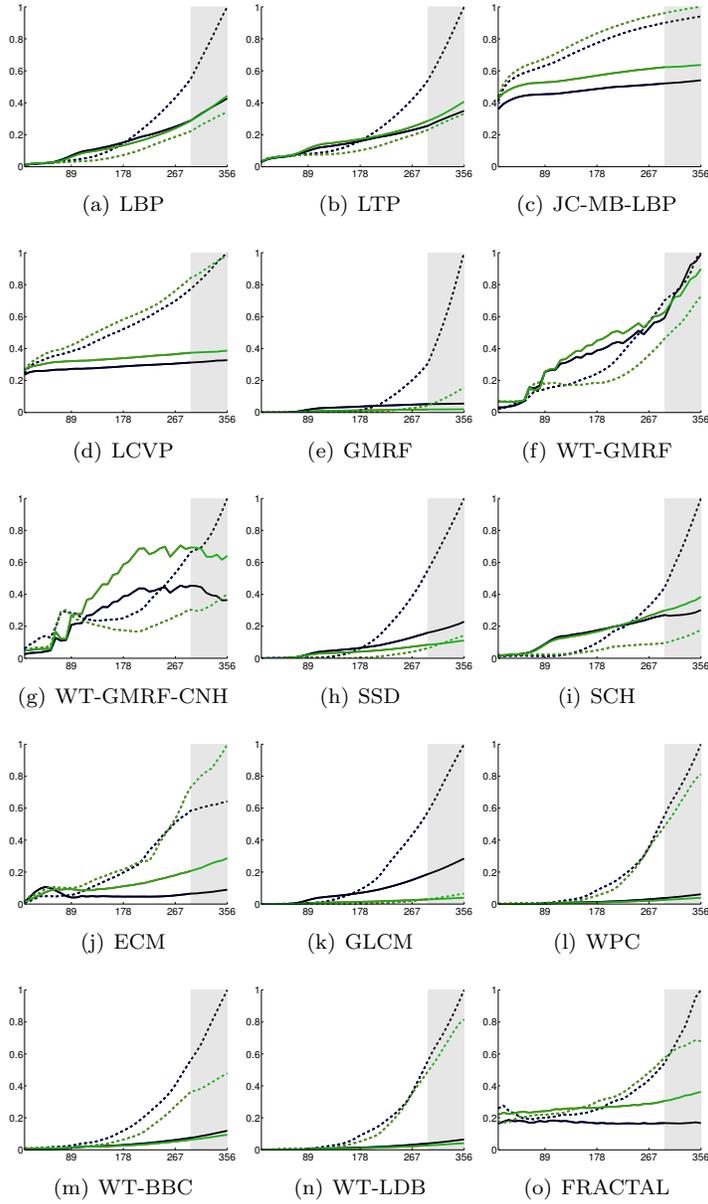


Figure 11: Normalized feature distances in dependence of the distance to the CoD (x-axis). The distances for the distorted and undistorted patches are shown as dotted lines and solid lines, respectively. Distances for non-celiac and celiac patches are shown in black and green, respectively (for interpretation of the references to color in this figure, the reader is referred to the web version of the article).

When comparing the distances for distorted and undistorted patches (dashed vs. solid lines) we notice that distortion correction leads to smaller feature distances for the majority of feature extraction methods. Hence, although the interpolation artifacts introduced by DC lead to higher distances for patches farther away from the CoD, the distances are nevertheless in most cases significantly lower as for the distorted patches (especially for higher distances to the CoD). But there are a few methods for which this applies to one image class only (LBP, LTP, WT-GMRF-CNH, SSD, and SCH). And for one method the distances are on average even getting higher (WT-GMRF).

The bottom line of this analysis is that the majority of the features used is indeed very sensitive to the level of distortion present in an image patch. In addition, we showed that, by employing a DC method, we are able to lower this effect to some extent for the majority of methods used.

### 7.2. Impact of the Level of Distortion on the Classification Accuracy

In the previous section we showed that the majority of features used is sensitive to barrel-type distortions and that we are able to cancel out this effect to some extent. In this section we want to investigate the impact of barrel-type distortions and distortion correction on the classification rates for our methods. More specifically we aim at answering the following questions: What is the distortion difference between patches found as nearest neighbors by the NN-classifier used? How do the distortion differences affect the classification rates? If there is an impact of distortions on the classification rates, we also want to examine whether we are able to lower the number of misclassifications by using a DC method. Hence, the following analysis is similar to the one already conducted in [10].

In case only patches equidistant to the CoD would be used for classification, an analysis regarding the CoD distances and the according levels of distortion of the patches would be meaningless. This is due to the roughly equal level of distortion the patches would exhibit in such a case. As we notice from Fig. 6 the majority of patches in our image database has been acquired with the endoscope GIF-Q165 (250 patches). Compared to this, the number of patches for the second endoscope is quite low (35). As a consequence we restrict our analysis to the first endoscope. It must be noted that the images are actually based on two GIF-Q165 endoscopes with slightly different CoDs. These differences are most probably caused by marginal differences in the optics used in the endoscopes. However, the distance between the two CoDs is minimal with about 4 pixels, hence we neglect it for the present analysis.

From Fig. 6(a) we notice that the patches are well distributed (except for the areas where patient related information is overlaid in the upper and lower left corner). Moreover, the distribution is sufficient to perform an analysis since this figure also shows that it is not the case that the set contains patches close to the CoD only.

In the distorted case as well as in the undistorted case the Euclidean distance of a patch to the CoD  $D(P_i, CoD)$  also reflects the level of distortion of the patch – in the distorted case due to the barrel-type distortion and in the undistorted

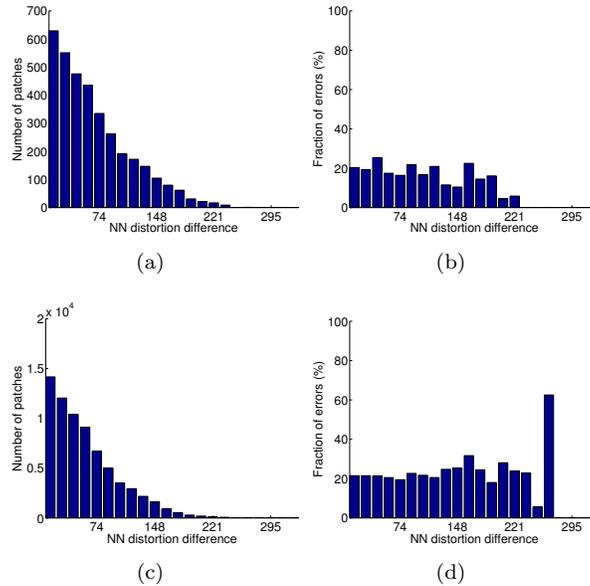


Figure 12: Distortion differences between patches and the chosen nearest neighbor patches (distorted (a) and undistorted case (c)) and the probabilities of misclassification depending on the distortion differences (distorted (b) and undistorted case (d)).

case due to the interpolation artifacts generated. Hence, it is reasonable to consider the difference  $|D(P_i, CoD) - D(NN(P_i), CoD)|$  to be the distortion difference between  $P_i$  and  $NN(P_i)$ , where  $NN(P_i)$  denotes the nearest neighbor patch for  $P_i$ . Figures 12(a) and 12(c) show the distribution of the distortion differences between each classified patch  $P_i$  and  $NN(P_i)$ . From these plots we notice that the NN-classifier favors patches as nearest neighbors which exhibit rather low distortion differences to the patch to be classified (in the distorted case as well as after applying a DC method). Figures 12(b) and 12(d) show the probability of classifying a patch  $P_i$  wrongly depending on the distortion difference to  $NN(P_i)$ . From these figures we notice that the probability for the misclassification of a patch does not really depend on the distortion difference between  $P_i$  and  $NN(P_i)$  (i.e. the fractions of errors in these plots remain roughly equal for different distortion differences). This applies to distorted patches as well as to undistorted patches. This may be a reasonable explanation why there is no clear trend visible in the classification results when applying distortion correction.

It must be noted that the plots in Fig. 12 are accumulated across all feature extraction methods used. For the undistorted images the plots are also accumulated over all combinations of DC methods and interpolation methods. The reason for accumulation is that otherwise the number of patches for the analysis would be rather low, resulting in noisy histograms without a chance for an interpretation.

## 8. Conclusion

In this work we compared various different feature extraction methods in the context of the correction of barrel-type distortions in endoscopic images. The special focus was a comparison of different combinations of DC methods and interpolation methods.

Our experiments showed that the classification results after a distortion correction are highly feature dependent. While for certain features we observe a consistent improvement of the classification results (e.g. WT-GMRF-CNH), there are also features for which the results always drop as soon as a DC method is applied to the imagery (e.g. ECM). We also saw, that some feature extraction methods are rather sensitive to the right combination of the DC method and interpolation method, while for other features the resulting variations in the classification rates between different combinations are not very high.

When comparing the results between the different DC methods, we notice that there is only one DC method which almost always yields comparable or even better results as compared to the other DC methods for each feature extraction method (i.e. the method by Melo et al.). This may be attributed to the fact that this method is the simplest one among all DC methods evaluated since only one distortion parameter has to be estimated. This in turn makes the estimation process more robust. A comparison of the different interpolation methods shows that there is no interpolation method which is consistently superior to the others in terms of the classification rates (among the interpolation methods evaluated).

The bottom line of our experiments is that it is not possible to make a clear recommendation on whether distortion correction in the context of an automated celiac disease diagnosis should be used or not, in order to improve the classification accuracies. This is due to the fact that we showed that the feature distances indeed depend on the distance of a patch to the CoD. But we also showed that the distortion difference between a patch and its nearest neighbor patch has no significant influence on whether the patch gets misclassified or not. Hence, at least for most of the feature extraction methods evaluated in this work and the underlying imagery it seems that in general we are not able to draw a benefit from distortion correction.

Since the present study is restricted to celiac disease images only, we plan to conduct similar studies on other image databases as well (e.g. high-definition colonoscopic images) in order to be able to assess if the results shown are generic or restricted to specific types of imagery. In addition, we also aim at investigating features, which can be computed from pixels located on an irregular grid, thus alleviating the need for an interpolation.

## 9. Acknowledgments

This work is partially funded by the Austrian Science Fund (FWF) under Project No. 24366.

## References

- [1] M. Liedlgruber, A. Uhl, A summary of research targeted at computer-aided decision support in endoscopy of the gastrointestinal tract, Tech. Rep. 2011-01, Department of Computer Sciences, University of Salzburg, Austria, <http://www.cosy.sbg.ac.at/research/tr.html> (2011).
- [2] G. Hanna, A. Cuschieri, Image display technology and image processing, *World Journal of Surgery* 25 (11) (2001) 1419–1427.
- [3] T. B. Borchardt, A. Conci, M. C. d’Ornellas, A warping based approach to correct distortions in endoscopic images, in: *Proceedings of the 22nd Brazilian Symposium on Computer Graphics and Image Processing (Sibgrapi’09)*, Rio de Janeiro, Brazil, 2009.
- [4] K. V. Asari, S. Kumar, D. Radhakrishnan, A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation, *IEEE Transactions on Medical Imaging* 18 (4) (1999) 345–354.
- [5] H.-X. Sun, Y.-H. Zhang, F.-L. Luo, A novel approach for nonlinear distortion correction of industrial endoscope images, in: *Proceedings of the 17th World Conference on Nondestructive Testing*, Shanghai, China, 2008, pp. 25–28.
- [6] H. Haneishi, Y. Yagihashi, Y. Miyake, A new method for distortion correction of electronic endoscope images, *IEEE Transactions on Medical Imaging* 14 (3) (1995) 548–55.
- [7] J. P. Helferty, C. Zhang, G. McLennan, W. E. Higgins, Videoendoscopic distortion correction and its application to virtual guidance of endoscopy, *IEEE Transactions on Medical Imaging* 20 (7) (2001) 605–617.
- [8] W. Li, S. Nie, M. Soto-Thompson, C.-I. Chen, Y. I. A-Rahim, Robust distortion correction of endoscope, in: *Proceedings of Medical Imaging 2008: Visualization, Image-guided Procedures, and Modeling*, Vol. 2008, 2008.
- [9] M. Gschwandtner, M. Liedlgruber, A. Uhl, A. Vécsei, Experimental study on the impact of endoscope distortion correction on computer-assisted celiac disease diagnosis, in: *Proceedings of the 10th International Conference on Information Technology and Applications in Biomedicine (ITAB’10)*, Corfu, Greece, 2010.
- [10] M. Liedlgruber, A. Uhl, A. Vécsei, Statistical analysis of the impact of distortion (correction) on an automated classification of celiac disease, in: *Proceedings of the 17th International Conference on Digital Signal Processing (DSP’11)*, Corfu, Greece, 2011.

- [11] M. Gschwandtner, J. Hämmerle-Uhl, Y. Höller, M. Liedlgruber, A. Uhl, A. Vécsei, Improved endoscope distortion correction does not necessarily enhance mucosa-classification based medical decision support systems, in: Proceedings of the IEEE International Workshop on Multimedia Signal Processing (MMSP'12), 2012.
- [12] J. Hämmerle-Uhl, Y. Höller, A. Uhl, A. Vécsei, Endoscope distortion correction does not (easily) improve mucosa-based classification of celiac disease, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI'12), Vol. 7512 of Springer Lecture Notes in Computer Science, 2012, pp. 574–581.
- [13] M. Marsh, Gluten, major histocompatibility complex, and the small intestine. a molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'), *Gastroenterology* 102 (1) (1992) 330–354.
- [14] G. Oberhuber, G. Granditsch, H. Vogelsang, The histopathology of coeliac disease: time for a standardized report scheme for pathologists, *European Journal of Gastroenterology and Hepatology* 11 (1999) 1185–1194.
- [15] M. Liedlgruber, A. Uhl, Computer-aided decision support systems for endoscopy in the gastrointestinal tract: A review, *IEEE Reviews in Biomedical Engineering* 4 (2012) 73–88.
- [16] G. Cammarota, L. Cuoco, P. Cesaro, L. Santoro, A. Cazzato, M. Montalto, R. L. Mura, L. M. Larocca, F. M. Vecchio, A. Gasbarrini, M. Salvagnini, G. Gasbarrini, A highly accurate method for monitoring histological recovery in patients with celiac disease on a gluten-free diet using an endoscopic approach that avoids the need for biopsy: a double-center study, *Endoscopy* 39 (1) (2007) 46–51.
- [17] R. Petroniene, E. Dubcenco, J. P. Baker, Given capsule endoscopy in celiac disease: evaluation of diagnostic accuracy and interobserver agreement, *The American Journal of Gastroenterology* 100 (3) (2005) 685–694.
- [18] J. Barreto, R. Swaminathan, J. Roquette, Non parametric distortion correction in endoscopic medical images, in: Proceedings of the 3DTV Conference, 2007, pp. 1–4.
- [19] Z. Zhang, A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (11) (2000) 1330–1334.
- [20] L. Alvarez, L. Gomez, J. R. Sendra, An algebraic approach to lens distortion by line rectification, *Journal of Mathematical Imaging and Vision* 35 (1) (2009) 36–50.
- [21] R. Hartley, S. Kang, Parameter-free radial distortion correction with center of distortion estimation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (8) (2007) 1309–1321.

- [22] R. Melo, J. P. Barreto, G. Falcao, A new solution for camera calibration and real-time image distortion correction in medical endoscopy-initial technical evaluation, *IEEE Transactions on Biomedical Engineering* 59 (3) (2012) 634–44.
- [23] A. W. Fitzgibbon, Simultaneous linear estimation of multiple view geometry and lens distortion, in: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01)*, 2001, pp. 125–132.
- [24] T. M. Lehmann, C. Gönner, K. Spitzer, Survey: Interpolation methods in medical image processing, *IEEE Transactions on Medical Imaging* 18 (1999) 1049–1075.
- [25] C. E. Duchon, Lanczos filtering in one and two dimensions, *Journal of Applied Meteorology* 18 (8) (1979) 1016–1022.
- [26] F. Wang, Y. Xu, Y. Zhao, F. Hu, A new nonlinear interpolation algorithm for edge preserving, in: *2010 International Conference on Multimedia Technology (ICMT'10)*, 2010, pp. 1–4.
- [27] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on feature distributions, *Pattern Recognition* 29 (1) (1996) 51–59.
- [28] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution Gray-Scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (7) (2002) 971–987.
- [29] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, in: *Analysis and Modelling of Faces and Gestures*, Vol. 4778 of LNCS, Springer, 2007, pp. 168–182.
- [30] R. M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, *IEEE Transactions on Systems, Man, and Cybernetics* 3 (1973) 610–621.
- [31] M. Gadermayr, M. Liedlgruber, A. Uhl, Shape curvature histogram: A shape feature for celiac disease diagnosis, Tech. Rep. 2012-06, Department of Computer Sciences, University of Salzburg, Austria, <http://www.cosy.sbg.ac.at/research/tr.html> (2012).
- [32] R. Rautakorpi, J. Iivarinen, A novel shape feature for image classification and retrieval, in: *Proceedings of the International Conference on Image Analysis and Recognition (ICIAR'04)*, 2004, pp. 753–760.
- [33] G. Ayala, J. Domingo, Spatial size distributions: Applications to shape and texture analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23 (12) (2001) 1430–1442.

- [34] M. Häfner, A. Gangl, M. Liedlgruber, A. Uhl, A. Vécsei, F. Wrba, Combining Gaussian Markov random fields with the discrete wavelet transform for endoscopic image classification, in: Proceedings of the 17th International Conference on Digital Signal Processing (DSP'09), Santorini, Greece, 2009, pp. 177–182.
- [35] S. Geman, D. Geman, Stochastic relaxation, gibbs distributions, and the bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6 (1984) 721–741.
- [36] M. Häfner, A. Gangl, M. Liedlgruber, A. Uhl, A. Vécsei, F. Wrba, Pit pattern classification using extended local binary patterns, in: Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine (ITAB'09), Larnaca, Cyprus, 2009, pp. 1–4.
- [37] M. Häfner, M. Liedlgruber, A. Uhl, A. Vécsei, F. Wrba, Color treatment in endoscopic image classification using multi-scale local color vector patterns, *Medical Image Analysis* 16 (1) (2012) 75–86.
- [38] A. Uhl, A. Vécsei, G. Wimmer, Fractal analysis for the viewpoint invariant classification of celiac disease, in: Proceedings of the 7th International Symposium on Image and Signal Processing (ISPA'11), 2011, pp. 727–732.
- [39] M. Häfner, M. Liedlgruber, F. Wrba, A. Gangl, A. Vécsei, A. Uhl, Pit pattern classification of zoom-endoscopic colon images using wavelet texture features, in: W. Sandham, D. Hamilton, C. James (Eds.), Proceedings of the International Conference on Advances in Medical Signal and Image Processing (MEDSIP'06), Glasgow, Scotland, UK, 2006.
- [40] M. Liedlgruber, A. Uhl, Statistical and structural wavelet packet features for Pit pattern classification in zoom-endoscopic colon images, in: P. Dondon, V. Mladenov, S. Impedovo, S. Cepisca (Eds.), Proceedings of the 7th WSEAS International Conference on Wavelet Analysis & Multirate Systems (WAMUS'07), Arcachon, France, 2007, pp. 147–152.
- [41] R. R. Coifman, M. V. Wickerhauser, Entropy based methods for best basis selection, *IEEE Transactions on Information Theory* 38 (2) (1992) 719–746.
- [42] N. Saito, R. R. Coifman, Local discriminant bases and their applications, *Journal of Mathematical Imaging and Vision* 5 (4) (1995) 337–358.
- [43] A. Gasbarrini, V. Ojetti, L. Cuoco, G. Cammarota, A. Migneco, A. Armuzzi, P. Pola, G. Gasbarrini, Lack of endoscopic visualization of intestinal villi with the immersion technique in overt atrophic celiac disease, *Gastrointestinal endoscopy* 57 (2003) 348–351.
- [44] S. Hegenbart, R. Kwitt, M. Liedlgruber, A. Uhl, A. Vécsei, Impact of duodenal image capturing techniques and duodenal regions on the performance

of automated diagnosis of celiac disease, in: Proceedings of the 6th International Symposium on Image and Signal Processing and Analysis (ISPA'09), Salzburg, Austria, 2009, pp. 718–723.

- [45] M. Liedlgruber, A. Uhl, Predicting pathology in medical decision support systems in endoscopy of the gastrointestinal tract, in: C. Jao (Ed.), Efficient Decision Support Systems – Practice and Challenges in Biomedical Related Domain, InTech, Rijeka, Croatia, 2011, pp. 195–214.