



ELSEVIER

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Survey on computer aided decision support for diagnosis of celiac disease

Sebastian Hegenbart ^{a,*}, Andreas Uhl ^a, Andreas Vécsei ^b

^a Department of Computer Sciences, University of Salzburg, Jakob-Haringer Strasse, 5020 Salzburg, Austria

^b St. Anna Children's Hospital, Medical University Vienna, 1090 Vienna, Austria

ARTICLE INFO

Article history:

Received 20 November 2014

Accepted 11 February 2015

Keywords:

Computer
Automated
Assisted
Diagnosis
Celiac
Disease
Endoscopy

ABSTRACT

Celiac disease (CD) is a complex autoimmune disorder in genetically predisposed individuals of all age groups triggered by the ingestion of food containing gluten. A reliable diagnosis is of high interest in view of embarking on a strict gluten-free diet, which is the CD treatment modality of first choice. The gold standard for diagnosis of CD is currently based on a histological confirmation of serology, using biopsies performed during upper endoscopy. Computer aided decision support is an emerging option in medicine and endoscopy in particular. Such systems could potentially save costs and manpower while simultaneously increasing the safety of the procedure. Research focused on computer-assisted systems in the context of automated diagnosis of CD has started in 2008. Since then, over 40 publications on the topic have appeared. In this context, data from classical flexible endoscopy as well as wireless capsule endoscopy (WCE) and confocal laser endomicroscopy (CLE) has been used. In this survey paper, we try to give a comprehensive overview of the research focused on computer-assisted diagnosis of CD.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Celiac disease (CD) is a multisystemic immune-mediated disease, which is associated with considerable morbidity and mortality [4,44,82]. The prevalence of CD in Europe and North America ranges from 1:80 to 1:300 [11,12,28,29,32,61,66]. In the pathogenesis of CD, a pivotal contribution is made by a dysregulated immune response directed against tissue transglutaminase type 2 (TG2) which has been identified as a prominent autoantigen of CD [27]. The exogenous trigger of this dysregulated immune response is gluten which is ingested along with grain-containing food. Gluten is the main protein of specific grains like wheat, barley and rye.

CD displays itself through a wide spectrum of clinical manifestations including gastrointestinal symptoms such as abdominal pain, chronic diarrhea, bloating, nausea and vomiting as well as more general features like failure to thrive, weight loss and fatigue. There are also numerous extraintestinal CD manifestations like dermatitis herpetiformis [6,97], alopecia areata and cerebellar ataxia [47] and various other neurologic and psychiatric diseases, iron deficiency [65] and premature osteopenia. It has been estimated that seven

out of eight patients with CD remain undiagnosed because most cases of CD only have minor gastrointestinal symptoms [13].

In untreated or inappropriately treated CD the inflammation caused by the dysregulated immune response can disrupt the intestinal mucosa thus leading to a total atrophy of the villi, which are finger-like projections of the mucosa. After embarking on a strict gluten-free diet (GFD), which is the CD treatment modality of first choice, the inflammation gradually subsides allowing for mucosal healing. Strict adherence to a GFD for life is required if acute and chronic complications are to be avoided [88]. To avoid the most severe complications of CD, early and reliable diagnosis of CD for commencing a strict GFD is of vital importance.

Besides for a small proportion of children fulfilling certain clinical, serological and genetic criteria [64], biopsy of the small intestine remains the gold standard for CD diagnosis and is usually recommended as CD requires a lifelong commitment to a strict GFD.

The mucosal alterations caused by CD are classified into different stages of severity. Oberhuber [77] modified a widely used staging of severity by Marsh [75] in 1999. This histological classification scheme identifies six classes, ranging from class Marsh-0 (no visible change of villi structure) up to class Marsh-3C (absent villi).

Unfortunately, the histological staging of biopsies is subject to a significant degree of intra- and inter-observer variability [76,1,85,96,23]. Therefore, observer independent diagnostic methods such as computer-assisted diagnosis systems are urgently needed. Furthermore, the whole diagnostic work-up of CD, including duodenoscopy

* Corresponding author. Tel.: +43 662 8044 6305; fax: +43 662 8044 172.

E-mail addresses: shegen@cosy.sbg.ac.at (S. Hegenbart), uhl@cosy.sbg.ac.at (A. Uhl), andreas.vecsei@stanna.at (A. Vécsei).

with biopsies, is time-consuming, cost-intensive, and rather invasive. Consequently, to save costs, time, and manpower and at the same time increase the safety of the procedure, a less invasive approach avoiding biopsies would be highly desirable. Recent studies by Cammarota et al. [10,8] investigating such endoscopic techniques report reliable results. Current diagnostic methods are entirely observer-dependent and require significant knowledge, expertise, and time.

An interesting field where computer-assisted video analysis could be useful is the application within follow-up endoscopies of celiac patients on a GFD, testing if a re-growth of villi has taken place. Another advantage of computer-assisted video analysis is quality improvement rendered possible by means of telemedicine, for example, by obtaining a second opinion after endoscopic video clips have been transmitted to institutions that have video analysis software at their disposal allowing for an observer independent objective diagnostic evaluation.

A further limitation of the current gold standard for the diagnosis of CD is due to the possibly patchy distribution of intestinal mucosa areas affected by CD in the midst of normal mucosa [5,62]. If, unfortunately, biopsies are taken only from areas of healthy mucosa within the duodenum the proper diagnosis of CD will be missed due to sampling error. A diagnostic tool based on computer-aided pattern analysis of endoscopic video clips could indicate areas that are damaged by CD thus improving the targeting accuracy of biopsy. Additionally, such a system would allow gastroenterologists who do not routinely take duodenal biopsies recognize mucosal alterations triggering a decision to biopsy. So even in a diagnostic scenario involving biopsies, computer-assisted video analysis contributes to a more reliable diagnosis.

In this work, we try to draw a comprehensive picture of the research focused on computer-assisted diagnosis of CD in endoscopic data. In Section 2 an overview of current endoscopic techniques used in automated diagnosis is given with a discussion of the different properties and requirements for computer-assisted systems. Section 3 provides a general overview of techniques devised for classification of CD. Concepts for handling image degradations in the challenging endoscopic environment are presented in Section 4. Common issues and flaws in the methodical evaluation of techniques and the available ground-truth in medical image classification are covered in Section 5. Finally, Section 6 concludes the paper.

2. Endoscopic techniques used in diagnosis of celiac disease

Besides standard upper endoscopy, several new endoscopic approaches for diagnosing CD have been applied [14] and used in research on computer-assisted diagnosis. The modified immersion

technique (MIT [7–9]) allows detailed scanning of the mucosal surface for villi. Technically, water is rapidly instilled into the duodenal lumen after evacuation of air by suction through the endoscope. Villi, if present, straighten up in water and appear as tiny finger-like structures. In a detailed evaluation, Hegenbart et al. [52] have presented strong empirical evidence, that MIT is superior for computer-aided diagnosis as compared to the conventional imaging technique.

Narrow band imaging (NBI [30]) uses specific blue (440–460 nm) and green (540–560 nm) wavelengths for illumination to enhance the contrast of vascular patterns on the mucosal surface. This imaging modality could be a promising technique for classification of celiac tissue. Valitutti et al. [90] recently proposed the use of NBI combined with the water immersion technique which could also be explored for automated diagnosis. At this point in time, a systematic assessment of all four imaging modalities in flexible endoscopy for computer-assisted diagnosis is still open.

Confocal laser endomicroscopy (CLE, also known as visual biopsy) is a novel technology allowing real-time in vivo microscopy. CLE allows the inspection of multiple mucosal layers, employing a laser at different focal points and has been shown to be a promising technique for diagnosis of CD in endoscopy [70].

A general drawback of endoscopy using flexible endoscopes is the limited range. As a way of inspecting a much larger area of the intestine, wireless capsule endoscopy (WCE [80]) is used. In WCE, a small capsule equipped with a camera is swallowed by the patient. The capsule records images of the mucosal tissue during its passage through the intestine, which is then analyzed by a clinician or potentially by an automated system for diagnosis.

Figs. 1 and 2 schematically illustrate the different endoscopic techniques used for computer-assisted diagnosis of CD.

2.1. Endoscopic markers and diagnosis of CD

The most prevalent endoscopic markers for CD include scalloped folds, mosaic patterns of the mucosa and a nodular mucosa [25]. After entering the duodenum with the endoscope, the endoscopist searches for visible CD-induced mucosal lesions indicated by these markers. All of these features become more pronounced and easier detectable when using MIT. Consequently, NBI is employed (if available) to specifically delineate the outline of the residual villous structures (if present). NBI is used because images captured using this modality allow a better assessment of the villous height and shape compared to the conventional white light endoscopy.

It has been reported that the prevalence of endoscopic markers is significantly lower for partial villous atrophy (58%) than for subtotal or total villous atrophy (82%) [26]. Consequently, systems

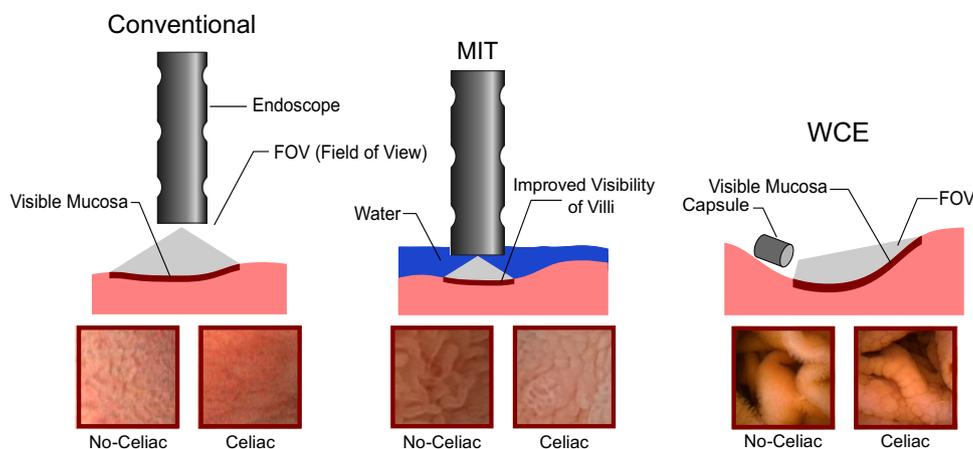


Fig. 1. Conventional endoscopic imaging, the modified immersion technique (MIT) as well as wireless capsule endoscopy (WCE).

based on visual endoscopic markers might be subject to error in diagnosis of CD due to missing features even in the high-risk population. CLE is a technique that could potentially be used for classification in accordance to the histologically described changes of the mucosal tissue instead of relying on visual endoscopic markers with low sensitivity.

Effort has been put into the identification of temporal features in CD such as the variations in brightness and texture [22,16,18,21] as well as changes in the duodenal wall motility [17]. This sort of features were mainly used in WCE and have shown promising results. They have not yet been evaluated in a large clinical study yet.

Fig. 3 illustrates the most prevalent endoscopic markers captured during flexible endoscopy.

2.2. Characteristics of endoscopic techniques

The natural features of endoscopic techniques used in the diagnosis of CD call for different approaches towards automated diagnosis. Data from flexible endoscopy is characterized by high image resolution but rapid non-monotonic movement. As a consequence, spatio-temporal features are rather inappropriate for this type of data and most work in this context was done using the approach from a more classical texture classification perspective. CLE combines microscopy with standard endoscopy, providing the highest image resolution of all endoscopic methods. Although classic endoscopic degradations are not expected during endomicroscopy, targeting of suspect areas for inspection is a vital part of the diagnosis chain and comparable to flexible endoscopy. Consequently, a significant number of challenges are shared between flexible endoscopy and CLE. The most significant difference of those two techniques compared to WCE is the possibility of interaction with a clinician during the procedure.

Data captured during WCE is characterized by slower, monotonic movement and lower resolutions. This is resulting from the passive propulsion of the camera through the intestine caused by peristalsis. Due to the nature of the data, spatio-temporal features are a much more promising concept as compared to flexible endoscopy and CLE. The missing possibility of interaction and the large amount of recorded data requires different concepts for computer-assisted systems as compared to flexible endoscopy and CLE. In WCE, automated systems focused on reducing the workload of a manual analysis and interpretation by clinicians and providing additional, discriminative statistics for diagnosis seem to be most valuable.

The properties of flexible endoscopy, CLE and WCE are quite divergent. Consequently, it is unlikely that methods developed for one endoscopic technique can be used in the other in a straightforward manner. Table 1 gives a summary of the most relevant properties of each endoscopic methodology for computer-assisted systems.

3. Image representations used in automated diagnosis of CD

The nature of data provided by endoscopy requires systems for computer-assisted diagnosis to perform visual analysis. Consequently, the main focus of research found in the literature is focused on methods for the visual classification of duodenal tissue.

3.1. Materials used in experimental studies

The unconstrained gastrointestinal environment provides a challenging scenario for visual analysis. Endoscopic image degradations such as blur, bubbles, specular reflections, non-opaque fluids as well as rapidly changing viewpoints and camera-distances complicate a systematic analysis of developed methodologies. To avoid unwanted side-effects caused by such degradations, current research focused on

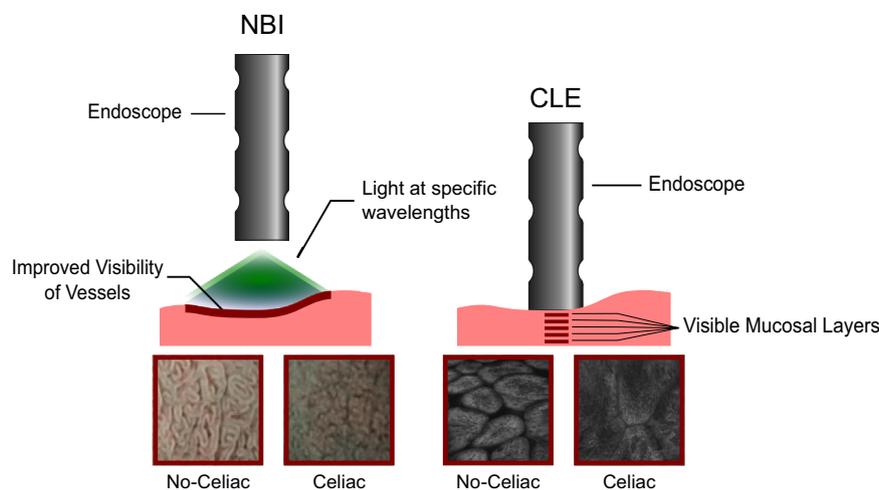


Fig. 2. Narrow band imaging (NBI) and confocal laser endomicroscopy (CLE).

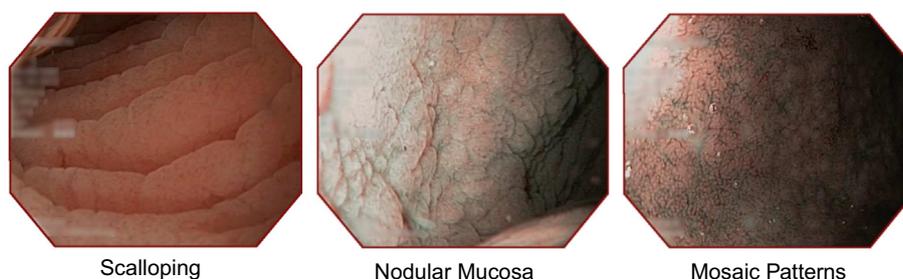


Fig. 3. The most prevalent endoscopic markers in CD.

the visual classification of intestinal mucosa affected by CD is based on a defined set of (possibly unrealistic) constraints. Due to the lack of publicly available data-sets experimental evaluation and development is performed using specific private data-sets.

WCE: The data used during experimentation and consequently all reported results in WCE [22] are based on a rather small number of patients (approximately 20). The sequences were captured in a follow up examination after the diagnosis of CD had been confirmed using standard flexible endoscopy including biopsies. All patients with positive histology (Marsh–Oberhuber) and serology were on a GFD. The authors report, that these patients are still considered to have active CD due to the fact that a period of multiple months is needed for the diet to cause a reduction in small intestinal villous atrophy. The entire sequences were interpreted by three experienced gastroenterologists and selected video clips with 200 image frames in length were exported as 576×576 pixel images.

Flexible endoscopy: The data used during experimentation in flexible endoscopy varied between publications. Initial work [95,94] was based on a rather small amount of data (approximately 200 images of the duodenal bulb and approximately 200 images of the second part of the duodenum) recorded with the conventional image capturing technique. Hegenbart et al. [52] compared MIT with the conventional technique and found MIT to be superior for computer-assisted diagnosis. Subsequently, all work in flexible endoscopy was based on images captured using MIT. The amount of data that was used was reported as 600 images from 170 patients in 2011 [93,56] and 1050 images with 320 patients in 2014 [67]. The results reported in this section are all based on comparable data-sets from 2011.

Table 1
Relevant properties of endoscopic techniques for computer-assisted diagnosis.

Property	Flexible endoscopy	WCE	CLE
Movement	Fast	Slow	None
Resolution	High	Low	High
Interactive	Yes	No	Yes
Modalities	Multiple	Single	Single
Field of view	120–170°	140–170°	500 × 500 μm

The flexible endoscopy data-set is based on video data captured during standard endoscopy (the majority using the MIT) in children suffering from signs and symptoms making upper

Table 2
Accuracies (OCR) of feature representations used in automated diagnosis of CD.

Flexible endoscopy			
Spatial Size Distribution	90	LDB	79.5–82.5
Wavelet-based LBP	88	MRF	78.5–80.5
Local Binary Patterns	83–86	Correlation signatures	82
Shape-Curvature-Histogram	85–87	Fourier statistics	82
Local Fractal Dimension - MR8	92	Multi-Fractal-Spectrum	89
Multiscale Blob Features	86	Dense SIFT	83.5
WCE			
Basis Image Statistics	71	Local Image Statistics	65
Volumetric statistics	64	Wall Motility	59
CLE			
Pyramidal LBP	93		

endoscopy necessary. The severity of the disease was classified based on histological findings (Marsh–Oberhuber) in children with positive CD serology. During the procedure, still images of areas with specific visual manifestations of CD or the absence of the disease were captured by the endoscopist. The full image frames were later manually segmented by a domain expert into possibly multiple, non-overlapping regions of size 128×128 pixels [52].

CLE: The CLE data-set used by Grisan et al. [45] in their initial work is still of rather small size (128 images from 30 patients). Subjects were recruited from a population with known CD, suspected CD and controls. CLE images and forceps biopsies of the same site were then taken sequentially at standardized locations.

A variety of different classification methodologies, including variations of support vector machines (SVM), k -nearest neighbor classifiers (k NN), Bayes classifiers and random forests were used in this context. According to the published literature, it seems clear that the most significant impact on the accuracy of automated diagnosis lies within the proposed image representations however. Consequently, the focus of the presented methodologies in the following subsections lies on the feature extraction methods used for classification of CD. Due to different data-sets and evaluation methodologies used across the literature, it is difficult to compare the obtained results. Consequently, herein we will only report the results of methods that were evaluated in a comparable fashion. The presented experiments were based on the data described in this section. Note that the data-sets used in experimentation was approximately balanced between cases and controls (in terms of samples) in this experiments. For the sake of consistency we present the classification accuracy (denoted as overall classification rate (OCR)) of each paper in the following section. Table 2 presents a summary of the reported results.

3.2. Spatial domain features

In flexible endoscopy, special attention was paid to image representations in the spatial-domain. A method particularly developed for classification of CD is the Shape-Curvature-Histogram (SCH) [34,36] (OCR 85–87) which is a robust shape based feature, describing local curvature along edges and is supposed to capture the shape of duodenal villi.

Ciacco [15] explicitly measured the fissure length in endoscopic images in a semi-automatic system (OCR 64). Based on manually selected sub-images, morphological skeletonization was performed. The total length of the skeletonized fissures was then used as one-dimensional feature.

Other, more general purpose image representations that were applied include the Edge Co-Occurrence Matrix [81,34] (OCR 86) describing edge information, the classical Haralick features [51,34] computed from gray-level Co-Occurrence Matrices (OCR 87) and features based on the autocorrelation between an image and a morphologically transformed version of the same image using the Distribution of Spatial Size (SSD) [2,34] (OCR 90).

A very promising class of spatial image representations was based on Local Binary Patterns (LBP [78]). LBP represents images as distributions of pixel neighborhoods, encoded as binary patterns and has been shown [55,53,58,41,43,45] to be very robust and accurate in classification of CD (OCR 83).

Various specialized LBP-variations have been developed for computer-assisted diagnosis of CD. Vécsei et al. [93] proposed two new LBP-based methodologies, particularly optimized for endoscopic environments. The WT-LBP (OCR 88) method was designed to combine different wavelet subbands with appropriate LBP-based operators, while the ELTP (OCR 86) method combines the benefits of the robust Local Ternary Patterns (LTP [86]) with the highly discriminative Extended Local Binary Patterns (ELBP [63]) using an adaptive thresholding, based on local image statistics.

Explicit frequency filtering using appropriate cut-off frequencies in combination with LBP-based image representations [53] has been explored (OCR 83-86) as well. Gadermayr et al. [41] employ LBP based on an adaptive neighborhood for implicit distortion correction while Hegenbart et al. [58] compute a scale- and orientation-adaptive LBP-representation, providing reliable features in scenarios with highly varying orientations and scales. Finally, Grisan et al. [45] used LBP with a pyramidal image decomposition of CLE imagery (OCR 93).

Spatial domain features seem to be very well suited for the visual classification of CD. In addition to providing a high classification accuracy, the majority of methods in this category are robust in terms of distortions and fast to compute. Both of these strengths are necessary requirements for real-time applications in endoscopy.

3.3. Transform domain features

Various feature extraction methods used for classification of CD in flexible endoscopy are based on image domain transformations. A benefit of computing features in transform domains is the possibility of employing different categories of information, such as multi-scale and multi-orientation information in wavelets subbands or frequency and phase information in the Fourier domain.

Wavelet-based methods have been used frequently for the classification of CD throughout literature. The most elementary features [49,34] are based on classical statistics such as the mean, the standard deviation or entropy of subband-coefficient energies (OCR 70). Due to the low discriminative power of this type of features, more sophisticated approaches based on the wavelet packet decomposition have been employed. Using the best basis centroids method (BBC) [71,95,93,34] (OCR 77-82) as well as the local discriminant basis (LDB) algorithm [49,83,95,93,34] (OCR 79.5-82.5) the discriminative power of elementary statistical features could be further improved.

Other wavelet-based image representations that have been applied include classical Gabor statistics [74,52,93] (OCR 80) as well as features based on Gaussian Markov Random Fields (MRF) in the wavelet domain [48,93,34] (OCR 78.5-80.5).

In addition to the standard wavelet transform, the dual-tree complex wavelet transform (DT-CWT, [84]) has been applied for classification of CD. Discriminative features include correlation signatures [91,93] (OCR 82) as well as the Weibull parameters of the marginal distribution of the complex wavelet coefficients [68,52,93] (OCR 77-82).

In a complementary fashion to the wavelet transform, Fourier based features have been proposed as well. Ring shaped filters with variable width, optimized in an evolutionary framework, were used to compute elementary statistics of the frequency domain coefficients [94,52,93] (OCR 82).

Comparing the classification performance achieved by features computed in transform domains (OCR 70-82.5) with the rates of spatial features (OCR 81-90), a lower accuracy was obtained when using transform domain features. This could potentially be caused by endoscopic distortions, negatively influencing the transformation process or a generally lower discriminative power of the image representations.

3.4. Scale-invariant features

Due to highly varying camera-perspectives and scales in endoscopic sequences, image representations invariant to scaling and rotation could be superior to features affected by such transformations. Consequently, scale-invariant image representations have been applied to data from flexible endoscopy [89,60,58]. Among the most promising scale-invariant features are the Multi-Fractal-

Spectrum [99,60] (OCR 89) and Local Fractal Features based on MR8 filtering [92,60] (OCR 92). Scale- and Orientation-Adaptive LBP [58] have been specifically designed for environments with highly varying camera-scales and orientations and are another promising image representation for endoscopy.

The majority of scale-invariant methods perform similarly to non-invariant image representations however. This is true for methods based on the DT-CWT and the D³T-CWT [89,72] (OCR 66-88), dense SIFT features [79,60,58] (OCR 83.5) as well as Multiscale Blob Features [98,60] (OCR 86).

Among the worst methods in terms of classification accuracy were features based on pulse coupled neural networks such as SCM [100,60] (OCR 64) and ICM [73,60] (OCR 67) as well as slide Matching of Gabor features [33,60] (OCR 71-75) and Local Affine Regions [69,60] (OCR 71).

The highly variable performance of scale-invariant features illustrates that such image representations are not always suited for difficult environments such as in endoscopy. Specular reflections, blur or bubbles are potentially misleading the identification of key-points or the estimation of affine shapes. Even more, the scale-invariance properties of a large number of methods are based on theoretical concepts and assumptions, which rarely hold true in practice. Consequently, the majority of features specifically designed for scale-invariance perform similarly to features affected by scaling. In general, scale-invariant image representations are computationally more demanding as compared to more elementary features. Therefore, the availability of such methods could potentially be restricted to offline scenarios in computer-assisted diagnosis.

3.5. Spatio-temporal features

Spatio-temporal features are a powerful tool for analyzing sequences of images or videos such as produced by WCE. The general benefits of this type of image representation are a higher robustness in terms of extraneous endoscopic degradations and the possibility of identifying CD even if the gastrointestinal manifestations of the disease are not visualized directly in the recorded data.

Initial work [16] (OCR 65) on spatio-temporal features, computed from WCE sequences, was based on elementary local statistics (such as pixel brightness and image texture) measured over 10×10 pixel subimages and then averaged for 56×56 subimages per frame. The authors reported spatio-temporal differences in brightness and texture of celiac images as compared to controls.

In celiac patients, gastrointestinal motility abnormalities were reported [87,3], which may result from a diminished number of enterochromaffine cells producing hormones that regulate intestinal motility. Ciaccio et al. [17] use dynamic estimates of wall motility for image classification by characterizing the motility according to frame-to-frame changes in lumen shape and position (OCR 59).

Based on the assumption that the reduced gastrointestinal motility in celiac disease affects peristalsis, Ciaccio et al. [22,18,21] (OCR 71) estimate the dominant period in a sequence of WCE images as the tallest peak in the ensemble average power spectrum, which is then used to compute salient information based on a set of basis images. The results show that celiac images had more detailed texture than controls and exhibited more variation in brightness. In celiacs, correlations existed between greater textural alterations vs. longer dominant periods.

Shape-from-shading is a technique for the reconstruction of the three dimensional structure of an object, based on illumination information. In [20] (OCR 64), shape-from-shading was applied to reconstruct the duodenal structure from WCE images for an analysis of luminal macro-architecture. Based on a syntactical analysis, protrusions were identified and used to compute volumetric statistics. Significant differences of these statistics between celiac patients and controls were reported.

Due to the relatively low accuracy but complementary characteristic of the presented features, a polling protocol for predicting CD in WCE analysis was presented, based on a majority voting scheme using a variety of features [19]. The experiments showed that a combination of complementary, low-discriminative feature can be used to construct an accurate classification system (OCR 88). This is impressive because no manual segmentation of WCE sequences was performed and endoscopic image degradations were handled implicitly. As a consequence of the small number of used sequences during experimentation, however the general performance of this system on independent data is still unclear.

4. The gastrointestinal-tract: a challenging environment

The gastrointestinal tract is a very challenging environment for computer-assisted diagnosis. As mentioned in Section 3, the majority of methods for automated diagnosis were developed based on constrained data-sets, allowing systematic and objective evaluations. Obviously, this is not a realistic scenario. As a consequence, the impact of common image degradations as well as concepts for handling challenges provided by the endoscopic environment was specifically studied in a significant number of publications.

The appearance of duodenal tissue is highly dependent on extrinsic camera parameters such as distance and angle towards the mucosa. As a combined effect of the geometric properties of the gastrointestinal tract with the large field of view of optical systems used in endoscopic hardware, certain areas of the visualized tissue are generally blurred or underexposed and noisy. The insufflation and suction of air during flexible endoscopy as well as the instillation of water into the lumen can lead to a significant amount of bubbles on the duodenal walls. Fluids on the mucosal surface can potentially cause specular reflections of the illuminating point light sources attached to the endoscope. These types of image degradations are likely to affect the feature extraction process and the performance of a computer-assisted system for diagnosis.

Due to the challenging environment, methods for informative frame identification and segmentation are generally applied across the literature in related fields of endoscopic image processing. Most of the proposed methodologies, however, are tailored toward specific applications or endoscopic scenarios and cannot be used in a straightforward manner for CD diagnosis.

An approach combining informative frame detection and segmentation in classification of CD was proposed by Gadermayr et al. [42]. The authors combine multiple elementary quality measures (brightness, contrast, blur, noise and reflections) to perform an automated segmentation of image frames from flexible endoscopy.

An alternative concept by Hegenbart et al. [54,57] was based on a one-class support vector machine (SVM), solely trained on celiac features for implicit handling of certain endoscopic image degradations.

In WCE, a way of handling extraneous features is based on robust spatio-temporal features computed from noise insensitive basis images [18,21].

4.1. The duodenal gang: blur, bubble, noise and reflection

It is safe to assume that endoscopic image degradations have a negative effect on the performance of computer-assisted diagnosis. The severity of these effects is unclear however. Consequently, Hegenbart et al. [54] systematically investigated the implicit handling of the most common image degradations (blur, bubble, noise and reflections) based on a one-class SVM. In an experimental evaluation using LBP-based methods for feature extraction, image degradations were simulated to allow a fine grained analysis (with corresponding ground-truth) of the individual effects on the proposed classification pipeline. Fig. 4 illustrates real and simulated endoscopic image degradations used by Hegenbart et al. [54].

The authors report that certain types of image degradations, such as bubbles and reflections, only affect a subset of LBP-based methods and can actually be compensated using the proposed approach. Blur and noise had the most impact on the accuracy of the system. The unconstrained classification of CD based on LBP using a one-class SVM is feasible to some degree. Extreme cases of image distortions might require an additional step of informative frame identification however. By relaxing the needs for informative frame identification to extreme cases, the general reliability of a fully automated system could possibly be increased.

Another, unorthodox approach is based on degradation adaptive texture classification [39]. Using a set of measures for the severity of image degradation (noise, blur, scale, illumination and contrast), the set of features used for training is adaptively adjusted in such a way that only images with a high similarity in terms of distortions to the image classified are used. The authors reported significantly improved classification accuracies (up to 5 percentage points) as compared to using the entire corpus of training data.

4.2. Gastrointestinal regions: Where am I?

Sequences in flexible endoscopy are commonly non-monotonic in terms of the passage through gastrointestinal regions. During the procedure, the endoscopic tip often traverses the stomach and the duodenum multiple times. Consequently, the visual appearance of esophageal and gastric tissue in endoscopic data is a

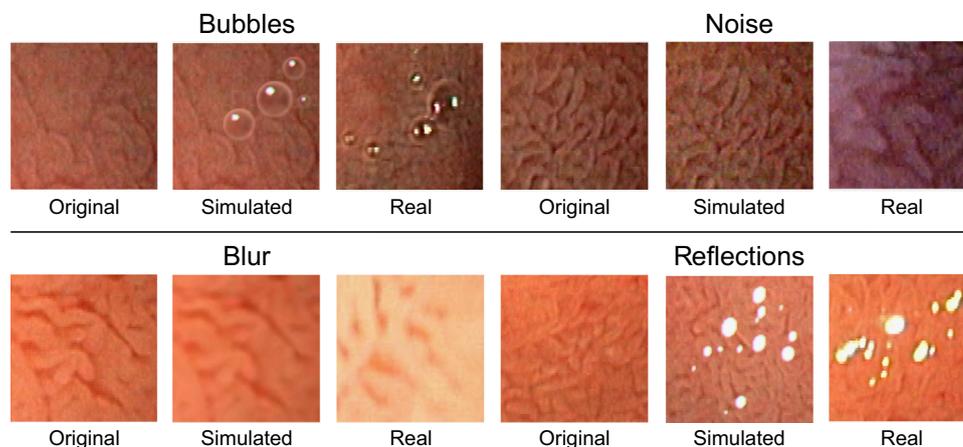


Fig. 4. Real and simulated endoscopic image degradations.

common thing. A classification model based on duodenal tissue is potentially subject to error if visualized tissue of other (unknown) gastrointestinal regions is provided. Consequently, Hegenbart et al. [57] studied if tissue visualized from other gastrointestinal regions (esophagus and stomach) would be misclassified as CD using their approach based on a one-class SVM.

The authors reported that the visualized appearance of esophageal and gastric tissue can be very close to duodenal tissue affected by villous atrophy (see Fig. 5). The experimental evaluation indicates that celiac tissue cannot be distinguished from esophageal and gastric tissue in such a classification approach. Consequently, it is likely that an additional mechanism for tracking the endoscope's location or identification of the intestinal region is required to handle non-monotonic transits and the visualization of unknown gastrointestinal regions in a fully automated system.

4.3. Extrinsic camera parameters: high variance

Highly varying extrinsic camera parameters such as distance and angle to the mucosa can potentially lead to a blurred visualization of the tissue (close distance) or to a missing visualization of small spatial structures such as villi (far distance) and could therefore be inappropriate for visual classification (see Fig. 6).

The effects of varying camera-scales were explicitly studied by Hegenbart et al. [57]. Experimental data indicated that the visualization of duodenal tissue at close and far distances is inappropriate for classification with features affected by scaling such as LBP.

Consequently, scale- and viewpoint-invariant methods have been employed [60] for the classification of CD. In this work, the authors showed in a comprehensive study that scale-invariant image representations often exhibit an intrinsically lower discriminative power as compared to features affected by scaling. Consequently, scale-invariant features often do not pose a significant



Fig. 5. Visual appearance of different GI-regions.

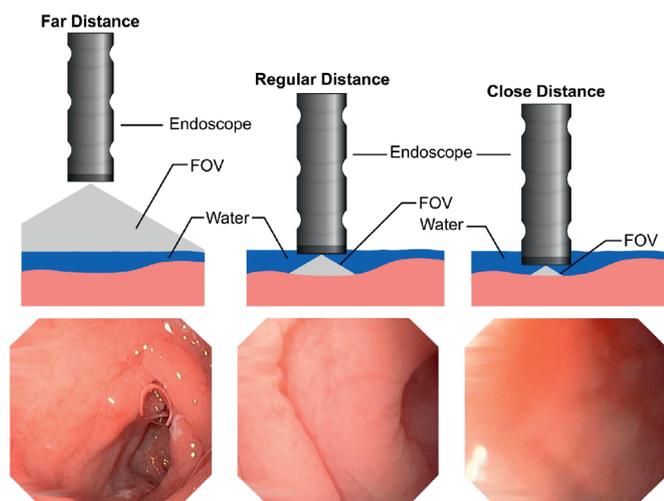


Fig. 6. Visualization of tissue at different camera distances.

benefit. A small subset of such methods (Fractal Dimension and affine-adaptive LBP) were promising however.

Recently, a scale- and rotation-invariant image representation based on highly discriminative LBP (SOA-LBP) has been proposed, particularly for scenarios with significantly varying camera-scales [58] such as endoscopy.

4.4. Interlacing

Interlaced scanning is a technique that has been widely used in video systems to double the perceived frame rate without increasing the required bandwidth. This technique is still in use by certain endoscopic video hardware today. Various specialized de-interlacing techniques have been developed over the last decades to reconstruct full frames from two interlaced half-frames. The impact of interlaced scanning and the benefits of applying suited de-interlacing techniques for computer-assisted systems was specifically studied [59].

The authors state that de-interlacing does not have a significant positive effect on the classification accuracy of endoscopic data with indication for CD. The benefits of applying de-interlacing were comparable to the effects of Gaussian filtering. Considering the accuracy of the system, the authors could not identify significant differences between simple and more complex de-interlacing techniques. Consequently, computer-assisted systems based on interlaced data could potentially be used without an explicit need for de-interlacing.

4.5. Do not believe your eyes: lens distortion

A special type of endoscopic image degradation, present in all endoscopic images, is a consequence of the wide-angle nature of the optical systems used in endoscopic hardware. Such barrel-type distortions introduce non-linear changes in the images, leading to a significantly smaller visualization of outer areas. Fig. 7 illustrates the visual manifestation of lens distortions in endoscopic imagery.

The effects of endoscopic lens distortions as well as distortion correction on computer-assisted diagnosis have been extensively studied [46,34,43,50]. The general consensus in the literature is that the effects of lens distortion and the benefits of distortion correction are highly dependent on the used type of features. Features based on low-frequency components of the images are merely affected by lens distortion while image representations based on high-frequency components clearly benefit from distortion correction. The largest improvements reported were in the range of 5–6 percentage points. In some cases the classification accuracy dropped by approximately 9 percentage points however.

Interpolation artifacts and varying camera-scales were identified as the main restrictions of distortion correction performance [35]. Consequently, distortion adaptive classification [38,40], distortion compensated features based on the Fourier transform [37] as well as intrinsic distortion correction [41] have been proposed.

A particular application for lens distortion correction is an improved visualization for clinicians during endoscopy [43]. It was shown in a study involving several gastroenterologists that the accuracy of diagnosis by domain experts using such a system is actually negatively affected. This is most likely a consequence of their medical training based on distorted imagery.

5. Evaluation of methods in medical image classification

The nature of medical data differs from the characteristics of data used in more classical texture classification scenarios. As a consequence of a potentially small number of available patients and an uneven distributions of healthy and unhealthy individuals

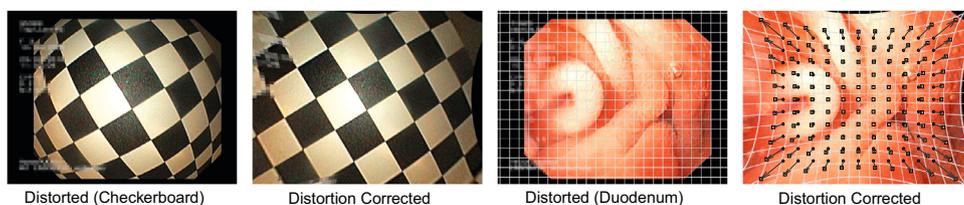


Fig. 7. Lens distortion and distortion correction applied to endoscopic images.

undergoing treatment, multiple samples per patient are often used to build balanced data sets of sufficient size. This often results in an intrinsic bias of data sets used in medical image classification.

To deal with restricted amounts of data available for experimentation, cross-validation techniques are used for predicting how well developed methodologies will generalize. Particularly in medical image classification, assumptions of such techniques are often violated by the inherent intrinsic bias of the created data-sets. The performance predicted by cross-validation is consequently subject to error.

Methods for feature extraction or classification are usually parametrized by a set of values such as thresholds, number of neighbors or filter coefficients. The optimal range of values of such parameters is often unclear from a theoretical standpoint and usually determined empirically. Consequently, parameter optimization is often performed in the process of development or evaluation. In a combination with cross-validation, optimization can be dangerous and can potentially lead to over-fitting, reducing the accuracy of performance prediction significantly.

A common flaw in the evaluation of methods is based on using the predicted classification accuracy of the cross-validation as the objective function for optimization. Hegenbart et al. [56] reported prediction errors when using this strategy of up to 7.5 percentage points (OCR 92 vs. OCR 84.5). The authors suggest to use nested cross-validations to avoid such errors, optimizing parameters in a separate cross-validation based on the current training fold.

Due to violated assumptions of cross-validation techniques, caused by the intrinsic bias of medical image sets, errors of up to 7 percentage points were reported (Leave-One-Out Cross Validation: OCR 90; Distinct Sets: OCR 83). The violation of assumptions in cross-validation techniques can be avoided by using a data partition based on patients instead of images.

The prediction error of nested-cross validations using patient-based partitions was reported to be only 1.4 percentage points as compared to the most realistic scenarios with two independent data-sets. Although nested-cross validations add a significant amount of computational complexity to the experimental setup, such schemes are mandatory for realistic predictions if optimization is performed.

5.1. Staging schemes and noisy ground-truths

In a revision of the Marsh-scheme for staging the severity of CD, Oberhuber et al. [77] classified Marsh type 3 lesions into three sub-groups. Unfortunately, the definitions of type Marsh-3A (mild or moderate villous shortening) and Marsh-3B (marked villous shortening) lack objective criteria, contributing to an increased intra- and inter-observer variability [31,76,1].

Only four publications [93,60,15,43] on computer-assisted diagnosis of CD are focused on a reduced 4-class Marsh-like scenario (including classes Marsh-0, Marsh-3A, Marsh-3B and Marsh-3C). The authors reported classification rates between 50 and 68% and point out problems with lacking distinction of visual appearance between classes of type Marsh-3. Vécsei et al. [93] suggest to use more simplified, visually focused systems such as proposed by Corazza and Villanaci [24] or Ensari [31] for automated diagnosis. Consequently, the majority of work throughout the literature has

been focused on the clinically most relevant two-class case (celiac disease vs. healthy).

As a side-effect of the potentially high intra- and inter-observer variability, the histological Marsh-Oberhuber based ground-truth used for developing computer-assisted systems is subject to a certain amount of noise. As a consequence of the patchy distribution of CD, discrepancies between areas with known histology and images used for developing computer-assisted systems could potentially exist.

A noisy ground-truth as basis for development and training of a computer-based system could be problematic and effectively reduce the accuracy of and confidence in the automated diagnosis. Recently, Kwitt et al. [67] were able to show that existing noise in ground-truths can be compensated by a sufficiently large corpus of training data. In a crowd-sourcing context, endoscopic data with indication for CD were labeled by non-experts and used to create a classification system that performed comparably to a system trained on data with histological ground-truth (and therefore a substantial smaller amount of label noise). This finding consequently indicates that a moderate amount of label noise (as present in histological ground-truths) can actually be compensated implicitly by a classification system based on a training corpus of sufficiently large size.

6. Conclusion

Since the first publication on computer-assisted diagnosis of CD in 2008, a lot of progress towards a fully automated system has been made. A large variety of image representations has been studied and a solid knowledge of the behavior of methods in an endoscopic environments has been acquired. Based on constrained data-sets from flexible endoscopy, CD can currently be classified with an accuracy of up to 92%. Based on a majority voting using multiple complementary spatio-temporal features, classification accuracies in sequences of WCE currently reach 88%. The small number of sequences used for evaluation, however, still leaves open questions about the general performance of this methodology. Recently reported results by Grisan et al. using data from CLE indicate that this imaging modality could potentially be promising for automated diagnosis (OCR 93). Again, the small data-set size in CLE is problematic.

Significant effort was put into the study and development of concepts for handling various challenges in endoscopic imagery such as image degradations, heavily varying extrinsic camera-parameters, interlacing and lens distortion. As an alternative approach to informative frame detection, a classification pipeline based on a one-class SVM has been proposed for the implicit handling of the most common endoscopic image degradations.

Robust features computed from degradation insensitive basis images have been used to implicitly handle extraneous features in WCE and in an unorthodox approach, degradation adaptive classification has been successfully performed.

Lens distortion has been shown to affect the accuracy of automated diagnosis. Consequently, distortion adaptive classification, distortion compensated features as well as intrinsic distortion correction have been used to improve the accuracy of classification.

A solid understanding of issues and common flaws in the evaluation of methods in a medical context has been gained. It was shown that medical data in particular has to be treated with extreme care if cross-validation is used in combination with any type of parameter optimization. Empirical data indicates that a moderate amount of label noise in the ground-truth can be compensated if a large amount of training data is available.

Although a big step towards the establishment of a fully automated system has been made, there are still a lot of challenges to be solved before a system for computer-assisted diagnosis can be used in clinical routine. One of the most stressing issues, especially in flexible endoscopy and CLE, is the development and establishment of robust and accurate methodologies for informative frame detection and segmentation. Although the implicit handling of various endoscopic image degradations is possible to some degree, images affected by blur and large camera distances are not very well handled in such an approach.

Combining spatio-temporal features with highly discriminative texture features as used in still image classification is another challenging but promising concept. It may be possible that other highly significant celiac markers could be integrated into such an approach, potentially improving the general accuracy and robustness of the systems.

Finally, it should be a general ambition to provide publicly available data-sets of CD. We are highly convinced that the research progress and interest on computer-assisted diagnosis of CD would thrive if public data-sets were available to fellow researchers.

Conflict of interest statement

None declared.

Acknowledgments

This work has been supported, in part, by the Austrian Science Fund (Project no. 24366).

References

- [1] C. Arguelles-Grande, C. Tennyson, S. Lewis, P. Green, G. Bhagat, Variability in small bowel histopathology reporting between different pathology practice settings: impact on the diagnosis of coeliac disease, *J. Clin. Pathol.* 65 (3) (2012) 242–247.
- [2] G. Ayala, J. Domingo, Spatial size distributions: applications to shape and texture analysis, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (12) (2001) 1430–1442.
- [3] G. Bassotti, G. Castellucci, C. Betti, C. Fusaro, M. Cavalletti, A. Bertotto, F. Spinozzi, A. Morelli, M. Pelli, Abnormal gastrointestinal motility in patients with celiac sprue, *Dig. Dis. Sci.* 39 (9) (1994) 1947–1954.
- [4] F. Biagi, G. Corazza, F. Biagi, G. Corazza, Mortality in celiac disease, *Nat. Rev. Gastroenterol. Hepatol.* 7 (3) (2010) 158–162.
- [5] M. Bonamico, P. Mariani, E. Thanasi, M. Ferri, R. Nenna, C. Tiberti, M. Mazzilli, F. Magliocca, Patchy villous atrophy of the duodenum in childhood celiac disease, *J. Pediatr. Gastroenterol. Nutr.* 38 (2) (2004) 204–207.
- [6] J. Brow, F. Parker, W. Weinstein, C. Rubin, The small intestinal mucosa in dermatitis herpetiformis. I. Severity and distribution of the small intestinal lesion and associated malabsorption, *Gastroenterology* 60 (3) (1971) 355–361.
- [7] G. Cammarota, A. Cazzato, O. Genovese, A. Pantanella, G. Ianiro, V. Giorgio, Water-immersion technique during standard upper endoscopy may be useful to drive the biopsy sampling of duodenal mucosa in children with celiac disease, *J. Pediatr. Gastroenterol. Nutr.* 49 (4) (2009) 411–416.
- [8] G. Cammarota, P. Cesaro, A. Martino, G. Zuccala, R. Cianci, E. Nista, L.M. Larocca, F.M. Vecchio, A. Gasbarrini, G. Gasbarrini, High accuracy and cost-effectiveness of a biopsy-avoiding endoscopic approach in diagnosing coeliac disease, *Aliment Pharmacol. Ther.* 23 (1) (2006) 61–69.
- [9] G. Cammarota, L. Cuoco, P. Cesaro, L. Santoro, A. Cazzato, M. Montalto, 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.
- [10] G. Cammarota, G. Pirozzi, A. Martino, G. Zuccala, R. Cianci, L. Cuoco, V. Ojetti, M. Landriscina, M. Montalto, F.M. Vecchio, G. Gasbarrini, A. Gasbarrini, Reliability of the “immersion technique” during routine upper endoscopy for detection of abnormalities of duodenal villi in patients with dyspepsia, *Gastrointest. Endosc.* 60 (2) (2004) 223–228.
- [11] A. Carlsson, D. Agardh, S. Borulf, E. Grodzinsky, I. Axelsson, S. Ivarsson, Prevalence of celiac disease: before and after a national change in feeding recommendations, *Scand. J. Gastroenterol.* 41 (5) (2006) 553–558.
- [12] L. Castano, R. Bharduni, L. Ortiz, J. Nunez, J. Bilbao, I. Rica, P. Martul, J.C. Vitoria, Prospective population screening for celiac disease: high prevalence in the first 3 years of life, *J. Pediatr. Gastroenterol. Nutr.* 39 (1) (2004) 80–84.
- [13] C. Catassi, I. Ratsch, E. Fabiani, M. Rossini, F. Bordicchia, F. Candela, Coeliac disease in the year 2000: exploring the iceberg, *Lancet* 343 (8891) (1994) 200–203.
- [14] N. Chand, A.A. Mihas, Celiac disease: current concepts in diagnosis and treatment, *J. Clin. Gastroenterol.* 40 (January (1)) (2006) 3–14.
- [15] E.J. Ciaccio, G. Bhagat, C.A. Tennyson, S.K. Lewis, L. Hernandez, P.H. Green, Quantitative assessment of endoscopic images for degree of villous atrophy in celiac disease, *Digest. Dis. Sci.* 56 (3) (2011) 805–811.
- [16] E.J. Ciaccio, C.A. Tennyson, G. Bhagat, S.K. Lewis, P.H. Green, Classification of videocapsule endoscopy image patterns: comparative analysis between patients with celiac disease and normal individuals, *BioMed. Eng. Online.* 9 (2010) 44–55.
- [17] E.J. Ciaccio, C.A. Tennyson, G. Bhagat, S.K. Lewis, P.H. Green, Quantitative estimates of motility from videocapsule endoscopy are useful to discern celiac patients from controls, *Digest. Dis. Sci.* 57 (11) (2012) 2936–2943.
- [18] E.J. Ciaccio, C.A. Tennyson, G. Bhagat, S.K. Lewis, P.H. Green, Transformation of videocapsule images to detect small bowel mucosal differences in celiac versus control patients, *Comput. Methods Progr. Biomed.* 108 (1) (2012) 28–37.
- [19] E.J. Ciaccio, C.A. Tennyson, G. Bhagat, S.K. Lewis, P.H. Green, Implementation of a polling protocol for predicting celiac disease in videocapsule analysis, *World J. Gastrointest. Endosc.* 5 (7) (2013) 313–322.
- [20] E.J. Ciaccio, C.A. Tennyson, G. Bhagat, S.K. Lewis, P.H. Green, Use of shape-from-shading to estimate three-dimensional architecture in the small intestinal lumen of celiac and control patients, *Comput. Methods Progr. Biomed.* 111 (3) (2013) 676–684.
- [21] E.J. Ciaccio, C.A. Tennyson, G. Bhagat, S.K. Lewis, P.H. Green, Use of basis images for detection and classification of celiac disease, *Bio-Med. Mater. Eng.* 24 (6) (2014) 1913–1923.
- [22] E.J. Ciaccio, C.A. Tennyson, S.K. Lewis, S. Krishnareddy, G. Bhagat, P.H. Green, Distinguishing patients with celiac disease by quantitative analysis of videocapsule endoscopy images, *Comput. Methods Progr. Biomed.* 100 (2010) 39–48.
- [23] G. Corazza, V. Villanacci, C. Zambelli, M. Milione, O. Luinetti, C. Vindigni, C. Chioda, L. Albarello, D. Bartolini, F. Donato, Comparison of the interobserver reproducibility with different histologic criteria used in celiac disease, *Clin. Gastroenterol. Hepatol.* 5 (7) (2007) 838–843.
- [24] G. Corazza, V. Villanacci, Coeliac disease, *J. Clin. Pathol.* 58 (6) (2005) 573–574.
- [25] W. Dickey, D. Hughes, Prevalence of celiac disease and its endoscopic markers among patients having routine upper gastrointestinal endoscopy, *Am. J. Gastroenterol.* 94 (8) (1999) 2182–2186.
- [26] W. Dickey, D. Hughes, Disappointing sensitivity of endoscopic markers for villous atrophy in a high-risk population: implications for celiac disease diagnosis during routine endoscopy, *Am. J. Gastroenterol.* 96 (2001) 2126–2128.
- [27] W. Dieterich, T. Ehnis, M. Bauer, P. Donner, U. Volta, E. Riecken, Identification of tissue transglutaminase as the autoantigen of celiac disease, *Nat. Med.* 3 (7) (1997) 797–801.
- [28] C. Dube, A. Rostom, R. Sy, A. Cranney, N. Saloojee, C. Garrity, M. Sampson, L. Zhang, F. Yazdi, V. Mamaladze, I. Pan, J. Macneil, D. Mack, D. Patel, D. Moher, The prevalence of celiac disease in average-risk and at-risk western european populations: a systematic review, *Gastroenterology* 128 (4 Suppl. 1) (2005) S57–67.
- [29] S. Dydensborg, P. Toftedal, M. Biaggi, S. Lillevang, D. Hansen, S. Husby, Increasing prevalence of coeliac disease in denmark: a linkage study combining national registries, *Acta Paediatr.* 101 (2) (2012) 179–194.
- [30] F. Emura, Y. Saito, H. Ikematsu, Narrow-band imaging optical chromocolonoscopy: advantages and limitations, *World J. Gastroenterol.* 14 (August 31) (2008) 4867–4872.
- [31] A. Ensari, Gluten-sensitive enteropathy (celiac disease): controversies in diagnosis and classification, *Arch. Pathol. Lab. Med.* 134 (2010) 826–836.
- [32] A. Fasano, I. Berti, T. Gerarduzzi, T. Not, R. Colletti, S. Drago, Y. Elitsur, P.H. Green, S. Guandalini, I.D. Hill, M. Pietzak, A. Ventura, M. Thorpe, D. Kryszak, F. Fornaroli, S.S. Wasserman, J.A. Murray, K. Horvath, Prevalence of celiac disease in at-risk and not-at-risk groups in the united states: a large multicenter study, *Arch. Intern. Med.* 163 (3) (2003) 286–292.
- [33] K. Fung, K. Lam, Rotation- and scale-invariant texture classification using slide matching of the Gabor feature, in: *ISAPACS*, January 2009, pp. 521–524.
- [34] M. Gadermayr, M. Liedlgruber, A. Uhl, A. Vécsei, Evaluation of different distortion correction methods and interpolation techniques for an automated classification of celiac disease, *Comput. Methods Progr. Biomed.* 112 (3) (2013) 694–712.
- [35] M. Gadermayr, M. Liedlgruber, A. Uhl, A. Vécsei, Problems in distortion corrected texture classification and the impact of scale and interpolation, in: *Image Analysis and Processing, Lecture Notes in Computer Science*, vol. 8156, ICIAP 2013, Springer, Berlin, Heidelberg, 2013, pp. 513–522.

- [36] M. Gadermayr, M. Liedlgruber, A. Uhl, A. Vécsei, Shape curvature histogram: a shape feature for celiac disease diagnosis, in: *Medical Computer Vision, Large Data in Medical Imaging, Lecture Notes in Computer Science*, Springer International Publishing, 2014, pp. 175–184.
- [37] M. Gadermayr, A. Uhl, A. Vécsei, Barrel-type distortion compensated fourier feature extraction, in: *Advances in Visual Computing, Lecture Notes in Computer Science*, vol. 8033, Springer, Berlin, Heidelberg, 2013, pp. 50–59.
- [38] M. Gadermayr, A. Uhl, A. Vécsei, Distortion adaptive image classification – an alternative to barrel-type distortion correction, in: *Advances in Visual Computing, Lecture Notes in Computer Science*, vol. 8034, Springer, Berlin, Heidelberg, 2013, pp. 465–474.
- [39] M. Gadermayr, A. Uhl, A. Vécsei, Degradation adaptive texture classification: a case study in celiac disease diagnosis brings new insight, in: *Proceedings of the International Conference on Image Analysis and Recognition (ICIAR'14)*, Lecture Notes in Computer Science, vol. 8815, Springer, 2014, pp. 263–273.
- [40] M. Gadermayr, A. Uhl, A. Vécsei, The effect of endoscopic lens distortion correction on physicians' diagnosis performance, in: *Bildverarbeitung für die Medizin 2014, Informatik aktuell*, Springer, Berlin, Heidelberg, 2014, pp. 174–179.
- [41] M. Gadermayr, A. Uhl, A. Vécsei, Feature extraction with intrinsic distortion correction in celiac disease imagery: no need for rasterization, in: *Medical Computer Vision. Large Data in Medical Imaging, Lecture Notes in Computer Science*, Springer International Publishing, 2014, pp. 196–204.
- [42] M. Gadermayr, A. Uhl, A. Vécsei, Getting one step closer to fully automatized celiac disease diagnosis, in: *Proceedings of the 4th IEEE International Conference on Image Processing Theory, Tools and Applications (IPTA'14)*, October 2014, pp. 13–17.
- [43] M. Gadermayr, A. Uhl, A. Vécsei, Is a precise distortion estimation needed for computer aided celiac disease diagnosis?, in: *Image and Signal Processing, Lecture Notes in Computer Science*, vol. 8509, Springer International Publishing, 2014, pp. 620–628.
- [44] J. Godfrey, T. Brantner, W. Brinjkij, K. Christensen, D. Brogan, C. Van Dyke, *Morbidity and mortality among older individuals with undiagnosed celiac disease, Gastroenterology 7 (3) (2010) 763–769.*
- [45] E. Grisan, H. Mirzaei, R. Leong, Computer-assisted automated image recognition of celiac disease using confocal endomicroscopy, in: *ISBI*, 2014, pp. 121–124.
- [46] M. Gschwandtner, M. Liedlgruber, A. Uhl, A. Vécsei, Experimental study on the impact of endoscope distortion correction on computer-assisted celiac disease diagnosis, in: *10th IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB)*, November 2010, pp. 1–6.
- [47] M. Hadjivassiliou, D. Sanders, N. Woodroffe, C. Williamson, R. Grunewald, *Gluten ataxia, Cerebellum 7 (3) (2008) 494–498.*
- [48] 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: *16th International Conference on Digital Signal Processing, July 2009*, pp. 1–6.
- [49] 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: *3rd International Conference on Advances in Medical, Signal and Information Processing, MEDSIP 2006, July 2006*, pp. 1–4.
- [50] 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)*, Lecture Notes in Computer Science, vol. 7512, Springer, Berlin, Heidelberg, 2012, pp. 574–581.
- [51] R. Haralick, K. Shanmugam, I. Dinstein, *Textural features for image classification, IEEE Trans. Syst., Man Cybern. SMC-3 (6) (1973) 610.*
- [52] 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 6th International Symposium on Image and Signal Processing and Analysis, (ISPA'09)*, 2009, pp. 718–723.
- [53] S. Hegenbart, S. Maimone, A. Uhl, A. Vécsei, G. Wimmer, Customised frequency pre-filtering in a local binary pattern-based classification of gastrointestinal images, in: *Medical Content-Based Retrieval for Clinical Decision Support, Lecture Notes in Computer Science*, vol. 7723, Springer, Berlin, Heidelberg, 2013, pp. 99–109.
- [54] S. Hegenbart, A. Uhl, A. Vécsei, Impact of endoscopic image degradations on lbp based features using one-class svm for classification of celiac disease, in: *Proceedings of the 7th International Symposium on Image and Signal Processing and Analysis (ISPA'11)*, Dubrovnik, Croatia, 2011, pp. 715–720.
- [55] S. Hegenbart, A. Uhl, A. Vécsei, Impact of histogram subset selection on classification using multi-scale lbp-operators, in: *Bildverarbeitung für die Medizin 2011, Informatik aktuell*, Springer, Berlin, Heidelberg, 2011, pp. 359–363.
- [56] S. Hegenbart, A. Uhl, A. Vécsei, Systematic assessment of performance prediction techniques in medical image classification, a case study on celiac disease, in: *Information Processing in Medical Imaging, Lecture Notes in Computer Science*, vol. 7512, 2011, pp. 498–509.
- [57] S. Hegenbart, A. Uhl, A. Vécsei, On the implicit handling of varying distances and gastrointestinal regions in endoscopic video sequences with indication for celiac disease, in: *Proceedings of the IEEE International Symposium on Computer-Based Medical Systems (CBMS'12)*, 2012, pp. 1–6.
- [58] S. Hegenbart, A. Uhl, A. Vécsei, A Scale- and Orientation-Adaptive Extension of Local Binary Patterns, Technical Report 2014-05, Department of Computer Sciences, University of Salzburg, Austria, 2014, (<http://uni-salzburg.at/index.php?id=38565>).
- [59] S. Hegenbart, A. Uhl, A. Vécsei, G. Wimmer, On the effects of de-interlacing on the classification accuracy of interlaced endoscopic videos with indication for celiac disease, in: *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems (CBMS'13)*, 2013, pp. 137–142.
- [60] S. Hegenbart, A. Uhl, A. Vécsei, G. Wimmer, *Scale invariant texture descriptors for classifying celiac disease, Med. Image Anal. 17 (4) (2013) 458–474.*
- [61] J. Henker, A. Losel, K. Conrad, T. Hirsch, W. Leupold, Prevalence of asymptomatic coeliac disease in children and adults in the dresden region of germany, *Dtsch. Med. Wochenschr.* 127 (28–29) (2002) 1511–1515.
- [62] A. Hopper, S. Cross, D. Sanders, *Patchy villous atrophy in adult patients with suspected gluten-sensitive enteropathy: is a multiple duodenal biopsy strategy appropriate? Endoscopy 40 (3) (2008) 219–224.*
- [63] X. Huang, S. Li, Y. Wang, Shape localization based on statistical method using extended local binary pattern, in: *Proceedings of the 3rd International Conference on Image and Graphics (ICIG'04)*, Hong Kong, China, 2004, pp. 1–4.
- [64] S. Husby, S. Koletzko, I. Korponay-Szabo, M. Mearin, A. Phillips, R. Shamir, R. Troncone, K. Giersiepen, D. Branski, C. Catassi, M. Lelegman, M. Mäki, C. Ribes-Koninckx, A. Ventura, K.P. Zimmer, European society for pediatric gastroenterology, hepatology, and nutrition guidelines for the diagnosis of coeliac disease, *J. Pediatr. Gastroenterol. Nutr.* 54 (1) (2012) 136–160.
- [65] A. Kalayci, Y. Kanber, A. Birinci, L. Yildiz, D. Albayrak, *The prevalence of coeliac disease as detected by screening in children with iron deficiency anaemia, Acta Paediatr. 94 (6) (2005) 678–681.*
- [66] W. Kratzer, M. Kibele, A. Akinli, M. Porzner, B. Boehm, W. Koenig, S. Oetzuerk, R. A. Mason, R. Mao, M.H. Haenle, Prevalence of celiac disease in germany: a prospective follow-up study, *World J. Gastroenterol.* 19 (17) (2013) 2612–2620.
- [67] R. Kwitt, S. Hegenbart, N. Rasiwasia, A. Vécsei, A. Uhl, Do we need annotation experts? A case study in celiac disease classification, in: *Medical Image Computing and Computer-Assisted Intervention (MICCAI'14)*, Lecture Notes in Computer Science, vol. 8674, Springer International Publishing, 2014, pp. 454–461.
- [68] R. Kwitt, A. Uhl, Modeling the marginal distributions of complex wavelet coefficient magnitudes for the classification of zoom-endoscopic images. In: *IEEE 11th International Conference on Computer Vision*, 2007, pp. 1–8.
- [69] S. Lazebnik, C. Schmid, J. Ponce, A sparse texture representation using affine-invariant regions, in: *CVPR*, vol. 2, June 2003, pp. II-319–II-324.
- [70] R. Leong, N.Q. Nguyen, C. Meredith, S. Al-Sohaily, D. Kucic, P. Delaney, E. Murr, J. Yong, N. Merret, A. Biankin, *In vivo confocal endomicroscopy in the diagnosis and evaluation of celiac disease, Gastroenterology 135 (6) (2008) 1870–1876.*
- [71] M. Liedlgruber, A. Uhl, Statistical and structural wavelet packet features for pit pattern classification in zoom-endoscopic colon images, in: *Proceedings of the 7th WSEAS International Conference on Wavelet Analysis and Multi-rate Systems, WAMUS'07, World Scientific and Engineering Academy and Society (WSEAS)*, Stevens Point, Wisconsin, USA, 2007, pp. 147–152.
- [72] E.H.S. Lo, M.R. Pickering, M.R. Frater, J.F. Arnold, Query by example using invariant features from the double dyadic dual-tree complex wavelet transform, in: *CIVR*, 2009, pp. 1–8.
- [73] Y. Ma, L. Liu, K. Zhan, Y. Wu, *Pulse coupled neural networks and one-class support vector machines for geometry invariant texture retrieval, Image Vis. Comput. 28 (11) (2010) 1524–1529.*
- [74] B. Manjunath, W. Ma, *Texture features for browsing and retrieval of image data, IEEE Trans. Pattern Anal. Mach. Intell. 18 (August (8)) (1996) 837–842.*
- [75] 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.*
- [76] A. Mubarak, P. Nikkels, R. Houwen, F. Ten Kate, Reproducibility of the histological diagnosis of celiac disease, *Scand. J. Gastroenterol.* 46 (9) (2011) 1065–1073.
- [77] G. Oberhuber, G. Granditsch, H. Vogelsang, *The histopathology of coeliac disease: time for a standardized report scheme for pathologists, Eur. J. Gastroenterol. Hepatol. 11 (10) (1999) 1185–1194.*
- [78] T. Ojala, M. Pietikäinen, D. Harwood, *A comparative study of texture measures with classification based on feature distributions, Pattern Recognit. 29 (January (1)) (1996) 51–59.*
- [79] F. Perronnin, C. Dance, Fisher kernels on visual vocabularies for image categorization, in: *CVPR*, 2007, pp. 1–8.
- [80] R. Petroniene, E. Dubcenco, J. Baker, *Given capsule endoscopy in celiac disease: evaluation of diagnostic accuracy and interobserver agreement, Am. J. Gastroenterol. 100 (3) (2005) 685–694.*
- [81] R. Rautkorpi, J. Iivarinen, A novel shape feature for image classification and retrieval, in: *Image Analysis and Recognition, Lecture Notes in Computer Science*, vol. 3211, Springer, Berlin, Heidelberg, 2004, pp. 753–760.
- [82] A. Rubio-Tapia, R. Kyle, E. Kaplan, D. Johnson, W. Page, F. Erdtmann, *Increased prevalence and mortality in undiagnosed celiac disease, Gastroenterology 137 (1) (2009) 88–93.*
- [83] N. Saito, R. Coifman, *Local discriminant bases and their applications, J. Math. Imaging Vis. 5 (4) (1995) 337–358.*
- [84] I. Selesnick, R. Baraniuk, N. Kingsbury, *The dual-tree complex wavelet transform, IEEE Signal Process. Mag. 22 (6) (2005) 123–151.*
- [85] J. Taavela, O. Koskinen, H. Huhtala, M. Lähdeaho, A. Popp, K. Laurila, P. Collin, K. Kaukinen, K. Kurppa, M. Mäki, *Validation of morphometric analyses of small-intestinal biopsy readouts in celiac disease, PLoS One 8 (10) (2013).*
- [86] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, in: *Analysis and Modelling of Faces and Gestures*, 2007, pp. 168–182.

- [87] A. Tursi, Gastrointestinal motility disturbances in celiac disease, *J. Clin. Gastroenterol.* 38 (8) (2004) 642–645.
- [88] A. Tursi, W. Elisei, G. Giorgetti, G. Brandimarte, F. Aiello, Complications in celiac disease under gluten-free diet, *Dig. Dis. Sci.* 54 (10) (2009) 2175–2182.
- [89] A. Uhl, A. Vécsei, G. Wimmer, Complex wavelet transform variants in a scale invariant classification of celiac disease, in: *Pattern Recognition and Image Analysis, Lecture Notes in Computer Science*, vol. 6669, Springer, Berlin, Heidelberg, 2011, pp. 742–749.
- [90] F. Valitutti, S. Oliva, D. Iorfida, M. Aloï, S. Gatti, C.M. Trovato, M. Montuori, A. Tiberti, S. Cucchiara, G. Di Nardo, Narrow band imaging combined with water immersion technique in the diagnosis of celiac disease, *Dig. Liver Dis.* 46 (12) (2014) 1099–1102.
- [91] G. Van de Wouwer, S. Livens, P. Scheunders, D. Van Dyck, Color texture classification by wavelet energy correlation signatures, in: A. Del Bimbo (Ed.), *Image Analysis and Processing, Lecture Notes in Computer Science*, vol. 1310, Springer, Berlin, Heidelberg, 1997, pp. 327–334.
- [92] M. Varma, A. Zissermann, A statistical approach to texture classification from single images, *Int. J. Comput. Vis.* 62 (1–2) (2005) 61–81.
- [93] A. Vécsei, G. Amann, S. Hegenbart, M. Liedlgruber, A. Uhl, Automated marsh-like classification of celiac disease in children using local texture operators, *Comput. Biol. Med.* 41 (6) (2011) 313–325.
- [94] A. Vécsei, T. Fuhrmann, M. Liedlgruber, L. Brunauer, H. Payer, A. Uhl, Automated classification of duodenal imagery in celiac disease using evolved fourier feature vectors, *Comput. Methods Progr. Biomed.* 95 (Suppl. 2) (2009) 68–78.
- [95] A. Vécsei, T. Fuhrmann, A. Uhl, Towards automated diagnosis of celiac disease by computer-assisted classification of duodenal imagery, in: *MEDSIP*, 2008, pp. 1–4.
- [96] B. Weile, B. Hansen, I. Hägerstrand, J. Hansen, P. Krasilnikoff, Interobserver variation in diagnosing coeliac disease. a joint study by danish and swedish pathologists, *Acta Pathol. Microbiol. Immunol. Scand.* 108 (5) (2000) 380–384.
- [97] W. Weinstein, J. Brow, F. Parker, C.E. Rubin, The small intestinal mucosa in dermatitis herpetiformis. II. Relationship of the small intestinal lesion to gluten, *Gastroenterology* 60 (3) (1971) 362–369.
- [98] Q. Xu, Y. Chen, Multiscale blob features for gray scale, rotation and spatial scale invariant texture classification, in: *18th International Conference on Pattern Recognition, ICPR 2006*, vol. 4, 2006, pp. 29–32.
- [99] Y. Xu, H. Ji, C. Fermüller, Viewpoint invariant texture description using fractal analysis, *Int. J. Comput. Vis.* 83 (1) (2009) 85–100.
- [100] K. Zhan, H. Zhang, Y. Ma, New spiking cortical model for invariant texture retrieval and image processing, *IEEE Trans. Neural Netw.* 20 (December (12)) (2009) 1980–1986.