Impact of Endoscopic Image Degradations on LBP based Features using One-Class SVM for Classification of Celiac Disease

Sebastian Hegenbart, Andreas Uhl
Department of Computer Sciences
Salzburg University, Austria

Andreas Vécsei
St.Anna Children's Hospital
Vienna, Austria

***Abstract*—The prevalence data of celiac disease have been continuously corrected upwards in the last years. An automated decision support system could improve the diagnosis and safety of the endoscopic procedure. An approach towards such a system is based on a one-class classifier (such as SVM) trained on celiac data only. By doing so, no special treatment of distorted image areas is needed. However, the performance of such a system is highly dependent on the discriminative power of the extracted features within an unconstrained environment such as the human bowel. Towards such a system we evaluate how well methods used in past work perform using a one-class SVM with images exhibiting common endoscopic image degradations such as blur, noise, light reflections and bubbles.**

I. Introduction

Most methods used for texture classification are developed for still images. Modern endoscopes however, transmit an entire stream of frames. For an automated decision support in endoscopic treatments, methods are needed to identify informative frames for classification. The standard approach towards such a system is therefore based on a stage of identification of informative frames followed by segmentation and classification [1]. Hence, the reliability of such as system is based on the quality of the recognition of informative frames.

An alternative approach is based on a one-class classifier (such as SVM) trained on celiac image data. By restricting the classification method to a single class, all frames of an endoscopic image stream can be treated as informative. No method for distinguishing between distorted and informative frames is needed. As a consequence, frames showing either distortions or no celiac specific markers are classified as no celiac and can therefore be ignored for further processing. This approach implicitly combines the informative frame identification with classification.

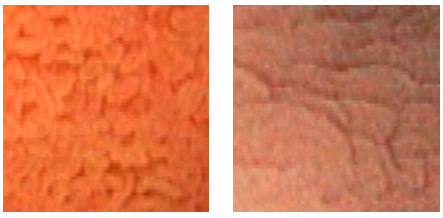
The accuracy of the second approach is heavily dependent on the discriminative power of the extracted features. The features are now required to be discriminative between the specific classes and to be able to compensate for image degradations caused by endoscopic distortions.

In recent work [2]–[4] we have shown that the automated classification of celiac disease based on endoscopic imagery is feasible using methods based on Local Binary Patterns (LBP) [5]. However, this work has been based on using a constrained image set with high quality [2]. Towards a more realistic scenario, this work is focused on the evaluation of how well methods based on LBP perform using a one-class SVM in an unconstrained environment. We focus on the most prominent types of endoscopic image degradations such as blur, noise, bubbles and specular reflections. In order to be able to assess how the specific methods are affected by certain types and levels of image degradations we simulate the common types of distortions.

In Section I-B we review the common endoscopic image degradations, Section II covers the simulation of the distortions in order to generate a dataset with known ground truth for evaluation. We discuss the details of feature extraction and classification in Section III and the details of experimentation in Section V. Finally, the results are discussed in Section VI while Section VII concludes the paper.

A. Celiac Disease

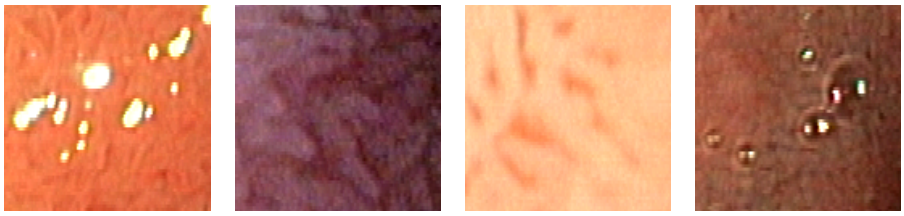
Celiac disease is a complex autoimmune disorder caused by the introduction of gluten containing materials such as wheat, rye and barley. During the course of the disease, hyperplasia of the enteric crypts occurs and the mucosa eventually looses its absorptive villi thus leading to a diminished ability to absorb nutrients. People with untreated celiac disease, are at risk for developing various associated complications like osteoporosis, infertility and other autoimmune diseases including type 1 diabetes, autoimmune thyroid disease and autoimmune liver disease.



(a) No-Celiac (b) Celiac

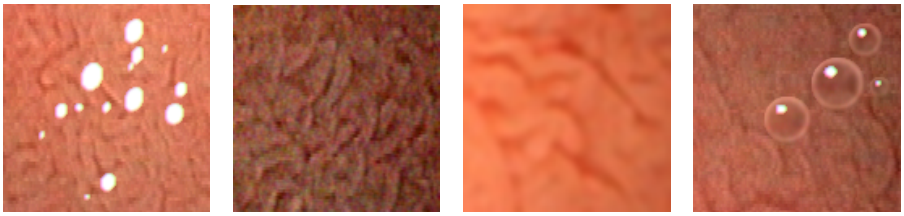
Fig. 1. Examples of Duodenal Image Patches

Figure 1 demonstrates two characteristic images showing healthy mucosal tissue on the left, and the effects of celiac disease on the right. The only treatment is a life long strict



(a) Reflections (b) Noise (c) Blur (d) Bubbles

Fig. 2. Examples of Image Degradations



(a) Reflections (b) Noise (c) Blur (d) Bubbles

Fig. 3. Examples of Simulated Image Degradations

gluten free diet which allows the tissue to heal, leading to a resolution of all symptoms in most cases.

B. Endoscopic Image Degradations

During the endoscopic procedure a small flexible tube (the endoscope) equipped with a camera and a point light source is introduced into the human bowel. The camera has a fixed focus, therefore all areas outside the focal plane appear blurred. Blur is a known problem in all areas of image processing, however in the specific case of classification of celiac disease, blur leads to another associated effect. During the course of the disease the mucosal villi are lost (villous atrophy). The length and form of the villi indicate the severeness of the disease. Depending on the strength of the blur, a healthy mucosa might be misinterpreted as being affected by celiac disease.

The bowel is illuminated using a point light source on the tip of the endoscope. Due to the geometric properties of the bowel a correct exposure can not always be guaranteed. Underexposure leads to an increase of amplifier noise (which is mainly based on thermal noise) within the digital image.

Finally, the third and fourth forms of common degradations are specular reflections and bubbles respectively, visible on the mucosal tissue. Light is reflected by moist tissue, while bubbles build up due to the instillation of water and insufflation of air into the bowel. Figure 2 shows examples of the most common degradations found in endoscopic imagery.

Another type of distortion is the strong lens distortion. The impact of this type of distortion on the automated classification of celiac disease was analyzed in a previous work [6].

II. Simulation of Image Degradations

In order to be able to quantitatively assess the impact of image degradations we need a known ground truth for the level of image degradations. Therefore the four aforementioned types of image degradations are simulated. We perform this simulation on the evaluation set of images as denoted in Table I.

TABLE I
Distribution of Image Data

	$\mathbf{Class_0}$	$\mathbf{Class_1}$	**Total**
		Images	
Training Set	-	157	157
Evaluation Set	151	149	300

Table I shows the distribution of the used images. Class_0 consist of images showing no villous atrophy (Marsh-0 type), while Class_1 is comprised of images showing mucosal tissue affected by celiac disease (Marsh-3 type). Figure 3 shows examples of simulated distortions.

A. Noise

Amplifier noise is primarily caused by thermal noise. Due to signal amplification in dark (or underexposed) areas of an image, thermal noise has a high impact on these areas. Additional sources contribute to the noise in a digital image such as shot noise, quantization noise and others. These additional noise sources however, only make up a negligible part of the noise and are therefore ignored during this work.

Let P be the set of all pixels in image $I \in \mathbb{N}^2$, $\omega = (\omega_p)_{p \in P}$, be a collection of independent identically distributed real-valued random variables following a Gaussian distribution with mean m and variance σ^2. We simulate thermal noise as additive Gaussian noise with $m = 0$, variance σ^2 for pixel p at position x, y as

$$N(x, y) = I(x, y) + \omega_p, \quad p \in P, \tag{1}$$

with N being the noisy image, for an original image I.

Our image data is extracted from an MPEG2 stream. The compression is based on a discrete cosine transform followed by a quantization step. The characteristics of thermal noise

are therefore changed due to the compression. We simulate this effect by applying a low-pass filter (Gaussian filter) to the simulated noise prior to adding it to the original image. Because the lowpass filtering used during compression only affects small high frequency components, this effect is neglected in case of the other types of distortions.

B. Blur

Out of focus blur is one of the most frequent distortions in endoscopic images. Blur is mainly caused by a wrong distance of the camera to the mucosa. Another type of blur is motion blur which is either caused by peristaltic or rapid movement of the endoscope. In this work we only consider out of focus blur. We simulate the point spread function of the blur as a Gaussian

$$f(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{2}$$

which is then convolved with the specific image.

C. Reflections

The light emitted by the endoscope is reflected by moist tissue. Reflections are usually seen as bright spots in certain areas of the image. We model the reflections as ellipsoids with maximum brightness and similar orientation. In natural images color distortions are seen along the contours of the reflections. This effect is caused by the arrangement of the RGB color filters on the CCD chips (Bayer pattern). In this arrangement the color of a single pixel is interpolated using a neighborhood of sensors. Therefore reflections lead to high intensity values in single color channels causing these color distortions along the borders. We simulate this effect by shifting the reflection in each color channel by a single pixel. The borders of reflections are not sharp. We therefore apply a blur to the reflection area to slightly smooth the borders. The mean length of the ellipsoids axes is set to 3 and 4 pixels respectively using a random scale factor with standard deviation of 1.2. This size relates to the common type of small spot reflections seen in endoscopic images. We chose to arrange the reflections such that no overlap happens. This was done to avoid a random factor in the degree of image degradation caused by the arrangement of reflections and does not influence the feature extraction directly.

D. Bubbles

We have shown [2] that the modified immersion technique for capturing images is beneficial to computer aided diagnosis. However, due to the instillation of water and the insufflation of air into the bowel bubbles can build up.

We simulate the appearance of bubbles using a template created in an image manipulation software resembling the visual properties of a bubble. As a simplification, bubbles are treated as circles. This was done to avoid a random factor and should have a negligible effect on the feature extraction. Due to soiled water in the bowel, mucosal tissue covered by bubbles is not clearly visible. We therefore apply a blur to the image area covered by a bubble. Bubbles influence the mucosal appearance instead of having a color. We therefore simulate this effect by considering the luminance component of an image (using the CIELAB color space). By interpolation of the image's luminance information using the bubble template, the mucosal color information is retained while the mucosal appearance resembles natural images containing bubbles.

One or more reflections can be observed on bubble surfaces, this effect is simulated using a single reflection positioned accordingly. These reflections are generated as discussed in Section II-C.

III. Feature Extraction and Classification

We use three LBP-based operators, which have shown to be promising in medical image classification in previous work [4]. The operators are LBP (Local Binary Patterns [5]), ELBP (Extended Local Binary Patterns [7]), and a modified version of the ELBP operator, the ELTP (Extended Local Ternary Patterns) operator [4].

For each color channel three scales (with the meaning of [8]) and filter orientations (in case of the extended LBP based operators: horizontal, vertical and diagonal) are used to compute the distribution of patterns. This results in 9 histograms for LBP and 27 histograms for ELBP and ELTP. For each histogram, only a subset of dominant patterns known as the uniform patterns [9], which make up the majority of discriminative patterns, is used

A. Local Binary Patterns (LBP)

For a radius r and the number of considered neighbors p, the LBP operator is defined as

$$LBP_{r,p}(x,y) = \sum_{k=0}^{p-1} 2^k \, s(I_k - I_c), \tag{3}$$

with I_k being the value of neighbor number k and I_c being the value of the corresponding center pixel. The function s acts as sign function, mapping to 1 if the difference is smaller or equal to 0 and mapping to 0 otherwise. The distribution of patterns is then used as feature for classification. Figure 4 demonstrates the calculation of a pattern.

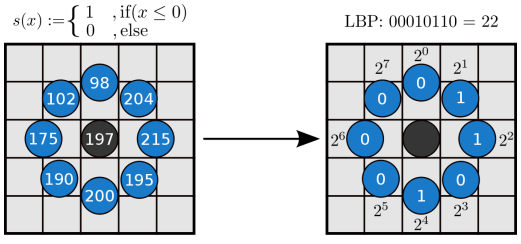
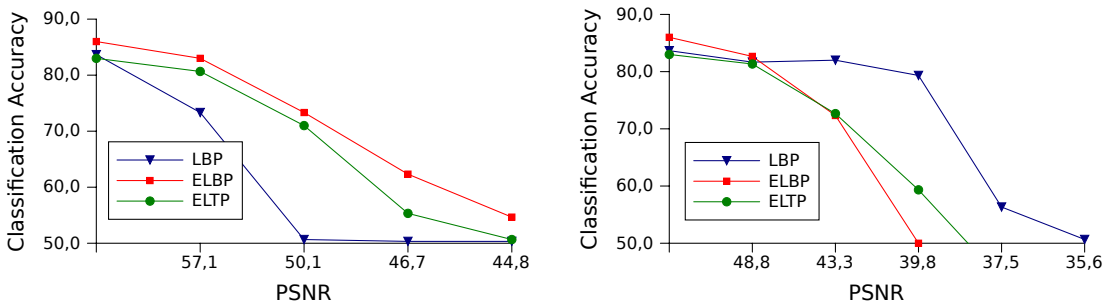


Fig. 4. Example of the Local Binary Pattern Operator

B. Extended Local Binary Patterns (ELBP)

Information extracted by the LBP method from the intensity function of a digital image can only reflect first derivative information. This might not be optimal, therefore Huang et al. [7] suggest using a gradient filtering before feature extraction



(a) Simulated Blur

(b) Simulated Noise

Fig. 5. Impact of Blur and Noise on the Classification Accuracy

and call this operator ELBP or extended LBP. By doing this, the velocity of local variation is described.

C. Extended Local Ternary Patterns (ELTP)

The extended LTP (ELTP) operator is defined consequently in perfect analogy to the ELBP operator. ELTP is based on the LTP (Local Ternary Patterns [10]) operator instead of the LBP operator to suppress unwanted noise in the gradient filtered data. The LTP operator is based on a thresholding mechanism which implicitly improves the robustness against noise. The LTP operator is used to ensure that pixel regions influenced by these kind of distortions do not contribute to the computed histograms. The LTP method is based on a thresholded sign function:

$$s(x) = \begin{cases} 1, & \text{if } x \geq T_h \\ 0, & \text{if } |x| < T_h \\ -1, & \text{if } x \leq -T_h. \end{cases} \tag{4}$$

The ternary decision leads to two separate histograms, one representing the distribution of the patterns resulting in -1, the other representing the distribution of the patterns resulting in 1. Both histograms are then concatenated and treated as a single histogram.

We apply an adaptive threshold based on the spatial image statistics to make sure that noisy regions do not contribute to the computed histograms while information present within high quality regions are not lost due to a threshold which was chosen too high. The calculation is based on an expected value for the standard deviation of the image (β). This value was found based on the training data used during experimentation and represents the average standard deviation of pixel intensity values within all training images. The value α is used as a weighting factor combined with the actual pixel standard deviation of the considered image (σ) and is used to adapt the threshold to match the considered image characteristics. During experimentation we used an α value of 0.05.

$$T_h = \begin{cases} \frac{\beta^{\frac{1}{2}}}{3} + \alpha\sigma, & \text{if } \sigma > \beta \\ \frac{\beta^{\frac{1}{2}}}{3} - \alpha\sigma, & \text{if } \sigma \leq \beta. \end{cases} \tag{5}$$

D. Classification

In this work a one-class Support Vector Machine [11] is used for classification. The classifier was trained using the data within the training set as depicted in Table I. We use parameters found in earlier experimentation. Hence, no further optimization of SVM parameters was performed.

IV. Experiments

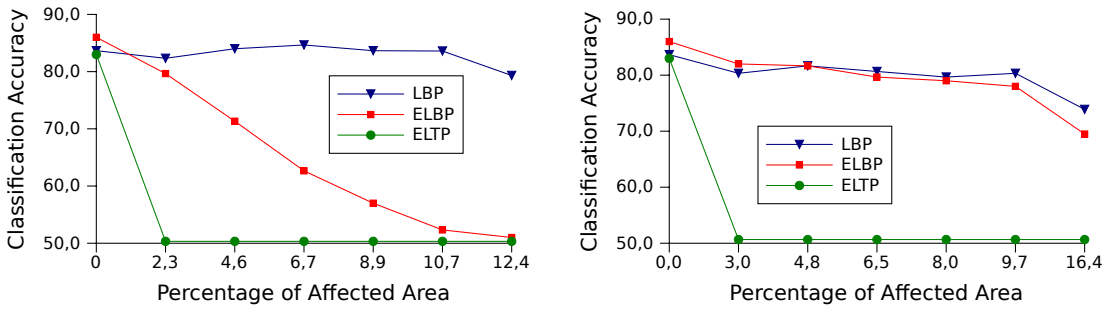
The results of the experiments are presented in Figures 5 and 6. The level of distortion by simulated specular reflections and bubbles is quantified by the percentage of the area of the original image that was affected by the distortion. In case of the reflections, 2.3 percent corresponds to 5 simulated reflections while 12.4 percent correspond to 30 reflections. A single bubble affects approximately 3 percent of the area of an image while 5 bubbles correspond to 16.4 percent. The blur was simulated using Gaussian filters with standard deviations ranging from 0.4 to 0.7. The noise was simulated using standard deviations in steps of 5 ranging from 5 to 25. In order to improve the readability we present the x-axes of Figure 5 labeled with the corresponding PSNR values.

The optimal feature subset for each texture operator was found by using the Sequential Forward Selection (SFS, [12]) algorithm. It must be noted that due to a limited number of image data and the nature of one-class SVM (only a single class to perform cross validation), the SFS algorithm was based on the classification accuracy of the undistorted evaluation set. Therefore the results might be slightly over fitted towards the evaluation data. Overall however, this should not have an impact on the analysis as the same feature set was used for all experiments regarding a specific method.

V. Results

Figure 5 shows how the methods behave when applied to noisy and blurred data. Considering the classification accuracies of the blurred images we already see an effect at a PSNR of 57 which corresponds to a Gaussian filter standard deviation of 0.4. We see that the standard operator (LBP) is most noticeably affected by this type of distortion while both, ELBP and ELTP, are better suited to handle blur.

Considering noise, the standard LBP operator is not significantly influenced by the distortion until a PSNR of 40



(a) Simulated Reflections

(b) Simulated Bubbles

Fig. 6. Impact of Reflections and Bubbles on the Classification Accuracy

(standard deviation of 15). The gradient based operators ELBP and ELTP are able to tolerate noise with standard deviations up to 5 (PSNR of 48.8). ELTP performs slightly better in terms of noise tolerance but both operators fail to achieve reasonable classification accuracies at higher noise levels.

Figure 6 shows the methods' classification accuracies on image data distorted with bubbles and specular reflections. The figures present the classification accuracies in relation to the percentage of the affected image area. Concerning reflections, we see that the accuracy of the standard LBP operator varies insignificantly. The classification rates are stable up to a distorted area of 10 percent. In case of ELBP we see a linear decrease in accuracy considering reflections. For distorted areas larger than 4.6 percent of the original image area the classification accuracy drops below 70 percent. The ELTP method fails completely to classify images distorted with reflections.

In case of simulated bubbles, we see that, in parallel to the specular reflection distortions, LBP is only slightly affected. We do not see classification rates below the 70 percent mark until an affected area of 9.7 percent. It is interesting that in contrary to the simulated reflections, the ELBP operator is as well only mildly affected by this type of distortion. The general behavior is similar to that of the LBP method. In contrast to that, the ELTP again fails completely to classify the distorted images.

VI. Discussion

In general we see from the results that classification in an unconstrained endoscopic environment using LBP based features is feasible. In general blur and noise had the highest impact on the classification accuracy. Especially blur had the most significant impact to classification accuracy. This can in general be explained by the characteristic markers of celiac disease which are lost due to the blur (blurred villi misinterpreted as villous atrophy). We also saw that the specific methods, although all based on LBP, react differently to certain types of image degradations.

A. Local Binary Patterns (LBP)

We saw that the LBP operator was heavily affected by blur. LBP considers neighborhoods of pixel intensity values which are all affected by blur. As a consequence, information useful to the method is lost due to this kind of distortion.

The method is suited best to handle noise. The low-frequency part of the noise (caused by quantization of DCT coefficients in the MPEG2 stream) affects entire pixel neighborhoods and therefore does not affect the intra neighborhood intensity values as much as high frequency noise would.

Bubbles and light reflections only had a small impact on the classification accuracy. This can be explained by the small distorted areas that actually affects the LBP operator. In case of reflections only the border of the reflections affect the distribution of patterns (inside the reflections all patterns are 255 which relates to no texture information and is ignored in our implementation). Therefore only a small part of distorted pixel neighborhoods actually affect the LBP method.

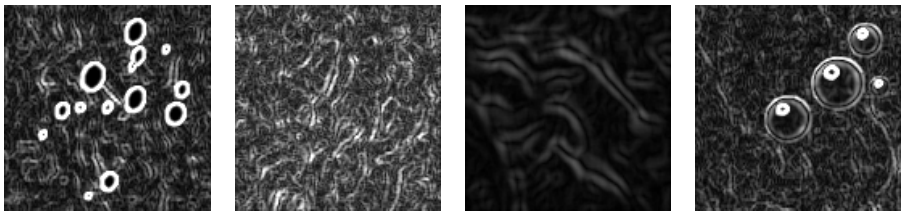
Bubbles had slightly more impact to the classification accuracy as compared to reflections. This is related to the blurred inner part of the bubbles which make up a larger area that negatively affects the operator due to the general information loss.

B. Extended Local Binary Patterns (ELBP)

Among all three methods, blur had the least impact on ELBP. Although small scale gradient information is lost due to blur, stronger gradients are retained. Due to the method's invariance in terms of monotonic grayscale changes, blur has a lower impact on ELBP as compared to LBP.

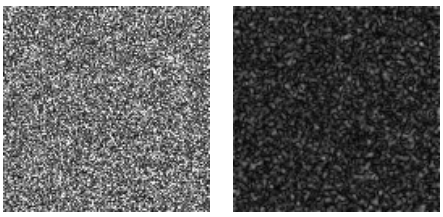
On the other hand, the method was significantly affected by noise. At a noise level of 5, the classification accuracy dropped rapidly. This can be explained by the properties of the low-frequency noise we used. Due to the Gaussian filtering of the noise, entire areas of the degraded image are affected by the same noise level (see Figure 8). Along the boundaries of these areas strong gradients exist which influence the operator's reliability in terms of classification accuracy.

In case of bubbles we see that the decrease in accuracy is almost linear. This is in contrast to the LBP method which was merely affected by this type of distortions and the ELTP operator which failed to classify images with this kind of distortions at the lowest level.



(a) Reflections (b) Noise (c) Blur (d) Bubbles

Fig. 7. Absolute Gradient Values of Distorted Images



(a) High-Frequency (b) Low-Frequency

Fig. 8. Example of Simulated Amplifier Noise

C. Extended Local Ternary Patterns (ELTP)

The ELTP method behaves similar to the ELBP method in case of images distorted with blur. The same also applies to noise with a small improvement as compared to ELBP due to the noise suppression by thresholding. However this improvement is only seen at low classification rates of approximately 70 percent.

In case of reflections and bubbles the method fails to classify images at very small levels of affected areas. This is due to the used thresholding. Considering Figure 7 we see that light reflections and bubbles introduce high-power gradients. In combination with thresholding, the smaller gradients caused by image texture get suppressed to some amount while the high-power gradients introduced by distortions all contribute significantly to the extracted features. This behavior could possibly be relaxed by introducing a second (upper) threshold to eliminate the contribution of high-power gradients.

VII. Conclusion

We evaluated the impact of common endoscopic image distortions using three LBP-based methods for feature extraction. A one-class support vector machines classifier, which was trained using celiac images, was used to classify endoscopic images with simulated distortions. We saw that the specific types of distortions have different effects on the methods.

We saw that distorted images can be accurately classified to some extent. It is interesting to note that bubbles and reflections have lesser impact on the classification rates than expected. Blur, the most common distortion, was best classified using gradient based methods while noise, bubbles and reflections could be handled well by the basic LBP operator.

We conclude, that the unconstrained classification of celiac disease based on LBP using a one-class SVM classifier is feasible to some degree. However, in extreme cases of image distortions an additional step of informative frame identification is unavoidable. By a possible relaxation of the demands on the frame identification method to extreme cases of distortions only, the general reliability could be increased.

We assume that by combining beneficial properties of the evaluated methods a more robust operator could be found to further improve the reliability of classification.

References

[1] M. Liedlgruber and A. Uhl, "A summary of research targeted at computer-aided decision support in endoscopy of the gastrointestinal tract," Department of Computer Sciences, University of Salzburg, Austria, http://www.cosy.sbg.ac.at/research/tr.html, Tech. Rep. 2011-01, 2011.

[2] S. Hegenbart, R. Kwitt, M. Liedlgruber, A. Uhl, and 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, Sep. 2009, pp. 718–723.

[3] S. Hegenbart, A. Uhl, and A. Véscei, "Systematic assessment of performance prediction techniques in medical image classification - a case study on celiac disease," in *Proceedings of the 22nd International Conference on Information Processing in Medical Imaging (IPMI'11)*, Monastery Irsee, Germany, July 2011, accepted.

[4] A. Vécsei, G. Amann, S. Hegenbart, M. Liedlgruber, and A. Uhl, "Automated marsh-like classification of celiac disease in children using an optimized local texture operator," *Computers in Biology and Medicine*, 2011, accepted.

[5] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, January 1996.

[6] M. Liedlgruber, A. Uhl, and 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, Jul. 2011, accepted.

[7] X. Huang, S. Li, and 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.

[8] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution Gray-Scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, July 2002.

[9] T. Mäenpää, T. Ojala, M. Pietikäinen, and M. Soriano, "Robust texture classification by subsets of local binary patterns," *Pattern Recognition, International Conference on*, vol. 3, p. 3947, 2000.

[10] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Analysis and Modelling of Faces and Gestures*, ser. LNCS, vol. 4778. Springer, oct 2007, pp. 168–182.

[11] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, p. 2001, 1999.

[12] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 153–158, 1997.