

Extended Gaussian-filtered Local Binary Patterns for colonoscopy image classification

Siyamalan Manivannan Ruixuan Wang Emanuele Trucco
CVIP, School of Computing, University of Dundee, UK.

{msiyamalan, ruixuanwang, manueltrucco}@computing.dundee.ac.uk

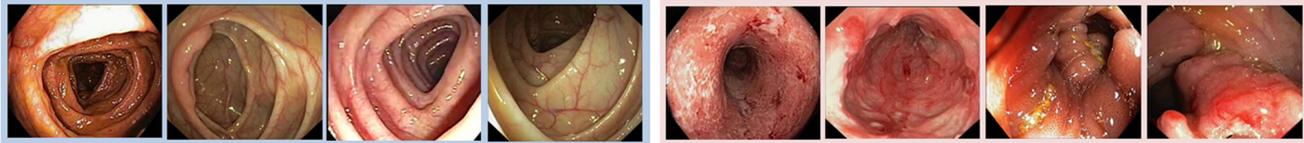


Figure 1: Some colonoscopy images: normal (left) and abnormal (right).

Abstract

Local Binary Patterns (LBP) and its variants are widely used for texture classification. In this paper we propose a new variant of LBP descriptor called the extended Gaussian filtered Local Binary Patterns (xGF-LBP) which is robust to illumination changes, noise and captures more informative edge-like features for classification. Experiments on a colonoscopy image dataset with 2100 images for binary ('normal' or 'abnormal') classification show that the proposed xGF-LBP descriptor significantly outperforms the standard LBP descriptor and its considered variants.

1. Introduction

Colorectal cancer is the second leading cause of cancer death in the world [1]. Adenoma detection rate (ADR), in terms of lesion detection, is a surrogate marker for quality of colonoscopy [15]. A reliable image processing system detecting abnormalities (including polyps, cancer, ulcers, etc.) in colonoscopy videos would be a useful screening tool to improve ADR by providing a consistent, repeatable and quantitative second opinion [13]. Here, we concentrate on normal-abnormal frame classification, a challenging task as abnormalities in colon vary in size, type, color, and shape (Figure 1).

Methods proposed for colonoscopy image classification are mainly focusing on feature design. Texture, color, shape and their combinations have been used in colonoscopy image analysis as features. Gray level co-occurrence matrices (GLCM) (e.g., detection of precancerous polyps in [14]), Local binary patterns (LBP) (e.g. normal vs ab-

normal colonoscopy image classification in [20]), Texture Spectrum (e.g. normal-cancer classification in [7]), etc. are used as texture features. Statistics of color values in different color channels (e.g. color histograms for bleeding detection in [8]), shape features (e.g., ellipses approximating contours for polyps detection in [6]) and combination of color, texture and/or shape features (e.g. Crohn disease classification in [9]) are also applied for colonoscopy image classification.

Local Binary Patterns (LBP)[17] is an efficient texture descriptor. Several variants of LBP have been proposed in the literature to improve its performance in classification. Due to its discriminative power and computational simplicity, LBP is proven to be a good descriptor for colonoscopy image classification and outperforms other considered features proposed in the literature for colonoscopy image analysis [20]. However, the standard LBP and its variants may not be robust to large illumination changes which often happens during colonoscopy with lighting source changed or adjusted. In this paper we propose a new variant of LBP called the extended Gaussian filtered Local Binary Patterns (xGF-LBP) which is robust to noise and illumination changes. In our formulation the standard LBP can be considered as a special case of xGF-LBP. We experimentally show that the proposed representation not only is robust to noise and illumination changes but also outperforms the LBP and its other variants on a colonoscopy dataset with 2100 images.

In the following, LBP and its relevant variants will be first introduced (Section 2) in order to explicitly compare with the proposed new feature (Section 3), with quantitative comparison shown later (Section 4).

2. Local Binary Patterns

LBP[17] as a powerful texture descriptor has been widely applied to texture classification, face recognition [5], and medical image classification [20], etc. While many variations of LBP have been proposed, e.g. uniform LBP[16] and BlockLBP[12], here we only review LBP and the most relevant variants.

LBP with N uniform sampling points on a circle of radius R around a 2D point \mathbf{p}_c in a gray image \mathbf{I} can be defined as:

$$LBP_{N,R}(\mathbf{p}_c) = \sum_{n=1}^N q_n \times 2^{n-1} \quad \text{where, } q_n = \begin{cases} 1 & I_n \geq I_c \\ 0 & I_n < I_c \end{cases} \quad (1)$$

I_c and I_n respectively represents the intensity values at the center point \mathbf{p}_c and the n -th sampled image point. I_n is bilinearly interpolated when the sampling point is not coincided with a pixel coordinate. Since this operator gives 2^N different labels an image can be represented as a histogram with 2^N bins.

To make LBP robust to noise, a three level thresholding is applied in *Local Ternary Patterns (LTP)* [22] by the introduction of a user specified threshold τ (Equation 2). The histogram representation of an image by LTP is obtained by splitting each LTP into two LBP and then concatenating the two LBP-based histograms.

$$q_n(\tau) = \begin{cases} 1 & I_n \geq I_c + \tau \\ 0 & |I_n - I_c| < \tau \\ -1 & I_n \leq I_c - \tau \end{cases} \quad (2)$$

However, LTP may be sensitive to illumination changes. When a scene illuminated by a single distant light source, the observed luminance image $I(x, y)$ at point (x, y) can be approximated as the product of the reflectance image $R(x, y)$ and the illuminance image $s(x, y)$ [2], i.e.,

$$I(x, y) = s(x, y)R(x, y) + G(x, y) \quad (3)$$

An ideal LBP should be robust to the change in illuminations s and the Gaussian noise G . While LTP is robust to noise with the introduction of the threshold τ , it is sensitive to the changes in illumination s , where the same threshold τ (Equation 2) may result in different LTP for an image under different illumination conditions (see Figure 2(c)).

To handle the variation in illumination, a variant of LTP called the *Scale Invariant Local Ternary Pattern (SILTP)* has been proposed in [11], i.e.,

$$SILTP_{N,R}(\mathbf{p}_c, w) = \oplus_{n=1}^N b_n(w) \quad \text{where, } b_n(w) = \begin{cases} 01 & I_n > (1+w)I_c \\ 10 & I_n < (1-w)I_c \\ 00 & \text{otherwise} \end{cases} \quad (4)$$

where w is the scale factor and \oplus denotes concatenation operator of the 2-bit binary strings b_n . Note that SILTP is not designed particularly for image classification. In fact, the ‘2-bit’ codes can be converted to ‘ternary’ patterns to generate a histogram representation for an image. In addition, SILTP does not consider the possible effect of noise, i.e., the threshold value is fixed to 0 as in LBP, which may lead to a representation sensitive to noise (Figure 2(d)).

Different from LBP and its variants, the proposed new LBP variant, called extended Gaussian filtered local binary patterns (xGF-LBP), is robust to both noise and illumination changes by capturing edge-like features.

3. Extended Gaussian-Filtered LBP

We first extend the standard LBP by adding a threshold parameter τ and a scale factor w , i.e.,

$$\text{x-LBP}_{N,R}(\mathbf{p}_c, w, \tau) = \sum_{n=1}^N q_n(w, \tau) \times 2^{n-1} \quad (5)$$

where,

$$q_n(w, \tau) = \begin{cases} 1 & w \times I_n - I_c \geq \tau \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

In the extended LBP (i.e., x-LBP), w can account for illumination changes and τ can account for noise (e.g., Figure 2(e)). The avoidance of ‘ternary pattern’ in x-LBP makes it easy and efficient to generate a histogram representation for an image as by the standard LBP. In fact, the standard LBP can be seen as a special case of x-LBP with $w = 1$ and $\tau = 0$.

3.1. Effect of parameters

Here we qualitatively show that the proposed x-LBP can capture edge-like features in an image by approximate parameter setting. Figure 3 shows a noisy CT brain image and the x-LBP codes obtained with different parameter settings. While the standard LBP (Figure 3(f)) codes are very noisy, changing the threshold values produces a better representation (Figures 3(e) and 3(g)). Changing both the threshold and the weights leads to an even better noise-reduced LBP (Figure 3(b)), with meaningful edge-like features preserved and noises reduced.

Figure 4 shows a colonoscopy image under two different illumination conditions (Figures 4(a) and 4(b)) as well as and its LBP and x-LBP representations. It is clear that the original LBP codes (second column) are very noisy compared to the x-LBP (last two columns). With appropriate parameter setting, x-LBP codes are also robust to illumination variations (e.g., the two very similar x-LBP representations in the last column).

| | | | | | | | | | | | | | | |
|-----|----|----|---|---|---|----|---|----|----|----|----|---|---|---|
| 70 | 40 | 26 | 1 | 1 | 0 | 1 | 1 | 0 | 01 | 01 | 10 | 1 | 1 | 0 |
| 10 | 30 | 32 | 0 | | 1 | -1 | | 0 | 10 | | 00 | 0 | | 0 |
| 35 | 30 | 31 | 1 | 1 | 1 | 1 | 0 | 0 | 01 | 00 | 00 | 0 | 0 | 0 |
| 140 | 80 | 52 | 1 | 1 | 0 | 1 | 1 | -1 | 01 | 01 | 10 | 1 | 1 | 0 |
| 20 | 60 | 68 | 0 | | 1 | -1 | | 1 | 10 | | 10 | 0 | | 0 |
| 65 | 56 | 58 | 1 | 0 | 0 | 1 | 0 | 0 | 00 | 00 | 00 | 0 | 0 | 0 |

(a) Image patches (b) LBP (c) LTP (d) SILTP (e) x-LBP

Figure 2: A demonstrative example for the effect of illumination change and noise on (b) LBP, (c) LTP ($\tau = 5$), (d) SILTP ($w = 0.1$), and the proposed (e) generalized LBP ($w = 0.9$, $\tau = 5$). (a) Top: original image patch with 3×3 pixels (i.e., $s = 1$ and $G = 0$ in Equation 3); Bottom: the transformed image patch with a different illumination ($s = 2$) with noise added to the pixels in red. LTP is robust to noise but not to illumination changes. LBP and SILTP are robust to illumination but not to noise. The proposed x-LBP is robust to both noise and illumination.

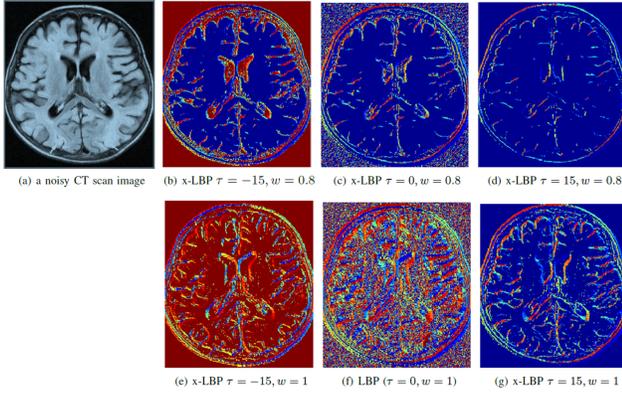


Figure 3: x-LBP of a CT image with different parameter settings. One LBP code (with range $[0, 255]$) is generated for each pixel and represented using a heat map, where color red indicates higher values and color blue indicates lower values.

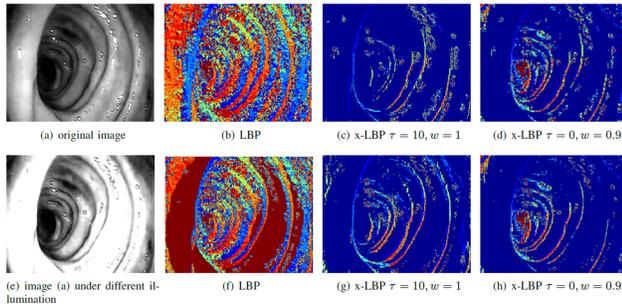


Figure 4: x-LBP of two colonoscopy images under different illuminations with different parameter settings. (the illumination changed image (e) is generated by applying a constant illumination field ($s = 2$) to the original image (a) according to Equation 3).

3.2. Extended Gaussian-Filtered LBP (xGF-LBP)

Gaussian filtered LBP (GF-LBP) has been proposed [19] to capture a larger neighborhood information and reduce

noise effect, where at each sampling point a Gaussian filter is applied to collect intensity information from an area larger than the original single pixel. Figure 5 shows an example for the generation of GF-LBP, where the local neighborhood is quantized radially into three resolutions (radii), and at each resolution a set of ($N = 8$) sampling regions (indicated as circles) are considered. A LBP code is constructed at each resolution by sampling the neighborhood at the centers of the solid circles, after a Gaussian filter with standard deviation proportional to the radius of the circle is applied to each circle's center [19]. Such a sampling pat-

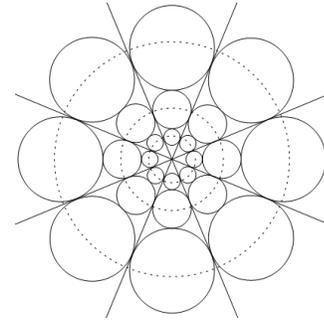


Figure 5: The Gaussian filtered sampling points.

tern of GF-LBP is similar to the spatial structure of receptive field of human retina, and has been widely adopted in recently developed visual features in computer vision, like FREAK [3], BRISK [10], and DAISY [23]. The combination of the sampling pattern with the proposed x-LBP is called xGF-LBP.

3.3. Image representation based on xGF-LBP

Denote by c a particular color channel (e.g., $c \in \{1, 2, 3\}$ for a color image), and by m a particular resolution (e.g., $m \in \{1, 2, 3\}$ in Figure 5), and by $\mathbf{h}(c, m)$ the x-LBP his-

gram of the an image I at the particular c and m (which can be generated as the standard LBP). Then the image can be easily represented by concatenating the histograms $\mathbf{h}(c, m)$ of x-LBP over all the resolutions and the color channels, resulting in the larger-dimensional histogram \mathbf{H} of xGF-LBP. The final image representation is then obtained as a normalized version of \mathbf{H} to make the features from different images to the same scale. Here we apply the L2 and power normalizations [18]:

$$\mathbf{H} \leftarrow \frac{\sqrt{\mathbf{H}}}{\|\mathbf{H}\|_2}, \quad (7)$$

where $\sqrt{\mathbf{H}}$ represent the square root operation on each component of \mathbf{H} . Unlike LTP where the final histogram representation of an image doubles the size of the GF-LBP histogram, in our case the dimensionality of the final feature vector is same as in the GF-LBP.

4. Experiments

This section describe the experimental settings and multiple experimental comparisons between the proposed xGF-LBP and relevant features in classification of colonoscopy images as normal or abnormal.

4.1. Experimental setup

The dataset contains 1050 normal and 1050 abnormal colonoscopy images, where the abnormal images contain various types of lesions or bleedings (Figure 1). Each image is rescaled by preserving its aspect ratio such that the maximum height is 300. For each image, xGF-LBP is extracted at every pixel in each color channel with the sampling pattern shown in Figure 5, where a three-resolution version was used with radii = $\{1, 2.43, 5.44\}$ and 8 sampling points were considered at each resolution, and 2D Gaussian filters of $\{(\text{window size} = 3 \times 3, \text{standard deviation} = 0.38), (\text{window size} = 7 \times 7, \text{standard deviation} = 0.85)\}$ were applied for the 2^{nd} and 3^{rd} resolutions respectively [19]. Gaussian filter is not applied for the first resolution level. Once the xGF-LBP is extracted, the histogram of xGF-LBP is used as the representation for the image (see Section 3.3). For all other LBP variants (including standard LBP and LTP), a similar process is applied to generate the corresponding histograms for each image. In all the cases the uniform pattern histogram representation [16] is considered as it captures some meaningful patterns such as edges, corners, etc. and reduces the dimensionality of the final histogram representation.

With the extracted histograms as image features, SVM classifier is trained (using the LibSVM toolbox [4]) with

the exponential chi-square kernel:

$$K(\mathbf{H}_1, \mathbf{H}_2) = \exp\left(-\frac{\gamma}{2} \sum_{i=1}^d \frac{(H_{1i} - H_{2i})^2}{H_{1i} + H_{2i}}\right) \quad (8)$$

where \mathbf{H}_1 and \mathbf{H}_2 are d -dimensional histograms to represent any two images, and H_{1i} and H_{2i} are the i -th component for the two histograms. The kernel parameter (γ) and the regularization parameter (C) of the soft margin SVM are learned based on a 5-fold cross validation on the training data. Classification performances are measured based the percentage of correctly classified images on the test dataset. Every classification result is based on the average over 10 experimental runs.

4.2. Classification using xGF-LBP

This section evaluates the classification performance of the proposed xGF-LBP, in comparison with the following baselines:

- (1) *LBP*: the standard LBP.
- (2) *LTP*: different threshold values $\tau \in \{5, 10, 20, 30\}$ were tried and the best classification performance was reported.
- (3) *DoG-LBP*: Since the proposed xGF-LBP captures edge-like features, we also applied DoG-LBP [21], where a difference-of-Gaussian (DoG) is applied to an image to emphasizes the edges and LBP is then computed from the DoG filtered image. For the two ('inner' and 'outer') Gaussians in the DoG filter, both the scale σ_1 of the inner Gaussian and the scale σ_2 of the outer Gaussian are chosen from the set $\Sigma = \{0.25, 0.5, 1, 2\}$, such that $\sigma_1 < \sigma_2$. Any of the two different window sizes $\{5 \times 5, 7 \times 7\}$ is tried for both Gaussians in DoG. The best classification performance was reported based on the different combinations of scales and window sizes.

For xGF-LBP, the best performance was reported based on the different combinations of the threshold $\tau \in \{-10, -5, 0, 5, 10\}$ and the weight $w \in \{1.4, 1.2, 1, 0.8, 0.6\}$. The size of the histogram features for LBP, DoG-LBP, and the proposed xGF-LBP is 531 (3 colors \times 3 resolutions \times 59 histogram bins), the histogram size for LTP doubles (i.e., 1062). For the classification, for each experimental run, a set of N training images were randomly selected from each of the two classes, and the remaining images were used for testing, where $N \in \{50, 100, 200, 300, 400, 500\}$

Figure 6 shows that the histogram based on the proposed xGF-LBP gives the best performance (with $w = 0.8$ and $\tau = 0$) and significantly outperforms other LBP versions. LTP gives a better performance (with $\tau = 10$) than LBP and DoG-LBP, and DoG-LBP gives the worse result. Table 1 shows the effect of the parameters (w and τ) when 500 images from each class are considered for training. It shows that varying the threshold value ($\tau \in -10, 0, 10$) when

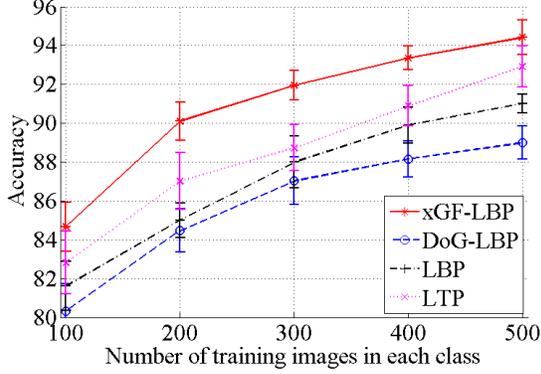


Figure 6: Experiments with the proposed xGF-LBP and other LBP variants.

$w = 1$ does not affect the classification performance. However, varying w results in different performance, with maximumly about 3% improvement compared to the standard LBP. Such improvement may result from the illumination robustness of the proposed xGF-LBP. During colonoscopy, the movement of the camera and lighting source inside the colon may lead to the imaging of the colon mucosa under different lighting conditions. The next experiment will further support the illumination robustness of xGF-LBP.

Table 1: The effect of w and τ in xGF-LBP (accuracy \pm std)

| $\tau \backslash w$ | 1.4 | 1.2 | 1 | 0.8 | 0.6 |
|---------------------|----------------|----------------|----------------|----------------|----------------|
| -10 | 92.1 \pm 0.7 | 92.0 \pm 0.5 | 91.3 \pm 0.5 | 90.6 \pm 0.7 | 91.7 \pm 1.2 |
| 0 | 93.9 \pm 1.3 | 93.9 \pm 0.8 | 91.0 \pm 0.5 | 94.4 \pm 0.9 | 93.1 \pm 0.6 |
| 10 | 93.0 \pm 0.9 | 91.5 \pm 0.7 | 91.0 \pm 0.7 | 92.1 \pm 0.6 | 91.1 \pm 0.6 |

4.3. xGF-LBP under different illuminations

This experiment investigates whether the proposed xGF-LBP is robust to illumination changes, where a set of images (500 normal and 500 abnormal) were used for classifier training, but the remaining images were artificially changed the illumination as the test images. For any original test image I , the illumination-changed image I_l is created by multiplying the intensity values of I by an illumination field s (a constant illumination field $s = \lambda$ was applied for the whole image) as in Equation 3. The intensity values of I_l which are greater than 255 are cropped to 255. Figure 8 shows some of the examples with different illumination factors λ .

With the new illumination-changed test images, the same experiment as in Section 4.2 were repeated and the results are shown in Figure 7. One can see that classification performance based xGF-LBP is robust (approximately constant) to different illuminations, while the performance based on LBP and LTP significantly degrades when the illumination decreases or increases. It is a bit surprise that

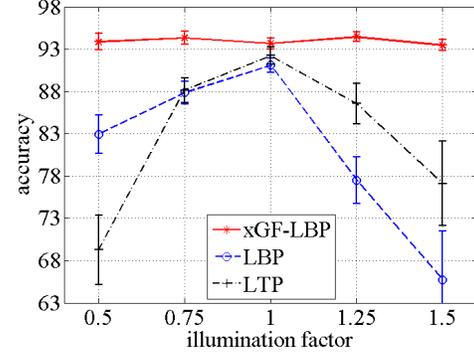


Figure 7: Classification performance of different methods under different illumination conditions.



(a) $\lambda = 0.5$ (b) $\lambda = 0.75$ (c) $\lambda = 1$ (d) $\lambda = 1.25$ (e) $\lambda = 1.5$

Figure 8: An image under different illumination conditions (original image is given by $\lambda = 1$).

LBP performs poorly with illumination changes because in general LBP is believed to be robust to illumination. Figure 9 uses an exemplar patch to illustrate the possible reason for the sensitivity of LBP to illumination changes. Figure 9(b) shows the patch 9(a) under a different illumination (the values are multiplied by 0.5). The pixel values are rounded to its integer representations to make sure they are in integer set $[0, 255]$. As a result of round-off errors, the the patches shown in Figure 9(a) and Figure 9(b) give different LBP representations lead to decrease in classification performance.

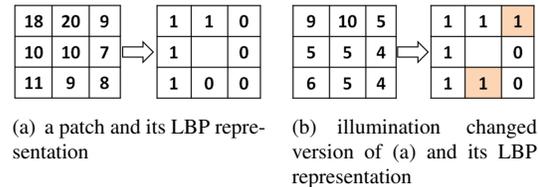


Figure 9: LBP codes of a patch under different illuminations. The shaded codes shows the codes which are affected by illumination.

5. Conclusions

In this paper we proposed a new variant of LBP, called xGF-LBP, which is robust to noise and illumination changes. Based on the experiments on a colonoscopy image dataset, we showed that the proposed method outperforms other LBP variants and robust to illumination changes in the

binary classification task. Future experiments will be focusing on testing the proposed feature on different datasets and on a dataset which is affected by both noise and illumination.

Acknowledgement This work is funded by 2011-2016 EU FP7 ERC project “CODIR: colonic disease investigation by robotic hydrocolonoscopy”, collaborative between the Universities of Dundee (PI Prof Sir A Cuschieri) and Leeds (PI Prof A Neville). Thanks to Dr. Adrian Hood (Leeds Institute of Molecular Medicine, University of Leeds, UK) for helping with the data annotations.

References

- [1] Cancer research UK. info.cancerresearchuk.org/cancerstats.
- [2] E. H. Adelson and A. P. Pentland. Perception as Bayesian inference. Cambridge University Press, 1996.
- [3] A. Alahi, R. Ortiz, and P. Vandergheynst. FREAK: Fast Retina Keypoint. In *IEEE Conf. on CVPR*, 2012.
- [4] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Trans. on IST*, 2, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [5] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen. Local binary patterns and its application to facial image analysis: A survey. *IEEE Trans. on SMC*, 41(6), 2011.
- [6] S. Hwang, J. H. Oh, W. Tavanpong, J. wong, and P. C. de Groen. Polyp detection in colonoscopy video using elliptical shape feature. *IEEE ICIP*, 2007.
- [7] S. Karkanis, K. Galousi, and D. Maroulis. Classification of endoscopic images based on texture spectrum. In *In Proceedings of Workshop on MLMA, Advance Course in AI*, 1999.
- [8] P. C. Khun, Z. Zhuo, L. Z. Yang, L. Liyuan, and L. Jiang. Feature selection and classification for wireless capsule endoscopic frames. *ICBPE*, 2009.
- [9] R. Kumar, Q. Zhao, S. Seshamani, G. Mullin, G. Hanger, and T. Dasopoulos. Assessment of Crohn’s disease lesions in wireless capsule endoscopy images. *Trans. on BEO*, 59, 2012.
- [10] S. Leutenegger, M. Chli, and R. Siegwart. Brisk: Binary robust invariant scalable keypoints. In *IEEE ICCV*, 2011.
- [11] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikainen, and S. Li. Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes. In *IEEE Conf. on CVPR*, 2010.
- [12] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S. Li. Learning multi-scale block local binary patterns for face recognition. In *Advances in Biometrics*, volume 4642 of *Lecture Notes in CS*. Springer Berlin Heidelberg, 2007.
- [13] M. Madhoun and W. Tierney. The impact of video recording colonoscopy on adenoma detection rates. *Gastrointestinal Endoscopy*, 75, 2012.
- [14] D. E. Maroulis, D. K. Iakovidis, S. A. Karkanis, and D. A. Karras. Cold: a versatile detection system for colorectal lesions in endoscopy video-frames. *Trans. on CMPB*, 70, 2003.
- [15] W. MB. Improving colorectal adenoma detection: technology or technique? *Gastroenterology*, 132, 2007.
- [16] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on PAMI*, 24(7), 2002.
- [17] T. Ojala, M. Pietikainen, and D. Harwood. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. *Proceedings of the 12th IAPR ICPR*, 1994.
- [18] F. Perronnin, J. Sánchez, and T. Mensink. Improving the Fisher kernel for large-scale image classification. In *ECCV*, 2010.
- [19] M. Pietikainen, A. Hadid, G. Zhao, and T. Ahonen. *Computer Vision Using Local Binary Patterns*. Springer, 2011.
- [20] S. Manivannan, R. Wang, E. Trucco, and A. Hood. Automatic normal-abnormal video frame classification for colonoscopy. In *IEEE ISBI*, 2013.
- [21] X. Tan and B. Triggs. Preprocessing and feature sets for robust face recognition. In *IEEE CVPR*, 2007.
- [22] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE TIP*, 19, 2010.
- [23] E. Tola, V. Lepetit, and P. Fua. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Trans. on PAMI*, 32(5), 2010.