Description and Classification of Confocal Endomicroscopic Images for the Automatic Diagnosis of Inflammatory Bowel Disease

Sara Couceiro[1], João P. Barreto[1], Paulo Freire[2], and Pedro Figueiredo[2]

[1] Institute for Systems and Robotics, University of Coimbra, Portugal
[2] Faculty of Medicine, University of Coimbra, Portugal
and Dept. of Gastroenterology, University Hospital of Coimbra, Portugal

Abstract. Confocal Endomicroscopy (CEM) is a newly developed diagnosis tool which provides *in vivo* examination of the gastrointestinal (GI) histological architecture, avoiding the traditional biopsy . The analysis of CEM images is a challenging task for experts, since there isn't a clearly defined taxonomy of the several disease stages. We aim at building an automatic on-the-fly classifier to provide useful clinical advices for diagnosis. In this work, we propose to make a split between two main subsets of our expert-annotated database: *low* and *high probability of pathology*. We focus on segmentation techniques to extract relevant histological structures, and then encode this information in a feature vector used for classification.

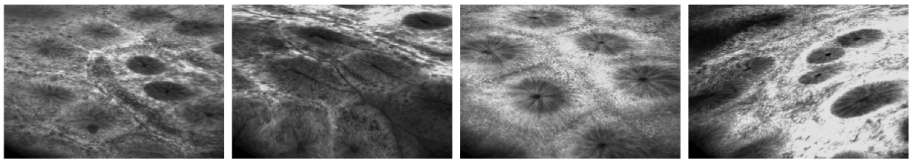
Keywords: Confocal Endomicroscopy, Image Classification, Support Vector Machine.

1 Introduction

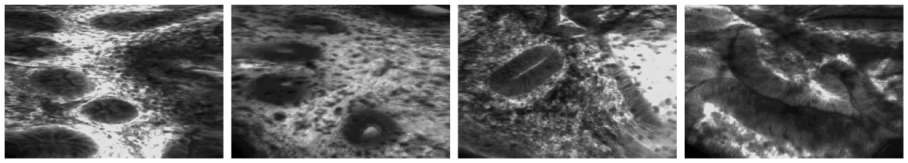
Inflammatory Bowel Disease (IBD) refers to a group of inflammatory conditions of the small intestine and the colon. Since most of the initial symptoms are undervalued by patients, IBD diseases are frequently noted at severe and chronic stages of illness. However, like in every clinical context, the diagnosis at early stages strongly increases the chances of successful treatment.

The common practice for the analysis of the histological architecture of the GI tract consists in removing the suspicious areas detected during endoscopy, and then sending them to lab analysis. Confocal Endomicroscopy (CEM) is a recently developed technique that allows the *in vivo* examination of the intestinal mucosa during ongoing endoscopy [5]. The basis behind this new diagnosis tool is the integration of a mini-confocal microscope with the distal tip of a conventional endoscope. The main advantage of the CEM over the traditional biopsy is the real-time examination of the intestinal histology, avoiding the need of tissue removal with increased risks of infection and bleeding, which require subsequent medical treatment. Unfortunately, since this is a quite recent technique, the taxonomy of CEM images has not yet been clearly defined. Physicians are still

F. Wang et al. (Eds.): MLMI 2012, LNCS 7588, pp. 144–151, 2012.



(a) data subset 1: *low probability of pathology*



(b) data subset 2: *high probability of pathology*

Fig. 1. Inter and intra class variability

exploring the complexity of these images and, even for them, their interpretation is still a challenging task [1].

This paper presents exploratory research toward building an on-the-fly automatic classification system to distinguish between two main IBD stages: *low* and *high probability of pathology*. The main challenge in this classification problem is related to the intra and inter class variability. As it is shown in Fig.1, besides the great variability between images from the same class, the differences between the two classes are not clearly evident.

In what concerns the histological architecture of the GI tract, the most evident structures are the intestinal crypts, which are responsible for the ongoing renewal of the epithelial tissue. According to experts feedback, the presence of crypts in the tissue is highly relevant to make a diagnosis. Their number, shape, appearance and distribution over the tissue are determinant to distinguish between the disease stages. Therefore, we focus on segmentation and description of these intestinal structures to detect abnormalities in the tissue.

Due the novelty of the CEM technique, the classification of the CEM images has not been very explored yet. Relevant references can be found in [1, 2], where André *et al* propose a content-based retrieval approach to perform classification of endomicroscopic images from an expert-annotated database. It starts by describing images using bi-scale dense Scale-Invariant Feature Transform (SIFT) features [8], that are quantized into visual words using k-means clustering. The classification is then performed by querying an image database with a histogram of visual words and retrieving the most similar images from the database. This approach requires large amounts of storage space, which may become unfeasible for large database sizes.

Unlike André's works [1, 2], that blindly apply techniques developed for conventional perspective cameras, we build our system based on the physicians interpretation of CEM images. We rely on tissue's histological properties to create a real-time classification system to assist doctors during endoscopy, while

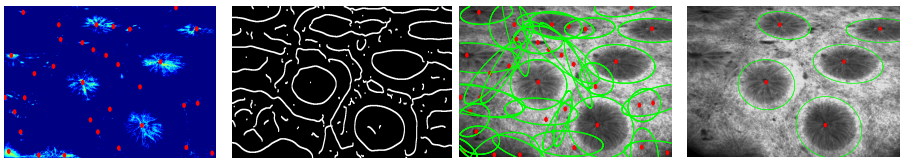
segmenting relevant structures and, therefore, provide helpful guidances for medical imaging analysis.

The paper outline is as follows: Section 2 describes the image segmentation scheme used to locate crypts in the tissue. In section 3, we refer to the chosen approach to face the classification task. Section 4 contains the experiments and discussion. In section 5, we present the actual research conclusions and refer to future work goals.

2 Detection and Segmentation of Crypts in CEM Images

Due to the great appearance variability of crypts, their segmentation is not a trivial task. As it can be observed in Fig. 1, crypts are irregular structures with markedly different texture patterns, and contour boundaries that are often poorly defined. In many cases, it is very difficult to accurately locate the exact frontier that separates the interior and exterior areas of a crypt, since there is a soft transition between these two s.

We propose an algorithm for the detection and segmentation of crypts that comprises three main steps, which are further explained in 2.1, 2.2 and 2.3. We start by performing the detection of crypts' centers by searching for local maxima in an energy of image symmetry (Fig.2(a)). Then, in the segmentation stage, we use a canny filter to locate the contours of these structures and fit an ellipse to the edge points (Fig. 2(b)). As we can see from Fig. 2(a) and 2(b), the detection step, not only catches the real centers of the crypts, but also several points that effectively denote symmetric structures in the image, but do not correspond to any crypt. Since there is no obvious solution to immediately discard the false detections, we decided to follow a strategy of overdetection. Centers detector is tuned to catch every high symmetric location, and the segmentation process is fully carried for all the detected keypoints. In the last stage (Fig. 2(c)), we use a binary Support Vector Machine (SVM) discriminator that will decide if each segmented corresponds to a crypt.



(a) Detection. (b) Segmentation. (c) Final result.

Fig. 2. Detection and Segmentation of Crypts in CEM images

2.1 Detection of Crypts' Centers Based on Symmetry Energy

Despite crypts great variability, their shapes range from circular to elongate, reminding elliptical structures, which are symmetric objects. We rely on this symmetry property of the crypts to identify them in intestinal tissue images.

Like in [7], we explore the bilateral symmetry of the local intensity levels of the image signal, by using a frequency-based approach. The extraction of local frequency information is done through a multi-resolution wavelet analysis, by using a bank of filters tuned at different scales. We use log-Gabor wavelets to obtain amplitude and phase information from the image signal [7].

The location of crypts centers is performed by computing local maxima of the symmetry energy (Fig. 3(a)).

2.2 Crypts' Shape Segmentation

Due to the elliptical shape of crypts, the most natural segmentation approach is to find an ellipse around each one of the detected keypoints. This is done by searching for boundary points around the center of the crypt and then fit a conic to them. The algorithm herein described was inspired and adapted from [10].

We first apply a Canny edge detector [3] over the original image signal to extract strong intensity transitions. Since the Canny algorithm relies on image local derivatives, it is highly sensible to noisy image data. To improve the final output of the Canny detector, we blur the image with a Gaussian kernel that is tuned to preserve crypts boundaries and eliminate other (irrelevant) image details. For each detected keypoint, we perform a radial search of boundary points, as proposed in [10]. We consider a searching circular patch in the original image whose center is the detected keypoint, and whose radius is large enough to include the whole crypt. The patch is mapped into a polar image by changing from Cartesian coordinates $\mathbf{x} = (x, y)^\mathsf{T}$ to polar coordinates $\chi = (\rho, \theta)^\mathsf{T}$. The ρ-dimension is scanned from the left to the right to find positive responses of the Canny filter, as show in Fig.3(b). Fig.3(c) shows the detected transitions mapped back from the polar space to the original Cartesian space. The parameters of the ellipse that best fits the detected edge points are determined using a RANdom SAmple Consensus (RANSAC) [4, 10]. The final result is shown in Fig 3(d).

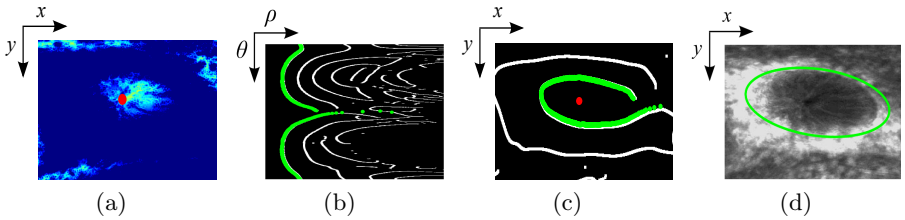
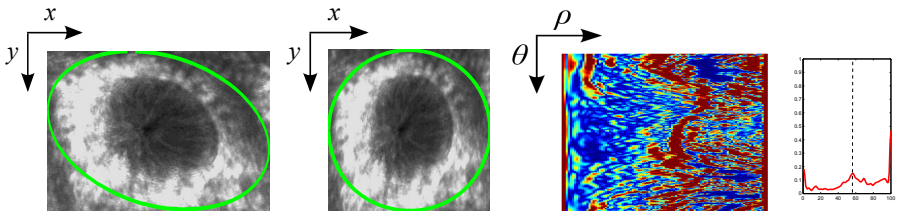


Fig. 3. Detailed steps of the detection and segmentation of a crypt in a CEM image. The search from left to right in the radial image (b) is substantially faster than in the Cartesian image (c).



(a) Segmented crypt. (b) Normalized isotropic patch. (c) Radial Gradients. (d) Radial Profile.

Fig. 4. Normalization step followed by feature extraction

2.3 Crypts Description and SVM Discriminator

In this step, we aim at building a discriminative local description scheme to encode crypts' appearance that enables a standard SVM classifier to distinguish between crypts and other symmetric structures.

Before the descriptors computation, we perform a normalization step of the segmented s to increase the resilience of the description schemes to image transformations, as shown in Fig.4. We use an affine transformation to map the segmented ellipses to equally sized circumferences. This corresponds to warp the image to a normalized isotropic patch in which the computed descriptors will be affine invariant by construction. This not only allows to normalize the shape of crypts, but also to reduce the slant effect that occurs if the endoscope is not positioned perpendicularly to the tissue surface during image acquisition.

We start by trying the SIFT descriptor [8] and texture analysis methods like Histogram Moments, Gray Level Co-occurence Matrix (GLCM) and Laws Energy [9], which have already been employed in the past in the context of medical image classification. In addition, we experiment with a specifically created radial descriptor, that uses image gradients for capturing small intensity and texture variations between crypts' interior and exterior s. The idea is to identify a transition that might correspond to the boundary of the crypts and, thus, encode the shape of these structures. Such a radial gradient profile is obtained by averaging the gradient magnitude values over 360 degrees around the center of each crypt (Fig.4(c)). The descriptor is normalized using the L2-norm for robustness against image contrast and brightness changes [8]. Fig.4(d) shows the normalized radial profile, in which it is possible to identify a slight peak of magnitude, marked by the dashed line, which corresponds to the boundary of the crypt.

3 Classification

As pointed out by medical experts, the spatial organization of crypts in CEM images is highly discriminative to distinguish between disease stages. Therefore, we propose to split the data into two main subsets: subset 1, in which there is a certain patterned arrangement of crypts, and subset 2, in which the tissue's

appearance is closer to a disordered stage. From a medical diagnosis point-of-view, this corresponds to a division between images with low (set 1) and high (set 2) probability of pathology. The diagnosis is accomplished by a standard SVM classifier that decides to which set the input image belongs. The SVM is applied on a feature vector that encodes the number of crypts in each image, provided by the crypts detector, and their arrangement in the tissue (*lattice*), computed using Delaunay Triangulation.

4 Experimental Validation

4.1 Dataset Specifications and Cross-Validation Scheme

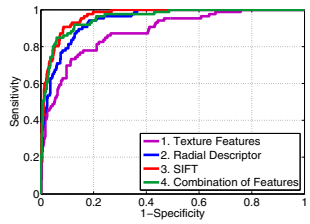
Our expert-annotated database contains a total of 192 CEM images, collected from 18 patients using a Pentax CEM device with a field-of-view (FOV) of 475 μm. The images are distributed over the two subsets as follows: 88 (46%) belonging to subset 1 and 104 (54%) to subset 2. Both in the SVM discriminator stage of crypts detection and in the binary SVM image classifier, we use the RBF kernel and 10-fold cross-validation to classify the whole dataset. The SVM parameters σ (spread of RBF kernel) and C (regularization term) were adjusted using a two-layer grid-search, as proposed in [6].

4.2 Evaluation of Crypts Detection and Segmentation

The success of centers detection is evaluated by quantifying the proportion of correctly detected centers. Thus, a local maxima is only accepted as a true detection if it is close enough to a real center. Both in subsets 1 and 2, we achieve high detection recalls (0.99). However, since we follow an overdetection approach, the precision values are very low (0.33 and 0.07 in subsets 1 and 2, respectively).

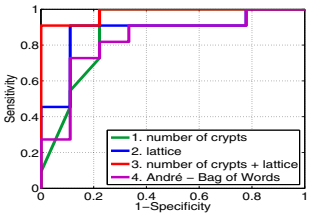
The success of the segmentation is quantified by metrics that compare the segmentation curves computed from crypts detector with the expert annotations. The criteria used in this comparison is based in the overlap area, the distance between the centers and the difference between the rotation angles. The segmentation algorithm performs quite well in subset 1, with a percentage of correctly segmented crypts of 78%, but falls short in subset 2, in which its performance is only about 42%. This is related to the tissue's appearance of the dataset images. In the first subset, there is a certain patterned arrangement of crypts. Besides, these structures show a more regular shape and their contours have stronger intensity transitions than in subset 2, which improves the canny's response.

In the SVM discrimination step, we compare the proposed radial gradient descriptor with textures features [9] and the SIFT descriptor [1,2,8]. The results are shown in Fig. 5. Our descriptor clearly outperforms the texture features and has very similar performance to the SIFT descriptor, with the advantage of being much simpler and easier to compute. It is also quite interesting to observe that, when using all the evaluated features, the precision of the detector increases (by reducing the number of spurious detections), but the recall is slightly decreased.



	ACC	SE	SP	PRC
1	0.80	0.80	0.80	0.37
2	0.86	0.86	0.86	0.47
3	0.89	0.88	0.89	0.55
4	0.92	0.73	0.94	0.67

Fig. 5. SVM Discriminator Scores. The graph shows a Receiver Operating Characteristic (ROC) curve obtained from one of the 10-fold cross validation rounds and the table shows the values of Accuracy (ACC), Sensitivity (SE), Specificity (SP) and Precision (PRC) computed by averaging the results over the rounds.



	ACC	SE	SP	PRC
1	0.83	0.86	0.80	0.83
2	0.86	0.85	0.88	0.90
3	0.89	0.90	0.88	0.90
4	0.71	0.82	0.58	0.71

Fig. 6. Image Classification Scores. The graph shows a ROC curve obtained from one of the 10-fold cross-validation rounds and the table shows the values of Accuracy (ACC), Sensitivity (SE), Specificity (SP) and Precision (PRC) computed by averaging the results over the rounds.

4.3 Image Classification Evaluation

In the image classification stage, we compare the features we proposed based on experts interpretation of CEM images with SIFT descriptor [1, 2, 8]. The results are shown in Fig. 6. As proposed in [1, 2], we quantize SIFT features into visual words and enconde them in image histograms. Then, instead of performing the retrieval, we use these histograms to train the SVM classifier. In the available dataset, our descriptor provides higher scores than André's. We consider this result quite satisfactory, since we are using conceptually simpler features, which reflect experts clinical evaluation of the GI histological appearance.

5 Conclusion

This paper presents preliminary research toward building an on-the-fly CEM image classification system, based on the physicians interpretation of the histological architecture of the GI tract, while segmenting relevant structures in the tissue and, thus, provide useful imaging analysis information. Given the complexity of the classification problem and the small size of the available dataset, we consider the current results encouraging toward further research. The current image classifier's performance shows that the histological appearance of

CEM images provide reliable information to train a discriminative and robust classifier. However, some technical improvements still need to be carried out to improve the crypts detector results. Currently, we are conducting research toward the enhancement of the discriminative power of the features used in the SVM discriminator. We aim at developing specific crypts descriptors to increase the SVM discrimination precision while maintaining high recall rates (Fig. 5). Future research directions will also focus on the improvement of the RANSAC-based segmentation both in terms of computational time and accuracy.

References

1. André, B., Vercauteren, T., Buchner, A.M., Wallace, M.B., Ayache, N.: A Smart Atlas for Endomicroscopy using Automated Video Retrieval. Medical Image Analysis 15(4), 460–476 (2011)
2. André, B., Vercauteren, T., Perchant, A., Buchner, A.M., Wallace, M.B., Ayache, N.: Introducing Space and Time in Local Feature-Based Endomicroscopic Image Retrieval. In: Caputo, B., Müller, H., Syeda-Mahmood, T., Duncan, J.S., Wang, F., Kalpathy-Cramer, J. (eds.) MCBR-CDS 2009. LNCS, vol. 5853, pp. 18–30. Springer, Heidelberg (2010)
3. Canny, J.: A Computational Approach to Edge Detection. IEEE Trans. Pattern Anal. Mach. Intell. 8(6), 679–698 (1986)
4. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24(6), 381–395 (1981)
5. Hoffman, A., Goetz, M., Vieth, M., Galle, P.R., Neurath, M.F., Kiesslich, R.: Confocal laser endomicroscopy: technical status and current indications. Endoscopy 38(12), 1275–1283 (2006)
6. Hsu, C.-W., Chang, C.-C., Lin, C.-J.: A practical guide to support vector classification. Technical report, Department of Computer Science, National Taiwan University (2003)
7. Kovesi, P.: Symmetry and Asymmetry from Local Phase. In: Sattar, A. (ed.) Canadian AI 1997. LNCS, vol. 1342, pp. 2–4. Springer, Heidelberg (1997)
8. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60(2), 91–110 (2004)
9. Manavalan, R., Thangavel, K.: Evaluation of textural feature extraction methods for prostate cancer trus medical images. International Journal of Computer Applications 36(12), 33–39 (2011)
10. Melo, R., Barreto, J.P., Falcao, G.: A new solution for camera calibration and real-time image distortion correction in medical endoscopy-initial technical evaluation. IEEE Trans. Biomed. Engineering 59(3), 634–644 (2012)
11. Sivic, J., Zisserman, A.: Video google: A text retrieval approach to object matching in videos. In: ICCV, pp. 1470–1477 (2003)