COMPUTER-ASSISTED AUTOMATED IMAGE RECOGNITION OF CELIAC DISEASE USING CONFOCAL ENDOMICROSCOPY.

Enrico Grisan[1], Hadis Mirzaei[2], Rupert Leong[2]

[1]Department of Information Engineering, University of Padova, Italy
[2]Faculty of Medicine, The University of New South Wales, Australia

ABSTRACT

Celiac disease (CD) is an immune-mediated enteropathy triggered by exposure to gluten and similar proteins which affects genetically susceptible persons, increasing their risk of different complications such as malignant lymphomas, small-bowel neoplasia, oropharyngeal tumors, osteoporosis, and bone fractures. In the clinical practice, the diagnosis of CD is made with a diagnostic intestinal biopsy and the concomitant presence of a positive celiac serology. Small bowels mucosa damage due to CD involves various degrees of endoscopically relevant lesions, that are not easily recognized: their overall sensitivity and positive predictive values are poor even when zoom endoscopy is used.

Confocal laser endomicroscopy (CLE) has recently revealed to be a useful technique for in vivo virtual histology of small bowels mucosa, but requires skilled and specifically trained operators.

We present a computer-based method for the automatic diagnosis of CD-related lesions on the basis of appearance features of confocal images. Comparing the automatic results with the histological gold standard, the proposed method classifies confocal images with accuracy comparable to human observer, suggesting the possibility of real time diagnosis of CD during endoscopy and a non-invasive method to replace biopsy

Index Terms— confocal endomicroscopy, computer aided diagnosis, celiac disease, random forests

1. INTRODUCTION

Celiac disease (CD) is an immune-mediated enteropathy triggered by exposure to gluten and similar proteins which affects genetically susceptible persons. Celiac disease is one the most frequent enteropathy and most of the celiac patients will remain undiagnosed during their life so it is a hidden epidemic. Exposure to gluten causes damage to the small bowel mucosa. The damage vary from mild, with only an increase in intraepithelial lymphocytes and Crypt Hyperplasia (CH), to severe form, which involves various degrees of endoscopically relevant lesions such as villous Atrophy (VA) [1] (see Fig. 1). VA and CH are important findings that could not be easily recognized during endoscopy; their overall sensitivity and positive predictive values are poor even when zoom endoscopy is used [2] [3]. In today practice, the diagnosis of CD is made with a diagnostic intestinal biopsy and the concomitant presence of a positive celiac serology [4]. Another post-treatment biopsy is sometimes necessary for those patients in whom the first biopsy and serologic test are inconclusive (e.g., seronegative enteropathy) or for patients who remain symptomatic in spite of having strict Gluten Free Diet (GDF) [5]. Introduction of a GFD can significantly improve the symptoms in symptomatic patients, as well as correct the abnormal biochemical measures, therefore improving in patient's quality of life.

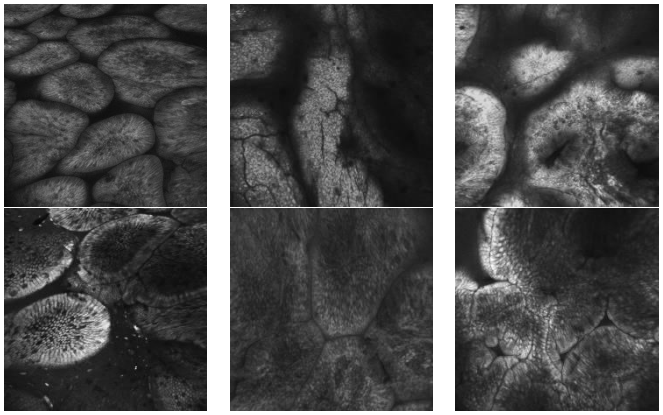

Figure 1 Representative images of villi appearance in the normal case (left column), in presence of VA (central column), and CH (right column)

Confocal laser endomicroscopy (CLE) is a novel endoscopic method that permits on-site microscopy of the gastrointestinal mucosa after the application of a fluorescent agent, allowing the experienced endoscopists to diagnose VA, CH as well as intraepithelial lymphocytes with high accuracy [6] [7] [8].

The ultimate goal of our study is to provide a system that has the potential to diagnose the CD at the time of endoscopy.

Preprint submitted to 2014 IEEE International Symposium on Biomedical Imaging.
Received October 10, 2013.

To date, few computer aided detection and diagnosis (CADD) systems have been proposed for confocal microendoscopic images, either for the analysis of salient features of images [9] [10] or for detecting subtle mucosal changes linked to pre-neoplastic and neoplastic tissues [11] [12] [13].

The classification of the state of colonic villi is made difficult by the variability in the appearance of relevant cues: common approaches resorting to the preliminary identification of single villi and their alterations are impractical. We exploit the growing evidence that the statistical properties of the image considered as a global entity, without any identification and analysis of its constituent objects, yield a rich set of cues, that are possibly sufficient for the identification of its correct semantic category (class) [14] [15].
To this end we extract a set of features from each image, which are then combined to represent an image signature. This signature is then fed to a classifier that identify the image as normal or diseased, at the same time assigning the image to the one of the classes: normal villi (NV), villous atrophy (VA), crypt hypertrophy (CH). This is achieved by first computing two scores based on the image signature: one represents the probability that the image presents villous atrophy (VAP), whereas the other represent the probability of the image showing crypt hypertrophy (CHP).
Finally each image, represented by its two dimensional score vector [VAP, CHP], is classified with a maximum a posteriori Bayesian classifier.

2. MATERIALS

In this retrospective study, 128 confocal images from 30 patients were obtained from previous clinical trial conducted at the Gastroenterology and Liver Services of the Bankstown-Lidcombe Hospital (Sydney, Australia) [6].
Subjects of that study were recruited with known CD, suspected CD, and controls. Each patients underwent a confocal gastroscopy (Pentax EC-3870FK, Pentax, Tokyo, Japan) under conscious sedation and with a IValiquots of fluorescein sodium and topical acriflavine hydrochloride to enhance images. CEM images and forceps biopsies of the same sites were taken sequentially at standardized locations at 5 small intestinal sites. Seven to 10 CEM images of different mucosal depths were collected from each site from the standardized locations for every forceps biopsy specimen. Small intestinal specimens were taken precisely matched to the CEM imaging sites and were assessed by 2 experienced blinded histopathologists independently and subsequently reviewed during a single session for internal consistency.

3. METHODS

3.1 Pyramidal decomposition

Multiscale representations [14] [16] of image characteristics have been shown to be powerful tools for scene classification [17] [18] [19]. The image I is represented by a Gaussian pyramid decomposition I^G of L levels, so that at level $l = 0, \dots, L-1$ the image component I_l^G can be obtained applying a Gaussian smoothing filter to the original image I, and downsampling the filtered image by a factor $\left(\frac{1}{2}\right)^l$. At the same time, the image is also represented by a Laplacian pyramidal decomposition I^Λ of L levels, by which the image component I_l^Λ can be obtained applying a Laplacian of Gaussian high-pass filter to the original image I, and downsampling the filtered image of a factor $\left(\frac{1}{2}\right)^l$.

3.2 Image signature

From each image I_l^G and I_l^Λ of the pyramidal decomposition of the original image I at level l, a set of features h_l^G and h_l^Λ are computed, representing the gray-level distribution, the gray-levels spatial organization (texture) and the presence of relevant bright or dark patches.
The intensity distribution is characterized by a vector of statistical descriptors (moments, percentiles, entropy) $h_{stat,l}$, whereas the texture are described evaluating rotation invariant Local Binary Pattern (LBP) as described in [20] [21]. Since LBP assigns to each pixel a binary code that represents the difference pattern between the pixel intensity and its neighboring pixels evaluated along a circle of radius R, it is particularly suited to accommodate the large intensity variability both among different CLE images and within the same image: the histogram $h_{LBP,l}$ of the binary codes assigned to each pixel can represent the textural characteristics of an image, regardless of its mean intensity, luminosity drift, and rotation.
Finally, image bright (dark) patches are evaluated thresholding the image with a set Θ of thresholds. For each threshold $\theta \in \Theta$, we identify the regions brighter (darker) than θ. A feature vector $h_{bright,l}$ (and corresponding $h_{dark,l}$) is extracted containing the mean area of the bright region (dark), the number of regions larger than half of the image area, the ratio of perimeter to area, and finally the fractal dimension of the bright (dark) pixels, computed with the box counting method [22].
All features vector for the images in the Gaussian and Laplacian pyramidal decomposition are then concatenated to create the image signature $h = [h^G, h^\Lambda]$, with $h^k = \bigcup_{l=0}^{L-1}[h^k_{stat,l}, h^k_{LBP,l}, h^k_{bright,l}, h^k_{dark,l}]$, interpreting the union operator as concatenation.

Preprint submitted to 2014 IEEE International Symposium on Biomedical Imaging.
Received October 10, 2013.

3.3 Image Classification

For each image, the probabilities that it presents VA or CH are computed as the score of two binary random forest classifiers [23] [24] applied to its signature. One binary random forest classifier provides the score of an image presenting VA, whereas the other provides the score of an image presenting CH. The scores are normalized in the range [0,1], and can be interpreted as probabilities, since they are computed as the fraction of classification trees in the random forest "voting" for the presence of VA or CH.

Finally, the image is classified as NV, VA, or CH by means of a Bayesian maximum a posteriori classifier, under the hypothesis of normal multivariate distribution for the three classes.

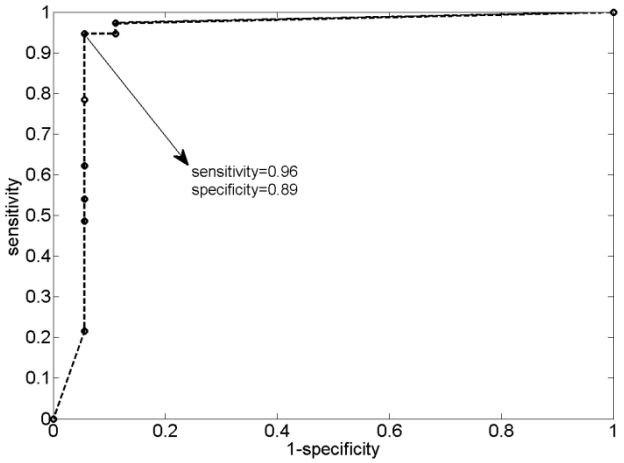

Figure 2 ROC analysis showing the specificity and sensitivity varying the classification threshold

4. EXPERIMENTS

The performance of the classifier has been evaluated using a leave-one-out procedure. At each round of the validation, out of the N=128 hystologically classified images, one image was used for testing, while the others N-1 (training set) were used for the classifier set-up.

The two random forests are grown with 500 classification trees each using Breiman algorithm [11], providing a two dimensional score vector [VAP, CHP] for each image in the training set. The score distribution for each class of interest (NV, VA, CH) is assumed to be fully described by the first two moments (mean vector and covariance matrix) of the score vectors of the images in the training set and belonging to the class. A maximum a posteriori classifier (naïve Bayes) is then used to assign each image to its final class. The a priori probability P_{normal} , P_{VA} , P_{CH} , of each class can be used to set the operating point of the classifier, so to weight differently the risk of lesion misdetection.

The test image is then classified: its scores VAP and CHP are computed and the image is finally classified as normal or presenting villous alterations (and which).

Considering as positive output are considered when a detection of either VA or CH or VACH results from the classification, and negative output when a normal image results, we can compute sensitivity (fraction of positive outputs over the number of positive samples in the dataset), specificity (fraction of negative outputs over the number of negative samples in the dataset), and accuracy of the classifier (fraction of images correctly classified).

To set the optimal operating point of the classifier, we performed a ROC analysis varying P_{normal} from 0 to 1 and setting correspondingly $P_{VA} = P_{CH} = P_{VACH} = (1 - P_{normal})/3$

5. RESULTS

On the 128 biopsied-matched images coming from the 30 subjects a leave one out validation scheme was used, coupled with a receiver operating characteristics (ROC) analysis to set the optimal value of the a priori probability P_{normal} (Fig. 2). The proposed method reached 96% sensitivity (probability of detecting images with either VA or CH or both) with 89% specificity (probability of detecting normal images). This method was successful in automatically identifying all four combinations of VA and CH presence of (1) no VA, no CH (normal mucosa), (2) prevalence of VA, (3) prevalence of CH. The AUC was 0.935 and the estimated classification error 0.07 (95% CI: 0.004-0.14) with accuracy 0.93 (95% CI: 0.86- 0.99). The naïve Bayes classifier bidimensional space (VAP, CHP) is shown in Fig.3 together with the posterior iso-probability lines at P=0.5.

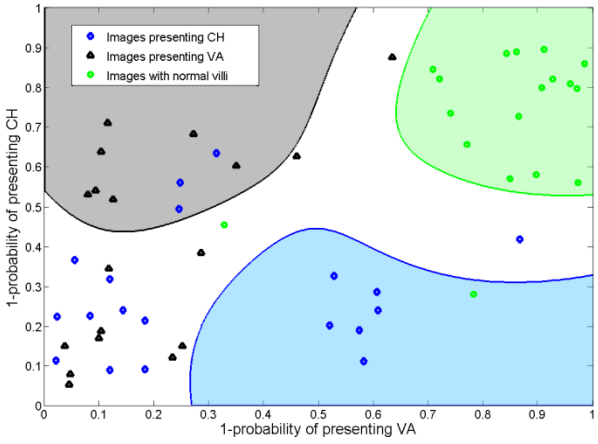

Figure 3 Naive Bayes classification space, with the posterior iso-probability lines at P=0.5 for the three classes superimposed (normal-green, VA-blue, CH-black). The lower left corner represent images showing both CH and VA

Preprint submitted to 2014 IEEE International Symposium on Biomedical Imaging.
Received October 10, 2013.

5. CONCLUSIONS

We have presented a computer-aided system for the automatic detection of small bowels mucosal lesions due to celiac disease on confocal microendoscopy images. The proposed method classifies confocal images with high accuracy with respect to the histological gold standard.
These result may provide a system for easier lesions detection/classification, allowingthe possibility of real time diagnosis of CD during endoscopy and a non-invasive method to replace biopsy

7. REFERENCES

[1] C. Mulder, S. van Weyenberg and M. Jacobs, "Celiac disease is not yet mainstream in endoscopy," *Endoscopy,* vol. 42, no. 3, pp. 218-9, 2010.

[2] D. Dewar and P. Ciclitira, "Clinical features and diagnosis of celiac disease," *Gastroenterology,* vol. 128, no. Suppl 1, pp. S19-S24, 2005.

[3] M. Bardella, G. Minoli, F. Radaelli, M. Quatrini, P. Bianchi and D. Conte, "Reevaluation of duodenal endoscopic markers in the diagnosis of celiac disease.," *Gastrointest Endosc.,* vol. 51, no. 6, pp. 714-6, 2000.

[4] A. Fasano and C. Catassi, "Current approaches to diagnosis and treatment of celiac disease: an evolving spectrum.," *Gastroenterology,* vol. 120, no. 3, pp. 636-51, 2001.

[5] C. Catassi and A. Fasano, "Celiac disease diagnosis: simple rules are better than complicated algorithms.," *Am J Med.,* vol. 123, no. 8, pp. 691-3, 2010.

[6] R. Leong, N. Nguyen, C. Meredith, S. Al-Sohaily, D. Kukic, P. Delaney, E. Murr, J. Yong, N. Merrett and A. Biankin, "In Vivo Confocal Endomicroscopy in the Diagnosis and Evaluation of Celiac Disease," *Gastroenterology,* vol. 135, no. 6, pp. 1870-6, 2008.

[7] A. A.-T. M. C. K. Venkatesh, C. Evans, S. Thomas, P. Oliver, C. Taylor and M. Thomson, "Role of confocal endomicroscopy in the diagnosis of celiac disease.," *J Pediatr Gastroenterol Nutr.,* vol. 51, no. 3, pp. 274-9, 2010.

[8] M. Goetz, A. Watson and R. Kiesslich, "Confocal laser endomicroscopy in gastrointestinal diseases.," *J Biophotonics,* vol. 4, no. 7-8, pp. 498-508, 2011.

[9] M. Häfner, M. Liedlgruber, A. Uhl, A. Vécsei and F. Wrba, "Color treatment in endoscopic image classification using multi-scale local color vector patterns," *Med Image Anal. ,* vol. 16, no. 1, pp. 75-86, 2012.

[10] V. Becker, M. Vieth, M. Bajbouj, R. Schmid and A. Meining, "Confocal laser scanning fluorescence microscopy for in vivo determination of microvessel density in Barrett's esophagus.," *Endoscopy,* vol. 40, no. 11, pp. 888-91, 2008.

[11] S. Gross, S. Palm, J. J. W. Tischendorf, A. Behrens, C. Trautwein and T. Aach, "Automated classification of colon polyps in endoscopic image data," in *SPIE Medical Imaging 2012: Computer-Aided Diagnosis*, 2012.

[12] B. André, T. Vercauteren, A. Buchner, M. Wallace and N. Ayache, "A smart atlas for endomicroscopy using automated video retrieval.," *Med Image Anal.,* vol. 15, no. 4, pp. 460-76, 2011.

[13] E. Veronese, E. Grisan, G. Diamantis, G. Battaglia, C. Crosta and C. Trovato, "Hybrid patch-based and image-wide classification of confocal laser endomicroscopy images in Barrett's esophagus surveillance," *2013 IEEE 10th International Symposium on Biomedical Imaging (ISBI),* pp. 362-365, 2013.

[14] S. Lazebnik, C. Schmid and J. Ponce, "Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2006.

[15] A. Oliva and A. Torralba, "Modeling the shape of the scene:holistic representation of the spatial envelope," *International Journal of Computer Vision,* vol. 42, no. 3, pp. 145-175, 2001.

[16] Z. Guo, D. Zhang and D. Zhang, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification," *IEEE Transactions on Image Processing,* vol. 19, no. 6, pp. 1657 - 1663, 2010.

[17] P. J. Burt, "The Pyramid as a Structure for Efficient Computation," *Multiresolution Image Processing and Analysis,* no. 12, pp. 6-35, 1984.

[18] E. Adelson, C. Anderson, J. Bergen, P. Burt and J. Ogden, "Pyramid methods in image processing," *RCA Engineer,* no. Nov/Dec, pp. 29-6, 1984.

[19] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *Journal of Applied Statistics,* vol. 21, no. 2, pp. 224-270, 1994.

[20] T. Ojala, M. Pietikainen and T. Maenpaa, "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, 2002.

[21] E. Hadjidemetriou, M. D. Grossberg and S. K. Nayar, "Multiresolution histograms and their use for recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 26, no. 7, pp. 831 - 847, 2004.

[22] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: a review," *Medical image analysis,* vol. 13, no. 4, pp. 634-649, 2009.

[23] Y. Amit and D. Geman, "Shape quantization and recognition with randomized trees," *Neural Computation,* vol. 9, p. 1545–1588, 1997.

[24] L. Breiman, "Random Forest," *Machine Learning,* vol. 45, pp. 5-32, 2001.

[25] T. Dietterich, "Ensemble methods in machine learning," *Multiple Classifier Systems,* vol. 1857, pp. 1-15, 2000.

[26] S. Lecleire, F. D. Fiore, M. Antonietti, G. Savoye, F. Lemoine, F. Le Pessot, E. Lerebours and P. Ducrotté, "Endoscopic markers of villous atrophy are not useful for the detection of celiac disease in patients with dyspeptic symptoms.," *Endoscopy,* vol. 38, pp. 696-701, 2006.

Preprint submitted to 2014 IEEE International Symposium on Biomedical Imaging.
Received October 10, 2013.