Motiro: a unified automatic framework for statistical analysis of probe-based confocal laser endomicroscopy videos of colorectal mucosa

Alan U. Sabino[1], **Adriana V. Safatle-Ribeiro**[2], **Fauze Maluf-Filho**[2] **and Alexandre F. Ramos**[1,3]

[1]Department of Radiology and Oncology
Cancer Institute of University of São Paulo Medical School
São Paulo, 01246-000, Brazil

[2]Department of Gastroenterology
Cancer Institute of University of São Paulo Medical School
São Paulo, 01246-000, Brazil

[3]School of Arts, Sciences and Humanities
University of São Paulo
São Paulo, 03828-000, Brazil

{alan.sabino, alex.ramos}@usp.br

Abstract. *Probe-based Confocal Laser Endomicrocopy (pCLE) enables imaging the colorectal mucosa for screening and surveillance of cancer. Analyzing acquired videos relies on subjectivity of the endomicroscopists. Quantitative criteria are needed to enhance the diagnostics obtained using pCLE. We present Motiro, an automatic framework to extract features of the colorectal mucosa imaged by pCLE. Morphometric features of the crypts of the healthy colorectal mucosa are analysed and their variability quantified using the Shannon entropy. Hellinger distance compares the statistics of a morphometric parameter in multiple mucosas (or mucosas' regions). Quantification of variability of the healthy mucosa is a prerequisite for pCLE-based early diagnostics of colorectal cancer.*

1. Introduction

The use of pCLE to image the colorectal mucosa enhanced screening and post-neoadjuvant surveillance of colorectal cancer (CRC) [Kiesslich et al. 2004]. pCLE videos aid endoscopists to classify population groups accordingly with their probabilities of developing a CRC[Gupta et al. 2020], e.g. by aberrant crypt foci quantification for inference of recurrence [Uchiyama et al. 2012] and neoplasms[Kowalczyk et al. 2020] chances. Computational tools to aid on the analysis of pCLE videos of the colorectal mucosa[Karstensen 2016] have been developed for classification of polyps[André 2012] and morphometric analysis of its crypts using a non-automatic set of programs based on softwares Icy and ImageJ[Quénéhervé et al. 2019]. In this short paper we present Motiro[1], a Python-based automatic framework for feature extraction and statistical characterization of the morphometry of the healthy colorectal mucosa imaged by pCLE at pre (R) or post (T) neoadjuvant treatment. Architectural variability of a healthy mucosa is

[1]From Tupi-guarani, the language of native Brazilians, meaning a reunion for building.

smaller than in earlier stages of CRC. Motiro computes the differential Shannon entropy (and Hellinger distance) to quantify the disorder of a feature's histogram (and overlay of histograms). Motiro enables automatic feeding of a database of morphological features of images of colorectal mucosa to be jointly used with clinical data[Kowalczyk et al. 2018] for: *i)* developing computer-based techniques for earlier diagnosis of CRC; *ii)* search for optimal screening and surveillance agenda for CRC patients[Gupta et al. 2020].

2. Methods and Materials

Motiro's input are single channel real-time *in vivo* pCLE videos of colorectal mucosa. pCLE detects fluorescence intensities of intravenously injected fluorescein protein. A probe is inserted through the endoscope's working channel into the rectum a few minutes after fluorescein injection to acquire $10^3\times$ magnified images of the colorectal mucosa of CRC patients at 12 frames/s. The study protocol was approved by the ethical committee of the University of São Paulo School of Medicine at the Cancer Institute of the State of São Paulo (registered at ClinicalTrials.gov: NCT02284802). All patients signed the informed consent. Videos of the healthy colorectal mucosa at R and T moments were used to build the mosaics used for automatic feature extraction and statistical characterization. Motiro combines tools from Open Source Computer Vision Library (OpenCV)[Bradski and Kaehler 2000], and ImageJ plugin Register Virtual Stack Slices[Arganda-Carreras et al. 2006] in three stages: *1.* it receives pCLE videos as input, dismantle the videos into frames that are stitched to generate a mosaic, that is preprocessed and segmented using k-means algorithm; *2.* the crypts of the mucosa are approximated by ellipses, morphometric analysis is done estimating their perimeters, area, axis ratio, elongation factor, roundness, sphericity, maximal Feret diameter, and crypt-to-crypt distances; *3.* the statistics of morphometry is done and the Shannon entropy and Hellinger distance between histograms of features built for R and T videos are computed.

3. Results and Discussion

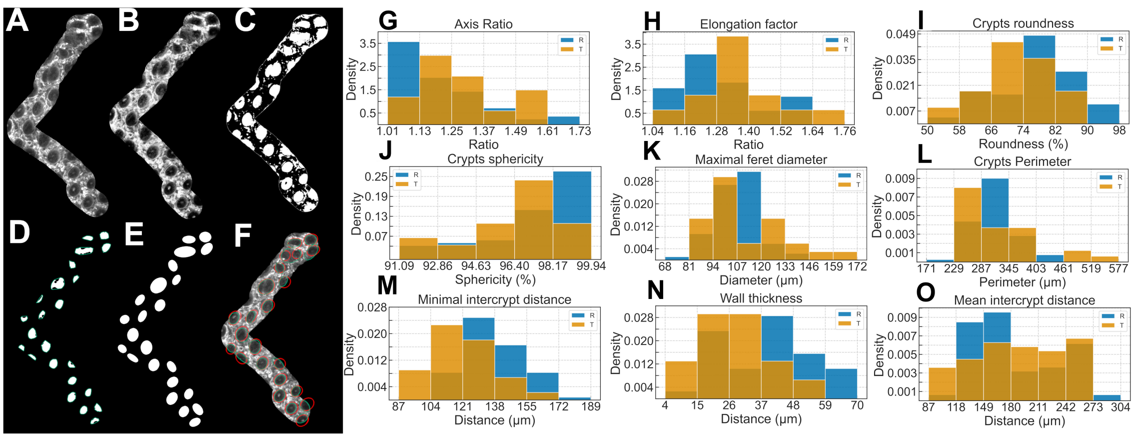

Figure 1. Panel of feature extraction process and morphometric data generated using Motiro. A-F show the automatic estimation of crypts' contours (green). G-O show the distribution of the assessed morphometric features of crypts from healthy mucosa of R and T mosaic images.

Fig. 1A shows a mosaic which irregular geometry results from the imaging acquisition process: the sensibility of the probe and absence of reference points causes a maneuvering variability. Application of contrast enhancement and noise removal highlights the crypts from background as shown in Fig. 1B. Fig. 1C shows a segmentation of the crypts and surrounding stroma. Application of morphological operations and application of convex hull to smooth the crypts' goblet cells is shown in Fig. 1D. Fig. 1E shows the elliptical contours estimation of the crypts' surrounding after the convex hull. Inspection of Figs. 1E and 1C indicates the similarity of the elliptical estimates to the segmented crypts after removal of noise of the stroma and irregularities of the crypts' boundaries. Fig. 1F shows a comparison between Motiro (red) and Icy's[de Chaumont et al. 2012] (green) contours estimation overlaid to the original mosaic. Figs. 1G-1O shows the statistical characterization of morphometric features of the crypts, the superposed blue and beige histograms summarize data for R and T mosaic images, respectively, after outliers' removal whereby the values bellow first quartile plus $1.5 \times$ the interquartile range (IQR) and above third quartile plus $1.5 \times$ IQR were indicated as outliers. The prevalence of single mode histograms and similarity between R and T mosaic images indicate the regular structure of the healthy mucosa. For the histograms of the features presented in Figs. 1G-1O, the calculated Hellinger distance between the histograms of R and T mosaic images ranges from 0.230 to 0.407 and it is fair to conclude that there is a good similarity on the morphometry of the healthy mucosa in both images. We compare the differential entropy of each histogram to their corresponding uniform and it is less than 0.848. That reinforces our conclusion that the evaluated healthy mucosa has a high degree of order despite their intrinsic variability. The morphometry analysis presented has been performed by Quénéhervé et al.(2019) (QA) using Icy and ImageJ softwares separately. Because the contours of the crypts are obtained manually in QA, we assume them gold-standard. The mean relative difference between estimates for the morphometric features obtained using Motiro and QA is 0.167. That demonstrates the viability of the automatic segmentation of pCLE videos of the healthy colon's mucosa. The learning curves of Motiro, executed from terminal, and QA, executed on graphical user interface (GUI), are similar. However, Motiro has new functionalities combined in one framework for pCLE videos analysis, is automatic and is $5.7 \times$ faster than QA with heuristic algorithm to evaluate crypts. The time complexity of the most expensive Motiro's function is $\mathcal{O}(n^2)$ and $\mathcal{O}(2^n)$ for heuristic and exhaustive search algorithm, respectively, where n is the number of crypts segmented. Icy is free to use but since its code is not available we could not analyse its time complexity, and the run-time difference is mainly explained by the user interaction with GUI.

Motiro brings a new automatic feature extraction method and significant improvements to statistically assess the morphometry of a healthy colorectal mucosa imaged by pCLE. It is ready to be employed on the analysis of multiple videos and to construct a large database. Motiro do not need a large quantity of pre-classified data to extract features and the morphometric properties have clear geometrical and biological interpretations. That may be useful on future application of classification algorithms. Additionally, the quantitative characterization of the architecture of the mucosa enables the establishment of additional standards to aid on human analysis and by reducing the role of subjectivity.

References

André, B. (2012). Software for automated classification of probe-based confocal laser endomicroscopy videos of colorectal polyps. *World Journal of Gastroenterology*, 18(39):5560.

Arganda-Carreras, I., Sorzano, C. O. S., Marabini, R., Carazo, J. M., Ortiz-de Solorzano, C., and Kybic, J. (2006). Consistent and Elastic Registration of Histological Sections Using Vector-Spline Regularization. In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Beichel, R. R., and Sonka, M., editors, *Computer Vision Approaches to Medical Image Analysis*, volume 4241, pages 85–95. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.

Bradski, G. and Kaehler, A. (2000). OpenCV. *Dr. Dobb's journal of software tools*, 3.

de Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., Meas-Yedid, V., Pankajakshan, P., Lecomte, T., Le Montagner, Y., Lagache, T., Dufour, A., and Olivo-Marin, J.-C. (2012). Icy: an open bioimage informatics platform for extended reproducible research. *Nature Methods*, 9(7):690–696.

Gupta, S., Lieberman, D., Anderson, J. C., Burke, C. A., Dominitz, J. A., Kaltenbach, T., Robertson, D. J., Shaukat, A., Syngal, S., and Rex, D. K. (2020). Recommendations for Follow-Up After Colonoscopy and Polypectomy: A Consensus Update by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology*, 158(4):1131–1153.e5. Publisher: Elsevier.

Karstensen, J. G. (2016). Evaluation of confocal laser endomicroscopy for assessment and monitoring of therapeutic response in patients with inflammatory bowel disease. *Danish Medical Journal*, 63(11).

Kiesslich, R., Burg, J., Vieth, M., Gnaendiger, J., Enders, M., Delaney, P., Polglase, A., McLaren, W., Janell, D., Thomas, S., Nafe, B., Galle, P. R., and Neurath, M. F. (2004). Confocal laser endoscopy for diagnosing intraepithelial neoplasias and colorectal cancer in vivo. *Gastroenterology*, 127(3):706–713.

Kowalczyk, M., Orlowski, M., Klepacki, L., Zinkiewicz, K., Kurpiewski, W., Kaczerska, D., Pesta, W., Zieliński, E., and Siermontowski, P. (2020). Rectal aberrant crypt foci (ACF) as a predictor of benign and malignant neoplastic lesions in the large intestine. *BMC Cancer*, 20(1):133.

Kowalczyk, M., Orłowski, M., Siermontowski, P., Mucha, D., Zinkiewicz, K., Kurpiewski, W., Zieliński, E., Kowalczyk, I., and Pedrycz, A. (2018). Occurrence of colorectal aberrant crypt foci depending on age and dietary patterns of patients. *BMC Cancer*, 18(1):213.

Quénéhervé, L., David, G., Bourreille, A., Hardouin, J. B., Rahmi, G., Neunlist, M., Brégeon, J., and Coron, E. (2019). Quantitative assessment of mucosal architecture using computer-based analysis of confocal laser endomicroscopy in inflammatory bowel diseases. *Gastrointestinal Endoscopy*, 89(3):626–636.

Uchiyama, T., Takahashi, H., Endo, H., Kato, S., Sakai, E., Hosono, K., Yoneda, M., Inamori, M., Hippo, Y., Nakagama, H., and Nakajima, A. (2012). Number of aberrant crypt foci in the rectum is a useful surrogate marker of colorectal adenoma recurrence. *Digestive Endoscopy: Official Journal of the Japan Gastroenterological Endoscopy Society*, 24(5):353–357.