Real-Time Polyp Detection in Colonoscopy Videos: A Preliminary Study For Adapting Still Frame-based Methodology To Video Sequences Analysis

Quentin Angermann, Jorge Bernal, Cristina Sánchez-Montes, Maroua Hammami, Gloria Fernández-Esparrach, Xavier Dray, Olivier Romain, F. Javier Sánchez, Aymeric Histace

Purpose:

Colorectal cancer is the second leading cause of cancer death in United States[1] when men and women are combined. Its incidence can be mitigated by detecting its precursor lesion, the polyp, before it develops into cancer. Colonoscopy is still the gold standard for colon screening though some polyps are still missed. This can be explained by technical limitations of colonoscopes (camera orientation, field of view, etc.), but also by human factors such as the number of exams already done, or the fact that one or more endoscopists are present during the exam. Several computational systems have already been proposed to assist clinicians in this task[2] but none of them is used in the exploration room due to not meeting real time constraints and not being tested under actual interventional sequences, compulsory to being of actual clinical use. A real-time system needs to process each image in less than 40 milliseconds. Such system aims to reduce the polyp miss rate by detecting region of interest in the image which could be a polyp.

In this abstract, we present a methodology to adapt and evaluate a frame-based method formerly introduced in [3], to the video real-time context, necessary step for a clinical use of proposed approaches. Adaptation involves the use of more computationally efficient feature descriptors and the incorporation of spatio-temporal stability in method's response.

Moreover, to assess quantitatively the performance of the proposed adapted approach, a full new annotated video database is introduced for the first time is this work.

Method:

The still frame detection system we chose as reference for this study and introduced in [3], was based on an active learning method, the training process being divided into two main steps:

- a cascade Adaboost learning step for computation of a classifier using patches extracted from the whole set of polyp images (polyp patches and non-polyp ones)
- a strengthening strategy based on active learning principle using hard negative examples reinjected into the training step.

The freely available CVC-Clinic[4] database was used for training and the CVC-Colon database for testing.

The adaptation of this method to video analysis is based on two main aspects: (i) influence of the local descriptors used for polyp candidate characterisation, and (ii) introduction of spatio-temporal coherence. Considering (i) if local binary pattern features were initially used, Haar-like features are also considered. About spatio-temporal coherence, main objective is to take advantage of the sequence of images by taking into account of previous detected area and consequently to reduce false detections. This step consists of performing a block fusion on the current frame with the two preceding images in a way such a detection in the current frame is only provided as actual system's output if it was a detection in the similar area in the two previous frames.

As said before, we assess the performance of our methodology using a new fully public annotated video database and under clinical and technical criteria. The validation was done on a brand new set of 18 videos from colonoscopy containing one polyp in each video and using two groups of metrics :

- The standard image/video metrics:

    - Precision: metric focused on the rate between true positives and false positives provided by the system.
    - Recall: metric focused on the rate between true positives and false negatives provided by the system.
    - F1-Score: metric combining Precision and Recall, weighting equally the impact of false positives and false negatives.
- ad hoc clinical metrics (assessing the clinical usability):
    - Polyp Detection Rate (PDR): checks if a method is able to detect polyp at least once in a sequence.
    - Mean Processing Time per frame (MPT): time to process one frame. It includes the detection time and the time to display result on the monitor.
    - Mean Number of False Positive per frame (MNFP): number of false positive alarms that can disturb the physician.
    - Reaction Time (RT): represents the delay between first appearance of a polyp in the sequence and the first correct detection provided by the method.

We will focus our attention on the influence of local features, on the improvement by the strengthening strategy related to the active learning process and on the influence of the spatio-temporal coherence.

Results:

Table 1 shows first the influence of local features (LBP and Haar-like) on the overall performance when considering usual metrics and, on secondly, how performance can be improved by considering a strengthening strategy. Here, we can see that Haar-like features are more interesting than LBP features (with respect to most of the metrics) and that the strengthening strategy improves results for the MNFP and the RT metrics, as well as MPT score when Haar features are considered. Most important, Table 1 shows that Haar-like features seems to be the best features to perform real-time processing with a good trade off in term of usual and clinical-based metrics.

Table 2 shows the influence of the spatio-temporal coherence. As expected, the MNFP is decreased, to the cost of an increasing of the RT metric. Nevertheless, the gain in clinical usability remains significant.

Conclusion:

Work presented in this abstract shows how a still-frame-based polyp detection method can be adapted to video analysis. We also introduced clinical usability metrics to measure the performance of the method and a fully new annotated video database for performance evaluation. From the experiments, we can conclude that :

- Whatever the configuration of the method, all compared setups lead to detect at least once all polyps in each video.
- Active learning strategy confirmed to be a very interesting way to improve the classifier performance for both LBP and Haar-like features.
- Haar-like features combined with active learning strategy and spatio-temporal coherence performed better than LBP, with a mean processing time per frame of only 21 ms.

The proposed method appears as an interesting alternative for real time polyp detection in clinical daily routine but there is room to improvement: image preprocessing, motion tracking of the camera to improve the spatio-temporal coherence or incorporate additional feature descriptors to improve overall performance.

Method	PDR	MPT	MNFP	Precision	Recall	F1	RT
LBPN0	100%	140 ms	1.9	16.25%	41.25%	23.31%	35.0 [1.4 sec]
LBPN1	100%	160 ms	1.1	27.11%	46.02%	34.12%	43.7 [1.7 sec]
LBPN2	100%	162 ms	0.7	29.88%	34.96%	32.22%	45.9 [1.8 sec]
HaarN0	100%	36 ms	0.9	27.02%	39.61%	32.12%	38.3 [1.5 sec]
HaarN1	100%	21 ms	0.6	39.14%	42.56%	40.78%	27.3 [1.1 sec]

Table 1. Comparison of overall performance results using N0, N1 and N2 classifiers for LBP and N0 and N1 classifiers for Haar-like features. In both cases, spatio-temporal coherence strategy is used.

Method	PDR	MPT	MNFP	Precision	Recall	F1	RT
LBPN0 with STC	100%	140 ms	1.9	16.25%	41.25%	23.31%	35.0 [1.4 sec]
LBPN0 without STC	100%	140 ms	3.5	12.42%	54.65%	20.24%	7.2 [0.3 sec]
HaarN0 with STC	100%	36 ms	0.9	27.02%	39.61%	32.12%	38.3 [1.5 sec]
HaarN0 without STC	100%	24 ms	1.4	23.29%	46.82%	31.10%	17.5 [0.7 sec]

Table 2. Overall performance results using N0 classifiers for LBP and Haar-like features, with and without spatio-temporal coherence (STC).

REFERENCES

1. ACS2016: Key statistics for colorectal cancer. Online (2016)

2. J. Bernal, N. Tajbakhsh, F. J. Sanchez, B. Matuszewski, H. Chen, L. Yu, Q. Angermann, O. Romain, B. Rustad, I, Balasingham, K. Pogorelov, S. Choi, Q. Debard, L. Maier-Hein, S. Speidel, D. Stoyanov, P. Brandao, H. Cordova, C. Sanchez-Montes, R. S. Gurudu, G. Fernandez-Esparrach, X. Dray, J. Liang, A. Histace, "Comparative Validation of Polyp Detection Methods in Video Colonoscopy: Results from the MICCAI 2015 Endoscopic

Vision Challenge", IEEE Trans. On Medical Images (TMI), *2017*. DOI: 10.1109/TMI.2017.2664042.
3. Angermann Q., Histace A., Romain O.: Active learning for real time detection of polyps in videocoloscopy. Procedia Computer Science 90, 182-187 (2016)
4. Bernal J., Sanchez F.J., Fernandez-Esparrach G, Gil D., Rodriguez C., Vilarino F. : Wm-dova maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physician. Computerized Medical Imaging and Graphicsn 43, 99-111 (2015).