Cyclope : vers une vidéocapsule intelligente pour la détection *in situ* des polypes colorectaux

Juan Silva[1], Aymeric Histace[1], Olivier Romain[1], Xavier Dray[2], Bertrand Granado[3], Andrea Pinna[3]

[1]ETIS UMR CNRS 8051
6 av. du Ponceau, 95014 Cergy-Pontoise, France

[2]iTECH, Paris 7, Hôpital Lariboisière
2 rue Ambroise Paré, 75010 Paris, France

[3]LIP6 UMR CNRS 7606
4 place Jussieu, 75252 Paris, France

aymeric.histace@u-cergy.fr, olivier.romain@u-cergy.fr,
xavier.dray@lrb.aphp.fr, bertrand.granado@lip6.fr, andrea.pinna@lip6.fr

Résumé − En France, la prévention du cancer colorectal est une priorité de santé publique. La détection d'adénomes coliques, lésions précancéreuses, fait l'objet d'un dépistage de masse par recherche de saignement digestif occulte. L'examen diagnostique de première intention est la vidéocoloscopie sous anesthésie générale. La capsule colique est un dispositif autonome qui permet une visualisation complète de la surface muqueuse colique, sans anesthésie ni hospitalisation. La capsule colique est une technique diagnostique prometteuse qui pourrait utilement remplacer la vidéocoloscopie de dépistage. Actuellement, la capsule colique présente des limites techniques qui rendent insuffisantes ses capacités de détection et de caractérisation des polypes (autonomie, sensibilité correcte mais médiocre spécificité). L'objectif scientifique du projet Cyclope porte sur l'idée de réaliser une détection *in situ* et en temps réel de polypes coliques lors d'un examen de type vidéocapsule. Nos travaux portent donc sur la conception des briques technologiques nécessaires à la fiabilisation de la détection et la reconnaissance (taille, caractère villeux, dysplasie) embarquées au sein de la vidéocapsule des polypes coliques. Pour cela, nous définissons les algorithmes nécessaires et compatibles avec les contraintes d'embarquement et la réalisation d'un démonstrateur de laboratoire. Dans cet article, nous présenterons plus particulièrement les premiers résultats obtenus en termes de détection 2D des polypes : les résultats quantitatifs les plus récents obtenus au moyen d'un processus d'apprentissage de type *boosting* fondé sur la texture des polypes et testé sur une base de données de 300 clichés de polypes se caractérisent par une sensibilité de 91,1%, une spécificité de 95,2% et enfin un taux de faux positifs de 4,8%.

Abstract − This paper presents a new embeddable method for polyp detections in Wireless Capsule Endoscopic - WCE images. this approach consists first of extracting candidate polyps within the image using geometric considerations about related shape, and second, in classifying (polyp/non-polyp) obtained candidates by a boosting-based method using texture features. The proposed approach has been designed in accordance with the hardware constraints related to FPGA implementation for integration within WCE imaging device. The classification performance of the method have been evaluated on a large dataset of 300 polyps, and 1200 non-polyps images. Experiments show interesting and promising performance: the boosting-based classification is characterized by a sensitivity of 91,1%, a specificity of 95,2% and a false detection rate of 4.8%. The performance of the boosting-based classification are in accordance with the most recent reference on this particular topic using the same dataset.

1 Introduction

La vidéocapsule endoscopique, introduite en 2003 par Given Imaging, est aujourd'hui une alternative crédible à la vidéocoloscopie. C'est un dispositif autonome et sans fils qui permet de prendre des images dans certains segments d'intérêt du tube digestif. Ces images sont transmises à un boîtier d'enregistrement porté en ceinture par le patient. Du fait de leur autonomie limitée (8h), il existe actuellement trois types de vidéocapsules endoscopiques, dont les caractéristiques ont été optimisées pour examiner respectivement l'œsophage, l'intestin grêle ou le côlon. Aucune ne permet un enregistrement continu de tout l'appareil digestif. L'examen est réalisé sur un mode ambulatoire, c'est-à-dire sans hospitalisation. Le patient ingère la vidéocapsule et le praticien récupère sous 24h le boitier d'enregistrement (type *holter*). Dans le cadre particulier du dépistage précoce du cancer colorectal, lors de l'examen par vidéocapsule du côlon, l'analyse par le praticien des 140 000 images acquises, sous forme d'un film d'environ

45 minutes, lui permet de détecter et caractériser diverses lésions notamment des polypes. Cependant, les vidéocapsules possèdent plusieurs limitations de par leur architecture minimaliste (capteurs, batterie, émetteur RF) :

- les images délivrées ne permettent d'estimer la taille des lésions qu'avec une précision de 50% ;
- l'autonomie de 8h est insuffisante pour assurer l'examen de l'ensemble du système digestif ;
- le temps de traitement des images par le praticien est important et fastidieux ;
- les outils disponibles sur le marché ne permettent pas d'interpréter de façon fiable le caractère précancéreux ou cancéreux des polypes.

Cyclope est un projet qui vise à étudier et à concevoir les briques technologiques (SW et HW) qui permettront de dépasser les limites actuelles des vidéocapsules en intégrant matériellement de l'intelligence afin de réaliser un diagnostic *in situ*. L'architecture proposée est illustrée par la Figure 1.

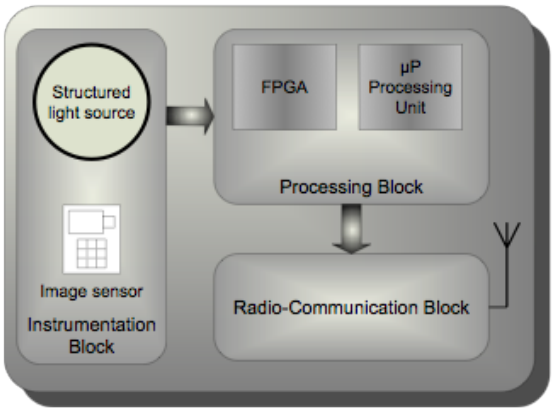


FIGURE 1 – Architecture de la capsule Cyclope

Dans [1] et [6], un premier prototype sous forme de démonstrateur principalement caractérisé par l'intégration d'un bloc d'instrumentation permettant la reconstruction des données en 3D au moyen du principe de stéréovision active (couplage entre une source de lumière cohérente (LASER) et un réseau de diffraction) a été proposé. Les critères de forme 3D extraits ont alors montré leur bonne adéquation pour la discrimination entre les polypes de type adénomateux (pouvant dégénérer vers un cancer) ou hyperplasique (bénin) : le processus de classification de type SVM (*Support Vector Machine*) intégré matériellement au sein d'un FPGA permet un taux de détection/classification global de 97% sur une base de données *in vitro* constituée de 111 polypes (40 adénomes and 81 polypes hyperplasiques) en silicone. Le passage aux données réelles a néanmoins montré les limites de cette approche quant à la capacité des descripteurs 3D a caractérisé la variété des types de polypes, principalement à un stade précoce où la structure peut alors être assez plate.

Cette article présente la mise en place d'une chaîne de traitements pour l'extraction de caractéristiques 2D dans le domaine du visible, compatible avec une implantation matérielle. L'objectif de ces traitements sont de détecter de manière fiable des structures de type polype et de renforcer les performances du détecteur précédemment proposé.

2 État de l'art et chaîne de traitements proposée

Il apparaît dans [2, 3, 5, 4] que les critères de forme et de texture restent les plus discriminants pour un étude 2D dans le domaine du visible des images acquises en vidéocoloscopie. Le tableau 1 récapitule pour chacune des approches les performances obtenues.

A noter qu'aucune de ces approches ne proposent d'utiliser les deux types de critères au sein d'une unique approche de détection. Il apparaît également que les bases de données d'images issues de vidéocapsules restent de tailles assez réduites et peu comparables entre elles. Seule la base mise librement à disposition par Figueiredo *et al.* [3] contient un nombre significatifs de polypes mais acquis en vidéocolonoscopie et non au moyen d'une WCE.

Enfin, ces méthodes sont développées afin de permettre un traitement *off-line* des données et sans prise en compte d'une possible intégration matérielle.

Pour pallier ces limitations, nous proposons donc une chaîne de traitements inspirée dans l'enchaînement des briques par l'étude psychovisuelle du gastroentérologue lors d'un examen :

1. Détection de contours de type Canny-Deriche ;
2. Détection des formes circulaires fondée sur une transformée de Hough multi-rayons ;
3. Classification des régions d'intérêt extraites à l'étape précédente au moyen de critères texturaux.

Concernant l'étape 3, le classifieur est construit au moyen d'un processus d'apprentissage de type *attentional boosting* [10] en utilisant comme caractéristiques texturales les paramètres d'Haralick (26) extraits des matrices de cooccurrence calculées sur une base de données d'apprentissage adaptée (voir paragraphe suivant). Ce choix particulier de descripteurs correspond à la démarche psychovisuelle du praticien qui cherche, une fois une région d'intérêt identifiée, des variations dans la texture de la muqueuse intestinale pouvant lui indiquer la présence d'une zone à risque.

A noter que chacune des étapes proposées présente la spécificité d'avoir d'ores et déjà été implantée matériellement sur un support de type FPGA au regard de la littérature, ce qui est un point essentiel en ce qui concerne le projet Cyclope dans l'objectif d'effectuer l'examen *in situ*. Ce point en particulier est discuté dans le dernier paragraphe de l'article.

Références	Performances de détection	Base de données
Bernal *et al.* [2]	Sensibilté 89% Spécificité 98%	300 images de polypes acquises en vidéocoloscopie (**libre d'accès**)
Figueiredo *et al.* [3]	Pas de performances indiquées	17 vidéos acquises en vidéocoloscopie de 100 images chacune, contenant 10 polypes, des lésions plates,, des diverticules, des bulles, et des matières.
Kodogiannis *et al.* [5]	Sensibilité 97% Spécificité 94%	140 images de vidéocapsules (70 polypes and 70 non-polypes)
Karkargyris *et al.* [4]	Sensibilité 100% Spécificité 67.5%	50 images de vidéocapsules (10 polypes and 40 non-polypes)

TABLE 1 – Synthèse des méthodes les plus récentes

3 Expérimentations

Afin de quantifier les performances de classification et de détection de la chaîne de traitements proposée, nous proposons d'utiliser la base de données de Bernal et al [2] bien qu'issue d'un videocolonoscope classique et non d'une WCE : cette base de données reste à ce jour la plus riche en termes d'exemples et la seule mise à disposition par les auteurs.

Afin d'adapter cette base au processus d'apprentissage de l'étape 3 de la chaîne de traitements proposée, sur chaque image de polype (300), une région d'intérêt rectangulaire est extraite autour du polype et quatre faux-polypes sont générés manuellement afin de constituer une base d'apprentissage représentative (voir Figure 2).

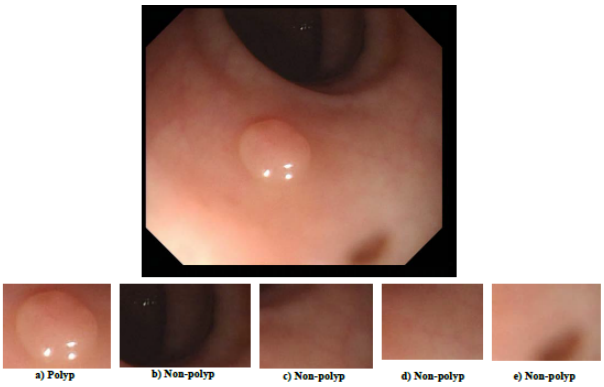


FIGURE 2 – Création de la base de donnée servant à l'étape d'apprentissage et de test de la méthode proposée

La base est alors scindée en deux parties : 200 polypes et 800 non-polypes servent à la phase d'apprentissage et le reste (100 polypes et 400 non-polypes) à l'évaluation des performances de classification selon les critères habituels de *Spécificité*, de *Sensibilité* et du *Taux de Faux Positif* (TFP).

Les paramètres texturaux sont calculés sur des matrices de cooccurrence ramenées sur 32 niveaux de gris significatifs, ce qui est habituellement le sous-échantillonnage tonal retenu pour préserver une validité statistique des paramètres qui en sont extraits (Homogénéité, Entropie, Moment d'ordres supérieurs...) tout en restreignant la taille des matrices et donc le temps de calcul.

Par ailleurs, afin d'obtenir des résultats significatifs statistiquement en termes de performance, une procédure de validation croisée est mise en place : les tests sont réalisés 50 fois, à chaque fois sur un tirage aléatoire des exemples constituant les bases d'apprentissage et de tests.

La Tableau 2 récapitule les résultats obtenus et les compare à des méthodes classiques (Adaboost, LVQ) ainsi qu'aux performances de la méthode de Bernal *et al.*

Type Adaboost	Sensibilité	Spécificité	TFP
LVQ classification	92,0%	86,0%	14,0%
Real Adaboost	77,1%	92,5%	7,5%
Cascade Adaboost	91,1%	95,2%	4,8%
[2]	89%	98%	2%

TABLE 2 – Performances comparées de différents types d'approches de classification dont celle proposée par [2].

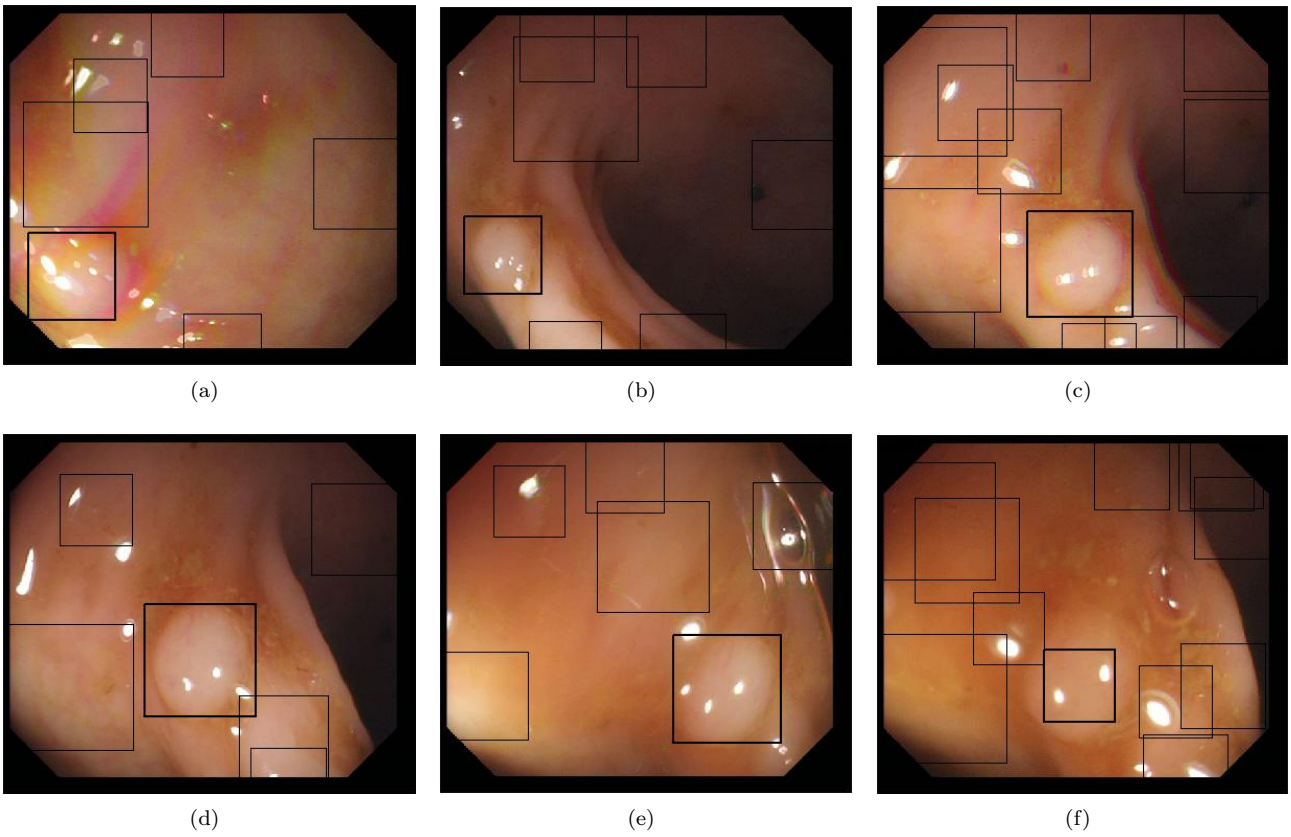
Ces résultats montrent que les critères texturaux issus des matrices de cooccurrence intégrés dans un processus de type *boosting* en cascade permettent une classification compétitive.

Enfin, la Figure 3 illustre sur quelques images extraites de la base de données les résultats obtenus au moyen de la chaîne de traitements proposée.

Pour chaque image de la Figure 3, les ROI carrées correspondent aux zones candidates détectées via la transformée de Hough circulaire. Les zones ensuite effectivement identifiées comme polypes par l'étape de classification sont indiquées par un carré au contour renforcé.

4 Conclusion et Discussion

Comme le montre les performances résumées dans le Tableau 2, ainsi que les résultats de détection présentés Figure 3, la chaîne de traitement proposée dans cet article est une solution tout à fait compétitive pour la mise en place d'une détection des polypes sur des images ac-



(a) (b) (c)

(d) (e) (f)

FIGURE 3 – Exemples de résultats de détection/classification obtenus sur quelques images extraites de la base de données.

quises en vidéoendoscopie. On notera en particulier que l'apprentissage permet d'effectuer la différence entre un polype et une bulle d'air (nombreuses lors des examens) comme illustré sur l'image de la Figure 3.(e). En outre, chacune des étapes proposées restent compatibles avec une implantation matérielle comme le montre la littérature associée [7, 8, 9]. L'étape de développement suivante consistera donc à développer les IP correspondantes à chacune des briques de la chaîne proposée. En parallèle, une base de données d'images issues d'examens par vidéocapsule sera mise en place.

Références

[1] J. Ayoub, B. Granado, Y. Mhanna, and O. Romain. SVM based colon polyps classifier in a wireless active stereo endoscope. In *2010 IEEE EMBC*, pages 5585 –5588, 2010.

[2] J. Bernal, J. Sanchez, and F. Vilari no. Towards automatic polyp detection with a polyp appearance model. *Pattern Recognition*, 45(9) :3166 – 3182, 2012.

[3] P. N. Figueiredo, I. N. Figueiredo, S. Prasath, and R. Tsai. Automatic polyp detection in pillcam colon 2 capsule images and videos : Preliminary feasibility report. *Diagnostic and Therapeutic Endoscopy*, 2011.

[4] A. Karargyris and N. Bourbakis. Identification of polyps in wireless capsule endoscopy videos using log gabor filters. In *IEEE Workshop LiSSA*, pages 143 –147, april 2009.

[5] V. Kodogiannis and M. Boulougoura. An adaptive neurofuzzy approach for the diagnosis in wireless capsule endoscopy imaging. *Int. J. of Information Technology*, 13 :46 – 56, 2007.

[6] A. Kolar, O. Romain, J. Ayoub, S. Viateur, and B. Granado. Prototype of Video Endoscopic Capsule With 3-D Imaging Capabilities. *Biomedical Circuits and Systems, IEEE Transactions on*, 4(4) :239 –249, 2010.

[7] J. Miteran, J. Matas, E. Bourennane, M. Paindavoine, and J. Dubois. Automatic hardware implementation tool for a discrete adaboost-based decision algorithm. *EURASIP Journal on Applied Signal Processing*, 2005 :1035–1046, 2005.

[8] L. Siéler, C. Tanougast, and A. Bouridane. A scalable and embedded FPGA architecture for efficient computation of grey level co-occurrence matrices and haralick textures features. *Microproc. and Microsys.*, 34(1) :14 – 24, 2010.

[9] S. Tagzout, K. Achour, and O. Djekoune. Hough transform algorithm for fpga implementation. *Signal Processing*, 81(6) :1295 – 1301, 2001.

[10] S. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of IEEE CVPR Conference*, pages 511–518, 2001.